

**KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

**DERİN ÖĞRENMEYE DAYALI KULAK ZARI OTOSKOP GÖRÜNTÜLERİ İÇİN TANISAL
BAĞIMSIZ BİR MOBİL UYGULAMANIN GELİŞTİRİLMESİ, TEST EDİLMESİ VE FİZİBİLİTESİ**

YÜKSEK LİSANS TEZİ

Furkancan DEMİRCAN

**TEMMUZ 2023
TRABZON**



KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

**DERİN ÖĞRENMEYE DAYALI KULAK ZARI OTOSKOP GÖRÜNTÜLERİ İÇİN
TANISAL BAĞIMSIZ BİR MOBİL UYGULAMANIN GELİŞTİRİLMESİ, TEST EDİLMESİ
VE FİZİBİLİTESİ**

Furkancan DEMİRCAN

ORCID : - - -

**Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsünce
"BİLGİSAYAR YÜKSEK MÜHENDİSİ"
Unvanı Verilmesi İçin Kabul Edilen Tezdir.**

Tezin Enstitüye Verildiği Tarih : 13 / 07 / 2023

Tezin Savunma Tarihi : 07 / 07 / 2023

Tez Danışmanı : Prof. Dr. Murat EKİNCİ

ORCID : - - -

İkinci Danışman : Doç. Dr. Zafer CÖMERT

Trabzon 2023

ÖNSÖZ

Çalışmalarım süresince bilgi, görüş ve önerileriyle bana yardımcı olan çok değerli danışman hocam sayın Prof. Dr. Murat EKİNCİ'ye teşekkürlerimi bir borç bilir ve şükranlarımı sunarım.

Çalışma süresinde vermiş olduğu değerli fikirler ve öneriler ile yol gösteren ikinci danışman hocam sayın Doç. Dr. Zafer CÖMERT'e teşekkürlerimi sunarım.

Yüksek lisans eğitimim süresince sabır, destek ve sevgisiyle, bu yaşıma kadar benim her anımda yanımda olan ve desteklerini hiçbir zaman esirgemeyen canım aileme ve çalışmalarımda her zaman bana destek olan değerli hocalarım ve arkadaşlarıma teşekkür eder şükranlarımı ve saygılarımı sunarım.

Furkancan DEMİRCAN

Trabzon 2023

TEZ ETİK BEYANNAMESİ

Yüksek Lisans Tezi olarak sunduğum “Derin Öğrenmeye Dayalı Kulak Zarı Otoskop Görüntüleri İçin Tanısal Bağımsız Bir Mobil Uygulamanın Geliştirilmesi, Test Edilmesi Ve Fizibilitesi” başlıklı bu çalışmayı baştan sona kadar danışmanım Prof. Dr. Murat EKİNCİ’nin sorumluluğunda tamamladığımı, verileri / örnekleri kendim topladığımı, deneyleri / analizleri ilgili laboratuvarlarda yaptığımı / yaptırdığımı, başka kaynaklardan aldığım bilgileri metinde ve kaynakçada eksiksiz olarak gösterdiğimi, çalışma sürecinde bilimsel araştırma ve etik kurallara uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul ettiğimi beyan ederim. 07/07/2023

Furkancan DEMİRCAN

İÇİNDEKİLER

	<u>Sayfa No</u>
ÖNSÖZ	iii
TEZ ETİK BEYANNAMESİ.....	iv
İÇİNDEKİLER.....	v
ÖZET	vii
SUMMARY	viii
ŞEKİLLER DİZİNİ.....	ix
TABLolar DİZİNİ.....	xiv
SEMBOLLER DİZİNİ	xvi
1. GENEL BİLGİLER.....	1
1.1. Giriş	1
1.2. Bekle ve Gör Yaklaşımı (Wait-And-See Approach)	6
1.3. Antibiyotik Tedavisi (Antibiotic Therapy)	7
1.4. Ameliyat.....	7
2. YAPILAN ÇALIŞMALAR	8
2.1. Veri Seti (Dataset)	13
2.2. Derin Öğrenme	18
2.2.1. Evrişim Katmanı	18
2.2.2. Havuzlama Katmanı.....	19
2.2.3. Ekleme Katmanı.....	19
2.2.4. Tam-bağlı Katman.....	20
2.3. Derin Öğrenme Modelleri.....	20
2.3.1. YOLO (You Only Look Once)	20
2.3.2. ResNet	23
2.3.3. Xception	26
2.3.4. InceptionResNet	28
2.3.5. DenseNet.....	32
2.3.6. Vision Transformer	34
2.3.7. LSTM	37
2.3.8. EfficientNet	38

2.4.	Ön İşleme (Preprocessing).....	41
2.5.	Performans Metrikleri.....	45
3.	BULGULAR VE TARTIŞMA	47
3.1.	Region Of Interest (ROI) - İlgili Alanın Çıkarılması / Zar Bölgesinin Tespiti	47
3.2.	Hazır Modeller	48
3.3.	Hibrit Modeller	48
3.4.	Modellerden Elde Edilen Bulgular	49
4.	SONUÇLAR	114
5.	ÖNERİLER	115
6.	KAYNAKLAR.....	116
	ÖZGEÇMİŞ.....	119

Yüksek Lisans Tezi

ÖZET

DERİN ÖĞRENMEYE DAYALI KULAK ZARI OTOSKOP GÖRÜNTÜLERİ İÇİN TANISAL
BAĞIMSIZ BİR MOBİL UYGULAMANIN GELİŞTİRİLMESİ, TEST EDİLMESİ VE
FİZİBİLİTESİ

Furkancan DEMİRCAN

Karadeniz Teknik Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı
Danışman: Prof. Dr. Murat EKİNCİ
2023, 118 Sayfa

Kulak hastalıkları günlük hayatımızı olumsuz etkileyen sağlık problemleridir. Bu hastalıkların tanısı uzmanlar tarafından otoskop cihazı kullanılarak gözlem ile koyulmaktadır ve bu tanı sisteminin doğruluğu uzman bilgisine bağlıdır. Bu çalışma kapsamında kulak hastalık tanılarının derin öğrenme modelleri kullanılarak koyulması araştırılmıştır. Yapılan çalışmada Ear Imagery ve Tympanic membrane veri setleri kullanılmıştır. Ayrıca bu iki veri seti birleştirilerek ayrı bir veri seti oluşturulmuştur. Kullanılan veri setleri eğitim, doğrulama ve test işlemleri için sırasıyla %70, %10 ve %20 olarak ayrılmıştır. Derin öğrenme modeli olarak ResNet50, InceptionV3, InceptionResNetV2 ve Xception modelleri seçilmiştir. Ayrıca DenseNet201, EfficientNetB5 ve InceptionResNetV2 özellik çıkarıcı modeller ile BiLSTM ve ConvBiLSTM sınıflandırıcı modellerden hibrit modeller oluşturulmuştur. Çalışma kapsamında sınıflandırma başarısına etkiyi araştırmak için sıfırdan eğitim, öğrenme aktarımı, HSV ve LUV renk uzayları kullanılmıştır. Yapılan çalışma sonucunda en yüksek doğruluk başarısı DenseNet201-BiLSTM ve ResNet50 modellerinden elde edilmiştir.

Anahtar Kelimeler: Kulak hastalıkları, Derin öğrenme, Sınıflandırma, Hazır modeller, Hibrit modeller

Master Thesis

SUMMARY

DEVELOPMENT, TESTING AND FEASIBILITY OF A DIAGNOSTICALLY INDEPENDENT MOBILE APPLICATION FOR EARDRUM OTOSCOPE IMAGES BASED ON DEEP LEARNING

Furkanca DEMIRCAN

Karadeniz Technical University
The Graduate School of Natural and Applied Sciences
Computer Engineering Graduate Program
Supervisor: Prof. Dr. Murat EKİNCİ
2023, 108 Page

Ear diseases are health problems that negatively affect our daily lives. These diseases are diagnosed by experts through observation using an otoscope device and the accuracy of this diagnosis system depends on expert knowledge. This study investigates the diagnosis of ear diseases using deep learning models. Ear Imagery and Tympanic membrane datasets were used in this study. In addition, a separate dataset was created by combining these two datasets. The datasets used were divided into 70%, 10% and 20% for training, validation and testing respectively. ResNet50, InceptionV3, InceptionResNetV2 and Xception models were selected as deep learning models. In addition, hybrid models were created from DenseNet201, EfficientNetB5 and InceptionResNetV2 feature extractor models and BiLSTM and ConvBiLSTM classifier models. Within the scope of the study, zero training, learning transfer, HSV and LUV color spaces were used to investigate the effect on classification success. As a result of the study, the highest accuracy was obtained from DenseNet201-BiLSTM and ResNet50 models.

Keywords: Ear disease, Deep learning, Classification, Available models, Hybrid models

ŞEKİLLER DİZİNİ

	<u>Sayfa No</u>
Şekil 1.1. Kulak içi yapısı[2].	1
Şekil 1.2. Otitis media hastalığı [4].	3
Şekil 1.3. Kulak mumu hastalığı.	4
Şekil 1.4. Timpanoskleroz ve miringoskleroz hastalığı[7].	5
Şekil 1.5. Otitis eksterna hastalığı[8].	6
Şekil 2.1. Ear Imagery veri seti Normal (a), CSOM (b), Earwax (c) ve Myringosclerosis (d).	14
Şekil 2.2. Tympanic membrane veri seti Normal (a), AOM (b), CSOM (c), Earwax (d) ve Otitis Externa (e).	16
Şekil 2.3. YOLOv1 mimarisinin çalışması.	21
Şekil 2.4. YOLOv5 mimarisi.	23
Şekil 2.5. ResNet atlama katmanı.	24
Şekil 2.6. VGG-19 (a) ve ResNet34 (b) mimarileri.	25
Şekil 2.7. Darboğaz bloğu.	26
Şekil 2.8. Xception mimarisi.	28
Şekil 2.9. InceptionResNetV1 ve V2 mimarisi.	30
Şekil 2.10. Inception-ResNet-A (a), Inception-ResNet-B (b) ve Inception-ResNet-C (c) blokları.	31
Şekil 2.11. DenseNet mimarisi.	33
Şekil 2.12. ViT mimarisi çalışma prensibi.	35
Şekil 2.13. Transformer bloğu.	36
Şekil 2.14. LSTM ağı hafıza hücreleri.	38
Şekil 2.15. EfficientNet mimarisi.	39
Şekil 2.16. Sıkıştırma ve uyarma bloğu.	40
Şekil 2.17. Artırma işlemi sonucunda elde edilen görüntüler.	41
Şekil 2.18. RGB görüntülerin HSV (a) ve LUV (b) renk uzayına çevrilmesi sonucunda elde edilen görüntüler	44
Şekil 3.1. Kulak zarı bölgesinin tespiti; orijinal görüntü (a), kulak zarı tespiti (b) ve tespit edilen alanın kesilmesi sonucunda elde edilen görüntü (c).	48
Şekil 3.2. Özellik çıkarımı ve özelliklerin sınıflandırılması için oluşturulan mimariler.	49

Şekil 3.3.	Modellere ait eğitim ve doğrulama grafikleri, (a) eğitim kaybı, (b) eğitim doğruluğu, (c) doğrulama kaybı ve (d) doğrulama doğruluğu.....	50
Şekil 3.4.	Modellere ait test işlemi karışıklık matrisi.....	51
Şekil 3.5.	ViT mimarisi ear imagery veri seti eğitim (a) ve doğrulama (b) grafiği.	52
Şekil 3.6.	ViT mimarisi ear imagery veri seti karışıklık matrisi.	52
Şekil 3.7.	Ear imagery veri seti öğrenme aktarımı ile eğitim sonucunda elde edilen eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).....	54
Şekil 3.8.	Merged_ds veri setine ait sıfırdan eğitim sonucunda elde edilen eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).....	55
Şekil 3.9.	Merged_ds veri setine ait öğrenme aktarımıyla eğitim sonucunda elde edilen eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).	56
Şekil 3.10.	Merged_ds (6 sınıf) veri setine ait sıfırdan eğitim sonucunda elde edilen eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).....	57
Şekil 3.11.	Merged_ds (6 sınıf) veri setine ait öğrenme aktarımıyla eğitim sonucunda elde edilen eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).....	58
Şekil 3.12.	Merged_ds (2 sınıf) veri setine ait sıfırdan eğitim sonucunda elde edilen eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).....	59
Şekil 3.13.	Merged_ds (2 sınıf) veri setine ait öğrenme aktarımıyla eğitim sonucunda elde edilen eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).....	60
Şekil 3.14.	Tympanic membrane veri setine ait sıfırdan eğitim sonucunda elde edilen eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).....	61
Şekil 3.15.	Tympanic membrane veri setine ait öğrenme aktarımıyla eğitim sonucunda elde edilen eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).	62
Şekil 3.16.	YOLOv5 uygulandıktan sonra tympanic membrane veri setine ait sıfırdan eğitim sonucunda elde edilen eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).....	63
Şekil 3.17.	YOLOv5 uygulandıktan sonra tympanic membrane veri setine ait öğrenme aktarımıyla eğitim sonucunda elde edilen eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d)... ..	64
Şekil 3.18.	Ear imagery veri setinde öğrenme aktarımı kullanılarak elde edilen karışıklık matrisi.	65

Şekil 3.19.	Merged_ds veri setinde sıfırdan eğitim kullanılarak elde edilen karışıklık matrisi.....	65
Şekil 3.20.	Merged_ds veri setinde öğrenme aktarımı kullanılarak elde edilen karışıklık matrisi.	66
Şekil 3.21.	Merged_ds (6 sınıf) veri setinde sıfırdan eğitim kullanılarak elde edilen karışıklık matrisi.	66
Şekil 3.22.	Merged_ds (6 sınıf) veri setinde öğrenme aktarımı kullanılarak elde edilen karışıklık matrisi.....	67
Şekil 3.23.	Merged_ds (2 sınıf) veri setinde sıfırdan eğitim kullanılarak elde edilen karışıklık matrisi.	67
Şekil 3.24.	Merged_ds (2 sınıf) veri setinde öğrenme aktarımı kullanılarak elde edilen karışıklık matrisi.....	68
Şekil 3.25.	Tympanic membrane veri setinde sıfırdan eğitim kullanılarak elde edilen karışıklık matrisi.	68
Şekil 3.26.	Tympanic membrane veri setinde öğrenme aktarımı kullanılarak elde edilen karışıklık matrisi.....	69
Şekil 3.27.	YOLOv5 uygulanmış tympanic membrane veri setinde sıfırdan eğitim kullanılarak elde edilen karışıklık matrisi.	69
Şekil 3.28.	YOLOv5 uygulanmış tympanic membrane veri setinde öğrenme aktarımı kullanılarak elde edilen karışıklık matrisi.....	70
Şekil 3.29.	Ear imagery veri setindeki görüntülerin HSV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).	74
Şekil 3.30.	Ear imagery veri setindeki görüntülerin LUV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).	75
Şekil 3.31.	Merged_ds veri setindeki görüntülerin HSV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).	76
Şekil 3.32.	Merged_ds veri setindeki görüntülerin LUV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).	77
Şekil 3.33.	Merged_ds (6 sınıf) veri setindeki görüntülerin HSV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).	78
Şekil 3.34.	Merged_ds (6 sınıf) veri setindeki görüntülerin LUV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).	79

Şekil 3.35.	Merged_ds (2 sınıf) veri setindeki görüntülerin HSV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).....	80
Şekil 3.36.	Merged_ds (2 sınıf) veri setindeki görüntülerin LUV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).....	81
Şekil 3.37.	Tympanic membrane veri setindeki görüntülerin HSV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).....	82
Şekil 3.38.	Tympanic membrane veri setindeki görüntülerin LUV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).....	83
Şekil 3.39.	YOLOv5 uygulanmış tympanic membrane veri setindeki görüntülerin HSV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).	84
Şekil 3.40.	YOLOv5 uygulanmış tympanic membrane veri setindeki görüntülerin LUV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).	85
Şekil 3.41.	Ear imagery veri seti HSV renk uzayı karışıklık matrisleri.	86
Şekil 3.42.	Ear imagery veri seti LUV renk uzayı karışıklık matrisleri.	86
Şekil 3.43.	Merged_ds veri seti HSV renk uzayı karışıklık matrisleri.	87
Şekil 3.44.	Merged_ds veri seti LUV renk uzayı karışıklık matrisleri.	87
Şekil 3.45.	Merged_ds (6 sınıf) veri seti HSV renk uzayı karışıklık matrisleri.	88
Şekil 3.46.	Merged_ds (6 sınıf) veri seti LUV renk uzayı karışıklık matrisleri.	88
Şekil 3.47.	Merged_ds (2 sınıf) veri seti HSV renk uzayı karışıklık matrisleri.	89
Şekil 3.48.	Merged_ds (2 sınıf) veri seti LUV renk uzayı karışıklık matrisleri.	89
Şekil 3.49.	Tympanic membrane veri seti HSV renk uzayı karışıklık matrisleri.	90
Şekil 3.50.	Tympanic membrane veri seti LUV renk uzayı karışıklık matrisleri.	90
Şekil 3.51.	YOLOv5 uygulanmış tympanic membrane veri seti HSV renk uzayı karışıklık matrisleri.	91
Şekil 3.52.	YOLOv5 uygulanmış tympanic membrane veri seti LUV renk uzayı karışıklık matrisleri.	91
Şekil 3.53.	Hibrit modellerin ear imagery veri seti üzerindeki eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c), doğrulama doğruluğu (d).	102
Şekil 3.54.	Hibrit modellerin merged_ds veri seti üzerindeki eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c), doğrulama doğruluğu (d).	103

Şekil 3.55.	Hibrit modellerin merged_ds (6 sınıf) veri seti üzerindeki eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c), doğrulama doğruluğu (d).	104
Şekil 3.56.	Hibrit modellerin merged_ds (2 sınıf) veri seti üzerindeki eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c), doğrulama doğruluğu (d).	105
Şekil 3.57.	Hibrit modellerin tympanic membrane veri seti üzerindeki eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c), doğrulama doğruluğu (d).	106
Şekil 3.58.	Hibrit modellerin YOLOv5 uygulanmış tympanic membrane veri seti üzerindeki eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c), doğrulama doğruluğu (d).	107

TABLolar DİZİNİ

	<u>Sayfa No</u>
Tablo 2.1. Tympanic Membrane veri setinde bulunan sınıflar ve görüntü sayıları	15
Tablo 2.2. Çalışmada kapsamında kullanılan veri setleri.....	17
Tablo 2.3. Artırma işlemi sonucunda elde edilen görüntü sayıları	42
Tablo 3.1. İlgili modellerden elde edilen performans metrikleri.	51
Tablo 3.2. Ear imagery veri seti ağırlık değerleri performans metrikleri.....	71
Tablo 3.3. Merged_ds veri seti ağırlık değerleri performans metrikleri.	71
Tablo 3.4. Merged_ds (6 sınıf) veri seti ağırlık değerleri performans metrikleri.	72
Tablo 3.5. Merged_ds (2 sınıf) veri seti ağırlık değerleri performans metrikleri.	72
Tablo 3.6. Tympanic membrane veri seti ağırlık değerleri performans metrikleri.	73
Tablo 3.7. YOLOv5 uygulanmış tympanic membrane veri seti ağırlık değerleri performans metrikleri.....	73
Tablo 3.8. Ear imagery veri seti renk uzayları performans metrikleri.	92
Tablo 3.9. Merged_ds veri seti renk uzayları performans metrikleri.....	92
Tablo 3.10. Merged_ds (6 sınıf) veri seti renk uzayları performans metrikleri.	93
Tablo 3.11. Merged_ds (2 sınıf) veri seti renk uzayları performans metrikleri.	93
Tablo 3.12. Tympanic membrane veri seti renk uzayları performans metrikleri.....	94
Tablo 3.13. YOLOv5 uygulanmış tympanic membrane veri seti renk uzayları performans metrikleri.....	94
Tablo 3.14. Artırma uygulanmış ear imagery veri seti ağırlık değerleri performans metrikleri.....	95
Tablo 3.15. Artırma uygulanmış ear imagery veri seti renk uzayları performans metrikleri.....	96
Tablo 3.16. Artırma uygulanmış merged_ds veri seti ağırlık değerleri performans metrikleri.....	96
Tablo 3.17. Artırma uygulanmış merged_ds veri seti renk uzayları performans metrikleri.....	97
Tablo 3.18. Artırma uygulanmış merged_ds (6 sınıf) veri seti ağırlık değerleri performans metrikleri.....	97
Tablo 3.19. Artırma uygulanmış merged_ds (6 sınıf) veri seti renk uzayları performans metrikleri.....	98
Tablo 3.20. Artırma uygulanmış merged_ds (2 sınıf) veri seti ağırlık değerleri performans metrikleri.....	98

Tablo 3.21.	Artırma uygulanmış merged_ds (2 sınıf) veri seti renk uzayları performans metrikleri.....	99
Tablo 3.22.	Artırma uygulanmış tympanic membrane veri seti ağırlık değerleri performans metrikleri.....	99
Tablo 3.23.	Artırma uygulanmış tympanic membrane veri seti renk uzayları performans metrikleri.....	100
Tablo 3.24.	Artırma ve YOLOv5 uygulanmış tympanic membrane veri seti ağırlık değerleri performans metrikleri.	100
Tablo 3.25.	Artırma ve YOLOv5 uygulanmış tympanic membrane veri seti renk uzayları performans metrikleri.....	101
Tablo 3.26.	Hibrit modellerin ear imagery veri seti performans metrikleri.	108
Tablo 3.27.	Hibrit modellerin merged_ds veri seti performans metrikleri.	108
Tablo 3.28.	Hibrit modellerin merged_ds (6 sınıf) veri seti performans metrikleri.....	109
Tablo 3.29.	Hibrit modellerin merged_ds (2 sınıf) veri seti performans metrikleri.....	109
Tablo 3.30.	Hibrit modellerin tympanic membrane veri seti performans metrikleri. ...	110
Tablo 3.31.	Hibrit modellerin YOLOv5 uygulanmış tympanic membrane veri seti performans metrikleri.....	110
Tablo 3.32.	Literatür karşılaştırılması.	112

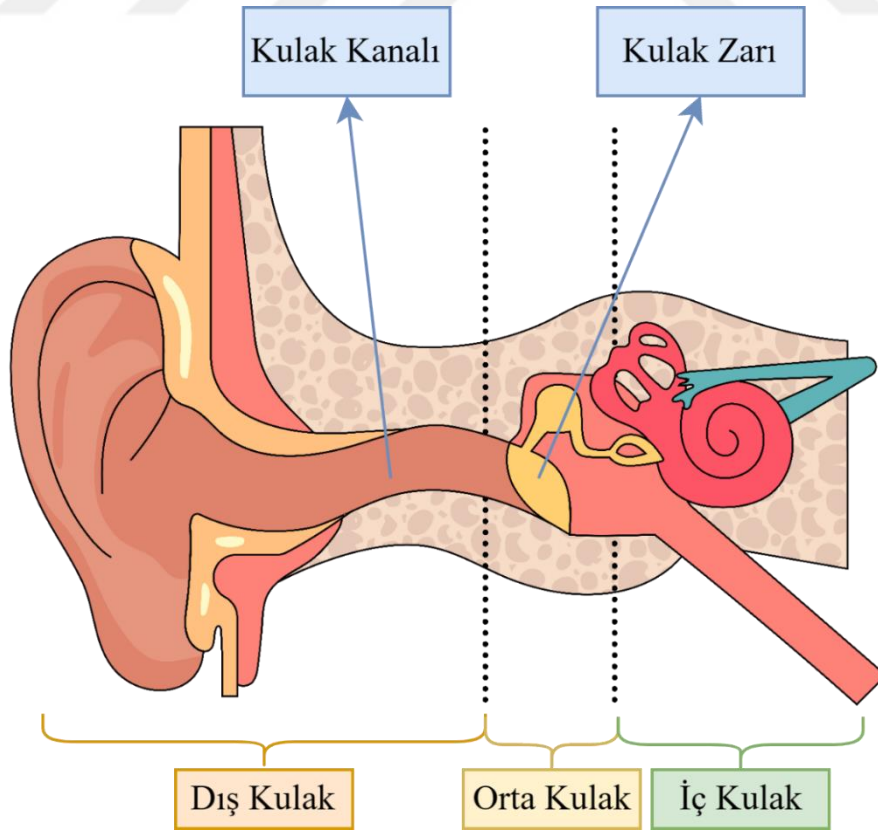
SEMBOLLER DİZİNİ

ResNet	: Residual Network (Kalıntı Ağı)
CNN	: Convolutional Neural Network (Evreşimsel Sinir Ağları)
MLP	: Multi-Layer Perceptron (Çok-Katmanlı Algılayıcılar)
FCN	: Fully Connected Layer (Tam-Bağlı Katman)
FPR	: False Positive Rate (Yanlış Pozitif Oranı)
TPR	: True Positive Rate (Doğru Pozitif Oranı)
ROC	: Receiver Operating Characteristic (İşlem Karakteristiği Eğrisi)
AUC	: Area Under the Curve (Eğri Altındaki Alan)
DNN	: Deep Neural Network (Derin Sinir Ağları)
LSTM	: Long Short-Term Memory (Uzun Kısa Süreli Hafıza)
BiLSTM	: Bidirectional Long Short Term Memory (Çift Yönlü Uzun Kısa Hafıza)
ConvBiLSTM	: Convolutional Bidirectional Long Short Term Memory (Evrişimsel Çift Yönlü Uzun Kısa Hafıza)
MHA	: Multi Head Attention (Çok Başlı Dikkat)

1. GENEL BİLGİLER

1.1. Giriş

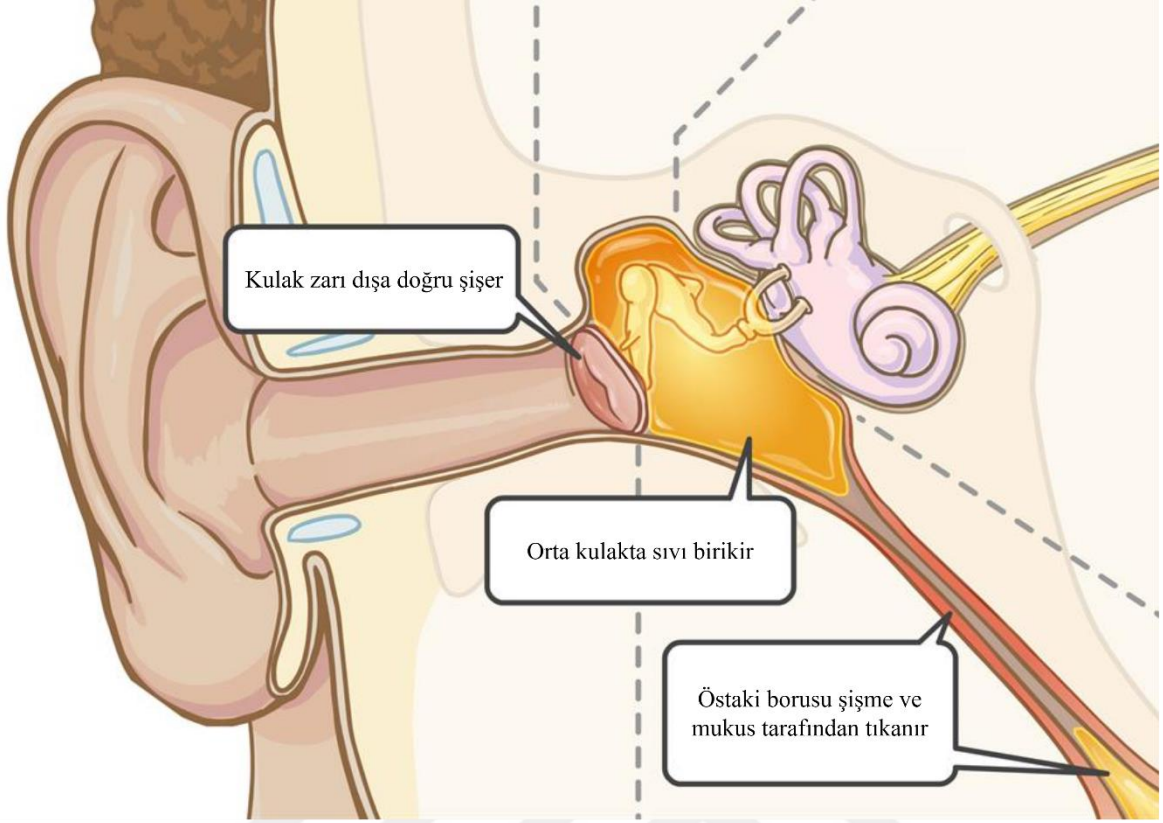
Canlı türlerinde kulak yapısında farklılıklar görülmektedir. Buna karşın sesin duyulma prensibi aynıdır. Kulak yapısı içerisinde üç ana bölüm bulunmaktadır; dış, orta ve iç. Kulak zarı, dış kulak ile orta kulağı ayıran ince, yarı saydam zardır. Ses dalgaları dış kulaktan içeri girdikten sonra kulak kanalından geçerek kulak zarına ulaşmaktadır. Ses dalgaları kulak zarını titreterek orta kulak içerisinde bulunan üç küçük kemiğe (malleus, incus ve stapes) aktarılır. Kemik yapısı titreşimi salyangoz benzeri içi sıvı dolu cochlea'ya aktarır. Sıvının dalgalanması sensörleri tetikler ve işitsel sinirler sinyalleri beyne taşıyarak sesin duyulmasını sağlar[1]. Şekil 1.1'de kulak içi yapıca gösterilmektedir. Genel olarak, kulak hastalıkları üç ana bölgeden birinde görülerek sadece ilgili alanı etkilemektedir. Bazı durumlarda hastalık geç tedavi edildiğinde veya tedavi edilmediğinde hastalık, bölümler arası geçişler veya yayılmalar göstermektedir.



Şekil 1.1. Kulak içi yapısı[2].

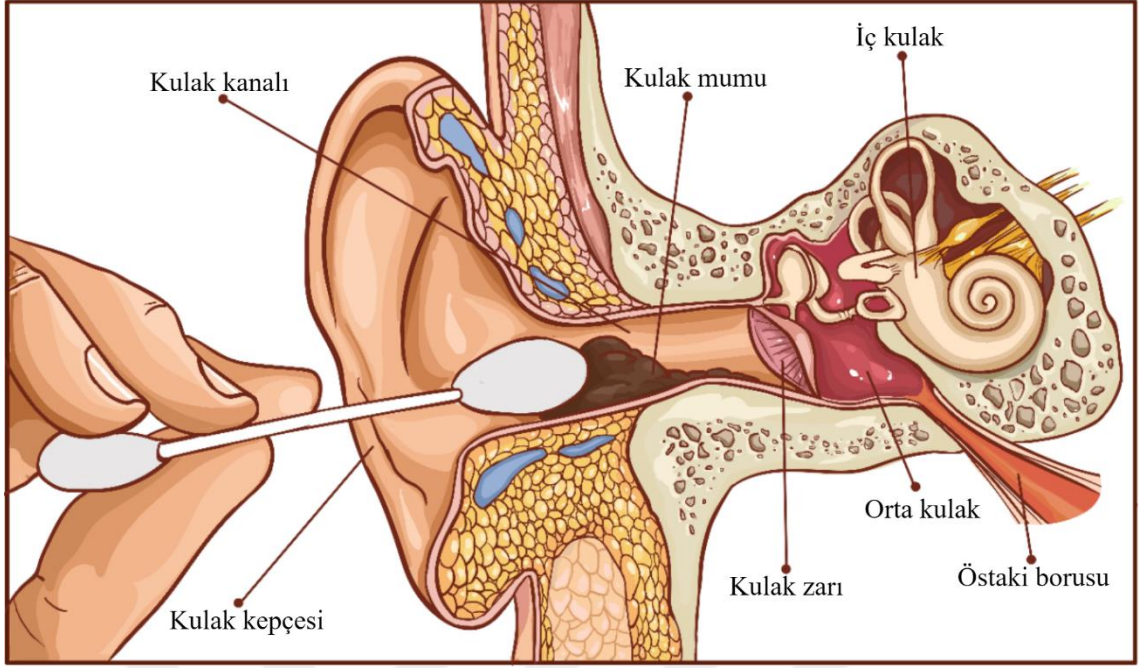
Kulak hastalıkları, insanların yaşadığı yaygın sağlık sorunlarından biridir. Kulak hastalıkları işitme mekanizmasının gerçekleştiği üç ana bölgeyi etkilemektedir. Bu hastalıklar kulağın yapı ve fonksiyonlarını etkileyerek işitme kaybı, kulak çınlaması, denge bozuklukları ve diğer semptomlara neden olmaktadır. Genetik, enfeksiyonlar, yaralanmalar, yüksek sese maruz kalma ve yaşlanma gibi faktörler kulak hastalıklarına neden olmaktadır.

Otitis Media (OM), orta kulak boşluğunda virüslerin veya bakterilerin enfeksiyona veya kulak zarı arkasında sıvı birikmesine neden olduğu hastalık türüdür. OM hastalığına neden olan en yaygın bakteriyel organizmalar; *Streptococcus pneumoniae*, tiplendirilemeyen *Haemophilus influenza* (NTHi) ve *Moraxella catarrhalis*. Bununla birlikte OM hastalığına neden olan en yaygın viral patojenler; respiratuvar sinsityal virüs (RSV), koronavirüsler, grip virüsler, adenovirüsler ve picornavirüslerdir. OM hastalığı genellikle 7 yaş altındaki çocuklarda görülmektedir. Bebeklerde hastalığın anlaşılması kendilerini ifade edemediklerinden dolayı daha zordur. Küçük yaştaki çocukların gelişim sürecinde görüldüğünden tedavi edilmediğinde işitme kaybına veya konuşma bozukluğuna neden olabilmektedir. OM hastalığı, enfeksiyonun süresine göre akut, subakut veya kronik olarak sınıflandırılmaktadır. Genellikle üç haftadan daha az süren orta kulak iltihapları akut otitis media (AOM), 3 aydan uzun süren orta kulak iltihapları kronik otitis media (CSOM) adı verilmektedir[3]. AOM hastalığında genel semptomlar ateş, kulak ağrısı, otore, kusma ve ishal şeklinde görülmektedir. CSOM hastalığının semptomları AOM hastalığına benzemektedir fakat şiddeti ve süresi daha uzundur. Ayrıca AOM hastalığından farklı olarak CSOM hastalığında kulak zarında yırtık veya delik oluşmaktadır. Şekil 1.2’de otitis media hastalığının etki alanı gösterilmektedir. Genellikle hastalık iç kulak ile orta kulak arasında gerçekleşmektedir.



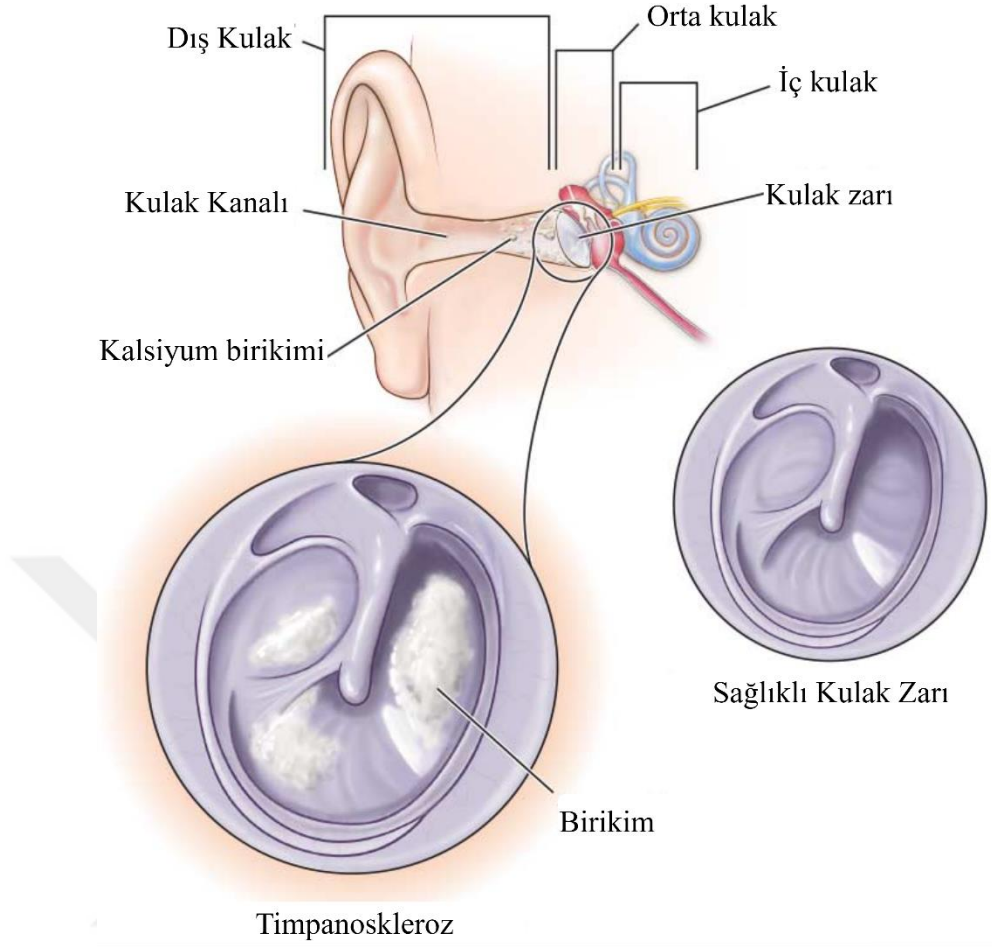
Şekil 1.2. Otitis media hastalığı [4].

Kulak mumu, kulak kanalı sağlığı için önemli savunma mekanizmasıdır. Kulak mumu, kulağın içerisinde toz, toprak birikimini ve enfeksiyon oluşumunu önlemektedir. Birçok faydasının yanında, kulak mumunun çevresel ve sağlıksal nedenlerden dolayı aşırı üretimi kulak mumu tıkaçı (*Cerumen impaction*) hastalığına sebep olmaktadır[5]. Kulak mumu tıkaçı, her yaş grubunda görülebileceği gibi Dünya genelinde çocukların %10'u ile yetişkinlerin %5'inde bulunmaktadır. Hastalık işitme cihazı, kulak tıkaçı, kulaklık kullanımında, kulak içerisinde aşırı kıl bulunmasında, egzama gibi cilt problemleri sonucunda orta çıkmaktadır. Kulak mumu tıkaçının semptomları, kulak içerisinde doluluk hissi, kulak ağrısı, zamanla kötüleşen işitme kaybı, kulak çınlaması, kulak kaşınması, kulak akıntısı ve baş dönmesi şeklindedir. Hastalık tedavi edilmediğinde bu semptomlar kötüye gitmektedir. Şekil 1.3'te hastalık şematize edilmektedir.



Şekil 1.3. Kulak mumu hastalığı.

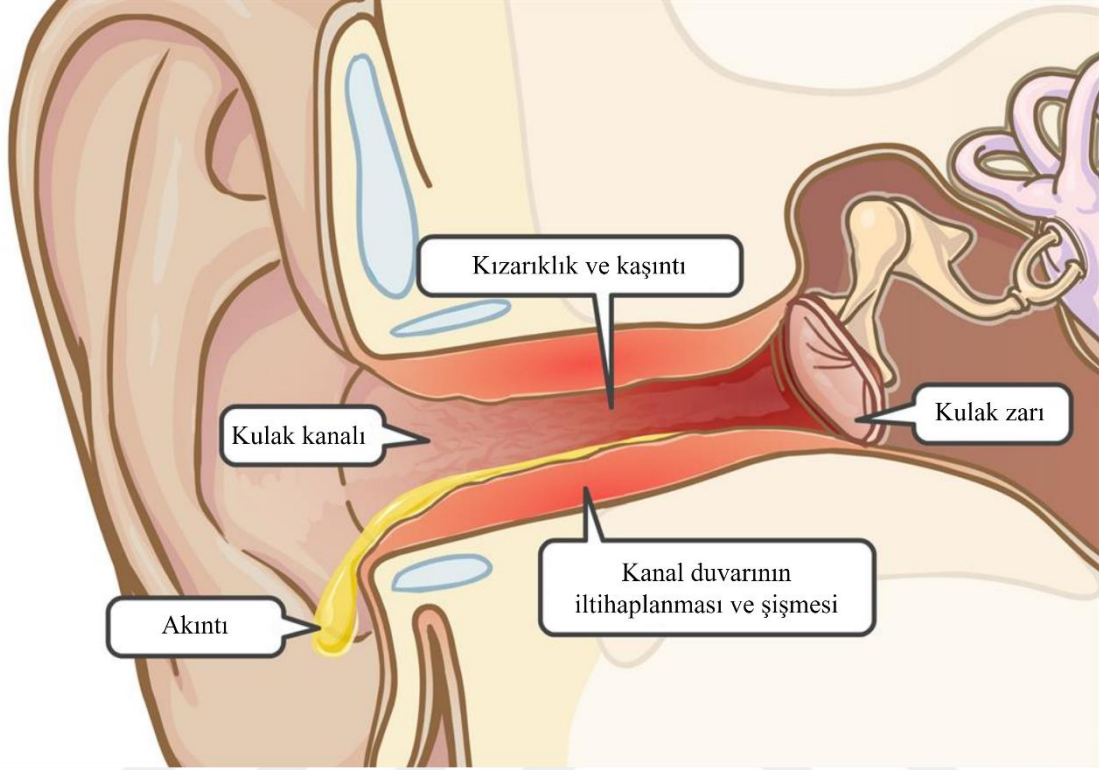
Kulak zarının hasar alması veya enfeksiyon kapması sonucunda (yaralanma veya ameliyat) kalsiyum birikmesi görülmektedir. Kalsiyumun birikimi kulak zarının kalınlaşmasına, esnekliğini kaybetmesine ve beyazlamasına neden olmaktadır. Bu hastalığa tympanosclerosis veya myringosclerosis denmektedir. Bu iki kavram arasındaki fark biriken kalsiyum miktarı ve biriktiği kulak bölümüdür. Bu hastalıkların semptomları, kulak zarının kireçli beyaz görünümü, işitme kaybı (sertleşme sonucunda titreminin azalmasıyla sesin iç kulağa aktarılması engellemektedir) ve kulak ağrısıdır[6]. Çocuklarda sıklıkla görülmesine de CSOM veya AOM sonucunda enfeksiyon kapmış kulağa timpanostomi tüpü yerleştirilmesi sonucunda %50'sinde miringoskleroz görülmektedir. Şekil 1.4'te timpanoskleroz ve miringoskleroz hastalıklarının kulak üzerindeki etkisi modellenmektedir.



Şekil 1.4. Timpanoskleroz ve miringoskleroz hastalığı[7].

Otitis Eksterna, dış kulak ile orta kulak arasında oluşan enfeksiyonel bir hastalıktır. Hastalık çoğunlukla yüzücülerde görüldüğünden dolayı “yüzücü kulağı” olarak da adlandırılmaktadır. Kulak kanalında nemin hapsolması sonucunda bakteri veya mantar büyümesi gerçekleşmektedir. Hastada ilerleyen süreçte bakteri veya mantar oluşumundan kaynaklı enfeksiyon gözlenmektedir. Otitis eksterna enfeksiyonuna yol açan diğer faktörler arasında cilt alerjisi, egzama gibi cilt rahatsızlıkları, kulak tıkacı, işitme cihazı veya kulaklık kullanımı yer almaktadır. Otitis eksterna semptomları hafif veya şiddetli gerçekleşmektedir. Hastalığın bazı semptomları kulak ağrısı, kulak içinde kaşıntı, kızarıklık, şişme, akıntı ve kulakta dolgunluk veya tıkanma hissidir. Şiddetli gerçekleştiğinde enfeksiyon çevre dokulara yayılarak boyundaki lenf düğümlerine sıçramaktadır. Tedavi edilmediğinde hastalık, iç kulağa ilerleyerek kulak zarının delinmesine ve üç küçük kemiği enfekte etmektedir. Otitis eksterna hastalığı antibiyotik veya antifungal kulak damlalarının kullanımı

ile tedavi edilmektedir. Şekil 1.5’te otitis eksterna hastalığı sonucunda oluşan enfeksiyonel akıntı gösterilmektedir.



Şekil 1.5. Otitis eksterna hastalığı[8].

Kulak hastalıklarının tanısı, Kulak Burun Boğaz (KBB) uzmanları tarafından otoskop kullanılarak gerçekleştirilmektedir. Pnömatik otoskop veya dijital otoskop, genellikle KBB uzmanlarının kulak hastalıklarını teşhis etmek için kullandığı tek araçtır. Bu alet kullanılarak kulak içinin gözlemlenmesi sonucunda uzmanlar hastalık teşhisi koymaktadır. Bazı hastalık tedavi yöntemleri şu şekildedir.

1.2. Bekle ve Gör Yaklaşımı (Wait-And-See Approach)

Kulak hastalığına ait semptomlar ilk günlerde artmakla birlikte çoğunlukla bir haftada kendiliğinden geçmektedir. Bundan dolayı herhangi bir tedavide bulunmadan önce uzmanlar semptomların gelişimini gözlerler. Eğer hastalık iki haftada gerileme göstermez ise tedavi uygulanmaktadır.

1.3. Antibiyotik Tedavisi (Antibiotic Therapy)

Gözlem sonucunda uzmanlar, enfeksiyonun giderilmesi için antibiyotik tedavisine başvurabilmektedir. 6 aylıktan küçük çocuklarda AOM hastalığı doğrulandığında gözlemsel bekleme süresi olmaksızın antibiyotik tedavisine başvurulur. CSOM hastalığında kulak zarında yırtılma veya delik oluştuğundan dolayı tedavisi daha zordur. Hastalığın tedavisi, kulak kanalından enfeksiyon emildikten sonra antibiyotik damlalarıyla gerçekleştirilir.

1.4. Ameliyat

Kulak hastalıklarının sonucunda oluşan işitme kaybı, enfeksiyon ve yapısal anormallikler kişinin günlük yaşamını veya yaşam kalitesini olumsuz etkilemektedir. Bu hastalıkların tedavisi için KBB uzmanları ameliyata başvurmaktadır. Bazı ameliyat yöntemleri miringoplasti, mastoidektomi ve koklear implant olarak verilebilmektedir. Miringoplasti, kulak zarındaki deliğin veya bozukluğun onarılması için uygulanmaktadır. Mastoidektomi, kulağın arkasındaki mastoid kemiğe yayılmış enfeksiyonun temizlenmesi için gerçekleştirilmektedir. Koklear implant, işitme kaybının giderilmesi için kulak zarının işlevini görecektir bir cihazın ameliyat ile kulak içine yerleştirilmektedir. Bu ameliyatlarda alanında uzman ve deneyimli kulak cerrahları tarafından gerçekleştirilmektedir.

2. YAPILAN ÇALIŞMALAR

Kulak hastalıkları insan hayatını olumsuz şekilde etkilemektedir. Hastalığın çocukluk döneminde görülmesinde tedavi çok önemlidir ve tedavi edilmeyen veya geç tedavi edilen hastalıklar ileriki yaşlarda kalıcı problemlere yol açmaktadır. Günümüze kadar kulak hastalıklarının tanısında kullanılmak üzere derin öğrenme ve makine öğrenmesi tabanlı otomatik sınıflandırma sistemleri geliştirilmiştir. Literatürde bu konu hakkında yapılmış çalışmalar bulunmaktadır.

Wu ve arkadaşları pediatrik orta kulak iltihaplarının kolay izlenmesi ve sınıflandırılmasında, arka planında sınıflandırma işlemi için ImageNet üzerinde önceden eğitilmiş Xception ve MobileNet-V2 evrimsel sinir ağı (CNN) mimarilerini kullanan bir mobil uygulama önermiştir. Önerilen metodun eğitim ve test işlemlerinde Shenzhen Çocuk Hastanesinden Ocak 2016 ile Aralık 2019 yılları arasında elde edilen üç tip otitis media hastalığını barındıran 12203 görüntü üzerinde uygulanmıştır. Yapılan çalışma sonucunda Xception ve MobileNet-V2 mimarileri sistem üzerindeki test görüntülerinde sırasıyla %97,82 ve %96,04; mobil uygulama üzerindeki test görüntülerinde ise sırasıyla %90,85 ve %88,89 doğruluk elde edilmiştir [9].

Sundgaard ve arkadaşları otoskop görüntülerindeki otitis media hastalığına otomatik tanı koyabilen ImageNet üzerinde önceden eğitilmiş Inception V3 mimarisinin önermiştir. Önerilen metodun eğitim ve test işlemlerinde Japonya Şizuoka şehrindeki Kamide ENT kliniğindeki 519 hastaya ait üç tip otitis media hastalığının 1136 görüntüsü üzerinde gerçekleştirilmiştir. Çalışma içerisinde beş farklı kayıp fonksiyonu karşılaştırılmış, en iyi başarı %85 doğruluk oranıyla üçlü kayıp fonksiyonundan elde edilmiştir [10].

Alhudhaif, Z. Cömert ve K. Polat beş tipteki otitis media hastalığının genelleştirilmiş karar mekanizması geliştirilmiş CNN mimarisini önermiştir. Önerilen mimarinin eğitim ve test işlemlerinde Özel Van Akdamar Hastanesinden Ekim 2018 ile Haziran 2019 yılları arasında uzmanlar tarafından sekiz tip hastalığa ayrılan 956 otoskop görüntüsünün bulunduğu açık erişimli veri seti kullanılmıştır. CNN mimarisinin karar mekanizması kanal ve uzamsal model (CBAM), artık bloklar ve hiper sütun tekniğinin birleşimi kullanılarak geliştirilmiştir. Aynı veri seti üzerinde AlexNet, VGG-16, VGG-19, GoogleNet, ResNet-50

ve önerilen metot dört sınıf için sırasıyla %81,40, %81,40, %84,88, %80,23, %80,23 ve %98,26 doğruluk göstermiştir [11].

Tran ve arkadaşları pediatrik otitis media hastalığına ait görüntülerinin otomatik tanı koyabilen sistem için segmentasyon işlemi seviye seti tabanlı aktif kontör metodu, özellik çıkarımı renk, HOG ve SSIM özellikleri ve hastalığın sınıflandırılması ortak seyrek temsil algoritması ile yapılması önerilmiştir. Önerilen metodun eğitim ve test işlemlerinde Genel Cathay Hastanesindeki çocuklardan dijital otoskop ile elde edilmiş uzmanlar tarafından iki tip hastalık altında tanı konulmuş 1230 görüntü kullanılmıştır. Önerilen metot test veri setinde %91,41 doğruluk göstermiştir [12].

Myburgh ve arkadaşları beş tip otitis media hastalığı için otomatik tanı koyabilen bulut tabanlı karar ağaçları ve sinir ağ yapısını kullanan mobil uygulamayı önermiştir. Önerilen metodun eğitim ve test işlemlerinde farklı kaynaklardan elde edilmiş beş tip otitis media hastalığını barındıran 562 video otoskop görüntüsü kullanılmıştır. Önerilen karar ağacı ve sinir ağı sırasıyla %81,58 ve %86,84 doğruluk göstermiştir [13].

Mohammed ve arkadaşları, CNN-EfficientNetB0 özellik çıkarıcısı, Bayes optimizasyon hiper parametre seçici ve CNN-BiLSTM sınıflandırıcı kullanılarak otomatik olarak otitis media hastalığına tanı koyabilen sistem önermiştir. CNN-BiLSTM mimarilerinin birlikte kullanılması yalnızca CNN mimarisine dayanan metotlara göre daha iyi performanslı ve daha hızlı çalışmıştır. Önerilen metodun eğitim ve test işlemlerinde Universidad de Chile Klinik Hastanesi Kulak Burun Boğaz bölümünden elde edilen dört tip hastalık barındıran 880 otoskop görüntüsü kullanılmıştır. Önerilen mimari %99,62 doğruluk göstermiştir [14].

Uçar ve arkadaşları otitis media hastalığına ait görüntülerin SURF anahtar noktaların VGG-16 mimarisine verilmesiyle beş farklı katmanından elde edilen hiper-parametrelerin BiLSTM ağına beslenerek sınıflandırılması önerilmiştir. Önerilen mimarinin eğitim ve test işlemlerinde Universidad de Chile Klinik Hastanesi Kulak Burun Boğaz bölümünden elde edilen 880 görüntü kullanılmıştır. SURF kullanarak anahtar noktalar elde edilmiştir ve anahtar noktalar sayesinde sistem renk uzayından bağımsızlaştırılmıştır. Önerilen mimari ortalama %99,06 doğruluk göstermiştir [15].

Wang ve arkadaşları kulak içi hastalıklarını sınıflandırmak için MESIC mimarisini önermişlerdir. Veri seti, Xiangya Hastanesindeki hastalara ait CT taramasından elde edilen üç sınıfa ait toplam 998 görüntüden oluşmaktadır. Önerilen mimaride görüntüler kullanılmadan önce ön işleme gerçekleştirilmiştir. Ön işlemden, 100x100 piksel boyutunda

ilgi alanı çıkarılmıştır. Önerilen mimaride, görüntülerden VGG16 modeliyle özellikleri çıkarılmıştır. Çıkarılan özelliklerden ilgi bloğuyla (attention block) üç yeni özellik haritası elde edilmiştir. Özellik haritalarının sınıflandırılmasında tam-bağlı katman ve softmax kullanılmıştır. Önerilen mimari, veri seti üzerinde %90,1'lik doğruluk göstermiştir [16].

Cha ve arkadaşları kulak içi hastalıklarının otomatik tanı konulması için farklı modellerin sınıflandırma performansları karşılaştırılmıştır. Veri seti, Severance Hastanesindeki hastalardan elde edilen altı sınıfa ait 10544 otoendoskopik görüntü kullanılmıştır. Görüntüler, eğitim ve test işlemleri için kullanılmak üzere sırasıyla %80 ve %20 olarak ayrılmıştır. Çalışma kapsamında açık derin öğrenme modellerinden, AlexNet, GoogLeNet, ResNet, Inception-V3, Inception-ResNet-V2, SqueezeNet ve MobileNet-V2 seçilmiştir. Veri setinin modeller üzerindeki uygulaması sonucunda en yüksek başarı gösteren iki model seçilmiştir. İki modelin çıktılarını birleştiren olasılık seçici önerilmiştir. Olasılık seçici iki modelden yüksek skora sahip olan etikete göre görüntüyü sınıflandırmaktadır. Çalışma sonucunda önerilen model %93,67 doğruluk göstermektedir [17].

Chen ve arkadaşları kulak içi hastalıkları sınıflandırılabilen AI-destekli mobil algoritma önermiştir. Veri seti olarak, Taipei Veteran General Hastanesinden elde edilen 10 sınıfa ait 2161 görüntü kullanılmıştır. Görüntüler, eğitim ve test işlemlerinde kullanılmak üzere sırasıyla %80 ve %20 olarak ayrılmıştır. Görüntülere ön işlem olarak “augmentation” işlemi uygulanmıştır. Derin öğrenme modeli olarak VGG19, VGG19, Xception, InceptionV3, NASNetLarge ve ResNet50 seçilmiştir. CAM (Class Activation Map) kullanılarak modellere ait özellik haritaları çıkarılmıştır. En yüksek doğruluğu %98 ile InceptionV3 modeli göstermiştir. Mobil sistemde kullanılması için NasNetMobile, MobileNetV2 ve DenseNet201 modellerine InceptionV3 modelinden elde edilen ağırlık değerleri verilmiştir. Bu çalışma sonundan mobil sistem için en yüksek başarı %97,6 ile MobileNetV2 modelinden elde edilmiştir. Önerilen çalışma, mobil sistemin uzmanlar kadar başarılı tanı koyabileceğini göstermektedir [18].

Çalışkan, otitis media hastalığının sınıflandırılması için VGG16 özellik çıkarıcı ve SVM sınıflandırıcı modelini önermiştir. Veri seti olarak Van Akdamar Hastanesindeki hastalardan elde edilen 956 iç kulak görüntüsü kullanılmıştır. Veri setindeki görüntüler normal ve normal olmayan (hastalıklı) olarak iki sınıfa ayrılmıştır. Önerilen modelde VGG16 mimarisinin farklı tam-bağlı katmanlarından elde edilen özellikler SVM modeliyle

sınıflandırılmıştır. En yüksek doğruluk değeri %82,17 ile fc6 tam-bağlı katmanı ve SVM modelinden elde edilmiştir (Çalışkan, 2022).

Sundgaard ve arkadaşları, otitis media hastalığını sınıflandırmak için geniş bant timpanometri (WBT) sinyal değerlerini kullanan mimari önermiştir. Geniş bant timpanometri, frekans sinyallerini kullanarak kulak kanalının akustik özelliklerini ölçmektedir. Sağlıklı kulak kanalı, hastalıklı olana göre daha yüksek emilim değerleri göstermiştir. Çalışmada veri seti olarak, Japonya'da Kamide KBB kliniğinde 2 aylık ile 12 yaşındaki hastalardan elde edilen 1014 WBT ölçümleri kullanılmıştır. Ölçümler Titan sistemi kullanılarak gerçekleştirilmiştir. Ölçümler sonucu 488 sağlıklı (NOE), 372 OME ve 154 AOM olmak üzere üç sınıf veri elde edilmiştir. Veri seti %80 eğitim ve %20 test için ayrılmıştır, eğitim veri seti tekrardan %20 ayrılarak doğrulama işlemi için kullanılmıştır. İlk olarak iki sınıf (sağlıklı ve OM) için AlexNet mimarisinden esinlenen model önerilmiştir. Önerilen model, PyTorch kütüphanesinde bulunan önceden eğitilmiş VGG, ResNet ve InceptionV3 ile karşılaştırılmıştır. Önceden eğitilen modeller, görüntü bazlı eğitildiğinden dolayı başarıları düşüktür. Önerilen modelin doğruluk başarısı %92,6'dır. İkinci yaklaşımda OM hastalığının ayrımı yapılmaya çalışılmıştır fakat ayrımı yapılamamıştır [20].

Tsutsumi ve arkadaşları, kulak hastalıklarının otomatik sınıflandırılması için web tabanlı derin öğrenme modeli önermiştir. Çalışmada kullanılmak üzere veri seti olarak internette açık olan ve Google Image görüntülerinin birleşimi kullanılmıştır. Veri seti, 116 AOM, 44 otitis eksterna, 23 CSOM ve 21 kulak kiri sıkışması olmak üzere toplamda 400 görüntüden oluşturulmuştur. Veri seti kategorik ve ikilik olarak iki farklı sınıf sayısına ayrılmıştır. Veri seti %75 eğitim ve %25 test için ayrılmıştır. Sınıflandırma için CNN mimarilerinden ResNet50, InceptionV3, InceptionResNetV2 ve MobileNetV2 seçilmiştir. Çalışma sonucunda ikili sınıf için en yüksek doğruluk başarısı %90 ile MobileNetV2 modelinden elde edilmiştir. Kategorik sınıf için MobileNetV2 modeli %91 doğruluk başarısı göstermiştir [21].

Wang ve arkadaşları, COM hastalığının tespiti için CT görüntülerini kullanan derin öğrenme modeli önermiştir. Veri seti olarak, 562 hastanın 672 CT taramasından elde edilen görüntüler kullanılmıştır. Görüntüler, 128-kanal çoklu-dedektör SOMATOM Definition Edge CT tarayıcıdan elde edilmiştir. Şakak kemiğinin aksenel görüntüsü iki kulağında resimlerde bulunmasını sağlamıştır ve toplamda 1344 kulak görüntüsü elde edilmiştir. İki kulağın uzmanlar tarafından incelenmesi sonucunda CSOM, normal ve kolesteatoma olmak üzere 3 sınıfa ait 1147 görüntüden oluşan son veri seti çıkarılmıştır. Veri setindeki

görüntülerin %85'i eğitim ve doğrulama, %15'i test için kullanılmıştır. Veri setinde kulak bölgesinin çıkarılması için Faster R-CNN modeli kullanılmıştır. Veri setindeki görüntülere artırma (augmentation) uygulanarak görüntü sayısı artırılmıştır. Özellik çıkarımı için InceptionV3 modeli ve sınıflandırma için softmax fonksiyonu kullanılmıştır. Önerilen model ve altı uzmanın test verisi üzerindeki performansları karşılaştırılmıştır. Önerilen modelin iki sınıf (CSOM ve Normal) üzerindeki ortalama doğruluk başarıları %86'dır. Uzmanların ortalama doğruluk başarıları %82'dir [22].

Byun ve arkadaşları, kulak içi hastalıkların derin öğrenme yöntemiyle sınıflandırılmasını önermiştir. Veri seti olarak Ocak 2015 ile Aralık 2018 tarihlerinde KBB uzmanlarını ziyaret etmiş hastalardan elde edilen 2272 görüntü kullanılmıştır. Veri setinde, normal, OME, COM veya kolesteatoma olmak üzere 4 sınıfa bulunmaktadır. Görüntüler otoskop cihazından elde edildiğinde dışında siyah çerçeveler bulunmaktadır ve görüntüler 270x270 boyutunda yeniden biçimlendirilerek siyah çerçeveler kesilmiştir. Rastgele dönüştürmelerle görüntü sayıları artırılmıştır. Önerilen model ResNet alt yapısına ve dikkat modülüne sahiptir. Model olarak ResNet18 ve ResNet50 kullanılmıştır. Dikkat modülü olarak CBAM ve 3D çıktı üreten Shuffle Attention Module önerilmiştir. Yapılan çalışmalar sonucunda ResNet18+Shuffle Attention Module %97,18 doğruluk başarıları göstermiştir [23].

Crowson ve arkadaşları, yanlış tanı sonucunda çocuklarda kalıcı etkilere yol açabilecek AOM ve COM hastalıklarının tanısı için derin öğrenme modeli önermiştir. Veri seti olarak Kasım 2019 ile Eylül 2020 yılları arasında hastalardan elde edilmiş 338 görüntüden oluşmaktadır. Veri seti, IMAGE1 HD kamera kafasına bağlı 0-derece sert fiberoptik endoskop ile elde edilmiştir ve 1920x1080 piksel çözünürlüktedir. Görüntüler normal (126 görüntü) ve normal olmayan (hastalıklı- 212 görüntü) olarak iki sınıfa ayrılmıştır ve %80'i eğitim, %20'si test için kullanılmıştır. Görüntülerin sınıflandırılması için ImageNet veri setinde önceden eğitilmiş ResNet34 mimarisi kullanılmıştır. Önerilen mimari %83,8 doğruluk başarıları göstermiştir [24].

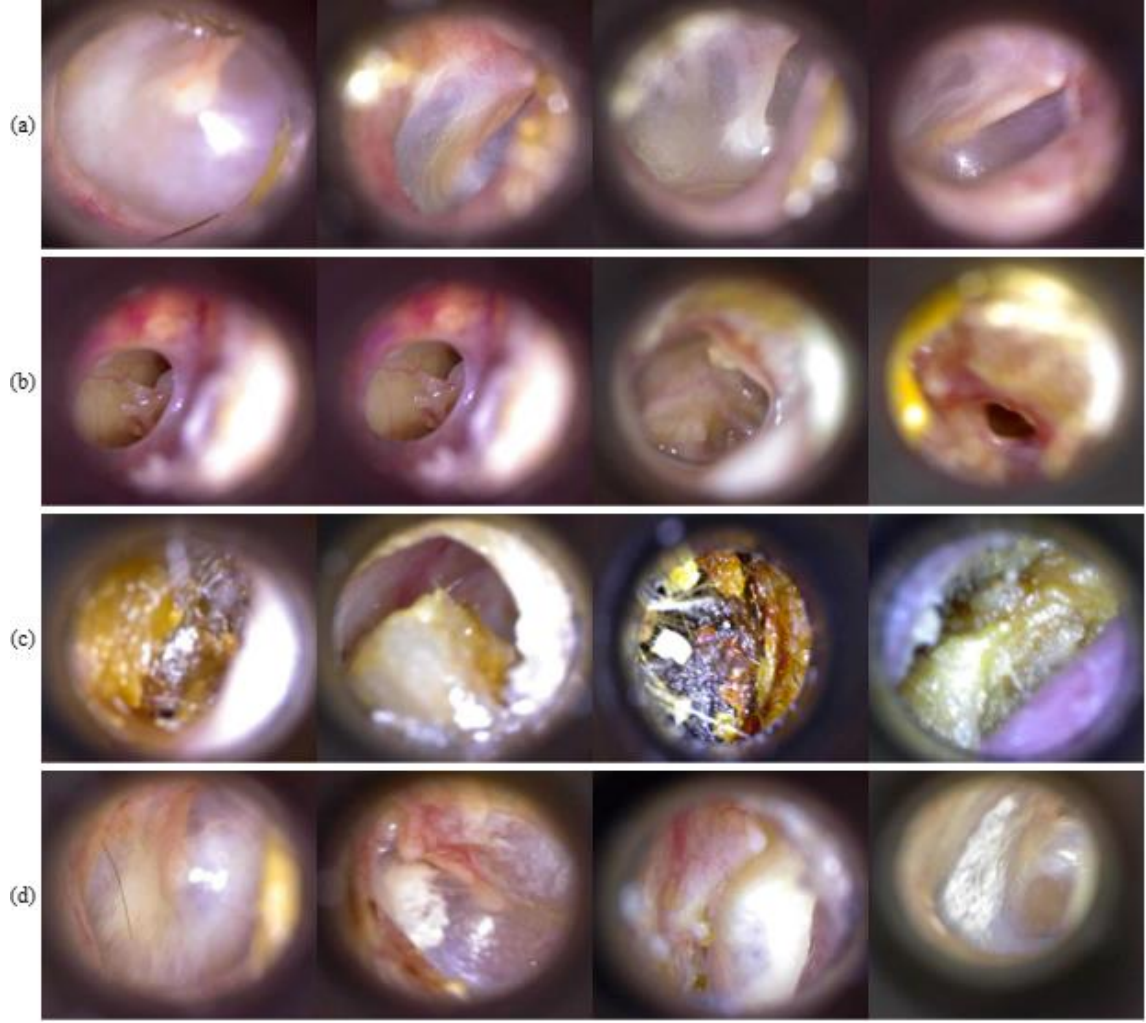
Livingstone ve arkadaşları, KBB uzmanlarına erişimi olmayan bölgelerde kullanılmak üzere derin öğrenme tabanlı otonom tanı koyan sistem önermiştir. Çalışmada kullanılan veri seti Alberta Sağlık Servisinde ENTiD dijital endoskopu kullanılarak hastalardan toplanan görseller, Google Image görselleri ve açık yayınlanmış veri setlerindeki görseller birleştirilerek oluşturulmuştur. Veri seti normal, AOM, kolesteatoma, serumen, SOM, eksostoz, miringit, miringoskleroz, otomikoz, otitis eksterna, delik (perforation), kulak zarı çekilmesi ve timpanostomi tüpü (mevcut ve pozisyonunda) olmak üzere 14 sınıfa ait toplamda

1366 görüntüden oluşmaktadır. Önerilen model olarak Google Cloud AutoML platformu kullanılmıştır. Önerilen model %90.9 başarı göstermiştir [25].

Kulak hastalıklarının tanısı KBB uzmanları tarafından otoskop kullanılarak kulak zarının incelenmesi sonucunda gözlem ile konulmaktadır. Tanının doğruluğu KBB uzmanının bilgi ve tecrübe seviyesine bağlıdır. Subjektif olarak sadece uzman gözlemiyle konulan tanımlar insan kaynaklı hatalara açıktır. Yanlış tanı hastalığın tedavisinin hatalı yapılmasına ve bunun sonucunda hasta üzerinde geri dönüşümlü problemlere yol açmaktadır. Önceki yapılan çalışmalarda derin öğrenme yöntemlerinin uzmanlara göre daha başarılı hastalık tanısı koyabildiği görülmüştür. Literatürde bulunan çalışmalar ışığında bu tez çalışmasında, kulak içi hastalıkların sınıflandırılmasını otomatik gerçekleştirebilecek bir karar destek sisteminin oluşturulması amaçlanmıştır. Karar destek sistemi için en uygun modelin seçilebilmesi için çalışma kapsamında farklı modellerin başarıları karşılaştırılmıştır.

2.1. Veri Seti (Dataset)

Çalışma kapsamında “Ear Imagery” ve “Tympanic Membrane” olmak üzere kulak hastalıklarını içeren iki farklı veri seti kullanılmıştır. Ear imagery veri seti, Universidad de Chile Klinik Hastanesi Kulak Burun Boğaz bölümünden elde edilen 7-65 yaş arasındaki 180 hastaya ait toplamda 880 görüntü içermektedir. Görüntüler bilgisayara bağlı 0,4 ve 0,5 mm spekulumlu DE500 Firefly dijital otoskop ile elde edilmiştir. Görüntüler 640x480 boyutundadır. Veri setinde normal, CSOM, miringoskleroz ve kulak mumu tıkaçı olmak üzere dört sınıf bulunmaktadır [26]. Şekil 2.1’de veri setindeki sınıflara ait örnek görüntüler verilmektedir.



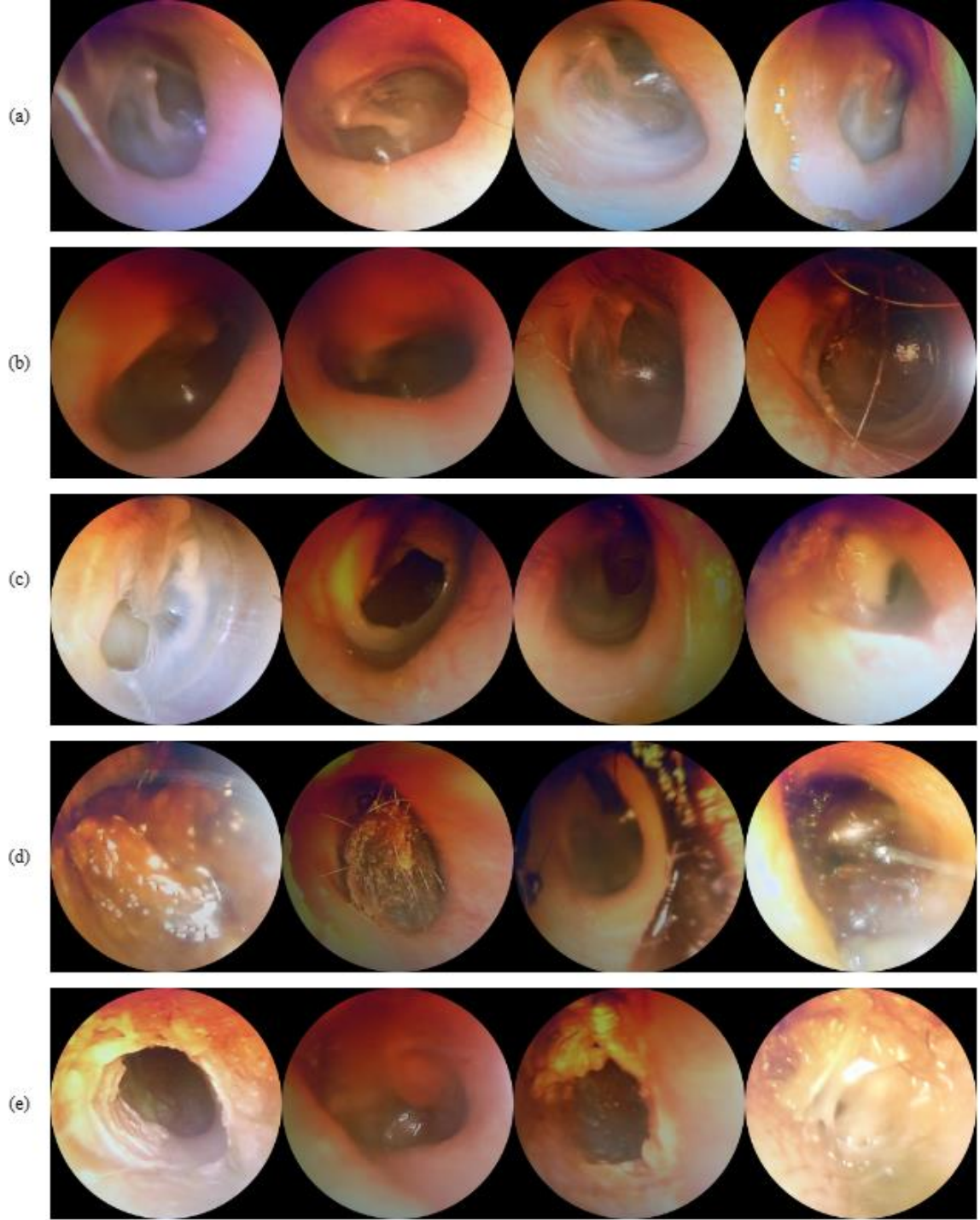
Şekil 2.1. Ear Imagery veri seti Normal (a), CSOM (b), Earwax (c) ve Myringosclerosis (d).

Tympanic membrane veri seti, Özel Akdamar Hastanesinde Ekim 2018 ile Haziran 2019 tarihleri arasında 2-71 yaş aralığındaki hastalardan elde edilmiştir. Görüntüler standart otoskop cihazıyla çekilerek ilgili hastalığın bulunduğu klasöre kaydedilmiştir. Görüntüler alanında uzman üç doktor tarafından incelenerek etiketlenmiştir. Veri setinde normal, AOM, CSOM, kulak mumu tıkaçı, kulak havalandırma tüpü, kulak dış nesne, otitis eksterna, yalancı kulak zarı ve timpanoskleroz olmak üzere 9 sınıf bulunmaktadır. Tablo 2.1’de veri setinde bulunan sınıflar ve ilgili sınıfa ait görüntü sayısı verilmektedir. Tabloda yalancı kulak zarı, timpanoskleroz, kulak dış nesne ve havalandırma tüpü sınıfları diğer olarak verilmektedir [27].

Tablo 2.1. Tympanic Membrane veri setinde bulunan sınıflar ve görüntü sayıları

Açıklama	Sınıf	Görüntü Sayısı
Sağlıklı	Normal	535
Hastalıklı	AOM	119
	CSOM	63
	Earwax	140
	Otitis Externa	41
	Diğer	58

Şekil 2.2’de ilgili veri setindeki sınıflara ait örnek görüntüler verilmektedir. Veri setinde toplamda 958 görüntü bulunmaktadır ve görüntüler 500x500 piksel boyutundadır.



Şekil 2.2. Tympanic membrane veri seti Normal (a), AOM (b), CSOM (c), Earwax (d) ve Otitis Externa (e).

Yapılan çalışmada, iki veri setinden farklı veri setleri türetilmiştir. Tablo 2.2’de türetilen yeni veri setleri, görüntü sayısı, sınıf sayısı ve sınıflar verilmiştir.

Tablo 2.2. Çalışmada kapsamında kullanılan veri setleri

Veri Seti Adı	Sınıf Sayısı	Sınıflar	Toplam Görüntü Sayısı
Merged_ds	9	Normal, Normal 2, AOM, CSOM, CSOM 2, Earwax, Earwax 2, Myringosclerosis, Otitis Externa	1778
Merged_ds (2 sınıf)	2	Normal, Abnormal	1778
Merged_ds (6 sınıf)	6	Normal, AOM, CSOM, Earwax, Myringosclerosis, Otitis Externa	1778
Tympanic Membrane	5	Normal, AOM, CSOM, Earwax, Otitis Externa	898
YOLOv5 uygulanmış Tympanic Membrane	5	Normal, AOM, CSOM, Earwax, Otitis Externa	898
Ear Imagery	4	Normal, CSOM, Earwax, Myringosclerosis	880

Tablo 2.2’de görüldüğü üzere veri setleri sınıf bazında kategorik (categorical) ve ikili (binary) olarak ikiye ayrılmıştır. İki veri seti birleştirilerek kategorik olarak altı ve dokuz sınıf bulunan, ikili olarak iki sınıf bulunan üç farklı veri seti elde edilmiştir. Altı sınıf olan veri setinde aynı etiketler birleştirilmiştir ve dokuz sınıf bulunan veri setinde iki veri seti ayrı olarak etiketlenerek kullanılmıştır. İkili olarak ayrılan veri setinde görüntüler normal ve hastalıklı (abnormal) olarak etiketlenmiştir.

2.2. Derin Öğrenme

Makine öğrenmesi, matematiksel modellerin tahmin yeteneğini oluşturabilmesi için eğitim verileri üzerinden bir örüntü çıkarması ve bu örüntünün kendi kendine geliştirilmesidir[28]. Bu öğrenme işleminde modelin doğru şekilde örüntüyü çıkarması için insan etkileşimine ihtiyaç duymaktadır. Derin öğrenme modellerinde öğrenme işlemi, modelin eski hatalar üzerinden kendini düzeltmesi ile sağlanmaktadır. Bundan dolayı makine öğrenmesine göre derin öğrenme modelleri daha uzun eğitim süresine ve daha fazla veriye ihtiyaç duymaktadır[29]. Derin öğrenmede kullanılan bazı katmanlar şunlardır:

2.2.1. Evrişim Katmanı

Evrişim katmanı, görüntülerden veya özellik haritalarından özellik öğrenme zorluğunu ele almak için tasarlanmıştır. Evrişim katmanı kullanılarak uzamsal bilginin korunurken veriye ait yerel özelliklerin çıkarılması sağlanmaktadır. Özellik çıkarımı için, giriş verisiyle filtreler arasında konvolüsyon işlemi gerçekleştirilmektedir. Konvolüsyon işlemi, filtrenin girdi verisi üzerinde kaydırılması ve noktasal çarpımının hesaplanması ile gerçekleştirilmektedir. Birden çok filtre uygulanarak girdi verisinin farklı türdeki özellikleri elde edilmektedir. Denklem (3.1)'de evrişim katmanının matematiksel gösterimi verilmektedir.

$$X_j^l = f\left(\sum_{i \in M_j} X_i^{l-1} * k_{ij}^l + b_j^l\right) \quad (3.1)$$

Burada, k_{ij}^l öğrenilebilir kernel değerini, b_j^l eğitilebilir bias değerini, $f(\cdot)$ aktivasyon fonksiyonunu, M_j giriş harita seçimini ve $*$ eleman-bazında çarpımı belirtmektedir. Eğitim sırasında kernel değeri ve bias değeri güncellenerek kayıp fonksiyonu minimize edilmektedir. Modeller, birden fazla evrişim katmanı yığınlayarak ve aktivasyon fonksiyonları uygulayarak, girdi hakkında tahminler yapmak için yararlı ve soyut özellikleri öğrenmektedir.

2.2.2. Havuzlama Katmanı

Havuzlama katmanı, çoğunlukla evrişim katmanından sonra özellik haritalarını altörnekleme de (downsampling) ve önemli özellikleri korurken boyutlarını küçültmek için kullanılmaktadır. Giriş özelliğinin haritasını, örtüşmeyen bölgelere veya “havuzlara” böldükten sonra her havuz için istatistik hesaplayarak çalışmaktadır. Denklem (3.2)’de havuzlama katmanının matematiksel gösterimi verilmektedir.

$$X_j^l = \text{down}(X_j^{l-1}) \quad (3.2)$$

Burada, $\text{down}(\cdot)$ fonksiyonu altörnekleme işlemini, X_j^{l-1} girdi verisini ve X_j^l katmanın çıktısını göstermektedir. Havuzlama katmanının bazı avantajları şu şekildedir. İlk olarak, girdi verilerinin uzamsal boyutunu azaltarak modeldeki parametre sayısını düşürür ve aşırı uyuma önlemeye yardımcı olur. İkinci olarak, modelin girdi verisi içerisinde bulunan kaymalarından ve bozulmalarından etkilenmesini önler. Son olarak, her havuzda bulunan önemli değerleri seçerek girdi verisi üzerinden önemli özelliklerin çıkarılmasını sağlar.

2.2.3. Ekleme Katmanı

Eleman-bazında ekleme katmanı olarak da bilinen ekleme katmanı, modellerde birden çok katmanın çıktılarını birleştirmek için kullanılır. Derin öğrenme modellerinde, ResNet mimarisinde kullanılan atlama katmanları gibi girdi ve çıktı arasında bağlantı kurmak için kullanılmaktadır. Eleman-bazında ekleme yapmak için verilen katmanların aynı boyutta olması gerekmektedir. Denklem (3.3)’de ekleme katmanının matematiksel gösterimi verilmektedir.

$$Z = \sum_{i=1}^N (X_i) \quad (3.3)$$

Burada X_i katmanların çıktılarını ve Z ekleme katmanı tarafından üretilen çıktıyı temsil etmektedir. Ekleme katmanı, modellerin performansını ve kararlılığını geliştirmektedir.

2.2.4. Tam-bağlı Katman

Tam-bağlı katman, derin öğrenme modellerinde verilerde bulunan karmaşık kalıpları ve ilişkileri öğrenmede yüksek derecede esneklik sağladıkları için kullanılmaktadır. Tam-bağlı katman ile, girdi ve çıktı arasındaki doğrusal olmayan bağlantı öğrenilmektedir. Doğrusal olmayan bağlantılar, doğrusal bağlantıların yakalayamayacağı özelliklerin çıkarılması için önemlidir. Katmandaki her nöron, önceki katmandaki tüm nöronlardan girdi almaktadır ve nöron bağlantılarında bulunan ağırlık değerlerini girdi ile çarparak sonuç üretmektedir. Ağırlık değerleri, nöron için girdinin önemini belirlemektedir. Denklem (3.4) ve Denklem (3.5)'de tam-bağlı katmanın matematiksel gösterimi verilmektedir.

$$Z = W * X + b \quad (3.4)$$

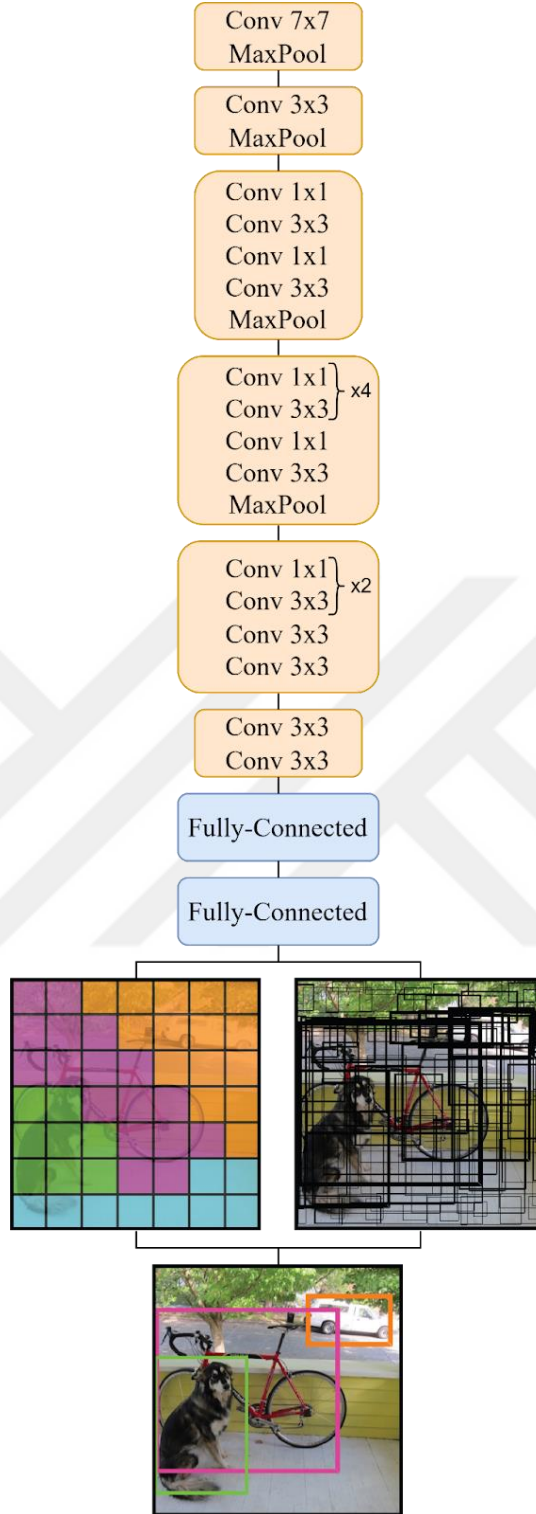
$$Y = f(Z) \quad (3.5)$$

Burada W nöronlar arasındaki bağlantıda bulunan ağırlık değerlerini, X nöronların girdilerini, b bias değerini, $*$ eleman-bazında çarpımı, $f(\cdot)$ aktivasyon fonksiyonunu, Z ve Y elde edilen çıktıları göstermektedir.

2.3. Derin Öğrenme Modelleri

2.3.1. YOLO (You Only Look Once)

You Only Look Once (YOLO), hızlı nesne tespiti için kullanılan bir modeldir. Model nesnenin sınıfını ve nesnenin bulunduğu alandaki sınırlayıcı kutucuğu bulmaya çalışmaktadır. YOLO modeli, nesnelerin tespitini gerçekleştirmek için girdi görüntüsünü CNN mimarisinde tek seferde çalıştırmaktadır. Bu modele hız kazandırırken modelin doğruluğunu düşürmektedir. YOLO modeli, ön işlemde verilen görüntüyü $S \times S$ 'lik ızgaralara bölmektedir ve her ızgara hücresinde bulunan nesnelerin etrafında sınırlayıcı kutu oluşturmaktadır. Her sınırlayıcı kutu için hesaplanan tahmin değerlerinde, nesnenin merkezinin koordinatları, genişliği, yüksekliği ve güven puanı bulunmaktadır. Güven puanı, modelin belirli sınırlayıcı kutuda nesne bulundurduğundan emin olmak için hesaplanmaktadır. Nesne bulunmayan kutularda güven puanı sıfırdır[30].



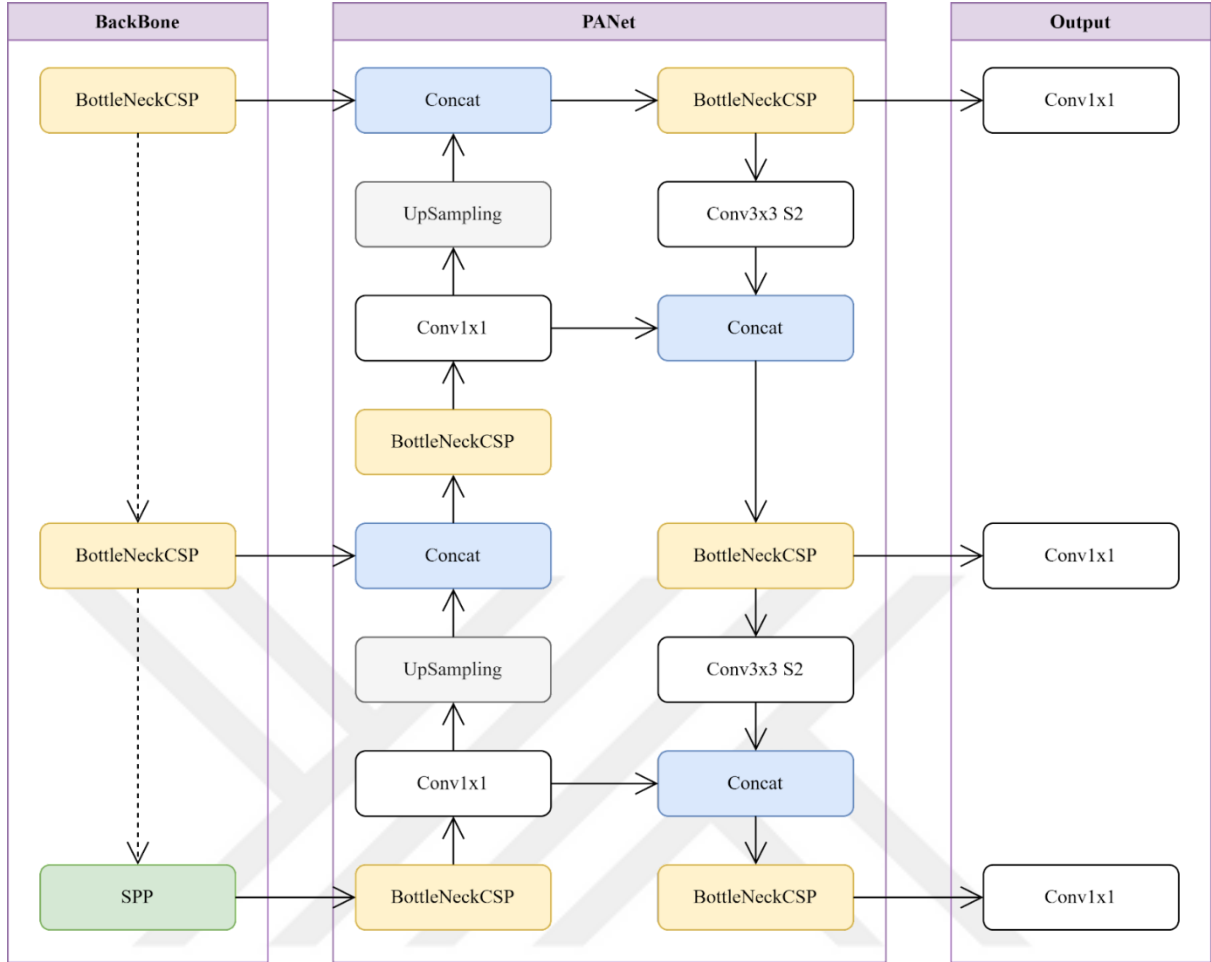
Şekil 2.3. YOLOv1 mimarisinin çalışması.

Ayrıca YOLO algoritması, sınırlayıcı kutuyla birlikte softmax fonksiyonunu kullanarak sınıf olasılıklarını hesaplamaktadır. Sınırlayıcı kutular ve sınıf olasılık değerlerinin birleşimi görüntü için son tahmini vermektedir. Şekil 2.3'te YOLOv1 mimarisinin çalışması gösterilmektedir. Ayrıca YOLOv2'de çapa kutular eklenmiştir. Bu

kutuların boyutu sınırlayıcı kutudan farklı olarak sabittir. Çapa kutuları, görüntü boyunca düzenli aralıklarla yerleştirilir ve görüntüyü oluşturan ızgara hücreleriyle ilişkilendirilir. Bu kutular sayesinde, görüntüde bulunan kısmen kapatılmış veya yalnızca bir kısmı görünen nesnelerin sınıflandırılabilmesi sağlanmaktadır. Genellikle, çapa kutuların boyutu eğitim verisinde bulunan nesnelerin en-boy oranlarına ve boyutlarına göre belirlenmektedir.

YOLOv5 algoritması, YOLOv3 ve YOLOv4 mimarilerin iyi özelliklerini geliştirerek yeni bir yaklaşım kullanmıştır. YOLOv5 algoritması, omurga (backbone), boyun (neck) ve baş (head) olmak üzere üç temel ağ yapısından oluşmaktadır[31]. Omurga ağı, giriş görüntüsünden üst düzey özelliklerin çıkarılmasından sorumludur. Omurga ağı olarak, EfficientNet mimarisinin değiştirilmiş bir versiyonuna dayanan daha güçlü bir mimariyi kullanmaktadır. Bu, modelin daha karmaşık ve soyut özellikleri öğrenmesini sağlayarak nesne algılama başarısını ve doğruluğunu arttırmaktadır. Boyun ağı, birden çok CNN ağını kullanarak farklı ölçek ve çözünürlükteki özellikleri birleştirmektedir. Omurga ağından gelen özelliklerden yerel özellik haritası çıkarır. Bu özellik haritası baş ağının girdisi olarak kullanılmaktadır. Baş ağı, görüntüdeki nesneler için sınırlayıcı kutuları ve sınıf olasılıklarını tahmin etmektedir. Farklı ölçeklerde özellik oluşturmak için SPP (Spatial Pyramid Pooling) yöntemini kullanmaktadır. Bu özellikler boyun ağından elde edilen özellikler ile birleştirilerek CNN algoritmasından geçirilmektedir.

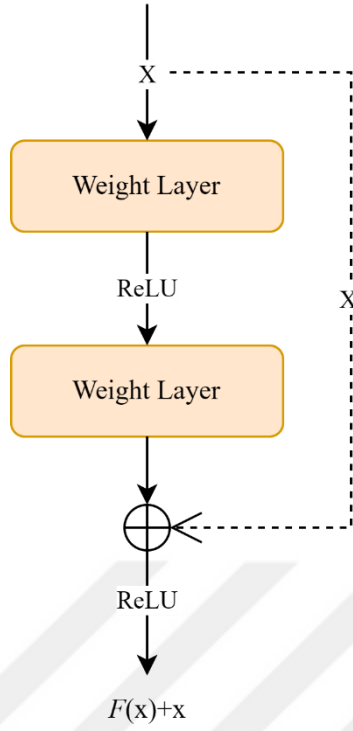
YOLOv5 nesne algılama için çapa tabanlı ve çapa içermeyen yöntemleri birleştirerek kullanmaktadır. Bu çapa kutuların boyutu ayarlamak için dinamik ölçeklendirme kullanmaktadır. Ayrıca model, “ölçeklendirilmiş yolo” yaklaşımını kullanarak donanım ve görev kısıtlamalarına bağlı olarak derinliğinin büyümesine ve küçülmesine izin vermektedir. Şekil 2.4’te YOLOv5 mimarisinin katmanları verilmektedir.



Şekil 2.4. YOLOv5 mimarisi.

2.3.2. ResNet

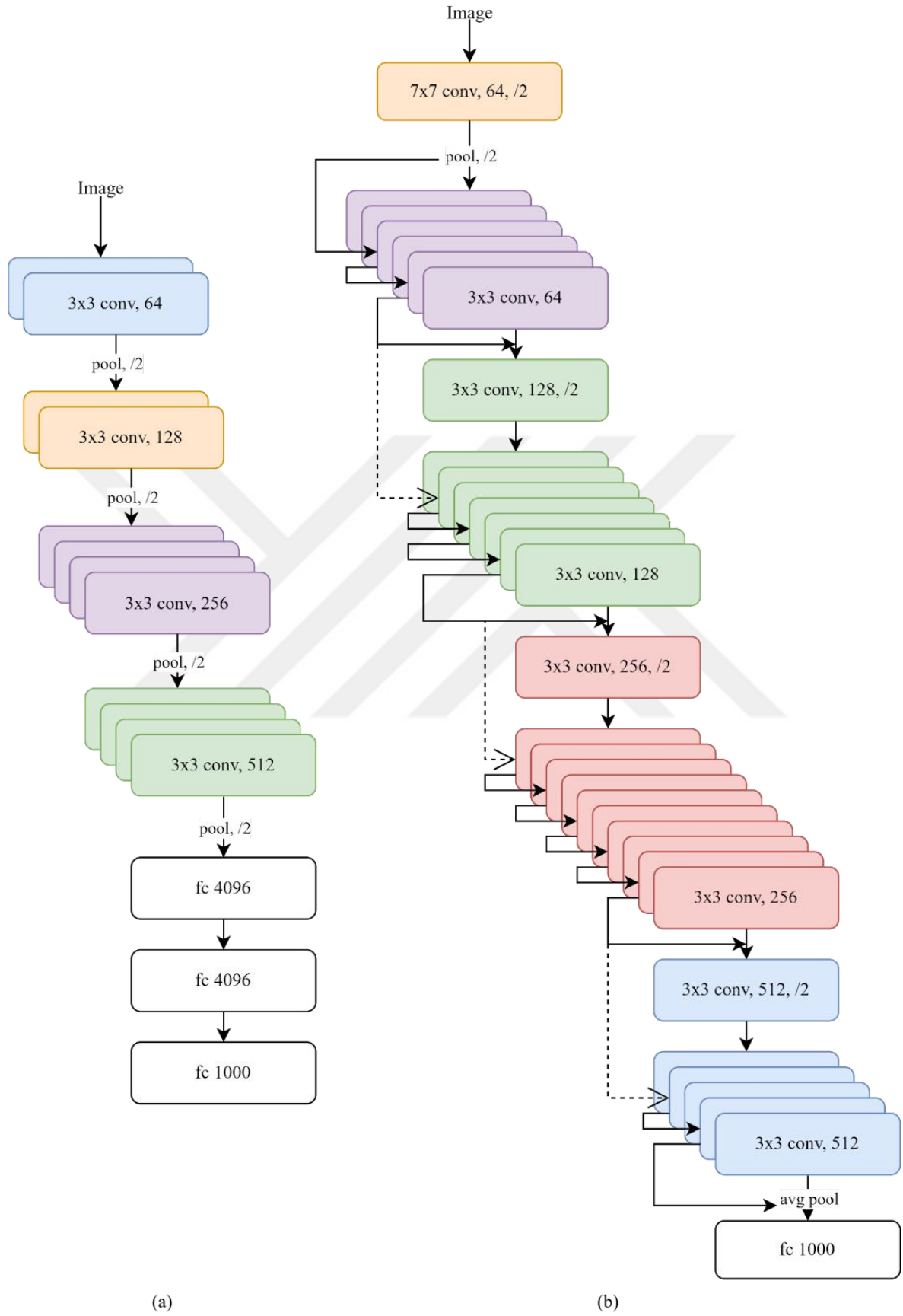
Derin öğrenme modellerinde katman sayısı arttıkça doğruluk değeri doygunluğa ulaşır ve doğruluk hızla düşmeye başlar. Bu probleme kaybolan gradyan denmektedir. ResNet mimarisinde bu problemi çözmek için atlama katmanları (skip layer) kullanmaktadır. ResNet modeli, VGG-19 mimarisine atlama katmanları eklenerek oluşturulmuştur. Şekil 2.5'te gösterildiği üzere atlama katmanları, işlem sonucunda elde edilen çıktılar ile bu katmanlara verilen girdilerin birleştirilmesiyle gerçekleştirilmiştir. Böylece işlem sonucunda kaybolan özellikler tekrar işleme dahil edilmektedir.



Şekil 2.5. ResNet atlama katmanı.

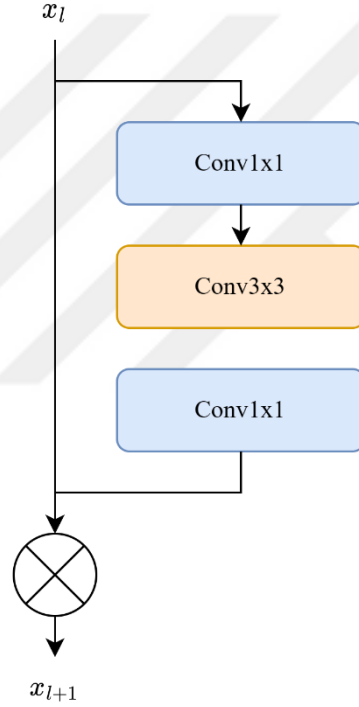
$$y = F(x) + x \quad (3.6)$$

Denklem (3.6)'da $F(x)$ katmanlardan elde edilen işlenmiş çıktıyı, x ise katmana verilen girdiyi göstermektedir. Toplama işleminin gerçekleştirilebilmesi için katmanın çıktısı ile girdinin boyutunun aynı olması gerekmektedir.



Şekil 2.6. VGG-19 (a) ve ResNet34 (b) mimarileri.

Bazı önemli ResNet mimarileri şu şekilde sıralanabilir. ResNet34 mimarisi, VGGNet'ten esinlenerek oluşturulan ilk mimaridir. 34 katmandan oluşmaktadır ve VGGNet'e göre daha az parametre bulundurmaktadır. Şekil 2.6'da ResNet34 ve VGG-19 mimarilerin karşılaştırılması verilmektedir. ResNet50 mimarisi, yapıcı ResNet34'e benzemekle birlikte her iki katmanlı bloklar üç katmanlı darboğaz bloklarıyla değiştirilmiştir. Şekil 2.7'de darboğaz katmanı gösterilmektedir. ResNet101 ve ResNet152 mimarileri, ResNet50 mimarisinin daha fazla darboğaz bloğu kullanılmış şeklidir. En derin ResNet modeli olan ResNet152, VGG-16 ve VGG-19'a göre daha az parametre içermektedir[32].



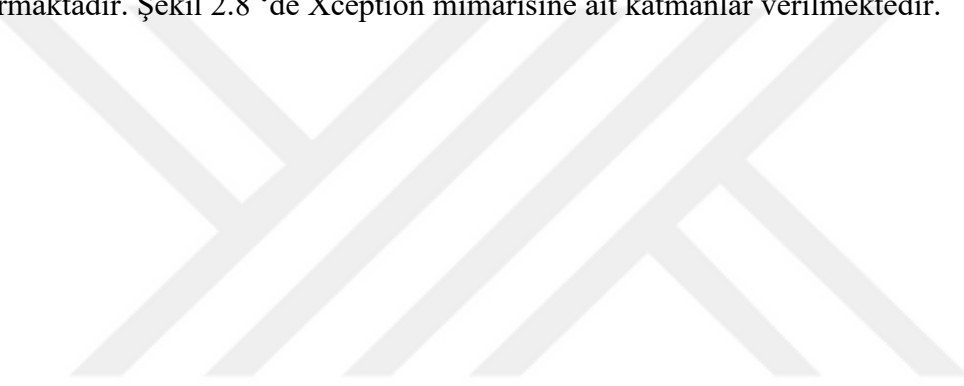
Şekil 2.7. Darboğaz bloğu.

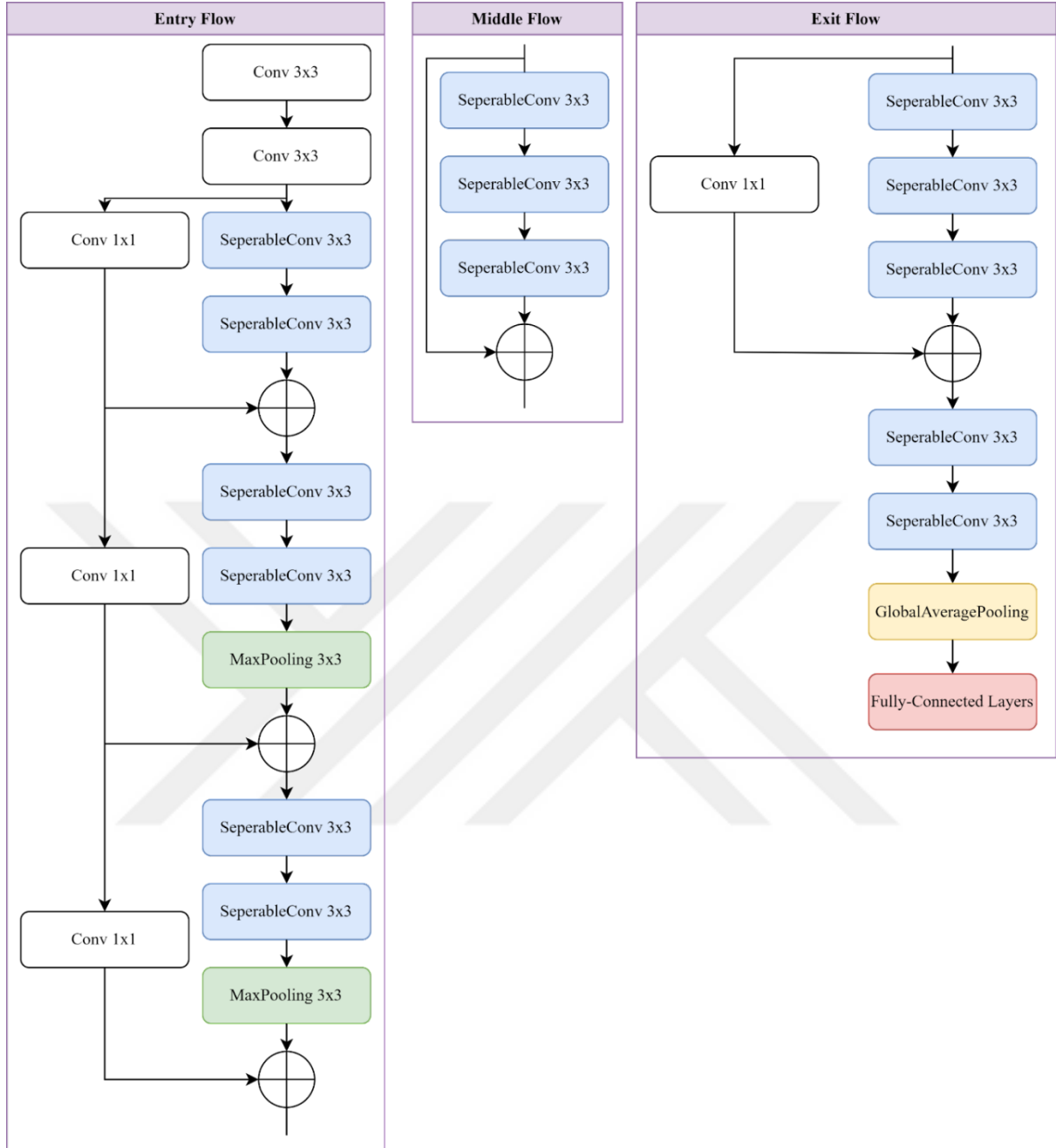
2.3.3. Xception

Xception mimarisi, Inception mimarisinin bir çeşididir. Model, Inception modüllerini derinlemesine ayrılabilir evrişimler ile değiştirmektedir. Derinlemesine ayrılabilir evrişim katmanları, derinlemesine evrişim ve noktasal evrişim olmak üzere iki işlemten oluşmaktadır. Derinlemesine evrişim her giriş kanalına ayrı ayrı tek bir filtre uygulamaktadır. Noktasal evrişim, derinlemesine evrişimin ürettiği çıktıları birleştirmek için

1x1 evriřim uygulamaktadır. Bu yaklařım, geleneksel evriřim katmanlarına gre daha az parametre kullanırken doęruluęu da artırmaktadır[33].

Kaybolan gradyan problemi, derin ęrenme aęlarının eęitiminde ortaya ıkan genel bir sorundur. Aęın aęırlıklarına ait kayıp fonksiyonunun gradyanları, geri besleme sırasında ok kk deęerlere ulařtıęında ortaya ıkmaktadır. Gradyanın ok kk deęerlerde olması, aęın aęırlıklarını gncellemesini yavařlatır veya aę aęırlaklarını hi gncelleyemez. Xception mimarisi, kaybolan gradyan probleminden kurtulmak iin ilk olarak ResNet mimarisinde tanıtılan atlama katmanlarını kullanmaktadır. Aędaki farklı derinliklerde atlama katmanlarını kullanarak aęın hem yzeysel hem de derin zellikleri kullanmasına olanak saęlar. Bu, kaybolan gradyan problemini zerken aęın doęruluęunu da arttırmaktadır. řekil 2.8 ‘de Xception mimarisine ait katmanlar verilmektedir.





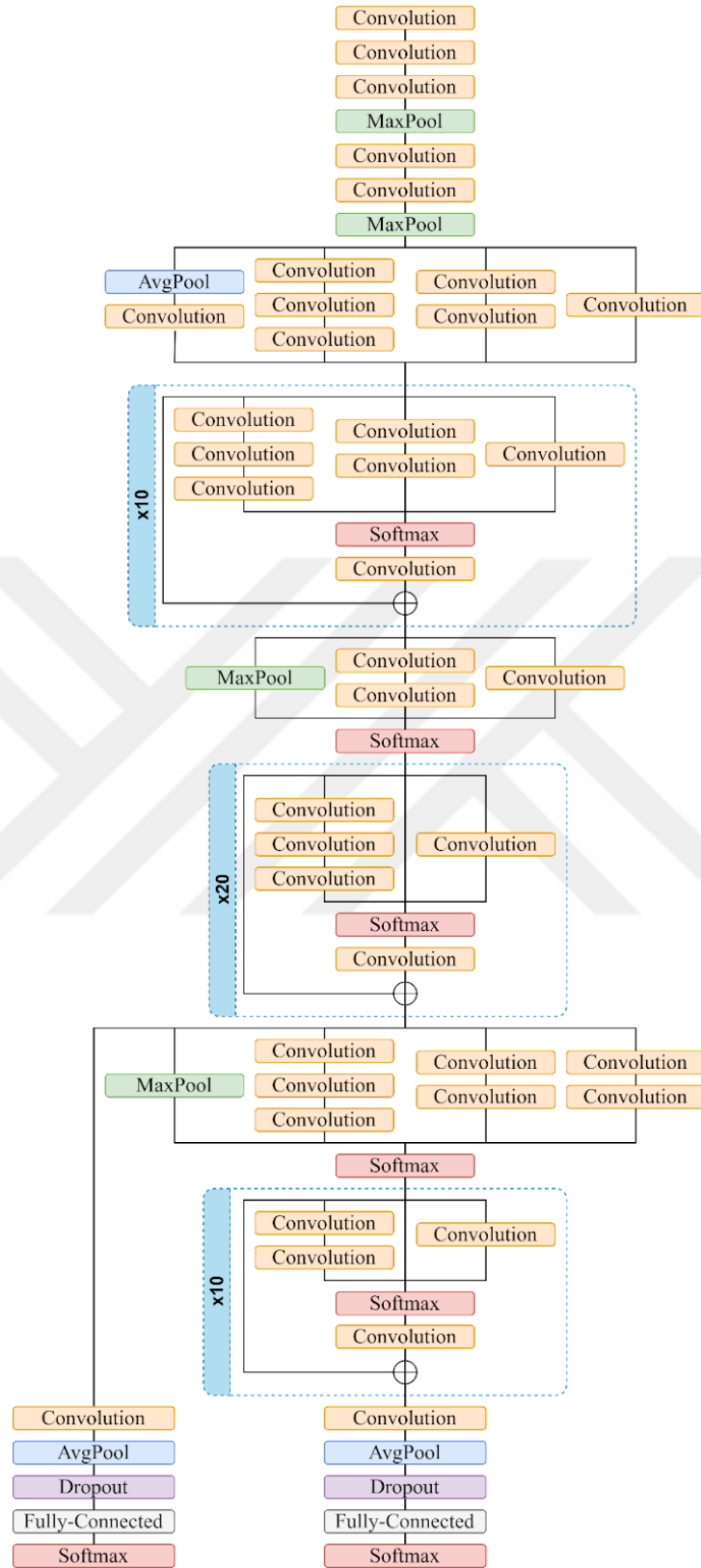
Şekil 2.8. Xception mimarisi.

2.3.4. InceptionResNet

InceptionResNet mimarisi, Inception ve ResNet mimarilerinin birleşimidir. Inception mimarisi, farklı evrişim katmanları kullanılarak çok ölçekli özelliklerin çıkarılması sağlanmıştır. Bu özellikler daha sonrasında birleştirilir ve sonraki katmana girdi olarak verilir. Bu ağın, tek bir filtrenin çıkaramadığı özellikleri çıkarmasına ve veride bulunan özelliklerin ağ tarafından daha iyi öğrenilmesini sağlamaktadır. Fakat mimarinin derinliği

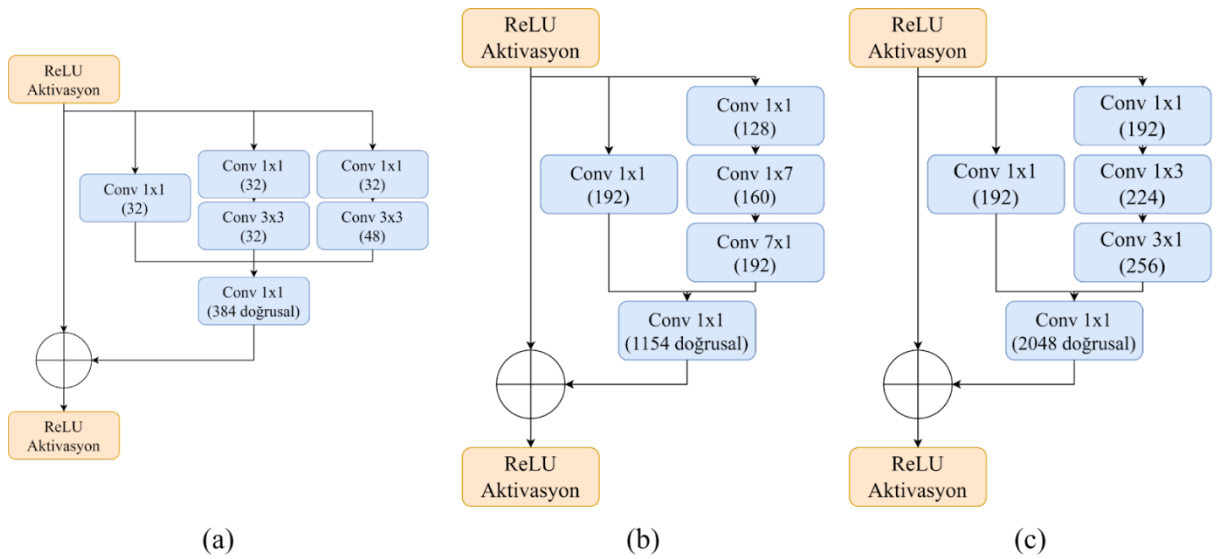
kaybolan gradyan sorununa yol açmaktadır. ResNet mimarisinde bulunan artık bağlantılar, katmanın girdisi ile çıktısını birleştirerek gradyanın ağ üzerindeki akışını hafifletir. Bu, kaybolan gradyan problemini çözmeye yardımcı olur ve ağın derin özellikleri öğrenmesine sağlar. Şekil 2.9’da InceptionResNetV1 ve Inception ResNetV2 mimarileri verilmektedir. V1 ve V2 mimarileri yapıca aynı olmakla birlikte kullandıkları hiperparametreler farklıdır.





Şekil 2.9. InceptionResNetV1 ve V2 mimarisi.

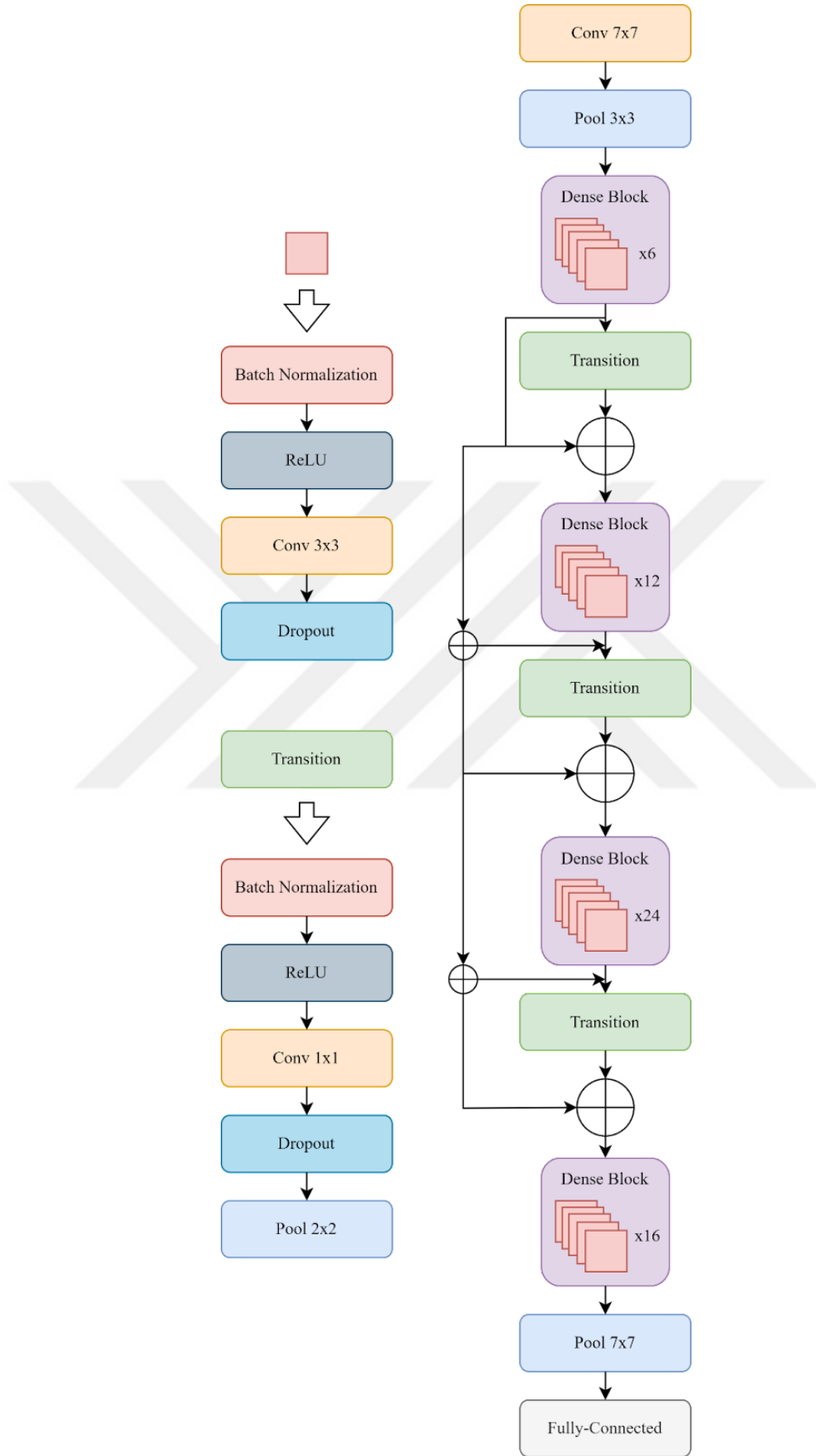
InceptionResNet mimarisi, Inception modüllerin ve artık bağlantıların avantajlarını birleştirmektedir. Mimari, Inception-ResNet bloklarını kullanmaktadır. Inception-ResNet blokları, 1x1, 3x3 ve 5x5 evrişim uygulayarak girdi görüntüsünden derinlemesine özellik çıkarmaktadır. Şekil 2.10'da Inception-ResNet-A, Inception-ResNet-B ve Inception-ResNet-C blokları verilmektedir. Inception-ResNet-A bloğunda, giriş özellik haritasını bağımsız olarak işleyen, her biri farklı filtre boyutlarına sahip dört paralel koldan oluşmaktadır. Kolların çıktıları, bir sonraki katmana girdi olarak verilmek için birleştirilir. Artık bağlantı, giriş özellik haritalarını çıkış özellik haritasına ekleyerek kullanmaktadır. Girdi özellik haritaları, hesaplama verimliliğini artırmak için giriş kanalını azaltan 1x1 evrişim katmanından geçirmektedir. İlk kol, girdinin alt düzey özelliklerini yakalayan tek 1x1 evrişim katmanından oluşmaktadır. İkinci kol, hem yerel hemde genel özelliklerin yakalanmasına yardımcı olan 3x3 ve 1x1 evrişim katmanlarını kullanmaktadır. Üçüncü kol, 1x1, 3x3 ve 3x3 evrişim katmanından oluşmaktadır. Bu kol, girdiden karmaşık özelliklerin ve kalıpların çıkarılmasından sorumludur. Dördüncü kol, üçüncü kola benzemektedir ve daha kompleks özelliklerin çıkarılmasını sağlar. Inception-ResNet-B bloğu, Inception-ResNet-A bloğunun geliştirilmiş versiyonudur. 1x1 evrişim ekleyerek bloğun hesaplama maliyeti azaltılmıştır. Inception-ResNet-C bloğu, Inception-ResNet-B bloğuna benzemektedir fakat Inception-ResNet-B bloğunda bulunan 1x7 ve 7x1 evrişim katmanları yerine 1x3 ve 3x1 evrişim katmanı kullanılmaktadır. Bu çıkarılan özellik sayısını arttırmaktadır.



Şekil 2.10. Inception-ResNet-A (a), Inception-ResNet-B (b) ve Inception-ResNet-C (c) blokları.

2.3.5. DenseNet

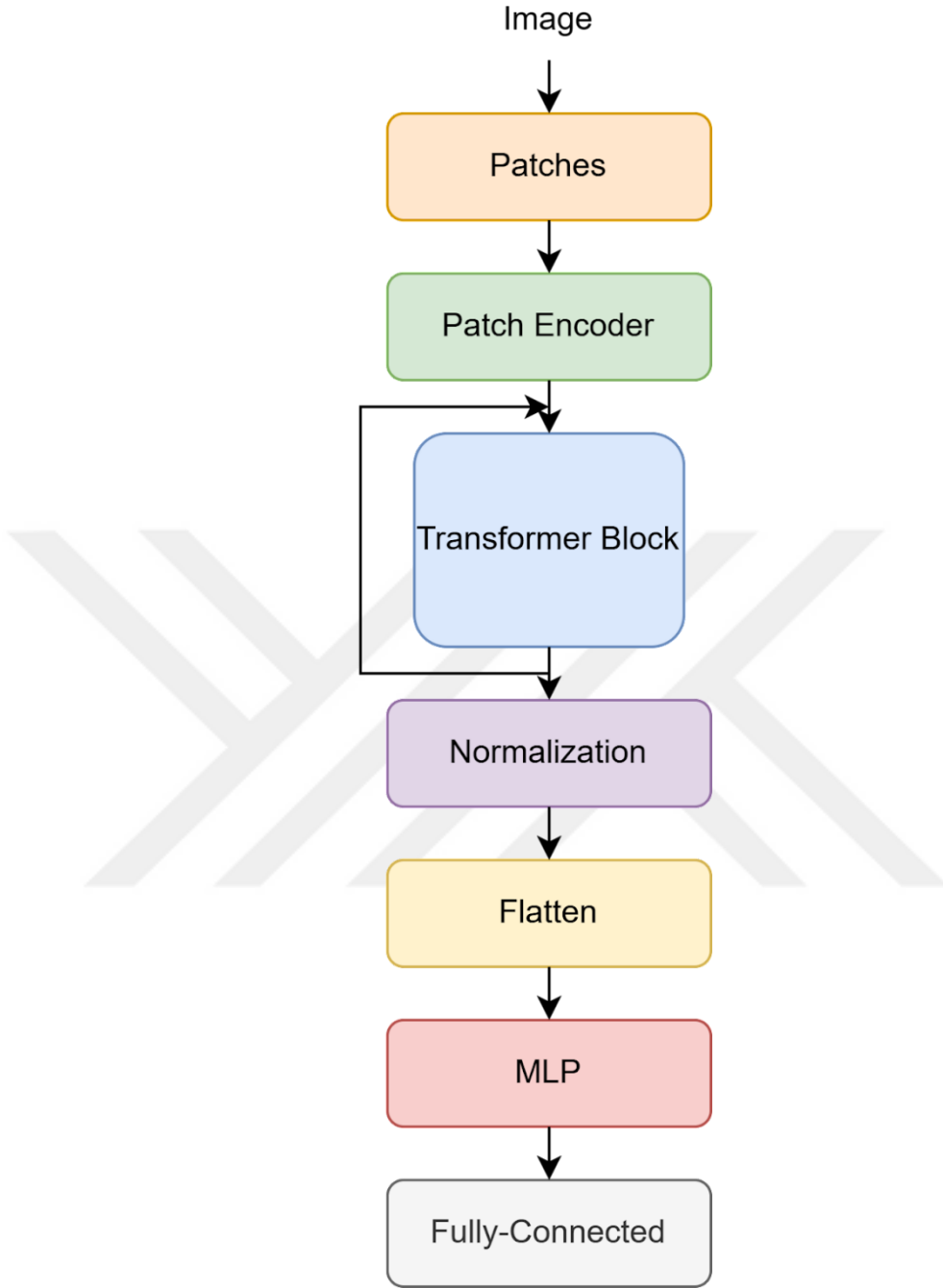
DenseNet, ağıdaki parametre sayısını azaltırken kaybolan gradyan sorununu hafifletmeyi amaçlayan bir evrişimli sinir ağı mimarisidir. ResNet ve Inception mimarilerinin aksine, DenseNet tüm katmanlardan özellik haritalarını birleştiren atlama bağlantılarından yararlanmaktadır. Bu mimari, tam-bağlı ve geçiş katmanlarından oluşturulmuştur. Önceki tüm katmanların çıktı özellik haritaları birleştirilerek bir sonraki katmana girdi olarak verilmektedir. Dense bloğunun çıktısı geçiş katmanına (transition layer) aktarılarak altörnekleme yapmak için 1×1 evrişim katmanı ve ardından 2×2 ortalama havuzlama katmanından geçirilmektedir. Bu işlem, özellik haritalarının uzamsal boyutlarını azaltır ve kanala sayısını artırarak özelliklerin yeniden kullanılmasına olanak sağlar. Dense bloklarında tam-bağlı katmanların kullanılması sonucunda, ağ daha geniş özellik seti öğrenmektedir ve girdi verileri hakkında daha ayrıntılı bilgi elde etmektedir. Şekil 2.11’de DenseNet mimarisinin katmanları verilmektedir.



Şekil 2.11. DenseNet mimarisi.

2.3.6. Vision Transformer

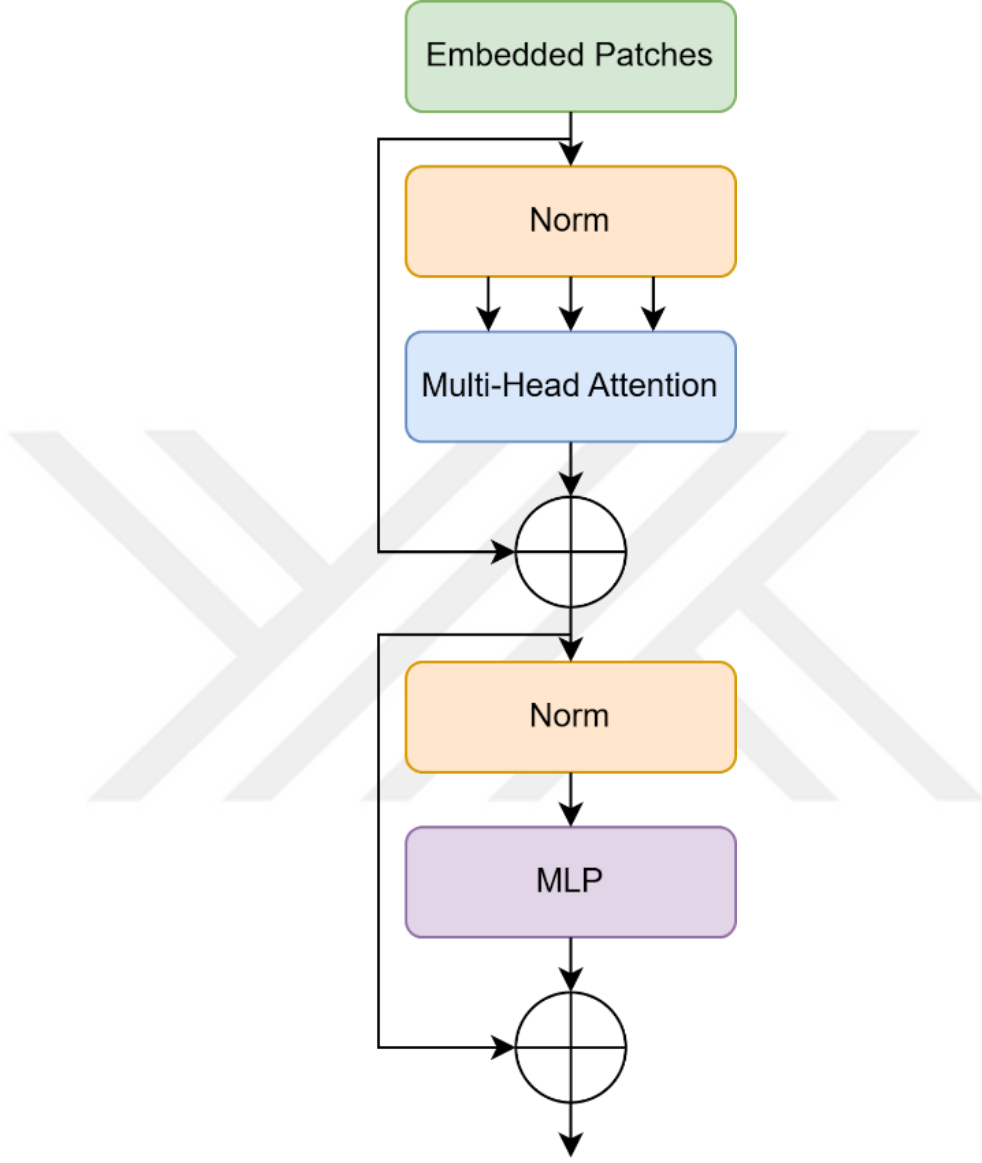
ViT olarak da bilinen Vision Transformer mimarisi, görüntülerden uzamsal özellikler çıkarmak için evrişim katmanlarını kullanan geleneksel evrişimli sinir ağlarının (CNN) aksine Transformer mimarisini kullanmaktadır. Transformer mimarisi ilk olarak, doğal dil işleme görevinde kullanılmak üzere geliştirilmiştir. ViT mimarisi, görüntünün farklı bölümlerine odaklanmak için Transformer mimarisinde bulunan öz-dikkat mekanizmasını kullanmaktadır. Öz-dikkat mekanizması, görüntüdeki farklı konumlardaki özelliklerin ağırlıklı toplamını hesaplayarak, modelin en ilgili özelliklere dikkat etmesini sağlamaktadır. ViT mimarisi, doğal dil işlemede bulunan token dizileri yerine görüntü yamalarını (patches) işlemektedir. Bu öz-dikkat mekanizmasının görüntünün belli bir noktasına odaklanmasını sağlamaktadır.



Şekil 2.12. ViT mimarisi çalışma prensibi.

Şekil 2.12’de görüldüğü üzere ViT mimarisi yama gömme katmanı (patch embedding layer) ve transformer şifreleme olmak üzere iki ana bileşenden oluşmaktadır. Yama gömme katmanı, girdi görüntüsünü örtüşmeyen yamalardan oluşan ızgaralara böler ve her yamayı vektörel temsilciye eşler. Daha sonrasında gömme yamalar, öz-dikkat ve ileri besleme katmanlarından oluşan transformer şifreleyiciye girdi olarak verilir. Öz-dikkat katmanı,

modelin girdi görüntüsünün farklı bölümlerine odaklanmasına izin verirken, ileri besleme katmanını girdi görüntüsünden özellik çıkarma işlemini gerçekleştirir[34].



Şekil 2.13. Transformer bloğu.

Şekil 2.13'te transformer bloğunun katmanları verilmektedir. Blok, çok-başlı öz-dikkat (MHA), ileri besleme ve normalleştirme olmak üzere üç katmandan oluşmaktadır. Çok-başlı öz-dikkat katmanı, gömül girdiyi çok başa ayırır ve bu çoklu dikkat mekanizmasının öğrenilmesini sağlamaktadır. MHA katmanı, gömül girdilerden sorgu, anahtar ve değer vektörleri oluşturur. Bu vektörler, girdi dizisindeki her token arasındaki dikkat ağırlıklarını hesaplamak için kullanılır.

$$Attention(Q, K, V) = softmax(\frac{QK^T}{\sqrt{d_k}})V \quad (3.7)$$

Denklem (3.7)'de tek-baş için dikkat hesabı verilmektedir. Burada Q sorgu matrisini, K anahtar matrisini, V değer matrisini ve d_k anahtar matrisi için saklı boyutu temsil etmektedir. Matris çarpımı QK^T , olası her sorgu ve anahtar çifti için iç çarpım hesaplamaktadır.

$$MultiHead(Q, K, V) = Concat(head_1, \dots, head_n)W^O \quad (3.8)$$

$$head_i = Attention(QW_i^Q, KW_i^K, VW_i^V) \quad (3.9)$$

Çok-başlı öz-dikkat hesabı için Denklem (3.8) ve Denklem (3.9) kullanılmaktadır. Burada $head_i$ gömülü yama için hesaplanan baş değerini, W^O izdüşüm matrisini göstermektedir. İzdüşüm matrisler kullanılarak dikkat ağırlıkları hesaplanmaktadır. İleri besleme katmanı, her simgeye (token) ileri yönlü doğrusal dönüşüm ve simge çıktılarına doğrusal olmayan aktivasyon fonksiyonu uygular. Bu katman, giriş dizisindeki daha karmaşık özelliklerin yakalanmasına yardımcı olur. Normalizasyon katmanı, modelin farklı girdi örneklerini genelleştirme yeteneğini güçlendirmek için kullanılmaktadır.

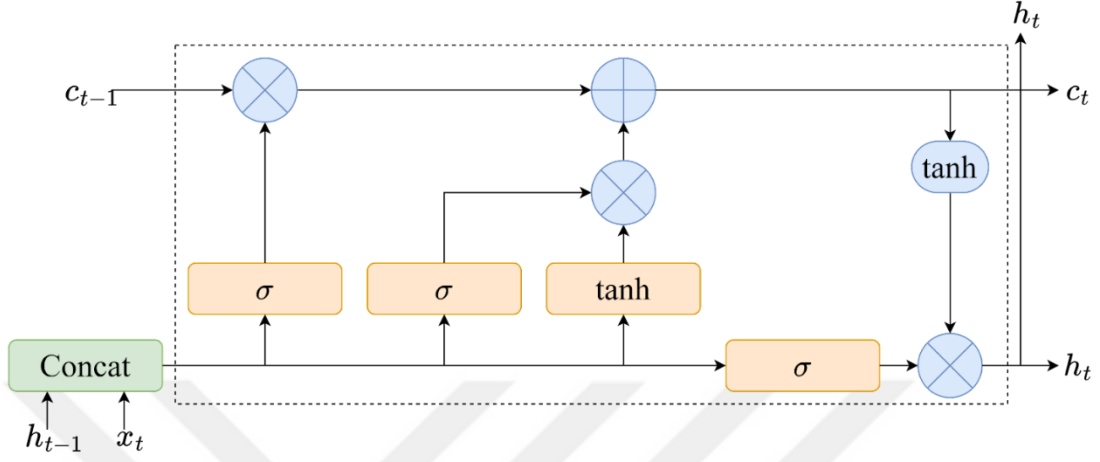
2.3.7. LSTM

LSTM (Long Short-Term Memory), ardışık verilerin modellenmesi gereken problemlerde kullanılan bir yapay sinir ağıdır. Kaybolan gradyan sorununa eğimli sinir ağlarının aksine, verilerdeki uzun vadeli bağımlılıkları yakalayarak bu problemi hafifletir. LSTM ağı, birbirine kapılarla bağlanan hafıza hücrelerini kullanmaktadır. Hücreler, uzun vadeli hafızayı yönetirken kapılar hücrelere giren ve çıkan bilgi akışını kontrol eder.

LSTM ağı, hücre durumu ve üç kapı bulunduran hafıza hücrelerinden oluşmaktadır. Üç kapı sırasıyla, giriş kapısı, unutma kapısı ve çıktı kapısıdır. Giriş kapısı yeni bilgilerin hücre durumuna eklenme miktarından, unutma kapısı eski bilgilerin silinme işleminden ve çıkış kapısı hücre durumdan çıkan bilgi miktarından sorumludur.

LSTM ağı her zaman adımını bir sonrakine besleyerek girdi dizilerini yinelemeli bir şekilde işler. Bu işlem, her zaman adımında hafıza hücresinin çıktısı bir sonraki hafıza hücresi için girdi verilerek gerçekleştirilir. Bu şekilde, LSTM ağı bellek hücrelerindeki

bilgileri seçerek depolamayı, girdi olarak seçmeyi ve değiştirmeyi öğrenmektedir. Şekil 2.14'te LSTM ağına hafıza hücreleri verilmektedir.

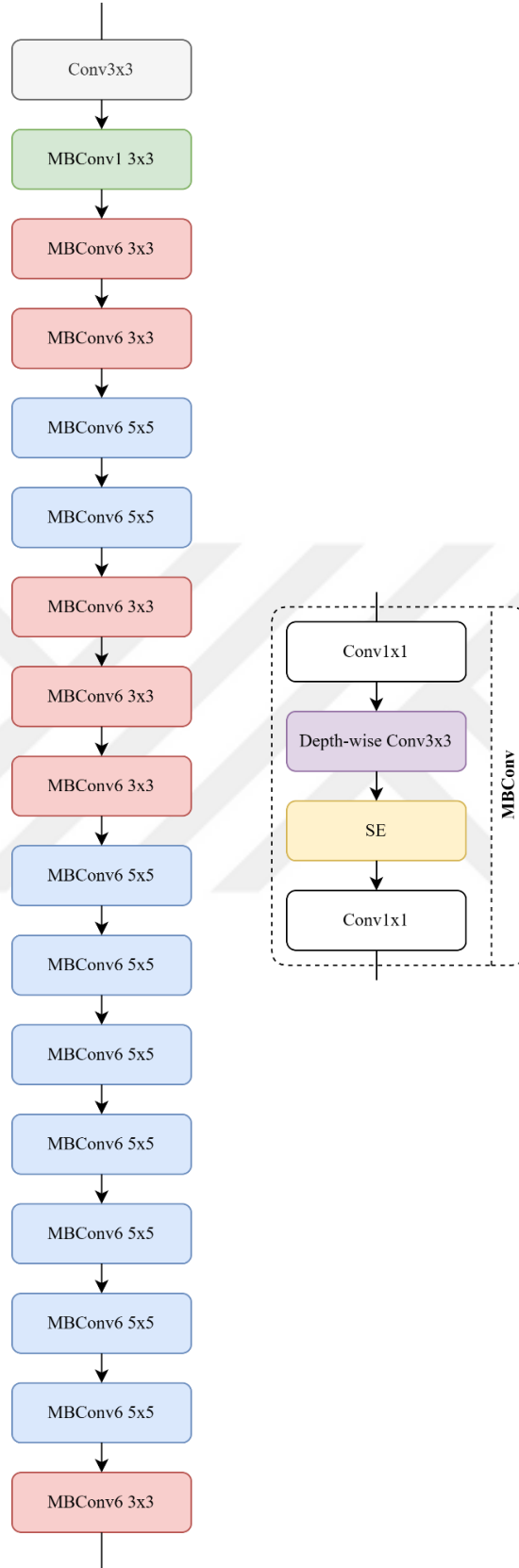


Şekil 2.14. LSTM ağı hafıza hücreleri.

2.3.8. EfficientNet

EfficientNet mimarisi, hesaplama verimliliğini korurken yüksek performans elde etmek için tasarlanmış bir CNN modelidir. Model boyutu, doğruluk ve hesaplama kaynakları arasında bir denge kurmak için tasarlanmıştır. EfficientNet ölçeklendirmeyi, parametre ve hesaplama sayısını azaltan derinlemesine ayrılabilir konvolüsyonlar (deep-wise separable convolution) ile ağı boyutunu ve karmaşıklığını kontrol eden bileşik ölçeklendirme katsayılarının kombinasyonunu kullanarak elde etmektedir. Mimari derinlik ölçeklendirme, genişlik ölçeklendirme ve çözünürlük ölçeklendirme olmak üzere üç ana fikir içermektedir. Derinlik mimarinin katman sayısını belirtmektedir. Daha derin ağlar girdiden daha fazla özellik yakalamaktadır. Genişlik her katmandaki kanal sayısını vermektedir. Ağı genişliğinin artırılması modelin çeşitli özellikleri yakalamasını sağlamaktadır. Çözünürlük, girdi görüntüsünün boyutunu ifade etmektedir. Yüksek çözünürlüklü görüntü daha fazla özellik barındırmaktadır fakat işlem maliyetini artırmaktadır.

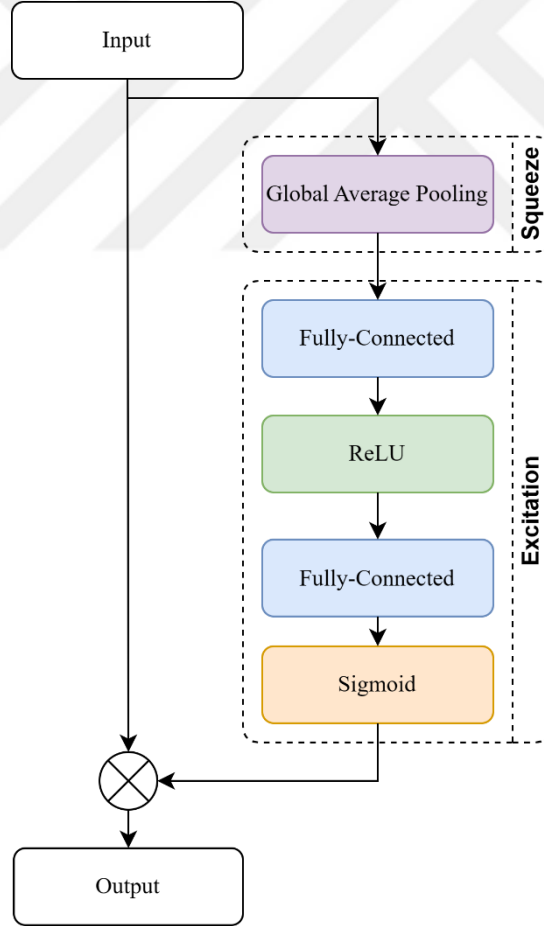
Mimari içerisinde, evrişim katmanları, derinlemesine ayrılabilir evrişimler, sıkıştırma ve uyarma blokları (Squeeze-and-Excitation block) ve küresel ortalama havuzlama gibi katmanların birleşiminden oluşturulmuştur. Şekil 2.15'te EfficientNet ağına mimarisi verilmektedir.



Şekil 2.15. EfficientNet mimarisi.

Model ölçeğine göre B0 ile B7 arasında adlandırılmıştır. EfficientNetB0 modeli ana mimariyi içerirken, EfficientNetB7 modeli en güçlü ve en derin mimariyi barındırmaktadır. EfficientNetB0 modeli, orta derinlik, genişlik ve çözünürlüğe sahip temel modeldir. EfficientNetB1-B6 modelleri, B0 modeline göre daha derin ölçeklidir. Ölçek sayısının artması modelin performansını, eğitim süresini ve donanım maliyetini artırmaktadır. EfficientNetB7 modeli, en derin ve en performansı yüksek modeldir. Bu modelin donanım maliyeti EfficientNet mimarileri arasında en yüksek seviyedir. Mimari seçiminde donanım kaynakları ve veri seti boyutu önemlidir.

Sıkıştırma ve uyarma blokları, EfficientNet mimarisinde ilk olarak kullanılmıştır. Bloklar, özelliklerden bilgilendirici olanları seçerek vurgular ve daha az alakalı olanları bastırır. Bu işlem, ağıın ayırt edici gücünü geliştirir. Şekil 2.16'da EfficientNet mimarisinde kullanılan sıkıştırma ve uyarma blokları verilmektedir.



Şekil 2.16. Sıkıştırma ve uyarma bloğu.

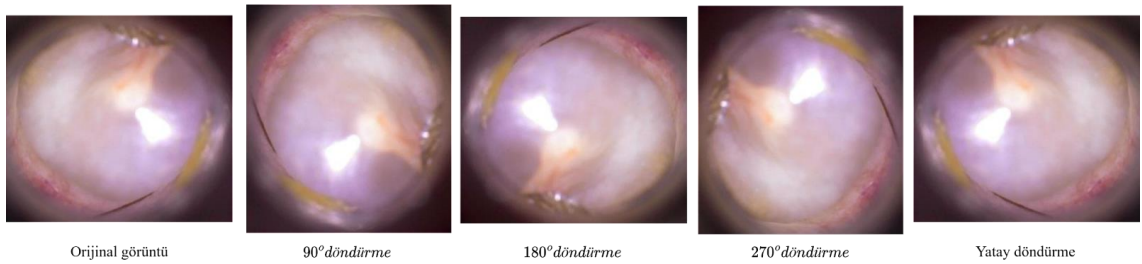
Sıkıştırma işleminde, özellik haritasının uzamsal boyutları üzerinde genel ortalama havuzlama işlemi gerçekleştirilerek genel bilgiler yakalanır. Bu süreç sonucunda, uzamsal

boyutlara ait kanal bazında istatistikler veya özellikler ağ tarafından yakalanır ve bunlara kanal bazında tanımlayıcılar adı verilmektedir. Sıkıştırma işleminden elde edilen kanal bazında tanımlayıcılar, uyarma işleminde küçük tam-bağlı sinir ağlarına girdi olarak verilir. Bu işlemin amacı, kanal bazında tanımlayıcılara ait olan bağımlılıkları yakalamaktır. Tam bağlı ağın çıktısı, her kanal için modülasyon veya dikkat değerlerini temsil etmektedir.

2.4.Ön İşleme (Preprocessing)

Derin öğrenme çalışmalarında veri setinin yetersiz geldiği durumlarda veri sayısının artırılması için arttırma (augmentation) işlemi uygulanmaktadır. Geometrik dönüşümler, renk uzayı dönüşümleri, kernel filtreleri, rastgele silmeler ve görüntülerin birleştirilmesi en genel görüntü arttırma işlemlerindendir. Veri arttırma işleminde uygulanan alana göre arttırma işleminin seçilmesi önemlidir. Örneğin X-ray görüntülerinde yatay aynalama işlemi sonucunda vücut anatomisi değişmektedir ve bu modelin başarısını olumsuz etkilenmektedir[35].

Çalışma kapsamında ön işleme adımı, kullanılan modellere uygun girdi boyutunda veri setlerinde bulunan görüntüler yeniden boyutlandırılmıştır. Ayrıca veri setindeki görsellere arttırma işlemi uygulanmıştır. Görüntüler üzerinde 90° , 180° ve 270° derece döndürme ve yatay aynalama uygulanmıştır. Ayrıca renk uzayının etkisinin incelenmesi için renkli RGB görüntüler OpenCV paketi kullanılarak HSV, LUV ve gri renk uzaylarına çevrilmiştir. Şekil 2.17’de arttırma işlemi uygulandıktan sonra elde edilen görüntüler verilmektedir.



Şekil 2.17. Arttırma işlemi sonucunda elde edilen görüntüler.

Tablo 2.3’te yapılan işlemler açıklanmaktadır ve işlemler sonucunda elde edilen görüntü sayıları verilmektedir. Arttırma işlemi sonucunda görüntü sayısı dört katına

çıkıştır. Bu işlem, modellerin daha fazla veri üzerinde eğitilmesini ve verilerde bulunan özellik haritalarını doğru şekilde çıkarmasını sağlamaktadır.

Tablo 2.3. Artırma işlemi sonucunda elde edilen görüntü sayıları

Veri Seti	Sınıf	Eğitim Görüntü Sayısı	Artırma Sonucunda Görüntü Sayısı
Tympanic Membrane	Normal	375	1500
	AOM	84	336
	CSOM	45	180
	Earwax	98	392
	Otitis Externa	29	116
Ear Imagery	Normal	154	616
	CSOM	154	616
	Earwax	154	616
	Myringosclerosis	154	616

HSV renk uzayı, ton (Hue), doygunluk (Saturation) ve değer (Value) alanlarından oluşmaktadır. Ton bileşeni, renk bilgisini renk çemberinde derece cinsinden temsil eder. Ton bileşeni için, 0o değeri kırmızı rengi, 120o değeri yeşil rengi ve 240o değeri mavi rengi vermektedir. Doymuluk bileşeni, rengin yoğunluğunu veya saflığını temsil etmektedir ve %0 gri görüntü, %100 tam renkli görüntü için kullanılmaktadır. Düşük doymuluğa sahip görüntünün rengi soluk, yüksek doymuluğa sahip görüntü parlak ve canlı görünmektedir. Değer bileşeni, rengin parlaklığını veya açıklığını temsil etmektedir ve %0 siyah ile %100 beyaz arasında yüzde olarak ölçülmektedir.

RGB renk uzayından HSV renk uzayına geçiş Denklem (3.10) ile Denklem (3.15) arasında verilen formüller kullanılarak gerçekleştirilmektedir.

$$R' = \frac{R}{255}, G' = \frac{G}{255}, B' = \frac{B}{255} \quad (3.10)$$

$$C_{max} = \max(R', G', B'), C_{min} = \min(R', G', B') \quad (3.11)$$

$$\Delta = C_{max} - C_{min} \quad (3.12)$$

$$H = \begin{cases} 60^\circ x \left(\frac{G' - B'}{\Delta} \bmod 6 \right), C_{max} = R' \\ 60^\circ x \left(\frac{B' - R'}{\Delta} + 2 \right), C_{max} = G' \\ 60^\circ x \left(\frac{R' - G'}{\Delta} + 4 \right), C_{max} = B' \end{cases} \quad (3.13)$$

$$S = \begin{cases} 0, C_{max} = 0 \\ \frac{\Delta}{C_{max}}, C_{max} \neq 0 \end{cases} \quad (3.14)$$

$$V = C_{max} \quad (3.15)$$

Burada R' , G' ve B' değerleri normalize edilmiş renk değerlerini temsil etmektedir. C_{max} ve C_{min} normalizasyon değerlerinden maksimum ve minimum değerleri vermektedir. H , S ve V değerleri renk uzayını tanımlamaktadır. Bu işlemler sonucunda RGB görüntü HSV görüntüye dönüştürülmektedir.

LUV renk uzayı, renkleri spektral özelliklerinden ziyade algısal özelliklerine göre temsil etmek üzere tasarlanmıştır. LUV renk uzayında diğer renklerde olduğu gibi üç kanal bulunmaktadır. İlk kanal olan L rengin parlaklığını, ikinci kanal U kırmızı-yeşil eksenini ve son kanal V mavi-sarı eksenini temsil eder. L için 0 siyah ve 100 beyaz rengi vermektedir. U ve V değerleri kromatik koordinatta -100 ile 100 arasında değer almaktadır.

RGB renk uzayından LUV renk uzayına dönüşüm Denklem (3.16) ile Denklem (3.21)'de verilmektedir. İlk olarak RGB renk uzayının XYZ renk uzayına çevrilmesi gerekmektedir ve dönüşüm formülü Denklem (3.16)'da verilmektedir.

$$\begin{aligned} X &= 0.412453 * R + 0.357580 * G + 0.180423 * B \\ Y &= 0.212671 * R + 0.715160 * G + 0.072169 * B \\ Z &= 0.019334 * R + 0.119193 * G + 0.950227 * B \end{aligned} \quad (3.16)$$

$$\begin{aligned} X_r &= 0.950456 \\ Y_r &= 1.0 \\ Z_r &= 1.088754 \end{aligned} \quad (3.17)$$

$$\begin{aligned} U_r &= \frac{4 * X_r}{X_r + 15 * Y_r + 3 * Z_r} \\ V_r &= \frac{9 * Y_r}{X_r + 15 * Y_r + 3 * Z_r} \end{aligned} \quad (3.18)$$

$$L = \begin{cases} 116 * Y^{\left(\frac{1}{3}\right)} - 16, Y > 0.008856 \\ 903.3 * Y, Y \leq 0.008856 \end{cases} \quad (3.19)$$

$$u' = \frac{4 * X}{X + 15 * Y + 3 * Z}$$

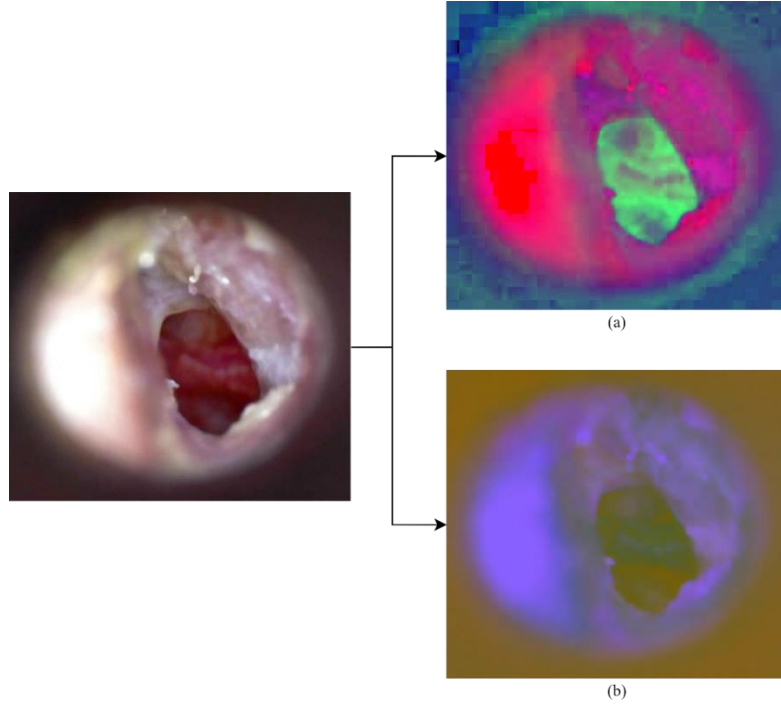
$$v' = \frac{9 * Y}{X + 15 * Y + 3 * Z} \quad (3.20)$$

$$U = 13 * L * (u' - U_r)$$

$$V = 13 * L * (v' - V_r) \quad (3.21)$$

Burada L , U ve V değerleri renk uzayını göstermektedir. U_r ve V_r değerleri belli noktadaki beyaz noktanın koordinatlarıdır. Y , koordinatı belli noktanın parlaklık değeridir. Bu işlemler uygulanarak RGB veya XYZ görüntünün rengi CIELUV formatına çevrilmektedir.

Şekil 2.18’de LUV ve HSV uygulandıktan sonra elde edilen görüntüler verilmektedir. Modeller farklı renk uzaylarında eğitilerek RGB renk uzayından farklı özellik haritaları çıkarmaktadır. Bu haritalar modellerin görüntülerdeki farklı tanımlayıcı bilgilerini sağlamaktadır.



Şekil 2.18. RGB görüntülerin HSV (a) ve LUV (b) renk uzayına çevrilmesi sonucunda elde edilen görüntüler.

2.5.Performans Metrikleri

Karışıklık matrisi, makine öğrenmesi ve derin öğrenme modellerinin performansını değerlendirmek için sınıflandırma görevlerinde kullanılan bir tablodur. Matris tablo şeklinde doğru ve yanlış sınıflandırılan tahmin sayısını vermektedir. Matrisin sütunları tahmin edilen sınıfı, satırlar ise gerçek sınıfı temsil etmektedir. Tabloda her satırda sırasıyla pozitif doğru (True Positive - DP), negatif doğru (False Negative - DN), pozitif yanlış (False Positive - YP) ve negatif yanlış (False Negative - DN) olmak üzere dört değer bulunmaktadır. Bu değerler kullanılarak modellerin doğruluk başarıları, duyarlılık, özgünlük, F-Skor, ortalama-doğruluk performans metrikleri hesaplanmaktadır. Denklem (3.22) ile Denklem (3.26)'da matematiksel formüller verilmektedir.

$$\text{Doğruluk (Accuracy)} = \frac{DP + DN}{DP + DN + YP + YN} \quad (3.22)$$

$$\text{Duyarlılık (Sensitivity)} = \frac{DP}{DP + YN} \quad (3.23)$$

$$\text{Özgünlük (Specificity)} = \frac{DN}{DN + YP} \quad (3.24)$$

$$F - \text{Skor (F - Score)} = \frac{2 \times DP}{2 \times DP + YP + YN} \quad (3.25)$$

$$\text{Ortalama - Doğruluk} = DP / (DP + DN) \quad (3.26)$$

Yapılan tez çalışmasında, DP ve DN değerleri model tarafından doğru sınıflandırılmış hastalık ve sağlıklı kulak zarı görüntülerinin sayısını vermektedir. Öte yandan, YP ve YN değerleri modelin yanlış sınıflandırdığı veya sınıfını belirleyemediği görüntülerin sayısını vermektedir. Burada dikkat edilmesi gereken unsur DP, DN, YP ve YN değerlerinin toplamının hesaplanan veri setindeki görüntü sayısına eşit olması gerekmektedir.

Doğruluk metriği, sınıflandırma modelinin başarısını değerlendirmek için kullanılan bir performans ölçüsüdür. Test veri setinde bulunan tüm görüntülerden modelin yüzdelik cinsinden doğru bulduğu etiket oranını vermektedir. Duyarlılık metriği, ikili sınıflandırma görevlerinde kullanılan bir performans ölçüsüdür. Modelin doğru şekilde tanımladığı gerçek pozitif örneklerin yüzdelik oranını vermektedir. Medikal uygulamalarda yüksek duyarlılık değeri, hastaları doğru şekilde teşhis koyabilmesi önemlidir. Özgünlük metriği, modelin doğru şekilde tanımladığı olumsuz durumların yüzdelik oranını vermektedir. Medikal

çalıřmalarda özgülñlñk deęerinin yüksek olması, hasta olmayan kiřilerin doęru bir řekilde ayırt edilmesi demektir. F-Skor metrięi, hassasiyet (Precision) ve geri çağırabilme (Recall) metriklerinin tek deęerde birleřtirerek her iki metrięe de eřit aęırlık vermektedir.

Çalıřmada kullanılan dięer bir performans metrięi de ROC grafięi ve AUC deęeridir. ROC (Receiver Operating Characteristic) grafięi, ikili sınıflandırma yapan modele ait performansın grafiksel temsilidir. Grafik farklı sınıflandırma eřiklerinde gerçek pozitif oranının (TPR) yanlış pozitif orana (FPR) karřı çizilmesiyle oluřturulmaktadır. TPR deęeri, duyarlılık metrięinden elde edilmektedir. Model tarafından doęru řekilde pozitif olarak tanımlanan gerçek pozitif örneklerin oranını ölçen bir performans ölçütüdür. FPR deęeri, yanlış bir řekilde pozitif olarak sınıflandırılan negatif örneklerin oranını ölçen ikili sınıflandırma modelinin performansını deęerlendirmek için kullanılmaktadır. Matematiksel formülü Denklem (3.27)'de verilmektedir.

$$FPR = \frac{YP}{DP + YP} \quad (3.27)$$

AUC (Area Under the Curve) deęeri, modelin pozitif ve negatif sınıfları ayırt etme yeteneęini ölçmektedir. ROC eęrisinin altındaki alan olarak hesaplanır ve 0 ile 1 arasında deęer almaktadır. Burada 0, tüm pozitif ve negatif örnekler için yanlış tahminler yapan modeli gösterirken 1 deęeri modelin tüm örnekleri doęru sınıflandırması anlamına gelmektedir.

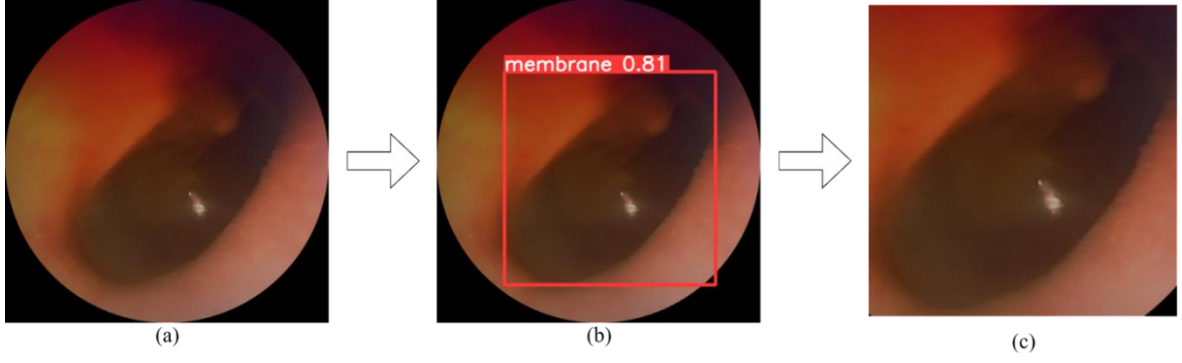
3. BULGULAR VE TARTIŞMA

Uzmanlar tarafından kulak zarının incelenmesi otoskop tarafından gerçekleştirilmektedir. Otoskopun kalitesine ve uzmanın bilgisine göre elde edilen kulak zarı görüntüleri farklılık göstermektedir. Bundan dolayı, yapılan çalışmada otoskop kalitesine ve kulak zarı görüntülerine bakılmaksızın her durumda tanı koyabilecek bir modelin oluşturulması amaçlanmıştır.

Tez çalışması kapsamında farklı derin öğrenme ağlarının kulak zarı hastalıklarını sınıflandırma performansları karşılaştırılmıştır. Kulak hastalıklarının sınıflandırılması için hazır modeller ve hibrit modeller kullanılmıştır. Modellerin performansı iki farklı veri seti üzerinde farklı senaryolar üzerinde test edilmiştir. Çalışma kapsamında oluşturulan veri setleri üzerinde performans metriklerince en yüksek değerlerin elde edildiği model başarılı olarak seçilmiştir. Ayrıca kulak zar bölgesinin seçilmesinin başarıya etkisini sınamak için zar bölgesini tespit edebilen model geliştirilmiştir.

3.1. Region Of Interest (ROI) - İlgili Alanın Çıkarılması / Zar Bölgesinin Tespiti

Ear imagery veri setiyle karşılaştırıldığında tympanic membrane veri seti, kulak içi görüntülerini daha uzak bir mesafeden elde ederek kulak kanalı ve kulak zarının geniş görünümünü yakalamıştır. Bu çalışmanın kapsamı, kulak zarını etkileyen hastalıkları sınıflandırma ile sınırlandırılmıştır, bu nedenle kulak zarı dışındaki alanlarla ilgilenilmeyecektir. Tympanic membrane veri setinde genel kulak içi yerine sadece kulak zarı bölgesinin çıkarılması için zar bölgesinin tespiti gerçekleştirilmiştir. Zar bölgesinin tespiti için YOLOv5 modeli kullanılmıştır. Model ile görüntülerdeki kulak zarı bölgeleri tanımlanmış ve bu bölgeler kesilerek yeni bir klasöre kaydedilmiştir. Sadece kulak zarı bölgesi arandığı için etiket sayısı bir olarak tutulmuştur. YOLOv5 modelinin eğitim işlemi, etiketlenmiş 300 görüntü üzerinde 50 eğitim döngüsü (epoch) olarak gerçekleştirilmiştir. Şekil 3.1'de orijinal görüntü ile model uygulandıktan sonra elde edilen kesilmiş görüntü verilmektedir.



Şekil 3.1. Kulak zarı bölgesinin tespiti; orijinal görüntü (a), kulak zarı tespiti (b) ve tespit edilen alanın kesilmesi sonucunda elde edilen görüntü (c).

3.2. Hazır Modeller

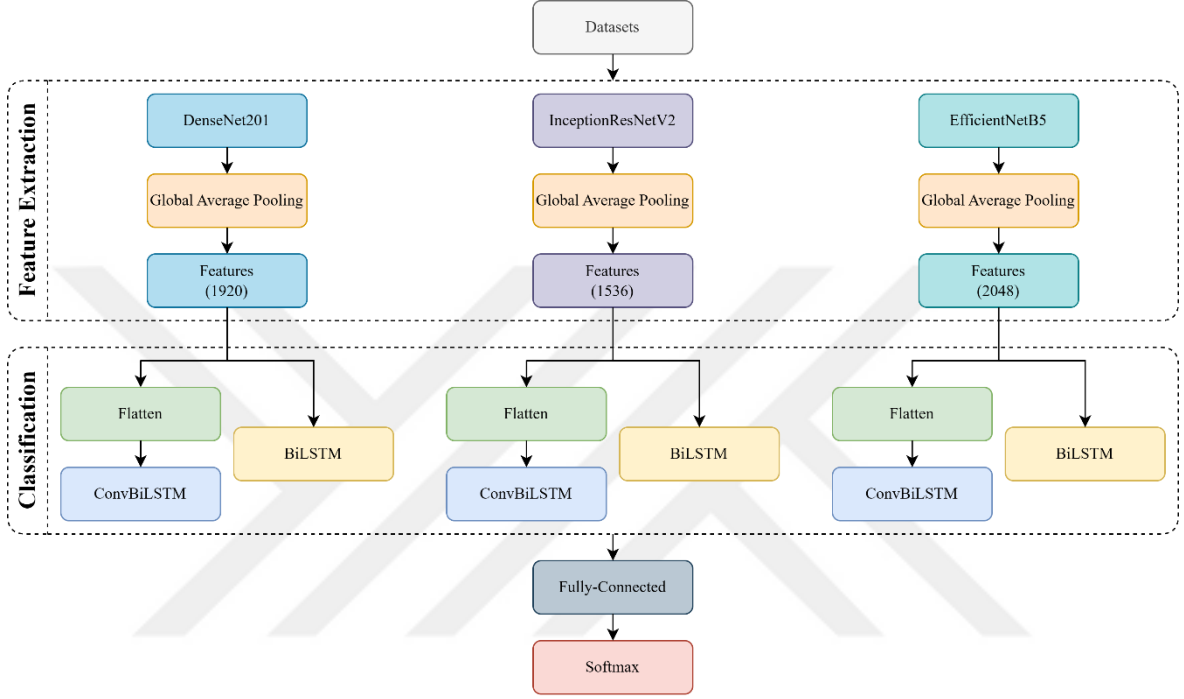
Hazır model olarak Google'ın Keras kütüphanesinde bulunan ResNet50, InceptionV3, InceptionResNetV2 ve Xception hazır modelleri veri setleri üzerinde uygulanmıştır. Hazır modeller sıfırdan eğitim ve öğrenme aktarımı (transfer learning) olmak üzere iki farklı şekilde eğitim işlemine tabi tutulmuştur. Sıfırdan eğitimde modellerin katmanlarının ağırlık değerleri rastgele bir değerle başlatmaktadır ve model girdi verisi üzerinden ağırlık değerlerini hesaplamaktadır. Öğrenme aktarımında modelin katmanlara ait ağırlık değerleri daha öncesinde ImageNet veri setinde eğitilerek hesaplanmıştır. Eğitim sonrasında modeller ayrılan test veri setlerinde uygulanarak görmediği veri üzerindeki performansları kıyaslanmıştır.

3.3. Hibrit Modeller

Tez çalışması kapsamında, veri setlerinden özellik çıkarımı ve özelliklerin sınıflandırılması işlemi için farklı modellerin birleşimi ile yeni modeller oluşturulmuştur. Bu işlem için hazır modellerin son sınıflandırma katmanları çıkarılmıştır ve son katman çıktı olarak çıkarılan özellikler haritalarını vermektedir. Çıkarılan özellikler uygun girdi formatına getirildikten sonra bir yapay sinir ağı veya makine öğrenme modeli tarafından sınıflandırılmıştır.

Özellik çıkarımı, verilerden tanımlayıcı bilgilerin belirlenmesi ve çıkarılması işlemidir. Özellik çıkarma sonucunda veri hakkında onu tanımlayan bilgi grupları elde edilmektedir. Görüntü işlemede, bir sınıfa ait görüntüyü diğer sınıflardan ayırt etmek için o sınıfa ait önemli öğelerin çıkarılması gerekmektedir. Özellik çıkarma sırasında, görsel

özelliklerin algılanmak ve çıkarmak için kenar algılama, damla algılama ve doku analizi gibi farklı teknikler kullanılmaktadır. Çalışma kapsamında DenseNet201, EfficientNetB5 ve InceptionResNetV2 modelleri kullanılarak özellik çıkarımı gerçekleştirilmiştir. Çıkarılan özellikler, BiLSTM ve ConvBiLSTM modelleri kullanılarak sınıflandırılmıştır. Oluşturulan modellerin mimarisi Şekil 3.2’de verilmektedir.



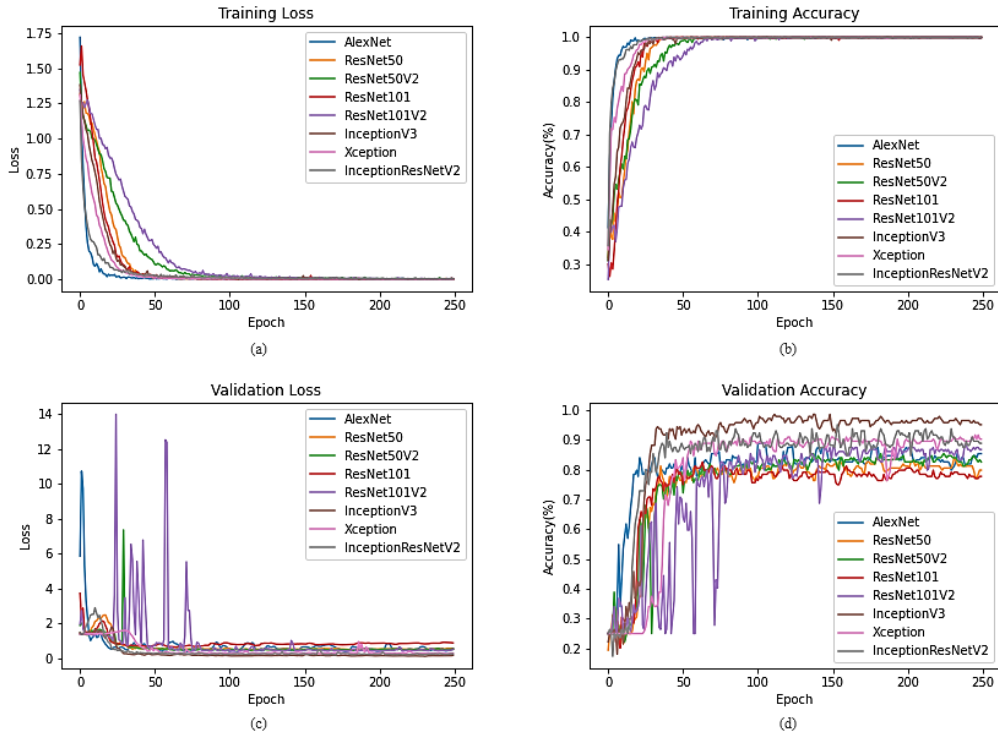
Şekil 3.2. Özellik çıkarımı ve özelliklerin sınıflandırılması için oluşturulan mimariler.

Şekil 3.2’de görüldüğü üzere her model sonucunda veri setlerinden farklı sayıda özellik haritaları elde edilmektedir. Bu özelliklerin ilk olarak global ortalama havuzlaması alınarak özellik indirme gerçekleştirilmiştir. Daha sonrasında LSTM ağına uygun formatta özellik haritası yeniden boyutlandırılmıştır. En son adımda verilerin sınıflandırılması için uygun sınıf kadar çıktı üreten Dense katmanı ve Softmax aktivasyon fonksiyonu kullanılmıştır. Çıkarılan özellikler ConvBiLSTM ağına verilmeden önce Flatten katmanı ile düzeltilmiştir. Daha sonrasında LSTM ağına kullanılan işlemler uygulanmıştır.

3.4. Modellerden Elde Edilen Bulgular

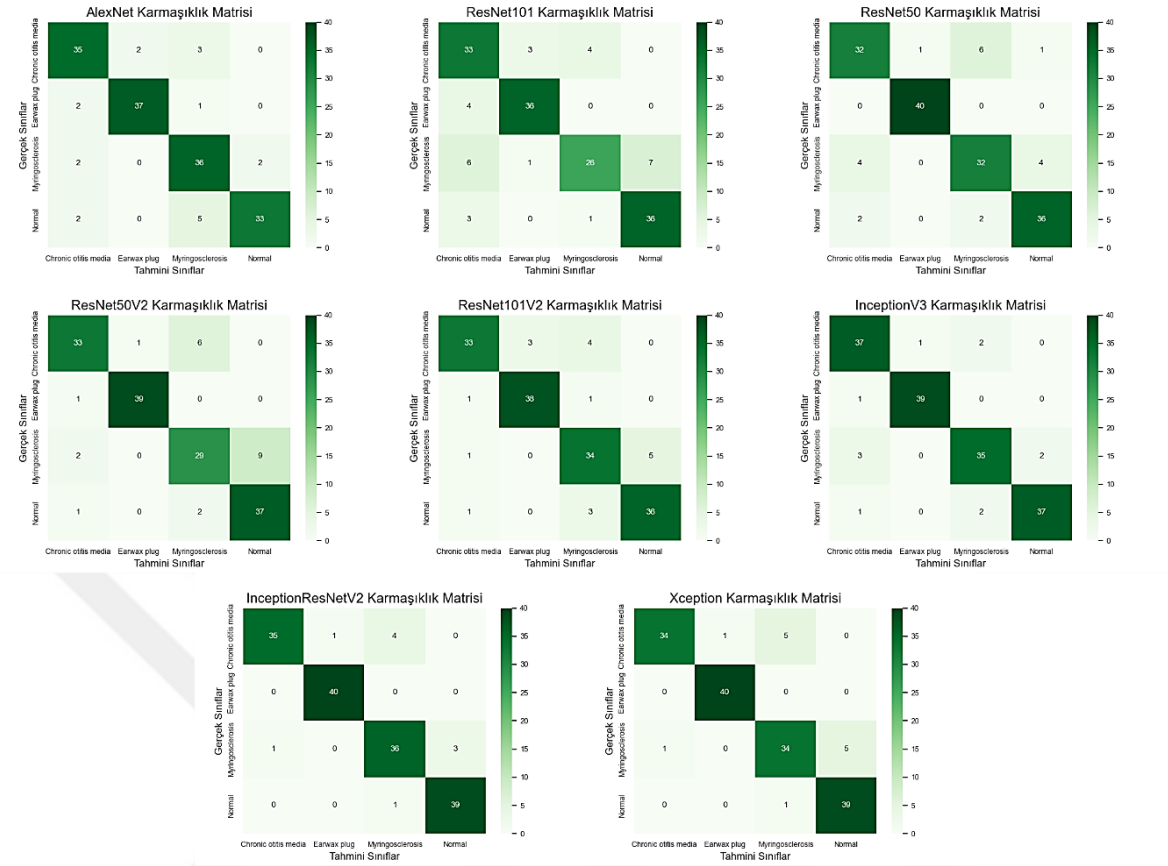
Çalışma kapsamında kullanılacak modellerin seçilmesi için öncelikle farklı modeller Ear Imagery veri setinde eğitilmiştir. Veri setindeki 880 görüntü eğitim ve test işlemleri için sırasıyla %82 (720) ve %18 (160) olarak ayrılmıştır. Eğitim kümesinde bulunan görüntülerin

%20'si doğrulama işlemi için kullanılmıştır. Eğitim işleminde SGD (Stochastic Gradient Descent) algoritması kullanılarak algoritmanın öğrenme oranı 0,0001 ve momentum değeri 0,8 olarak ayarlanmıştır. Optimizasyon sırasında modellerin katmanlarına L2 düzenleyiciyle (L2 regularizer) 0,001 oranında penaltı uygulanarak modellerin aşırı öğrenmesi engellenmiştir. Model olarak AlexNet, ResNet50, ResNet50V2, ResNet101, ResNet101V2, InceptionV3, InceptionResNetV2 ve Xception mimarileri kullanılmıştır. Sıfırdan eğitim sonucunda elde edilen eğitim ve doğrulama işlemlerine ait doğruluk ve kayıp grafikleri Şekil 3.3'de verilmektedir.



Şekil 3.3. Modellere ait eğitim ve doğrulama grafikleri, (a) eğitim kaybı, (b) eğitim doğruluğu, (c) doğrulama kaybı ve (d) doğrulama doğruluğu.

Şekil 3.3'de görüldüğü üzere ResNet50V2 ve ResNet101V2 mimarilerine ait doğrulama işleminin kayıp ve doğruluk değerlerinde zıplamalar bulunmaktadır. Ayrıca Şekil 3.4'de modellerin test işlemi sonucunda elde edilen karışıklık matrisi verilmektedir.



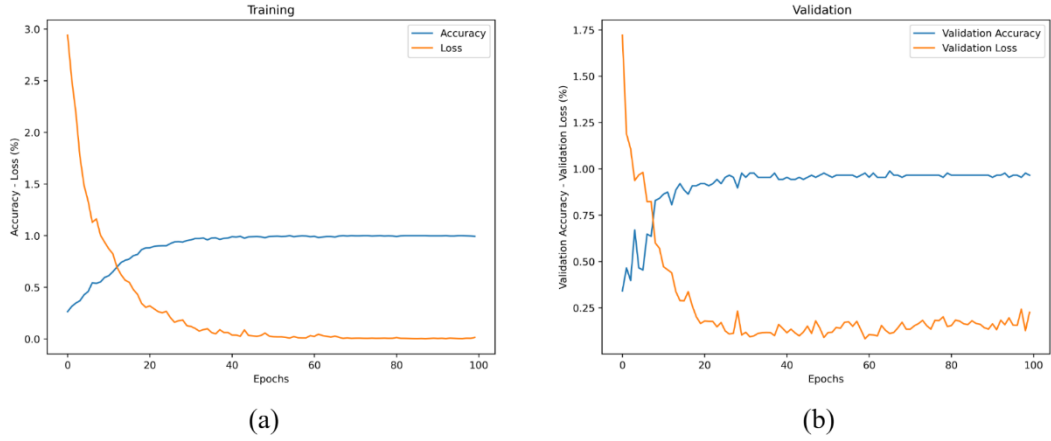
Şekil 3.4. Modellere ait test işlemi karışıklık matrisi.

Tablo 3.1. İlgili modellerden elde edilen performans metrikleri.

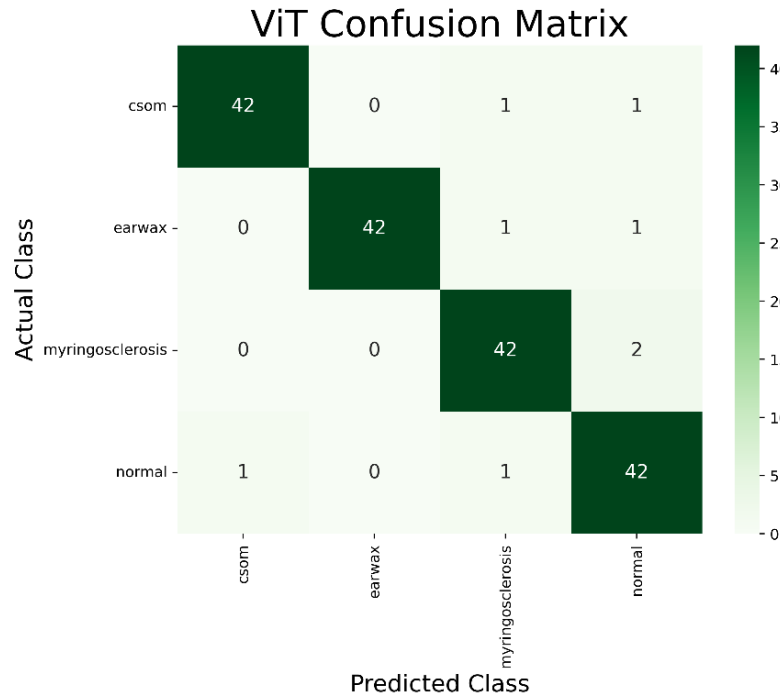
Modeller	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-SCOR (%)	Ortalama - Doğruluk (%)
AlexNet	88	88	96	88	89
ResNet50	88	88	95	87	87
ResNet50V2	86	86	95	86	86
ResNet101	82	82	94	82	82
ResNet101V2	88	88	96	88	88
InceptionV3	93	93	97	93	93
Xception	92	92	97	92	92
InceptionResNetV2	94	94	98	94	94

Test işleminde en yüksek başarıyı gösteren dört model yapılacak çalışmalarda kullanılmak üzere seçilmiştir. Tablo 3.1’de görüldüğü üzere bu modeller ResNet50,

InceptionV3, Xception ve InceptionResNetV2 mimarileridir. ResNet50, ResNet101V2 ve AlexNet mimarileri aynı başarıya sahip olmasına karşın ResNet50 modeli seçilmiştir. Bunun en büyük nedeni bu model hem küçük ölçeklidir hem de atlama katmanlarıyla kaybolan gradyandan daha az etkilenmektedir. Bu mimarilere ek olarak günümüzde son teknoloji olan Vision Transformer (ViT) mimarisi kullanılmıştır. Ear Imagery veri setinde sıfırdan eğitim sonucunda elde edilen eğitim ve doğrulama grafiği Şekil 3.5'te verilmektedir. Şekil 3.6'da test işlemi sonucunda elde edilen karışıklık matrisi verilmektedir.



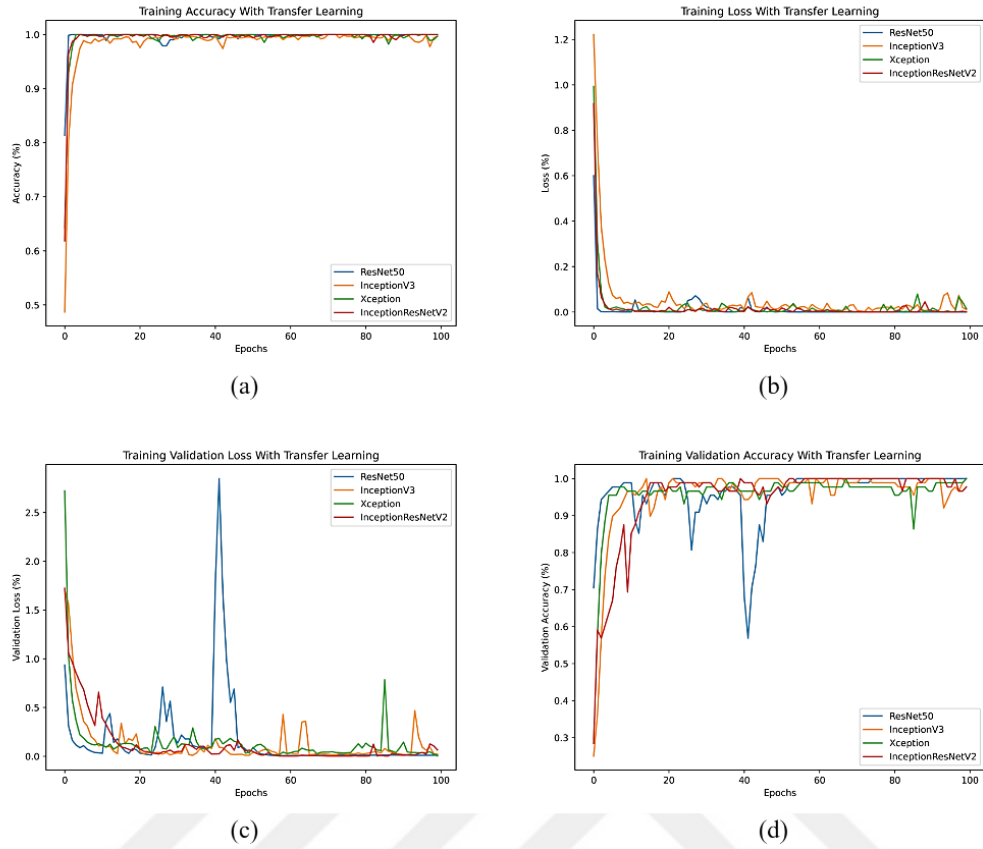
Şekil 3.5. ViT mimarisi ear imagery veri seti eğitim (a) ve doğrulama (b) grafiği.



Şekil 3.6. ViT mimarisi ear imagery veri seti karışıklık matrisi.

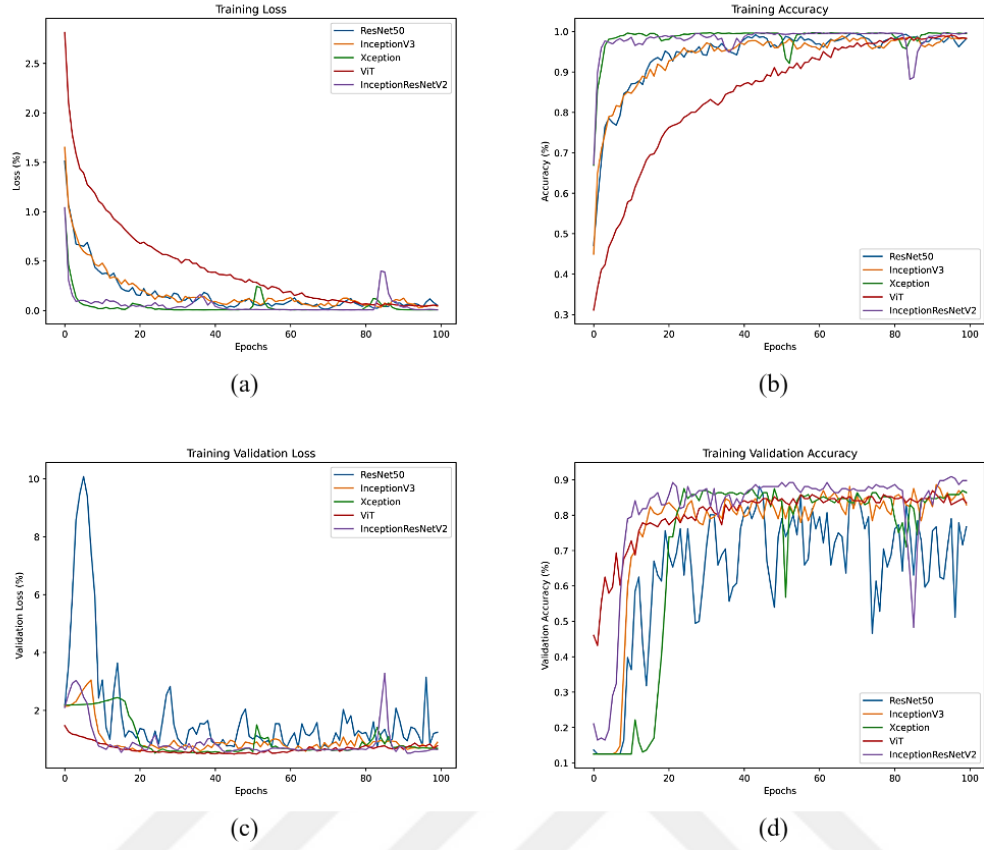
Şekil 3.5 ve 3.6’da görüldüğü üzere son teknoloji Vision Transformer modeli günümüzde çok kullanılan CNN modellerine yaklaşan sınıflandırma performansı göstermektedir. Seçilen modeller Ear Imagery ve Tympanic Membrane veri setlerinden oluşturulan veri setleri üzerinde renk uzayı ve artırma işlemlerinin başarıya etkisini kıyaslamak için kullanılmaktadır. Yapılan çalışmada tüm modellerin hiperparametre değerleri karşılaştırılabilmeleri için eş tutulmuştur. Eğitim işlemi 100 epoch olarak gerçekleştirilmiştir. AdamW (Adam with Weight Decay) optimizasyon algoritması kullanılarak gerçekleştirilmiştir. AdamW optimizasyon algoritmasında ağırlık kaybı değeri Adam algoritmasından farklı olarak ağırlık güncellemeleri yerine ağırlıkların kendisine uygulanmaktadır. Öğrenme katsayısı ve ağırlık kaybı değerleri 0,0001 olarak alınmıştır. Tüm işlemler NVIDIA RTX 3060 6GB ekran kartına sahip dizüstü bilgisayar üzerinde CUDA mimarisi kullanılarak gerçekleştirilmiştir.

Ayrıca yapılan çalışmada sıfırdan ve öğrenme aktarımı (transfer learning) kullanılarak eğitim işlemleri gerçekleştirilmiştir. Öğrenme aktarımı için modellerin ImageNet veri setinde önceden eğitilmiş ağırlık değerleri kullanılmıştır. Bu, yetersiz veri sayısı bulunan çalışmalarda başarıyı artırmaktadır. Öğrenme aktarımı RGB görüntüler üzerinde kullanıldığı için HSV ve LUV renk uzaylarına ait çalışma değerlerinde kullanılmamıştır. Şekil 3.7-3.15’de RGB renk uzayında sıfırdan ve öğrenme aktarımı kullanılarak elde edilen eğitim grafikleri verilmektedir. ViT mimarisinin öğrenme aktarımı için kullanılabilecek ImageNet üzerinde önceden eğitilmiş modeli bulunmadığından dolayı sadece sıfırdan eğitim işlemlerinde dahil edilmiştir.

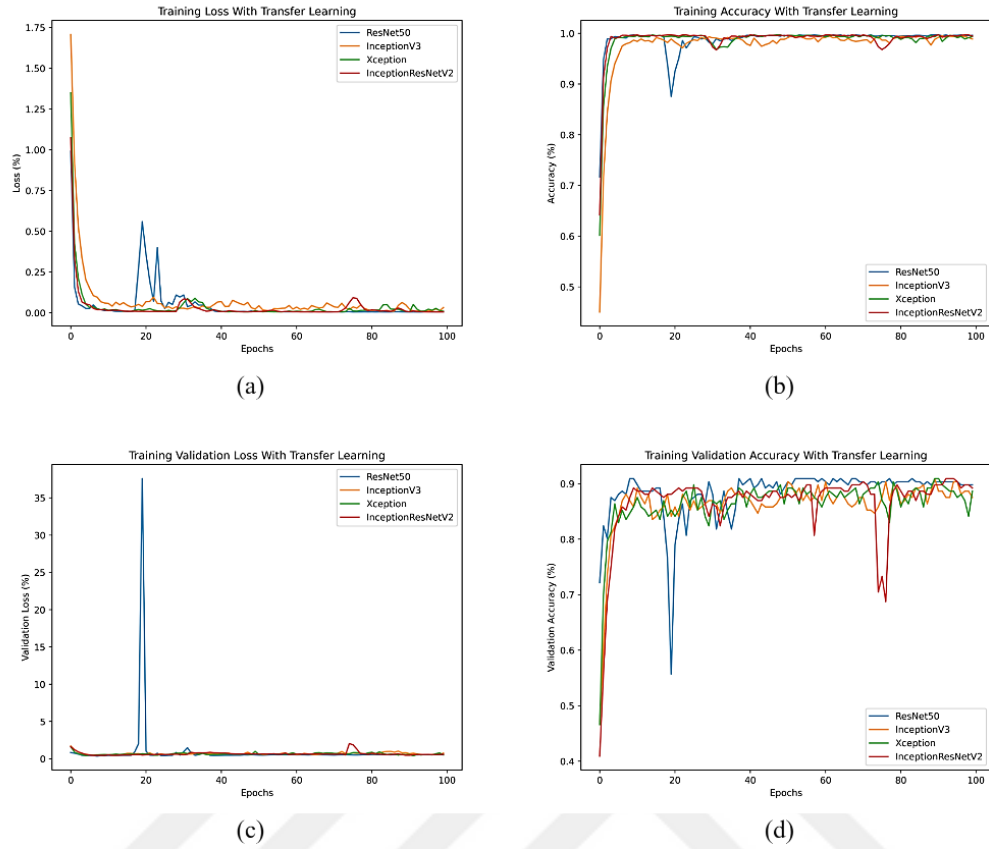


Şekil 3.7. Ear imagery veri seti öğrenme aktarımı ile eğitim sonucunda elde edilen eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).

Şekil 3.7’de görüldüğü üzere sıfırdan eğitime göre öğrenme aktarımı eğitimde ve doğrulamada daha yüksek doğruluk başarısına erişmektedir. ResNet50 mimarisi doğrulama işleme sırasında düşüşler yaşamaktadır fakat eğitim ilerleyen aşamalarında model ağırlık değerlerini iyileştirerek başarısını artırmıştır.

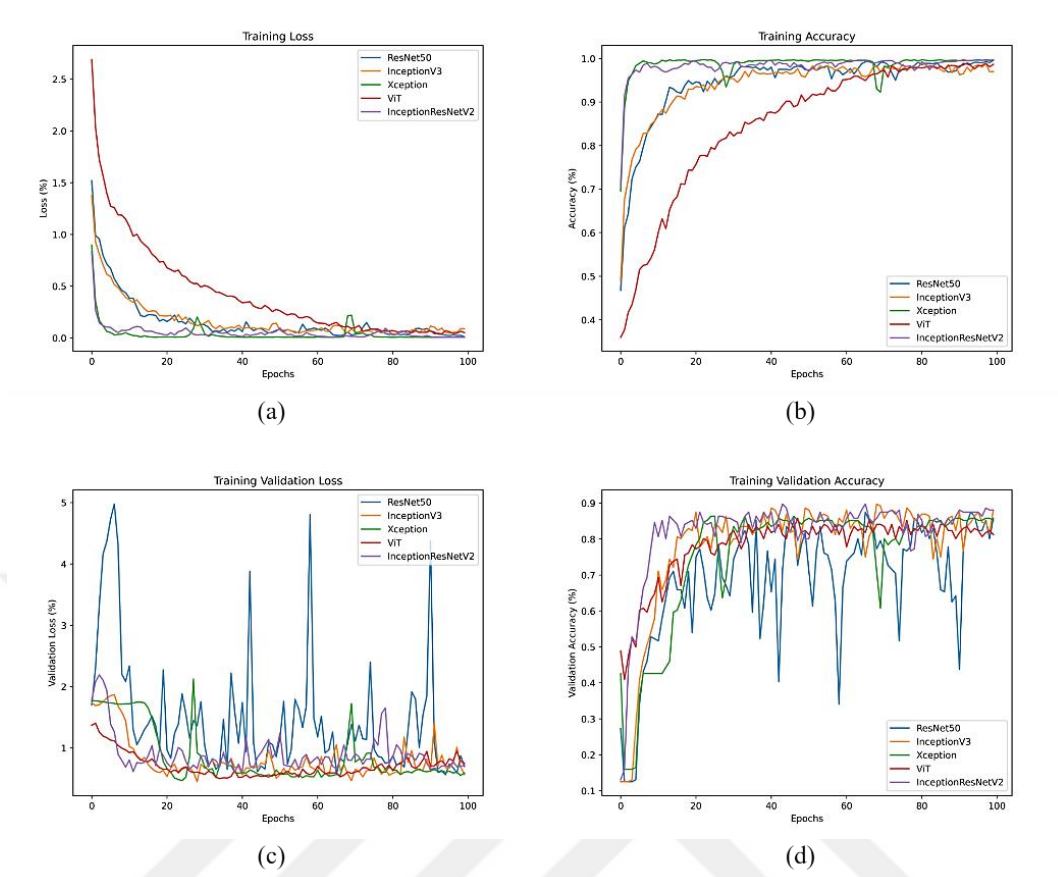


Şekil 3.8. Merged_ds veri setine ait sıfırdan eğitim sonucunda elde edilen eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).

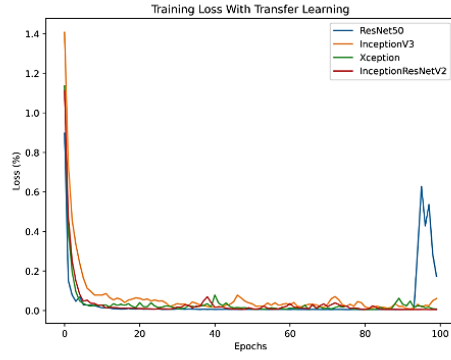


Şekil 3.9. Merged_ds veri setine ait öğrenme aktarımıyla eğitim sonucunda elde edilen eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).

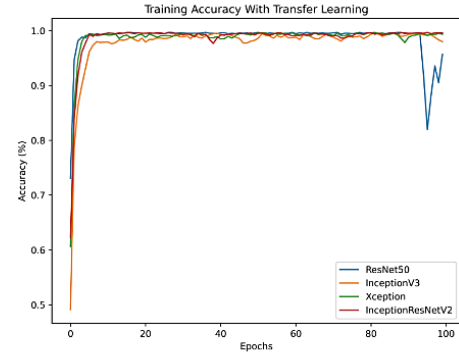
Şekil 3.8 ve 3.9'dan görüldüğü üzere sıfırdan eğitimde görülen sapmalar öğrenme aktarımında kaybolmaktadır ve hem eğitim hem de doğrulama işlemi daha düzenli ilerlemektedir. ResNet50 mimarisi sıfırdan eğitimde çok fazla sapma yaşarken öğrenme aktarımı kullanıldığında bu sapmalar kaybolmaktadır. Ayrıca tüm modeller eğitim işleminde %100 doğruluk başarısına sıfırdan eğitime göre daha hızlı erişmektedir. Bu modellerin eğitim süresini azalmaktadır.



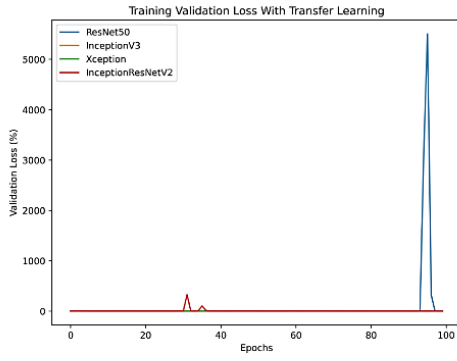
Şekil 3.10. Merged_ds (6 sınıf) veri setine ait sıfırdan eğitim sonucunda elde edilen eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).



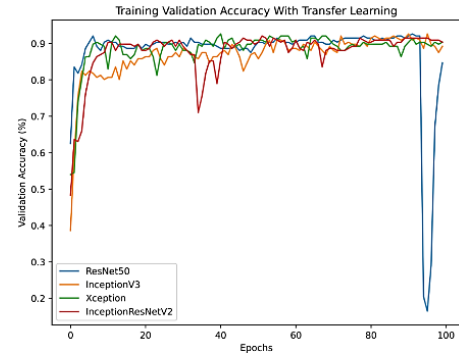
(a)



(b)

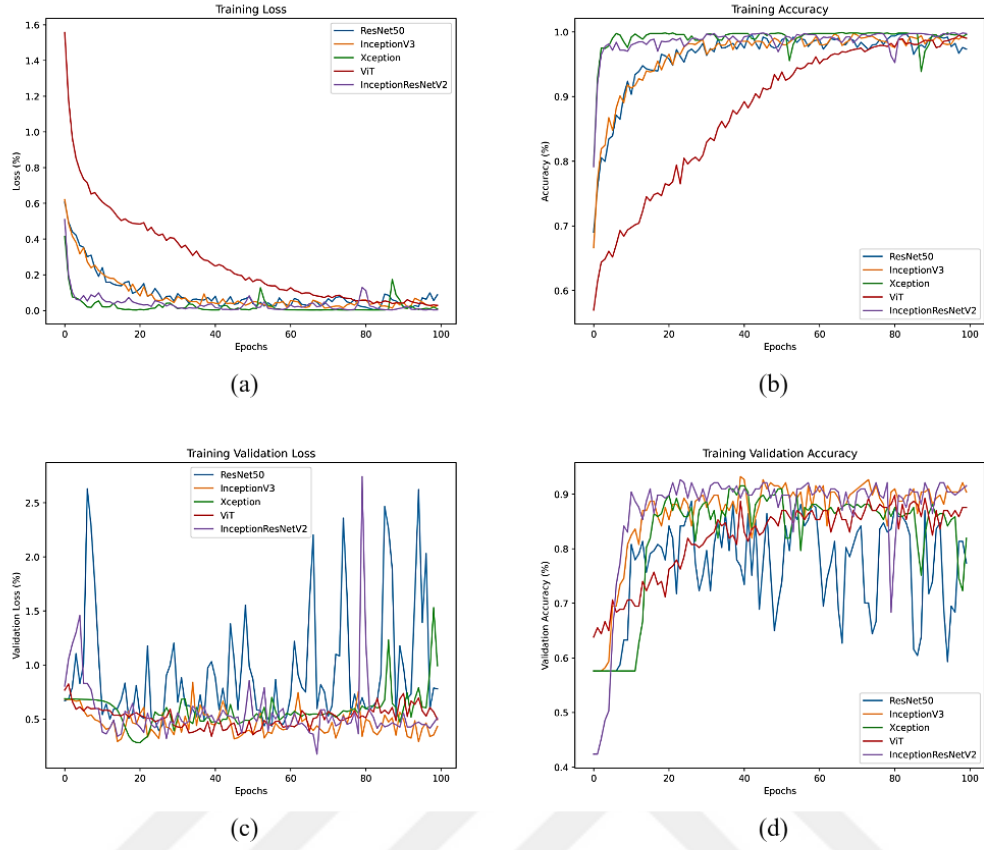


(c)

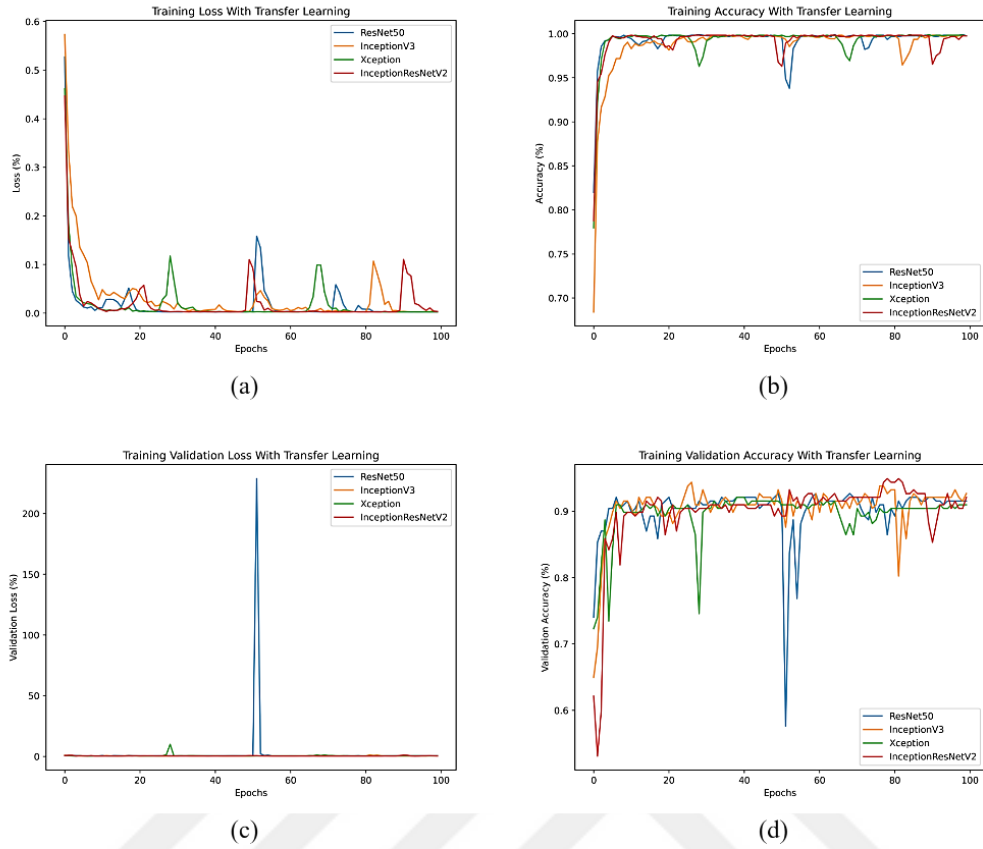


(d)

Şekil 3.11. Merged_ds (6 sınıf) veri setine ait öğrenme aktarımıyla eğitim sonucunda elde edilen eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).

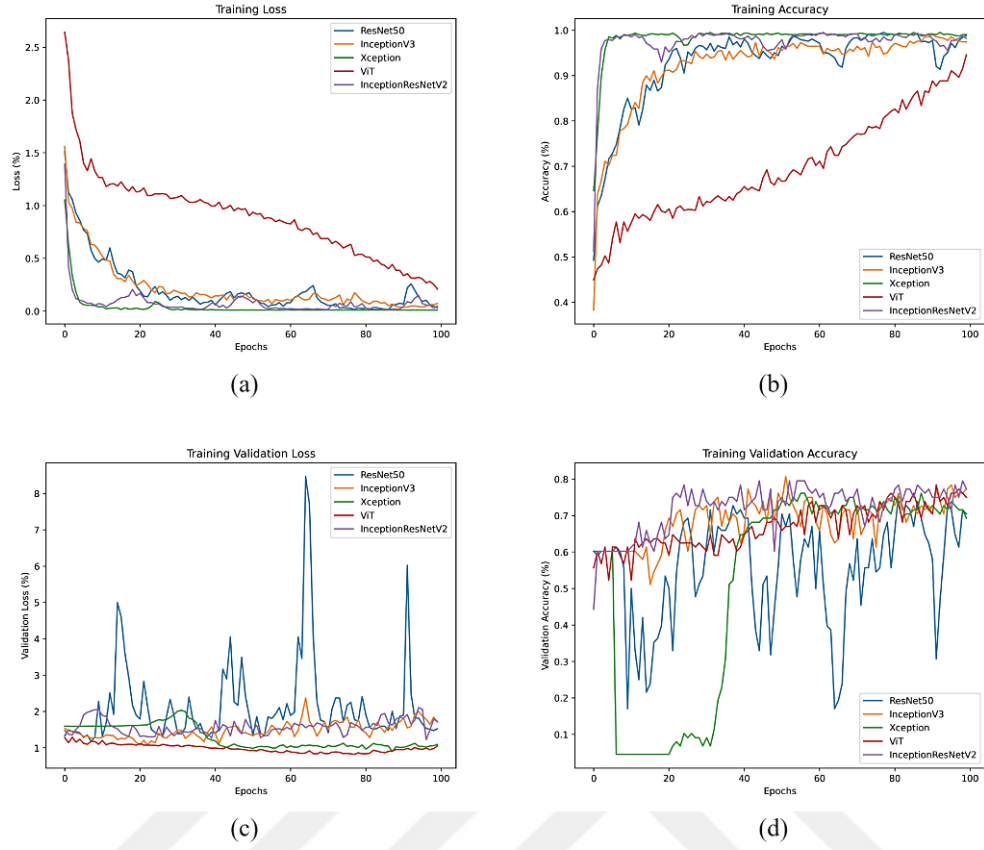


Şekil 3.12. Merged_ds (2 sınıf) veri setine ait sıfırdan eğitim sonucunda elde edilen eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).

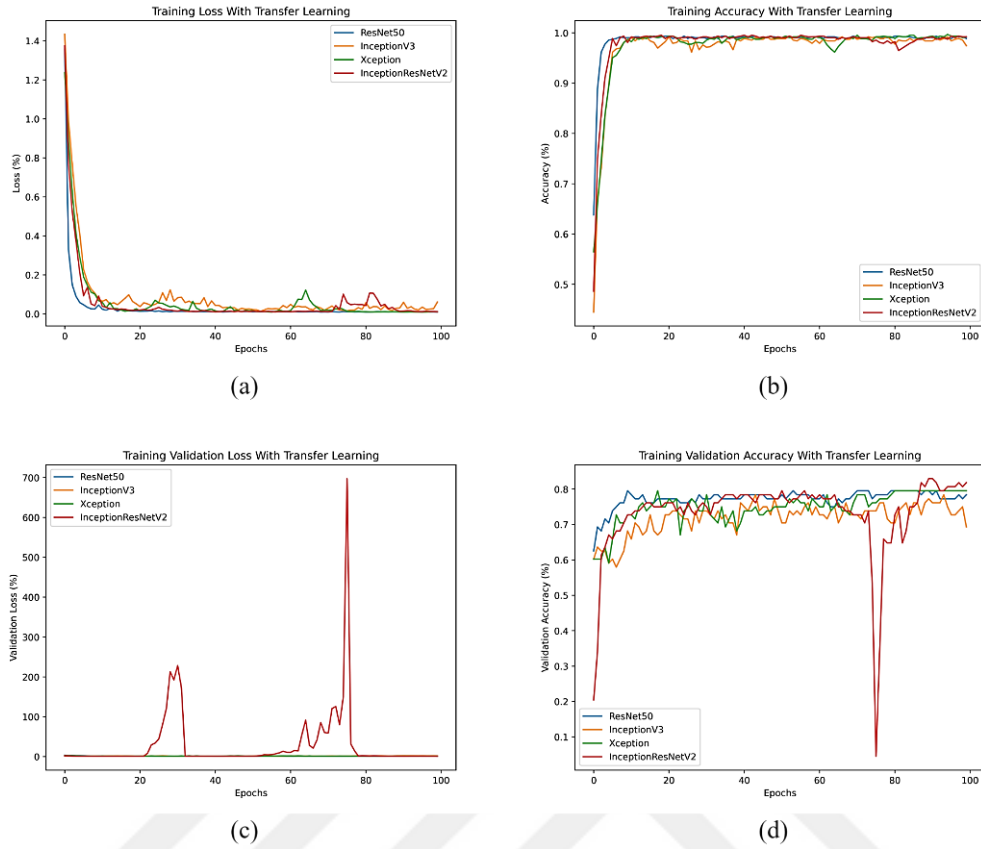


Şekil 3.13. Merged_ds (2 sınıf) veri setine ait öğrenme aktarımıyla eğitim sonucunda elde edilen eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).

Şekil 3.10-3.13’de görüldüğü üzere dokuz sınıf için elde edilen bulgular altı sınıf ve iki sınıf içinde geçerlidir. ResNet50 modeli eğitim sırasında yüksek hata değerlerine çıkmasına karşın model eğitimin ilerleyen sürelerinde ağırlık değerlerini düzeltmiştir.



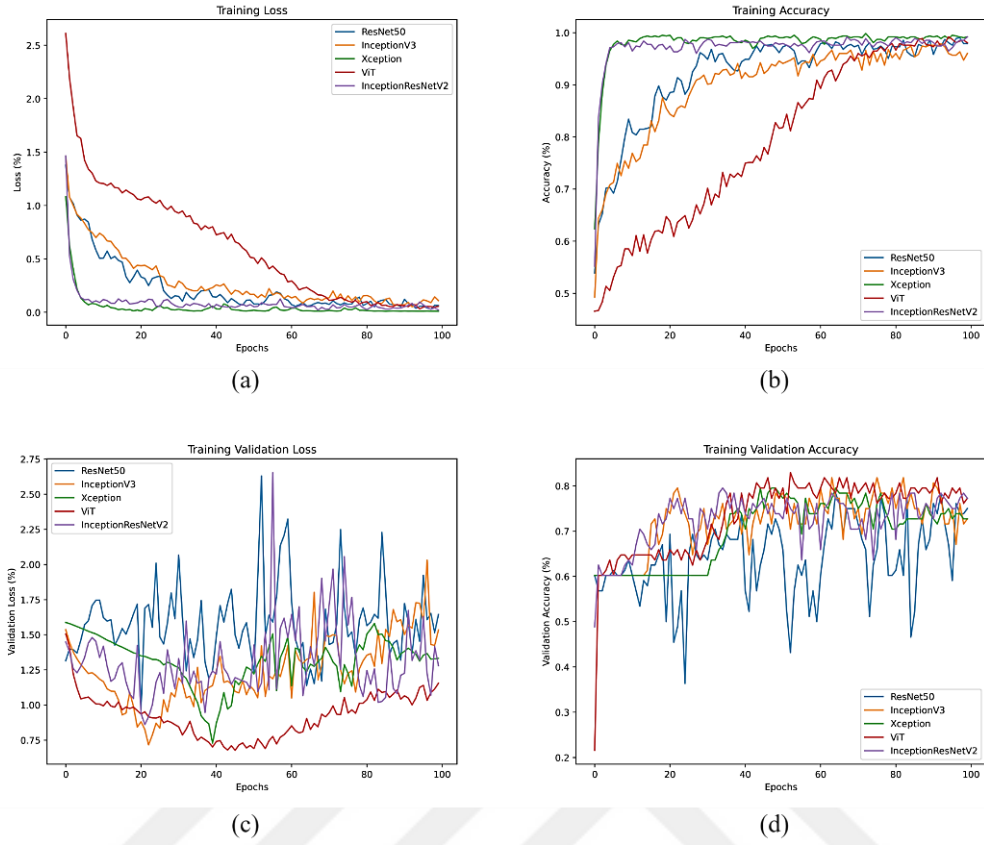
Şekil 3.14. Tympanic membrane veri setine ait sıfırdan eğitim sonucunda elde edilen eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).



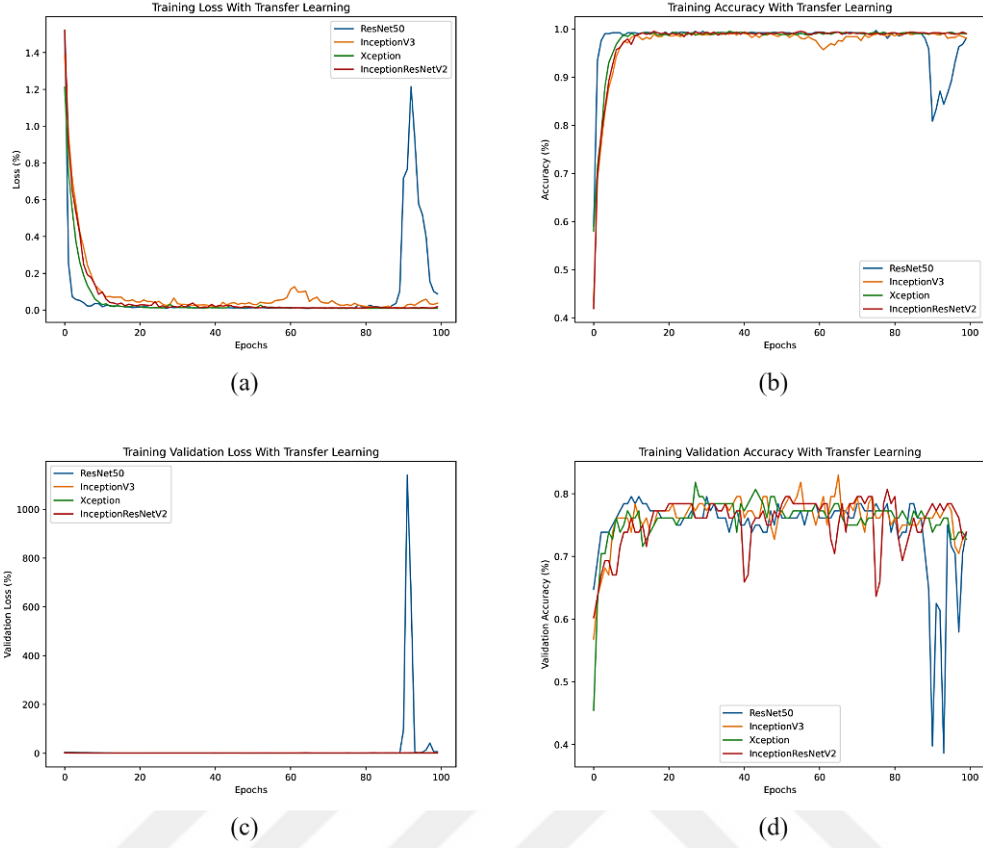
Şekil 3.15. Tympanic membrane veri setine ait öğrenme aktarımıyla eğitim sonucunda elde edilen eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).

Şekil 3.14 ve 3.15'te görüldüğü üzere modeller Tympanic Membrane veri setinde Ear Imagery veri setine göre daha düşük sonuçlar elde edilmektedir. Öğrenme aktarımı kullanıldığında modellerin yaşadığı sapma azalmaktadır. InceptionResNetV2 mimarisi doğrulama işleminde yüksek sapmalar yaşamaktadır. Bu problem, modellerin aşırı öğrenmesi sonucunda gerçekleşmektedir. Eğitimin erken bitirilmesi bu problemi çözebilmektedir.

Ayrıca çalışma kapsamında Tympanic Membrane veri setinde kulak zarı bölgesinin YOLOv5 ile kesilmesi sonucunda ayrı bir veri seti oluşturulmuştur. Bu veri setinde modelleri çalıştırdıktan sonra elde edilen grafikler Şekil 3.16 ve 3.17'de verilmektedir.

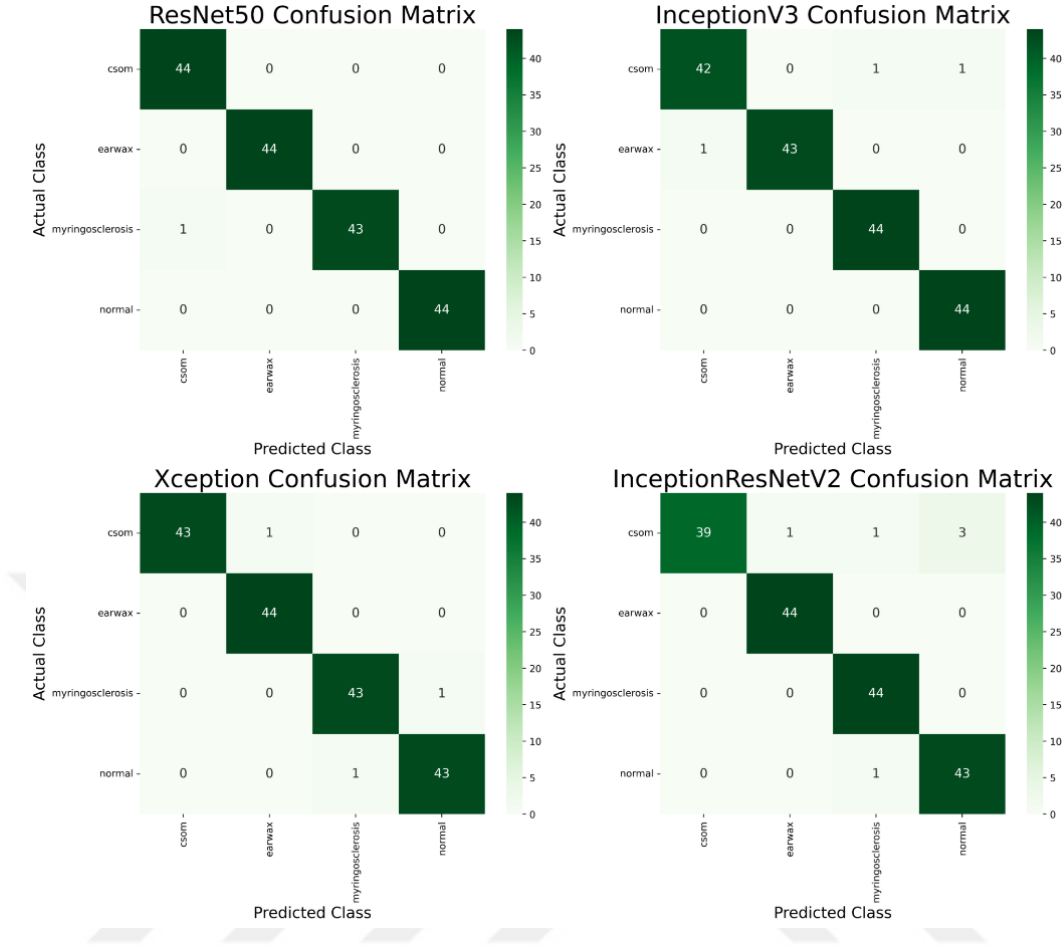


Şekil 3.16. YOLOv5 uygulandıktan sonra tympanic membrane veri setine ait sıfırdan eğitim sonucunda elde edilen eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).

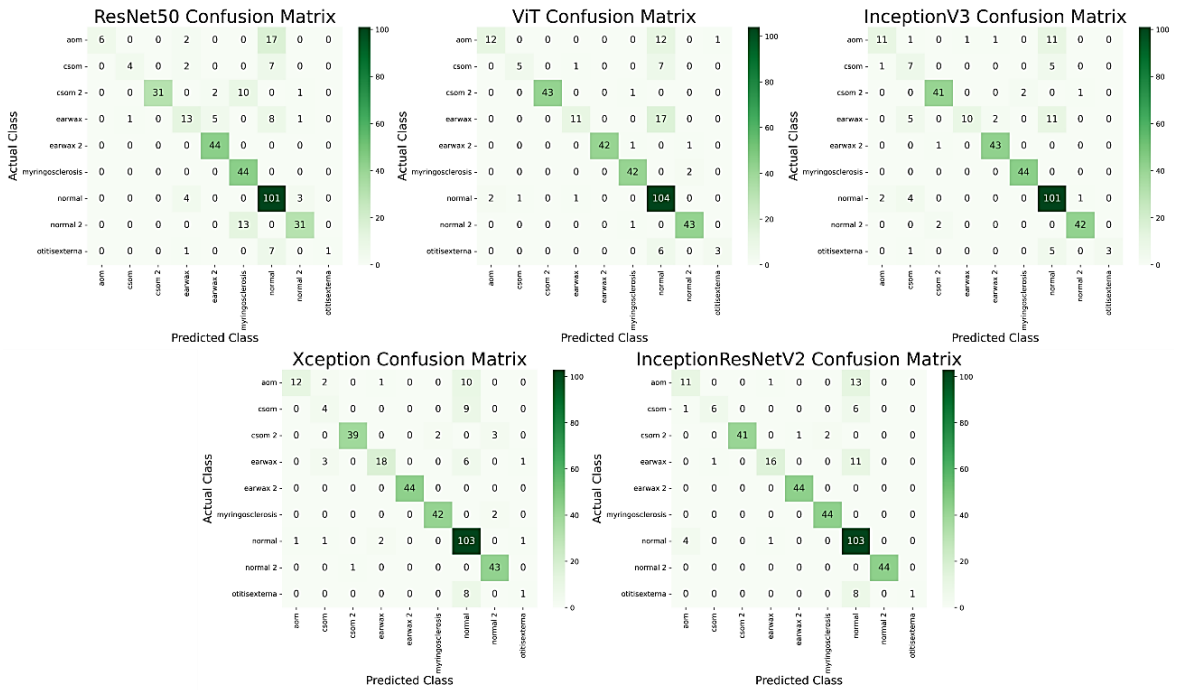


Şekil 3.17. YOLOv5 uygulandıktan sonra tympanic membrane veri setine ait öğrenme aktarımıyla eğitim sonucunda elde edilen eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).

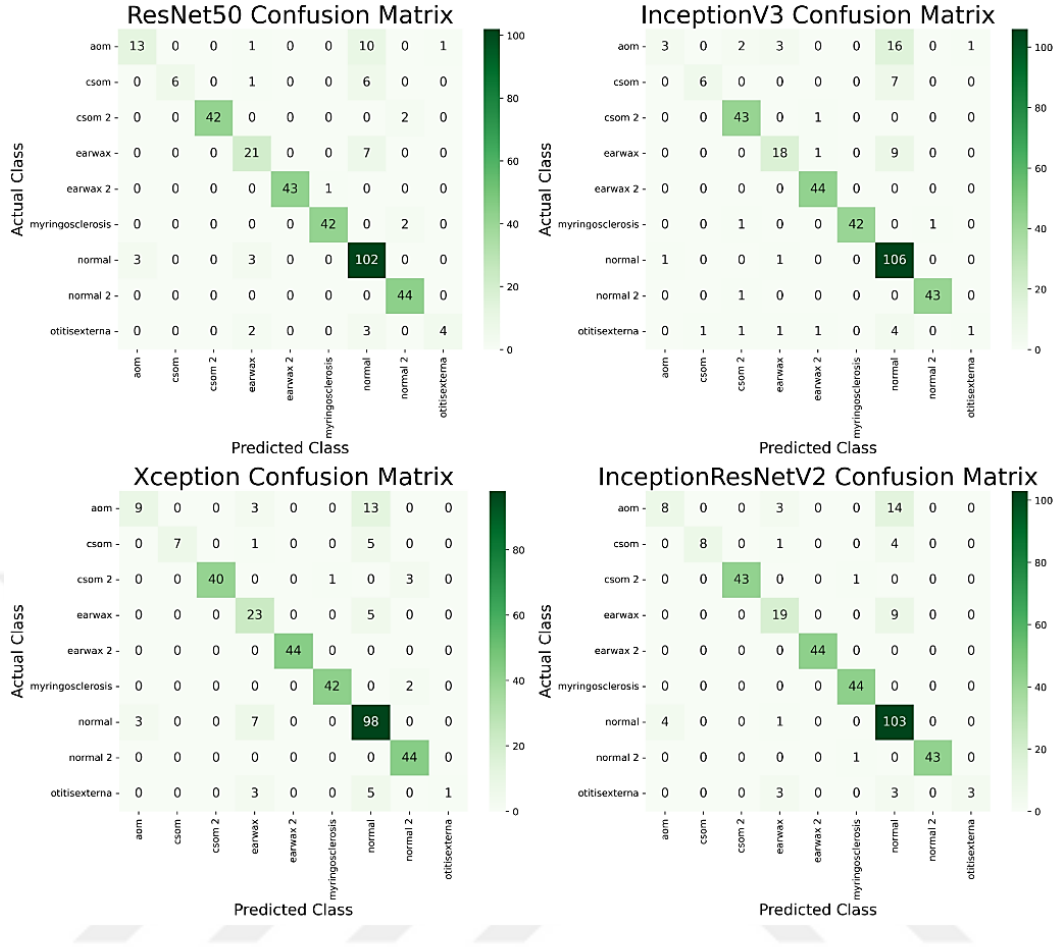
Seçilen modeller, eğitiminden sonra test veri seti üzerinde uygulanmıştır. Test veri seti toplam görüntülerin %20'si seçilerek oluşturulmuştur. Sınıflandırma işleminde sınıf bazında ve hastalıklı/sağlıklı olarak iki farklı etikete göre gerçekleştirilmiştir. Modellerim sırasıyla Ear Imagery, Merged_ds, Merged_ds (6 sınıf), Merged_ds (2 sınıf), Tympanic Membrane ve YOLOv5 uygulanmış Tympanic Membrane veri setlerine uygulandıktan sonra elde edilen karışıklık matrisleri Şekil 3.18-3.28'de verilmektedir. Verilen karışıklık matrisleri şekillerinde ağırlık değerlerinin sıfırdan eğitilmesi ve öğrenme aktarımı kullanıldıktan sonra elde edilen değerler bulunmaktadır.



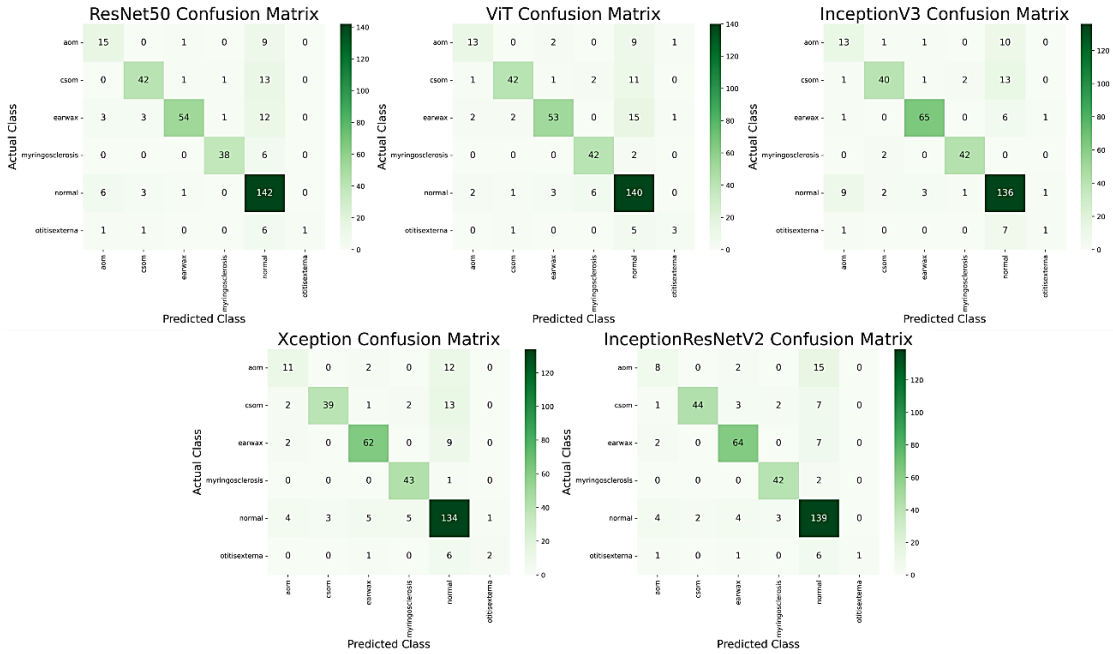
Şekil 3.18. Ear imagery veri setinde öğrenme aktarımı kullanılarak elde edilen karışıklık matrisi.



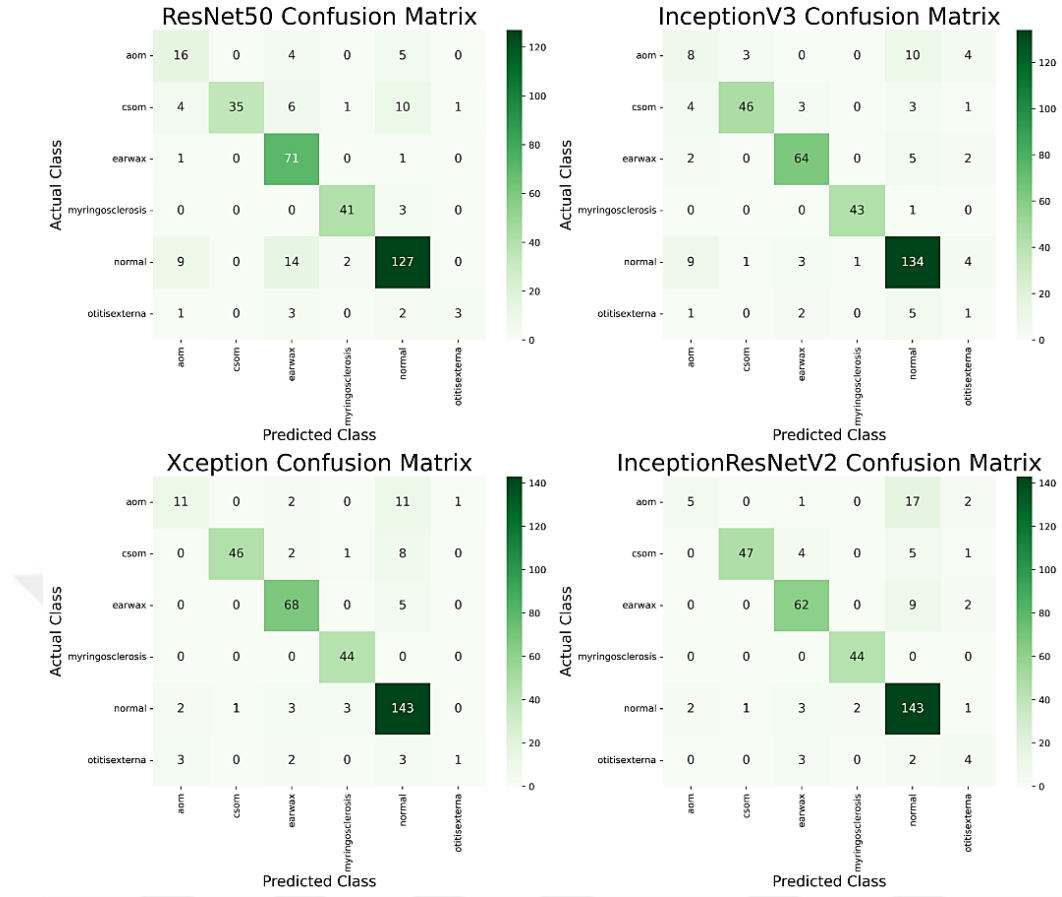
Şekil 3.19. Merged_ds veri setinde sıfırdan eğitim kullanılarak elde edilen karışıklık matrisi.



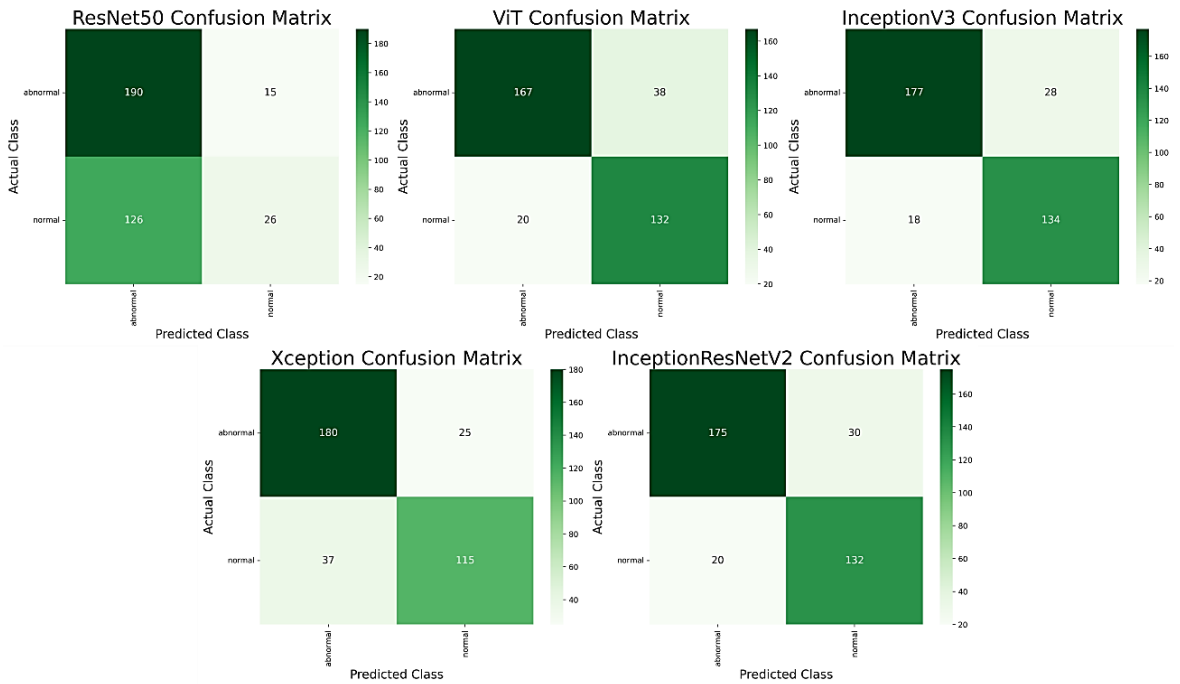
Şekil 3.20. Merged_ds veri setinde öğrenme aktarımı kullanılarak elde edilen karışıklık matrisi.



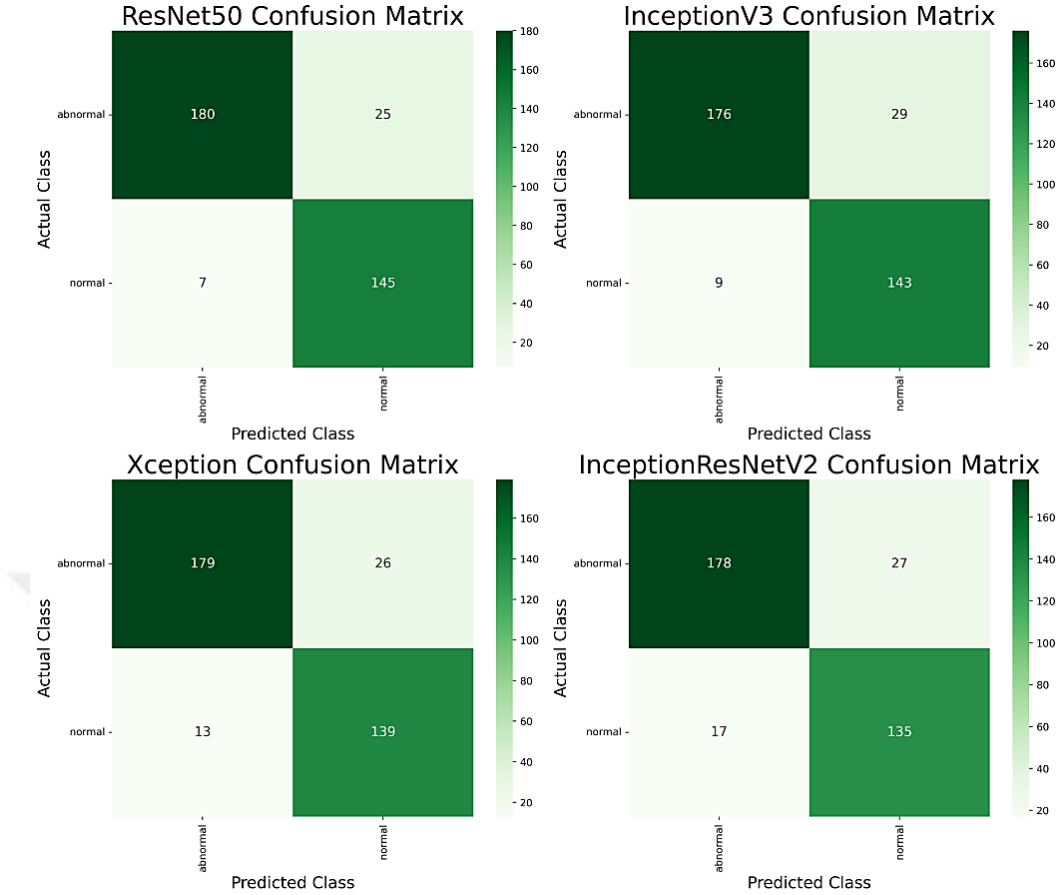
Şekil 3.21. Merged_ds (6 sınıf) veri setinde sıfırdan eğitim kullanılarak elde edilen karışıklık matrisi.



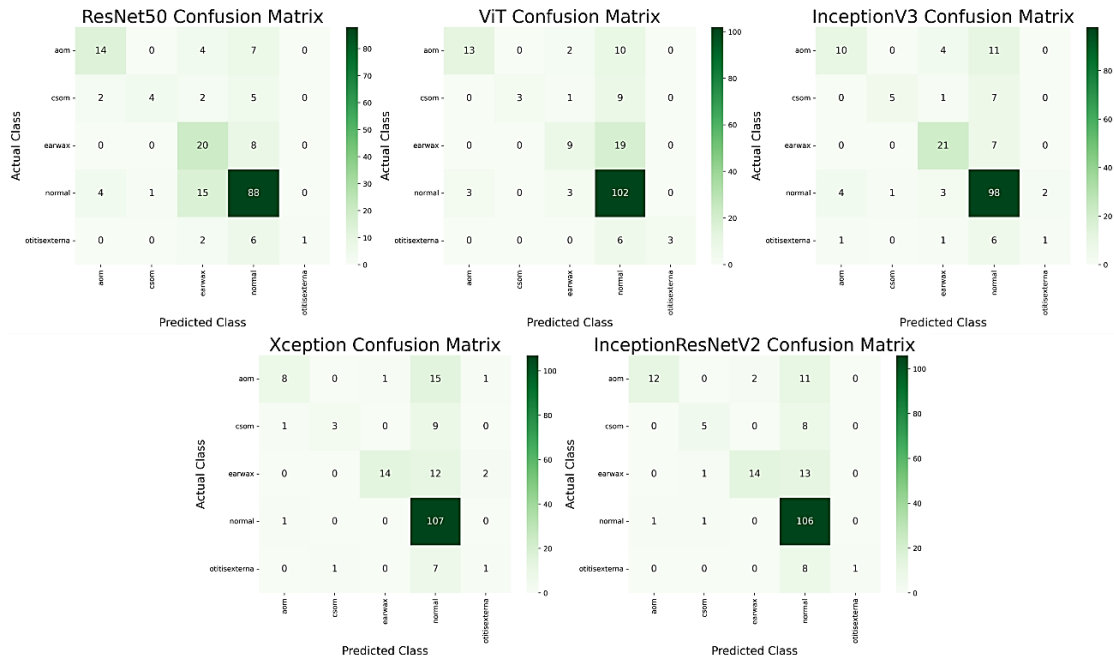
Şekil 3.22. Merged_ds (6 sınıf) veri setinde öğrenme aktarımı kullanılarak elde edilen karışıklık matrisi.



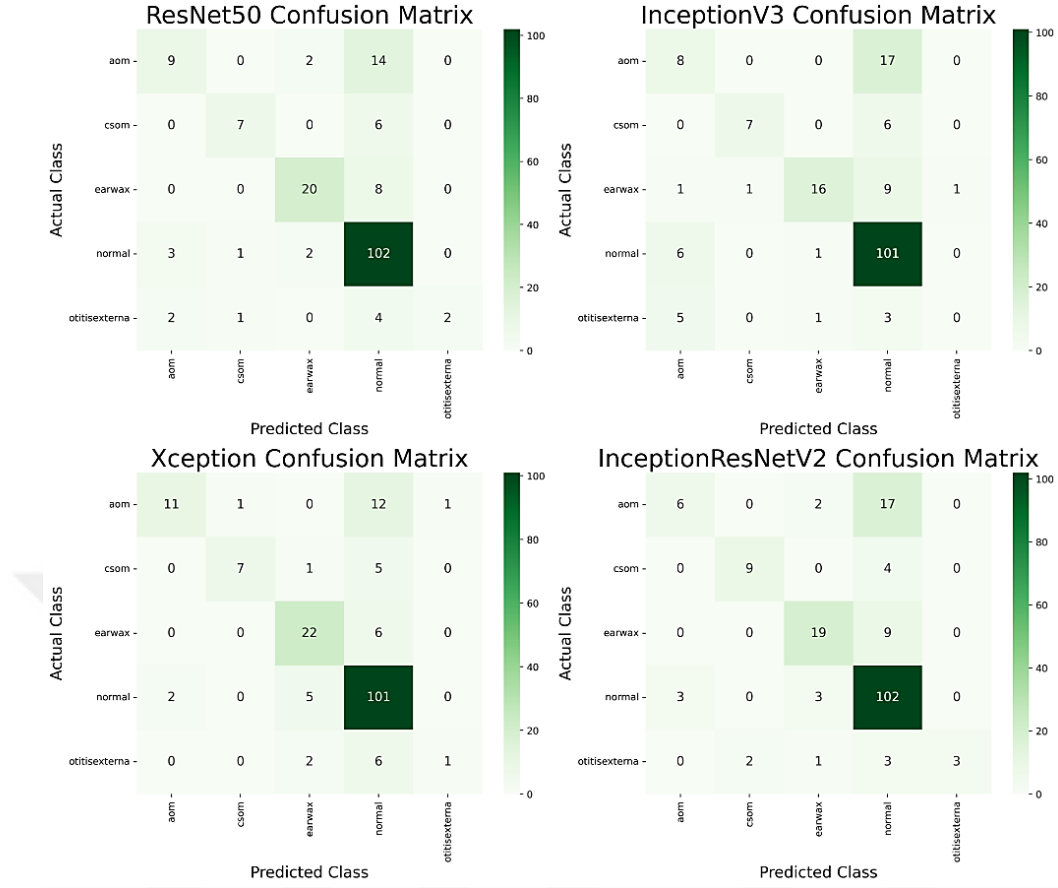
Şekil 3.23. Merged_ds (2 sınıf) veri setinde sıfırdan eğitim kullanılarak elde edilen karışıklık matrisi.



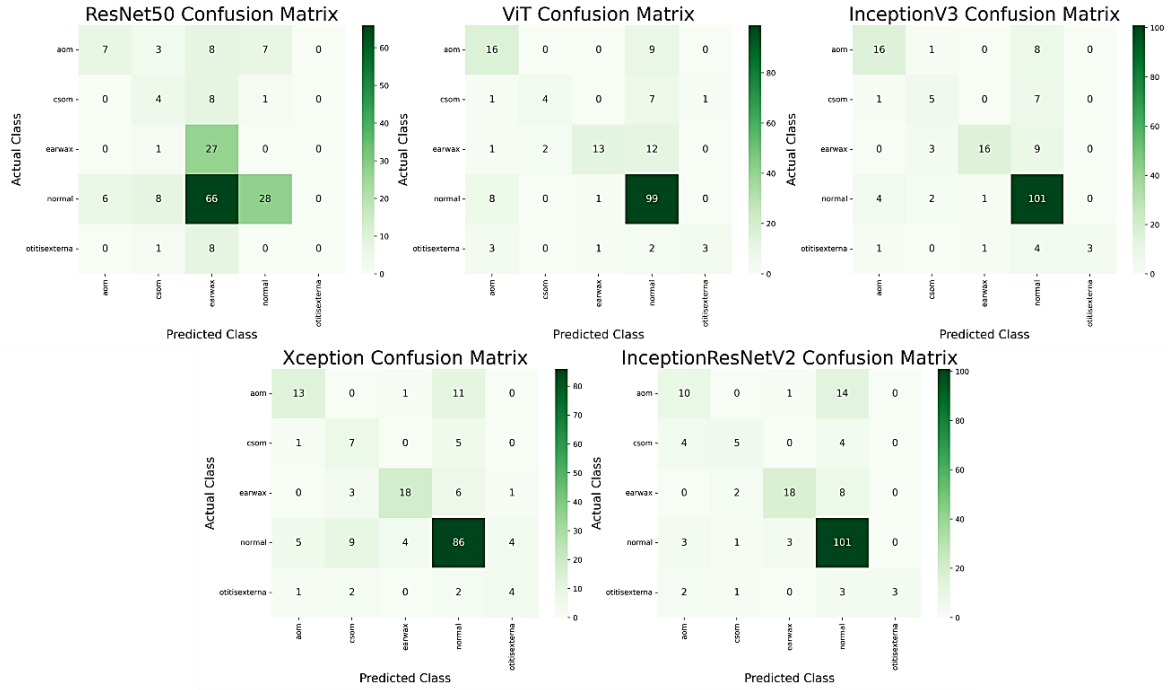
Şekil 3.24. Merged_ds (2 sınıf) veri setinde öğrenme aktarımı kullanılarak elde edilen karışıklık matrisi.



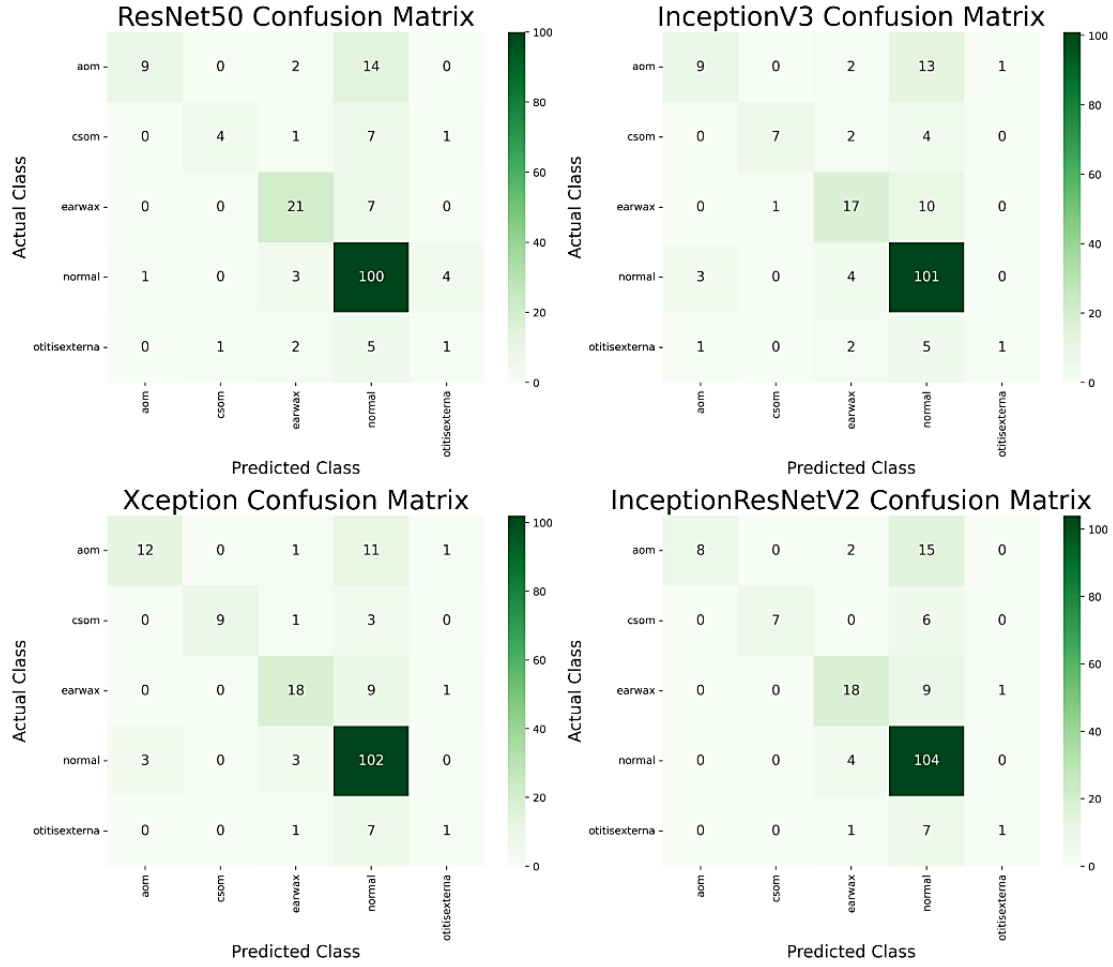
Şekil 3.25. Tympanic membrane veri setinde sıfırdan eğitim kullanılarak elde edilen karışıklık matrisi.



Şekil 3.26. Tympanic membrane veri setinde öğrenme aktarımı kullanılarak elde edilen karışıklık matrisi.



Şekil 3.27. YOLOv5 uygulanmış tympanic membrane veri setinde sıfırdan eğitim kullanılarak elde edilen karışıklık matrisi.



Şekil 3.28. YOLOv5 uygulanmış tympanic membrane veri setinde öğrenme aktarımı kullanılarak elde edilen karışıklık matrisi.

Şekil 3.18-3.28’de verilen karışıklık matrislerinden elde edilen sonuçlara göre öğrenme aktarımı kullanımı modellerin genel başarısını artırmaktadır. Ear Imagery veri setinde genel başarı Tympanic Membrane veri setine göre daha yüksektir. Bu kullanılan otoskop cihazı veya kulak zarının görünürlüğüyle ilgili olabilmektedir. Ayrıca iki veri setinin birleştirilmesi ile oluşturulan Merged_ds veri setinde otitis eksterna hastalığına ait görüntü sayısının az olmasından dolayı genelde modeller bu hastalığı sınıflandıramamaktadır. YOLOv5 kullanımı modellerin kulak zarı bölgesine odaklanmasını sağlamasına karşın görüntü boyundaki küçülmeden dolayı etkisi düşük olmaktadır.

Modellerin veri setleri üzerindeki test işlemleri sonucunda elde edilen performans metrikleri Tablo 3.2-3.7’de verilmektedir.

Tablo 3.2. Ear imagery veri seti ağırlık değerleri performans metrikleri.

Modeller	Ağırlık Değerleri	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	Sıfırdan Eğitim	88	88	98	88	89
ResNet50	ImageNet	99	99	99	99	99
InceptionV3	Sıfırdan Eğitim	93	93	97	93	93
InceptionV3	ImageNet	98	98	99	98	98
Xception	Sıfırdan Eğitim	92	92	97	92	92
Xception	ImageNet	98	98	99	98	98
InceptionResNetV2	Sıfırdan Eğitim	94	94	98	94	94
InceptionResNetV2	ImageNet	97	97	98	97	97
Vision Transformer	Sıfırdan Eğitim	95	95	98	95	95

Tablo 3.3. Merged_ds veri seti ağırlık değerleri performans metrikleri.

Modeller	Ağırlık Değerleri	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	Sıfırdan Eğitim	77	61	96	63	77
ResNet50	ImageNet	88	78	98	82	88
InceptionV3	Sıfırdan Eğitim	84	72	97	74	84
InceptionV3	ImageNet	85	69	97	71	85
Xception	Sıfırdan Eğitim	85	70	97	73	85
Xception	ImageNet	86	73	97	76	86
InceptionResNetV2	Sıfırdan Eğitim	86	72	97	75	86
InceptionResNetV2	ImageNet	88	76	98	80	88
Vision Transformer	Sıfırdan Eğitim	85	71	97	76	85

Tablo 3.4. Merged_ds (6 sınıf) veri seti ağırlık değerleri performans metrikleri.

Modeller	Ağırlık Değerleri	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	Sıfırdan Eğitim	81	66	95	69	81
ResNet50	ImageNet	81	72	95	73	81
InceptionV3	Sıfırdan Eğitim	82	68	95	69	82
InceptionV3	ImageNet	82	66	96	67	82
Xception	Sıfırdan Eğitim	81	68	95	70	81
Xception	ImageNet	87	71	96	73	87
InceptionResNetV2	Sıfırdan Eğitim	83	66	95	68	83
InceptionResNetV2	ImageNet	85	71	96	72	85
Vision Transformer	Sıfırdan Eğitim	81	70	95	73	81

Tablo 3.5. Merged_ds (2 sınıf) veri seti ağırlık değerleri performans metrikleri.

Modeller	Ağırlık Değerleri	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	Sıfırdan Eğitim	61	55	93	50	61
ResNet50	ImageNet	91	92	98	91	91
InceptionV3	Sıfırdan Eğitim	87	87	95	87	87
InceptionV3	ImageNet	89	90	97	89	89
Xception	Sıfırdan Eğitim	83	82	95	82	83
Xception	ImageNet	89	89	96	89	89
InceptionResNetV2	Sıfırdan Eğitim	86	86	95	86	86
InceptionResNetV2	ImageNet	88	88	95	87	88
Vision Transformer	Sıfırdan Eğitim	84	84	95	84	84

Tablo 3.6. Tympanic membrane veri seti ağırlık değerleri performans metrikleri.

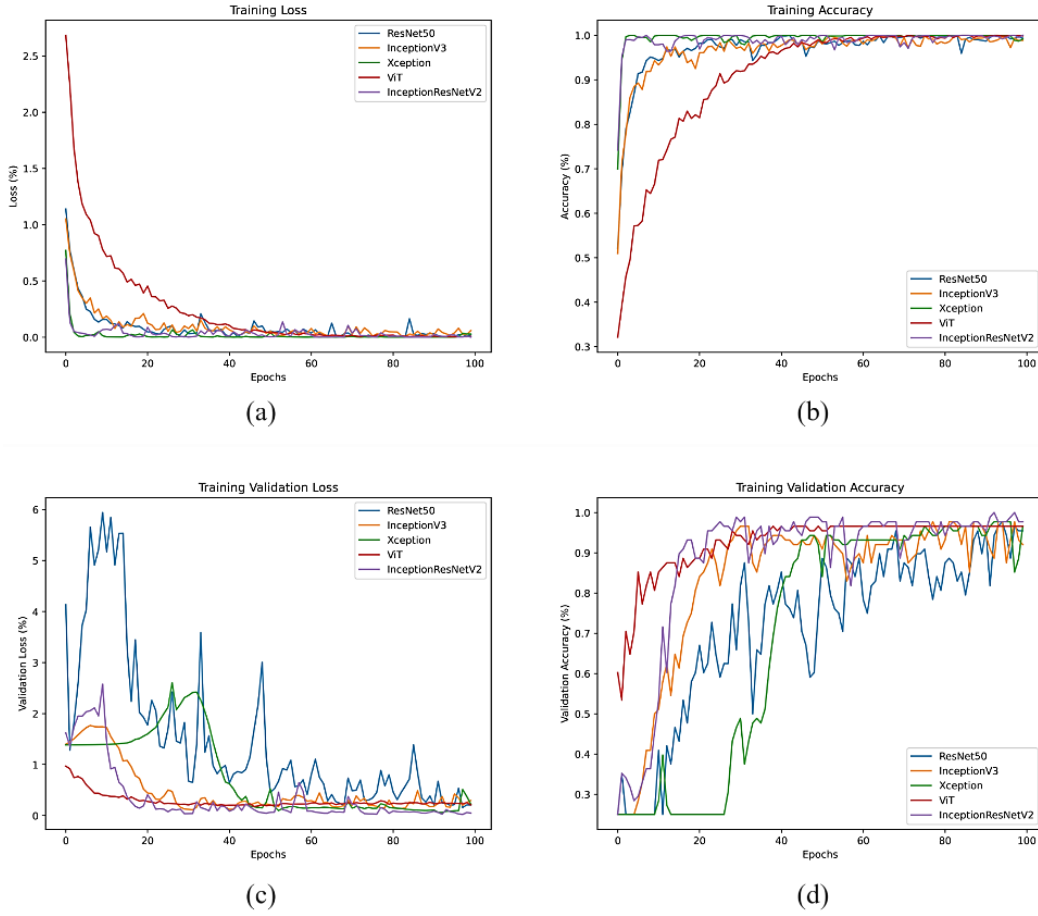
Modeller	Ağırlık Değerleri	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	Sıfırdan Eğitim	69	50	89	52	69
ResNet50	ImageNet	77	56	90	61	77
InceptionV3	Sıfırdan Eğitim	74	51	89	55	74
InceptionV3	ImageNet	72	47	88	51	72
Xception	Sıfırdan Eğitim	73	43	87	49	73
Xception	ImageNet	78	56	90	61	78
InceptionResNetV2	Sıfırdan Eğitim	75	49	88	56	75
InceptionResNetV2	ImageNet	76	58	89	63	76
Vision Transformer	Sıfırdan Eğitim	71	47	87	55	71

Tablo 3.7. YOLOv5 uygulanmış tympanic membrane veri seti ağırlık değerleri performans metrikleri.

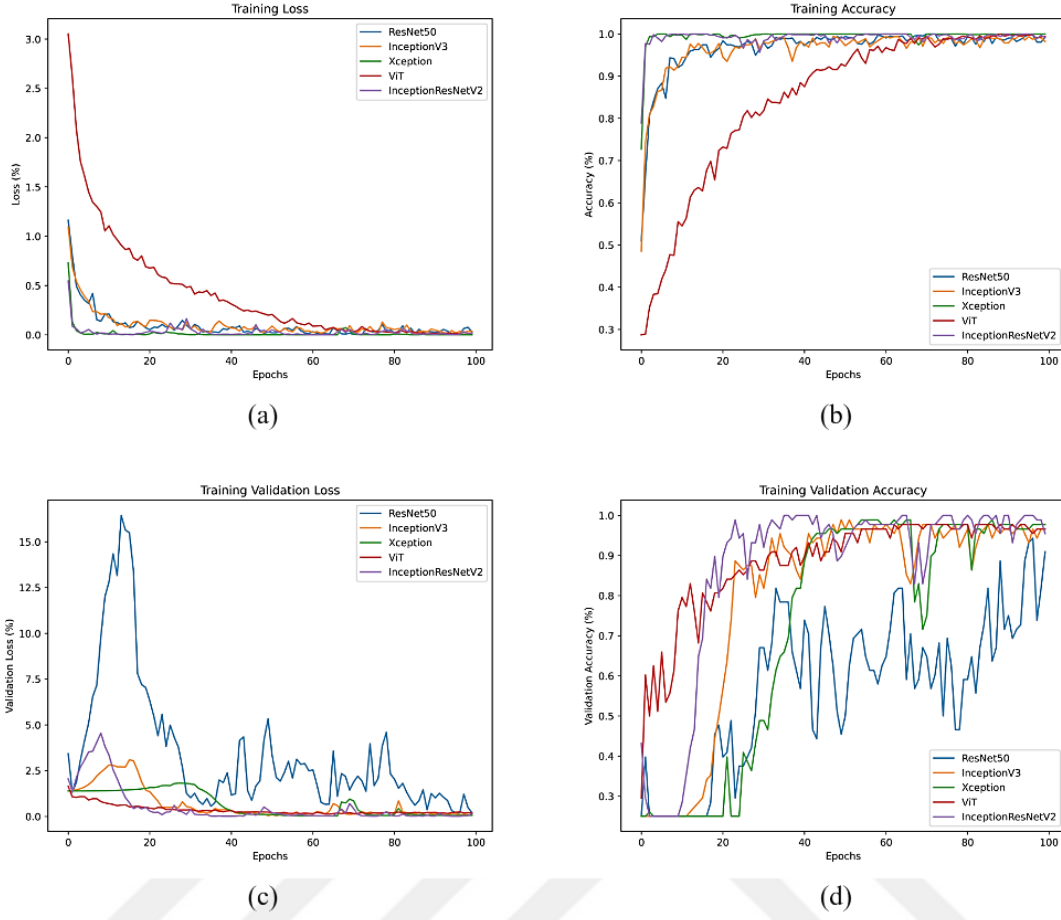
Modeller	Ağırlık Değerleri	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	Sıfırdan Eğitim	75	53	88	61	75
ResNet50	ImageNet	74	49	89	53	74
InceptionV3	Sıfırdan Eğitim	65	43	86	45	65
InceptionV3	ImageNet	74	51	89	56	74
Xception	Sıfırdan Eğitim	75	52	89	58	75
Xception	ImageNet	80	60	92	65	80
InceptionResNetV2	Sıfırdan Eğitim	73	47	88	52	73
InceptionResNetV2	ImageNet	75	52	89	58	75
Vision Transformer	Sıfırdan Eğitim	77	56	90	63	77

Tablo 3.1-3.6'dan elde edilen sonuçlar dikkate alınarak her veri seti için en yüksek başarıya sahip modeller sırasıyla, ResNet50 (Ear Imagery), öğrenme aktarımlı ResNet50 ve InceptionResNetV2 (Merged_ds), öğrenme aktarımlı Xception (Merged_ds 6 sınıf), öğrenme aktarımlı ResNet50 (Merged_ds 2 sınıf), öğrenme aktarımlı Xception (Tympanic Membrane) ve öğrenme aktarımlı Xception (YOLOv5 uygulanmış Tympanic Membrane) modelleridir. Genel olarak ResNet50 ve Xception modelleri kulak hastalıklarını sınıflandırmada başarılıdır. Modellerin sıfırdan eğitime göre ImageNet veri setinde eğitimi sonucunda elde edilen ağırlık değerlerinin kullanımı modellerin başarılarını artırmaktadır.

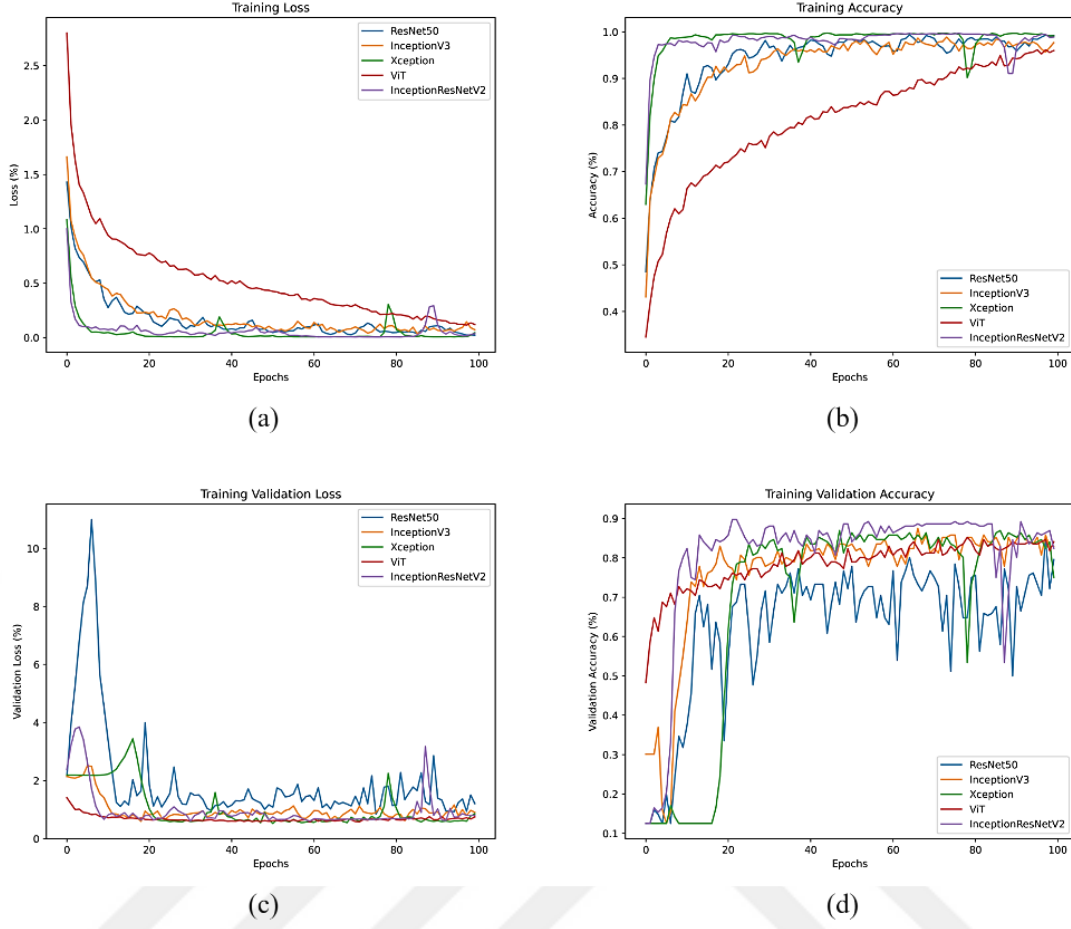
Ayrıca çalışma kapsamında farklı renk uzaylarının modellerinin sınıflandırma doğruluğu başarısına etkisi araştırılmıştır. Bu kapsamda veri setlerindeki görüntüler HSV ve LUV renk uzaylarına çevrilmiştir. Şekil 3.29-3.40'da modellerin eğitimlerine ait grafikler verilmektedir. Öğrenme aktarımında modeller RGB renk uzayındaki görüntüler üzerinde eğitildiğinden dolayı farklı bir renk uzayı kullanımı modellerin performansını düşürmektedir. Bu yüzden modeller sıfırdan eğitilmiştir.



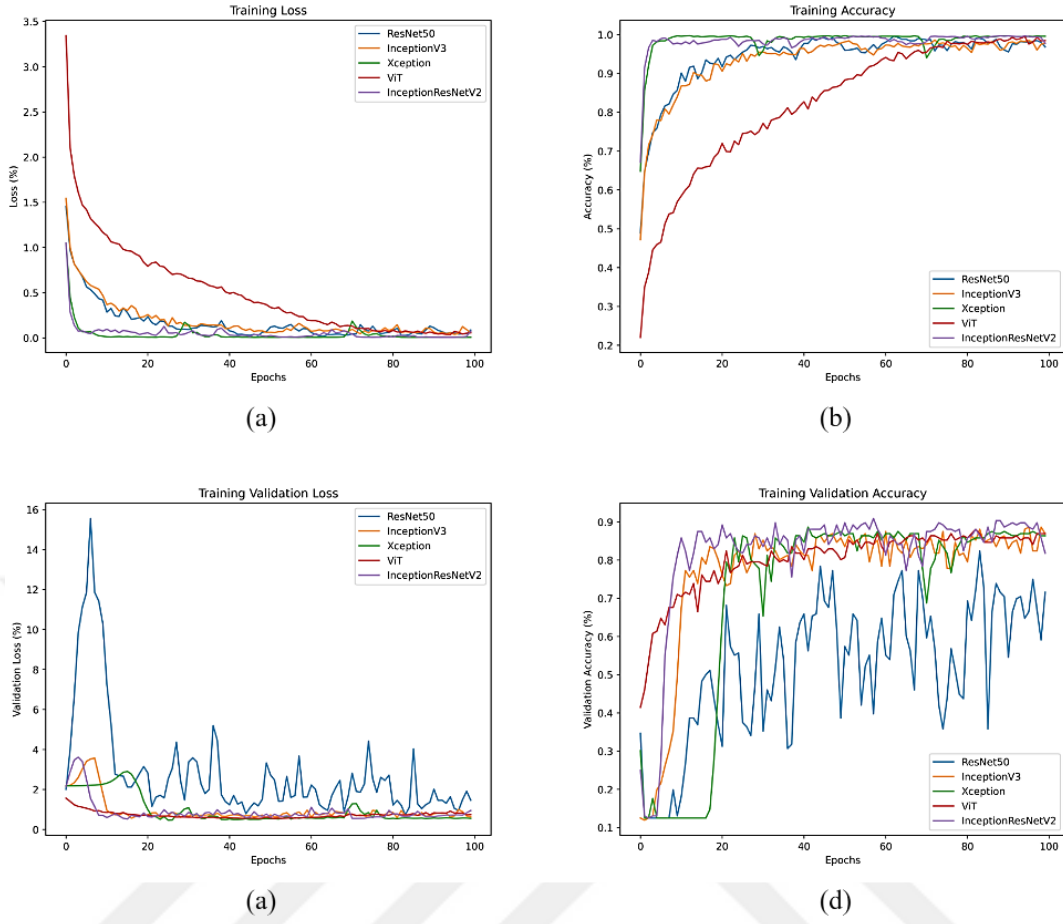
Şekil 3.29. Ear imagery veri setindeki görüntülerin HSV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).



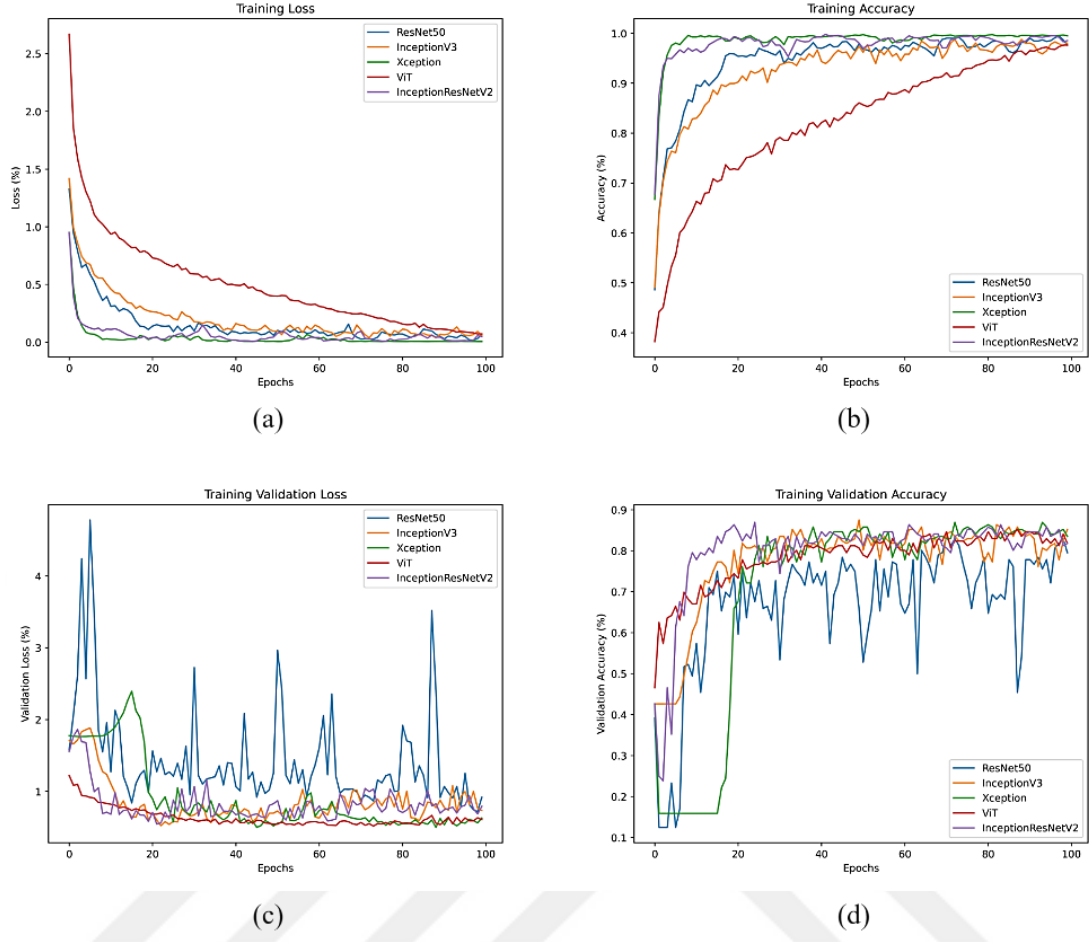
Şekil 3.30. Ear imagery veri setindeki görüntülerin LUV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).



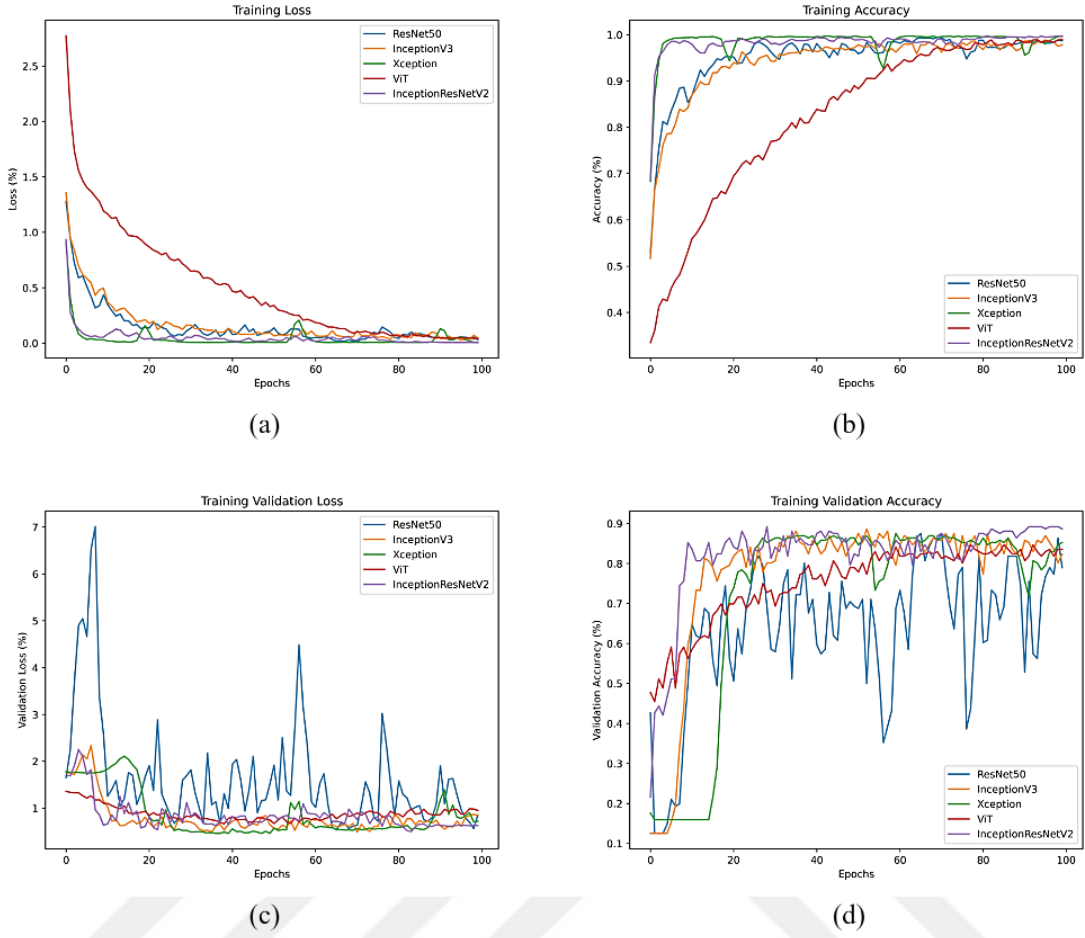
Şekil 3.31. Merged_ds veri setindeki görüntülerin HSV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).



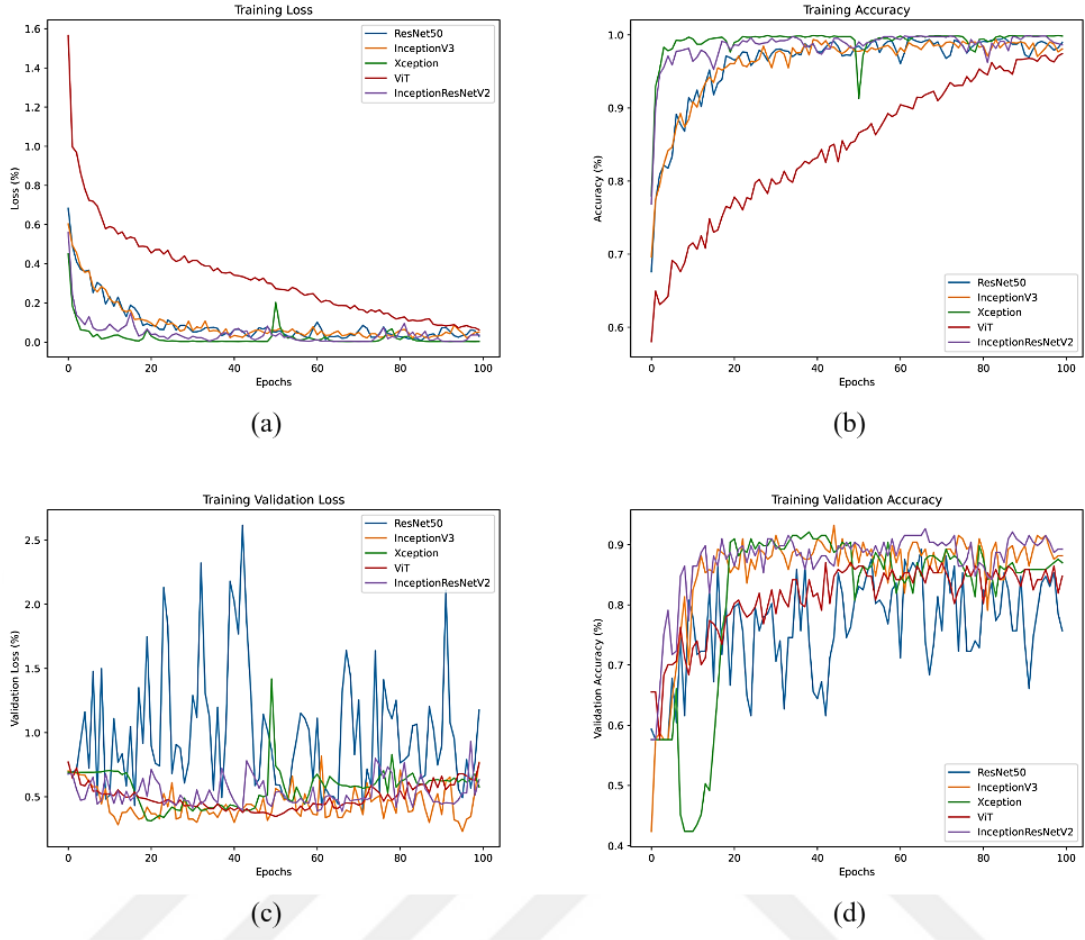
Şekil 3.32. Merged_ds veri setindeki görüntülerin LUV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).



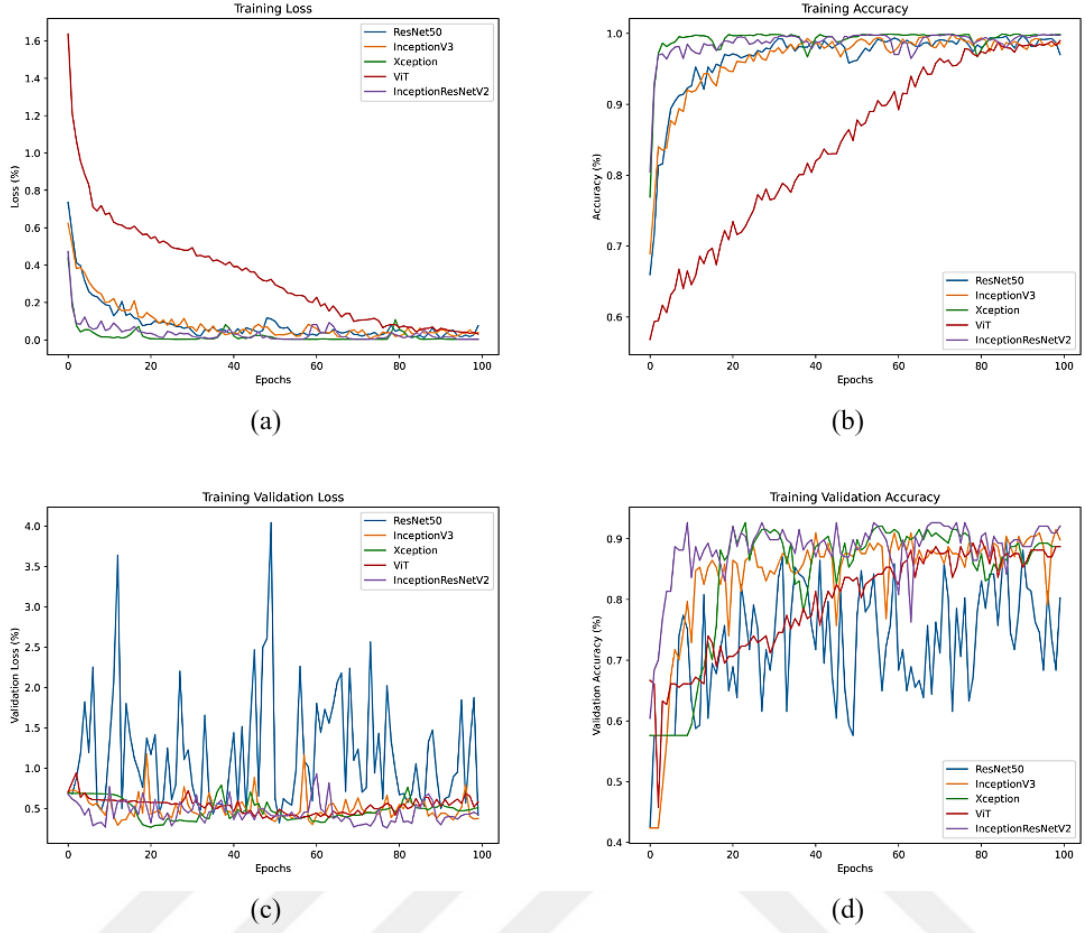
Şekil 3.33. Merged_ds (6 sınıf) veri setindeki görüntülerin HSV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).



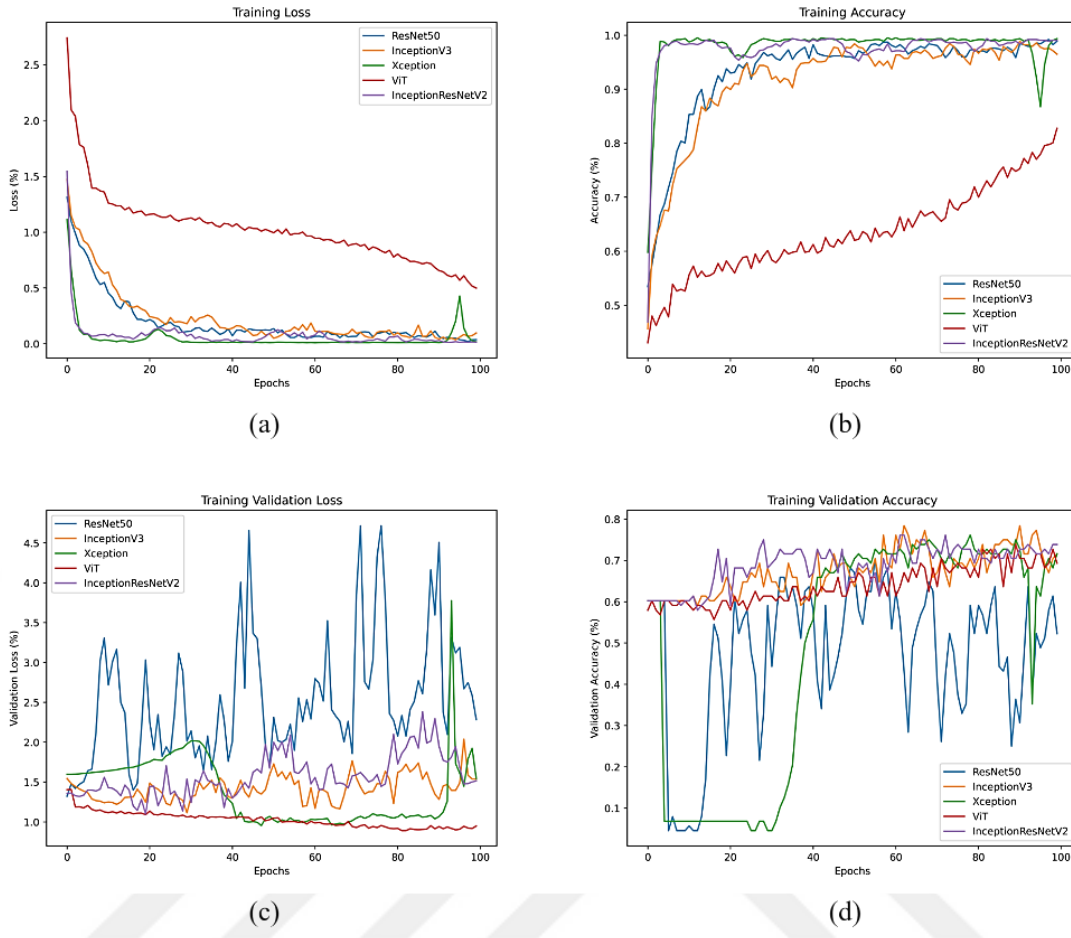
Şekil 3.34. Merged_ds (6 sınıf) veri setindeki görüntülerin LUV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).



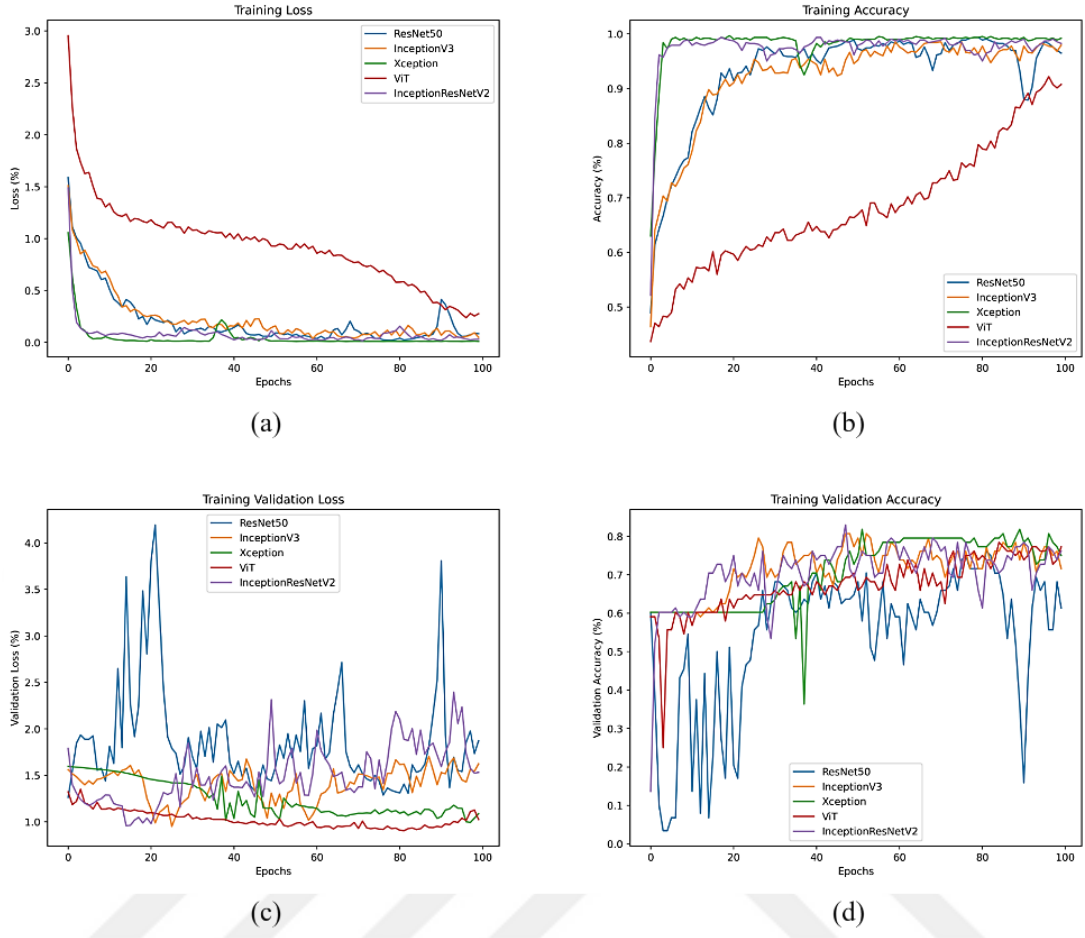
Şekil 3.35. Merged_ds (2 sınıf) veri setindeki görüntülerin HSV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).



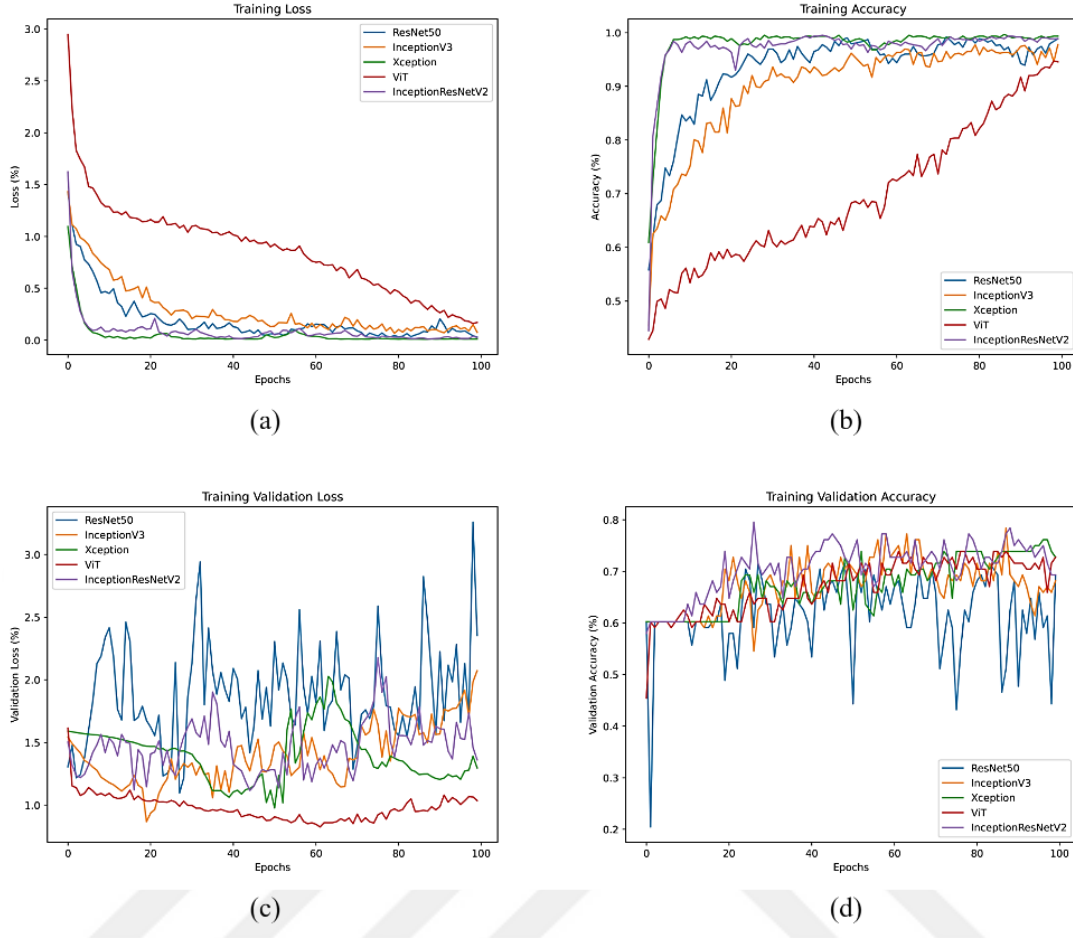
Şekil 3.36. Merged_ds (2 sınıf) veri setindeki görüntülerin LUV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).



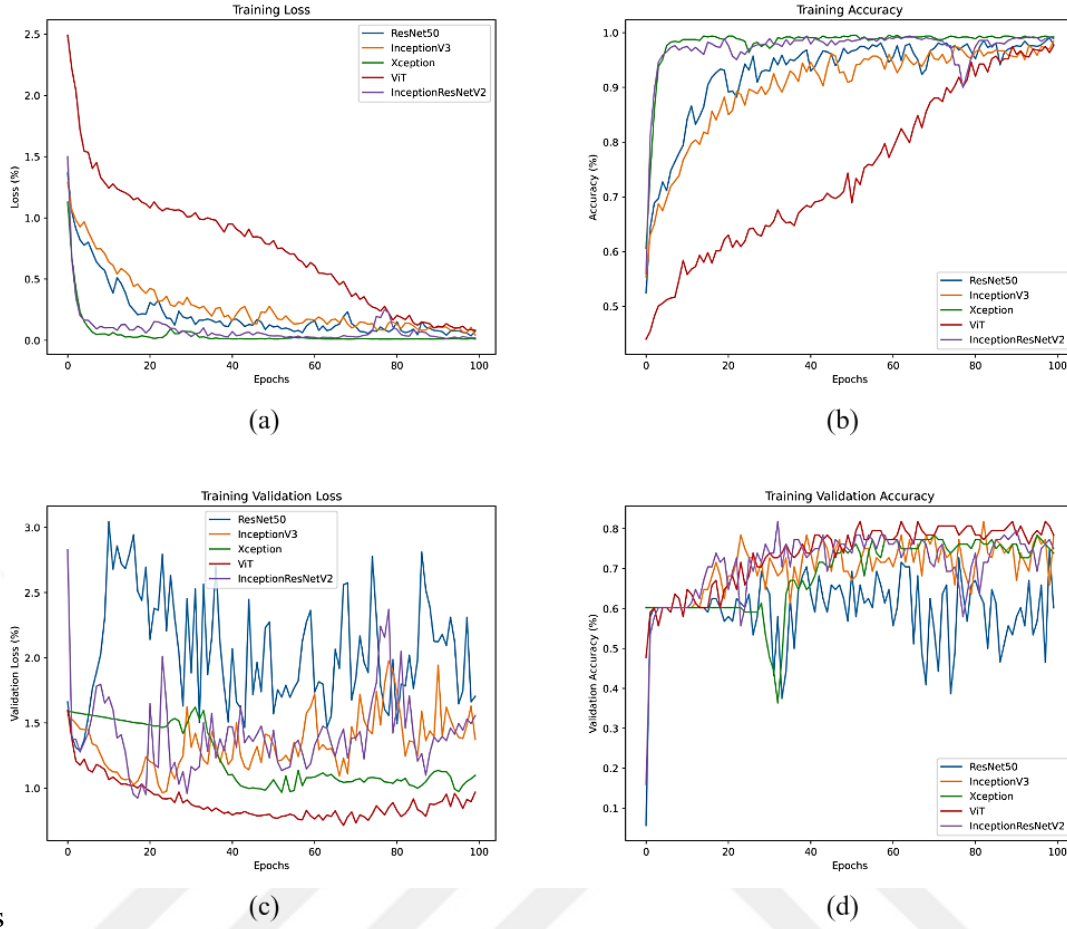
Şekil 3.37. Tympanic membrane veri setindeki görüntülerin HSV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).



Şekil 3.38. Tympanic membrane veri setindeki görüntülerin LUV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).

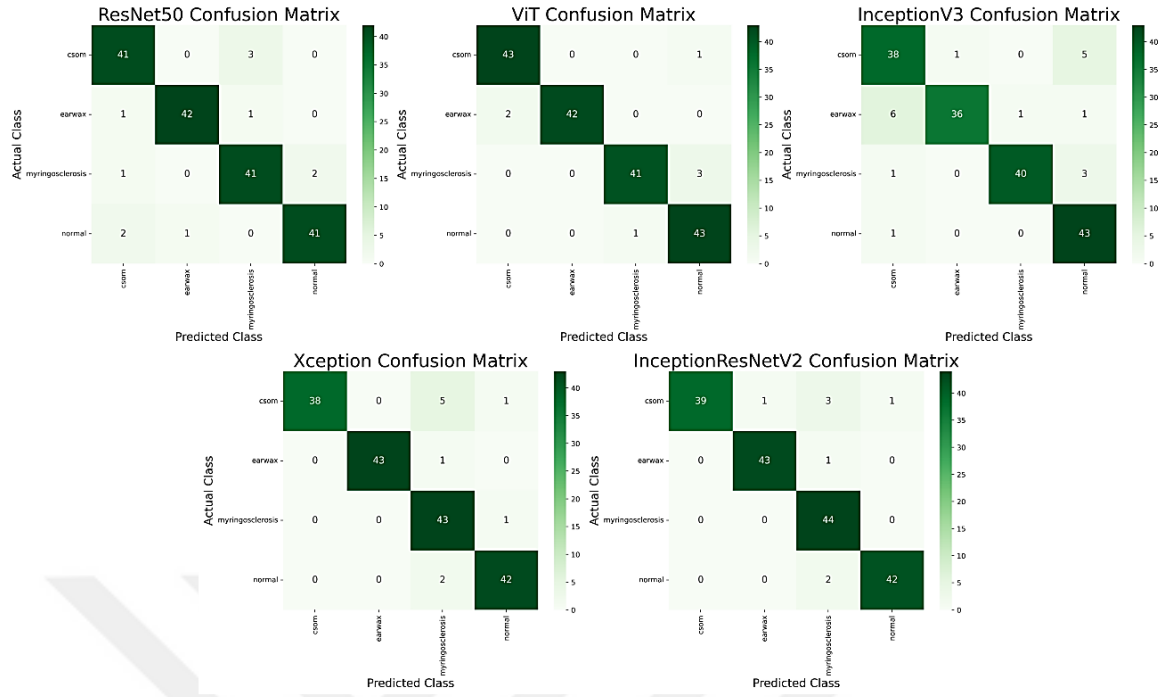


Şekil 3.39. YOLOv5 uygulanmış tympanic membrane veri setindeki görüntülerin HSV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).

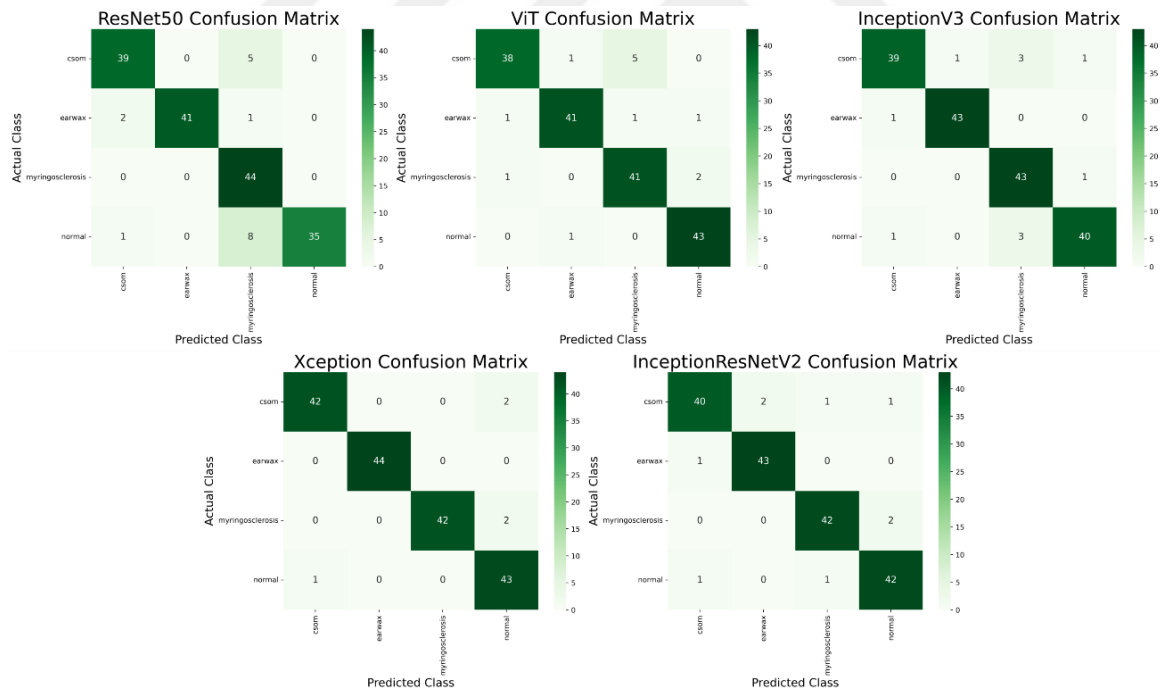


Şekil 3.40. YOLOv5 uygulanmış tympanic membrane veri setindeki görüntülerin LUV renk uzayı için eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c) ve doğrulama doğruluğu (d).

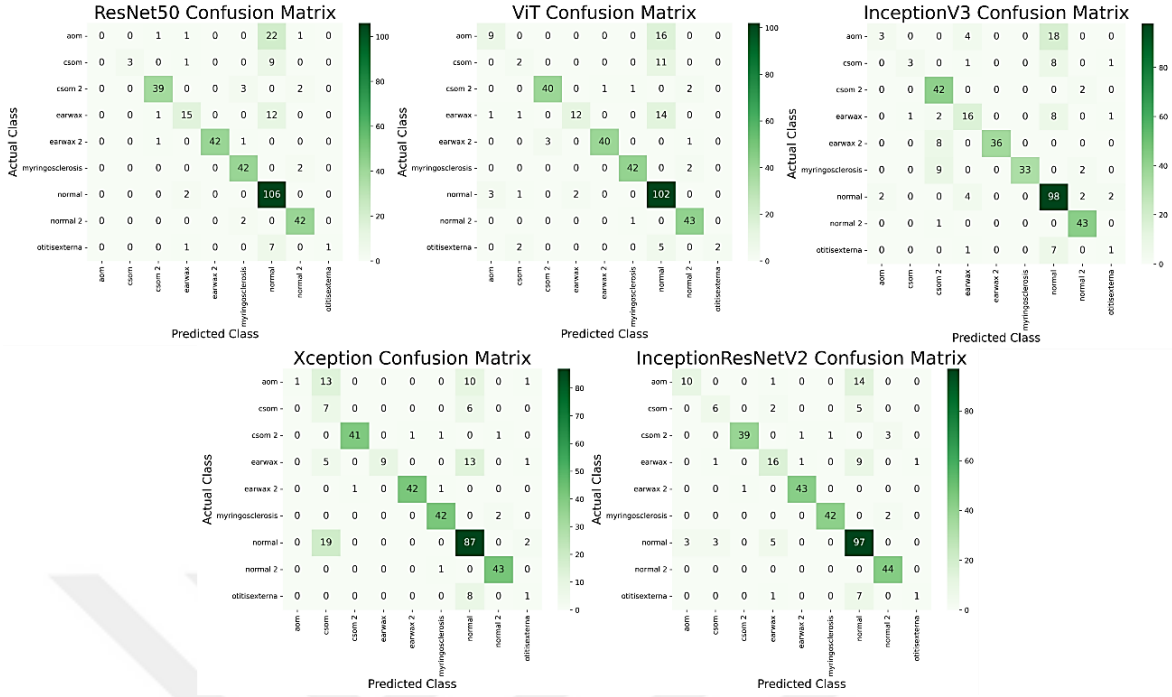
Şekil 3.29-3.40'da görüldüğü üzere farklı renk uzayı kullanımı modellerin eğitim performanslarını etkilemektedir. RGB renk uzayında kullanılan öğrenme aktarımının farklı renk uzayında kullanılamaması başarıyı düşürmektedir. Veri sayısının az olması eğitim süresinin artmasına neden olmaktadır. Ayrıca bazı modeller doğrulama işleminde farklı renk uzayında RGB renk uzayına göre daha kötü performans göstermektedir. Test veri seti üzerindeki modellere ait karışıklık matrisleri Şekil 3.41-3.52'de verilmektedir.



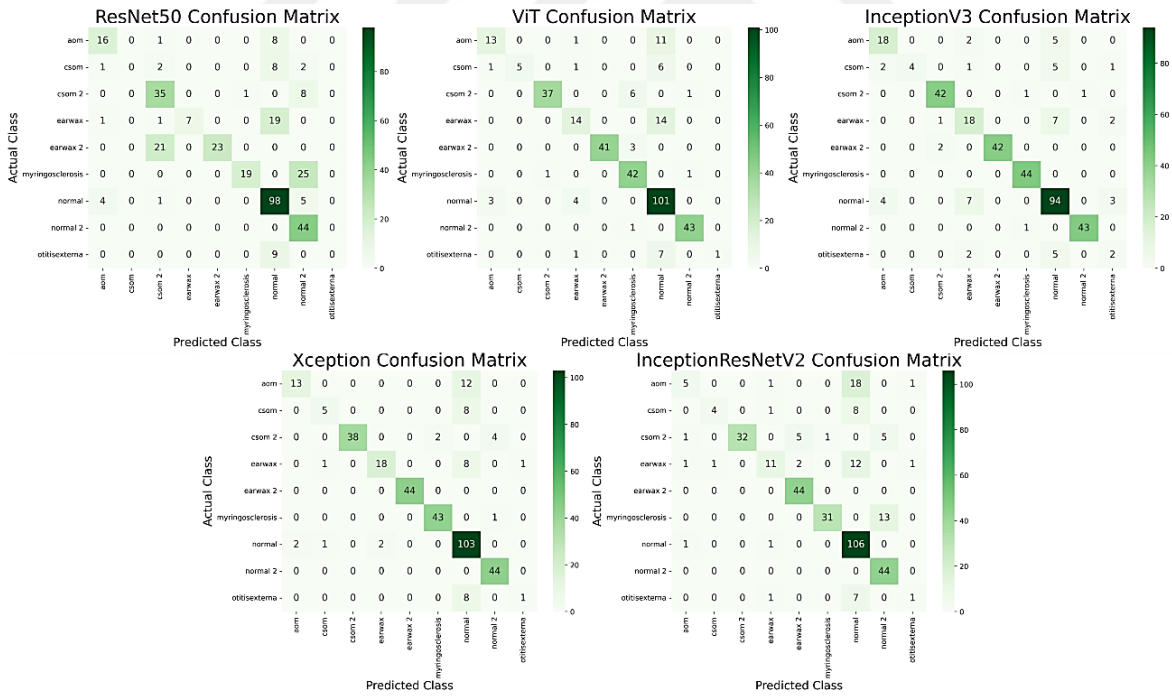
Şekil 3.41. Ear imagery veri seti HSV renk uzayı karışıklık matrisleri.



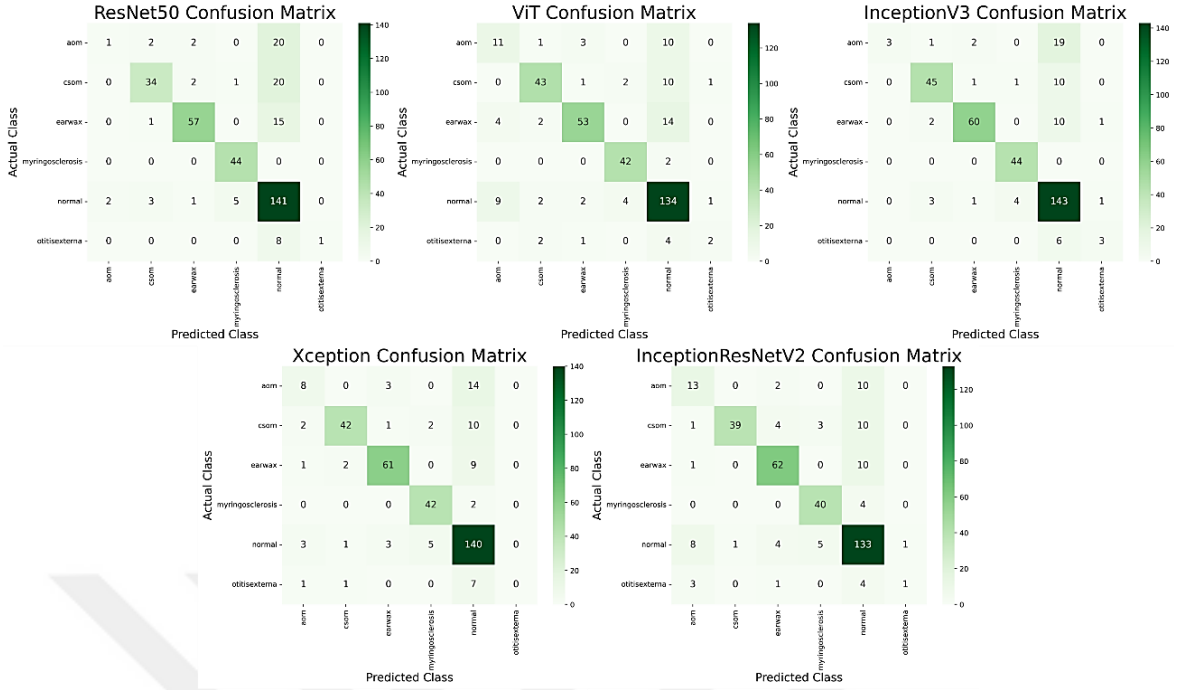
Şekil 3.42. Ear imagery veri seti LUV renk uzayı karışıklık matrisleri.



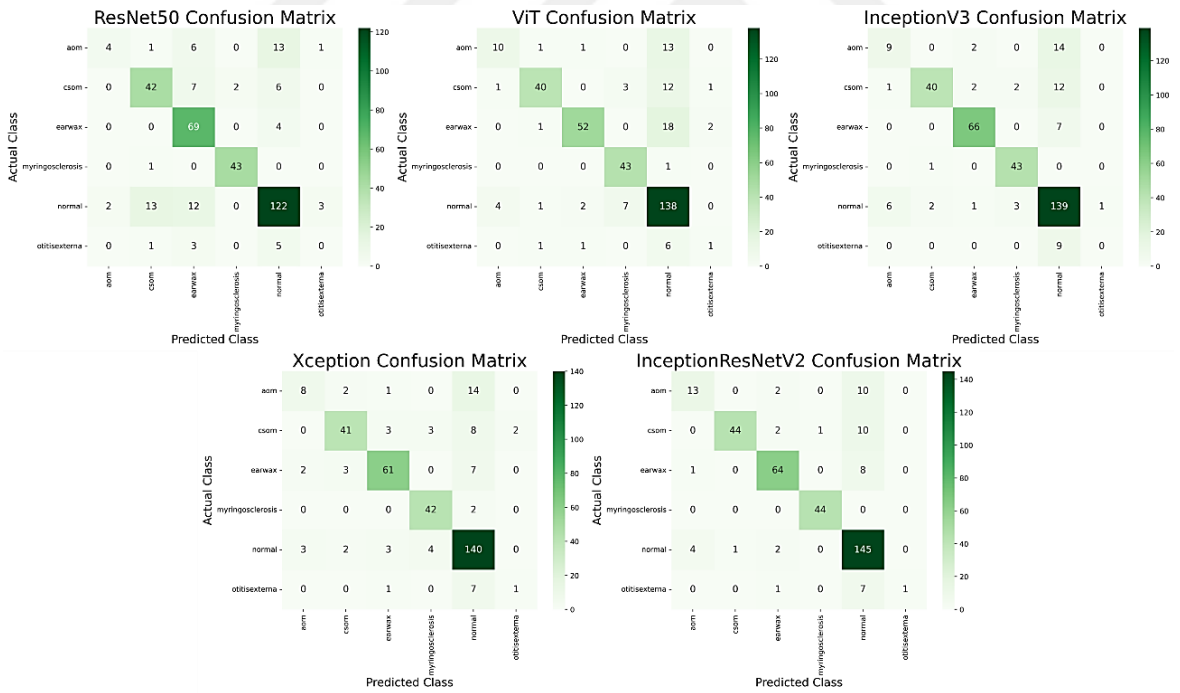
Şekil 3.43. Merged_ds veri seti HSV renk uzayı karışıklık matrisleri.



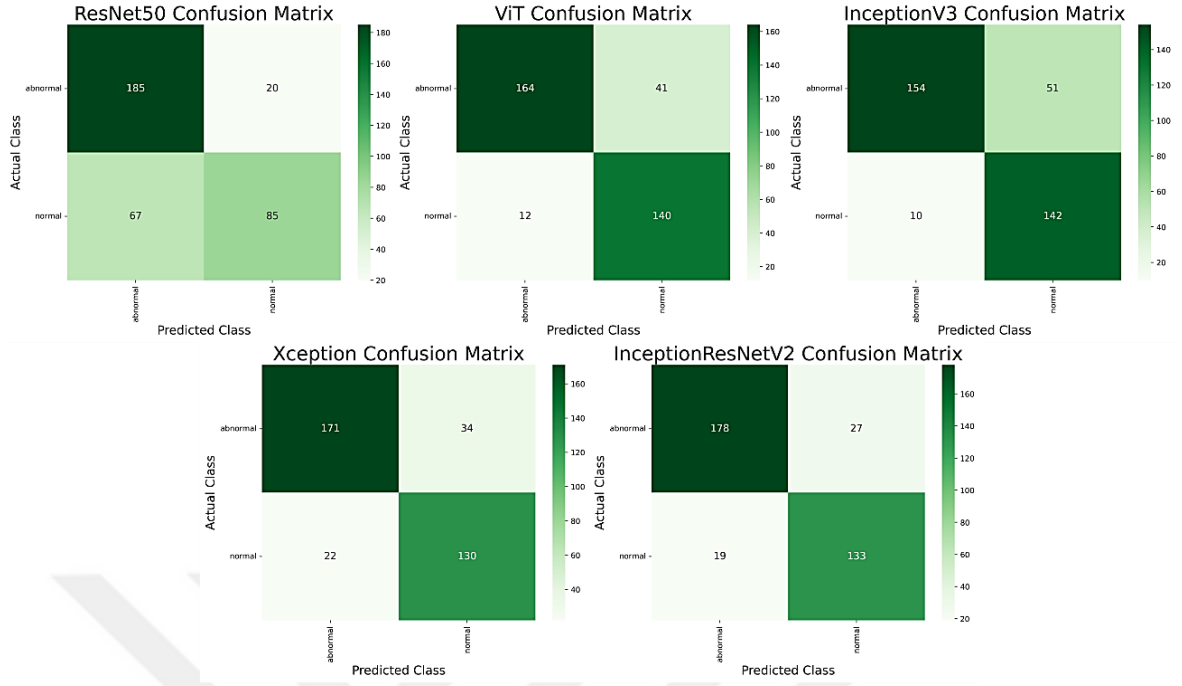
Şekil 3.44. Merged_ds veri seti LUV renk uzayı karışıklık matrisleri.



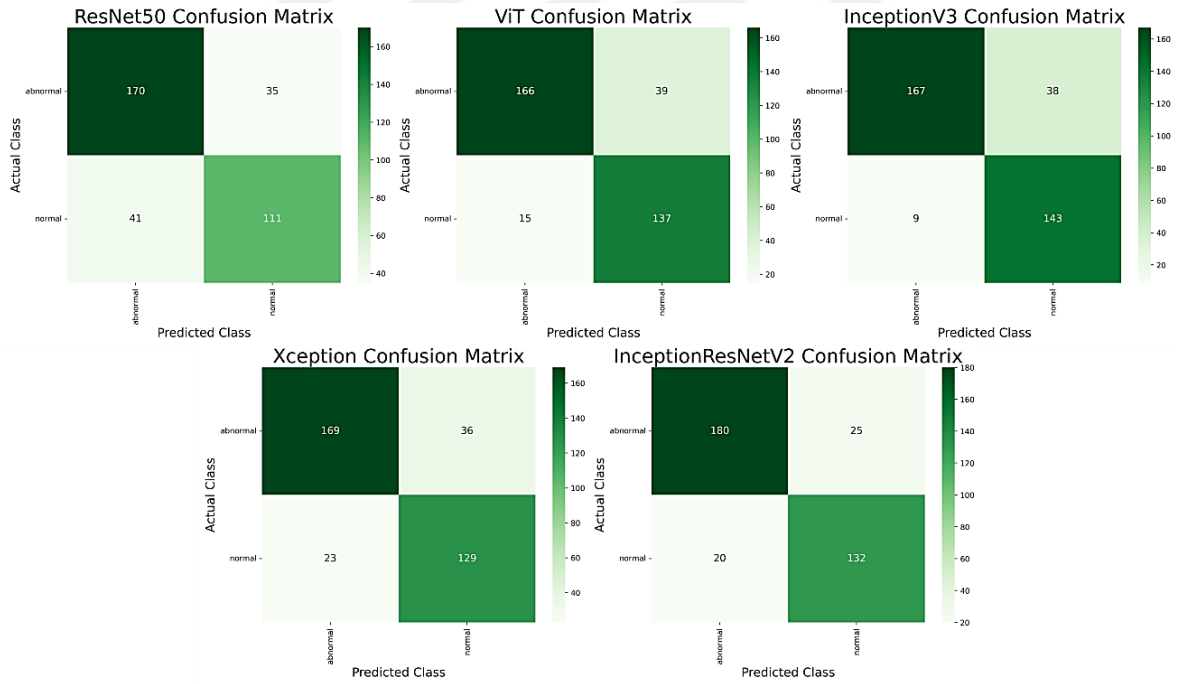
Şekil 3.45. Merged_ds (6 sınıf) veri seti HSV renk uzayı karışıklık matrisleri.



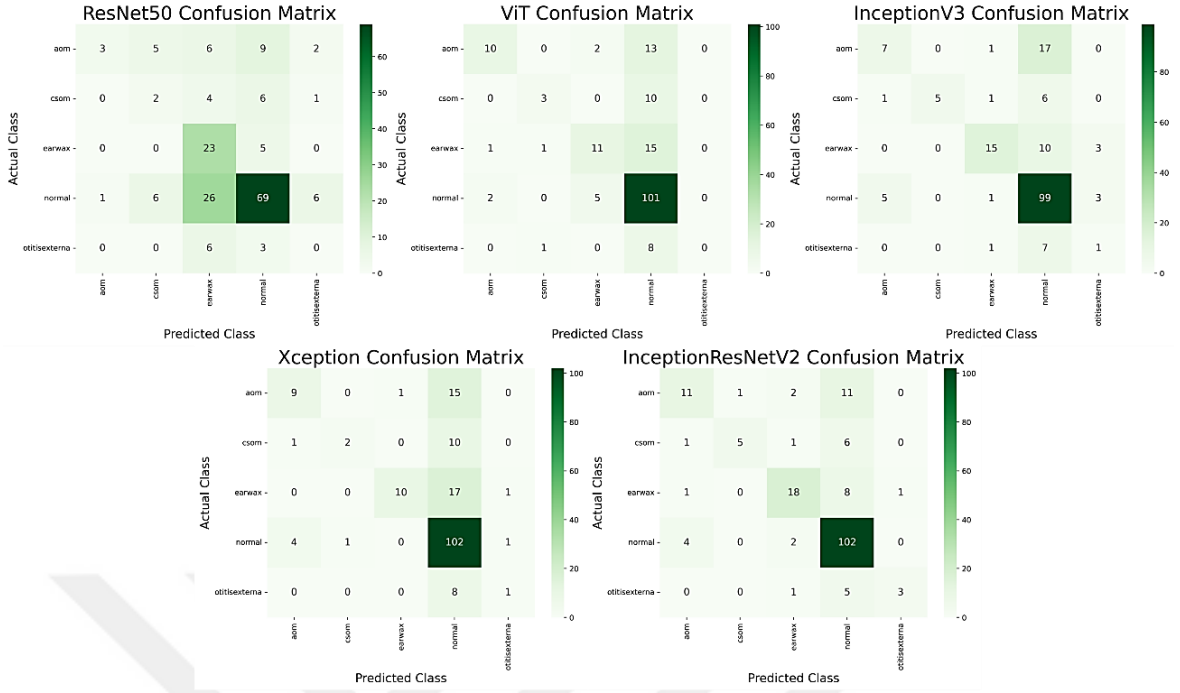
Şekil 3.46. Merged_ds (6 sınıf) veri seti LUV renk uzayı karışıklık matrisleri.



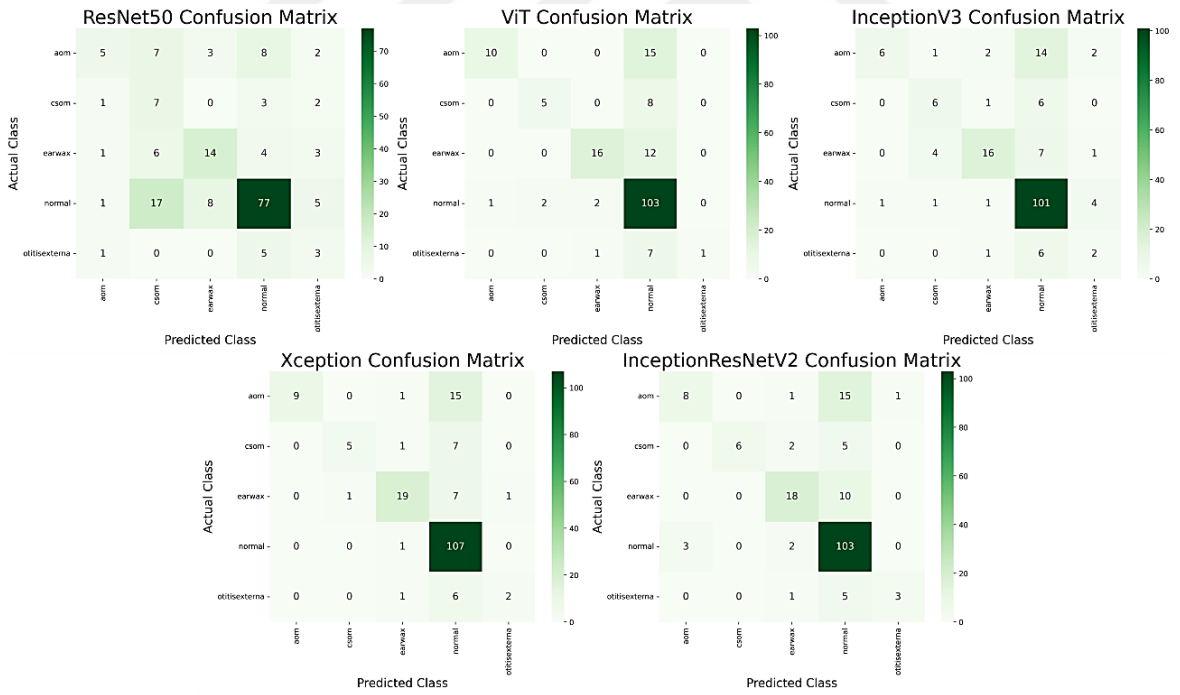
Şekil 3.47. Merged_ds (2 sınıf) veri seti HSV renk uzayı karışıklık matrisleri.



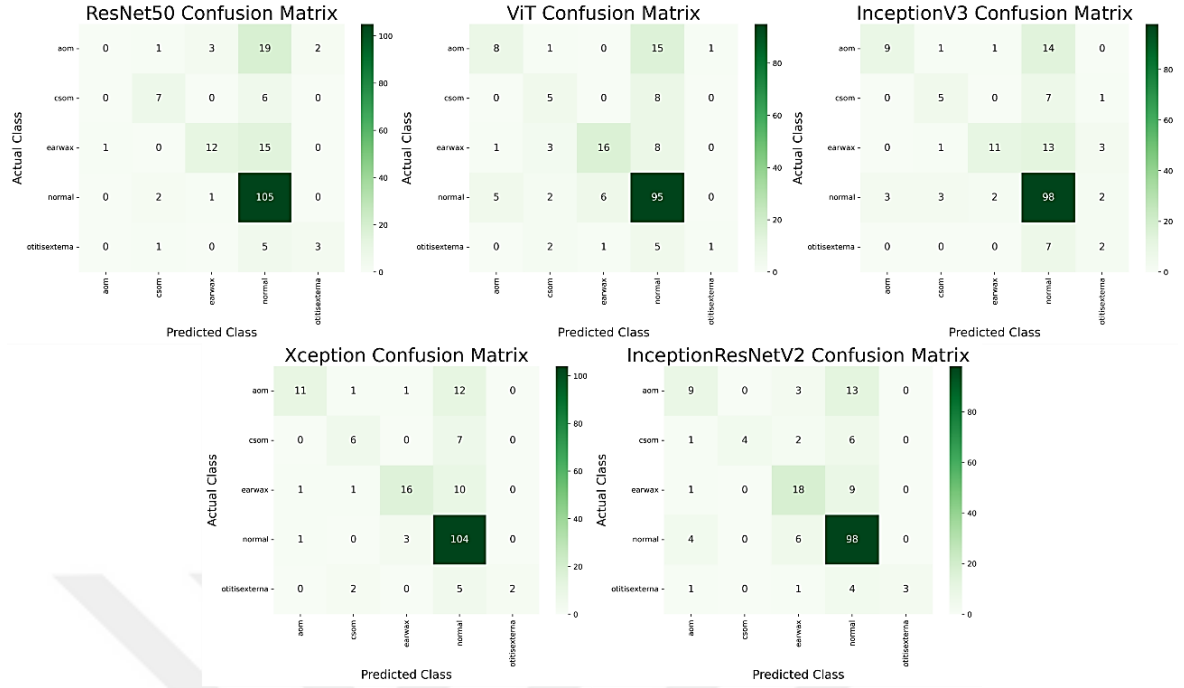
Şekil 3.48. Merged_ds (2 sınıf) veri seti LUV renk uzayı karışıklık matrisleri.



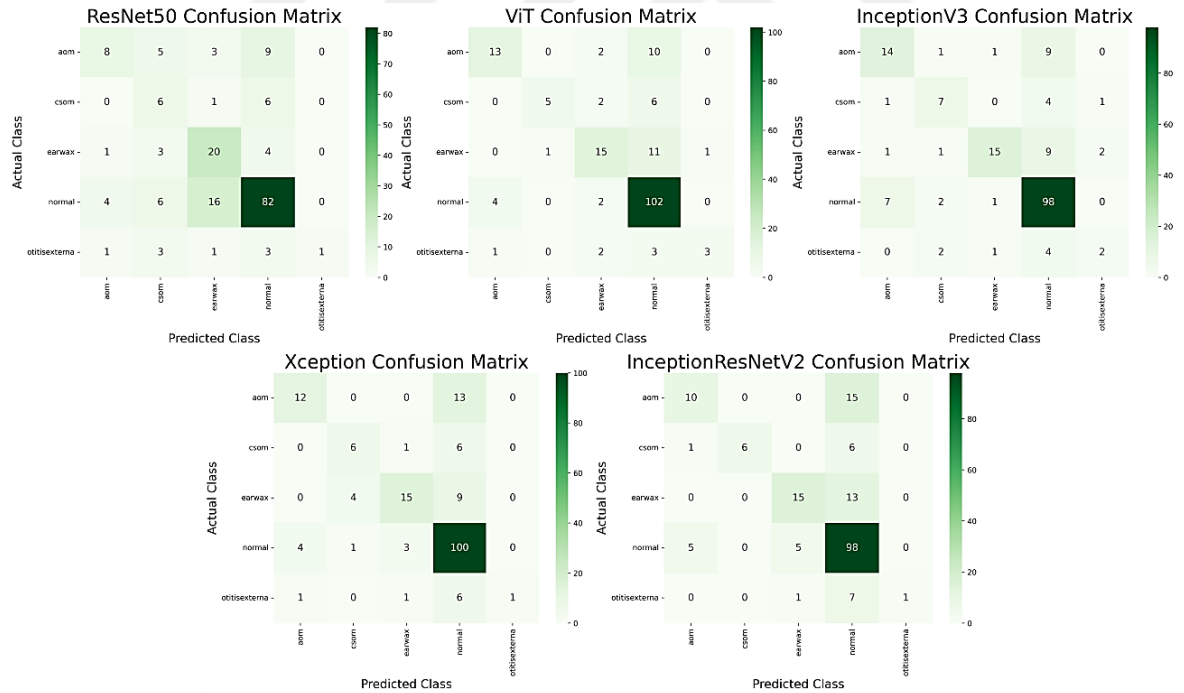
Şekil 3.49. Tympanic membrane veri seti ISV renk uzayı karışıklık matrisleri.



Şekil 3.50. Tympanic membrane veri seti LUV renk uzayı karışıklık matrisleri.



Şekil 3.51. YOLOv5 uygulanmış tympanic membrane veri seti HSV renk uzayı karışıklık matrisleri.



Şekil 3.52. YOLOv5 uygulanmış tympanic membrane veri seti LUV renk uzayı karışıklık matrisleri.

Şekil 3.41-3.52’den elde edilen sonuçlara dayanarak bazı modellere ait performans değerleri RGB renk uzayına göre değişmemiştir. Öte yandan modellerden bazılarının

performans değerleri düşmüştür. Genel olarak yüksek veri sayısına sahip sağlıklı kulak görüntüleri modeller tarafından başarılı şekilde sınıflandırılmıştır. Bu, veri sayısının artırılmasının modellerin performansını olumlu etkileyeceğini göstermektedir. Tablo 3.8-3.13’de modellere ait performans metrikleri verilmektedir.

Tablo 3.8. Ear imagery veri seti renk uzayları performans metrikleri.

Modeller	Renk Uzaı	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	HSV	94	94	97	94	94
ResNet50	LUV	90	90	96	91	89
InceptionV3	HSV	89	89	96	89	89
InceptionV3	LUV	94	94	97	94	93
Xception	HSV	94	94	98	94	94
Xception	LUV	97	97	99	97	97
InceptionResNetV2	HSV	95	95	98	95	95
InceptionResNetV2	LUV	95	95	98	95	95
Vision Transformer	HSV	96	96	98	96	96
Vision Transformer	LUV	93	93	97	93	92

Tablo 3.9. Merged_ds veri seti renk uzayları performans metrikleri.

Modeller	Renk Uzaı	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	HSV	81	62	97	64	80
ResNet50	LUV	67	51	95	50	66
InceptionV3	HSV	77	60	96	62	76
InceptionV3	LUV	86	74	98	75	85
Xception	HSV	76	63	96	61	76
Xception	LUV	86	72	97	75	85
InceptionResNetV2	HSV	83	70	97	72	82
InceptionResNetV2	LUV	77	60	96	63	76
Vision Transformer	HSV	81	65	97	69	80
Vision Transformer	LUV	83	68	97	72	82

Tablo 3.10. Merged_ds (6 sınıf) veri seti renk uzayları performans metrikleri.

Modeller	Renk Uzayı	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	HSV	77	58	93	59	77
ResNet50	LUV	78	60	94	60	76
InceptionV3	HSV	83	67	95	69	81
InceptionV3	LUV	82	64	95	65	81
Xception	HSV	81	63	95	64	80
Xception	LUV	81	64	95	66	80
InceptionResNetV2	HSV	80	66	95	67	80
InceptionResNetV2	LUV	86	71	96	74	85
Vision Transformer	HSV	79	66	94	68	78
Vision Transformer	LUV	79	64	94	66	78

Tablo 3.11. Merged_ds (2 sınıf) veri seti renk uzayları performans metrikleri.

Modeller	Renk Uzayı	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	HSV	76	73	95	74	75
ResNet50	LUV	79	78	94	78	78
InceptionV3	HSV	83	84	96	83	82
InceptionV3	LUV	87	88	98	87	86
Xception	HSV	84	84	96	84	83
Xception	LUV	83	84	96	83	82
InceptionResNetV2	HSV	87	87	98	87	87
InceptionResNetV2	LUV	87	87	98	87	87
Vision Transformer	HSV	85	86	97	85	85
Vision Transformer	LUV	85	86	96	85	84

Tablo 3.12. Tympanic membrane veri seti renk uzayları performans metrikleri.

Modeller	Renk Uzaı	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	HSV	53	35	85	31	50
ResNet50	LUV	58	46	87	42	58
InceptionV3	HSV	69	45	87	50	68
InceptionV3	LUV	72	49	88	51	71
Xception	HSV	68	39	85	44	67
Xception	LUV	78	53	89	60	78
InceptionResNetV2	HSV	76	55	90	61	73
InceptionResNetV2	LUV	75	54	89	61	75
Vision Transformer	HSV	68	39	86	43	67
Vision Transformer	LUV	74	48	88	55	73

Tablo 3.13. YOLOv5 uygulanmış tympanic membrane veri seti renk uzayları performans metrikleri.

Modeller	Renk Uzaı	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	HSV	69	44	86	46	67
ResNet50	LUV	64	47	88	46	63
InceptionV3	HSV	68	45	87	49	67
InceptionV3	LUV	74	55	90	58	73
Xception	HSV	76	53	89	60	75
Xception	LUV	73	50	89	54	73
InceptionResNetV2	HSV	72	51	89	57	70
InceptionResNetV2	LUV	71	48	87	55	70
Vision Transformer	HSV	68	45	87	48	67
Vision Transformer	LUV	75	54	90	61	74

Tablo 3.8-3.13'den elde edilen bulgular dikkate alınarak renk uzayının değiştirilmesi modellerin RGB renk uzayında çıkaramadıkları özellikleri yakalamasını sağlamaktadır. Bu, modellerin veri setleri üzerindeki performansını olumlu yönde etkilemiştir. Renk uzayının değiştirilmesinin dezavantajı öğrenme aktarımının kullanılamamasından dolayı modellerin sıfırdan eğitimini gerektirmektedir. Medikal uygulamalarda veri bulmak zor olduğundan bu işlem için artırma kullanılarak veri sayısı artırılmıştır.

Çalışma kapsamında verilerin artırılması için eğitim setinde bulunan görüntülere artırma işlemi uygulanmıştır. Modellerin eğitim işlemi artırılmış veriler üzerinde gerçekleştirilerek performansları artırılmaya çalışılmıştır. Modeller artırılmış veri setinde hem sıfırdan hem de öğrenme aktarımı kullanılarak eğitilmiştir. Ayrıca renk uzayı dönüşümü sonrasında elde edilen HSV ve LUV görüntülerde de artırma işlemi uygulanmıştır. Tablo 3.14-3.25'de artırılmış eğitim veri setinde eğitilmiş modellerin performans metrikleri verilmektedir.

Tablo 3.14. Artırma uygulanmış ear imagery veri seti ağırlık değerleri performans metrikleri.

Modeller	Ağırlık Değerleri	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	Sıfırdan Eğitim	99	99	99	99	99
ResNet50	ImageNet	99	99	99	99	99
InceptionV3	Sıfırdan Eğitim	99	99	99	99	99
InceptionV3	ImageNet	99	99	99	99	99
Xception	Sıfırdan Eğitim	93	93	97	93	93
Xception	ImageNet	99	99	99	99	99
InceptionResNetV2	Sıfırdan Eğitim	98	98	99	98	98
InceptionResNetV2	ImageNet	99	99	99	99	99
Vision Transformer	Sıfırdan Eğitim	95	95	98	95	94

Tablo 3.15. Artırma uygulanmış ear imagery veri seti renk uzayları performans metrikleri.

Modeller	Renk Uzayı	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	HSV	93	93	97	93	93
ResNet50	LUV	78	78	92	77	77
InceptionV3	HSV	97	97	99	97	97
InceptionV3	LUV	95	95	98	95	95
Xception	HSV	92	92	97	92	92
Xception	LUV	84	84	94	83	84
InceptionResNetV2	HSV	98	98	99	98	98
InceptionResNetV2	LUV	97	97	98	97	97
Vision Transformer	HSV	95	95	98	95	95
Vision Transformer	LUV	94	94	97	94	94

Tablo 3.16. Artırma uygulanmış merged_ds veri seti ağırlık değerleri performans metrikleri.

Modeller	Ağırlık Değerleri	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	Sıfırdan Eğitim	82	64	97	66	81
ResNet50	ImageNet	87	72	98	76	87
InceptionV3	Sıfırdan Eğitim	84	71	97	72	83
InceptionV3	ImageNet	79	68	97	70	78
Xception	Sıfırdan Eğitim	81	65	97	67	80
Xception	ImageNet	87	75	98	78	87
InceptionResNetV2	Sıfırdan Eğitim	82	67	97	70	81
InceptionResNetV2	ImageNet	88	76	98	78	88
Vision Transformer	Sıfırdan Eğitim	83	68	97	70	83

Tablo 3.17. Artırma uygulanmış merged_ds veri seti renk uzayları performans metrikleri.

Modeller	Renk Uzayı	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	HSV	77	67	96	68	76
ResNet50	LUV	60	53	94	49	59
InceptionV3	HSV	79	67	97	68	78
InceptionV3	LUV	75	62	96	63	74
Xception	HSV	81	67	97	68	80
Xception	LUV	77	61	96	63	77
InceptionResNetV2	HSV	86	71	97	75	85
InceptionResNetV2	LUV	86	72	98	74	85
Vision Transformer	HSV	82	68	97	71	81
Vision Transformer	LUV	84	70	97	74	83

Tablo 3.18. Artırma uygulanmış merged_ds (6 sınıf) veri seti ağırlık değerleri performans metrikleri.

Modeller	Ağırlık Değerleri	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	Sıfırdan Eğitim	61	55	91	54	60
ResNet50	ImageNet	83	69	96	71	82
InceptionV3	Sıfırdan Eğitim	75	64	94	62	74
InceptionV3	ImageNet	88	72	86	74	87
Xception	Sıfırdan Eğitim	80	62	95	64	80
Xception	ImageNet	86	70	96	72	85
InceptionResNetV2	Sıfırdan Eğitim	82	67	95	71	81
InceptionResNetV2	ImageNet	87	71	96	73	87
Vision Transformer	Sıfırdan Eğitim	81	66	94	69	80

Tablo 3.19. Artırma uygulanmış merged_ds (6 sınıf) veri seti renk uzayları performans metrikleri.

Modeller	Renk Uzayı	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	HSV	77	59	94	60	76
ResNet50	LUV	81	66	95	68	80
InceptionV3	HSV	80	67	95	69	79
InceptionV3	LUV	83	66	95	67	81
Xception	HSV	79	65	94	67	78
Xception	LUV	74	55	93	55	74
InceptionResNetV2	HSV	82	64	95	66	81
InceptionResNetV2	LUV	68	45	91	48	67
Vision Transformer	HSV	79	68	94	70	78
Vision Transformer	LUV	83	69	95	71	83

Tablo 3.20. Artırma uygulanmış merged_ds (2 sınıf) veri seti ağırlık değerleri performans metrikleri.

Modeller	Ağırlık Değerleri	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	Sıfırdan Eğitim	57	50	85	37	56
ResNet50	ImageNet	88	88	97	88	88
InceptionV3	Sıfırdan Eğitim	85	86	94	85	84
InceptionV3	ImageNet	90	90	98	89	90
Xception	Sıfırdan Eğitim	80	78	90	78	79
Xception	ImageNet	88	89	97	88	88
InceptionResNetV2	Sıfırdan Eğitim	86	85	95	85	85
InceptionResNetV2	ImageNet	87	88	97	87	87
Vision Transformer	Sıfırdan Eğitim	84	85	94	84	83

Tablo 3.21. Artırma uygulanmış merged_ds (2 sınıf) veri seti renk uzayları performans metrikleri.

Modeller	Renk Uzayı	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	HSV	80	79	91	79	79
ResNet50	LUV	69	72	88	68	68
InceptionV3	HSV	85	84	95	85	84
InceptionV3	LUV	71	74	89	71	71
Xception	HSV	82	84	93	82	81
Xception	LUV	85	85	94	85	84
InceptionResNetV2	HSV	87	87	97	86	86
InceptionResNetV2	LUV	83	81	93	82	82
Vision Transformer	HSV	84	84	95	84	83
Vision Transformer	LUV	85	85	94	68	84

Tablo 3.22. Artırma uygulanmış tympanic membrane veri seti ağırlık değerleri performans metrikleri.

Modeller	Ağırlık Değerleri	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	Sıfırdan Eğitim	75	50	89	56	74
ResNet50	ImageNet	74	56	90	61	74
InceptionV3	Sıfırdan Eğitim	55	47	86	45	54
InceptionV3	ImageNet	77	55	90	59	76
Xception	Sıfırdan Eğitim	73	49	88	53	72
Xception	ImageNet	77	60	91	63	76
InceptionResNetV2	Sıfırdan Eğitim	72	44	86	50	71
InceptionResNetV2	ImageNet	79	58	91	62	78
Vision Transformer	Sıfırdan Eğitim	74	50	89	55	73

Tablo 3.23. Artırma uygulanmış tympanic membrane veri seti renk uzayları performans metrikleri.

Modeller	Renk Uzayı	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	HSV	39	35	84	29	38
ResNet50	LUV	45	41	85	34	44
InceptionV3	HSV	67	50	88	48	66
InceptionV3	LUV	73	49	89	52	73
Xception	HSV	72	45	88	49	70
Xception	LUV	72	43	87	48	71
InceptionResNetV2	HSV	72	52	88	58	71
InceptionResNetV2	LUV	72	50	88	55	71
Vision Transformer	HSV	69	41	86	46	68
Vision Transformer	LUV	77	55	89	60	76

Tablo 3.24. Artırma ve YOLOv5 uygulanmış tympanic membrane veri seti ağırlık değerleri performans metrikleri.

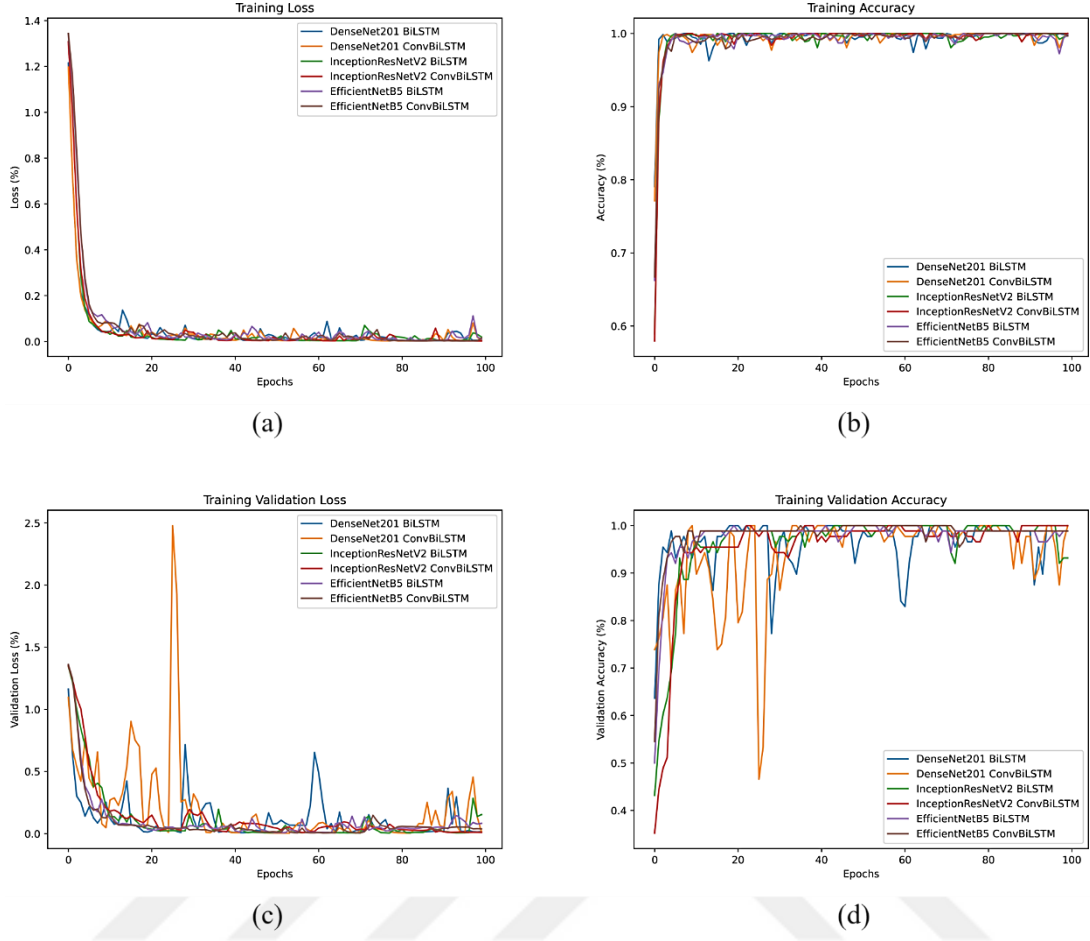
Modeller	Ağırlık Değerleri	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	Sıfırdan Eğitim	36	36	83	28	35
ResNet50	ImageNet	78	55	91	59	77
InceptionV3	Sıfırdan Eğitim	77	57	90	63	76
InceptionV3	ImageNet	75	53	89	58	74
Xception	Sıfırdan Eğitim	70	59	89	59	69
Xception	ImageNet	78	57	90	63	78
InceptionResNetV2	Sıfırdan Eğitim	75	54	90	60	75
InceptionResNetV2	ImageNet	80	59	91	63	79
Vision Transformer	Sıfırdan Eğitim	74	53	89	58	73

Tablo 3.25. Artırma ve YOLOv5 uygulanmış tympanic membrane veri seti renk uzayları performans metrikleri.

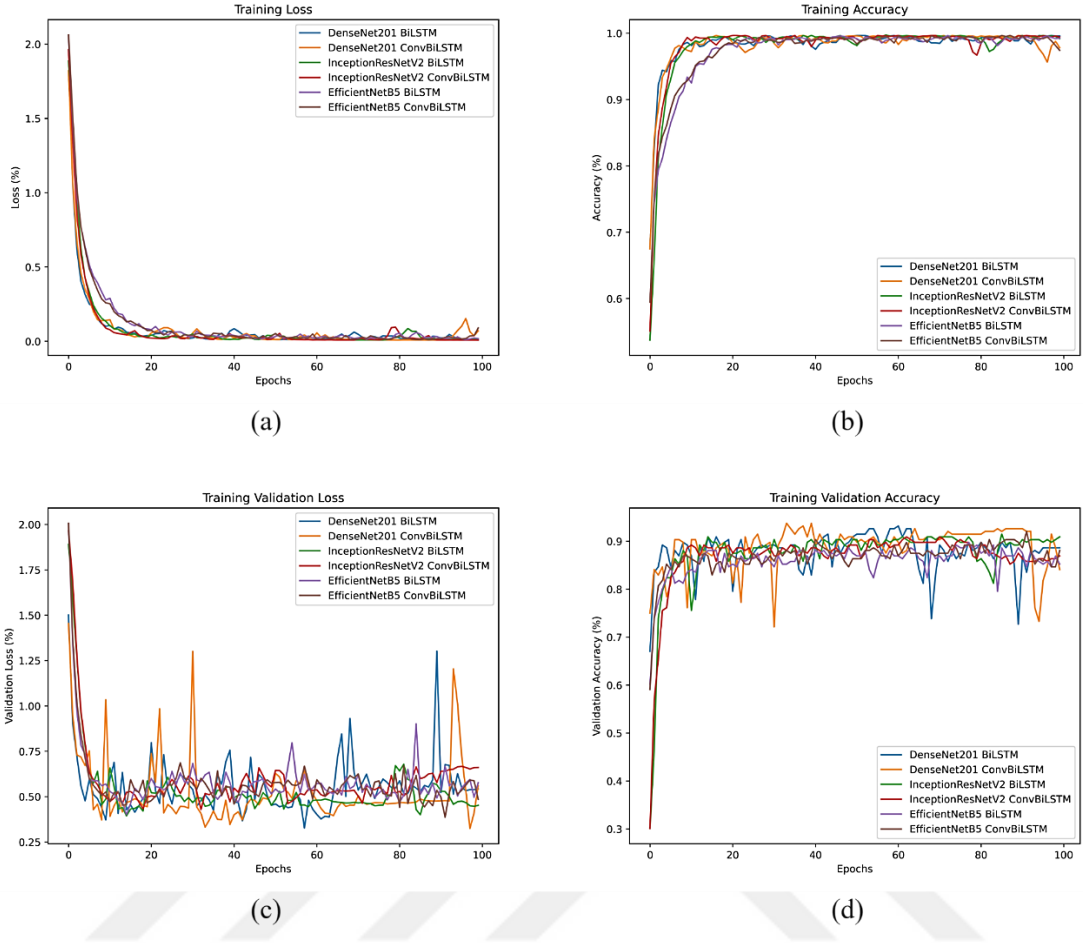
Modeller	Renk Uzaı	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
ResNet50	HSV	33	28	83	19	33
ResNet50	LUV	21	27	81	17	20
InceptionV3	HSV	75	55	89	60	74
InceptionV3	LUV	72	43	87	47	70
Xception	HSV	66	42	87	44	65
Xception	LUV	73	54	90	55	72
InceptionResNetV2	HSV	72	52	88	57	71
InceptionResNetV2	LUV	73	50	89	53	72
Vision Transformer	HSV	73	49	88	56	72
Vision Transformer	LUV	70	52	88	56	69

Tablo 3.14-3.25’den elde edilen sonuçlara göre genel olarak artırma işlemi kullanılması modellerin performansını olumlu etkilemektedir. Veri setinde bulunan veri sayısının artırılması modellerin sınıflara ait özellik haritalarını daha doğru çıkarmasını sağlamaktadır.

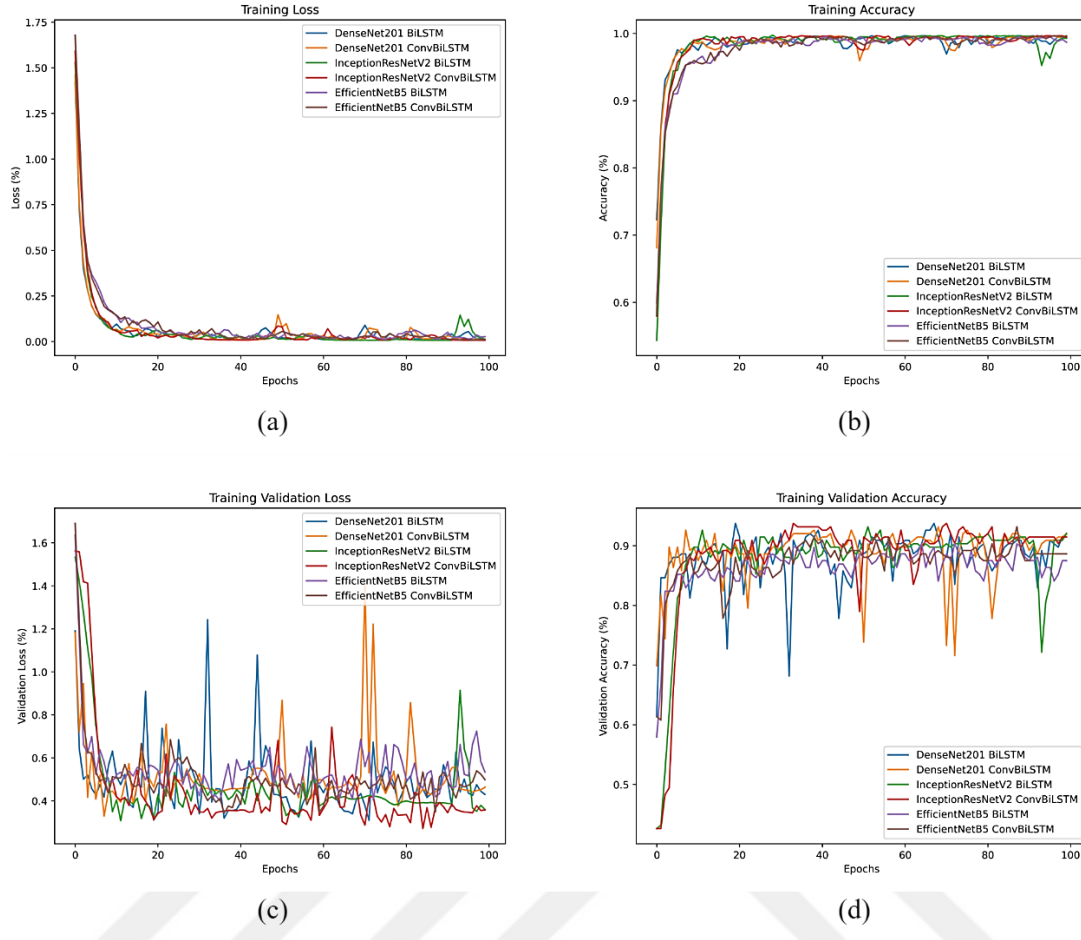
Çalışmada hazır modellerin yanı sıra hibrit oluşturulan modeller de kullanılmıştır. Hibrit modellerde özellik çıkarımı ve sınıflandırma olarak iki farklı işlem gerçekleştirilmektedir. Özellik çıkarımı işlemi için DenseNet201, EfficientNetB5 ve InceptionResNetV2 modelleri seçilmiştir. Modellerden çıkartılan özellik haritaları BiLSTM ve ConvBiLSTM modelleri tarafından ikili ve kategorik şekilde sınıflandırılmaktadır. Şekil 3.53-3.58’de hibrit modellerin eğitim grafikleri verilmektedir. Bu modellerde ImageNet ağırlık değerleri kullanıldığından dolayı HSV ve LUV renk uzaylarında bu modeller uygulanmamıştır.



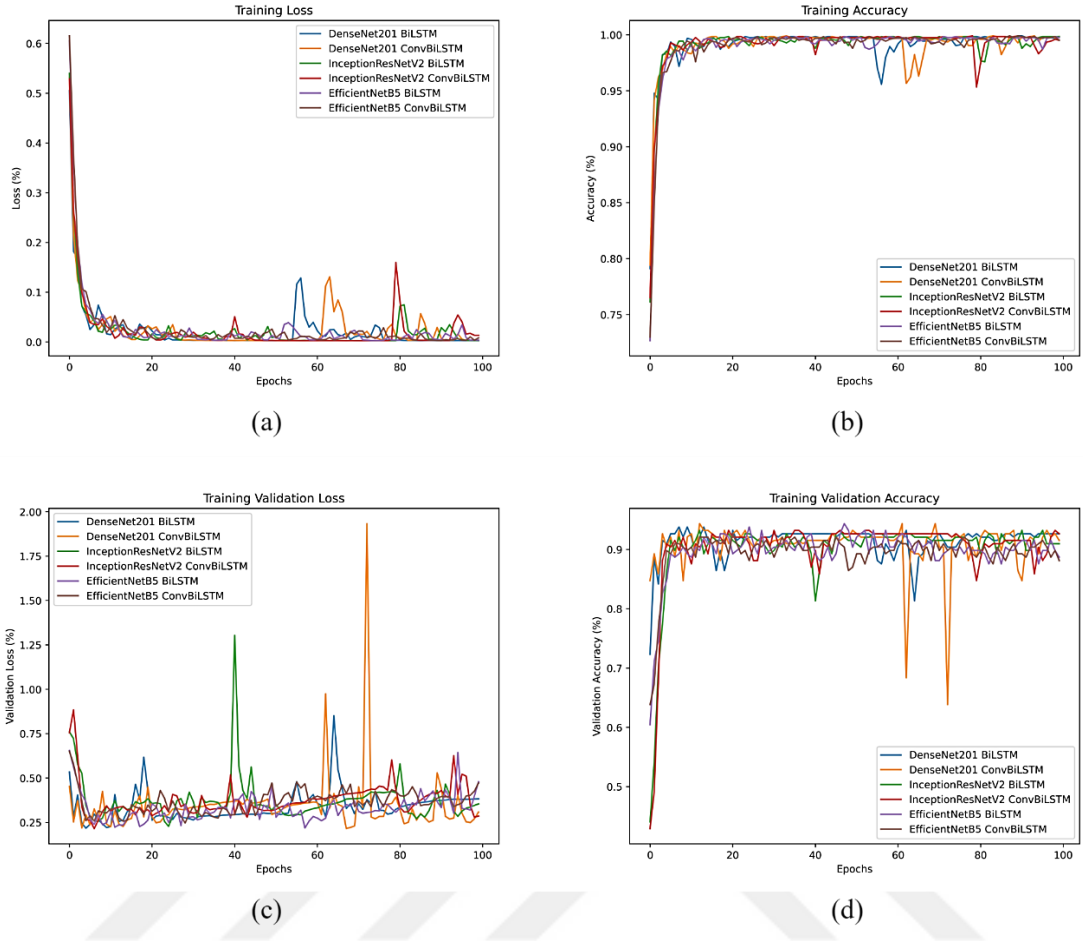
Şekil 3.53. Hibrit modellerin ear imagery veri seti üzerindeki eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c), doğrulama doğruluğu (d).



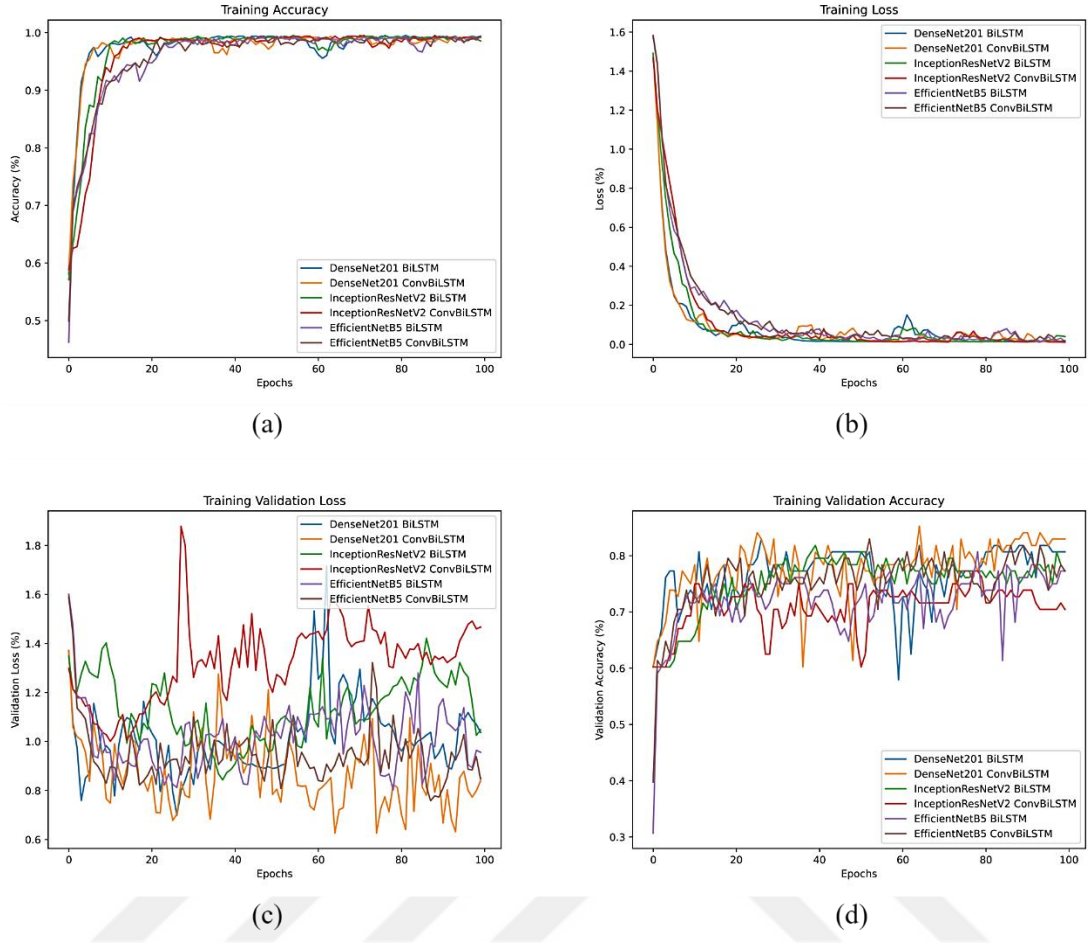
Şekil 3.54. Hibrit modellerin merged_ds veri seti üzerindeki eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c), doğrulama doğruluğu (d).



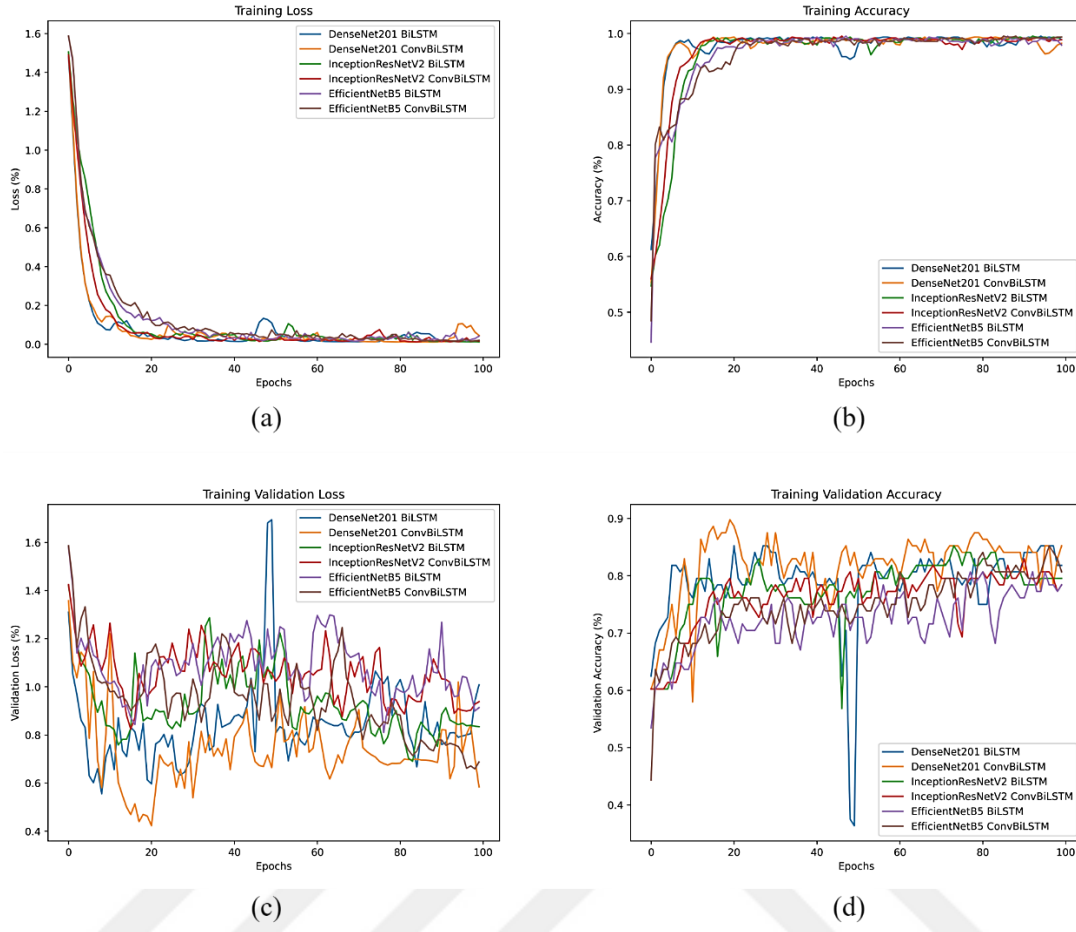
Şekil 3.55. Hibrit modellerin merged_ds (6 sınıf) veri seti üzerindeki eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c), doğrulama doğruluğu (d).



Şekil 3.56. Hibrit modellerin merged_ds (2 sınıf) veri seti üzerindeki eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c), doğrulama doğruluğu (d).



Şekil 3.57. Hibrit modellerin tympanic membrane veri seti üzerindeki eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c), doğrulama doğruluğu (d).



Şekil 3.58. Hibrit modellerin YOLOv5 uygulanmış tympanic membrane veri seti üzerindeki eğitim kaybı (a), eğitim doğruluğu (b), doğrulama kaybı (c), doğrulama doğruluğu (d).

Şekil 3.53-3.58'den elde edilen bulgulara görüldüğü üzere modellerin eğitiminde doğruluk başarısı %100'e ulaşmıştır. Bazı veri setlerinde doğrulama işleminde başarıda yüksek düşüşler görülmektedir. Eğitim işleminin düşüşlerden önce durdurulması modellerin başarılarını olumlu etkileyecektir. Tablo 3.25-3.30'da modellerin performans metrikleri verilmektedir.

Tablo 3.26. Hibrit modellerin ear imagery veri seti performans metrikleri.

Modeller	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
DenseNet201 + BiLSTM	98	98	99	98	98
DenseNet201 + ConvBiLSTM	98	98	99	98	98
EfficientNetB5 + BiLSTM	98	98	99	98	98
EfficientNetB5 + ConvBiLSTM	99	99	99	99	99
InceptionResNetV2 + BiLSTM	94	94	98	94	94
InceptionResNetV2 + ConvBiLSTM	97	97	99	97	97

Tablo 3.27. Hibrit modellerin merged_ds veri seti performans metrikleri.

Modeller	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
DenseNet201 + BiLSTM	89	78	98	82	88
DenseNet201 + ConvBiLSTM	80	70	97	70	80
EfficientNetB5 + BiLSTM	86	78	98	78	86
EfficientNetB5 + ConvBiLSTM	86	72	97	74	86
InceptionResNetV2 + BiLSTM	87	75	98	77	87
InceptionResNetV2 + ConvBiLSTM	87	75	79	98	87

Tablo 3.28. Hibrit modellerin merged_ds (6 sınıf) veri seti performans metrikleri.

Modeller	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
DenseNet201 + BiLSTM	88	78	96	81	88
DenseNet201 + ConvBiLSTM	87	73	96	76	87
EfficientNetB5 + BiLSTM	84	68	96	70	84
EfficientNetB5 + ConvBiLSTM	84	74	96	74	84
InceptionResNetV2 + BiLSTM	88	76	97	78	88
InceptionResNetV2 + ConvBiLSTM	86	73	96	75	86

Tablo 3.29. Hibrit modellerin merged_ds (2 sınıf) veri seti performans metrikleri.

Modeller	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
DenseNet201 + BiLSTM	90	91	99	90	90
DenseNet201 + ConvBiLSTM	89	90	98	89	89
EfficientNetB5 + BiLSTM	87	87	96	87	87
EfficientNetB5 + ConvBiLSTM	89	88	97	88	89
InceptionResNetV2 + BiLSTM	89	90	98	89	89
InceptionResNetV2 + ConvBiLSTM	90	90	99	90	90

Tablo 3.30. Hibrit modellerin tympanic membrane veri seti performans metrikleri.

Modeller	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
DenseNet201 + BiLSTM	82	66	92	71	81
DenseNet201 + ConvBiLSTM	80	60	91	68	79
EfficientNetB5 + BiLSTM	74	60	91	60	73
EfficientNetB5 + ConvBiLSTM	77	60	91	64	76
InceptionResNetV2 + BiLSTM	78	59	91	65	77
InceptionResNetV2 + ConvBiLSTM	78	61	91	66	77

Tablo 3.31. Hibrit modellerin YOLOv5 uygulanmış tympanic membrane veri seti performans metrikleri.

Modeller	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-Skor	Ortalama – Doğruluk (%)
DenseNet201 + BiLSTM	80	64	92	69	80
DenseNet201 + ConvBiLSTM	79	64	92	69	79
EfficientNetB5 + BiLSTM	77	63	92	62	77
EfficientNetB5 + ConvBiLSTM	80	63	92	66	80
InceptionResNetV2 + BiLSTM	78	55	90	60	78
InceptionResNetV2 + ConvBiLSTM	73	56	90	56	73

Tablo 3.25-3.30'dan elde edilen sonuçlara göre hibrit modellerin sınıflandırma başarısı hazır modellere göre daha yüksektir. Hibrit modeller ilk olarak görüntülerde bulunan özellik haritalarını çıkarmak için derin öğrenme modellerinden yararlanmaktadır. Farklı veri setlerindeki yapılan uygulamalarda genel başarısı en yüksek model DenseNet201 modelidir. Bu model 1920 boyutunda özellik haritaları çıkarmaktadır. Çıkarılan özellik haritaları BiLSTM ve ConvBiLSTM modelleri tarafından sınıflandırılmaktadır. BiLSTM modelinin sınıflandırma başarısı ConvBiLSTM modelinden daha yüksektir.

Bu çalışmada kulak içi hastalıkların derin öğrenme modelleriyle sınıflandırılması incelenmiştir. Uzmanlar tarafından yapılan sınıflandırma veya tanımlama işlemi uzmanın deneyimine ve otoskop modeline bağlıdır. Deneyimli uzmanın ve kaliteli kullanılan otoskop ile koyulan hastalık tanısı daha doğrudur. Bu kapsamda, gözlem ile koyulan tanının derin öğrenme modelleriyle gerçekleştirilmesi önerilmiştir. Çalışma kapsamında farklı veri setlerinin farklı modellerdeki sınıflandırma başarıları karşılaştırılmıştır. Veri seti olarak Ear Imagery ve Tympanic Membrane veri setleri kullanılmıştır. Bu veri setlerinde sırasıyla 1 sağlıklı, 3 ve 4 sınıf hastalıklı kulak zarı görüntüleri bulunmaktadır. Bu veri setlerinden 6 sınıf veri seti oluşturulmuştur. Bu veri setlerinde farklı etiketler bulundurularak modellerin performansları değerlendirilmiştir. Derin öğrenme modeli olarak hazır modeller ve hibrit modeller kullanılmıştır. Hazır modeller seçilirken Ear Imagery veri setinde sınıflandırma başarıları karşılaştırılmıştır ve en başarılı 4 model oluşturulan veri setlerinde kullanılmıştır. Hibrit modellerde özellik çıkarma ve sınıflandırma işlemleri ayrı modeller tarafından gerçekleştirilmektedir. Özellik çıkarma için DenseNet201, EfficientNetB5 ve InceptionResNetV2 modelleri seçilmiştir. Sınıflandırma işlemi için BiLSTM ve ConvBiLSTM modelleri kullanılmıştır. Literatürde yapılan çalışmalar ile bu çalışmanın karşılaştırılması Tablo 3.32'de verilmektedir.

Tablo 3.32. Literatür karşılaştırılması.

Yazarlar	Kullanılan Veri Seti	Sınıf Sayısı	Önerilen Model	Doğruluk Başarısı (%)
Wu vd.[9]	Gizli	3	Xception	97,82
Sundgaard vd.[10]	Gizli	3	InceptionV3	85
Alhudhaif vd.[11]	Tympanic Membrane	4	CNN, CBAM, residual blok ve hiper-sütun tekniği	98,26
Tran vd.[12]	Gizli	2	Ortak seyrek temsil algoritması	91,41
Myburgh vd.[13]	Gizli	5	NN	86,84
Mohammed vd.[14]	Ear Imagery	4	CNN-BiLSTM	99,62
Uçar vd.[15]	Ear Imagery	4	SURF + VGG16	99,06
Wang vd.[16]	Gizli	3	MESIC	90,1
Cha vd.[17]	Gizli	6	Modeller + Olasılık seçici	93,67
Chen vd.[18]	Gizli	10	MobileNetV2	97,6
Çalışkan vd.[19]	Tympanic Membrane	2	VGG16 + SVM	82,17
Sundgaard vd.[20]	Gizli	2	AlexNet tabanlı mimari	92,6
Tsutsumi vd.[21]	Karışık	2, 4	MobileNetV2	90, 91
Wang vd.[22]	Gizli	2	FasterRCNN + InceptionV3	86
Byun vd.[23]	Gizli	4	ResNet18+Shuffle Attention Module	97,18
Crowson vd.[24]	Gizli	2	ResNet34	83,8
Livingstone vd.[25]	Karışık	14	Google Cloud AutoML	90,9
Bu çalışma	Ear Imagery, Tympanic Membrane, Karışık	2, 4, 5, 6, 9	Hazır ve Hibrit modeller	91, 99, 82-80, 88, 89

Tablo 3.3'den elde edilen bulgulara göre literatürde yapılan çalışmalara yakın doğruluk başarısı elde edilmiştir. Çalışmada en yüksek doğruluk başarısı ImageNet'te önceden eğitilmiş ResNet50 ve DenseNet201+BiLSTM modellerinden elde edilmiştir. Literatürdeki çalışmalarda yapılan çalışmadan farklı olarak veri setleri sıfırdan toplanarak çalışma yürütülmüştür. Ayrıca bu çalışmada bulunmayan kulak hastalıkları da bulunmaktadır. Tympanic membrane veri setinde kullanılan otoskopa göre ear imagery veri setinde kullanılan otoskop cihazı daha başarılı kulak zarı görüntüleri üretmektedir. İleriki yapılacak çalışmalarda görüntülerin çözünürlüğünün ve sayısının artırılması doğruluk performansını olumlu yansıtmaktadır. Renk uzayının değiştirilmesi genel anlamda model başarısını olumsuz etkilemiştir.



4. SONUÇLAR

Bu çalışma kapsamında kulak zarı görüntülerin sınıflandırılması yapılmıştır. Yapılan çalışmada veri seti olarak iki farklı kulak hastalıkları bulunan veri seti kullanılmıştır. Bu veri setlerinden birleştirilerek yeni veri setleri oluşturulmuştur. Ayrıca YOLOv5 modeli kullanılarak kulak zarı bölgelerinin tespit edilerek kesilmesi sonucunda elde edilen görüntülerden ayrı bir veri seti oluşturulmuştur. Oluşturulan 6 veri setinde hazır ve hibrit modeller kullanılarak ikili ve kategorik olarak sınıflandırılmaları gerçekleştirilmiştir.

Ear imagery, Tympanic membrane, YOLOv5 uygulanmış Tympanic membrane, Merged_ds, Merged_ds (6 sınıf) ve Merged_ds (2 sınıf) veri setlerinde en yüksek başarı değerleri sırasıyla %99 (ImageNet eğitilmiş ResNet50), %82 (ImageNet eğitilmiş DenseNet201+BiLSTM), %80 (ImageNet eğitilmiş DenseNet201+BiLSTM), %89 (ImageNet eğitilmiş DenseNet201+BiLSTM), %88 (ImageNet eğitilmiş DenseNet201+BiLSTM) ve %91 (ImageNet eğitilmiş ResNet50)'dir. Genel başarı olarak ImageNet eğitilmiş DenseNet201+BiLSTM modeli tüm veri setlerinde yüksek başarı göstermiştir.

Kulak zarı hastalıklarının tanısı uzmanlar tarafından gözlem ile koyulmaktadır. Aynı işlemi derin öğrenme modelleri kullanılarak gerçekleştirmek mümkündür. Bu çalışmada karar destek sistemi geliştirmek için en uygun model araştırılmıştır. Geliştirilecek bir karar destek sistemi ile insan kaynaklı hatalar azaltılabilir, uzmanlara ikinci bir görüş sunulabilir ve uzmana erişimin zor olduğu kesimlerde hastalıkların tanısı otonom bir şekilde gerçekleştirilebilir.

5. ÖNERİLER

Bu çalışma kapsamında kulak içi hastalıkların sınıflandırılması için derin öğrenme modelleri önerilmiştir. Bu çalışmanın konusu yapılacak farklı çalışmalara yol göstereceği düşünülmektedir.

- Bu tez kapsamında farklı veri setlerindeki hastalık çeşitleri kullanılmıştır. Bu kapsamda farklı kulak hastalıkları bulunmaktadır. Ayrıca çalışma kapsamında orta ve iç kulak kanalındaki hastalıklara yönelik deneyler gerçekleştirilmiştir. Daha detaylı bir veri seti kullanılarak derin kulak hastalık sınıflandırıcı geliştirilebilir.
- Yapılan çalışmada farklı hastalıklara ait veri sayısı dengesizdir. Bu, modellerin performansını etkilemektedir. Artırma yöntemlerinin kullanılması veri setindeki başarıyı artırmasına karşın gerçek dünya problemlerinde problemler doğurmaktadır ve gerçek başarıyı yansıtmamaktadır. Dengeli sayıda görüntüye sahip bir veri setinin kullanılması gerçek dünya sonuçlarını ve modellerin başarısını olumlu yansıyacaktır.
- Günümüzde son teknoloji modeller yüksek donanım ihtiyacı duymaktadır. Bu modellerin kulak hastalıkları sınıflandırmadaki başarısı araştırılabilir.

6. KAYNAKLAR

1. <https://www.nidcd.nih.gov/health/how-do-we-hear> How Do We Hear?. 03 Nisan 2023.
2. <https://www.who.int> Primary Ear And Hearing Care Training Manual. 15 Nisan 2023.
3. <https://www.mayoclinic.org/diseases-conditions/ear-infections/diagnosis-treatment/drc-20351622> Ear Infection (Middle Ear) Diagnosis And Treatment. 03 Nisan 2023.
4. <https://www.aboutkidshealth.ca/earinfection> Ear Infection (Otitis Media) In Children. 20 Nisan 2023.
5. Horton, G. A., Simpson, M. T. W., Beyea, M. M. ve Beyea, J. A., Cerumen Management: An Updated Clinical Review And Evidence-Based Approach For Primary Care Physicians, J Prim Care Community Health, 11 (2020).
6. Ağrı, İ., Ağrı, A. E., Özdemir, D. ve Özgür, A., Cape And Tympanosclerosis, Polyphenols: Mechanisms of Action in Human Health and Disease, (2018) 421-430.
7. <https://my.clevelandclinic.org/health/diseases/24265-tympanosclerosis> Cleveland Clinic, Tympanosclerosis: Causes, Symptoms & Treatment. 23 Nisan 2023.
8. <https://www.aboutkidshealth.ca/swimmersear> Swimmer's ear (otitis externa). 25 Nisan 2023.
9. Wu, Z. vd., Deep Learning for Classification of Pediatric Otitis Media, Laryngoscope, 131, 7, (2021) 2344-2351.
10. Sundgaard, J. V. vd., Deep Metric Learning for Otitis Media Classification, Med Image Anal, (2021).
11. Alhudhaif, A., Cömert, Z. ve Polat, K., Otitis Media Detection Using Tympanic Membrane Images With A Novel Multi-Class Machine Learning Algorithm, PeerJ Comput Sci, 7, (2021) 1-22.
12. Tran, T. T., Fang, T. Y., Pham, V. T., Lin, C., Wang, P. C. ve Lo, M. T., Development of An Automatic Diagnostic Algorithm for Pediatric Otitis Media, Otology and Neurotology, 39, 8, (2018) 1060-1065.
13. Myburgh, H. C., Jose, S., Swanepoel, D. W. ve Laurent, C., Towards Low Cost Automated Smartphone And Cloud-Based Otitis Media Diagnosis, Biomed Signal Process Control, 39, (2018) 34-52.
14. Mohammed, K. K., Hassanien, A. E. ve Afify, H. M., Classification of Ear Imagery Database Using Bayesian Optimization Based On CNN-LSTM Architecture, J Digit Imaging, 35, 4, (2022) 947-961.

15. Uçar, M., Akyol, K., Atila, Ü. ve Uçar, E., Classification of Different Tympanic Membrane Conditions Using Fused Deep Hypercolumn Features and Bidirectional LSTM, IRBM, 43, 3, (2022) 187-197.
16. Wang Z. vd., Structure-Aware Deep Learning for Chronic Middle Ear Disease, Expert Syst Appl, 194, (2022).
17. Cha, D., Pae, C., Seong, S. B., Choi, J. Y. ve Park, H. J., Automated Diagnosis of Ear Disease Using Ensemble Deep Learning With A Big Otoendoscopy Image Database, EBioMedicine, 45, (2019) 606-614.
18. Chen, Y. C. vd., Smartphone-Based Artificial Intelligence Using A Transfer Learning Algorithm for The Detection And Diagnosis of Middle Ear Diseases: A Retrospective Deep Learning Study, EClinicalMedicine, 51, 201, (2022) 101543.
19. Çalışkan, A., Classification of Tympanic Membrane Images based on VGG16 Model, Kocaeli Journal of Science and Engineering, (2022).
20. Sundgaard, J. V. vd., A Deep Learning Approach for Detecting Otitis Media From Wideband Tympanometry Measurements, IEEE J Biomed Health Inform, 26, 7, (2022) 2974-2982.
21. Tsutsumi, K. vd., A Web-Based Deep Learning Model for Automated Diagnosis of Otoscopic Images, Otol Neurotol, 42, 9, (2021) 1382-1388.
22. Wang, Y. M. vd., Deep Learning in Automated Region Proposal and Diagnosis of Chronic Otitis Media Based on Computed Tomography, Ear Hear, (2020) 669-677.
23. Byun, H. vd., An Assistive Role of A Machine Learning Network In Diagnosis of Middle Ear Diseases, J Clin Med, 10, 15, (2021).
24. Crowson, M. G. vd., Machine Learning for Accurate Intraoperative Pediatric Middle Ear Effusion Diagnosis, Pediatrics, 147, 4, (2021).
25. Livingstone, D. ve Chau, J., Otoscopic Diagnosis Using Computer Vision: An Automated Machine Learning Approach, Laryngoscope, 130, 6, (2020) 1408-1413.
26. Viscaino, M., Maass, J. C., Delano, P. H., Torrente, M., Stott, C. ve Auat Cheein, F., Computer-Aided Diagnosis of External and Middle Ear Conditions: A Machine Learning Approach, PLoS One, 15, 3, (2020).
27. Cömert, Z., Fusing Fine-Tuned Deep Features for Recognizing Different Tympanic Membranes, Biocybern Biomed Eng, 40, 1, (2020) 40-51.
28. <https://monkeylearn.com/machine-learning/> An Introduction to Machine Learning. 03 Haziran 2023.

29. <https://www.coursera.org/articles/ai-vs-deep-learning-vs-machine-learning-beginners-guide> Deep Learning vs. Machine Learning: Beginner's Guide. 03 Haziran 2023.
30. <https://www.baeldung.com/cs/yolo-algorithm> What Is YOLO Algorithm?. 03 Haziran 2023.
31. <https://blog.roboflow.com/yolov5-improvements-and-evaluation/> What is YOLOv5? A Guide for Beginners. 19 Mayıs 2023.
32. <https://neurohive.io/en/popular-networks/resnet/> ResNet (34, 50, 101): Residual CNNs for Image Classification Tasks. 03 Haziran 2023.
33. <https://maelfabien.github.io/deeplearning/xception/#i-what-is-an-xception-network> Xception Model and Depthwise Separable Convolutions, 03 Mayıs 2023.
34. <https://www.v7labs.com/blog/vision-transformer-guide> Vision Transformer: What It Is & How It Works [2023 Guide]. 15 Mayıs 2023.
35. <https://www.datacamp.com/tutorial/complete-guide-data-augmentation> A Complete Guide to Data Augmentation. 27 Mayıs 2023.

ÖZGEÇMİŞ

İlkokul, ortaokul ve lise eğitimimi Trabzon’da tamamladı. 2016 yılında Avrasya Üniversitesi Bilgisayar Mühendisliği Bölümüne başladı ve 2020 yılında mezun oldu. 2020 yılında Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı’nda yüksek lisans öğrenimine başladı. 2020 yılından itibaren Samsun Üniversitesi Mühendislik Fakültesi Yazılım Mühendisliği dalında araştırma görevlisi olarak görev almakta.

