



**BULANIK DALGACIK ÇEKİRDEK TABANLI
AŞIRI ÖĞRENME MAKİNELERİ YAKLAŞIMI
İLE SINIFLANDIRMA ANALİZİ**

Doktora Tezi

Aslı KAYA KARAKÜTÜK

Eskişehir 2023

**BULANIK DALGACIK ÇEKİRDEK TABANLI AŞIRI ÖĞRENME
MAKİNELERİ YAKLAŞIMI İLE SINIFLANDIRMA ANALİZİ**

Ash KAYA KARAKÜTÜK

Doktora Tezi

İstatistik Anabilim Dalı

İstatistik Teorisi Bilim Dalı

Danışman: Doç. Dr. Özer ÖZDEMİR

Eskişehir

Eskişehir Teknik Üniversitesi

Lisansüstü Eğitim Enstitüsü

Ağustos 2023

Bu tez çalışması BAP Komisyonu tarafından kabul edilen 21DRP054 no.lu proje kapsamında desteklenmiştir.

JÜRİ VE ENSTİTÜ ONAYI

Aslı KAYA KARAKÜTÜK'ün BULANIK DALGACIK ÇEKİRDEK TABANLI AŞIRI ÖĞRENME MAKİNELERİ YAKLAŞIMI İLE SINIFLANDIRMA ANALİZİ başlıklı çalışması 17/08/2023 tarihinde aşağıdaki jüri tarafından değerlendirilerek “Eskişehir Teknik Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliği”nin ilgili maddeleri uyarınca, İstatistik Anabilim dalında Doktora Tezi olarak kabul edilmiştir.

Jüri Üyeleri

Unvan Adı Soyadı

İmza

Üye	: Doç. Dr. Özer ÖZDEMİR	
Üye	: Doç. Dr. Alper Kürşat UYSAL	
Üye	: Doç. Dr. Şükrü ACITAŞ	
Üye	: Dr. Öğr. Üyesi Gültekin ATALIK	
Üye	: Dr. Öğr. Üyesi Mustafa ÇAVUŞ	

Prof. Dr. Semra KURAMA

Lisansüstü Eğitim Enstitüsü Müdürü

17/08/2023

DANIřMAN ONAYI

Daniřmanlıęını yrttęm Doktora ęrencisi Aslı KAYA KARAKTK, BULANIK DALGACIK EKİRDEK TABANLI AřIRI ęRENME MAKİNELERİ YAKLAřIMI İLE SINIFLANDIRMA ANALİZİ bařlıklı tez alıřmasını tamamlamıřtır. Hazırlamıř olduęu tez tarafımca incelenmiř ve ęrencinin tez savunma sınavına alınması bilimsel ve etik aıdan uygun grlmřtir.

Tez Daniřmanı

Do. Dr. zer ZDEMİR

ÖZET

BULANIK DALGACIK ÇEKİRDEK TABANLI AŞIRI ÖĞRENME MAKİNELERİ YAKLAŞIMI İLE SINIFLANDIRMA ANALİZİ

Aslı KAYA KARAKÜTÜK

İstatistik Anabilim Dalı

İstatistik Teorisi Bilim Dalı

Eskişehir Teknik Üniversitesi, Lisansüstü Eğitim Enstitüsü, Ağustos 2023

Danışman: Doç. Dr. Özer ÖZDEMİR

Sınıflandırma problemleri, günlük hayatta farklı veri tiplerinde sıklıkla karşılaşılan problemlerden biridir. Son zamanlarda bu tarz problemlerin giderilmesi için çeşitli makine öğrenmesi algoritmaları önerilmiştir ve bu algoritmaların kendine özgü bazı avantajları bulunmaktadır. Dalgacıkların zaman-frekans yerelleştirme özellikleri, yapay sinir ağının öğrenme yetenekleri, bulanık çıkarım sisteminin yaklaşık muhakeme özelliği ve aşırı öğrenme makinesi algoritmasının son derece yüksek öğrenme hızında tek geçişli öğrenme ve iyi genelleme performansı gibi avantajları sınıflandırma problemlerinde etkili bir çözüm ortaya koyacak özellikleri sergileyebilmektedir. Bu bağlamda, bu tez çalışması bu avantajları birleştirerek bulanık dalgacık çekirdek tabanlı aşırı öğrenme makinesi (FWKELM) adı verilen yeni bir yapı sunmaktadır. FWKELM'nin ana hedefleri, doğrusal öğrenme parametrelerinin sayısını azaltarak ağ karmaşıklığını önemli ölçüde azaltmak ve kabul edilebilir doğruluk ve genelleme performansları korumaktır. Önerilen yapıda, her bulanık kural bir alt dalgacık sinir ağına karşılık gelmektedir ve farklı genişleme ve ötelemelere sahip dalgacıklardan oluşmaktadır. Çalışma kapsamında, öncelikle bir Mexican Hat dalgacığının kabul edilebilir bir aşırı öğrenme makinesi algoritması (ELM) çekirdeği olduğu gösterilmiş ve daha sonra bir bulanık model ile çekirdek tabanlı aşırı öğrenme makinesine eşdeğerliği kanıtlanmıştır. FWKELM'nin sınıflandırma doğruluğunu değerlendirmek için farklı tip sınıflandırma problemleri üzerinde çeşitli popüler yöntemler ile karşılaştırılmıştır. Simülasyon sonuçları, önerilen yaklaşımın daha robust yapıda olduğu ve dikkate değer verimliliğini göstermiştir.

Anahtar Sözcükler: Bulanık dalgacık, Çekirdek yapısı, Aşırı öğrenme makineleri, Sınıflandırma.

ABSTRACT

CLASSIFICATION ANALYSIS WITH FUZZY WAVELET KERNEL EXTREME LEARNING MACHINES APPROACH

Aslı KAYA KARAKÜTÜK

Department of Statistic

Programme in Statistical Theory

Eskişehir Technical University, Institute of Graduate Programs, August 2023

Supervisor: Doç. Dr. Özer ÖZDEMİR

Classification problems are one of the problems frequently encountered in different data types in daily life. Recently, various machine learning algorithms have been proposed to solve such problems, and these algorithms have some unique advantages. The time-frequency localization properties of wavelets, the learning capabilities of the artificial neural network, the approximate reasoning feature of the fuzzy inference system, and the advantages of the extreme learning machine algorithm such as one-pass learning at an extremely high learning speed and good generalization performance can present an effective solution for classification problems. In this context, this thesis combines these advantages and presents a new structure called fuzzy wavelet kernel-based extreme learning machine (FWKELM). The main goals of FWKELM are to significantly reduce network complexity by reducing the number of linear learning parameters and maintain acceptable accuracy and generalization performances. In the proposed structure, each fuzzy rule corresponds to a sub-wavelet neural network and consists of wavelets with different dilation and translation. Within the scope of the study, firstly, it is shown that a Mexican Hat wavelet is an acceptable extreme learning machine algorithm kernel, and then its equivalence to a kernel extreme learning machine is proven with a fuzzy model. To evaluate the classification accuracy of FWKELM, it is compared with various popular methods on different type classification problems. The simulation results show that the proposed approach is more robust and has remarkable efficiency.

Keywords: Fuzzy wavelet, Kernel, Extreme learning machine, Classification.

TEŞEKKÜR

Bu tez Eskişehir Teknik Üniversitesi Bilimsel Araştırma Projeleri tarafından mali olarak desteklenmiştir. Proje: 21DRP054. Bu tez aynı zamanda Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK), Bilim İnsanı Destek Programları Başkanlığı (BİDEB), 2211-C Yurt İçi Öncelikli Alanlar Doktora Burs Programı tarafından mali olarak desteklenmektedir.

Doktora sürecim boyunca bana yön gösteren, destek ve emeğini esirgemeyen, beni yüreklendiren, öğrencisi olmaktan gurur duyduğum değerli hocam Doç. Dr. Özer Özdemir'e teşekkür ederim. Ayrıca tez yürütme komitesi üyelerim; Doç. Dr. Şükrü Acıtaş ve Doç. Dr. Alper Kürşat Uysal'a, jüri üyelerim Dr. Öğr. Üyesi Gültekin Atalık ve Dr. Öğr. Üyesi Mustafa Çavuş'a değerli görüşleri ve yapıcı eleştirileri için teşekkür ederim.

Çalışmamın veri toplama sürecinde harikulade derecede destekleyici olan tüm arkadaşlarıma da teşekkürü borç bilirim.

Hayatımın her aşamasında bana yardımcı olan, her zaman her türlü desteklerini benden esirgemeyen en kıymetlilerime; eşim Ferdi Karakütük'e, kardeşim Sinem Kaya'ya, babam Ahmet Kaya'ya ve annem Gülsen Kaya'ya en içten teşekkürlerimi sunarım.

Aslı KAYA KARAKÜTÜK

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ

Bu tezin bana ait, özgün bir çalışma olduğunu; çalışmamın hazırlık, veri toplama, analiz ve bilgilerin sunumu olmak üzere tüm aşamalarında bilimsel etik ve kurallara uygun davrandığımı; bu çalışma kapsamında elde edilen tüm veri ve bilgiler için kaynak gösterdiğimi ve bu kaynaklara kaynakçada yer verdiğimi; bu çalışmanın Eskişehir Teknik Üniversitesi tarafından kullanılan “bilimsel intihal tespit programı”yla tarandığını ve hiçbir şekilde “intihal içermediğini” beyan ederim. Herhangi bir zamanda, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçları kabul ettiğimi bildiririm.

Aslı KAYA KARAKÜTÜK

İÇİNDEKİLER

Sayfa

BAŞLIK SAYFASI	I
JÜRİ VE ENSTİTÜ ONAYI.....	II
DANIŞMAN ONAYI	III
ÖZET	IV
ABSTRACT.....	V
TEŞEKKÜR	VI
ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ.....	VII
İÇİNDEKİLER	VIII
TABLolar DİZİNİ	X
ŞEKİLLER DİZİNİ	XI
SİMGELER VE KISALTMALAR DİZİNİ	XII
1. GİRİŞ	1
2. LİTERATÜR TARAMASI	5
2.1. Sınıflandırma Problemleri için Literatür Taraması (Aşırı Öğrenme Makineleri ve Çekirdek Tabanlı Aşırı Öğrenme Makineleri)	5
2.2. Sınıflandırma Problemleri için Literatür Taraması (Dalgacık ve Bulanık Dalgacık Yapay Sinir Ağları)	8
3. BULANIK DALGACIK ÇEKİRDEK TABANLI AŞIRI ÖĞRENME MAKİNESİ ALGORİTMASI YAKLAŞIMI.....	13
3.1. Aşırı Öğrenme Makineleri (ELM).....	13
3.2. Çekirdek Tabanlı Aşırı Öğrenme Makineleri (KELM)	17
3.3. Bulanık Dalgacık Yapay Sinir Ağları (FWNN).....	19
3.3.1. Dalgacık ve tarihsel gelişimi	19
3.3.2. Dalgacık yapay sinir ağları (WNN)	23
3.3.3. Bulanık dalgacık yapay sinir ağları (FWNN)	27

3.4. Hibrit Algoritma Önerisi (Fuzzy Wavelet Kernel Extreme Learning Machine) (FWKELM)	32
3.4.1. FWKELM algoritmasının matematiksel yapısı.....	38
4. BULGULAR.....	41
4.1. Uygulama 1	41
4.2. Uygulama 2	45
4.3. Uygulama 3	53
5. SONUÇ VE ÖNERİLER.....	61
5.1. Sonuç	61
5.2. Öneriler	63
KAYNAKÇA.....	64
EKLER	
ÖZGEÇMİŞ	

TABLÖLER DİZİNİ

Sayfa

Tablo 3.1. ELM’de aktivasyon fonksiyonları	15
Tablo 3.2. Bulanık sinir ağlarının sınıflandırılması	27
Tablo 3.3. FW ile çekirdek tabanlı ELM yapısının eşitliği.....	37
Tablo 4.1. UCI veri tabanından alınan veri setlerine ilişkin bilgiler.....	41
Tablo 4.2. Sınıflandırma problemleri için performans karşılaştırması	43
Tablo 4.3. LBP+FWKELM sınıflandırıcısı kullanılarak elde edilen karışıklık matrisi	48
Tablo 4.4. HOG+FWKELM sınıflandırıcısı kullanılarak elde edilen karışıklık matrisi	48
Tablo 4.5. Farklı yaklaşımların sınıflandırma oranının karşılaştırmalı sonuçları	49
Tablo 4.6. FER2013 veri seti için FWKELM sınıflandırıcısı kullanılarak elde edilen karışıklık matrisi.....	52
Tablo 4.7. Orijinal veri seti için normalize edilmiş karışıklık matrisi	57
Tablo 4.8. Eğitim ve test sürelerinin ELM, WNN ve FWKELM algoritmalarında karşılaştırılması	60

ŞEKİLLER DİZİNİ

Sayfa

Şekil 1.1. Önerilen yaklaşımın temelinde yer alan algoritmalar ve birleşimleri	3
Şekil 3.1. ELM'nin yapısı.....	14
Şekil 3.2. Dalgacık sinir ağı algoritmasının akış şeması	25
Şekil 3.3. Bulanık dalgacık sinir ağının yapısı	29
Şekil 4.1. Morlet fonksiyonunun yapısı.....	42
Şekil 4.2. Yüz ifadesi için akış diyagramı	45
Şekil 4.3. JAFFE veri seti görüntü örnekleri	46
Şekil 4.4. Orijinal görüntüler için ön işleme gösterimi	46
Şekil 4.5. Yerel ikili örüntü işleme örneği.....	47
Şekil 4.6. FER2013 veri seti görüntü örnekleri	49
Şekil 4.7. Eğitim veri seti duygu dağılımı	50
Şekil 4.8. Test veri seti duygu dağılımı	50
Şekil 4.9. Gabor çekirdeklerinin beş ölçek ve sekiz yönelimdeki gerçek kısmı	51
Şekil 4.10. Bir yüz görüntüsü üzerinde Şekil 4.9'da gösterilen kıvrımlı 40 filtrenin sonuçları.....	52
Şekil 4.11. Orijinal yüz veri seti görüntü örnekleri	54
Şekil 4.12. Orijinal resim- gri işlemeye tabi tutulmuş resim- yüz tespiti	56
Şekil 4.13. Yönlendirilmiş gradyanların histogramları ile öznetelik çıkarımı	56
Şekil 4.14. FWKELM yöntemiyle eğitilmiş orijinal veri setinden doğru sınıflandırılmış ve yanlış sınıflandırılmış görüntüleri örnekleri	58
Şekil 4.15. Orijinal görüntülerde algoritmaların sınıflandırma doğrulukları	59

SİMGELER VE KISALTMALAR DİZİNİ

α	: Alpha
β	: Beta
ε	: Epsilon
η	: Eta
γ	: Gamma
ξ	: Ksi
λ	: Lamda
μ	: Mü
ω	: Omega
ϕ	: Phi
Ψ	: Psi
A-FWNC	: Adaptive Fuzzy Wavelet Network Controller (Uyarlanabilir Bulanık Dalgacık Ağ Denetleyicisi)
ANFIS	: Adaptive-Network Based Fuzzy Inference Systems (Uyarlamalı Ağ Tabanlı Bulanık Çıkarım Sistemi)
ANN	: Artificial Neural Network (Yapay Sinir Ağı)
APSO	: Active Operators Particle Swam Optimization Algorithm (Aktif Operatörler Partikül Yüzmesi Optimizasyon Algoritması)
BP	: Backpropagation (Geri Yayılım)
BPNN	: Backpropagation Neural Network (Geri Yayılımlı Yapay Sinir Ağı)
ELM	: Extreme Learning Machine (Aşırı Öğrenme Makineleri)
FEBANN	: Feedforward Error Backpropagation Artificial Neural Networks (İleri Beslemeli Hata Geri Yayılımlı Yapay Sinir Ağı)
FIS	: Fuzzy Inference Systems (Bulanık Çıkarım Sistemleri)
FLD	: Fisher's Linear Discriminant Analysis (Fisher'in Lineer Diskriminant Analizi)
FLNN	: Functional Link Neural Network (Fonksiyonel Bağlantı Sinir Ağı)
FM	: Fuzzy Models (Bulanık Modeller)
FNN	: Fuzzy Neural Network (Bulanık Sinir Ağı)
FWELM	: Fuzzy Wavelet Extreme Learning Machine (Bulanık Dalgacık Aşırı Öğrenme Makineleri)

FWKELM	: Fuzzy Wavelet Kernel Extreme Learning Machine (Bulanık Dalgacık Çekirdek Tabanlı Aşırı Öğrenme Makineleri)
FWNN	: Fuzzy Wavelet Neural Network (Bulanık Dalgacık Yapay Sinir Ağı)
HOG	: Histogram Of Oriented Gradient (Yönlendirilmiş Gradyanların Histogramları)
JAFFE	: The Japanese Female Facial Expression
KELM	: Kernel Extreme Learning Machine (Çekirdek Tabanlı Aşırı Öğrenme Makineleri)
KKT	: Karush – Kuhn – Tucker
kNN	: k-Nearest Neighbors (k- En Yakın Komşuluk)
LBP	: Local Binary Pattern (Yerel İkili Örüntüler)
LDA	: Linear Discriminant Analysis (Linear Diskriminant Analizi)
LS-SVM	: Least-Squares Support Vector Machine (En Küçük Kareler Destek Vektör Makineleri)
MHWKELM	: Mexican Hat Wavelet Kernel Extreme Learning Machine (Mexican Hat Dalgacık Çekirdek Tabanlı Aşırı Öğrenme Makinesi)
MK-ELM	: Multi-Kernel Extreme Learning Machine (Çoklu Çekirdek Aşırı Öğrenme Makinesi)
MSE	: Mean Squared Error (Ortalama Karesel Hata)
OAA	: One-Against-All (Bire Karşı Hepsi)
OAo:	: One-Against-One (Bire Karşı Bir)
OLS	: Orthogonal Least Square (Ortogonal En Küçük Kareler)
ORL	: Our Database of Faces
OSFuzzyELM	: Online Sequential Fuzzy Extreme Learning Machine (Çevrimiçi Sıralı Bulanık Aşırı Öğrenme Makinesi)
PCA	: Principal Components Analysis (Temel Bileşenler Analizi)
PSO	: Particle Swarm Optimization (Parçacık Sürüsü Optimizasyonu)
R-ELM	: Regularized Extreme Learning Machine (Düzenleştirilmiş Aşırı Öğrenme Makineleri)
RBF	: Radial Basis Function (Radyal Tabanlı Fonksiyon)
SLFN	: Single Hidden Layer Feedforward Neural Network (Tek Gizli Katmanlı İleri Beslemeli Yapay Sinir Ağı)
SVM	: Support Vector Machine (Destek Vektör Makinesi)
SVR	: Support Vector Regression (Destek Vektör Regresyonu)
TLBO	: Teaching-Learning Based Optimization (Öğretme Öğrenmeye Dayalı Optimizasyon)
TSK	: Takagi Sugeno Kang Model
UCI	: UC Irvine Machine Learning Repository

VJ	:	Viola Jones Algoritması
WA	:	Wavelet Analysis (Dalgacık Analizi)
WNN	:	Wavelet Neural Network (Dalgacık Yapay Sinir Ağı)



1. GİRİŞ

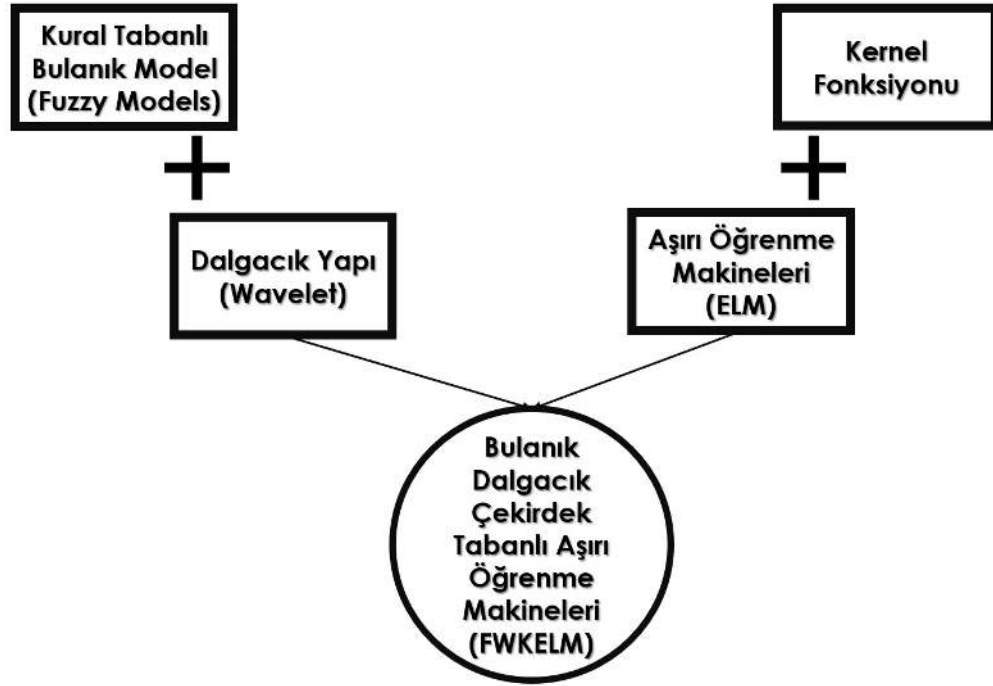
Bilginin uygun şekilde işlenmesiyle yüzleşmek için gelişmiş hesaplama algoritmalarının geliştirilmesi, tahmin yetenekleri ve gelişmiş performansı ile biyo-ilhamlı hesaplama modellerin kurulmasını teşvik etmiştir. İnsan zihninin belirsizlik ve kesinlikten uzak bir ortamda muhakeme etme ve öğrenme konusundaki olağanüstü yeteneğine paralel bir bilgi işlem yaklaşımı olan esnek hesaplama, zaman polinomunda kesin bir çözüm elde edilemeyen parametrik olmayan karmaşık problemlerin çözümü gibi hesaplama açısından zor görevlerde kesin olmayan çözümlerin kullanılmasıyla karakterize edilmektedir.

Yapay sinir ağları (YSA/ANN) [1], Dalgacık yapılar (WA) [2] ve Bulanık Modeller (BM/FM) [3] gibi esnek hesaplama teknikleri, yapay zekâ, örüntü tanıma, robotik, sistem tanımlama vb. [4–6] gibi çok çeşitli alanlarda uygulanmaktadır. ANN'ler arasında ileri beslemeli sinir ağı, en çok kullanılan öğrenme mekanizmalarından biridir [7,8]. ANN'ler öğrenme ve genelleme yeteneklerine sahip olmalarına rağmen, çok sayıda nörona ihtiyaç duyabilecekleri için ağ karmaşıklığından muzdariptirler [9]. Sinir ağı ve bulanık çıkarım sistemi (FIS) birleştirilerek, bulanık sinir ağı (FNN) geliştirilmiştir [10]. FNN, belirsiz veya kötü tanımlanmış faktörleri içeren karmaşık problemlerle başa çıkmak üzere bir teknik sunmak için sinir ağının öğrenme yeteneğini ve FIS'in yaklaşık muhakeme özelliklerini kullanmaktadır [11]. Dalgacık, sınırlı uzunluğa ve sıfır ortalamaya sahip bir dalgadır. Dalgacık dönüşümü, belirli bir dalgacık fonksiyonunu çevirdikten sonra, farklı ölçeklerde sinyal ile iç çarpımı elde etmek anlamına gelmektedir. Dalgacıkların çoklu çözünürlük ve zaman frekansı yerleştirme özellikleri matematik, fizik ve mühendislik alanlarındaki zor problemleri çözmek için güçlü yaklaşımlar haline getirmektedir [12]. Yapay sinir ağı, bulanık teori ve dalgacık dönüşümünden ilham alınarak, bulanık dalgacık sinir ağı (FWNN) oluşturulmuştur [13]. FWNN'ler, nöro-bulanık sistemlerin model doğruluğunu, genelleme yeteneğini ve hesaplama gücünü geliştirebilmektedir [14]. FWNN'lerde, her bulanık kuralın sonuç kısmı, lineer sonlu bir dalgacık fonksiyonları kümesini içermektedir. Bir bulanık dalgacık ağı üç tip parametreden oluşur: dalgacık katsayıları, üyelik fonksiyonlarının parametreleri ve çıktı katmanı ağırlıkları. Bu parametreler uygun bir yöntem kullanılarak başlatılmalı ve ardından ağın eğitim adımlarında güncellenmelidir [15]. Genel olarak, FWNN'lerde parametre başlatma ve yinelemeli öğrenme adımı için verimli yöntemler bulmak ana konulardan biridir. Bu faktörler sadece

eğitim algoritmalarının hızını belirlemek için önemli değildir, aynı zamanda modelin doğruluğunu da etkilemektedir.

Yukarıda belirtilen tüm sinir ağlarının yinelemeli öğrenme yöntemleriyle eğitildiği ve çoğunun geri yayılım (BP) gibi gradyan tabanlı öğrenme algoritmaları kullandığı belirtilmelidir. Bu algoritmalar tarafından sağlanan yüksek doğruluğa rağmen, parametreleri ayarlamak ve güncellemek için uygun bir yöntem bulmak, uzun hesaplama süresi ve yerel optimumlar gibi çeşitli dezavantajlardan mustarıptirler [16]. Bu sorunların üstesinden gelmek için, [16]'da Huang-Bin'e aşırı öğrenme makinesi (ELM) adı verilen tek bir gizli katmanlı ileri beslemeli sinir ağı (SLFN) eğitimi için bir yöntem sunulmuştur. BP gibi geleneksel öğrenme algoritmalarının aksine, ELM' de, gizli katman parametreleri ve gizli katman ile giriş katmanı arasındaki ağırlıkların ayarlanması gerekmemekte ve ağırlıklar bir kez rastgele atanmakta ve öğrenme prosedürü boyunca sabit kalmaktadır. Gizli katman ile çıktı katmanını birbirine bağlayan çıktı ağırlıkları analitik olarak belirlenmektedir. Geleneksel sinir ağı teorilerine göre, SLFN'ler, ağıın tüm parametrelerinin ayarlanabilir olmasına izin verildiğinde evrensel yaklaşım özelliklerine sahiptir. Sonuç olarak, SLFN'ler, rastgele seçilmiş gizli düğümler ve ayarlanabilir çıktı ağırlıkları ile evrensel yaklaşımdayıcılar olarak da çalışabilmektedir. Dalgacık yapay sinir ağları (WNN') ve ELM'nin çeşitli kombinasyonları prognostik, sağlık izleme, tahmin, tanımlama sorunları vb. [17] için önerilmiştir.

Yukarıda belirtilen sorunların üstesinden gelmek için, bu tez çalışması, dalgacık dönüşümlerinin, bulanık kavramların ve çekirdek tabanlı aşırı öğrenme makinelerinin çoklu çözünürlük analizi teorisine dayanan yeni bir bulanık dalgacık modeli yapısı sunmaktadır. Bu yöntemle bulanık dalgacık çekirdek tabanlı aşırı öğrenme makinesi (FWKELM) denir. Önerilen yaklaşımda yer alan algoritmalar ve birleşimleri Şekil 1.1 ile özetlenmiştir.



Şekil 1.1. Önerilen yaklaşımın temelinde yer alan algoritmalar ve birleşimleri

Önerilen yaklaşımda, dalgacıkların zaman-frekans yerelleştirme özellikleri, ANN'nin öğrenme yetenekleri, bulanık çıkarım sisteminin yaklaşık muhakeme özellikleri ve ayrıca aşırı öğrenme makinelerinin avantajları birleştirilmiş olup birçok uygulamada etkili bir çözüm ortaya koyacak özelliklerini sergileyebileceği düşünülmüştür. Bulanık dalgacık ağı ve KELM kombinasyonunun amacı, ağın karmaşıklığını azaltmak ve performans doğruluğunu korurken rastgele başlatma prosedürünün etkisini azaltmaktır. KELM algoritmasını önerilen bulanık dalgacık modeline uygulamak için öncelikle bir çekirdek tabanlı ELM'nin bir FW modeli ile eşdeğerliği kanıtlanmaktadır. Ayrıca sunulan modelde öncelikle üyelik fonksiyonlarının tüm parametreleri ve dalgacık katsayıları rasgele atanmakta ve daha sonra tek geçişli öğrenme yöntemi kullanılarak en küçük kareler çözümü ile her bir kuralın sonuç kısmındaki çıktı ağırlıkları analitik olarak belirlenmektedir. FWKELM'de, her bulanık kural, EĞER (IF) bölümünde sigmoid üyelik fonksiyonunu ve İSE (THEN) bölümünde Mexican hat dalgacık fonksiyonunu içermektedir.

FWKELM'nin performansını değerlendirmek için, gerçek hayat sınıflandırma problemlerinde, yüz ve yüz ifadesi tanıma probleminde gerçek makine öğrenimi kıyaslama problemlerini kullanan diğer iyi bilinen yöntemlerle karşılaştırılmıştır. Bu

bağlamda bu çalışmanın ikinci bölümünde detaylı literatür taramasından bahsedilecektir. Metodoloji bölümünde önerilen KELM algoritmasının FW modeli ile denkliği gösterilip hibrit algoritma olan FWKELM algoritması tanıtılacaktır. Dördüncü bölümde FWKELM algoritması çeşitli veri seti ve gerçek yüz ifadesi veri setinde uygulanacak ve diğer algoritmalar ile karşılaştırılacak ve yorumlanacaktır. Son bölümde ise önerilen yaklaşımın sahip olduğu avantaj ve dezavantajlar üzerinde durulacaktır.



2. LİTERATÜR TARAMASI

Bu bölümde, tezin amacı doğrultusunda önerilen hibrit algoritmanın temelinde yer alan Çekirdek Tabanlı Aşırı Öğrenme Makineleri ve Bulanık Dalgacık Yapay Sinir Ağları algoritmalarının sınıflandırma problemlerinde kullanılması kapsamında gerçekleştirilmiş literatür taraması özetlenmiştir.

2.1. Sınıflandırma Problemleri için Literatür Taraması (Aşırı Öğrenme Makineleri ve Çekirdek Tabanlı Aşırı Öğrenme Makineleri)

2004 yılında Huang ve Siew tarafından “Extreme Learning Machine: RBF Network Case” başlıklı çalışma ile aşırı öğrenme makinesi (ELM) yapısı radyal tabanlı fonksiyon (RBF) ağ durumuna göre genişletilmiştir [18].

2006 yılında Huang, Zhu ve Siew tarafından yapılan çalışmada geleneksel uygulamalardan farklı olarak, gizli düğümleri rastgele seçen ve SLFN'lerin çıktı ağırlıklarını analitik olarak belirleyen tek gizli katmanlı ileri beslemeli sinir ağları (SLFN'ler) için aşırı öğrenme makinesi (ELM) adı verilen yeni bir öğrenme algoritması önermişlerdir [16].

2011 yılında Zong ve Huang tarafından yapılan çalışmada çok etiketli yüz tanıma uygulamalarında sınıflandırma için bire karşı hepsi (OAA) ve bire karşı bir (OAO) aşırı öğrenme makinelerinin performansı incelenmiştir. Performans, dört karşılaştırmalı işaretleme yüz görüntüsü veri kümesiyle doğrulanmıştır [19].

2011 yılında Zong ve arkadaşları tarafından yapılan çalışmada yüz tanıma problemleri için en küçük kare çözümlerine dayanan orijinal aşırı öğrenme makinesinin çekirdeklerle uygulanabileceğini gösterilmiştir. Kernel ELM'nin performansını geleneksel en küçük kareler destek vektör makineleri (LS-SVM) algoritmasıyla karşılaştırmalı bir yüz tanıma veri seti üzerinden yapılmıştır. Deney sonuçları, Kernel ELM'nin hem tanıma tahmini doğruluğu hem de eğitim hızı açısından LS-SVM'den daha iyi performans gösterdiği belirtilmiştir [20].

2012 yılında Yang ve arkadaşları tarafından yapılan çalışmada geleneksel öğrenme yöntemlerinin zorluklarının üstesinden gelmek için parmak izi eşleştirmeye ilk olarak Aşırı Öğrenme Makineleri ve Düzenleştirilmiş Aşırı Öğrenme Makineleri (R-ELM) algoritmalarını uygulayarak yeni bir parmak izi tanıma sistemi önerilmiştir. Deneysel sonuçlar, önerilen yöntemlerin daha yüksek bir eşleştirme doğruluğuna sahip olduğunu ve daha az zaman harcadığını göstermektedir; bu nedenle gerçek zamanlı işlemeye

uygundurlar. Geleneksel yöntemleri içeren diğer karşılaştırmalı çalışmalar da ELM ve R-ELM ile önerilen yöntemlerin geleneksel yöntemlerden daha iyi performans gösterdiğini göstermektedir [21].

2013 yılında Liu ve arkadaşları tarafından yapılan çalışmada çekirdek tabanlı aşırı öğrenme makinesi ile yapılan uygulamalarda az önem verilen iki husustan -çekirdek seçimini optimize etmek ve sınıflandırma için birden fazla heterojen veri kaynağını entegre etmek için ELM için genel bir çerçeve olmaması- yola çıkarak çoklu çekirdek aşırı öğrenme makineleri (MK-ELM) olarak adlandırılan genel bir öğrenme çerçevesi önerilmiştir. Önerilen MK-ELM'de, ELM'nin optimum çekirdek kombinasyon ağırlıkları ve yapısal parametreleri birlikte optimize edilmiştir [22].

2013 yılında Pal, Maxwell ve Warner tarafından yapılan çalışmada hem çoklu hem de hiperspektral uzaktan algılama verilerini kullanan bir arazi örtüsü sınıflandırması için yeni bir çekirdek tabanlı aşırı öğrenme makinesi (ELM) algoritmasının etkinliği değerlendirilmiştir. Deneysel sonuçlar, en yaygın kullanılan algoritma olan destek vektör makineleri (SVM) ile karşılaştırılmıştır. Sonuçlar, yeni algoritmanın, sınıflandırma doğruluğu açısından SVM'ye benzer veya ondan daha doğru olduğunu, dikkate değer şekilde daha düşük hesaplama maliyetine sahip olduğunu ve çok sınıflı bir stratejinin uygulanmasını gerektirmediğini göstermiştir [23].

2014 yılında Chen ve arkadaşları tarafından yapılan çalışmada, hiperspektral görüntü sınıflandırması için spektral-uzamsal bilgileri entegre etmeyi ve çekirdek tabanlı ELM (KELM) sınıflandırıcı için uzamsal özellikleri kullanmanın avantajlarından yararlanma önerilmiştir. Önerilen Gabor filtrelemeye dayalı KELM sınıflandırıcısı ve çoklu hipotez tahminine dayalı KELM sınıflandırıcısı, iki gerçek hiperspektral veri kümesinde doğrulanmıştır. Sınıflandırma sonuçları, önerilen yöntemlerin, zorlu küçük eğitim örneği boyutu koşullarında, geleneksel piksel bazlı sınıflandırıcıların yanı sıra Gabor filtreleme tabanlı destek vektör makinesinden (SVM) ve çoklu hipotez tahminine dayalı SVM'den daha iyi performans gösterdiğini göstermektedir [24].

2015 yılında Iosifidis, Tefas ve Pitas tarafından yapılan çalışmada çoklu görünüm eylem tanıma için Aşırı Öğrenme Makinesi algoritmasını kullanılmıştır. Ağırlık parametrelerinin hesaplanması ve optimize edilmiş ağırlıkların belirlenmesi için çoklu eylem temsillerinden ve dağılım bilgilerinden yararlanabilen Aşırı Öğrenme Algoritmasının bir uzantısı önerilmiştir. Önerilen yaklaşımın yaygın olarak kullanılan üç video gösterimi kombinasyon yaklaşımı ve ilgili sınıflandırma şemaları ile deneysel

karşılaştırması, denetimli bir görünüm kombinasyon şeması kullanan ELM ağlarının genellikle denetimsiz kombinasyon yaklaşımlarından yararlananlardan daha iyi performans gösterdiğini belirtmişlerdir [25].

2016 yılında Lu ve arkadaşları tarafından yapılan çalışmada Kernel Aşırı Öğrenme Makineleri için optimum bir başlangıç parametreleri seti elde etmek için aktif operatörler partikül yüzmesi optimizasyon algoritması (APSO) kullanılmış ve böylece APSO-KELM olarak adlandırılan optimal bir KELM sınıflandırıcısı oluşturulmuştur. Standart genetik veri kümeleri üzerinde yapılan deneyler, APSO-KELM'in mevcut ELM, KELM ile karşılaştırıldığında daha yüksek sınıflandırma doğruluğuna sahip olduğu ve güvenilir, etkili bir sınıflandırma algoritması olduğu gösterilmiştir [26].

2016 yılında Parikh ve Shah tarafından yapılan çalışmada çekirdek tabanlı aşırı öğrenme makinesi ile bakteriyel enfeksiyon, mantar enfeksiyonu, egzama ve uyuz gibi yaygın cilt bozukluklarını tanımlama hedeflenmiştir. Radial Basis Function, Polinomial çekirdek, Üstel Ki-Kare çekirdek gibi çeşitli çekirdek fonksiyonlarını cilt veri setine uygulanmıştır. Deneyde Aşırı Öğrenme Makineleri (ELM) ve Destek Vektör Makinesi (SVM) ile bu çekirdekler kullanılarak elde edilen doğruluk ve eğitim süresi ölçülmüştür. Tüm bu öğrenme algoritmalarının doğruluk ve öğrenme süresinin karşılaştırmalı bir çalışması yapılmış ve ELM'nin üstel ki-kare çekirdeği kullanarak iyi bir sınıflandırma doğruluğu ile optimum öğrenme süresi verdiği görülmüştür [27].

2016 yılında Zhang ve arkadaşları tarafından yapılan çalışmada optimal genelleme performansını elde etmek için parçacık sürüsü optimizasyonu (PSO) kullanılarak Gauss çekirdeği ile KELM parametrelerini optimize edilmiştir. Önerilen yöntemin etkinliğini doğrulamak için, Grid algoritmasıyla optimize edilmiş KELM ile karşılaştırılmıştır [28].

2017 yılında Kongsorot ve arkadaşları tarafından yapılan çalışmada çok etiketli sınıflandırma problemleri için Kernel Aşırı Öğrenme Makinesi (KELM) algoritmasının sınırlaması olduğunu ve bu sorunu çözmek için bulanık küme teorisine dayanan çekirdek ile gelişmiş bir ELM önerilmiştir. DeneySEL sonuçlar, önerilen yöntemin çok etiketli problemler için ELM algoritma ailesinin yanı sıra son teknoloji çok etiketli sınıflandırma algoritmalarından daha iyi performansa sahip olduğunu göstermiştir [29].

2017 yılında Mahmud ve Maun tarafından yapılan çalışmada Aşırı Öğrenme Makinesine dayalı bir yüz ifadesi tanıma modeli önerilmiştir. Bir yüz görüntüsünden kaşlar, gözler, ağız ve burun gibi göze çarpan yüz özelliği segmentleri algılanmış ve daha sonra bu özellik segmentleri, özellik vektörleri oluşturmak için morfolojik görüntü işleme

operasyonu ve kenar algılama tekniği kullanılarak çıkarılmıştır. Aşırı Öğrenme Makinesi, giriş yüzlerinin ifadelerini mutlu, üzgün, sürpriz, kızgın, tiksinti ve korku gibi altı temel kategoriye ayırmak için kullanılmıştır. Yüz ifadesi tanıma sisteminin deneyleri JAFFE yüz ifadesi veri tabanı üzerinde gerçekleştirilmiş ve deneysel sonuçların performansları analiz edilmiştir [30].

2018 yılında Anam ve Al-Jumaliy tarafından yapılan çalışmada parçacık sürüşü optimizasyonunun (PSO) hibridizasyonu ve bir dalgacık fonksiyonu kullanılarak RBF-ELM parametrelerinin optimizasyonu önerilmiştir. Önerilen bu sistemler, elektromiyografi sinyallerini kullanarak ampüteler ve güçlü vücutlu denekler üzerindeki parmak hareketlerini sınıflandırmak için kullanılmıştır. Deneysel sonuçlar SW-RBF-ELM'nin, destek vektör makinesi (SVM), doğrusal ayırma analizi (LDA) ve k-en yakın komşu (kNN) gibi diğer iyi bilinen sınıflandırıcılardan daha iyi bir doğruluk elde ettiğini ortaya koymuştur [31].

2.2. Sınıflandırma Problemleri için Literatür Taraması (Dalgacık ve Bulanık Dalgacık Yapay Sinir Ağları)

2004 yılında Zhang ve arkadaşları yaptıkları çalışmada iki özneliği olan (1. yüz görüntülerinin iki boyutlu (2-D) dalgacık alt bant katsayıları ile temsil edilmesi ve 2. çekirdek ilişkisel belleğe dayalı modüler, kişiselleştirilmiş bir sınıflandırma yöntemi ile tanıma modeller) etkili bir yüz tanıma şeması önermişlerdir. PCA projeksiyonları ve düşük çözünürlüklü "başparmak-çivi" görüntü temsilleriyle karşılaştırıldığında, dalgacık alt bant katsayıları hesaplama karmaşıklığını düşük tutarken önemli yüz özelliklerini verimli bir şekilde yakalayabilir olduğunu vurgulamışlardır [32].

2006 yılında Subaşı ve arkadaşları tarafından yapılan çalışmada ileri beslemeli hata geri yayımlı yapay sinir ağları (FEBANN) ve dalgacık sinir ağları (WNN) tabanlı sınıflandırıcılar geliştirilmiş ve EMG sinyallerinin sınıflandırılmasındaki doğrulukları açısından karşılaştırmışlardır. WNN tabanlı sınıflandırıcı, FEBANN muadilinden daha iyi performans göstermiştir [33].

2006 yılında Wang, Li ve Chen tarafından yapılan çalışmada bulanık dalgacık yapay sinir ağı algoritması (FWNN) ile bir parmak izi sınıflandırma algoritması önerilmiştir. Parmak izi görüntüsü yön alanından çıkarılan özelliklere dayanmaktadır. Yönlü alanın doğruluğunu iyileştirmek için verimli bir tahmin yaklaşımı geliştirilmiştir. Geliştirilmiş yön alanına bağlı olarak, tekil noktalar ve ilgili özellikler, parmak izi

sınıflandırıcının giriş özelliklerini oluşturmak için çıkarılmıştır. Giriş özelliklerini kodladıktan sonra, beş sınıf problemi için parmak izlerini sınıflandırmak için bulanık bir dalgacık sinir ağı tabanlı sınıflandırıcı uygulanmıştır. Deneysel sonuçlar, önerilen algoritmanın mükemmel bir sınıflandırma performansını göstermektedir [34].

2008 yılında Sahoolizadeh ve arkadaşları yaptıkları çalışmada boyutsallık azaltma ve doğrusal ayırım analizi gerçekleştirebilen ve ayırt etme yeteneğini artırabilen Gabor dalgacığı ile hem yüz hareketi tanıma hem de yüz tanımlama için etkili bir yaklaşımı önermişlerdir. Yaptıkları deneylerde, önerilen Gabor dalgacık herhangi bir ön işleme aşaması olmaksızın ORL veri setinde %93 maksimum doğru tanıma oranına ulaşabildiklerini belirtmişlerdir [35].

2008 yılında Zekri ve arkadaşları tarafından yapılan çalışmada dalgacık dönüşümleri ve bulanık kavramları, çok çözünürlüklü analizi teorisinden esinlenerek, doğrusal olmayan afın sistemlerin kontrolü için yeni bir uyarlanabilir bulanık dalgacık ağ denetleyicisi (A-FWNC) sunulmuştur. Doğrudan uyarlamalı yaklaşımdan kaynaklanan önerilen uyarlamalı kazanç denetleyicisi, gerçek zamanlı işlem sırasında her bulanık kuralın THEN kısmındaki uyarlama parametresini ayarlama yeteneğine sahiptir. Her bulanık kural, bir alt dalgacık sinir ağına (alt-WNN) ve bir uyarlama parametresine karşılık gelmektedir. Her bir alt WNN, belirli bir genişleme değerine sahip dalgacıklardan oluşmaktadır. Ortogonal en küçük kareler (OLS) yöntemi, bulanık kural sayısını belirlemek ve her bir alt WNN için dalgacıkları saflaştırmak için kullanılmaktadır. OLS yönteminde kullanılan etkin dalgacık seçme prosedürü giriş boyutuna çok duyarlı olmadığından, yaklaşık fonksiyonun boyutu FWNN oluşturmak için darboğaza neden olmamaktadır. Önerilen A-FWNC'nin performansı, ikinci dereceden doğrusal olmayan ters çevrilmiş bir sarkaç sistemi uygulanarak gösterilmiş ve daha önce yayınlanmış yöntemlerle karşılaştırılmaktadır. Deney sonuçlarıyla, önerilen kontrol algoritmasının dikkate değer yeteneklerinin olduğu gösterilmiştir [36].

2010 yılında Yılmaz ve Sevcan tarafından yapılan çalışmada doğrusal olmayan dinamik sistemlerin tahmini ve tanımlanması için bulanık dalgacık sinir ağı (FWNN) modelleri sunulmuştur. Önerilen FWNN modelleri, geleneksel TSK (Takagi-Sugeno-Kang) bulanık sisteminden, bulanık kuralların THEN kısmı hem zaman hem de frekans alanında yerelleştirme yeteneğine sahip dalgacık tabanlı fonksiyonlarla değiştirilerek elde edilmiştir. Bulanık kuralların IF kısmında Gauss tipi aktivasyon fonksiyonları kullanılmıştır. Hızlı gradyan tabanlı bir eğitim algoritması, yani Broyden-Fletcher-

Goldfarb-Shanno yöntemi, FWNN modellerinin bilinmeyen parametreleri için en uygun değerleri bulmak için kullanılmıştır. Modellerin etkinliğini literatürde bilinen diğer yöntemlerle karşılaştırmak için simülasyon örnekleri de verilmiştir. Simülasyon sonuçlarına göre, önerilen FWNN modellerinin etkileyici bir genelleme yeteneğine sahip olduğu görülmüştür [37].

2011 yılında Shankar, Meher, Ghosh tarafından yapılan çalışmada çok bantlı uzaktan algılama görüntülerinde kara örtülerinin bulanık sınıflandırması için dalgacık özelliği tabanlı denetimli bir şema önerilmiştir. Önerilen şema, esnek bir hesaplama yaklaşımı olan dalgacık-bulanık hibridizasyon çerçevesinde geliştirilmiştir [38].

2015 yılında Abidin ve Alamsyah tarafından yapılan çalışmada geri yayımlı yapay sinir ağları (BPNN), yüz ifadesi görüntülerini (öfke, iğrenme, korku, mutluluk, üzüntü, nötr ve şaşkınlık olmak üzere) yedi sınıflı ifadelerle ayırmak sınıflandırıcı olarak kullanılmıştır. Öznitelik çıkarma görevleri amacıyla, görüntüleri ayrıştırmak için Haar dalgacık, Daubechies dalgacık ve Coiflet dalgacık olmak üzere üç ayrı dalgacık dönüşümü kullanılmıştır. Önerilen yöntemi analiz etmek için bir yüz ifadesi tanıma sistemi oluşturulmuştur. Önerilen yöntem, JAFFE veri tabanından alınan statik görüntüler üzerinde test edilmiştir [39].

2017 yılında Dharani ve Vibhute tarafından yapılan çalışmada otomatik yüz tanıma sistemini gerçekleştirmek için hibrit yaklaşım dalgacık yüz algılama ve sinir ağı tabanlı yüz tanıma önermişlerdir. Ön işleme ve özellik çıkarma adımları için dalgacık dönüşümü ve PCA-FLD'nin bir kombinasyonunu uygulanmıştır. Sınıflandırma aşamasında, geniş yüz varyasyonlarının varlığında sağlam bir karar vermek için sinir ağı araştırılmıştır [40].

2018 yılında Güzel ve Bilgin tarafından yapılan çalışmada görüntülerin dokusal özelliklerini çıkarmak için yerel ikili örüntü, yönlü gradyanların histogramları ve Gabor filtreler kullanılmıştır. Elde edilen dokusal özelliklere önermiş oldukları iki seviyeli özellik birleştirme yöntemi uygulanmıştır. Füzyon edilen özellikler üzerinde çekirdek tabanlı aşırı öğrenme makinesi kullanılarak hiperspektral görüntüler üzerindeki sınıflama başarılarının artırılması amaçlanmıştır. Bu amaçla aşırı öğrenme makinelerinden oluşan kararların birleştirilmesini de önermişlerdir [41].

2018 yılında Golestaneh, Zekri ve Sheikholeslam tarafından yapılan çalışmada bulanık dalgacık aşırı öğrenme makinesi (FW-ELM) adı verilen yeni bir yapı sunulmuştur. FW-ELM'nin ana hedefleri, doğrusal öğrenme parametrelerinin sayısını azaltarak ağ karmaşıklığını önemli ölçüde azaltmak ve kabul edilebilir doğruluk ve

genelleme performansları korunurken rastgele başlatma prosedürüne duyarlılığı azaltmaktır. Bu modelde, ağ karmaşıklığı ile performans doğruluğu arasında bir denge sağlamak için, her bulanık kuralın THEN bölümünde, her iki girdi için bir katsayı dikkate alınmaktadır. FW-ELM'yi değerlendirmek için OS-Fuzzy-ELM, Simpl_eTS, ANFIS gibi popüler bulanık modeller ve çeşitli kıyaslama veri kümeleri üzerinde ELM, BP ve SVR gibi diğer birkaç ilgili algoritma ile karşılaştırılmıştır. Deney sonuçları, önerilen yaklaşımdan kaynaklanan dikkate değer verimliliğini göstermiştir [42].

2018 yılında Sailaja ve Deepthi tarafından yapılan çalışmada dalgacık tabanlı öznitelik çıkarma teknikleri ışık yönündeki büyük değişikliklere ve poz varyasyonlarına karşı duyarsız olduğundan, Log Gabor dalgacık filtreleri kullanılarak yüz öznitelikleri çıkarılmıştır. Daha sonra destek vektör makinesine (SVM) dayalı sınıflandırma yapılmıştır. İfade tanıma sistemini test etmek için JAFFE standart veri tabanı kullanılmıştır [43].

2018 yılında Shi ve Jiang tarafından yapılan çalışmada dalgacık sinir ağının bulanık sinir ağına gömülü olduğu bir bulanık dalgacık sinir ağı (FWNN) yapısı önerilmiştir. Öznitelik çıkarma aşamasında, yüz görüntülerinin zaman ve frekans alanındaki özellikleri entegre edilerek daha kusursuz bilgi elde edilmiştir. Eğitim ve tanımlama aşamasında sınıflandırma için bulanık dalgacık ağını benimsemişler ve eğitim hızını artırmak için Levenberg-Marquardt algoritması kullanmışlardır. Deney, altı duyguyu tanımlamak için bir yüz duygu veri tabanında test edilmiştir; nötr, üzüntü, öfke, neşe, tiksinti, şaşkınlık, korku. Deneysel sonuçlar, modelin daha iyi sonuçlar elde edebileceğini göstermiştir [44].

2019 yılında Alneamy ve arkadaşları tarafından yapılan çalışmada, öğretme öğrenmeye dayalı optimizasyon (TLBO) algoritması ve fonksiyonel bağlantı sinir ağı (FLNN) ile bulanık dalgacık sinir ağı kombinasyonuna dayanan yeni bir sınıflandırma tekniği önerilmiştir. Ek olarak, TLBO algoritması, yeni hibrit fonksiyonel bulanık dalgacık sinir ağı'nı eğitmek ve ağırlıklar, genişleme ve öteleme olan öğrenme parametrelerini optimize etmek için kullanılmıştır. Önerilen yöntemin performansını değerlendirmek için beş standart tıbbi veri seti kullanılmıştır: Göğüs Kanseri, Kalp Hastalığı, Hepatit, Pima-Hint diyabeti ve Apandisit. Önerilen yöntemin verimliliği, ortalama kare hatası, sınıflandırma doğruluğu, çalışma süresi, duyarlılık, özgüllük ve kappa açısından 5 kat çapraz doğrulama ve 10 kat çapraz doğrulama kullanılarak değerlendirilmiştir. Önerilen yöntemin, ilgili önceki çalışmalarda bulunan diğer yöntemlerin performansına göre verimli performansa sahip olduğu görülmüştür [45].

2021 yılında Ahmedî ve arkadaşları tarafından yaptıkları çalışmada sinir ağı biçiminde bulanık mantık ve dalgacık dönüşümüne dayalı yeni bir sınıflandırıcı sunmuşlardır. Bu sınıflandırıcı, etiketlere veya sınıflara karşılık gelen sayısal özelliği tahmin etmek için bir katman içermektedir. Sunulan sınıflandırıcı, beyin tümörü teşhisinde uygulanmıştır. Özellik çıkarımı için, dört Gauss fonksiyonuna sahip bir fraktal model kullanılmıştır. Deney sonuçlarına göre önerilen sınıflandırıcı (FWNNet), fraktal özellik çıkarma yöntemi ve MRI görüntülerine dayalı beyin tümörü teşhisi ile %100'lük en yüksek doğruluğu göstermektedir. Sonuçlara göre, beyin tümörünün teşhisi için en iyi sınıflandırıcı FWNNet mimarisidir [46].



3. BULANIK DALGACIK ÇEKİRDEK TABANLI AŞIRI ÖĞRENME MAKİNESİ ALGORİTMASI YAKLAŞIMI

Bu bölümde önerilen hibrit algoritma ve algoritmanın temelinde yer alan metotlardan detaylı olarak bahsedilecektir.

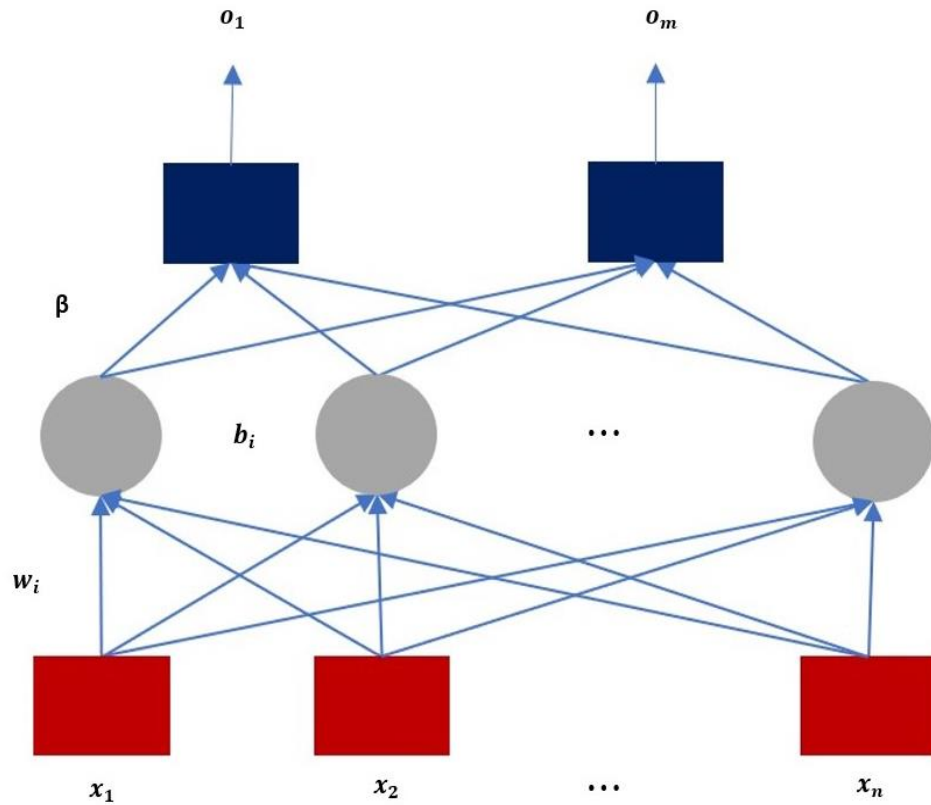
3.1. Aşırı Öğrenme Makineleri (ELM)

Sinir ağlarını temel alan yapay sinir ağları, 70 yılı aşkın bir geçmişe sahiptir [47]. İnsan beyinde bulunan sinir sistemini temel yapısı olarak öykünen yapay sinir ağları, tipik bir sinir ağı nöronlardan, bağlantılardan ve ağırlıklardan oluşan bir tür makine öğrenimi yöntemidir. Sinir ağında, algı [48] olarak da adlandırılan bir nöron, önceki nöronlardan giriş sinyallerini almakta ve sonraki nöronlara çıkış sinyalleri göndermektedir. Bir sinir ağı içindeki nöronlar genellikle farklı katmanlarda düzenlenmektedir. Bağlantı, farklı katmanlardaki nöronlar arasındaki ilişki olarak tanımlanmaktadır. Bir nöron, önceki katmanındaki ve sonraki katmanındaki nöronlarla çok sayıda bağlantıya sahip olabilmektedir ve her bağlantı, ağırlık adı verilen önemli bir parametre sunmaktadır. Bir nöronun çıkış sinyallerini vermek için giriş sinyallerini nasıl işlediğine karar veren ağırlıklardır.

Ağı eğitmek, veriye en iyi şekilde yaklaşmasını sağlayan katmanlarındaki ağırlıkları belirlemektir. Geleneksel eğitim algoritmaları genellikle geri yayılım (BP) [49] gibi eğime dayalıdır. BP, anlaşılması kolay ve uygulanması basit olduğu için birçok alanda yaygın olarak kullanılmaktadır. Bununla birlikte, gradyan tabanlı algoritmalar her zaman küresel en iyi çözümü getirmemektedir. Çözümleri, parametre başlatmaya ve özellik uzayının karmaşıklığına bağlıdır, dolayısıyla yerel ekstremumda yakınsama olasılığı daha yüksektir. Guang-Bin ve Qin-Yu [16] tarafından tek gizli katmanlı ileri beslemeli ağları (SLFN'ler) eğitmeyi amaçlayan aşırı öğrenme makinesi (ELM) önerilmiştir. Gradyan tabanlı yöntemlerden farklı olarak ELM, girdi ile gizli katman arasındaki ağırlıklara ve gizli katmandaki sapmalara rastgele değerler atamakta ve bu parametreler eğitim sırasında dondurulmaktadır. Gizli katmandaki doğrusal olmayan aktivasyon fonksiyonları, sistem için doğrusal olmayanlık sağlamaktadır. O halde lineer olmayan bir sistem olarak kabul edilebilmektedir. Öğrenilmesi gereken tek parametre, gizli katman ile çıktı katmanı arasındaki ağırlıktır. Bu nedenle ELM, yineleme olmadan öğrendiği için geleneksel algoritmalarından çok daha hızlı birleşmektedir. Ek olarak, rastgele gizli düğümler, evrensel yaklaşım yeteneği vaat etmektedir. Teorik analiz, ELM'nin, eğitilecek tüm parametrelere sahip geleneksel ağlara göre rastgele

parametrelerle global optimal çözüme ulaşma olasılığının daha yüksek olduğunu göstermiştir [50]. Üstün eğitim hızı ve iyi genelleme yeteneği nedeniyle [51] ELM, sınıflandırma, regresyon, kümeleme ve özellik haritalama gibi çeşitli öğrenme problemlerinde yaygın olarak uygulanmaktadır. ELM, belirli uygulamalar için stabilitesini ve genelleştirmesini daha da artırmak için geliştirilmiştir [52, 53, 54, 54]. Son on yılda, daha az manuel müdahale, daha yüksek sınıflandırma doğruluğu ve daha az eğitim süresi amaçlarını sağlayabilmek adına ELM üzerine kapsamlı araştırmalar gerçekleştirilmiştir. ELM'nin [56] genelleştirme ve küresel yaklaşım yeteneğini araştırmak için teorik analiz de yapılmıştır.

ELM, en yaygın kullanılan yapay sinir ağı yapısı olan tek gizli katmanlı ileri beslemeli sinir ağları (SLFN) eğitmek için icat edilmiştir. Geleneksel bir SLFN, giriş katmanı, gizli katman ve çıkış katmanı olmak üzere üç katmandan oluşmaktadır (Şekil 3.1). x ve o , giriş ve çıkış vektörünü belirtmektedir. w ve b , girdiden gizli katmana olan ağırlığı ve gizli katmanın sapmasını temsil etmektedir. β çıkış ağırlığını belirtmektedir. Ağı eğitmek, optimum çözüme ulaşan bu parametrelere karar vermektir.



Şekil 3.1. ELM'nin yapısı

Bir eğitim seti verildiğinde $S = \{(x_i, t_i) | x_i = (x_{i1}, x_{i2}, \dots, x_{in})^T \in R^n, t_i = (t_{i1}, t_{i2}, \dots, t_{im})^T \in R^m\}$ burada x_i giriş değerini ve t_i hedefi temsil etmektedir. Bir ELM'nin o çıkışı $\widehat{N}$ gizli nöronlar Denklem 3.1 ile ifade edilebilir:

$$\sum_{i=1}^{\widehat{N}} \beta_i g(w_i x_j + b_i) = o_j, \quad j = 1, \dots, N \quad (3.1)$$

Burada $g(x)$, gizli katmandaki aktivasyon fonksiyonu anlamına gelmektedir. ELM'de aktivasyon fonksiyonları, sistem için doğrusal olmayan haritalama sağlamak için doğrusal olmayan işlevlerdir. Tablo 3.1 yaygın olarak kullanılan çeşitli aktivasyon fonksiyonlarını listelemektedir.

Tablo 3.1. ELM'de aktivasyon fonksiyonları

Fonksiyon	Formül
Sigmoid fonksiyon	$G(a, b, x) = \frac{1}{1 + \exp(-(a \cdot x + b))}$
Hiperbolik fonksiyon	$G(a, b, x) = \frac{1 - \exp(-(ax + b))}{1 + \exp(-(ax + b))}$
Radyal tabanlı fonksiyon	$G(a, b, x) = \exp(-b\ x - a\)$
Hard- limit transfer fonksiyonu	$G(a, b, x) = \begin{cases} 1, & a \cdot x + b \leq 0 \\ 0, & d. d \end{cases}$
Kosinüs fonksiyonu	$G(a, b, x) = \cos(a \cdot x + b)$

Eğitimin amacı, hedef ile ELM'nin çıktısı arasındaki hatayı en aza indirmektir. En sık kullanılan nesne işlevi, karesel ortalama hatadır (MSE) ve Denklem 3.2 ile ifade edilir.

$$MSE = \sum_{i=1}^N (t_{ij} - o_{ij})^2, \quad j = 1, \dots, m \quad (3.2)$$

Denklem 3.2'de N , gizli düğüm sayısıdır ve i ve j , eğitim örneği ve çıktı katmanı düğümü için dizinlerdir. N gizli düğüm sayısı sonsuza yaklaştığında, SLFN'nin tüm eğitim örneklerine yaklaşabildiği kanıtlanabilir:

$$\sum_{j=1}^N \|o_j - t_j\| = 0 \quad (3.3)$$

Bu duruma evrensel yaklaşma yeteneği denmektedir. Bu nedenle, yeterli olan bir w_i, b_i ve β_i dizisi olmalıdır.

$$\sum_{i=1}^N \beta_i g(w_i x_j + b_i) = t_j, \quad j = 1, \dots, m \quad (3.4)$$

3.4 ile ifade edilen formül Denklem 3.5 ile gösterilecek şekilde kısaltılabilir;

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{T} \quad (3.5)$$

Denklem 3.5'teki $\mathbf{H}$, $\boldsymbol{\beta}$, $\mathbf{T}$ matrisleri Denklem 3.6 ve 3.7 ile gösterildiği gibidir;

$$H(w_1, \dots, w_N, b_1, \dots, b_N, x_1, \dots, x_N) = \begin{bmatrix} g(w_1 x_1 + b_1) & \dots & g(w_N x_1 + b_N) \\ \vdots & \ddots & \vdots \\ g(w_1 x_N + b_1) & \dots & g(w_N x_N + b_N) \end{bmatrix} \quad (3.6)$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_N^T \end{bmatrix}_{N \times m}, \quad \mathbf{T} = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N \times m} \quad (3.7)$$

Bu nedenle, SLFN'yi eğitmek en iyi w_i , b_i ve β_i parametrelerini bulmaktır.

ELM'nin temel eğitimi rastgele başlatma ve doğrusal parametre çözümü iki adım olarak kabul edilmektedir. İlk olarak, ELM kendi gizli katmanında rastgele w_i ve b_i parametreleri kullanmakta ve parametreler tüm eğitim süreci boyunca dondurulmaktadır. Girdi vektörü, eğitilmiş parametrelerden daha verimli olacak şekilde rastgele ayarlamak ve doğrusal olmayan aktivasyon fonksiyonları ile rastgele bir özellik uzayına eşlenmektedir. Doğrusal olmayan parçalı sürekli aktivasyon fonksiyonları ile ELM, evrensel yaklaşım kabiliyetine sahiptir [16, 57]. İkinci adımda, β_i , $\mathbf{H}\boldsymbol{\beta} = \mathbf{T}$ lineer bir problem olduğu için Moore-Penrose tersi (Denklem 3.8) ile elde edilmektedir.

$$\boldsymbol{\beta} = \mathbf{H}^\dagger \mathbf{T} \quad (3.8)$$

ELM, yinelemeli olarak ayarlanmış gizli parametreler olmadan daha iyi genelleme yeteneği sağlayabilmektedir. ELM teorisi, Guang-Bin ve Qin-Yu [16] tarafından titizlikle kanıtlanmıştır. ELM, yeterli sayıda gizli düğüm ile keyfi karmaşık sınıflandırma sınırına yaklaşabilmektedir. ELM algoritmasının öğrenme prosedürü aşağıdaki adımlarda yürütülmektedir [54, 55]:

- ✓ **Adım 1:** Gizli düğümlerin giriş ağırlığı w_i ve bias b_i 'yi rastgele dağıtınız.
 $i = 1, 2, \dots, N$.
- ✓ **Adım 2:** Gizli katman $\mathbf{H}$ çıktı matrisini Denklem 3.6 yardımı ile hesaplayınız.
- ✓ **Adım 3:** $\boldsymbol{\beta}$ çıkış ağırlığını Denklem 3.8'e dayalı olarak hesaplayınız.

3.2. Çekirdek Tabanlı Aşırı Öğrenme Makineleri (KELM)

Geleneksel sinir ağı öğrenme algoritmaları, eğitim hatalarını en aza indirerek ampirik riskleri de en aza indirmektedir. Doğrusal olmayan bir işleve uyması için giriş ve çıkış verilerini kullanmaktadır. Bununla birlikte, bu geleneksel sinir ağı öğrenme algoritmaları genellikle daha fazla eğitim örneği ve daha uzun öğrenme süresi gerektirir, bu da genellikle öğrenme modellerinin aşırı uyumuna yol açmaktadır.

ELM'deki eğitim süreci sadece bir kerelik tamamlama gerektirir; çünkü ELM'nin gizli katmanları için parametreleri yeniden ayarlaması gerekmez ve tek gizli katman ileri beslemeli ağların parametrelerini asgari norm kullanarak doğrusal denklem gruplarını çözme problemine dönüştürür ve ağ çıkışının ağırlığı olarak en küçük kareler çözümünden yararlanır. Böylece, eğitim hızı büyük ölçüde hızlanır ve genelleme performansı iyileştirilir. Geleneksel öğrenme algoritmalarının aksine, ELM'nin amacı hem eğitim hatalarını hem de çıktı ağırlığı normunu aynı anda en aza indirmektir [58].

Çekirdek aşırı öğrenme makinesi (KELM), ELM'ye dayalı olarak Huang ve arkadaşları tarafından geliştirilmiştir. Geleneksel ELM modeli, Denklem 3.1 ile tek gizli katmanlı ileri beslemeli sinir ağı (SLFN) mimarisine sahip iken Çekirdek aşırı öğrenme makinesi algoritmasının matematiksel modeli aşağıda verilmiştir;

N keyfi ayrı $(x_i, t_i) \in R^n \times R^m$ eğitim verisi için; burada $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T$ ve $t_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T$ şu şekilde ifade edilebilir:

$$\min \sum \|\beta \cdot h(x_i) - t_i\|^2 \text{ ve } \min \|\beta\| \quad (3.9)$$

Denklem 3.9'da β çıktı ağırlık matrisidir, $h(x_i)$ gizli katman çekirdek eşlemesidir ve t_i örneklerin etiketidir. Standart optimizasyon teorisi bakış açısından, yukarıdaki optimizasyon problemi, amaç fonksiyonu denklem 3.10 gibi yeniden yazılarak basitleştirilmiş bir kısıtlı optimizasyon problemi olarak çözülebilmektedir:

$$\min L_p = \frac{1}{2} \|\beta\|^2 + \frac{1}{2} C \sum_{i=1}^N \xi_i^2 \quad (3.10)$$

Kısıt;

$$h(x)\beta = t_i - \xi_i, \quad i = 1, \dots, N \quad (3.10.a)$$

3.10 ile ifade edilen denklemde ξ_i eğitim hatasını belirtir, C ceza parametresidir ve N örnek sayısıdır. KKT (Karush – Kuhn – Tucker) teorisine [26] dayanarak, ELM'nin eğitimi aşağıdaki ikili optimizasyon probleminin çözülmesine eşdeğerdir.

$$L_{PELM} = \frac{1}{2} \|\beta\|^2 + \frac{1}{2} C \sum_{i=1}^N \xi_i^2 - \sum_{i=1}^N \alpha_i (h(x_i)\beta - t_i + \xi_i) \quad (3.11)$$

Denklem 3.11’de her Lagrangian operatörü α_i i. örneğe karşılık gelir. KKT için aşağıdaki optimal koşullar elde edilebilir:

$$\frac{\partial L_{PELM}}{\partial \beta} = 0 \Rightarrow \beta = \sum_{i=1}^N \alpha_i (h(x_i))^T = H^T \alpha \quad (3.12)$$

$$\frac{\partial L_{PELM}}{\partial \xi_i} = 0 \Rightarrow \alpha_i \xi_i = 0 \quad i = 1, 2, \dots, N \quad (3.13)$$

$$\frac{\partial L_{PELM}}{\partial \alpha_i} = 0 \Rightarrow h(x_i)\beta - t_i + \xi_i = 0 \quad i = 1, 2, \dots, N \quad (3.14)$$

Küçük eğitim örnekleri için yukarıdaki formül şu şekilde yeniden yazılmıştır:

$$\left(\frac{I}{C} + HH^T \right) \alpha = T \quad (3.15)$$

Denklem 3.15 ile ifade edilen denklemde I birim matrisi, C ceza parametresidir ve H gizli katman çıktı matrisidir. ELM'in çıktı ağırlığı şu şekilde hesaplanır:

$$\beta = \left(\frac{I}{C} + HH^T \right)^{-1} H^T T \quad (3.16)$$

KELM'in çekirdek matrisini aşağıdaki şekilde tanımlamak için Mercer'in koşulunun uygulanması Denklem 3.17 ile gösterilmektedir. Mercer koşulu, yüksek boyutta çalışılırken çekirdek fonksiyonların her zaman iki girdi vektörünün iç çarpımı şeklinde ifade edilmesini sağlamaktadır.

$$\Omega_N = H_N H_N^T: \Omega_{N_{ij}} = h(x_i) \cdot h(x_j) = K(x_i, x_j) \quad (3.17)$$

Denklem 3.17’de H_N gizli katman çıktı matrisi, $\Omega_{N_{ij}}$ kernel matrisi, $K(x_i, x_j)$ kernel fonksiyonu ve x_i, x_j Ω_N kernel matrisinde sırasıyla konumlandırılmış i. ve j. elemanlardır. KELM modelinin çıktısı şu şekilde hesaplanır:

$$f(x_p) = \begin{bmatrix} K(x_p, x_1) \\ \vdots \\ K(x_p, x_N) \end{bmatrix}^T \alpha_N \quad (3.18)$$

$$\alpha_N = \left(\frac{I_N}{C} + \Omega_N \right)^{-1} T_N \quad (3.19)$$

Denklem 3.19 ile ifade edilen denklemde α_N KELM’nin çıktı ağırlığıdır.

3.3. Bulanık Dalgacık Yapay Sinir Ağları (FWNN)

3.3.1. Dalgacık ve tarihsel gelişimi

Tezin bu bölümünde bulanık dalgacık yapay sinir ağlarına değinilmeden önce dalgacık ifadesinin ne olduğu ve yapay sinir ağlarına nasıl entegre edildiğine dair kısa bir bilgilendirme yapılacaktır.

Dalgacıklar, verileri farklı frekans bileşenlerine bölen ve ardından her bileşeni ölçeğine uygun bir çözünürlükle inceleyen matematiksel işlevlerdir. Dalgacıkların arkasındaki temel fikir, ölçeğe göre analiz etmektir. Aslında, dalgacık alanındaki bazı araştırmacılar, dalgacıkları kullanarak veri işlemede tamamen yeni bir zihniyet benimsediğini düşünmektedir. Dalgacıklar, belirli matematiksel gereksinimleri karşılayan ve verileri veya diğer işlevleri temsil etmede kullanılan işlevlerdir.

Dalgacık fikri yeni değildir. Fonksiyonların süperpozisyonunu kullanan yaklaşım, Joseph Fourier'nin diğer fonksiyonları temsil etmek için sinüsleri ve kosinüsleri üst üste bindirebileceğini keşfettiği 1800'lerin başından beri var olmuştur. Ancak dalgacık analizinde, verilere bakmak için kullandığımız ölçek özel bir rol oynamaktadır. Dalgacık algoritmaları, verileri farklı ölçeklerde veya çözünürlüklerde işlemektedir. Büyük bir "pencereye" sahip bir sinyale bakılırsa, kaba özellikler fark edilir. Benzer şekilde, küçük bir "pencereye" sahip bir sinyale bakılırsa, küçük özellikler fark edilir. Dalgacık analizinde sonuç, tabiri caizse hem ormanı hem de ağaçları görmektir. Bu, dalgacıkları ilginç ve kullanışlı kılmaktadır. Bilim insanları, dalgalı sinyalleri yaklaşık olarak tahmin etmek için Fourier analizinin temellerini oluşturan sinüs ve kosinüslerden daha uygun fonksiyonlar istemişlerdir [59]. Tanımları gereği, bu işlevler yerel değildir (ve sonsuza kadar uzanır). Bu nedenle, keskin sivri uçlara yaklaşma konusunda çok kötü bir iş çıkarmaktadırlar. Ancak dalgacık analiziyle, sonlu alanlarda düzgün bir şekilde yer alan yaklaşık fonksiyonlar kullanılabilir. Dalgacıklar, keskin süreksizliklere sahip verilere yaklaşmak için çok uygundur.

Dalgacık analizi prosedürü, dalgacığı veya ana dalgacık olarak adlandırılan bir dalgacık prototip fonksiyonunu benimsemektir [60]. Frekans analizi, aynı dalgacığın genişlemiş, düşük frekanslı bir versiyonu ile gerçekleştirilirken, zamansal analiz prototip dalgacığın kısa, yüksek frekanslı bir versiyonu ile gerçekleştirilir. Orijinal sinyal veya fonksiyon bir dalgacık genişlemesi cinsinden temsil edilebildiğinden (dalgacık fonksiyonlarının lineer bir kombinasyonundaki katsayılar kullanılarak), veri işlemleri sadece karşılık gelen dalgacık katsayıları kullanılarak gerçekleştirilebilir. Ayrıca, verilere

uyarlanmış en iyi dalgacıklar seçilirse veya katsayılar bir eşiğin altına düşürülürse, veriler seyrek olarak temsil edilebilir. Bu seyrek kodlama, dalgacıkları veri sıkıştırma alanında mükemmel bir araç haline getirmektedir.

Dalgacıklar yalnızca birkaç parlak keşiften değil çeşitli alanlarda halihazırda var olan kavram ve teorilerden geliştirilmiştir. Dalgacık yapısının gelişiminde bazı önemli kilometre taşlarının tarihi turları kısaca anlatılmaktadır.

Matematik tarihinde, dalgacık analizi birçok farklı köken göstermektedir [60]. İşin çoğu 1930'lu yıllarda yapılmasına rağmen ayrı çabalar tutarlı bir teorenin parçaları gibi görünmemektedir. 1930'dan önce, matematiğin dalgacıklara götüren ana dalı, Joseph Fourier (1807) ile şimdilerde genellikle Fourier analiz olarak anılan frekans analizi teorileriyle başlamıştır. Fourier, herhangi bir 2π -periyodik $f(x)$ fonksiyonunun;

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \sin nx \quad (3.20)$$

Fourier serisinin toplamı olduğunu iddia etmiştir [61]. a_0, a_n ve b_n katsayıları sırasıyla Denklem 3.21, Denklem 3.22 ve Denklem 3.23 yardımıyla hesaplanmaktadır.

$$a_0 = \frac{1}{2\pi} \int_0^{2\pi} f(x) dx \quad (21)$$

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(nx) dx \quad n = 0, 1, 2, \dots \quad (22)$$

$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(nx) dx \quad n = 0, 1, 2, \dots \quad (23)$$

Fourier'nin iddiası, matematikçilerin fonksiyonlar hakkındaki fikirlerinin evriminde önemli bir rol oynamış ve yeni bir işlevsel evrenin kapısını açmıştır. 1807'den sonra matematikçiler, fonksiyonların, Fourier seri yakınsamalarının ve ortogonal sistemlerin anlamlarını keşfederek, kademeli olarak önceki frekans analizi kavramlarından ölçek analizi kavramına yönelmişlerdir.

Dalgacıklardan ilk defa, A. Haar'ın tezinin ekinde bahsedilmektedir [62]. Haar tezinde $[0,1]$ üzerinde tanımlanan fonksiyonların başka bir $h_0, h_1, \dots, h_n, \dots$ ortonormal sistemi var mı, öyle ki $[0,1]$ aralığında tanımlı herhangi bir sürekli f fonksiyonu için, seri

$$\langle h_0, f \rangle h_0(x) + \langle h_1, f \rangle h_1(x) + \dots + \langle h_n, f \rangle h_n(x) \quad (3.24)$$

$[0,1]$ aralığında düzgün bir şekilde $f(x)$ 'e yakınsar sorusuna cevap aramış ve sonsuz sayıda çözümü olan probleme Haar en basit çözümü keşfetmiş ve böylelikle dalgacıklara giden yollardan birini açmıştır [62].

Haar dalgacığının bir özelliği, kompakt desteğe sahip olmasıdır, bu da sonlu bir aralığın dışında kaybolması anlamına gelmektedir. Ne yazık ki, Haar dalgacıkları, sürekli olarak türevlenemezler ve bu durum uygulamalarını bir şekilde sınırlamaktadır.

1930'larda, bağımsız olarak çalışan birkaç grup, ölçekle değişen temel işlevleri kullanarak fonksiyonların temsiliğini araştırmıştır. Temel işlevler ve ölçekle değişen temel işlevler kavramlarını anlamak, dalgacıkları anlamının anahtarıdır. 1930'larda fizikçi olan Paul Levy, Haar temel işlevi adı verilen ölçekle değişen bir temel işlevi kullanarak, bir tür rastgele sinyal (süreç) olan Brownian hareketini araştırmıştır [63]. Brownian hareketindeki küçük karmaşık detayları inceleyerek Haar temel fonksiyonunu Fourier temel fonksiyonlarından daha üstün bulmuştur.

Littlewood, Paley ve Stein tarafından 1930'larda yapılan başka bir araştırma çalışması, bir $f(x)$ fonksiyonunun enerjisini lokalize etmeyi (Denklem 3.25) içermektedir:

$$enerji = \frac{1}{2} \int_0^{2\pi} |f(x)|^2 dx \quad (3.25)$$

Araştırmacılar, ölçek olarak değişebilen ve işlevsel enerjiyi hesaplarken enerji tasarrufu sağlayabilen bir işlev keşfetmişlerdir. Çalışmaları, David Marr'a 1980'lerin başında dalgacıklar kullanarak sayısal görüntü işleme için etkili bir algoritma sağlamıştır.

1960 ve 1980 yılları arasında, matematikçiler Guido Weiss ve Ronald R. Coifman, ortak bir işlev için atomları bulmak ve hepsinin yeniden oluşturulmasına izin veren "birleştirme kurallarını" bulmak amacıyla, Lusin'in teorisini atomlar ve atomik ayrışmalar açısından yorumlamış ve bir işlev uzayının atom adı verilen en basit öğelerini incelemişlerdir [64].

1980'de, bir fizikçi ve bir mühendis olan Grossman ve Morlet, dalgacıkları kuantum fiziği bağlamında geniş bir şekilde tanımlamışlardır. Bu iki araştırmacı, dalgacıklar için fiziksel sezgiye dayalı bir düşünme yolu sağlamışlardır [65].

1985'te Stephane Mallat, dijital sinyal işleme alanındaki çalışmalarıyla dalgacıklara ek bir sıçrama sağlamıştır. Mallat, kareleme ayna filtreleri, piramit algoritmaları ve ortonormal dalgacık tabanları arasında bazı ilişkiler keşfetmiştir [66]. Kısmen bu sonuçlardan ilham alan Meyer, önemsiz olmayan ilk dalgacıkları oluşturmuştur. Haar

dalgacıklarının aksine, Meyer dalgacıkları sürekli olarak türevlenebilir; ancak kompakt destekleri yoktur [60]. Birkaç yıl sonra, Ingrid Daubechies, Mallat'ın çalışmalarını, belki de en şık olan ve bugün dalgacık uygulamalarının mihenk taşı haline gelen bir dizi dalgacık ortonormal temel fonksiyon oluşturmak için kullanmıştır [67]. Daubechies'in kompakt destekli dalgacık tabanlarının keşfi dalgacık teorisinin gelişimi için bir başka kilometre taşıdır. Daubechies'in tabanları pürüzsüzlükte çok yönlü ve sonraki genellemeler ve teorik tavsiyeler için başlama noktası sunmaktadır.

Dalgacık teorisi birçok disiplinde kullanılarak gelişmektedir. Dalgacıklar zaman-ölçek tipli problemlerin uygulanması için kullanışlı ve ilgi çekici araçlar sağlamaktadır. Sinyallerin analizinde dalgacık gösterimi, ölçek bileşenleri arasında bir zaman alanı evrimini görmemizi sağlamaktadır.

Dalgacık ψ , ortalama sıfır değerine sahip, etkin sınırlı süreli bir dalga biçimidir. WA prosedürü, ana dalgacık adı verilen belirli bir dalgacık işlevini benimsemektedir. Bir dalgacık ailesi, kompakt olarak desteklenen bir ölçekleme fonksiyonunun φ (veya baba dalgacık) ve bir dalgacık fonksiyonunun ψ (veya ana dalgacık) genişlemesi ve çevrilmesi ile oluşturulan bir dizi ortogonal temel fonksiyondur.

Baba dalgacıklar φ ve anne dalgacıklar ψ tatmin edilir:

$$\int_{-\infty}^{\infty} \varphi(t) dt = 1 \quad (3.26)$$

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \quad (3.27)$$

Dalgacık ailesi, genişleyen ve anne dalgacıklarının çevrilmiş formları olan dalgacık çocuklarından oluşur:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (3.28)$$

Denklemler 3.28'te a ölçek veya dilatasyon parametresidir. Bu parametredeki değişim dalgacığı yatay eksen boyunca belli bir oranda uzatır veya büzer. Burada $|a| < 1$ iken dalgacık büzülür, $|a| > 1$ iken uzar. b ise kaydırma veya öteleme parametresidir.

Çoğu durumda, a ve b değerleri seçimi ayrı bir küme kullanarak sınırlandırılır, çünkü mümkün olan her ölçekte dalgacık katsayılarını hesaplamak hesaplama açısından yoğundur. Temporal analiz, ana dalgacıkların daraltılmış yüksek frekanslı versiyonu ile gerçekleştirilirken, frekans analizi aynı ana dalgacıklarının genişletilmiş, düşük frekanslı

versiyonu ile gerçekleştirilir. Başka bir deyişle, Fourier analizi bir sinyali çeşitli frekanslardaki sinüs dalgalarına ayırmaktan ibaretken, dalgacık analizi, bir sinyalin orijinal (veya ana) dalgacıkların kaydırılmış ve ölçeklenmiş versiyonlarına bölünmesidir.

Dalgacık dönüşümleri sonsuz bir kümeden oluşmaktadır. Farklı dalgacık aileleri, baz fonksiyonlarının uzayda ne kadar kompakt bir şekilde yerleştirildiği ve bu baz fonksiyonlarının ne kadar düzgün olduğu arasında farklı alternatifler meydana getirmektedir. Farklı özelliklere sahip ve yaygın olarak bilinen Haar dalgacıkları, Daubechies dalgacıkları, vb. gibi bazı dalgacıklar mevcuttur.

Son yıllarda, dalgacık dönüşümü veri madenciliğinde veri sıkıştırmanın önemli bir adım olduğu sinyal işleme metodolojilerinde geniş uygulama alanı kazanmıştır. Özellikle kısa zamanlı Fourier dönüşümünün aksine gerçek zamanlı uygulamalardan gelen düzensiz sinyallerin hesaplamalarında dalgacık dönüşümü daha hızlı ve güvenilirdir. Dalgacık dönüşümü bir ana dalgacık olarak bilinen temel fonksiyonunun ölçeklenmiş ve çevrilmiş versiyonları olan bir aile fonksiyonları ile orijinal bir sinyali ilişkilendirmenin sonucu olarak bir dizi katsayı vermektedir. En çekici özelliği, spektrumdaki zamansal bilgileri kısa zamanlı Fourier dönüşümünden daha iyi bir şekilde tanımlama yeteneğidir.

3.3.2. Dalgacık yapay sinir ağları (WNN)

Geleneksel olarak, ANN ve dalgacık dönüşümleri veri işlemede ayrı disiplinler olmuştur. Son zamanlarda dalgacık sinir ağları (WNN) ismiyle anılan yeni bir sinir ağı sınıfı ortaya çıkmıştır ve araştırma ve uygulamada gittikçe büyümektedir. Günümüzde, bu ağların potansiyeli mühendislik, kimya, ekonomi ve sosyal bilim uygulamalarında zengindir ve hala artmaktadır.

Dalgacık analizinin (WA) en büyük dezavantajı, küçük girdi boyutundaki uygulamalarla sınırlı olmasıdır. Bunun nedeni, giriş vektörünün boyutsallığı görece yüksek olduğunda, bir dalgacık temeli oluşturmanın hesaplama açısından pahalı olmasıdır [2]. Öte yandan, yapay sinir ağları (ANN), sürecin doğasına ilişkin çok az bilgi ve varsayım bulunmayan herhangi bir deterministik doğrusal olmayan sürece yaklaşma yeteneğine sahiptir. Bununla birlikte, klasik sigmoid yapay sinir ağının bir dizi dezavantajı bulunmaktadır. Tipik olarak, yapay sinir ağları ağırlıklarının başlangıç değerleri rastgele seçilmektedir. Rastgele ağırlıkların başlatılmasına genellikle uzun eğitim süreleri eşlik etmektedir. Ek olarak, aktivasyon fonksiyonu sigmoidal tipte olduğunda, eğitim algoritmasının yerel minimuma yakınsayacağı konusunda her zaman

önemli bir değişiklik bulunmaktadır. Son olarak, bir sigmoidal aktivasyon fonksiyonunun spesifik parametrelendirmesi ile optimal ağ mimarisi, yani model karmaşıklığı arasında teorik bir bağlantı yoktur.

[68]'de dalgacık ayrışımı açısından ileri beslemeli bir ANN'nin teorik bir formülasyonunu oluşturmanın mümkün olduğu gösterilmiştir. [69]'da dalgacık yapay sinir ağları, her yöntemle ilişkili yukarıda belirtilen zayıflıkları hafifletecek olan ileri beslemeli yapay sinir ağlarına bir alternatif olarak önerilmiştir. Dalgacık yapay sinir ağlarının arkasında yatan konsept dalgacık ve yapay sinir ağlarının birleşimi ile ilişkilidir. Bu başarılı teori sentezi dalgacık yapay sinir ağları ya da basit dalgacık ağları diye çağrılan yeni ağ sınıfı üretmiştir. Bu ağların yeniliği ise klasik yapay sinir ağlarında gizli katman aktivasyon fonksiyonu olarak dalgacık fonksiyonunu tercih etmesidir.

Dalgacık yapay sinir ağları modeli genellikle tek gizli katmanlı geriye dönük yapay sinir ağıdır ve aktivasyon fonksiyonları ortonormal dalgacık ailesinden çekilmektedir. WNN'nin bazı uygulamaları örüntü tanıma ve sıkıştırma, konuşma segmentasyonu, görüntü işleme (yüz tanıma), sanal gerçekliğin kullanımı için gerçek çevre karakterizasyonu, kaotik sinyalin tahmini ve klasik robot kontrol problemi ile ilişkilidir.

Bir dalgacık ağı genellikle üç katmanlı bir ağ biçimine sahiptir. İlk katman giriş katmanını, orta katman gizli katmanı ve üst katman çıktı katmanını temsil etmektedir. Üç katmanın bağlanma ve etkileşim kurma şekli, ağın yapısını tanımlamaktadır. Giriş katmanında, açıklayıcı değişkenler modele eklenmekte ve dalgacıklara dönüştürülmektedir. Gizli katman, dalgacıklardan veya gizli birimlerden oluşmaktadır. Gizli katmanda giriş değişkenleri, ana dalgacıkların genişletilmiş ve çevrilmiş bir versiyonuna dönüştürülmektedir. Son olarak, tüm dalgacıklar, hedef değerın yaklaşık y_p olan çıkış katmanında ağın çıktısını $\hat{y}_p$ üretmek için birleştirilir.

$g(x)$ dalgacık sinir ağı çıkış verileri, $\Psi(t)$ dalgacığa uygun olarak mevcuttur.

$$g(x) = \sum_{k=1}^T \omega_k \Psi\left(\frac{x - b_k}{\alpha_k}\right) \quad (3.29)$$

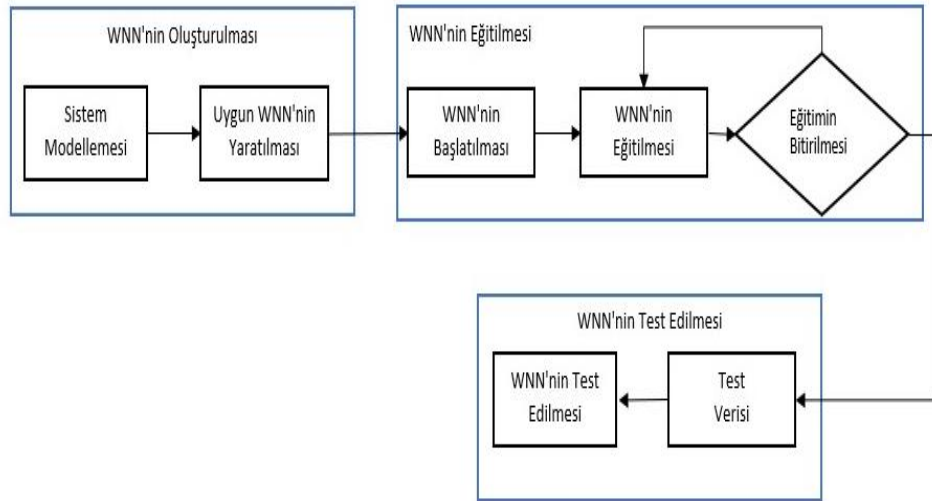
Denklem 3.29'da $g(x)$ uyumlanmış veridir, ω_k ağırlıklardır, α_k dalgacığın genişleme faktörüdür, b_k dalgacığın çeviri faktörüdür; T dalgacık sayısıdır. Ağın giriş örneklerinin sayısı m , giriş düğümü sayısının T olduğu varsayıldığında, ilk T örneğinin ve ilk n düğümün çıktısı şu şekilde ifade edilebilir:

$$g_t(x) = \sum_{k=1}^T \omega_k \Psi \left(\frac{\sum_{i=1}^m r_{ik} x_t(i) - b_k}{\alpha_k} \right) \quad (3.30)$$

Denklem 3.30'da k gizli katman sayısı, ω_k ağırlar arası ağırlıklardır. Ağın çalışması gereken parametreler ω_k ağırlıklar, genleşme faktörü α_k , öteleme faktörü b_k ve gizli katman N (ω_k, α_k, b_k için $k = 1, 2, \dots, N$) sayısıdır ve enerji fonksiyonunu minimize ederek parametrelerin optimizasyonunu gerçekleştirebilir. Minimum ortalama karesel hata (MMSE) fonksiyon biçimi aşağıda ifade edildiği gibidir:

$$E = \frac{1}{2} \sum_{i=1}^T [g_t(x) - y_t]^2 \quad (3.31)$$

Denklem 3.31'de, $g_t(x)$ ağın gerçek çıktısı, y_t ise ağın beklenen çıktısıdır. Dalgacık sinir ağının özel eğitim algoritması Şekil 3.2 ile gösterildiği gibidir [70]:



Şekil 3.2. Dalgacık sinir ağı algoritmasının akış şeması

Adım 1: Ağ parametrelerinin rastgele başlatılması ve ağın oluşturulması. (Dalgacık fonksiyonu dilatasyon faktörü α_k , öteleme faktörü b_k , ve ağ bağlantısının ağırlığı w_k , öğrenme oranı η , momentum katsayısı α , izin verilen hata oranı ε vb.).

Adım 2: Örnek sınıflandırma. (Numuneler, eğitim numunesi ve test numunesi ve ağın eğitimi için kullanılan eğitim numunesi, ağın tahmin doğruluğunu test etmek için kullanılan test numuneleri olarak ayırın.)

Adım 3: Model çıktısının hesaplanması. (Ağın gerçek çıktısını $g_t(x)$ hesaplamak için mevcut ağ parametrelerini kullanarak eğitim örneğini x_k ve karşılık gelen beklenen çıktıyı y_k girin, ardından gerçek çıktı ile istenen çıktı arasındaki E hatasını hesaplayın.)

Adım 4: Ağırlıkların gözden geçirilmesi. (E hatasına göre ağ ağırlıklarını ve dalgacık fonksiyonu parametrelerini gözden geçir, yeni parametre değerini yeni nesil öğrenme sürecine koyunuz. Bu nedenle, ağ tahmini değeri beklenen değere yaklaşıncaya kadar hatayı azaltmaya devam edin). Ağ parametrelerinin değişimi,

$$\begin{cases} \Delta\omega_k^{(i+1)} = -\eta \frac{\partial E}{\partial \omega_k^{(i)}} + \alpha \Delta\omega_k^{(i)} \\ \Delta\alpha_k^{(i+1)} = -\eta \frac{\partial E}{\partial \alpha_k^{(i)}} + \alpha \Delta\alpha_k^{(i)} \\ \Delta b_k^{(i+1)} = -\eta \frac{\partial E}{\partial b_k^{(i)}} + \alpha \Delta b_k^{(i)} \\ \Delta r_{ik}^{(i+1)} = -\eta \frac{\partial E}{\partial r_{ik}^{(i)}} + \alpha \Delta r_{ik}^{(i)} \end{cases} \quad (3.32)$$

Yenilenen ağ parametreleri Denklem 3.33'teki gibidir:

$$\begin{aligned} \omega_k^{(i+1)} &= \omega_k^{(i)} + \Delta\omega_k^{(i+1)} \\ \alpha_k^{(i+1)} &= \alpha_k^{(i)} + \Delta\alpha_k^{(i+1)} \\ b_k^{(i+1)} &= b_k^{(i)} + \Delta b_k^{(i+1)} \\ r_{ik}^{(i+1)} &= r_{ik}^{(i)} + \Delta r_{ik}^{(i+1)} \end{aligned} \quad (3.33)$$

Adım 5: Algoritmanın sonuna kadar gidip gitmediğinin tespit edilmesi. (E , izin verilen hatadan küçük olduğunda veya belirlenmiş yineleme sürelerine ulaştığında, öğrenme süreci sona erer; Aksi takdirde, hata geri yayılmaya gider, ardından $E = 0$ yapınız, 3. adıma dönün.)

Dalgacık yapay sinir ağının eğitiminde farklı öğrenme metotları uygulanmaktadır. Bazı çalışmalar stokastik gradyan algoritması, eşlenik gradyan metodu, geri beslemeli algoritma ve çoğunlukla örneklem teorisini tercih etmektedirler. Geri beslemeli yöntemlerle birleştirilmiş ortogonal en küçük kare geriye eliminasyon yöntemi de ağ yapısının seçiminde kullanılmaktadır. Dalgacık yapay sinir ağlarında başlangıç parametrelerini ayarlanması önemli hususlardan biridir. Dalgacık ağlarında, sigmoid işlevlerini kullanan sinir ağlarının aksine, dilatasyon ve çevirme parametrelerinin başlangıç değerlerini rastgele seçmek uygun olmayabilir [71]. Dalgacık, ortalama sıfır değerine ve yerleştirilmiş özelliklere sahip, etkin sınırlı süreli bir dalga biçimidir; bu

nedenle, rastgele bir başlatma, sıfır değerine sahip dalgacıklara yol açabilir. Rastgele başlatma ile gradyan inişi gibi eğitim algoritmaları verimsizdir [72], çünkü rastgele başlatma eğitimin hızını etkiler ve yerel minimum kayıp fonksiyonuna yol açabilir [73]. Ayrıca, sigmoid sinir ağlarında, kayıp fonksiyonunun bir minimizasyonu rastgele başlatma ile kopyalanabilse de ağırlıkların değerleri her seferinde değişecektir [74].

3.3.3. Bulanık dalgacık yapay sinir ağları (FWNN)

Son zamanlarda, bulanık sinir ağları için çeşitli modeller önerilmiştir. Ağların girdi, çıktı ve ağırlıklarındaki farka göre bulanık sinir ağları Tablo 3.2'de gösterildiği gibi üç kategoriye ayrılır [75].

Tablo 3.2. *Bulanık sinir ağlarının sınıflandırılması*

	Ağırlık	Girdiler	Çıktılar
Geleneksel NN	Reel sayı	Reel sayı	Reel sayı
FNN1	Reel sayı	Bulanık Sayı	Bulanık Sayı
FNN2	Bulanık Sayı	Reel sayı	Bulanık Sayı
FNN3	Bulanık Sayı	Bulanık Sayı	Bulanık Sayı

Tablo 3.2'de FNN1, bulanık sayı değerli doğrusal olmayan fonksiyonların ve doğrusal olmayan bulanık regresyon analizinin yaklaşımlayıcıları olarak kullanılabilir. FNN2, bulanık eğer-ise kurallarının interpolasyonunda, bulanık model vektörlerinin sınıflandırılmasında ve tamamlanmamış eğitim verilerinden öğrenmede uygulanabilmektedir. FNN3, eğer-ise kurallarının interpolasyonunda ve bulanık uzman sistemi modellemede uygulanabilmektedir.

Bulanık sinir ağı modellerine yönelik çeşitli yaklaşımlarda, bunların çoğu, gerçek girdileri ve/veya gerçek çıktıları ve/veya gerçek ağırlıkları bulanık sayılara genişleten çok katmanlı algılayıcı modelinin doğrudan bulanıklaştırılmasına dayanmaktadır ve öğrenme algoritmaları çoğunlukla BP öğrenme algoritmasının uzantısına dayanmaktadır. BP algoritmasının kendine has özellikleri olduğundan, algoritmanın yerel minimuma düşmesi kolaydır ve yakınsama oranı düşüktür. Böyle bir modelin yeteneği ve uygulaması büyük ölçüde sınırlı olmaktadır. Pek çok gerçek dünya problemlerinde, bu tür bir durum sıklıkla ortaya çıkabilmektedir. Bu tür bir durum genellikle girdi bulanık olmayan bir değer olduğunda (örneğin, bir maliyet veya yayılan güç verildiğinde) ve hedef gerçek yaşam deneyiminden elde edilen bulanık bir değer olduğunda ortaya çıkar. Bulanık

kontrolde, bulanık denetleyicinin önünde bir bulanıklaştırıcı vardır ve bulanıklaştırıcının işlevi, gerçek bir değeri bulanık bir değere dönüştürmektir. Bu tür durumlar genellikle girdi bulanık olmayan bir değer olduğunda ve hedef gerçek yaşam deneyiminden gelen bulanık bir değer olduğunda ortaya çıkmaktadır [75].

Literatürde yukarıdaki durumla başa çıkabilmek için çeşitli FWNN modelleri kurulmuştur. Tez çalışması kapsamında bulanık çıkarım sistemi (TSK modeli) WNN ile entegre edilerek FWNN' nin yapısı oluşturulmuştur. Bulanık sistemin çekirdeği, sistemin girdi-çıkı veri noktalarından oluşan bulanık bilgi tabanının dilbilimsel bulanık kurallara yorumlanmasıdır. TSK tipi bulanık EĞER-İSE (IF-THEN) kurallarının sonuç bölümleri, bir sabit veya bir fonksiyon ile temsil edilmektedir. Çoğu bulanık ve nöro-bulanık modelde, işlev, girdi değişkenlerinin doğrusal bir kombinasyonu artı bir sabit terimdir. TSK tipi sistemler, genel olarak, belirli sayıda kural kullanarak karmaşık süreçleri istenen doğrulukta modelleyemezler. Bu sistemlerin yerelleştirilebilirliği yoktur, bunun yerine sürecin genel özelliklerini modellerler ve yakınsamaları genellikle yavaştır. Bu TSK tipi bulanık ağların sonuç kısmı, tam eşleme yetenekleri sağlamaz ve karmaşık doğrusal olmayan süreçlerin modellenmesi durumunda, istenen doğruluğu elde etmek için çok sayıda kural gerektirebilir. Kural sayısının artması, ağın gizli katmanındaki nöron sayısının artmasına neden olmaktadır. Nöro-bulanık sistemin hesaplama gücünü geliştirmek için (doğrusal yerine) dalgacık fonksiyonlarının kullanılması önerilmektedir. Bu nedenle kullanılan kurallar aşağıdaki forma sahiptir [37]:

$$\text{Eğer } x_1, A_{11}; x_2, A_{12}; \dots; x_m, A_{1m} \text{ ise } y_1 = \sum_{i=1}^m w_{i1} (1 - z_{i1}^2) e^{-z_{i1}^2/2} \quad (3.34)$$

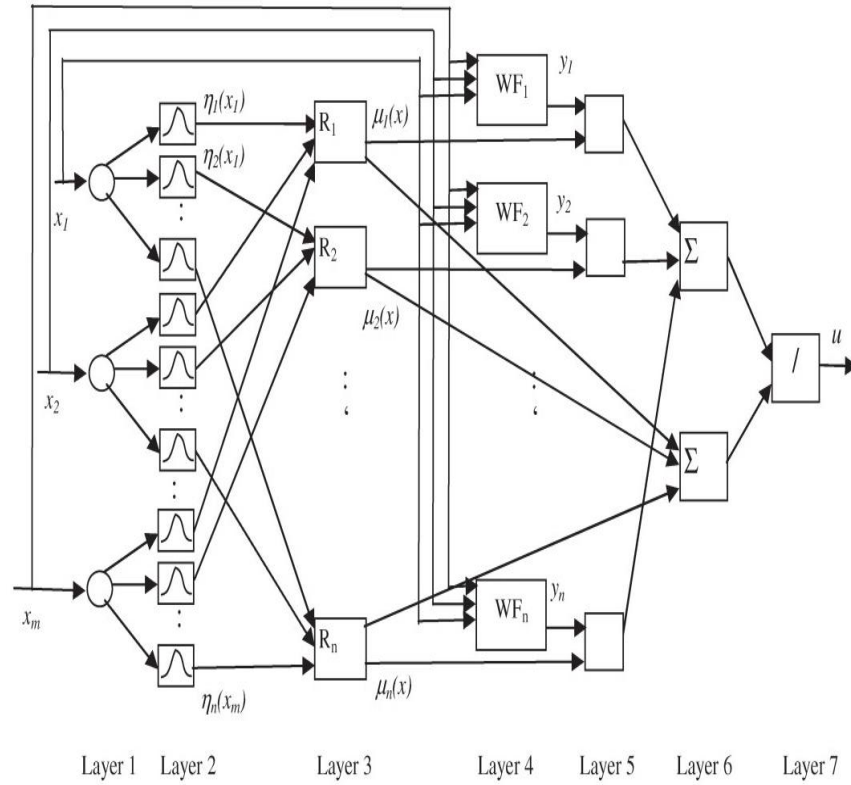
$$\text{Eğer } x_1, A_{21}; x_2, A_{22}; \dots; x_m, A_{2m} \text{ ise } y_2 = \sum_{i=1}^m w_{i2} (1 - z_{i2}^2) e^{-z_{i2}^2/2} \quad (3.34a)$$

$$\text{Eğer } x_1, A_{n1}; x_2, A_{n2}; \dots; x_m, A_{nm} \text{ ise } y_n = \sum_{i=1}^m w_{in} (1 - z_{in}^2) e^{-z_{in}^2/2} \quad (3.34b)$$

Denklem 34'te $x_1, x_2, \dots, x_m$ girdi değişkenlerini, $y_1, y_2, \dots, y_n$ çıktı değişkenlerini, A_{ij} , bir Gauss fonksiyonu olarak tanımlanan j'inci girdinin i'inci kuralı için bir üyelik fonksiyonunu temsil etmektedir. Çözüm kuralların bazı bölümleri Mexican Hat dalgacıklarını içermektedir.

Farklı genişleme ve öteleme değerlerine sahip dalgacıkların kullanılması, bu bulanık kurallar altında doğrusal olmayan modelin temel özelliklerini ve farklı

davranışları yakalamayı sağlamaktadır. EĞER-İSE kuralları tarafından tanımlanan uygun bulanık model, sonuç bölümlerinin dilatasyon ve öteleme parametrelerini ve öncül bölümlerin üyelik fonksiyonunun parametrelerini öğrenerek elde edilebilmektedir. Dalgacıkların kullanılması nedeniyle, FWNN'nin hesaplama gücü ve genelleme yeteneği iyileştirilmekte ve FWNN, lineer olmayan süreci istenen doğrulukla tanımlayabilmektedir.



Şekil 3.3. Bulanık dalgacık sinir ağının yapısı

FWNN'nin yapısı Şekil 3.3. ile gösterilmektedir. Yedi katmanı bulunmaktadır. İlk katmanda, düğüm sayısı giriş sinyali sayısına eşittir. Bu düğümler giriş sinyallerini dağıtmak için kullanılmaktadır. İkinci katmanda, her düğüm bir dilsel terime karşılık gelmektedir. Sisteme giren her giriş sinyali için, giriş değerinin ait olduğu bulanık kümeye üyelik derecesi hesaplanmaktadır. Dilsel terimleri tanımlamak için, Gauss üyesi Denklem 3.35 yardımı ile hesaplanmaktadır:

$$\eta_j(x_i) = e^{-\frac{(x_i - c_{ij})^2}{\sigma_{ij}^2}}, \quad i = 1, \dots, m; j = 1, \dots, n \quad (3.35)$$

Denklem 3.35'te m , harici giriş sinyallerinin sayısı, n , bulanık kuralların (üçüncü katmandaki gizli nöronlar) sayısıdır, c_{ij} ve σ_{ij} , sırasıyla i . giriş değişkeninin j . teriminin Gauss üyelik fonksiyonlarının merkezi ve genişliğidir ve $\eta_j(x_i)$ j . terim için i . giriş değişkeninin üyelik fonksiyonudur.

Üçüncü katmanda, düğüm sayısı $R_1, R_2, \dots, R_n$ kural sayısına karşılık gelmektedir. Her düğüm bir bulanık kuralı temsil etmektedir. Burada katmanın çıkış sinyallerinin değerlerini hesaplamak için AND (min) işlemi kullanılır. Π minimum işlemi temsil etmektedir ve aşağıda sunulmuştur:

$$\mu_j(x) = \prod_i \eta_j(x_i), \quad i = 1, \dots, m; j = 1, \dots, n \quad (3.36)$$

Bu $\mu_j(x)$ sinyalleri, bir sonraki katman olan ardıl katman için giriş sinyalleridir. Bu katman bir sonuç katmanıdır. n adet dalgacık yapay sinir ağlarını içermektedir. Beşinci katmanda, üçüncü katmanın çıkış sinyalleri WNN'lerin çıkış sinyalleri ile çarpılır. j . dalgacığın çıktısı denklem 3.37 ile hesaplanmaktadır:

$$y_j = w_j \Psi_j(z); \quad \Psi_j(z) = \sum_{i=1}^m \frac{1}{\sqrt{|a_{ij}|}} (1 - z_{ij}^2) e^{-z_{ij}^2/2}. \quad (3.37)$$

Denklem 3.37'de $z_{ij} = \frac{x_i - b_{ij}}{a_{ij}}$ şeklinde hesaplanmaktadır ve a_{ij} ve b_{ij} parametreleri ($i = 1, \dots, m$), ($j = 1, \dots, n$) WNN'lerin i . girdi ve j . çıktısının arasındaki dalgacık fonksiyonudur. Altıncı ve yedinci katmanlarda tüm ağın çıktısını hesaplamak için durulaştırma (defuzzification) yapılmaktadır. Bu katmanda, her WNN'nin FWNN'nin çıktısına katkısı Denklem 3.38 ile belirlenmektedir.

$$u = \sum_{j=1}^n \mu_j(x) y_j / \sum_{j=1}^n \mu_j(x). \quad (3.38)$$

Denklem 3.38'de y_j , WNN'lerin çıkış sinyalleridir. FWNN'nin çıkış sinyali hesaplandıktan sonra ağın eğitimi başlamaktadır. Başlangıçta, FWNN'nin parametreleri rastgele oluşturulmaktadır. Parametreler, ağın ikinci katmanındaki dilsel değerlerin üyelik fonksiyonu ve WNN'nin parametreleridir. Uygun bir FWNN modeli oluşturmak için parametrelerin eğitimi gerçekleştirilmektedir. Eğitim, öncül kısımda c_{ij} ve σ_{ij} ($i = 1, \dots, m; j = 1, \dots, n$) üyelik fonksiyonlarının parametre değerlerinin ve sonuç kısmında WNN'lerin $w_j(t)$, $a_{ij}(t)$, $b_{ij}(t)$ ($i = 1, \dots, m; j = 1, \dots, n$) parametre değerlerinin ayarlanmasını içermektedir. Literatürde FWNN algoritmasının eğitimi

kapsamında uyarlanabilir bir öğrenme oranıyla gradyan öğrenme önerilmektedir. Uyarlanabilir öğrenme oranı, yakınsamayı garanti etmekte ve ağıın öğrenmesini hızlandırmaktadır. Ek olarak, momentum öğrenme süreçlerini hızlandırmak için kullanılmaktadır [9]. İlk olarak ağıın çıktısında hata fonksiyonunun değeri Denklem 3.39 ile hesaplanmaktadır.

$$E = \frac{1}{2} \sum_{i=1}^o (u_i^d - u_i)^2 \quad (3.39)$$

Denklem 3.39’da O ağıın çıkış sinyallerinin sayısı, u_i^d ve u_i sırasıyla istenen ve ağıın mevcut çıkış değerleridir. WNN’nin $w(t)$, $a_{ij}(t)$, $b_{ij}(t)$ ($i = 1, \dots, m; j = 1, \dots, n$) parametreleri ve nöro-bulanık yapının üyelik fonksiyonu c_{ij} ve σ_{ij} ($i = 1, \dots, m; j = 1, \dots, n$) parametreleri Denklem 3.40 ve Denklem 3.41 kullanılarak ayarlanmaktadır:

$$\begin{aligned} w_j(t+1) &= w_j(t) + \gamma \frac{\partial E}{\partial w_j} + \lambda (w_j(t) - w_j(t-1)), \\ a_{ij}(t+1) &= a_{ij}(t) + \gamma \frac{\partial E}{\partial a_{ij}} + \lambda (a_{ij}(t) - a_{ij}(t-1)), \\ b_{ij}(t+1) &= b_{ij}(t) + \gamma \frac{\partial E}{\partial b_{ij}} + \lambda (b_{ij}(t) - b_{ij}(t-1)), \\ i &= 1, \dots, m; j = 1, 2, \dots, n \end{aligned} \quad (3.40)$$

$$\begin{aligned} c_{ij}(t+1) &= c_{ij}(t) + \gamma \frac{\partial E}{\partial c_{ij}}, \quad \sigma_{ij}(t+1) = \sigma_{ij}(t) + \gamma \frac{\partial E}{\partial \sigma_{ij}}, \\ i &= 1, \dots, m; j = 1, 2, \dots, n \end{aligned} \quad (3.41)$$

Burada γ öğrenme hızı, λ momentum, m ağıın giriş sinyallerinin sayısı (giriş nöronları) ve n bulanık kuralların (gizli nöronlar) sayısıdır. Denklem 3.40’te türevler Denklem 3.42 kullanılarak hesaplanmaktadır.

$$\begin{aligned} \frac{\partial E}{\partial w_j} &= \frac{\partial E}{\partial u} \frac{\partial u}{\partial y_j} \frac{\partial y_j}{\partial w_j} = \frac{(u(t) - u^d(t)) \cdot \mu_j \cdot \psi_j(z)}{\sum_{j=1}^n \mu_j}, \\ \frac{\partial E}{\partial a_{ij}} &= \frac{\partial E}{\partial u} \frac{\partial u}{\partial y_j} \frac{\partial y_j}{\partial \psi_j} \frac{\partial \psi_j}{\partial z_{ij}} \frac{\partial z_{ij}}{\partial a_{ij}} = \frac{\delta_j (3.5z_{ij}^2 - z_{ij}^4 - 0.5)e^{-z_{ij}^2/2}}{\left(\sqrt{a_{ij}^3}\right)}, \\ \frac{\partial E}{\partial b_{ij}} &= \frac{\partial E}{\partial u} \frac{\partial u}{\partial y_j} \frac{\partial y_j}{\partial \psi_j} \frac{\partial \psi_j}{\partial z_{ij}} \frac{\partial z_{ij}}{\partial b_{ij}} = \frac{\delta_j (3z_{ij} - z_{ij}^3)e^{-z_{ij}^2/2}}{\left(\sqrt{a_{ij}^3}\right)} \end{aligned} \quad (3.42)$$

Denklem 3.42’de yer alan δ_j ;

$$\delta_j = \frac{\partial E}{\partial u} \frac{\partial u}{\partial y_j} \frac{\partial y_j}{\partial \psi_j} = \left(u(t) - u^d(t) \right) \cdot \mu_j \cdot w_j / \sum_{j=1}^n \mu_j, \quad i = 1, \dots, m; \quad j = 1, \dots, n. \quad (3.43)$$

Denklem 3.41'deki türevler, Denklem 3.44 ve Denklem 3.45 kullanılarak belirlenmektedir.

$$\frac{\partial E}{\partial c_{ij}} = \sum_j \frac{\partial E}{\partial u} \frac{\partial u}{\partial \mu_j} \frac{\partial \mu_j}{\partial c_{ij}} \quad (3.44)$$

$$\frac{\partial E}{\partial \sigma_{ij}} = \sum_j \frac{\partial E}{\partial u} \frac{\partial u}{\partial \mu_j} \frac{\partial \mu_j}{\partial \sigma_{ij}} \quad (3.45)$$

burada,

$$\begin{aligned} \frac{\partial E}{\partial u} &= u(t) - u^d(t), \\ \frac{\partial u}{\partial \mu_j} &= \frac{y_j - u}{\sum_{j=1}^n \mu_j}, \quad i = 1, \dots, m; j = 1, \dots, n \end{aligned} \quad (3.46)$$

$$\begin{aligned} \frac{\partial \mu_j(x_i)}{\partial c_{ij}} &= \begin{cases} \mu_j(x_i) \frac{2(x_i - c_{ij})}{\sigma_{ij}^2} & \text{eğer } j \text{ düğümü } j \text{ kural düğümüne bağlıysa} \\ 0 & d.d \end{cases}, \\ \frac{\partial \mu_j(x_i)}{\partial \sigma_{ij}} &= \begin{cases} \mu_j(x_i) \frac{2(x_i - c_{ij})^2}{\sigma_{ij}^3} & \text{eğer } j \text{ düğümü } j \text{ kural düğümüne bağlıysa} \\ 0 & d.d \end{cases} \end{aligned} \quad (3.47)$$

(3.42) ve (3.47) numaralı denklemleri kullanarak (3.40) ve (3.41) numaralı denklemlerdeki türevler hesaplanmakta ve FWNN parametrelerinin düzeltilmesi yapılmaktadır.

3.4. Hibrit Algoritma Önerisi (Fuzzy Wavelet Kernel Extreme Learning Machine) (FWKELM)

Bulanık dalgacık model, her bulanık kuralın THEN bölümünün sonlu dalgacık fonksiyonları kümesinin doğrusal bir kombinasyonunu içerdiği bir FIS (Bulanık Çıkarım Sistemi) modelidir (TSK modeli). Bu bölümde, Mexican Hat dalgacık çekirdek fonksiyonu kullanılacak olup Mexican Hat dalgacık fonksiyonunun kabul edilebilir bir ELM çekirdeği olduğu ve genelleştirilmiş bir çekirdek tabanlı ELM'nin yapısına göre, bir çekirdek tabanlı ELM'nin bir FW modeli ile eşdeğerliği kanıtlanmaktadır.

$\psi(x)$ ana dalgacık olsun. a ve c sırasıyla dilatasyon ve öteleme parametresi ve $x, a, c \in \mathbb{R}$ 'dir. Eğer $x_i, x_j \in \mathbb{R}^D$ o zaman nokta çarpım dalgacık çekirdeği Denklem 3.48'deki gibi gösterilir.

$$K(x_i, x_j) = \prod_{d=1}^D \psi\left(\frac{x_{di} - c}{a}\right) \psi\left(\frac{x_{dj} - c}{a}\right). \quad (3.48)$$

Denklem 3.48, Translation-invariant kernel (öteleme ile değişmez çekirdek) teoremini karşılıyorsa, aşağıdaki ötelemeyle değişmez çekirdek işlevi elde edilebilir:

$$K(x_i, x_j) = \prod_{d=1}^D \psi\left(\frac{x_{di} - x_{dj}}{a}\right). \quad (3.49)$$

Ana dalgacık olarak Mexican Hat dalgacığı kullanıldığında Mexican Hat dalgacık çekirdek fonksiyonu türetilir.

$$\psi(x) = (1 - x^2) \exp\left(-\frac{x^2}{2}\right) \quad (3.50)$$

$$K(x_i, x_j) = \prod_{d=1}^D \left[1 - \left(\frac{x_{di} - x_{dj}}{a}\right)^2\right] \exp\left[-\frac{1}{2} \left(\frac{x_{di} - x_{dj}}{a}\right)^2\right] \quad (3.51)$$

Lemma 1: Bir tür translation-invariant çekirdek işlevi olarak, Mexican Hat dalgacık kabul edilebilir bir ELM çekirdeğidir.

Kanıt 1: İlk olarak, Mexican Hat dalgacığının Fourier dönüşümünün negatif olmadığı kanıtlanmalıdır.

$$F(w) = (2\pi)^{D/2} \prod_{d=1}^D \int_{-\infty}^{\infty} \exp(-jw_d x_d) \left(1 - \frac{x_d^2}{a^2}\right) \exp\left(-\frac{x_d^2}{2a^2}\right) dx_d \geq 0 \quad (3.52)$$

Mexican Hat dalgacığının Fourier dönüşümünün negatif olmadığını kanıtlamak için öncelikle Fourier dönüşümünün negatif olmadığının gösterilmesi gerekmektedir.

$$F(w) = (2\pi)^{D/2} \int \exp(-j(w \cdot x)) K(x) dx \quad (3.53)$$

$$K(x) = \prod_{i=1}^D h\left(\frac{x_i}{a}\right) = \prod_{i=1}^D \cos\left(\frac{1.75x_i}{a}\right) e^{-\left(\frac{\|x_i\|^2}{2a^2}\right)} dx \quad (3.54)$$

$K(x)$ denklem 3.54'te yerine yazıp integral çözülürse;

$$\begin{aligned} & \int \exp(-j(w \cdot x)) K(x) dx \\ &= \int_{-\infty}^{\infty} \exp(-j(w \cdot x)) \left(\prod_{i=1}^D \cos\left(\frac{1.75x_i}{a}\right) e^{-\left(\frac{\|x_i\|^2}{2a^2}\right)} dx \right) \end{aligned} \quad (3.55)$$

$$= \prod_{i=1}^D \int_{-\infty}^{\infty} \exp(-j(w \cdot x)) \cdot \left(\frac{\exp\left(\frac{j1.75x_i}{a}\right) + \exp\left(\frac{-j1.75x_i}{a}\right)}{2} \right) e^{-\left(\frac{\|x_i\|^2}{2a^2}\right)} dx_i \quad (3.55. a)$$

$$= \prod_{i=1}^D \frac{1}{2} \int_{-\infty}^{\infty} \left(\exp\left(-\frac{\|x_i\|^2}{2a^2} + \left(\frac{j1.75}{a} - jw_i a\right)x_i\right) + \exp\left(-\frac{\|x_i\|^2}{2a^2} - \left(\frac{j1.75}{a} - jw_i a\right)x_i\right) \right) dx_i \quad (3.55. b)$$

$$= \prod_{i=1}^D \frac{|a|\sqrt{2\pi}}{2} \left(\exp\left(-\frac{(1.75 - w_i a)^2}{2}\right) + \exp\left(-\frac{(1.75 + w_i a)^2}{2}\right) \right). \quad (3.55. c)$$

Denklem 3.55.c'yi denklem 3.53'te yerine koyarak Fourier dönüşümünü elde edilmektedir;

$$F[K](w) = \prod_{i=1}^D \left(\frac{|a|}{2} \right) \left(\exp\left(-\frac{(1.75 - w_i a)^2}{2}\right) + \exp\left(-\frac{(1.75 + w_i a)^2}{2}\right) \right). \quad (3.54)$$

Eğer $a \neq 0$ ise $F[K](w) \geq 0$ olur. Bir başka deyişle, Fourier dönüşümü negatif değildir. Bu kanıta paralel olarak Mexican hat dalgacığının Fourier dönüşümünün negatif olmadığını göstermek istendiğinde Denklem 3.52, bir dizi integral eşitsizliklere ayrıştırılabilir ve türev işlemi;

$$F(\omega) = (2\pi)^{D/2} a^D \int_{-\infty}^{\infty} \left\{ \exp\left[-jwa\left(\frac{x_d}{a}\right) - \frac{1}{2}\left(\frac{x_d}{a}\right)^2\right] - \left(\frac{x_d}{a}\right)^2 \exp\left[-jwa\left(\frac{x_d}{a}\right) - \frac{1}{2}\left(\frac{x_d}{a}\right)^2\right] \right\} d\left(\frac{x_d}{a}\right) \quad (3.55)$$

$$F(\omega) = (2\pi)^{D/2} a^D \prod_{d=1}^D \exp\left(-\frac{1}{2}a^2 w_d^2\right) \left\{ \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}(x_d + ja\omega_d)^2\right] dx_d - \int_{-\infty}^{\infty} x_d^2 \exp\left[-\frac{1}{2}(x_d + ja\omega_d)^2\right] dx_d \right\} \quad (3.56)$$

Denklem 3.56'da bulunan integral terim Denklem 3.57'teki gibi yazılabilir:

$$I = \prod_{d=1}^D (F_1^{(d)}(\omega) - F_2^{(d)}(\omega)), \quad (3.57)$$

Denklem 3.57'de bulunan I notasyonu, Denklem 3.56'daki integral terimdir.

$$F_1^{(d)}(\omega) = \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}(x_d + ja\omega_d)^2\right] dx_d, \quad (3.58)$$

$$F_2^{(d)}(\omega) = \int_{-\infty}^{\infty} x_d^2 \exp \left[-\frac{1}{2} (x_d + ja\omega_d)^2 \right] dx_d. \quad (3.59)$$

İntegralin öteleme değişmezliğine göre, kısmi integrasyon yöntemini kullanarak Denklem 3.60'ı elde etmek kolaydır.

$$\begin{aligned} F_1^{(d)}(\omega) &= (2\pi)^{1/2} \\ F_2^{(d)}(\omega) &= (2\pi)^{1/2} (1 - a^2 \omega^2). \end{aligned} \quad (3.60)$$

Denklem 3.60, Denklem 3.57'de yerine konulursa;

$$\prod_{d=1}^D (F_1^{(d)}(\omega) - F_2^{(d)}(\omega)) = (2\pi^{D/2}) a^{2D} \omega^{2D} \quad (3.61)$$

elde edilir. Daha sonra, Denklem 3.61, Denklem 3.52'de yerine konulursa,

$$F(\omega) = (2\pi)^D a^{3D} \exp \left(-\frac{a}{2} \sum_{d=1}^D \omega_d^2 \right) \prod_{d=1}^D \omega_d^2. \quad (3.62)$$

Fourier dönüşümü elde edilebilir.

Denklem 3.54'ten, $a \geq 0$, $F(\omega) \geq 0$ olduğu bilinmektedir. Bu nedenle, öteleme değişmez çekirdek teoremine göre, Mexican Hat dalgacığı kabul edilebilir bir ELM çekirdeğidir. İkili sınıflandırma problemleri için kernel fonksiyonu ile Mexican Hat dalgacığı ile değiştirilebilir.

$$f(x) = \text{sgn} \left\{ \begin{bmatrix} K(x, x_1) \\ \vdots \\ K(x, x_N) \end{bmatrix}^T \left(\frac{I}{C} + \Omega_{ELM} \right)^{-1} T \right\}. \quad (3.63)$$

Ayrıca bu sınıflandırıcı çok sınıflı sınıflandırma problemlerinde de kullanılabilir ve çıkış fonksiyonu;

$$f(x) = \arg \max \left\{ \begin{bmatrix} K(x, x_1) \\ \vdots \\ K(x, x_N) \end{bmatrix}^T \left(\frac{I}{C} + \Omega_{ELM} \right)^{-1} T \right\}. \quad (3.64)$$

şeklinde hesaplanmaktadır.

Lemma 2: Bir bulanık dalgacık (FW) modeli çekirdek tabanlı bir tek gizli katmanlı ileri beslemeli sinir ağına eş değerdir.

Kanıt 2: n girdili ve m çıktılı bir TSK modeli Denklem 3.65'teki gibi tanımlanmaktadır:

$$R^i: \text{Eğer } x_1, A_1^i; x_2, A_2^i; \dots; x_n, A_n^i \text{ ise } (y_1, \gamma_1^i), \dots, (y_m, \gamma_m^i) \quad (3.65)$$

Denklem 3.65'te;

- $x = [x_1, x_2, \dots, x_n]^T$ girdi vektörü,

- $y = [y_1, y_2, \dots, y_m]^T$ çıktı vektörü,
- A_j^i ($j = 1, 2, \dots, n; i = 1, 2, \dots, L$) i. kurala ait bulanık üyelik fonksiyonu
- L = bulanık kural sayısı,
- $\gamma_i = [\gamma_1^i, \gamma_2^i, \dots, \gamma_m^i]$ girdi değişkenlerin lineer kombinasyonudur ve

Denklem 3.66 yardımı ile hesaplanır:

$$\gamma_k^i = q_{ik,0} + q_{ik,1}x_1 + \dots + q_{ik,n}x_n \quad (3.66)$$

Bir FW modeli FIS modeli olduğunda 3.65 ve 3.66 denklemleri 3.67 ve 3.68 denklemleri gibi ifade edilmektedir:

$$R^i: \text{Eğer } x_1, A_1^i; x_2, A_2^i; \dots; x_n, A_n^i \text{ ise } O \text{ zaman } (y_1, s\beta_1^i), \dots, (y_m, s\beta_m^i) \quad (3.67)$$

$$s\beta_k^i = q_{ik,0} + q_{ik,1}\Psi_{i1}(x_1) + \dots + q_{ik,n}\Psi_{in}(x_n) \quad (3.68)$$

Burada $s\beta_k^i$ dalgacık fonksiyonların lineer kombinasyonudur. $s\beta_k^i$ 'de S standart FW yapısını, Ψ_j^i ($j = 1, 2, \dots, n; i = 1, 2, \dots, L$) i. kuralın j. girdinin dalgacık fonksiyonunu ve A_j^i ($j = 1, 2, \dots, n; i = 1, 2, \dots, L$) i. kurala ait bulanık üyelik fonksiyonunu göstermektedir. Denklem 3.68'de T-norm ve S-norm bulanık mantık operatörleri kullanılmaktadır. Böylece i. kuralın ateşleme gücü şu şekilde tanımlanır:

T-norm operatörü:

$$\mu_i(x; c_i, \alpha_i) = A_1^i(x_1, c_{1i}, \alpha_{1i}) \otimes \dots \otimes A_n^i(x_n, c_{ni}, \alpha_{ni}) \quad (3.69)$$

S-norm operatörü:

$$\mu_i(x; c_i, \alpha_i) = A_1^i(x_1, c_{1i}, \alpha_{1i}) \oplus \dots \oplus A_n^i(x_n, c_{ni}, \alpha_{ni}) \quad (3.70)$$

x girdisi için FW modelinin çıktısı her bir normalleştirilmiş kuralın ağırlıklandırılmış toplam çıktılarıdır ve Denklem 3.71 ile hesaplanmaktadır;

$$\hat{y} = \frac{\sum_{i=1}^L [\prod_{j=1}^n A_j^i(x_j)] y_i}{\sum_{i=1}^L [\prod_{j=1}^n A_j^i(x_j)]} = \sum_{i=1}^L s\beta_i G_{FW}(x; c_i, \alpha_i) \quad (3.71)$$

Burada $G_{FW}(x; c_i, \alpha_i) = \frac{\mu_i(x; c_i, \alpha_i)}{\sum_{i=1}^L \mu_i(x; c_i, \alpha_i)}$ R^i kuralı için normalleştirilmiş ateşleme gücüdür.

Tablo 3.3. FW ile çekirdek tabanlı ELM yapısının eşitliği

FW Model	Çekirdek Tabanlı ELM
$\sum_{i=1}^L s\beta_i G_{FW}(x; c_i, \alpha_i)$	$\sum_{i=1}^{\tilde{N}} \beta_i G(w_i, x_j + b)$ $\beta = H^T \left(HH^T + \frac{I}{C} \right)^{-1} T.$
Kural sayısı (L)	Gizli düğüm sayısı ($\tilde{N}$)
$G_{FW}(\cdot)$ her kural için normalleştirilmiş ateşleme gücü	$G(\cdot)$ gizli katman çıktı fonksiyonu

Denklem 3.71 ile Denklem 3.1 kıyaslandığında, çekirdek tabanlı ELM ile FW modelinin eşdeğer olduğu görülmüştür. Bir çekirdek tabanlı ELM' nin daha fazla ayrıntı içeren bir FW modeliyle eşdeğerliği Tablo 3.3'te gösterilmektedir. Sonuç olarak, ELM, gizli düğümlerin çıktı fonksiyonunun $G_{FW}(\cdot)$ ve $s\beta_k^i$ 'nin analitik olarak belirlenmesi gereken çıktı ağırlık vektörü olduğu FW modeline uygulanabilir. Denklem 3.1 ve Denklem 3.71'de verildiği gibi, bir ELM'deki gizli düğümler için çıktı fonksiyonları, FW modelinin üyelik fonksiyonlarına dayanmaktadır.

Bulanık dalgacık yapısı standart FW yapısı Denklem 3.67 ve Denklem 3.68'e dayanmaktadır, her bulanık kuralın THEN kısmında ise dalgacık üyelik fonksiyonu kullanılmaktadır. Bu yapıdaki doğrusal öğrenme parametrelerinin sayısının A yapısına eşittir [43]. Normalleştirilmiş Meksika şapka dalgacık fonksiyonu bir üyelik fonksiyonu (A_i^j) olarak kullanılabilir ve Denklem 3.72'deki gibi tanımlanır:

$$A_i^j = \frac{\left(1 - \left(\frac{x_j - b_{ji}}{d_{ji}}\right)^2\right) \exp\left(0.5 \left(\frac{x_j - b_{ji}}{d_{ji}}\right)^2\right) + \varepsilon}{1 + \varepsilon} \quad (3.72)$$

Denklem 3.72'de $\varepsilon = 0,4461$ olarak literatürde belirlenmiştir [43]. Önerilen FW yapısı şu şekilde tanımlanır:

$$R^i: \text{Eğer } x_1, A_1^i; x_2, A_2^i; \dots; x_n, A_n^i \text{ ise } (y_1, s\beta_1^i), \dots, (y_m, s\beta_m^i) \quad (3.73)$$

β_i^k 'nin yapısı, n'nin tek veya çift olmasına bağlıdır.

- Eğer n çift ise,

$$\beta_k^i = q_{ik,0} + q_{ik,1} \sum_{j=1}^2 \Psi_{ij}(x_j) + q_{ik,2} \sum_{j=3}^4 \Psi_{ij}(x_j) \dots + q_{ik, \lfloor \frac{n}{2} \rfloor} \sum_{j=n-1}^n \Psi_{ij}(x_j) \quad (3.74)$$

- Eğer n tek ise,

$$\begin{aligned} \beta_k^i = & q_{ik,0} + q_{ik,1} \sum_{j=1}^2 \Psi_{ij}(x_j) + q_{ik,2} \sum_{j=3}^4 \Psi_{ij}(x_j) \dots \\ & + q_{ik, \lfloor \frac{n}{2} \rfloor - 1} \sum_{j=n-2}^{n-1} \Psi_{ij}(x_j) + q_{ik, \lfloor \frac{n}{2} \rfloor} \Psi_{in}(x_n) \end{aligned} \quad (3.75)$$

Denklem 3.74 ve Denklem 3.75'te görüldüğü gibi, her bir bulanık kuralın (doğrusal öğrenme parametreleri) THEN kısmındaki lineer katsayıların sayısı $(\frac{n}{2} + 1)$ 'dir, burada n , n 'nin tavanını temsil eder, bu en küçük tam sayıdır veya n 'ye eşittir. Denklem 3.73'te, kuralların EĞER (IF) kısmında Sigmoid üyelik fonksiyonu Denklem 3.76'daki gibi kullanılmaktadır.

$$A(x) = sig(x; a, c) = \frac{1}{1 + \exp(-a(x - c))} \quad (3.76)$$

Denklem 3.76'da c merkez noktası, $x = c$ geçiş noktasında a eğim noktasıdır.

FW modelin sonuç kısmında, Mexican hat dalgacık fonksiyonu kullanılır ve denklem 3.77'teki gibi ifade edilir:

$$\psi(x) = (1 - x^2) \exp\left(-\frac{x^2}{2}\right) \quad (3.77)$$

3.4.1. FWKELM algoritmasının matematiksel yapısı

Keyfi (x_i, t_i) eğitim veri seti verilsin. Burada $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T$ girdi vektörü, $t_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T$ çıktı vektörü, L = bulanık kural sayısıdır. 3 parametre içeren (β_i, c_i, α_i) FW matematiksel modeli Denklem 3.78) ile tanımlanır:

$$f_L(x_j) = \sum_{i=1}^L \beta_i G_{FW}(x; c_i, \alpha_i) = t_j, \quad j = 1, 2, \dots, N. \quad (3.78)$$

$$\beta_i = \Psi_{ie}^T(x_j) q_{iFW} \quad (3.79)$$

Denklem 3.79’da β_i sonlu dalgacık kümesinin lineer kombinasyonuna dayanmaktadır. (3.74) ve (3.75)’e göre, n (tek veya çift)’e bağlı olarak $\psi_{ie}(x)$ vektörü (3.80) ve (3.81) denklemleri ile tanımlanabilir:

Eğer n çift ise;

$$\psi_{ie}(x) = \left[1, \sum_{j=1}^2 \psi_{ij}(x_j), \dots, \sum_{j=n-1}^n \psi_{ij}(x_j) \right]^T \quad (3.80)$$

Eğer n tek ise;

$$\psi_{ie}(x) = \left[1, \sum_{j=1}^2 \psi_{ij}(x_j), \dots, \sum_{j=n-2}^{n-1} \psi_{ij}(x_j), \psi_{in}(x_n) \right]^T \quad (3.81)$$

Denklem 3.79’da i. kural için q_{iFW} ($i = 1, \dots, L$) Denklem 3.83 ile ifade edilir.

$$q_{iFW} = \begin{bmatrix} q_{i1,0} & \dots & q_{im,0} \\ \vdots & \vdots & \vdots \\ q_{i1, \lfloor \frac{n}{2} \rfloor} & \dots & q_{im, \lfloor \frac{n}{2} \rfloor} \end{bmatrix}_{(\lfloor \frac{n}{2} \rfloor + 1) \times m} \quad (3.83)$$

Aslında, q_{iFW} matrisinin k. sütunu Denklem 3.73 veya Denklem 3.74’teki doğrusal parametreleri içermektedir. Sonuç olarak, önerilen FW yapısında yukarıdaki matrisin boyutu $\left(\left\lfloor \frac{n}{2} \right\rfloor + 1\right) \times m$ ’dir.

Denklem (3.71), (3.78), (3.79)’ye göre; önerilen FW modeli için toplam çıktı şu şekilde yazılabilir:

$$f_L(x_j) = \sum_{i=1}^L \psi_{ie}(x_j) q_{iFW} G_{FW}(x; c_i, \alpha_i) = t_j, \quad j = 1, 2, \dots, N. \quad (3.84)$$

Yukarıdaki N denklemleri kompakt olarak şu şekilde yazılabilir:

$$HQ = T. \quad (3.85)$$

$$H(c_1, \dots, c_L, a_1, \dots, a_L; b_1, \dots, b_L, d_1, \dots, d_L; x_1, \dots, x_N) \\ = [\psi_{ie}^T(x_j; b_1, d_1) G_{FW}(x_j; c_1, a_1), \dots, \psi_{ie}^T(x_j; b_L, d_L) G_{FW}(x_j; c_L, a_L)] \quad (3.86)$$

$$Q = \begin{bmatrix} q_{1FW} \\ \vdots \\ q_{LFW} \end{bmatrix}_{L(\lfloor \frac{n}{2} \rfloor + 1) \times m} \quad \text{ve} \quad T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N \times m} \quad (3.87)$$

Denklem 3.86’da H gizli birim matrisidir ve bulanık dalgacık modeli için parametre matrisidir. Denklem 3.87’de Q ’nun elemanları lineer parametrelerdir.

Denklem 3.86'da $(c_i = [c_{i1}, \dots, c_{in}]^T, \alpha_i = [\alpha_{i1}, \dots, \alpha_{in}]^T, i = 1, \dots, L)$ Sigmoid üyelik fonksiyonunun parametreleridir ve $(b_i = [b_{i1}, \dots, b_{in}]^T, d_i = [d_{i1}, \dots, d_{in}]^T, i = 1, \dots, L)$ dalgacık katsayılarıdır.

Bir çekirdek tabanlı ELM ve bir FW modeli arasındaki eşdeğerliğin kanıtına göre, KELM önerilen FW modeline uygulanabilir. Ortaya çıkan algoritmaya FWKELM denir. Denklem 86'da üyelik fonksiyonlarının ve dalgacık fonksiyonlarının $(c_i, \alpha_i$ ve b_i, d_i $i = 1, \dots, L)$ parametreleri dahil tüm doğrusal olmayan parametreler rastgele oluşturulmaktadır. Lineer parametreler (Q elemanları), tek geçişli FWKELM algoritması ile KELM'ye göre hesaplanmaktadır. Prosedür aşağıdaki gibi özetlenebilir:

Keyfi (x_i, t_i) eğitim veri seti verilsin. Burada $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T$ girdi vektörü, L = bulanık kural sayısı ve $t_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T$ çıktı vektörü olsun.

- Adım 1: Üyelik fonksiyonu ve dalgacık katsayılarının rastgele atanan parametreleri atayınız. $(c_i = [c_{i1}, \dots, c_{in}]^T, \alpha_i =, b_i = [b_{i1}, \dots, b_{in}]^T, d_i = [d_{i1}, \dots, d_{in}]^T, i = 1, \dots, L)$
- Adım 2: Denklem 3.86'ya göre H gizli matrisini hesaplayınız.
- Adım 3: (3.88)'e göre, Q matrisine karar verin.

$$Q = H^+T. \quad (3.88)$$

4. BULGULAR

Bu tez çalışması kapsamında önerilen hibrit algoritmanın sınıflandırma başarısını ölçebilmek amacıyla;

- Küçük/büyük boyutlu geleneksel sınıflandırma veri setleri,
- Literatürde yer alan yüz ve yüz ifadeleri veri setleri ve
- Orijinal yüz ifadesi veri seti

üzerinde algoritma çalıştırılmıştır.

4.1. Uygulama 1

Bu uygulamada, FWKELM'in performansı küçük ve büyük boyutlu dengesiz veri setlerinde ölçülmektedir. Ölçüm sonucu geleneksel ELM, KELM, WNN, MHWKELM sınıflandırıcısı sonuçları ile karşılaştırılmaktadır. Bahsi geçen tüm algoritmalar MATLAB R2021b yazılımı üzerinde çalıştırılmaktadır. İşletim ortamı Core-i7, 1.80 GHz CPU, 8 GB RAM'dir. Deneyde kullanılan veri setleri UCI veri tabanından [http-1] alınmaktadır. Bunlar sırasıyla Abalone, Auto MPG, Wine, Wine Quality, Glass, Breast Cancer, Zoo ve Page Blocks'dur. Bu 8 veri setinin temel özellikleri Tablo 4.1'de gösterilmektedir.

Tablo 4.1. UCI veri tabanından alınan veri setlerine ilişkin bilgiler

Veri seti adı	Gözlem Sayısı		Öznitelik Sayısı	Yanıt Değişkeninin Kategori Sayısı
	Eğitim	Test		
Abalone	3342	835	8	3
Auto MPG	318	80	16	2
Wine	117	61	13	3
Wine Quality	5198	1299	11	7
Glass	128	86	9	2
Breast Cancer	172	114	9	2
Zoo	61	40	16	7
Page Blocks	4378	1095	5	10

Tablo 4.1'de verilen veri setlerini kullanarak beş algoritmanın çalışma süresini ve test doğruluğunun ölçülmesi hedeflenmektedir. Tüm veri setleri incelendiğinde ve girdi değerlerine bakıldığında bu değerlerin *max* ve *min* değerleri birbirinden oldukça

farklıdır. Dağılımı normalleştirmek ve başarı oranını artırmak için öncelikle normalizasyon uygulanmalıdır. Normalleştirme için özellik ölçeklendirme yöntemi kullanılmaktadır. Normalleştirme Denklem 4.1 aracılığıyla ifade edilmektedir;

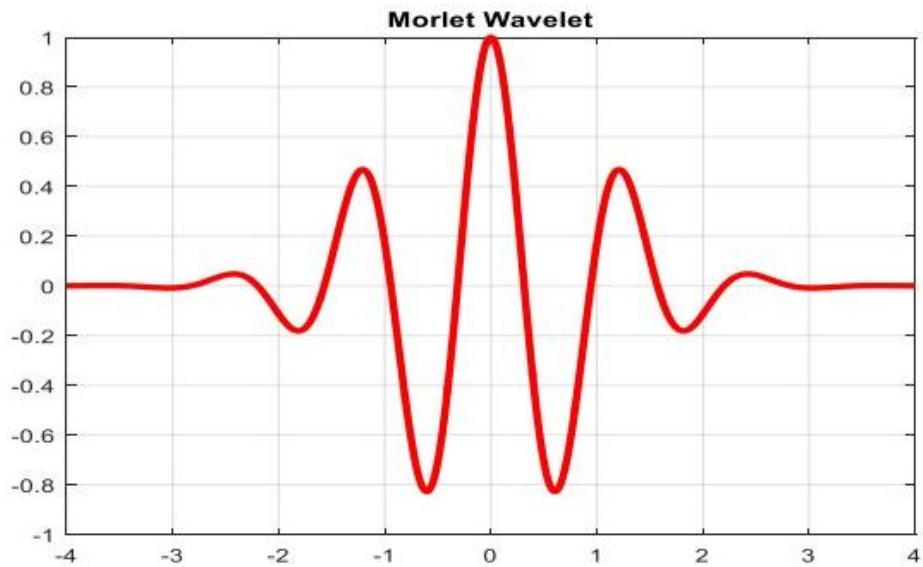
$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (4.1)$$

Normalleştirme işleminden sonra veri setlerinden rasgele eğitim ve test kümeleri üretilmektedir. Wine, Glass, Zoo ve Breast Cancer veri seti hariç tüm veri setleri için test ve eğitim kümeleri 80:20 oranı olarak oluşturulmaktadır. Wine, Glass, Zoo ve Breast Cancer veri setlerinde nispeten boyut daha küçük olduğundan test ve eğitim kümeleri 70:30 oranı olarak oluşturulmuştur. Her veri seti, her algoritma tarafından 1000 kez tekrar edilmektedir. Her seferinde, eğitim örneği toplam örnekten rastgele seçilmiştir. Aşırı öğrenme makinesi algoritmasında ve önerilen hibrit algorithmada genelleştirilmiş Moore-Penrose tersi “*pinv*” kodu yardımı ile hesaplanmıştır. Çekirdek tabanlı aşırı öğrenme makinesi algoritması için Gauss kerneli tercih edilmiştir. Dalgacık yapay sinir ağı algoritması için gizli katman aktivasyon fonksiyonu olarak Morlet fonksiyonu ele alınmaktadır. Morlet fonksiyonuna yapısı ise Şekil 4.1. ile gösterilmektedir. Ayrıca Morlet fonksiyonu;

function[çıktı] = morlet(x)

çıktı = cos(1.75 * x).* exp(-(x.^2)/2);

kod yapısı ile MATLAB platformunda kullanılmıştır.



Şekil 4.1. Morlet fonksiyonunun yapısı

Önerilen FWKELM yapısında yer alan “EĞER-İSE” kurallının EĞER kısmında Sigmoid aktivasyon fonksiyonu kullanılırken İSE kısmında Mexican Hat dalgacık fonksiyonu yer almaktadır. Sigmoid aktivasyon fonksiyonu, Matlab platformunda aşağıdaki şekilde çağırılmıştır.

```
function j = sigmoid(z)
j = 1/(1 + exp(-z));
end
```

Ayrıca Mexican Hat dalgacığı;

```
function w = wavelet( x, a, b )
z = (x - b) / a;
z = z^2;
w = (1 - z) * exp(-z/2) / sqrt(abs(a));
```

Kod yapısı MATLAB platformunda tanımlanmıştır. Algoritmelerde başlangıç değerleri (öğrenme oranı parametresi (η), momentum değeri (α), vb.) kullanıcı giriřli tanımlanmıştır. Bu parametrelerin kullanıcı giriřli tanımlanmasının sebebi önerilen yaklaşımın rastgelelik durumuna göre nasıl performans sergileyeceğinin araştırılmak istenmesidir. Algoritmalar çalıştırıldıktan sonra elde edilen sonuçlar Tablo 4.2 ile sunulmaktadır.

Tablo 4.2. Sınıflandırma problemleri için performans karşılaştırması

Veri setleri	Algoritmalar	Ortalama Başarı Oranı (%)	Standart Sapma	Çalışma Süresi (sn)	Nöron Sayısı/Kural Sayısı
Abalone	ELM	71,25	0,04	0,59	10
	KELM	78,42	0,26	0,65	21
	WNN	72,42	0,91	2,49	300
	MHWKELM	79,75	0,44	2,25	100
	FWKELM	81,81	0,0045	5,89	8
Auto MPG	ELM	73,55	3,24	0,05	8
	KELM	74,88	2,89	0,07	20
	WNN	74,77	2,25	0,67	256
	MHWKELM	82,24	1,22	0,77	8
	FWKELM	82,82	0,51	0,98	6

Tablo 4.2. (Devam) Sınıflandırma problemleri için performans karşılaştırması

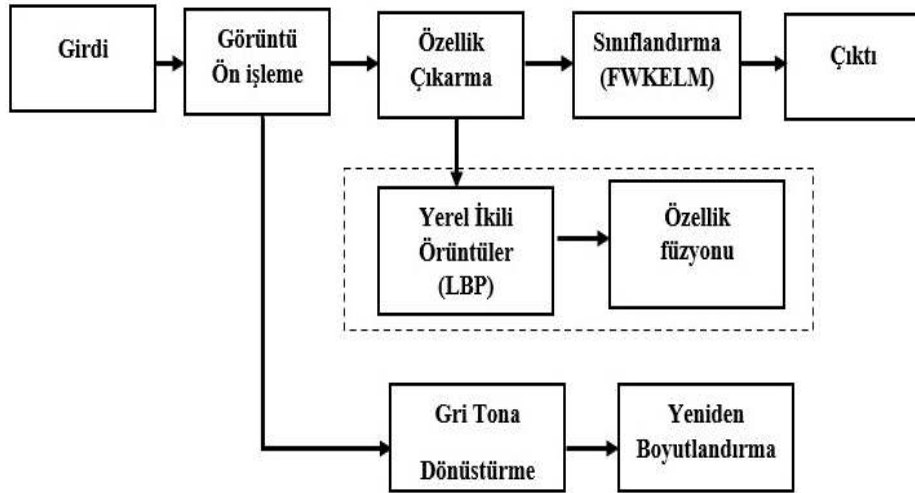
Veri Setleri	Algoritmalar	Ortalama Başarı Oranı (%)	Standart Sapma	Çalışma Süresi (sn)	Nöron Sayısı/Kural Sayısı
Wine	ELM	68,18	2,93	0,05	12
	KELM	83,64	1,28	0,21	18
	WNN	80,28	1,13	0,32	42
	MHWKELM	88,25	0,81	0,35	18
	FWKELM	92,82	0,37	1,01	15
Wine Quality	ELM	45,7	0,87	1,19	50
	KELM	52,14	0,28	1,13	50
	WNN	45,26	0,52	1,27	74
	MHWKELM	49,69	0,35	3,12	55
	FWKELM	50,58	1,17	5,28	30
Glass	ELM	75,12	0,79	0,024	18
	KELM	73,24	0,03	0,09	22
	WNN	72,07	1,00	0,17	38
	MHWKELM	75,76	0,06	0,67	24
	FWKELM	77,23	0,02	0,98	12
Breast Cancer	ELM	66,84	0,29	1,37	11
	KELM	69,01	0,41	1,75	20
	WNN	65,7	0,33	1,74	256
	MHWKELM	66,78	0,72	2,1	20
	FWKELM	71,85	0,02	3,74	9
Zoo	ELM	90,13	0,14	0,059	15
	KELM	90,07	1,21	0,08	24
	WNN	89,86	0,81	0,066	52
	MHWKELM	90,05	1,53	0,062	24
	FWKELM	91,26	0,12	1,135	10
Page blocks	ELM	95,17	0,21	1,82	15
	KELM	95,19	0,036	1,88	10
	WNN	94,32	0,455	1,01	14
	MHWKELM	94,55	0,027	3,5	14
	FWKELM	96,16	0,004	10,49	7

Tablo 4.2 incelendiğinde hemen hemen tüm veri setleri için önerilen algoritmanın sınıflandırma başarısı bünyesinde barındırdığı alitmadan daha başarılı olduğu

belirlenmiştir. Aynı zamanda elde edilen sınıflar yine birbirine benzeyen yapıda yani bir başka ifade ile homojendir. Çalışma süresi parametresine göre önerilen algoritma, daha uzun sürelerde döngüsünü tamamlamıştır. Önerilen hibrit algoritması bulanık kural tabanlı ve aynı zamanda çekirdek yapısını işlettiğinden diğer algoritmalara göre öğrenme hızı yavaş kaldığı söylenebilir. Nöron/kural sayısı parametresi incelendiğinde önerilen algoritmanın ağ yapısının karmaşıklığını diğer algoritmalara göre daha da azalttığı gözlemlenmiştir.

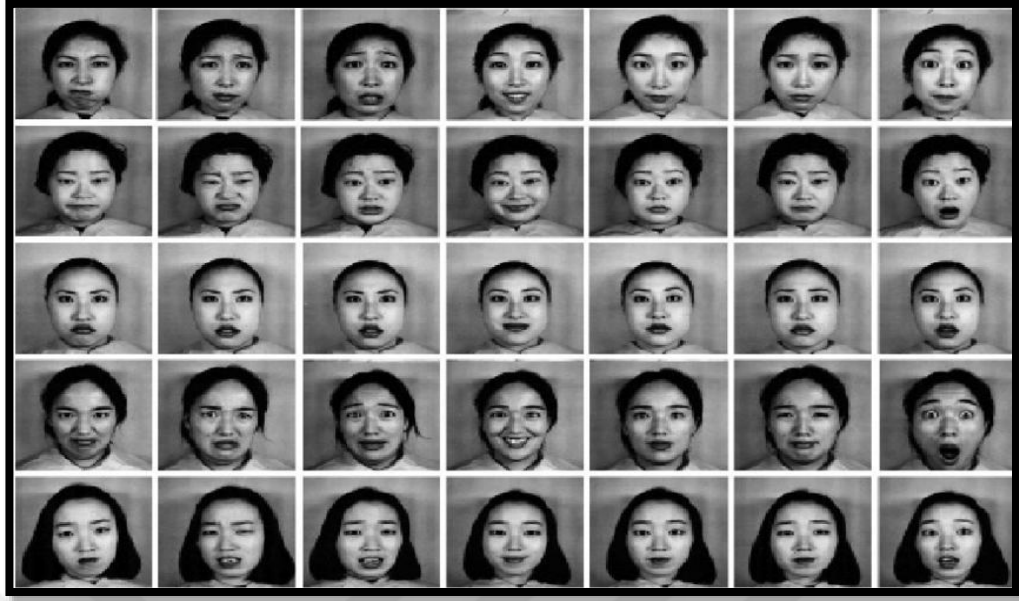
4.2. Uygulama 2

İkinci uygulamada, literatürde var olan ve yüz ifadesi sınıflandırma problemlerinde sıklıkla kullanılan iki farklı yüz ifadelerini barındıran veri tabanlarından yararlanarak deneyler yapılmıştır. İlk veri tabanı olan JAFFE veri tabanı, her biri 7 temel yüz ifadesine sahip 10 farklı kadının 213 görüntüsünü içermektedir ki bu ifadeler; öfke, iğrenme, korku, mutlu, üzgün, sürpriz ve nötr olarak sıralanabilmektedir. JAFFE veri tabanındaki görüntüler 70:30 oranında (140 görüntü eğitim seti ve kalan 73 görüntü ise test veri seti) ayarlanmıştır. Yüz ifadesi sınıflandırma problemlerinde kullanılmak üzere önerilen algoritmanın iş akışı Şekil 4.2’de gösterilmiştir.



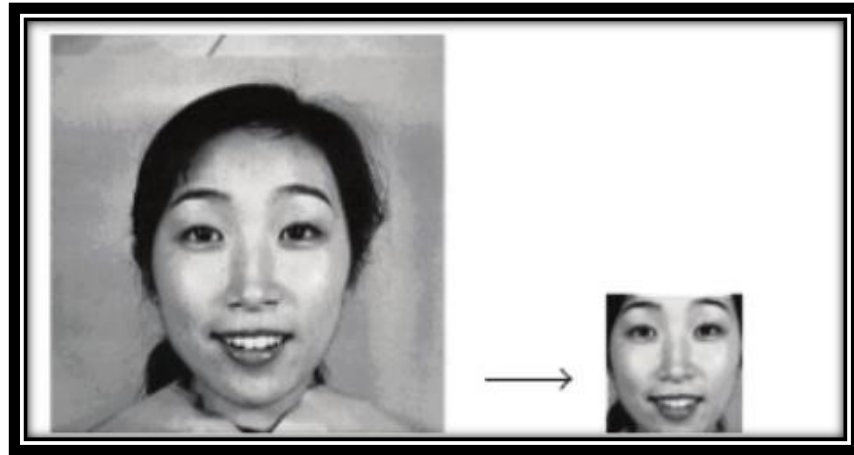
Şekil 4.2. Yüz ifadesi için akış diyagramı

Görüntülerin çözünürlüğü 256×256 pikseldir. JAFFE veri tabanından alınan görüntü örnekleri Şekil 4.3'te gösterilmektedir.



Şekil 4.3. JAFFE veri seti görüntü örnekleri

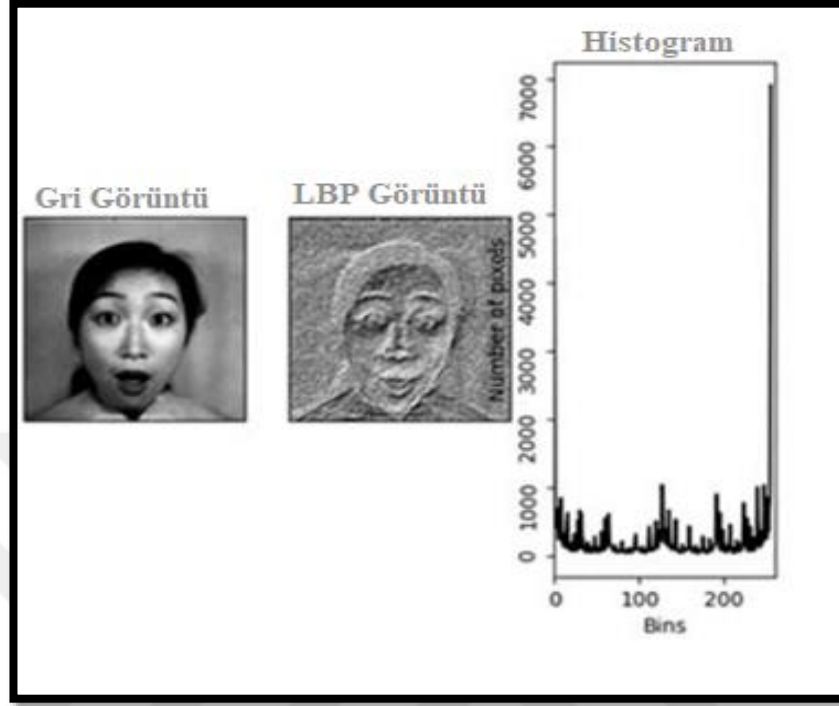
Ön işleme adımında görüntü yoğunlukları normalleştirilmiştir ve sadece duyguyu ifade eden bir yüzden oluşan görüntü ön işleme tabi tutulduktan sonra yeniden boyutlandırılmaktadır (Şekil 4.4).



Şekil 4.4. Orijinal görüntüler için ön işleme gösterimi

Yüz ifadesini tanımanın diğer adımı, yüz özelliklerini bir görüntüden çıkarmaktır. Bu yerel tanımlayıcılar, özellik çıkarma sürecinin temel yapı taşlarıdır. Popüler tanımlayıcılardan biri olan Yerel İkili Örüntüler (Modeller) (LBP) kullanılarak bu görüntülerden özellikler çıkarılmıştır. LBP operatörü, gri seviyeden bağımsız bir doku

ölçümü yöntemidir. LBP operatörü görüntünün her pikseli için bir etiket oluşturmaktadır ve bu etiketler birler ve sıfırlardan oluşmaktadır.



Şekil 4.5. Yerel ikili örüntü işleme örneği

Çıktı görüntüsünde bir başka ifade ile Şekil 4.5'te görüldüğü üzere, LBP dönüştürülmüş görüntü, gelişmiş yüz işareti sınırlarına sahip olmaktadır. Ek olarak, yüzdeki duygu tespiti için kaşların, gözlerin ve ağzın piksel konumlarındaki varyasyonları belirli bir referans noktasından anlamak çok önemlidir.

Yüz ifade tanıma probleminde sınıflandırma aşaması için önerilen algoritmanın mimarisi (ağ yapısı) oluşturulması gereklidir. Bu doğrultuda, ağın gizli katmandaki nöron sayısı 5, öğrenme oranı ise 0,25 olarak belirlenmiştir. Ayrıca minimum hata 0.001 olarak belirlenmiştir. JAFFE veri tabanındaki görüntü sayısı sınırlı olduğundan, ortalama sınıflandırma oranını elde etmek için deney 10 kat çapraz doğrulama ile gerçekleştirilmiştir. Veri seti eğitildikten sonra, test görüntüsü önerilen hibrit algoritma sınıflandırıcısı kullanılarak sınıflandırılmıştır.

Karışıklık matrisi, bir sınıflandırma algoritmasının performansını özetlemek için kullanılan bir tekniktir. Her sınıfta eşit olmayan sayıda gözlem varsa veya veri setinde ikiden fazla sınıf varsa, tek başına sınıflandırma doğruluğu yanıltıcı olabilir. Bir karışıklık matrisinin hesaplanması, sınıflandırma modelinin neyi doğru yaptığı ve ne tür hatalar

yaptığı konusunda daha iyi bir fikir verebilmektedir. Bu deney için elde edilen karışıklık matrisi Tablo 4.3. ile sunulmuştur.

Tablo 4.3. *LBP+FWKELM sınıflandırıcısı kullanılarak elde edilen karışıklık matrisi*

	Mutlu	Üzgün	Öfke	Korku	İğrenme	Şaşkın	Nötr
Mutlu	96,13	1,29	0	0	0,65	0	1,93
Üzgün	2,58	94,96	0	0,55	0,62	0	1,29
Öfke	0	0	94,0	0,65	2,67	0	2,67
Korku	0	2,5	0,63	91,89	1,88	3,1	0
İğrenme	0	0,69	0	2,76	96,55	0	0
Şaşkın	0	0,37	0	0,67	0	97,63	1,33
Nötr	2,76	5,27	0,62	0,67	0	0	90,67

Mutluluk ifadesi için sınıflandırıcının performansı %96,13'tür. Korku ifadesi için performans %91,89 ve şaşkın ifadesi için ise %97,63'tür. Önerilen algoritma tüm ifadeler için %90 üzerinde doğru sınıflandırma yapmış olduğu belirlenmiştir. Ortalama sınıflandırma doğruluğu %94,55 olarak hesaplanmıştır.

Paralel olarak, aynı veri seti üzerinde yüz özelliklerini çıkarma adımıyla farklı yöntemden yararlanılarak deney tekrar edilmiştir. Böylelikle önerilen algoritmanın farklı yüz özellik çıkarma algoritmalarında nasıl performans sergilediği gözlemlenmek istenmiştir. Deney kapsamında yüz özelliklerini çıkarma aşamasında yönlendirilmiş gradyanların histogramları (HOG) kullanılmıştır. Sınıflandırma aşamasında ise önerilen algoritma FWKELM algoritması uygulanmış ve karışıklık matrisi elde edilmiştir. Bu deney için elde edilen karışıklık matrisi ise Tablo 4.4 ile sunulmuştur.

Tablo 4.4. *HOG+FWKELM sınıflandırıcısı kullanılarak elde edilen karışıklık matrisi*

	Mutlu	Üzgün	Öfke	Korku	İğrenme	Şaşkın	Nötr
Mutlu	97,78	1,11	0	0	0	0	1,11
Üzgün	0	95,65	3,24	0	0	1,11	0
Öfke	0	0	97,78	1,11	1,11	0	0
Korku	0	1,11	0	98,89	0	0	0
İğrenme	0	0	0	0	84,22	15,78	0
Şaşkın	2,22	0	0	4,44	0	90,01	3,33
Nötr	3,33	0	1,11	0	0	0	95,56

Tablo 4.4 incelendiğinde, Mutluluk ifadesi için sınıflandırıcının performansı %97,78, korku ifadesi için performans %98,89 ve iğrenme ifadesi için ise %84,22'dir.

Yönlendirilmiş gradyanların histogramları ile çalıştırılan algoritma iğrenme ifadesi hariç diğer ifadeler için %90 üzerinde doğru sınıflandırma yapmış olduğu belirlenmiştir. Ortalama sınıflandırma doğruluğu %94,27 olarak hesaplanmıştır.

Elde edilen sonuçlar aynı veri setini kullanan literatürde mevcut diğer yöntemlerle de karşılaştırılmış ve sonuçlar Tablo 4.5 ile sunulmuştur.

Tablo 4.5. Farklı yaklaşımların sınıflandırma oranının karşılaştırmalı sonuçları

Yöntemler	Ortalama Tanıma Oranları
LBP+FWKELM	%94,55
HOG+FWKELM	%94,27
Fisherface + BPNN [76]	%86,85
Morphological image processing+ ELM [30]	%90,16
HOG+ELM (bagging method) [77]	%94,37
PCA-FLD+ WNN [41]	%90,00

Tablo 4.5 incelendiğinde, önerilen algoritmanın ve öznetelik tespitinde kullanılan yerel ikili örüntü (LBP) algoritmasının birlikte çalışarak daha verimli sonuçlar elde ettiği görülmüştür.

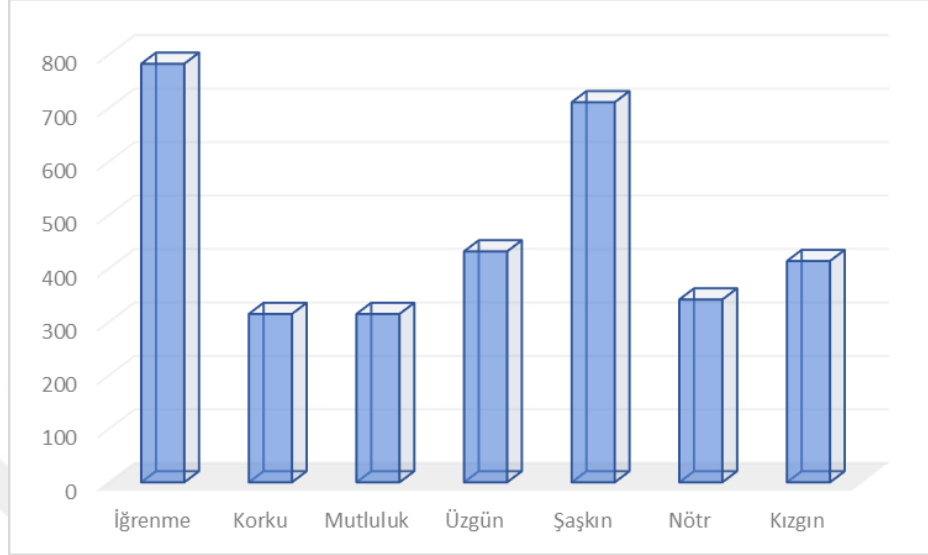
İkinci veri seti olarak, Kaggle tarafından 2013 yılında yayınlanan Facial Expression Recognition veri seti (FER2013) kullanılmıştır [http-2]. Veri tabanından alınan görüntü örnekleri Şekil 4.6 ile gösterilmektedir.



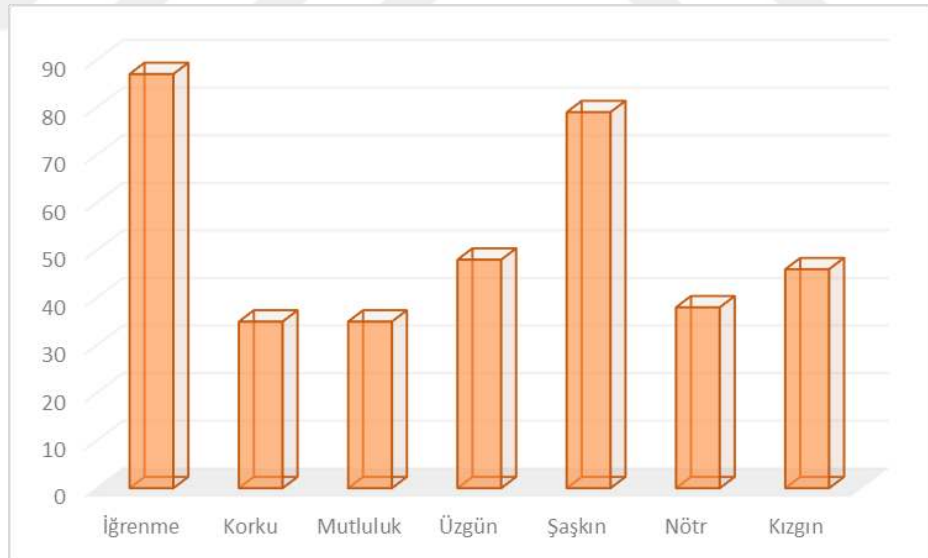
Şekil 4.6. FER2013 veri seti görüntü örnekleri

Çalışmada kullanılan FER2013 veri kümesindeki klasör yalnızca eğitim klasörüdür. Eğitim klasöründe, 870 iğrenme sınıfı verisi, 350 korku sınıfı verisi, 350 mutlu sınıf verisi, 460 kızgın, 480 üzgün sınıf verisi, 790 şaşkın sınıf verisi, 380 nötr detayıyla toplam 3680 görüntü bulunmaktadır. Eğitim veri seti 3312 görüntüyü kapsayacak şekilde ayarlanmıştır ve bunların her biri 10 katlı çapraz doğrulama

kullanılarak dağıtılmıştır. Veriler 7 ifade sınıfına bölünmüştür. Eğitim ve Test veri setlerine ilişkin bilgiler Şekil 4.7 ve Şekil 4.8 ile verilmiştir.



Şekil 4.7. Eğitim veri seti duygu dağılımı



Şekil 4.8. Test veri seti duygu dağılımı

Eğitim aşamasında, her bir yüz görüntüsü, yüz görüntülerinin gri tonlamaya dönüştürüldüğü ve tek tip bir boyuta sahip olması için 48x48 piksele yeniden boyutlandırıldığı yüz hizalama adımına tabi tutulmuştur. Özellik çıkarma, bir nesneyi

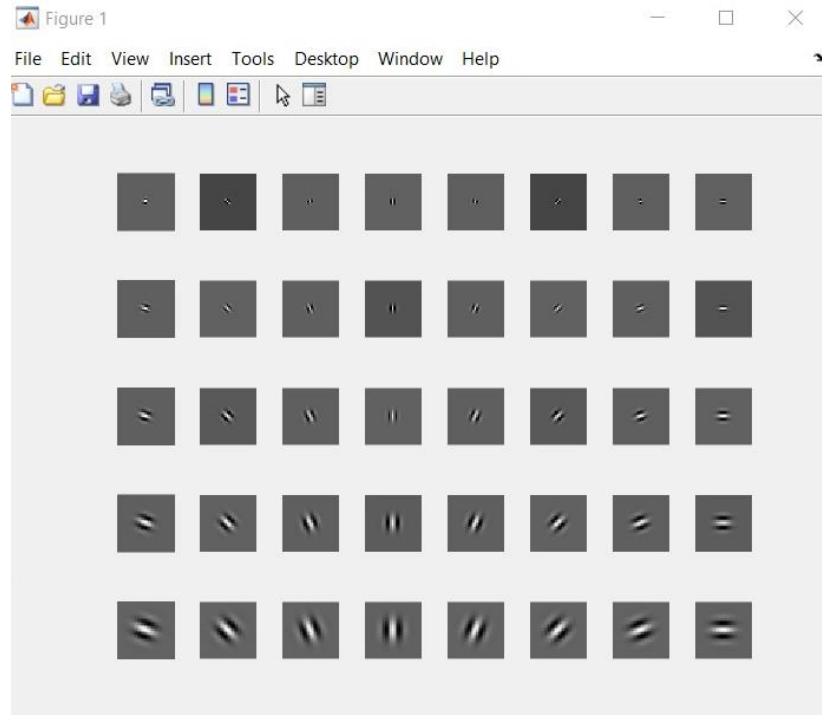
algılamak için çok önemli bir işlemdir. Daha sonra, hizalanmış yüz görüntülerinden öznitelikleri çıkarmak için Gabor tabanlı öznitelik çıkarımı kullanılmıştır. Gabor filtresi Denklem 4.2 ile hesaplanmaktadır.

$$\psi(x, y, \bar{\omega}, \theta) = \frac{1}{2\pi\sigma^2} e^{-\left(\frac{x'^2 + y'^2}{2\sigma^2}\right)} \left[e^{-i\bar{\omega}x'} - e^{-\frac{\bar{\omega}^2\sigma^2}{2}} \right]$$

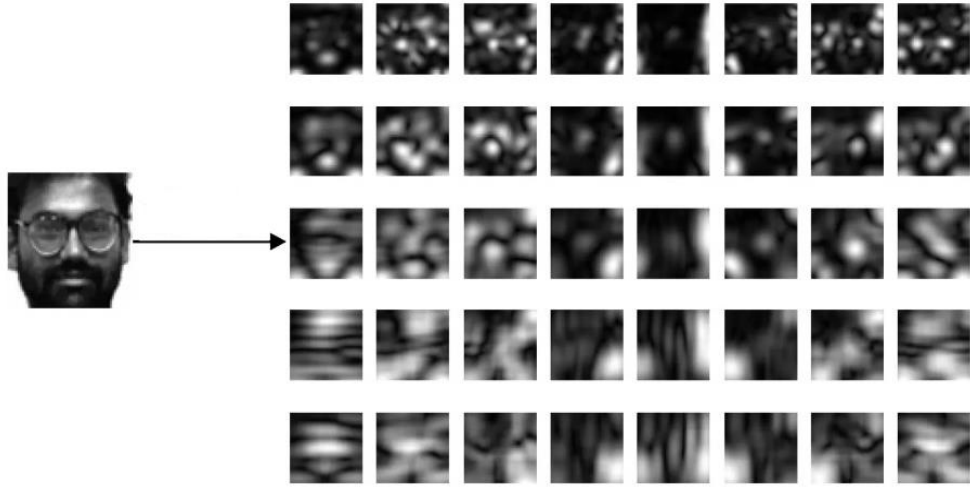
$$x' = x\cos\theta + y\sin\theta, \quad y' = -x\sin\theta + y\cos\theta \quad (4.2)$$

Denklem 4.2’de x ve y eksenleri uzamsal alandaki piksel konumu, ω radyal merkez frekansı, θ Gabor filtresinin yönü ve σ x ve y eksenleri boyunca yuvarlak Gauss fonksiyonunun standart sapmasını temsil etmektedir.

Öznitelik çıkarma aşamasında ilk olarak Şekil 4.9’da gösterildiği gibi, sekiz yön ve beş ölçek içeren bir Gabor filtre bankası oluşturulmuştur. Ardından, Gabor filtre bankasındaki her bir filtre hizalanmış yüz görüntüsüyle birleştirilmiştir. Şekil 4.10, Şekil 4.9’da gösterilen kırk filtreyle birleştirildikten sonra yüz veri kümesinden alınan bir yüz görüntüsünün sonuçlarını göstermektedir. Her görüntü kırk Gabor dalgacı (5x8) ile ilişkilidir. Böylece 128x96’lık bir görüntü 128x96x5x8’e (491.520) dönüştürülmektedir. Bir görüntüdeki bitişik pikseller çoğu durumda yakından ilişkili olduğundan, gereksiz bilgiler Gabor filtrelerinden kaynaklanan görüntüler aşağı örneklemeyle azaltılmıştır.



Şekil 4.9. Gabor çekirdeklerinin beş ölçek ve sekiz yönelimdeki gerçek kısmı



Şekil 4.10. Bir yüz görüntüsü üzerinde Şekil 4.9'da gösterilen kıvrımlı 40 filtrenin sonuçları

Bir önceki çalışmada olduğu gibi bu ağ yapısının mimarisi oluşturulurken başlangıç parametreleri (gizli katman nöron sayısı, kural sayısı, Kernel parametresi (C), minimum hata ve maksimum iterasyon sayısı, vb.) kullanıcı girişi olarak tanımlanmıştır. Elde edilen karışıklık matrisi Tablo 4.6 ile gösterilmektedir.

Tablo 4.6. FER2013 veri seti için FWKELM sınıflandırıcısı kullanılarak elde edilen karışıklık matrisi

	Mutlu	Üzgün	Öfke	Korku	İğrenme	Şaşkın	Nötr
Mutlu	60,00	5,71	0,00	0,00	11,43	14,29	8,57
Üzgün	4,17	37,50	4,17	14,58	14,58	10,42	14,58
Öfke	4,35	15,22	52,17	2,17	8,70	8,70	8,70
Korku	2,86	0,00	2,86	88,57	0,00	5,71	0,00
İğrenme	0,00	4,60	0,00	0,00	95,40	0,00	0,00
Şaşkın	3,80	6,33	5,06	2,53	7,59	68,35	6,33
Nötr	2,63	26,32	10,53	5,26	10,53	10,53	34,21

Mutluluk ifadesi için sınıflandırıcının performansı %60'tır. Korku ifadesi için performans %88,57, şaşkın ifadesi için ise %68,35'tir. Ortalama sınıflandırma doğruluğu %62,32 olarak hesaplanmıştır. Veri setindeki yüz ifadelerinin ve yüz açılarının tutarlılığı, verilerin doğruluk değerini etkilemektedir. FER2013 veri kümesi, bir ön yüz koşulu kullanmamanın yanı sıra, her etikette tutarlı ifadeler kullanmaz, bu da bazı sınıfların belirsiz olmasına neden olur. Belirsiz ifadeler, farklı ifade anlamlarına sahip olduğundan,

ifadelerin belirlenmesine karar vermede şüphe uyandıran çift anlamlı ifadelerdir. Bu nedenle, doğruluk değerinin daha düşük çıktığı söylenebilir.

4.3. Uygulama 3

Son uygulama ise 38 kişiden alınmış orijinal görüntülerin sınıflandırılması hususuna dayanmaktadır. Süreçte ilk olarak veri setini oluşturma adımında Etik Kurul Onayı (Ek-1) alınmıştır. Onayın alınması ile birlikte her biri 6 temel yüz ifadeyi (mutlu, üzgün, korku, kızgın, şaşkın ve nötr) gösteren 17 erkek ve 21 kadın olmak üzere 38 farklı denek 504 görüntüden oluşan bir kapsamlı veri tabanı oluşturulmuştur. Deneklerden görüntüler elde edilirken Gönüllü Katılım Formu (Ek-2) aracılığıyla gerekli kişisel veri paylaşım rızaları alınmıştır. Yüz ifadeleri bazı deneklerden üç farklı açıdan üç kez fotoğraflandırılmıştır. Böylelikle önerilen algoritmanın farklı açıların yüz ifadesi sınıflandırmasında nasıl performans sergilediği ortaya koymak istenmiştir. Orijinal görüntülerin boyutu 768x1024 pikseldir. 6 ifadenin tümünü gösteren bir resim örneği Şekil 4.11 ile gösterilmektedir.



Şekil 4.11. Orijinal yüz veri seti görüntü örnekleri

Yüz ifadesi sınıflandırma için FWKELM algoritmasının sözde kodu aşağıda verilmiştir;

Girdi: Data= {Görüntüler, Etiketler}

Adım 1: Başlangıç

Fonksiyonların tanımlanması

Başlangıç parametre değerlerinin tanımlanması

Görüntüler, etiketler, öznitelikler>CreateDataset (Data)

Adım 2: Görüntü Önileme

For $i=1$: size(görüntü)

Gri Dönüştürme (Histogram eşitleme)

Viola Jones Yüz Tespiti ()

Yeniden Boyutlandırma (görüntü büyüklüğü)

end

Adım 3: Öznitelik Çıkarma

$X_{\text{öznitelik}} = \text{HOG}(\text{görüntüler}, \text{noktolar})$

Adım 4: Boyut Azaltma

$X = \text{LDA}(X, \text{etiketler})$

Adım 5: Sınıflandırma

$X_{\text{egitim}}, X_{\text{test}}, y_{\text{egitim}}, y_{\text{test}} = \text{EgitimTestOlusturma}(X, \text{etiketler}, \%70,)$

$y_{\text{sınıflandırma}} = \text{FWKELM}(X_{\text{egitim}}, X_{\text{test}}, y_{\text{egitim}})$

Adım 6: Değerlendirme

Sınıflama Doğruluğu= Doğruluk ($y_{\text{sınıflandırma}}, y_{\text{test}}$)

Karışıklık Matrisi = Karışıklık Matrisi ($y_{\text{sınıflandırma}}, y_{\text{test}}$)

Return Sınıflandırma Doğruluğu, Karışıklık Matrisi, tamamlanma süresi

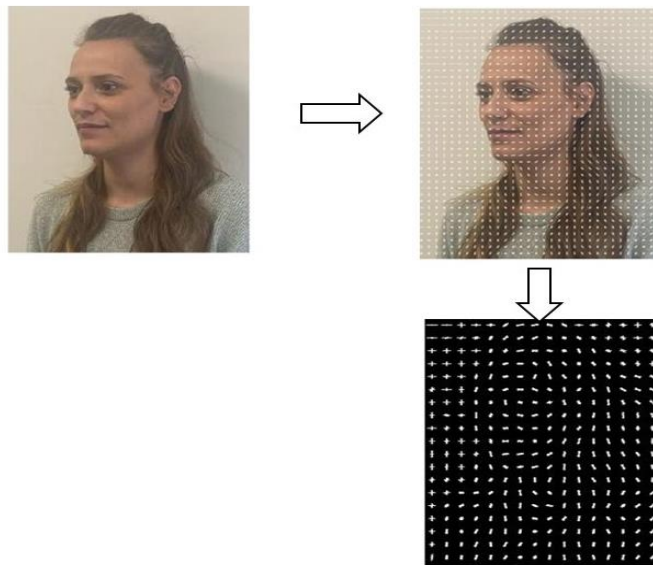
Yüz tespiti işleminden önce, görüntüler normalizasyon işlemine tabii tutulmuştur. Bir yandan da görüntüler aynı boyutta standartlaştırılmıştır. Öte yandan, yüz görüntülerinde tekdüze gri işleme kullanılarak ışık efekti ortadan kaldırılmıştır. Gri işleme için histogram eşitleme uygulanmıştır. Yüz tespiti için VJ (Viola Jones) algoritması kullanılarak görüntülerdeki belirli özellikler kırılarak görüntünün yüz alanı alınmıştır. VJ algoritması, görüntülerdeki yüzleri ve yüz özelliklerini tespit etmek için Haar benzeri öznitelikler ve AdaBoost kullanma fikrine dayanmaktadır. VJ algoritması, bir görüntüdeki birden fazla ölçek ve konum üzerinden hesaplanan geniş bir Haar benzeri öznitelikler kümesi oluşturmaktadır. Algoritma daha sonra kümeden en iyi özellikleri

seçmek ve bu özellikleri kullanarak bir sınıflandırıcıyı eğitmek için AdaBoost algoritmasını kullanmaktadır. Eğitilmiş sınıflandırıcı, bir pencereyi görüntü boyunca kaydırmak ve mevcut pencerede bir nesnenin varlığını değerlendirmek için kullanılmaktadır. Bir nesnenin varlığı tespit edilirse pencere yeniden boyutlandırılmakta ve nesnenin konumu doğru bir şekilde bulununcaya kadar işlem tekrarlanmaktadır. Bir örneği Şekil 4.12 ile gösterilmiştir.



Şekil 4.12. Orijinal resim- gri işlemeye tabi tutulmuş resim- yüz tespiti

Ortaya çıkan işlenmiş yüz görüntüleri daha sonra 128x128 boyutuna ölçeklendirilmiştir. Görüntü özellikleri daha sonra 16 filtreden oluşan bir yönlendirilmiş gradyanların histogramları (HOG) kullanılarak çıkarılmıştır ve Şekil 4.13 ile gösterilmiştir.



Şekil 4.13. Yönlendirilmiş gradyanların histogramları ile öznitelik çıkarımı

HOG, genellikle nesne algılama için kullanılan bir özellik tanımlayıcısıdır. Bir HOG, yoğunluk gradyanlarının veya kenar yönlerinin dağılımına sahip olmak için bir görüntü içindeki nesnelerin özelliğine güvenmektedir. Degradeler, blok başına bir görüntü içinde hesaplanmaktadır. Bir blok, blok içindeki piksellerin yoğunluklarındaki değişimin büyüklüğü ve yönünden gradyanların oluşturulduğu bir piksel ızgarası olarak düşünülmektedir. Ardından doğrusal diskriminant analizi (LDA), önceki aşamadan çıkarılan özelliklerin sayısını düşürmek üzere kullanılmıştır. Ağ mimarisi için kullanıcı girişli parametreler rasgele olarak belirlenmiş ve yüz ifadeleri sınıflandırılmıştır. Sınıflandırma sürecinde eğitim ve test veri kümeleri 70:30 olacak şekilde ayrılmıştır. Bu oran, algoritmanın doğruluğunu test etmek için 150 görüntünün nispeten yeterli olduğu gerçeğiyle birlikte sinir ağını maksimum veri ile eğitmek amaçlandığından seçilmiştir. Ortalama sınıflandırma oranını elde etmek için deney 5 kat çapraz doğrulama ile gerçekleştirilmiştir. Elde edilen karışıklık matrisi Tablo 4.7 ile sunulmuştur.

Tablo 4.7. Orijinal veri seti için normalize edilmiş karışıklık matrisi

	Mutlu	Üzgün	Kızgın	Korku	Şaşkın	Nötr
Mutlu	0,88	0	0	0	0,12	0
Üzgün	0	0,89	0,05	0,06	0	0
Kızgın	0	0,08	0,76	0	0,16	0
Korku	0	0	0,06	0,88	0,06	0
Şaşkın	0	0	0	0,11	0,79	0,1
Nötr	0,03	0	0	0	0	0,97

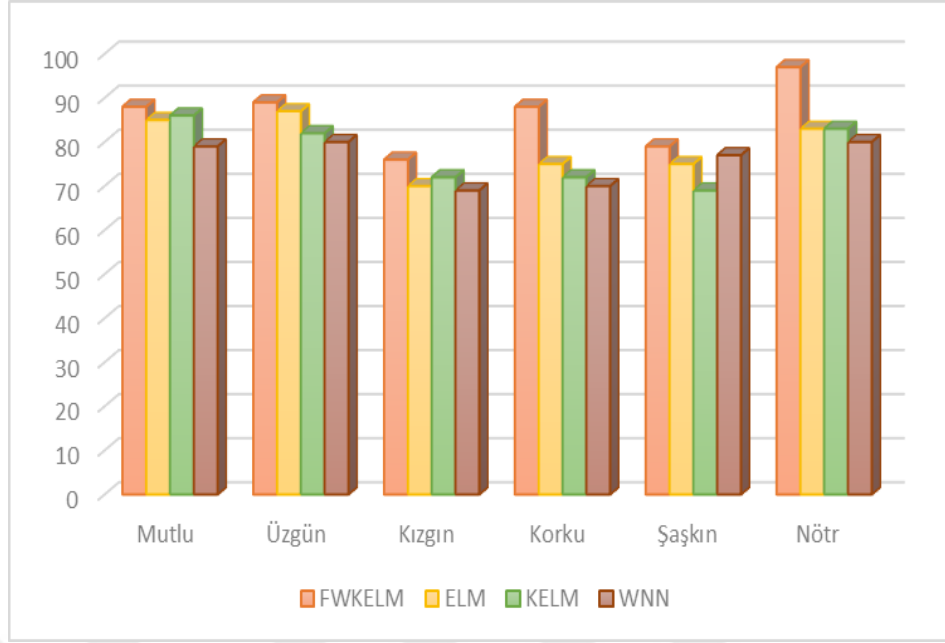
Tablo 4.7 incelendiğinde en yüksek sınıflandırma doğruluğu üzgün ifadelere aittir. Mutluluk ve korku ifadeleri için sınıflandırıcının performansı %88'dir. Ortalama sınıflandırma doğruluğu %86,17 olarak hesaplanmıştır. Yüz ifadeleri genellikle duyguların birleşiminden oluşur ve bir ifadeyi temsil etmek için yalnızca bir etiket kullanmak zor olabilir. Şaşkınlık ifadesi genellikle etiketlenmesi zor olan ifade türlerinden biridir. Yine yapılan deney sonucunda önerilen algoritma şaşkınlık ifadesinde de benzer başarıyı yakaladığı görülmektedir.

Şekil 4.14'te FWKELM yöntemiyle orijinal yüz ifadesi veri setinden doğru şekilde sınıflandırılmış ve yanlış sınıflandırılmış bazı örnek görüntüler bulunmaktadır.



Şekil 4.14. FWKELM yöntemiyle eğitilmiş orijinal veri setinden doğru sınıflandırılmış ve yanlış sınıflandırılmış görüntüleri örnekleri

Şekil 4.14'te “p” tahmin ve “t” doğru etiket içindir. Önerilen algoritmanın sınıflandırma performansını karşılaştırmak adına Orijinal veri seti üzerinde geleneksel makine öğrenmesi algoritmalarından ELM, KELM ve WNN yöntemleri uygulanmıştır. Algoritmaların sınıflandırma doğrulukları Şekil 4.15 ile gösterilmiştir.



Şekil 4.15. Orijinal görüntülerde algoritmaların sınıflandırma doğrulukları

Şekil 4.15'ten görülebileceği gibi, diğer üç yöntemle karşılaştırıldığında, önerilen yöntem en yüksek ortalama sınıflandırma doğruluğuna sahiptir. Bir başka ifade ile önerilen yöntem daha robust bir yöntemdir. Bunun nedeni, hibrit yöntemin tüm kategorilerde genel olarak daha iyi ve istikrarlı performansıdır.

Ek olarak, dalgacık dönüşümü sayesinde FWKELM'nin, problemin belirli özelliklerini etkili bir şekilde ele alabilir ve zaman-uzay frekansı için konum analizi yapabilir. Yüz görüntüleri için çok ölçekli iyileştirme yapmak üzere öteleme faktörünü ve ölçek faktörünü kullanarak, yüksek frekansta zaman altbölümünü ve düşük frekansta frekans altbölümünü gerçekleştirebilir ve sinyallerin zaman-frekans analizine uyum sağlar. Ayrıca, FWKELM görüntünün herhangi bir detayına odaklanabilir ve bu nedenle daha yüksek sınıflandırma oranına sahip olur. Ayrıca, eğim düzeltmesi nedeniyle, bileşik özellikte farklı ifadeler arasındaki farklar daha belirgindir. Örneğin, önerilen yöntem korku ve nötr ifadeleri üzerinde diğer yöntemlere göre çok daha iyi performans göstermektedir. Elde edilen sonuçların eğitim ve test sürelerini değerlendirmek amacıyla gerçekleştirilen deneyler sonucunda elde edilen bilgiler Tablo 4.8 ile gösterilmiştir.

Tablo 4.8. Eğitim ve test sürelerinin ELM, WNN ve FWKELM algoritmalarında karşılaştırılması

Yöntemler	Eğitim Süresi (s)	Test Süresi (s)
Aşırı Öğrenme Makineleri (ELM)	150	2.85
Çekirdek Tabanlı Aşırı Öğrenme Makineleri (KELM)	150	2,82
Dalgacık Yapay Sinir Ağı (WNN)	221	5.00
Önerilen algoritma (FWKELM)	351	5.26

Deney, FWKELM'nin doğruluğu açısından dalgacık sinir ağı algoritmasından üstün olduğunu gösterse de yakınsama süresi olarak dalgacık yapay sinir ağı daha üstündür. Gizli katmanın aktivasyon fonksiyonu olarak dalgacık temel fonksiyonunu kullanan WNN algoritması, algoritmaya dahil edilen öteleme faktörü ve ölçek faktörü ile yüksek frekansta iyi zaman doğruluğu ve düşük frekansta iyi frekans doğruluğu içerir ve karmaşık fonksiyonların özelliklerini daha titizlikle sunabilmektedir. Önerilen algoritmanın her ne kadar ELM aşaması öğrenme süresini önemli ölçüde azaltsa da bulanık kural tabanının ağır işlem yükü önerilen algoritmanın daha yavaş öğrenme hızına neden olmaktadır.

5. SONUÇ VE ÖNERİLER

5.1. Sonuç

Bu tez çalışması, bulanık kuralların, dalgacık dönüşümünün çok çözünürlüklü analizi teorisinin ve KELM'nin bir kombinasyonuna dayanan bulanık dalgacık çekirdek tabanlı aşırı öğrenme makinesi (FWKELM) yapısını sunmaktadır. Spesifik olarak, bu yapının en önemli avantajları, ağır karmaşıklığını verimli bir şekilde azaltabilmesi, yalnızca bir geçişli öğrenme adımına sahip olması ve kabul edilebilir doğruluk ve genelleme performansı sağlayabilmesidir. Tez çalışması kapsamında ilk olarak algoritma yapısının matematiksel modellemesi ortaya konmuş; bir bulanık dalgacık yapay sinir ağına nasıl çekirdek tabanlı aşırı öğrenme makineleri algoritmasına entegre edileceği hususu tartışılmıştır. Aynı zamanda bulanık çıkarım modeli yapısının İSE (THEN) yapısında kullanılan Mexican hat dalgacığının uyarlanabilir bir aşırı öğrenme makinesi çekirdeği olduğu da gösterilmiştir.

Algoritmanın yeteneği, sınıflandırma problemlerine konu olan farklı veri tipleri için (küçük/büyük boyutlu veri seti, yüz ifadesi veri seti vb.) kıyaslamalı gerçek veri kümelerinin sınıflandırılması üzerinde FWKELM modeli uygulanarak doğrulanmıştır. Algoritmanın uygulanmasında MATLAB programından yararlanılmıştır. Önerilen algoritma ve diğer tüm karşılaştırma algoritmalarına ilişkin kodlamalar MATLAB (R2021b) uygulamasında yapılmıştır. Küçük boyutlu veri kümelerinde sınıflandırma için, FWKELM'nin performansı, KELM, WNN ve Mexican Hat KELM ile karşılaştırılabilir ve bildirilen diğer çalışmalardan daha iyi ve standart sapmaların yani değişkenliğin ise daha küçük olduğu belirlenmiştir.

İkinci araştırma çalışmasında sunulan yüz ifadesi sınıflandırma sistemi, davranışsal özelliklerin fizyolojik biyometrik özelliklerle eşlenmesine dayalı esnek bir yüz ifadesi belirleme modeline katkıda bulunmaktadır. İnsan yüzünün mutluluk, üzüntü, korku, öfke, şaşkınlık ve nötr gibi çeşitli ifadelerle ilgili fizyolojik özellikleri, sınıflandırma sistemi için temel eşleştirme şablonu olarak geri yüklenen geometrik yapılarla ilişkilendirilir. Yüz ifadesi veri setlerinin sınıflandırılması sürecinde yer alan öznitelik çıkarma algoritmaları (Yerel ikili örüntü, Yönlendirilmiş gradyanların histogramları, Gabor filtresi) ile önerilen algoritmanın birlikte kullanımı yüksek başarılı sınıf oluşturmaya meyillidir. Literatürde yapılan çalışmalardan örneklerle karşılaştırıldığında, benzer veri setlerinde aynı öznitelik çıkarma algoritmaların diğer sınıflandırıcılarla kombinasyonuna dayalı sonuçların daha az doğruluğu sahip olduğu görülmüştür.

Paralel olarak 38 farklı denekten farklı açılardan farklı poz tekrarları ile 6 temel yüz ifadeyi içerecek şekilde oluşturulan yüz görüntüleri de önerilen algoritmanın sınıflandırma başarısını ölçümlemek amacıyla kullanılmıştır. Burada yüz tespiti için diğer uygulamadan farklı olarak VJ algoritması ile yapılmıştır. Tespit edilen ve yeniden boyutlandırılan görüntü, ön işleme adımları ve önerilen algoritmanın kombinasyonu ile sınıflandırılmıştır. Kombinasyonun sınıflandırma başarısı hem ortalama hem de yüz ifadesi kategorilerinde yüksek olduğu görülmüştür.

Elde edilen bulgulara dayanarak ve insan duygularının ifadesinin ve tanınmasının, insanların kendileri için bile belirsizlikle içsel olduğunu bilerek, gerçek bir belirsizlik ölçüsü istendiğinde FWKELM'nin seçilmesi önerilmektedir. Algoritma nispeten daha uzun sürede tamamlanıyor olsa dahi nispeten yüksek bir doğruluk rafine bir model ortaya koymaktadır.

Üç uygulama beraber incelendiğinde önerilen algoritma bünyesinde bulanık kural tabanı ve aynı zamanda çekirdek yapısını bulundurması hem avantaj hem de dezavantaj sağladığı görülmüştür.

Avantajları;

- ✓ Önerilen algoritmanın sınıflandırma problemlerinde robust sonuçlar ürettiği görülmektedir.
- ✓ Her bir bulanık kuralın İSE (THEN) bölümünde, her iki girdi için bir katsayı dikkate alınan yeni bir bulanık dalgacık modeli yapısı ağ karmaşıklığı ile performans doğruluğu arasında uygun bir denge kurmaktadır.
- ✓ Yapay sinir ağları, bulanık çıkarım sistemi, dalgacık teorisi ve çekirdek tabanlı aşırı öğrenme makinesinin avantajları birleştirildiğinde FWKELM iyi bir performans doğruluğuna sahip olurken, bunun sonucunda ortaya çıkan ağırlık parametrelerinin sayısı önemli ölçüde azalmaktadır.
- ✓ Bu algoritmada, TSK modelindeki gibi sabitler veya doğrusal denklemler kullanmak yerine, her bir bulanık kural, farklı öteleme ve genişleme değerlerine sahip alt dalgacıklara karşılık gelir. Bu nedenle, öğrenme fonksiyonlarının farklı davranışlarını (global veya yerel) yakalamak için farklı genişleme değerlerindeki (yani farklı çözünürlük seviyelerindeki) alt dalgacık yapay sinir ağlarını kullanılmaktadır.
- ✓ Yinelemeli öğrenme algoritmaların aksine (ağın parametreleri yinelemeli olarak ayarlanmalıdır), önerilen yapı, bulanık kuralların ardışık bölümlerinin doğrusal

öğrenme parametrelerini analitik olarak belirlemek için yalnızca tek geçişli öğrenme algoritması kullanmaktadır.

Dezavantajları ise;

- ✓ Önerilen algortmada dalgacık katsayıları rastgele üretılmekte ve sabit kalmaktadır. Ek olarak, algortmada başlangıç parametrelerin hepsi kullanıcı girişli ve rastgele tanımlanmaktadır.
- ✓ FWKELM yapısı bulanık çıkarım sistemi ve dalgacık modeli iş yükü hacmini artırmasından dolayı klasik sınıflandırma algortmalarına göre öğrenme hızı yavaş kalmaktadır.

5.2. Öneriler

Bu tez çalışmasında önerilen yapının dezavantajı, başlangıç parametrelerinin rastgele seçiliyor olmasıdır. Parametre değeri kullanıcı girişli tanımlamak yerine optimal parametre değeri sezgisel algortmalar yardımı ile belirlenebilir. Ayrıca yine yapının bir diğeri dezavantajı uygulama zamanının çok olmasıdır. Öznitelik çıkarımı işlemini hızlandırmak için farklı yöntemler geliştirilebilir. Daha iyi bir ön işleme yaklaşımı ile yüz olma olasılığı olan alanın küçültülmesi sağlanabilir. Sınıflandırma basamağında, farklı sinir ağıları geliştirilip bu çalışmada kullanılmış ileri beslemeli yapay sinir ağıları ile karşılaştırılarak performans değeri değerlendirilmesi yapılabilir.

KAYNAKÇA

- [1] Haykin, S. (1994). *Neural Networks: A Comprehensive Foundation*, McMillan, New York.
- [2] Zhang, Q. (1997). Using wavelet networks in nonparametric estimation, *IEEE Trans. Neural Netw.* 8, 227–236.
- [3] Wang, L. (1994). *Adaptive Fuzzy Systems and Control*, Prentice-Hall, Englewood Cliffs, NJ.
- [4] Wang, J., Wang, C., Chen, C. (2014). The bounded capacity of fuzzy neural networks (FNNs) via a new fully connected neural fuzzy inference system (F-CONFIS) with its applications, *IEEE Trans. Fuzzy Syst.* 22(6), 1373–1386.
- [5] Lin, C. K. (2009). H_{∞} reinforcement learning control of robot manipulators using fuzzy wavelet networks, *Fuzzy Sets Syst.* 160(12), 1765–1786.
- [6] Reyneri, L. (1999). Unification of neural and wavelet networks and fuzzy systems, *IEEE Trans. Neural Netw.* 10(4), 801–814.
- [7] Daqi, G., Genxing, Y. (2003). Influences of variable scales and activation functions on the performances of multilayer feed forward neural networks, *Pattern Recognit.* 36(4), 869–878.
- [8] Svozil, D., Kvasnicka, V., Pospichal, J. (1997). Introduction to multi-layer feed-forward neural networks, *Chemom. Intell. Lab. Syst.* 39(1), 43–62.
- [9] Abiyev, R. H., Kaynak, O. (2008). Fuzzy wavelet neural networks for identification and control of dynamic plants—A novel structure and a comparative study, *IEEE Trans. Ind. Electron.* 55(8), 3133–3140.
- [10] Takagi, T., Sugeno, M. (1985). Fuzzy identification of systems and its applications for modeling and control, *IEEE Trans. Syst. Man Cybern.* SMC-15(1), 116–132.
- [11] Mendel, J. (1995). Fuzzy logic systems for engineering: a tutorial, *Proc. IEEE* 83(3), 345–377.
- [12] Zhang, J., Walter, G. G., Miao, Y., Wayne L.W. (1995). Wavelet neural networks for function learning, *IEEE Trans. Signal Process.* 43(6).
- [13] Ho, D., Zhang P.A., Xu, J. (2001). Fuzzy wavelet networks for function learning, *IEEE Trans. Fuzzy Syst.* 9(1), 200–211.
- [14] Lu, C. (2011). Wavelet fuzzy neural networks for identification and predictive control of dynamic systems, *IEEE Trans. Ind. Electron.* 58(7), 3046–3058.
- [15] Lin, F., Huang, Y., Ruan, K. (2014). An intelligent second-order sliding-mode control for an electric power steering system using a wavelet fuzzy neural network, *IEEE Trans. Fuzzy Syst.* 22(6), 1598–1611.
- [16] Huang, G.B., Zhu, Q.Y., Siew, C.K. (2006). Extreme learning machine: theory and applications, *Neurocomputing* 70(1–3), 489–501.
- [17] Javed, K., Gouriveau, R., Zerhouni, N. (2014). SW-ELM: a summation wavelet extreme learning machine algorithm with a priori parameter initialization, *Neurocomputing* 123, 299–307.

- [18] Huang, G.B., Siew, C.K. (2004). Extreme learning machine: RBF network case. ICARCV 2004 8th International Conference on Control, Automation, Robotics and Vision Conference, C. 2, ss. 1029–1036.
- [19] Zong, W., Huang, G.B. (2011). Face Recognition Based on Extreme Learning Machine, *Neurocomputing* 74, pp. 2541–2551.
- [20] Zong, W., Zhou, H., Huang, G.B., Zhiping, L. (2011). Face Recognition Based on Kernelized Extreme Learning Machine, Springer-Verlag Berlin Heidelberg, pp. 263–272, 2011.
- [21] Yang, J., Xie, S., Yoon, S., Park, D., Fang, Z., Yang, S. (2012). Fingerprint Matching Based on Extreme Learning Machine, *Extreme Learning Machine's Theory & Application, Neural Comput & Applic*, pp. 22:435–445.
- [22] Liu, X., Wang, L., Huang, G.B., Zhang, J., Yin, J. (2013). Multiple Kernel Extreme Learning Machine. *Neurocomputing*. 149.
- [23] Pal, M., Maxwell, A.E., Warner, T.A. (2013). “Kernel-Based Extreme Learning Machine for Remote-Sensing Image Classification.” *Remote Sensing Letters* 4 (9): 853–862. doi:10.1080/2150704X.2013.805279.
- [24] Chen, C., Li, W., Su, H., Liu, K. (2014). Spectral-Spatial Classification of Hyperspectral Image Based on Kernel Extreme Learning Machine, *Remote Sens.* 2014, 6, 5795-5814.
- [25] Iosifidis, A., Tefas, A., Pitas, I. (2015). Human Action Recognition Based on Multi-View Regularized Extreme Learning Machine, *International Journal on Artificial Intelligence Tools* Vol. 24, No. 5.
- [26] Liu, H., Du, B., Liu, J., Xia, H., Yeap, W.K. (2016). A Kernel Extreme Learning Machine Algorithm Based on Improved Particle Swarm Optimization, *Memetic Comp., Springer*.
- [27] Parikh, K., Shah, T. (2016). Kernel Based Extreme Learning Machine in Identifying Dermatological Disorders, *International Journal of Innovative Science, Engineering & Technology*, Vol. 3 Issue 10, pp. 370-375.
- [28] Zhang, L., Wang, H., Zhao, C., Zhang, Y. (2016). Parameters Selection of Kernel Based Extreme Learning Machine Using Particle Swarm Optimization, *Rev. Téc. Ing. Univ. Zulia*. Vol. 39, No 7, pp. 100 – 106.
- [29] Kongsorot, Y., Horota, P., Musikawan, P., Sunat, K. (2017). Kernel Extreme Learning Machine Based on Fuzzy Set Theory for Multi-Label Classification, *International Journal of Machine Learning and Cybernetics*, Springer-Verlag GmbH Germany, part of Springer Nature.
- [30] Mahmud, F., Maun, M. A. (2017). Facial Expression Recognition System Using Extreme Learning Machine, *International Journal of Scientific & Engineering Research*, Volume 8, Issue 3, pp.26-30.
- [31] Anam, K., Al-Jumaily, A. (2018). Optimized Kernel Extreme Learning Machine for Myoelectric Pattern Recognition, *International Journal of Electrical and Computer Engineering* 8(1):483-496.

- [32] Zhang, B.L., Zhang, H., Ge, S. S. (2004). Face Recognition by Applying Wavelet Subband Representation and Kernel Associative Memory, *Ieee Transactions on Neural Networks*, VOL. 15, NO. 1.
- [33] Subaşı, A., Alkan, A., Köklükaya, A., Kiymik, M.K. (2006). Wavelet Neural Network Classification of EEG Signals by Using AR Model with MLE Preprocessing, *Neural Networks Volume 18, Issue 7*, pp. 985-997.
- [34] Wang, W., Li, J., Chen, W. (2006). Fingerprint Classification Using Improved Directional Field and Fuzzy Wavelet Neural Network, *Proceedings of the 6th World Congress on Intelligent Control and Automation*, June 21 - 23, Dalian, China.
- [35] Sahoolizadeh, H., Sarikhanimoghadam, D., Dehghani, H. (2008). Face Detection using Gabor Wavelets and Neural Networks, *World Academy of Science, Engineering and Technology*, Vol. 45, pp552- 554.
- [36] Zekri, M., Sadri, S., Sheikholeslam, F. (2008). Adaptive fuzzy wavelet network control design for nonlinear systems, *Fuzzy Sets and Systems*, Volume 159, Issue 20, Pages 2668-2695.
- [37] Yilmaz, S., Oysal, Y. (2010). Fuzzy wavelet neural network models for prediction and identification of dynamical systems, *IEEE Trans. Neural Netw.* 21(10, 1599–1609.
- [38] Shankar, B.U., Meher, S., Ghosh, A. (2011). Wavelet-Fuzzy Hybridization: Feature-Extraction and Land-Cover Classification Of Remote Sensing Images. *Appl. Soft Comput.*, 11, 2999-3011.
- [39] Abidin, Z., Alamsyah, A. (2015). Wavelet Based Approach for Facial Expression Recognition, *International Journal of Advances in Intelligent Informatics*, Vol 1, No 1, pp. 7-14.
- [40] Dharani, P., Vibhute, A. S. (2017). Face Recognition Using Wavelet Neural Network. *International Journal of Advanced Research in Computer Science and Software Engineering*, Volume 7, Issue 5, 101-107.
- [41] Güzel, K., Bilgin, G. (2018). Hiperpektral Görüntülerin Sınıflanması için Dokusal Özellik Çıkarımı ve Aşırı Öğrenme Makinelerinin Topluluğu” *IEEE 26. Sinyal İşleme ve Uygulamaları Kurultayı- IEEE 26th Conf. on Signal Processing and Communications Applications*, Çeşme-İzmir, Türkiye.
- [42] Golestaneh, P., Zekri, M., Sheikholeslam, F. (2018). Fuzzy Wavelet Extreme Learning Machine, *Fuzzy Sets and Systems* 342, pp. 90–108.
- [43] Sailaja, G., Deepthi, V.H. (2018). Wavelet Based Feature Extraction for Facial Expression Recognition, *International Journal of Pure and Applied Mathematics*, Volume 118, No.24.
- [44] Shi, W., Jiang, M. (2018). Fuzzy Wavelet Network with Feature Fusion and LM Algorithm for Facial Emotion Recognition, *IEEE International Conference of Safety Produce Informatization (IICSPI)*, Chongqing, China, 2018, pp. 582-586.
- [45] Alneamy, M.J.S., Alnaish, H.A.Z., Hashim, M.S.Z, Alnaish, H.A.R. (2019). Utilizing Hybrid Functional Fuzzy Wavelet Neural Networks with A Teaching Learning-Based Optimization Algorithm for Medical Disease Diagnosis, *Computers in Biology and Medicine*, Volume 112.

- [46] Ahmadi, M., Ahangar, F.D., Astaraki, N., Abbasi, M., Babaei, B. (2021). "FWNNNet: Presentation of a New Classifier of Brain Tumor Diagnosis Based on Fuzzy Logic and the Wavelet-Based Neural Network Using Machine-Learning Methods", *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 8542637, 13 pages, 2021.
- [47] McCulloch W.S., Pitts W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bull Mathematical Biophys* 5(4):115–133.
- [48] Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol Rev* 65(6):386–408.
- [49] Werbos, P. (1974). Beyond regression: new tools for prediction and analysis in the behavioral sciences. Ph. D. dissertation, Harvard University.
- [50] Leung, H.C., Leung, C.S., Wong, E.W.M. (2019). Fault and Noise Tolerance in the Incremental Extreme Learning Machine. *IEEE Access* 7:155171–155183.
- [51] Li, H.T., Chou, C.Y., Chen, Y.T., Wang, S.H., Wu, A.Y. (2019). Robust and Lightweight Ensemble Extreme Learning Machine Engine Based on Eigenspace Domain for Compressed Learning. *IEEE Trans Circuits Syst I: Regular Papers* 66(12):4699–4712.
- [52] Deng W.Y., Bai Z., Huang, G.B., Zheng, Q.H. (2016). A Fast SVD-Hidden-nodes based Extreme Learning Machine for Large-Scale Data Analytics. *Neural Network* 77:14–28
- [53] Huang, J., Yu, Z.L., Gu, Z. (2018). A clustering method based on extreme learning machine. *Neurocomputing* 277:108–119
- [54] Li, X., He, S., Wei, Z., Wu, L. (2019). Improved Online Sequential Extreme Learning Machine: A New Intelligent Evaluation Method for AZ-Style Algorithms. *IEEE Access* 7:124891–124901
- [55] Zhang, M., Liu, X., Zhang, Z. (2016). A soft sensor for industrial melt index prediction based on evolutionary extreme learning machine. *Chin J Chem Eng* 24(8):1013–1019
- [56] Huang, G.B., Ding, X., Zhou, H. (2010). Optimization method based extreme learning machine for classification. *Neurocomputing* 74(1-3):155–163.
- [57] Huang, G.B., Zhu, Q.Y., Siew, C.K. (2004). Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks. *Proc Intl Joint Confer Neural Networks (IJCNN2004)* 2:985–990.
- [58] Liu, J.F., Chen, Y.Q., Liu M.J. et al. (2011). SELM: semi-supervised ELM with application in sparse calibrated location estimation. *Neurocomputing* 74, 2566–2572.
- [59] Crandall, R. (1994). *Projects in Scientific Computation*. Springer-Verlag, New York, pp. 197-198, 211-212.
- [60] Meyer, Y. (1993). *Wavelets: Algorithms and Applications*. Society for Industrial and Applied Mathematics, Philadelphia, pp. 13-31, 101-105.
- [61] Graps, A. (1995). *An Introduction to Wavelets*. IEEE Computational Science and Engineering, Vol. 2, No. 2 pp. 50-61.

- [62] Haar, A. (1910). Zur theorie der orthogonalen funktionensysteme. *Mathematische Annalen*, 69 (3), 331-371.
 - [63] Levy, P. (1940). Le mouvement brownien plan. *Amer. J. Math.* 62, 487–550.
 - [64] Coifman, R.R., Weiss, G. (1977). Extensions Of Hardy Spaces and Their Use in Analysis. *Bulletin of The American Mathematical Society*, Volume 83, Number 4, 569-645.
 - [65] Grossman, A., Morlet, J. (1984). Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape. *SIAM Journal on Mathematical Analysis*, 15, 723-736.
 - [66] Mallat, S. (1989). A Theory of Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Trans. Pattern Anal. Machine Intell.* 11, 674-693. *Pattern Analysis and Machine Intelligence, IEEE Transactions on.* 11. 674- 693.
 - [67] Daubechies, I. (1988). Orthonormal Bases of Compactly Supported Wavelets, *Comm. Pure Appl. Math.*, Vol 41, pp. 906-966.
 - [68] Pati, Y. C., Krishnaprasad, P. S. (1993). Analysis and Synthesis of Feedforward Neural Networks Using Discrete Affine Wavelet Transforms. *IEEE Trans. on Neural Networks*, 4(1), 73-85.
 - [69] Zhang, Q., Benveniste, A. (1992). Wavelet Networks. *IEEE Trans. Neural Networks*, 3(6), 889-898.
 - [70] Yang, H. Z., Zhao, W. G., Chen, W., Chen, X. Q. (2011). The Study of Prediction Model Based on Morlet Wavelet Neural Network. *Applied Mechanics and Materials*, 121-126, 4847–4851.
 - [71] Oussar, Y., Rivals, I., Presonnaz, L., Dreyfus, G. (1998). Training Wavelet Networks for Nonlinear Dynamic Input Output Modelling. *Neurocomputing*, 20, 173–188.
 - [72] Zhang, Q. (1993). Regressor Selection and Wavelet Network Construction, Technical Report, INRIA.
 - [73] Postalcioglu, S., Becerikli, Y. (2007). Wavelet Networks for Nonlinear System Modelling, *Neural Computing and Applications*, 16, 434–441.
 - [74] Anders, U., Korn, O. (1999). Model selection in neural networks. *Neural networks the official journal of the International Neural Network Society*, 12 2, 309-323.
 - [75] Szu, H. H., Telfer, B., Kadambe, S. (1992). Neural Network Adaptive Wavelets for Signal Representation and Classification. *Opt. Eng.*, Vol. 31, 1907-1916.
 - [76] Abidin, Z., Harjoko, A. (2012). A Neural Network based Facial Expression Recognition using Fisherface. *International Journal of Computer Applications* (0975 – 8887) Volume 59– No.3
 - [77] Ghimire, D., Lee, J. (2014). Extreme Learning Machine Ensemble Using Bagging for Facial Expression Recognition. *J Inf Process Syst*, Vol.10, No.3, pp.443-458.
- http-1: <https://archive.ics.uci.edu/> (Eriřim Tarihi: 15.11.2022)
- http-2: <https://www.kaggle.com/datasets> (Eriřim Tarihi: 15.04.2023)

EKLER

Ek 1: Etik Kurul Onay Yazısı



T.C.
ESKİŞEHİR TEKNİK ÜNİVERSİTESİ REKTÖRLÜĞÜ
Etik Kurulları Sekreteryası

Sayı : E-87914409-050.03.04-2200020290
Konu : Etik Kurul Kararı

21.11.2022

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ MÜDÜRLÜĞÜNE

İlgi : 20.09.2022 tarihli ve E-60850919-100-2200008467 sayılı yazınız.

İlgi yazınıza istinaden Rektörlüğümüze gönderilen Doç. Dr. Özer ÖZDEMİR'in yürütücülüğünü yaptığı "Geliştirilmiş Aşırı Öğrenme Makineleri ve Büyük Veri Yaklaşımları ile Sınıfları Analizi" başlıklı Lisansüstü Tez çalışmasına ilişkin başvurusu Fen ve Mühendislik Bilimleri Bilimsel Araştırma ve Yayın Etiği Kurulu tarafından incelenmiş olup etik açıdan uygun bulunmuştur.

Bilgilerinizi ve uygulama dosyasının hazırlanmasında, ilgili kurumun bulunması halinde Etik Kurulu Yönergesinin dikkate alınması konusunda gereğini rica ederim.

Prof. Dr. Onur KAYA
Fen Bil. ve Müh. Bil. Bilimsel Araş. ve Yay. Etiği Kurulu Bşk.

Ek: Etik Kurul Kararı

Ek 2: Araştırma Gönüllü Katılım Formu Örneği

ARAŞTIRMA GÖNÜLLÜ KATILIM FORMU

Bu çalışma, Geliştirilmiş Aşırı Öğrenme Makineleri ve Büyük Veri Yaklaşımları ile Sınıflandırma Analizi başlıklı bir araştırma çalışması olup doktora tez amacını taşımaktadır. Çalışma, İstatistik Anabilim Dalı Doktora Öğrencisi Aslı KAYA tarafından yürütülmekte ve sonuçları ile yüz sınıflandırma ve tanıma probleminin gelişimine ışık tutulacaktır.

- Bu çalışmaya katılımınız gönüllülük esasına dayanmaktadır.
- Çalışmanın amacı doğrultusunda, görüntü kaydı yapılarak sizden veriler toplanacaktır.
- İsminizi yazmak ya da kimliğinizi açığa çıkaracak bir bilgi vermek zorunda değilsiniz/araştırmada katılımcıların isimleri gizli tutulacaktır.
- Araştırma kapsamında toplanan veriler, sadece bilimsel amaçlar doğrultusunda kullanılacak, araştırmanın amacı dışında ya da bir başka araştırmada kullanılmayacak ve gerekmesi halinde, sizin (yazılı) izniniz olmadan başkalarıyla paylaşılmayacaktır.
- İstemeniz halinde sizden toplanan verileri inceleme hakkınız bulunmaktadır.
- Sizden toplanan veriler istatistiksel programlarda dosyalama yöntemi ile korunacak ve araştırma bitiminde arşivlenecek veya imha edilecektir.
- Veri toplama sürecinde/süreçlerinde size rahatsızlık verebilecek herhangi bir soru/talep olmayacaktır. Yine de katılımınız sırasında herhangi bir sebepten rahatsızlık hissederseniz çalışmadan istediğiniz zamanda ayrılabilirsiniz. Çalışmadan ayrılmanız durumunda sizden toplanan veriler çalışmadan çıkarılacak ve imha edilecektir.

Gönüllü katılım formunu okumak ve değerlendirmek üzere ayırdığınız zaman için teşekkür ederim. Çalışma hakkındaki sorularınızı Eskişehir Teknik Üniversitesi Lisansüstü Eğitim Enstitüsü İstatistik Anabilim Dalı'ndan Aslı Kaya' ya yöneltebilirsiniz.

Araştırmacı Adı: Aslı KAYA
Adres : ESTÜ-Rektörlük

İş Tel : +90 222 321 35 50
Cep Tel : [REDACTED]

Bu çalışmaya tamamen kendi rızamla, istediğim takdirde çalışmadan ayrılabileceğimi bilerek verdiğim bilgilerin bilimsel amaçlarla kullanılmasını kabul ediyorum.
(Lütfen bu formu doldurup imzaladıktan sonra veri toplayan kişiye veriniz.)

Katılımcı Ad ve Soyadı:
İmza:
Tarih:

ÖZGEÇMİŞ

ORCID NO: 0000-0003-2155-9391

Ad Soyad : Aslı KAYA KARAKÜTÜK

Yabancı Dil : İngilizce

Eğitim ve Mesleki Geçmişi:

- 2015-2018, Anadolu Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı (Yüksek Lisans)
- 2011-2016, Anadolu Üniversitesi, İşletme Bölümü (Lisans)
- 2009-2014, Anadolu Üniversitesi, İstatistik Bölümü (Lisans)
- 2021-Halen, Koordinatör Yardımcısı, Eskişehir Teknik Üniversitesi, Kurumsal Gelişim ve Planlama Koordinatörlüğü

Yayınları ve/veya Bilimsel/Sanatsal Faaliyetleri:

- Özdemir, Ö., Kaya, A. (2019). Comparison of FCM, PCM, FPCM and PFCM Algorithms in Clustering Methods. Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi, Cilt: 19 Sayı: 1,92-102.
- Kaya, A., Özdemir, Ö., Karakütük, F. (2021). Investigation of Life Satisfaction Data by Data Mining and Machine Learning Techniques, International Journal of Cultural Heritage, Volume 6(1), 63-68.
- Kaya, A., Çemrek, F., Özdemir, Ö. (2021). Forecasting Analysis of Covid-19 Cases with Wavelet Neural Network and Time Series Approach, Volume 8(2), 6-11.

Ödülleri:

- TÜBİTAK, 2211-C Yurt İçi Öncelikli Alanlar Doktora Bursu
- YÖK 100/2000 Doktora Bursu