

KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

KONUŞMACI TANIMADA GERÇEKÇİ SİNİR AĞLARININ PERFORMANSININ
İNCELENMESİ

DOKTORA TEZİ

Zübeyir ÖZCAN

AĞUSTOS 2024

TRABZON



KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

KONUŞMACI TANIMADA GERÇEKÇİ SİNİR AĞLARININ
PERFORMANSININ İNCELENMESİ

Zübeyir ÖZCAN

Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsünde
"DOKTOR (ELEKTRONİK MÜHENDİSLİĞİ)"
Unvanı Verilmesi İçin Kabul Edilen Tezdir.

Tezin Enstitüye Verildiği Tarih : 08 / 08 / 2024

Tezin Savunma Tarihi : 23 / 08 / 2024

Tez Danışmanı : Prof. Dr. Temel KAYIKÇIOĞLU

Trabzon 2024

ÖNSÖZ

Bu tez, Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Elektrik-Elektronik Mühendisliği Anabilim Dalı, Elektronik Mühendisliği Bilim Dalı Doktora Programı'nda hazırlanmıştır. Bu tez çalışmasında; gerçekçi sinir ağları tabanlı bir örüntü tanıma yaklaşımı önerilmiş, test edilmiş ve otomatik konuşmacı tanımda performansı incelenmiştir.

Akademik hayata başladığım dönemden itibaren bilimsel desteği ve yönlendirici fikirleriyle bana yardımcı olan ve yol gösteren danışmanım Prof. Dr. Temel KAYIKÇIOĞLU'na teşekkür ederim. Tez çalışmam süresince değerli katkıları için tez izleme komitesi hocalarım Prof. Dr. Ali GANGAL ve Prof. Dr. Cemal KÖSE'ye teşekkür ederim.

Bu zorlu sürecin her anında, yanımda olan ve bana inanan aileme sonsuz teşekkürlerimi sunarım. Özellikle annem Sıddıka ÖZCAN ve babam İmran ÖZCAN'a, bana verdikleri maddi ve manevi destek için minnettarım. Bu süreçte; sevgisi, sabrı ve hoşgörüsüyle her daim yanımda olan sevgili nişanım Tuğba AĞI'ya da ayrıca minnettarım. Varlıkları ve destekleri, en büyük güç ve motivasyon kaynağım olmuştur.

Bu süreçte bana arkadaşlık eden, fikir ve yardımlarıyla bana destek olan değerli arkadaşlarım Dr. Taha Yasin HATAY, Y. Müh. Alper Tunga AKIN, Sıla SEÇKİN, Halil KILIÇ, Oğulcan YADİGAROĞLU ve Cemal KALKAN'a da teşekkür ederim. Uzun çalışma gecelerimde sabır ve anlayışla yanımda bekleyen Kaktüs Kafe sahibi Sinan SARUHAN'a ayrıca teşekkür ederim. Onların dostluğu ve destekleri, bu çalışmayı tamamlamamda büyük bir katkı sağlamıştır.

Son olarak, doktora sürecimde karşılaştığım her zorluğu bir öğrenme fırsatına dönüştüren tüm kişilere teşekkür eder, bu çalışmanın alana katkı sağlamasını umut ederim.

"Bilginin efendisi olmak için çalışmanın kölesi olmak gerekir." – Honore de Balzac

Zübeyir ÖZCAN

Trabzon 2024

TEZ ETİK BEYANNAMESİ

Doktora Tezi olarak sunduğum “Konuşmacı Tanımada Gerçekçi Sinir Ağlarının Performansının İncelenmesi” başlıklı bu çalışmayı baştan sona kadar danışmanım Prof. Dr. Temel KAYIKÇIOĞLU’nun sorumluluğunda tamamladığımı, verileri/örnekleri kendim topladığımı, deneyleri/analizleri ilgili laboratuvarlarda yaptığımı/yaptırdığımı, başka kaynaklardan aldığım bilgileri metinde ve kaynakçada eksiksiz olarak gösterdiğimi, çalışma sürecinde bilimsel araştırma ve etik kurallara uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul ettiğimi beyan ederim.

23/08/2024

Zübeyir ÖZCAN

İÇİNDEKİLER

	<u>Sayfa No</u>
ÖNSÖZ.....	III
TEZ ETİK BEYANNAMESİ.....	IV
İÇİNDEKİLER.....	V
ÖZET	III
SUMMARY	IV
ŞEKİLLER DİZİNİ.....	V
TABLolar DİZİNİ.....	III
SEMBOLLER VE KISALTMALAR DİZİNİ	III
1. GENEL BİLGİLER	1
1.1. Biyolojik Sinir Ağları.....	3
1.1.1. Hücre Zarında İyonlar ve İyon Kanalları	4
1.1.2. Aksiyon Potansiyeli	6
1.1.3. Biyolojik Nöron	7
1.1.4. Biyolojik Nöronlarda Sinaptik Mekanizma ve Plastisite.....	8
1.1.5. Biyolojik Nöronlarda Dendritik Yapı	11
1.1.6. Piramidal Nöronlar.....	12
1.1.7. Biyolojik Nöronlarda Nöral Kodlama.....	13
1.2. Yapay Sinir Ağları.....	14
1.2.1. Yapay Nöron Modeli	16
1.2.2. Yapay Sinir Ağlarında Öğrenme.....	18
1.3. Gerçekçi Sinir Ağları.....	21
1.3.1. Nernst Denklemi ve Hücre Zarının Pasif Modeli	22
1.3.2. Aktif İyon Kanallarının Modellenmesi	26
1.3.3. Sızıntılı Toplam Ateşle Nöron Modeli.....	27
1.3.4. Hodgkin&Huxley Nöron Modeli	28
1.3.5. Izhikevich Nöron Modeli	30
1.3.6. Gerçekçi Nöronlarda Sinaptik Mekanizma ve Plastisite.....	31
1.3.7. Nöronların Modellenmesinde Bölmeli Yaklaşım	35
1.3.8. Gerçekçi Nöronlarda Nöral Kodlama	38
1.4. Otomatik Konuşmacı Tanıma	41

1.4.1.	Mel Frekansı Kepstrum Katsayıları	42
1.4.2.	Konuşmacı Tanıma Yaklaşımları.....	45
1.4.3.	Konuşmacı Tanımadaki Performans Ölçütleri.....	51
1.5.	Literatür Taraması	53
1.5.1.	Genel Ses Tanıma	54
1.5.2.	Otomatik Konuşma Tanıma	55
1.5.3.	Otomatik Konuşmacı Tanıma	56
1.6.	Tezin Kapsamı ve Katkıları.....	57
2.	YAPILAN ÇALIŞMALAR.....	60
2.1.	Veri Kümeleri.....	63
2.1.1.	Yapay Veri Kümesi.....	64
2.1.2.	Karşılaştırma Veri Kümeleri	65
2.1.3.	Librispeech Veri Kümesi	67
2.2.	Gerçekçi Modeller İçin Seçilen Parametreler	69
2.2.1.	Basit Nöron Modelleri	70
2.2.2.	Piramidal Nöron Modeli	72
2.2.3.	Sinaps Modeli	74
2.3.	MFKK Esaslı Öznitelik Çıkarımı.....	76
2.4.	Gerçekçi Sinir Ağı Modeli	78
2.4.1.	Öznitelik-Spike Dönüşümü.....	80
2.4.2.	Aynalama Yaklaşımı.....	85
2.4.3.	Ağ Kurulumu ve Sinaptik Bağlantılar	86
2.4.4.	Öğrenme Yöntemi.....	88
2.4.5.	Eğitim.....	90
2.4.6.	Test.....	90
2.4.7.	Yazılım.....	93
2.5.	Hibrit Yaklaşım için Derin Sinir Ağı Mimarisi.....	95
2.5.1.	Analog-Analog Dönüşüm	97
2.5.2.	Analog-İkili Dönüşüm	97
2.6.	Deneysel Çalışmalar.....	98
2.6.1.	D1: Otomatik Konuşmacı Tanımadaki En Gelişkin Yöntemlerin Performansının İncelenmesi.....	99
2.6.2.	D2: Öznitelik Spike Dönüşümü Yaklaşımının Karşılaştırmalı İncelenmesi	103

2.6.3.	D3: Sinaptik Bağlantı Tiplerinin Sınıflandırma Performansına Etkisinin İncelenmesi	106
2.6.4.	D4: Önerilen Gerçekçi Sinir Ağı Modelinin Konuşmacı Tanıma Performansının İncelenmesi.....	108
2.6.5.	D5: Piramidal Hücrenin Konuşmacı Tanıma Performansının İncelenmesi	111
3.	BULGULAR VE TARTIŞMA	115
3.1.	Otomatik Konuşmacı Tanımaya Dair Öncül Bulgular.....	116
3.1.1.	Otomatik Konuşmacı Tanıma En Gelişkin Yöntemlerin Performansı ve Parametrelerin Etkisi	118
3.1.2.	Eğitim Verisi Uzunluğunun Performansa Etkisi.....	122
3.1.3.	Test Verisi Uzunluğunun Performansa Etkisi.....	124
3.1.4.	MFKK Sayısının Performansa Etkisi.....	126
3.1.5.	Tartışma	128
3.2.	Klasik Makine Öğrenimi Yöntemlerinde Öznitelik Spike Dönüşümü Aşamasının İncelenmesi.....	132
3.2.1.	Kodlama Yönteminin Etkisi.....	133
3.2.2.	Kuantalama Tekniğinin Etkisi	139
3.2.3.	Bit Uzunluğunun Etkisi.....	142
3.2.4.	Tartışma	145
3.3.	Gerçekçi Sinir Ağı Modeli Performansının Karşılaştırmalı İncelenmesi ve Sinaptik Bağlantı Tipinin Etkisi	148
3.3.1.	Yapay Veri Kümesinde Öbek Alan Haritaları	149
3.3.2.	IRIS Veri Kümesinde Sınıflandırma Performansı	151
3.3.3.	MNIST Veri Kümesinde Sınıflandırma Performansı	153
3.3.4.	Tartışma	155
3.4.	Otomatik Konuşmacı Tanıma Uygulamasında Gerçekçi Sinir Ağlarının Performansı	157
3.4.1.	VK6 Veri Kümesinde Uygun Parametrelerin Belirlenmesi.....	157
3.4.2.	MFKK Özniteliklerinde Gerçekçi Sinir Ağlarının Performansları.....	163
3.4.3.	Analog-Analog Dönüşümlü Özniteliklerde Gerçekçi Sinir Ağlarının Performansları.....	166
3.4.4.	Analog-İkili Dönüşümlü Özniteliklerde Gerçekçi Sinir Ağlarının Performansları	170
3.4.5.	Piramidal Nöron ve Alt Bölgelerinin Otomatik Konuşmacı Tanıma Performansları.....	173
3.4.6.	Tartışma	176

4. SONUÇLAR VE ÖNERİLER.....	186
5. KAYNAKLAR	191
6. EKLER.....	207
ÖZGEÇMİŞ	



Doktora Tezi

ÖZET

KONUŞMACI TANIMADA GERÇEKÇİ SİNİR AĞLARININ PERFORMANSININ İNCELENMESİ

Zübeyir ÖZCAN

Karadeniz Teknik Üniversitesi
Fen Bilimleri Enstitüsü
Elektrik-Elektronik Mühendisliği Anabilim Dalı
Danışman: Prof. Dr. Temel KAYIKÇIOĞLU
2024, 206 Sayfa, 31 Sayfa Ek

Bu çalışmada, işitsel sistem esinli bir otomatik konuşmacı tanıma sisteminin performansı incelenmiştir. Bu sistemde, kulak, işitsel sinirler ve beyin, bilgi işleme, analiz ve sınıflandırma yetenekleri, oldukça gerçekçi bir yaklaşımla taklit edilmiş ve performansı otomatik konuşmacı tanımada incelenmiştir. Bulgular, önerilen yaklaşımın işitsel sinyallerin sınıflandırılmasında önemli potansiyelini vurgulamaktadır. Buna göre, literatürde sunulan en gelişkin yöntemlere kıyasla, önerilen yaklaşım otomatik konuşmacı tanıma görevini yüksek başarıyla yerine getirebilmektedir. Ayrıca, çalışmada önerilen hibrit yaklaşım bulgularına göre derin sinir ağlarının ve gerçekçi sinir ağlarının ardışık kullanılmasının sınıflandırma problemlerinde performansı artırıcı etkiye sahip olduğu gösterilmiştir. Sonuç olarak, biyo-esinli sistemlerin sınıflandırmada kullanılmasıyla, en gelişkin yöntemlerden daha yüksek performans elde edilebileceği ortaya koyulmuştur.

Anahtar Kelimeler: Modelleme, Gerçekçi Sinir Ağları, Piramidal Nöron, Dendritik Hesaplama, Sinaps, Sınıflandırma, Konuşma, Otomatik Konuşmacı Tanıma, Evrensel Arka Plan Modeli, Derin Öğrenme

PhD. Thesis

SUMMARY

INVESTIGATING THE PERFORMANCE OF REALISTIC NEURAL NETWORKS IN SPEAKER RECOGNITION

Zübeyir ÖZCAN

Karadeniz Technical University
The Graduate School of Natural and Applied Sciences
Electrical and Electronics Engineering Graduate Program
Supervisor: Prof. Dr. Temel KAYIKÇIOĞLU
2024, 206 Pages, 31 Pages Appendix

In this study, the performance of an auditory system-inspired automatic speaker recognition system was investigated. In this system, the information processing, analysis and classification abilities of the ear, auditory nerves and the brain are simulated with a very realistic approach and their performance is investigated in automatic speaker recognition. The findings highlight the significant potential of the proposed approach in classifying auditory signals. Accordingly, compared to the state-of-the-art methods presented in the literature, the proposed approach can fulfill the task of automatic speaker recognition with high success rate. In addition, according to the findings of the hybrid approach proposed in the study, it has been shown that the sequential use of deep neural networks and realistic neural networks has a performance-enhancing effect on classification problems. As a result, it has been revealed that using bio-inspired systems in classification can achieve higher performance than state-of-the-art methods.

Key Words: Modelling, Realistic Neural Networks, Pyramidal Neuron, Dendritic Computation, Synapse, Classification, Speech, Automatic Speaker Recognition, Universal Background Model, Deep Learning

ŞEKİLLER DİZİNİ

Sayfa No

Şekil 1. Hücre zarında iyonların dağılımı ve iyon kanallarının temsili gösterimi.....	5
Şekil 2. Aksiyon potansiyeli ve aşamaları.....	6
Şekil 3. Biyolojik nöron	8
Şekil 4. Sinaps anatomisi.....	9
Şekil 5. Biyolojik nöronların morfolojik çeşitliliği (Georgiev vd., 2020).....	11
Şekil 6. Piramidal nöronda dendritik zayıflama (Özcan vd., 2018)	13
Şekil 7. Yapay nöron modeli	17
Şekil 8. Gerçekçi nöron modeli	22
Şekil 9. Pasif iyon kanallarının hücre zarında elektriksel eşdeğer devresi.....	24
Şekil 10. Pasif kanallar için indirgenmiş eşdeğer devre	25
Şekil 11. HH nöron modelinin temsili ve eşdeğer devresi	29
Şekil 12. IZH nöron modelinde parametreler ve dinamikleri (Izhikevich, 2006)	31
Şekil 13. Farklı zaman sabitlerine göre sinaptik iletkenlik değişimi.....	33
Şekil 14. SZBP ile öğrenmede sinaptik ağırlığın zamana göre değişimi	34
Şekil 15. Bölmeli nöron modeli ve bölmelere ait örnek eşdeğer devreler	36
Şekil 16. Bölmeli bir nöronun elektriksel eşdeğeri	37
Şekil 17. MFKK hesaplama akış diyagramı	43
Şekil 18. GKM yaklaşımında konuşmacı sınıflandırma akış diyagramı	46
Şekil 19. GKM-EAPM yaklaşımında konuşmacı sınıflandırma akış diyagramı.....	48
Şekil 20. Öznitelik dönüştürücülerin eğitimi ve kullanımına dair akış diyagramı.....	50
Şekil 21. Biyolojik konuşmacı tanıma süreci ve önerilen yaklaşımlar.....	61
Şekil 22. Yapay veri kümesi öbekleri ve iki boyutlu uzayda dağılımları	65
Şekil 23. IRIS veri kümesi ve iki boyutta sınıflarının dağılımı.....	66
Şekil 24. MNIST veri kümesi ve iki boyutta sınıflarının dağılımı.....	67
Şekil 25. Çalışmada kullanılan piramidal nöron modeli	72
Şekil 26. MFKK hesaplama adımları	77
Şekil 27. Tasarlanan gerçekçi sinir ağı modelinin akış diyagramı	79
Şekil 28. Örnek bir sayı dizisiyle kuantalama teknikleri.....	81
Şekil 29. Tekdüze kuantalanmış bir sayı dizisinin ikili formda kod karşılıkları.....	83
Şekil 30. Piramidal nöron için sinaptik bağlantıların yerleşimi	87

Şekil 31. Tekil örüntü tipi veriler için gerçekçi sinir ağı çıktıları	91
Şekil 32. Çoğul örüntü tipi veriler için gerçekçi sinir ağı çıktıları	93
Şekil 33. Önerilen öznitelik dönüştürücü derin sinir ağının mimarisi.....	96
Şekil 34. Önerilen öznitelik-spike dönüştürme yönteminin basit yapısı	98
Şekil 35. D2 için belirlenen deney senaryosunun akış diyagramı	105
Şekil 36. D4 için kullanılan öznitelik vektörleri.....	110
Şekil 37. D5 için hibrit sinir ağı mimarisi	112
Şekil 38. En gelişkin yöntemlerin OKT performansları.....	119
Şekil 39. Eğitim verisi uzunluğunun OKT performansına etkisi	123
Şekil 40. Test verisi uzunluğunun OKT performansına etkisi	125
Şekil 41. MFKK sayısının OKT performansına etkisi	127
Şekil 42. Kodlama yöntemine göre öbek alan haritaları	134
Şekil 43. Kodlama yöntemine göre IRIS veri kümesinde performanslar.....	136
Şekil 44. Kodlama yöntemine göre MNIST veri kümesinde performanslar	137
Şekil 45. Kuantalama tekniğine göre öbek alan haritaları.....	140
Şekil 46. Kuantalama tekniğine göre IRIS veri kümesinde performanslar	141
Şekil 47. Bit uzunluğuna göre öbek alan haritaları	143
Şekil 48. Bit uzunluğuna göre IRIS veri kümesinde performanslar.....	144
Şekil 49. Sinaptik bağlantı tiplerine göre öbek alan haritaları	150
Şekil 50. Sinaptik bağlantı tiplerine göre IRIS veri kümesinde eğitim hataları.....	152
Şekil 51. Sinaptik bağlantı tiplerine göre IRIS veri kümesinde performanslar.....	153
Şekil 52. Sinaptik bağlantı tiplerine göre MNIST veri kümesinde eğitim hataları	154
Şekil 53. Sinaptik bağlantı tiplerine göre MNIST veri kümesinde performanslar.....	155
Şekil 54. ÖSD seçeneklerine göre VK6 veri kümesinde eğitim hataları.....	158
Şekil 55. ÖSD seçeneklerine göre VK6 veri kümesinde performanslar	160
Şekil 56. Sinaptik bağlantı tiplerine göre VK6 veri kümesinde eğitim hataları.....	162
Şekil 57. Sinaptik bağlantı tiplerine göre VK6 veri kümesinde performanslar.....	163
Şekil 58. MFKK öznitelikleri kullanıldığında eğitim hataları ve performanslar	165
Şekil 59. A1 öznitelikleri kullanıldığında eğitim hataları ve performanslar	167
Şekil 60. A2 öznitelikleri kullanıldığında eğitim hataları ve performanslar	169
Şekil 61. B16 öznitelikleri kullanıldığında eğitim hataları ve performanslar	171
Şekil 62. B32 öznitelikleri kullanıldığında eğitim hataları ve performanslar	172
Şekil 63. Piramidal nöron için eğitim hataları ve performanslar.....	175
Şekil 64. Belirlenen şartlar altında en gelişkin yöntemler ve GSA performansları	184

TABLÖLAR DİZİNİ

	<u>Sayfa No</u>
Tablo 1. Karışıklık matrisi.....	51
Tablo 2. LASRC veri tabanı ve bileşenleri (Panayotov vd., 2015)	68
Tablo 3. Konuşmacılar ve çalışmanın veri kümeleri	69
Tablo 4. Basit nöron modelleri ve parametreleri	71
Tablo 5. Piramidal nöron için bölgeler ve kullanılan parametreler	73
Tablo 6. Sinaps modelleri ve kullanılan parametreler	75
Tablo 7. Kodlama yöntemleri ve 16 seviye için kod karşılıkları.....	82
Tablo 8. D1 için kullanılan parametre dizileri.....	101
Tablo 9. Literatürde otomatik konuşmacı tanıma performansları	177
Tablo 10. Belirlenen şartlar altında en gelişkin yöntemlerin OKT performansları.....	179
Tablo 11. Belirlenen şartlar altında GSA'ların OKT performansları	183

SEMBOLLER VE KISALTMALAR DİZİNİ

AFD	: Ayrık Fourier Dönüşümü
AKD	: Ayrık Kosinüs Dönüşümü
AMPA	: α -Amino-3-Hidroksi-5-Metil-4-Izoksazolpropiyonik Asit
AP	: Aksiyon Potansiyeli
ATP	: Adenozin Trifosfat
BN	: Biyolojik Nöron
BSA	: Biyolojik Sinir Ağı
DÖ	: Derin Öğrenme
DSA	: Derin Sinir Ağı
DVM	: Destek Vektör Makineleri
EAPM	: Evrensel Arka Plan Modeli
EVU	: Eğitim Verisi Uzunluğu
Exc	: Excitatory, Uyarıcı
GABA	: Gama Aminobütirik Asit
GHK	: Goldman-Hodgkin-Katz
GK	: Gray Kodlama
GKM	: Gauss Karışım Modeli
GN	: Gerçekçi Nöron
GSA	: Gerçekçi Sinir Ağı
HFD	: Hızlı Fourier Dönüşümü
HH	: Hodgkin&Huxley
Inh	: Inhibitory, Baskılayıcı
IZH	: Izhikevich
İK	: İkili Kodlama
K-EYK	: K-En Yakın Komşuluk
LAA	: Lineer Ayırıcı Analizi
LASRC	: Librispeech Automatic Speech Recognition Corpus
MÇSA	: Maksimum Çıktı Sinir Ağı
MFKK	: Mel Frekansı Kepstrum Katsayıları
MÖ	: Makine Öğrenimi

MP	: McCulloch & Pitts
NB	: Naive Bayes
NMDA	: N-Metil-D-Aspartat
OKH	: Ortalama Karesele Hata
OKT	: Otomatik Konuşmacı Tanıma
ÖAH	: Öbek Alan Haritası
ÖSD	: Öznitelik Spike Dönüşümü
PN	: Piramidal Nöron
ReLU	: Doğrultulmuş Lineer Birim
SB	: Sinaptik Bağlantı
STA	: Sızıntılı Topla&Ateşle
STS	: Superior Temporal Sulcus
SZBP	: Spike Zamanına Bağlı Plastisite
TKP	: Tek Katmanlı Perceptron
TÖ	: Transfer Öğrenme
t-SKG	: t-Dağıtılmış Stokastik Komşu Gömme
TVU	: Test Verisi Uzunluğu
USD	: Uzun Süreli Depresyon
USP	: Uzun Süreli Potansiyasyon
VK	: Veri Kümesi
YK	: Yoğunluk Kodlama
YN	: Yapay Nöron
YSA	: Yapay Sinir Ağı
YZ	: Yapay Zeka
ZGSA	: Zaman Gecikmeli Sinir Ağı

1. GENEL BİLGİLER

Teknolojinin temel amacı insan yaşamını basitleştirmek ve kalitesini artırmaktır. Tarih boyunca teknolojiye ilerlemeler sürekli olarak yükleri hafifletmeyi, süreçleri kolaylaştırmayı ve bireylerin daha az çabayla daha fazlasını başarmalarını sağlayacak kolaylıklar yaratmayı amaçlamıştır. Çağdaş zamanlarda teknoloji, insan yaşamını iyileştirme hedefinin yanı sıra, daha önce insanlara ayrılmış belirli görevleri yerine getirebilecek makineler geliştirmeyi de odağına almıştır. Yapay zeka (YZ), robot teknolojisi ve otomasyondaki ilerlemeler, makinelerin tekrarlayan, emek yoğun veya özel görevleri yüksek doğruluk ve hızla yerine getirmesinin yolunu açmıştır (Wang & Siau, 2019). Bu evrim, belirli bağlamlarda maliyetlerin azaltılması ve üretkenliğin artması dahil olmak üzere çok sayıda fayda sunmaktadır (Narciso & Martins, 2020).

Bu bağlamda, makine öğrenimi (MÖ), belirli insan yeteneklerini ve işlevlerini taklit etmek veya artırmak için tasarlanmış sistemlerin geliştirilmesine olanak tanıyan güçlü bir araçtır. Algoritmalar ve istatistiksel modelleri kullanan MÖ yöntemleri, bilgisayarların verilerden öğrenmesine, kalıpları belirlemesine ve her görev için tahminlerde bulunmasına olanak tanımaktadır (Zhou, 2021). Görüntü ve konuşma tanıma, tıbbi teşhis, finansal tahmin ve otonom araçlar gibi çeşitli alanlarda, MÖ algoritmaları, çoğu zaman insan performansına rakip olan, hatta onu aşan olağanüstü performanslar sergilemiştir (Fujiyoshi vd., 2019; Sidey-Gibbons & Sidey-Gibbons, 2019; Dixon vd., 2020; Bhangale & Mohanaprasad, 2021).

Günümüzde işlenecek veri hacmi genişledikçe, klasik MÖ algoritmaları önemli zorluklar ve kısıtlarla karşı karşıya kalmaktadır. Klasik algoritmalar, genellikle, veri dağılımı ve hesaplama kaynaklarıyla ilgili belirli varsayımlar altında çalışmaktadır. Bu algoritmalar, çok büyük miktarda veriyle karşı karşıya kaldıklarında genelleme yapma konusunda zorluk yaşamaktadırlar. Ayrıca, klasik yöntemler işlemsel yük açısından ağır kalabilmektedir. Bu da işlem sürelerinin ve kaynak taleplerinin artmasına neden olmaktadır.

Klasik MÖ yöntemlerinin yetersiz kalması, araştırma odaklarının insan beyninden esinlenen yaklaşımlara yönelmesine neden olmuştur. Bu nedenle derin öğrenme (DÖ) ve yapay sinir ağları (YSA'lar) gibi daha gelişmiş teknikler, araştırmacıların dikkatini daha çok çekmektedir. Bu modern yaklaşımlar, büyük veri kümelerini işleme, karmaşık

örüntüleri çıkarma ve görüntü ve konuşma tanıma gibi görevlerde üstün performans elde etme konusunda öne çıkmaktadır (Dong vd., 2021). Buna rağmen, klasik algoritmaların belirli senaryolar ve uygulamalarda hala etkili olduklarını söylemek mümkündür.

İnsan beyninden ilham alınarak ortaya atılmış olmalarına rağmen derin sinir ağları (DSA'lar), hala insan beyninin kabiliyetinden ve veriminden yoksundurlar (Yamazaki vd., 2022). Yapay nöronların (YN'lerin) biyolojik nöronlara (BN'lere) göre basit birer temsil olması, bu ağların beynin tüm kabiliyetine sahip olmasının önünde bir engel teşkil etmektedir. Öncelikle, DSA'ların karmaşık yapıları belirli problemlerin çözümünde gereğinden fazla veriye ihtiyaç duymalarına ve eğitim aşamasında, ağır işlemler için uzun zaman harcamalarına neden olmaktadır. Bunun yanında, DSA'lar verileri ezberlemeye yatkın olduklarından test aşamalarında başarısız olabilmektedirler (Castiglioni vd., 2021).

DSA'ların bu eksiklikleri, çalışmaları beyinden daha çok ilham alan sinir ağlarına yönlendirmiştir. Bu ağlar beynin bilinen en gerçekçi temsilleri olan gerçekçi sinir ağlarıdır (GSA'lardır). Literatürde ateşlemeli sinir ağları olarak da tanımlanabilen bu ağlar, doğrusal olmayan hesaplama kabiliyetleri ve beyinden ilham almaları nedeniyle büyük ilgi görmüştür (Yamazaki vd., 2022).

GSA'ların en büyük avantajları enerji verimlilikleridir. Bunun yanında çalışmalarda YSA'lara göre daha az eğitim turunda problemleri çözebildikleri gösterilmiştir. Diğer yandan GSA'larda aynı problemin YSA'lara göre daha az sayıda nöron kullanılarak çözülebildiği belirtilmiştir (Wang vd., 2020). Tüm bu özellikler, GSA'ların gelecekte zorlayıcı problemleri daha kolay çözebileceğini göstermektedir.

Bu tez çalışmasında, otomatik konuşmacı tanıma (OKT) uygulanmasında kullanılmak üzere GSA tabanlı yaklaşımlar önerilmiş ve performansları karşılaştırmalı olarak incelenmiştir. Bu konudaki genel bilgileri içeren bu bölüm, altı alt başlıktan oluşmaktadır. İlk alt başlıkta, biyolojik sinir ağlarının (BSA'ların) temellerine değinilmiş ve moleküler seviyeden karmaşık yapılara kadar uzanan nörobiyolojik altyapı tanıtılmıştır. İkinci alt başlıkta, literatürde yaygın olarak kullanılan YSA'ların temellerine değinilmiş ve YN modellerinden öğrenme algoritmalarına kadar tüm süreç anlatılmıştır. Üçüncü alt başlıkta, BSA'ların bilinen en gerçekçi modelleri olan GSA'ların temellerine değinilmiş ve moleküler seviyeden karmaşık sinir ağlarına kadar uzanan modelleme yaklaşımları tanıtılmıştır. Dördüncü alt başlıkta, OKT alanında kullanılan güncel öznetelikler ve konuşmacı modelleme yaklaşımlarına değinilmiştir. Beşinci alt başlıkta, tez çalışması ile

ilgili literatürde yapılan çalışmalar tanıtılmıştır. Altıncı ve son alt başlıkta ise tez çalışmasının kapsamı tanıtılmış ve literatüre sunduğu katkılar sıralanmıştır.

1.1. Biyolojik Sinir Ağları

Sinir sistemi; nöral iletim, refleksler, duyuşsal algılama, motor hareketler ve bilişsel süreçlerin yönetilmesinde canlı vücutlarındaki en büyük aktördür. Özellikle beyin, canlıların dış dünyayı algılaması ve bunlara karşı tepkiler vermesi konusunda hayati bir öneme sahiptir. Beyin ilk bakışta üç temel kısımdan meydana gelmektedir. İç ve orta beyin olarak bilinen beynin kısımları, daha çok canlıların hayatta kalması için gerekli olan bilinçdışı süreçlerden sorumluyken; beynin dış kabuğu, duyuşsal algılama, duyuşsal işleme, motor hareketler ve bilişsel karar verme gibi karmaşık ve bilinçli süreçlerden sorumludur (Nowinski, 2011). Bu yanıyla dış kabuk, insanlar ve hayvanlar arasındaki zeka ve bilinç gibi özellikler bakımından gözlemlenen farklılıkların kaynağıdır.

Beyin, bilinçli veya bilinçsiz tüm görevlerini, yapısında bulunan nöral ağlar sayesinde yerine getirmektedir. BSA olarak bilinen bu yapılar, içeriğinde bulunan BN'ler ve bu nöronların birbirleriyle iletişim sağlamasına yarayan sinaptik bağlantılar (SB'ler) sayesinde karmaşık süreçleri yönetmektedirler. Bu yönüyle BSA'lar, beynin ve sinir sisteminin mekanizmasının anlaşılmasında yapılan araştırmaların temelini oluşturmaktadır (Sterratt vd., 2023).

BSA'ların yapısı incelendiğinde, BN'ler ve SB'ler olmak üzere temel iki yapı göze çarpmaktadır. BN'ler, ağ içinde iletişimde kullanılacak olan aksiyon potansiyellerinin (AP'lerin) üretilmesinden sorumluyken, SB'ler ise bu AP'lerin nöronlar arasında iletiminden sorumludurlar. Ayrıca, SB'ler BSA'ların değişen koşullara uyum sağlamasında kritik öneme sahip olan öğrenme süreçlerinin de merkezidir.

Bu bölümde, BSA'ların genel anatomisi, bileşenleri ve çalışma mekanizmalarına değinilmiştir. Bölüm yedi alt başlıktan meydana gelmektedir. İlk alt başlıkta, hücre zarında bulunan iyonlar ve iyon kanallarından bahsedilmiştir. İkinci alt başlıkta, AP'nin üretim süreci anlatılmıştır. Üçüncü alt başlıkta, BN genel olarak tanıtılmıştır. Dördüncü alt başlıkta, BN'lerin iletişimde baş rolü oynayan SB'lerin mekanizması ve plastisite tanıtılmıştır. Beşinci alt başlıkta, BN'nin dendritik çeşitliliğine dikkat çekilirken; altıncı alt başlıkta, beyin dış kabuğunda yaygın olarak rastlanılan piramidal nöronlar (PN'ler) ve

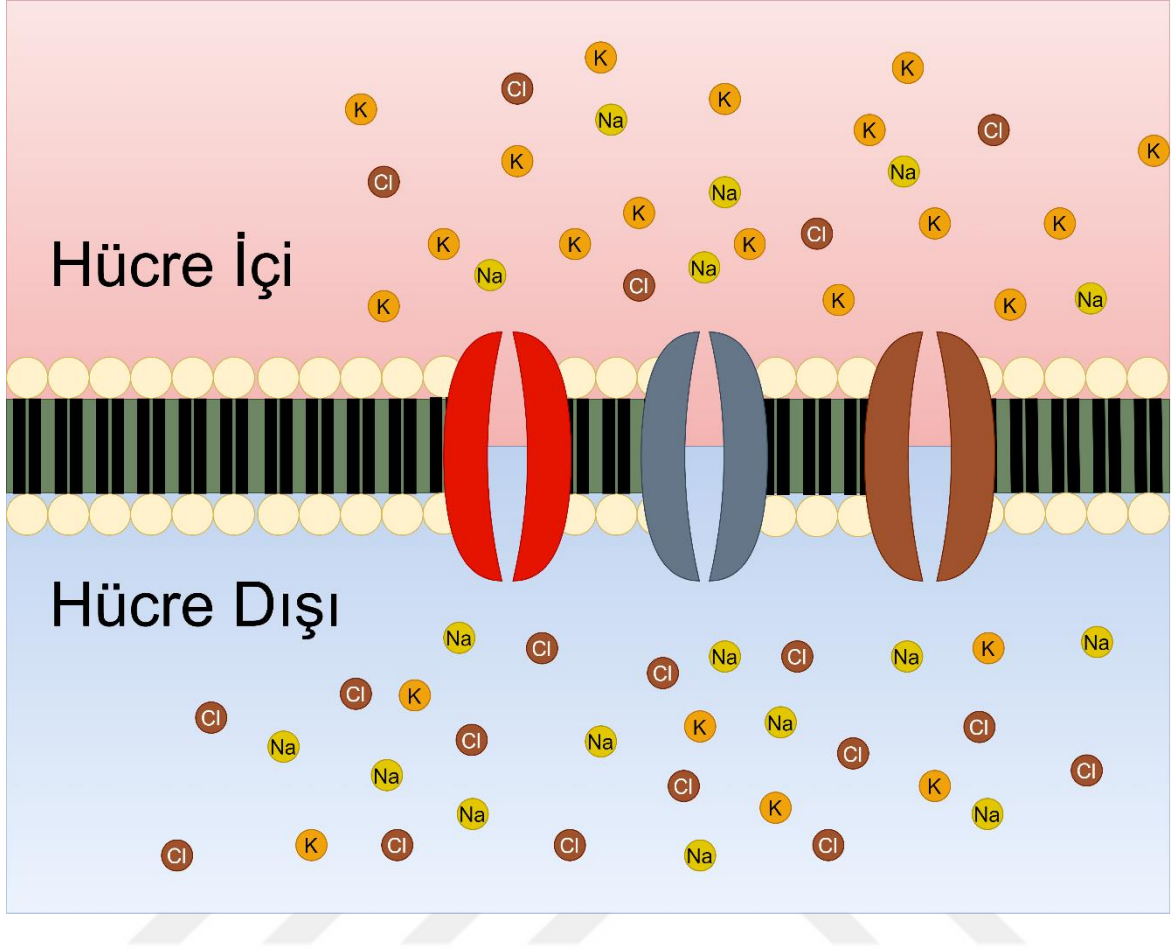
dendritik morfolojileri tanıtılmıştır. Son olarak, yedinci ve son alt başlıkta ise BN’lerde nöral kodlamaya değinilmiştir.

1.1.1. Hücre Zarında İyonlar ve İyon Kanalları

Hücre zarı, çeşitli iyonların bulunduğu yoğun bir ortamda konumlanmaktadır. Bu iyonlar arasında Sodyum, Potasyum, Klorür, Kalsiyum ve Magnezyum bulunmaktadır. Hücre zarında Sodyum iyonları (Na^+), voltaj kapılı sodyum kanalları vasıtasıyla, AP’lerin oluşumunu tetikleyerek, sinaptik iletişim ve bilgi işleme mekanizmalarının temelini oluşturmaktadır. Benzer şekilde, Potasyum iyonları (K^+), hücre zarında AP’nin repolarizasyonunda etkili olup, hücre zarının dinlenme potansiyeline geri dönmesini sağlamaktadır. Klorür iyonları (Cl^-) ise hücre içi sıvı dengesinin korunmasında ve sinaptik iletimde görev alırken; Kalsiyum iyonları (Ca^{2+}), nöro-iletici salınımı ve sinaptik plastisiteyi düzenleyerek, öğrenme ve hafıza mekanizmalarının temelini oluşturmaktadır. Son olarak, Magnezyum iyonları (Mg^{2+}) *N-metil-D-aspartat* (NMDA) reseptörleri üzerinde düzenleyici etkiye sahip olmakla beraber sinaptik plastisitede de görev almaktadır (Sterratt vd., 2023).

Bu bağlamda, iyonların hücre zarındaki dengenin korunmasındaki katkıları ve BN’lerin elektrokimyasal süreçlerinin düzenlenmesindeki görevleri, hücre zarında iyon hareketleri ve mekanizmalarına dair araştırmaların önemini ortaya koymaktadır. Bu nedenle, nörobiyolojik süreçlerin anlaşılmasında ilgi odağının merkezinde iyon hareketleri ve bu iyonların transferinden sorumlu iyon kanalları bulunmaktadır.

İyon kanalları, hücre zarında bulunan ve iyon hareketinin düzenlenmesinde temel işlevlere sahip olan protein yapılarıdır. Bu kanallar, iyonların hücre zarı üzerinde pasif veya aktif taşınmasını sağlamaktadırlar. Pasif taşıma, özellikle difüzyon yoluyla, iyonların hücre zarı boyunca iyonların konsantrasyon farkları doğrultusunda hareketlerini ifade etmektedir (Taju & Ou, 2019). Aktif taşıma ise, iyonların konsantrasyon farkına zıt yönde hareket etmelerini içeren bir süreç olup, genellikle adenosin trifosfat (ATP) hidrolizi aracılığıyla enerji bağımlıdır (Kim vd., 2023). İyon kanallarının seçici geçirgenliği ve çeşitliliği, BN’lerde, sinaptik iletişimden elektrokimyasal dengenin düzenlenmesine kadar birçok sürecin çeşitli iyonik akış mekanizmaları ile gerçekleştirilmesine olanak tanımaktadır. Şekil 1, hücre içi ve dışında bulunan iyonların dağılımını ve hücre zarındaki iyon kanallarını temsili olarak görselleştirmektedir. Şekil 1’de görselleştirilen iyon

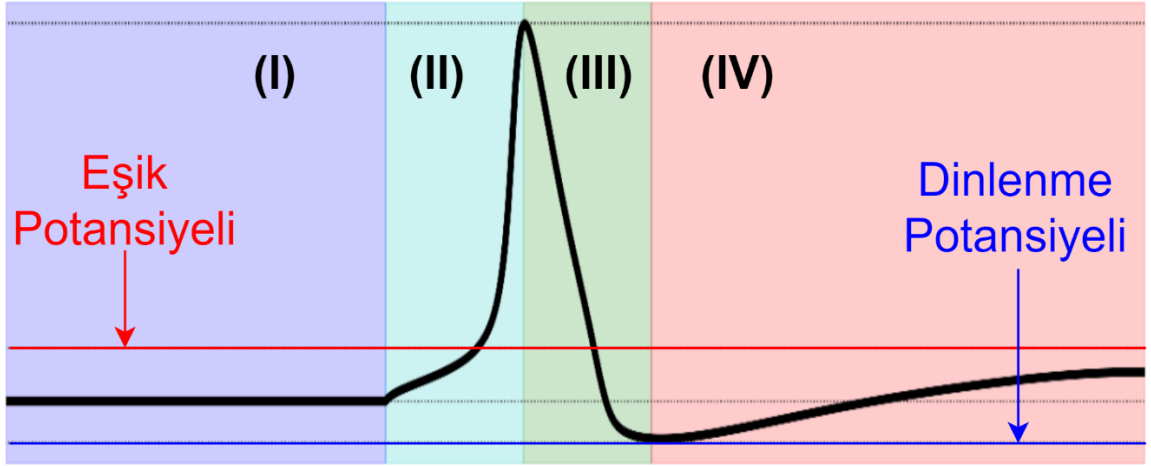


Şekil 1. Hücre zarında iyonların dağılımı ve iyon kanallarının temsili gösterimi

kanalları sayesinde iyonların hücre içi ve dışı arasındaki transferi düzenlenmekte ve hücre içi ve dışı denge sağlanmaktadır.

Pasif iyon kanalları, hücre zarında iyonların konsantrasyon farkına bağlı olarak transferini gerçekleştirmektedir. Bu kanallarda iyon akışı difüzyon yoluyla gerçekleştiğinden, pasif iyon kanallarının herhangi bir enerji ihtiyacı bulunmamaktadır. Potasyum, Sodyum, Klorür ve Kalsiyum gibi hücre zarında bulunan temel iyonlar, bu kanallar aracılığıyla geçiş yaparak hücre içi ve hücre dışı iyon konsantrasyonlarının dengelenmesini sağlamaktadırlar (Kondratskyi vd., 2017). Dolayısıyla, pasif iyon kanalları, hücre içi ve dışı iyonik dengenin anlaşılmasında araştırmaların odağındadır.

Aktif iyon kanalları, iyonların hücre zarı üzerindeki konsantrasyon farkına karşı taşınmasında görevlidir. Bu kanallar, iyonların transferi sırasında konsantrasyon farkının tersine hareketi sağlamak amacıyla, ATP moleküllerinden elde edilen enerjiyi kullanmaktadırlar. Aktif iyon kanallarının enerji bağımlı mekanizmaları, hücre zarının



Şekil 2. Aksiyon potansiyeli ve aşamaları

dinlenme potansiyelinin dengesinin bozulmasına sebep olmakta ve BN'lerin bilgi işleme mekanizmalarının temelinde konumlanan AP'lerin üretilmesini sağlamaktadır.

1.1.2. Aksiyon Potansiyeli

BN'ler, bilgi işlem ve iletim fonksiyonlarını gerçekleştirmek için AP olarak bilinen spesifik elektriksel sinyal formlarına bağımlıdır. Hücre içi ve dışı iyonların hareket dinamiklerine bağlı olarak, hücre zarında yer alan aktif ve pasif iyon kanallarının aktivasyonu ile bu spesifik potansiyel değişiklikler gerçekleşmektedir. AP'nin oluşum sürecini, Şekil 2'de gösterildiği gibi dört temel aşamada ele almak mümkündür: Dinlenme evresi (I), depolarizasyon evresi (II), repolarizasyon evresi (III) ve refrakter periyot (IV).

BN, dinlenme durumunda iken hücre zarındaki aktif iyon kanalları kapalı durumdadır. Bu durumda, Sodyum ve Potasyum iyonlarının zar üzerinden geçişi sadece pasif iyon kanalları yoluyla olmakla beraber kısıtlıdır. Hücre içerisinde, negatif yüklü moleküllerin baskın olması nedeniyle, hücrenin içinde negatif bir potansiyel mevcuttur. Bu durum, BN'nin dinlenme potansiyeli olarak adlandırılmaktadır.

Bir dış uyarı etkisi altında, hücre zarındaki aktif Sodyum kanalları hızla açılmakta ve hücre içine Sodyum iyonlarının girişi sağlanmaktadır. Sodyum iyonlarının hücre içine girmesiyle hücre zarındaki elektrokimyasal denge bozulmakta ve hücre içinde pozitif yüke yönelik bir değişiklik meydana gelmektedir. Bu hızlı pozitif değişikliğe depolarizasyon adı verilmektedir. Dış uyarı etkisinin devam etmesi sonucu, hücre zar potansiyeli pozitif zirveye ulaşmaktadır.

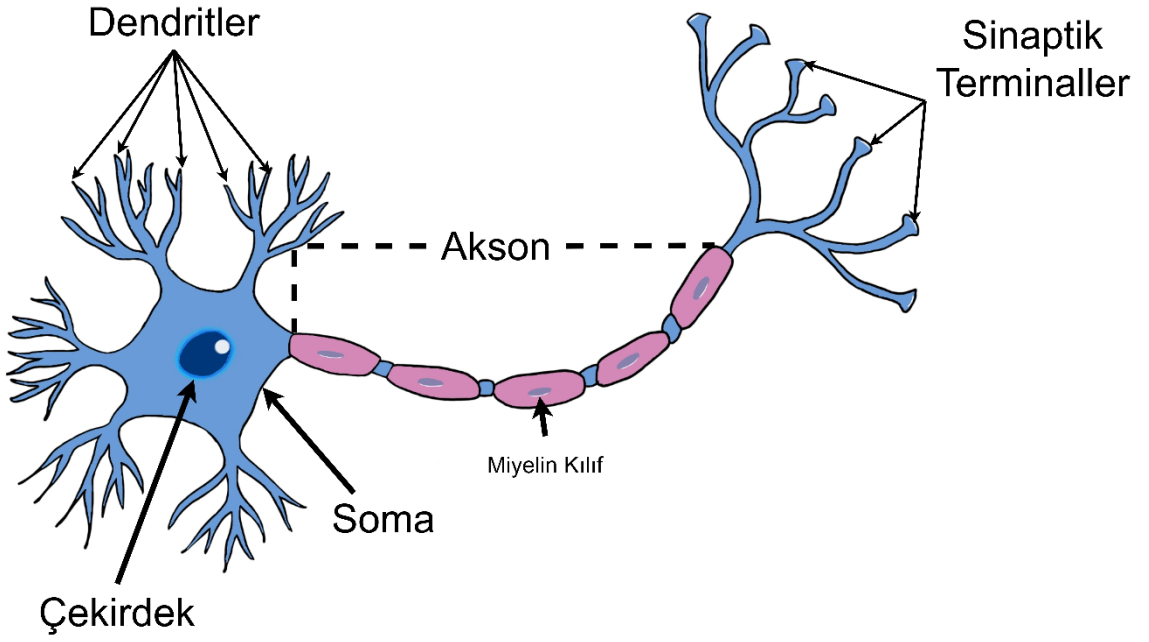
Depolarizasyonun gerçekleşmesini takiben, hücre zarında bulunan aktif Potasyum kanalları hızla açılmaktadır. Hücre içinde meydana gelen aşırı pozitif yükün hücre dışına aktarılmasını sağlamak amacıyla çalışan bu aktif iyon kanalları, hücre zarının dinlenme potansiyeline geri dönmesini mümkün kılmaktadır. Bu potansiyel değişim repolarizasyon olarak isimlendirilmektedir. Bu süreç, hücre zar potansiyeli negatif zirveye ulaşana kadar devam etmektedir.

Repolarizasyon evresinin ardından hücre zar potansiyeli, dinlenme potansiyelinden de düşük olmak üzere, negatif zirveye ulaşmış durumdadır. Bu negatif zirvenin dinlenme potansiyeli seviyesine ulaşması için tüm aktif iyon kanalları kapalı duruma geçmektedir ve pasif iyon kanalları sayesinde hücre zarının dinlenme potansiyeli seviyesinde bir dengeye ulaşması sağlanmaktadır. Bu süreç, refrakter periyot olarak isimlendirilmektedir. Refrakter periyot bir AP'nin ardından nöronun dış uyarılara geçici olarak tepki veremediği bir zaman dilimini ifade etmektedir. Refrakter periyotta tüm aktif iyon kanallarının kapalı olması nedeniyle, BN'nin gereğinden daha sık ardışık AP üretmesi engellemekte ve BN'nin istikrarlı ve düzenli bir şekilde çalışması sağlanmaktadır (Sterratt vd., 2023).

1.1.3. Biyolojik Nöron

Beyin mimarisinin temelini oluşturan BN'ler, sinir sistemindeki bilgi işleme süreçlerinin merkezi yapı taşlarıdır. BN'lerin karmaşık yapıları, sinir sinyallerinin iletimini mümkün kılan çeşitli temel bileşenlerden meydana gelmektedir. Bir BN'nin temel bileşenleri arasında dendritler, nöronal hücre gövdesi (soma) ve akson bulunmaktadır. Söz konusu anatomi Şekil 3'te temsili bir görsel ile verilmektedir. Dendritler, nöronal bağlantılarda diğer BN'lerden gelen sinirsel sinyalleri alarak bu sinyalleri hücre gövdesine yönlendiren dallanmış yapılar olarak tanımlanmaktadır. Hücre gövdesi veya soma, BN'nin metabolik aktivitelerinin gerçekleştirildiği bölgedir. Bu bölge, çekirdek ve diğer hücresel organelleri içermektedir. Son olarak, akson, hücre gövdesinden başlayarak uzanan ve elektriksel uyarıları diğer BN'lere, kaslara veya salgı bezlerine ileten uzun ve ince bir yapıdır (Kirch & Gollo, 2020).

BN'ler, sinir sisteminde duyuşal algılama, motor koordinasyon ve bilişsel işlevler gibi çeşitli biyolojik süreçlerin gerçekleşmesinde görev almaktadırlar (Lewis vd., 2020). Beyinde BN'ler, sinaptik iletim mekanizmasıyla birbirleriyle etkileşimde bulunmaktadır. Buna göre elektriksel sinyaller, BN'ler arasındaki SB noktasında kimyasal sinyallere



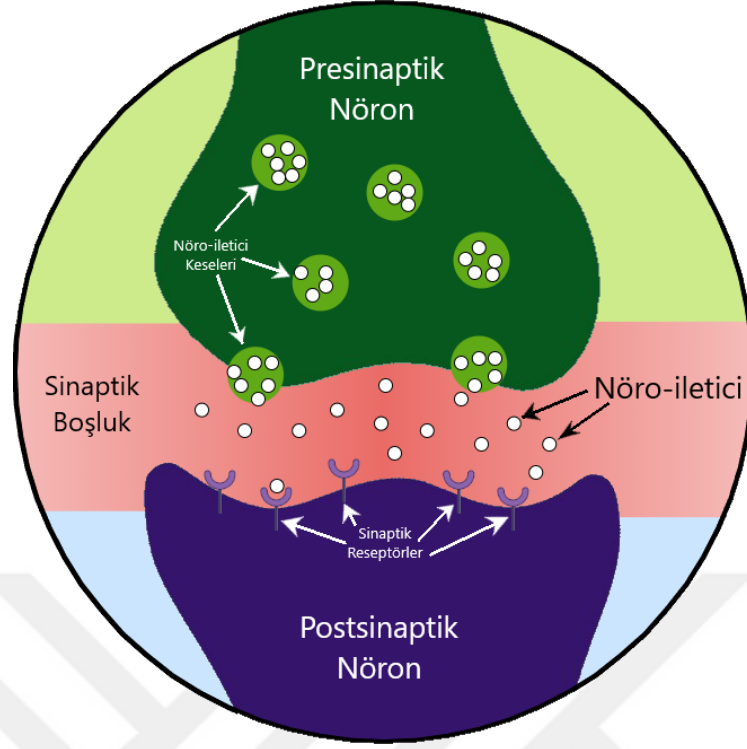
Şekil 3. Biyolojik nöron

dönüştürülerek iletim gerçekleştirilmektedir. Söz konusu kimyasal sinyal, hedef BN üzerindeki reseptörler aracılığıyla bir elektriksel uyarıyı tetikleyerek bilginin aktarımını sağlamaktadır. Sinir sisteminin içerisinde bulunan BSA'lar, sinyal iletiminin optimizasyonunu sağlayarak duyuşsal algılama ve bilişsel işlevler gibi birçok görevi yönetmektedirler.

BN'lerin anatomik yapısı, özellikle dendritik ve aksonal çeşitlilikleri, nöronal devrelerin karmaşık mimarisini ve işlevsel kapasitelerini gözler önüne sermektedir. Dendritik morfolojinin özgün yapısı, BN'lerin çeşitli kaynaklar ve bölgelerden gelen girdileri alabilmesini ve bu girdileri işleyebilmesini mümkün kılmaktadır. BN'lerin dinamik ve karmaşık yapısı sayesinde sinir sistemi yüksek bilgi işlem kabiliyeti kazanmaktadır (Kirch & Gollo, 2020).

1.1.4. Biyolojik Nöronlarda Sinaptik Mekanizma ve Plastisite

Sinapslar, sinir sisteminin iletişimdeki temel bileşenleridir. Bu özelleşmiş bağlantılar, presinaptik BN'ler tarafından salgılanan nöro-iletici moleküller aracılığıyla iletimi sağlamaktadırlar. Nöro-ileticiler, postsinaptik BN üzerindeki özel reseptörlere bağlanarak postsinaptik BN'nin zar potansiyelini etkilemektedirler.



Şekil 4. Sinaps anatomisi

Sinapslar genel olarak iki temel kategoriye ayrılmaktadır: Uyarıcı (Excitatory, Exc) ve Baskılayıcı (Inhibitory, Inh). Exc sinapslar, postsinaptik BN'nin zar potansiyelini depolarize ederek AP oluşumunu teşvik etmektedir. Bu mekanizma; genellikle glutamat gibi belirli nöro-iletici moleküllerin postsinaptik reseptörlere spesifik bir etkileşimle bağlanmasıyla, özellikle sodyum iyonlarının postsinaptik BN içerisine transferi sonucu gerçekleşmektedir. Böylece, presinaptik BN ile postsinaptik BN arasında bilgi alışverişi mümkün hale gelmektedir.

Öte yandan, Inh sinapslar, postsinaptik BN'nin zar potansiyelini hiperpolarize ederek AP oluşumunu baskılamaktadır. Bu baskılama; genellikle *gama aminobütirik asit* (GABA) gibi spesifik nöro-iletici moleküllerin postsinaptik reseptörlere bağlanmasıyla, Klorür iyonlarının hücre içine veya Potasyum iyonlarının hücre dışına transferi sonucu gerçekleşmektedir. Bu sayede, BN'nin aşırı uyarılması kontrollü bir şekilde sınırlanmakta ve genel dengenin sürdürülmesi sağlanmaktadır.

Sinaptik mekanizma, BN'lerin arasındaki iletişimi sağlayan temel yapı ve işlevsel süreçleri kapsamaktadır. Bu etkileşim, bir sinaptik boşlukta birbirleriyle buluşmuş presinaptik ve postsinaptik BN'ler arasında gerçekleşmektedir. Söz konusu bu anatomi, Şekil 4'te temsili olarak görselleştirilmiştir.

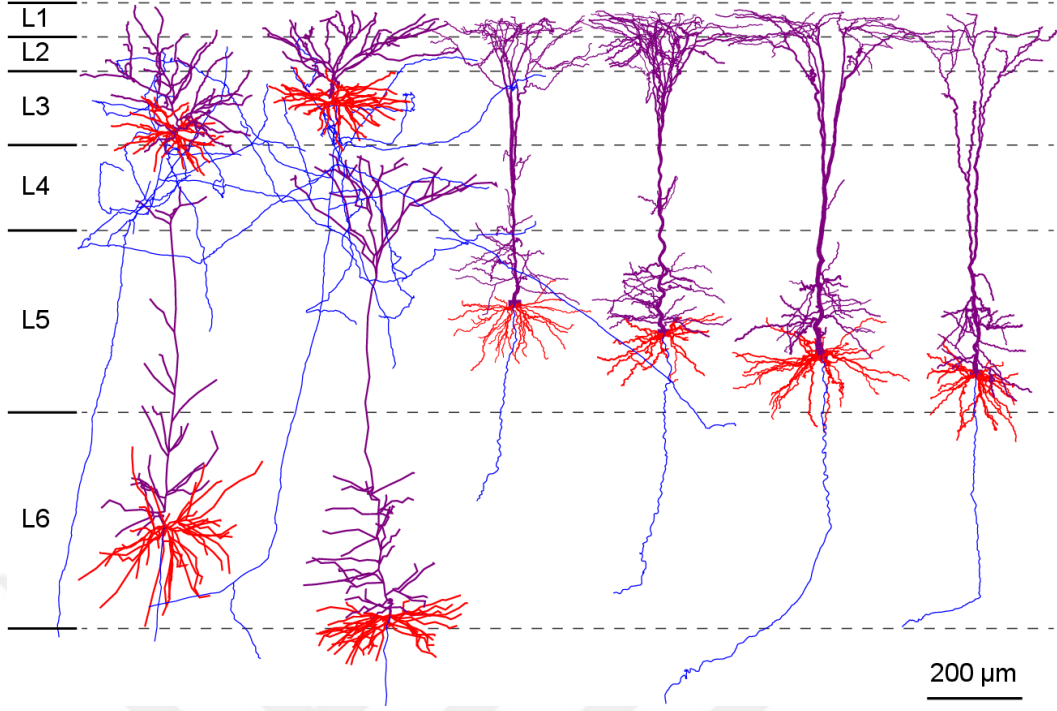
Presinaptik BN, AP'nin oluřtuđu BN'yi ifade etmektedir. AP, presinaptik BN boyunca ilerledikçe, son konum olarak presinaptik terminal adı verilen bir bölgeye ulaşmaktadır. Bu terminalde, nöro-iletici keseleri adı verilen yapılar, nöro-ileticileri depolamaktadır. Nöro-ileticiler, elektrokimyasal iletişimin gerçekleştirilmesinde ana bileşenlerdir. Sinaptik boşluk, iki BN arasında bulunan ve nöro-ileticilerin serbest bırakılacağı özel bir alandır. Postsinaptik BN, sinaptik iletimde alıcı olarak işlev görmektedir ve üzerinde belirli nöro-ileticilere karşılık gelen spesifik sinaptik reseptörler mevcuttur. Bu reseptörler, uygun nöro-ileticinin varlığında, postsinaptik BN içindeki elektrokimyasal süreçleri tetiklemektedirler.

Sinaptik iletim, presinaptik BN'nin AP üretmesiyle başlamaktadır. Üretilen AP, presinaptik BN'nin aksonu boyunca ilerleyerek akson terminaline doğru hareket etmektedir. Akson terminaline eriştiğinde, AP, aktif Kalsiyum kanallarının açılmasına neden olmaktadır. Böylece, Kalsiyum iyonlarının presinaptik terminale giriři başlamaktadır. Kalsiyum iyonlarının giriři, nöro-iletici keselerinin hücre zarına yaklaşmasına neden olmaktadır. Ardından, nöro-ileticiler sinaptik boşluđa serbest bırakılmaktadır. Nöro-ileticiler, postsinaptik BN'nin hücre zarında özel reseptörlere bağlanmaktadır. Bununla birlikte, postsinaptik BN'nin hücre zarında bulunan aktif Sodyum, Potasyum, Kalsiyum ve Klorür kanalları reseptörün uyarıcı veya baskılayıcı doğasına bağlı olarak ilgili iyonların transferlerini gerçekleştirmektedirler. Bu iyonların hareketi, postsinaptik nöronda zar potansiyelinin deđişimine neden olmaktadır.

Sinapslar, BN'ler arası bağlantı ve iletişimin temelini oluřtururken, aynı zamanda öğrenme ve hafıza süreçlerinin merkezi bir parçasıdır. Presinaptik BN'de meydana gelen bir AP, farklı postsinaptik BN'lerde farklı şekillerde etkileşime yol açabilmektedir. Bu etkileşimler, SB'lerin gücünü ve etkinliğini göstermektedir. Bazı SB'lerin diğerklerine göre daha güçlü olması sinaptik plastisitenin varlığına işaret etmektedir. Sinaptik plastisite, SB'lerin gücünün ve yapısının zamanla deđişebilirliğini ifade etmektedir.

Yapılan arařtırmalar, öğrenme ve hafıza oluřumunda sinaptik plastisitenin önemini vurgulamaktadır (Magee & Grienberger, 2020). Buna göre, yoğun tekrar ile bir bilgi öğrenilirken, ilgili SB'ler güçlenmekte ve bilginin daha uzun süreli olarak hafızada kalmasını sağlanmaktadır. Öte yandan, bilginin zamanla kullanılmaması sonucu, ilgili SB'ler zayıflamakta ve bilgi unutulmaktadır.

Sinaptik plastisite, uzun süreli potansiyasyon (USP) ve uzun süreli depresyon (USD) olmak üzere iki ana türde gerçekleşmektedir (Bear & Malenka, 1994). Bir SB'nin



Şekil 5. Biyolojik nöronların morfolojik çeşitliliği (Georgiev vd., 2020)

etkinliğinin veya ağırlığının artması USP olarak tanımlanırken, bir SB'nin zamanla etkinliğinin azalması ise USD olarak isimlendirilmektedir.

USP genellikle yoğun tekrar ve öğrenme deneyimleri sonucunda ortaya çıkmaktadır. Bu süreç, uzun süreli hafızanın oluşumuna katkıda bulunmaktadır. Öte yandan, USD belirli bir SB'nin etkinliğinin zamanla azalmasıyla ilişkilidir. Bu süreç, gereksiz veya az kullanılan SB'lerin zayıflatılmasını ve bilgilerin unutulmasını sağlamaktadır. Bu iki plastisite türü, BN'lerin ve BSA'ların çevresel değişikliklere adaptasyonunu sağlayarak öğrenmeyi sağlamaktadır.

1.1.5. Biyolojik Nöronlarda Dendritik Yapı

BN'ler vücutta bulunan diğer hücrelerden farklı olarak, zamanla değişen morfolojilere ve çeşitli biyofiziksel özelliklere sahip olabilmektedir. Gerek BN'lerin birbirlerine bağlanmak için fiziksel olarak yakınlaşmak zorunda olmaları gerekse hücre zarında bulunan iyon kanallarının BN boyunca heterojen bir dağılım göstermesi, BN'lerin morfolojik ve biyofiziksel çeşitliliğinin temel nedenleri arasındadır.

BN'ler, beyin boyunca çeşitli katmanlarda bulunan BN'ler ile iletişim haline olabilmektedirler. BN'lerin beyinde mekânsal olarak dağınık durumları, BN'ler arası

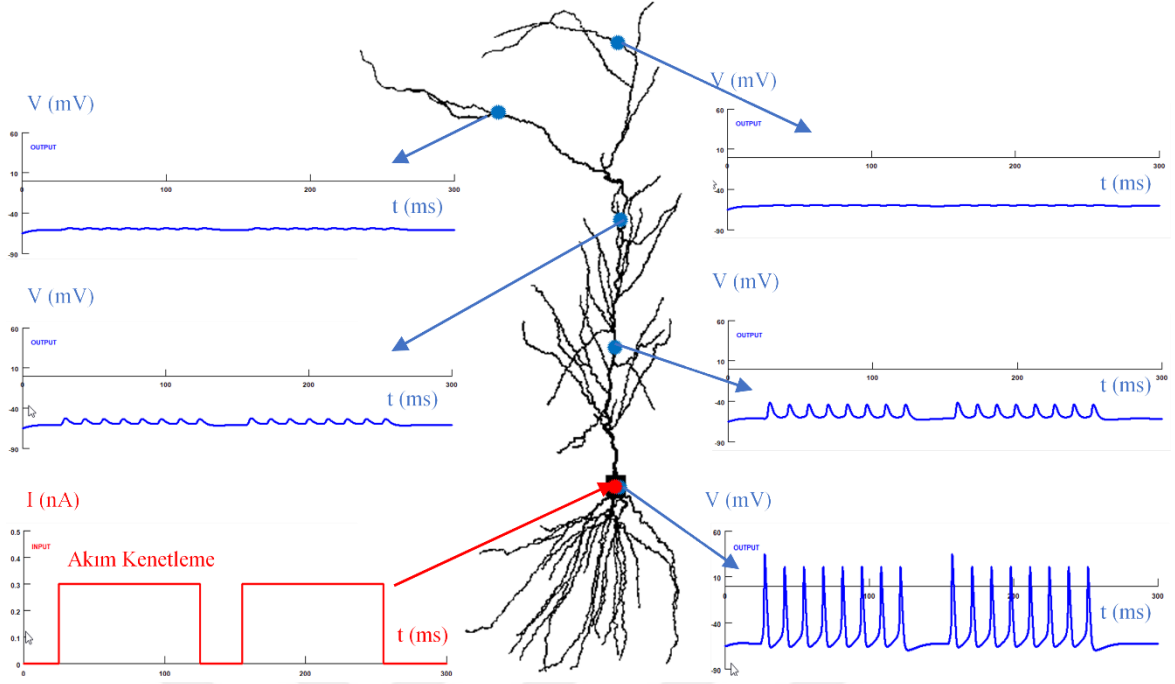
bağlantılar için birbirlerine doğru fiziksel ulaşım zorunluluğunu beraberinde getirmektedir. Bu nedenle, zamanla formlarını değiştirebilen dendritik kıvrımlara sahiptirler. Dendritler BN'lerin bir nevi kolları olarak bilinmekte ve gerek çeşitli katmanlara ulaşmaları gerekse her katman için çeşitli dinamiklere ihtiyaç duymaları nedeniyle özgün morfolojik ve biyofiziksel özelliklere sahiptirler (Kath, 2005). Her BN sorumlu olduğu görevden ötürü kendine özgü morfoloji ve biyofiziğe sahiptir. Şekil 5'te BN'lerin morfolojik çeşitlilikleri gösterilmektedir (Georgiev vd., 2020). Şekil 5'te gösterilen BN'lerin morfolojik çeşitliliği ve özgünlükleri BSA'ları güçlü yapan yönlerden biri olarak kabul edilmektedir (Jones vd., 2009).

BN'lerin morfolojik çeşitliliklerinin yanında bir diğer ayırıcı özellikleri biyofizikleridir. İyon kanalları BN boyunca homojen dağılıma sahip değildir. Bunun nedeni, beynin her katmanının kendine özgü bir iyon konsantrasyona sahip olmasıdır. Ayrıca, beyin içinde bulunan BSA'lar belirli görevlerde uzmanlaşacak şekilde evrilmektedirler. Bu nedenle, BSA içinde bulunan BN'ler uzmanlaştıkları konuya özgü iyon kanalları ve biyofiziksel özelliklere sahip olmalıdırlar. Tüm bunlar göz önünde bulundurulduğunda BN'lerin morfolojik ve biyofiziksel çeşitlilikleri ayrıca araştırma konusu olmaktadır. Araştırmalara göre beyinde en çok bulunan BN morfolojileri PN'lerdir (Spruston, 2008).

1.1.6. Piramidal Nöronlar

PN'ler, beynin dış kabuğunun temel bileşenlerinden biridir. Özellikle bilgi işleme ve bilişsel fonksiyonlar üzerinde önemli bir etkiye sahiptir. PN'ler, piramit şeklindeki somaları, apikal dendritleri ve basilar dendritik ağaçları ile spesifik bir forma sahiptir. Geniş SB ağları ve çeşitli elektrofizyolojik özellikleri sayesinde kortikal fonksiyonlarda önemli görevler üstlenmektedirler (Mihaljević vd., 2021).

PN'ler, duyuşsal algıdan motor kontrolüne kadar çeşitli bilişsel süreçlere katılımlarıyla bilinmektedirler (Mihaljević vd., 2021). Çalışmalar, bireysel PN'lerin dendritik ve aksonal mimarisinin kortikal katmanlara, türlere ve beyin bölgelerine göre değiştiğini göstermiştir (Mohan vd., 2015). PN çeşitliliğinin anlaşılması, BSA'ların ve kortikal alanların fonksiyonel uzmanlaşmasının anlaşılmasına katkıda bulunmaktadır (Luebke, 2017).



Şekil 6. Piramidal nöronda dendritik zayıflama (Özcan vd., 2018)

Çalışmalar, PN'lerin dendritik AP'ler ve dendritik sodyum spike oluşumu gibi karmaşık elektrofizyolojik özellikleri sergilediğini göstermektedir (Gidon vd., 2020; Gooch vd., 2022). PN'lerin dendritik AP oluşturma yetenekleri ve dendritik ağaçlarında zar potansiyellerinin mesafeye bağlı zayıflaması, karmaşık sinyal işleme kapasitelerini yansıtmaktadır (Gooch vd., 2022). Şekil 6'da bir PN'nin somasında meydana gelen AP'nin dendritler üzerinde zayıflaması gösterilmiştir (Özcan vd., 2018).

PN'lerin bilişsel süreçlere katılımları ve nörolojik bozukluklardaki rolleri önemlerini daha da vurgulamaktadır. Kronik şizofrenide PN'lerin transkripsiyonel profilinin değiştiğini gösteren çalışmalar, PN'lerin hastalık patofizyolojisindeki ne denli etkin olduğunu göstermektedir (Wu vd., 2020). Ayrıca, Alzheimer hastalığı gibi nörodejeneratif hastalıklarda belirli PN türlerinin kaybolması, PN'lerin hastalık patolojisindeki rolünün anlaşılmasının önemini vurgulamaktadır (Singleton vd., 2022).

1.1.7. Biyolojik Nöronlarda Nöral Kodlama

BN'ler görevlerini, sinir sisteminin temel iletişim mekanizması olan AP'ler (spike) aracılığı ile gerçekleştirmektedirler. Bu bağlamda, çevresel uyarıcıların analog formdaki

verileri, beyin tarafından işlenebilir hale getirmek için spike dizilerine dönüştürülmesi gereklidir.

Görsel algılama sürecinde, retinanın fotoreseptör hücreleri tarafından algılanan analog ışık ve yoğunluk değerleri, kompleks nöral devreler tarafından spike dizilerine kodlanarak beyne iletilmektedir. Bu dönüşüm süreci, çevresel ışık yoğunluğu, rengi ve hareketini temsil eden spesifik nöral paternlerin oluşturulmasıyla gerçekleşmektedir (Milosavljevic vd., 2018).

Koku algısında ise, çevredeki moleküllerin doğrudan anlamlı bilgi sağlamadığı görülmektedir. Bunun yerine, koku moleküllerinin özgül reseptör hücreleri tarafından algılanması ve bu bilginin nöronal ağlar içerisinde spike dizileri şeklinde kodlanmasıyla, beyin farklı kokuları ayırt edebilir hale gelmektedir (Seshadri vd., 2022).

İşitme sistemi için benzer bir mekanizma söz konusudur. Analog ses dalgaları, iç kulağın koklea yapısında bulunan kıl hücreleri tarafından algılandığında, bu mekanik uyarılar nöral reseptörler tarafından spike dizilerine dönüştürülmektedir. Bu sayede, sesin frekansı, şiddeti ve yönü gibi özellikleri, beyin tarafından anlamlı işitsel deneyimlere dönüştürülmektedir (Ishimoto vd., 2015).

BN'ler, çevresel dünyadan aldıkları analog sinyalleri spike dizisi formatına dönüştürerek, beyin tarafından işlenebilir ve yorumlanabilir bilgilere dönüştürme yeteneğine sahip olmalıdır. Bu mekanizma, sinir biliminin karmaşık ve etkileyici bir yönünü oluşturmaktadır. Çünkü çevresel verilerin nöral kodlaması, insan deneyimini şekillendirmekte ve anlamlandırmaktadır.

1.2. Yapay Sinir Ağları

YSA'ların evrimi, insan beyninin karmaşık yapısından ilham alarak başlamış ve matematiksel modellerle ve bilgisayar bilimleriyle birleşerek gelişmiştir. Bu gelişim süreci, araştırmacıların BN ve SB'lerin işleyişini anlamaya yönelik yıllarca süren çalışmalarıyla şekillenmiştir. YSA'ların temelini oluşturan bu YN ve SB'ler, beyin ve sinir sisteminin işleyişini taklit eden yapay yapılar oluşturmak için kullanılmıştır. Wang ve arkadaşlarının (2020) gerçekleştirdiği kapsamlı çalışma, YSA'ların bu uzun evrim sürecini üç ana nesil üzerinden detaylı bir şekilde incelemiştir. Bu nesiller, YSA'ların farklı özellikler kazanması, daha karmaşık hale gelmesi ve teknolojik olarak daha gelişmiş çözümler sunabilmesi açısından önemli dönüm noktalarını temsil etmektedir.

McCulloch ve Pitts'in (1943) tanıttığı ilk YN modeli, nörobilim ve YZ alanlarında devrim niteliğinde bir adım olarak kabul edilmektedir. Bu model, BN'lerin basit bir matematiksel temsilini oluşturarak, araştırmacılara nöronların ve sinir ağlarının karmaşık işleyişini anlamaları için bir çerçeve sunmuştur. Bu YN modeli, girdi olarak aldığı sinyalleri, belirli ağırlıklarla çarparak toplamaktadır. Ağırlıklar, girdi sinyallerinin önem derecesini ve etkisini belirleyen parametrelerdir. McCulloch ve Pitts (1943), bu ağırlıkların doğru bir şekilde ayarlanmasının, YN'nin istenen çıktıyı üretmesini sağlayacağını vurgulamışlardır.

McCulloch ve Pitts'in (MP) YN modeli, nöral ağların ve YZ modellemesinin temelini atmış olsa da SB ağırlıklarının nasıl ayarlanacağına dair spesifik bir eğitim mekanizması sunmamaktaydı. Bu eksiklik, modelin gerçek dünyadaki problemleri çözebilmesi önünde bir engel teşkil etmekteydi. Donald Hebb (1949) çalışmasında, YN'lerin eşzamanlı olarak ateşlemesi durumunda, SB'lerin güçlendiğini ve bu SB'lerin daha da güçlü hale gelebileceğini öne süren bir öğrenme kuralı ortaya atmıştır. Hebbian öğrenme kuralı, YSA'ların eğitiminde ağırlıkların nasıl güncelleneceğine dair ilk temel yaklaşımlardan biridir.

Frank Rosenblatt (1958), perceptron adında bir model tanıtarak, bu modelin eğitimi için spesifik bir öğrenme kuralı önermiştir. Perceptron, girdileri alıp çıktı üreten ve ağırlıklarını güncelleyen bir YN modelidir. Perceptron öğrenme kuralı, bir hata fonksiyonunu minimize ederek ağırlıkların güncellenmesini sağlamaktadır. Bu yaklaşım, YSA'nın eğitiminde ve öğrenmesinde önemli bir adım olmuştur. Perceptron öğrenme kuralı, YSA'ların ve YZ'nin gelişiminde önemli bir dönüm noktası olmuştur. MP nöron modeli ve perceptron öğrenme kuralının birleşimi, birinci nesil sinir ağlarının temelini oluşturmuştur. Bu birleşim, YN'lerin temel işlevlerini ve öğrenme mekanizmalarını tanımlayarak, ilk YSA'ların oluşturulmasını sağlamıştır.

Birinci nesil sinir ağları, basit yapılarıyla birçok temel problemi başarıyla çözebilmiştir. Ancak, XOR problemi gibi bazı spesifik ve doğrusal olmayan problemler karşısında yetersiz kalmışlardır. XOR problemi, tek bir katman kullanılarak doğru bir şekilde çözülememektedir. Bu sınırlamalar, tek katmanlı sinir ağlarının doğrusal olarak ayrılabilen problemler dışındaki karmaşık yapıları çözmedeki kısıtlamalarını ortaya koymuştur.

Bu kısıtlamaların farkında olan araştırmacılar, ikinci nesil sinir ağlarına, yani çok katmanlı YSA'lara yönelmeye başlamışlardır. Çok katmanlı YSA'lar, birden fazla katman

ve bu katmanlar arasında gelişmiş öğrenme algoritmaları ile daha karmaşık ve doğrusal olmayan problemleri modelleyebilme yeteneğine sahip olmuştur. Ancak perceptron öğrenme kuralı çok katmanlı sinir ağlarının eğitilmesinde yetersiz kalmıştır. Bu nedenle geri yayılım algoritması ikinci nesil sinir ağlarının eğitilmesinde tercih edilmeye başlanmıştır (Baldi, 1995). Ancak, MP nöron modeli geri yayılım algoritması için yetersiz kalmaktaydı. MP nöron modelinin klasik ikili çıkış üretme kapasitesi türevlenebilir bir fonksiyon olmadığından, bu model aktivasyon fonksiyonları gibi türevlenebilir transfer fonksiyonları ile genişletilmiştir. Geri yayılım algoritmasının ortaya çıkışı bu bakımdan, YSA'ların eğitim metodolojisinde devrim yaratan bir başka adım olmuştur. Bu algoritmanın entegrasyonu, çok katmanlı perceptron yapılarının karmaşık ve doğrusal olmayan problemler üzerindeki etkinliğini artırmış, özellikle XOR gibi zorlu problemleri çözme kapasitesini önemli ölçüde yükseltmiştir (Lisboa & Perantonis, 1991).

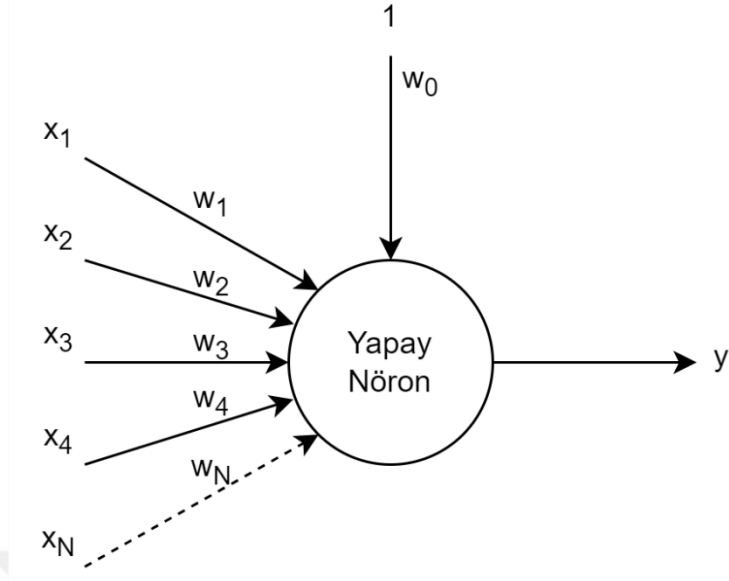
Bu gelişmelerin ötesinde, aktivasyon fonksiyonlarının dahil edilmesiyle, oluşturulan YSA'lar biyolojik açıdan daha gerçekçi bir yapı kazanmıştır. Bu yapılar, birinci nesil sinir ağlarına göre daha ileri seviyede bir modelleme ve öğrenme yeteneği sergilemiştir. Bugün bilinen DSA'lar ve mimarileri de ikinci nesil sinir ağlarının en etkin uygulamaları olarak ortaya çıkmaktadır (Zhou, 2021).

Bu bölümde, YSA'lar genel olarak tanıtılmıştır. Bölüm iki alt başlıktan meydana gelmektedir. İlk alt başlıkta, YN modeli ve aktivasyon fonksiyonları tanıtılmıştır. İkinci alt başlıkta ise YSA'larda öğrenme mekanizmalarına değinilmiştir.

1.2.1. Yapay Nöron Modeli

MP nöron modeli (McCulloch & Pitts, 1943), sinirbilim ve YZ disiplinlerinde önemli bir rol oynamakta ve literatürde geniş kabul görmektedir. BN'lerin işlevselliğinden ilham alınarak geliştirilen bu model, YN'lerin hesaplamalı temsillerini sunmaktadır. MP modelinin dikkate değer bir karakteristiği, BN'lerin karmaşık etkileşimlerini doğrusal eşik fonksiyonlarıyla basit ve matematiksel olarak ifade edebilme yeteneğine sahip olmasıdır. Bu matematiksel sadeleştirme yaklaşımı, YSA'nın evrimine katkı sağlamakta ve DÖ ile bilişsel bilim alanlarında bilimsel ilerlemelere katkıda bulunmaktadır (Özdemir, 2016).

Şekil 7'de verilen nöron yapısında, girdi vektörü $X = [1, x_1, x_2, x_3, x_3, \dots, x_N]$ ve ağırlık vektörü $W = [w_0, w_1, w_2, w_3, w_3, \dots, w_N]$ olarak tanımlanmaktadır. Bu kapsamda, YN'nin ağırlıklı toplam değeri Denklem (1)'de matematiksel olarak ifade edilmektedir.



Şekil 7. Yapay nöron modeli

$$y = W \times X^T = \sum_{i=1}^N w_i x_i + w_0 \quad (1)$$

Bu denklemin içeriği, YN'nin ateşleme mekanizmasını tam anlamıyla yansıtmamaktadır. Bu sebeple, y değeri, belirli bir aktivasyon fonksiyonu tarafından işlenmektedir. Söz konusu aktivasyon fonksiyonu, Denklem (2)'de matematiksel olarak tanımlanmıştır.

$$çıktı = \varphi(y) \quad (2)$$

Bu aktivasyon fonksiyonu, YN'nin çıktı değerini ve tetikleme eşliğini belirleyerek, YN'nin bir sonraki katmana veya çıkışa gönderilen sinyalin düzenlenmesini sağlamaktadır. Aktivasyon fonksiyonları, spesifik YN yapılarına bağlı olarak değişiklik gösterebilmekte ve bu fonksiyonlar, Denklemler (3), (4), (5) gibi matematiksel ifadelerle temsil edilebilmektedir.

$$\varphi_1(y) = \begin{cases} 1, & y \geq \theta \\ 0, & y < \theta \end{cases} \quad (3)$$

$$\varphi_2(y) = \frac{1}{1 + e^{-y}} \quad (4)$$

$$\varphi_3(y) = \begin{cases} y, & y \geq \theta \\ 0, & y < \theta \end{cases} \quad (5)$$

Denklem (3), MP nöron modeline özel aktivasyon fonksiyonunu ifade ederken, Denklem (4) sigmoid fonksiyonunu, Denklem (5) ise doğrultulmuş lineer birim (ReLU) aktivasyon fonksiyonunu temsil etmektedir. Bu denklemler, YN modellerinin temel işlevselliğini ve karakteristik özelliklerini belirten matematiksel ifadelerdir.

Aktivasyon fonksiyonu, aslında bir BN'nin ateşleme veya ateşlememe davranışını temsil etmektedir. Ağırlıkların ve eşik değerinin belirlenmesi, MP nöron modeline dayalı YSA'ların öğrenme süreçleri ve karar verme mekanizmaları için önemlidir. YN, sinirbilim araştırmalarında önemli bir paradigmaya dönüşmüş ve beyindeki nöronal işleyişin temelini sunmuştur. YZ literatüründe YN, özellikle örüntü tanıma, karar destek mekanizmaları ve MÖ gibi çeşitli disiplinlerde kullanım alanı bulmuştur (Zhou, 2021). Ayrıca, YN'nin teorik sinirbilim ve bilgisayar bilimi alanlarında kullanılması, disiplinler arası uygulama potansiyelini de gözler önüne sermektedir.

1.2.2. Yapay Sinir Ağlarında Öğrenme

YSA'larda öğrenme, YZ ve DÖ alanlarının temel taşıdır. Bu özelliği sayesinde sistemler, verilerden bilgi çıkarabilmekte, yeni görevleri öğrenebilmekte ve optimize edilebilmektedir. Öğrenme mekanizmaları, YSA'ların adaptif ve esnek olmasını sağlamaktadır. Böylece değişen koşullara ve veri kümelerine (VK'lere) uyum sağlayabilmektedirler. Öğrenme süreci, modelin hatalarını azaltma, doğruluk oranını artırma ve genelleme yeteneğini geliştirme kapasitesiyle, YSA'ların performansını sürekli olarak iyileştirmektedir. Ayrıca öğrenme; YSA'ların karmaşık görevleri, örüntüleri ve ilişkileri tanıma ve kavrama yeteneğini artırmaktadır. Bu da gerçek dünya uygulamalarında, bilgi işleme, tahminleme ve karar verme süreçlerinde daha etkili ve doğru sonuçlar elde etmeyi sağlamaktadır (Zhou, 2021).

Öğrenme süreci, ağırlık matrisinin belirlenen bir optimum değere ulaşması amacıyla uygulanan adımlar dizisidir. Temel yaklaşım, her bir ağırlığın değişimini hesaplamak ve bu hesaplamaları temel alarak ağırlıkların güncellenmesini sağlamaktır. Bu süreç, Denklem (6)'da modellenmiştir.

$$w_i^{Yeni} \leftarrow w_i^{Eski} + \Delta w_i \quad (6)$$

Burada Δw_i , i 'inci ağırlığın değişimini; w_i^{Eski} , ağırlığın güncellenmeden önceki değerini ve w_i^{Yeni} ise ağırlığın güncelleme sonucu aldığı değeri sembolize etmektedir. Öğrenme yöntemleri temel olarak Δw_i değerinin nasıl hesaplanacağına dair yöntemlerdir.

Hebbian öğrenme paradigması, nörobiyolojik temellere sahip bir öğrenme teorisi olarak kabul edilmektedir ve Donald Hebb (1949) tarafından ortaya atılmıştır. Bu teori, BN'lerin eşzamanlı olarak aktivasyon gösterdiklerinde aralarındaki SB'nin kuvvetlenmesi prensibine dayanmaktadır. Temelde, bir BN diğerini aktive ettiğinde ve bu eşzamanlı aktivasyon tekrar edildiğinde, aralarındaki SB'de artan bir güçlenme meydana gelmektedir. Bu paradigmanın beyin temeli öğrenme mekanizması olduğu kabul edilmektedir. Denklem (7), Hebbian öğrenme mekanizması çerçevesinde ağırlık değişiminin nasıl hesaplandığını ifade eden matematiksel formülü ifade etmektedir.

$$\Delta w_i = \eta \times çıktı \times x_i \quad (7)$$

Bu denklemde; Δw_i , ağırlık değişimini; x_i , i 'inci girdinin değerini; $çıktı$, YN çıktısını ve η ise öğrenme hızını temsil etmektedir.

Hebbian öğrenme, özünde bir öğrenme metodu olarak kabul edilmesine rağmen, danışmansız bir öğrenme yöntemi olarak sınırlı kalmıştır. Denklem (7), belirli bir hedef çıktıya yönelik bilgi içermemektedir. Bu yüzden danışmanlı öğrenme senaryolarında eksik kalmaktadır.

Rosenblatt (1958) tarafından tanıtilen perceptron öğrenme kuralı, danışmanlı öğrenme problemlerinde etkili bir biçimde kullanılabilir. Bu öğrenme kuralı, bir YN'nin çıktısı ile beklenen çıktı arasındaki farkın en aza indirgenmesi ilkesine dayanmaktadır. Ağırlık güncellemesi de bu farkın temel alındığı bir hesaplamayla gerçekleştirilmektedir ve bu süreç Denklem (8)'de matematiksel olarak gösterilmektedir.

$$\Delta w_i = \eta \times (d - çıktı) \times x_i \quad (8)$$

Bu denklemde; Δw_i , ağırlık değişimini; x_i , i 'inci girdinin değerini; $çıktı$, YN'nin çıktısını; d , YN'nin beklenen çıktısını; η ise öğrenme hızını ifade etmektedir.

Perceptron öğrenme kuralı, YSA'ların eğitiminde uzun yıllar boyunca temel bir yöntem olarak kabul edilmiştir. Ancak, doğrusal olarak ayıramayan VK'ler ve çok katmanlı sinir ağlarının eğitiminde sınırlılıkları olduğu tespit edilmiştir. Bu tespitler sonrasında, geri yayılım algoritması perceptron öğrenme kuralının evrimine katkıda bulunarak bu sınırlılıkları aşmayı mümkün kılmıştır. Bu algorithmada, ağırlık güncellemesi, belirli bir hata fonksiyonunun minimize edilmesi prensibine dayanmaktadır. Bu prensibi matematiksel olarak ifade eden Denklem (9) aşağıda verilmektedir.

$$\Delta w_i = -\eta \times \frac{\partial E}{\partial w_i} \quad (9)$$

Bu denklemde; Δw_i , ağırlık değişimini; η , öğrenme hızını; $\frac{\partial E}{\partial w_i}$ ise hatanın i 'inci ağırlığa göre türevini temsil etmektedir. Hata değeri YN'lerde, Denklem (10)'da tanımlandığı şekliyle hesaplanmaktadır.

$$E = \frac{1}{2} (d - out)^2 \quad (10)$$

Bu denklemde; E , hatayı; $çıkıtı$, YN'nin çıktısını; d ise beklenen çıktı değerini simgelemektedir.

Denklem (9) için $\frac{\partial E}{\partial w_i}$ ifadesini elde etmek adına, hata değerinin belirli bir ağırlığa göre türevinin hesaplanması gereklidir. Ancak, hata $çıkıtı$ ve d arasındaki bir fonksiyon olarak belirtilmiştir. Dolayısıyla doğrudan ağırlığın fonksiyonu değildir. Bu tür bir türevin hesaplanması için zincir kuralı veya zincir yöntemi kullanılmaktadır. Bu yaklaşım, hatanın türevini doğrudan ağırlığa göre hesaplamayı mümkün kılmaktadır ve yöntem Denklem (11)'de sunulmuştur.

$$\frac{\partial E}{\partial w_i} = \frac{\partial E}{\partial \varphi(y)} \times \frac{\partial \varphi(y)}{\partial y} \times \frac{\partial y}{\partial w_i} \quad (11)$$

Bu denklemde; $\varphi(y)$, YN'nin çıktısını; y , YN'nin aktivasyon fonksiyonuna girmeden önceki değerini; w_i ise ağırlığı temsil etmektedir. Tüm terimlerin türevleri alındığında, bu süreci açıkça gösteren Denklem (12) elde edilmektedir.

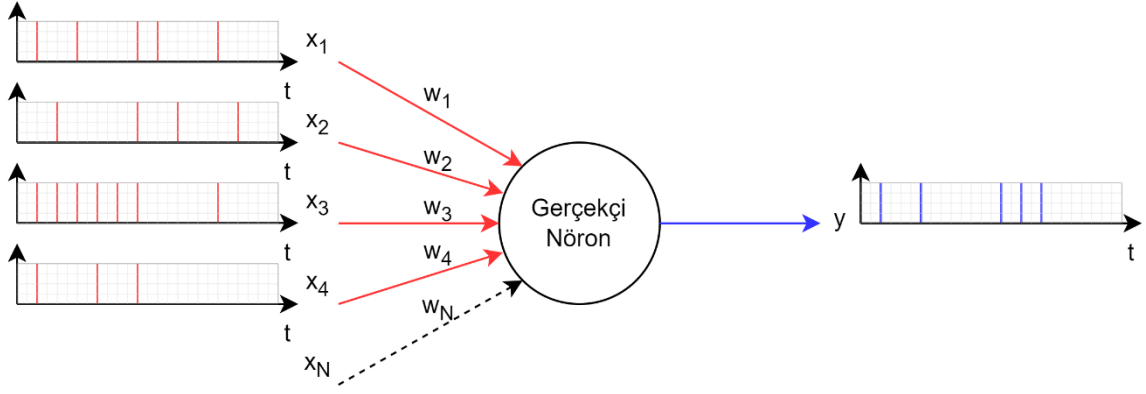
$$\frac{\partial E}{\partial w_i} = (çıktı - d) \times (çıktı \times (1 - çıktı)) \times x_i \quad (12)$$

Denklem (12) sayesinde $\frac{\partial E}{\partial w_i}$ ifadesi; YN'nin girdisi x_i , YN'nin çıktısı $çıktı$ ve beklenen çıktı değeri d ile ifade edilebilir hale gelmiştir. Denklem (12) incelendiğinde, $(çıktı \times (1 - çıktı))$ ifadesinin sigmoid aktivasyon fonksiyonunun türevi olduğu anlaşılmalıdır. Bu ifade değiştirilerek, başka bir aktivasyon fonksiyonunun türevi kullanılarak yöntem genişletilebilmektedir. Önemli bir not olarak, Denklem (3)'te tanımlanan MP nöronunun aktivasyon fonksiyonunun türevi tüm girdi değerleri için "0" olacağından, bu türev değeri geri yayılım algoritması için uygun değildir. Bu durum, MP nöronunun eğitiminde geri yayılım algoritmasının kullanılamayacağı anlamına gelmektedir.

1.3. Gerçekçi Sinir Ağları

Sinir ağlarının evrim süreci içerisinde, ilk ve ikinci nesil sinir ağlarının bazı kısıtlamaları ve dezavantajları ortaya çıkmıştır. Örneğin, ilk nesil sinir ağları, basit yapıları nedeniyle karmaşık ve doğrusal olmayan problemleri etkili bir şekilde çözmede sınırlı kalmışlardır. İkinci nesil sinir ağları, çok katmanlı yapılara sahip olmalarına rağmen, eğitim ve optimizasyon süreçlerinde çeşitli zorluklarla karşılaşmışlardır. Özellikle, geri yayılım algoritması gibi tekniklerin eksiklikleri, çok katmanlı perceptron yapılarının eğitilmesini karmaşık hale getirmiştir. Ayrıca, bu yapılar aşırı öğrenme eğilimli olabilmektedir. Bu nedenle, birinci ve ikinci nesil sinir ağlarının genelleme yeteneklerinin kısıtlı olduğu söylenebilmektedir (Hawkins, 2004). Bu bağlamda, üçüncü nesil sinir ağları olan GSA yaklaşımı, bu dezavantajlara çözüm sunmayı hedeflemiştir. GSA, BN'lerin gerçek zamanlı tepkilerini ve dinamiklerini simüle ederek, YZ sistemlerine daha gerçekçi, esnek ve etkili bir çözüm sunma potansiyeline sahiptir.

Geleneksel YN modellerinin aksine, gerçekçi nöronlar (GN'ler), BN'lerin gerçek zamanlı AP'lerini (spike) Şekil 8'de gösterildiği gibi taklit ederek işlem yapma kapasitesine sahiptir. Spike işleme tabanlı yaklaşım, BSA'ların dinamik tepkilerini, sinaptik dinamikleri ve zaman bağımlı etkileşimleri daha etkili bir şekilde simüle etmeye olanak tanımaktadır. Tüm bu özellikler GSA'ları biyolojik gerçekçilik ve işlevsellik konusunda ön plana çıkarmaktadır.



Şekil 8. Gerçekçi nöron modeli

GN'ler, geleneksel YN modellerine kıyasla bir dizi üstünlüğe sahiptir. Öncelikle, GN'ler gerçek zamanlı işlem kapasitesi sunması, özellikle zamanın önemli olduğu uygulamalarda ve gerçek zamanlı veri işleme gereksinimlerinde avantajlar sağlamaktadır (Lobo vd., 2020). Ayrıca, enerji verimliliği açısından GN'ler, AP üretme sürecinde daha az enerjiye ihtiyaç duymaktadır. Bu yönüyle GN'lerin enerji sınırlı sistemler ve mobil cihazlar için önemli bir potansiyeli olduğu vurgulanmaktadır (Han & Roy, 2020).

Bu bölümde, GSA'lar ve modelleme yaklaşımları genel olarak tanıtılmıştır. Bölüm sekiz alt başlıktan meydana gelmektedir. İlk alt başlıkta, Nernst denklemi ve Goldman-Hodgkin-Katz (GHK) denklemi tanıtılmıştır. İkinci alt başlıkta, iyon kanallarının modellenmesine dair yaklaşım açıklanmıştır. Üçüncü, dördüncü ve beşinci alt başlıklarda, literatürde en çok kullanılan basit GN modelleri olarak, sırasıyla Sızıntılı Topla&Ateşle (STA), Hodgkin&Huxley (HH) ve Izhikevich (IZH) nöron modelleri tanıtılmıştır. Altıncı alt başlıkta, GSA'larda sinaptik mekanizma ve modeller tanıtılmıştır. Yedinci alt başlıkta, GN'lerin modellenmesinde bölme yaklaşımı anlatılmıştır. Sekizinci ve son alt başlıkta ise GN'lerde nöral kodlama hakkında yaklaşımlardan bahsedilmiştir.

1.3.1. Nernst Denklemi ve Hücre Zarının Pasif Modeli

Nernst denklemi, elektrokimyasal bağlamda, yarı geçirgen bir zar üzerinde bir iyonun denge potansiyelini ifade etmektedir. Bu denklem, BN'lerin iyonik mekanizmalarının derinlemesine anlaşılmasında merkezi konumdadır. Denklemin tanımladığı denge potansiyelleri, hücre zarında bulunan Na^+ , K^+ , Cl^- ve Ca^{2+} gibi spesifik iyonlar için matematiksel bir açıklama ifade etmektedir (del Olmo vd., 2020).

Şekil 1, hücre içi ve hücre dışı ortamdaki iyonların nasıl konumlandığını görselleştirmektedir. Bu görselden anlaşıldığı üzere, dinlenme potansiyeli sırasında iyonlar, hücre içi ve dışı arasında elektrokimyasal bir dengeye sahip olacak şekilde dağılım göstermektedir. Bir iyon türü için Nernst denklemi, Denklem (13)'te verilmiştir.

$$E_{iyon} = \frac{RT}{zF} \ln \left(\frac{[İyon]_o}{[İyon]_i} \right) \quad (13)$$

Bu bağlamda; E_{iyon} , iyonun denge potansiyelini; R , ideal gaz sabitini; T , mutlak sıcaklığı; z , iyonun elektrik yükünü; F , Faraday sabitini ve $[İyon]_i$ ile $[İyon]_o$ sırasıyla hücrenin içi ve dışındaki iyon konsantrasyonlarını temsil etmektedir.

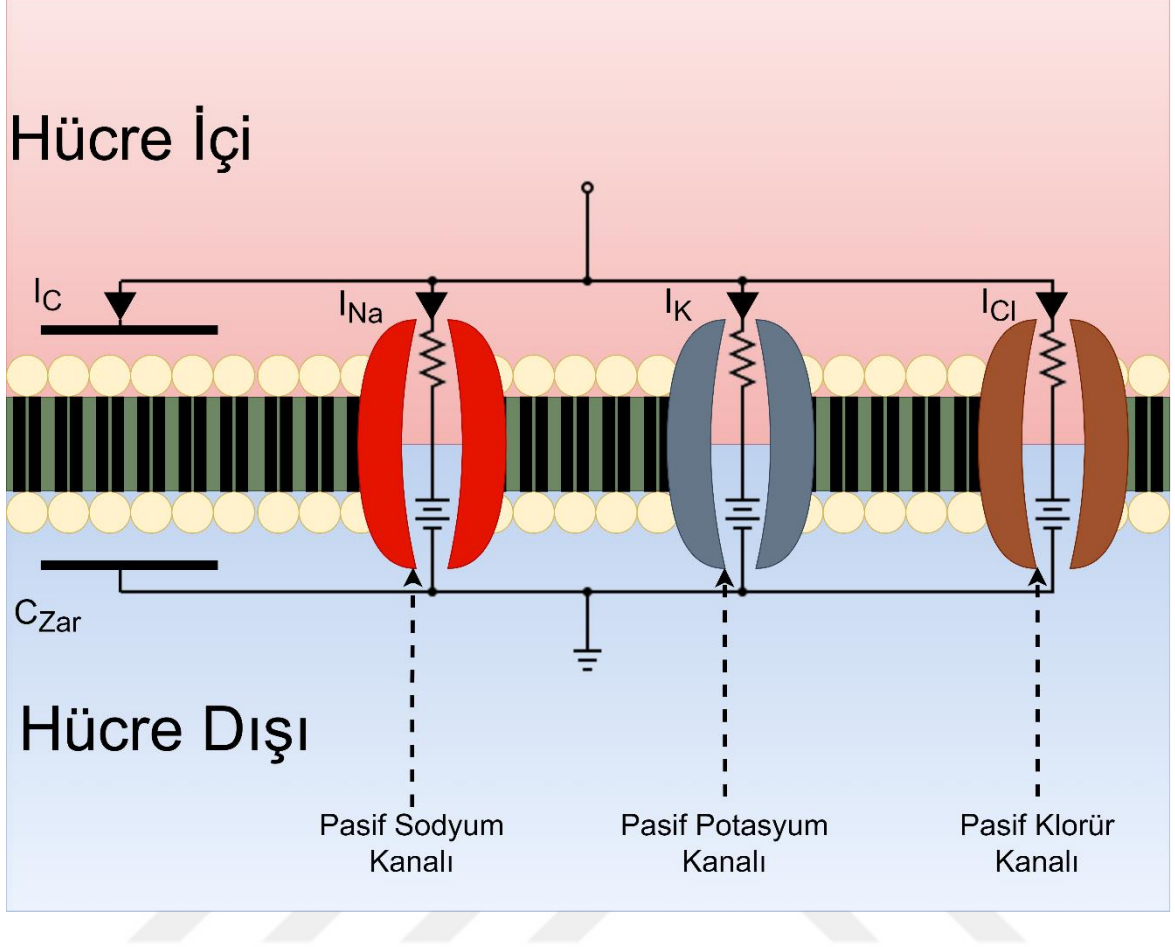
Nernst denklemi, kendi başına bir iyon türü için denge potansiyelini ifade etse de tüm iyonların hücre zarındaki toplu etkisini ifade etmede yetersiz kalmaktadır. Bu nedenle, tüm iyonların bir arada göz önünde bulundurulduğu durumda Nernst denklemi genişletilmekte ve Denklem (14)'te verilen GHK denklemi elde edilmektedir.

$$V_{zar} = \frac{RT}{F} \ln \left(\frac{P_K[K^+]_o + P_{Na}[Na^+]_o + P_{Cl}[Cl^-]_o}{P_K[K^+]_i + P_{Na}[Na^+]_i + P_{Cl}[Cl^-]_i} \right) \quad (14)$$

Bu bağlamda; V_{zar} , hücre denge potansiyelini; R , ideal gaz sabitini; T , mutlak sıcaklığı; P_{iyon} , iyonun zar geçirgenliğini; F , Faraday sabitini ve $[İyon]_i$ ile $[İyon]_o$ sırasıyla hücrenin içi ve dışındaki iyon konsantrasyonlarını temsil etmektedir. GHK'de tanımlanan denge potansiyeli, nöronal hücre zarı boyunca spesifik iyonların konsantrasyon farkı ile ilişkilendirilmektedir. Bu konsantrasyon farkları, nöronların elektrokimyasal özelliklerinin ve pasif iyonik taşıma mekanizmalarının anlaşılmasında fayda sağlamaktadır.

Pasif iyon kanalları, hücre zarında iyonların konsantrasyon farkı yönünde hareketini sağlayan kanallardır. Enerji gereksinimine sahip olmadan iyon transferini gerçekleştiren bu kanallar için Şekil 1'de verilen temsili gösterimden yola çıkıldığında elde edilen iyon kanallarına ait elektriksel eşdeğer Şekil 9'da verilmiştir.

Şekil 9'da verilen eşdeğer devre için, hücre zarının yalıtkan olmasından kaynaklı bir potansiyel farkın ortaya çıktığı göz önünde bulundurulduğunda, kapasitif bir eleman hücre zarının karakteristiğinin modellenmesi için uygun olmaktadır. Pasif iyon kanalları, her iyon transferi için spesifik olmak üzere iyonların zar geçirgenliğini temsilen bir iletkenlik



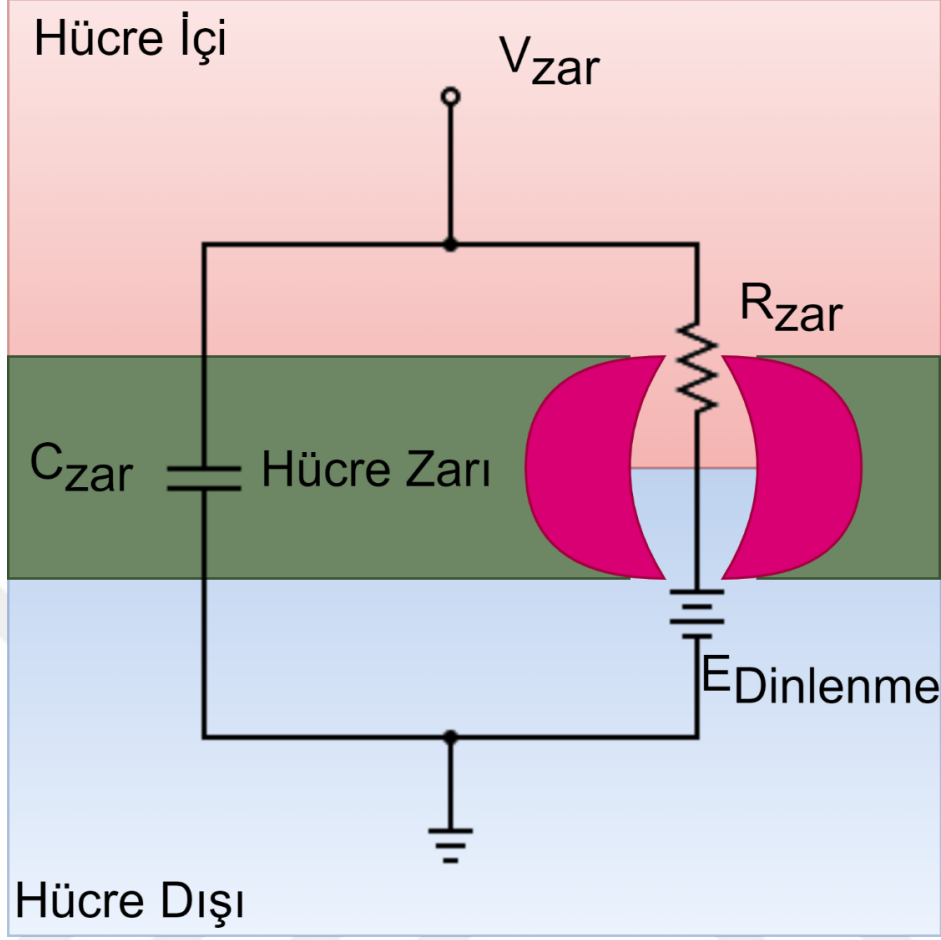
Şekil 9. Pasif iyon kanallarının hücre zarında elektriksel eşdeğer devresi

elemanı, iyon denge potansiyelini temsilen bir gerilim elemanı ile temsil edilmektedir. Eşdeğer devrenin analizi yapıldığında Denklem (15) elde edilmektedir.

$$I_C = -(I_{Na} + I_K + I_{Cl}) \quad (15)$$

Bu bağlamda; I_C , kapasitif akımı; I_{iyon} ise iyon kaynaklı pasif akımı ifade etmektedir. Bu akım ifadeleri, GHK denkleminde yola çıkılarak, kanal iletkenlik elemanları ve iyonlara özgü denge potansiyelleri cinsinden ifade edilebilecek şekilde genişletilebilmektedir. Akım ifadeleri içi değişkenler genişletildiğinde Denklem (16) elde edilmektedir.

$$C_{zar} \frac{dV_{zar}}{dt} = -g_{Na}(V_{zar} - E_{Na}) - g_K(V_{zar} - E_K) - g_{Cl}(V_{zar} - E_{Cl}) \quad (16)$$



Şekil 10. Pasif kanallar için indirgenmiş eşdeğer devre

Burada; C_{zar} , hücre zar sığasını; g_{iyon} , iyon iletkenliğini; E_{iyon} , ND ile tanımlanan iyon denge potansiyelini; V_{zar} ise zar potansiyelini ifade etmektedir.

Denklem (16)'da verilen ifadenin işlemsel yükünün ağır olması ve tüm kanalların benzer pasif etkiyi göstermesi nedeniyle Şekil 9'da verilen karmaşık yapının Şekil 10'da verilen indirgenmiş eşdeğer devre olarak ifade edilmesi mümkündür. Verilen indirgenmiş eşdeğer devre benzer şekilde analiz edildiğinde Denklem (17) elde edilmektedir.

$$C_{zar} \frac{dV_{zar}}{dt} = - \frac{V_{zar} - E_{Dinlenme}}{R_{zar}} \quad (17)$$

Bu bağlamda; R_{zar} , pasif kanallar için tek bir direnç ifadesini; $E_{Dinlenme}$ ise GHK denkleminde tanımlanan dinlenme potansiyelini temsil etmektedir. Bu değişkenler, Denklem (16)'da verilen değişkenler cinsinden ifade edilmek istendiğinde Denklem (18) ve (19) ortaya çıkmaktadır.

$$E_{Dinlenme} = \frac{g_{Na}E_{Na} + g_K E_K + g_{Cl}E_{Cl}}{g_{Na} + g_K + g_{Cl}} \quad (18)$$

$$R_{zar} = \frac{1}{g_{Na} + g_K + g_{Cl}} \quad (19)$$

Pasif iyon kanalları hücre zar potansiyelinin dinlenme anında bir dengede kalmasında görevlidirler. İyonlar pasif kanallar boyunca enerji bağımsız şekilde hareket etmektedirler. Bu durumun bir sızıntıyı andırması nedeniyle pasif kanal akımları aynı zamanda sızıntı akımları olarak da isimlendirilmektedir. Bu bağlamda, pasif kanallar hücre zarının denge sağlama araçlarıdır (Ermentrout & Terman, 2010).

1.3.2. Aktif İyon Kanallarının Modellenmesi

Aktif iyon kanalları hücre zarında ATP kullanmak suretiyle enerji bağımlı iyon akışlarının sağlandığı kanallardır. Literatürde birçok iyon kanalı tanımlanmış olmakla birlikte her iyon kanalı transferinden sorumlu olduğu iyonun akışını düzenlemek adına özgün iç dinamiklere sahiptir. Ermentrout ve Terman (2010) aktif iyon kanallarının Denklem (20)'de verilen ifadeyle genel olarak modellenebileceğini ifade etmektedirler.

$$I_{kanal} = \bar{g}_{kanal}([Ca^{+2}], ...)m^p n^q h^r (V_{zar} - E_{iyon}) \quad (20)$$

Bu bağlamda; I_{kanal} , iyon kanalının hücre zarında sağladığı akımı; $\bar{g}_{kanal}$, iyon kanalının maksimum iletkenliğini; m , n ve h değerleri gerilime bağlı kapı değişkenlerini; p , q ve r değerleri, deneysel olarak belirlenmiş tam sayı değerlerini; E_{iyon} , iyon kanalının sorumlu olduğu iyonun denge potansiyelini; $[iyon]$, iyon kanalının çalışmasını etkileyen bir başka iyon konsantrasyonunu; V_{zar} ise hücre zar potansiyelini temsil etmektedir. Denklem (20)'de verilen m , n ve h değişkenleri Denklem (21) veya Denklem (22)'de verilen ifadelerinden biriyle genel olarak ifade edilebilmektedir.

$$\frac{dx}{dt} = \alpha_x(1 - x) - \beta_x x \quad x = m, n, h \quad (21)$$

$$\frac{dx}{dt} = \frac{(x_{\infty} - x)}{\tau_x} \quad x = m, n, h \quad (22)$$

Burada; α_x , β_x , x_∞ ve τ_x değerleri hücre zar potansiyeline göre anlık değişen dinamikleri temsil etmektedir.

1.3.3. Sızıntılı Topla&Ateşle Nöron Modeli

STA nöron modeli (Lapicque, 1907), BN'lerin karmaşık davranışlarını matematiksel olarak basit bir formda ifade edebilme yeteneği sayesinde, sinirbilim ve YZ disiplinlerinde kullanılan önemli bir araçtır. Bu GN yapıları, zaman bağımlı sinaptik girdileri "toplama" (Integrate) kapasitesiyle tanımlanırken, hücre zar potansiyeli belirli bir eşiğe ulaştığında AP oluşturma (Fire) yeteneği sergilemektedirler. "Sızıntı" (Leak) kavramı, hücre zarının pasif elektrofizyolojik özelliklerini belirtmekte ve kademeli bir elektriksel sızıntı akımını mümkün kılmaktadır. Bu GN modeli, sinaptik sinyal entegrasyonu, AP üretimi ve yerel zar potansiyel farkının azalması gibi özellikleri nedeniyle bilimsel literatürde geniş bir ilgi alanı bulmuştur (Zhong vd., 2023).

Şekil 10'da sunulan eşdeğer devre modeli göz önünde bulundurulduğunda, hücre zarına dış kaynaklı bir akım uyarımı söz konusu ise ilgili ifadenin Denklem (23) olarak güncellenmesi mümkündür.

$$\tau_{zar} \frac{dV}{dt} = -(V - E_{sızıntı}) + R_{zar} I_{giriş} \quad (23)$$

Bu bağlamda; V , hücre zar potansiyelini; $E_{sızıntı}$, dinlenme hücre zar potansiyelini; R_{zar} , pasif kanal direncini; $I_{giriş}$, dış kaynaklı uygulanan akımı; $\tau_{zar} = R_{zar} \times C_{zar}$ ise hücre zarı zaman sabitini sembolize etmektedir.

Bu denklemle, hücrenin zar potansiyeli dış kaynaklı bir uyarı akımı ile uyarılabilir hale gelmektedir. Bununla birlikte, bu modeldeki hücre zarı yaklaşımı AP oluşturma mekanizmasını tam olarak açıklayamamaktadır. Bu nedenle; hücre zar potansiyelinin önceden tanımlanmış bir eşik değerini aştığında, GN'nin önce AP oluşturduğu; ardından belirlenmiş bir sıfırlama gerilimine geri döneceği varsayılmaktadır. Bu mekanizma Denklem (24)'te tanımlanan şartlar sağlandığında AP oluşturabilir hale gelmektedir.

$$V \leftarrow \begin{cases} V_{Tepe} & V > V_{Eşik} \\ V_{Sıfırlama} & V \geq V_{Tepe} \end{cases} \quad (24)$$

Bu bağlamda; V_{Tepe} , GN ateşlendiğinde ulaştığı maksimum zar potansiyelini; $V_{Eşik}$, GN'nin ateşlenmesi için aşması gereken kritik eşik değerini; $V_{Sıfırlama}$ ise GN'nin ateşlendikten sonra zar potansiyelinin sıfırlandığı gerilimi sembolize etmektedir. Ek olarak, bu denklemin refrakter periyodu temsil etmesi amacıyla bazı modellerde, GN'nin belirli bir zaman diliminde $V_{Sıfırlama}$ seviyesinde tutulması gerekmektedir. STA nöron modeli, AP oluşturma mekanizmasını tanımlamak için önemli bir araç olarak kabul edilmekte ve hesaplamalı sinirbilim literatüründe geniş bir kabul görmektedir.

1.3.4. Hodgkin&Huxley Nöron Modeli

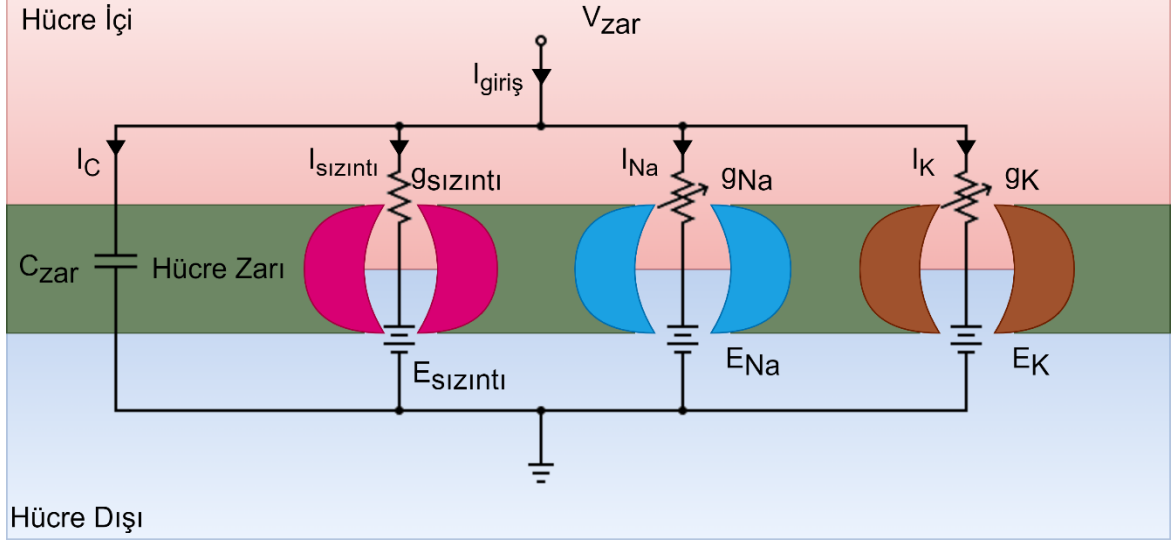
HH nöron modeli (Hodgkin & Huxley, 1952), sinirbilimde kendine geniş kullanım alanı bulan önemli bir hesaplamalı modeldir. Bu model, AP'lerin üretilmesinde derin bir yaklaşım sunmaktadır. HH modelinin analitik zenginliği, iyon kanallarının dinamiklerini ve zar potansiyeli değişikliklerini biyofiziksel olarak gerçekçi bir şekilde temsil etme yeteneğiyle öne çıkmaktadır. Özellikle, BSA'ların hesaplamalı temsillerinin geliştirilmesinde çok önemli bir rolü üstlenmektedir (Shirahata, 2016).

HH nöron modeli, Şekil 9'da gösterilen hücre zar modelinin geliştirilmiş bir versiyonuna dayanmaktadır. Bu modelde, Şekil 9'da sunulan pasif iyon kanalları Şekil 10'da gösterildiği gibi indirgenmiş, ek olarak aktif iyon kanalları da dahil edilmiştir. Bu sayede hücre zarının doğrusal olmayan özellikler kazanması sağlanmıştır. Bu genişletilmiş hücre zarı modeli Şekil 11'de detaylandırılmıştır. Matematiksel olarak, bu model Denklem (25)'te analiz edilmiştir.

$$I_{giriş} = I_C + I_{Na} + I_K + I_{sızıntı} \quad (25)$$

Bu denklemde; $I_{giriş}$, dışarıdan uygulanan toplam akımı; I_C , kapasitif akımı; I_{Na} , Sodyum kanal akımlarını; I_K , Potasyum kanal akımlarını; $I_{sızıntı}$ ise sızıntı (kaçak) akımlarını temsil etmektedir. Bu ifadenin daha detaylı versiyonu, devre analizi kullanılarak Denklem (26)'da verildiği gibi elde edilebilmektedir.

$$C_{zar} \frac{dV}{dt} = I_{giriş} - \bar{g}_{Na} m^3 h (V - E_{Na}) - \bar{g}_K n^4 (V - E_K) - \bar{g}_{sızıntı} (V - E_{sızıntı}) \quad (26)$$



Şekil 11. HH nöron modelinin temsili ve eşdeğer devresi

Burada; C_{zar} , hücre zarının sığasını; V , hücre zarı potansiyelini; $\bar{g}_{Na}$, Sodyum iyon iletkenliğini; $\bar{g}_K$, Potasyum iyon iletkenliğini; $\bar{g}_{sızıntı}$, sızıntı akımları iletkenliğini; E_{Na} , Sodyum iyon potansiyelini; E_K , Potasyum iyon potansiyelini; $E_{sızıntı}$ ise kaçak potansiyelini ifade etmektedir.

Sodyum ve Potasyum kanallarının iletkenlik ifadeleri; m , n , ve h olarak adlandırılan değişkenlerle tanımlanan doğrusal olmayan dinamiklere sahiptir. Kapı değişkenleri olarak bilinen bu değerler Denklem (21) kullanılarak anlık olarak hesaplanmaktadır. Denklem (21)'de verilen kapı değişkenlerine ait α ve β ifadeleri aşağıda verilen (27), (28), (29), (30), (31) ve (32) numaralı denklemler kullanılarak anlık olarak hesaplanmaktadır (Ermentrout & Terman, 2010).

$$\alpha_m(V) = 0.1 \frac{V + 40}{\left(1 - e^{\left(-\frac{V+10}{10}\right)}\right)} \quad (27)$$

$$\beta_m(V) = 4e^{\left(-\frac{V+65}{18}\right)} \quad (28)$$

$$\alpha_n(V) = 0.01 \frac{V + 55}{\left(1 - e^{\left(-\frac{V+55}{10}\right)}\right)} \quad (29)$$

$$\beta_n(V) = 0.125e^{\left(-\frac{V+65}{80}\right)} \quad (30)$$

$$\alpha_h(V) = 0.07e^{\left(-\frac{V+65}{20}\right)} \quad (31)$$

$$\beta_h(V) = \frac{1}{\left(1 - e^{\left(-\frac{V+35}{10}\right)}\right)} \quad (32)$$

HH nöron modeli, voltaja bağlı iletkenlik özelliklerini temel alarak Sodyum ve Potasyum gibi spesifik iyonların hücre zarı üzerindeki hareketini tanımlamıştır. Bu iyon kanallarının dinamiklerinin doğru bir şekilde temsil edilmesi, BN'lerdeki AP'nin başlangıcı ve yayılması konusundaki anlayışı derinleştirmiştir. Ayrıca bu model, sonraki modellerinin geliştirilmesi için bir temel oluşturmaktadır.

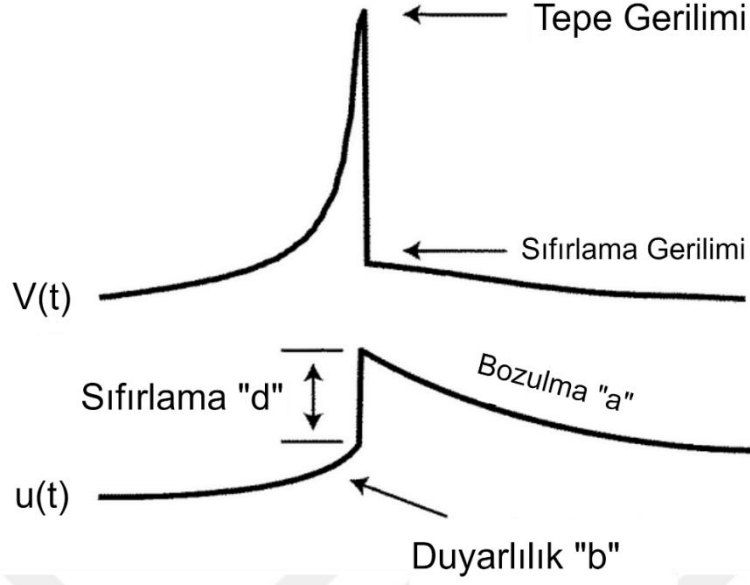
1.3.5. Izhikevich Nöron Modeli

IZH nöron modeli (Izhikevich, 2006), sinirbilim alanında önemli bir model olarak kabul görmektedir. Modelin en belirgin avantajlarından biri, minimum hesaplama karmaşıklığıyla geniş bir nöronal davranış spektrumunu temsil edebilme kapasitesidir (Qu & Wang, 2017). Sadece iki diferansiyel denkleme ve birkaç parametre dizisine sahip olan IZH modeli, BN'lerin karmaşık ateşleme dinamiklerini matematiksel olarak modelleyebilmektedir. Ek olarak, IZH modelinin nöromorfik hesaplama ve hesaplamalı sinirbilim gibi multidisipliner alanlarda geniş bir uygulama potansiyeli vardır. Bu, beyin esintili bilgi işlem sistemlerinin gelişimine katkıda bulunduğunu göstermektedir (Osborn vd., 2018). IZH nöron modelinin matematiksel formu, Denklem (33)'te verilmiştir.

$$C \frac{dV}{dt} = I_{giriş} + k(V - V_{sıfırlama})(V - V_{eşik}) - u \quad (33)$$

Burada; V , zar potansiyelini; $I_{giriş}$, dış kaynaklı akımı; $V_{sıfırlama}$, sıfırlama gerilimini; $V_{eşik}$, eşik gerilimini; u ise geri besleme akımını ifade etmektedir. Denklem (33) içerisindeki u terimi, zamana bağlı bir diferansiyel denklem ile ifade edilmekte ve Denklem (34) yardımıyla hesaplanmaktadır.

$$\frac{du}{dt} = a(b(V - V_{sıfırlama}) - u) \quad (34)$$



Şekil 12. IZH nöron modelinde parametreler ve dinamikleri (Izhikevich, 2006)

Denklemler (33) ve (34) zar potansiyelinin hesaplamasında kullanılırken, modelin BN'nin ateşleme yeteneğine sahip olabilmesi için Denklem (35)'te tanımlanan şartların sağlanması gerekmektedir:

$$\text{Eğer } (V > V_{Tepe}) \text{ ise } (V \leftarrow c) \text{ ve } (u \leftarrow u + d) \quad (35)$$

IZH modeli, sadece iki diferansiyel denklem vasıtasıyla BN'nin ateşleme dinamiklerini etkili bir şekilde temsil edebilme kapasitesine sahiptir. Bu temsil, sadece birkaç parametre dizisinin ayarlanmasıyla gerçekleştirilmektedir. Bu belirlenmiş parametreler sayesinde, model BSA'larda sıkça gözlemlenen ateşleme ve patlama davranışlarını başarılı bir biçimde yansıtabilmektedir. IZH nöron modelinde parametrelerin AP üretimindeki etkileri Şekil 12'de açıkça gösterilmektedir.

1.3.6. Gerçekçi Nöronlarda Sinaptik Mekanizma ve Plastisite

BN'ler arası iletişimde sinapsların merkezi bir rol oynaması, bu mekanizmanın matematiksel modellenmesinin önemini vurgulamaktadır. Matematiksel bir yaklaşımla, ardışık AP (spike) üreten bir presinaptik BN'nin spike dizisi, Denklem (36)'da tanımlanmıştır.

$$f(t) = \sum_i \delta(t - t_i) \quad (36)$$

Bu bağlamda; $f(t)$ ifadesi, spike dizisini; $\delta(t)$, birim basamak fonksiyonunu; t_i ise i 'nci AP'nin oluştuğu anı temsil etmektedir. Bu matematiksel ifade, sinaptik iletim sürecinde, postsinaptik GN'nin sinaptik iletkenliğini belirli bir ağırlık faktörü altında düzenlemektedir. Eğer postsinaptik GN herhangi bir sinaptik girdi almazsa, sinaptik iletkenlik sıfırdır. Ancak, sinaptik bir uyarı söz konusu olduğunda, bu uyarının etkisi sinaptik ağırlıkla doğru orantılı olarak artmakta ve bu artış, Denklem (37)'de tanımlanan değişim formülüyle hesaplanmaktadır.

$$\tau_{sinaps} \frac{dg_{sinaps}(t)}{dt} = - \left(g_{sinaps}(t) + \omega_{sinaps} f(t) \right) \quad (37)$$

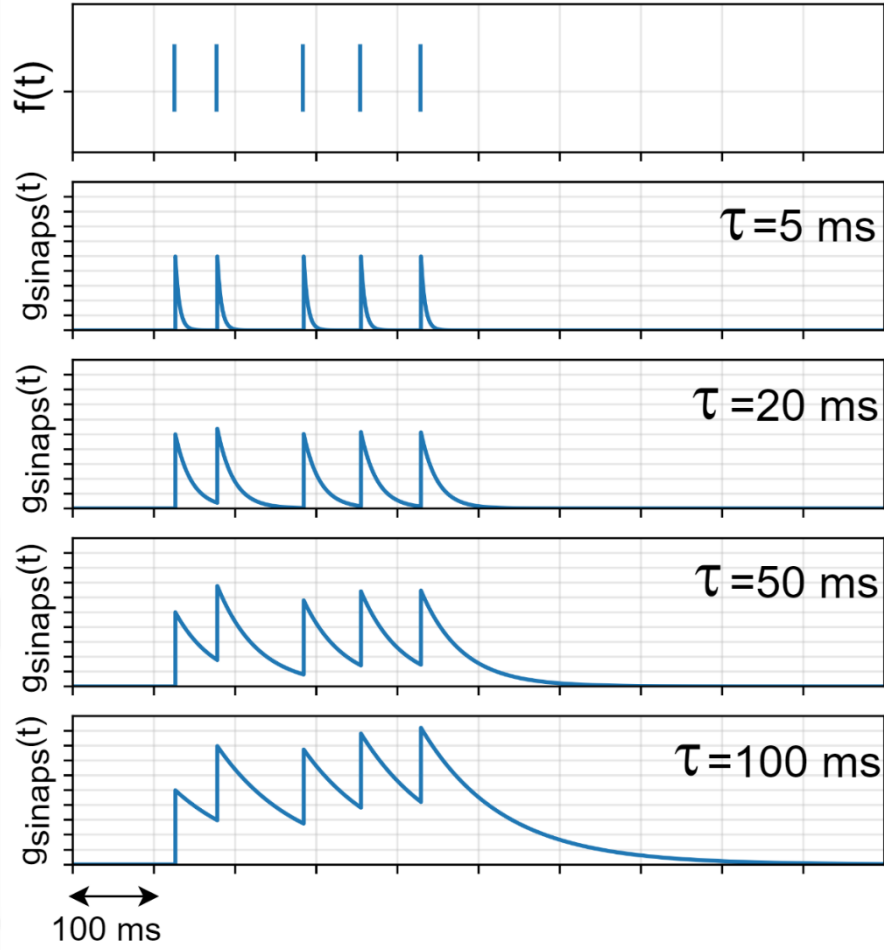
Bu bağlamda; τ_{sinaps} ifadesi, sinaptik iletkenliğin zaman sabitini; $g_{sinaps}(t)$, sinaptik iletkenliği; ω_{sinaps} ise sinaptik ağırlığı temsil etmektedir. Sinaptik ağırlık, iki GN arasındaki SB'nin kuvvetini belirtirken, τ_{sinaps} ise sinaptik etkinin ne kadar süreyle devam edeceğini gösteren bir ölçüttür. Şekil 13, τ_{sinaps} değişkeninin etkisini detaylıca göstermektedir.

Öte yandan, Denklem (37), Exc ya da Inh sinapslar hakkında bilgi sunmamaktadır. Bu tür bir bilgiyi elde etmek için, SB'nin zar potansiyeli üzerindeki etkisi göz önünde bulundurulmalıdır. Zar potansiyeli dikkate alındığında, SB'den kaynaklanan toplam sinaptik akım, Denklem (38)'de tanımlanan formülle hesaplanmaktadır.

$$I_{sinaps}(t) = \sum_i g_{sinaps_i}(t) \times (V_{post} - E_{sinaps_i}) \quad (38)$$

Burada; $I_{sinaps}(t)$, toplam sinaptik akımı; $g_{sinaps_i}(t)$, i 'inci sinapsın iletkenlik değerini; E_{sinaps_i} ise i 'inci sinapsın zar potansiyelini sembolize etmektedir.

Sinaptik iletimde, bir sinapsın Exc veya Inh işlevini belirlemek için Denklem (38) kriter olarak kabul edilmektedir. Eğer sinaptik potansiyelin değeri, postsinaptik GN'nin dinlenme potansiyelinden yüksekse, söz konusu sinapsın Exc nitelikte olduğu söylenmektedir. Diğer taraftan, E_{sinaps} değeri dinlenme potansiyelinden daha düşükse, bu

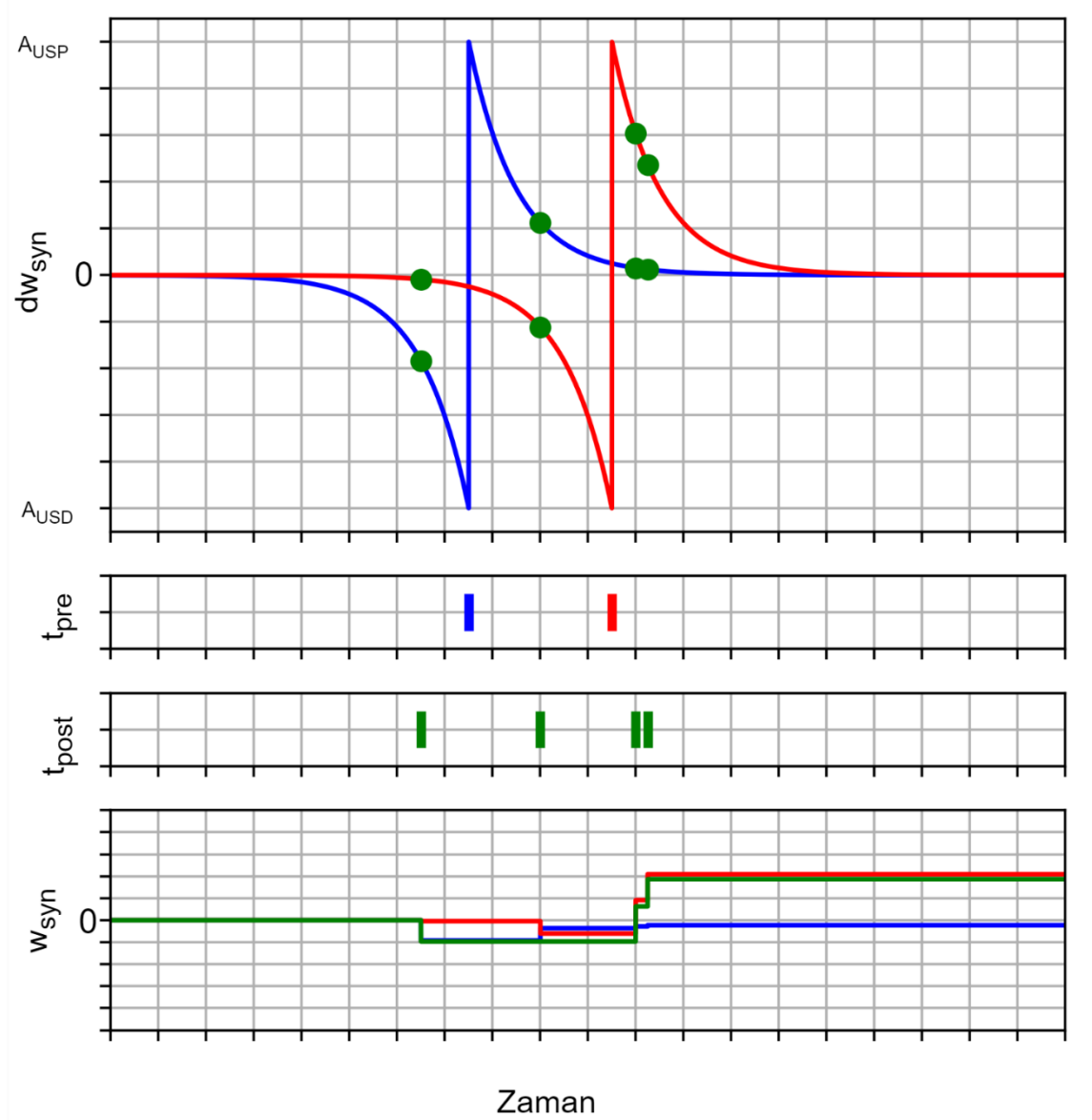


Şekil 13. Farklı zaman sabitlerine göre sinaptik iletkenlik değişimi

sinapsın Inh özellikte olduğu kabul edilmektedir. SB'lerin çeşitliliği ise τ_{sinaps} değişkeni üzerinden belirlenmektedir.

Spike zamanına bağlı plastisite (SZBP), sinir bilimi alanında beyindeki öğrenme ve hafıza süreçlerinin temelini oluşturan mekanizmadır. Presinaptik BN'de oluşan bir AP ile postsinaptik BN'de oluşan AP arasındaki zaman farkına dayanan uzun vadeli bir sinaptik plastisite biçimidir (Sgritta vd., 2017). SZBP, sinaptik plastisite mekanizmaları arasında en önemli öğrenme kurallarından biri olarak tanımlanmıştır. Bu plastisite biçimi kısa sayılabilecek bir zaman ölçeğinde çalışmakta ve yapay sinapslarda taklit edilebilmektedir (Du vd., 2015). SZBP, sinir sisteminde yaygın bir plastisite mekanizması olarak kabul edilmiş olup, sinir devrelerinin dinamiklerini anlamadaki önemi vurgulanmaktadır.

SZBP öğrenme kuralı, birlikte ateşlenen GN'lerin birbirleriyle nedensel bir ilişki içinde olduğunu ve bu nedenle de aralarındaki SB'nin güçlenmesi gerektiği varsayımına



Şekil 14. SZBP ile öğrenmede sinaptik ağırlığın zamana göre değişimi

dayanmaktadır. Eğer GN'lerin ateşlenme zamanları arasında nedensel bir ilişki söz konusu değilse de SB'nin zayıflaması gerekmektedir. Bu nedenle bu öğrenme kuralı presinaptik GN ile postsinaptik GN'de oluşan AP'ler arasındaki zaman farkını gözetmektedir.

Şekil 14'te SZBP öğrenme kuralının bir simülasyonu gösterilmektedir. En üstte presinaptik GN'de oluşan iki AP için zamana göre SB'nin ağırlık değişimi verilmiştir. Postsinaptik GN'de oluşan AP de grafik üzerinde işaretlenmiştir. En altta ise iki GN arasındaki SB ağırlığının zamana göre değişimi verilmiştir. Mavi eğri, presinaptik GN'deki mavi ile gösterilen AP'ye göre SB ağırlığının zamanla değişimini, kırmızı eğri ise kırmızı

AP'ye göre SB ağırlığının zamanla değişimini göstermektedir. Yeşil eğri ise toplamsal olarak SB ağırlığının zamana göre değişimini vermektedir.

Şekil 14'te görüleceği üzere, eğer presinaptik GN'de oluşan bir AP'nin hemen ardından postsinaptik GN'de bir AP oluşuyorsa bu GN'ler arasında nedensel bir ilişki vardır demektir. Bu nedenle aradaki zaman farkına bağlı olarak SB ağırlığı artırılmaktadır. Aradaki zaman farkı arttıkça SB ağırlık değişimi de azalmaktadır. Eğer presinaptik GN'de bir AP oluşmadan postsinaptik GN'de AP oluşuyorsa bu iki GN arasında nedensel bir ilişki yoktur demektir. Bu nedenle presinaptik GN'de oluşan ilk AP'ye göre SB ağırlığı azaltılmaktadır. Aradaki zaman farkı arttıkça da SB ağırlık değişimi de azalmaktadır. Bu süreç Denklem (39)'da matematiksel olarak ifade edilmektedir.

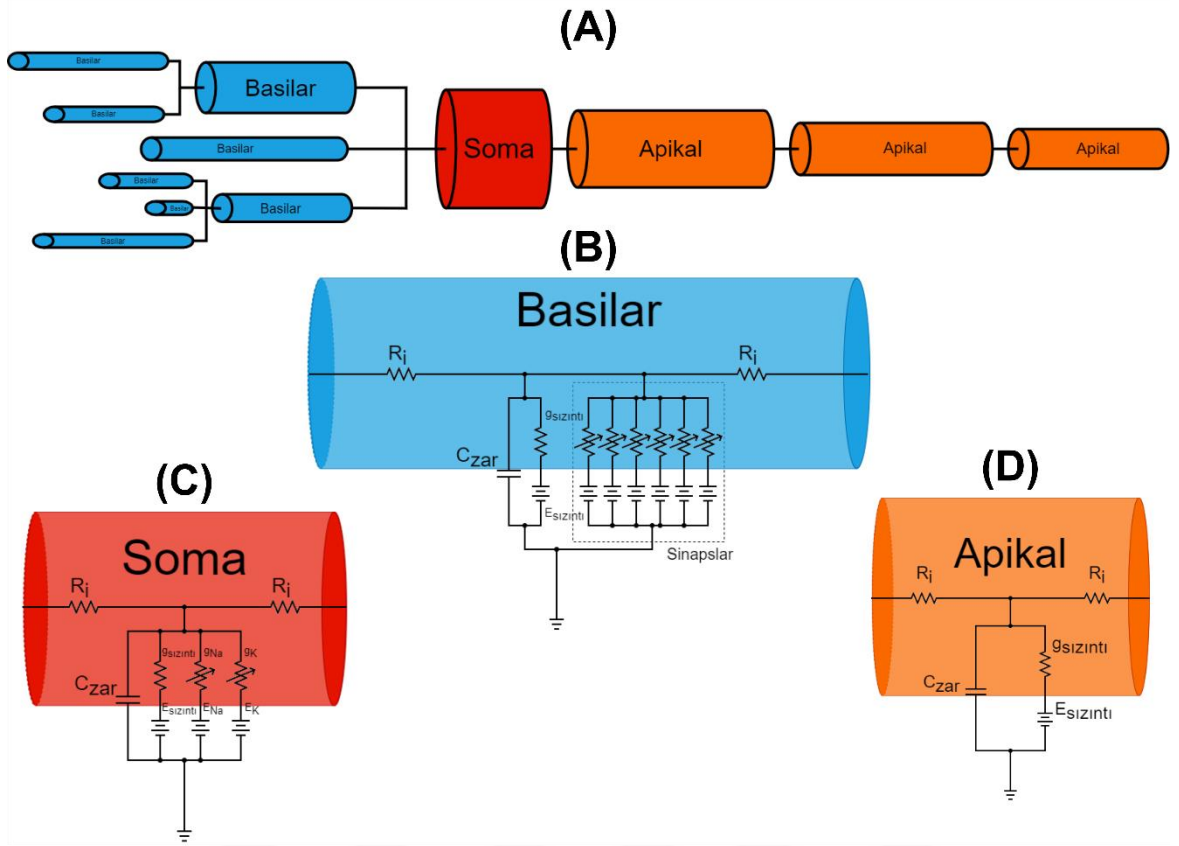
$$\Delta\omega = \begin{cases} A_{USP} \times e^{-\frac{t_{post}-t_{pre}}{\tau_{USP}}} & t_{post} - t_{pre} \geq 0 \\ A_{USD} \times e^{-\frac{t_{post}-t_{pre}}{\tau_{USD}}} & t_{post} - t_{pre} < 0 \end{cases} \quad (39)$$

Burada; $\Delta\omega$, SB ağırlık değişimini; A_{USP} , USP oluşması gerektiğinde SB ağırlık değişiminin pozitif yönde maksimum değerini; A_{USD} , USD oluşması gerektiğinde SB ağırlık değişiminin negatif yönde maksimum değerini; τ_{USP} ve τ_{USD} , sırasıyla USP ve USD zaman sabitlerini; t_{pre} , presinaptik AP'nin oluşma anını; t_{post} , postsinaptik AP'nin oluşma anını sembolize etmektedir.

1.3.7. Nöronların Modellenmesinde Bölmeli Yaklaşım

STA, HH ve IZH gibi tek bölmeli GN modelleri, basitlikleri ve hesaplamalı verimlilikleri sayesinde sinirbilim literatüründe yaygın bir kullanıma sahiptir. Bununla birlikte, bu modellerin birtakım kısıtlılıkları ve eksiklikleri mevcuttur. Bu modeller, BN'lerin karmaşık dendritik ağaç yapılarını ve çok bölmeli morfolojilerini tam olarak yansıtmada yetersiz kalmaktadırlar. Bu yetersizlik, GN'lerin sinaptik aktivite ve dendritik spike oluşumunu gibi süreçleri simüle etme kapasitelerini düşürmektedir (Beaulieu-Laroche vd., 2018). Ek olarak, tek bölmeli GN modelleri, iyon kanalı ve reseptörlerin uzamsal dağılımlarını aşırı basit bir şekilde ele alarak, BN'ler için önemli biyofiziksel detayların göz ardı edilmesine yol açabilmektedir (Eyal vd., 2016).

Literatürde tanımlanan bilgiler ışığında, BN'lerin daha kapsamlı bir şekilde modellenmesi için çok bölmeli GN modellerinin gerekliliği vurgulanmaktadır (Sterratt vd.,

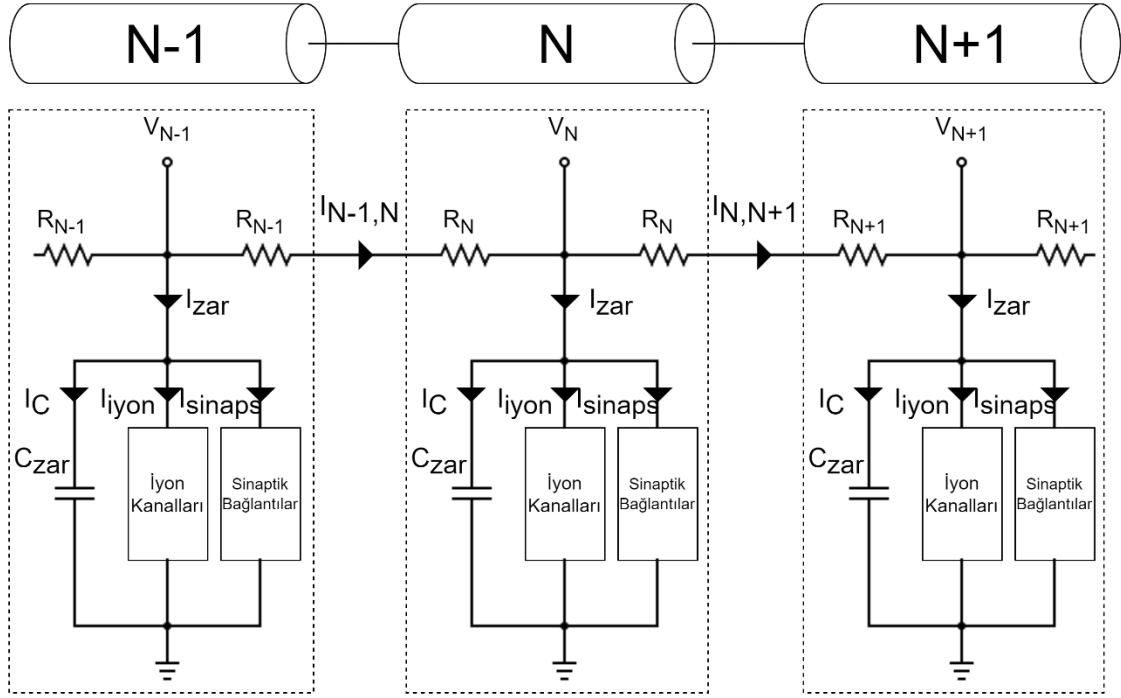


Şekil 15. Bölmeli nöron modeli ve bölmelere ait örnek eşdeğer devreler

2023). Bu modellemeler, kompleks bir BN'nin çeşitli silindirik bölmelere ayrıldığını ve bölmelerin ayrıca modellenip birleştirildiğini öngörmektedir. Bu yaklaşım, BN'lerin daha detaylı ve biyolojik gerçekliğe daha uygun bir şekilde temsil edilmesini amaçlamaktadır.

Şekil 15(A)'da gösterilen GN modeli, merkezinde bir soma olacak şekilde tasarlanmıştır. İçerisinde soma, basilar ve apikal dendritler bulunmaktadır. Bu bölmeler arasında, GN'nin uzamsal düzlemde farklı iyon kanallarına sahip olabileceği varsayılarak biyofiziksel özelliklerinin çeşitlilik gösterdiği dikkate alınmıştır. Şekil 15(B), GN'nin basilar dendrit kısmına ait elektriksel eşdeğer devre modelini göstermektedir.

Şekil 15(C)'de, GN'nin soma bölgesine ait elektriksel eşdeğer devre gösterilmektedir. Bu bölme, basilar dendritlerden gelen uyarıları, sahip olduğu aktif iyon kanalları aracılığıyla işleyerek apikal dendritlere iletmekle görevlidir. Ayrıca, Şekil 15(D), apikal dendrit bölmelerinin eşdeğer devresini tanımlamaktadır. Apikal dendritler, somadan aldığı bilgileri iletmek için tasarlanmıştır. Bu tür bir detaylı GN modellemesi ve simülasyonları, Özcan ve arkadaşları (2021) tarafından geliştirilen eğitim yazılımı



Şekil 16. Bölmeli bir nöronun elektriksel eşdeğeri

platformunda başarıyla gerçekleştirilebilmektedir. İlgili eğitim yazılımı sayesinde Şekil 15(A)'da verilen GN'nin hem morfolojik hem de biyofiziksel özelliklerinin etkisi incelenebilmektedir.

GN'lerin detaylı bir şekilde bölmelere ayrılması, biyolojik gerçekçiliği artırmakla birlikte, işlemsel karmaşıklığı da beraberinde getirmektedir. Nöronun yapısındaki karmaşıklık arttıkça, bu modellerin işlem yükü de artmaktadır. Bu tür hesaplamaların gerçekleştirilmesi için, Şekil 16'da verilen eşdeğer devre gibi modellerin kapsamlı bir analizi gerekmektedir.

Üç bölmeli bir GN'nin elektriksel eşdeğer modeli için, hesaplamaların ilk adımı, bir bölmeye giren toplam akımın belirlenmesidir. Bu hesaplama, Denklem (40)'da ifade edilmektedir.

$$I_{zar,N} = I_{N-1,N} - I_{N,N+1} \quad (40)$$

Burada; $I_{zar,N}$, bir bölmeye giren toplam akımı; $I_{N-1,N}$ ve $I_{N,N+1}$ ise sırasıyla komşu bölmelerden gelen akımları sembolize etmektedir. Ayrıca, bu toplam akım, HH nöron

modelinden aşına olunduğu üzere Denklem (41)'de gösterildiği gibi de ifade edilebilmektedir.

$$I_{zar,N} = I_{C,N} + I_{iyon,N} + I_{sinaps,N} \quad (41)$$

Bu ifadeler birleştirildiğinde ve komşu bölmelerden gelen akımlar dikkate alındığında, Denklem (42) elde edilmektedir.

$$C_{zar,N} \frac{dV_N}{dt} + I_{iyon,N} + I_{sinaps,N} = \frac{V_{N-1} - V_N}{R_{N-1} + R_N} - \frac{V_N - V_{N+1}}{R_N + R_{N+1}} \quad (42)$$

Bu denklemde özellikle vurgulanması gereken husus, hesaplamalar sırasında $I_{iyon,N}$ ve $I_{sinaps,N}$ akımlarının dinamik ve değişken olabilmesidir. Sinaptik uyarıyı temsil eden $I_{sinaps,N}$ yalnızca bir sinaptik girdinin varlığında dikkate alınırken, $I_{iyon,N}$ ise bölmenin biyofiziksel özelliklerine bağlı olarak farklılık gösterebilmektedir. Şekil 15(A)'da verilen eşdeğer devre modellerinde, her bir bölmeye özgü $I_{iyon,N}$ ifadeleri dikkate alınarak hesaplamalar gerçekleştirilmektedir.

1.3.8. Gerçekçi Nöronlarda Nöral Kodlama

GSA'ların geliştirilmesi sürecinde, BN'lerin işlevselliğini taklit etme amacıyla, giriş olarak alınan analog verilerin spike dizisi veya örüntüsü formatına dönüştürülmesi gerekmektedir. Spike dönüşüm süreci, GSA'nın BSA aktivitesini en yakın şekilde simüle edebilmesi için önemli bir aşamayı temsil etmektedir. Spike üretme yöntemleri, genel olarak üç ana kategoride sınıflandırılabilir: Orana dayalı yöntemler, spike zamanına dayalı yöntemler ve popülasyon örüntüsüne dayalı yöntemler (Auge vd., 2021). Orana dayalı ve spike zamanına dayalı yöntemler analog girişlerin zamana yayılmış spike dizileri olarak kodlanmasına karşılık gelirken, popülasyon örüntüsüne dair yöntemler analog verinin zamansal ve/veya mekânsal olarak yayılmış spike örüntüleri olarak kodlanmasına karşılık gelmektedir.

Orana dayalı spike üretme yöntemleri, dış ortamdan gelen analog girdilerin değerinin artmasıyla buna karşılık gelen spike dizisinin belirli bir aralıktaki spike sayısının artacağı prensibine dayanmaktadır. Bu prensip ilk olarak Adrian (1926) tarafından ortaya atılmıştır. Adrian'ın (1926) çalışmasına göre bir kasa uygulanan kuvvet arttıkça o kasta bulunan

BN'nin ateşleme frekansı artmaktadır. İlerleyen dönemlerde, Gawne ve arkadaşları (1996) da benzer sonuçlar elde etmiştir. Çalışmanın sonuçlarına göre bir BN'nin dış ortamdan aldığı uyarımın gücü arttıkça, BN daha yüksek frekansta spike üretmektedir. Bu bilgiden yola çıkılarak, girdilerin genliğinin yüksek frekans olarak ifade edildiği orana dayalı spike üretim yöntemleri ortaya atılmıştır. Bu yaklaşım ayrıca frekans kodlama olarak da bilinmektedir. Zamana göre değişen bir dış ortam sinyali spike dizisine dönüştürülürken bir ifade üzerinden anlık frekans ifadesine dönüştürülmektedir ve bu ifade Denklem (43)'te gösterildiği gibi modellenebilmektedir.

$$f_{anlık}(t) = \alpha \times x(t) + \beta \quad (43)$$

Burada; $x(t)$, dış ortam sinyali; $f_{anlık}(t)$, anlık spike frekansına; α ve β değerleri ise doğrusal dönüşümde kullanılacak olan keyfi parametrelere karşılık gelmektedir. Bu aşamadan itibaren spike dizileri $f_{anlık}(t)$ özelliklerini gösterecek şekilde üretilmelidir.

Orana dayalı spike üretim yöntemleri istatistiksel ve sinyal işleme tabanlı olmak üzere iki ana grupta incelenebilmektedir. İlk grup olan istatistiksel yöntemler genelde spike üretim mekanizmasının bir Poisson süreci olduğu varsayımına dayanmaktadır. Bu varsayıma göre spike dizileri Denklem (44)'te verilen yaklaşımla üretilmektedir.

$$S[n] = \begin{cases} 1, & c_{random} < \frac{f_{anlık}[n]}{f_s} \\ 0, & c_{random} > \frac{f_{anlık}[n]}{f_s} \end{cases} \quad (44)$$

Burada; $S[n]$, spike dizisini; $f_{anlık}[n]$, anlık frekans değerlerini; f_s , örnekleme frekansını; c_{random} ise tekdüze dağılıma sahip 0 ile 1 arası değerler alabilen rastgele üretilmiş bir sayıya karşılık gelmektedir. Bu yaklaşım, Poisson sürecine uygun spike dizileri üretilmesinde kullanılmaktadır.

Bir diğer orana bağlı spike üretme yöntemi ters katlama (deconvolution) temeline dayanmaktadır. Bu yaklaşıma göre bir spike dizisinin anlık frekansı bir pencere fonksiyonu ile katlanarak (convolution) bulunabilmektedir. Bir spike dizisinin anlık frekansı Denklem (45)'te verilmektedir.

$$f_{anlık}(t) = S(t) * W(t) = \int_{-\infty}^{\infty} S(\tau)W(t - \tau)d\tau \quad (45)$$

Burada; $S(t)$, spike dizisini; $W(t)$, pencere fonksiyonunu; $f_{anlık}(t)$ ise anlık frekansı ifade etmektedir. Bu ifadenin tersi alındığında, yani ters katlama yapıldığında, spike dizilerini elde etmek mümkündür. Bu yaklaşım, ilk olarak Hough ve arkadaşları (1999) tarafından önerilmiş, daha sonrasında Schrauwen ve van Campenhout (2003) tarafından da geliştirilmiştir.

Thorpe ve arkadaşları (1996) insan görme sistemini incelediklerinde, BN'lerin dış ortam uyaranlarını sadece oran tabanlı olarak değil aynı zamanda spike dizilerinin kesin zamanlamasında kodladıklarını göstermişlerdir. Bulgularına göre, dış ortamdan gelen ışığın şiddeti arttıkça BN daha erken zamanda spike üretmektedir. Bu bulgu, spike dizisi üretme yöntemlerine ilham olmuştur. Temporal kodlama olarak da bilinen bu yaklaşıma göre, dış ortam uyaranının genliğine göre spike zamanları belirlenmektedir (Auge vd., 2021).

Temporal kodlama yöntemleri tek spike yöntemleri ve çoklu spike yöntemleri olarak çeşitlilik göstermektedir. Tek spike yöntemlerinde spike dizisinden ziyade tek bir spike içeren dönüşümler yapılmaktadır. Bu yöntemlerde dış ortam uyaranı ile spike üretilme anı arasında ters ilişkili olacak şekilde bir ilişki kurulmakta ve genlik arttıkça daha erken zamanda spike üreten kodlara dönüştürülmektedir. Bu yöntemlere örnek olarak Sıralama kodlama (Thorpe & Gautrais, 1998) ve İlk Spike Zamanı kodlama (Gollisch & Meister, 2008) yöntemleri gösterilebilmektedir.

Çoklu spike yöntemlerinde ise global bir referans üzerinden spike dizileri üretilmektedir. Faz kodlama olarak da adlandırılan bu yaklaşım zamana göre değişen dış ortam uyaranına göre spike dizisinin fazının değiştirilmesi prensibine dayanmaktadır. Bu yöntemlere örnek olarak Faz kodlama (Montemurro vd., 2008) ve Gecikme kodlama (Storchi vd., 2012) gösterilebilmektedir.

BN'ler beyinde çeşitli katmanlardan aldıkları bilgileri işlemekten sorumludurlar. Özellikle, PN'ler gibi birden fazla katmandan girdi alan BN'ler, her katman için ayrı bir görevden sorumlu olabilmektedirler. Bu nedenle, bir BN, aldığı bilginin içeriğinin yanında hangi katman geldiğine göre de işlem yapmak zorundadır. Bu da girdilerin konumlarının da önemli olduğunu göstermektedir. Bu nedenle dış ortam uyaranlarının spike dilinde ifade edilebilmesi için girdi sahibi BN'lerin birbirlerine göre ilişkileri, konumları ve önemleri

etkilidir. Bu olgular, beyinde bilginin spike dizilerinin yanında popülasyonların birlikte hareket etmesiyle de taşınabileceğinin delillerini sunmaktadır.

Popülasyon kodlama bilginin bir grup BN tarafından ifade edildiği bir kodlama yaklaşımıdır. Özellikle PN gibi karmaşık yapılara sahip, girdi konumunun BN çıktısında etkili olduğu morfolojik modellerde popülasyon kodlama önemli bir yaklaşım olarak görülmektedir. Bu yaklaşım zaman-mekân tabanlı ve sadece mekân tabanlı yöntemler olarak iki ana grupta incelenebilmektedir.

Zaman-mekân tabanlı popülasyon kodlamada, bilgi bir grup GN tarafından kodlanmaktadır. Bu kodlama yönteminde her bir girdi GN'si, dış uyarıyı kendi iç dinamiklerine göre spike dizilerine dönüştürerek hedef GN'ye iletmektedir. Girdi GN'lerinin farklı noktalara farklı türde girdi uygulamaları sonucu, bilgi daha geniş bir perspektifte ifade edilmekte ve hedef GN'nin işlemine sunulmaktadır (Eliasmith & Anderson, 2003). Burada grup GN'lerinin SB konumları ayrı bir kodlama yöntemi, her bir GN'nin spike dizileri ise ayrı bir kodlama yöntemi sayılacağından bu yöntemler zaman-mekân tabanlı popülasyon kodlama olarak isimlendirilmektedir. Sadece mekân tabanlı kodlama yöntemlerinde ise önemli olan sadece GN'lerin konumlarıdır. Bu kodlama yönteminde, popülasyon bilgileri sadece ikili kod olarak ifade edilmektedir. Bu kodlamaya örnek olarak Moldwin ve Segev'in (2020) çalışmalarında kullandığı yaklaşım gösterilebilmektedir.

1.4. Otomatik Konuşmacı Tanıma

OKT, bireyin ses özelliklerine dayalı olarak konuşmacıyı otomatik olarak doğrulama ve/veya kimliklendirme amacı güden bir biyometrik tanıma sistemidir. Bu sistem, konuşmacıya özgü bilgileri konuşma sinyallerinde tarayarak bilinmeyen bir konuşmacıyı bir dizi referans konuşmacı arasından ayırt etme görevini üstlenmektedir (Campbell, 1997).

OKT; ticari, adli, biyometrik kimlik doğrulama ve teknolojik ilerlemeler dahil olmak üzere çeşitli alanlarda kendine yer bulmaktadır. Ticari uygulamalarda; sesli mesaj sistemleri, telefon bankacılığı ve biyometrik kimlik doğrulama gibi alanlarda kullanılmaktadır (Algabri vd., 2017). Ayrıca; adli uygulamalarda, OKT, suç soruşturmaları sırasında telefon dinlemeleri veya ambiyans kayıtlarından elde edilen konuşma verilerini analiz etmek için kullanılmaktadır (Beritelli & Spadaccini, 2012). Teknoloji, aynı zamanda, sesin konuşmacı doğrulama için bir özellik olarak kullanıldığı biyometrik kimlik

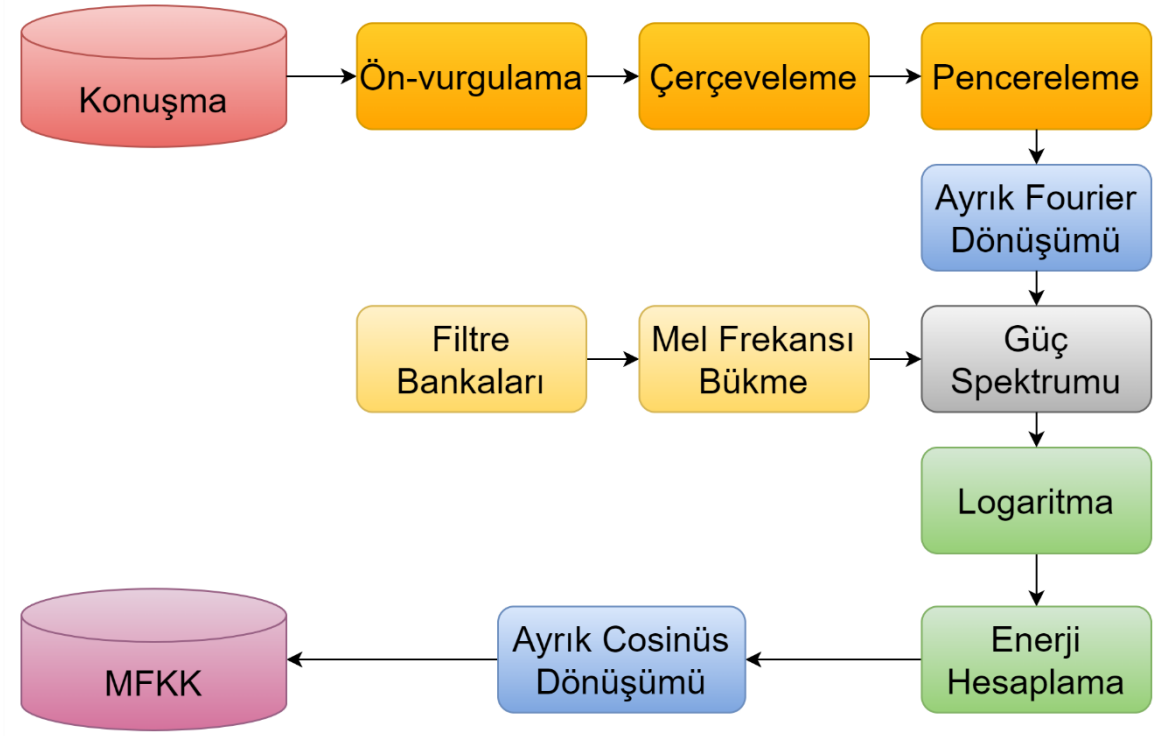
doğrulama sistemlerinde de yer almaktadır (Poddar vd., 2017). Son olarak, OKT, farklı alanlara adaptasyon yeteneğini göstererek mobil adli uygulamalara entegre edilmiştir. Hassas işlemlerde sahtekarlığı engellemek için otomatik ses tabanlı tanıma sistemleri gibi teknolojik ilerlemelerde de kullanılmıştır (Hamza vd., 2020).

Bu bölümde, OKT uygulamalarında kullanılan yöntemlerden bahsedilmiştir. Bölüm üç alt başlıktan meydana gelmektedir. İlk alt başlıkta, konuşma sinyallerinin analizinde kullanılan yöntemlerden bahsedilmiş ve bu yöntemler arasında en yaygın kullanılan öznitelik çıkartma yöntemi olan Mel Frekanslı Kepstrum Katsayıları (MFKK) tanıtılmıştır. İkinci alt başlıkta, konuşmacıların sınıflandırılmasında kullanılan yöntemlerin tarihi sürecinden bahsedilmiştir. Son alt başlıkta ise OKT uygulamalarında geliştirilen sistemlerin performanslarının ölçümünde kullanılan ölçütler detaylandırılmıştır.

1.4.1. Mel Frekanslı Kepstrum Katsayıları

OKT, ses sinyallerinden çıkarılan çeşitli öznitelikler aracılığıyla bireyleri özgün ses karakteristiklerinin belirlenmesi ve saptanması amacını taşımaktadır. Bu alanda, çeşitli öznitelik çıkarma yöntemleri öne çıkmaktadır. MFKK, konuşmanın zaman-frekans dinamiklerini hassas bir şekilde analiz ederken (Zheng vd., 2001); doğrusal öngörülü kodlama, sesin spektral detaylarını derinlemesine incelemektedir (Caucheteux vd., 2023). Benzer şekilde, algısal doğrusal tahmin, insan algısında öne çıkan ses niteliklerine odaklanmaktadır (Hermansky, 1990). Bu öznitelikler, OKT sistemlerinin konuşmacı kimliklendirme ve doğrulama probleminde üst düzey doğruluk elde etmesini sağlamaktadır (Beigi, 2011).

Ses, iletildiği arayüzde basınç değişimleri olarak kulağa ulaşmaktadır. Dış kulak bu basınç değişimlerini yakalamakta ve kulak zarına ileterek zamana göre değişen basınç değişimlerini mekanik harekete dönüştürülmesini sağlamaktadır. Kulak kemikleri, kulak zarındaki bu mekanik hareketleri algılayarak yükseltmekte ve koklea (salyangoz) olarak isimlendiren organa aktarmaktadır. Salıngoz bu mekanik hareketleri, içinde bulunan sıvı sayesinde yine basınç değişimlerine dönüştürmekte ve içinde farklı frekansları tespit etmekle sorumlu olan kıl hücreleri yardımıyla analiz etmektedir. Analiz sonucunda her bir kıl hücresi sorumlu olduğu frekans bileşenini algılamakta ve gerekli nöral kodlamanın üretilmesini sağlamak amacıyla işitsel sinirlere iletmektedir (Litovsky, 2015).



Şekil 17. MFKK hesaplama akış diyagramı

Kulaktaki işitme sistemi incelendiğinde, bu sürecin bir zaman-frekans analizi olduğu savunulmaktadır. Bu özellik, konuşma analizi ve özniteliklere de ilham kaynağı olmuştur. MFKK öznitelikleri konuşma tanıma ve OKT uygulamalarında yaygın bir yöntemdir. Literatürdeki çalışmalar incelendiğinde, OKT uygulamalarında MFKK özniteliklerinin kullanımının diğer yöntemlere nazaran ağır bastığı görülmüştür (Mohd Hanifa vd., 2021; Beigi, 2011). Bunun nedeni; MFKK özniteliklerinin kulak ve işitme sisteminin çalışma prensibinden esinlenerek, konuşma sinyallerinin zaman-frekans dinamiklerini başarılı bir şekilde temsil etmesidir. MFKK özniteliklerinin hesaplanma süreci ardışık bazı adımların takip edilmesidir ve süreç Şekil 17’de bir akış halinde gösterilmektedir.

Gerçek dünyada kaydedilmiş bir konuşma sinyali, $x(t)$, örnekleme frekansı f_s olacak şekilde örneklendiğinde $x[n]$ sinyali elde edilmektedir. Bu sinyal için MFKK hesaplama aşamasında yüksek frekans bileşenlerinin etkisinin artırılması amacıyla bir ön-vurgulama işlemi gerçekleştirilmektedir. Bu işlem için Denklem (46) kullanıldığında ön-vurgulanmış konuşma sinyali edilmektedir.

$$x_{vur}[n] = x[n] - \alpha \times x[n - 1] \quad (46)$$

Burada; $x[n]$, örneklenmiş konuşma sinyalini; $x_{vur}[n]$, ön-vurgulanmış konuşma sinyalini; α ise ön-vurgulama katsayısını ifade etmektedir. MFKK hesaplanmasında ön-vurgulama işlemi bir zorunluluk olmamakla birlikte, literatürdeki çalışmalar ön-vurgulama işleminin performansı önemli ölçüde artırdığını göstermektedir (Kinnunen & Li, 2010). Ön-vurgulama işleminin ardından konuşma sinyali kısa zamanlı çerçevelere ayrılmaktadır. Bu çerçevelerin keskin kenarlarının etkisinin azaltılması amacıyla Denklem (47)'de ifade edildiği gibi bir pencereleme işlemi yapılmaktadır.

$$x_{pen}[n] = x_{cer}[n] \times h_{pen}[n] \quad (47)$$

Burada; $x_{cer}[n]$, konuşma sinyaline ait tek bir çerçeveyi; $h_{pen}[n]$, çerçevenin yumuşatılması için kullanılan pencere fonksiyonunu; $x_{pen}[n]$ ise pencere fonksiyonu uygulanan çerçeve parçasını ifade etmektedir. Pencereleme işlemi de ön-vurgulama gibi, bir zorunluluk olmamakla beraber, pencerelemenin ve kullanılan pencere fonksiyonunun performansı önemli ölçüde artırdığı literatürde bildirilmiştir (Kinnunen & Li, 2010). Bundan sonraki adımlar sadece bir çerçeve üzerinden açıklanmış olup, bu adımların konuşma sinyalinin her bir çerçevesine ayrı ayrı uygulandığı göz önünde bulundurulmalıdır.

Bu aşamada elde, bir konuşma sinyaline ait çerçeve parçaları bulunmaktadır. Sonraki adım bu çerçevelerin frekans bileşenlerinin tespit edilmesidir. Denklem (48)'de verilen Ayırık Fourier Dönüşümü (AFD) uygulanarak çerçevelerin frekans içerikleri bulunabilmektedir.

$$X[k] = \sum_{n=0}^{N-1} x_{pen}[n] e^{-j \frac{2\pi kn}{N}} \quad (48)$$

Burada; $x_{pen}[n]$, çerçeve parçasını ve $X[k]$ ise çerçevenin frekans genlik spektrumunu ifade etmektedir.

Kulak içinde bulunan salyangoz, frekans bileşenlerini doğrusal bir formda algılamamaktadır. Bu nedenle frekans uzayının salyangozun algılayabileceği ölçüde bükülmesi gerekmektedir. MFKK öznitelikleri de frekans spektrumunun Mel frekansına bükülmesi adımıyla sahiptir (Abdul & Al-Talabani, 2022). Denklem (49), Mel frekansının elde edilmesinde kullanılan ifadeyi göstermektedir.

$$Mel(f) = 2595 \times \log_{10} \left(1 + \frac{f}{700} \right) \quad (49)$$

Salyangoz içinde bulunan kıl hücreleri her bir frekans bileşenini ayrı ayrı analiz etmek yerine filtre bankaları gibi çalışmakta ve bu filtre bankalarının enerjilerini göz önünde bulundurmaktadır. Bu nedenle, Mel frekansı uzayında eşit aralıklarla dizilmiş filtre bankaları oluşturulmakta ve filtre bankalarının enerjileri hesaplanmaktadır. Filtre bankaları çeşitli formlara sahip olsa da literatürde en çok kullanılan filtre bankası üçgen filtre bankalarıdır (Beigi, 2011). Filtre bankalarının logaritmik tabanda enerjilerinin hesaplanmasında kullanılan ifade Denklem (50)'de, filtre bankası olarak kullanılan üçgen ifadesi ise Denklem (51)'de verilmiştir.

$$E[m] = \log \left(\sum_{k=0}^M H_m[k] \times |X[k]|^2 \right) \quad (50)$$

$$H_m[k] = 1 - \left| \frac{k - \frac{N}{2}}{\frac{N}{2}} \right|, \quad 0 \leq k \leq N \quad (51)$$

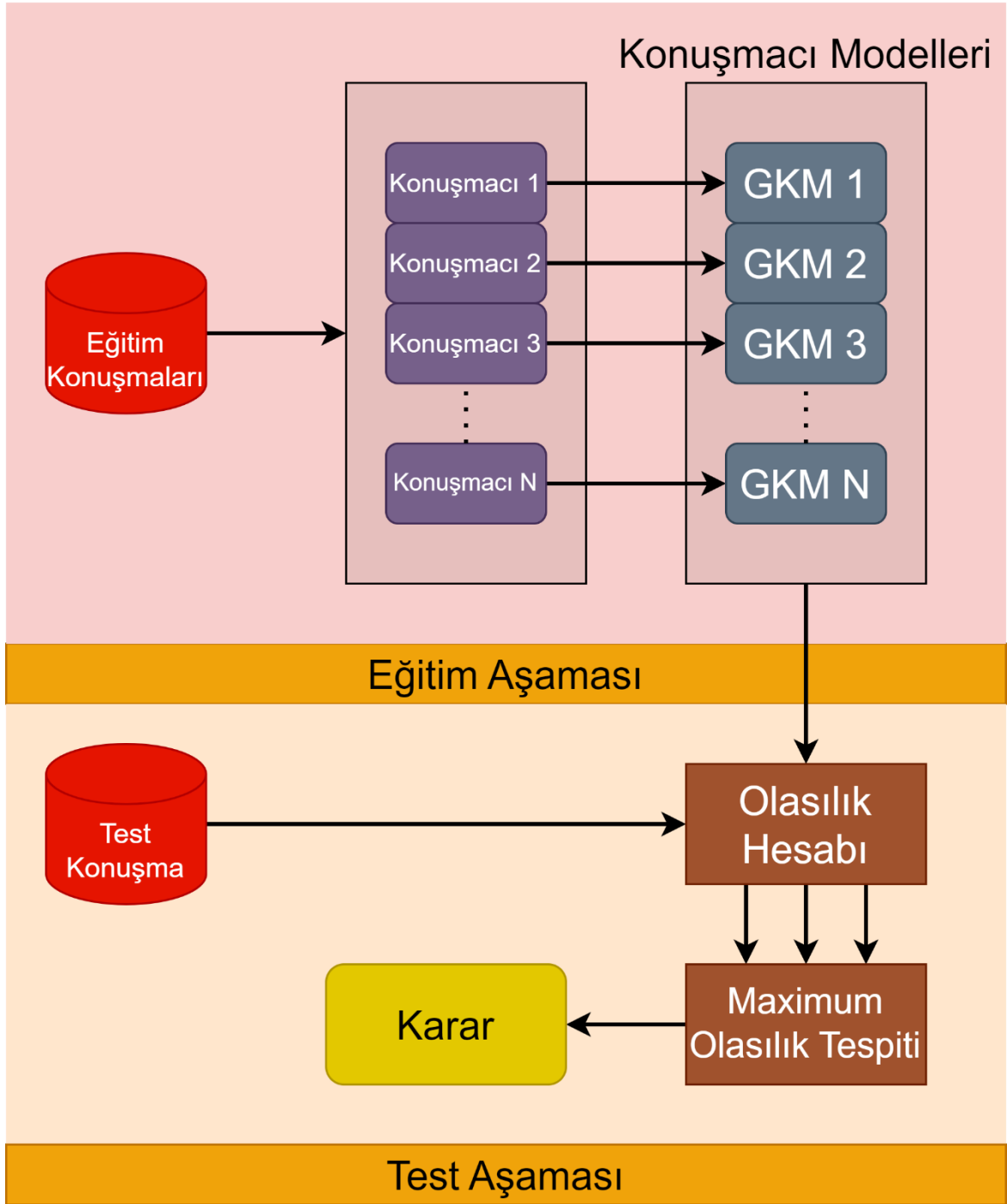
Burada; $E[m]$, m 'inci filtre bankasının logaritmik enerjisini; m , filtre bankası numarasını; $H_m[k]$ ise m 'inci filtre bankasının ifadesini göstermektedir. Filtre bankası enerjilerinin bulunmasının ardından, bu enerji dizisine Ayrık Kosinüs Dönüşümü (AKD) uygulanarak MFKK öznitelikleri hesaplanabilmektedir. AKD ifadesi Denklem (52)'de verilmiştir.

$$MFKK(i) = \sqrt{\frac{2}{N}} \sum_{m=1}^N E[m] \cos \left(\frac{\pi i}{N} \left(m - \frac{1}{2} \right) \right) \quad (52)$$

Burada; $E[m]$, filtre bankalarının logaritmik enerji değerlerini, $MFKK(i)$ ise i 'inci MFKK öznitelik değerini ifade etmektedir.

1.4.2. Konuşmacı Tanıma Yaklaşımları

OKT sistemlerinde konuşmacıların modellenmesi önemli bir adımdır. Kullanılan modellerin başarısı ve kapasitesi sistemin doğruluğunu, hızını ve donanımsal ihtiyaçlarını



Şekil 18. GKM yaklaşımında konuşmacı sınıflandırma akış diyagramı

doğrudan etkilemektedir (Özcan & Kayıkçıoğlu, 2021). Literatürde sunulan konuşmacı sınıflandırma yöntemleri istatistikî yöntemler ve DÖ tabanlı yöntemler olarak ikiye ayrılabilir. Bu bölümde konuşmacı modelleme yaklaşımlarında en gelişkin (state-of-the-art) yöntemlere değinilmiştir.

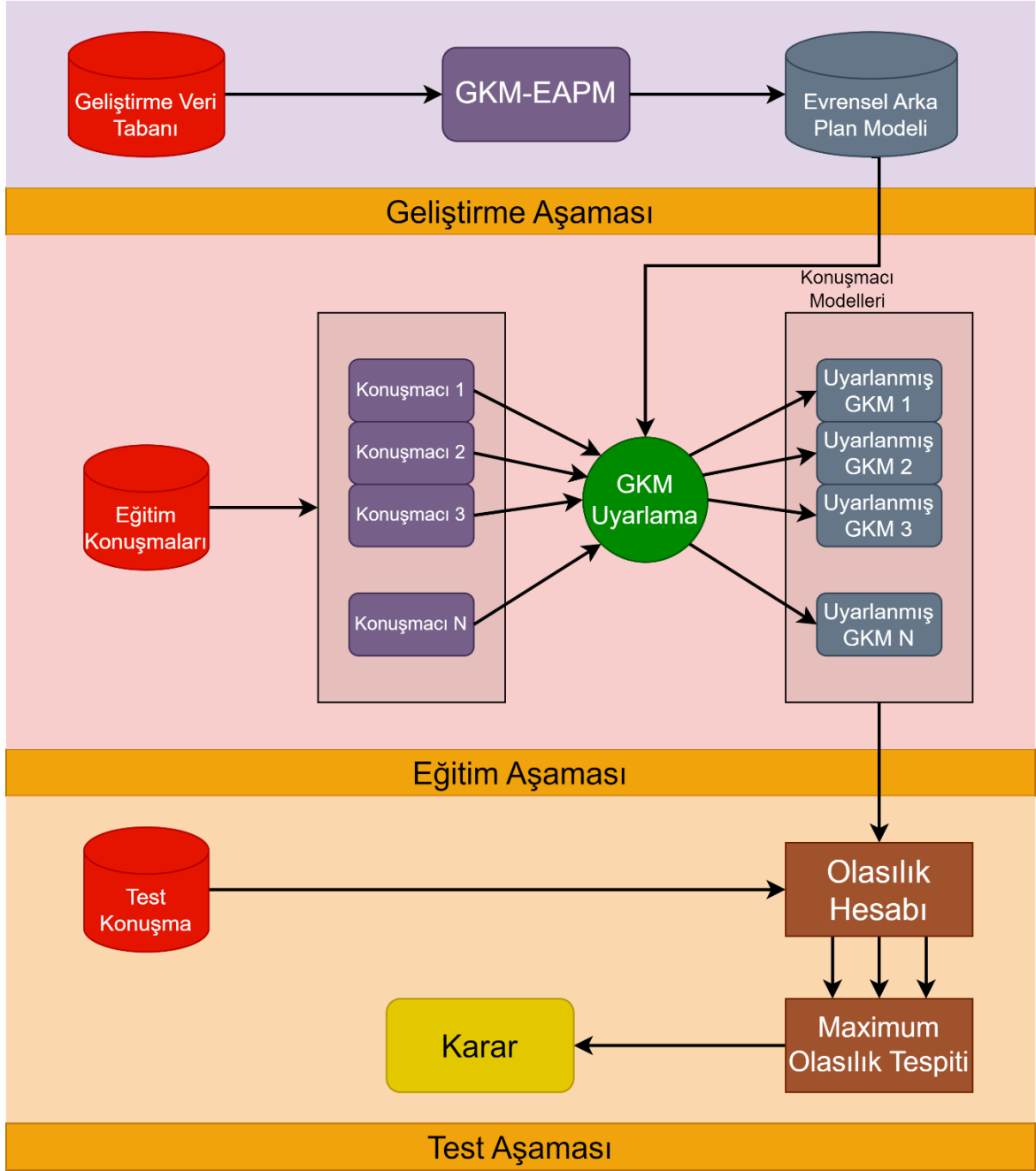
Literatürde ilk OKT modelleme yöntemi Reynolds (1995) tarafından önerilmiştir. Çalışmalarında Gauss karışım modellerini (GKM'leri) kullanmışlardır (Redner & Walker, 1984). Yöntemlerinin hem konuşmacı kimliklendirme hem de doğrulama konusunda gösterdiği başarı, GKM yaklaşımını literatürde önemli bir yerde tutmuştur. Çalışmalarında kullandıkları yaklaşım çeşitli varyasyonlara sahip olsa da basitçe Şekil 18'de gösterilmektedir.

GKM yaklaşımında problemin konuşmacı kimliklendirme veya doğrulama olmasına bağlı olarak farklı yaklaşımlar mevcuttur. Öncelikle konuşma kayıtlarından çerçeve tabanlı öznitelikler çıkartılmaktadır. Her bir çerçeve tek bir örnekmiş gibi etiketlenmektedir. Öznitelikler bireysel konuşmacı kimlikleri gözetilerek gruplanmaktadır. Bireysel konuşmacı öznitelikleri önceden karar verilmiş bir karışım miktarınca GKM kullanılarak modellenmektedir. Modelleme sonucunda GKM öbekleri ağırlık ve ortalama matrisleri halinde elde edilmektedir. Eğitim sonunda bireysel konuşmacı sayısı kadar GKM modeli elde edilmektedir.

Test aşamasında eğitim aşamasında olduğu gibi test konuşmasında çerçeve tabanlı öznitelikler çıkartılmaktadır. Bu öznitelikler eğitim aşamasında modellenen GKM'lere girdi olarak verilmekte ve her bir GKM modeli için ardıl olasılıklar hesaplanmaktadır. Buradan itibaren üç farklı yaklaşım uygulanabilir durumdadır. Eğer problem bir konuşmacı kimliklendirme işlemi ise her bir GKM modelinin ardıl olasılıklarının zamana göre ortalamaları alınmaktadır. Test konuşması için elde edilen olasılık değeri hangi konuşmacı GKM'sinde daha yüksek ise test konuşma kaydının o konuşmacıya ait olduğuna kanaat getirilmektedir.

Eğer problem bir konuşmacı doğrulama işlemi ise yine iki yaklaşım mevcuttur. İlk yaklaşım yine konuşmacı kimliklendirme problemindeki adımları takip edilmesiyle başlamaktadır. Eğer test konuşma kaydı iddia edilen konuşmacı GKM'sinde en yüksek olasılık değeri edilmiş ise konuşmacı doğrulanmaktadır, aksi durumda konuşmacının yanlışlanması gerekmektedir. Bir diğer konuşmacı doğrulama yaklaşımında ise test konuşma kaydı sadece iddia edilen konuşmacı GKM'sine girdi olarak verilmektedir. Ardıl olasılıkların ortalaması belirlenen bir eşikten büyük ise konuşmacı doğrulanmakta, aksi durumda yanlışlanmaktadır (Kinnunen & Li, 2010; Beigi, 2011).

Reynolds ve arkadaşları (2000) çalışmalarında, GKM yaklaşımını genişleterek daha başarılı sonuçlar elde etmişlerdir. Bu yaklaşıma göre, Şekil 18'de gösterilen akış Şekil 19'da gösterildiği gibi güncellenmiştir. Bu yaklaşımda eğitim ve test aşamalarına ek olarak



Şekil 19. GKM-EAPM yaklaşımında konuşmacı sınıflandırma akış diyagramı

geliştirme aşaması olarak bilinen bir aşama eklenmiştir. Öncelikle problemten bağımsız, içinde olabildiğince çok konuşma kaydına sahip, konuşmacı çeşitliliğinin yüksek olduğu bir VK kullanılmaktadır. Geliştirme veri tabanı olarak isimlendirilen bu VK genel bir konuşmacı modeli elde edilmesi amacıyla kullanılmaktadır.

Bu yaklaşıma göre geliştirme veri tabanından öznitelikler çıkartılmaktadır. Tüm öznitelikler konuşmacı kimliği göz ardı edilerek tek bir konuşmacıymış gibi GKM ile

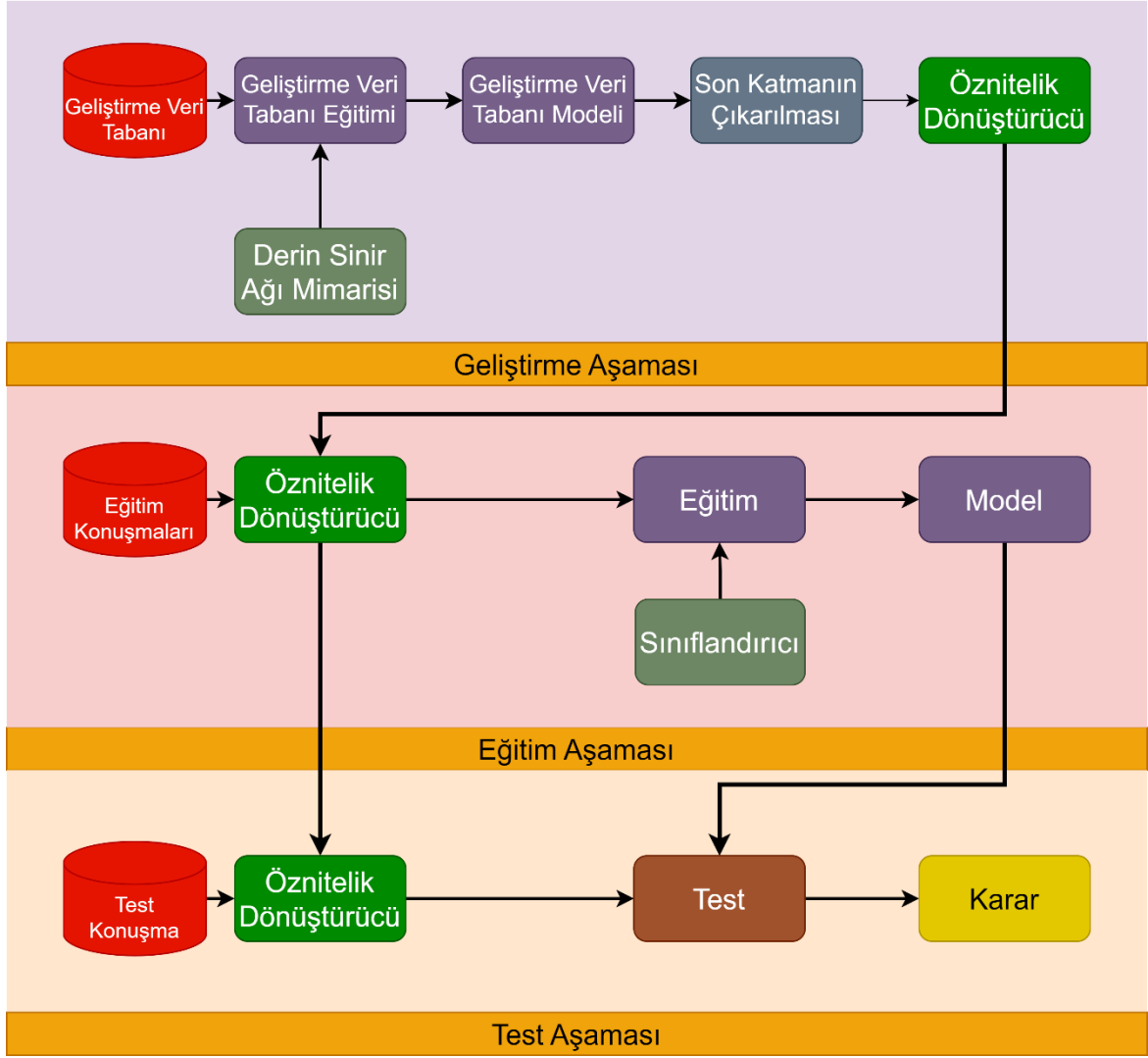
modellenmektedir. Oluşturulan bu model evrensel arka plan modeli (EAPM) olarak isimlendirilmektedir. Bu modelin genel bir konuşmacı modeli olduğu kabul edilmektedir.

Eğitim aşamasında, GKM modelleri bireysel konuşmacı modellerinin oluşturulması hedeflenmektedir. Fakat GKM-EAPM yaklaşımında konuşmacı modelleri eğitim kayıtlarından çıkartılmak yerine, EAPM'nin eğitim kayıtlarından yararlanılarak uyarlanması yoluyla elde edilmektedir. Uyarlama işlemi MAP uyarlaması ismi verilen bir süreçle yapılmakta ve elde edilen GKM modelleri uyarlanmış GKM olarak isimlendirilmektedir. Test aşaması GKM yaklaşımıyla aynı şekilde yapılmaktadır. Bu yaklaşım genel bir konuşmacı modelinin hesaba katılması sayesinde OKT performansını artırmıştır.

Campbell ve arkadaşları (2006) EAPM yaklaşımına yeni bir yorum getirmişlerdir. Bu yorum, uyarlanmış GKM modellerinin parametrelerinin öznitelik olarak kullanılmasına dayanmaktadır. Çalışmalarında, uyarlanmış GKM modellerinin ağırlık ve ortalama değerlerini öznitelik olarak kullanmış ve bu öznitelikleri süpervektörler olarak isimlendirmişlerdir. Daha sonra süpervektörleri sınıflandırmak için destek vektör makinelerini (DVM'yi) kullanmışlardır (Cortes & Vapnik, 1995). Bu nedenle, literatüre bu yaklaşım GKM-DVM yöntemi olarak anılmaktadır.

Dehak ve arkadaşları (2011) süpervektör fikrinden yola çıkarak kendi çalışmalarını önermişlerdir. Yaklaşımlarına göre eğitim kayıtlarının özniteliklerine faktör analizi uygulayarak kimlik vektörleri, i-vektörler, olarak isimlendirilen özniteliklerine ulaşmışlardır. Yaklaşımları faktör analizinin güçlü yanlarından faydalanmanın yanında, çok boyutlu süpervektörlerin de daha az boyutlarda ifade edilmesini sağlamaktadır. Böylece, OKT sistemlerinin başarımını artırmasının yanında işlemsel yükü hafif uygulamalara dönüşmesinin önü açılmıştır.

DÖ, klasik MÖ algoritmalarından daha iyi performans göstermeye başladığında, OKT alanında DÖ dönemi başlamıştır. DÖ yaklaşımlarının hemen hemen hepsinin amacı i-vektörlerin yerini alacak öznitelikleri hesaplamaktır. Variani ve arkadaşları (2014) maksimum çıktı sinir ağlarını (MÇSA'ları) kullanarak d-vektörler olarak isimlendirdikleri kendi özniteliklerini tanıtmışlardır. Chen ve arkadaşları (2015) bu yaklaşımı konuşulan içeriğin de etkisini ekleyerek genişletmiş ve j-vektörler olarak isimlendirilen kendi özniteliklerini tanıtmışlardır. Snyder ve arkadaşları (2018) çalışmalarında, zaman gecikmeli sinir ağlarını (ZGSA'ları) kullanarak, bugün x-vektörler olarak bilinen öznitelikleri önermişlerdir. Tüm bu yaklaşımlara ek olarak, uçtan uca sistemler (Heigold



Şekil 20. Öznitelik dönüştürücülerin eğitimi ve kullanımına dair akış diyagramı

vd., 2016; Rohdin vd., 2018), derin inanç ağları (Ali vd., 2016; Banerjee vd., 2018), derin düzeltici öğrenme ağları (Wen vd., 2018) ve konvolüsyonel sinir ağları (Ravanelli & Bengio, 2018; Hajavi & Etemad, 2019) gibi yaklaşımlar da OKT uygulamalarına kendi yaklaşımlarını getirmiştir. Yakın zamanlı çalışmalara göre DÖ yöntemleri OKT konusunda muazzam bir potansiyele sahiptir.

Literatürde önerilen DÖ tabanlı yöntemler genelde geliştirme aşamasında bir DSA yapısının eğitilmesine dayanmaktadır ve süreç Şekil 20’de gösterilmiştir. Bir DSA mimarisi seçilmektedir. Bir geliştirme veri tabanı kullanılarak bu DSA eğitilmektedir. Bu aşamadan itibaren elde olan DSA sadece geliştirme veri tabanını sınıflandırma kapasitesine sahiptir. Fakat DSA’nın son katmanı ortadan kaldırıldığında gizli katmanların başarılı bir öznitelik çıkarma işlemi yaptığı düşünülmektedir. Bu nedenle, son katmanın çıkarılmasında sonra artı kalan DSA öznitelik dönüştürücü olarak kullanılmaktadır.

Tablo 1. Karışıklık matrisi

Gerçek Tahmin \	Pozitif	Negatif
Pozitif	TP	FP
Negatif	FN	TN

Eğitim aşamasında konuşma kayıtlarından elde edilen öznitelikler öznitelik dönüştürücü DSA'ya girdi olarak verilmekte ve çıktıları dönüştürülmüş öznitelik olarak kabul edilmektedir. Bu öznitelikler bir sınıflandırıcı yardımıyla eğitilmekte ve eğitim modeli elde edilmektedir. Test aşamasında aynı adımlar test konuşma kaydı için de yapılmakta ve sınıflandırıcı modelinde test edilmektedir.

1.4.3. Konuşmacı Tanımada Performans Ölçütleri

OKT, esasen bir sınıflandırma problemidir ve her sınıflandırma problemi gibi performansının güçlü ölçütlerle ölçülmesi gerekmektedir. Literatürde birçok performans ölçütü bulunmaktadır. Bunlardan en yaygın olanları doğruluk (accuracy), kesinlik (precision), duyarlılık (recall) ve F1-Skordur. Bu ölçütlerin en önemli özelliği çoklu sınıf problemlerinde önemli bilgiler sağlamasıdır. Bu bölümde bu ölçütler tanıtılmıştır.

Öncelikle hedef bir sınıf belirlenmektedir. Bu sınıf pozitif olarak isimlendirilirken, geri kalan tüm sınıflar negatif olarak isimlendirilmektedir. Gerçek etiketler ve tahmin edilen etiketler göz önünde bulundurularak, iki sınıflı karışıklık (confusion) matrisi Tablo 1'de gösterildiği gibi elde edilmektedir. Tablo 1 için alanların belirlenmesinde aşağıdaki özellikler gözetilmektedir:

- TP: Gerçek etiketi pozitif olan ve tahmin sonucu pozitif olarak doğru sınıflandırılan toplam örnek sayısı

- FP: Gerçek etiketi negatif olan ama tahmin sonucu pozitif olarak yanlış sınıflandırılan toplam örnek sayısı
- FN: Gerçek etiketi pozitif olan ama tahmin sonucu negatif olarak yanlış sınıflandırılan toplam örnek sayısı
- TN: Gerçek etiketi negatif olan ve tahmin sonucu negatif olarak doğru sınıflandırılan toplam örnek sayısı

Tablo 1’de verilen karışıklık matrisi belirlendikten sonra birçok performans ölçütü bu matris üzerinden hesaplanabilmektedir. Aşağıda bu performans ölçütlerinin ifadeleri ve anlamları verilmiştir.

- Doğruluk: Toplam başarılı tahmin edilen örneklerinin oranını ifade etmektedir ve Denklem (53)’te gösterildiği gibi hesaplamaktadır.

$$\text{Doğruluk} = \frac{TP + TN}{TP + TN + FP + FN} \quad (53)$$

- Kesinlik: Bir sınıf için gerçek etiketlerin doğru tahmin edilme oranıdır ve Denklem (54)’te gösterildiği gibi hesaplamaktadır.

$$\text{Keskinlik} = \frac{TP}{TP + FP} \quad (54)$$

- Duyarlılık: Bir sınıf açısından incelendiğinde tahmin edilen etiketlerin gerçek etiketlerle uyumunun oranıdır ve Denklem (55)’te gösterildiği gibi hesaplamaktadır.

$$\text{Duyarlılık} = \frac{TP}{TP + FN} \quad (55)$$

- F1-Skor: Kesinlik ve duyarlılık ölçütlerinin harmonik ortalamasıdır ve Denklem (56)’da gösterildiği gibi hesaplamaktadır.

$$\text{F1 Skor} = \frac{2 \times \text{Kesinlik} \times \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}} = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (56)$$

1.5. Literatür Taraması

İşitsel sinyallerin biyo-esinli işlenmesi ve tanınması, hem canlıların bu süreçleri yüksek başarı ile gerçekleştirebilmesi hem de bu süreçte çok düşük enerji gereksinimine sahip olmaları nedeniyle araştırmacıların ilgisini çekmektedir. Bu nedenle, araştırmacılar, işitsel sistem ve beyin çalışma mekanizmalarını sinyal işleme ve MÖ uygulamalarına entegre etme çabası içindedirler.

Erken sayılabilecek bir dönemde, Näger ve arkadaşları (2002), Lyon (1982) tarafından önerilen salyangoz (koklea) modelinden elde ettikleri spike dizilerini işlemek üzere, STA nöronlarının kullanıldığı bir GSA tasarlamışlardır. GSA modelini konuşma sinyalleri kullanarak eğitmelerinin ardından, GSA'nın İngilizce "Yes" ve "No" sözcüklerini birbirinden ayırabilme kapasitesinin bulunduğunu göstermişlerdir.

Sonrasında, Gosztolya ve Tóth (2008) çalışmalarında STA nöronlarını, konuşma sinyallerinde hece sınırlarını belirlemek için kullanmışlardır. Çalışmalarına göre hedef, hece geçişlerinde, eğitilmiş bir STA nöronun spike üretmesidir. Çalışmaları STA nöronun bu görevi %96.42 doğruluk ile gerçekleştirdiğini göstermektedir. Ayrıca, STA nöronun mevcut yöntemlere göre %35.72 daha hızlı olduğunu göstermişlerdir.

DellaFerrera ve arkadaşları (2020) bir konuşma sinyalinde aktif ve basit bölgeleri tespit etmek üzere, STA nöronlarından kurulmuş bir GSA modeli önermişlerdir. Çalışmaları sonucunda; önerilen yöntemin mevcut yöntemlerle yarışabilecek performansının yanında, bu görevi gerçekleştirirken $3.8 \mu W$ gibi çok düşük enerji gereksinimine sahip olmasına dikkat çekmişlerdir. Wu ve arkadaşları da (2021) aynı problemi çözmek adına, akım tabanlı STA nöronlardan kurulu bir GSA modeli önermişlerdir. Çalışmalarına göre GSA modelleri sayesinde bir konuşma sinyalindeki aktif ve pasif bölgelerin tespiti, yarı toplam hata oranı %2.65 olmak üzere, mümkündür.

Bu bölümde, literatürde işitsel sinyallerin işlenmesi ve tanınması konusunda biyolojik gerçekçilikten faydalanan çalışmalardan bahsedilmiştir. Çalışmalar üç alt başlıkta altında incelenmiştir. İlk alt başlıkta, genel ses tanıma uygulamalarından; ikinci alt başlıkta, otomatik konuşma tanıma uygulamalarından ve son alt başlıkta ise otomatik konuşmacı tanıma uygulamalarından bahsedilmiştir.

1.5.1. Genel Ses Tanıma

Bu bölümde; otomatik konuşma veya konuşmacı tanıma uygulamaları haricinde kalan, işitsel sinyallerin GSA modelleri yardımıyla işlenmesi ve tanınması ile ilgili literatürde bulunan çalışmalardan bahsedilmektedir.

Coath ve arkadaşları (2005) önerdikleri bir GSA modelini eğitirken, Gammatone filtre bankalarının çıktılarını kullanmışlardır. Test sonuçlarına göre önerdikleri GSA modeli, yönetilen bir konuşma kaydında konuşulan cümlelerin bir soru cümlesi mi yoksa açıklama cümlesi mi olduğunu %88 doğrulukla tespit edebilmektedir.

Larson ve arkadaşları (2010) bir kuş türüne, aynı kuş türünün şarkılarını rastgele dinleterek, denek canlı bireyinin hücre kayıtlarını invazif yöntemlerle ölçmüşlerdir. Hücre kayıtlarından elde edilen spike dizilerini sınıflandırmak üzere STA nöronlardan kurulmuş bir GSA modeli kullanmışlardır. Çalışmalarına göre önerilen GSA modeli, kuş türünün şarkılarından elde edilen spike dizilerini %92.4 ile %98.3 arası doğruluklarla sınıflandırabilmektedir.

Newton ve Smith (2012) çalışmalarında müzikal çalgı seslerinin sınıflandırılması amacıyla üç farklı GSA modeli önermişlerdir. Çalışmalarında önerdikleri modellere göre filtre bankaları enerjileri, MFKK öznitelikleri ve salyangoz modeli olmak üzere üç farklı giriş kullanmışlardır. Sonuçlarına göre, GSA modelleri, 5 sınıfta gruplanan müzikal çalgıları, modellere göre farklılık göstermekle birlikte, %75.6, %76.3 ve %72.1 doğruluklarla tanıyabilmektedir.

Dennis ve arkadaşları (2015) toplam 50 sınıftan oluşan ses olayı sınıflandırma problemi için bir GSA modeli kullanmışlardır. Önerdikleri spike kodlama yöntemini test etmek için kullandıkları GSA modelinde STA nöronlarını tempotron öğrenme kuralı ile eğitmişlerdir. Sonuçlarına göre GSA modeli 50 sınıflı problemde ses olaylarını ortalama %94.1 doğruluk ile tanımıştır.

Wu ve arkadaşları (2018) geliştirdikleri GSA modelini test etmek adına çeşitli ses kayıtları kullanmışlardır. Öznitelik olarak MFKK kullanmış ve STA nöronlardan oluşturdukları GSA modelini eğitmek için tempotron öğrenme kuralını kullanmışlardır. Sonuçlarına göre önerdikleri GSA modeli, 10 sınıftan meydana gelen ve çevre seslerinin bulunduğu veri kümesini %99.6 doğrulukla sınıflandırmıştır. Aynı veri kümesini kullanan GSA çalışmalarında, Dennis ve arkadaşları (2013) kendi modellerinin %98.50 doğruluk performansına, Xiao ve arkadaşları (2017) ise kendi modellerinin %97.50 doğruluk

performansına sahip olduğunu bildirmişlerdir. Daha sonrasında, aynı veri kümesini kendi önerdikleri GSA modelinde kullanan Miao ve arkadaşlarının (2018) sonuçlarına göre önerilen GSA modeli 9 sınıflı problem için %99.9 doğruluğa sahiptir.

Mohanty ve arkadaşları (2020) kuş türlerinin çağrılarında faydalananarak, tür tespiti problemini ele almışlardır. Çalışmalarında STA nöronlarının kullanıldığı üç katmanlı bir GSA modelini eğitmek için MFKK özniteliklerini spike dizilerine dönüştürmüşlerdir. Toplam 14 türün sınıflandırılmasını içeren çalışmalarında, GSA modelinin kuş türlerini %90.1 doğrulukla tanıdığını bildirmişlerdir. Ayrıca çalışmalarının diğer çalışmalara göre daha hızlı çalıştığına dikkat çekmişlerdir.

1.5.2. Otomatik Konuşma Tanıma

Bu bölümde, GSA modellerinin bir konuşma sinyalinde aktarılacak istenen mesajın tespit edilmesinin amaçlandığı otomatik konuşma tanıma uygulamalarında kullanım alanlarından ve literatürdeki çalışmalardan bahsedilmiştir.

Coath ve arkadaşları (2005) önerdikleri bir GSA modelini eğitirken, Gammatone filtre bankalarının çıktıları kullanmışlardır. Deneylerinden birinde İngiliz alfabesinde bulunan harflere ait ses kayıtlarının sınıflandırılmasını amaçlamışlardır. Kullanılan GSA modeli 26 sınıflı bu problemde %80'den fazla doğruluk performansı göstermiştir.

Wade ve arkadaşları (2010) geliştirdikleri bir GSA modelinin performansını ölçmek adına 0 ila 9 rakamlarının ses kayıtlarının toplandığı TI46 veri kümesini kullanmışlardır. STA nöron modelleri kullanılan GSA modelleri 10 sınıflı problemde ortalama %95.25 doğruluk performansı göstermiştir. Aynı veri kümesinin kullanıldığı diğer GSA çalışmalarında, Verstraeten ve arkadaşları (2005) kendi GSA modellerinin %97.5, Zhang ve arkadaşları (2015) ise %92.3 doğruluk gösterdiğini bildirmişlerdir.

Wu ve arkadaşları (2018) geliştirdikleri GSA modelini test etmek adına 0 ila 9 rakamlarının ses kayıtları dahil olmak üzere konuşma kayıtlarının içerildiği TIDIGITS veri kümesini kullanmışlardır. Öznitelik olarak MFKK kullanmış ve STA nöronlardan oluşturdukları GSA modelini eğitmek için tempotron öğrenme kuralını kullanmışlardır. Sonuçlarına göre önerdikleri GSA modeli 11 sınıflı konuşma kayıtlarının bulunduğu veri kümesini %97.4 doğrulukla sınıflandırmıştır. Aynı veri kümesini kullanan diğer GSA çalışmaları için; Tavanaei ve Maida (2017) kendi modellerinin %96 doğruluğa, Zhang ve arkadaşları (2015) kendi modellerinin %92.3 doğruluğa, Anumula ve arkadaşları (2018)

kendi modellerinin %97.9 doğruluğa, Dibazar ve arkadaşları (2004) kendi modellerinin %85.1 doğruluğa, Dong ve arkadaşları (2018) ise kendi modellerinin %97.5 doğruluğa ulaştığını bildirmişlerdir.

Wu ve arkadaşları (2021) geliştirdikleri insan-robot işitsel arayüzünün verilen komutları tanıyabilmesi adına akım tabanlı STA nöronlardan kurulu bir GSA modeli önermişlerdir. Çalışmalarına göre GSA modelleri sayesinde konuşma kayıtlarında geçen “Yes”, “No”, “Up”, “Down”, “Left”, “Right”, “On”, “Off”, “Stop” ve “Go” gibi komutlar %92.55 doğrulukla tanınabilmektedir.

Guo ve arkadaşları (2023) geliştirdikleri GSA modelini test etmek için bir konuşma tanıma uygulaması yapmışlardır. IZH nöron modeli kullanarak kurdukları GSA modelini eğitmek için konuşma sinyalleri için Lyon (1982) tarafından önerilen salyangoz modelinin çıktılarını kullanmışlardır. Çalışmaları sonucunda GSA modeli konuşma tanıma görevini %96.73 doğrulukla yerine getirmiştir.

Sun ve arkadaşları (2023), DÖ ve GSA modellerini birlikte kullandıkları bir konuşma tanıma çalışmasında, NTIDIDIGITS ve SHD veri kümelerini test amacıyla kullanmışlardır. Spike tepki modeli olarak isimlendirilen nöron modellerinden kurulmuş hibrit ağ modelleri; NTIDIDIGITS veri kümesinde %95.22, SHD veri kümesinde ise %93.25 doğruluk performansı göstermiştir. Çalışmalarının en çarpıcı tarafı önerdikleri hibrit modelin mevcut yöntemlere göre daha az parametre sayısına sahip olmasına karşın, diğer yöntemlerle yarışabilecek performans göstermesidir.

1.5.3. Otomatik Konuşmacı Tanıma

Bu bölümde, literatürde otomatik konuşmacı tanıma alanında GSA modellerinin kullanıldığı çalışmalar göz önünde bulundurulmuştur.

Coath ve arkadaşları (2005) önerdikleri bir GSA modelini eğitirken, Gammatone filtre bankalarının çıktılarını kullanmışlardır. Deneylerinden birinde, 4 farklı konuşmacı kaydına sahip kayıtlardan konuşmacıların kimliklerini tespit etmeyi amaçlamışlardır. Kullanılan GSA modeli konuşmacı tanıma uygulamasında %89 doğruluk performansı göstermiştir. Ayrıca, bir başka deneylerinde, konuşma kayıtlarından konuşmacıların cinsiyetini tespit etmeyi amaçlamışlar ve GSA modelleri bu görevi %95 doğruluk ile gerçekleştirmiştir.

Wysoski ve arkadaşları (2007) çalışmalarında bir GSA modelinin konuşmacı doğrulama problemindeki performansını incelemişlerdir. Çalışmalarında 43 farklı konuşmacı kaydına sahip VidTimit veri kümesini kullanmışlar ve konuşma kayıtlarında öznitelik olarak MFKK kullanmışlardır. Çalışmaları sonucunda GSA modelinin vektör kuantalama yöntemine kıyasla başarılı bir performans olarak sayılabilecek %20 hata oranı yakalamışlardır.

Timoszczuk ve Cabral (2007) bir konuşmacı tanıma uygulamasında GSA modelini kullanmışlardır. Spike tepki model olarak bilinen bir nöron modelini GSA tasarlamak için kullanan yazarlar, öznitelik olarak Mel filtre bankalarının enerji değerlerini kullanmışlardır. Toplam 10 farklı konuşmacı kaydının kullanıldığı uygulamada GSA modeli %82 doğrulukla konuşmacı kimliklerini tanımıştır. Klasik çok katmanlı perceptron modelinin %96 doğrulukla aynı görevi yerine getirebilmesine karşın yazarlar, GSA modelinin düşük enerji gereksinimi ve hızına dikkat çekmişlerdir.

Ceolini ve arkadaşları (2019) ateşlemeli silikon koklea ile elde edilmiş spike dizilerini konuşmacı doğrulamada kullanmışlardır. Kullanılan sınıflandırma modelleri arasında herhangi bir GSA modeli olmasa da spike dizilerin konuşmacı doğrulama için kullanılması biyolojik olarak gerçekçi bir yaklaşım olarak düşünülmüştür. Çalışmaları sonucunda konuşmacı doğrulama görevinde ortalama %88 doğruluğa ulaşmışlardır.

1.6. Tezin Kapsamı ve Katkıları

Bu tez çalışmasında, otomatik konuşmacı tanıma uygulamaları için, konuşma sinyalinin işlenmesi sürecinin kulaktan beyne kadar olan aşamalarının biyo-esinli bir yaklaşımla modellendiği, beyinde gerçekleşen süreçlerin ise gerçekçi sinir ağı modelleriyle temsil edildiği bir yaklaşım önerilmiş ve performansı incelenmiştir.

Bu tez çalışmasının literatüre katkıları; önerilen özgün yaklaşımlar, ilk kez yapılan çalışmalar ve literatürde bulunan bilgileri derinleştiren katkılar olmak üzere üç grupta incelenebilmektedir.

Tez çalışmasında önerilen özgün yaklaşımlar ve detayları aşağıdaki gibi sıralanabilmektedir:

- Bu tez çalışmasında, analog değerlerin zaman-mekân bağımlı bir yaklaşımla spike örüntülerine dönüştürüldüğü bir öznitelik-spike dönüştürücü önerilmiştir. Bu

dönüştürücünün parametre ve seçenek olarak sunulan faktörlerinin sınıflandırma performansına etkisi incelenmiştir.

- Bu tez çalışmasında, derin sinir ağları ve gerçekçi sinir ağlarının birlikte kullanıldığı hibrit bir yaklaşım önerilmiştir.
- Bu tez çalışmasında, ikili formda ifade edilen spike örüntüleri için ağırlık değerlerinin pozitif olma kısıtının bulunduğu durumlarda, ağırlıkların bu kısıttan kurtarılması adına aynalama yaklaşımı önerilmiştir.

Literatürde daha önce değinilmemiş ve tez çalışması sayesinde ilk kez gerçekleştirilmiş katkılar aşağıdaki gibi sıralanabilmektedir:

- Otomatik konuşmacı tanıma uygulamalarında ilk kez piramidal nöronların performansı incelenmiştir.
- Piramidal nöronların sınıflandırma performansı ilk kez zaman-mekân bağımlı öznetelik vektörleri kullanılarak incelenmiştir.
- Piramidal nöronlarda sinaptik bağlantı sayısının sınıflandırma performansını nasıl etkilediği ilk kez incelenmiştir.
- Inh sinaptik bağlantıların gerçekçi sinir ağlarında sınıflandırma performansını nasıl etkilediği ilk kez incelenmiştir.
- Perceptron öğrenme kuralı da ilk kez Inh sinaptik bağlantıları da kapsayacak şekilde genişletilmiştir.

Tez çalışması sayesinde literatürde bulunan bilgilerin derinleştirildiği ve detaylandırıldığı katkılar aşağıdaki gibi sıralanabilmektedir:

- Otomatik konuşmacı tanıma uygulamalarında eğitim veri uzunluğu, test verisi uzunluğu ve MFKK sayısı gibi faktörlerin sınıflandırma performansına etkisi en gelişkin yöntemlerde incelenmiştir.
- Öznetelik-spike dönüşümü yaklaşımının sınıflandırma performansına etkisi klasik makine öğrenimi yöntemleriyle karşılaştırmalı olarak incelenmiştir.
- Önerilen gerçekçi sinir ağı modelinin sınıflandırma performansı çeşitli veri kümeleri ve klasik makine öğrenimi algoritmaları kullanılarak incelenmiştir.
- Aynalama yaklaşımının sınıflandırma performansına etkisi basit gerçekçi nöron modelleri ile karşılaştırmalı olarak incelenmiştir.
- Sinaptik bağlantı tiplerinin sınıflandırma performansına etkisi basit gerçekçi nöron modelleri ile karşılaştırmalı olarak incelenmiştir.

- Hibrit yaklaşımın sınıflandırma performansına etkisi basit gerçekçi nöron modelleri ile karşılaştırmalı olarak incelenmiştir.
- Gerçekçi sinir ağlarının veri miktarına göre eğitim performansının nasıl etkilendiği gösterilmiştir.
- Gerçekçi sinir ağlarının çoklu sınıf problemlerinde sınıflandırma performansının nasıl etkilendiği basit gerçekçi nöron modelleri ile karşılaştırmalı olarak incelenmiştir.



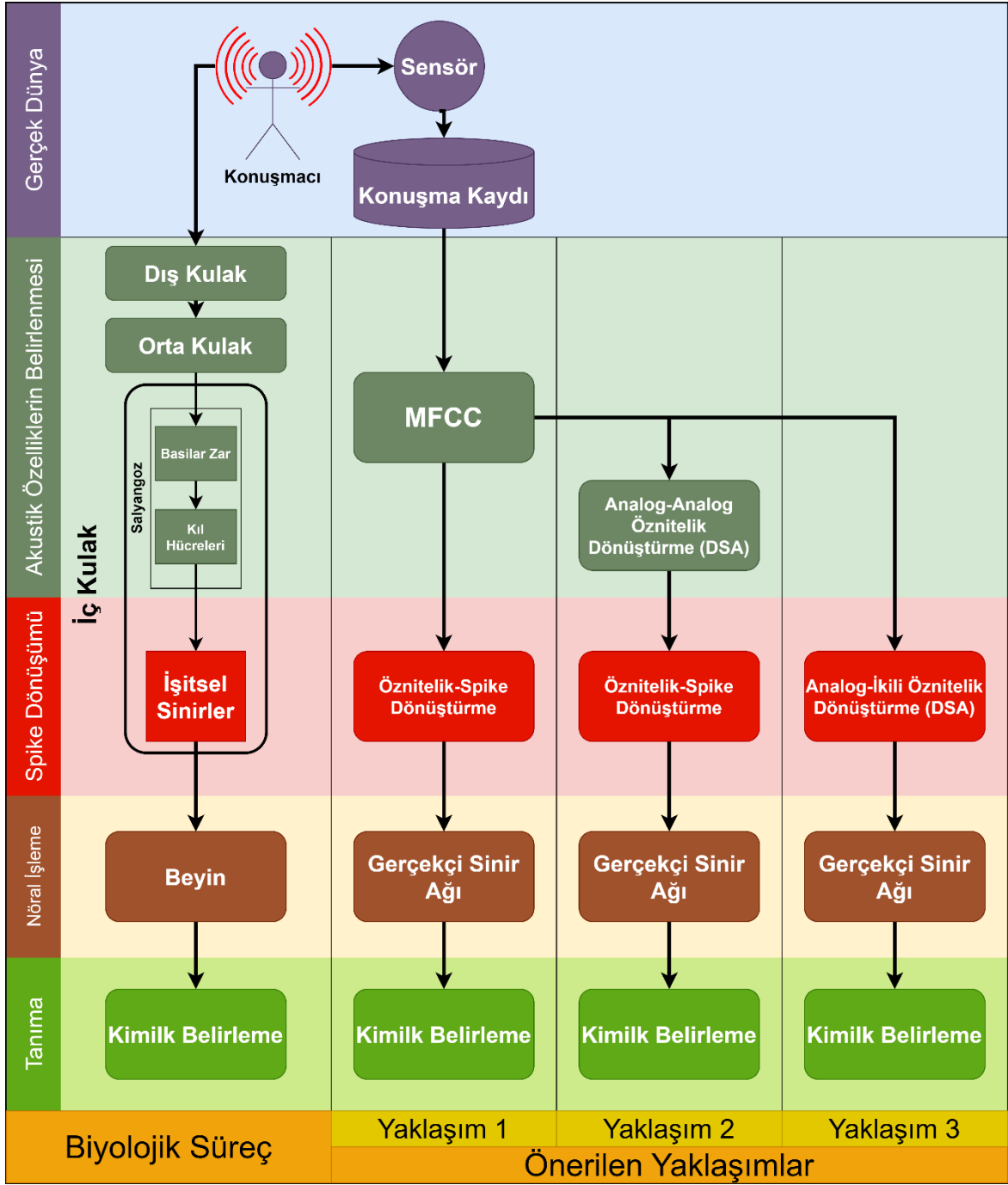
2. YAPILAN ÇALIŞMALAR

Konuşma, insanların birbirleriyle iletişimde birincil kaynak olarak kabul edilmektedir. Konuşma sinyalinin temel amacı içerdiği mesajı alıcıya ulaştırmaktır. Fakat konuşma içeriğinde iletilmek istenen mesajın yanında konuşmacıya dair çeşitli bilgiler de aktarılmaktadır (Beigi, 2011). Bunlar; konuşulan dil (Ambikairajah vd., 2011), bölgesel ağız (Zaidan & Callison-Burch, 2014) ve kelime dağarcığı (Fendji vd., 2022) olarak sıralanabilmektedir. Ayrıca konuşma sinyalleri; konuşmacının yaşı (Bahari vd., 2014), cinsiyeti (Wang vd., 2015), duygudurumu (Kim vd., 2009) ve kimliği (Kinnunen & Li, 2010) gibi konuşmacıya dair bilgileri de taşımaktadır. Konuşma sinyalleri için temel çalışmalar iletilen mesajın tespit edilmesine yönelik olsa da dil ve konuşmacıya özgü bilgilerin tanınması da ayrıca araştırma konularıdır.

Otomatik konuşmacı tanıma (OKT), bir kişinin kimliğinin konuşma sinyalinden faydalanılarak tespit edilmesidir. OKT uygulamaları konuşmacı kimliklendirme ve konuşmacı doğrulama olmak üzere iki temel yaklaşımla gerçekleştirilmektedir. Konuşmacı kimliklendirme uygulamaları bir konuşma sinyalinde “Konuşan kişi kimdir?” sorusuna cevap ararken, konuşmacı doğrulama uygulamaları konuşma sinyalinde “Konuşan kişi X midir?” sorusuna cevap aramaktadır. Bu nedenle, uygulamalar bu yaklaşımlara uygun olacak şekilde tasarlanmaktadır. OKT uygulamalarında başlıca zorluklar, konuşma kayıtlarının kalitesi ile ilgili kısıtlardır. Bunun yanında; konuşma sinyalindeki mesaj, dil ve konuşmacıya özgü diğer bilgiler; OKT uygulamalarında performansı önemli ölçüde etkileyen faktörler olarak bilinmektedir (Beigi, 2011; Kinnunen & Li, 2010).

Akustik konuşma sinyallerinin işlenmesi kulakta başlamaktadır. Konuşma sinyallerinin kulağa ulaşmasıyla, içerikte bulunan bilgilerin elde edilmesine yönelik bir dizi işlem sürdürülmektedir. OKT perspektifinden bakıldığında, bu işlemler sayesinde konuşma sinyali içeriğindeki kimlik bilgileri elde edilmektedir. Biyolojik konuşmacı tanıma süreci birçok noktada OKT uygulamalarının esin kaynağı olsa da mevcut yöntemler hala biyolojik gerçekçilik kaygısı gütmeyen yaklaşımlar içermektedir.

Şekil 21, biyolojik konuşmacı tanıma sürecinin basit bir akışını göstermektedir. Akustik konuşma sinyalleri ilk olarak dış kulağa ulaşmaktadır. Dış kulak, kulak kepçesi sayesinde konuşma sinyallerini toplamakta ve kulak kanalı vasıtasıyla kulak zarına ve dolayısıyla orta kulağa aktarmaktadır (Vollandri vd., 2015).



Şekil 21. Biyolojik konuşmacı tanıma süreci ve önerilen yaklaşımlar

Kulak zarı, aldığı akustik sinyale göre titreşerek, konuşma sinyalini mekanik titreşime dönüştürmekte ve yükseltmek üzere kulak kemiklerine aktarmaktadır. Kulak kemiklerinde, yükseltme ve bir tür ön-vurgulama işleminden geçtikten sonra; sinyal üzeni kemiğinin oval pencere aracılığıyla salyangozun içinde bulunan sıvıyı titreştirmesiyle iç kulağa aktarılmaktadır (Parveen vd., 2021).

Salyangozun içinde bulunan basilar zar, anatomisi gereği oval pencereden uzaklaştıkça daha düşük frekans bileşenleri tespit etmekten sorumlu kıl hücreleri ile donatılmıştır (Meaud & Grosh, 2010). Salyangoz içinde titreşen sıvı, salyangozun özgün anatomisi sayesinde mekâna göre bir filtreleme işleminden geçerek, frekans bileşenlerinin ilgili kıl hücrelerine ulaştırılmasında görev almaktadır. Kıl hücrelerinin özgün frekansları tespit etmekten sorumlu olması, bu sürecin konuşma sinyalleri için akustik özelliklerin frekansa bağlı analizi olduğunu göstermektedir (Wichmann & Moser, 2015).

Beynin nöral dili, spike dizileri veya örüntüleri; alfabesi ise aksiyon potansiyelleridir (AP). Bu bağlamda, kıl hücrelerinin yaptığı dönüşüm sonucu ortaya çıkan sinyaller beynin işleyebileceği forma uygun değildir. Bu nedenle, anatomik olarak kıl hücrelerinin hemen ardında bulunan işitsel sinirler, kıl hücrelerinden gelen sinyalleri uygun spike dizilerine dönüştürerek beyne aktarılmasını sağlamaktadır (Sendin vd., 2014). Bu aşamadan itibaren, gerekli işleme ve bilgilerin elde edilmesi beynin görevidir.

Yapılan çalışmalar, işitsel sinyallerin beyinde bulunan işitsel kortekste ilk olarak frekansa özgü bir yaklaşımla işlendiğini göstermektedir (Lee vd., 2020). Daha sonrasında, konuşma ile ilgili bölgelere aktarılmaktadır. Buna göre konuşma sinyallerinin işlenmesinden görevli ilk bölüm Wernicke bölgesidir (DeWitt & Rauschecker, 2013). Bu bölge, konuşma sinyallerinden gerekli mesajın elde edilmesi konusunda temel işlem merkezidir. Konuşmacının kimliğinin belirlenmesi ve konuşmacı sesinin algılanması ile ilgili çalışmalar ise beyinde konuşma sinyalleri ile ilgili daha çok bölgenin olduğunu göstermektedir. Buna göre, Latinus ve Belin (2011) konuşmacının sesinin algılanması ile ilgili başat bölümün Superior Temporal Sulcus (STS) bölgesi olduğunu göstermiştir.

Biyolojik konuşmacı tanıma süreci, OKT'nin geliştirilmesi ve performansının artırılması konusunda önemli bir gelecek vaat etmektedir. Bu tez çalışmasında, OKT performansı biyolojik konuşmacı tanıma sürecine oldukça benzer bir yaklaşımla incelenmiştir. Temelde üç farklı yaklaşımın önerildiği bu tez çalışmasında, önerilen yaklaşımlar Şekil 21'de verilen akışta gösterilmektedir. Buna göre, dış, orta ve iç kulakta gerçekleştirilen akustik özelliklerin belirlenebilmesi için üç yaklaşımda da Mel Frekansı Kepstrum Katsayıları (MFKK) esaslı öznitelikler kullanılmıştır. Bu öznitelikler, literatürde kulakta gerçekleşen zaman-frekans analizine oldukça benzemesiyle bilinmektedirler. Özniteliklerin işlenmesi noktasında önerilen üç yaklaşım birbirinden ayrılmaktadır.

İlk yaklaşımda, MFKK öznitelikleri orijinal halleri kullanılarak, bir gerçekçi sinir ağının (GSA'nın) konuşmacı kimliklerini tespit etmesi sağlanmıştır. MFKK

özniteliklerinin analog yapısı, GSA'lar için uygun olmadığından, bu tez çalışmasında, ayrıca bir öznitelik spike dönüşümü (ÖSD) yaklaşımı önerilmiştir. Özniteliklerin spike örüntülerine dönüştürülmesinin ardından, GSA'ların OKT uygulamasında performansı ölçülmüştür. İkinci yaklaşımda, MFKK öznitelikleri önceden eğitilmiş bir analog-analog öznitelik dönüştürücü derin sinir ağı (DSA) kullanılarak yeni öznitelik vektörlerine dönüştürülmüştür. Yine ÖSD kullanılarak elde edilen spike örüntüleri, GSA'ların performansının incelenmesinde kullanılmıştır. Üçüncü yaklaşımda, ÖSD yöntemi atlanarak, önceden eğitilmiş bir analog-ikili öznitelik dönüştürücü DSA yardımıyla MFKK öznitelikleri doğrudan spike örüntülerine dönüştürülmüştür. Bu spike örüntüleri yine GSA'ların OKT uygulamasındaki performansının incelenmesi için kullanılmıştır.

Bu bölümde, Şekil 21'de gösterilen yaklaşımlar için kullanılan materyal ve yöntemler tanıtılmıştır. Bölüm, altı alt başlıktan meydana gelmektedir. İlk alt başlıkta, çalışmada kullanılan veri kümeleri (VK'ler) tanıtılmıştır. İkinci alt başlıkta, GSA'nın oluşturulmasında kullanılan gerçekçi modellere dair parametre dizleri verilmiştir. Üçüncü bölümde, konuşma sinyallerinde kullanılmak üzere MFKK esaslı özniteliklerin hesaplanmasındaki detaylar açıklanmıştır. Dördüncü bölümde, GSA'nın genel yapısı tanıtılmış ve bu yapı içinde bulunan adımlar detaylı bir şekilde açıklanmıştır. Beşinci bölümde, Şekil 21'de gösterilen öznitelik dönüştürücü DSA modellerinden bahsedilmiş ve kullanılan yaklaşımın detaylarına değinilmiştir. Altıncı ve son bölümde ise beş farklı deney tasarlanmış, bu deneylerin amaçları, adımları ve detaylarından bahsedilmiştir.

2.1. Veri Kümeleri

Önerilen bir yöntemin performansının ölçülmesinde çeşitli yaklaşımlar önerilmiştir. Bunlar arasında en çok tercih edilenlerden biri, yöntemin avantaj ve dezavantajlarının kolaylıkla gözlemlenebileceği yapay bir VK'de test edilmesidir. Yapay VK'nin kullanımı, bir yöntemin kontrollü şartlar altında nasıl performans gösterdiği konusunda bilgi sağlaması nedeniyle, önemli bir yaklaşım olarak kabul edilmektedir. Bu sayede, yöntemin kısıtları ve kapasitesi önemli ölçüde belirlenebilmektedir. Fakat bu VK'lerin sentetik yapıları her ne kadar performans ölçümünde kolaylık sağlasa da gerçek bir dünya probleminde yöntemin nasıl performans göstereceği konusunda yeterli bilgiyi sağlamamaktadır (Arora & Arora, 2023). Bu nedenle, gerçek dünya VK'lerinin kullanımı da ayrıca gerekmektedir.

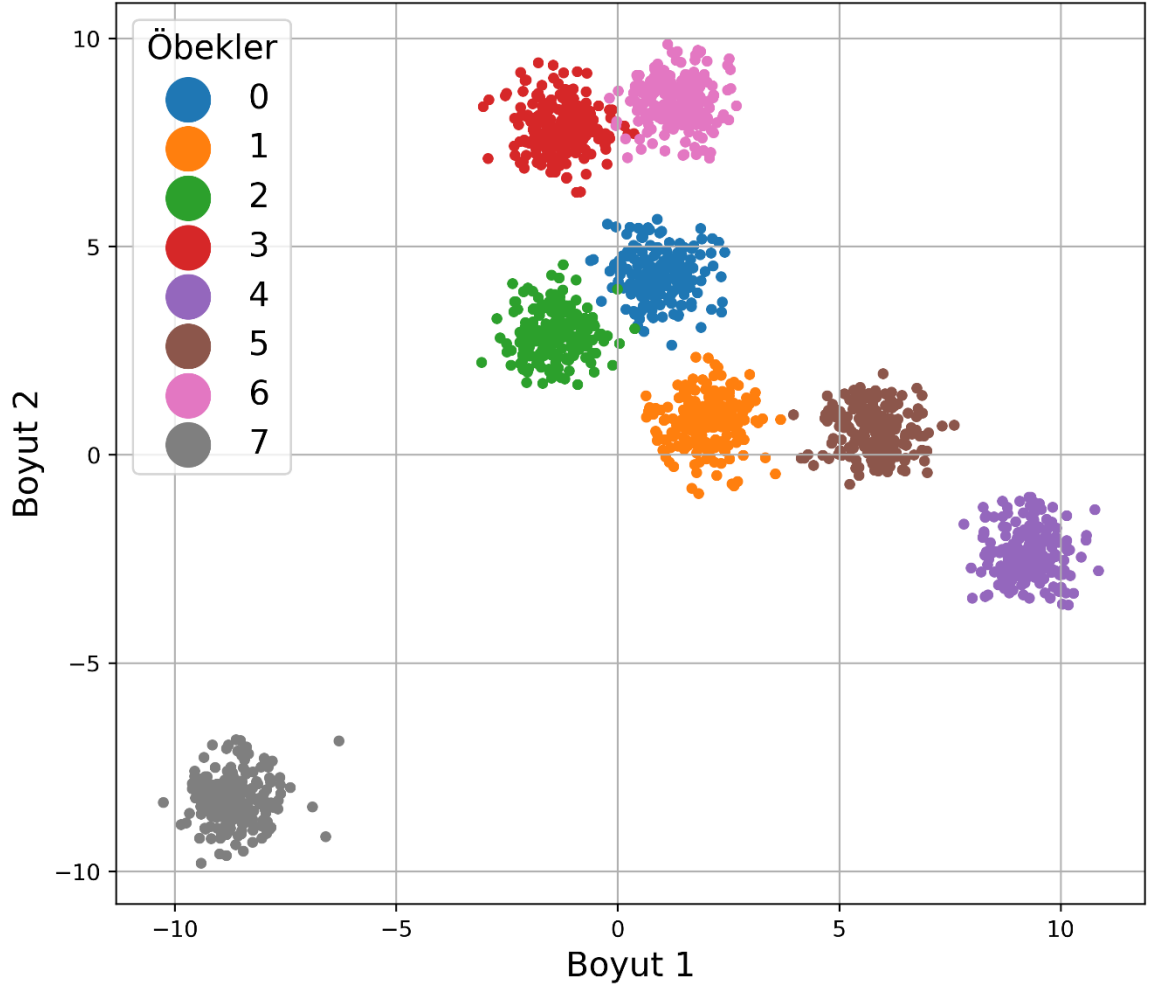
Bir diğer yaklaşım ise yöntemin; literatürde sıklıkla kullanılan, birçok paralel yöntemde performanslarının defalarca ölçülmüş ve doğrulanmış olduğu hazır VK'lerde test edilmesidir. Bu yaklaşıma göre, karşılaştırma VK veya benchmark dataset olarak bilinen VK'ler performans ölçümünde kullanılmaktadır. Bu VK'lerin paralel yöntemlerde performansları hali hazırda bilindiğinden, önerilen yöntemin performansı kolaylıkla karşılaştırmalı olarak belirlenebilmektedir (Thiyagalingam vd., 2022). Bu yaklaşım performans konusunda her ne kadar önemli bilgiler sağlasa da yapay VK'lerden elde edilebilecek bilgileri sağlamamaktadır. Ayrıca bu yaklaşım, bir yöntemin tüm VK'ler için başarılı olduğunu göstermemekte, sadece kullanılan VK özelinde performansı işaret etmektedir. Bu nedenle, önerilen bir yöntemin hedeflenen problem için özel olarak değerlendirilmesi de ayrıca gerekmektedir.

Bu çalışmada önerilen yöntemin performansının ölçülmesi ile ilgili deneylerde dört farklı VK kullanılmıştır. Bu VK'lerin ilki, yapay olarak üretilmiş bir VK'dir. Bu VK, önerilen yöntemin performansının görselleştirilmesi amacıyla çalışmaya dahil edilmiştir. IRIS ve MNIST, karşılaştırma VK'leri olarak kullanılmış ve yöntemin performansının evrensel olarak gösterilebilmesi için çalışmaya dahil edilmiştir (Fisher, 1936; Deng, 2012). Son olarak, OKT probleminde kullanılmak üzere Librispeech Automatic Speech Recognition Corpus (LASRC) veri tabanı çalışmaya dahil edilmiştir (Panayotov vd., 2015). Yapay VK ve karşılaştırmalı VK'lerden elde edilen tecrübeler ile konuşmacı tanıma probleminde yüksek performans elde edilmesi amaçlanmıştır.

2.1.1. Yapay Veri Kümesi

Bu çalışmada kullanılmak üzere, sanal olarak iki boyutlu ve sekiz öbekten meydana gelen bir yapay VK üretilmiştir. VK, her öbekte 200 adet örnek olmak üzere toplam 1600 örnekten meydana gelmektedir. VK'nin her bir öbeği, normal dağılıma sahiptir ve her iki boyutta da öbek içi örnekler 0,6 standart sapma değerine sahiptir.

VK'nin iki boyutlu olması, örnek dağılımlarının ve karar sınırlarının görselleştirilmesine olanak tanımakta ve sekiz öbekli olması da VK'nin çok sınıflı problemlerde, sınıflandırma performansının incelenmesinde önemli bir araç olmasını sağlamaktadır. VK'nin dağılımı Şekil 22'de görselleştirilmiştir. Bu VK, Python dilinde makine öğrenimi (MÖ) uygulamaları için yaygın olarak kullanılan scikit-learn kütüphanesinin `make_blobs` yöntemi ile üretilmiştir (Kramer, 2016).

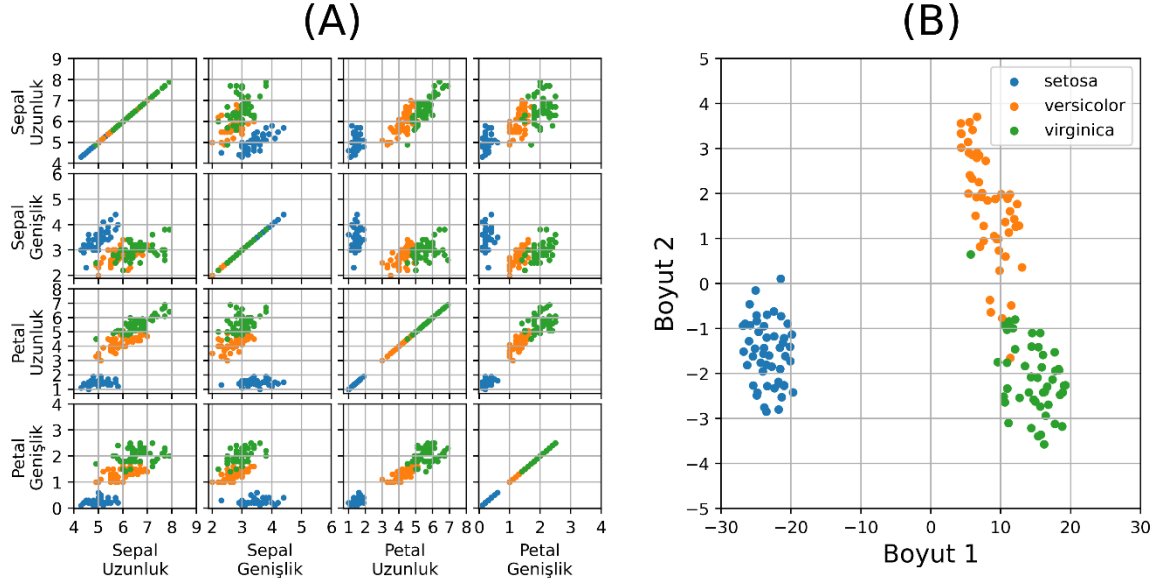


Şekil 22. Yapay veri kümesi öbekleri ve iki boyutlu uzayda dağılımları

2.1.2. Karşılaştırma Veri Kümeleri

Karşılaştırma VK'leri bir yöntemin performansının incelenmesinde sıklıkla kullanılmaktadırlar. Bu çalışmada, karşılaştırma VK'leri hem yöntemin performansının ölçülmesinde kullanılmış hem de hedef VK'de performansın artırılması için uygun parametrelerin belirlenmesi konusunda öngörü sağlamıştır.

Çalışmada kullanılan ilk karşılaştırma VK olan IRIS, ismini aldığı çiçeklerinin türünün belirlenmesi için kullanılan üç sınıflı ve dört öznitelik boyutuna sahip bir VK olarak paylaşılmıştır (Fisher, 1936). Bu VK'de öznitelikler, iris çiçeklerinin çanak (sepal) ve taç (petal) yapraklarının uzunlukları ve genişlikleridir. Bu VK'de hedef, verilen bu özniteliklerden faydalanılarak *setosa*, *versicolor* ve *virginica* türlerinin doğru bir şekilde

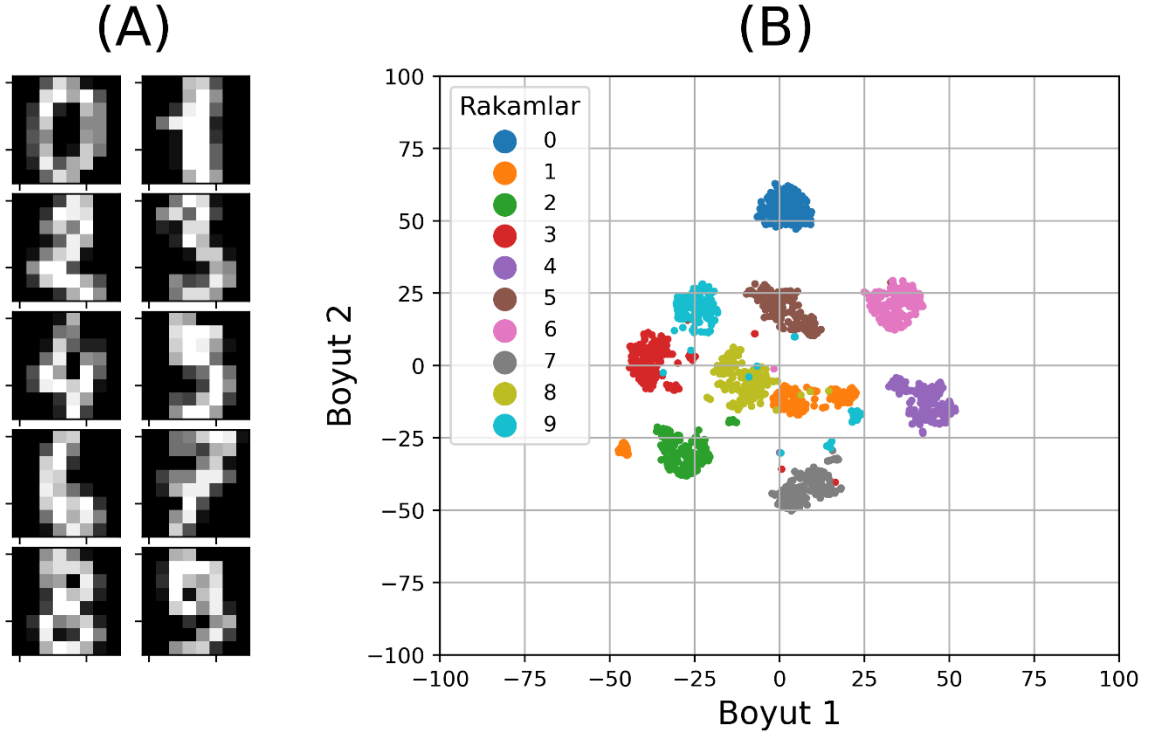


Şekil 23. IRIS veri kümesi ve iki boyutta sınıflarının dağılımı

sınıflandırılmasıdır. Şekil 23(A)'da IRIS VK'nin sınıfların dağılımı, öz nitelikler ikili halde eşleştirilerek verilmiştir. Şekil 23(A), VK'nin sınıflarının ayrılabilirliği konusunda yeterli bilgiyi sağlamadığından, VK, t-Dağıtılmış Stokastik Komşu Gömme (t-SKG) yöntemi kullanılarak iki boyuta indirgenmiştir (van der Maaten & Hinton, 2008). Şekil 23(B) bu indirgenmiş dağılımı göstermektedir.

Çalışmada kullanılan bir diğer karşılaştırma VK olan MNIST el yazısı rakamları, sıfırdan dokuzaya kadar olan rakamların el yazısı ile yazılmış görüntülerini içermektedir (Deng, 2012). Bu VK'nin elde edildiği scikit-learn kütüphanesinde, her biri 8x8 çözünürlükte görüntüler olmak üzere toplam 1797 örnek bulunmaktadır. Bu VK'de sınıf başına örnekler, dengeli sayılabilecek şekilde ayrılmış olup yaklaşık 180 adettir. Şekil 24 (A), her bir sınıftan bir adet olmak üzere VK'den örnekler göstermektedir. Her bir örnek 8x8 görüntü formatından düzleştirilerek 64 adet öz niteliğe sahip bir forma dönüştürüldüğünde çalışmada kullanılmak üzere uygun hale getirilmiştir. Düzleştirilen bu VK'ye t-SKG uygulanarak iki boyuta indirgendiğinde sınıflar Şekil 24(B)'de gösterildiği gibi dağılmaktadır.

Şekil 24(B)'den anlaşılacağı üzere on sınıflı MNIST problemi, ayrılması zor sınıflara sahiptir. MNIST'in çözülmesi daha zor ve daha yüksek miktarda sınıf sayısına sahip olması IRIS'ten ayrıldığı ilk özelliktir. Bunun yanında, MNIST görüntüleri 0 ila 15 arası değişen kuantalanmış piksel değerlerine sahiptir. Ayrıca, görüntüler düzleştirildikten sonra



Şekil 24. MNIST veri kümesi ve iki boyutta sınıflarının dağılımı

çok daha fazla öznitelik sayısına sahiptir. Çalışmada kullanılan IRIS ve MNIST, Python dili scikit-learn kütüphanesinden kolaylıkla elde edilebilir durumdadır (Kramer, 2016).

2.1.3. Librispeech Veri Kümesi

OKT uygulamalarında; NIST SRE (Sadjadi vd., 2022), VoxCeleb (Nagrani vd., 2017) ve LASRC (Panayotov vd., 2015) olmak üzere birçok VK kullanılmaktadır. Bu VK'lerin kullanılmasında konuşma miktarı, konuşmacı sayısı, konuşma kayıtlarının kalitesi, işaret-gürültü oranı ve ulaşılabilirliği gibi çeşitli özelliklere dikkat edilmekte; bu özelliklere göre hangi VK'nin kullanılması gerektiği belirlenmektedir (Sztahó vd., 2021). NIST SRE, telefon görüşmelerinden elde edilen konuşma kayıtlarından meydana gelirken; VoxCeleb, YouTube videolarından elde edilmiş kayıtları içermektedir. Bu nedenle, OKT uygulamalarında bu VK'ler, konuşma kayıtlarının kalitesinin ve işaret-gürültü oranlarının düşük olması nedeniyle ilk tercih olmaktan uzaktırlar. Sztahó ve arkadaşları (2021) çalışmalarında, LASRC'nin OKT uygulamalarında kullanılabilecek en iyi VK olduğu savunmuşlardır. Bunu nedeni olarak, bu VK'deki kayıtların, yüksek kalitede ve işaret

Tablo 2. LASRC veri tabanı ve bileşenleri (Panayotov vd., 2015)

Subset	Kadın Sayısı	Erkek Sayısı	Toplam
dev-clean	20	20	40
test-clean	20	20	40
dev-other	16	17	33
test-other	17	16	33
train-clean-100	125	126	251
train-clean-360	439	482	921
train-other-500	564	602	1166
Toplam			2484

gürültü oranının ise yüksek olmasını öne sürmüşlerdir. Bu öneriden referansla bu tez çalışmasında LASRC veri tabanı kullanılmıştır.

LASRC her biri İngilizce sesli kitaplardan elde edilen 16 KHz ile örneklenmiş 1000 saatlik konuşma kayıtlarından meydana gelmektedir. Toplam 2484 kişilik konuşmacı havuzuna sahip olan bu VK alt kümelerine ayrılmış olarak tanıtılmıştır. Bu alt kümeler ve konuşmacı sayıları Tablo 2’de detaylı olarak verilmektedir (Panayotov vd., 2015).

LASRC her ne kadar geniş bir konuşmacı havuzuna sahip olsa da bu çalışma için ihtiyaç duyulan konuşmacı miktarı kısıtlıdır. Bu nedenle, VK’nin alt kümelerinden biri olan dev-clean çalışmada ana VK olarak kullanılmıştır. Daha sonrasında bu VK, çeşitli deneylerde kullanılmak üzere daha küçük dört ayrı alt kümeye ayrılmıştır. Bu VK’ler oluşturulurken rastgele konuşmacılar seçilmiş ve her VK için eşit sayıda kadın ve erkek konuşmacının olmasına özen gösterilmiştir. Bu alt kümeler, Tablo 3’te detaylı olarak gösterilmiştir.

Tablo 3’te, dev-clean konuşma kayıtlarından rastgele seçilmiş 16 konuşmacının, veri tabanında geçen isim, ID, cinsiyet ve toplam konuşma kayıt uzunlukları verilmiştir. Seçilen 16 konuşmacı, farklı deneysel senaryolarda kullanılmak üzere dört alt kümeye ayrılmıştır. Bu alt kümeler sırasıyla dört, altı, sekiz ve on altı konuşmacıya sahip olacak şekilde ayrılmıştır. Konuşmacı sayısı yüksek olan alt kümeler, kendinden küçük alt kümelerin üstüne yeni konuşmacılar eklenmesiyle oluşturulmuştur.

Tablo 3, oluşturulan alt kümelerde seçilen konuşmacıların bilgilerini de içermektedir. Çalışmanın ana VK’si dev-clean olarak seçilmiş olsa da gerekli görülen deneylerde Tablo

Tablo 3. Konuşmacılar ve çalışmanın veri kümeleri

No	İsim	ID	Cinsiyet	Toplam Konuşma Uzunluğu (sn)	Veri Kümesi			
					VK 4	VK 6	VK 8	VK 16
1	Mark Nelson	251	Erkek	482,4	√	√	√	√
2	President Lethe	422	Erkek	502,8				√
3	E. Tavano	1462	Kadın	482,4	√	√	√	√
4	Sharon Bautista	2035	Kadın	486,6				√
5	calystra	2412	Kadın	483,6			√	√
6	Brianvon Dedenroth	3000	Erkek	481,8	√	√	√	√
7	VOICEGUY	3170	Erkek	486				√
8	Arielle Lipshaw	3536	Kadın	489				√
9	SR Colon	5338	Kadın	484,2				√
10	badey	6241	Erkek	483				√
11	Michael Packard	6295	Erkek	482,4			√	√
12	Jean Bascom	6345	Kadın	484,2				√
13	Jill Engle	7850	Kadın	483,6	√	√	√	√
14	Jennifer Rutters	7976	Erkek	487,8		√	√	√
15	David Mecionis	8297	Erkek	482,4				√
16	Mary J	8842	Kadın	486		√	√	√
Toplam					4	6	8	16

2 verilen diğer VK'ler de kullanılmıştır. Buna göre, test-clean, dev-other ve train-clean-100 VK'leri deneysel çalışmalar içinde çeşitli yerlerde kullanılmıştır. Bu VK'lerin kullanım nedenleri ve kullanım şekillerine, Bölüm 2.6'da tanıtılan deneysel çalışmaların detaylarından ulaşılabilmektedir.

2.2. Gerçekçi Modeller İçin Seçilen Parametreler

Bu bölümde, çalışma boyunca kullanılan hesaplamalı modellere dair seçilen parametreler tanıtılmış ve değerleri verilmiştir. Bölüm, üç alt başlıktan meydana gelmektedir. İlk alt başlıkta; kullanılan basit hücre modelleri olan Sızıntılı Topla&Ateşle (STA), Hodgkin&Huxley (HH) ve Izhikevich (IZH) modellerine dair parametreler verilmiştir. İkinci alt başlıkta, kullanılan piramidal nöron (PN) modelinin bölgeleri ve

dendritik organizasyonu tanıtılmış ve modele dair morfolojik ve biyofiziksel parametreler verilmiştir. Üçüncü ve son alt başlıkta ise kullanılan sinaptik bağlantı (SB) modelleri ve parametreleri verilmiştir.

2.2.1. Basit Nöron Modelleri

Basit nöron modelleri işlemsel yüklerinin PN gibi detaylı gerçekçi nöron (GN) modellerine kıyasla hafif oluşu ile bilinmektedirler. Bu nedenle, GSA'ların karmaşık simülasyonunda sıklıkla tercih edilmektedir (Gerstner vd., 2014). Bu GN'lerin sınıflandırma performansının ölçülmesi önerilen tez çalışmasının hedeflerinden biridir. Bu nedenle, basit nöron modelleri çalışmaya dahil edilmiştir. Çalışmada üç farklı basit nöron modeli kullanılmıştır.

Çalışmada kullanılan ilk basit GN modeli STA nöronudur. STA nöronu, Bölüm 1.3.3'te detaylıca tanıtılmıştır. Literatürde sıkça tercih edilen bir GN modeli olan STA nöronun en dikkat çekici özelliği işlem yükünün çok düşük olmasıdır (Gerstner vd., 2014). Bu nedenle, donanım imkanlarının kısıtlı olan uygulamalarda basit çözümler sunmaktadır. Bu çalışmada, Denklem (23) ve (24)'te verilen STA nöron modelinden faydalanılmıştır. Denklem (23) için $\tau_{zar} = R_{zar} \times C_{zar}$ ve $g_{zar} = 1/R_{zar}$ olduğu göz önünde bulundurulduğunda, Denklem (57) elde edilmektedir. Tablo 4, STA nöronu için Denklem (57) ve (24)'te verilen modelin çözümünde kullanılan parametre dizilerini göstermektedir.

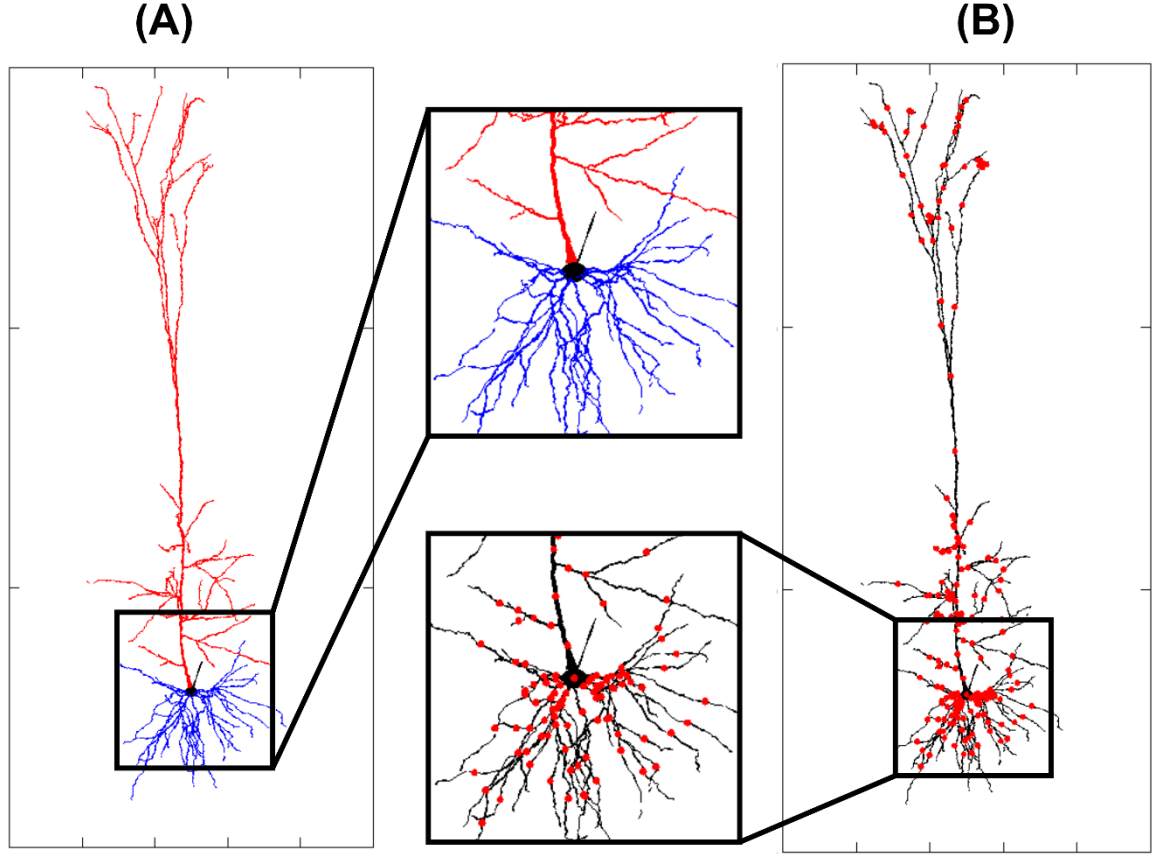
$$C_{zar} \frac{dV}{dt} = -g_{zar}(V - E_{Dinlenme}) + I_{giriş}(t) \quad (57)$$

Çalışmada kullanılan ikinci basit nöron modeli HH nöronudur. HH nöronu, Bölüm 1.3.4'te detaylıca tanıtılmıştır. Bu GN, literatürdeki tek bölmeli GN modelleri arasında en gerçekçi model olarak kabul edilmektedir (Gerstner vd., 2014). Bunun nedeni, hem aktif ve pasif iyon kanallarını hesaba katması hem de zar özelliklerini göz önünde bulundurmasıdır. Fakat biyolojik nöronun (BN) modellenmesinde bu kadar detaylı bir yaklaşım, aynı zamanda işlemsel yükü de beraberinde getirmektedir. Bu nedenle, HH nöronu STA nöron ile karşılaştırıldığında işlem yükü nispeten daha ağırdır. Bu çalışmada, Denklem (26)'da verilen HH nöron modelinden faydalanılmıştır. Tablo 4, HH nöronu için Denklem (26)'nın çözümünde kullanılan parametre dizilerini göstermektedir.

Tablo 4. Basit nöron modelleri ve parametreleri

Nöron Modeli	Parametre	Değer	Birim
STA	$V_{Dinlenme}$	-60	mV
	$V_{Eşik}$	-45	mV
	V_{Reset}	-60	mV
	V_{Tepe}	35	mV
	C_{zar}	0.45	mF
	g_{zar}	0.001	μS
HH	E_{Na}	55.17	mV
	E_K	-72.14	mV
	$E_{Sızıntı}$	-49.42	mV
	g_{Na}	1.2	μS
	g_K	0.36	μS
	$g_{Sızıntı}$	0.003	μS
	C_{zar}	0.01	mF
IZH	$V_{Dinlenme}$	-60	mV
	$V_{Eşik}$	-40	mV
	V_{Tepe}	35	mV
	C_{zar}	1	mF
	k	0.7	nS/mV
	a	0.03	$1/ms$
	b	-2	nS
	c	-50	mV
	d	100	pA

Çalışmada kullanılan üçüncü basit nöron modeli IZH nöronudur. IZH nöronu, Bölüm 1.3.5'te detaylıca tanıtılmıştır. Bu GN, diğer GN'lere kıyasla en güncel GN modelidir. STA nöronun BN'nin dinamiklerini yansıtmada eksik kalması, HH nöronun ise ağır işlemsel yükü, bu GN'nin kullanım alanının genişlemesine neden olmuştur. Literatürde IZH nöronu hem basit ve hızlı çözüme sahip bir model olması hem de BN'nin dinamiklerini başarılı bir şekilde yansıtmayla bilinmektedir (Izhikevich, 2006). Bu çalışmada, Denklem (33), (34) ve (35)'te verilen IZH nöron modelinden faydalanılmıştır. Tablo 4, IZH nöronu için modelin çözümünde kullanılan parametre dizilerini göstermektedir.



Şekil 25. Çalışmada kullanılan piramidal nöron modeli

2.2.2. Piramidal Nöron Modeli

PN'ler beynin çeşitli bölgelerinde sıkça rastlanılan ve bilişsel süreçlerde önemli roller oynadığı gösterilen BN'lerdir (Spruston, 2008). Bu BN'ler kendilerine has morfolojik ve biyofiziksel özellikleri ile bilinmektedirler. Sinir sisteminde sıkça bulunmaları ve detaylı morfolojik özellikleri sayesinde doğrusal olmayan problemleri çözebilme yetenekleri, bu BN'lerin sınıflandırma performanslarının merak edilmesine neden olmaktadır (Polsky vd., 2004). Bu nedenle, çalışmada sınıflandırma performansını ölçebilmek adına bir PN modeli kullanılmıştır.

Çalışmada kullanılan PN, Hay ve arkadaşlarının (2011) çalışmasından alınan kortikal bir PN modelidir. İlgili çalışmada, üç farklı PN kullanılmakla beraber, bu çalışmaya sadece Şekil 25'te gösterilen model dahil edilmiştir. Bu model aynı zamanda Moldwin ve Segev (2020) tarafından da sınıflandırma amacıyla kullanılmış ve yapay bir VK'de performansı incelenmiştir.

Tablo 5. Piramidal nöron için bölgeler ve kullanılan parametreler

	Parametre	Somatik	Apikal	Basilar	Aksonal	Birim
Morfoloji	<i>Bölme Sayısı</i>	1	109	84	2	<i>Adet</i>
	<i>Uzunluk</i>	23,17	68,26	61,11	30	μm
	<i>Çap</i>	13,4	0,89	0,67	1	μm
Zar	R_a	100	100	100	100	$\Omega \times cm$
Parametreleri	C_{zar}	1	2	2	1	$\mu F/cm^2$
İyon Potansiyelleri	$E_{Sızıntı}$	-90	-90	-90	-90	mV
	E_K	-85	-85	-85	-85	mV
	E_{Na}	50	50	50	50	mV
Pasif Kanal	$g_{Sızıntı}$	$3,38 \times 10^{-5}$	$5,89 \times 10^{-5}$	$4,67 \times 10^{-5}$	$3,25 \times 10^{-5}$	S/cm^2
K1	g_{Kanal}	$2,00 \times 10^{-4}$	$2,53 \times 10^{-4}$	2×10^{-4}	-	S/cm^2
	E_{Kanal}	-45	-45	-45	-	mV
K2	g_{Kanal}	2,04	$2,13 \times 10^{-2}$	-	-	S/cm^2
K3	g_{Kanal}	$1,72 \times 10^{-3}$	-	-	-	S/cm^2
K4	g_{Kanal}	-	$6,75 \times 10^{-5}$	-	-	S/cm^2
K5	g_{Kanal}	$6,93 \times 10^{-1}$	$2,61 \times 10^{-4}$	-	-	S/cm^2
K6	g_{Kanal}	$8,12 \times 10^{-2}$	-	-	-	S/cm^2
K7	g_{Kanal}	$2,23 \times 10^{-3}$	-	-	-	S/cm^2
K8	g_{Kanal}	$4,41 \times 10^{-2}$	$1,20 \times 10^{-3}$	-	-	S/cm^2
	γ	$5,01 \times 10^{-4}$	$5,09 \times 10^{-4}$	-	-	<i>Birimsiz</i>
	τ_{decay}	460	122	-	-	ms
	<i>Depth</i>	0,1	0,1	-	-	μm
	$[Ca^{2+}]_{min}$	$1,00 \times 10^{-4}$	$1,00 \times 10^{-4}$	-	-	mM
K9	g_{Kanal}	$3,43 \times 10^{-3}$	$1,87 \times 10^{-4}$	-	-	S/cm^2
K10	g_{Kanal}	$9,92 \times 10^{-4}$	$5,55 \times 10^{-5}$	-	-	S/cm^2

Kullanılan GN, memeli neokorteksinin 5B katmanında bulunan bir PN'nin detaylı modelidir. Bu yüzden L5BPC olarak isimlendirilmektedir. Bu GN'nin en önemli özelliği neokorteksin altı katmanı ile de bağlantılara sahip olmasıdır (Hay vd., 2011). Bu özelliği, karmaşık hesaplamaları yapabilme yetisi kazandığı anlamına gelmektedir.

Hay ve arkadaşlarının (2011) çalışmasında bahsedilen PN, dört ana bölge halinde incelenmektedir. Bu bölgeler somatik, apikal, basilar ve aksonal bölgeler olarak belirlenmiştir. PN, 1'i somatik bölgede, 2'si aksonal bölgede, 84'ü basilar bölgede ve 109'u apikal bölgede olmak üzere toplam 196 bölmeden meydana gelmektedir. Şekil 25(A)'da; bu bölgeler apikal bölge kırmızı, basilar bölge mavi ve geri kalanı ise siyah olacak şekilde gösterilmiştir.

Ayrıca, Hay ve arkadaşları (2011), PN'nin biyofiziksel özelliklerini de belirlemişler ve deneysel çalışmalardan elde edilen kayıtlardan faydalanarak parametreleri

kestirmişlerdir. Daha yakın zamanda, Shai ve arkadaşları (2015) aynı PN için kendi deneysel verilerini kullanmış ve biyofiziksel parametreleri yeniden kestirmişlerdir. Bu tez çalışmasında, Moldwin ve Segev'in (2020) çalışmasından referansla, Shai ve arkadaşlarının (2015) önerdiği biyofiziksel parametreler kullanılmıştır. Bu parametreler Tablo 5'te her bölge için ayrı ayrı verilmiştir. Tablo 5'te verilen parametrelerde, bir değerin koyu yazılması, o değerin ortalama bir değer olduğunu göstermektedir.

Tablo 5 incelendiğinde, biyofiziksel özelliklerin her bölge için çeşitlilik gösterdiği görülmektedir. Aksonal bölgede, sadece pasif kanallar bulunmaktadır. Bunun nedeni, PN'nin morfolojisinin tespit edilmesinde aksonal bölge detaylarının göz ardı edilmesidir. Basilar bölgede, sadece K1 kanalı ve pasif kanallara sahip olduğu görülmektedir. Somatik bölgede, K4 hariç olmak üzere aktif iyon kanalları ve pasif kanallar mevcuttur. Apikal bölgede, K1, K2, K4, K5, K8, K9, K10 aktif iyon kanalları ve pasif kanallar mevcuttur. Tablo 5'te verilen aktif kanalların detayları ve parametrelerin kullanıldığı denklemler, EK-1 başlığı altında verilmektedir. Somatik ve apikal bölgeler incelendiğinde her iki bölgede de aktif kanalların mevcut olduğu görülebilmektedir. Fakat apikal bölgedeki aktif kanallar kendi başlarına AP üretebilecek yeteneğe sahip değilken, somatik bölge AP üretimini mümkün kılan aktif kanallara sahiptir.

2.2.3. Sinaps Modeli

Sinapslar BN'lerin dış dünyadan bilgi alıp işlemesini sağlayan reseptörlerdir. Sinapslar hakkında detaylı bilgi, Bölüm 1.1.4'te verilmiştir. Ayrıca, Bölüm 1.3.6, sinapsların nasıl modellendikleri konusunda detayları göstermektedir. Bu çalışmada, GN'lerin spike örüntülerini alıp işleyebilmeleri adına SB modelleri kullanılmıştır. Bu bölümde, sinapsların hücrelere nasıl yerleştirildiği, nasıl çalıştıkları ve türlerinden bahsedilmektedir.

GN modelleri başlangıçta herhangi bir sinaptik reseptöre sahip değildirler. Bu nedenle, sinaptik reseptörlerin GN yüzeyine yerleştirilmesi gerekmektedir. Bu yerleştirme işlemi seçilen nöronun basit bir GN modeli mi yoksa çok bölmeli bir model mi olduğuna göre farklılık göstermektedir. Eğer seçilen nöron, bir basit GN modeli ise nöron modelinin tek bir bölmeden meydana geldiği anlaşılmalıdır. Bu nedenle, GN'nin yüzeyindeki her bir nokta teorik olarak aynı noktaya karşılık gelmektedir. Bu nedenle, çalışmada eğer bir basit

Tablo 6. Sinaps modelleri ve kullanılan parametreler

Bağlantı Tipi	Sinaps	τ_{syn}	E_{syn}
Exc	AMPA	5 ms	0 mV
	NMDA	150 ms	0 mV
Inh	GABA _A	6 ms	-70 mV
	GABA _B	150 ms	-90 mV

GN modeli seçildiyse bu GN modelinin merkezine tek bir sinaptik reseptör yerleştirilmektedir.

Seçilen nöronun çok bölmeli bir model olan PN modeli olması durumunda, sinaptik reseptör konumunun GN performansını etkilediği gösterilmiştir (Limbacher & Legenstein, 2020). Bunun nedeni, PN'nin asıl işlemini gerçekleştirdiği somadan uzaklaştıkça, girdilerin etkinliğinin uzaklığa bağlı olarak değişiyor olmasıdır. Sinaptik parametreler ve ağırlıkların eşit olduğu bir durumda, somaya daha yakın olan sinaptik bağlantı PN'nin AP üretmesi konusunda daha yüksek etkiye sahiptir. Bu nedenle, sinaptik reseptörler PN yüzeyine uzamsal olarak serpiştirilmelidir.

PN'nin bir bölmesindeki her bir noktanın teorik olarak aynı nokta olduğu göz önünde bulundurulduğunda, işlemsel yükün azaltılması adına, her bir bölme için bir sinaptik reseptör yerleştirilmesi yeterli olacaktır. Bu nedenle, çalışmada kullanılan PN'nin her bir bölmesine girdi sayısından bağımsız olacak şekilde, birer sinaptik reseptör doğrudan yerleştirilmiştir. Kullanılan PN'nin toplam 196 bölme sahip olduğu göz önünde bulundurulduğunda, PN'ye toplam 196 sinaptik reseptör yerleştirilmiştir. Şekil 25(B), yerleştirilen sinaptik reseptörlerin konumunu göstermektedir.

Sinaptik reseptörlerin konumunun yanında içerikleri de ayrıca önemlidir. Denklem (38), sinaptik reseptörlerin bağlı oldukları bölme için toplam akım girdilerinin modelini göstermektedir. Bu denkleme göre, SB'ler bir bölme içinde ayrı ayrı bulunsalar dahi tek bir toplamsal ifade ile gösterilebilmektedir. Bu bağlamda, sinaptik reseptörlerin birden fazla SB'yi ifade edebilecekleri anlaşılabilmektedir. Bu nedenle, çalışmada her bir sinaptik reseptörde ikisi Exc ve diğer ikisi Inh olmak üzere toplam dört SB türü toplamsal olarak kullanılmıştır. SB'lerin toplam akım katkıları Denklem (38)'de gösterildiği gibi hesaplanmış ve g_{sinaps} değerleri Denklem (37) yardımıyla elde edilmiştir.

Bu çalışmada iki farklı Exc SB modeli kullanılmıştır. Bunlardan ilki, hızlı Exc sinaps olarak da bilinen *α -Amino-3-Hidroksi-5-Metil-4-Izoksazolpropiyonik asit* (AMPA) sinaps modelidir (Kessels & Malinow, 2009). Bu SB, zaman sabitinin düşük olması nedeniyle bu isimle anılmaktadır. İkinci Exc SB ise NMDA modelidir (Malenka & Nicoll, 1993). Bir diğer ismi yavaş Exc olan bu SB, zaman sabitinin yüksek olması nedeniyle sinaptik etkinin postsinaptik GN’de uzunca bir şekilde etkinlik göstermesine neden olmaktadır.

Bu çalışmada, özgün bir yaklaşım olarak Inh SB’ler da kullanılmış ve performansları incelenmiştir. Inh SB olarak GABA sinapslarının iki türevi kullanılmıştır (Mody vd., 1994). Bu SB’lerin ilki olan A tipi, kısa zaman sabiti sayesinde hızlı bir etki göstermekte, diğer türev olan B tipi ise uzun bir süre boyunca postsinaptik GN’yi baskılamasıyla ön plana çıkmaktadır. Çalışmada kullanılan SB modellerine ait parametreler Tablo 6’da verilmiştir.

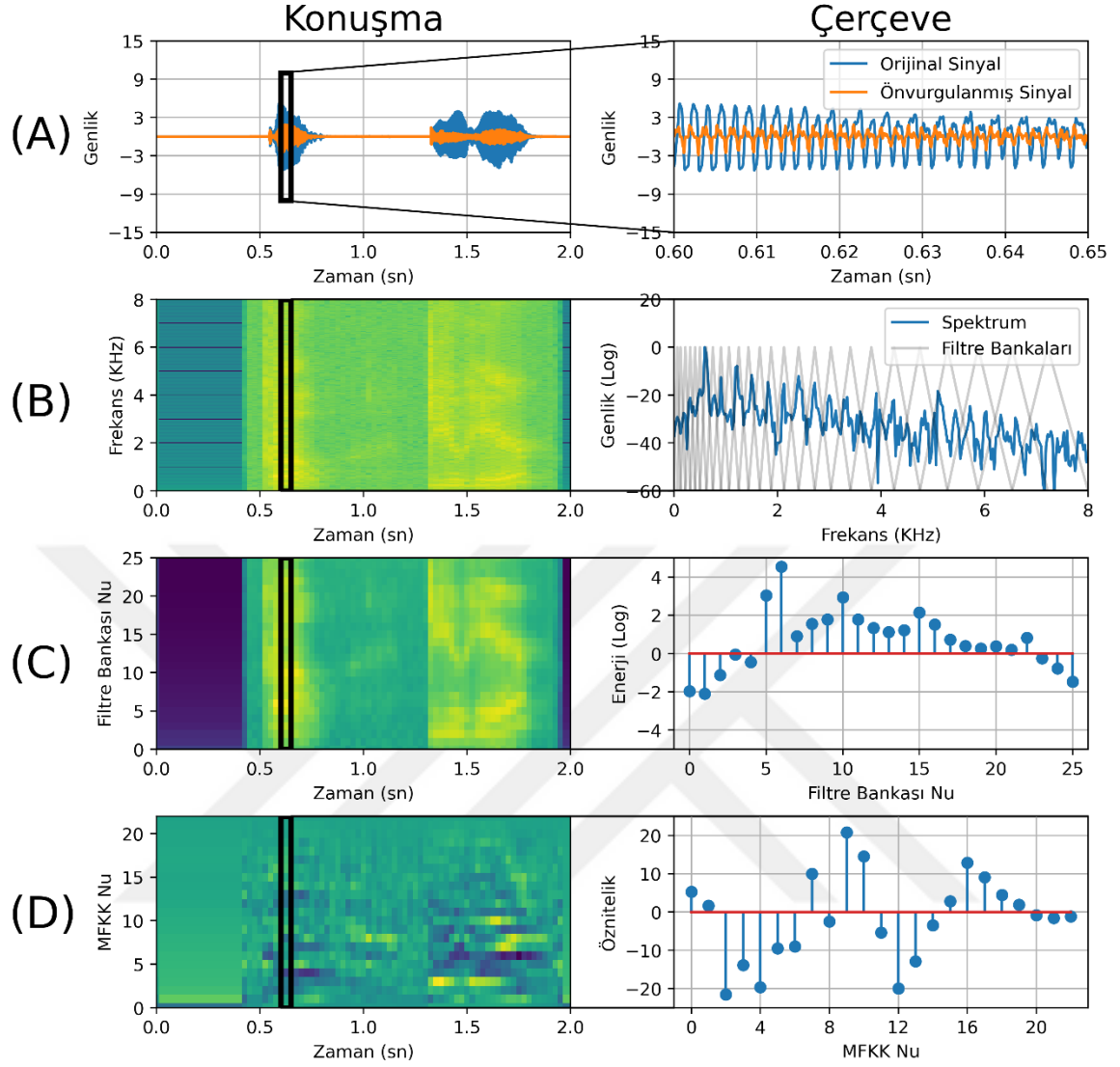
2.3. MFKK Esaslı Öznitelik Çıkarımı

Bu çalışmada, konuşma kayıtlarından MFKK esaslı öznitelikler elde edilmiş ve kullanılmıştır. MFKK yönteminin matematiksel temelleri, Bölüm 1.4.1’de detaylıca anlatılmıştır. Bu bölümde, MFKK çıkartma adımları boyunca çalışmada kullanılan parametreler verilmiş ve adımlar gösterilmiştir.

Örnekleme frekansı 16 KHz olan, 2 saniye uzunluğunda bir $x[n]$ konuşma sinyali göz önünde bulundurulduğunda, Şekil 26 konuşma sinyali üzerinden MFKK özniteliklerinin hesaplaması için kullanılan adımları göstermektedir. Şeklin sol tarafında konuşma sinyalinin tamamı için adımlar gösterilirken, sağ tarafında ise bir çerçeve için adımlar gösterilmiştir.

İşlem, Denklem (46)’da verilen ifadeden faydalanılarak sinyalin ön-vurgulanmasıyla başlamaktadır. Bu çalışmada ön-vurgulama katsayısı $\alpha = 0.97$ olarak seçilmiştir. Ön-vurgulanmış sinyal daha sonra çerçevelere ayrılmaktadır. Bu çalışmada çerçeve uzunluğu 50 ms olarak seçilmiştir. Bunun nedeni GN’lerin çıkış üretebilmek için en az 50 ms işlem yapmak zorunda olmalarıdır.

Sonrasında, çerçevelere Denklem (47)’de gösterildiği gibi bir pencere fonksiyonu uygulanması adımı gelmektedir. Bu çalışmada bu adım atlanmış, bir başka ifadeyle $h_{pen}[n] = u[n]$ olduğu kabul edilmiştir. Bu işlemin sonucu olarak 2 saniye uzunluğundaki konuşma sinyalinden 40×800 boyutlarında bir vektörler matrisi ele edilmiştir. Şekil 26(A)



Şekil 26. MFKK hesaplama adımları

orijinal ve ön-vurgulanmış sinyalleri göstermektedir. Sol taraf konuşma sinyalinin tamamını, sağ taraf ise seçilmiş bir çerçeveyi göstermektedir. Şeklin sonraki adımları aynı konuşma sinyali ve çerçevesi ile elde edilmiştir.

Sonraki adım, çerçevelerin Denklem (48)'de verilen ifade kullanılarak Ayrık Fourier Dönüşümünün (AFD) alınmasıdır. Bu çalışmada, AFD yerine, Hızlı Fourier Dönüşümü (HFD) algoritması kullanılmıştır. HFD algoritmasında hesaplanacak spektrum için çözünürlük 512 olarak seçilmiştir. Bu aşamada konuşma sinyali 40×512 boyutundadır. Şekil 26(B), konuşma sinyali ve çerçeve için HFD sonucunu logaritmik tabanda göstermektedir. Sol taraf, Kısa Zamanlı Fourier Dönüşümü olarak da bilinmektedir. Sonraki adım, Mel frekans uzayında eşit aralıklarla yerleştirilmiş olan filtre bankalarının

enerjilerinin hesaplanmasıdır. Şekil 26(B)'nin sağ tarafında bir çerçeve için filtre bankaları gösterilmektedir. Filtre bankası fonksiyonu olarak Denklem (51)'de verilen üçgen kullanılmıştır.

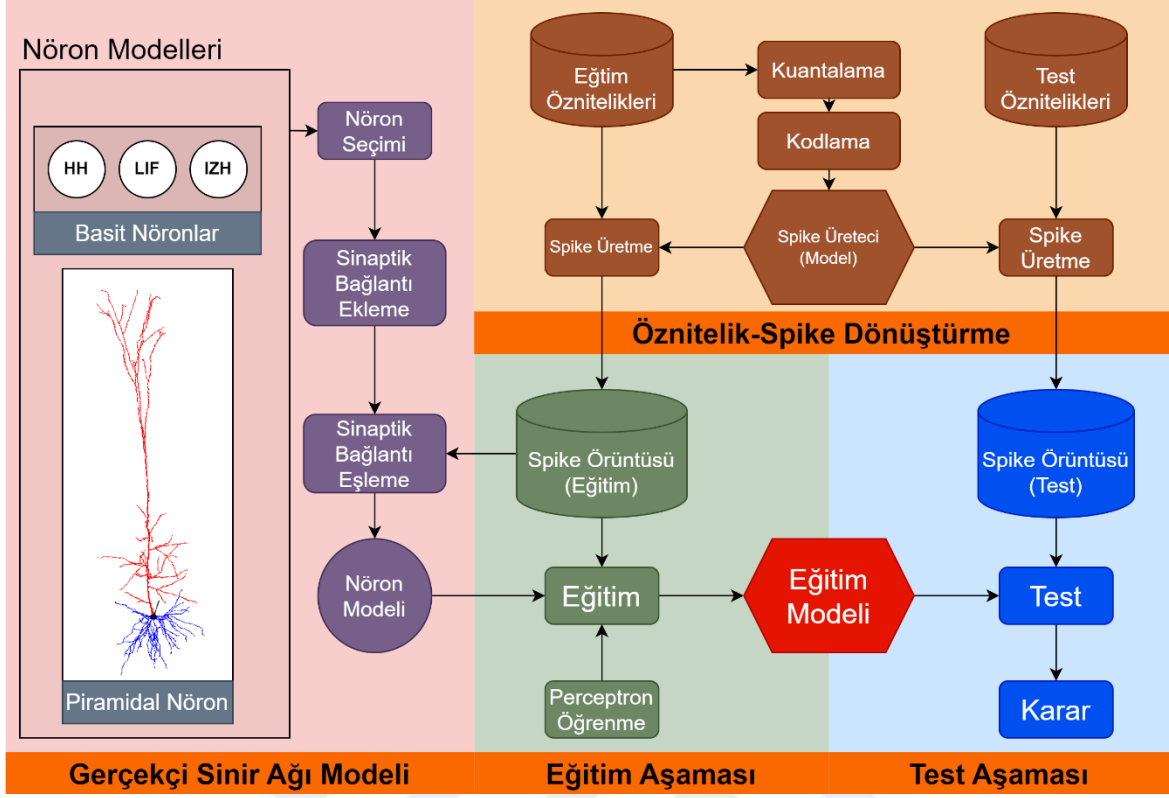
Filtre bankalarının enerjilerinin hesaplanmasında Denklem (50) kullanılmış ve sonuçlar Şekil 26(C)'de gösterilmiştir. Toplam filtre banka sayısı 26 olarak seçilmiştir. Bu aşamada konuşma sinyali 40×26 boyutuna getirilmiş durumdadır. Son olarak, enerji vektörleri Denklem (52)'de verilen Ayrık Kosinüs Dönüşümü (AKD) kullanılarak MFKK özniteliklerine dönüştürülmüştür. MFKK sayısı 23 seçilmiştir. Bu aşamada konuşma sinyali 40×23 boyutuna gelmiştir. Şekil 26(D) konuşma sinyali ve çerçeve için MFKK özniteliklerini göstermektedir. Bu adımlar tüm konuşma kayıtları için aynen uygulanmaktadır.

2.4. Gerçekçi Sinir Ağı Modeli

Bu çalışmada önerilen GSA modeli, öznitelik vektörleri olarak ifade edilen verileri sınıflandırması amacıyla tasarlanmıştır. Eğitim ve test aşamalarından önce VK'nin eğitim ve test öznitelikleri olarak ayrılmış olması gerekmektedir. Bu öznitelik vektörleri, Şekil 27'de gösterildiği gibi eğitim ve test aşamalarına tabii tutulmakta ve performans incelenmektedir. Başlangıçta, öznitelik vektörleri analog değerlere sahiptir ve GSA'nın işleyebilmesi için ikili değerler halinde ifade edilmelidirler. Burada, 1 değeri GSA'ya spike girişi olacağını, 0 değeri ise spike girişi olmayacağını ifade etmektedir. Bu nedenle, analog özniteliklerin ikili formda ifade edilebileceği bir ÖSD aşamasına ihtiyaç duyulmaktadır.

ÖSD aşaması, Şekil 27'de gösterildiği gibi iki ana adımdan meydana gelmektedir. Bu adımlardan ilki olan kuantalama adımına göre, analog değerler için bir kuantal seviye sayısı seçilmektedir. Sonrasında bu seviye sayısına göre analog değerlerin kuantal seviye sınırları, seçilen bir kuantalama tekniğine göre elde edilmektedir. Sonrasında, analog değerlerin karşılıkları, kuantal seviyelere göre elde edilmekte ve öznitelik vektörlerinde kuantalama aşaması sonuçlandırılmaktadır.

ÖSD aşamasının ikinci adımı olan kodlama aşamasında ise kuantalanmış verilerin ikili forma dönüştürülmesi sağlanmaktadır. Buna göre seçilen bir kodlama yönteminde bir önceki adımda seçilen kuantal seviye sayısına göre karşılıkları elde edilmektedir. Kuantalanmış değerlerin seviye karşılıklarına seçilen kodlama yönteminin seviye karşılıkları atanmaktadır. Bu aşamada, analog öznitelik vektörleri, ikili formda spike



Şekil 27. Tasarlanan gerçekçi sinir ağı modelinin akış diyagramı

örüntüleri olarak ifade edilmiş olmaktadır. Son olarak, spike örüntüleri düzleştirilmekte ve GSA için uygun forma dönüştürülmüş olmaktadır. ÖSD aşamasında seçilen ve elde edilen tüm özellikler daha sonra test özneliklerinin dönüşümünde kullanılmak üzere kaydedilmektedir.

GSA'lar, spike örüntülerini işlemek üzere uzmanlaşmış yapılardır. Spike örüntülerinin işlenmesi adına uygun GSA modelinin tasarlanması, bu nedenle, önemli bir aşamadır. Buna göre, Şekil 27'de gösterildiği gibi bir GN seçilmektedir. GN'ler başlangıçta sinaptik reseptörlere sahip değildirler. Bu nedenle GN yüzeyine spike örüntülerini işleyebilmeleri için uygun sayı ve özellikte sinaptik reseptörler yerleştirilmektedir. Sonrasında, sinaptik reseptörler ile spike örüntülerinde bulunan her bir eleman birbirleriyle eşleştirilmektedir. Bu sayede bir GN için spike örüntüsü ve sinaptik reseptörlerin birbirilerine bağlantıları tamamlanmış olmaktadır. Bu aşamada henüz GSA modeli tamamlanmamış olmakta ve elde sadece bir GN modeli bulunmaktadır. GSA modelinin oluşturulması adına GN modeli sınıflandırma probleminin sınıf sayısı miktarınca kopyalanmakta ve paralel olarak konumlandırılmaktadır. Sonrasında her bir sınıf, bir GN modeliyle eşleştirilmektedir. Buna göre eğer bir sınıfa ait spike örüntüsü girdi

olarak verildiyse, o sınıfı tespit etmekten sorumlu GN spike üretmeli, diğer GN'ler suskun kalmalıdır. Bu aşamadan itibaren tek katmanlı bir GSA modeli tasarlanmış olmaktadır.

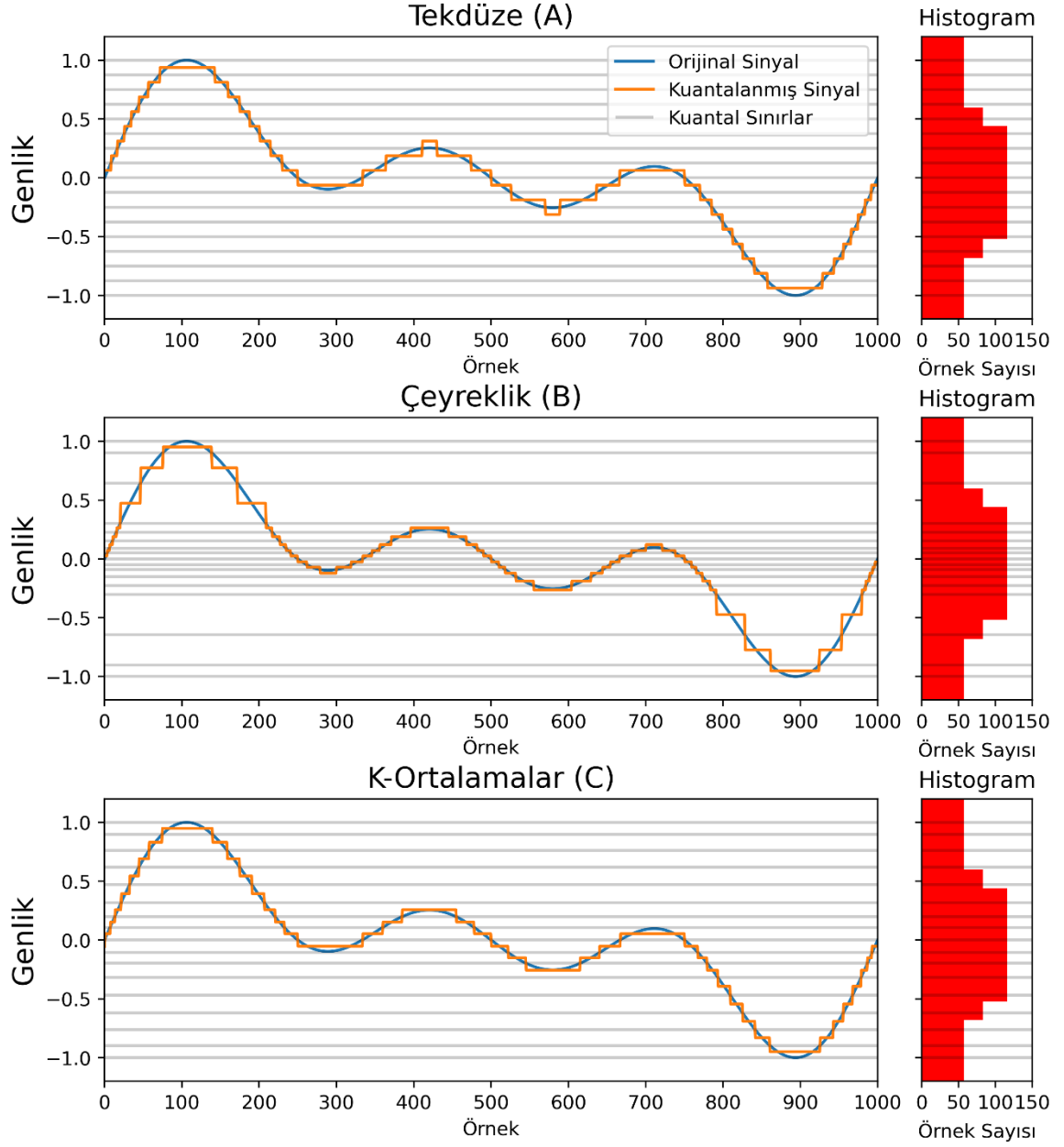
GSA modelinin tasarlanmasından sonraki aşama eğitimidir. Eğitim aşamasında spike örüntüleri teker teker GN'lere girdi olarak verilmektedir. Her bir spike örüntüsü için GN'lerin çıktıları kaydedilmektedir. Kaydedilen çıktılara göre perceptron öğrenme yöntemi kullanılarak SB'lerin ağırlıkları güncellenmekte ve bu işlem GSA'nın hatası minimum olana kadar veya belirlenen bir epoch sayısının sonuna kadar sürdürülmektedir. Eğitim sonunda GSA test edilmeye hazır duruma gelmektedir.

Test aşamasına gelindiğinde, bir test örneği seçilmektedir. Bu örneğe ait öznitelik vektörü, eğitim özniteliklerinin spike örüntülerine dönüştürülmesi süresince seçilen ve belirlenen özelliklere bağlı kalınarak spike örüntüsüne dönüştürülmektedir. Bu spike örüntüsü GSA modeline girdi olarak verilmekte ve GN'lerinin çıktıları kaydedilmektedir. Bu çıktıların özelliklerine göre test örneğinin hangi sınıfa ait olduğu belirlenmektedir.

Bu bölümde tasarlanan GSA modeline dair detaylar verilmektedir. Bölüm, yedi alt başlıktan meydana gelmektedir. İlk alt başlıkta, ÖSD aşaması tanıtılmıştır. İkinci alt başlıkta, aynalama yaklaşımı detaylandırılmıştır. Üçüncü alt başlıkta, GSA modelinin tasarlanması, sinaptik reseptörlerin yerleştirilmesi ve SB'lerin oluşturulması süreci detaylı olarak açıklanmıştır. Dördüncü alt başlıkta, eğitim aşamasında kullanılan öğrenme yöntemi ve detayları anlatılmıştır. Beşinci alt başlıkta, GSA'nın eğitim aşaması; altıncı alt başlıkta ise test aşamasına dair detaylar verilmiştir. Yedinci ve son alt başlıkta, GSA modelinin oluşturulduğu platform ve geliştirilen yazılıma dair detaylar sunulmuştur.

2.4.1. Öznitelik-Spike Dönüşümü

Bu çalışmada, analog değerlerin uzamsal olarak serpiştirilmiş SB'lerde tek spike ile ifade edildiği popülasyon tabanlı bir kodlama stratejisi izlenmiştir. Bu stratejinin ilk adımı kuantalamadır. Kuantalama adımına başlamadan önce bir bit uzunluğu seçilmelidir. Bit uzunluğunun n olduğu varsayıldığında kuantal seviye sayısı 2^n olarak elde edilmektedir. Kuantalama, analog değerlerin ayrık seviyeler olarak ifade edilmesidir. Buna göre analog değerler alabilen bir sayısı dizisinin elemanları, değerleri de gözetilerek önceden seçilmiş seviyeler cinsinden ifade edilmektedir. Kuantal seviyeler olarak isimlendirilen bu seviyeler için farklı yaklaşımlar mevcuttur.



Şekil 28. Örnek bir sayı dizisiyle kuantalama teknikleri

Bu çalışmada, üç farklı yaklaşımla kuantalama işlemi yapılmaktadır. Şekil 28, bu yaklaşımları 1000 örnekten oluşan bir sayı dizisi üzerinden göstermektedir. Şeklin sol tarafında sayı dizisinin değerleri gösterilirken, sağ tarafta bu dizinin histogram dağılımı gösterilmiştir. Şekil 28, ayrıca, yöntemlerin uygulanması sonucu elde edilen kuantal sınırları ve kuantalanmış sayı dizisini de göstermektedir.

İlk yaklaşım, kuantal seviyelerin tekdüze (uniform) olarak eşit aralıklarla bölünmesidir. Bu yaklaşıma göre sayı dizisinin maksimum ve minimum değerleri

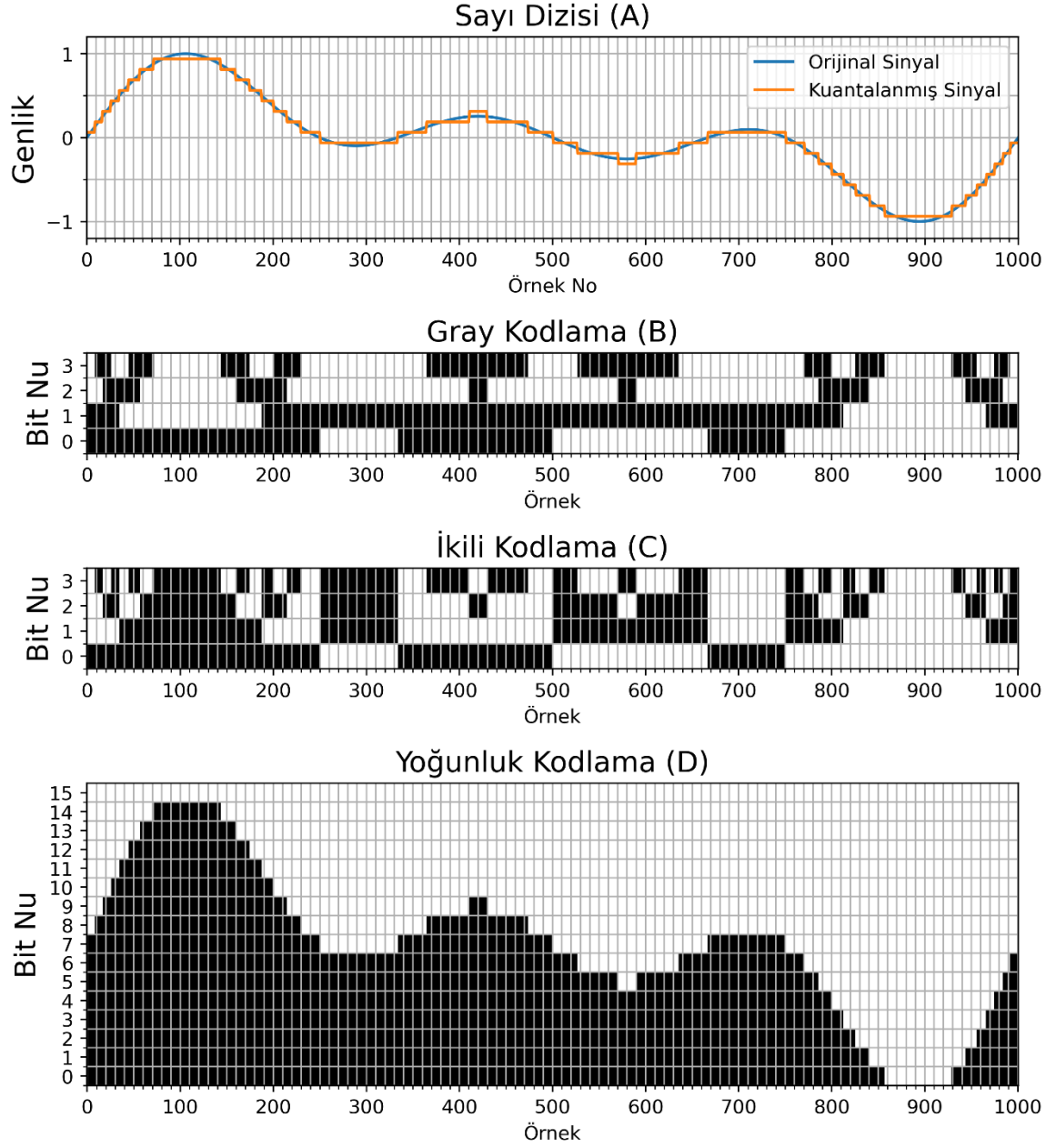
Tablo 7. Kodlama yöntemleri ve 16 seviye için kod karşılıkları

Seviye	İkili Kodlama	Gray Kodlama	Yoğunluk Kodlama
0	0000	0000	0000000000000000
1	0001	0001	0000000000000001
2	0010	0011	0000000000000011
3	0011	0010	0000000000000111
4	0100	0110	0000000000001111
5	0101	0111	0000000000011111
6	0110	0101	0000000000111111
7	0111	0100	0000000001111111
8	1000	1100	0000000011111111
9	1001	1101	0000000111111111
10	1010	1111	0000001111111111
11	1011	1110	0000011111111111
12	1100	1010	0000111111111111
13	1101	1011	0001111111111111
14	1110	1001	0011111111111111
15	1111	1000	0111111111111111

belirlenmekte ve aralarındaki değer mesafesi eşit aralıklı olacak şekilde bölünmektedir. Şekil 28(A) örnek sayı dizisinin orijinalini, tekdüze kuantal aralıkları ve kuantalanmış sayı dizisini göstermektedir.

Bir diğer yaklaşım ise çeyreklik (quantile) yaklaşımıdır. Bu yaklaşıma göre sayı dizisinin histogram dağılımı ele alınmaktadır. Sayısı dizisinin histogramında, her bir aralığa eşit sayıda örnek düşecek şekilde kuantal sınırlar belirlenmektedir. Bu sayede daha çok değere sahip aralıklar daha detaylı ifade edilmektedir. Sonrasında, belirlenen bu kuantal sınırlara göre kuantalama işlemi yapılmaktadır. Şekil 28(B) örnek sayı dizisinin orijinalini, çeyreklik kuantal aralıkları ve kuantalanmış sayı dizisini göstermektedir.

Son yaklaşım ise kuantal aralıkların K-Ortalamlar (K-Means) gibi kümeleme yöntemi kullanılarak belirlenmesidir. Bu yaklaşıma göre sayı dizine, seçilen kuantal seviye sayısı kadar öbek oluşturmak kaydıyla kümeleme işlemi yapılmaktadır. Bu sayede kuantal aralıklar belirlenmekte ve kuantalama işlemi yapılmaktadır. Şekil 28(C) örnek sayı



Şekil 29. Tekdüze kuantalanmış bir sayı dizisinin ikili formda kod karşılıkları

dizisinin orijinalini, K-Ortalamlar kümeleme yöntemiyle elde edilmiş kuantal aralıkları ve kuantalanmış sayı dizisini göstermektedir.

Bu çalışmada, kuantalama işlemi öznitelik vektörlerine uygulanmıştır. Toplam N adet örnek ve M adet öznitelik sayısına sahip $N \times M$ boyutunda bir öznitelik vektörünün olduğu varsayıldığında, her öznitelik cinsi kendi aralarında kuantalanacak şekilde ayrılmaktadır. Toplam N adet örneğin i 'inci öznitelikleri kendi içinde bir sayı dizisi olarak kabul edilmektedir ve kuantal sınırları belirlenmektedir. Bu kuantal sınırlara göre i 'inci

öznitelik dizisi kuantalanmaktadır. Bu işlem, M sayıdaki tüm öznitelik cinsleri için ayrı ayrı yapılmaktadır.

ÖSD'nin ikinci adımı kodlamadır. Bu adımda kuantalanmış örnekler ikili formda ifade edilebilecekleri kodlara dönüştürülmektedir. Bu çalışmada üç farklı yaklaşımla kodlama işlemi yapılmaktadır.

Çalışmada kullanılan ilk yaklaşım ikili (binary) kodlamadır (İK'dir). Bu yaklaşıma göre seçilen bit uzunluğunda, yani n uzunluğunda, kodlar üretilmektedir. Buna göre her bir kuantal seviye ikili tabandaki karşılıkları ile ifade edilmektedir.

Diğer yaklaşım ise Gray kodlamadır (GK'dir). Bu yaklaşıma göre de bit uzunluğunda kodlar üretilmektedir. Gray kodlamanın seviyelerin ikili formda ifade edilmesi sırasında, seviyelerin değerlerini de gözetmesi nedeniyle önemli bir kodlama yöntemidir. Bu özelliğinden yararlanılmak amacıyla çalışmaya dahil edilmiştir.

Son yaklaşım ise yoğunluk (intensity) kodlamadır (YK'dir). Bu yaklaşıma göre kuantal seviye sayısı uzunluğunda, yani 2^n uzunluğunda, kodlar üretilmektedir. Bu yaklaşıma göre, toplam 2^n seviye ile kuantalanmış bir sayı dizisinin bir elemanı, 2^n sayıda ikili kodla ifade edilmektedir. Eğer bir eleman i 'inci seviyede ise, o elemanın yoğunluk kodlamadaki, ilk i sayıdaki değerler 1, geri kalanların ise 0 olarak ifade edilmesiyle bulunmaktadır. Tablo 7, tanıtılan üç farklı kodlama yaklaşımında, bit uzunluğunun 4 olduğu durumdaki kuantal seviyelerin karşılıklarını göstermektedir.

Şekil 28'de verilen sayı dizisinin Şekil 28(A)'da gösterildiği gibi tekdüze yaklaşımıyla kuantalanmış olduğu varsayıldığında; sayı dizisinin tanıtılan yaklaşımlarla kodlandığı durumdaki karşılıkları Şekil 29'da gösterilmiştir. Görüldüğü üzere, her yaklaşım kendine özgü kodlar üretmektedir.

İkili kodların elde edilmesinden sonra, iki boyutlu olan spike örüntüleri düzleştirilerek tek boyutta ifade edilmektedir. Tek boyuta düzleştirilmiş spike örüntüleri için, kod karşılığı 1 olan elemanlarda spike girişi olduğu, 0 olan elemanlarda ise spike girişinin olmadığı varsayılmaktadır. Bu aşamada $N \times M$ boyutlarındaki bir öznitelik vektörü kodlama işleminin sonunda, kodlama yöntemine bağlı olarak $N \times K$ boyutlarında spike örüntüleri olarak elde edilmektedir. Burada K , ikili veya Gray kodlama yapıldıysa $K = nM$, yoğunluk kodlama yapıldıysa $K = 2^n M$ olarak bulunabilmektedir.

2.4.2. Aynalama Yaklaşımı

Geleneksel YN yapılarında, çıktılar girdilerin ağırlıklı toplamlarından elde edilmektedir ve ağırlıkların pozitif veya negatif olması genellikle bir sorun teşkil etmemektedir. Ancak, bazı çalışmalara göre eğitim süresince ağırlıkların pozitif değer alması gerektiği durumlar ortaya çıkabilmektedir. Moldwin ve Segev (2020) gibi araştırmacılar, çalışmalarında ağırlıkların pozitif olmasının gerektiği durumlarla karşılaştıklarını belirtmişlerdir.

Ağırlık kısıtlamalarının etkisi daha önce Amit ve arkadaşları (1989) tarafından da ele alınmıştır. Chapeton ve arkadaşları (2012) tarafından yapılan bir çalışmada, bu kısıtlamanın sınıflandırma kapasitesini önemli ölçüde azalttığı gözlemlenmiştir. Bu sorunu aşmak için Kayıkcıoğlu Bozkır ve arkadaşları (2023), aynalama yönteminin etkinliğini araştırmışlardır. Bu yaklaşımda, ikili x_i girdileri için eşlenik olan $x_i^* = 1 - x_i$ girdileri de nörona ek girdiler olarak sunulmaktadır. Bu eklenen girdiler göz önünde bulundurulduğunda Denklem (1), Denklem (58) olarak güncellenebilmektedir.

$$y = w_0 + \sum_{i=1}^N w_i x_i + \sum_{i=1}^N w_i^* x_i^* \quad (58)$$

Bu denklemde; x_i ve x_i^* 'inci girdinin orijinal ve eşlenik değerlerini temsil ederken, w_i ve w_i^* bu girdilere ait ağırlık değerlerini ifade etmektedir. Özellikle, w_i ve w_i^* değerlerinin pozitif olma kısıtlaması dikkate alınmalıdır. Denklem (58) düzenlendiğinde, Denklem (59) ve (60)'ta belirtilen adımlar izlenerek Denklem (61) elde edilmektedir.

$$y = w_0 + \sum_{i=1}^N w_i x_i + \sum_{i=1}^N w_i^* (1 - x_i) \quad (59)$$

$$y = w_0 + \sum_{i=1}^N w_i x_i - \sum_{i=1}^N w_i^* x_i + \sum_{i=1}^N w_i^* \quad (60)$$

$$y = \sum_{i=1}^N (w_i - w_i^*) x_i + \sum_{i=1}^N w_i^* + w_0 \quad (61)$$

Burada Denklem (61)'de verilen $(w_i - w_i^*)$ ifadesi $w_i^{Geniş}$ ve $\sum_{i=1}^N w_i^* + w_0$ ifadesi $w_0^{Geniş}$ olarak yeniden isimlendirilirse, Denklem (62) elde edilmektedir.

$$y = \sum_{i=1}^N w_i^{Geniş} x_i + w_0^{Geniş} \quad (62)$$

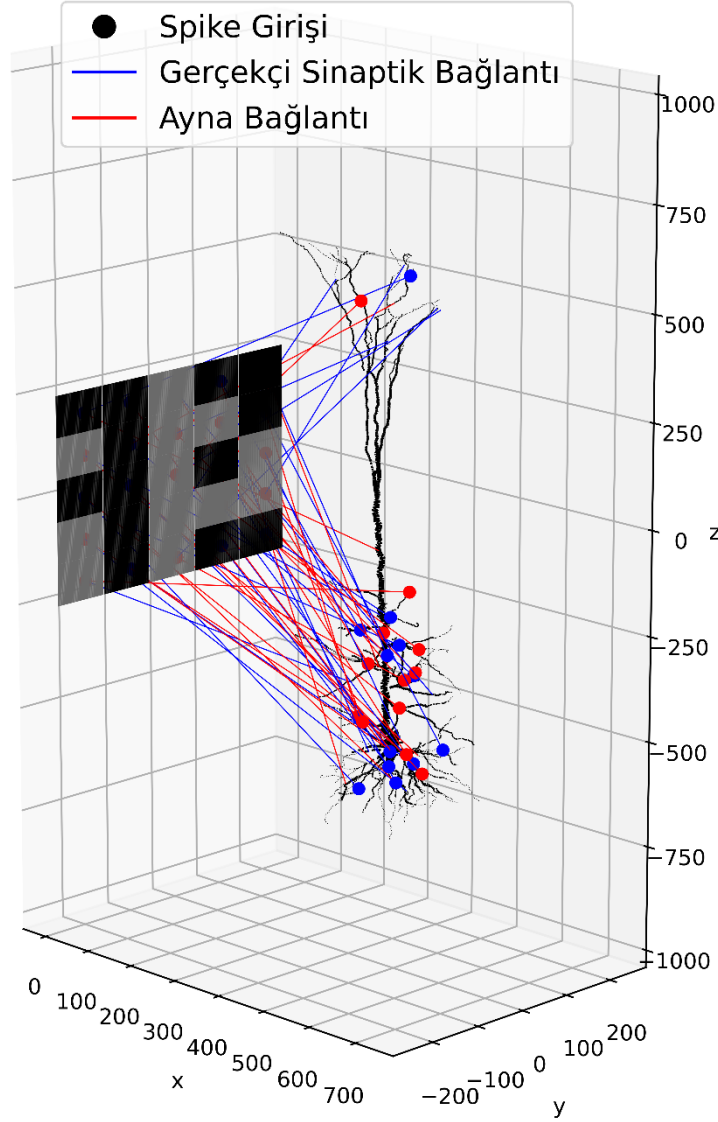
Bu denkleme göre w_i ve w_i^* değerleri pozitif olma kısıtına sahip olsa da $w_i^{Geniş}$ değeri bu kısıttan muaftır. Denklem (1) için ağırlıkların pozitif olma kısıtı varken, Denklem (62) aynı forma sahip olmasına rağmen bu kısıttan aynalama yaklaşımı sayesinde kurtarılmıştır.

2.4.3. Ağ Kurulumu ve Sinaptik Bağlantılar

ÖSD'nin ardından, GSA modelinin tasarlanması aşamasına geçilmektedir. Ancak tasarım aşamasından önce GSA'ya dair bazı özelliklerin seçilmesi gerekmektedir. Bu özelliklerin ilki, aynalama yaklaşımıdır. Bu yaklaşıma göre, SB ağırlıklarının pozitif değer alma zorunluluğundan dolayı ortaya çıkan kısıtların aşılması adına, ikili değer alan öznitelik değerlerinin eşleniklerin de ayrıca GN'ye giriş olarak verilmektedir (Kayıkcıoğlu Bozkır vd., 2023). Aynalama yaklaşımının varlığı GSA tasarımını doğrudan etkilemektedir ve en başından belirlenmelidir.

İkinci özellik ise spike örüntülerindeki her bir girdinin GN yüzeyinde kaç farklı noktaya bağlanacağıdır. Literatürdeki çalışmalar göstermiştir ki BN'ler birbirlerine birden fazla SB aracılığıyla bağlanabilmektedirler (Sporns, 2007). Bu sayede, SB'lerin konumu itibarıyla oluşabilecek kısıtlar ortadan kalkmaktadır. Örneğin, bir BN aynı anda hem GN'nin apikal bölgesine hem basilar bölgesine ve hatta apikal bölgesinde başka bir konuma bağlanabilmektedir (Spruston, 2008). Bu aşamada seçilmesi gereken parametre, spike örüntülerindeki her bir elemanın GN yüzeyinde kaç farklı noktaya bağlanacağını sayıdır. SB tekrarı olarak da ifade edilebilecek bu seçenek, GSA tasarımından önce belirlenmelidir.

Bölüm 2.4.1'de anlatıldığı üzere $N \times M$ boyutundaki öznitelik vektörleri ÖSD sonrası $N \times K$ boyutundaki spike örüntüleri olarak ifade edilmektedir. GSA tasarımından önce ayna bağlantı yapılacağı varsayıldığında spike örüntülerinin eşlenik versiyonları elde edilmekte ve sanki öznitelik vektörünün devamıymış gibi yanına eklenmektedir. Ardından, seçilen bir SB tekrarı miktarınca çoğaltılmakta ve bu kopyalar da spike örüntüleri ile birleştirilmektedir. Bu durumda $N \times K$ boyutundaki spike örüntüleri; $N \times S$ boyutunda, genişletilmiş spike örüntüleri olarak ifade edilmektedir. Burada; X , SB tekrarı sayısı olmak üzere $S = 2KX$ olduğu göz önünde bulundurulmalıdır.



Şekil 30. Piramidal nöron için sinaptik bağlantıların yerleşimi

Bu aşamada, GSA, $N \times S$ boyutlu spike örüntüsü için tasarlanabilir durumdadır. Öncelikle bir GN modeli seçilmektedir. Başlangıçta, GN modeli sinaptik reseptörlere sahip değildir. Sinaptik reseptörlerin GN yüzeyine yerleştirilmesi süreci Bölüm 2.2.3'te anlatıldığı gibi yapılmaktadır. Bu aşamada, S uzunluğundaki her bir spike örüntüsü için çok az sayıda sinaptik reseptör bulunduğu göz önünde bulundurulmalıdır. Girdi ve reseptör eşleştirme işlemi, gerektiğinde bir reseptöre birden fazla girdi denk düşecek şekilde yapılmaktadır. Sonuçta, bir GN modeli için sinaptik reseptörler ve girdilerle eşleştirilmesi tamamlanmış ve sonuç olarak SB'ler belirlenmiş olmaktadır. GSA modelinin tasarımı için GN'ler sınıf sayısı miktarınca kopyalanmakta ve yan yana konumlandırılmaktadır.

Böylece, GSA tasarımı makro boyutta tamamlanmış olsa da mikro boyutta hala ayarlanması gereken bir özellik daha bulunmaktadır. Bu da sinaptik reseptörlerde bulunan SB modellerinin seçimidir.

Bölüm 2.2.3'te çalışmada kullanılan SB modellerinin yapısı tanıtılmıştır. Bu yapıya göre her bir sinaptik reseptör ikisi Exc, ikisi Inh yapıda olmak üzere dört farklı SB'den meydana gelmektedir. Her bir SB kendine ait bir ağırlık parametresiyle ifade edilmektedir. Bağlantı yapıldığı anda tüm SB'lerin ağırlıkları sıfır olarak ayarlanmaktadır. Bu çalışmanın bir diğer hedefi ise SB tiplerinin etkisini incelemektir. Bu aşamada sinaptik reseptörlerde bulunan SB'lerin hangilerinin dahil edileceğinin kararının verilmesi gerekmektedir. Seçilen SB modelleri sinaptik reseptör işlemlerine dahil edilirken, dışta tutulan SB modellerinin ağırlıkları sıfır olacak şekilde dondurulmakta ve eğitim süresince güncellenmeleri engellenmektedir. Son olarak, seçilen GN modeli eğer basit nöron ise merkezine, PN ise somasına sabit bir akım kaynağı eklenmektedir. Böylece beyin içinde bulunan BN'lerin diğer BN'lerden aldığı ortalama sinaptik akımlar modellenmiş olmaktadır. Sonuç olarak, GSA modelinin tasarımı tamamlanmış olmaktadır.

2.4.4. Öğrenme Yöntemi

Sınıflandırmada kullanılan bir GN için, $X = [x_0, x_1, x_2, x_3, x_3, \dots, x_N]$ sınıflandırılması istenen bir spike örüntüsünü, $W = [w_0, w_1, w_2, w_3, w_3, \dots, w_N]$ ise bu spike örüntüsündeki her bir girdinin SB ağırlığını ifade etmektedir. Ayrıca bir GN, çıktısında en az bir spike üretiyorsa 1, çıktısında hiç spike üretmiyorsa 0 olarak etiketlenmektedir. Bu yaklaşıma göre eğer bir örnek GN'nin tespit etmekten sorumlu olduğu sınıfa ait ise, GN'nin o örneğe ait spike örüntüsü için en az bir spike üretmesi beklenmektedir. Eğer örnek hedef sınıfa ait değilse de GN'nin suskun kalması beklenmektedir. GN eğer spike üretmesi gereken bir spike örüntüsü için spike üretmiyorsa veya spike üretmemesi gereken bir spike örüntüsü için spike üretiyorsa, GN'nin SB ağırlıklarının düzeltilmesi gerektiği anlamına gelmektedir.

SB ağırlıklarının güncellenmesi adına bu çalışmada öğrenme yöntemi olarak, Bölüm 1.2.2'de anlatılan perceptron kuralı kullanılmıştır. Perceptron kuralı için ağırlık değişimi Denklem (6) ve (8)'de verilmiştir.

$$w_i^{Yeni} \leftarrow w_i^{Eski} + \Delta w_i \quad (6)$$

$$\Delta w_i = \eta \times (d - \text{çıktı}) \times x_i \quad (8)$$

Bu problem özelinde *çıktı*, GN'nin çıktısını; *d*, beklenen çıktı değerini; x_i *i*'inci girdi değerini ve Δw_i ise *i*'inci girdinin SB ağırlık değişimini göstermektedir. Bu öğrenme kuralına göre bir GN'nin SB ağırlıkları, ürettiği ve beklenen çıktı arasındaki fark üzerinden Denklem (6)'da gösterildiği gibi güncellenebilmektedir. Yapay nöronlar (YN'ler) ağırlık konusunda herhangi bir kısıta sahip olmadıklarından, bu öğrenme yöntemi YN'ler için Rosenblatt (1958) tarafından önerilen hali ile kullanılabilir.

Moldwin ve Segev (2020) çalışmalarında GN'lerin SB ağırlıklarının pozitif olma zorunluluğundan bahsetmişlerdir. GN'lerde SB ağırlıkları eğer negatif olursa, Denklem (37)'de ifade edilen sinaptik iletkenliğin negatif değer alması anlamına gelmektedir. Bu nedenle Moldwin ve Segev (2020) çalışmalarında perceptron öğrenme kuralını Denklem (63)'te ifade edildiği gibi güncellemişlerdir.

$$w_i^{Yeni} \leftarrow \text{maksimum} \left(0, (w_i^{Eski} + \Delta w_i) \right) \quad (63)$$

Burada amaç, eğitim süresince SB ağırlık değerlerinin negatif olması durumunda Denklem (6)'da verilen w_i değerinin kısıtlanmasıdır. Bu sayede, SB ağırlığı eğer negatif değer alması gereken bir durumda ise sifıra eşitlenmektedir.

Moldwin ve Segev'in (2020) yaklaşımları her ne kadar ağırlık kısıtlama işlemini yapabilse de sadece Exc SB'ler için geçerli olduğundan bu çalışma için yetersiz kalmaktadır. Bu çalışmada ayrıca Inh SB'ler kullanıldığından Denklem (63)'te verilen ifade, SB'lerin tipleri de göz önünde bulundurularak Denklem (64) olarak güncellenmiştir.

$$w_i^{Yeni} \leftarrow \begin{cases} \text{maksimum} \left(0, (w_i^{Eski} + \Delta w_i) \right), & \text{Exc Sinapslar için} \\ \text{maksimum} \left(0, (w_i^{Eski} - \Delta w_i) \right), & \text{Inh Sinapslar için} \end{cases} \quad (64)$$

Denklem (64) sadece Exc ve Inh SB'ler için geçerlidir. Bir diğer girdi olan sabit $I_{giriş}$ ifadesi için ağırlık güncellemesi Denklem (8)'de gösterildiği gibi yapılmaktadır.

2.4.5. Eğitim

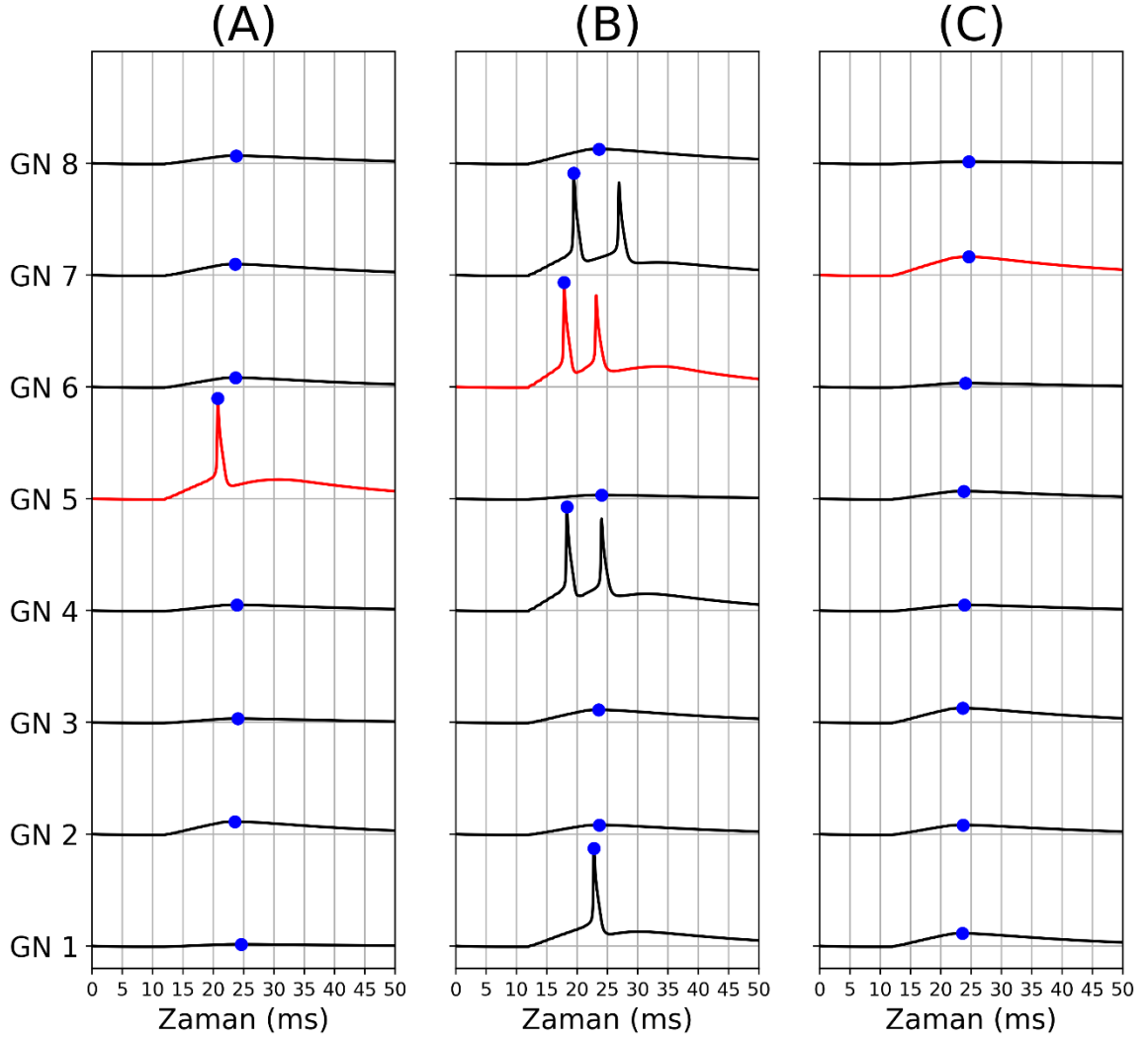
Bu bölümde, GSA modelinin eğitim sürecinden bahsedilmiştir. Öncelikle, eğitim örneklerinden ilki seçilmekte ve spike örüntüsü elde edilmektedir. Bu spike örüntüsü her bir girdisi belirli SB'lerle eşleştirilmiş durumdadır. Girdilerin eşleştirilmesinin ardından GSA, 50 ms boyunca simüle edilmektedir. Simülasyonun 10'uncu ms anında, spike örüntüsünde girdi değeri 1 olan girdiler için bağlı oldukları SB'lerde bir spike girişi uygulanmaktadır. Simülasyon sonunda GSA içinde bulunan GN'lerin her biri için zar potansiyelleri kaydedilmektedir. Zar potansiyelleri incelendiğinde en az bir spike üreten GN'lerin çıktıları 1, spike üretmeyen GN'lerin çıktıları ise 0 olarak etiketlenmektedir. Bu etiketler seçilen örnek için beklenen sınıf etiketi ile karşılaştırılmakta ve uyumsuzluk söz konusu ise 2.4.4'te anlatılan perceptron kuralı yardımıyla SB ağırlıkları güncellenmektedir.

Bir eğitim örneği için işlem tamamlandığında sıradakine geçilmekte ve bu işlem tüm örnekler eğitimde kullanılabilecek kadar devam etmektedir. Tüm örnekler bir kez eğitimde kullanıldığında, bu bir epok olarak isimlendirilmektedir. Eğer GSA hala istenen çıktıları vermiyorsa bu işlem bir epok daha tekrar edilmektedir. Bu süreç GSA istenen çıktıları verene kadar veya sınırlanan bir epok sayısının sonuna kadar tekrar edilmektedir. Eğitim sonunda ağırlık güncelleme özelliği dondurulmakta ve test aşamasına geçilmektedir.

2.4.6. Test

Test aşaması iki adımdan meydana gelmektedir. İlk adımda GSA bir test örneği için simüle edilmektedir, ikinci bölümde ise GSA'nın içinde bulunan GN çıktıları değerlendirilerek test örneğinin sınıf etiketi belirlenmektedir. Bu bölümde, bu işlemlerin detaylarına yer verilmiştir.

Test aşaması eğitime benzer olarak bir test örneğinin spike örüntüsüne dönüştürülmesiyle başlamaktadır. Bu aşamada, eğitimde seçilen ve belirlenen özellikler göz önünde bulundurularak spike örüntüleri elde edilmektedir. Bir test spike örüntüsünün her bir girdisi eğitimde belirlenen SB'ler ile eşleştirilmektedir. GSA, eğitim aşamasındaki gibi 50 ms boyunca simüle edilmekte ve 10'uncu ms anında test spike örüntüsü içinde 1 değerine sahip girdilerden için ilgili SB'lere spike girişi uygulanmaktadır. Simülasyon sonunda tüm GN'lerin zar potansiyelleri kaydedilmekte ve sınıf tespiti için bir sonraki adıma aktarılmaktadır.



Şekil 31. Tekil örüntü tipi veriler için gerçekçi sinir ağı çıktıları

Test sonunda, bir test örneğinin sınıf etiketinin belirlenmesinde iki farklı yöntem bulunmaktadır. Bu yöntem sınıflandırılan örneğin tipine göre belirlenmektedir. Eğer bir test örneği sadece tek bir spike örüntüsünden meydana geliyorsa sınıf belirleme adımında tekil örüntü sınıf belirleme yöntemi, eğer bir test verisi birbirlerine bağlı birden fazla spike örüntülerinden meydana geliyorsa, sınıf belirleme adımında çoğul örüntü sınıf belirleme yöntemi kullanılmaktadır.

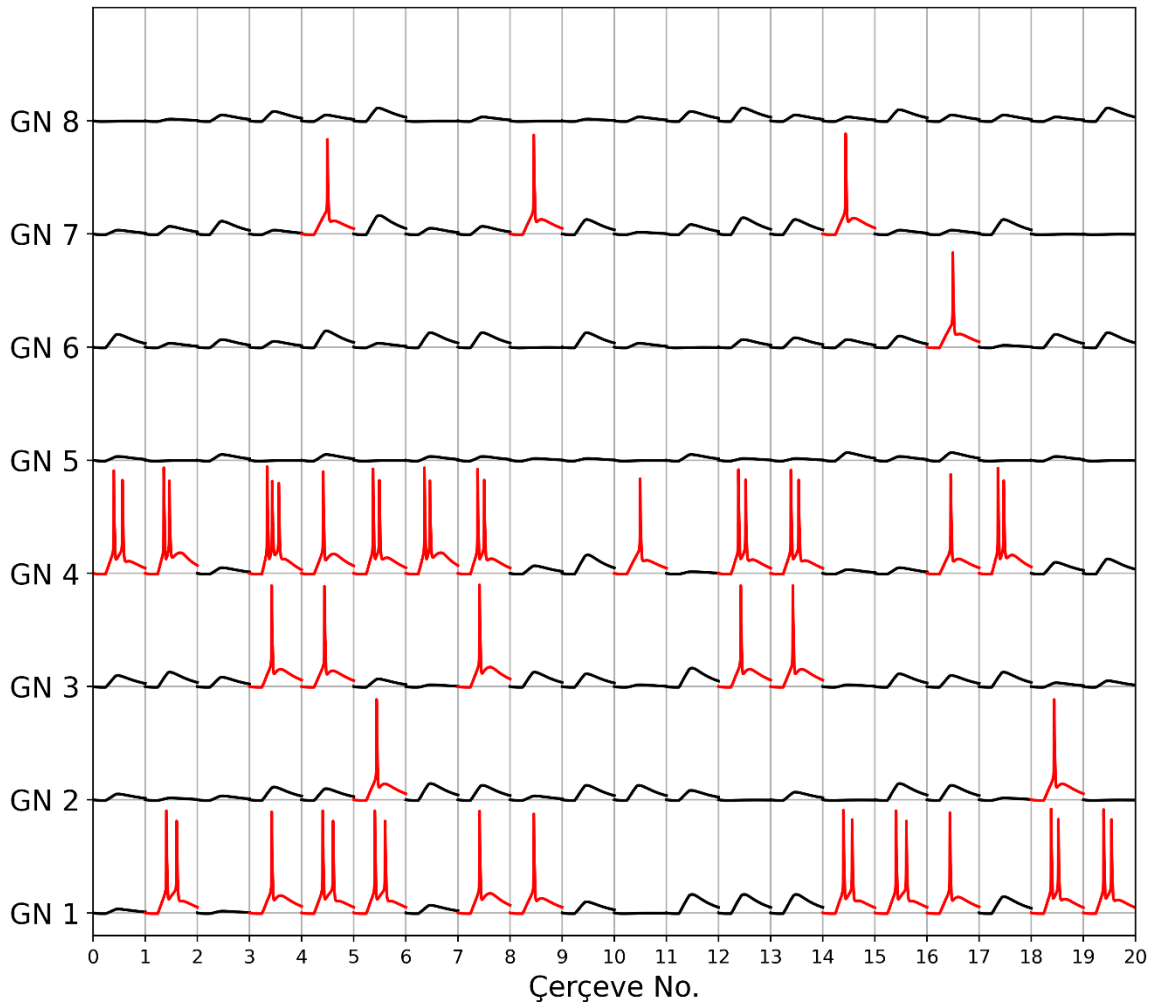
Tekil örüntü sınıf belirleme yöntemi bu çalışmada; yapay VK, IRIS ve MNIST gibi, öznelitliklerin zamandan bağımsız olduğu ve her bir örneğin tek bir spike örüntüsü ile ifade edildiği verilerde kullanılmıştır. Buna göre her bir örnek tek başına değerlendirmeye alınmış ve sınıf etiketi belirlenmiştir.

Simülasyon sonucunda GN'lerin zar potansiyelleri için üç farklı olasılık ortaya çıkabilmektedir ve bu olasılıklar Şekil 31'de gösterilmektedir. Bu olasılıklar ile karşılaşıldığında, aşağıdaki kurallar takip edilerek sınıf etiketleri belirlenmiştir.

- Birinci olasılık Şekil 31(A)'da gösterildiği gibidir. Bu durumda çıkış katmanındaki GN'lerden sadece biri spike üretmektedir. Bu hedeflenen durumdur ve örneğin sınıfı, spike üreten GN'nin sınıfı olarak atanmaktadır. Burada, Şekil 31(A)'da verilen test örneği için sonuçlar, test örneğinin beşinci GN'nin tespit etmekten sorumlu olduğu sınıfa ait olduğunu göstermektedir.
- İkinci olasılık Şekil 31(B)'de gösterildiği gibidir. Bu durumda çıkış katmanında birden fazla GN, spike üretmektedir. Bu istenmeyen bir durumdur fakat yine de örneğin hangi sınıfa ait olduğu belirlenmesi gerekmektedir. Bu olasılıkta, test örneğinin, spike üreten GN'ler arasında, spike sayısına bakılmaksızın, zamana göre ilk spike üreten GN'nin tespit etmekten sorumlu olduğu sınıfa ait olduğu kabul edilmektedir. Burada, Şekil 31(B)'de verilen test örneği için sonuçlar, örneğin altıncı GN'nin tespit etmekten sorumlu olduğu sınıfa ait olduğunu göstermektedir.
- Üçüncü olasılık Şekil 31(C)'de gösterildiği gibidir. Bu durumda çıkış katmanında hiçbir GN spike üretmemektedir. Bu istenmeyen bir başka durumdur fakat yine de örneğin hangi sınıfa ait olduğu belirlenmesi gerekmektedir. Böyle bir çıktı ile karşılaşıldığında zar potansiyellerinin maksimum değerleri (V_{max}) incelenmektedir. Bu olasılıkta, test örneğinin, V_{max} değerleri arasında en yüksek değere sahip GN'nin tespit etmekten sorumlu olduğu sınıfa ait olduğu kabul edilmektedir. Burada, Şekil 31(C)'de verilen test örneği için sonuçlar, test örneğinin yedinci GN'nin tespit etmekten sorumlu olduğu sınıfa ait olduğunu göstermektedir.

Çoğul örüntü sınıf belirleme yöntemi bu çalışmada konuşma kayıtları gibi özniteliklerin zamana bağlı ve her bir örneğin ardışık spike örüntüleriyle ifade edildiği verilerde kullanılmıştır. Konuşma kayıtları çerçeveler olarak ifade edildiği için her kayıt birden fazla ve ardışık spike örüntüleri olarak ifade edilmektedir. Bu nedenle spike örüntüleri ardışık ve bir bütün olarak değerlendirmeye alınmış ve sınıf etiketleri belirlenmiştir.

Test aşamasında her bir çerçeveye ait spike örüntülerinin GN çıktıları yukarıda verildiği gibi bulunmuş ve kaydedilmiştir. GN'lerin çıktı zar potansiyelleri Şekil 32'de gösterildiği üzere çerçeveler gibi ardışık olarak sıralanmıştır. GN çıktıları, spike sayısına bakılmaksızın en az bir spike üretenler 1, spike üretmeyenler 0 olacak şekilde



Şekil 32. Çoğul örüntü tipi veriler için gerçekçi sinir ağı çıktıları

etiketlenmiştir. Şekil 32, bu yüzden, 1 olarak etiketlenen çıktıların kırmızı, 0 olarak etiketlenen çıktıların ise siyah olarak gösterileceği şekilde düzenlenmiştir. Sonuç olarak, sınıf etiketi en fazla 1 çıktısı üreten GN'ye göre belirlenmektedir. Burada, Şekil 32'de verilen test örneği için sonuçlar, test örneğinin dördüncü GN'nin tespit etmekten sorumlu olduğu sınıfa ait olduğunu göstermektedir.

2.4.7. Yazılım

Bölüm 2.4'te tanıtilan GSA modeli ile ilgili tüm simülasyon çalışmaları NEURON platformunda gerçekleştirilmektedir. NEURON, sinirbilim araştırmalarında GN'lerin ve GSA'ların modellenmesinde yaygın bir kullanıma sahiptir (Hines & Carnevale, 1997). Platform, GN modellerinin tasarımı ve bu modellerin fonksiyonel analizini gerçekleştirme

kapasitesi sunmaktadır. Migliore ve arkadaşları (2006), NEURON platformunun paralel ağ simülasyonlarını destekleme yeteneğini vurgulayarak, büyük ölçekli GSA'ların modellenmesindeki etkinliğini göstermektedir. Ayrıca, NEURON platformunun *Python* ile uyumluluğu, GSA modellemeye esnekliğini artırmaktadır (Hines, 2009).

Bu çalışmada, ilgili GSA modelinin kurulabilmesi adına NEURON platformunun kullanıldığı bir Python kütüphanesi geliştirilmiştir. Bu kütüphane, Şekil 27'de detayları verilen GSA modeli ve sınıflandırma akışını otomatik olarak gerçekleştirmektedir. Bir sınıflandırma işlemi yapılmak istendiğinde, kullanıcı giriş olarak öznitelik vektörlerini vermekte, program dahilinde hangi bit uzunluğunda, hangi kuantalama tekniğini kullanarak, hangi kodlama yöntemiyle spike örüntülerini oluşturacağını belirleyebilmektedir. Ayrıca kullanıcı, GSA dahilinde aynalama yaklaşımının olup olmayacağını veya hangi tip SB'lerin kaç tekrar ile kullanılacağını belirleyebilmektedir. Sonuç olarak bir kullanıcı gerekli materyali sağlaması ve gerekli ayarlamaları yapması durumunda ilgili kütüphaneyi rahatlıkla kullanabilmektedir.

Geliştirilen yazılımda ihtiyaç duyulan parametreler kullanıcı tarafından seçilebilir durumdadır. İlk seçenek kullanılacak GN modelidir. Yazılımda varsayılan GN modeli IZH olarak seçilmiş olsa da kullanıcı HH, IZH, STA ve PN olmak üzere dört farklı GN modelini kullanabilmektedir. Kullanıcının PN modelini seçtiği varsayıldığında bu modele özgü bir dizi seçenek sunulmaktadır. Bu seçenekler PN modeli içinde ayrılmış olan bölgelerin hangilerinin sinaptik reseptör yerleştirilmesinde kullanılacağıdır. Varsayılan olarak somatik, apikal ve basilar bölgeleri kullanılmak üzere seçilmiş, aksonal bölge seçilmemiştir. Yazılımda GSA modelinin aynalama yaklaşımını kullanıp kullanmayacağı da seçime tabiidir ve varsayılan olarak aynalama yaklaşımının dahil edildiği kabul edilmiştir.

Yazılımda toplam beş adet SB tipi seçilebilir durumdadır. Varsayılan olarak AMPA, NMDA $GABA_A$ ve $GABA_B$ sinapsları dahil edilmiş, Opsin dahil edilebilir olmak üzere varsayılan olarak dahil edilmemiştir. Yazılımda SB tekrar sayısı bir seçenek olarak sunulmuştur. Kullanıcı tarafından pozitif tam sayı olarak seçilebilmekte ve varsayılan olarak 1 bağlantı yapıldığı kabul edilmiştir. Yazılım ÖSD aşamasında bir dizi yöntemi ve parametreyi seçenek olarak sunmaktadır. Bit uzunluğu pozitif tam sayı olarak seçilebilmekle beraber, varsayılan olarak 5 olduğu kabul edilmiştir. Kuantalama tekniği için yazılımda tekdüze, çeyreklik ve K-Ortalamlar yöntemleri seçenek olarak sunulmuş ve varsayılan olarak tekdüze yöntemi kabul edilmiştir. Kodlama yöntemi olarak kullanıcıya

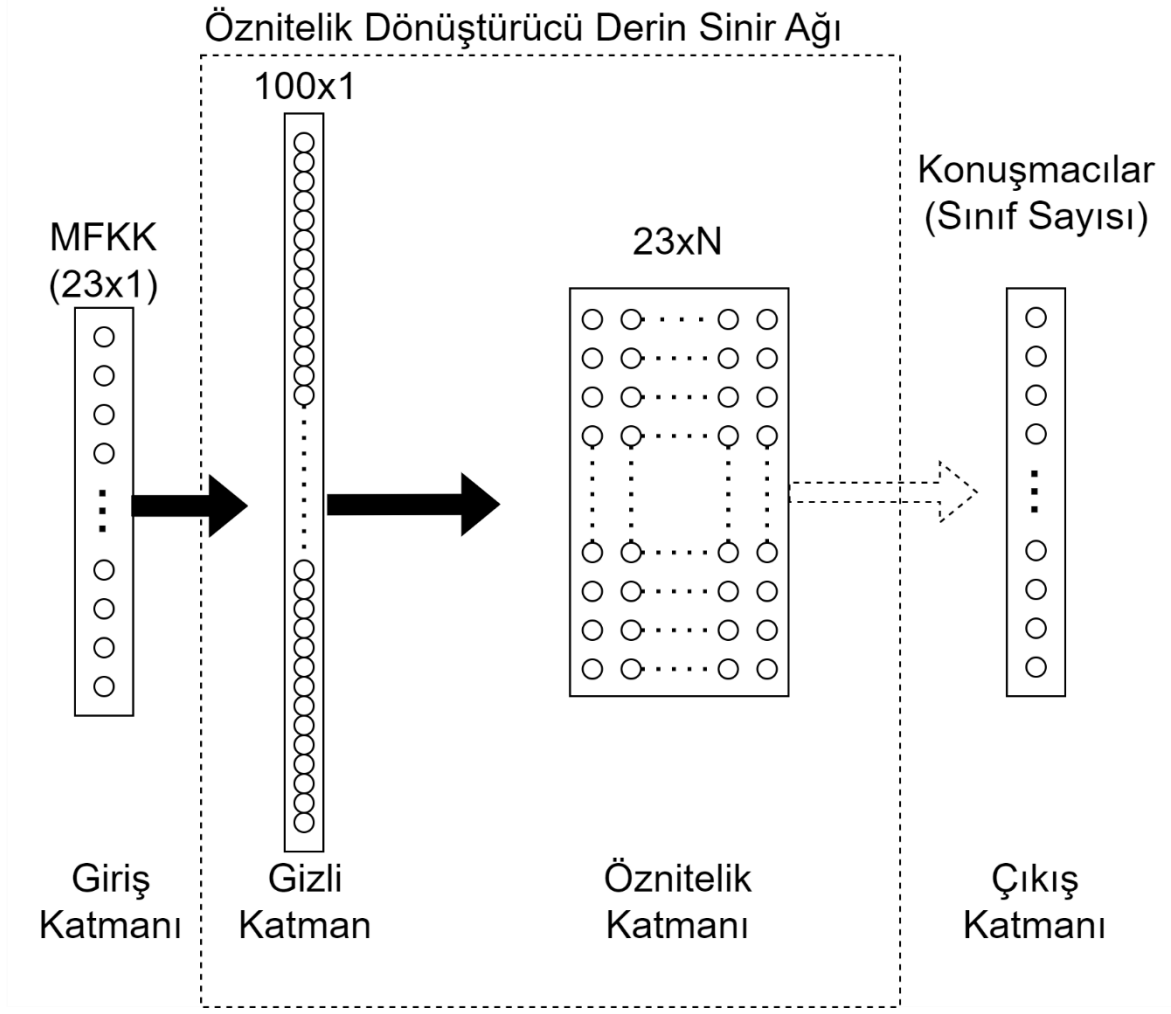
ikili, Gray ve yoğunluk kodlama yöntemleri seçenek olarak sunulmuş ve varsayılan olarak yoğunluk kodlamanın yapılacağı kabul edilmiştir. Kullanıcı ayrıca öğrenme hızı ve GSA modelinin toplam simülasyon süresi gibi çeşitli parametreleri de yazılım üzerinden belirleyebilmektedir.

2.5. Hibrit Yaklaşım için Derin Sinir Ağı Mimarisi

Transfer öğrenme (TÖ), yeni modellerin eğitiminde önceden eğitilmiş modellerin kullanıldığı bir derin öğrenme (DÖ) tekniğidir (LeCun vd., 2015). Bu önceden eğitilmiş model, genellikle, büyük bir VK üzerinde eğitilmektedir ve verilerden güçlü öznitelikler çıkarmayı öğrendiği kabul edilmektedir. Eğitim sürecini sıfırdan başlatmak yerine, TÖ, önceden eğitilmiş modelde öğrenilen özniteliklerden yararlanılmasına ve ilgili bir görevde veya VK'de kullanılmasına olanak tanımaktadır. Transfer öğrenme kullanılarak, önceden eğitilmiş model tarafından öğrenilen, optimize edilmiş ve güçlü özniteliklerden yararlanılabilmektedir. Bunun yanında, TÖ sayesinde zamandan tasarruf edilebilmekte ve yeni modellerin eğitiminde performans artırılabilir (Weiss vd., 2016).

DÖ algoritmalarının sınıflandırma uygulamalarında ağırlığını artırmasıyla birlikte OKT uygulamalarında TÖ tabanlı yaklaşımlar da ortaya çıkmaktadır. Şekil 20, bu yaklaşımların genel bir akışını sunmaktadır. Bu yaklaşımların ortak noktaları, elde edilmiş öznitelik vektörlerinin sınıflar arası ayrımı artırabilecek yeni vektörlere dönüştürülmesidir. Variani ve arkadaşları (2014), filtre bankalarının enerjilerini d-vektörlere; Chen ve arkadaşları (2015), MFKK özniteliklerini j-vektörlere; Snyder ve arkadaşları (2018), MFKK özniteliklerini x-vektörlere TÖ yaklaşımıyla dönüştürmüşlerdir. Literatürdeki bu çalışmalar TÖ yaklaşımının MFKK tabanlı OKT uygulamalarının performansını artırmada etkili bir araç olduğunu göstermektedir.

Bu çalışmada MFKK öznitelikleri, TÖ uygulamalarından esinlenerek oluşturulmuş bir DSA mimarisi yardımıyla sınıflar arası ayrımın daha belirgin olacağı düşünülen yeni vektörlere dönüştürülür. Bu yaklaşıma göre Şekil 33'te verilen DSA mimarisi, nispeten daha çok konuşmacı sayısına sahip bir VK ile eğitilmiştir. Eğitimin sonunda konuşmacı kimliklerini tespit etmekten sorumlu olan çıkış katmanı DSA'dan ayrılmış ve Şekil 33'te gösterilen gizli katman ve öznitelik katmanı MFKK özniteliklerinin dönüştürülmesinde kullanılmıştır. Bu mimari ile 23×1 boyutundaki bir MFKK öznitelik vektörü alınarak önce



Şekil 33. Önerilen öznitelik dönüştürücü derin sinir ağının mimarisi

bir gizli katman yardımıyla 100×1 boyutlu bir diziye dönüştürülmekte, ardından N önceden seçilen bir sayı olmak üzere $23 \times N$ boyutunda bir vektöre dönüştürmektedir.

Eğitilen DSA tabanlı öznitelik dönüştürme modeli daha sonra GSA için kullanılmaktadır. Genel bir bakışla, OKT uygulaması için elde edilen MFKK öznitelikleri önce bir DSA modeliyle dönüştürülmekte ve ardından dönüştürülen bu vektörler GSA'ya giriş olarak verilmektedir. Bu bağlamda, önerilen bu yaklaşımın hem DSA'nın hem de GSA'nın özelliklerinden ardışık olarak yararlanan hibrit bir yaklaşım olduğu savunulabilmektedir.

Öznitelik dönüştürmek için önerilen DSA mimarisi, çeşitli yaklaşımlarla çeşitli vektörler elde edilmesi amacıyla dört farklı türde eğitilmiştir. Bu türlerden ikisi analog öznitelikleri analog vektörlere dönüştürürken, diğer iki türde ise analog öznitelikler spike

örüntüsü olarak kullanılmak üzere ikili kodlanmış vektörlere dönüştürülmüştür. Takip eden iki alt başlıkta, önerilen DSA yapıları tanıtılmıştır.

2.5.1. Analog-Analog Dönüşüm

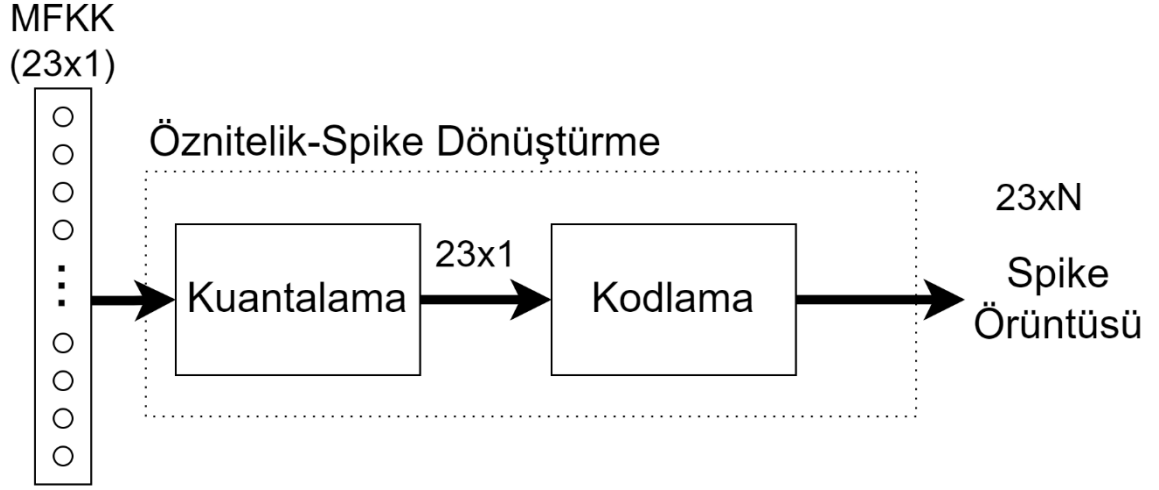
Literatürde kullanımı göz önünde bulundurulduğunda, Şekil 33'te verilen DSA öznitelik dönüştürücüsünün, analog MFKK özniteliklerini işledikten sonra, çıkışında analog vektörler vermesi beklenmektedir. Bu özelliklerden yola çıkılarak, çalışmada iki farklı analog-analog dönüşüm yapan DSA modeli eğitilmiştir. Bu iki DSA modelini birbirinden ayıran özellik, öznitelik katmanında bulunan YN sayısıdır.

Öncelikle MFKK öznitelik sayısının her çerçeve için 23 olduğu göz önünde bulundurulursa, DSA'nın bu 23 adet öznitelikten oluşan vektörü işledikten sonra yine 23 adet elemandan oluşan bir vektör çıkışı vermesi istenmiştir. Bu nedenle, Şekil 33'te verilen DSA modeli için öznitelik katmanında bulunan YN sayısı 23 olarak, yani $N = 1$, seçilmiştir. Bir diğer yaklaşım ile çıkış katmanında giriş katmanının iki katı uzunluğunda vektör çıktısı vermesidir. Bu nedenle öznitelik katmanında bir seçenek olarak da 46 adet, yani $N = 2$, YN kullanılmıştır.

DSA'nın analog-analog dönüştürücü amacıyla eğitilmesi sırasında gizli katmanda ve öznitelik katmanında bulunan YN'ler için aktivasyon fonksiyonları Denklem (5)'te verilen ReLU olarak seçilmiştir. Böylece çıkış vektörlerinde bulunan elemanlar sadece pozitif değerler alabileceklerdir. Eğitimin sonucunda analog-analog dönüşüm DSA'ları için, eğer $N = 1$ olarak seçildiyse bu vektörler Analog 1 anlamında A1, eğer $N = 2$ olarak seçildiyse bu vektörler Analog 2 anlamında A2 olarak isimlendirilmiştir. Bu vektörler Şekil 27'de verilen GSA akışında ile verilen akışta eğitim ve test öznitelikleri yerine kullanılmaktadır.

2.5.2. Analog-İkili Dönüşüm

Bölüm 2.4.1'de anlatılan ÖSD yöntemi Şekil 34'te gösterildiği gibi özetlenebilmektedir. Buna göre analog formdaki öznitelik vektörleri önce seçilen bir kuantalama tekniği uyarınca kuantalanmakta, ardından seçilen bir kodlama yöntemi uyarınca ikili forma dönüştürülmektedir. Genel bir bakışla, önerilen yöntemin seçilen kuantalama tekniklerinin ve kodlama yöntemlerinin kısıtlarından etkilenebileceği söylenebilmektedir. Bu nedenle, ÖSD aşamasında DSA tabanlı bir yaklaşım, sistemin performansını artırma konusunda önemli bir adaydır.



Şekil 34. Önerilen öznitelik-spike dönüştürme yönteminin basit yapısı

Bu çalışmada, Şekil 33'te verilen DSA modeli, girişinde aldığı MFKK özniteliklerini çıkışında ikili spike örüntüleri üretecek şekilde eğitilmiştir. Modelin mimari yapısı aynı tutulmuş ve gizli katmanda bulunan YN'lerin aktivasyon fonksiyonları Denklem (5)'te verilen ReLU olarak seçilmiştir. Spike dönüşüm için özel olarak, öznitelik katmanında bulunan YN'lerin aktivasyon fonksiyonları Denklem (4)'te verilen sigmoid fonksiyonu olarak seçilmiştir. Seçilen sigmoid fonksiyonu sayesinde öznitelik katmanında çıkışlar ikili forma yakın bir şekilde elde edilmektedir. Son olarak, çıkışlar en yakın bit değerine yuvarlanarak spike örüntüleri elde edilmiştir.

Model, analog-analog dönüşüm modellerine benzer olarak iki farklı türde eğitilmiştir. İlk türde; öznitelik katmanında 23×16 boyutunda spike örüntüleri, ikinci türde ise öznitelik katmanında 23×32 boyutunda spike örüntüleri çıktısı meydana gelmektedir. Bu çıktılar Binary 16 ve Binary 32 anlamında, sırasıyla B16 ve B32 olarak isimlendirilmiştir. İkili formdaki çıktılar, Şekil 27'de verilen akışta ÖSD adımı atlanarak spike örüntüsü yerine kullanılmaktadır.

2.6. Deneysel Çalışmalar

Bu bölümde çalışmada gerçekleştirilen deneysel adımlar açıklanmış ve detayları verilmiştir. Bölüm, beş alt başlıktan meydana gelmektedir. İlk alt başlıkta, OKT uygulamalarında kullanılan en gelişkin sınıflandırma yöntemlerinin performanslarının incelendiği deneysel senaryo tanıtılmıştır. İkinci alt başlıkta, tez çalışmasında önerilen

ÖSD yaklaşımının sınıflandırma performansına etkisinin çeşitli VK'ler ve klasik makine öğrenimi yöntemleri kullanılarak incelendiği deneysel senaryoya dair detaylar verilmiştir. Üçüncü alt başlıkta, GSA modelinde SB tiplerinin sınıflandırma performansına etkisinin incelenmesi amacıyla tasarlanmış deneysel senaryo tanıtılmıştır. Dördüncü alt başlıkta, GSA ve DSA modelleri hibrit olarak kullanıldığı durumlarda, OKT performansının nasıl etkilendiğinin incelendiği deneysel senaryonun adımları verilmiştir. Beşinci ve son alt başlıkta ise PN'lerin OKT performansının incelenmesine dair tasarlanan deneylere dair detaylar verilmektedir.

2.6.1. D1: Otomatik Konuşmacı Tanımda En Gelişkin Yöntemlerin Performansının İncelenmesi

Bir OKT uygulamasının gerçekleştirildiği varsayıldığında; toplam N adet konuşmacının tanınması istenen bir OKT görevinde, sistemin L saniye uzunluğundaki konuşmalardan konuşmacı kimliğini mümkün olduğunca doğru bir şekilde belirlemesi beklenmektedir. Bu amaca ulaşmak için sistem tasarımcısı, T saniye uzunluğunda eğitim konuşmalarını bir sınıflandırıcıyı eğitmek için kullanmaktadır. Öncelikle eğitim; konuşma verilerinin akustik özellikleri, çoğunlukla MFKK, hesaplanmaktadır ve ardından M elemanlı öznitelik vektörleri elde edilmektedir. Bu öznitelik vektörleri, eğitim aşamasında X sınıflandırıcının eğitilmesinde kullanılmaktadır. Sınıflandırıcıya bağlı olarak tasarımcının uygun parametreleri de ayrıca seçmesi gerekmektedir. Örneğin, bir Gauss karışım modeli (GKM) modeli kullanılacaksa karışım sayısının seçilmesi gerekirken, bir DSA modeli kullanılacaksa gizli katmanların sayısının seçilmesi gerekmektedir. Örnekte bahsedilen sınıflandırıcıların haricinde başka sınıflandırıcılar da kullanılabilmekte fakat her sınıflandırıcı için belirli parametrelerin seçilmesi zorunluluğu sabit kalmaktadır. Eğitim aşamasının sonunda sistemden, L saniye uzunluğundaki test konuşmalarından konuşmacıların kimliklerini bulması beklenmektedir. Test aşamasının sonunda sistemin performansı çoğunlukla Bölüm 1.4.3'te tanıtılan sınıflandırma ölçütleri olmak üzere çeşitli performans ölçütleri ile değerlendirilmektedir.

Bu senaryo, olası parametre ve yöntemlerin çeşitliliğini ortaya koymaktadır. Özellikle parametreler, tasarımcı tarafından seçilen parametreler veya tasarımcıdan bağımsız parametreler olarak iki ayrı grupta incelenebilmektedir. MFKK sayısı, karışım sayısı ve gizli katman sayısı gibi parametreler VK ile ilgisi olmadığı için seçilebilirken, eğitim verisi uzunluğu (EVU) ve test verisi uzunluğu (TVU) gibi parametreler çoğunlukla

VK'nin nitelikleriyle sınırlıdır. Bu bağlamda tasarımcılar, en iyi performansı elde edebilecek, en uygun değerleri seçme konusunda zorluklarla karşı karşıyadır, çünkü kötü seçilmiş parametreler ve/veya yöntemler performansı önemli ölçüde etkileyebilmektedir (Özcan & Kayıkçıoğlu, 2021).

Bu deneyde, literatürdeki OKT uygulamaları arasında performansları bakımından en gelişkin yöntemler olarak kabul edilen yaklaşımların performansları LASRC veri tabanı kullanılarak incelenmiştir. Deney üç ana amaca hizmet etmektedir:

1. En gelişkin OKT yöntemlerinin, LASRC veri tabanında eşit şartlarda karşılaştırılması
2. OKT uygulamalarında seçilmesi gereken parametrelerin ve VK kısıtlarının performans üzerindeki etkisinin incelenmesi
3. GSA modeli ile OKT uygulaması yapıldığında, seçilen parametreler dahilinde önerilen model ve en gelişkin yöntemlerin karşılaştırılabilmesi

Bu deney iki adımda gerçekleştirilmiştir. İlk adımda, LASRC'nin alt kümeleri kullanılmıştır. Konuşmacı sayısının performansa etkisinin incelenebilmesi adına en az 40 farklı konuşmacı kayıtlarına ihtiyaç duyulmaktadır. Bu nedenle, Tablo 2'de verilen LASRC veri tabanında, 40'ar adet konuşmacı sayısına sahip olan dev-clean ve test-clean alt kümeleri birleştirilmiş ve toplam 80 sınıfa sahip bir konuşmacı havuzu elde edilmiştir. Bu havuz, ihtiyaç duyulduğunda gerekli sayıda konuşmacı kayıtlarının rastgele seçilmesi için kullanılmıştır. Yöntemler arasında geliştirme veri tabanına ihtiyaç duyanlar için Tablo 2'de verilen alt kümeler arasından 251 farklı konuşmacı kaydına sahip olan train-clean-100 alt kümesi kullanılmıştır.

İkinci adımda, en gelişkin yöntemler ve önerilen GSA modelinin karşılaştırılabilmesi adına, GSA modelinde kullanılacak olan parametre dizileri ile en gelişkin yöntemlerin performansı incelenmiştir. Buna göre; geliştirme aşamasında dev-other ve yöntemlerin performansının ölçülmesinde ise dev-clean alt kümeleri kullanılarak, en gelişkin yöntemler ile önerilen GSA modelinin VK bakımından aynı şartlar altında değerlendirilebilmeleri adına deneyler tekrar edilmiştir. Birinci ve ikinci adımlarda kullanılan VK'ler ve detayları Tablo 8'de verilmektedir.

Deneyde, yedi farklı sınıflandırma yaklaşımı kullanılmış ve karşılaştırılmıştır. Bu yaklaşımlar istatistiksel yaklaşımlar ve DÖ tabanlı yaklaşımlar olarak iki gruba ayrılabilir. Kullanılan istatistiksel yaklaşımlardan, Gauss karışım modeli (GKM) yaklaşımı için detaylı bilgiye Reynolds'un (1995) çalışmasından, Gauss karışım modeli-

Tablo 8. D1 için kullanılan parametre dizileri

Seenekler	Birinci Adım Deęerleri			İkinci Adım Deęerleri		
	Düşük	Orta	Yüksek	Düşük	Orta	Yüksek
Konuşmacı Sayısı	10	20	40	4	8	16
MFKK Sayısı	5	10	20		23	
EVU (sn)	5	10	20		5	
TVU (sn)	1	2	4		1	
Karışım Sayısı (GKM için)	16	32	64	16	32	64
Gizli Katman Sayısı (DSA için)	2	-	4	2	-	4
Segment (ZGSA için)	a ve b			a ve b		
Veri Tabanı (Eęitim ve Test)	dev-clean + test-clean (80 Konuşmacı)			dev-clean (40 Konuşmacı)		
Veri Tabanı (Geliştirme)	train-clean-100 (251 Konuşmacı)			dev-other (33 Konuşmacı)		

evrensel arka plan modeli (GKM-EAPM) yaklaşımın için detaylı bilgiye Reynolds ve arkadaşlarının (2000) çalışmalarından, Gauss karışım modeli-destek vektör makineleri (GKM-DVM) yaklaşımı için detaylı bilgiye Campbell ve arkadaşlarının (2006) çalışmalarından ulaşılabilmektedir. DÖ tabanlı yaklaşımlar için ise d-vektörler yaklaşımı için detaylı bilgiye Variani ve arkadaşlarının (2014) çalışmalarından, x-vektörler yaklaşımı için Snyder ve arkadaşlarının (2018) çalışmalarından ulaşılabilmektedir. Kullanılan DSA ve DSA-TÖ modelleri ile ilgili detaylı bilgiye ise Özcan ve Kayıkçıoęlu'nun (2021) çalışmasından ulaşılabilmektedir. Deneyde, ayrıca kullanılan yaklaşımlar için seçilmesi gereken parametrelerin de performansa etkileri incelenmiştir. Bu parametreler GKM tabanlı yaklaşımlar için karışım sayısına karşılık gelirken, DSA ve DSA-TÖ yaklaşımları için ise gizli katman sayısını ifade etmektedir. Ayrıca, x-vektörler yaklaşımında seçenek olarak hem “a” segmenti hem de “b” segmentinin performansları incelenmiştir.

Deneyin ana amaçlarında biri yukarıda sıralanan en gelişkin OKT yaklaşımlarının LASRC ile performansının incelenmesidir. Bir dięer amaç ise OKT uygulamalarında seçilen parametreler ve VK kısıtlarının performansa etkisinin incelenmesidir. Bu nedenle, ilk adımda deney süresince önemli olan parametreler deęiştirilerek performansın nasıl etkilendięi incelenmiştir. Bu parametreler ve deneyde kullanılan deęerler Tablo 8’de verilmektedir. Buna göre parametreler iki grupta incelenebilmektedir. İlk grupta kullanıcı

tarafından bir ön-kabul olarak seçilmesi gereken parametrelerdir. Kullanılacak MFKK özniteliklerinin sayısı, GKM sınıflandırıcısının karışım sayısı ve DSA modelinde kullanılacak gizli katman sayıları ilk grupta sayılabilmektedir. İkinci grupta ise kullanıcının belirleyemediği, daha çok VK'nin kısıtları veya uygulamanın zorluğu ile ilintili parametrelerdir. EVU, TVU ve OKT uygulamasındaki konuşmacı sayısı bunlar arasında sayılabilmektedir. Bu deneyin birinci adımında ilgili parametreler Tablo 8'de verildiği gibi düşük, orta ve yüksek değerler olmak üzere üç seviyede incelenmiştir. İkinci adımda, birinci adımdan elde edilen tecrübeler doğrultusunda, GSA modelinin performansının ölçülmesi için seçilen parametreler kullanılarak birinci adımda yapılan işlemler tekrar edilmiştir. Bu parametre değerleri Tablo 8'de ayrıca gösterilmektedir. Buna göre EVU değeri 5 saniye, TVU değeri 1 saniye ve MFKK sayısı 23 seçilerek, senaryo, konuşmacı sayıları 4, 8 ve 16 olacak şekilde tekrar edilmiştir.

Deneye başlamadan önce, yöntemler arasında geliştirme aşamasına sahip olanlar için gerekli modellerin eğitilmesi gerekmektedir. Buna göre, GKM-DVM ve GKM-EAPM yaklaşımları için EAPM'lerin, DSA-TÖ, x-vektör ve d-vektör yaklaşımları içinse DSA modellerinin geliştirme veri tabanları kullanılarak eğitilmesi gerekmektedir. Birinci ve ikinci adımlarda, Tablo 8'de gösterilen veri tabanları kullanılarak, geliştirme modelleri belirtilen parametreler gözetilerek ayrı ayrı eğitilmiş ve kaydedilmiştir. Geliştirme aşamasında ilgili VK'den her konuşmacı için 1'er saniyelik veriler kullanılmıştır. Geliştirme aşamasının ardından eğitim aşamasına geçilmiştir. DSA modellerinin eğitiminde tüm modeller için epok sayısı 200 olarak seçilmiştir.

Eğitim aşamasında, kullanılan yöntemlerin her biri için ilgili çalışmalarında verilen prosedürler aynen uygulanmıştır. Birinci adım için, öncelikle Tablo 8'de verilen konuşmacı sayısı kadar sınıf verisi, oluşturulan VK içinden rastgele seçilmiştir. Bu konuşmacılar için, Tablo 8'de verilen TVU değerlerinden biri seçilerek eğitim kümesi oluşturulmuş ve bu eğitim kümesi örneklerinden Tablo 8'de verilen MFKK sayısı değerlerinden seçilen biri kadar öznitelik çıkarılmıştır. Elde edilen öznitelik vektörleri her yöntem için ilgili çalışmalarında detaylandırılan prosedürler takip edilerek eğitilmiştir. Buraya kadar bahsedilen kısım, sadece bir parametre dizisi için geçerlidir. Tablo 8'de verilen parametrelerin her biri birbirleriyle en az bir kez eşleşecek sayıda aynı işlem tekrar edilmiş ve eğitim modelleri oluşturulmuştur. Güvenilir sonuçlar elde etmek adına, her parametre ve yöntem dizileri için eğitim 50'şer kez tekrar edilmiş ve modeller kaydedilmiştir. İkinci adımda ise EVU ve MFKK sayısının sırasıyla 5 saniye ve 23 olduğu

kabul edilerek aynı işlemler konuşmacı sayısı 4, 8 ve 16 olacak şekilde tekrarlanmıştır. Birinci ve ikinci adımlarda parametreler harici tüm aşamalar birebir aynıdır.

Test aşamasına gelindiğinde eğitilen modellerin performansı, Tablo 8’de gösterilen TVU değerleri gözetilerek elde edilen konuşma kayıtları üzerinden ölçülmüştür. Birinci adımda TVU değerleri üç seviye de değiştirilmiş olsa da ikinci adımda TVU değeri 1 saniye olacak şekilde kabul edilmiştir. Bu deneyde, performans ölçütü olarak F1-Skor kullanılmıştır. Deneyin birinci adımına dair daha detaylı bilgi ve sonuçlara bu tez çalışması uyarınca gerçekleştirilmiş olan çalışmadan ulaşılabilmektedir (Özcan & Kayıkçıoğlu, 2021). Bu deney için, geliştirme, eğitim ve test aşamalarının tümü, Python dilinde scikit-learn ve TensorFlow kütüphaneleri kullanılarak gerçekleştirilmiştir (Kramer, 2016; Abadi vd., 2015).

2.6.2. D2: Öznitelik Spike Dönüşümü Yaklaşımının Karşılaştırmalı İncelenmesi

Bölüm 2.4’te bir GSA modeli önerilmiştir. Ancak bu GSA modeli, bir sınıflandırma uygulamasında kullanılmadan önce test edilmeye, kapasite ve kısıtlarının detaylı olarak belirlenmesine muhtaçtır. Geliştirilen bir sınıflandırma modelinin test edilmesinde iki yaklaşım bulunmaktadır. Bunlardan ilki modelin aynı problemi çözmekten sorumlu olduğu diğer sınıflandırıcılarla karşılaştırmalı olarak incelenmesidir. Bu sayede modelin paralel yöntemler karşısında performansı incelenebilmektedir. Bir diğer yaklaşım ise modelin bilinen VK’ler ile test edilmesidir. Bu yaklaşımın ilk adımı yapay bir VK kullanılması ve basit bir VK ile kısıtlarının tespit edilmesidir. İkinci adımı ise genel kabul görmüş kısıtları performansı yaygın olarak bilinen karşılaştırmalı VK’lerde test edilmesidir (Malakar vd., 2018). Kullanılan test yönteminden bağımsız olarak, bu çalışmada önerilen GSA modelinin kısıtlarının ve kapasitesinin ölçülmesi elzemdir.

Önerilen GSA modeli öznitelik vektörlerini analog formda işleyememekte ve analog formdaki öznitelikleri ikili forma dönüştürmek üzere biz dizi teknik kullanmaktadır. Bu tekniklerin performansa nasıl etki ettiğinin ölçülmesi de GSA’nın kısıtlarının ve kapasitesinin belirlenmesinde önemli bir adımdır. Kayıkçıoğlu Bozkır ve arkadaşları (2023) çalışmalarında tekdüze ve Gray kodlama kullandıklarını bildirmişlerdir. Fakat bu tez çalışmasıyla önerilen GSA modelinde, seçilebilir durumda olan diğer yöntemlerin performansı nasıl etkilediği literatürde cevaplanmamış durumdadır. Ayrıca, ÖSD aşamasında özniteliklerin bilgi içeriklerinin nasıl değiştiği de ayrıca cevaplanması gereken

bir sorudur. Bu nedenle, analog değerlere sahip öznitelik vektörlerinden spike örüntüleri üretme sürecinde performansın ne yönde etkilendiği de belirlenmelidir.

Bu deneyde, önerilen GSA modelinin performansı çeşitli veri kümeleri ile incelenmiştir. Deney üç ana amaca hizmet etmektedir:

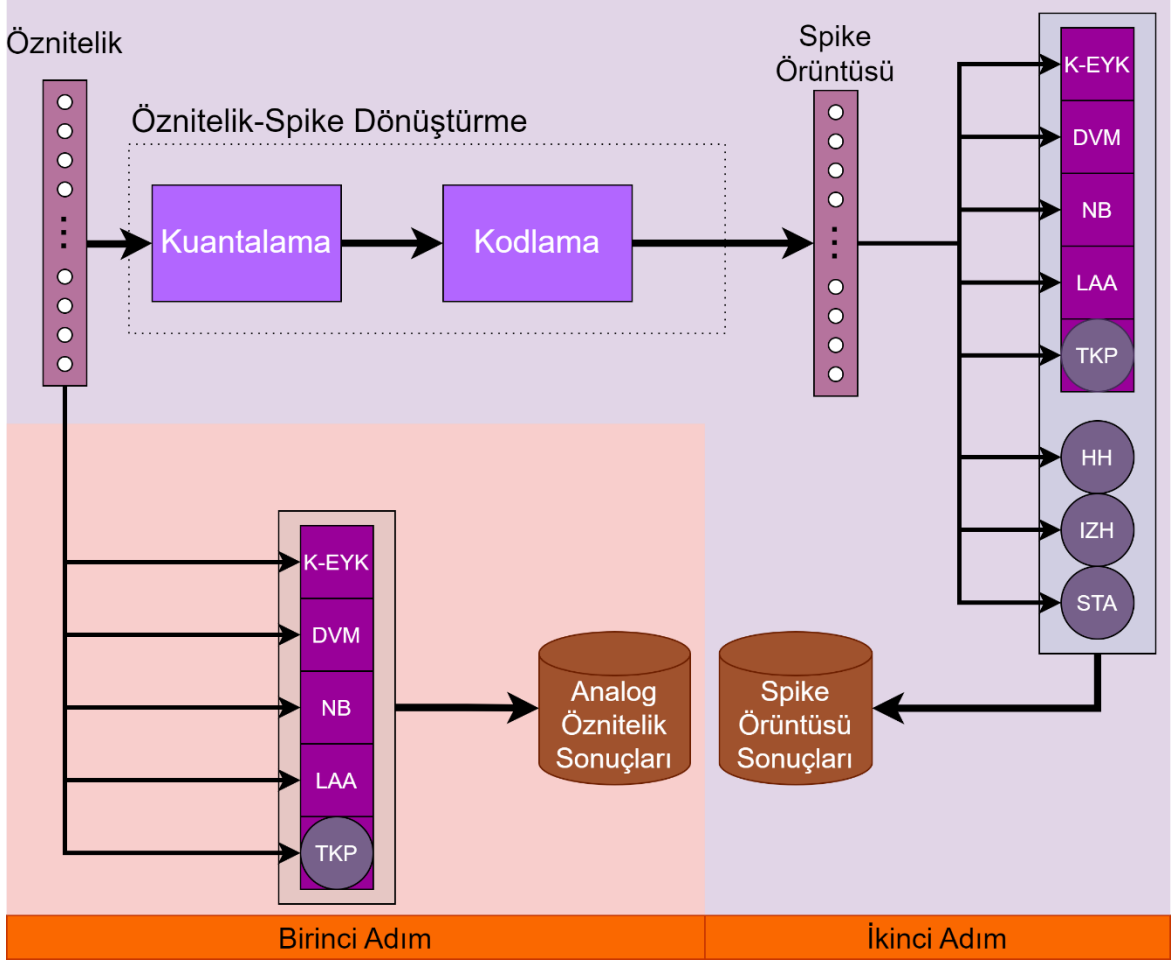
1. GSA modelinin bir yapay VK, iki karşılaştırmalı VK ve 6 sınıflı bir konuşmacı tanıma VK'sinde performansının incelenmesi
2. GSA modelinin geleneksel MÖ algoritmalarıyla karşılaştırılması
3. GSA modelinde bit uzunluğu, kuantalama tekniği ve kodlama yöntemi gibi seçilmesi istenen özelliklerin performansa etkisinin incelenmesi

Bu çalışmada kullanılan VK'ler amaca yönelik üç ana grupta sıralanabilmektedir. Kullanılan ilk VK, Bölüm 2.1.1'de tanıtılan yapay veri kümesidir. Bu VK'nin kullanım amacı GSA modelinin kısıtlarının ve kapasitesinin, öznitelikleri iki boyutlu uzayda dağılmış, çok sınıflı bir problemde performansının görselleştirilmesidir. Böylece yapılan çalışmaların etkileri görselleştirilebilir bir VK üzerinden incelenebilmektedir.

Kullanılan ikinci grup VK'ler, Bölüm 2.1.2'de tanıtılan IRIS ve MNIST veri kümeleridir (Fisher, 1936; Deng, 2012). Bu VK'ler, yapılan çalışmaların etkilerinin literatürde verilen yöntemlerle karşılaştırmalı olarak incelenebilmelerine olanak tanımaktadırlar. Bu sayede, GSA modelinin performansı evrensel manada karşılaştırılabilir hale gelmektedir.

Son VK ise Tablo 3'te detayları verilen VK6 konuşmacı tanıma veri kümesidir. VK6 için hesaplanan MFKK öznitelikleri için Bölüm 2.3'te açıklanan adım ve parametrelerin takip edilmesiyle elde edilmiştir. Bu tez çalışmasının OKT problemine GSA tabanlı bir yaklaşım önerme çabası içinde olması, bu VK'nin deneye dahil edilmesinin önemli olacağını göstermektedir. Deneyde, ayrıca, ÖSD aşamasında kullanılan yöntemlerin performansı ölçülmektedir. İlerleyen deneylerde hangi yöntemin seçileceğinin belirlenmesinde GSA modelinin VK8'deki performansı göz önünde bulundurulmuştur.

Bu deneyde, GSA modelinde basit GN modelleri olan HH, IZH ve STA nöronları ayrı ayrı seçilerek performansları ölçülmüştür. Ayrıca, karşılaştırma amacıyla 5 klasik MÖ yöntemi kullanılmıştır. Bu yöntemler; k-en yakın komşuluk (K-EYK), destek vektör makineleri (DVM), Naive Bayes, (NB), lineer ayırıcı analizi (LAA) ve tek katmanlı perceptron (TKP) sınıflandırıcılarıdır (Cover & Hart, 1967; Cortes & Vapnik, 1995; Cheeseman vd., 1988; Xanthopoulos vd., 2013). K-EYK, DVM, NB ve LAA sınıflandırıcıları Python programlama dilinde scikit-learn kütüphanesinde varsayılan



Şekil 35. D2 için belirlenen deney senaryosunun akış diyagramı

parametre dizileri ile kullanılmış (Kramer, 2016), TKP ise Kayıkçıoğlu Bozkır ve arkadaşlarının (2023) çalışmasında önerildiği gibi kullanılmıştır.

Deney, elde edilen sonuçların niteliği, kullanılan VK'ye göre değişmek üzere iki adımdan meydana gelmektedir. İlk adımda, yukarıda sıralanan klasik MÖ yöntemleri kullanılarak VK'lerin performansı analog öznitelik vektörlerine göre incelenmiştir. D2 ile ilgili genel senaryo Şekil 35'te verilmektedir.

Yapay VK, klasik MÖ yöntemleriyle eğitilmiş ve iki boyutlu uzayda sınıfların karar sınırları çizdirilmiştir. Benzer şekilde, IRIS ve MNIST VK'leri kullanılarak MÖ yöntemleri eğitilmiş ve performansları F1-Skor değerleri üzerinden ölçülmüştür. VK8 için konuşmacı başına EVU değeri 5 saniye ve TVU değeri 1 saniye olmak üzere konuşma kayıtları oluşturulmuştur. Eğitim aşamasında konuşma kayıtlarının çerçeveleri sanki ayrı ayrı örneklermiş gibi eğitilmiştir. Test aşamasında kayıtların tüm çerçeveleri için etiketler elde edilmiş ve çerçevelerin meydana getirdiği konuşma kayıtlarının etiketleri sahip olduğu

çerçevelerin etiketlerine göre belirlenmiştir. Performans ölçütü olarak F1-Skor kullanılmıştır. Deneyin ilk adımı sayesinde analog öznitelik vektörlerinin performansı klasik MÖ yöntemleri ile belirlenmiştir. Bu adımda basit nöron modelleri kullanılmamıştır.

Deneyin ikinci adımında tüm VK'ler Bölüm 2.4.1'de detaylandırılan ÖSD aşamasına tabii tutulmuştur. Spike örüntüleri için, ilk adımda yapılan işlemlerin aynısı tekrarlanmış ve aynı performans ölçümleri yapılmıştır. Bu bölümde ayrıca HH, IZH ve STA nöron modelleri de dahil edilmiştir. GSA modelinin eğitiminde, tüm modeller için epok sayısı 500 olarak seçilmiştir. Önerilen GSA modeli için geliştirilen yazılımda; ÖSD için varsayılan bit uzunluğu 5 olarak, varsayılan kuantalama tekniği tekdüze olarak ve varsayılan kodlama yöntemi ise yoğunluk kodlama olarak seçilmektedir. Bu adımda, ayrıca, ÖSD aşamasında seçilen parametrelerin performansa etkisi de incelenmiştir. Buna göre diğer seçenekler Bölüm 2.4.7'de sıralanan varsayılan seçenekler olarak belirlenmesi kaydıyla, bit uzunluğu, kuantalama tekniği ve kodlama yönteminin etkisi ayrı ayrı incelenmiştir. MNIST özelinde özniteliklerin hali hazırda 16 seviye ile kuantalanmış olmasından dolayı, bu VK, kuantalama tekniği ve bit uzunluğunun etkisinin incelendiği deneylerden muaf tutulmuştur.

2.6.3. D3: Sinaptik Bağlantı Tiplerinin Sınıflandırma Performansına Etkisinin İncelenmesi

Sinapslar BN'ler arası bağlantının yegâne yoludur. Bu nedenle, BN'lerin iletişimde kritik rolü üstlenerek BN'lerin işlemesi gereken bilgilerin elde edilmesinde görev almaktadırlar. GN'lerin de BN'lerin gerçekçi birer modelleri olduğu göz önünde bulundurulduğunda, SB modellerinin de GSA'lardaki rolünün önemi ortaya çıkmaktadır. Bölüm 1.1.4 ve 1.3.6'da tanıtılan biyolojik ve gerçekçi sinapslar incelendiğinde SB'lerin temelde iki farklı grupta incelenebileceği görülmektedir. Exc ve Inh SB'ler olarak bilinen bu türlerde; Exc SB'ler, GN'lerin spike üretmesi konusunda uyarıcı etkiye sahipken; Inh SB'ler ise baskılayıcı etkiye sahiptir. Bir sınıflandırma probleminde Exc ve Inh SB'lerin performansı nasıl etkilediği literatürde hala tam olarak belirlenememiştir (Du vd., 2015). Özellikle, çalışmaların Inh SB'leri genellikle ihmal etmeleri, Inh SB'ler konusunda bilgilerin ayrıca yetersizliğini gözler önüne sermektedir. Bu nedenle, SB tiplerinin bir sınıflandırma probleminde performansı ne denli etkileyeceğinin belirlenmesi önem arz etmektedir.

GN modellerine entegre edilen SB modellerinde, SB ağırlık değerlerinin doğaları gereği negatif değer almamaları gerekmektedir. Bu nedenle GSA modelleme çalışmalarında bu ağırlıklara sadece pozitif değerler atanmaktadır. Fakat bir sınıflandırma problemünde ağırlıkların güncellenmesi anında SB ağırlıklarının negatif değere yakınsamaları söz konusu olabilmektedir. Bu nedenle, Moldwin ve Segev (2020) çalışmalarında eğitim algoritmalarını ağırlıkların negatif değer alması durumunda, ağırlık değerini 0 olmaya zorlayarak bir kısıt ortaya koymuşlardır. Fakat Kayıkçıoğlu Bozkır ve arkadaşları (2023) çalışmalarında bunun sınıflandırma performansına olumsuz yönde etki ettiğini ve önerdikleri aynalama yaklaşımı sayesinde ağırlık kısıtlamadan kaynaklı etkilerin ortadan kalktığını ispatlamışlardır. Fakat çalışmalarında Inh SB kullanmadıklarından, aynalama yönteminin Inh SB bakımından nasıl bir etki yaptığı cevapsız kalmaktadır.

Bu deneyde, önerilen GSA modelinde SB tiplerinin sınıflandırma performansına etkisi çeşitli veri kümeleri ile incelenmiştir. Deney iki ana amaca hizmet etmektedir:

1. GSA modelinde aynalama yaklaşımı kullanılmasının sınıflandırma performansına etkisinin incelenmesi
2. Sinaptik reseptörlerde kullanılan Exc ve Inh SB'lerin sınıflandırma performansına etkisinin incelenmesi

Bu çalışmada, Bölüm 2.1.1'de tanıtilan yapay VK; Bölüm 2.1.2'de tanıtilan karşılaştırma VK'leri ve detayları Tablo 3 verilen, 6 konuşmacı sayısına sahip VK6; GSA modelinin performansının karşılaştırmalı olarak incelenmesinde kullanılmıştır. Yapay VK'nin deneyde kullanılmasının amacı, GSA modelinin kısıtlarının ve kapasitesinin, öznelilikleri iki boyutlu uzayda dağılmış, çok sınıflı bir problemde performansının görselleştirilmesidir. Karşılaştırmalı VK'lerin kullanılmasının nedeni ise GSA modelinin performansının bilinen VK'lerde incelenmesidir. Son olarak, VK6, ilerleyen deneylerde bir konuşmacı tanıma uygulamasında GSA parametrelerinin nasıl seçilmesi gerektiğinin belirlenmesi amacıyla deneye dahil edilmiştir. Yapay ve karşılaştırmalı VK'ler tekil örüntü içeren öznelilik vektörlerine sahip olduğundan aynen kullanılmış, VK6 için her konuşmacı için 5 saniyelik eğitim konuşmaları kullanılarak, 1'er saniyelik test konuşmaları üzerinden performansları ölçülmüştür.

Bu deneyde, SB tiplerinin sınıflandırma performansına etkisinin incelenmesi adına, GSA tasarımında aynalama yaklaşımının varlığı, sadece Exc SB'lerin varlığı, sadece Inh SB'lerin varlığı ve hem Exc hem de Inh SB'lerin varlığı olmak üzere toplam altı farklı durumda sınıflandırma performansı incelenmiştir. Tüm durumlar için basit GN modelleri

olan HH, IZH ve STA nöronları kullanılmıştır. Sinaptik reseptörlerde bulunan ikisi Exc ve ikisi Inh SB'ler olmak üzere toplam dört SB tipinden; sadece Exc SB'ler kullanıldığında, AMPA ve NMDA sinapsları aktif edilmiş; GABA sinapslarının ağırlık güncellenmeleri engellenmiştir. Benzer olarak; sadece Inh SB'ler kullanıldığında, GABA sinapsları aktif hale getirilmiş, AMPA ve NMDA sinapslarının ağırlık güncellemeleri engellenmiştir. Exc ve Inh SB'lerin kullanıldığı durumlarda ise dört SB modeli de aktif hale getirilmiştir. Deney süresince seçilmesi gereken diğer tüm özellikler yazılımda verilen varsayılan değerler olarak seçilmiştir. GSA modelinin eğitiminde, tüm modeller için epok sayısı 500 olarak seçilmiştir. Bu deneyde, performans ölçütü olarak F1-Skor kullanılmıştır.

2.6.4. D4: Önerilen Gerçekçi Sinir Ağı Modelinin Konuşmacı Tanıma Performansının İncelenmesi

OKT, kişilerin kimliğinin kayıtlı konuşma kayıtları ile tespit edilmesidir. Bu amaçla, literatürde çeşitli yöntemler önerilmiştir. Bu yöntemler istatistiksel ve DÖ tabanlı yöntemler olarak gruplanabilmektedir (Beigi, 2011; Kinnunen & Li, 2010). İstatistiksel yöntemler daha çok GKM kümeleme yönteminin çeşitli uygulamalara dahil edilmesini içermektedir. DÖ tabanlı yöntemler ise iki ayrı gruba ayrılabilir. İlk grupta klasik bir MÖ yaklaşımıyla eğitim verileri ile bir DSA modelinin eğitilmesi ve test verilerinin bu modelin performansının ölçümünde kullanılması sayılabilir. Bir diğer DÖ tabanlı yaklaşım ise nispeten büyük veri tabanları kullanılarak eğitilen DSA modellerinin sınıfları tespit etmekten sorumlu olan son katmanının çıkarılmasıyla güçlü öznitelikler elde etmek üzere kullanılan modellerdir.

DÖ tabanlı yöntemlerin gösterdiği başarılar ve DSA'ların YN'lerden meydana gelmesi; OKT uygulamalarında BN'lerin daha gerçekçi versiyonları olan GN'lerin ve GN'lerden meydana gelen GSA'ların performansının merak edilmesine neden olmaktadır. Literatürde, OKT uygulamalarında GSA modellerinin kullanılması konusunda çalışmalar sınırlı kalmaktadır. Bu nedenle, GSA'ların OKT uygulamalarındaki performansının incelenmesi literatüre önemli bir katkı olarak sayılabilir.

OKT uygulamalarında çeşitli öznitelik çıkarma yöntemleri kullanılmaktadır. Bu öznitelikler konuşmacıların ayırıcı özelliklerinin ortaya çıkarılabilmesi adına kritik öneme sahiptirler. Literatürde, MFKK özniteliklerinin kullanımının diğer yöntemlere kıyasla ağır bastığı söylenebilir. Bunun yanında, MFKK özniteliklerinden faydalanan TÖ esaslı yaklaşımlar sayesinde konuşmacı kimliğinin daha ayırt edilebilir olduğu öznitelikler de

elde edilmektedir. Bu tez çalışmasında önerilen DSA modeli, hem daha güçlü vektörler elde etmek adına hem de GSA modelinin ihtiyaç duyduğu ikili formdaki spike örüntülerini üretebilme adına önemli bir potansiyel sunmaktadır. Bu potansiyelin ortaya koyulabilmesi adına önerilen öznitelik vektörlerinin test edilmesi gerekmektedir.

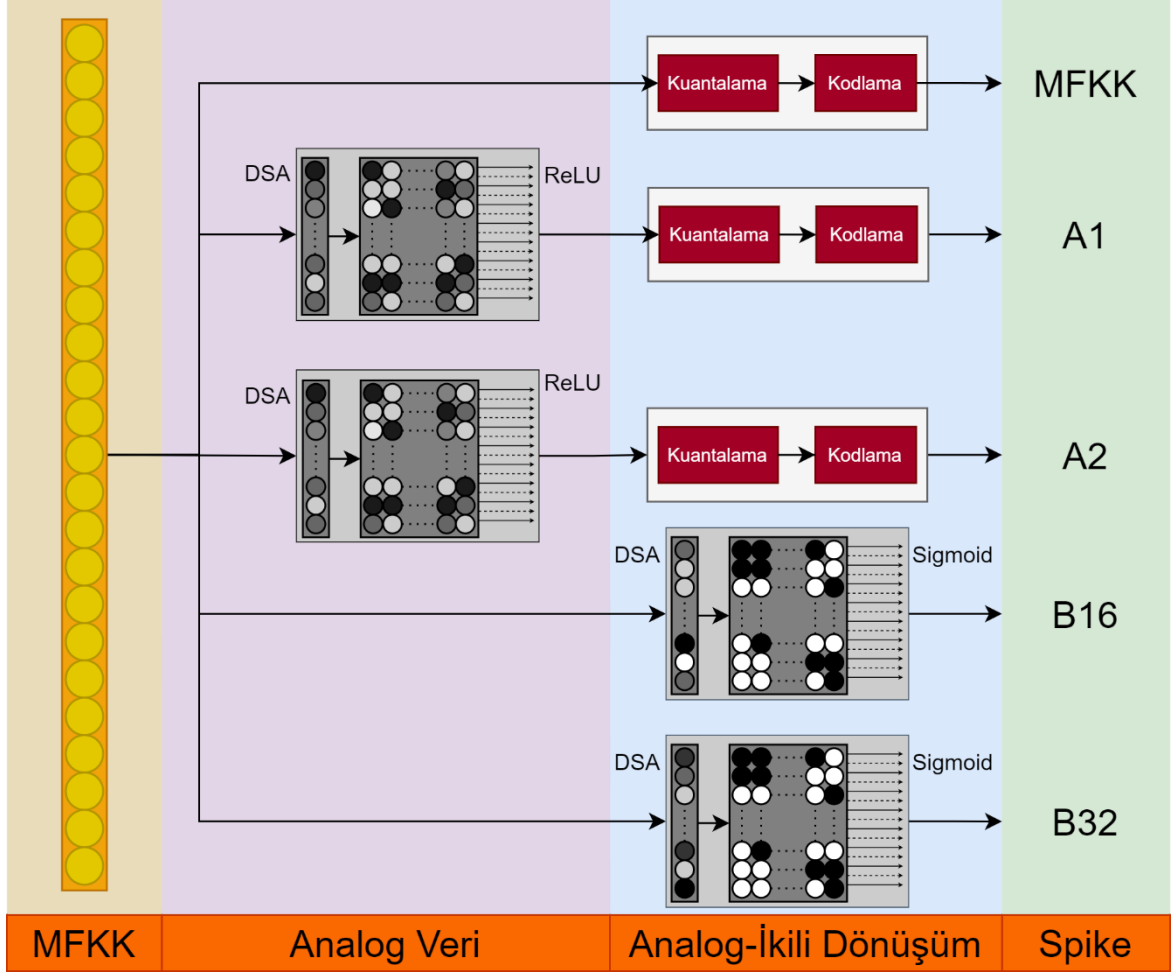
Bu deneyde, önerilen GSA modelinin OKT uygulamasında sınıflandırma performansı çeşitli öznitelik vektörleri ile incelenmiştir. Deney iki ana amaca hizmet etmektedir:

1. Önerilen GSA modelinin OKT uygulamasındaki performansının incelenmesi
2. GSA ve DSA yöntemlerinin bir arada kullanıldığı bir hibrit yaklaşımın performansının incelenmesi
3. GSA modelinin sınıf sayısı ve eğitim verisi miktarına göre performansının incelenmesi

Bu deneyde, Tablo 2 ile verilen konuşmacı tanıma VK'leri arasından VK4, VK8 ve VK16 kümeleri kullanılmıştır. VK8'den EVU değeri 5 saniye ve 10 saniye olmak üzere iki farklı eğitim kümesi oluşturulmuş, diğer VK'ler için ise EVU değeri 5 saniye olarak belirlenmiştir. Performans ölçümünde kullanılmak üzere tüm VK'ler için 1'er saniyelik test kayıtları oluşturulmuştur. Bu deneyde, ayrıca, DSA modellerinin bir geliştirme veri tabanı yardımıyla eğitilmesi söz konusudur. Şekil 36'da gösterilen DSA modellerinin eğitilmesinde Tablo 2'de verilen dev-other alt kümesi, geliştirme veri tabanı olarak kullanılmıştır. Geliştirme aşamasında konuşmacı başına 1'er saniyelik konuşma kayıtları kullanılmıştır.

Bu deneyde MFKK özniteliklerinin yanında, ayrıca, Bölüm 2.5'te tanıtılan hibrit yaklaşımdan elde edilen öznitelik vektörleri de çalışmaya dahil edilmiş ve sınıflandırma performansına etkileri incelenmiştir. Bu öznitelik vektörlerinin elde edilme akışları Şekil 36'da gösterilmektedir. İlk olarak MFKK öznitelikleri analog değerlere sahip olduğundan GSA dahilinde bulunan ÖSD yöntemi uygulanarak spike örüntülerine dönüştürülmüştür. Benzer şekilde, A1 ve A2 vektörleri, öznitelik dönüştürücü DSA modelinden elde edilmesine rağmen analog değerlere sahiptir. Bu nedenle, GSA dahilinde spike örüntülerine dönüştürülmüştür. Farklı olarak, DSA modelinden elde edilen B16 ve B32 öznitelikleri ikili formdadır ve GSA uygulamasında aynen kullanılmıştır.

Deneye başlamadan önce öznitelik dönüştürüşü DSA'ların eğitilmesi gerekmektedir. Öncelikle, Şekil 33'te verilen DSA mimarisi, A1 öznitelikleri için öznitelik katmanında bulunan YN boyutlarının 23×1 , A2 öznitelikleri için öznitelik katmanında bulunan YN



Şekil 36. D4 için kullanılan öznitelik vektörleri

boyutlarının 23×2 , B16 öznitelikleri için öznitelik katmanında bulunan YN boyutlarının 23×16 , B32 öznitelikleri için öznitelik katmanında bulunan YN boyutlarının 23×32 olacağı şekilde oluşturulmuştur. Öznitelik katmanında bulunan YN'lerin aktivasyon fonksiyonları A1 ve A2 vektörleri için ReLU, B16 ve B32 vektörleri için sigmoid olacak şekilde seçilmiştir. Sonrasında geliştirme veri tabanından elde edilen MFKK öznitelikleri ile bu modeller 200 epok boyunca eğitilmiştir. Son olarak sınıfları tespit etmekten sorumlu olan çıkış katmanı modelden ayrılmış ve DSA öznitelik dönüştürücü olarak kullanılmıştır. Tüm bu işlemler Python dilinde geliştirilmiş TensorFlow kütüphanesinde gerçekleştirilmiştir (Abadi vd., 2015).

Eğitim aşamasında, öncelikle, oluşturulan eğitim kümelerinden biri seçilmiştir. Bu eğitim kümesinde bulunan konuşma kayıtlarında detayları Bölüm 2.3'te verilen MFKK öznitelikleri çıkarılmıştır. Bu öznitelikler eğer MFKK öznitelikleri kullanılmak isteniyorsa

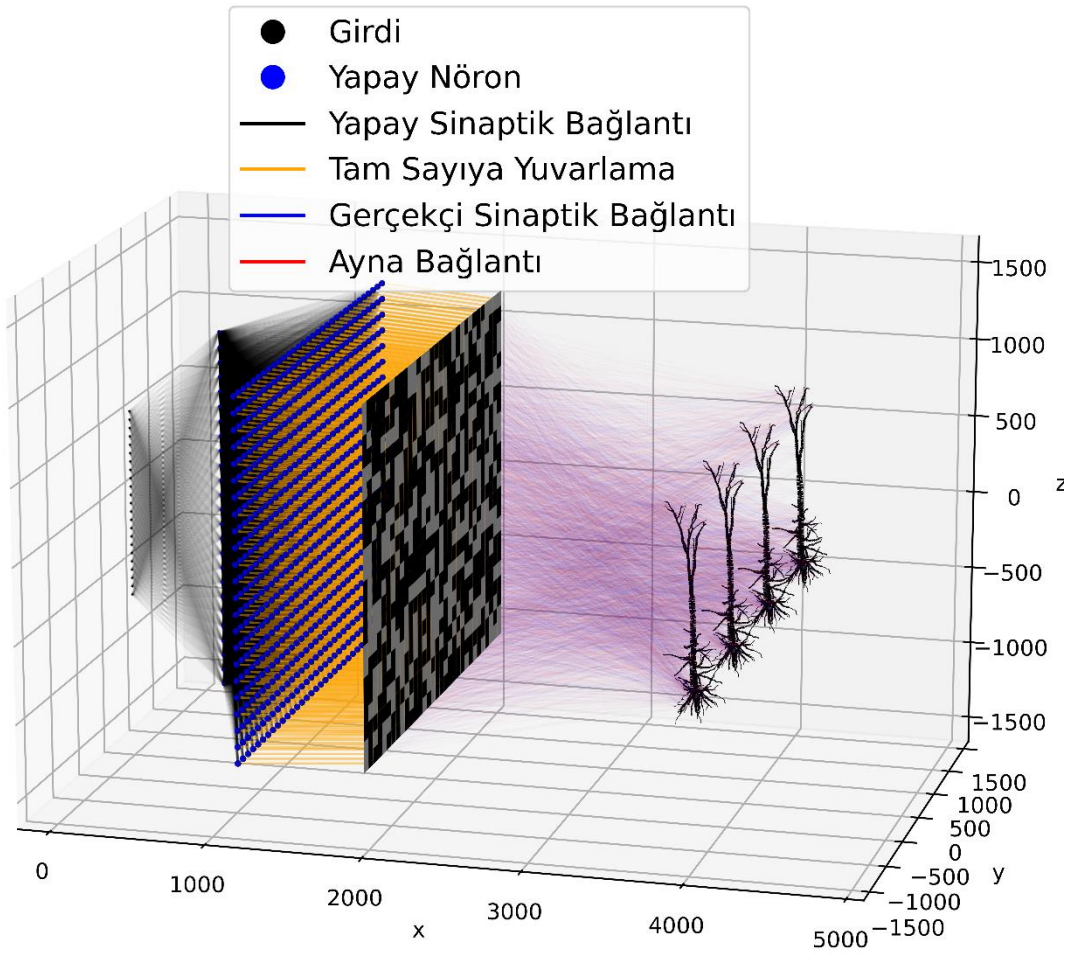
Şekil 27’de akışı verilen GSA’ya eğitim öznitelikleri olarak; eğer A1 veya A2 öznitelikleri kullanılacaksa önce geliştirilen ilgili DSA modeline giriş olarak verilmiş, çıkışları da Şekil 27’de akışı verilen GSA’ya eğitim öznitelikleri olarak verilmiştir. B16 veya B32 vektörlerinin sınıflandırma için kullanıldığı düşünülürse, MFKK öznitelikleri önce geliştirilen ilgili DSA modeline giriş olarak verilmiş, çıkışları ise Şekil 27’de akışı verilen GSA’ya eğitim spike örüntüleri olarak verilmiştir. Deneyde kullanılan GSA modelinde basit GN modelleri olan HH, IZH ve STA nöron modelleri ayrı ayrı kullanılmıştır. GSA’nın eğitimi sırasında epok sayısı 500 olarak seçilmiştir.

Test aşamasında test konuşmalarının MFKK öznitelikleri TVU değeri 1 saniye olmak üzere elde edilmiştir. Bu öznitelikler kullanılan yaklaşıma bağlı olarak, GSA’ya giriş olarak verilmiş ve etiketleri belirlenmiştir. Bu deneyde performans ölçütü olarak F1-Skor kullanılmıştır. Deney, Bölüm 2.4.7’de tanıtılan yazılım kullanılarak gerçekleştirilmiştir.

2.6.5. D5: Piramidal Hücrenin Konuşmacı Tanıma Performansının İncelenmesi

Piramidal nöronlar sahip oldukları soma, akson ve dendritlerde bulunan doğrusal olmayan kanallar sayesinde, aldıkları yoğun sinaptik girdileri karmaşık ağaç yapılarında işleyebilmektedirler (Spruston, 2008). PN’lerin çeşitli ve karmaşık morfolojileri spike işlemede olağanüstü esneklik ve çok yönlülük sağlamaktadır (Galloni vd., 2020). Bu nöronların zamanla ve komşu BN’lere göre morfolojilerini değiştirme özellikleri, PN’lerin dinamik ve karmaşık yapısını ortaya koymaktadır. Ayrıca, PN’lerin beyinde bulunan birden fazla katmandan aldıkları girdileri işleyebilmesi, bilgi işleme konusunda karmaşık doğalarını gözler önüne sermektedir.

PN’ler sinaptik girdileri hem yapılarında bulunan doğrusal olmayan iyon kanalları sayesinde hem de her biri kendine özgü karmaşık dendritik ağaç yapıları sayesinde işleyebilmektedirler. Bu nedenle, PN’lerin dendritik hesaplama yapabilme gücü araştırmaların konusu olmaktadır. Buna rağmen dendritik hesaplama konusundaki bilgiler, dendritik yapıların karmaşık yapısı ve sinaptik konumların çeşitlilik göstermesi nedeniyle sınırlı kalmaktadır (Polsky vd., 2004). Poirazi ve Mel (2001) çalışmalarında, PN’lerin morfolojik yapısı ve yerel sinaps konumlarının, PN’ye dikkate değer bir hesaplama kapasitesi sunduğunu belirtmişlerdir. Poirazi ve arkadaşları (2003) CA1 PN’lerin ancak iki katmanlı bir yapay sinir ağı (YSA) ile ifade edilebileceğini öne sürmüşlerdir. PN’lerin



Şekil 37. D5 için hibrit sinir ağı mimarisi

doğrusal olmayan özelliklerinin yansıtılabilmesi adına gizli katmanlarda doğrusal olmayan aktivasyon fonksiyonları kullanmış olsalar da bu çalışma, dendritik yapıların karmaşık hesaplama gücünün anlaşılmasında yetersiz kalmıştır. Takip eden yıllarda, çeşitli çalışmalar PN'lerin dendritik yapısının ancak çok katmanlı YSA modelleri ile ifade edilebileceğini gösteren kanıtlar sunmuşlardır (Katz vd., 2009; Ujfalussy vd., 2018). Dendritik yapılardaki öğrenme mekanizmasının henüz tam olarak aydınlatılmamış olmasına rağmen bu çalışmalar, PN'lerin birden fazla YN'nin görevlerini tek başına yapabilmesi nedeniyle, biyolojik olarak gerçekçi PN modellerinden kurulu GSA'ların daha az GN ile daha büyük YSA yapılarıyla aynı işlem gücüne sahip olduğunu göstermektedir.

Bölüm 1.1.4'te detaylandırıldığı üzere, sinaptik plastisitenin moleküler seviyede detayları nispeten anlaşılmış durumdadır. Fakat bu dinamiklerin dendritik yapıya sahip GN'lerde nasıl kullanılabileceği konusu hala süregelen bir araştırma konusudur (Legenstein vd., 2005; Gütiğ & Sompolsky, 2006). Bu motivasyonla, Legenstein ve

Maass (2011), PN'lerin doğrusal olmayan bir problemi çözmelerinde kullanılmak üzere, sinaptik plastisite ve dendritik dalların potansiyellerinin bir arada gözetildiği bir dendritik hesaplama modeli önermişlerdir. Limbacher ve Legenstein (2020), bu yaklaşımı bir adım öteye götürerek sinaptik bağlantıların sayısının artırılmasına yönelik bir öğrenme yaklaşımı önermişlerdir. Çalışmaları, ayrıca, dendritik hesaplamada sinaptik reseptörlerin konumlarının ve sayısının da önemini göstermektedir.

PN'lerin dendritik hesaplamada kullanılması birçok çalışmayı da bu odağa taşımıştır. Gidon ve arkadaşları (2020) insan beyninin 2/3 katmanında bulunan bir kortikal GN kullanarak, yalnızca çok katmanlı YSA kullanılarak çözülebilecek olan XOR problemini çözmüşlerdir. Moldwin ve Segev (2020), bir PN modelini perceptron öğrenme kuralı kullanarak eğitmiş ve PN'nin alt bölgelerinin her birinin işlem kapasitesini ortaya koymuşlardır. Kayıkçıoğlu Bozkır ve arkadaşları (2023) ise bu yaklaşımı gerçek dünya verisinde test etmiş ve ağırlık kısıtlamanın PN'nin sınıflandırma performansını düşürdüğünü göstermişlerdir. Çalışmalarında sinaptik girdilerin genişletilmesine dayanan bir yaklaşım önermişler ve bunun PN'nin performansının arttığını göstermişlerdir. Bicknell ve Häusser (2021), aktif dendritlere sahip olan bir GN modelinin doğrusal olmayan bir problemi çözebildiğini ve bunun için Gütiğ ve Sompolsky (2006) tarafından önerilen tempotron kuralını modifiye ettiklerini bildirmişlerdir.

Çalışmalar karmaşık dendritik yapıya sahip PN'lerin sınıflandırma ve hesaplama konusundaki potansiyelini ortaya koymaktadır. Literatürdeki kanıtlar göz önünde bulundurulduğunda, PN'lerin bir sınıflandırma problemleri için önemli bir gelecek vaat ettiği söylenebilmektedir. Ayrıca, sinaptik reseptörlerin dendritik bölgelerde çeşitli konumlarda ve sayılarda bulunabilmesi de PN'lerin potansiyelini daha da artırmaktadır. Bu nedenle, PN'lerin dendritik yapısının, sinaptik organizasyonunun ve bağlantı çeşitliliğinin sınıflandırma problemleri için önemi araştırılmaya değer bir konudur.

Bu deneyde, önerilen GSA modeli içinde sunulan PN'nin konuşmacı tanıma probleminde sınıflandırma performansı incelenmiştir. Deney iki ana amaca hizmet etmektedir:

1. PN'nin ve bölgelerinin ayrı ayrı konuşmacı tanıma probleminde performansının incelenmesi
2. Sinaptik bağlantı tekrarının sınıflandırma performansına etkisinin incelenmesi

Bu deneyde, Tablo 2'de verilen konuşmacı tanıma VK'leri arasından VK4 kümesi kullanılmıştır. VK4'ten, EVU değeri 5 saniye olan bir eğitim kümesi oluşturulmuştur.

Performans ölçümünde kullanılmak üzere VK4 için TVU değerleri 1 saniye olacak şekilde test kümeleri oluşturulmuştur. D4'te kullanılan B32 vektörlerinin en başarılı öznitelikler olduğu göz önünde bulundurularak bu deneyde sadece B32 vektörleri kullanılmıştır. Bu vektörlerin elde edilmesinde D4'te oluşturulan B32 öznitelik dönüştürücü DSA kullanılmıştır.

Deneyde kullanılan PN modeli; Tablo 5'te görüleceği üzere; somatik, apikal, basilar ve aksonal olmak üzere dört ana bölgeden oluşmaktadır. Bu bölgeler sadece soma, sadece apikal, sadece basilar, tüm dendritler ve tüm bölgeler olmak üzere toplam beş adet gruba ayrılmıştır. Aksonal bölge, Hay ve arkadaşlarının (2011) çalışmalarında morfolojik ve biyofiziksel olarak detaylı olarak modellenmediği için ihmal edilmiş ve sinaptik bağlantılardan muaf tutulmuştur.

Gruplar somaya uzaklık ve biyofiziksel detaylar olmak üzere iki özellik bakımından birbirlerinden ayrılmaktadır. Sadece soma olarak isimlendirilen grup hem aktif hem pasif iyon kanalları içermektedir. Sadece apikal olarak isimlendirilen grup somaya oldukça uzaktır. Bunun yanında içeriğinde hem aktif hem pasif iyon kanalları bulundurmasına rağmen kendi başına AP üretecek kanallara sahip değildir. Basilar grup somaya apikal bölgeye göre nispeten daha yakın olsa da içeriğinde aktif kanal bulundurmamaktadır. Tüm dendritler olarak isimlendirilen grup apikal ve basilar bölgelerinin birlikte düşünüldüğü gruptur ve dendritik aktivitenin bütüncül olarak incelenebilmesi adına deneye dahil edilmiştir. Tüm bölgeler ise somatik, apikal ve basilar bölgelerin birleşimini ve PN'ye bütüncül bir yaklaşımı ifade etmektedir.

Deney, VK4 eğitim kümesinden MFKK özniteliklerinin hesaplanmasıyla başlamaktadır. Bu öznitelikler B32 vektörlerinin elde edilmesi amacıyla ilgili öznitelik dönüştürücü DSA modeline giriş olarak verilmiş ve çıkışları Şekil 27'de gösterilen akışta spike örüntüleri olarak kullanılmıştır. GSA modelinin tasarlanması aşamasında geri kalan tüm seçenekler varsayılan olarak belirlenmek kaydıyla sinaptik bağlantı tekrar sayısı 1, 2, 4 ve 8 olacak şekilde değiştirilmiş ve deney bu parametreler için ayrı ayrı gerçekleştirilmiştir. Elde edilen B32 vektörleri ile GSA modeli eğitilmiş ve eğitim modeli elde edilmiştir. PN'lerin işlem yükünün ağır olması nedeniyle, eğitim süresinin kabul edilebilen bir zaman aralığında bitmesi adına, tüm modeller epok sayısı 50 olacak şekilde eğitilmiştir. Test aşamasında, TVU değerleri sırasıyla 1 saniye olan konuşma kayıtları, eğitilen model ile etiketlenmiştir. Performans ölçümünde F1-Skor ölçütü kullanılmıştır. Tüm GSA modelleri önerilen yazılım aracılığıyla oluşturulmuş, eğitilmiş ve test edilmiştir.

3. BULGULAR VE TARTIŞMA

Şekil 21’de verilen yöntem akışı incelendiğinde, bu tez çalışmasında önerilen yaklaşımların ardışık aşamalara sahip olduğu görülebilmektedir. Bu aşamaların her biri sistem performansını birbirinden bağımsız olarak etkileyen faktörlerdir. Bu nedenle, sistem performansı etkileyen faktörlerin öncelikle ne derece etkili oldukları belirlenmeli ve performansları ortaya koyulmalıdır. Sonrasında sistem aşamalarında seçilmesi gereken parametreler belirlenerek sistem performansı incelenmelidir. Böylece hem probleme uygun parametreler belirlenmiş hem de sistem performansı maksimize edilmiş olacaktır. Ayrıca, elde edilen bulgular karşılaştırmalı olarak performans ölçütlerinin başarısını da ortaya koyacaktır.

Bu tez çalışmasında, D1, D2, D3, D4 ve D5 olarak isimlendirilmiş olan toplam beş deneysel senaryo, önerilen yöntemin performansının aşamalar halinde ortaya koyulması amacıyla tasarlanmıştır. Her senaryoda, belirlenen bir aşamaya odaklanılmış ve elde edilen bulgular bu bölüm altında sunulmuştur.

D1’de, tez çalışmasında kullanılan veri kümesinin (VK’nin) sınıflandırma performansı, literatürde önerilen en gelişkin otomatik konuşmacı tanıma (OKT) yöntemleriyle, çeşitli faktörlerin etkisi altında incelenmiş ve tez çalışmasında kullanılacak olan parametreler bu deney bulgularından yararlanılarak seçilmiştir. Ayrıca, deney, seçilen parametreler kullanılarak tekrar edilmiş ve elde edilen performans bulguları sonraki deneyler için referans olarak kabul edilmiştir.

D2’de, önerilen öznitelik spike dönüşümü (ÖSD) aşamasında seçilen parametrelerin performansı nasıl etkilediği klasik makine öğrenimi (MÖ) yöntemleri kullanılarak incelenmiştir. Buna göre kodlama yöntemlerinin, kuantalama tekniklerinin ve bit uzunluğunun performans üzerinde ne derece etkili oldukları 3 farklı VK ile karşılaştırmalı olarak incelenmiştir. Deney sonucunda elde edilen bulgulardan yararlanılarak sonraki deneylerde hangi parametrelerin kullanılacağı belirlenmiştir.

D3’te, önerilen gerçekçi sinir ağı (GSA) modelinde kullanılan sinaptik bağlantı (SB) tiplerinin performansı ne derece etkilediği 3 farklı VK ile karşılaştırmalı olarak incelenmiştir. Buna göre aynalama yaklaşımının, excitatory (Exc) ve inhibitory (Inh) SB’lerin performans etkisi ortaya koyulmuştur. Deney sonucunda elde edilen bulgulara göre sonraki deneylerde hangi SB tiplerinin dahil edileceği belirlenmiştir.

D4'te; D1, D2 ve D3 senaryolarından elde edilen bulgulara göre konuşma VK, ÖSD ve GSA'lara dair parametreler belirlenmiş ve sistem performansı incelenmiştir. Ayrıca, hibrit yaklaşım sonucu elde edilen öznitelikler de deneye dahil edilerek performansın nasıl etkilendiği incelenmiştir. Bu deneyde elde edilen bulgulara dayanılarak, en yüksek performansı sağlayan öznitelikler sonraki deneyde kullanılmıştır.

D5'te, piramidal nöronların (PN'lerin) OKT performansları incelenmiştir. PN'lerin ağır işlemsel yükü göz önünde bulundurularak, önceki deneylerden elde edilen bulgular ışığında en yüksek performansı sağlayan parametre ve seçimler kullanılmıştır. Sonuç olarak, Şekil 21'de verilen yöntem akışının performansının karşılaştırmalı olarak incelenebilmesi mümkün olmuştur.

Bu bölümde, sıralanan deneysel senaryolar sonucu elde edilen bulgular sistematik olarak sunulmuştur. Bölüm, dört alt başlıktan meydana gelmektedir. İlk alt başlıkta, en gelişkin OKT yöntemlerinin performanslarına dair bulgular verilmiş ve tartışılmıştır. İkinci alt başlıkta, önerilen ÖSD aşamasının ve yöntemde sunulan seçeneklerin performansı nasıl etkilendiği, klasik MÖ yöntemleri kullanılarak incelenmiş ve tartışılmıştır. Üçüncü alt başlıkta, önerilen GSA modelinde dahil edilen SB tiplerinin performansa etkisi karşılaştırmalı olarak incelenmiş ve tartışılmıştır. Beşinci ve son alt başlıkta, OKT uygulamasında GSA'ların performansına dair bulgular detaylı olarak verilmiş ve tartışılmıştır.

3.1. Otomatik Konuşmacı Tanımaya Dair Öncül Bulgular

Bu bölümde, Bölüm 2.6.1'de detayları verilen D1 isimli deneysel senaryonun ilk adımına dair bulgulara odaklanılmıştır. D1'de, literatürde yer alan en gelişkin OKT yaklaşımlarının performanslarını çeşitli faktörler ve parametreler altında incelenmesi amaçlanmaktadır. Bu inceleme, Gauss karışım modeli (GKM) tabanlı, derin öğrenme (DÖ) tabanlı ve vektör tabanlı yaklaşımlar olmak üzere üç ana gruba ayrılmıştır.

GKM tabanlı yaklaşımlar içerisinde üç ana yöntem bulunmaktadır. Bunlar; GKM yaklaşımı (Reynolds, 1995), Gauss karışım modeli-evrensel arka plan modeli (GKM-EAPM) yaklaşımı (Reynolds vd., 2000) ve Gauss karışım modeli-destek vektör makineleri (GKM-DVM) yaklaşımı (Campbell vd., 2006) olarak sıralanmaktadır. DÖ tabanlı yaklaşımlar ise iki farklı yöntemi içermektedir. Bunlar; derin sinir ağları (DSA) ve transfer öğrenme (TÖ) ile güçlendirilmiş DSA olarak sıralanmaktadır (Özcan & Kayıkçıoğlu,

2021). Vektör tabanlı yaklaşımlar, iki farklı yöntemi içermektedir. Bunlar; x-vektörler (Snyder vd., 2018) ve d-vektörler (Variani vd., 2014) olarak sıralanmaktadır.

D1 süresince, 1296 farklı parametre ve yöntem dizisini içeren deneyler 50'şer kez tekrarlanmış, sonuç olarak toplamda 64800 denemeye ilişkin veriler toplanmıştır. Her bir denemeye ait bulgular için sınıf içi F1-Skorları hesaplanmış ve bir denemenin performans ölçütü, bu F1-Skorlarının ortalaması alınarak belirlenmiştir. Aynı parametre ve yöntem dizisine sahip denemelerin performans ölçütlerinin ortalamaları ve standart sapmaları hesaplanmış ve bu kapsamda elde edilen tüm bulgular EK-2'de tablolar halinde sunulmuştur. EK-2 altında verilen Ek Tablo 1'den Ek Tablo 16'ya kadar olan tablolar, D1 deneyinin ilk aşamasında elde edilen tüm ortalama ve standart sapma değerlerini içermektedir.

Tüm bulgular değerlendirildiğinde, denemelerin ortalama performansı 81.9 ± 13.7 olarak hesaplanmıştır. Bu ortalama değer, minimum %21.7 ve maksimum %99.1 arasında değişen 1296 farklı parametre ve yöntem dizisinin performanslarını temsil etmektedir. Ayrıntılı bir analiz yapıldığında, sonuçların kullanılan yöntem, yöntem parametreleri, konuşmacı sayısı, eğitim verisi uzunluğu (EVU), test verisi uzunluğu (TVU) ve MFKK sayısı gibi faktörlere bağlı olarak değiştiği görülmektedir. Bu nedenle, bulguların bu parametreler ve yöntemler çerçevesinde ayrı ayrı incelenmesi, daha sağlıklı ve doğru sonuçlara ulaşılmasını sağlayacaktır.

D1 sonucunda elde edilen bulguların çok sayıda ve çok boyutlu olması, bu bulguların görselleştirilmesini zorlaştırmaktadır. Bu nedenle, ilerleyen deneylerle daha ilişkili olan bulgular "daraltılmış bulgular" olarak adlandırılmış, ayrıca incelenmiş ve görselleştirilmiştir. OKT uygulaması yapılan sonraki deneylerde, EVU değerinin 5 saniye, TVU değerinin 1 saniye ve MFKK sayısının 23 olarak seçildiği dikkate alınarak, bu bölümde EVU ve TVU değerlerinin sırasıyla 5 ve 1 saniye, MFKK sayısının ise 20 olduğu denemelere ait bulgular referans bulgular olarak kabul edilmiştir.

Daraltılmış bulgular, kutu grafiği tekniği ile görselleştirilmiştir. Bu grafiklerde, aynı koşullar altında 50 kez tekrarlanan gözlemlerin performans bulguları sınıf içi ortalama F1-Skor ölçütü cinsinden gösterilmektedir. Kutu grafikleri, gözlem dağılımlarına dair beş temel özelliği özetleyen grafiklerdir. Her kutu grafiğinde, alt dikey çizginin alt sınırı gözlemlerin minimum değerini, üst dikey çizginin üst sınırı ise maksimum değerini göstermektedir. Gözlemler aritmetik olarak sıralandığında, alt dikey çizgi gözlemlerin ilk çeyreğini, üst dikey çizgi ise son çeyreğini temsil etmektedir. Merkezdeki renkli kutu,

gözlemlerin ikinci ve üçüncü çeyreklerini içerirken, kutunun ortasındaki yatay çizgi ortanca değeri göstermektedir. Son olarak, kutu grafiklerinin üst ve alt kısımlarında baklava dilimi şeklinde işaretlenen noktalar, aşırı uç gözlemleri belirtmektedir.

Bu bölüm, beş alt başlıktan meydana gelmektedir. İlk alt başlıkta, en gelişkin OKT yöntemlerinin performansları karşılaştırmalı olarak incelenmiştir. İkinci, üçüncü ve dördüncü alt başlıklarda sırasıyla EVU değerinin, TVU değerinin ve MFKK sayısının OKT performansını nasıl etkilediği araştırılmıştır. Beşinci ve son alt başlıkta ise tüm bulgular bir arada değerlendirilerek tartışılmıştır.

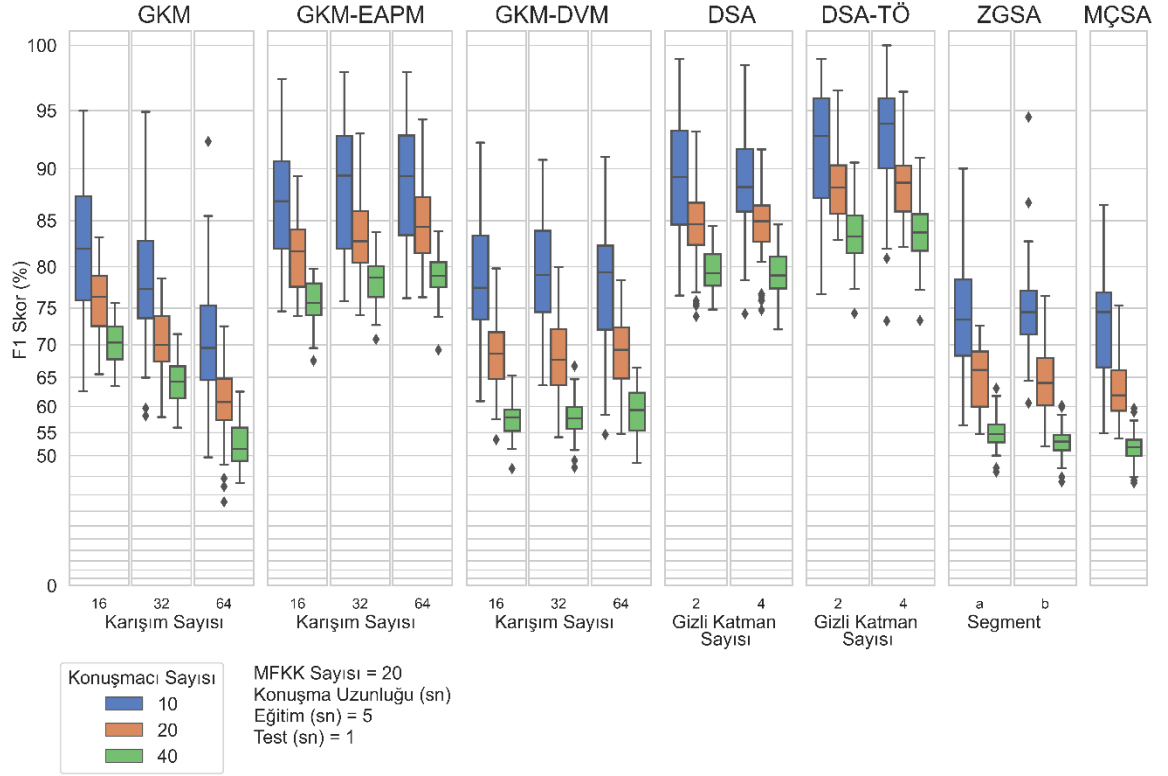
3.1.1. Otomatik Konuşmacı Tanımda En Gelişkin Yöntemlerin Performansı ve Parametrelerin Etkisi

Tüm bulgular değerlendirildiğinde, GKM tabanlı yaklaşımlardan ilki olan GKM yaklaşımının ortalama performansı $\%83.4 \pm 12.2$ olarak hesaplanmıştır. Diğer GKM tabanlı yöntemlerden GKM-EAPM için bu değer $\%87.7 \pm 8.2$, GKM-DVM için ise $\%79.4 \pm 12.5$ olarak belirlenmiştir. Daha ayrıntılı bir inceleme, bu performans ölçütlerinin yöntem parametrelerine bağlı olarak değiştiğini göstermektedir.

GKM yaklaşımı için, karışım sayısı 16 olduğunda ortalama performans $\%85.9 \pm 9.6$, 32 olduğunda $\%84.2 \pm 11.4$ ve 64 olduğunda $\%80.1 \pm 14.6$ olarak hesaplanmıştır. GKM-EAPM yaklaşımında, karışım sayısı 16 iken ortalama performans $\%86.3 \pm 8.6$, 32 iken $\%88.0 \pm 8.0$ ve 64 iken $\%88.8 \pm 7.7$ olarak belirlenmiştir. GKM-DVM yaklaşımında ise, karışım sayısı 16 olduğunda ortalama performans $\%78.0 \pm 12.9$, 32 olduğunda $\%79.4 \pm 12.3$ ve 64 olduğunda $\%80.6 \pm 12.2$ olarak hesaplanmıştır.

DÖ tabanlı yaklaşımlar incelendiğinde, DSA yönteminin ortalama performansı $\%87.2 \pm 10.3$, DSA-TÖ yönteminin ise $\%87.4 \pm 10.0$ olarak hesaplanmıştır. Ayrıntılı bir analiz, bu performans ölçütlerinin de yöntem parametrelerine bağlı olarak değiştiğini ortaya koymaktadır.

DSA ve DSA-TÖ yaklaşımları incelendiğinde, gizli katman sayısı 2 olarak seçildiğinde DSA yaklaşımında elde edilen ortalama performans $\%87.4 \pm 10.2$ olarak hesaplanmıştır. Aynı şekilde, gizli katman sayısı 4 olduğunda ortalama performans $\%87.1 \pm 10.6$ olarak ölçülmüştür. DSA-TÖ yaklaşımında ise, gizli katman sayısı 2 olduğunda ortalama performans $\%87.5 \pm 10.0$ olarak belirlenirken, bu sayı 4 olduğunda ortalama performans $\%87.4 \pm 10.1$ olarak tespit edilmiştir.



Şekil 38. En gelişkin yöntemlerin OKT performansları

Vektör tabanlı yaklaşımlar değerlendirildiğinde, x-vektörler yönteminin ortalama performansı 68.0 ± 18.2 , d-vektörler yönteminin ise 73.9 ± 11.5 olarak hesaplanmıştır. Detaylı bir incelemeyle, x-vektörlerin "a" segmenti kullanıldığında ortalama performansı 67.6 ± 18.4 olarak belirlenirken, "b" segmenti kullanıldığında bu performans 68.4 ± 18.1 olarak hesaplanmıştır.

Şekil 38, aynı yöntem ve parametrelerin performans üzerindeki etkisini kutu grafiği olarak sunmaktadır. Bu grafik, EVU değerinin 5 saniye, TVU değerinin 1 saniye ve MFKK sayısının 20 olduğu deneye ait bulguları içermektedir. Grafikte, performansın konuşmacı sayısına olan etkisini gözlemlemek amacıyla konuşmacı sayıları 10, 20 ve 40 olarak seçilmiş ve bu gözlemler farklı kutu renkleriyle temsil edilmiştir.

Daraltılmış bulgulara göre, farklı yaklaşımların performanslarına ilişkin çeşitli bulgular elde edilmiştir. GKM yaklaşımı ortalama 69.0 ± 9.2 performans sergilerken, GKM-EAPM yaklaşımı 82.6 ± 4.4 ve DSA yaklaşımı 84.1 ± 4.3 ile daha yüksek bir performans göstermiştir. En yüksek performans DSA-TÖ yaklaşımıyla elde edilmiştir. Buna göre, 87.9 ± 3.7 ile diğer yöntemlerden belirgin şekilde ayrılmıştır. GKM-DVM yaklaşımı ise 68.1 ± 8.6 ortalama performans değeriyle diğer yöntemlere göre daha

düşük bir performans göstermiştir. Öte yandan, x-vektörler ve d-vektörler yaklaşımlarının performansları ise sırasıyla 64.1 ± 9.1 ve 62.1 ± 10.3 olmak üzere en düşük değerlere sahiptir. Bu sonuçlar, farklı yaklaşımların konuşmacı tanıma performansı üzerindeki etkisini açıkça göstermektedir.

Şekil 38’de sunulan daraltılmış bulgular, kullanılan farklı parametrelerin konuşmacı tanıma performansı üzerindeki etkisini açıkça göstermektedir. Bu bulgulara göre, GKM yaklaşımının 16, 32 ve 64 olmak üzere üç farklı karışım sayısı değerinde ortalama performansı sırasıyla 75.8 ± 5.8 , 70.8 ± 6.9 ve 60.6 ± 8.8 olarak belirlenmiştir. Benzer şekilde, GKM-EAPM yaklaşımı da farklı parametrelerle incelenmiş ve karışım sayısı 16, 32 ve 64 olduğu durumlarda sırasıyla 81.1 ± 5.3 , 83.1 ± 4.8 ve 83.7 ± 4.6 ortalama performans değerleri elde edilmiştir. GKM-DVM yaklaşımının ise karışım sayısının 16, 32 ve 64 olarak seçildiği durumlarda ortalama performansı 68.0 ± 10.0 , 68.1 ± 10.5 ve 68.2 ± 9.2 olarak hesaplanmıştır.

DSA yaklaşımı için gizli katman sayısının 2 ve 4 olduğu durumlarda sırasıyla 84.2 ± 4.8 ve 84.0 ± 4.9 ortalama performans değerleri elde edilmiştir. DSA-TÖ yaklaşımı ise 2 ve 4 gizli katman sayısı için sırasıyla 87.9 ± 4.1 ve 88.0 ± 4.3 ortalama performans göstermiştir. Son olarak, x-vektörler yaklaşımı “a” ve “b” segmentlerinde sırasıyla 64.4 ± 9.4 ve 63.7 ± 10.9 ortalama performans değerlerine sahiptir. Bu bulgular, kullanılan farklı parametrelerin ve yaklaşımların konuşmacı tanıma sistemlerinin performansına olan etkisini ayrıntılı bir şekilde açıklamaktadır.

Genel olarak tüm bulgular incelendiğinde, konuşmacı sayısının artmasıyla performansta bir düşüş yaşandığı gözlemlenmektedir. Konuşmacı sayısının iki katına çıkması durumunda, ortalama %5.5’lik bir performans düşüşü yaşanmıştır. Dört katına çıkması durumunda ise, ortalama %11.7’lik bir performans düşüşü gözlenmiştir.

Kullanılan yöntemlere bağlı olarak, konuşmacı sayısının iki katına çıkmasıyla performansta farklı düşüşler gözlemlenmiştir. GKM yaklaşımında ortalama %5.0, GKM-EAPM yaklaşımında ortalama %4.1, GKM-DVM yaklaşımında ortalama %6.8, DSA yaklaşımında ortalama %4.3, DSA-TÖ yaklaşımında ortalama %4.1, x-vektörler yaklaşımında ortalama %8.4 ve d-vektörler yaklaşımında ise ortalama %7.9 performans düşüşü yaşanmıştır.

Konuşmacı sayısının dört katına çıkmasıyla ise, GKM yaklaşımında ortalama %9.5, GKM-EAPM yaklaşımında ortalama %8.2, GKM-DVM yaklaşımında ortalama %14.0, DSA yaklaşımında ortalama %9.8, DSA-TÖ yaklaşımında ortalama %10.6, x-vektörler

yaklaşımında ortalama %18.0 ve d-vektörler yaklaşımında ise ortalama %16.7 performans düşüşü gözlenmiştir. Bu bulgular, kullanılan yöntemlerin konuşmacı sayısındaki artışa nasıl tepki verdiğini göstermektedir.

Şekil 38’de sunulan daraltılmış bulgular, konuşmacı sayısının iki katına çıkmasıyla birlikte kullanılan farklı yöntemlerde performansta meydana gelen değişiklikleri detaylı bir şekilde göstermektedir. GKM yaklaşımında, 16 karışım kullanıldığında ortalama performans düşüşü %6.0, 32 karışım kullanıldığında %7.6, 64 karışım kullanıldığında ise %9.9 olmuştur. GKM-EAPM yaklaşımında ise 16 karışım kullanıldığında ortalama performans düşüşü %5.1, 32 karışım kullanıldığında %4.5, 64 karışım kullanıldığında ise %3.7 olarak belirlenmiştir. GKM-DVM yaklaşımında konuşmacı sayısına olan duyarlılığın arttığı görülmekle birlikte, 16 karışım kullanıldığında ortalama performans düşüşü %9.2, 32 karışım kullanıldığında %10.8, 64 karışım kullanıldığında ise %8.8 olarak hesaplanmıştır.

DSA yaklaşımında, 2 gizli katman kullanıldığında ortalama performans düşüşü %4.9, 4 gizli katman kullanıldığında ise %4.2 olarak gözlenmiştir. DSA-TÖ yaklaşımında ise 2 gizli katman kullanıldığında ortalama performans düşüşü %3.3, 4 gizli katman kullanıldığında ise %3.6 olarak saptanmıştır.

Öte yandan, x-vektörler yaklaşımında “a” segmenti kullanıldığında ortalama performans düşüşü %8.8, “b” segmenti kullanıldığında ise %10.5 olarak belirlenmiştir. Son olarak, d-vektörler yaklaşımı için ortalama performans düşüşü %9.5 olarak hesaplanmıştır.

Konuşmacı sayısının dört katına çıkmasıyla birlikte, farklı yöntemlerdeki performans değişiklikleri oldukça çeşitlenmiştir. Örneğin, GKM yaklaşımında 16 karışım kullanıldığında ortalama performans düşüşü %11.6 olarak belirlenmiştir. Aynı yöntemde 32 karışım kullanımında bu düşüş %13.8’e yükselmiş, 64 karışım kullanımında ise %17.4’e çıkmıştır. GKM-EAPM yaklaşımında, 16 karışım kullanımında ortalama performans düşüşü %10.6, 32 karışım kullanımında %9.6 ve 64 karışım kullanımında ise %9.1 olarak hesaplanmıştır. GKM-DVM yaklaşımında ise, 16 karışım kullanımında ortalama %20.1, 32 karışım kullanımında %21.0 ve 64 karışım kullanımında ise %18.4 performans düşüşü gözlenmiştir.

DSA ve DSA-TÖ yaklaşımlarında da benzer bir eğilim gözlenmiş olup, her iki yöntemde de konuşmacı sayısının artmasıyla performans düşüşü yaşanmıştır. Örneğin, DSA yaklaşımında 2 ve 4 gizli katman kullanımlarında sırasıyla %9.6 ve %9.7’lik performans düşüşleri tespit edilmiştir. DSA-TÖ yaklaşımında da 2 gizli katman

kullanımında ortalama performans düşüşü %8.1, 4 gizli katman kullanımında ise %8.5 olarak belirlenmiştir.

Öte yandan, x-vektörler yaklaşımında, “a” ve “b” segmentlerinin kullanımlarında performans düşüşü sırasıyla %18.9 ve %21.7 olarak ölçülmüştür. Son olarak, d-vektörler yaklaşımında ise ortalama %20.5'lik bir performans düşüşü gözlenmiştir. Bu bulgular, konuşmacı sayısındaki artışın farklı yöntemlerde nasıl bir performans değişikliğine neden olduğunu ayrıntılı bir şekilde açıklamaktadır.

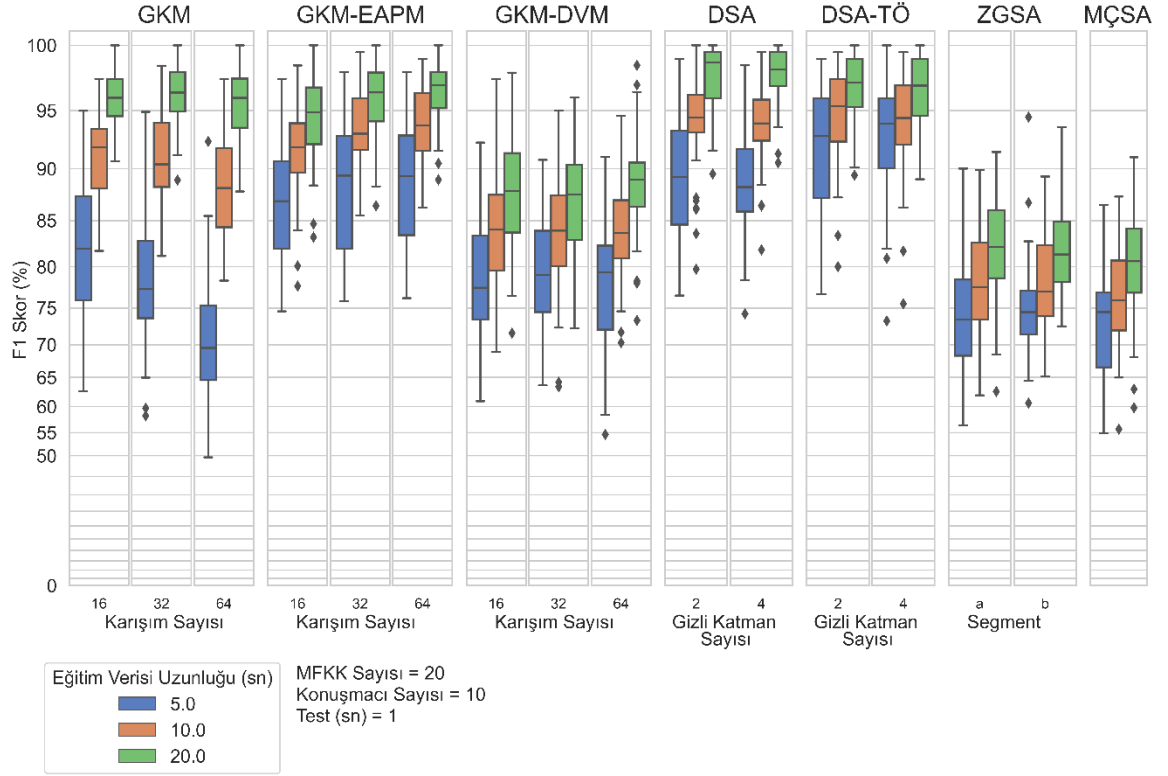
3.1.2. Eğitim Verisi Uzunluğunun Performansa Etkisi

Genel olarak tüm bulgular incelendiğinde, eğitim verisinin uzunluğunun artmasıyla birlikte performansta bir artış gözlemlenmektedir. EVU değerinin iki katına çıkması durumunda, yaklaşımlarda ortalama %7.4'lük bir performans artışı yaşanmıştır. Dört katına çıkması durumunda ise, ortalama %11.9'luk bir performans artışı gözlenmiştir. Bu bulgular, EVU değerinin konuşmacı tanıma performansı üzerinde önemli bir faktör olduğunu açıkça ortaya koymaktadır.

Kullanılan farklı yöntemlere göre, EVU değerinin iki katına çıkmasıyla performansta farklı artışlar gözlemlenmiştir. Örneğin, GKM yaklaşımında ortalama performans %15.5 artış gösterirken, GKM-EAPM yaklaşımında bu artış %5.0 olmuştur. GKM-DVM yaklaşımında ise ortalama performans %7.2 artmıştır. DSA ve DSA-TÖ yaklaşımlarında ise, EVU değerinin iki katına çıkmasıyla sırasıyla %6.0 ve %4.0'lık performans artışları tespit edilmiştir. Son olarak, x-vektörler yaklaşımında performans artışı %5.6, d-vektörler yaklaşımında ise %4.4 olarak ölçülmüştür.

EVU değerinin dört katına çıkmasıyla birlikte, farklı yöntemlerde performansta gözlenen artışlar çeşitlenmiştir. GKM yaklaşımında ortalama performans %22.5 artarken, GKM-EAPM yaklaşımında bu artış %8.4 olarak belirlenmiştir. GKM-DVM yaklaşımında ise, ortalama performans %11.9 artmıştır. DSA ve DSA-TÖ yaklaşımlarında, EVU değerinin dört katına çıkmasıyla sırasıyla %10.2 ve %6.9'luk performans artışları gözlenmiştir. Son olarak, x-vektörler yaklaşımında performans artışı %9.8, d-vektörler yaklaşımında ise %7.8 olarak ölçülmüştür. Bu bulgular, farklı yöntemlerin eğitim verisinin uzunluğundaki artışa nasıl tepki verdiğini detaylı bir şekilde göstermektedir.

Şekil 39'da çeşitli yöntem parametrelerinin performans üzerindeki etkisi kutu grafiği ile gösterilmektedir. Bu grafiğin oluşturulmasında, eğitim verisi uzunluğunun performansa



Şekil 39. Eğitim verisi uzunluğunun OKT performansına etkisi

etkisinin anlaşılabilmesi için EVU değerinin 5, 10 ve 20 saniye olarak seçildiği gözlemler, farklı kutu renkleriyle temsil edilmiştir. Bu şekilde, en gelişkin yöntemlerin performansı ve EVU değerinin etkisi arasındaki ilişki net bir şekilde görülebilmektedir.

Daraltılmış bulgulara göre, EVU değerinin iki katına çıkmasıyla, yöntemlerde performans artışları gözlenmiştir. Örneğin, GKM (16) yaklaşımında EVU değerinin iki katına çıkmasıyla ortalama performans %9.2 artmıştır. Benzer şekilde, GKM (32) yaklaşımında bu artış %12.6'ya yükselmiş, GKM (64) yaklaşımında ise artış %18,0 olarak ölçülmüştür. GKM-EAPM (16) ve GKM-EAPM (32) yaklaşımlarında sırasıyla %5.2 ve %5.7'lik artışlar gözlenirken, GKM-EAPM (64) yaklaşımında bu artış %5.7'ye ulaşmıştır. GKM-DVM (16) yaklaşımında ortalama %5.8, GKM-DVM (32) yaklaşımında ortalama %4.8, ve GKM-DVM (64) yaklaşımında ortalama %6.9'luk performans artışları kaydedilmiştir.

DSA (2) ve DSA (4) yaklaşımlarında sırasıyla %4.9 ve %5.3'lük performans artışları tespit edilmiştir. DSA-TÖ (2) ve DSA-TÖ (4) yaklaşımlarında ise, sırasıyla %2.7 ve %1.9'luk artışlar gözlenmiştir. Öte yandan, x-vektörler (a) yaklaşımında ortalama

performans %4.3 artarken, x-vektörler (b) yaklaşımında %3,6'lık bir artış belirlenmiştir. Son olarak, d-vektörler yaklaşımında ise, ortalama performans %3.7 artmıştır.

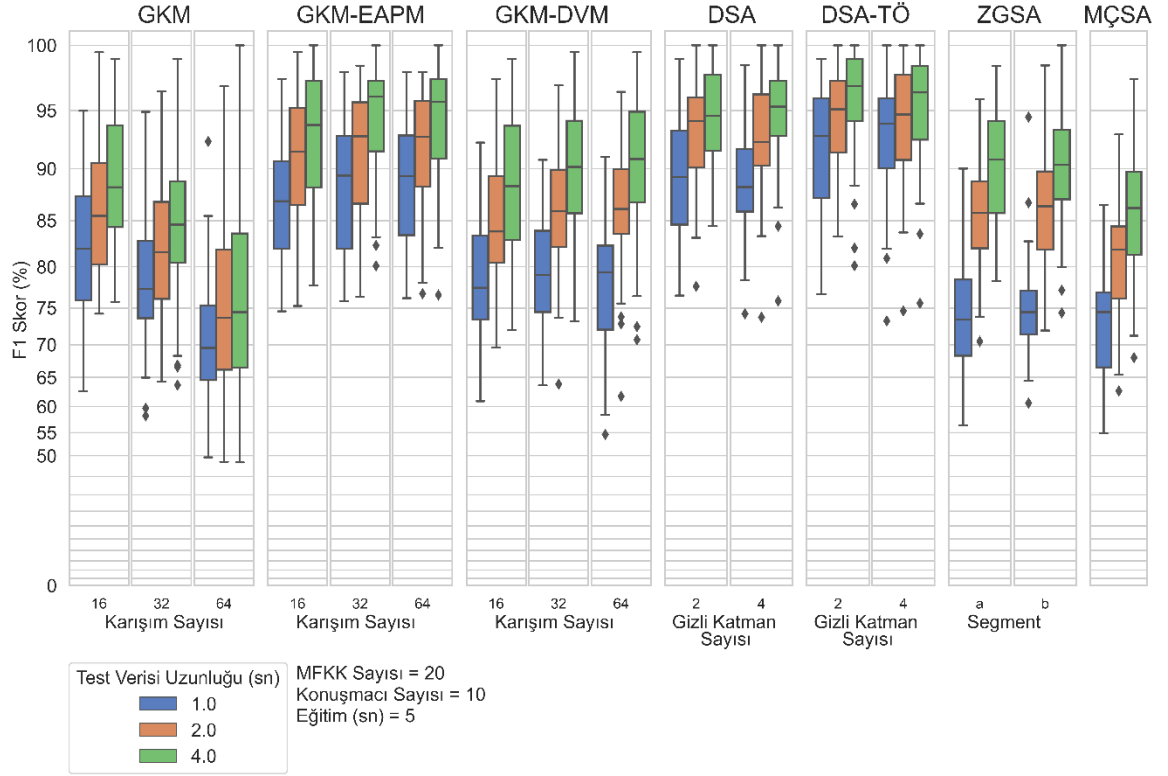
EVU değerinin dört katına çıkmasıyla, performansta benzer artışlar gözlemlenmiştir. Örneğin, GKM (16) yaklaşımında bu artış %14.2 olarak ölçülmüştür. Benzer şekilde, GKM (32) yaklaşımında bu artış %18.2'ye yükselmiş, GKM (64) yaklaşımında ise %25.7 seviyesine ulaşmıştır. GKM-EAPM (16) ve GKM-EAPM (32) yaklaşımlarında sırasıyla %7.9 ve %7.8'lik artışlar kaydedilmiştir. GKM-EAPM (64) yaklaşımında ise bu artış %8.5'e çıkmıştır. GKM-DVM (16) yaklaşımında ortalama %9.4, GKM-DVM (32) yaklaşımında ortalama %7.7 ve GKM-DVM (64) yaklaşımında ortalama %11.1'lik performans artışları gözlenmiştir.

DSA (2) ve DSA (4) yaklaşımlarında sırasıyla %8.5 ve %9.1'lik performans artışları tespit edilmiştir. DSA-TÖ (2) ve DSA-TÖ (4) yaklaşımlarında ise, sırasıyla %5.2 ve %4.7'lik artışlar gözlenmiştir. Öte yandan, x-vektörler (a) yaklaşımında ortalama performans %8.2 artarken, x-vektörler (b) yaklaşımında %7.3'lük bir artış belirlenmiştir. Son olarak, d-vektörler yaklaşımında ise, ortalama performans %7.9 artmıştır. Bu bulgular, yaklaşımların EVU değerindeki değişikliklere karşı performanslarındaki değişimi vurgulamaktadır.

3.1.3. Test Verisi Uzunluğunun Performansa Etkisi

Tüm bulgular göz önüne alındığında, test verisi uzunluğunun artmasıyla performansın arttığı tespit edilmiştir. Test verisinin uzunluğunun iki katına çıkmasıyla ortalama performansta %8.1'lik bir artış gözlemlenmiştir. Bu uzunluğun dört katına çıkmasıyla ise ortalama performansta %12.9'luk bir artış yaşanmıştır. Bu bulgular, test verisinin artan uzunluğunun performans üzerinde olumlu bir etkisi olduğunu göstermektedir.

Tüm bulgular değerlendirildiğinde, TVU değerinin artmasının farklı yaklaşımlarda farklı performans artışlarına neden olduğu belirlenmiştir. Örneğin, TVU değerinin iki katına çıkması, GKM yaklaşımında ortalama %5.2, GKM-EAPM yaklaşımında ortalama %5.6 ve GKM-DVM yaklaşımında ortalama %10.3 artışa yol açmıştır. Benzer şekilde, DSA yaklaşımında ortalama %5.9, DSA-TÖ yaklaşımında ortalama %5.2, x-vektörler yaklaşımında ortalama %16.2 ve d-vektörler yaklaşımında ise ortalama %11.6 performans artışı sağlanmıştır.



Şekil 40. Test verisi uzunluğunun OKT performansına etkisi

TVU değerinin dört katına çıkması durumunda da benzer bir durum ile karşılaşmıştır. Buna göre, GKM yaklaşımında ortalama %8.4, GKM-EAPM yaklaşımında ortalama %9.1 ve GKM-DVM yaklaşımında ortalama %16.8'lik artışlar meydana gelmiştir. Öte yandan, DSA yaklaşımında ortalama %9.8, DSA-TÖ yaklaşımında ortalama %8.8, x-vektörler yaklaşımında ortalama %23.8 ve d-vektörler yaklaşımında ise ortalama %18.6 performans artışı olması TVU değerinin performansı artırma yönündeki etkisini ortaya koymaktadır.

Şekil 40'ta çeşitli yöntem parametrelerinin etkisi altında, en gelişkin yöntemlerin performansları kutu grafiği halinde görselleştirilmiştir. Burada, TVU değerinin performansa etkisinin görselleştirilebilmesi adına, TVU değerinin 1, 2 ve 4 saniye olarak seçildiği gözlemler farklı kutu renkleri ile gösterilmiştir.

Şekil 40'ta verilen bulgular değerlendirildiğinde, TVU değerinin iki katına çıkması durumunda GKM (16), GKM (32) ve GKM (64) yaklaşımlarında performans artışı sırasıyla %4.1, %3.1 ve %3.5 olarak kaydedilmiştir. TVU değerinin dört katına çıkmasıyla birlikte ise, GKM (16), GKM (32) ve GKM (64) yaklaşımlarında performans artışı sırasıyla %6.8, %6.2 ve %5.2 olmuştur. GKM-EAPM yaklaşımında karışım sayıları 16, 32

ve 64 olduğunda TVU değerinin iki katına çıkmasıyla performans artışı sırasıyla %4.3, %3.6 ve %3.4 olarak gözlemlenmiştir. TVU değerinin dört katına çıkmasıyla birlikte ise performans artışı sırasıyla %5.9, %5.9 ve %6.0 olarak belirlenmiştir. GKM-DVM yaklaşımında karışım sayıları 16, 32 ve 64 olduğunda TVU değerinin iki katına çıkmasıyla performans artışı %6.5, %6.7 ve %8.5 olarak; TVU değerinin dört katına çıkmasıyla birlikte ise, performans artışı sırasıyla %10.2, %10.9 ve %13.0 olarak hesaplanmıştır.

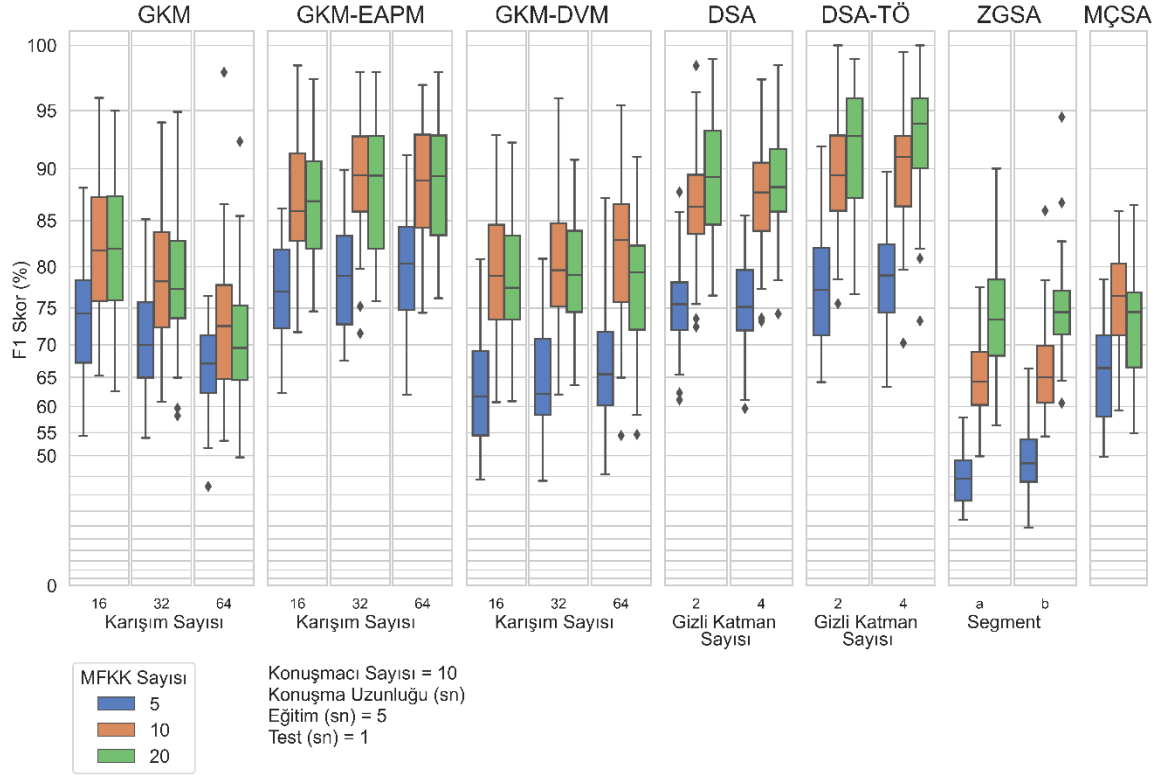
Şekil 40'a göre, TVU değerinin iki katına çıkması durumunda DSA (2) ve DSA (4) yaklaşımlarında sırasıyla performans artışı %3.4 ve %3.7 olarak belirlenmiştir. TVU değerinin dört katına çıkmasıyla birlikte ise, DSA (2) ve DSA (4) yaklaşımlarında performans artışı sırasıyla %5.2 ve %5.8 olarak gözlemlenmiştir. Benzer şekilde, TVU değerinin iki katına çıkması durumunda DSA-TÖ (2) ve DSA-TÖ (4) yaklaşımlarında sırasıyla performans artışı %2.1 ve %1.8 olarak belirlenmiştir. TVU değerinin dört katına çıkmasıyla birlikte ise, DSA-TÖ (2) ve DSA-TÖ (4) yaklaşımlarında performans artışı sırasıyla %4.0 ve %3.0 olarak gözlemlenmiştir.

Son olarak, TVU değerinin iki katına çıkmasıyla x-vektörler (a) ve x-vektörler (b) yaklaşımlarında sırasıyla performans artışı %11.7 ve %11.4 olarak kaydedilmiştir. TVU değerinin dört katına çıkmasıyla birlikte ise, x-vektörler (a) ve x-vektörler (b) yaklaşımlarında performans artışı sırasıyla %16.0 ve %15.4 olarak belirlenmiştir. Öte yandan, iki kat TVU değeri için d-vektörler yaklaşımında performans artışı %7.9 olarak belirlenmiştir. TVU değerinin dört katına çıkmasıyla birlikte ise, d-vektörler yaklaşımında performans artışı %12.8 olarak gözlemlenmiştir.

3.1.4. MFKK Sayısının Performansa Etkisi

Tüm bulgular göz önüne alındığında, MFKK sayısının artmasıyla performansta bir artış olduğu gözlenmiştir. MFKK sayısının iki katına çıkması durumunda ortalama %12.2 performans artışı görülürken, MFKK sayısının dört katına çıkması durumunda ise ortalama %13.6 performans artışı kaydedilmiştir. Bu sonuçlar, MFKK sayısının artmasının modelin performansını olumlu yönde etkilediğini işaret etmektedir.

Kullanılan yöntemlere bağlı olarak, MFKK sayısının iki katına çıkmasıyla performansta çeşitli artışlar gözlemlenmiştir. GKM yaklaşımında ortalama %8.0, GKM-EAPM yaklaşımında ortalama %9.3, GKM-DVM yaklaşımında ortalama %14.7, DSA yaklaşımında ortalama %12.3, DSA-TÖ yaklaşımında ortalama %12.7, x-vektörler



Şekil 41. MFKK sayısının OKT performansına etkisi

yaklaşımında ortalama %20.7 ve d-vektörler yaklaşımında ise ortalama %7.8 performans artışı yaşanmıştır.

MFKK sayısının dört katına çıkmasıyla ise, GKM yaklaşımında ortalama %8.1, GKM-EAPM yaklaşımında ortalama %10.6, GKM-DVM yaklaşımında ortalama %11.9, DSA yaklaşımında ortalama %15.1, DSA-TÖ yaklaşımında ortalama %15.7, x-vektörler yaklaşımında ortalama %29.0 ve d-vektörler yaklaşımında ise ortalama %6.3 performans artışı gözlenmiştir. Bu bulgular, kullanılan yöntemlerin MFKK sayısındaki artışa nasıl tepki verdiğini göstermektedir.

Şekil 41’de çeşitli yöntem parametrelerinin etkisi altında, en gelişkin yöntemlerin performansları kutu grafiği halinde görselleştirilmiştir. Burada, MFKK sayısının performansa etkisinin görselleştirilebilmesi adına, MFKK sayısının 5, 10 ve 20 olarak seçildiği gözlemler farklı kutu renkleri ile gösterilmiştir.

Şekil 41’de sunulan daraltılmış bulgular, MFKK sayısının iki katına çıkmasıyla, kullanılan farklı yöntemlerde performansta meydana gelen değişiklikleri detaylı bir şekilde göstermektedir. GKM yaklaşımında, 16 karışım kullanıldığında ortalama performans artışı %7.5, 32 karışım kullanıldığında %8.1, 64 karışım kullanıldığında ise %5.4 olmuştur.

GKM-EAPM yaklaşımında ise 16 karışım kullanıldığında ortalama performans artışı %9.9, 32 karışım kullanıldığında %10.4, 64 karışım kullanıldığında ise %9.3 olarak belirlenmiştir. GKM-DVM yaklaşımında, 16 karışım kullanıldığında ortalama performans artışı %16.3, 32 karışım kullanıldığında %15.4, 64 karışım kullanıldığında ise %15.3 olarak hesaplanmıştır.

DSA yaklaşımında, 2 gizli katman kullanıldığında ortalama performans artışı %10.4, 4 gizli katman kullanıldığında ise %11.6 olarak gözlenmiştir. DSA-TÖ yaklaşımında ise 2 gizli katman kullanıldığında ortalama performans artışı %11.8, 4 gizli katman kullanıldığında ise %11.4 olarak saptanmıştır.

Öte yandan, x-vektörler yaklaşımında “a” segmenti kullanıldığında ortalama performans artışı %20.7, “b” segmenti kullanıldığında ise %17.4 olarak belirlenmiştir. Son olarak, d-vektörler yaklaşımı için ortalama performans artışı %9.8 olarak hesaplanmıştır.

MFKK sayısının dört katına çıkmasıyla birlikte, farklı yöntemlerdeki performans değişiklikleri oldukça çeşitlenmiştir. GKM yaklaşımında, 16 karışım kullanıldığında ortalama performans artışı %8.2, 32 karışım kullanıldığında %7.8, 64 karışım kullanıldığında ise %3.6 olmuştur. GKM-EAPM yaklaşımında ise 16 karışım kullanıldığında ortalama performans artışı %9.8, 32 karışım kullanıldığında %9.4, 64 karışım kullanıldığında ise %9.1 olarak belirlenmiştir. GKM-DVM yaklaşımında, 16 karışım kullanıldığında ortalama performans artışı %15.5, 32 karışım kullanıldığında %14.7, 64 karışım kullanıldığında ise %11.6 olarak hesaplanmıştır.

DSA yaklaşımında, 2 gizli katman kullanıldığında ortalama performans artışı %13.6, 4 gizli katman kullanıldığında ise %13.4 olarak gözlenmiştir. DSA-TÖ yaklaşımında ise 2 gizli katman kullanıldığında ortalama performans artışı %14.3, 4 gizli katman kullanıldığında ise %13.7 olarak saptanmıştır.

Öte yandan, x-vektörler yaklaşımında “a” segmenti kullanıldığında ortalama performans artışı %29.5, “b” segmenti kullanıldığında ise %26.0 olarak belirlenmiştir. Son olarak, d-vektörler yaklaşımı için ortalama performans artışı %6.7 olarak hesaplanmıştır.

3.1.5. Tartışma

Şekil 38’den, Şekil 41’e kadar verilen bulgulara göre OKT performansı aynı anda birden çok faktörden etkilenmektedir. Bunların başında konuşmacı sayısı gelmektedir. Konuşmacı sayısı, tanınması gereken sınıf sayısını ifade ettiğinden, artmasıyla birlikte

problemin karmaşıklığı da artmaktadır. Bu durum, performansın önemli ölçüde düşmesine neden olmuştur. Çok sınıflı sınıflandırma problemlerinde bu durum kaçınılmazdır.

Diğer bir önemli faktör, konuşmacı başına düşen eğitim kayıtlarının uzunluğudur. Tüm yaklaşımlar için, EVU değerinin artmasıyla performans önemli ölçüde artmıştır. Ancak, performansın artış miktarı yöntem bazında farklılık göstermektedir. GKM ve DSA yaklaşımları diğer yaklaşımlara kıyasla EVU değerinin artmasından daha fazla etkilenmiştir. Bu yaklaşımları diğer yöntemlerden ayıran özellik, başlangıçta bir geliştirme aşamasına sahip olmamaları ve model eğitiminde sadece eğitim verilerini kullanmalarıdır. Bu açıdan, geliştirme aşamasının genelleme kapasitesini artırdığı ve eğitimde kullanılacak veri ihtiyacını düşürdüğü söylenebilmektedir. Ayrıca, bu yöntemlerin parametre değerlerinin artmasıyla performans, düşük parametre değerlerine göre daha ivmeli bir artış göstermiştir. Sınıflandırıcı parametre değerlerinin yükselmesinin yöntem karmaşıklığını artırdığı düşünüldüğünde, yöntemlerin karmaşıklığının artmasıyla eğitim verisi ihtiyacının da arttığı sonucuna varılabilmektedir. GKM ve DSA yaklaşımlarının haricinde kalan yaklaşımlar için bu yönde bir bulguya rastlanılmamıştır.

Öte yandan, test aşamasında kullanılan konuşma kayıtlarının uzunluğunun da performansı önemli ölçüde etkilediği görülmektedir. Tüm yaklaşımlar için, TVU değerinin artması performansın artmasına neden olmuştur. Ancak yöntem bazında değerlendirildiğinde; GKM-DVM, x-vektörler ve d-vektörler yaklaşımları TVU değerine diğerlerine göre daha duyarlıdır. Bu yaklaşımları diğerlerinden ayıran özellik, test aşamasında çerçeveleri değerlendirmek yerine çerçeve ortalamalarını değerlendirmeleridir. Diğer yaklaşımlar, test aşamasında MFKK özniteliklerinin tüm çerçevelerinin sınıf etiketlerini ayrı ayrı belirlemekte ve konuşmacı kimliğini konuşma kaydına ait çerçevelerin en çok atandığı konuşmacı olarak tayin etmektedir. GKM-DVM, x-vektörler ve d-vektörler yaklaşımında ise her konuşma kaydı için MFKK özniteliklerinin çerçevelerinden elde edilen vektör tek boyutludur. Bu nedenle, TVU değerinin artmasıyla, tek boyutlu özniteliklerin konuşmacıyı temsil etme kabiliyeti artmakta ve performans artışı gerçekleşmektedir. Tüm yöntemler için TVU değeri performansı pozitif yönde etkilese de GKM-DVM, x-vektörler ve d-vektörler yaklaşımlarının performansları, uzun test verisi ihtiyaçları nedeniyle, daha ivmeli bir şekilde artmaktadır.

D1 senaryosunun sonuçlarına göre, otomatik konuşmacı tanıma için kullanılan MFKK sayısı da performansı etkileyen önemli bir faktördür. Ancak, bu etki MFKK sayısının artış hızına ve kullanılan yönteme bağlıdır. MFKK sayısının 5'ten 10'a çıkması

tüm yaklaşımların performansını pozitif yönde etkilemiştir, ancak 10'dan 20'ye çıkmasıyla performans yöntemler arasında çeşitlilik göstermiştir. MFKK sayısının etkisinin incelendiği deneylerde, MFKK sayısının 5 olarak seçilmesi OKT uygulaması için yetersiz kalmıştır. Bu nedenle, MFKK sayısının 10 olarak seçilmesi tüm yaklaşımlarda pozitif etki ile sonuçlanmıştır. Öte yandan, MFKK sayısının 10'dan 20'ye çıkması durumunda DSA, DSA-TÖ ve x-vektör yaklaşımlarının performansları pozitif yönde etkilenirken, diğer yöntemler için performansın düştüğü gözlenmiştir. Bu yaklaşımların DÖ tabanlı yöntemler olduğu düşünüldüğünde, öznitelik sayısının artmasıyla performansın artmış olması, DÖ tabanlı yöntemlerin veri miktarının yanı sıra, verilerin ifade edildiği öznitelik vektörlerinin de detaylı olmasına ihtiyaç duyduğunu göstermektedir. GKM, GKM-EAPM ve GKM-DVM yaklaşımlarında, MFKK sayısının artmasının model karmaşıklığını artırdığı ve bu nedenle performansta düşüşe neden olduğu görülmektedir.

Sınıflandırma yaklaşımları açısından yapılan incelemelerde, yaklaşımların belirli özelliklerinin performanslarını ve faktörlere olan duyarlılıklarını etkilediği gözlemlenmiştir. GKM yaklaşımında herhangi bir geliştirme aşamasının olmaması ve eğitimde sadece belirli konuşmacılara ait verilerin kullanılması, bu yaklaşımın genelleme kapasitesinin düşük olmasına neden olmaktadır. Bu durum, yaklaşımın eğitim verisi ihtiyacının GKM-EAPM ve GKM-DVM yaklaşımlarına göre daha yüksek olmasına yol açmaktadır. Ayrıca, düşük eğitim verisi miktarlarında, model karmaşıklığını artıran parametrelerin artması, performansın önemli ölçüde düşmesine neden olmaktadır. GKM yaklaşımının bir EAPM yardımıyla geliştirme aşamasına tabi tutulması ve elde edilen GKM modellerinin EAPM modelinin adapte edilmesiyle elde edilmesi, modelin eğitim verisi ihtiyacını azaltmakta ve performansın yanı sıra performansın sağlamlığını da artırmaktadır. Bu açıdan, geliştirme aşamasının yaklaşımların genelleme kabiliyetini artırdığı ve performansı yükselttiği söylenebilmektedir.

Benzer bir durum, DSA ve DSA-TÖ yaklaşımlarının incelenmesiyle de gözlemlenmektedir. DSA yönteminin eğitim verisinin miktarına olan duyarlılığı, DSA-TÖ yaklaşımına göre daha yüksektir. Bununla birlikte, DSA-TÖ yaklaşımının performansı ve performans sağlamlığı, DSA yaklaşımına göre daha yüksektir. Ayrıca DSA-TÖ yöntemi, DSA yaklaşımına göre model karmaşıklığından daha az etkilenmektedir. DSA yaklaşımında model karmaşıklığı arttıkça performans düşme eğilimindeyken, DSA-TÖ yaklaşımında performans artma eğilimindedir.

GKM yaklaşımı ile GKM-EAPM ve GKM-DVM yaklaşımlarının, DSA yaklaşımı ile DSA-TÖ yaklaşımının performansları karşılaştırıldığında geliştirme aşamasının performans yönünden pozitif etki gösterdiği görülmektedir. Geliştirme aşamasına sahip olan diğer yaklaşımlar olan x-vektörler ve d-vektörler yaklaşımlarında, tüm deneylerde en düşük ortalama performans değerleri gözlemlenmiştir. Bunun başlıca nedeni, bu yaklaşımların test aşamasında çerçeve bazlı bir yaklaşım yerine ortalama vektörlerin etiketlerinin belirlenmesi yaklaşımına sahip olmasıdır. Deneyde belirlenen TVU değerlerinin düşük olması, bu yaklaşımların performansını dramatik bir şekilde etkilemiştir. Ayrıca, bu yaklaşımlar, DSA ve DSA-TÖ yaklaşımları ile karşılaştırıldığında oldukça karmaşık bir yapıya sahiptirler. Nispeten daha karmaşık yapıları, yaklaşımların genelleme yapma kapasitesini negatif yönde etkilemiştir. Daha yüksek EVU ve TVU değerlerinde kullanıldığı, konuşmacı sayısının daha fazla olduğu uygulamalarda bu yöntemler diğer yaklaşımlardan daha yüksek performans göstermeye adaydırlar.

EK-2'de verilen tablolar incelendiğinde, ortalama performansın en düşük olduğu denemelerin, EVU değerinin 5 saniye, TVU değerinin 1 saniye ve MFKK sayısının 5 olarak seçildiği denemeler olduğu görülmüştür (57.1 ± 14.3). Diğer taraftan, ortalama performansın en yüksek olduğu denemelerin ise, EVU değerinin 20 saniye, TVU değerinin 4 saniye ve MFKK sayısının 20 olarak seçildiği denemeler olduğu belirlenmiştir (95.8 ± 3.5). Ayrıca, tüm denemelerin ortalama performansı 81.9 ± 13.7 olarak hesaplanmıştır. Referans bulguları olarak belirlenen denemelerde ise EVU değerinin 5 saniye, TVU değerinin 1 saniye ve MFKK sayısının 20 olarak seçildiği durumlar için ortalama performans 74.6 ± 11.5 olarak hesaplanmıştır.

Referans bulgularının ortalama performans değeri göz önünde bulundurulduğunda, parametre dizilerinin ortalamanın üstünde zorluk derecesine sahip olduğu görülmektedir. Bu nedenle, bu tez çalışmasının sıradaki deneyleri için TVU değeri, EVU değeri ve MFKK sayısının sırasıyla 5 saniye, 1 saniye ve 23 olduğu düşünüldüğünde, deneylerin oldukça zorlayıcı şartlar altında gerçekleştirildiği savunulabilmektedir.

D1 senaryosunda elde edilen sonuçlara göre OKT uygulamalarının başarısı, konuşmacıları ne kadar doğru tespit ettiği ile ölçülmektedir. Ancak, gerçek dünyada konuşmacıların tespit edilmesinden sorumlu bir modelin, yüksek doğruluğun yanında bir dizi özelliklere de sahip olması gerekmektedir. Bu özellikler arasında, modelin eğitiminde ihtiyaç duyulan toplam süre, bir konuşma kaydının kimliğini tespit ederken gereken süre ve modelin hafızada işgal ettiği alan bulunmaktadır. Bir OKT modelinin yüksek

performans sergilemesi, o modelin etkin bir model olduğunu göstermemektedir. Aynı zamanda, hızlı bir eğitim aşamasına sahip olması, konuşma kayıtlarını hızlı bir şekilde sınıflandırması ve hafızada az yer kaplaması, gerçek dünyada kullanılabilirliği açısından da önemlidir. Bu tez çalışmasında, D1 senaryosu olarak tanıtılmış olan bulgular, sıralanan kriterler bakımından genişletilmiş ve “Expert Systems with Applications” dergisinde “Evaluating MFCC-based speaker identification systems with data envelopment analysis” başlıklı makalede yayımlanmıştır (Özcan & Kayıkçıoğlu, 2021).

3.2. Klasik Makine Öğrenimi Yöntemlerinde Öznitelik Spike Dönüşümü Aşamasının İncelenmesi

Bu bölümde, detaylarıyla birlikte Bölüm 2.6.2'de tanıtılan D2 isimli deneysel senaryoya dair bulgular verilmiştir. D2 senaryosunda, çeşitli VK'ler, önce öznitelik vektörleri herhangi bir işleme tabi tutulmadan klasik MÖ yöntemleriyle sınıflandırılmış ve değerlendirilmiştir. Sonrasında, aynı VK'ler, önerilen ÖSD yaklaşımı ile ikili formda spike örüntülerine dönüştürülmüş ve yine aynı klasik makine öğrenimi yöntemleriyle ve gerçekçi nöron (GN) modelleri ile sınıflandırılmış ve değerlendirilmiştir.

Burada sunulan bulgular, ÖSD yaklaşımının sınıflandırma performansını nasıl etkilediğini ortaya koymak amacıyla organize edilmiştir. D2'ye dair detaylı bulgulara EK-3 başlığı altında ulaşılabilmektedir. Bu bölümde, üç farklı VK'ye dair bulgular sunulmuştur. Bunlar; biri yapay olarak üretilmiş bir VK, diğeri IRIS çiçekleri VK (Fisher, 1936) ve sonuncusu MNIST el yazısı rakamlar VK (Deng, 2012) olarak sıralanmıştır. Yapay VK, ÖSD aşamasında öbekler arası karar sınırlarının gözlemlenebilmesi için kullanılmıştır. Karşılaştırma VK'leri ise ÖSD aşamasının, yaygın kullanılan VK'lerde sınıflandırma performansını ölçebilmek adına kullanılmıştır. ÖSD aşamasının etkisinin gözlemlenebilmesi adına, klasik MÖ yöntemleri olarak k-en yakın komşuluk (K-EYK), destek vektör makineleri (DVM), tek katmanlı perceptron (TKP), Naive Bayes (NB) ve lineer ayırıcı analizi (LAA) olmak üzere toplam beş yöntem kullanılmıştır (Cover & Hart, 1967; Cortes & Vapnik, 1995; Cheeseman vd., 1988; Xanthopoulos vd., 2013).

Yapay veri kümeleri incelenirken, eğitim sonucu oluşan öbek alan haritaları (ÖAH) göz önünde bulundurulmuş ve görselleştirilmiştir. Şekillerin değerlendirilmesi aşamasında, karşılaştırmalı bir yaklaşım takip edilmiştir. Öncelikle, bir referans ÖAH seçilmiştir. Sonrasında, referans ÖAH ile karşılaştırmada kullanılacak bir ÖAH arasında örtüşme oranı

belirlenmiştir. Örtüşme oranı belirlenirken, karşılaştırılan iki ÖAH için öbek alanlarının ne kadarının çakıştığı yüzdesel olarak hesaplanmış ve değerlendirilmiştir.

Karşılaştırma VK'leri olarak kullanılan IRIS ve MNIST VK için tüm bulgular, sınıf bazlı olarak F1-Skoru cinsinden hesaplanmıştır. Buna göre, elde edilen gerçek ve tahmin edilen etiketler kullanılarak VK'lerin tüm sınıfları için ayrı ayrı F1-Skor ölçütleri belirlenmiştir. Kullanılan yöntemin genel performansı, sınıflar için hesaplanan F1-Skor değerlerinin ortalama ve standart sapmasının hesaplanmasıyla elde edilmiştir. Ayrıca, yöntem performansları sınıf içi performanslarının da ortaya koyulabilmesi amacıyla radar grafiklerle görselleştirilmiştir. Radar grafiklerde, sınıf içi performans değerleri bir çember içinde, belirli açılarla belirli sınıflara atandığı bir yaklaşımla görselleştirilmektedir. Çemberin merkezi %0 performansa, dış sınırı ise %100 performansa karşılık gelmektedir. Çember içinde gösterilen iç çemberler ise ara değerleri ifade etmektedir.

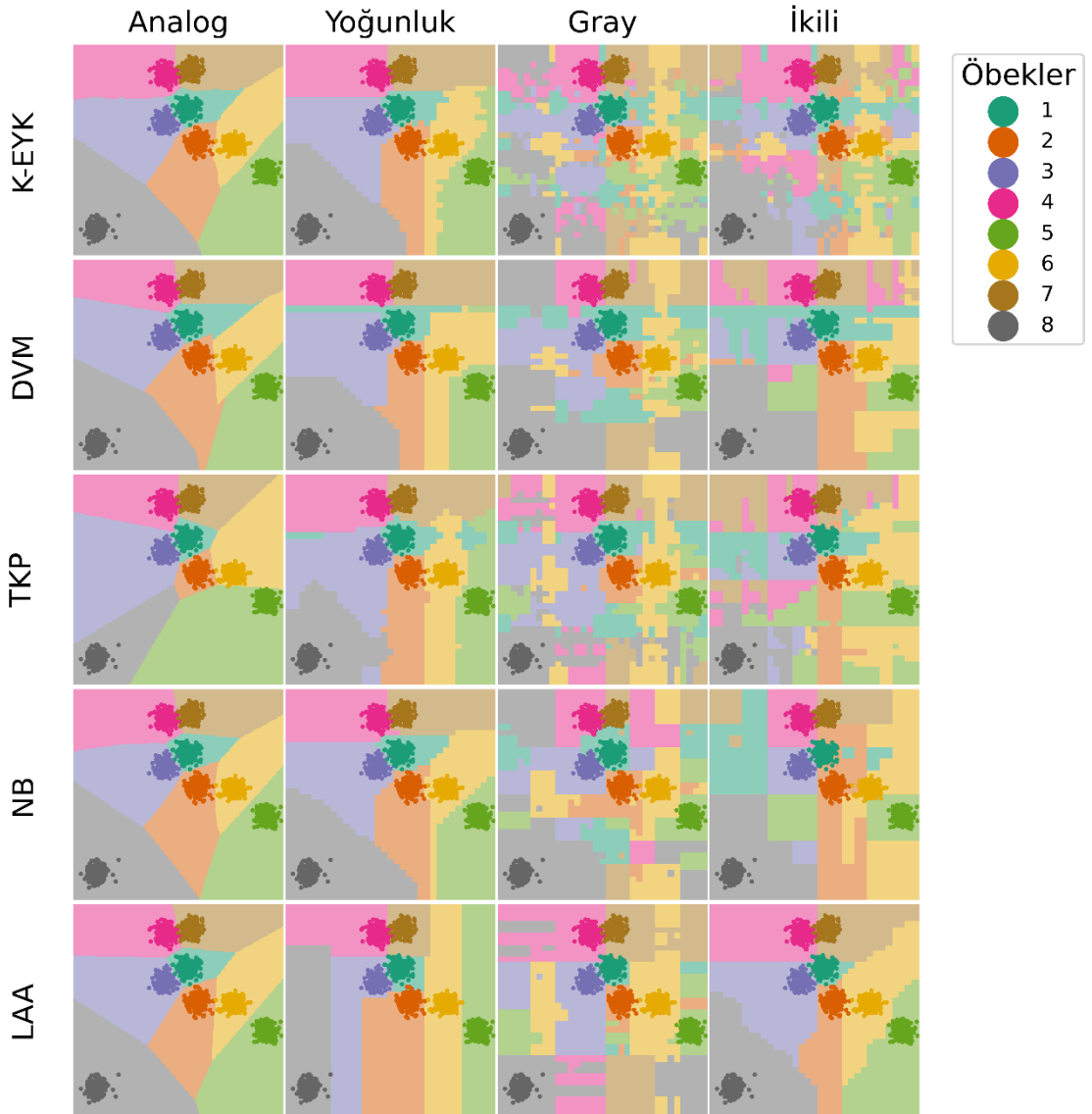
Yapay veri kümesi için analog öznitelikler kullanıldığında, eğitim sonunda elde edilen ÖAH'ler incelendiğinde, K-EYK yönteminin toplam 9 örneği hatalı etiketlediği tespit edilmiştir. DVM yöntemi için bu değer 10 olarak bulunurken, TKP yöntemi toplam 13 örneği hatalı etiketlemiştir. NB yöntemi toplam 11 hatalı etiketleme yaparken, LAA yöntemi için bu sayı 10 olarak belirlenmiştir. Bu deneyde, en iyi performansı K-EYK yöntemi gösterirken, TKP yöntemi en başarısız sınıflandırıcı olarak tespit edilmiştir.

Analog özniteliklerin kullanıldığı durumda, IRIS ve MNIST VK'ler için de performanslar incelenmiştir. IRIS VK için, K-EYK yöntemi 93.9 ± 4.4 , DVM yöntemi 98.8 ± 0.9 , TKP yöntemi 90.2 ± 7.0 , NB yöntemi 95.2 ± 3.4 ve LAA yöntemi 98.8 ± 0.9 performans sergilemiştir. Öte yandan, MNIST VK için, K-EYK yöntemi 97.0 ± 2.4 , DVM yöntemi 96.5 ± 3.1 , TKP yöntemi 93.4 ± 3.3 , NB yöntemi 79.7 ± 15.9 ve LAA yöntemi 93.9 ± 3.6 performans göstermiştir.

Bu bölüm, dört alt başlıktan meydana gelmektedir. İlk alt başlıkta, kodlama yönteminin; ikinci alt başlıkta, kuantalama tekniğinin; üçüncü alt başlıkta, bit uzunluğunun; klasik MÖ yöntemlerinin performansını nasıl etkilediğine dair bulgular verilmiştir. Dördüncü ve son alt başlıkta ise bulgular değerlendirilmiş ve tartışılmıştır.

3.2.1. Kodlama Yönteminin Etkisi

Bu bölümde, ÖSD aşamasında bir seçenek olarak sunulan kodlama yöntemlerinin sınıflandırıcı performanslarını nasıl etkilediği, klasik MÖ yöntemleri kullanılarak



Şekil 42. Kodlama yöntemine göre öbek alan haritaları

incelenmiştir. Toplam üç farklı VK'den ilki olan yapay VK, bu bölümde sınıflandırıcıların eğitim sonunda öbekler için belirlediği karar alan ve sınırlarının görselleştirilmesi amacıyla kullanılmıştır. Diğer iki VK ise kodlama yönteminin sınıflandırma performansını nasıl etkilediğinin gözlemlenebilmesi için kullanılmıştır.

Klasik MÖ yöntemlerinin yapay VK için kodlama yöntemlerine bağlı olarak belirlediği ÖAH'ler Şekil 42'de görselleştirilmiştir. Burada; sütunlar kodlama yöntemini, satırlar ise klasik MÖ yöntemlerini temsil etmektedir. "Analog" olarak isimlendirilmiş sütunda öznitelik vektörlerinin kodlanmamış haline ait bulgular, diğer sütunlarda ise analog özniteliklerin ilgili yöntemlerle kodlanmış hallerine ait bulgular gösterilmektedir.

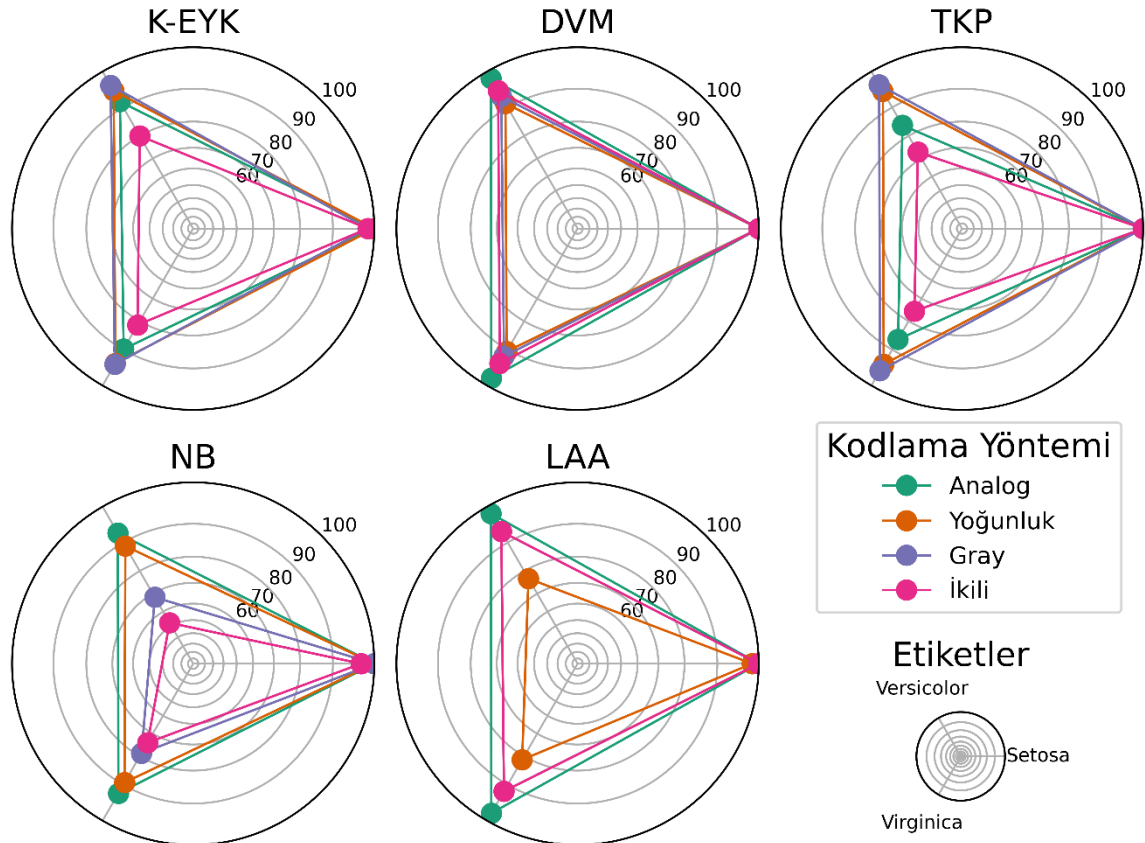
ÖSD aşamasındaki diğer faktörler olan kuantalama tekniği ve bit uzunluğu, bu bölüm özelinde sırasıyla "tekdüze" ve "5" olarak seçilmiştir.

Şekil 42 incelendiğinde, MÖ yöntemlerinin ÖAH bulguları kodlama yöntemlerine göre farklılık göstermektedir. Kodlama yöntemleri değerlendirildiğinde, yoğunluk kodlama (YK) uygulanmış özniteliklerin öbek alanlarının bir bütünlük arz ettiği, Gray kodlama (GK) ve ikili kodlama (İK) uygulandığında ise öbek alanlarının parçalı bir yapıya sahip olduğu görülebilmektedir. Bu durumun tek istisnası, LAA yönteminin ikili kodlama yapılmış öznitelikler için belirlediği ÖAH'lerdir. LAA yöntemi burada bütünlük arz eden öbek alanlarını belirlemiştir.

K-EYK yöntemi incelendiğinde, YK uygulanmış özniteliklerde toplam 14 adet örnek hatalı etiketlenmiştir. GK ve İK için ise bu değerler sırasıyla 18 ve 17 olarak belirlenmiştir. ÖAH'ler değerlendirildiğinde, K-EYK yönteminde YK uygulandığında elde edilen ÖAH %89.9, GK veya İK kullanıldığında ise elde edilen ÖAH'ler %54.5 oranında, analog ÖAH ile örtüşmektedir. DVM yönteminde, özniteliklere YK uygulandığında toplam 14 adet örneğin etiketleri hatalı belirlenmiştir. GK ve İK için bu değerlerin sırasıyla 15 ve 17 olduğu görülmüştür. ÖAH örtüşme oranları ise YK, GK ve İK için sırasıyla %85.3, %56.3 ve %66.6 olarak belirlenmiştir.

TKP yönteminde; analog öznitelikler kullanıldığında toplam hatalı etiket miktarı 13 olarak belirlenirken, YK uygulanmış öznitelikler için bu değer 12, GK için 15 ve İK için ise 64 olarak belirlenmiştir. Burada, TKP yönteminin YK uygulanmış özniteliklerde, analog özniteliklere göre daha az hata yaptığı görülmüştür. Ayrıca, TKP yönteminin İK uygulanmış özniteliklerde özellikle "Öbek 1" örneklerini doğru etiketleme kapasitesi düşmüştür. ÖAH'ler değerlendirildiğinde; YK uygulanmış özniteliklere ait ÖAH ile analog öznitelik ÖAH arasında örtüşme oranı %61.9 olarak hesaplanmıştır. GK ve İK için bu değerlerin sırasıyla %49.1 ve %51.7 olduğu saptanmıştır. ÖAH örtüşme oranlarına göre TKP yönteminin kodlama aşamasından oldukça etkilendiği görülmüştür.

NB yöntemi için kodlama sonucu performansta düşüş gözlenmesine rağmen en uygun kodlama yönteminin YK olduğu saptanmıştır. Buna göre, NB yöntemi YK uygulanan özniteliklerde toplam 18 adet hatalı etiketleme yapmıştır. GK ve İK uygulanmış özniteliklerde ise sırasıyla 93 ve 292 adet hata yapmıştır. ÖAH'lere göre, YK uygulanmış öznitelikler ile analog öznitelik ÖAH'leri %91.7 oranında örtüşmektedir. GK ve İK uygulanmış özniteliklerin ÖAH'lerine göre ise sırasıyla %51.3 ve %53.2 örtüşme oranları saptanmıştır.

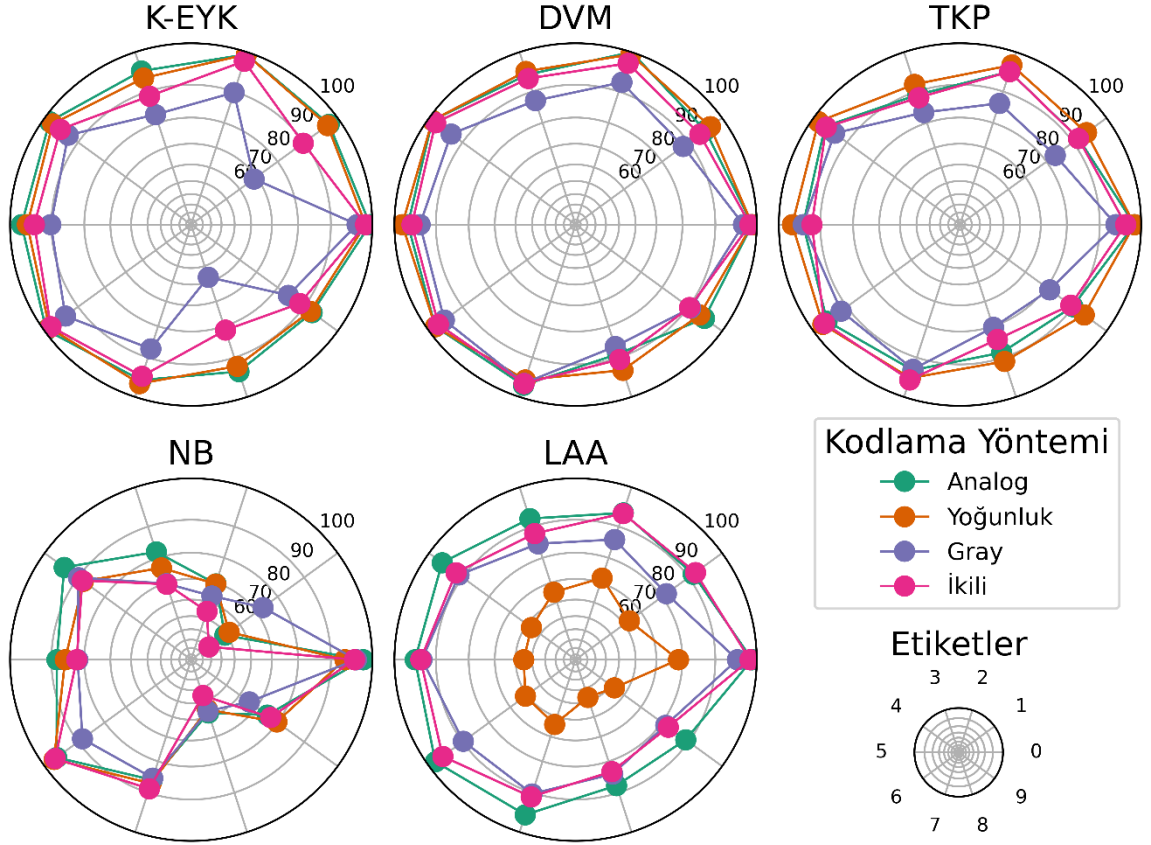


Şekil 43. Kodlama yöntemine göre IRIS veri kümesinde performanslar

Son olarak, LAA yönteminin kodlama aşamasından oldukça etkilendiği görülmüştür. Analog özniteliklerde toplam 10 adet hatalı etiketleme yapan LAA yöntemi için bu değerler YK için 32, GK ve İK için ise 75'er olarak belirlenmiştir. ÖAH'ler incelendiğinde ise diğer yöntemlerin aksine LAA yöntemi en yüksek örtüşme oranını %85.9 ile İK uygulanmış özniteliklerde göstermiştir. YK için bu oran %60.8, GK için ise %51.5 olarak hesaplanmıştır.

Yapay VK için ÖAH bulguları, kodlama yönteminin sınıflandırma performansına etkileri hakkında birçok bilgi sağlamaktadır. Ancak kodlama yönteminin gerçek dünya verilerinde performansın ne kadar değişeceğine dair yeteri kadar bilgi sağlamamaktadır. Bu nedenle, IRIS ve MNIST gibi gerçek dünya VK'lerinin kullanıldığı deney sonuçlarının önemi artmaktadır. İlk olarak IRIS incelenmiştir. IRIS için klasik MÖ yöntemlerinde kodlama yönteminin etkisi, Şekil 43'te verilen radar grafiklerde gösterilmektedir.

Şekil 43'te verilen radar grafiklere göre, kodlama yöntemlerinin performansını en az etkilediği sınıflandırıcı DVM yöntemidir. DVM yönteminde analog öznitelikler



Şekil 44. Kodlama yöntemine göre MNIST veri kümesinde performanslar

kullanıldığında sınıflandırma performansı 98.8 ± 0.9 olarak hesaplanmıştır. Öznitelikler kodlandığında ise YK için 4.9 , GK için 3.7 ve İK için ise 2.5 'lik performans düşüşleri kaydedilmiştir.

Kodlama yöntemlerinin performansını en çok etkilediği sınıflandırıcı ise NB yöntemidir. NB yönteminde analog öznitelikler kullanıldığında performans 95.2 ± 3.4 olarak kaydedilirken, özniteliklere YK uygulandığında 2.5 'lik bir düşüş saptanmıştır. GK ve İK uygulandığı durumlarda ise performans değerlerinde 12.9 ve 20.6 'lık düşüş saptanmıştır.

Diğer sınıflandırıcıların aksine TKP yönteminde kodlama yöntemleri performansta pozitif etki meydana getirmiştir. YK uygulanmış özniteliklerde performansın 6.2 arttığı görülürken, GK için bu artış 7.4 olarak kaydedilmiştir. Ancak İK uygulanan özniteliklerde performansın 7.2 düşmesi, TKP yönteminde İK yönteminin performansta negatif etkiye sahip olduğunu göstermektedir. Benzer şekilde, K-EYK yöntemi için

özniteliklere YK uygulandığında performansta %2.4 artış, GK uygulandığında %2.6 artış, İK uygulandığında ise %7.2 düşüş meydana gelmiştir.

Öte yandan, LAA yöntemi için YK uygulandığında performansında %13.2 düşüş saptanmıştır. GK için bu değer %3.7 düşüşe karşılık gelirken, İK uygulandığında LAA performansı %3.7 düşmüştür. LAA yönteminin yapay VK'de YK için gösterdiği negatif duyarlılığın IRIS için de geçerli olduğu görülmektedir.

Karşılaştırma VK'lerinden bir diğeri olan MNIST için sınıflandırıcı performansları Şekil 44'te gösterilmektedir. Buna göre NB yönteminin kodlama yönteminden bağımsız olarak MNIST için başarısız bir sınıflandırıcı olduğu söylenebilmektedir. Gerek ortalama performansının %79.7 olması gerek sınıf içi performansların dengesiz yapısı NB yönteminin analog özniteliklerde dahi değerlendirilmesinin önünde bir engel teşkil etmektedir.

DVM yöntemi, MNIST için kodlama yönteminden en az etkilenen sınıflandırıcı olmuştur. YK uygulanmış özniteliklerde performansta %0.4 artış görülürken, İK için %1.5 düşüş belirlenmiştir. GK uygulanan özniteliklerde performansın %4.2 düşmesi ve sınıf içi performans değerlerinin dalgalanması, DVM yöntemi için GK'nin diğer kodlama yöntemlerine göre daha başarısız olduğunu göstermektedir. TKP yöntemi için de benzer bir durum söz konusudur. YK uygulanmış özniteliklerde TKP yöntemi için performans %2.1 artmış, İK yönteminde ise %0.7 düşmüştür. GK için performansın %5.1 düşmüş olması GK yönteminin MNIST özniteliklerinde diğer kodlama yöntemlerine göre daha yüksek bozulmaya neden olduğunu göstermektedir.

K-EYK yöntemi için tüm kodlama yöntemleri performansta düşüşe neden olsa da YK, aralarında en başarılı kodlama yöntemi olmuştur. Buna göre YK uygulanmış özniteliklerde performans %0.7, İK uygulanmış özniteliklerde ise %4.2 düşmüştür. GK uygulanmış özniteliklerde performansın %12.6 düşmüş olması, MNIST için GK yönteminin uygun olmadığını göstermektedir. Son olarak, LAA yöntemi için en uygun kodlama yönteminin İK olduğu görülmüştür. Buna göre, İK uygulanmış özniteliklerde LAA yönteminin performansı %2.7 düşmüştür. GK yöntemi için bu değer %6.0, YK için ise %32.2 düşüşe karşılık geldiği düşünüldüğünde, LAA ile YK arasında negatif, LAA ile İK arasında görece pozitif bir uyum olduğu söylenebilmektedir.

3.2.2. Kuantalama Tekniğinin Etkisi

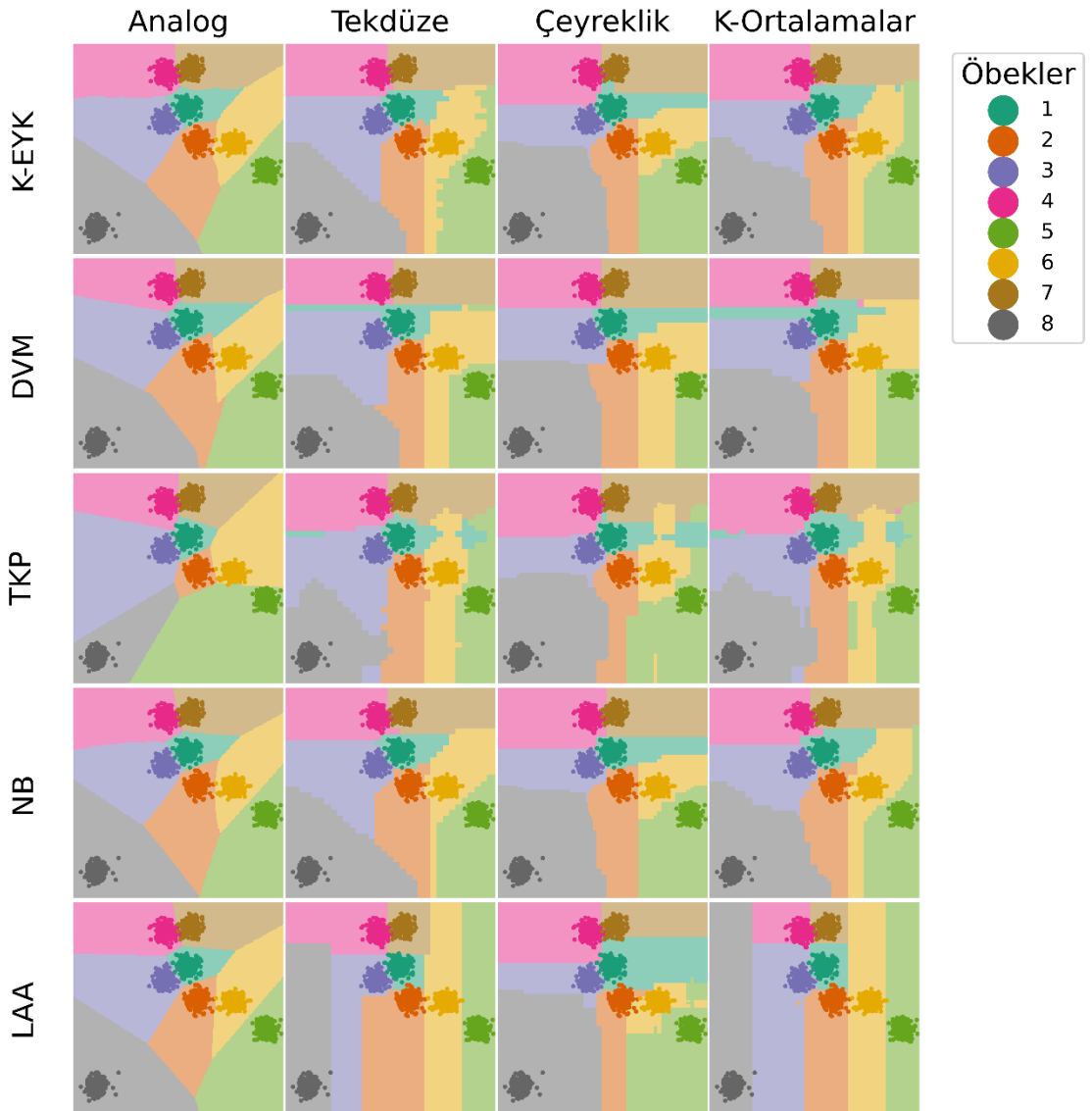
Bu bölümde, ÖSD aşamasında bir seçenek olarak sunulan kuantalama teknikleri sınıflandırıcı performanslarını nasıl etkilediği klasik MÖ yöntemleri kullanılarak incelenmiştir. Toplam üç farklı VK'den ilki olan yapay VK, bu bölümde, sınıflandırıcıların eğitim sonunda öbekler için belirlediği karar alan ve sınırlarının görselleştirilmesi için kullanılmıştır. Diğer iki VK ise kuantalama tekniğinin sınıflandırma performansını nasıl etkilediğinin gözlemlenebilmesi adına kullanılmıştır.

Klasik MÖ yöntemlerinin yapay VK için kuantalama tekniğine bağlı olarak belirlediği ÖAH'ler Şekil 45'te görselleştirilmiştir. Buna göre sütunlar kuantalama tekniğini, satırlar ise klasik MÖ yöntemlerini temsil etmektedir. “Analog” olarak isimlendirilmiş sütunda öznitelik vektörlerinin kodlanmamış haline ait sonuçlar, diğer sütunlarda ise analog özniteliklerin ilgili yöntemler ile kuantalanmış ve kodlanmış hallerine ait ÖAH'ler gösterilmektedir. ÖSD aşamasındaki diğer faktörler olan kodlama yöntemi ve bit uzunluğu, bu bölüm özelinde sırasıyla “Yoğunluk” ve “5” olarak seçilmiştir.

K-EYK yöntemi incelendiğinde, K-Ortalamlar ile kuantalanmış özniteliklerde 12 adet örnek hatalı etiketlenmiştir. Tekdüze ve çeyreklik için bu değerler, 14 ve 10 olarak belirlenmiştir. ÖAH'ler değerlendirildiğinde ise K-EYK yönteminde, Tekdüze kuantalanmış özniteliklerin ÖAH'si analog öznitelikler ile %89.9 örtüşmüştür. Bu değerler; K-Ortalamlar için %86.8, çeyreklik için ise %83.6 olarak hesaplanmıştır.

DVM yönteminde, tekdüze kuantalama yapıldığında toplam 14 adet örneğin etiketleri hatalı belirlenmiştir. K-Ortalamlar için bu değer 10 olarak belirlenirken; çeyreklik ile kuantalanmış özniteliklerde DVM yönteminin 6 örneği hatalı etiketlemesi, analog özniteliklere göre daha başarılı bir sınıflandırma yaptığını göstermektedir. ÖAH örtüşme oranları ise tekdüze, çeyreklik ve K-Ortalamlar için sırasıyla %85.3, %80.4 ve %82.9 olarak belirlenmiştir.

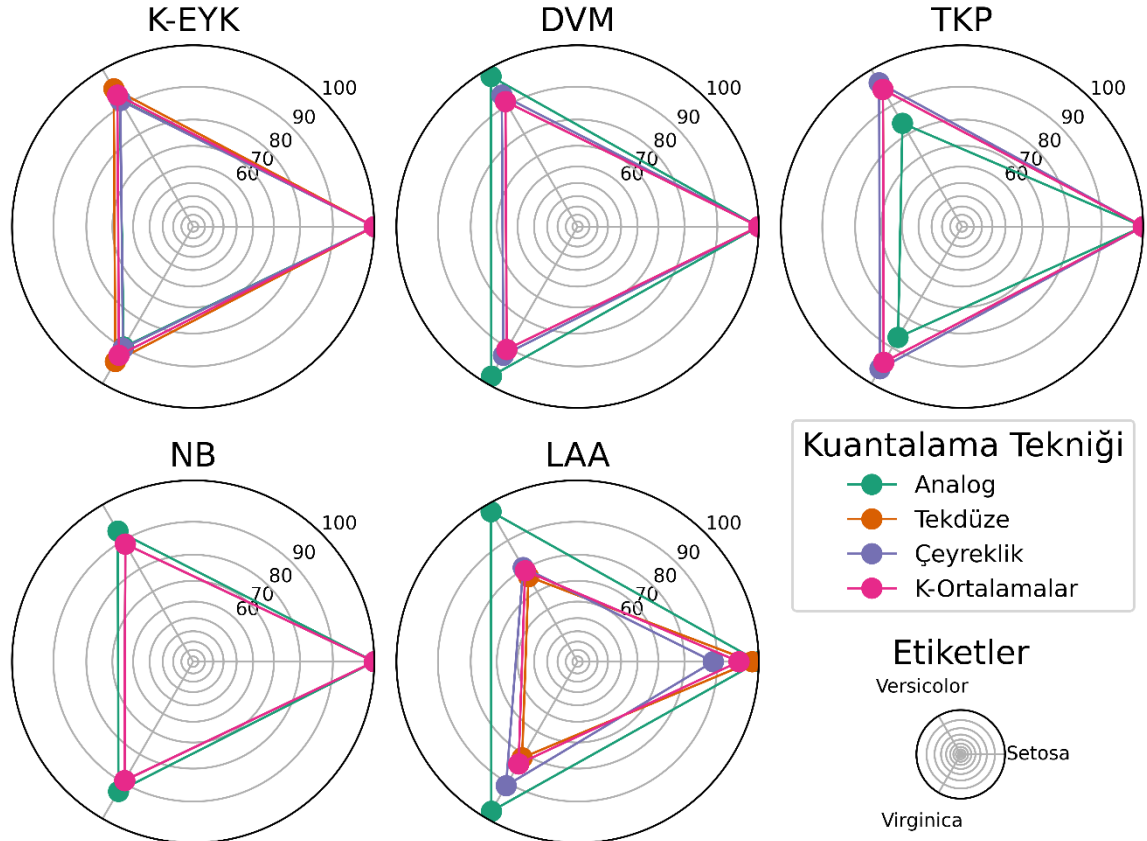
TKP yönteminde; kuantalama tekniğine tabii tutulmuş tüm öznitelikler analog öznitelik sonuçlarına kıyasla daha az hatalı etiketleme yapmıştır. Buna göre, çeyreklik kuantalanmış özniteliklerde TKP yöntemi toplam 6 örneğin etiketini belirlerken hata yapmıştır. Tekdüze ve K-Ortalamlar için bu değerler 12 ve 10 olarak saptanmıştır. TKP yönteminin analog öznitelikleri etiketlerken toplam 13 hata yapmış olması, kullanılan kuantalama ve kodlama yöntemlerinin TKP performansına pozitif etki yaptığı



Şekil 45. Kuantalama tekniğine göre öbek alan haritaları

görülebilmektedir. Ancak, ÖAH örtüşme oranları bu yönde bir bulgu sağlamamaktadır. Tekdüze kodlanmış öznitelikler için örtüşme oranı %61.9 iken, çeyreklik ve K-Ortalamlar için bu değerler %67.0 ve %63.4 olarak hesaplanmıştır.

NB yönteminde özniteliklere çeyreklik kuantalama uygulanarak kodlandığında toplam 15 adet örnekte hatalı etikete rastlanmıştır. Tekdüze ve K-Ortalamlar için bu değerler 18 olmak üzere eşit çıkmıştır. ÖAH örtüşme oranları incelendiğinde ise, tekdüze için %91.7, çeyreklik için %85.8, K-Ortalamlar için ise %89.3 analog öznitelik sonuçlarıyla örtüşme saptanmıştır.



Şekil 46. Kuantalama tekniğine göre IRIS veri kümesinde performanslar

Son olarak, LAA yönteminde, tekdüze, çeyreklik ve K-Ortalamlar ile kuantalama yapıldığında, sınıflandırıcı sırasıyla 32, 68 ve 36 adet örnekte hatalı etiketleme yapmıştır. LAA yönteminin YK uygulanmış öznitelikleri sınıflandırmadaki başarısızlığı, bu sonuçların kuantalama tekniğinden çok kodlama yöntemiyle alakalı olarak düşük çıktığını göstermektedir. Bu durum, ÖAH örtüşme oranlarından da görülmektedir. Buna göre tekdüze kuantalanmış özniteliklerde örtüşme oranı %60.8 çıkarken, çeyreklik için %75.2, K-Ortalamlar için ise %54.7 olarak hesaplanmıştır.

Kuantalama tekniğinin performans üzerindeki etkisinin IRIS için incelendiği ikinci kısımda, klasik MÖ yöntemlerine ait sonuçlar Şekil 46'da verilmiştir. Genel olarak LAA yöntemi hariç tutulduğunda, kuantalama tekniğinin sınıflandırıcıların IRIS'teki performansını önemli ölçüde etkilemediği görülmektedir. LAA yönteminin YK yöntemine duyarlılığı ve bu deney özelinde kuantalanmış özniteliklerin YK uygulanarak spike örüntülerine dönüştürüldüğü göz önünde bulundurulduğunda, bu beklenen bir sonuçtur. Bu

nedenle LAA yönteminden ziyade diğer sınıflandırıcıların performansının değerlendirilmesi sağlıklı olacaktır.

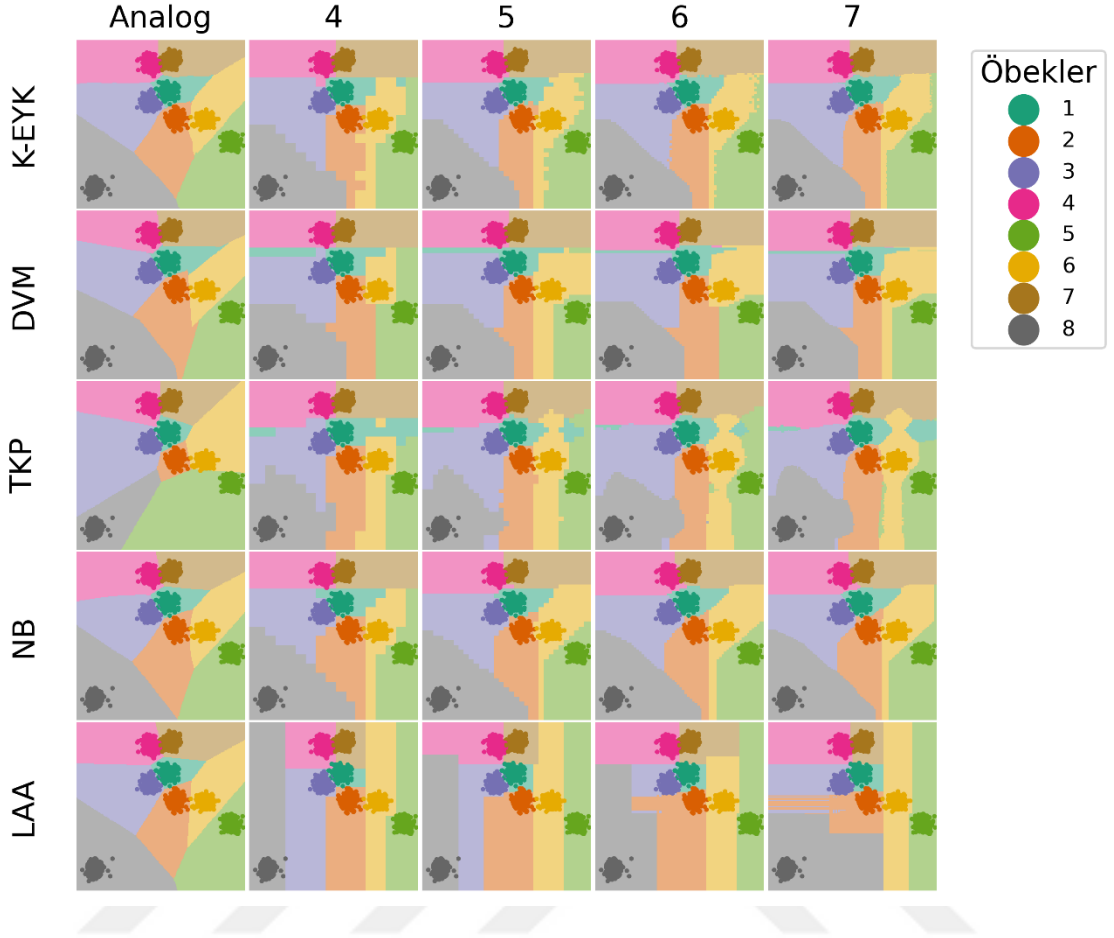
K-EYK yönteminin kuantalama teknikleri arasındaki farklardan etkilenmediği görülmektedir. Yöntem, tekdüze ve K-Ortalamlar ile kuantalanmış özniteliklerde analog özniteliklere göre sırasıyla %2.4 ve %1.2 performans artışı göstermiştir. Çeyreklik kuantalanmış özniteliklerde ise performans değişimi saptanamamıştır. Benzer durum NB yöntemi için de geçerlidir. Buna, tüm kuantalama teknikleri uygulanmış özniteliklerde performansta %2.4 düşüşe neden olmuştur.

TKP yöntemi için tüm kuantalama teknikleri performansı artıran sonuçlar doğurmuştur. Tekdüze, çeyreklik ve K-Ortalamlar ile kuantalanmış özniteliklerde performans sırasıyla %6.2, %7.4 ve %6.2 artmıştır. Burada, TKP yöntemi için ÖSD aşamasının performansı artıran bir etkiye sahip olduğu savunulabilmektedir. Buna karşın DVM yönteminde performans tüm kuantalama tekniklerinde düşmüştür. Buna göre tekdüze kuantalanmış özniteliklerde performans %4.9 düşerken, Çeyreklik kuantalanmış özniteliklerde bu düşüş %3.7 olarak kaydedilmiştir. K-Ortalamlar için ise performans %4.9 düşerek tekdüze kuantalanmış özniteliklere benzer bir sonuç elde edilmiştir.

3.2.3. Bit Uzunluğunun Etkisi

Bu bölümde, ÖSD aşamasında bir seçenek olarak sunulan bit uzunluğunun sınıflandırıcı performanslarını nasıl etkilediği klasik MÖ yöntemleri kullanılarak incelenmiştir. Toplam üç farklı VK'den ilki olan yapay VK, bu bölümde, sınıflandırıcıların eğitim sonunda öbekler için belirlediği karar alan ve sınırlarının görselleştirilmesi için kullanılmıştır. Diğer iki VK ise bit uzunluğunun sınıflandırma performansını nasıl etkilediğinin gözlemlenebilmesi adına kullanılmıştır.

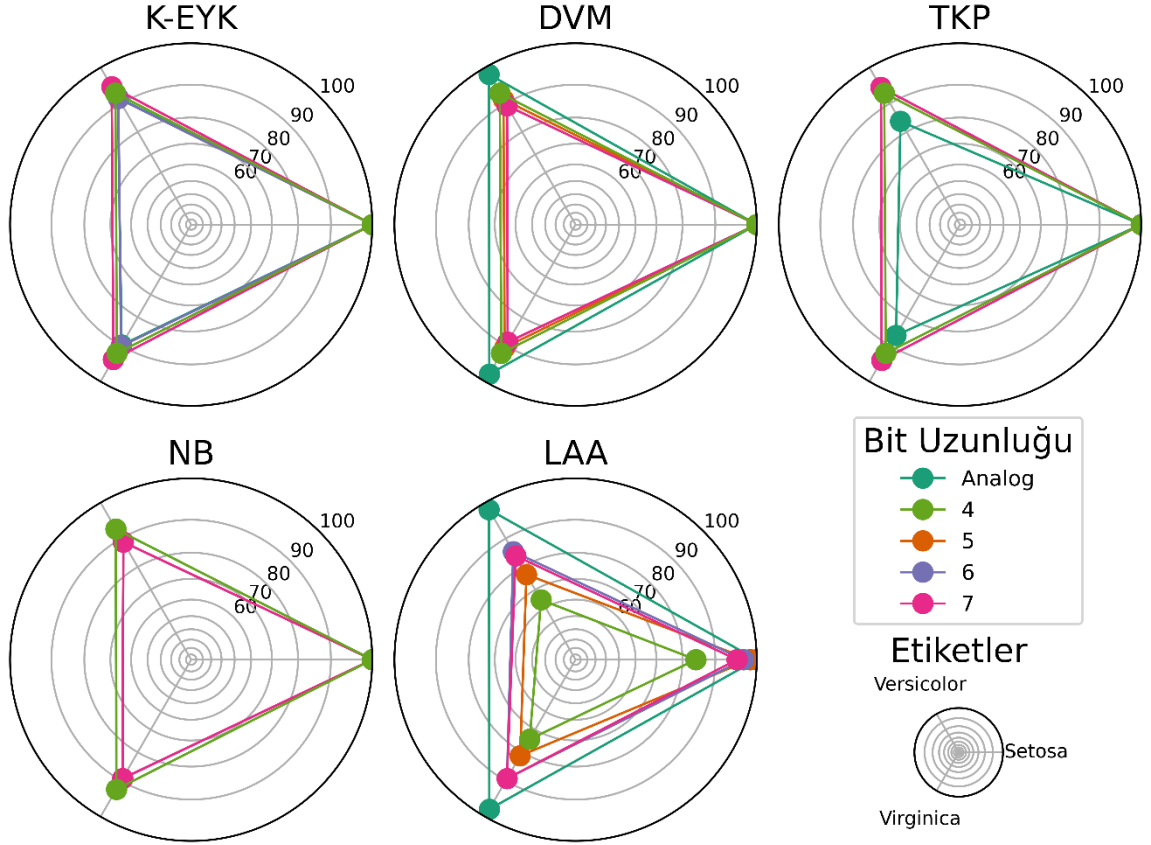
Klasik MÖ yöntemlerinin yapay VK için bit uzunluğuna bağlı olarak belirlediği ÖAH'ler Şekil 47'de görselleştirilmiştir. Burada, sütunlar bit uzunluğunu, satırlar ise klasik MÖ yöntemlerini temsil etmektedir. “Analog” olarak isimlendirilmiş sütunda öznitelik vektörlerinin kodlanmamış haline ait sonuçlar, diğer sütunlarda ise analog özniteliklerin ilgili bit uzunluklarında kuantalanmış ve kodlanmış hallerine ait ÖAH'ler gösterilmektedir. ÖSD aşamasındaki diğer faktörler olan kodlama yöntemi ve kuantalama tekniği, bu gösterim özelinde sırasıyla “Yoğunluk” ve “Tekdüze” olarak seçilmiştir.



Şekil 47. Bit uzunluğuna göre öbek alan haritaları

Şekil 47 detaylı incelendiğinde, bit uzunluğunun artmasıyla öbek alanları arasındaki sınırların geçişlerinin yumuşadığı görülmektedir. Bit uzunluğunun kuantal seviye artışına karşılık geldiği göz önünde bulundurulduğunda, bu oldukça beklenen bir durumdur. Kodlanmış ÖAH'ler ile analog öznitelik ÖAH'leri aralarındaki örtüşme oranı göz önünde bulundurulduğunda ise spike örüntüsü ÖAH'lerinin bit uzunluğunun artmasıyla birlikte analog ÖAH'lere daha çok benzediği görülmektedir.

K-EYK yöntemi özelinde bit uzunluğunun 4 ile 7 arası değişimi sırasıyla %86.0, %89.9, %92.1 ve %92.9 olduğu görülmüştür. DVM yöntemi için de benzer bir durum söz konusu olmakla birlikte, bu değerler sırasıyla %82.6, %85.3, %88.1 ve %88.1 olarak hesaplanmıştır. NB yöntemi de K-EYK ve DVM yöntemleriyle paralellik gösteren sonuçlar vermiş ve ÖAH örtüşme oranları sırasıyla %86.6, %91.7, %92.4 ve %92.5 olarak saptanmıştır.



Şekil 48. Bit uzunluğuna göre IRIS veri kümesinde performanslar

TKP ve LAA yöntemleri için de bit uzunluğunun artması örtüşme oranlarının artmasıyla sonuçlansa da örtüşme oran değerleri sınırlı kalmıştır. Bu yöntemler için bir uzunluğun 4 seçildiği durumda örtüşme oranları sırasıyla %60.1 ve %57.8 olarak hesaplanırken, bir uzunluğu 7 olarak seçildiğinde ise TKP yöntemi %64.1 ve LAA yöntemi %68.9 olan birer üst sınırdaki kısıtlanmıştır. Bu yönden bakıldığında, TKP ve LAA yöntemlerinin yapay VK için örtüşme oranlarının ÖSD aşaması ile oldukça düştüğü sonucuna varılmaktadır.

Bit uzunluğunun artması klasik MÖ yöntemlerinin ayrıca etiketleme hatası yapma miktarını da azaltmaktadır. K-EYK yönteminde bit uzunluğu 4 seçildiğinde toplam hatalı etiketleme miktarı 59 iken, bit uzunluğunun 6 ve 7 seçildiği durumda analog öznitelik sonuçları ile aynı sayıda olmak üzere toplam 9'ar adet hata yapılmıştır.

DVM, TKP, NB yöntemlerinde ise bit uzunluğunun artmasıyla analog öznitelikler kullanıldığı duruma göre daha başarılı etiketleme yapmanın mümkün olduğu görülmüştür. Analog özniteliklerin kullanıldığı durumda, toplam 10 hatalı etiketleme yapılan DVM

yönteminde, bit uzunluğu 4 seçildiğinde 15 hatalı etiketleme yapılsa da bit uzunluğunun 6 ve 7 seçildiği durumlarda etiketleme hatası sırasıyla 9 ve 7'ye düşmüştür.

TKP yönteminde bit uzunluğu 7 seçildiğinde toplam 8 adet örneğin hatalı olarak etiketlenmesi, analog özniteliklerde aynı değerin 13 olduğu göz önünde bulundurulduğunda, bir performans artışına karşılık gelmektedir. NB yöntemi için de benzer bir durum söz konusudur. Buna göre sınıflandırıcı bit uzunluğunun 6 ve 7 seçildiği durumlarda, analog özniteliklere göre sırasıyla 2 ve 3 adet daha fazla örneği doğru etiketlemiştir.

LAA yönteminin genel performans düşüklüğü önceki deneylerde sabit olsa da bit uzunluğunun artması, LAA yönteminin de etiketleme hatasını düşürmüştür. Bit uzunluğunun 4 seçildiği durumda toplam 40 adet örneği hatalı etiketleyen LAA yöntemi, bir uzunluklarının 5, 6 ve 7 seçildiği durumlarda sırasıyla 32, 17 ve 19 örneği etiketlerken hata yapmıştır. Analog özniteliklerde toplam 10 örneğin etiketlenmesinde hata yaptığı göz önünde bulundurulduğunda, LAA yönteminin ÖSD aşamasında performansının düşmesine karşın bit uzunluğunun artmasıyla performans düşüklüğü kısıtlanmıştır.

IRIS'ten elde edilen bulgular da bit uzunluğunun performansı artırdığını göstermektedir. Şekil 48'de verilen IRIS bulguları değerlendirildiğinde, bit uzunluğunun artmasıyla özellikle LAA yönteminde performansın arttığı görülmektedir. Bu artışı tüm sınıflandırıcılar için değişen miktarlarda görmek mümkündür.

TKP yönteminde, ÖSD aşamasının performansı artırması nedeniyle, bit uzunluğunun artmasıyla performans artışı yüksek bit uzunluklarında sınırlanmıştır. Bit uzunluğunun 4 olduğu durumda TKP yöntem performansı %4.9 artış gösterirken, daha yüksek bit uzunluklarında bu artış %6.2 olmak üzere sınırlanmıştır. Diğer tüm yöntemlerde, ÖSD aşamasının genel olarak bir performans düşüşüne neden olduğu göz önünde bulundurulmak kaydıyla, bit uzunluğunun artışı kodlanmış özniteliklerin kullanıldığı durumlar için performansta genel bir artışa neden olmuştur.

3.2.4. Tartışma

Şekil 42'den, Şekil 48'e kadar verilen bulgulara göre ÖSD aşamasında yapılan seçimler sınıflandırıcılar açısından çeşitli performans dalgalanmalarına neden olmaktadır. Buna bağlamda, ÖSD sonucu elde edilen öznitelikler kullanıldığında bazı sınıflandırıcıların performansı artarken, bazılarının ise performansı negatif yönde

etkilanmaktadır. Detaylı bir incelemeyle, her sınıflandırıcının ÖSD aşamasında yapılan seçimlerden farklı etkilendiği söylenebilmektedir.

ÖAH'lerin değerlendirilmesinde iki farklı ölçüt kullanılmıştır. İlk ölçüt, referans olarak seçilen analog öznitelikler için ÖAH'ler ve aynı özniteliklerin spike örüntülerine dönüştürülmüş versiyonlarına ait ÖAH'ler arasındaki örtüşme oranlarıdır. Bu ölçüt, ÖSD aşamasında öznitelik vektörlerinde bozulmanın ortaya koyulması amacıyla kullanılmıştır. Ancak özniteliklerin bozulmuş olması sınıflar arası ayrılabilirliğin azaldığı anlamına gelmemektedir. Bu nedenle kullanılan diğer ölçüt ise sınıflandırıcıların eğitim sonunda ne kadar hatalı etiketleme yaptığıdır. Bu ölçüt sayesinde ÖSD aşamasının performansı ne kadar etkilendiği belirlenmiştir.

ÖAH'ler her ne kadar ÖSD aşamasında özniteliklerin dönüştürülmesinin eğitimde ne gibi etkiye neden olduğunu aydınlatır da bir test aşamasında performansın ne kadar etkilendiğinin belirlenmesi ayrıca önemlidir. Test performansı, bir sınıflandırıcının genelleme başarısının göstergesidir. Test aşamasına sahip olmayan yapay VK deneyleri genelleme performansının belirlenmesi için yeterli değildir. Bu nedenle gerçek dünya VK'leri olan IRIS ve MNIST hem genelleme performansının ölçümünde hem de karşılaştırma amacıyla kullanılmıştır.

ÖAH örtüşme oranları, toplam yapılan hatalı etiket miktarları ve karşılaştırma VK bulgularına göre, kodlama yöntemi ÖSD aşamasında performansı en çok etkileyen faktör olmuştur. ÖAH'ler incelendiğinde, kodlama yönteminin değişmesiyle, öbek alanlarının çeşitli formlarda belirlendiği görülmüştür. YK kullanılan özniteliklerde sınıflandırıcılar analog öznitelik sınırlarına daha benzer sınırlar belirlemişlerdir. Buna göre YK kullanılan ÖAH'ler, analog ÖAH'ler ile ortalama %77.9 örtüşmektedir. LAA yönteminin genel olarak düşük performans gösterimi nedeniyle, LAA yöntemi hariç tutulduğunda bu oran %82.2'ye yükselmektedir. GK ve IK ile kodlanmış öznitelikler için bu değerler sırasıyla %52.8 ve %56.5 olarak hesaplanmıştır. Bu bulgulara göre özniteliklerin bozulmasının minimuma indirmek için yapay VK için YK kullanmak gerekmektedir.

Etiketleme hatasına göre ise sınıflandırıcılar analog özniteliklerde ortalama 10.8 hatalı etiketleme yaparken, YK kullanıldığında ortalama 14.5, GK kullanıldığında ise ortalama 35 adet örneği hatalı etiketlemiştir. Bu değer, İK için ise 97.8 adet olarak belirlenmiştir. Bu bağlamda, yapay VK performansına göre kodlama için kullanılacak en başarılı yöntemin yoğunluk kodlama olduğu söylenebilmektedir.

Karşılaştırma VK bulguları YK yönteminin diğer yöntemlere göre üstünlüğünü ayrıca göstermektedir. IRIS için sınıflandırıcıların analog özniteliklerde %94.5 başarı gösterdiği göz önünde bulundurulduğunda; YK için performansın %94.8, GK için %92.9 ve İK için %85.2 olması, YK yönteminin en başarılı kodlama yöntemi olduğunu vurgulamaktadır. Benzer durumu MNIST bulgularında da görmek mümkündür. Analog özniteliklerde ortalama %91.7'luk F1-Skor değerini yakalayan sınıflandırıcılar, özniteliklerin YK ile kodlanması sonucu ortalama %91.8 başarı göstermiştir. GK ve İK için bu değerler, %85.6 ve %88.5 olarak tespit edilmiştir. Sonuç olarak, bu bölümde sunulan bulgulara dayanarak çalışmanın ilerleyen deneylerinde kodlama yöntemi olarak yoğunluk kodlamanın kullanılması önerilmektedir.

Kuantalama yöntemi de performansı ayrıca etkilemektedir. ÖAH'ler incelendiğinde; tekdüze kuantalama kullanıldığında, sınıflandırıcılar için örtüşme oranının ortalama 82.2 olduğu; bu oranların çeyreklik için ortalama %79.2 ve K-Ortalamlar için ise ortalama %80.6 olduğu görülmektedir. Etiketleme hatası incelendiğinde ise tekdüze kuantalanmış öznitelikler için ortalama 14.5, çeyreklik için ortalama 9.3 ve K-Ortalamlar için ise ortalama 12.5 adet örnek hatalı etiketlenmiştir. Burada örtüşme oranları ve etiketleme hataları arasında çeşitlilik bulunmaktadır. Örtüşme oranlarına göre en başarılı kuantalama tekniği tekdüze iken, etiketleme hatalarına göre en başarılı teknik çeyreklik olarak tespit edilmiştir.

IRIS için kuantalama tekniğinin sınıflandırıcıların performansı önemli ölçüde etkilemediği görülmüştür. Buna göre, sınıflandırıcıları analog özniteliklerde %94.5 başarı gösterdiği IRIS için tekdüze kuantalama kullanılmış spike örüntüleri için ortalama performans %94.8 olarak hesaplanmıştır. Çeyreklik için bu değer %94.8 iken, K-Ortalamlar için ise %94.5 olarak hesaplanmıştır. Yapay VK ve karşılaştırma VK bulguları göz önünde bulundurulduğunda, çalışmanın ilerleyen aşamalarında kuantalama tekniği olarak tekdüze yaklaşımının kullanılması önerilmektedir.

Bit uzunluğunun artmasıyla performansın genel olarak arttığı tespit edilmiştir. Buna göre bit uzunluğunun 4, 5, 6 ve 7 seçildiği durumlarda alan örtüşme oranları sırasıyla ortalama %78.8, %82.2, %84.3 ve %84.4 olarak hesaplanmıştır. Buna göre bit uzunluğunun artmasıyla özniteliklerin daha az bozulmaya maruz kaldığı söylenebilmektedir. Etiketleme hataları incelendiğinde de benzer bir durum ile karşılaşmıştır. Buna göre bir uzunluğunun 4 ile 7 arası değişen değerler alması durumunda sınıflandırıcılar sırasıyla ortalama 33.3, 14.5, 9.3 ve 8 adet örneği hatalı

etiketlemişlerdir. Yapay VK bulguları bit uzunluğunun artması ile performans arasında pozitif bir ilişki olduğunu göstermektedir.

IRIS bulguları incelendiğinde ise bit uzunluğu ile performans arasında anlamlı bir ilişkiye rastlanılmamıştır. Bunun başlıca nedeni olarak IRIS için performansın üst sınırına ulaşılmış olması gösterilebilmektedir. Buna göre IRIS hali hazırda olabilecek en başarılı şekilde sınıflandırıldıktan sonra bit uzunluğunun artması performansın daha da artmasını sağlayamayacaktır. Bu nedenle, IRIS bulguları bu konuda yeterli bilgiyi sağlamaktadır.

Burada sunulan bulgular, klasik MÖ yöntemlerinin ÖSD aşamasında ne kadar etkilendiğinin ortaya koymaktadır. Klasik MÖ yöntemlerinin GN modellerinin aksine analog değerlere sahip özniteliklerle de sınıflandırma yapabilmesi nedeniyle bu bölüm sadece MÖ yöntem performanslarını içermektedir. Klasik MÖ bulguları göz önünde bulundurulduğunda, en başarılı kodlama yönteminin YK, kuantalama tekniğinin tekdüze olduğu görülmüştür. Aynı deneyler, Hodgkin&Huxley (HH), Izhikevich (IZH) ve Sızıntılı Topla&Ateşle (STA) nöron modellerinden kurulu GSA'lar ile de gerçekleştirilmiş ve deneylere dair bulgular EK-3 başlığı altında verilmiştir.

D2 senaryosunda performansa etkisi incelenmiş olan ÖSD aşamasının öncül bir versiyonu, “Journal of Computational Neuroscience” dergisinde “Improving a cortical pyramidal neuron model’s classification performance on a real-world ecg dataset by extending inputs” başlıklı makalede yayımlanmıştır (Kayıkcıoğlu Bozkır vd., 2023). İlgili makalede gerçek dünya Elektrokardiyogram verilerinin spike örüntüsüne dönüştürülmesinde bit uzunluğu 4, kuantalama tekniği tekdüze ve kodlama yöntemi de Gray kodlama olarak seçilmiştir. Bu tez çalışması için ÖSD aşaması genişletilmiş, sıralanan seçenekler eklenmiş ve performansa etkileri incelenmiştir.

3.3. Gerçekçi Sinir Ağı Modeli Performansının Karşılaştırmalı İncelenmesi ve Sinaptik Bağlantı Tipinin Etkisi

Bu bölümde, detayları Bölüm 2.6.3’te verilen, D3 isimli deneysel senaryoya dair bulgulara yer verilmiştir. D3’te çeşitli VK’ler için gerçekçi nöron modellerinin performansları, çeşitli sinaptik bağlantı tiplerinin varlığında incelenmiştir. D3 deneysel senaryosuna dair detaylı bulgulara EK-4 başlığı altında ulaşılabilmektedir.

Bu senaryoda, gerçekçi nöron modelleri olarak HH, IZH ve STA olmak üzere toplam üç adet GN modeli kullanılmıştır. Senaryoda, ayrıca, sinaptik bağlantı tiplerinin

performansı nasıl etkilediğinin incelenmesi adına seçenekler çeşitlendirilmiştir. Buna göre aynalama yaklaşımının kullanıldığı veya kullanılmadığı, Exc SB'lerin kullanıldığı veya kullanılmadığı, Inh SB'lerin kullanıldığı veya kullanılmadığı modeller olarak toplam altı kombinasyon ayrı ayrı incelenmiş ve bu bağlantı tiplerinin etkisi ortaya koyulmuştur.

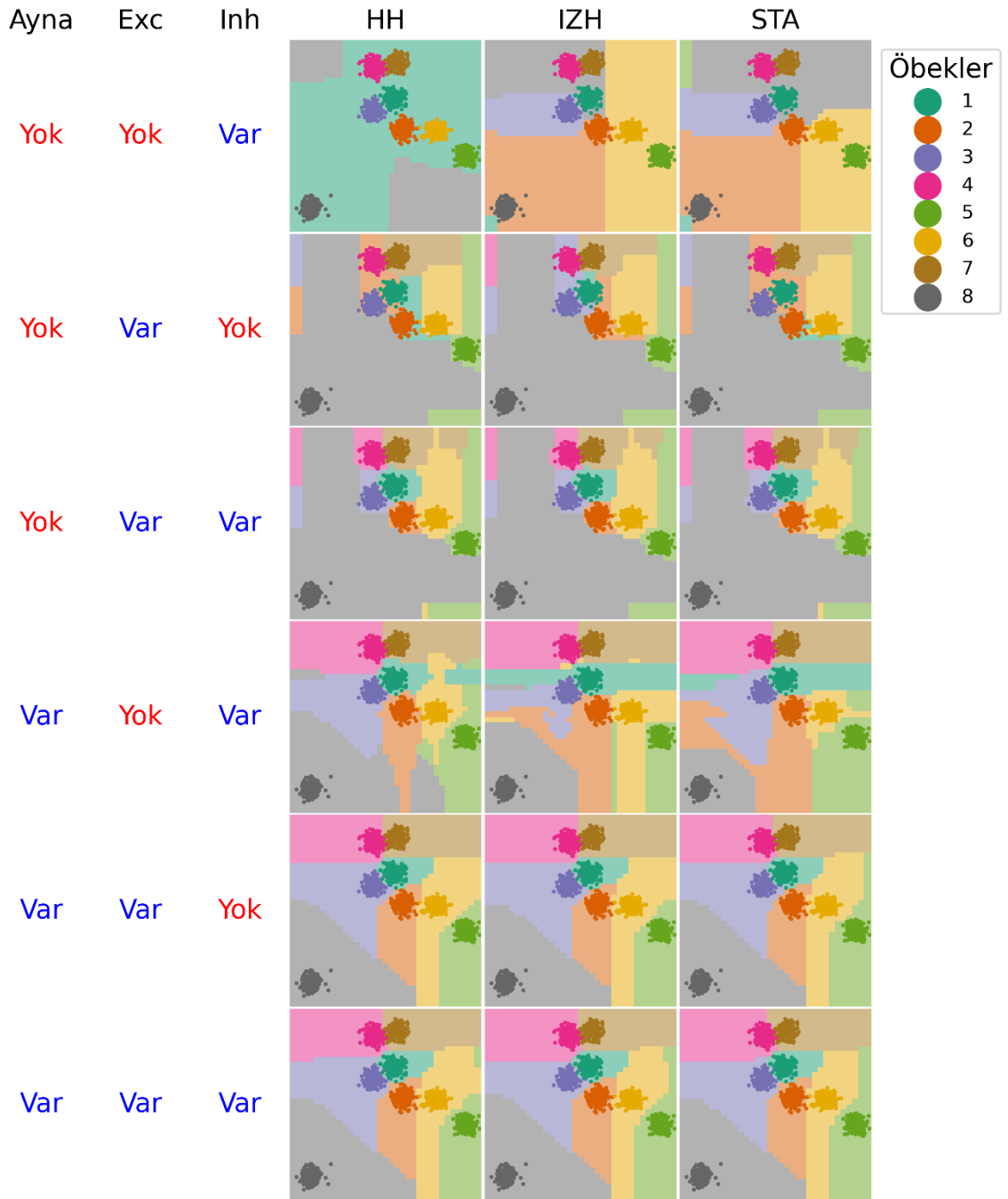
Bu bölüm dört alt başlıktan meydana gelmektedir. İlk alt başlıkta, yapay VK; ikinci alt başlıkta, IRIS çiçekleri VK; üçüncü alt başlıkta ise MNIST el yazısı rakamlar VK'lere dair bulgular verilmiştir. Dördüncü ve son alt başlıkta ise bulgular değerlendirilmiş ve tartışılmıştır.

3.3.1. Yapay Veri Kümesinde Öbek Alan Haritaları

Bu bölümde, SB tiplerinin varlığına dair tüm kombinasyonlar yapay VK kullanılarak incelenmiştir. Deneyler sonucu elde edilen ÖAH'ler Şekil 49'da gösterilmektedir. İlk bakışta, aynalama yaklaşımının varlığının ÖAH'leri doğrudan etkilediği görülmektedir. Buna göre, aynalama yaklaşımının varlığında öbek alanları daha dengeli bir yapıya sahipken; aynalama yaklaşımının yokluğunda bazı öbek alanların diğerlerine göre fazla yayıldığını göstermektedir. Öte yandan, Exc SB'lerin yokluğu ile GSA performanslarının dramatik bir şekilde düştüğü tespit edilmiştir.

Aynalama olmadığı durumda, sadece Inh SB'lerin kullanılması GN'lerin başarılı bir ÖAH belirlemede yeterli olmamaktadır. Buna göre, hiçbir GN modeli Öbekler 1, 4, 6 ve 8 için doğru öbek alanları belirleyememiştir. HH modeli özelinde ise GSA neredeyse tüm alanı Öbek 3 olarak belirlemiştir. Bu yönden Inh SB'lerin yalnız başına kullanılması GSA modelleri için yeterli olmamaktadır. Sadece Exc SB'lerin kullanılması durumunda ise GSA'ların performansı açık bir şekilde yükselse de GN modelleri hala Öbek 4 için doğru bir alan belirleyememiştir. Ayrıca, HH ve STA modellerinde Öbek 4'ün yanında Öbek 1 için de doğru bir alan belirlenememiştir.

Exc ve Inh SB'ler beraber kullanıldığında ise bu SB'lerin tekil kullanımındaki başarısızlığın önüne geçildiği görülmektedir. Buna göre, GN modelleri toplam 1600 adet örnekten, HH için 19, IZH için 19 ve STA için 18 örneği etiketlerken hata yapmıştır. Öbekler kendi arasında değerlendirildiğinde, tüm öbekler için ortalama 2 ila 3 örnekte hata yapılmış olmasına karşın, Öbek 4 için toplam 13 etiketleme hatası yapılmış olması; Öbek 4'te kaydedilen başarısızlığın ortadan kalkmadığını, sadece performansın yükseldiğine işaret etmektedir.



Şekil 49. Sinaptik bağlantı tiplerine göre öbek alan haritaları

Aynalama yaklaşımının çalışmaya dahil edilmesiyle performansın önemli ölçüde arttığı görülmektedir. Sadece Inh SB'lerin kullanıldığı durumda, HH modeli toplam 14, IZH modeli toplam 44, STA modeli ise toplam 30 adet örneği hatalı etiketlemiştir. HH modeli için etiketlemede yapılan hatalar öbek bazlı değerlendirildiğinde, hataların öbeklere dengeli dağıldığını söylemek mümkündür. Ancak IZH ve STA modellerinin özellikle Öbek

3 ve Öbek 4 örneklerini etiketlerken diğer öbeklere göre daha fazla hata yapmış olması, yine de sadece Inh SB'lerin kullanımını, aynalama yaklaşımına rağmen, yetersiz kılmaktadır.

Sadece Exc SB'lerin kullanıldığı durumda ise GSA'lar, sadece Inh SB'lerin kullanıldığı duruma göre daha başarılı bir performans göstermiştir. Tüm GN modelleri etiketleme yaparken toplam 18'er adet örnekte hata yapmıştır. Aynı performans tüm SB tiplerinin aynı anda kullanıldığı durumda da ortaya çıkmaktadır. Öbek bazlı değerlendirildiğinde ise tüm sınıflandırıcılarda hataların 13'er adeti Öbek 4 örneklerinde yapılmıştır.

3.3.2. IRIS Veri Kümesinde Sınıflandırma Performansı

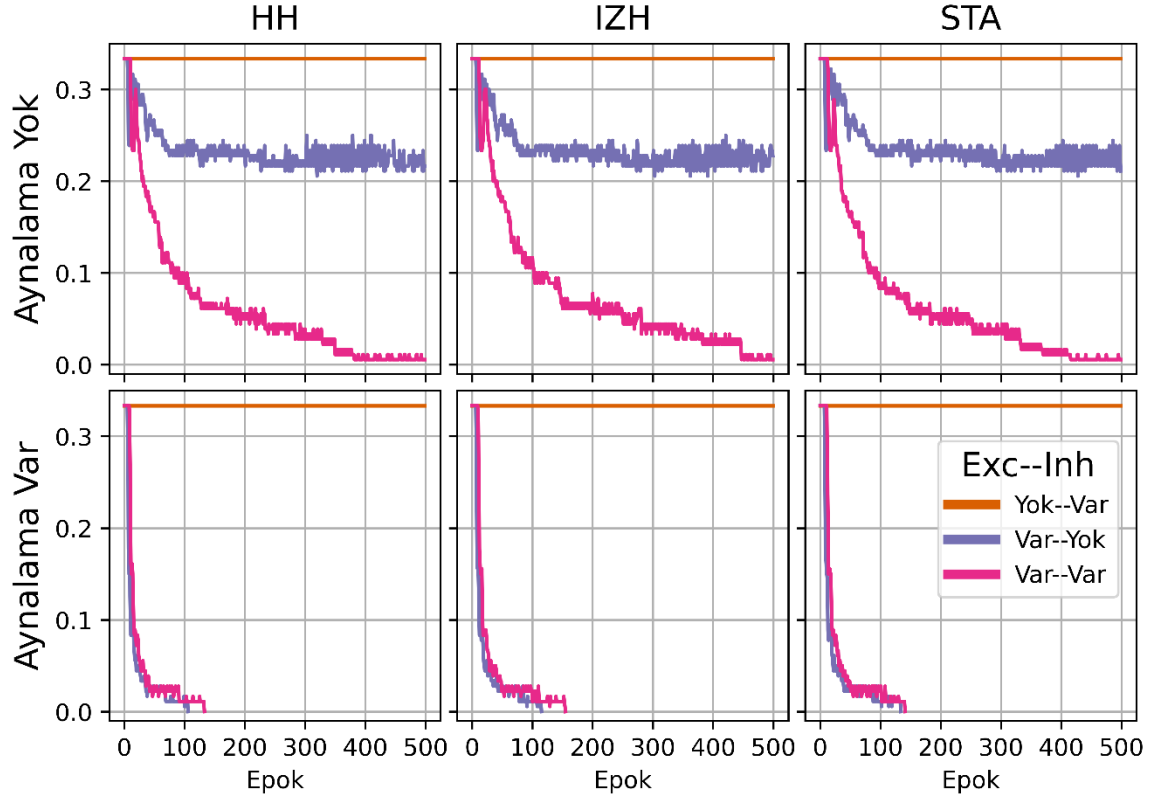
Bu bölümde sinaptik bağlantı tiplerinin varlığına dair tüm kombinasyonlar IRIS kullanılarak incelenmiştir. Önce seçilen kombinasyon şartları altında GSA modelleri eğitilmiş ve eğitim hatalarının seyri Şekil 50'de ortalama karesel hata (OKH) cinsinden verilmiştir. Ardından, GSA modelleri test edilmiş ve performansları Şekil 51'de verilmiştir.

Şekil 50 değerlendirildiğinde, tüm GN modelleri için Exc SB'lerin olmadığı durumda, GSA modellerinin eğitiminde OKH değeri başlangıçtaki değerde sabit kalmış ve model herhangi bir çözüme yakınsamamıştır. Bu bağlamda, Exc SB'lerin GSA modellerinin sınıflandırma performansındaki önemleri ön plana çıkmaktadır.

Bulgular, ayrıca, Inh SB'lerin GSA modellerinin eğitimini nasıl etkilediğini de göstermektedir. Aynalama olmadığı durumda ve sadece Exc SB'ler kullanıldığında, GSA modelleri 500 epok sonunda bir çözüme yakınsamamıştır. Exc ve Inh SB'lerin birlikte kullanıldığı durumda ise GSA modellerinin OKH'lerinde erken epoklardan itibaren gözle görülür bir düşüşün olduğu tespit edilmiştir. Bu yönüyle Inh SB'lerin de eğitim için önemli bir faktör olduğu görülmektedir.

IRIS için GSA modellerinin eğitim aşaması, aynalama yaklaşımının da önemini göstermektedir. Aynalama yaklaşımının dahil edildiği deneylerde GSA modelleri, dahil edilmeyen deneylere göre daha erken epoklarda bir çözüme yakınsamıştır. Bu yönüyle, aynalama yaklaşımının eğitim hızını artırdığı söylenebilmektedir. Aynalama yaklaşımının ayrıca Inh SB'lerin yokluğunu da telafi edebildiği görülmektedir.

Şekil 51'de verilen performans bulguları değerlendirildiğinde ise Exc SB'lerin kullanılmadığı durumda, GSA modellerinin eğitimdeki başarısızlıkları performansa

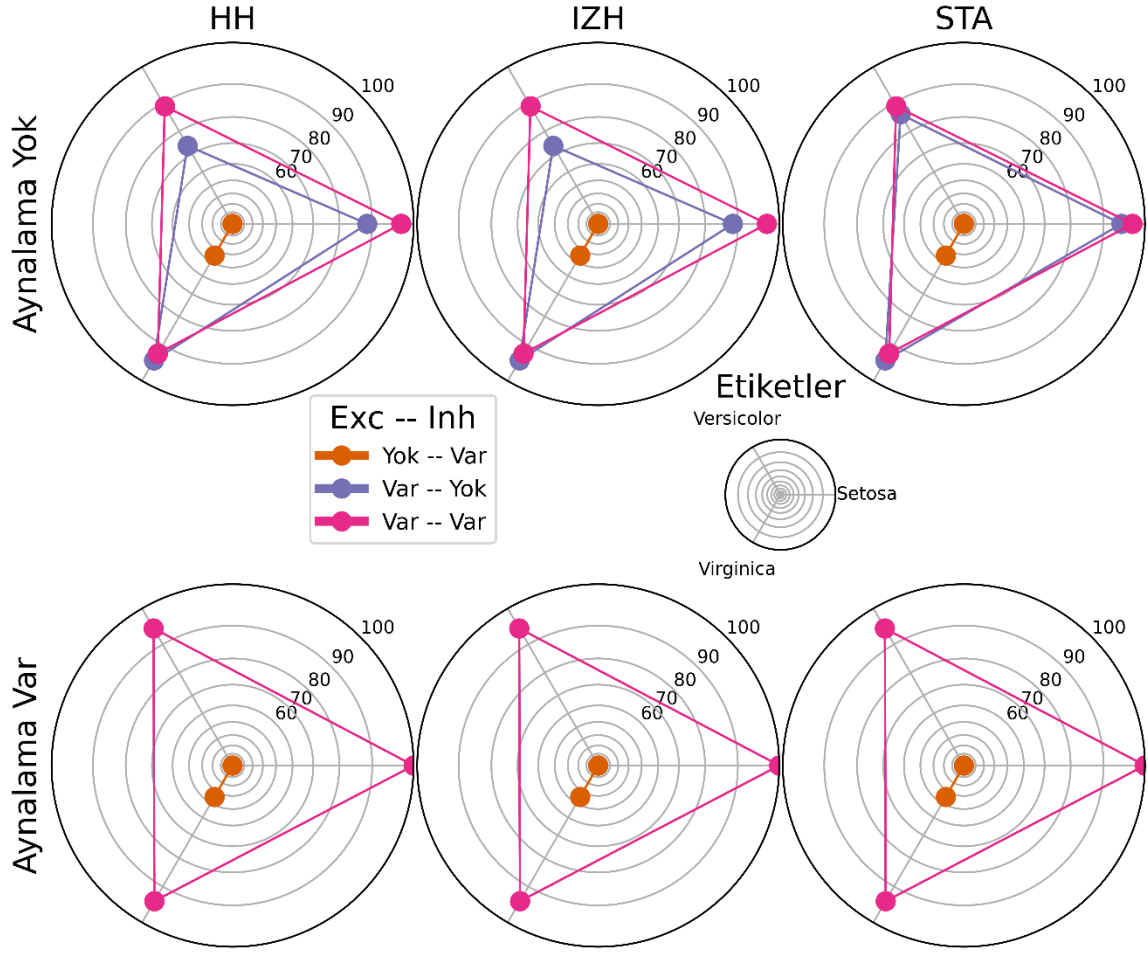


Şekil 50. Sinaptik bağlantı tiplerine göre IRIS veri kümesinde eğitim hataları

yansımıştır. Exc SB'lerin dahil edilmediği durumlarda, GSA modelleri IRIS örneklerini sınıflandıramamıştır.

Inh SB'lerin yokluğu, GSA modellerini kullanılan GN modelleri bazında farklı etkilemiştir. STA modeli için aynalama yaklaşımının dahil edilmediği durumda, performans %1.1 düşerken, HH ve IZH modelleri için bu düşüş %7.2 olarak kaydedilmiştir. HH ve IZH modellerinin STA modelinden daha fazla etkilenmesinin sebebi olarak versicolor ve setosa sınıfları özelinde gösterdikleri performans düşüşü gösterilebilmektedir.

GSA modellerinin performansları değerlendirildiğinde, aynalama yaklaşımının önemi açıkça görülmektedir. GSA'lar aynalamanın varlığında hem Inh SB'lerin eksikliğini telafi etmiş, hem de performansı önemli ölçüde artırmıştır. Buna göre tüm GN modelleri aynalamanın yokluğunda %92.9 başarı gösterirken, aynalamanın dahil edilmesiyle bu başarı %96.4'e yükselmiştir.

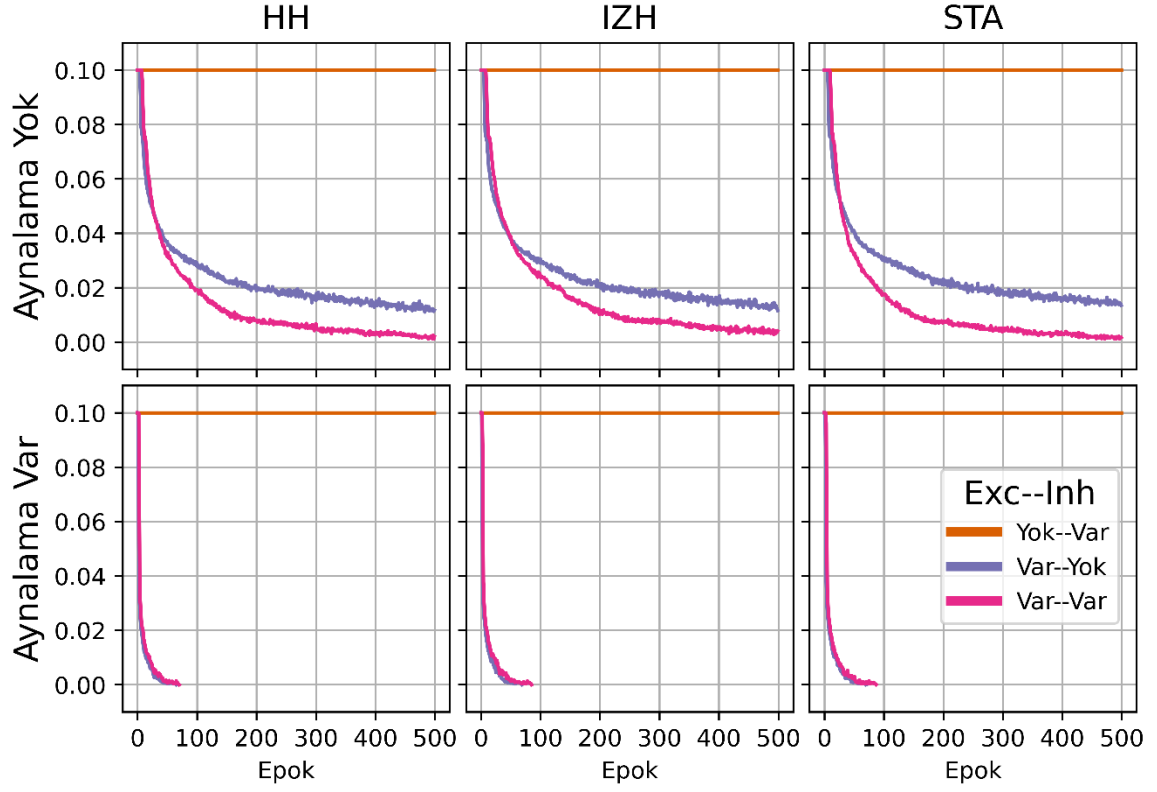


Şekil 51. Sinaptik bağlantı tiplerine göre IRIS veri kümesinde performanslar

3.3.3. MNIST Veri Kümesinde Sınıflandırma Performansı

Bu bölümde, sinaptik bağlantı tiplerinin varlığına dair tüm kombinasyonlar MNIST kullanılarak incelenmiştir. Önce seçilen kombinasyon şartları altında GSA modelleri eğitilmiş ve eğitim hatalarının seyri Şekil 52’de OKH cinsinden verilmiştir. Ardından, GSA modelleri test edilmiş ve performansları Şekil 53’te verilmiştir.

Şekil 52 değerlendirildiğinde, tüm GN modelleri için Exc SB’lerin olmadığı durumda, GSA modellerinin eğitiminde OKH değeri başlangıçtaki değerde sabit kalmış ve model herhangi bir çözüme yakınsamamıştır. Bu bağlamda, Exc SB’lerin GSA modellerinin sınıflandırma performansındaki etkinliğinin MNIST için de geçerli olduğu söylenebilmektedir.

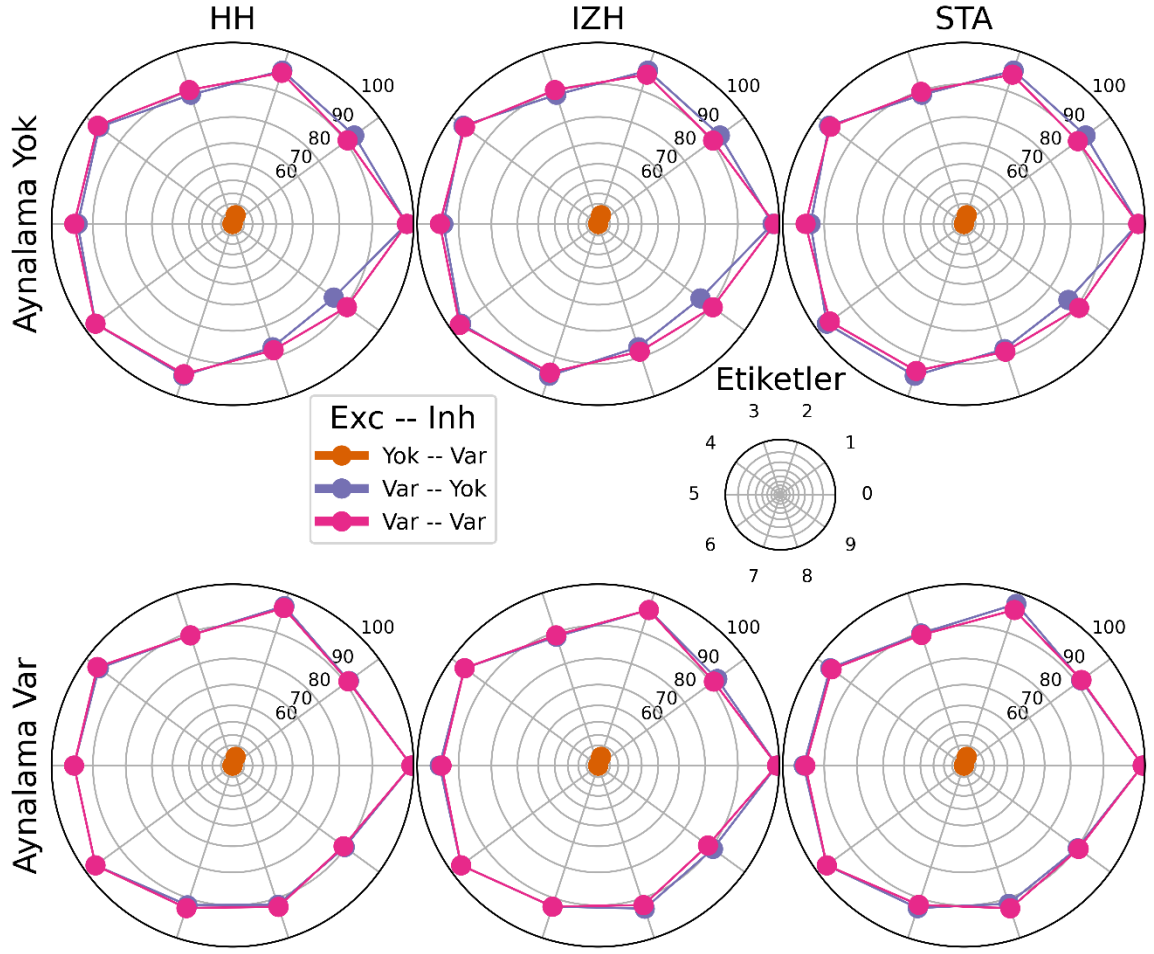


Şekil 52. Sinaptik bağlantı tiplerine göre MNIST veri kümesinde eğitim hataları

Öte yandan, MNIST için Inh SB'lerin etkisi de IRIS sonuçlarıyla paralellik göstermiştir. Aynalama yaklaşımının kullanılmadığı durumda açıkça görülebilen bu etkiye göre; Inh SB'lerin Exc SB'lerle beraber kullanılması, GSA'ların sadece Exc SB'lerin kullanıldığı duruma göre daha erken epoklarda bir çözüme yakınsamasına katkı sağlamıştır. Son olarak, aynalama yaklaşımının varlığı ise eğitim aşamasını pozitif yönde etkilemiştir. Aynalama yaklaşımı kullanıldığı durumda, GSA modellerinin hepsi bir çözüme yakınsamak için 100 epoktan daha az bir değere ihtiyaç duymuştur.

Şekil 53'te verilen performans bulguları değerlendirildiğinde ise Exc SB'lerin kullanılmadığı durumda, GSA modellerinin eğitimdeki başarısızlıkları performansa yine yansımıştır. Exc SB'lerin dahil edilmediği durumlarda, GSA modelleri MNIST örneklerini de sınıflandıramamıştır. Bu bağlamda karşılaştırma VK bulguları, Exc SB'lerin GSA modellerinde vazgeçilmez bir bileşen olduğunu göstermektedir.

Inh sinaptik bağlantıların varlığı, sınıflandırma performansını eğitimdeki kadar etkilememiştir. Inh sinaptik bağlantıların kullanılmadığı durumda, GSA modelleri için eğitimde 500 epok yeterli olmamasına rağmen, bu durum performansa beklenen ölçüde



Şekil 53. Sinaptik bağlantı tiplerine göre MNIST veri kümesinde performanslar

yansımamıştır. Inh SB'lerin varlığı performansı anlamlı bir ilişki olacak şekilde etkilememiştir. Ancak aynalama yaklaşımının kullanılmadığı durumda, 1 ve 9 rakamlarının sınıflandırılmasındaki performans düşüşü, Inh SB'lerin önemini göstermektedir.

MNIST sonuçları aynalama yaklaşımı özelinde diğer VK'ler ile aynı önemi işaret etmektedir. Aynalama yaklaşımının kullanılmadığı durumlarda GSA modellerinin sınıflandırma performansı yaklaşık %0.1 ila %0.9 arası artmıştır. Buna göre, diğer tüm VK'ler için de görülebileceği üzere, aynalama yaklaşımı sınıflandırma performansının artması yönünde doğrudan etkilidir.

3.3.4. Tartışma

Şekil 49'dan, Şekil 53'e kadar verilen bulgulara göre sinaptik bağlantı tiplerinin sınıflandırma performansını önemli ölçüde etkilediği görülmektedir. Bu konudaki ilk bulgu

Exc sinaptik bağlantıların GSA modellerinin bir sınıflandırma probleminde kullanılması konusunda hayati önemidir. Exc sinaptik bağlantıların kullanılmadığı durumda GSA modelleri eğitim aşamasını bile tamamlayamamıştır.

İkinci bulgu, Inh sinaptik bağlantıların eklenmesiyle, GSA performanslarının artmasıdır. Inh sinaptik bağlantıların GSA modelleri için önemi nörobiyolojik olarak detaylıca araştırılmış olsa da bu bağlantıların sınıflandırma problemlerini nasıl etkilediği literatürde hala yeterli ilgiyi görmemiş bir konudur. GSA modelleri ile sınıflandırma çalışmaları genelde Inh sinaptik bağlantıları ihmal etmektedirler. Bu nedenle sinaptik ağırlıkların pozitif olma kısıtına sahip olan GN modelleri için Inh sinaptik bağlantıların yokluğu, Chapeton ve arkadaşlarının (2012) da değindiği üzere, sınıflandırma kapasitesini düşürmektedir. Burada elde edilen bulgulara göre Inh sinaptik bağlantıların eklenmesiyle bu kapasite düşüşü engellenmiştir. Bu yönüyle Inh sinaptik bağlantıların GSA modelleri için önemi ortaya koyulmuştur.

Üçüncü ve son bulgu ise aynalama yaklaşımının GSA performansının artırılmasındaki etkin rolüdür. Buna göre aynalama yaklaşımının kullanılmasıyla GSA modelleri hem eğitim aşamasında daha düşük epok değerlerinde bir çözüme yakınsamış hem de modellerin sınıflandırma performansları gözle görülür artışlar göstermiştir. Bulgular bu yanıyla incelendiğinde; Moldwin ve Segev'in (2020) çalışmasında aynalama yaklaşımının kullanılmaması nedeniyle; GSA modelinin olması gerekenden daha düşük performans gösterdiği savunulabilmektedir.

Aynalama yaklaşımının bir diğer katkısı ise Inh sinaptik bağlantıların yokluğunda baskılayıcı özellikleri kompanse etmesidir. Inh bağlantılar ve ayna yaklaşımının birbirleriyle aynı mekanizmaya sahip olmamasına rağmen, Inh bağlantıların eksikliğinin ayna yaklaşımının dahil edilmesiyle giderilebileceği görülmüştür. Chapeton ve arkadaşlarının (2012) ortaya koydukları, sinaptik ağırlık kısıtının sınıflandırma kapasitesinde yarattığı düşüş; aynalama yaklaşımı sayesinde telafi edilebilmektedir.

D3 senaryosunda performansa etkisi incelenmiş olan aynalama yaklaşımı, "Journal of Computational Neuroscience" dergisinde "Improving a cortical pyramidal neuron model's classification performance on a real-world ecg dataset by extending inputs" başlıklı makalede önerilmiş ve matematiksel temelleri aydınlatılmıştır (Kayikcioglu Bozkir vd., 2023). İlgili makalede, aynalama yaklaşımının perceptron öğrenme kuralının performansını artıracak matematiksel olarak ispatlanmış ve sadece Exc sinaptik

bağlantıların kullanıldığı bir piramidal nöron modelinin sınıflandırma performansını nasıl etkilediği incelenmiştir.

Esasında, aynalama yaklaşımına benzer bir yaklaşım daha önce Ji ve arkadaşlarının (2016) çalışmasında önerilmiştir. Ancak yazarlar ters bağlantı olarak isimlendirdikleri bu bağlantı tipinin performansı nasıl etkilediği konusunda matematiksel temellere değinmemişlerdir. Bu tez çalışması için önerilen ve yayınlanan çalışmanın perceptron öğrenme kuralında aynalama yaklaşımının performansa nasıl etki ettiğini matematiksel olarak göstermesi, literatürdeki bu boşluğu tamamlar niteliktedir. Ancak yazarların öneriler kısmında da bahsettiği üzere, çalışmada aynalama yaklaşımının Inh sinaptik bağlantıların varlığında performansı nasıl etkilediği aydınlatılmamıştır. Bu tez çalışmasında, Inh sinaptik bağlantılar GSA modellerinde kullanılarak ve performansı nasıl etkiledikleri incelenerek literatürdeki bu boşluk da doldurulmuştur.

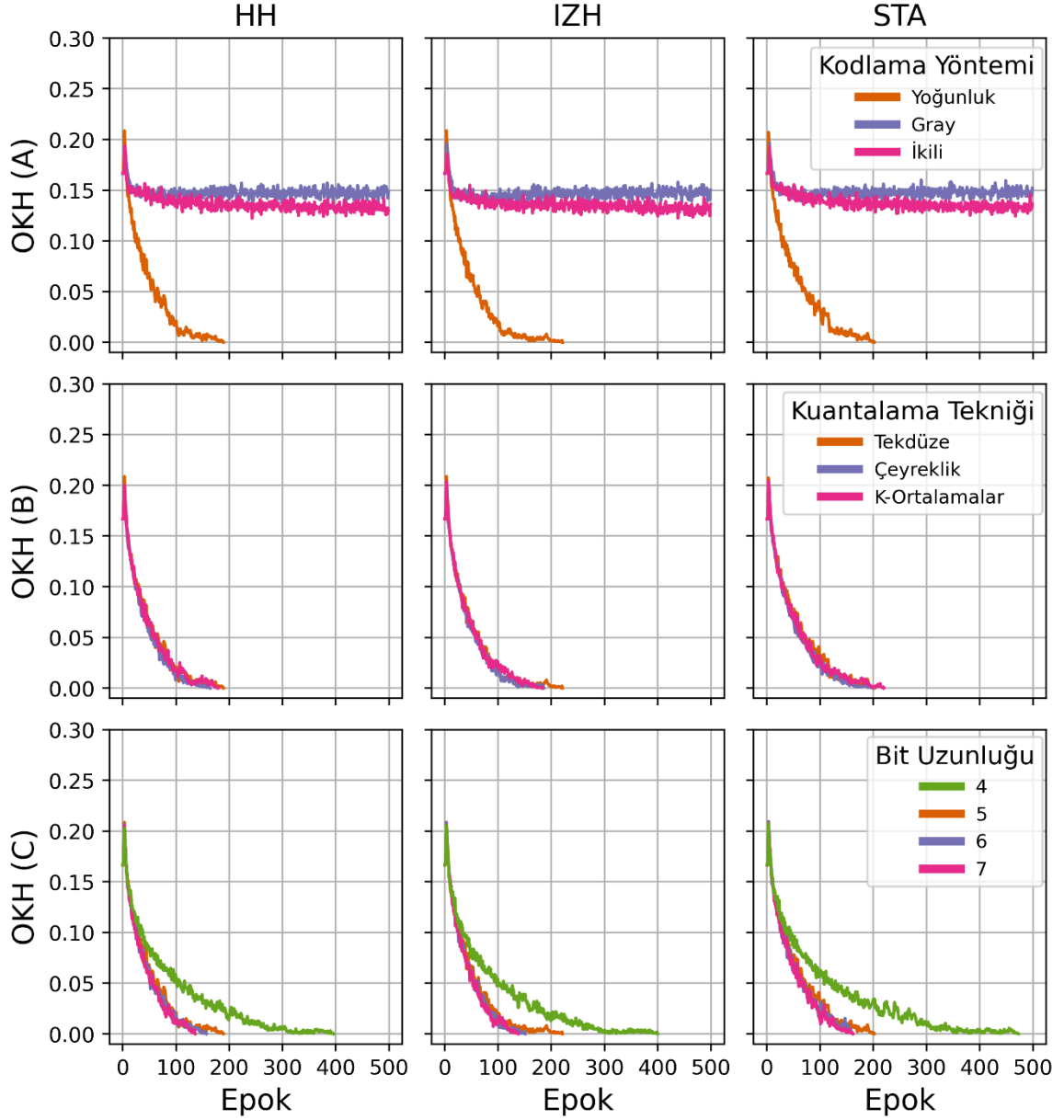
Sinaptik bağlantı tiplerinin performansa etkisi göz önünde bulundurulduğunda, Inh ve Exc sinaptik bağlantıların performansa pozitif etki ettiği burada sunulan incelemelerde belirlenmiştir. Ayrıca aynalama yaklaşımının da performansı önemli ölçüde artırdığı görülmüştür. Bulgular ilerleyen deneylerin tüm sinaptik bağlantı tiplerinin kullanıldığı durumlarda gerçekleştirilmesini önermektedir.

3.4. Otomatik Konuşmacı Tanıma Uygulamasında Gerçekçi Sinir Ağlarının Performansı

Bu bölüm altı alt başlıktan meydana gelmektedir. İlk alt başlıkta, ÖSD aşaması ve GSA kurulumunda seçilmesi gereken parametrelerin etkisi VK6 kümesinde incelenmiş ve uygun olan parametreler seçilmiştir. İkinci alt başlıkta, MFKK öznitelikleri kullanıldığında GSA'ların OKT performanslarına dair bulgular verilmiştir. Üçüncü ve dördüncü alt başlıklarda, hibrit yaklaşımla elde edilen özniteliklerin GSA'ların OKT performansını ne derece etkilediğine dair bulgulara yer verilmiştir. Beşinci alt başlıkta, PN'lerin OKT performansları incelenmiştir. Altıncı ve son alt başlıkta ise tüm bulgular bir arada değerlendirilerek tartışılmıştır.

3.4.1. VK6 Veri Kümesinde Uygun Parametrelerin Belirlenmesi

Bölüm 3.2 ve 3.3'te, çeşitli VK'ler için ÖSD aşamasının ve SB tiplerinin sınıflandırma performansını nasıl etkiledikleri incelenmiştir. Bulgular, ÖSD aşamasında



Şekil 54. ÖSD seçeneklerine göre VK6 veri kümesinde eğitim hataları

seçilmesi gereken parametrelerin ve GSA modelinde dahil edilmesi gereken SB tiplerinin VK'ye bağlı olduğunu göstermektedir. Bu nedenle, bir OKT uygulaması yapmadan önce konuşma VK'leri için uygun parametre ve yöntemlerin belirlenmesi gerekmektedir. Bu bölümde, D2 ve D3 senaryolarında VK6 kümesine ait bulgular verilmiş ve uygun parametreler araştırılmıştır.

Bölüm 3.2'de sunulan bulgulara göre önerilen ÖSD aşamasında uygun parametrelerin seçilmesiyle, VK'lerin başarılı bir şekilde sınıflandırılması mümkündür. Yapay VK bulgularına göre ÖAH'ler için en yüksek örtüşme oranı özniteliklerin çeyreklik

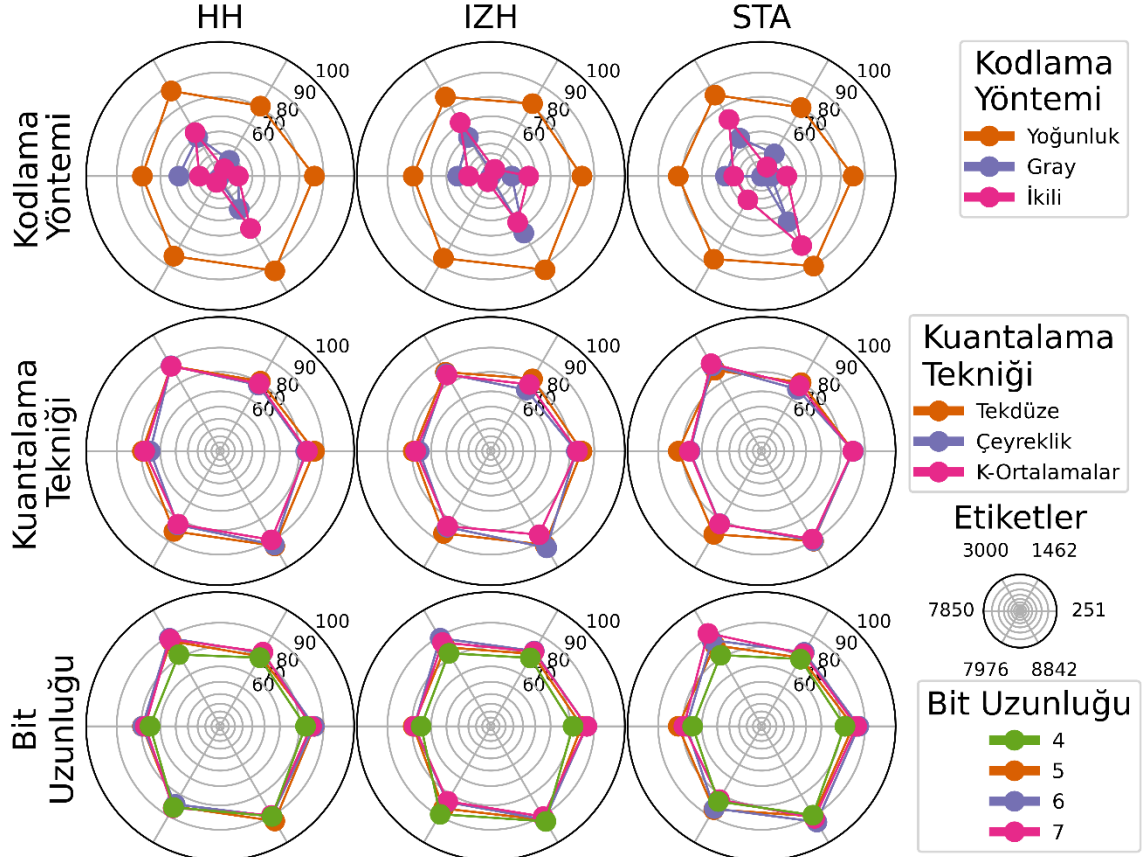
tekniki ile 5 bit uzunluğunda kuantalandığı ve YK kullanılarak kodlanması sonucu elde edilen spike örüntülerinde görülmüştür. IRIS bulguları, en yüksek performansın elde edildiği spike örüntülerinin, özniteliklerin 5 bit uzunluğunda tekdüze kuantalanması ve GK ile kodlanması sonucu elde edildiğini göstermektedir. MNIST bulguları ise en uygun kodlama yönteminin YK olduğuna işaret etmektedir. Bulgular, ÖSD aşamasında seçilmesi gereken bit uzunluğu, kuantalama tekniği ve kodlama yönteminin hem sınıflandırıcıya hem de VK'ye bağlı olduğunu göstermektedir. Şekil 54 ve Şekil 55'te GSA modellerinin VK6 kümesi performansları, ÖSD aşamasında seçilmesi gereken parametrelere göre karşılaştırmalı olarak verilmiştir.

Şekil 54(A)'da kodlama yöntemlerine göre eğitim hatalarının seyri gösterilmektedir. Buna göre YK yönteminin diğer yöntemlere göre açık bir üstünlüğe sahip olduğunu söylemek mümkündür. MFKK özniteliklerinin YK ile kodlanmasıyla elde edilen spike örüntüleri için GSA modelleri tüm GN modellerinde 200 epoktan önce sıfır hataya ulaşmışlardır. GK ve İK yöntemleri için ise 500 epok sayısı, GSA modellerinin bir çözüme yakınsaması için yeterli olmamıştır. Bu bağlamda MFKK öznitelikleri için en uygun kodlama yönteminin YK olduğu söylenebilmektedir.

Şekil 54(B)'de kuantalama tekniklerine göre eğitim hatalarının seyri gösterilmektedir. Buna göre bulgular teknikler arası herhangi bir üstünlüğü işaret etmemektedir. Tüm kuantalama teknikleri için GSA modelleri, 0.3 OKH değerinden başlayarak 200 epoktan önce sıfır hataya ulaşmışlardır. OKH eğrilerinin yaklaşık olarak üstü üste binmesi, yöntemler arasında eğitim aşaması için bir anlamlı farklılığın olmadığını göstermektedir.

Şekil 54(C)'de bit uzunluğuna göre eğitim hatalarının seyri gösterilmektedir. Buna göre bit uzunluğunun 4'te büyük olduğu durumlarda, OKH eğrilerinin yaklaşık olarak üst üste bindiği saptanmıştır. Ancak bit uzunluğunun 4 olarak seçildiği durumda, GSA daha geç epoklarda sıfır hataya ulaşmıştır. OKH eğrileri, bit uzunluğunun 4'ten büyük seçilmesi gerektiğine aksi durumda eğitim için ihtiyaç duyulan zamanın uzayacağına işaret etmektedir.

Şekil 55'te verilen performans bulguları incelendiğinde ise eğitim bulgularıyla paralel sonuçlar elde edilmiştir. GK ve İK kullanılan özniteliklerin eğitim aşamasında bir çözüme yakınsamamış olmasının sonucu, test performansları bu durumdan negatif etkilenmiştir. Buna göre GK ve İK yöntemiyle kodlanmış özniteliklerde test performansı HH model için sırasıyla 41.8 ± 18.2 ve 40.6 ± 20.3 olarak hesaplanmıştır. IZH ve STA



Şekil 55. ÖSD seçeneklerine göre VK6 veri kümesinde performanslar

modelleri için de benzer bir durumla karşılaşmıştır. IZH modeli için GK kodlama uygulandığında sınıflandırma performansının 39.1 ± 25.8 , İK kullanıldığında ise 44.3 ± 21.7 olduğu saptanmıştır. STA modeli için ise aynı değerlerin sırasıyla 40.8 ± 22.8 ve 51.8 ± 19.4 olduğu belirlenmiştir.

GK ve İK yöntemlerinin düşük performansına karşın, YK yöntemi uygulanmış özniteliklerde test performansının oldukça yüksek olduğu görülmüştür. Bu bağlamda YK yöntemi ile kodlanmış özniteliklerde HH modeli 85.4 ± 4.3 sınıflandırma performansı gösterirken, IZH ve STA modelleri için bu değerler 85.0 ± 3.9 ve 85.2 ± 3.3 olarak hesaplanmıştır. Sonuç olarak, MFKK özniteliklerinin kodlanmasında en uygun yöntemin YK olduğu sonucuna varılmıştır.

Kuantalama tekniği açısından test performansları incelendiğinde ise en yüksek performansa tekdüze kuantalama kullanılmış spike örüntülerinde rastlanmıştır. K-Ortalamlar ile kuantal seviyelerin belirlendiği durumda HH modeli 83.5 ± 4.1 test performansı gösterirken, IZH modeli için bu değer 82.5 ± 3.1 , STA modeli için ise 83.5 ± 4.9 olarak kaydedilmiştir. Çeyreklik yaklaşımıyla kuantalanmış spike

örüntülerinde ise HH, IZH ve STA modelleri sırasıyla 83.3 ± 5.4 , 82.7 ± 5.7 ve 83.1 ± 5.4 olarak hesaplanmıştır. Öte yandan, tekdüze kuantalanmış spike örüntülerinde test performansı diğerlerine göre ortalama yaklaşık %2 artarak HH modeli için 85.4 ± 4.3 , IZH modeli için 85.0 ± 3.9 ve STA modeli için ise 85.2 ± 3.3 olduğu tespit edilmiştir. Bulgular, MFKK özniteliklerinin spike örüntülerine dönüştürülmesinde en uygun kuantalama tekniğinin tekdüze olduğunu göstermiştir.

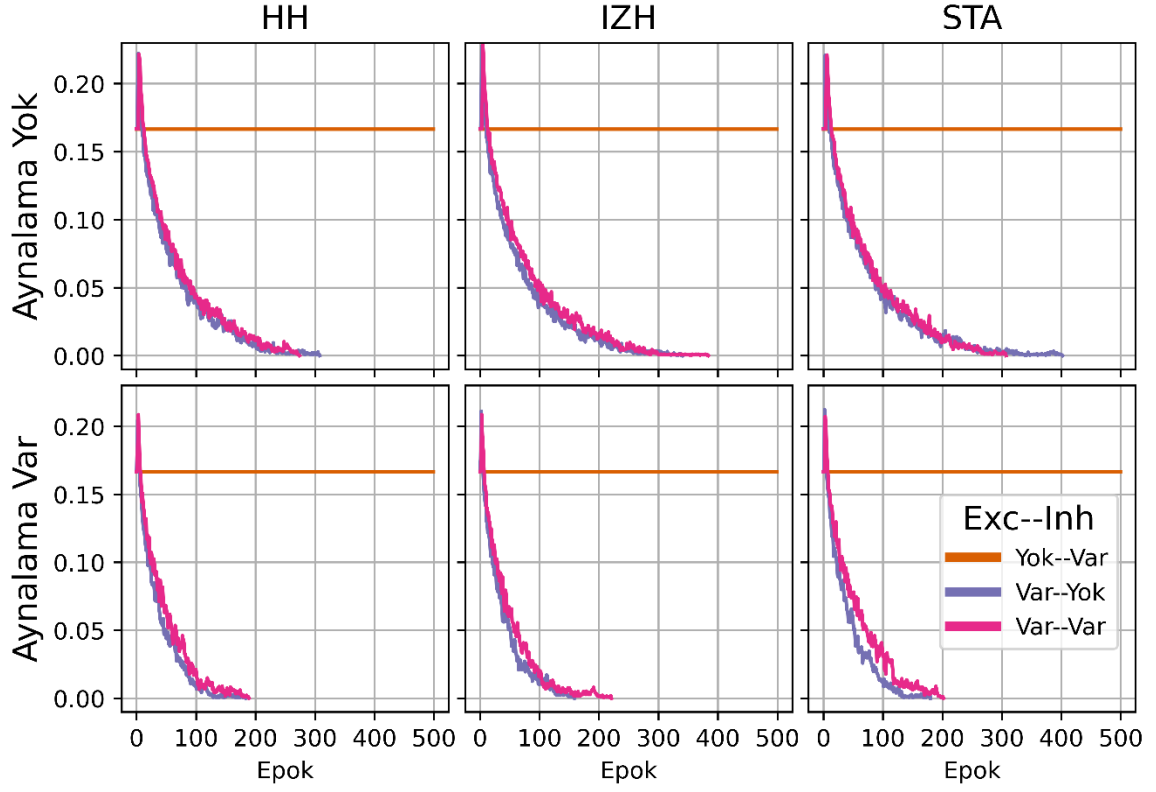
Bit uzunluğunun test performansını etkileyen bir diğer parametre olduğu görülmüştür. Şekil 55'te verilen bulgulara göre bit uzunluğunun 5 ve üzeri seçildiği durumlarda, sınıflandırma performanslarının bit uzunluğundan bağımsız olarak minimum değeri 85.0 ± 3.9 ve maksimum değeri ise 86.1 ± 4.5 olan bir aralıkta değiştiği tespit edilmiştir. Buna karşın, bit uzunluğunun 4 olarak seçildiği durumda ise HH modelinin 82.8 ± 4.6 , IZH modelinin 83.4 ± 5.7 ve STA modelinin 81.8 ± 4.4 test performansları göstermiş olması, bit uzunluğu için 4 değerinin yetersiz kaldığını göstermektedir. Bu nedenle, MFKK özniteliklerinin spike örüntülerine dönüştürülmesi sürecinde bit uzunluğunun en az 5 olarak seçilmesi gerektiği sonucuna varılmıştır.

Bölüm 3.3'te sunulan bulgular, tüm SB tiplerinin GSA performansını artırma yönünde etkilerinin olduğunu göstermektedir. Buna göre Exc SB'lerin yokluğunda sınıflandırma mümkün olmamakta, Inh SB'lerin yokluğunda ise sınıflandırma performansının nispeten düşük olduğu saptanmıştır. Ayrıca aynalama yaklaşımı, sınıflandırma performansının artmasında önemli bir etkiye sahiptir. Ancak, bulgular SB'lerin etkisinin VK'ye bağlı olacak şekilde değiştiğini de göstermektedir. Şekil 56 ve Şekil 57'de GSA modellerinin VK6 kümesi performansları, kullanılan sinaptik bağlantı tiplerine göre karşılaştırmalı olarak verilmiştir.

Şekil 56'da Exc SB'lerin GSA modelleri için hayati önemi açıkça görülmektedir. Buna göre Exc SB'lerin GSA modellerine dahil edilmediği durumlarda, GSA'lar bir çözüme yakınsamamış hatta OKH değerleri 500 epok boyunca değişmemiştir. Buna karşın Exc SB'lerin dahil edilmesiyle GSA modellerinin eğitim hataları bir çözüme yakınsama eğilimindedir.

Inh SB'lerin GSA modellerinin eğitim aşamasında anlamlı bir fark yarattığını söylemek mümkün değildir. Inh SB'lerin dahil edilmesine bağlı olmaksızın GSA modellerinin OKH eğrileri üst üste çakışmış ve sıfır hata değerlerine ulaşmışlardır.

Aynalama yaklaşımının önemi GSA modellerinin eğitim hatalarının seyrinde açıkça görülmektedir. Buna göre aynalama yaklaşımının dahil edilmesiyle GSA modellerinin sıfır

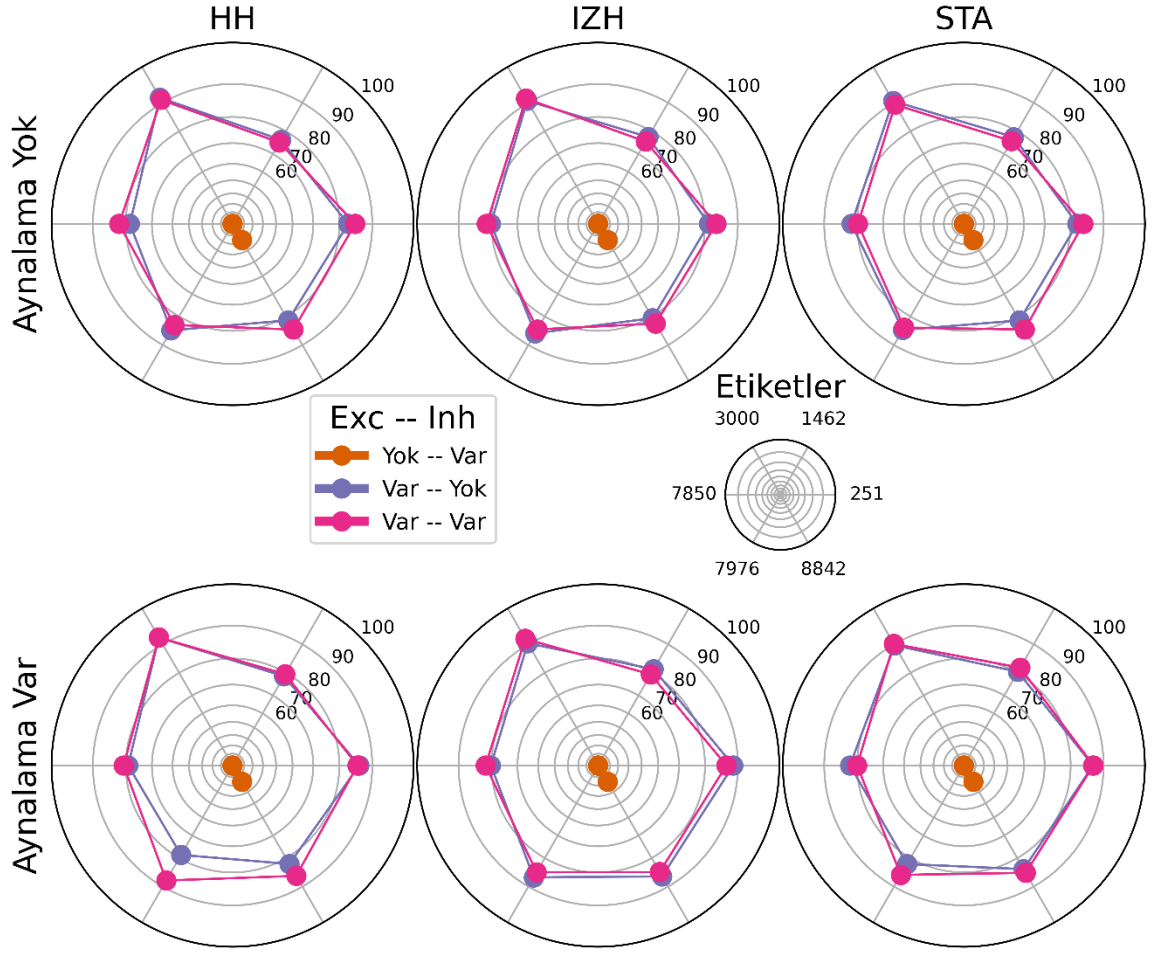


Şekil 56. Sinaptik bağlantı tiplerine göre VK6 veri kümesinde eğitim hataları

hataya ulaşması için ihtiyaç duyulan epok sayısı 100 ile 200 arası düşmüştür. Buna göre aynalama yaklaşımının eğitim performansını artırdığını söylemek mümkündür.

GSA modellerinin Şekil 57’de verilen test performansları incelendiğinde ise Exc SB’lerin dahil edilmediği durumlarda, eğitim aşamasına bir çözüme yakınsamayan GSA modellerinin test performanslarının da başarısız olduğu görülmüştür. Buna göre Exc sinaptik bağlantıların kullanılmadığı GSA modelleri VK6’yı sınıflandıramamıştır.

Eğitim aşamasında anlamlı bir etkisi gözlenmeyen Inh SB’lerin etkisi test performanslarında ortaya çıkmıştır. Inh SB’lerin dahil edilmesiyle; aynalamanın kullanılmadığı durumlarda, HH modeli için performansta %0.9, IZH modeli için ise %0.5 artış saptanmıştır. STA modeli için ise performansta %0.1 gibi küçük bir düşüş meydana gelmiştir. Aynalama yaklaşımının kullanıldığı durumlarda ise HH modelinde %2.5, STA modelinde ise %1.1’lik bir performans artışı gözlenmiştir. Ancak IZH modeli için ise %0.7’lik bir performans düşüşü belirlenmiştir. Buna göre Inh SB’lerin varlığının performansı artırmaya yönelik bir etkisi olabileceği söylenebilmektedir.



Şekil 57. Sinaptik bağlantı tiplerine göre VK6 veri kümesinde performanslar

Aynalama yönteminin, eğitim performansında olduğu gibi, test performansını da artırdığı sonucuna varılmıştır. Buna göre aynalama yaklaşımının dahil edilmesiyle HH modeli için $\% 83.6 \pm 4.6$ olan test performansı $\%85.4 \pm 4.3$ 'e yükselmiştir. Benzer şekilde IZH modeli için test performansı $\%83.3 \pm 4.6$ 'dan $\%85.0 \pm 3.9$ 'a, STA modeli için ise $\%83.0 \pm 4.2$ 'den $\%85.2 \pm 3.3$ 'e yükselmiştir. Test performansları incelendiğinde aynalama yaklaşımının sınıflandırma performansını artırma yönünde etkisinin yanında sınıflar arası performans dalgalanmalarını da azalttığı görülmektedir. Buna göre aynalama yaklaşımının dahil edilmesiyle performans ölçütlerinin standart sapmaları da düşmüştür.

3.4.2. MFKK Özniteliklerinde Gerçekçi Sinir Ağlarının Performansları

Bölüm 3.4.1'de MFKK özniteliklerinin spike örüntülerine dönüştürülmesi ve GSA modellerinin tasarımı için uygun parametreler araştırılmıştır. Elde edilen bulgulara göre

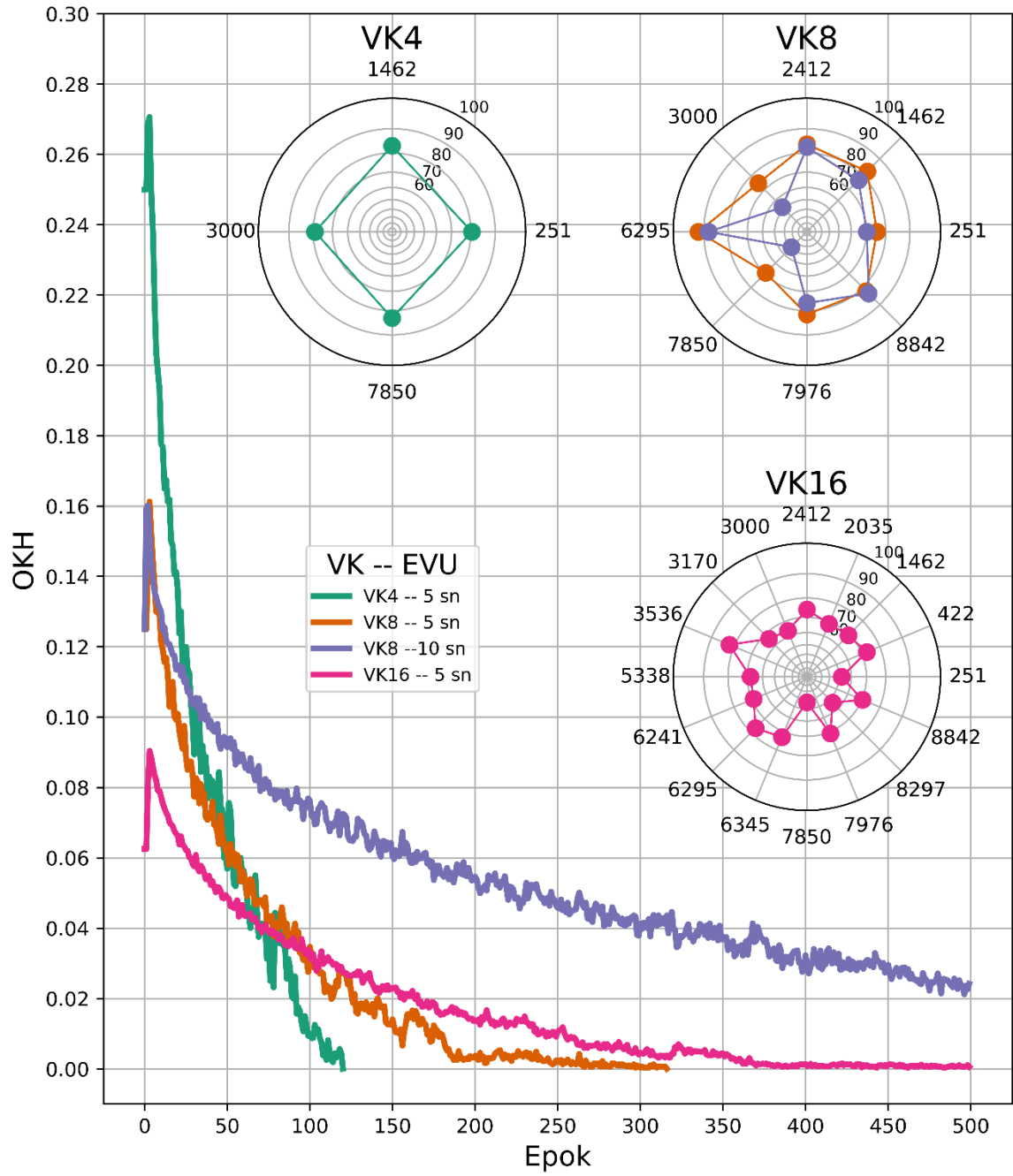
MFKK özniteliklerinin spike örüntülerine dönüştürülmesinde en uygun kodlama yönteminin yoğunluk kodlama, kuantalama tekniğinin tekdüze ve bit uzunluğunun ise 5 olduğu sonucuna varılmıştır. Ayrıca, GSA modellerinin tasarımında Exc ve Inh SB'lerin birlikte kullanılması önerilmiş ve aynalama yaklaşımının performansı önemli ölçüde artırdığı gözlemlenmiştir. Bu bulgulardan yola çıkılarak, D4 deneysel senaryosu tasarlanmıştır.

Bu bölümde, detayları Bölüm 2.6.4'te verilen, D4 isimli deneysel senaryonun ilk adımına dair bulgulara yer verilmiştir. D4'ün ilk adımında, konuşmacı sayısı itibariyle birbirinden farklı üç adet VK'den elde edilen MFKK öznitelikleri Bölüm 3.4.1'de önerilen şartlar altında spike örüntülerine dönüştürülmüş ve GSA modelleri ile sınıflandırılmıştır. Tüm VK'ler için eğitimde konuşmacı başına 5'er saniyelik kayıtlar kullanılmış ve eğitilen modellerin 1'er saniyelik konuşma kayıtlarında konuşmacı kimliklerini tanıması hedeflenmiştir. VK8 özelinde ise konuşmacı başına eğitim kayıtlarının uzunluğu iki katına çıkarılarak deney tekrar edilmiş ve sonuçları ayrıca değerlendirilmiştir. D4'e dair detaylı sonuçlar EK-5 başlığı altında paylaşılmıştır.

Şekil 58'de VK4, VK8 ve VK16 kümelerinden elde edilen MFKK öznitelikleri için GSA modellerinin eğitim hatalarının seyri ve test performansları verilmiştir. Şekil 58 oluşturulurken sadece HH nöronlarından kurulu GSA modellerinin bulguları kullanılmıştır. Ana grafik eğitim hatalarının seyrini ifade eden eğrileri gösterirken, şeklin içindeki radar grafikler ise konuşmacılar bazında test performanslarını göstermektedir.

Bulgulara göre eğitim aşamasında konuşmacı başına 5'er saniyelik kayıtlar kullanıldığında, VK4 için oluşturulan GSA modellerinin OKH'leri HH modeli için 121, IZH modeli için 171 ve STA modeli için ise 177 olmak üzere ortalama 156 epok sonunda 0 değerine ulaşmıştır. VK8 için oluşturulan GSA modelleri için ise konuşmacı sayısının artmasıyla sınıflandırma probleminin zorluk derecesi artmış ve OKH'lerin sıfır hataya ulaşmaları, HH modeli için 317, IZH modeli için 278 ve STA modeli için ise 459 olmak üzere ortalama 351 epok sürmüştür.

VK16 için eğitimin zorluğu daha da artmış ve konuşmacı sayısının artmasıyla birlikte 500 epok sayısı GSA modellerinin sıfır hataya ulaşmaları için yeterli olmamıştır. Buna göre 500'üncü epok sonunda HH nöronlarında kurulu GSA modelinin OKH değeri 5.4×10^{-4} 'te kalmıştır. Bu değer IZH ve STA nöronları için sırasıyla 1.6×10^{-4} ve 5.9×10^{-4} olarak kaydedilmiştir.



Şekil 58. MFKK öznelikleri kullanıldığında eğitim hataları ve performanslar

EVU değerinin iki katına çıkmasıyla, MFKK özneliklerinin kullanıldığı durumda GSA modellerinin eğitim performanslarının düştüğü tespit edilmiştir. Bu bağlamda eğitim sonunda hiçbir GN modeli için OKH değerinin sıfır hataya ulaşmadığı görülmektedir. EVU değerinin 5 saniyeden 10 saniyeye çıkmasıyla, sıfır hataya 200 epoktan önce ulaşan GN modelleri için eğitimde ihtiyaç duyulan epok sayısı artmıştır. Buna göre 500'üncü epok sonunda OKH değerleri HH modeli için 2.4×10^{-2} iken, IZH için 2.5×10^{-2} değerinde

kalmıştır. Aynı değer STA nöronlarından kurulu GSA modeli için 2.9×10^{-2} olarak kaydedilmiştir.

Eğitim aşamasının sonunda GSA modelleri 1'er saniyelik konuşma kayıtlarıyla test edilmiştir. Konuşmacı başına 5 saniyelik kayıtlarıyla eğitilmiş olan GSA modellerinde, VK4 için HH modelinde $\%81.5 \pm 1.8$, IZH modelinde $\%83.7 \pm 2.1$ ve STA modelinde ise $\%83.6 \pm 1.6$ olmak üzere test performansı ortalama $\%82.9$ olarak hesaplanmıştır. Konuşmacı sayısının artmasıyla VK8 için ise bu değer HH modelinde $\%80.3 \pm 6.5$, IZH modelinde $\%79.9 \pm 6.8$ ve STA modelinde $\%81.2 \pm 5.5$ olmak üzere ortalama $\%80.5$ olduğu saptanmıştır.

VK4 ve VK8 için performanslar karşılaştırıldığında ortalama performans değerinde yaklaşık $\%2$ 'lik bir düşüş olduğu görülmektedir. Bunun yanında sınıflar arası performans dalgalanmalarını ifade eden standart sapma değerleri artmıştır. Buna göre VK4 için test performansının standart sapması $\%1.6$ ila $\%2.1$ arası değişirken, VK8 için aynı değerler $\%5.5$ ila $\%6.8$ arası değişmektedir. Bu bağlamda sınıf sayısının artmasıyla sınıf bazlı performansların dalgalandığı görülmüştür. VK16 için test bulguları bunu destekler niteliktedir. Buna göre HH nöronlarının kullanıldığı GSA modelleri $\%67.3 \pm 9.5$ performans göstermiştir. Aynı değerler IZH ve STA nöronları için sırasıyla $\%68.0 \pm 11.2$ ve $\%64.5 \pm 11.1$ olarak hesaplanmış olup, tüm GSA modellerinin ortalama performansının $\%66.6$ olduğu saptanmıştır.

EVU değerinin iki katına çıkmasıyla eğitimde yaşanan performans düşüşü test performanslarını da etkilemiştir. Buna göre 5 saniyelik konuşma kayıtlarıyla eğitilmiş olan GSA modellerinin VK8 için gösterdikleri performans, EVU değerinin 10 olarak seçilmesiyle düşmüştür. Özellikle 3000 ve 7850 ID numaralarına sahip konuşmacılar için test performansının düşmesiyle tüm GSA modellerinin ortalama performansı $\%71.0$ olarak hesaplanmıştır. Bu bağlamda EVU değerinin iki katına çıkmasının ortalama performansta yaklaşık $\%10$ 'luk bir düşüşe neden olduğu belirlenmiştir.

3.4.3. Analog-Analog Dönüşümlü Özniteliklerde Gerçekçi Sinir Ağlarının Performansları

Bölüm 3.4.2'de verilen bulgulara göre MFKK özniteliklerinin kullanılmasıyla VK4 kümesinde ortalama $\%82.9$ test performansı kaydedilmiştir. Bu değerlerin VK8 ve VK16 kümeleri için ise sırasıyla $\%80.5$ ve $\%66.6$ olduğu göz önünde bulundurulduğunda, konuşmacı sayısının artmasıyla birlikte performansta önemli ölçüde düşüş saptanmıştır.



Ayrıca VK8 kümesi için EVU değerinin iki katına çıkmasıyla, test performansı yaklaşık %10 düşmüştür. Buna göre eğitim verisinin uzunluğunun artmasının MFKK öznitelikleri için zorlayıcı bir faktör olduğu savunulabilmektedir.

D1 senaryosunun ilk adımında elde edilen bulgulara göre geliştirme aşamasına sahip olan yaklaşımlar eğitim aşamasında diğer yaklaşımlara üstünlük kurmuştur. Bu bulgulara göre gerçekçi sinir ağlarına bir geliştirme aşamasının dahil edilmesiyle performansın

artabileceği hipotezi ortaya çıkmaktadır. Bu bağlamda D4 senaryosunda ikinci adıma geçilmiştir. Bu adıma göre analog değerlere sahip olan MFKK öznitelikleri bir analog-analog dönüştürücüye giriş olarak verilmiş ve çıkışları dönüştürülmüş öznitelikler olarak kabul edilmiştir. Dönüştürülmüş öznitelikler bir önceki adımda kullanılan MFKK öznitelikleri yerine kullanılmıştır.

D4'ün ikinci adımında, analog-analog dönüşümde A1 ve A2 öznitelikleri olmak üzere iki farklı öznitelik dönüştürme yaklaşımı kullanılmıştır. Bir çerçeve için 23 adet MFKK özniteliklerini giriş olarak kabul edip 23 adet öznitelik çıkışı veren A1 için bulgular Şekil 59'da verilmiştir. Aynı girişler için 46 adet öznitelik çıkışı veren A2 için bulgular Şekil 60'ta verilmiştir.

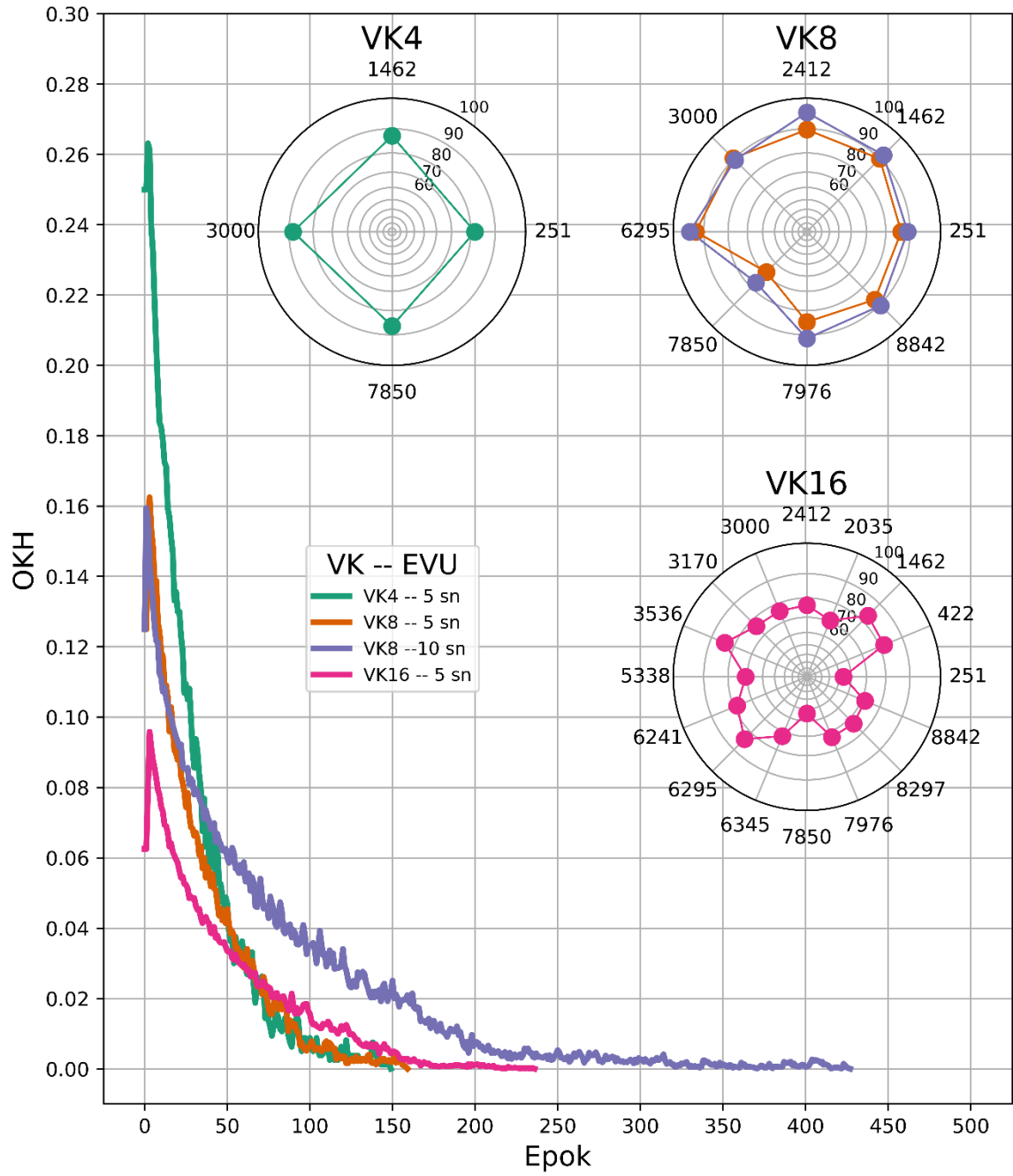
Şekil 59'a göre A1 öznitelikleri kullanıldığında GSA modellerinin sıfır hataya ulaşmaları için ortalama 191 epok yeterli olmuştur. VK16 için eğitim bulguları incelendiğinde ise GSA modelleri için eğitim sonunda OKH değerleri sıfır hataya ulaşmamışlardır. Buna göre 500'üncü epok sonunda GSA modelleri için ortalama OKH değeri 3.6×10^{-3} olarak hesaplanmıştır.

VK8 için EVU değeri 5 saniye olarak seçildiğinde GSA modelleri ortalama 363 epokta birer çözüm noktasına yakınsamışlardır. EVU değeri 10 saniye olarak seçildiğinde ise GSA modelleri için eğitimde 500 epok yeterli olmamıştır. Buna göre eğitim sonunda GSA modelleri için ortalama OKH değeri 3.6×10^{-2} olarak hesaplanmıştır.

A1 öznitelikleri bulgularına göre VK4 için test performansı MFKK özniteliklerine kıyasla önemli bir artış göstermiştir. Buna göre test performansında %5'lik bir artış kaydedilmiştir. Ancak benzer bir performans artışına VK16 için rastlanılmamıştır. Bu bağlamda A1 özniteliklerinin kullanılmasıyla test performansı MFKK özniteliklerine kıyasla yaklaşık %6 düşmüştür.

VK8 kümesi için A1 öznitelikleri kullanıldığında, 5'er saniyelik konuşma kayıtlarıyla eğitilmiş GSA modellerinde ortalama %80.5 test performansı saptanmıştır. Ancak EVU değerinin iki katına çıkmasıyla bu performans %71.0'a düşmüştür. A1 özniteliklerine dair tüm bulgular incelendiğinde, MFKK özniteliklerinden daha başarılı sonuçlar doğurmadığı ve MFKK öznitelikleri yerine kullanılmasının önerilemeyeceği söylenebilmektedir.

Öte yandan, Şekil 60'a göre A2 öznitelikleri A1 özniteliklerine kıyasla daha yüksek bir potansiyel sunmaktadır. A2 özniteliklerinin VK4 eğitiminde kullanılmasıyla, GSA modellerinin sıfır hataya ulaşmaları ortalama 156 epok sürmüştür. VK16 kümesi için A1



Şekil 60. A2 öznelikleri kullanıldığında eğitim hataları ve performanslar

özneliklerinin kullanılmasının eğitim performansı önemli ölçüde artırdığı görülmüştür. Buna göre MFKK ve A1 öznelikleri kullanıldığında GSA modelleri 500'üncü epok sonunda sıfır hataya ulaşamamışken, A2 özneliklerinin kullanılmasıyla ortalama 263 epokta modeller birer çözüme yakınsamışlardır. Bu bağlamda A2 özneliklerinin eğitim aşamasında ihtiyaç duyulan epok sayısını düşürdüğü söylenebilmektedir.

VK8 kümesinde A2 öznitelikleri kullanıldığı durumda bulgular incelendiğinde, 5 saniyelik konuşma kayıtlarıyla eğitilen GSA modellerinde eğitim ortalama 176 epoch sürmüş ve modeller birer çözüme yakınsamışlardır. EVU değerinin iki katına çıkmasıyla GSA modelleri MFKK ve A1 bulgularının aksine ortalama 400 epok sonunda sıfır hataya ulaşmışlardır. Bu yanılarıyla, A2 özniteliklerinin MFKK ve A1 özniteliklerine göre eğitim aşamasındaki üstünlükleri göze çarpmaktadır.

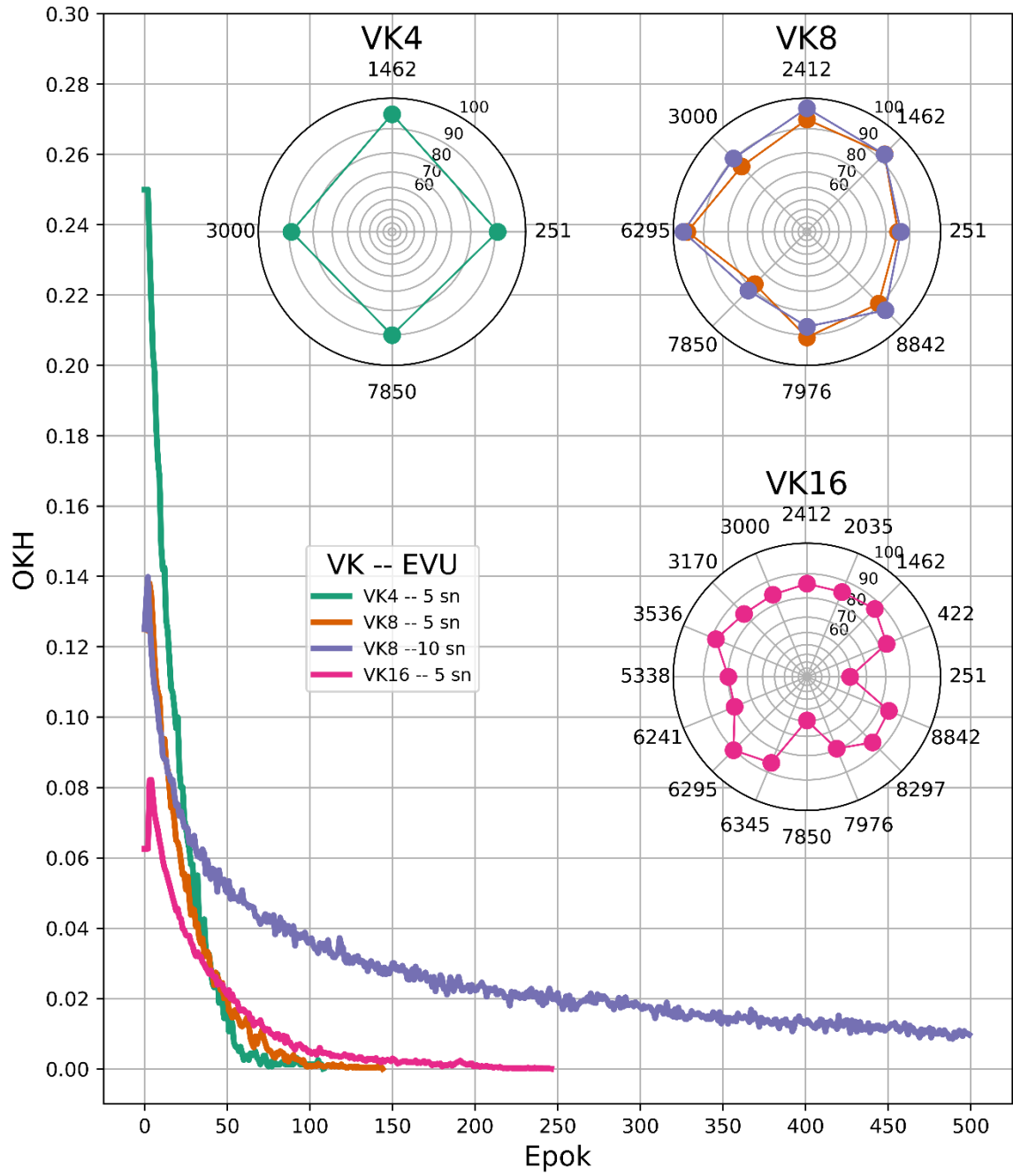
Test aşamasında; VK4 kümesinden elde edilen A1 öznitelikleriyle eğitilen GSA modelleri, ortalama %87.7 ile MFKK özniteliklerine göre yaklaşık %5'lik bir performans artışı sergilemiştir. Benzer artış VK16 için eğitilen GSA modelleri için de geçerlidir. A2 özniteliklerinin kullanılması MFKK özniteliklerine göre VK16 için yaklaşık %7.7 performans artışını mümkün kılmıştır.

Son olarak; VK8 için EVU değeri 5 saniye seçildiğinde elde edilen %85.6'lık performans, EVU değerinin 10 saniye seçilmesiyle %89.4'e yükselmiştir. MFKK ve A1 özniteliklerinde EVU değerinin artmasıyla performansın düştüğü göz önünde bulundurulduğunda, A2 özniteliklerinde EVU değerinin artması bir performans artışı ile sonuçlanmıştır.

3.4.4. Analog-İkili Dönüşümlü Özniteliklerde Gerçekçi Sinir Ağlarının Performansları

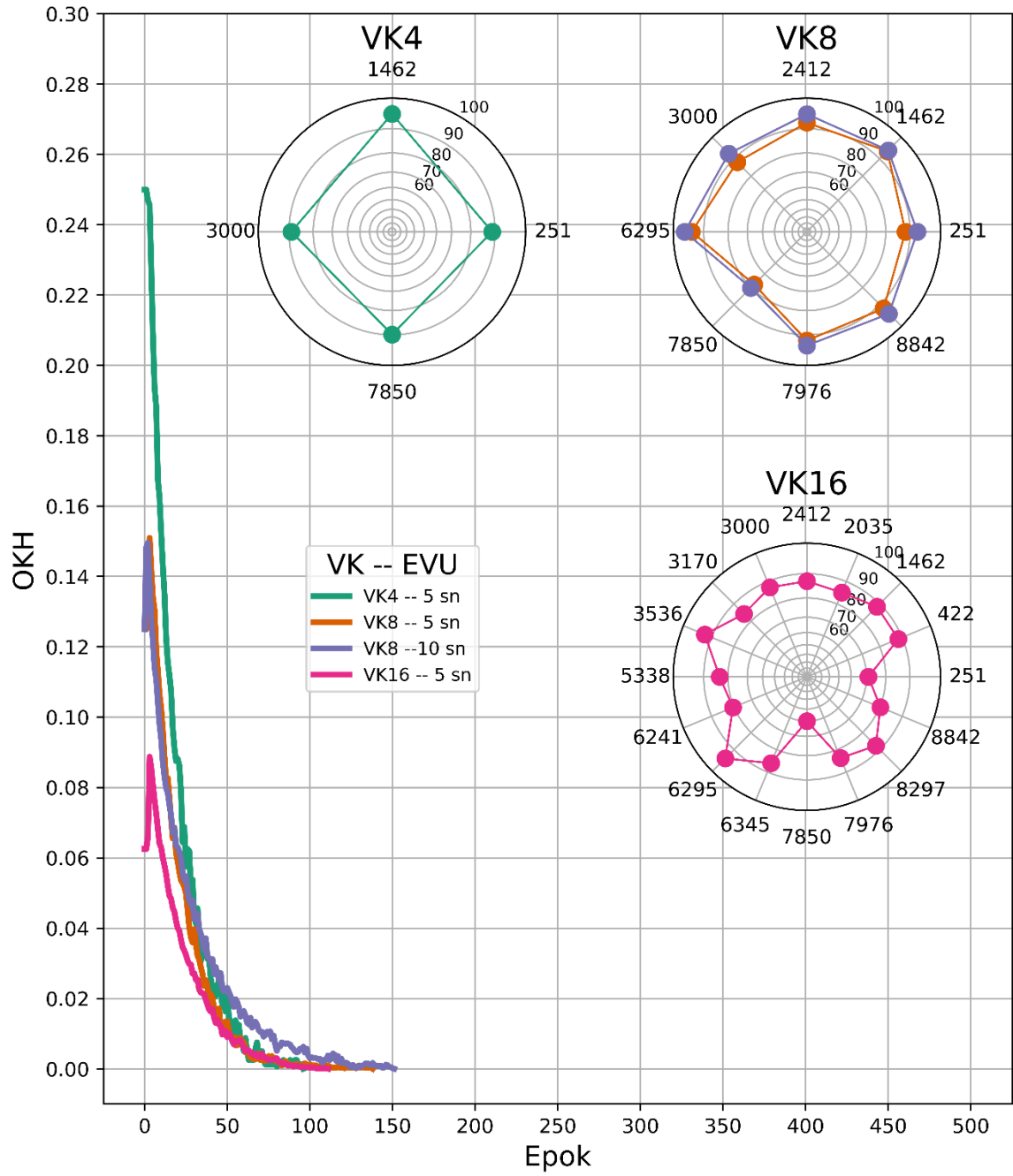
Bölüm 3.4.3'te elde edilen bulgular, geliştirme aşamasıyla gerçekçi sinir ağlarının OKT performansının artmasının mümkün olduğunu göstermiştir. Öte yandan, Bölüm 3.2'de elde edilen bulgulara göre özniteliklerin spike örüntülerine dönüştürülmesinde performansın düştüğü saptanmıştır. Bu iki bilgiden yola çıkıldığında, MFKK özniteliklerinden bir DSA yardımıyla doğrudan spike örüntüsü elde edilmesi, OKT performansını artırmaya aday bir yaklaşımdır.

D4 senaryosunun üçüncü adımında, DSA'nın öznitelik katmanında bulunan YN'lerin aktivasyon fonksiyonları sigmoid olarak seçilerek spike örüntülerinin elde edilmesi hedeflenmiştir. Bu bağlamda, elde edilen B16 ve B32 öznitelikleri doğrudan spike örüntüleri olarak kabul edilmiş, herhangi bir ÖSD aşamasına tabii tutulmadan GSA modellerine giriş olarak verilmiştir. Buna göre 4 bit uzunluğunda yoğunluk kodlama uygulanmış özniteliklere denk B16 öznitelikleri için bulgular Şekil 61'de verilmiştir. Benzer şekilde, 5 bit uzunluğunda yoğunluk kodlama uygulanmış özniteliklere denk B32 öznitelikleri için bulgular ise Şekil 62'de verilmiştir.



Şekil 61. B16 öznitelikleri kullanıldığında eğitim hataları ve performanslar

Şekil 61'e göre B16 öznitelikleri kullanıldığında, VK4 kümesinin eğitiminde OKH'leri ortalama 110 epokta sıfır değerine ulaşmıştır. Bu sayı MFKK öznitelikleri ile karşılaştırıldığında, GSA modellerinin yaklaşık 45 epok daha erken sıfır hataya ulaştığı görülmektedir. VK16 için ise GSA modelleri ortalama 260 epokta sıfır hataya ulaşmışlardır. MFKK öznitelikleri için VK16 kümesinin eğitiminde 500 epok değerinin



Şekil 62. B32 öznitelikleri kullanıldığında eğitim hataları ve performanslar

yeterli olmadığı düşünüldüğünde B16 özniteliklerinin eğitim aşamasını hızlandığı görülmektedir.

VK8 bulgularına göre EVU değeri 5 saniye seçilerek eğitilen GSA modelleri ortalama 141 epok sonunda sıfır hataya ulaşmışlardır. EVU değerinin iki katına çıkmasıyla, ihtiyaç duyulan epok sayısı artmış ve 500'üncü epok sonunda sıfır hataya ulaşılmadan eğitim sonlanmıştır. Ancak eğitim sonunda; ortalama OKH değeri 2.6×10^{-2}

olan MFKK öznitelikleri bulgularına kıyasla, B16 öznitelikleri için ortalama OKH değerinin 9.6×10^{-3} olması; B16 özniteliklerinin üstünlüğünü göstermektedir.

Test aşamasında, VK4 kümesi için GSA modellerinin ortalama performansı MFKK özniteliklerine kıyasla yaklaşık %8 artmış ve %91.0 olarak hesaplanmıştır. Bu artışın VK16 kümesi için daha da belirginleştiği ve yaklaşık %21 olduğu saptanmıştır. Bu bakımdan B16 özniteliklerinin OKT performansını önemli ölçüde artırdığı görülmüştür.

VK8 için test bulguları incelendiğinde, EVU değerinin 5 saniye olarak belirlendiği durumda, OKT performansı tüm GN modellerinde ortalama %89.4 olarak hesaplanmıştır. EVU değerinin iki katına çıkmasıyla ortalama performansın %90.0'a yükseldiği görülmüştür. Bulgular, B16 özniteliklerinin kullanımının EVU değerine duyarlılığı azalttığını göstermektedir.

Öte yandan B32 öznitelikleri kullanıldığında GSA modellerinin eğitim performansları önemli ölçüde artmıştır. Şekil 62'de verilen bulgulara göre tüm durumlarda GSA modelleri 160 epoktan önce sıfır hataya ulaşmışlardır. Bu bağlamda B32 özniteliklerinin kullanıldığı durumda GSA modellerinin eğitim hatalarının seyri hem konuşmacı sayısından hem de EVU değerinden daha az etkilenmiştir.

Test performansları, B32 özniteliklerinin başarısını daha da ön plana çıkarmıştır. VK4 kümesinde B32 özniteliklerinin kullanılmasıyla ortalama test performansı %91.6 olarak ölçülmüştür. Bu değer, diğer özniteliklere kıyasla VK4 kümesi için en yüksek performansa karşılık gelmektedir. VK16 kümesinde ise test performansının ortalama %83.6 olduğu tespit edilmiştir. Bu bulguya göre VK16 kümesi için de B32 özniteliklerinin kullanımı en yüksek ortalama performansa neden olmuştur.

VK8 bulgularına göre 5'er saniyelik konuşma kayıtlarıyla eğitilen GSA modellerinin test performansları ortalama %89.2 olarak hesaplanmıştır. EVU değerinin iki katına çıkmasıyla test performansının %3.3 arttığı ve %92.5 olduğu saptanmıştır. Buna göre B32 öznitelikleri sayesinde eğitim veri miktarının artmasıyla performansta artış meydana gelmiştir.

3.4.5. Piramidal Nöron ve Alt Bölgelerinin Otomatik Konuşmacı Tanıma Performansları

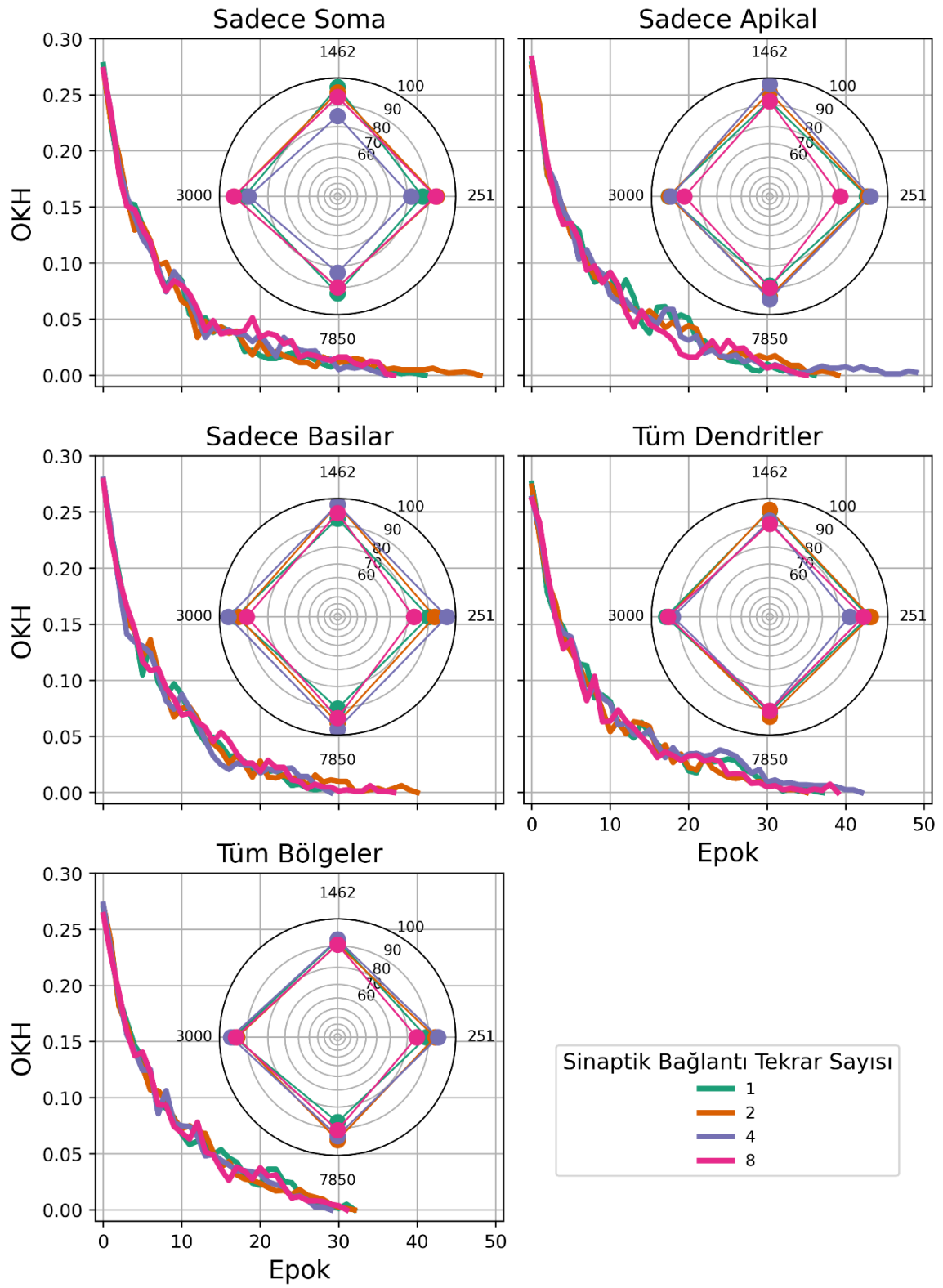
Bu bölümde, detaylarıyla birlikte Bölüm 2.6.5'te tanıtılan D5 isimli deneysel senaryoya dair bulgular verilmiştir. D5 senaryosunda; Bölüm 3.4.1, 3.4.2, 3.4.3 ve 3.4.4'te elde edilen bulgular göz önünde bulundurularak; PN'lerden kurulu GSA modellerinin en

uygun şartlar altında OKT performansları incelenmiştir. Bölüm 3.4.1’de verilen bulgulara göre, GN modellerinde Exc ve Inh SB’lerin birlikte bulunduğu durumda sınıflandırma performansı artmış, aynalama yaklaşımıyla da SB ağırlıklarının pozitif olma kısıtından kaynaklı dezavantajlar ortadan kalkmıştır. Bu nedenle; D5, tüm SB tiplerinin dahil edildiği şartlar altında gerçekleştirilmiştir. Bölüm 3.4.2, 3.4.3 ve 3.4.4’te verilen bulgulara göre B32 özniteliklerinin spike örüntüleri olarak kullanıldığı durumda, OKT performansları en yüksek değerlere ulaşmıştır. Bu nedenle, D5’te B32 öznitelikleri kullanılmıştır.

D5’te kullanılan PN, basit GN modellerine kıyasla karmaşık morfolojiye ve çeşitli aktif iyon kanallarına sahiptir. Ağır işlemsel yükleri, simülasyonlarının basit GN modellerine göre daha uzun zamanda sonlanmasına neden olmaktadır. Bu nedenle, D5 önceki senaryolara göre daha kısıtlı şartlar altında gerçekleştirilmiştir. Buna göre OKT uygulaması için sadece VK4 kümesi kullanılmış ve eğitimde epok sayısı 50 olarak sınırlandırılmıştır. Ayrıca, SB konumlarının ve tekrar sayısının etkisinin ortaya koyulması amacıyla SB sayısının sınıflandırmaya etkisi ve PN alt bölgelerinin performansları incelenmiştir. PN bulgularının karşılaştırılabilmesi adına D5’e basit GN’lerden kurulu GSA modelleri de dahil edilmiştir. PN alt bölgelerinin OKT performansına dair detaylı bulgular EK-6 başlığı altında Ek Tablo 25’te verilmiştir.

D5 için basit GN’lerden kurulu GSA modellerinin performansları incelendiğinde, HH modeli için eğitim aşamasının 97 epok sonunda sıfır hataya ulaştığı ve test performansının 90.8 ± 2.6 olduğu saptanmıştır. IZH modeli için ise GSA modelinin OKH değeri 107 epok sonunda sıfır olmuş ve 90.2 ± 1.8 test performansı sergilemiştir. Son olarak; STA nöronlarından kurulu GSA modeli için, 103 epok sonunda sıfır hataya ulaşan model test aşamasında 93.9 ± 1.5 ’lik performans göstermiştir.

Şekil 63, PN’lerden kurulu GSA modellerinin eğitim ve test performanslarını göstermektedir. Eğitim bulgularına göre PN, VK4 kümesinin eğitiminde basit GN’lere kıyasla daha az miktarda epok sayısına ihtiyaç duymaktadır. Buna göre neredeyse tüm gruplar için GSA modelleri 50 epoktan önce sıfır hataya ulaşmış ve eğitim sonlanmıştır. Basit GN’lerden kurulu modeller için eğitimde ihtiyaç duyulan ortama epok sayısının 102 olması nedeniyle PN’lerin iki kat daha hızlı bir eğitim aşamasına sahip olduğunu söylemek mümkündür. Ayrıca SB tekrar sayısının eğitim aşamasında anlamlı bir etkiye sahip olmadığı da belirlenmiştir.



Şekil 63. Piramidal nöron için eğitim hataları ve performanslar

Detaylı bir incelemeyle, eğitimde ihtiyaç duyulan epok sayısının bölgeler arasında çeşitlilik gösterdiğini saptamak mümkündür. Buna göre sadece basilar bölgesine spike girişi uygulandığı durumda GSA modelleri ortalama 35 epokta sıfır hataya ulaşmıştır. Somatik ve apikal bölgeler için ise bu değerlerin sırasıyla 42 ve 41 olduğu tespit edilmiştir. Apikal ve basilar bölgelerinin gruplanarak kullanıldığı durumda, GSA'lar ortalama 39 epokta sıfır hataya ulaşmışken; tüm bölgelerin aynı anda kullanıldığı durumda, bu sayı 32 olarak kaydedilmiştir.

Test aşaması incelendiğinde, performans bulguları ile SB tekrar sayısı arasında anlamlı bir ilişkiye rastlanılmamıştır. Buna göre ortalama 93.7 ± 3.1 test performansı ile basilar bölge, en yüksek performansın elde edildiği alt bölge olmuştur. Apikal bölgeye konumlandırılmış SB'lerde ortalama test performansı 92.0 ± 3.3 olarak hesaplanırken, somatik bölge 90.2 ± 3.9 ile en düşük test performansı gösteren alt bölge olmuştur.

Öte yandan, alt bölgelerin gruplanması eğitim aşamasında performansta iyileşmeye neden olsa da test performansı için bu durum söz konusu olmamıştır. Tüm dendritlere konumlandırılmış SB'ler için ortalama test performansı 92.5 ± 2.0 iken, tüm bölgeler bir arada kullanıldığında bu performans ortalama 91.9 ± 1.8 olarak ölçülmüştür. Buna göre PN için bölgelerin bir arada kullanılmasının ortalama performansı düşürdüğü saptanmıştır. Ayrıca, SB tekrar sayısı ile test performansı arasında bir ilişki saptanmamıştır. Bu bağlamda PN'lerin SB bağlantı tekrar sayısından etkilenmediğini söylemek mümkündür.

3.4.6. Tartışma

Bu tez çalışmasında, işitsel sistemden esinlenilmiş bir OKT yaklaşımı önerilmiştir. Bu yaklaşımda, konuşma kayıtlarından konuşmacı kimliklerinin tespit edilmesi amaçlanmıştır. OKT, literatürde de önemli bir alan olarak göze çarpmaktadır. Tablo 9'da literatürde yer alan OKT çalışmalarının kimliklendirme performansları verilmiştir. Literatür incelendiğinde, çalışmaların hem konuşma VK'leri hem de konuşmacı sayısı bakımından çeşitli olduğu görülmektedir. Ayrıca kullanılan VK'ler için eğitim ve test aşamasında kullanılan konuşma uzunluklarının da çeşitlilik gösterdiği saptanmıştır. Bu bakımdan elde edilen performans bulgularının karşılaştırma amacıyla değerlendirilmesinin önünde birtakım engeller mevcuttur. Bu nedenle literatürde verilen çalışmalarla karşılaştırma yapılmak istendiğinde eşit şartlar altındaki performansların elde edilmesi gerekmektedir.

Tablo 9. Literatürde otomatik konuşmacı tanıma performansları

Yıl	Yazarlar	Veri Kümesi	Konuşmacı Sayısı	Yöntem	Doğruluk (%)
2003	Wildermoth ve Paliwal	TIMIT	630	GKM+MFKK	81,36
2012	Elmir ve arkadaşları	ELSDSR	22	DVM	100
2013	Bhardwaj ve arkadaşları	Voxforge	100	GFM	93
2015	Assaad ve Serpen	ELSDSR	22	Ensemble sınıflandırıcı	99,2
2015	Ahmad ve arkadaşları	Voxforge	100	PNN	94
2016	Lukic ve arkadaşları	TIMIT	630	CNN+spectrogram	97
2017	Al-Kaltakchi ve arkadaşları	NIST-2008	120	GKM-EAPM	95,83
2017	Al-Kaltakchi ve arkadaşları	NIST-2008	120	GKM-EAPM	95,83
2017	Al-Kaltakchi ve arkadaşları	NIST-2008	120	i-vector	96,97
2017	Al-Kaltakchi ve arkadaşları	NIST-2008	120	GKM-EAPM	95,83
2017	Chang ve Wang	NIST-2008	395	DSA	96,65
2017	Nagrani ve arkadaşları	Voxceleb1	1251	CNN	80,5
2018	Okabe ve arkadaşları	ELSDSR	22	GKM + MFKK + Kalman Filtresi	81,34
2018	Banerjee ve arkadaşları	ELSDSR	22	GKM-EAPM+MFKK	35
2018	Banerjee ve arkadaşları	ELSDSR	22	DBF+GKM-EAPM	95
2018	Sun ve arkadaşları	NIST-2008	395	CNN	96,14
2018	Sun ve arkadaşları	NIST-2008	395	CNN + FBN	97,8
2018	Wen ve arkadaşları	NIST-2008	395	CLNET	97,08
2018	Ravanelli ve Bengio	TIMIT	630	SincNet	93,18

Önerilen yaklaşım dört adımdan meydana gelmektedir. İlk adım, kullanılacak VK ve VK'ye dair özelliklerin belirlenmesidir. Bu bağlamda VK4, VK8 ve VK16 olarak isimlendirilen üç farklı VK; OKT uygulamasında kullanılmıştır. Ancak VK'lere dair iki

Tablo 9'un devamı

Yıl	Yazarlar	Veri Kümesi	Konuşmacı Sayısı	Yöntem	Doğruluk (%)
2019	Dhokal ve arkadaşları	ELSDSR	22	DVM+MFKK	91,66
2019	Dhokal ve arkadaşları	ELSDSR	22	Rastgele Orman+MFKK	93,37
2019	Dhokal ve arkadaşları	ELSDSR	22	DSA+MFKK	92,85
2019	Gupta ve arkadaşları	ELSDSR	22	GKM	99,9
2019	Dhokal ve arkadaşları	ELSDSR	22	Gabor filtresi+CNN+DSA	94,8
2019	Wang ve arkadaşları	VCTK	110	DSA+PNL	98,63
2019	Ribas ve Vincent	NIST-2008	262	EAPM + i-vector	90,38
2019	Nunes ve arkadaşları	TIMIT	630	AM- SincNet	96,24
2020	Medikonda ve arkadaşları	Voxforge	100	Hanman sınıflandırıcı (IHC)	100
2020	Medikonda ve arkadaşları	VCTK	110	Hanman sınıflandırıcı (IHC)	98,9
2020	Nunes ve arkadaşları	TIMIT	630	AM Mobilenet 1D	97,62
2020	Chowdhury ve arkadaşları	TIMIT	630	ALL-SincNet	94,59
2021	Ghezaiel ve arkadaşları	TIMIT	630	Hybrid dalgacık dönüşümü+CNN	98,12
2021	Li	TIMIT	630	LINENET	94,06
2022	Karhikeyan ve Suja	ELSDSR	22	Hibrit Rastgele Orman-DVM+MFKK	93,58
2022	Prachi ve arkadaşları	TIMIT	630	CNN+LSTM	97,15
2024	Karhikeyan ve Suja	ELSDSR	22	CNN	92,25
2024	Karhikeyan ve Suja	ELSDSR	22	MSA-CNN	97,68
2024	Karhikeyan ve Suja	NIST-2008	395	MSA-CNN	98,12

özelliğin seçilmesi gerekmektedir. Bunlar, eğitim ve test aşamalarında kullanılacak olan kayıtların uzunluğudur.

Tablo 10. Belirlenen şartlar altında en gelişkin yöntemlerin OKT performansları

Yıl	Yazarlar	Öznitelik	Yöntem (Parametre)	Konuşmacı Sayısı		
				4	8	16
1995	Reynolds	MFKK	GKM (16)	88,2 ± 9,5	82,5 ± 7,5	77,7 ± 4,9
			GKM (32)	86,2 ± 11,1	78,9 ± 8,6	73,1 ± 5,3
			GKM (64)	83,4 ± 11,0	73,4 ± 11,1	65,9 ± 6,8
2000	Reynolds ve arkadaşları	MFKK	GKM-EAPM (16)	91,9 ± 8,1	84,7 ± 6,7	82,2 ± 3,9
			GKM-EAPM (32)	92,1 ± 6,7	85,3 ± 5,9	83,1 ± 3,9
			GKM-EAPM (64)	91,8 ± 8,0	85,6 ± 7,1	83,8 ± 3,5
2006	Campbell ve arkadaşları	Supervektörler	GKM-DVM (16)	83,1 ± 9,5	73,7 ± 6,7	63,5 ± 4,8
			GKM-DVM (32)	83,2 ± 9,4	76,9 ± 8,3	65,6 ± 5,6
			GKM-DVM (64)	83,4 ± 9,0	74,9 ± 7,7	66,0 ± 5,5
2014	Variani ve arkadaşları	d-vektör	MÇSA	74,8 ± 11,3	58,4 ± 7,9	46,6 ± 3,5
2018	Snyder ve arkadaşları	x-vektör (a)	ZGSA	83,6 ± 9,5	74,3 ± 9,0	64,5 ± 5,0
		x-vektör (b)	ZGSA	81,2 ± 8,9	73,3 ± 7,6	63,0 ± 5,6
2021	Özcan ve Kayıkçıoğlu	MFKK	DSA (2)	94,9 ± 5,3	90,3 ± 5,2	86,6 ± 3,6
			DSA (4)	94,0 ± 6,3	90,5 ± 5,3	87,0 ± 3,7
2021	Özcan ve Kayıkçıoğlu	MFKK+TÖ	DSA-TÖ (2)	95,5 ± 5,3	91,1 ± 5,8	88,0 ± 4,0
			DSA-TÖ (4)	95,7 ± 4,1	90,2 ± 6,0	85,7 ± 4,1

İkinci adım ise konuşma kayıtlarından özniteliklerin çıkarımıdır. Bu adımda, orta ve iç kulaktaki zaman-frekans analizine yakın olan ve literatürde yaygın olarak kullanılan MFKK öznitelikleri kullanılmıştır. Ancak bu adımda da çıkarılacak olan MFKK özniteliklerinin sayısının belirlenmesi gerekmektedir.

İlk iki adımda belirlenmesi gereken MFKK sayısı, EVU ve TVU değerleri OKT performansını doğrudan etkileyen faktörlerdir. Bu tez çalışmasının bir performans incelemesi olduğu düşünüldüğünde, bu faktörlerin etkilerinin detaylı bir yaklaşımla belirlenmesi gerekmektedir. Ayrıca belirlenecek şartlar altında önerilen yaklaşımın gösterdiği performansın başarısının karşılaştırmalı olarak belirlenmesi, yaklaşımın değerini belirlemektedir.

Bu bağlamda sıralanan faktörlerin etkilerinin incelendiği D1 tasarlanmış ve senaryonun ilk adımına dair bulgular Bölüm 3.1 altında paylaşılmıştır. Elde edilen bulgulara göre MFKK sayısı, EVU ve TVU değerleri ile performanslar arasındaki ilişki aydınlatılmıştır.

Bu tez çalışmasında EVU değerinin 5 saniye, TVU değerinin 1 saniye ve MFKK sayısının 23 olduğu şartlar kabul edilmiştir. D1’de sıralanan faktörlerin performansa etkisi incelenmiş olsa da bulgular VK4, VK8 ve VK16 kümelerinde en gelişkin OKT yöntemlerinin performanslarını aydınlatmamıştır. Bu nedenle, tez için belirlenen şartlar altında, seçilen konuşma VK’leri için D1 tekrar edilmiştir. Bu adıma dair bulgular Tablo 10’da verilmektedir. Tablo 10, GSA modellerinin performanslarının karşılaştırılmasında referans olarak kabul edilmiştir.

Önerilen OKT yaklaşımına göre üçüncü adım, iç kulakta işitsel sınırlar tarafından gerçekleştirilen nöral kodlamadır. Konuşma sinyallerinden elde edilen MFKK özniteliklerinin analog formda olması ve beyin bilgiyi işleyebilmek için spike formunda veriye ihtiyaç duyması nedeniyle özniteliklerin spike formunda kodlanması gerekmektedir. Bu bağlamda tez çalışmasında popülasyon kodlama tabanlı bir ÖSD yaklaşımı önerilmiştir. Bu yaklaşıma göre öznitelikler belirlenmesi gereken bit uzunluğunda, belirlenmesi gereken bir kuantalama tekniği ile kuantalanması ve belirlenmesi gereken bir kodlama yöntemiyle ikili forma dönüştürülmesi gerekmektedir.

Nöral kodlama aşaması için önerilen ÖSD yaklaşımın hem belirlenmesi gereken birden fazla faktöre sahip olması hem de performansı ne derece etkileyeceğinin bilinmemesi, bu yaklaşımının test edilmesini gerekmektedir. Yaklaşımın performansının belirlenmesi için bilinen VK özniteliklerinin spike örüntülerine dönüştürülmesi ve bilinen yöntemlerle sınıflandırılması stratejisi izlenmiştir. Bu bağlamda D2 tasarlanmış ve bulguları Bölüm 3.2’de verilmiştir.

D2’de elde edilen bulgulara göre önerilen dönüşümün karşılaştırma VK’lerinde performansı aydınlatılmıştır. Buna göre belirlenmesi gereken faktörlerin performansa etkisi hem sınıflandırıcıya hem de kullanılan VK’ye bağlıdır. Bu bağlamda D2, VK6 kümesinde tekrar edilmiş ve konuşma VK’ler için en uygun spike örüntülerinin 5 bit uzunluğunda tekdüze kuantalanması ve yoğunluk kodlama uygulanmasıyla elde edildiği saptanmıştır. Belirlenen bu faktörler tez çalışmasında kullanılmıştır.

Tez çalışmasının dördüncü ve son adımında elde edilen spike örüntülerinin önerilen bir GSA yaklaşımı kullanılarak sınıflandırılmasıdır. Bu yaklaşımda, spike örüntüleri SB’ler üzerinden GN’lerden kurulu GSA modellerine giriş olarak verilmekte ve GSA modellerinin çıktıları üzerinden sınıf etiketleri belirlenmektedir. Önerilen GSA yaklaşımında; üçü basit GN modelleri, biri PN olmak üzere dört farklı GN modeli kullanılabilmektedir. Öte yandan SB tipleri hem Exc hem de Inh olmak üzere iki farklı tip

olup seçenek olarak sunulmuştur. Ayrıca bu tez çalışmasında önerilen aynalama yaklaşımı da dahil edilebilir olarak tasarlanmıştır. Tüm bu seçenekler düşünüldüğünde önerilen GSA yaklaşımında da faktörlerin performansa etkisinin aydınlatılması gerekmektedir. Yaklaşımın performansının incelenmesi için bilinen VK'lerin spike örüntülerine dönüştürüldüğü ve GSA modelleriyle sınıflandırıldığı bir strateji takip edilmiştir. Bu bağlamda D3 tasarlanmış ve bulguları Bölüm 3.3'te verilmiştir.

D3'te elde edilen bulgulara göre Exc ve Inh SB'lerin GSA modellerindeki önemi aydınlatılmıştır. Ayrıca önerilen ayna yaklaşımının da performansı artırma yönündeki etkili gücü ortaya çıkmıştır. D3'ün son adımında ise deney, VK6 kümesi için tekrarlanmış ve konuşma VK'leri için uygun şartlar belirlenmiştir. Buna göre deneylerde tüm SB tiplerinin kullanılacağı kabul edilmiştir.

D1, D2 ve D3 senaryoları önerilen OKT yaklaşımının tüm adımlarında performansı etkileyen faktörlerin belirlenmesi, performans ile faktörler arasında nasıl bir ilişki olduğu ve konuşma VK'ler için belirlenmesi gereken faktörlerin seçilmesi için gerekli bulguları sunmuştur. Deneyler sonunda; belirlenmesi gereken tüm parametre ve faktörler belirlenmiş, önerilen OKT yaklaşımının bütüncül performansının incelendiği deneylere geçilmiştir.

Bütüncül performansın incelenmesi adına D4 tasarlanmıştır. İlk olarak; D1, D2 ve D3 senaryolarında belirlenen şartlar altında VK4, VK8 ve VK16 kümeleri GSA modelleri yardımıyla sınıflandırılmıştır. Buna göre, MFKK öznitelikleri spike örüntülerine dönüştürülmüş, sonra üç farklı basit GN modellerinden kurulu GSA modelleri ile ayrı ayrı sınıflandırılarak performansları incelenmiştir. İlgili deneylere dair performans bulguları Tablo 11'de toplanmıştır. Tablo 11'de verilen bulgulara göre MFKK özniteliklerinin doğrudan kullanılması sonucu elde edilen performans değerleri, Tablo 10'da performans bulguları verilen en gelişkin yöntemlere kıyasla yetersiz kalmıştır.

MFKK özniteliklerinin doğrudan kullanılmasında elde edilen yetersiz bulgular, önerilen OKT yaklaşımının geliştirilmesinin gerekliliğini ortaya koymuştur. Bu bağlamda, önerilen ÖSD yaklaşımı ve GSA modeli için hali hazırda en uygun faktörlerin seçildiği düşünülerek geliştirmeye açık noktanın MFKK özniteliklerinin kullanıldığı ikinci adım olduğu sonucuna varılmıştır. D1'de ortaya koyulduğu ve Bölüm 3.1.5'te tartışıldığı üzere, önerilen yaklaşıma bir geliştirme aşamasının dahil edilmesiyle performansın artırılmasının mümkün olduğu görülmüştür. Buna göre hibrit bir yaklaşım olarak; MFKK özniteliklerinin analog girişleri alıp, çıkışında güçlendirilmiş analog öznitelikler sunan bir DSA modeli

önerilmiştir. Bu yaklaşıma göre MFKK öznitelikleri, A1 ve A2 olmak üzere iki farklı öznitelik vektörlerine dönüştürülmekte ve deneyler tekrar edilmektedir. D4'ün ikinci adımı olarak belirlenen bu deneylere dair bulgular, Tablo 11'de verilmiştir. Bulgular, DSA ve GSA modellerinin birlikte kullanıldığı hibrit yaklaşım sayesinde performansın artabileceğini göstermiştir. Ancak; A2 özniteliklerinin kullanılması performansta artışa neden olurken, A1 özniteliklerinin kullanılmasıyla performansta düşüşler meydana gelmemiştir.

Analog-analog dönüşüm yapan DSA'ların kullanıldığı deneylerde performansta artışların gözlemlenmesi, DSA ve GSA'ların birlikte kullanıldığı hibrit yaklaşımın potansiyelinin yüksek olduğunu göstermektedir. Bu bağlamda DSA modellerinin kullanıldığı başka bir yaklaşım daha önerilmiştir. Bölüm 3.2.4'te tartışıldığı üzere, önerilen ÖSD yaklaşımının bazı sınıflandırıcılar için performansı düşürücü etkisi saptanmıştır. Bu nedenle, özniteliklerin spike örüntülerine dönüştürülmesinde DÖ tabanlı bir yaklaşıma odaklanılan üçüncü adıma geçilmiştir. Bu yaklaşıma göre analog MFKK öznitelikleri önceden eğitilmiş bir analog-ikili öznitelik dönüştürücü DSA modeline giriş olarak verilmekte ve çıkışları spike örüntüleri olarak GSA modeline giriş olarak verilmektedir. Bir başka ifadeyle, önerilen ÖSD yaklaşımı yerine bir DSA modeli kullanılmaktadır. Bu adıma dair bulgular Tablo 11'de verilmiştir.

Analog-ikili dönüşüm yapan DSA'ların kullanıldığı deneylerde, hibrit yaklaşımın OKT performansının önemli ölçüde arttığı ve Tablo 10'da verilen en gelişkin OKT yöntemleriyle yarışabilecek seviyelere çıktığı görülmüştür. Buna göre DSA'ların özniteliklerin spike örüntülere dönüştürülmesinde kullanılmasıyla, performansta %15'e varan artışlar meydana gelmiştir. Hem B16 öznitelikleri hem de B32 özniteliklerinde rastlanan bu artışlar sonucu, DSA ve GSA modellerinin birlikte kullanıldığı hibrit yaklaşımın performansı artırma yönündeki gücü ortaya çıkmıştır. Özellikle, B32 özniteliklerinde elde edilen bulgular, önerilen yaklaşımın literatürdeki en gelişkin yöntemler arasına girebilmesinin önünü açmıştır.

OKT uygulaması için basit GN modellerinden meydana gelen GSA modellerinin performansının incelenmesinin ardından, PN ve alt bölgelerinin OKT performansının incelenmesi adına D5 tasarlanmıştır. PN'lerin ağır işlemsel nedeniyle D5 kısıtlı şartlar altında gerçekleştirilmiştir. Buna göre en yüksek performansın elde edildiği B32 öznitelikleri kullanılarak, VK4 kümesi sınıflandırılmıştır. Ayrıca PN'lerin dendritik yapılarının da performansa etkisinin incelenebilmesi adına hem PN'lerin alt bölgeleri ayrı

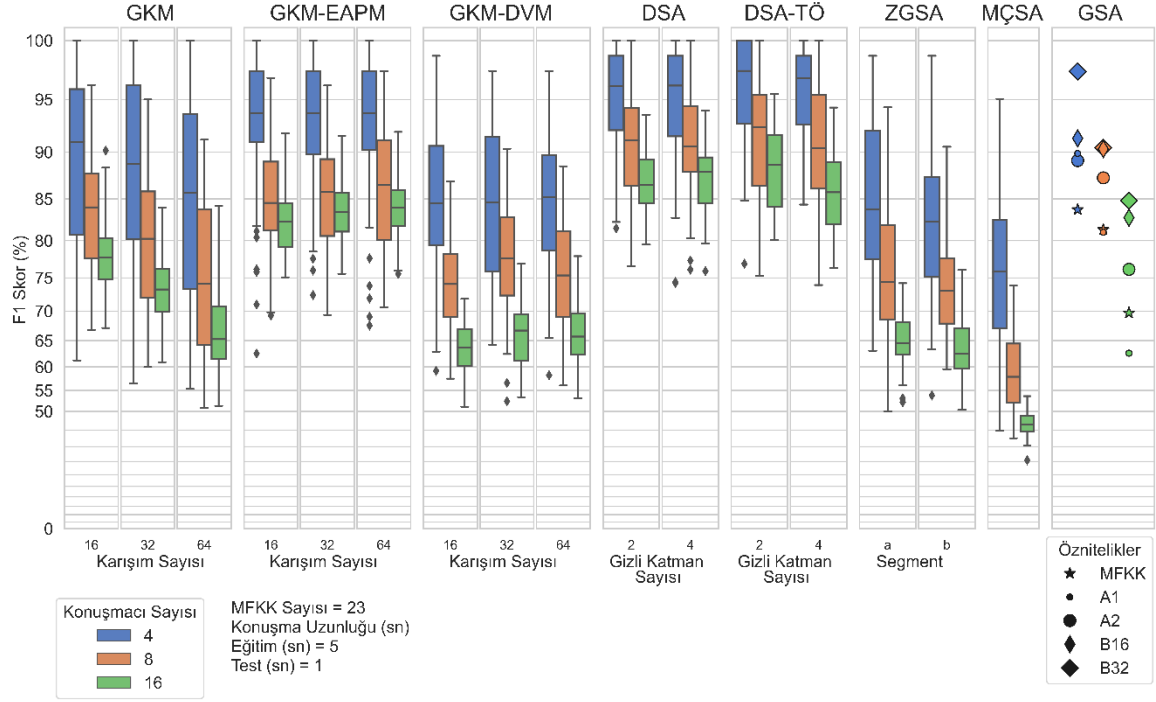
Tablo 11. Belirlenen şartlar altında GSA'ların OKT performansları

Öznitelik	GN (Bölge)	Veri Kümesi		
		VK4	VK8	VK16
B32	PN (Sadece Soma)	93,3		
	PN (Sadece Apikal)	94,9		
	PN (Sadece Basılar)	97,4		
	PN (Tüm Dendritler)	94,4		
	PN (Tüm Bölgeler)	93,9		
	HH	90,8	89,8	83,9
	IZH	90,2	89,0	83,7
	STA	93,9	88,8	83,2
B16	HH	91,3	88,8	81,5
	IZH	90,3	89,3	81,7
	STA	91,3	90,0	81,7
A2	HH	86,1	86,3	73,9
	IZH	88,6	85,0	75,2
	STA	88,4	85,5	73,9
A1	HH	89,7	78,4	62,4
	IZH	87,8	79,2	59,1
	STA	86,4	78,2	60,7
MFKK	HH	81,5	80,3	67,3
	IZH	83,7	79,9	68,0
	STA	83,6	81,2	64,5

ayrı incelenmiş hem de SB bağlantı tekrar sayısının etkisinin gözlemlenebilmesi adına deney genişletilmiştir. D5'te elde edilen bulgular, Tablo 11'de verilmiştir.

Elde edilen bulgulara göre PN'ler basit GN'lere göre daha yüksek OKT performansı göstermiştir. Ayrıca, SB tekrar sayısının sınıflandırma performansına etkisi olmadığı saptanmıştır. Bulgular, PN'lerin sınıflandırma problemlerindeki başarısını ortaya koymaktadır.

Genel bir bakışla, bu tez çalışmasında elde edilen bulgular literatüre önemli katkılar sunmaktadır. Öncelikle, önerilen OKT yaklaşımının işitsel sistemden esinlenmesi ve yüksek performansı, yaklaşımın literatürde önemli bir yere koymaktadır. Öte yandan, önerilen ÖSD ve GSA yaklaşımları, bu tez çalışmasında test edilmiş ve literatürde önemli bir yer tutacağı gösterilmiştir. Ayrıca bulgular, DSA ve GSA modellerinin birlikte



Şekil 64. Belirlenen şartlar altında en gelişkin yöntemler ve GSA performansları

kullanıldığı hibrit yaklaşımının yüksek potansiyelini ortaya koyulmuştur. Bu bakımdan, biyo-esinli bu yaklaşımın gelecekte birçok çalışmaya ışık tutacağını söylemek mümkündür.

Literatürde benzer çalışmalarla karşılaştırıldığında, önerilen yaklaşımın başarısı daha da belirginleşmektedir. Coath ve arkadaşlarının (2005) önerdiği GSA modeli 4 sınıflı bir OKT uygulamasında %89 doğruluk göstermiştir. Bu tez çalışmasında ise VK4 kümesinde hibrit yaklaşımla elde edilen B16 ve B32 özniteliklerinde minimum %90.1'lik F1-Skor başarısı elde edilmiştir. Timoszczuk ve Cabral (2007) çalışmalarında 10 sınıflı bir OKT uygulaması gerçekleştirmişler ve %82 doğruluk performansı bildirmişlerdir. Bu tez çalışmasında ise benzer performansı 16 sınıflı VK16 kümesinde elde etmek mümkün olmuştur.

Önerilen yaklaşım aynı zamanda literatürde en gelişkin yöntemler olarak bilinen yaklaşımlar ile de yarışacak seviyede performanslar göstermiştir. Şekil 64 aynı şartlar altında, en gelişkin yöntemler ve tez çalışmasında elde edilen en yüksek performans değerlerini bir arada göstermektedir.

Reynolds'un (1995) önerdiği GKM yaklaşımı ile karşılaştırıldığında, elde edilen en yüksek GSA bulgularının ortalamasının üzerinde bir performansa sahip olduğu görülmüştür. Buna göre 4, 8 ve 16 sınıflı OKT uygulamaları için B32 öznitelikleri kullanıldığında elde

edilen performans değeri GKM bulgularında Q1 seviyesine karşılık gelmektedir. Burada Q1, Q2, Q3 ve Q4 seviyeleri; ilgili kutu grafiğinde sırasıyla birinci, ikinci, üçüncü ve dördüncü çeyrekleri ifade etmektedir. Reynolds ve arkadaşlarının (2000) önerdiği GKM-EAPM yaklaşımı için de benzer bir durum söz konusudur. Buna göre 4 ve 8 sınıflı OKT uygulamaları için B32 öznitelikleri kullanıldığında elde edilen performans değeri GKM-EAPM bulgularında Q1 seviyesine karşılık gelmektedir. Konuşmacı sayısı 16 olduğunda ise performans Q2 seviyesine gerilemiştir.

Campbell ve arkadaşlarının (2006) GKM-DVM yaklaşımı bulgularına göre önerilen yöntem açıkça üstün durumdadır. Buna göre, tüm konuşmacı sayılarında B32 öznitelikleri, GKM-DVM yaklaşımında elde edilen en yüksek performans değerinden daha yüksek çıkmıştır. Benzer şekilde, önerilen OKT yaklaşımı hem Snyder ve arkadaşlarının (2018) hem de Variani ve arkadaşlarının (2014) çalışma performanslarının önüne geçmiştir. Tüm konuşmacı sayılarında B32 öznitelikleri kullanıldığı durumda GSA modelleri en yüksek x-vektör ve d-vektör performanslarından daha yüksek performans göstermiştir.

Önerilen OKT yaklaşımı bulgularıyla, Özcan ve Kayıkçıoğlu'nun (2021) çalışmalarında kullandıkları DSA ve DSA-TÖ yaklaşımlarının bulguları karşılaştırıldığında, B32 performanslarının ortalama civarında olduğu görülmüştür. Buna göre B32 öznitelikleri bulguları, 4 ve 8 sınıflı OKT uygulaması için Q2 seviyesine, 16 sınıflı için ise Q3 seviyesine karşılık gelmektedir.

Sonuç olarak; bulgular önerilen OKT yaklaşımının literatürdeki diğer çalışmalara kıyasla yüksek bir potansiyele sahip olduğunu vurgulamaktadır. Bu bağlamda konuşma sinyallerinin işitsel sistem esinli incelenmesi, analizi ve sınıflandırılması; gelecek çalışmalar için yeni ufuklar açmaktadır.

4. SONUÇLAR VE ÖNERİLER

Bu tez çalışmasında önerilen bir gerçekçi sinir ağı modelinin otomatik konuşmacı tanıma problemindeki performansı incelenmiştir. İncelemeler dört ardışık adımda gerçekleştirilmiştir.

İlk inceleme adımında, kullanılan konuşma veri kümesinde elde edilebilecek performans sınırları belirlenmiştir. Buna göre literatürde önerilen en gelişkin yöntemler kullanılarak seçilen veri kümesinde mümkün olan performanslar belirlenmiştir. Ayrıca, konuşma veri kümesi için konuşmacı sayısı, MFKK sayısı, eğitim aşamasında kullanılacak olan konuşmacı başına düşen konuşma uzunluğu ve test konuşma kayıtlarının uzunluğunun performansı nasıl etkilediği incelenmiştir. Bu adım sonunda hem önerilen yöntemin literatürdeki en gelişkin yöntemlerle karşılaştırılabilmesi mümkün olmuş hem de tez çalışmasında kullanılan parametrelere göre problemin zorluk derecesi belirlenmiştir. Bu adımda elde edilen bulgulara göre aşağıda sıralanan sonuçlara ulaşmak mümkündür:

- Konuşmacı sayısının artmasıyla sistem performansının düştüğü belirlenmiştir.
- MFKK sayısının artmasıyla sınıflandırma performansı genel olarak artmıştır.
- Eğitimde kullanılan konuşma kayıtlarının uzunluğu ile sınıflandırma performansı arasında pozitif bir ilişki bulunmaktadır.
- Test için kullanılan konuşma kayıtlarının süresi uzadıkça sınıflandırma performansı artmaktadır.
- Geliştirme aşamasına sahip olan yöntemlerin eğitim aşamasında ihtiyaç duydukları veri miktarı, geliştirme aşamasına sahip olmayanlara göre daha azdır.

Bu adımda, tez çalışmasında seçilen parametreler için problemin oldukça zorlayıcı olduğu görülmüştür. Bunun yanında, geliştirme aşamasının performans üzerindeki pozitif etkisi, bu tez çalışmasında önerilen hibrit yaklaşıma ilham kaynağı olmuştur.

İkinci inceleme adımında, önerilen öznitelik spike dönüşümü yaklaşımının sınıflandırma performansına ne derece etki ettiği incelenmiştir. Buna göre üç farklı veri kümesi için özniteliklerin dönüştürülmesinde çeşitli kuantalama teknikleri, kodlama yöntemleri ve bit uzunlukları seçilerek spike örüntüleri elde edilmiştir. Spike örüntüleri için performanslar beş farklı klasik makine öğrenimi yöntemiyle ölçülmüş ve değerlendirilmiştir. Bu adımda elde edilen bulgulara göre aşağıda sıralanan sonuçlara ulaşmak mümkündür:

- Yoğunluk kodlama analog özniteliklerde en az bozulmaya neden olan kodlama yöntemidir.
- Gray ve ikili kodlama yöntemleri bazı veri kümesi ve sınıflandırıcı eşleşmelerinde başarı sağlamış olsa da genel olarak performansları düşüktür.
- Spike örüntülerinin elde edilmesinde kullanılacak kuantalama tekniğinin başarısı kullanılan veri kümesinin öznitelik dağılımına bağlıdır. Bu nedenle kuantalama tekniği veri kümesine özel olarak belirlenmelidir.
- Bit uzunluğunun artmasıyla sınıflandırma performansı artmaktadır. Fakat bu artış bit uzunluğunun belirli değerlere ulaşmasıyla sınırlanmaktadır.
- Bazı sınıflandırıcılar için öznitelik spike dönüşümünün performansı artırdığı saptanmıştır. Bu nedenle, spike örüntülerinin makine öğrenimi uygulamalarında ayrıca performansı artırma potansiyeli vardır.

Bu adım, önerilen öznitelik spike dönüşümü aşamasının etkisinin incelenmesinde önemli bilgiler sağlamıştır. Özellikle karşılaştırma veri kümelerinin kullanımı sayesinde sonuçların evrensel olarak değerlendirilmesi sağlanmıştır. Ayrıca incelemelerin klasik makine öğrenimi yöntemleri kullanılarak yapılması sayesinde, performanslar gerçekçi sinir ağlarından kaynaklı olabilecek etkilerden arındırılmıştır.

İkinci inceleme adımının geliştirilmeye açık birçok yönü bulunmaktadır. Örneğin, kodlama yöntemi olarak sadece üç farklı yöntem incelenmiştir. İlerleyen çalışmalarda, farklı kodlama yöntemlerinin incelenmesi faydalı olacaktır. Benzer durum kuantalama teknikleri için de geçerlidir. İlerleyen çalışmalarda farklı kuantalama tekniklerinin dahil edilmesiyle daha detaylı bilgilere ulaşılabileceğini söylemek mümkündür. Ayrıca bu adımda spike örüntülerinin performansı üç farklı veri kümesi için beş farklı klasik makine öğrenimi yönteminde incelenmiştir. Daha detaylı incelemeler için hem veri kümeleri çeşitlendirilmeli hem de başka sınıflandırıcılar kullanılarak yöntemin performansı değerlendirilmelidir.

Üçüncü inceleme adımında, önerilen gerçekçi sinir ağı modelinde kullanılan sinaptik bağlantı tiplerinin sınıflandırma performansına etkisi incelenmiştir. Buna göre üç farklı veri kümesi spike örüntülerine dönüştürülmüş ve gerçekçi sinir ağlarına giriş olarak verilmiştir. Performans değerlendirmeleri gerçekçi sinir ağlarının tasarımında aynalama yaklaşımının, Exc ve Inh sinaptik bağlantıların kullanılma durumuna göre yapılmıştır. Gerçekçi sinir ağlarında gerçekçi nöronlar olarak Hodgkin&Huxley, Izhikevich ve Sızıntılı

Topla&Ateşle modelleri kullanılmıştır. Bu adımda elde edilen bulgulara göre aşağıda sıralanan sonuçlara ulaşmak mümkündür:

- Exc sinaptik bağlantıların yokluğunda sınıflandırma mümkün değildir.
- Inh sinaptik bağlantılar hem eğitim aşamasını hızlandırmakta hem de test performansını artırmaktadır.
- Önerilen aynalama yaklaşımı sayesinde, sinaptik bağlantı ağırlıklarının pozitif olma kısıtının sınıflandırma kapasitesini düşürme etkisi ortadan kalkmıştır.
- Aynalama yaklaşımı ile gerçekçi sinir ağlarının hem eğitim hem de test performansları önemli ölçüde artmaktadır.
- Aynalama yaklaşımı, Inh sinaptik bağlantılarının yokluğunu kompanse edebilmektedir.

Bu adımda önemli bilgiler edinilmiş olsa da birçok yönden geliştirilmeye açık noktalar bulunmaktadır. Öncelikle önerilen gerçekçi sinir ağında üç farklı gerçekçi nöron modeli kullanılmıştır. Gerçekçi nöron modellerinin çeşitlerinin artırılmasıyla daha detaylı sonuçların elde edilmesi mümkündür. Ayrıca kullanılan sinaptik bağlantı tipleri bireysel olarak değil gruplar halinde incelenmiştir. Buna göre Exc sinaptik bağlantılarda AMPA ve NMDA sinaptik bağlantıların birlikte bulunduğu varsayılmıştır. Bu sinaptik bağlantıların birbirinden ayrı olarak incelenmesiyle bireysel etkilerin değerlendirilmesi mümkündür. Aynı durum Inh sinaptik bağlantılar için de geçerlidir. Buna göre GABA sinapslarının A ve B türlerinin ayrı ayrı incelenmesi Inh sinaptik bağlantıların bireysel etkilerinin aydınlatılması konusunda katkılar sağlayacaktır.

Dördüncü ve son adımda, otomatik konuşmacı tanımada gerçekçi sinir ağlarının performansı üç farklı yaklaşım ile incelenmiştir. İlk olarak, konuşma sinyallerinden elde edilen MFKK öznitelikleri spike örüntülerine dönüştürülerek gerçekçi sinir ağlarına giriş olarak verilmiş ve performansları incelenmiştir. İkinci olarak, MFKK öznitelikleri analog-analog öznitelik dönüştürücü bir derin sinir ağı ile dönüştürülmüştür. Dönüştürülen öznitelikler MFKK özniteliklerinde olduğu gibi spike örüntülerine dönüştürülerek gerçekçi sinir ağlarına giriş olarak verilmiş ve performansları incelenmiştir. Üçüncü ve son olarak, MFKK öznitelikleri analog-ikili öznitelik dönüştürücü bir derin sinir ağı ile spike örüntülerine dönüştürülmüştür. Dönüştürülen bu spike örüntüleri doğrudan gerçekçi sinir ağlarına giriş olarak verilmiş ve performansları incelenmiştir. İncelemeler sırasında gerçekçi sinir ağlarında gerçekçi nöronlar olarak Hodgkin&Huxley, Izhikevich ve Sızıntılı Topla&Ateşle modellerine ek olarak bir piramidal nöron modeli de kullanılmış ve

performansları değerlendirilmiştir. Bu adımda elde edilen bulgulara göre aşağıda sıralanan sonuçlara ulaşmak mümkündür:

- Konuşma sinyalleri için en uygun kodlama yöntemi yoğunluk kodlamadır.
- MFKK öznitelikleri gerçekçi sinir ağlarında otomatik konuşmacı tanıma uygulaması için yetersiz kalmaktadır.
- Önerilen hibrit yaklaşım sayesinde gerçekçi sinir ağı performansı en gelişkin yöntemlerle yarışabilecek seviyelere yükselmiştir.
- Piramidal nöronlar otomatik konuşmacı tanıma için başarılı performanslar sunmuştur.
- Piramidal nöronun dendritik alt bölgelerinin sınıflandırmada benzer performans sergilemeleri, morfolojik ve biyofiziksel çeşitliliğe rağmen piramidal nöronların bölgesel tutarlılığını göstermektedir.

Bu adım sonunda, otomatik konuşmacı tanıma görevinde gerçekçi sinir ağlarının sınıflandırma performansları aydınlatılmıştır. Ancak çalışmada belirlenen kısıtlar nedeniyle geliştirilmeye açık birçok nokta bulunmaktadır. Aşağıda sıralanan öneriler sayesinde çalışmanın detaylandırılması mümkündür:

- Bu çalışmada, LASRC veri tabanının küçük sayılabilecek bir parçası kullanılmıştır. Ayrıca karşılaştırma veri kümeleri olarak IRIS ve MNIST veri kümeleri kullanılmıştır. Gelecekteki çalışmalarda başka veri kümelerinin kullanılmasıyla gerçekçi sinir ağlarının performansı daha detaylı incelenebilir.
- Çalışmada sadece MFKK öznitelikleri için deneyler gerçekleştirilmiştir. Literatürde otomatik konuşmacı tanıma için önerilmiş diğer öznitelikler ile gerçekçi sinir ağlarının performansı yeniden değerlendirilebilir.
- Önerilen gerçekçi sinir ağı modelinde basit nöron modelleri olarak Hodgkin&Huxley, Izhikevich ve Sızıntılı Topla&Ateşle nöron modelleri, piramidal nöron modeli olarak da memeli neokorteksinde bulunan bir model kullanılmıştır. İlerleyen çalışmalarda hem literatürde bulunan diğer basit nöron modelleri dahil edilebilir, hem de çeşitli katmanlarda bulunan piramidal nöronlar ve morfolojileri kullanılabilir.
- Çalışmada öğrenme yöntemi olarak perceptron kuralı kullanılmıştır. Spike zamanına bağlı plastisite gibi biyolojiden daha çok esinlenen veya gradyan iniş gibi daha gelişmiş öğrenme yöntemleriyle performansın artırılması sağlanabilir.

- Önerilen hibrit yaklaşımın performansı önemli ölçüde artırdığı görülmüştür. Bu yaklaşım takip edilerek hem daha etkili derin sinir ağı mimarileri kullanılabilir, hem de geliştirme aşamasında kullanılan veri kümesi genişletilerek öznitelik dönüşümünün sağlamlığı artırılabilir.
- Çalışmada öznitelik spike dönüşümünde kısıtlı sayıda kodlama yöntemi ve kuantalama tekniği kullanılmıştır. Bu seçenekler artırılarak gerçekçi sinir ağının performansı tekrar değerlendirilebilir.
- Bu çalışmada özniteliklerin spike örüntülerine dönüştürülmesinde zamandan bağımsız, popülasyon tabanlı bir yaklaşım takip edilmiştir. İlerleyen çalışmalarda frekans veya faz tabanlı yaklaşımlarla gerçekçi sinir ağlarının performansını artırmak mümkün olabilir.
- Bu çalışmada spike örüntüleri giriş olarak verilmiş ve sadece çıkış katmanı kullanılmıştır. İlerleyen çalışmalarda çok katmanlı yapılar sayesinde performans artışı sağlanabilir.
- Konuşma sinyallerinden elde edilen MFKK öznitelikleri zaman-mekan tabanlı bir yapıya sahiptir. Ancak çalışmada çerçeve bazlı bir yaklaşımla zaman özellikleri ihmal edilmiştir. Gelecek çalışmalarda zaman-mekan tabanlı bir yaklaşımla performans artışı mümkün olabilir.

Sonuç olarak, bu çalışmada gerçekçi sinir ağlarının otomatik konuşmacı tanıma için yüksek performans sergileyebileceği ortaya koyulmuştur. Önerilen öznitelik spike dönüşümü yöntemi sayesinde analog değerlerin başarılı bir şekilde spike kodlarına dönüştürülmesi sağlanmıştır. Önerilen aynalama yaklaşımı sayesinde gerçekçi sinir ağlarının sınıflandırma kapasitelerinin önemli ölçüde arttığı gösterilmiştir. Ayrıca derin sinir ağları ve gerçekçi sinir ağların birlikte çalışmasıyla performansın önemli ölçüde arttığı gösterilmiştir. Bu bağlamda tez çalışmasının literatüre önemli katkılar sağladığını söylemek mümkündür.

5. KAYNAKLAR

- Abadi M., Agarwal A., Barham P., Brevdo E., Chen Z., Citro C., Corrado G. S., Davis A., Dean J., Devin M., Ghemawat S., Goodfellow I., Harp A., Irving G., Isard M., Yangqing J., Jozefowicz R., Kaiser L., Kudlur M., Levenberg J., Mané D., Monga R., Moore S., Murray D., Olah C., Schuster M., Shlens J., Steiner B., Sutskever I., Talwar K., Tucker P., Vanhoucke V., Vasudevan V., Viégas F., Vinyals O., Warden P., Wattenberg M., Wicke M., Yu Y., & Zheng X. (2016). TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. *arXiv preprint arXiv:1603.04467*.
- Abdul, Z. K., & Al-Talabani, A. K. (2022). Mel Frequency Cepstral Coefficient and Its Applications: A Review. *IEEE Access*, 10, 122136-122158.
- Adams, P. R., Brown, D. A., & Constanti, A. (1982). M-Currents and Other Potassium Currents in Bullfrog Sympathetic Neurones. *The Journal of Physiology*, 330(1), 537-572.
- Adrian, E. D. (1926). The Impulses Produced by Sensory Nerve Endings: Part I. *The Journal of Physiology*, 61(1), 49.
- Ahmad, K. S., Thosar, A. S., Nirmal, J. H., & Pande, V. S. (2015). A Unique Approach in Text Independent Speaker Recognition Using MFCC Feature Sets and Probabilistic Neural Network. *International Conference on Advances in Pattern Recognition*, s. 1-6.
- Algabri, M., Mathkour, H., Bencherif, M. A., Alsulaiman, M., & Mekhtiche, M. A. (2017). Automatic Speaker Recognition for Mobile Forensic Applications. *Mobile Information Systems*, 2017(1), 6986391.
- Ali, H., Tran, S. N., Benetos, E., & d'Avila Garcez, A. S. (2016). Speaker Recognition with Hybrid Features from a Deep Belief Network. *Neural Computing and Applications*, 29, 13–19.
- Al-Kaltakchi S, M. T., Woo, W. L., Dlay, S., & Chambers, J. A. (2017). Evaluation of a Speaker Identification System with and without Fusion Using Three Databases in the Presence of Noise and Handset Effects. *Journal on Advances in Signal Processing*, 2017, 1–17.
- Ambikairajah, E., Li, H., Wang, L., Yin, B., & Sethu, V. (2011). Language Identification: A Tutorial. *IEEE Circuits and Systems Magazine*, 11, 82–108.
- Amit, D. J., Wong, K. Y., & Campbell, C. (1989). Perceptron Learning with Sign-Constrained Weights. *Journal of Physics A: Mathematical and General*, 22, 2039–2045.
- Anumula, J., Neil, D., Delbruck, T., & Liu, S. C. (2018). Feature Representations for Neuromorphic Audio Spike Streams. *Frontiers in Neuroscience*, 12, 308889.

- Arora, A., & Arora, A. (2023). Machine Learning Models Trained on Synthetic Datasets of Multiple Sample Sizes for the Use of Predicting Blood Pressure from Clinical Data in a National Dataset. *Plos One*, 18(3), e0283094.
- Assaad, F. S., & Serpen, G. (2015). Transformation Based Score Fusion Algorithm for Multi-Modal Biometric User Authentication through Ensemble Classification. *Procedia Computer Science*, 61, 410–415.
- Auge, D., Hille, J., Mueller, E., & Knoll, A. (2021). A Survey of Encoding Techniques for Signal Processing in Spiking Neural Networks. *Neural Processing Letters*, 53, 4693–4710.
- Avery, R. B., & Johnston, D. (1996). Multiple Channel Types Contribute to the Low-Voltage-Activated Calcium Current in Hippocampal CA3 Pyramidal Neurons. *The Journal of Neuroscience*, 16, 5567–5582.
- Bahari, M. H., McLaren, M., Van hamme, H., & van Leeuwen, D. A. (2014). Speaker Age Estimation Using I-Vectors. *Engineering Applications of Artificial Intelligence*, 34, 99–108.
- Baldi, P. (1995). Gradient Descent Learning Algorithm Overview: A General Dynamical Systems Perspective. *IEEE Transactions on Neural Networks*, 6, 182–195.
- Banerjee, A., Dubey, A., Menon, A., Nanda, S., & Nandi, G. C. (2018). Speaker Recognition Using Deep Belief Networks. *arXiv preprint arXiv:1805.08865*.
- Bear, M. F., & Malenka, R. C. (1994). Synaptic Plasticity: LTP and LTD. *Current Opinion in Neurobiology*, 4, 389–399.
- Beaulieu-Laroche L., Toloza E. H. S., van der Goes M. S., Lafourcade M., Barnagian D., Williams Z. M., Eskandar E. N., Frosch M. P., Cash S. S., & Harnett M. T. (2018). Enhanced Dendritic Compartmentalization in Human Cortical Neurons. *Cell*, 175, 643–651.
- Beigi, H. (2011). *Speaker Recognition*. Springer US.
- Beritelli, F., & Spadaccini, A. (2012). Performance Evaluation of Automatic Speaker Recognition Techniques for Forensic Applications. *New Trends and Developments in Biometrics*, 129–148.
- Bhangale, K. B., & Mohanaprasad, K. (2021). A Review on Speech Processing Using Machine Learning Paradigm. *International Journal of Speech Technology*, 24, 367–388.
- Bhardwaj, S., Srivastava, S., Hanmandlu, M., & Gupta, J. R. (2013). GFM-Based Methods for Speaker Identification. *IEEE Transactions on Cybernetics*, 43, 1047–1058.
- Bicknell, B. A., & Häusser, M. (2021). A Synaptic Learning Rule for Exploiting Nonlinear Dendritic Computation. *Neuron*, 109, 4001–4017.

- Campbell, J. P. (1997). Speaker Recognition: A Tutorial. *Proceedings of the IEEE*, 85, 1437–1462.
- Campbell, W. M., Campbell, J. P., Reynolds, D. A., Singer, E., & Torres-Carrasquillo, P. A. (2006). Support Vector Machines for Speaker and Language Recognition. *Computer Speech & Language*, 20(2-3), 210–229.
- Castiglioni I., Rundo L., Codari M., Di Leo G., Salvatore C., Interlenghi M., Gallivanone F., Cozzi A., D’Amico N. C., & Sardanelli F. (2021). AI Applications to Medical Images: From Machine Learning to Deep Learning. *Physica Medica*, 83, 9–24.
- Caucheteux, C., Gramfort, A., & King, J. R. (2023). Evidence of a Predictive Coding Hierarchy in the Human Brain Listening to Speech. *Nature Human Behaviour*, 7, 430–441.
- Ceolini, E., Anumula, J., Braun, S., & Liu, S. C. (2019). Event-Driven Pipeline for Low-Latency Low-Compute Keyword Spotting and Speaker Verification System. *IEEE International Conference on Acoustics, Speech and Signal Processing*, s. 7953–7957.
- Chang, J., & Wang, D. (2017). Robust Speaker Recognition Based on DNN/I-Vectors and Speech Separation. *IEEE International Conference on Acoustics, Speech and Signal Processing*, s. 5415–5419.
- Chapeton, J., Fares, T., LaSota, D., & Stepanyants, A. (2012). Efficient Associative Memory Storage in Cortical Circuits of Inhibitory and Excitatory Neurons. *Proceedings of the National Academy of Sciences*, 109, 3614–3622.
- Cheeseman, P. C., Self, M., Kelly, J., Taylor, W., Freeman, D., & Stutz, J. C. (1988). Bayesian Classification. *AAAI*, 88, s. 607–611.
- Chen, N., Qian, Y., & Yu, K. (2015). Multi-Task Learning for Text-Dependent Speaker Verification. *Annual Conference of the International Speech Communication Association*.
- Chowdhury, L., Zunair, H., & Mohammed, N. (2020). Robust Deep Speaker Recognition: Learning Latent Representation with Joint Angular Margin Loss. *Applied Sciences*, 10, 7522.
- Coath, M., Brader, J. M., Fusi, S., & Denham, S. L. (2005). Multiple Views of the Response of an Ensemble of Spectro-Temporal Features Support Concurrent Classification of Utterance, Prosody, Sex and Speaker Identity. *Network: Computation in Neural Systems*, 16, 285–300.
- Colbert, C. M., & Pan, E. (2002). Ion Channel Properties Underlying Axonal Action Potential Initiation in Pyramidal Neurons. *Nature Neuroscience*, 5, 533–538.
- Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, 20, 273–297.

- Cover, T., & Hart, P. (1967). Nearest Neighbor Pattern Classification. *IEEE Transactions on Information Theory*, 13, 21–27.
- Dehak, N., Kenny, P. J., Dehak, R., Dumouchel, P., & Ouellet, P. (2011). Front-End Factor Analysis for Speaker Verification. *IEEE Transactions on Audio, Speech, and Language Processing*, 19, 788–798.
- del Olmo, D., Pavelka, M., & Kosek, J. (2020). Open-Circuit Voltage Comes from Non-Equilibrium Thermodynamics. *Journal of Non-Equilibrium Thermodynamics*, 46, 91–108.
- Dellaferrera, G., Martinelli, F., & Cernak, M. (2020). A Bin Encoding Training of a Spiking Neural Network Based Voice Activity Detection. *IEEE International Conference on Acoustics, Speech and Signal Processing*, s. 3207–3211.
- Deng, L. (2012). The MNIST Database of Handwritten Digit Images for Machine Learning Research. *IEEE Signal Processing Magazine*, 29, 141–142.
- Dennis, J., Huy Dat, T., & Li, H. (2015). Combining Robust Spike Coding with Spiking Neural Networks for Sound Event Classification. *IEEE International Conference on Acoustics, Speech and Signal Processing*, s. 176–180.
- Dennis, J., Yu, Q., Tang, H., Tran, H. D., & Li, H. (2013). Temporal Coding of Local Spectrogram Features for Robust Sound Recognition. *IEEE International Conference on Acoustics, Speech and Signal Processing*, s. 803–807.
- DeWitt, I., & Rauschecker, J. P. (2013). Wernicke’s Area Revisited: Parallel Streams and Word Processing. *Brain and Language*, 127, 181–191.
- Dhakal, P., Damacharla, P., Javaid, A. Y., & Devabhaktuni, V. (2019). A near Real-Time Automatic Speaker Recognition Architecture for Voice-Based User Interface. *Machine Learning and Knowledge Extraction*, 1, 504–520.
- Dibazar, A. A., Song, D., Yamada, W., & Berger, T. W. (2004). Speech Recognition Based on Fundamental Functional Principles of the Brain. *IEEE International Joint Conference on Neural Networks*, 4, s. 3071–3075.
- Dixon, M. F., Halperin, I., & Bilokon, P. (2020). *Machine Learning in Finance*. New York: Springer International Publishing.
- Dong, M., Huang, X., & Xu, B. (2018). Unsupervised Speech Recognition through Spike-Timing-Dependent Plasticity in a Convolutional Spiking Neural Network. *Plos One*, 13, e0204596.
- Dong, S., Wang, P., & Abbas, K. (2021). A Survey on Deep Learning and Its Applications. *Computer Science Review*, 40, 100379.
- Du N., Kiani M., Mayr C. G., You T., Bürger D., Skorupa I., Schmidt O. G., & Schmidt H. (2015). Single Pairing Spike-Timing Dependent Plasticity in Bifeo3 Memristors with a Time Window of 25 ms to 125 μ s. *Frontiers in Neuroscience*, 9, 227.

- Du Z., Ben-Dayan Rubin D. D., Chen Y., He L., Chen T., Zhang L., Wu C., & Temam O. (2015). Neuromorphic Accelerators: A Comparison between Neuroscience and Machine-Learning Approaches. *Proceedings of the 48th International Symposium on Microarchitecture*, s. 494–507.
- Eliasmith, C., & Anderson, C. H. (2003). *Neural Engineering: Computation, Representation, and Dynamics in Neurobiological Systems*. MIT press.
- Elmir, Y., Elberrichi, Z., & Adjoudj, R. (2012). Score Level Fusion Based Multimodal Biometric Identification (fingerprint & voice). *International Conference on Sciences of Electronics, Technologies of Information and Telecommunications*, s. 146–150.
- Ermentrout, G. B., & Terman, D. H. (2010). *Mathematical Foundations of Neuroscience*. New York: Springer New York.
- Eyal G., Verhoog M. B., Testa-Silva G., Deitcher Y., Lodder J. C., Benavides-Piccione R., Morales J., DeFelipe J., de Kock C. P. J., Mansvelder H. D., & Segev I. (2016). Unique Membrane Properties and Enhanced Signal Processing in Human Neocortical Neurons. *eLife*, 5, e16553.
- Fendji, J. L., Tala, D. C., Yenke, B. O., & Atemkeng, M. (2022). Automatic Speech Recognition Using Limited Vocabulary: A Survey. *Applied Artificial Intelligence*, 36, 2095039.
- Fisher, R. A. (1936). The Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics*, 7, 179–188.
- Fujiyoshi, H., Hirakawa, T., & Yamashita, T. (2019). Deep Learning-Based Image Recognition for Autonomous Driving. *IATSS Research*, 43, 244–252.
- Galloni, A. R., Laffere, A., & Rancz, E. (2020). Apical Length Governs Computational Diversity of Layer 5 Pyramidal Neurons. *eLife*, 9, e55761.
- Gawne, T. J., Kjaer, T. W., & Richmond, B. J. (1996). Latency: Another Potential Code for Feature Binding in Striate Cortex. *Journal of Neurophysiology*, 76, 1356–1360.
- Georgiev, D. D., Kolev, S. K., Cohen, E., & Glazebrook, J. F. (2020). Computational Capacity of Pyramidal Neurons in the Cerebral Cortex. *Brain Research*, 1748, 147069.
- Gerstner, W., Kistler, W. M., Naud, R., & Paninski, L. (2014). *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition*. Cambridge University Press.
- Ghezaiel, W., Brun, L., & Lézoray, O. (2021). Hybrid Network for End-To-End Text-Independent Speaker Identification. *International Conference on Pattern Recognition*, s. 2352–2359.
- Gidon A., Zolnik T. A., Fidzinski P., Bolduan F., Papoutsis A., Poirazi P., Holtkamp M., Vida I., & Larkum M. E. (2020). Dendritic Action Potentials and Computation in Human Layer 2/3 Cortical Neurons. *Science*, 367, 83–87.

- Gollisch, T., & Meister, M. (2008). Rapid Neural Coding in the Retina with Relative Spike Latencies. *Science*, 319, 1108–1111.
- Gooch H. M., Bluett T., Perumal M. B., Vo H. D., Fletcher L. N., Papacostas J., Jeffree R. L., Wood M., Colditz M. J., McMillen J., Tsahtsarlis T., Amato D., Campbell R., Gillinder L., & Williams S. R. (2022). High-Fidelity Dendritic Sodium Spike Generation in Human Layer 2/3 Neocortical Pyramidal Neurons. *Cell Reports*, 41.
- Gosztolya, G., & Tóth, L. (2008). Detection of Phoneme Boundaries Using Spiking Neurons. *International Conference of Artificial Intelligence and Soft Computing*, s. 782–793.
- Guo, L., Guo, M., Wu, Y., & Xu, G. (2023). Specific Neural Coding of fMRI Spiking Neural Network Based on Time Coding. *Chaos, Solitons & Fractals*, 174, 113821.
- Gupta, M., Bharti, S. S., & Agarwal, S. (2019). Gender-Based Speaker Recognition from Speech Signals Using GMM Model. *Modern Physics Letters B*, 33, 1950438.
- Gütig, R., & Sompolinsky, H. (2006). The Tempotron: A Neuron That Learns Spike Timing–Based Decisions. *Nature Neuroscience*, 9, 420–428.
- Hajavi, A., & Etemad, A. (2019). A Deep Neural Network for Short-Segment Speaker Recognition. *Interspeech*.
- Hamza, M., Khodadadi, T., & Palaniappan, S. (2020). A Novel Automatic Voice Recognition System Based on Text-Independent in a Noisy Environment. *International Journal of Electrical and Computer Engineering*, 10, 3643.
- Han, B., & Roy, K. (2020). Deep Spiking Neural Network: Energy Efficiency through Time Based Coding. *European Conference on Computer Vision*, s. 388–404.
- Hawkins, D. M. (2004). The Problem of Overfitting. *Journal of Chemical Information and Computer Sciences*, 44(1), 1-12.
- Hay, E., Hill, S., Schürmann, F., Markram, H., & Segev, I. (2011). Models of Neocortical Layer 5b Pyramidal Cells Capturing A wide Range of Dendritic and Perisomatic Active Properties. *Plos Computational Biology*, 7, e1002107.
- Hebb, D. O. (1949). The First Stage of Perception: Growth of the Assembly. *The Organization of Behavior*, 4, 78–60.
- Heigold, G., Moreno, I., Bengio, S., & Shazeer, N. (2016). End-To-End Text-Dependent Speaker Verification. *IEEE International Conference on Acoustics, Speech and Signal Processing*, s. 5115–5119.
- Hermansky, H. (1990). Perceptual Linear Predictive (PLP) Analysis of Speech. *The Journal of the Acoustical Society of America*, 87, 1738–1752.
- Hines, M. (2009). NEURON and Python. *Frontiers in Neuroinformatics*, 3, 391.

- Hines, M. L., & Carnevale, N. T. (1997). The NEURON Simulation Environment. *Neural Computation*, 9, 1179–1209.
- Hodgkin, A. L., & Huxley, A. F. (1952). A Quantitative Description of Membrane Current and Its Application to Conduction and Excitation in Nerve. *Bulletin of Mathematical Biology*, 52, 25–71.
- Hough, M., De Garis, H., Korkin, M., Gers, F., & Nawa, N. E. (1999). SPIKER: Analog Waveform to Digital Spiketrain Conversion in Atr's Artificial Brain (cam-brain) Project. *International Conference on Robotics and Artificial Life*, 92.
- Ishimoto, H., Izawa, M., & Torikai, H. (2015). A Novel Cochlea Partition Model Based on Asynchronous Bifurcation Processor. *Nonlinear Theory and Its Applications*, 6, 207–225.
- Izhikevich, E. M. (2006). *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*. The MIT Press.
- Ji, J., Gao, S., Cheng, J., Tang, Z., & Todo, Y. (2016). An Approximate Logic Neuron Model with a Dendritic Structure. *Neurocomputing*, 173, 1775–1783.
- Jones, A. R., Overly, C. C., & Sunkin, S. M. (2009). The Allen Brain Atlas: 5 Years and Beyond. *Nature Reviews Neuroscience*, 10, 821–828.
- Karthikeyan, V., & Suja Priyadharsini, S. (2024). A Stacked Convolutional Neural Network Framework with Multi-Scale Attention Mechanism for Text-Independent Voiceprint Recognition. *Pattern Analysis and Applications*, 27, 1–15.
- Karthikeyan, V., & Suja, P. S. (2022). Hybrid Machine Learning Classification Scheme for Speaker Identification. *Journal for Forensic Sciences*, 67, 1033–1048.
- Kath, W. L. (2005). Computational Modeling of Dendrites. *Journal of Neurobiology*, 64, 91–99.
- Katz, Y., Menon, V., Nicholson, D. A., Geinisman, Y., Kath, W. L., & Spruston, N. (2009). Synapse Distribution Suggests a Two-Stage Model of Dendritic Integration in CA1 Pyramidal Neurons. *Neuron*, 63, 171–177.
- Kayikcioglu Bozkir, I., Ozcan, Z., Kose, C., Kayikcioglu, T., & Cetin, A. E. (2023). Improving a Cortical Pyramidal Neuron Model's Classification Performance on a Real-World ECG Dataset by Extending Inputs. *Journal of Computational Neuroscience*, 51, 329–341.
- Kessels, H. W., & Malinow, R. (2009). Synaptic AMPA Receptor Plasticity and Behavior. *Neuron*, 61, 340–350.
- Kim, E. H., Hyun, K. H., Kim, S. H., & Kwak, Y. K. (2009). Improved Emotion Recognition with a Novel Speaker-Independent Feature. *IEEE/ASME Transactions on Mechatronics*, 14, 317–325.

- Kim, J. S., Kim, J., Ahn, J., Chung, S., & Han, C. S. (2023). Artificial Action Potential and Ionic Power Device Inspired by Ion Channels and Excitable Cell. *Advanced Science*, 10, 2301037.
- Kinnunen, T., & Li, H. (2010). An Overview of Text-Independent Speaker Recognition: From Features to Supervectors. *Speech Communication*, 52, 12–40.
- Kirch, C., & Gollo, L. L. (2020). Spatially Resolved Dendritic Integration: Towards a Functional Classification of Neurons. *PeerJ*, 8, e10250.
- Kole, M. H., Hallermann, S., & Stuart, G. J. (2006). Single Ih Channels in Pyramidal Neuron Dendrites: Properties, Distribution, and Impact on Action Potential Output. *Journal of Neuroscience*, 26, 1677–1687.
- Kondratskyi, A., Kondratska, K., Skryma, R., Klionsky, D. J., & Prevarskaya, N. (2017). Ion Channels in the Regulation of Autophagy. *Autophagy*, 14, 3–21.
- Korngreen, A., & Sakmann, B. (2000). Voltage-Gated K⁺ Channels in Layer 5 Neocortical Pyramidal Neurones from Young Rats: Subtypes and Gradients. *The Journal of Physiology*, 525, 621–639.
- Köhler, M., Hirschberg, B., Bond, C. T., Kinzie, J. M., Marrion, N. V., Maylie, J., & Adelman, J. P. (1996). Small-Conductance, Calcium-Activated Potassium Channels from Mammalian Brain. *Science*, 273, 1709–1714.
- Kramer, O. (2016). *Machine Learning for Evolution Strategies*. Heidelberg: Springer.
- Lapicque, L. M. (1907). Recherches Quantitatives sur L'excitation Electrique des Nerfs. *J Physiol Paris*, 9, 620–635.
- Larson, E., Perrone, B. P., Sen, K., & Billimoria, C. P. (2010). A Robust and Biologically Plausible Spike Pattern Recognition Network. *The Journal of Neuroscience*, 30, 15566–15572.
- Latinus, M., & Belin, P. (2011). Human Voice Perception. *Current Biology*, 21, 143–145.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. *Nature*, 521, 436–444.
- Lee, J. H., Wang, X., & Bendor, D. (2020). The Role of Adaptation in Generating Monotonic Rate Codes in Auditory Cortex. *Plos Computational Biology*, 16, e1007627.
- Legenstein, R., & Maass, W. (2011). Branch-Specific Plasticity Enables Self-Organization of Nonlinear Computation in Single Neurons. *Journal of Neuroscience*, 31, 10787–10802.
- Legenstein, R., Naeger, C., & Maass, W. (2005). What Can a Neuron Learn with Spike-Timing-Dependent Plasticity. *Neural Computation*, 17, 2337–2382.
- Lewis, C. M., Ni, J., Wunderle, T., Jendritza, P., Diester, I., & Fries, P. (2020). Gamma-Rhythmic Input Causes Spike Output. *bioRxiv*, 2020–03.

- Li, W. (2021). Speaker Identification from Raw Waveform with Linenet. *arXiv preprint arXiv:2105.14826*.
- Limbacher, T., & Legenstein, R. (2020). Emergence of Stable Synaptic Clusters on Dendrites through Synaptic Rewiring. *Frontiers in Computational Neuroscience*, 14, 57.
- Lisboa, P. J., & Perantonis, S. J. (1991). Complete Solution of the Local Minima in the XOR Problem. *Network: Computation in Neural Systems*, 2, 119–124.
- Litovsky, R. (2015). Development of the Auditory System. *Handbook of Clinical Neurology*, 129, 55–72.
- Lobo, J. L., Del Ser, J., Bifet, A., & Kasabov, N. (2020). Spiking Neural Networks and Online Learning: An Overview and Perspectives. *Neural Networks*, 121, 88–100.
- Luebke, J. I. (2017). Pyramidal Neurons Are Not Generalizable Building Blocks of Cortical Networks. *Frontiers in Neuroanatomy*, 11, 11.
- Lukic, Y., Vogt, C., Dürr, O., & Stadelmann, T. (2016). Speaker Identification and Clustering Using Convolutional Neural Networks. *International Workshop on Machine Learning for Signal Processing*, s. 1–6.
- Lyon, R. (1982). A Computational Model of Filtering, Detection, and Compression in the Cochlea. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 7, s. 1282–1285.
- Magee, J. C., & Grienberger, C. (2020). Synaptic Plasticity Forms and Functions. *Annual Review of Neuroscience*, 43, 95–117.
- Magistretti, J., & Alonso, A. (1999). Biophysical Properties and Slow Voltage-Dependent Inactivation of a Sustained Sodium Current in Entorhinal Cortex Layer-II Principal Neurons: A Whole-Cell and Single-Channel Study. *The Journal of General Physiology*, 114, 491–509.
- Malakar, P., Balaprakash, P., Vishwanath, V., Morozov, V., & Kumaran, K. (2018). Benchmarking Machine Learning Methods for Performance Modeling of Scientific Applications. *IEEE/ACM Performance Modeling, Benchmarking and Simulation of High Performance Computer Systems*, s. 33–44.
- Malenka, R. C., & Nicoll, R. A. (1993). NMDA-Receptor-Dependent Synaptic Plasticity: Multiple Forms and Mechanisms. *Trends in Neurosciences*, 16, 521–527.
- McCulloch, W. S., & Pitts, W. (1943). A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biology*, 52, 99–115.
- Meaud, J., & Grosh, K. (2010). The Effect of Tectorial Membrane and Basilar Membrane Longitudinal Coupling in Cochlear Mechanics. *The Journal of the Acoustical Society of America*, 127, 1411–1421.

- Medikonda, J., Bhardwaj, S., & Madasu, H. (2020). An Information Set-Based Robust Text-Independent Speaker Authentication. *Soft Computing*, 24, 5271–5287.
- Miao, Y., Tang, H., & Pan, G. (2018). A Supervised Multi-Spike Learning Algorithm for Spiking Neural Networks. *International Joint Conference on Neural Networks*, s. 1–7.
- Migliore, M., Cannia, C., Lytton, W. W., Markram, H., & Hines, M. L. (2006). Parallel Network Simulations with NEURON. *Journal of Computational Neuroscience*, 21, 119–129.
- Mihaljević, B., Larrañaga, P., & Bielza, C. (2021). Comparing the Electrophysiology and Morphology of Human and Mouse Layer 2/3 Pyramidal Neurons with Bayesian Networks. *Frontiers in Neuroinformatics*, 15, 580873.
- Milosavljevic, N., Storch, R., Eleftheriou, C. G., Collins, A., Petersen, R. S., & Lucas, R. J. (2018). Photoreceptive Retinal Ganglion Cells Control the Information Rate of the Optic Nerve. *Proceedings of the National Academy of Sciences*, 115, 11817–11826.
- Mody, I., De Koninck, Y., Otis, T. S., & Soltesz, I. (1994). Bridging the Cleft at GABA Synapses in the Brain. *Trends in Neurosciences*, 17, 517–525.
- Mohan H., Verhoog M. B., Doreswamy K. K., Eyal G., Aardse R., Lodder B. N., Goriounova N. A., Asamoah B., Brakspear A. B. C., Groot C., van der Sluis S., Testa-Silva G., Obermayer J., Boudewijns Z. S. R. M., Narayanan R. T., Baayen J. C., Segev I., Mansvelder H. D., & de Kock C. P. J. (2015). Dendritic and Axonal Architecture of Individual Pyramidal Neurons across Layers of Adult Human Neocortex. *Cerebral Cortex*, 25, 4839–4853.
- Mohanty, R., Mallik, B. K., & Solanki, S. S. (2020). Normalized Approximate Descent Used for Spike Based Automatic Bird Species Recognition System. *International Journal of Speech Technology*, 25, 57–65.
- Mohd Hanifa, R., Isa, K., & Mohamad, S. (2021). A Review on Speaker Recognition: Technology and Challenges. *Computers & Electrical Engineering*, 90, 107005.
- Moldwin, T., & Segev, I. (2020). Perceptron Learning and Classification in a Modeled Cortical Pyramidal Cell. *Frontiers in Computational Neuroscience*, 14, 33.
- Montemurro, M. A., Rasch, M. J., Murayama, Y., Logothetis, N. K., & Panzeri, S. (2008). Phase-Of-Firing Coding of Natural Visual Stimuli in Primary Visual Cortex. *Current Biology*, 18, 375–380.
- Näger, C., Storck, J., & Deco, G. (2002). Speech Recognition with Spiking Neurons and Dynamic Synapses: A Model Motivated by the Human Auditory Pathway. *Neurocomputing*, 44–46, 937–942.
- Nagrani, A., Chung, J. S., & Zisserman, A. (2017). Voxceleb: A Large-Scale Speaker Identification Dataset. *arXiv preprint arXiv:1706.08612*.

- Narciso, D. A., & Martins, F. G. (2020). Application of Machine Learning Tools for Energy Efficiency in Industry: A Review. *Energy Reports*, 6, 1181–1199.
- Newton, M. J., & Smith, L. S. (2012). A Neurally Inspired Musical Instrument Classification System Based upon the Sound Onset. *The Journal of the Acoustical Society of America*, 131, 4785–4798.
- Nowinski, W. L. (2011). Introduction to Brain Anatomy. *Biomechanics of the Brain*, 5–40.
- Nunes, J. A., Macêdo, D., & Zanchettin, C. (2019). Additive Margin Sincnet for Speaker Recognition. *International Joint Conference on Neural Networks*, s. 1–5.
- Nunes, J. A., Macêdo, D., & Zanchettin, C. (2020). Am-Mobilenet1D: A Portable Model for Speaker Recognition. *International Joint Conference on Neural Networks*, s. 1–8.
- Okabe, K., Koshinaka, T., & Shinoda, K. (2018). Attentive Statistics Pooling for Deep Speaker Embedding. *arXiv preprint arXiv:1803.10963*.
- Osborn, L. E., Dragomir, A., Betthausen, J. L., Hunt, C. L., Nguyen, H. H., Kaliki, R. R., & Thakor, N. V. (2018). Prosthesis with Neuromorphic Multilayered E-Dermis Perceives Touch and Pain. *Science Robotics*, 3, 3818.
- Özcan, Z., & Kayıkçıoğlu, T. (2021). Evaluating MFCC-Based Speaker Identification Systems with Data Envelopment Analysis. *Expert Systems with Applications*, 168, 114448.
- Özcan, Z., Kayıkçıoğlu, İ., Yıldırım, Ö., Köse, C., & Kayıkçıoğlu, T. (2018). Cell Master: A Versatile and User-Friendly Educational Software for Simulation of Neuronal Dynamics. *Medical Technologies National Congress*, s. 1–4.
- Özcan, Z., Yıldırım, Ö., & Kayıkçıoğlu, T. (2021). Broomycell: An Introductory Educational Software for Neuroscience. *Signal Processing and Communications Applications Conference*, s. 1–4.
- Özdemir, A. (2016). Only One Neuron Either N-Bit Parity Rule Based Modified Translated Multiplicative or Mcculloch-Pitts Models for Some Machine Learning Problems. *International Journal of Intelligent Systems and Applications in Engineering*, 4, 67–72.
- Panayotov, V., Chen, G., Povey, D., & Khudanpur, S. (2015). Librispeech: An ASR Corpus Based on Public Domain Audio Books. *IEEE International Conference on Acoustics, Speech and Signal Processing*, s. 5206–5210.
- Parveen, S., Jain, S., Kumar, S., Acharya, S., & Talwar, D. (2021). Evolution of Middle Ear Modelling Techniques: A Review. *Cureus*, 13.
- Poddar, A., Sahidullah, M., & Saha, G. (2017). Novel Quality Metric for Duration Variability Compensation in Speaker Verification Using I-Vectors. *International Conference on Advances in Pattern Recognition*, s. 1–6.

- Poirazi, P., & Mel, B. W. (2001). Impact of Active Dendrites and Structural Plasticity on the Memory Capacity of Neural Tissue. *Neuron*, 29, 779–796.
- Poirazi, P., Brannon, T., & Mel, B. W. (2003). Pyramidal Neuron As Two-Layer Neural Network. *Neuron*, 37, 989–999.
- Polsky, A., Mel, B. W., & Schiller, J. (2004). Computational Subunits in Thin Dendrites of Pyramidal Cells. *Nature Neuroscience*, 7, 621–627.
- Prachi, N. N., Nahiyan, F. M., Habibullah, M., & Khan, R. (2022). Deep Learning Based Speaker Recognition System with CNN and LSTM Techniques. *Interdisciplinary Research in Technology and Management*, s. 1–6.
- Qu, J., & Wang, R. (2017). Collective Behavior of Large-Scale Neural Networks with GPU Acceleration. *Cognitive Neurodynamics*, 11, 553–563.
- Ravanelli, M., & Bengio, Y. (2018). Speaker Recognition from Raw Waveform with Sincnet. *IEEE Spoken Language Technology Workshop*, s. 1021–1028.
- Redner, R. A., & Walker, H. F. (1984). Mixture Densities, Maximum Likelihood and the EM Algorithm. *SIAM Review*, 26, 195–239.
- Rettig J., Wunder F., Stocker M., Lichtinghagen R., Mastiaux F., Beckh S., Kues W., Pedarzani P., Schröter K. H., & Ruppersberg J. P. (1992). Characterization of a Shaw-Related Potassium Channel Family in Rat Brain.. *The EMBO Journal*, 11, 2473–2486.
- Reuveni, I., Friedman, A., Amitai, Y., & Gutnick, M. J. (1993). Stepwise Repolarization from Ca^{2+} Plateaus in Neocortical Pyramidal Cells: Evidence for Nonhomogeneous Distribution of HVA Ca^{2+} Channels in Dendrites. *The Journal of Neuroscience*, 13, 4609–4621.
- Reynolds, D. A. (1995). Speaker Identification and Verification Using Gaussian Mixture Speaker Models. *Speech Communication*, 17, 91–108.
- Reynolds, D. A., Quatieri, T. F., & Dunn, R. B. (2000). Speaker Verification Using Adapted Gaussian Mixture Models. *Digital Signal Processing*, 10, 19–41.
- Ribas, D., & Vincent, E. (2019). An Improved Uncertainty Propagation Method for Robust I-Vector Based Speaker Recognition. *IEEE International Conference on Acoustics, Speech and Signal Processing*, s. 6331–6335.
- Rohdin, J., Silnova, A., Diez, M., Plchot, O., Matejka, P., & Burget, L. (2018). End-To-End DNN Based Speaker Recognition Inspired by I-Vector and PLDA. *IEEE International Conference on Acoustics, Speech and Signal Processing*, s. 4874–4878.
- Rosenblatt, F. (1958). The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain.. *Psychological Review*, 65, 386–408.

- Sadjadi, S. O., Greenberg, C., Singer, E., Mason, L., & Reynolds, D. (2022). The 2021 NIST Speaker Recognition Evaluation. *arXiv preprint arXiv:2204.10242*.
- Schrauwen, B., & Van Campenhout, I. (2003). BSA, a Fast and Accurate Spike Train Encoding Scheme. *Proceedings of the International Joint Conference on Neural Networks*, 4, s. 2825–2830.
- Sendin, G., Bourien, J., Rassendren, F., Puel, J. L., & Nouvian, R. (2014). Spatiotemporal Pattern of Action Potential Firing in Developing Inner Hair Cells of the Mouse Cochlea. *Proceedings of the National Academy of Sciences*, 111, 1999–2004.
- Seshadri, A., Gandhi, H. A., Wellawatte, G. P., & White, A. D. (2022). Why Does That Molecule Smell. *ChemRxiv*.
- Sgritta, M., Locatelli, F., Soda, T., Prestori, F., & D'Angelo, E. U. (2017). Hebbian Spike-Timing Dependent Plasticity at the Cerebellar Input Stage. *The Journal of Neuroscience*, 37, 2809–2823.
- Shai, A. S., Anastassiou, C. A., Larkum, M. E., & Koch, C. (2015). Physiology of Layer 5 Pyramidal Neurons in Mouse Primary Visual Cortex: Coincidence Detection through Bursting. *Plos Computational Biology*, 11, e1004090.
- Shirahata, T. (2016). Quantitative Evaluations of the Contribution of the Excitatory Ionic Conductance to Repetitive Spiking in a Mathematical Model of Medial Vestibular Nucleus Neurons. *Acta Biologica Hungarica*, 67, 215–219.
- Sidey-Gibbons, J. A., & Sidey-Gibbons, C. J. (2019). Machine Learning in Medicine: A Practical Introduction. *BMC Medical Research Methodology*, 19, 1–18.
- Singleton E. H., Pijnenburg Y. A. L., Gami-Patel P., Boon B. D. C., Bouwman F., Papma J. M., Seelaar H., Scheltens P., Grinberg L. T., Spina S., Nana A. L., Rabinovici G. D., Seeley W. W., Ossenkoppele R., & Dijkstra A. A. (2022). The Behavioral Variant of Alzheimer's Disease Does Not Show a Selective Loss of Von Economo and Phylogenetically Related Neurons in the Anterior Cingulate Cortex. *Alzheimer's Research & Therapy*, 14, 1–8.
- Snyder, D., Garcia-Romero, D., Sell, G., Povey, D., & Khudanpur, S. (2018). X-Vectors: Robust DNN Embeddings for Speaker Recognition. *IEEE International Conference on Acoustics, Speech and Signal Processing*, s. 5329–5333.
- Sporns, O. (2007). Brain Connectivity. *Scholarpedia*, 2, 4695.
- Spruston, N. (2008). Pyramidal Neurons: Dendritic Structure and Synaptic Integration. *Nature Reviews Neuroscience*, 9, 206–221.
- Sterratt, D., Graham, B., Gillies, A., Einevoll, G., & Willshaw, D. (2023). *Principles of Computational Modelling in Neuroscience*. Cambridge University Press.
- Storchi, R., Bale, M. R., Biella, G. E., & Petersen, R. S. (2012). Comparison of Latency and Rate Coding for the Direction of Whisker Deflection in the Subcortical Somatosensory Pathway. *Journal of Neurophysiology*, 108, 1810–1821.

- Sun, C., Yang, Y., Wen, C., Xie, K., & Wen, F. (2018). Voiceprint Identification for Limited Dataset Using the Deep Migration Hybrid Model Based on Transfer Learning. *Sensors*, 18, 2399.
- Sun, P., Chua, Y., Devos, P., & Botteldooren, D. (2023). Learnable Axonal Delay in Spiking Neural Networks Improves Spoken Word Recognition. *Frontiers in Neuroscience*, 17, 1275944.
- Sztahó, D., Szaszák, G., & Beke, A. (2021). Deep Learning Methods in Speaker Recognition: A Review. *Periodica Polytechnica Electrical Engineering and Computer Science*, 65, 310–328.
- Taju, S. W., & Ou, Y. Y. (2019). DeepIon: Deep Learning Approach for Classifying Ion Transporters and Ion Channels from Membrane Proteins. *Journal of Computational Chemistry*, 40, 1521–1529.
- Tavanaei, A., & Maida, A. (2017). Bio-Inspired Multi-Layer Spiking Neural Network Extracts Discriminative Features from Speech Signals. *Neural Information Processing: 24th International Conference*, s. 899–908.
- Thiyagalingam, J., Shankar, M., Fox, G., & Hey, T. (2022). Scientific Machine Learning Benchmarks. *Nature Reviews Physics*, 4, 413–420.
- Thorpe, S., & Gautrais, J. (1998). Rank Order Coding. *Computational Neuroscience: Trends in Research*, s. 113–118.
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of Processing in the Human Visual System. *Nature*, 381, 520–522.
- Timoszczuk, A. P., & Cabral, E. F. (2007). Speaker Recognition Using Pulse Coupled Neural Networks. *International Joint Conference on Neural Networks*, s. 1965–1969.
- Ujfalussy, B. B., Makara, J. K., Lengyel, M., & Branco, T. (2018). Global and Multiplexed Dendritic Computations under In-vivo-like Conditions. *Neuron*, 100, 579–592.e5.
- van der Maaten, L., & Hinton, G. (2008). Visualizing Data Using t-SNE. *Journal of Machine Learning Research*, 9.
- Variani, E., Lei, X., McDermott, E., Moreno, I. L., & Gonzalez-Dominguez, J. (2014). Deep Neural Networks for Small Footprint Text-Dependent Speaker Verification. *IEEE International Conference on Acoustics, Speech and Signal Processing*, s. 4052–4056.
- Verstraeten, D., Schrauwen, B., Stroobandt, D., & Van Campenhout, J. (2005). Isolated Word Recognition with the Liquid State Machine: A Case Study. *Information Processing Letters*, 95, 521–528.
- Volandri, G., Carmignani, C., Puccio, F. D., & Forte, P. (2015). Finite Element Formulations Applied to Outer Ear Modeling. *Strojniški Vestnik – Journal of Mechanical Engineering*, 60, 363–372.

- Wade, J. J., McDaid, L. J., Santos, J. A., & Sayers, H. M. (2010). SWAT: A Spiking Neural Network Training Algorithm for Classification Problems. *IEEE Transactions on Neural Networks*, 21, 1817–1830.
- Wang, J., Wang, K. C., Law, M. T., Rudzicz, F., & Brudno, M. (2019). Centroid-Based Deep Metric Learning for Speaker Recognition. *IEEE International Conference on Acoustics, Speech and Signal Processing*, s. 3652–3656.
- Wang, M., Chen, Y., Tang, Z., & Zhang, E. (2015). I-Vector Based Speaker Gender Recognition. *IEEE Advanced Information Technology, Electronic and Automation Control Conference*, s. 729–732.
- Wang, W., & Siau, K. (2019). Artificial Intelligence, Machine Learning, Automation, Robotics, Future of Work and Future of Humanity: A Review and Research Agenda. *Journal of Database Management (JDM)*, 30, 61–79.
- Wang, X., Lin, X., & Dang, X. (2020). Supervised Learning in Spiking Neural Networks: A Review of Algorithms and Evaluations. *Neural Networks*, 125, 258–280.
- Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016). A Survey of Transfer Learning. *Journal of Big Data*, 3, 1–40.
- Wen, Y., Zhou, T., Singh, R., & Raj, B. (2018). A Corrective Learning Approach for Text-Independent Speaker Verification. *International Conference on Acoustics, Speech and Signal Processing*, s. 4894–4898.
- Wichmann, C., & Moser, T. (2015). Relating Structure and Function of Inner Hair Cell Ribbon Synapses. *Cell and Tissue Research*, 361, 95–114.
- Wildermoth, B. R., & Paliwal, K. K. (2003). GMM Based Speaker Recognition on Readily Available Databases. *Microelectronic Engineering Research Conference*, 7, s. 55.
- Wu, J., Chua, Y., Zhang, M., Li, H., & Tan, K. C. (2018). A Spiking Neural Network Framework for Robust Sound Classification. *Frontiers in Neuroscience*, 12, 379777.
- Wu, J., Liu, Q., Zhang, M., Pan, Z., Li, H., & Tan, K. C. (2021). HuRAI: A Brain-Inspired Computational Model for Human-Robot Auditory Interface. *Neurocomputing*, 465, 103–113.
- Wu X., Shukla R., Alganem K., Depasquale E., Reigle J., Simmons M., Hahn C. G., Haroutunian V., Meller J., Meador-Woodruff J., & McCullumsmith R. (2020). Transcriptional Profile of Pyramidal Neurons in Chronic Schizophrenia Reveals Lamina-Specific Dysfunction of Neuronal Immunity. *Biological Psychiatry*, 87, S347–S348.
- Wysoski, S. G., Benuskova, L., & Kasabov, N. (2007). Text-Independent Speaker Authentication with Spiking Neural Networks. *ICANN*, s. 758–767.
- Xanthopoulos, P., Pardalos, P. M., Trafalis, T. B., Xanthopoulos, P., Pardalos, P. M., & Trafalis, T. B. (2013). Linear Discriminant Analysis. *Robust Data Mining*, 27–33.

- Xiao, R., Yan, R., Tang, H., & Tan, K. C. (2017). A Spiking Neural Network Model for Sound Recognition. *Cognitive Systems and Signal Processing: Third International Conference*, s. 584–594.
- Yamazaki, K., Vo-Ho, V. K., Bulsara, D., & Le, N. (2022). Spiking Neural Networks and Their Applications: A Review. *Brain Sciences*, 12, 863.
- Zaidan, O. F., & Callison-Burch, C. (2014). Arabic Dialect Identification. *Computational Linguistics*, 40, 171–202.
- Zhang, Y., Li, P., Jin, Y., & Choe, Y. (2015). A Digital Liquid State Machine with Biologically Inspired Learning and Its Application to Speech Recognition. *IEEE Transactions on Neural Networks and Learning Systems*, 26, 2635–2649.
- Zheng, F., Zhang, G., & Song, Z. (2001). Comparison of Different Implementations of MFCC. *Journal of Computer Science and Technology*, 16, 582–589.
- Zhong, L., Xie, W., Yin, J., & Jie, W. (2023). Chemical-Vapor-Deposited 2D VSe₂ Nanosheet with Threshold Switching Behaviors for Boolean Logic Calculations and Leaky Integrate-and-Fire Functions. *Journal of Materials Chemistry C*, 11, 5032–5038.
- Zhou, Z. H. (2021). *Machine Learning*. Springer Singapore.

6. EKLER

EK-1: Kullanılan Aktif İyon Kanalları için Matematiksel İfadeler

Aşağıda, Shai ve arkadaşlarının (2015) çalışmalarında piramidal nöronun modellenmesinde kullandıkları aktif iyon kanallarına ait matematiksel arka plan verilmiştir.

K1: Hiperpolarizasyonda aktif olan katyon akımları (Kole vd., 2006):

$$I_{kanal} = \bar{g}_{kanal} m (V_{zar} - E_{kanal}) \quad (65)$$

$$\alpha_m = 6,43 \times 10^{-3} \times \frac{V_{zar} + 155}{e^{\left(\frac{V_{zar} + 155}{12}\right)} - 1} \quad (66)$$

$$\beta_m = 193 \times 10^{-3} \times e^{\left(\frac{V_{zar}}{33}\right)} \quad (67)$$

K2: A tipi Sodyum kanalı (Colbert & Pan, 2002):

$$I_{kanal} = \bar{g}_{kanal} m^3 h (V_{zar} - E_{Na}) \quad (68)$$

$$\alpha_m = 1,82 \times 10^{-1} \times \frac{V_{zar} - (-38)}{1 - e^{-\left(\frac{V_{zar} - (-38)}{6}\right)}} \quad (69)$$

$$\beta_m = -1,24 \times 10^{-1} \times \frac{V_{zar} + 38}{1 - e^{-\left(\frac{V_{zar} - (-38)}{6}\right)}} \quad (70)$$

K3: Sürekli Sodyum kanalı (Magistretti & Alonso, 1999):

$$I_{kanal} = \bar{g}_{kanal} m^3 h (V_{zar} - E_{Na}) \quad (71)$$

$$\alpha_m = 1,82 \times 10^{-1} \times \frac{V_{zar} - (-38)}{1 - e^{-\left(\frac{V_{zar} - (-38)}{6}\right)}} \quad (72)$$

$$\beta_m = -1,24 \times 10^{-1} \times \frac{V_{zar} + 38}{1 - e^{-\left(\frac{V_{zar} - (-38)}{6}\right)}} \quad (73)$$

$$\alpha_h = -2,88 \times 10^{-6} \times \frac{V_{zar} + 17}{1 - e^{\left(\frac{V_{zar}+17}{4,63}\right)}} \quad (74)$$

$$\beta_h = 6,94 \times 10^{-6} \times \frac{V_{zar} + 64,4}{1 - e^{-\left(\frac{V_{zar}+64,4}{2,63}\right)}} \quad (75)$$

K4: M tipi Potasyum kanalı (Adams vd., 1982):

$$I_{kanal} = \bar{g}_{kanal} m (V_{zar} - E_K) \quad (76)$$

$$\alpha_m = 3,3 \times 10^{-3} \times e^{0,1(V_{zar}-(-35))} \quad (77)$$

$$\beta_m = 3,3 \times 10^{-3} \times e^{-0,1(V_{zar}-(-35))} \quad (78)$$

K5: Shaw ilişkili Potasyum kanalı (Rettig vd., 1992):

$$I_{kanal} = \bar{g}_{kanal} m (V_{zar} - E_K) \quad (79)$$

$$m_{\infty} = \frac{1}{1 + e^{\left(\frac{V_{zar}-18,7}{-9,7}\right)}} \quad (80)$$

$$m_{\tau} = \frac{4}{1 + e^{\left(\frac{V_{zar}-(-46,56)}{-44,14}\right)}} \quad (81)$$

K6: Hızlı Potasyum kanalı (Korngreen & Sakmann, 2000):

$$I_{kanal} = \bar{g}_{kanal} m^4 h (V_{zar} - E_K) \quad (82)$$

$$m_{\infty} = \frac{1}{1 + e^{-\left(\frac{V_{zar}+0}{19}\right)}} \quad (83)$$

$$m_{\tau} = 0,34 + 0,92 \times e^{-\left(\left(\frac{V_{zar}+71}{59}\right)^2\right)} \quad (84)$$

$$h_{\infty} = \frac{1}{1 + e^{-\left(\frac{V_{zar}+66}{-10}\right)}} \quad (85)$$

$$h_{\tau} = 8 + 49 \times e^{-\left(\left(\frac{V_{zar}+73}{23}\right)^2\right)} \quad (86)$$

K7: Yavaş Potasyum kanalı (Korngreen & Sakmann, 2000):

$$I_{kanal} = \bar{g}_{kanal} m^2 h (V_{zar} - E_K) \quad (87)$$

$$m_{\infty} = \frac{1}{1 + e^{-\left(\frac{V_{zar}+1}{12}\right)}} \quad (88)$$

$$m_{\tau} = \begin{cases} 1,25 + 175,03 \times e^{0,026 \times V_{zar}} & V_{zar} < -50 \text{ mV} \\ 1,25 + 13 \times e^{-0,026 \times V_{zar}} & V_{zar} \geq -50 \text{ mV} \end{cases} \quad (89)$$

$$h_{\infty} = \frac{1}{1 + e^{-\left(\frac{V_{zar}+54}{-11}\right)}} \quad (90)$$

$$h_{\tau} = 360 + (1010 + 24 \times (V_{zar} + 55)) \times e^{-\left(\left(\frac{V_{zar}+75}{48}\right)^2\right)} \quad (91)$$

K8: SK tipi Kalsiyum bağımlı Potasyum kanalı (Köhler vd., 1996):

$$I_{kanal} = \bar{g}_{kanal} z (V_{zar} - E_K) \quad (92)$$

$$z_{\infty} = \frac{1}{1 + \left(\frac{4,3 \times 10^{-4}}{[Ca^{2+}]_i}\right)^{4,8}} \quad (93)$$

$$z_{\tau} = 1 \quad (94)$$

$$\frac{d[Ca^{2+}]_i(t)}{dt} = \frac{\gamma \times I_{Ca}(t)}{2 \times Depth} - \frac{[Ca^{2+}]_i(t) - [Ca^{2+}]_{min}}{\tau_{decay}} \quad (95)$$

K9: Düşük gerilimde aktif olan Kalsiyum kanalı (Avery & Johnston, 1996):

$$I_{kanal} = \bar{g}_{kanal} m^2 h (V_{zar} - E_{Ca}) \quad (96)$$

$$m_{\infty} = \frac{1}{1 + e^{-\left(\frac{V_{zar}-(-30)}{-6}\right)}} \quad (97)$$

$$m_{\tau} = 5 + \frac{20}{1 + e^{-\left(\frac{V_{zar}-(-25)}{5}\right)}} \quad (98)$$

$$h_{\infty} = \frac{1}{1 + e^{-\left(\frac{V_{zar}-(-80)}{6,4}\right)}} \quad (99)$$

$$h_{\tau} = 20 + \frac{50}{1 + e^{\left(\frac{V_{zar} - (-40)}{7}\right)}} \quad (100)$$

K10: Yüksek gerilimde aktif olan Kalsiyum kanalı (Reuveni vd., 1993):

$$I_{kanal} = \bar{g}_{kanal} m^2 h (V_{zar} - E_{Ca}) \quad (101)$$

$$\alpha_m = -5,5 \times 10^{-2} \times \frac{V_{zar} + 27}{e^{-\left(\frac{V_{zar} + 27}{3,8}\right)} - 1} \quad (102)$$

$$\beta_m = 0,94 \times e^{-\left(\frac{V_{zar} + 75}{17}\right)} \quad (103)$$

$$\alpha_h = -4,57 \times 10^{-4} \times e^{-\left(\frac{V_{zar} + 13}{50}\right)} \quad (104)$$

$$\beta_h = \frac{6,5 \times 10^{-3}}{e^{-\left(\frac{V_{zar} + 15}{28}\right)} + 1} \quad (105)$$

EK-2: D1 için Ek Bulgular

Aşağıda, "D1: Otomatik Konuşmacı Tanımda En Gelişkin Yöntemlerin Performansının İncelenmesi" başlığı altında detayları verilen deneysel senaryoda ilk adıma dair detaylı bulgular tablolar halinde verilmiştir. Tablolarda verilen bulgular çeşitli parametre ve yöntemlere göre organize edilmiştir. Tüm parametre ve yöntemler için deneyler ellışer kez tekrar edildiğinden bulguların istatistikleri "Ortalama $\pm$ Standart Sapma" formatında verilmiştir. Tüm bulgular yüzdelik F1-Skor ölçütüyle paylaşılmıştır.

Ek Tablo 1. GKM deneylerinde karışım sayısı 16 seçildiğinde performanslar

EVU	TVU	#MFKK	Konuşmacı Sayısı		
			10	20	40
5	1	5	73,41 $\pm$ 7,71	62,93 $\pm$ 5,21	54,61 $\pm$ 3,08
		10	80,94 $\pm$ 7,34	74,72 $\pm$ 4,03	68,09 $\pm$ 3,80
		20	81,60 $\pm$ 7,10	75,63 $\pm$ 4,70	70,05 $\pm$ 3,15
	2	5	79,26 $\pm$ 7,78	70,26 $\pm$ 5,47	63,02 $\pm$ 3,44
		10	85,88 $\pm$ 6,80	80,79 $\pm$ 4,54	75,73 $\pm$ 3,50
		20	85,65 $\pm$ 6,54	81,26 $\pm$ 4,58	77,13 $\pm$ 3,25
	4	5	83,40 $\pm$ 8,13	75,92 $\pm$ 5,49	68,96 $\pm$ 4,33
		10	88,77 $\pm$ 6,77	85,42 $\pm$ 4,17	80,87 $\pm$ 3,88
		20	88,40 $\pm$ 6,28	85,01 $\pm$ 4,82	81,45 $\pm$ 3,43
10	1	5	83,77 $\pm$ 5,55	75,05 $\pm$ 4,31	67,00 $\pm$ 3,22
		10	90,66 $\pm$ 5,35	85,20 $\pm$ 4,17	81,28 $\pm$ 2,61
		20	90,83 $\pm$ 4,04	87,23 $\pm$ 3,82	83,05 $\pm$ 2,40
	2	5	88,18 $\pm$ 5,93	83,31 $\pm$ 3,78	77,24 $\pm$ 2,93
		10	93,80 $\pm$ 4,61	90,21 $\pm$ 3,72	88,36 $\pm$ 2,62
		20	93,91 $\pm$ 4,04	90,96 $\pm$ 3,64	88,91 $\pm$ 2,22
	4	5	91,58 $\pm$ 5,33	88,28 $\pm$ 3,48	84,69 $\pm$ 2,56
		10	96,08 $\pm$ 3,79	92,90 $\pm$ 3,54	92,21 $\pm$ 2,19
		20	95,68 $\pm$ 3,89	93,71 $\pm$ 3,41	92,04 $\pm$ 2,28
20	1	5	89,14 $\pm$ 4,60	83,24 $\pm$ 3,32	76,50 $\pm$ 2,54
		10	94,54 $\pm$ 3,52	91,87 $\pm$ 2,34	88,50 $\pm$ 2,12
		20	95,84 $\pm$ 2,34	93,12 $\pm$ 2,46	90,41 $\pm$ 1,83
	2	5	93,84 $\pm$ 4,15	90,30 $\pm$ 3,07	86,35 $\pm$ 2,16
		10	96,88 $\pm$ 3,06	95,13 $\pm$ 2,18	93,71 $\pm$ 2,13
		20	97,81 $\pm$ 1,78	96,14 $\pm$ 1,65	94,30 $\pm$ 1,35
	4	5	96,36 $\pm$ 2,78	94,47 $\pm$ 2,57	92,11 $\pm$ 1,80
		10	97,92 $\pm$ 2,47	97,08 $\pm$ 1,75	96,07 $\pm$ 1,79
		20	98,49 $\pm$ 1,88	97,64 $\pm$ 1,52	96,65 $\pm$ 1,13

Ek Tablo 2. GKM deneylerinde karışım sayısı 32 seçildiğinde performanslar

EVU	TVU	#MFKK	Konuşmacı Sayısı		
			10	20	40
5	1	5	70,16 ± 7,32	59,04 ± 5,44	49,82 ± 3,12
		10	78,22 ± 7,92	71,26 ± 5,17	63,09 ± 3,86
		20	77,92 ± 7,61	70,31 ± 4,74	64,12 ± 3,54
	2	5	76,37 ± 7,46	65,54 ± 6,03	56,78 ± 3,87
		10	81,37 ± 8,62	76,51 ± 5,52	70,14 ± 4,47
		20	81,03 ± 7,70	75,99 ± 5,10	70,56 ± 4,30
	4	5	79,53 ± 8,47	70,77 ± 6,69	62,27 ± 4,36
		10	84,68 ± 8,52	80,07 ± 5,51	74,41 ± 4,68
		20	84,12 ± 7,96	78,93 ± 5,29	74,40 ± 4,59
10	1	5	82,88 ± 5,58	73,87 ± 3,86	66,74 ± 3,05
		10	90,69 ± 4,50	84,70 ± 3,51	81,22 ± 2,31
		20	90,53 ± 4,30	86,29 ± 4,30	82,52 ± 2,29
	2	5	87,73 ± 6,38	81,82 ± 3,59	76,12 ± 2,95
		10	93,90 ± 3,94	89,52 ± 3,43	87,36 ± 2,29
		20	93,71 ± 3,93	90,30 ± 3,82	87,93 ± 1,91
	4	5	91,48 ± 5,55	86,64 ± 3,75	82,81 ± 2,84
		10	95,71 ± 3,77	92,72 ± 3,49	91,20 ± 2,12
		20	95,65 ± 3,80	93,37 ± 3,27	91,36 ± 2,08
20	1	5	89,51 ± 4,85	83,89 ± 3,83	77,26 ± 2,57
		10	95,45 ± 3,71	92,33 ± 2,63	89,48 ± 2,06
		20	96,09 ± 2,42	93,46 ± 2,09	91,27 ± 1,75
	2	5	94,41 ± 3,41	90,77 ± 2,57	86,71 ± 2,30
		10	97,33 ± 2,64	95,67 ± 2,21	94,04 ± 1,73
		20	97,75 ± 1,84	96,31 ± 1,72	94,86 ± 1,52
	4	5	96,32 ± 2,92	94,67 ± 2,09	92,26 ± 1,79
		10	98,36 ± 1,81	97,31 ± 1,71	96,28 ± 1,65
		20	98,77 ± 1,51	97,80 ± 1,29	96,82 ± 1,38

Ek Tablo 3. GKM deneylerinde karışım sayısı 64 seçildiğinde performanslar

EVU	TVU	#MFKK	Konuşmacı Sayısı		
			10	20	40
5	1	5	66,15 ± 7,19	52,84 ± 6,06	42,47 ± 2,98
		10	71,52 ± 9,11	62,90 ± 4,95	54,31 ± 4,36
		20	69,72 ± 9,01	59,78 ± 7,12	52,28 ± 4,90
	2	5	70,81 ± 7,99	57,78 ± 7,20	48,10 ± 4,04
		10	75,51 ± 9,61	68,12 ± 5,96	59,35 ± 5,11
		20	73,22 ± 10,87	63,85 ± 8,96	56,67 ± 5,50
	4	5	73,37 ± 9,44	61,30 ± 8,21	51,78 ± 4,73
		10	78,70 ± 10,23	70,92 ± 6,33	62,90 ± 5,72
		20	74,90 ± 11,19	65,58 ± 10,01	59,44 ± 6,48
10	1	5	80,52 ± 6,75	70,78 ± 3,93	62,24 ± 2,86
		10	89,51 ± 4,40	81,50 ± 4,84	77,33 ± 3,03
		20	87,72 ± 5,02	82,92 ± 4,49	78,60 ± 2,96
	2	5	84,94 ± 7,47	78,48 ± 4,52	71,40 ± 2,99
		10	92,26 ± 4,89	86,47 ± 4,22	83,91 ± 3,16
		20	90,99 ± 4,44	87,73 ± 4,41	84,86 ± 2,67
	4	5	88,84 ± 6,54	83,68 ± 4,68	78,05 ± 3,02
		10	94,30 ± 4,44	89,71 ± 3,93	88,40 ± 2,85
		20	93,61 ± 4,10	91,09 ± 4,23	88,51 ± 2,93
20	1	5	88,83 ± 5,07	82,91 ± 3,31	76,58 ± 2,54
		10	94,88 ± 3,54	91,80 ± 2,56	89,32 ± 2,05
		20	95,45 ± 2,88	93,05 ± 2,33	90,86 ± 1,55
	2	5	93,49 ± 3,59	90,06 ± 3,24	85,88 ± 2,16
		10	97,29 ± 2,59	95,40 ± 2,09	93,69 ± 1,93
		20	97,67 ± 2,19	96,00 ± 1,79	94,77 ± 1,34
	4	5	95,59 ± 3,22	93,76 ± 2,66	91,06 ± 1,97
		10	98,09 ± 2,84	96,93 ± 1,86	96,11 ± 1,56
		20	98,56 ± 2,01	97,53 ± 1,61	96,59 ± 1,11

Ek Tablo 4. GKM-EAPM deneylerinde karışım sayısı 16 seçildiğinde performanslar

EVU	TVU	#MFKK	Konuşmacı Sayısı		
			10	20	40
5	1	5	76,50 ± 6,24	66,31 ± 5,01	55,80 ± 3,88
		10	86,36 ± 6,08	80,24 ± 4,24	73,53 ± 2,88
		20	86,29 ± 5,69	81,22 ± 4,35	75,66 ± 2,95
	2	5	83,40 ± 7,04	74,70 ± 5,66	66,91 ± 3,68
		10	90,51 ± 5,80	85,95 ± 3,95	81,43 ± 3,25
		20	90,54 ± 5,73	86,63 ± 4,39	82,62 ± 2,90
	4	5	87,77 ± 6,23	80,74 ± 6,03	75,45 ± 3,82
		10	93,04 ± 5,47	89,68 ± 4,20	86,43 ± 3,51
		20	92,15 ± 5,99	90,40 ± 4,25	87,24 ± 2,99
10	1	5	81,73 ± 5,67	72,49 ± 4,83	62,65 ± 2,88
		10	90,15 ± 4,93	85,21 ± 4,08	80,33 ± 2,60
		20	91,49 ± 3,98	87,50 ± 4,48	82,88 ± 2,67
	2	5	86,83 ± 5,88	81,40 ± 4,35	74,65 ± 2,74
		10	93,26 ± 4,78	90,11 ± 4,03	87,19 ± 2,48
		20	95,31 ± 3,35	91,50 ± 3,81	88,71 ± 2,67
	4	5	91,64 ± 4,94	86,86 ± 4,12	82,76 ± 2,78
		10	94,76 ± 4,21	92,68 ± 4,11	91,19 ± 2,40
		20	96,69 ± 3,19	94,08 ± 3,61	92,44 ± 2,27
20	1	5	84,24 ± 5,12	76,46 ± 4,28	67,71 ± 2,73
		10	93,13 ± 3,71	89,56 ± 2,96	84,72 ± 2,78
		20	94,17 ± 3,44	91,49 ± 2,94	88,19 ± 2,55
	2	5	90,67 ± 3,95	84,93 ± 4,07	79,31 ± 2,79
		10	95,91 ± 3,14	93,77 ± 2,65	90,80 ± 2,32
		20	96,54 ± 2,85	95,36 ± 2,24	93,09 ± 2,02
	4	5	93,52 ± 4,15	90,05 ± 3,47	86,65 ± 2,50
		10	97,13 ± 3,21	96,04 ± 2,45	93,98 ± 1,94
		20	98,21 ± 2,37	97,17 ± 1,63	95,98 ± 1,78

Ek Tablo 5. GKM-EAPM deneylerinde karışım sayısı 32 seçildiğinde performanslar

EVU	TVU	#MFKK	Konuşmacı Sayısı		
			10	20	40
5	1	5	78,47 ± 6,26	67,95 ± 5,15	58,87 ± 3,30
		10	88,85 ± 5,47	82,88 ± 4,23	76,28 ± 3,12
		20	87,84 ± 6,12	83,38 ± 4,02	78,19 ± 2,73
	2	5	85,08 ± 5,76	77,03 ± 5,28	69,57 ± 3,85
		10	91,34 ± 5,45	87,45 ± 4,05	83,53 ± 3,14
		20	91,46 ± 5,11	88,44 ± 3,95	84,84 ± 3,00
	4	5	89,29 ± 5,41	82,90 ± 5,36	77,83 ± 3,62
		10	93,78 ± 4,69	90,84 ± 3,84	87,96 ± 3,37
		20	93,76 ± 5,14	91,32 ± 3,60	88,82 ± 2,78
10	1	5	82,89 ± 6,12	75,17 ± 4,69	66,23 ± 2,96
		10	91,53 ± 4,67	87,52 ± 3,65	83,16 ± 2,73
		20	93,56 ± 3,04	88,88 ± 4,22	85,46 ± 2,59
	2	5	89,04 ± 5,44	83,16 ± 4,47	76,97 ± 3,04
		10	93,83 ± 4,85	91,27 ± 3,81	89,09 ± 2,21
		20	95,85 ± 2,87	92,67 ± 3,63	90,84 ± 2,04
	4	5	92,48 ± 4,92	88,14 ± 4,08	84,89 ± 2,69
		10	95,82 ± 3,80	93,60 ± 3,86	92,11 ± 2,21
		20	97,19 ± 2,59	94,47 ± 3,38	93,83 ± 2,12
20	1	5	86,45 ± 4,66	78,94 ± 4,58	70,97 ± 2,40
		10	93,70 ± 4,08	91,39 ± 3,18	87,91 ± 2,15
		20	95,67 ± 3,19	93,68 ± 2,24	90,76 ± 1,95
	2	5	91,73 ± 5,09	87,27 ± 3,23	81,63 ± 2,78
		10	96,38 ± 3,32	94,98 ± 2,20	92,63 ± 1,85
		20	97,41 ± 2,66	96,25 ± 2,10	94,53 ± 1,97
	4	5	95,31 ± 4,21	91,67 ± 3,22	88,33 ± 2,53
		10	97,47 ± 2,97	96,36 ± 2,15	95,04 ± 1,99
		20	98,54 ± 2,07	97,57 ± 1,51	96,60 ± 1,77

Ek Tablo 6. GKM-EAPM deneylerinde karışım sayısı 64 seçildiğinde performanslar

EVU	TVU	#MFKK	Konuşmacı Sayısı		
			10	20	40
5	1	5	78,93 ± 6,47	68,85 ± 5,21	60,04 ± 3,44
		10	88,27 ± 5,92	82,77 ± 4,53	77,75 ± 2,84
		20	87,99 ± 5,82	84,28 ± 3,94	78,89 ± 2,76
	2	5	85,70 ± 5,78	78,04 ± 5,28	70,71 ± 3,53
		10	92,25 ± 5,00	87,81 ± 3,92	84,49 ± 3,23
		20	91,41 ± 5,45	88,95 ± 3,68	85,75 ± 2,66
	4	5	89,88 ± 5,43	84,30 ± 5,05	78,78 ± 3,80
		10	94,08 ± 4,56	91,10 ± 3,95	89,11 ± 3,25
		20	94,00 ± 5,05	91,62 ± 3,65	89,54 ± 2,66
10	1	5	83,75 ± 5,72	76,97 ± 4,64	68,00 ± 2,70
		10	92,14 ± 4,66	88,26 ± 3,25	84,57 ± 2,47
		20	93,65 ± 3,16	89,95 ± 3,64	87,10 ± 2,42
	2	5	89,76 ± 5,61	84,48 ± 4,45	78,89 ± 2,90
		10	94,64 ± 3,95	91,92 ± 3,68	89,84 ± 2,14
		20	95,92 ± 3,09	93,17 ± 3,32	91,55 ± 2,28
	4	5	93,17 ± 4,87	89,08 ± 4,35	85,92 ± 2,79
		10	95,80 ± 3,57	93,96 ± 3,73	92,96 ± 2,27
		20	97,78 ± 2,53	95,05 ± 2,93	94,30 ± 2,06
20	1	5	87,73 ± 4,86	80,57 ± 4,21	73,32 ± 2,97
		10	94,85 ± 3,34	92,18 ± 2,54	89,04 ± 2,06
		20	96,50 ± 2,51	94,12 ± 2,41	92,19 ± 2,03
	2	5	92,30 ± 4,56	87,96 ± 3,89	83,39 ± 2,35
		10	96,59 ± 2,98	95,77 ± 2,03	93,41 ± 2,18
		20	98,00 ± 2,11	96,49 ± 2,10	95,36 ± 1,90
	4	5	95,28 ± 4,19	92,45 ± 3,51	90,19 ± 2,21
		10	97,85 ± 2,41	96,66 ± 2,24	95,91 ± 1,57
		20	99,05 ± 1,51	97,80 ± 1,67	97,27 ± 1,57

Ek Tablo 7. GKM-DVM deneylerinde karışım sayısı 16 seçildiğinde performanslar

EVU	TVU	#MFKK	Konuşmacı Sayısı		
			10	20	40
5	1	5	62,22 ± 9,08	49,86 ± 5,02	37,77 ± 2,59
		10	78,48 ± 7,78	70,31 ± 4,40	61,41 ± 3,52
		20	77,72 ± 6,80	68,56 ± 5,11	57,67 ± 3,50
	2	5	73,94 ± 8,29	62,58 ± 5,07	51,20 ± 3,57
		10	86,45 ± 7,21	80,34 ± 4,84	73,00 ± 3,32
		20	84,24 ± 6,34	77,00 ± 5,14	67,53 ± 3,53
	4	5	82,11 ± 8,01	71,59 ± 5,38	62,20 ± 3,65
		10	90,16 ± 6,90	85,91 ± 4,56	80,52 ± 3,61
		20	87,93 ± 6,44	82,47 ± 4,33	74,41 ± 3,78
10	1	5	68,04 ± 7,68	55,58 ± 5,29	44,99 ± 2,55
		10	85,20 ± 5,36	77,64 ± 4,24	69,57 ± 2,66
		20	83,51 ± 5,61	74,76 ± 5,12	65,46 ± 3,05
	2	5	80,18 ± 7,50	69,93 ± 5,54	61,07 ± 3,57
		10	91,60 ± 4,90	86,47 ± 3,79	81,06 ± 2,88
		20	89,07 ± 5,07	82,98 ± 4,29	76,10 ± 2,93
	4	5	88,15 ± 6,83	81,33 ± 4,71	74,52 ± 3,15
		10	95,05 ± 3,98	91,43 ± 3,19	88,19 ± 2,92
		20	93,12 ± 4,40	88,09 ± 4,45	83,13 ± 2,58
20	1	5	73,43 ± 7,10	60,69 ± 4,24	49,36 ± 2,78
		10	88,44 ± 5,61	82,69 ± 4,14	74,99 ± 2,70
		20	87,15 ± 5,69	80,30 ± 3,89	73,15 ± 2,84
	2	5	85,18 ± 6,37	75,67 ± 4,00	67,74 ± 2,91
		10	94,17 ± 4,19	90,48 ± 3,26	86,31 ± 2,67
		20	92,34 ± 5,04	88,95 ± 3,24	83,15 ± 2,96
	4	5	91,91 ± 4,64	86,15 ± 3,66	81,63 ± 2,57
		10	96,61 ± 2,89	94,52 ± 2,93	92,36 ± 2,14
		20	95,87 ± 3,79	93,17 ± 2,94	89,60 ± 2,28

Ek Tablo 8. GKM- DVM deneylerinde karışım sayısı 32 seçildiğinde performanslar

EVU	TVU	#MFKK	Konuşmacı Sayısı		
			10	20	40
5	1	5	64,02 ± 8,42	52,08 ± 4,66	40,22 ± 2,85
		10	79,44 ± 7,72	72,36 ± 4,78	62,96 ± 3,12
		20	78,71 ± 7,05	67,91 ± 5,28	57,73 ± 3,83
	2	5	75,93 ± 7,35	64,20 ± 6,09	54,11 ± 3,76
		10	86,10 ± 7,16	81,22 ± 4,68	74,19 ± 3,65
		20	85,42 ± 6,70	77,57 ± 4,88	69,17 ± 4,23
	4	5	83,00 ± 7,78	73,19 ± 5,48	64,59 ± 3,70
		10	90,52 ± 6,07	86,77 ± 3,94	81,29 ± 3,32
		20	89,64 ± 5,79	83,56 ± 4,72	76,34 ± 4,26
10	1	5	71,80 ± 7,20	59,34 ± 5,01	48,20 ± 3,30
		10	86,53 ± 5,98	78,96 ± 4,14	71,10 ± 3,02
		20	83,50 ± 6,22	75,10 ± 4,86	65,72 ± 3,09
	2	5	82,33 ± 6,94	73,33 ± 4,84	64,32 ± 3,82
		10	91,84 ± 4,97	87,35 ± 3,46	82,02 ± 3,03
		20	89,69 ± 5,73	84,28 ± 4,58	77,41 ± 3,15
	4	5	89,34 ± 6,16	82,69 ± 4,22	76,48 ± 3,32
		10	94,95 ± 4,09	92,07 ± 3,17	88,48 ± 2,85
		20	94,17 ± 4,18	90,14 ± 4,01	84,92 ± 3,06
20	1	5	76,13 ± 6,23	64,39 ± 4,42	54,00 ± 3,14
		10	89,80 ± 5,16	84,80 ± 3,72	77,87 ± 2,97
		20	86,39 ± 5,69	79,47 ± 4,41	70,87 ± 3,20
	2	5	86,94 ± 6,43	78,32 ± 4,59	70,68 ± 3,19
		10	94,79 ± 3,32	91,87 ± 3,15	87,64 ± 2,41
		20	93,55 ± 3,80	88,45 ± 3,65	83,45 ± 3,26
	4	5	93,55 ± 4,47	87,72 ± 3,80	83,84 ± 2,89
		10	97,29 ± 2,49	94,88 ± 2,46	92,69 ± 2,20
		20	96,63 ± 2,94	93,77 ± 3,05	90,17 ± 2,60

Ek Tablo 9. GKM- DVM deneylerinde karışım sayısı 64 seçildiğinde performanslar

EVU	TVU	#MFKK	Konuşmacı Sayısı		
			10	20	40
5	1	5	65,75 ± 8,66	53,35 ± 4,55	41,46 ± 3,43
		10	81,07 ± 8,11	73,34 ± 4,63	64,57 ± 3,80
		20	77,34 ± 8,19	68,42 ± 5,58	58,93 ± 4,44
	2	5	76,86 ± 6,84	65,23 ± 5,62	54,56 ± 3,41
		10	86,94 ± 6,76	82,59 ± 4,55	75,39 ± 3,16
		20	85,84 ± 6,44	78,87 ± 5,67	71,68 ± 4,34
	4	5	83,32 ± 7,92	74,61 ± 5,14	65,16 ± 4,37
		10	90,94 ± 6,19	87,58 ± 4,58	82,50 ± 3,37
		20	90,29 ± 6,33	85,51 ± 4,57	79,51 ± 4,07
10	1	5	72,29 ± 8,14	61,05 ± 4,59	49,33 ± 3,29
		10	87,12 ± 5,56	80,32 ± 4,58	73,22 ± 2,84
		20	84,23 ± 5,55	77,37 ± 5,34	68,66 ± 3,49
	2	5	83,25 ± 6,99	74,27 ± 5,44	64,97 ± 3,11
		10	92,52 ± 4,68	88,67 ± 3,65	83,50 ± 3,28
		20	91,09 ± 5,15	86,88 ± 4,68	80,86 ± 3,33
	4	5	89,81 ± 5,94	83,75 ± 4,06	76,61 ± 3,29
		10	95,47 ± 4,06	92,59 ± 3,10	89,03 ± 2,62
		20	94,88 ± 4,57	91,78 ± 4,38	88,16 ± 2,64
20	1	5	77,90 ± 7,17	65,98 ± 4,56	56,26 ± 3,06
		10	90,63 ± 4,50	86,17 ± 3,32	79,65 ± 2,71
		20	88,48 ± 4,83	80,84 ± 5,55	73,75 ± 4,10
	2	5	87,97 ± 5,54	80,10 ± 3,93	72,64 ± 2,76
		10	95,49 ± 3,49	92,67 ± 2,95	88,86 ± 2,11
		20	94,77 ± 3,81	90,95 ± 3,61	87,13 ± 2,90
	4	5	94,07 ± 4,40	88,27 ± 3,40	84,48 ± 2,71
		10	97,48 ± 2,25	95,31 ± 2,60	93,38 ± 2,03
		20	97,98 ± 2,65	95,48 ± 2,61	93,10 ± 2,09

Ek Tablo 10. DSA deneylerinde gizli katman sayısı 2 seçildiğinde performanslar

EVU	TVU	#MFKK	Konuşmacı Sayısı		
			10	20	40
5	1	5	75,47 ± 5,43	63,21 ± 4,50	50,47 ± 2,91
		10	85,89 ± 5,51	80,67 ± 4,09	72,88 ± 2,65
		20	89,05 ± 5,02	84,12 ± 4,12	79,48 ± 2,56
	2	5	82,83 ± 4,81	73,05 ± 4,44	62,46 ± 3,18
		10	90,04 ± 5,78	87,73 ± 3,63	82,24 ± 2,76
		20	92,50 ± 4,99	88,83 ± 3,88	86,74 ± 2,40
	4	5	87,46 ± 4,75	80,49 ± 4,76	71,95 ± 3,17
		10	92,95 ± 5,45	91,33 ± 3,66	88,19 ± 2,76
		20	94,20 ± 4,32	91,78 ± 3,69	91,22 ± 2,48
10	1	5	81,71 ± 6,01	71,59 ± 3,86	58,72 ± 2,99
		10	92,24 ± 3,27	87,59 ± 3,52	82,34 ± 2,94
		20	93,90 ± 3,93	91,17 ± 3,29	88,40 ± 2,50
	2	5	87,14 ± 5,81	81,03 ± 3,47	70,33 ± 3,08
		10	95,19 ± 3,40	92,09 ± 2,77	88,92 ± 2,68
		20	95,80 ± 3,74	94,17 ± 3,05	92,90 ± 2,10
	4	5	92,19 ± 4,82	87,87 ± 3,46	79,25 ± 3,14
		10	96,87 ± 3,28	94,70 ± 2,60	92,50 ± 2,33
		20	96,96 ± 3,21	96,05 ± 2,81	95,39 ± 2,01
20	1	5	87,20 ± 4,37	77,74 ± 3,96	65,77 ± 2,99
		10	95,87 ± 2,78	92,94 ± 2,34	88,01 ± 2,09
		20	97,55 ± 2,51	96,04 ± 1,90	93,80 ± 1,64
	2	5	92,15 ± 4,42	85,64 ± 4,05	77,12 ± 2,58
		10	97,53 ± 2,65	95,60 ± 2,17	93,12 ± 1,73
		20	98,74 ± 1,71	97,69 ± 1,44	96,17 ± 1,48
	4	5	95,20 ± 3,58	91,05 ± 3,35	84,42 ± 2,85
		10	98,67 ± 1,70	97,29 ± 2,08	95,60 ± 1,56
		20	99,05 ± 1,59	98,46 ± 1,49	97,64 ± 1,39

Ek Tablo 11. DSA deneylerinde gizli katman sayısı 4 seçildiğinde performanslar

EVU	TVU	#MFKK	Konuşmacı Sayısı		
			10	20	40
5	1	5	75,28 ± 5,90	62,39 ± 3,68	47,94 ± 2,89
		10	86,86 ± 5,43	79,88 ± 4,08	69,80 ± 3,19
		20	88,66 ± 5,03	84,46 ± 4,08	78,96 ± 2,78
	2	5	82,80 ± 5,47	72,79 ± 4,72	60,74 ± 3,32
		10	90,34 ± 5,37	86,65 ± 3,88	80,11 ± 3,21
		20	92,52 ± 4,84	88,98 ± 3,64	86,58 ± 2,79
	4	5	87,70 ± 5,02	81,43 ± 4,39	71,39 ± 3,47
		10	93,10 ± 5,25	90,92 ± 3,55	86,84 ± 2,86
		20	94,45 ± 4,48	91,83 ± 3,49	90,95 ± 2,38
10	1	5	81,09 ± 7,00	70,12 ± 4,26	57,49 ± 2,43
		10	92,05 ± 3,82	86,02 ± 3,27	79,99 ± 2,87
		20	93,95 ± 3,38	91,26 ± 3,08	87,07 ± 2,30
	2	5	87,44 ± 5,94	80,79 ± 3,58	69,90 ± 3,21
		10	95,35 ± 2,96	91,49 ± 3,04	88,20 ± 2,61
		20	95,49 ± 3,55	94,32 ± 2,92	92,37 ± 2,19
	4	5	92,29 ± 4,59	87,57 ± 3,56	79,70 ± 2,91
		10	96,71 ± 2,96	94,46 ± 2,69	92,72 ± 2,24
		20	96,75 ± 3,24	96,36 ± 2,84	94,96 ± 1,97
20	1	5	86,37 ± 5,33	76,84 ± 3,56	65,28 ± 2,60
		10	95,02 ± 3,14	92,11 ± 2,48	88,49 ± 1,65
		20	97,71 ± 2,17	95,76 ± 2,00	93,60 ± 1,47
	2	5	91,84 ± 4,59	86,00 ± 3,42	77,31 ± 2,25
		10	97,11 ± 2,81	95,51 ± 2,16	93,58 ± 1,58
		20	98,71 ± 1,63	97,54 ± 1,64	96,34 ± 1,40
	4	5	95,11 ± 3,65	91,75 ± 3,08	85,83 ± 2,38
		10	98,54 ± 2,03	97,39 ± 1,72	96,29 ± 1,27
		20	99,05 ± 1,42	98,57 ± 1,30	97,77 ± 1,30

Ek Tablo 12. DSA-TÖ deneylerinde gizli katman sayısı 2 seçildiğinde performanslar

EVU	TVU	#MFKK	Konuşmacı Sayısı		
			10	20	40
5	1	5	77,37 ± 6,89	65,24 ± 4,85	51,48 ± 2,80
		10	89,21 ± 5,29	83,56 ± 3,84	75,38 ± 2,90
		20	91,64 ± 5,62	88,36 ± 3,24	83,57 ± 3,33
	2	5	85,32 ± 6,68	74,33 ± 4,72	62,80 ± 3,42
		10	92,71 ± 5,58	88,93 ± 4,02	82,93 ± 2,79
		20	93,74 ± 4,87	91,47 ± 3,13	88,52 ± 3,21
	4	5	89,08 ± 5,94	81,15 ± 5,29	72,12 ± 4,15
		10	94,22 ± 4,80	92,26 ± 3,62	88,20 ± 2,81
		20	95,60 ± 4,50	94,10 ± 2,64	91,55 ± 2,76
10	1	5	81,95 ± 5,54	71,76 ± 4,85	58,06 ± 2,57
		10	91,73 ± 5,36	88,59 ± 3,51	81,76 ± 2,72
		20	94,29 ± 4,23	92,88 ± 3,07	88,53 ± 2,64
	2	5	88,74 ± 6,14	81,01 ± 4,65	69,30 ± 2,87
		10	94,30 ± 4,49	92,55 ± 3,04	88,10 ± 2,64
		20	95,67 ± 4,09	94,95 ± 2,51	92,23 ± 2,31
	4	5	91,95 ± 5,31	86,62 ± 3,82	77,29 ± 3,20
		10	95,66 ± 4,14	95,02 ± 2,78	92,18 ± 2,41
		20	96,53 ± 4,07	96,59 ± 2,30	94,52 ± 2,28
20	1	5	86,06 ± 4,61	76,56 ± 3,77	62,85 ± 2,55
		10	94,98 ± 3,14	91,87 ± 2,77	85,62 ± 2,14
		20	96,80 ± 2,61	95,55 ± 2,13	92,09 ± 1,97
	2	5	91,45 ± 4,17	84,98 ± 4,05	74,14 ± 2,81
		10	96,24 ± 3,20	94,90 ± 2,38	91,35 ± 1,88
		20	97,26 ± 2,47	96,89 ± 1,92	94,84 ± 1,66
	4	5	94,57 ± 3,70	90,33 ± 3,48	81,50 ± 3,06
		10	97,67 ± 2,55	96,89 ± 1,93	94,51 ± 1,49
		20	98,19 ± 2,19	98,02 ± 1,71	96,60 ± 1,46

Ek Tablo 13. DSA-TÖ deneylerinde gizli katman sayısı 4 seçildiğinde performanslar

EVU	TVU	#MFKK	Konuşmacı Sayısı		
			10	20	40
5	1	5	78,21 ± 6,30	64,73 ± 4,95	51,12 ± 2,75
		10	89,64 ± 5,59	83,59 ± 4,23	75,21 ± 2,92
		20	92,07 ± 5,49	88,48 ± 3,34	83,55 ± 3,60
	2	5	84,70 ± 6,53	74,19 ± 4,84	62,00 ± 3,43
		10	92,14 ± 5,52	88,62 ± 4,07	83,01 ± 2,83
		20	93,86 ± 5,07	91,85 ± 3,31	88,46 ± 3,35
	4	5	88,99 ± 6,22	81,33 ± 4,69	71,38 ± 3,80
		10	94,22 ± 5,04	92,68 ± 3,52	87,51 ± 2,98
		20	95,11 ± 4,70	93,69 ± 2,99	91,33 ± 3,18
10	1	5	82,04 ± 6,08	71,10 ± 4,62	57,81 ± 2,47
		10	91,35 ± 5,26	88,39 ± 3,58	81,14 ± 2,58
		20	93,97 ± 4,69	92,28 ± 3,48	88,17 ± 2,60
	2	5	88,62 ± 5,82	80,29 ± 4,79	68,91 ± 2,49
		10	94,13 ± 4,83	92,45 ± 3,01	87,77 ± 2,49
		20	95,41 ± 4,59	94,95 ± 2,47	91,75 ± 2,54
	4	5	92,41 ± 4,80	86,10 ± 4,45	77,59 ± 3,22
		10	95,81 ± 3,76	94,98 ± 2,86	91,84 ± 2,47
		20	96,47 ± 3,78	96,62 ± 2,28	94,66 ± 2,12
20	1	5	86,25 ± 4,78	76,57 ± 3,70	62,73 ± 2,64
		10	94,68 ± 3,62	91,12 ± 2,82	85,13 ± 2,09
		20	96,73 ± 2,66	95,34 ± 2,04	91,54 ± 2,20
	2	5	90,96 ± 4,77	85,05 ± 3,74	73,38 ± 3,03
		10	96,45 ± 2,82	94,46 ± 2,31	90,66 ± 1,92
		20	97,40 ± 2,58	97,11 ± 2,11	94,67 ± 1,95
	4	5	94,68 ± 3,60	90,60 ± 3,25	81,19 ± 3,04
		10	97,73 ± 2,31	96,59 ± 1,90	94,30 ± 1,51
		20	98,22 ± 2,10	97,87 ± 1,90	96,64 ± 1,51

Ek Tablo 14. ZGSA deneylerinde “a” segmenti seçildiğinde performanslar

EVU	TVU	#MFKK	Konuşmacı Sayısı		
			10	20	40
5	1	5	44,12 ± 6,48	33,35 ± 4,65	21,70 ± 2,32
		10	64,82 ± 6,62	53,26 ± 5,73	41,12 ± 2,97
		20	73,62 ± 6,77	64,86 ± 5,19	54,77 ± 3,35
	2	5	57,41 ± 8,72	47,44 ± 6,00	33,44 ± 2,45
		10	77,74 ± 7,22	70,37 ± 6,46	60,52 ± 3,68
		20	85,30 ± 5,75	79,82 ± 5,09	72,58 ± 3,30
	4	5	65,88 ± 9,66	56,75 ± 6,26	42,35 ± 3,30
		10	83,40 ± 6,71	78,43 ± 5,50	70,25 ± 3,57
		20	89,61 ± 5,79	85,24 ± 4,73	79,33 ± 3,72
10	1	5	49,57 ± 7,92	37,45 ± 4,22	24,36 ± 2,39
		10	67,96 ± 7,30	58,44 ± 4,54	46,46 ± 2,84
		20	77,90 ± 5,85	71,05 ± 3,74	60,37 ± 2,42
	2	5	65,03 ± 8,13	54,20 ± 5,22	40,21 ± 3,34
		10	82,68 ± 6,28	76,57 ± 4,42	67,85 ± 2,79
		20	89,07 ± 5,68	85,50 ± 4,56	78,99 ± 2,54
	4	5	73,54 ± 8,81	64,34 ± 5,81	51,30 ± 3,77
		10	88,09 ± 5,81	84,51 ± 4,58	78,05 ± 3,01
		20	92,43 ± 5,16	90,01 ± 3,89	85,36 ± 2,68
20	1	5	52,11 ± 8,18	39,68 ± 3,70	26,84 ± 2,17
		10	72,56 ± 6,00	62,91 ± 4,61	49,49 ± 2,80
		20	81,78 ± 5,81	74,32 ± 3,64	64,95 ± 2,88
	2	5	68,94 ± 8,39	58,58 ± 4,71	45,06 ± 2,94
		10	86,02 ± 5,65	82,05 ± 4,01	72,18 ± 3,07
		20	92,63 ± 4,41	88,63 ± 3,44	84,17 ± 2,97
	4	5	77,96 ± 8,01	70,54 ± 4,99	58,22 ± 3,31
		10	91,26 ± 4,89	89,22 ± 3,96	82,56 ± 3,12
		20	95,31 ± 3,95	92,18 ± 3,44	89,99 ± 2,88

Ek Tablo 15. ZGSA deneylerinde “b” segmenti seçildiğinde performanslar

EVU	TVU	#MFKK	Konuşmacı Sayısı		
			10	20	40
5	1	5	48,48 ± 6,89	32,85 ± 4,05	23,03 ± 2,20
		10	65,87 ± 6,43	53,49 ± 4,69	41,76 ± 2,89
		20	74,48 ± 5,64	63,94 ± 5,38	52,78 ± 3,36
	2	5	62,30 ± 8,49	48,61 ± 5,89	35,90 ± 3,40
		10	79,93 ± 7,75	70,18 ± 5,47	61,08 ± 3,19
		20	85,86 ± 5,82	79,48 ± 5,21	71,06 ± 3,93
	4	5	69,68 ± 9,54	56,23 ± 7,55	44,89 ± 4,06
		10	85,14 ± 7,44	78,18 ± 5,53	70,12 ± 3,59
		20	89,87 ± 5,30	85,65 ± 5,05	78,49 ± 4,01
10	1	5	52,38 ± 6,85	37,27 ± 3,81	25,55 ± 2,32
		10	70,84 ± 6,78	58,38 ± 4,77	46,94 ± 2,77
		20	78,04 ± 5,76	68,58 ± 4,93	58,98 ± 3,81
	2	5	69,25 ± 7,48	54,59 ± 5,93	41,84 ± 3,33
		10	84,54 ± 6,45	76,72 ± 5,35	67,85 ± 2,87
		20	89,85 ± 4,73	83,45 ± 4,66	78,14 ± 3,32
	4	5	78,22 ± 6,98	65,05 ± 6,31	53,86 ± 4,30
		10	89,92 ± 6,30	84,05 ± 5,51	77,87 ± 3,39
		20	92,75 ± 4,72	88,99 ± 4,43	84,73 ± 3,47
20	1	5	55,83 ± 6,52	39,07 ± 3,81	28,04 ± 2,64
		10	73,75 ± 6,16	62,40 ± 4,66	51,10 ± 2,86
		20	81,82 ± 4,28	73,72 ± 4,33	63,69 ± 2,68
	2	5	72,80 ± 6,64	59,62 ± 4,86	47,31 ± 3,31
		10	87,81 ± 5,51	81,07 ± 4,58	73,51 ± 3,08
		20	91,81 ± 4,14	89,13 ± 3,28	83,08 ± 2,86
	4	5	82,57 ± 6,42	71,56 ± 5,81	61,05 ± 3,46
		10	92,81 ± 5,14	88,92 ± 4,56	83,86 ± 2,92
		20	95,00 ± 4,22	93,66 ± 3,19	89,43 ± 2,39

Ek Tablo 16. MÇSA deneylerinde performanslar

EVU	TVU	#MFKK	Konuşmacı Sayısı		
			10	20	40
5	1	5	65,43 ± 8,09	54,50 ± 4,94	42,59 ± 3,15
		10	75,23 ± 6,86	64,37 ± 5,22	53,54 ± 3,56
		20	72,12 ± 6,77	62,67 ± 4,62	51,58 ± 3,48
	2	5	75,68 ± 8,26	67,09 ± 5,76	56,03 ± 4,04
		10	83,57 ± 6,56	74,99 ± 5,03	65,85 ± 3,35
		20	79,98 ± 6,57	73,01 ± 5,73	64,72 ± 3,67
	4	5	81,76 ± 7,31	74,65 ± 6,19	65,52 ± 3,93
		10	87,66 ± 6,42	81,36 ± 4,55	73,69 ± 3,71
		20	84,91 ± 6,45	79,01 ± 5,30	72,52 ± 3,49
10	1	5	70,20 ± 6,46	57,91 ± 5,05	45,90 ± 3,60
		10	78,49 ± 6,82	68,11 ± 4,97	57,46 ± 3,28
		20	75,79 ± 6,79	66,54 ± 4,53	55,25 ± 3,03
	2	5	79,83 ± 6,83	71,39 ± 5,32	61,24 ± 3,37
		10	86,96 ± 7,15	79,13 ± 5,42	71,08 ± 2,85
		20	84,64 ± 7,27	77,43 ± 4,49	69,84 ± 3,43
	4	5	86,96 ± 5,76	80,00 ± 5,46	71,96 ± 3,80
		10	90,34 ± 6,52	85,59 ± 4,99	79,40 ± 3,10
		20	89,19 ± 6,62	84,35 ± 4,28	77,44 ± 2,83
20	1	5	72,04 ± 6,70	61,93 ± 5,37	49,23 ± 2,79
		10	80,43 ± 6,22	70,47 ± 4,77	60,70 ± 3,06
		20	80,01 ± 6,52	69,52 ± 4,71	60,08 ± 3,13
	2	5	83,44 ± 6,07	75,37 ± 4,30	65,63 ± 3,36
		10	88,34 ± 5,31	81,45 ± 4,44	74,94 ± 2,93
		20	88,27 ± 5,67	80,93 ± 4,03	74,28 ± 3,55
	4	5	89,45 ± 5,27	84,07 ± 4,48	77,15 ± 3,12
		10	92,54 ± 5,23	88,31 ± 4,13	84,12 ± 2,78
		20	92,15 ± 5,77	87,50 ± 4,06	83,26 ± 2,83

EK-3: D2 için Ek Bulgular

Aşağıda, "D2: Öznitelik Spike Dönüşümü Yaklaşımının Karşılaştırmalı İncelenmesi" başlığı altında detayları verilen deneysel senaryoya dair detaylı bulgular sunulmuştur. Bulgular, çeşitli parametre ve yöntemlere göre organize edilmiştir. Tablolarda verilen bulgular, ayrı ayrı belirlenen sınıf bazlı performansların "Ortalama $\pm$ Standart Sapma" formatında verilmiş değerleridir ve F1-Skor ölçütüyle paylaşılmıştır.

Ek Tablo 17. Makine öğrenimi yöntemlerinin kodlama yöntemine göre performansları

Veri Kümesi	Yöntem	Kodlama Yöntemi			
		Analog	Yoğunluk	Gray	İkili
IRIS	K-EYK	93,9 $\pm$ 4,4	96,3 $\pm$ 2,6	96,5 $\pm$ 1,7	86,7 $\pm$ 8,4
	DVM	98,8 $\pm$ 0,9	93,9 $\pm$ 4,3	95,1 $\pm$ 3,5	96,3 $\pm$ 2,6
	TKP	90,2 $\pm$ 7,0	96,4 $\pm$ 2,6	97,6 $\pm$ 1,7	83,0 $\pm$ 12,1
	NB	95,2 $\pm$ 3,4	92,7 $\pm$ 5,2	82,3 $\pm$ 13,2	74,6 $\pm$ 18,2
	LAA	98,8 $\pm$ 0,9	85,6 $\pm$ 9,3	95,1 $\pm$ 3,5	95,1 $\pm$ 3,5
MNIST	K-EYK	97,0 $\pm$ 2,4	96,3 $\pm$ 2,6	84,4 $\pm$ 11,7	92,8 $\pm$ 5,3
	DVM	96,5 $\pm$ 3,1	96,9 $\pm$ 2,1	92,3 $\pm$ 3,7	95,3 $\pm$ 2,9
	TKP	93,4 $\pm$ 3,3	95,5 $\pm$ 2,5	88,3 $\pm$ 5,5	92,7 $\pm$ 4,1
	NB	79,7 $\pm$ 15,9	78,6 $\pm$ 14,5	77,2 $\pm$ 12,5	73,2 $\pm$ 21,2
	LAA	93,9 $\pm$ 3,6	61,7 $\pm$ 8,8	87,9 $\pm$ 4,6	91,2 $\pm$ 4,9
VK6	K-EYK	90,7 $\pm$ 3,5	89,0 $\pm$ 4,3	78,3 $\pm$ 5,3	63,1 $\pm$ 6,2
	DVM	88,1 $\pm$ 3,9	91,8 $\pm$ 3,3	77,7 $\pm$ 4,0	80,3 $\pm$ 5,1
	TKP	89,3 $\pm$ 3,3	87,4 $\pm$ 3,3	77,6 $\pm$ 4,3	82,3 $\pm$ 4,1
	NB	78,4 $\pm$ 5,2	31,5 $\pm$ 19,2	63,3 $\pm$ 4,5	64,3 $\pm$ 8,3
	LAA	89,6 $\pm$ 2,5	41,8 $\pm$ 4,7	79,4 $\pm$ 4,2	85,2 $\pm$ 2,9

Ek Tablo 18. Makine öğrenimi yöntemlerinin kuantalama tekniğine göre performansları

Veri Kümesi	Yöntem	Kuantalama Tekniği			
		Analog	Tekdüze	Çeyreklik	K-Ortalama
IRIS	K-EYK	93,9 $\pm$ 4,4	96,3 $\pm$ 2,6	93,9 $\pm$ 4,3	95,1 $\pm$ 3,5
	DVM	98,8 $\pm$ 0,9	93,9 $\pm$ 4,3	95,1 $\pm$ 3,5	93,9 $\pm$ 4,3
	TKP	90,2 $\pm$ 7,0	96,4 $\pm$ 2,6	97,6 $\pm$ 1,7	96,4 $\pm$ 2,6
	NB	95,1 $\pm$ 3,4	92,7 $\pm$ 5,1	92,7 $\pm$ 5,1	92,7 $\pm$ 5,1
	LAA	98,8 $\pm$ 0,9	85,6 $\pm$ 9,3	86,8 $\pm$ 4,4	86,2 $\pm$ 6,9
VK6	K-EYK	90,7 $\pm$ 3,5	89,0 $\pm$ 4,3	89,5 $\pm$ 3,9	89,5 $\pm$ 4,7
	DVM	88,1 $\pm$ 3,9	91,8 $\pm$ 3,3	90,6 $\pm$ 2,8	91,5 $\pm$ 2,7
	TKP	89,3 $\pm$ 3,3	87,4 $\pm$ 3,3	86,6 $\pm$ 3,2	87,8 $\pm$ 2,6
	NB	78,4 $\pm$ 5,2	31,5 $\pm$ 19,2	52,4 $\pm$ 20,7	40,5 $\pm$ 19,9
	LAA	89,6 $\pm$ 2,5	41,8 $\pm$ 4,7	50,0 $\pm$ 5,4	48,1 $\pm$ 7,9

Ek Tablo 19. Makine öğrenimi yöntemlerinin bit uzunluğuna göre performansları

Veri Kümesi	Yöntem	Bit Uzunluğu				
		Analog	4	5	6	7
IRIS	K-EYK	93,9 $\pm$ 4,4	95,1 $\pm$ 3,5	96,3 $\pm$ 2,6	93,9 $\pm$ 4,3	96,3 $\pm$ 2,6
	DVM	98,8 $\pm$ 0,9	95,1 $\pm$ 3,5	93,9 $\pm$ 4,3	92,7 $\pm$ 5,1	92,7 $\pm$ 5,1
	TKP	90,2 $\pm$ 7,0	95,1 $\pm$ 3,5	96,4 $\pm$ 2,6	96,4 $\pm$ 2,6	96,4 $\pm$ 2,6
	NB	95,1 $\pm$ 3,4	95,1 $\pm$ 3,4	92,7 $\pm$ 5,1	92,7 $\pm$ 5,1	92,7 $\pm$ 5,1
	LAA	98,8 $\pm$ 0,9	74,5 $\pm$ 8,1	85,6 $\pm$ 9,3	90,7 $\pm$ 4,7	89,7 $\pm$ 4,6
VK6	K-EYK	90,7 $\pm$ 3,5	90,1 $\pm$ 3,6	89,0 $\pm$ 4,3	89,6 $\pm$ 4,0	89,5 $\pm$ 4,1
	DVM	88,1 $\pm$ 3,9	90,9 $\pm$ 3,1	91,8 $\pm$ 3,3	90,9 $\pm$ 3,0	91,5 $\pm$ 2,9
	TKP	89,3 $\pm$ 3,3	86,5 $\pm$ 3,8	87,4 $\pm$ 3,3	87,8 $\pm$ 2,9	87,6 $\pm$ 2,7
	NB	78,4 $\pm$ 5,2	26,9 $\pm$ 19,4	31,5 $\pm$ 19,2	36,7 $\pm$ 17,9	36,8 $\pm$ 15,9
	LAA	89,6 $\pm$ 2,5	64,5 $\pm$ 8,0	41,8 $\pm$ 4,7	74,8 $\pm$ 4,2	79,8 $\pm$ 3,7

Ek Tablo 20. Gerçekçi nöron modellerinin kodlama yöntemine göre performansları

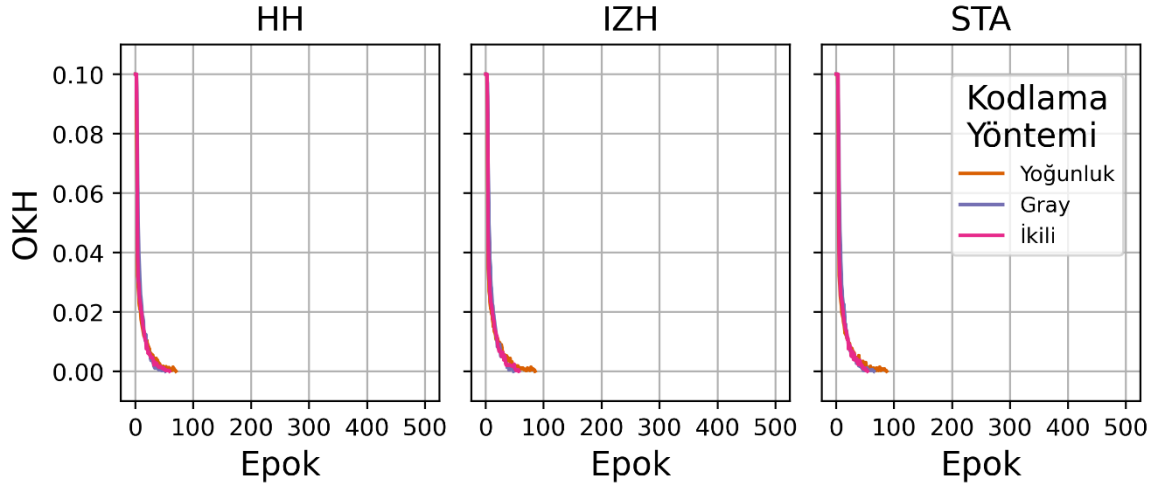
Veri Kümesi	Yöntem	Kodlama Yöntemi		
		Yoğunluk	Gray	İkili
IRIS	HH	96,4 $\pm$ 2,6	97,6 $\pm$ 1,7	79,6 $\pm$ 8,9
	IZH	96,4 $\pm$ 2,6	97,6 $\pm$ 1,7	79,6 $\pm$ 8,9
	STA	96,4 $\pm$ 2,6	97,6 $\pm$ 1,7	81,8 $\pm$ 8,2
MNIST	HH	94,0 $\pm$ 3,4	89,2 $\pm$ 5,6	91,6 $\pm$ 4,7
	IZH	93,7 $\pm$ 3,4	88,7 $\pm$ 5,3	91,7 $\pm$ 4,5
	STA	94,0 $\pm$ 3,1	88,6 $\pm$ 6,1	91,9 $\pm$ 4,2
VK6	HH	85,4 $\pm$ 4,3	41,8 $\pm$ 18,2	40,6 $\pm$ 20,3
	IZH	85,0 $\pm$ 3,9	39,1 $\pm$ 25,8	44,3 $\pm$ 21,7
	STA	85,2 $\pm$ 3,3	40,8 $\pm$ 22,8	51,8 $\pm$ 19,4

Ek Tablo 21. Gerçekçi nöron modellerinin kuantalama tekniğine göre performansları

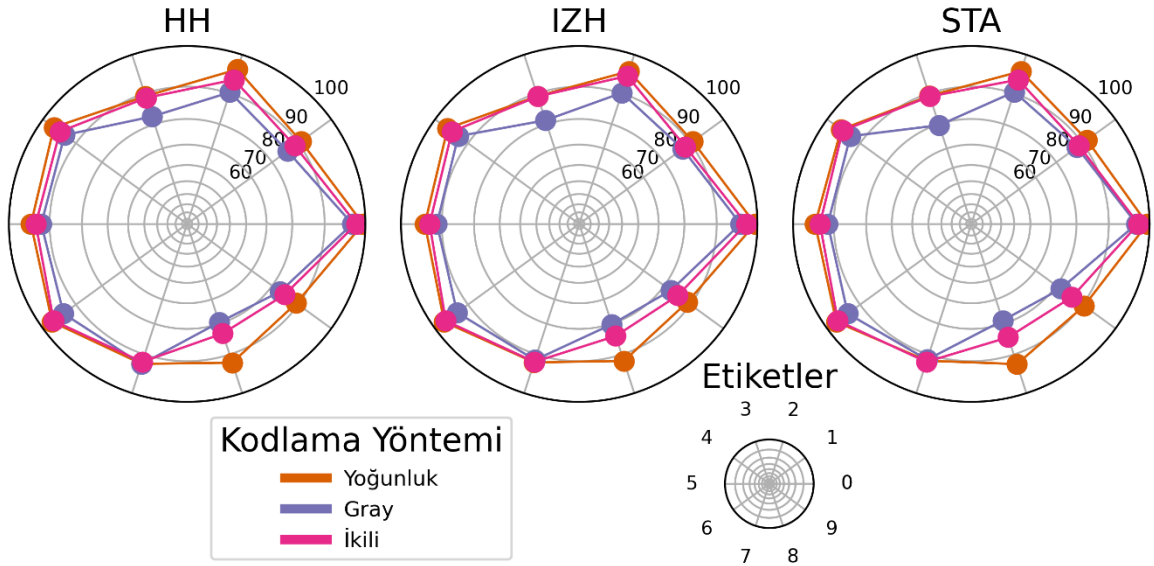
Veri Kümesi	Yöntem	Kuantalama Tekniği		
		Tekdüze	Çeyreklik	K-Ortalama
IRIS	HH	96,4 $\pm$ 2,6	96,4 $\pm$ 2,6	96,4 $\pm$ 2,6
	IZH	96,4 $\pm$ 2,6	96,4 $\pm$ 2,6	96,4 $\pm$ 2,6
	STA	96,4 $\pm$ 2,6	96,4 $\pm$ 2,6	96,4 $\pm$ 2,6
VK6	HH	85,4 $\pm$ 4,3	83,3 $\pm$ 5,4	83,5 $\pm$ 4,1
	IZH	85,0 $\pm$ 3,9	82,7 $\pm$ 5,7	82,5 $\pm$ 3,1
	STA	85,2 $\pm$ 3,3	83,1 $\pm$ 5,4	83,5 $\pm$ 4,9

Ek Tablo 22. Gerçekçi nöron modellerinin bit uzunluğuna göre performansları

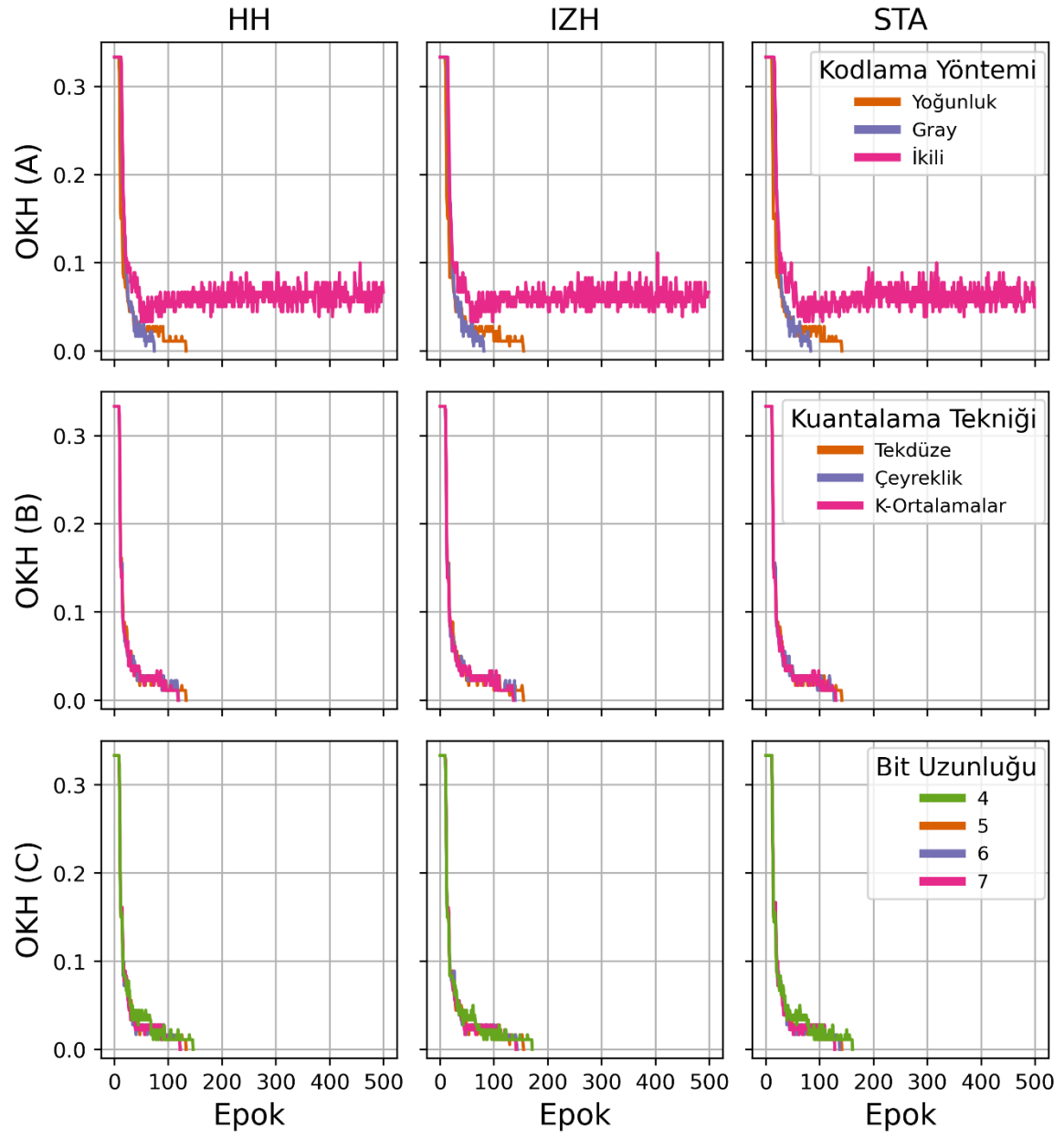
Veri Kümesi	Yöntem	Bit Uzunluğu			
		4	5	6	7
IRIS	HH	96,4 $\pm$ 2,6	96,4 $\pm$ 2,6	96,4 $\pm$ 2,6	96,4 $\pm$ 2,6
	IZH	96,4 $\pm$ 2,6	96,4 $\pm$ 2,6	96,4 $\pm$ 2,6	96,4 $\pm$ 2,6
	STA	95,1 $\pm$ 3,4	96,4 $\pm$ 2,6	96,4 $\pm$ 2,6	96,4 $\pm$ 2,6
VK6	HH	82,8 $\pm$ 4,6	85,4 $\pm$ 4,3	85,4 $\pm$ 3,7	85,2 $\pm$ 3,9
	IZH	83,4 $\pm$ 5,7	85,0 $\pm$ 3,9	85,7 $\pm$ 4,2	85,0 $\pm$ 3,7
	STA	81,8 $\pm$ 4,4	85,2 $\pm$ 3,3	86,1 $\pm$ 4,5	85,5 $\pm$ 4,6



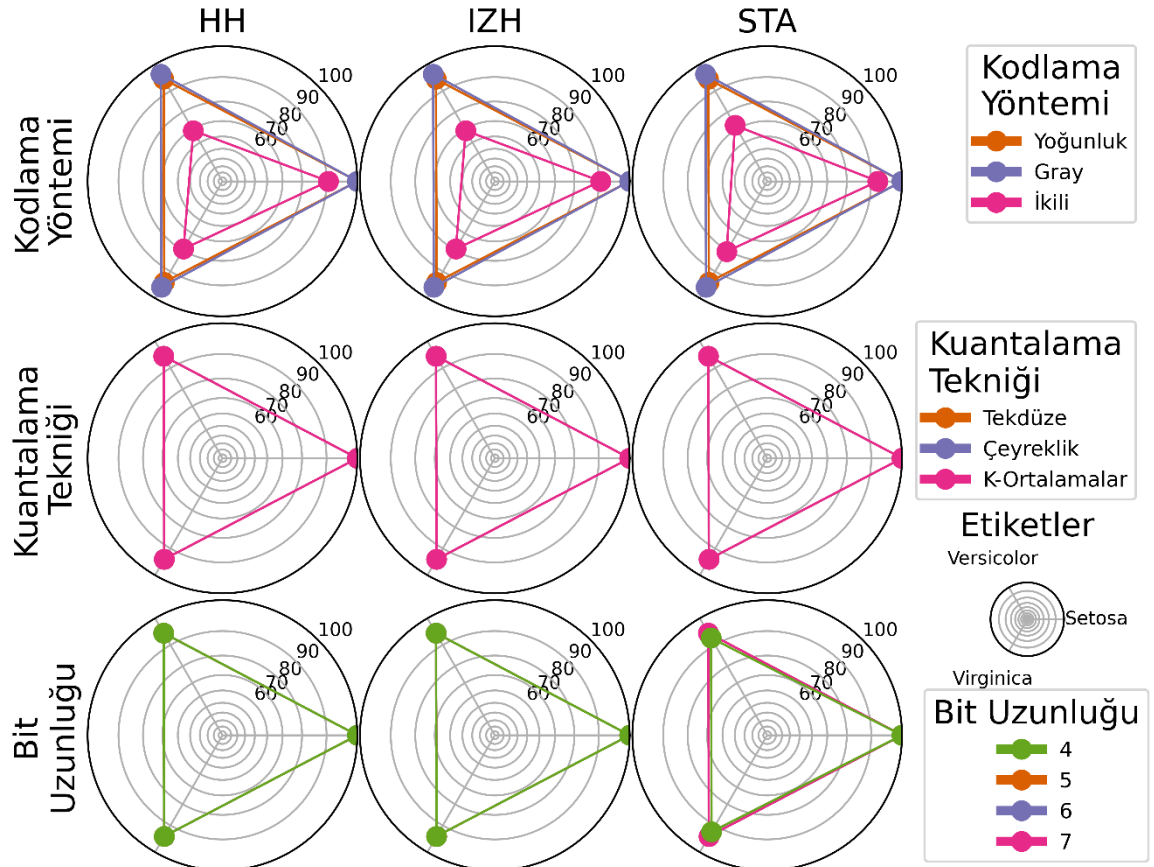
Ek Şekil 1. IZH nöronlarının MNIST veri kümesinde eğitim hatalarının seyri



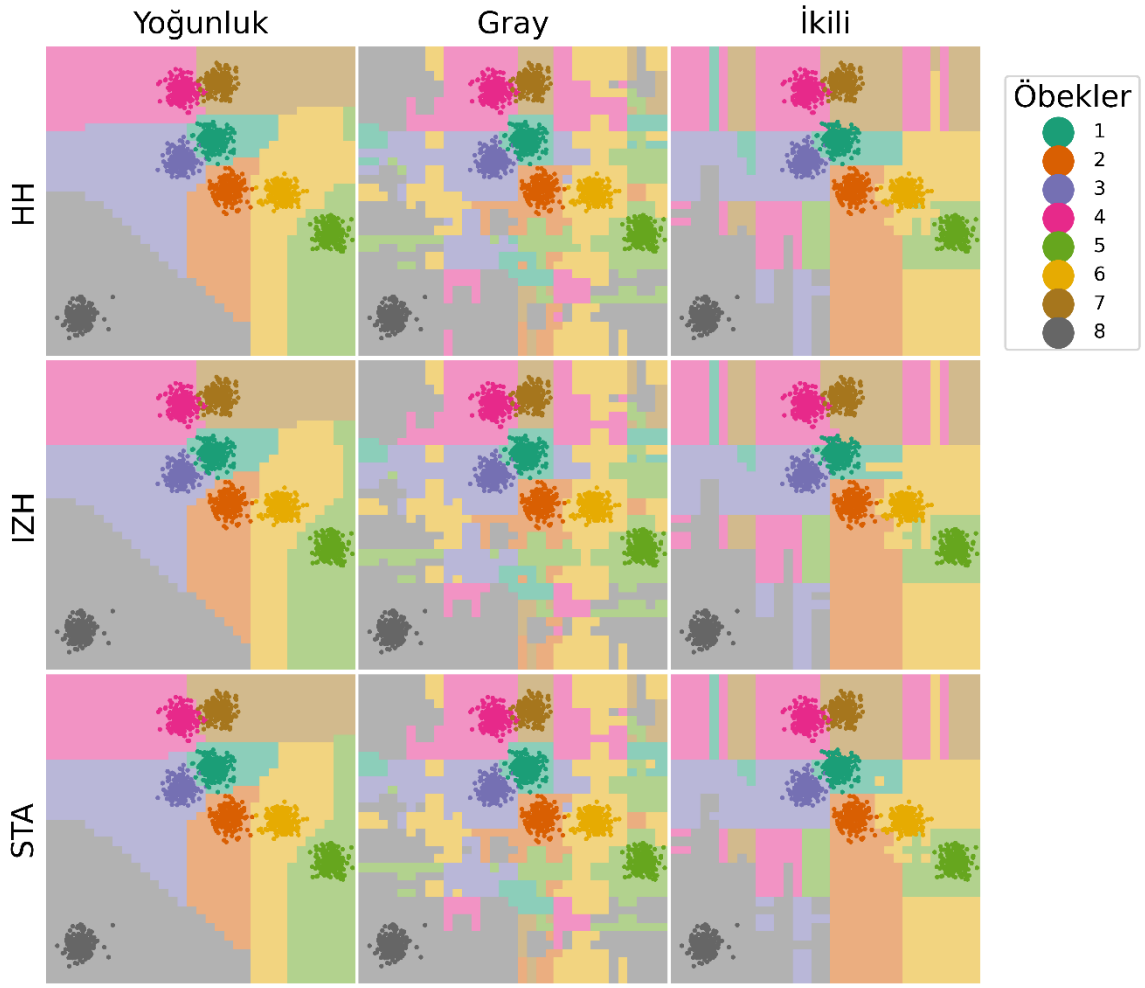
Ek Şekil 2. IZH nöronlarının MNIST veri kümesinde performansı



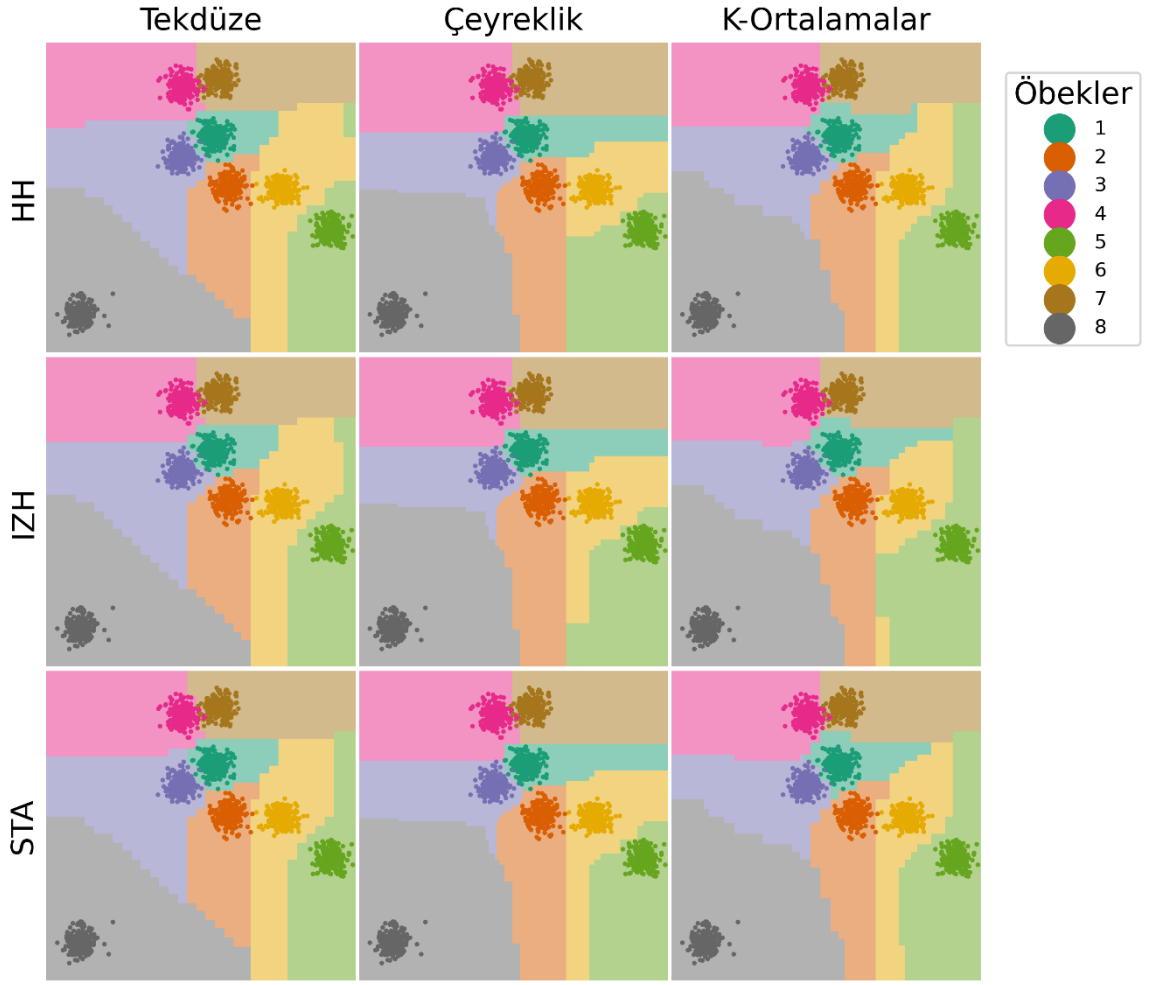
Ek Şekil 3. IZH nöronlarının IRIS veri kümesinde eğitim hatalarının seyri



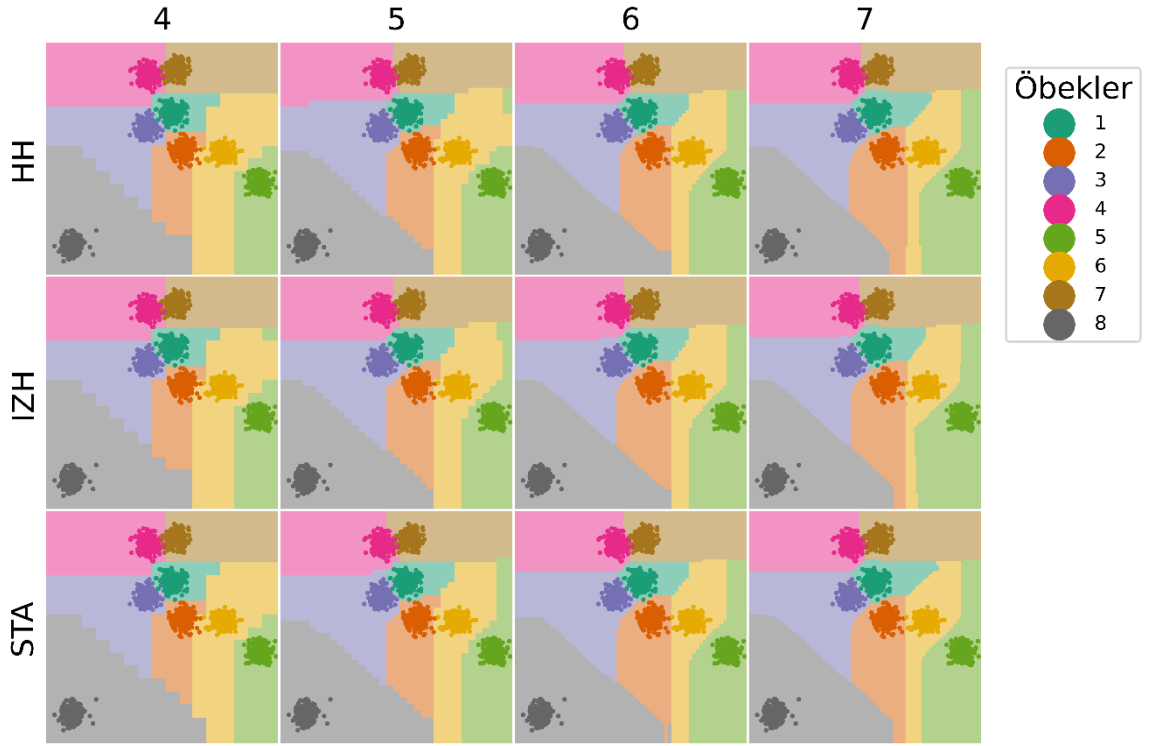
Ek Şekil 4. IZH nöronlarının IRIS veri kümesinde performansı



Ek Şekil 5. Gerçekçi nöron modellerinin kodlama yöntemine göre öbek alanları



Ek Şekil 6. Gerçekçi nöron modellerinin kuantalama tekniğine göre öbek alanları



Ek Şekil 7. Gerçekçi nöron modellerinin bit uzunluğuna göre öbek alanları

EK-4: D3 için Ek Bulgular

Aşağıda, "D3: Sinaptik Bağlantı Tiplerinin Sınıflandırma Performansına Etkisinin İncelenmesi" başlığı altında detayları verilen deneysel senaryoya dair bulgular sunulmuştur. Bulgular, çeşitli parametre ve yöntemlere göre organize edilmiştir. Tablolarda verilen bulgular, ayrı ayrı belirlenen sınıf bazlı performansların "Ortalama $\pm$ Standart Sapma" formatında verilmiş değerleridir ve F1-Skor ölçütüyle paylaşılmıştır.

Ek Tablo 23. Gerçekçi sinir ağlarında sinaptik bağlantı tiplerinin performansa etkisi

Veri Kümesi	Ayna Bağlantı	Exc Bağlantı	Inh Bağlantı	HH	IZH	STA
IRIS	Yok	Yok	Var	14,9 $\pm$ 21,1	14,9 $\pm$ 21,1	14,9 $\pm$ 21,1
		Var	Yok	85,7 $\pm$ 8,7	85,7 $\pm$ 8,7	91,8 $\pm$ 3,9
			Var	92,9 $\pm$ 3,4	92,9 $\pm$ 3,4	92,9 $\pm$ 3,4
	Var	Yok	Var	14,9 $\pm$ 21,1	14,9 $\pm$ 21,1	14,9 $\pm$ 21,1
		Var	Yok	96,4 $\pm$ 2,6	96,4 $\pm$ 2,6	96,4 $\pm$ 2,6
			Var	96,4 $\pm$ 2,6	96,4 $\pm$ 2,6	96,4 $\pm$ 2,6
MNIST	Yok	Yok	Var	1,6 $\pm$ 4,9	1,6 $\pm$ 4,9	1,6 $\pm$ 4,9
		Var	Yok	93,1 $\pm$ 4,2	93,2 $\pm$ 4,2	93,3 $\pm$ 3,9
			Var	93,6 $\pm$ 3,4	93,6 $\pm$ 3,4	93,4 $\pm$ 3,2
	Var	Yok	Var	1,6 $\pm$ 4,9	1,6 $\pm$ 4,9	1,6 $\pm$ 4,9
		Var	Yok	93,9 $\pm$ 3,4	94,1 $\pm$ 3,2	94,2 $\pm$ 3,3
			Var	94,0 $\pm$ 3,4	93,7 $\pm$ 3,4	94,0 $\pm$ 3,1
VK6	Yok	Yok	Var	4,8 $\pm$ 10,6	4,8 $\pm$ 10,6	4,8 $\pm$ 10,6
		Var	Yok	82,7 $\pm$ 4,9	82,8 $\pm$ 4,3	83,1 $\pm$ 4,0
			Var	83,6 $\pm$ 4,6	83,3 $\pm$ 4,6	83,0 $\pm$ 4,2
	Var	Yok	Var	4,8 $\pm$ 10,6	4,8 $\pm$ 10,6	4,8 $\pm$ 10,6
		Var	Yok	82,9 $\pm$ 4,9	85,7 $\pm$ 3,7	84,3 $\pm$ 3,1
			Var	85,4 $\pm$ 4,3	85,0 $\pm$ 3,9	85,2 $\pm$ 3,3

EK-5: D4 için Ek Bulgular

Aşağıda, "D4: Önerilen Gerçekçi Sinir Ağı Modelinin Konuşmacı Tanıma Performansının İncelenmesi" başlığı altında detayları verilen deneysel senaryoya dair bulgular sunulmuştur. Tablolarda verilen bulgular, ayrı ayrı belirlenen sınıf bazlı performansların "Ortalama $\pm$ Standart Sapma" formatında verilmiş değerleridir ve F1-Skor ölçütüyle paylaşılmıştır.

Ek Tablo 24. Kullanılan öz niteliklere göre gerçekçi nöronların performansları

Veri Kümesi	Öznitelik	EVU	HH	IZH	STA
VK4	A1	5	89,7 $\pm$ 2,7	87,8 $\pm$ 2,7	86,4 $\pm$ 5,4
	A2	5	86,1 $\pm$ 2,5	88,6 $\pm$ 4	88,4 $\pm$ 5,7
	B16	5	91,3 $\pm$ 2,3	90,3 $\pm$ 1,8	91,3 $\pm$ 2,1
	B32	5	90,8 $\pm$ 2,6	90,2 $\pm$ 1,8	93,9 $\pm$ 1,5
	MFKK	5	81,5 $\pm$ 1,8	83,7 $\pm$ 2,1	83,6 $\pm$ 1,6
VK8	A1	5	78,4 $\pm$ 9,4	79,2 $\pm$ 9,7	78,2 $\pm$ 10,1
		10	61,7 $\pm$ 27,5	60,7 $\pm$ 25,8	60,9 $\pm$ 23,6
	A2	5	86,3 $\pm$ 7,1	85 $\pm$ 7,9	85,5 $\pm$ 9,1
		10	89,9 $\pm$ 5,5	88,4 $\pm$ 7,2	89,9 $\pm$ 5,6
	B16	5	88,8 $\pm$ 5,4	89,3 $\pm$ 3,9	90 $\pm$ 3,4
		10	90,6 $\pm$ 5	90,7 $\pm$ 4,3	88,6 $\pm$ 6,1
	B32	5	89,8 $\pm$ 4,9	89 $\pm$ 5,4	88,8 $\pm$ 5
		10	92,5 $\pm$ 4,8	92,6 $\pm$ 5,3	92,5 $\pm$ 5,2
	MFKK	5	80,3 $\pm$ 6,5	79,9 $\pm$ 6,8	81,2 $\pm$ 5,5
		10	71,3 $\pm$ 15,9	78,7 $\pm$ 7,3	63 $\pm$ 19,1
VK16	A1	5	62,4 $\pm$ 11,5	59,1 $\pm$ 17,9	60,7 $\pm$ 13,4
	A2	5	73,9 $\pm$ 8,8	75,2 $\pm$ 8,1	73,9 $\pm$ 9,1
	B16	5	81,5 $\pm$ 8,9	81,7 $\pm$ 8,4	81,7 $\pm$ 8,7
	B32	5	83,9 $\pm$ 8	83,7 $\pm$ 8	83,2 $\pm$ 8,9
	MFKK	5	67,3 $\pm$ 9,5	68 $\pm$ 11,2	64,5 $\pm$ 11,1

EK-6: D5 için Ek Bulgular

Aşağıda, "D5: Piramidal Hücrenin Konuşmacı Tanıma Performansının İncelenmesi" başlığı altında detayları verilen deneysel senaryoya dair bulgular sunulmuştur. Tablolarda verilen bulgular, ayrı ayrı belirlenen sınıf bazlı performansların "Ortalama $\pm$ Standart Sapma" formatında verilmiş değerleridir ve F1-Skor ölçütüyle paylaşılmıştır.

Ek Tablo 25. Piramidal nöronun otomatik konuşmacı tanıma performansı

TVU	Bölge	Bağlantı Sayısı			
		1	2	4	8
1	Sadece Soma	91,9 $\pm$ 3,4	93,3 $\pm$ 2,1	84,9 $\pm$ 2,7	92,8 $\pm$ 1,9
	Sadece Apikal	91,9 $\pm$ 1,8	93,9 $\pm$ 0,8	94,9 $\pm$ 1,9	87,4 $\pm$ 4,3
	Sadece Basilar	91,8 $\pm$ 1,7	94,9 $\pm$ 2,2	97,4 $\pm$ 0,5	90,5 $\pm$ 4,6
	Tüm Dendritler	93,9 $\pm$ 1,5	94,4 $\pm$ 1,0	90,1 $\pm$ 3,0	91,8 $\pm$ 1,4
	Tüm Bölgeler	90,9 $\pm$ 2,9	92,9 $\pm$ 1,5	93,9 $\pm$ 1,3	90,1 $\pm$ 3,4
Çerçeve	Sadece Soma	63,2 $\pm$ 4,3	63,6 $\pm$ 3,8	61,9 $\pm$ 6,2	62,6 $\pm$ 5,3
	Sadece Apikal	62,6 $\pm$ 5,0	62,9 $\pm$ 4,5	63,6 $\pm$ 3,3	62,3 $\pm$ 5,6
	Sadece Basilar	63,3 $\pm$ 4,7	62,7 $\pm$ 4,1	64,2 $\pm$ 3,4	62,7 $\pm$ 4,4
	Tüm Dendritler	62,3 $\pm$ 2,5	63,4 $\pm$ 3,5	62,3 $\pm$ 4,5	62,7 $\pm$ 4,4
	Tüm Bölgeler	62,1 $\pm$ 3,0	63,2 $\pm$ 4,3	62,9 $\pm$ 2,8	62,5 $\pm$ 4,6

ÖZGEÇMİŞ

İlk ve orta öğrenimini Hasan Ali Yücel İlköğretim Okulu, lise öğrenimini Kuleli Askeri Lisesi'nde tamamladı. 2010 yılında Karadeniz Teknik Üniversitesi Mühendislik Fakültesi Elektrik-Elektronik Mühendisliği Bölümü'nde lisans programına başladı ve 2014 yılında bu bölümden mezun oldu. Aynı yıl Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü Elektrik-Elektronik Mühendisliği Anabilim Dalı'nda yüksek lisans öğrenimine başladı ve 2017 yılında bu programdan mezun oldu. Takip eden 2017 yılının güz yarıyılında aynı programda doktora programına başladı. 2015 yılı bahar döneminden itibaren Elektrik-Elektronik Mühendisliği bölümünde araştırma görevlisi olarak çalışmaktadır.

2'si SCI/Expanded indeksli dergilerde, 7'si ulusal ve uluslararası konferanslarda olmak üzere toplam 9 yayını mevcuttur. Yabancı dil olarak İngilizce bilmektedir. Yayınları aşağıda verilmiştir.

Uluslararası Hakemli Dergilerde Yayımlanan Makaleler (SCI/SCI-E)

Kayıkcioglu Bozkir, I., Ozcan, Z., Kose, C., Kayıkcioglu, T., & Cetin, A. E. (2023). Improving a cortical pyramidal neuron model's classification performance on a real-world ecg dataset by extending inputs. *Journal of Computational Neuroscience*, 1-13.

Özcan, Z., Kayıkçıoğlu, T. (2021). Evaluating MFCC-based speaker identification systems with data envelopment analysis. *Expert Systems with Applications*, 168, 114448.

Hakemli Konferans/Sempozyumların Bildiri Kitaplarında Yer Alan Yayınlar

Özcan, Z., Yıldırım, Ö., & Kayıkçıoğlu, T. (2021). Broomycell: An Introductory Educational Software for Neuroscience. *Signal Processing and Communications Applications Conference*, s. 1-4.

Özcan, Z., Kayıkçıoğlu, İ., Yıldırım, Ö., Köse, C., Kayıkçıoğlu, T. (2019). Simulation of morphologically and biophysically realistic single nerve cells using user-friendly software. *Anatomy: International Journal of Experimental & Clinical Anatomy*, 13.

- Yıldırım, Ö., Özcan, Z., Turhan, H., Kayıkçıoğlu, T. (2019). Simulation of biological networks. *Anatomy: International Journal of Experimental & Clinical Anatomy*, 13.
- Şimşek, H., Kayıkçıoğlu, T., Özcan, Z. (2019). Biyolojik Nöronların Farklı Akım Uyarımlarında Dinamiğinin İncelenmesi. 2. *Uluslararası 19 Mayıs Yenilikçi Bilimsel Yaklaşımlar Kongresi*, s. 893-900.
- Yıldırım, Ö., Özcan, Z., Turhan, H., Kayıkçıoğlu, T. (2018). Network Master: A Versatile and User-Friendly Educational Software for Simulation of Neural Networks. *Medical Technologies National Congress*, s. 1-4.
- Özcan, Z., Kayıkçıoğlu, İ., Yıldırım, Ö., Köse, C., Kayıkçıoğlu, T. (2018). Cell Master: A Versatile and User-Friendly Educational Software for Simulation of Neuronal Dynamics. *Medical Technologies National Congress*, s. 1-4.
- Özcan, Z., Kayıkçıoğlu, T. (2017). A speaker identification performance comparison based on the classifier, the computation time and the number of MFCC. *Signal Processing and Communications Applications Conference*, s. 1-4.