



**T.C.
BATMAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİLGİ TEKNOLOJİLERİ ANA BİLİM DALI**

YÜKSEK LİSANS TEZİ

**DERİN ÖĞRENME YAKLAŞIMLARIYLA ARAÇ KAPORTA
BOYALARININ ORJİNALLİK TESPİTİ**

Mehmet Mehdi TUNÇYÜZLÜ

**Ağustos-2025
BATMAN**

**T.C.
BATMAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİLGİ TEKNOLOJİLERİ ANA BİLİM DALI**

YÜKSEK LİSANS TEZİ

**DERİN ÖĞRENME YAKLAŞIMLARIYLA ARAÇ KAPORTA
BOYALARININ ORJİNALLİK TESPİTİ**

Mehmet Mehdi TUNÇYÜZLÜ

**Danışman
Dr. Hafzullah İŞ**

Diğer Jüri Yeleri

Doç. Dr. Mustafa KAYA Doç. Dr. Abidin ÇALIŞKAN

**Ağustos-2025
BATMAN**

TEZ KABUL VE ONAYI

Mehmet Mehdi TUNÇYÜZLÜ tarafından hazırlanan “Derin Öğrenme Yaklaşımlarıyla Araç Kaporta Boyalarının Orijinallik Tespiti” adlı tez çalışması 01/08/2025 tarihinde aşağıdaki jüri tarafından oy birliği ile Batman Üniversitesi Lisansüstü Eğitim Enstitüsü Bilgi Teknolojileri Ana Bilim Dalı’nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Jüri Üyeleri

İmza

Başkan

Doç. Dr. Mustafa KAYA

.....

Danışman

Dr. Hafzullah İŞ

.....

Üye

Doç. Dr. Abidin ÇALIŞKAN

.....

Yukarıdaki sonucu onaylarım.

Dr. Öğr. Üyesi Ömer Murat ÖTER
Lisansüstü Eğitim Enstitüsü Müdürü

ETİK BEYANI

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını beyan eder, aksinin ortaya çıkması durumunda her türlü yasal sorumluluğu kabullendiğimi bildiririm.

ETHICAL DECLARATION

I declare that all the information in this thesis has been obtained within the framework of ethical behavior and academic rules, and that the source of any statements and information that do not belong to me in this study prepared in accordance with the thesis writing rules has been fully cited, and I declare that I accept all kinds of legal responsibility in case of any contrary situation.

Mehmet Mehdi TUNÇYÜZLÜ
Tarih: 01/08/2025

ÖZET

YÜKSEK LİSANS TEZİ

DERİN ÖĞRENME YAKLAŞIMLARIYLA ARAÇ KAPORTA BOYALARININ ORİJİNALLİK TESPİTİ

Mehmet Mehdi TUNÇYÜZLÜ

Batman Üniversitesi Lisansüstü Eğitim Enstitüsü

Bilgi Teknolojileri Ana Bilim Dalı

Danışman: Dr. Hafzullah İŞ

2025, 66 Sayfa

Otomotiv endüstrisi, yüksek kaliteli ve güvenilir araç üretiminde boyanın önemli bir rol oynadığını kabul eder. Araç kaporta boyası, sadece estetik bir unsur olmanın ötesinde, aynı zamanda aracın dış yüzeyini çevresel faktörlere karşı koruma görevini üstlenir. Bu nedenle, araç kaporta boyalarının orijinallik ve kalitesi, hem üreticiler hem de tüketiciler için büyük önem taşır. Günümüzde, araç kaporta boyalarının orijinallikini tespit etme işlemi genellikle araç ekspertizlerinde boya kalınlığını ölçen mikron cihazı ya da araçların kaporta ustalarının el yordamı ve deneyime dayalı yöntemlerle yapılmaktadır. Ancak, bu yöntemler hem zaman alıcıdır hem de insan hatasına açıktır. Teknolojinin hızla gelişmesiyle birlikte, otomotiv sektöründe kalite kontrol süreçlerinde daha yenilikçi ve güvenilir yöntemlere olan ihtiyaç artmaktadır. Bu bağlamda, derin öğrenme teknolojileri, araç kaporta boyalarının orijinallikini tespit etmek için umut verici bir çözüm olarak ortaya çıkmaktadır. Derin öğrenme, büyük veri kümeleri üzerinde öğrenme yeteneği sayesinde, karmaşık desenleri ve ilişkileri tespit edebilme kapasitesine sahiptir. Görüntü işleme teknikleri ile birleştğinde, derin öğrenme modelleri, araç kaporta boyalarının orijinallikini ve kalitesini yüksek doğrulukla analiz edebilir. Bu tezde, derin öğrenme tabanlı bir sistemin araç kaporta boyalarının orijinallikini tespit etmedeki etkinliği incelenmiştir. Çalışmada, araç kaporta boyalarının birincil kaynağından olup olmadığını görüntü analizleri kullanarak tespit etmek hedeflenmiştir. Bu sayede, araç kaporta boyalarının orijinallik tespitinde bilinen yöntemlere alternatif bir yaklaşım sunulması amaçlanmaktadır. Görüntü işleme ve derin öğrenme yöntemleri kullanılarak oluşturulan bir ESA modeli, fabrika çıkışlı boyalı ve sonradan boyanmış araçlardan alınmış toplam 2000 görüntü örnekleri üzerinde eğitilmiştir. Veri kümesi, %80-%20 eğitim ve doğrulama setlerine bölünüp ResNet50, VGG16, InceptionV3, EfficientNetB0, MobileNetV2 ESA modelleri ile test edilerek en yüksek doğruluk oranı InceptionV3 ile %98 olarak hesaplanmıştır. Bu sonuçlar, derin öğrenmenin araç boyalarının orijinallikini tespit etmek için kullanılabilir olduğunu göstermektedir, ancak modelin doğruluğunu artırmak için daha fazla veri ve farklı model yapılarına ihtiyaç duyulabilir.

Anahtar Kelimeler: Derin öğrenme, ESA, Araç Kaporta Boya Orijinallik

ABSTRACT
MASTER THESIS
**ORIJINALITY DETERMINATION OF VEHICLE BODY PAINTS WITH DEEP
LEARNING APPROACHES**

Mehmet Mehdi TUNÇYÜZLÜ

Batman University Graduate Education Institute

Department of Information Technologies

Advisor: Dr. Hafzullah İŞ

Year, 66 Pages

The automotive industry acknowledges that paint plays a crucial role in the production of high-quality and reliable vehicles. Automotive body paint is not merely an aesthetic component but also serves the essential function of protecting the vehicle's exterior surface against environmental factors. Therefore, the originality and quality of automotive body paint hold significant importance for both manufacturers and consumers. Today, the process of detecting the originality of automotive body paint is generally performed during vehicle inspections using micron gauges that measure paint thickness, or through manual methods relying on the skill and experience of body repair experts. However, these methods are both time-consuming and prone to human error. With the rapid advancement of technology, the automotive sector faces a growing demand for more innovative and reliable methods in quality control processes. In this context, deep learning technologies have emerged as a promising solution for determining the originality of automotive body paint. Thanks to its ability to learn from large datasets, deep learning has the capacity to identify complex patterns and relationships. When combined with image processing techniques, deep learning models can analyze the originality and quality of automotive body paint with high accuracy. In this thesis, the effectiveness of a deep learning-based system in detecting the originality of automotive body paint is examined. The study aims to determine whether automotive body paint originates from its primary source through image analysis. In doing so, it seeks to provide an alternative approach to the commonly used methods for detecting paint originality. By employing image processing and deep learning methods, an ESA model was developed and trained on a dataset of 2,000 images obtained from factory-painted and repainted vehicles. The dataset was divided into 80% training and 20% validation sets, and the ESA models ResNet50, VGG16, InceptionV3, EfficientNetB0, and MobileNetV2 were tested. The highest accuracy rate, 98%, was achieved with InceptionV3. These results indicate that deep learning can be effectively used to determine the originality of automotive paint; however, larger datasets and different model architectures may be required to further improve accuracy.

Keywords: Deep Learning, ESA, Vehicle body paint originality

ÖN SÖZ

Yüksek Lisans tezimin hazırlanması ve oluşturulmasında, değerli bilgilerini benimle paylaşan, bana her zaman destek olan, tecrübe ve tavsiyeleriyle bana yol gösteren, kendisine ne zaman danışsam bana kıymetli zamanını ayırıp sabırla ve büyük bir ilgiyle bana faydalı olmaya çalışan kıymetli danışman hocam Dr. Hafzullah İŞ' e teşekkürlerimi sunuyorum.

Ayrıca hiçbir zaman maddi ve manevi desteklerini benden esirgemeyen biricik eşim Meral TUNÇYÜZLÜ'ye ve sevgili aileme teşekkür ediyor ve onlara minnettar olduğumu dile getirmek istiyorum.

Mehmet Mehdi TUNÇYÜZLÜ
BATMAN-2025

İÇİNDEKİLER

ÖZET	iv
ABSTRACT.....	v
ÖN SÖZ	vi
İÇİNDEKİLER	vii
TABLolar LİSTESİ	ix
ŞEKİLLER LİSTESİ	x
SİMGELER VE KISALTMALAR.....	xi
1. GİRİŞ.....	1
1.1. Amaç	3
1.2. Kapsam	3
2. LİTERATÜR ÇALIŞMALARI.....	5
3. METARYAL VE YÖNTEM	9
3.1. Materyal.....	9
3.1.1. Fotoğraf çekim koşulları.....	9
3.1.2. Orijinal fotoğraflar.....	9
3.1.3. Boyalı fotoğraflar.....	11
3.1.4. Veri setinin hazırlanması	12
3.2. Yöntem.....	12
3.2.1. Transfer öğrenme yöntemleri.....	13
3.2.2. Evrişimsel sinir ağları	14
3.2.2.1. ResNet50.....	25
3.2.2.2. VGG16.....	27
3.2.2.3. InceptionV3	29
3.2.2.4. EfficientNetB0	31
3.2.2.5. MobileNetV2	33
3.2.3. Performans ölçütleri.....	35
4. BULGULAR.....	37
4.1. ResNet50 Bulguları.....	38
4.2. VGG16 Bulguları.....	42
4.3. InceptionV3 Bulguları	46
4.4. EfficientNetB0 Bulguları.....	50
4.5. MobileNetV2 Bulguları	54
5. TARTIŞMA.....	58
6. SONUÇLAR VE ÖNERİLER.....	59

6.1. Sonuçlar	59
6.2. Öneriler	59
KAYNAKLAR	62

TABLÖLER LİSTESİ

Tablo 3.1. Veri setindeki görüntü sınıfları ve sayısı	9
Tablo 4.1. Modeller için kullanılan standart parametreler	37
Tablo 4.2. Resnet50 ile sınıflandırma metrikleri	40
Tablo 4.3. VGG16 ile sınıflandırma metrikleri.....	44
Tablo 4.4. InceptionV3 ile sınıflandırma metrikleri	48
Tablo 4.5. EfficientNetB0 ile sınıflandırma metrikleri.....	52
Tablo 4.6. MobileNetV2 ile sınıflandırma metrikleri	56

ŞEKİLLER LİSTESİ

Şekil 3.1. Orijinal boyalı araç fotoğraf örnekleri.....	10
Şekil 3.2. Orijinal araç boyalarının renk dağılımı	10
Şekil 3.3. Sonradan boyanmış araç fotoğraf örnekleri	11
Şekil 3.4. Sonradan boyanmış araç boyalarının renk dağılımı	12
Şekil 3.5. ESA ağ yapısı (Hidaka & Kurita, 2017).....	15
Şekil 3.6. Evrişim işlemi.....	17
Şekil 3.7. Maksimum ve ortalama havuzlama İşlemleri	19
Şekil 3.8. Tam bağlantılı katman.....	20
Şekil 3.9. Aktivasyon fonksiyonları (Kalyoncu, 2020).....	22
Şekil 3.10. Standart sinir ağı ve dropout uygulanan sinir ağı (Hinton vd. 2012)	23
Şekil 3.11. Düzleştirme katmanı.....	24
Şekil 3.12. ResNet50 mimarisi (Habek vd. 2024).....	26
Şekil 3.13. VGG16 mimarisi katman blok diyagramı (Neurohive, 2018)	29
Şekil 3.14. InceptionV3 mimarisi (SB Shetty vd. 2022).....	31
Şekil 3.15. EfficientNetB0 mimarisi (Chen vd. 2024)	32
Şekil 3.16. MobileNetV2 mimarisi (Shahoveisi vd. 2023)	34
Şekil 3.17. Karışıklık Matrisi	35
Şekil 4.1. ResNet50 mimarisinin eğitim ve doğrulama grafikleri.....	38
Şekil 4.2. ResNet50 mimarisinin karışıklık matrisi.....	39
Şekil 4.3. ResNet50 mimarisi ROC grafiği	40
Şekil 4.4. ResNet50 ile örnek tahmin görüntüleri	41
Şekil 4.5. VGG16 mimarisinin eğitim ve doğrulama grafikleri	42
Şekil 4.6. VGG16 mimarisinin karışıklık matrisi.....	43
Şekil 4.7. VGG16 mimarisi ROC grafiği	44
Şekil 4.8. VGG16 ile örnek tahmin görüntüleri	45
Şekil 4.9. InceptionV3 mimarisinin eğitim ve doğrulama grafikleri.....	46
Şekil 4.10. InceptionV3 mimarisinin karışıklık matrisi	47
Şekil 4.11. InceptionV3 mimarisi ROC grafiği.....	48
Şekil 4.12. InceptionV3 ile örnek tahmin görüntüleri	49
Şekil 4.13. EfficientNetB0 mimarisinin eğitim ve doğrulama grafikleri	50
Şekil 4.14. EfficientNetB0 mimarisinin karışıklık matrisi	51
Şekil 4.16. EfficientNetB0 ile örnek tahmin görüntüleri.....	53
Şekil 4.17. MobileNetV2 mimarisinin eğitim ve doğrulama grafikleri	54
Şekil 4.18. MobileNetV2 mimarisinin karışıklık matrisi	55
Şekil 4.19. MobileNetV2 mimarisi ROC grafiği.....	56
Şekil 4.20. MobileNetV2 ile örnek tahmin görüntüleri.....	57

SİMGELER VE KISALTMALAR

Simgeler

Açıklama

b	Bias
f	Aktivasyon fonksiyonu
X	Giriş verisi
W	Filtre ağırlıkları
Z	Özellik Haritası

Kısaltmalar

Açıklama

ACO	Ant Colony Optimization, Karınca Kolonisi Optimizasyonu
ADAM	Adaptive Moment Estimation, Uyarlanabilir Anlık Tahminleme
AI	Artificial Intelligence, Yapay Zeka
ANN	Artificial Neural Network, Yapay Sinir Ağları
APF	Artificial Potential Field, Yapay Potansiyel Alan
AUC	Area Under the Curve , Eğri Altındaki Alan
CNN	Convolutional Neural Networks, Evrişimsel Sinir Ağları
DL	Deep Learning, Derin Öğrenme
DPI	Dots Per Inch, İnç Başına Nokta Sayısı
ESA	Evrişimsel Sinir Ağı
FC	Fully Connected Layer, Tam Bağlantılı Katman
FPS	Frames Per Second, Saniyedeki Kare Sayısı
HSV	Hue- Saturation-Value , Renk tonu-Doygunluk- Parlaklık
ISO	International Organization for Standardization, Uluslararası Standardizasyon Örgütü
LLM	Large Language Model, Büyük Dil Modeli
mAP	Mean Average Precision, Ortalama Hassasiyet
MB	Megabyte, Veri Depolama Birimi
ML	Machine Learning, Makine Öğrenimi
ResNet	Residual Neural Networks, Artık Sinir Ağları
RGB	Kırmızı-Green-Blue, Kırmızı-Yeşil-Mavi
OAA	One-Against-All, Bire Karşı Tümü
SSD	Single Shot MultiBox Detector, Tek Adımda Çoklu Kutu Algılama
SSE	Spatial Squeeze and Excitation, Uzamsal Sıkıştırma ve Uyarım
SVM	Support Vector Machine, Destek Vektör Makinesi
ROC	Receiver Operating Characteristic ,Alıcı İşletim Karakteristiği Eğrisi

1. GİRİŞ

Otomotiv endüstrisi, teknolojik yeniliklerin hızla ilerlediği ve tüketici taleplerinin sürekli olarak değiştiği dinamik bir sektördür. Araçların güvenliği, performansı ve estetiği, tüketici tercihlerini etkileyen önemli faktörler arasında yer almaktadır. Özellikle araçların dış görünüşü, ilk izlenimi oluşturmaya ve aracın genel durumunu yansıtması açısından büyük bir öneme sahiptir. Bu bağlamda, araç kaporta boyası, sadece estetik bir unsur değil, aynı zamanda aracın dış etkenlerden korunması için kritik bir rol oynayan bir bileşendir. Kaporta boyası, araç gövdesini paslanma ve diğer çevresel hasarlara karşı korurken, aynı zamanda aracın değerini ve ömrünü de doğrudan etkiler. Ancak, araç kaporta boyasının orijinalliğinin ve kalitesinin doğru bir şekilde tespit edilmesi, özellikle ikinci el araç piyasasında büyük bir önem taşımaktadır. Orijinal boya, aracın fabrikadan çıktığı haliyle korunmuş olmasını ifade ederken, sonradan yapılan boyama işlemleri genellikle hasar veya tamir geçmişine işaret eder. Bu nedenle, boyanın orijinalliğinin ve kalitesinin doğru bir şekilde değerlendirilmesi, hem araç sahipleri hem de potansiyel alıcılar için kritik bir öneme sahiptir. Geleneksel yöntemlerle yapılan incelemeler ve gözlemsel değerlendirmeler, insan faktörüne bağlı olarak hatalara açık olabilir ve bu da subjektif sonuçlara yol açabilir. Bu yöntemler, kaportadaki boya kalınlığını ölçen cihazlar veya görsel inceleme teknikleri kullanılarak gerçekleştirilir. Ancak, bu teknikler genellikle sınırlı hassasiyete sahiptir ve uzman olmayan kişiler tarafından uygulandığında yanlış sonuçlar verebilir. Ayrıca, geleneksel yöntemlerin uygulanması zaman alıcı olabilir ve büyük ölçüde kullanıcı deneyimine bağlıdır. Bu nedenle, bu yöntemler her zaman güvenilir sonuçlar vermez ve araç kaporta boyasının gerçek durumu hakkında eksik bilgi sunabilir.

Son yıllarda, yapay zeka ve özellikle derin öğrenme alanındaki gelişmeler, otomotiv endüstrisinde de önemli ilerlemeler kaydedilmesini sağlamıştır. Derin öğrenme, büyük veri setleri üzerinde eğitim alarak karmaşık görevleri yerine getirebilen çok katmanlı sinir ağları kullanır. Bu ağlar, görüntü işleme, ses tanıma ve doğal dil işleme gibi çeşitli alanlarda başarıyla uygulanmaktadır. Özellikle evrimsel sinir ağları (ESA), görüntü tabanlı verilerin işlenmesi ve analizi konusunda güçlü yeteneklere sahiptir.

ESA'lar, farklı katmanlar arasında öznitelik çıkarımı yaparak, görüntüdeki anlamlı özellikleri öğrenir ve bu sayede yüzey kusurlarını tespit etmede yüksek doğruluk sağlar.

Araç kaporta boya kusurları tespitine yönelik araştırmalar, otomotiv endüstrisinde halen sınırlıdır. Araç gövdesindeki yüzey kusurlarını görselleştirmek için optik üç boyutlu tarayıcı sistemler kullanılmıştır (Eichhorn, vd., 2005). Kusur yerleri ihmal edilebilir, nitelikli ve niteliksiz olmak üzere üç kategoriye ayrılarak sınıflandırmada yüksek doğruluk oranlarına ulaşılmıştır (Chung, vd., 2006). Diğer çalışmalar ise görüntü işleme teknikleri kullanarak kusur tespitinde önemli ilerlemeler kaydetmiştir.

Derin öğrenme ve özellikle evrişimsel sinir ağları, görüntü sınıflandırma ve kusur tespiti alanında geniş bir uygulama alanına sahiptir. ESA doküman tanıma için kullanılmaya başlamıştır (LeCun, vd., 1998). Krizhevsky, vd. (2012) ImageNet veri setinde ESA kullanarak elde ettikleri başarılı sonuçlarla devam etmiştir. Szegedy, vd. (2015) tarafından geliştirilen GoogleNet ve He, vd., (2016) geliştirdiği ResNet modelleri, derin öğrenmenin gücünü ve potansiyelini göstermiştir. Bu çalışmalar, ESA'ların görüntü sınıflandırma ve obje tespitindeki etkinliğini ortaya koymaktadır. Özellikle otomotiv endüstrisinde, boya kusurlarının tespitine yönelik yapılan çalışmalarda ESA'lar önemli bir rol oynamaktadır. Ayrıca, Zheng (2021), insan davranışları tanıma ve Kaur, vd. (2022), araç plakası tanıma gibi farklı alanlarda ESA'ların başarıyla uygulandığını göstermişlerdir.

Bu çalışmada, araç gövdesi boyalarının orijinal veya sonradan boyalı olup olmadığını belirlemek amacıyla derin öğrenme tabanlı bir sınıflandırma yaklaşımı geliştirilmiştir. Özellikle ESA mimarileri kullanılarak, araç kaporta boya yüzeylerinin sınıflandırılması hedeflenmiştir. Yapılan literatür incelemeleri, ESA tabanlı yöntemlerin görsel sınıflandırma problemlerinde yüksek doğruluk oranları sağladığını ve otomotiv endüstrisinde de yüzey kontrolü ve kusur tespiti gibi uygulamalarda başarıyla kullanılmaya başlandığını göstermektedir. Bu bağlamda önerilen modelin, araç gövdesi boyalarının orijinal üretim standartlarına uygunluğu ile sonradan yapılan boya işlemleri arasındaki farkı daha yüksek doğrulukla tespit edebileceği ve mevcut geleneksel yöntemlere kıyasla daha tutarlı ve otomatik bir çözüm sunacağı öngörülmektedir.

1.1. Amaç

Teknolojinin hızla ilerlemesi, birçok endüstride olduğu gibi otomotiv sektöründe de önemli değişiklikler ve yenilikler getirmiştir. Özellikle araç kaportalarındaki boya orijinalliğinin tespiti, araçların değerinin belirlenmesinde ve ikinci el pazarında önemli bir rol oynamaktadır. Geleneksel inceleme yöntemleri, bu tespiti yaparken zaman alıcı ve subjektif olabilir. Bu durum, özellikle büyük ölçekli işlemler için verimliliği azaltabilir ve insan faktörüne bağlı olarak hata payını artırabilir.

Bu çalışmanın amacı, derin öğrenme yöntemlerinden biri olan ESA modelini kullanarak araç kaportalarındaki orijinal ve boyalı bölgeleri sınıflandırmaktır. ESA, özellikle görüntü işleme ve nesne tanıma gibi alanlarda büyük başarılar elde etmiştir ve bu çalışmada araç kaportalarında boya orijinalliğini tespit etmek için kullanılacaktır. Araç boya orijinalliği tespitinde çeşitli ESA modelleri eğitilecek ve bu modellerin araç kaporta boyasını sınıflandırma başarısı üzerindeki etkileri değerlendirilecektir. Amaç, geleneksel yöntemlere kıyasla daha hızlı, daha doğru ve tutarlı sonuçlar elde ederek, boya orijinallik tespitinde hataları minimize etmek ve maliyetleri düşürmektir. Bu bağlamda, çalışmada elde edilecek bulgular, otomotiv endüstrisinde kalite kontrol süreçlerinin iyileştirilmesine ve müşteri memnuniyetinin artırılmasına katkı sağlayacaktır. Ayrıca, yapay zeka tabanlı bu yeni yaklaşım, insan hatalarını minimize ederek daha objektif bir değerlendirme sunacak ve kullanıcıların doğru kararlar almasına yardımcı olacaktır. Bu, özellikle araç alım-satım işlemlerinde şeffaflık ve güvenliği artıracaktır.

1.2. Kapsam

Otomotiv sektörü, araçların estetik değerlerinin ve işlevselliğinin korunmasında büyük önem taşımaktadır. Özellikle ikinci el araç piyasasında, araçların kaporta boyalarının orijinalliği, araçların geçmişi ve değerini belirlemede kritik bir faktördür. Ülkemizde, araç kaporta boya orijinalliği, araç alım-satım işlemlerinde önemli bir kriter olarak kabul edilir ve bu nedenle, araç sahipleri ve alıcıları arasında büyük bir hassasiyet konusu olmuştur. Bu konunun doğruluğu ve güvenilirliği, araç alıcılarının mağduriyetini önlemek için hayati öneme sahiptir.

Bu çalışmanın kapsamı, araç kaporta boya orijinalliği tespit yöntemlerini incelemek ve bu yöntemler arasında en etkili olanları belirlemektir. Geleneksel yöntemler, genellikle manuel incelemelere dayanmakta ve bu da insan faktörüne bağlı olarak sonuçların değişken olmasına neden olmaktadır. Ancak, yapay zeka

teknolojilerinin gelişmesiyle birlikte, bu tür subjektif yöntemlerin yerine daha objektif ve güvenilir yöntemlerin kullanılması mümkün hale gelmiştir. Bu çalışmada, birbirine çok yakın benzerlikte olan görüntülerin yapay zeka tabanlı ESA ağı kullanılarak sınıflandırılması hedeflenmiştir. Özellikle ResNet50, VGG16, InceptionV3, EfficientB0 ve MobileNetV2 gibi ESA modelleri test edilip sınıflandırma başarımları karşılaştırılacaktır. Bu çalışma, mobil cihazlardan elde edilen görüntülerin ne derece tutarlı bir şekilde sınıflandırılabilirliğini araştırarak, günlük kullanımda pratik bir çözüm sunmayı amaçlamaktadır. Bu bağlamda, çalışma hem akademik literatüre hem de endüstriyel uygulamalara katkıda bulunacak ve yeni teknolojilerin otomotiv sektöründe nasıl kullanılabileceği konusunda önemli bulgular sağlayacaktır. Ayrıca, geliştirilen sistemin pratikte uygulanabilirliği ve kullanıcı dostu olması, gelecekteki uygulamalar için önemli bir referans noktası olacaktır.

2. LİTERATÜR ÇALIŞMALARI

Bu bölümde, araç kaporta boya orijinalliği tespiti çalışmamıza ışık tutabilecek başta otomotiv endüstrisinde ESA ile yapılmış çeşitli görüntü sınıflandırma çalışmalarına başvurulmuştur.

Araç kaporta boya filmi kusurlarının tespiti, otomotiv endüstrisinde kalite kontrol süreçlerinin kritik bir parçasıdır. Yüzey kusurlarının zamanında ve doğru bir şekilde tespit edilmesi, üretim sürecinin verimliliğini artırmakta ve nihai ürünün kalitesini sağlamaktadır. Bu çalışmada, derin öğrenme teknikleri ve transfer öğrenme stratejileri kullanılarak araç kaporta boya orijinalliğinin tespit edilmesi hedeflenmiştir.

Otomobil kaporta boyası filmi kusur tespiti alanında son yıllarda birçok yöntem önerilmiştir. Kamani, vd. (2011) , boya kusurlarının tespiti ve sınıflandırılması için geliştirilmiş destek vektör makinesini (SVM) kullanmışlardır. İlk olarak, kusur konumu yerel varyans rotasyonu değişmez ölçüsüyle bulunmuş ve ardından tüm kusur konumları bire-karşı hepsi destek vektör makinesi (OAA-SVM) ile sınıflandırılmıştır. Cheng, vd. (2017), görüntü işlemeye dayalı otomobil boya filminin görüntü tanınması ve sınıflandırılması için bir yöntem önermişlerdir. Kusur konumunu belirlemek için grafik teorisine dayalı morfolojik işleme ve görüntü bölümlleme algoritması kullanılmış ve tanıma oranı %90 olarak rapor edilmiştir. Edris, vd. (2015), araç gövdesinin yüzeyinden üç boyutlu bir görüntü elde etmek, düzensiz konumu işaretlemek ve sinir ağını kullanarak işaretlenen alanı tanımlamak için üç boyutlu bir tarayıcı kullanmışlardır. Zhang, vd. (2020), araç gövde üretim boya kusurlarının otomatik tespiti için bilgisayarlı görü ve derin öğrenme yöntemleri kullanılmıştır. Küçük veri örneklerinden kaynaklanan aşırı uyum sorununu çözmek için yeni bir veri iyileştirme algoritması önerilmiştir. MobileNet-SSD ağına özellik katmanı iyileştirilmiş ve sınırlayıcı kutu eşleştirme stratejisi optimize edilmiştir. Deneysel sonuçlar, geliştirilen algoritmanın %95'in üzerinde bir doğrulukla araç boya kusurlarını tespit edebildiğini ve geleneksel SSD algoritmasından %10 daha hızlı olduğunu göstermiştir. Böylece araç boya kusurlarının gerçek zamanlı ve doğru bir şekilde tespit edilebilmesi sağlanmıştır.

Derin öğrenme, kusur tespitinde yaygın olarak kullanılmaktadır. Jeyaraj, vd. (2019), ileri öğrenme algoritmasını kullanarak bilgisayar destekli bir kumaş hatası tespit ve sınıflandırma sistemi tasarlamış ve geliştirmişlerdir. Bu sistemde temel

malzemelerdeki altı çeşit kusur, derin ESA ile sınıflandırılmakta ve sınıflandırmanın ortalama doğruluğu %96,55'tir. Kang, (2024) tarafından yapılan çalışmada, kumaş kusur tespiti konusunda mevcut zorlukların ele alındığı ve bu sorunlara yönelik hafif bir ağ tabanlı yaklaşım önerildiği vurgulanmaktadır. Özellikle, tekstil sektöründeki yüksek gerçek zamanlı gereksinimler ve derin öğrenme modellerinin karmaşıklığının, endüstriyel sahalarda uygulanmasını sınırladığı belirtilmiştir. Bu bağlamda, yazarlar YOLOv5s modelini temel alarak, evrişim blok dikkat modülü ve özellik geliştirme modülü gibi bileşenleri entegre eden bir yöntem geliştirmişlerdir. Ayrıca, YOLOv5s'nin kayıp fonksiyonu CIOU_Loss olarak değiştirilmiş ve bu sayede modelin bilgi çıkarma yeteneği ile tahmin doğruluğu artırılmıştır. Deneysel sonuçlara göre, geliştirilen model ham kumaş ve desenli kumaş veri kümelerinde ortalama doğruluk hassasiyeti (mAP) açısından sırasıyla %86,4 ve %75,8'e ulaşmış, orijinal YOLOv5s algoritmasına göre %7,6 ve %1,7 oranında bir iyileşme sağlamıştır. Ayrıca, modelin ortalama tespit hızı 102 FPS olup, endüstriyel alanlarda gerçek zamanlı gereksinimleri karşılayabilecek kapasitededir. Zhao, vd. (2020), sürekli döküm ürünlerinin yüzey kusurlarını tespit etmek için tamamen bağlı ağı ve tamamen kıvrımlı ağı birleştiren bir füzyon derinliği ESA modeli tasarlayarak yeni bir iki aşamalı kusur tespit stratejisi önermektedir. Hu, vd. (2014) yaptıkları çalışmada, çelik şerit yüzey görüntülerindeki kusurları sınıflandırmak için SVM kullanılmıştır. Görüntülerin ikili yapılmasının ardından, kusur hedef görüntüsü ve buna karşılık gelen ikili görüntü birleştirilerek geometrik, gri seviye ve şekil özellikleri gibi üç tür görüntü özelliği çıkarılmıştır. SVM tabanlı sınıflandırma modeli için Gauss radyal bazlı çekirdek fonksiyonu kullanılmış, model parametreleri çapraz doğrulama yöntemiyle belirlenmiş ve çok sınıflı sınıflandırıcı olarak one-versus-one yöntemi uygulanmıştır. Deney sonuçları, SVM modelinin, ortalama sınıflandırma doğruluğunda geleneksel geri yayılım sinir ağına dayalı sınıflandırma modelinden daha iyi performans gösterdiğini ortaya koymuştur. Wei, vd. (2020), yüzey kusurlarını tespit edebilen ve dört tür kusuru doğru bir şekilde tanımlayabilen PDDNet'i önermişlerdir. Aşınma derinliği tahmin sonuçları sahada ölçülen verilerle karşılaştırıldığında, sonuçlar yöntemin yüksek tahmin doğruluğuna sahip olduğunu göstermektedir.

Xu, vd. (2020) tarafından yapılan çalışmada, araç gövde boyası kusurlarının tespiti için bilgisayarla görme temelli bir yöntem önerilmiştir. Araştırmacılar, geleneksel kenar tespit algoritmalarının sınırlamalarını aşmak amacıyla, otomotiv boya özelliklerine dayalı bir karınca kolonisi optimizasyonu kenar tespit algoritması (APF-ACO) geliştirmişlerdir. Bu yöntem, ACO algoritmasında hem genel keşif hem de yerel inceleme

süreçlerini bir arada kullanıyor. Böylece hem geniş alanda en iyi çözümleri arıyor hem de detaylı olarak ince ayar yaparak daha doğru (yüksek hassasiyetli) kusur tespiti sağlıyor. Ayrıca, HSV (Hue, Saturation, Value) renk uzayını kullanarak yansıtıcı alanları tespit eden ve bu alanları elimine eden bir algoritma da geliştirilmiştir. Deneysel sonuçlar, önerilen sistemin saç, çizik, damla, boya katlanması ve yağmur damlası gibi beş farklı kusur türünü %97,76 doğruluk oranıyla tespit edebildiğini göstermektedir. Bu sonuçlar, geliştirilen APF-ACO algoritmasının hem zaman açısından verimli olduğunu hem de yüksek doğruluk sağladığını ortaya koymaktadır.

Xu, vd. (2021) tarafından yapılan çalışmada, araç gövdesi ve benzeri metalik yüzeylerin boya filminin kusurlarını tespit etmek için derin öğrenme tabanlı bir sistem önerilmiştir. Çalışmada, SSD (Single Shot MultiBox Detector) ve Faster R-CNN (Region-based Convolutional Neural Networks) algoritmaları kullanılarak yüzey kalitesi değerlendirilmiştir. Araştırmacılar, yüksek çözünürlüklü görüntüler kullanarak 2000 görüntüden oluşan bir veri seti oluşturmuş ve bu veri seti üzerinde SSD ve Faster R-CNN yöntemlerini eğitmişlerdir. Elde edilen sonuçlar, bu derin öğrenme algoritmalarının, özellikle manuel olarak tasarlanması gereken özelliklere ihtiyaç duymadan, yüzey kusurlarını tespit etmede daha sağlam olduğunu göstermiştir. Deneysel veriler, bu yöntemlerin çeşitli sanayi uygulamalarında, örneğin ahşap kusur tespiti, parça yüzey kusur tespiti ve kaynak kusur tespiti gibi alanlarda uygulanabilirliğini ortaya koymuştur. Çalışma, derin öğrenme algoritmalarının kullanımıyla ilgili önemli bulgular sunmuş ve bu yöntemlerin geleneksel görsel tespit yöntemlerine göre daha yüksek doğruluk ve hız sağladığını ortaya koymuştur.

Araç gövde yüzeylerindeki kusurların tespiti, otomotiv endüstrisinde kalite kontrol sürecinin önemli bir parçasıdır. Geleneksel manuel muayene yöntemleri, insan hatalarına karşı savunmasızdır ve bu durum, kusurların doğru bir şekilde tespit edilmesini zorlaştırabilir. Molina, vd. (2017), aynasal araç gövde yüzeylerindeki spekül yüzeylerdeki küçük kusurların tespiti için yeni bir algoritma sunulmuştur. Çalışma, spekül yüzeylerde 0.2 mm çapa kadar olan kusurların tespiti için deflektometri ve görüntü birleştirme tekniklerine dayanan bir yaklaşım önermektedir. Önerilen sistem, kusurları kenarlar, köşeler ve derin içbükeylikler gibi zorlu alanlarda bile tespit edebilmektedir. Bu yöntem, geleneksel yöntemlerin aksine tüm yüzeyi kapsamayı amaçlayarak, stil çizgileri ve diğer karmaşık geometrilere sahip alanlarda bile etkin bir tespit sunmaktadır. Algoritma, Mercedes-Benz'in İspanya'daki Vitoria fabrikasında başarıyla uygulanmış ve kusurların tespiti sürecini büyük ölçüde otomatikleştirerek, daha

hızlı ve güvenilir sonuçlar elde edilmesini sağlamıştır. Wang vd. (2024) tarafından yapılan çalışmada Otomotiv yapıştırıcı hatalarının tespitinde, manuel inceleme düşük verimlilik ve insan görüşündeki kör noktalar nedeniyle zorluklar yaşar, bu da parçaların performansını olumsuz etkiler. Bu nedenle, otomatik tespit yöntemleri büyük önem taşır. Otomatik yapıştırma süreçlerinde yapıştırıcı hatalarının üretim üzerindeki olumsuz etkilerini gidermek amacıyla, bu çalışmada otomotiv uygulamaları için iyileştirilmiş YOLOv8 tabanlı bir yapıştırıcı hata tespit yöntemi önerilmiştir. İlk olarak, modelimizin önemli özelliklere seçici olarak odaklanabilmesini sağlamak için modelin omurga kısmında SSE (spatial squeeze and excitation) dikkat mekanizması kullanılmıştır. Daha sonra, orijinal sınır kutusu kayıp fonksiyonu yerine WIoU kayıp fonksiyonu eklenmiştir. Deneysel sonuçlar, bu yöntemin orijinal YOLOv8n modelinin mAP50 değerini %3.25 oranında artırdığını ve ortalama olarak görüntü başına 7.9 ms tespit süresiyle, saniyede 126.58 kare hızına (FPS) ulaştığını göstermektedir. Bu da gerçek zamanlı hata tespiti gereksinimlerini karşılamaktadır. Park vd. (2016) çalışmasında, makine öğrenimi (ML) tabanlı görüntüleme sistemleriyle yüzey kusurlarının otomatik tespitinde önemli ilerlemeler kaydedildiği belirtilmiştir. Özellikle, ESA kullanımı ile geleneksel yöntemlerin ötesinde yüksek doğruluk ve verimlilik sağlanmıştır. Bu çalışma, derin öğrenme modellerinin yüzey kusuru denetiminde uygulanabilirliğini göstermiş ve çeşitli yüzeylerdeki kusurların tespitinde enerji ve maliyet tasarrufu sağladığını kanıtlamıştır.

Derin öğrenmeye (DL) dayalı otomatik yüzey kusurları tespiti, farklı görsel inceleme uygulamaları üzerinde büyük bir etkiye sahip olan ve hızla gelişen bir araştırma alanıdır. Derin öğrenme yöntemleri, bu zorlu görevi yerine getirmek için etkili çözümler sunar. Ancak, ayna benzeri malzemelerdeki yüzey kusurları tespiti, üretim süreçlerinde hala önemli bir zorluktur. Geleneksel yöntemler, yüzey kusurlarını tespit etmek için yaygın olarak kullanılır, ancak bu yöntemler genellikle el ile çıkarılan özelliklere dayanır ve belirli aydınlatma koşullarına ihtiyaç duyar, bu da tespit yeteneklerini sınırlı hale getirir. Karangwa vd. (2020) yaptıkları çalışmada ayna benzeri malzemelerdeki yüzey kusurlarının tespitini otomatikleştirmek için daha hızlı bir RCNN modeli önerilmektedir. Önerilen model, özellik çıkarıcı olarak VGG16 mimarisini kullanarak filtrelerin ve özellik haritasının yararlı özellikleri yakalamasını sağlar. Seramik bir nesne üzerinde, çatlak, mikro çatlak, iğne deliği, kir, çukur ve leke gibi altı farklı yüzey kusuru türü incelenmiştir. Deneysel sonuçlar, önerilen modelin %94 doğruluk oranına ulaşarak geleneksel yöntemlere kıyasla üstün performans sergilediğini göstermiştir.

3. METARYAL VE YÖNTEM

3.1. Materyal

Bu çalışmada, araç kaportalarındaki boya orijinalliğinin tespiti amacıyla orijinal ve boyalı yüzeylere sahip araçların fotoğrafları kullanılmıştır. Toplamda 2000 fotoğraf, 1000 adet orijinal ve 1000 adet boyalı yüzeylere sahip araçlardan elde edilmiştir. Fotoğrafların çekiminde kullanılan ekipman, çekim koşulları ve verilerin toplanma süreci aşağıda ayrıntılı olarak açıklanmıştır.

Tablo 3.1. Veri setindeki görüntü sınıfları ve sayısı

Sıra	Sınıf	Görüntü Sayısı
1	Boyalı	1000
2	Orijinal	1000

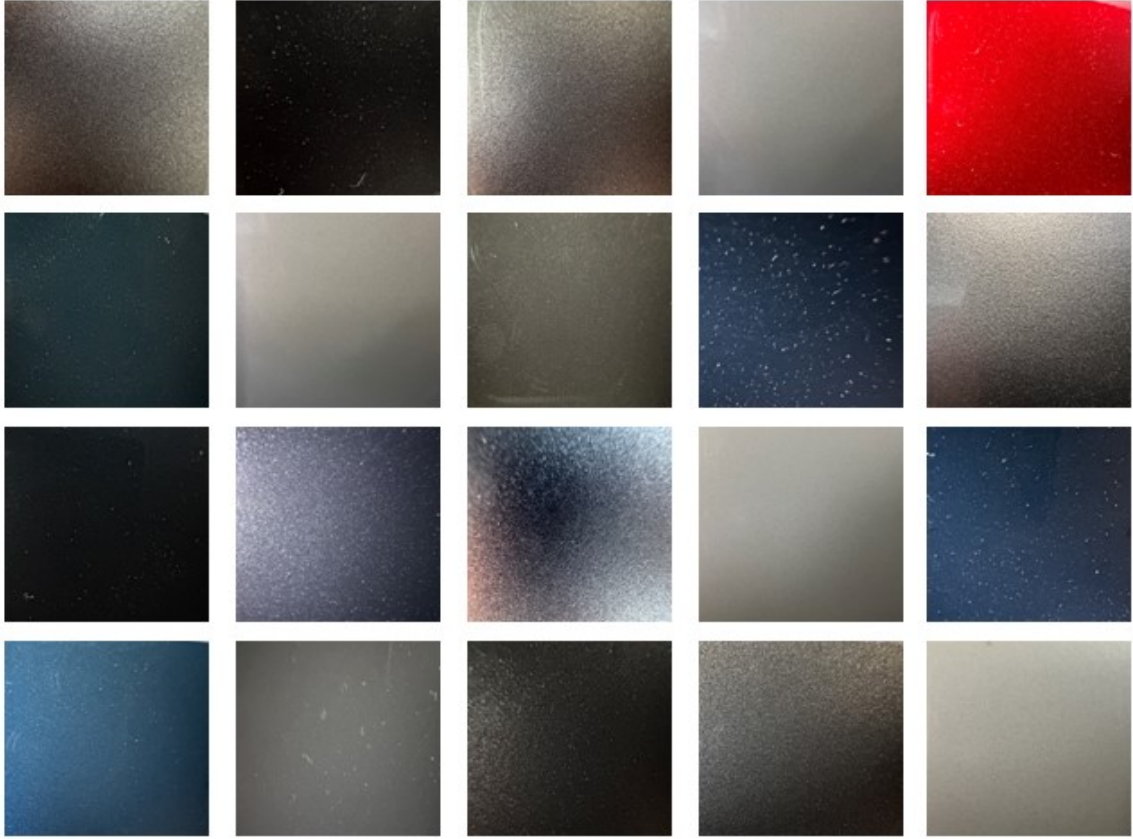
3.1.1. Fotoğraf çekim koşulları

Fotoğraflar, gündüz saatlerinde ve yansımanın minimum düzeyde olduğu gölgeli alanlarda çekilmiştir. Bu durum, görüntülerin kalitesini artırmak ve yansımaların tespit sürecine olumsuz etkilerini en aza indirmek amacıyla tercih edilmiştir. Çekimler sırasında doğal ışığın homojen dağılımı sağlanmış ve gereksiz parlamalar önlenmiştir. Ayrıca çekimi yapılan yüzeyler önceden temizlenip çekime hazır hale getirilmiştir. Kullanılan çekim cihazı Apple iPhone 14 Pro Max'tir. Cihazın kamerası, f/2.2 diyafram açıklığı, (1/15)-(1/60) aralığında saniye enstantane hızı ve ISO 50-2500 gibi ayarlarla kullanılmıştır. Bu ayarlar, düşük ışık koşullarında bile yeterli pozlama sağlarken, parazitlenmeyi minimize etmek için optimize edilmiştir. Fotoğraflar, 72 DPI çözünürlük ve 24 bit renk derinliği ile çekilmiş olup, her bir görüntü en az 1 MB en fazla 22 MB boyutundadır. Çekim mesafesi boya görüntüsü alınan yüzey ile çekim cihazı arasında 3-7 cm değişmektedir. Bu özellikler, fotoğrafların yüksek kalitede ve detaylı olmasını sağlamıştır.

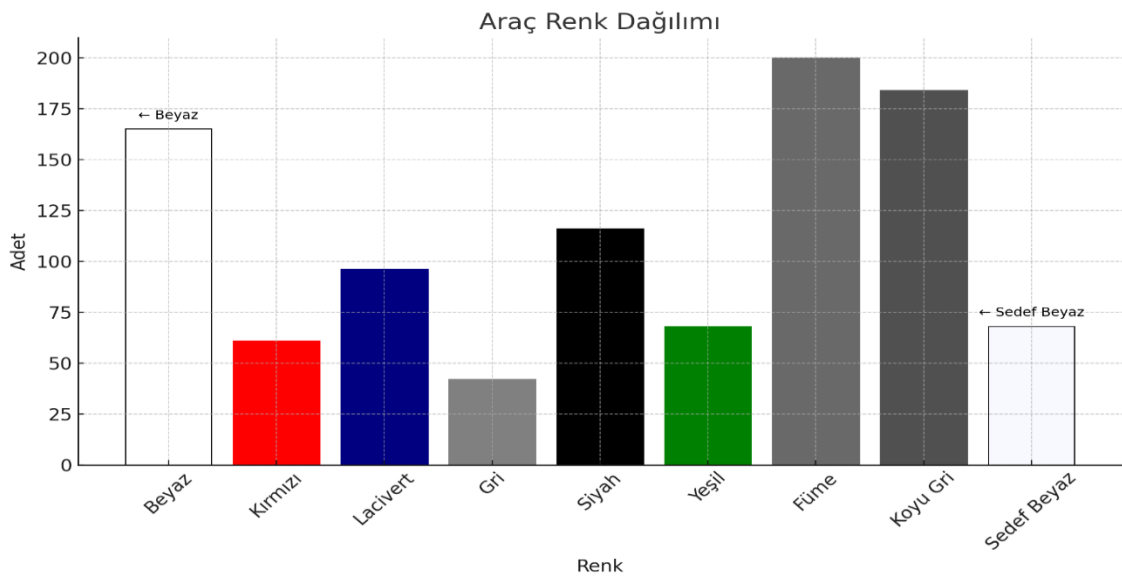
3.1.2. Orijinal fotoğraflar

Orijinal boyalı araç fotoğrafları, yalnızca tek bir üretici firmadan temin edilmiştir. Bu seçim, orijinal boya uygulamalarının tutarlılığını sağlamak ve verilerin homojenliğini

korumak amacıyla yapılmıştır. Fotoğraflanan araçlar, üretici tarafından belirlenen fabrika standartlarına uygun olarak orijinal boyalı araçlardır. Her araç, farklı açılardan ve çeşitli kaporta bölgelerinden çekilerek, boyanın orijinal yapısını ve kusursuzluğunu belgelemek üzere özenle seçilmiştir.



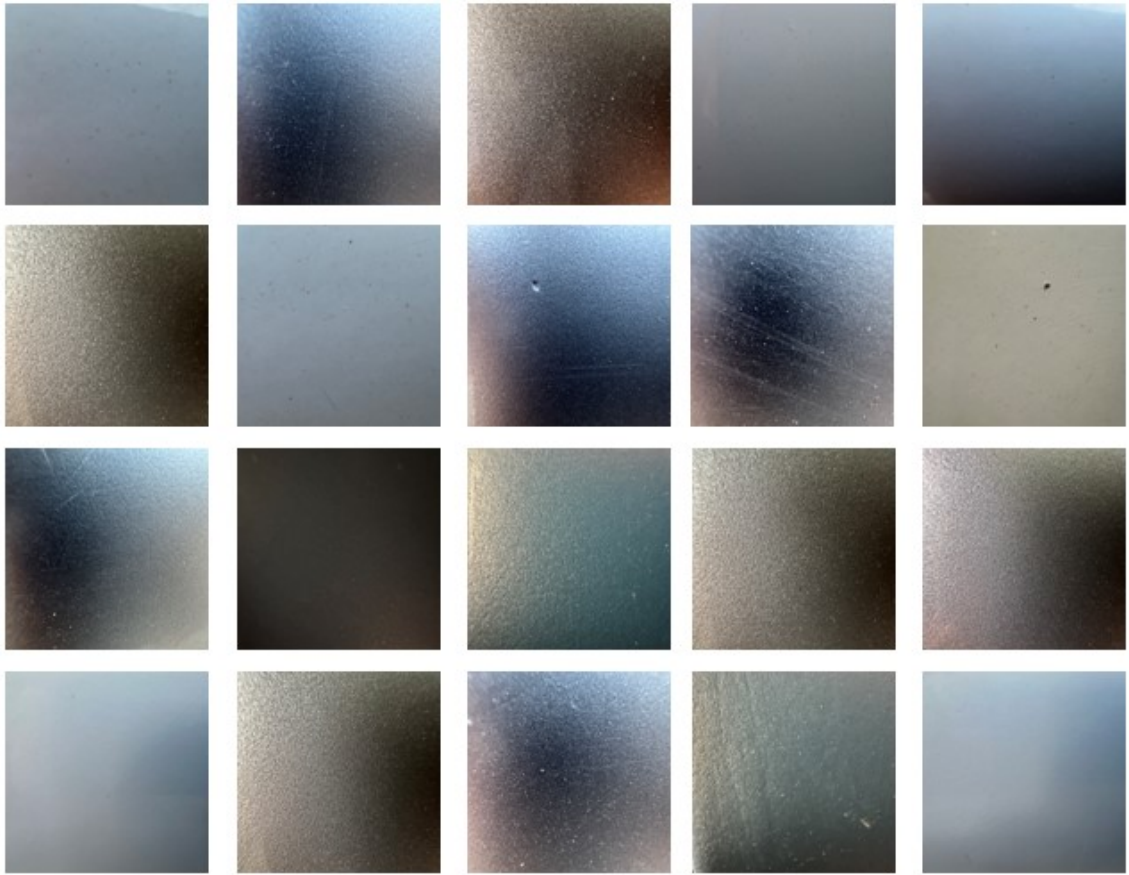
Şekil 3.1. Orijinal boyalı araç fotoğraf örnekleri



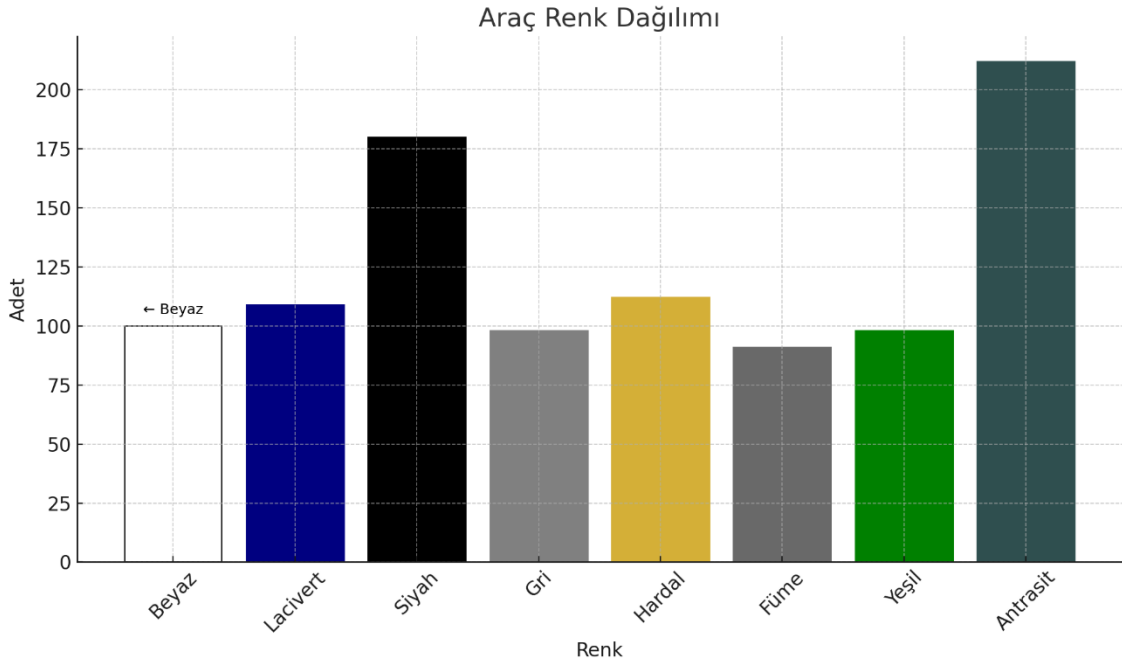
Şekil 3.2. Orijinal araç boyalarının renk dağılımı

3.1.3. Boyalı fotoğraflar

Boyalı araç fotoğrafları, ikinci el araç pazarından ve farklı boyama işlemlerinden elde edilmiştir. Bu araçlar, orijinal boya dışında ek boyama işlemlerine tabi tutulmuş olup, genellikle tamir veya modifikasyon sonrası tekrar boyanmış araçlardır. Çeşitli boyama teknikleri ve boya türleri kullanılmış olup, bu çeşitlilik, farklı kusur türlerinin tespit edilmesi ve sınıflandırılması için önemlidir. Boyalı araçların fotoğrafları da orijinal araçlarda olduğu gibi çeşitli açılardan çekilmiş ve her bir bölgenin detaylı görüntüsü alınmıştır.



Şekil 3.3. Sonradan boyanmış araç fotoğraf örnekleri



Şekil 3.4. Sonradan boyanmış araç boyalarının renk dağılımı

3.1.4. Veri setinin hazırlanması

Çekilen tüm fotoğraflar, yüksek çözünürlükte ve kalite kaybı olmadan dijital ortamda saklanmıştır. Fotoğrafların dosya isimlendirilmesi, çekim sırasına göre yapılmış ve her bir dosya, orijinal veya boyalı olarak etiketlenmiştir. Yansımaların etkisini en aza indirmek amacıyla, fotoğraflar üzerinde görüntü işleme teknikleri kullanılarak gerekli düzenlemeler yapılmış ve yansıma azaltma algoritmaları uygulanmıştır. Bu yöntemlerle elde edilen veri seti, hem orijinal hem de boyalı yüzeylerin detaylı analizine olanak tanımakta ve boya orijinalliğinin tespitine yönelik geliştirilmiş modellerin eğitim ve doğrulama süreçlerinde kullanılmaktadır.

3.2. Yöntem

Bu çalışmada, deneyler 11. nesil Intel(R) Core(TM) i7-1165G7 işlemciye sahip, 2.80 GHz hızında çalışan ve 16 GB RAM donanımına sahip bir bilgisayarda gerçekleştirilmiştir. Derin öğrenme modellerinin uygulanmasında python programlama dilinin tensorflow.keras kütüphanesi kullanılmıştır. Çalışmada, temel olarak beş farklı ESA modeli kullanılmıştır.

3.2.1. Transfer öğrenme yöntemleri

Transfer öğrenme, bir modelin önceden edinmiş olduğu bilgi ve deneyimi, farklı ancak ilgili bir görevde yeniden kullanmasını sağlayan bir öğrenme paradigmasıdır. Geleneksel makine öğrenimi sistemleri, her yeni görev için sıfırdan model eğitimi gerektirirken; transfer öğrenme, mevcut modellerin parametrelerini, yapısını ya da temsillerini yeniden kullanarak daha hızlı ve etkili sonuçlar elde etmeyi amaçlar. Özellikle sınırlı veri setine sahip alanlarda, transfer öğrenme önemli bir avantaj sağlar (Pan & Yang, 2010). Transfer öğrenme yöntemleri genel olarak görev benzerliği, veri etiket durumu ve özellik dağılımları gibi etkenlere göre sınıflandırılır. Bu sınıflandırmalar sayesinde, uygulama alanına en uygun transfer stratejisi seçilebilir. Aşağıda transfer öğrenmenin temel türleri sistematik biçimde açıklanmıştır.

Tümevarımsal Transfer Öğrenme (Inductive Transfer Learning): Bu yaklaşımda, hedef görevde etiketli veriler mevcuttur ve kaynak ile hedef görev genellikle farklıdır. Öğrenme süreci, hedef görevdeki tahmin performansını artırmayı amaçlar. Derin öğrenme alanında yaygın bir uygulaması, büyük bir veri kümesi üzerinde (örneğin ImageNet) eğitilmiş bir ESA modelinin, başka bir görsel sınıflandırma görevine aktarılmasıdır. Bu tür aktarımlar genellikle ince ayar (fine-tuning) veya özellik çıkarımı (feature extraction) şeklinde uygulanır (Yosinski, vd. 2014).

Geçişli Transfer Öğrenme (Transductive Transfer Learning): Transductive öğrenmede, kaynak ve hedef görev aynıdır, ancak veri dağılımları farklıdır. Yani model aynı görevi çözmeye çalışırken farklı domenden gelen verileri işler. Bu tür transfer, özellikle domain adaptation (alan uyarlaması) problemlerinde önem kazanır. Örneğin, bir metin sınıflandırma modeli İngilizce verilerle eğitildikten sonra, benzer görevleri Almanca ya da Fransızca veri üzerinde gerçekleştirmek istendiğinde bu yaklaşım kullanılır. Burada amaç, kaynak alanda öğrenilen bilgiyi hedef alana uygun hale getirmektir (Ben-David vd. 2010).

Gözetimsiz Transfer Öğrenme (Unsupervised Transfer Learning): Bu yöntemde hem kaynak hem de hedef alanlar için etiketli veri bulunmaz. Genellikle özellik çıkarımı, boyut indirgeme veya kümeleme gibi görevlerde uygulanır. Gözetimsiz transfer öğrenme, temsil öğrenme (representation learning) kavramıyla yakın ilişkilidir. Bu tür senaryolar,

medikal görüntüleme, ağ analizi veya doğal dil işleme gibi çok boyutlu ve az etiketli veri ortamlarında öne çıkar.

Few-Shot ve Zero-Shot Learning: Modern derin öğrenme modellerinde sıkça karşılaşılan bir başka strateji ise çok az örnekle (few-shot) ya da hiç örnek görmeden (zero-shot) hedef sınıfların öğrenilmesini sağlayan yöntemlerdir. Bu teknikler, genellikle büyük dil modelleri (LLM) ya da meta-öğrenme yaklaşımlarıyla birlikte kullanılır. Örneğin, bir GPT modelinin daha önce görmediği bir görevde mantıklı yanıtlar üretebilmesi, zero-shot transferin tipik bir örneğidir (Brown, vd. 2020).

Multi-Task ve Continual Learning Yaklaşımları: Transfer öğrenmeye dayalı daha ileri yöntemler arasında çok görevli öğrenme (multi-task learning) ve sürekli öğrenme (continual/lifelong learning) yaklaşımları yer alır. Multi-task öğrenmede model, birden fazla görevi eş zamanlı olarak öğrenir ve ortak temsiller sayesinde bilgi paylaşımı yapılır. Sürekli öğrenmede ise model, zamana yayılan ardışık görevlerde öğrendiklerini unutmadan yeni bilgileri edinmeyi hedefler (Parisi, vd. 2019).

Transfer öğrenme, makine öğrenimi sistemlerinin verimliliğini artıran güçlü bir araçtır. Özellikle derin sinir ağlarının karmaşık yapıları, bu bilgiyi farklı görevlere aktarmaya uygun hale getirmiştir. Doğru seçilen transfer stratejileri, yalnızca eğitim süresini ve veri gereksinimini azaltmakla kalmaz, aynı zamanda genelleme kapasitesini de önemli ölçüde artırır. Günümüzde transfer öğrenme, sağlık, güvenlik, doğal dil işleme, otonom sistemler ve finans gibi birçok alanda yaygın biçimde kullanılmaktadır.

3.2.2. Evrişimsel sinir ağları

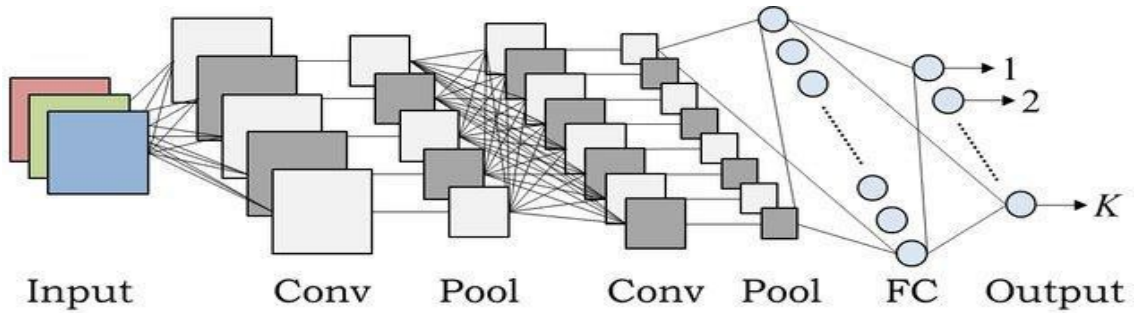
Evrişimsel sinir ağları , LeCun vd. (1998) tarafından önerilen ve Krizhevsky vd. (2012) tarafından AlexNet modeli ile popüler hale getirilen derin öğrenme mimarileridir. Evrişimli Sinir Ağları , derin öğrenme tabanlı yapay sinir ağlarının bir alt sınıfını oluşturmaktadır ve özellikle görüntü işleme, video analizi ve nesne tanıma gibi alanlarda önemli başarılar elde etmiştir. ESA mimarisi, biyolojik görsel korteksten esinlenerek geliştirilen ve görüntü verilerindeki mekânsal ve zamansal ilişkileri modelleme kapasitesine sahip bir yapıdır. Bu ağlar, özellikle yüksek boyutlu verilerde öz nitelik çıkarımı ve sınıflandırma görevlerini otomatikleştirerek klasik yöntemlere göre önemli avantajlar sağlamaktadır.

ESA mimarisi genellikle üç temel bileşenden oluşmaktadır: evrişim katmanları (convolutional layers), havuzlama katmanları (pooling layers) ve tam bağlantılı katmanlar

(fully connected layers). Evrişim katmanları, öğrenilebilir filtreler aracılığıyla girdi verisinden çeşitli özellik haritaları (feature maps) oluşturur. Bu filtreler sayesinde görüntülerdeki kenar, doku ve diğer temel görsel özellikler etkin biçimde yakalanır. Aktivasyon fonksiyonları, genellikle ReLU (Rectified Linear Unit) kullanılarak doğrusal olmayanlık kazandırılır. Bunu takip eden havuzlama katmanları, öznitelik haritalarının boyutlarını azaltarak modelin hesaplama yükünü hafifletmekte ve konumsal genelleme yeteneğini artırmaktadır. Son aşamada yer alan tam bağlantılı katmanlar ise sınıflandırma veya regresyon gibi görevlerde nihai karar mekanizmasını oluşturmaktadır.

ESA'ların en önemli avantajlarından biri, özellik mühendisliği ihtiyacını ortadan kaldırmasıdır. Geleneksel yöntemlerde manuel olarak belirlenmesi gereken öznitelikler, ESA'lar tarafından otomatik olarak öğrenilmektedir. Ayrıca, parametre paylaşımı sayesinde ağ yapısı karmaşık görüntü verileri üzerinde daha verimli ve daha düşük hesaplama maliyetiyle işlem yapabilmektedir.

Bu yapıların başarıları, LeNet-5, AlexNet, VGGNet, ResNet ve MobileNet gibi çeşitli mimarilerin geliştirilmesini teşvik etmiş; günümüzde medikal görüntüleme, otonom sürüş sistemleri, yüz tanıma, endüstriyel kalite kontrol gibi pek çok uygulama alanında yaygın olarak kullanılmaktadır. ESA'lar, yalnızca sınıflandırma problemlerinde değil; aynı zamanda segmentasyon, nesne tespiti ve stil transferi gibi farklı bilgisayarla görü görevlerinde de üstün performans sergilemektedir. Aşağıda ESA temel yapısı gösterilmiştir.



Şekil 3.5. ESA ağ yapısı (Hidaka & Kurita, 2017)

Giriş katmanı (Input layer) : Giriş katmanı, modelin giriş verilerini alır. Örneğin, bir görüntü sınıflandırma probleminde, giriş katmanı genellikle RGB renk kanallarına sahip görüntüleri kabul eder. Girdi verisi, genellikle üç boyutlu bir matris (yükseklik x genişlik x kanal sayısı) olarak temsil edilir (Goodfellow, vd. 2016). Giriş katmanı, modelin geri kalan katmanlarına anlamlı bir biçimde veri sağlayabilmesi için girdilerin boyutunu ve yapısını tanımlar. Bu katman, genellikle model mimarisi tasarlanırken veri

tipine göre özelleştirilir. Örneğin, gri tonlu görüntülerde giriş boyutu iki boyutlu (yükseklik x genişlik) bir matris olurken, RGB renkli görüntülerde üç boyutlu bir tensör yapısı kullanılır. Ayrıca, ön işleme aşamasında giriş verileri genellikle normalize edilerek (0-1 aralığına ya da -1 ile 1 arasına) ölçeklendirilir. Bu işlem, modelin daha hızlı ve kararlı bir şekilde öğrenmesine yardımcı olur (LeCun, Bottou, Orr, & Müller, 2012). Giriş katmanı aynı zamanda modelin aldığı veri boyutlarının sabitlenmesini sağlar. Eğitim sırasında tüm görüntülerin aynı boyutta olması gerektiğinden, farklı boyutlardaki resimler çoğu zaman yeniden boyutlandırılarak modele verilir. Böylece evrişimli katmanlar (convolutional layers) üzerinde homojen bir veri akışı sağlanır. Özellikle derin öğrenme modellerinde, giriş katmanı sonradan gelen katmanların parametre sayısını ve hesaplama gereksinimini doğrudan etkileyerek modelin genel karmaşıklığını belirler (Dumoulin & Visin, 2016).

Evrişim katmanları (Convolutional layers): Evrişim katmanları, giriş verileri üzerinde küçük boyutlu filtreler (kernel) kullanarak özellik haritaları (feature maps) üretir. Bu filtreler, görüntünün belirli bölgelerini tarayarak düşük seviye özellikleri (örneğin kenarlar, köşeler, dokular) otomatik olarak öğrenir. Evrişim işlemi, filtre ağırlıkları ile giriş verisinin yerel pencerelerde çarpılması ve sonuçların toplanması yoluyla gerçekleştirilir (Fukushima, 1980). Bu işlem sırasında filtre ağırlıkları öğrenilebilir parametrelerdir ve modelin eğitim sürecinde güncellenerek veri setine uygun en iyi özellik çıkarımını sağlar.

Evrişim katmanlarının çıktısı olan özellik haritaları, giriş görüntüsünün uzamsal boyutlarını (genişlik ve yükseklik) genellikle stride ve padding ayarlarına bağlı olarak azaltır ve daha derin (çok kanallı) temsillere dönüştürür. Bu sayede model, daha karmaşık ve soyut özellikleri hiyerarşik olarak öğrenebilir. Evrişim katmanlarının temel matematiksel formülü genel olarak şu şekilde ifade edilir:

$$Z = f(X * W + b) \quad (3.1)$$

Burada Z , özellik haritasını, X giriş verisini, W filtre ağırlıklarını, b bias terimini ve f aktivasyon fonksiyonunu temsil eder. Filtreler, giriş verisi üzerinde kaydırılarak yerel bölgelerden görüntüler çıkarır. Tek bir çıktı birimi için daha ayrıntılı olarak şu şekilde hesaplanır:

$$Z_{i,j} = f\left(\sum_{m=1}^M \sum_{n=1}^N X_{(i+m),(j+n)} \cdot W_{m,n} + b\right) \quad (3.2)$$

Çok kanallı girişler (ör. RGB görüntüler) için formül şu şekilde genişletilir:

$$Z_{i,j} = f\left(\sum_{c=1}^C \sum_{m=1}^M \sum_{n=1}^N X_{(i+m),(j+n)} \cdot W_{c,m,n} + b\right) \quad (3.3)$$



Şekil 3.6. Evrişim işlemi

Yukarıdaki şekilden anlaşıldığı gibi gösterilen örnekte, 5x5 boyutunda bir görüntü matrisi üzerinde 3x3 boyutlu bir filtre uygulanmaktadır. Filtre, belirli bir bölgedeki piksel değerleriyle çarpılarak ve toplanarak tek bir çıktı değeri üretir. Bu işlem, giriş görüntüsünün tüm bölgelerinde aynı ağırlıklar ile kaydırılarak tekrarlanır. Böylece, giriş görüntüsünün boyutu küçülürken özellik bilgisi korunur ve yeni bir özellik haritası (feature map) elde edilir. Özellik haritası, orijinal görüntünün yerel örüntülerinin (kenarlar, köşeler, dokular gibi) temsili bir formudur.

Şekil 3.6’da görüldüğü gibi ilk kaydırma konumundaki 3x3 pencere seçilmiş ve filtre ile element bazlı çarpım-toplam işlemi yapılmıştır. Görüntünün sol üst bölgesinden alınan 3x3 blok şu değerlerden oluşmaktadır: [1, 2, 0], [1, -1, 2], [1, 2, 3]. Filtre matrisi ise [1, 0, 2], [2, 1, 3], [2, 0, 1] şeklindedir. Evrişim işlemi sırasında, karşılık gelen elemanlar çarpılır ve tüm sonuçlar toplanır ve özellik haritasındaki ilgili konumdaki sonuç ortaya çıkar. Bu sonuç, özellik haritasının ilgili konumundaki değeri oluşturur. Filtre, tüm görüntü üzerinde belirlenen stride ayarına göre kaydırılarak işlem tekrarlandığında tam bir özellik haritası elde edilir. Özellik haritası, giriş görüntüsünün belirli örüntülerini (kenarlar, köşeler, dokular) öne çıkaran ve daha küçük boyutlu, özetlenmiş bir temsildir.

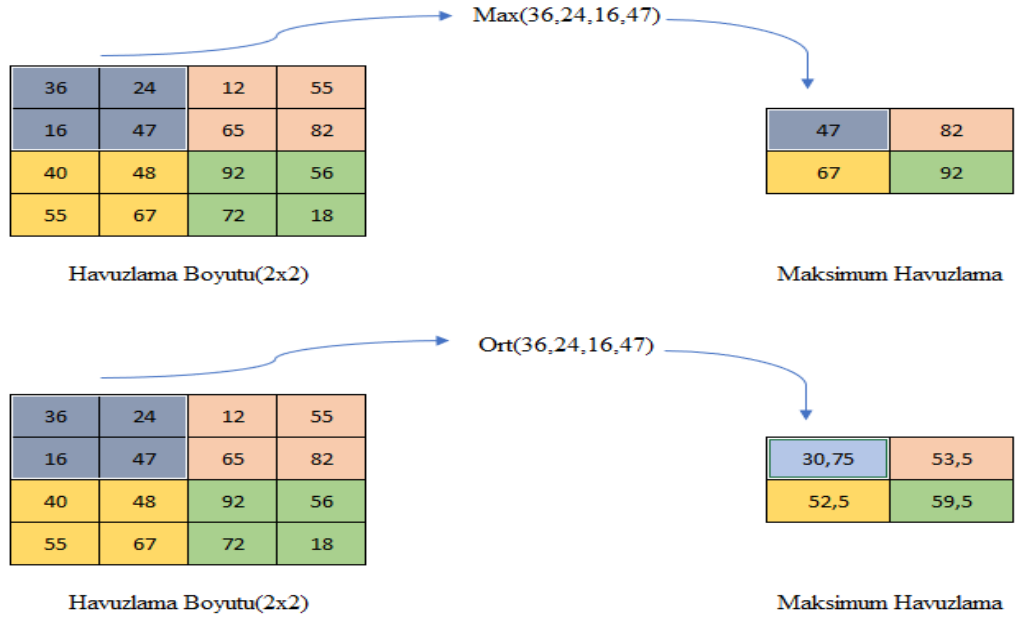
Evrişim katmanlarının önemli bir özelliği parametre paylaşımıdır: aynı filtre ağırlıkları tüm görüntüde kullanılır. Bu yaklaşım modelin parametre sayısını azaltır ve genelleme yeteneğini artırır. Eğitim sürecinde bu filtre ağırlıkları öğrenilerek, model veri setine uygun örüntüleri algılamayı optimize eder.

Sonuç olarak, evriřim iřlemi derin öğrenme modellerinde yerel özelliklerin çıkarılmasını sağlar ve görüntülerin daha soyut, yüksek seviyeli temsillerini üretir. Bu süreç, sınıflandırma, nesne tanıma ve segmentasyon gibi bilgisayarla görü görevlerinde başarının temel taşlarından biridir.

Havuzlama katmanları (Pooling layers): Havuzlama katmanları, evriřimli sinir ağı mimarilerinin temel bileřenlerinden biri olarak, uzamsal boyutları küçültmek, hesaplama yükünü azaltmak ve modelin aşırı öğrenmesini önlemek amacıyla kullanılır. Özellik haritalarının boyutlarını azaltırken önemli bilgileri korumayı hedefler. Bu katmanlar, giriş özellik haritaları üzerinde belirli bir pencere (örneğin 2x2 veya 3x3) kaydırarak özet bir temsil çıkarır.

En yaygın kullanılan havuzlama yöntemi max pooling'dir. Bu yöntemde, pencere içinde yer alan tüm değerler incelenir ve maksimum değer seçilerek çıktı özellik haritasına yazılır. Örneğin, 2x2 boyutunda bir pencere kullanıldığında, her 4 piksel değerinden yalnızca en yüksek olanı çıkışta yer alır. Bu yaklaşım, kenar veya doku gibi baskın özellikleri korurken detaylı konum bilgisinin bir kısmını atar, bu da modelin çevresel kaymalara karşı daha dayanıklı olmasını sağlar (Riesenhuber ve Poggio, 1999).

Max pooling iřlemi, özellik haritasının genişlik ve yükseklik boyutlarını genellikle stride parametresine baėlı olarak yarıya indirir. Örneğin, 2x2 pencere ve 2 birimlik stride kullanıldığında, çıkış haritası orijinal boyutun dörtte biri olur. Bu boyut küçültme, modelin parametre sayısını azaltır ve eğitim süresini kısaltır. Ayrıca, havuzlama katmanları modelin yerel örüntülere karşı esnekliėini artırarak, farklı konumlarda benzer özellikleri tanıyabilmesini kolaylaştırır. Bunun yanı sıra average pooling adı verilen başka bir havuzlama türü de vardır. Average pooling, pencere içindeki değerlerin ortalamasını alarak boyut küçültür. Bu yaklaşım, daha yumuřak ve genelleřtirici bir temsil sağlar ancak max pooling kadar belirgin kenar veya köőe bilgilerini korumayabilir. Bazı modern ESA mimarilerinde ise global average pooling kullanılarak tüm uzamsal boyutlar ortalama bir değere indirgenir ve sınıflandırma katmanına bu özet özellik vektörü aktarılır.



Şekil 3.7. Maksimum ve ortalama havuzlama İşlemleri

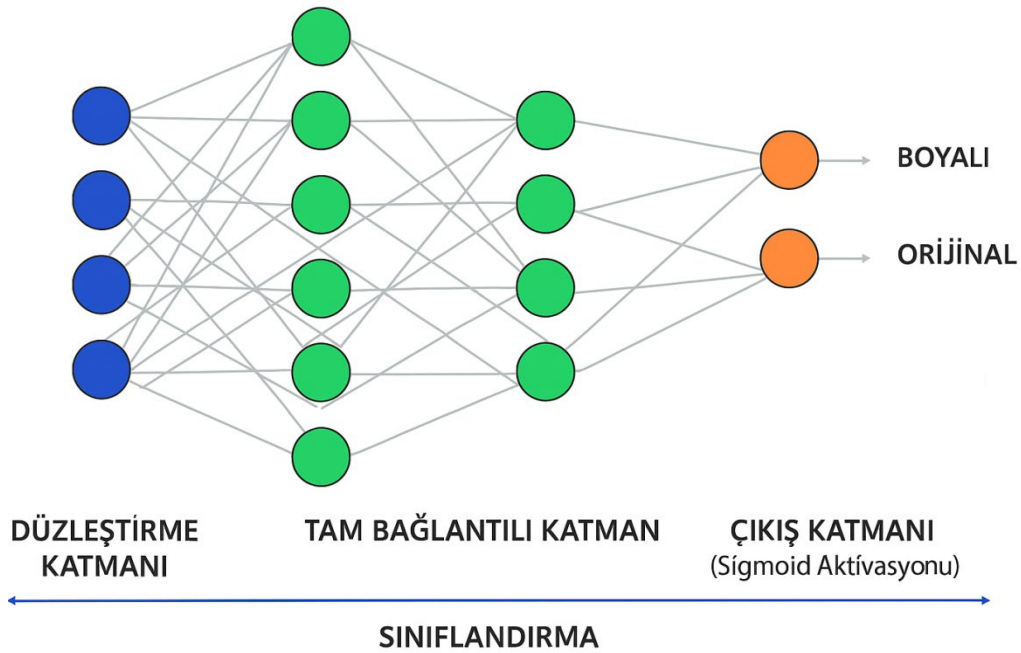
Havuzlama katmanları, evrişim katmanlarıyla birlikte kullanılarak ağın derinliğini artırırken, parametre yükünü kontrol altında tutar. Katmanlar arasında bilgi aktarımını daha soyut ve genel hale getirerek yüksek seviyeli özelliklerin çıkarılmasına olanak tanır. Sonuç olarak, havuzlama işlemi, ESA mimarilerinin hem verimli hem de güçlü bir şekilde öğrenmesini sağlayan temel adımlardan biridir.

Tam bağlantılı katmanlar (Fully connected layers): Tam bağlantılı katmanlar, klasik yapay sinir ağı (ANN) katmanlarıdır. Bu katman, daha önceki evrişimsel ve havuzlama katmanlarında çıkarılan özellikleri tek bir düzleştirilmiş vektör halinde alır. Özellik haritasının düzleştirilmesi (flatten işlemi), çok boyutlu verilerin tek boyutlu bir vektöre dönüştürülmesini sağlar ve modelin tüm önemli bilgileri yoğun bir temsil halinde kullanabilmesine imkân tanır. Tam bağlantılı katmanda her bir nöron, bir önceki katmandaki tüm nöronlardan giriş alır. Bu yapı sayesinde model, tüm çıkarılmış özellikleri aynı anda değerlendirme yeteneğine sahip olur. Katman içindeki ağırlıklar ve bias terimleri, eğitim sürecinde güncellenerek, giriş özellikleri ile doğru sınıf arasında en iyi ilişkiyi kurmayı öğrenir. Böylece, modelin daha önce öğrendiği kenar, renk, doku gibi düşük seviyeli özellikler, daha soyut ve yüksek seviyeli temsil haline dönüştürülür.

Model mimarisinde kullanılan tam bağlantılı katmanda ReLU (Rectified Linear Unit) aktivasyon fonksiyonu tercih edilmiştir. ReLU fonksiyonu, doğrusal olmayan yapıyla negatif değerleri sıfırlayarak modelin karmaşık örüntüleri öğrenmesini sağlar ve gradyan sönmesi problemini azaltır. Bu sayede ağı derinliği arttıkça bile öğrenme kabiliyeti korunur. Tam bağlantılı katmanın ReLU aktivasyonu, çıkarılan özellikler daha anlamlı ve ayırt edilebilir bir forma dönüştürerek sınıflandırma başarısını destekler.

Çıkış katmanında ise sigmoid aktivasyon fonksiyonu kullanılmıştır. Sigmoid fonksiyonu, modelin çıktısını 0 ile 1 arasında normalize ederek iki sınıflı (boyalı ve orijinal) problem için olasılık tahmini yapmasına olanak verir. Bu sayede model, verilen bir araç boya görüntüsünün "boyalı" sınıfına ait olma olasılığını üretir. Sınıflandırma kararı ise belirlenen bir eşik değerine (genellikle 0.5) göre yapılır.

Sonuç olarak, tam bağlantılı katman, evrimsel katmanlarda çıkarılan bilgileri en yüksek düzeyde yorumlayarak sınıflandırma kararının verilmesinde köprü işlevi görür. Kullanılan ReLU aktivasyonu öğrenmenin doğrusal olmayan karmaşık örüntüleri yakalamasını sağlarken, çıkıştaki sigmoid aktivasyonu modelin ikili sınıflandırma problemine uygun, yorumlanabilir ve pratik bir olasılık tahmini üretmesine katkıda bulunur. Şekil 3.8'de sınıflandırmamda kullandığım temsili tam bağlantılı katmana ait görsel bulunmaktadır.



Şekil 3.8. Tam bağlantılı katman

Aktivasyon fonksiyonları: Derin öğrenme modellerinde, özellikle Evrimsel Sinir Ağları mimarilerinde aktivasyon fonksiyonları, modelin karmaşık ve doğrusal olmayan ilişkileri temsil etmesini sağlayan temel bileşenlerdir. Bu fonksiyonlar, her katmandaki doğrusal dönüşümler arasına eklenerek modele doğrusal olmayanlık kazandırır ve sinir ağının sıradan bir doğrusal regresyon sisteminden çok daha güçlü bir yaklaşımlayıcı haline gelmesini sağlar. Aktivasyon fonksiyonlarının başlıca görevi, ağdaki nöronların aldığı girdileri dönüştürerek hangi bilgilerin iletileceğini ve hangi bilgilerin bastırılacağını belirlemektir. Bu işlem, ağın öğrenebileceği örüntülerin karmaşıklığını artırır ve farklı problem tiplerine uygun çözümler üretmesine imkan tanır. Özellikle görüntü işleme, dil modelleme ve ses tanıma gibi karmaşık görevlerde doğru aktivasyon fonksiyonunun seçilmesi, modelin başarısı üzerinde doğrudan etkili olur (Nwankpa vd., 2018). Aşağıda, ESA mimarilerinde sıkça kullanılan başlıca aktivasyon fonksiyonları ve özellikleri özetlenmiştir:

Sigmoid Fonksiyonu (Logistic Sigmoid): Sigmoid fonksiyonu, giriş değerlerini 0 ile 1 arasında bir aralığa dönüştürerek çıktıyı olasılık yorumu yapılabilecek bir değere çevirir. Bu özelliği nedeniyle özellikle ikili sınıflandırma problemlerinde çıkış katmanında yaygın olarak tercih edilir. Ancak büyük pozitif ya da negatif girişlerde gradyan değerinin çok küçülmesi, vanishing gradient problemine yol açabilir (Han & Moraga, 1995).

Softmax Fonksiyonu: Çok sınıflı sınıflandırma problemlerinde yaygın olarak kullanılan Softmax fonksiyonu, tüm sınıf skorlarını normalize ederek bir olasılık dağılımına dönüştürür. Böylece her sınıfa ait çıktı değeri, toplamda 1 olacak şekilde normalize edilir. Bu, modelin hangi sınıfa daha yüksek olasılık atadığını yorumlamayı kolaylaştırır).

ReLU (Rectified Linear Unit): ReLU fonksiyonu modern derin öğrenme mimarilerinin temel yapı taşlarından biridir. Negatif girişleri sıfıra eşitlerken, pozitif girişleri olduğu gibi geçirir. Bu basit ama etkili yaklaşım, modelin hesaplama maliyetini azaltır ve gradyan seyrelmesi sorununu önemli ölçüde hafifletir. Ancak tüm negatif değerlerin sıfırlanması, bazı nöronların tamamen etkinliğini kaybetmesine (dying ReLU) yol açabilir (Nair & Hinton, 2010).

Tanh (Hyperbolic Tangent): Tanh fonksiyonu sigmoid'e benzer S-şeklinde bir eğriye sahiptir ancak çıktıları -1 ile +1 arasında değişir. Bu merkezlenmiş yapı, öğrenme sürecinde ağırlıkların daha dengeli güncellenmesini sağlar. Genellikle sigmoid'e göre

daha iyi performans gösterir çünkü çıkışın ortalaması sifira daha yakındır (LeCun vd., 2012).

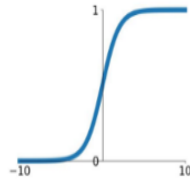
ELU (Exponential Linear Unit): ELU fonksiyonu, ReLU'nun bir varyantı olarak tasarlanmıştır ve negatif değerlerde sıfır olmayan, yumuşak bir çıkış sağlar. Negatif taraftaki bu eğim, ağırlıkların sifira yakınsamasını hızlandırabilir ve modelin genelleme yeteneğini artırabilir. ELU'nun bu özelliği, gradyan akışını daha dengeli tutar ve öğrenme sürecini iyileştirir (Clevert vd., 2016).

Leaky ReLU: Leaky ReLU, ReLU'nun “ölü nöron” problemini çözmek için geliştirilmiştir. Negatif girişler için küçük ama sabit bir eğim sağlar. Bu sayede gradyanlar sıfırlanmaz ve öğrenme sırasında tüm nöronlar etkin kalabilir. Bu özelliği sayesinde özellikle derin ağlarda daha istikrarlı bir öğrenme sağlar (Maas vd., 2013).

MaxOut: MaxOut aktivasyonu, girişteki bir grup nöron arasında maksimum değeri seçerek çalışır. Bu yaklaşım, farklı doğrusal bölgelere uyum sağlayarak modelin daha karmaşık ilişkileri öğrenmesine olanak tanır. Parametre sayısını artırsa da temsil kapasitesini de önemli ölçüde yükseltir (Goodfellow vd., 2013). Aktivasyon fonksiyonlarına ait grafikler Şekil 3.9'da verilmiştir.

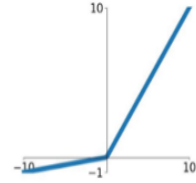
Sigmoid

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



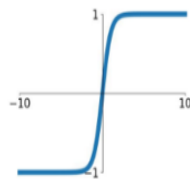
Leaky ReLU

$$\max(0.1x, x)$$



tanh

$$\tanh(x)$$

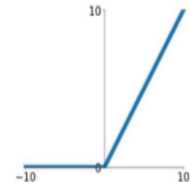


Maxout

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

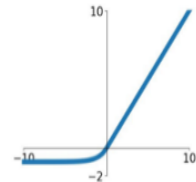
ReLU

$$\max(0, x)$$



ELU

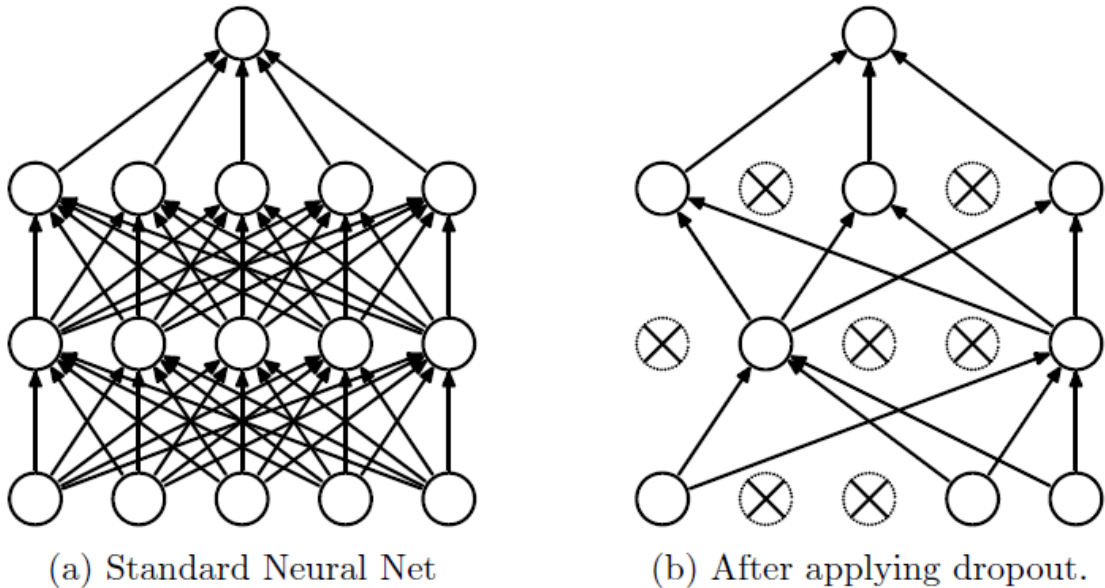
$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$



Şekil 3.9. Aktivasyon fonksiyonları (Kalyoncu, 2020)

Dropout katmanı: Derin öğrenme modellerinin eğitimi sırasında karşılaşılan en önemli sorunlardan biri aşırı öğrenme problemidir. Aşırı öğrenme, modelin eğitim verisini ezberlemesi ve dolayısıyla yeni, görülmemiş veriler üzerinde düşük genelleme performansı göstermesiyle sonuçlanır. Bu sorunun çözümü için geliştirilen yöntemlerden biri dropout katmanıdır.

Dropout, Hinton ve arkadaşları tarafından 2012 yılında önerilen bir düzenleme tekniğidir (Hinton vd., 2012). Bu yöntemde, eğitim sırasında modelin belirli bir oranla (örneğin %20–%50) rastgele seçilen nöronlarının aktivasyonu sıfırlanır. Bu işlem her eğitim adımında farklı bir alt ağ (subnetwork) oluşturulmasına olanak tanır. Böylece model, özelliklerin tümüne bağımlı hale gelmek yerine daha genelleştirilebilir temsiller öğrenir. Dropout katmanının uygulanması sırasında, model eğitimi esnasında belirlenen dropout oranına göre düğümler geçici olarak devre dışı bırakılır. Ancak model test aşamasında tüm nöronlar aktif hale getirilir ve çıkışlar, eğitim sırasında yapılan dropout'un telafisi için ölçeklendirilir. Bu yaklaşım, modelin karmaşıklığını azaltmadan aşırı uyumu sınırlandırır. Bir başka avantajı ise çoğul (ensemble) öğrenmeye benzer bir etki yaratmasıdır. Dropout, tek bir büyük model yerine çok sayıda daha küçük ve farklı modelin birleşik bir versiyonunu eğitmiş gibi davranır.

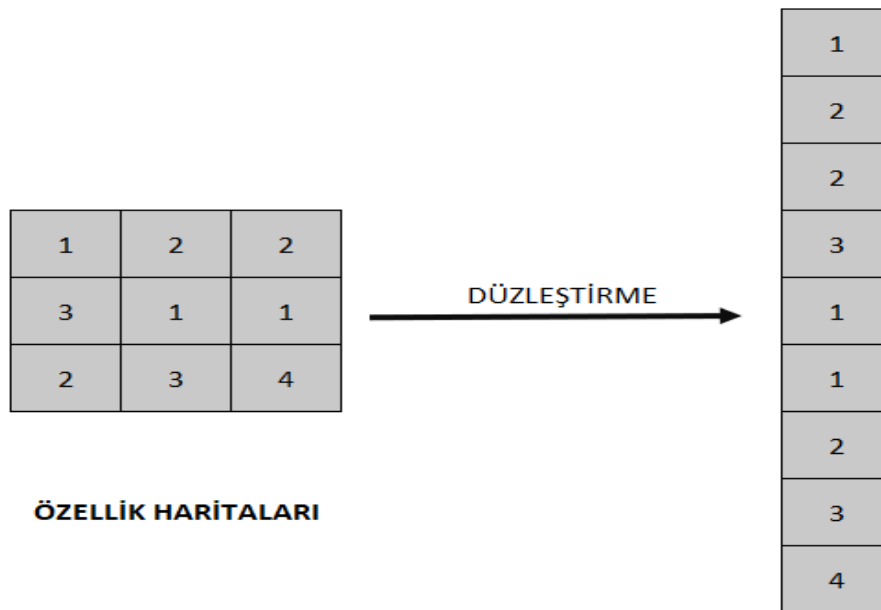


Şekil 3.10. Standart sinir ağı ve dropout uygulanan sinir ağı (Hinton vd. 2012)

Şekil 3.10’da standart bir sinir ağının dropout uygulandığında bazı sinir düğümlerinin rastgele sıfırlandığı görülmektedir. Sonuç olarak, model daha sağlam ve yeni veri üzerinde daha başarılı hale gelir. Dropout oranı hiperparametre olarak belirlenir ve genellikle 0.2 ile 0.5 arasında seçilir. Daha yüksek oranlar daha agresif bir düzenleme sağlar ancak aşırı yüksek değerler modelin öğrenmesini zorlaştırabilir.

Düzleştirme (Flatten) katmanı: Düzleştirme (Flatten) katmanı, derin öğrenme mimarilerinde yer alan ve evrimsel sinir ağlarının önemli bir bileşeni olarak öne çıkan bir yapıdır. Bu katmanın temel işlevi, evrişim ve havuzlama katmanlarından elde edilen çok boyutlu özellik haritalarını (feature maps) tek boyutlu bir vektör biçimine dönüştürmektir. Böylece, ağın sınıflandırma aşamasında kullanılan tam bağlantılı (fully connected) katmanlara uygun bir giriş yapısı sağlanır.

Evrışim ve havuzlama katmanları genellikle uzamsal (spatial) boyutları ve kanal bilgilerini içeren üç boyutlu tensör çıktıları üretir. Ancak tam bağlantılı katmanlar, bu veriyi yalnızca tek boyutlu vektör formatında işleyebilir. Düzleştirme katmanı bu dönüşümü gerçekleştirerek modelin son katmanlarındaki sınıflandırma işlemlerinin uygulanabilmesini mümkün kılar. Düzleştirme katmanı parametre içermeyen bir yapıdadır ve yalnızca tensörün şekil dönüşümünü gerçekleştirir; bu süreçte veri içeriğinde herhangi bir kayıp veya değişiklik oluşmaz. Modelin hesaplama yüküne anlamlı bir etkisi bulunmaz.



Şekil 3.11. Düzleştirme katmanı

Öğrenme süreci: Evrişimli Sinir Ağları, görüntü verilerinden otomatik öznitelik çıkarımı ve sınıflandırma amacıyla tasarlanmış derin öğrenme mimarileridir. Öğrenme süreci, konvolüsyon katmanlarının giriş verisi üzerinde öğrenilebilir filtreler aracılığıyla yerel örüntüleri ve temel görsel özellikleri yakalamasıyla başlar. Bu katmanlarda kullanılan aktivasyon fonksiyonları, modelin doğrusal olmayan ilişkileri temsil etmesine olanak tanıyarak karmaşık örüntülerin öğrenilmesini mümkün kılar.

Havuzlama katmanları, öznitelik haritalarının uzamsal boyutlarını azaltarak modelin hesaplama maliyetini düşürür ve konumsal genelleme yeteneğini artırır. Bu katmanlar, girişteki bilgiyi daha kompakt ve özet bir temsil haline getirirken, önemli öznitelikleri korumaya odaklanır.

Düzleştirme (Flatten) katmanı, çok boyutlu öznitelik haritalarını tek boyutlu bir vektör yapısına dönüştürerek tam bağlantılı katmanlara uygun bir giriş oluşturur. Tam bağlantılı katmanlar, bu öznitelikleri bütüncül bir biçimde değerlendirerek sınıflandırma kararını üretir. Ayrıca dropout katmanları, eğitim sırasında nöronların belirli oranlarda rastgele devre dışı bırakılması yoluyla aşırı öğrenmeyi (overfitting) azaltır ve modelin genelleme yeteneğini güçlendirir.

Öğrenme süreci, ileri besleme aşamasıyla başlar; burada giriş verisi katmanlar boyunca ardışık dönüşümlerden geçirilerek modelin çıktısı hesaplanır. Üretilen çıktı, gerçek etiketle karşılaştırılarak bir kayıp fonksiyonu aracılığıyla hata değeri elde edilir. Ardından, geriye yayılım (backpropagation) algoritması kullanılarak bu hata değeri ağırlıkların güncellenmesi için ağ boyunca geriye doğru iletilir. Optimizasyon algoritmaları yardımıyla ağ parametreleri adım adım güncellenerek modelin hatası minimize edilir. Bu süreç sonucunda ESA, karmaşık görsel örüntüleri öğrenerek yeni ve görülmemiş veriler üzerinde yüksek doğrulukta tahmin yapma yeteneği kazanır.

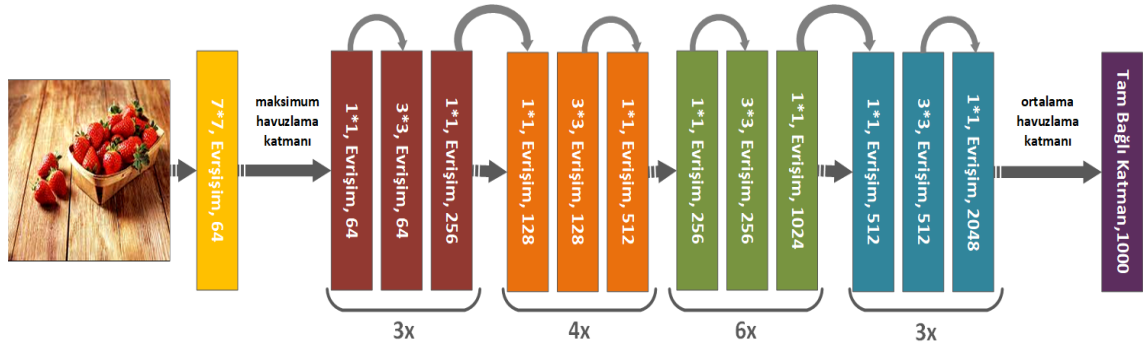
3.2.2.1. ResNet50

ResNet50, Kaiming He ve çalışma arkadaşları tarafından geliştirilen ve derin öğrenme mimarilerinde artık bağlantılar (residual connections) kavramını sistemli biçimde kullanan bir evrişimsel sinir ağı modelidir. Artık bağlantılar, bir evrişim katmanının çıktısını doğrudan sonraki katmanın girdisine ekleyerek iletir. Bu yapı, derin ağlarda karşılaşılan gradyan sönmesi problemini azaltır, eğitim sürecini daha kararlı hale getirir ve modelin karmaşık öznitelikleri daha etkili biçimde öğrenmesini sağlar.

ResNet50 mimarisi beş ana evreden oluşur. İlk evrede 7x7 boyutunda ve 64 filtrelili bir evrişim katmanı bulunur ve bunu 2 adım uzunluğunda 3x3 boyutlu bir maksimum havuzlama katmanı takip eder. İkinci evrede, üçlü tekrarlardan oluşan bloklar yer alır: her blokta sırasıyla 64 kanallı 1x1 evrişim, 64 kanallı 3x3 evrişim ve 256 kanallı 1x1 evrişim katmanları bulunur.

Üçüncü evrede 128 kanallı 1x1 evrişim, ardından 128 kanallı 3x3 evrişim ve 512 kanallı 1x1 evrişim katmanları ardışık şekilde yer alır ve bu yapı belirli sayıda tekrarlanır. Dördüncü evrede ise 256 kanallı 1x1 evrişim, 256 kanallı 3x3 evrişim ve 1024 kanallı 1x1 evrişim katmanlarından oluşan bloklar ardışık olarak uygulanır. Beşinci ve son evrede ise 512 kanallı 1x1, 512 kanallı 3x3 ve 2048 kanallı 1x1 evrişim katmanlarından oluşan yapı bulunur. Tüm bu evrişim aşamalarının ardından bir global ortalama havuzlama katmanı uygulanır ve sınıflandırma işlemi için tam bağlantılı katman yer alır.

Ayrıca, ResNet50 mimarisi transfer öğrenme yaklaşımlarında yaygın olarak tercih edilmektedir. Artık bağlantılar sayesinde önceden eğitilmiş ağların farklı veri kümelerine uyarlanabilmesi ve sınırlı etiketli veriyle bile yüksek doğruluk sağlanması mümkündür. Örneğin, Reddy ve Juliet (2019), ResNet50'yi transfer öğrenme yöntemiyle sıtma hücre görüntülerinin sınıflandırılmasında başarıyla kullanarak bu mimarinin farklı görsel sınıflandırma problemlerinde etkinliğini göstermiştir.



Şekil 3.12. ResNet50 mimarisi (Habek vd. 2024)

Şekil 3.12’de ResNet50 mimarisi katman yapısıyla birlikte gösterilmektedir. Farklı aşamalarda yer alan 1x1 ve 3x3 evrişim katmanları, artık bağlantılarla birleştirilerek özellik haritalarının derinlemesine öğrenilmesine olanak tanırken, maksimum ve ortalama havuzlama katmanları uzamsal boyutları azaltarak modelin parametre verimliliğini artırmaktadır (Rojas-Aranda vd. 2020).

3.2.2.2. VGG16

VGG16, Oxford Üniversitesi Görsel Geometri Grubu (Visual Geometry Group) tarafından Simonyan ve Zisserman (2014) öncülüğünde geliştirilmiş derin bir evrişimsel sinir ağı mimarisidir. ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2014 yarışmasında öne çıkan bu mimari, özellikle sadeliği ve sistematik tasarımıyla derin öğrenme alanında önemli bir kilometre taşı olarak kabul edilir. VGG16'nın en belirgin özelliği, tüm evrişim katmanlarında sabit boyutlu (3x3) küçük çekirdekler (filtreler) kullanarak ağı derinleştirmesidir. Bu yaklaşım, daha büyük alana sahip alıcı bölgeler (receptive fields) oluşturarak görüntüdeki karmaşık örüntüleri daha iyi modelleme imkânı sağlar (Simonyan & Zisserman, 2014).

VGG16 mimarisi, toplamda 16 ağırlıklı katmandan (13 evrişim + 3 tam bağlantılı) oluşur. Modelin tasarımında 3x3 boyutlu evrişim filtreleri ve 2x2 boyutlu maksimum havuzlama katmanları ardışık olarak yerleştirilmiştir. Küçük çekirdek boyutları sayesinde model, parametre sayısını aşırı artırmadan derinliği artırabilir ve dolayısıyla daha karmaşık ve çok katmanlı örüntüleri öğrenebilir. Örneğin, ardışık iki adet 3x3 evrişim katmanı, 5x5 bir filtreye eşdeğer bir alıcı alan sağlar ancak daha az parametreyle bu hesaplamayı yapar. Bu sistematik yapı, modelin anlaşılabilirliğini ve uygulanabilirliğini de kolaylaştırır.

Mimari detaylara bakıldığında, VGG16 mimarisi beş evrişimsel bloktan oluşur. İlk iki blokta 64 ve 128 filtreli 3x3 evrişim katmanları yer alırken, üçüncü blokta 256 filtreli katmanlar bulunur. Dördüncü ve beşinci bloklarda ise 512 filtreli evrişim katmanları ardışık olarak uygulanır. Her blok sonunda 2x2 boyutlu maksimum havuzlama katmanı ile uzamsal boyutlar yarıya indirilir. Evrişim bloklarının ardından gelen katmanlar, tam bağlantılı yapıda olup, iki adet 4096 nöronlu katman ve bir adet 1000 sınıflı softmax çıkış katmanından oluşur. Bu yapı ImageNet veri setindeki 1000 sınıflı sınıflandırmak için tasarlanmıştır (Simonyan & Zisserman, 2014).

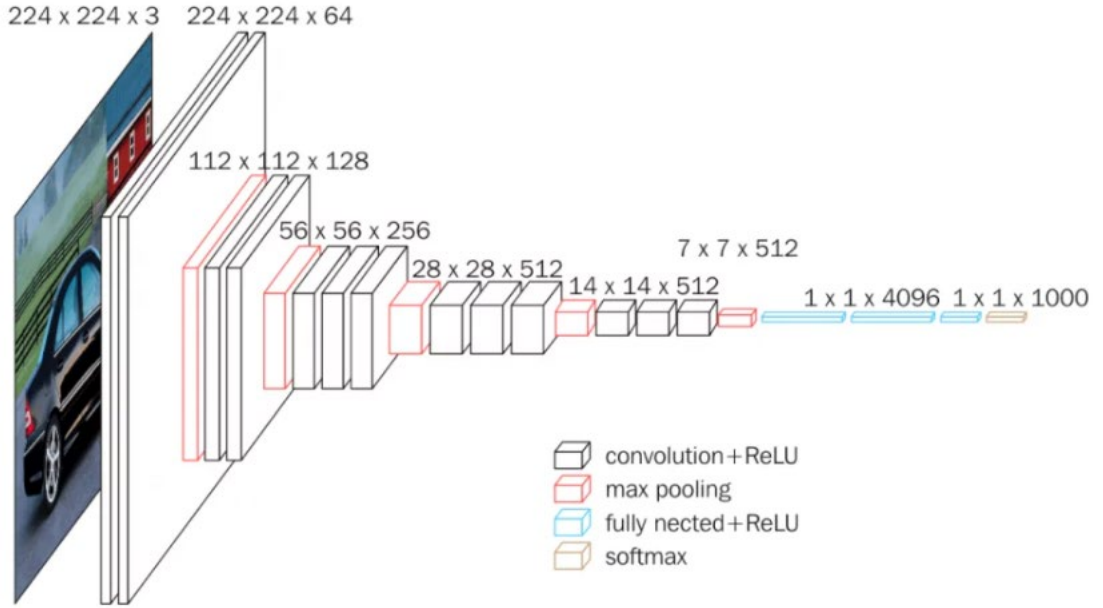
VGG16 mimarisinin bir diğer önemli özelliği, tüm katmanlarda ReLU aktivasyon fonksiyonunun kullanılmasıdır. ReLU, negatif değerleri sıfıra eşitleyerek gradyan sönmesi problemini azaltır ve modelin daha hızlı ve kararlı bir şekilde öğrenmesini sağlar. Ayrıca, modelin eğitiminde dropout gibi düzenleme teknikleri kullanılarak aşırı öğrenme riski azaltılmıştır. Eğitim sırasında 224x224 çözünürlükte RGB görüntüler kullanılmış ve giriş görüntüleri normalize edilmiştir.

Teknik açıdan VGG16 tasarımının en önemli avantajlarından biri, tasarım sadeliği ve modüler yapısıdır. Tüm evrişim katmanlarının aynı boyuttaki filtrelerle tasarlanması, mimarinin farklı veri kümelerine ve problem türlerine kolayca uyarlanmasına olanak sağlar. Bu modüler yapı sayesinde araştırmacılar, modelin derinliğini veya genişliğini değiştirerek yeni varyantlar geliştirebilir. Örneğin, VGG19 modeli, VGG16'ya benzer yapıyı daha da derinleştirerek 19 ağırlıklı katmana sahiptir.

VGG16'nın başarısı, transfer öğrenme uygulamalarında da etkili olmasına yol açmıştır. ImageNet üzerinde önceden eğitilmiş VGG16 ağırlıkları, sınırlı veri içeren farklı görevlerde yeniden kullanılabilir. Özellikle medikal görüntü analizi, nesne tespiti, yüz tanıma ve endüstriyel kalite kontrol gibi alanlarda VGG16 tabanlı transfer öğrenme yöntemleri yaygınca uygulanmıştır (Shin vd., 2016). Bu yaklaşımda, önceden eğitilmiş konvolüsyon katmanları sabit tutulabilir veya ince ayar yapılabilir. Böylece, veri kümesi küçük olsa bile yüksek doğruluklu sınıflandırma performansı elde edilebilir.

Bir diğer önemli unsur, VGG16'nın parametre sayısının yüksek olmasıdır. Yaklaşık 138 milyon parametre içeren bu mimari, özellikle belleği sınırlı cihazlarda doğrudan uygulanması zor olan bir yapıdadır. Bu nedenle mobil ve gömülü sistemler için daha hafif mimari varyantları (ör. MobileNet, EfficientNet) geliştirilmiştir. Bununla birlikte, VGG16'nın eğitilebilirliği, modülerliği ve transfer öğrenme performansı sayesinde derin öğrenme literatüründe en çok referans verilen modellerden biri olmaya devam etmektedir (Chollet, 2017).

Sonuç olarak, VGG16, küçük çekirdek boyutlarının ardışık kullanımına dayalı basit ama derin bir mimari tasarım ortaya koyar. Bu yaklaşım, daha karmaşık örüntüleri katmanlı yapılar içinde hiyerarşik olarak öğrenmeye imkân tanır. Sadeliği, güçlü transfer öğrenme yetenekleri ve modüler tasarımı sayesinde VGG16, günümüzde bilgisayarla görü ve görüntü sınıflandırma alanında temel yapı taşlarından biri olarak kabul edilmektedir. Şekil 3.13'te VGG16 model mimarisi görülmektedir.



Şekil 3.13. VGG16 mimarisi katman blok diyagramı (Neurohive, 2018)

3.2.2.3. InceptionV3

InceptionV3, Szegedy ve çalışma arkadaşları (2016) tarafından geliştirilen ve GoogleNet mimarisi üzerine inşa edilen gelişmiş bir derin evrişimsel sinir ağı modelidir. Inception mimarileri, geleneksel evrişimsel tasarımlara göre farklı ölçeklerdeki filtrelerin aynı katmanda paralel olarak uygulanması prensibine dayanır. Bu tasarım, çok ölçekli özellik çıkarımı sağlayarak modelin hem düşük seviyeli hem de yüksek seviyeli örüntüleri aynı anda yakalayabilmesine imkân tanır.

Inception mimarisinin temel fikri, farklı boyutlardaki evrişim filtrelerini (ör. 1x1, 3x3, 5x5) aynı seviyede uygulayarak ve ardından bu çıktıları kanal eksenine boyunca birleştirerek çok ölçekli bir temsile ulaşmaktır. Böylece model, farklı büyüklükteki alıcı bölgelerden (receptive fields) gelen bilgileri bütünleştirir ve daha zengin bir özellik haritası üretir (Szegedy vd., 2015).

InceptionV3, önceki Inception sürümlerine göre önemli geliştirmeler içerir. En önemli yeniliklerden biri faktörize evrişimler kavramıdır. Büyük boyutlu evrişimler daha küçük boyutlara ayrılarak (örneğin bir 5x5 evrişim → iki ardışık 3x3 evrişime bölünerek) hesaplama maliyeti düşürülür. Ayrıca, asimetrik evrişimler kullanılarak örneğin 3x3 evrişimler 1x3 ve 3x1 şeklinde ayrıştırılır. Bu yaklaşım parametre sayısını azaltırken öğrenme kapasitesini korur ve modelin daha derin katmanlar içermesine olanak sağlar (Szegedy vd., 2016). Bunun yanı sıra Auxiliary Classifiers adı verilen yardımcı sınıflandırıcılar, modelin orta seviyelerinde yer alır ve etiket bilgisini erken katmanlara

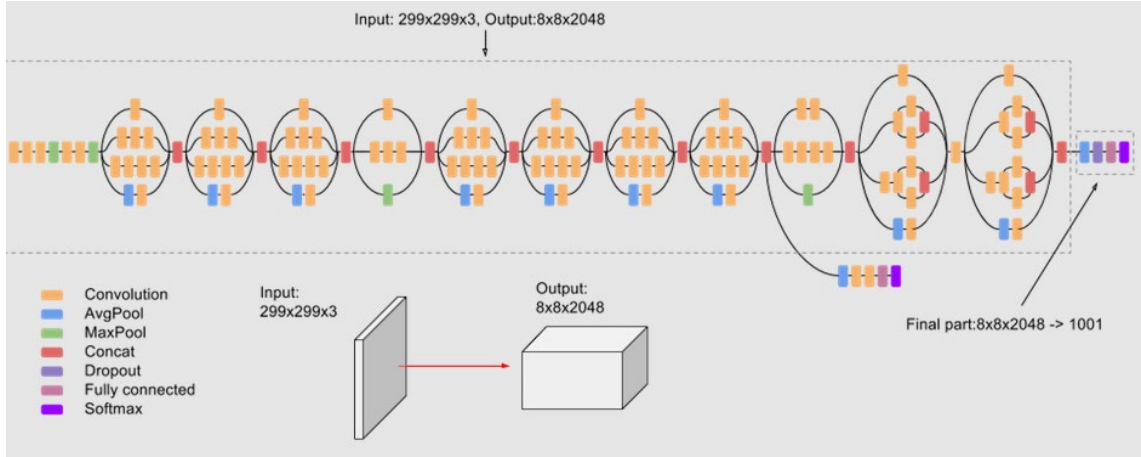
taşıır. Eğitim sırasında bu yardımcı sınıflandırıcılar gradyan akışını iyileştirerek daha istikrarlı bir öğrenme süreci sağlar ve gradyan sönmesi (vanishing gradient) problemini azaltır. Eğitim aşamasında toplam kayıp fonksiyonuna bu yardımcı kayıplar da dahil edilir, ancak çıkarım sırasında yalnızca ana sınıflandırıcı kullanılır.

InceptionV3 mimarisi toplamda yaklaşık 42 katmandan oluşur ve optimize edilmiş inception modülleri içerir. Bu modüller, hesaplama verimliliğini artırmak için 1x1 evrişimler ile boyut indirgeme gerçekleştirir. Böylece, kanal sayısı azaltılarak büyük boyutlu evrişimlerin parametre maliyeti düşürülür. Bu prensip, modelin hem daha derin hem de daha geniş yapıda olmasını sağlar ve bu yapı, farklı veri kümelerinde ve görevlerde yüksek genelleme performansı gösterir.

InceptionV3 modeli, ImageNet veri kümesi üzerinde eğitilmiş önceden eğitilmiş ağırlıklarıyla transfer öğrenme uygulamalarında sıklıkla kullanılmaktadır. Özellikle sınırlı veri içeren problemler için önceden eğitilmiş InceptionV3 ağırlıkları dondurularak veya ince ayar yöntemiyle adapte edilerek yüksek doğruluklu sınıflandırma modelleri geliştirilebilir. Medikal görüntü analizi, nesne tanıma, yüz tanıma ve endüstriyel kalite kontrol gibi birçok alanda bu model başarıyla uygulanmaktadır (Chollet, 2017).

Bir diğer önemli avantaj, Inception mimarisinin modüler yapısıdır. Her inception bloğu bağımsız olarak tasarlanabilir ve farklı evrişim tipleri paralel dallar halinde yerleştirilebilir. Bu modüler tasarım, farklı donanım gereksinimlerine ve uygulama senaryolarına göre özelleştirilebilir. GoogleNet'ten InceptionV3'e evrim süreci boyunca bu modüler yapı daha verimli ve derin tasarımlara dönüşmüştür.

Sonuç olarak, InceptionV3 mimarisi, çok ölçekli özellik çıkarımı, faktörize evrişimler, asimetrik evrişimler ve yardımcı sınıflandırıcılar gibi modern derin öğrenme tekniklerini bir araya getirerek hesaplama verimliliği ile yüksek performansı dengeler. Bu özellikleri sayesinde InceptionV3, günümüzde transfer öğrenme tabanlı bilgisayarla görü uygulamalarında yaygın olarak tercih edilen bir ESA mimarisi haline gelmiştir.



Şekil 3.14. InceptionV3 mimarisi (SB Shetty vd. 2022)

Şekil 3.14'te SB Shetty ve arkadaşlarının (2022) cilt hastalıkları tespitinde kullandığı InceptionV3 model mimarisinin diyagramı gösterilmiştir.

3.2.2.4. EfficientNetB0

EfficientNetB0, Tan ve Le (2019) tarafından önerilen ve evrimsel sinir ağı mimarilerinde ölçeklendirme problemine sistematik bir çözüm sunan EfficientNet ailesinin temel modelidir. Derin öğrenme mimarilerinde daha yüksek doğruluk elde etmek için genellikle ağı daha derin (depth), daha geniş (width) veya daha yüksek çözünürlüklü girişler (resolution) ile ölçeklendirilmesi gerekir. Ancak bu üç ölçeklendirme boyutu geleneksel olarak rastgele veya tek boyutlu olarak uygulanıyordu ve hesaplama maliyetleri çok hızlı büyüyordu (Tan & Le, 2019).

EfficientNet mimarisi, compound scaling adı verilen bir yöntemle bu üç boyutu birlikte ve dengeli bir şekilde ölçeklendirmeyi önerir. Compound scaling, sabit bir kaynak bütçesi altında derinlik, genişlik ve çözünürlük parametrelerini optimize ederek en iyi performansı sağlar. EfficientNet ailesinin tüm modelleri, bu prensibi kullanarak EfficientNetB0 üzerine kuruludur. B0 modeli, otomatik makine öğrenimi (AutoML) arama teknikleri ile optimize edilmiş ve daha büyük versiyonlar için temel yapı taşı oluşturulan referans mimaridir (Tan & Le, 2019).

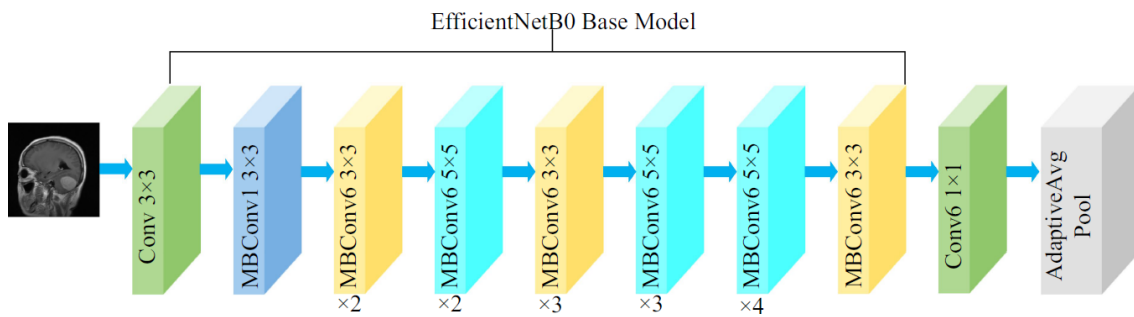
EfficientNetB0 mimarisi, temel olarak Mobile Inverted Bottleneck Convolution (MBConv) bloklarından oluşur. Bu bloklar ilk olarak MobileNetV2 mimarisinde önerilmiş ve EfficientNet'te genişletilmiştir. MBConv blokları, derin ayrık (depthwise separable) evrimsimleri ve sıkıştırma-genişletme (squeeze-and-excitation)

mekanizmalarını içerir. Bu yapı sayesinde parametre sayısı azaltılırken modelin temsil kapasitesi artırılır (Sandler vd., 2018; Hu vd., 2018).

EfficientNetB0, yaklaşık 5.3 milyon parametreye sahiptir ve ImageNet veri kümesi üzerinde son derece rekabetçi doğruluk elde eder. Bu parametre verimliliği, onu mobil ve gömülü sistemlerde kullanıma uygun hale getirir. Model, daha derin ve daha büyük EfficientNet versiyonlarının (B1–B7) oluşturulmasında referans ölçekleme noktası olarak görev yapar. EfficientNet mimarisi, yalnızca doğruluk açısından değil, aynı zamanda FLOPs (Floating Point Operations) bakımından da daha verimlidir. Tan ve Le (2019) çalışmalarında, benzer doğrulukta modellerden daha az hesaplama kaynağı kullanarak daha yüksek verimlilik sağladıklarını göstermişlerdir.

EfficientNetB0'nun bir diğer önemli özelliği, MBConv bloklarındaki sıkıştırma-genişletme (SE) katmanlarıdır. SE modülleri, kanal boyutunda dikkati öğrenerek önemli özelliklerin vurgulanmasını sağlar ve modelin genel performansını artırır (Hu vd., 2018). Ayrıca, EfficientNet mimarisi, AutoML NAS (Neural Architecture Search) yöntemleriyle optimize edilmiş blok ve katman yapılarına sahiptir, bu sayede el ile tasarlanmış geleneksel CNN mimarilerine göre daha dengeli ve verimli bir tasarım sunar.

Sonuç olarak EfficientNetB0, dengeli ölçeklendirme, parametre verimliliği ve modern blok tasarımı ile öne çıkar. Hem akademik araştırmalarda hem de endüstriyel uygulamalarda yaygın olarak kullanılmaktadır. Özellikle transfer öğrenme senaryolarında önceden eğitilmiş EfficientNetB0 ağırlıkları, sınırlı veri içeren problemler için etkili bir başlangıç noktası sunar.



Şekil 3.15. EfficientNETB0 mimarisi (Chen vd. 2024)

Şekil 3.15'te EfficientNetB0 mimarisinin genel yapısı gösterilmektedir. Model, başlangıçta standart evrişim katmanıyla giriş verisini işler ve ardından bir dizi MBConv bloğu aracılığıyla derin özellik çıkarımı gerçekleştirir. Farklı büyüklükteki çekirdek boyutlarına (3×3, 5×5) sahip MBConv katmanları, özellikleri çok ölçekli olarak yakalamayı hedefler. Katman sayılarının farklı aşamalarda tekrarlanması, modelin

derinliğini ve temsil kapasitesini artırır. Son aşamada yer alan adaptif ortalama havuzlama katmanı ise sınıflandırmaya hazır, sabit boyutta bir özellik vektörü üretir.

3.2.2.5. MobileNetV2

MobileNetV2, Sandler ve çalışma arkadaşları (2018) tarafından önerilen ve mobil cihazlarda veya hesaplama kısıtlı ortamlarda derin öğrenme modellerinin verimli çalışmasını amaçlayan bir ESA mimarisidir. MobileNetV2'nin temel tasarım felsefesi, daha az hesaplama maliyeti ve daha düşük parametre sayısı ile yüksek doğruluk sağlamaktır. Bu amaçla, derin ayrık evrişimler (depthwise separable convolutions) ve inverted residual blok yapısı geliştirilmiştir (Sandler vd., 2018).

Depthwise separable convolution, klasik konvolüsyon işlemini iki aşamaya ayırarak parametre ve işlem maliyetini önemli ölçüde azaltır. İlk aşama depthwise convolution, her giriş kanalında ayrı ayrı evrişim uygulanarak uzamsal özellik çıkarır. İkinci aşama olan pointwise convolution (1x1 konvolüsyon), kanallar arası kombinasyonları öğrenir. Bu ayırım sayesinde klasik evrişimlere göre yaklaşık sekiz ila dokuz kat daha az hesaplama gerektirir (Howard vd., 2017).

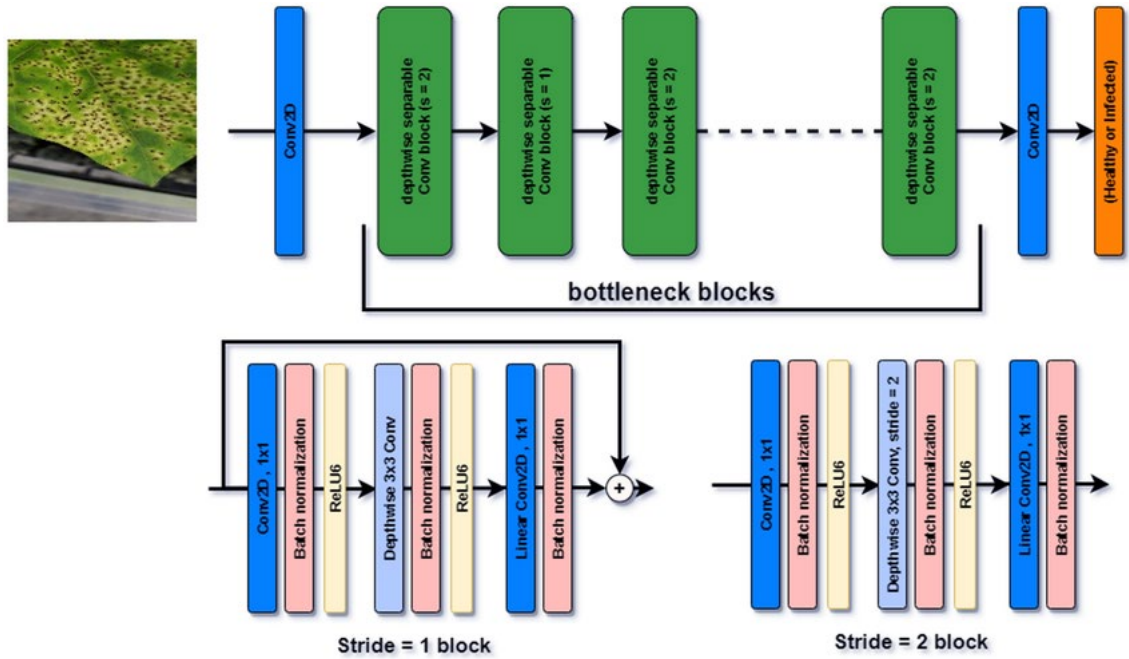
MobileNetV2'nin en önemli yeniliği, inverted residual blok tasarımıdır. Geleneksel “residual block” yapılarında dar bir yapıdan geniş bir temsil alanına geçiş görülürken, inverted residual bloklarda genişleterek (expand) başlamak ve tekrar daraltarak (project) sonlandırmak esas alınır. Bu bloklar, linear bottleneck adı verilen katmanlar içerir. Linear bottleneck, aktivasyon fonksiyonunun bilgi kaybını önlemek için doğrusal özellik haritaları üretir ve daha ince temsillerde bilgiyi korur. Böylece model, yüksek seviyeli özellikleri korurken parametre verimliliğini de sağlar (Sandler vd., 2018).

MobileNetV2 mimarisinde tüm yapı bu blokların ardışık bir dizilimiyle inşa edilmiştir. İlk katmanda klasik bir 3x3 evrişim yer alır. Ardından sırasıyla birden fazla inverted residual blok yer alır. Bu bloklar içerisinde ReLU6 gibi sınırlandırılmış aktivasyon fonksiyonları kullanılarak sayısallaştırma (quantization) uyumluluğu artırılmış ve mobil donanımlarda verimli çalışması sağlanmıştır. Ayrıca, her blokta shortcut (skip connection) kullanılarak gradyan akışının daha iyi gerçekleşmesi ve derin yapıların eğitilebilirliği artırılmıştır.

Parametre verimliliği, MobileNetV2'nin başlıca avantajlarından biridir. Örneğin, ImageNet veri kümesinde eğitilmiş bir MobileNetV2 modeli yaklaşık 3.4 milyon parametre içerirken, çok daha büyük modellerle karşılaştırılabilir doğruluk sağlar. Bu hafif yapı, akıllı telefonlar, IoT cihazları ve gömülü sistemler gibi sınırlı kaynaklı

ortamlarda çalışmaya uygundur. Ayrıca, MobileNetV2 genellikle transfer öğrenme senaryolarında önceden eğitilmiş ağırlıklarla kullanılabilir ve sınırlı veri bulunan özel sınıflandırma problemlerinde yüksek performans sunar (Howard vd., 2019).

Sonuç olarak, MobileNetV2 mimarisi, inverted residual blok yapısı, depthwise separable convolution stratejisi ve linear bottleneck katmanları ile hesaplama açısından verimli, düşük gecikmeli ve yüksek doğruluklu çözümler sunar. Özellikle mobil cihazlarda gerçek zamanlı görüntü sınıflandırma uygulamaları için tasarlanmış modern, hafif ve optimize edilmiş bir ESA mimarisi olarak öne çıkar.

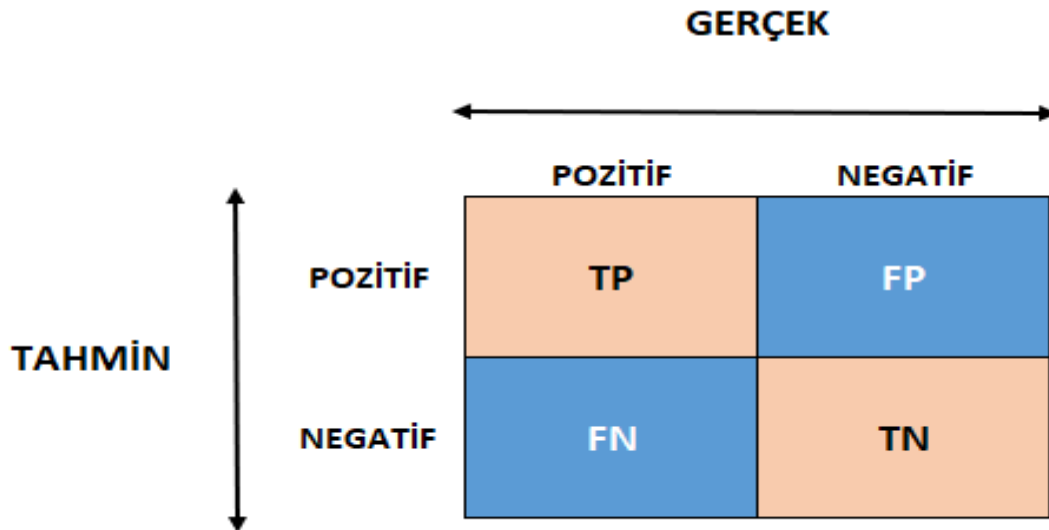


Şekil 3.16. MobileNetV2 mimarisi (Shahoveisi vd. 2023)

Şekil 3.16’da MobileNetV2 mimarisi, giriş katmanından başlayarak örnek bir görüntünün sınıflandırma sürecindeki tüm aşamalarını içeren bir şemayla sunulmuştur. İlk aşamada, 32 filtreye sahip bir evrişim katmanı temel kenar ve desen bilgilerini çıkarır. Sonrasında yer alan 19 adet residual bottleneck katmanı, modelin derinlemesine özellik öğrenmesini sağlar ve bilgi kaybını azaltır. Bu katmanlar, düşük parametre sayısıyla karmaşık yapıları tanıyabilmeyi mümkün kılar. Son aşamada ise tam bağlantılı bir katman yer alarak sınıflandırma sonucunu üretir. Görselde bu işlem akışı, transfer öğrenme yaklaşımında önceden eğitilmiş ağı farklı bir probleme adapte etmenin bir örneği olarak gösterilmiştir.

3.2.3. Performans ölçütleri

Bu bölümde, geliştirilen sınıflandırma modelinin başarımını değerlendirmek amacıyla kullanılan performans ölçütleri detaylı olarak açıklanmaktadır. Modelin sınıflandırma yeteneğini nesnel olarak ölçmek için doğruluk (accuracy), duyarlılık (recall), kesinlik (precision), F1-skoru, karışıklık matrisi ve ROC eğrisi gibi metriklerden yararlanılmıştır. Bu ölçütler, modelin farklı hata türlerini nasıl yönettiğini ortaya koyarak genel ve sınıf bazlı başarımı kapsamlı bir şekilde değerlendirmeye olanak tanır. Şekil 3.17’de öncelikle modelin sınıflandırma performansını temel düzeyde özetleyen karışıklık matrisi sunulmakta ve gerçek etiketlerle model tahminleri arasındaki doğru ve yanlış sınıflamaların dağılımı detaylı olarak gösterilmektedir.



Şekil 3.17. Karışıklık Matrisi

TP (True Pozitif): Gerçek sınıfı “boyalı” olan bir araç boyası örneğinin model tarafından doğru bir şekilde “boyalı” olarak tahmin edilmesidir. Yani hem gerçek hem tahmin etiketi boyalıdır.

TN (True Negatif): Gerçek sınıfı “orijinal” olan bir aracın doğru şekilde “orijinal” olarak tahmin edilmesidir. Modelin doğru şekilde boyalı olmadığını anlamasıdır.

FP (False Pozitif): Gerçek sınıfı “orijinal” olan bir araç boyası örneğinin model tarafından yanlış bir şekilde “boyalı” olarak tahmin edilmesidir. Model burada yanlış alarm üretmiştir.

FN (False Negatif): Gerçek sınıfı “boyalı” olan bir aracın model tarafından yanlışlıkla “orijinal” olarak tahmin edilmesidir. Bu durum, modelin boyalı aracı tanıyamayıp atlamasıdır.

Bu çalışmada sınıflandırma başarımı, aşağıdaki metrikler yardımıyla değerlendirilmiştir. Başarı (Accuracy), Kesinlik (Precision), Duyarlılık (Recall) ve F-Ölçütü (F-score) formülleri sırasıyla aşağıda verilmiştir:

$$Başarı = \frac{TP+TN}{TP+TN+FN+FP} \quad (3.4)$$

$$Kesinlik = \frac{TP}{TP+FP} \quad (3.5)$$

$$Duyarlılık = \frac{TP}{TP+FN} \quad (3.6)$$

$$F - score = 2 * \frac{Kesinlik * Duyarlılık}{Kesinlik + Duyarlılık} \quad (3.7)$$

ROC (Receiver Operating Characteristic): ROC eğrisi, modelin sınıflandırma performansını farklı karar eşiklerinde değerlendirmek için kullanılan bir araçtır. İkili sınıflandırma problemlerinde modelin duyarlılık (True Positive Rate, TPR) ile yanlış pozitif oranı (False Positive Rate, FPR) arasındaki ilişkiyi görselleştirir. TPR ve FPR şu formüllerle tanımlanır:

$$TPR = \frac{TP}{TP+FN} \quad (3.8)$$

$$FPR = \frac{FP}{FP+TN} \quad (3.9)$$

ROC eğrisi, modelin karar eşiği değiştikçe bu iki oran arasındaki dengeyi ortaya koyar. Eğrinin altında kalan alan (AUC – Area Under Curve) değeri, modelin sınıflar arasında ayırt edebilme başarısını özetleyen tek bir ölçüt olarak yorumlanır. AUC değeri 1'e yaklaştıkça modelin sınıflandırma yeteneği artar, 0.5 civarında olması ise rastgele tahmin düzeyini ifade eder.

4. BULGULAR

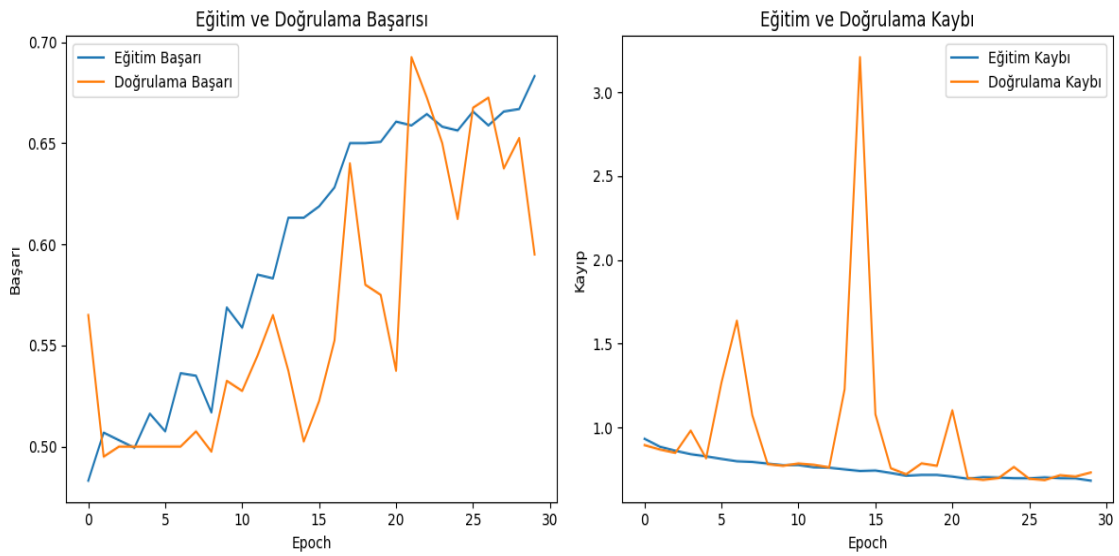
Bu çalışmada, araç boya yüzeylerinin orijinal ve boyalı olarak sınıflandırılması amacıyla derin öğrenme tabanlı transfer öğrenme yaklaşımları kullanılarak kapsamlı bir model karşılaştırması gerçekleştirilmiştir. MobileNetV2, InceptionV3, ResNet50, VGG16 ve EfficientNetB0 gibi önceden eğitilmiş evrimsel sinir ağı mimarileri, veri artırma teknikleri ve ince ayar (fine-tuning) stratejileri ile eğitilerek test edilmiştir. Eğitim-doğrulama bölünmesi %80-%20 oranında yapılmış ve model performansı doğruluk (accuracy), doğrulama kaybı (val_loss), ROC eğrisi ve karışıklık matrisi gibi metriklerle değerlendirilmiştir. Sonuçlar, özellikle InceptionV3, MobileNetV2 ve VGG16 modellerinin doğrulama setinde %90 üzeri başarı oranları elde ederek karmaşık ve parlak metalik yüzey farklılıklarını diğer modellere göre daha iyi ayırt edebildiğini göstermiştir. ResNet50 ve EfficientNetB0 modellerinin ise sınıf dengesizliğine ve düşük veri çeşitliliğine daha duyarlı olduğu ve doğrulama başarımlarının %50-%68 seviyelerinde kaldığı belirlenmiştir. Ayrıca, uygulanan veri artırma stratejileri ve katman dondurma sayıları modellerin genelleme yeteneğini artırmada önemli rol oynamış, doğru öğrenme oranı (learning rate) ayarlamaları ile aşırı öğrenme etkileri azaltılmıştır. Eğitim ve doğrulama kaybı grafikleri incelendiğinde, iyi performans gösteren modellerin kayıplarının istikrarlı şekilde azaldığı ve doğrulama başarısının doygunluğa ulaştığı gözlemlenmiştir. ROC eğrileri ve AUC değerleri de sınıflandırma ayırım yeteneğinin modellenmesine katkı sağlamıştır. Bu bulgular, derin öğrenme tabanlı transfer öğrenme modellerinin araç boya orijinallliği sınıflandırmasında uygulanabilirliğini ortaya koyarken, daha dengeli ve çeşitli veri kümeleri ile eğitildiklerinde gerçek dünyada daha yüksek doğruluk ve genelleme yeteneği sağlayabileceklerini göstermektedir. Tablo 4.1’de tüm modeller için kullanılan standart ayarlar gösterilmektedir.

Tablo 4.1. Modeller için kullanılan standart parametreler

Parametreler	Değerler
Eğitim-Doğrulama	%80-%20
Batch_Size	16
Dropout	0.5
Epoch	30
Optimizer	Adam
Learning_rate	0.0001

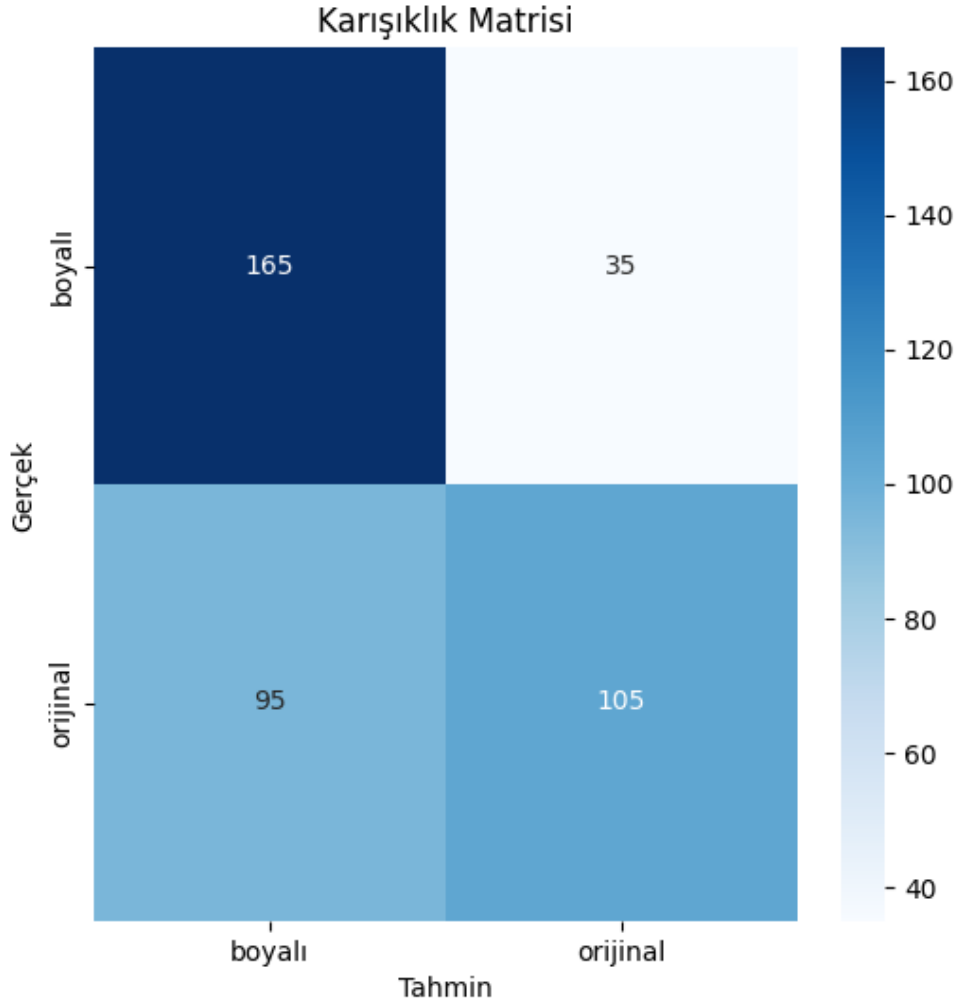
4.1. ResNet50 Bulguları

ResNet50 mimarisi kullanılarak yapılan sınıflandırma eğitiminde modelin doğrulama başarıları yaklaşık 0.68 seviyelerine ulaşmıştır. Eğitim sırasında doğrulama ve eğitim başarı eğrileri birbirine paralel bir artış göstermiş, aşırı öğrenme eğilimi sınırlı düzeyde kalmıştır. Şekil 4.1’de kayıp değerlerinde özellikle doğrulama kaybında bazı dalgalanmalar ve ani sıçramalar gözlenmiştir. Bu durum, verinin karmaşıklığı ve sınıf içi çeşitlilik nedeniyle modelin bazı dönemlerde daha zor genelleme yapabildiğini göstermektedir.



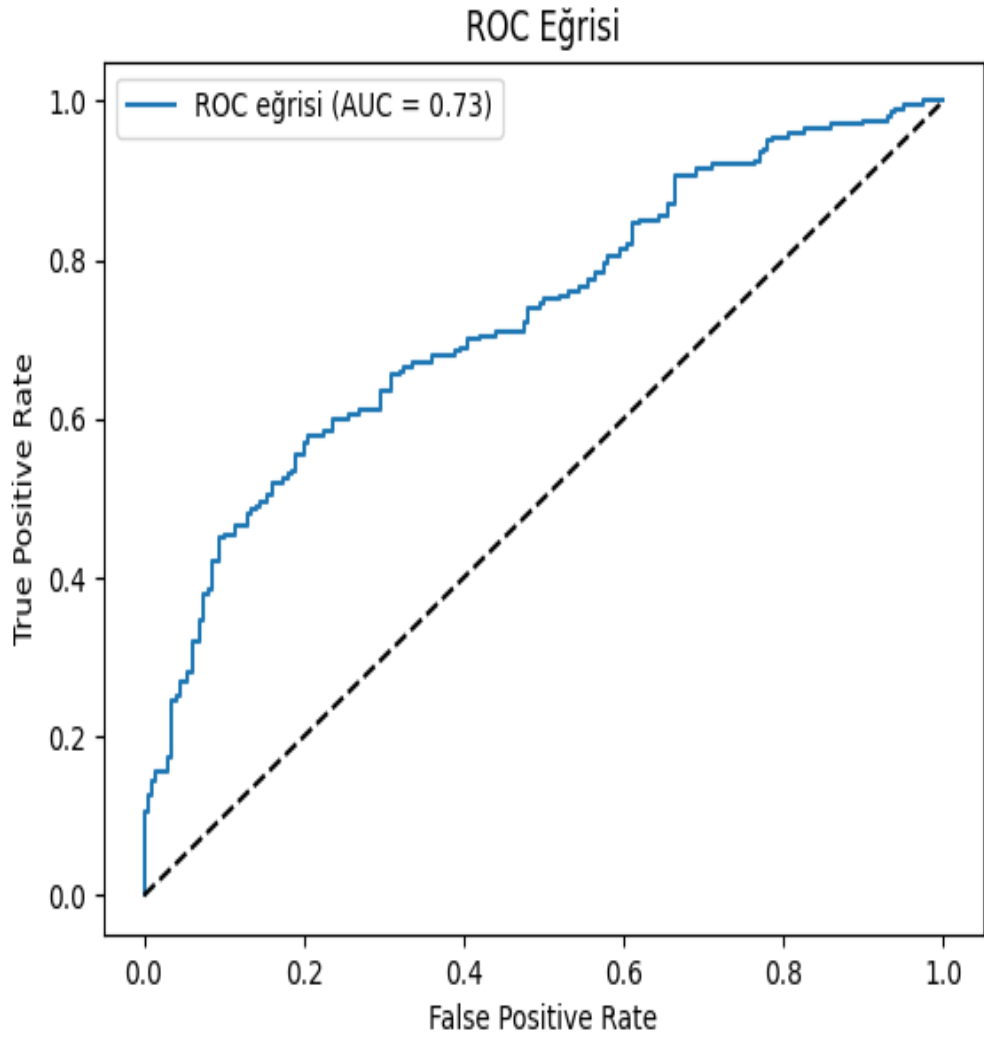
Şekil 4.1. ResNet50 mimarisinin eğitim ve doğrulama grafikleri

Şekil 4.2’de karışıklık matrisine göre model, boyalı sınıfı 165 doğru ve 35 yanlış şekilde tahmin etmiş; orijinal sınıfı ise 105 doğru ve 95 yanlış sınıflandırmıştır. Boyalı yüzeylerde kesinlik değeri 0.63 ve duyarlılık değeri 0.82 olarak elde edilmiştir. Orijinal sınıfta ise kesinlik 0.75 seviyesinde iken duyarlılık değeri 0.53’te kalmıştır. Bu sonuçlar, modelin boyalı yüzeyleri daha başarılı ayırt edebildiğini, ancak orijinal yüzeylerdeki varyasyonlar ve detaylı yapılar nedeniyle daha fazla hata yaptığını ortaya koymaktadır.



Şekil 4.2. ResNet50 mimarisinin karışıklık matrisi

Şekil 4.3'te ROC eğrisinin altında kalan alan (AUC) yaklaşık 0.73 olarak hesaplanmıştır. Bu değer, modelin sınıfları ayırt etme becerisinin makul ancak mükemmel olmadığını gösterir. AUC 0.5'ten büyük olduğu için modelin rastgele tahminden anlamlı ölçüde daha iyi performans gösterdiği söylenebilir. ROC eğrisinin kademeli yükselmesi sınıf ayrımının belirli eşik değerlerinde daha başarılı olabileceğine işaret eder.



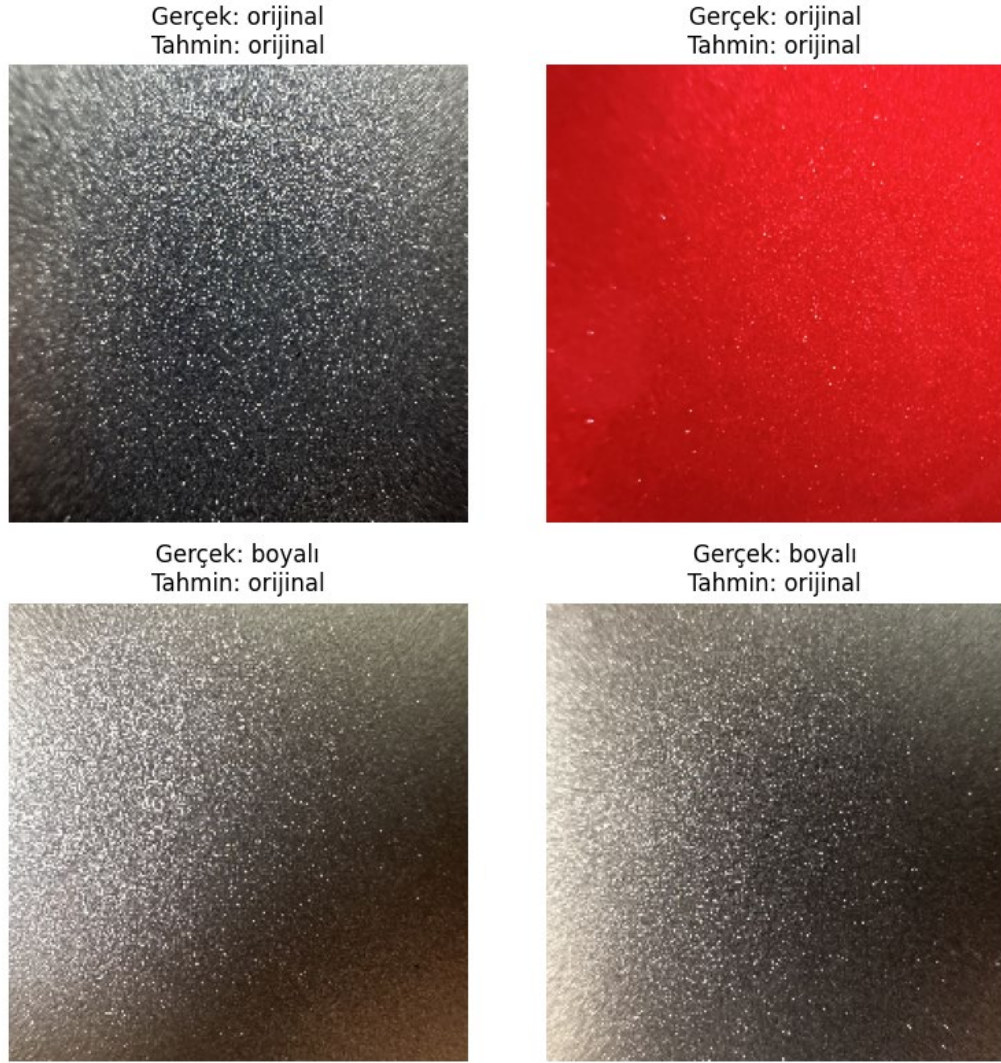
Şekil 4.3. ResNet50 mimarisi ROC grafiği

Tablo 4.2’de %80-%20 oranında ayrılan eğitim ve doğrulama bölmesinde doğrulama setindeki 400 görüntü için, boyalı ve orijinal sınıflarına ait kesinlik, duyarlılık ve F-Score metrikleri sunulmuştur. Modelin doğrulama başarısı ise 0.68 olarak elde edilmiştir.

Tablo 4.2. Resnet50 ile sınıflandırma metrikleri

Sınıflar	Kesinlik	Duyarlılık	F-score	Veri Sayısı
Boyalı	0.63	0.82	0.72	200
Orijinal	0.75	0.53	0.62	200

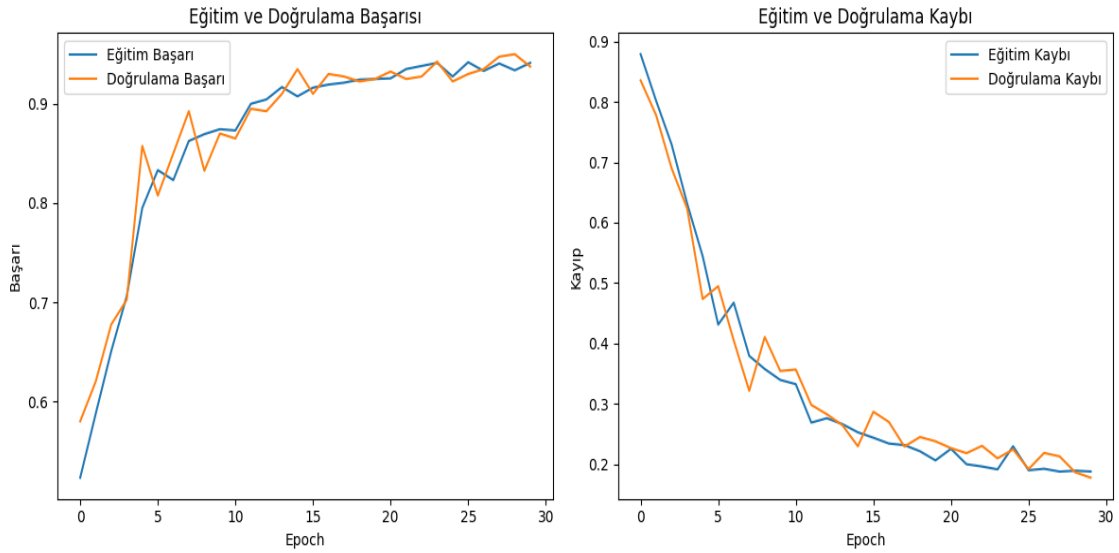
Görsel tahmin örnekleri incelendiğinde modelin orijinal yüzeyleri genelde doğru tahmin ettiği ancak bazı boyalı yüzeyleri orijinal olarak etiketlediği görülmektedir. Bu yanlış sınıflandırmalar, boyalı yüzeyin renk tonu, ışık yansıması veya metalik yapıdaki benzerliklerden kaynaklanabilir.



Şekil 4.4. ResNet50 ile örnek tahmin görüntüleri

4.2. VGG16 Bulguları

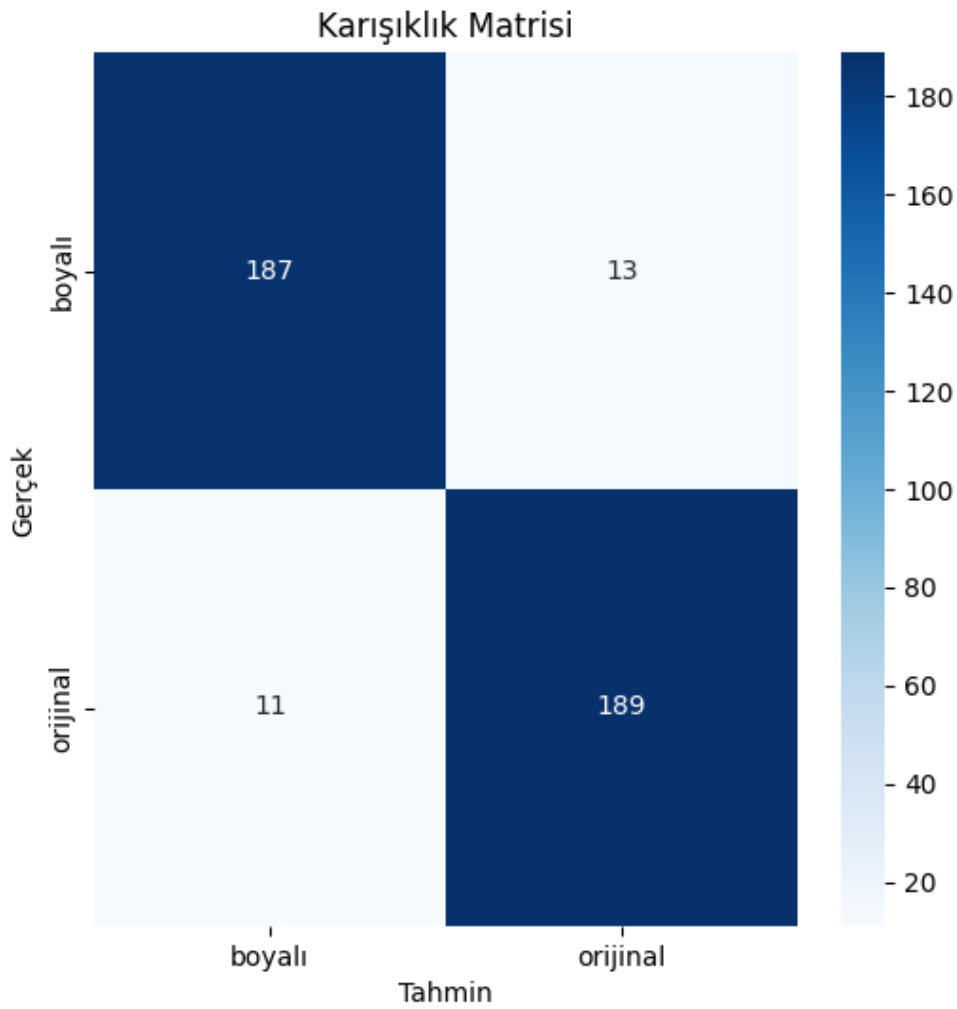
Bu çalışmada araç boyalarının orijinal ve boyalı yüzey olarak sınıflandırılması için transfer öğrenme yaklaşımı kapsamında VGG16 mimarisi tercih edilmiştir. VGG16, katı ve düzenli yapıdaki ardışık evrişim ve havuzlama katmanlarıyla derin özellik çıkarımı sağlayan, yaygınca kullanılan bir evrişimsel sinir ağıdır. Bu modelde ince ayar stratejisi olarak son 6 katmanı serbest bırakılmış, bu sayede önceden ImageNet üzerinde öğrenilen evrensel özelliklerin korunması sağlanırken, problem özelinde belirli bir uyum geliştirilmesi hedeflenmiştir. Bu yaklaşım, modelin hem genelleme yeteneğini hem de veri kümesine özgü ayırt edici özellikleri öğrenmesini desteklemektedir.



Şekil 4.5. VGG16 mimarisinin eğitim ve doğrulama grafikleri

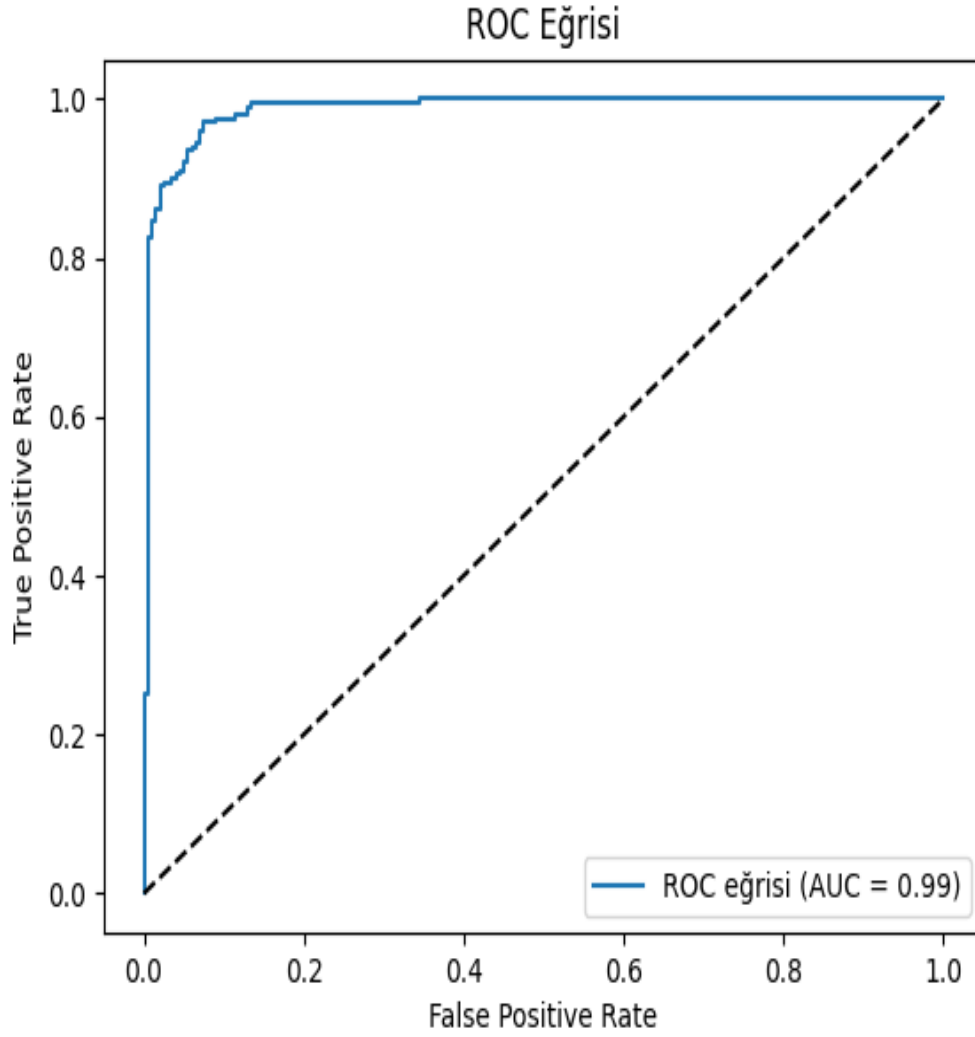
Şekil 4.5'te modelin eğitim ve doğrulama başarıları ile kayıplarının epoch sayısına göre değişimi gösterilmiştir. Eğitim ve doğrulama başarılarının yaklaşık 0.94 seviyesine kadar yükseldiği ve birbirine yakın seyredererek aşırı öğrenme riskinin düşük olduğu gözlemlenmektedir. Kayıp fonksiyonları ise düzenli bir şekilde azalarak modelin öğrenme sürecinde istikrarlı bir iyileşme sağladığını ortaya koymaktadır. Son 6 katmanın serbest bırakılarak ince ayar yapılması, modelin önceden öğrenilmiş özellikleri korurken veri setine özgü ayırt edici detayları öğrenmesini sağlamış ve yüksek doğruluk seviyelerine ulaşılmasına katkıda bulunmuştur.

Şekil 4.6’da yer alan karışıklık matrisi, modelin boyalı ve orijinal sınıfları ayırt etme performansını detaylı bir şekilde göstermektedir. Boyalı sınıfta 187 doğru tahmin, 13 yanlış tahmin; orijinal sınıfta ise 189 doğru tahmin ve 11 yanlış tahmin gerçekleştirilmiştir. Bu sonuçlar, sınıflar arasında oldukça dengeli ve başarılı bir ayırım yapıldığını göstermekte ve modelin sınıflar arası karışıklığı minimuma indirdiğini ortaya koymaktadır.



Şekil 4.6. VGG16 mimarisinin karışıklık matrisi

Şekil 4.7’de verilen ROC eğrisi altında kalan alan AUC değeri 0.99 olarak hesaplanmıştır. Bu yüksek AUC değeri, modelin sınıfları ayırma gücünün oldukça yüksek olduğunu ve tüm eşik değerlerinde dengeli bir duyarlılık ve özgüllük performansı sunduğunu göstermektedir. Bu sonuç, transfer öğrenme mimarisinin güçlü ayrıştırma kabiliyetini ve veri setine uyum sağladığını desteklemektedir.

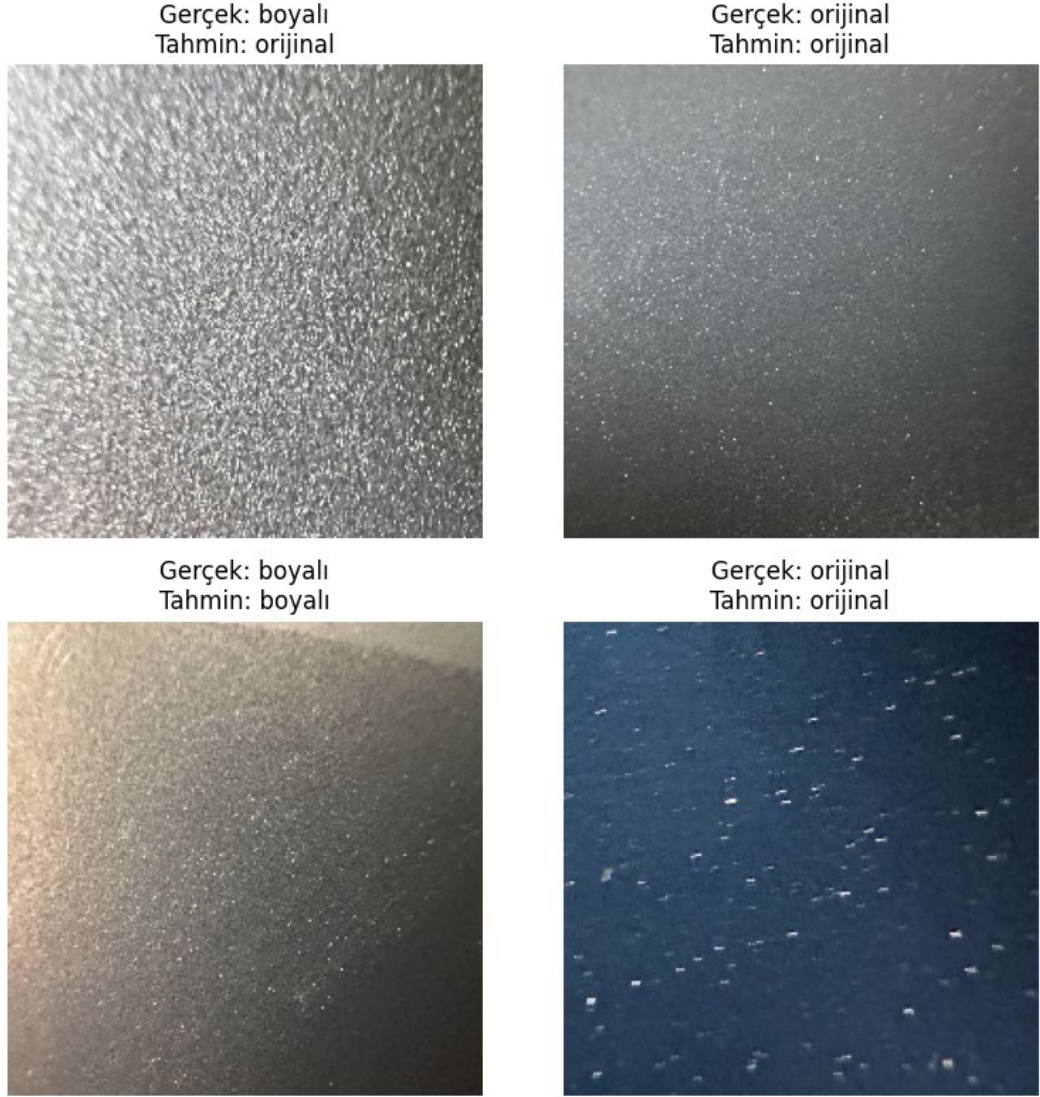


Şekil 4.7. VGG16 mimarisi ROC grafiği

Tablo 4.3. VGG16 ile sınıflandırma metrikleri

Sınıflar	Kesinlik	Duyarlılık	F-score	Veri Sayısı
Boyalı	0.94	0.94	0.94	200
Orijinal	0.94	0.94	0.94	200

Tablo 4.3'te modelin sınıf bazlı kesinlik, duyarlılık ve F-score değerleri gösterilmiştir. Her iki sınıf için kesinlik, duyarlılık ve F-score değerlerinin 0.94 olduğu gözlenmiştir. Bu sonuç, modelin boyalı ve orijinal sınıfları yüksek doğrulukla ayırt edebildiğini ve sınıflar arası dengesizliğin model performansını olumsuz etkilemediğini ortaya koymaktadır. Modelin doğrulama başarısı 0.94 olarak elde edilmiştir.

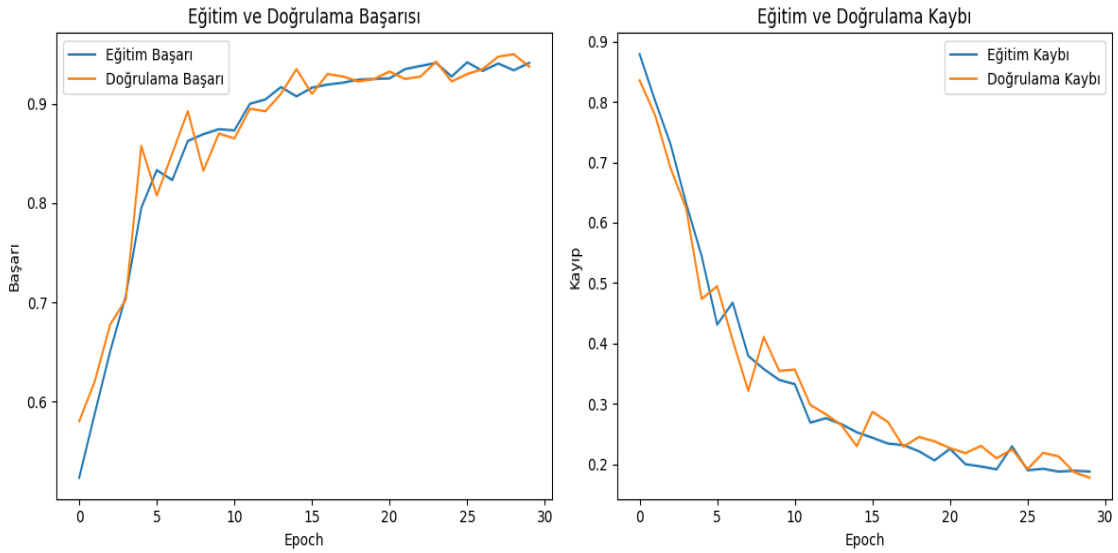


Şekil 4.8. VGG16 ile örnek tahmin görüntüleri

Şekil 4.8’de modelin doğrulama setinden rastgele seçilen bazı örneklerdeki tahmin sonuçları yer almaktadır. Görsellerde gerçek sınıf ve model tahmini karşılaştırıldığında, modelin yüzey dokusu ve parlaklık gibi detaylı özellikleri başarılı bir şekilde ayrıştırabildiği görülmektedir. Son 6 katmanın serbest bırakılması sayesinde transfer öğrenme modeli, önceden öğrenilmiş genel özelliklerin üzerine boyalı ve orijinal yüzeyler arasındaki ince farkları da adapte edebilmiştir.

4.3. InceptionV3 Bulguları

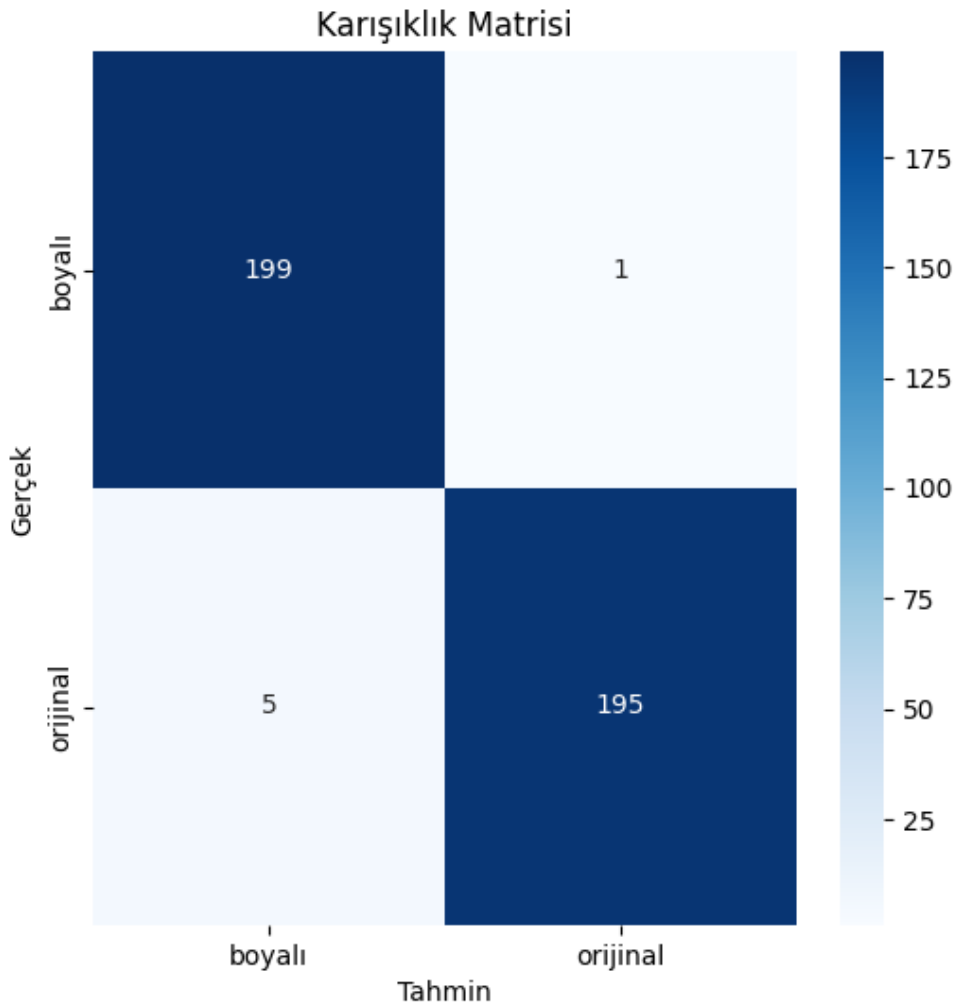
Bu çalışmada, araç boya yüzeylerinin orijinal veya boyalı olup olmadığını sınıflandırmak amacıyla InceptionV3 mimarisi tercih edilmiştir. InceptionV3, Google tarafından geliştirilmiş derin ve kompleks bir evrimsel sinir ağı mimarisi olup, farklı boyutlardaki filtreleri aynı anda kullanarak detaylı öznitelik çıkarımı sağlar. Modelin daha iyi genelleme yapabilmesi ve transfer öğrenmeden yararlanabilmesi için önceden ImageNet üzerinde eğitilmiş ağırlıkları kullanılmış ve son 100 katmanı serbest bırakılarak ince ayar yapılmıştır. Böylece hem önceden öğrenilen genel özellikler korunmuş hem de problemimize özgü detaylar öğrenilmiştir.



Şekil 4.9. InceptionV3 mimarisinin eğitim ve doğrulama grafikleri

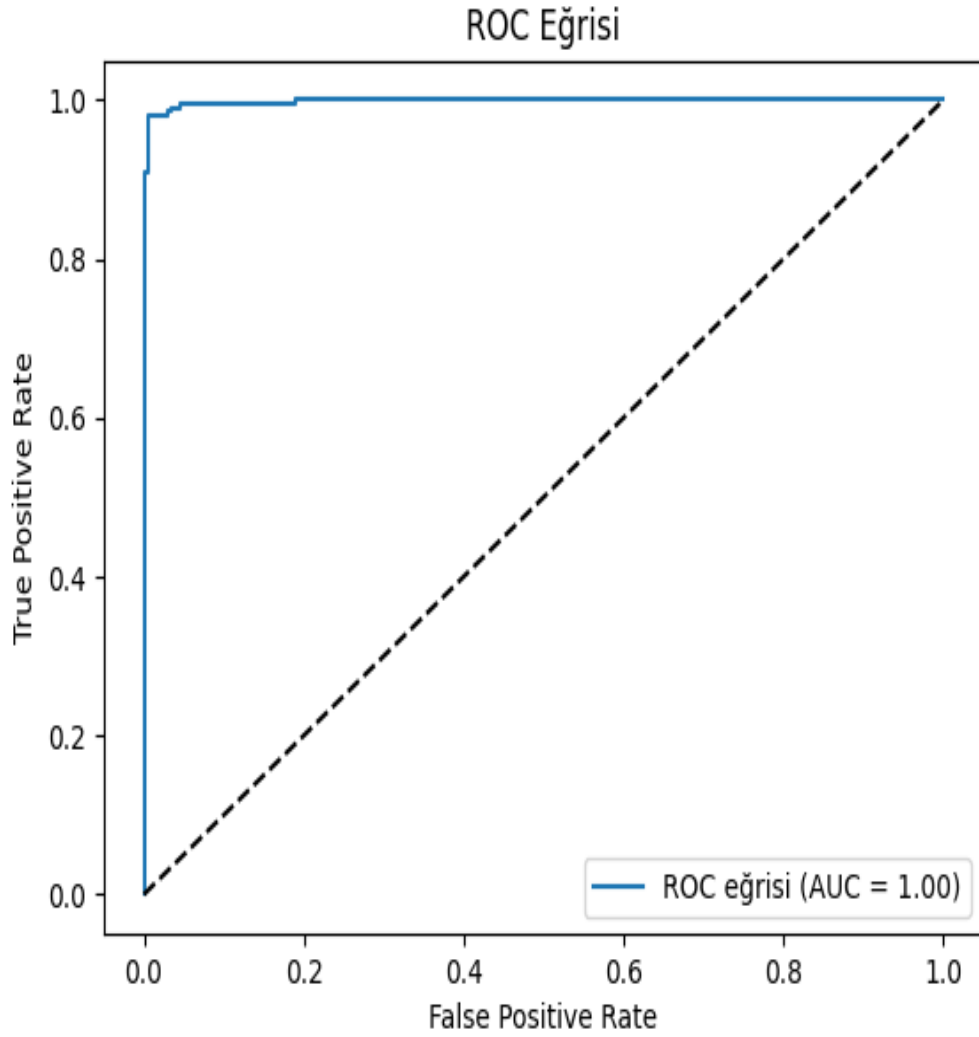
Şekil 4.9'de eğitim ve doğrulama başarı/kayıp değerlerinin epoch sayısına göre değişimi gösterilmiştir. Eğitim ve doğrulama başarı oranlarının yaklaşık 0.98 seviyelerine ulaştığı gözlenmektedir. Bu, modelin hem eğitim hem de doğrulama verisi üzerinde güçlü bir ayırt edici performans geliştirdiğini göstermektedir. Kayıp fonksiyonlarının da zamanla azalarak stabil hale gelmesi, aşırı öğrenme riskinin kontrol altında tutulduğunu ve modelin yüksek doğrulukla genelleme yapabildiğini göstermektedir. Bu başarıda son 100 katmanın serbest bırakılarak veri setine uyumlu şekilde yeniden eğitilmesinin etkili olduğu düşünülmektedir.

Şekil 4.10’da yer alan karışıklık matrisi modeli daha ayrıntılı inceleme imkanı sunmaktadır. Orijinal sınıfa ait 200 görüntüden 195’i doğru, boyalı sınıfa ait 200 görüntüden 199’u doğru sınıflandırılmıştır. Yanlış sınıflandırılan örnek sayısı son derece düşüktür. Bu sonuç, modelin iki sınıfı dengeli şekilde ayırt edebildiğini ve yüksek duyarlılık ile kesinlik değerlerine ulaştığını göstermektedir.



Şekil 4.10. InceptionV3 mimarisinin karışıklık matrisi

Şekil 4.11’de yer alan ROC eğrisi ve AUC değeri, sınıflandırma modelinin olağanüstü bir ayırt edici güce sahip olduğunu göstermektedir. Eğrinin 1’e yakın bir şekilde yükselmesi ve AUC değerinin 1.00 olması, yanlış pozitif oranını en aza indirirken doğru pozitif oranını maksimize eden bir karar sınırı geliştirdiğini ortaya koymaktadır.

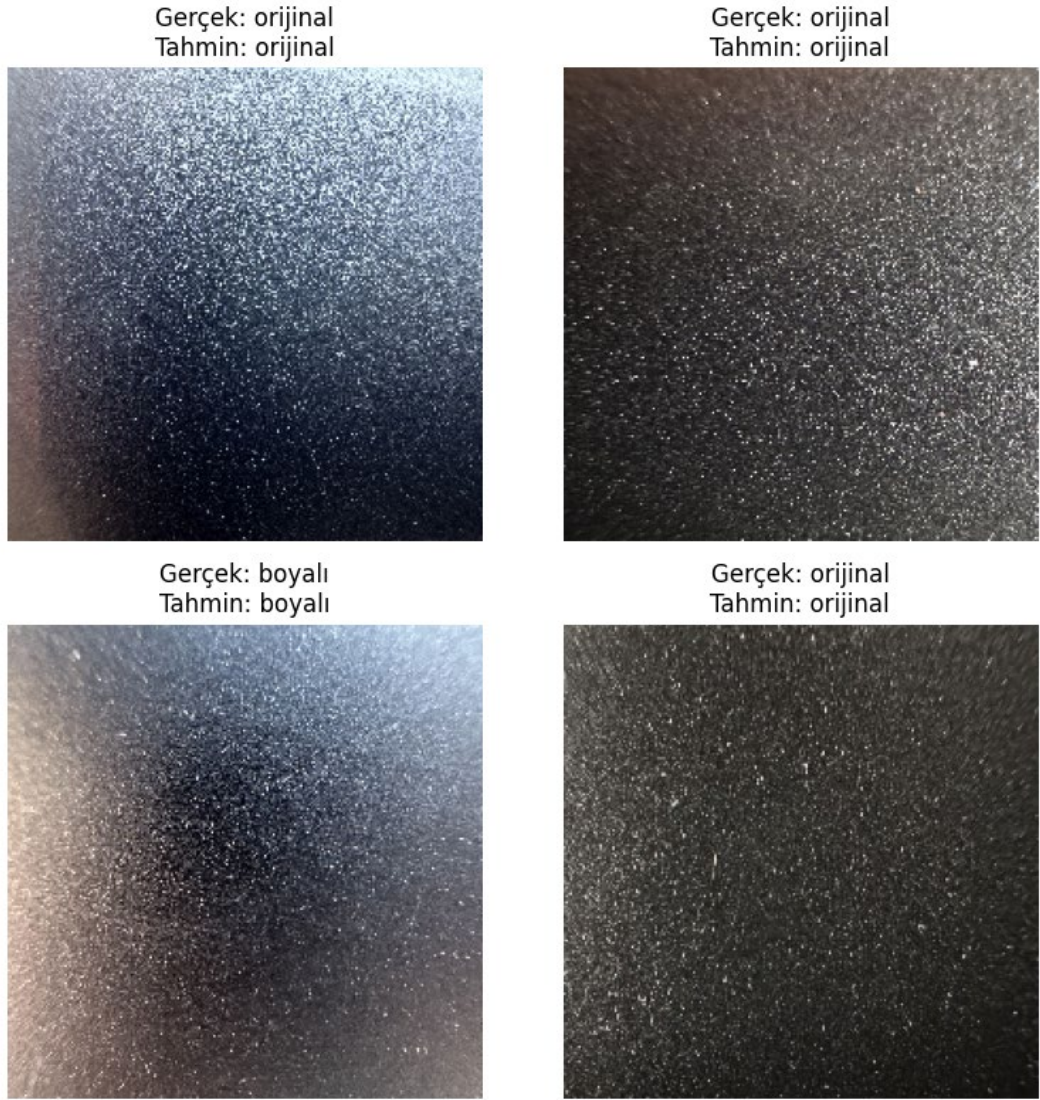


Şekil 4.11. InceptionV3 mimarisi ROC grafiđi

Tablo 4.4. InceptionV3 ile sınıflandırma metrikleri

Sınıflar	Kesinlik	Duyarlılık	F-score	Veri sayısı
Boyalı	0.98	0.99	0.99	200
Orijinal	0.99	0.97	0.98	200

Tablo 4.4’de yer alan sınıflandırma raporu, her sınıf için kesinlik, duyarlılık, F-score deđerlerini sunmaktadır. Hem boyalı hem de orijinal sınıf için kesinlik ve duyarlılık deđerleri 0.97–0.99 aralığında ölçölmüştür. F-score her iki sınıf için de 0.99 seviyesine ulaşmıştır. Genel başarı ise 0.98 olarak hesaplanmıştır. Bu sonuçlar, modelin iki sınıf için de dengeli ve yüksek doğrulukta tahminler yaptığını göstermektedir.

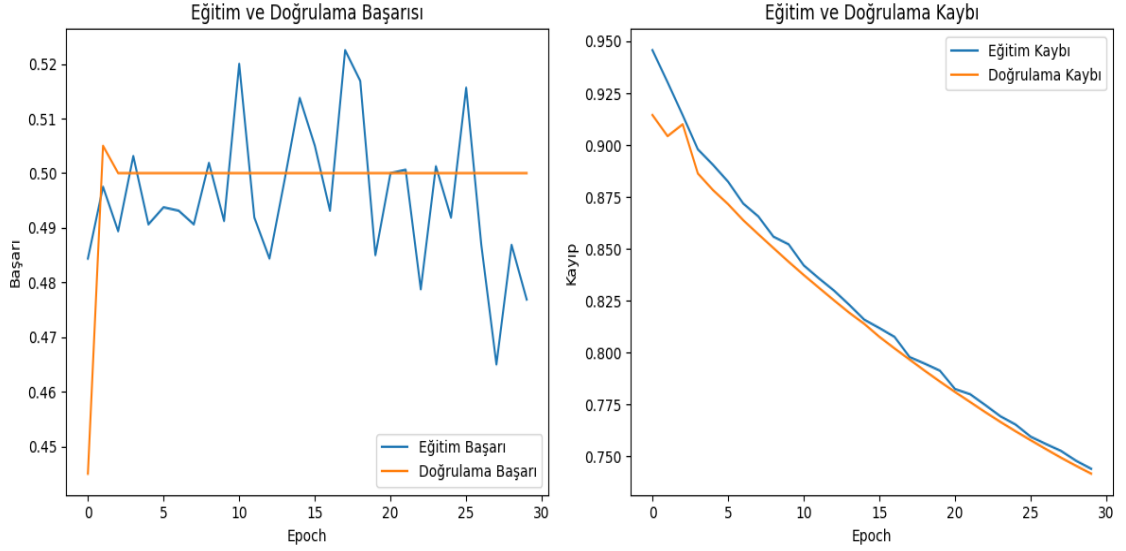


Şekil 4.12. InceptionV3 ile örnek tahmin görüntüleri

Şekil 4.12’de modelin doğrulama seti üzerinde gerçekleştirdiği örnek tahminler görselleştirilmiştir. Gerçek etiketler ile tahmin edilen etiketlerin uyumlu olduğu bu örnekler, modelin gerçek dünyada da anlamlı tahminler yapabildiğini göstermektedir. Boya yüzeyindeki renk tonu, doku ve parlaklık gibi özelliklerin başarılı şekilde öğrenildiği ve transfer öğrenmenin önemli bir avantaj sunduğu görülmektedir.

4.4. EfficientNetB0 Bulguları

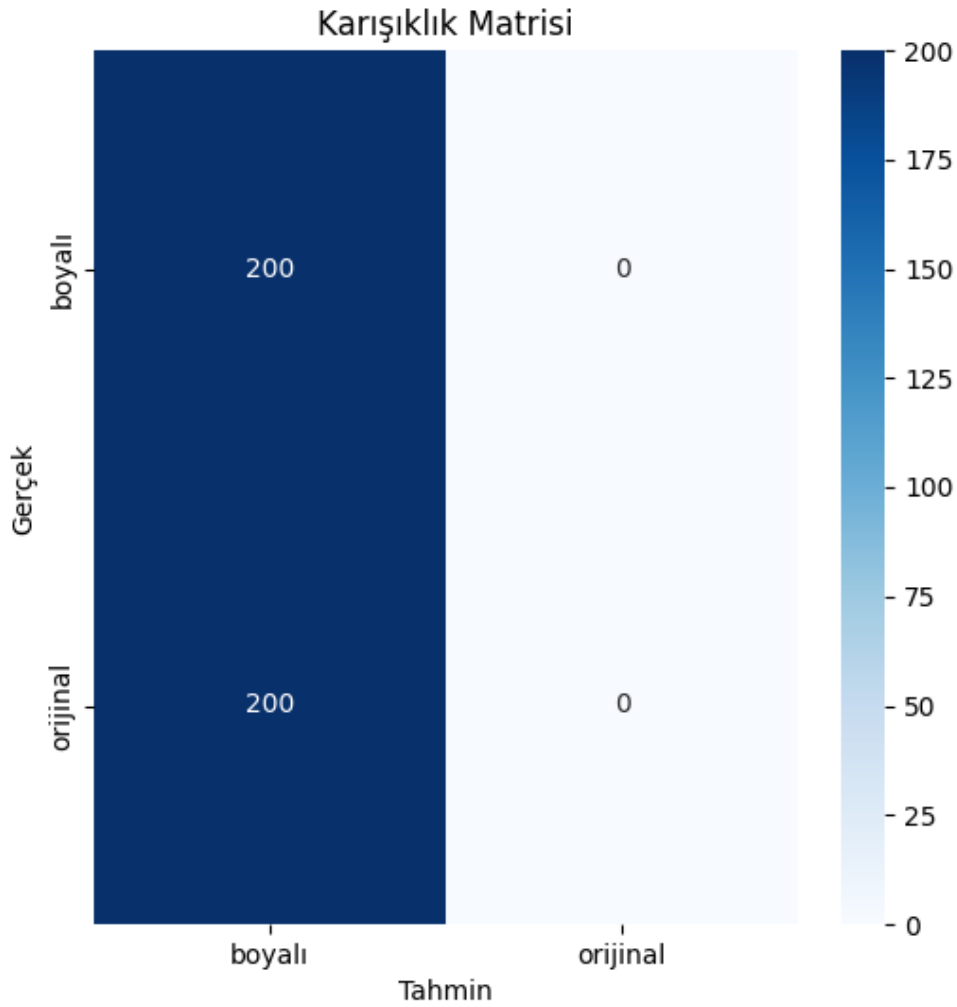
Bu çalışmada, transfer öğrenme yaklaşımıyla EfficientNetB0 mimarisi kullanılarak araç boya yüzeylerinin orijinal ve boyalı olarak sınıflandırılması amaçlanmıştır. EfficientNetB0, parametre verimliliği açısından optimize edilmiş bir ESA mimarisi olup genelde daha az parametreyle yüksek doğruluk sağlayabilir. Bu çalışmada ince ayar stratejisi olarak son 50 katman serbest bırakılarak, modelin probleme özgü detaylı özellikleri öğrenmesi hedeflenmiştir. Ancak sonuçlar, bu ayarın veri setindeki sınıflar arası ayrımı yeterince yakalayamadığını ve modelin istenen sınıflama başarımını gösteremediğini ortaya koymuştur. Bu durum, veri çeşitliliğinin ve çekim protokollerinin geliştirilmesi gerekliliğini vurgulamaktadır.



Şekil 4.13. EfficientNetB0 mimarisinin eğitim ve doğrulama grafikleri

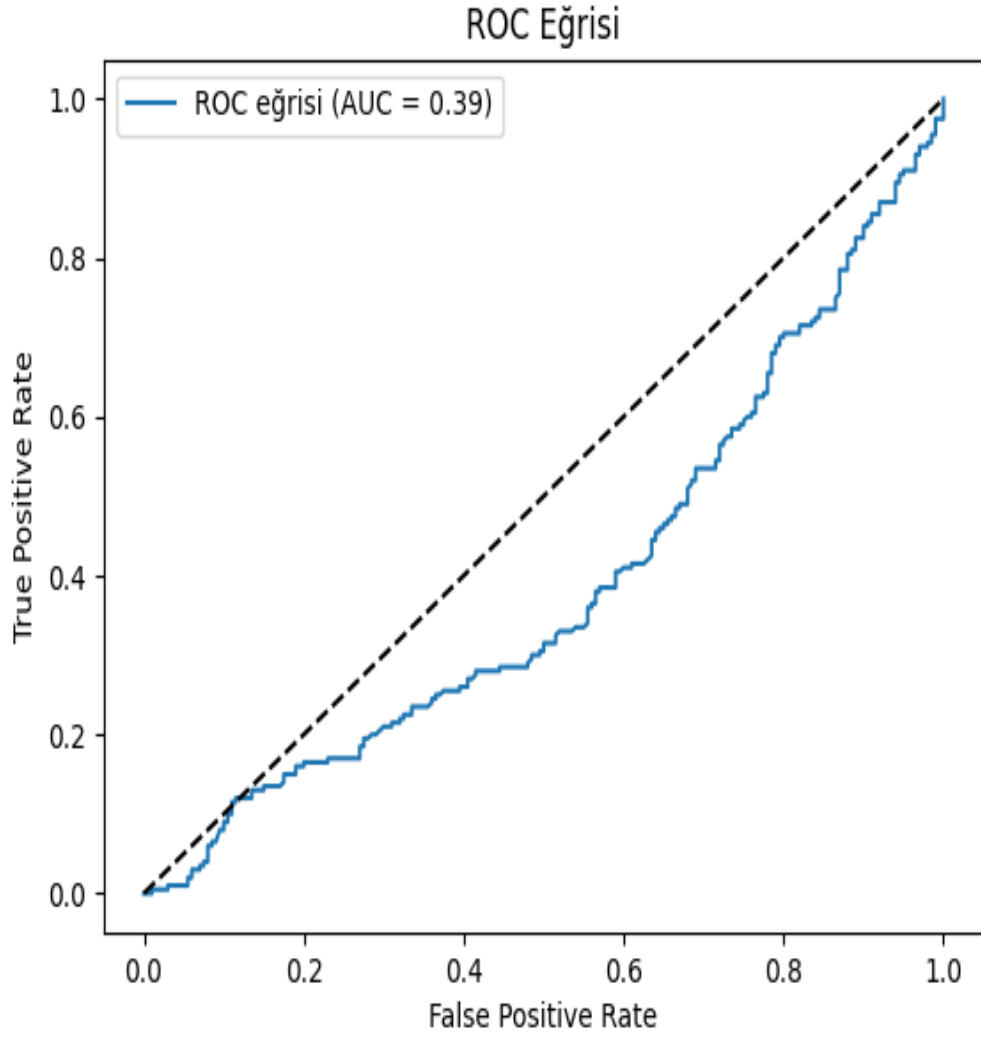
Şekil 4.13’de eğitim ve doğrulama başarı/kayıp grafikleri sunulmuştur. Eğitim sürecinde hem eğitim hem de doğrulama başarısının 0.50 civarında sabit kaldığı, belirgin bir öğrenme gerçekleşmediği görülmektedir. Kaybın ise her iki sette de yavaşça azaldığı ancak başarının artmadığı dikkat çekmektedir. Bu durum, modelin veri setindeki sınıfları ayırt etmekte zorlandığını ve ayırıcı özellikleri öğrenemediğini göstermektedir.

Şekil 4.14'te gösterilen karışıklık matrisi, modelin tüm örnekleri tek bir sınıfı (boyalı) tahmin ettiğini ortaya koymaktadır. 200 gerçek boyalı örneğin tamamı doğru tahmin edilirken, 200 gerçek orijinal örnek boyalı olarak sınıflandırılmıştır. Bu durum, modelin sınıflar arası ayırıcı özellikleri öğrenemediğini ve sınıflandırma görevinde başarısız kaldığını göstermektedir.



Şekil 4.14. EfficientNetB0 mimarisinin karışıklık matrisi

Şekil 4.15'te modelin ROC eğrisi verilmiştir. Elde edilen AUC değeri 0.39 olup rastgele tahmin AUC=0.5 seviyesinin bile altında kalmıştır. Bu sonuç, modelin sınıfları ayırma yetisinin zayıf olduğunu ve pratik kullanımda güvenilir sonuçlar vermeyeceğini göstermektedir. Sonuç olarak model, gerçek pozitif oranını düşük tutarken yanlış pozitif oranını yüksek tutmaktadır.



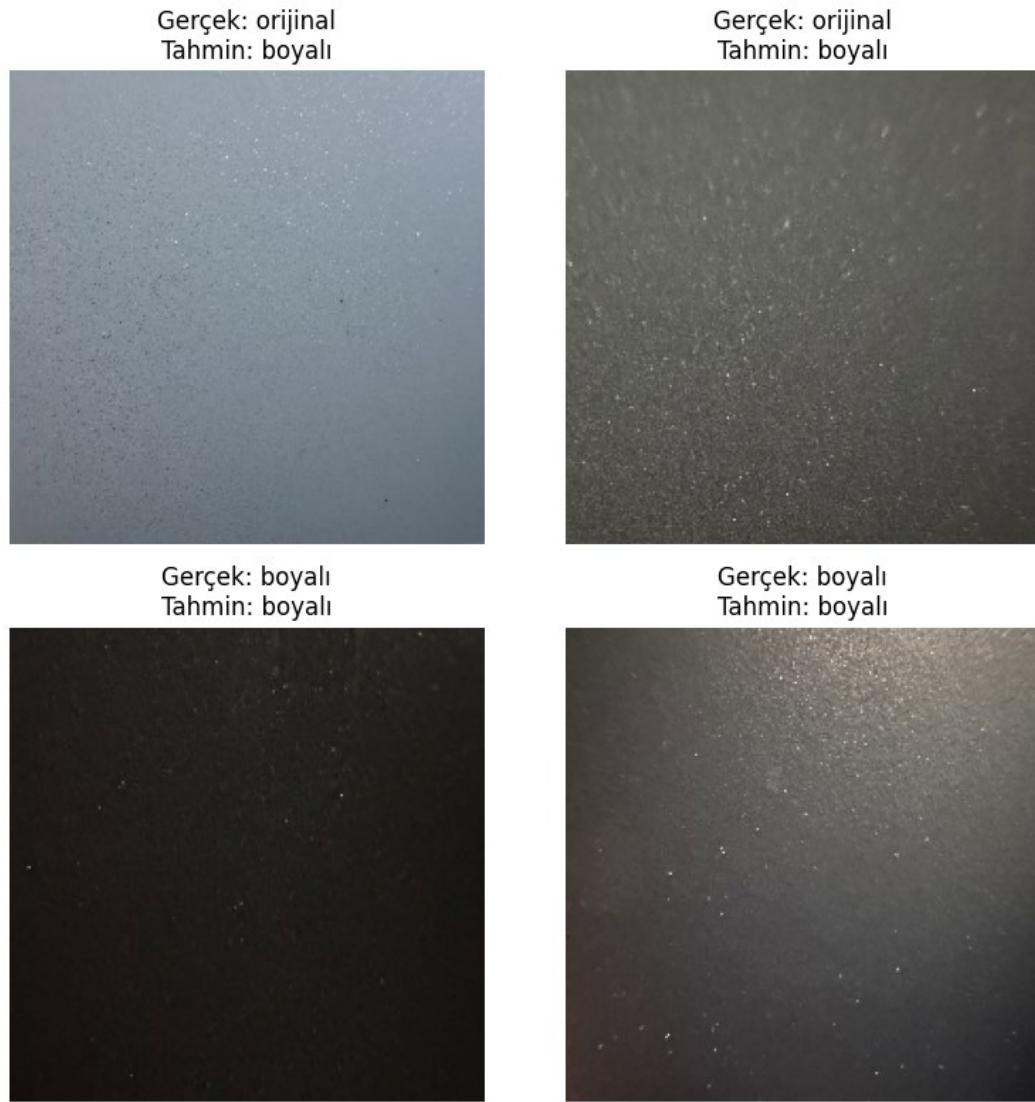
Şekil 4.15. EfficientNetB0 mimarisi ROC grafiđi

Tablo 4.5. EfficientNetB0 ile sınıflandırma metrikleri

Sınıflar	Kesinlik	Duyarlılık	F- score	Veri sayısı
Boyalı	0.50	1.00	0.67	200
Orijinal	0.00	0.00	0.00	200

Tablo 4.5'te yer alan sınıf bazlı metrik raporunda, boyalı sınıf için kesinlik 0.50, duyarlılık 1.00 ve F-score 0.67; orijinal sınıf için ise kesinlik 0.00, duyarlılık 0.00 ve F-score 0.00 elde edilmiştir. Bu tablo, modelin sadece boyalı sınıfını öğrenip orijinal sınıfı tamamen göz ardı ettiđini göstermektedir. Genel başarı 0.50 olarak kalmıştır ki bu dengeli iki sınıfta tamamen rastgele tahmin anlamına gelir.

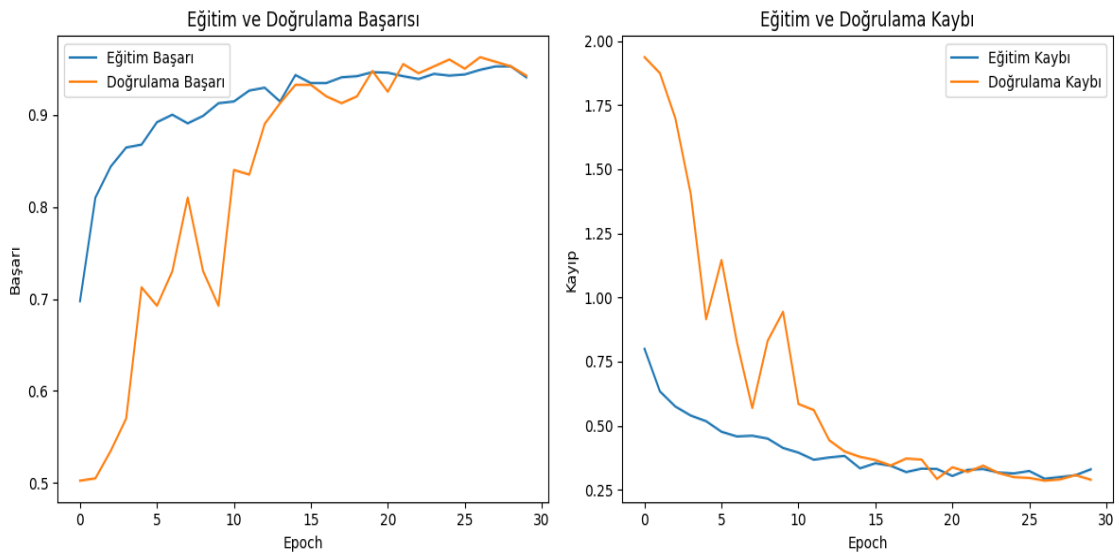
Şekil 4.16’da modelin örnek tahmin sonuçları gösterilmiştir. Tüm örneklerde model orijinal örnekleri boyalı olarak tahmin etmiştir. Görsellerin detaylı incelenmesi, boyalı ve orijinal yüzeylerin bazı durumlarda oldukça benzer görsel özelliklere sahip olduğunu ve modelin renk, doku gibi ayrımları yeterince öğrenemediğini ortaya koymaktadır. Bu durum, veri setinin çeşitliliğinin artırılması ve çekim protokollerinin standartlaştırılması gereğini vurgulamaktadır.



Şekil 4.16. EfficientNetB0 ile örnek tahmin görüntüleri

4.5. MobileNetV2 Bulguları

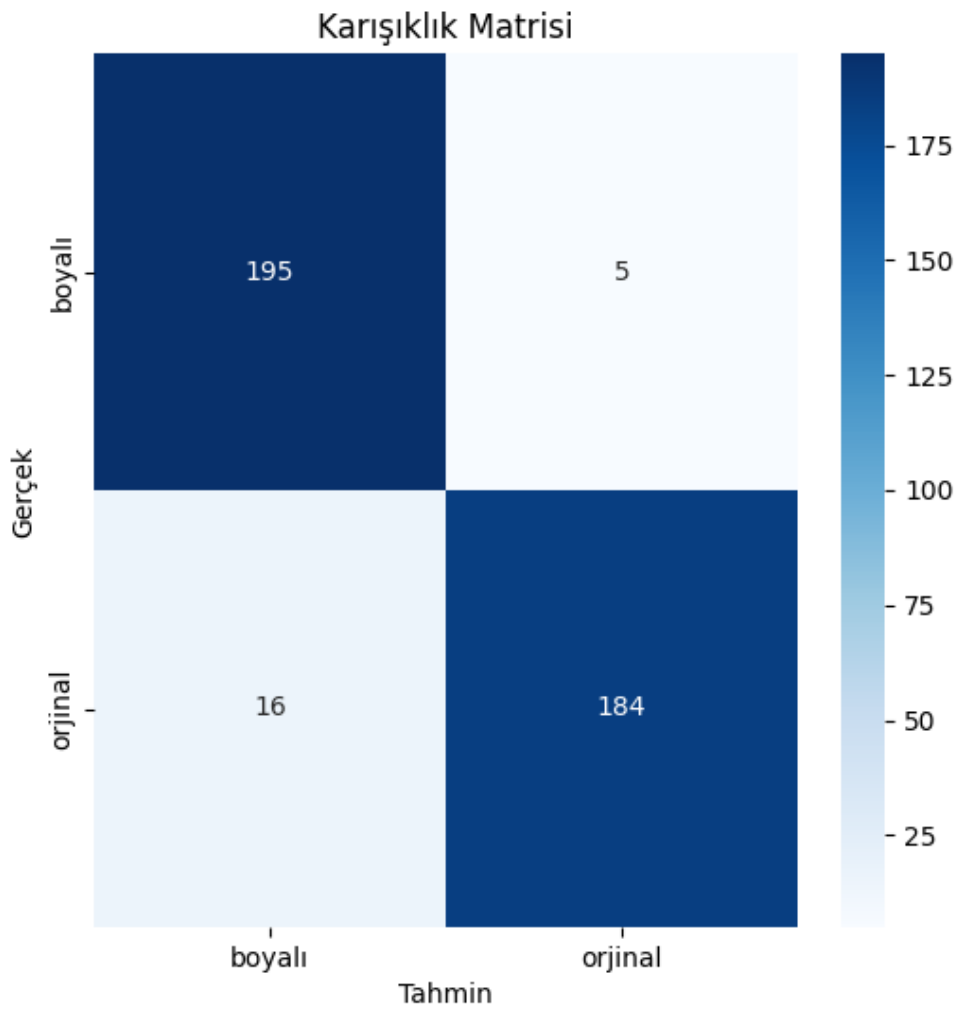
Bu çalışmada, transfer öğrenme yaklaşımıyla MobileNetV2 mimarisi kullanılarak araç boyalarının orijinal veya boyalı olup olmadığının sınıflandırılması hedeflenmiştir. MobileNetV2, mobil ve gömülü cihazlarda çalışmaya uygun hafif ama güçlü bir evrişimli sinir ağı mimarisidir. Modelin önceden ImageNet üzerinde öğrendiği genellemeler korunurken, ince ayar stratejisiyle son 30 katmanı serbest bırakılmıştır. Böylece, modelin araç boya yüzeyinin incelikli dokusal farklılıklarına daha iyi uyum sağlaması ve probleme özgü özellikler öğrenmesi amaçlanmıştır.



Şekil 4.17. MobileNetV2 mimarisinin eğitim ve doğrulama grafikleri

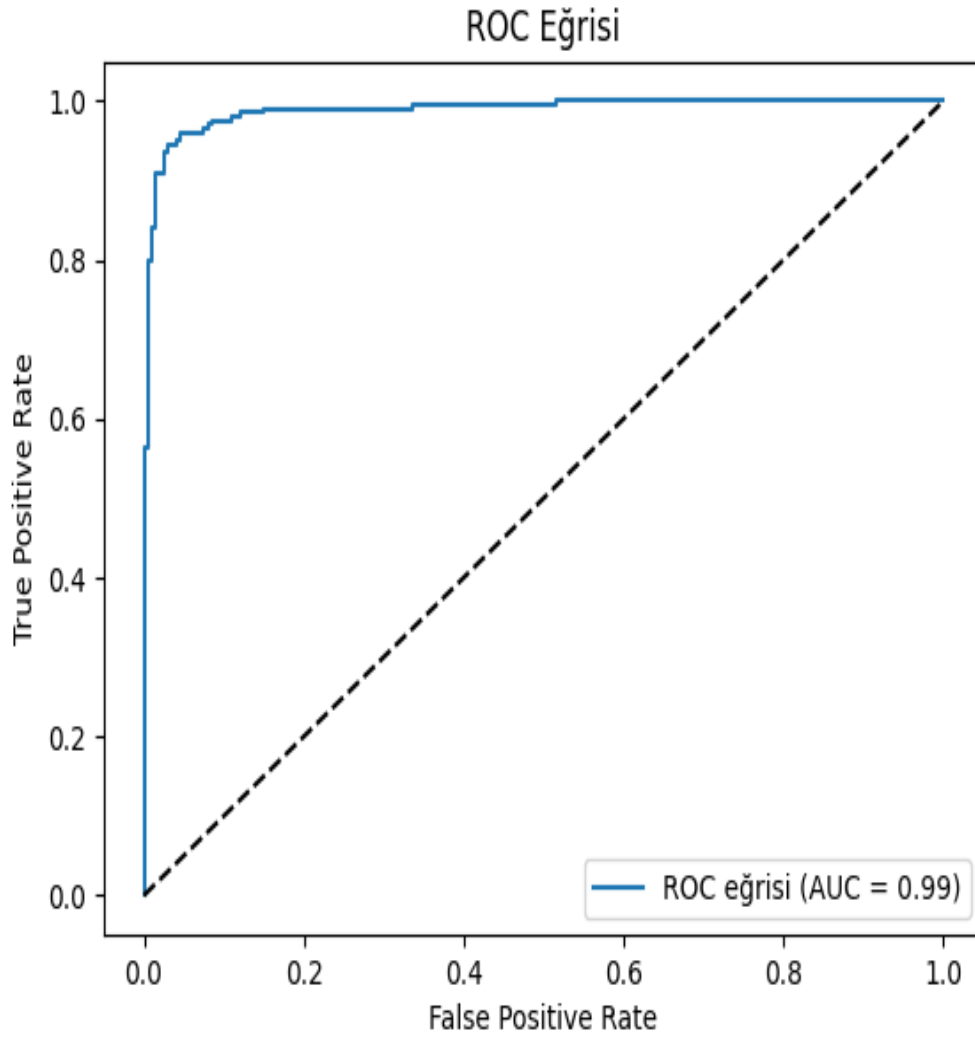
Şekil 4.17'de eğitim ve doğrulama aşamalarında elde edilen başarı ve kayıp değerlerinin epoch bazlı değişimi görülmektedir. Eğitim ve doğrulama başarıları, eğitim süreci boyunca 0.95 seviyelerine kadar yükselmiştir. Grafikler, eğitim ve doğrulama başarılarının birbirine yakın seyrettiğini ve belirgin bir aşırı öğrenme eğilimi göstermediğini ortaya koymaktadır. Kaybın hem eğitim hem de doğrulama verilerinde azalarak dengeli bir şekilde düşmesi, modelin hem öğrenme kabiliyetinin yüksek olduğunu hem de genelleme yeteneğinin iyi seviyede olduğunu göstermektedir.

Şekil 4.18’ de yer alan karışıklık matrisine göre, model boyalı sınıfı 195 kez doğru ve 5 kez yanlış tahmin etmiş; orijinal sınıfı ise 184 kez doğru ve 16 kez yanlış tahmin etmiştir. Bu sonuçlar, modelin her iki sınıf için de yüksek bir ayırt edebilme yeteneğine sahip olduğunu göstermektedir. Yanlış sınıflandırmalar düşük sayılarda kalmış ve sınıflar arası denge korunmuştur. Bu durum, modelin boyalı ve orijinal yüzeylerin özelliklerini iyi ayırt edebildiğine işaret etmektedir.



Şekil 4.18. MobileNetV2 mimarisinin karışıklık matrisi

Şekil 4.19’da verilen ROC eğrisi, modelin sınıflandırma başarısını eşik değerlerine göre değerlendirmektedir. AUC skoru 0.99 olarak elde edilmiştir. Bu değer, modelin neredeyse mükemmel bir ayırım gücüne sahip olduğunu göstermektedir. ROC eğrisinin üst sol köşeye yakın ilerlemesi, hem yüksek duyarlılık hem de düşük yanlış pozitif oranı sağlandığını ifade eder.



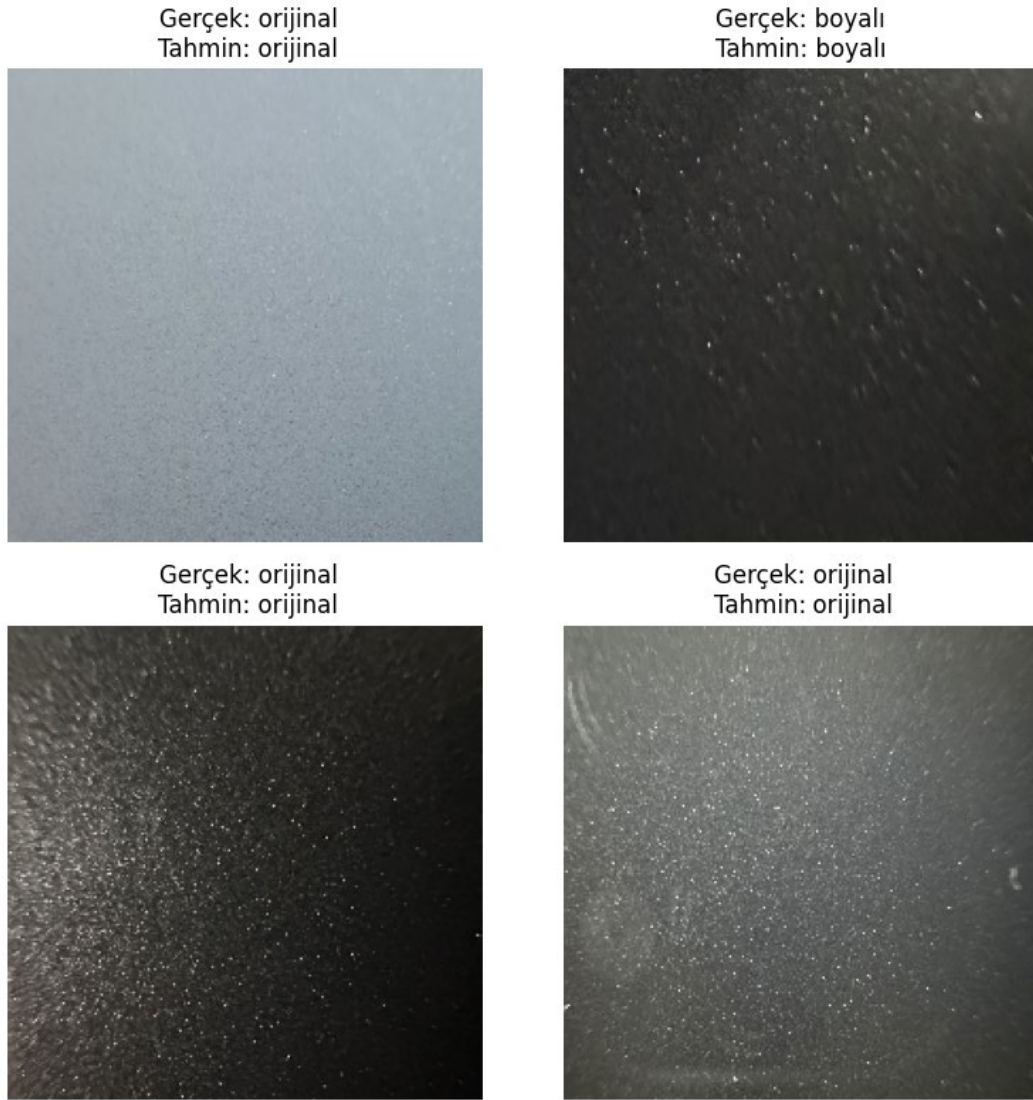
Şekil 4.19. MobileNetV2 mimarisi ROC grafiği

Tablo 4.6. MobileNetV2 ile sınıflandırma metrikleri

Sınıflar	Kesinlik	Duyarlılık	F-score	Veri sayısı
Boyalı	0.92	0.97	0.95	200
Orijinal	0.97	0.92	0.95	200

Tablo 4.6’da yer alan sınıflandırma raporu, modelin kesinlik, duyarlılık ve F-score gibi metriklerini ortaya koymaktadır. Boyalı sınıf için kesinlik 0.92, duyarlılık 0.97 ve F-score 0.95 olarak elde edilmiştir. Orijinal sınıfta ise kesinlik 0.97, duyarlılık 0.92 ve F-score yine 0.95 seviyesindedir. Genel doğruluk (accuracy) oranı 0.95 olarak elde edilmiştir. Bu sonuçlar, modelin her iki sınıfı da dengeli ve yüksek başarıyla sınıflandırabildiğini ortaya koymaktadır.

Şekil 4.20’de örnek doğrulama görüntüleri ve modelin tahminleri yer almaktadır. Görsellerde modelin gerçek etiketlerle uyumlu tahminler yaptığı gözlemlenmektedir. Hem orijinal hem de boyalı yüzeylerdeki tahminlerin doğru olması, modelin görsel dokuları ve renk farklılıklarını başarılı bir şekilde öğrenip ayırt edebildiğini göstermektedir.



Şekil 4.20. MobileNetV2 ile örnek tahmin görüntüleri

5. TARTIŞMA

Otomotiv endüstrisinde araç kaporta boya orijinalliğinin tespiti, ikinci el araç alım-satım işlemlerinde şeffaflık ve güvenilirlik sağlamada kritik bir rol oynamaktadır. Geleneksel yöntemler, genellikle uzman ustaların gözlemleri veya boya kalınlığı ölçüm cihazları gibi manuel yaklaşımlara dayanmakta, bu da insan hatasına ve yorum farklılıklarına açık sonuçlar doğurmaktadır (Park vd., 2016). Bu çalışmada önerilen derin öğrenme tabanlı sistem, bu sınırlamaları aşmayı hedefleyerek, görsel veriler üzerinden araç kaporta boyasının orijinal mi yoksa sonradan boyalı mı olduğunu yüksek doğrulukla belirleyebilen bir çözüm sunmuştur.

Literatürde yapılan araştırmalar, ESA temelli yaklaşımların görüntü sınıflandırma ve yüzey kusuru tespiti gibi görevlerde üstün performans sağladığını göstermektedir (Krizhevsky vd., 2012; He vd., 2016; Szegedy vd., 2015). Çalışmada kullanılan ResNet50, VGG16, InceptionV3, EfficientNetB0 ve MobileNetV2 modelleri, transfer öğrenme stratejileriyle birlikte kullanılarak, nispeten sınırlı sayıda (2000 görüntü) yerel veri üzerinde oldukça yüksek sınıflandırma doğruluğu elde etmiştir. Özellikle InceptionV3 modeli %98 doğruluk oranı ile öne çıkmıştır. Bu sonuçlar, transfer öğrenmenin sınırlı veri senaryolarında etkinliğini doğrular (Pan & Yang, 2010).

Diğer yandan, bu çalışmanın bulguları, ESA modellerinin yalnızca otomotiv sektöründe değil, kumaş kusuru tespiti (Jeyaraj & Nadar, 2019) veya endüstriyel yüzey kontrolü (Zhang vd., 2020) gibi farklı uygulama alanlarında da benzer şekilde başarılı sonuçlar verebileceğini göstermektedir. Böylece, önerilen yaklaşım yalnızca araç boya orijinalliğinin tespitinde değil, geniş bir endüstriyel kalite kontrol problem alanına genellenebilir bir yöntem olarak değerlendirilebilir.

Sonuç olarak, tezde geliştirilen derin öğrenme tabanlı sınıflandırma sistemi, geleneksel yöntemlerin zayıf yönlerini azaltarak, kullanıcıdan bağımsız ve daha tutarlı bir kalite kontrol aracı olarak öne çıkmaktadır. Bununla birlikte, daha geniş ve dengeli veri kümeleriyle eğitim yapılması ve farklı aydınlatma, kamera ayarı gibi koşulların da dikkate alınması, modelin pratikteki dayanıklılığını ve genellenebilirliğini daha da artıracaktır.

6. SONUÇLAR VE ÖNERİLER

6.1. Sonuçlar

Bu deneysel çalışmada, araç boya yüzeylerinin orijinal ve boyalı olarak sınıflandırılmasında derin öğrenme tabanlı transfer öğrenme yaklaşımları değerlendirilmiştir. MobileNetV2, InceptionV3, VGG16, ResNet50 ve EfficientNetB0 mimarileri kullanılarak gerçekleştirilen deneyler sonucunda, özellikle InceptionV3 (%98 doğruluk), MobileNetV2 (%95 doğruluk) ve VGG16 (%94 doğruluk) modellerinin yüksek performans sergilediği belirlenmiştir. Bu modeller, boyalı ve orijinal yüzeyler arasındaki ince dokusal ve renk farklılıklarını başarıyla ayırt edebilmiştir. Buna karşın, ResNet50 (%68 doğruluk) ve EfficientNetB0 (%50 doğruluk) mimarileri düşük performans göstermiş, özellikle veri sayısı ve veri çeşitliliğinin yetersizliği karşısında genelleme yetenekleri sınırlı kalmıştır. Bulgular, veri artırma ve katman dondurma stratejilerinin model performansını artırmada kritik rol oynadığını, ayrıca doğru öğrenme oranı ayarlarının aşırı öğrenme riskini azalttığını ortaya koymuştur.

Sonuç olarak, transfer öğrenme yöntemlerinin araç boya orijinalliği tespitinde etkin bir şekilde kullanılabileceği görülmüştür. Ancak daha geniş ve çeşitli veri kümeleri ile desteklendiğinde, bu modellerin gerçek dünyadaki güvenilirliği ve genelleme gücü daha da artırılabilir.

6.2. Öneriler

Bu çalışmanın temel hedefi, araç kaporta boya orijinalliğinin görsel veriler üzerinden otomatik olarak sınıflandırılmasını sağlamaktır. Ancak çalışma sınırlı sayıda ve belirli bir üretici grubuna ait veri setiyle yürütülmüş, çekim koşullarındaki farklılıklar, ışık değişimi ve kamera açıları gibi etmenler kontrol altında tutulamamıştır. Modelin genelleme yeteneğini artırmak amacıyla daha çeşitli ve dengeli veri kümeleriyle çalışmak, farklı çekim senaryolarını kapsamak ve gerçek zamanlı uygulama olanaklarını geliştirmek gelecekteki çalışmalar için hedeflenmektedir. Aşağıda çeşitli kısıtlamalar ve çözüm önerileri sunulmuştur:

Veri seti kısıtlamaları: Çalışmada kullanılan görüntü veri seti, her ne kadar sınıflandırma için uygun olacak şekilde etiketlenmiş olsa da, gerçek dünya koşullarını tam olarak temsil etmeyebilir. Çekim açıları, ışık koşulları, yüzey temizliği ve kamera kalitesi gibi faktörler, modelin eğitiminde kullanılan görsellerle gerçek uygulamadaki görseller arasında farklılıklar yaratabilir. Ayrıca veri setinin boyutu, derin öğrenme modellerinin öğrenme kapasitesi açısından sınırlı kalabilir. Bu kısıtlamalar, modelin gerçek ortamda beklenen genelleme performansını olumsuz yönde etkileyebilir.

Sınıf dengesizliği: Veri setindeki sınıfların ("boyalı" ve "orijinal") dağılımının dengeli olması sınıflandırma performansı için kritik öneme sahiptir. Sınıflar arasında örnek sayısı açısından bir dengesizlik olduğunda model, çoğunluk sınıfına aşırı uyum sağlayarak azınlık sınıftaki örnekleri doğru sınıflandıramaz. Çalışmada sınıf dengesini sağlamak amacıyla veri artırma (data augmentation) tekniklerinden yararlanılmış olsa da, dengesizlik tamamen ortadan kaldırılamamış olabilir. Bu durum, özellikle gerçek dünyadaki nadir boyalı yüzeylerin tespitinde yanlış sınıflama riskini artırabilir.

Eğitim–test oranı: Bu çalışmada veri seti eğitim (train) ve doğrulama (validation) olarak ikiye ayrılmıştır. Eğitim ve doğrulama oranları, modelin öğrenme yeteneğini optimize etmek için %80-%20 olarak belirlenmiştir. Bu ayırım, modelin aşırı öğrenme ve eksik öğrenme problemlerini azaltmak için kritik bir parametre olup, modelin yeni verilere uyum yeteneğini değerlendirmek amacıyla doğrulama sürecinde kullanılmıştır.

Hiperparametre seçimi: Model mimarisi ve eğitim süreci boyunca öğrenme oranı (learning rate), batch size, dropout oranı ve optimizasyon algoritması gibi hiperparametreler deneysel olarak ayarlanmıştır. Uygun hiperparametre seçimi model başarısını doğrudan etkiler. Bu çalışmada öğrenme oranı kademeli azaltılmış, dropout katmanları aşırı öğrenmeyi engellemek için eklenmiş ve Adam optimizasyon algoritması tercih edilmiştir. Ancak hiperparametre seçimi, veri setine özgü en uygun değerlerin belirlenmesi için daha ileri hiperparametre arama teknikleriyle geliştirilebilir.

Üretici firma boya çeşitliliği: Bu çalışmada kullanılan orijinal örneklerin büyük kısmı aynı üretici firmaya ait araçlardan elde edilmiştir. Bu durum, boya formülasyonu, pigment yapısı ve uygulama teknikleri açısından görece homojen bir veri kümesi oluşmasına yol açmıştır. Farklı üreticiler, kendi boya teknolojileri, renk seçenekleri ve yüzey finisajlarıyla önemli çeşitlilik gösterir. Dolayısıyla model, yalnızca bu üreticiye özgü özellikleri daha iyi öğrenirken, farklı marka ve modellerdeki boya varyasyonlarına genelleme yeteneği sınırlı kalabilir. Gelecek çalışmalarda, çok sayıda üretici firmaya ait

araçlardan toplanan orijinal ve boyalı örneklerle eğitilen bir modelin, üreticiden bağımsız daha güçlü bir sınıflandırma yeteneğine sahip olması hedeflenmelidir.

Sınıflandırma doğruluğunun artırılması: Sınıflandırma doğruluğunu artırmak için veri artırma stratejileri, önceden eğitilmiş modellerin katmanlarının uygun şekilde dondurulup açılması ve öğrenme oranı planlaması kullanılmıştır. İleri aşamalarda, daha büyük ve daha çeşitli veri setleri, otomatik hiperparametre optimizasyonu ve çoklu öğrenme gibi yöntemlerle başarı oranı daha da artırılabilir.

Genelleme yeteneğinin iyileştirilmesi: Mevcut veri seti, belirli bir coğrafi veya ticari segmentte toplanmış olabilir ve boya türleri bakımından sınırlı kalabilir. Farklı renk tonları, yaşlanma seviyeleri, tamir sonrası yüzeyler ve farklı çekim cihazlarıyla oluşturulmuş veri setlerinin entegre edilmesi, modelin sınıflandırma başarımını artırmak için önemlidir. Gelecek çalışmalarda bu çeşitliliğin sağlanması hedeflenmelidir.

KAYNAKLAR

- Ben-David, S., Blitzer, J., Crammer, K., Kulesza, A., Pereira, F., & Vaughan, J. W. (2010). A theory of learning from different domains. *Machine Learning*, 79(1–2), 151–175.
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *NeurIPS*, 33, 1877–1901.
- Chen, W., Tan, X., Zhang, J., Du, G., Fu, Q., & Jiang, H. (2024). A robust approach for multi-type classification of brain tumor using deep feature fusion. *Frontiers in Neuroscience*, 18, 1288274.
- Cheng, P., Cui, A., Yang, Y., Luo, Y., & Sun, W. (2017). Recognition and classification of coating film defects on automobile body based on image processing. *Proceedings of the 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics*, 1–5.
- Chollet, F. (2017). *Deep learning with Python*. Manning Publications.
- Chung, Y. C., & Chang, M. (2006). Visualization of subtle defects of car body outer panels. *SICE-ICASE International Joint Conference 2006*, 4639–4642.
- Clevert, D. A., Unterthiner, T., & Hochreiter, S. (2016). Fast and accurate deep network learning by exponential linear units (ELUs). *arXiv preprint arXiv:1511.07289*. <https://doi.org/10.48550/arXiv.1511.07289>
- Dumoulin, V., & Visin, F. (2016). A guide to convolution arithmetic for deep learning. *arXiv preprint arXiv:1603.07285*.
- Edris, M. Z. B., Jawad, M. S., & Zakaria, Z. (2015). Surface defect detection and neural network recognition of automotive body panels. *2015 IEEE International Conference on Control System, Computing and Engineering*, 117–122.
- Eichhorn, A., Girimonte, D., Klose, A., & Kruse, R. (2005). Soft computing for automated surface quality analysis of exterior car body panels. *Applied Soft Computing Journal*, 5, 301–313.
- Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36(4), 193–202.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Regularization for deep learning. In *Deep learning* (pp. 216–261).

- Goodfellow, I. J., Warde-Farley, D., Mirza, M., Courville, A., & Bengio, Y. (2013). Maxout networks. *Proceedings of the 30th International Conference on Machine Learning (ICML-13)*, 1319–1327. <https://proceedings.mlr.press/v28/goodfellow13.pdf>
- Habek, G. C., Tasdemir, S., Basciftci, F., & Yılmaz, A. (2024). Transfer derin öğrenme teknikleri ile görüntü sınıflandırmada aktivasyon fonksiyonlarının performans üzerindeki etkisi. *Afyon Kocatepe Üniversitesi Fen Ve Mühendislik Bilimleri Dergisi*, 24(2), 294-307.
- Han, J., & Moraga, C. (1995). The influence of the sigmoid function parameters on the speed of backpropagation learning. In *From Natural to Artificial Neural Computation* (pp. 195–201). Springer. https://doi.org/10.1007/3-540-59497-3_175
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778.
- Hidaka, A., & Kurita, T. (2017). Consecutive dimensionality reduction by canonical correlation analysis for visualization of convolutional neural networks. *Proceedings of the ISCIE International Symposium on Stochastic Systems Theory and its Applications*, 160–167. <https://doi.org/10.5687/sss.2017.160>
- Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2012). Improving neural networks by preventing co-adaptation of feature detectors. *arXiv preprint arXiv:1207.0580*. <https://arxiv.org/abs/1207.0580>
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., ... & Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*. <https://arxiv.org/abs/1704.04861>
- Howard, A., Sandler, M., Chu, G., Chen, L. C., Chen, B., Tan, M., ... & Le, Q. V. (2019). Searching for MobileNetV3. *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 1314–1324. <https://doi.org/10.1109/ICCV.2019.00140>
- Hu, H., Li, Y., Liu, M., & Liang, W. (2014). Classification of defects in steel strip surface based on multiclass support vector machine. *Multimedia Tools and Applications*, 69(1), 199–216.
- Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 7132–7141. <https://doi.org/10.1109/CVPR.2018.00745>
- Jeyaraj, P. R., & Samuel Nadar, E. R. (2019). Computer vision for automatic detection and classification of fabric defect employing deep learning algorithm. *International Journal of Clothing Science and Technology*, 31(4), 510–521.
- Kamani, P., Afshar, A., Towhidkhah, F., & Roghani, E. (2011). Car body paint defect inspection using rotation invariant measure of the local variance and one-against-

- all support vector machine. *2011 First International Conference on Informatics and Computational Intelligence*, 244–249.
- Kang, X. (2024). Research on fabric defect detection method based on lightweight network. *Journal of Engineered Fibers and Fabrics*, 19.
- Karangwa, J., Kong, L., You, T., & Zheng, J. (2020). Automated surface defects detection on mirrorlike materials by using Faster R-CNN. *2020 7th International Conference on Information Science and Control Engineering*, 2288–2294.
- Kaur, P., Kumar, Y., Ahmed, S., Alhumam, A., Singla, R. et al. (2022). Automatic license plate recognition system for vehicles using a CNN. *Computers, Materials & Continua*, 71(1), 35–50.
- Kalyoncu, S. (2020). Deep learning networks for stock market analysis. İstanbul Sabahattin Zaim Üniversitesi.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25, 1097–1105.
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324.
- LeCun, Y., Bottou, L., Orr, G. B., & Müller, K. R. (2012). Efficient BackProp. In G. Montavon, G. B. Orr, & K. R. Müller (Eds.), *Neural Networks: Tricks of the Trade* (pp. 9–48). Springer.
- Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013). Rectifier nonlinearities improve neural network acoustic models. *Proceedings of the 30th International Conference on Machine Learning (ICML)*.
https://ai.stanford.edu/~amaas/papers/relu_hybrid_icml2013_final.pdf
- Molina, J., Solanes, J. E., Arnal, L., & Tornero, J. (2017). On the detection of defects on specular car body surfaces. *Robotics and Computer-Integrated Manufacturing*, 48, 263–278.
- Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted Boltzmann machines. *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, 807–814.
- Neurohive. (2018). VGG16 – Convolutional network for classification and detection. Retrieved from <https://neurohive.io/en/popular-networks/vgg16/>
- Nwankpa, C., Ijomah, W., Gachagan, A., & Marshall, S. (2018). Activation functions: Comparison of trends in practice and research for deep learning. *arXiv preprint arXiv:1811.03378*. <https://doi.org/10.48550/arXiv.1811.03378>
- Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359.

- Parisi, G. I., Kemker, R., Part, J. L., Kanan, C., & Wermter, S. (2019). Continual lifelong learning with neural networks: A review. *Neural Networks*, 113, 54–71.
- Park, J. K., Kwon, B. K., Park, J. H., & Kang, D. J. (2016). Machine learning-based imaging system for surface defect inspection. *International Journal of Precision Engineering and Manufacturing-Green Technology*, 3(3), 303–310.
- Reddy, A. S. B., & Juliet, D. S. (2019). Transfer learning with ResNet-50 for malaria cell-image classification. *2019 International Conference on Communication and Signal Processing (ICCSP)*, 945–949. <https://doi.org/10.1109/ICCSP.2019.8697909>
- Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, 2(11), 1019–1025.
- Rojas-Aranda, J. L., Nunez-Varela, J. I., Cuevas-Tello, J. C., & Rangel-Ramirez, G. (2020). Fruit classification for retail stores using deep learning. In A. E. Bayro-Corrochano & E. Hancock (Eds.), *Pattern recognition. MCPR 2020. Lecture Notes in Computer Science* (Vol. 12241, pp. 3–13). Springer. https://doi.org/10.1007/978-3-030-49076-8_1
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4510–4520. <https://doi.org/10.1109/CVPR.2018.00474>
- Shahoveisi, F., Gorji, H. T., Shahabi, S., Hosseinirad, S., Markell, S., & Vasefi, F. (2023). Application of image processing and transfer learning for the detection of rust disease. *Scientific Reports*, 13(1), 5133. <https://doi.org/10.1038/s41598-023-31942-9>
- Shin, H. C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., ... & Summers, R. M. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Transactions on Medical Imaging*, 35(5), 1285–1298. <https://doi.org/10.1109/TMI.2016.2528162>
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*. <https://arxiv.org/abs/1409.1556>
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1–9. <https://doi.org/10.1109/CVPR.2015.7298594>
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2818–2826. <https://doi.org/10.1109/CVPR.2016.308>

- Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. *International Conference on Machine Learning (ICML)*, 6105–6114. <https://arxiv.org/abs/1905.11946>
- Wang, C., Sun, Q., Dong, X., & et al. (2024). Automotive adhesive defect detection based on improved YOLOv8. *SIViP*, 18, 2583–2595.
- Wei, X., Jiang, S., Li, Y., Li, C., Jia, L., & Li, Y. (2020). Defect detection of pantograph slide based on deep learning and image processing technology. *IEEE Transactions on Intelligent Transportation Systems*, 21(3), 947–958.
- Xu, F., Liu, Y., Zi, B., & Zheng, L. (2021). Application of deep learning for defect detection of paint film. *2021 6th International Conference on Intelligent Computing and Signal Processing*, 1118–1121.
- Xu, J., Zhang, J., Zhang, K., Li, Y., & Zhang, Y. (2020). An APF-ACO algorithm for automatic defect detection on vehicle paint. *Multimedia Tools and Applications*, 79(33), 25315–25333.
- Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks? *NeurIPS*, 3320–3328.
- Zhang, J., Xu, J., Zhu, L., Ren, Y., Qiu, Z., & Zhao, L. (2020). An improved MobileNet-SSD algorithm for automatic defect detection on vehicle body paint. *Multimedia Tools and Applications*, 79, 23367–23385.
- Zhao, L., Li, F., Zhang, Y., Xu, X., Xiao, H., & Feng, Y. (2020). A deep-learning-based 3D defect quantitative inspection system in CC products surface. *Sensors*, 20(4), 980.
- Zheng, G. (2021). A novel attention-based convolution neural network for human activity recognition. *IEEE Sensors Journal*, 21(23), 27015–27025.