



**DERİN ÖĞRENME İLE UZAKTAN
ALGILAMA GÖRÜNTÜLERİNDE SAHNE
SINIFLANDIRMASI**

Yüksek Lisans Tezi

Damla DALGIÇ

Eskişehir 2025

**DERİN ÖĞRENME İLE UZAKTAN ALGILAMA GÖRÜNTÜLERİNDE
SAHNE SINIFLANDIRMASI**

Damla DALGIÇ

Yüksek Lisans Tezi

Bilgisayar Mühendisliği Anabilim Dalı

Bilgisayar Yazılımı Bilim Dalı

Danışman: Dr. Öğr. Üyesi Sevcan YILMAZ GÜNDÜZ

Eskişehir

Eskişehir Teknik Üniversitesi

Lisansüstü Eğitim Enstitüsü

Ocak 2025

JÜRİ VE ENSTİTÜ ONAYI

Damla DALGIÇ'ın DERİN ÖĞRENME İLE UZAKTAN ALGILAMA GÖRÜNTÜLERİNDE SAHNE SINIFLANDIRMASI başlıklı çalışması 15/01/2025 tarihinde aşağıdaki jüri tarafından değerlendirilerek "Eskişehir Teknik Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliği"nin ilgili maddeleri uyarınca, Bilgisayar Mühendisliği Anabilim dalında Yüksek Lisans Tezi olarak kabul edilmiştir.

Jüri Üyeleri

Unvan Adı Soyadı

İmza

Üye

: Dr. Öğr. Üyesi Sevcan YILMAZ
GÜNDÜZ

Üye

: Dr. Öğr. Üyesi Selcan KAPLAN
BERKAYA

Üye

: Dr. Öğr. Üyesi Sinem BOZKURT
KESER

Prof. Dr. Semra KURAMA

Lisansüstü Eğitim Enstitüsü Müdürü

15/01/2025

DANIŐMAN ONAYI

DaniŐmanlıđını yurttuđuđum Yůksek Lisans ۆđrencisi Damla DALGIÇ, DERİN RENME İLE UZAKTAN ALGILAMA GÖRÜNTÜLERİNDE SAHNE SINIFLANDIRMASI baŐlıklı tez alıŐmasını tamamlamıŐtır. HazırlamıŐ olduđu tez tarafımca incelenmiŐ ve ۆđrencinin tez savunma sınavına alınması bilimsel ve etik aıdan uygun görölmüŐtür.

Tez DaniŐmanı

Dr. đr. Üyesi Sevcan YILMAZ GÜNDÜZ

ÖZET

DERİN ÖĞRENME İLE UZAKTAN ALGILAMA GÖRÜNTÜLERİNDE SAHNE SINIFLANDIRMASI

Damla DALGIÇ

Bilgisayar Mühendisliği Anabilim Dalı

Bilgisayar Yazılımı Bilim Dalı

Eskişehir Teknik Üniversitesi, Lisansüstü Eğitim Enstitüsü, Ocak 2025

Danışman: Dr. Öğr. Üyesi Sevcan YILMAZ GÜNDÜZ

Derin öğrenme, birçok alanda önemli bir rol oynadığı gibi uzaktan algılama görüntü sahnesi sınıflandırmasında da oldukça popülerdir. Uzaktan algılama görüntüleri, hava ve uydu platformlarından elde edilen verilerle Dünya yüzeyinin çeşitli kategorilerde otomatik olarak sınıflandırılmasını sağlar. Son yıllarda, uzaktan algılama görüntülerinden sahne sınıflandırması için çeşitli veri kümeleri geliştirilmiş ve çeşitli yaklaşımlar sunulmuştur. Ancak, mevcut veri kümeleri sınıf çeşitliliği ve ölçek açısından sınırlamalara sahiptir. Bu durum, yeni derin öğrenme tabanlı yaklaşımların geliştirilmesini zorlaştırmaktadır. Bu çalışmada, sahne sınıflandırması için önceden eğitilmiş bir EfficientNet Versiyon 2 Küçük (EfficientNetV2S) derin öğrenme ağı kullanılmıştır. Modelin son katmanları bir transfer öğrenme yöntemiyle değiştirilerek EffiSceneNet adı verilen yeni bir derin öğrenme modeli önerilmiş ve literatürdeki en iyi sonuçları sunmuştur. Bunun yanı sıra orijinal EfficientNetV2S ve EfficientNet Versiyon 2 Orta (EfficientNetV2M) modelleri de çalıştırılmış ve EffiSceneNet modelinin üstün performansı gözlemlenmiştir. Ayrıca önerilen yöntem ile orijinal 2 model topluluk öğrenme yöntemiyle birleştirilmiştir. Bu yöntem, modellerin bireysel performanslarını bir araya getirerek genel doğruluk oranını artırmış ve daha güçlü bir sınıflandırma yeteneği sağlamıştır. Son olarak, bu çalışmada Görsel Dönüştürücü (ViT) modeli de sahne sınıflandırma amacıyla kullanılmış ve üstün performans sergilemiştir. Çalışma, büyük ve çeşitli veri kümeleri üzerinde gerçekleştirilmiş olup, literatüre önemli katkılar sağlamıştır.

Anahtar Sözcükler: Derin öğrenme, Görüntü sahnesi sınıflandırması, Uzaktan algılama, Evrişimli sinir ağları, EfficientNetV2S.

ABSTRACT

SCENE CLASSIFICATION IN REMOTE SENSING IMAGES WITH DEEP LEARNING

Damla DALGIÇ

Department of Computer Engineering

Programme in Computer Software

Eskişehir Technical University, Institute of Graduate Programs, January 2025

Supervisor: Asst. Prof. Dr. Sevcan YILMAZ GÜNDÜZ

Deep learning is popular in remote sensing image scene classification as well as playing an important role in many fields. Remote sensing images provide automatic classification of the Earth's surface into various categories with data obtained from air and satellite platforms. In recent years, various approaches have been presented for scene classification. However, existing datasets have limitations in terms of class diversity and scale, which makes it difficult to develop new deep learning-based approaches. In this study, a pre-trained EfficientNet Version 2 Small (EfficientNetV2S) deep learning network was used for scene classification. By replacing the last layers of the model with a transfer learning method, a new deep learning model called EffiSceneNet was proposed and presented the best results in the literature. In addition to proposed method, the original EfficientNetV2S and EfficientNet Version 2 Medium (EfficientNetV2M) models were also run and the superior performance of the EffiSceneNet model was observed. Three models were combined with ensemble learning method with the proposed method. This method increased the overall accuracy rate by combining the individual performances of the models. Finally, in this study, Vision Transformer (ViT) model was also used for scene classification and showed performance. The study was carried out on large and diverse datasets and made significant contributions to the literature.

Keywords: Deep learning, Remote sensing classification, Remote sensing, Convolutional neural networks, EfficientNetV2S.

TEŐEKKÜR

Tez alıŐması s¼recimde ilgi ve alakası baŐta olmak üzere deęerli fikirleri ile beni doęru y¼nlendiren ve desteklerini esirgemeyen tez danıŐmanım Dr. Öğr. Üyesi Sevcan YILMAZ G¼ND¼Z'e içtenlikle teŐekk¼r ederim. T¼m eęitim hayatım boyunca beni her zaman destekleyen ve benimle her zaman gurur duyan biricik aileme minnettarım. Son olarak s¼rete manevi olarak yanımda olan b¼t¼n sevdiklerime teŐekk¼rlerimi sunarım.

Damla DALGI



ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ

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

Damla DALGIÇ

İÇİNDEKİLER

	<u>Sayfa</u>
BAŞLIK SAYFASI	I
JÜRİ VE ENSTİTÜ ONAYI.....	II
DANIŞMAN ONAYI	III
ÖZET	IV
ABSTRACT.....	V
TEŞEKKÜR	VI
ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ.....	VII
İÇİNDEKİLER	VIII
TABLolar DİZİNİ	X
ŞEKİLLER DİZİNİ	XI
SİMGELER VE KISALTMALAR DİZİNİ	XII
1. GİRİŞ	1
1.1. Tezin Literatüre Katkısı	3
1.2. Tezin Organizasyonu	4
2. LİTERATÜR TARAMASI	5
3. MATERYAL VE METOT.....	12
3.1. Materyal	12
3.1.1. Veri kümesi	12
3.1.1.1. <i>NWPU-RESISC45</i> veri kümesi	13
3.1.1.2. <i>UCM</i> veri kümesi	15
3.1.1.3. <i>AID</i> veri kümesi	17
3.2. Metot.....	18
3.2.1. Evrişimsel sinir ağları	18
3.2.2. Transfer öğrenimi.....	20
3.2.3. EfficientNetV2S derin öğrenme modeli.....	21
3.2.4. EfficientNetV2M derin öğrenme modeli	22

3.2.5. EffiSceneNet modeli	23
3.2.6. Topluluk öğrenmesi.....	25
3.2.6.1. Çoğunluk oylaması	26
3.2.6.2. Ağırlıklı oylama	28
3.2.6.3. Topluluk öğrenmesinin uygulanması	29
3.2.7. Görsel dönüştürücü	30
4. DENEY SONUÇLARI	33
4.1. Veri Hazırlama ve Arttırma.....	33
4.2. Eğitim Detayları	35
4.3. NWPU-RESISC45 Veri Kümesi	35
4.4. UCM Veri Kümesi.....	40
4.5. AID Veri Kümesi.....	45
4.6. Test ve Eğitim Süreleri	49
5. SONUÇ	51
KAYNAKÇA.....	53

TABLULAR DİZİNİ

Sayfa

Tablo 3.1. Uzaktan algılama görüntü veri kümeleri	12
Tablo 4.1. Orijinal ve arttırılmış veri kümelerindeki görüntü sayıları.....	34
Tablo 4.2. NWPU-RESISC45 veri kümesinin sonuçları	36
Tablo 4.3. EffiSceneNet'in NWPU-RESISC45 üzerinde son teknoloji performansla karşılaştırılması.....	40
Tablo 4.4. UCM veri kümesinin sonuçları.....	41
Tablo 4.5. EffiSceneNet'in UCM veri kümesi üzerinde son teknoloji performansla karşılaştırılması.....	44
Tablo 4.6. AID veri kümesinin sonuçları.....	45
Tablo 4.7. EffiSceneNet'in AID veri kümesi üzerinde son teknoloji performansla karşılaştırılması	48
Tablo 4.8. Modellerin kaydedilen eğitim ve test süreleri.....	50

ŞEKİLLER DİZİNİ

Sayfa

Şekil 3.1. NWPU-RESISC45 veri kümesinin sınıfları ve sayıları	14
Şekil 3.2. NWPU-RESISC 45 veri kümesine ait görüntü örnekleri (Han, Cheng, & Lu, 2017).....	15
Şekil 3.3. UCM veri kümesinin sınıfları ve sayıları	16
Şekil 3.4. UCM veri kümesine ait görüntü örnekleri (Yang & Newsam, 2010)	16
Şekil 3.5. AID veri kümesinin sınıfları ve sayıları	17
Şekil 3.6. AID veri kümesine ait görüntü örnekleri (Xia, ve diğerleri, 2017).....	18
Şekil 3.7. CNN mimarisi	19
Şekil 3.8. EfficientNetV2S'in mimarisi	22
Şekil 3.9. Önerilen yöntemin iş akışı.....	23
Şekil 3.10. Orijinal EfficientNetV2S ile önerilen EffiSceneNet modellerinin farkı	24
Şekil 3.11. Sert oylama.....	26
Şekil 3.12. Yumuşak oylama.....	27
Şekil 3.13. Ağırlıklı oylama	29
Şekil 3.14. Topluluk modeli	30
Şekil 3.15. ViT'in mimarisi.....	32
Şekil 4.1. Artırılmış görüntü örnekleri. (a) Orijinal, (b) Yansıma, (c) Döndürme, (d) Yakınlaştırma	34
Şekil 4.2. NWPU-RESISC45'nin eğitim ve doğrulama değerleri. (a) Kayıp grafiği, (b) Doğruluk grafiği	37
Şekil 4.3. NWPU-RESISC45 veri kümesi sınıflarının F-1 skorları	38
Şekil 4.4. UCM'nin eğitim ve doğrulama değerleri. (a) Kayıp grafiği, (b) Doğruluk grafiği.....	42
Şekil 4.5. UCM veri kümesi sınıflarının F-1 skorları.....	43
Şekil 4.6. AID'in eğitim ve doğrulama değerleri. (a) Kayıp grafiği, (b) Doğruluk grafiği.....	46
Şekil 4.7. AID veri kümesi sınıflarının F-1 skorları.....	47

SİMGELER VE KISALTMALAR DİZİNİ

AID	: Aerial Image Dataset (Hava Görüntüsü Veri Kümesi)
CapsNet	: Capsule Network (Kapsül Ağları)
CNN	: Convolutional Neural Network (Evrşimsel Sinir Ağları)
ConvNeXt	: Convolutional Network Evolution (Evrşimsel Ağ Evrimi)
CPU	: Central Processing Unit (Merkezî İşlem Birimi)
CycleGAN	: Cycle-Consistent Generative Adversarial Networks (Döngü Tutarlı Üretken Çekişmeli Ağlar)
DenseNet121	: Densely Connected Convolutional Network with 121 Layers (Yoğun Bağlantılı 121 katmanlı Evrşimsel Ağ)
DWConv	: Depthwise Separable Convolution (Derinlemesine Ayrılabilir Konvolüsyon)
EfficientNetV2S	: EfficientNet Version 2 Small (EfficientNet Versiyon 2 Küçük)
EfficientNetV2M	: EfficientNet Version 2 Medium (EfficientNet Versiyon 2 Orta)
Fused MBConv	: Fused Mobile Inverted Bottleneck Convolution (Birleştirilmiş Mobil Ters Darboğaz Konvolüsyonu)
GPU	: Graphics Processing Unit (Grafik İşlem Birimi)
H-GCN	: Hierarchical Graph Convolutional Network (Hiyerarşik Grafik Evrşimsel Ağ)
InceptionV3	: Inception Version 3 (Başlangıç Versiyon 3)
KAN	: Kolmogorov-Arnold Network (Kolmogorov-Arnold Ağı)
KSA	: Kingdom of Saudi Arabia Remote Sensing Dataset (Suudi Arabistan Krallığı Uzaktan Algılama Veri Kümesi)
LRC	: Logistic Regression Classifier (Lojistik Regresyon Sınıflandırıcı)
MBConv	: Mobile Inverted Bottleneck Convolution (Mobil Ters Darboğaz Konvolüsyonu)
MobileNet	: Mobile Network (Mobil Ağ)
MobileNetV2	: Mobile Network Version 2 (Mobil Ağ Versiyon 2)
MS2AP	: Multi-scale Staking Attention Pooling (Çok Ölçekli Hisse Senedi Dikkat Havuzu)
NAS	: Neural Architecture Search (Nöral Mimari Arama)
NWPU-RESISC45	: Northwestern Polytechnical University Remote Sensing Image Scene Classification Dataset (Northwestern Polytechnical Üniversitesi Uzaktan Algılama Görüntü Sahne Sınıflandırma Veri Kümesi)
OPTIMAL-31	: Optimal Landcover Classification Dataset (Optimal Arazi Örtüsü Sınıflandırma Veri Kümesi)
ReLU	: Rectified Linear Unit (Doğrultulmuş Lineer Ünite)

ResNet	: Residual Network (Artık Ağ)
ResNet101	: Residual Network 101-Layer (Artık Ağ 101 Katmanı)
ResNet34	: Residual Network 34-Layer (Artık Ağ 34 Katmanı)
ResNet50	: Residual Network 50-Layer (Artık Ağ 50 Katmanı)
RSSCN7	: Remote Sensing Scene Classification Dataset with 7 Classes (Uzaktan Algılama 7 Sınıflı Sahne Sınıflandırma Veri Kümesi)
SIRI-WHU	: Sentinel-2 Imagery for Remote Sensing and Image Analysis at Wuhan University (Wuhan Üniversitesi'nde Uzaktan Algılama ve Görüntü Analizi için Sentinel-2 Görüntüleri)
SVM	: Support Vector Machine (Destek Vektör Makinesi)
UCM	: UC Merced Land Use Dataset (UC Merced Arazi Kullanım Veri Kümesi)
USGS	: United States Geological Survey (Amerika Birleşik Devletleri Jeolojik Araştırması)
ViT	: Vision Transformer (Görsel Dönüştürücü)
VGG	: Visual Geometry Group (Görsel Geometri Grubu)
VGG-16	: Visual Geometry Group 16-Layer (Görsel Geometri Grubu 16 Katmanlı)
VGG-19	: Visual Geometry Group 19-Layer (Görsel Geometri Grubu 19 Katmanlı)
WHU-RS19	: Wuhan University Remote Sensing 19 (Wuhan Üniversitesi Uzaktan Algılama 9)
Xception	: Extreme Inception (Aşırı Başlangıç)
YOLOv5	: You Only Look Once Version 5 (Sadece Bir Kez Bakarsın Versiyon 5)

1. GİRİŞ

Derin öğrenme teknikleri, bilgisayarlı görü alanında önemli bir paradigma değişimine yol açarak birçok endüstri alanında hızla benimsenmiştir. Bu teknikler, büyük veri kümelerini daha etkin bir şekilde işleme ve daha karmaşık modeller oluşturma yeteneğine sahip olmaları nedeniyle, geleneksel yöntemlere kıyasla önemli performans artışı sağlayabilmektedir. Evrimsel sinir ağları (CNN) ve diğer derin öğrenme tekniklerinin tanıtılması, görüntü sahnesi sınıflandırmasının doğruluğunda büyük gelişmelere katkıda bulunmuş, karmaşık sahneler üzerinde daha doğru ve otomatik sınıflandırma sistemleri geliştirmeye olanak tanımıştır. CNN tabanlı yöntemler, büyük ölçekli görüntü veri kümelerini kullanarak tahminlerde bulunmak amacıyla tasarlanmıştır. Bu süreçte, giderek daha karmaşık ve derin mimariye sahip modeller önerilmiştir. Örneğin, Görsel Geometri Grubu (VGG) (Simonyan & Zisserman, 2015) GoogLeNet (Szegedy, ve diğerleri, 2015) Residual Network (He, Zhang, Ren, & Sun, 2016) ve EfficientNet (Tan & Le, 2019) gibi modeller, bu amaç doğrultusunda tasarlanmış ve görüntü sınıflandırma performansını artırmak amacıyla geliştirilmiştir.

Uzaktan algılama görüntülerini kullanarak derin öğrenme ile sahne sınıflandırması, yer gözlemi, çevresel izleme, afet yönetimi ve tarımsal üretim gibi alanlarda geniş uygulama potansiyeline sahip olduğu için son yıllarda ilgi gören bir konu olmuştur. Bu yöntem, hava ve uydu platformlarından toplanan görüntülerden otomatik olarak bilgi çıkarmak için güçlü bir araç sunmaktadır. Özellikle coğrafi bilginin otomatik olarak çıkarılması ve Dünya yüzeyinin temsili açısından kritik bir öneme sahiptir. Bu işlem, görüntüdeki piksellerin mekansal ve spektral özniteliklerini analiz ederek, farklı arazi alanlarını otomatik olarak sınıflandırmayı hedefler. Bu sayede yerleşim alanları, tarım arazileri, ormanlık alanlar, su kütleleri ve kayalık alanlar gibi kategoriler tanımlanabilir. Sahne sınıflandırması, uzaktan algılama görüntülerinin etkin bir şekilde işlenmesi sayesinde sadece arazi kullanımının belirlenmesiyle sınırlı kalmayıp, şehir planlama, çevresel değişimlerin izlenmesi ve afet yönetimi gibi kritik uygulamalarda da önemli bir rol oynamaktadır. Doğru ve güvenilir bir sahne sınıflandırma sistemi, bu alanlarda karar alma süreçlerini iyileştirerek sürdürülebilir ve verimli çözümler geliştirilmesine katkı sağlamaktadır. Geleneksel yöntemler belirli özellikleri elle tanımlama zorunluluğuyla sınırlıyken, derin öğrenme tabanlı yaklaşımlar, görüntüdeki mekansal ve spektral ilişkileri daha kapsamlı bir şekilde öğrenerek yüksek doğruluklu tahminlerde bulunabilmektedir. Bu bağlamda, sahne sınıflandırmasının başarısı, coğrafi bilgi

sistemleri, savunma, tarım ve doğal afet yönetimi gibi birçok alanda etkili çözümler sunarak büyük bir katma değer sağlamaktadır.

Uzaktan algılama görüntü sahnesi sınıflandırma araştırmalarında sıklıkla kullanılan popüler veri kümeleri arasında UC Merced Arazi Kullanım Veri Kümesi (UCM) (Yang & Newsam, 2010), Hava Görüntüsü Veri Kümesi (AID) (Xia, ve diğerleri, 2017) ve Northwestern Polytechnical Üniversitesi Uzaktan Algılama Görüntü Sahne Sınıflandırma Veri Kümesi (NWPU-RESISC45) (Han, Cheng, & Lu, 2017) yer almaktadır. Özellikle UCM ve AID veri kümelerinde elde edilen sınıflandırma sonuçları, yüksek performans göstermiştir (Cheng, Xie, Han, Guo, & Xia, 2020). Bu veri kümeleri, çeşitli sahne türlerini ve yüksek çözünürlüklü görüntüleri içermeleri açısından avantaj sağlamaktadır. Bununla birlikte, NWPU-RESISC45 veri kümesi, sahne sınıflarının sayısı, sınıf başına düşen görüntü sayısı ve toplam görüntü sayısı açısından çok daha büyük ölçekli bir yapıya sahiptir. Bu durum, mevcut sınıflandırma modellerinin öznelik ayırt etme yeteneğini zorlayabilir. Mevcut yöntemlerin çoğu, genellikle UCM ve AID gibi daha küçük veri kümeleri üzerinde test edilmiştir. Ayrıca NWPU-RESISC45 veri kümesinde uygulanan modeller, UCM ve AID veri kümelerinden alınan sonuçlar kadar başarılı olamamıştır. Bundan dolayı daha büyük ve daha çeşitli veri kümeleri üzerinde test edilmek üzere yeni sınıflandırma modellerinin geliştirilmesine ihtiyaç duyulmaktadır.

Bu tez çalışmasında, uzaktan algılama görüntülerinde sahne sınıflandırması problemini çözmek amacıyla derin öğrenme tabanlı yaklaşımlar geliştirilmiştir. Çalışmanın temel hedefi, mevcut yöntemlerin sınırlamalarını aşarak daha yüksek doğruluk oranlarına ulaşan ve genelleştirilebilir bir model önermektir. Bu kapsamda, önceden eğitilmiş EfficientNetV2S modeli temel alınarak transfer öğrenme yöntemiyle yeni bir model olan EffiSceneNet geliştirilmiştir. EffiSceneNet modeli, modelin son katmanlarının optimize edilmesiyle sınıflandırma performansını artırmakta ve farklı veri kümeleri üzerinde üstün sonuçlar sunmaktadır. Ayrıca, EfficientNetV2S ve EfficientNetV2M modelleri ile EffiSceneNet modeli topluluk öğrenme yöntemleriyle birleştirilmiş, bu yöntem sayesinde modellerin bireysel başarıları optimize edilerek genel doğruluk oranı daha da artırılmıştır. Çalışmada ayrıca, ViT modeli sahne sınıflandırma amacıyla uygulanmış ve derin öğrenme tabanlı modellerle karşılaştırılmıştır. Önerilen yöntemlerin performansları UCM, AID ve NWPU-RESISC45 gibi geniş ve çeşitli veri kümeleri üzerinde test edilmiş; sonuçlar literatürdeki güncel yaklaşımlarla kıyaslanarak elde edilen gelişmeler ortaya konmuştur. Bu tez, uzaktan algılama görüntülerinin sahne

sınıflandırması alanında derin öğrenme tabanlı yaklaşımların etkinliğini artırmaya yönelik önemli katkılar sunmaktadır.

1.1. Tezin Literatüre Katkısı

Bu tez çalışmasının literatüre katkıları aşağıda sıralanmıştır.;

- Uzaktan algılama görüntü sahnesi sınıflandırmasında, büyük ve çeşitli veri kümeleri üzerinde test edilmesi gereken yeni derin öğrenme yöntemlerinin geliştirilmesine odaklanılmıştır. Literatürdeki çoğu çalışmada küçük-orta ölçekli veri kümelerinde çalışmalar bulunurken; uzaktan algılama görüntü veri kümelerinden en büyük ölçekli olanlardan NWPU-RESISC45'te literatürdeki en iyi sonuçlar elde edilmiştir.
- EfficientNetV2S CNN modeli özelleştirilerek temel EfficientNetV2S ve EfficientNetV2M CNN modellerine kıyasla belirgin bir performans artışı sağlanmıştır.
- Daha az parametreyle daha verimli işlemler gerçekleştiren yeni bir derin öğrenme modeli olan EffiSceneNet önerilmiştir.
- EffiSceneNet modeli, karmaşık sahneleri daha iyi sınıflandırma yeteneği göstermiştir. NWPU-RESISC45, AID ve UCM veri kümelerinde elde edilen sınıflandırma doğrulukları, literatürle karşılaştırıldığında en iyi sonuçlar sunmuştur.
- Önerilen model, EfficientNetV2S ve EfficientNetV2M modelleriyle topluluk öğrenme yöntemine entegre edilmiştir ve önerilen modelin başarısı sayesinde geliştirilen topluluk modellerinde de rekabet edici sonuçlar elde edilmiştir.
- ViT mimarisi, uzaktan algılama görüntü sahnesi sınıflandırması için uyarlanmış ve transformer tabanlı modellerin dikkat mekanizmalarının başarısı vurgulanmıştır. ViT'in küresel bağlamı etkili bir şekilde yakalaması sayesinde, büyük veri kümelerinde CNN tabanlı yaklaşımlarla rekabet edebilecek sonuçlar elde edilmiştir.
- Uzaktan algılama görüntülerinde farklı derin öğrenme uygulamalarının birbirleri ile karşılaştırılması sunulmuştur.

1.2. Tezin Organizasyonu

Tezin geri kalanı Őu Őekilde organize edilmiŐtir. Blm 2’de uzaktan algılama sahne sınıflandırması alanındaki alıŐmaları detaylı bir Őekilde inceleyen literatr taramasına yer verilmiŐtir. alıŐmada kullanılan veri kmelerinin detaylı bir aıklaması ve kullanılan yntemlerin tanıtılması ve nerilen yntemin detaylı aıklaması Blm 3’te yapılmıŐtır. Daha sonra Blm 4’te nerilen sınıflandırma tekniğinden elde edilen deney sonuları detaylı bir Őekilde sunulmuŐtur. Son olarak Blm 5’te bu alıŐmayla elde edilen sonular anlatılmıŐtır.



2. LİTERATÜR TARAMASI

CNN birçok alanda büyük başarılar elde etmiştir. Görüntü tanıma ve sınıflandırma, nesne tespiti ve segmentasyon gibi bilgisayarla görme görevlerinde yaygın olarak kullanılmaktadır (YOUNESI, ve diğerleri, 2024). CNN'ler hastalık tespiti ve tedavi planlaması gibi kritik görevler örnek olmak üzere tıbbi görüntü analizi alanından; otonom sürüş teknolojilerinde çevresel algılama ve yol işaretlerinin tanınmasına kadar geniş bir kullanım yelpazesine sahiptir. Doğal dil işleme alanında, metin sınıflandırma ve duygu analizi gibi görevlerde CNN'ler etkili sonuçlar elde etmiştir. Örneğin Turki vd.'nin yaptığı çalışmada CNN'ler kullanılarak tıbbi görüntülerde hastalık tespiti üzerine önemli bulgular sunulmuştur (Turki, Al-Sharif, & Taguchi, 2021). Benzer şekilde, Mahaur vd.'nin otonom araçlar için gerçek zamanlı nesne tespiti uygulamalarında CNN algoritmalarından olan Sadece Bir Kez Bakarsın Versiyon 5 (YOLOv5)'in başarısını göstermişlerdir (Mahaur & Mishra, 2023). Ayrıca, Jang vd. ise doğal dil işleme alanında CNN'lerin metin sınıflandırma performansı değerlendirmişlerdir (Jang, Kim, Harerimana, Kang, & Kim, 2020).

CNN'lerin görüntülerden hiyerarşik öznitelikleri öğrenme ve çıkarma yetenekleri, çok sayıda farklı uygulamanın önünü açmaya devam etmektedir. Bu tür umut verici alanlardan biri, Dünya yüzeyi hakkında ayrıntılı verilerin uydu veya hava görüntüleri aracılığıyla yakalanan uzaktan algılama görüntüleridir. Uzaktan algılama görüntüleri, Dünya yüzeyi hakkında kapsamlı bilgi sağlar ve çeşitli özelliklerinin karmaşık ayrıntılarını yakalar. CNN'lerin bilgisayarlı görüşteki başarısından yararlanan uzaktan algılama uzmanları, bu ağları uzaktan algılama görüntülerinden hızlı ve doğru öznitelik çıkarma için uygulamaya başlamıştır. CNN'leri uzaktan algılama verileriyle entegre ederek, araştırmacılar coğrafi bilgilerin zamanında ve hassas bir şekilde analiz edilmesinde önemli ilerlemeler elde etmişlerdir. Bu yaklaşım, CNN'lerin arazi örtüsü türlerini belirleme ve sınıflandırma, değişiklikleri algılama ve çevresel olayları izleme konusundaki etkinliğini gösteren çok sayıda çalışma tarafından desteklenmektedir.

Datla vd.'i sahne sınıflandırması için tüm sahnelerden gelen gereksiz öznitelikleri içeren bir özellik vektörü oluşturmuşlardır. Bu vektördeki gereksiz öznitelikleri azaltmak için sahne vektörleri önermişlerdir. Sahne özellikleri evrimsel özellikleri kullanarak Gauss karışım modeli eğitmişlerdir. Önerilen yöntem mevcut en iyi yöntemleri geride bırakmıştır (Datla, Perveen, & C., 2024).

Bi vd.'i temelinde özellik haritası çıkararak sahne sınıflandırmasında özellik temsil yeteneğini arttıran bir yöntem önermişlerdir. Bir dikkat modülü ile hem kanal hem de mekansal alanda yukarıdan aşağıya ve aşağıdan yukarıya özellikleri birleştirilerek farklı ölçeklerde haritalar çıkarmışlardır. Bu özellik haritaları, dikkat modülünden alınan ağırlıklarla bir havuzlama işlemine tutulmuştur (Bi, Zhang, & Qin, 2021).

Petrovska vd.'nin yaptığı çalışmada önceden eğitilmiş Artık Ağ 50 Katmanı (ResNet50), Başlangıç Versiyon 3 (InceptionV3), Aşırı Başlangıç (Xception) ve Yoğun Bağlantılı 121 katmanlı Evrişimsel Ağ (DenseNet121) CNN modellerinden yararlanan uzaktan algılama görüntü sınıflandırma mimarisi kurulmuştur. Sinir ağlarının aktivasyonlarından sahneler için en iyi özellik çıkarımını elde etmeye çalışılmıştır. Modeller tamamen ImageNet veri kümesi üzerinde eğitilmiştir. Özellik çıkarımı ve özellik birleştirmenin ardından sınıflandırma için Destek Vektör Makinesi (SVM) ve Lojistik Regresyon Sınıflandırıcı (LRC) önermişlerdir. Elde edilen sınıflandırma performansı en son teknoloji yaklaşımlarla eşdeğerdir. Son olarak Görüntü sahnesi sınıflandırmasında kullanılan görüntü öznelikleri açısından önceden eğitilmiş CNN modelleri InceptionV3, ResNet50, Xception ve DenseNet121 modelleri karşılaştırılmıştır. UCM ve Wuhan Üniversitesi Uzaktan Algılama 9 (WHU-RS19) veri kümelerinde test ederek %98,33 ile %97,76 doğruluk oranı elde edilmiştir (Petrovska, ve diğerleri, 2020).

Zhang vd.'nin yaptığı çalışmada CNN ve kapsül ağları (CapsNet) mimarilerini birlikte kullanarak iki bölümden oluşan CNN-CapsNet adında yeni bir mimari önerilmiştir. İlk bölümde önceden eğitilmiş bir model olan Görsel Geometri Grubu 16 katmanlı (VGG-16) ImageNet, veri kümesi üzerinde eğitilmiş ve ardından sahneleri etiketlemek için CapsNet mimarisi kullanılmıştır. AID (%96,85), UCM (%99,05) ve NWPU-RESISC45 (%92,6) veri kümelerinden elde ettikleri deneysel sonuçlar diğer yöntemlerle karşılaştırıldığında daha rekabetçi olduklarını göstermektedir (Zhang, Tang, & Zhao, 2019).

Thirumaladevi vd.'nin yaptığı çalışmada transfer öğrenimi yöntemini kullanmanın en iyi sonuçları elde etmede yardımcı olduğu savunulmuştur. Bu yüzden UCM ve Wuhan Üniversitesi'nde Uzaktan Algılama ve Görüntü Analizi için Sentinel-2 Görüntüleri (SIRI-WHU) veri kümeleri kullanılarak, önceden eğitilmiş Alex Net, Görsel Geometri Grubu 19 katmanlı (VGG-19) ve VGG-16 mimarileri ile transfer öğrenme uygulanmış; sahneleri

sınıflandırmak için SVM 'den yararlanılmıştır. UCM veri kümesi için %93,57, %94,08 ve %95 ve SIRI-WHU için %91,34, %92,78 ve %93,4'lük doğruluk oranları elde edilmiştir (Thirumaladevi, Swamy, & Sailaja, 2023).

Zhao vd.'nin yaptığı çalışmada CNN modellerinin öznelik çıkarma yeteneğinden etkili bir şekilde yararlanmak için, kendi kendini denetleyen öğrenme ve sınıflandırma görevlerini birleştiren bir Multi-Task Learning (MTL) çerçevesi önerilmiştir. Önerilen çerçevede ağırlıkları dinamik olarak ayarlamak için karıştırma kaybı stratejisi kullanılmış; bu strateji CNN modellerinin önemli özellik bilgilerini etkili bir şekilde çıkararak sınıflandırma zorluklarını daha da azaltmasına yardımcı olmuştur. NWPU-RESISC45, AID, UCM ve WHU-RS19 veri kümelerinde sırasıyla %94,21, %96,89, %99,11 ve %98,98 oranında doğruluk elde edilmiştir (Zhao, Luo, Li, Chen, & Piao, 2020).

Gao vd.'nin yaptığı çalışmada CNN özelliklerinin ayırt edici yeteneğini geliştirmek için dikkat modülünü ve ayrıca anlamsal sınıf bağımlılıklarını modellemek için Hiyerarşik Grafik Evrimsel Ağ (H-GCN)'ını kullanan yeni bir uzaktan algılama sahne sınıflandırma yöntemi önerilmiştir. Yalnızca CNN ile özellik öğrenimi yoluyla görüntü içeriğindeki temel bileşenlere odaklanmakla kalmayıp, aynı zamanda anlamsal sınıf bağımlılıklarını temsil eden grafiğin bitişiklik matrisindeki karıştırma güçlerini kullanarak daha bilgilendirici bir özellik gösterimi elde ettiklerini savunmuşlardır. NWPU-RESISC45, AID ve UCM veri kümelerinde sırasıyla %93,62, %95,78, %99 oranında doğruluk elde edilmiştir (Gao, Shi, Li, & Wang, 2021).

Cheon vd.'nin yaptığı çalışmada, uzaktan algılama sahne sınıflandırma görevleri için Kolmogorov-Arnold Ağı'nı (KAN) önceden eğitilmiş çeşitli CNN modelleriyle entegre ederek KCN adını verdikleri yeni bir yaklaşım önerilmiştir. VGG-16, Mobil Ağ Versiyon 2 (MobileNetV2), EfficientNet, Evrimsel Ağ Evrimi (ConvNeXt), Artık Ağ 101 Katmanı (ResNet101) ve ViT dahil olmak üzere birden fazla CNN tabanlı model kullanılmış ve bunların KAN ile eşleştirildiğinde performansları değerlendirilmiştir. Elde edilen sonuçlarda yaklaşımlarının daha az eğitim dönemi ve daha az parametreyle yüksek doğruluk elde ettiği görülmüştür. KAN ve CNN modellerinin gelişmiş yeteneklerini birleştirerek bu alanda yeni bir alternatif sunulmuştur. Ancak en büyük sınırlamalardan biri KAN'ın yeni bir algoritma olduğu göz önüne alındığında, performansını optimize

etmede ve çeşitli uzaktan algılama uygulamalarına entegrasyonunu sağlamada eksik kalmıştır (Chen, ve diğerleri, 2024)

Alhichri vd.'nin yaptığı çalışmada, CNN modeline dikkat mekanizması ekleyerek; EfficientNet-B3-Attn-2 adını verdikleri yeni bir sahne sınıflandırması modeli önerilmiştir. Önerilen model 6 farklı uzaktan algılama veri kümesi üzerinde test edilmiştir. UCM, Suudi Arabistan Krallığı Uzaktan Algılama Veri Kümesi (KSA), Optimal Arazi Örtüsü Sınıflandırma Veri Kümesi (OPTIMAL-31), (Uzaktan Algılama 7 Sınıflı Sahne Sınıflandırma Veri Kümesi (RSSCN7), AID ve WHU-RS19 veri kümelerinde elde ettikleri sonuçlar her birinde %95'in üzerindedir. Özellikle en iyi sonuç aldıkları UCM de %99,21; AID de %96,56 elde edilmiştir (Alhichri, Alswayed, Bazi, Ammour, & Alajlan, 2021).

Sinaga vd.'nin yaptığı çalışmada transfer öğrenmesi mimarisinde Ağırlıklı Kayıp, Seyrek Düzenleme ve Budama kullanılarak uzaktan algılama görüntüleri sınıflandırması için ResNet-50, InceptionV3 ve EfficientNet CNN modelleri üzerinde karşılaştırmalı bir çalışma yapılmıştır. Üç model ile de AID veri kümesinde %95'in üzerinde doğrulukla performans elde etmişlerdir ancak EfficientNet-B7 mimarisi %97,46 ile en yüksek doğruluğa ulaşmıştır. Bir diğer dikkat çektikleri kısım ise transfer öğrenmesi eğitimi sırasında modele Dondurulmuş Katman, Ağırlıklı Kayıp ve Seyrek Düzenleme gibi yöntemler uygulandığında doğruluğun arttığı gözlemlenmiştir (Sinaga, Yudistira, & Santoso, 2024).

Kwak vd.'nin yaptığı çalışmada, uzaktan algılama görüntülerinin arazi örtüsü sınıflandırması için yarı-denetimli bir yaklaşım sunulmuştur. Yöntem, Döngü Tutarlı Üretken Çekişmeli Ağlar (CycleGAN) ve EfficientNet modellerini birleştirerek, etiketlenmiş verilerin sınırlı olduğu durumlarda bile yüksek doğrulukla arazi örtüsü sınıflandırması yapmayı sağlamaktadır. EfficientNet; ağ genişliği, ağ derinliği ve çözünürlük olmak üzere üç faktörün bileşik model ölçeklemesi yoluyla üstün sınıflandırma sonuçları göstermiştir. Geleneksel yöntemlerin bu üç faktörden yalnızca birini ölçeklediği düşünülürse EfficientNet'in bu faktörlerin sabit katsayılar kümesiyle düzgün bir şekilde ölçeklenebileceği görülmüştür (Kwak & Kim, 2023).

CNN algoritmaları büyük veri kümelerinde daha iyi doğruluk elde etmek amacıyla karmaşık, anlaşılabilir ve maliyetli stratejiler kullanır. Bu sorun üzerine Song'un yaptığı çalışmada EfficientNetV2 CNN modelini kullanarak FST EfficientNetV2 adını verdiği transfer öğrenme stratejisi içeren ve hızlı eğitim özelliklerine karşılık gelen yeni ve etkili bir yöntem önerilmiştir. Önerilen yöntem ile Grafik İşlem Birimi (GPU) donanım maliyetleri %75-%81 oranında bir azalma, eğitim süresi maliyetlerinde ise %60-%80 oranında bir azalma gözlemlenmiştir. Diğer CNN algoritmalarına göre kıyasla da sınıflandırma görevinde %1,7-%6,2 oranında doğruluk artışı elde edilmiştir (Song, 2023).

Zhang vd.'nin yaptığı çalışmada uzaktan algılama görüntü sınıflandırması için ince ayar stratejisine sahip önceden eğitilmiş EfficientNet modellerine dayalı bir transfer öğrenme yöntemi sunulmuştur. Yöntemlerini birden fazla uzaktan algılama veri kümesi üzerinde doğrulamışlardır, deneysel sonuçlar uzaktan algılama görüntülerinde sahne sınıflandırması için önerilen yöntemin etkinliğini ve üstünlüğünü göstermiştir. Üç veri kümesindeki önceden eğitilmiş EfficientNet yöntemlerinin (AID, UCM ve NWPU) klasik önceden eğitilmiş CNN'lere kıyasla çok daha iyi sınıflandırma doğrulukları (sırasıyla %97,55, %98,81, %96,19) elde ettiğini görmüşlerdir. Deneysel sonuçların analizinden, EfficientNet'in uzaktan algılama görüntü veri kümesindeki genelleme ve transfer öğrenme yeteneğinin çok iyimser olduğu sonucuna varabiliriz (Zhang, Liu, & Shi, 2020).

Yin vd.'nin yaptığı çalışmada, EfficientNet algoritmasının verimliliğini ve doğruluğunu doğrulamak için UCM veri kümesini EfficientNet'e göre sınıflandırarak EfficientNet serisinin uzaktan algılama görüntü sınıflandırmasında uygulanması sistematik olarak incelenmiştir. Deneysel sonuçlar, VGGNet, Artık Ağ (ResNet) ve Mobil Ağ (MobileNet) ile karşılaştırıldığında, EfficientNet ağının bileşik parametreleri tanıttığını ve aynı anda derinliği, genişliği ve çözünürlüğü ölçeklediğini göstermiştir. EfficientNetV2S'nin doğrulama kümesindeki doğruluğu, VGG, MobileNetV2, Artık Ağ 34 Katmanı (ResNet34) ve EfficientNet-b0'dan sırasıyla %16,5, %5,2, %1,8 ve %1,7 daha yüksektir ve bu da EfficientNet ağının uzaktan algılama görüntü sınıflandırma veri kümesindeki verimliliğini ve doğruluğunu göstermiştir. Ancak yalnızca küçük bir veri kümesinde uygulanmıştır (Ying, Yang, & Lu, 2022).

ViT modellerinin uzaktan algılama verileri üzerindeki etkinliğini gösteren akademik çalışmalar mevcuttur. Hiperspektral görüntü sınıflandırması için özel olarak tasarlanmış hiyerarşik spektral ViT mimarisi üzerine bir çalışma, ViT'nin bu alandaki potansiyelini ortaya koymaktadır (Liu, Prasad, & Crawford, 2024).

Bir diğer çalışmada yazarlar ViT tabanlı bir model önererek, çeşitli uzaktan algılama veri kümeleri üzerinde deneyler gerçekleştirmişlerdir. Sonuçlar ViT modellerinin, özellikle büyük veri kümeleri üzerinde eğitildiklerinde, yüksek sınıflandırma doğruluğu elde edebildiklerini göstermektedir. UCM veri setinde %98,49, AID veri setinde %95,86, Optimal31 veri setinde %95,56 ve NWPU veri setinde %93,83 doğruluk oranlarına ulaşılmıştır (Bazi, Bashmal, Rahhal, Dayil, & Ajlan, 2021).

Deng vd.'i sahne sınıflandırmasında, CNN'leri ve ViT'i birleştirerek CTNet adını verdikleri yüksek performanslı bir ortak çerçeve önerilmiştir. CTNet yöntemi, ViT akışı (T akışı) ve CNN akışı (C akışı) olmak üzere iki modül içerir. T akışı için, düzeltilmiş görüntü yamaları, HRRS görüntülerindeki anlamsal öznitelikleri çıkarmak için önceden eğitilmiş ViT modeline gönderilir. T akışını tamamlamak için, önceden eğitilmiş CNN, C akışındaki yerel yapısal öznitelikleri çıkarmak üzere aktarılır. Daha sonra, bilinmeyen örneklerin etiketlerini tahmin etmek için anlamsal öznitelikler ve yapısal öznitelikler birleştirilir. CTNet yöntemi ile AID ve NWPU-RESISC45 veri kümelerinde doğruluklar sırasıyla %97,70 ve %95,49'dur (Deng, Xu, & Huang, 2021).

Pengyuan vd.'nin yaptığı çalışmada sahne sınıflandırmasını geliştirmek amacıyla bir ViT modeli olan SCViT önerilmiştir. Önerilen SCViT modeli, görüntülerin hem mekansal hem de kanalsal özniteliklerini yakalamak ve öğrenmek için özel bir mimari sunar. Bu mimari için görüntünün pikseller arasındaki mekansal ilişkilerini koruyan bir modül entegre edilmiştir. İkinci olarak farklı kanallar arasındaki önemli bilgilerin korunması için özel bir kanal dikkat mekanizması kullanılmıştır. ViT mimarisini temel alarak, büyük miktarda bağlam bilgisi ve karmaşık görsel ilişkileri öğrenebilme kapasitesi vurgulanmıştır. En son sahne sınıflandırma yöntemleriyle karşılaştırıldığında, yapılan bu çalışma ile deneysel sonuçlar sahne sınıflandırmasında ViT modellerinin kullanılma potansiyelini doğrulamıştır (Lv, Wu, Zhong, Du, & Zhang, 2022).

Zongyao vd.'nin çalışmasında, uzaktan algılama görüntülerinde sahne sınıflandırmasını iyileştirmek için çoklu örnek öğrenme paradigmasıyla uyumlu bir ViT modeli önerilmiştir. Bu model, uzaktan algılama görüntülerinin genellikle heterojen özelliklere sahip karmaşık sahneler içerdiği gerçeğinden yola çıkarak tasarlanmıştır. Yeni bir çoklu örnek görüntü dönüştürücüsü (MITformer), bir görüntüyü birden fazla örnek şeklinde temsil eder ve her örneğin mekansal ve bağlamsal özniteliklerini etkin bir şekilde çıkararak bütünsel bir sahne sınıflandırması gerçekleştirir. Transformer tabanlı dikkat mekanizması, farklı örnekler arasındaki uzun menzilli ilişkileri öğrenirken; çoklu örnek öğrenme çerçevesi, modelin yalnızca en bilgilendirici örnekler üzerinde yoğunlaşmasını sağlar. Bu hibrit yaklaşım hem detaylı lokal özellikleri hem de geniş bağlam ilişkilerini yakalayarak, uzaktan algılama sahne sınıflandırmasında geleneksel yöntemlere kıyasla üstün performans sergilemiştir (Sha & Li, 2022).

3. MATERYAL VE METOT

Bu bölümde, tez kapsamında gerçekleştirilen arařtırmalar için kullanılan materyal ve yöntemler detaylandırılmıřtır. Materyal bölümü içerisinde kullanılan veri kümeleri açıklanmıř; örnek görüntüler verilmiřtir. Metot bařlığında ise tez kapsamında kullanılan yöntemler ve uygulamaları adım adım açıklanmıřtır.

3.1. Materyal

Tezin bu bölümünde derin öğrenme modelinin eğitim ve test ařamalarında kullanılan veri kümelerinden bahsedilmiřtir.

3.1.1. Veri kümesi

Sahne sınıflandırması alanında otomatik sınıflandırma ve etiketleme yöntemlerinin geliştirilmesi ve iyileřtirilmesi için kamuya açık yüksek çözünürlüklü fazla sayıda farklı veri kümeleri bulunmaktadır. Bu bölümde önerilen modeli deęerlendirmek için kullanılan veri kümeleri üzerinde durulmuř ve veri kümelerinin özniteliklerinden bahsedilmiřtir. Tablo 3.1'de uzaktan algılama görüntülerini içeren veri kümeleri hakkında bilgi verilmiřtir.

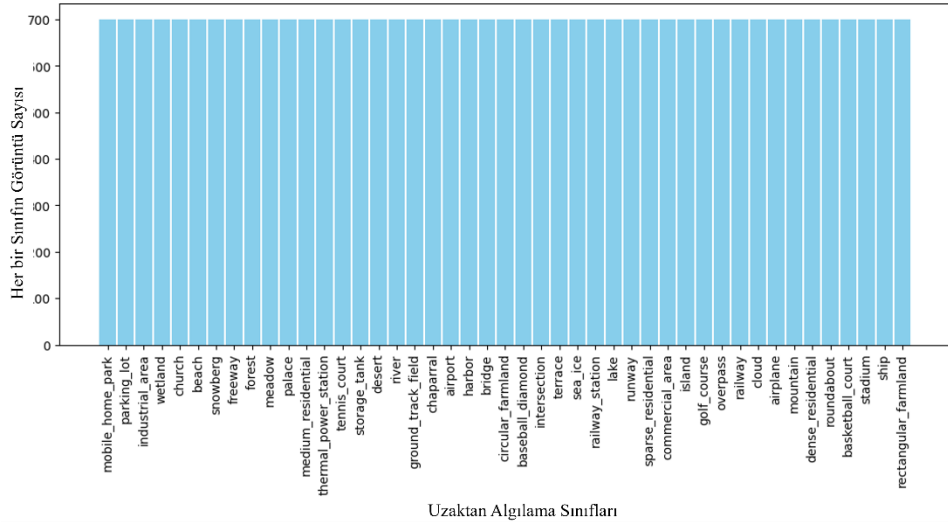
Tablo 3.1. *Uzaktan algılama görüntü veri kümeleri*

Veri kümesi	Bir sınıftaki görüntü sayısı	Sınıf sayısı	Toplam görüntü sayısı	Görüntü boyutu
UCM	100	21	2100	256 x 256
WHU-RS19	50 ~ 60	19	1500	600 x 600
AID	220 ~ 420	30	10000	600 x 600
NWPU-RESISC45	700	45	31500	256 x 256
OPTIMAL-31	60	31	1860	256 X 256
SRI-WHU	200	12	2400	200 x 200
PatternNet	800	38	30400	256 x 256
KSA	250	13	3250	256 x 256
RSCN7	400	7	2800	400 x 400
RSI-CB256	yaklařık 690	35	24000	256 x 256
RSI-CB128	yaklařık 800	45	36000	128 x 128
RSC11	yaklařık 100	11	1232	512 x 512

Bu çalışmada, NWPU-RESISC45, UCM ve AID veri kümeleri kullanılmıştır. Bu veri setlerinin seçilmesinin birkaç önemli nedeni bulunmaktadır. Söz konusu veri kümeleri, literatürde yaygın olarak kullanılan veri setleri arasında yer almakta olup, farklı sahne kategorileri içermeleri, görüntü sayısı açısından çeşitlilik göstermeleri ve farklı çözünürlüklerde olmaları nedeniyle bu veri kümeleri ile çalışılmıştır. Bu veri kümeleri üzerinde gerçekleştirilen testler, çalışmanın mevcut literatürle kıyaslanabilir olmasını sağlamak ve elde edilen sonuçların önceki çalışmalarla karşılaştırılmasına olanak tanımaktadır. NWPU-RESISC45, sahne sınıfı sayısı ve toplam görüntü miktarı bakımından en büyük ölçekli veri kümelerinden biri olup, sunduğu yüksek çeşitlilik ve zorluk seviyeleri sayesinde modelin genelleme yeteneğini değerlendirmek açısından önemli bir veri kümesi niteliğindedir. UCM veri kümesi, literatürde standart bir kıyaslama veri kümesi olarak yaygın şekilde kullanılan ve uzaktan algılama görüntülerinin sınıflandırılmasına yönelik temel test ortamlarından birini oluşturan bir veri kümesidir. Çalışmada farklı veri kümelerinin kullanılması, modelin farklı çözünürlüklerde ve içeriklerdeki görüntüler üzerindeki performansını kapsamlı bir şekilde değerlendirmeyi sağlamış ve sınıflandırma modelinin çeşitli sahnelerde genelleme kapasitesini artırmaya yönelik bir analiz ortamı sunmuştur.

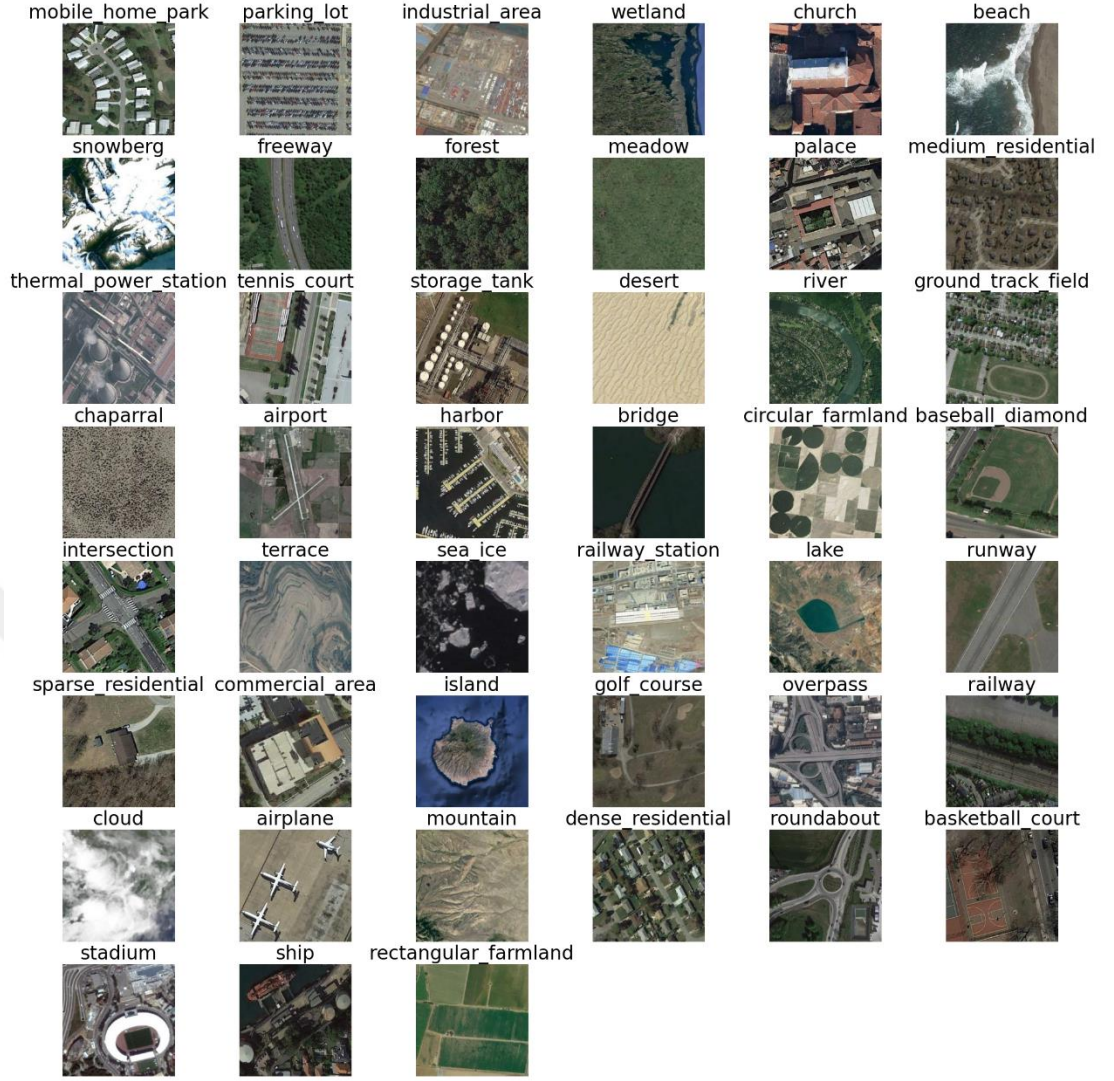
3.1.1.1. NWPU-RESISC45 veri kümesi

NWPU-RESISC45 veri kümesi, diğer çoğu uzaktan algılama veri kümelerine göre daha karmaşıktır ve Şekil 3.1'de gösterildiği gibi her sınıfında 700 görüntü içeren 45 sahne sınıfına bölünmüş toplam 31.500 uzaktan algılama görüntüsünden oluşmaktadır. Her bir görüntü, kırmızı yeşil mavi (RGB) renk uzayında 256×256 piksel boyutlarındadır. Bu veri kümesi, alanında uzman kişiler tarafından Google Earth'ten toplanmıştır. Dünya genelinde gelişmekte olan, geçiş aşamasındaki ve oldukça gelişmiş ekonomilere sahip 100'den fazla ülke ve bölgeyi kapsamaktadır.



Şekil 3.1. NWPU-RESISC45 veri kümesinin sınıfları ve sayıları

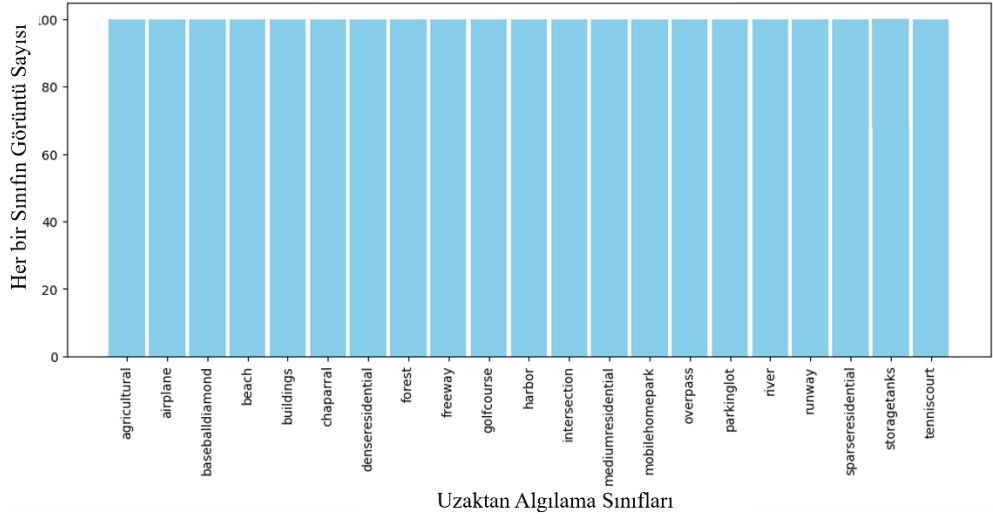
Oldukça geniş kapsamlı olan bu veri kümesi toplam görüntü sayısı, sahne sınıflarının sayısı ve her bir sınıfa düşen görüntü sayısı açısından uzaktan algılama görüntü sınıflandırmasına yönelik en büyük veri kümesidir. Diğer bir yandan görüntülerin çözünürlük, ışık, aydınlatma gibi görüntü farklılıkları diğer veri kümelerine göre başarılıdır. Şekil 3.2'de bu veri kümesindeki her sınıftan birer örnek gösterilmiştir.



Şekil 3.2. NWPU-RESISC 45 veri kümesine ait görüntü örnekleri (Han, Cheng, & Lu, 2017)

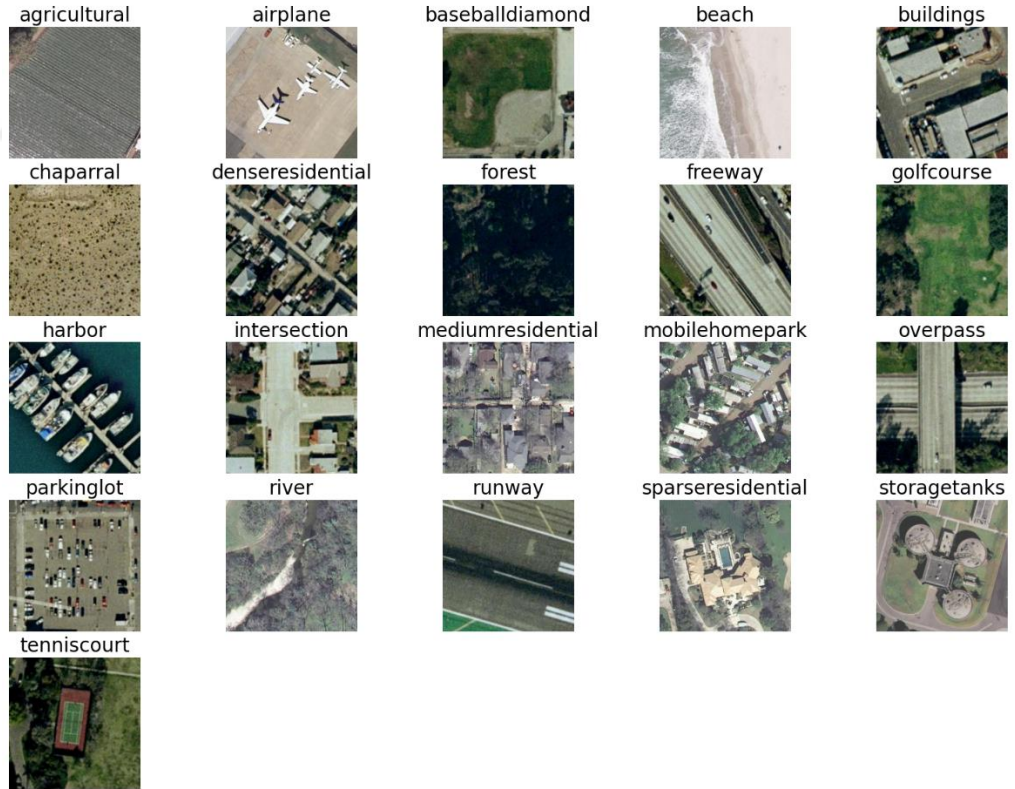
3.1.1.2. UCM veri kümesi

UCM veri kümesi, 21 arazi kullanım sınıfını gösteren hava görüntülerinden oluşan bir veri kümesidir. Şekil 3.3'te sınıflar ve her bir sınıfın 100 görselden oluştuğu gösterilmiştir. Bu veri kümesi ABD bölgelerini kapsayan Amerika Birleşik Devletleri Jeolojik Araştırması (USGS) Ulusal Harita Kentsel Alan Görüntüleri koleksiyonundan elde edilen bir veri kümesidir.



Şekil 3.3. UCM veri kümesinin sınıfları ve sayıları

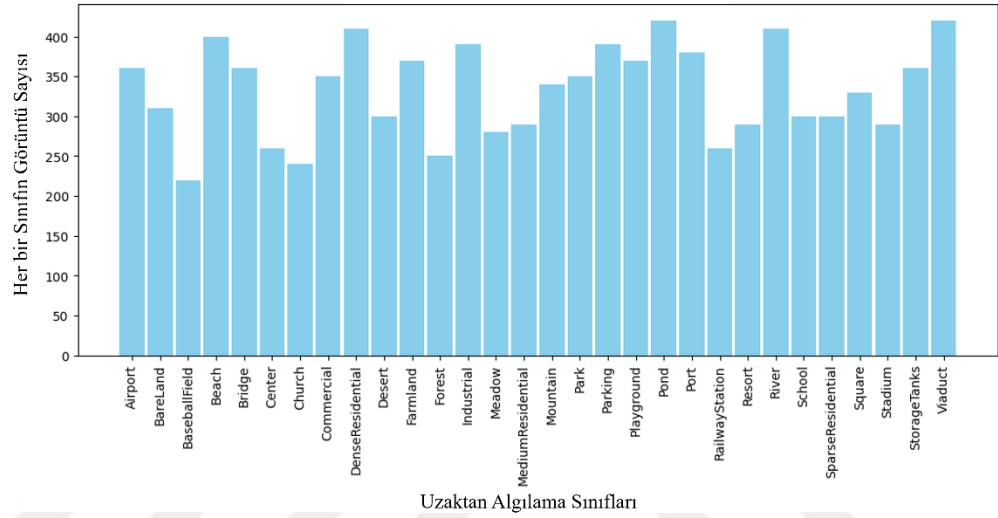
Şekil 3.4'te bu veri kümesindeki her sınıftan birer örnek gösterilmektedir.



Şekil 3.4. UCM veri kümesine ait görüntü örnekleri (Yang & Newsam, 2010)

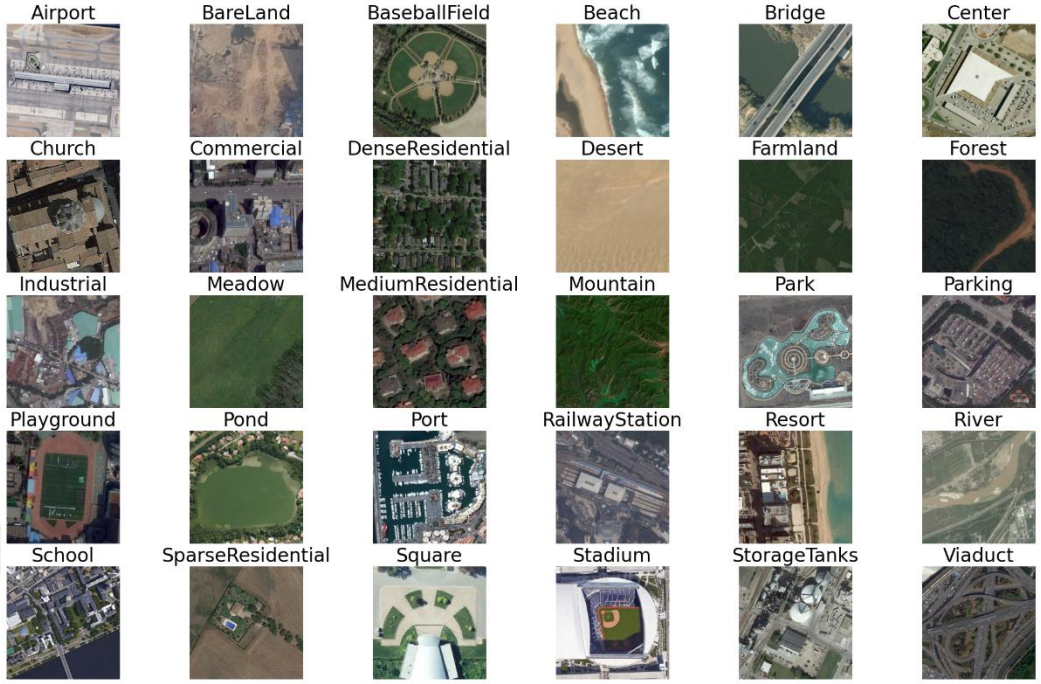
3.1.1.3. AID veri kümesi

AID veri kümesi, Şekil 3.5'te gösterildiği üzere 30 arazi kullanım sahnesi sınıfından toplam 10.000 hava görüntüsünden oluşan bir veri kümesidir. UCM veri kümesi ile karşılaştırıldığında, AID çok daha fazla görüntü içermekte ve daha geniş bir sahne kategorisi yelpazesini kapsamaktadır. AID, Google Earth Görüntülerinden elde edilen bir veri kümesidir.



Şekil 3.5. AID veri kümesinin sınıfları ve sayıları

Şekil 3.6'da bu veri kümesindeki her sınıftan birer örnek gösterilmektedir.



Şekil 3.6. AID veri kümesine ait görüntü örnekleri (Xia, ve diğerleri, 2017)

3.2. Metot

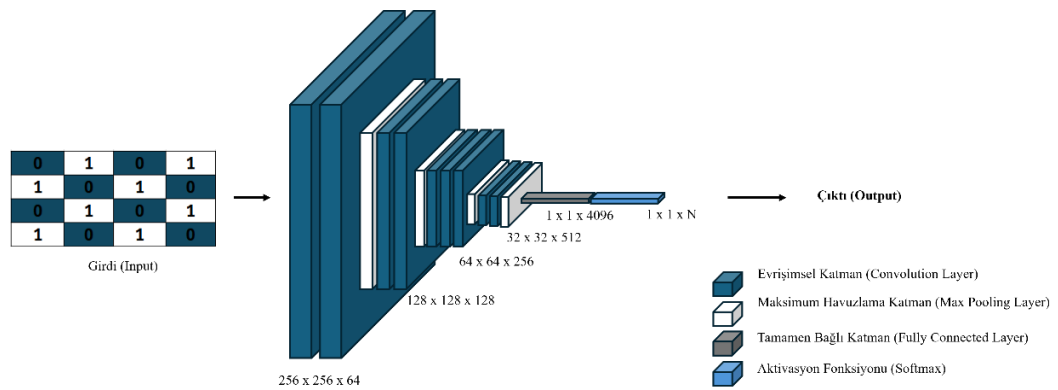
Tezin bu bölümünde kurulan derin öğrenme modelinden ve uygulanan yöntemlerden bahsedilmektedir. Transfer öğrenimi ve EfficientNetV2S ile EfficientNetV2M CNN'leri kısaca açıklanacak, ViT mimarisi tanıtılacak daha sonra önerilen mimari ve yapılan deneyler detaylandırılacaktır.

3.2.1. Evrimsel sinir ağları

CNN, modern derin öğrenme dünyasında görsel verilerin işlenmesinde görsel veri analizinde devrim yaratan bir yapay sinir ağı türüdür. Özellikle görüntü ve video tanıma, nesne tespiti ve sınıflandırma gibi alanlarda üstün performans sergileyen etkili bir derin öğrenme modelidir.

Temelde, bir görüntü üzerinde evrişim işlemleri uygulayarak verinin özniteliklerini çıkarmak için kullanılırlar. CNN'lerin üç ana bölümü Şekil 3.7'de gösterilmiştir. Konvolüsyon katmanlar giriş görüntüsünden öznitelik haritaları çıkararak temel görsel özniteliklerini tanır. Konvolüsyon katmanları, veriyi düşük seviyeli özniteliklere dönüştürmek için giriş verisini küçük parçalara bölerek her bir parçada filtreler (kernels)

uygular. Bu işlem, bir veya birden fazla filtre kullanarak yapılır. Her filtre, girdi verisinde belirli öznitelikleri (örneğin kenarlar, köşeler, desenler) tespit etmek için kullanılır. Çoğunlukla her konvolüsyon katmanının ardından, doğrusal olmayan Doğrultulmuş Lineer Ünite (ReLU) fonksiyonu kullanılır. Bu fonksiyon, ağı doğrusal olmayan ilişkileri öğrenmesini ve daha karmaşık modeller oluşturmasını sağlar. Havuzlama Katmanları konvolüsyon katmanlarından gelen özellik haritalarını daha küçük, özetlenmiş versiyonlarına indirir. Havuzlama, özellik haritasını küçültmek ve önemli özellikleri vurgulamak amacıyla kullanılır. En yaygın kullanılan teknikler maksimum havuzlama (max pooling) ve ortalama havuzlamadır (average pooling). Bu katmanlar, modelin hesaplama yükünü azaltırken, aynı zamanda özelliklerin konumuna olan duyarlılığı da azaltır. Örneğin, maksimum havuzlama, her bir bölge içindeki en büyük değeri alarak boyutu azaltır ve önemli özellikleri korur. Bir dizi konvolüsyon ve havuzlama katmanlarından sonra ağ genellikle bir veya daha fazla tam bağlantı katmanı içerir. Tam bağlantılı katmanlar, ağırlıklar ve önyargılar aracılığıyla tüm özellik haritalarını birleştirir. Bu katmanlar, özellik haritalarını birleştirdikten sonra düzleştirir ve sınıflandırma, tespit veya regresyon gibi görevleri gerçekleştirmek için kullanılır. Son katman genellikle bir softmax fonksiyonu kullanarak, ağı farklı sınıflar için olasılık puanları üretmesini sağlar (Goodfellow, Bengio, & Courville, 2016).



Şekil 3.7. CNN mimarisi

CNN'lerin önemli bir özelliği, ağırlık paylaşımıdır. Yani, bir filtre bir katmanda belirli bir özelliği öğrenirse, bu filtre diğer katmanlarda da aynı özelliği tanımak için kullanılabilir. Bu, ağırlık genelde daha az parametreye sahip olmasını sağlar ve böylece eğitim sürecini hızlandırır. CNN'ler genellikle sınıflandırma, nesne tespiti, yüz tanıma, dil işleme ve daha birçok uygulamada kullanılmaktadır. Örneğin, görüntülerdeki nesnelere sınıflandırmak için yaygın olarak kullanılan bir yapıdır ve görüntülerdeki örüntüyü yakalayabilmede oldukça başarılıdır. Bu, bir nesnenin ne olduğunu tanımak için ağırlık sonunda bir sınıflandırma katmanı eklenerek gerçekleştirilir. Sonuç olarak, CNN'ler, bilgisayarlı görüş ve resim işleme alanında büyük başarı elde etmiş, özellikle büyük veri kümeleri üzerinde karmaşık özniteliklerin tespiti ve sınıflandırılması gibi görevlerde oldukça etkili olan derin öğrenme modelleridir.

3.2.2. Transfer öğrenimi

Transfer öğrenmesi, makine öğrenmesi ve özellikle derin öğrenme alanında önemli bir yaklaşımdır. Transfer öğrenmesi, bir modelin belirli bir görevde kazandığı bilgiyi başka bir görevde kullanarak performansını artırmayı amaçlayan bir makine öğrenimi tekniğidir. Geleneksel makine öğrenme yöntemleri, genellikle her yeni görev için sıfırdan bir model eğitmeyi gerektirir. Ancak bu yöntem mevcut bir modelin önceki bilgi ve deneyimlerinden yararlanarak yeni görevlerde daha hızlı ve daha verimli bir şekilde performans göstermesini sağlar. Bu süreç, modelin belirli katmanlarını yeniden eğiterek veya tamamen yeni bir model kullanarak gerçekleştirilebilir. Modelin daha önce öğrenilen bilgileri yeni görevlere uygulaması hem zaman hem de hesaplama maliyetlerinden tasarruf sağlar.

Temel çalışma prensibi, önceden eğitilmiş modelin öznitelik çıkarıcı olarak kullanılması ve bu özniteliklerin yeni görevde bir sınıflandırıcıya veya regresyon modeline girdi olarak verilmesidir. Önceden eğitilmiş bu modelin tüm katmanları veya belirli katmanları, hedef görevdeki veri kümesiyle yeniden eğitilir. Bu, modelin hedef göreve özel bilgileri öğrenmesini sağlar. Son olarak transfer öğrenmesi, birçok alanda başarıyla uygulanmaktadır. Bu yöntem, özellikle görüntü tanıma, doğal dil işleme ve konuşma tanıma gibi alanlarda geniş çapta kullanılmaktadır. Örneğin, ImageNet veri kümesinde eğitilmiş derin sinir ağları, daha sonra tıbbi görüntüleme veya otonom sürüş gibi özel alanlarda ince ayar yapılarak kullanılmıştır (Shin, ve diğerleri, 2016).

3.2.3. EfficientNetV2S derin öğrenme modeli

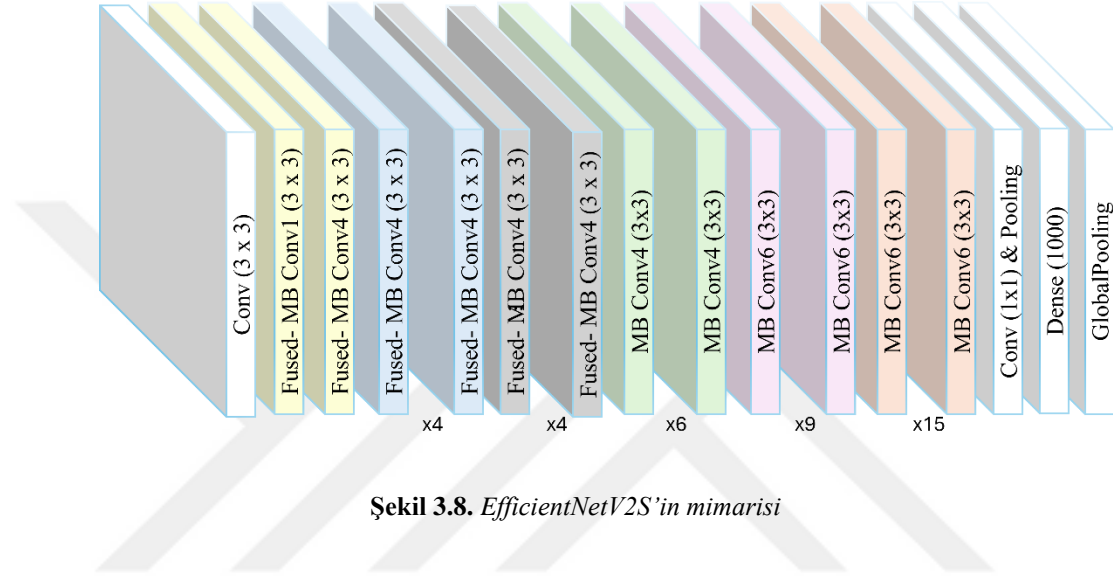
EfficientNetV2S, Google tarafından 2021'de geliştirilen ve EfficientNet ailesinin bir üyesi olan güçlü bir görüntü sınıflandırma modelidir. ImageNet veri kümesi üzerinde eğitilen bu model ailesi, çeşitli görüntü sınıflandırma görevlerinde üstün performans sergilemiştir. EfficientNet B0-B7 modelleri, diğer yaygın CNN mimarileri ile karşılaştırıldığında daha yüksek doğruluk oranları sunmuş ve hesaplama açısından daha verimli olduğu gösterilmiştir. (Tan & Le, 2019)

EfficientNetV2S, farklı ölçeklerde kullanılabilir şekilde tasarlanmış olup, hızlı eğitim süresi ve verimli parametre kullanımı ile öne çıkmaktadır. Modelin mimarisi, Mobil Ters Darboğaz Konvolüsyonu (MBConv), Birleştirilmiş Mobil Ters Darboğaz Konvolüsyonu (Fused MBConv), Swish aktivasyon fonksiyonu, BatchNorm normalleştirme katmanı ve stokastik derinlik atlama bağlantıları gibi bileşenlerden oluşmaktadır. Geleneksel derin öğrenme modelleri yüksek hesaplama gücü ve bellek gerektirirken, EfficientNetV2S, hesaplama maliyetini düşüren optimize edilmiş konvolüsyonel blokları ile bu soruna çözüm sunmaktadır. (Tan & Le, 2021)

EfficientNetV2S'nin temel yapı taşları MBConv ve Fused MBConv bloklarıdır. MBConv, 1×1 genişleme evrişimi, Derinlemesine Ayrılabilir Konvolüsyon (DWConv) ve 1×1 sıkıştırma evrişimi kullanarak hesaplama maliyetini azaltır. DWConv sayesinde her kanal için ayrı evrişim uygulanarak mekânsal ve kanal bazlı özellik çıkarımı daha verimli hale getirilir. Fused MBConv, modelin erken aşamalarında hesaplama sürecini hızlandırmak amacıyla geliştirilmiştir. Bu blok, DWConv yerine doğrudan standart evrişim kullanarak daha güçlü mekânsal özellikler yakalar ve işlem süresini optimize eder. MBConv blokları, modelin orta ve ileri aşamalarında verimli özellik çıkarımı sağlarken, Fused MBConv blokları, erken aşamalarda işlem hızını artırarak toplam eğitim süresini kısaltır. Bu iki yapı taşının birleşimi, EfficientNetV2'nin hem hızlı eğitilmesini hem de yüksek doğruluk oranı elde etmesini sağlar.

EfficientNetV2, ağ derinliği, genişliği ve çözünürlüğünü bileşik ölçekleme stratejisi ile optimize ederek diğer modellerden farklı bir yaklaşım benimser. Bileşik ölçekleme, kaynak kullanımını minimize ederken doğruluğu korumak için model boyutları arasında hassas bir denge kurar. Bu model, Nöral Mimari Arama (NAS) yöntemi kullanılarak tasarlanmış ve hesaplama verimliliği ile model performansı arasında optimal bir denge sağlamıştır.

NWPU-RESISC45 veri kümesi, farklı çözünürlük ve karmaşıklığa sahip geniş çaplı uzaktan algılama görüntüleri içerdiğinden, eğitim süreleri genellikle uzundur. Ancak bu çalışmada, EfficientNetV2S kullanılarak eğitim süresi önemli ölçüde azaltılmıştır. Şekil 3.8, EfficientNetV2S modelinin genel şemasını sunarak bu stratejinin uygulanmasını görselleştirmektedir.



Şekil 3.8. *EfficientNetV2S'in mimarisi*

3.2.4. EfficientNetV2M derin öğrenme modeli

EfficientNetV2 model ailesindeki varyantlardan bir diğeri olan EfficientNetV2M modeli, derin öğrenme alanında, özellikle görüntü sınıflandırma görevlerinde yüksek doğruluk sunan yenilikçi bir mimariye sahiptir (Tan & Le, 2021). Bu modelin temel özelliği, büyük ve karmaşık veri kümesi için optimize edilmiş olan yapısıdır. MBConv ve Fused MBConv gibi yapı taşlarına dayanarak çalışır. Bu bloklar, hesaplama kaynaklarını daha etkili kullanarak daha az parametreyle daha çok işlem yapabilme avantajı sağlar. EfficientNetV2M, bileşik ölçekleme stratejisini kullanarak ağır derinliği, genişliği ve çözünürlüğünü optimize eder. NAS yöntemiyle tasarlanan model, hesaplama maliyetleri ile model performansı arasındaki dengeyi sağlamak amacıyla geliştirilmiştir.

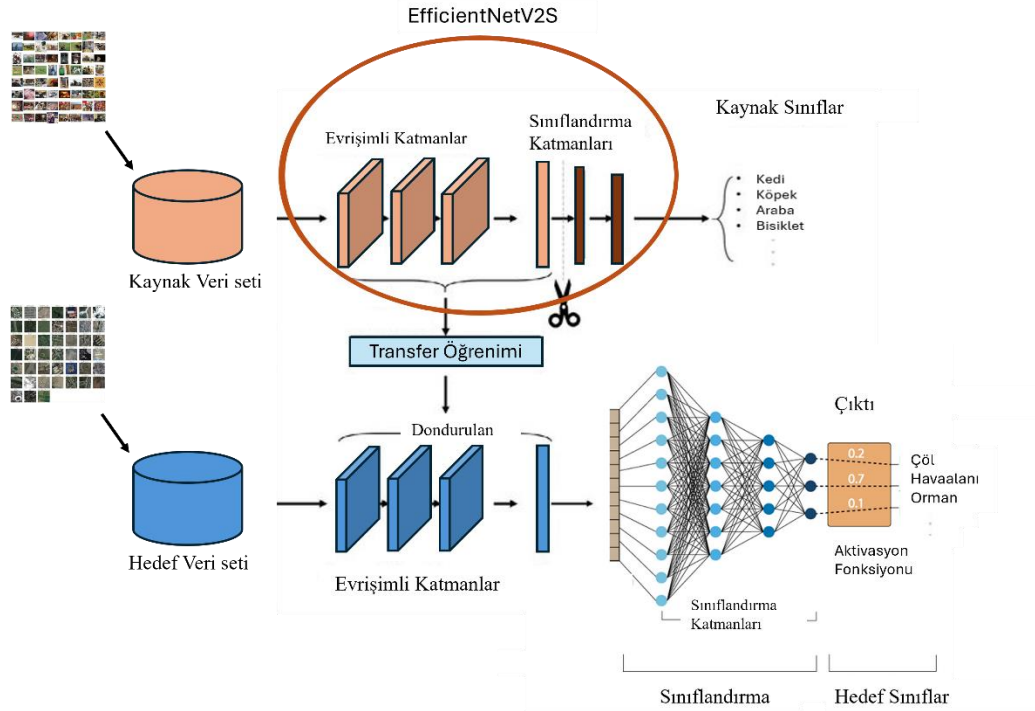
EfficientNetV2 model ailesinin EfficientNetV2S ve EfficientNetV2M varyantları, model boyutu, hesaplama maliyeti ve kullanım senaryoları açısından farklılık göstermektedir. Modelin mimari tasarımında, EfficientNetV2M, daha büyük ve karmaşık bir model olduğu için EfficientNetV2S'e göre daha fazla katman derinliği ve MBConv blokları içerir. Ek olarak katmanlarının genişlikleri (filtre sayısı) de artırılmıştır. Bu

durum, modelin daha fazla parametreye sahip olması ve dolayısıyla daha karmaşık görevlerde daha yüksek performans göstermesi anlamına gelir. Bununla birlikte, bu genişletilmiş yapı, daha yüksek bellek ve hesaplama kaynakları gerektirir. EfficientNetV2S, düşük çözünürlüklü görüntülerle hızlı eğitim süreçleri sunarken, EfficientNetV2M, daha yüksek çözünürlüklü veri üzerinde detaylı analizler ve üstün performans elde edebilmek için optimize edilmiştir.

3.2.5. EffiSceneNet modeli

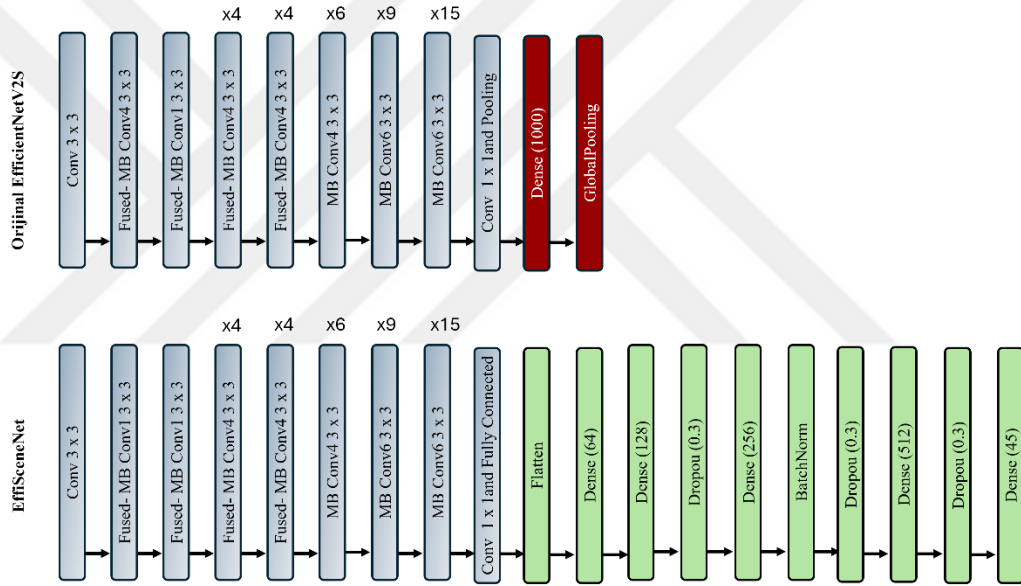
Transfer öğrenme yöntemi kullanarak görüntü sınıflandırma görevini gerçekleştirmek üzere EfficientNetV2S tabanlı bir derin öğrenme modeli önerilmiştir. Bu yöntemde, EfficientNetV2S modelinin önceden eğitilmiş ağırlıklarından yararlanılmış ve bir yapay sinir ağı ile entegre edilerek EffiSceneNet adında yeni bir model geliştirilmiştir.

EfficientNetV2S modelinde son katmanları (sınıflandırma katmanı) çıkarılarak transfer öğrenimi uygulanmıştır. Modelin ImageNet veri kümesi üzerinde önceden eğitilmiş ağırlıkları kullanılır ve EfficientNetV2S modelinin katman parametreleri (örneğin, ağırlıklar ve biaslar) geri yayılım sırasında güncellenerek, eğitime dahil edilmiştir. Önerilen yöntemin mimarisi Şekil 3.9’da gösterilmiştir.



Şekil 3.9. Önerilen yöntemin iş akışı

Transfer öğrenimi kapsamında, EfficientNetV2S modeli önceden ImageNet veri kümesi ile eğitilmiş olup, transfer öğrenimi sırasında "include_top=False" hiperparametresi kullanılarak modelin orijinal sınıflandırma katmanları çıkarılmıştır. Böylece, EfficientNetV2S modelinin önceden eğitilmiş ağırlıkları korunmuş, ancak sınıflandırma süreci önerilen EffiSceneNet modeline özgü hale getirilmiştir. Orijinal modelin son sınıflandırma katmanları yerine veri kümesine uygun şekilde özelleştirilmiş yeni sınıflandırma katmanları entegre edilmiştir. Şekil 3.10, orijinal EfficientNetV2S modelinden çıkarılan katmanları ve önerilen EffiSceneNet modeline eklenen yeni katmanları görselleştirmektedir.



Şekil 3.10. Orijinal EfficientNetV2S ile önerilen EffiSceneNet modellerinin farkı

EfficientNetV2S modeline ek olarak, önerilen EffiSceneNet modelinde ek sınıflandırma katmanları kullanılmıştır. Flatten katmanı, modelin çıktısını tek boyutlu bir vektöre dönüştürerek tam bağlantılı katmanlara giriş sağlamıştır. Bunu takiben, 64, 128, 256 ve 512 nöronlu tam bağlantılı (Dense) katmanlar eklenmiş ve her biri ReLU aktivasyon fonksiyonu ile modellenmiştir. Tam bağlantılı katmanlar, modelin daha karmaşık özellikleri öğrenmesine ve sahne sınıflandırma görevinde daha iyi genelleştirme yapmasına yardımcı olmaktadır. Ayrıca, Dropout katmanları (%30 oranında), modelin eğitim sürecinde aşırı öğrenmeyi önlemek amacıyla eklenmiştir. Dropout, her eğitim

iterasyonunda belirli oranlarda nöronları rastgele devre dışı bırakarak, modelin belirli ağırlıklara aşırı bağımlı hale gelmesini engeller ve genelleştirme yeteneğini artırır. Batch Normalization katmanı ise modelin iç dağılımlarını stabilize etmek, eğitim sürecini hızlandırmak ve gradyan kaybolma sorununu azaltmak için dahil edilmiştir. Bu sayede model, daha hızlı ve stabil bir şekilde öğrenim sürecini tamamlamaktadır. Son olarak, modelin çıkış katmanı, çalışılan veri kümesine uygun sınıf sayısı kadar nörondan oluşmakta olup, softmax aktivasyon fonksiyonu kullanılarak her sınıf için olasılık değerleri üretilmiştir. Böylece, modelin çıktısı çok sınıflı bir olasılık dağılımına dönüştürülmüş ve en yüksek olasılığa sahip olan sınıf, modelin tahmini olarak belirlenmiştir.

3.2.6. Topluluk öğrenmesi

Topluluk öğrenmesi, makine öğrenimi modellerinin performansını artırmak amacıyla birden fazla modeli bir araya getiren bir yöntemdir. Bu yaklaşımın temelinde, bireysel modellerin farklı zayıf yönlerini telafi ederken güçlü yönlerini birleştirerek daha genel ve doğru tahminler elde etme fikri yatmaktadır. Topluluk öğrenme yöntemleri, genellikle model çeşitliliğini ve tahmin doğruluğunu optimize etmek için farklı algoritmaların, hiper parametrelerin veya veri alt kümelerinin kullanıldığı modellerden oluşur.

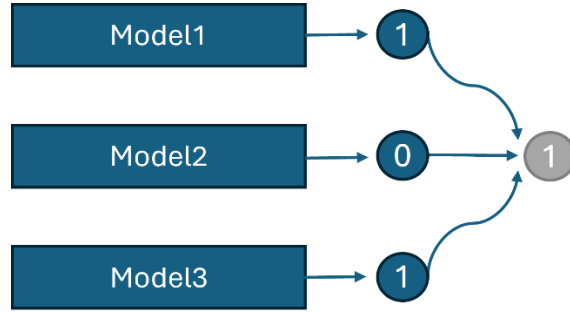
Topluluk yöntemleri arasında yaygın olarak kullanılan stratejilerden biri oylama yöntemidir. Oylama, birden fazla modelin tahminlerini birleştirerek nihai bir karar verir. Bu yöntem, çoğunluk oylaması ve ağırlıklı oylama üzere genellikle iki şekilde uygulanır. Çoğunluk oylaması, modellerin yaptığı tahminler arasından en fazla oy alan tahmini seçerken, ağırlıklı oylama her modele belirli bir ağırlık atayarak tahminlerin etkisini modelin başarımına göre ölçeklendirir (Sarmah, Borah, & Bhattacharyya, 2024).

Oylama yönteminin tercih edilmesinin başlıca nedenleri arasında, uygulanmasının nispeten basit ve doğrudan olması ile güçlü genellenebilirlik sağlaması yer alır. Oylama, özellikle farklı modellerin bağımsız olarak çalıştığı durumlarda etkili bir stratejidir. Bu yöntem, belirli bir modelin eksik kaldığı durumları diğer modellerin telafi etmesine olanak tanır, böylece daha dengeli ve güvenilir sonuçlar elde edilir.

Arařtırmalar, topluluk yöntemlerinin bireysel modellerin sınırlarını aşarak daha tutarlı ve genellenebilir sonuçlar sunduğunu göstermiştir. Bu nedenle, topluluk öğrenmesi hem akademik hem de endüstriyel alanlarda geniş bir uygulama yelpazesıyla modern yapay zeka sistemlerinin ayrılmaz bir parçası haline gelmiştir (Mienye & Sun, 2022).

3.2.6.1. Çoğunluk oylaması

Çoğunluk oylaması, birden fazla modelin tahminlerini birleştirerek bir karara varmayı hedefler. Bu yöntemde, modellerin her birinden alınan tahminlerin sonucunda en fazla oy alan sınıf seçilir. Çoğunluk oylaması, sert oylama ve yumuşak oylama olarak ikiye ayrılır. Sert oylama, her bir modelin yaptığı sınıf tahminine bakar ve en çok tahmin edilen sınıf yeni etiket olarak seçilir. Sert oylamanın temel avantajı, uygulamasının basit ve hesaplama açısından verimli olmasıdır. Ancak, modellerin sınıf tahmin olasılıklarını dikkate almaması nedeniyle güvenilirlik düzeyleri arasında bir ayrım yapmaz. Şekil 3.11’de, örnek olarak üç farklı modelin (Model 1, Model 2, Model 3) tahmin süreci gösterilmektedir; Model 1 ve Model 3, sınıf 1 tahmini yapmıştır. Model 2, sınıf 0 tahmini yapmıştır. Topluluk yöntemi, bu tahminler arasında çoğunluğu belirleyerek en fazla oy alan sınıf 1’i nihai tahmin olarak seçmiştir.



Şekil 3.11. Sert oylama

Matematiksel olarak sert oylama şu şekilde ifade edilir:

$$y_{final} = \arg \max_c \sum_{i=1}^n \delta(y_i, c) \quad (3.1)$$

Yumuşak oylama modellerin her bir sınıf için ürettiği olasılık değerlerini kullanarak karar verilen bir yöntemdir. Olasılık değerleri birleştirilir ve en yüksek toplam olasılığa sahip sınıf etiketi seçilir. Bu yöntem, modellerin tahminlerindeki belirsizliği yumuşatarak genellenebilirliği artırır.

Şekil 3.12’de, üç farklı modelin (Model 1, Model 2, Model 3) yumuşak oylama yöntemiyle nasıl birleştirildiği gösterilmektedir.

Model 1, sınıf 0 için 0.1 ve sınıf 1 için 0.9 olasılık tahmini yapmıştır.

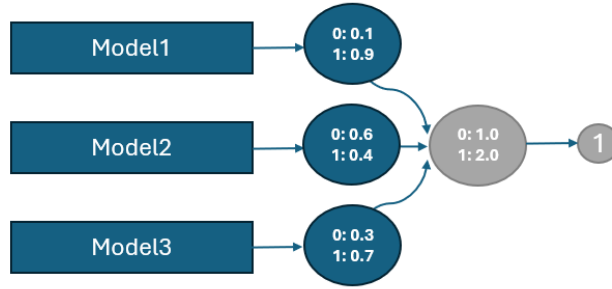
Model 2, sınıf 0 için 0.6 ve sınıf 1 için 0.4 olasılık tahmini yapmıştır.

Model 3, sınıf 0 için 0.3 ve sınıf 1 için 0.7 olasılık tahmini yapmıştır.

Her sınıf için toplam olasılıklar şu şekilde hesaplanmıştır:

- Sınıf 0 için toplam olasılık: $P(0)=0.1+0.6+0.3=1.0$
- Sınıf 1 için toplam olasılık: $P(1)=0.9+0.4+0.7=2.0$

Toplam olasılıkların karşılaştırılması sonucunda, sınıf 1’in toplam olasılığı daha yüksek olduğundan yeni tahmin 1 olarak belirlenmiştir.



Şekil 3.12. Yumuşak oylama

Matematiksel olarak yumuşak oylama şu şekilde ifade edilir:

$$y_{final} = \arg \max_c \sum_{i=1}^n P_i(c) \quad (3.2)$$

3.2.6.2. Ağırlıklı oylama

Ağırlıklı oylama, modellerin tahmin ettiği olasılıkların veya sınıf etiketlerinin, her modele atanan bir ağırlık faktörü ile çarpılarak birleştirilmesini sağlar. Her sınıf için ağırlıklandırılmış olasılıklar toplanır ve en yüksek toplam ağırlıklı olasılığa sahip sınıf seçilir. Ağırlıklar, genellikle modellerin performansına göre bir orantıyla belirlenir. Sonuçta, daha güvenilir modele daha yüksek bir ağırlık atanarak tahminlere katkısı artırılır. Bu durumda modellerin farklı başarı oranlarına göre etkisini ayarlayarak daha hassas tahminler elde edilir ve daha az performans gösteren modellerin tahmine etkisini azaltmış, daha güçlü modellerin etkisini arttırmış olur.

Şekil 3.13'te, üç farklı modelin (Model 1, Model 2, Model 3) tahmin olasılıkları, ağırlıklandırma yöntemiyle nasıl birleştirildiği gösterilmektedir:

Model 1, sınıf 0 için 0.1; sınıf 1 için 0.9 olasılık tahmini yapmıştır.

Model 2, sınıf 0 için 0.6; sınıf 1 için 0.4 olasılık tahmini yapmıştır.

Model 3, sınıf 0 için 0.3; sınıf 1 için 0.7 olasılık tahmini yapmıştır.

Her sınıf için toplam ağırlıklı olasılıklar şu şekilde hesaplanmıştır:

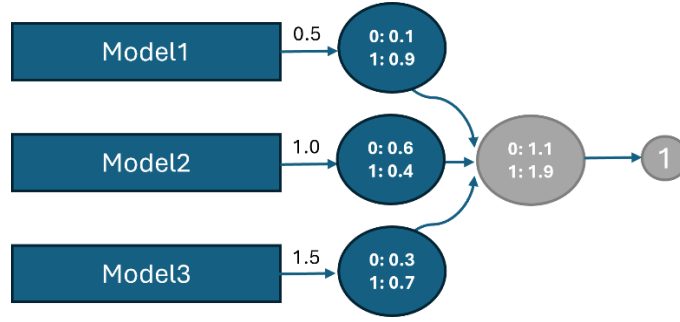
Sınıf 0 için toplam ağırlıklı olasılık:

$$P(0)=(0.5 \times 0.1)+(1.0 \times 0.6)+(1.5 \times 0.3)=0.05+0.6+0.45=1.1$$

Sınıf 1 için toplam ağırlıklı olasılık:

$$P(1)=(0.5 \times 0.9)+(1.0 \times 0.4)+(1.5 \times 0.7)=0.45+0.4+1.05=1.9$$

Toplam ağırlıklı olasılıkların karşılaştırılması sonucunda, sınıf 1'in toplam ağırlıklı olasılığı daha yüksek olduğundan yeni tahmin 1 olarak belirlenmiştir.



Şekil 3.13. Ağırlıklı oylama

Matematiksel olarak ağırlıklı oylama şu şekilde ifade edilir:

$$y_{final} = \arg \max_c \sum_{i=1}^n w_i \cdot P_i(c) \quad (3.3)$$

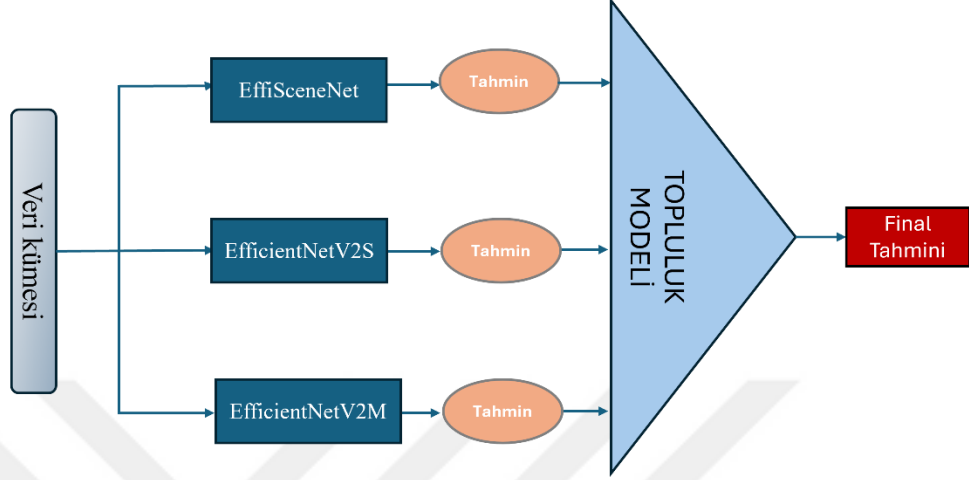
3.2.6.3. Topluluk öğrenmesinin uygulanması

Bu çalışmada EffiSceneNet, EfficientNetV2S ve EfficientNetV2M olmak üzere üç farklı derin öğrenme modeli kullanılarak topluluk öğrenme modeli geliştirilmiştir. Topluluk modellerinin oluşturulmasında sert oylama, yumuşak oylama ve ağırlıklı oylama olmak üzere üç farklı topluluk öğrenme stratejisi uygulanmıştır. Şekil 3.14'te modelleme süreci gösterilmiştir.

Çalışma prensibi;

- Veri kümesi her bir model için işlenmiş ve eğitilmiş modeller kaydedilmiştir. Modeller, farklı mimariler ve öğrenme kapasitelerine sahip olduğundan, veri kümesindeki sınıflar arasında çeşitli öznelikleri öğrenmişlerdir.
- Her bir model, sınıf bazında tahminler yapmış ve bu tahminler topluluk modeline girdi olarak sunulmuştur. Bireysel modellerin tahminleri, topluluk modeli tarafından birleştirilmiştir. Bu süreç, sert oylama, yumuşak oylama ve ağırlıklı oylama gibi farklı birleştirme yöntemleriyle gerçekleştirilmiştir.
- Topluluk modeli, bireysel modellerin güçlü yönlerini bir araya getirerek genel tahmin doğruluğunu artırmayı hedeflemiştir.

- Topluluk modelinin çıktısı, birleştirilmiş tahminler sonucunda oluşan nihai tahmin final tahmini olarak belirlenmiştir.



Şekil 3.14. Topluluk modeli

Topluluk modeli, bu çalışmada bireysel modellerin performansını artırmak ve daha sağlam tahminler yapmak amacıyla kullanılmıştır. Sert oylama yöntemi, basit bir çoğunluk kuralını izlerken; yumuşak oylama, olasılık dağılımlarını dikkate almış ve daha sofistike bir değerlendirme sağlamıştır. Ağırlıklı oylama ise modellerin güvenilirlik düzeylerini hesaba katarak tahmin sürecini optimize etmiştir.

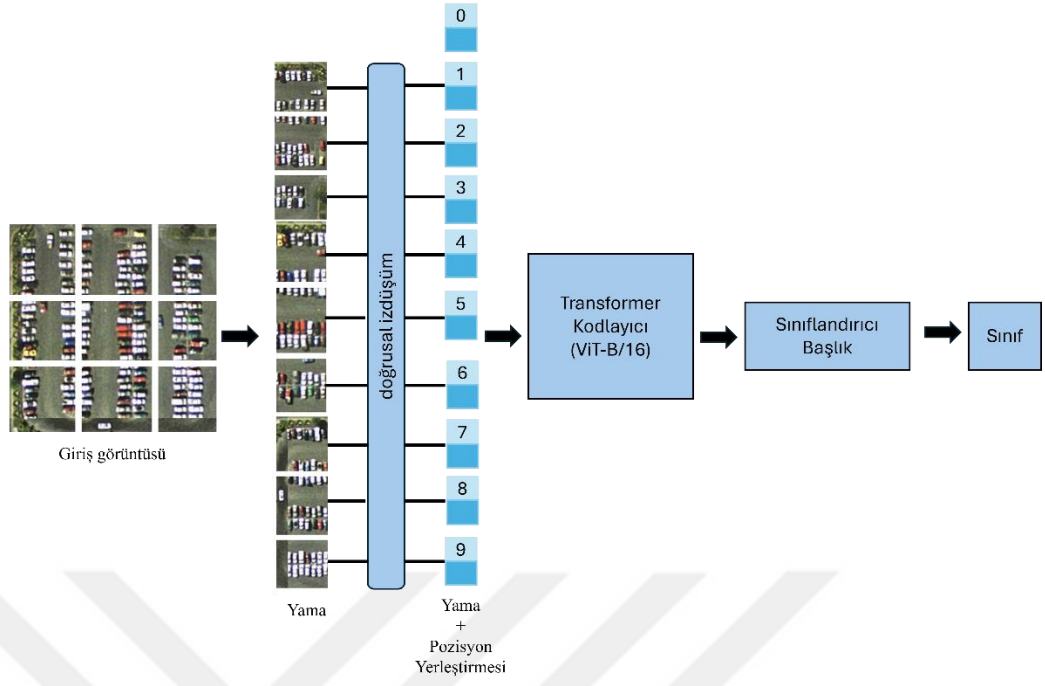
3.2.7. Görsel dönüştürücü

ViT, görüntü işleme görevleri için derin öğrenme alanında yenilikçi bir yaklaşım sunan bir yöntemdir. Transformer mimarisine dayanan ViT, doğal dil işleme modellerinde yaygın olarak kullanılan dikkat mekanizmasını görsel verilere uyarlayarak çalışır. ViT, büyük ölçekli veri kümelerinde üstün başarı sağlayarak modelin küçük görüntü yamalarını giriş olarak kullanır ve saf Transformer tabanlı bir yaklaşımı CNN'le yarışır hale getirmiştir (Dosovitskiy, ve diğerleri, 2020).

ViT'in dikkat mekanizması, giriş verisindeki her bir piksel grubunun diğer tüm öğelerle olan bağımlı öğrenir ve bu ilişkilerden önemli olanları vurgular. Bu mekanizma, modelin tüm bağlamı anlamasına ve uzun menzilli bağımlılıkları etkili bir şekilde öğrenmesine olanak tanır. ViT, giriş görüntüsünü sabit boyutlu parçalara (patch) böler ve her bir parçayı bir dizi öğe olarak işler. Bu parçalar, pozisyonel kodlama ile modele

aktarılır ve dikkat mekanizması katmanları sayesinde analiz edilir. Geleneksel evrişimli sinir ağlarının yerel bilgiye odaklanmasına karşın ViT, tüm bağlamın daha etkili bir şekilde yakalanmasını sağlar ve özellikle büyük ölçekli veri kümeleriyle eğitildiğinde yüksek performans sergiler.

Şekil 3.15'te ViT'in temel olarak kurulan dört ana bileşenin mimarisi gösterilmiştir. İlk olarak, giriş ve parçalama (input and patch embedding) aşamasında, giriş görüntüsü sabit boyutlu küçük parçalara (patch) bölünür ve her bir parça doğrusal bir katman kullanılarak sabit boyutlu bir vektöre dönüştürülür. Bu vektörlere pozisyonel kodlama eklenerek, parçaların mekânsal düzenlemelerinin modele iletilmesi sağlanır. İkinci bileşen olan transformer katmanları, dikkat mekanizması sayesinde her bir parçanın diğer tüm parçalarla olan ilişkisini öğrenir ve küresel bağlamı anlama yeteneği sağlar. Bu katmanlarda normalize ve besleme ileri ağı, katman normalizasyonu ve kalıntı bağlantıları gibi yapılar bulunur. Bu çalışmada Transformer kodlayıcı olarak önceden eğitilmiş bir ViT modeli olan vit base 16 kullanılmıştır. Bu model, büyük ölçekli görüntü veri kümeleri üzerinde önceden eğitilmiş bir temel yapıya sahiptir ve transfer öğrenimi yöntemi ile özelleştirilmiştir. Modelin parametreleri transfer öğrenimi yöntemine uygun olacak şekilde düzenlenmiş ve tüm temel katman parametreleri güncellenebilir hale getirilmiştir. Patch size 16x16 olarak ayarlanmıştır. Kodlayıcıda gizli boyutu 768 olan 12 katman ve dikkat katmanında ise 12 başlık kullanılmıştır. Son olarak, sınıflandırıcı başlık, transformer katmanlarından gelen çıktıyı tam bağlantılı bir katman üzerinden geçirir ve sınıflandırma yapılacak sınıf sayısına göre özelleştirilir.



Şekil 3.15. ViT'in mimarisi

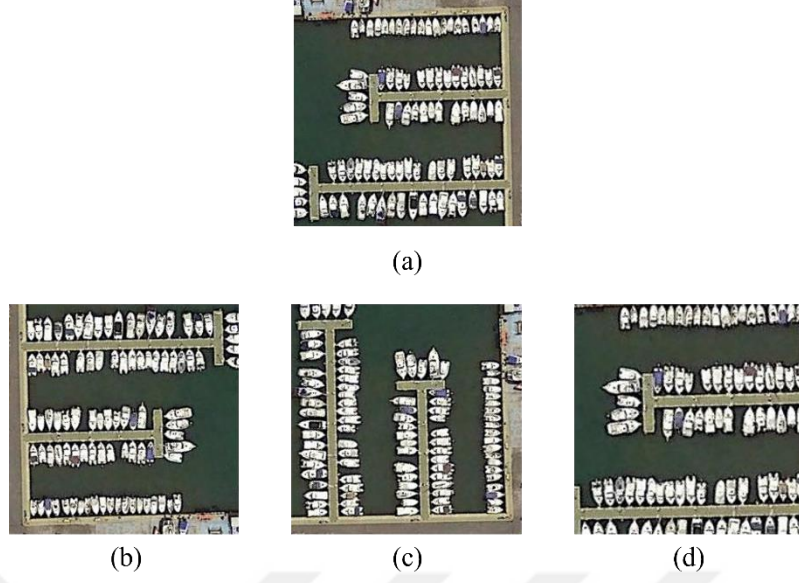
4. DENEY SONUÇLARI

Bu bölümde, önerilen yöntemin sınıflandırma performansı, modelin genelleştirilebilirliğini değerlendirmek amacıyla NWPU-RESISC45, AID ve UCM veri kümeleri üzerinde test edilmiş ve elde edilen sonuçlar özetlenmiştir. Tüm veri kümelerinde elde edilen doğruluk ve kayıp grafikleri incelendiğinde, eğitim ve doğrulama verileri arasında anlamlı bir fark olmadığı ve doğrulama oranının oldukça yüksek olduğu gözlemlenmiştir. Bu durum modelin genelleştirilebilir bir yapıya sahip olduğunu ortaya koymaktadır. Ayrıca, üç veri kümesi kapsamında, elde edilen sonuçlar literatürdeki en güncel yöntemlerle karşılaştırılmıştır.

4.1. Veri Hazırlama ve Arttırma

Veri hazırlama, görüntülerin gerekli girdi boyutuna uyacak şekilde yeniden boyutlandırılmasını içerir. AID veri kümesindeki tüm görüntülerin, eğitim aşamasında bellek taşması hatası alındığından dolayı orijinal 600×600 piksel boyutundan 224×224 piksele yeniden boyutlandırılmıştır. UCM ve NWPU-RESISC45 veri kümelerinde de 256×256 piksel boyutundan 224×224 piksele yeniden boyutlandırılarak çalışılmıştır. Yapılan deneylerde, veri kümeleri ilk aşamada %80-%20 olmak üzere eğitim/test olarak bölünmüştür. Eğitim ve test kümesi dışında modelin her epoch sonunda doğrulama (validation) performansını değerlendirmek için kullanılan doğrulama seti de mevcuttur. Eğitim için ayrılan setten %20 oranında da doğrulama seti için bölünmüştür.

Veri artırma, eğitim veri kümesinin çeşitliliğini artırmak, modelin genelleme yeteneğini iyileştirmek ve aşırı öğrenmeyi (overfitting) önlemek amacıyla kullanılmıştır. Bu çalışmada, eğitim verileri üzerinde belirli oranlarda döndürme, kaydırma, yakınlaştırma ve yansıma ayarlamaları yapılmıştır. Orijinal görüntü ve uygulanan teknikler ile ortaya çıkan görüntülere örnek Şekil 4.1'de gösterilmiştir. Bir görüntü üreticisi, veri kümesindeki görüntüleri rastgele dönüşümlerle değiştirerek yeni, benzer görüntüler oluşturmuştur. Yeni oluşan görüntüler veri kümesini fiziksel olarak büyütmemiştir. Sadece eğitim sırasında her bir batch için farklı veri arttırmalar uygulayarak görüntüleri çeşitlendirmiş ve o an kullanmıştır. Bu çalışmada iki farklı görüntü üreticisi oluşturulmuş ve her birine farklı teknikler uygulanmıştır; ardından bu üretilen veriler birleştirilmiştir. Bu kombinasyon, modele farklı perspektiflerden daha zengin ve çeşitli bir eğitim verisi sağlamıştır.



Şekil 4.1. Artırılmış görüntü örnekleri. (a) Orijinal, (b) Yansıma, (c) Döndürme, (d) Yakınlaştırma

Veri arttırımı sonrasında veri kümelerinin içerdiği toplam görüntü sayıları ile arttırılmış görüntü sayıları Tablo 4.1’de sunulmaktadır. UCM ve AID veri kümelerinin eğitim parçasında her sınıf için 6 kat artırım yapılmıştır. NWPU-RESISC45 veri kümesinin eğitim parçasında ise her sınıf için 3 kat artırım yapılmıştır.

Tablo 4.1. Orijinal ve arttırılmış veri kümelerindeki görüntü sayıları

Veri Kümesi	Veri Parçası	Orijinal	Arttırılmış
UCM	Eğitim	1344	8064
UCM	Doğrulama	336	2016
UCM	Test	420	420
AID	Eğitim	6400	38400
AID	Doğrulama	1600	9600
AID	Test	2000	2000
NWPU-RESISC45	Eğitim	20160	60480
NWPU-RESISC45	Doğrulama	5040	15120
NWPU-RESISC45	Test	6300	6300

4.2. Eğitim Detayları

Bu çalışmada, önerilen yöntemi uygulamak için TensorFlow derin öğrenme uygulama çerçevesi kullanılmıştır. Eğitim aşamasında kullanılan hiper parametreler veri kümelerinin özniteliklerine göre analiz ederek belirlenmiştir.

Modelin derlenmesi sırasında Adam optimize kullanılmış ve öğrenme oranı 0.001 olarak belirlenmiştir. Kategorik çapraz entropi kayıp fonksiyonu kullanılarak model derlenmiş ve modelin performansı doğruluk (accuracy) metriği ile değerlendirilmiştir. Modelin derlenmesi sırasında bir geri çağırma (callback) fonksiyonu olan erken durdurma (early stopping) kullanılmıştır. Bu fonksiyon modelin doğrulama (validation) kaybını veya doğrulama başarımını izleyerek, belirli bir süre boyunca iyileşme gözlenmezse eğitim sürecini durdurmuştur. Bu, modelin aşırı öğrenme (overfitting) yapmasını engelleyerek ve eğitim süresini kısaltmıştır. Önerilen yöntem de erken durdurma kullanımı, modelin doğrulama kaybı üzerinden izleme yaparak, iyileşme gözlenmediğinde eğitimi durdurarak çeşitli faydalar sağlamıştır;

- Daha kısa eğitim süreleri ile hesaplama kaynaklarından ve enerjiden tasarruf edilmiştir.
- Modelin en iyi performansı gösterdiği noktada eğitimi durdurduğundan dolayı, doğrulama verileri üzerindeki performans en iyi şekilde korunmuştur.
- Model, erken durdurma sayesinde gereksiz yere eğitilmemiş. Bu, eğitim süresinin kısaltılmasını sağlamıştır.

4.3. NWPU-RESISC45 Veri Kümesi

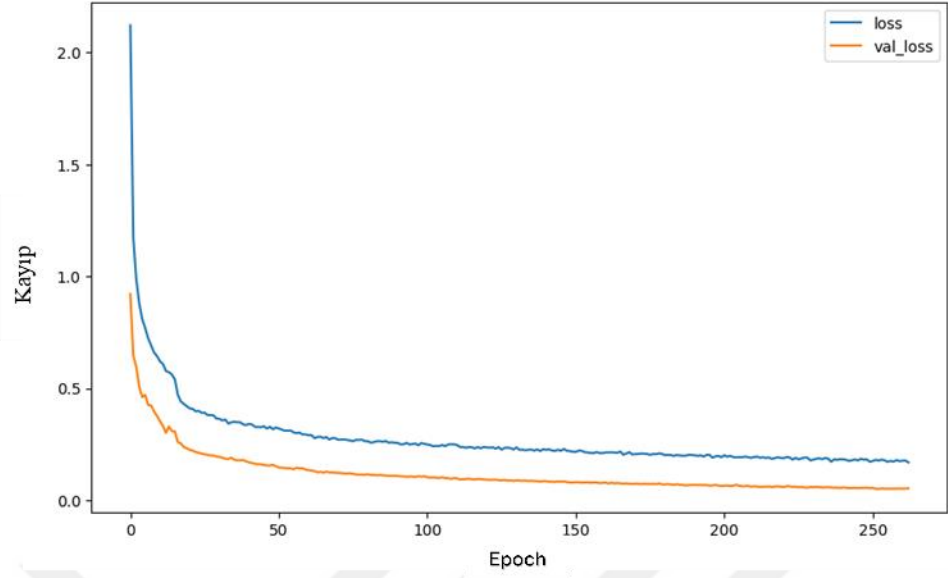
Tablo 4.2'de NWPU-RESISC45 veri kümesi üzerinde gerçekleştirilen deneylerin sınıflandırma başarısı, kesinlik (precision), duyarlılık (recall) ve F-1 skoru sonuçları özetlenmektedir. EfficientNetV2S ve EfficientNetV2M modelleri sırasıyla %91,54 ve %89,24 sınıflandırma başarısı gösterirken, geliştirilen EffiSceneNet modeli %98,02'lik sınıflandırma başarısıyla belirgin bir performans artışı sağlamıştır. Bu sonuç, önerilen modelin uzaktan algılama görüntülerinde daha etkin özellik çıkarımı gerçekleştirdiğini ortaya koymaktadır. Ayrıca, kesinlik ve duyarlılık değerleri de 0,98'in üzerinde olarak yüksek seviyede tutarlı bir performans sergilediğini göstermektedir. ViT modeli ise %94,68 sınıflandırma başarısı ile EffiSceneNet'in gerisinde kalmış, ancak yine de EfficientNetV2M modelinden daha iyi sonuçlar elde etmiştir. Farklı oylama stratejileriyle yapılan topluluk yöntemleri, EffiSceneNet modeline benzer şekilde %97'nin üzerinde

sınıflandırma başarısı sağlamış olup, özellikle ağırlıklı oylama yöntemi %97,44 ile en yüksek performansa yakın bir sonuç vermiştir. Tüm modeller arasında F-1 skoru değerlerinin 0,8915 ile 0,9806 arasında değiştiği görülmekte, bu da modellerin dengeli bir şekilde hassasiyet ve geri çağırma metriklerinde başarılı olduğunu göstermektedir. Sonuçlar genel olarak EffiSceneNet'in ve oylama tabanlı topluluk modellerinin, geleneksel EfficientNet ve ViT modellerine kıyasla daha üstün performans sergilediğini ortaya koymaktadır.

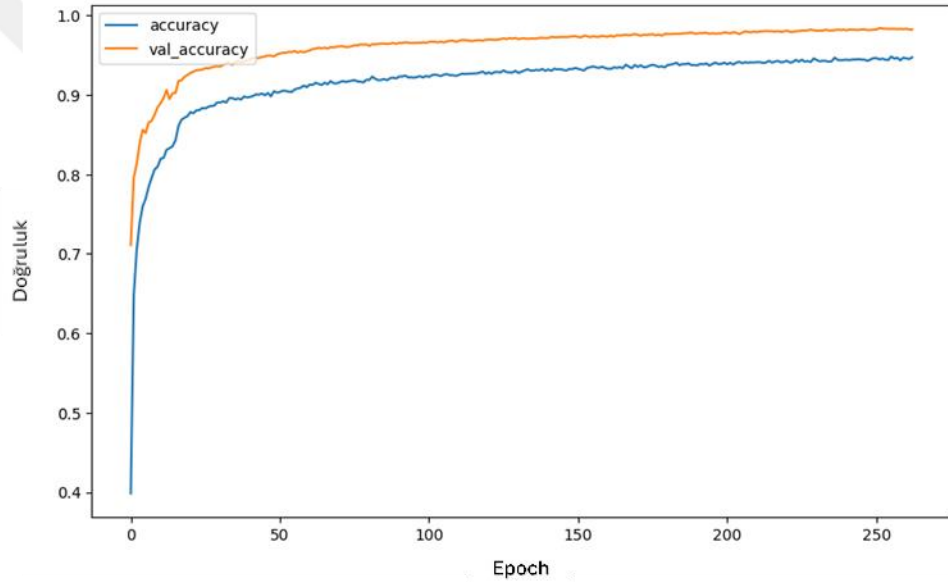
Tablo 4.2. *NWPU-RESISC45 veri kümesinin sonuçları*

Model	Sınıflandırma başarısı (%)	Kesinlik (precision)	Duyarlılık (recall)	F-1 skor
EfficientNetV2S	%91,54	0.9103	0.9112	0.9107
EfficientNetV2M	%89,24	0.8921	0.8909	0.8915
EffiSceneNet	%98,02	0.9801	0.9805	0.9803
ViT	%94,68	0.9501	0.9499	0.9500
Sert oylama	%97,32	0.9804	0.9806	0.9805
Yumuşak oylama	%97,42	0.9804	0.9807	0.9806
Ağırlıklı oylama	%97,44	0.9805	0.9808	0.9806

Şekil 4.2 her epoch için doğruluk ve kayıp grafiklerini göstermektedir. Grafikler her 50 yineleme için sabit bir temel öğrenme oranıyla gerçekleştirilen eğitim sürecini özetlemektedir. Model, uzun eğitim süresine rağmen oldukça başarılı bir performans sergilemiştir. Düşük kayıp değerleri ve doğruluk oranlarındaki kararlılık, modelin verileri başarıyla öğrendiğini ve güçlü bir performansa sahip olduğunu gösterir.



(a)



(b)

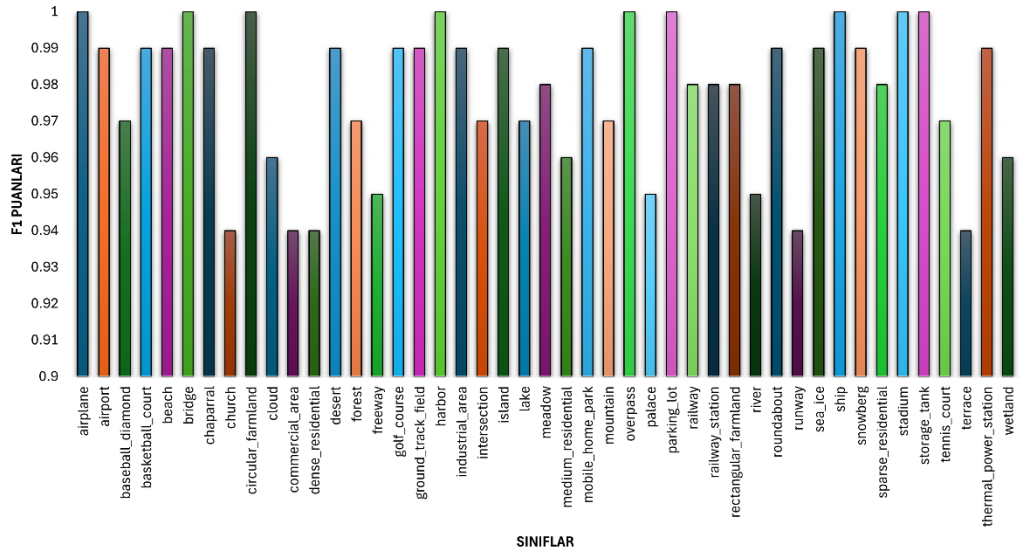
Şekil 4.2. NWPU-RESISC45'nin eğitim ve doğrulama değerleri. (a) Kayıp grafiği, (b) Doğruluk grafiği

Daha ileri analize bakıldığında, bu çalışma sonucunda NWPU-RESISC45 veri kümesinin tüm sınıflarında test edilen görüntü, çoğunlukla doğru sınıfla eşleşmiştir. Ancak, veri kümesi çok sınıflı olduğundan ve sınıflar benzer yapılar içerdiğinden, benzer görüntü dağılımlarına sahip sınıflar arasında yanlış tahminler meydana gelmiştir. Özellikle yoğun yerleşim (dense residential) ile ortalama yerleşim (medium residential) sınıfları arasında belirgin bir karışıklık gözlemlenmektedir. Bu sınıflar mekânsal yoğunluk ve yapı özellikleri bakımından benzer görüntülere sahiplerdir. Bundan dolayı

her iki sınıfın görsel benzerlikler taşıması ve uzaktan algılama görüntülerinde yapı yoğunluğu ile desen benzerliklerinin model tarafından ayırt edilmesinde zorluk yaşanmasına sebep olmuştur. Bunun yanı sıra, nehir (river) ve göl (lake) sınıfları da sınırlı sayıda yanlış sınıflandırma ile birbirine karıştırılmıştır. Bu durum, su kütlelerinin uzaktan algılama görüntülerindeki görsel benzerliklerinden dolayı modelin bu iki sınıfı ayırt etmede zaman zaman güçlük çektiğini göstermektedir.

Genel olarak, EffiSceneNet'in sınıflandırma performansı oldukça başarılıdır ve yüksek kesinlik, duyarlılık ve F-1 skoru değerleriyle desteklenmektedir. Ancak, benzer görsel özelliklere sahip sınıflar arasında yaşanan karışıklıklar, modelin daha derin özellik çıkarımı yapacak şekilde optimize edilmesiyle veya veri ön işleme adımlarında iyileştirmelerle azaltılabilir. Bu tür iyileştirmeler, modelin karmaşık sınıflar arasındaki ayırt edici özellikleri daha iyi öğrenmesini sağlayacaktır.

Şekil 4.3, EffiSceneNet modelinin NWPU-RESISC45 veri kümesindeki sınıfların F1 puanlarını gösterir. Uçak (airplane), köprü (bridge), dairesel tarım arazisi (circular farmland), liman (harbor), üst geçit (overpass), otopark (parking lot), gemi (ship), stadyum (stadium) ve depolama tank (storage tank) sınıfları için doğruluk oranı %100'e ulaşırken diğer tüm sınıflar %94'ün üzerindedir.



Şekil 4.3. NWPU-RESISC45 veri kümesi sınıflarının F-1 skorları

Tablo 4.3'te, geliştirilen EffiSceneNet modelinin NWPU-RESISC45 veri kümesi üzerinde literatürde yer alan güncel yöntemlerle karşılaştırmalı performans sonuçları sunulmaktadır. Daha önce önerilmiş DNE, ADSSM, Siamese ResNet50, SCCOV, CapsNet, Hydra, ADFE, EfficientNet, SAFF, H-GCN, AC Net ve CNN gibi çeşitli derin öğrenme tabanlı yaklaşımlar ile elde edilen sınıflandırma başarıları %87,86 ile %96,19 arasında değişmektedir. Ancak, önerilen EffiSceneNet modeli %98,02'lik sınıflandırma başarısıyla mevcut en iyi sonucu elde ederek literatürdeki tüm modelleri geride bırakmıştır. Ayrıca, önerilen topluluk öğrenme yöntemi olan ağırlıklı oylama stratejisi %97,44'lük bir başarı sağlamış ve EffiSceneNet'in performansına oldukça yakın sonuçlar vermiştir. Bu durum, topluluk yoluyla elde edilen model birleştirme stratejisinin, sınıflandırma doğruluğunu artırmada etkili olduğunu göstermektedir. Sonuçlar, EffiSceneNet'in yalnızca temel model performansında değil, aynı zamanda topluluk öğrenme yaklaşımlarıyla da üstün bir genel başarı sağladığını ortaya koymaktadır. Bu bulgular, geliştirilen modelin uzaktan algılama görüntü sınıflandırma alanında güçlü bir potansiyele sahip olduğunu desteklemektedir. NWPU-RESISC45 veri kümesi, kapsamlı sınıf sayısı ve sınıf başına önemli miktarda görüntü olması nedeniyle en zorlu veri kümelerinden biri olarak kabul edilir. Bu sonuçlar önerilen modelin karmaşık sınıflandırma görevlerini ele almadaki sağlamlığını ve verimliliğini vurgulamıştır.

Tablo 4.3. *EffiSceneNet*'in *NWPU-RESISC45* üzerinde son teknoloji performansla karşılaştırılması

Yazar	Algoritma	Sonuç
Dede vd. (2018)	DNE	%96,01
Zhu vd. (2018)	ADSSM	%94,29
Liu vd. (2019)	Siamese ResNet50	%92,28
He vd. (2019)	SCCOV	%92,10
Zhang vd. (2019)	CapsNet	%92,6
Minetto vd. (2019)	Hydra	%94,51
ZHU vd. (2019)	ADFF	%91,91
Zhao vd. (2020)	CNN	%94,21
Zhang vd. (2020)	EfficientNet	%96,19
Cao vd. (2020)	SAFF	%87,86
Gao vd. (2021)	H-GCN	%93,62
Tang vd. (2021)	AC Net	%92,42
Wang vd. (2023)	CNN	%94,57
Song (2023)	FST-EfficientNetV2	%95,75
Keyan vd. (2024)	RSMamba-H	%95,22
Önerilen yöntem	EffiSceneNet	%98,02
Önerilen topluluk öğrenmesi	Ağırlıklı oylama	%97,44

4.4. UCM Veri Kümesi

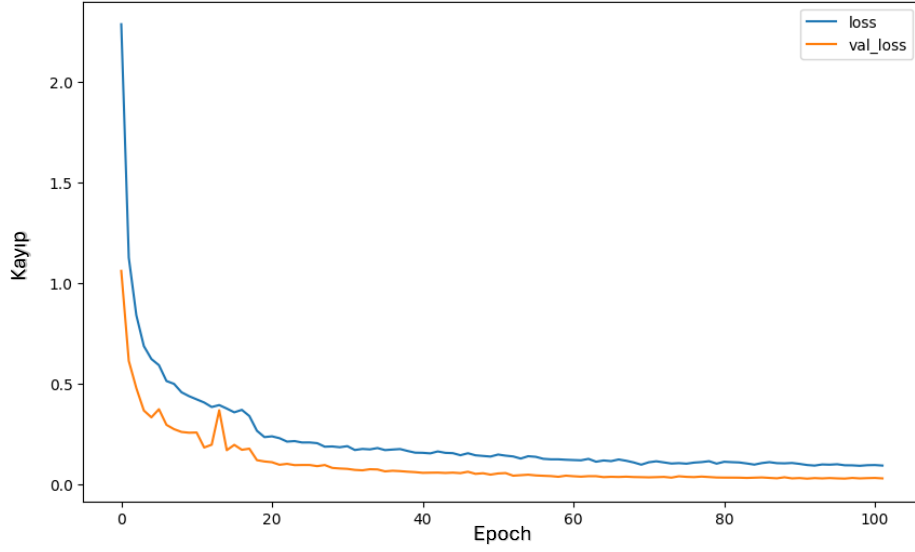
Tablo 4.4'te, UCM veri kümesi üzerinde gerçekleştirilen deneylerin sınıflandırma başarısı, kesinlik, duyarlılık ve F-1 skoru açısından karşılaştırmalı sonuçları sunulmaktadır. EfficientNetV2S ve EfficientNetV2M modelleri sırasıyla %95,92 ve %95,00 sınıflandırma başarıları ile istikrarlı bir performans göstermiştir. Ancak, geliştirilen EffiSceneNet modeli %99,69'luk sınıflandırma başarısı ile dikkat çekici bir performans sergileyerek bu modelleri önemli ölçüde geride bırakmıştır. EffiSceneNet'in kesinlik, duyarlılık ve F-1 skoru değerlerinin tümü 1,00' e çok yakın değerler olarak ölçülmüş, bu da modelin sınıflandırma görevinde mükemmel yakın bir doğrulukla çalıştığını ortaya koymaktadır. Sert, yumuşak ve ağırlıklı oylama stratejileri de yüksek performans göstererek sırasıyla %97,95, %98,22 ve %98,39 doğruluk oranları elde etmiştir. Bu oylama yöntemleri, sınıflandırma başarısını artırmada etkili olsa da EffiSceneNet'in bireysel performansının altında kalmıştır. Öte yandan, ViT modeli %99,52 sınıflandırma başarısı ile oldukça yüksek bir performans göstermiştir, ancak EffiSceneNet ile kıyaslandığında küçük bir farkla geride kalmıştır. Sonuçlar,

EffiSceneNet'in UCM veri kümesinde hem kesinlik hem de duyarlılık açısından tutarlı ve üstün bir performans sergilediğini ve uzaktan algılama görüntülerinin sınıflandırılmasında güçlü bir alternatif sunduğunu göstermektedir.

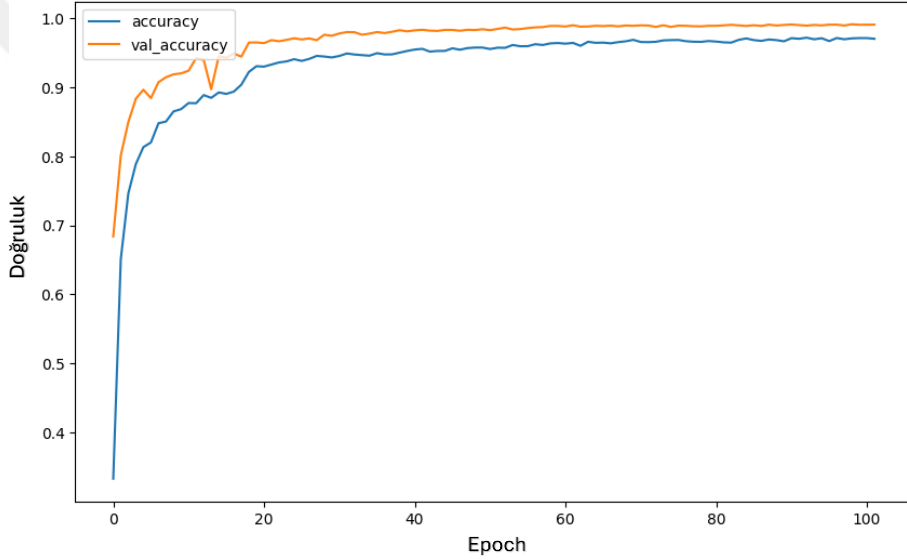
Tablo 4.4. UCM veri kümesinin sonuçları

Model	Sınıflandırma başarısu (%)	Kesinlik (precision)	Duyarlılık (recall)	F-1 skor
EfficientNetV2S	%95,92	0,9321	0,9235	0,9278
EfficientNetV2M	%95,00	0,9314	0,9287	0,9300
EffiSceneNet	%99,69	0,9996	0,9998	0,9997
Sert oylama	%97,95	0,9632	0,9617	0,9624
Yumuşak oylama	%98,22	0,9645	0,9654	0,9648
Ağırlıklı oylama	%98,39	0,9711	0,9674	0,9692
ViT	%99,52	0,9989	0,9995	0,9992

Şekil 4.4, her epoch için doğruluk ve kayıp grafiklerini göstermektedir. Grafikler, her 10 yinelemede sabit bir temel öğrenme oranı ile gerçekleştirilen eğitim sürecini özetlemektedir. Grafikler incelendiğinde; kayıp değerlerindeki kademeli düşüş ve doğruluk oranlarındaki ilk artış, modelin öğrenme sürecinin başarılı olduğunu göstermektedir. Eğitim ve doğrulama eğrilerinin birbirine yakın olması, modelde aşırı uyum riski olmadığını açıklamaktadır.



(a)

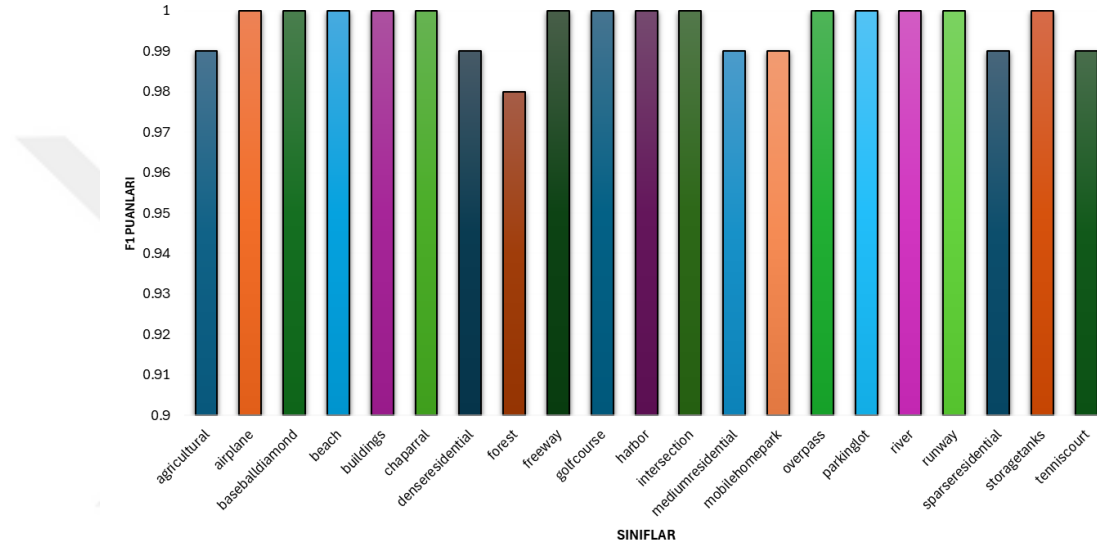


(b)

Şekil 4.4. UCM'nin eğitim ve doğrulama değerleri. (a) Kayıp grafiği, (b) Doğruluk grafiği

Yapılan sınıflandırma performansı analizi sonucunda, modelin genel doğruluk oranının %99,69 olduğu ve çoğu sınıfta yüksek başarı sağladığı gözlemlenmiştir. Ancak, belirli sınıflarda hataların yoğunlaştığı tespit edilmiştir. Özellikle orta konut (medium residential) ile yoğun konut (sparse residential) kategorilerinde modelin bazı örnekleri yanlış tahmin ettiği görülmüştür. Bu hataların temel nedeni, görüntülerin benzer görsel özelliklere sahip olmasıdır.

Şekil 4.5, EffiSceneNet modelinin UCM veri kümesindeki sınıfların F1 puanlarını göstermektedir. Doğruluk oranı, uçak (airplane), beyzbol sahası (baseball diamond), sahil (beach), binalar (buildings), çalılık (chaparral), otoban (freeway), golf sahası (golfcourse), liman (harbor), kavşak (intersection), üst geçit (overpass), otopark (parkinglot), nehir (river), pist (runway) ve depolama tankları (storage tanks) sınıfları için %100'e ulaşırken diğer tüm sınıflar %98'in üzerindedir.



Şekil 4.5. UCM veri kümesi sınıflarının F-1 skorları

Tablo 4.5'te, geliştirilen EffiSceneNet modelinin UCM veri kümesi üzerindeki performansı, literatürde önerilen çeşitli derin öğrenme tabanlı yöntemlerle karşılaştırılmaktadır. Önceki çalışmalarda kullanılan D-CNN, MCNN, 4-layer SVM, CapsNet, FACNN, SF-CNN, Transfer Learning, EfficientNet, AlexNet, VGG-19 ve ARCNet-VGG16 gibi algoritmaların sınıflandırma başarıları %93,48 ile %99,86 arasında değişmektedir. Ancak, önerilen EffiSceneNet modeli %99,69'luk sınıflandırma başarısı ile bu modele oldukça yakın bir performans sergilemiş ve literatürdeki pek çok yöntemi geride bırakmıştır. Ayrıca, önerilen topluluk öğrenmesi yöntemi olan ağırlıklı oylama stratejisi %98,39 doğruluk oranıyla yine yüksek bir başarı göstermiştir. EffiSceneNet'in üstün performansı, modelin uzaktan algılama görüntülerindeki karmaşık desenleri ve özellikleri etkili bir şekilde öğrenme yeteneğine sahip olduğunu göstermektedir. Literatürde yer alan diğer yöntemler arasında %99,50 ve üzeri sonuçlara ulaşan modeller

olmasına rağmen, EffiSceneNet'in bu modellerle rekabet edebilecek düzeyde bir doğruluk oranı yakalaması, modelin genelizabilite gücünü ortaya koymaktadır. Bu bulgular, önerilen modelin UCM veri kümesi üzerindeki etkinliğini kanıtlamakta ve uzaktan algılama görüntü sınıflandırma alanında önemli bir katkı sunduğunu göstermektedir.

Tablo 4.5. *EffiSceneNet'in UCM veri kümesi üzerinde son teknoloji performansla karşılaştırılması*

Yazar	Algoritma	Sonuç
Luus vd. (2015)	D-CNN	%93,48
Liu vd. (2018)	D-CNN	%98,93
Zhu vd. (2018)	MCNN	%99,76
Muhammad vd. (2018)	4-layer SVM	%99,52
Zhang vd. (2019)	CapsNet	%99,05
Xie vd. (2019)	FACNN	%99,05
Fang vd. (2019)	SF-CNN	%99,05
Ali vd. (2019)	CNN-CapsNet	%99,05
Biserka vd. (2020)	Transfer Learning	%98,33
Zhao vd. (2020)	CNN	%99,11
Özyurt vd. (2020)	AlexNet CNN	%98,8
Cao vd. (2020)	SAFF	%97,02
Zhang vd. (2020)	EfficientNet	%98,81
Gao vd. (2021)	H-GCN	%99
Tang vd. (2021)	AC Net	%99,76
Kareem vd. (2021)	AlexNet	%94,20
Bi vd. (2021)	Çok Ölçekli Hisse Senedi Dikkat Havuzu (MS2AP)	%99,09
Alhicri vd. (2021)	EfficientNet-B3-Attn-2	%99,21
Thirumaladevi vd. (2023)	Alex Net	%93,57
Thirumaladevi vd. (2023)	VGG-19	%94,08
Thirumaladevi vd. (2023)	VGG-16	%95
Wang vd. (2023)	CNN	%99,76
Wang vd. (2023)	ARCNet-VGG16	%99,12
Keyan vd. (2024)	RSMamba-H	%95,47
Datla vd. (2024)	Scene-vector	%99,88
Önerilen Yöntem	EffiSceneNet	%99,69
Önerilen topluluk öğrenmesi	Ağırlıklı oylama	%98,39

4.5. AID Veri Kümesi

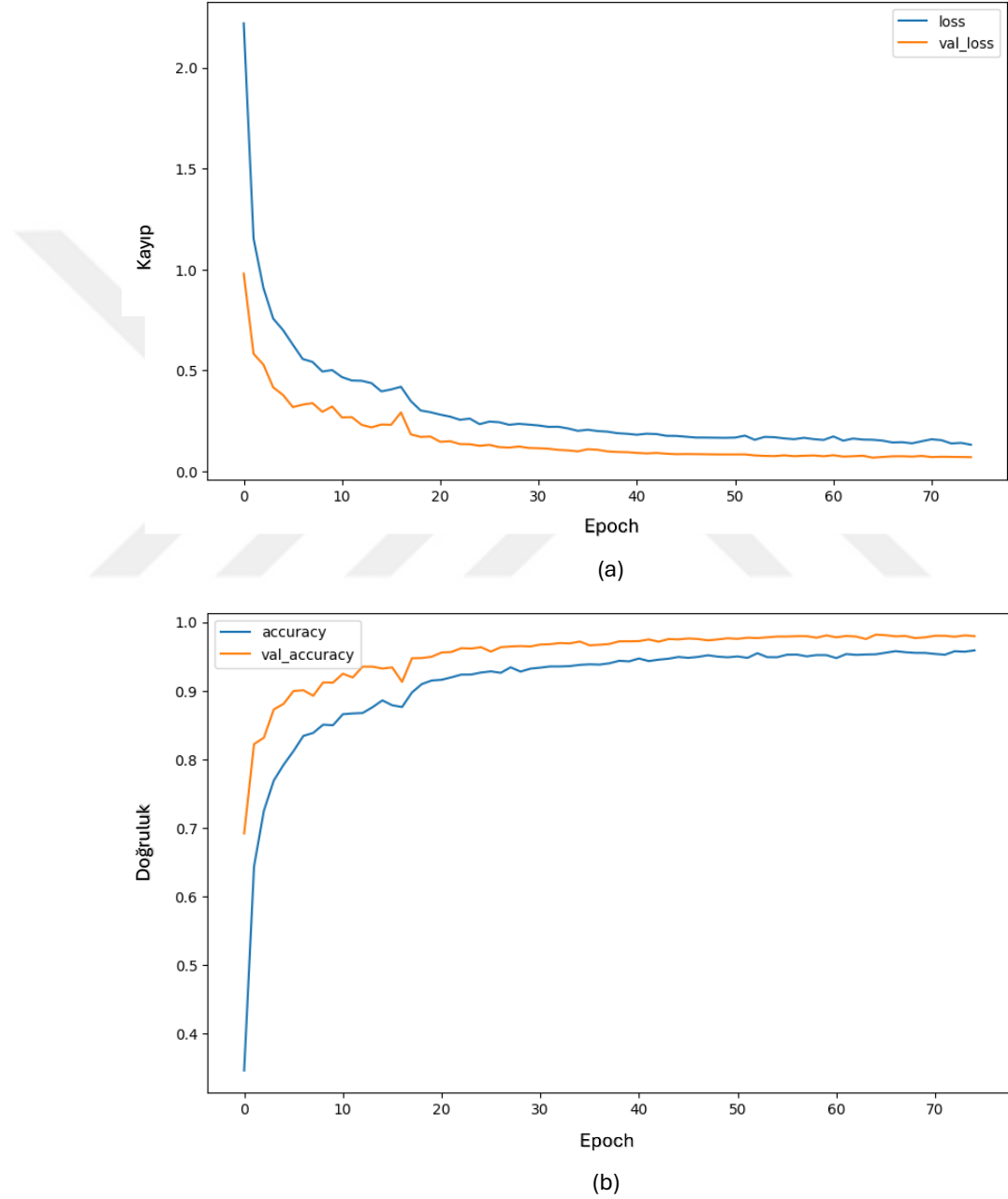
Tablo 4.6’da AID veri kümesi üzerinde gerçekleştirilen deneylerin sınıflandırma başarısı, kesinlik, duyarlılık ve F-1 skoru açısından karşılaştırmalı sonuçları sunulmaktadır. EfficientNetV2S ve EfficientNetV2M modelleri sırasıyla %92,39 ve %90,69 sınıflandırma başarıları ile makul düzeyde bir performans göstermiştir. Ancak, geliştirilen EffiSceneNet modeli %99,12’lik sınıflandırma başarısı ile dikkat çekici bir performans sergilemiş, bu modelleri önemli ölçüde geride bırakmıştır. Ayrıca kesinlik, duyarlılık ve F-1 skoru değerlerinin tümü 0,99’un üzerinde değerler olarak kaydedilmiş, bu da modelin hem doğru sınıflandırmalar yapma hem de yanlış sınıflandırmaları minimumda tutma konusundaki başarısını ortaya koymaktadır. Sert, yumuşak ve ağırlıklı oylama stratejileri de sırasıyla %98,24, %98,44 ve %98,68 doğruluk oranları ile güçlü performans sergileyerek neredeyse EffiSceneNet’in bireysel performansı kadar başarı elde etmişlerdir. ViT modeli %97,20 sınıflandırma başarısı ile yüksek bir performans göstermiştir, ancak EffiSceneNet’in ulaştığı doğruluk seviyesine erişememiştir. Bu sonuçlar, EffiSceneNet’in AID veri kümesindeki yüksek genel başarı oranı ve tutarlı performansı ile uzaktan algılama görüntü sınıflandırma görevlerinde önemli bir avantaj sunduğunu ortaya koymaktadır.

Tablo 4.6. AID veri kümesinin sonuçları

Model	Sınıflandırma başarısı (%)	Kesinlik (precision)	Duyarlılık (recall)	F-1 skor
EfficientNetV2S	%92,39	0.9214	0.9187	0.9200
EfficientNetV2M	%90,69	0.9032	0.9075	0.9053
EffiSceneNet	%99,12	0.9931	0.9904	0.9917
Sert oylama	%98,24	0.9523	0.9515	0.9519
Yumuşak oylama	%98,44	0.9537	0.9541	0.9539
Ağırlıklı oylama	%98,68	0.9578	0.9556	0.9567
ViT	%97,20	0.9711	0.9694	0.9702

Şekil 4.6, her bir epoch için doğruluk ve kayıp grafiklerini göstermektedir. Grafikler, her 10 yinelemede sabit bir temel öğrenme oranı ile gerçekleştirilen eğitim sürecini özetlemektedir. Grafikler analiz edildiğinde, ilk dönemlerdeki kayıptaki azalma ve doğruluktaki artış, modelin veri kümesinden öznelikleri hızla öğrendiğini

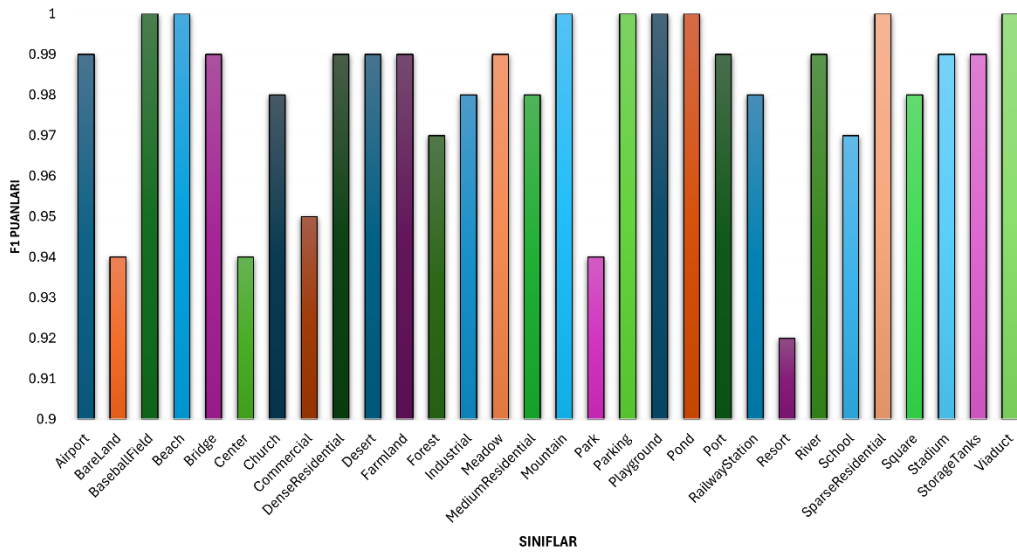
göstermektedir. Eğrilerin sonraki dönemlerde yatay ve paralel olması, modelin artık öğrenmeyi tamamladığını ve performansını optimize ettiğini açıklamaktadır. Bu veri kümesinde eğitim ve doğrulama kümeleri arasındaki performans farkı düşük olduğundan, modelin aşırı öğrenmediği sonucuna varabiliriz.



Şekil 4.6. AID'in eğitim ve doğrulama değerleri. (a) Kayıp grafiği, (b) Doğruluk grafiği

Yapılan sınıflandırma performansı analizi sonucunda, modelin genel doğruluk oranının %99,12 olduğu ve çoğu sınıfta yüksek başarı sağladığı gözlemlenmiştir. Ancak, belirli sınıflarda hataların yoğunlaştığı tespit edilmiştir. Özellikle reklam (commercial) ile endüstriyel (industrial); park ile tatil yeri (resort) kategorilerinde modelin bazı örnekleri yanlış tahmin ettiği görülmüştür. Bu hataların temel nedeni, görüntülerin benzer görsel özelliklere sahip olmasıdır.

Şekil 4.7, EffiSceneNet modelinin AID veri kümesindeki sınıfların F1 puanlarını göstermektedir. Beyzbol sahası (baseball field), sahil (beach), dağ (mountain), otopark (parking), oyun alanı (playground), gölet (pond), seyrek konut (sparse residential) ve viyadük (viaduct) sınıfları için doğruluk oranı %100'e ulaşırken diğer tüm sınıflar %92'nin üzerindedir.



Şekil 4.7. AID veri kümesi sınıflarının F-1 skorları

Tablo 4.7’de, geliştirilen EffiSceneNet modelinin AID veri kümesi üzerindeki performansı, literatürde önerilen çeşitli son teknoloji derin öğrenme algoritmalarıyla karşılaştırılmaktadır. Daha önce sunulan yöntemler arasında CNN-CAPSNET, HW-CNN, DCNNS, CNN-WD, SCCOV, RSFIR, SF-CNN, EfficientNet, H-GCN ve AC Net gibi modeller yer almakta olup, bu yöntemlerin sınıflandırma başarıları %89,64 ile %98,41 arasında değişmektedir. Buna karşılık, önerilen EffiSceneNet modeli %99,12

sınıflandırma başarısı ile oldukça yüksek bir doğruluk oranı yakalamış ve literatürdeki pek çok güçlü modeli geride bırakmıştır. Ayrıca, önerilen topluluk öğrenmesi yöntemi olan ağırlıklı oylama stratejisi, %97,68 doğruluk oranı ile güçlü bir performans göstermiştir. Bu sonuç, topluluk öğrenmesiyle modellerin birleştirilmesinin sınıflandırma doğruluğunu artırmada etkili olduğunu göstermektedir, ancak bireysel olarak EffiSceneNet modelinin daha üstün bir performans sergilediği görülmektedir. EffiSceneNet'in yüksek doğruluk oranı, modelin uzaktan algılama görüntülerindeki karmaşık desenleri ve özellikleri etkili bir şekilde öğrenme kapasitesine sahip olduğunu göstermektedir.

Tablo 4.7. *EffiSceneNet'in AID veri kümesi üzerinde son teknoloji performansla karşılaştırılması*

Yazar	Algoritma	Sonuç
Crews vd. (2014)	CNN-CAPSNET	%96,32
He vd. (2018)	HW-CNN	%96,98
He vd. (2018)	DCNNS	%96,89
Liu vd. (2018)	CNNS-WD	%97,24
He vd. (2019)	SCCOV	%96,10
Fang vd. (2019)	RSFJR	%96,81
Xie vd. (2019)	SF-CNN	%96,66
Zhang vd. (2019)	CapsNet	%96,85
Cao vd. (2020)	SAFF	%93,83
Zhang vd. (2020)	EfficientNet	%97,55
Zhao vd. (2020)	CNN	%96,89
Gao vd. (2021)	H-GCN	%95,78
Alhicri vd. (2021)	EfficientNet-B3-Attn-2	%96,56
Tang vd. (2021)	AC Net	%95,38
Kareem vd. (2021)	VGG-VD-16	%89,64
Bi vd. (2021)	MS2AP	%95,42
Wang vd. (2023)	CNN	%97,18
Song (2023)	FST-EfficientNetV2	%97,79
Sinaga vd. (2024)	EfficientNet-B7	%97,46
Keyan vd. (2024)	RSMamba-H	%92,97
Datla vd. (2024)	Scene-vector	%98,41
Önerilen Yöntem	EffiSceneNet	%99,12
Önerilen topluluk öğrenmesi	Ağırlıklı oylama	%97,68

4.6. Test ve Eğitim Süreleri

Tablo 4.8’de, farklı veri kümeleri üzerinde uygulanan modellerin parametre sayıları ile eğitim ve test süreleri karşılaştırılmaktadır. Sonuçlar, modellerin hesaplama karmaşıklığı ve verimliliği açısından önemli bilgiler sunmaktadır. En yüksek parametre sayısına sahip model olan ViT, tüm veri kümelerinde yaklaşık 85 milyon parametreye sahiptir. Bu durum, ViT modelinin daha karmaşık bir yapıya sahip olduğunu ve daha fazla hesaplama gücü gerektirdiğini göstermektedir. Buna karşın, EffiSceneNet modeli yaklaşık 24-25 milyon parametre ile EfficientNetV2M’den (53 milyon) daha az karmaşıklığa sahiptir ve parametre verimliliği açısından daha optimize bir yapı sunmaktadır. Eğitim süreleri incelendiğinde, en kısa süre UCM veri kümesi üzerinde ViT modeli ile elde edilmiştir (1.809,89 sn), bu da ViT’in daha az parametre güncellemesi gerektiren yapısal avantajlarından kaynaklanmaktadır. Bununla birlikte, daha büyük veri kümelerinde (AID ve NWPU-RESISC45), eğitim süresi önemli ölçüde artmaktadır. EffiSceneNet, EfficientNetV2 modellerine kıyasla daha kısa eğitim sürelerine sahiptir. Örneğin, NWPU-RESISC45 veri kümesinde EfficientNetV2M modeli için eğitim süresi 62.130,00 sn iken, EffiSceneNet için bu süre 45.150,00 sn’ye düşmektedir. Bu durum, EffiSceneNet’in daha verimli bir öğrenme sürecine sahip olduğunu göstermektedir. Test süreleri açısından, en hızlı model ViT olup, özellikle UCM veri kümesinde 4,003 sn ile en düşük test süresini sunmaktadır. EffiSceneNet, test sürelerinde de rekabetçi bir performans göstererek EfficientNetV2 modellerine kıyasla daha hızlı sonuçlar üretmektedir. Örneğin, NWPU-RESISC45 veri kümesinde EfficientNetV2M’in test süresi 254,52 sn iken, EffiSceneNet’in test süresi 103,62 sn’dir. Bu da EffiSceneNet’in gerçek zamanlı uygulamalarda daha avantajlı olabileceğini göstermektedir. Sonuçlar genel olarak değerlendirildiğinde, EffiSceneNet modeli eğitim ve test süreleri açısından oldukça dengeli bir performans sergilemekte ve EfficientNetV2 modellerine göre daha verimli çalışmaktadır. Parametre sayısı ile hesaplama süresi arasındaki dengeyi iyi kuran EffiSceneNet, hem yüksek doğruluk hem de daha düşük işlem süresi ile optimize edilmiş bir model olduğunu kanıtlamaktadır. Özellikle büyük veri kümelerinde sağladığı hesaplama avantajı, modelin uzaktan algılama ve görüntü sınıflandırma gibi yoğun veri işleme gerektiren uygulamalarda tercih edilebilirliğini artırmaktadır.

Tablo 4.8. Modellerin kaydedilen eğitim ve test süreleri

Veri kümesi	Model	Parametre sayısı	Eğitim süresi (sn)	Test süresi (sn)
UCM	EfficientNetV2S	20.358.261	3.186,94	8,66
UCM	EfficientNetV2M	53.177.289	4.142,68	12,57
UCM	EffiSceneNet	24.530.229	3.010,25	6,36
UCM	ViT	85.814.805	1.809,89	4,003
AID	EfficientNetV2S	20.369.790	15.171,42	32,80
AID	EfficientNetV2M	53.188.818	19.723,80	45,95
AID	EffiSceneNet	24.534.846	14.333,33	38,02
AID	ViT	85.821.726	8.614,28	16,002
NWPU-RESISC45	EfficientNetV2S	20.389.005	47.790,00	93,47
NWPU-RESISC45	EfficientNetV2M	53.208.033	62.130,00	254,52
NWPU-RESISC45	EffiSceneNet	24.542.541	45.150,00	103,62
NWPU-RESISC45	ViT	85.833.261	27.135,00	45,006

Yapılan deneylerin çalıştırıldığı bilgisayarın donanımsal özellikleri: Evrişimsel Ağ Evrimi (CPU): 13th Gen Intel(R) Core(TM) i9-13980HX CPU @ 2.20 GHz, 16 GB RAM ve NVIDIA GeForce RTX 4070 GPU. Tüm Kod uygulamaları Python yazılım dilinin 3.10 versiyonunda gerçekleştirilmiştir. Ek olarak Tensorflow-gpu:2.10.0, Cuda-11.2 ve Cudnn-8.1.0 yazılım paketleri kullanılmıştır.

5. SONUÇ

Bu tezde uzaktan algılama görüntülerinin sahne sınıflandırması için EffiSceneNet adında yeni bir derin öğrenme modeli önerilmektedir. EfficientNetV2S CNN modelinin transfer öğrenme yöntemi kullanılarak nasıl özelleştirileceğini ve bir yapay sinir ağı ile entegre edilerek görüntü sınıflandırma görevinde nasıl kullanılacağını göstermektedir. Bu yöntem, önceden eğitilmiş modellerin güçlü özellik çıkarıcılarını kullanarak modelin genel performansını arttırmıştır. Önerilen yöntem UCM, AID ve NWPU-RESISC45 olmak üzere 3 farklı uzaktan algılama veri kümesi ile test edilmiştir; alınan sonuçlar literatür ile karşılaştırıldığında başarı oranı olarak birinci sıraya yükselmiştir. Mevcut yöntemler genellikle küçük ölçekli veri kümeleri ile test edilmiştir. Çoğunlukla tercih edilen UCM ve AID veri kümelerine ek olarak bu çalışmada en son teknolojileri daha da iyileştirmek için özellikle büyük ölçekli NWPU-RESISC45 veri kümesinde de deneyler test edilmiştir. Oldukça büyük ölçekli bir veri kümesi olması, çok sayıda sahne sınıfına sahip olması önerilen yöntem ile değerlendirildiğinde önerilen modelin başarısını kanıtlamıştır. Ayrıca, temel EfficientNetV2S ve EfficientNetV2M modelleri ile başarı oranları karşılaştırıldığında, EffiSceneNet modelinin belirgin bir performans üstünlüğü sağladığı görülmüştür. EffiSceneNet'in temel avantajı, daha düşük parametre sayısı ile yüksek doğruluk oranları sunmasıdır. Ancak EffiSceneNet modelinin yüksek doğruluk oranları sunmasına rağmen, işlem süreleri açısından ilerleyen çalışmalarda daha fazla optimizasyona ihtiyaç duyabilir.

Bu tezde, EffiSceneNet modeli, EfficientNetV2S ile EfficientNetV2M CNN modelleri olmak üzere üç farklı derin öğrenme modeli kullanılarak topluluk öğrenme yöntemleriyle topluluk modelleri geliştirilmiştir. Yapılan analizler, topluluk öğrenme yöntemlerinin bireysel modellerin sınırlamalarını telafi ederek genel sınıflandırma başarısını önemli ölçüde artırdığını göstermiştir. Özellikle yumuşak oylama yöntemi, farklı modellerin tahmin olasılıklarını birleştirerek UCM ve NWPU-RESISC45 veri kümelerinde en yüksek doğruluk oranlarına ulaşmıştır. Elde edilen sonuçlar, topluluk modellerinin veri çeşitliliğine karşı daha genellenebilir ve güvenilir bir yapı sunduğunu ortaya koymuştur. Bu sayede önerilen topluluk modelleri karmaşık veri kümelerinde yüksek doğruluk oranlarıyla üstün performans sağlamıştır. EffiSceneNet modelinin bireysel olarak en güçlü performansa sahip olması topluluk modellerin başarısına önemli bir katkı sağlamıştır.

İlerleyen çalışmalarda, önerilen EffiSceneNet modelinin geliştirilmesi ve uygulama alanlarının genişletilmesi için çeşitli yönlere odaklanılabilir. Öncelikle, modelin genellenebilirliğini artırmak amacıyla farklı coğrafi bölgelerden özellikle sensörlerden elde edilen daha çeşitli ve büyük ölçekli veri kümeleriyle performansı değerlendirilebilir. Ayrıca, modelin işlem sürelerini azaltmak için özel donanımlara yönelik optimizasyon yapılabilir. Örneğin, NVIDIA TensorRT, Intel OpenVINO veya Apple Metal gibi platformlar kullanılarak model hızlandırılabilir. Son olarak da ViT modelindeki hesaplama maliyetinin katlarca düşük olmasından yararlanılarak daha iyi doğruluk oranları elde edilmeye çalışılabilir.



KAYNAKÇA

- Alhichri, H., Alswayed, A. S., Bazi, Y., Ammour, N., & Alajlan, N. A. (2021). Classification of remote sensing images using EfficientNet-B3 CNN model with attention. *IEEE access* 9, s. 14078-14094.
- Ali, N., Zafar, B., Iqbal, M. K., Sajid, M., Younis, M. Y., Dar, S. H., . . . Lee, I. H. (2019). Modeling global geometric spatial information for rotation invariant classification of satellite images. *PloS one*, 14(7), e0219833.
- Bazi, Y., Bashmal, L., Rahhal, M. M., Dayil, R. A., & Ajlan, N. A. (2021). Vision Transformers for Remote Sensing Image Classification. *Remote Sensing*, 13(3), 516.
- Bazi, Y., Melgani, F., Othman, E., Alhichri, H., Alajlan, N., & Zuair, M. (2017). Domain Adaptation Network for Cross-Scene Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(8), 4441-4456.
- Bi, Q., Zhang, H., & Qin, K. (2021). Multi-scale stacking attention pooling for remote sensing scene. *Neurocomputing*, 436, 147-161.
- Cao, R., Fang, L., Lu, T., & He, N. (2020). Self-Attention-Based Deep Feature Fusion for Remote Sensing Scene Classification. *IEEE Geoscience and Remote Sensing Letters*, 18(1), 43-47.
- Chen, K., Chen, B., Liu, C., Li, W., Zou, Z., & Shi, Z. (2024). Rsmamba: Remote sensing image classification with state space model. *IEEE Geoscience and Remote Sensing Letters*.
- Cheng, G., Xie, X., Han, J., Guo, L., & Xia, G.-S. (2020). Remote Sensing Image Scene Classification Meets Deep Learning: Challenges, Methods, Benchmarks, and Opportunities. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, s. 3735-3756.
- Cheon, M. (2024). KOLMOGOROV-ARNOLD NETWORK FOR SATELLITE IMAGE CLASSIFICATION IN REMOTE SENSING. *arXiv preprint arXiv:2406.00600*.
- Datla, R., Perveen, N., & C., K. M. (2024). Learning scene-vectors for remote sensing image scene classification. *Neurocomputing*, 587, 127679.
- DEDE, M. A., Aptoula, E., & Genc, Y. (2018). Deep network ensembles for aerial scene classification. *IEEE Geoscience and Remote Sensing Letters*, 16(5), 732-735.
- Deng, P., Xu, K., & Huang, H. (2021). When CNNs Meet Vision Transformer: A Joint Framework for Remote Sensing Scene Classification. *IEEE Geoscience and Remote Sensing Letters*, 19, 1-5.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., . . . Houlsby, N. (2020). AN IMAGE IS WORTH 16X16 WORDS: TRANSFORMERS FOR IMAGE RECOGNITION AT SCALE. *arXiv preprint arXiv:2010.11929*.

- Fang, J., Yuan, Y., Lu, X., & Feng, Y. (2019). Robust Space–Frequency Joint Representation for Remote Sensing Image Scene Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(10), 7492-7502.
- Gao, Y., Shi, J., Li, J., & Wang, R. (2021). Remote Sensing Scene Classification Based on High-Order Graph Convolutional Network. *European Journal of Remote Sensing*, 54(sup1), s. 141-155.
- Goodfellow, L., Bengio, Y., & Courville, A. (2016). *Deep Learning*. The MIT Press.
- Han, J., Cheng, G., & Lu, X. (2017). Remote Sensing Image Scene Classification: Benchmark and State of the Art. *Proceedings of the IEEE*, 105(10), s. 1865-1883.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*, s. 770-778.
- He, N., Fang, L., Li, S., Plaza, A., & Plaza, J. (2018). Remote Sensing Scene Classification Using Multilayer Stacked Covariance Pooling. *IEEE Transactions on Geoscience and Remote Sensing*, 56(12), 6899-6910.
- He, N., Fang, L., Li, S., Plaza, J., & Plaza, A. (2019). Skip-Connected Covariance Network for Remote Sensing Scene Classification. *IEEE transactions on neural networks and learning systems*, 31(5), 1461-1474.
- Jang, B., Kim, M., Harerimana, G., Kang, S.-u., & Kim, J. W. (2020). Bi-LSTM Model to Increase Accuracy in Text Classification: Combining Word2vec CNN and Attention Mechanism. *Applied Sciences*, 10(17), 5841.
- Kareem, R. S., Ramanjineyulu, A. G., Rajan, R., Setiawan, R., Sharma, & D., Gupta, M. K., . . . Sengan, S. (2021). Multilabel land cover aerial image classification using convolutional neural networks. *Arabian Journal of Geosciences*, 14, 1-18.
- Kwak, T., & Kim, Y. (2023). Semi-supervised land cover classification of remote sensing imagery using CycleGAN and EfficientNet. *KSCE Journal of Civil Engineering*, 27(4), s. 1760-1773.
- Li, H., Dou, X., Tao, C., Hou, Z., Chen, J., Peng, J., . . . Zhao, L. (2017). RSI-CB: A Large-Scale Remote Sensing Image Classification Benchmark via Crowdsourced Data. *arXiv preprint arXiv:1705.10450*.
- Liu, W., Prasad, S., & Crawford, M. (2024). Investigation of Hierarchical Spectral Vision Transformer Architecture for Classification of Hyperspectral Imagery. *IEEE Transactions on Geoscience and Remote Sensing*.
- Liu, X., Zhou, Y., Zhao, J., Yao, R., Liu, B., & Zheng, Y. (2019). Siamese Convolutional Neural Networks for Remote Sensing Scene Classification. *IEEE Geoscience and Remote Sensing Letters*, 16(8), 1200-1204.
- Liu, Y., Zhong, Y., & Qin, Q. (2018). Scene Classification Based on Multiscale Convolutional Neural Network. *IEEE Transactions on Geoscience and Remote Sensing*, 56(12), 7109-7121.

- Luus, F. P., Salmon, B. P., Bergh, F. v., & Maharaj, B. T. (2015). Multiview Deep Learning for Land-Use Classification. *IEEE Geoscience and Remote Sensing Letters*, 12(12), 2448-2452.
- Lv, P., Wu, W., Zhong, Y., Du, F., & Zhang, L. (2022). SCViT: A Spatial-Channel Feature Preserving Vision Transformer for Remote Sensing Image Scene Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1-12.
- Mahaur, B., & Mishra, K. (2023). Small-object detection based on YOLOv5 in autonomous driving systems. *Pattern Recognition Letters*, s. 115-122.
- Mienye, I. D., & Sun, Y. (2022). A Survey of Ensemble Learning: Concepts, Algorithms, Applications, and Prospects. *IEEE Access*, 10, 99129-99149.
- Minetto, R., Segundo, M. P., & Sarkar, S. (2019). Hydra: An ensemble of convolutional neural networks for geospatial land classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(9), 6530-6541.
- Mishra, N. B., & Crews, K. A. (2014). Mapping vegetation morphology types in a dry savanna ecosystem: integrating hierarchical object-based image analysis with Random Forest. *International Journal of Remote Sensing*, 35(3), 1175-1198.
- Muhammad, U., Wang, W., Chattha, S. P., & Ali, S. (2018, August). Pre-trained VGGNet Architecture for Remote-Sensing Image Scene Classification. In *2018 24th International Conference on Pattern Recognition (ICPR)* , s. 1622-1627.
- Özyurt, F., Avci, E., & Sert, E. (2020). UC-Merced image classification with CNN feature reduction using wavelet entropy optimized with genetic algorithm. *Traitement du Signal*.
- Petrovska, B., Zdravevski, E., Lameski, P., Corizzo, R., Štajduhar, I., & Lerga, J. (2020). Deep Learning for Feature Extraction in Remote Sensing: A Case-Study of Aerial Scene Classification. *Sensors*, 20(14), 3906.
- Sarmah, U., Borah, P., & Bhattacharyya, D. K. (2024). Ensemble Learning Methods: An Empirical Study. *SN Computer Science*, 5(7), 924.
- Sha, Z., & Li, J. (2022). MITformer: A Multiinstance Vision Transformer for Remote Sensing Scene Classification. *IEEE Geoscience and Remote Sensing Letters*, 19, 1-5.
- Shin, H.-C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., . . . Summers, R. M. (2016). Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE transactions on medical imaging*, 35(5), 1285-1298.
- Simonyan, K., & Zisserman, A. (2015). VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION. *arXiv preprint arXiv:*, s. 1409-1556.
- Sinaga, K. B., Yudistira, N., & Santoso, E. (2024). Efficient CNN for high-resolution remote sensing imagery understanding. *Multimedia Tools and Applications* 83.22, s. 61737-61759.

- Song, H. (2023). FST-EfficientNetV2: Exceptional Image Classification for Remote Sensing. *Computer Systems Science & Engineering*, 46(3).
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., . . . Rabinovich, A. (2015). Going Deeper with Convolutions. *Proceedings of the IEEE conference on computer vision and pattern recognition*, s. 1-9.
- Tan, M., & Le, Q. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In *International conference on machine learning*, s. 6105-6114.
- Tan, M., & Le, Q. (2021). EfficientNetV2: Smaller Models and Faster Training. In *International conference on machine learning*, s. 10096-10106.
- Tang, X., Ma, Q., Zhang, X., Liu, F., Ma, J., & Jiao, L. (2021). Attention Consistent Network for Remote Sensing Scene Classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 2030-2045.
- Thirumaladevi, S., Swamy, K. V., & Sailaja, M. (2023). Remote Sensing Image Scene Classification by Transfer Learning to Augment the Accuracy Author Links Open Overlay Panel. *Measurement: Sensors*, 25, 100645.
- Turki, T., Al-Sharif, A., & Taguchi, Y.-h. (2021). End-to-End Deep Learning for Detecting Metastatic Breast Cancer in Axillary Lymph Node from Digital Pathology Images. *Intelligent Data Engineering and Automated Learning—IDEAL 2021: 22nd International Conference, IDEAL 2021, Manchester, UK, November 25–27, 2021, Proceedings 22*, s. 343-353.
- Wang, J., Li, W., Zhang, M., Tao, R., & Chanussot, J. (2023). Remote Sensing Scene Classification via Multi-Stage Self-Guided Separation Network. *Network, M. S. S. G. S.*
- Wang, Q., Liu, S., Chanussot, J., & Li, X. (2018). Scene Classification With Recurrent Attention of VHR Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, 57(2), 1155-1167.
- Xia, G.-S., Hu, J., Hu, F., Shi, B., Bai, X., Zhong, Y., & Zhang, L. (2017). AID: A Benchmark Dataset for Performance Evaluation of Aerial. *IEEE Transactions on Geoscience and Remote Sensing*, s. 3965-3981.
- Xia, G.-S., Yang, W., Delon, J., Gousseau, Y., Sun, H., & Maître, H. (2010). Structural high-resolution satellite image indexing. *Structural high-resolution satellite image indexing*, s. 298-303.
- Xie, J., He, N., Fang, L., & Plaza, A. (2019). Scale-Free Convolutional Neural Network for Remote Sensing Scene Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(9), 6916-6928.
- Yang, Y., & Newsam, S. (2010, Kasım). Bag-Of-Visual-Words and Spatial Extensions for Land-Use. *Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems*, s. 270-279.
- Ying, H., Yang, C., & Lu, J. (2022, Nisan). Research on Remote Sensing Image Classification Algorithm Based on EfficientNet. In *2022 7th international conference on intelligent computing and signal processing (ICSP)* (pp. 1757-1761).

- YOUNESI, A., ANSARI, M., FAZLI, M., EYLALI, A., EYLALI, A., & HENKEL, J. (2024). A Comprehensive Survey of Convolutions in Deep Learning: Applications Challenges, and Future Trends. *IEEE Access*, s. 41180-41218.
- Zhang, D., Liu, Z., & Shi, X. (2020). Transfer learning on efficientnet for remote sensing image classification. In *2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE)* (pp. 2255-2258). IEEE.
- Zhang, W., Tang, P., & Zhao, L. (2019). Remote sensing image scene classification using CNN-CapsNet. *Remote Sensing*, 11(5), 494.
- Zhao, B., Zhong, Y., Xia, G.-S., & Zhang, L. (2015). Dirichlet-Derived Multiple Topic Scene Classification Model for High Spatial Resolution Remote Sensing Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 54(4), 2108-2123.
- Zhao, L., Tang, P., & Huo, L. (2016). Feature significance-based multibag-of-visual-words model for remote sensing image scene classification. *Journal of Applied Remote Sensing*, 10(3), 035004-035004.
- Zhao, Z., Luo, Z., Li, J., Chen, C., & Piao, Y. (2020). When Self-Supervised Learning Meets Scene Classification: Remote Sensing Scene Classification Based on a Multitask Learning Framework. *Remote Sensing*, 12(20), 3276.
- Zhou, W., Newsam, S., Li, C., & Shao, D. Z. (2018). PatternNet: A benchmark dataset for performance evaluation of remote sensing image retrieval. *ISPRS journal of photogrammetry and remote sensing*, 145, 197-209.
- Zhu, Q., Zhong, Y., Zhang, L., & Li, D. (2018). Adaptive Deep Sparse Semantic Modeling Framework for High Spatial Resolution Image Scene Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 56(10), 6180-6195.
- Zhu, R., Yan, L., Mo, N., & Liu, Y. (2019). RETRACTED: Attention-Based Deep Feature Fusion for the Scene Classification of High-Resolution Remote Sensing Images. *Remote Sensing*, 11(17), 1996.
- Zou, Q., Ni, L., Zhang, T., & Wang, Q. (2015). Deep learning based feature selection for remote sensing scene classification. *IEEE Geoscience and remote sensing letters*, 12(11), 2321-2325.

ÖZGEÇMİŞ

ORCID NO: 0009-0003-9760-5005

Ad Soyad : DAMLA DALGIÇ

Yabancı Dil : İNGİLİZCE

Doğum Yeri ve Yılı :

E-Posta :

Eğitim ve Mesleki Geçmişi:

- 2022, Eskişehir Osmangazi Üniversitesi, Mühendislik Mimarlık Fakültesi, Bilgisayar Mühendisliği
- 2023 – halen, İş Süreç Analisti, ALP Havacılık
- 2022 – 2023, Yapay Zeka Mühendisi, Inovako