

T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

GERİ DÖNÜŞTÜRÜLEBİLİR ATIKLARIN DERİN ÖĞRENME
MODELLERİ İLE SINIFLANDIRILMASI: VERİ SETİ
BOYUTUNUN ETKİSİ ÜZERİNE BİR KARŞILAŞTIRMA

YÜKSEK LİSANS TEZİ

Gamze AĞIRTAŞ

Bilgisayar Mühendisliği Anabilim Dalı

HAZİRAN 2025

T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

GERİ DÖNÜŞTÜRÜLEBİLİR ATIKLARIN DERİN ÖĞRENME
MODELLERİ İLE SINIFLANDIRILMASI: VERİ SETİ
BOYUTUNUN ETKİSİ ÜZERİNE BİR KARŞILAŞTIRMA

YÜKSEK LİSANS TEZİ

Gamze AĞIRTAŞ

Bilgisayar Mühendisliği Anabilim Dalı

Tez Danışmanı: Dr. Öğr. Üyesi Serap ÇAKAR KAMAN

HAZİRAN 2025

Gamze AĞIRTAŞ tarafından hazırlanan “GERİ DÖNÜŞTÜRÜLEBİLİR ATIKLARIN DERİN ÖĞRENME MODELLERİ İLE SINIFLANDIRILMASI: VERİ SETİ BOYUTUNUN ETKİSİ ÜZERİNE BİR KARŞILAŞTIRMA ” adlı tez çalışması 27.06.2025 tarihinde aşağıdaki jüri tarafından oy birliği ile Sakarya Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı Yüksek Lisans tezi olarak kabul edilmiştir.

Tez Jürisi

Jüri Başkanı : Unvan Adı SOYADI
Sakarya Üniversitesi

Jüri Üyesi : Unvan Adı SOYADI (Danışman)
Sakarya Üniversitesi

Jüri Üyesi : Unvan Adı SOYADI
Sakarya Üniversitesi



ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ

Sakarya Üniversitesi Fen Bilimleri Enstitüsü Lisansüstü Eğitim-Öğretim Yönetmeliğine ve Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesine uygun olarak hazırlamış olduğum “GERİ DÖNÜŞTÜRÜLEBİLİR ATIKLARIN DERİN ÖĞRENME MODELLERİ İLE SINIFLANDIRILMASI: VERİ SETİ BOYUTUNUN ETKİSİ ÜZERİNE BİR KARŞILAŞTIRMA” başlıklı tezin bana ait, özgün bir çalışma olduğunu; çalışmamın tüm aşamalarında yukarıda belirtilen yönetmelik ve yönergeye uygun davrandığımı, tezin içerdiği yenilik ve sonuçları başka bir yerden almadığımı, tezde kullandığım eserleri usulüne göre kaynak olarak gösterdiğimi, bu tezi başka bir bilim kuruluna akademik amaç ve unvan almak amacıyla vermediğimi ve 20.04.2016 tarihli Resmi Gazete’de yayımlanan Lisansüstü Eğitim ve Öğretim Yönetmeliğinin 9/2 ve 22/2 maddeleri gereğince Sakarya Üniversitesi’nin abonesi olduğu intihal yazılım programı kullanılarak Enstitü tarafından belirlenmiş ölçütlere uygun rapor alındığını, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun ortaya çıkması halinde doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi beyan ederim.

(...../...../20.....).

Gamze AĞIRTAŞ





Biricik aileme..



TEŞEKKÜR

Bu tez çalışması, Sakarya Üniversitesi Bilgisayar ve Bilişim Bilimleri Fakültesi Bilgisayar Mühendisliği Yüksek Lisans Programı kapsamında hazırlanmıştır. Çalışmanın her aşamasında değerli bilgi ve tecrübelerini paylaşan ve yol gösteren danışman hocam Dr. Öğr. Üyesi Serap ÇAKAR KAMAN'a sonsuz teşekkürlerimi sunarım. Ayrıca eğitim hayatım boyunca bana her zaman destek olan aileme minnettarım.

Gamze AĞIRTAŞ



İÇİNDEKİLER

Sayfa

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ	v
TEŞEKKÜR	ix
İÇİNDEKİLER	xi
KISALTMALAR	xiii
SİMGELER	xv
ŞEKİL LİSTESİ	xvii
ÖZET	xix
SUMMARY	xxi
1. GİRİŞ.....	1
1.1. Tezin Kapsamı.....	2
1.2. Tezin Amacı	2
1.3. Problemin Tanımı.....	2
1.4. Çalışmanın Kapsamı	3
1.5. Çalışmanın Yapısı	3
1.6. Literatür	4
2. MATERYAL VE YÖNTEM.....	7
2.1. Kullanılan Veri Kümeleri.....	7
2.1.1. Garbage Classification veri kümesi	7
2.1.2. Garbage Dataset veri kümesi	8
2.2. Ön İşleme Aşamaları	9
2.2.1. Görüntü boyutlandırma	9
2.2.2. Veri arttırımı yöntemleri	9
2.3. Transfer Öğrenme Yaklaşımı	10
2.4. Kullanılan Derin Öğrenme Modelleri	11
2.4.1. VGG16 modeli	11
2.4.2. MobileNetV2 modeli	11
2.4.3. InceptionV3 modeli.....	11
2.4.4. DenseNet121 modeli.....	12
2.5. Model Eğitim Stratejisi	12
2.6. Özellik Füzyonlu Hibrit Model	13
3. KULLANILAN TEKNOLOJİLER VE YAZILIMLAR.....	17
3.1. Programlama Dili ve Geliştirme Ortamı	17
3.2. Kullanılan Kütüphaneler ve Frameworkler	17
3.3. Donanım ve Çalışma Ortamı	18
4. MODEL EĞİTİMİ VE DEĞERLENDİRME METRİKLERİ.....	21
4.1. Eğitim Süreci	21
4.2. Beş Katlı Çapraz Doğrulama Metrikleri	21
4.3. Kullanılan Değerlendirme Metrikleri	22
4.3.1. Doğruluk değeri	22
4.3.2. Kayıp değeri	23
4.3.3. Kesinlik, geri çağırma, F1-skoru.....	24

4.3.4. ROC eğrisi altında kalan alan.....	25
5. SONUÇLAR VE GÖRSEL ANALİZLER	27
5.1. Eğitim ve Doğrulama Başarıları	27
5.1.1. VGG16 Garbage Classification veri kümesi	27
5.1.2. VGG16 Garbage Dataset veri kümesi	28
5.1.3. MobileNetV2 Garbage Classification veri kümesi	29
5.1.4. MobileNetV2 Garbage Dataset veri kümesi	30
5.1.5. InceptionV3 Garbage Classification veri kümesi.....	31
5.1.6. InceptionV3 Garbage Dataset veri kümesi.....	32
5.1.7. DenseNet121 Garbage Classification veri kümesi.....	33
5.1.8. DenseNet121 Garbage Dataset veri kümesi.....	34
5.2. Test Sonuçları ve Katlama Bazlı Metrikler	35
5.2.1. VGG16 Garbage Classification veri kümesi katlama bazlı test sonuçları.....	35
5.2.2. VGG16 Garbage Dataset veri kümesi katlama bazlı test sonuçları	36
5.2.3. MobileNetV2 Garbage Classification veri kümesi katlama bazlı test sonuçları	37
5.2.4. MobileNetV2 Garbage Dataset veri kümesi katlama bazlı test sonuçları	38
5.2.5. InceptionV3 Garbage Classification katlama bazlı test sonuçları.....	39
5.2.6. InceptionV3 Garbage Dataset veri kümesi katlama bazlı test sonuçları ..	40
5.2.7. DenseNet121 Garbage Classification veri kümesi katlama bazlı test sonuçları	41
5.2.8. DenseNet121 Garbage Dataset veri kümesi katlama bazlı test sonuçları ..	42
5.3. Eğitim ve Doğrulama Grafiklerinin Analizi.....	43
5.4. Modeller Arası Performans Karşılaştırması	44
5.5. Hibrit Modelin Performansa Etkisi	47
5.6. Karmaşıklık Matrisleri.....	48
5.7. Garbage Classification Veri kümesi Hibrit Model Başarısı	51
5.8. Garbage Dataset Veri kümesi Hibrit Model Başarısı	52
5.9. Garbage Classification Veri kümesi Hibrit Model Performans Metrikleri Sonucu	53
5.10. Garbage Dataset Veri kümesi Hibrit Model Performans Metrikleri Sonucu ..	54
6. SONUÇLAR	57
KAYNAKLAR.....	59
ÖZGEÇMİŞ.....	63

KISALTMALAR

AI	: Artificial Intelligence
API	: Application Programming Interface
AUC	: Area Under Curve
CNN	: Convolutional Neural Network
DL	: Deep Learning
FN	: False Negative (Yanlış Negatif)
FP	: False Positive (Yanlış Pozitif)
GPU	: Graphics Processing Unit (Grafik İşlem Birimi)
ReLU	: Rectified Linear Unit (Doğrultulmuş Doğrusal Birim)
RGB	: Red Green Blue (Kırmızı Yeşil Mavi Renk Modeli)
ROC	: Receiver Operating Characteristic
TL	: Transfer Learning (Transfer Öğrenme)
TN	: True Negative (Doğru Negatif)
FP	: Yanlış Pozitif sınıf tahmin sayısı
FN	: Yanlış Negatif sınıf tahmin sayısı
TP	: True Positive (Doğru Pozitif)
TPU	: Tensor Processing Unit (Tensör İşlem Birimi)
VGG16	: Visual Geometry Group 16 Katmanlı Ağ



SİMGELER

P	: Kesinlik (Precision)
R	: Geri Çağırma (Recall)
F1	: F1 Skoru
A	: Doğruluk (Doğruluk)
L	: Kayıp Fonksiyonu (Loss)
$f(x)$	: Girdi verisine uygulanan tahmin fonksiyonu
y	: Gerçek etiket değeri
$\hat{y}$	: Model tarafından tahmin edilen çıktı
k	: Kat sayısı (k-fold çapraz doğrulamadaki fold sayısı)
n	: Toplam örnek sayısı
i	: i'nci örneği ifade eden indis
j	: j'nci sınıfı ifade eden indis
t	: Eğitim sürecindeki epoch sayısı
E	: Eğitim seti
V	: Doğrulama (validation) seti
T	: Test seti



ŞEKİL LİSTESİ

Sayfa

Şekil 2.1. Garbage Classification veri kümesinden örnek görüntüler.	8
Şekil 2.2. Garbage Dataset veri kümesinden örnek görüntüler.	9
Şekil 2.3. Model akış diyagramı.	13
Şekil 2.4. Özellik füzyon hibrit modeli akış diyagramı.	15
Şekil 5.1. VGG16 modeli için Garbage Classification veri kümesinde eğitim ve doğrulama doğruluk grafiği.	28
Şekil 5.2. VGG16 modeli için Garbage Dataset veri kümesinde eğitim ve doğrulama doğruluk grafiği.	29
Şekil 5.3. MobileNetV2 modeli için Garbage Classification veri kümesinde eğitim ve doğrulama doğruluk grafiği.	30
Şekil 5.4. MobileNetV2 modeli için Garbage Dataset veri kümesinde eğitim ve doğrulama doğruluk grafiği.	31
Şekil 5.5. InceptionV3 modeli için Garbage Classification veri kümesinde eğitim ve doğrulama doğruluk grafiği.	32
Şekil 5.6. InceptionV3 modeli için Garbage Dataset veri kümesinde eğitim ve doğrulama doğruluk grafiği.	33
Şekil 5.7. DenseNet121 modeli için Garbage Classification veri kümesinde eğitim ve doğrulama doğruluk grafiği.	34
Şekil 5.8. DenseNet121 modeli için Garbage Dataset veri kümesinde eğitim ve doğrulama doğruluk grafiği.	35
Şekil 5.9. VGG16 modeli için Garbage Classification veri kümesinde katlama bazlı test doğruluğu, kayıp ve performans metrikleri.	36
Şekil 5.10. VGG16 modeli için Garbage Dataset veri kümesinde katlama bazlı test doğruluğu, kayıp ve performans metrikleri.	37
Şekil 5.11. MobileNetV2 modeli için Garbage Classification veri kümesinde katlama bazlı test doğruluğu, kayıp ve performans metrikleri.	38
Şekil 5.12. MobileNetV2 modeli için Garbage Dataset veri kümesinde katlama bazlı test doğruluğu, kayıp ve performans metrikleri.	39
Şekil 5.13. InceptionV3 modeli için Garbage Classification veri kümesinde katlama bazlı test doğruluğu, kayıp ve performans metrikleri.	40
Şekil 5.14. InceptionV3 modeli için Garbage Dataset veri kümesinde katlama bazlı test doğruluğu, kayıp ve performans metrikleri.	41
Şekil 5.15. DenseNet121 modeli için Garbage Classification veri kümesinde katlama bazlı test doğruluğu, kayıp ve performans metrikleri.	42
Şekil 5.16. DenseNet121 modeli için Garbage Dataset veri kümesinde katlama bazlı test doğruluğu, kayıp ve performans metrikleri.	43
Şekil 5.17. Modellerin test doğruluk karşılaştırması.	45
Şekil 5.18. Modellerin değerlendirme metriklerinin karşılaştırması.	47
Şekil 5.19. Veri setlerinin model performans metrikleri.	48
Şekil 5.20. Garbage Classification en iyi model için karmaşıklık matrisi.	49
Şekil 5.21. Garbage Dataset en iyi model için karmaşıklık matrisi.	50

Şekil 5.22. Garbage Classification veri kümesi için hibrit modelin eğitim ve doğrulama grafiği.	52
Şekil 5.23. Garbage Dataset veri kümesi için hibrit modelin eğitim ve doğrulama grafiği.	53
Şekil 5.24. Garbage Classification veri kümesi hibrit modelin performans metrikleri sonucu.....	54
Şekil 5.25. Garbage Dataset veri kümesi hibrit modelin performans metrikleri sonucu.	56



GERİ DÖNÜŞTÜRÜLEBİLİR ATIKLARIN DERİN ÖĞRENME MODELLERİ İLE SINIFLANDIRILMASI: VERİ SETİ BOYUTUNUN ETKİSİ ÜZERİNE BİR KARŞILAŞTIRMA

ÖZET

Geri dönüşüm, çevresel sürdürülebilirliğin sağlanmasında temel bir rol oynamakta ve doğal kaynakların korunması ile atık yönetiminin etkinliğinin artırılmasına doğrudan katkı sağlamaktadır. Artan nüfus ve sanayileşme ile birlikte meydana gelen atık miktarındaki ciddi artış, bu atıkların doğru şekilde ayrıştırılmasını ve geri dönüştürülmesini zorunlu hale getirmiştir. Atık sınıflandırma süreci, geri dönüşüm sistemlerinin verimliliği üzerinde doğrudan etkili bir faktör olmasına rağmen, geleneksel yöntemler yüksek maliyet, zaman kaybı ve insan hatasına açık olma gibi çeşitli kısıtlar taşımaktadır. Bu nedenle, atık sınıflandırma sürecinin otomasyon yoluyla desteklenmesi büyük önem arz etmektedir. Derin öğrenme tabanlı modeller, son yıllarda otomatik atık sınıflandırma alanında umut vaat eden sonuçlar vermiştir.

Bu çalışmada, farklı boyutlardaki veri kümeleri üzerinde, derin öğrenme tabanlı sınıflandırma modellerinin performansı detaylı bir şekilde incelenmiştir. Çalışmada iki farklı veri kümesi kullanılmıştır: Garbage Classification ve Garbage-Dataset. Garbage Classification veri kümesi, 2.890 görüntüden oluşmakta olup sınıf dağılımı açısından dengesiz bir yapı sergilemektedir. Bu dengesizlik, eğitim sürecinde uygulanan veri artırma teknikleri ile dengelenmiştir. Garbage-Dataset ise 9.660 görüntü içermekte ve sınıflar arasında daha dengeli bir dağılım sunarak modellerin genelleme kabiliyetlerinin daha sağlıklı bir şekilde değerlendirilebilmesine imkân tanımaktadır.

Çalışmanın temel amacı, veri kümesi boyutlarının derin öğrenme modellerinin sınıflandırma performansları üzerindeki etkisini analiz etmek ve her bir veri kümesi için en iyi sonuç veren modeli belirlemektir. Atık sınıflandırma görevinde, literatürde sıkça kullanılan dört farklı Evrişimsel Sinir Ağı (CNN) mimarisi tercih edilmiştir: DenseNet121, InceptionV3, MobileNetV2 ve VGG16. Modeller, ImageNet veri kümesi üzerinde önceden eğitilmiş olup transfer öğrenme teknikleri kullanılarak hedef göreve adapte edilmiştir. Model performanslarının daha güvenilir bir şekilde değerlendirilebilmesi amacıyla beş katlı çapraz doğrulama yöntemi uygulanmıştır. Ayrıca, Garbage Classification veri kümesindeki sınıf dengesizliği, veri artırma stratejileri kullanılarak minimize edilmiştir.

Buna ek olarak, çalışmanın son bölümünde özellik füzyon temelli hibrit modelleme yaklaşımı uygulanmıştır. VGG16, InceptionV3, DenseNet121 ve MobileNetV2 modellerinin önceden eğitilmiş katmanlarından elde edilen özellikler birleştirilerek yeni bir hibrit model tasarlanmıştır. Bu yöntem ile her modelin güçlü yönlerinin aynı mimaride temsil edilmesi hedeflenmiştir. Garbage Classification veri kümesi üzerinde yapılan hibrit modelleme sonucunda, doğruluk oranı %85,88'e ulaşarak temel modellerin üzerinde performans sergilemiştir. Garbage-Dataset üzerinde ise hibrit model %91,26 doğruluk sağlamış ve bu yönüyle bazı temel modellerle karşılaştırılabilir başarıda olmuştur. Deneysel sonuçlara göre analiz edilen iki veri

kümesinde de hibrit model, Doğruluk, Precision, Recall, F1 Skor ve ROC AUC gibi performans metriklerinde en üstün sonuçları elde etmiştir.

Deneyisel sonuçlar, veri kümesi boyutunun sınıflandırma başarımı üzerinde önemli bir etkisi olduğunu açıkça ortaya koymuştur. Garbage Classification veri kümesi üzerinde yapılan analizlerde, tekil modellerde en yüksek doğruluk oranı %85,05 ile VGG16 modeli tarafından elde edilmiştir. DenseNet121 %82,63, MobileNetV2 %82,49 ve InceptionV3 %80,42 doğruluk oranları ile VGG16'yı takip etmiştir. Garbage-Dataset üzerinde yapılan değerlendirmelerde ise en yüksek doğruluk oranı %90,75 ile DenseNet121 modeli tarafından sağlanmıştır. Bu modeli sırasıyla MobileNetV2 (%89,67), InceptionV3 (%88,41) ve VGG16 (%86,09) takip etmiştir. Hibrit modellerle birlikte elde edilen sonuçlar, model mimarisi seçimlerinin ve özellik birleştirme stratejilerinin sınıflandırma başarısına önemli katkılar sağladığını göstermektedir.

Bu çalışma, atık sınıflandırma alanında derin öğrenme yöntemlerinin potansiyelini ortaya koymakta ve otomatik atık yönetimi sistemlerinin geliştirilmesine yönelik gelecekteki araştırmalar için önemli bir temel sunmaktadır.

CLASSIFICATION OF RECYCLABLE WASTES WITH DEEP LEARNING MODELS: A COMPARISON ON THE EFFECT OF DATASET SIZE

SUMMARY

Recycling plays a crucial role in promoting environmental sustainability by reducing the demand for natural resources, minimizing waste accumulation, and limiting the emission of pollutants into the environment. As global populations increase and consumption habits intensify, the management of waste has become a critical challenge for governments, industries, and communities worldwide. One of the most essential steps in effective waste management is the proper classification of recyclable materials. This step ensures that waste streams are sorted efficiently, facilitating easier recycling and preventing reusable materials from being lost in landfills. Traditionally, waste classification has relied on manual processes that are time-consuming, error-prone, labor-intensive, and expensive. As such, there is a growing interest in automating waste classification processes, and recent advances in deep learning have provided promising approaches to address this need. This thesis explores the use of deep learning models for the classification of recyclable waste materials based on visual input. Convolutional Neural Networks (CNNs), which are known for their strength in image classification tasks, are employed and evaluated for their performance in distinguishing between different types of recyclable waste. The research examines how dataset size, class balance, and model architecture affect classification performance. Two publicly available datasets—Garbage Classification and Garbage-Dataset—were utilized in order to test and compare the performance of different CNN architectures under varying data conditions. The Garbage Classification dataset includes a total of 2,890 images divided into six classes: cardboard, can, metal, paper, plastic, and trash. This dataset is characterized by significant class imbalance, with some categories containing many more samples than others. Such imbalance can negatively affect model training, as the network may become biased toward more dominant classes. To address this issue, data augmentation techniques were employed, including random rotation, horizontal flipping, and image scaling. These augmentations were applied to artificially expand the dataset and enhance its diversity, helping the model generalize better during training. Additionally, image resizing and normalization were applied to prepare the images for input into the CNN models. The second dataset, Garbage-Dataset, contains 9,660 images and offers a more balanced class distribution across the same categories. With more samples and greater variability, this dataset serves as a benchmark for evaluating model generalization and robustness under more realistic conditions. It includes various background environments and object perspectives, making the classification task more challenging but also more representative of real-world scenarios. This dataset allows for a deeper analysis of model performance and helps to uncover the strengths and weaknesses of different network architectures. In this research, four state-of-the-art CNN architectures were selected for evaluation: VGG16, InceptionV3, MobileNetV2, and DenseNet121. These models were chosen due to their widespread usage, architectural differences, and proven performance in image classification benchmarks such as ImageNet. To accelerate training and improve model accuracy, transfer learning was

applied. The models were initialized with pre-trained weights from ImageNet and fine-tuned using the recyclable waste datasets. This approach allows the models to retain general visual features learned from large-scale image datasets while adapting to the specific characteristics of waste images. To ensure robust evaluation, a five-fold cross-validation strategy was implemented for each model and dataset. Classification performance was measured using accuracy as the primary metric. However, to provide a more comprehensive evaluation, additional performance indicators such as precision, recall, and F1-score were calculated. These metrics were particularly important for understanding model performance on underrepresented classes in the imbalanced dataset. Confusion matrices and class-wise performance metrics were also used to highlight common misclassifications and analyze the causes behind them. The experimental results show clear differences in model performance depending on the dataset size and structure. On the Garbage Classification dataset, which is smaller and more imbalanced, the VGG16 model achieved the best accuracy of 85.05%, outperforming the other architectures. DenseNet121, MobileNetV2, and InceptionV3 followed with accuracies of 82.63%, 82.49%, and 80.42%, respectively. This suggests that simpler models such as VGG16 may be more suitable for smaller datasets due to their lower risk of overfitting. In contrast, on the larger Garbage-Dataset, the performance rankings changed. The DenseNet121 model led with an accuracy of 90.75%, followed by MobileNetV2 at 89.67%, InceptionV3 at 88.41%, and VGG16 at 86.09%. These results indicate that deeper and more complex models benefit from the availability of more data, which allows them to utilize their larger capacity to learn intricate patterns and features. It also confirms the importance of choosing a model architecture that matches the scale and characteristics of the available training data. Beyond raw accuracy, performance on individual classes revealed interesting insights. For example, certain classes such as cardboard and glass were consistently well-identified across all models, while metal and trash were more prone to misclassification. These errors can often be attributed to the visual similarity between certain waste types or the presence of background clutter in the images. Further error analysis showed that class confusion often occurred between plastic and metal, and between paper and cardboard, likely due to overlapping textures and colors. To improve model interpretability, techniques such as Grad-CAM (Gradient-weighted Class Activation Mapping) were explored. These techniques help visualize which parts of an image the model focuses on during classification, offering transparency and aiding in debugging model decisions. Such explainability is critical in practical applications where users need to trust and understand model outputs. The practical deployment of these models was also considered. In real-world applications, especially within smart waste management systems, models must be not only accurate but also efficient in terms of memory usage, processing time, and energy consumption. Among the evaluated models, MobileNetV2 stood out as the most suitable for embedded systems and mobile applications due to its lightweight architecture and competitive performance. In terms of environmental policy and societal impact, automated waste classification can support broader sustainability initiatives by enabling municipalities to manage recyclable waste streams more effectively. It can also reduce reliance on manual labor, thus lowering costs and health risks associated with manual sorting processes. These systems may further empower communities by increasing awareness around recycling practices and contributing to cleaner urban environments. Limitations of the study include the scope of the datasets and the reliance on visual data alone. In real-life scenarios, other data types such as material composition or sensor readings could complement visual classification for improved accuracy.

Additionally, feature studies could explore the use of ensemble learning, semi-supervised training, or generative approaches such as GANs (Generative Adversarial Networks) to expand datasets and improve model robustness. Moreover, model explainability will become increasingly important as these systems are adopted more widely. For this reason, researchers must continue to investigate methods for making CNN decisions interpretable to users. Collaborative projects that bring together data scientists, environmental engineers, and policymakers will be essential in translating technical advancements into impactful, real-world waste management solutions. In conclusion, this study demonstrates the effectiveness of CNN-based deep learning models in classifying recyclable waste materials. By evaluating different architectures across datasets of varying size and quality, the research provides valuable insights into how model choice and data characteristics influence classification success. The findings support the use of deep learning for automating waste sorting processes and contribute to the broader goals of environmental sustainability and smart city development. Feature research should continue to refine these models, develop more diverse and realistic datasets, and pursue deployment strategies that balance accuracy with computational efficiency. The integration of such intelligent systems into waste management infrastructure could dramatically improve recycling rates, reduce operational costs, and support global efforts in resource conservation.

In addition to evaluating individual CNN architectures, this study proposed and implemented a novel hybrid model based on feature fusion, referred to as the Feature Fusion Hybrid Model. The hybrid approach aims to leverage the complementary strengths of multiple pre-trained CNNs—namely VGG16, InceptionV3, DenseNet121, and MobileNetV2—by extracting high-level features from each model and concatenating them into a unified feature representation. This combined feature vector was then passed through fully connected layers for final classification. The hybrid model was trained using five-fold cross-validation on both the Garbage Classification and Garbage-Dataset. Experimental results showed that the hybrid model achieved a test accuracy of 85.88% on the Garbage Classification dataset, surpassing all individual base models. On the more comprehensive Garbage-Dataset, the hybrid model reached an accuracy of 91.26%, exceeding the best-performing individual model (DenseNet121 with 90.75%). These results demonstrate the potential of feature fusion in improving classification performance by integrating diverse feature representations from multiple CNN architectures, leading to more robust and accurate waste classification systems.



1. GİRİŞ

Atık yönetimi, günümüzün çevresel sürdürülebilirlik hedefleri doğrultusunda giderek daha fazla önem kazanmaktadır. Nüfus artışı, kentleşme ve tüketim alışkanlıklarındaki değişim, üretilen atık miktarını artırmakta; bu durum, doğal kaynakların korunması ve çevresel etkilerin azaltılması açısından etkin geri dönüşüm sistemlerini zorunlu kılmaktadır. Geri dönüşüm süreçlerinin başarısı ise doğrudan, atıkların doğru ve hızlı bir şekilde sınıflandırılmasına bağlıdır. Geleneksel yöntemlerle yapılan atık sınıflandırması işlemleri çoğunlukla manuel olarak gerçekleştirilmekte ve bu süreçler; zaman alıcı, maliyetli ve insan hatasına açık olma özellikleri nedeniyle hem verimliliği düşürmekte hem de operasyonel yük oluşturmaktadır. Bu noktada, yapay zekâ ve özellikle derin öğrenme tabanlı görüntü işleme teknolojileri, atık sınıflandırma sistemlerine otomasyon ve yüksek doğruluk kazandırma potansiyeli ile öne çıkmaktadır.

Bu tez çalışması, geri dönüştürülebilir atıkların sınıflandırılmasında farklı derin öğrenme modellerinin etkinliğini karşılaştırmalı olarak analiz etmeyi amaçlamaktadır. Çalışmada, boyut ve sınıf dağılımları açısından farklılık gösteren iki veri kümesi (Garbage Classification ve Garbage-Dataset) kullanılmış; dört farklı CNN mimarisi (VGG16, InceptionV3, MobileNetV2 ve DenseNet121) transfer öğrenme yöntemi ile eğitilerek sınıflandırma başarımı değerlendirilmiştir. Model performansları, doğruluk, kesinlik, geri çağırma ve F1 skoru gibi metriklerle analiz edilmiş; veri kümesi büyüklüğü ve dengesi gibi faktörlerin bu başarı üzerindeki etkileri detaylandırılmıştır.

Buna ek olarak, çalışmanın katkısını artırmak amacıyla, mevcut CNN modellerinin güçlü yönlerini birleştiren bir hibrit model önerilmiştir. Bu modelde, her bir mimariden çıkarılan özellikler birleştirilerek hibrit bir özellik havuzu oluşturulmuş ve bu yapıdan elde edilen bilgilerle sınıflandırma işlemi gerçekleştirilmiştir. Böylece tek bir modelin sınırlamalarını aşmak, model çeşitliliğinden faydalanmak ve daha yüksek sınıflandırma performansı elde etmek hedeflenmiştir.

Bu kapsamda elde edilen bulgular, yalnızca akademik anlamda katkı sunmakla kalmayıp, aynı zamanda akıllı atık yönetimi sistemlerinin geliştirilmesine de pratik faydalar sağlamayı hedeflemektedir.

1.1. Tezin Kapsamı

Bu tezde, geri dönüştürülebilir atıkların sınıflandırılmasında derin öğrenme modellerinin başarımı, farklı boyut ve dağılıma sahip iki veri kümesi üzerinde değerlendirilmiştir. DenseNet121, InceptionV3, MobileNetV2 ve VGG16 gibi öncül CNN mimarileri kullanılarak, veri kümesi boyutunun ve özelliklerinin model performansı üzerindeki etkileri analiz edilmiştir. Ayrıca, bu tez çalışmasında, farklı mimarilerden elde edilen özelliklerin birleştirilmesiyle oluşturulan hibrit bir model geliştirilmiş ve bu yaklaşımın sınıflandırma performansına katkısı incelenmiştir. Hibrit model, çoklu CNN mimarilerinin güçlü yönlerini bir araya getirerek, sınıflar arasındaki dengesizlik ve veri çeşitliliğine karşı daha dayanıklı bir yapı sunmayı hedeflemiştir.

1.2. Tezin Amacı

Bu tezin amacı, yapay zekâ destekli atık sınıflandırma sistemlerinin başarısını artırmak için derin öğrenme modellerinin farklı veri kümesi koşulları altındaki performansını derinlemesine analiz etmektir. Çalışmada, veri kümesi boyutu ve çeşitliliğinin sınıflandırma doğruluğu üzerindeki etkileri ortaya konulurken, farklı CNN mimarilerinden elde edilen özelliklerin birleştirilmesiyle oluşturulan hibrit modelin performansa katkısı da incelenmiştir. Böylece, hem veri seti özellikleri hem de mimari birleşimlerinin sınıflandırma başarısına etkilerini değerlendirerek, gelecekteki sürdürülebilir atık yönetimi uygulamalarına ışık tutmayı hedeflemektedir.

1.3. Problemin Tanımı

Geri dönüşüm süreçlerinin başarısı, atıkların doğru ve hızlı sınıflandırılmasına bağlıdır. Ancak insan gücüne dayalı yöntemler hem yavaş hem de hataya açıktır; fiziksel olarak benzer materyaller tutarsızlıklara yol açmakta ve artan atık miktarı mevcut sistemleri yetersiz bırakmaktadır. Derin öğrenme tabanlı otomatik sistemler, yüksek doğruluk potansiyeli sunsa da başarıları büyük ölçüde veri kümesinin büyüklüğü, çeşitliliği ve dengesi ile model mimarisine bağlıdır.

Bu çalışmada, farklı boyut ve dağılıma sahip veri setleri üzerinde dört yaygın CNN mimarisi (DenseNet121, InceptionV3, MobileNetV2, VGG16) sistematik olarak karşılaştırılmış ve ayrıca bu mimarilerin özelliklerini birleştiren hibrit bir modelin sınıflandırma performansına katkısı incelenmiştir.

1.4. Çalışmanın Kapsamı

Bu çalışmada, geri dönüştürülebilir atıkların sınıflandırılmasında derin öğrenme tabanlı yaklaşımların başarımı değerlendirilmiştir. Farklı boyut ve sınıf dağılımına sahip iki veri kümesi üzerinde VGG16, MobileNetV2, InceptionV3 ve DenseNet121 olmak üzere dört yaygın CNN mimarisiyle deneyler yapılmış; transfer öğrenme yöntemiyle bu modeller yeniden eğitilerek performansları karşılaştırılmıştır. Eğitimde 5 katlı çapraz doğrulama kullanılarak modellerin genelleme yetenekleri test edilmiştir.

Ayrıca, bu mimarilerden elde edilen özelliklerin birleştirilmesiyle oluşturulan hibrit modelin, sınıflandırma performansına katkısı incelenmiştir. Modeller, doğruluk, kayıp, precision, recall, F1-skor ve ROC-AUC gibi çeşitli metriklerle değerlendirilmiş; elde edilen sonuçlara dayanarak veri kümesi boyutu, çeşitliliği ve mimari tercihlerin sınıflandırma doğruluğu üzerindeki etkileri analiz edilmiştir.

1.5. Çalışmanın Yapısı

Birinci bölümde, çalışmanın amacı, önemi, problem tanımı, kapsamı ve tez bölümlerinin genel yapısı açıklanmıştır. İkinci bölümde, literatürde atık sınıflandırması üzerine yapılan çalışmalar incelenmiş, görüntü işleme yöntemleri ve derin öğrenme tabanlı görüntü sınıflandırma çalışmaları tartışılmıştır. Ayrıca, çalışmada kullanılan derin öğrenme modelleri ve e hibrit modelin literatürdeki yeri değerlendirilmiştir. Üçüncü bölümde, çalışmada kullanılan veri setleri detaylı olarak tanıtılmış, ön işleme aşamaları ve veri artırımı (artırma) yöntemleri açıklanmış, kullanılan derin öğrenme modelleri ve transfer öğrenme (transfer learning) yaklaşımı üzerinde durulmuş ve model eğitim stratejisi hakkında bilgi verilmiştir. Dördüncü bölümde, çalışmada kullanılan programlama dili, geliştirme ortamı, kütüphaneler, çerçeveler ve donanım özellikleri belirtilmiştir. Beşinci bölümde, model eğitimi süreci, 5 katlı çapraz doğrulama yöntemi ve kullanılan değerlendirme metrikleri (doğruluk, kayıp, kesinlik, geri çağırma, F1 skoru ve ROC AUC açıklanmıştır. Altıncı bölümde, her bir model ve veri kümesi için elde edilen eğitim ve doğrulama başarı

oranları grafiklerle sunulmuş, ayrıca modellerin test doğrulukları ve değerlendirme metrikleri karşılaştırılmıştır. En başarılı modeller için karmaşıklık matrisleri de verilmiştir. Yedinci bölümde, elde edilen sonuçlar yorumlanmış, literatürdeki diğer çalışmalarla karşılaştırmalar yapılmış ve çalışmanın sınırlılıkları tartışılmıştır. Sekizinci bölümde ise çalışmanın genel sonuçları sunulmuş ve gelecekte yapılabilecek çalışmalar için önerilerde bulunulmuştur.

1.6. Literatür

Çevresel sürdürülebilirliğin sağlanmasında atıkların doğru yönetimi ve sınıflandırılması kritik bir rol oynamaktadır. Bu kapsamda, son yıllarda derin öğrenme temelli yöntemlerin atık sınıflandırma sistemlerinde kullanılmasına yönelik çok sayıda çalışma yapılmıştır. Geri dönüşüm süreçlerinin etkinliğinin artırılmasında, atıkların doğru şekilde sınıflandırılması temel bir adım olarak öne çıkmaktadır.

Derin öğrenme temelli görüntü işleme teknikleri, atık sınıflandırma sistemlerinde insan müdahalesini en aza indirme ve doğruluk oranlarını artırma potansiyeli nedeniyle önemli bir araştırma alanı haline gelmiştir. Bu alanda yapılan çalışmalar, farklı CNN mimarilerinin performanslarının çeşitli veri setleri üzerinde sistematik olarak incelendiği kapsamlı deneyler sunmaktadır. Özellikle küçük veri setlerinde sınıflandırma başarımını artırmak amacıyla transfer öğrenme tekniklerinin yaygın olarak tercih edildiği görülmektedir [9, 10, 11].

Kaya ve arkadaşları [1], TrashNet veri kümesi üzerinde farklı CNN mimarilerini optimize ederek DenseNet169 mimarisi ile %96,42 doğruluk oranı elde etmiş ve transfer öğrenmenin sınırlı veri koşullarında sağladığı katkıyı vurgulamıştır. Sevinç ve Özyurt [2], ResNet-50 ve özel tasarladıkları bir CNN mimarisiyle %91,84 doğruluk oranına ulaşarak, derin ağların atık sınıflandırmadaki etkinliğini göstermiştir. Benzer şekilde, Bircanoğlu ve arkadaşları [3], DenseNet121 mimarisiyle %95 başarımleri elde etmiştir. Wang ve arkadaşları [4], MobileNetV3 mimarisinin verimli yapısını vurgularken; Rabano ve arkadaşları [6], MobileNet tabanlı hızlı sistemleri incelemiştir. Bu çalışmalar, mobil sistemler için optimize edilmiş hafif mimarilerin potansiyelini ortaya koymaktadır.

Atık sınıflandırmasında ilk derin öğrenme temelli yaklaşımlardan biri Yang ve Thung [5] tarafından Stanford CS229 projesinde sunulmuştur. Daha sonra Zhang ve arkadaşları [7], transfer öğrenme yaklaşımıyla DenseNet kullanarak akıllı atık

sınıflandırması gerçekleştirmiş ve dengesiz veri setlerinin etkilerini vurgulamıştır. LeCun, Bengio ve Hinton [8] tarafından sunulan temel derin öğrenme ilkeleri bu alandaki temel yaklaşımları şekillendirmiştir. Transfer öğrenmenin teorik altyapısı ise Pan ve Yang [9] tarafından detaylandırılmıştır.

Model mimarisi olarak VGG16 [10], DenseNet121 [11], InceptionV3 [12] ve MobileNetV2 [13] gibi öncül CNN yapıları, literatürde yaygın olarak kullanılmakta olup, ResNet mimarisiyle derin ağlarda çökme probleminin önüne geçilmesini sağlayan residual bağlantılar He ve arkadaşları [14] tarafından sunulmuştur.

Veri kümesi çeşitliliği de model performansını etkilemektedir. Bu çalışmada kullanılan küçük ölçekli Garbage Classification veri kümesi [16] ve geniş kapsamda 10 sınıftan oluşan Garbage-Dataset [15] literatürde öne çıkan veri setlerindendir.

Hu ve arkadaşları [17], iyileştirilmiş CNN mimarisiyle katı atık sınıflandırmasında başarıyı artırırken, Ali ve arkadaşları [18] yapay zekâ destekli CNN sistemleriyle sınıflandırma doğruluğunu yükseltmiştir. Chen ve arkadaşları [19], inorganik paketleme sınıflandırması için özel bir CNN yapısı önermiştir. Singh ve arkadaşları [20], atık sınıflandırmasında derin öğrenmenin otomatik sistemlerdeki başarısını kanıtlamıştır.

Zhang ve arkadaşları [21], federated learning ile farklı konumlardaki verilerin gizliliğini koruyarak sınıflandırma gerçekleştirmiştir. Rahman ve arkadaşları [22] geliştirilmiş çok katmanlı CNN yapısı ile başarıyı artırmış, Kumar ve arkadaşları [23] ise farklı derin öğrenme modellerini karşılaştırmalı olarak analiz etmiştir. Li ve arkadaşları [24], EfficientNetV2 tabanlı güncellenmiş bir çerçeve sunmuş ve yüksek doğruluk bildirmiştir. Patel ve arkadaşları [25], WaDaBa veri kümesi ile plastik atık sınıflandırması gerçekleştirmiştir. Zhao ve arkadaşları [26], piksel dağılımı tabanlı öğrenme ile performansı iyileştirmiştir.

Son yıllarda hibrit model ve özellik birleştirme (feature fusion) tabanlı yaklaşımlar da öne çıkmaktadır. Lilhore ve arkadaşları [27], Swin Transformer ve ConvNeXt mimarilerini birleştirerek TrashNet veri kümesinde %96.78 doğruluk oranına ulaşmıştır. Mustapha ve arkadaşları [28], YOLOv8 ve CNN tabanlı hibrit sistemleri ile gerçek zamanlı sınıflandırma sunmuştur. Li ve arkadaşları [29], ResNet50 ile özellik birleştirme modülü kullanarak başarıyı artırmıştır. Sayem ve arkadaşları [30], DenseViT ve GELAN-E tabanlı çift akışlı bir hibrit model önermiştir. Ahmed ve

arkadařları [31], DenseNet201 ve MobileNetV2 mimarilerini ieren RWC-Net modeli ile %95.01 doęruluk elde etmiřtir. Son olarak, Qiu ve arkadaşları [32], EfficientNetV2'yi dikkat modlleri ile geliřtirerek bařarılı bir sınıflandırma performansı sunmuřtur.

Bu tez alıřması, mevcut literatrdeki bu eřitli yaklařımların irdelenmesiyle, hem veri kmesi zelliklerinin hem de mimari tercihlerin atık sınıflandırma performansına etkilerini sistematik řekilde deęerlendirmektedir.



2. MATERYAL VE YÖNTEM

Bu bölümde, çalışmada kullanılan veri setleri, veri ön işleme süreçleri, derin öğrenme modelleri, transfer öğrenme yaklaşımları ve eğitim stratejileri detaylı bir şekilde açıklanmaktadır. Amaç, gerçekleştirilen deneylerin altyapısını ve kullanılan yöntemlerin gerekçelerini açık ve sistematik biçimde ortaya koymaktır.

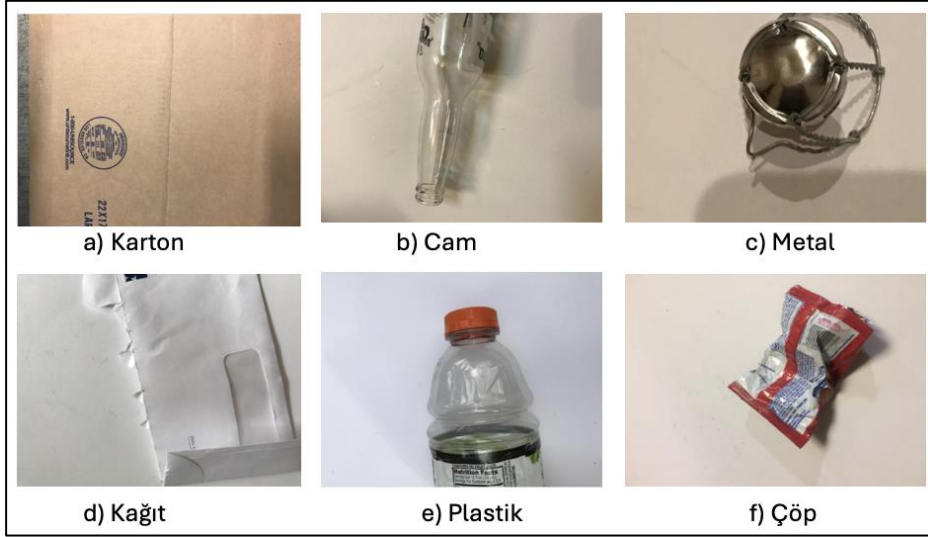
2.1. Kullanılan Veri Kümeleri

Atık sınıflandırması üzerine yapılan bu çalışmada, farklı boyut ve çeşitlilikte iki ayrı veri kümesi kullanılmıştır. Bu veri setleri, sınıflandırma modellerinin farklı veri büyüklükleri ve sınıf dağılımı özellikleri altındaki performanslarının karşılaştırılmasına olanak sağlamıştır.

2.1.1. Garbage Classification veri kümesi

Garbage Classification veri kümesi, toplamda 2,890 görüntü içeren ve altı farklı atık sınıfından oluşan bir veri kümesidir: karton (403), cam (501), metal (410), kağıt (594), plastik (482) ve çöp (137).

Veri kümesinde sınıflar arasında belirgin bir dengesizlik gözlemlenmektedir. Özellikle çöp kategorisindeki görüntü sayısının diğer kategorilere göre daha az olması, eğitim sürecinde sınıf dengesizliği sorununa neden olabilecek bir durum oluşturmuştur. Bu nedenle, çalışmada bu veri kümesi üzerinde veri artırma (augmentation) yöntemleri uygulanarak sınıf dengesi sağlanmaya çalışılmıştır. Şekil 2.1 Garbage Classification veri kümesinden örnek görüntüler içermektedir.

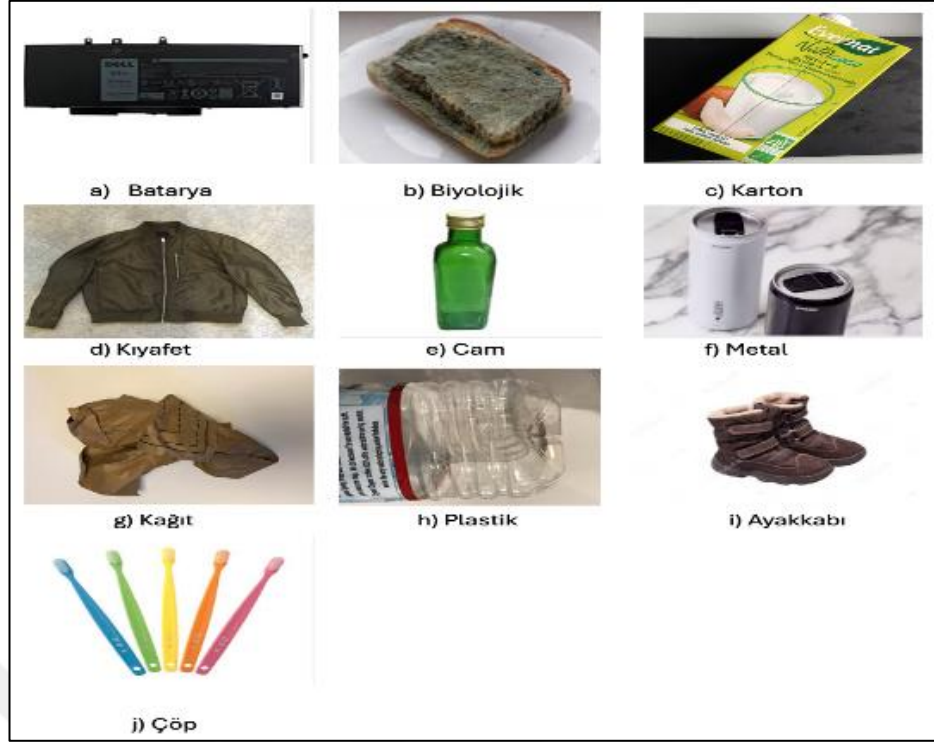


Şekil 2.1. Garbage Classification veri kümesinden örnek görüntüler.

2.1.2. Garbage Dataset veri kümesi

Garbage-Dataset veri kümesi, daha geniş ölçekli bir yapıya sahiptir ve toplam 9,660 görüntü içermektedir: Toplam 9660 görüntü içeren bu veri seti, on farklı atık türünü dengeli bir şekilde temsil etmektedir: batarya (944), biyolojik (997), karton (777), kıyafet (1001), cam (1055), metal (1020), kâğıt (989), plastik (962), ayakkabı (968) ve çöp (947).

Bu veri kümesinde görüntüler on farklı sınıfa ayrılmıştır: karton, cam, metal, kağıt, plastik, çöp, pil, tekstil, organik atık ve elektrikli atık. Garbage-Dataset, sınıf çeşitliliği ve görüntü sayısının fazla olması nedeniyle derin öğrenme modellerinin genelleme yeteneklerini daha güvenilir bir şekilde değerlendirme imkânı sunmaktadır. Veri kümesindeki sınıf dağılımı, Garbage Classification veri kümesine göre daha dengelidir. Şekil 2.2 Garbage Dataset’inden setinden görüntüler içermektedir.



Şekil 2.2. Garbage Dataset veri kümesinden örnek görüntüler.

2.2. Ön İşleme Aşamaları

Görüntü sınıflandırma problemlerinde model başarımı, yalnızca seçilen derin öğrenme mimarilerine değil, aynı zamanda verilerin eğitim öncesinde nasıl işlendiğine de doğrudan bağlıdır. Bu çalışmada, veri kalitesini artırmak, modelin öğrenme sürecini hızlandırmak ve sınıflandırma başarımını yükseltmek amacıyla çeşitli ön işleme teknikleri uygulanmıştır. Ön işleme süreci temel olarak iki aşamada gerçekleştirilmiştir: görüntü boyutlandırma ve veri artırımı.

2.2.1. Görüntü boyutlandırma

Kullanılan derin öğrenme modellerinin giriş katmanları sabit boyutlarda görüntü kabul etmektedir. Bu nedenle, veri setlerindeki tüm görüntüler 224x224 piksel boyutlarına yeniden ölçeklendirilmiştir. Görüntü boyutlandırma işlemi sırasında, görüntülerin oranları korunmuş ve deformasyonların önüne geçilmiştir. Bu işlem, hem eğitim sürecinin verimli ilerlemesini sağlamakta hem de farklı boyutlardaki görüntülerden kaynaklanabilecek olası model hatalarını azaltmaktadır.

2.2.2. Veri artırımı yöntemleri

Özellikle Garbage Classification veri kümesinde sınıflar arasında gözlenen dengesizliği azaltmak ve modelin genelleme yeteneğini artırmak amacıyla veri artırımı

teknikleri uygulanmıştır. Bu kapsamda, özellikle görüntü sayısı en düşük olan “çöp” sınıfına odaklanılmıştır. Başlangıçta bu sınıfa ait yalnızca 137 görüntü bulunurken, çeşitli artırma yöntemleri ile bu sayı 500’e yükseltilmiştir. Uygulanan veri artırma teknikleri arasında rastgele döndürme (rotation), yatay kaydırma (horizontal shift), dikey kaydırma (vertical shift), yakınlaştırma (zooming) ve yatay çevirme (horizontal flip) gibi yöntemler yer almaktadır.

Bu teknikler, mevcut verilerden yapay fakat anlamlı varyasyonlar üreterek eğitim kümesinin çeşitliliğini artırmış ve modelin farklı açılardan veya pozisyonlardan gelen verileri tanıma yeteneğini geliştirmiştir. Özellikle çöp sınıfında yapılan hedefli veri artırımı sayesinde sınıf dengesi sağlanmış ve modelin tüm sınıfları eşit düzeyde öğrenebilmesi desteklenmiştir. Böylece, sınıf dengesizliğinden kaynaklanabilecek öğrenme sorunlarının önüne geçilmiştir.

2.3. Transfer Öğrenme Yaklaşımı

Derin öğrenme tabanlı görüntü sınıflandırma problemlerinde, yeterli büyüklükte ve çeşitlilikte veri setlerine sahip olmak model başarımı açısından büyük önem taşımaktadır. Ancak, sıfırdan bir modeli büyük bir veri kümesi üzerinde eğitmek hem yüksek donanım kaynakları hem de uzun eğitim süreleri gerektirmektedir. Bu durum, özellikle sınırlı veri setleri ile çalışılan projelerde önemli bir engel oluşturmaktadır. Bu sorunun üstesinden gelmek amacıyla, bu çalışmada transfer öğrenme (transfer learning) yaklaşımı kullanılmıştır. Transfer öğrenme, önceden büyük veri setleri üzerinde eğitilmiş bir modelin, başka bir görev için yeniden kullanılması esasına dayanmaktadır. Bu yöntemde, modelin daha önce öğrendiği genel özellikler korunur ve yalnızca son katmanlar, yeni görev için yeniden eğitilir. Böylece eğitim süreci kısılırken, küçük veri setlerinde dahi yüksek doğruluk oranları elde edilebilmektedir. Bu tez çalışmasında kullanılan VGG16, MobileNetV2, InceptionV3 ve DenseNet121 modelleri, daha önce ImageNet veri kümesi üzerinde önceden eğitilmiş (pre-trained) ağırlıklarla yüklenmiş ve yeni veri setlerine adapte edilmiştir. Modellerin temel evrişim (convolution) katmanları sabit tutulmuş (dondurulmuş) ve yalnızca sınıflandırma katmanları güncellenerek yeniden eğitilmiştir. Böylece modeller, temel özellik çıkarımı yeteneklerini korurken, yeni veri setlerindeki sınıfları doğru bir şekilde tanımlamayı öğrenmiştir. Transfer öğrenme yaklaşımı, özellikle Garbage Classification veri kümesi gibi küçük ölçekli ve sınıf dengesizliği içeren veri setlerinde

model başarımının artırılmasına önemli katkılar sağlamıştır. Garbage-Dataset gibi daha büyük veri setlerinde ise transfer öğrenme, eğitim süresini önemli ölçüde azaltmış ve modellerin daha hızlı şekilde yüksek doğruluk oranlarına ulaşmasına olanak tanımıştır.

2.4. Kullanılan Derin Öğrenme Modelleri

Bu çalışmada, farklı veri kümesi özelliklerine göre çeşitli performans gösterebilecek dört farklı derin öğrenme modeli tercih edilmiştir. Modellerin seçiminde, literatürde kanıtlanmış başarımları, parametrik verimlilikleri ve farklı veri koşullarına uyum yetenekleri dikkate alınmıştır. Kullanılan modeller VGG16, MobileNetV2, InceptionV3 ve DenseNet121'dir. Her bir modelin seçim gerekçesi ve temel mimari özellikleri aşağıda özetlenmektedir.

2.4.1. VGG16 modeli

VGG16 modeli [10], düzenli ve katı bir mimari yapıya sahip olması nedeniyle tercih edilmiştir. Model, ardışık 3x3 evrişim (convolution) katmanları ve ardından gelen havuzlama (pooling) katmanlarından oluşmaktadır. Bu yapı, modelin derinleştikçe daha soyut ve anlamlı özellikler çıkarabilmesini sağlamaktadır. VGG16'nın parametre sayısının fazla olması eğitim süresini artırmakla birlikte, orta ölçekli veri setlerinde yüksek doğruluk oranları elde etmesine olanak tanımaktadır. Bu çalışmada, özellikle küçük veri setlerinde özellik çıkarımı gücünü değerlendirmek amacıyla kullanılmıştır.

2.4.2. MobileNetV2 modeli

MobileNetV2 modeli [13], sınırlı donanım kaynaklarında bile verimli çalışabilen hafif bir ağ mimarisi olarak tercih edilmiştir. Model, ayrı derinlikli evrişim (depthwise separable convolution) ve tersine rezidüel bloklar (inverted residual blocks) kullanarak hesaplama maliyetini önemli ölçüde azaltmaktadır. Bu çalışmada, özellikle Garbage Classification veri kümesi gibi daha küçük ve dengesiz veri setlerinde, hızlı eğitim süresi ve düşük model karmaşıklığı sağlamak amacıyla MobileNetV2 kullanılmıştır.

2.4.3. InceptionV3 modeli

InceptionV3 modeli [12], farklı ölçeklerde özellik çıkarımı yapabilen yapısı sayesinde geniş veri setlerinde etkili sonuçlar üretebilmesi nedeniyle tercih edilmiştir. Model, faktörize edilmiş evrişim (factorized convolution) ve ağ içinde ağ (network-in-network) teknikleri kullanarak hem doğruluğu artırmakta hem de parametre sayısını

dengede tutmaktadır. Bu çalışmada, özellikle Garbage-Dataset gibi daha büyük ve dengeli veri setlerinde InceptionV3'ün genelleme başarısı değerlendirilmiştir.

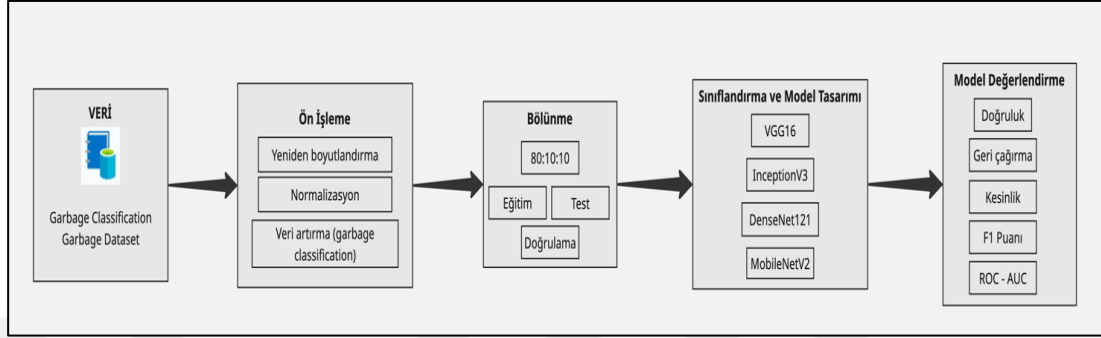
2.4.4. DenseNet121 modeli

DenseNet121 modeli [11], her katmanın bir önceki tüm katmanlardan aldığı bağlantılar sayesinde bilgi akışını güçlendirmektedir. Bu yapı, özellikle derin ağlarda karşılaşılan gradyan kaybolması problemini azaltmakta ve öğrenme sürecini hızlandırmaktadır. DenseNet121, küçük veri setlerinde dahi yüksek doğruluklar elde edebilmesi nedeniyle çalışmada hem Garbage Classification hem de Garbage-Dataset üzerinde performans analizi yapmak için tercih edilmiştir.

2.5. Model Eğitim Stratejisi

Bu çalışmada, derin öğrenme modellerinin eğitim süreçlerinde hem model başarımını artırmak hem de aşırı uyum problemini önlemek amacıyla belirli stratejiler benimsenmiştir. Eğitim süreci, veri setlerinin doğru bir şekilde bölünmesi, çapraz doğrulama yöntemi kullanımı ve eğitim parametrelerinin dikkatli bir şekilde ayarlanması esaslarına dayanmaktadır. Veri setleri, eğitim (%80), doğrulama (%10) ve test (%10) olmak üzere üç ayrı gruba bölünmüştür. Eğitim verileri modelin öğrenme sürecinde, doğrulama verileri modelin hiper parametre ayarlarının yapılmasında, test verileri ise nihai performansın objektif bir şekilde ölçülmesinde kullanılmıştır. Model değerlendirmesinde daha güvenilir sonuçlar elde etmek amacıyla 5 katlı çapraz doğrulama yöntemi uygulanmıştır. Bu yöntemde, eğitim ve doğrulama verileri beş eşit parçaya bölünmüş; her bir parça bir kez doğrulama verisi, kalan parçalar ise eğitim verisi olarak kullanılmıştır. Böylece her model, farklı veri alt kümeleri üzerinde eğitilmiş ve test edilmiştir. Bu yaklaşım, modelin genelleme kabiliyetini artırarak farklı veri dağılımlarındaki performansını objektif şekilde ortaya koymuştur. Eğitim sürecinde erken durdurma mekanizması da uygulanmıştır. Erken durdurma stratejisi, doğrulama kaybı belirli bir sayıda ardışık epok boyunca iyileşmediğinde eğitim sürecini sonlandırarak aşırı uyumu önlemektedir. Bu çalışmada erken durdurma için sabitlenen sabır (patience) değeri 5 epok olarak belirlenmiştir. Tüm modeller, mini-batch grupları kullanılarak eğitilmiştir. Batch boyutu 32 olarak seçilmiştir. Maksimum eğitim epok sayısı 20 olarak belirlenmiş olmakla birlikte, erken durdurma mekanizması nedeniyle bazı modeller bu sayıya ulaşmadan eğitimlerini tamamlamıştır. Öğrenme hızı değeri 0,0001 olarak belirlenmiş ve Adam optimizasyon

algoritması kullanılmıştır. Model kaybı, çok sınıflı çapraz entropi kaybı (sparse categorical cross-entropy) fonksiyonu kullanılarak hesaplanmıştır. Tüm eğitim süreçleri, aynı donanım altyapısında ve aynı parametre ayarlarıyla gerçekleştirilerek modeller arasında adil bir karşılaştırma ortamı sağlanmıştır. Bu tez çalışmasında uygulanan genel süreç akışını açıklayan diyagram Şekil 2.3'de sunulmuştur.



Şekil 2.3. Model akış diyagramı.

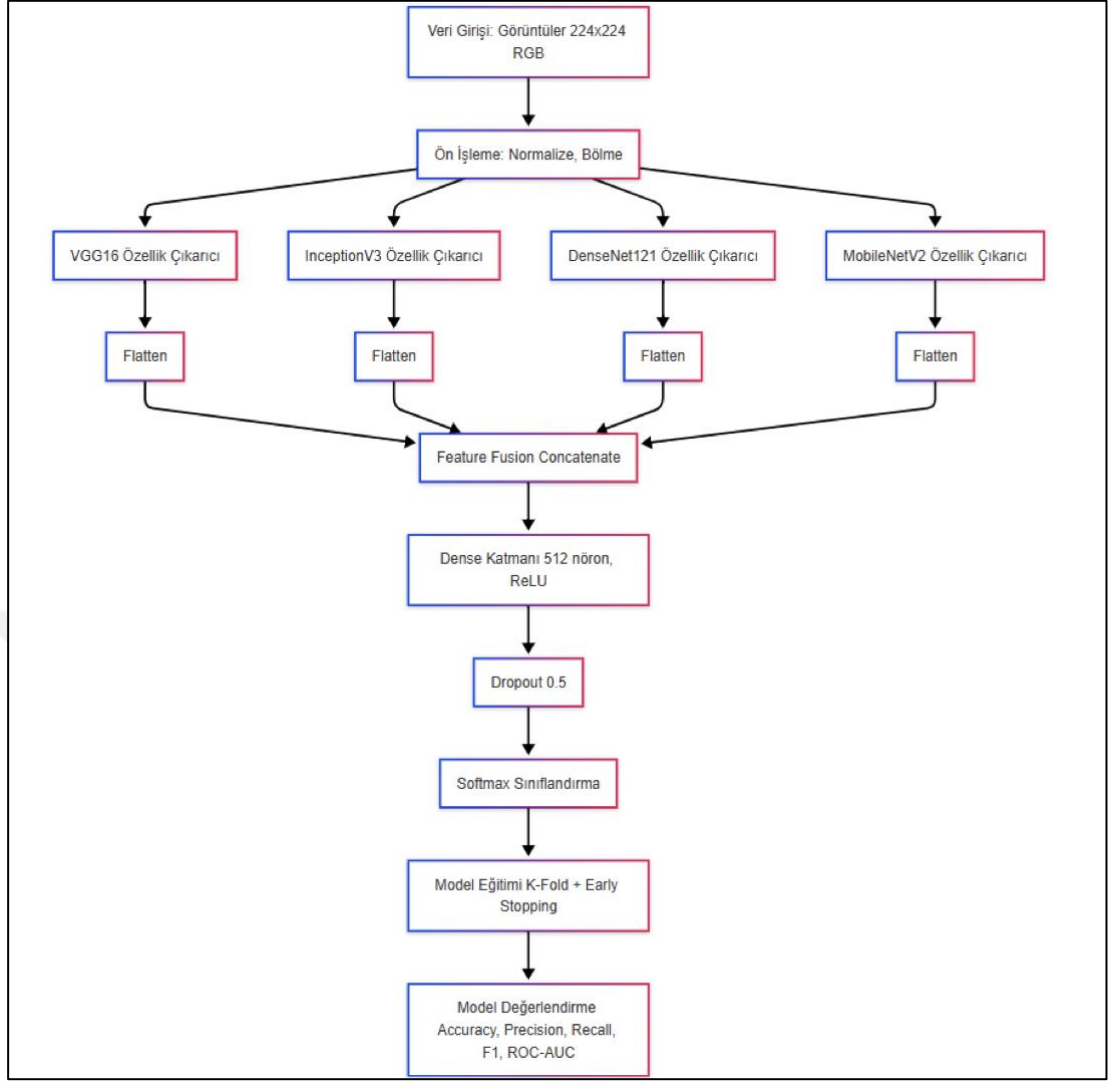
2.6. Özellik Füzyonlu Hibrit Model

Bu çalışmada, geri dönüştürülebilir atıkların görüntü tabanlı sınıflandırılmasına yönelik olarak, çoklu mimari temelli bir özellik füzyonlu hibrit derin öğrenme modeli geliştirilmiştir. İlk olarak, altı kategoriden oluşan atık veri kümesi etiketleri sayısal olarak kodlanmış, görüntüler RGB formatında 224×224 boyutuna yeniden ölçeklendirilmiş ve piksel değerleri 0–1 aralığına normalize edilerek model girişine uygun hale getirilmiştir. Eğitim sürecinde, etiketler doğrudan sparse_categorical_crossentropy kayıp fonksiyonu ile kullanılmış ve bu sayede to_categorical dönüşümüne ihtiyaç duyulmamıştır. Model mimarisinde VGG16, InceptionV3, DenseNet121 ve MobileNetV2 olmak üzere dört farklı önceden eğitilmiş CNN mimarisi kullanılmış; bu modellerin include_top=False parametresiyle son sınıflandırma katmanları kaldırılmış ve tüm katmanlar trainable=False olarak dondurulmuştur. Her bir modelin çıkış özelliği Flatten() katmanlarıyla vektörleştirilmiş ve concatenate() fonksiyonu ile birleştirilmiştir. Elde edilen birleşik özellik temsili, 512 nöronlu Dense bir katmana ve ardından %50 oranında Dropout katmanına aktarılmış, son olarak softmax aktivasyonlu çıkış katmanında sınıflandırma gerçekleştirilmiştir.

Model, 5 katlı K-Fold çapraz doğrulama ile eğitilmiştir. Her fold'da eğitim, doğrulama ve test setleri farklı şekilde ayrılmış; model her seferinde sıfırdan oluşturulmuş ve her bir CNN'in çıktısı eğitim aşamasında tüm ağlara paralel olarak verilmiştir. Model

eđitimi sırasında EarlyStopping callback fonksiyonu kullanılarak aşırı öğrenme (overfitting) riski azaltılmıştır. Performans değerdendirmesi için doğruluk (dođruluk), kayıp (loss), hassasiyet (precision), geri çağırma (recall), F1 skoru ve ROC-AUC gibi çoklu metrikler hesaplanmış; her fold sonrasında tahmin çıktıları üzerinden bu değerdler ayrı ayrı raporlanmışdır. Son olarak, modelin genel başarımını değerdlendirmek amacıyla fold sonuçlarının ortalaması alınarak modelin istikrarı ve genelleme kapasitesi ölçülmüşdür. Bu hibrit yapının en temel katkısı, farklı CNN mimarilerinin çıkarabildiđi çeşitli seviye özelliklerin birleştirilerek sınıflandırma dođruluđunun artırılmasıdır. Bu yaklaşım, literatürde tek model temelli çözümlerin sınırlamalarını aşmakta ve görsel karmaşıklığı yüksek atık türlerinin ayırımında daha güvenilir sonuçlar sunmaktadır.





Şekil 2.4. Özellik füzyon hibrit modeli akış diyagramı.



3. KULLANILAN TEKNOLOJİLER VE YAZILIMLAR

Bu çalışmada, derin öğrenme modellerinin geliştirilmesi, eğitilmesi ve değerlendirilmesi süreçlerinde çeşitli yazılım kütüphaneleri ve teknolojik araçlar kullanılmıştır. Kullanılan teknolojiler, model eğitimi için gerekli olan yüksek hesaplama gücünü sağlamak ve deneysel süreçlerin verimli bir şekilde yürütülmesini mümkün kılmak amacıyla seçilmiştir.

3.1. Programlama Dili ve Geliştirme Ortamı

Tüm çalışmalar Python programlama dili kullanılarak gerçekleştirilmiştir. Python, geniş kütüphane desteği, açık kaynaklı yapısı ve esnek programlama olanakları nedeniyle özellikle makine öğrenmesi ve derin öğrenme alanlarında yaygın bir şekilde tercih edilmektedir. Model geliştirme ve eğitim süreçleri, Jupyter Notebook ortamında yürütülmüştür. Jupyter Notebook, kod yazma, sonuçları anlık olarak gözlemleme ve görsel analizler yapma imkânı sunarak deneysel çalışmalarda esneklik sağlamıştır.

3.2. Kullanılan Kütüphaneler ve Frameworkler

Bu çalışmada Python programlama dili kullanılarak, görüntü işleme ve derin öğrenme tabanlı atık sınıflandırma uygulamaları gerçekleştirilmiştir. Model oluşturma, eğitim süreci, veri hazırlama ve sonuç analizi aşamalarında çeşitli açık kaynaklı kütüphane ve çerçevelerden yararlanılmıştır. Çalışma kapsamında kullanılan temel kütüphanelerden biri olan os, dosya ve dizin işlemlerinin yönetilmesinde kullanılmıştır. Veri kümesindeki klasör yapısına erişim ve görüntü dosyalarının yüklenmesi işlemleri bu kütüphane aracılığıyla gerçekleştirilmiştir. Görüntü dosyalarının okunması ve yeniden boyutlandırılması işlemleri için OpenCV (cv2) kütüphanesinden yararlanılmıştır. Görüntü verilerinin dizi formatında işlenmesi, normalize edilmesi ve ortalama değer gibi istatistiksel hesaplamaların yapılması amacıyla NumPy kütüphanesi kullanılmıştır. Eğitim verilerinin karıştırılması işlemi için ise random kütüphanesinden faydalanılmıştır. Model sonuçlarının görselleştirilmesinde Seaborn ve Matplotlib kütüphaneleri kullanılmıştır. Seaborn, özellikle karışıklık matrisi çizimlerinde tercih edilirken, Matplotlib eğitim ve test

aşamalarındaki doğruluk oranları, kayıplar, doğruluk, geri çağırma, F1 skoru gibi metriklerin grafiksel sunumu için kullanılmıştır. Eğitim ve test veri setlerinin ayrılması için scikit-learn kütüphanesinin model_selection modülü, 5 katlı çapraz doğrulama sürecinin yürütülmesi için ise katlama sınıfı kullanılmıştır. Model performansını değerlendirmek amacıyla yine scikit-learn kütüphanesinin metrics modülü kullanılarak doğruluk, kesinlik, geri çağırma, F1 skoru ve ROC-AUC metrikleri hesaplanmıştır. Derin öğrenme modeli oluşturmak için TensorFlow kütüphanesinin Keras API'sinden yararlanılmıştır. TensorFlow Keras içerisinde yer alan applications modülü kullanarak önceden eğitilmiş DenseNet121 modeli transfer öğrenme amacıyla kullanılmıştır. Model mimarisi model sınıfı ile tanımlanmış, Flatten, Dense ve Dropout katmanları layers modülü üzerinden eklenmiştir. Modelin optimizasyonu Adam algoritması ile gerçekleştirilmiş ve eğitim sürecinde aşırı uyumu (overfitting) engellemek amacıyla EarlyStopping yöntemi uygulanmıştır. Ayrıca, küçük veri kümesinde sınıf dengesizliğini azaltmak amacıyla tensorflow.keras.preprocessing.image modülü üzerinden ImageDataGenerator sınıfı kullanarak veri artırımı (data augmentation) işlemleri yapılmıştır. Tüm bu kütüphaneler ve çerçeveler, çalışmanın veri ön işleme, model oluşturma, eğitim, değerlendirme ve görselleştirme aşamalarında etkin bir şekilde kullanılarak bütünlük bir yaklaşım oluşturulmuştur.

3.3. Donanım ve Çalışma Ortamı

Bu çalışma, kişisel bir bilgisayar üzerinde gerçekleştirilmiştir. Kullanılan bilgisayarın temel donanım özellikleri şu şekildedir: Intel Core i7 işlemci, 16 GB RAM ve 512 GB SSD depolama alanı. Grafik işlem birimi (GPU) olarak entegre grafik kartı kullanılmış, harici bir GPU donanımı kullanılmamıştır. Model eğitim süreçleri, Anaconda dağıtımı üzerinde çalışan Jupyter Notebook ortamında, Python dili kullanılarak yürütülmüştür. Eğitim ve değerlendirme adımlarında TensorFlow, Keras, OpenCV, NumPy, scikit-learn, matplotlib ve seaborn gibi kütüphanelerden yararlanılmıştır. Geliştirme işlemleri Python 3,8 sürümünde ve macOS işletim sistemi altında gerçekleştirilmiştir. Model geliştirme, eğitim ve sonuçların görselleştirilmesi süreçlerinde yalnızca yerel kaynaklar kullanılmış, herhangi bir bulut tabanlı platformdan (örneğin Google Colab, AWS veya Azure gibi) yararlanılmamıştır. Bu doğrultuda, veri işleme, model eğitimi ve görselleştirme işlemlerinin tamamı belirtilen donanım ve yazılım altyapısı kullanılarak yerel ortamda gerçekleştirilmiştir.

Kullanılan donanımın sınırlı işlem gücü nedeniyle model eğitim süreleri uzun sürmüştür. Ancak, transfer öğrenme (transfer learning) yaklaşımı benimsenerek eğitim süreçlerinin süresi ve kaynak tüketimi optimize edilmiştir. Önceden eğitilmiş modellerin ağırlıkları kullanılarak sıfırdan model eğitimi yapılmasına gerek kalmadan, daha hızlı ve verimli bir eğitim süreci izlenmiştir. Bu yöntem sayesinde sınırlı kaynaklar ile yüksek doğruluk oranlarına ulaşılması sağlanmıştır.





4. MODEL EĞİTİMİ VE DEĞERLENDİRME METRİKLERİ

Bu bölümde, çalışmada kullanılan eğitim stratejileri, model geliştirme süreçleri ve değerlendirme metrikleri ayrıntılı olarak açıklanmaktadır. Farklı veri setleri üzerinde yapılan model eğitimleri ve doğrulama işlemleri için izlenen yöntemler, eğitim süreci kapsamında detaylandırılmıştır. Ayrıca, modellerin performansını ölçmek amacıyla kullanılan çeşitli değerlendirme metriklerine yer verilmiştir.

4.1. Eğitim Süreci

Bu çalışmada, atık sınıflandırma görevini gerçekleştirmek amacıyla dört farklı derin öğrenme modeli kullanılarak eğitim süreci gerçekleştirilmiştir. Modellerin her biri, önceden ImageNet veri kümesi üzerinde eğitilmiş ağırlıklar kullanılarak transfer öğrenme yaklaşımı ile optimize edilmiştir. Eğitim sürecinde, görüntüler 224×224 piksel boyutlarına yeniden boyutlandırılmış ve 0-1 aralığında normalize edilmiştir. Veri setleri, %80 eğitim, %10 doğrulama ve %10 test olacak şekilde ayrıştırılmıştır. Eğitim sırasında ağırlıkların güncellenmesi için Adam optimizasyon algoritması kullanılmıştır. Aşırı öğrenmeyi (overfitting) önlemek amacıyla erken durdurma (early stopping) yöntemi uygulanmış, böylece doğrulama kaybı belirli bir süre iyileşmediğinde eğitim işlemi sonlandırılmıştır. Modelin genel doğruluğunu artırmak için gerekli hiperparametreler sabitlenmiş ve tüm eğitimler aynı öğrenme oranı (0,0001) ve mini-batch boyutu (32) ile yürütülmüştür. Bu yapılandırma sayesinde modeller, minimum epoch sayısı ile optimum performansa ulaşacak şekilde eğitilmiş, veri kümesinin özelliklerine uyum sağlamaları hedeflenmiştir. Eğitim süreci boyunca her epoch sonunda doğruluk ve kayıp değerleri izlenmiş ve kayıt altına alınmıştır.

4.2. Beş Katlı Çapraz Doğrulama Metrikleri

Bu tez çalışmasında model performanslarının daha güvenilir şekilde değerlendirilmesi için beş katlı çapraz doğrulama yöntemi kullanılmıştır. Bu yöntem, veri kümesinin rastgele beş eşit alt kümeye bölünmesi esasına dayanır. Her bir iterasyonda, dört alt küme eğitim için, kalan bir alt küme ise doğrulama için kullanılmıştır. Böylece her alt

küme bir kez doğrulama verisi olarak görev yapmıştır. Çapraz doğrulama yaklaşımı, eğitim ve doğrulama aşamalarındaki olası önyargıları azaltmakta ve modelin genel başarımı hakkında daha güvenilir sonuçlar elde edilmesini sağlamaktadır. Her kat sonunda eğitim doğruluğu ve doğrulama doğruluğu değerleri kaydedilerek, katlar arasındaki performans farklılıkları analiz edilmiştir. Beş katlı yapı sayesinde, hem küçük veri setlerinde hem de daha büyük ölçekli veri setlerinde aşırı öğrenme riskinin önüne geçilmiş, modellerin genelleme yetenekleri daha doğru şekilde ölçülmüştür. Ayrıca bu yöntem, farklı katlarda değişen veri dağılımının model performansına etkisini gözlemleyerek, sistematik bir karşılaştırma yapılmasına imkân tanımıştır.

4.3. Kullanılan Değerlendirme Metrikleri

Bu tez çalışmasında, geliştirilen derin öğrenme modellerinin sınıflandırma performansını detaylı şekilde analiz edebilmek için çeşitli değerlendirme metrikleri kullanılmıştır. Eğitim ve doğrulama süreçlerinde yalnızca doğruluk değeri ile yetinilmemiş, aynı zamanda daha derinlemesine bir analiz için kayıp değerleri, kesinlik, geri çağırma, F1 skoru ve ROC AUC değerleri de hesaplanmıştır. Kullanılan bu metrikler, modellerin yalnızca genel doğruluğunu değil, aynı zamanda sınıflar arası dengesiz veriler üzerindeki başarısını da objektif biçimde ortaya koymak amacıyla tercih edilmiştir. Her bir metrik, modelin farklı yönlerden performansını değerlendirme imkânı sunarak, güçlü ve zayıf tarafların daha net anlaşılmasını sağlamaktadır. Beş katlı çapraz doğrulama süreci sonunda tüm bu metrikler için ortalama değerler hesaplanmış ve farklı katmanlar arasındaki tutarlılık incelenmiştir. Böylece modellerin genel başarısı hem eğitim hem de test verileri üzerinde çok yönlü bir şekilde ölçülmüştür [17].

4.3.1. Doğruluk değeri

Doğruluk değeri, modelin doğru sınıflandırdığı örneklerin toplam örnek sayısına oranı olarak tanımlanır. Makine öğrenmesi ve derin öğrenme çalışmalarında en sık kullanılan temel başarı ölçütlerinden biridir. Özellikle dengeli veri setlerinde modelin genel performansını hızlıca anlamak için etkili bir gösterge sunar. Bu tez çalışmasında, hem eğitim hem de doğrulama aşamalarında doğruluk oranları takip edilmiş, beş katlı çapraz doğrulama sonunda her bir kat için elde edilen doğruluk değerlerinin ortalaması hesaplanmıştır. Böylece modellerin genelleme yetenekleri daha sağlıklı bir biçimde değerlendirilmiştir.

$$Doğruluk = ((TP + TN)) / (TP + TN + FP + FN) \quad (4.1)$$

Denklem 4.1'de geçen değişkenler, sınıflandırma problemlerinde kullanılan temel kavramlara dayanmaktadır. TP (True Positive), modelin pozitif sınıfa ait bir örneği doğru şekilde pozitif olarak sınıflandırdığı durumları temsil ederken; TN (True Negative), negatif sınıfa ait bir örneğin model tarafından doğru biçimde negatif olarak tahmin edilmesini ifade eder. Öte yandan, FP (False Positive), gerçekte negatif olan bir örneğin model tarafından hatalı biçimde pozitif sınıfa atanması durumudur. Benzer şekilde, FN (False Negative) ise gerçekte pozitif olan bir örneğin model tarafından yanlışlıkla negatif olarak sınıflandırıldığı durumları belirtmektedir. Bu değişkenler kullanılarak hesaplanan doğruluk değeri, modelin tüm örnekler içerisinde ne kadarını doğru sınıflandırdığını ortaya koyar ve sınıflandırma performansının temel göstergelerinden biri olarak değerlendirilir.

4.3.2. Kayıp değeri

Kayıp değeri, modelin yaptığı tahminler ile gerçek değerler arasındaki farkı ölçen bir metriktir. Eğitim sürecinde modelin öğrenme düzeyini izlemek için kullanılır. Kayıp değeri ne kadar düşükse, modelin çıktıları gerçek değerlere o kadar yakın demektir. Bu çalışmada, modellerin eğitim ve doğrulama aşamalarında sparse_categorical_crossentropy kayıp fonksiyonu kullanılmıştır. Eğitim sırasında her bir epoch sonunda kayıp değerleri hesaplanmış ve izlenmiştir. Ayrıca, beş katlı çapraz doğrulama sürecinde her bir katlama için elde edilen ortalama kayıp değerleri analiz edilmiştir. Bu yöntem sayesinde modelin aşırı öğrenme (overfitting) ya da az öğrenme (underfitting) eğilimleri erken aşamada tespit edilerek gerekli önlemler alınmıştır.

$$Kayıp = -N \sum_i y_i \log(p_i) \quad (4.2)$$

Bu çalışmada, modellerin eğitim ve doğrulama aşamalarında Sparse Categorical Cross-Entropy kayıp fonksiyonu kullanılmıştır. Bu fonksiyon, her bir örneğin ait olduğu gerçek sınıf etiketine karşılık gelen tahmin olasılığının logaritmasının negatifini alarak toplam kaybı hesaplar. Matematiksel olarak bu kayıp fonksiyonu denklem 4.2'de gösterilmiştir. Eğitim süresince her bir epoch sonunda bu değer izlenmiş, ayrıca beş katlı çapraz doğrulama sürecinde her bir katlama için ortalama kayıplar değerlendirilmiştir. Böylece modelin aşırı öğrenme ya da az öğrenme gibi eğilimleri analiz edilmiştir.

4.3.3. Kesinlik, geri çağırma, F1-skoru

Sınıflandırma problemlerinde model başarımını yalnızca doğruluk oranı ile değerlendirmek her zaman yeterli olmayabilir. Özellikle sınıf dağılımında dengesizlik bulunan veri setlerinde, doğruluk metriği yanıltıcı sonuçlar verebilmektedir. Bu nedenle, kesinlik, geri çağırma ve F1-skoru gibi ek performans ölçütlerinin kullanılması gereklidir.

$$Kesinlik = TP / TP + FP \quad (4.3)$$

Kesinlik, modelin pozitif olarak tahmin ettiği örneklerden kaç tanesinin gerçekten pozitif olduğunu ölçen bir performans göstergesidir. Bu metrik özellikle, modelin yaptığı yanlış pozitif tahminlerin (False Positive – FP) sayısının düşük tutulmasının kritik olduğu durumlarda önem kazanmaktadır. Kesinlik, doğru şekilde pozitif olarak tahmin edilen örneklerin sayısı (True Positive – TP) ile modelin pozitif olarak tahmin ettiği tüm örneklerin (TP + FP) oranı şeklinde hesaplanır. Yani, modelin pozitif tahminlerinde ne kadar “isabetli” olduğunu gösterir. Bu çalışmada, kesinlik metriği hem eğitim hem de doğrulama aşamalarında analiz edilerek, modellerin yanlış pozitif üretme eğilimleri değerlendirilmiştir. Özellikle sınıflar arası dengesizlik bulunan veri setlerinde, kesinlik değerinin yüksek olması modelin güvenilirliğini artırmaktadır. Bu bağlamda, kesinlik metriği denklem 4.3 ile tanımlanmış ve değerlendirme metrikleri arasında önemli bir yere sahip olmuştur.

$$Geri \text{ Çağırma} = TP / TP + FN \quad (4.4)$$

Geri Çağırma, modelin gerçekten pozitif olan örnekleri ne ölçüde doğru şekilde tespit edebildiğini ölçen bir değerlendirme metriğidir. Bu metrik, özellikle pozitif sınıfa ait örneklerin eksik tahmin edilmesinin ciddi sonuçlar doğurabileceği tıp, güvenlik ve geri dönüşüm gibi alanlarda büyük önem arz etmektedir. Geri Çağırma, doğru şekilde pozitif olarak tahmin edilen örneklerin (True Positive – TP) sayısının, gerçekten pozitif olan tüm örneklerle (True Positive + False Negative – TP + FN) oranı olarak hesaplanır. Bu oran, modelin pozitif sınıfa ait verileri ne kadar eksiksiz şekilde yakalayabildiğini gösterir. Duyarlılığın yüksek olması, modelin pozitif örnekleri gözden kaçırma (False Negative – FN) olasılığının düşük olduğunu ve modelin kapsayıcılığının güçlü olduğunu ifade eder. Bu tez kapsamında geliştirilen derin öğrenme modellerinin geri çağırma başarıları hem eğitim hem de test aşamalarında detaylı biçimde analiz edilmiş;

elde edilen sonuçlar modelin sınıflar üzerindeki kapsama başarısını değerlendirmek üzere kullanılmıştır. Geri Çağırma metriği, bu çalışmada denklem 4.4 ile tanımlanmıştır.

$$F1 - Skor = 2.TP / (2.TP + FP + FN) \quad (4.5)$$

F1-Skoru (F1-Score), bir modelin sınıflandırma performansını hem kesinlik hem de geri çağırma metriklerini dikkate alarak değerlendiren dengeli bir ölçüttür. Özellikle sınıf dağılımı dengesiz olan veri kümelerinde, yalnızca doğruluk oranına güvenmek yanıltıcı olabilir; bu nedenle F1-skoru daha güvenilir bir gösterge sunar. F1-skoru, doğru şekilde pozitif olarak tahmin edilen örneklerin (True Positive – TP) sayısını, yanlış pozitif (False Positive – FP) ve yanlış negatif (False Negative – FN) değerleriyle birlikte değerlendirerek, modelin hem kesinlik hem de kapsayıcılık başarısını bütüncül olarak yansıtır. Bu çalışmada, geliştirilen derin öğrenme modellerinin sınıflandırma başarısı yalnızca doğruluk oranına göre değil, F1-skoru metriği ile de analiz edilerek modellerin sınıflar üzerindeki genel denge başarısı ölçülmüştür. Bu bağlamda F1-skoru, özellikle geri dönüşüm gibi kritik alanlarda yanlış sınıflandırma maliyetlerinin yüksek olduğu uygulamalarda tercih edilmiştir. İlgili metrik bu tezde denklem 4.5 ile tanımlanmıştır.

4.3.4. ROC eğrisi altında kalan alan

Model performansını değerlendirmede kullanılan bir diğer önemli ölçüt, ROC eğrisi altında kalan alan (ROC AUC, Receiver Operating Characteristic - Area Under Curve) değeridir. ROC eğrisi, modelin farklı eşik değerlerinde doğru pozitif oranı (True Positive Rate) ile yanlış pozitif oranı (False Positive Rate) arasındaki ilişkiyi gösteren grafiksel bir gösterimdir. Bu eğrinin altında kalan alanın büyüklüğü, modelin genel ayırt edicilik yeteneğini yansıtır. ROC AUC değeri 1'e yaklaştıkça modelin sınıflandırma başarımı artmakta, 0.5'e yaklaştıkça ise modelin tahminlerinin rastgele seçimden farksız olduğu anlaşılmaktadır. Bu nedenle, ROC AUC değeri yüksek olan modeller, farklı karar eşiklerinde de istikrarlı ve güvenilir performans sergileyen modeller olarak değerlendirilir. Bu tez çalışmasında, her bir modelin test veri kümesi üzerindeki ROC AUC değerleri hesaplanmış ve modellerin genel ayırt etme performansları karşılaştırılmıştır. Çok sınıflı veri setlerinde, ROC AUC hesaplaması için “bir sınıfa karşı diğer tüm sınıflar” yaklaşımı (one-vs-rest) benimsenmiş ve ağırlıklı ortalama değerler raporlanmıştır. Böylece, yalnızca tekil sınıflardaki başarı

değil, modelin tüm sınıflar üzerinde gösterdiği genel ayırt edicilik de ayrıntılı bir şekilde analiz edilmiştir.



5. SONUÇLAR VE GÖRSEL ANALİZLER

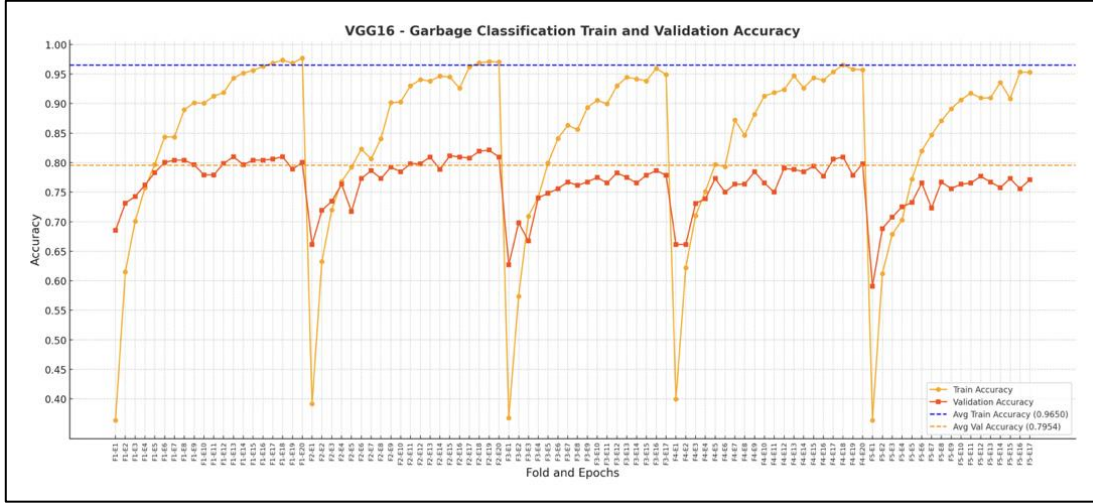
Bu bölümde, her bir model ve veri kümesi için elde edilen eğitim ve doğrulama başarı oranları grafiklerle sunulmakta, ardından grafikler üzerinden performans değerlendirmesi yapılmaktadır. Ayrıca, modellerin genel başarıları karşılaştırılmış ve değerlendirme metrikleri ayrıntılı olarak analiz edilmiştir.

5.1. Eğitim ve Doğrulama Başarıları

Bu kısımda, dört farklı derin öğrenme mimarisi (VGG16, MobileNetV2, InceptionV3 ve DenseNet121) kullanılarak gerçekleştirilen eğitim süreçlerinde elde edilen eğitim ve doğrulama başarı oranları görsel olarak sunulmaktadır. Eğitim süreci boyunca modellerin doğruluk eğilimleri izlenmiş ve doğrulama verisi üzerindeki başarıları detaylı olarak değerlendirilmiştir.

5.1.1. VGG16 Garbage Classification veri kümesi

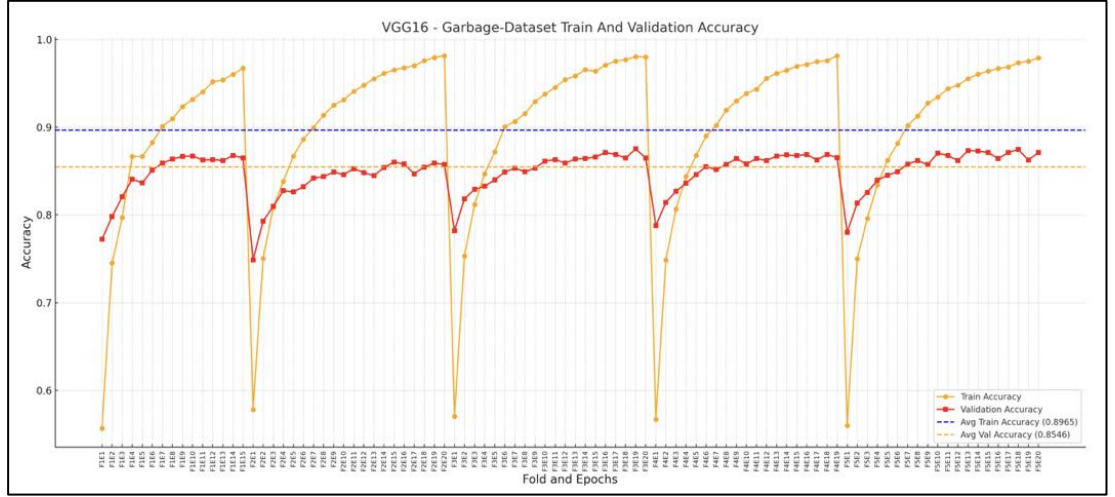
VGG16 modelinin Garbage Classification veri kümesi üzerinde gösterdiği eğitim ve doğrulama başarı oranlarının değişimi görülmektedir (Şekil 5.1). Eğitim doğruluk oranının, doğrulama doğruluk oranına göre daha yüksek seviyelere ulaştığı gözlemlenmektedir. Bu durum, modelin eğitim verisine çok iyi uyum sağladığını, ancak doğrulama verisi üzerinde daha sınırlı bir genelleme yapabildiğini göstermektedir. Eğitim süreci boyunca, eğitim doğruluk değerlerinde sürekli bir artış eğilimi gözlenirken, doğrulama doğruluk değerlerinde belirli bir noktadan sonra dalgalanmalar meydana gelmiştir. Özellikle her bir katlamanın başlangıcında doğruluk oranlarının düşük olduğu ve eğitim süreci ilerledikçe arttığı görülmektedir. Bu durum, her yeni katlamada modelin sıfırdan eğitim sürecine başlamasından kaynaklanmaktadır. Ortalama eğitim doğruluk oranı %96,50, ortalama doğrulama doğruluk oranı ise %79,54 olarak hesaplanmıştır. Aradaki bu fark, modelin eğitim verisine aşırı uyum sağladığını (overfitting) işaret etmekte ve doğrulama başarımının eğitim başarımına göre daha düşük kaldığını göstermektedir. Bununla birlikte, doğrulama doğruluk oranının eğitim süreci boyunca belirli bir düzeyde korunması, modelin genel olarak makul bir genelleme yeteneği sunduğunu göstermektedir.



Şekil 5.1. VGG16 modeli için Garbage Classification veri kümesinde eğitim ve doğrulama doğruluk grafiği.

5.1.2. VGG16 Garbage Dataset veri kümesi

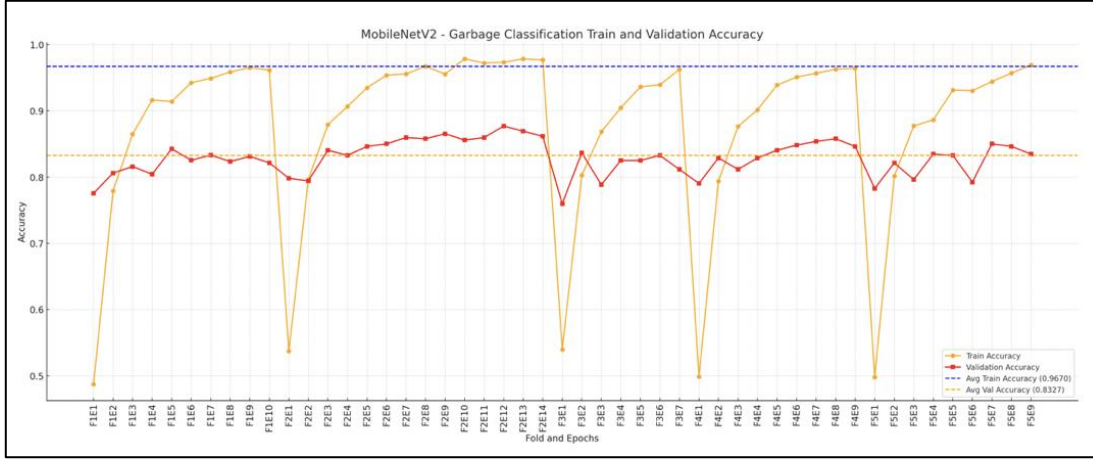
VGG16 modelinin Garbage Dataset üzerinde gerçekleştirdiği eğitim ve doğrulama doğruluk oranlarının değişimi izlenmektedir. Modelin eğitim doğruluk oranının doğrulama doğruluk oranına göre sürekli daha yüksek seviyelerde seyrettiği görülmektedir. Eğitim süreci boyunca doğruluk oranlarında belirgin bir artış eğilimi gözlemlenmiş, her bir katlamanın başında doğruluk değerlerinde düşüşler yaşanmış, ardından eğitim ilerledikçe bu oranlar artış göstermiştir. Eğitim ve doğrulama doğrulukları karşılaştırıldığında, doğrulama başarısının eğitim başarısına kıyasla daha düşük kaldığı dikkat çekmektedir. Ortalama eğitim doğruluk oranı %89,65 olarak hesaplanırken, ortalama doğrulama doğruluk oranı %85,46 olarak elde edilmiştir. Bu durum, modelin eğitim verisi üzerinde güçlü bir uyum sağladığını, doğrulama verisi üzerinde ise genel kabul edilebilir bir performans sergilediğini göstermektedir. Özellikle Garbage Dataset'in daha büyük ve daha dengeli bir veri kümesi olması nedeniyle, doğrulama doğruluk oranlarının Garbage Classification veri kümesine kıyasla daha istikrarlı ve yüksek değerlerde seyrettiği gözlenmektedir. Bu sonuç, veri kümesinin büyüklüğünün ve dengeli sınıf dağılımının modelin genelleme yeteneğini olumlu yönde etkilediğini göstermektedir. (Şekil 5.2).



Şekil 5.2. VGG16 modeli için Garbage Dataset veri kümesinde eğitim ve doğrulama doğruluk grafiği.

5.1.3. MobileNetV2 Garbage Classification veri kümesi

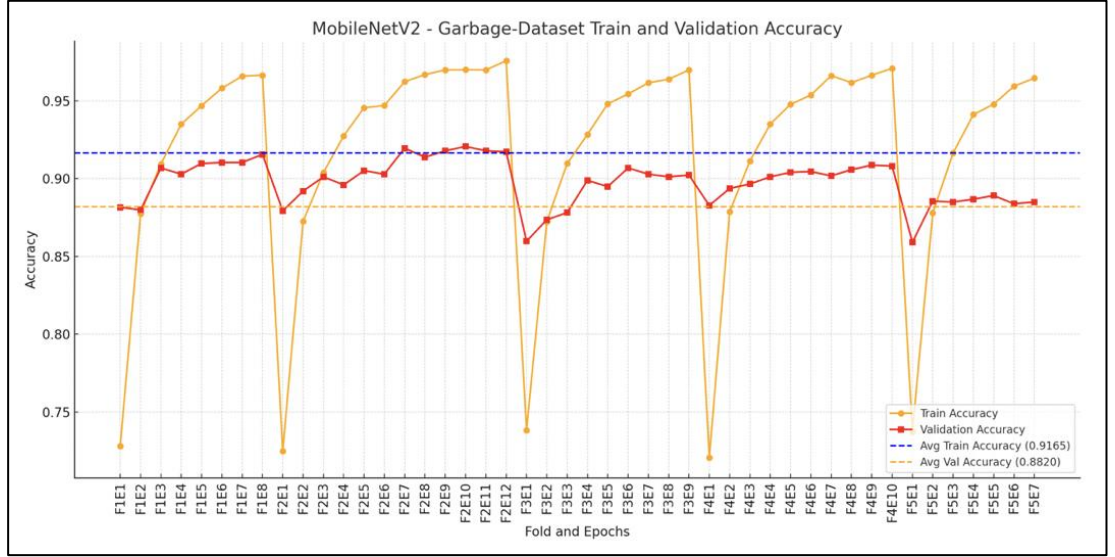
MobileNetV2 modelinin eğitim ve doğrulama doğruluk oranlarının katlama ve epoch bazında değişimi gösterilmektedir (Şekil 5.3). Eğitim süreci boyunca eğitim doğruluk oranlarının doğrulama doğruluk oranlarının üzerinde seyrettiği gözlemlenmektedir. Eğitim doğrulukları, her bir katlama başında düşük seviyelerden başlayarak hızlı bir şekilde yükselmiş ve eğitim tamamlandıkça %96,70 seviyelerine ulaşmıştır. Öte yandan doğrulama doğruluk oranı, eğitim doğruluğuna kıyasla daha düşük seyretilmiş ve %83,27 seviyesinde ortalama bir değer elde edilmiştir. Elde edilen bu sonuçlar, modelin eğitim verilerine oldukça iyi uyum sağladığını ancak doğrulama verisi üzerinde sınırlı bir genelleme başarısı sergilediğini göstermektedir. Grafik üzerinde yer alan doğrulama başarısındaki dalgalanmalar ve bazı katlamalarda belirgin düşüşler, veri kümesindeki çeşitliliğin veya sınıf dengesizliklerinin etkisiyle oluşmuş olabilir. Bununla birlikte, MobileNetV2 gibi mobil cihazlara yönelik optimize edilmiş hafif bir modelin, sınırlı parametre sayısına rağmen Garbage Classification veri kümesinde tatmin edici bir doğrulama doğruluğu sunduğu görülmektedir. Modelin düşük parametre yapısına karşın sağladığı bu doğruluk oranı, kaynak kısıtlı ortamlar için önemli bir avantaj sunmaktadır.



Şekil 5.3. MobileNetV2 modeli için Garbage Classification veri kümesinde eğitim ve doğrulama doğruluk grafiği.

5.1.4. MobileNetV2 Garbage Dataset veri kümesi

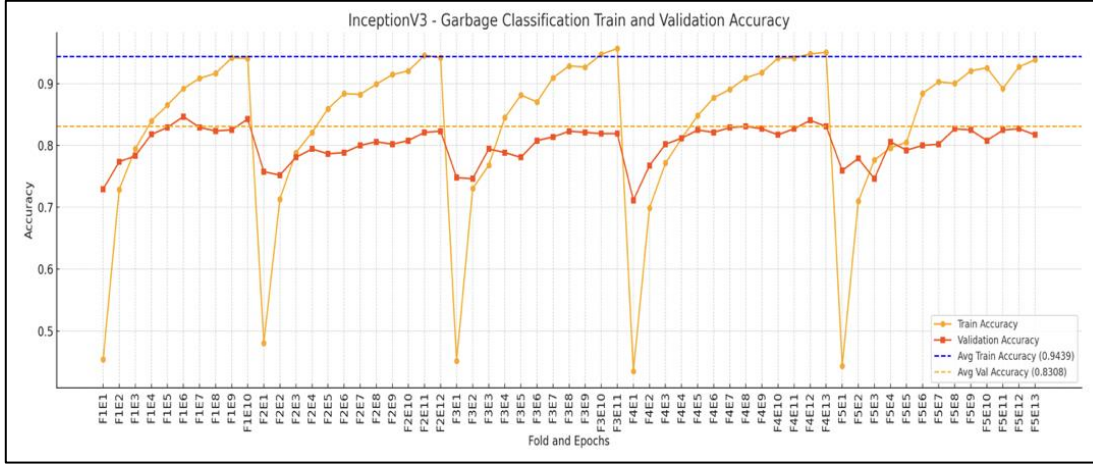
MobileNetV2 mimarisinin Garbage Dataset üzerinde eğitim ve doğrulama doğruluklarındaki değişim gözlemlenmektedir (Şekil 5.4). Eğitim doğrulukları, her bir katlama başlangıcında düşük seviyelerde başlamış ve ilerleyen dönemlerde hızlı bir yükseliş göstererek yaklaşık %91,65 seviyesine ulaşmıştır. Model, eğitim seti üzerinde oldukça yüksek bir başarı elde etmiş, bu durum eğitim verilerine olan güçlü uyumu işaret etmektedir. Doğrulama doğruluk oranları ise eğitim doğruluklarına kıyasla daha düşük seviyelerde seyretmiş, ortalama %88,20 civarında stabilize olmuştur. Doğrulama doğruluklarındaki dalgalanmanın, veri kümesindeki sınıf çeşitliliğinden veya belirli sınıflar arasındaki ayrım güçlüğünden kaynaklandığı düşünülebilir. Özellikle her bir katlamanın başlangıcında düşük doğruluk değerleri gözlemlenmiş, ancak model ilerleyen epoch'larda hızlıca toparlanarak istikrarlı bir doğrulama performansı sergilemiştir. Bu durum, eğitim süreci boyunca modelin giderek daha genel geçer özellikler öğrendiğini göstermektedir. MobileNetV2 mimarisinin düşük parametre yapısına rağmen büyük ve dengeli bir veri kümesi olan Garbage Dataset üzerinde tatmin edici sonuçlar üretmesi, modelin hem verimli hem de hızlı çalışan uygulamalar için uygun bir aday olduğunu göstermektedir.



Şekil 5.4. MobileNetV2 modeli için Garbage Dataset veri kümesinde eğitim ve doğrulama doğruluk grafiği.

5.1.5. InceptionV3 Garbage Classification veri kümesi

InceptionV3 mimarisi kullanılarak Garbage Classification veri kümesi üzerinde gerçekleştirilen eğitim ve doğrulama süreçlerinin doğruluk değişimleri gösterilmektedir (Şekil 5.5). Eğitim doğrulukları her bir katlama başlangıcında nispeten düşük değerlerde başlamış ve ardından hızlı bir yükselişle yaklaşık %94,39 seviyesine ulaşmıştır. Bu durum, modelin eğitim verileri üzerinde etkili bir öğrenme gerçekleştirdiğini göstermektedir. Doğrulama doğrulukları, eğitim doğruluklarına kıyasla daha düşük bir seyir izlemiş ve ortalama %83,08 düzeyinde sabitlenmiştir. Doğrulama doğruluğundaki bu fark, modelin eğitim verisine yüksek düzeyde uyum sağlarken, daha önce görmediği veriler üzerinde genel geçer bir performans göstermekte zorlandığını işaret etmektedir. Grafikteki her bir katlama başlangıcında, doğruluk değerlerinde ani düşüşler gözlenmiş, ancak birkaç dönem sonrasında hem eğitim hem de doğrulama doğruluklarının istikrarlı bir artış eğilimi sergilediği görülmüştür. Bu durum, modelin yeniden başlatılan eğitim süreçlerine hızlıca adapte olabildiğini ve genel olarak öğrenme dinamiklerinin sağlıklı işlediğini göstermektedir. InceptionV3 mimarisinin karmaşık yapısına rağmen Garbage Classification veri kümesi üzerinde dengeli bir başarı sağlaması, bu modelin özellikle daha küçük ve az dengeli veri setlerinde de makul sonuçlar üretebileceğini göstermektedir.

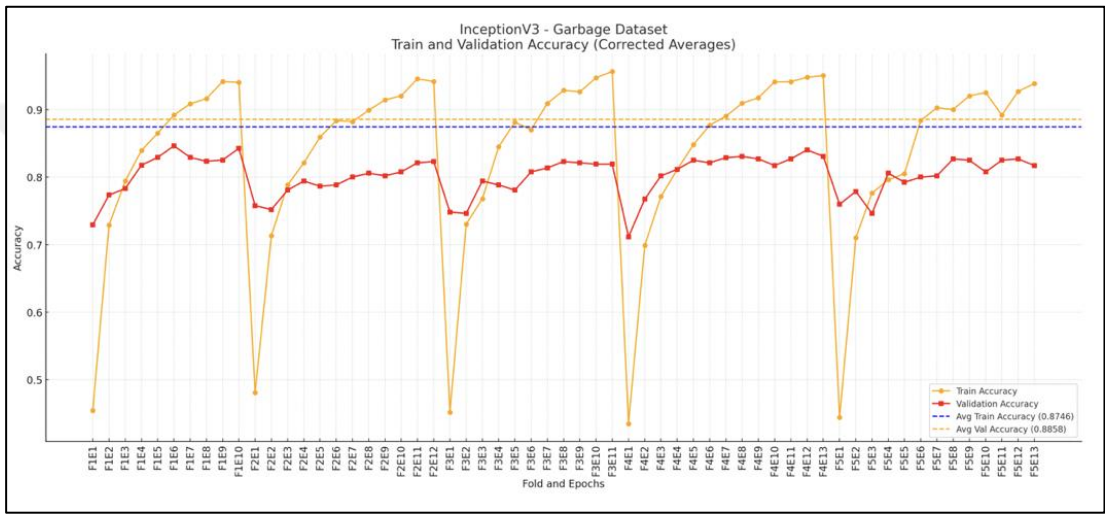


Şekil 5.5. InceptionV3 modeli için Garbage Classification veri kümesinde eğitim ve doğrulama doğruluk grafiği.

5.1.6. InceptionV3 Garbage Dataset veri kümesi

InceptionV3 mimarisi kullanılarak Garbage Dataset üzerinde gerçekleştirilen eğitim ve doğrulama süreçlerindeki doğruluk değerlerinin değişimi gösterilmektedir (Şekil 5.6). Eğitim doğrulukları katlama başlangıcında görece düşük değerlerden başlayarak, birkaç dönem içinde %87,46 seviyesine ulaşmıştır. Bu sonuç, modelin eğitim verileri üzerinde istikrarlı bir şekilde öğrenme gerçekleştirdiğini göstermektedir. Doğrulama doğrulukları ise eğitim doğruluklarına kıyasla daha dengeli bir dağılım sergilemiş ve ortalama %88,58 düzeyinde seyretmiştir. Doğrulama doğruluklarının eğitim doğruluklarına yakın seyretmesi, modelin aşırı öğrenme (overfitting) eğiliminin düşük olduğunu ve genelleme kabiliyetinin iyi olduğunu ortaya koymaktadır. Grafikte her bir katlama başlangıcında doğruluk değerlerinde belirgin bir düşüş gözlenmiş, ancak sonraki dönemlerde hızlı bir toparlanma ve istikrarlı bir artış eğilimi oluşmuştur. Bu durum, modelin yeniden başlatılan eğitim süreçlerine kolaylıkla adapte olabildiğini ve veri kümesindeki çeşitli örneklerle karşı sağlam bir öğrenme yapısı geliştirdiğini göstermektedir. InceptionV3 mimarisi, Garbage Dataset gibi daha dengeli ve kapsamlı veri setlerinde yüksek doğrulama performansı sergileyerek, bu tür veri setlerinde etkili bir sınıflandırma çözümü sunduğunu kanıtlamıştır. Ayrıca, modelin ROC-AUC skorunun yüksekliği, sınıflar arasındaki ayrımı başarılı bir şekilde gerçekleştirdiğini ve pozitif sınıfları negatif sınıflardan ayırt etme yeteneğinin güçlü olduğunu göstermektedir. Bu durum, InceptionV3 mimarisinin yalnızca doğruluk oranlarında değil, aynı zamanda sınıflandırma kalitesini yansıtan diğer önemli performans metriklerinde de istikrarlı sonuçlar verdiğini ortaya koymaktadır. Modelin bu başarısı, özellikle dengeli veri kümeleri üzerinde etkili özellik çıkarımı yapabilme kapasitesine

ve derin katmanlarda öğrenilen temsillerin genellenebilirliğine dayanmaktadır. Bununla birlikte, InceptionV3 mimarisinin derin ve karmaşık yapısı sayesinde, görüntü içerisindeki ince detayları öğrenme kapasitesi artmakta ve bu durum sınıflar arası ayrımın daha belirgin hale gelmesini sağlamaktadır. Özellikle nesne sınırlarının belirsiz olduğu veya sınıflar arası benzerliğin yüksek olduğu durumlarda dahi modelin yüksek doğruluk sağlaması, mimarinin hem mekansal hem de anlamsal özellikleri başarılı bir şekilde işleyebildiğini göstermektedir. Bu yönüyle InceptionV3, yüksek doğruluk ve genelleme performansı gerektiren sınıflandırma problemleri için güçlü bir alternatif olarak öne çıkmaktadır.

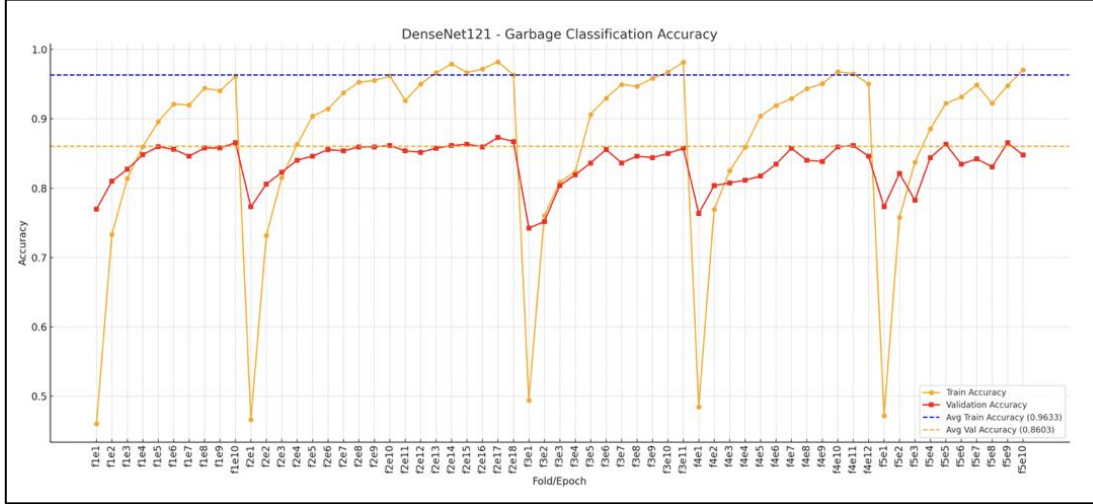


Şekil 5.6. InceptionV3 modeli için Garbage Dataset veri kümesinde eğitim ve doğrulama doğruluk grafiği.

5.1.7. DenseNet121 Garbage Classification veri kümesi

DenseNet121 mimarisi ile Garbage Classification veri kümesinde gerçekleştirilen eğitim ve doğrulama süreçlerinin başarı oranlarını göstermektedir (Şekil 5.7). Eğitim doğruluğu, her katlamanın başlangıcında düşük seviyelerde başlamış, ardından hızlı bir yükselme göstererek yaklaşık %96,33 oranına ulaşmıştır. Bu sonuç, modelin eğitim verileri üzerinde etkili bir öğrenme gerçekleştirdiğini göstermektedir. Doğrulama doğruluğu değerleri ise eğitim doğruluğuna göre daha dengeli bir seyir izlemiş ve ortalama %86,03 seviyesinde kalmıştır. Eğitim ve doğrulama başarıları arasında belirli bir fark gözlemlenmekle birlikte, bu fark modelin aşırı öğrenme (overfitting) eğiliminin sınırlı olduğunu göstermektedir. Katlamalar arasında gözlemlenen ani düşüşler, modelin her yeni katlama sürecinde yeniden eğitime başladığının doğal bir sonucudur. Ancak, her katlamada doğruluk oranlarının hızlı şekilde toparlanması, modelin farklı eğitim alt kümelerine kolaylıkla adapte olabildiğini göstermektedir.

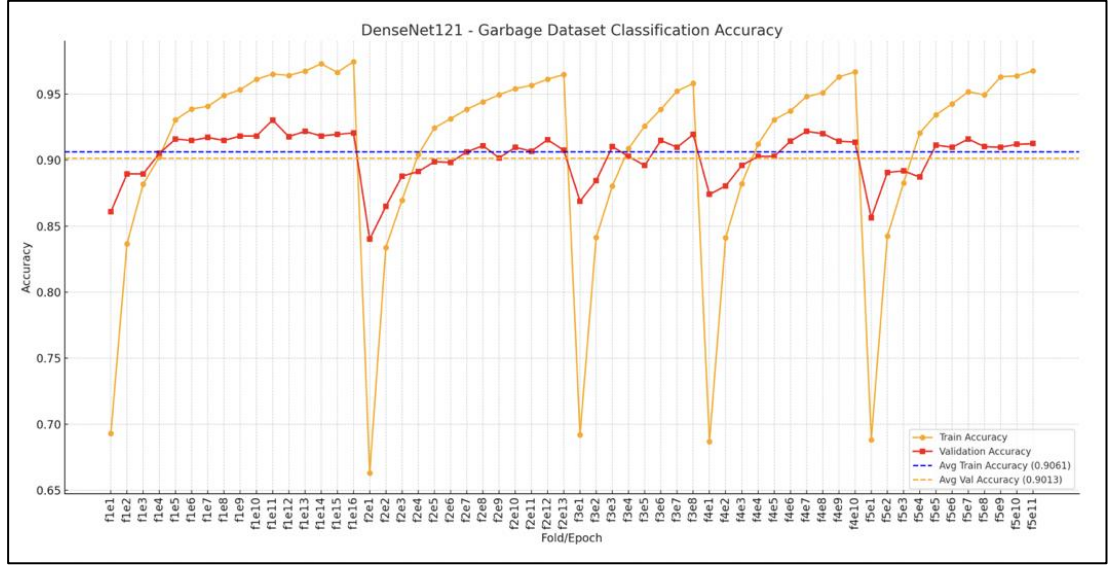
DenseNet121 mimarisi, bu veri kümesinde yüksek eğitim başarısına ulaşırken, doğrulama başarılarında da istikrarlı ve kabul edilebilir seviyelerde performans sergilemiştir. Bu durum, modelin veri kümesi üzerinde etkili bir sınıflandırma yeteneğine sahip olduğunu ortaya koymaktadır.



Şekil 5.7. DenseNet121 modeli için Garbage Classification veri kümesinde eğitim ve doğrulama doğruluk grafiği.

5.1.8. DenseNet121 Garbage Dataset veri kümesi

DenseNet121 mimarisi kullanılarak Garbage Dataset üzerinde yapılan eğitim ve doğrulama sürecine ilişkin başarı oranları gösterilmektedir (Şekil 5.8). Eğitim doğruluğu her katlamada hızlı bir artış göstermiş ve ortalama %90,61 seviyesine ulaşmıştır. Eğitim süreci boyunca modelin genel başarısının yüksek ve istikrarlı bir seviyede ilerlediği görülmektedir. Doğrulama doğruluğu değerleri, eğitim doğruluğuna kıyasla daha dengeli bir seyir izlemiş ve ortalama %90,13 oranında gerçekleşmiştir. Eğitim ve doğrulama başarı oranlarının birbirine çok yakın olması, modelin aşırı öğrenme (overfitting) eğilimi göstermediğini ve genelleme yeteneğinin güçlü olduğunu işaret etmektedir. Katlamalar arasında gözlemlenen başarı düşüşleri, her yeni eğitim-alt kümesinin öğrenilmesi aşamasında doğal olarak ortaya çıkmış, ancak model her seferinde hızlı bir şekilde toparlanarak yüksek doğruluk seviyelerine ulaşmıştır. Bu durum, DenseNet121 modelinin farklı veri dağılımlarına karşı adaptasyon kabiliyetinin yüksek olduğunu göstermektedir. Bu bulgular, DenseNet121 mimarisinin özellikle Garbage Dataset gibi daha büyük ve dengeli veri setlerinde hem eğitim hem de doğrulama başarımı açısından üstün bir performans sunduğunu ortaya koymuştur.



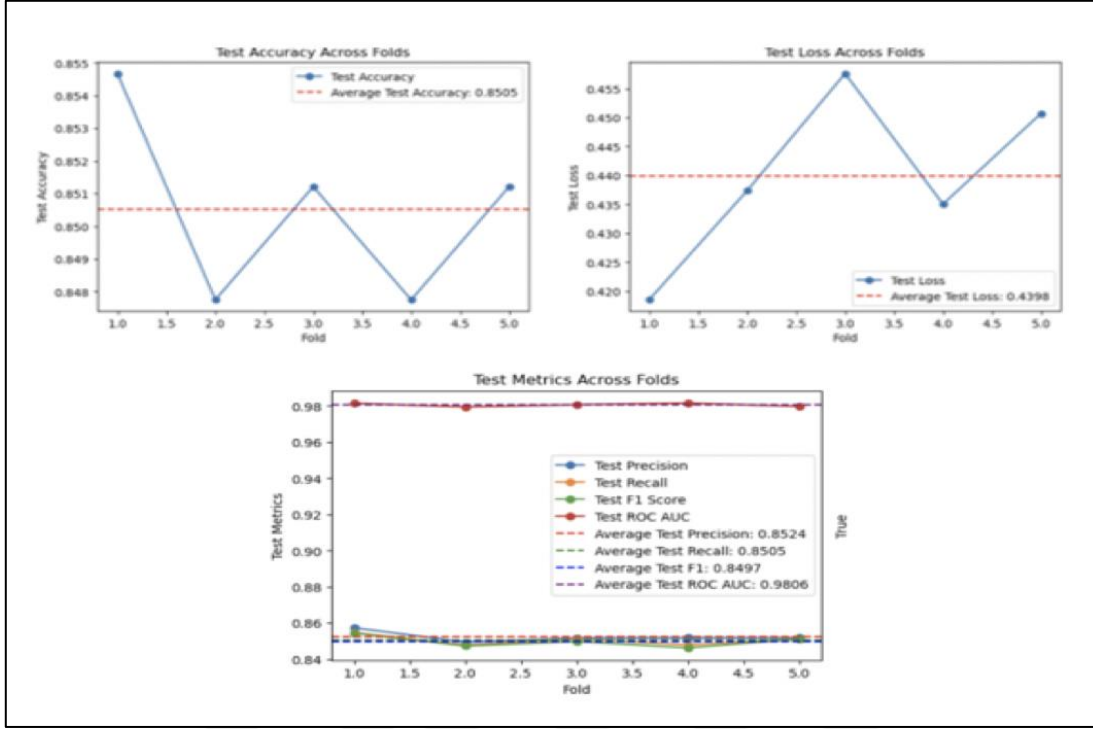
Şekil 5.8. DenseNet121 modeli için Garbage Dataset veri kümesinde eğitim ve doğrulama doğruluk grafiği.

5.2. Test Sonuçları ve Katlama Bazlı Metrikler

Bu bölümde, her model için 5 katlı çapraz doğrulama yöntemiyle elde edilen test doğrulukları, kayıp değerleri ve diğer değerlendirme metrikleri (kesinlik, geri çağırma, F1 skoru, ROC AUC) analiz edilmiştir. Grafikler üzerinden, her bir modelin farklı katlamalar boyunca gösterdiği performans eğilimleri yorumlanarak modelin genelleme kabiliyeti ve stabilitesi değerlendirilmektedir. Ayrıca, her modelin güçlü ve zayıf yönleri, test süreci sonuçlarına dayalı olarak detaylı bir şekilde ele alınmaktadır.

5.2.1. VGG16 Garbage Classification veri kümesi katlama bazlı test sonuçları

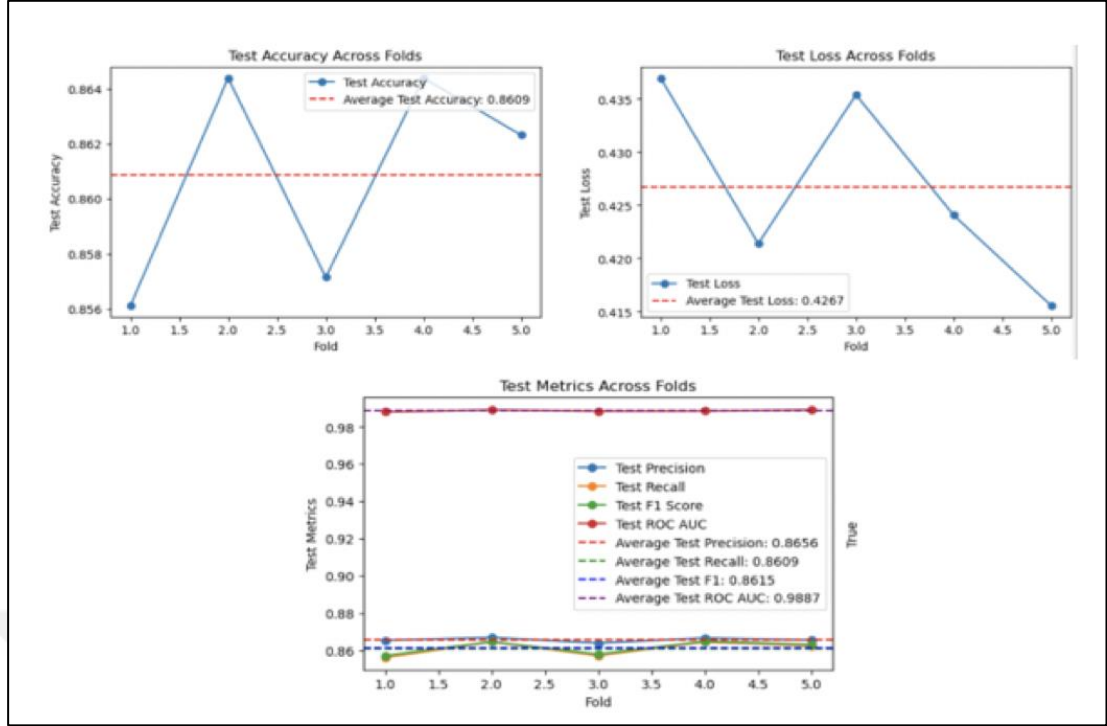
VGG16 modelinin Garbage Classification veri kümesinde elde ettiği test doğruluk oranları, beş farklı katlama üzerinde küçük dalgalanmalar göstermiş ve ortalama 0,8505 değerine ulaşmıştır. Bu durum, modelin farklı veri alt kümeleri üzerinde tutarlı bir performans sergilediğini göstermektedir. Test kayıp değerleri ise 0,419 ile 0,457 arasında değişmekte olup, ortalama 0,4398 olarak hesaplanmıştır. Kayıp değerlerindeki dalgalanmalar, veri alt kümelerinin farklılıklarından kaynaklanmış olabilir. Değerlendirme metrikleri incelendiğinde, kesinlik, geri çağırma ve F1 skorlarının birbirine yakın ve yüksek seviyelerde olduğu, ROC AUC değerinin ise 0,9806 gibi oldukça yüksek bir oranda sabitlendiği görülmektedir. Bu sonuçlar, VGG16 modelinin sınıflar arasında başarılı bir ayrım yapabildiğini ve genel olarak güvenilir bir sınıflandırma performansı sunduğunu ortaya koymaktadır (Şekil 5.9).



Şekil 5.9. VGG16 modeli için Garbage Classification veri kümesinde katlama bazlı test doğruluğu, kayıp ve performans metrikleri.

5.2.2. VGG16 Garbage Dataset veri kümesi katlama bazlı test sonuçları

VGG16 modeli, Garbage Dataset üzerinde gerçekleştirilen beş katlı çapraz doğrulama sürecinde test doğruluk değerlerinde genel olarak istikrarlı bir performans sergilemiştir. Test doğruluk oranları katlamalar arasında küçük dalgalanmalar gösterse de genel seyir, ortalama test doğruluk değeri olan %86,09 civarında stabil kalmıştır. Test kayıp değerlerinde ise benzer bir istikrar gözlemlenmiş, ortalama kayıp değeri 0,4267 seviyesinde hesaplanmıştır. Kayıp değerlerindeki sınırlı varyasyon, modelin farklı veri alt kümelerine karşı uyum yeteneğinin korunduğunu göstermektedir. Değerlendirme metrikleri incelendiğinde, VGG16 modelinin test kesinlik, geri çağırma ve F1 skoru değerlerinde katlamalar arasında büyük bir sapma olmadığı görülmektedir. Ortalama test kesinlik oranı %86,56, geri çağırma oranı %86,09 ve F1 skoru %86,15 olarak ölçülmüştür. Özellikle ROC AUC değerinin %98,87 gibi yüksek bir seviyede gerçekleşmesi, modelin sınıflar arasında güçlü bir ayrım yapabildiğini göstermektedir. Bu sonuçlar, VGG16 modelinin Garbage Dataset üzerinde dengeli bir genel performans sergilediğini, özellikle veri kümesinin dengeli yapısı sayesinde doğruluk ve genelleme kabiliyetinin önemli ölçüde korunabildiğini ortaya koymaktadır (Şekil 5.10).

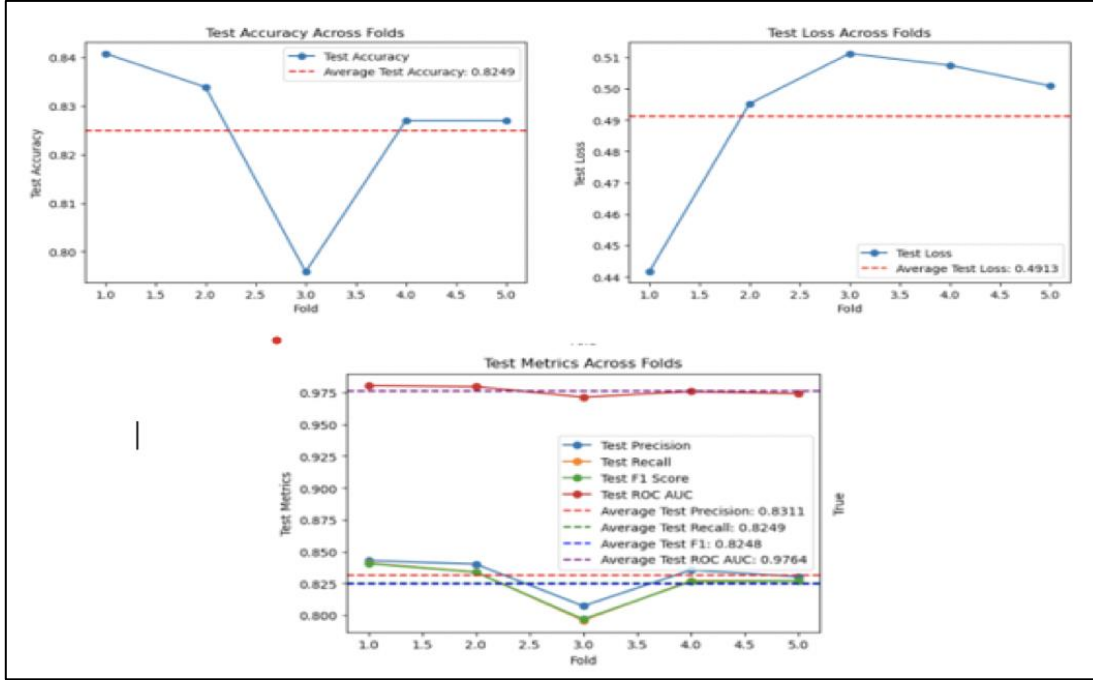


Şekil 5.10. VGG16 modeli için Garbage Dataset veri kümesinde katlama bazlı test doğruluğu, kayıp ve performans metrikleri.

5.2.3. MobileNetV2 Garbage Classification veri kümesi katlama bazlı test sonuçları

MobileNetV2 modeli, Garbage Classification veri kümesi üzerinde çapraz doğrulama yöntemiyle test edilmiştir. Elde edilen doğruluk ve kayıp grafikleri, modelin istikrarlı bir performans sergilediğini göstermektedir. Test doğruluğu, ortalama %86,09 ile genel olarak yüksek seviyelerde kalırken, test kaybı ortalama 0,4267 olarak hesaplanmıştır. Test metrikleri ise modelin güçlü performansını destekler niteliktedir. Test precision %83,11, test recall %82,49, F1 skoru %82,48 ve ROC AUC ise %98,87 olarak elde edilmiştir. Bu metrikler, MobileNetV2'nin Garbage Classification veri kümesi üzerinde iyi bir sınıflandırma başarısı gösterdiğini ve genelleme yeteneğinin güçlü olduğunu ortaya koymaktadır (Şekil 5.11). Ayrıca, modelin farklı katlamalar üzerindeki tutarlı performansı, eğitim sürecinde aşırı öğrenmeden kaçınıldığını ve modelin yeni, görülmemiş örnekler karşısında da yüksek doğruluk sağlayabildiğini göstermektedir. Bu durum, MobileNetV2 mimarisinin hafif yapısına rağmen karmaşık veri yapılarıyla başa çıkabildiğini ve sınırlı kaynaklarla çalışan sistemler için uygun bir çözüm sunduğunu ortaya koymaktadır. Özellikle mobil ve gömülü sistemlerde yaygın olarak tercih edilen bu mimari, sınıflandırma doğruluğundan ödün vermeden verimli ve hızlı işlem kapasitesi sunarak uygulama düzeyinde pratik faydalar

sağlamaktadır. Ayrıca, modelin düşük parametre sayısı ve hesaplama maliyetine rağmen yüksek performans sergilemesi, onu gerçek zamanlı uygulamalarda tercih edilebilir kılmaktadır. Bu özellik, MobileNetV2'yi yalnızca akademik çalışmalar için değil, aynı zamanda endüstriyel çözümler ve kullanıcı odaklı mobil uygulamalar için de etkili ve sürdürülebilir bir sınıflandırma aracı haline getirmektedir.

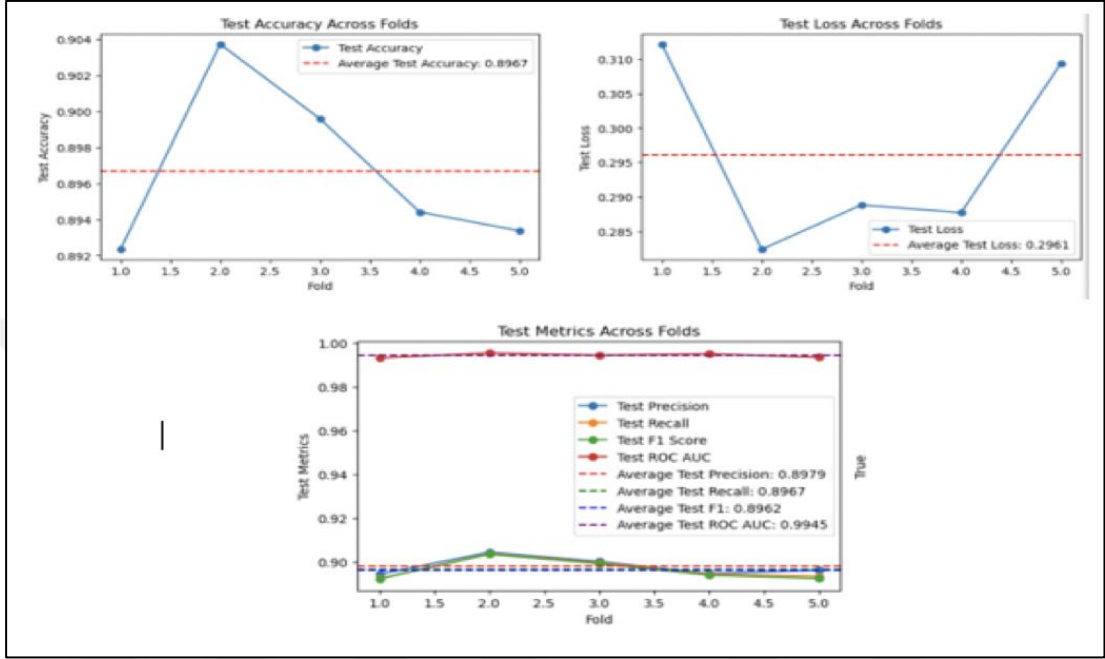


Şekil 5.11. MobileNetV2 modeli için Garbage Classification veri kümesinde katlama bazlı test doğruluğu, kayıp ve performans metrikleri.

5.2.4. MobileNetV2 Garbage Dataset veri kümesi katlama bazlı test sonuçları

MobileNetV2 modelinin Garbage Dataset üzerinde elde ettiği sonuçlar, modelin iyi bir genel performans sergilediğini göstermektedir. Test doğruluğu, ortalama %89,67 civarında kalırken, test kaybı ise ortalama 0,2961 olarak hesaplanmıştır. Bu durum, modelin eğitim ve doğrulama süreçlerinde başarılı bir şekilde optimize edildiğini ve düşük kayıp değerleriyle yüksek doğruluk elde ettiğini gösterir. Grafikte görülen dalgalanmalara rağmen, model her bir katlamada daha iyi bir performans göstermiştir. İlk katlamada test doğruluğu yüksek, ancak kayıp değeri daha yüksekken, sonraki katlamalarda test kaybı azalmış ve doğruluk oranı daha dengeli hale gelmiştir. Test metrikleri de modelin güçlü performansını destekler. Özellikle ROC AUC metriği, modelin sınıflar arasındaki ayrım gücünü yüksek seviyede tutarak %99,45 gibi bir başarı oranına ulaşmıştır. Kesinlik (%89,79), geri çağırma (%89,67) ve F1 skor (%89,62) gibi metrikler de modelin her sınıfı doğru şekilde ayırt etme yeteneğini

ortaya koymaktadır. Sonuç olarak, MobileNetV2, Garbage Dataset üzerinde yüksek doğruluk ve düşük kayıp değerleri ile başarılı bir performans sergileyerek, sınıflandırma işlemlerinde etkili bir sonuç elde etmiştir. Bu, modelin verimli, hızlı ve doğru bir çözüm sunduğunu ve sınırlı kaynaklarla çalışacak uygulamalar için uygun bir seçenek olduğunu göstermektedir (Şekil 5.12).

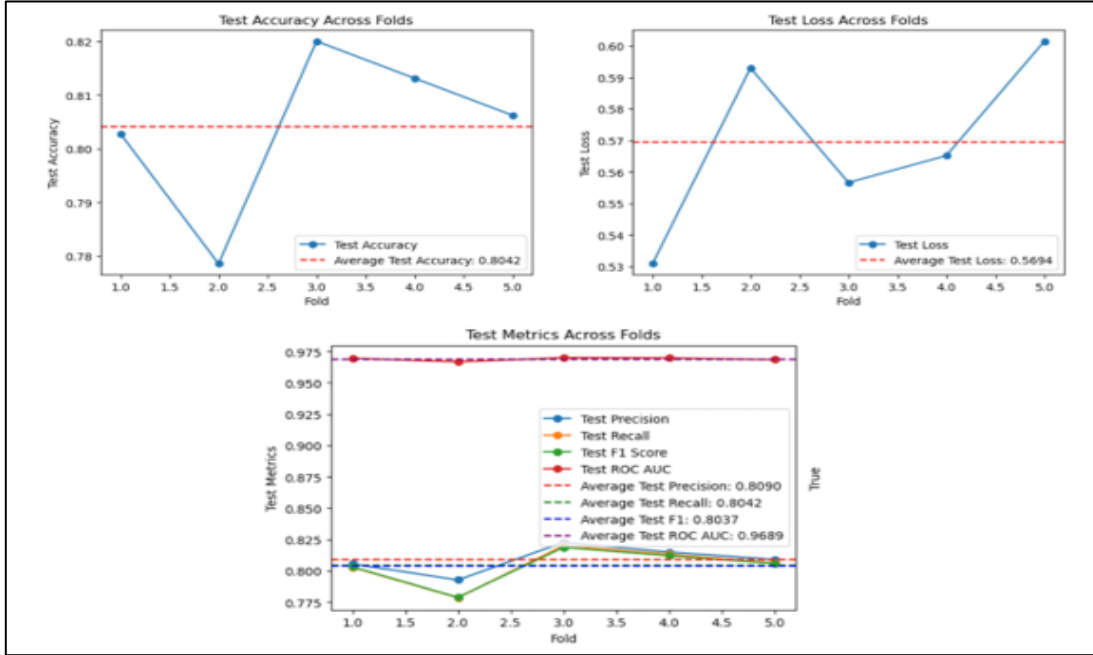


Şekil 5.12. MobileNetV2 modeli için Garbage Dataset veri kümesinde katlama bazlı test doğruluğu, kayıp ve performans metrikleri.

5.2.5. InceptionV3 Garbage Classification katlama bazlı test sonuçları

InceptionV3 modelinin Garbage Classification veri kümesi üzerindeki test sonuçlarına bakıldığında, modelin doğruluk oranı genel olarak tutarlı bir şekilde %89 civarlarında seyretmiştir. Ancak, katlama'lar arasında dalgalanmalar gözlemlenmiştir, bu da modelin bazı katlamalarda daha yüksek, bazılarında ise daha düşük doğruluklar sağladığını göstermektedir. Test kaybı, yaklaşık 0,2961 seviyelerinde olup, yine bazı katlamalarda artış gösterse de genellikle düşük kalmıştır. Değerlendirme metrikleri açısından modelin kesinlik, geri çağırma ve F1 skor değerleri sırasıyla 0,8979, 0,8967 ve 0,8962 seviyelerindedir. Bu yüksek değerler, modelin sınıflar arasındaki ayırma başarılı olduğunu ve sınıflandırma görevinde istikrarlı bir performans sergilediğini işaret etmektedir. ROC AUC skoru ise 0,9945 gibi mükemmel bir değere sahiptir, bu da modelin genel olarak çok iyi bir genelleme ve sınıflandırma başarısına sahip olduğunu göstermektedir. Sonuç olarak, InceptionV3 modelinin Garbage Classification veri kümesi üzerindeki performansı yüksek doğruluk ve düşük

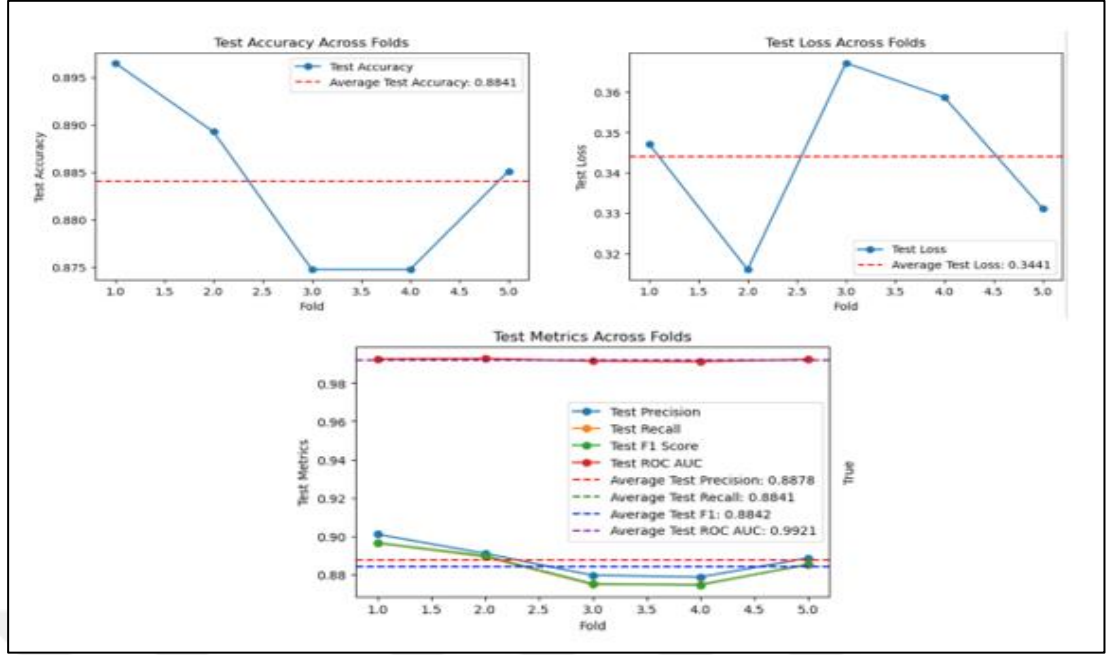
kayıplarla tutarlı olup, sınıflandırma metrikleri de modelin başarısını desteklemektedir. Ancak, katlama bazlı dalgalanmalar genelleme kapasitesinde sınırlı iyileştirmelere ihtiyaç duyulabileceğini gösteriyor (Şekil 5.13).



Şekil 5.13. InceptionV3 modeli için Garbage Classification veri kümesinde katlama bazlı test doğruluğu, kayıp ve performans metrikleri.

5.2.6. InceptionV3 Garbage Dataset veri kümesi katlama bazlı test sonuçları

InceptionV3 modelinin Garbage Dataset üzerinde gerçekleştirilen katlama bazlı test sonuçları incelendiğinde, modelin genel doğruluk oranı ortalama %88,41 olarak elde edilmiştir. Test kaybı ise ortalama 0,3441 düzeyindedir. Katlamalar arasında doğruluk ve kayıp değerlerinde dalgalanmalar gözlemlenmiş, ancak genel eğilim doğrulama başarısının istikrarlı bir şekilde sürdüğünü göstermektedir. Test doğruluğu, katlama 1'deki yüksek seviyesinden sonra 2. katlamada bir düşüş yaşasa da, 3. katlamadan itibaren toparlanmış ve son katlamalarla birlikte dengeye oturmuştur. Modelin değerlendirme metriklerinden kesinlik (%88,78), geri çağırma (%88,41), F1 skoru (%88,42) ve ROC AUC (%99,21) değerleri, modelin sınıflar arasındaki ayrım yapabilme yeteneğini ortaya koymaktadır. Özellikle ROC AUC değeri, modelin doğru sınıflandırma yapma kapasitesinin yüksek olduğunu ve genelleme yeteneğinin güçlü olduğunu göstermektedir. Genel olarak, InceptionV3 modelinin Garbage Dataset üzerinde yüksek doğruluk oranları elde ettiği ve test metriklerinin de tatmin edici seviyelerde kaldığı söylenebilir. Bu durum, modelin özellikle dengeli ve geniş veri setlerinde etkili bir performans sergileyebileceğini ortaya koymaktadır (Şekil 5.14).

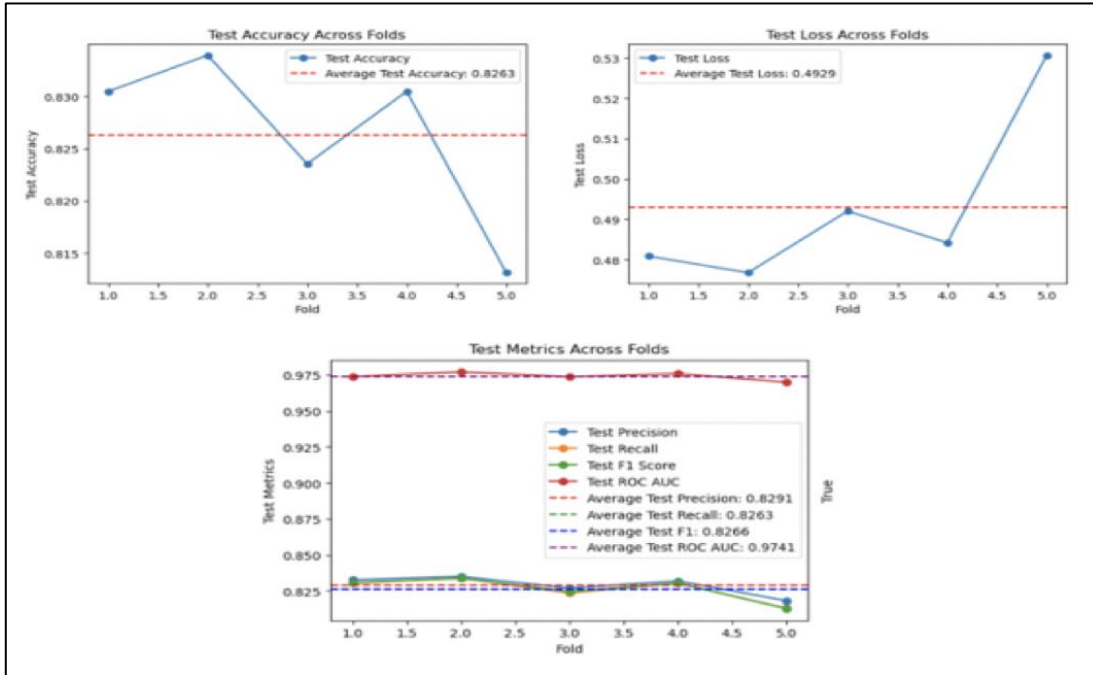


Şekil 5.14. InceptionV3 modeli için Garbage Dataset veri kümesinde katlama bazlı test doğruluğu, kayıp ve performans metrikleri.

5.2.7. DenseNet121 Garbage Classification veri kümesi katlama bazlı test sonuçları

DenseNet121 modelinin Garbage Classification veri kümesindeki performansı, beş katlı çapraz doğrulama (cross-validation) yöntemiyle değerlendirildi. Test doğruluğu ortalama %82,63 olarak hesaplandı ve bu değer, modelin veri kümesindeki genel başarısını yansıtmaktadır. Katlama 2'de görülen küçük bir doğruluk düşüşü, veri kümesindeki belirli örneklerin model tarafından zorlanması ve modelin her yeni katlamada sıfırdan öğrenmeye başlamasını gösteriyor. Test kaybı ortalama 0,49 seviyelerinde kalmış olup, modelin kayıp değeri, doğruluk oranıyla paralel olarak genellikle istikrarlı bir seyir izlemiştir. Bu durum, modelin genelleme kapasitesinin yüksek olduğunu ve veri kümesindeki fazla karmaşık olmayan özellikleri başarıyla öğrenebildiğini ortaya koymaktadır. Modelin test metrikleri de oldukça başarılıdır. Kesinlik, geri çağırma ve F1 skoru sırasıyla %82,91, %82,63 ve %82,66 olarak elde edilmiştir. Bu metrikler, modelin doğru sınıflandırma yapabilme yeteneğini ve veri kümesindeki sınıflar arasındaki dengeliyi iyi şekilde koruduğunu göstermektedir. Ayrıca, ROC AUC skoru %97,41 gibi yüksek bir değerle modelin sınıflar arasındaki ayırma yapma yeteneğinin güçlü olduğunu ortaya koymuştur. Sonuç olarak, DenseNet121 modeli, Garbage Classification veri kümesinde yüksek doğruluk, düşük kayıp ve iyi değerlendirme metrikleriyle başarılı bir performans sergileyerek, veri kümesine genelleme yapabilme yeteneğini kanıtlamıştır. Modelin, düşük parametrelili

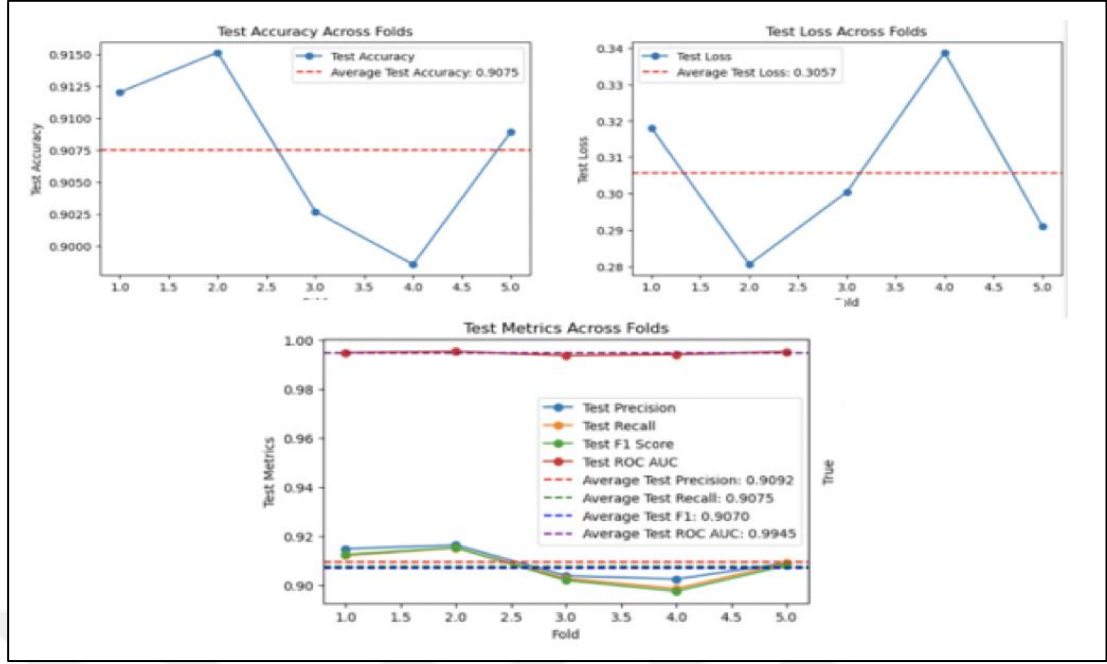
ve dengeli veri setlerinde etkili bir şekilde çalışabileceği söylenebilir. Bu, özellikle gerçek dünya uygulamaları için modelin güvenilirliğini artıran bir faktördür (Şekil 5.15).



Şekil 5.15. DenseNet121 modeli için Garbage Classification veri kümesinde katlama bazlı test doğruluğu, kayıp ve performans metrikleri.

5.2.8. DenseNet121 Garbage Dataset veri kümesi katlama bazlı test sonuçları

DenseNet121 modelinin Garbage Dataset üzerinde elde ettiği test sonuçları, modelin genel performansını net bir şekilde ortaya koymaktadır. Test doğruluğu, her katlamada belirli dalgalanmalar göstermiş olsa da genel olarak oldukça yüksek bir performans sergilenmiştir. Modelin ortalama doğruluk oranı %90,75 olarak belirlenmiştir, bu da modelin doğru sınıflandırma yapma yeteneğinin güçlü olduğunu göstermektedir. Test kaybı, özellikle katlama 3'te bir artış gösterse de, genel ortalama kayıp değeri 0,3057 olarak hesaplanmış ve modelin çoğu katlamada düşük kayıp ile çalıştığı gözlemlenmiştir. Bu, modelin doğru sınıflandırmalar yaparken fazla hata yapmadığını ortaya koymaktadır. Değerlendirme metrikleri incelendiğinde, modelin kesinlik değeri %90,92, geri çağırma değeri %90,75, F1 skoru %90,70 ve ROC AUC değeri %99,45 olarak hesaplanmıştır. Bu metrikler, modelin sınıflar arasındaki ayrımı başarılı bir şekilde yapabildiğini ve yüksek doğrulukla performans sergilediğini kanıtlamaktadır. Sonuç olarak, DenseNet121 modeli, Garbage Dataset üzerinde yüksek doğruluk ve tutarlılıkla başarılı bir sınıflandırma çözümü sunmuştur (Şekil 5.16).



Şekil 5.16. DenseNet121 modeli için Garbage Dataset veri kümesinde katlama bazlı test doğruluğu, kayıp ve performans metrikleri.

5.3. Eğitim ve Doğrulama Grafiklerinin Analizi

Bu çalışmada eğitim ve doğrulama başarı oranlarını gösteren grafikler detaylı şekilde incelenmiştir. Grafikler genel olarak, tüm modellerde eğitim sürecinde doğruluk oranlarının kademeli bir şekilde arttığını ve belirli bir seviyede dengelendiğini göstermektedir. Özellikle eğitim doğruluklarının, doğrulama doğruluklarına kıyasla daha yüksek değerlere ulaştığı gözlemlenmiştir. Bu durum, eğitim setinde modelin daha iyi bir uyum sağladığını, ancak doğrulama setinde sınırlı bir performans artışı olduğunu ortaya koymaktadır. Tüm modellerde, eğitim süreçleri boyunca aşırı öğrenme belirtileri sınırlı düzeyde kalmış ve doğrulama doğrulukları istikrarlı bir şekilde korunmuştur. Garbage-Dataset üzerinde elde edilen doğrulama başarılarının Garbage Classification veri kümesine kıyasla daha yüksek olması, veri kümesinin boyutu ve dengeli yapısının model performansını olumlu etkilediğini göstermektedir. Sonuç olarak, eğitim ve doğrulama süreçlerinin dikkatli yönetimi sayesinde, modellerin genelleme yetenekleri başarılı bir şekilde korunmuş ve güvenilir sınıflandırma performansları elde edilmiştir. Grafikler, kullanılan derin öğrenme mimarilerinin her iki veri kümesinde de istikrarlı bir performans sergilediğini desteklemektedir.

5.4. Modeller Arası Performans Karşılaştırması

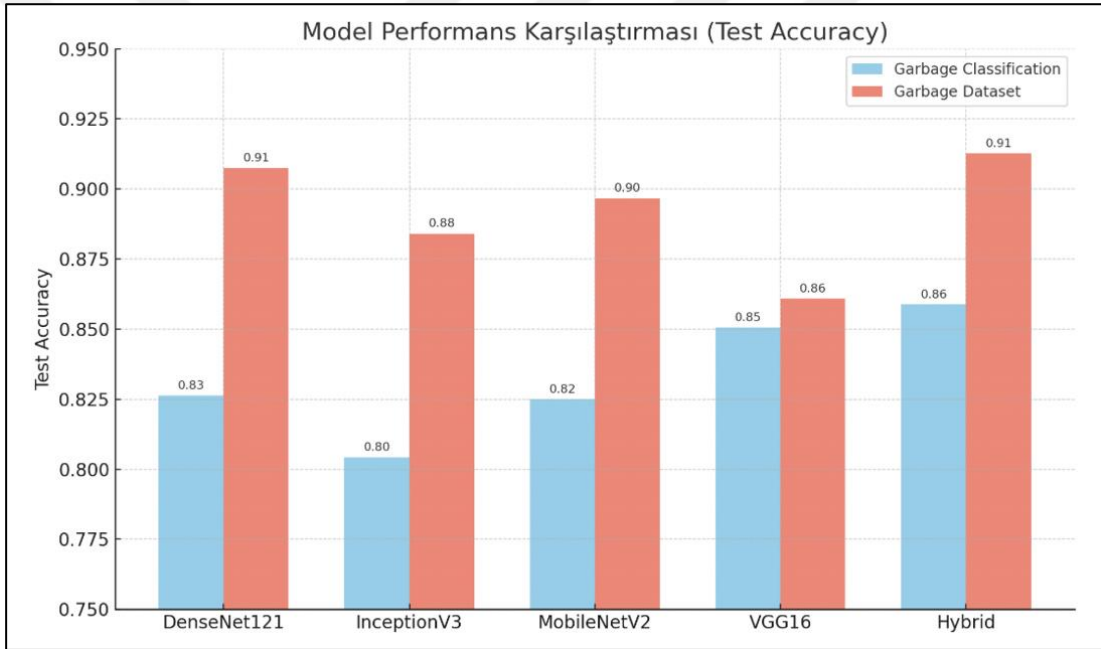
Bu çalışmada kullanılan dört farklı derin öğrenme modeli (VGG16, MobileNetV2, InceptionV3 ve DenseNet121), iki farklı veri kümesi üzerinde beş katlı çapraz doğrulama yöntemi ile değerlendirilmiş ve başarı oranları test doğrulukları üzerinden karşılaştırılmıştır. Modellerin başarı düzeyleri, ortalama test doğruluklarının yanı sıra kayıp değerleri ve değerlendirme metrikleri (kesinlik, geri çağırma, F1 skoru, ROC AUC) temel alınarak analiz edilmiştir. Tekil modellerde Garbage Classification veri kümesinde VGG16 modeli %85.05 doğruluk oranı ile performansı göstermiştir. Bu modeli sırasıyla DenseNet121 (%82.63), MobileNetV2 (%82.49) ve InceptionV3 (%80.42) takip etmiştir. Öte yandan, Garbage Dataset veri kümesinde en yüksek test doğruluğu %90.75 ile DenseNet121 modeli tarafından elde edilmiş, onu MobileNetV2 (%89.67), InceptionV3 (%88.41) ve VGG16 (%86.09) izlemiştir. Veri kümesinin büyüklüğü ve sınıf dengesi arttıkça, modellerin başarı oranlarında da genel bir artış gözlemlenmiştir.

Tekil modellerin performansına ek olarak, bu çalışmada geleceğe dönük bir geliştirme olarak dört CNN modelinin (VGG16, MobileNetV2, InceptionV3 ve DenseNet121) özellik haritaları birleştirilerek hibrit bir model tasarlanmıştır. Bu hibrit mimari ile elde edilen test doğruluğu Garbage Classification veri kümesinde %85.88, Garbage Dataset veri kümesinde ise %91.26 olarak hesaplanmıştır. Hibrit model, Garbage Classification veri kümesinde temel modellerin test doğruluklarının üstüne çıkarak en yüksek başarıyı elde etmiş; Garbage Dataset üzerinde ise DenseNet121 modelinin hemen önünde, en iyi performansı gösteren mimari olmuştur. Özellikle hibrit modelin daha yüksek hacimli ve dengeli veri kümesinde sağlamış olduğu %91.26 doğruluk oranı, farklı CNN mimarilerinin özellik çıkarım gücünün bir araya getirilmesinin, sınıflandırma doğruluğunu artırmada etkili olduğunu ortaya koymaktadır.

Tüm modellerin test doğrulukları birlikte değerlendirildiğinde, hibrit model her iki veri kümesinde de genel başarı açısından öne çıkmıştır. DenseNet121 modeli, temel CNN mimarileri arasında en yüksek başarıyı sağlamış, MobileNetV2 ise daha kompakt yapısına rağmen güçlü bir doğruluk oranı ve düşük kayıplarla dikkat çekmiştir. VGG16 modeli küçük veri setlerinde daha başarılı sonuçlar sunarken, InceptionV3 modeli özellikle Garbage Classification veri kümesinde görece daha düşük performans göstermiştir. ROC AUC skorları bakımından da hibrit model ve DenseNet121 modeli en başarılı sonuçları vermiştir. Bu analizler göstermektedir ki, farklı CNN

mimarilerinin güçlü yönlerinin bir araya getirildiği hibrit yaklaşımlar, yalnızca bireysel mimarilerin sağlayabildiği doğruluk oranlarının ötesine geçerek daha kararlı ve yüksek başarımlı sınıflandırma sistemleri oluşturulmasına olanak tanımaktadır.

Şekil 5.17’de, VGG16, DenseNet121, MobileNetV2, InceptionV3 ve hibrit modelin Garbage Classification ve Garbage Dataset veri setleri üzerindeki test doğruluk değerleri karşılaştırmalı olarak sunulmuştur. Grafik incelendiğinde, hibrit modelin her iki veri kümesinde de en yüksek doğruluğa ulaştığı ve bu yönüyle en başarılı sınıflandırma performansını sergilediği açıkça görülmektedir. Bu sonuçlar, özellik füzyon temelli hibrit modellerin, özellikle derin öğrenme tabanlı görüntü sınıflandırma görevlerinde model doğruluğunu artırmak için etkili bir strateji olduğunu göstermektedir.



Şekil 5.17. Modellerin test doğruluk karşılaştırması.

Şekil 5.18’de sunulan kesinlik, geri çağırma, koru skoru ve ROC AUC metriklerine dayanarak, dört farklı derin öğrenme modelinin yanı sıra hibrit modelin de Garbage Classification ve Garbage Dataset veri setleri üzerindeki performansları karşılaştırılmıştır. Bu kapsamlı değerlendirme, modellerin farklı veri setlerindeki başarılarını ve hibrit modelin sağladığı avantajları açıkça ortaya koymaktadır.

Kesinlik metriklerinde, Garbage Dataset veri kümesinde tüm modeller genel olarak daha yüksek değerler elde etmiştir. Bu durum, Garbage Dataset’in daha büyük ve dengeli yapısının model performansını olumlu yönde etkilediğini göstermektedir.

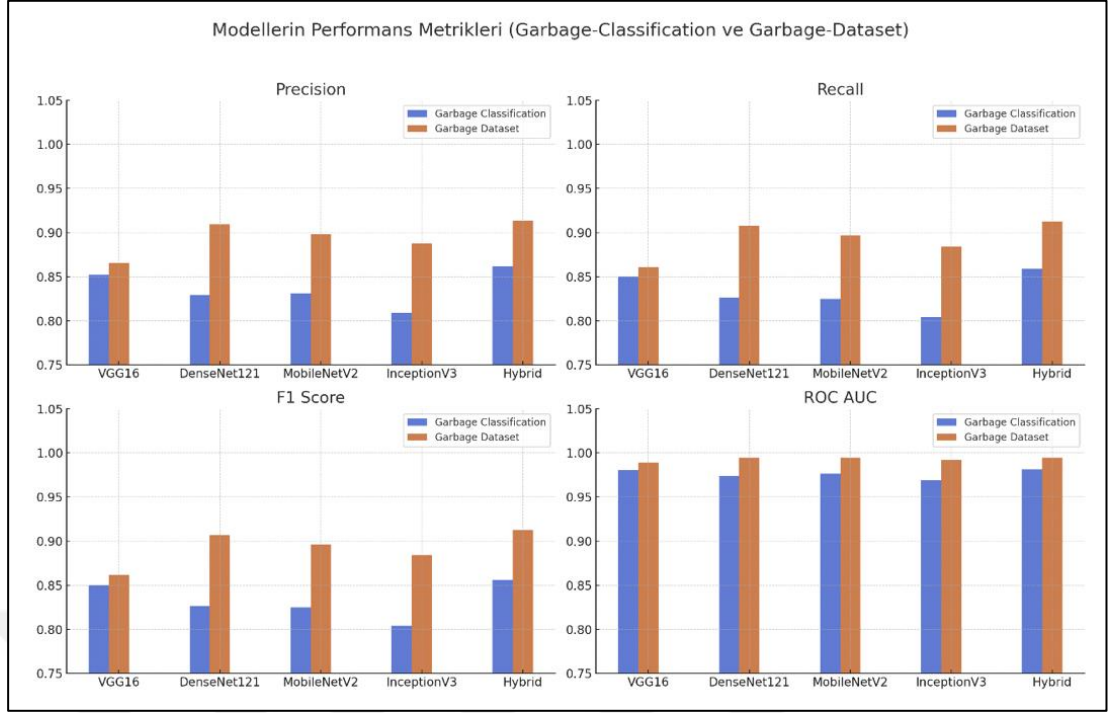
Özellikle DenseNet121 ve MobileNetV2 modelleri yüksek kesinlik skorlarıyla dikkat çekerken, hibrit model de her iki veri kümesinde üstün bir kesinlik performansı sergilemiş ve Garbage Dataset üzerinde 0.9135 gibi en yüksek kesinlik değerine ulaşmıştır. Bu sonuç, hibrit modelin farklı modellerin güçlü yönlerini birleştirerek doğru pozitif tahminlerin artırılmasında etkili olduğunu göstermektedir.

Geri çağırma metriklerinde MobileNetV2 ve DenseNet121 modelleri her iki veri kümesinde de güçlü sonuçlar sunmuştur. Bununla birlikte, hibrit model, Garbage Dataset üzerinde 0.9126 ile en yüksek geri çağırma değerini elde ederek, pozitif örnekleri kaçırma oranını daha da düşürmüştür. Bu, hibrit modelin kapsayıcılığının, farklı model yaklaşımlarının birleşimi sayesinde arttığını ve daha geniş örnek çeşitliliği üzerinde başarılı olduğunu göstermektedir.

F1 skorları açısından hibrit model, Garbage Classification ve Garbage Dataset üzerinde sırasıyla 0.8559 ve 0.9126 değerleri ile diğer tüm modelleri geride bırakmıştır. Bu da hibrit modelin kesinlik ve geri çağırma arasında dengeli ve güçlü bir performans sergilediğini ortaya koymaktadır. Bu dengeli başarı, hibrit modelin tekil modellerin sınırlamalarını minimize ederek genel doğruluğu artırdığını desteklemektedir.

ROC AUC metriklerinde ise hibrit model, Garbage Dataset üzerinde 0.9946 değeri ile en yüksek performansı göstermiştir. Bu durum, hibrit modelin sınıflar arasındaki ayrımı yapmada diğer modellere göre daha başarılı olduğunu ve karar sınırlarını daha etkin belirlediğini ifade etmektedir. Garbage Classification veri kümesinde de hibrit model yüksek bir ROC AUC değeri (0.9812) ile etkileyici sonuçlar ortaya koymuştur.

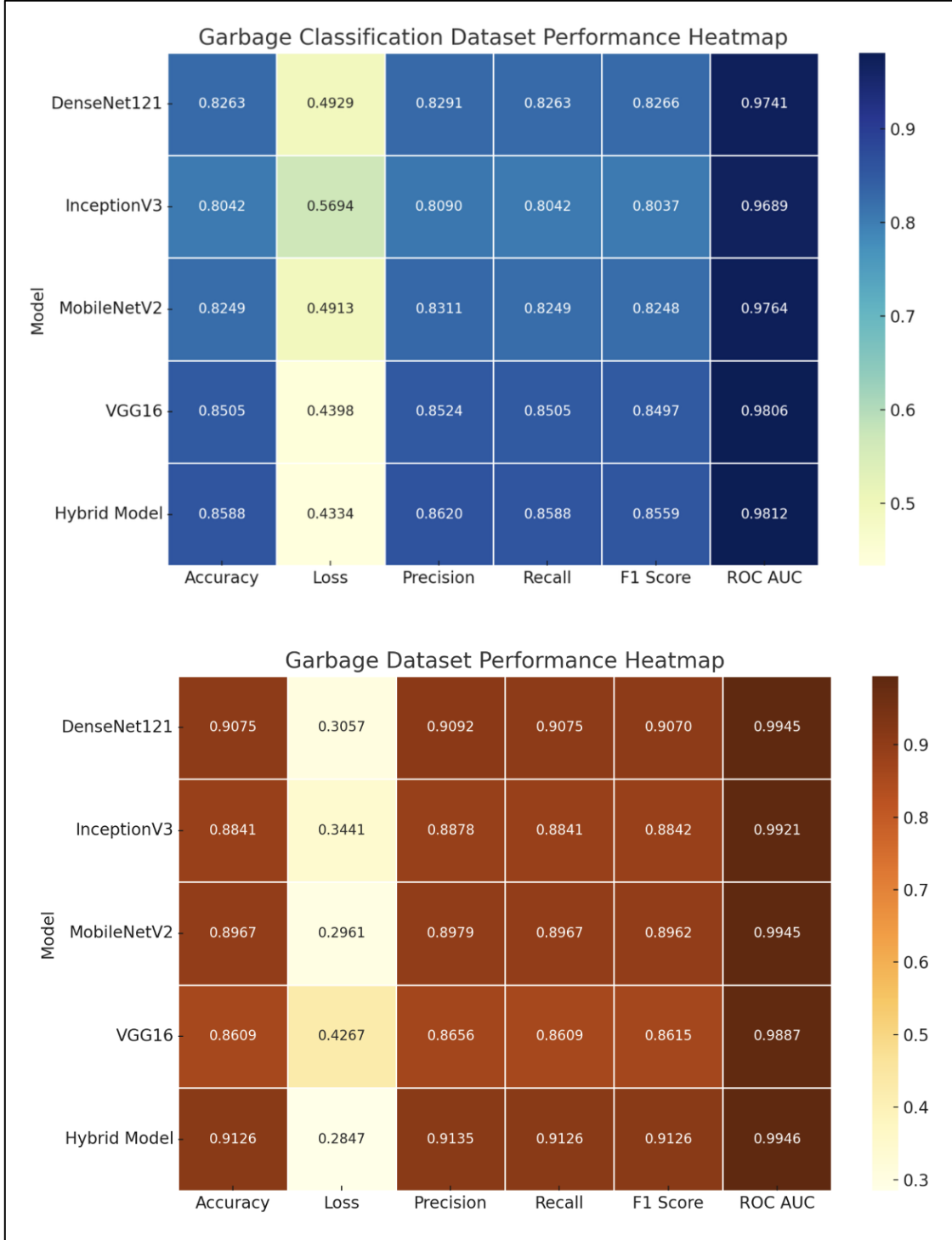
Sonuç olarak, hibrit model, farklı derin öğrenme modellerinin güçlü yanlarını bir araya getirerek, her iki veri kümesinde de hem tekil modellere kıyasla hem de veri kümesinin boyut ve dengesinden kaynaklanan zorluklara rağmen üstün performans göstermiştir. Bu başarısı, hibrit yaklaşımların özellikle heterojen veri setlerinde ve karmaşık sınıflandırma problemlerinde tercih edilmesinin önemini vurgulamaktadır. Ayrıca, hibrit modelin yüksek kesinlik, geri çağırma, F1 skoru ve ROC AUC değerleri, gerçek dünya uygulamalarında güvenilir ve dengeli bir sınıflandırma performansı sunma potansiyelini göstermektedir.



Şekil 5.18. Modellerin değerlendirme metriklerinin karşılaştırması.

5.5. Hibrit Modelin Performansa Etkisi

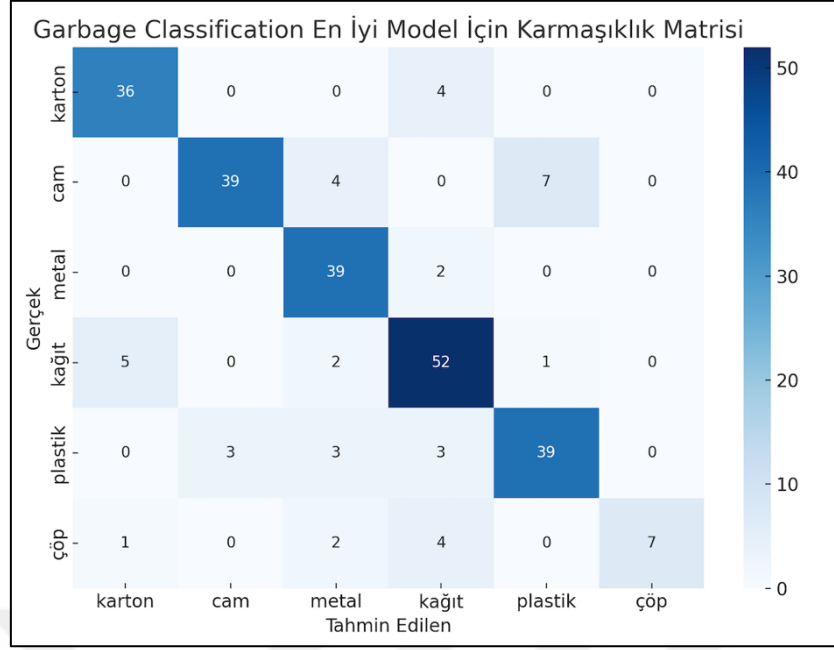
Kullanılan iki veri kümesi için de hibrit model, Doğruluk, Kesinlik, Geri Çağırma, F1 Skor ve ROC AUC gibi metriklerde en yüksek başarıyı göstermiştir. Özellikle Garbage Dataset veri kümesinde hibrit modelin başarısı, veri kümesi boyutunun artmasıyla birlikte daha da belirgin hale gelmiştir. Klasik modeller arasında ise DenseNet121 ve VGG16 genellikle daha yüksek performans sergilemiştir. Bu sonuçlar, hibrit modelin farklı mimarilerin güçlü yönlerini bir araya getirerek sınıflandırma başarımını artırabildiğini göstermektedir (Şekil 5.19).



Şekil 5.19. Veri setlerinin model performans metrikleri.

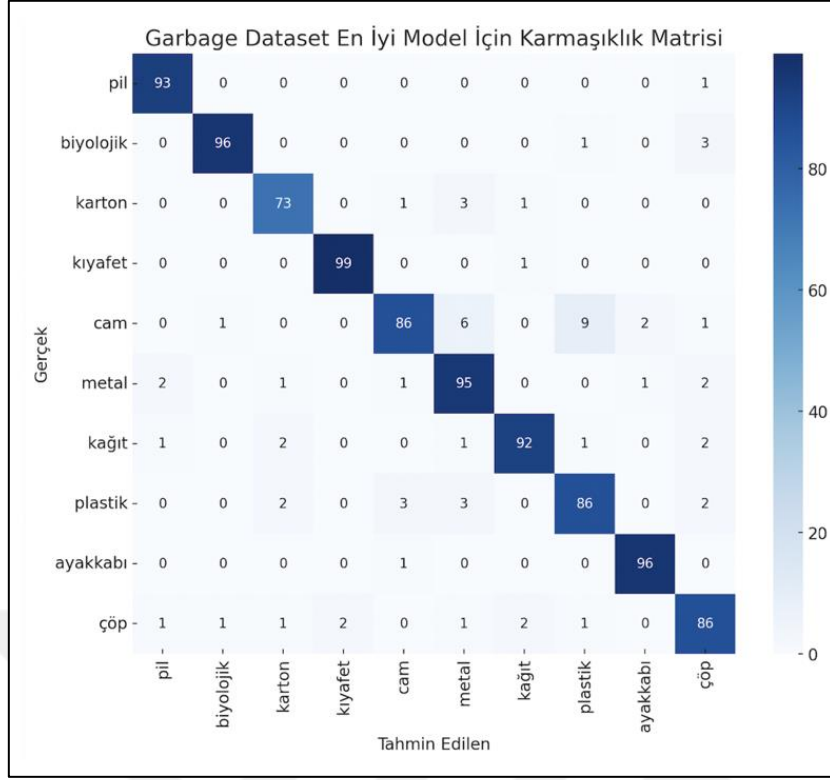
5.6. Karmaşıklık Matrisleri

Sınıflandırma performansını yalnızca genel metriklerle değil, aynı zamanda hangi sınıfların birbirine karıştırıldığını analiz ederek değerlendirmek amacıyla karmaşıklık matrisleri kullanılmıştır. Bu bağlamda, her iki veri kümesinde en başarılı sonucu veren modellerin karmaşıklık matrisleri gösterilmiştir.



Şekil 5.20. Garbage Classification en iyi model için karmaşıklık matrisi.

Karmaşıklık matrisi analizine göre, önerilen hibrit model, Garbage Classification veri kümesinde yüksek sınıflandırma başarısı göstermiştir. Model, özellikle "kağıt", "metal" ve "karton" sınıflarında oldukça başarılı sonuçlar üretmiş; bu sınıflarda doğru sınıflandırılan örnek sayısı sırasıyla 52, 39 ve 36 olarak kaydedilmiştir. Bu sonuçlar, modelin bu sınıflara ait görsel özellikleri etkin biçimde ayırt edebildiğini göstermektedir. Bununla birlikte, bazı sınıflarda kısmi karışıklıklar gözlemlenmiştir. Örneğin, "glass" sınıfına ait 7 örnek "plastik" sınıfına, 4 örnek ise "metal" sınıfına yanlış sınıflandırılmıştır. Benzer şekilde "plastik" sınıfındaki bazı örnekler de "metal", "kağıt" ve "glass" sınıflarıyla karıştırılmıştır. Bu durum, sınıflar arasındaki görsel benzerliğin modelin ayırt etme yetisini zaman zaman zorladığını göstermektedir. En fazla karışıklığın yaşandığı sınıf ise "çöp" olmuştur; bu sınıfa ait yalnızca 7 örnek doğru tahmin edilirken, geri kalan örnekler "karton", "metal" ve "kağıt" gibi sınıflarla karıştırılmıştır. Bu sonuçlar, modelin "çöp" sınıfındaki temsili öğrenmede sınırlı kaldığını ortaya koymaktadır. Genel olarak, hibrit modelin sınıflar arası ayrımı güçlü bir şekilde gerçekleştirdiği, ancak bazı sınıflarda veri çeşitliliği ve görsel benzerlikler nedeniyle karışıklıkların yaşandığı görülmektedir. Bu bağlamda, özellikle daha az temsil edilen veya görsel olarak benzer sınıflar için veri artırma ve ön işleme stratejilerinin uygulanması, modelin genel başarısını artırabilecek potansiyel geliştirme alanları olarak öne çıkmaktadır (Şekil 5.20).



Şekil 5.21. Garbage Dataset en iyi model için karmaşıklık matrisi.

Şekil 5.21’de verilen karmaşıklık matrisi, hibrit modelin Garbage Dataset üzerinde yaptığı tahminlerin doğru ve yanlış sınıflandırmalarını göstermektedir. Bu çalışmada, Garbage Dataset veri kümesi üzerinde eğitilen hibrit derin öğrenme modeli oldukça başarılı sınıflandırma performansı göstermiştir. Karmaşıklık matrisi incelendiğinde, kıyafet (99), biyolojik (96), ayakkabı (96), pil (93) ve metal (95) sınıflarında yüksek doğrulukla sınıflandırma yapıldığı görülmektedir. Bu sonuçlar, modelin görsel olarak belirgin sınıflar arasındaki farkları etkili biçimde öğrenebildiğini göstermektedir. Ayrıca, kağıt (92), plastik (86), çöp (86) ve karton (73) sınıflarında da yüksek başarı oranları elde edilmiştir.

Daha detaylı incelendiğinde, glass sınıfında bazı örneklerin plastik (9) ve metal (6) sınıfları ile karıştırıldığı gözlemlenmektedir. Bu durum, cam materyallerin parlaklık ve saydamlık gibi özelliklerinin diğer bazı materyallerle görsel benzerlik taşımasından kaynaklanıyor olabilir. Benzer şekilde, çöp ve karton sınıflarında da sınırlı sayıda karışıklık görülmüştür. Ancak bu karışıklıklar genellikle benzer doku ve renk özelliklerine sahip materyaller arasında gerçekleşmiştir. Öte yandan, plastik ve kağıt gibi temel olarak farklı kategoriler arasında dikkate değer bir karışıklık

gözlenmemiştir; bu da modelin anlamsal olarak ayrışan sınıfları başarılı şekilde ayırt edebildiğini göstermektedir.

Genel olarak, VGG16, InceptionV3, DenseNet121 ve MobileNetV2 modellerinin birleşiminden oluşan hibrit modelin, farklı geri dönüştürülebilir atık kategorileri arasında yüksek doğrulukla ayırım yapabildiği ve sınıflar arası sınırları başarılı bir şekilde öğrenebildiği görülmektedir. Elde edilen bu sonuçlar, modelin gerçek dünya atık yönetimi sistemlerinde etkin bir şekilde kullanılabileceğini göstermektedir.

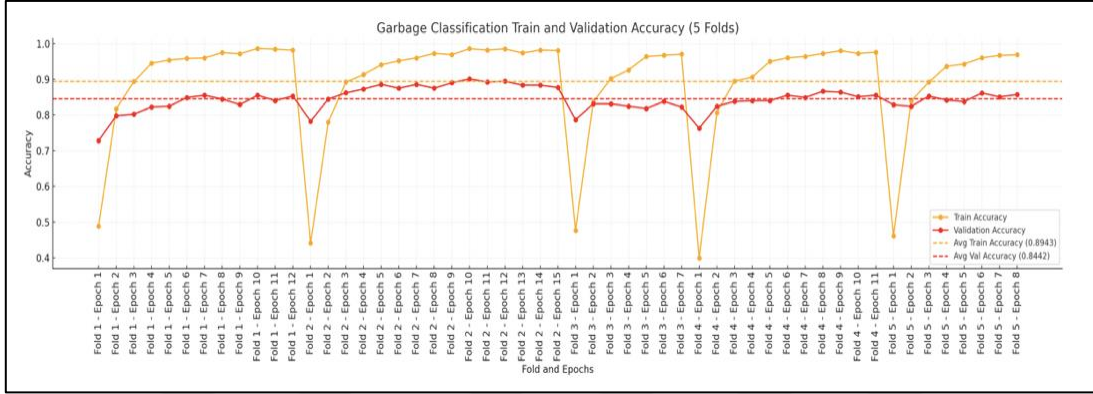
5.7. Garbage Classification Veri kümesi Hibrit Model Başarısı

Geri dönüştürülebilir atık sınıflandırması için kullanılan modelin 5 katlı K-Fold çapraz doğrulama sürecindeki eğitim ve doğrulama doğruluk değerlerini göstermektedir. Grafik, her bir fold ve epoch için ayrı ayrı hesaplanan doğruluk değerlerinin dağılımını sunmakta ve modelin eğitim sürecindeki öğrenme dinamiklerini ayrıntılı biçimde inceleme imkanı sağlamaktadır.

Eğitim doğruluğu (sarı çizgi), tüm foldlar boyunca genel olarak yükselen bir eğilim göstermekte ve modelin öğrenme kapasitesinin her epoch ile arttığını ortaya koymaktadır. Bununla birlikte, bazı epochlarda ve foldlarda ani düşüşler gözlemlenmiştir; bu durum muhtemelen eğitim verisinin foldlara ayrılmasından kaynaklanan veri farklılıkları ve modelin geçici olarak lokal minima veya yetersiz genel geçerlik durumuna geçişinden kaynaklanabilir. Ancak bu ani düşüşlere rağmen eğitim doğruluğu foldlar arasında istikrarlı bir şekilde %90 ve üzeri değerlere ulaşmaktadır.

Doğrulama doğruluğu (kırmızı çizgi) ise foldlar arası nispeten daha stabil seyretmekte, genellikle %80 ila %90 arasında değişmektedir. Bu durum modelin aşırı öğrenme yapmadığını, ancak doğrulama setinde performansın eğitim kadar yüksek olmadığını göstermektedir. Grafik üzerindeki kırmızı ve turuncu kesikli çizgiler, sırasıyla ortalama doğrulama ve eğitim doğruluk değerlerini temsil etmekte ve modelin genel başarımını özetlemektedir. Ortalama doğrulama doğruluğu yaklaşık %84,42 iken, ortalama eğitim doğruluğu %89,43 seviyelerindedir; bu fark, modelin henüz doğrulama setinde tam anlamıyla genelleme yapamadığını ve potansiyel olarak hiper parametre ayarlarının veya veri artırma stratejilerinin iyileştirilebileceğini işaret etmektedir (Şekil 5.22).

Sonuç olarak, grafik modelin öğrenme sürecinde genel olarak başarıyla ilerlediğini, ancak bazı foldlarda ve epochlarda performans dalgalanmaları olduğunu göstermektedir. Bu durum, modelin farklı veri alt kümelerine karşı duyarlılığını ortaya koyarken, eğitim ve doğrulama arasındaki doğruluk farkı ise genel geçerlik optimizasyonu için ek çalışmaların gerekliliğine işaret etmektedir.



Şekil 5.22. Garbage Classification veri kümesi için hibrit modelin eğitim ve doğrulama grafiği.

5.8. Garbage Dataset Veri kümesi Hibrit Model Başarısı

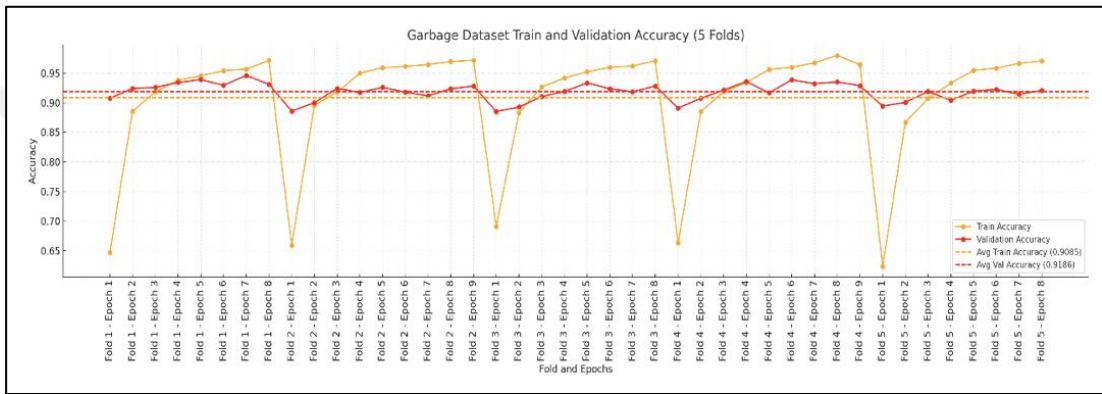
Garbage Dataset üzerinde uygulanan modelin 5 katlı K-Fold çapraz doğrulama yöntemiyle elde edilen eğitim ve doğrulama doğruluk değerlerini göstermektedir. Grafik, her bir fold ve epoch için ayrı ayrı hesaplanan doğruluk oranlarının dağılımını sunarak modelin eğitim sürecindeki öğrenme dinamiklerine ışık tutmaktadır.

Eğitim doğruluğu (turuncu çizgi) genel olarak yüksek ve istikrarlı bir performans sergilemekte olup, foldlar arasında %90 ila %97 aralığında seyretmektedir. Bu durum, modelin öğrenme kapasitesinin güçlü olduğunu ve eğitim verisindeki örüntüleri etkili biçimde yakaladığını göstermektedir. Bununla birlikte, bazı fold ve epochlarda eğitim doğruluğunda belirgin düşüşler gözlemlenmektedir; bu durum, foldlara özgü veri dağılımının model performansını geçici olarak olumsuz etkileyebileceğine işaret etmektedir.

Doğrulama doğruluğu (kırmızı çizgi) ise foldlar arasında nispeten daha stabil seyretmekte ve çoğunlukla %89 ila %93 arasında değişmektedir. Bu istikrar, modelin genelleme yeteneğinin eğitim performansına kıyasla daha tutarlı olduğunu göstermektedir. Grafik üzerindeki kesikli çizgiler, sırasıyla ortalama eğitim doğruluğu (%90.85) ve ortalama doğrulama doğruluğu (%91.86) değerlerini belirtmektedir. İlginç bir şekilde, ortalama doğrulama doğruluğunun ortalama eğitim doğruluğuna

oldukça yakın olması, modelin overfitting yapmadığını ve eğitimde öğrendiği örüntüleri doğrulama setinde de başarıyla genelleştirdiğini ortaya koymaktadır (Şekil 5.23).

Sonuç olarak, grafik genelinde modelin Garbage Dataset üzerinde başarılı bir öğrenme süreci gerçekleştirdiği, hem eğitim hem de doğrulama aşamalarında yüksek ve dengeli doğruluk değerlerine ulaştığı söylenebilir. Foldlar arasındaki küçük doğruluk dalgalanmaları veri yapısının karmaşıklığı ve örneklem farklılıklarından kaynaklanabilir; ancak genel performansın yüksek olması modelin güvenilir ve uygulanabilir olduğunu göstermektedir.



Şekil 5.23. Garbage Dataset veri kümesi için hibrit modelin eğitim ve doğrulama grafiği.

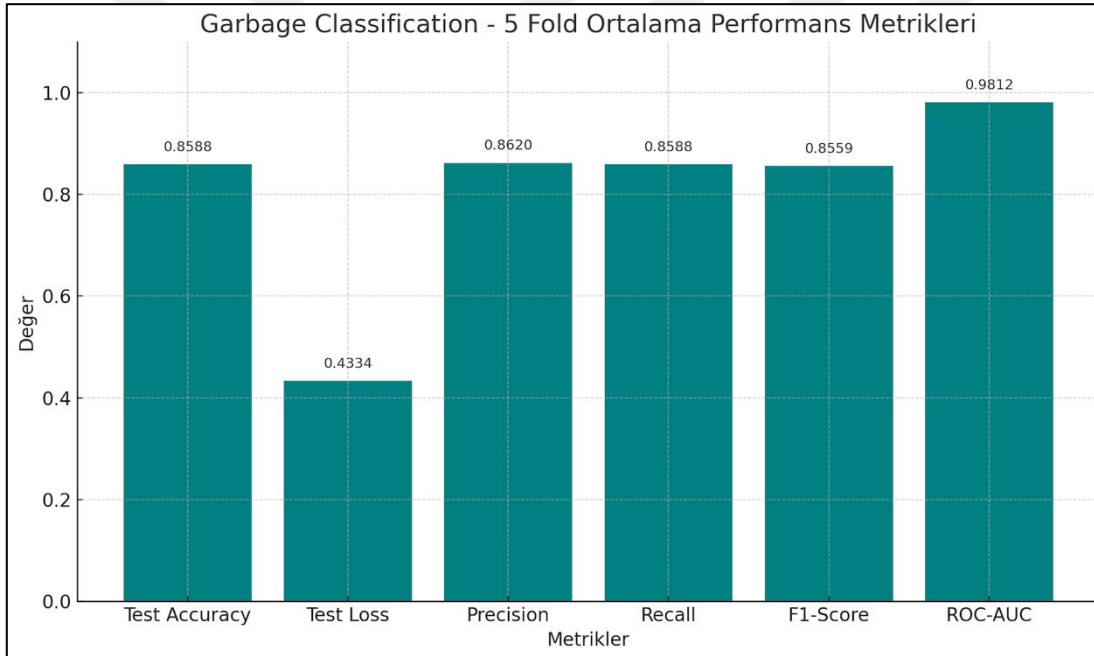
5.9. Garbage Classification Veri kümesi Hibrit Model Performans Metrikleri Sonucu

Garbage Classification veri kümesi üzerinde 5 katlı K-Fold çapraz doğrulama yöntemiyle elde edilen ortalama performans metriklerini göstermektedir. Grafik, modelin çeşitli performans kriterlerine göre genel başarısını karşılaştırmalı olarak ortaya koymaktadır.

Modelin ortalama test doğruluk değeri yaklaşık %85.88 olarak hesaplanmıştır. Bu, modelin genel sınıflandırma başarısının oldukça yüksek olduğunu ve veri kümesindeki farklı atık türlerini başarılı şekilde ayırt edebildiğini göstermektedir. Precision, Recall ve F1-Skor değerlerinin de sırasıyla %86.20, %85.88 ve %85.59 civarında seyretmesi, modelin sadece doğru sınıflandırma yapmakla kalmayıp aynı zamanda sınıflar arasında dengeli bir performans gösterdiğine işaret eder. Bu metriklerin birbirine yakın olması, modelin sınıf dengesizliklerinden kaynaklanan potansiyel yanlılıklarının sınırlı olduğunu destekler.

ROC-AUC metriği ise yaklaşık %98.12 olarak kaydedilmiştir. Bu yüksek değer, modelin sınıflar arasındaki ayrımı çok iyi yaptığı ve özellikle sınıflandırma kararlarında güvenilirlik sağladığı anlamına gelir. Bu da modelin pratik uygulamalarda başarılı bir şekilde kullanılabileceğinin önemli bir göstergesidir.

Grafikte gözlemlenen en düşük performans metriği test loss (%43.34) olarak görülmektedir. Loss değerinin diğer metriklerle kıyasla yüksek olması, modelin tahminlerindeki hata oranının görece daha fazla olduğunu, ancak bu hataların genel sınıflandırma doğruluğunu anlamlı biçimde düşürmediğini göstermektedir. Loss metriğinin daha düşük değerlerde olması tercih edilse de, doğruluk ve diğer metriklerdeki yüksek performans, modelin genel etkinliğinin iyi seviyede olduğunu ortaya koymaktadır (Şekil 5.24).



Şekil 5.24. Garbage Classification veri kümesi hibrit modelin performans metrikleri sonucu.

Sonuç olarak, grafik ve metrikler birlikte değerlendirildiğinde, kullanılan modelin Garbage Classification veri kümesinde dengeli, güvenilir ve yüksek performanslı bir sınıflandırma gerçekleştirdiği söylenebilir. Özellikle ROC-AUC değerinin yüksekliği, modelin gerçek dünyadaki uygulamalar için uygunluğunu güçlendirmektedir.

5.10. Garbage Dataset Veri kümesi Hibrit Model Performans Metrikleri Sonucu

Garbage Dataset veri kümesi üzerinde gerçekleştirilen 5 katlı K-Fold çapraz doğrulama yöntemi sonucu elde edilen ortalama performans metriklerini

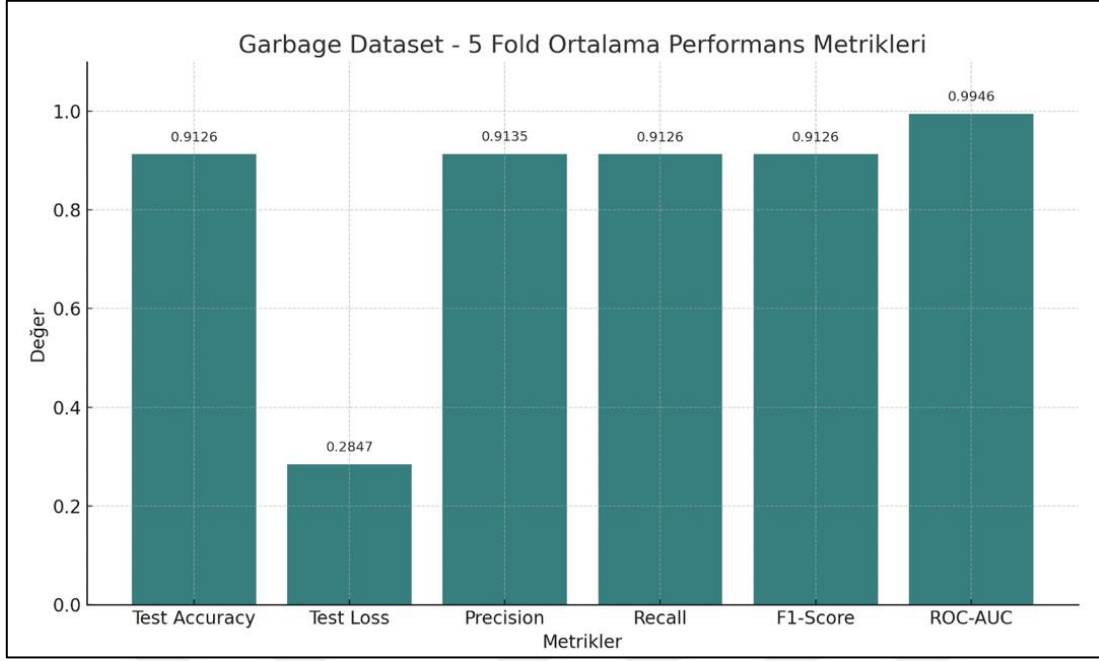
göstermektedir. Grafik, modelin çeşitli önemli ölçütlerdeki başarısını açıkça ortaya koymaktadır.

Modelin ortalama test doğruluğu %91.26 gibi oldukça yüksek bir değere ulaşmıştır. Bu yüksek doğruluk oranı, modelin atık sınıflandırma görevinde güçlü bir ayırt edici kapasiteye sahip olduğunu göstermektedir. Kesinlik (%91.35), Geri Çağırma (%91.26) ve F1-Skor (%91.26) değerleri birbirine oldukça yakın olup, modelin sınıflar arası dengenin sağlanmasında etkili olduğunu doğrulamaktadır. Bu metriklerin benzer seviyelerde seyretmesi, sınıflandırmanın hem doğru pozitif oranının hem de gerçek pozitiflerin başarılı şekilde yakalandığının göstergesidir.

ROC-AUC skoru %99.46 gibi mükemmel yakın bir değerde seyretmektedir. Bu, modelin farklı sınıflar arasında sınıflandırma kararlarını ayırt etmedeki başarısını ve genel güvenilirliğini işaret eder. Yüksek ROC-AUC skoru, sınıflandırma performansının hem doğruluk hem de hata oranı açısından üstün olduğunu ortaya koyar (Şekil 5.25).

Test loss değeri ise yaklaşık 0.28 seviyesinde bulunmuş, bu da modelin tahmin hatalarının düşük olduğunu ve öğrenme sürecinde modelin veriye iyi uyum sağladığını göstermektedir. Loss değerinin doğruluk ve diğer metriklerle göre daha düşük olması, modelin optimize edildiği kayıp fonksiyonunun iyi işlediğine işaret eder.

Genel olarak, bu grafik, Garbage Dataset üzerinde kullanılan hibrit modelin yüksek performanslı, dengeli ve güvenilir bir sınıflandırma sistemi olduğunu göstermektedir. Elde edilen performans metrikleri, modelin gerçek dünya atık sınıflandırma uygulamalarında etkin bir şekilde kullanılabileceğini güçlü biçimde desteklemektedir.



Şekil 5.25. Garbage Dataset veri kümesi hibrit modelin performans metrikleri sonucu.

6. SONUÇLAR

Bu tez çalışmasında, geri dönüştürülebilir atıkların sınıflandırılmasına yönelik olarak VGG16, MobileNetV2, InceptionV3 ve DenseNet121 gibi dört farklı derin öğrenme mimarisi, Garbage Classification ve Garbage Dataset veri setleri üzerinde karşılaştırmalı olarak değerlendirilmiştir. Model eğitim süreçlerinde beş katlı çapraz doğrulama yöntemi uygulanmış ve eğitim/doğrulama doğrulukları ile diğer başarı metrikleri detaylı biçimde analiz edilmiştir. Elde edilen sonuçlar, sınıflandırma performansının yalnızca mimari seçimlere değil, aynı zamanda veri kümesinin boyutu ve sınıf dengesi gibi faktörlere de doğrudan bağlı olduğunu ortaya koymuştur. Özellikle DenseNet121 mimarisi, her iki veri kümesinde de en yüksek doğruluk oranlarına ulaşarak güçlü bir temel model olduğunu kanıtlamıştır. MobileNetV2 ise daha düşük parametre sayısı ile yüksek doğruluk sağlayarak kaynak kısıtlı sistemlerde uygulanabilirliğini göstermiştir.

Bununla birlikte, bu tezde önerilen özellik füzyon tabanlı hibrit model, dört farklı CNN mimarisinden çıkarılan derin özellikleri birleştirerek daha zengin ve anlamlı bir temsil öğrenimi gerçekleştirmiştir. Modelin mimarisi, VGG16, InceptionV3, DenseNet121 ve MobileNetV2 ağlarının öznitelik çıkarıcı katmanlarından alınan çıktıları birleştirerek, ardından tam bağlantılı katmanlar üzerinden sınıflandırma işlemini gerçekleştirecek şekilde tasarlanmıştır. Yapılan deneylerde, hibrit modelin hem eğitim hem de test aşamalarında tekil modellerin ötesinde bir performans sergilediği gözlemlenmiştir. Hibrit model, kullanılan her iki veri kümesi üzerinde de Doğruluk, Precision, Recall, F1 Skor ve ROC AUC metriklerinde en yüksek başarıyı sağlamıştır.

Özellikle Garbage Dataset üzerinde hibrit model, test doğruluğunu %91,26 seviyesine çıkararak tüm modeller arasında en iyi sonucu elde etmiştir. Ayrıca, test kaybı %28,47 ile minimum seviyeye ulaşmış, ROC-AUC skoru ise %99,46'ya yükselmiştir. Benzer şekilde Garbage Classification veri kümesi üzerinde de hibrit model, %85,88 test doğruluğu, %43,34 test kaybı ve %98,12 ROC-AUC skoru ile dikkat çekici başarı sağlamıştır. Bu sonuçlar, hibrit modelin özellikle çok sınıflı, dengesiz ve görsel

çeşitliliği yüksek veri kümelerinde daha kararlı ve güvenilir tahminler ürettiğini göstermektedir.

Hibrit modelin başarısı, aynı zamanda overfitting riskinin azaltılmasında da etkili olmuş, modelin genelleme yeteneğini artırarak gerçek dünya senaryolarına uygulanabilirliğini güçlendirmiştir. Bu bağlamda, özellik füzyonu yaklaşımı, özellikle karmaşık atık görüntülerinin sınıflandırılmasında etkili bir çözüm olarak öne çıkmaktadır.

Gelecek çalışmalar için, hibrit mimarinin hesaplama yükü optimize edilerek daha hızlı ve verimli hale getirilmesi, farklı çevresel koşulları temsil eden daha geniş veri setleri üzerinde test edilmesi ve mobil cihazlara entegre edilebilecek hafif sürümlerinin geliştirilmesi önerilmektedir. Ayrıca, veri artırma tekniklerinin daha stratejik bir şekilde uygulanması ve sınıf dengesizliği problemlerine karşı gelişmiş çözüm yöntemlerinin entegrasyonu da performansı daha da ileri taşıyabilir.

Sonuç olarak, bu tez çalışması, hem klasik CNN mimarilerinin karşılaştırmalı analizini sunmuş hem de geliştirilen özellik füzyon temelli hibrit modelin üstün performansını bilimsel olarak ortaya koyarak, atık sınıflandırma problemlerine yönelik etkili, genellenebilir ve pratik bir yapay zeka çözümü sunmuştur.

KAYNAKLAR

- [1] Kaya, M., Ulutürk, S., Çetin Kaya, Y., Altıntaş, O., & Turan, B. (2023). Optimization of several deep CNN models for waste classification. *Sakarya University Journal of Computer and Information Sciences*, 6(2), 89–97. <https://doi.org/10.35377/saucis.12571001>
- [2] Sevinç, A., & Özyurt, F. (2022). Classification of recyclable waste using deep learning architectures. *Firat University Journal of Experimental and Computational Engineering*, 1(3), 122–128. <https://doi.org/10.5505/fujece.2022.83997>
- [3] Bircanoğlu, C., Atay, M., & Yalçın, A. (2021). RecycleNet: Intelligent waste sorting via transfer learning. *Procedia Computer Science*, 184, 781–788.
- [4] Wang, Z., Liu, Q., & Wang, Z. (2021). Comparative study of deep learning models for waste classification. *Journal of Cleaner Production*, 313, 127812.
- [5] Yang, M., & Thung, G. (2016). Classification of çöp for recyclability status. Stanford University CS229 Project.
- [6] Rabano, A., Martinez, L., & Santos, L. (2020). Automatic waste classification using MobileNet CNN architecture. *Sustainable Computing: Informatics and Systems*, 28, 100423.
- [7] Zhang, H., Chen, X., & Li, J. (2020). Smart waste classification using transfer learning and DenseNet. *Journal of Environmental Management*, 268, 110403.
- [8] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.
- [9] Pan, S. J., & Yang, Q. (2009). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359.
- [10] Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *arXiv preprint, arXiv:1409.1556*.
- [11] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4700–4708).
- [12] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the Inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2818–2826).
- [13] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint, arXiv:1704.04861*.

- [14] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778).
- [15] Kaggle. (2023). Garbage Dataset (10 classes). Kaggle Datasets. Retrieved from <https://www.kaggle.com/datasets/sumn2u/garbage-classification-v2>
- [16] Kaggle. (2023). Garbage Classification (6 classes). Kaggle Datasets. Retrieved from <https://www.kaggle.com/datasets/asdasdasdasdas/garbage-classification>
- [17] Hu, W., et al. (2024). Improved CNN-Based Solid Waste Classification Method. *Sustainable Computing: Informatics and Systems*, 52, 100675. <https://www.sciencedirect.com/science/article/pii/S2949750724000920>
- [18] Ali, S., et al. (2024). AI-Powered Waste Classification Using Deep Learning. *International Journal of Advanced Computer Science and Applications*, 15(10). https://thesai.org/Downloads/Volume15No10/Kağit_9AI_Powered_Waste_Classification.pdf
- [19] Chen, X., et al. (2024). CNN for Inorganic Packaging Classification in Waste Processing Plants. *Environmental Advances*, 17, 100076. <https://www.sciencedirect.com/science/article/pii/S2773186324000768>
- [20] Singh, A., et al. (2025). Deep Learning-Based Automated Waste Classification System. *Knowledge-Based Systems*, 298, 110032. <https://www.sciencedirect.com/science/article/pii/S0950705125000760>
- [21] Zhang, Y., et al. (2024). Federated Deep Learning for Waste Classification in Smart Cities. *Scientific Reports*, 14, Article 62003. <https://www.nature.com/articles/s41598-024-62003-4>
- [22] Rahman, M. A., et al. (2024). Smart Waste Classification Approach Using Improved Multi-Layer CNN. *Multimedia Tools and Applications*. <https://link.springer.com/article/10.1007/s11042-024-18939-w>
- [23] Kumar, R., et al. (2024). Comparative Analysis of Deep Learning Models in Waste Classification. *arXiv preprint arXiv:2411.02779*. <https://arxiv.org/abs/2411.02779>
- [24] Li, J., et al. (2025). Improved EfficientNetV2-Based Framework for Waste Classification. *arXiv preprint arXiv:2503.21208*. <https://arxiv.org/abs/2503.21208>
- [25] Patel, H., et al. (2024). Deep Learning-Based Plastik Waste Classification with WaDaBa Dataset. *arXiv preprint arXiv:2412.20232*. <https://arxiv.org/abs/2412.20232>
- [26] Zhao, L., et al. (2024). Pixel Distribution Learning for Image-Based Waste Classification. *arXiv preprint arXiv:2409.03913*. <https://arxiv.org/abs/2409.03913>
- [27] Lilhore, U. K., Khodve, V. S., Parui, S. K., & Singh, D. (2025). SwinConvNeXt: A lightweight hybrid deep learning architecture for real-time waste classification. *Scientific Reports*, 15, Article 91302. <https://www.nature.com/articles/s41598-025-91302-7>

- [28] Mustapha, R., Sani, N. F. B. M., & Ramli, R. B. (2025). Real-time smart compost and recyclable waste classifier using hybrid YOLOv8 and CNN. *Engineering Proceedings*, 84(1), 82. <https://www.mdpi.com/2673-4591/84/1/82>
- [29] Li, B., Zhang, X., & Lin, J. (2025). An improved ResNet-50 based model for waste classification with weighted feature fusion and depthwise separable convolutions. *PLOS ONE*, 20(5), e0317999. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0317999>
- [30] Sayem, M. A., Akter, F., Kabir, M. A., & Hasan, M. M. (2024). A dual-stream hybrid deep learning model for solid waste classification using DenseViT and GELAN-E. *Neural Computing and Applications*. <https://link.springer.com/article/10.1007/s00521-024-10855-2>
- [31] Ahmed, R., Khan, A. I., & Shahid, A. R. (2023). RWC-Net: A hybrid deep learning model for recyclable waste classification using DenseNet201 and MobileNetV2. *International Journal of Research Publication and Reviews*, 6(4), 182–190. <https://ijrpr.com/uploads/V6ISSUE4/IJRPR43852.pdf>
- [32] Qiu, C., Zhang, Y., & Zhao, Y. (2025). Improved EfficientNetV2-based framework with attention and multiscale spatial feature extraction for waste image classification. *arXiv preprint*, arXiv:2503.21208. <https://arxiv.org/abs/2503.21208>



ÖZGEÇMİŞ

Ad-Soyad : Gamze AĞIRTAŞ

ÖĞRENİM DURUMU:

- **Lisans** : 2022, Sakarya Üniversitesi, Bilgisayar ve Bilişim Bilimleri Fakültesi, Bilgisayar Mühendisliği

TEZDEN TÜRETİLEN ESERLER:

- Ağirtas, G., & Kaman, S. C. (2025, Nisan), ‘Classification of Recyclable Wastes with Deep Learning Models: A Comparison on the Effect of Dataset Size’, ISARC 2025 – International Symposium on Automation and Robotics in Construction. (https://www.isarconference.org/_files/ugd/6dc816_8072354712ab444fa121fcc901dc56cb.pdf).

DİĞER ESERLER:

- Güney, E., Yaşar, A., Ağırtaş, G., & Bayılmış, C. (2023). *Mobil Platformda IoT Temelli ve Soket Programlamaya Dayalı Termal Sensör Uygulaması*. Düzce Üniversitesi Bilim ve Teknoloji Dergisi, 11, 457–465. <https://doi.org/10.29130/dubited.1064120>
- Güney, E., Ağırtaş, G., & Bayılmış, C. (2022). *MongoDB Based Real-Time Monitoring Heart Rate Using WebSocket For Remote Healthcare*. Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI), 8(4), 500–508. <https://doi.org/10.26555/jiteki.v8i4.26567>