

KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

HARİTA MÜHENDİSLİĞİ ANABİLİM DALI

UZAKTAN ALGILANMIŞ GÖRÜNTÜLERDEN YENİ BİR DERİN
ÖĞRENME MİMARİSİ İLE BİNA ÇIKARIMI

YÜKSEK LİSANS TEZİ

Feride Seçil YILDIRIM

OCAK 2024

TRABZON



KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

HARİTA MÜHENDİSLİĞİ ANABİLİM DALI

UZAKTAN ALGILANMIŞ GÖRÜNTÜLERDEN YENİ BİR DERİN
ÖĞRENME MİMARİSİ İLE BİNA ÇIKARIMI

Feride Seçil YILDIRIM

Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsünde

“HARİTA YÜKSEK MÜHENDİSİ”

Unvanı Verilmesi İçin Kabul Edilen Tezdir.

Tezin Enstitüye Verildiği Tarih : 15 / 12 / 2023

Tezin Savunma Tarihi : 08 / 01 / 2024

Tez Danışmanı : Prof. Dr. Fevzi KARSLI

Trabzon 2024

ÖNSÖZ

“Uzaktan Algılanmış Görüntülerden Yeni Bir Derin Öğrenme Mimarisi İle Bina Çıkarımı” isimli bu tez Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü Harita Mühendisliği Anabilim Dalı, Yüksek Lisans Programı’nda hazırlanmıştır.

Tez çalışmam boyunca destek ve emeklerini esirgemeyen, cesaret veren, tecrübesi ile yoluma ışık tutan değerli danışman hocam Prof. Dr. Sayın Fevzi KARSLI’ya teşekkürlerimi sunarım. Ayrıca değerli jüri üyelerim Prof. Dr. Sayın Taşkın KAVZOĞULU ve Doç. Dr. Mustafa DİHKAN’a fikirleri ve yorumları için teşekkür ederim.

Fotogrametri ve Uzaktan Algılama Laboratuvarındaki çalışma arkadaşlarım Harita Yüksek Mühendisi Sayın Merve YILDIRIM, Harita Yüksek Mühendisi Sayın Murat BAHADIR, Harita Mühendisi Sayın Batuhan GÜMRÜKÇÜ ve Harita Yüksek Mühendisi Sayın Ahmad Shakib SAHAK’a yapıcı eleştirileri, önerileri ve teknik destekleri ile tezime yaptıkları büyük katkılarından ötürü teşekkürlerimi sunarım.

Son olarak, her kararında yanımda olan, daimî destekleri ve fedakârlıkları ile bugüne ulaşabilmemde en büyük paya sahip olan annem Sema YILDIRIM, babam Cemal YILDIRIM ve babaannem Feride YILDIRIM’a teşekkür eder ve minnettarlığımı sunarım.

Hazırladığım bu tez çalışmasının bundan sonra yapılacak bilimsel çalışmalara katkı sağlamasını ümit ve temenni ederim.

Feride Seçil YILDIRIM

Trabzon 2024

TEZ ETİK BEYANNAMESİ

Yüksek Lisans Tezi olarak sunduğum “Uzaktan Algılanmış Görüntülerden Yeni Bir Derin Öğrenme Mimarisi İle Bina Çıkarımı” başlıklı bu çalışmayı baştan sona kadar danışmanım Prof. Dr. Fevzi KARSLI’nın sorumluluğunda tamamladığımı, verileri/örnekleri kendim topladığımı, deneyleri/analizleri ilgili laboratuvarlarda yaptığımı/yaptırdığımı, başka kaynaklardan aldığım bilgileri metinde ve kaynakçada eksiksiz olarak gösterdiğimi, çalışma sürecinde bilimsel araştırma ve etik kurallara uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul ettiğimi beyan ederim. 08/01/2024

Feride Seçil YILDIRIM

İÇİNDEKİLER

	<u>Sayfa No</u>
ÖNSÖZ	III
TEZ ETİK BEYANNAMESİ.....	IV
İÇİNDEKİLER.....	V
ÖZET	VII
SUMMARY	VIII
ŞEKİLLER DİZİNİ.....	IX
TABLolar DİZİNİ.....	XI
SEMBOLLER VE KISALTMALAR DİZİNİ	XII
1. GENEL BİLGİLER	1
1.1. Giriş.....	1
1.2. Çalışmanın Motivasyonu.....	2
1.3. Tezin Ana Hattı	2
1.4. Çalışmanın Amacı	3
1.5. Literatür Araştırması	4
1.6. Temel Kavramlar.....	10
1.6.1. Fotogrametri.....	10
1.6.2. Hava Fotoğrafı Üretim Teknikleri	13
1.6.3. Dijital Hava Kameraları ve Özellikleri	14
1.6.4. Uzaktan Algılanmış Verilerden Bina Çıkarımı.....	16
1.6.5. Matematiksel Morfolojik Operasyonlar.....	18
1.6.6. Görüntü Segmentasyonu ve Segmentasyon Türleri.....	21
1.6.7. Yapay Zekâ	23
1.6.8. Makine Öğrenimi	25
1.6.9. Yapay Sinir Ağları	26
1.6.10. Derin Öğrenme (Derin Yapay Sinir Ağları)	29
1.6.11. Evrişimli (Evrişimsel) Sinir Ağları	30
1.6.12. U-Net Mimarisi	31
1.6.13. Pürüzsüz Karıştırma (Smooth Blending) Yöntemi	34
1.6.14. Doğruluk Değerlendirmesi.....	36
1.7. Derin Öğrenmenin Temel Kavramları	38

1.7.1. Evrişim Katmanı	38
1.7.2. Havuzlama Katmanı.....	39
1.7.3. Seyreltme Katmanı.....	40
1.7.4. Toplu Normalleştirme Katmanı	41
1.7.5. Birleştirme Katmanı	43
1.7.6. Aktivasyon Fonksiyonları	43
1.7.7. Sınıflandırıcı Katman	45
1.7.8. Maliyet/Kayıp Fonksiyonu	46
1.7.9. Öğrenme Oranı.....	46
1.7.10. Optimizasyon Algoritmaları	47
2. YAPILAN ÇALIŞMALAR.....	50
2.1. Çalışma Alanı ve Veri Seti.....	52
2.2. Veri Seti Ön Hazırlık İşlemleri	53
2.2.1. Maske Veri Setindeki Binalara Sınır Çizdirilmesi.....	53
2.3. Eğitim, Doğrulama ve Test Veri Setinin Oluşturulması	54
2.4. Yeni Bir Derin Öğrenme Mimarisi Tasarımı ve Eğitimi	56
2.4.1. Tasarlanan Yeni Mimari ile İki Sınıflı Semantik Segmentasyon	60
2.4.2. Tasarlanan Yeni Mimari ile Üç Sınıflı Semantik Segmentasyon	62
2.4.3. U-Net Mimarisi ile İki Sınıflı Semantik Segmentasyon	64
2.4.4. U-Net Mimarisi ile Üç Sınıflı Semantik Segmentasyon	65
2.5. Eğitilen Mimarilerin Test Edilmesi	68
3. BULGULAR VE TARTIŞMA	70
3.1. İki Sınıflı ve Üç Sınıflı Segmentasyon ile Bina Çıkarım Sonuçlarının Değerlendirilmesi	70
3.2. Önerilen Mimari ve U-Net Mimari Sonuçlarının Karşılaştırılması	72
3.3. Test Görüntüleri İçin Doğruluk Analizi Sonuçlarının Değerlendirilmesi.....	76
3.4. Test Görüntüleri Üzerindeki İki ve Üç Sınıflı Segmentasyon Doğruluk Değerlendirme Sonuçlarının Karşılaştırılması	80
4. SONUÇLAR.....	84
5. ÖNERİLER.....	86
6. KAYNAKLAR	88
ÖZGEÇMİŞ	

ÖZET

Yüksek Lisans Tezi

UZAKTAN ALGILANMIŞ GÖRÜNTÜLERDEN YENİ BİR DERİN ÖĞRENME MİMARİSİ İLE BİNA ÇIKARIMI

Feride Seçil YILDIRIM

Karadeniz Teknik Üniversitesi
Fen Bilimleri Enstitüsü
Harita Mühendisliği Anabilim Dalı
Danışman: Prof. Dr. Fevzi Karşlı
2024, 95 Sayfa

Haritanın temel unsurlarından biri olan bina katmanı, çeşitli nizam ve mimarilere sahip olup; kentsel planlama, emlak yönetimi, kaçak yapı tespiti ve afet analizi gibi birçok haritacılık faaliyetinde kullanılmaktadır. Bu sebeple bina obje bilgilerine hızlı ve otomatize bir biçimde erişmek önem arz etmekle birlikte son derece popüler bir konu haline gelmiştir. Bu tez çalışmasında, Massachusetts bina veri seti kullanılarak yüksek çözünürlüklü dijital hava fotoğrafları üzerinden otomatik bir biçimde bina çıkarımı hedeflenmiş ve bu amaç için yeni bir derin öğrenme mimarisi tasarlanmıştır. Bunun yanı sıra veri setindeki maskelerde bulunan bina temsillerine morfolojik operasyonların uygulanmasıyla bina sınır bilgisi eklenerek, derin öğrenme ağının bina objesini öğrenmesi hususunda, bina sınır bilgisinin katkısı araştırılmıştır. Tasarlanan mimarinin test verileri üzerindeki performansı; Genel Doğruluk (OA), Doğruluk (Cr), Tamlık (Cp), F1-Puanı, Birleşimlerin Kesişimi (IoU) doğruluk metrikleri için sırasıyla %91, %80, %73, %76 ve %62 olarak hesaplanmıştır. Önerilen mimarinin karşılaştırılması için aynı veri seti ile eğitilip, aynı test verileri üzerinde değerlendirilen, literatürde mevcut U-Net mimarisine göre; Genel Doğruluk metriği için eşit, diğer metrikler için ise ortalama %2'lik daha düşük performans sergilediği gözlemlenmiştir. Bu duruma karşın tasarlanan mimarinin eğitim işlemini, U-Net mimarisine göre 400 dakika daha kısa sürede tamamlaması önemli bir üstünlük sağlamaktadır. Sonuç olarak tasarlanan mimarinin, literatürde başarısı kabul görmüş olan U-Net mimari performansına yaklaşmış olması ve eğitim süresindeki avantajı sebebiyle kullanılabilirlik açısından hız ve kolaylık sağlayacağı öngörülmektedir.

Anahtar Kelimeler: Derin Öğrenme, Evrişimli Sinir Ağları, Bina Çıkarımı, U-Net Mimarisi, Massachusetts Bina Veri Seti

SUMMARY

MSc. Thesis

BUILDING EXTRACTION FROM REMOTELY SENSED IMAGES WITH A NOVEL DEEP LEARNING ARCHITECTURE

Feride Seçil YILDIRIM

Karadeniz Technical University
The Graduate School of Natural and Applied Sciences
Geomatics Engineering Graduate Program
Supervisor: Prof. Fevzi KARSLI
2024, 95 Pages

The building layer, which is one of the basic elements of the map, has various architectures and is used in many mapping activities such as urban planning, real estate management, illegal building detection and disaster analysis. For this reason, it is important to access building object information quickly and automatically. In this thesis study, automatic building extraction from high-resolution digital aerial images was aimed using the Massachusetts building dataset and a new deep learning architecture was designed for this purpose. In addition, building boundary information was added to the building representations in the masks in the data set by applying morphological operations, and the contribution of the building boundary information to the deep learning network's learning of the building object was investigated. The performance of the designed architecture on test data; for Overall Accuracy, Correctness, Completeness, F1-Score, Intersection over Union accuracy metrics was calculated as 91%, 80%, 73%, 76% and 62%, respectively. It has been observed that the proposed architecture has equal performance for the Overall Accuracy metric and an average of 2% lower performance for other metrics than the existing U-Net architecture in the literature. Despite this situation, the designed architecture has a significant advantage in completing the training process in 400 minutes less than the U-Net architecture. As a result, it is estimated that the designed architecture will approach the U-Net architectural performance, which is accepted in the literature, and will provide speed and convenience in terms of usability due to its advantage in training time.

Key Words: Deep Learning, Convolutional Neural Networks, Building Extraction, U-Net Architecture, Massachusetts Building Dataset

ŞEKİLLER DİZİNİ

	<u>Sayfa No</u>
Şekil 1. Fotogrametrik veri alım geometrisi.....	12
Şekil 2. Analog ve dijital kameralar ile elde edilen ortofotolar (a) Analog, (b) Dijital	14
Şekil 3. Matematiksel morfolojik operasyonlar için örnek yapı elemanları.....	19
Şekil 4. Aşınma işlemi.....	20
Şekil 5. Genleşme işlemi	20
Şekil 6. Görüntülerden nesne çıkarımı (a) Obje tespiti, (b) Segmentasyon	22
Şekil 7. Segmentasyon türleri.....	23
Şekil 8. Yapay zekâ alt dalları	24
Şekil 9. Makine öğreniminin geleneksel yaklaşıma göre tanımı.....	25
Şekil 10. Makine öğrenimi ve derin öğrenme farkı.....	26
Şekil 11. Yapay sinir ağlarının bileşenleri ve çalışma prensibi.....	28
Şekil 12. Basit yapay sinir ağı ve derin yapay sinir ağı karşılaştırma	29
Şekil 13. Evrişimli sinir ağı ile görüntü özelliklerinin çıkarılması.....	31
Şekil 14. U-Net mimarisi.....	33
Şekil 15. Spline eğrileri için süreklilik örneği.....	35
Şekil 16. Pürüzsüz karıştırma yöntemi (a) Görüntüde süreksizlik örneği, (b) Pürüzsüz karıştırma yöntemi ile görüntüdeki süreksizliğin ortadan kaldırılması	36
Şekil 17. DP, YP, YN ve DN metriklerinin belirlenmesi (a) Referans Pikseller, (b) Tahmin Edilen Pikseller.....	37
Şekil 18. Evrişim katmanının görüntüler üzerindeki işlemi	39
Şekil 19. Ortalama ve maksimum havuzlama katmanları	40
Şekil 20. Seyreletme katmanı işlevi	41
Şekil 21. Doğrusal regresyon	44
Şekil 22. Aktivasyon fonksiyonları	45
Şekil 23. Öğrenme oranının model üzerindeki etkisi	47
Şekil 24. Yerel ve küresel minimum	48
Şekil 25. Optimizasyon algoritmalarının performans karşılaştırması	49
Şekil 26. Genel iş akış şeması	51
Şekil 27. Çalışma alanı	52

Şekil 28.	Massachusetts bina veri seti örneği (a) Hava görüntüsü, (b) Bina maske	53
Şekil 29.	Morfolojik operasyonlar ile bina sınırlarının oluşturulması	54
Şekil 30.	Bindirme oranı ile tüm binaların algılanması.....	55
Şekil 31.	Girdi verisi örnekleri (a) İki sınıflı segmentasyon için görüntü ve maske, b) Üç sınıflı segmentasyon için görüntü ve maske	57
Şekil 32.	Tasarlanan mimarinin detayları (a) Tasarlanan mimarinin görsel temsili, (b) Tasarlanan mimari katmanlarının detayları	58
Şekil 33.	Tasarlanan mimarinin iki sınıflı semantik segmentasyon için eğitim doğrulama grafikleri (a) Doğruluk grafiği, (b) Kayıp grafiği.....	61
Şekil 34.	Tasarlanan mimarinin iki sınıflı semantik segmentasyon örnek tahmin sonuçları.....	61
Şekil 35.	Tasarlanan mimarinin üç sınıflı semantik segmentasyon için eğitim-doğrulama grafikleri (a) Doğruluk grafiği, (b) Kayıp grafiği.....	63
Şekil 36.	Tasarlanan mimarinin üç sınıflı semantik segmentasyon örnek tahmin sonuçları.....	63
Şekil 37.	U-Net mimarisinin iki sınıflı semantik segmentasyon için eğitim doğrulama grafikleri (a) Doğruluk grafiği, (b) Kayıp grafiği.....	64
Şekil 38.	U-Net mimarisinin iki sınıflı semantik segmentasyon örnek tahmin sonuçları.....	65
Şekil 39.	U-Net mimarisinin üç sınıflı semantik segmentasyon için Eğitim-Doğrulama grafikleri (a) Doğruluk grafiği, (b) Kayıp grafiği.....	66
Şekil 40.	U-Net mimarisinin üç sınıflı semantik segmentasyon örnek tahmin sonuçları.....	67
Şekil 41.	Pürüzsüz karıştırmasız ve pürüzsüz karıştırma yöntemi ile yapılan tahmin sonuçlarının karşılaştırılması.....	69
Şekil 42.	Doğrulama verileri üzerinde önerilen mimari ile yapılan iki ve üç sınıflı segmentasyon sonuçlarının karşılaştırılması	71
Şekil 43.	Tasarlanan mimari ve U-Net mimarisi.....	72
Şekil 44.	İki sınıflı segmentasyon için U-Net mimarisi ve önerilen mimarinin doğrulama verileri üzerindeki tahminleri	74
Şekil 45.	Üç sınıflı segmentasyon için U-Net mimarisi ve önerilen mimarinin doğrulama verileri üzerindeki tahminleri	75
Şekil 46.	İki sınıflı eğitilen U-Net mimarisi ve tasarlanan yeni mimarinin test görüntüsü üzerindeki sonuçları.....	78
Şekil 47.	Üç sınıflı eğitilen U-Net mimarisi ve tasarlanan yeni mimarinin test görüntüsü üzerindeki sonuçları.....	79
Şekil 48.	Tasarlanan mimarinin test görüntüsü üzerindeki iki ve üç sınıflı segmentasyonu.....	82

TABLolar DİZİNİ

	<u>Sayfa No</u>
Tablo 1. Örnek hava kameraları ve özellikleri	16
Tablo 2. Yapay sinir ağı elemanlarının biyolojik sinir sistemindeki karşılıkları	27
Tablo 3. Karışıklık matrisi.....	37
Tablo 4. Eğitim, doğrulama ve test veri seti miktarları	56
Tablo 5. Bilgisayar özellikleri ve kullanılan yazılımlar	70
Tablo 6. Tasarlanan mimari ve U-Net mimarisi teknik karşılaştırma	73
Tablo 7. U-Net mimarisi ve tasarlanan mimarinin iki sınıflı segmentasyon için doğruluk değerlendirme sonuçları	77
Tablo 8. U-Net mimarisi ve tasarlanan mimarinin üç sınıflı segmentasyon için doğruluk değerlendirme sonuçları	77
Tablo 9. Önerilen mimari ile test görüntüleri üzerindeki iki ve üç sınıflı segmentasyon sonuçlarının karşılaştırılması	81
Tablo 10. Test görüntüleri üzerinde üç sınıf ile eğitilmiş önerilen mimarinin, iki sınıf ile eğitilmiş mimariye göre doğruluk oranı değişimi	81

SEMBOLLER VE KISALTMALAR DİZİNİ

2D/B	: İki Boyutlu
2D/B	: Üç Boyutlu
CBS	: Coğrafi Bilgi Sistemleri
KBS	: Kent Bilgi Sistemi
GPS	: Global Konum Belirleme Sistemi (Global Positioning System)
INS	: Atalet Navigasyon Sistemi (Inertial Navigation System)
IMU	: Atalet Ölçü Birimi (Inertial Measurement Unit)
GSD/YÖA	: Yer Örnekleme Aralığı (Ground Sampling Distance)
CCD	: Yük Bağlamalı Düzen (Charged-Coupled Device)
CMOS	: Tamamlayıcı Yarı İletken Metal Oksit (Complementary Metal Oxide Semiconductor)
LiDAR	: Işık Tespiti ve Ölçümü (Light Detection and Ranging)
VDVI	: Görünür Işık Farkı Bitki Örtüsü İndeksi (Visible Light Difference Vegetation Index)
AI	: Yapay Zekâ (Artificial Intelligence)
CNN	: Evrişimli Sinir Ağı (Convolutional Neural Network)
RNN	: Tekrarlayan Sinir Ağı (Recurrent Neural Network)
RBM	: Kısıtlı Boltzmann Makineleri (Restricted Boltzmann Machines)
DBN	: Derin İnanç Ağları (Deep Belief Networks)
FCN	: Tamamen Evrişimli Ağ (Fully Convolutional Network)
SAM	: Her Şeyin Segmentasyonu (Segment Anything Model)
INRIA	: Ulusal Bilgisayar Bilimi ve Otomasyon Araştırma Enstitüsü (Institut National de Recherche en Informatique et en Automatique)
WV2	: WorldView-2
WV3	: WorldView-3

1. GENEL BİLGİLER

1.1. Giriş

1858’de Paris’in hemen dışındaki Petit-Bicêtre’nin (Petit-Clamart) yaklaşık 80 m (262 feet) yukarısından bir sıcak hava balonu ile ilk başarılı hava fotoğrafının çekilmesinden bu yana insanlığın yeryüzüne dair havadan bilgi edinme merakı artmıştır (Waxman, 2018). Teknolojinin gelişimiyle hava fotoğrafları ve uydu görüntülerinin konumsal ve spektral çözünürlükleri, kullanılan sensörler vasıtasıyla ileri düzeylere ulaşmıştır. Bununla birlikte havadan yeryüzü hakkında bilgi toplama işlemi günümüze kadar gelen süreçte daha fazla yaygınlaşarak, popülerite kazanmıştır.

Hava ve uydu görüntülerinin zengin spektral ve konumsal özellikleri sayesinde yeryüzünden; orman envanteri, maden tespiti, iklim değişikliği, hava/su kirliliği ve hassas tarım uygulamalarına yönelik çeşitli bilgiler sağlanabilmektedir. Tüm bunların yanı sıra yaklaşık son otuz yılda; kentsel planlama, emlak yönetimi, kaçak yapı tespiti, deprem analizi ve afet yardımı uygulamalarına olanak sağlamasıyla bilinen otomatik bina tespitine yönelik çok çeşitli yöntemler önerilmiştir (Samui vd., 2017).

Uzaktan algılanmış görüntülerden bina çıkarım yöntemleri klasik ve derin öğrenme tabanlı yöntemler olmak üzere ikiye ayrılmaktadır. Klasik yöntemde ilk olarak binaların geometrik, spektral ve gölge özellikleri çıkarılıp, ardından bu özellikler bina ve bina olmayan sınıfları ayırmak için tasarlanan sınıflandırıcıya gönderilmektedir (Zha vd., 2003; Liu vd., 2005; Awrangjeb vd., 2011; Huang & Zhang, 2011; Huang & Zhang, 2012). Klasik yöntemde büyük miktarda ön bilgiye ihtiyaç duyulur ve bu yöntemler yoğun iş gücü gerektirmesiyle birlikte otomasyonu düşüktür. Derin öğrenme yöntemleri ise orijinal görüntüde bulunan özellikleri örüntüler yardımıyla otomatik olarak öğrenip yorumlayabilmektedir. Bu durum, uzaktan algılanmış görüntülerden; karmaşık spektral, dokusal ve geometrik özelliklerin elde edilmesi yönünde büyük bir avantaj sağlamaktadır (Yuan vd., 2020).

Gerçekleştirilen tez çalışmasında derin öğrenme yöntemleri ile yüksek çözünürlüklü hava fotoğraflarından bina çıkarımına yönelik; Massachusetts bina veri seti (Mnih, 2013) kullanılarak, yeni bir derin öğrenme mimarisi tasarlanmıştır. Literatürden farklı olarak, özgün bir derin öğrenme mimarisi tasarımının yanında, Massachusetts bina veri setinde

bulunan bina maske görüntülerine, morfolojik operasyonların (genleşme, aşınma) (Gonzalez & Woods, 2004) uygulanmasıyla otomatik olarak bina sınırı katmanı eklenmiş ve ilave edilen bina sınır bilgisinin, bina çıkarımına yönelik katkısı araştırılmıştır.

Önerilen yeni mimari ile bina çıkarımına yönelik uygulama sonuçları literatürde kullanılabilirliği kabul görmüş bir derin öğrenme ağı olan U-Net mimarisi (Ronneberger vd., 2015) sonuçları ile karşılaştırılarak, tasarlanan mimarinin, günümüzde mevcut ve başarılı kanıtlanmış derin öğrenme ağlarına; doğruluk, katman sayısı ve eğitim süresi bakımından bir alternatif olabileceği ortaya konulmuştur.

1.2. Çalışmanın Motivasyonu

Çalışmanın başlıca motivasyonu; çeşitli görüntü işleme yöntemlerinin kullanıldığı, yüksek miktarda iş gücü gerektiren, otomasyonu düşük klasik yöntemler ile hava görüntülerinden bina çıkarım uygulamalarına karşın, derin öğrenme yöntemleri kullanılarak daha hızlı ve az iş gücü gereksinimi ile otomasyonu yüksek bir şekilde bina katmanı bilgisine erişim işlemi gerçekleştirme isteğidir. İlave olarak literatürde derin öğrenme tabanlı olarak gerçekleştirilen bina çıkarım tekniklerinden farklı olarak, sadece yapılan çalışmaya özgü yeni bir derin öğrenme mimarisi tasarlanmasıyla hava görüntülerinden yüksek doğruluklu bina çıkarım uygulamasının gerçekleştirilmesi amaçlanmıştır.

1.3. Tezin Ana Hattı

Genel bilgiler bölümünün giriş kısmında yapılan çalışma hakkında ön bilgiler verilmesinin ardından çalışmanın motivasyonundan bahsedilmiştir. Devamında çalışmanın genel amacına değinilerek, literatür araştırması bölümünde, mevcut çalışmanın konusu olan, uzaktan algılanmış görüntülerden bina çıkarımı hakkında yapılan klasik ve çoğunlukla derin öğrenme tabanlı yöntemlere dair araştırmalara yer verilmiştir.

Temel kavramlar bölümünde; fotogrametri bilim dalı, dijital hava kameraları ve özellikleri, hava fotoğrafı üretim teknikleri ve uzaktan algılanmış verilerden bina çıkarım yöntemlerinden kısaca bahsedilmiştir. Ardından çalışmanın içerisinde yer alan temel metotlar ve kavramlar olan; matematiksel morfolojik operasyonlar, görüntü segmentasyonu, yapay zekâ ve yapay zekânın alt dalları olan; makine öğrenimi, yapay sinir ağları, derin öğrenme ve evrişimli sinir ağları ifadelerinin hiyerarşik ve kronolojik sıraya uygun bir şekilde anlatımı sağlanmıştır. Devamında çalışmada bahsi geçen; U-Net mimarisi, Pürüzsüz

Karıştırma (Smooth Blending) yöntemi ve yapılan uygulamanın performansını değerlendirmek için kullanılan doğruluk metrikleri gibi diğer temel kavramlara yer verilmiştir. Bu bölümün sonrasında ise derin öğrenme yöntemlerine dayanan ve çalışma dahilinde kullanılan; yapay sinir ağı katmanları, fonksiyonları ve ifadeleri gibi önemli bilgilere yer verilmiştir. Mevcut tez çalışmasını oluşturan uygulamanın anlatıldığı yapılan çalışmalar başlığı altında; çalışma alanı ve kullanılan veri seti, veri setine uygulanan ön hazırlık işlemleri ve veri setinin; eğitim, doğrulama ve test olmak üzere üç kategoriye ayrılması, tasarlanan yeni mimarinin özellikleri ve bu mimari ile yapılan iki (bina, arka plan) ve üç (bina, arka plan, bina sınırı) sınıflı segmentasyon eğitimleri ve son olarak; önerilen mimarinin karşılaştırılması amacıyla kullanılacak olan U-Net mimarisinin de aynı sınıflar ve şartlar altında eğitilmesi işlemleri detaylı bir biçimde anlatılmıştır. Uygulamanın son aşamasında ise kullanılan mimarilerin eğitilmesi sonucu kaydedilen dört adet modelin değerlendirilebilmesi için yapılan test işleminden bahsedilmiştir. Bulgular ve irdeleme başlığı altında iki ve üç sınıflı segmentasyonun bina çıkarımı üzerindeki farklarının araştırılması için yapılan değerlendirmeler ve tasarlanan yeni mimarinin kalitesini ortaya koymak adına U-Net mimarisi ile karşılaştırma sonuçları sunulmuştur. Tez yazım işlemi; sonuçlar, öneriler, kaynaklar, son olarak özgeçmiş bölümleri ile sonlandırılmıştır.

1.4. Çalışmanın Amacı

Yapılan çalışmanın ana amacı; minimum iş gücü ve mümkün olabildiğince zaman tasarrufu sağlayarak uzaktan algılanmış görüntülerden bina çıkarımını derin öğrenme temelli metotlar ile gerçekleştirmektir. Bu amaç için literatürde mevcut, genel geçer ve başarısı kanıtlanmış, genellikle görüntü segmentasyonunda kullanılan derin öğrenme mimarilerine; katman sayısı, eğitim süresi ve doğruluk oranı bakımından alternatif olabilecek, görüntü segmentasyon görevini başarıyla gerçekleştirebilecek yeni bir derin öğrenme mimarisi tasarlanmıştır. Ek olarak, çalışmada derin öğrenme ağının eğitiminde kullanılan bina veri setindeki hava fotoğraflarının, siyah-beyaz (ikili/binary) formattaki maskelerinde bulunan bina temsillerine morfolojik operasyonlar ile iki piksel genişlikli bina sınırı çizdirilerek, eğitilecek mimariye arka plan ve bina sınıflarının yanı sıra bina sınır bilgisinin de ilave edilmesinin uzaktan algılanmış görüntülerden bina çıkarımı işlemine katkısının araştırılması da çalışmanın amaçları dahilindedir. Özetle literatüre, veri seti üzerine morfolojik operasyonlar yardımıyla görüntü ön işleme metotlarının uygulanması ve görüntü

segmentasyonu hususunda başarılı yeni bir derin öğrenme mimarisinin tasarımı ile hava fotoğraflarından bina çıkarımına yönelik alternatif yeni bir yöntem kazandırılması hedeflenmiştir.

1.5. Literatür Araştırması

Uzaktan algılanmış görüntülerin konumsal ve spektral çözünürlüklerinin gelişmesiyle birlikte yersel ölçümler yapılmaksızın hava fotoğraflarından binaların tespiti, bina sınır bilgilerinin elde edilmesi ve kontrolü; ekonomik, sosyal ve sürdürülebilir kalkınma açısından büyük önem arz etmektedir. Literatürde hava fotoğrafları ve uydu görüntüleri kullanılarak; kentsel planlama, emlak yönetimi, kaçak yapı tespiti, deprem analizi ve afet yardımı uygulamalarına olanak sağlamasıyla bilinen otomatik bina tespitine yönelik çok çeşitli yöntemler önerilmiştir (Samui vd., 2017).

Uzaktan algılanmış görüntülerden bina çıkarım yöntemleri literatürde klasik ve derin öğrenme tabanlı olmak üzere ikiye ayrılmaktadır. Klasik yöntemlerin uygulanması için büyük miktarda ön bilgiye ihtiyaç duyulur ve bu yöntemler zaman alıcı ve otomasyonu düşük olmakla birlikte genelleştirilebilir değildir. Klasik yöntemler ile binaların çıkarımı için, görüntülerin spektral ve geometrik koşullarının iyi tanınması ve kullanılmasıyla, ilk olarak bitki ve gölge özelliklerinin çıkarılması gerekmektedir. Ardından geriye kalan özellikler bina ve bina olmayan sınıfları ayırmak için geliştirilen sınıflandırıcıya gönderilmektedir (Zha vd., 2003; Liu vd., 2005; Awrangjeb vd., 2011; Huang & Zhang, 2011; Huang & Zhang, 2012).

Akbulut vd., (2018) yaptıkları çalışmada bina sınırlarının tespiti için başlangıç kontur konumlarının otomatik oluşturulmasını hedefleyerek bölge bazlı aktif kontur modelini tercih etmişlerdir. Bu işlem için morfolojik operasyonlar, bant oranı ve eşikleme yöntemlerini kullanmışlardır. Önerdikleri yaklaşımı konumsal çözünürlüğü 8 cm olan yüksek çözünürlüklü renkli kızılötesi hava görüntüleri üzerinde test etmişlerdir. Önerilen yöntem ile test görüntüleri üzerindeki binaların tamamı çıkarılmıştır. Ancak görüntülerde bulunan yollar, bitki örtüsü ve gölge efektlerinden ötürü tespit edilen bazı binaların sınırları pürüzlü hale gelmiş, bazı bina sınırları ise spektral olarak benzer özelliklere sahip bina dışı alanlara doğru genişlemiştir. Ayrıca yapılan bu çalışmada kontur başlangıç noktasının otomatik belirlenmesi hususunda, binalar için elde edilen sınırlayıcı kutu merkezinin bina dışına

gelebileceği örnekler üzerinde denemeler yapılmamış olması yöntemin genelleştirilebilirliğini kısıtlamaktadır.

Gavankar ve Ghosh (2018) tarafından klasik yöntemlerin kullanımıyla yapılan bina çıkarım uygulaması, morfolojik Top-hat filtresine ve K-ortalamlar (K-means) algoritmasına dayanmaktadır. Öncelikle görüntünün karanlık ve parlak bölgelerini birbirinden ayırmak için morfolojik işlemleri kullanan Top-hat filtresinden yararlanılmıştır. Ardından K-ortalamlar algoritması kullanılarak aydınlık, karanlık ve ara spektral özellikli nesneler için üç eşik değeri belirlenmiş ve nesnelerin ağırlık merkezleri tespit edilmiştir. Devamında medyan filtre kullanılarak görüntüdeki tuz-biber etkisi giderilmiş ve Bağlantılı Bileşen Etiketleme tekniği ile tespit edilen her nesneye bir etiket atanmıştır. Bağlantılı Bileşen tekniği ile etiketlenen nesneler arasında aday bina, yol, park yeri gibi bina dışı objeler bulunmaktadır. Bina dışı objelerin elimine edilmesi adına eksen oranı, dışmerkezlilik ve bölüm alanı gibi parametreler kullanılmıştır. Önerilen metodolojinin performansını, ABD'nin Santa Ana, California bölgesinden alınan 0.60 m çözünürlüğe sahip IKONOS pankromatik görüntüsü üzerinde test edilmiştir. Elde edilen sonucun başarılı olmasıyla birlikte, az sayıda da olsa gerçek binalar ile benzer yansıma değerleri ve yapısal özellikler gösteren bina dışı özelliklerin, bina sınıfında sınıflandırıldığı gözlemlenmiştir.

Chaudhuri vd. (2016), sadece yüksek çözünürlüklü pankromatik uydu görüntüleri için konumsal ve spektral bilgileri kullanan bir bina çıkarma stratejisi önermişlerdir. Yapılan çalışmada sırasıyla morfolojik görüntü iyileştirme (açma-kapama), Dahili Gri Varyansı (IGV) kullanan çok tohumlu kümeleme, gölge tespiti, bina kenarları ve gölgelerin konum bilgisini kullanarak yanlış bina tespit durumlarının azaltılması ve son olarak Uyarlanabilir Eşik Bazlı Segmentasyon Tekniği yöntemleri kullanılmıştır. Önerilen yöntemin, ön sınıflandırma veya herhangi bir eğitim seti gerektirmeden otomatik olarak çalışmakta olduğunu, fakat bazı başlangıç algoritma parametrelerinin kullanıcı tarafından ayarlanması gerektiği vurgulanmıştır. Yöntem, IKONOS ve QuickBird uydularından alınan çeşitli pankromatik görüntüler kullanılarak değerlendirilmiş ve önerilen algoritmanın literatürdeki diğer klasik metotlara göre hem daha doğru hem de daha verimli sonuçlar verdiği gözlemlenmiştir.

Ahmady vd. (2008) yapmış oldukları uzaktan algılanmış görüntülerden bina çıkarımına dair çalışmada aktif kontur modeline dayalı yeni bir yöntem sunmuşlardır. İlk olarak girdi görüntüyü Gauss Çekirdeği yöntemiyle yumuşatmış, ardından binaların içinde bir tohum noktası belirleyerek aktif kontur modeli ile bina sınır bilgilerine ulaşmışlardır.

Yeni aktif kontur modelinin, klasik aktif kontur modeline kıyasla binanın sınırlarını daha doğru bir şekilde tespit edebildiğini fakat görüntüde bulunan sokaklara benzer spektral özellik gösteren binaların çıkarılamaması sorunu ile karşılaşıldığı belirtilmiştir.

Ahmadi vd. (2010), tarafından yapılan çalışmada hava görüntülerinden otomatik bina çıkarımı için aktif kontur modelinin optimize edilmiş bir versiyonunu önerilmiştir. Yöntemin ilk aşamasında girdi görüntüye geometrik ve radyometrik düzeltmeler uygulanmış, sonraki adımda bina sayısına ve arka plan sınıflarına bağlı olarak, binaların çevresindeki belirli noktalar tanıtılmış ve önerilen aktif kontur modeli etkinleştirilerek bina sınırları çıkartılmıştır. Daha sonra, çıkarılan bina sınırlarında, dik açılar ve düz çizgiler kullanılarak pürüzlü çizgilerin yumuşatılması sağlanmıştır. Önerilen yöntemde kullanılan parametrelerinin optimum değerleri deneme yanılma yaklaşımı kullanılarak deneysel olarak belirlenmiş ve bir görüntü için parametreler belirlendikten sonra diğer test görüntüleri için parametrelerde değişiklikler yapılması gerektiği belirtilmiştir.

İncelenen, uzaktan algılanmış hava görüntülerinden bina çıkarımına yönelik, klasik yöntemlere dayanan çalışmaların genelinde kullanılan yöntemlerin genelleştirilebilir ve otomatize olması hususunda eksiklikler mevcuttur.

Derin öğrenme temelli yöntemler ise orijinal görüntüde bulunan özellikleri görüntüler yardımıyla otomatik olarak öğrenmekte ve yorumlayabilmektedir. Bu sayede uzaktan algılanmış görüntülerde bulunan karmaşık spektral, dokusal ve geometrik özelliklerin elde edilmesinde avantaj sağlanmaktadır (Yuan vd., 2020). Literatürde hava görüntülerinden bina çıkarımına yönelik olarak yapılan derin öğrenme temelli çalışmalara aşağıdaki açıklamalarda yer verilmiştir.

Uzaktan algılanmış görüntülerden derin öğrenme tabanlı olarak bina çıkarım işlemini gerçekleştirmek üzere Abdollahi ve Pradhan (2021), tarafından AIRS veri setinin (Chen vd., 2018) kullanımıyla, orijinal U-Net ağının geliştirilmiş bir versiyonu olan MultiRes-UNet isimli yeni bir derin öğrenme ağı ortaya çıkarılmıştır. U-Net mimarisinde birleştirme katmanı ile birleştirilen iki özellik grubu arasında anlamsal bir boşluk olduğunu ve bunun sebebinin; U-Net mimarisinde bulunan kodlayıcı bölümdeki katmanların düşük seviyeli özellikleri hesaplarken, kod çözücü bölümdeki katmanların ise dikkate değer üst düzey özellikleri hesaplamasıyla birlikte kodlayıcı ve kod çözücü katmanlarının direkt olarak birleştirilmesinin öğrenme sürecinde tutarsızlıklara neden olması şeklinde yorumlanmıştır. Bu sorun dikkate alınarak yapılan çalışmada, Res yolu adı verilen ve U-Net mimarisine, kodlayıcı ve kod çözücü özellikleri arasındaki farkları azaltmak için, kullanılan birleştirme

katmanlarıyla birlikte birkaç evrişim katmanının dahil edilmesini önerilmiştir. Bu şekilde kodlayıcı ve kod çözücü özellikleri arasındaki anlamsal boşlukların, ekstra doğrusal olmayan işlemler (evrişim) kullanılarak azaltılması sağlanmıştır.

Zhu vd. (2021) tarafından yapılan çalışmada ISPRS Vaihingen 2D ve Inria (Maggiori vd., 2017) veri setleri kullanılarak hava görüntülerinden bina çıkarımını gerçekleştirmek amacıyla segmentasyon işlemi için Kenar Detay Ağı (Edge-Detail-Network (E-D-Net)) adlı yeni bir anlamsal segmentasyon sinir ağı önerilmiştir. Önerilen E-D-Net, E-Net ve D-Net olmak üzere iki alt ağdan oluşmaktadır. E-Net, binaları arka plandan ayırmak ve bina ile arka plan arasındaki pikselleri kenar olarak sınıflandırmak için tasarlanmıştır. D-Net ise bir iyileştirme modelidir ve görevi, E-Net'in sonuçlarını iyileştirmeye yardımcı olmaktır.

Wang vd. (2020) tarafından yapılan çalışmada Massachusetts ve WHU bina veri setleri kullanılarak U-şekilli, kodlayıcı-kod çözücü yapısından ve asimetrik piramit adlı geliştirilmiş bir yerel olmayan bloktan (APNB) (Zhu vd., 2019) oluşan hem küçük hem de büyük binalarda bina çıkarımının hassasiyetini arttırmak için Yerel Olmayan U-şekilli Ağ (ENRU-Net) adlı yeni bir derin öğrenme ağı önerilmiştir. Yerel olmayan blok (APNB) kodlayıcı-kod çözücü yapısına ait özellik haritalarını dikkatlice çıkarmak ve geri yüklemek için benimsenmiş ve omurga ile son katman olan 1×1 boyutlu evrişim katmanından oluşan sınıflandırma katmanı arasına yerleştirilmiştir. Geliştirdikleri bu simetrik tasarımlı yapı, yüksek çözünürlüklü hava görüntülerinden binaları çıkarırken küçük alanlı bina sınırlarını tanıma yeteneğini önemli ölçüde iyileştirdiği gözlemlenmiştir.

Abriha ve Szabó (2023), Macaristan'ın ikinci büyük şehri olan Debrecen'deki kentsel alana ait WorldView-2 (WV2), WorldView-3 (WV3) uydu görüntüleri ve ortofoto (ORTHO) verileri ile birden fazla kaynaktan alınan uzaktan algılama verilerini kullanarak bina çıkarımı için U-Net ile derin öğrenme segmentasyon tekniklerinin az sayıda veriyle öğrenmenin etkinliğini araştırmışlardır. Eğitim için en iyi sonuçlar, ortofoto ve WV2 kaynaklarından alınan görüntülerin birleştirilmesi durumunda elde edilmiştir. Sonuç olarak U-Net'in farklı kaynaklardan gelen eğitim verileriyle dahi bina çıkarımında yüksek doğruluklar sağlayabildiği, ancak binaların çıkarılması istenilen hedef test görüntüsünden, eğitim verisine minimum sayıda görüntü eklenmesine ihtiyaç duyulduğu sonucuna ulaşılmıştır.

Feng vd. (2022) tarafından küçük bina ihmali, bina içi süreksizlik, bulanık sınır ve düzensiz bina görünümü gibi var olan birçok sorunu gidermek için Küresel Bir Özellik Yakalama ve Katmanlar Arası Bilgi Etkileşim Ağı (GCCINet) geliştirilmiştir. Yapılan

çalışmada ilk olarak küresel çok ölçekli bilgi elde etmek ve binaların sürekliliğini iyileştirmek için kodlayıcı ve kod çözücü arasında kullanılmak üzere CAFE modülü tasarlanmıştır. Ardından görüntülerden üst düzey özellikler elde edilirken küçük binaların göz ardı edilmesi sorununu çözmek için, yüksek-düşük özellikleri verimli bir şekilde birleştirmek ve yüksek düzeydeki özellikler tarafından sağlanan semantik bilgiler verimli bir şekilde kullanılırken, yerel ayrıntıları korumak amacıyla küresel bir yüksek-düşük özellik çapraz füzyon modülü (GFC) önerilmiştir. Son olarak, farklı özellik seviyelerinin iyi bir şekilde birleştirilmesi ve bina sınırlarının iyileştirilmesi amacıyla kodlayıcı ve kod çözücü katmanları arasındaki bilgileri birbirine bağlamak için Çapraz Katman Rafine Bir Füzyon ve Sınır İyileştirme Modülü (CFBR) kullanılmıştır. Ortaya çıkardıkları GCCINet derin öğrenme ağının etkinliğini doğrulamak için WHU (Ji vd.,2019) ve Inria Bina Veri Setleri (Maggiori vd., 2017) kullanılmıştır. Doğrulama sonuçlarından, kullanılan yöntemin, literatürdeki diğer yöntemlere göre göz ardı edilen veya yanlış çıkarma alanlarını içeren bölgelerin nispeten daha az olduğunu ve çıkarma sonuçlarının daha doğru olduğunu tespit edilmiştir.

Wang vd. (2022), tarafından yapılan çalışmada; U-Net'in, düşük seviyeli özelliklerin kullanımını geliştirmek için kullandığı birleştirme bağlantılarına rağmen düşük ve yüksek seviyeli özellikler arasındaki anlamsal boşluğu göz ardı etmesi sebebiyle bina çıkarım performansını sınırlaması gibi durumların çözümü hedeflenmiştir. Bu amaç için; toplu mekânsal dikkat (SA) birimlerine, küresel özellik bilgisi farkındalığı (GFIA) ve çapraz seviye özellik yeniden kalibrasyon (CLFR) isimli modüllere dayalı, B-FGC-Net isimli yeni bir bina çıkarma ağı önerilmiştir. SA birimi, bina özelliği temsilinin konumsal düzeyini vurgular, GFIA modülü bağlamsal bilgileri yakalar, CLFR modülü, düşük ve yüksek seviyeli özellikler arasındaki anlamsal boşluğu tamamen göz önünde bulundurarak, özellik bilgilerinin etkili bir şekilde birleştirilmesini sağlar, düşük seviyeli özelliklere ait gereksiz bilgileri bastırır ve modelin bina çıkarma performansını artırır. Yapılan çalışmada, önerilen yöntemi eğitmek ve değerlendirmek için WHU ve INRIA bina veri setleri kullanılarak, B-FGC-Net mimarisinin, özellikle küçük ve büyük ölçekli binalarda yüksek doğruluklu çıkarım etkisine sahip olduğu ve ağaç ile bina gölgeleri etkilerinin üstesinden geldiği gözlemlenmiştir.

Wang vd. (2022) kırsal bölgelerde bulunan binaları çıkarmak amacıyla insansız hava aracından (İHA) elde edilen 3 bantlı (RGB) görüntülerin kullanılmasıyla, Mask R-CNN (He vd., 2018) derin öğrenme ağının tanıma doğruluğunu iyileştirmek için ilgili mimariyi,

Görünür Işık Farkı Bitki Örtüsü İndeksi (VDVI) özellikleri ve Sobel kenar algılama özellikleriyle birleştirmiştir.

Zhao vd. (2022), tarafından yapılan çalışmada UNet++ derin öğrenme mimarisine dayanarak, binaları uzaktan algılanmış hava görüntülerinden otomatik olarak çıkarmak için AEUNet++ isimli yeni bir ağ tasarlanmıştır. Önerilen derin öğrenme ağında çıkarılan özellik haritalarının daha kullanışlı bir biçimde birleştirilmesi için mekansal dikkat ve kanal dikkat yapılarını içeren Evrişimli Blok Dikkat Modülü (CBAM) (Woo vd., 2018) kullanmıştır. Bu sayede UNet++ mimarilerindeki özellik haritalarının ortalamasını alarak, daha iyi özellik haritalarının bastırılması hususundaki dezavantajları önemli ölçüde ortadan kaldırılmıştır. Ayrıca çıkarılan binaların sınır hassasiyetini geliştirmek için, binanın sınır geometrik bilgisi, önermiş oldukları mesafe sınıfı haritası ile AEUNet++ mimarisine dahil edilmiştir. Önerilen mimari, Massachusetts ve WHU bina veri setleri üzerinde test edilmiş ve ağı bina çıkarma konusunda U-Net ve UNet++ mimarilerine göre daha doğru sonuçlar ürettiği sonucuna varılmıştır.

Hu vd. (2019) tarafından uydu görüntülerini kullanarak yapılan çalışmada Derin Otomatik Bina Çıkarma Ağı (DABE_Net) isimli yeni bir derin yapay sinir ağı önerilmiş ve ayrıca segmentasyon doğruluğunu arttırmak amacıyla ağa bir dikkat mekanizması eklenilmiştir. Qiu vd. (2022) ise mevcut bina çıkarma mimarilerinin çoğunda hesaplama karmaşıklığı olduğu ve yüksek verimlilik gerektiren pratik uygulamalardaki kullanımlarının sınırlı olduğu düşüncesine dayanılarak, bu durumun önüne geçmek adına, MSL-Net adı verilen, kodlayıcı-kod çözücü yapısına sahip bir derin öğrenme ağı önermişlerdir. Tasarlanan MSL-Net mimarisinde, MobileNetV2 (Sandler vd. 2018) derin öğrenme ağı, orijinal görüntü özelliklerini çıkarmak için omurga olarak tanıtılmıştır. Ağ karmaşıklığını azaltmak için Derinlemesine Ayrılabilir Evrişim (DSC) (Sifre & Mallat, 2014) işlemi ve yüksek seviyeli özellik haritalarını çıkarmak için kod çözücü bölümde bir ASPP modülü kullanılmıştır. Binaların şekillerinin çeşitliliği sebebiyle normal bir evrişimli ağı binaların şekil özelliklerini çıkarmakta zorluk çekmesi nedeniyle, Deforme Olabilen Evrişim Ağı (DCN) (Zhu vd., 2019) kullanılarak bir görüntüdeki özellikler verimli bir şekilde çıkarılmıştır. Önerilen yöntem WHU, Inria ve Massachusetts veri setleri üzerinde değerlendirilmiştir. Geliştirilen metodun farklı şehirlerdeki kentleşmeyi izlerken iyi performans gösterebileceği tespit edilmiştir. Fakat karmaşık görüntülere sahip Inria veri kümesi üzerinde elde edilen sonuçların yeterince üstün olmadığı gözlemlenmiştir.

Lin vd. (2019) tarafından mevcut çalışmaların yüksek doğruluk için çok miktarda parametre ve işlem gerektirdiği bu durumun yüksek bellek tüketimi ve yavaş çıkarım hızına sebep olduğu problemine dayanılarak, düşük hesaplama maliyeti ve bellek tüketimi ile yüksek doğruluk elde etmeyi amaçlayan ESFNet adında yeni bir derin öğrenme ağı tasarlanmıştır. Önerilen mimaride standart evrişim işlemi yerine SFRB (Ayrılabilir Faktörize Artık Blok) ismi verilen modül kullanılmıştır. Bu sayede mimari boyutu ve hesaplama maliyeti büyük ölçüde azaltılmıştır.

Literatürde klasik ve derin öğrenme tabanlı olarak yapılan bina çıkarma çalışmalarının incelenmesi sonucu derin öğrenme tabanlı yöntemlerin, klasik yöntemlere karşı daha otomatize ve genelleştirilebilir olduğu sonucuna ulaşılmıştır. Fakat derin öğrenme tabanlı yöntemlerin olumlu yönlerinin yanında, yüksek miktarda veri ihtiyacı ve bellek tüketimi, karmaşık ve çok katmanlı derin öğrenme ağları sebebiyle mimari eğitimleri için uzun zaman harcanması gibi dezavantajların olduğu gözlemlenmiştir. Bu sebeplerden ötürü derin öğrenmeye dayalı yöntemlerde her zaman, daha başarılı derin sinir ağlarının ve metotların üretilmesi hedeflenmektedir.

Yapılan tez çalışmasında ise minimum düzeyde karmaşıklığa sahip bir derin öğrenme ağı tasarlanması hedeflenerek, eğitim işleminin kısaltılması amaçlanmıştır. Bunun yanı sıra uzaktan algılanmış görüntülerden bina çıkarımı konusunda kullanılan veri setinden maksimum düzeyde verim alınabilmesi için tasarlanan derin öğrenme ağına bina ve arka plan bilgilerine ilave olarak, morfolojik operasyonlar ile otomatik bir şekilde oluşturulan bina sınır bilgileri dahil edilmiştir. Bina sınır bilgilerinin dahil edilmesiyle gerçekleştirilen üç sınıflı segmentasyon eğitimi sonucunda bina çıkarım performansının arttığı gözlemlenmiştir.

1.6. Temel Kavramlar

1.6.1. Fotogrametri

Fotogrametri, fotoğrafları ölçüm için temel araç olarak kullanan, fotoğraflardan; yeryüzüne ait nesneler, detaylar ve arazi özelliklerinin üç boyutlu hassas koordinatlarının elde edilebildiği ve yaklaşık olarak fotoğrafın kendisi kadar eski olan bir ölçüm tekniğidir (Aber vd., 2010; Siegel & Sukko, 2012). Fotogrametri bilimi yersel ve hava fotogrametrisi olmak üzere ikiye ayrılmaktadır. Yersel fotogrametride fotoğraflar, konumu ve yönelimi

yatay olacak şekilde doğrudan pozlama anında yer tabanlı kameralarla çekilir. Hava fotogrametrisinde ise fotoğraflar düşey veya eğik olarak sınıflandırılır. Düşey fotoğraflar, fotoğraf düzlemi ile referans düzleminin paralel olması durumunda ortaya çıkan fotoğraflardır. Fakat uygulamada, uçuş esnasında kaçınılmaz bir biçimde ortaya çıkan eğiklikler nedeniyle kamera eksenini nadiren tam olarak düşey olabilmektedir. Kamera ekseninin istenmeden düşey yönden 5° ve üzeri bir değerde saptığı durumda ortaya çıkan fotoğrafa eğik fotoğraf adı verilmektedir (Gomasasca, 2009). Küçük ölçekli hava fotogrametrisi yaygın olarak topografik ve tematik haritaların üretimi, sayısal arazi modellerinin elde edilmesi işlemlerinde; büyük ölçekli hava fotogrametrisi ise Kent Bilgi Sistemleri (KBS), 3B şehir modellerinin oluşturulması uygulamalarında kullanılabilmektedir.

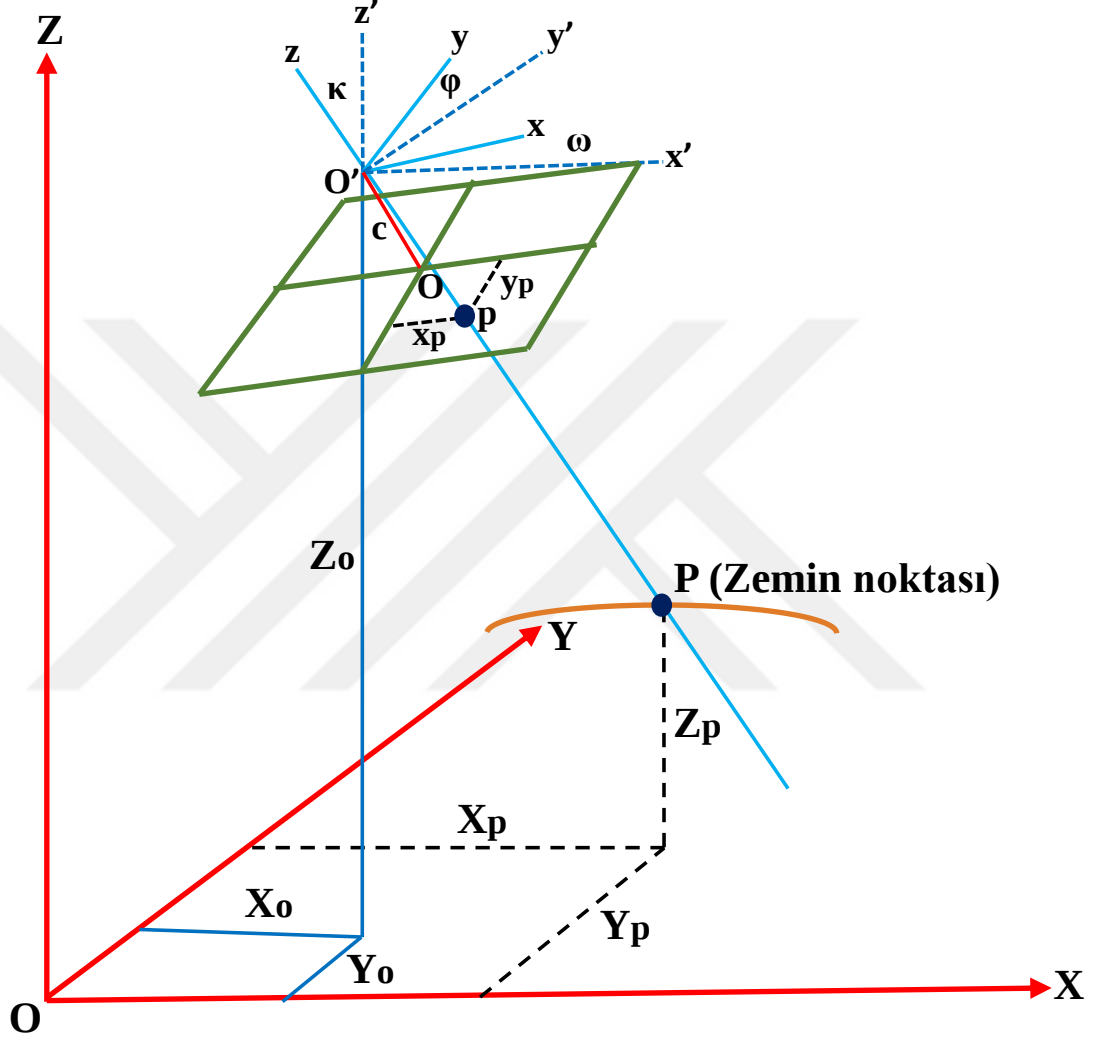
Topografik haritalama için fotogrametri kullanımına ilişkin ilk gerçek deneyler, 1849 yılında Fransız Ordusu Mühendisler Birliği'nden Albay Aimé Laussedat'ın yönetimi altında gerçekleştirilmiştir. Albay Laussedat deneylerinde havadan fotoğraf çekmek için uçurtma ve balonlar kullanmıştır. Fakat havadan fotoğraf elde etmede karşılaşılan zorluklar nedeniyle araştırma alanını kısıtlanmış ve çalışmalarını yersel fotoğraflarla haritalama üzerinde yoğunlaştırmıştır. Albay Laussedat, Bu konudaki öncü çalışmaları ile “Fotogrametrinin Babası” unvanını elde etmiştir.

1861'de üç renkli fotoğraf süreci başlamış, 1891'de rulo film geliştirilmiş, 1909'da ise Almanya'dan Dr. Carl Pulfrich, bindirmeli fotoğraf çiftleriyle denemeler yapmaya başlamıştır. Dr. Carl Pulfrich'in çalışmaları, günümüzde kullanılan birçok aletli fotogrametrik haritalama tekniğinin geliştirilmesinin temelini oluşturmaktadır.

Uçağın 1903 yılında Wright kardeşler tarafından icadı, modern hava fotogrametrisinin ortaya çıkışına büyük ivme kazandırmış ve uçak ilk kez 1913 yılında haritalama amacıyla fotoğraf elde etmek için kullanılmıştır. Birinci Dünya Savaşı sırasında hava fotoğraflarından özellikle keşif amaçlı olarak yoğun bir şekilde yararlanılmış, İkinci Dünya Savaşı sırasında ise haritalara yönelik büyük talebi karşılamak için fotogrametrik tekniklerin yaygın olarak kullanıldığı bir dönem yaşanmıştır. Teknoloji alanında ve fotogrametri tekniklerindeki gelişmeler, 20. yüzyılın geri kalanında devam etmiş ve 21. yüzyılda da hızlı bir şekilde devam etmektedir. (Gomasasca, 2009).

Fotogrametride kullanılan temel prensip üçgenleme yöntemidir. En az iki farklı konumdan fotoğraf çekilerek, her kameradan 3B bilgi elde edilmek istenen nesne üzerindeki noktalara ‘görüş hatları’ adı verilen çizgiler oluşturulabilmektedir. Bu görüş hatları (Şekil

1), istenilen noktaların üç boyutlu koordinatlarını üretmek için matematiksel olarak keşismektedir (Horswell, 2013). Bu bilgilere dayanarak, bir detayın art arda çekilmiş en az iki fotoğrafta bulunması durumunda 3B koordinat bilgisine ulaşılabilceği sonucuna varılmaktadır.



Şekil 1. Fotogrametrik veri alım geometrisi

Fotoğraf ile gerçek dünyada bulunan detaylar arasındaki ilişki izdüşüm (kollinearite) denklemleri ile sağlanmaktadır (Denklem 1 ve 2). İzdüşüm denklemleri Şekil 1’de gösterildiği üzere, bir sensörün izdüşüm merkezini, görüntü noktasını ve yerdeki bir nesneyi birbirine bağlayan bir ışın demetinin geometrisini vermektedir. Bu ilişkiyi kurabilmek adına öncelikle iç (x_0, y_0, c) ve dış $(\omega, \phi, \kappa, X_0, Y_0, Z_0)$ yöneltme parametre bilgilerine erişilmesi gerekmektedir.

$$x = x_0 - c \frac{m_{11}(X - X_0) + m_{12}(Y - Y_0) + m_{13}(Z - Z_0)}{m_{31}(X - X_0) + m_{32}(Y - Y_0) + m_{33}(Z - Z_0)} \quad (1)$$

$$y = y_0 - c \frac{m_{21}(X - X_0) + m_{22}(Y - Y_0) + m_{23}(Z - Z_0)}{m_{31}(X - X_0) + m_{32}(Y - Y_0) + m_{33}(Z - Z_0)} \quad (2)$$

Denklemlerde bulunan x ve y görüntü/resim koordinatlarını, x_0 ve y_0 asal nokta koordinatlarını, m_{ij} dönüklük matrisi elemanlarını, m_{ij} içerisindeki ω , ϕ , κ değişkenleri açısal dönüklük ve eğiklik elemanlarını, X_0, Y_0, Z_0 izdüşüm merkezi (fotoğraf çekim konumu) koordinatlarını, c fotoğraf düzlemi ile izdüşüm merkezi arasındaki uzaklık olan asal uzaklığı (kamera sabiti) ve son olarak X, Y, Z ise fotoğrafta bulunan ve 3B koordinatları hesaplanmak istenen bir detay noktasının arazi koordinatlarını temsil etmektedir.

1.6.2. Hava Fotoğrafı Üretim Teknikleri

Hava fotoğrafçılığı, uzaktan algılamanın en eski şeklidir. Günümüzde havadan çekilmiş film fotoğrafları ve dijital görüntüler, uzaktan algılama verilerinin en güvenilir ve yaygın olarak kullanılan kaynaklarından biri olmaya devam etmektedir (Fagan & DeFries, 2022). Hava fotoğrafları; uçak, helikopter, drone ve uydu gibi platformlara yerleştirilmiş dijital ve analog kameraların kullanımıyla objelerden yansıyan ışığın kaydedilmesi ile elde edilen görüntülerdir.

Analog hava kameralarında fotoğraf düzleminde genellikle bir film bulunur. Hızlı ve otomatik biçimde arka arkaya fotoğraf çekebilir. Her yeni fotoğraf çekimine geçiş süresi içinde pozlanan filmin sarılması, yeni filmin tam düz hale getirilmesi, arzu edilen bindirmeye karşılık gelecek şekilde uygun zaman aralığında ve istenen poz süresi kadar kısa bir süre içinde objektifin açılması gerekmektedir. Analog hava kameraları, fotoğraf boyutları ve asal uzaklık arasındaki orana göre ortaya çıkan görüş açısına yönelik olarak sınıflandırılır. Kamera görüş açısı büyüdükçe bir fotoğrafın arazide kapladığı alan genişlemektedir. Analog fotoğraflarda ölçek kavramı mevcuttur ve fotoğrafın çözünürlüğü 1 mm'de ayırt edilebilen maksimum çizgi çifti sayısı ile belirlenir. Teknolojinin gelişmesi ve yeni tekniklerin ortaya çıkmasıyla birlikte analog fotogrametriden dijital fotogrametriye geçiş dönemi başlamıştır. Bu gelişim süreci ile analog hava kameraları yerini dijital hava kameralarına bırakmıştır.

Dijital hava kameralarının kullanımıyla birlikte, analog fotogrametrinin beraberinde getirdiği film banyo işlemleri sona ermiş, görüntü kayıt sistemi olarak CCD/CMOS ve her fotoğraf pozlama anında kamera konumunu belirleyen GPS/IMU sistemleri kullanılmaya başlanmıştır. Çözünürlük kavramı ise bir pikselin yeryüzünde kapladığı alan, Yer Örnekleme Aralığı (YÖA) ifadesi ile tanımlanmıştır. Analog ve dijital hava kameraları ile elde edilen ortofotolar arasındaki çözünürlük farkı Şekil 2’de açık bir biçimde gözükmemektedir.

Dijital hava fotoğrafları dijital kameralar ile elde edilebileceği gibi analog hava fotoğraflarının sayısallaştırılması vasıtasıyla da elde edilebilmektedir. Bu dönüşümün en önemli sebepleri; nesne tanıma, çıkarma ve veri toplamada otomasyon sağlanması, dijital olarak elde edilen verilerin diğer sistemlerde analiz, yorumlama ve girdi verisi olarak kullanımının mümkün olmasına ek olarak bellek maliyetinin düşük olmasıdır (Özbalımcı, 1996).



Şekil 2. Analog ve dijital kameralar ile elde edilen ortofotolar (a) Analog, (b) Dijital (Erol & Öner, 2011)

1.6.3. Dijital Hava Kameraları ve Özellikleri

Hava Kameraları; metrik kamera veya ölçüm kamerası olarak da adlandırılan, hava fotogrametrisi için gerekli resimlerin çekiminde kullanılan, özel yapıdaki bir uçağa, helikoptere veya bir drona monte edilerek kullanılan, metrik özellikleri ve iç yöneltme

elemanları hassas olarak bilinen, ilave olarak görüntü alımı anında, kameranın konumu ve dönüklük parametrelerinden oluşan dış yöneltme elemanlarının belirlenebilmesi için platformlarında GPS/INS (IMU) sistemleri bulundurulmuş kameralardır (URL-1, 2023). Hava kameraları ile elde edilen hava fotoğrafları; haritacılıkta topografik harita ve ortofotoların üretiminde, arazi kullanım planlamasında ve arkeolojide kullanılmaktadır.

Hava fotogrametrisi için kullanılan kameralar dijital ve analog olmak üzere iki sınıfa ayrılmaktadır. Fotogrametri alanında yapılan teknolojik gelişmeler çoğunlukla doğruluğun artırılması ve maliyetin azaltılması yönündedir. Analog kameralardan, dijital hava kameralarına geçiş ile fotogrametrik doğruluğun artırılmasının yanında proje maliyetlerinin düşmesi yönünde de iyileştirmeler sağlanmıştır (Kiracı vd., 2010).

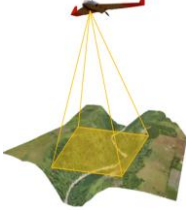
Dijital kameralar çerçeve ve satır tarayıcılı olmak üzere görüntü geometrisi açısından ikiye ayrılmaktadır. Kullanımı çoğunlukla tercih edilen çerçeve kameralar CCD ve CMOS adında alıcılara sahiptir. CMOS alıcıların radyometrik çözünürlükleri (bit derinliği) düşüktür ve küçük piksel boyu ile maksimum sayıda piksel üretirler. CCD (Charge Couple Device) alıcılar ise CMOS'a göre daha kaliteli görüntü alımı gerçekleştirebilmektedir. Dijital hava kameraları; kamera gövdesi, mercek, taşıyıcı (CCD veya CMOS), diyafram ve obtüratör bileşenlerinden oluşmaktadır. Dijital görüntü için en önemli kamera bölümü olan CCD veya CMOS sensörleri, elektronları sensör üzerinden herhangi bir deformasyona uğramadan aktarabilmeleri için üretimlerinde özel bir teknik kullanılmaktadır ve bu sayede yüksek kaliteli ve konumsal çözünürlüğü yüksek görüntüler elde edilebilmektedir (Makineci vd., 2020).

Dijital kameraların kullanımıyla birlikte piksellere atanan renk değerleriyle gerçek dünyadaki cisimleri temsil eden dijital fotoğraflar kullanılmaya başlanmıştır. Dijital fotoğraflarda uzamsal/konumsal, radyometrik ve spektral olmak üzere üç çeşit çözünürlük kavramı mevcuttur. Uzamsal çözünürlük, bir dijital fotoğrafın gerçek dünyadaki karşılığını temsil eden piksel sayısıdır. Piksel sayısı arttıkça uzamsal çözünürlük artmakta ve detaylar daha net bir şekilde ortaya çıkmaktadır. Radyometrik çözünürlük, kullanılan dijital kameranın cinsine bağlı olarak fotoğrafların temsil edildiği bit derinliğini (gri değer hassasiyeti) ifade etmektedir. Spektral çözünürlük ise görüntünün sahip olduğu bant/kanal sayısı ile ifade edilmektedir. Günümüzde tek bantlı pankromatik, çok bantlı (1-10) multispektral ve çok daha fazla bant sayısına (>10) sahip hiperspektral hava görüntüleri kullanılmaktadır.

Dijital hava kameraları uçak, drone gibi platformlar vasıtasıyla yeryüzü hakkında veri sağlamaktadır. 2001-2023 yılları arasında üretilip, son güncellemeleri 2010-2023 yılları

arasında yapılan örnek dijital hava kameraları ve özellikleri Tablo 1’de sunulmuştur (URL-2, 2023; URL-3, 2023).

Tablo 1. Örnek hava kameraları ve özellikleri

						
Kameralar ve Özellikleri	IGI Urbanmapper-2	Phase One IXM-RS280F	Ultracam Eagle 4.1	Ultracam Merlin 4.1 3020	CityMapper Airborne Hybrid Sensor	Ultracam Condor 4.1
Piksel Boyutu	3.76 μm	3.76 μm	3.76 μm	3.76 μm	5.2 μm	3.76 μm
Pankromatik Görüntü Boyutu	34500×14100	20150×14118	28110×18060	-	-	20488×14040
Renkli Görüntü Boyutu	30460×14100	14204×10652	9.370×6020	31300×14016	10320×7752	48462×6150
Mevcut Spektral Bantlar	RGB (Oblik) NIR (Nadir) CIR PAN	RGB NIR CIR PAN	RGB NIR PAN	RGB NIR	RGB NIR	RGB NIR PAN
Renk Derinliği	16 Bit	14 Bit	14 Bit	14 Bit	14-16 Bit	14 Bit
Kullanılan Sensör	BSI-CMOS	CMOS	CMOS	CMOS	CCD	CMOS
RGB: Kırmızı-Yeşil-Mavi, NIR: Yakın kızılötesi, PAN: Pankromatik, CIR: Renkli kızılötesi (yakın kızılötesi-kırmızı-yeşil). BSI-CMOS: Arkadan aydınlatmalı CMOS sensör						

1.6.4. Uzaktan Algılanmış Verilerden Bina Çıkarımı

Uzaktan algılama teknolojilerinin gelişimi ve birçok alanda kullanımına başlanması sayesinde yeryüzünde bulunan tüm detayların yoğun iş gücü, zaman ve maliyet gerektiren yersel ölçüm teknikleri yerine; hızlı, otomatize ve daha az maliyetle uzaktan algılama verileri kullanılarak tespit, ölçüm ve yorumlanma işlemleri gerçekleştirilebilmektedir. Yeryüzünde

en fazla yüzey alanına sahip objelerden biri olması sebebiyle, bilgisine en çok ihtiyaç duyulan nesnelerden biri olan bina katmanının, uzaktan algılama verilerinden tespiti, çeşitli büyük ölçekli haritaların üretimine ve projelere katkı sağlamaktadır.

Bina objelerinin tespiti ve sınır bilgilerinin elde edilmesi işlemi; sıklıkla LiDAR nokta bulutları, yüksek çözünürlüklü uydu görüntüleri, hava fotoğrafları ve LiDAR nokta bulutları ile hava görüntülerinin entegre bir biçimde kullanımı üzerinden çeşitli yöntemler ile gerçekleştirilmektedir.

Hava fotoğrafları üzerinden bina objelerinin bilgisine ulaşma işlemi klasik ve derin öğrenme tabanlı yöntemlerin kullanımı şeklinde ikiye ayrılmaktadır. Klasik yöntemlerin kullanımı için yoğun miktarda ön bilgiye ihtiyaç duyulmasıyla birlikte otomasyonu düşüktür. El yordamı ile binaların geometrik, spektral ve gölge özellikleri çıkarılıp, ardından geliştirilen çeşitli sınıflandırma teknikleri ile bina ve bina olmayan sınıflar birbirinden ayrılmaktadır. Literatürde, klasik yöntemler ile binaların çıkarılması amacına yönelik, doğrudan veya dolaylı olarak; aktif kontur modeli, kenar algılama teknikleri, matematiksel morfolojik operasyonlar, bulanık mantığa dayalı sınıflandırma yöntemleri, Hough dönüşümü, Rastgele Orman Sınıflandırıcısı, Su Havzası (Watershed) algoritması ve eğitim analizi yöntemleri kullanılmıştır.

Yılanlar (snakes) olarak bilinen aktif kontur modeli görüntülerde nesnelerin sınır tespiti için kullanılan bir tekniktir. Nesnenin sınırına yakın bir başlangıç konturu oluşturmayı ve nesnenin kenarlarına tam olarak uyacak şekilde konumunu yinelemeli olarak ayarlamayı içerir. Aktif kontur segmentasyonu, eşikleme veya kenar tespiti gibi yöntemlerin başarısız olabileceği durumlarda özellikle kullanışlıdır (Kass vd., 1988).

Nesnelerin tanımlanması için sınırlarının belirlenmesi önemli olduğu için kenar bulma, görüntü işleme için en önemli aşamalardan biridir. Görüntü içinde belirgin geçişler nesne kenarlarını gösterir ve kenar bulma algoritmaları da bu geçişleri bulmayı amaçlamaktadır. Literatürde temel kenar belirleme algoritmalarından bazıları; Prewitt (Prewitt, 1970), Sobel (Sobel, 1970), Canny (Canny, 1986), Marr ve Hildreth'dir (Marr & Hildreth, 1980).

Klasik mantıktaki küme kavramı $[0, 1]$ küme elemanları ile adlandırılırken, bulanık mantık ise bu klasik mantığın yapısını içine alarak sadece iki değil sonsuz derecede doğruluk değerinin olduğunu ve $[0, 1]$ aralığındaki sayısal değerleri de ilişkilendiren fonksiyon olarak ortaya çıkarmaktadır (Zadeh, 1965).

Hough Dönüşümü, görüntülerdeki doğru ve daireleri tespit etmeyi kolaylaştıran, görüntü işleme alanlarında kullanılan oylama mantığı ile çalışan bir algoritmadır. Görüntülerde şekillerdeki süreksizlikler şekil tespitini zorlaştırmaktadır. Bu noktada hough dönüşümü ile görüntünün tamamının görülebilir olmadığı durumlarda da olası şekillerin tespiti mümkün olmaktadır (Duda & Hart, 1972).

Rastgele Orman Sınıflandırıcısı tek bir sonuca ulaşmak için birden fazla karar ağacının çıktısını birleştiren bir algoritmadır (Breiman, 2001).

Su Havzası algoritması, gri tonlamalı bir görüntü üzerinde tanımlanan bir dönüşümdür. Uygulandığı görüntüde her noktanın parlaklığını, bir yüksekliği temsil ettiği bir topografik harita gibi ele alır ve sırtların tepeleri boyunca uzanan çizgileri bulur. Bu algoritma genellikle bir görüntüdeki üst üste binen nesnelerin tespit edilmesine yardımcı olur (Najman vd., 1994).

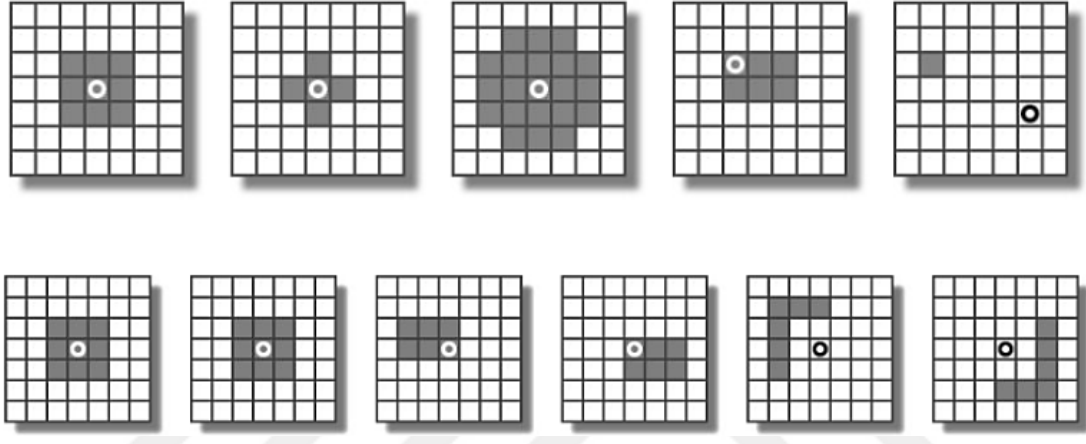
Eğim analizi yöntemleri ise sayısal arazi modeli veya yükseklik modeli üzerinden hesaplanan eğim değerlerinin istenilen tematik aralıklarda gösterilmesine dayanmaktadır.

Klasik yöntemlere karşın, derin öğrenme yöntemleri orijinal görüntüde bulunan özellikleri görüntüler yardımıyla otomatik olarak öğrenebilmekte ve yorumlayabilmektedir. Ayrıca derin öğrenme, klasik yöntemlere nispeten daha genelleştirilebilir bir yöntemdir. Bu bilgiler ışığında derin öğrenme yöntemleri ile daha az iş gücü ve ön bilgi ihtiyacıyla otomatize bir şekilde bina objelerinin ayak izi bilgisine ulaşılabilirdiği gözlemlenmektedir. Literatürde derin öğrenme yöntemleri ile hava görüntülerinden bina katman bilgisine, hızlı ve gerçeğe en yakın biçimde ulaşmak amacıyla birçok yeni yöntem ve derin öğrenme ağı tasarlanmıştır.

1.6.5. Matematiksel Morfolojik Operasyonlar

“Morfoloji”, Yunancada “şekil, biçim” anlamında kullanılan *morf* kelimesinden türemiştir. Matematiksel morfolojik operasyonlar; topoloji, küme teoremi, rastgele fonksiyonlar ve kafes kuramı kavramlarının görüntü işleme ve analizinde uygulanmasıdır. Bu operasyonlar sıklıkla dijital görüntüler üzerinde kenar belirleme, bölge ayırma ve görüntü netleştirme gibi amaçlar için kullanılmaktadır (Özşen, 2002). Morfolojik operasyonlar temel olarak genişleme ve aşınma isimli iki işlemten oluşmaktadır ve gri tonlu görüntüler üzerinde kullanılabileceği gibi genellikle 0 ve 1 renk değerlerinden oluşan siyah-beyaz ikili (binary) formattaki görüntüler üzerinde uygulanmaktadır (Gonzalez & Woods, 2004).

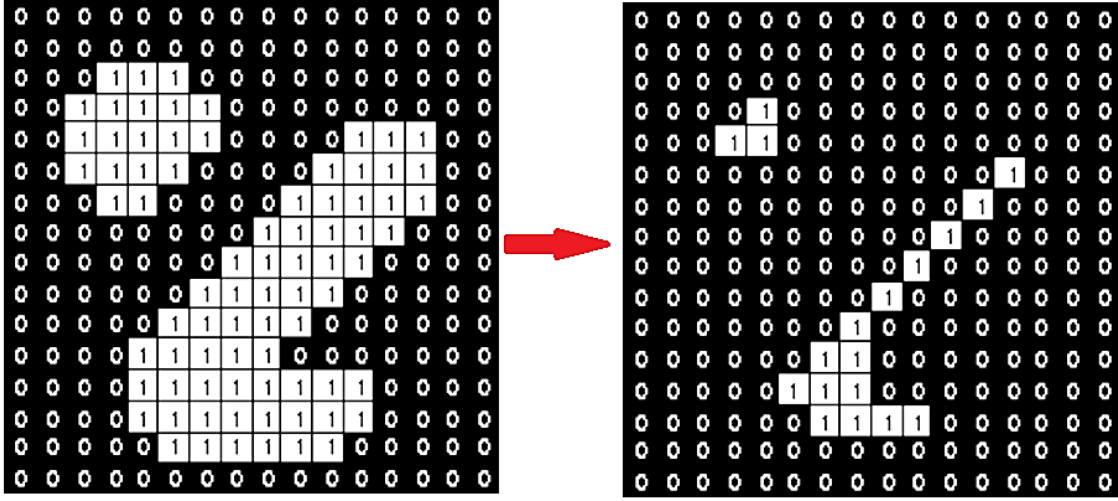
Morfolojik operatörlerin iki adet girdi verisi bulunmaktadır. Bunlar; üzerinde genleşme ya da aşınma uygulanacak görüntü ve bu işlemlerin uygulanmasında araç olarak kullanılan, genleşme ve aşınma işlemlerinin şeklini belirleyen yapı elemanıdır. Yapı elemanları, farklı şekil ve büyüklüklerde olabilmekte ve bir merkez noktaları mevcut olup, işlem uygulanacak görüntünün her bir pikseli bu nokta ile üst üste getirilerek uygulama yapılmaktadır (Acar & Bayram, 2009). Kullanılan yapı eleman matrislerinden bazıları görsel olarak Şekil 3’te sunulmuştur.



Şekil 3. Matematiksel morfolojik operasyonlar için örnek yapı elemanları (Acar & Bayram, 2009)

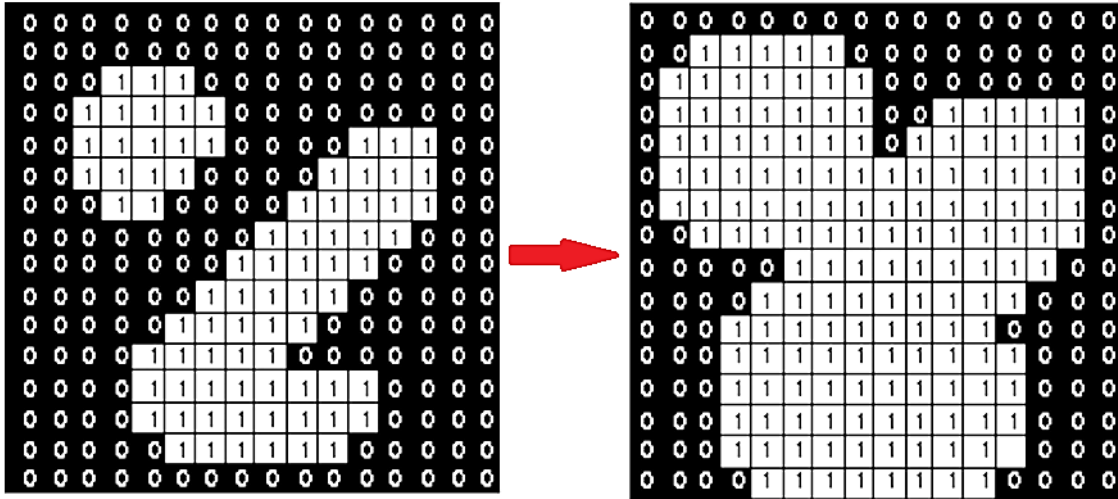
Morfolojik görüntü işlemenin iki temel işleminden biri olan aşınma işleminde F görüntü ögesinin, S yapı elemanı tarafından aşınımı $F \ominus S$ ile gösterilir. Yapı elemanı olarak seçilen S 'nin merkez noktası görüntüdeki (x, y) koordinatlı merkez üzerine yerleştirildiğinde, yapı elemanı ile karşılaştırıldığı görüntü piksellerinin uyum sağlaması durumunda, yapı eleman merkezinin üzerinde olduğu piksele 1, diğer piksellere ise 0 renk değeri atanır. Bu kuralın matematiksel gösterimi Denklem 3'te, örnek görsel üzerinde gösterimi ise Şekil 4'te yer almaktadır.

$$g(x, y) = \begin{cases} 1 & \text{eğer } S \text{ eşleşir } F \\ 0 & \text{eğer } S \text{ eşleşmez } F \end{cases} \quad (3)$$



Şekil 4. Aşınma işlemi (URL-4, 2023)

Temel işlemlerden bir diğeri olan genleşme işleminde ise F görüntüsünün, S yapı elemanı tarafından genişletilmesi $F \oplus S$ ile gösterilir. Genleşme işleminin temel etkisi görüntüyü kenarlarından genişletmektir. Yapı elemanının merkez noktasının üzerine uygulandığı görüntü parçası ile aynı değeri alması durumunda, yapılandırma elemanı matrisi renk değerleri olduğu gibi görüntü parçasına aktarılır. Genleşme işleminin görsel üzerindeki temsili Şekil 5'te gösterilmiştir.



Şekil 5. Genleşme işlemi (URL-4, 2023)

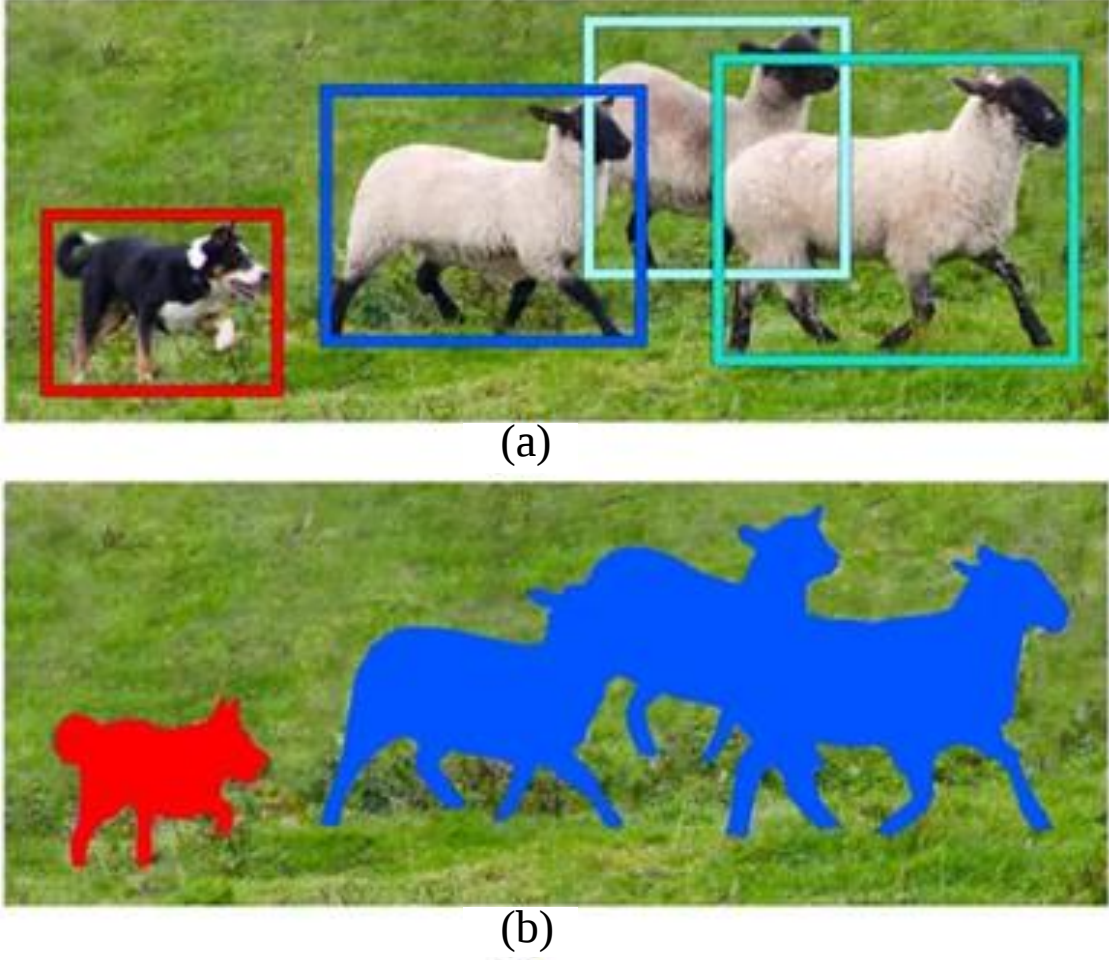
Basit matematiksel morfolojik operasyonlar olan aşınma ve genleşme işlemlerinin birlikte kullanılmasıyla gerçekleştirilen açma ve kapama isimli birleşik morfolojik

operasyonlar ortaya çıkmıştır. Açma işlemi görüntüye aşınma işlemi uygulandıktan sonra genleşme işleminin uygulanmasıdır. Kapama işlemi ise genleşme işlemi takip eden aşınma işleminden oluşmaktadır. Açma işleminin matematiksel gösterimi $F \circ S = (F \ominus S) \oplus S$ ve kapama işleminin matematiksel gösterimi ise $F \bullet S = (F \oplus S) \ominus S$ şeklindedir (Gonzalez & Woods, 2004).

1.6.6. Görüntü Segmentasyonu ve Segmentasyon Türleri

Uzaktan algılama teknolojilerinin gelişmesiyle birlikte yüksek çözünürlüklü uydu görüntüleri ve hava fotoğraflarından yeryüzü hakkında bilgi sağlama çalışmaları hız kazanmıştır. Bu sebeple yeryüzünde bulunan nesnelerin uydu görüntüleri ve hava fotoğraflarındaki karşılıklarının çıkarımı uzaktan algılama alanında önemli bir yer edinmiştir. Fotoğraflara uygulanan nesne tanıma algoritmaları çıktı olarak görselde bulunan objelere ait sınıfları ve objelerin sınırlayıcı kutuları (bounding box) ile vermektedir (Şekil 6a). Fakat bu sınırlayıcı kutular ile detayların sınırlarının tespit edilememesi sebebiyle bazı durumlarda yeterince işe yaramamaktadır. Bunun yanı sıra klasik piksel-tabanlı sınıflandırma metotları, yüksek çözünürlüklü görüntülerdeki nesnelerin genellikle farklı, heterojen spektral özniteliklere sahip piksellerden oluşması sebebiyle detayların segmentlere ayrılması aşamasında yetersiz kalmaktadırlar (Çölkesen & Kavzoğlu, 2017).

Segmentasyon ise görüntüdeki belirleyici özellikleri temel alarak gerçek yeryüzü nesnelere karşılık gelen detayların segmentlerini oluşturma işlemidir. Segmentasyon işlemi (Şekil 6b), nesne tabanlı sınıflandırmanın en temel adımını oluşturmakla birlikte görüntü üzerinde segmentlere ayrılan bölgeler, nesne tabanlı analizin çekirdek yapılarıdır (Benz vd., 2004; Blascke, 2010). Sınıflara ayrılan segmentlerin doğruluğunu değerlendirmek için referans veri seti ve çeşitli segmentasyon değerlendirme metrikleri üretilmiştir (Tonbul & Kavzoğlu, 2017).

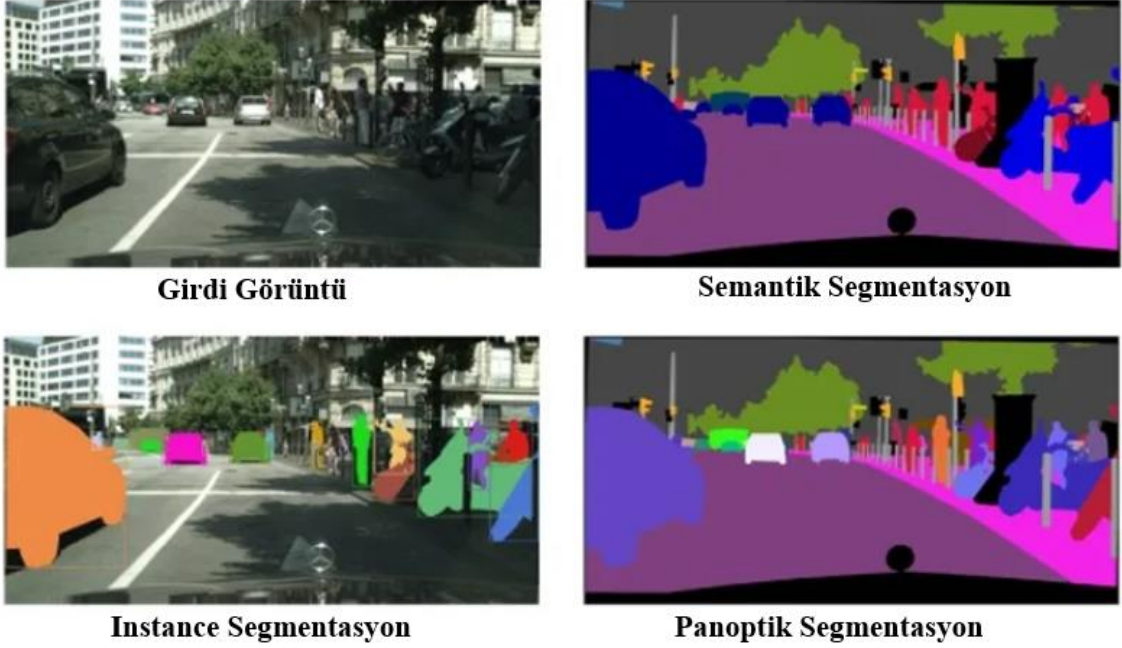


Şekil 6. Görüntülerden nesne çıkarımı (a) Obje tespiti, (b) Segmentasyon (URL-5, 2023)

Sağlık, jeoloji, çevre, harita, uzaktan algılama gibi birçok alanda görüntü segmentasyonu kullanılmaktadır. Görüntü segmentasyonu temelde *semantik segmentasyon* ve *instance segmentasyon* adında iki farklı bölütleme tekniğinden oluşmaktadır. İlave olarak semantik ile instance segmentasyonun birleşiminden oluşan *panoptik segmentasyon* çeşidi de bulunmaktadır (Kale & Thorat, 2021).

Semantik segmentasyon, bir görüntüdeki her pikseli önceden tanımlanmış birkaç sınıftan birine atamaktadır. Semantik segmentasyon ile her piksele ait bir sınıf etiketinin bulunduğu, ancak aynı sınıfın farklı örneklerinin birbirinden ayırt edilemediği bir etiket haritası oluşturulur. Instance segmentasyon, semantik segmentasyondan farklı olarak görüntüdeki her pikseli sınıflandırmanın yanı sıra, aynı sınıfın nesnelerini de ayırır. Panoptik segmentasyon ise piksellerin atandığı sınıflara benzersiz bir örnek etiketi vererek, bir görüntünün eksiksiz ve birleşik bir segmentasyonunu sağlamaktadır. Literatürde bulunan Her Şeyin Segmentasyonu (Segment Anything Model (SAM)) (Kirillov vd., 2023)

uygulaması bir panoptik segmentasyon örneğidir. Bahsedilen segmentasyon çeşitleri arasındaki farklar görsel olarak Şekil 7’de verilmiştir. Tüm segmentasyon türlerinde pikseller ve sınıflar arasındaki ilişkileri öğrenmek için genellikle Evrişimli Sinir Ağı (CNN) mimarileri kullanılmaktadır.



Şekil 7. Segmentasyon türleri (URL-6, 2023)

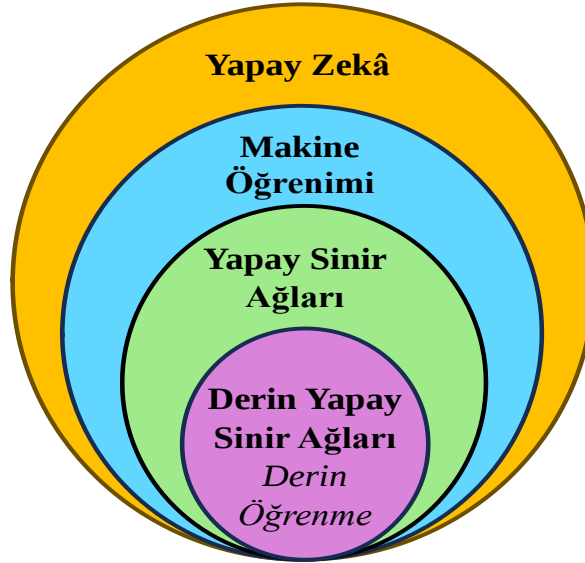
1.6.7. Yapay Zekâ

Yapay Zekâ (AI), insanın algılama yeteneklerini taklit etmeyi amaçlayan; matematiksel mantık, istatistik, olasılık, hesaplamalı nörobiyoloji ve bilgisayar bilimi dahil olmak üzere bir dizi teknikten oluşan bir disiplindir. Makine öğrenmesi ve derin öğrenme kavramlarını kapsayan geniş bir alan olan yapay zekâ sistemleri, kullanıcıların büyük veriyi işleyerek daha doğru ve etkili kararlar almalarına yardımcı olabilecek bilişsel düzenleyicilerdir (Coşkun & Gülleroğlu, 2021).

Yapay zekanın temelleri ve bu alanda atılan ilk adımlar 17. yüzyıla dayandırılmaktadır. Dönemin ünlü filozoflarından Descartes (1596-1650), insanı saat benzeri bir düzenekle çalışan makinelere benzetmiştir. İnsana ait birçok davranışın taklit edilmeye başlandığı bu dönemden sonra İngiliz matematikçi Charles Babbage insana ait fiziksel özelliklerin yerine bilişsel özelliklerinin taklit edilmesini hedeflemiştir.

Modern anlamda yapay zekâ çalışmalarının önem kazanması, İkinci Dünya Savaşı'nın hemen sonrasında bilgisayarların gelişimiyle yakından bağlantılı olarak yaşanmıştır. Yapay zekanın fikir babası Alan Turing olarak kabul edilse de “yapay zekâ” terimi ilk defa 1956 yılında Dartmouth Koleji'nde John McCarthy tarafından yapay zekâ üzerine düzenlenen bir çalıştayda kullanılmış ve bu etkinliğin ardından yapay zekâ çalışmaları hız kazanmıştır. 2010 yılı itibariyle de esas olarak bilgisayarların bilgi işlem gücündeki iyileşme ve büyük miktarda veriye erişim nedeniyle bu disiplinde ciddi bir yayılım yaşanmıştır (Coşkun & Gülleroğlu, 2021). Başlangıçta insanlar tarafından kodlanan kurallara yönelik yöntemlere dayanan yapay zekâ, basit mantıksal sorunları çözmek için uygun olsa da makinenin kendi başına öğrenme işlemini gerektiren; görüntü segmentasyonu, sınıflandırma, ses tanıma, yüz tanıma gibi daha kompleks sorunları çözme konusunda yetersiz kalmıştır (Chollet, 2021). Bu sorunları çözebilmek için yapay zekanın alt kümesi olan “Makine Öğrenimi” ve makine öğreniminin alt dalı olan “Yapay Sinir Ağları” ve “Derin öğrenme” isimli yaklaşımlar ortaya çıkmıştır (Şekil 8).

Özetle yapay zekâ; makineler tarafından taklit edilen insan zekâsı, makine öğrenimi; yapay zekaya ulaşmak için ortaya çıkan teknik bir yaklaşım, yapay sinir ağları; canlıların beyinde yer alan sinir ağlarının modellenmesi ile oluşturulan makine öğrenme metodu, derin öğrenme ise makine öğrenimini uygulamak için kullanılan bir tekniktir (URL-7, 2023; URL-8, 2023).

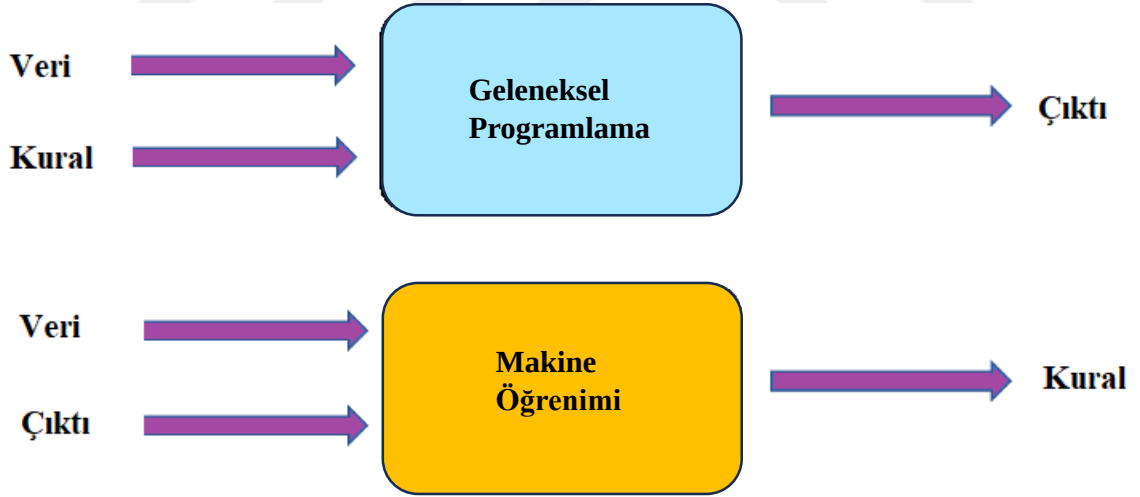


Şekil 8. Yapay zekâ alt dalları

1.6.8. Makine Öğrenimi

Makine öğrenimi, bir makinenin akıllı insan davranışını taklit etme yeteneği olarak tanımlanan yapay zekânın ve bilgisayar biliminin, insanların öğrenme şeklini taklit etmek için veri ve algoritmaların kullanımına odaklanan bir alt dalıdır. Makine öğrenimi 1950’lerde yapay zekânın öncülerinden biri olan Arthur Samuel tarafından “bilgisayarların programlanmadan öğrenme yeteneği” olarak tanımlanmıştır.

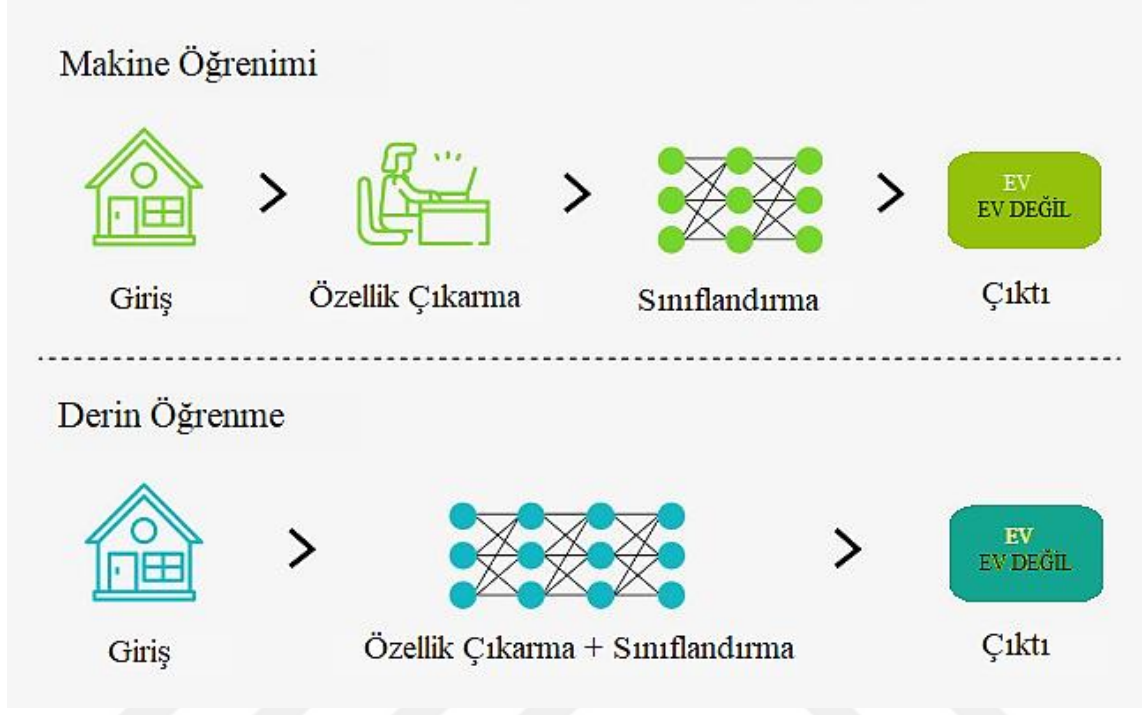
Tipik bir makine öğrenimi algoritması kabaca üç bileşenden oluşmaktadır. Bunlar; karar süreci, hata fonksiyonu, güncelleme veya optimizasyon sürecidir. Karar süreci alınan verilere göre algoritmanın ne tür bir model (kural) bulmak istediğini tahmin eden bir hesaplama sürecidir. Hata fonksiyonu yapılan tahminin ne kadar iyi olduğunu mevcut örneklerle karşılaştırarak ölçmeye yönelik bir yöntemdir. Güncelleme ve optimizasyon süreci ise algoritmanın gözden kaçırdığı ve ardından karar sürecinin nihai ulaştığı kararı güncellediği, böylece bir dahaki sefere kaybın azaltılmasının hedeflendiği bir yöntemdir. Makine öğrenme yaklaşımının, klasik yaklaşıma göre kısaca tanımlanması Şekil 9’da gösterilmiştir.



Şekil 9. Makine öğreniminin geleneksel yaklaşıma göre tanımı

Ham veriler üzerinde insan etkisinin varlığı veya yokluğuna göre sınıflandırılan makine öğrenimi modeli türleri mevcuttur. Bunlardan bazıları denetimli öğrenme, denetimsiz öğrenme, yarı denetimli öğrenme ve takviyeli öğrenmedir. Ve son olarak, insanların kuralları veya bilgileri eklemeyen, veri kümelerinden otomatik olarak öğrenme işlemi gerçekleştirebilen, makine öğreniminin daha yeni bir alanı olan derin öğrenme

kavramı mevcuttur (Şekil 10). Derin öğrenme için işlenmek üzere çok büyük miktarlarda ham veri gereklidir. Veri miktarı arttıkça tahmine dayalı model de o kadar iyileşme göstermektedir.



Şekil 10. Makine öğrenimi ve derin öğrenme farkı (URL-9, 2023)

1.6.9. Yapay Sinir Ağları

Yapay sinir ağları canlı organizmaların beyninden esinlenilerek geliştirilen, insan beyninin en önemli işlevleri olan; öğrenme ve öğrenme yolu ile yeni bilgiler türetebilme, oluşturabilme ve keşfedebilme gibi yetenekleri, herhangi bir insan müdahalesi olmadan otomatik olarak gerçekleştirebilme amacı ile geliştirilen bilgisayar sistemleridir.

Canlıların davranışlarını inceleyip, matematiksel olarak modelleme işlemi *sibernetik* olarak adlandırılmaktadır. Öğrenme işlemini sinir ağına girdi olarak sunulan örnekler yardımı ile gerçekleştirirler. Ağlar birbirine bağlı katmanlardan (yapay sinir hücrelerinden) oluşmaktadır ve bu bağlantılar yapay sinir ağının sahip olduğu bilgileri saklayan ağırlık değerlerine sahiptir.

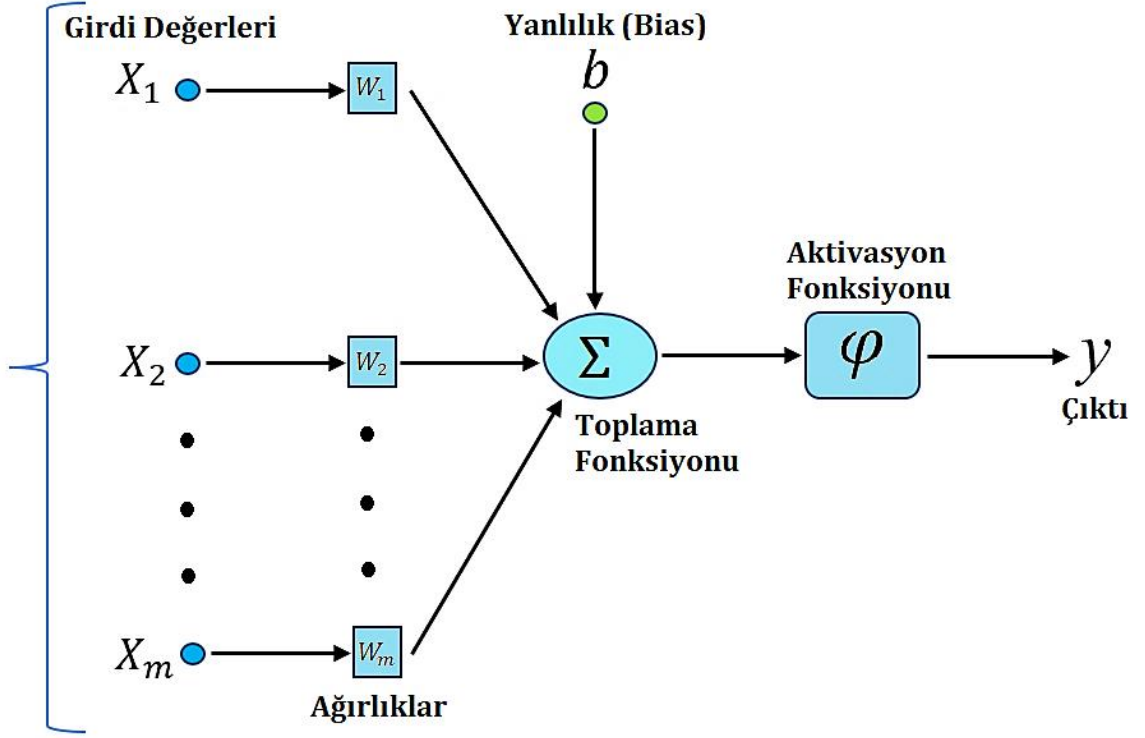
Yapay sinir ağları, biyolojik sinir sisteminden esinlenilerek geliştirilmiştir. Biyolojik sinir hücreleri sinapslar vasıtası ile birbirleriyle iletişim kurarlar. Bir sinir hücresi üzerinde işlem yaptığı bilgileri aksonlar vasıtasıyla diğer sinir hücrelerine gönderir. Bunun taklidi

şeklinde yapay sinir hücreleri dışarıdan gelen bilgileri bir toplama fonksiyonu ile toplar ve aktivasyon fonksiyonundan geçirerek çıktı bilgiyi üretip, diğer yapay sinir hücrelerine (işlemci elemanlarına) gönderir (Öztemel, 2012). Yapay sinir ağlarını birbirlerine bağlayan bağlantıların değerlerine ağırlık değerleri denmektedir. Biyolojik sinir sistemi elemanlarının yapay sinir ağlarındaki karşılıkları Tablo 2’de sunulmuştur.

Tablo 2. Yapay sinir ağı elemanlarının biyolojik sinir sistemindeki karşılıkları (Öztürk & Şahin, 2018).

Biyolojik Sinir Sistemi	Yapay Sinir Ağları
Nöron (Sinir Hücresi)	İşlemci Elemanı
Dentrit	Toplama Fonksiyonu
Hücre Gövdesi	Aktivasyon Fonksiyonu
Akson	Yapay Nöron Çıkışı
Sinapslar	Ağırlık Değerleri

İşlemci elemanları üç katman halinde bir araya gelerek bir ağ oluştururlar. Bunlar; girdi katman, ara katman ve çıktı katmandır. Bilgiler yapay sinir ağına girdi katman ile iletilir. Ara katmanlar, ağın ağırlık değerlerini kullanarak, ağa gelen bilgilerden bir sonuç üretir ve çıktı katmana gönderirler. Ağın girdiler için doğru sonuçları üretebilmesi hususunda ağırlık değerlerinin doğru olması gerekmektedir. Doğru ağırlık değerlerinin bulunması işlemi, ağın eğitilmesi olarak adlandırılır. Bu durum sağlandıktan sonra test setindeki örnekler eğitilen ağa sunulur ve ağ test setindeki örneklere doğru cevaplar verirse ağ eğitilmiş olarak kabul edilir (Öztemel, 2012). Bir yapay sinir ağını oluşturan ana bileşenler Şekil 11’de sunulmuştur.



Şekil 11. Yapay sinir ağlarının bileşenleri ve çalışma prensibi

Şekil 11’de Sembolik olarak X_m ’li ifadeler girdi değerlerinin; W_m ifadeler ise girdi değerlerine ait ağırlıkların temsilidir. Yanlılık parametresi (b) ise, çıktı bilginin oluşturulmasına katkı sağlayan bir etkidir. Girdi değerleri ağırlıklar ile çarpılıp, yanlılık (bias) değeri ile toplanarak (Σ) ağırlıklı toplamın elde edilmesi sağlanır. Son olarak ağırlıklı toplam, aktivasyon fonksiyonundan (φ) geçirilerek çıktı (y) elde edilir. Bahsi geçen işlemlerin matematiksel ifadesi Denklem 4’te sunulmuştur. En genel haliyle yapay sinir ağı modeli yapıları yapılan temel işlem; modelin en iyi doğruluğu vereceği ağırlık değerlerinin (W) ve yanlılık (b) parametresinin hesabını yapmaktır.

$$s = \Sigma(X_{1..m} * W_{1..m}) + b$$

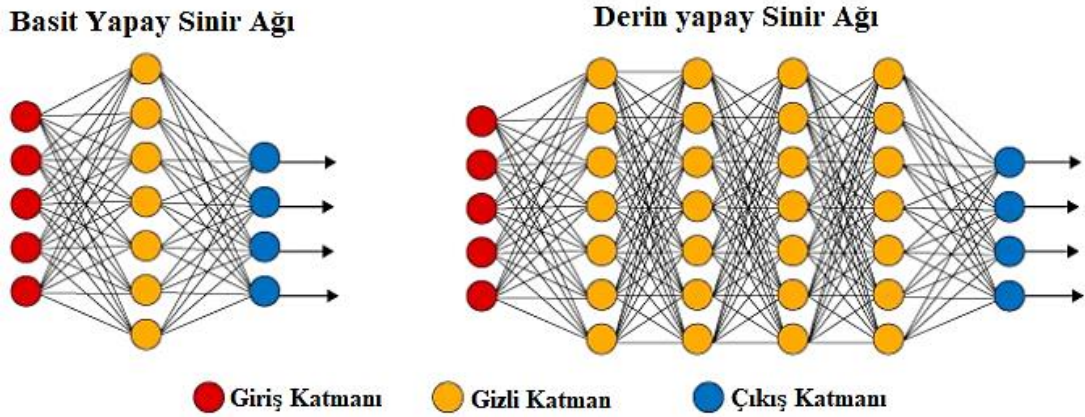
$$y = \varphi(s) \quad (4)$$

Bir yapay sinir ağı zekasının saklandığı ağırlıklar belirlendikten sonra her bir ağırlığın ne anlama geldiği bilinmemektedir. Bu sebeple yapay sinir ağlarına *kara kutu* yakıştırması yapılmaktadır.

1.6.10. Derin Öğrenme (Derin Yapay Sinir Ağları)

Derin öğrenme, verileri işlemede insan beyninin işleyiş sistemini taklit ederek çalışan görüntü tanımlama/sınıflandırma, ses tanımlama, tahmin yapma gibi görevleri yerine getiren makine öğrenim türüdür (Goodfellow vd., 2016). İnsan denetimi olmadan, girdi olarak verilen bir veri kümesi ile çıktıları tahmin eden yapay zekanın eğitilmesine olanak sağlar. Derin öğrenme, çok katmanlı yapay sinir ağlarının geri yayılım (backpropagation) (Rumelhart vd., 1986) adında bir algoritma ile eğitilmesine verilen isimdir. Art arda katmanlardan oluştuğu için bu modellere “Derin Yapay Sinir Ağları”, bu ağların geri yayılım algoritması ile eğitilmesine de “Derin Öğrenme” ismi verilmiştir.

Derin yapay sinir ağları; giriş katmanı, gizli katmanlar ve çıkış katmanı olmak üzere üç katmandan oluşmaktadır. Derin Öğrenme tabirindeki “Derin”, birden fazla gizli katmana sahip olmayı ifade eder (Şekil 12). Çözülen problemin niteliğine göre, gizli katmanların sayısı ile bu katmanlarda bulunan nöron sayıları değişebilmektedir (Ser & Bati, 2019).



Şekil 12. Basit yapay sinir ağı ve derin yapay sinir ağı karşılaştırma (URL-10, 2023)

Derin öğrenme mimarilerinin oluşturulma aşamasında kullanılacak algoritmalar büyük önem taşımaktadır. Bu algoritmalar verilerin çeşidine, boyutuna, yapısına, hacmine ve kullanılacak parametrelere göre farklılık göstermektedir. Algoritmalarından bazıları; Tekrarlayan Sinir Ağı/Recurrent Neural Network (RNN), Kısıtlı Boltzmann Makineleri/Restricted Boltzmann Machines (RBM), Derin İnanç Ağları/Deep Belief Nets (DBN) ve Evrişimli Sinir Ağları/Convolutional Neural Network (CNN) şeklindedir.

Tekrarlayan sinir ağları (RNN), düğümler arasındaki bağlantıların bir döngü oluşturabildiği ve bazı düğümlerden gelen çıktıların, aynı düğümler için sonraki girdileri

etkilemesine izin veren bir yapay sinir ağıları sınıfıdır. RNN'ler genelde bir sonraki adımı tahmin etmek için kullanılan derin öğrenme yapılarıdır. RNN uygulama alanlarının başında kelime tahmini gelmektedir. Bunun yanı sıra; konuşma tanıma, zaman serisi anomali tespiti, borsa tahmini, film altyazı tahmini uygulamalarında da kullanılmaktadır (Schmidt, 2019).

Kısıtlı Boltzmann Mankineleri sınıflandırma, regresyon ve özellik öğrenimi işlemlerini yapan giriş veri seti üzerinde olasılıksal dağılımları öğrenebilen bir sinir ağıdır. RBM'ler görünür ve gizli katmanlar olmak üzere aralarında simetrik bağlantı bulunan iki parçalı yapıya sahiptirler. Katmanlar içindeki düğümler hesaplamaların gerçekleştiği yerdir. Düğümler birbirine katmanlar halinde bağlanır, ancak aynı katmandaki iki düğüm birbirine bağlanmaz. Bundan dolayı katman içi iletişim yoktur. RBM'lerin kısıtlaması budur (Ackley vd., 1985).

Derin inanç ağıları, art arda eklenen Kısıtlı Boltzman Makineleri katmanlarından oluşan bir sinir ağı yaklaşımıdır. Kısıtlı boltzman makinelerinin sırasıyla eğitilerek öğrenilmesiyle gerçekleşir. DBN'ler görüntü tanımlama, hareket yakalama, konuşma tanıma gibi uygulamalarda kullanılmaktadır (Hinton vd., 2006).

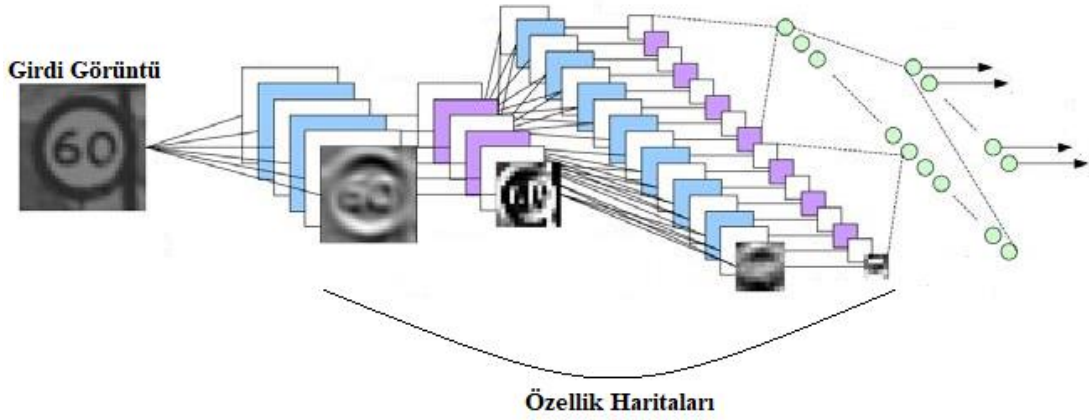
Evrişimli Sinir Ağları ise filtreler yardımıyla elde edilen özellik haritaları yoluyla sınıflandırma ve bilgisayarlı görme görevleri için kullanılmaktadır (O'Shea & Nash, 2015).

1.6.11. Evrişimli (Evrişimsel) Sinir Ağları

En popüler derin sinir ağılarından biri olan Evrişimli Sinir Ağları (Convolutional Neural Network-CNN), makine öğreniminin bir alt kümesidir. Bu ismi evrişim adı verilen matrisler arasındaki matematiksel işlemden almaktadır (Albawi vd., 2017). Evrişimli sinir ağıları, giriş verisi olarak görüntüleri kullanıp, filtreler vasıtasıyla görüntüdeki çeşitli nesnelere önem derecesi (ağırlık değeri ve yanlılık parametresi) atayan, detayların birbirinden ayırt edilmesi ve aralarındaki ilişkinin tespiti gibi görüntü özelliklerini çıkarabilen bir derin öğrenme algoritmasıdır. Genel anlamıyla evrişimli sinir ağıları görüntü işleme, sınıflandırma ve segmentasyonu için kullanılan bir araçtır (Chauhan vd., 2018).

Yapay sinir ağıları karmaşık veriler üzerinde başarı sağlarken, evrişimli sinir ağlarına ihtiyaç duyulmasının birtakım sebepleri mevcuttur. Bunlardan ilki; normal sinir ağlarındaki en büyük sorunun, ölçeklenebilirlik eksikliği olmasıdır. Daha az spektral banda sahip ve daha küçük görüntüler için normal bir sinir ağı tatmin edici sonuçlar üretebilir. Ancak görüntünün boyutu ve karmaşıklığı arttıkça, hesaplama gücü ve kaynaklara olan ihtiyaç

artmaktadır. Ek olarak görüntü verileri kullanıldığı için sinir ağının girdileri piksel değerleri halindedir. Böylece mimari bir görüntü için piksel değerlerini okur, özellik çıkarımı ve sınıflandırma yapar. Ancak bir görüntünün özellikleri sadece piksel değerleri değildir. Görüntüde bulunan kenarlar, çizgiler, köşeler veya temelde bir görüntünün şekilleri ve desenleri pikseller arasındaki ilişkilerin sağladığı görüntü özellikleridir. Yapay sinir ağı ile sınıflandırma yapılabilmesi için ilgili veriyi temsil eden özelliklerin yapay sinir ağına girdi olarak sunulması gerekir. Buna karşılık evrişimli sinir ağlarında, evrişim katmanı özellik seçimini otomatik olarak yapmaktadır. Sinir ağına, evrişim katmanı eklenir ve ağ artık evrişimli sinir ağı olarak adlandırılır. Evrişim katmanının amacı bir görüntüden özellik çıkarmaktır (Şekil 13). Ayrıca bu ağların en faydalı özelliği, yapay sinir ağlarındaki parametre sayısını azaltmasıdır (Albawi vd., 2017).



Şekil 13. Evrişimli sinir ağı ile görüntü özelliklerinin çıkarılması (Peemen vd., 2016)

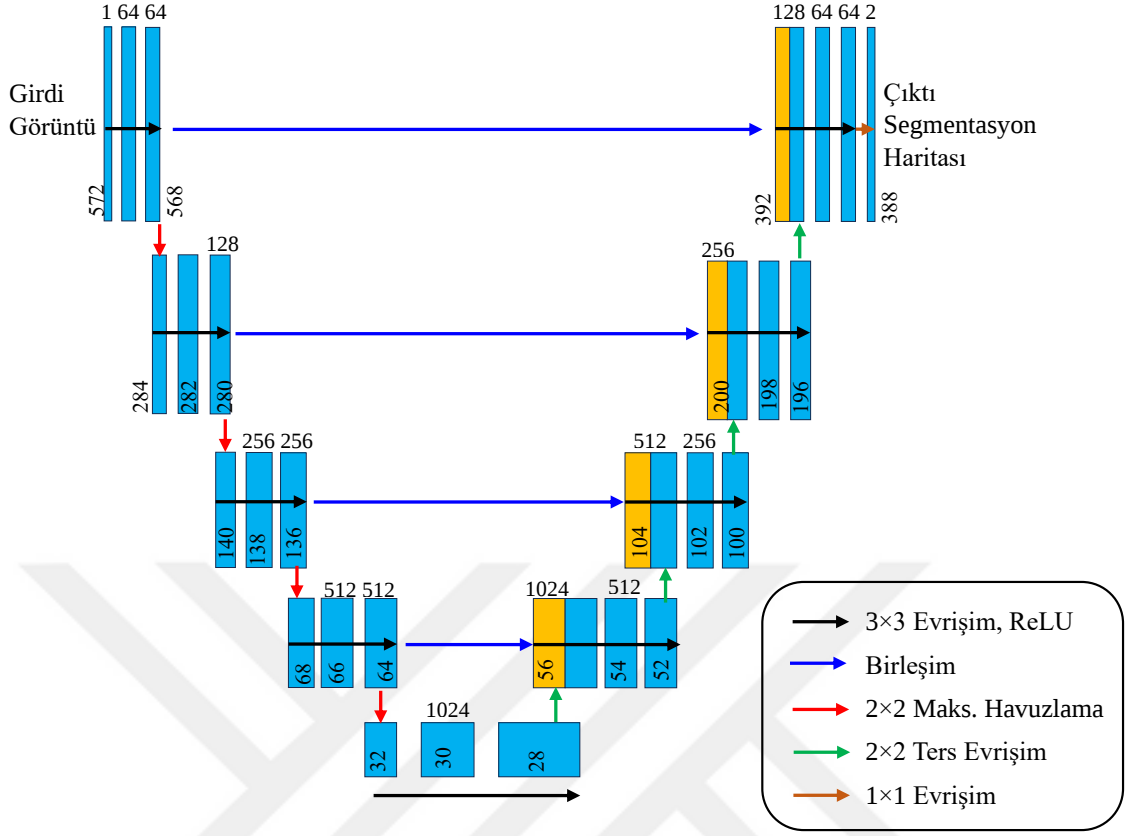
Evrişimli sinir ağları; evrişim katmanı, havuzlama katmanı ve tam bağlantılı katman olmak üzere temelde üç katmandan oluşur. Tez çalışması dahilinde kullanılan tüm katmanlar 1.7. başlığı altında detaylandırılmıştır.

1.6.12. U-Net Mimarisi

U-Net mimarisi, özellikle tıbbi görüntü segmentasyonu için geliştirilmiş bir derin öğrenme ağıdır (Ronneberger vd., 2015). Bir tür evrişimli sinir ağı yaklaşımı olan U-Net mimarisi ilk olarak Freiburg Üniversitesi Bilgisayar Bilimleri Bölümünde biyomedikal görüntü segmentasyonu için geliştirilmiştir. U-Net derin öğrenme ağının geliştirilmesinde Long, Shelhamer ve Darrell (2015) tarafından Tamamen Evrişimli Ağların (Fully

Convolutional Network (FCN)) (Long vd., 2015) önerilmesi önemli bir kilometre taşıdır. FCN mimarisi her pikseli etiketlemek için, bir kodlayıcı/kod çözücü yapısı kullanmakta olduğundan ve tamamen bağlı katmanlar (fully connected layer) içermediğinden, herhangi bir boyuttaki görüntüler girdi olarak seçilebilmektedir. Ronneberger vd. (2015) FCN'ye dayanarak U-Net mimarisini önermişlerdir. Bahsedilen iki mimari arasındaki temel farklardan biri, U-Net'in daha iyi sonuçlar elde etmek için yapısında birleştirme katmanları bulunmasıdır. Ayrıca, U-Net, daha az eğitim verisi ile hassas segmentasyon elde edilebilecek şekilde geliştirilmiştir (Bardis vd., 2020).

U-Net mimarisi; 23 katman ve bir köprü aracılığıyla bağlanan, kodlayıcı-kod çözücü olarak tanımlanmış iki ana bölümden oluşan U şeklinde bir mimaridir (Şekil 14). Mimarinin sol yarısı olan ve dört adet kodlayıcı bloktan oluşan bölümde görüntünün özellik haritaları elde edilir kısacası görüntünün “Ne” olduğu öğrenilir. Sağ tarafında ise dört adet kod çözücü blok mevcuttur ve bu aşamada görüntü içerisinde bulunan özelliklerin “Nerede” olduğu öğrenilir. Kodlayıcı bölüm, ağ parametrelerinin sayısını azaltmak amacıyla görüntü boyutlarını yarıya indiren ve filtre sayısını iki katına çıkaran, evrişim ve maksimum havuzlama katmanlarının bir birleşimidir. Kod çözücü aşamada ise ters evrişim katmanları kullanılarak fotoğrafların boyutları iki katına çıkarılır ve filtre sayısı yarıya indirilir. Bu sayede görüntülerde boyut arttırma ile simetrik genişleme işlemi gerçekleştirilerek girdi ve çıktı görüntü çözünürlüklerinin/boyutlarının aynı olması sağlanmaktadır.



Şekil 14. U-Net mimarisi

Orijinal U-Net mimarisi ile gerçekleştirilen işlemler maddeler halinde aşağıda sunulmaktadır.

Kodlayıcı;

- Girdi görüntü boyutları 572×572 olarak alınmıştır.
- Evrişim katmanı ile 64 adet filtre uygulanır (Boyut: 568×568×64).
- Maksimum havuzlama katmanı ile 2×2 matris kullanılarak fotoğraflardaki en büyük piksel değerleri kullanılır (Boyut: 284×284×64).
- Evrişim katmanı ile 128 adet filtre daha uygulanır (Boyut: 280×280×128).
- Bu işlemler iki kez daha devam ettikten sonra 64×64×512 boyutlu fotoğraflar oluşur.

Darboğaz;

- Kodlayıcı ve Kod çözücü blokları birleştirilirken oluşan elips yapıya “Darboğaz” (Bottleneck) denir.
- Darboğazda maksimum havuzlama ile oluşan 32×32×512 boyutlu fotoğraflara iki kez evrişim katmanı uygulanır (Boyut: 28×28×1024).

Kod çözücü;

- Ters evrişim katmanı ile fotoğraflara 2×2 matrisler uygulanır (Boyut: $56 \times 56 \times 1024$).
- Kod çözücü bloğunun her bir katmanında, öncesinde o katmanla aynı hizada olan kodlayıcı bloğundaki katman birleştirilir ve bu işleme “Birleştirme” (Concatenate) denir.
- Evrişim katmanı ile 512 adet filtre uygulanır (Boyut: $52 \times 52 \times 512$).
- Dört kez kod çözücü bloğu gerçekleştirildikten sonra $388 \times 388 \times 2$ (2 segmentasyondaki sınıf sayısıdır) boyutlu çıktı görüntüler elde edilir (URL-11, 2023).

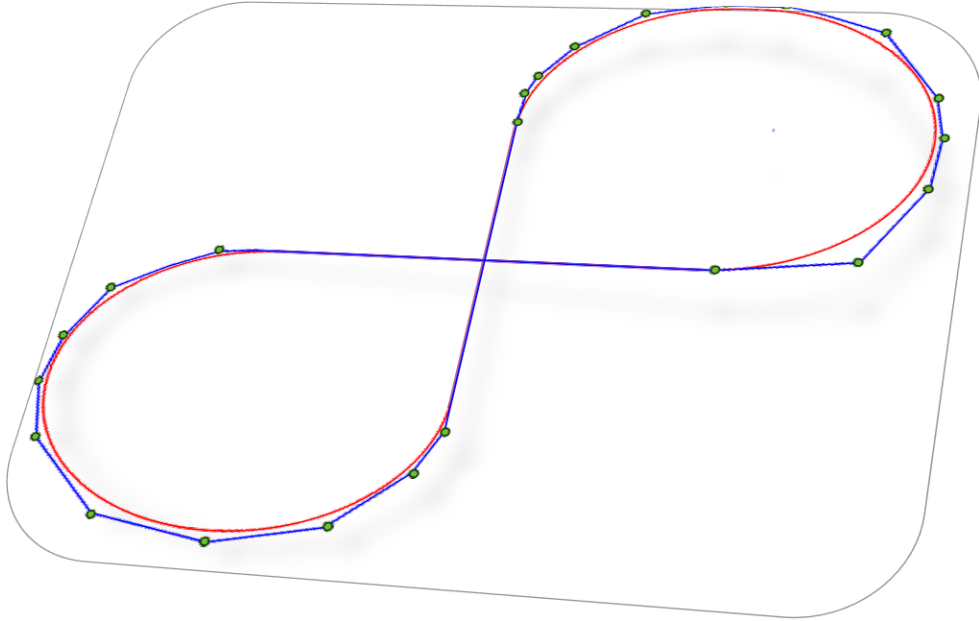
Klasik bir oto kodlayıcı mimarisinde, U-Net ağında olduğu gibi ilerleyen katmanlar boyunca, girdi görüntülerin boyutu azaltılır. Ardından kod çözücü kısımda doğrusal öznitelik gösterimi öğrenilir ve boyut giderek artar. Mimarinin sonunda çıkış boyutu ile giriş boyutu eşit olur ancak tüm özniteliklerin iletilemediği bir durum meydana gelir. Bu noktada U-Net mimarisinin avantajı ortaya çıkar. U-Net, kod çözücü (ikinci yarı) aşamasında ters evrişim işlemi gerçekleştirir ve buna ek olarak mimarinin kodlayıcı tarafından gelen bağlantılar sayesinde özniteliklerin kaybolmasına sebep olan ‘darboğaz’ probleminin üstesinden gelebilmektedir (URL-12, 2023). Bu bilgilere dayanarak mimarinin asıl önemli kısmının bağlantı katmanları aracılığıyla, kodlayıcı blokların, kod çözücü bloklar ile evrişim içinde olmasında saklı olduğu söylenebilmektedir.

1.6.13. Pürüzsüz Karıştırma (Smooth Blending) Yöntemi

Pürüzsüz karıştırma yöntemi, görüntüler ile eğitilen derin öğrenme mimarilerinin test edilmesi aşamasında kullanılan bir tekniktir. Kaydedilen modelin, mimarinin eğitimi sırasında kullanılan görüntü boyutundan daha büyük görüntüler üzerinde uygulanması aşamasında çıktı görüntüde süreklilik sağlanması amacıyla tercih edilen bir yöntemdir. Bu yöntemle ihtiyaç duyulmasının en önemli sebebi; test aşamasında büyük görüntüler üzerinde tahmin yapılırken, mimarinin eğitildiği görüntü boyutlarına parçalanarak parçalanmış görüntüler üzerinde tahmin işlemi gerçekleştirilip, ardından görüntü parçalarının birleştirilmesidir. Örneğin bir mimari 256×256 boyutunda görüntüler ile eğitilmiş ve test aşamasında 1000×1000 boyutlu görüntülerin kullanımı söz konusu ise, öncelikle 1000×1000 boyutlu görüntüler 256×256 ’lık parçalara ayrılarak yama görüntüler üzerinde tahmin

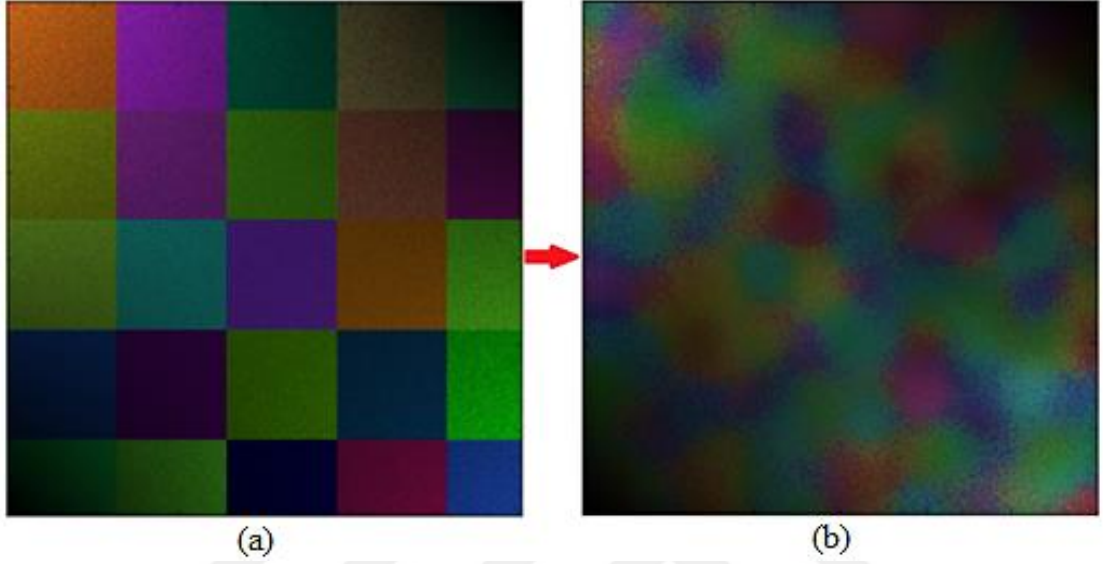
gerçekleştirilir. Bu aşamada pürüzsüz karıştırma yöntemi kullanılmadığı takdirde çıktı görüntü 256×256 'nin 1000×1000 'den küçük maksimum katı olan 768×768 boyutlarında olup, görüntü içerisinde süreksizlik durumları ve boyuttan kaynaklı bilgi kayıpları mevcut olacaktır. Bu durumun ana sebebi; derin öğrenme modellerinin, görüntü yamaları olan bu küçük yerel pencereler üzerinde, yamaların sınırına yakın veriler olmadan tahminlerde bulunmasıdır. Bu nedenle, pencerelerin dışına yakın yerlerde yapılan tahminlerde yüksek hatalar mevcut olabilmektedir. Bu hatalara sahip tahmin yamalarının direkt olarak birleştirilmesi süreksizliğin olduğu bir sonuç görüntüye sebep olmaktadır. Sonuç olarak test aşamasında, hem büyük boyutlu görüntünün tamamı üzerinde tahmin yapılamayıp hem de tahmin görüntüsü üzerinde süreksizlikler mevcut olacaktır (Şekil 16a). Bu noktada bahsi geçen sorunların çözümü için pürüzsüz karıştırma yönteminin kullanımı gerekmektedir.

Algoritma girdi olarak ana görüntüyü ve boyutunu, pencerelerin boyutunu ve bu pencereler için istenilen örtüşme miktarını kullanıcı girişi olarak alır. Ardından yapılan tahminleri bir araya getirmek için basit ikinci dereceden spline eğrilerinin pencereler üzerindeki işleviyle enterpolasyon uygulanır. Spline eğrileri parçalar arasında süreklilik şartını sağlayan bileşik eğrilerdir (Şekil 15).



Şekil 15. Spline eğrileri için süreklilik örneği

Birleştirilmiş pencereler arasındaki örtüşme miktarı, istenilen sonucun alınması durumuna göre kullanıcı tarafından değiştirilebilmektedir. Pürüzsüz karıştırma işlemine ait örnek bir görsel Şekil 16b’de sunulmaktadır.



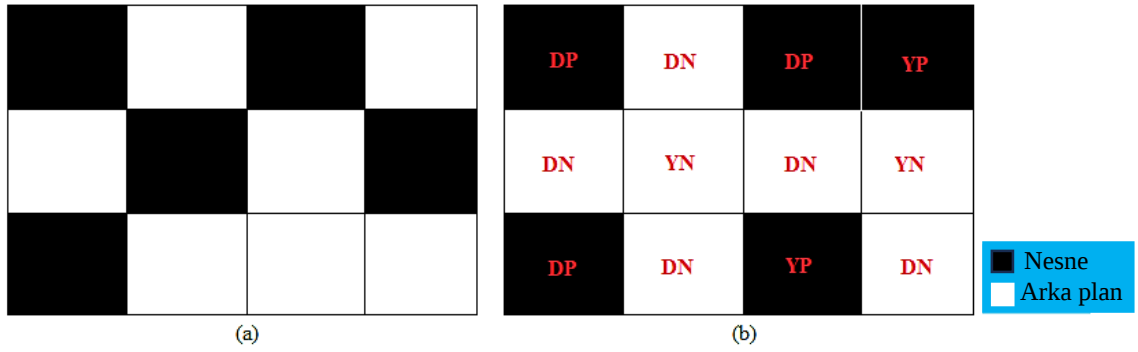
Şekil 16. Pürüzsüz karıştırma yöntemi (a) Görüntüde süreksizlik örneği, (b) Pürüzsüz karıştırma yöntemi ile görüntüdeki süreksizliğin ortadan kaldırılması (URL-13, 2023)

Pürüzsüz karıştırma yönteminin işlevine dair Şekil 16 görseli incelendiğinde Şekil 16a’da bulunan görüntü süreksizliği; eğitilen bir mimarinin, test görüntüsü üzerinde yaptığı tahminlerin birleştirilmesi sonucu elde edilen büyük görüntüdeki süreksiz görünüme örnek olarak gösterilebilir. Bu negatif durumun önüne geçmek adına enterpolasyon işlemi temel alan pürüzsüz karıştırma yönteminin uygulanması sonucunda Şekil 16b örneğinde olduğu gibi süreksiz görünüm giderilebilmektedir.

1.6.14. Doğruluk Değerlendirmesi

Uzaktan algılanmış görüntüler üzerinde sınıflandırma ve segmentasyon gibi işlemler gerçekleştirildiğinde üretilen sonuç sınıflandırma ya da segmentasyon haritalarının ne derecede doğru olduğu sorusuna cevap bulmak, işlemin gerçekleştirildiği yöntemin güvenilirliğinin anlaşılması açısından önem arz etmektedir. Doğruluk analizi, üretilen sonuç ile referans verinin karşılaştırılması üzerine elde edilen başarı düzeyi olarak

tanımlanmaktadır. Literatürde raster verilerin kullanımı ile geliştirilen tekniklerin güvenilirliğinin test edilebilmesi için en yaygın kullanılan yöntem piksel bazlı doğruluk analizidir. Piksel bazlı doğruluk analizi yönteminin gerçekleştirilmesi için ilk şart üzerinde doğruluk metriklerinin hesaplanacağı sonuç görüntüsüne karşılık gelen, aynı boyutlarda bir referans veri olmasıdır. Sonuç ve referans görüntülerinin pikselleri arasında birebir karşılaştırma işlemi gerçekleştirilmelidir. Bu karşılaştırma sonucunda tahmin görsellerinin piksellerine Doğru Pozitif (DP), Doğru Negatif (DN), Yanlış Pozitif (YP) ve Yanlış Negatif (YN) olmak üzere dört adet etiket atanmaktadır. DP, görüntüdeki bir objenin cinsinin ve konumunun, DN ise obje olmayan bir bölgenin doğru olarak tahmin edildiği pikselleri temsil etmektedir. YP obje olmayan bir bölgede hatalı bir biçimde obje tespit etme durumu iken, YN ise obje olan bir bölgede objeyi tespit edememe durumudur (Şekil 17). Yapılan bir sınıflandırma veya modelinin sonuçlarını yorumlayabilmek, gerçek ve tahmin edilen değerler arasında yer alan ilişkideki hataları inceleyebilmek için içerisinde bahsi geçen etiket bilgilerinin yer aldığı karışıklık matrisi (confusion matrix) (Tablo 3) kullanılır.



Şekil 17. DP, YP, YN ve DN metriklerinin belirlenmesi (a) Referans Pikseller, (b) Tahmin Edilen Pikseller

Tablo 3. Karışıklık matrisi

Karışıklık Matrisi		Referans Veri Sonuçları		Toplam
		Nesne	Arka plan	
Tahmin Sonuçları	Nesne	DP (Doğru Pozitif)	YP (Yanlış Pozitif)	Pozitif Piksel Sayısı
	Arka Plan	YN (Yanlış Negatif)	DN (Doğru Negatif)	Negatif Piksel Sayısı

Tablo 3’te bir adet nesne ve arka plandan oluşan iki sınıflı bir karışıklık matrisi örneği gösterilmektedir. Bir görüntü verisine ait karışıklık matrisinde bulunan tüm DP, YP, YN ve DN sayı değerlerinin toplamı o görüntüdeki piksel sayısını vermek zorundadır.

Karışıklık matrisinde bulunan değerler yardımıyla birçok doğruluk metriği hesaplanabilmektedir. Bunlardan bazıları; Genel Doğruluk (Overall Accuracy/OA), Doğruluk (Correctness/Cr), Tamlık (Completeness/Cp), F1-Puanı (F1-Score), Birleşimlerin Kesişimi (Intersection Over Union/IoU) ve Kalite (Quality/Q) doğruluk metrikleri olarak literatürde mevcuttur (Denklem 5-10).

$$OA = \frac{DP + DN}{DP + YP + DN + YN} \quad (5)$$

$$Cr = \frac{DP}{DP + YP} \quad (6)$$

$$Cp = \frac{DP}{DP + YN} \quad (7)$$

$$IoU = \frac{DP}{DP + YP + YN} \quad (8)$$

$$Q = \frac{DP}{DP + YP + YN} \quad (9)$$

$$F1 = 2 \times \frac{Cr \times Cp}{Cr + Cp} \quad (10)$$

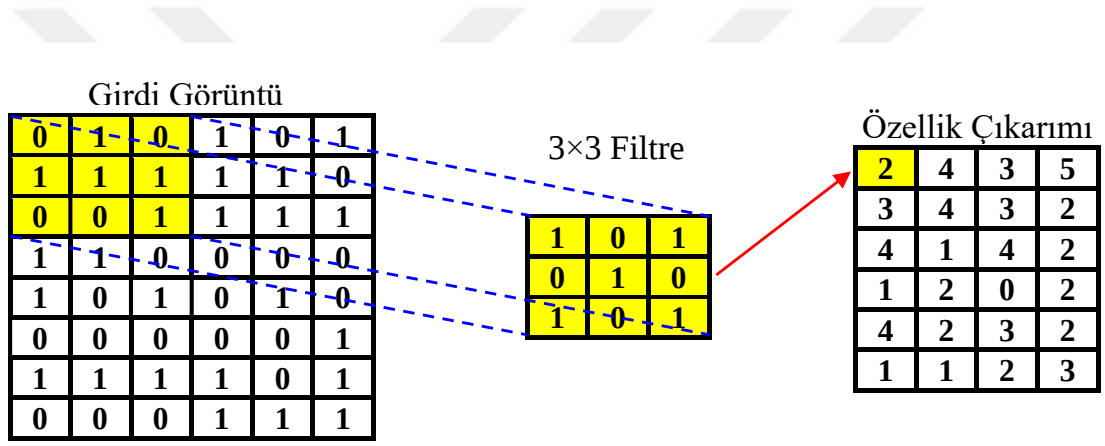
İlgili metriklerle ait piksel adetleri ilgili denklemlere yazılarak doğruluk oranları hesaplanmaktadır.

1.7. Derin Öğrenmenin Temel Kavramları

1.7.1. Evrişim Katmanı

Evrişim katmanı, bir evrişimli sinir ağının en önemli yapı taşıdır. Bir ağın eğitimi boyunca bir dizi filtre (çekirdek) ile parametrelerinin öğrenilmesi işlemini gerçekleştirir. Her

filtre görüntü üzerine uygulanarak özellik haritaları oluşturulur. Evrişim katmanındaki filtre boyutları, girdi görüntü boyutuna göre seçilir ve boyutları genellikle 3×3 , 5×5 ve 7×7 şeklinde olmaktadır. İlk evrişim katmanları görüntülerin düşük seviyeli özelliklerini (çizgiler) yakalarken, daha sonraki katmanlar yüksek seviyeli özelliklerin (şekiller ve detaylar) çıkarımını sağlar (Mostafa & Wu, 2021). Evrişim katmanları bir, iki ve üç boyutlu olabilmektedirler. Literatürde kullanılan en yaygın evrişim türü 2B evrişim katmanıdır ve genellikle *Conv2D* olarak kısaltılır. Evrişim katmanındaki bir filtre, giriş görüntüsü üzerinde adım adım kayar ve filtrelerinin ağırlıkları ile giriş görüntüsünün değeri arasında çarpma işlemi gerçekleştirir (Şekil 18), Bu durum özellik haritasının üretilmesiyle sonuçlanır (Singh vd., 2020).

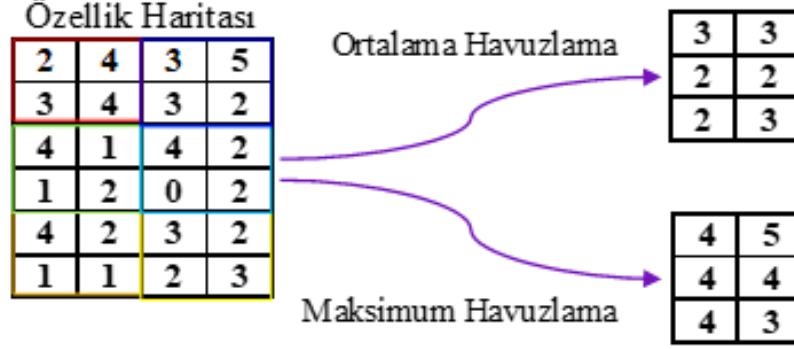


Şekil 18. Evrişim katmanının görüntüler üzerindeki işlemi

1.7.2. Havuzlama Katmanı

Havuzlama katmanında öğrenilen herhangi bir parametre yoktur. Bu katmanda evrişim katmanından çıkan özellik haritası, giriş verisi olarak alınarak bu verinin yükseklik ve genişlik bilgisi indirgenir. Bu katman sayesinde hesaplama karmaşıklığı azaltılır. Maksimum ve ortalama olmak üzere iki adet havuzlama yöntemi mevcuttur ancak genellikle maksimum havuzlama yöntemi daha sık kullanılmaktadır. Maksimum ve ortalama havuzlama işlemini yapmak için öncelikle adım boyutu belirlenir. Ardından görüldüğü haliyle maksimum havuzlama için giriş görseli üzerinde belirlenen adım boyutunun içerdiği piksel değerleri üzerinde sırayla ilerlenerek, mevcut grup içinde kalan piksellerden maksimum renk değerine sahip olan piksel seçilir. Ortalama havuzlamada ise grup içinde

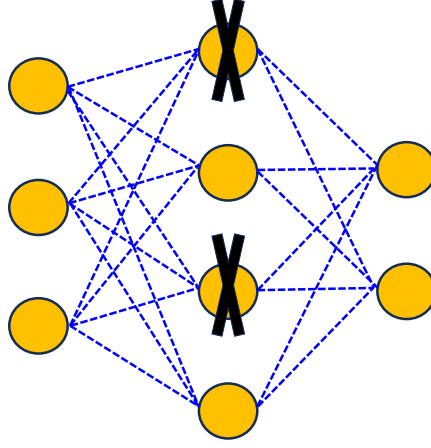
kalan piksellerin renk değerlerinin ortalaması alınarak tek piksele indirgenir. Bu sayede girdi görüntü boyutunun indirgenmesi sağlanmış olur (Şekil 19).



Şekil 19. Ortalama ve maksimum havuzlama katmanları

1.7.3. Seyreltme Katmanı

Seyreltme (Dropout) katmanı, sinir ağlarında ezberlemeyi (overfitting-aşırı öğrenme) önlemek için kullanılan bir katmandır. Aşırı öğrenme durumu genellikle mimarinin, kullanılan veriler için fazla karmaşık bir yapıya sahip olması durumunda meydana gelir. Sorun, eğitilen mimarinin test verileri üzerinde çok zayıf genelleme yapması, yani yeni veriler için yetersiz tahminler sağlamasıdır. Öte yandan eğitim veri setindeki performansın iyi olduğu, bu nedenle yüksek bir model kalitesi varmış gibi sonuçlar vermesidir. Bu durumu önlemek için seyreltme katmanı kullanımı ile önceki katmandan çıkan belirli sonuçlar sıfır değerine ayarlanarak ağdan çıkarılır (Şekil 20). Bu sayede derin sinir ağlarında, nöronlar kendi aralarında sözde bir adaptasyon oluşturmadıkları, verideki daha derin yapıları tanıdıkları için aşırı öğrenme tehlikesi azaltılır ve her aşamada yeni, biraz değiştirilmiş bir ağ mimarisi oluşturulur. Böylece ağ belirli girdiler olmadan iyi tahminler üretmeyi öğrenir. Ayrıca seyreltme katmanı kullanılırken, seyreltme oranının (dropout rate) yani bir katmandaki çıktı verilerin yüzde kaçının bir sonraki katmana aktarılmayacağını belirtilmesi gerekmektedir.



Şekil 20. Seyreltme katmanı işlevi

1.7.4. Toplu Normalleştirme Katmanı

Bir derin öğrenme ağı eğitimi sırasında geri yayılım yolu ile her bir katman kendi hatasını (ağırlıklarını ve önyargılarını) düzeltmeye çalışır. Katmanlar birbirlerinin çıktıları ile beslendikleri için bir katmandaki parametre değişikliği sonraki katmanların girdilerinin dağılımında değişime yol açmaktadır. Bu durum literatürde, Ortak Değişken Kayması (Internal Covariate Shift) olarak tanımlanmaktadır (Ioffe & Szegedy, 2015). Her bir katmandaki bu kaymalar nedeniyle hesaplamalar için zaman gittikçe artar. Çünkü bir katman kendi işlemine başlamak için kendinden önceki katmanın işleminin bitmesini beklemek durumundadır. Bu kaymaları azaltmak için toplu normalleştirme uygulanır ve bu işlem sayesinde ağıdaki katmanlar, önceki katmanın öğrenmesini beklemek zorunda kalmadan eş zamanlı öğrenime olanak sağlar. Bu avantajlı durum ise eğitimin hızlanmasını sağlamaktadır.

Toplu normalleştirme katmanını bir derin öğrenme ağına eklemenin gerekliliğinin ve bu işlemin çalışma prensibinin anlaşılması için öncelikle normalleştirme bir diğer adıyla ölçekleme kavramının anlamı ve kullanım sebebi anlaşılmalıdır. Normalleştirme (ölçekleme), verileri standartlaştırmak için kullanılan bir ön işleme tekniğidir. Bunun anlamı, aynı değer aralıklarında farklı veri kaynaklarına sahip olmayı sağlamaktır. Bu duruma, görüntü verilerinin bulunduğu bir veri setinde her bir görüntünün sahip olduğu renk değerlerinin, aynı sayı değerleri aralığına (0-1) ayarlanması işlemi örnek olarak gösterilebilir. Kullanılan verileri eğitimden önce normalleştirmemek derin öğrenme ağında; bilgi kaybı, öğrenme hızının yavaşlaması ile eğitimin zorlaşması gibi sorunlara neden

olmaktadır. Verileri normalleştirmek için kullanılan iki ana yöntem mevcuttur. Yaygın olarak kullanılan, en basit yöntem Denklem 11'deki formül ile ifade edilmektedir.

$$X_n = \frac{X - m}{X_{max} - X_{min}} \quad (11)$$

Denklem 11'de X ifadesi veri kümesinde bulunan işlenmemiş bir değeri, m veri kümesinin ortalamasını, X_{max} ve X_{min} veri kümesinde bulunan maksimum ve minimum değerleri, son olarak X_n ifadesi normalleştirilmiş X değerini temsil etmektedir.

Verileri normalleştirmek için kullanılan diğer bir teknik ise Denklem 12'deki eşitlik ile sağlanmaktadır.

$$X_n = \frac{X - m}{s} \quad (12)$$

Denklem 12'de de aynı şekilde X_n , normalleştirilmiş X değerini, X , veri kümesinde bulunan işlenmemiş bir değeri, m , veri kümesinin ortalamasını temsil ederken, s ifadesi ise veri kümesinin standart sapma değerini göstermektedir. Bu işlemin veri kümesine uygulanması halinde her veri noktası standart bir normal dağılımı taklit edecektir. Ayrıca her iki denklemde bulunan işlemlerden birinin ölçeklemede kullanılmak üzere tercih edilmesi halinde veri kümesindeki sayı değerlerinin 0 ve 1 değerleri arasına sıkıştırılarak, normalleştirilmiş olması sağlanacaktır.

Toplu normalleştirme ise, ham veriler yerine derin öğrenme ağının katmanları arasında yapılan bir normalleştirme tekniğidir. Bütün bir veri seti yerine, gruplar üzerinde uygulanır. Normalleştirme tekniği takip edilerek toplu normalleştirme formülü Denklem 13 ile tanımlanmaktadır.

$$Z^n = \frac{Z - m_Z}{s_Z} \quad (13)$$

Denklemde bulunan Z ; bir katmandan çıkan nöron değeri, Z^n ; işleme tabi tutulan nöronun normalleştirilmiş değeri, m_Z ; ilgili katmanın nöron değerlerinin ortalaması ve s_Z ise standart sapmasıdır (URL-14, 2023).

Toplu normalleştirme işlemi eğitim sürecini hızlandırmak adına daha yüksek öğrenme oranları kullanımına olanak sağlamakla birlikte evrişim katmanlarının ardından kullanılıp, katmanın ağırlıklarını normalleştirerek aşırı öğrenmenin azaltılmasına ve modelin genelleştirilebilme yeteneğinin geliştirilmesine yardımcı olabilmektedir.

1.7.5. Birleştirme Katmanı

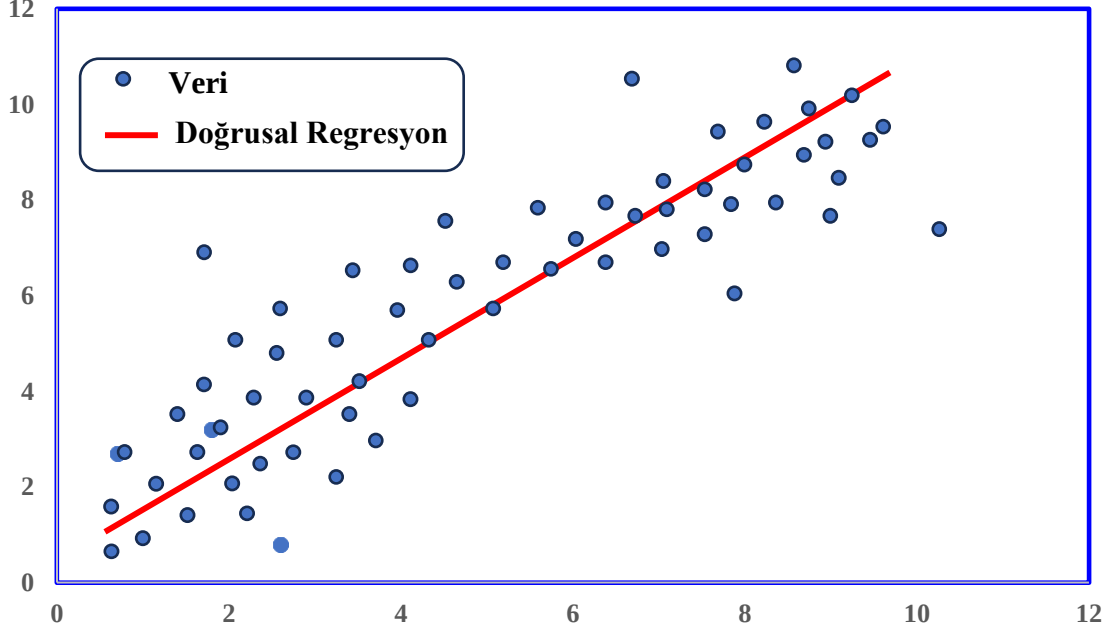
Evrişimli sinir ağlarının temel bileşenlerinden biri, özellikleri çıkarmak için giriş görüntüsüne bir dizi filtre uygulayan evrişim katmanıdır. Ancak çoğu durumda filtreler ile elde edilen özellikler, bir görüntünün doğru şekilde sınıflandırılması için yeterli olmamakla birlikte, ağı ilerleyen süreçlerinde başlangıçta öğrenilen özniteliklerin kaybolması problemi ortaya çıkarabilmektedir. Bu durumda, paralel tarzda eğitilen ağlar ile doğruluğu artırmak için birden fazla evrişim katmanının sonucu olan özellik haritalarını birleştirerek eşzamanlı bir şekilde ağı başka bir katmanına girdi verisi olarak sunmak amacıyla birleştirme katmanı kullanılmaktadır (Noreen vd, 2020). Birleştirme katmanı, her iki ağı boyutlarının aynı olması gerektiği kuralına sahiptir.

Bir derin öğrenme ağında ilgili katmanın kullanımının avantaj ve dezavantajı mevcuttur. Avantajlı yönü, ağ tarafından daha fazla parametre öğrenildiğinde sonuçların iyileştirilebilmesidir. Dezavantajı ise bir ağı etkili bir şekilde performans gösterememesi durumunda sonuçların tamamen etkilenmesi olabilmektedir. Ancak bu prosedürün literatürde birçok durumda etkili ve başarılı olduğu kanıtlanmıştır (Gupta vd., 2022).

1.7.6. Aktivasyon Fonksiyonları

Derin sinir ağı mimarilerinde bilginin bir sonraki nörona aktarılıp aktarılmayacağına karar vermek ve öğrenme işlemi gerçekleştirebilmek için aktivasyon fonksiyonlarına ihtiyaç duyulmaktadır. Derin öğrenme algoritmalarında öğrenme sürecinin sağlıklı bir şekilde gerçekleştirilmesi, aşırı öğrenmenin (ezberleme) önlenmesi, doğruluğun artırılması, hesaplama maliyetinin azaltılması gibi özellikler dikkate alınarak aktivasyon fonksiyonları geliştirilmiştir. Aktivasyon fonksiyonları olmadan çalışan bir sinir ağında kaç katman olduğunun bir önemi kalmamaktadır çünkü her nöron, ağırlıkları ve yanlılıkları kullanarak girdiler üzerinde yalnızca doğrusal bir dönüşüm gerçekleştirecektir. Bunun anlamı tüm katmanların aynı şekilde davranarak modelin yalnızca bir doğrusal regresyondan ibaret

olmasıdır (Şekil 21). Bu durumda mimarinin herhangi karmaşık bir görevi öğrenmesi imkânsız olacaktır.

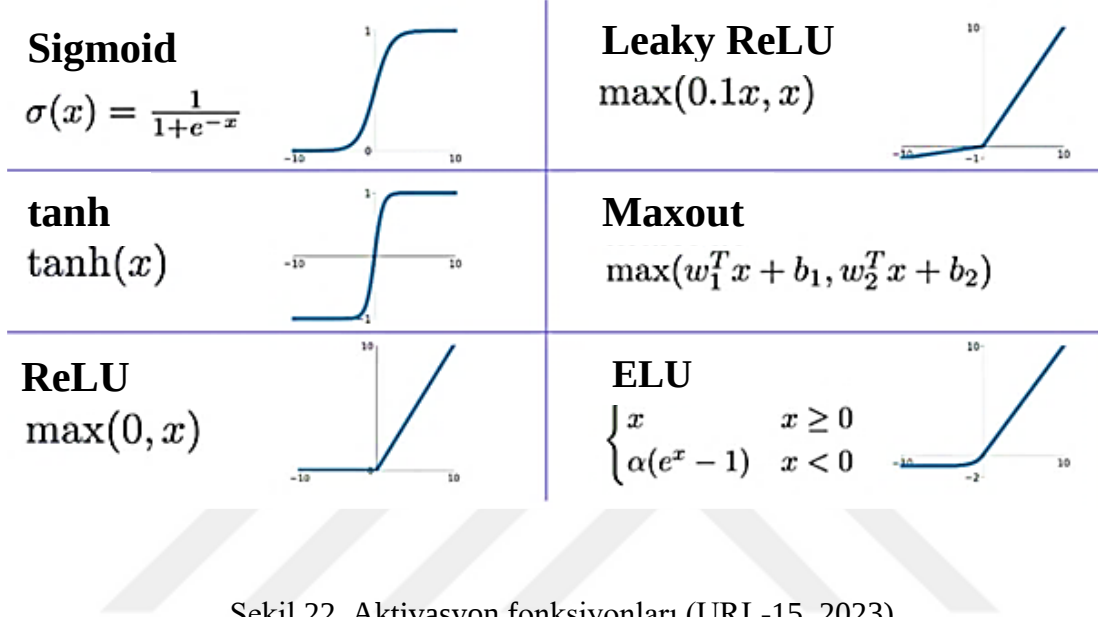


Şekil 21. Doğrusal regresyon

Aktivasyon fonksiyonları, derin öğrenme modellerinin daha iyi çalışmasını ve öğrenmesini sağlamaktadır fakat ilgili derin öğrenme modelleri için uygun aktivasyon fonksiyonları seçilmeden eğitim işlemi başlatılırsa sınırlı öğrenme sebebiyle istenilen başarı elde edilemez (Kılıçarslan vd., 2021).

Aktivasyon fonksiyonları sabit (doğrusal) ve eğitilebilir (doğrusal olmayan) olmak üzere iki şekilde sınıflandırılmaktadır (Scardapane vd., 2019). Derin öğrenme mimarilerinde nöronlar arasındaki öğrenme sürecini gerçekleştirmek için geriye yayılım algoritması kullanılmaktadır. Geri yayılım algoritması mimarilerdeki parametrelerin türevlerini alarak ağırlık ve yanlılıkların güncellenmesi işlemini gerçekleştirir. Bu bilgiye dayanarak aktivasyon fonksiyonlarının türevlenebilir olması önemlidir. Doğrusal aktivasyon fonksiyonlarında geriye yayılım algoritması yardımıyla türev gerçekleştirilirken sabit bir değer döndürülür. Bunun anlamı doğrusal aktivasyon fonksiyonlarının türevlenebilir olmaması ve mimari üzerinde öğrenme işlemini gerçekleştiremeyeceğidir. Bu sebeple derin sinir ağlarında, sürekli türevlenebilir özelliğine sahip doğrusal olmayan aktivasyon fonksiyonlarının kullanımı tercih edilmektedir. Özetle derin öğrenme mimarilerinde

kullanılan aktivasyon fonksiyonlarının türevlenebilir olması, doğrusal olmaması, yerel optimuma takılmadan global optimuma ulaşması gibi özelliklere sahip olması beklenmektedir (Kılıçarslan vd., 2021). Literatürde bahsi geçen amaçlar için kullanılan; Tanh, Softmax, Sigmoid, Swish, ReLU, Leaky ReLU gibi birçok aktivasyon fonksiyonu (Şekil 22) mevcuttur (Sharma vd., 2020).



Şekil 22. Aktivasyon fonksiyonları (URL-15, 2023)

1.7.7. Sınıflandırıcı Katman

Sınıflandırıcı katman derin öğrenme ağlarının son katmanıdır. Mimaride önceki katmanlardan veri alır ve bu verileri istenilen çıktıya dönüştürür. Sınıflandırıcı katman çıktısı, girdi veri seti için etiketlerin bulunduğu son form olmakla birlikte tüm derin öğrenme ağının yaptığı tahminleri içermektedir. Bu katmanda tahminlerin nihai formunun oluşturulması için uygun aktivasyon fonksiyonu seçimi yapılmalıdır.

Son katman olan sınıflandırıcı katman için genellikle *Sigmoid* ve *Softmax* aktivasyon fonksiyonları tercih edilmektedir. Sigmoid; doğrusal olmayan, çıktıların hangi sınıfa ait olduğuna dair karar vermeye yönelik, değer aralığı (0-1) olan, ikili sınıflandırma veya segmentasyon işlemleri göz önüne alındığında iki sınıfa etkili bir şekilde ayırt edebilen olasılıksal bir yaklaşım ve aktivasyon fonksiyonudur (Minai & Williams, 1993). *Çoklu Sigmoid* olarak da bilinen *Softmax* ise etiketlerdeki sınıfların ikiden fazla olduğu durumlarda, en yüksek olasılığa sahip sınıfın çok verimli bir şekilde çıkarılabileceği ve tüm öngörülebilir sınıfların olasılıklarını veren çok katmanlı sınıflandırma veya segmentasyon

problemlerinin çözümü adına mimarilerin çıkış katmanında kullanımı uygun bir aktivasyon fonksiyonudur (Bridle, 1989; Nwankpa vd., 2021).

1.7.8. Maliyet/Kayıp Fonksiyonu

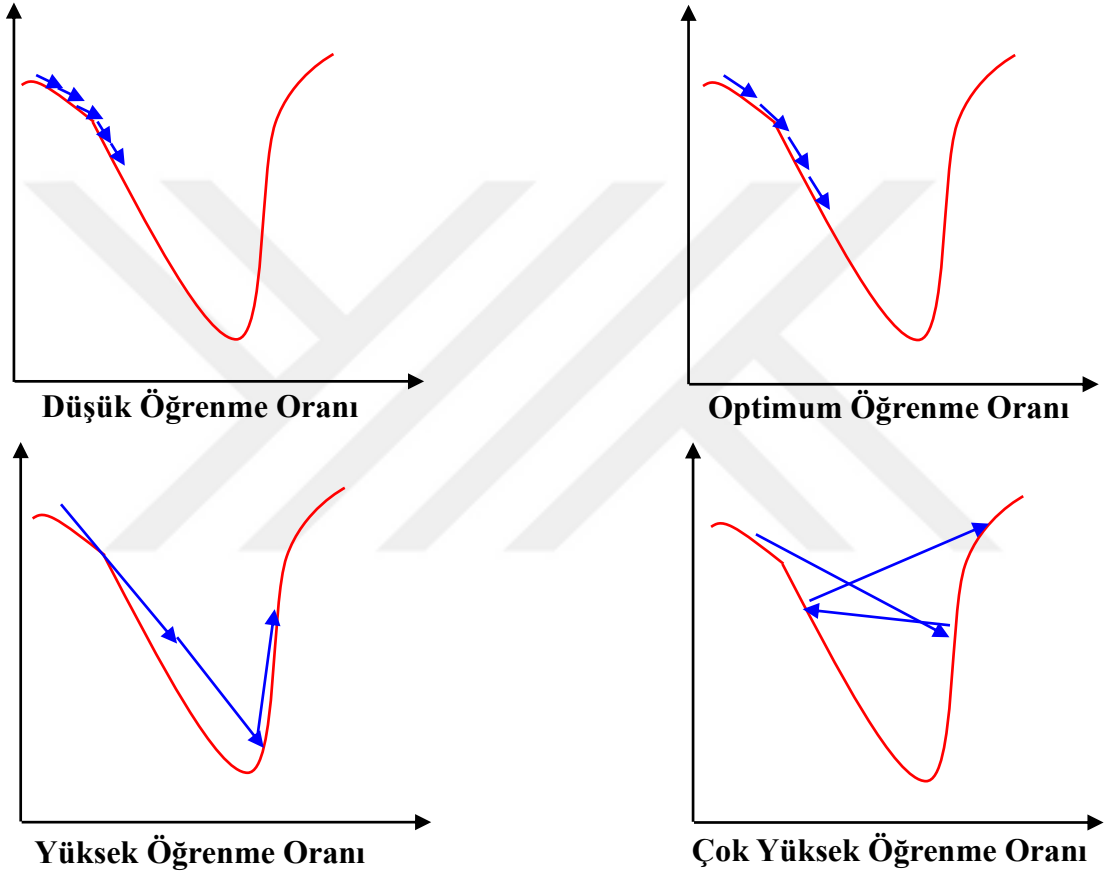
Derin öğrenmede “kayıp fonksiyonu” ve “maliyet fonksiyonu” sıklıkla birbirinin yerine kullanılmaktadır. Her ikisi de bir derin öğrenme modelinde tahmin edilen ve gerçek değerler arasındaki farkı ölçen matematiksel bir fonksiyondur. Modelin performansını ölçerek, oluşturulan mimarinin verilere ne kadar uyduğu konusunda geri bildirim sağlayarak optimizasyon sürecine yön vermektedirler. Mimarinin yaptığı tahminlerin gerçek değerlere en yakın şekilde olması amacıyla minimize veya maksimize etmek istediğimiz fonksiyona amaç fonksiyonu veya kriteri, amaç fonksiyonunu minimize ettiğimizde buna maliyet fonksiyonu veya kayıp fonksiyonu denilmektedir (Goodfellow vd., 2016). Bu bağlamda yüksek bir doğruluk oranı elde edebilmek için modelin eğitim süreci sırasında maliyet veya kayıp fonksiyonun minimum düzeyde olması gerekmektedir.

Literatürde farklı amaçlar için kullanılan birçok kayıp fonksiyonu bulunmaktadır ve kayıp fonksiyonunun seçimi, kullanılan mimarideki katmanların çıktıları dahi etkileyebilmektedir. Bu sebeple ilgili problemin iyi tanınması ve probleme uygun kayıp fonksiyonunun seçilmesi gerekmektedir. Örneğin derin öğrenme ile çözülmek istenen bir regresyon problemi için Karesel Ortalama Hata (Mean Square Error) kayıp fonksiyonu kullanılırken, iki sınıflı sınıflandırma problemleri için İkili Çapraz Entropi (Binary Cross Entropy) ve çok sınıflı sınıflandırma problemleri için ise Kategorik Çapraz Entropi (Categorical Cross Entropy) kayıp fonksiyonları kullanılmaktadır (Mao vd., 2023).

1.7.9. Öğrenme Oranı

Bir sinir ağının ağırlıklarının, stokastik gradyan inişi adı verilen ampirik bir optimizasyon prosedürü yoluyla keşfedilmesi gerekmektedir. Stokastik gradyan inişi, eğitim veri kümesindeki örnekleri kullanarak modelin mevcut durumu için hata gradyanını tahmin eden ve ardından geri yayılımı kullanarak modelin ağırlıklarını güncelleyen bir optimizasyon algoritmasıdır. Öğrenme oranı ise hata gradyanının düşüşüne göre ağırdaki ağırlıkları ayarlamayı kontrol eden bir hiper parametredir. Optimum ağırlıklara doğru, ne derece hızlı veya yavaş ilerleneceğini belirleyerek, sinir ağının öğrendiği kavramları güncelleme hızını tanımaktadır. Öğrenme oranının genellikle 0 ile 0,1 arasında küçük bir

pozitif değere sahip olması beklenilmektedir. Bir sinir ağı için karar verilen öğrenme oranı ağıın yararlı bir şeye yaklaşmasına yetecek kadar düşük, ancak makul bir sürede eğitilebilmesine yetecek kadar yüksek olmalıdır. Küçük öğrenme oranları daha fazla eğitim dönemi gerektirir ve sürecin küresel bir minimumda takılıp kalmasına neden olabilir. Daha büyük öğrenme oranları ise modelin optimal olmayan bir çözüme çok hızlı bir şekilde yakınsamasına neden olabilmektedir (Şekil 23).



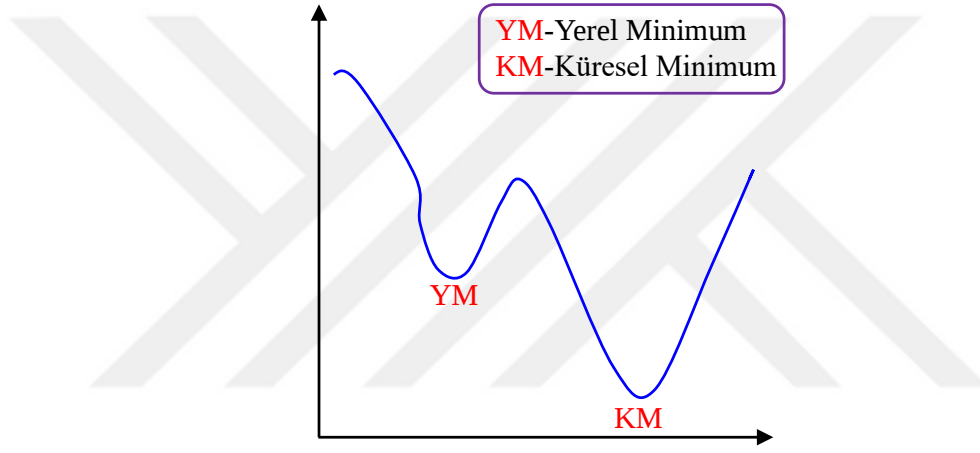
Şekil 23. Öğrenme oranının model üzerindeki etkisi

1.7.10. Optimizasyon Algoritmaları

Optimizasyon algoritmaları, derin öğrenmede mimarinin bir veri seti üzerinde ne kadar iyi tahminler yapabildiğini ölçen ve kayıp fonksiyonunu minimum düzeye indirecek şekilde model parametrelerini ayarlamaktadır. Literatürde, Stokastik Gradyan İnişi (Stochastic Gradient Descent-SGD), İtme (Momentum), Adagrad, Adadelta, RMSprop ve Adamax gibi çeşitli optimizasyon algoritması mevcuttur. Belirli bir veri seti ve mimari için seçilecek

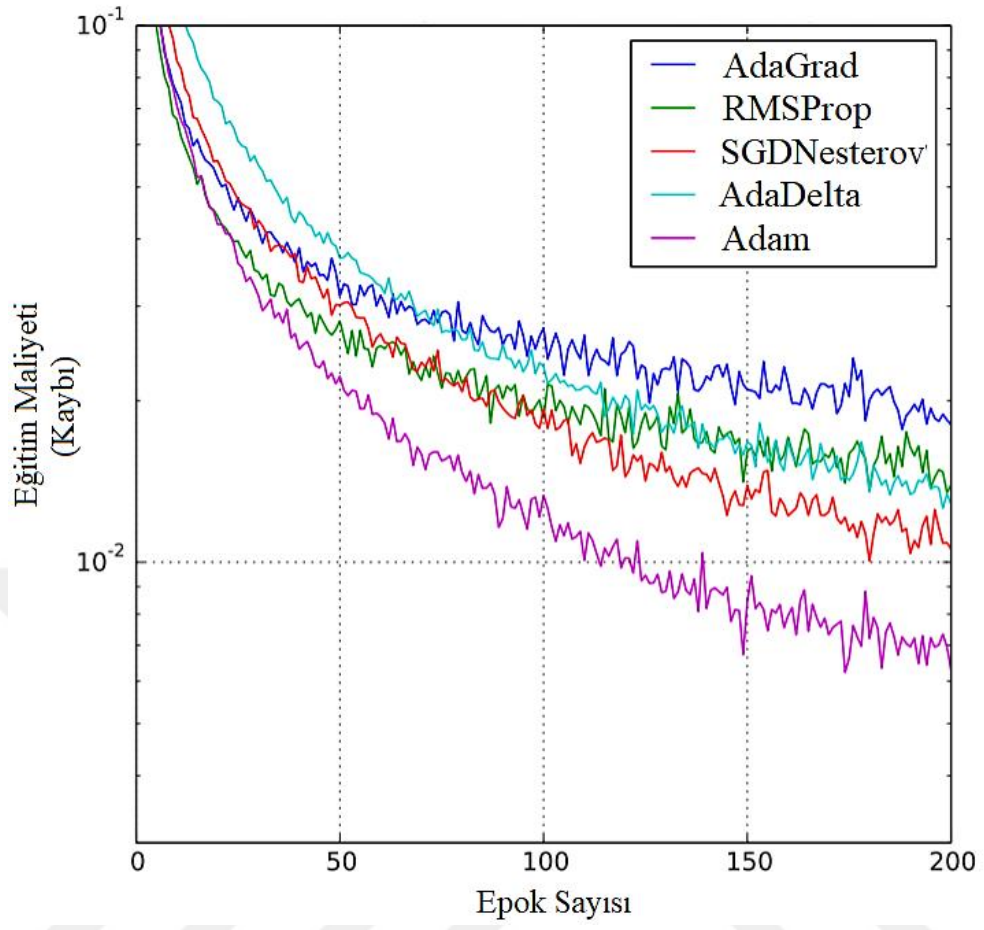
optimize edici, modelin eğitim performansını ve hızını büyük ölçüde etkileyebilmektedir. Literatürde mevcut optimizasyon algoritmaları arasında, İtme (Momentum) ve Ortalama Kare Yayılımın Kökü (Root Mean Square Propagation/RMSP) isimli iki gradyan iniş algoritmasının bir kombinasyonunu içermesi ve daha az bellek gerektirmesiyle birlikte Adam optimizasyon algoritması genellikle derin öğrenmedeki en etkili optimize edicilerden biri olarak kabul edilmektedir (Kingma & Ba, 2015).

Bu algoritmaların birleşimini içermesi ile yol boyunca yerel minimum engellerini aşacak kadar büyük adımlar atarken, küresel minimuma en verimli şekilde ulaşmak amacıyla minimum salınım olacak şekilde ilerlemesi sağlanacaktır (Şekil 24).



Şekil 24. Yerel ve küresel minimum

Önceki algoritmaların güçlü yönlerini temel alan Adam optimize edici, genellikle diğerlerinden daha yüksek performans sağlamaktadır. Adam optimizasyon algoritmasının, eğitim maliyeti ve performans açısından diğer optimize edicilerden farkı Şekil 25'te gösterilmektedir.

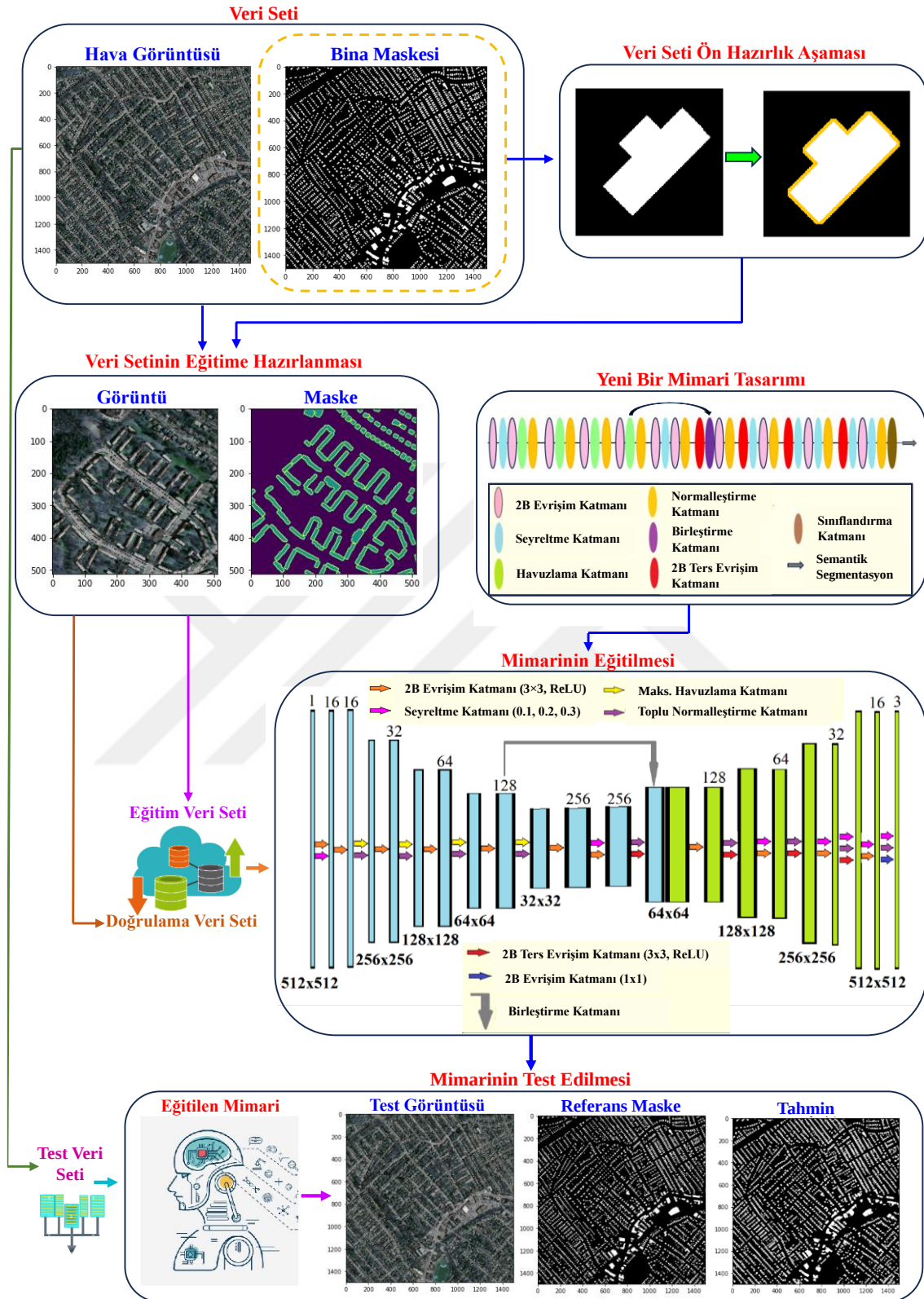


Şekil 25. Optimizasyon algoritmalarının performans karşılaştırması (Ruder, 2016).

2. YAPILAN ÇALIŞMALAR

Bu tez çalışmasında derin öğrenme yöntemleri ile uzaktan algılanmış görüntülerden bina çıkarımı uygulamasını gerçekleştirmek adına, yeni bir derin öğrenme mimarisi tasarlanmış ve bunun yanında bina çıkarım performansını arttırmak için veri setindeki hava fotoğraflarının maskelerinde bulunan bina temsillerine morfolojik operasyonların kullanımıyla otomatik bir biçimde bina sınırı çizdirme işlemi gerçekleştirilmiştir. Elde edilen bina sınır bilgileri derin öğrenme ağına girdi verisi olarak dahil edilmiştir. Maskelerdeki binalara sınır çizdirilerek bina ve arka plandan oluşan iki sınıflı segmentasyon işlemine ilave olarak; bina, arka plan ve bina sınırı olmak üzere üç sınıflı segmentasyon uygulaması, tasarlanan mimarinin eğitilmesiyle gerçekleştirilmiştir. İki sınıf (bina, arka plan) ve üç sınıf (bina, arka plan ve bina sınırı) kullanılarak eğitilen mimarinin bina çıkarım performansı hem görsel hem de doğruluk metriklerinden elde edilen sayısal sonuçların karşılaştırılmasıyla değerlendirilmeye tabi tutulmuştur.

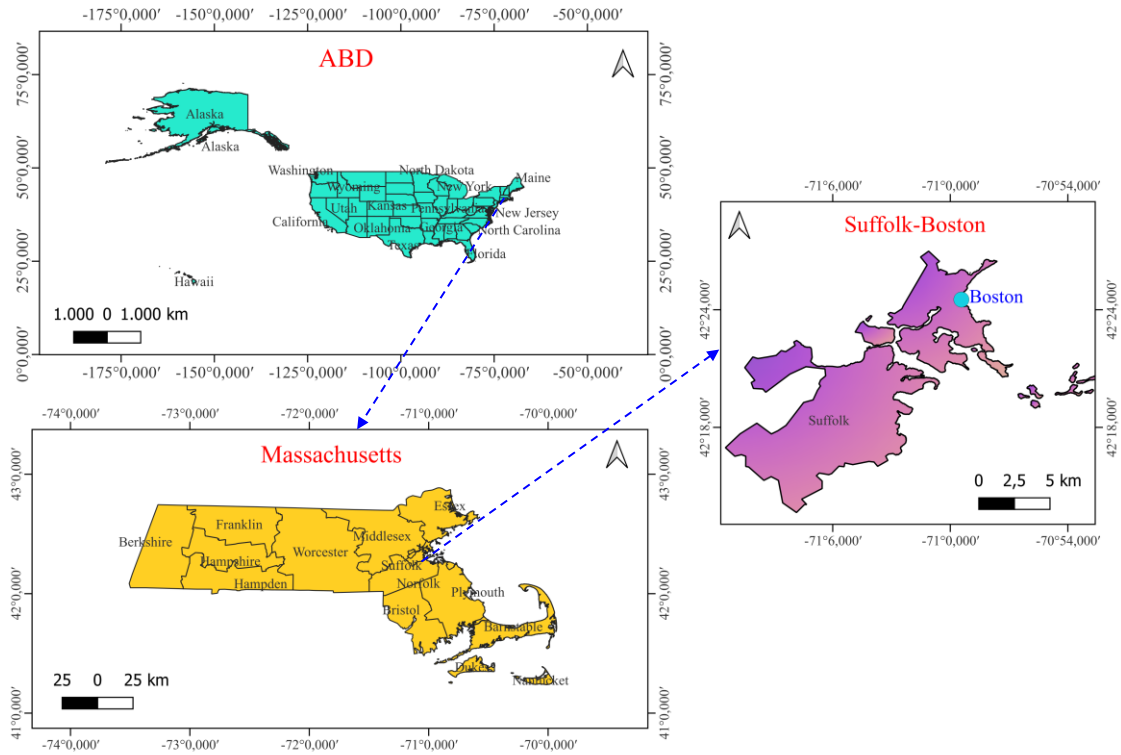
Tasarlanan yeni mimarinin performansının kanıtlanması hususunda, literatürde mevcut ve kullanılabilirliği kanıtlanmış bir derin öğrenme ağı olan U-Net mimarisinin, aynı veri setleri (iki sınıflı ve üç sınıflı olmak üzere) ile aynı şartlarda (mimarilere gözlemlenen en uygun ayar parametreleri, ortak parti boyutu ve epok sayısı) eğitilmesiyle, her iki mimarinin bina çıkarımı için görüntü segmentasyon performansları karşılaştırılmıştır. Gerçekleştirilen çalışmanın iş akış şeması Şekil 26'da sunulmaktadır.



Şekil 26. Genel iş akış şeması

2.1. Çalışma Alanı ve Veri Seti

Uzaktan algılanmış görüntülerden bina çıkarımı uygulamasında kullanılmak üzere açık kaynak olarak sunulmayan *Massachusetts* bina veri seti (Mnih, 2013), Kaggle çevrimiçi platformundan kısıtlı olarak elde edilmiştir. *Massachusetts* bina veri seti görüntüleri, Amerika Birleşik Devletleri'nin *Massachusetts* Eyaletinin başkenti ve en büyük şehri olup, $42^{\circ} 21' 29''$ Kuzey enlemi ve $71^{\circ} 3' 49''$ Batı boylamında bulunan Boston şehrine aittir (Şekil 27).

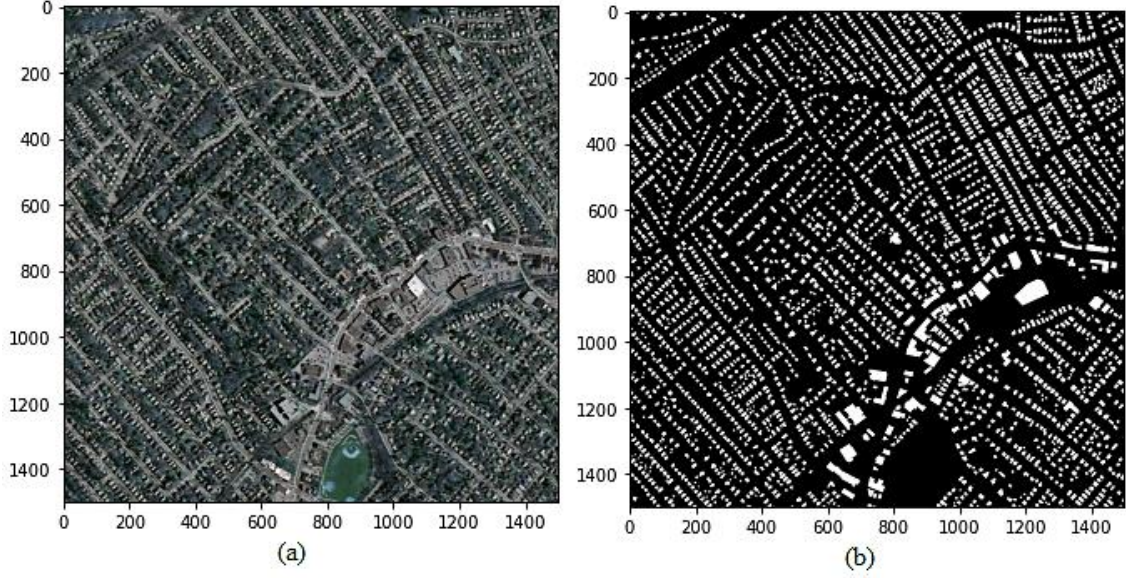


Şekil 27. Çalışma alanı

Veri seti 1 m konumsal çözünürlüğe sahip 151 adet hava görüntüsü ve bu görüntülere ait bina maskelerinden oluşmaktadır. Görüntülerin her biri 2.25 km^2 alanı 1500×1500 piksel ile temsil etmektedir (Şekil 28). Dolayısıyla, tüm veri seti yaklaşık olarak 340 km^2 lik bir alanı kapsamaktadır.

Massachusetts bina veri setinin, açık kaynaklı olmamasından ötürü, bu çalışmada veri setinde bulunan 151 adet hava görüntüsü ve maskesinden, sadece 115 adet görüntü ve maskeye ait, yaklaşık 260 km^2 lik alanı kapsayan bir bölgeye erişim sağlanabilmektedir. Veri

seti, çoğunlukla kentsel ve kentlere çok yakın (banliyö) alanları, münferit evler ve garajlar da dahil olmak üzere farklı mimari ve nizamda her büyüklükteki binayı içerisinde barındırmaktadır.



Şekil 28. Massachusetts bina veri seti örneği (a) Hava görüntüsü, (b) Bina maskesi (Mnih, 2013)

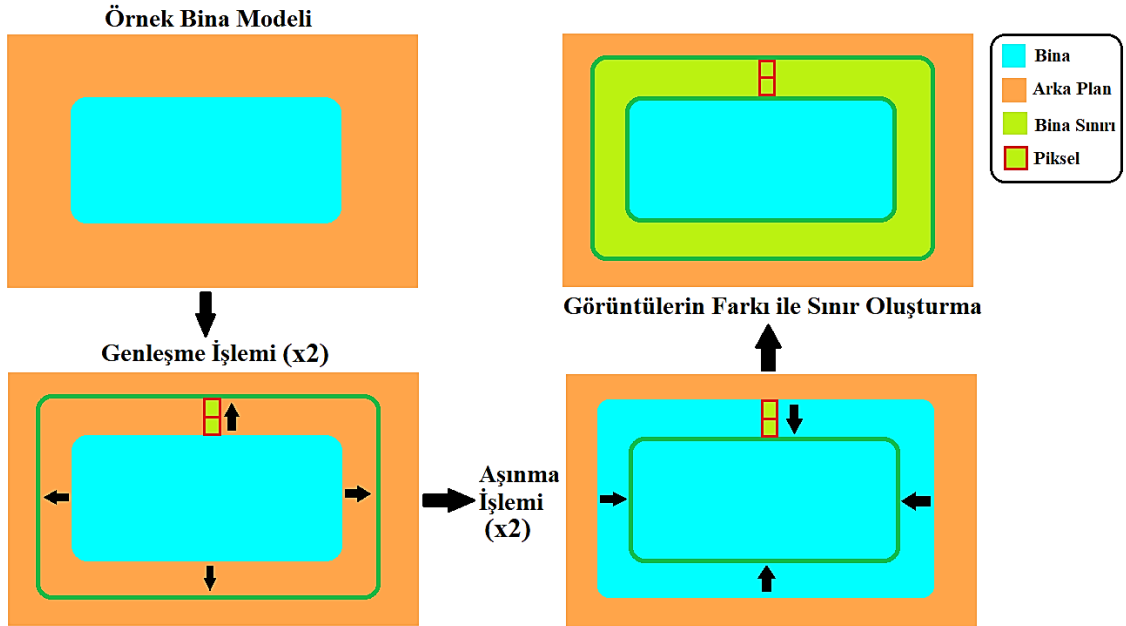
2.2. Veri Seti Ön Hazırlık İşlemleri

2.2.1. Maske Veri Setindeki Binalara Sınır Çizdirilmesi

Çalışmada veri setini iki (bina, arka plan) ve üç (bina, arka plan ve bina sınırı) sınıflı segmentasyon işlemine hazırlamak adına birtakım işlemler gerçekleştirilmiştir. Bunlardan ilki; hava görüntülerinin, ikili (binary/siyah-beyaz) formattaki maske verilerinde bulunan, 1 gri değerine sahip bina etiketlerine sınır oluşturmak için morfolojik operasyonlar kapsamındaki genişleme ve aşınma metotlarını kullanmaktır. Maske verilerinde üçüncü sınıf olan bina sınırlarını oluşturmak için ikili (binary) formattaki siyah-beyaz görüntülere önce [1] matrisi yapılandırma elemanı kullanılarak iki kez genişleme işlemi uygulanmıştır. Ardından genişleme işlemi uygulanan görüntü üzerinde, aynı şekilde [1] matrisi yapılandırma elemanı kullanılarak iki kez aşınma işlemi gerçekleştirilmiştir. Genişleme ve aşınma işlemi uygulanan görüntülerin farkı alınarak, arada kalan iki piksel genişliğindeki fark kısmına 0.5 renk değeri (gri renk) atanmış ve bu renk değerine sahip bölgeler bina sınırı

olarak tanımlanmıştır (Şekil 29). Bina sınırı oluşturma işlemi 1500×1500 boyutundaki 115 adet maske verisinin tümü için otomatik bir şekilde Python ortamında gerçekleştirilmiştir.

Bu işlemlerde [1] matrisi yapılandırma elemanı kullanılmasının tek ve en önemli nedeni; ana maske verisindeki bina kenarlarının orijinallliğini bozmaksızın bina sınırlarını oluşturmaktır. Ayrıca bir diğer detay olan art arda iki kez genişleme ve ardından aşınma işlemi uygulanmasının sebebi; iki piksel genişliğinde bina sınırı oluşturmak istenmesidir. Yapılan denemeler sonucunda, bir piksel genişliğinde elde edilen bina sınırı ile gerçekleştirilen üç sınıflı segmentasyon işleminde derin öğrenme mimarisinin bina sınırlarını net olarak algılayamadığı ve iki pikselden daha geniş bina sınırları kullanıldığında ise yan yana bulunan bina sınırlarının çakışabilme olasılığı gözlemlenmiştir. Elde edilen bu çıkarımlar sebebiyle, tez kapsamında maske verilerinde bulunan bina temsillerine iki piksel genişlikli sınır çizdirilmesi uygun görülmüştür.

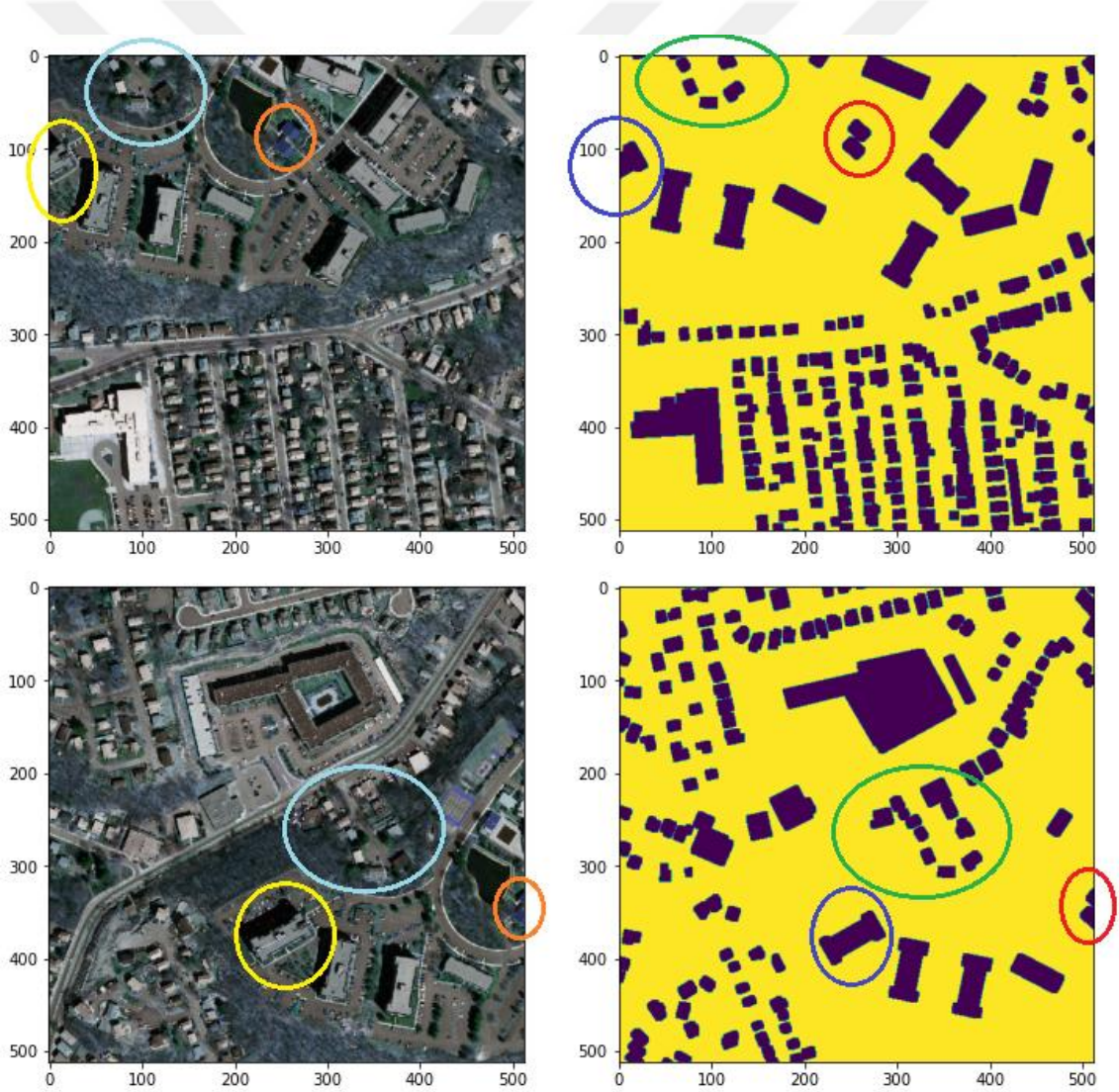


Şekil 29. Morfolojik operasyonlar ile bina sınırlarının oluşturulması

2.3. Eğitim, Doğrulama ve Test Veri Setinin Oluşturulması

Girdi verinin derin öğrenme mimarisinin eğitiminde kullanılmak üzere eğitim, doğrulama ve test veri seti olmak üzere üç bölüme ayrılması gerekmektedir. Bu hususta mevcut çalışma için kullanılan 115 adet hava fotoğrafı ve bina maskesinden, 1500×1500 boyutuna sahip 5 adet hava fotoğrafı ve bina maskesi, test veri seti olarak seçilip, eğitim

işlemine tabi tutulmamıştır. Geriye kalan 110 adet 1500×1500 boyutundaki hava fotoğrafları ve maskeleri ise Python ortamında bulunan *Patchify* kütüphanesi kullanılarak, %50 enine ve boyuna bindirme oranlarıyla parçalanmıştır. Görüntülerin yamalara ayrılması aşamasında enine ve boyuna olacak şekilde %50 bindirme oranı kullanılmasının iki ana sebebi mevcuttur. Bunlardan ilki; daha fazla veri sayısı ile eğitim işlemini gerçekleştirme hedefi iken bir diğer sebep ise Şekil 30’da gösterildiği üzere bir görüntünün kenar kısmına denk gelerek tamamı algılanmayan bir binanın bindirme oranı kullanılarak kırılması üzerine diğer bir görüntüde tamamının algılanmasını sağlamaktır. Böylelikle görüntülerde bulunan tüm binalar eğitim işlemine dahil olabilecek ve derin öğrenme mimarisi tarafından daha fazla bina türü öğrenilebilecektir.



Şekil 30. Bindirme oranı ile tüm binaların algılanması

Görüntü kırpma işlemleri sonucunda 512×512 boyutlu 990 adet hava görüntüsü ve bina maske yaması elde edilmiştir. İşlemin devamında ise 990 adet görüntünün yaklaşık %9'u (90 adet) doğrulama veri seti, geri kalanı (900 adet) ise eğitim veri seti olarak derin öğrenme modelinde kullanılmak üzere kaydedilerek eğitim işlemine hazır hale getirilmiştir (Tablo 4).

Tablo 4. Eğitim, doğrulama ve test veri seti miktarları

Veri Seti	Hava Fotoğrafı ve Maske Sayısı	Boyut
Eğitim Veri Seti	900	512×512
Doğrulama Veri Seti	90	512×512
Test Veri Seti	5	1500×1500
Toplam Veri Seti	115	1500×1500

Eğitim işlemine başlanmadan önce, heterojen olarak 0-255 gri değerleri arasında renk değerlerine sahip (8 bit) hava görüntüsü yamaları üzerinde ölçekleme işlemi yapılarak [0-1] aralığında gri değerlere sahip olmaları sağlanmış ve homojen bir veri yapısı ile bilgi kaybının önüne geçilmesi hedeflenmiştir. Maske verileri halihazırda [0-1] aralığındaki renk değerlerine sahip oldukları için bu verilere ölçekleme işlemi uygulanmasına gerek duyulmamıştır.

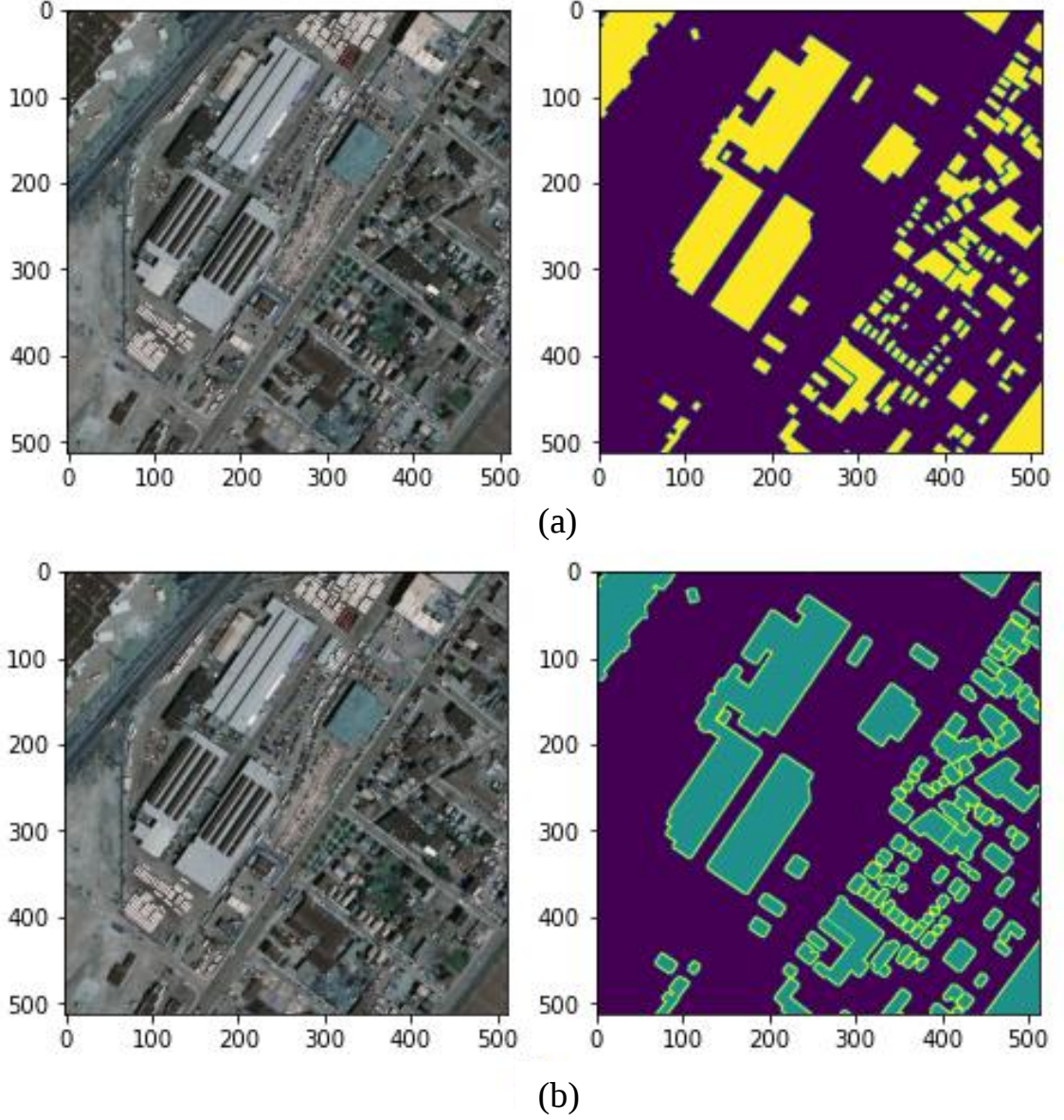
Bahsi geçen işlem adımları, tüm hava fotoğrafları ile hem işlenmemiş hem de daha önce bina sınırlarını oluşturmak adına ön hazırlığa tabi tutulmuş maske veri setlerine uygulanarak, iki (bina, arka plan) ve üç (bina, arka plan ve bina sınırı) sınıflı semantik segmentasyon işlemleri için eğitim ve doğrulama veri setleri hazır hale getirilmiştir.

2.4. Yeni Bir Derin Öğrenme Mimarisi Tasarımı ve Eğitimi

Tez çalışmasında, uzaktan algılanmış görüntülerden bina çıkarımı uygulamasının, hızlı ve yüksek doğruluğa sahip bir şekilde yapılması hususunda literatürde bulunan ve iki veya çok sınıflı görüntü segmentasyonu işlemini gerçekleştiren mevcut derin öğrenme mimarilerine alternatif olabilecek yeni bir derin öğrenme mimarisi tasarlanmıştır.

Eğitim işlemi için kullanılan girdi veri $512 \times 512 \times 3$ boyutuna sahip görüntülerden (RGB) oluşmaktadır. Tasarlanan mimari, iki (bina, arka plan) ve üç (bina, arka plan ve bina sınırı) sınıflı segmentasyon yapılması hususunda iki şekilde eğitilmiş ve eğitim sonucu elde

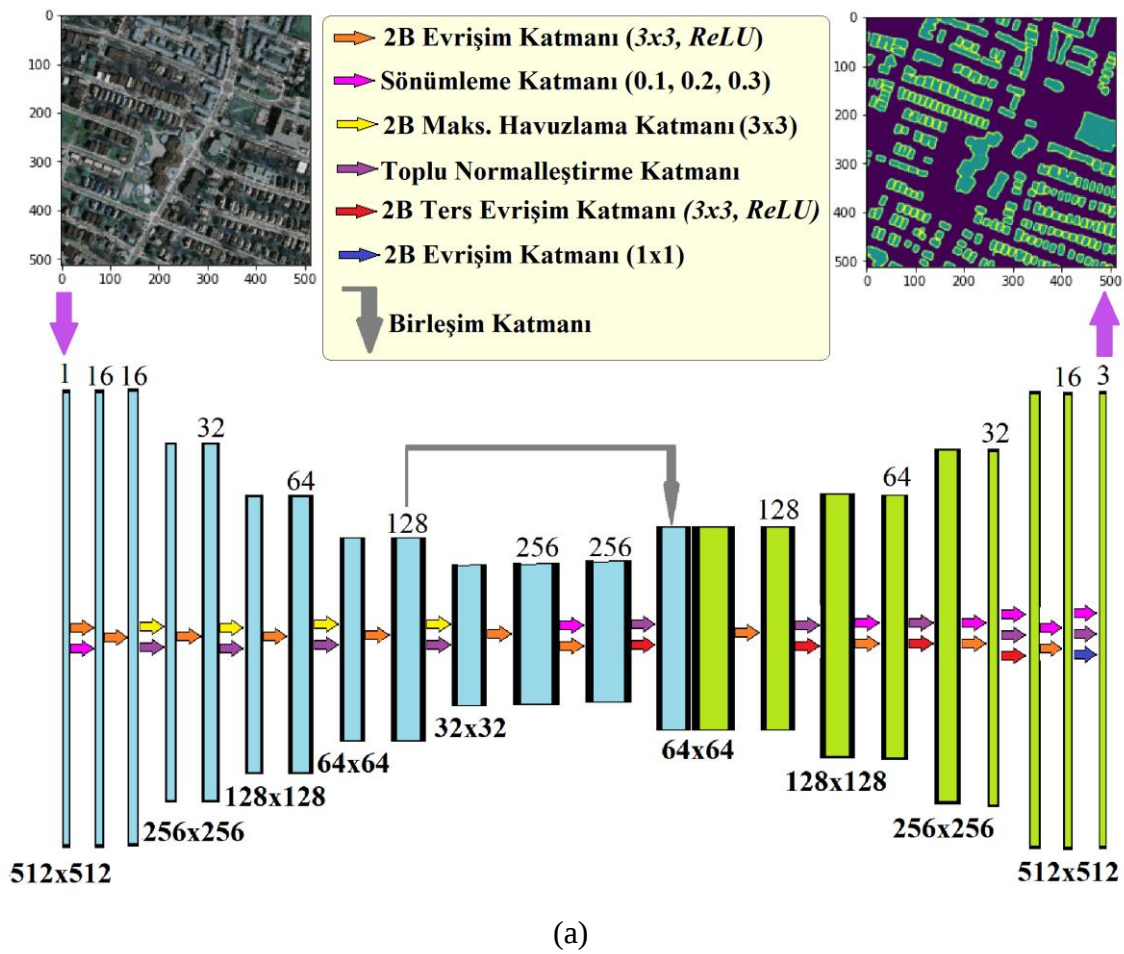
edilen modeller test aşamasında kullanılmak üzere kaydedilmiştir. İki ve üç sınıflı segmentasyon için kullanılan girdi verilere ait örnek görseller Şekil 31’de gösterilmiştir.



Şekil 31. Girdi verisi örnekleri (a) İki sınıflı segmentasyon için görüntü ve maske, b) Üç sınıflı segmentasyon için görüntü ve maske

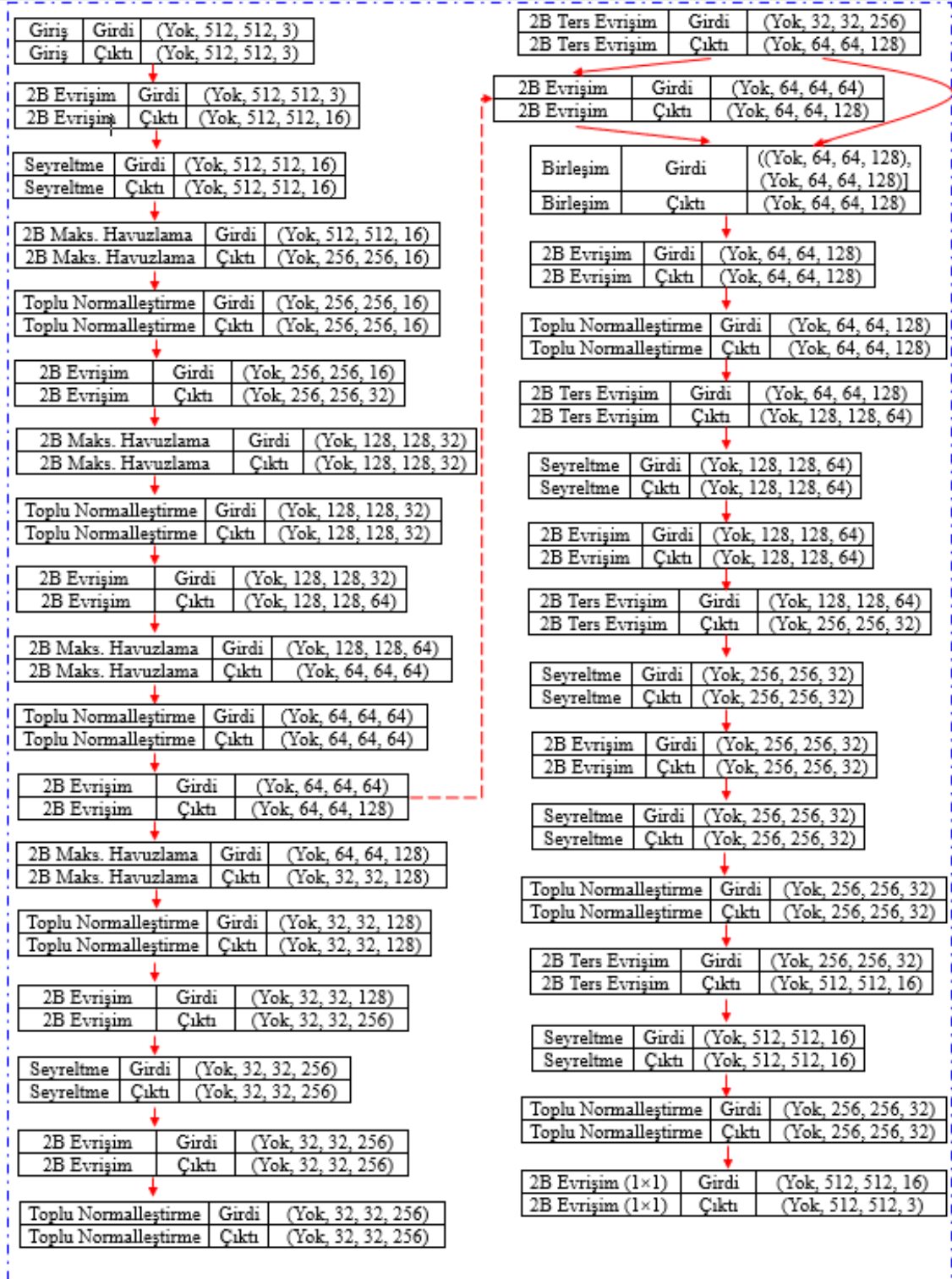
Önerilen yeni derin öğrenme mimarisi kodlayıcı ve kod çözücü olmak üzere iki bölümden oluşmakta ve 16 katman bulundurmaktadır. Mimarinin kodlayıcı bölümünde 3×3 boyutlu filtrelere sahip 7 adet 2B Evrişim katmanı, kod çözücü bölümünde ise 3×3 boyutlu filtrelere sahip 4 adet 2B Evrişim ve Ters Evrişim katmanı bulunmaktadır. Son katman olan sınıflandırma katmanında ise 1 adet 1×1 boyutunda filtresi bulunan 2B Evrişim katmanı

kullanılmıştır. Mimaride sınıflandırma katmanı (son katman) dışındaki tüm katmanlarda aktivasyon fonksiyonu olarak *ReLU*, sınıflandırma katmanında ise 2 sınıflı segmentasyon için *Sigmoid*, 3 sınıflı segmentasyon için *Softmax* aktivasyon fonksiyonları tercih edilmiştir. Ayrıca mimari 0.001 öğrenme oranıyla, Adam optimizasyon algoritması ve parti boyutunun (batch size) 8 olarak seçilmesiyle 100 epokta eğitilmiştir. Tasarlanan mimarinin görsel temsili Şekil 32a ve her katman için girdi ve çıktı boyutlarının detayları Şekil 32b’de gösterilmiştir.



Şekil 32. Tasarlanan mimarinin detayları (a) Tasarlanan mimarinin görsel temsili, (b) Tasarlanan mimari katmanlarının detayları

Şekil 32'nin devamı

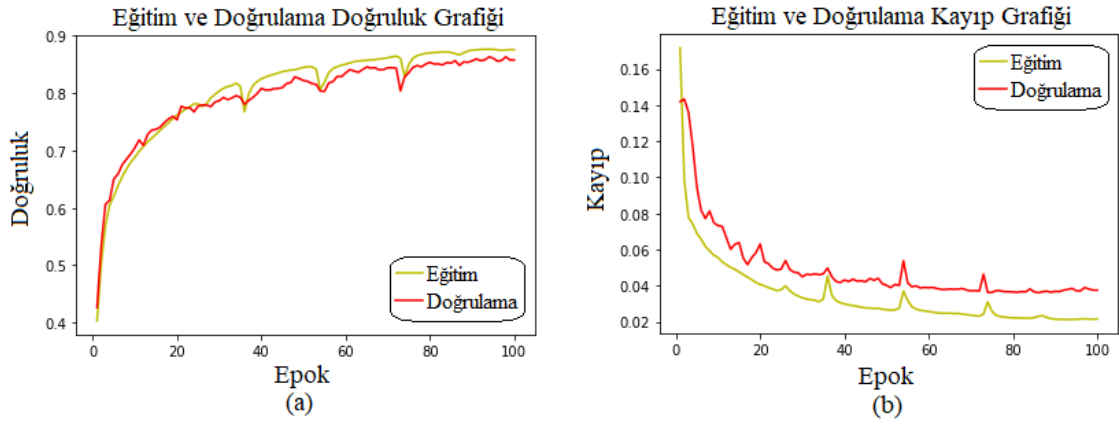


(b)

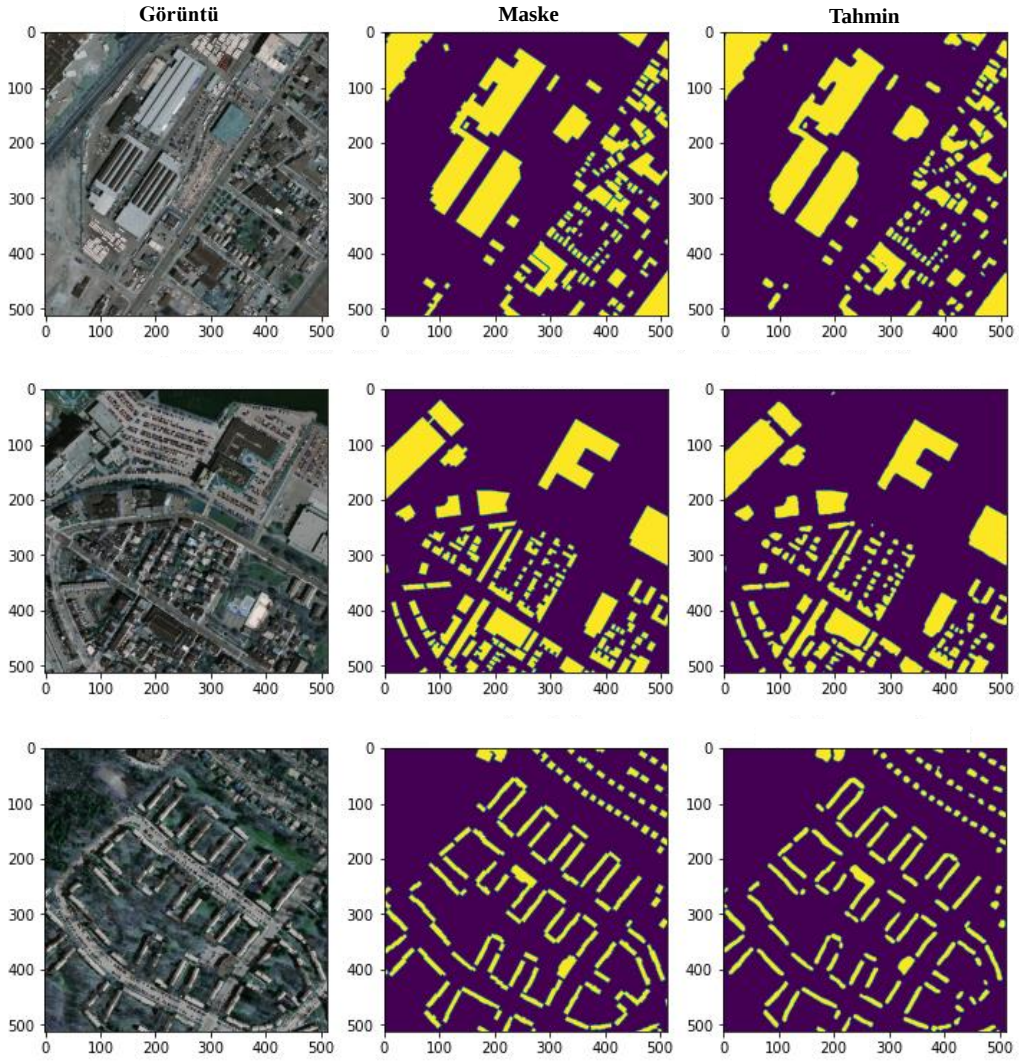
2.4.1. Tasarlanan Yeni Mimari ile İki Sınıflı Semantik Segmentasyon

Çalışma içeriğinde derin öğrenme yöntemleri ile iki sınıflı semantik segmentasyon uygulaması, önerilen mimari vasıtasıyla uzaktan algılanmış verilerde bulunan bina dışı objeleri arka plan sınıfı olarak; değişken çatı tipi, nizam ve yüksekliğe sahip binaların tümünü ise bina sınıfı olarak sınıflandırma işleminin makineye öğretilmesi vasıtasıyla gerçekleştirilmiştir. İşlemden eğitim için girdi verisi olarak daha önce 512×512 boyutunda parçalara ayrılan 900 adet, doğrulama için ise aynı boyuta sahip, 90 adet hava fotoğrafı ve maske yamaları kullanılmıştır. Kaydedilen veri seti hava görüntüleri ve bu görüntülere ait ikili (binary) veri tipindeki bina maske verileri *.npy uzantılı olarak Python ortamına aktarılmıştır. Ardından maske verisinde bulunan bina ve arka plan özelliklerinin tanıtımı, sınıfların sahip olduğu renk değerlerinin kodları yardımı ile #000000 (siyah) renk koduna sahip arka plan sınıfına 0 rakamı, #ffffff (beyaz) renk koduna sahip bina sınıfına ise 1 rakamı atanarak etiketleme işlemi gerçekleştirilmiştir. Dolayısıyla artık makinenin bina sınıfını 1 rakamı ile, bina dışı objelerin bulunduğu arka plan sınıfını ise 0 rakamı ile eşleştirilmesini sağlamaktadır.

Bina çıkarımı amacına yönelik iki sınıflı semantik segmentasyon eğitimi, yeni mimarinin son katmanı olan sınıflandırma katmanında *Sigmoid* aktivasyon fonksiyonunun kullanımı ile birlikte 0.001 öğrenme hızıyla ve 100 epokta gerçekleştirilmiştir. Eğitim işleminin sonlanması ile doğrulama verisi için ortalama IoU doğruluk metriği değeri 0.8526 olarak hesaplanmış olup, eğitim yaklaşık 1130 dakikada tamamlanmıştır. Eğitim işlemine ait doğruluk-kayıp grafiği Şekil 33'te ve mimarinin doğrulama verileri üzerinde yaptığı tahmin sonuçlarına yönelik örnek görseller Şekil 34'te sunulmuştur. Ağırlık bilgilerinin depolandığı model, test aşamasında kullanılmak üzere kaydedilmiştir.



Şekil 33. Tasarlanan mimarinin iki sınıflı semantik segmentasyon için eğitim-doğrulama grafikleri (a) Doğruluk grafiği, (b) Kayıp grafiği



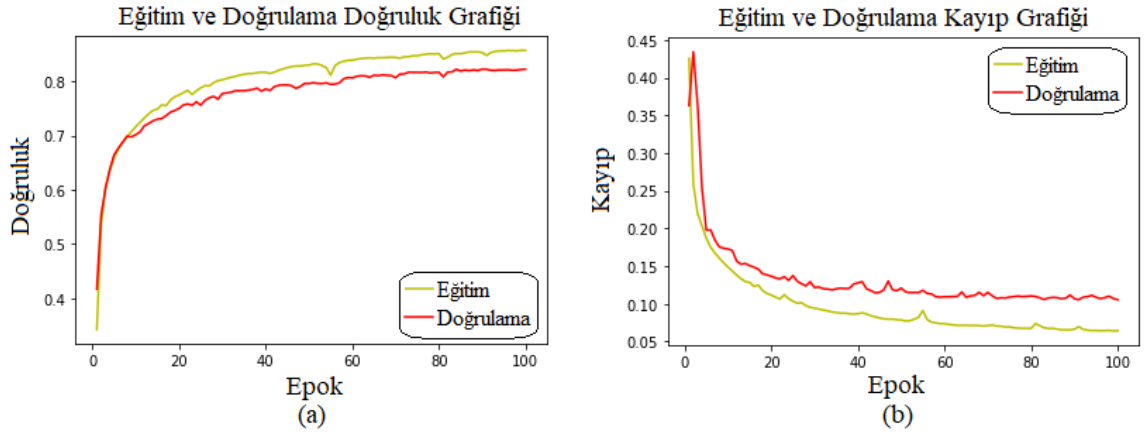
Şekil 34. Tasarlanan mimarinin iki sınıflı semantik segmentasyon örnek tahmin sonuçları

2.4.2. Tasarlanan Yeni Mimari ile Üç Sınıflı Semantik Segmentasyon

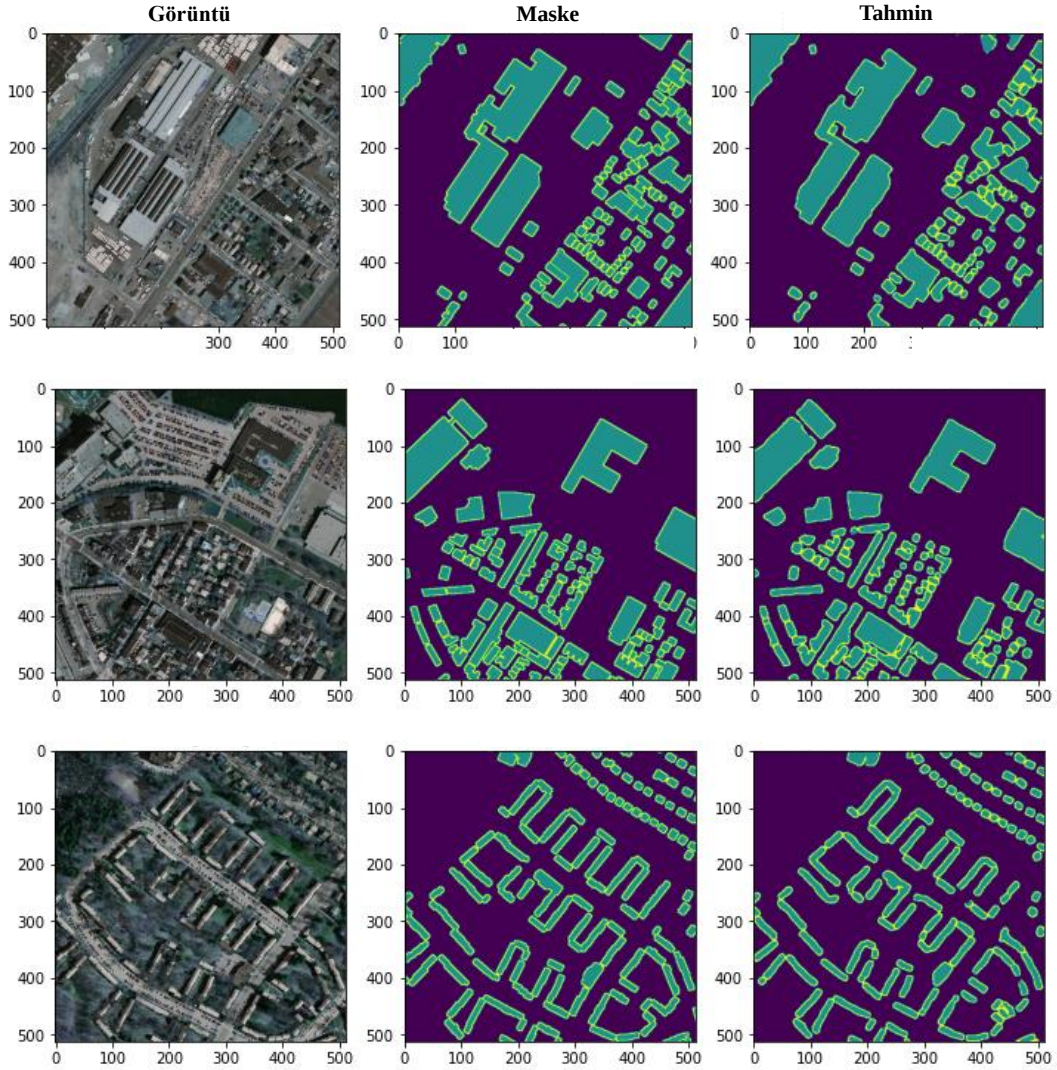
Bu aşamada, iki sınıflı semantik segmentasyona ilave olarak, çalışmanın ön hazırlık aşamasında oluşturulan bina sınır sınıfının eklenmesiyle üç sınıflı semantik segmentasyon işlemi gerçekleştirilmiştir. Eğitim aşamasında ek olarak bina sınırlarının kullanılması ile binaların çıkarılmasına yönelik bina sınırı katmanının sağlayacağı katkı derecesinin değerlendirilmesi hedeflenmiştir. Eklenen yeni sınıf ile bina çıkarım işleminin olumlu yönde etkileneceği öngörülmüştür. Bunun sebebi derin öğrenme mimarisine, bina sınır bilgisinin de kullanılmasıyla daha fazla bilgi sunulmasının sağlanmış olmasıdır.

Üç sınıflı segmentasyon işleminde de aynı şekilde eğitim için girdi verisi olarak 512×512 boyutlu parçalara ayrılan 900 adet, doğrulama için ise aynı boyutlara sahip olan 90 adet hava fotoğrafı ve bina maske yamaları kullanılmıştır. Üç sınıflı semantik segmentasyonda kullanılan maske verileri (0-1) aralığında renk değerine sahip double formatta görüntü verileridir. Maskelerde bulunan #000000 (siyah) renk kodu arka planı, #ffffff (beyaz) renk kodu binaları ve #7b7b7b (gri) renk kodu ise bina sınırlarını temsil etmektedir. Bu renk kodları kullanılarak arka plan 0, bina 1 ve bina sınırı ise 2 rakamı ile etiketlenmiştir. Maske verilerindeki objeler için genellikle her sınıfa bir rakam tanımlanmasıyla etiket atanması yapılmaktadır. Bunun anlamı, makinenin artık öğrenme işleminde ve ardından test için kullanılan hava görüntülerinin semantik segmentasyonunda 0,1 ve 2 rakamlarını; bina, arka plan ve bina sınırını tanımlamak için kullanacak olmasıdır.

Önerilen mimari ile üç sınıflı semantik segmentasyon işleminin eğitimi, sınıflandırma katmanında *Softmax* aktivasyon fonksiyonu ve öğrenme hızının 0.001 olarak kullanılmasıyla, 100 epokta gerçekleştirilip, model test aşamasında kullanılmak amacıyla kaydedilmiştir. Eğitim sonucu doğrulama verileri için hesaplanan ortalama IoU doğruluk metriği değeri 0.7439 olarak hesaplanmış ve eğitimin yaklaşık olarak 1318 dakikada tamamlandığı tespit edilmiştir. Eğitim işlemine ait doğruluk-kayıp grafiği Şekil 35'te ve mimarinin doğrulama verileri üzerinde yaptığı tahmin sonuçlarına yönelik örnek görseller Şekil 36'da gösterilmiştir.



Şekil 35. Tasarlanan mimarinin üç sınıflı semantik segmentasyon için eğitim-doğrulama grafikleri (a) Doğruluk grafiği, (b) Kayıp grafiği

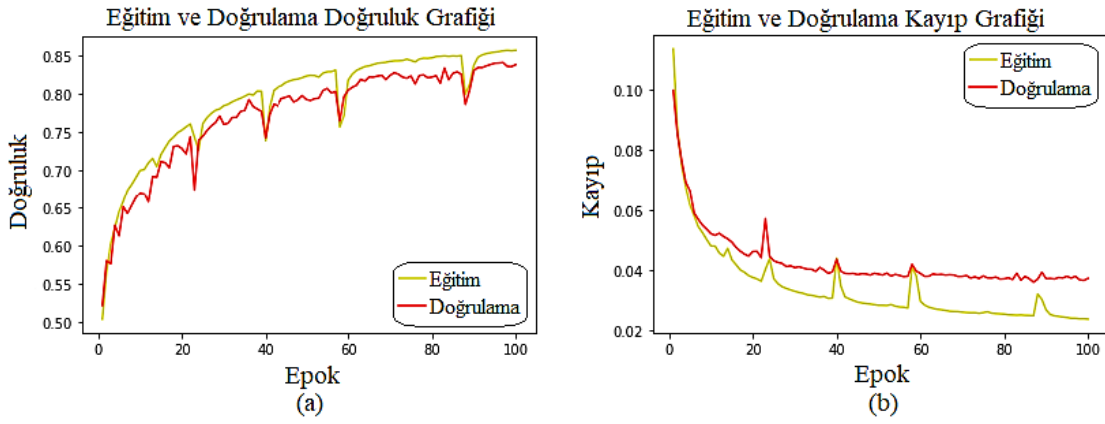


Şekil 36. Tasarlanan mimarinin üç sınıflı semantik segmentasyon örnek tahmin sonuçları

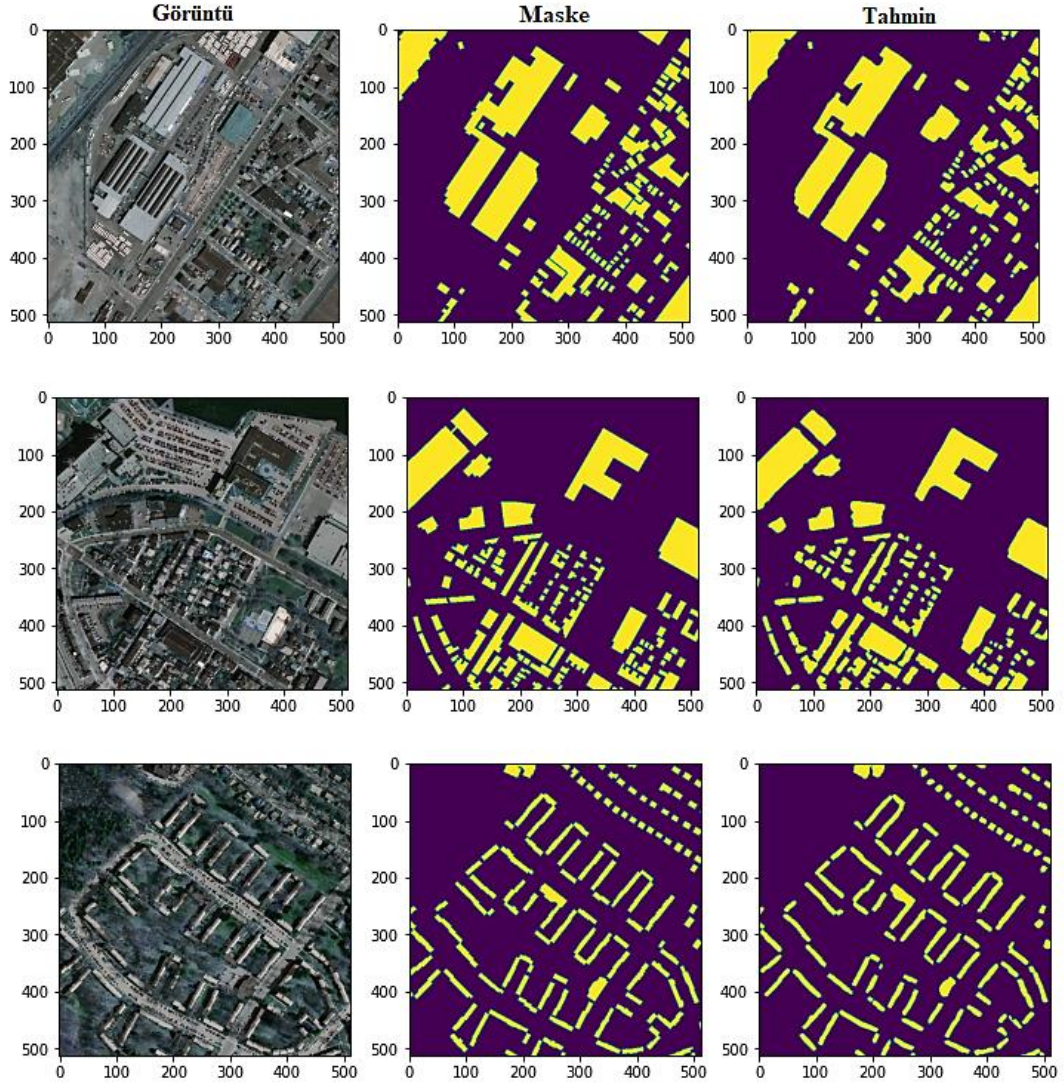
2.4.3. U-Net Mimarisi ile İki Sınıflı Semantik Segmentasyon

Söz konusu çalışmada tasarlanmış olan yeni mimarinin semantik segmentasyon konusundaki etkinliğini gösterebilmek adına, karşılaştırma maksadıyla literatürde bulunan ve başarısı kanıtlanmış bir mimari olan U-Net ile aynı veri setleri kullanılarak, aynı şartlar altında eğitim işlemi gerçekleştirilmiştir. Bu aşamada da iki sınıflı semantik segmentasyon için daha önce bahsedilen bina ve arka plan renk kodları kullanılarak, 0 ve 1 rakamları ile etiketleme işlemi yapılmıştır.

U-Net mimarisinin son katmanı olan sınıflandırma katmanında da yeni mimari ile yapılan iki sınıflı segmentasyonda olduğu gibi *Sigmoid* aktivasyon fonksiyonunun kullanımı ile birlikte 0.001 öğrenme hızı ve 100 epokta eğitim gerçekleştirilmiştir. Doğrulama verileri için ortalama IoU doğruluk metriği değeri 0.8604 olarak hesaplanmış ve yaklaşık 1506 dakikada tamamlanan eğitimin modeli kaydedilmiştir. Eğitim işlemine ait doğruluk-kayıp grafiği Şekil 37’de ve mimarinin doğrulama verileri üzerinde yaptığı tahmin sonuçlarına yönelik örnek görseller ise Şekil 38’de gösterilmiştir.



Şekil 37. U-Net mimarisinin iki sınıflı semantik segmentasyon için eğitim-doğrulama grafikleri (a) Doğruluk grafiği, (b) Kayıp grafiği



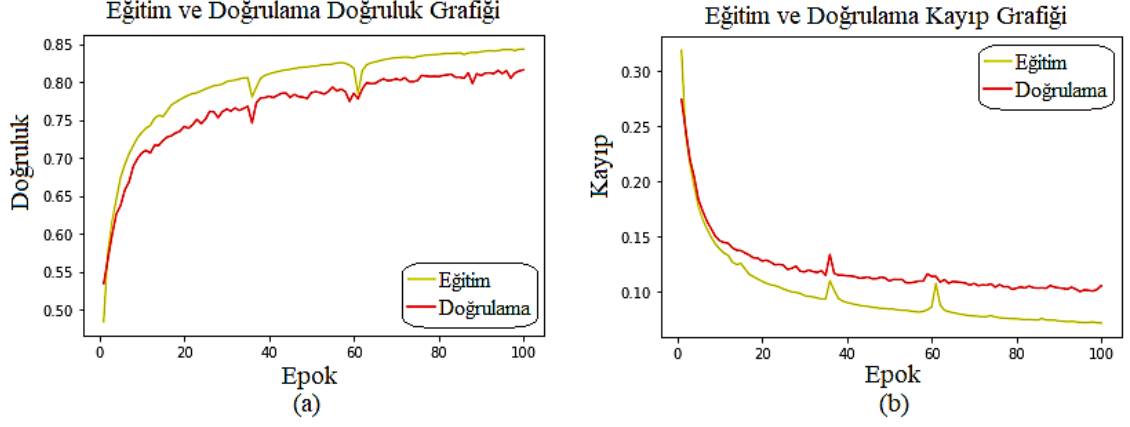
Şekil 38. U-Net mimarisinin iki sınıflı semantik segmentasyon örnek tahmin sonuçları

2.4.4. U-Net Mimarisi ile Üç Sınıflı Semantik Segmentasyon

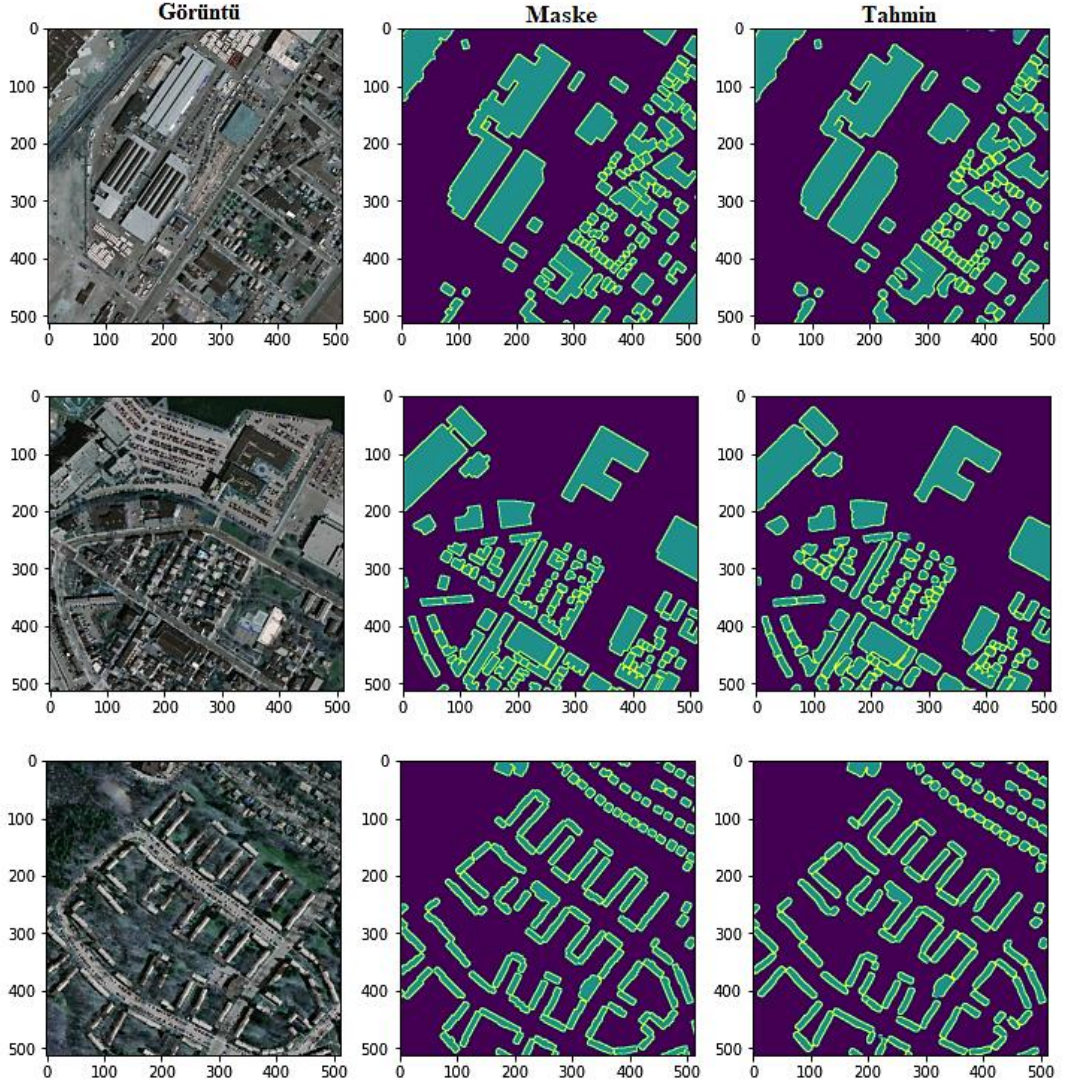
Bu bölümde de tasarlanan mimari ile üç sınıflı semantik segmentasyon işlemi sonucunu kıyaslamak amacıyla aynı veriler ve koşullar kullanılarak U-Net mimarisi ile üç sınıflı semantik segmentasyon eğitimi işlemi gerçekleştirilmiştir. Aynı şekilde bina, arka plan ve bina sınırı sınıfları için ilgili rakamlar (0,1 ve 2) ile etiketleme yapılmıştır.

U-Net mimarisinin son katmanı olan sınıflandırma katmanında da yeni mimari ile yapılan üç sınıflı segmentasyonda olduğu gibi *Softmax* aktivasyon fonksiyonu kullanılarak, 0.001 öğrenme hızıyla ve 100 epokta eğitim gerçekleştirilmiştir. Eğitim işleminin süresi yaklaşık 1695 dakikaya tekabül etmektedir. Doğrulama verisi için ortalama IoU doğruluk metriği değeri 0.7387 olarak hesaplanmış ve model kaydedilmiştir. Eğitim işlemine ait

doğruluk-kayıp grafiği Şekil 39’da ve mimarinin doğrulama verileri üzerinde yaptığı tahmin sonuçlarına yönelik örnek görseller ise Şekil 40’da gösterilmiştir.



Şekil 39. U-Net mimarisinin üç sınıflı semantik segmentasyon için Eğitim-Doğrulama grafikleri (a) Doğrululuk grafiği, (b) Kayıp grafiği



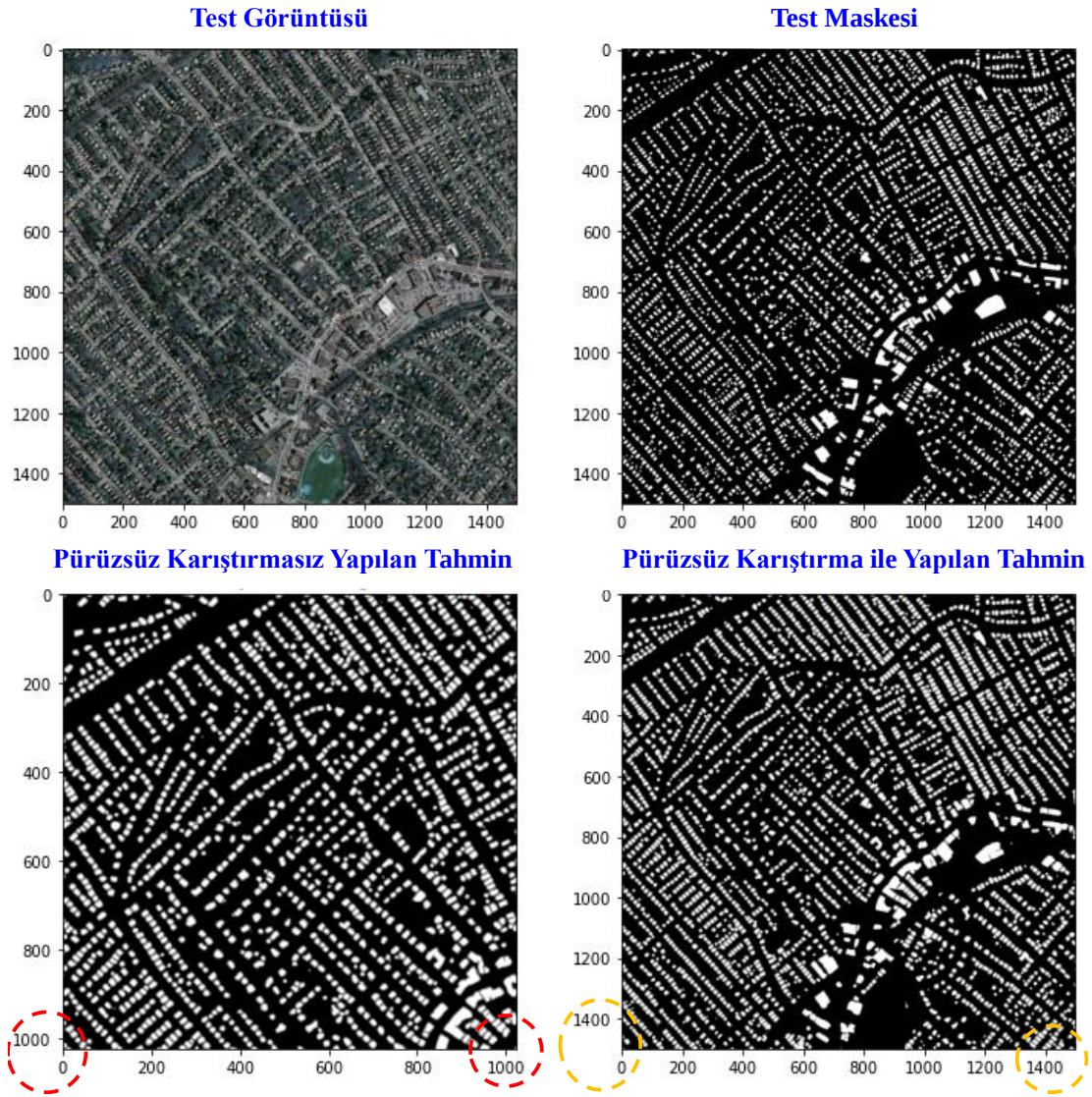
Şekil 40. U-Net mimarisinin üç sınıflı semantik segmentasyon örnek tahmin sonuçları

Yapılan tüm eğitimlerin sonucu olan, doğruluk-kayıp grafiklerinde bulunan doğruluk değerleri IoU, kayıp değerleri ise Weighted Categorical Cross Entropy (Ağırlıklı Kategorik Çapraz Entropi) metrikleri ile hesaplanmıştır. Ağırlıklı Kategorik Çapraz Entropi işlemi için öncelikle tüm sınıflar için rastgele olarak başlangıç ağırlıkları belirlenmektedir. Yapılan çalışmada başlangıç ağırlıkları tüm sınıflar için 0.1666 olarak atanmıştır. Mimari veriler ile eğitildikçe ağırlıklar güncellenir ve belirli sınıflara daha fazla önem verilerek ağırlıkları yükseltilir ve bu şekilde öne çıkarılan sınıfların bilinmesiyle model performansı arttırabilmektedir.

2.5. Eğitilen Mimarilerin Test Edilmesi

Toplam iki mimariden (Önerilen mimari ve U-net mimarisi) oluşan, farklı sınıf sayıları (iki ve üç) ile eğitilmiş dört adet modelin testi, daha önce bu işlem için ayrılan beş adet 1500×1500 boyutuna sahip uzaktan algılanmış hava görüntüleri üzerinde ve Pürüzsüz Karıştırma (Smooth Blending) yöntemi yardımıyla gerçekleştirilmiştir. Test aşamasında bu yöntemin kullanılmasının en önemli sebebi eğitim işleminde kullanılan görüntü boyutunun katları şeklinde olmayan test görüntülerinin tahmini yapılırken boyuttan kaynaklı bilgi kaybının engellenmesidir. Eğitim işleminde 512×512 boyutunda girdi görüntü parçaları kullanılmış ve bu durumla bağlantılı olarak Pürüzsüz Karıştırma yönteminin test aşamasında kullanılmadığı takdirde, test işleminin uygulandığı 1500×1500 boyutlu hava görüntülerine dair, 512'nin katı olan ve sol üste köşeden başlayacak şekilde toplam 1024×1024 boyutlu bir alan için tahmin işlemi gerçekleştirilebilmiştir. Pürüzsüz karıştırma yönteminin kullandığı test işleminde ise görüntünün tümü için yani 1500×1500 boyutunda tahmin görselinin elde edilmesi sağlanabilmiştir. Klasik tahmin yöntemine nazaran çok daha faydalı olan bu yöntemin kullanımı boyuttan kaynaklı bilgi kaybının önüne geçmiştir.

Beş adet 1500×1500 boyutlu test görüntüsü için daha önce kaydedilmiş; iki ve üç sınıflı olarak eğitilen yeni mimari ve U-Net mimarisi kullanılarak yapılan tahminlerde, karşılaştırma yapabilmek adına hem klasik yöntem hem de Pürüzsüz Karıştırma yöntemi kullanılmıştır. Klasik yöntem ile yapılan tahmin sonucu ile 1024×1024 boyutunda tahmin görselleri elde edilirken, pürüzsüz karıştırma yöntemi ile görüntünün tamamı için yani 1500×1500 boyutlu tahmin görselleri elde edilmiştir. Bu karşılaştırmaya yönelik örnek görseller Şekil 41'de gösterilmiştir.



Şekil 41. Pürüzsüz karıştırmasız ve pürüzsüz karıştırma yöntemi ile yapılan tahmin sonuçlarının karşılaştırılması

Son aşama olan doğruluk analizi işleminde ise 1500×1500 boyutlu tahmin görselleri, ana veride bulunan aynı boyuttaki maskeler ile karşılaştırılarak; OA, Cr, Cp IoU-Q, Genel ve F1-Puanı doğruluk değerlendirme metrikleri, beş adet test görüntüsü için hesaplanmış ve metriklerin sayısal değerleri Bulgular bölümünde sunulmuştur.

3. BULGULAR VE TARTIŞMA

Bu bölümde, uzaktan algılanmış görüntülerden derin öğrenme yöntemi ile bina çıkarımı işleminin gerçekleştirilmesine yönelik tasarlanan mimari ve U-Net mimarisinin kullanımıyla yapılan tüm uygulamaların sonuçları irdelenip, gerekli karşılaştırmalar yapılmıştır. Görüntülerden derin öğrenme yöntemi ile bina çıkarımı işlemi yapılırken, üçüncü bir sınıf olan bina sınırı sınıfının oluşturulma sebebi ve bu sınıfın eklenmesinin bina çıkarımına olan katkıları nitel ve nicel yorumlamalar yardımıyla ifade edilmiştir.

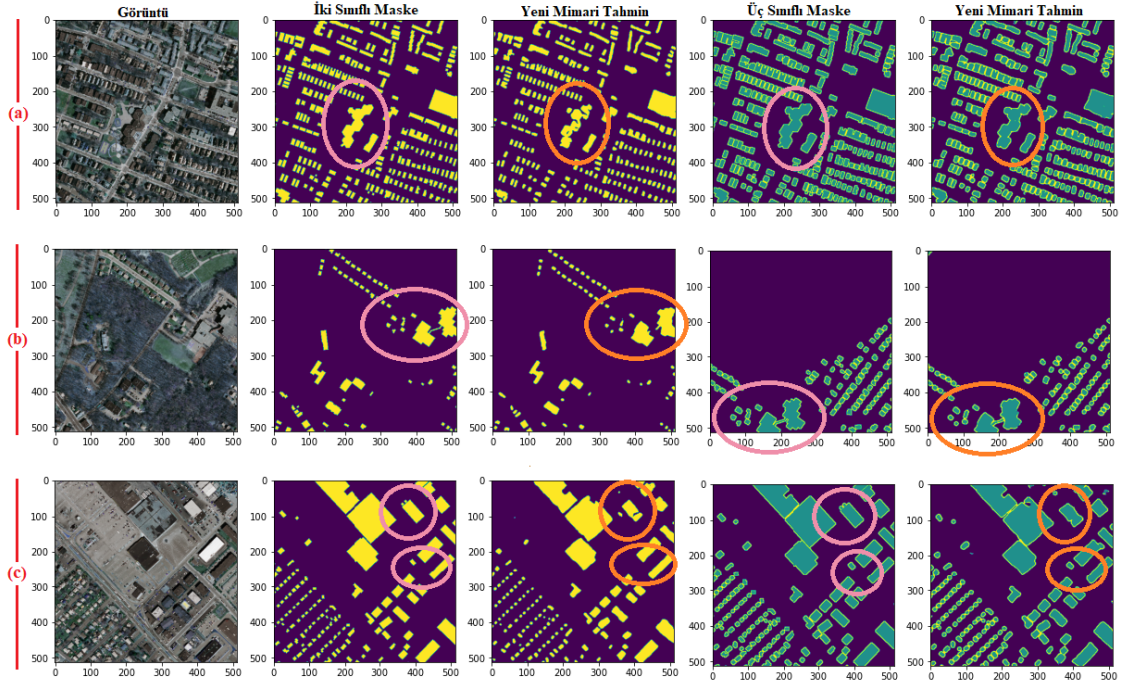
İlave olarak literatürde bulunan ve başarısı kanıtlanmış bir mimari olan U-Net sonuçları, mevcut başarılı derin öğrenme mimarilerine alternatif olması amacıyla çalışma kapsamında tasarlanan ve tüm işlemleri daha kısa süre içerisinde tamamlayabilen yeni mimari sonuçları ile karşılaştırılarak, yapılan çalışma için iki mimarinin birbirlerine göre avantajlı ve dezavantajlı yönleri tartışılmıştır. Tez çalışmasında bahsedilen tüm işlemler, Tablo 5'te sunulan özellikteki bilgisayar ortamında ve ilgili yazılımlar kullanılarak gerçekleştirilmiştir.

Tablo 5. Bilgisayar özellikleri ve kullanılan yazılımlar

İşlemci	Intel(R) Core (TM) i7-7700HQ CPU 2.80GHz
RAM	24 GB
Yazılımlar	Python & MATLAB

3.1. İki Sınıflı ve Üç Sınıflı Segmentasyon ile Bina Çıkarım Sonuçlarının Değerlendirilmesi

Derin öğrenme ile uzaktan algılanmış görüntülerden bina çıkarımı işlemi için, morfolojik operatörler ile sonradan bina maske verilerine eklenen bina sınırı katmanı bilgisinin, derin öğrenme mimarisine sunulmasının bina çıkarım başarısına etkisinin incelenmesi bu başlık altında gerçekleştirilmiştir. Bahsi geçen amaca yönelik olarak 16 katman ile tasarlanan yeni mimarinin, doğrulama verileri için yapılan iki ve üç sınıflı segmentasyon görsel sonuçları karşılaştırılmıştır (Şekil 42). Karşılaştırma yapabilmek için aynı binaların bulunduğu sonuç görseller tercih edilmiştir.

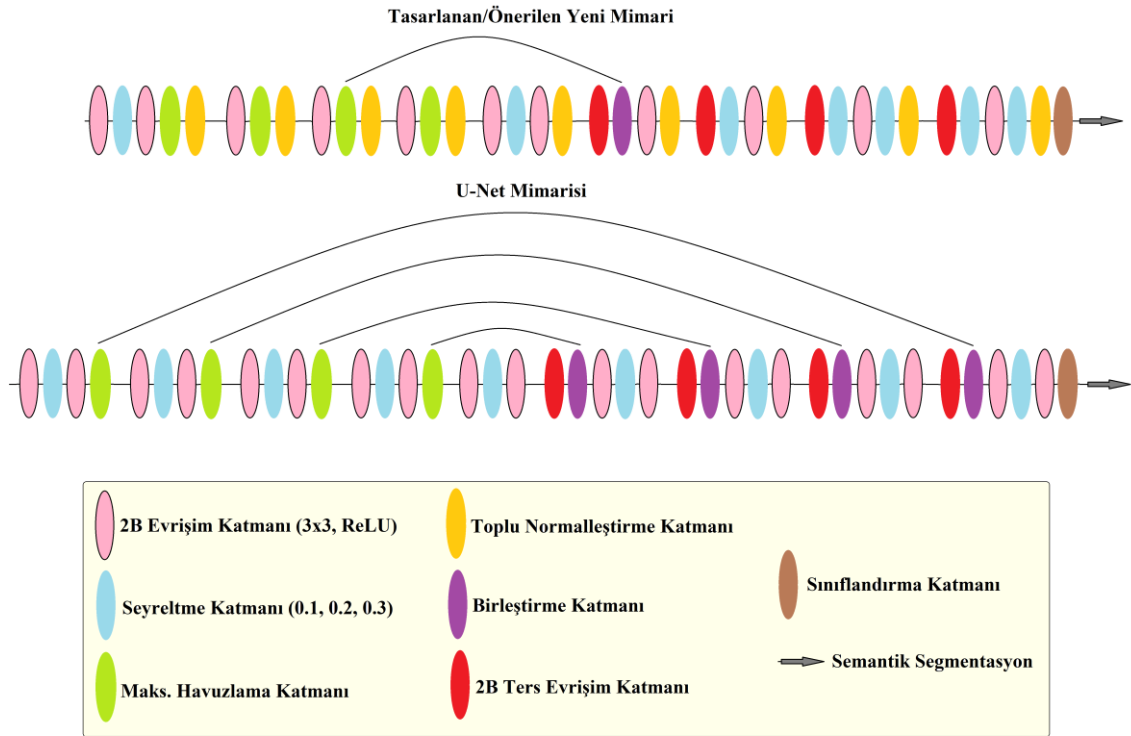


Şekil 42. Doğrulama verileri üzerinde önerilen mimari ile yapılan iki ve üç sınıflı segmentasyon sonuçlarının karşılaştırılması

Şekil 42’de sunulan iki ve üç sınıflı segmentasyon sonuçlarına yönelik karşılaştırma görseli incelendiğinde, Şekil 42 (a) örneğinde bulunan doğrulama görüntüsünde işaretli binanın, iki sınıf ile eğitilen yeni mimari ile elde edilen tahmininde bina içi süreksizliklerin olduğu ve bu sorunun, bina sınır bilgisinin dahil edildiği üç sınıf ile eğitilen yeni mimari tahmin görselinde giderildiği net bir biçimde gözlemlenmektedir. Bir diğer örnek olan (b) örneğinde görüntüde vurgulanan binada ise iki sınıf ile eğitilen mimari tahmininde binalar arasındaki bağlantının tespit edilemediği, üç sınıf ile eğitilen mimari tahmini ile bu bağlantının net bir şekilde ortaya çıkarılabildiği ortaya konulmuştur. Son örnek olan (c) örneğinde bulunan görselde iki sınıf ile eğitilen mimari tahmininde üç farklı sorun tespit edilmiştir. Bunlar; bina olmayan bölgede bir bina algılanması, bina içi sürekliliğin sağlanamaması ve sağ alt kısımda vurgulanan küçük binanın ihmal edilmesidir. Üç sınıf ile eğitilen mimari tahmininde bu üç sorun büyük ölçüde giderilmiştir. Tüm bu örneklerin incelenmesi sonucu sınır bilgisinin dahil edildiği segmentasyon sonucunun uzaktan algılanmış görüntülerden bina çıkarımı için daha başarılı etkiler oluşturduğu kanısına varılmıştır.

3.2. Önerilen Mimari ve U-Net Mimari Sonuçlarının Karşılaştırılması

Tez çalışması dahilinde geliştirilen yeni mimarinin etkinliğini inceleyebilmek adına, karşılaştırma yapmak için derin öğrenme yöntemleri ile görüntü segmentasyonu işlemlerinde sıkça kullanılan ve başarısı kanıtlanmış bir mimari olması sebebiyle U-Net tercih edilmiştir (Şekil 43). Bu bölümde öncelikle mimarilerde kullanılan parametreleri, katman sayısı ve yapılan işlemlerin gerçekleştirme süreleri gibi teknik özellikler karşılaştırılarak Tablo 6’da sunulmuştur.

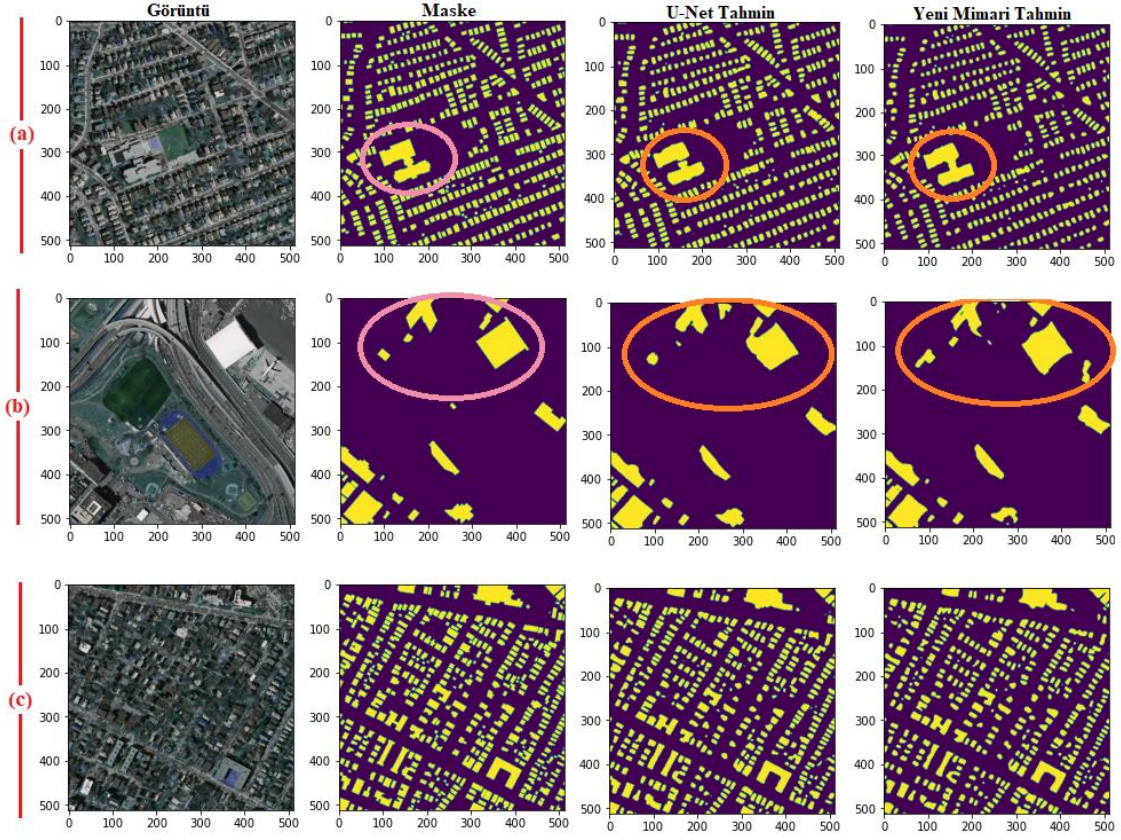


Şekil 43. Tasarlanan mimari ve U-Net mimarisi

Tablo 6. Tasarlanan mimari ve U-Net mimarisi teknik karşılaştırma

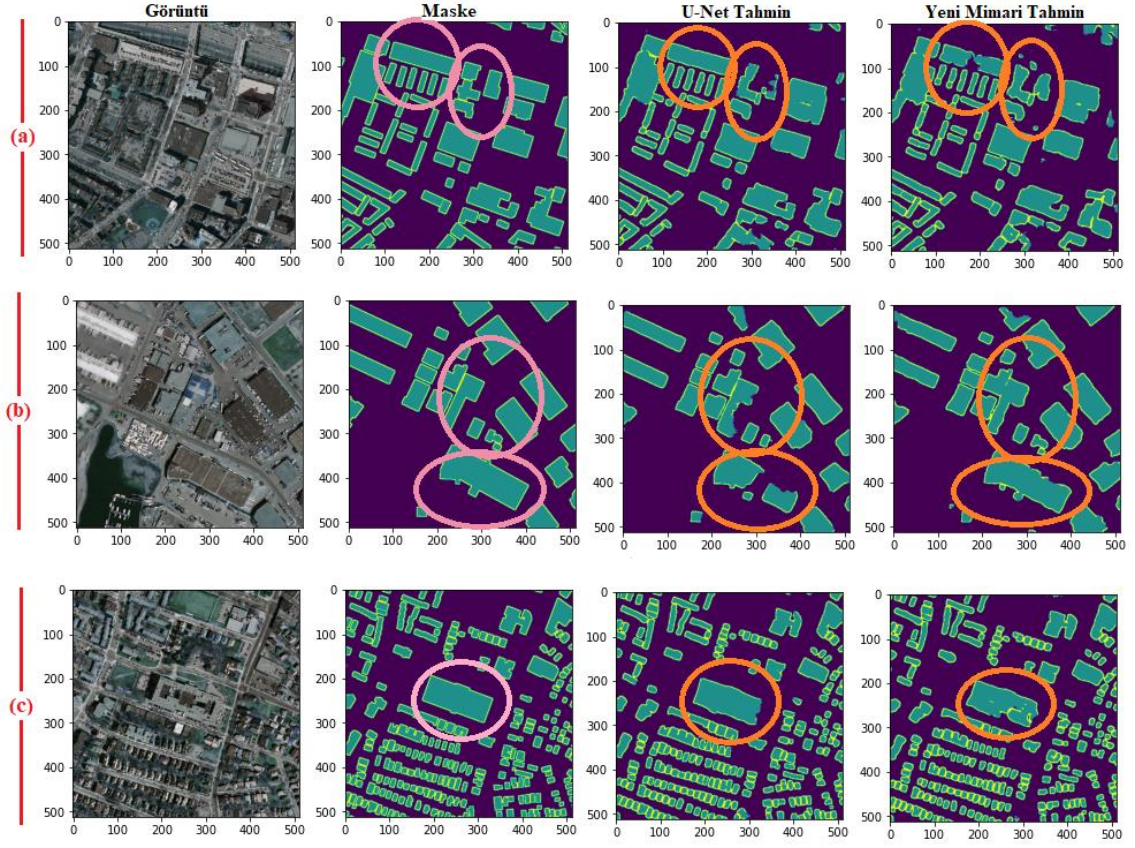
Mimari	İki Sınıflı		Üç sınıflı	
	U-Net	Tasarlanan Mimari	U-Net	Tasarlanan Mimari
Katman Sayısı	23	16	23	16
İşlem Süresi (dk)	1506	1130	1695	1318
Epok Sayısı	100	100	100	100
Öğrenme Hızı	0.001	0.001	0.001	0.001
Parti Boyutu	8	8	8	8
Optimize Edici	Adam	Adam	Adam	Adam
Aktivasyon Fonksiyonu	ReLU/Sigmoid	ReLU/Sigmoid	ReLU/Softmax	ReLU/Softmax
Ortalama IoU	0.8604	0.8526	0.7387	0.7439

Tablo 6'daki verilerde görüldüğü üzere U-Net ve tasarlanan mimari aynı şartlar altında eğitim işlemine tabi tutulmuştur. Bu şartlar her iki mimarinin de tespit edilebilen en yüksek performansı gösterebildiği ayar parametrelerinden oluşmaktadır. Her iki mimari adına doğrulama verisi için hesaplanan ortalama IoU doğruluk metriği değerlerindeki eşitliğe karşın, eğitim işlemini gerçekleştirme süreleri arasında tasarlanan mimari lehine yaklaşık 400 dakika fark bulunmaktadır. Bu farkın sebebi tasarlanan mimarinin daha az katman sayısı ile işlemleri daha hızlı bir şekilde bitirebilmesidir. Her iki mimari için aynı sonuçların üretilmesi ve işlem süresinin daha kısa sürmesi sebebiyle önerilen mimari, U-Net mimarisine nazaran daha avantajlı bir konumda yer almaktadır. Her iki mimari ile görüntülerden bina çıkarımı uygulaması adına yapılan iki ve üç sınıflı segmentasyon görsel sonuçları karşılaştırmalı olarak Şekil 44 ve 45'te gösterilmiştir.



Şekil 44. İki sınıflı segmentasyon için U-Net mimarisi ve önerilen mimarinin doğrulama verileri üzerindeki tahminleri

Şekil 44'te iki sınıf ile eğitilen U-Net ve yeni mimarinin doğrulama görselleri üzerindeki tahminleri karşılaştırmalı olarak sunulmuştur. Görüntüde ilk olarak (a) örneğindeki hava görüntü parçasında vurgulanan bina için U-Net tahmini incelendiğinde bina içinde bir süreksizlik gözlemlenmiştir. Yeni mimari ise süreksizlik olmaksızın binayı daha doğru bir şekilde tespit edebilmiştir. Bir diğer örnek ise (b) örneğinde vurgulanan geniş bölgedir. Bu bölgede hem U-Net'in hem de yeni mimarinin avantajlı ve dezavantajlı olduğu durumlar mevcuttur. Örneğin U-Net yaptığı tahminde, görselde bulunan kare bina ve üzerindeki küçük binayı bir bütün şeklinde hatalı olarak ortaya çıkarmıştır. Aynı bina için daha doğru bir sonuca erişebilen yeni mimari ise bölgedeki bazı kısımlarda gürültü şeklinde hatalı tespitler yapmıştır. Son görüntünün bulunduğu (c) örneğinde yoğun ve küçük binaların bulunduğu hava görüntüsü parçasında her iki mimari birbirine benzer ve başarılı bir sonuç ortaya çıkarmıştır.



Şekil 45. Üç sınıflı segmentasyon için U-Net mimarisi ve önerilen mimarinin doğrulama verileri üzerindeki tahminleri

Şekil 45'te üç sınıf ile eğitilen U-Net ve yeni mimarinin doğrulama görselleri üzerindeki tahminleri karşılaştırmalı olarak gösterilmiştir. İlk olarak (a) örneğindeki hava görüntü parçasında vurgulanan iki adet bölgedeki binalar incelendiğinde U-Net mimarisinin yaptığı tahminde sağ kısımdaki binanın tespit edilemediği, buna karşın yeni mimarinin ilgili bina için daha doğru bir tahmin gerçekleştirdiği gözlemlenmektedir. Fakat aynı satırda sol bölgede işaretlenen binayı ise U-Net mimarisinin daha iyi temsil edebildiği yeni mimarinin ise bina içi süreksizliklere sebep olduğu fark edilmiştir. Bir diğer örnek ise (b) örneğinde vurgulanan iki adet bölgedir. İlgili görüntü için yapılan bina tahmininde işaretlenen her iki bölgede de yeni mimari, U-Net mimarisine nazaran daha iyi tahminlerde bulunmuştur. İşaretli iki alandan birinde U-Net mimarisi bina sınırını hatalı olarak daha geniş bir şekilde tespit etmiş, diğer alanda ise bina içi süreksizliğe neden olması sebebiyle bina geometrisini ortaya çıkaramamıştır. Son görüntünün bulunduğu (c) örneğinde yoğun miktarda, küçük boyutlara sahip binaların bulunduğu bölgeler için her iki mimari başarılı sonuçlar vermiştir.

Fakat işaretlenen alandaki büyük boyutlu bina için U-Net mimarisi daha doğru bir sınır bilgisi elde etmiştir.

Genel hatları ile ilgili sonuçlar yorumlandığında U-Net ve tasarlanan mimarinin, doğrulama verisi görsellerinde bulunan aynı binalar üzerindeki görsel tahmin sonuçları karşılaştırıldığında genel olarak benzer sonuçlar alınmasının yanı sıra, bazı binaların tespitinde minimal düzeyde farklar olduğu sonucuna ulaşılmıştır.

3.3. Test Görüntüleri İçin Doğruluk Analizi Sonuçlarının Değerlendirilmesi

Toplamda eğitilmiş iki mimariden oluşan, dört adet modelin performansını farklı ve büyük boyutlu görüntüler üzerinde de değerlendirebilmek adına beş adet 1500×1500 boyutlu test görüntüsü üzerinde Pürüzsüz Karıştırma yöntemi yardımıyla test işlemi gerçekleştirilmiştir. Ardından görüntüler üzerinde yapılan tüm tahminlerin doğruluk oranını sayısal olarak belirleyebilmek için aynı boyuttaki maske verileri ile birebir karşılaştırma yapılarak piksel bazlı doğruluk analizi işlemi gerçekleştirilmiştir. Doğruluk analizinde öncelikle tüm görüntüler için maske görüntüleriyle birebir karşılaştırma yapmak suretiyle DP, DN, YP ve YN değerleri hesaplanarak, Genel Doğruluk (OA), Doğruluk (Cr), Tamlık (Cp), F1-Puanı (F1) ve Birleşimlerin Kesişimi-Kalite (IoU-Q) doğruluk metrikleri hesaplanmıştır. Hesaplanan doğruluk metriklerinin iki ve üç sınıflı segmentasyon için her iki mimarideki karşılaştırmalı sonuçları Tablo 7 ve Tablo 8’de sunulmuştur.

Tablo 7. U-Net mimarisi ve tasarlanan mimarinin iki sınıflı segmentasyon için doğruluk değerlendirme sonuçları

Doğruluk Metrikleri	Segmentasyon	Test-1	Test-2	Test-3	Test-4	Test-5
OA	U-Net	0.9208	0.9266	0.9399	0.9011	0.8694
	Yeni Mimari	0.9124	0.9220	0.9394	0.8930	0.8495
Cr	U-Net	0.8243	0.8068	0.8559	0.8224	0.9113
	Yeni Mimari	0.8657	0.8382	0.8905	0.8475	0.9305
Cp	U-Net	0.6888	0.6237	0.6542	0.6858	0.7025
	Yeni Mimari	0.5843	0.5470	0.6158	0.6099	0.6250
F1	U-Net	0.7505	0.7035	0.7416	0.7479	0.7934
	Yeni Mimari	0.6977	0.6620	0.7281	0.7093	0.7477
IoU-Q	U-Net	0.6006	0.5426	0.5893	0.5973	0.6576
	Yeni Mimari	0.5357	0.4947	0.5724	0.5495	0.5971

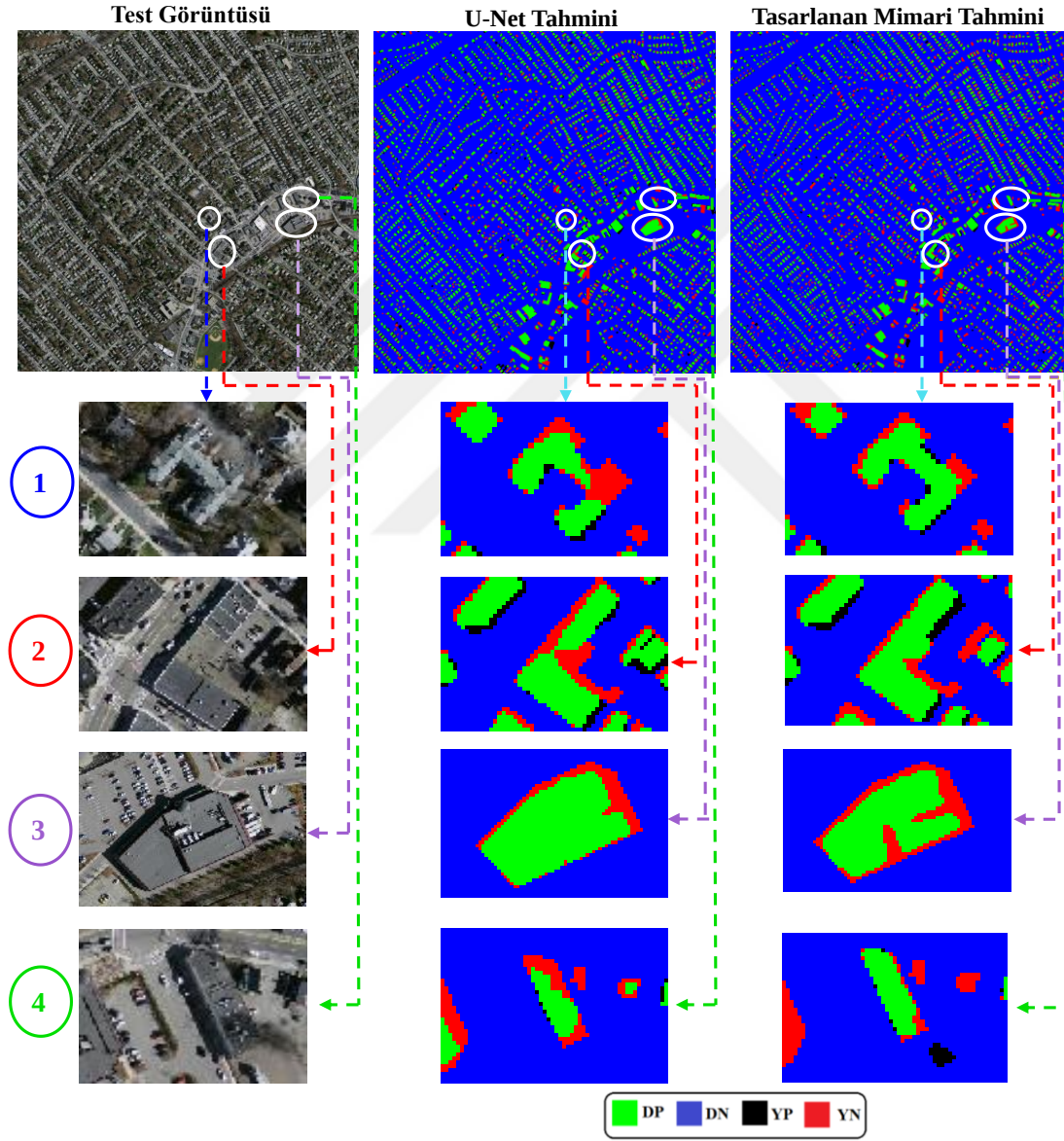
Tablo 8. U-Net mimarisi ve tasarlanan mimarinin üç sınıflı segmentasyon için doğruluk değerlendirme sonuçları

Doğruluk Metrikleri	Segmentasyon	Test-1	Test-2	Test-3	Test-4	Test-5
OA	U-Net	0.9220	0.9279	0.9404	0.9027	0.8833
	Yeni Mimari	0.9173	0.9259	0.9454	0.8978	0.8739
Cr	U-Net	0.7772	0.7448	0.7996	0.7674	0.8644
	Yeni Mimari	0.8010	0.7602	0.8316	0.7802	0.8321
Cp	U-Net	0.7694	0.7361	0.7672	0.7825	0.7982
	Yeni Mimari	0.6943	0.6858	0.7343	0.7275	0.8101
F1	U-Net	0.7733	0.7404	0.7831	0.7748	0.8300
	Yeni Mimari	0.7438	0.7211	0.7800	0.7230	0.8210
IoU-Q	U-Net	0.6304	0.5878	0.6435	0.6324	0.7094
	Yeni Mimari	0.5921	0.5638	0.6393	0.6038	0.6963

İki sınıf ile eğitilen yeni mimari ile 5 adet test görüntüsü üzerinde yapılan tahminlerin Genel Doğruluk sonuçları sırasıyla; **0.9124, 0.9220, 0.9394, 0.8930, 0.8495** ve U-Net mimarisi için ise sırasıyla; 0.9208, 0.9266, 0.9399, 0.9011, 0.8694 oranları şeklinde elde edilmiştir. Bunun yanı sıra üç sınıf ile eğitilen yeni mimarinin aynı test görüntüleri üzerinde yaptığı tahminlerin Genel Doğruluk değerleri sırasıyla; **0.9173, 0.9259, 0.9454, 0.8978, 0.8739** ve U-Net mimarisi için ise sırasıyla; 0.9220, 0.9279, 0.9404, 0.9027, 0.8833 olarak elde edilmiştir.

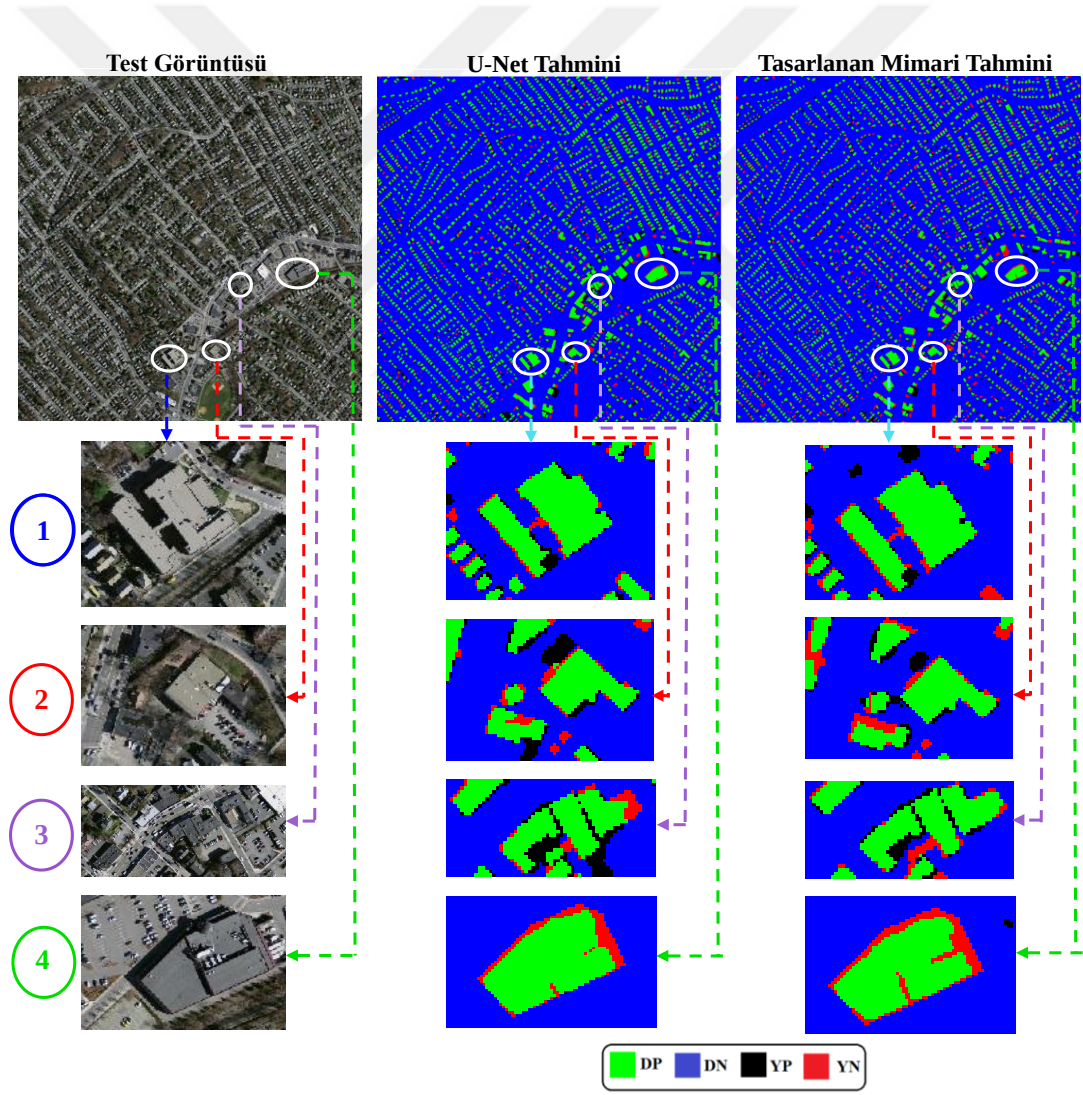
Tablo 7 ve 8’de test görüntüleri üzerinden bina çıkarımı için U-Net ve yeni mimariye ait karşılaştırmalı iki ve üç sınıflı segmentasyon sonuçları incelendiğinde; 23 katmanlı U-Net mimarisi ve 16 katmanlı yeni mimarinin birbirlerine çok yakın doğruluk sonuçları

verdiği gözlemlenmektedir. Ayrıca Genel Doğruluk (OA) metriğinin diğer doğruluk metriklerine kıyasla daha yüksek doğruluk oranı sonucu verdiği görülmektedir. Bu durumun temel sebebi; Genel Doğruluk (OA) metriğinde DP, DN, YP ve YN değerlerinin tümünün işleme katılmasıdır. İki ve üç sınıflı olarak eğitilen U-Net mimarisi ve yeni mimariden elde edilen modellerin bir adet test görüntüsü üzerindeki görsel tahmin sonuçları Şekil 46 ve Şekil 47’de gösterilmiştir.



Şekil 46. İki sınıflı eğitilen U-Net mimarisi ve tasarlanan yeni mimarinin test görüntüsü üzerindeki sonuçları

Eğitilen mimarilerin test edilmesi için Şekil 46'daki örnek test görüntüsünde de olduğu gibi yoğun miktarda ve farklı boyutlardaki binaların bulunduğu hava fotoğrafları tercih edilmiştir. İki sınıflı olarak eğitilen yeni mimari ve U-Net mimarisinin Şekil 46'da bulunan test görüntüsü üzerindeki tahmin sonuçları incelendiğinde her iki mimarinin yaklaşık olarak aynı sonuçları verdiği gözlemlenmektedir. Hava görüntüsünde en fazla farklılığın mevcut olduğu dört adet bölge seçilerek işaretlenmiştir. Seçilen bina örnekleri incelendiğinde yeni mimarinin binaları daha iyi bir şekilde temsil ettiği gözlemlenmektedir. Fakat bu duruma karşın üçüncü örnekte, yeni mimarinin gölgeleri bina sınırı olarak algılaması sebebiyle, tasarlanan mimarinin gölgeleri ayırt etme yeteneğinin, U-Net mimarisine göre daha zayıf kaldığı sonucuna varılmıştır.



Şekil 47. Üç sınıflı eğitilen U-Net mimarisi ve tasarlanan yeni mimarinin test görüntüsü üzerindeki sonuçları

Üç sınıflı olarak eğitilen yeni mimari ve U-Net mimarisinin test görüntüsü üzerindeki bina tahmin sonuçları Şekil 47’de sunulmuştur. İşaretlenerek vurgulanan dört adet binaya dair sonuçlar genel hatları ile incelendiğinde birinci örnekte her iki mimariden yaklaşık olarak benzer sonuçlar alınmıştır. Fakat ikinci ve üçüncü örneklerdeki binalar için U-Net tahmininde, yoğun miktarda YP etiketi tespit edilmiştir. Bu durum U-Net mimarisinin bina sınırlarını yeni mimariye göre daha hatalı bir biçimde tespit etmesine neden olmaktadır. Dördüncü örnekte ise Şekil 46’da belirtilen, gölgelerin bina sınırı olarak algılanması sorunu iki mimari için de bir problem olarak tekrar ortaya çıkmıştır.

3.4. Test Görüntüleri Üzerindeki İki ve Üç Sınıflı Segmentasyon Doğruluk Değerlendirme Sonuçlarının Karşılaştırılması

Bu bölümün amacı veri setinde bulunan maske görüntülerindeki bina ve arka plan sınıflarına ek olarak bina sınırı sınıfının eklenmesi durumunun uzaktan algılanmış görüntülerden derin öğrenme yöntemleri ile bina çıkarma işlemi için sağladığı katkıyı araştırmaktır. Bu hedef doğrultusunda tasarlanan yeni mimarinin iki ve üç sınıf ile eğitilmesiyle ortaya çıkan modeller vasıtasıyla, 5 adet test görüntüsü üzerinde yapılan tahminler arasındaki farklar incelenmiştir. Yeni mimari ile iki ve üç sınıflı olarak eğitilen modellerin test görüntüleri üzerindeki tahmin sonuçlarının doğruluk değerleri Tablo 9’da verilmiştir. Bununla birlikte üç sınıf ile eğitilen yeni mimarinin iki sınıf ile eğitilen mimariye göre, test görüntülerinde bina tahmini üzerinde sağladığı ortalama pozitif ve negatif katkılar tüm doğruluk metrikleri bakımından hesaplanarak Tablo 10’da listelenmiştir.

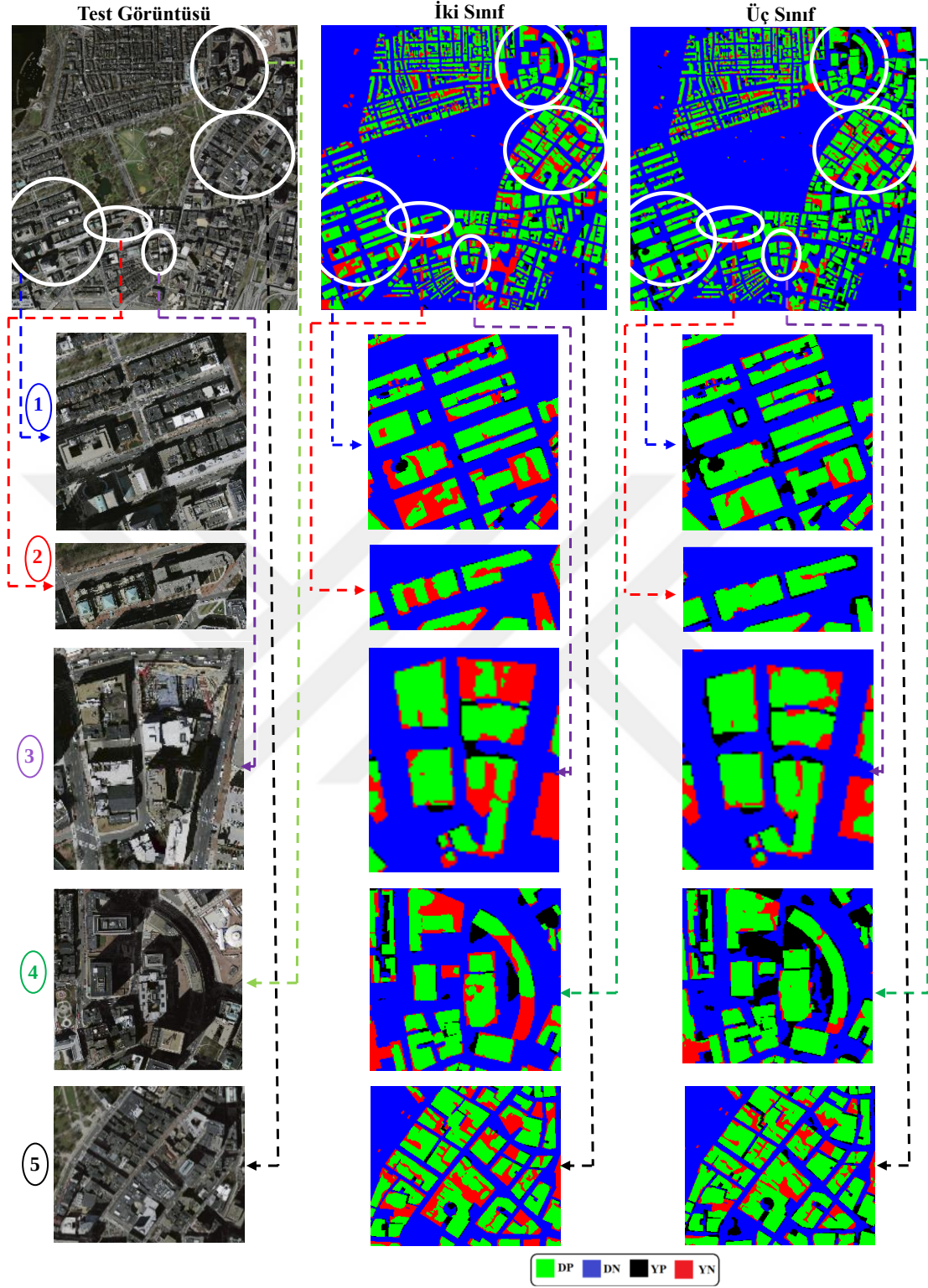
Tablo 9. Önerilen mimari ile test görüntüleri üzerindeki iki ve üç sınıflı segmentasyon sonuçlarının karşılaştırılması

Doğruluk Metrikleri	Segmentasyon	Test-1	Test-2	Test-3	Test-4	Test-5
OA	2 Sınıflı	0.9124	0.9220	0.9394	0.8930	0.8495
	3 Sınıflı	0.9173	0.9259	0.9454	0.8978	0.8739
Cr	2 Sınıflı	0.8657	0.8382	0.8905	0.8475	0.9305
	3 Sınıflı	0.8010	0.7602	0.8316	0.7802	0.8321
Cp	2 Sınıflı	0.5843	0.5470	0.6158	0.6099	0.6250
	3 Sınıflı	0.6943	0.6858	0.7343	0.7275	0.8101
F1	2 Sınıflı	0.6977	0.6620	0.7281	0.7093	0.7477
	3 Sınıflı	0.7438	0.7211	0.7800	0.7230	0.8210
IoU-Q	2 Sınıflı	0.5357	0.4947	0.5724	0.5495	0.5971
	3 Sınıflı	0.5921	0.5638	0.6393	0.6038	0.6963

Tablo 10. Test görüntüleri üzerinde üç sınıf ile eğitilmiş önerilen mimarinin, iki sınıf ile eğitilmiş mimariye göre doğruluk oranı değişimi

OA	Cr	Cp	F1	IoU-Q
+ %0.880	- %7.346	+ %13.400	+ %2.882	+ %6.918

Tablo verilerinin incelenmesiyle aynı şartlar altında yeni mimarinin kullanılmasıyla yapılan eğitimlerde, bina sınır bilgisinin eklenmiş olduğu üç sınıf ile eğitilen mimarinin iki sınıf ile eğitilen mimariye karşın bina çıkarımı uygulamasında; Genel Doğruluk, Tamlık, F1-Puanı ve Birleşimlerin Kesişimi-Kalite metrikleri için sırasıyla %0.880, %13.400, %2.882 ve %6.918 oranlarında bir artış, Doğruluk metriğinde ise %7.346 oranında bir düşüş gözlemlenmiştir. En genel haliyle bina sınır bilgisinin dahil edildiği üç sınıflı segmentasyonun bina çıkarım hususunda iki sınıflı segmentasyona karşın daha fazla başarı sağladığı kanısına varılmıştır. Bahsi geçen durum, örnek bir test görüntüsü üzerinde yapılan tahmin sonuçlarının hata haritaları kullanılarak, farkın en büyük olduğu bölgeler işaretlenerek Şekil 48’de gösterilmiştir.



Şekil 48. Tasarlanan mimarinin test görüntüsü üzerindeki iki ve üç sınıflı segmentasyonu

Şekil 48’de sunulan görselde yeni mimarinin iki ve üç sınıf ile eğitilmesi sonucu örnek bir test görüntüsü üzerinde yapılan bina tahminleri sunulmuş ve farkın en fazla olduğu 5 adet

bölge işaretlenerek ön plana çıkarılmıştır. Farklı sınıf sayısı ile eğitilen mimarının testi için yoğun sayıda, farklı nizam ve mimarilere sahip binaların bulunduğu ve gölgelik alanların fazla olduğu bir test görüntüsü tercih edilmiştir.

Test görüntüsünden alınan tahmin sonuçları incelendiğinde bina sınır bilgisinin dahil edilmesi ile eğitilen mimarının, sınır bilgisinin dahil edilmemesi durumunda eğitilen mimariye kıyasla bina çıkarma konusunda daha başarılı olduğu gerçeği ortaya konulmuştur. İşaretli bölgelerde üç sınıf ile eğitilen mimarının bina içi sürekliliği daha iyi bir şekilde sağlayarak binaları daha net bir şekilde ortaya çıkarabildiği gözlemlenmiştir. Fakat bu olumlu gelişmelere karşın birinci ve dördüncü örnekte gözlemlenebildiği üzere üç sınıf ile eğitilen mimari gölgelik alanları da bina sınıfı olarak sınıflandırmıştır. Gölgelemlerin yoğun olmadığı bölgelerde bina çıkarma hususunda üç sınıf ile eğitilen mimarının daha başarılı olduğu sonucuna varılmıştır.

4. SONUÇLAR

Yüksek lisans tez çalışması kapsamında Massachusetts veri seti kullanılarak, uzaktan algılanmış görüntülerden bina çıkarımı konusunda yapılmış derin öğrenme temelli çalışmada, literatürde mevcut ve görüntü bölütleme uygulamalarında sıklıkla kullanılan derin öğrenme ağlarına alternatif olabilecek 16 katmanlı bir mimari tasarlanmıştır. Bu çalışmada tasarlanan mimariye ek olarak bir diğer özgün değer ise kullanılan veri setindeki ikili (binary) formata sahip maske verilerinde bulunan bina temsillerine morfolojik operatörler yardımıyla iki piksel genişliğinde bina sınırı çizdirilmesi ile hava görüntülerinden binaların çıkarımı hususundaki başarının artırılması hedefidir. Bu hedef doğrultusunda tasarlanan mimari, Massachusetts veri setinde bulunan görüntülerin %50 bindirmeli olarak 512×512 boyutunda parçalanması ile hem iki sınıflı (bina, arka plan) hem de üç sınıflı (bina, arka plan ve bina sınırı) maske verileri kullanılarak ayrı ayrı eğitilmiştir. Ardından tasarlanan mimarinin performansını karşılaştırmak maksadıyla literatürde sıkça kullanılmış ve başarısı kanıtlanmış bir mimari olan 23 katmanlı U-Net mimarisi de aynı veri setleriyle iki ve üç sınıflı olarak eğitilmiştir. Yapılan eğitimler sonucunda doğrulama verileri için ortalama Birleşimlerin Kesişimi doğruluk metriği sonuçları, iki sınıf ile eğitilen U-Net ve yeni mimari için sırasıyla 0.8604, 0.8526 ve üç sınıf ile eğitilen U-Net ve yeni mimari için 0.7387, 0.7439 şeklinde elde edilmiştir. Ayrıca yeni mimarinin katman sayısının daha az olması sebebiyle iki ve üç sınıflı olarak yapılan eğitimleri U-Net'e göre yaklaşık 400 dakika daha az süreyle bitirmeyi başarmıştır.

Eğitilen mimarilerin, veri setinde sunulan bilgileri ne derecede öğrendiklerini belirleyebilmek adına daha önce veri seti içerisinden seçilerek eğitime dahil edilmeyen 5 adet 1500×1500 boyutunda test görüntüleri üzerinde, kaydedilen 4 adet model ile Pürüzsüz Karıştırma yöntemi vasıtasıyla bina çıkarım hususunda tahminler yapılmıştır. Eğitilen 4 mimari dahilinde teste tabi tutulan 5 adet görsel için Genel Doğruluk (OA), Doğruluk (Cr), Tamlik (Cp), F1-Puanı (F1) ve Birleşimlerin Kesişimi-Kalite (IoU-Q) doğruluk metrikleri hesaplanmıştır. Test görüntüleri üzerindeki ortalama Genel Doğruluklar iki sınıf ile eğitilen yeni mimari ve U-Net mimarisi için sırasıyla 0.9033, 0.9115; üç sınıf ile eğitilen yeni mimari ve U-Net mimarisi için ise sırasıyla 0.9121, 0.9152 değerleri şeklindedir.

Sonuç olarak tüm bu doğruluk değerleri incelendiğinde tasarlanan mimarinin U-Net mimarisi ile aynı düzeyde performans göstermesi ve eğitim işlemini daha hızlı

gerçekleştirebilmesi sebebiyle literatürdeki başarılı mimarilere alternatif olması konusunda umut vadetmektedir.

Ayrıca üç sınıflı olarak eğitilen mimarilerin test görüntüleri üzerinden bina çıkarım konusunda, iki sınıflı eğitilen mimarilere göre daha başarılı olduğu doğruluk analizi sonuçlarıyla kanıtlanmaktadır. Bunun anlamı; görüntülerden bina çıkarımı için yapılan derin öğrenme temelli eğitim uygulamalarında, veride bulunan maske görüntülerindeki binaların çevresine 2 piksel genişliğinde bina sınırı çizdirilmesinin bina çıkarım doğruluğuna katkıda bulunacağıdır.



5. ÖNERİLER

Uzaktan algılanmış hava görüntülerinden bina çıkarımı maksadıyla derin öğrenme temelli olarak yapılan çalışma kapsamında tasarlanan mimari, Massachusetts veri seti ile eğitilerek, test görüntüleri üzerinde yapılan tahminler ile binalar başarılı bir şekilde elde edilmiştir. Fakat yapılan çalışmanın daha genelleştirilebilir olması bakımından, farklı bölgelerden, şehirlerden elde edilmiş ve daha fazla miktarda hava görüntülerinden oluşan büyük boyutlu veri setlerinin kullanımı bina çıkarımı uygulamasındaki başarıyı artırma hususunda gereklilik arz etmektedir. Ayrıca yapılan çalışmada Massachusetts veri seti ile eğitilerek, bu veri seti dışında başka bir bölgeye ait bir hava fotoğrafı üzerinde bina tahmini işlemi gerçekleştirilmek istendiğinde hedeflenen başarı elde edilememiştir. Bu durum kullanılan mimarinin başarısı ile ilgili olmayıp, eğitim veri setinin sadece belirli bölgelerden (Boston şehri) elde edilmiş hava fotoğraflarından oluşması durumu ile ilişkilidir. Bu konunun üstesinden gelmek üzere yapılan araştırmalar ve denemeler sonucunda, üzerinde bina tahmini yapılmak istenilen hava fotoğrafından birkaç yama elde edilerek, bu yama görüntülerin eğitim veri setine eklenmesi ile istenilen bir bölgeden elde edilen hava fotoğrafı üzerinde yapılan, bina tahmini başarısının büyük oranda sağlanacağı öngörülmüştür.

Gerçekleştirilen çalışmanın eleştiriye açık bir diğer yönü ise veri seti içerisinde eğitim ve doğrulama işlemi için ayrılan görüntü parçalarının veri setindeki görüntü sayısının arttırmak amacıyla enine ve boyuna olmak üzere %50 bindirme oranı ile parçalara ayrılmış olmasıdır. Buna bağlı olarak eğitim işleminde her binanın birden fazla kez öğrenme işlemine tabi tutulması ile mimarinin aşırı öğrenmeye maruz kalma riski bulunmaktadır. Daha fazla hava görüntüsüne sahip veri setlerine, daha düşük enine ve boyuna bindirme oranlarının uygulanmasıyla bahsi geçen negatif durumun önüne geçilebileceği öngörülmektedir.

Yapılan tez çalışmasından elde edilen çıktıları dayanarak, derin öğrenme mimarileri kullanarak hava görüntülerinden elde edilen binaların, projelerde ve büyük ölçekli haritalarda direkt olarak kullanılamayacağı sonucuna varılmıştır. Bu durumun ana sebebi; derin öğrenme mimarilerinden çıktı olarak alınan tahmin bina sınırlarının genellikle düzgün olmayıp, zikzaklı bir yapıda olmasıdır. Bahsedilen problemin üstesinden gelme konusunda ise derin öğrenme mimarileri ile elde edilen bina tahminlerinin projelerde ve büyük ölçekli haritalarda kullanımı sağlanmadan önce, bina sınırlarına bir düzgünleştirme işlemi uygulanarak gerçeğe en yakın geometrideki sınırlar elde edilmelidir.

Yapılan gözlemler ve arařtırmalar sonucunda elde edilen tecrübeler ile oluşturulan önerilerin sağlanması durumunda hava görüntülerinden derin öğrenme yöntemleri ile binaların çıkarımı konusunda istenilen başarı ve hedeflenen amaçların ulaşılabilirliği sağlanacaktır.



6. KAYNAKLAR

- Aber, J.S., Marzolff, I., & Ries, J.B., (2010). Small-Format Aerial Photography. *Elsevier Science*, Amsterdam.
- Abdollahi, A., & Pradhan, B. (2021). Integrating Semantic Edges And Segmentation Information For Building Extraction From Aerial Images Using UNet. *Machine Learning with Applications*, 6, 1-10.
- Abriha, D., & Szabó, S. (2023). Strategies in training deep learning models to extract building from multisource images with small training sample sizes. *International Journal of Digital Earth*, 16(1), 1707–1724.
- Acar, U. & Bayram, B. (2009). Morfolojik Görüntü Filtreleri ile İkonos Görüntülerinden Otomatik Bina Çıkarımı. TMMOB Harita ve Kadastro Mühendisleri Odası 12. Türkiye Harita Bilimsel ve Teknik Kutultayı, Ankara.
- Ackley, D. H., Hinton, G. E., & Sejnowski, T. J. (1985). A Learning Algorithm for Boltzmann Machines. *Cognitive Science* 9, 147-169.
- Ahmady, S., Ebadi, H., Zouj, M.J., Moghaddam, H.A., & Toosi, K.N. (2008). Automatic Building Extraction From High Resolution Aerial Images Using Active Contour Model.
- Ahmadi, S., Zoej, M. J. V., Ebadi, H., Moghaddam, H. A., & Mohammadzadeh, A. (2010). Automatic Urban Building Boundary Extraction From High Resolution Aerial Images Using An Innovative Model Of Active Contours. *International Journal of Applied Earth Observation and Geoinformation*, 12(3), 150–157.
- Akbulut, Z., Özdemir, S., Acar, H., Dihkan, M., & Karsli, F. (2018). Automatic Extraction of Building Boundaries from High Resolution Images with Active Contour Segmentation. *International Journal of Engineering and Geoscience (IJEG)*, 3(1), 36-42.
- Albawi, S., Mohammed, T. A., & Al-Zawi, S. (2017). Understanding of a Convolutional Neural Network. *2017 International Conference on Engineering and Technology (ICET)*, 1–6.
- Awrangjeb, M., Zhang, C., Fraser, C.S. (2011). Improved Building Detection Using Texture Information. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38, 143–148.
- Bardis, M., Houshyar, R., Chantaduly, C., Ushinsky, A., Glavis-Bloom, J., Shaver, M., Chow, D., Uchio, E., & Chang, P. (2020). Deep Learning with Limited Data: Organ Segmentation Performance by U-Net. *Electronics*, 9(8), 1199.
- Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I., & Heynen, M. (2004). Multi-Resolution, Object-Oriented Fuzzy Analysis of Remote Sensing Data for GIS-Ready

- Information. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(3-4), 239-258.
- Blaschke, T. (2010). Object Based Image Analysis for Remote Sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1), 2-16.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5–32.
- Bridle, J. (1989). Training Stochastic Model Recognition Algorithms as Networks can Lead to Maximum Mutual Information Estimation of Parameters. *Proceedings of the 2nd International Conference on Neural Information Processing Systems*, 211–217
- Canny, J. (1986). A Computational Approach to Edge Dedection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8, 679-700.
- Chaudhuri, D., Kushwaha, N. K., Samal, A., & Agarwal, R. C. (2016). Automatic Building Detection From High-Resolution Satellite Images Based on Morphology and Internal Gray Variance. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(5), 1767–1779.
- Chauhan, R., Ghanshala, K. K., & Joshi, R. C. (2018). Convolutional Neural Network (CNN) for Image Detection and Recognition. *2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)*, 278–282.
- Chen, Q., Wang, L., Wu, Y., Wu, G., Guo, Z., & Waslander, S. L. (2019). TEMPORARY REMOVAL: Aerial imagery for roof segmentation: A large-scale dataset towards automatic mapping of buildings. *ISPRS Journal of Photogrammetry and Remote Sensing*, 147, 42-55.
- Chollet, F. (2021). *Deep learning with Python*. Manning Publications.
- Coşkun, F. & Gülleroğlu, H. D. (2021). Yapay Zekanın Tarih İçindeki Gelişimi ve Eğitimde Kullanılması. *Ankara University Journal of Faculty of Educational Sciences (JFES)*, 54 (3), 947-966.
- Çölkesen, İ. & Kavzoğlu, T. (2017). Kanonik Korelasyon Orman Algoritması ile Uzaktan Algılanmış Görüntülerin Sınıflandırılması. *Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, 17(1), 102-111
- Duda, R. O., & Hart, P. E. (1972). Use of the Hough Transformation to Detect Lines and Curves in Pictures. *Communications of the ACM*, 15, 11-15.
- Erol, O., & Öner, E. Fiziki Coğrafyada Hava Fotoğrafları (Fotojeoloji-Fotojeomorfoloji). *Ege Üniversitesi Yayınları*, 171, İzmir.
- Fagan, M. E., & DeFries, R. S. (2022). *Remote Sensing and Image Processing*. Reference Module in Life Sciences.
- Feng, D., Chen, H., Xie, Y., Liu, Z., Liao, Z., Zhu, J., & Zhang, H. (2022). GCCINet: Global Feature Capture and Cross-Layer Information Interaction Network for Building

- Extraction from Remote Sensing Imagery. *International Journal of Applied Earth Observation and Geoinformation*, 114, 103046.
- Gavankar, N. L., & Ghosh, S. K. (2018). Automatic Building Footprint Extraction From High-Resolution Satellite Image Using Mathematical Morphology. *European Journal of Remote Sensing*, 51(1), 182–193.
- Gomarasca, M. A. (2009). *Elements of Photogrammetry*. In Basics of Geomatics, 79-121.
- Gonzalez, R. C., & Woods, R. E. Digital Image Processing. *II. Edition, Prentice Hall*, 519-560 & 617-621.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*
- Gupta, A., Nair, S., Joshi, R., & Chitre, V. (2022). Residual-Concatenate Neural Network with Deep Regularization Layers for Binary Classification.
- He, K., Gkioxari, G., Dollar, P., & Girshick, R. (2020). Mask R-CNN. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(2), 386–397.
- Hinton, G. E., Osindero, S., & Teh, Y.W. (2006). A Fast Learning Algorithm for Deep Belief Nets. *Neural Computation*, 18, 1527-1554.
- Horswell, J. (2013). *Recording*. Encyclopedia of Forensic Sciences, 368-371.
- Hu, Q., Zhen, L., Mao, Y., Zhou, X., & Zhou, G. (2021). Automated Building Extraction Using Satellite Remote Sensing Imagery. *Automation in Construction*, 123, 103509.
- Huang, X., Zhang, L., 2011. A Multidirectional and Multiscale Morphological Index for Automatic Building Extraction from Multispectral Geoeye-1 Imagery. *Photogrammetric Engineering And Remote Sensing*, 77 (7), 721–732.
- Huang, X., Zhang, L., 2012. Morphological Building/Shadow Index for Building Extraction From High-Resolution Imagery Over Urban Areas. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* 5, 161–172.
- Ioffe, S., & Szegedy, C. (2015). Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.
- Ji, S., Shen, Y., Lu, M., & Zhang, Y. (2019). Building Instance Change Detection from Large-Scale Aerial Images using Convolutional Neural Networks and Simulated Samples. *Remote Sensing*, 11(11), 1343.
- Kale, R. S., & Thorat, D. S. (2021). Image Segmentation Techniques with Machine Learning. *International Journal of Scientific Research in Computer Science Engineering and Information Technology*, 7(6), 232-235.
- Kass, M., Witkin, A., & Terzopoulos, D. (1988). Snakes: Active Contour Models. *International Journal of Computer Vision*, 321-331.
- Kingma, D., & Ba, J. (2014) Adam: A method for stochastic optimization. *Comput Sci*.

- Kiracı, A. C., İşcan, L., Eker, O., Yilmaz, A., Özerbil, T., Hakan, H., Firat, O., Atila, O., Harita, A., Komutanlığı, G., Dairesi, F., & Dikimevi, A. Sayısal Hava Kamerası ve Fotogrametriye Kazandırdıkları, *Harita Dergisi*, 143.
- Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y., Dollár, P., & Girshick, R. (2023). Segment Anything.
- Kılıçarslan, S., Adem, K., & Çelik, M. (2021). An Overview of the Activation Functions Used in Deep Learning Algorithms. *Journal of New Results in Science*, 10(3), 75–88.
- Lin, J., Jing, W., Song, H., & Chen, G. (2019). ESFNet: Efficient Network for Building Extraction From High-Resolution Aerial Images. *IEEE Access*, 7, 54285-54294.
- Liu, Z., Wang, J., Liu, W., 2005. Building Extraction from High Resolution Imagery Based on Multiscale Object Oriented Classification and Probabilistic Hough Transform. *2005 IEEE International Geoscience and Remote Sensing Symposium*, 2250–2253.
- Long, J., Shelhamer, E., Darrell, T., 2015. Fully Convolutional Networks for Semantic Segmentation. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 3431-3440.
- Maggiori, E., Tarabalka, Y., Charpiat, G., & Alliez, P. (2017). Can Semantic Labeling Methods Generalize to Any City? The Inria Aerial Image Labeling Benchmark. *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 3226-3229.
- Mao, A., Mohri, M., & Zhong, Y. (2023). Cross-Entropy Loss Functions: Theoretical Analysis and Applications.
- Makineci, H. B., Karasaka, L. & Şahin, D. (2020). Fotogrametrik Amaçlı Kamera Kalibrasyonu Yazılımlarının Kullanımı Üzerine Bir Araştırma. *Türkiye Fotogrametri Dergisi*, 2 (1), 14-21.
- Marr, D., & Hildreth, E. (1980). Theory of Edge Detection. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 207(1167), 187-217.
- Minai, A. A., & Williams, R. D. (1993). On the derivatives of the sigmoid. *Neural Networks*, 6(6), 845-853.
- Mnih, (2013). *Machine Learning For Aerial Image Labeling*. Ph.D. Dissertation, University of Toronto.
- Mostafa, S., & Wu, F.-X. (2021). Diagnosis of Autism Spectrum Disorder with Convolutional Autoencoder and Structural MRI Images. *In Neural Engineering Techniques for Autism Spectrum Disorder*, 23–38.
- Najman, L., & Schmitt, M. (1994). Watershed of a Continuous Function. *Signal Processing*, 38(1), 99-112.

- Noreen, N., Palaniappan, S., Qayyum, A., Ahmad, I., Imran, M., & Shoaib, M. (2020). A Deep Learning Model Based on Concatenation Approach for the Diagnosis of Brain Tumor. *IEEE Access*, 8, 55135-55144.
- Nwankpa, C. E., Ijomah, W., Gachagan, A., & Marshall, S. (2021). Activation Functions: Comparison of Trends in Practice and Research for Deep Learning. *II. International Conference on Computational Sciences and Technology*, Jamshoro, Pakistan. 124 - 133.
- O'Shea, K. & Nash, R. (2015). An Introduction to Convolutional Neural Net-Works.
- Özbalmumcu, M. (1996). Sayısal (Digital) Fotogrametri ve Sayısal Fotogrametrik Stereo Çalışma İstasyonları (Softcopy Plotter). *Harita Dergisi*, 116, 62-74.
- Özşen, Ö. (2002). *Matematiksel Morfoloji Kullanılarak Göğüs Kanserinin Erken Teşhisi*. A.Ü., Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Eskişehir.
- Öztemel, E. (2012). *Yapay Sinir Ağları*.
- Öztürk, K. & Şahin, M. E. (2018). Yapay Sinir Ağları ve Yapay Zekâ'ya Genel Bir Bakış. *Takvim-i Vekayi*, 6 (2), 25-36.
- Peemen, M., Shi, R., Lal, S., Juurlink, B., Mesman, B., & Corporaal, H. (2016). The Neuro Vector Engine: Flexibility to Improve Convolutional Net Efficiency for Wearable Vision. *Proceedings of the 2016 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 1604–1609.
- Prewitt, J. (1970). Object Enhancement and Extraction. *Elsevier Science*, 75-149.
- Qiu, Y., Wu, F., Yin, J., Liu, C., Gong, X., & Wang, A. (2022). MSL-Net: An Efficient Network for Building Extraction from Aerial Imagery. *Remote Sensing*, 14(16), 3914.
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. 234–241.
- Ruder, S. (2016). An overview of gradient descent optimization algorithms.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning Representations by Back-Propagating Errors. *Nature*, 323(6088), 533–536.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4510–4520.
- Samui, P., Roy, S. S., & Balas, V. E. (2017). *Handbook of Neural Computation*.
- Scardapane, S., Van Vaerenbergh, S., Totaro, S., & Uncini, A. (2017). Kafnets: kernel-based non-parametric activation functions for neural networks.

- Schmidt, R. M. (2019). Recurrent Neural Network (RNNs): A gentle Introduction and Overview.
- Ser, G. & Bati, C. T. (2019). Derin Sinir Ağları ile En İyi Modelin Belirlenmesi: Mantar Verileri Üzerine Keras Uygulaması . *Yuzuncu Yıl University Journal of Agricultural Sciences*, 29 (3), 406-417.
- Sharma, S., Sharma, S., & Athaiya, A. (2020). Activation Functions in Neural Networks. *International Journal of Engineering Applied Sciences and Technology*, 04(12), 310-316.
- Siegel, J. A., & Saukko P. J., (2012). *Encyclopedia of Forensic Sciences*.
- Sifre, L. & Mallat, S. (2014). Rigid-Motion Scattering for Image Classification. Ph.D. Thesis, École Polytechnique, Paris, France.
- Singh, S. A., Meitei, T. G., & Majumder, S. (2020). Short PCG Classification Based on Deep Learning. *In Deep Learning Techniques for Biomedical and Health Informatics*, 141–164.
- Sobel, I. (1970). Camera Models and Perception. Ph.D. Thesis, Stanford University, Stanford.
- Tonbul, H. & Kavzoğlu, T. (2017). Nesne-Tabanlı Sınıflandırmada Segmentasyon (Bölütleme) Kalitesinin Sınıflandırma Doğruluğu Üzerine Etkisinin İncelenmesi. *Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, 17, 118-125.
- URL-1. (2023, Eylül 20). Dijital Hava Kameraları: <https://terim.rehberim.gen.tr/terim/hava-kamerasi-nedir> adresinden alınmıştır.
- URL-2. (2023, Eylül 20). Dijital Hava Kameraları ve Özellikleri: <https://geo-matching.com/digital-aerial-cameras> adresinden alınmıştır.
- URL-3. (2023, Eylül 21). Dijital Hava Kameraları ve Özellikleri: <https://ataymuhendislik.com/sayisal-hava-kameralari/> adresinden alınmıştır.
- URL-4. (2023, Eylül 22). Aşınma ve Genleşme İşlemi: www.yalova.edu.tr/Files/UserFiles/70/aliiskurt/IMAGE_PROCESS/morfolojik_30Ekim.pdf adresinden alınmıştır.
- URL-5. (2023, Eylül 24). Obje Tespiti ve Segmentasyon Farkı: <https://medium.com/move-on-ai/g%C3%B6r%C3%BCnt%C3%BCsegmentasyonu-ba193b168a1b> adresinden alınmıştır.
- URL-6. (2023, Eylül 25). Segmentasyon Türleri: <https://www.dataguess.com/post/semantic-segmentation-ve-instance-segmentation-> adresinden alınmıştır.
- URL-7. (2023, Eylül 26). Yapay Zeka: <https://proente.com/makine-ogrenimi-derin-ogrenme-ve-yapay-zeka-arasindaki-fark/> adresinden alınmıştır.
- URL-8. (2023, Eylül 26). Yapay Zeka: <https://soyleki.com/yapay-zeka-nedir-nasil-calisir-insan-gibi-dusunen-makine/> adresinden alınmıştır.

- URL-9. (2023, Eylül 27). Makine Öğrenimi ve Derin Öğrenme Farkı: <https://www.turhost.com/blog/deep-learning-nedir/#serp> adresinden alınmıştır.
- URL-10. (2023, Eylül 29). Basit Yapay Sinir Ağı ve Derin Yapay Sinir Ağı: <https://towardsdatascience.com/a-conversation-about-deep-learning-9a915983107> adresinden alınmıştır.
- URL-11. (2023, Eylül 29). U-Net Mimarisi: <https://www.binamod.com/post/u-net-ile-g%C3%B6r%C3%BCnt%C3%BCsegmentasyonu> adresinden alınmıştır.
- URL-12. (2023, Ekim 1). U-Net Mimari Özellikleri: <https://ayyucekizrak.medium.com/g%C3%B6r%C3%BCnt%C3%BCb%C3%B6l%C3%BCtleme-segmentasyon-i%C3%A7in-derin-%C3%B6l%C4%9Frenme-u-net-3340be23096b> adresinden alınmıştır.
- URL-13. (2023, Ekim 1). Pürüzsüz Karıştırma Yöntemi: <https://github.com/Vooban/Smoothly-Blend-Image-Patches> adresinden alınmıştır.
- URL-14. (2023, Ekim 2). Toplu Normalleştirme: <https://www.baeldung.com/cs/batch-normalization-cnn> adresinden alınmıştır.
- URL-15. (2023, Ekim 3). Aktivasyon Fonksiyonları: <https://datakapital.com/blog/yapay-sinir-aglari-nedir-ve-nasil-calisir/> adresinden alınmıştır.
- Wang, S., Hou, X., & Zhao, X. (2020). Automatic Building Extraction From High-Resolution Aerial Imagery via Fully Convolutional Encoder-Decoder Network With Non-Local Block. *IEEE Access*, 8, 7313–7322.
- Wang, Y., Li, S., Teng, F., Lin, Y., Wang, M., & Cai, H. (2022). Improved Mask R-CNN for Rural Building Roof Type Recognition from UAV High-Resolution Images: A Case Study in Hunan Province. *Remote Sensing*, 14(2), 265.
- Wang, Y., Zeng, X., Liao, X., & Zhuang, D. (2022). B-FGC-Net: A Building Extraction Network from High Resolution Remote Sensing Imagery. *Remote Sensing*, 14(2), 269.
- Woo, S., Park, J., Lee, J.-Y., & Kweon, I. S. (2018). CBAM: Convolutional Block Attention Module. 3-19.
- Waxman, O. B. (2018, Mayıs 31). *Aerial Photography's Surprising Role in History*. <https://time.com/longform/aerial-photography-drones-history/> adresinden alınmıştır.
- Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., Jiang, Y., Xu, H., Tan, W., Yang, Q., Wang, J., Gao, J., & Zhang, L. (2020). Deep Learning in Environmental Remote Sensing: Achievements and Challenges. *Remote Sensing of Environment*, 241, 111716.
- Zadeh, L. A. (1965). Fuzzy Sets. *Information and Control*, 8, 338-353.
- Zha, Y., Gao, J., Ni, S. (2003). Use of Normalized Nifference Built up Index in Automatically.

- Zhao, H., Zhang, H., & Zheng, X. (2022). A Multiscale Attention-Guided UNet++ with Edge Constraint for Building Extraction from High Spatial Resolution Imagery. *Applied Sciences*, 12(12), 5960.
- Zhu, X., Hu, H., Lin, S., & Dai, J. (2019). Deformable ConvNets V2: More Deformable, Better Results. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 9300–9308.
- Zhu, Y., Liang, Z., Yan, J., Chen, G., & Wang, X. (2021). E-D-Net: Automatic Building Extraction From High-Resolution Aerial Images With Boundary Information. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 4595–4606.
- Zhu, Z., Xu, M., Bai, S., Huang, T., & Bai, X. (2019). Asymmetric Non-local Neural Networks for Semantic Segmentation.



ÖZGEÇMİŞ

Feride Seçil YILDIRIM, İlköğrenimini Yalova Saffet Çam İlköğretim Okulu'nda bitirdikten sonra lise eğitimini 2017 yılında Yalova Anadolu Lisesi'nde tamamladı. 2017-2018 öğretim yılında Karadeniz Teknik Üniversitesi Mühendislik Fakültesi Harita Mühendisliği bölümünde lisans eğitimine başladı. Lisans tezi kapsamında "MODIS Uydu Verilerinden Yağışa Dönüştürülebilir Su Buharı Hesabı" konusu üzerinde çalıştı. 2242-Üniversite Öğrencileri Araştırma Proje Yarışması'nda "Hiperspektral Uydu Görüntüleri Yardımıyla Yapay Resif Uygulaması" isimli projeyle final aşamasına ilerledi. 2021 yılında lisans eğitimini, bölüm birincisi (Yüksek Onur Öğrencisi) olarak tamamladı. 2022 yılında Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Harita Mühendisliği Anabilim dalında tezli yüksek lisans programına başladı. Tez danışmanı Prof. Dr. Fevzi KARSLI ve Harita Yüksek Mühendisi Murat BAHADIR ile yapmış oldukları "Görüntü Eşleştirme ve LiDAR Kaynaklı Nokta Bulutundan Orman Metriklerinin Otomatik Belirlenmesi" isimli çalışma ile VIII. UZAL-CBS 2022 (Ankara) Sempozyumunda en iyi sözlü bildiri ikincilik ödülünü kazandılar.

Bahadır, M., Yıldırım, F. S., & Karşlı, F. (2022). Görüntü Eşleştirme ve LiDAR Kaynaklı Nokta Bulutundan Orman Metriklerinin Otomatik Belirlenmesi. *VIII. Uzaktan Algılama-CBS Sempozyumu*, 17-19 Kasım 2022, Ankara.

Özdemir, E., Bahadır, M., Yıldırım, F. S., Karşlı, F. (2023). Düzgünleştirilmiş Bina Sınır Doğruluğunun Piksel ve Nokta Bazlı Doğruluk Analizi Yöntemleriyle İrdelenmesi. *XII. Türkiye Ulusal Fotogrametri ve Uzaktan Algılama Birliği (TUFUAB) Teknik Sempozyumu*, 24-26 Mayıs 2023, Sivas.

Yıldırım, F. S. & Tunç Görmüş, E. (2022). Landsat ETM+ Uydu Görüntüleri ile Metal Kaynaklı Bitki Stresinin Araştırılması. *Turkish Journal of Remote Sensing and GIS*, 3 (2), 183-190.