



**GÖRÜNTÜ İŞLEME TEKNİKLERİ VE DERİN
ÖĞRENME METOTLARI KULLANILARAK
YIĞMA TAŞ DUVAR SEGMENTASYONU**

Rüstem Muhammed KARADEMİR

Danışman: Doç. Dr. Tolga AYDIN

Yüksek Lisans Tezi

Bilgisayar Mühendisliği Ana Bilim Dalı

2025

(Her hakkı saklıdır.)

T.C.
ATATÜRK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

**GÖRÜNTÜ İŞLEME TEKNİKLERİ VE DERİN ÖĞRENME METOTLARI
KULLANILARAK YIĞMA TAŞ DUVAR SEGMENTASYONU**
(Masonry Stone Wall Segmentation Using Image Processing Techniques and Deep Learning
Methods)

YÜKSEK LİSANS TEZİ

Rüstem Muhammed KARADEMİR

Danışman: Doç. Dr. Tolga AYDIN

Erzurum
Ağustos, 2025

KABUL VE ONAY TUTANAĐI

Rüstem Muhammed KARADEMİR tarafından hazırlanan “Görüntü İşleme Teknikleri ve Derin Öğrenme Metotları Kullanılarak Yığma Taş Duvar Segmentasyonu” başlıklı çalışması 08 / 09 / 2025 tarihinde yapılan tez savunma sınavı sonucunda başarılı bulunarak jürimiz tarafından Bilgisayar Mühendisliği Ana Bilim Dalı, Bilgisayar Mühendisliği Bilim Dalında yüksek lisans tezi olarak kabul edilmiştir.

Jüri Başkanı: Doç. Dr. Ferhat BOZKURT

Atatürk Üniversitesi

Aslı Islak İmzalıdır

Danışman: Doç. Dr. Tolga AYDIN

Atatürk Üniversitesi

Aslı Islak İmzalıdır

Jüri Üyesi: Dr. Öğr. Üyesi Işıl KARABEY AKSAKALLI

Erzurum Teknik Üniversitesi

Aslı Islak İmzalıdır

Bu tezin Atatürk Üniversitesi Lisansüstü Eğitim ve Öğretim Yönetmeliği'nin ilgili maddelerinde belirtilen şartları yerine getirdiğini onaylarım.

Prof. Dr. Alper NUHOĞLU

Enstitü Müdürü

Aslı Islak İmzalıdır

ETİK BİLDİRİM VE İNTİHAL BEYAN FORMU

Yüksek Lisans Tezi olarak *Doç. Dr. Tolga AYDIN* danışmanlığında sunulan “**Görüntü İşleme Teknikleri ve Derin Öğrenme Metotları Kullanılarak Yığma Taş Duvar Segmentasyonu**” başlıklı çalışmanın tarafımızdan bilimsel etik ilkelere uyularak yazıldığını, yararlanılan eserlerin kaynakçada gösterildiğini, Fen Bilimleri Enstitüsü tarafından belirlenmiş olan Turnitin Programı benzerlik oranlarının aşılmadığını ve aşağıdaki oranlarda olduğunu beyan ederiz.

Tez Bölümleri	Tezin Benzerlik Oranı (%)	Maksimum Oran (%)
Giriş	2	30
Kuramsal Temeller	4	30
Materyal ve Metot	4	35
Araştırma Bulguları ve Tartışma	1	20
Sonuçlar ve Öneriler	0	20
Tezin Geneli	2	25

Not: Yedi kelimeye kadar benzerlikler ile Başlık, Kaynakça, İçindekiler, Teşekkür, Dizin ve Ekler kısımları tarama dışı bırakılabilir. Yukarıdaki azami benzerlik oranları yanında tek bir kaynaktan olan benzerlik oranlarının %5'den büyük olmaması gerekir.

Sunulan bilgilerin doğru olduğunu, aksi halde doğacak hukuki sorumlulukları kabul ettiğimizi beyan ederiz.

Tez Yazarı (Öğrenci)	Tez Danışmanı
Rüstem Muhammed Karademir	Doç. Dr. Tolga AYDIN
11.9.2025	11.9.2025
İmza: Aslı Islak İmzalıdır	İmza: Aslı Islak İmzalıdır

* Tez ile ilgili YÖKTEZ’de yayınlamasına ilişkin bir engelleme var ise aşağıdaki alanı doldurunuz.

☐ Tezle ilgili patent başvurusu yapılması / patent alma sürecinin devam etmesi sebebiyle Enstitü Yönetim Kurulunun/....../.... tarih ve sayılı kararı ile teze erişim 2 (iki) yıl süreyle engellenmiştir.

☐ Enstitü Yönetim Kurulunun/....../.... tarih ve sayılı kararı ile teze erişim 6 (altı) ay süreyle engellenmiştir.

TEŞEKKÜR

Tez yazım sürecim boyunca her adımda içtenlikle yardım eden, bana kıymetli bilgiler veren ve rehberlik sunan değerli danışman hocam Doç. Dr. Tolga AYDIN’a, eğitim sürecimde ve tez aşamasında yol gösteren ve desteklerini esirgemeyen sayın Doç. Dr. Ferhat BOZKURT’a, tezime fikirleriyle destek sağlayan sayın Doç. Dr. İrfan KOCAMAN’a en içten teşekkürlerimi sunarım.

Bu zorlu süreçte desteklerini hiçbir zaman esirgemeyen aileme, beni anlayışla karşılayıp sabırla yol gösteren dostlarıma, tüm süreçte maddi manevi ve psikolojik desteği ile motivasyonumu sağlayan sevgili eşime en içten teşekkürlerimi sunarım. Karşılaştığım tüm zorlukları aşmamda hep yanımda olan ve desteklerini eksik etmeyen bu özel insanlara minnettarım.

Rüstem Muhammed KARADEMİR

ÖZET

YÜKSEK LİSANS TEZİ

GÖRÜNTÜ İŞLEME TEKNİKLERİ VE DERİN ÖĞRENME METOTLARI KULLANILARAK YIĞMA TAŞ DUVAR SEGMENTASYONU

Rüstem Muhammed KARADEMİR

Danışman: Doç. Dr. Tolga AYDIN

Amaç: Bu çalışmanın amacı, tarihi yığma taş yapıların korunması ve yapısal analizlerinde kullanılmak üzere, duvar görsellerindeki taş ve harç bölgelerini piksel düzeyinde otomatik olarak segmente eden derin öğrenme ve görüntü işleme tabanlı hibrit bir yöntem geliştirmektir. Çalışma ile geleneksel manuel ölçüm yöntemlerine kıyasla daha hızlı, nesnel ve yüksek doğruluklu bir analiz altyapısı sunulması hedeflenmektedir.

Yöntem: Araştırmada kamuya açık bir veri seti ile Erzurum ilindeki tarihi yapılardan toplanan yerel görüntülerin birleştirildiği hibrit bir veri seti kullanılmıştır. Yüksek çözünürlüklü görüntüler, detay kaybını önlemek amacıyla 256x256 piksellik daha küçük parçalara (patch) ayrılarak toplam 108194 görüntü maske çifti (85747 eğitim, 22447 doğrulama) modellere girdi olarak hazırlanmıştır. U-Net, FPN, LinkNet ve PSPNet mimarileri; VGG16, ResNet34/50, MobileNetV1/V2, EfficientNetB0 ve DenseNet121 gibi önceden eğitilmiş yedi farklı omurga (backbone) ile transfer öğrenme yöntemi kullanılarak eğitilmiştir. Model performansları Piksel Doğruluğu, IoU, Dice Katsayısı, Kesinlik, Duyarlılık ve Özgüllük gibi segmentasyon metrikleri ile değerlendirilmiştir. Derin öğrenme modellerinin ham segmentasyon çıktıları, morfolojik işlemler, Watershed algoritması ve bağlantılı bileşen analizi gibi son işlem adımlarıyla iyileştirilmiştir. Nihai segmentasyon maskeleri üzerinden taş boyutları hesaplanmıştır.

Bulgular: Deneysel bulgular, geliştirilen derin öğrenme modellerinin yığma taş duvarlardaki taş ve harç bölgelerini yüksek bir başarıyla ayrıştırabildiğini göstermiştir. Karşılaştırılan modeller arasında U-Net/MobileNetV1 modeli %97,87 IoU ve %98,92 Dice katsayısı ile en yüksek genel segmentasyon performansını sergilemiş, %99,09 duyarlılık ve %82,51 özgüllük değerlerine ulaşmıştır. Harç bölgelerinin tespitinde kritik bir metrik olan özgüllük açısından %83,03 ile en başarılı model U-Net/EfficientNetB0 olmuştur. Hesaplama verimliliği (eğitim süresi ve parametre sayısı) ve performans dengesi açısından ise LinkNet/MobileNetV1 modeli en uygun çözümlerden birini sunmuş ve pratik uygulamalar için cazip bir seçenek olduğu belirlenmiştir. Uygulanan son işlem adımlarının, modellerin eksik veya hatalı tahminlerini önemli ölçüde düzelttiği ve taşların geometrik özelliklerinin hassas bir şekilde hesaplanmasına katkı sağladığı tespit edilmiştir.

Sonuç: Bu çalışma, derin öğrenme ve görüntü işleme tekniklerinin bir arada kullanıldığı hibrit yaklaşımın, yığma taş duvarların karmaşık yapılarının segmentasyonunda hızlı, nesnel ve yüksek doğruluklu bir yöntem sunduğunu kanıtlamıştır. Gelecekteki çalışmalar için veri seti genişletme, daha gelişmiş mimariler, gerçek zamanlı uygulamalar için optimizasyon ve detaylı hata analizi gibi alanlarda potansiyel araştırma yönleri belirlenmiştir. Geliştirilen bu yaklaşım, tarihi yapıların korunması ve restorasyonu alanında çalışan uzmanlara, yapısal analizler için temel bir girdi olan taş ve harç dağılımını otomatik olarak belirleme imkânı sunarak geleneksel yöntemlere kıyasla önemli bir avantaj sağlamaktadır.

Anahtar Kelimeler: yığma taş duvar, semantik segmentasyon, derin öğrenme, U-Net, görüntü işleme

Ağustos 2025, 94 sayfa

ABSTRACT

MASTER'S THESIS

MASONRY STONE WALL SEGMENTATION USING IMAGE PROCESSING TECHNIQUES AND DEEP LEARNING METHODS

Rüstem Muhammed KARADEMİR

Supervisor: Assoc. Prof. Dr. Tolga AYDIN

Purpose: The aim of this study is to develop a hybrid method based on deep learning and image processing that automatically segments stone and mortar areas in wall images at the pixel level for use in the preservation and structural analysis of historic masonry structures. The study seeks to provide an analysis infrastructure that is faster, objective, and highly accurate compared to traditional manual measurement methods.

Method: The research utilized a hybrid dataset created by combining a publicly available dataset with local images collected from historical structures in the province of Erzurum. To prevent the loss of fine details, high-resolution images were partitioned into smaller 256x256 pixel patches, generating a total of 108194 image-mask pairs (85747 for training, 22447 for validation) as input for the models. Four main architectures U-Net, FPN, LinkNet, and PSPNet were trained using transfer learning with seven different pre-trained backbones: VGG16, ResNet34/50, MobileNetV1/V2, EfficientNetB0, and DenseNet121. Model performance was evaluated using segmentation metrics, including Pixel Accuracy, Intersection over Union (IoU), Dice Coefficient, Precision, Recall (Sensitivity), and Specificity. The raw segmentation outputs from the deep learning models were subsequently refined through post-processing steps such as morphological operations, the Watershed algorithm, and connected component analysis (CCA). Finally, stone dimensions were calculated from the finalized segmentation masks.

Findings: The experimental results demonstrate that the developed deep learning models successfully segmented stone and mortar regions in masonry walls with high accuracy. Among the compared models, the U-Net/MobileNetV1 architecture exhibited the best overall segmentation performance, achieving an IoU of 97,87% and a Dice coefficient of 98,92%, with a sensitivity of 99,09% and a specificity of 82,51%. In terms of specificity, a critical metric for the accurate detection of mortar regions, the U-Net/EfficientNetB0 model proved to be the most successful with a score of 83,03%. When considering the trade-off between computational efficiency (training time and parameter count) and performance, the LinkNet/MobileNetV1 model emerged as one of the most suitable solutions, identifying it as an attractive option for practical applications. It was also determined that the applied post-processing steps significantly corrected the models' incomplete or erroneous predictions and contributed to the precise calculation of the geometric properties of the stones.

Results: This study confirms that a hybrid approach combining deep learning and image processing techniques offers a fast, objective, and highly accurate method for the segmentation of complex masonry wall structures. Potential future research directions have been identified, including dataset expansion, the exploration of more advanced architectures, optimization for real-time applications, and detailed error analysis. The developed approach provides experts in the conservation and restoration of historical structures with a powerful tool for automatically determining the distribution of stone and mortar, which is a fundamental input for structural analysis. This offers a significant advantage over conventional methods.

Keywords: masonry wall, semantic segmentation, deep learning, U-Net, image processing

August 2025, 94 pages

İÇİNDEKİLER

KABUL VE ONAY TUTANAĞI	i
ETİK BİLDİRİM VE İNTİHAL BEYAN FORMU	ii
TEŞEKKÜR	iii
ÖZET	iv
ABSTRACT	v
İÇİNDEKİLER.....	vi
TABLolar DİZİNİ	viii
ŞEKİLLER DİZİNİ.....	ix
KISALTMALAR VE SİMGELER DİZİNİ	xi
GİRİŞ	1
KURAMSAL TEMELLER	6
Temel Kavramlar.....	6
Yapay zekâ	6
Makine öğrenmesi	7
Yapay sinir ağları.....	8
Derin öğrenme.....	8
Görüntü işleme	9
Görüntü segmentasyonu.....	10
Literatür Taraması	10
MATERYAL VE METOT	17
Veri Seti	17
Veri Ön İşleme.....	20
Segmentasyon Mimarileri	20
U-Net.....	21
FPN (Feature Pyramid Network).....	21
LinkNet.....	22
PSPNet (Pyramid Scene Parsing Network).....	23
VGG16	24
ResNet34/50.....	25
DenseNet121.....	26
MobileNetV1/V2	26
EfficientNetB0.....	28
Transfer Öğrenme.....	30

Performans Metrikleri ve Parametreler	31
<i>Karmaşıklık matrisi</i>	32
<i>Piksel doğruluğu (Pixel Accuracy)</i>	32
<i>Kesişim/Birleşim oranı (Intersection over union-IoU) / Jaccard indeksi</i>	32
<i>Dice katsayısı (Dice Coefficient)</i>	33
<i>Kesinlik (Precision)</i>	33
<i>Duyarlılık (Recall)</i>	33
<i>F1 Skoru (F1 Score)</i>	33
<i>Özgüllük (Specificity)</i>	34
<i>Erken durdurma (Early Stopping)</i>	34
Metot	34
ARAŞTIRMA BULGULARI VE TARTIŞMA	38
Uygulama Detayları	38
Değerlendirme Metrikleri	39
Deneysel Sonuçlar	39
<i>U-Net modellerine ait eğitim ve doğrulama grafikleri</i>	39
<i>FPN modellerine ait eğitim ve doğrulama grafikleri</i>	43
<i>LinkNet modellerine ait eğitim ve doğrulama grafikleri</i>	47
<i>PSPNet modellerine ait eğitim ve doğrulama grafikleri</i>	51
Görüntü İşleme ve Son İşlem (Post-Processing) Sonuçları	59
Sonuçların Karşılaştırılması ve Tartışılması	67
SONUÇLAR VE ÖNERİLER	69
KAYNAKÇA	72
ÖZGEÇMİŞ.....	80

TABLÖLAR DİZİNİ

Tablo 1. İki Sınıflı Karmaşıklık Matrisi.....	32
Tablo 2. Modellerin Test Sonuçları.....	56
Tablo 3. Önerilen Yaklaşımın Benzer Çalışmalarla Karşılaştırılması	67



ŞEKİLLER DİZİNİ

Şekil 1. Yığma Taş Yapılar: (a) Moloz taş, (b) Kesme taş, (c) İri Kesme Taş (Dabanlı, 2008)	1
Şekil 2. Yakutiye Medresesi ve ölçüm örneği (Kocaman vd., 2018)	2
Şekil 3. Yığma taş duvar segmentasyon örneği	3
Şekil 4. Yapay zekâ küme yapısı (Goodfellow vd., 2016)	6
Şekil 5. Çalışma akış diyagramı	17
Şekil 6. Veri setindeki örnek duvar görselleri (a) Rastgele moloz taş duvar, (b) Kare moloz duvar, (c) Kuru moloz duvar, (d) İnce kesme taş, (e) Kaba kesme taş (İbrahim vd., 2020)	18
Şekil 7. Yerel veri seti örnek duvar görselleri ve etiketleri	19
Şekil 8. U-Net mimarisi (Ronneberger vd., 2015)	21
Şekil 9. FPN mimarisi (Lin vd., 2017)	22
Şekil 10. LinkNet mimarisi (Chaurasia & Culurciello, 2017)	22
Şekil 11. PSPNet mimarisi (Zhao vd., 2017)	23
Şekil 12. VGG16 model mimarisi (Loukadakis vd., 2018)	24
Şekil 13. Artık Blok Yapısı (He vd., 2016)	25
Şekil 14. ResNet34 mimarisi (He vd., 2016)	25
Şekil 15. DenseNet mimarisi (Huang vd., 2017)	26
Şekil 16. Derinlik bazlı ayrılabilir evrişim katman yapısı (Zhao vd., 2024)	27
Şekil 17. MobilenetV2 katman şeması (Sandler vd., 2018)	28
Şekil 18. Bileşik ölçeklendirme (Tan ve Le, 2019)	29
Şekil 19. EfficientNet model mimarisi (Hisaria vd., 2024)	30
Şekil 20. Transfer öğrenme	31
Şekil 21. U-Net/VGG16 modeli eğitim ve doğrulama grafiği	40
Şekil 22. U-Net/ResNet34 modeli eğitim ve doğrulama grafiği	40
Şekil 23. U-Net/ResNet50 modeli eğitim ve doğrulama grafiği	41
Şekil 24. U-Net/MobileNetV1 modeli eğitim ve doğrulama grafiği	41
Şekil 25. U-Net/MobileNetV2 modeli eğitim ve doğrulama grafiği	42
Şekil 26. U-Net/EfficientNetB0 modeli eğitim ve doğrulama grafiği	42
Şekil 27. U-Net/DenseNet121 modeli eğitim ve doğrulama grafiği	43
Şekil 28. FPN/VGG16 modeli eğitim ve doğrulama grafiği	44
Şekil 29. FPN/ResNet34 modeli eğitim ve doğrulama grafiği	44
Şekil 30. FPN/ResNet50 modeli eğitim ve doğrulama grafiği	45
Şekil 31. FPN/MobileNetV1 modeli eğitim ve doğrulama grafiği	45

Şekil 32. FPN/MobileNetV2 modeli eğitim ve doğrulama grafiği.	46
Şekil 33. FPN/EfficientNetB0 modeli eğitim ve doğrulama grafiği.	46
Şekil 34. FPN/DenseNet121modeli eğitim ve doğrulama grafiği.	47
Şekil 35. LinkNet/VGG16 modeli eğitim ve doğrulama grafiği.	48
Şekil 36. LinkNet/ResNet34 modeli eğitim ve doğrulama grafiği.	48
Şekil 37. LinkNet/ResNet50 modeli eğitim ve doğrulama grafiği.	49
Şekil 38. LinkNet/MobileNetV1 modeli eğitim ve doğrulama grafiği.	49
Şekil 39. LinkNet/MobileNetV2 modeli eğitim ve doğrulama grafiği.	50
Şekil 40. LinkNet/EfficientNetB0 modeli eğitim ve doğrulama grafiği.	50
Şekil 41. LinkNet/DenseNet121modeli eğitim ve doğrulama grafiği.	51
Şekil 42. PSPNet/VGG16 modeli eğitim ve doğrulama grafiği.	52
Şekil 43. PSPNet/ResNet34 modeli eğitim ve doğrulama grafiği.	52
Şekil 44. PSPNet/ResNet50 modeli eğitim ve doğrulama grafiği.	53
Şekil 45. PSPNet/MobileNetV1 modeli eğitim ve doğrulama grafiği.	53
Şekil 46. PSPNet/MobileNetV2 modeli eğitim ve doğrulama grafiği.	54
Şekil 47. PSPNet/EfficientNetB0 modeli eğitim ve doğrulama grafiği.	54
Şekil 48. PSPNet/DenseNet121modeli eğitim ve doğrulama grafiği.	55
Şekil 49. U-Net/MobileNetV1 modeli tahmin maske örnekleri.	60
Şekil 50. Image_25 tahmin histogramı.	61
Şekil 51. Image_25 Otsu ve klasik eşikleme çıktısı.	61
Şekil 52. Image_25 son işlem sonrası maske.	62
Şekil 53. Image_25 son işlem sonrası taş alanları.	62
Şekil 54. Image_25 hatalı bölgelerin tespiti ve temizliği.	63
Şekil 55. Image_25 sonuç görseli.	63
Şekil 56. PSPNet/MobilenetV2 modeli örnek sonuç görselleri.	64
Şekil 57. U-Net/MobilenetV1 modeli örnek sonuç görselleri.	65
Şekil 58. LinkNet/MobilenetV1 modeli örnek sonuç görselleri.	66

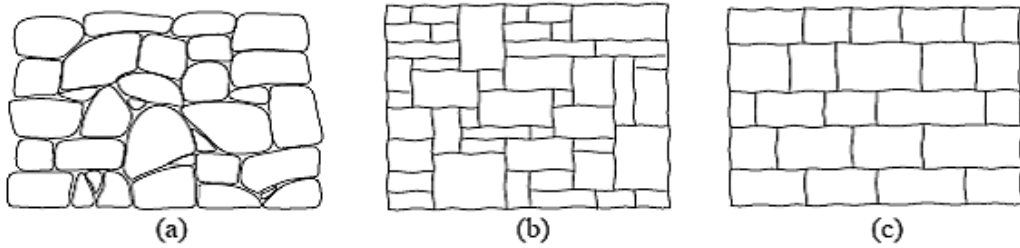
KISALTMALAR VE SİMGELER DİZİNİ

Adam	: Adaptive Moment Estimation (Uyarlanabilir Moment Tahmini)
AI	: Artificial Intelligence (Yapay Zekâ)
ANN	: Artificial Neural Networks (Yapay Sinir Ağları)
AP	: Average Precision (Ortalama Hassasiyet)
CCA	: Connected Component Analysis (Bağlantılı Bileşen Analizi)
CNN	: Convolutional Neural Networks (Evrişimli Sinir Ağları)
CPU	: Central Processing Unit (Merkezi İşlem Birimi)
DL	: Deep Learning (Derin Öğrenme)
FCN	: Fully Convolutional Networks (Tam Evrişimli Ağlar)
FLOPS	: Floating Point Operations Per Second (Saniyedeki Kayan Nokta İşlemleri)
FN	: False Negative (Yanlış Negatif)
FP	: False Positive (Yanlış Pozitif)
FPN	: Feature Pyramid Network (Özellik Piramit Ağı)
GAN	: Generative Adversarial Networks (Üretici Çekişmeli Ağlar)
GPU	: Graphics Processing Unit (Grafik İşlem Birimi)
IDE	: Integrated Development Environment (Bütünleşik Geliştirme Ortamı)
IoU	: Intersection over Union (Kesişim/Birleşim Oranı)
MBConv	: Mobile Inverted Bottleneck Convolution (Mobil Ters Darboğaz Konvolüsyonu)
mIoU	: Mean Intersection over Union (Ortalama Kesişim/Birleşim Oranı)
ML	: Machine Learning (Makine Öğrenmesi)
NAS	: Neural Architecture Search (Sinirsel Mimari Arama)
PPM	: Pyramid Pooling Module (Piramit Havuzlama Modülü)
PSPNet	: Pyramid Scene Parsing Network (Piramit Sahne Ayrıştırma Ağı)
ReLU	: Rectifier Linear Unit (Doğrultulmuş Doğrusal Birim)
RNN	: Recurrent Neural Networks (Tekrarlayan Sinir Ağları)
SAM	: Segment Anything Model (Genel Amaçlı Segmentasyon Modeli)
SE	: Squeeze and Excitation (Sıkıştırma ve Uyarma)
SOM	: Self-Organizing Maps (Öz-düzenleyici Haritalar)
TN	: True Negative (Doğru Negatif)
TP	: True Positive (Doğru Pozitif)
U-Net	: U-Net Evrişimli Sinir Ağı Mimarisi
U	: Küme Birleşimi
∩	: Küme Kesişimi

GİRİŞ

Tarihi yapılar toplumların kültürel, mimari ve sosyal tarihini günümüze taşıyan ve geleceğe aktaran değerli miras unsurlarıdır. Bu yapıların korunması kültürel kimliğin devamlılığı, tarihsel belleğin canlı tutulması ve turizme katkı sunması açısından büyük önem taşımaktadır (Icomos, 2013; Açıcı vd., 2017; Tekin, 2021). Türkiye gibi tarihi zenginliği yüksek olan ülkelerde mimari mirasın korunması daha da kritik bir hal almaktadır. Özellikle Selçuklu ve Osmanlı dönemlerinden günümüze ulaşan çok sayıda geleneksel taş yapı ve anıtsal eser ülkemizin kültürel mirasının önemli bir parçasını oluşturmaktadır (Kocaman vd., 2024). Erzurum gibi zengin bir tarihi mirasa sahip şehirlerde sıkça rastlanan kale, han, hamam, medrese ve cami gibi tarihi yığma taş yapılar hem mimari hem de kültürel mirasın önemli birer tanığıdır (Zaman vd., 2018).

Yığma taş duvarlar tarih boyunca yaygın bir inşaat tekniği olarak kullanılmıştır. Yapım tekniği açısından Şekil 1’de gösterildiği gibi genellikle taşların birbirine harç yardımıyla tutturulmasıyla oluşturulmaktadır. Yapısal olarak taşların boyutları, dizilimleri, şekilleri, harç kalitesi ve taş-harç oranı gibi pek çok parametre duvarın dayanıklılığını ve genel stabilitesini belirlemektedir (Tomazevic, 1999). Bu yapılar hem taşıyıcı sistem hem de estetik açıdan özgün nitelikler taşımakta ancak çeşitli faktörlerin etkisiyle zaman içerisinde deforme olmakta ve özgün yapısal özelliklerini kaybetmektedir.

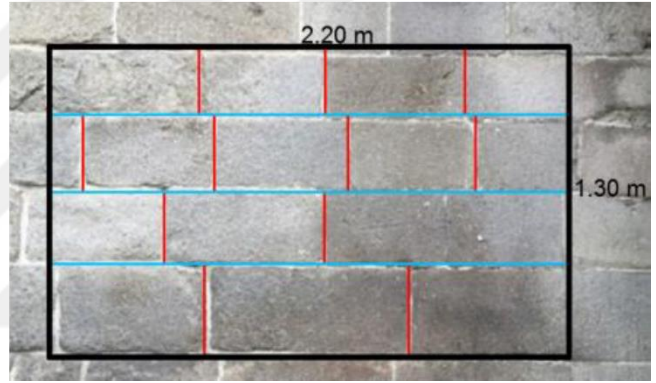


Şekil 1. Yığma Taş Yapılar: (a) Moloz taş, (b) Kesme taş, (c) İri Kesme Taş (Dabanlı, 2008)

Türkiye'nin aktif deprem kuşağında yer alması bu yapıların deprem dayanıklılığının bilimsel yöntemlerle değerlendirilmesi ve korunmasını zorunlu hâle getirmiştir (Dabanlı, 2008). Depremler ve diğer doğal afetler, yığma taş duvarların yapısal bütünlüğünü tehdit eden başlıca unsurlardan biridir (Milani, 2018). Bu nedenle yığma taş duvarların deprem güvenliğinin artırılması ve mevcut durumunun doğru şekilde tespit edilmesi hem akademik hem de uygulamaya yönelik çalışmalarda öncelikli araştırma alanlarından biri hâline gelmektedir (Tercan, 2018; Kutlu, 2024). Özellikle 2023 yılında Türkiye'de yaşanan büyük depremler

sonrasında tarihi yapıların dayanıklılığının değerlendirilmesi ve korunması konusu daha da büyük bir önem kazanmıştır (Kocaman, 2023).

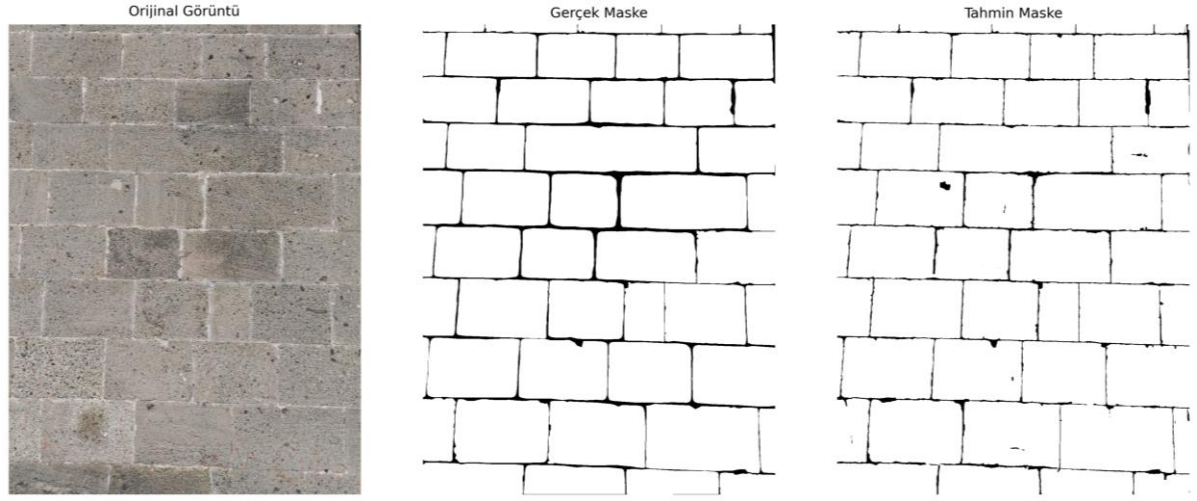
Tarihi yapıların değerlendirilmesinde ve restorasyon süreçlerinde duvarların mevcut durumunu objektif, doğru ve hızlı bir şekilde belirlemek hayati önem arz etmektedir. Geleneksel yöntemlerle yapılan duvar analizi genellikle uzmanların elle yaptıkları ölçümlere ve görsel değerlendirmelere dayanmaktadır. Bu yöntemler subjektif sonuçlara açık olduğu gibi zaman alıcı ve yüksek maliyetlidir (İbrahim vd., 2020). Ayrıca elde edilen veriler çoğunlukla sınırlı bir alanda uygulanmakta ve tüm yapının genelini yansıtmakta yetersiz kalabilmektedir. Kocaman vd. (2018) tarafından yapılan bir çalışmada Yakutiye Medresesi'nin duvarından hesaplama için 2,2 m x 1,3 m örnek alınmıştır. Ancak yapının ana boyutları 33,4 m x 24,4 m'dir. Bu durum yapılan analizin yapının genel davranışını temsil etme konusunda yetersiz kalabileceğinin somut bir göstergesidir.



Şekil 2. Yakutiye Medresesi ve ölçüm örneği (Kocaman vd., 2018)

Dijitalleşme ve yapay zekâ teknolojilerindeki hızlı gelişmeler tarihi yapıların belgelenmesi, analizi ve korunmasında yeni fırsatlar sunmaktadır. Özellikle görüntü işleme teknikleri ile makine öğrenimi ve derin öğrenme tabanlı yaklaşımlar, karmaşık yapısal problemlerin çözümünde manuel ve öznel yöntemlerin yerine nesnel, otomatik ve tekrarlanabilir analiz olanakları sağlamaktadır (Ronneberger vd., 2015; Wang vd., 2021).

Görüntü işleme bir resimdeki nesnelerin veya bölgelerin piksellerine dayalı olarak tanımlanması, ayrıştırılması ve analiz edilmesi olarak tanımlanabilir (Gonzalez ve Woods, 2018). Görüntü segmentasyonu ise bu sürecin en kritik adımlarından biridir ve bir görüntüyü anlamlı ve analiz edilebilir alt bölgelere veya nesnelere ayırma işlemi olarak tanımlanabilir (Long vd., 2015). Yığma taş duvarlar söz konusu olduğunda segmentasyon Şekil 3'te gösterildiği gibi her bir taşı ve aralarındaki harç dolgusunu (derz) dijital ortamda hassas bir şekilde ayırt etmeyi ifade eder. Bu işlem duvar dayanımının değerlendirilmesi, hasar tespiti, çatlak analizi ve restorasyon çalışmalarında kullanılacak malzeme miktarının belirlenmesi gibi birçok mühendislik uygulaması için temel bir girdi oluşturur (Valero vd., 2018).



Şekil 3. Yığma taş duvar segmentasyon örneği.

Son yıllarda yapay zekânın bir alt dalı olan makine öğrenmesi ve özellikle de derin öğrenme görüntü işleme alanında bir devrim yaratmıştır. Çok katmanlı yapay sinir ağlarını temel alan derin öğrenme modelleri geleneksel görüntü işleme tekniklerinin zorlandığı karmaşık desenleri ve özellikleri büyük veri kümelerinden otomatik olarak öğrenebilme yeteneğine sahiptir (LeCun vd., 2015; Goodfellow vd., 2016). Derin öğrenme algoritmalarının görüntü segmentasyonu alanında gösterdiği başarılar yapı mühendisliği ve mimari koruma alanlarında da yeni uygulamaların önünü açmıştır. Zhang vd. (2016) tarafından yapılan çalışmada Evrişimli Sinir Ağları'nın (CNN) yol ve beton yüzeylerde çatlak tespitinde yüksek performans sergilediği gösterilmiştir. Benzer şekilde Fan vd. (2018) yol yüzeylerinde çatlakların otomatik olarak tespiti için yapılandırılmış tahmin yöntemini ve CNN modellerini kullanmıştır. Çalışma klasik görüntü işleme yöntemlerine kıyasla CNN tabanlı modellerin çatlak tespitinde çok daha başarılı ve nesnel sonuçlar verdiğini göstermektedir. Bu ve benzeri çalışmalar derin öğrenme tabanlı segmentasyon yaklaşımlarının yapı elemanlarının analizinde ne kadar etkili olabileceğini göstermektedir.

Yığma taş duvarlarda süreklilik ve süreksizlik alanlarının (taş ve harç bölgeleri ile çatlakların) doğru tespiti, dayanım analizlerinde ve risk değerlendirmelerinde önemli bir parametre olarak öne çıkmaktadır (Valero vd., 2018). Tarihi yığma taş duvarların segmentasyonu taşların düzensiz geometrileri, farklı boyutları, yüzey dokularındaki çeşitlilik, taş ve harç malzemelerinin renk ve doku benzerliği, aşınma ve yıpranma nedeniyle oluşan bozulmalar, aydınlatma koşullarındaki değişkenlikler ve veri setlerinin yetersizliği gibi nedenlerle oldukça zorlu bir problemdir. Ayrıca farklı dönemlere ait yapılarda kullanılan malzeme ve tekniklerin çeşitliliği genel geçer bir segmentasyon modelinin geliştirilmesini zorlaştırmaktadır (İbrahim vd., 2020; Kassotakis ve Sarhosis, 2021).

Geleneksel görüntü işleme metotları bu tür karmaşık ve düzensiz yapılarda istenilen doğruluk seviyesine ulaşmakta yetersiz kalabilmektedir. Bu noktada derin öğrenme metotları bu zorlukların üstesinden gelmek için güçlü bir potansiyel sunmaktadır. Son yıllarda transfer öğrenme, veri zenginleştirme ve hibrit modeller gibi yaklaşımlar önerilmektedir (Karimi vd., 2024). Ayrıca görüntülerin yüksek çözünürlüklü ve heterojen yapıda olması gerçek dünyadan toplanan verilerde segmentasyon başarısını etkileyen faktörlerdendir.

Geliştirilen yeni metodolojiler geleneksel tekniklere göre daha hızlı, güvenilir ve kullanıcıdan bağımsız (subjektif etkiden arınmış) sonuçlar sunabilmektedir. Son çalışmalar derin öğrenme mimarilerinin transfer öğrenme, veri zenginleştirme ve değerlendirme metrikleri ile desteklenmesinin taş, harç ve çatlak ayırımında yüksek başarı sağladığını göstermektedir (Dorafshan vd., 2018; Loverdos ve Sarhosis, 2022).

Shamsabadi vd. çalışmalarında derin öğrenme temelli dönüştürücü (Transformer) mimarisiyle yığma duvarlarda çatlakların otomatik ve hassas segmentasyonu gerçekleştirilmiştir. Modelin yüksek genellebilirlik gösterdiği farklı duvar tipleri ve veri kümelerinde de etkin performans sergilediği raporlanmıştır (Shamsabadi vd., 2022).

Chen vd. çalışmalarında bina cephelerindeki çatlakların tespiti için derin öğrenme tabanlı otomatik bir segmentasyon yöntemi geliştirmiştir. Çalışmada önerilen modelin insan müdahalesi olmadan hızlı ve güvenilir çatlak tespiti sağladığı gösterilmiştir. Ayrıca bina cephelerinin rutin bakım ve izlenmesinde otomatik derin öğrenme tabanlı sistemlerin kullanılabilirliği ifade edilmiştir (Chen vd., 2021).

Literatürdeki benzer çalışmalar ile makine öğrenmesi ve derin öğrenme algoritmaları ile yığma duvarlarda tuğla, taş, harç ve çatlakların otomatik olarak segmentasyonunun başarı ile uygulandığı gösterilmiştir. Ancak Türkiye'ye özgü tarihi yapıların çok çeşitlilik göstermesi, açık ve büyük ölçekli veri setlerinin azlığı, yöntemlerin farklı mimari tipler üzerinde yeterince test edilmemiş olması ve segmentasyon sonuçlarının sahadaki doğruluklarının net olarak ortaya konmamış olması literatürdeki temel boşluklar arasında yer almaktadır. Özellikle Türkiye'deki yığma taş yapıların özgün karakteristiklerini dikkate alan, yerel malzeme ve yapım tekniklerine uyarlanmış modellerin geliştirilmesi konusunda önemli bir araştırma boşluğu bulunmaktadır.

Bu tez çalışmasının temel amacı ülkemizde bulunan yığma taş duvar yapım tekniği ile inşa edilmiş tarihi yapıların, deprem dayanımının ölçülmesinde kullanılan tekniklere görüntü işleme ve derin öğrenme temelli hibrit yaklaşımlar (derin öğrenme modellerinin geleneksel görüntü işleme filtreleriyle birleştirilmesi gibi) entegre ederek ölçümlerin doğruluğunu ve uygulanabilirliğini artırmaktır. Çalışmanın özgün katkısı ise ülkemizdeki yapılara özgü taş ve harç dokularını içeren yerel veri seti kullanılarak, U-Net, LinkNet, FPN, PSPNet gibi güncel

segmentasyon mimarilerinin; ResNet34/50, DenseNet121, MobileNetV1/V2, VGG16 ve EfficientNetB0 gibi önceden eğitilmiş omurgalar (backbone) ile transfer öğrenme yöntemiyle eğitilmesidir. Modellerin segmentasyon sonuçlarının piksel doğruluğu, Dice katsayısı ve kesişim/birleşim oranı (Intersection over Union-IoU) gibi metrikler üzerinden performanslarını ortaya koymak, ayrıca geliştirilen bu modelleri çeşitli görüntü işleme teknikleri ile destekleyerek taş boyutlarının hesaplanmasına katkı sağlamaktır.

Araştırmanın temel soruları şu şekildedir:

- Yığma taş duvarlarda taş ve harç bölgeleri derin öğrenme ve görüntü işleme yöntemleriyle ne kadar hassas ve güvenilir bir şekilde ayrılabilir?
- Farklı derin öğrenme modellerinin segmentasyon başarısı hangi metrikler ve koşullar altında farklılık göstermektedir?
- Kısıtlı hesaplama kaynakları olan uygulamalarda (mobil cihazlar, gömülü sistemler vb.) hangi model mimarisi en uygun hız-başarı dengesini sunmaktadır?
- Yerel yığma taş yapıları için oluşturulan veri seti ve geliştirilen yöntemler literatürdeki yaklaşımlara göre ne tür yenilikler sunmaktadır?

Bu tez kapsamında elde edilen bulguların yığma taş duvar segmentasyonu alanında özgün bir yöntem, veri seti ve değerlendirme çerçevesi sunarak gelecekte yapılacak çalışmalara rehberlik etmesi beklenmektedir. Çalışmanın sonuçları hem akademik literatüre katkı sağlayacak hem de tarihi yapıların korunması ve restorasyonu alanında çalışan uzmanlara pratik çözümler sunacaktır. Çalışma ile yerelde Erzurum'un genelde Türkiye'nin zengin mimari mirasını koruma çabalarına bilimsel ve teknolojik bir katkı sunulması hedeflenmektedir.

KURAMSAL TEMELLER

Temel Kavramlar

Bu bölümde araştırma konusunun kuramsal temelleri, tarihsel gelişimi ve literatürdeki önemli çalışmalar incelenecektir. Öncelikle yapay zekâ, derin öğrenme, görüntü işleme ve segmentasyon gibi temel kavramlar tarihsel gelişimleriyle birlikte açıklanacak ardından konuyla ilgili yapılmış çalışmalar, kullanılan veri setleri, yöntemler ve elde edilen sonuçlar aktarılacaktır.

Yapay zekâ ve ilişkili teknolojiler günümüz dünyasını dönüştüren en temel dinamiklerden biridir. Bu alan geniş bir kavram olan yapay zekâdan başlayarak onun bir alt dalı olan makine öğrenmesine, makine öğrenmesinin özel bir modeli olan yapay sinir ağlarına ve nihayetinde bu ağların en gelişmiş hali olan derin öğrenmeye uzanan hiyerarşik bir yapı sergiler (LeCun vd., 2015; Goodfellow vd., 2016). İç içe geçen bu küme yapısı Şekil 4’te gösterilmiştir.



Şekil 4. Yapay zekâ küme yapısı (Goodfellow vd., 2016)

Yapay zekâ

Yapay zekâ (Artificial Intelligence-AI) insan benzeri akıl yürütme, öğrenme, algılama ve problem çözme yeteneklerine sahip bilgisayar sistemlerinin geliştirilmesini amaçlayan çok disiplinli bir bilim alanıdır (Russell & Norvig, 2021). Alan Turing’in 1950’de önerdiği “Turing Testi” ile başlayan bu alan, 1956’da John McCarthy vd. tarafından Dartmouth Konferansı’nda resmen isimlendirilmiştir (Turing, 1950; McCarthy vd., 2006). Yapay zekâ; problem çözme (sezgisel arama algoritmaları), doğal dil işleme (metin anlama ve üretme), uzman sistemler (belirli alanlarda uzman bilgisini modelleyen sistemler), bilgisayarla görme (görsel verileri

anlama), robotik (fiziksel dünyada hareket ve manipölasyon) ve otomatik planlama (karmaşık görevleri sıralama) gibi çok sayıda alt alanı içerir (Nilsson, 2009).

Gelişiminde 1950-1970'lerde sembolik ve mantıksal yaklaşımlar, 1980'lerde bilgi tabanlı sistemler, 1990'lardan itibaren ise istatistiksel yöntemler ve makine öğrenmesi yaklaşımları öne çıkmıştır. Günümüzde yapay zekâ tıp, finans, ulaşım, eğitim ve sanayi gibi birçok alanda yaygın biçimde kullanılmaktadır. Tıbbi görüntü analizinde kanser tespiti (McKinney vd., 2020), doğal dil işlemede makine çevirisi ve içerik üretimi (Brown vd., 2020), bankacılık sistemlerinde sahtekarlık/usulsüzlük tespiti (Keskenler vd., 2021) ve müşteri hizmetlerinde otomatik yanıt sistemleri yapay zekânın güncel uygulama alanlarına örnektir. Yapay zekâda başarı veri kalitesi ve kullanılan algoritmanın yetkinliği bağlıdır. Ayrıca etik, güvenilirlik, açıklanabilirlik ve mahremiyet gibi sosyal konular da giderek önem kazanmaktadır (Floridi & Cowls, 2022). İlk dönemlerinde kural tabanlı (eğer-ise) sistemlerle ilerleyen bu alanın nihai hedefi, insan müdahalesi olmadan karmaşık görevleri yerine getirebilen otonom sistemler yaratmaktır. Bu hedefe ulaşmak için geliştirilen en güçlü ve yaygın yaklaşımlardan biri sistemlerin veriden öğrenmesini sağlayan Makine Öğrenmesi'dir.

Makine öğrenmesi

Makine öğrenmesi (Machine Learning-ML) bilgisayarların açıkça programlanmadan deneyim yoluyla öğrenmelerini ve gelişmelerini sağlayan algoritmik bir yaklaşımdır (Mitchell, 1997). Arthur Samuel'in 1959'daki deneyimden öğrenme yaklaşımı ile verideki örüntüleri istatistiksel tekniklerle tespit ederek tahmin veya karar verme görevlerini yerine getiren dama oyunu çalışmasıyla makine öğrenmesinin temellerini atmıştır (Samuel, 1959). Makine öğrenmesi; denetimli öğrenme (etiketli veriyle tahmin ve sınıflandırma), denetimsiz öğrenme (kümeleme, boyut indirgeme) ve pekiştirmeli öğrenme (ödül-ceza ile öğrenme) olmak üzere üç temel kategoriye ayrılır (Sutton & Barto, 1998; Bishop, 2006).

Denetimli öğrenme e-posta filtreleme ve hastalık tahmini gibi görevlerde kullanılırken denetimsiz öğrenme veri kümesindeki gizli yapıları keşfetmeye yöneliktir. Pekiştirmeli öğrenme ise oyun oynama ve robotik gibi alanlarda başarılı sonuçlar vermektedir. Günümüzde sağlık alanında hastalık teşhisi (Esteva vd., 2021), finans sektöründe risk değerlendirmesi (López de Prado, 2020), tarımda verim optimizasyonu (Liakos vd., 2018) ve enerji sektöründe tüketim tahmini (Wang vd., 2018) gibi çok çeşitli alanlarda yaygın biçimde kullanılmaktadır. Makine öğrenmesi modern yapay zekâ uygulamalarının temelini oluştursa da özellikle görüntü ve ses gibi karmaşık verilerdeki örüntüleri tanıma görevini, insan beyninin çalışma prensibinden ilham alan Yapay Sinir Ağları adlı modeller son derece başarılı bir şekilde gerçekleştirebilmektedir.

Yapay sinir ağıları

Yapay sinir ağıları (Artificial Neural Networks-ANN) insan beynindeki biyolojik nöron ağlarından ilham alınarak geliştirilen birbirine bağlı çok sayıda işlem biriminden (yapay nöron) oluşan hesaplama modelleridir (Haykin, 2009). Her bir nöron kendisine gelen ağırlıklı girişleri toplar ve bir aktivasyon fonksiyonu aracılığıyla bir çıktı üretir. Bu yapı karmaşık örüntüleri tanıma, sınıflandırma ve tahmin etme yetenekleriyle öne çıkmaktadır. İlk model McCulloch ve Pitts (1943) tarafından önerilmişse de Rosenblatt'ın algılayıcı (perceptron) modeli (1958) ile öğrenen ağların temeli atılmıştır. Tek katmanlı perceptron'un sınırlamaları 1986'da Rumelhart vd. tarafından geliştirilen ve alandaki en büyük devrim olarak kabul edilen "geri yayılım" (backpropagation) algoritmasıyla aşılmıştır (Rumelhart vd., 1986). Bu algoritma temelde ağı çıktısı ile beklenen çıktı arasındaki hatanın, ağırlıkların güncellenmesi için ağ boyunca geriye doğru yayılması prensibine dayanmaktadır. Bu yapısı ile çok katmanlı sinir ağlarının verimli bir şekilde eğitilmesinin önü açılmış oldu.

Yapay sinir ağıları; İleri Beslemeli Sinir Ağları (Feedforward Neural Networks-FNN), Tekrarlayan Sinir Ağları (Recurrent Neural Networks-RNN), Evrişimli Sinir Ağları (Convolutional Neural Networks-CNN) ve Öz-düzenleyici Haritalar (Self-Organizing Maps-SOM) gibi farklı mimarilere sahiptir. CNN'ler, LeNet, AlexNet, VGG, ResNet gibi mimarilerle görüntü işleme alanında devrim yaratmıştır (LeCun vd., 2002; Krizhevsky vd., 2012). Son dönemde transformer tabanlı yapılar özellikle doğal dil işleme ve bilgisayarla görme alanlarında geleneksel sinir ağı yapılarının çok ötesine geçen performanslar sunmaktadır (Vaswani vd., 2017; Dosovitskiy vd., 2020). Sinir ağlarının karmaşıklığı arttıkça parametre sayısı ve hesaplama ihtiyacı artmakta bu da model sıkıştırma ve verimlilik üzerine araştırmaları artırmaktadır (Hinton vd., 2015). Geri yayılım algoritması sayesinde ağlardaki katman sayısını artırarak çok daha karmaşık ve hiyerarşik temsilleri öğrenen derin modellerin yani Derin Öğrenmenin önü açılmıştır.

Derin öğrenme

Derin öğrenme (Deep Learning-DL) çok katmanlı yapay sinir ağıları kullanılarak veriden öğrenen bir makine öğrenmesi yöntemidir. Geleneksel makine öğrenmesi yöntemlerinin aksine derin öğrenme modelleri görüntü, ses ve metin gibi karmaşık verilerden özelliklerin manuel olarak çıkarılmasına gerek kalmadan bu özellikleri doğrudan veriden hiyerarşik bir şekilde öğrenir (LeCun vd., 2015). Örneğin bir görüntü tanıma sisteminde ilk katmanlar kenarlar ve dokular gibi basit özellikleri orta katmanlar şekiller ve desenler gibi daha karmaşık yapıları son katmanlar ise nesneler ve sahneler gibi yüksek seviyeli kavramları öğrenir. Bu hiyerarşik öğrenme süreci insan beyninin görsel işleme mekanizmasıyla benzerlik

göstermektedir (Yamins ve DiCarlo, 2016). Büyük veri setlerinin ve güçlü hesaplama kaynaklarının özellikle grafik işlem birimlerinin (GPU) yaygınlaşmasıyla yeni bir dönemin önu açılmıştır. AlexNet modelinin 2012'deki ImageNet yarışmasındaki ezici başarısı modern derin öğrenme devrimini başlatmıştır (Krizhevsky vd., 2012).

Günümüzde Transformer, U-Net, ResNet ve Üretici Çekişmeli Ağlar (Generative Adversarial Networks-GAN) gibi gelişmiş mimariler sayesinde derin öğrenmeye dayalı modeller; doğal dil işleme (Vaswani vd., 2017), tıbbi görüntü analizi (Irmak vd., 2022), meyve sebze sınıflandırma (Karademir vd., 2025), biyometrik tanıma (He vd., 2016), otonom araçlar (Grigorescu vd., 2020) ve kaza tespiti (Nusari vd., 2024) gibi sayısız alanda insan performansına ulaşan ve hatta onu aşan sonuçlar üretmektedir. U-Net, SegNet, FCN ve DeepLab gibi segmentasyon mimarileri özellikle tıbbi görüntü analizinde ve otonom sistemlerde piksel seviyesinde sınıflandırma için kullanılmaktadır (Chen vd., 2018). Bu yetenekleri sayesinde derin öğrenme modern yapay zekânın lokomotifi olarak kabul edilmektedir (LeCun vd., 2015).

Görüntü işleme

Görüntü işleme (Image Processing) dijital görüntüler üzerinde matematiksel işlemler uygulayarak analiz, manipölasyon ve anlamlı bilgiler çıkarılması veya otomasyon amacıyla yapılan işlemler bütünüdür (Nixon ve Aguado, 2019). Bu alan bilgisayar bilimi, elektrik mühendisliğı, matematik ve optik gibi disiplinlerin kesişiminde yer almaktadır. Görüntü işleme sayesinde bir görüntüden gürültü giderme, kenar tespiti, boyutlandırma, renk düzeltme, nesne algılama gibi çok sayıda analiz ve dönüşüm işlemi gerçekleştirilebilir (Sonka vd.; 2013; Gonzalez ve Woods, 2018). Görüntü işleme genellikle bir görüntüyü girdi olarak alıp çıktıda o görüntünün geliştirilmiş bir versiyonunu veya o görüntüden çıkarılmış karakteristik özellikleri üreterek insan algısını iyileştirmek veya otonom bir makinenin yorumlaması için görüntüyü daha uygun hale getirmeyi amaçlar (Karakoç, 2012; Szeliski, 2022).

Dijital görüntü işlemenin kökenleri 1960'larda NASA'nın uzay araştırmaları için geliştirilen ilk dijital görüntü işleme tekniklerine dayanır. Bu dönemde Ay görüntülerinin iyileştirilmesi için çoğunlukla matematiksel operasyonlar ile manuel özellik çıkarımına dayalı çeşitli algoritmalar geliştirilmiştir (Castleman, 1996). 1970'lerde Fourier dönüşümü, filtreleme ve kenar algılama gibi temel algoritmalar ortaya çıkmış ve tıbbi görüntüleme alanındaki ilk uygulamalara başlamıştır (Russ, 2006). Görüntü işleme tekniklerinden uzamsal alan teknikleri doğrudan piksel değerleri üzerinde çalışırken, frekans alanı teknikleri Fourier dönüşümü gibi yöntemlerle görüntünün frekans bileşenlerinin analiz edilmesine dayanır (Jain, 1989).

Günümüzde tıbbi görüntüleme (MRI, CT, ultrason), uzaktan algılama (uydu görüntüleri), endüstriyel kalite kontrol, güvenlik sistemleri, otomotiv uygulamaları ve dijital sanat gibi çok çeşitli alanlarda yaygın olarak kullanılmaktadır (Sonka vd., 2013). Son yıllarda derin öğrenme tabanlı yaklaşımlar geleneksel görüntü işleme tekniklerinin yerini almaya başlamıştır. Ancak klasik görüntü işleme teknikleri hala özellikle hesaplama kaynaklarının sınırlı olduğu durumlarda ön işleme aşamasında veya derin öğrenme modellerinin performansını artırmak için hibrit yaklaşımlarda kullanılmaktadır (Wang vd., 2020).

Görüntü segmentasyonu

Görüntü segmentasyonu (Image Segmentation) bir görüntünün analiz edilebilir alt bölgelere, nesnelere ya da sınıflara karşılık gelen çoklu segmentlere (piksel kümelerine) bölme işlemidir (Stockman ve Shapiro, 2001). Segmentasyonun amacı görüntüyü anlamlı parçalara ayırarak nesneleri, yapıları veya belirli bölgeleri otomatik olarak tanımlamak ve analiz edilmesi daha kolay bir forma dönüştürmektir (Pham vd., 2000). Piksel bulutlarına etiket ataması yaparak nesne tespiti için Örnek Segmentasyonu (Instance Segmentation), her piksele ayrı ayrı etiket ataması yaparak kategorik tespit için Semantik Segmentasyon (Semantic Segmentation) kullanılır (Wang vd., 2018). Görüntü işlemenin en temel ve zorlu problemlerinden biri segmentasyon olarak kabul edilmektedir. Nesne tanıma, sınıflandırma ve analiz gibi daha karmaşık görevlerin ön adımını oluşturur (Szeliski, 2022).

Görüntü segmentasyonu teknikleri; kenar tabanlı, bölge tabanlı, kümeleme, eşige dayalı, model tabanlı ve makine öğrenmesi (özellikle derin öğrenme) temelli yaklaşımlar olarak gruplandırılır (Long vd., 2015; Ronneberger vd., 2015). Son yıllarda U-Net, LinkNet, FCN ve DeepLabV3Plus gibi derin öğrenme tabanlı segmentasyon algoritmaları klasik yöntemlere göre çok daha yüksek doğruluk sağlamaktadır (Minaee vd., 2021).

Segmentasyon tıbbi görüntülerde kanserli hücreler, tümör ve organ ayrıştırılması (Bozkurt, 2023), uydu fotoğraflarında arazi sınıflandırması, otonom araçlarda yol/engel algılama (Janai vd., 2020), endüstriyel kalite kontrol (Zhang vd., 2016) ve artırılmış gerçeklik gibi pek çok uygulama alanında yaygın olarak kullanılmaktadır.

Literatür Taraması

Sithole (2015) lazer tarama ile elde edilen yüksek yoğunluklu 3D nokta bulutu verilerinden taş duvarlardaki tuğlaları otomatik olarak tespit etmek için bir segmentasyon yöntemi geliştirmiştir. Geliştirilen yöntem ile önce nokta bulutu verileri düzlemsel yüzeylere ayrılmakta ardından her bir düzlemsel yüzeyin taş ve harç bölgelerinin segmentasyonu yapılmaktadır. Çalışma sonucunda elde edilen görsellerde manuel doğrulama ile

karşılaştırıldığında tuğla sıralarının ve yerleşiminin doğru olarak tespit edildiği gösterilmiştir. Yöntemin özellikle düzgün örülmüş ve net derzlere sahip duvarlarda başarılı olduğu fakat harçların silikleştiği veya duvarda bozulmaların bulunduğu durumlarda hassasiyetin düştüğü vurgulanmıştır. Sonuçlar lazer tarama nokta bulutlarından tuğla tespiti ve duvar segmentasyonu için bu tür yaklaşımların kullanılabilirliğini ortaya koymuştur.

Ronneberger vd. (2015) ISBI hücre izleme yarışmasında biyomedikal görüntü segmentasyonu için U-Net mimarisi temelli bir derin öğrenme modeli geliştirmişlerdir. Araştırmacılar sınırlı veri setlerinde yüksek performans elde etmek için veri zenginleştirme (data augmentation) tekniklerini yoğun şekilde kullanmışlardır. Geliştirilen U-Net modeli ile “PhC-U373” veri setinde %92,03 ve “DIC-HeLa” veri setinde %77,56 doğruluk oranları elde edilmiş ve rakiplerine ciddi bir fark atarak yarışmada birinci olmuşlardır. U-Net mimarisi küçük veriyle bile güvenilir segmentasyon sağlamıştır. Özellikle tıbbi görüntü segmentasyonu alanında altın standart haline gelmiş ve sonraki birçok çalışmada temel mimari olarak kullanılmıştır.

Long vd. (2015) görüntü segmentasyonu için ilk kez tam evrişimli ağırları (Fully Convolutional Networks-FCN) önermişlerdir. Çalışmada PASCAL VOC 2011-2012, NYUDv2 ve SIFT Flow veri setleri kullanılmıştır. Araştırmacılar klasik CNN mimarilerine kıyasla görüntünün uzamsal bilgisini koruyabilen ve piksel seviyesinde sınıflandırma yapabilen bir mimari geliştirmişlerdir. FCN mimarisi VGG16, AlexNet ve GoogLeNet gibi önceden eğitilmiş ağırları kullanarak transfer öğrenme yaklaşımını benimsemiş ve VGG16 modeli ile segmentasyon performansını yüksek oranda artırmıştır. Çalışma sonucunda PASCAL VOC 2011 veri setinde %90,3 piksel doğruluğu ve %62,7 ortalama IoU (mean Intersection over Union-mIoU), NYUDv2 veri setinde %65,4 piksel doğruluğu ve %34,0 mIoU, SIFT Flow veri setinde %85,2 piksel doğruluğu ve %39,5 mIoU değerleri elde edilmiştir.

Zhang vd. (2016) çalışmalarında yollardaki çatlakların tespiti için 500 adet yüksek çözünürlüklü yol yüzeyi görüntüsünü, 99x99 piksellik daha küçük parçalara ayırarak bir veri seti oluşturmuş ve Derin Evrişimli Sinir Ağı (ConvNet) temelli bir model geliştirmişlerdir. Geliştirdikleri ConvNet modeli %86,96 doğruluk ve %89,65 F1 skor, SVM modeli ise %81,12 doğruluk ve %73,59 F1 skor elde edilmiştir. Bu çalışma ile ConvNet tabanlı derin öğrenme modellerinin zorlu koşullar altında başarılı sonuçlar ürettiği gösterilmiştir.

Chaurasia ve Culurciello (2017) çalışmalarında Cityscapes veri seti (2975 eğitim, 500 doğrulama ve 1525 test görüntüsü) ve CamVid veri seti (367 eğitim, 101 doğrulama ve 233 test görüntüsü) ile gerçek zamanlı segmentasyon için LinkNet mimarisini kullanmışlardır. Araştırmacılar U-Net'e benzer bir yapıya sahip olan ancak daha az parametre ile daha hızlı

alışan modelleri iin kodlayıcı olarak ResNet18 mimarisini kullanmış ve zellikle gerek zamanlı uygulamalar iin uygun bir mimari sunmuşlardır. alışma sonucunda Cityscapes veri setinde %76,4 IoU ve CamVid veri setinde %68,3 IoU performansı elde edilmiştir. LinkNet mimarisi zellikle sınırlı hesaplama kaynaklarına sahip cihazlarda (mobil cihazlar, gömülü sistemler) kullanılabilecek etkili bir segmentasyon özümü sunmuştur ve hız-doğruluk dengesi aısından önemli avantajlar sağlamıştır.

Cha vd. (2017) beton yüzeylerdeki atlakları tespit etmek iin CNN tabanlı bir yaklaşım önermişlerdir. alışmada araştırmacılar tarafından 332 adet yüksek özünürlüklü (4928×3264 piksel) görüntü daha küçük paralara ayrılarak 40000'den fazla görüntü ieren bir veri seti oluşturulmuştur. Bu veri seti farklı aydınlatma koşulları, arka planlar ve atlak türleri ieren gerek beton yüzey görüntülerinden oluşmaktadır. Araştırmacılar AlexNet mimarisine dayalı bir CNN modeli geliştirmiş, görüntüyü 256×256 piksel boyutunda paralara ayrılarak her bir parada atlak olup olmadığını sınıflandırmaktadır. alışma sonucunda 32 bin görüntü ile yapılan eğitimde %98,22 doğruluk, 8 bin görüntü ile yapılan eğitimde %97,95 doğruluk elde edilmiştir. Ayrıca modelin gerek dünya koşullarında farklı aydınlatma ve arka planlarda bile atlakları başarıyla tespit edebildiği gösterilmiştir.

Chen vd. (2018) alışmalarında PASCAL VOC 2012 veri seti (1464 eğitim, 1449 doğrulama ve 1456 test görüntüsü) ve Cityscapes veri seti üzerinde DeepLabV3Plus mimarisini kullanmıştır. Araştırmacılar kodlayıcı olarak Xception ve ResNet101 gibi farklı mimarileri kullanarak zellikle nesne sınırlarının daha kesin bir şekilde segmente edilmesini sağlamışlardır. alışma sonucunda PASCAL VOC 2012 veri setinde %89,0 mIoU ve Cityscapes veri setinde %82,1 mIoU deęerleri elde edilmiştir. DeepLabV3Plus mimarisi zellikle nesne sınırlarının hassas bir şekilde segmente edilmesi gereken uygulamalar iin etkili bir özüm sunmuştur ve günümüzde hala en iyi performans gösteren segmentasyon mimarilerinden biri olarak kabul edilmektedir.

Yang vd. (2018) beton yapılarıdaki atlakları tespit etmek ve ölçmek iin FCN tabanlı bir model geliştirmişlerdir. alışmada araştırmacılar tarafından oluşturulan farklı genişliklerde atlaklar ieren ve piksel seviyesinde etiketlenmiş 800'den fazla beton yüzey görüntüsü ieren bir veri seti kullanılmıştır. Araştırmacılar geliştirdikleri modeli transfer öğrenme yaklaşımıyla her bir pikseli atlak ya da sağlam olarak sınıflandırabilen ve atlak genişliğini otomatik ölçebilen şekilde eğitmişlerdir. alışma sonucunda %97,96 doğruluk, %79,95 F1 skor, atlak uzunluęunda -48,03% ile 177,79% arasında ve atlak genişliği ölçümünde ise -13,27% ile 24,01% arasında hata oranlarında sonuçlar elde edilmiştir. Ayrıca modelin gerek zamanlı

uygulamalar için yeterince hızlı çalıştığı ve mobil cihazlarda bile kullanılabileceği gösterilmiştir.

Dorafshan vd. (2018) beton yapılarda çatlakların tespiti için AlexNet DCNN mimarisine dayalı bir derin öğrenme tabanlı algoritma kullanmıştır. Araştırmada beton köprülerin, duvarlarının ve kaldırımlarının yüzeylerindeki çatlakları içeren 56000 görüntüden oluşan SDNET2018 veri seti kullanılmıştır. Sıfırdan eğitilen ve transfer öğrenimi ile eğitilen yöntemler karşılaştırılmıştır. Köprü, duvar ve kaldırımlarda sıfırdan eğitilen model ile sırasıyla %90,45, %87,54, %94,86 transfer yöntemi ile eğitilen modelde %91,92, %89,31, %95,52 doğruluk oranları elde edilmiştir. Çalışma transfer öğrenimi modellerinin sıfırdan eğitilen modellere kıyasla gerek eğitim süresi gerek doğruluk oranlarında daha başarılı sonuçlar sunduğunu göstermektedir.

Özgenel ve Sorguç (2018) çalışmalarında beton çatlaklarının tespiti ve bina yüzeylerinde yapısal sağlık izlemesi için görüntü tabanlı önceden eğitilmiş derin öğrenme modellerinin hangisinin daha etkili olduğunu belirlemek amacıyla farklı modelleri karşılaştırmışlardır. ODTÜ Kampüsü'ndeki çeşitli beton binaların duvar ve zeminlerinden alınan 40000 adet 224x224 piksel görüntü içeren veri seti üzerinde yapılan testlerde oluşturulan model ile %90,9 ortalama hassasiyet (Average Precision-AP) ve %91,9 F1 skor elde edilmiştir. Bu sonuç ile FCN tabanlı segmentasyon modellerinin çatlak tespiti ve hasarın geometrisini (genişlik, uzunluk) anlamada başarılı olduğu kanıtlanmıştır.

İbrahim vd. (2019) yaptıkları çalışmada farklı tip (rastgele, kare, harçsız, düzgün ve kaba kesme taş) duvar görüntüsünü içeren ve elle etiketlenmiş 162 adet görsel ile oluşturulan bir veri seti kullanmıştır. Model eğitim performansını arttırmak için döndürme, kaydırma, kırpma ve yakınlaştırma gibi veri zenginleştirme işlemleri uygulamışlardır. Çalışmada U-Net tabanlı model ile referans alınan iki metot (Riveiro ve Osés) karşılaştırılmıştır. Geliştirilen model ile %81,57 F1 skor elde edilmiş ve diğer yöntemlerin (%23,65, %22,58) oldukça ilerisinde sonuçlar alınmıştır. Model tahminlerindeki eksik veya belirsiz harç bölgelerinin giderilmesi için ayrıca havza (Watershed) algoritması uygulayarak hibrit bir yaklaşım geliştirilmiştir. Klasik bağlantı bileşen analizi (CCA) ile %78,63 F1 skor elde edilirken bu yaklaşım ile %81,23 F1 skor elde edilmiş ve modelin başarılı olduğu kanıtlanmıştır.

Dung (2019) çalışmasında beton çatlaklarının piksel düzeyinde segmentasyonu için VGG-16 mimarisini temel alan bir FCN modeli önermiştir. 4032x3024 piksellik 458 adet yüksek çözünürlüklü fotoğraftan oluşan halka açık veri seti, 40000 adet 227x227 piksellik görüntüye çevrilerek yapılan testlerde oluşturulan model ile %89,3 ortalama hassasiyet elde

edilmiştir. Bu sonuç ile FCN tabanlı segmentasyon modellerinin çatlak tespiti ve hasarın geometrisini (genişlik, uzunluk) anlamada başarılı olduğu kanıtlanmıştır.

Valero vd. (2019) tarihi duvarlardaki bozulmaları hem tekil taşlar hem de genel duvar deseni düzeyinde otomatik olarak tespit etmek için makine öğrenmesi tabanlı bir yöntem geliştirmiştir. İskoçya'daki Stirling Kalesi'nde bulunan Kraliyet Şapeli (Chapel Royal) binasının güney cephesine ait duvarlarından lazer tarama ile yüksek çözünürlüklü 3D nokta bulutu verileri toplanmış ve fotoğraf makinası ile çekilmiş 347 adet görüntü tarama verisiyle birleştirilerek veri seti oluşturulmuştur. Çalışma sonucunda lojistik regresyon modeli kullanılarak yapılan segmentasyon ile taşların sınırlarının belirlenmesinde ve kusurlu alanların (erozyon, delaminasyon, mekanik hasar gibi) tespitinde başarılı sonuçlar elde edildiği çeşitli görseller ile gösterilmiştir.

Ibrahim vd. (2020) yaptıkları çalışmada farklı duvar tiplerinde 512x512 piksel boyutunda 532 adet sentetik ve gerçek duvar görüntüsünü içeren kendi oluşturdukları veri setini kullanmıştır. Çalışmanın ana amacı duvar görüntülerindeki engelleri ve çeşitli nesneleri (bitki örtüsü, tabela) kaldırıp bulundukları bölgeleri sanal olarak tamamlayarak duvarın orijinal yapısını tahmin etmek ve yüksek doğrulukta tuğla harç ayrımı yapmaktır. İlk olarak U-Net mimarisine dayalı geliştirilmiş modeli kullanılarak giriş görüntüsü tuğla, harç ve engellenmiş bölge olarak üç ana sınıfa ayrılmıştır. Sonrasında iki aşamalı Üretken Çekişmeli Ağ (Generative Adversarial Network-GAN) modeli kullanılarak görüntüdeki engellenmiş bölgeler kaldırılmış ve gözlemlenen duvar desenine uygun şekilde doldurulmuştur. Son aşamada watershed algoritmasına dayalı bir segmentasyon işlemi gerçekleştirilerek ham görüntüden tuğla harç ayrımı yapılmıştır. Gerçek ve engelsiz görüntülerde %74,18 F1 skor ve %63,78 IoU, gerçek ve engelli görüntülerde %67,44 F1 skor ve %54,04 IoU değerlerine ulaşılmıştır. Ayrıca dolgu aşamasındaki başarı kıyası için kendi modellerinin ve EdgeConnect yaklaşımının sonuçlarını içeren 93 kişinin katıldığı görsel anket gerçekleştirmiştir. Elde edilen %58,59'a karşı %41,41 sonucuna göre yöntemlerinin başarılı olduğu ifade edilmiştir.

Hatır ve İnce (2021) yaptıkları çalışmada tarihi yığma yapılara ait farklı taş türlerinin sınıflandırmasına yönelik Mask R-CNN/ResNet101 tabanlı bir model geliştirmiştir. Eğitim ve test veri seti için Konya ilinde bulunan çeşitli tarihi yığma yapılardan topladıkları 1800 adet 800x600 piksellik görüntüler kullanılmıştır. Çalışma sonucunda farklı taş türleri için %89,10 ile %100 arasında ortalama hassasiyet değerine ulaşıldığı ifade edilmiştir.

Mousavi ve Bakhshi (2022) yığma yapılarda çatlak tespiti için 115 görüntü içeren kendi oluşturdukları veri setini kullanmıştır. Dice katsayısı kayıp (Dice-Coefficient loss) fonksiyonu kullanarak eğitim yapılan U-Net/EfficientNetB3 modeli ile test verileri üzerinde %81,44

hassasiyet, %71,41 duyarlılık ve %75,36 F1 skor elde edilmiştir. Sınırlı eğitim verisi ve karmaşık görsellere rağmen başarılı sonuçlar alındığı ifade edilmiştir.

Loverdos ve Sarhosis (2022) çalışmalarında duvar analizini önce tuğla harç ayrımı yapıp sonra çatlakları tespit ederek iki temel aşamada gerçekleştiren bütünsel bir yaklaşım sunmaktadır. Kendi çektikleri 107 adet duvar fotoğrafından 224x224 piksel boyutunda 2814 adet görsel ile veri seti oluşturup U-Net, DeepLabV3Plus, LinkNet ve FPN gibi derin öğrenme mimarilerine dayalı modeller geliştirmiştir. Tuğla segmentasyonu için DeepLabV3Plus/Xception modeli ile %96,86 doğrulama doğruluğu ve %96,29 F1 skor değerlerine ulaşarak basit görüntü işleme (eşikleme) yöntemlerine göre özellikle aydınlatma ve renk değişimlerinin olduğu karmaşık görüntülerde çok daha üstün sonuçlar verdiği gösterilmiştir. Çatlak tespitinde U-Net/MobileNet tabanlı model ile %79,6 F1 skor, %79,9 duyarlılık (recall) ve %81,4 kesinlik (precision) değerlerine ulaşmıştır.

Ravichand vd. (2022) yaptıkları çalışmada taş duvarlardaki çatlak tespiti için yüksek çözünürlüklü duvar görüntülerinden elde edilen 96x96 piksellik 6002 adet görsel kullanmıştır. 2 konvolüsyon katmanı ve aralarında maksimum havuzlama ile çıkışında tam bağlantılı katman kullanılarak oluşturulan CNN modeli ve bu modelin çıkış katmanına SVM uygulanarak iki farklı model geliştirilmiştir. Test verilerinde CNN modeli ile %67,50 doğruluk ve %73,00 F1 skor, CNN+SVM modeli ile %74,90 doğruluk ve %78,00 F1 skor elde edilmiştir.

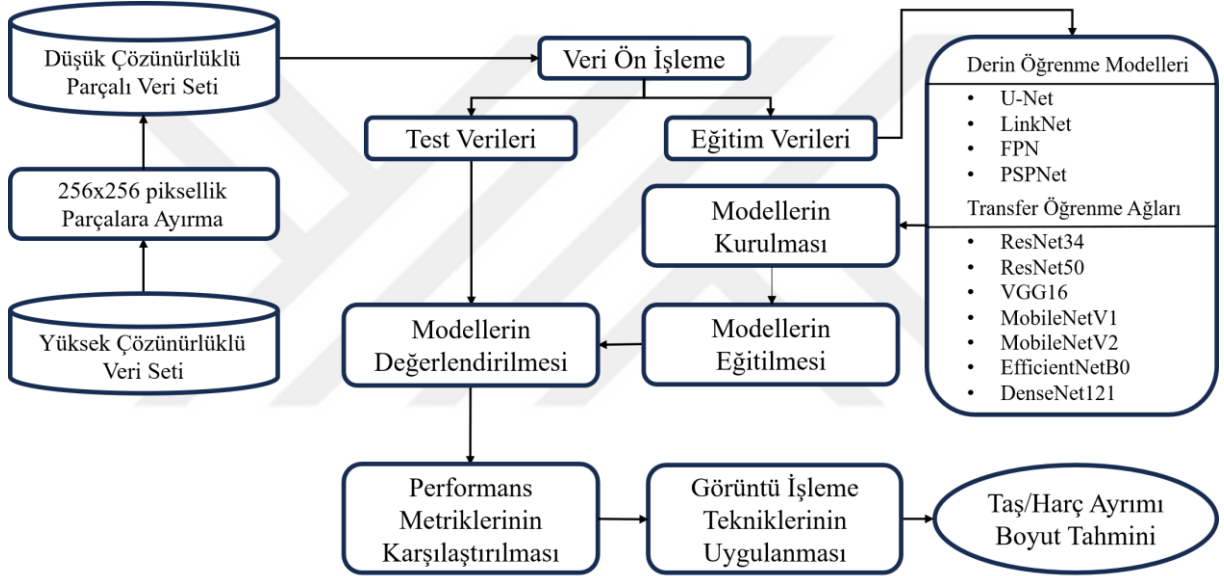
Bu tez çalışması, Türkiye'deki tarihi yığma taş yapıların özgün morfolojik özelliklerini dikkate alan ve bu alandaki yerel veri seti eksikliğini hedef alan bir yaklaşımla literatürdeki önemli bir boşluğu doldurmaktadır. Literatürdeki benzer çalışmalar genellikle tuğla veya betonarme yapılar üzerine odaklanmakta ya da Türkiye'nin zengin ve çeşitli mimari mirasını temsil eden yerel veri setlerinden yoksun kalmaktadır. Bu çalışmanın en temel farkı ve avantajı, kamuya açık bir veri setini Erzurum ilindeki tarihi yapılardan toplanan yerel ve yüksek çözünürlüklü görüntülerle birleştirerek hibrit bir veri seti oluşturması ve bu veri seti üzerinde U-Net, FPN, LinkNet, PSPNet gibi dört farklı mimariyi yedi farklı omurga (backbone) ile sistematik olarak karşılaştırmasıdır. Ayrıca derin öğrenme modellerinin ham çıktılarını morfolojik işlemler, Watershed algoritması ve bağlantılı bileşen analizi gibi geleneksel görüntü işleme teknikleriyle iyileştiren hibrit bir metodoloji sunması ile çalışma literatürdeki pek çok yaklaşımdan ayrılmaktadır. Bu son işlem adımları model tahminlerindeki gürültü ve süreksizlik gibi kusurları önemli ölçüde gidererek segmentasyon doğruluğunu artırmaktadır. Bir diğer avantajı ise model performansını değerlendirirken, taş ve harç pikselleri arasındaki sayısal dengesizliği göz önünde bulundurarak, genel metriklerin yanı sıra harç tespitindeki başarıyı ölçen özgüllük (specificity) gibi sınıfa özgü metriklere de odaklanmasıdır. Çalışmanın

dezavantajı veya kısıtlılığı, yerel veri setinin Erzurum ilindeki yapılarla sınırlı olması nedeniyle geliştirilen modellerin farklı coğrafi bölgelerdeki taş yapılar üzerindeki genelleme performansının test edilmemiş olmasıdır. Buna rağmen çalışma, yerel mimari özelliklere odaklanarak yüksek başarıya ulaşan modeller sunmasıyla literatüre özgün bir katkı sağlamaktadır.



MATERYAL VE METOT

Bu bölümde yığma taş duvar görüntülerinden taş ve harç bölgelerinin segmentasyonu amacıyla yürütülen araştırmanın metodolojisi ve kullanılan materyaller açıklanmaktadır. Araştırma için hazırlanan görüntü veri seti üzerinde derin öğrenme tabanlı farklı modellerin eğitilmesi, segmentasyon performanslarının nicel metriklerle değerlendirilmesi ve sonrasında uygulanan görüntü işleme teknikleri ile taş boyutlarının hesaplanmasını içeren deneysel ve tanımlayıcı bir nitelik taşımaktadır. Çalışmada kullanılan yöntemle ait akış diyagramı Şekil 5'te gösterilmiştir.



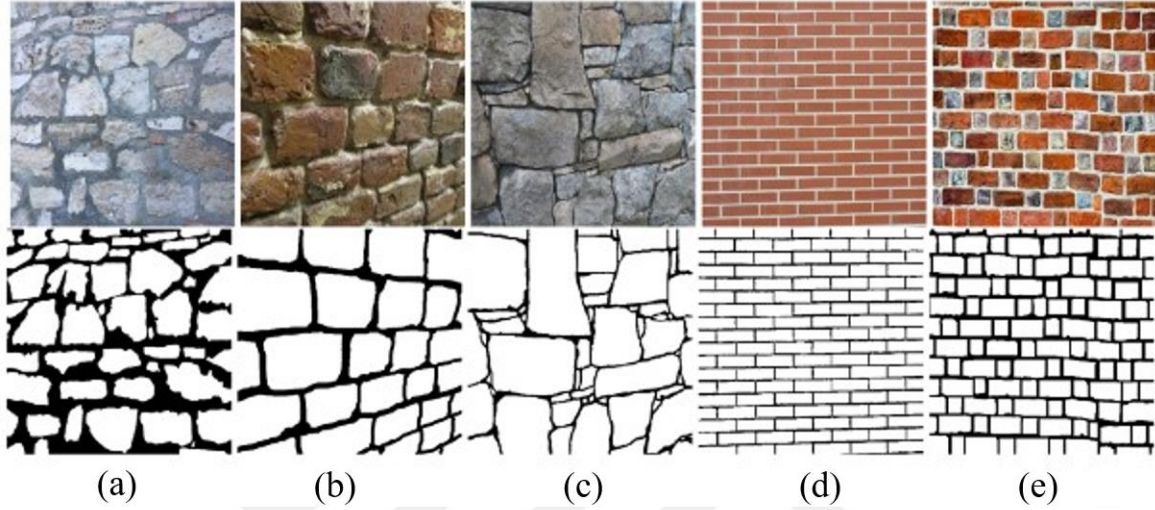
Şekil 5. Çalışma akış diyagramı.

Derin öğrenme modellerinin eğitim sürecindeki başarısı ve hassasiyeti büyük ölçüde kullanılan veri setlerinin kalitesine bağlıdır (Minaee vd., 2021). Özellikle yüksek doğruluk ve hassasiyet gerektiren taş/harç segmentasyonu gibi zorlu görevlerde, zengin içerikli ve hassas bir şekilde etiketlenmiş veri setlerinin kullanılması kritik bir öneme sahiptir. Duvar görüntüsü verileriyle ilgili yaygın zorluklar arasında yetersiz görüntü çözünürlüğü, etiketleme hataları ve veri setinin yeterli çeşitliliğe sahip olmaması yer almaktadır. Çeşitlilikten yoksun bir veri seti ile eğitilen modellerin farklı mimari veya çevresel koşullara sahip veriler üzerinde genelleme yapma yeteneği oldukça kısıtlanmakta ve modelin başarısı düşmektedir (Örs ve Özçelik, 2025).

Veri Seti

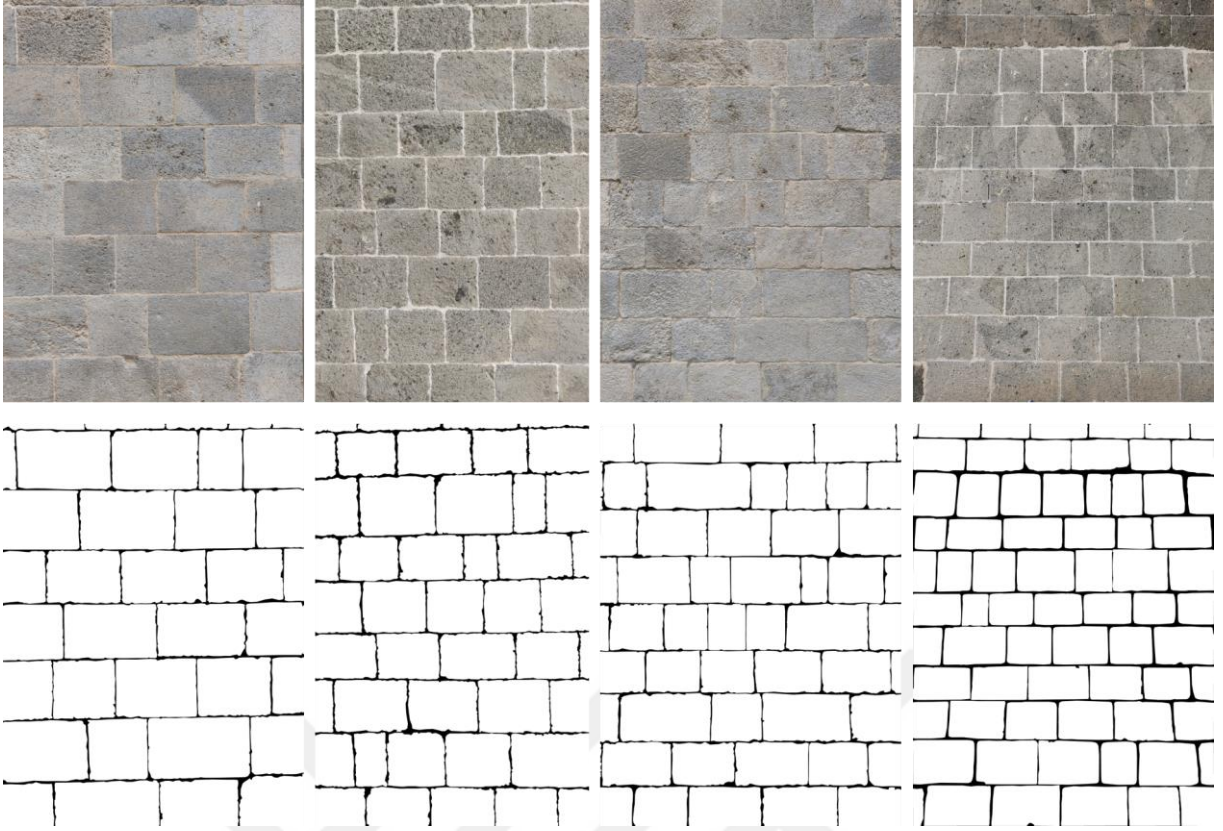
Çalışmada modellerin performansını ve genelleme yeteneğini arttırmak için iki farklı veri seti bir arada kullanılmıştır. İlk olarak Ibrahim vd. (2020) tarafından yayımlanan çalışmaya

ait “Masonry Wall Image Analysis Benchmark” isimli veri seti kullanılmıştır. Bu veri seti farklı açılardan çekilmiş çeşitli taş türlerini içeren yığma taş duvar görüntülerinden oluşmaktadır. Veriler yığma duvar görüntüsü ve bu görüntülerin maskelerini içerir, eğitim için 230 adet görüntü ve maske çifti, doğrulama için 123 adet görüntü ve maske çifti olmak üzere etiketlenmiş toplam 353 görüntü ve maske çiftinden oluşmaktadır. Veri setindeki görsellere ait örnekler Şekil 6’da gösterilmiştir.



Şekil 6. Veri setindeki örnek duvar görselleri (a) Rastgele moloz taş duvar, (b) Kare moloz duvar, (c) Kuru moloz duvar, (d) İnce kesme taş, (e) Kaba kesme taş (İbrahim vd., 2020)

Kullanılacak veri setini yerel örnekler ile zenginleştirmek için ikinci olarak Erzurum ilinde bulunan (Yakutiye Medresesi, Çifte Minareli Medrese, Erzurum Ulu Camii ve Lala Mustafa Paşa Camisi) çeşitli tarihi yapıların yığma taş duvarlarına ait farklı açılardan çekilmiş yüksek çözünürlüklü dijital görüntülerden ve bu görüntülerdeki taş alanlarının manuel olarak etiketlendiği maske görüntülerinden oluşturulan veri seti kullanılmıştır. Veri setindeki görsellere ait örnekler Şekil 7’de gösterilmiştir.



Şekil 7. Yerel veri seti örnek duvar görselleri ve etiketleri.

Görüntülerin etiketli maskelerinin oluşturulması için web tabanlı “SuperAnnotate” uygulamasından faydalanılmıştır. Öncelikle otomatik olarak taş sınırları belirlenmiş ancak detayların istenilen seviyede olmaması sebebiyle manuel olarak düzeltmeler yapılarak gerçeğe en yakın ve doğruluğu yüksek taş/harç maskeleri hazırlanmıştır. Maskeler taş bölgeleri beyaz (255) ve harç bölgeleri siyah (0) olacak şekilde ikili maske formatında oluşturulmuştur. Eğitim için 134 adet görüntü ve maske çifti, doğrulama için 34 adet görüntü ve maske çifti, test için 8 adet görüntü ve maske çifti olmak üzere etiketlenmiş toplam 176 yüksek çözünürlüklü görüntü ve maske çiftinden oluşmaktadır.

Kullanılacak modellerin etkin bir şekilde eğitilebilmesi için giriş katmanlarına uygun görüntü boyutlarının seçilmesi gerekmektedir. Uygun boyutlar belirlenirken segmentasyon için kritik öneme sahip detayların kaybolmamasına ve işlem ihtiyacını minimum düzeyde arttıracak değerler seçilmesi gerekir. Çalışmada kullanılacak modellerin giriş katmanına uygun çözünürlükte 256x256 boyutunda görseller kullanılmıştır. Ancak veri setinde bulunan yüksek çözünürlüklü (6144x8192, 4182x6144 vb.) orijinal duvar görüntüleri ve maskeleri uygun boyutlara indirilmesi esnasında taş/harç ayrımında kritik öneme sahip bilgilerin kaybolmaması için, Zhang vd. tarafından önerilen yaklaşım ile 256x256 piksel boyutlarında daha küçük parçalara (patch) ayrılmıştır (Zhang vd., 2022). Bu işlem sonucunda veri ve detay kaybı yaşanmadan toplam 108194 görüntü ve maske çifti elde edilmiştir. Oluşturulan parçalı veri setinin %80’i eğitim, %20’si ise doğrulama aşamasında kullanılmıştır.

Veri Ön İşleme

Makine öğrenimi ve derin öğrenme modellerinin geliştirilmesinde en kritik ve gerekli aşamalardan biri ham verinin işlenerek modelin kullanımına uygun hale getirildiği veri hazırlama aşamasıdır. Elde edilecek sonuçların güvenilirliği ve doğruluğu açısından temel bir öneme sahiptir (Maharana vd., 2022). Bu çalışmada geliştirilen segmentasyon modellerinin performansını iyileştirmek ve farklı veri setleri üzerindeki genelleme yeteneğini güçlendirmek amacıyla çeşitli ön işleme ve veri artırma tekniklerinden faydalanılmıştır. Bu dönüşümler için açık kaynak kodlu “alumentations” kütüphanesi aracılığıyla veri işleme adımları (pipeline) tanımlanmış ve eğitim esnasında anlık olarak uygulanmıştır.

Ön işleme yöntemlerinden biri de görüntüyü standart hale getirmektir. İlk adımda tüm resimleri aynı boyuta indirgemek için görseller 256x256 piksel olarak yeniden ölçeklendirilmiştir. Sonrasında piksel değerleri 0-255 arasında tamsayılarla RGB görüntülerde sunulduğundan modellerin eğitim hızını artırmak ve daha stabil çalışması için tüm görüntü verilerinin ve maskelerinin piksel değerleri ikili (binary) formatta 0 ile 1 aralığına normalize edilmiştir. Eğitim veri setindeki çeşitliliği yapay olarak artırmak (data augmentation) ve modelin aşırı öğrenmesini (overfitting) önlemek amacıyla yatay-dikey çevirme, açısız döndürme, ölçekleme, kaydırma, yakınlaştırma, parlaklık ve kontrast ayarı gibi bazı geometrik ve fotometrik dönüşümler rastgele olarak uygulanmıştır. Ayrıca modelin genelleme yeteneğini arttırmak için karıştırma (shuffle) yapılmıştır. Bu teknikler sayesinde model aynı görüntünün farklı varyasyonlarını görerek çeşitli özellikleri daha başarılı öğrenme yeteneği kazanmıştır (Shorten ve Khoshgoftaar, 2019). Doğrulama verilerine ise modelin ham performansını ölçmek amacıyla veri artırma uygulanmamıştır.

Segmentasyon Mimarileri

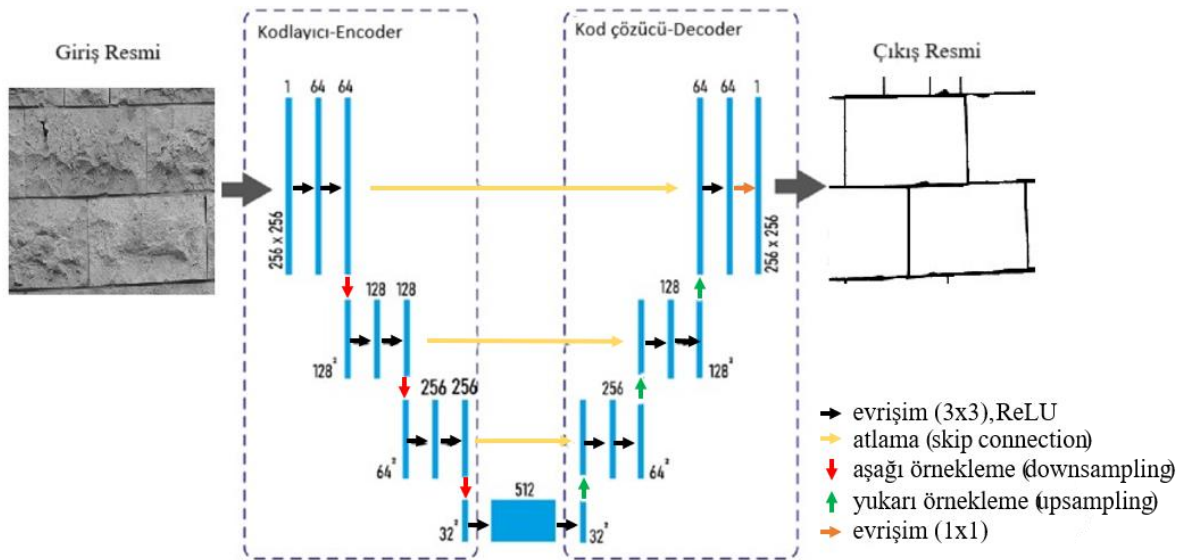
Bu çalışmada yığma taş duvarlardaki taş ve harç bölgelerini otomatik olarak ayırmak için güncel ve yaygın olarak kullanılan ayrıca farklı yaklaşımlar sunan modern derin öğrenme tabanlı segmentasyon mimarileri kullanılmıştır. Modelin görüntüden noktasal ve bütünsel özellikleri en verimli şekilde çıkarmasını sağlayan kodlayıcı (encoder) ve bu özellikleri kullanarak piksel bazında sınıflandırma haritası üreten kod çözücü (decoder) bölümlerinden oluşan bu mimariler farklı stratejilerle optimize edilmiştir. Çalışma kapsamında U-Net, FPN, LinkNet ve PSPNet olmak üzere dört farklı segmentasyon mimarisi değerlendirilmiştir.

Bu mimarilerde transfer öğrenmeden yararlanmak amacıyla çeşitli önceden eğitilmiş (pre-trained) omurga (backbone) ağları kullanılmıştır. Bu mimarilerin performansını ve özellik çıkarma kapasitesini belirleyen kodlayıcı kısımları için ResNet, DenseNet, MobileNet, VGG

ve EfficientNet gibi literatürde başarısını kanıtlamış farklı modeller kullanılmıştır. Bu çeşitlilik kıyaslama yaparak görev için en uygun mimari kombinasyonunu belirlemeyi amaçlamaktadır.

U-Net

U-Net özellikle biyomedikal görüntü segmentasyonunda popülerleşen bir derin öğrenme mimarisidir ve adını "U" harfine benzeyen simetrik yapısından alır. Çalışmada kullanılan U-Net modelinin yapısı Şekil 8’de gösterilmiştir.



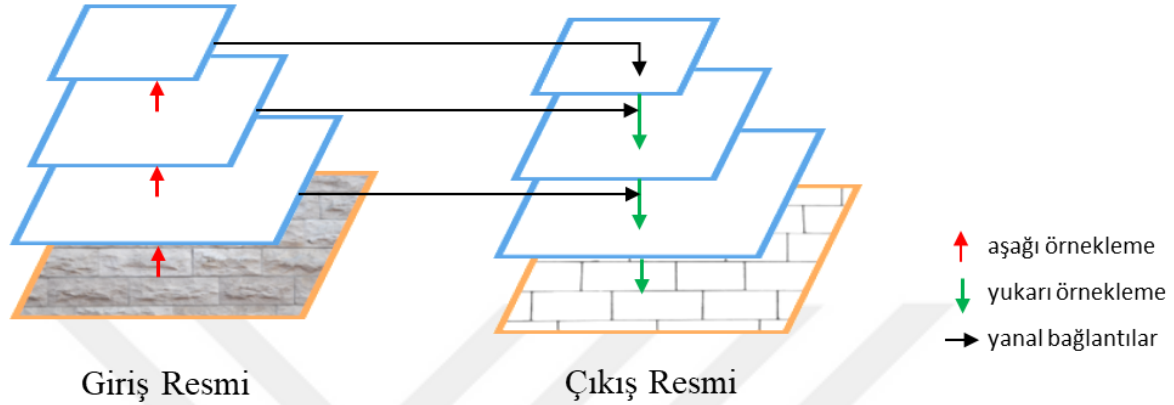
Şekil 8. U-Net mimarisi (Ronneberger vd., 2015)

Bir kodlayıcı (encoder) ve bir kod çözücü (decoder) bölümünden oluşur (Ronneberger vd., 2015). Kodlayıcı standart bir evrişimli sinir ağı gibi çalışarak görüntüyü sıkıştırıp düşük çözünürlüklü ve yüksek öznitelikli bir hale getirir ve anlamsal özellik haritaları çıkarır. Kod çözücü ise bu sıkıştırılmış özellik haritasını aşamalı olarak büyütür ve orijinal görüntü boyutunda bir segmentasyon haritası oluşturur. U-Net'in en ayırt edici özelliği kodlayıcı ve kod çözücü katmanları arasındaki atlama bağlantılarıdır (skip connections). Bu bağlantılar kodlayıcıdaki aşağı örnekleme (downsampling) aşamasında yüksek çözünürlüklü ve ince detayları içeren öznitelikleri doğrudan kod çözücüdeki yukarı örnekleme (upsampling) aşamasında karşılık gelen katmana taşıyarak segmentasyon sınırlarındaki hassasiyeti artırır ve bilgi kaybını en aza indirir. Bu yapıyla az sayıda veriyle ve detaylı segmentasyon gereken durumlarda başarılı sonuçlar üretebilir (Zhou vd., 2018; Falk vd., 2019).

FPN (Feature Pyramid Network)

Özellik Piramit Ağı (Feature Pyramid Network- FPN) nesneleri farklı ölçeklerde tespit ve segmente etme görevlerinde çok ölçekli özelliklerin etkin bir şekilde kullanılmasını sağlayan bir mimaridir (Lin vd., 2017). Şekil 9’da gösterildiği üzere geleneksel yaklaşımların aksine tek bir ölçekteki özellik haritasını kullanmak yerine bir özellik piramidi oluşturur. Kodlayıcı ağın

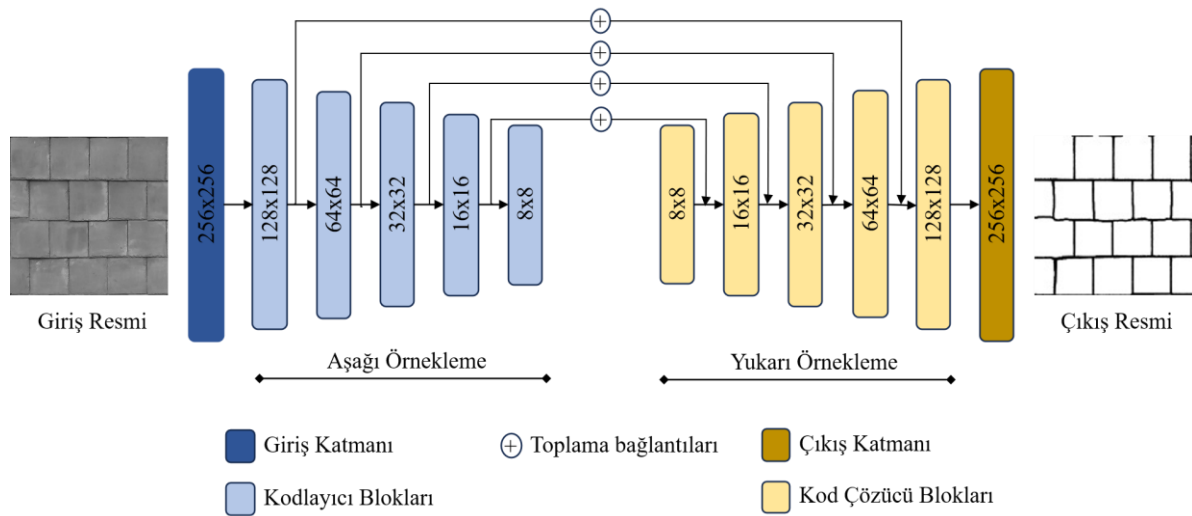
farklı derinliklerinden (düşük seviye ve yüksek seviye) alınan özellik haritalarını birleştirir. Bu işlemleri temelde bir encoder ve ona paralel yukarıdan aşağıya (top-down) bir yol ve yanal bağlantılar (lateral connections) kullanarak yapar (Seferbekov vd., 2018). Bu sayede hem düşük seviyeli (kenarlar, dokular) hem de yüksek seviyeli (anlamsal bağlam) zengin özellik bilgisini tüm ölçeklerde koruyarak taş ve tuğla gibi farklı boyutlardaki nesnelerin segmentasyonunda daha tutarlı sonuçlar üretir.



Şekil 9. FPN mimarisi (Lin vd., 2017)

LinkNet

Özellikle hız ve verimlilik gerektiren gerçek zamanlı uygulamalar için tasarlanmış hafif bir mimaridir. U-Net gibi bir kodlayıcı-kod çözücü (encoder-decoder) yapısına dayanır (Chaurasia & Culurciello, 2017). Ancak LinkNet mimarisinin temel farkı kod çözücündeki her bir bloğun kodlayıcıdaki karşılık gelen bloğun doğrudan girdisini almasıdır. Modelin mimari yapısı Şekil 10’da gösterilmiştir.



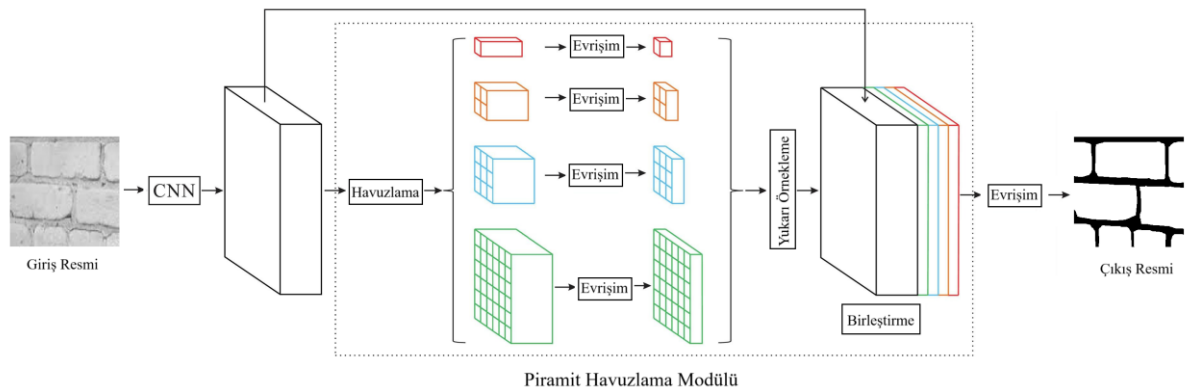
Şekil 10. LinkNet mimarisi (Chaurasia & Culurciello, 2017)

Bu basit ama etkili bağlantı şeması sayesinde gradyan akış kolaylaşır ve U-Net'in atlama bağlantılarındaki gibi özellik birleştirme (concatenation) işlemi yerine toplama (add) yaparak

parametre sayısı azaltılır. Bu sayede gerçek zamanlı segmentasyon senaryoları için hafif, hızlı ve az parametrelili tasarımıyla CamVid ve Cityscapes gibi veri setlerinde yüksek performans sağlayabilir (Chaurasia & Culurciello, 2017).

PSPNet (Pyramid Scene Parsing Network)

Piramit Sahne Ayırıştırma Ağı ilk olarak Zhao vd. tarafından 2017 yılında yayımlanan "Pyramid Scene Parsing Network" başlıklı makale ile sunulmuştur. Özellikle karmaşık sahnelerin anlaşılması ve piksel bazında etiketlenmesi (scene parsing) için piramit havuzlama modülü (Pyramid Pooling Module-PPM) kullanan bir anlamsal segmentasyon mimarisidir. Mimarinin temel amacı geleneksel Tam Evrişimli Ağların (FCN) küresel bağlam bilgilerinin yetersiz kullanımı sorununu çözmektir (Zhao vd., 2017). Farklı boyut ve görünümlere sahip nesnelerin bir arada bulunduğu karmaşık sahnelerde sadece yerel bilgiye dayalı tahminler yapmak nesnelerin yanlış sınıflandırılmasına neden olabilir. PSPNet mimarisinde bu sorun piramit havuzlama modülü (PPM) yaklaşımı ile çözülmektedir (Yang ve Guo, 2022).



Şekil 11. PSPNet mimarisi (Zhao vd., 2017)

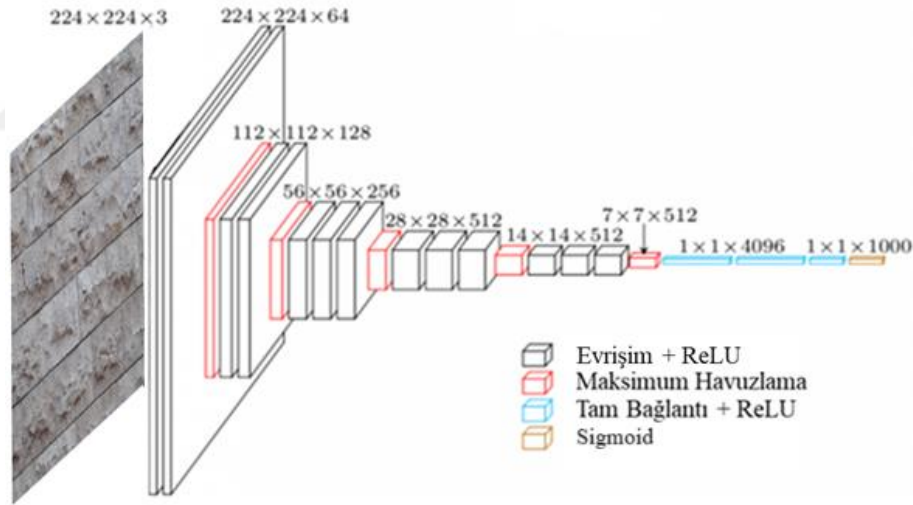
Şekil 11’de gösterilen yapıda öncelikle encoder aşamasında CNN kullanılarak öznetelik haritası çıkarılır. PPM modülünde bu özellik haritasına farklı katmanlarda 1x1, 2x2, 3x3 ve 6x6 gibi çok ölçekli paralel piramit havuzlama (pyramid pooling) işlemi uygulanır. Elde edilen sonuçlar 1x1evrişim katmanından geçirilerek boyutları azaltılıp parametre sayısı dengelenir. Farklı ölçekteki öznetelikler orijinal görüntü boyutuna yükseltilir (upsample) ve kanal bazında birleştirilerek (concat) son evrişim katmanına gönderilir. Araştırmacılar derin ağlarda gradyanların daha etkili bir şekilde geriye yayılmasına yardımcı olarak modelin eğitim sürecini optimize eden yardımcı kayıp (auxiliary loss) fonksiyonunu geliştirmiştir. Bu yaklaşımın daha derin ağlarda daha fazla fayda sağlayacağı belirtilmiştir (Zhao vd., 2017).

Piramit Havuzlama Modülü sayesinde bir sahnedeki nesneler hem yerel detaylarıyla hem de sahnenin genel bağlamı içindeki farklı ölçeklerdeki ilişkileriyle birlikte analiz edilebilir

(Durmuş vd., 2024). Bu yapıyla PSPNet özellikle karmaşık ve kalabalık sahnelerin segmentasyonunda oldukça güçlü ve başarılı bir mimari sunmaktadır (Hisaria vd., 2024).

VGG16

Oxford Üniversitesi'nin Visual Geometry Group (VGG) laboratuvarında Simonyan ve Zisserman tarafından 2014 ILSVRC yarışması için geliştirilen ve derin öğrenmenin yaygınlaşmasında öncü rol oynayan klasik bir CNN mimarisidir (Simonyan ve Zisserman, 2014). VGG16'nın temel yaklaşımı karmaşık büyük filtreler yerine yalnızca 3x3 boyutunda konvolüsyon katmanları ve aralara yerleştirilen maksimum havuzlama (max pooling) işlemlerini art arda kullanarak derinliği artırmaktır. İki adet 3x3 evrişim katmanının üst üste kullanılması 5x5 boyutunda tek bir evrişim katmanı ile aynı etkin alıcı alana (effective receptive field) sahiptir. Benzer şekilde üç adet 3x3'lük katman 7x7'lik bir katmana eşdeğerdir. Ancak bu şekilde katmanlara ayırmak aralara daha fazla doğrultulmuş doğrusal birim (ReLU: Rectified Linear Unit) aktivasyon fonksiyonu eklenmesine olanak tanıyarak modelin öğrenme kapasitesini artırır ve 2x2'lik maksimum havuzlama (max-pooling) katmanı ile parametre sayısı düşürülür (Simonyan ve Zisserman, 2014). Model mimarisi Şekil 12'de gösterilmiştir.

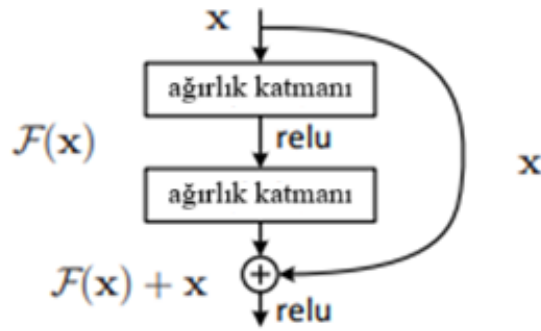


Şekil 12. VGG16 model mimarisi (Loukadakis vd., 2018)

VGG16'nın oldukça düzenli ve homojen yapısı daha kolay optimize edilmesine olanak tanır bu sayede transfer öğrenmeye oldukça yatkın bir modeldir. Ancak tam bağlantılı katmanları ve derinlikli yapısı sebebiyle çok yüksek parametre sayısına ulaşır. Parametre sayısının fazlalığı hem eğitim hem de tahmin süreçlerinde yüksek hesaplama gücü (FLOPS) gerektirir ve modeli hantal bir yapıya sokar (Loukadakis vd., 2018). Toplamda 16 ağırlıklı katmandan oluşur (13 konvolüsyon + 3 tam bağlı (fully connected) katman) bu sebeple VGG16 ismi verilmiştir.

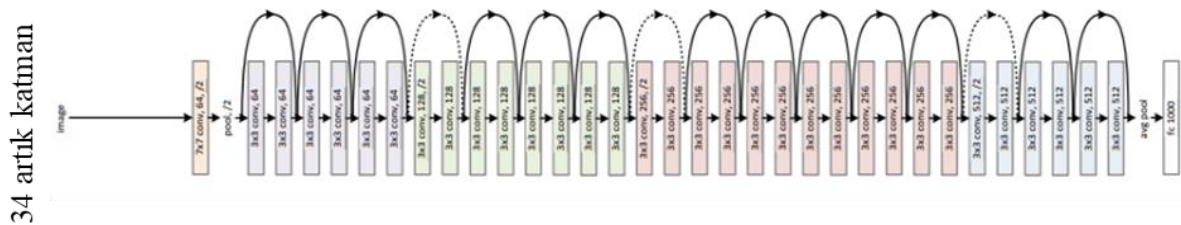
ResNet34/50

Artık Ağlar (Residual Networks, ResNet) He vd. tarafından 2016 yılında sunulan "Deep Residual Learning for Image Recognition" makalesi ile tanıtılmıştır ve derin öğrenme tarihinde bir dönüm noktası olarak kabul edilir. CNN eğitiminde karşılaşılan temel sorunlardan biri olan kaybolan gradyan (vanishing gradient) problemine çözüm getirmek amacıyla geliştirilmiştir (He vd., 2016). Konvansiyonel sinir ağlarında katman sayısı arttıkça gradyanlar geri yayılım sırasında hızla sıfıra yaklaşarak öğrenmeyi zorlaştırır ve modelin performansını sınırlar. ResNet mimarisinin en önemli yeniliği olan artık bloklar (residual blocks) ve bu bloklarda yer alan atlama bağlantıları (skip connections) sayesinde bu sorun aşılmıştır (Szegedy vd.,2017).



Şekil 13. Artık Blok Yapısı (He vd., 2016)

Artık bloklarda bir katmanın çıktısı aynı zamanda ilerideki katmanların girişine atlama bağlantıları ile doğrudan eklenir. Şekil 13'te yapısı gösterilen bu yöntem sayesinde her katman doğrudan bir haritalama öğrenmek yerine girdisi ile çıktısı arasındaki farkı yani "artık" kısmı öğrenmeye odaklanır. Böylece ağ derinleşse bile gradyanlar doğrudan daha erken katmanlara iletilerek bilgi kaybı minimize edilir ve aynı zamanda öğrenme hızlanır. ResNet mimarisi ile artık öğrenme (residual learning) adı verilen bu yenilikçi yaklaşım deneysel olarak 100'den fazla katmana sahip ağların dahi stabil ve yüksek doğrulukla eğitilebileceği gösterilmiştir (Huang vd., 2016). ResNet34 model mimarisi Şekil 14'te gösterilmiştir.



Şekil 14. ResNet34 mimarisi (He vd., 2016)

ResNet34 ve ResNet50 sırasıyla 34 ve 50 katmanlı iki popüler varyasyondur. ResNet34 daha hafif ve hızlıdır sınırlı hesaplama gücü veya daha küçük veri setlerinde başarılıdır. ResNet50 ise daha derin yapısı sayesinde karmaşık görsel özellikler daha iyi öğrenilebilir ve büyük veri setlerinde yüksek performans gösterir. Bu ağlar ImageNet gibi büyük görüntü veri

setleri üzerinde eğitildikleri için zengin ve genel geçer özellikler öğrenmişlerdir. Bu özellikler segmentasyon gibi farklı görevlerde transfer öğrenimi yaklaşımı ile modelin daha hızlı yakınsamasını ve daha iyi performans göstermesini sağlar (Tajbakhsh vd., 2016). ResNet mimarisinin sunduğu bu yenilikler DenseNet ve EfficientNet gibi modern ve karmaşık mimarinin temelini oluşturan bir ilham kaynağı olmuştur.

DenseNet121

Yoğun Bağlantılı Evrişimsel Ağlar (DenseNet) Huang vd. tarafından "Densely Connected Convolutional Networks" makalesi ile ilk olarak 2017 yılında tanıtılmıştır. ResNet'in getirdiği atlama bağlantısı fikrini daha da ileriye taşıyan bir yaklaşımdır. DenseNet mimarisinin getirdiği temel yenilik ağ içindeki her katmanın kendisinden önceki tüm katmanların çıktısına doğrudan bağlantı sağlamasıdır. Bu yoğun bağlantı yapısı özelliklerin yeniden kullanımını en üst düzeye çıkararak bilgi akışını güçlendirir ve kaybolan gradyan sorununu daha da azaltır. Özelliklerin tekrar tekrar kullanılması imkân veren bu tasarım ağı daha az parametre ile daha kompakt ve verimli olmasını sağlar (Huang vd., 2017).



Şekil 15. DenseNet mimarisi (Huang vd., 2017)

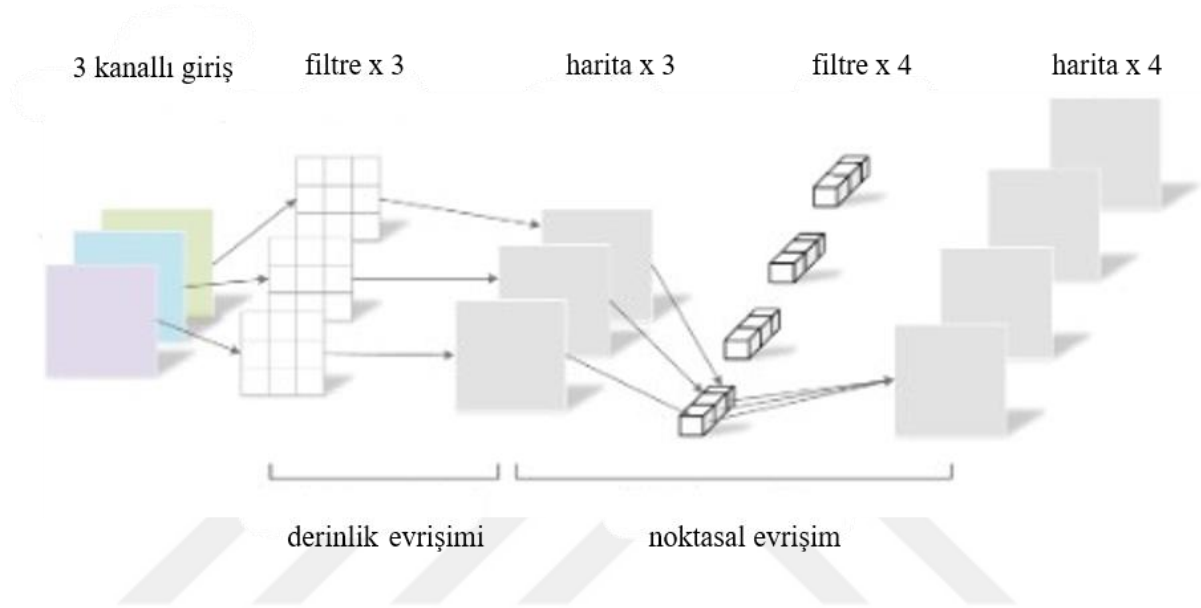
Şekil 15'te gösterildiği gibi DenseNet'in temel yapı taşı yoğun blok (dense block) olarak adlandırılır. Bir yoğun blok içinde her bir katman kendisinden önceki tüm katmanların çıktılarını girdi olarak alır. ResNet'teki gibi kendisinden bir önceki katmandan öznetelikleri toplamak yerine bu özellik haritaları kanal boyutunda birleştirilir (concatenation). Yoğun bloklar arasında özellik haritalarının boyutunu küçültmek ve kanal sayısını kontrol altında tutmak için geçiş katmanları (transition layers) bulunur. Bu katmanlar genellikle 1x1 evrişim (kanal sayısını azaltmak için darboğaz katmanı) ve ardından 2x2 ortalama havuzlama (average pooling) katmanından oluşur.

DenseNet mimarisi ile CIFAR-10, CIFAR-100, SVHN ve ImageNet gibi çeşitliliği yüksek veri setleri üzerinde başarılı sınıflandırma sonuçları alındığı gösterilmiştir (Huang vd., 2017). DenseNet121 bu mimarinin 121 katmanlı popüler bir versiyonudur ve genellikle ResNet'e kıyasla benzer doğruluk seviyelerine daha az parametre ile ulaşabilir.

MobileNetV1/V2

Howard vd. tarafından 2017'de tanıtılan MobileNet mimarisi derin öğrenme modellerini yüksek hesaplama gücü gerektiren sunuculardan çıkartarak, akıllı telefonlar ve gömülü

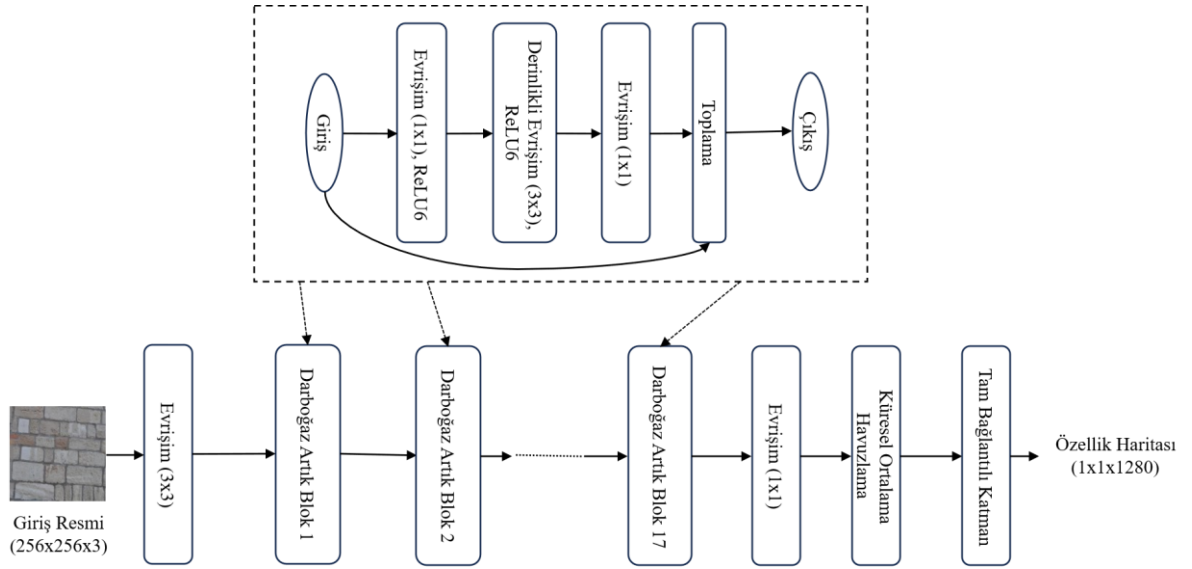
sistemler gibi kaynakları kısıtlı platformlara taşımayı ve gerçek zamanlı çalışmaya uygun hesaplama verimliliği yüksek modeller geliştirmeyi hedeflemiştir (Howard vd., 2017). Bu modellerin temel amacı model doğruluğundan minimum düzeyde ödün vererek, model boyutu ve gecikme süresi gibi hesaplama maliyetini arttıran metrikler arasında optimal bir denge kurmaktır. MobileNet mimarisinin bu dengeyi kurmak için getirdiği yenilik klasik evrişim işlemini daha az maliyetli iki ayrı adıma bölerek gerçekleştiren derinlik bazlı ayrılabilir evrişim (depthwise separable convolution) tekniğidir. Katman yapısı Şekil 16’da gösterilmiştir.



Şekil 16. Derinlik bazlı ayrılabilir evrişim katman yapısı (Zhao vd., 2024)

Standart evrişimde hem kanallar arası hem de kanal boyutunda ağırlık paylaşımı varken derinlik bazlı ayrılabilir evrişimde önce her kanala ayrı ve bağımsız şekilde bir filtre uygulanır (depthwise) ve çıkarılan özellikler diğer kanallara aktarılmaz. Ardından kanal bazında birleştirici noktasal evrişim (pointwise) uygulanır ve bu adımda kanallar arasındaki bilgi birleştirilerek yeni özellik haritaları oluşturulur (Howard vd., 2017). Bu yaklaşım parametre sayısını ve hesaplama maliyetini çok ciddi oranda azaltır ancak katmanlar arasındaki iletişimin ve özellik çeşitliliğinin azalması doğruluğu düşürebilir (Zhao vd., 2024).

2018 yılında tanıtılan MobileNetV2 mimarisi ile orijinal MobileNet üzerine ek olarak iki temel yenilik, ters artık bloklar (inverted residuals) ve doğrusal darboğazlar (linear bottlenecks) kavramları getirilmiştir (Sandler vd., 2018).



Şekil 17. MobilenetV2 katman şeması (Sandler vd., 2018)

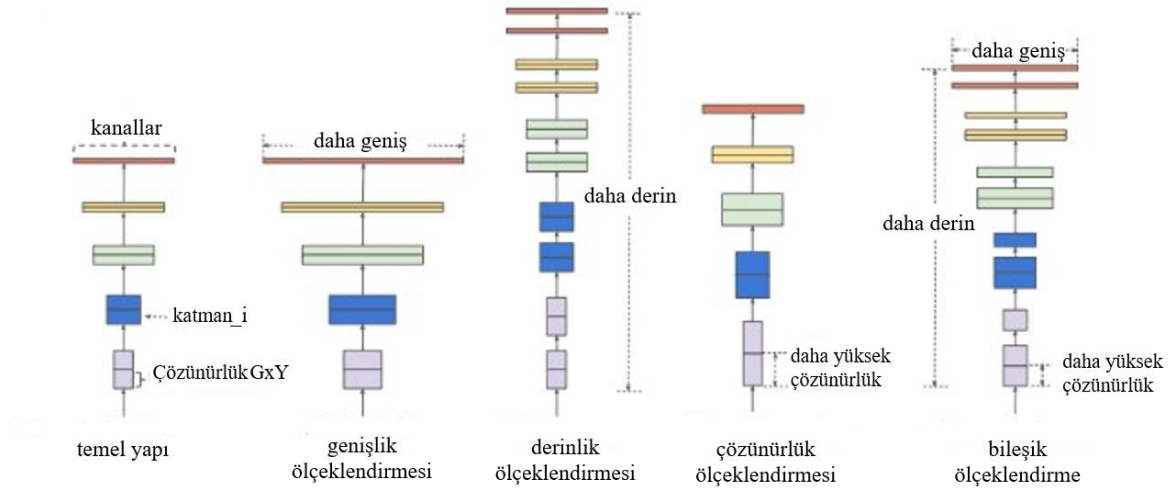
ResNet'teki artık bloklar önce kanal sayısını azaltan (darboğaz) sonra işlem yapan ve son olarak kanal sayısını artıran bir "dar-geniş-dar" yapıya sahiptir (He vd., 2016). MobileNetV2 ise bu yapıyı "geniş-dar-geniş" olacak şekilde tersine çevirir. Modelin katman şeması Şekil 17'de gösterilmiştir. Blok önce 1x1'lik bir evrişimle kanal sayısını artırır (genişletme katmanı) ardından 3x3'lük derinlik-bazlı evrişimi bu genişletilmiş alan üzerinde uygular ve son olarak 1x1'lik bir evrişimle kanal sayısını tekrar azaltır. Bu yaklaşım özelliklerin daha yüksek boyutlu bir uzayda işlenmesine olanak tanıyarak modelin öğrenme gücünü artırır ve ResNet'teki gibi atlama bağlantılarıyla birleştirildiğinde daha iyi performans gösterir (Dong vd., 2020).

Geleneksel olarak evrişim katmanlarından sonra ReLU aktivasyon fonksiyonları kullanılır. Ancak geliştiriciler ReLU'nun düşük boyutlu uzaylarda (dar katmanlarda) bilgi kaybına neden olabildiğini gözlemlemiştir. Bu nedenle MobileNetV2 dar katmanlardan sonraki projeksiyon katmanlarının ReLU aktivasyonunu kaldırır ve bilginin katmanlar arasında kaybolmasını önleyerek modelin genel doğruluğunu artırır. Bu sayede daha az kayıpla daha hızlı ve verimli bir ağ oluşturulur (Sandler vd., 2018). MobileNet mimarileri yenilikçi yaklaşımları ile sınırlı kaynaklara sahip sistemlerde hız ve verimliliğin kritik olduğu birçok görev için endüstri standardı haline gelmiştir (Khasoggi vd., 2019).

EfficientNetB0

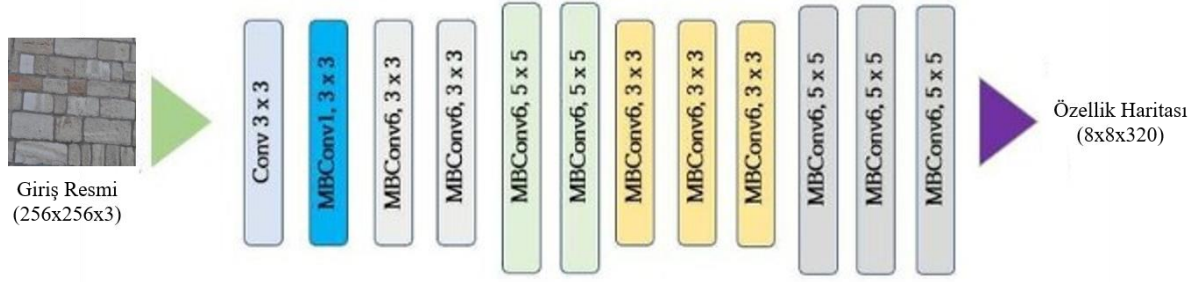
EfficientNet ailesi model doğruluğu ve verimliliği arasında daha iyi bir denge kurmayı amaçlayan ve ölçeklendirme (scaling) stratejisine yenilikçi bir yaklaşım getiren bir mimaridir (Tan ve Le, 2019). Önceki yaklaşımlarda modellerin performansını yükseltmek için genellikle ağırlık üç boyutundan biri; modelin genişliği (width) yani katmanlardaki kanal sayısı, modelin derinliği (depth) yani katman sayısı veya giriş çözünürlüğü (resolution) ayrı ayrı artırılmaktadır.

Derinliğin artması ResNet gibi çok derin ve karmaşık özellikli ağlarda kaybolan gradyan sorunu ve optimizasyon zorlukları nedeniyle modellerin eğitilmesini zorlaştırır (He vd., 2016). Genişliğin artması modellerin daha hassas özellikler öğrenmesini sağlar ve eğitilmesini kolaylaştırır. Ancak aşırı geniş ve sık ağlar yüksek seviyeli soyut özellikleri yakalamakta zorlanır. Çözünürlüğün artması modelin daha ince detayları tespit etmesine yardımcı olur. Ancak bu hesaplama maliyetini (FLOPS) karesel olarak artırır ve getirdiği doğruluk kazancı hızla azalır (Liu vd., 2022). EfficientNet'in tasarımında bileşik ölçeklendirme (compound scaling) olarak adlandırılan yöntem ile birbirinden bağımsız olmadığı ifade edilen bu üç unsur dengeli ve orantılı bir şekilde ölçeklendirilir. Geliştirilen bu yenilikçi yöntemde boyutları dengeli ölçeklendirmek için harcanabilir hesaplama kaynağına bağlı grid arama ile belirlenen sabit katsayılar kullanılır. Bileşik ölçeklendirme yöntemi Şekil 18'de gösterilmiştir.



Şekil 18. Bileşik ölçeklendirme (Tan ve Le, 2019)

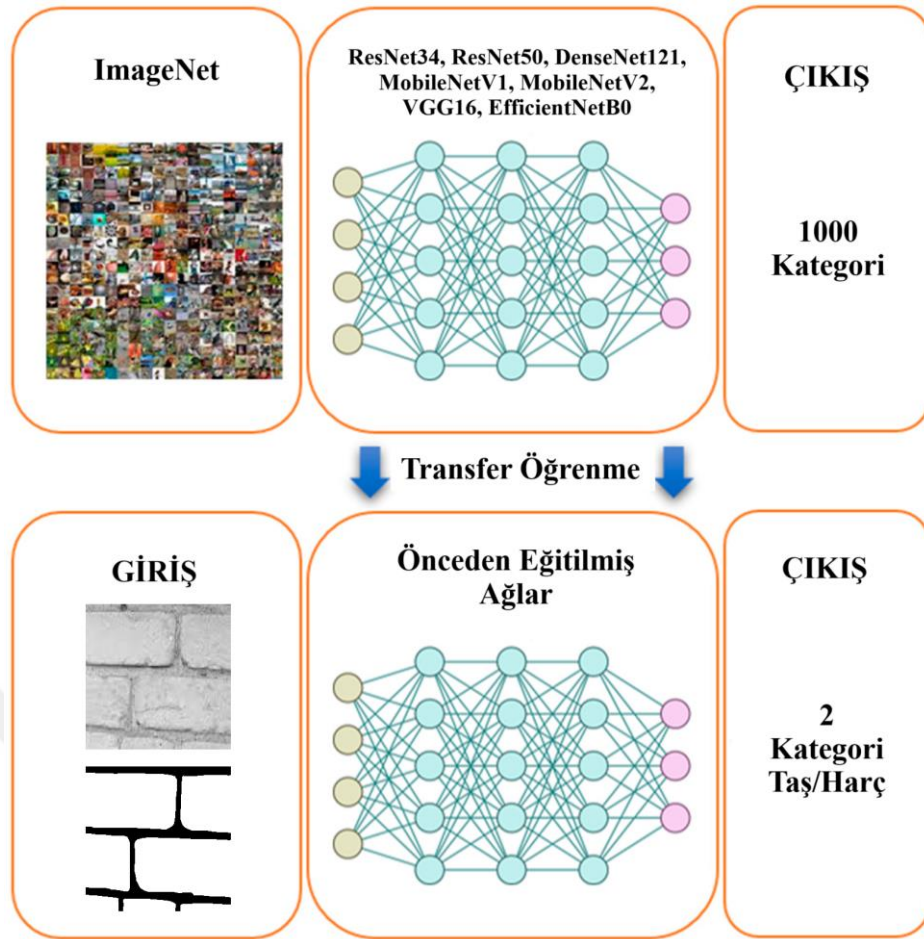
EfficientNetB0 bu ailenin en temel ve kompakt modelidir. Ağ mimarisi hem doğruluk hem de verimliliği (FLOPS cinsinden) optimize etmek için sinirsel mimari arama (Neural Architecture Search-NAS) kullanılarak oluşturulmuştur. Şekil 19'da gösterildiği şekilde EfficientNet mimarisinde bileşik ölçeklendirme tekniği ile genişlik, derinlik ve çözünürlük belirlenmiş katsayılarla dengeli şekilde ölçeklendirilir. Mobil ters darboğaz evrişim (Mobile Inverted Bottleneck Convolution-MBConv) blokları kullanılır ve bloklara sıkıştırma ve uyarma (Squeeze and Excitation-SE) optimizasyonu uygulanır (Tan ve Le, 2019; Hisaria vd., 2024). Geliştirilen bu mimariler ile parametre sayısı ve hesaplama maliyeti minimize edilirken yüksek doğruluk elde edilebilir. Örneğin ImageNet veri seti üzerinden yapılan karşılaştırmada EfficientNet-B7 önceki en iyi model olan Gpipe'tan 8,4 kat daha küçük ve 6,1 kat daha hızlıyken benzer bir doğruluk oranı elde etmiştir (Tan ve Le, 2019). Bu verimlilik onu hem araştırma hem de endüstriyel uygulamalar için son derece popüler bir omurga mimarisi haline getirmiştir.



Şekil 19. EfficientNet model mimarisi (Hisaria vd., 2024)

Transfer Öğrenme

Makine öğrenmesi ve derin öğrenme yaklaşımlarında her bir görev için modellerin sıfırdan rastgele başlatılmış ağırlıklarla eğitilmesi, özellikle modelin milyonlarca parametre içerdiği derin öğrenme uygulamalarında iki temel zorluk ortaya çıkarır; büyük miktarda etiketli veri ihtiyacı ve bu verilerle eğitim yapabilmek için yüksek hesaplama yapabilen donanım ihtiyacı yani yüksek hesaplama maliyeti. Ancak birçok gerçek problemde binlerce veya milyonlarca büyüklükte ve etiketli bir veri setine ulaşmak mümkün değildir. Az sayıda veriyle derin sinir ağlarını sıfırdan eğitmeye çalışmak ise modelin veriyi ezberlemesine ve yeni veriler karşısında genelleme yapamamasına sonuç olarak aşırı öğrenmeye (overfitting) yol açar (Pan ve Yang, 2009). Bu noktada transfer öğrenme yaklaşımı daha önce büyük ve kapsamlı ImageNet gibi veri kümeleri üzerinde eğitilmiş bir modelin daha küçük veya özel bir veri kümesinin eğitimde başlangıç noktası olmasını sağlar. Bu yöntem ile kullanılacak model önceden öğrenilen temel görsel öznitelikleri (kenar, doku, renk ve temel şekiller vb.) sıfırdan öğrenme maliyetinden kurtarılır ve doğrudan daha karmaşık özellikleri öğrenmesi sağlanarak sınırlı veri setlerinde bile başarılı sonuçlar gösterebilir (Shin vd., 2016). Hacıfendioğlu vd. (2023) yaptıkları çalışmada CNN modeli ile %80 başarıya ulaşıldığını, aynı modeli transfer öğrenme yaklaşımı ile eğiterek başarı oranının %95'e yükseldiğini göstermiştir.



Şekil 20. Transfer öğrenme.

Görüntü segmentasyonu işlemlerinde transfer öğrenme yaygın olarak kodlayıcı omurgaları (backbone) olarak isimlendirilen önceden eğitilmiş ağların kullanılması şeklinde uygulanır. Tez çalışması kapsamında ImageNet veri seti üzerinde önceden eğitilmiş ResNet34, ResNet50, DenseNet121, MobileNetV1, MobileNetV2, VGG16 ve EfficientNetB0 backbone olarak kullanılmıştır. Şekil 20’de gösterilen transfer öğrenme mimarisi ile eğitimde kullanılan modellere (U-Net, LinkNet, FPN, PSPNet) taş ve harç dokularını ayırt etmek için gerekli olan temel görsel desenleri (doku, gölgeler, renk geçişleri vb.) sıfırdan öğretmek yerine mevcut bilgiler aktarılmıştır. Nispeten az örnekle (176 yüksek çözünürlüklü görüntü ve maske çifti) oluşturulan yerel veri seti ile eğitim süresini önemli ölçüde kısaltarak yüksek performanslı ve genelleme kabiliyeti güçlü segmentasyon modelleri eğitmek mümkün olmuştur.

Performans Metrikleri ve Parametreler

Bu çalışmada segmentasyon modellerinin performansı tahmin edilen maskelerin gerçek maskeler ile nicel olarak karşılaştırılması yoluyla değerlendirilmiştir. Değerlendirme sürecinde iki maske arasındaki mekânsal örtüşme ve sınırsal benzerlik gibi piksel seviyesindeki özellikleri matematiksel olarak ifade eden ve literatürde yaygın olarak kabul gören standart metriklerden yararlanılmıştır (Hossin ve Sulaiman, 2015; Powers, 2020).

Karmaşıklık matrisi

Sınıflandırma algoritmalarının performansını nicel olarak analiz etmek amacıyla kullanılan temel metriklerden biri karmaşıklık matrisidir (confusion matrix). Bir modelin tahmin ettiği sınıf etiketleri ile gerçek sınıf etiketleri arasındaki ilişkiyi sistematik bir biçimde özetler. Matris sunduğu ayrıştırılmış veriler ile modelin isabet oranının (accuracy) ve yaptığı hata türlerinin (Tip I ve Tip II), niteliğini ve dağılımını anlamak için kritik bir zemin oluşturur (Sokolova ve Lapalme, 2009).

Tablo 1. İki Sınıflı Karmaşıklık Matrisi

	Tahmin Edilen Sınıf: Pozitif	Tahmin Edilen Sınıf: Negatif
Gerçek Sınıf: Pozitif	TP	FN
Gerçek Sınıf: Negatif	FP	TN

- Doğru Pozitif (True Positive - TP): Gerçek sınıfı pozitif (taş) olan ve model tarafından da pozitif (taş) olarak sınıflandırılan örneklerin sayısıdır. Modelin pozitif sınıfı doğru tanıma yeteneğini gösterir.
- Yanlış Pozitif (False Positive - FP): Gerçek sınıfı negatif (harç) olan ancak model tarafından hatalı bir şekilde pozitif (taş) olarak sınıflandırılan örneklerin sayısıdır (Tip I Hata).
- Yanlış Negatif (False Negative - FN): Gerçek sınıfı pozitif (taş) olan ancak model tarafından hatalı bir şekilde negatif (harç) olarak sınıflandırılan örneklerin sayısıdır (Tip II Hata).
- Doğru Negatif (True Negative - TN): Gerçek sınıfı negatif (harç) olan ve model tarafından da negatif (harç) olarak sınıflandırılan örneklerin sayısıdır. Modelin negatif sınıfı doğru tanıma kapasitesini ifade eder.

Piksel doğruluğu (Pixel Accuracy)

Toplam piksellerin ne kadarının (taş veya harç) doğru sınıflandırıldığını gösterir. Denklem 1’de gösterildiği gibi doğru tahmin edilen piksellerin toplamının tüm piksel sayısına oranı ile hesaplanır. Başarı için genel bir fikir verse de sınıflar arasında dengesizlik olduğunda yanıltıcı olabilir.

$$\text{Piksel Doğruluğu (Pixel Accuracy)} = \frac{\text{Doğru Sınıflandırılan Piksel Sayısı}}{\text{Toplam Piksel Sayısı}} = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

Kesişim/Birleşim oranı (Intersection over union-IoU) / Jaccard indeksi

Tahmin edilen segmentasyon maskesi ile gerçek maske arasındaki örtüşme oranını ölçer. Denklem 2’de gösterildiği gibi tahmin edilen bölge ile gerçek bölgenin kesişiminin birleşimine bölünmesiyle hesaplanır. Segmentasyon görevlerinde en yaygın ve önemli

metriklerden biridir ayrıca Jaccard indeksi olarak da adlandırılır. Değer 0 ile 1 arasında değişir 1'e yakın değerler daha iyi örtüşme olduğunu gösterir.

$$IoU = \frac{Alan(Gerçek \cap Tahmin)}{Alan(Gerçek \cup Tahmin)} = \frac{TP}{TP + FP + FN} \quad (2)$$

Dice katsayısı (Dice Coefficient)

IoU'ya benzer şekilde iki kümeyi karşılaştıran bir benzerlik ölçütüdür. Denklem 3'te gösterildiği gibi kesişim alanının iki katının toplam alanlarına bölünmesiyle hesaplanır. Alanların örtüşme derecesini verir ve segmentasyon kalitesi için yaygın olarak kullanılır.

$$Dice \text{ katsayısı} = \frac{2 \times |Gerçek \cap Tahmin|}{|Gerçek| + |Tahmin|} = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (3)$$

Kesinlik (Precision)

Modelin pozitif (taş) olarak tahmin ettiği piksellerin ne kadarının gerçekten pozitif (taş) olduğunu gösterir. Denklem 4'te gösterildiği gibi modelin doğru tahmin ettiği gerçek pozitiflerin doğru tahmin edilen tüm piksel sayısına bölünmesi ile hesaplanır.

$$Kesinlik (Precision) = \frac{TP}{TP + FP} \quad (4)$$

Duyarlılık (Recall)

Gerçek pozitif (taş) piksellerin ne kadarının model tarafından pozitif (taş) tahmin edildiğini gösterir. Denklem 5'te gösterildiği gibi modelin doğru tahmin ettiği gerçek pozitiflerin, gerçek tüm pozitif piksel sayısına bölünmesi ile hesaplanır.

$$Duyarlılık (Recall) = \frac{TP}{TP + FN} \quad (5)$$

F1 Skoru (F1 Score)

Kesinlik ve duyarlılık metriklerinin denklem 6'da gösterildiği gibi harmonik ortalamasıdır. Bu iki metrik arasında dengeyi gösterir 1'e yaklaştıkça başarı artar. Sınıf dengesizliği olan durumlarda doğruluk değerinden daha güvenilir bir performans göstergesidir.

$$F1 \text{ Skor} = 2 \times \frac{Kesinlik \times Duyarlılık}{Kesinlik + Duyarlılık} \quad (6)$$

Özgüllük (Specificity)

Gerçek negatif (harç) piksellerin ne kadarının model tarafından negatif (harç) tahmin edildiğini gösterir. Denklem 7’de gösterildiği gibi modelin doğru tahmin ettiği gerçek negatiflerin, gerçek tüm negatif piksel sayısına bölünmesi ile hesaplanır.

$$\text{Özgüllük (Specificity)} = \frac{TN}{TN + FP} \quad (7)$$

Erken durdurma (Early Stopping)

Erken durdurma (early stopping) derin öğrenme modellerinin eğitim sürecinde aşırı öğrenmenin (overfitting) önüne geçmek için yaygın olarak kullanılan etkili bir regularizasyon yöntemidir. Bu yöntem ile modelin eğitim verisi üzerinde aşırı uyum göstermeye başladığı noktada doğrulama kaybı (validation loss) veya başka bir performans metriği izlenerek eğitimin otomatik olarak sonlandırılması sağlanır. Belirlenen bir sabır süresi (patience) boyunca doğrulama kaybında iyileşme görülmezse eğitim durdurulur ve genellikle doğrulama kaybının en düşük olduğu noktadaki ağırlıklar kaydedilir (Goodfellow vd., 2016). Böylece modelin genelleme yeteneği korunur ve gereksiz yere fazla parametre güncellemesi yapılmaz. Ayrıca erken durdurmanın modelin eğitim süresini kısaltması ve donanım kaynaklarının daha verimli kullanılmasını sağlaması da pratik açıdan önemli bir avantajdır (Prechelt, 2002).

Metot

Bu çalışmada yığma taş duvarların anlamsal segmentasyonu için derin öğrenme tabanlı bir yaklaşım geliştirilmiştir. Çalışmanın temel amacı yüksek çözünürlüklü duvar görsellerindeki taş ve harç alanlarını piksel düzeyinde doğru bir şekilde sınıflandırmaktır. Görüntülerin yüksek çözünürlüklü olması standart derin öğrenme modellerinin GPU bellek kısıtlamaları nedeniyle doğrudan işlenmesini zorlaştırmaktadır (Szeliski, 2022). Görüntülerin tamamını modellerin girdi boyutuna küçültmek taşlar arasındaki ince harç derzleri gibi kritik detayların kaybına yol açacağından bu yaklaşımdan kaçınılmıştır. Bunun yerine veri bütünlüğünü korumak ve bellek kullanımını optimize etmek amacıyla literatürde sıklıkla kullanılan parça tabanlı eğitim tercih edilerek tüm yüksek çözünürlüklü görseller kullanılan modeller için tercih edilen giriş boyutu olan 256x256 piksellik daha küçük parçalara (patch) ayrılmıştır (Zhang vd., 2016). Bu parça tabanlı yaklaşım modelin hem eğitim hem de tahmin aşamalarında yerel doku ve desen özelliklerini daha etkin bir şekilde öğrenmesini sağlamıştır. Modelin genelleme yeteneğini artırmak ve aşırı öğrenmeyi (overfitting) engellemek amacıyla eğitim verilerine uygulanan literatürde etkinliği kanıtlanmış veri artırma (data augmentation) teknikleri veri ön işleme bölümünde açıklanmıştır (Shorten & Khoshgoftaar, 2019).

Model geliştirme sürecinde literatürde başarısını kanıtlamış ve yaygın olarak kullanılan segmentasyon mimarilerinden faydalanılmıştır. Bu kapsamda esnekliği ve modüler yapısıyla araştırmalarda sıklıkla tercih edilen “segmentation_models” kütüphanesi aracılığıyla U-Net (Ronneberger vd., 2015), LinkNet (Chaurasia ve Culurciello, 2017), FPN (Lin vd., 2017) ve PSPNet (Zhao vd., 2017) gibi farklı kodlayıcı-çözücü (encoder-decoder) mimarilerine dayanan önceden eğitilmiş modeller kullanılmıştır. Bu modellerin performansını ve genelleme yeteneğini artırmak için transfer öğrenme (transfer learning) tekniğinden yararlanılmıştır (Pan ve Yang, 2009). Bu sayede modeller temel görsel özellikleri (kenar, köşe, taş yüzeyi, harç vb.) sıfırdan öğrenmek yerine önceden edinilmiş zengin bir bilgi birikimiyle eğitime başlanmıştır. Farklı karmaşıklık ve verimlilik seviyelerine sahip olan VGG16, ResNet34, ResNet50, DenseNet121, MobileNetV1, MobileNetV2 ve EfficientNetB0 gibi çeşitli omurga mimarileri sistematik olarak test edilmiştir. İnce ayar (fine-tuning) işlemi için önceden eğitilmiş modellerin ilk katmanlarındaki düşük seviyeli özellikler transfer edilirken, oluşturulan modellerin son katmanları problem özelinde veri setine ait yüksek seviyeli özellikler ile yeniden eğitilmiştir (Kornblith vd., 2019).

Modelin derlenmesi aşamasında ikili segmentasyon için tek sınıflı çıktı (0, 1) ve çıkış aktivasyonu için “sigmoid” fonksiyonu kullanılmıştır. Parametre optimizasyonu için adaptif öğrenme oranı ayarlaması sebebiyle uyarlanabilir moment tahmini (Adaptive Moment Estimation-Adam) algoritması tercih edilmiştir. Adam algoritması özellikle segmentasyon gibi yüksek boyutlu ve karmaşık problemlerde hızlı yakınsama sunar (Kingma ve Ba, 2015). İkili (binary) segmentasyon için literatürde sıklıkla tercih edilen ikili çapraz entropi (binary cross-entropy-BCE) kayıp fonksiyonu kullanılmıştır. Bu fonksiyon modelin ürettiği olasılık haritası ile gerçek maske arasındaki farkı ölçerek modelin hatalı tahminleri üzerinden ödül-ceza yöntemiyle ağırlıkların bu hatayı minimize edecek yönde güncellenmesini sağlamaktadır (Jadon, 2020). Modelin eğitim başarısını değerlendirmek için literatürde yaygın şekilde kabul gören doğruluk/kayıp, IoU, Dice katsayısı ve piksel doğruluğu gibi çeşitli metrikler hesaplanmıştır (Wang vd., 2020). En iyi modelin belirlenmesi için doğrulama kaybı (validation loss) takip edilmiş ve en düşük olduğu durumda test aşamalarında kullanılmak üzere model ağırlıkları kaydedilmiştir. Aşırı öğrenmenin (overfitting) önüne geçmek için erken durdurma (early stopping) ile doğrulama kaybı üst üste 5 turdan sonra değişmediğinde eğitim otomatik olarak sonlandırılmıştır. Eğitimler 30 tur (epoch) üzerinden gerçekleştirilmiştir. Epoch sayısının az olması durumunda model problemin çözümünü öğrenemeyebilir ve başarısı düşük olur, fazla olması durumunda model problemin çözümünü ezberleyebilir ve başarısı yine düşük olur. Optimum seviyede bir değerin belirlenmesi model başarısı için önemlidir (Skansi, 2018). Eğitim sürecinde model ağırlıkları her bir iterasyonda tüm veri seti yerine veri setinden seçilen

daha küçük bir alt küme kullanılarak güncellenir. Bu alt kümedeki örnek sayısı mini-yığın boyutu (batch size) olarak adlandırılır (Goodfellow vd., 2016). Modellerinin eğitimi sırasında tüm verinin aynı anda eğitim işlemine alınması bellek ve işlem süresi nedeniyle maliyetli bir iştir. Kullanılan farklı model mimarilerinin karmaşıklığı ve mevcut donanım kaynaklarının işlem kapasitesi gibi faktörler göz önüne alınarak batch size değerleri 8, 64, 128, 256 gibi farklı seviyelerde kullanılmıştır.

Eğitim süreci tamamlandıktan sonra modelin yeni ve daha önce görmediği veriler üzerindeki tahmin yeteneğini ölçmek için çeşitli metrikler ve görsel analiz yöntemleri kullanılır. Bu değerlendirme aşaması modelin güçlü ve zayıf yönlerinin anlaşılması, potansiyel iyileştirme alanlarının belirlenmesi ve nihayetinde modelin gerçek dünya senaryolarında ne kadar başarılı olabileceğinin öngörülmesini sağlar. Çalışmada kullanılan modellerin performanslarının karşılaştırılması için tahmin maskeleri üzerinden piksel doğruluğu, Dice katsayısı, IoU, kesinlik, duyarlılık, F1 skor ve özgüllük metrikleri hesaplanmıştır (Taha ve Hanbury, 2015).

Eğitilen modellerin test aşamasında elde edilen tahmin maskeleri (olasılık haritaları) her pikselin taş a ait olma olasılığını gösteren 0 ile 1 arasında değerler içeren bir olasılık haritasıdır. Bu haritayı net bir ikili segmentasyon maskesine (0: harç, 1: taş) dönüştürmek için eşikleme (thresholding) yöntemleri kullanılır (Sezgin ve Sankur, 2004). Çalışmada klasik ve Otsu eşikleme yöntemleri uygulanmıştır. Klasik eşikleme yaklaşımında 0,7 sabit eşik değeri (treshold), Otsu eşikleme yönteminde ise görüntü histogramına bağlı olarak otomatik belirlenen global eşik değeri tercih edilmiştir. Test aşamasında her iki yöntemin benzer sonuçlar üretmesi nedeniyle çalışmaya Otsu eşikleme yöntemi ile devam edilmiştir. Eşikleme sonrası elde edilen ham ikili maskelerdeki kusurları gidermek amacıyla aşamalı bir son işleme (post-processing) prosedürü geliştirilmiştir.

İlk aşamada küçük nesneler temizlendikten sonra kapama (closing) işlemi ile taşlarda oluşan küçük delikler doldurulmuştur. Erozyon (erode) ile kaplama sonrası genişleyen harç ve taş sınırları düzenlenerek maskenin bütünlüğü artırılmıştır (Gonzalez ve Woods, 2018). Eksik kalan veya kopuk olan harç çizgilerinin tamamlanması için yatay ve dikey doldurma uygulanmıştır. Mesafe dönüşümü (distance transform) uygulanarak lokal maksimumlar belirlenmiş ve taş merkezleri tespit edilmiştir. Watershed algoritması kullanılarak harç çizgileri belirlenmiş ve birleşik taş bölgeleri birbirinden ayrılmıştır (Vincent ve Soille, 1991). Morfolojik genişletme (dilation) uygulanarak taşların arasındaki sınır çizgileri (harç) kalınlaştırılmış ve görsel olarak belirginleştirilmiştir (Gonzalez ve Woods, 2018).

Sonraki aşamada ise taşların alan hesaplamalarını veya görsel analizleri olumsuz etkileyebilecek, morfolojik işlemler veya modelin kendisi tarafından taşların yüzeyinde

oluřturulan küçük ve hatalı harç adacıkları tespit edilmiřtir. Hatalı bölgelerin tespiti için ana harç bloğundan (kesintisiz en büyük harç bölgesi) ayrı genellikle taşların içinde kalan parçalar, bağlantılı bileřen analizi (Connected Component Analysis-CCA) ve dikdörtgensellik (alan/řekil) tabanlı filtreleme ile tespit edilip temizlenerek nihai segmentasyon maskesi elde edilmiřtir (Aksoy, 2020).

Son ařamada temizlenmiř ve son haline getirilmiř maskeler üzerinden nicel analizler yapılmıřtır. Her bir maske ierisindeki taşlar ayrı ayrı etiketlenerek alan, genişlik, yükseklik ve en/boy oranı gibi geometrik özellikleri hesaplanmıřtır. Ayrıca tahmin edilen her bir taşın merkez noktası (centroid) ile gerek maskesindeki karřılıėı eřleřtirilerek alan örtüşmesine dayalı bireysel bir doėruluk skoru hesaplanmıř ve görselleřtirilmiřtir. Tüm bu geometrik ve model bařarısını ölçen metrikler (piksel doėrulu, IoU, Dice katsayısı vb.) ile sürecin her ařamasından elde edilen görsel ıktılar, farklı model ve iřleme adımlarının segmentasyon bařarısını kapsamlı biimde karřılařtırmak amacıyla bulgular kısmında tablolar ve görseller halinde sunulmuřtur.

ARAŞTIRMA BULGULARI VE TARTIŞMA

Bu bölümde yığma taş yapıların dış duvar yüzeylerinde taş ve harç ayırımının otomatik olarak yapılabilmesi ve taş boyutlarının belirlenebilmesi amacıyla kullanılan derin öğrenme tabanlı segmentasyon modellerinin ve uygulanan görüntü işleme tekniklerinin deneysel bulguları sunulmuştur. Çalışma kapsamında U-Net, FPN, LinkNet ve PSPNet gibi güncel mimariler ile VGG16, ResNet34, ResNet50, DenseNet121, MobileNetV1, MobileNetV2 ve EfficientNetB0 omurgalarına sahip farklı derin öğrenme modellerinin segmentasyon performansları karşılaştırılmıştır. Deneyler Intel(R) Core(TM) i7-12700H CPU ve 16 GB RAM'e sahip bir bilgisayar üzerinde PyCharm 2024.2.3 IDE'si kullanılarak gerçekleştirilmiş; donanımın yetersiz kaldığı durumlarda ise Google Colab platformundaki L4-GPU sanal makinaları ile model eğitimleri desteklenmiştir. Kullanılan veri seti yüksek çözünürlüklü 176 adet duvar görüntüsünün, 256x256 piksellik parçalara (patch) ayrılmasıyla toplam 108194 adet görüntü-maske çiftinden oluşmaktadır.

Araştırmanın bu bölümünde karşılaştırmalı segmentasyon modellerinin eğitim ve test süreçlerine ilişkin nicel sonuçlar tablolar ve grafikler aracılığıyla sunulmaktadır. Ayrıca uygulanan görüntü işleme adımları sonucunda elde edilen nihai segmentasyon maskelerine ilişkin görsel analizlere yer verilmektedir. Elde edilen bulgular taş ve harç ayrımı başta olmak üzere model seçimlerinin uygulama başarısına etkisi bakımından ayrıntılı olarak tartışılacaktır.

Uygulama Detayları

Bu çalışmada yığma taş yapıların dış duvar yüzeylerinde taş ve harç ayırımının otomatik olarak yapılabilmesi ve taş boyutlarının hassas biçimde belirlenebilmesi amacıyla çeşitli derin öğrenme tabanlı semantik segmentasyon modelleri ve gelişmiş görüntü işleme teknikleri bir arada kullanılmıştır. Deneyler Intel(R) Core (TM) i7-12700H CPU, 16 GB RAM'e sahip bir bilgisayarda PyCharm 2024.2.3 ortamında yürütülmüş, ayrıca donanımın yetersiz kaldığı durumlarda ise Google Colab platformundaki L4-GPU sanal makinaları ile model eğitimleri desteklenmiştir. Kullanılan veri seti yüksek çözünürlüklü (6144x8192, 4182x6144 vb.) taş duvar görsellerinin 256x256 piksellik parçalara (patch) bölünmesiyle oluşturulmuş olup, toplamda 108194 adet görüntü-maske çiftinden (eğitim için %80, doğrulama için %20) meydana gelmektedir.

Değerlendirme Metrikleri

Model performansını değerlendirmek amacıyla piksel doğruluğu (Pixel Accuracy), IoU (Intersection over Union), Dice katsayısı (Dice Coefficient), kesinlik (Precision), duyarlılık (Recall), F1 Skor ve özgüllük (Specificity) gibi literatürde sıklıkla tercih edilen metrikler hesaplanmıştır. Ayrıca segmentasyon sonrası her bir taşın geometrik özelliklerinin (alan, genişlik, yükseklik, en/boy oranı) çıkarılmasına yönelik analizler de yapılmıştır.

DeneySEL Sonuçlar

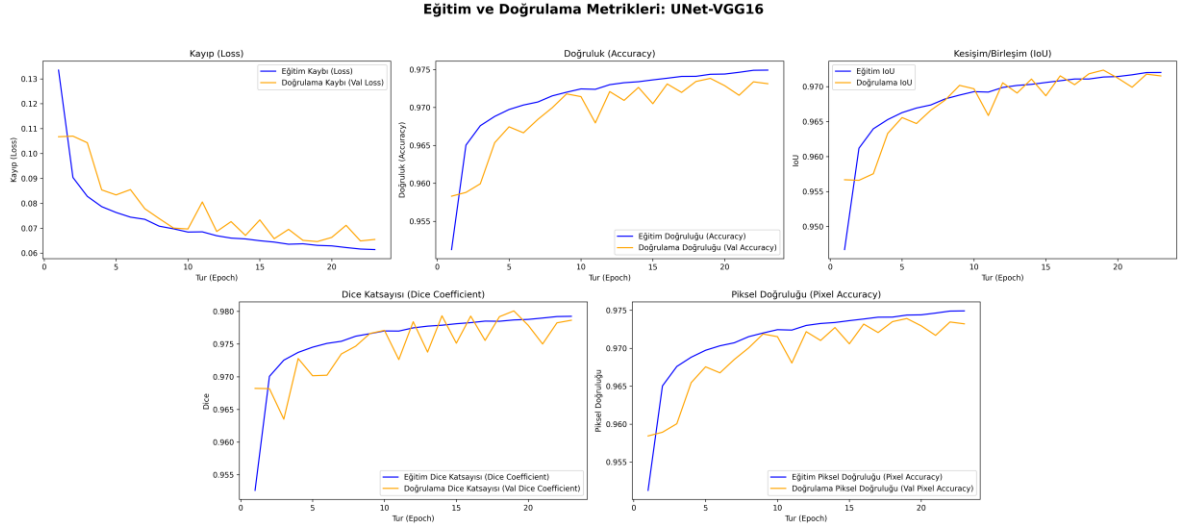
Çalışmada U-Net, FPN, LinkNet ve PSPNet mimarilerine dayalı VGG16, ResNet34, ResNet50, DenseNet121, MobileNetV1, MobileNetV2 ve EfficientNetB0 omurga yapısına sahip toplamda 28 farklı model ile segmentasyon deneyleri gerçekleştirilmiştir. Model eğitimlerinde transfer öğrenme uygulanarak önceden eğitilmiş ağırlıklar kullanılmış ve veri artırma teknikleri ile modellerin genelleme yeteneği güçlendirilmiştir. Modellerin eğitim süreçlerinde erken durdurma (early stopping) yöntemi ile aşırı öğrenmenin önüne geçilmiş ve gereksiz donanım kullanımı (hesaplama maliyeti) azaltılarak eğitim süreleri kısaltılmıştır. Bu stratejilere ek olarak modellerin öğrenme sürecini doğrudan etkileyen optimizasyon algoritması, öğrenme oranı, kayıp fonksiyonu ve tur sayısı (epoch) gibi temel parametrelerde bulunmaktadır. Karşılaştırılan tüm modellerin başarısını objektif ve nesnel bir şekilde değerlendirebilmek için tüm modellerde aynı parametreler ve teknikler kullanılmıştır.

Modellerin eğitim ve doğrulama süreçlerine ait kayıp, doğruluk, kesişim/birleşim, Dice katsayısı ve piksel doğruluğu grafikleri ilgili başlıklar altında sunulmuştur. Karşılaştırmalı analizi kolaylaştırmak için grafiklerde eğitim verilerine ait sonuçlar mavi, doğrulama verilerine ait sonuçlar ise turuncu renkle belirtilmiştir. Bu yaklaşım her modelin eğitim sürecindeki kararlılığını izleme ve modelleri doğrulama performansları açısından nesnel olarak kıyaslama olanağı sunmaktadır.

U-Net modellerine ait eğitim ve doğrulama grafikleri

Bu bölümde U-Net mimarisinin farklı kodlayıcıları (backbone) kullanarak transfer öğrenme ile eğitilmesi sırasında elde edilen eğitim ve doğrulama metriklerinin grafikleri (Şekil 21-27) analiz edilmektedir. Bu grafikler modellerin öğrenme hızı, kararlılığı ve genelleme kapasitesi hakkında önemli bilgiler sunmaktadır.

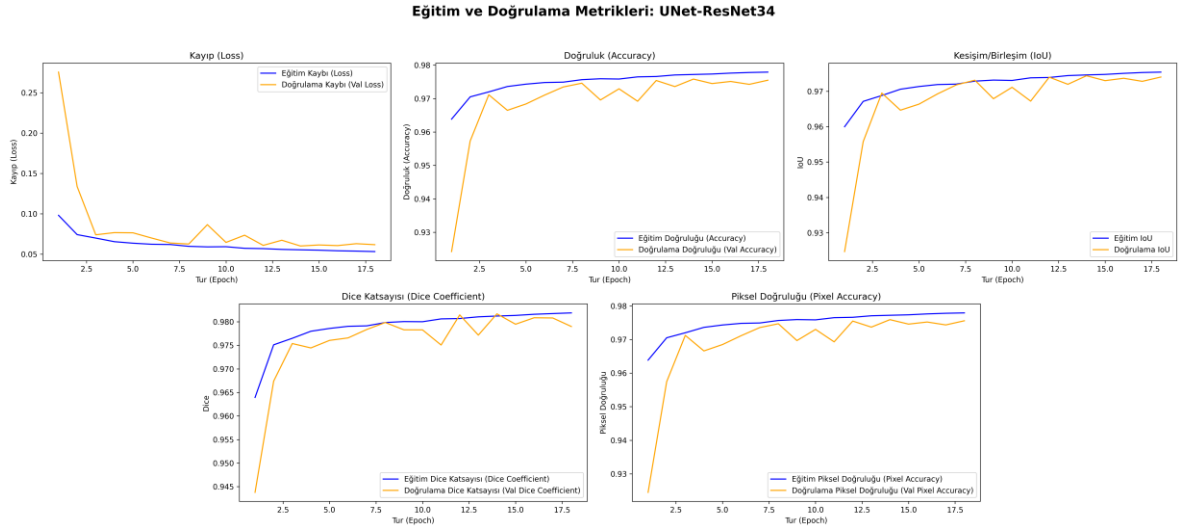
VGG16



Şekil 21. U-Net/VGG16 modeli eğitim ve doğrulama grafiği.

Şekil 21’de gösterilen U-Net/VGG16 modeline ait eğitim ve doğrulama metriklerinde hızlı bir yakınsama ile doğrulama eğrilerindeki küçük dalgalanmalara rağmen modelin stabil bir öğrenme gerçekleştirdiği görülmektedir. Doğruluk, IoU, Dice katsayısı ve piksel doğruluğu değerlerinin yüksek seviyelere ulaştığı, eğitim ve doğrulama eğrileri arasındaki farkların az olduğu gözlemlenmektedir. Eğriler arasındaki farkın az olması modelin aşırı öğrenmeye (overfitting) yatkın olmadığını ve güçlü bir genelleme yeteneği oluşturduğunu göstermektedir.

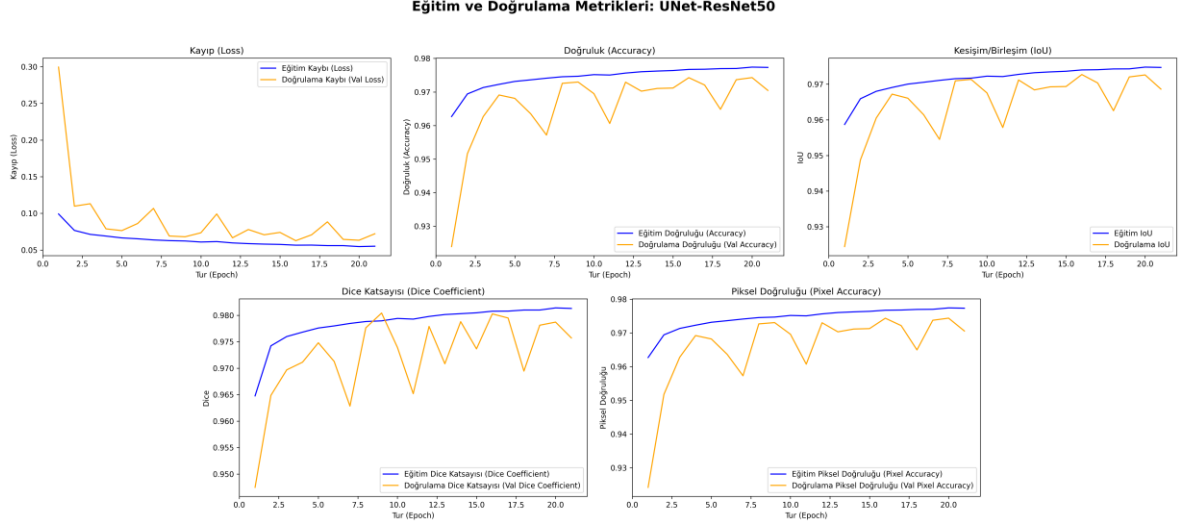
ResNet34



Şekil 22. U-Net/ResNet34 modeli eğitim ve doğrulama grafiği.

Şekil 22’de gösterilen U-Net/ResNet34 modeli eğitiminde hem kayıp hem de doğruluk metrikleri kısa sürede stabil değerler almış, doğrulama metrikleri de eğitim değerlerine yakın seyretmiştir. Doğrulama kaybı zaman zaman yükselse de genel olarak yüksek bir doğruluk ve genelleme yeteneği gözlemlenmektedir.

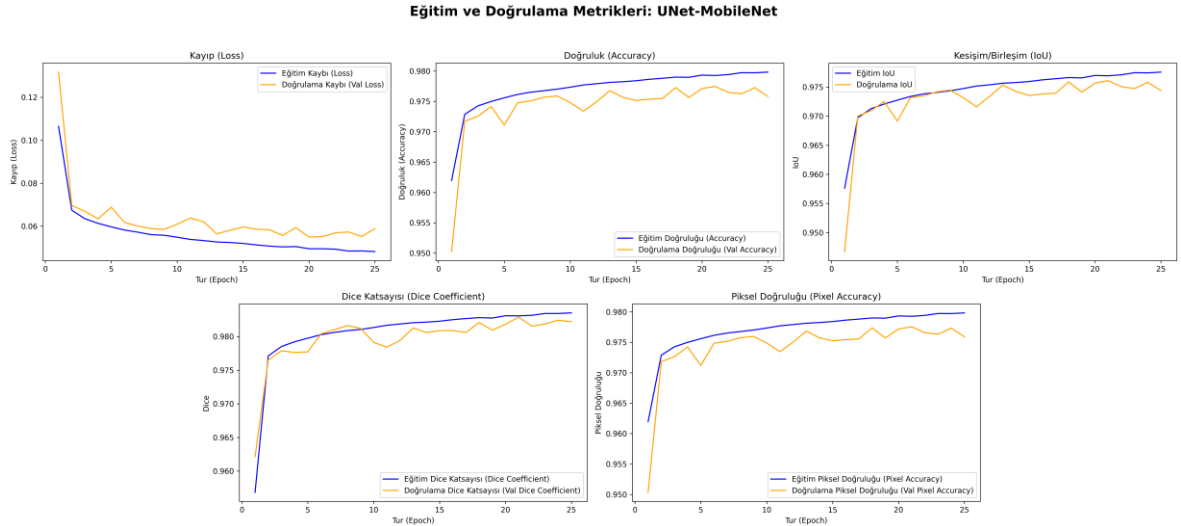
ResNet50



Şekil 23. U-Net/ResNet50 modeli eğitim ve doğrulama grafiği.

Şekil 23'te gösterilen U-Net/ResNet50 modeli eğitim eğrilerinde düzenli bir artış ve iyileşme gözlemlenmiştir. Ancak doğrulama eğrilerinde gözlenen keskin dalgalanmalar veri集中的 çeşitlilik veya öğrenme oranı kaynaklı olabilir. Bu durum modelin genelleme yeteneğinin zayıf olduğuna yorumlanabilir. Eğitim değerinin yüksek olması güçlü bir öğrenme gerçekleştirdiği gösterse de doğrulama eğrilerindeki dalgalanmalar modelin kararsız bir yapıda olduğunu düşündürmektedir.

MobileNetV1

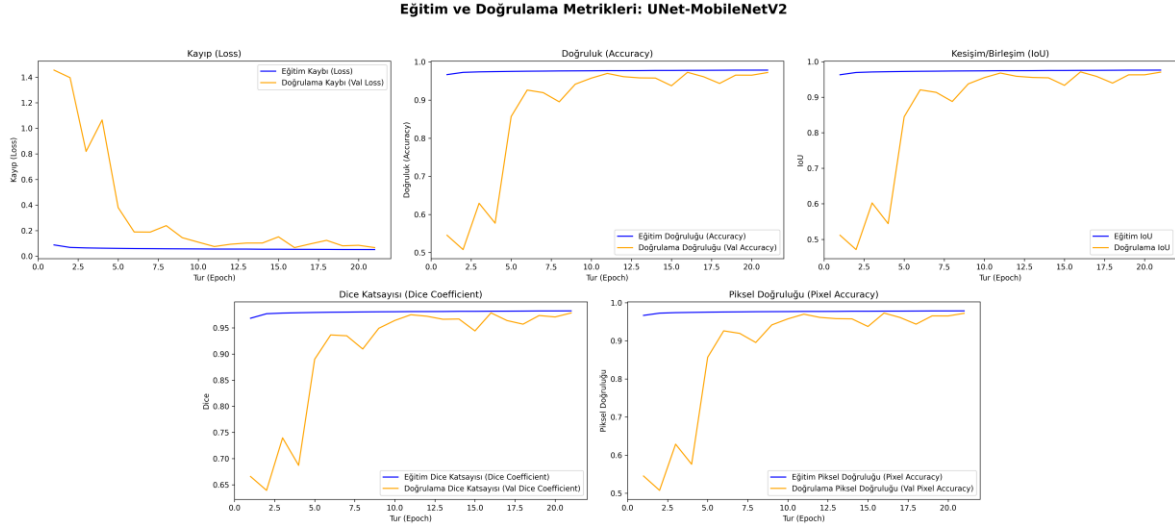


Şekil 24. U-Net/MobileNetV1 modeli eğitim ve doğrulama grafiği.

Şekil 24'te gösterilen U-Net/MobileNetV1 modeli eğitim süreci boyunca hem eğitim hem de doğrulama metriklerinde yüksek ve dengeli bir performans sergilediği görülmektedir. Piksel doğruluğu, IoU ve Dice katsayısı metrikleri her iki veri kümesinde de benzer eğilimler göstererek yüksek değerlere ulaştığı gözlemlenmiştir. Doğrulama eğrileri eğitim eğrilerine

yakın seyrederek aşırı uyum (overfitting) riskinin düşük olduğunu ve doğrulama setindeki dalgalanmaların minimal düzeyde kalması modelin genelleme yeteneğinin yüksek olduğunu göstermektedir.

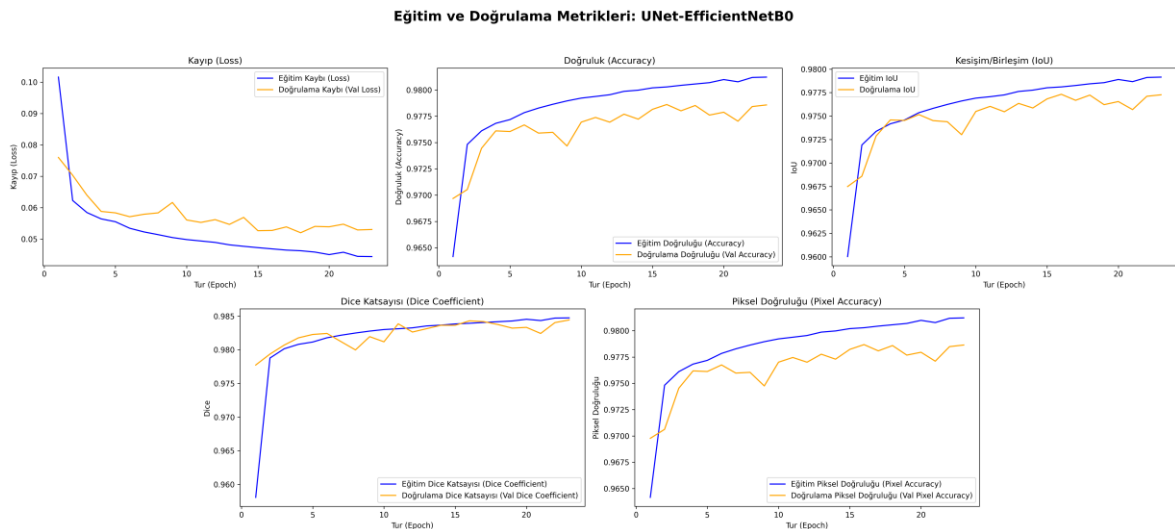
MobileNetV2



Şekil 25. U-Net/MobileNetV2 modeli eğitim ve doğrulama grafiği.

Şekil 25'te gösterilen U-Net/MobileNetV2 modeline ait eğitim eğrilerinde hızlı bir iyileşme ve kısa sürede plato (düz bir seyir) oluşumu görülmüştür. Ancak doğrulama eğrilerinde ilk turlarda dalgalanmalar görülmesi modelin genelleme kapasitesinin eğitimin başında sınırlı olduğuna yorumlanabilir. İlerleyen turlarda dalgalanmalar azalarak genel olarak başarılı ve stabil bir öğrenme sağlandığı gözlemlenmektedir.

EfficientNetB0



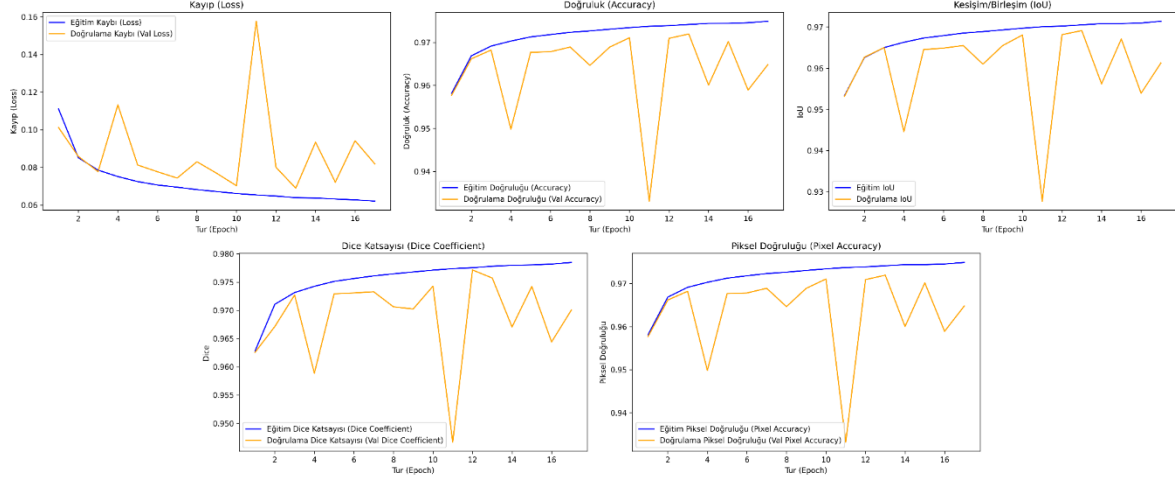
Şekil 26. U-Net/EfficientNetB0 modeli eğitim ve doğrulama grafiği.

Şekil 26'da gösterilen U-Net/EfficientNetB0 modeli eğitim kayıp eğrilerinde düzenli bir azalma ve doğrulukta istikrarlı artış gözlenirken, doğrulama eğrilerinde minimal seviyelerde

de olsa dalgalanmalar mevcuttur. Metriklerin yüksek olması modelin genel segmentasyon performansının başarılı olduğuna işaret etmektedir.

DenseNet121

Eğitim ve Doğrulama Metrikleri: UNet-DenseNet121



Şekil 27. U-Net/DenseNet121 modeli eğitim ve doğrulama grafiği.

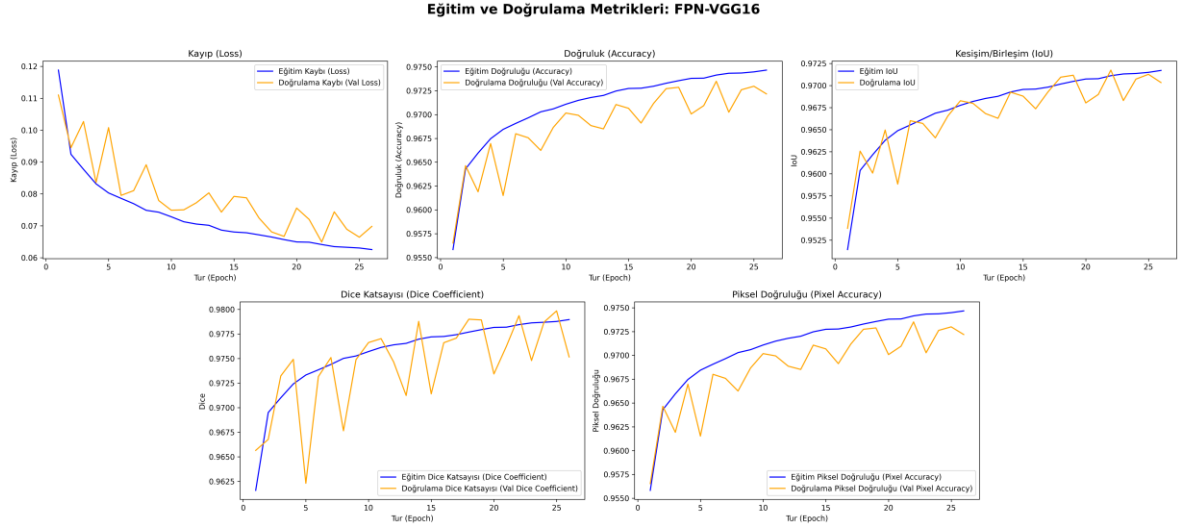
Şekil 27’de gösterilen U-Net/DenseNet121 modeli eğitim eğrilerinde hızlı bir yakınsama ve yüksek doğruluk elde edilmiş; fakat doğrulama eğrilerinde zaman zaman ani düşüş ve yükselmeler görülmektedir. Özellikle doğrulama kaybı ve Dice katsayısı eğrisindeki dalgalanmalar modelin bazı turlarda aşırı öğrenmeye veya doğrulama setindeki zorlu örneklerle karşı hassas olduğuna yorumlanabilir.

Tüm U-Net modelleri eğitim sürecinde genel olarak piksel doğruluğu, Dice katsayısı ve IoU gibi metriklerde başarılı sonuçlar üretmiştir. Özellikle VGG16, ResNet34 ve MobileNetV1 tabanlı modeller, eğitim ve doğrulama metriklerinde hızlı yakınsama ve yüksek değerlerde stabil seyir sergilemiştir. ResNet50, MobileNetV2 ve DenseNet121 mimarilerinde doğrulama eğrilerindeki dalgalanmalar modelin veri setindeki çeşitliliğe karşı daha hassas olabileceğini göstermektedir. EfficientNetB0 ise genel olarak istikrarlı bir öğrenme sağlarken doğrulama ve eğitim eğrilerinde yer yer ayrılmalar görülse de nihai model başarısı yüksektir.

FPN modellerine ait eğitim ve doğrulama grafikleri

Bu bölümde FPN (Feature Pyramid Network) mimarisinin farklı kodlayıcıları (backbone) kullanarak transfer öğrenme ile eğitilmesi sırasında elde edilen eğitim ve doğrulama metriklerinin grafikleri (Şekil 28-34) üzerinden incelenmektedir. Bu analiz her bir modelin öğrenme kararlılığını, yakınsama hızını ve genelleme performansını değerlendirmektedir.

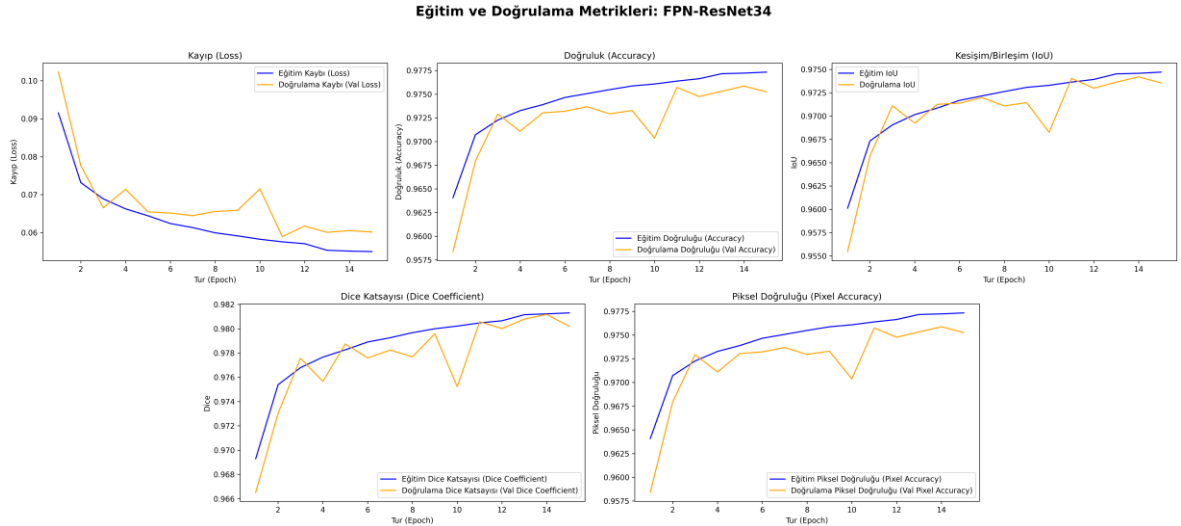
VGG16



Şekil 28. FPN/VGG16 modeli eğitim ve doğrulama grafiği.

Şekil 28’de gösterilen FPN/VGG16 modeline ait eğitim eğrileri istikrarlı biçimde iyileşirken, doğrulama metriklerinde dikkat çekici dalgalanmalar gözlemlenmiştir. Bu dalgalanmalar modelin validasyon setindeki çeşitlilik ve bazı zor örnekler (çok ince harçlar gibi) karşısında duyarlılığını artırmasından kaynaklanabilir. Model genel olarak yüksek IoU ve Dice katsayısı değerlerine ulaşmış olsa da doğrulama metriklerindeki dengesizlik genelleme açısından sınırlı bir başarıya işaret ettiğini göstermektedir.

ResNet34

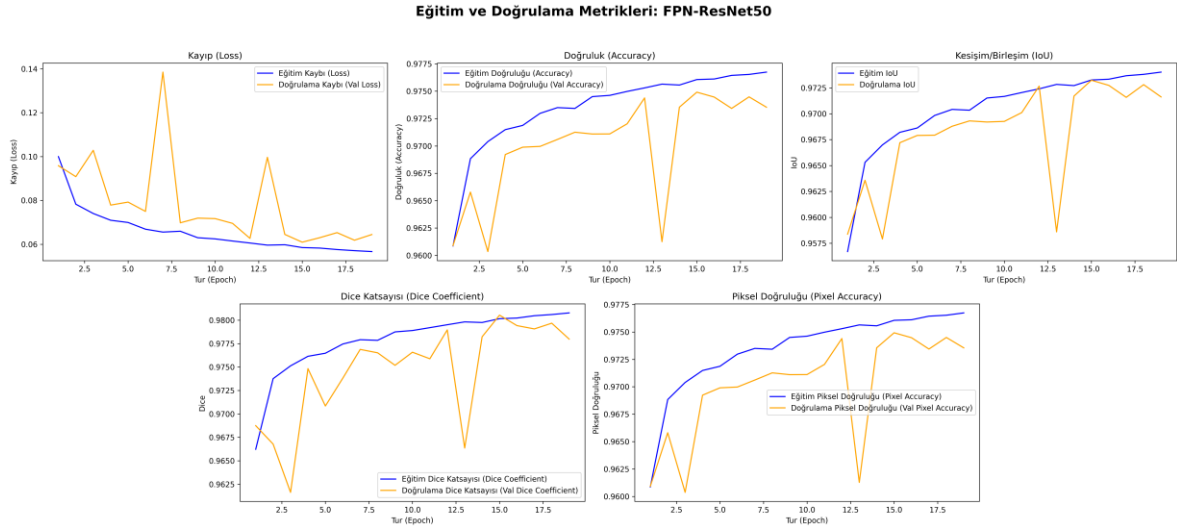


Şekil 29. FPN/ResNet34 modeli eğitim ve doğrulama grafiği.

Şekil 29’da gösterilen FPN/ResNet34 modeli eğitiminde başlangıçta hızlı ve stabil bir yakınsama gözlemlenmiş, eğitim ve doğrulama metrikleri birbirine paralel ve yakın ilerlemiştir. Bu durum modelin hızlı öğrendiğini ve yüksek genelleme kabiliyetine sahip olduğuna

yorumlanabilir. Doğrulama eğrilerinde zaman zaman hafif dalgalanmalar olsa da modelin tutarlı ve yüksek doğruluk değerlerine ulaştığı görülmektedir.

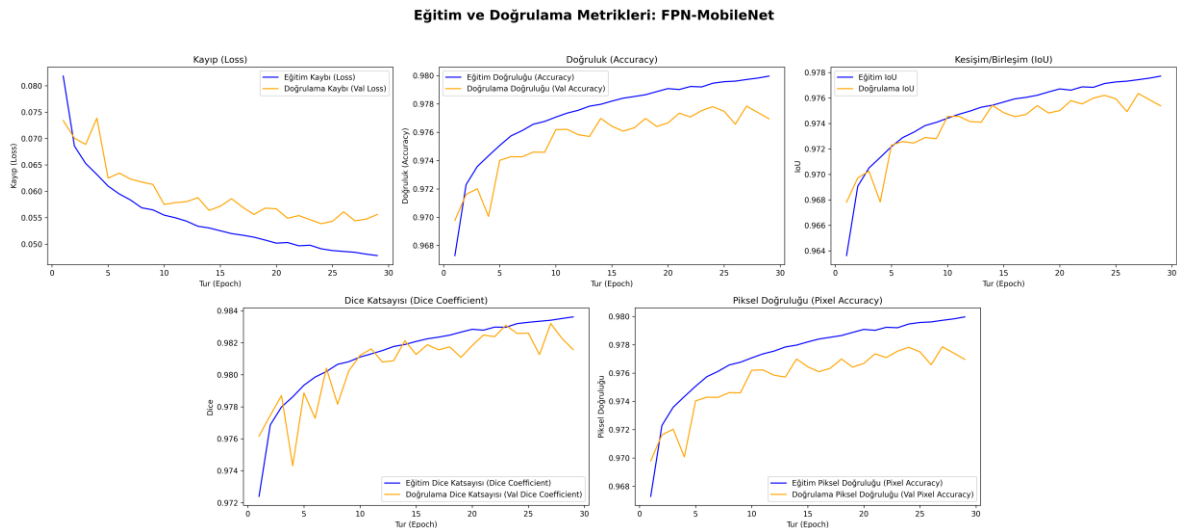
ResNet50



Şekil 30. FPN/ResNet50 modeli eğitim ve doğrulama grafiği.

Şekil 30’da gösterilen FPN/ResNet50 modeli eğitim eğrileri düzgün ilerlerken doğrulama metriklerde anlık ani yükselmeler ve düşüşler gözlemlenmiştir. Bu dalgalanmalar modelin doğrulama setindeki zorlayıcı ve aykırı örnekler karşısında adaptasyon sürecini yansıtmaktadır. Ancak turlar ilerledikçe doğrulama metriklerindeki artış trendi ile model genel olarak yüksek doğrulama değerlerine ulaştığı gözlemlenmiştir.

MobileNetV1

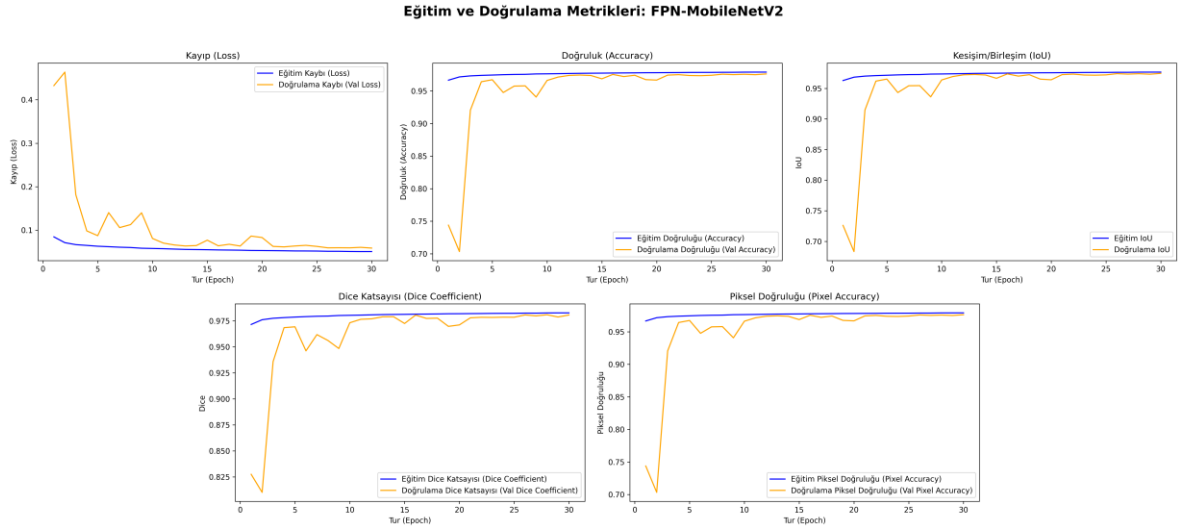


Şekil 31. FPN/MobileNetV1 modeli eğitim ve doğrulama grafiği.

Şekil 31’de gösterilen FPN/MobileNetV1 modelinin eğitim eğrileri oldukça düzgün ve kararlı ilerlemiş ancak doğrulama eğrilerinde dalgalanmalar gözlenmiştir. Bu durum modelin bazı validasyon örneklerine karşı hassasiyetini yansıtsa da genel başarıda belirgin bir düşüşe

yol açmamıştır. Eğitim sürecindeki dalgalanmalara rağmen modelin eğitim ve doğrulama değerlerinin yüksek seviyelere ulaştığı gözlemlenmiştir.

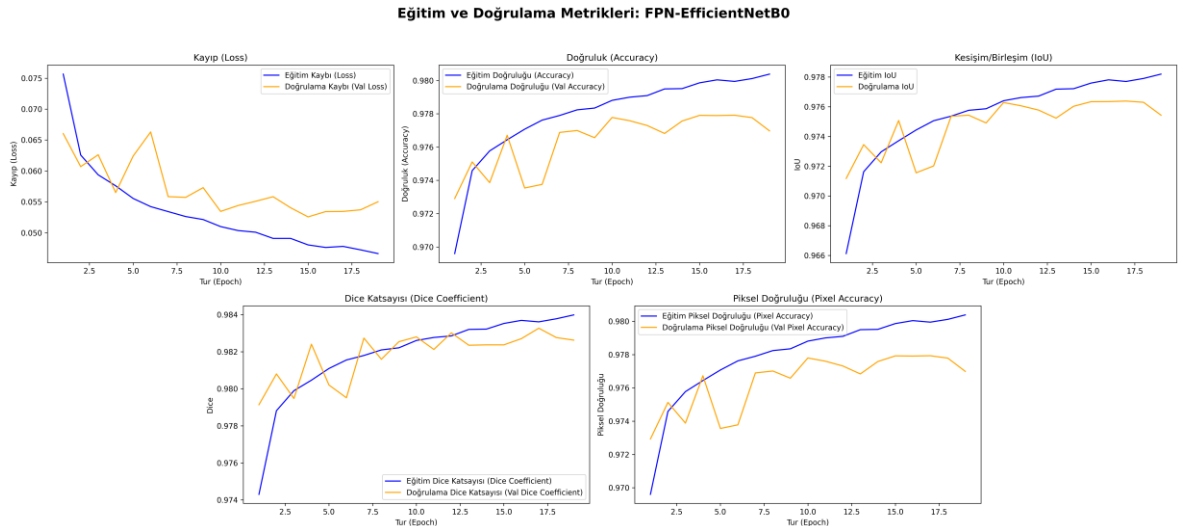
MobileNetV2



Şekil 32. FPN/MobileNetV2 modeli eğitim ve doğrulama grafiği.

Şekil 32’de gösterilen FPN/MobileNetV2 modeline ait hem eğitim hem de doğrulama metriklerinde hızlı bir iyileşme ve plato oluşumu görülmüştür. Eğitim sürecinin başında doğrulama ve eğitim eğrileri arasında oluşan fark ilerleyen turlarda hızla azalmıştır. Doğrulama Dice katsayısı eğrisinde küçük dalgalanmalar olsa da model aşırı öğrenmeye düşmeden etkili ve hızlı bir öğrenme ile genel olarak yüksek ve stabil başarı gösterdiği gözlemlenmiştir.

EfficientNetB0



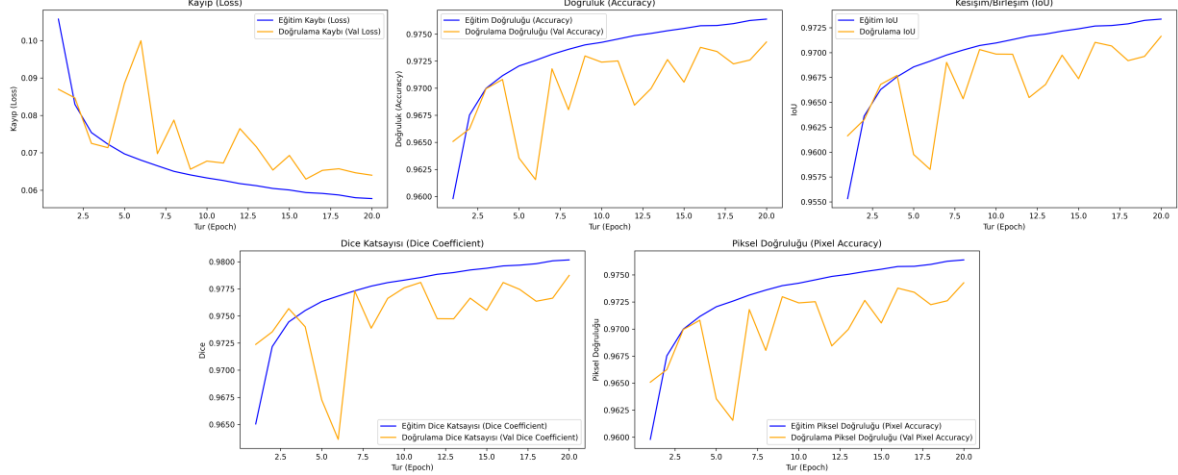
Şekil 33. FPN/EfficientNetB0 modeli eğitim ve doğrulama grafiği.

Şekil 33’te gösterilen FPN/EfficientNetB0 modeli eğitim eğrileri boyunca düzenli ve sürekli bir iyileşme gözlenmiş, doğrulama metriklerinde ise dalgalanmalar olmuştur. Özellikle Dice katsayısı ve IoU gibi metriklerdeki dalgalanmalar ile piksel doğruluğu grafiğindeki

ayrışmalar modelin doğrulama verisi üzerindeki genelleme kabiliyetinin azaldığına ve modelin veriye aşırı uyum gösterdiğine (overfitting) yorumlanabilir.

DenseNet121

Eğitim ve Doğrulama Metrikleri: FPN-DenseNet121



Şekil 34. FPN/DenseNet121 modeli eğitim ve doğrulama grafiği.

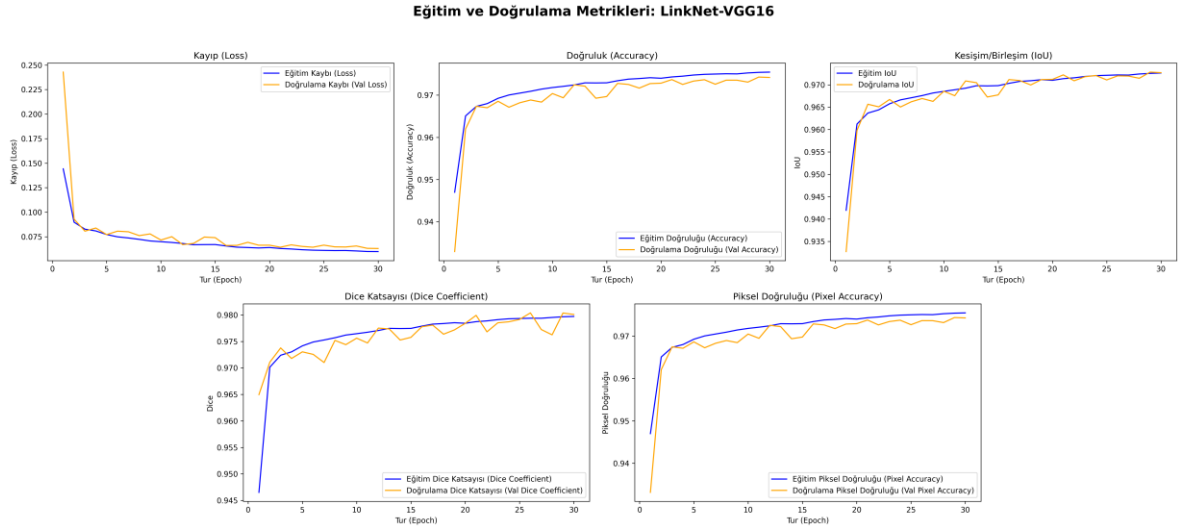
Şekil 34’te gösterilen FPN/DenseNet121 modeli eğitim eğrileri hızlı bir artış ve stabil bir seyir izlerken doğrulama eğrilerinde zaman zaman ani düşüş ve yükselmeler görülmektedir. Bu dalgalanmalar modelin aşırı öğrenmeye ve bazı farklı ya da zor doğrulama örneklerine karşı hassas olduğunu ve genelleme başarısının diğer modellere göre düşük olduğunu göstermektedir. Ancak genel metrik düzeyi yüksek olup, piksel doğruluğu ve Dice katsayısı değerlerinin üst seviyelere ulaştığı gözlemlenmiştir.

FPN tabanlı modellerin tamamı doğrulama metriklerinde yüksek değerlere ulaşmıştır. Ancak hemen her modelde özellikle doğrulama IoU ve Dice katsayısı değerlerinde farklı seviyelerde dalgalanmalar gözlemlenmiştir. Bu dalgalanmalar validasyon setinin karmaşıklığına, veri çeşitliliğine, modelin bazı örnek türlerine karşı adaptasyon sürecine veya olası veri artırma işlemlerinin etkilerine bağlanabilir. Eğitim eğrilerinin stabil ve düzgün ilerlemesi modellerin temel öğrenme kapasitesinin güçlü olduğunu, doğrulama eğrilerindeki anlık dalgalanmaların ise genellikle geçici ve model başarısını temelden etkilemeyen türde adaptasyon sürecini yansıtan anlık değişimler olduğunu göstermektedir.

LinkNet modellerine ait eğitim ve doğrulama grafikleri

Bu bölümde LinkNet mimarisinin farklı kodlayıcıları (backbone) kullanarak transfer öğrenme ile eğitilmesi sırasında elde edilen eğitim ve doğrulama metriklerinin grafikleri (Şekil 35-41) analiz edilmektedir. Bu grafikler modellerin öğrenme hızı, kararlılığı ve genelleme kapasitesi hakkında önemli bilgiler sunmaktadır.

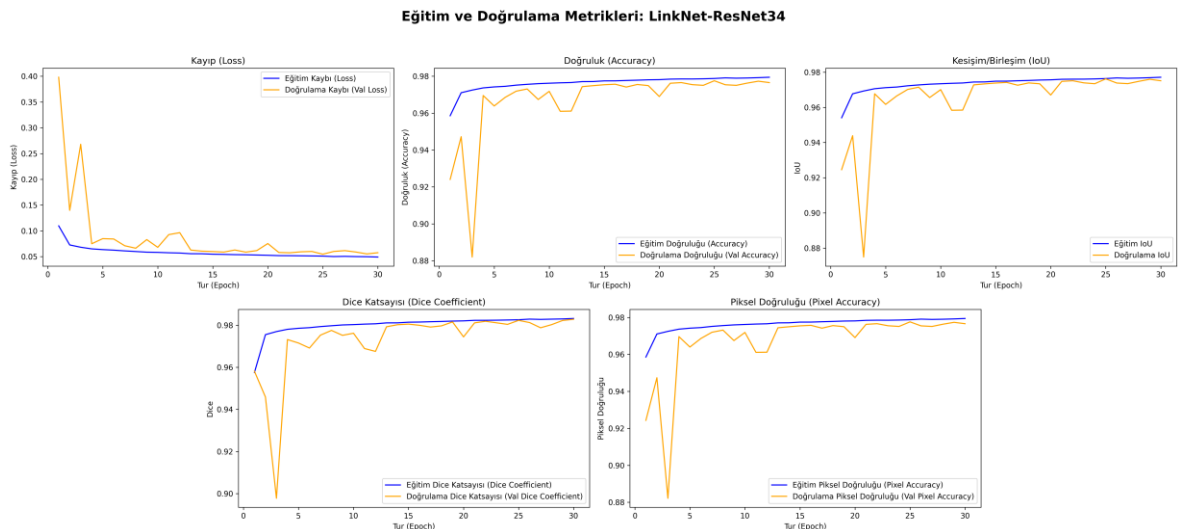
VGG16



Şekil 35. LinkNet/VGG16 modeli eğitim ve doğrulama grafiği.

Şekil 35'te gösterilen LinkNet/VGG16 modeline ait eğitim ve doğrulama kayıp değerlerinin turlar ilerledikçe istikrarlı bir şekilde azaldığı ve benzer değerlere yakınsadığı gözlenmiştir. Doğruluk, IoU, Dice katsayısı ve piksel doğruluğu metriklerinde eğitim ve doğrulama değerleri yüksek bir uyum içinde olup, eğitimin başlarında doğrulama metriklerinde küçük dalgalanmalar görülse de ilerleyen turlarla birlikte kararlılık sağlanmıştır. Metriklerdeki tutarlı yükseliş ile modelin genel olarak eğitim verilerine uyum sağladığı ve yüksek performans sergilediği söylenebilir.

ResNet34

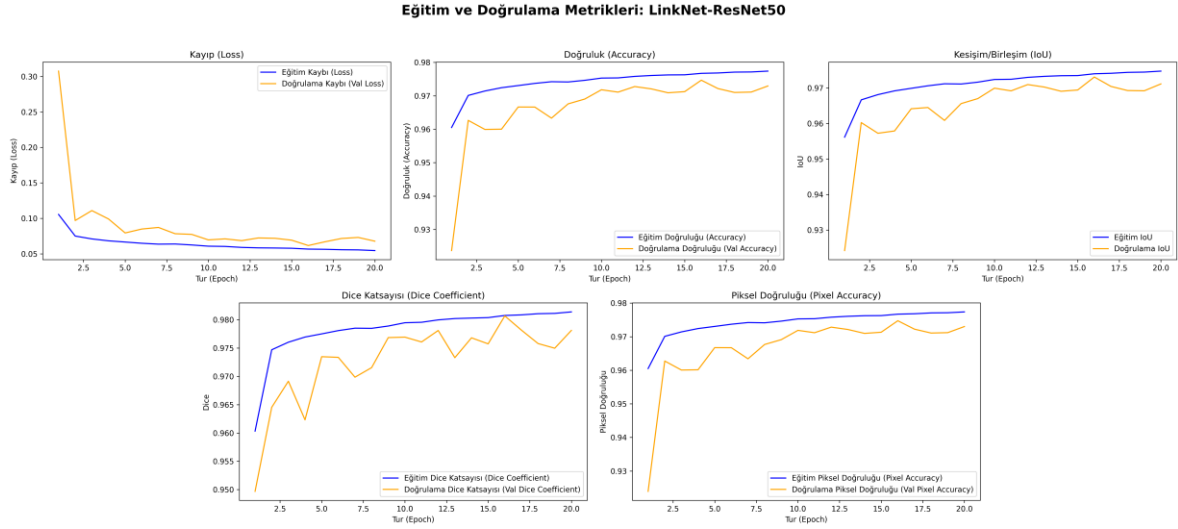


Şekil 36. LinkNet/ResNet34 modeli eğitim ve doğrulama grafiği.

Şekil 36'da gösterilen LinkNet/ResNet34 modeli eğitim eğrilerinde hızlı bir iyileşme ve stabil bir plato gözlemlenmiştir. Doğrulama metrikleri özellikle erken turlarda belirgin dalgalanmalar göstermiş olsa da ilerleyen aşamalarda eğitim eğrilerine yakın değerlerde

seyretmiş ve aşırı uyum (overfitting) eğilimi gözlenmemiştir. IoU, Dice katsayısı ve piksel doğruluğu metriklerinde eğitim ve doğrulama değerleri üst seviyede bir yakınsama göstermiştir. Bu durum modelin genelleme yeteneğinin yüksek olduğunu, eğitim sürecinin başarılı ve istikrarlı bir şekilde gerçekleştiğini göstermektedir.

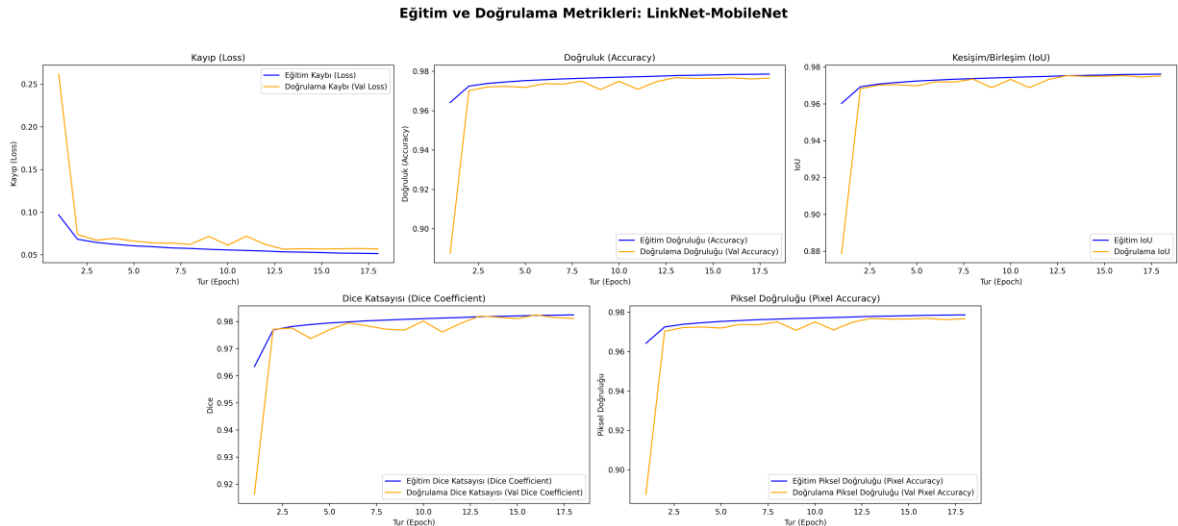
ResNet50



Şekil 37. LinkNet/ResNet50 modeli eğitim ve doğrulama grafiği.

Şekil 37’de gösterilen LinkNet/ResNet50 modeli eğitim sürecinde genel olarak başarılı olsa da doğrulama performansında belirgin dalgalanmalar gözlemlenmiştir. Eğitim kaybı düzenli şekilde azalırken, doğrulama kaybında belirgin dalgalanmalar olması modelin genelleme yeteneğinde bazı kısıtlamaların olabileceğine yorumlanabilir. Ancak turlar ilerledikçe doğrulama metrikleri artış göstermekte ve model genel olarak yüksek doğruluk ve IoU değerlerine ulaşmaktadır.

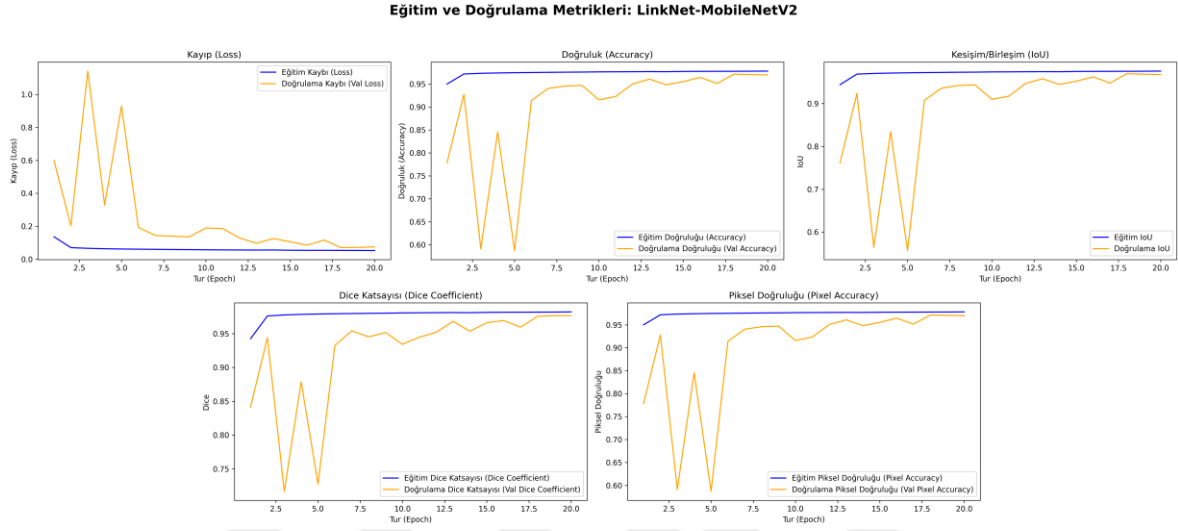
MobileNetV1



Şekil 38. LinkNet/MobileNetV1 modeli eğitim ve doğrulama grafiği.

Şekil 38’de gösterilen LinkNet/MobileNetV1 modelinin eğitim ve doğrulama kayıpları kısa süre içerisinde kararlı hale gelmiş ve diğer metriklerdeki değerlerin de hızlıca yüksek seviyelere ulaştığı gözlemlenmiştir. Modelin doğrulama ve eğitim seti performansı birbirine oldukça yakındır. Bu durum modelin genelleme yeteneğinin ve segmentasyon başarısının yüksek olduğunu göstermektedir.

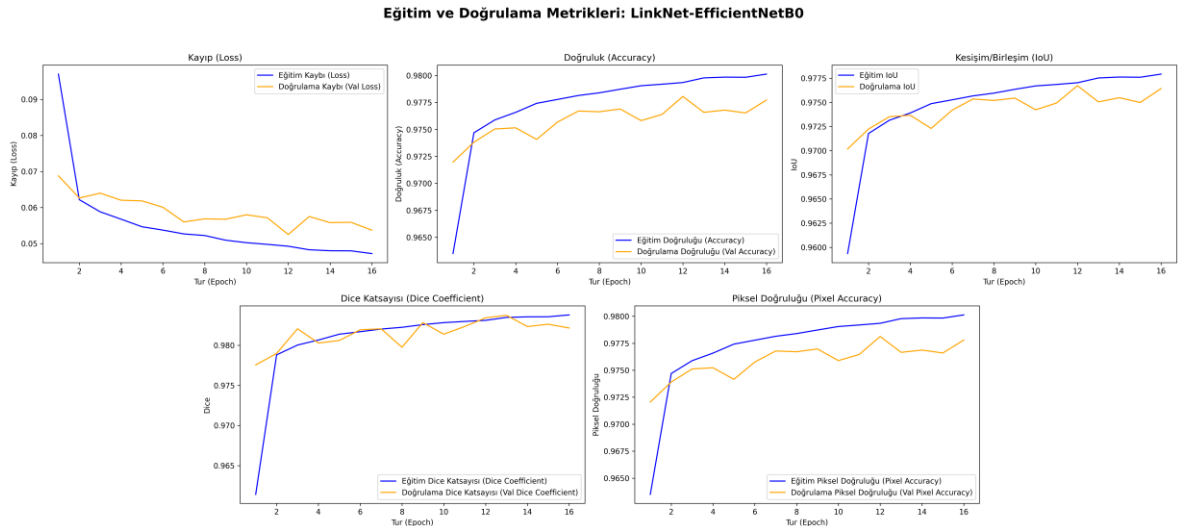
MobileNetV2



Şekil 39. LinkNet/MobileNetV2 modeli eğitim ve doğrulama grafiği.

Şekil 39’da gösterilen LinkNet/MobileNetV2 modeli eğitim eğrilerinde hızlı bir iyileşme ve plato oluşumu görülmüştür. Ancak doğrulama metriklerinde özellikle eğitimin başlarında belirgin dalgalanmalar oluşmuştur. İlerleyen turlarda metrikler yüksek seviyelere ulaşırsa da doğrulama verileri üzerinde görülen bu düzensiz yapı, modelin aşırı öğrenmeye eğilimli olduğuna ve doğrulama setinde kararsız performans gösterdiğine yorumlanabilir.

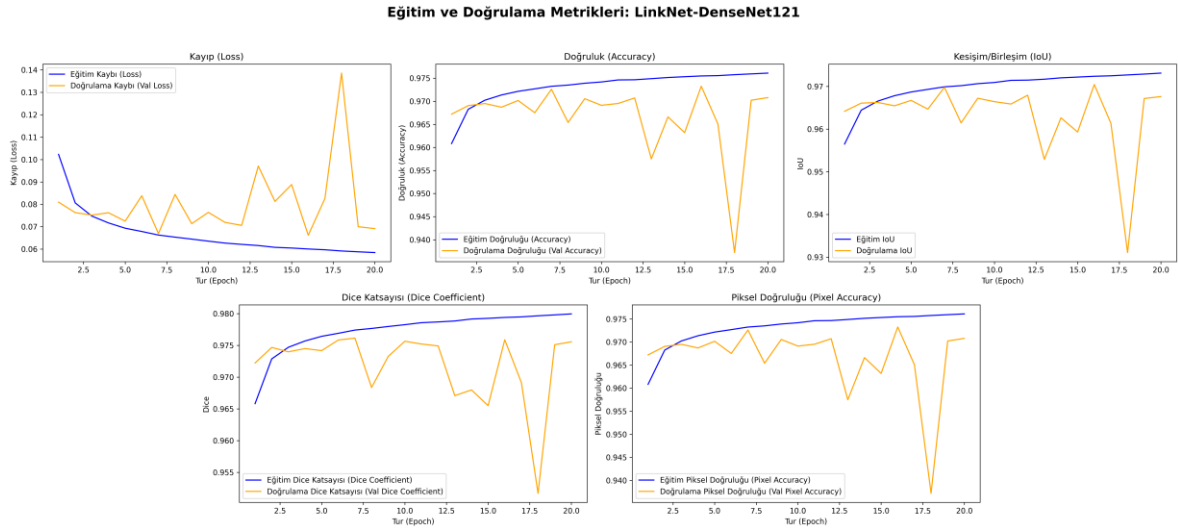
EfficientNetB0



Şekil 40. LinkNet/EfficientNetB0 modeli eğitim ve doğrulama grafiği.

Şekil 40'ta gösterilen LinkNet/EfficientNetB0 modeli eğitim eğrileri boyunca düzenli ve sürekli bir iyileşme gözlenmiş ancak doğrulama metriklerinde ise küçük dalgalanmalar olmuştur. Eğitim ve doğrulama eğrilerinin paralel seyretmesi modelin öğrenme kapasitesinin güçlü olduğunu ancak dalgalanmaların oluşması genelleme yeteneğinin kısıtlı olabileceğini göstermektedir.

DenseNet121



Şekil 41. LinkNet/DenseNet121 modeli eğitim ve doğrulama grafiği.

Şekil 41'de gösterilen LinkNet/DenseNet121 modeli eğitim eğrileri hızlı bir artış ve stabil bir seyir izlerken doğrulama eğrilerinde belirgin ve güçlü dalgalanmalar görülmektedir. Bu dalgalanmalar modelin aşırı öğrenmeye ve bazı doğrulama örneklerine karşı hassas olduğunu ve genelleme başarısının kısıtlı olabileceğini göstermektedir. Eğitim metriklerinin yüksek olmasına rağmen doğrulama verileri üzerindeki bu aşırı değişkenlik modelin güvenilirliğinin düşük olduğuna yorumlanabilir.

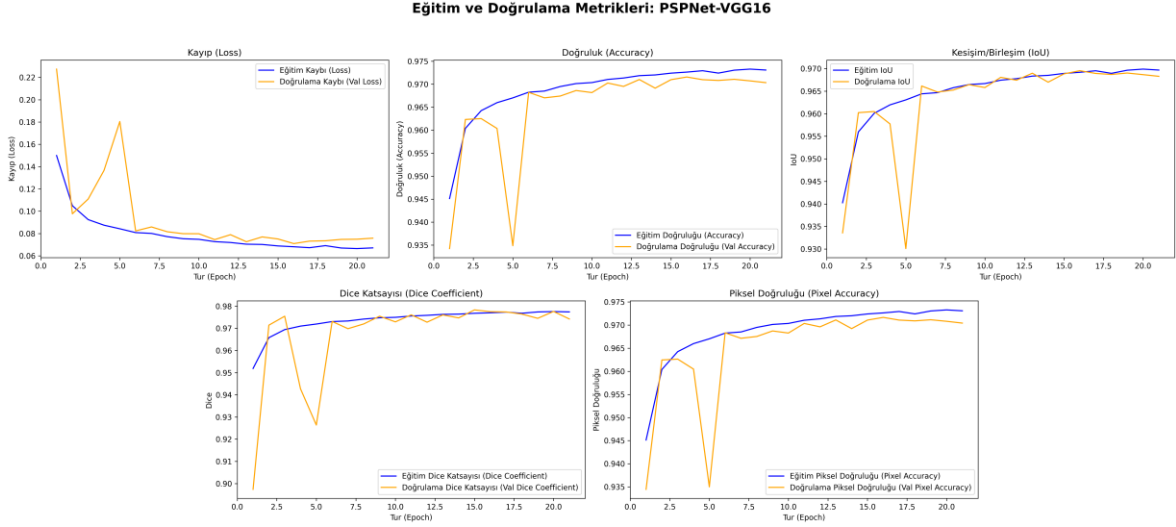
LinkNet modellerinin karşılaştırmalı değerlendirmesinde, VGG16 ve MobileNetV1 modellerinin eğitim ve doğrulama performansları bakımından en başarılı ve kararlı modeller olduğu gözlenmiştir. Bu modeller hem hızlı yakınsama hem de yüksek genelleme yeteneği ile öne çıkmaktadır. ResNet34 ve EfficientNetB0 modelleri de genel anlamda başarılı olmakla birlikte doğrulama verilerinde küçük dalgalanmalar sergilemiştir. Buna karşılık ResNet50 ve MobileNetV2 modellerinde doğrulama verileri üzerinde ciddi dalgalanmalar ile özellikle DenseNet121 modelinde kayda değer tutarsızlıklar görülmüştür.

PSPNet modellerine ait eğitim ve doğrulama grafikleri

Bu bölümde PSPNet (Pyramid Scene Parsing Network) mimarisinin farklı kodlayıcıları (backbone) kullanarak transfer öğrenme ile eğitilmesi sırasında elde edilen eğitim ve doğrulama

metriklerinin grafikleri (Şekil 42-48) üzerinden incelenmektedir. Bu analiz her bir modelin öğrenme kararlılığını, yakınsama hızını ve genelleme performansını değerlendirmektedir.

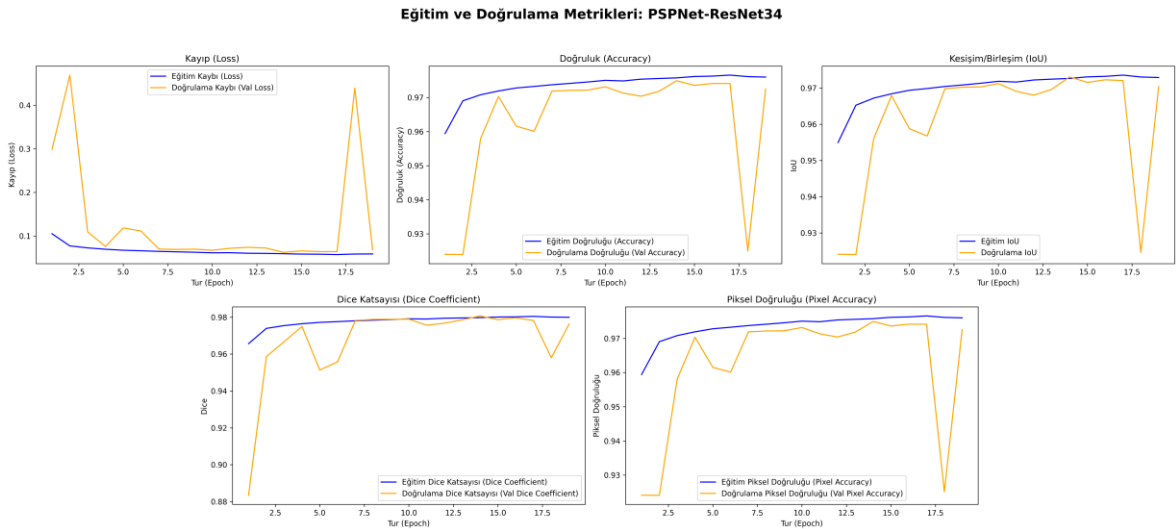
VGG16



Şekil 42. PSPNet/VGG16 modeli eğitim ve doğrulama grafiği.

Şekil 42’de gösterilen PSPNet/VGG16 modeline ait eğitim eğrilerinde kayıp ve doğruluk metrikleri istikrarlı şekilde iyileşirken, doğrulama eğrilerinde özellikle ilk birkaç turda ani dalgalanmalar gözlemlenmiştir. Bu dalgalanmalar validasyon setindeki bazı zorlu örneklerle modelin uyum sağlamaya çalışmasından yani adaptasyon sürecinden kaynaklanabilir. Tur (epoch) ilerledikçe plato oluşumu görülmekte ve hem IoU hem de Dice katsayısı değerleri doğrulama setinde istikrarlı biçimde artış göstermektedir.

ResNet34

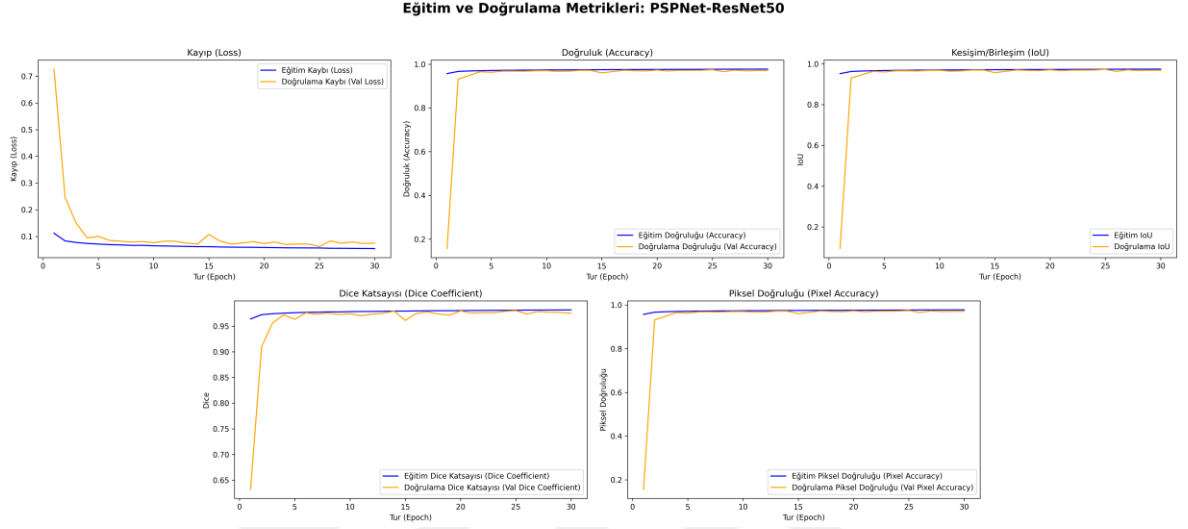


Şekil 43. PSPNet/ResNet34 modeli eğitim ve doğrulama grafiği.

Şekil 43’te gösterilen PSPNet/ResNet34 modeli eğitim kaybı ve metriklerinde hızlı bir yakınsama ve stabil seyir izlenirken, doğrulama eğrilerinde özellikle eğitimin başında ve

sonunda ani düşüşler ve yükselmeler görülmektedir. Bu ani değişimler validasyon setinde bulunan istisnai veya zor örneklerin model performansını geçici olarak etkilemesinden kaynaklanabilir. Ancak genel olarak model eğitimde başarı metrikleri yüksek düzeyde seyrettiği gözlemlenmiştir.

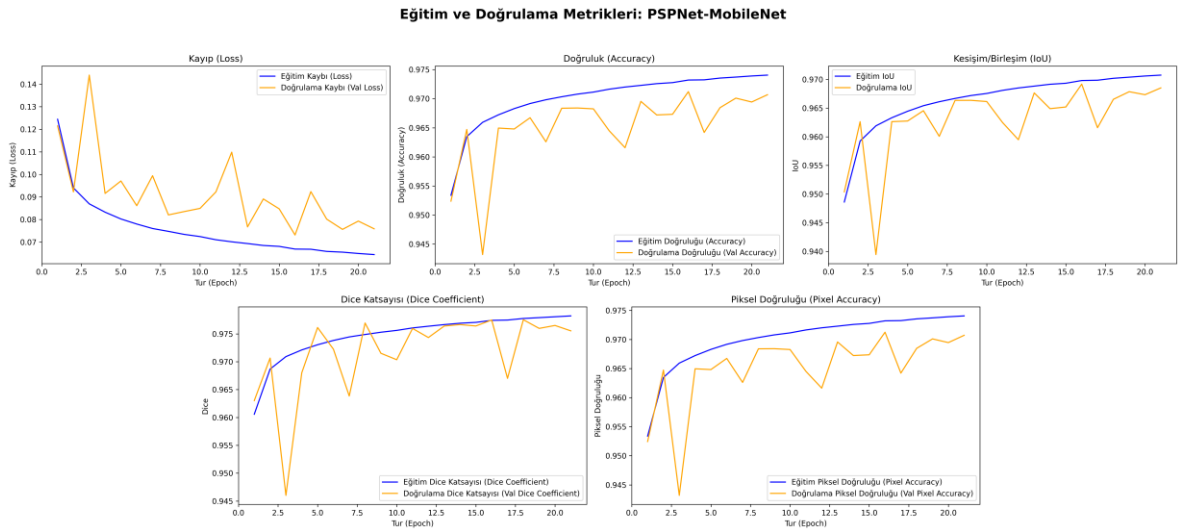
ResNet50



Şekil 44. PSPNet/ResNet50 modeli eğitim ve doğrulama grafiği.

Şekil 44'te gösterilen PSPNet/ResNet50 modeli hem eğitim hem de doğrulama eğrilerinde oldukça hızlı ve kararlı bir yakınsama gözlenmiştir. Başarı metrikleri neredeyse tamamen düz bir platoya ulaşmış, doğrulama ve eğitim eğrileri birbirine çok yaklaşmıştır. Bu durum modelin öğrenme sürecinin hem verimli hem de son derece kararlı olduğuna ve yüksek genelleme başarısına sahip olduğuna yorumlanabilir.

MobileNetV1

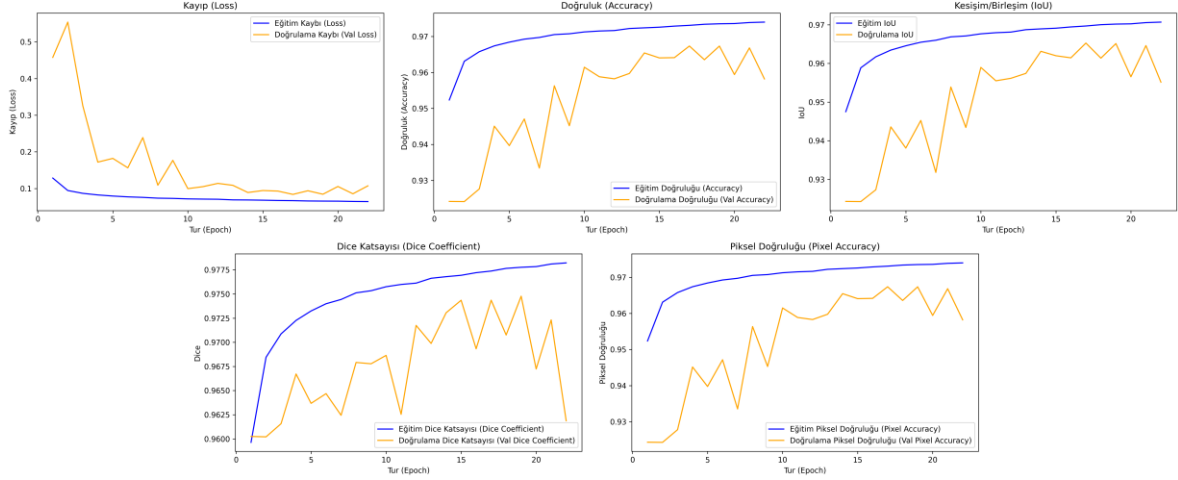


Şekil 45. PSPNet/MobileNetV1 modeli eğitim ve doğrulama grafiği.

Şekil 45'te gösterilen PSPNet/MobileNetV1 modelinin eğitim eğrileri düzenli bir iyileşme gösterirken, doğrulama eğrilerinde özellikle ilk turlarda dikkat çekici dalgalanmalar görülmektedir. IoU ve Dice katsayısı gibi metriklerde yüksek değerlere ulaşılsa da doğrulama metriklerindeki dalgalanmalar sebebiyle modelin genelleme başarısında tam bir tutarlılık sağlandığı söylenemez.

MobileNetV2

Eğitim ve Doğrulama Metrikleri: PSPNet-MobileNetV2

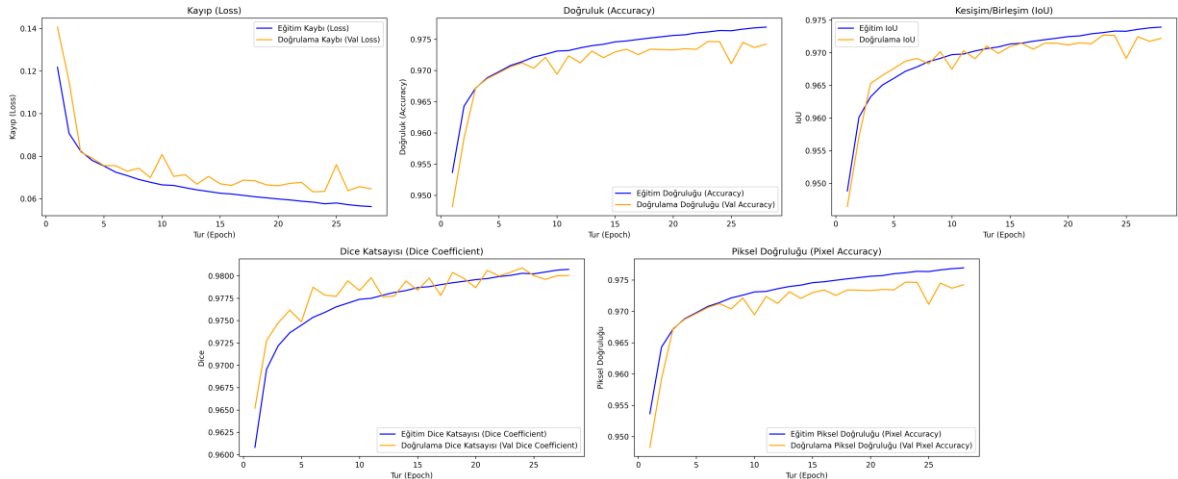


Şekil 46. PSPNet/MobileNetV2 modeli eğitim ve doğrulama grafiği.

Şekil 46'da gösterilen PSPNet/MobileNetV2 modeline ait eğitim eğrilerinde hızlı bir yakınsama gözlemlenirken doğrulama eğrilerinde ise uzun süreli ve görece yüksek seviyede dalgalanmalar devam etmiştir. Bu dalgalanmalar modelin validasyon setindeki zor ve çeşitli örneklerde genelleme yapmakta zorlandığını gösterebilir. Eğitimin ilerleyen turlarında IoU ve doğruluk metriklerinde artışlar görülse de modelin kararsız yapıda olduğu gözlemlenmiştir.

EfficientNetB0

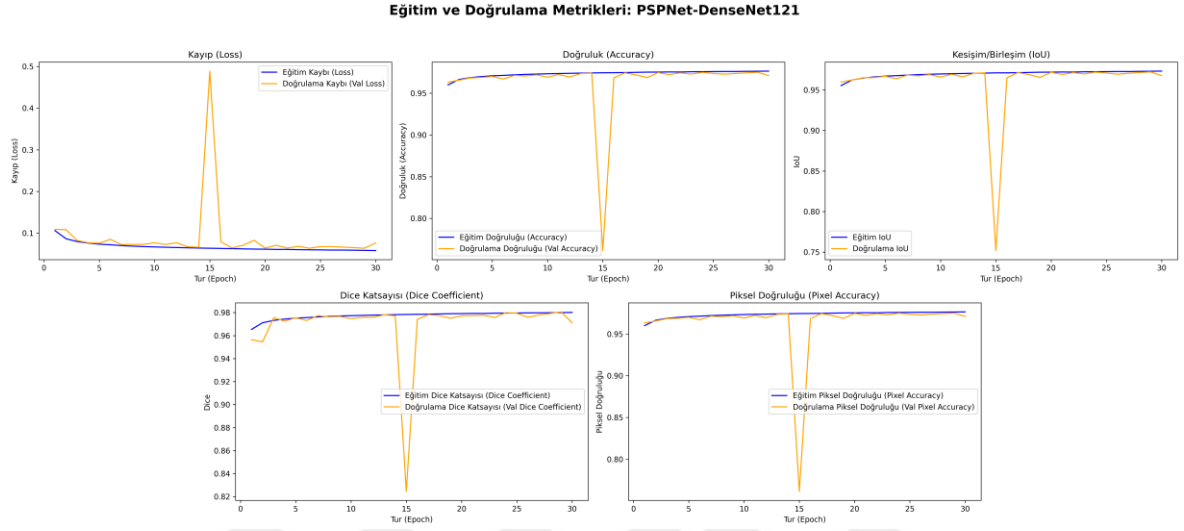
Eğitim ve Doğrulama Metrikleri: PSPNet-EfficientNetB0



Şekil 47. PSPNet/EfficientNetB0 modeli eğitim ve doğrulama grafiği.

Şekil 47’de gösterilen PSPNet/EfficientNetB0 modeli eğitim eğrileri boyunca stabil ve sürekli bir iyileşme izlenmekte doğrulama eğrilerinde ise bazı turlarda küçük dalgalanmalar görülmektedir. Eğitim ve doğrulama eğrilerinin birbirine yakın olması modelin güçlü bir genelleme kapasitesi sunduğunu göstermektedir. Eğitim ilerledikçe doğrulama metrikleri de yüksek ve stabil seviyelere ulaşmıştır.

DenseNet121



Şekil 48. PSPNet/DenseNet121 modeli eğitim ve doğrulama grafiği.

Şekil 48’de gösterilen PSPNet/DenseNet121 modeli eğitim eğrileri oldukça stabil ilerlerken, doğrulama eğrilerinde özellikle 15’inci turda (epoch) ani bir düşüş yaşandığı görülmektedir. Bu durum doğrulama setine ait bazı örneklerin model için aşırı zorlayıcı olması veya örnekleme sırasında dengesizlik yaşanmasından kaynaklanabilir. Bu dalgalanmaya rağmen genel olarak plato oluşumu gerçekleştiği ve istikrarlı bir eğitim sürecinde metriklerin genel seviyesinin yüksek olduğu gözlemlenmiştir.

PSPNet tabanlı modeller incelendiğinde eğitim sürecinde yüksek doğruluk, Dice katsayısı ve IoU değerlerine ulaşıldığı ancak doğrulama eğrilerinde farklı düzeylerde dalgalanmalar olduğu gözlemlenmiştir. VGG16, ResNet34, MobileNetV1 ve MobileNetV2 modellerinde doğrulama kaybı ve diğer metriklerde ani düşüş ve yükselişler dikkat çekmektedir. Bu durum modelin bazı doğrulama örneklerine karşı hassasiyet geliştirdiğini ve aşırı öğrenme eğilimi gösterebileceğini düşündürmektedir. Buna karşılık ResNet50, EfficientNetB0 ve DenseNet121 modelleri daha stabil öğrenme eğrileri sergileyerek hem eğitim hem de doğrulama setinde benzer başarı metriklerine ulaşmıştır.

Tablo 2. Modellerin Test Sonuçları

Model		Piksel Doğruluğu (Pixel Accuracy) %	IoU (Intersection over Union) %	Dice Katsayısı (Dice Coefficient) %	Keskinlik (Precision) %	Duyarlılık (Recall) %	Özgüllük (Specificity) %	Eğitim Süresi (s:d)	Veri Yığını Boyutu (Batch Size)	Parametre Sayısı (Milyon)
U-Net	VGG16	97,65	97,52	98,74	98,58	98,91	79,51	8:04	128	23,7
	ResNet34	97,83	97,69	98,83	98,71	98,95	81,75	3:18	128	24,4
	ResNet50	97,72	97,59	98,77	98,46	99,10	78,26	5:25	128	32,5
	MobileNetV1	97,99	97,87	98,92	98,75	99,09	82,51	3:58	128	8,3
	MobileNetV2	97,58	97,44	98,70	98,60	98,81	80,55	3:49	128	8
	EfficientNetB0	97,90	97,77	98,87	98,77	98,98	83,03	4:25	128	10,1
	DenseNet121	97,60	97,47	98,72	98,28	99,17	75,80	12:04	8	12,1
FPN	VGG16	97,62	97,49	98,72	98,44	99,02	77,75	15:22	64	17,5
	ResNet34	97,72	97,58	98,77	98,68	98,87	80,94	6:16	64	23,9
	ResNet50	97,70	97,57	98,77	98,44	99,10	77,60	9:15	64	26,9
	MobileNetV1	97,74	97,60	98,78	98,79	98,79	82,38	11:31	64	6,1
	MobileNetV2	97,79	97,67	98,81	98,52	99,12	79,05	12:24	64	5,2
	EfficientNetB0	97,87	97,74	98,86	98,80	98,92	82,66	8:03	64	7
	DenseNet121	97,75	97,62	98,79	98,50	99,09	78,19	21:50	8	9,9
LinkNet	VGG16	97,69	97,56	98,76	98,56	98,97	79,53	10:50	128	20,3
	ResNet34	97,93	97,80	98,89	98,55	99,23	79,06	4:20	128	21,6
	ResNet50	97,74	97,61	98,79	98,49	99,09	78,38	4:55	128	28,7
	MobileNetV1	97,87	97,74	98,86	98,80	98,92	82,75	2:20	128	4,5
	MobileNetV2	97,45	97,30	98,63	98,79	98,48	83,01	3:10	128	4,1
	EfficientNetB0	97,67	97,54	98,75	98,54	98,98	80,64	2:46	128	6,1
	DenseNet121	97,72	97,58	98,77	98,64	98,92	80,58	10:08	8	8,3
PSPNet	VGG16	97,51	97,37	98,67	98,28	99,06	75,27	7:20	128	10
	ResNet34	97,71	97,58	98,77	98,61	98,94	80,21	2:46	256	2,7
	ResNet50	97,64	97,50	98,73	98,54	98,94	77,89	5:37	256	3,8
	MobileNetV1	97,40	97,26	98,61	98,25	98,97	75,05	3:05	256	1,8
	MobileNetV2	97,06	96,92	98,43	97,63	99,26	65,25	3:15	256	1,6
	EfficientNetB0	97,16	97,01	98,48	98,11	98,86	71,39	4:06	256	1,7
	DenseNet121	97,81	97,69	98,83	98,46	99,20	78,01	12:31	8	3,1

Modellerin test verileri üzerindeki tahmin performansları farklı örnekler için (belirgin ve kalın harç yüzeyleri olup olmaması gibi) yüksek veya düşük değerlerde olduğu gözlemlenmiştir. Bu sebeple daha adil ve anlamlı bir değerlendirme için model tahmin

performanslarının ortalama deęerleri Tablo 2’de sunulmuştur. Modellerin deęerlendirilmesi genel segmentasyon kalitesini yansıtan metrikler (IoU, Dice katsayısı) ile sınıfa özgü başarıyı ölçen metrikler (duyarlılık, özgüllük) ve hesaplama verimlilikleri üzerinden gerçekleştirilmiştir.

Model başarısını belirleyen en temel metriklerden olan kesişim/birleşim (IoU) ve Dice katsayısı incelendiğinde, birbirine çok yakın deęerler olsa da U-Net mimarisinin dięer mimarilere kıyasla üstünlük sergiledięi gözlemlenmektedir. En yüksek performansı gösteren model, U-Net mimarisi ile MobileNetV1 kodlayıcısının birleşimi olmuş ve %97,87 IoU ile %98,92 Dice katsayısı skorları elde edilmiştir. Ayrıca %99,09 duyarlılık deęeri ile taş piksellerini gözden kaçırma oranının çok düşük olduęu, %82,51 özgüllük deęeri ile harç piksellerinin de büyük oranda doğru sınıflandırıldıęı gözlemlenmiştir. MobileNetV1 düşük parametre sayısı (8,3 Milyon) ve nispeten kısa eğitim süresi (4 saat 38 dakika) ile pratik uygulamalar için cazip bir seçenek olduęunu göstermektedir. U-Net/EfficientNetB0 modelinde %97,77 IoU, %98,87 Dice katsayısı, %98,98 duyarlılık ve %83,03 özgüllük deęeri ile karşılaştırılan modeller arasında en iyi harç segmentasyonu elde edilmiştir. U-Net mimarisinde parametre sayısı yüksek olmasına rağmen, benzer veri yığını boyutları ile eğitilen modellere kıyasla eğitim süreleri görece yüksek değildir. Bu bulgular U-Net mimarisinin atlama bağlantıları (skip-connections) aracılığıyla hem düşük seviyeli mekânsal detayları hem de yüksek seviyeli anlamsal bağlamı etkin bir şekilde birleştirme kapasitesinin yüksek olduęunu ve bu segmentasyon görevi özelinde en başarılı sonuçları ürettiğini göstermektedir.

Segmentasyon görevlerinde özellikle sınıf dengesizlięi (taş pikselleri sayısının harç pikselleri sayısından bariz fazla olduęu) olduęunda, duyarlılık (taş piksellerinin doğru tespiti) ve özgüllük (harç piksellerinin doğru tespiti) metrikleri başarı kıyaslamasında öne çıkmaktadır. PSPNet/MobileNetV2 modeli %99,26’lık bir duyarlılık deęeri ile yüksek performans sergilemiştir. Bu yüksek duyarlılık modelin taş piksellerini kaçırma oranının son derece düşük olduęunu ve neredeyse tüm gerçek taş piksellerini başarıyla tanımladıęını göstermektedir. PSPNet’in piramit ayrıştırma modülü sayesinde küresel bağlamı etkili bir şekilde yakalaması ve MobileNetV2’nin verimli özellik çıkarma yeteneęi bu başarının arkasındaki temel faktörler olabilir. Ancak aynı model %65,25 özgüllük deęeri ile harç piksellerini tespit etme aşamasında tüm modellerin gerisinde kalmıştır. Bu durum modelin taş piksellerini kaçırmamak için daha geniş bir alanı taş olarak işaretleme eğiliminde olduęunu ancak bu durumun harç piksellerinin yanlış sınıflandırılmasına yol açabileceğini düşündürmektedir. Eğitilen dięer PSPNet modelleri genel olarak başarı metriklerinde yüksek skora ulaşırsa da özgüllük deęerleri dięer mimarilere (U-Net, FPN ve LinkNet) dayanan modellerin gerisinde kalmıştır. Çalışmada karşılaştırılan

modellerin duyarlılık değerleri (%98,48-%99,26) yani taşları doğru tespit etme oranları oldukça yüksektir. Buna karşın özgüllük değerleri (%65,25-%83,03) yani harçları doğru tespit etme oranları daha düşük kalmıştır. Bunun ana nedeni etiketli verilerde taş yüzeylerinin (taş piksel sayıları) harç yüzeylerinden (harç piksel sayıları) çok daha fazla olmasından kaynaklandığı düşünülmektedir. FPN/MobileNetV1 modeli %82,38 ve FPN/EfficientNetB0 modeli %82,66 özgüllük değerleri ile harç segmentasyonundaki başarıları ile öne çıkmaktadır. Ayrıca modellerin diğer performans metrikleri (IoU, Dice katsayısı, piksel doğruluğu, duyarlılık gibi) karşılaştırılan modellere yakın değerlerde ölçülmüştür.

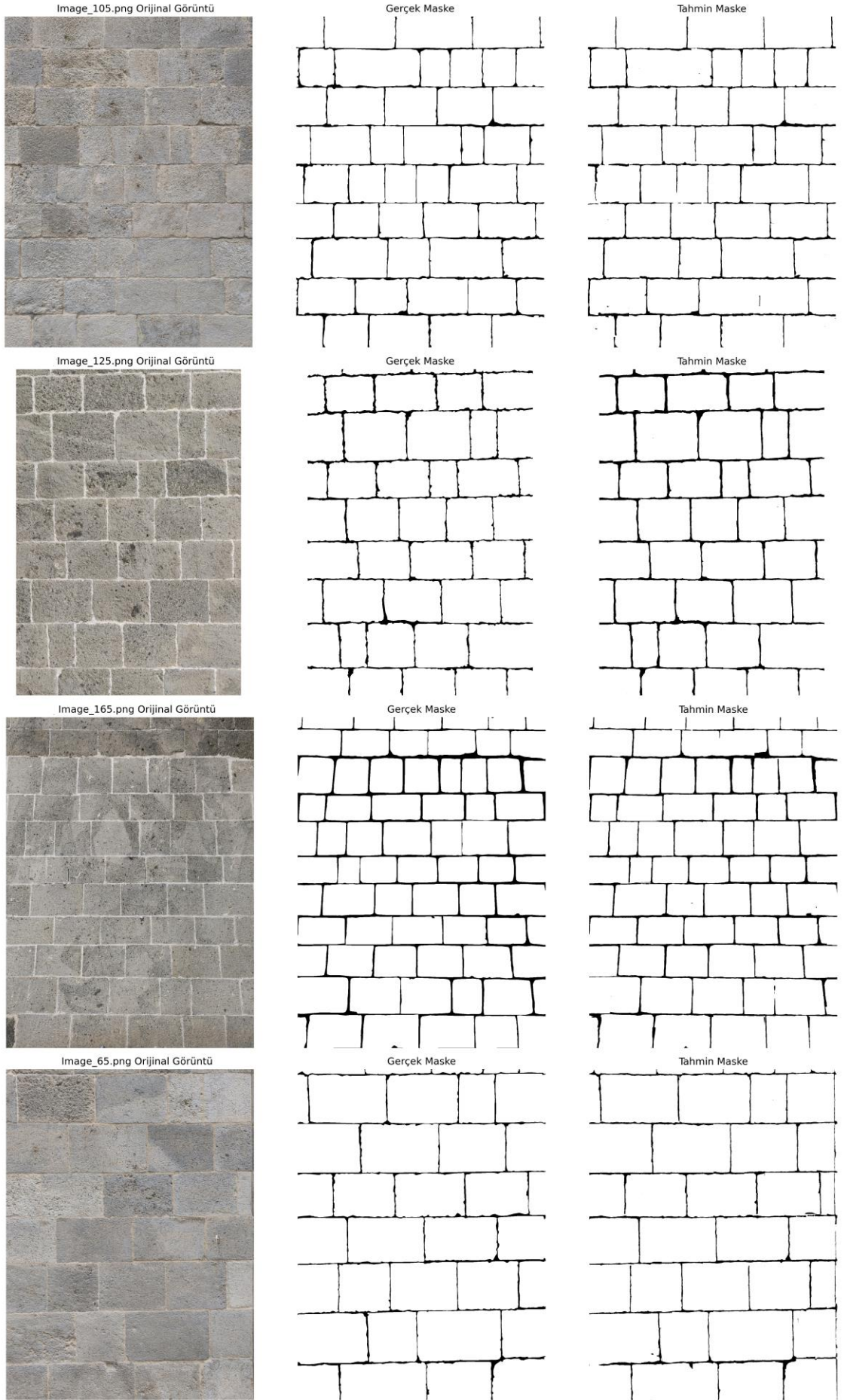
Eğitim hızı ve parametre sayıları incelendiğinde, LinkNet modelleri hem eğitim hızı hem görece düşük parametre sayıları ile karşılaştırılan modellere kıyasla daha iyi sonuçlar vermiştir. Düşük eğitim sürelerine rağmen genel olarak performans metrikleri yüksek seviyelere ulaşmıştır. Özellikle LinkNet/MobileNetV1 modeli en düşük eğitim süresi (2 saat 20 dakika) ve %97,74 IoU, %98,96 Dice katsayısı, %82,75 özgüllük değerleri ile yüksek başarı göstermiştir. Ayrıca LinkNet/MobileNetV2 modeli 3 saat 10 dakika eğitim süresi ve %83,01 özgüllük değeri ile hızlı bir eğitim ve harç tespitinde yüksek başarı oranlarına ulaşmıştır. Kısıtlı donanıma sahip sistemlerde (mobil cihazlar, gömülü sistemler vb.) hafif ve hızlı yapısıyla LinkNet/MobileNetV1 modeli pratik çözümler sunabilir.

DenseNet121 modelleri donanımsal kısıtlar sebebiyle düşük veri yığını boyutu (8) ile eğitilmiştir. Elde edilen segmentasyon performans metrikleri diğer modellere yakın olsa da eğitim süreleri (21 saat 50 dakika, 12 saat 31dakika) ciddi oranda artmıştır. Küçük veri yığını boyutları kullanımı daha fazla ağırlık güncellemeleri ve gradyan tahminleri nedeniyle eğitim süresini uzatmıştır. Ayrıca doğrulama grafiklerinde oluşan dalgalanmalar ile eğitim ve doğrulama eğrilerinin birbirinden ayrılması sebebiyle modellerin genelleme performansının düşük olduğu gözlemlenmiştir. Çalışmada kullanılan yüksek veri yığını boyutları (128, 256) ile genel olarak yüksek performansa karşı daha kısa eğitim süreleri elde edilebilmiştir. Düşük batch size değerlerine sahip modellerin eğitim ve doğrulama grafiklerinin bazılarında (U-Net/DenseNet121) dalgalanmaların olduğu bazılarında (PSPNet/DenseNet121) ise daha stabil eğrilerin olduğu gözlemlenmiştir. Aynı şekilde yüksek batch size değerlerine sahip modellerin eğitim ve doğrulama grafiklerinin bazılarında (PSPNet/MobileNetV2) dalgalanmaların olduğu bazılarında (LinkNet/MobileNetV1) ise daha stabil eğrilerin olduğu gözlemlenmiştir. Ancak veri yığını boyutlarının yüksek veya düşük olmasının modellerin genelleme performansı üzerindeki etkisine dair net bulgular tespit edilememiştir (Masters ve Lusch, 2018).

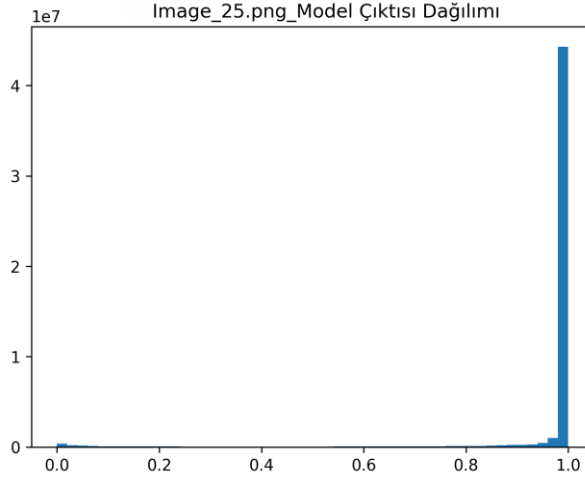
Görüntü İşleme ve Son İşlem (Post-Processing) Sonuçları

Derin öğrenme modelleri tarafından üretilen ham segmentasyon maskeleri, yüksek doğruluk oranlarına sahip olsalar da genellikle küçük gürültüler ve taş yüzeylerinde hatalı harç (delik) oluşumları içerirler. Bu nedenle maskelerin sonraki adımlarda (taşların tek tek etiketlenmesi ve boyutlarının ölçülmesi) güvenilir bir şekilde kullanılabilmesi için bir dizi son işlem (post-processing) adımından geçirilmesi kritik öneme sahiptir. Bu bölümde uygulanan iyileştirme adımları ve sonuçları görsel olarak analiz edilmektedir. Segmentasyon sonrası elde edilen ham maskelerde morfolojik işlemler, kapama (closing), erozyon (erode), harç yüzeylerini kalınlaştırma, mesafe dönüşümü ve havza (Watershed) algoritması gibi teknikler uygulanmıştır. Özellikle Watershed algoritması ile birleşik taş bölgeleri başarıyla ayrılmış, küçük ve hatalı harç adacıkları ise bağlantılı bileşen analizi ile elimine edilmiştir. Nihai segmentasyon maskeleri üzerinde yapılan ölçümler ile taş/harç ayrımının görsel ve sayısal olarak başarılı şekilde gerçekleştirilebildiği gösterilmektedir.

Performans metriklerine göre en başarılı olan U-Net/MobileNetV1 modelinin test verilerine ait tahmin maskesi örnekleri Şekil 49'da gösterilmiştir. Örnek olarak seçilen Image_25.png test görüntüsü üzerinde yapılan görüntü işleme adımlarına ait çıktılar ve yapılan işlem sonuçları aşağıda sunulmuştur.

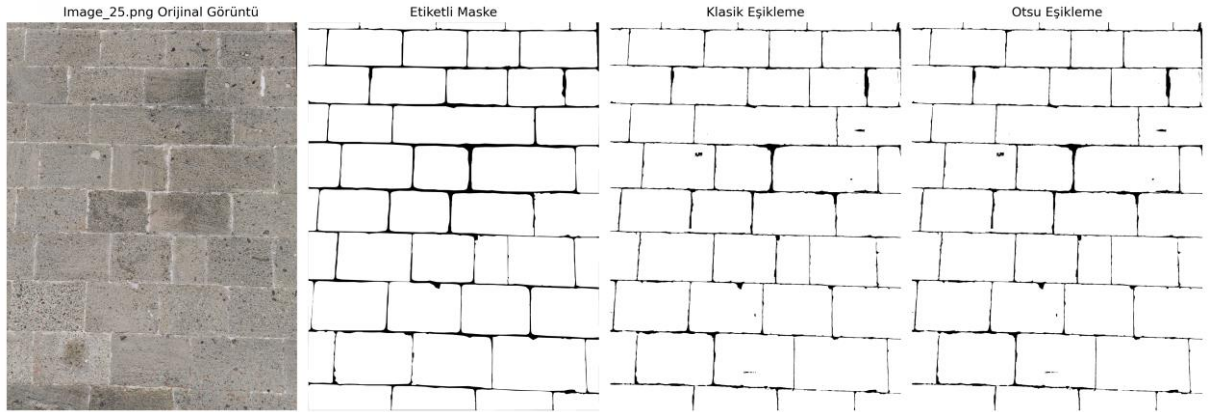


Şekil 49. U-Net/MobileNetV1 modeli tahmin maske örnekleri.



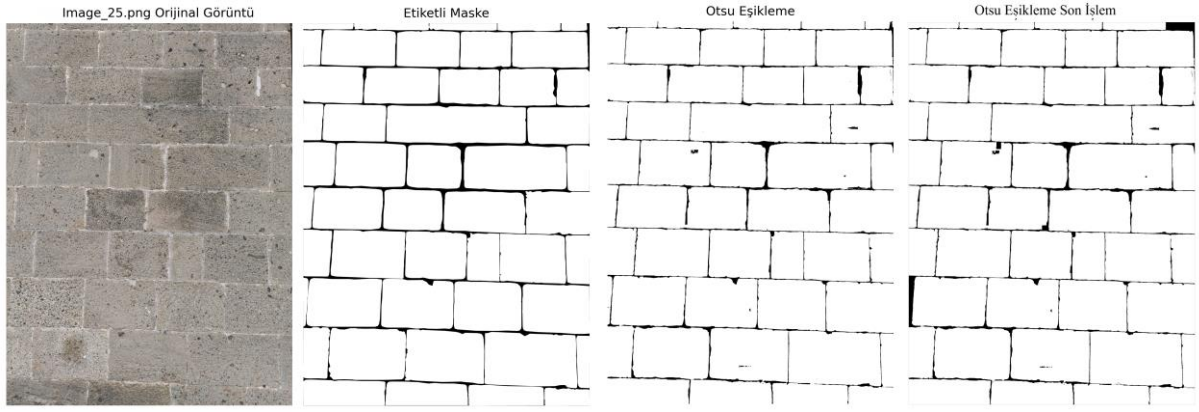
Şekil 50. Image_25 tahmin histogramı.

İlk olarak modelin örnek görüntü üzerindeki tahmin maskesinde piksel olasılık dağılımı (histogram) incelendiğinde, taş (1) piksel sayısının harç (0) piksel sayısından çok daha fazla olduğu Şekil 50’de görülmektedir. Dağılımın uçlara (0 ve 1) yığılması modelin çoğu pikselde kararlı davrandığını, taş ve harç hatlarının beyaz ve siyah olarak keskin şekilde ayrıldığına işaret eder.



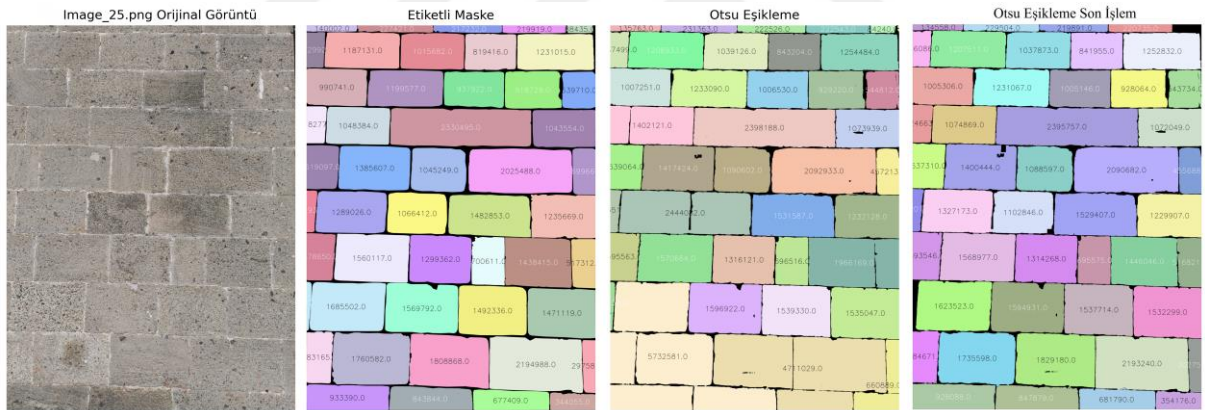
Şekil 51. Image_25 Otsu ve klasik eşikleme çıktısı.

Modelin tahmin çıktısına uygulanan klasik ve Otsu eşikleme yöntemlerine ait ikili maske sonuçları Şekil 51’de gösterilmiştir. Eşikleme sonuçları karşılaştırıldığında her iki eşikleme yönteminin de etiketli gerçek maskeye (ground truth) oldukça yakın sonuçlar ürettiği, ancak harç derzlerinde küçük kopukluklar ve taş yüzeylerinde minimal gürültüler içerdiği gözlemlenmektedir. Histogram grafiğinde görüldüğü üzere pikseller arasında siyah ve beyaz olarak net ve keskin bir ayrımın olması, klasik ve Otsu eşikleme sonuçlarının birbirine çok yakın olmasının nedeni olabilir. Sonuçların birbirine yakın olması sebebiyle örnek görsellerde Otsu eşikleme yöntemine ait çıktılar kullanılmıştır.



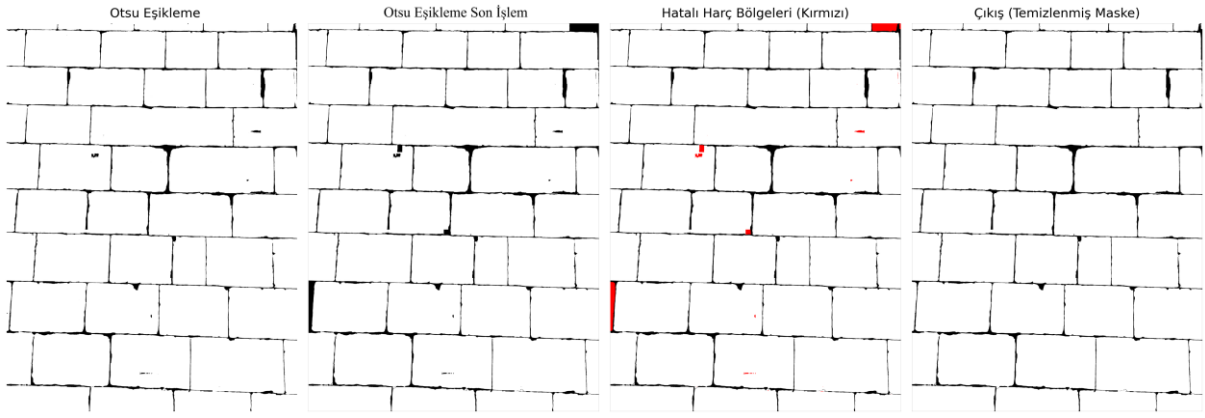
Şekil 52. Image_25 son işlem sonrası maske.

Elde edilen ham ikili maskeler üzerinde harç hatlarını belirginleştirmek ve gürültüyü ortadan kaldırmak için uygulanan morfolojik işlemlere ait sonuçlar Şekil 52’de gösterilmiştir. Kapama (closing) işlemiyle harç çizgilerindeki kopukluklar giderilmiş, genişleme/erozyon (dilation/erosion) adımlarıyla harç konturları kalınlık açısından dengelenmiştir. Harç kalınlaştırma adımıyla ince veya kopuk derzler güçlendirilmiş, böylece taş sınırları daha belirgin hale getirilmiştir. Son olarak Watershed algoritması ile yapışık taşların merkezleri tespit edilerek birbirinden ayrılması sağlanmıştır.



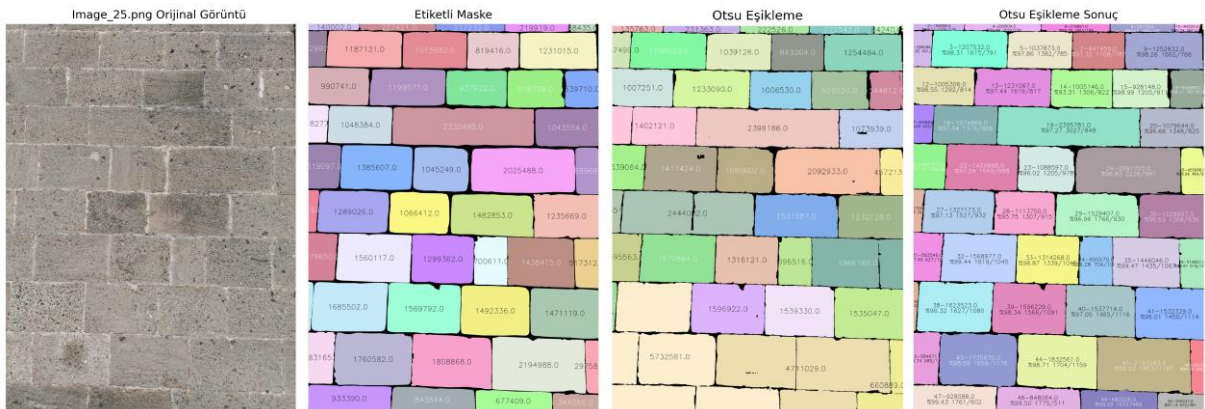
Şekil 53. Image_25 son işlem sonrası taş alanları.

Yapılan morfolojik işlemler sonucu taş sınırlarının belirlenmesindeki gelişimler taş alanları renklendirilerek Şekil 53’te gösterilmiştir. Ham maske üzerindeki otsu eşikleme ile taş sınırlarının keskin bir şekilde belirlenememesi sonucunda taş alanları yanlış tespit edilerek bazı büyük ve bitişik taşların tek bir renk (tek bir etiket) ile gösterildiği görülmektedir. Ancak yapılan morfolojik işlemler ile taş sınırlarının keskin bir şekilde ayrılması ve alanlarının belirlenmesinde etiketli gerçek maskeye çok daha yakın sonuçlar elde edildiği görülmektedir.



Şekil 54. Image_25 hatalı bölgelerin tespiti ve temizliği.

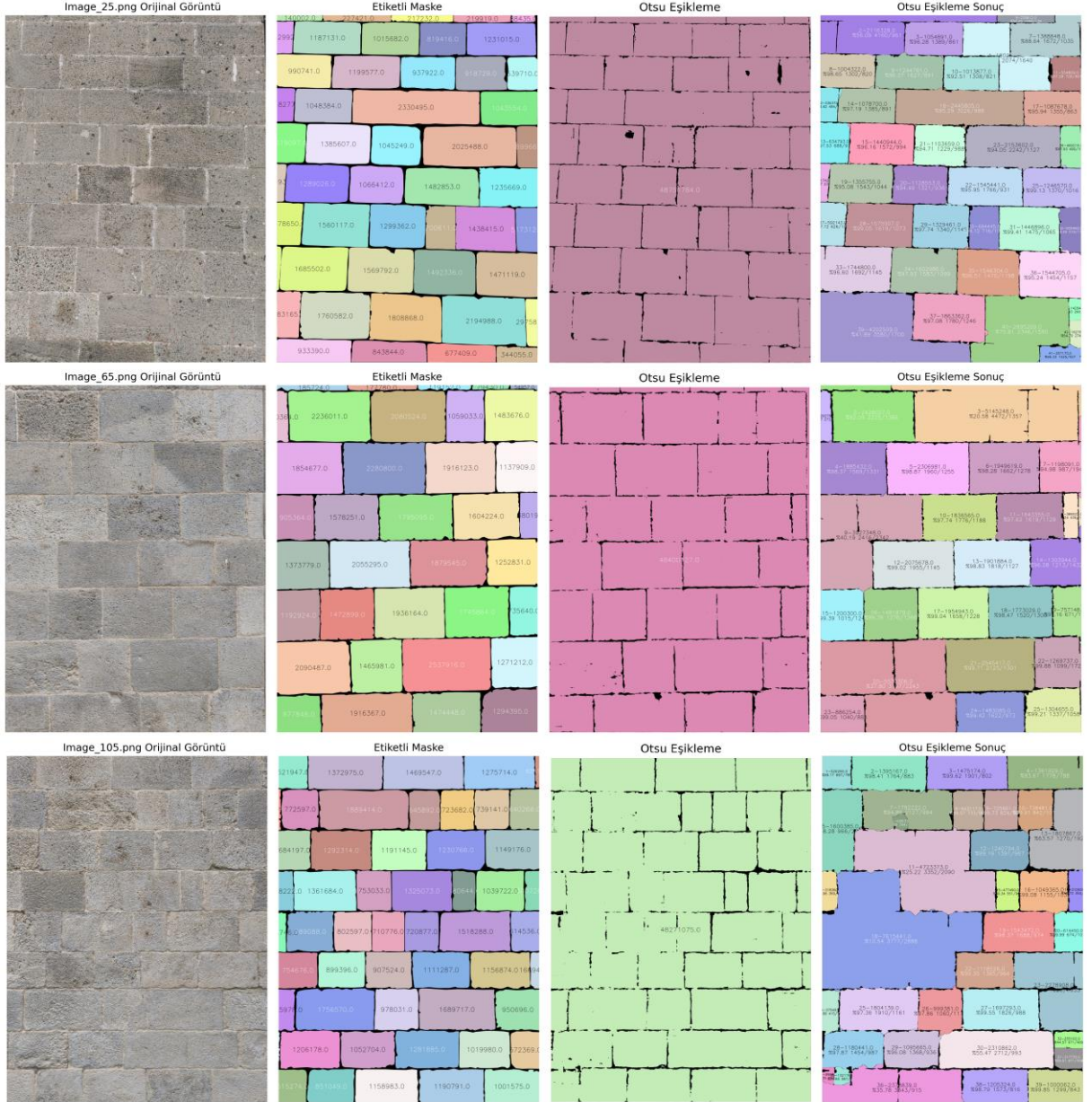
Modelin tahmin çıktısı veya morfolojik işlemler sonucunda ana harç ağına ait olmayan, taş yüzeylerinde izole şekilde kalan ve aslında taş bölgesine ait olduğu halde ‘harç adacığı’ olarak yanlış sınıflandırılan küçük alanlar gözlemlenmiştir. Bağlantılı bileşen analizi ile küçük/hatalı harç adacıklarının tespiti ve eliminasyonun yapıldığı işlem adımları Şekil 54’te gösterilmiştir. İlk olarak “Hatalı Harç Bölgeleri (Kırmızı)” sütununda gürültü olarak tanımlanan küçük harç adacıkları temizleme öncesi görsel olarak tespit edilmiş ve kırmızı renkle vurgulanmıştır. Görüldüğü gibi bu hatalı bölgeler genellikle taşların yüzeyinde yer alan küçük lekeler veya modelin yarattığı küçük gürültülerdir. Gürültü olarak etiketlenen bu kırmızı bölgeler elimine edilip taş olarak yeniden sınıflandırılarak temizlenmiş maske elde edilmiştir. Bu işlem sonucunda taş yüzeylerindeki lekelerin temizlenerek pürüzsüz hale getirildiği ve ana harç derz ağının korunduğu görülmektedir. Yapılan işlemler ile taşların içindeki yanlış siyah lekelerin (FN) ve harçtaki beyaz boşlukların (FP) anlamlı ölçüde azaltılması sonucunda özgüllük (specificity) değerinin artmasına, taş sınırlarının belirginleşerek kesintisizleşmesi IoU ve Dice katsayısı değerlerinin yükselmesine doğrudan katkı sağladığı gözlemlenmiştir.



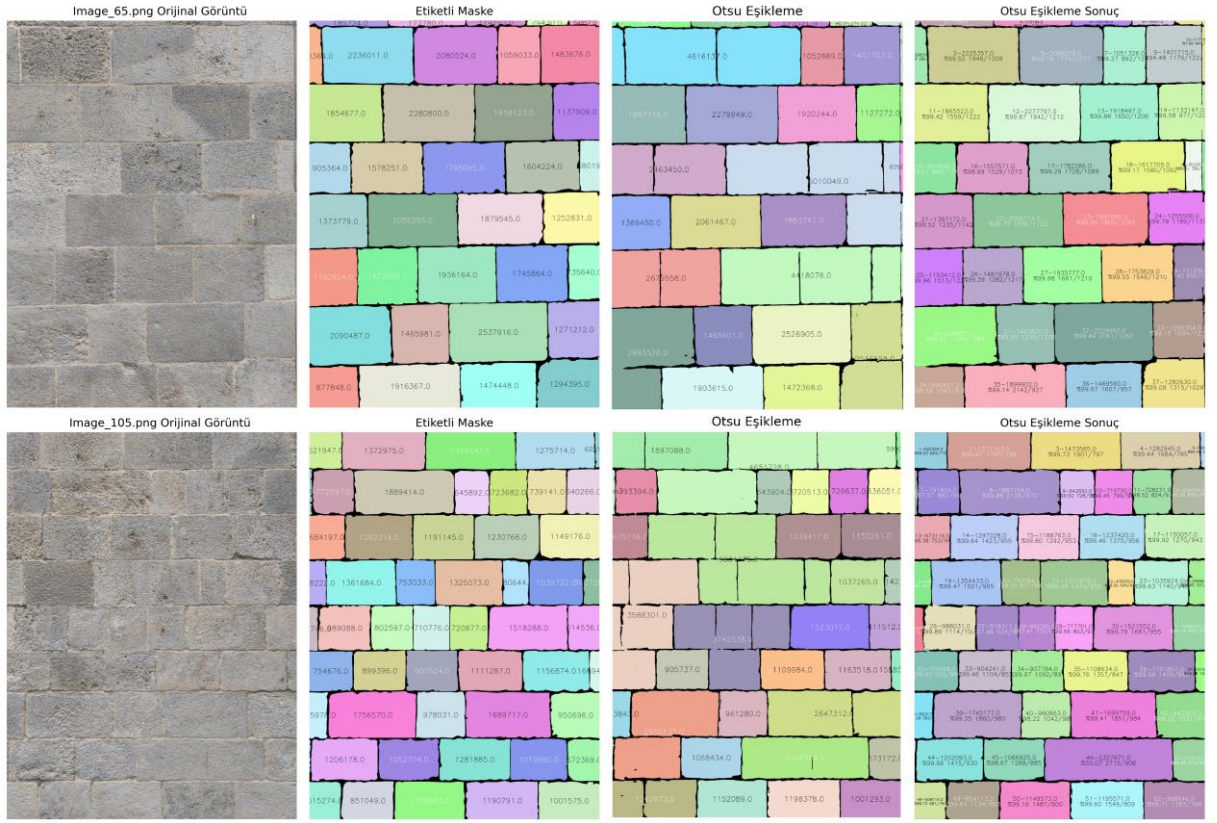
Şekil 55. Image_25 sonuç görseli.

Uygulanan morfolojik işlemler, bağlantı bileşen analizi, Watershed algoritması ve hatalı harçların temizliği sayesinde birleşik bölgeler başarılı bir şekilde ayrılarak her bir taşın farklı bir renk (farklı bir etiket) aldığı ve etiketli gerçek maskeye oldukça yakın sonuçlar ürettiği Şekil

55’te görülmektedir. Bu görüntü uygulanan son işlem adımlarının, birleşik taşları ayırma ve her bir taşı ayrı bir nesne olarak tanımlama gibi karmaşık görevler için etkili bir yöntem olduğunu kanıtlamaktadır. Taşların doğru bir şekilde ayrıştırılmasının ardından taş bazında alan ölçümü analizi yapılmış ve yüzdesel olarak gerçek maskeyle uyuma oranları gösterilmiştir. Ayrıca taş bazında ölçülen genişlik, yükseklik ve alan bilgisi değerleri hesaplamalarda kullanılmak üzere Excel formatında kaydedilmiştir.



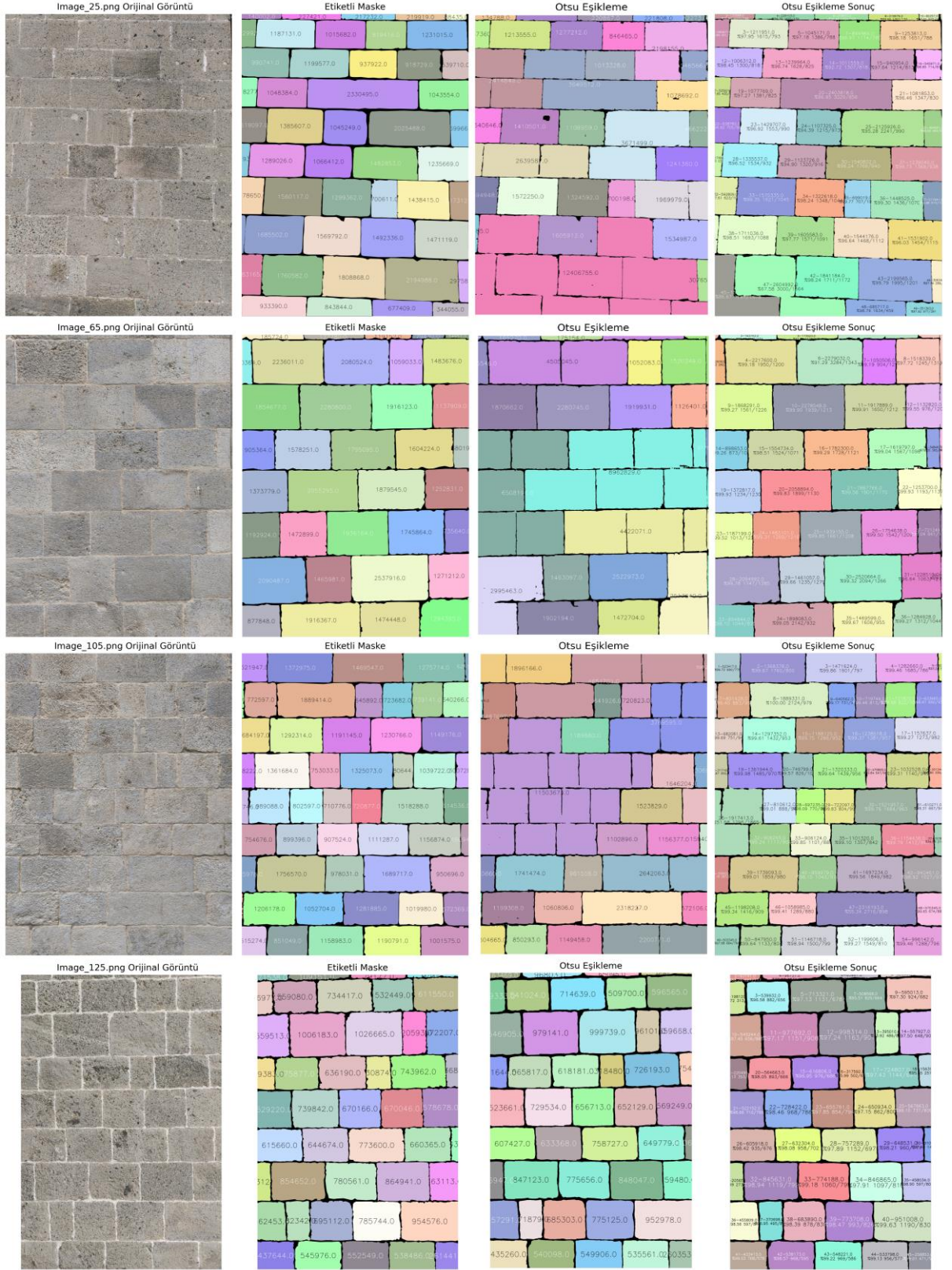
Şekil 56. PSPNet/MobilenetV2 modeli örnek sonuç görselleri.



Şekil 57. U-Net/MobilenetV1 modeli örnek sonuç görselleri.

Şekil 56’da PSPNet/MobilenetV2 ve Şekil 57’de U-Net/MobilenetV1 modellerinin test verileri üzerindeki örnek çıktılarına yer verilmiştir. Bu görsellerde ilk sütunda orijinal duvar görüntüsü, ikinci sütunda taş ve harç bölgelerini gösteren etiketlenmiş gerçek maske, üçüncü sütunda modelin ham tahmini ve dördüncü sütunda ise tahmin sonuçlarına uygulanan görüntü işleme teknikleri sonrasındaki nihai çıktılar sunulmaktadır. Görüldüğü üzere her iki modelin ham tahminlerinde özellikle harç bölgelerinin sürekliliğinin bozulduğu, bazı ince harç çizgilerinin tamamen kaybolduğu ve kimi bölgelerde taş/harç sınırlarının belirsizleştiği dikkat çekmektedir. Bu eksiklikler taş bölgelerinin farklı renklerle renklendirilmesi sayesinde daha görünür hale getirilmiştir. Modellerin ham segmentasyon çıktılarına, görüntü işleme teknikleri uygulanarak yapılan iyileştirmeler sonucunda taş sınırlarının gerçek maskeye çok daha yakın seviyede olduğu görülmektedir. Bununla birlikte görüntü işleme tekniklerinin başarısı doğrudan modelin ürettiği tahmin kalitesine bağlıdır, yani model ne kadar doğru ve tutarlı bir tahmin üretirse, son işlem adımlarının elde edeceği iyileştirme de o denli başarılı olmaktadır. Modeller arasında karşılaştırma yapıldığında, PSPNet/MobileNetV2 modelinin gerek görsel kalite gerekse Tablo 2’de raporlanan performans metrikleri açısından diğer modellere göre geride kaldığı açıkça görülmektedir. Bu modelde taş ve harç sınırlarının daha fazla belirsizlik içerdiği, harç bölgelerinin çoğu yerde kesintili ve eksik tahmin edildiği tespit edilmiştir. Görüntü işleme teknikleri PSPNet/MobileNetV2 modelinin çıktılarında belirgin bir iyileşme

sağlamış olsa da elde edilen nihai maskelerin genel benzerlik performansı gerçek maskelere kıyasla yetersiz kalmıştır.



Şekil 58. LinkNet/MobilenetV1 modeli örnek sonuç görselleri.

Karşılaştırılan modellere kıyasla daha düşük parametre sayısı ve eğitim süresine rağmen yüksek başarı skorlarına ulaşan LinkNet/MobileNetV1 modeline ait segmentasyon örnekleri

Şekil 58’de gösterilmiştir. Modelin ham tahmin sonuçları incelendiğinde taş ve harç sınırlarının bazı bölgelerde belirsizleştiği, ince harç çizgilerinin sürekliliğinin bozulduğu ve birleşik taş bölgelerinin oluştuğu görülmektedir. Ancak uygulanan morfolojik işlemler ve Watershed algoritması gibi görüntü işleme adımlarıyla yapılan iyileştirmeler ile etiketli gerçek maskeye yakın sonuçlar elde edilmiştir. Bu iyileştirmeler sonucunda elde edilen segmentasyon çıktılarının, gerçek maskeyle yüksek oranda örtüştüğü ve taş boyutlarının tutarlı şekilde belirlendiği görülmektedir. Bu karşılaştırmalı örnekler derin öğrenme modellerinin segmentasyon kalitesinin, yalnızca ham tahmin metriklerinde değil aynı zamanda son işlem sonrası elde edilecek nihai başarının da belirleyici unsuru olduğunu göstermektedir.

Sonuçların Karşılaştırılması ve Tartışılması

Tablo 3. Önerilen Yaklaşımın Benzer Çalışmalarla Karşılaştırılması

Çalışma	Yöntem	Veri Seti	IoU (%)	F1 Skor / Dice Katsayısı (%)	Doğruluk (%)
Ibrahim vd. (2019) taş segmentasyonu	U-Net	162 adet 512x512 çözünürlüklü görüntü	74,88	83,98	-
Ibrahim vd. (2020) taş segmentasyonu	U-Net + Çekişmeli Ağlar (GAN) + Watershed	532 adet 512x512 çözünürlüklü görüntü	63,78	74,18	-
Idjaton vd. (2021) taş segmentasyonu	DeepLabV3Plus	245 adet 256x256 çözünürlüklü görüntü	-	98,00	-
Dang vd. (2022) çatlak tespiti	DeepLabV3Plus + ResNet50	2670 adet 224x224 çözünürlüklü görüntü	97,00	98,00	98,00
Loverdos ve Sarhosis (2022) taş segmentasyonu	DeepLabV3Plus + Xception	2814 adet 224x224 çözünürlüklü görüntü	-	96,29	96,86
Haciefendioğlu vd. (2023) çatlak tespiti	CNN + ResNet50	502 adet 128x128 çözünürlüklü görüntü	-	99,00	98,44
Lucho vd. (2024) taş segmentasyonu	SAM + SVM	245 adet 256x256 çözünürlüklü görüntü	76,37	85,89	96,44
Yue vd. (2025) çatlak tespiti	DeepLabV3Plus	3290 adet 512x512 çözünürlüklü görüntü	63,50	76,70	74,90
Önerilen Yaklaşım	U-Net + MobileNetV1	108194 adet 256x256 çözünürlüklü görüntü	97,87	98,92	97,99

Tablo 3’te sunulan örnek çalışmalarda yerel ve çalışmaya özel veri setleri oluşturulmuştur. Bu sebeple modelleri doğrudan karşılaştırmak adil ve anlamlı sonuçlar vermeyecektir. Ancak taş segmentasyonu ve çatlak tespiti amacıyla gerçekleştirilen bu çalışmalarda çeşitli derin öğrenme modelleri kullanılmıştır. Çalışmalarda birden çok model geliştirilmiş ve en yüksek başarıya ulaştığı raporlanan modellere ait sonuçlar paylaşılmıştır.

Görüldüğü üzere literatürdeki çeşitli örneklerde çalışılan alanın ve veri seti örneklerinin farklı olması, farklı modellerin başarılı sonuçlar üretmesini sağlamıştır. Genel olarak U-Net ve DeepLabV3Plus mimarilerinin transfer öğrenimi ile eğitildiği modellerde alanına göre yüksek başarı oranlarına ulaşıldığı görülmüştür. Yapılan çalışmada bu yaklaşımdan yola çıkarak oluşturulan yerel veri seti ile çeşitli mimariler ve farklı omurgalardan oluşan modeller eğitilerek taş/harç segmentasyonu yapılmıştır. Örnek çalışmalardan farklı olarak model başarıları daha detaylı metrikler ile karşılaştırılmıştır. Örneğin harç doğruluğunu gösteren özgüllük değeri hesaplanarak, özellikle taş ve harç piksel sayısının dengesiz olduğu etiketlerde model başarısı sadece IoU, Dice katsayısı ve doğruluk metrikleri ile sınırlandırılmamıştır. Görüntüdeki taş piksellerinin harç piksellerine kıyasla çok daha fazla alana sahip olması değerlendirme metriklerinde önemli bir yanlılığa yol açabilmektedir. Özellikle taş bölgelerinin tespit edilmesi görece daha kolay olduğundan modellerin genel doğruluk, IoU, Dice katsayısı ve duyarlılık değerleri yüksek çıkabilmektedir. Ancak harç bölgelerinin sınırlı ve daha zor ayırt edilebilir yapısı bu sınıfın yanlış sınıflandırılma olasılığını artırmakta ve buna bağlı olarak özgüllük değerlerinin diğer metriklerle kıyasla oldukça düşük kalmasına neden olmaktadır. Bu analiz segmentasyon modellerinin performansını değerlendirirken sadece genel doğruluk metriklerine değil aynı zamanda uygulamanın özel ihtiyaçlarına göre duyarlılık ve özgüllük gibi sınıf bazlı metriklerin de dikkate alınması gerektiğini vurgulamaktadır. Bu sayede gerçek dünya problemlerine en uygun ve güvenilir çözümler geliştirilebilir.

SONUÇLAR VE ÖNERİLER

Bu çalışmada tarihi yığma taş yapıların dış duvarlarındaki taş boyutlarının belirlenmesi ve yapısal analizlerine destek oluşturmaya amacıyla, taş ve harç bölgelerinin otomatik olarak segmentasyonu için derin öğrenme ve görüntü işleme tekniklerine dayalı hibrit bir yaklaşım geliştirilmiştir. Çalışmada U-Net, FPN, LinkNet ve PSPNet gibi dört farklı segmentasyon mimarisi; VGG16, ResNet34/50, DenseNet121, MobileNetV1/V2 ve EfficientNetB0 gibi yedi farklı önceden eğitilmiş omurga (backbone) kullanılarak transfer öğrenme yöntemiyle eğitilmiştir. Modellerin performansları hem kamuya açık bir veri seti hem de Erzurum ilindeki tarihi yapılardan elde edilen yerel bir veri setinin birleştirilmesiyle elde edilen çeşitli görseller üzerinde değerlendirilmiştir. Böylece modellerin farklı malzeme türleri ve duvar dokularındaki genelleme kapasitesinin artırılması hedeflenmiştir. Yüksek çözünürlüklü duvar görsellerinin parçalara (patch) bölünmesiyle oluşturulan geniş ölçekli veri seti ile derin öğrenme modellerinin eğitiminde önemli bir avantaj sağlamış ve taş/harç ayırımında detay kaybının önüne geçilmiştir.

Deneysel bulgular geliştirilen derin öğrenme modellerinin yığma taş duvarlardaki taş ve harç bölgelerini yüksek bir başarıyla ayırtılabildiğini göstermiştir. Modellerin performans metrikleri incelendiğinde, U-Net mimarisinin özellikle MobileNetV1 omurgası ile birleştirildiğinde diğer mimarilere kıyasla daha iyi sonuçlar verdiği tespit edilmiştir. U-Net/MobileNetV1 modeli %97,87 IoU ve %98,92 Dice katsayısı gibi metriklerde en yüksek değerler ile %82,51 özgüllük değeri elde ederek segmentasyon doğruluğu açısından en başarılı model olarak öne çıkmıştır. Bu başarının arkasında U-Net mimarisinin atlama bağlantıları (skip-connections) sayesinde hem düşük seviyeli mekânsal detayları hem de yüksek seviyeli anlamsal bağlamı etkin bir şekilde birleştirme kapasitesi yatmaktadır.

Sadece genel doğruluk oranlarının tek başına model başarısını yansıtmadığı, özellikle harç bölgelerinin sınırlı yapısı nedeniyle duyarlılık (recall) ve özgüllük (specificity) gibi sınıf bazlı metriklerin de dikkate alınmasının gerektiği belirlenmiştir. Çalışmada hesaplanan metrikler, modellerin taş bölgelerinde yüksek başarıya ulaşmasına karşın harç bölgelerinde görece daha düşük performans sergilediğini göstermektedir. Sınıf dengesizliğinin (taş piksellerinin harç piksellerine oranla fazla olması) getirdiği zorlukları aşmada, modellerin harç tespit başarısını ölçen özgüllük metriği kritik bir rol oynamıştır. Bu bağlamda U-Net/EfficientNetB0 modeli %83,03 ile en yüksek özgüllük değerine ulaşarak harç bölgelerini en doğru şekilde tespit eden model olmuştur. Bu durum modelin taş ve harç gibi ince yapıları ayırt etmedeki hassasiyetini göstermektedir.

Hesaplama verimliliği ve eğitim süresi açısından LinkNet/MobileNetV1 modeli en düşük eğitim süresi (2 saat 20 dakika) ve düşük parametre sayısına rağmen yüksek segmentasyon performansı (%97,74 IoU, %98,86 Dice katsayısı, %82,75 özgüllük) sergileyerek pratik uygulamalar için en uygun maliyetli ve etkin çözümlerden birini sunmuştur.

Derin öğrenme modellerinden elde edilen ham segmentasyon maskeleri üzerinde uygulanan morfolojik işlemler, Watershed algoritması ve bağlantılı bileşen analizi gibi son işleme (post-processing) adımlarının, modelin eksik veya hatalı tahminlerini (gürültü, kopuk harç hatları vb.) önemli ölçüde iyileştirdiği gözlemlenmiştir. Bu hibrit yaklaşım taşların birleşik alanlarını başarılı bir şekilde ayırmış ve her bir taşın geometrik özelliklerinin (alan, en-boy oranı vb.) hassas bir şekilde hesaplanmasına olanak tanımıştır. Böylece nihai segmentasyon maskelerinin mühendislik uygulamalarına daha uygun hale geldiği gözlemlenmiştir.

Sonuç olarak bu çalışma, Türkiye'deki yerel yığma taş yapıların özgün taş ve harç dokularını içeren özel bir veri seti oluşturulmasına, farklı derin öğrenme mimarilerinin karşılaştırmalı olarak değerlendirilmesine ve segmentasyon sonrası taş boyutlarının hesaplanmasına odaklanılması açısından literatüre özgün katkılar sağlamaktadır. Ayrıca derin öğrenme ve görüntü işleme tekniklerinin bir arada kullanılmasının, yığma taş duvarların karmaşık ve değişken yapılarının segmentasyonunda hızlı, nesnel ve yüksek doğruluklu bir yöntem sunduğunu kanıtlamıştır. Geliştirilen yaklaşım restorasyon ve koruma alanında çalışan uzmanlara, yapısal analizler için temel bir girdi olan taş ve harç dağılımını otomatik olarak belirleme imkânı sunarak geleneksel yöntemlere kıyasla önemli bir avantaj sağlamaktadır.

Araştırma sürecinde elde edilen bulgular göz önüne alınarak gelecekte yapılacak çalışmalara yönelik şu öneriler sunulabilir:

1. **Veri Setinin Zenginleştirilmesi:** Bu çalışmada kullanılan veri seti Erzurum ilindeki yapıların özgün niteliklerini barındırsa da farklı coğrafi bölgelere, dönemlere ve yapım tekniklerine ait daha çeşitli duvar örnekleriyle genişletilebilir. Ayrıca farklı aydınlatma koşulları, mevsimsel etkiler ve yüzey bozulmaları (bitki örtüsü, nem vb.) içeren görsellerin eklenmesi modellerin genelleme kabiliyetini daha da artıracaktır.
2. **Farklı Derin Öğrenme Mimarilerinin İncelenmesi:** Derin öğrenme mimarilerinin geleneksel görüntü işleme yöntemleriyle entegrasyonu başarılı sonuçlar doğurmuştur. Gelecek çalışmalarda GAN (Generative Adversarial Networks) tabanlı veri üretimi, Transformer tabanlı segmentasyon modelleri, SAM (Segment Anything Model) gibi büyük ölçekli önceden eğitilmiş modeller veya daha yeni nesil anlamsal segmentasyon modelleri ile hibrit yaklaşımlar denenebilir. Bu tür modellerin küresel bağlamı anlama yetenekleri karmaşık taş desenlerinin çözümünde başarılı sonuçlar sunabilir.

3. **Hasar Tespiti ve Analizi:** Geliştirilen segmentasyon modellerinde her ne kadar taş/harç ayırımına odaklanılmışsa da duvar yüzeylerindeki çatlak, dökülme, malzeme kaybı, yüzey bozulmaları gibi yapısal hasarların otomatik olarak tespiti ve sınıflandırılması için bir temel olarak da kullanılabilir. Modeller taş ve harç bölgelerine ek olarak "hasar" sınıfını da tanıyacak şekilde yeniden eğitilerek yapısal sağlık izleme açısından önemli katkılar sunabilir.
4. **Gerçek Zamanlı Uygulamalar ve Mobil Entegrasyon:** Özellikle LinkNet ve MobileNet gibi hafif ve verimli modeller, mobil cihazlara (cep telefonları, insansız hava araçları vb.) veya gömülü sistemlerde gerçek zamanlı çalışacak uygulamalara entegre edilebilir. Bu sayede hızlı ve gerçek zamanlı segmentasyon ile tarihi yapıların periyodik bir şekilde taranarak izlenmesi sağlanabilir. Böylece restorasyon uzmanları, mühendisler veya araştırmacılar yapı yüzeylerini yerinde analiz edebilir, koruma ve bakım süreçlerinde önemli pratik faydalar sağlanabilir.
5. **Hata Analizi ve İyileştirme Stratejileri:** Modellerin yanlış sınıflandırma yaptığı durumların detaylı bir hata analizi yapılması, gelecekteki model geliştirmeleri için kritik bilgiler sağlayacaktır. Özellikle harç piksellerinin düşük özgüllükle sınıflandırıldığı durumlar üzerinde yoğunlaşarak, bu hataların nedenleri (ince harç derzleri, düşük kontrast, gölgelenme vb.) belirlenebilir ve bu sorunları gidermeye yönelik özel veri artırma teknikleri veya kayıp fonksiyonları geliştirilebilir. Uygulanan görüntü işleme adımları başarılı olsa da daha gelişmiş morfolojik filtreler veya makine öğrenmesi tabanlı gürültü temizleme algoritmaları ile son işlem sürecinin daha da otomatik ve hassas hale getirilmesi üzerine çalışmalar yapılabilir.

KAYNAKÇA

- Açııcı, F. K., Ertaş, Ş., & Sönmez, E. (2017). Sürdürülebilir turizm: Kültür turizmi ve kültürel miras. *Akademia Disiplinlerarası Bilimsel Araştırmalar Dergisi*, 3(1), 52-66.
- Aksoy B. (2020). *Python ile imgeden veriye görüntü işleme ve uygulamaları*. Ankara: Nobel Yayıncılık.
- Aydın, T., Bozkurt, F., & Karademir, R. M. (2025). Makine Öğrenimi ve Derin Öğrenme Metotları Kullanılarak Hurma Meyvesi Türü Tespiti. *Journal of the Institute of Science and Technology*, 15(2), 382-395.
- Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. 4, No. 4, p. 738). New York: springer.
- Bozkurt, F. (2023). Skin lesion classification on dermatoscopic images using effective data augmentation and pre-trained deep learning approach. *Multimedia Tools and Applications*, 82(12), 18985-19003.
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, 33, 1877-1901.
- Castleman, K. R. (1996). *Digital image processing*.
- Cha, Y. J., Choi, W., & Büyüköztürk, O. (2017). Deep learning- based crack damage detection using convolutional neural networks. *Computer- Aided Civil and Infrastructure Engineering*, 32(5), 361-378.
- Chatar, K., Kitamura, K., & Cho, M. (2024). Onboard data prioritization using multi-class image segmentation for nanosatellites. *Remote Sensing*, 16(10), 1729.
- Chaurasia, A., & Culurciello, E. (2017). Linknet: Exploiting encoder representations for efficient semantic segmentation. In *2017 IEEE visual communications and image processing (VCIP)* (pp. 1-4). IEEE.
- Chen, K., Reichard, G., Xu, X., & Akanmu, A. (2021). Automated crack segmentation in close-range building façade inspection images using deep learning techniques. *Journal of Building Engineering*, 43, 102913.
- Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 801-818).
- Dabanli, Ö. (2008). *Tarihi yapıların deprem performansının belirlenmesi* (Master's thesis, Fen Bilimleri Enstitüsü).
- Dang, L. M., Wang, H., Li, Y., Nguyen, L. Q., Nguyen, T. N., Song, H. K., & Moon, H. (2022). Deep learning-based masonry crack segmentation and real-life crack length measurement. *Construction and Building Materials*, 359, 129438.
- De Prado, M. M. L. (2020). *Machine learning for asset managers*. Cambridge University Press.
- Dong, K., Zhou, C., Ruan, Y., & Li, Y. (2020, December). MobileNetV2 model for image classification. In *2020 2nd International Conference on Information Technology and Computer Application (ITCA)* (pp. 476-480). IEEE.
- Dorafshan, S., Thomas, R. J., & Maguire, M. (2018). Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete. *Construction and Building Materials*, 186, 1031-1045.

- Dorafshan, S., Thomas, R. J., & Maguire, M. (2018). SDNET2018: An annotated image dataset for non-contact concrete crack detection using deep convolutional neural networks. *Data in brief*, 21, 1664-1668.
- 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*.
- Dung, C. V. (2019). Autonomous concrete crack detection using deep fully convolutional neural network. *Automation in Construction*, 99, 52-58.
- Durmuş, M., Ergen, B., Çelebi, A., & Türkoğlu, M. (2024). ResNet Tabanlı PSPNet Kullanarak Panoramik Görüntülerde Gömülü Diş Segmentasyon Analizi. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 36(1), 159-166.
- Esteva, A., Chou, K., Yeung, S., Naik, N., Madani, A., Mottaghi, A., ... & Socher, R. (2021). Deep learning-enabled medical computer vision. *NPJ digital medicine*, 4(1), 5.
- Falk, T., Mai, D., Bensch, R., Çiçek, Ö., Abdulkadir, A., Marrakchi, Y., ... & Ronneberger, O. (2019). U-Net: deep learning for cell counting, detection, and morphometry. *Nature methods*, 16(1), 67-70.
- Fan, Z., Wu, Y., Lu, J., & Li, W. (2018). Automatic pavement crack detection based on structured prediction with the convolutional neural network. *arXiv preprint arXiv:1802.02208*.
- Floridi, L., & Cowls, J. (2022). A unified framework of five principles for AI in society. *Machine learning and the city: Applications in architecture and urban design*, 535-545.
- Gonzalez, R. C., & Woods, R. E. (2018). *Digital Image Processing* (4th ed.). Pearson.
- Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1, No. 2). Cambridge: MIT press.
- Grigorescu, S., Trasnea, B., Cocias, T., & Macesanu, G. (2020). A survey of deep learning techniques for autonomous driving. *Journal of field robotics*, 37(3), 362-386.
- Hacıfendioğlu, K., Altunışık, A. C., & Abdioğlu, T. (2023). Deep learning-based automated detection of cracks in historical masonry structures. *Buildings*, 13(12), 3113.
- Hatir, M. E., & Ince, I. (2021). Lithology mapping of stone heritage via state-of-the-art computer vision. *Journal of Building Engineering*, 34, 101921.
- Haykin, S. (2009). *Neural networks and learning machines*, 3/E. Pearson Education India.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).
- Hinton, G., Vinyals, O., & Dean, J. (2015). Distilling the knowledge in a neural network. *arXiv preprint arXiv:1503.02531*.
- Hisaria, S., Sharma, P., Gupta, R., & Sumalatha, K. (2024). An Analysis of Multi-Criteria Performance in Deep Learning-Based Medical Image Classification: A comprehensive review.
- Hossin, M., & Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. *International journal of data mining & knowledge management process*, 5(2), 1.

- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., ... & Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.
- Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).
- Huang, G., Sun, Y., Liu, Z., Sedra, D., & Weinberger, K. Q. (2016). Deep networks with stochastic depth. In *European conference on computer vision* (pp. 646-661). Cham: Springer International Publishing.
- Ibrahim, Y., Nagy, B., & Benedek, C. (2019). CNN-based watershed marker extraction for brick segmentation in masonry walls. In *International Conference on Image Analysis and Recognition* (pp. 332-344). Cham: Springer International Publishing.
- Ibrahim, Y., Nagy, B., & Benedek, C. (2020). Deep learning-based masonry wall image analysis. *Remote Sensing*, 12(23), 3918.
- ICOMOS. (2013). Türkiye Mimari Mirası Koruma Bildirgesi. http://www.icomos.org.tr/Dosyalar/ICOMOSTR_tr0784192001542192602.pdf (22.08.2025).
- Idjaton, K., Desquesnes, X., Treuillet, S., & Brunetaud, X. (2021). Stone-by-stone segmentation for monitoring large historical monuments using deep neural networks. In *International Conference on Pattern Recognition* (pp. 235-248). Cham: Springer International Publishing.
- Irmak, M. C., Aydin, T., & Yağanoğlu, M. (2022). Monkeypox skin lesion detection with MobileNetV2 and VGGNet models. In *2022 medical technologies congress (TIPTEKNO)* (pp. 1-4). IEEE.
- Jadon, S. (2020). A survey of loss functions for semantic segmentation. In *2020 IEEE conference on computational intelligence in bioinformatics and computational biology (CIBCB)* (pp. 1-7). IEEE.
- Jain, A. K. (1989). *Fundamentals of digital image processing*. Prentice-Hall, Inc.
- Janai, J., Güney, F., Behl, A., & Geiger, A. (2020). Computer vision for autonomous vehicles: Problems, datasets and state of the art. *Foundations and trends® in computer graphics and vision*, 12(1-3), 1-308.
- Karakoç, M. (2012). Görüntü işleme, teknolojiler ve uygulamaları. *Akademik Bilişim*, 1-3.
- Karimi, N., Mishra, M., & Lourenço, P. B. (2024). Deep learning-based automated tile defect detection system for Portuguese cultural heritage buildings. *Journal of Cultural Heritage*, 68, 86-98.
- Kassotakis, N., & Sarhosis, V. (2021, August). Employing non-contact sensing techniques for improving efficiency and automation in numerical modelling of existing masonry structures: A critical literature review. In *Structures* (Vol. 32, pp. 1777-1797). Elsevier.
- Keskar, N. S., Mudigere, D., Nocedal, J., Smelyanskiy, M., & Tang, P. T. P. (2016). On large-batch training for deep learning: Generalization gap and sharp minima. *arXiv preprint arXiv:1609.04836*.
- Keskenler, Mustafa Furkan, Deniz Dal, and Tolga Aydin. "Yapay zeka destekli ÇOKS yöntemi ile kredi kartı sahtekarlığının tespiti." *El-Cezeri* 8.2 (2021): 1007-1023.

- Khasoggi, B., Ermatita, E., & Sahmin, S. (2019). Efficient mobilenet architecture as image recognition on mobile and embedded devices. *Indonesian Journal of Electrical Engineering and Computer Science*, 16(1), 389-394.
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Kocaman, İ., Kazaz, İ., & Okuyucu, D. (2018). TARİHİ ERZURUM YAKUTIYE MEDRESESI'NİN YAPISAL DAVRANIŞININ İNCELENMESİ. *Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen Ve Mühendislik Dergisi*, 20(58), 36-51.
- Kocaman, İ. (2023). The effect of the Kahramanmaraş earthquakes (Mw 7.7 and Mw 7.6) on historical masonry mosques and minarets. *Engineering Failure Analysis*, 149, 107225.
- Kocaman, İ., Mercimek, Ö., Gürbüz, M., Erbaş, Y., & Anıl, Ö. (2024). The effect of Kahramanmaraş earthquakes on historical Malatya Yeni Mosque. *Engineering Failure Analysis*, 161, 108310.
- Kornblith, S., Shlens, J., & Le, Q. V. (2019). Do better imagenet models transfer better?. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 2661-2671).
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.
- Kutlu, İ. (2024). Yığma Yapılarda Güçlendirme Tekniklerinin Koruma İlkeleri ile İlişkilendirilmesi ve Mardin Tarihi Dokusunda Değerlendirilmesi. *Gazi Üniversitesi Fen Bilimleri Dergisi Part C: Tasarım ve Teknoloji*, 12(4), 979-1005.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436-444.
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (2002). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.
- Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, 18(8), 2674.
- Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2117-2125).
- Liu, Z., Mao, H., Wu, C. Y., Feichtenhofer, C., Darrell, T., & Xie, S. (2022). A convnet for the 2020s. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 11976-11986).
- Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3431-3440).
- Loukadakis, M., Cano, J., & O'Boyle, M. (2018). Accelerating deep neural networks on low power heterogeneous architectures. In *11th International Workshop on Programmability and Architectures for Heterogeneous Multicores (MULTIPROG-2018)*.
- Loverdos, D., & Sarhosis, V. (2022). Automatic image-based brick segmentation and crack detection of masonry walls using machine learning. *Automation in Construction*, 140, 104389.
- Lucho, S., Treuillet, S., Desquesnes, X., Leconge, R., & Brunetaud, X. (2024). Weakly Supervised SVM-Enhanced SAM Pipeline for Stone-by-Stone Segmentation of the Masonry of the Loire Valley Castles. *Journal of Imaging*, 10(6), 148.

- Maharana, K., Mondal, S., & Nemade, B. (2022). A review: Data pre-processing and data augmentation techniques. *Global Transitions Proceedings*, 3(1), 91-99.
- Masters, D., & Luschi, C. (2018). Revisiting small batch training for deep neural networks. *arXiv preprint arXiv:1804.07612*.
- McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. *AI magazine*, 27(4), 12-12.
- McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5, 115-133.
- McKinney, S. M., Sieniek, M., Godbole, V., Godwin, J., Antropova, N., Ashrafian, H., ... & Shetty, S. (2020). International evaluation of an AI system for breast cancer screening. *Nature*, 577(7788), 89-94.
- Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., & Terzopoulos, D. (2021). Image segmentation using deep learning: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 44(7), 3523-3542.
- Mitchell, T. M., & Mitchell, T. M. (1997). *Machine learning* (Vol. 1, No. 9). New York: McGraw-hill.
- Mousavi, M., & Bakhshi, A. (2022). Crack detection in masonry structures using computer vision based on deep learning.
- Nilsson, N. J. (2009). *The quest for artificial intelligence*. Cambridge University Press.
- Nixon, M., & Aguado, A. (2019). *Feature extraction and image processing for computer vision*. Academic press.
- Nusari, A. N., Alawi, A. E. B., Bozkurt, F., & Özbek, İ. Y. (2024). Comparison of YOLO Algorithms for Vehicle Accident Detection and Classification. In *2024 4th International Conference on Emerging Smart Technologies and Applications (eSmarTA)* (pp. 1-8). IEEE.
- Örs, M. E., & Özçelik, Z. (2025). Görüntü İşleme ve Derin Öğrenme ile Kaynak Hatalarının Gerçek Zamanlı Tespiti. *DCE Doğa Bilimleri Dergisi*, 5(2), 83-97.
- Pan, S. J., & Yang, Q. (2009). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10), 1345-1359.
- Pham, D. L., Xu, C., & Prince, J. L. (2000). Current methods in medical image segmentation. *Annual review of biomedical engineering*, 2(1), 315-337.
- Powers, D. M. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061*.
- Prechelt, L. (2002). Early stopping-but when?. In *Neural Networks: Tricks of the trade* (pp. 55-69). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Ravichand, M., Kumar, R., Hazela, B., & Suthar, T. (2022, December). Crack on brick wall detection by computer vision using machine learning. In *2022 6th International Conference on Electronics, Communication and Aerospace Technology* (pp. 1017-1020). IEEE.
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III* 18 (pp. 234-241). Springer international publishing.

- Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6), 386.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *nature*, 323(6088), 533-536.
- Russ, J. C. (2006). *The image processing handbook*. CRC press.
- Russell, S. J., & Norvig, P. (2016). *Artificial intelligence: a modern approach*. pearson.
- Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, 3(3), 210-229.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4510-4520).
- Seferbekov, S., Iglovikov, V., Buslaev, A., & Shvets, A. (2018). Feature pyramid network for multi-class land segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (pp. 272-275).
- Sezgin, M., & Sankur, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13(1), 146-165.
- Shamsabadi, E. A., Xu, C., & Dias-da-Costa, D. (2022). Robust crack detection in masonry structures with Transformers. *Measurement*, 200, 111590.
- 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.
- Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of big data*, 6(1), 1-48.
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Sithole, G. (2008). Detection of bricks in a masonry wall. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37, 567-72.
- Skansi, S. (2018). *Introduction to Deep Learning: from logical calculus to artificial intelligence*. Springer.
- Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information processing & management*, 45(4), 427-437.
- Sonka, M., Hlavac, V., & Boyle, R. (2013). *Image processing, analysis and machine vision*. Springer.
- Stockman, G., & Shapiro, L. G. (2001). *Computer vision*. Prentice Hall PTR.
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1, No. 1, pp. 9-11). Cambridge: MIT press.
- Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. (2017, February). Inception-v4, inception-resnet and the impact of residual connections on learning. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 31, No. 1).
- Szeliski, R. (2022). *Computer vision: algorithms and applications*. Springer Nature.
- Taha, A. A., & Hanbury, A. (2015). Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC medical imaging*, 15(1), 29.

- Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., Hurst, R. T., Kendall, C. B., Gotway, M. B., & Liang, J. (2016). Convolutional neural networks for medical image analysis: Full training or fine tuning?. *IEEE transactions on medical imaging*, 35(5), 1299-1312.
- Tan, M., & Le, Q. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning* (pp. 6105-6114). PMLR.
- Tekin, G. (2021). Tarihi Kent Merkezlerinde Yeniden Canlandırmanın Kültürel Boyutu. *Milli Folklor*, 17(130), 84-95.
- Tercan, B. (2018). Koruma Politikaları: Tarihi, Kültür ve Doğa Varlıklarının Afetlere Karşı Korunması. *Yüzüncü Yıl Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (40), 299-318.
- Tomazevic, M. (1999). *Earthquake-resistant design of masonry buildings* (Vol. 1). World Scientific.
- Turing, A. M. (2009). *Computing machinery and intelligence* (pp. 23-65). Springer Netherlands.
- Valero, E., Bosché, F., & Forster, A. (2018). Automatic segmentation of 3D point clouds of rubble masonry walls, and its application to building surveying, repair and maintenance. *Automation in Construction*, 96, 29-39.
- Valero, E., Bosché, F., & Forster, A. (2018). Automatic segmentation of 3D point clouds of rubble masonry walls, and its application to building surveying, repair and maintenance. *Automation in Construction*, 96, 29-39.
- Valero, E., Forster, A., Bosché, F., Hyslop, E., Wilson, L., & Turmel, A. (2019). Automated defect detection and classification in ashlar masonry walls using machine learning. *Automation in construction*, 106, 102846.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- Vincent, L., & Soille, P. (1991). Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 13(06), 583-598.
- Wang, C., Antos, S. E., & Triveno, L. M. (2021). Automatic detection of unreinforced masonry buildings from street view images using deep learning-based image segmentation. *Automation in Construction*, 132, 103968.
- Wang, P., Chen, P., Yuan, Y., Liu, D., Huang, Z., Hou, X., & Cottrell, G. (2018). Understanding convolution for semantic segmentation. In *2018 IEEE winter conference on applications of computer vision (WACV)* (pp. 1451-1460). Ieee.
- Wang, Y., Chen, Q., Hong, T., & Kang, C. (2018). Review of smart meter data analytics: Applications, methodologies, and challenges. *IEEE Transactions on smart Grid*, 10(3), 3125-3148.
- Wang, Z., Chen, J., & Hoi, S. C. (2020). Deep learning for image super-resolution: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 43(10), 3365-3387.
- Wang, Z., Wang, E., & Zhu, Y. (2020). Image segmentation evaluation: a survey of methods. *Artificial Intelligence Review*, 53(8), 5637-5674.
- Yamins, D. L., & DiCarlo, J. J. (2016). Using goal-driven deep learning models to understand sensory cortex. *Nature neuroscience*, 19(3), 356-365.

- Yang, C., & Guo, H. (2022). [Retracted] A Method of Image Semantic Segmentation Based on PSPNet. *Mathematical problems in engineering*, 2022(1), 8958154.
- Yang, X., Li, H., Yu, Y., Luo, X., Huang, T., & Yang, X. (2018). Automatic pixel- level crack detection and measurement using fully convolutional network. *Computer- Aided Civil and Infrastructure Engineering*, 33(12), 1090-1109.
- Yue, Y., Donà, M., Liu, X., Secco, M., Liu, H., & da Porto, F. (2025). Automatic Detection of Damage in Renaissance Walls Using a Deep Learning Model. *International Journal of Architectural Heritage*, 1-15.
- Zaman, M., Sevindi, C., & Birinci, S. (2018). Tarihi yolların bulunduğu Erzurum şehrindeki beşeri turistik eserler. *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 581-601.
- Zhang, L., Yang, F., Zhang, Y. D., & Zhu, Y. J. (2016). Road crack detection using deep convolutional neural network. In *2016 IEEE international conference on image processing (ICIP)* (pp. 3708-3712). IEEE.
- Zhang, L., Zhang, L., & Du, B. (2016). Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geoscience and remote sensing magazine*, 4(2), 22-40.
- Zhang, Z., Liu, Q., & Wang, Y. (2018). Road extraction by deep residual u-net. *IEEE Geoscience and Remote Sensing Letters*, 15(5), 749-753.
- Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). Pyramid scene parsing network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2881-2890).
- Zhao, X., Wang, L., Zhang, Y., Han, X., Deveci, M., & Parmar, M. (2024). A review of convolutional neural networks in computer vision. *Artificial Intelligence Review*, 57(4), 99.
- Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., & Liang, J. (2018). Unet++: A nested u-net architecture for medical image segmentation. In *International workshop on deep learning in medical image analysis* (pp. 3-11). Cham: Springer International Publishing.

ÖZGEÇMİŞ

Kişisel Bilgiler	
Adı Soyadı:	Rüstem Muhammed KARADEMİR
Doğum tarihi:	
Doğum Yeri:	
Uyruğu:	
Adres:	
Tel:	
E-posta:	
Eğitim	
Lise:	Samsun Anadolu Lisesi
Lisans:	Kırıkkale Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü
Yüksek lisans:	Atatürk Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı
Yabancı Dil Bilgisi	
İngilizce:	İyi
Üye Olunan Mesleki Kuruluşlar	
Tezden Üretilmiş Yayınlar	
1. Karademir, R. M., Aydın, T., 2025. Masonry Stone Wall Segmentation Using Image Processing Techniques and Deep Learning Methods. 22th International Istanbul Scientific Research Congress on Life, Engineering, Architecture and Mathematical Sciences, İstanbul.	