

**BAŞKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİMDALİ
BİLGİSAYAR MÜHENDİSLİĞİ TEZLİ YÜKSEK LİSANS
PROGRAMI**

**UYDU GÖRÜNTÜLERİNDEN EVRİŞİMSEL SİNİR AĞLARI İLE
BİNA BÖLGE TESPİTİ VE ÇATI TİPİ SINIFLANDIRMASI**

HAZIRLAYAN

ERDAL ÖZTÜRK

YÜKSEK LİSANS TEZİ

ANKARA – 2022

**BAŞKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİMDALİ
BİLGİSAYAR MÜHENDİSLİĞİ TEZLİ YÜKSEK LİSANS
PROGRAMI**

**UYDU GÖRÜNTÜLERİNDEN EVRİŞİMSEL SİNİR AĞLARI İLE
BİNA BÖLGE TESPİTİ VE ÇATI TİPİ SINIFLANDIRMASI**

HAZIRLAYAN

ERDAL ÖZTÜRK

YÜKSEK LİSANS TEZİ

TEZ DANIŞMANI

DR. ÖĞR. ÜYESİ EMRE SÜMER

ANKARA – 2022

BAŞKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Bilgisayar Mühendisliği Anabilim Dalı Bilgisayar Mühendisliği Tezli Yüksek Lisans Programı çerçevesinde Erdal Öztürk tarafından hazırlanan bu çalışma, aşağıdaki jüri tarafından Yüksek Lisans Tezi olarak kabul edilmiştir.

Tez Savunma Tarihi: 11 / 08 / 2022

Tez Adı: Uydu Görüntülerinden Evrişimsel Sinir Ağları ile Bina Bölge Tespiti ve Çatı Tipi Sınıflandırması

Tez Jüri Üyeleri (Unvanı, Adı - Soyadı, Kurumu)

İmza

Dr. Öğr. Üyesi Tülin Erçelebi Ayyıldız, Başkent Üniversitesi

.....

Dr. Öğr. Üyesi Emre Sümer, Başkent Üniversitesi

.....

Dr. Öğr. Üyesi Berk Anbaroğlu, Hacettepe Üniversitesi

.....

ONAY

Prof. Dr. Faruk Elaldı
Fen Bilimleri Enstitüsü Müdürü
Tarih: ... / ... /

BAŞKENT ÜNİVERSİTESİ
FEN BİLİMLER ENSTİTÜSÜ
YÜKSEK LİSANS ÇALIŞMASI ORJİNALLİK RAPORU

Tarih: 24 / 08 / 2022

Öğrencinin Adı, Soyadı : Erdal Öztürk

Öğrencinin Numarası : 21810095

Anabilim Dalı : Bilgisayar Mühendisliği

Programı : Bilgisayar Mühendisliği Tezli Yüksek Lisans

Danışmanın Unvanı/Adı, Soyadı : Dr. Öğr. Üyesi Emre Sümer

Tez Başlığı : Uydu Görüntülerinden Evrişimsel Sinir Ağları ile Bina Bölge Tespiti ve Çatı Tipi Sınıflandırması

Yukarıda başlığı belirtilen Yüksek Lisans tez çalışmamın; Giriş, Ana Bölümler ve Sonuç Bölümünden oluşan, toplam 70 sayfalık kısmına ilişkin, 24 / 08 / 2022 tarihinde tez danışmanım tarafından Turnitin adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı % 15 'tir. Uygulanan filtrelemeler:

1. Kaynakça hariç
2. Alıntılar hariç
3. Beş (5) kelimeden daha az örtüşme içeren metin kısımları hariç

“Başkent Üniversitesi Enstitüleri Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Usul ve Esaslarını” inceledim ve bu uygulama esaslarında belirtilen azami benzerlik oranlarına tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Öğrenci İmzası:

ONAY

24 / 08 / 2022

Öğrenci Danışmanı

Dr. Öğr. Üyesi Emre Sümer

TEŞEKKÜR

Çalışma sürecini ve benden beklenenleri açıklayan, karşılaşılan güçlüklerin aşılmasında her zaman yardımcı olan ve çalışmanın sonuca ulaştırılmasında değerli katkılarından dolayı tez danışmanım Sayın Dr. Öğr. Üyesi Emre SÜMER'e sonsuz teşekkürlerimi ve saygılarımı sunarım.

Bu süreçte her zaman yanımda olan ve destekleyen aileme, Derin öğrenme konusunda teknik desteklerinden dolayı Türk Havacılık ve Uzay Sanayi şirketine İHA Grubu Yazılım Müdürü Sn. Semih DADANLAR'a, her zorlukta motivasyonumu sağlayan sevgili eşim Derya ÖZTÜRK'e çok teşekkür ederim.

ÖZET

Erdal ÖZTÜRK

UYDU GÖRÜNTÜLERİNDEN EVRİŞİMSEL SİNİR AĞLARI İLE BİNA BÖLGE TESPİTİ VE ÇATI TİPİ SINIFLANDIRMASI

Başkent Üniversitesi Fen Bilimleri Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

2022

Coğrafi Bilgi Sistemleri (CBS); verimlilik, planlama, hasar tespiti ve daha fazlası dahil olmak üzere bir bina modelinin birçok yönünü analiz eder. CBS uygulamalarının ihtiyaç duyduğu temel bileşenlerden biri, bir binanın çatı tipi ve geometrisidir. Bu nedenle, çatıların tanımlanması ve sınıflandırılması CBS'de araştırılan önemli bir konudur. Gelişen teknoloji sayesinde geniş alanların yüksek çözünürlüklü uydu görüntülerini elde etmek kolay ve ucuz bir hale gelmiştir. Uydu görüntüleri üzerinde yapılabilecek sınıflandırma yöntemi ile geniş alanlar hızlı bir şekilde analiz edilebilmektedir. Bu sayede zaman, maliyet ve işçilik konusunda önemli derecede kazanç sağlanabilmektedir. Bazı nesne algılama yöntemleri, günümüzde uydu görüntülerinde bina çatılarının konumlarını belirleyebilmektedir. Derin öğrenme, nesne tespiti için kullanılan bu yöntemlerden biridir. Klasik makine öğrenme metodu bu sorunu çözmede başarılı sonuçlar elde etmiş olsa da, derin öğrenme üzerine yapılan son araştırmalar doğruluk oranlarındaki başarının daha yüksek seviyelere çıkabileceğini göstermiştir. Bu çalışmada, derin öğrenme yöntemleri kullanılarak bina bölgelerine ait çatı konumlarının alansal olarak belirlenmesinin yanı sıra ilgili çatı tipleri de sınıflandırılmıştır. Bunu başarmak için, CNN yönteminin performansı, orta seviye çözünürlüklü görüntüler kullanılarak test edilmiştir. Farklı bölgelere ait elde edilen uydu görüntülerinin test sonuçları incelendiğinde, AlexNet'in %92,30, GoogleNet'in ise %87,18 doğrulukla çatı bölgelerini tespit ettiği görülmektedir. Ayrıca, incelenen üç mimariden bina çatı tiplerinin sınıflandırılması için en uygun modelin AlexNet olduğu sonucuna varılmıştır.

ANAHTAR KELİMELEER: Bina atı Tipi, Uydu Görüntüleri, Nesne Sınıflandırma, Evrişimsel Sinir Ağı (ESA)



ABSTRACT

Erdal ÖZTÜRK

BUILDING ZONE DETECTION AND ROOF TYPE CLASSIFICATION FROM SATELLITE IMAGES USING CONVOLUTIONAL NEURAL NETWORKS

Baskent University Institute of Science and Engineering

Department of Computer Engineering

2022

Geographic Information Systems (GIS) analyzes many aspects of a building model, including efficiency, planning, damage assessment and more. One of the key components required by GIS applications is the roof type and geometry of a building. Therefore, the identification and classification of roofs is an important topic explored in GIS. With developing technology, obtaining high-resolution satellite images of large areas can be easy and inexpensive. Large areas can be analyzed quickly with the classification method that can be done on satellite images. In this way, significant savings in time, cost and labor can be achieved. Some object detection methods can now determine the positions of building roofs in satellite images. Deep learning is one of these methods used for object detection. Although the classical machine learning method has achieved successful results in solving this problem, recent research on deep learning has shown that the success in accuracy rates can reach higher levels. In this study, besides the regional determination of the roof locations of the building regions by using deep learning methods, the related roof types are also classified. To achieve this, the performance of the CNN method is tested using medium resolution images. When the test results of the satellite images obtained from different regions are examined, it is seen that AlexNet produced successful results in region detection with 92.30% accuracy and GoogleNet produced 87.18% accuracy. In addition, it was concluded that AlexNet was the most suitable model for the classification of building roof types among the three architectures examined.

KEYWORDS: Building Roof Type, Satellite Images, Object Classification, Convolutional Neural Network (CNN)

İÇİNDEKİLER

TEŞEKKÜR.....	i
ÖZET	ii
ABSTRACT	iv
İÇİNDEKİLER.....	v
TABLolar LİSTESİ	viii
ŞEKİLLER LİSTESİ	1
1. GİRİŞ.....	1
2. TEMEL KAVRAMLAR.....	6
2.1. Uzaktan Algılama	6
2.1.1. Uzaktan Algılama nedir?	6
2.1.1.1. Platform.....	7
2.1.1.1.1. Yer Platformu	7
2.1.1.1.2. Hava Platformu	8
2.1.1.1.3. Uydu Platformu.....	8
2.1.1.2. Kullanılan Enerji.....	10
2.1.1.3. Algılayıcı Sistemler.....	10
2.1.1.3.1. Pankromatik Görüntüleme	10
2.1.1.3.2. Multi-Spektral Görüntüleme	11
2.1.1.3.3. Hiper-Spektral Görüntüleme.....	11
2.1.2. Uydu Sistemleri ile Uzaktan Algılama.....	12
2.1.3. Optik Algılayıcılar	14
2.1.3.1. Konumsal Çözünürlük.....	15
2.1.3.1.1. Düşük Çözünürlüklü Uydu Görüntüsü	15
2.1.3.1.2. Orta Çözünürlüklü Uydu Görüntüleri	16

2.1.3.1.3. Yüksek Çözünürlüklü Uydu Görüntüleri.....	17
2.1.3.2. Radyometrik Çözünürlük.....	18
2.1.3.3. Zamansal Çözünürlük	19
2.1.3.4. Spektral Çözünürlük	20
2.1.4. Radar	21
2.2. Derin Öğrenme.....	22
2.2.1. Derin Öğrenme (Deep Learning) Nedir?.....	22
2.2.2. Derin Öğrenmenin Tarihi.....	24
2.2.3. Derin Öğrenmenin Kullanım Alanları	28
2.2.3.1. Görüntü İşleme.....	28
2.2.3.2. Sınıflandırma	29
2.2.3.3. Nesne Tespiti.....	30
2.2.3.4. Ses Verilerinin Sınıflandırılması.....	30
2.2.3.5. Medikal Alanda Teşhis ve Tanı	31
2.2.3.6. Endüstriyel Alanda Tespit ve Sınıflandırma	31
2.3. Evrimsel Sinir Ağları (CNN).....	31
2.3.1. Convolutional Layer.....	33
2.3.1.1. Stride	34
2.3.1.2. Padding.....	35
2.3.2. ReLu.....	35
2.3.3. Pooling	36
2.3.4. Flattening.....	37
2.3.5. Fully Connected Layer	37
2.4. R-CNN, Fast R-CNN ve Faster R-CNN.....	38
2.4.1. R-CNN	38
2.4.2. Fast R-CNN	40
2.4.3. Faster R-CNN	40

2.4.3.1.	Region Proposal Network (RPN)	40
3.	MATERYAL VE METOT	42
3.1.	Model Eğitimi	42
3.1.1.	AlexNet	42
3.1.2.	GoogleNet	43
3.1.3.	DarkNet53	46
3.2.	Arayüz Tasarımı	53
3.2.1.	Test.....	58
4.	DENEYSEL SONUÇLAR	61
5.	SONUÇ VE DEĞERLENDİRME	66
	KAYNAKLAR.....	71

TABLÖLER LİSTESİ

Tablo 1. Uydu Sistemleri ve Özellikleri.....	13
Tablo 2. Çeşitli Algılayıcılar için Farklı Çözünürlüklerdeki Değerler.....	21
Tablo 3. AlexNet Hata Matrisi	62
Tablo 4. GoogleNet Hata Matrisi	63
Tablo 5. DarkNet53 Hata Matrisi	64
Tablo 6. Mikro F1 Skor Tablosu	65
Tablo 7. Top-1 ve Top-2 Değerlerine Göre Sınıflandırma Değerleri.....	66



ŞEKİLLER LİSTESİ

Şekil 2.1. Uzaktan Algılama Sistemi.....	6
Şekil 2.2. Örnek Yer Platformları.....	7
Şekil 2.3. Örnek Bir Hava Platformu Görünümü	8
Şekil 2.4. Örnek Bir Uydu Platformu Görünümü.....	9
Şekil 2.5. Platformlar ve Elde Edilen Örnek Görüntü Verileri	10
Şekil 2.6. Multi-Spektral Görüntü	11
Şekil 2.7. Hiper–spektral Görüntü.....	12
Şekil 2.8 Farklı Konumsal Çözünürlükte Optik Uydu Görüntüleri	15
Şekil 2.9. Düşük Çözünürlüklü Sentinel-2 Görüntüsü	16
Şekil 2.10. Orta Çözünürlüklü Sentinel-2 Görüntüsü	17
Şekil 2.11. Yüksek Çözünürlüklü PlanetScope Görüntüsü	18
Şekil 2.12. Yüksek Çözünürlüklü SkySat Görüntüsü	18
Şekil 2.13. Radyometrik çözünürlüğe ait görüntüler.....	19
Şekil 2.14. Spektral Çözünürlüğe Ait Bant ve Kanal Değerleri.....	20
Şekil 2.15. Derin Öğrenme ve Yapay Zekâ İlişkisi.....	23
Şekil 2.16. Derin Öğrenme Sistemi	23
Şekil 2.17. Yapay Sinir Ağı Modeli	24
Şekil 2.18. Perceptron Model	25
Şekil 2.19. Çok Katmanlı Ağ Modeli	26
Şekil 2.20. Destek Vektör Makineleri	27
Şekil 2.21. Yapay Sinir Ağı.....	32
Şekil 2.22. 3 x 3 Convolutional Layer Filtresi	33
Şekil 2.23. Konvolüsyonel Filtrenin Çalışma Prensibi	34
Şekil 2.24. Konvolüsyonel Çıktı	34
Şekil 2.26. ReLu Değer Grafiği.....	36
Şekil 2.27. Pooling İşlemi	36
Şekil 2.28. Flattening İşlemi.....	37
Şekil 2.29. Fully Connected İşlemi	38
Şekil 2.30. R-CNN Bölgesel İşaretleme	39
Şekil 2.31. R-CNN İşlem Adımları	39

Şekil 2.32. Faster R-CNN yapısı	41
Şekil 3.1 AlexNet Katman Yapısı	43
Şekil 3.2. GoogleNet Katman Yapısı	45
Şekil 3.3. DarkNet53 Katman Yapısı	47
Şekil 3.4. Eğitim Parametre Değerleri.....	48
Şekil 3.5. Düz Çatı Tipine Ait Veri Seti Örneği.....	49
Şekil 3.6. Üçgen Çatı Tipine Ait Veri Seti Örneği.....	50
Şekil 3.7. Sarkık Çatı Tipine Ait Veri Seti Örneği.....	50
Şekil 3.8. Sınıflara Ait Eğitim Veri Seti Dağılımları	51
Şekil 3.9. AlexNet Eğitim Sonuçları	52
Şekil 3.10. GoogleNet Eğitim Sonuçları	52
Şekil 3.11. DarkNet53 Eğitim Sonuçları.....	53
Şekil 3.12. İşlem Akış Diyagramı	54
Şekil 3.13. Tarama Hareket Yönü	55
Şekil 3.14. Görüntü Tarama Algoritması	56
Şekil 3.15. Bina Bölge Tespiti ve Çatı Sınıflandırılmasına Ait İşlemler	56
Şekil 3.16. Test Görüntüleri Üzerinde Bina Bölgelerinin Tespiti ve Sadeleştirilmesi.....	57
Şekil 3.17. DarkNet53 Sınıflandırma Örneği	58
Şekil 3.18. Çatı Renk Çeşitliliği İçeren Test Görüntü Örneği.....	59
Şekil 3.19. Test Görüntü Örneği.....	60
Şekil 4.1. DarkNet53 Test Sonucu	67
Şekil 4.2. GoogleNet Test Sonucu	67
Şekil 4.3. AlexNet Test Sonucu.....	68

1. GİRİŞ

Günümüzde çeşitli uçak veya uydulardan görüntü almanın maliyeti düşerken, elde etme kolaylığı artmıştır. Uydu görüntüleri, şehirlerin, binaların ve nesnelerin tespiti ve analizi için vazgeçilmez bir veri kaynağı haline gelmiştir. Uydu görüntüleri, şehirlerde yıllar içinde meydana gelen değişiklikleri, yapısal analizleri ve doğal afetlerden kaynaklanan hasarları belirlemek için yaygın olarak kullanılmaktadır. Derin öğrenme yöntemi, derin evrimsel sinir ağlarının kullanımına dayanan bir makine öğrenmesi yöntemidir ve bu yöntemin kullanıldığı birçok çalışmada yüksek hassasiyetli nesne algılama ve sınıflandırma sonuçları elde edilmiştir. Bina tespiti ve sınıflandırmada evrimsel sinir ağlarının kullanılması bu başarı oranını daha da artırmaktadır. Yapı enerji verimliliği, kentsel yapılaşma, şehir planlama ve doğal afet hasar tespiti gibi çeşitli analizlerin yapılabilmesi konusunda binalara ait çatı tipi, renk ve geometri gibi verilere ihtiyaç vardır. Bu sebeple binalara ait çatıların tespiti ve sınıflandırılması, CBS alanında kritik öneme sahip problemlerden biri haline gelmiştir. Evrimsel sinir ağlarında en önemli veri kaynağı görüntülerdir. Günümüz teknolojisinde uydu görüntülerine çok kolay erişim sağlanabilmektedir. Evrimsel sinir ağları kullanılarak geniş alanlara ait binaların çatı sınıflandırılmasının kısa süre içerisinde ve düşük maliyetle yapılabilmesi mümkün olmaktadır. Özellikle doğal afetlerde durum tespiti, kentsel yerleşimde planlama ve istatistiksel çalışmalar için verilerin oluşturulabilmesi konusunda büyük fayda sağlamaktadır. Son zamanlarda yapılan benzer çalışmalarda elde edilen yüksek doğruluk oranları, bu yöntemi daha cazip kılmaktadır.

Ayrıca, doğal afetlerden sonra meydana gelen hasarların tespitinin yapılması ve hızlı müdahale edilebilmesi amacıyla kısa sürede yüksek doğrulukta tespit gereklidir. Geniş alanların analiz edilebilmesinin zaman alıcı olması ve yüksek maliyeti, tespit ve sınıflandırma için evrimsel sinir ağlarının kullanılmasının diğer önemli bir nedenidir. Günümüzde evrimsel sinir ağlarına dayalı çeşitli yöntemler önerilmiştir. Görüntü bölütleme, maskeleme ve YOLO (You Only Look Once) [1] bu yöntemlerden bazılarıdır. Yakın zamanda yapılan bir çalışmada amaç, R-CNN mimarisi ile bir maskeleme yöntemi kullanarak bina algılama başarı oranını artırmaktır [2]. Yatay ve dikey kaydırma, kontrast, gürültü değerleri ve parlaklık ayarlanarak bina algılama üzerindeki etkileri gözlemlenmiştir. Başka bir çalışmada ise depremde hasar gören binalar tespit edilmiştir [3]. Bu çalışmada

DarkNet derin öğrenme mimarisine dayalı olarak YOLO-V3 yönteminin tespit doğruluğunu iyileştirmeye yönelik bir çalışma yapılmıştır. Çalışma, YOLO-V3 kullanılarak gözlemlenen %5.24'lük bir performans artışı ile yaklaşık %91'lik bir başarı oranı elde etmiştir. Özyurt (2019) yapmış olduğu çalışmada WHU-RS19 veri setinden elde edilen uzaktan algılama görüntülerinin sınıflandırılması için farklı derin öğrenme modellerinden alınan özneliklerin komşuluk bileşen analizi (KBA) ile indirgenip Destek Vektör Makinesi (DVM) ile sınıflandırmasını yapmıştır. CNN modellerinden AlexNet, VGG-16 ve GoogleNet'i kullanarak her mimarinin son tam bağlı katmanından 1000'er adet öznelik elde etmiştir. Ayrıca üç mimariden elde edilen öznelikler birleştirilerek komşuluk bileşen analizi yöntemiyle 1000 özneliğe indirgenmiştir. Havaalanı, plaj, köprü, ticari alan, çöl, tarım arazileri, futbol sahası, orman, sanayi bölgesi, çayır, dağ, park, park yeri, gölet, liman, tren istasyonu, konut, alan, nehir ve viyadük olmak üzere toplamda 20 sınıf oluşturulmuştur. Aynı veriyi kullanan diğer çalışmalar ile kıyaslama yapılabilmesi için veri setinin %60 ve %40'ı kullanılarak eğitimi DVM ile gerçekleştirilmiştir. Her sınıftan yaklaşık 50 görüntü olmak üzere toplamda 1005 görüntü veri seti olarak alınmıştır. Çalışma sonucunda verilerin %60'ı eğitim olarak kullanıldığında %98.75, %40'ı eğitim olarak kullanıldığında ise %97.01 oranında başarı elde edilmiştir [4]. CNN kullanılarak yapılan diğer bir çalışmada fotoğraf ve video görüntülerine ait veri seti kullanılarak görüntü içerisindeki bardak görüntüsünün tespiti yapılmıştır [5]. Bu çalışmayı gerçekleştirmek için Faster R-CNN adında bir derin öğrenme tekniği ve Object Detection API kullanılmıştır. Bardak ve kupa resimlerinden oluşan 502 adet görüntü veri setini etiketlemek için "LabelImg" kullanılmıştır. Bu çalışmada, 18 farklı deneyde 502 görüntüden oluşturulan bir veri seti üzerinde derin öğrenme teknikleri kullanılarak Faster R-CNN kullanılarak nesne tespiti gerçekleştirilmiştir. Toplam 100.000 adımlık eğitimde %97.83 başarı ve %2.16 kayıp oranı elde edilmiştir. Sınıflandırma üzerine diğer bir çalışmada CNN modelleri kullanılarak çiçeklerin sınıflandırılması yapılmıştır [6]. CNN modellerinden AlexNet, VGG-16 ve VGG-19 kullanılarak, 4326 adet görüntüden oluşan veri seti ile çalışma gerçekleştirilmiştir. Çiçek görüntülerinden elde edilen özellikler destek vektör makineleri (DVM) ile sınıflandırılmış ve elde edilen sonuçlar karşılaştırılmıştır. Elde edilen bulgulara göre en iyi sonucu %86,56 doğruluk oranıyla VGG-16 sağlamıştır. AlexNet ise %86,29 ile çok yakın bir değer üreterek başarılı olduğunu göstermiştir. Sonraki aşamada CNN mimarilerinin son tam bağlantılı katmanından elde edilen 1000 özellik birleştirilerek 3000 özellik seti oluşturulmuştur. Ardından, özellik seçim yöntemlerinden; Maksimum Bilgi Katsayısı (MBK), Ridge regresyonu ve Özyinelemeli

Özellik Eleme (ÖÖE) yöntemleri kullanılarak özellik sayısı 300'e düşürülmüştür. DVM yöntemi ile tekrar sınıflandırma yapılmıştır. Sınıflandırma sonucu başarı oranı yaklaşık %4,54 artarak %91,10 olmuştur.

Literatürde bina çatılarının sınıflandırılmasına yönelik önemli çalışmalar yer almaktadır. CNN yöntemiyle yapılan bir çalışmada binaların çatı tipleri sınıflandırılmıştır [7]. Bu çalışma için Düz (Flat), Üçgen (Gable) ve Sarkık (Hip) tiplerinde üç farklı çatı sınıfı tanımlanmıştır. Görüntü içerisinde tek bir bina görüntüsü olacak şekilde veri seti oluşturulmuştur. Bu çalışma temel olarak iki aşamadan oluşmaktadır. İlk olarak, bir CNN modeli; ağaç, yol ve bina gibi kentsel nesneleri çıkarmak için eğitilmiştir. Daha sonra, bir diğer CNN modeli kullanılarak farklı çatı tiplerine göre sınıflandırılması gerçekleştirilmiştir. Değerlendirme sonuçları, tespit ve tanıma adımlarında sırasıyla yaklaşık %97 ve %92 başarı oranları ile önerilen yöntemin etkinliğini kanıtlamıştır. Diğer bir çalışma ise yüksek çözünürlüklü uydu görüntülerinden bina çatılarının tespit edilmesidir [8]. Bu çalışmada, eğitim veri kümesinin oluşturulması, model eğitimi, görüntü segmentasyonu ve bina çatıları sınıflandırılması gibi birkaç ana adım söz konusudur. İlk olarak; ağaç, yol, meydan ve bina gibi kentsel nesneler, nesne yönelimli sınıflandırma yaklaşımı ile rastgele orman algoritmasına göre sınıflandırılmış, yapı bölgeleri görsel olarak yorumlama ve düzeltme yardımı ile diğer sınıflardan ayrılmıştır. Sonrasında ağırlıklı olarak renk ve boyut bilgilerine göre kategorize edilen farklı bina çatıları, CNN kullanılır. Son olarak, endüstriyel ve konut bina çatıları ayrı ayrı tanınmış ve sonuçlar ayrı ayrı doğrulanmıştır. Değerlendirme sonuçları, önerilen yöntemin etkinliğini sırasıyla yaklaşık %91 ve %88 başarı oranları ile kanıtlamaktadır. Bir diğer çalışmada, görüntü verileri kullanarak çatı şekillerinin tanınmasına ilave olarak bina tespiti için Evrişimli Sinir Ağları'nın başarısı araştırılmıştır [9]. Çalışmadaki işlem adımları, veri seti oluşturma, modelin eğitimi, görüntü segmentasyonu, bina algılama ve çatı şeklini tanımayı içerir. İlk olarak, bir CNN mimarisi kullanılarak; ağaç, yol ve bina gibi kentsel nesneleri çıkarmak için eğitim yapılmıştır. Daha sonra, ikinci eğitilmiş CNN kullanılarak farklı çatı tiplerinin düz, üçgen ve sarkık şekillerine göre sınıflandırılması gerçekleştirilmiştir. Çalışmaya ait değerlendirme sonuçları, tespit ve tanıma adımlarında sırasıyla yaklaşık %97 ve %92 kalite oranları ile önerilen yöntemin etkinliğini kanıtlamaktadır. Başka bir çalışmada, Çin uydusu Gaofen-2'nin çok yüksek çözünürlüklü (VHR) görüntülerinden Evrişimsel Sinir Ağı kullanarak bina çatıları çıkarılmaya çalışılmıştır [8]. CNN'nin, nesne tabanlı sınıflandırma yaklaşımıyla sınıflandırılmayan ayrıntılı özellikleri tanıma konusunda daha yüksek bir kapasiteye sahip

olduğu kanıtlanmıştır. Bu çalışmada, eğitim veri kümesinin oluşturulması, model eğitimi, görüntü segmentasyonu ve bina çatılarının tanınması gibi birkaç ana adım söz konusudur. İlk olarak, ağaç, yol, meydan ve bina gibi kentsel nesneler nesne yönelimli sınıflandırma yaklaşımı ile rastgele orman algoritmasına göre sınıflandırılmış, yapı bölgeleri görsel olarak yorumlama ve düzeltme yardımı ile diğer sınıflardan ayrılmıştır. Daha sonra, ağırlıklı olarak renk ve boyut bilgilerine göre kategorize edilen farklı bina çatıları, eğitilmiş CNN kullanılarak eğitilmiştir. Son olarak, endüstriyel ve konut bina çatıları ayrı ayrı tanınmış ve sonuçlar ayrı ayrı doğrulanmıştır. Değerlendirme sonuçları, önerilen yöntemin, endüstriyel ve konut bina çatılarının tespitinde sırasıyla yaklaşık %91 ve %88 kalite oranları ile etkinliğini kanıtlamaktadır. Bu, CNN yaklaşımının binaları çok daha yüksek bir doğrulukla tespit etmede başarı gösterdiği anlamına gelmektedir. Diğer bir çalışmada uydu görüntüleri üzerinden bina yerini tespit eden ve binanın çatılarını sınıflandıran 43 katmanlı bir CNN mimarisi kullanılmıştır [10]. Binanın çatı materyalleri sınıflandırılmış, eksik ve tamamlanmamış olan binaların tespiti yapılarak bir afet durumunda meydana gelecek hasar artışını önlenmek amaçlanmıştır. Kullanılan mimari ile GoogleNet mimarisi kıyaslanmış ve %5 ila %7 oranında bir performans artışı gözlemlenmiştir. BRI-CNN model yaklaşımı tasarlanan diğer bir çalışmada, çok ölçekli çekirdekleri paralel olarak kullanarak iç, kenar ve bitişik özellikleri birleştirmek için yeni bir yaklaşım sunulmuştur [11]. İç spektral özellikler nedeniyle bina çatısı da dahil olmak üzere arazi örtüsü türlerinin tanınmasında en uygun olan 1×1 öznitelik matrisi seçilmiştir. 3×3 boyutundaki çekirdeklerin bina çatısının kenarında iyi çalıştığı ve 5×5 boyutundaki çekirdeklerin ise bitişik binaları ayırmak için verimli olduğu bulunmuştur. Yaklaşım Indian Pines (IP) veri seti kullanılarak test edilmiş ve bu karşılaştırmalı çalışma, BRI-CNN modelinin kapsül ağ modelinden en az %0,2 ve CNN modellerinden %2'den fazla daha yüksek genel doğruluk elde ettiğini göstermiştir. Başka bir çalışmada, görsel özelliklere sahip derin öğrenme ağlarının birleşimi, görünür farklılık bitki örtüsü indeksi (VDVI), Sobel kenar algılama özellikleri ve UAV görünür görüntülerinin farklı özellik kombinasyonlarının kırsal bina çatı tiplerinin model tanıma üzerindeki etkisini analiz etmek için değerlendirilmiştir [12]. Bu amaca ilaveten geçiş öğrenmeli ResNet152 öznitelik çıkarma ağı kullanılarak bina çatılarının farklı görüntü türlerinin daha karmaşık özelliklerini öğrenmek için geliştirilmiş bir Mask R-CNN modeli tasarlanmıştır. Üçgen, düz, kırma, karmaşık ve tek eğimli çatılar gibi kırsal alana ait çatı sınıfları belirlenmiştir. Çalışma sonunda iki sonuç elde edilmiştir. Elde edilen sonuçlarından birincisi, Sobel kenar algılama özelliklerini içeren RGB görüntülü modelin en yüksek doğruluğa sahip olmasıdır. Ayrıca,

farklı morfolojideki kırsal binaların çatı tiplerini giderek daha doğru bir şekilde tanımak ve RGB görüntülerine kıyasla model tanıma doğruluğu ortalama olarak 0,115 oranında iyileştirilmesidir. Çıkarılan ikinci sonuç, Orijinal Mask R-CNN, U-Net, DeeplabV3 ve PSPNet derin öğrenme modelleri ile karşılaştırıldığında, geliştirilmiş Mask R-CNN modeli, F1 puanı, KC ve KC ile kırsal binaların çatı tiplerini tanımada en yüksek doğruluğa sahip olduğunun gözlemlenmesidir. Başka bir çalışmada, GaoFen-2 görüntülerine dayalı DeeplabV3+ anlamsal segmentasyon sinir ağı, Çin'in önemli imalat sanayi üssü olan Nanhai bölgesinde Foshan'daki (Shishan, Guicheng, Dali ve Lishui kasabaları dahil) mavi çelik çatıların miktarını ve mekansal dağılımını analiz etmek için kullanılmıştır [13]. Elde edilen sonuçlara göre DeeplabV3+, maksimum olabilirlik sınıflandırmasından %92 sonucu ile daha yüksek bir genel doğrulukla tatmin edici bir performans göstermiştir.

Tez çalışması kapsamında yapılan araştırmanın temel sorularından birisi, evrimsel sinir ağları kullanılarak orta seviye çözünürlüklü veri seti içerisindeki bina bölgelerinin tespiti ve çatı sınıflandırılmasının yapılmasında elde edilecek başarı düzeyidir. Tez raporunun sıradaki bölümünde uzaktan algılama, derin öğrenme ve evrimsel sinir ağları hakkında bilgilere yer verilmiştir. Takip eden materyal ve metot bölümünde ise modelin eğitilmesi, arayüzün tasarlanma algoritması ve test aşamaları açıklanmıştır. Tez çalışmasının diğer bölümünde ise sırasıyla deneysel sonuçlar, genel sonuçlar ve değerlendirmeler yer almaktadır.

2. TEMEL KAVRAMLAR

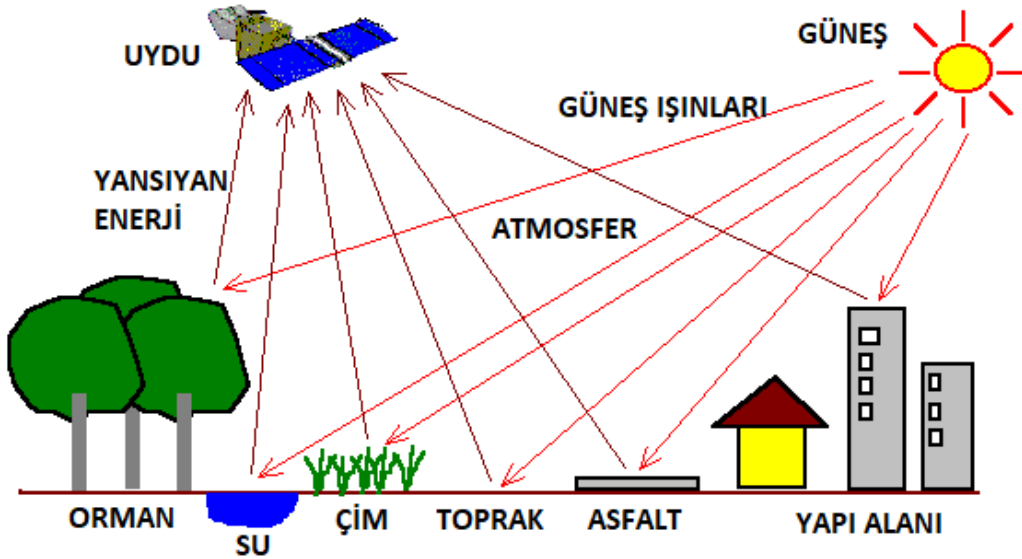
2.1. Uzaktan Algılama

Uzaktan algılama günümüzde birçok alanda kullanılan ve çokça duyduğumuz bir konudur. Uzaktan algılama belirli bir alana ait görüntü işleme çalışmalarında veri kaynağı olarak kritik bir öneme sahiptir. Yeryüzü ve geniş alanlara ait veri kaynakları bu yöntem aracılığıyla elde edilerek birçok çalışmada veri seti olarak kullanılmaktadır.

2.1.1. Uzaktan Algılama nedir?

Uzaktan algılama, arada fiziksel bir temas olmaksızın cisimler hakkında uzaktan bilgi toplanması veya arada mekanik bir temas olmaksızın bir cisimden yayılan ışının nitelik ve nicelik yönünden değerlendirilmesi ile cismin özelliklerinin uzaktan ortaya konması ve ölçülmesi şeklinde tanımlanabilir [14]. Diğer bir deyişle fiziksel bir temas olmadan belli bir mesafeden nesneler hakkında bilgi edinme yöntemidir.

Uzaktan algılamanın çalışma prensibini en temel seviyede anlatacak olursak, yeryüzüne giren aktif ve pasif ışık kaynakları nesneler üzerinden yansarak geri dönerler. Bu yansımalar bize ilgili nesne hakkında bilgi verir.



Şekil 2.1. Uzaktan Algılama Sistemi [15]

Uzaktan algılama tanımını biraz daha daraltacak olursak, geniş alanlarda yeryüzü ve üzerinde bulunan nesneler hakkında bilgi toplama amacıyla ışık kaynağı kullanılarak temas olmadan yapılan bir yöntemdir. Temel olarak çalışma prensibinden bahsetmek istersek uzaktan algılamada; algılayıcı, nesneler ve ışık kaynağı olmak üzere üç ana bileşen bulunmaktadır. Yeryüzü ve üzerindeki nesnelere gelen ışık ışınları nesne üzerinden yansır. Bu yansıyan ışınlar algılayıcılar tarafından algılanır. Algılanan bu ışınlar bize nesne hakkında bilgi edinmemize imkân tanır.

Uzaktan algılama sistemini Platform, Kullanılan Enerji ve Algılayıcı Sistem olmak üzere üç ana başlık altında inceleyebiliriz.

2.1.1.1. Platform

Uzaktan algılamada görüntü alanın büyüklüğü, görüntülere ait çözünürlük, görüntü elde etme maliyeti gibi birçok sebeplerden dolayı farklı platformlar kullanılmaktadır. Bu platformlar sırasıyla yer platformu, hava araçları platformu ve uydu platformudur.

2.1.1.1.1. Yer Platformu

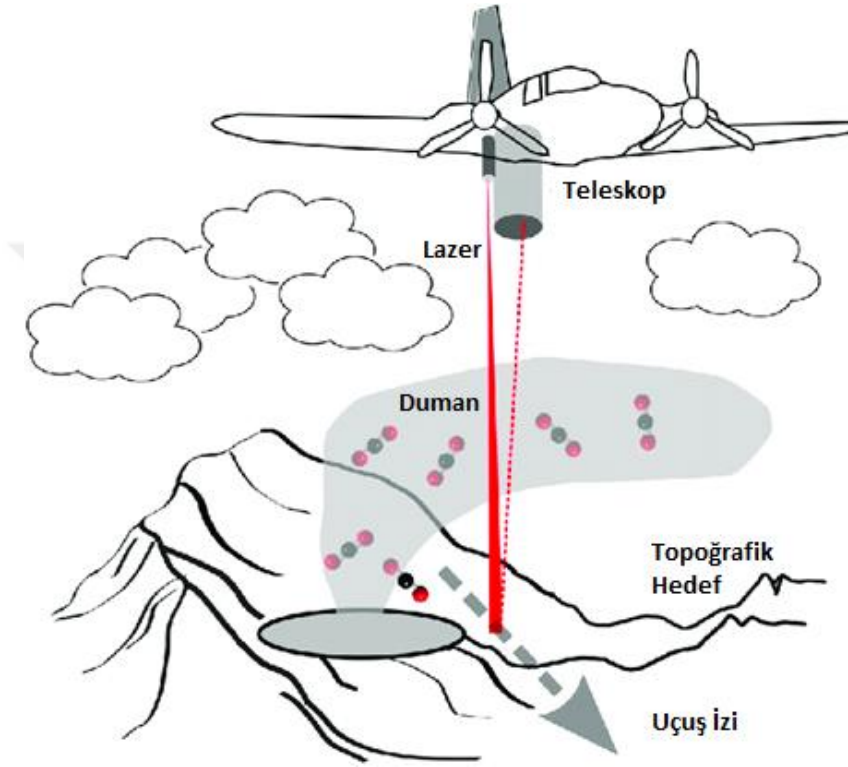
Yer platformları, yüksek çözünürlük ve ayrıntı gerektiren çalışmalarda kullanılmaktadır. Görüntü aldıkları alan küçüktür. Büyük alanlarda ve yüksek çözünürlük gerektirmeyen çalışmalar için yüksek maliyete sebep olur. Yer platformlarına ait bazı örneklere Şekil 2.2’de yer verilmiştir.



Şekil 2.2. Örnek Yer Platformları

2.1.1.1.2. Hava Platformu

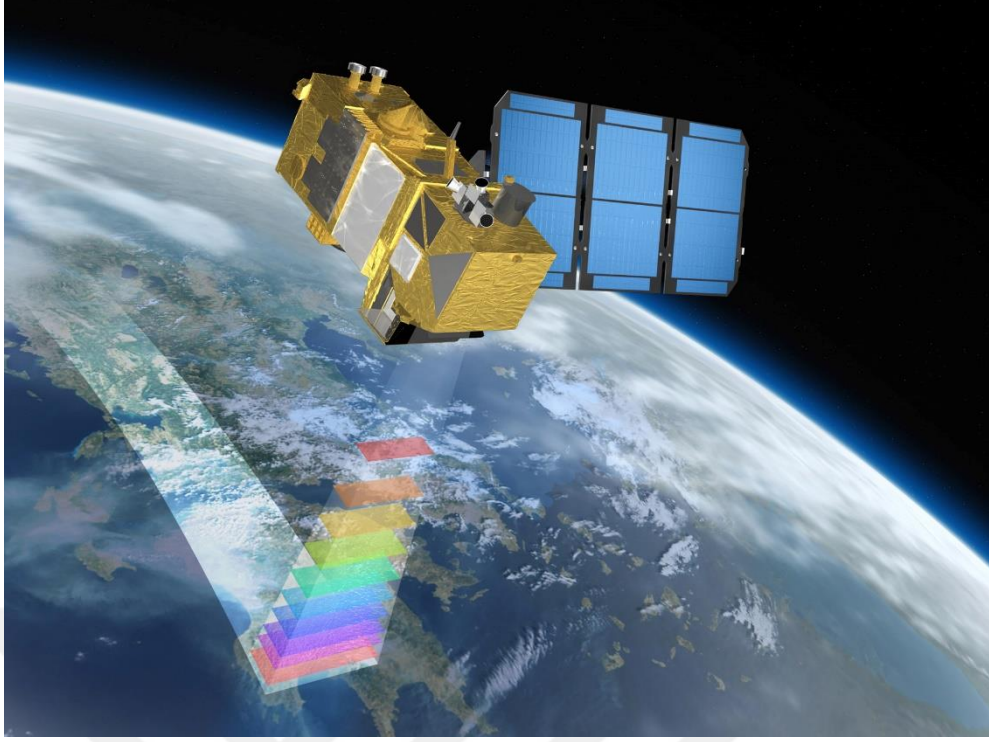
Uzaktan algılama sensörlerinin hava araçlarına monte edilmesiyle yapılan uzaktan algılama modelidir. Yer platformlarına göre daha düşük çözünürlükte görüntü elde edilir ancak daha geniş alanlara ait görüntüleme yapılabilir. Temsili bir hava platformu görüntüsüne Şekil 2.3'te yer verilmektedir.



Şekil 2.3. Örnek Bir Hava Platformu Görünümü [16]

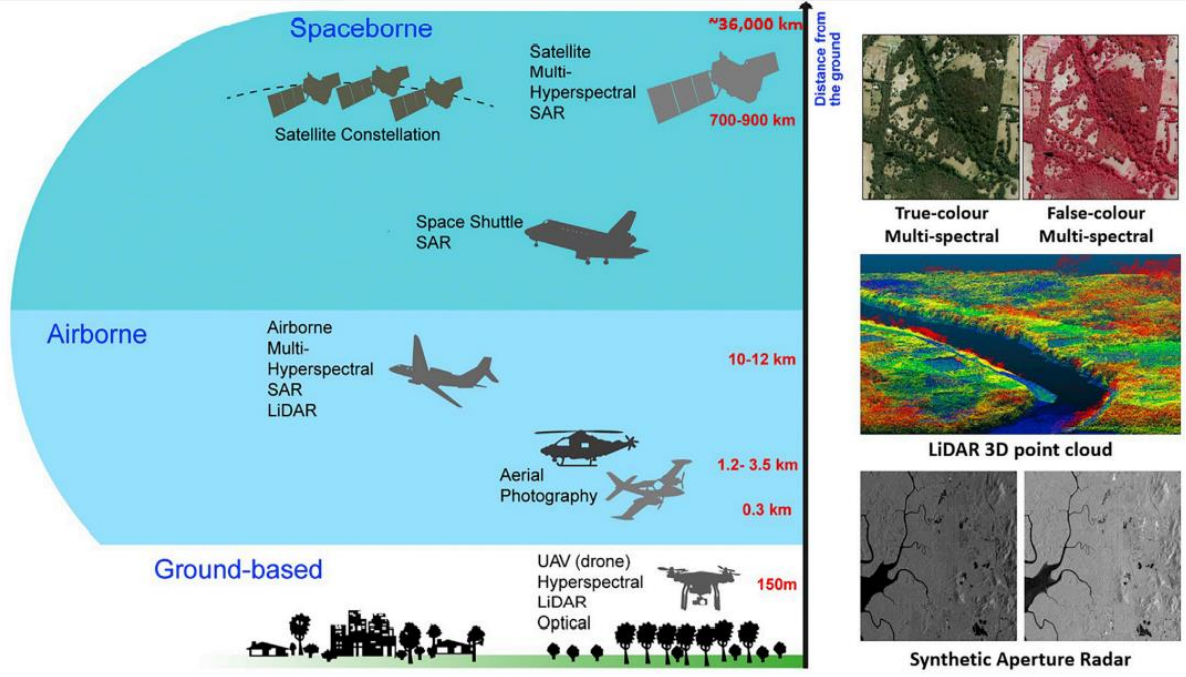
2.1.1.1.3. Uydu Platformu

Uzay platformlarına monte edilen sensörlerden oluşur. Çözünürlüğü diğer iki platforma göre en düşük olanıdır ancak bunun yanı sıra tarama alanı çok geniştir (Şekil 2.4).



Şekil 2.4. Örnek Bir Uydu Platformu Görünümü

Yapılacak olan uzaktan algılama çalışmasında; istenilen alan, çözünürlük, detay seviyesi gibi faktörlere göre hangi platformun seçilmesi gerektiği belirlenerek optimum maliyet tasarlanır. Bütün platform çeşitleri ve elde edilen görüntü örnekleri Şekil 2.5'te gösterilmektedir.



Şekil 2.5. Platformlar ve Elde Edilen Örnek Görüntü Verileri [17]

2.1.1.2. Kullanılan Enerji

Uzaktan algılama sistemleri kullandıkları enerji kaynaklarına göre ikiye ayrılır. Bunlar pasif ve aktif enerji kaynaklarıdır. Pasif enerji kaynağı güneş ışınlarıdır. Uzaktan algılama sensörleri güneş ışınlarının cisim üzerindeki yansımalarını algılar. Enerji kaynağının güneş olması sebebiyle görüntüleme sadece gündüz yapılabilir. Sensörler mor ötesi ve mikro dalga frekansları arasında algılama yapar. Aktif enerji kaynağı ise kendi elektromanyetik ışık kaynağına sahip olan algılama sistemleridir. LİDAR ve RADAR bunlara birer örnek olarak verilebilir. Aktif sistemler pasif sistemlere göre daha uzun dalga boyunda çalışırlar. Bu sayede karanlık, bulut ve sis gibi engellerde bile görüntüleme yapabilirler.

2.1.1.3. Algılayıcı Sistemler

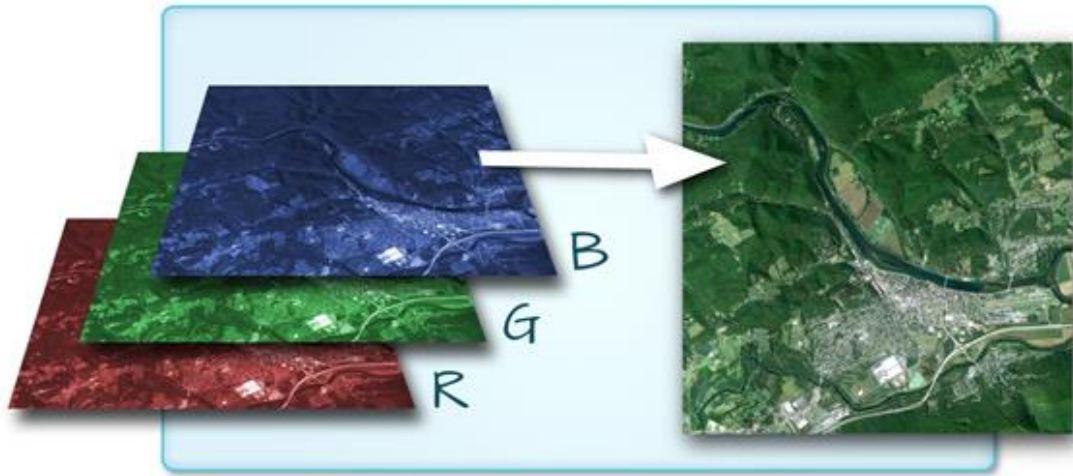
2.1.1.3.1. Pankromatik Görüntüleme

Pankromatik görüntü, elektromanyetik spektrumun geniş bir bölümünden yansıyan enerjiyi ölçen algılayıcı sistemler tarafından elde edilir. Pankromatik algılayıcılar için

bantlar, spektrumun görünen ile yakın kızıl ötesi bölümünü içerirler. Elde edilen bu görüntü veriler siyah-beyaz olurlar [18].

2.1.1.3.2. Multi-Spektral Görüntüleme

Multi-spektral görüntüleme sistemi pankromatik görüntüleme sisteminin aksine aynı alanı farklı dalga boylarında görüntüleme yaparak ve bu katmanları birleştirerek çoklu bir görüntüleme sağlar. Landsat 7 uydusu multi-spektral sisteme bir örnektir. Landsat 7 uydusunun TM sensörü mavi (bant1), yeşil (bant2), kırmızı (bant3), yakın-kızılötesi (bant4), orta-kızılötesi (bant5 ve 7) ve uzak-kızılötesi (bant6) olmak üzere toplam 7 bantta yansıyan ya da emilen ışınımı kaydederek multi-spektral tarama yapabilmektedir. Örnek bir multi-spektral görüntüye Şekil 2.6’da yer verilmiştir.

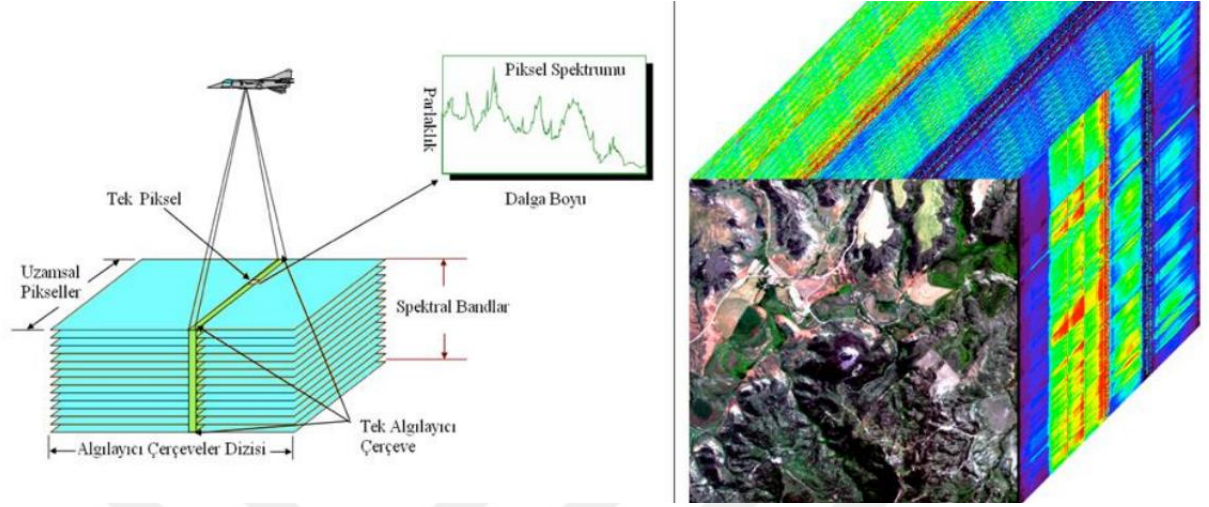


Şekil 2.6. Multi-Spektral Görüntü

2.1.1.3.3. Hiper-Spektral Görüntüleme

Hiper-spektral görüntüler, multi-spektral görüntülerin gelişmiş şekli olarak düşünülebilir. Aynı yüzeyin, elektromanyetik spektrumun morötesi, görünür bölge ve

kızılötesi dalga boylarında ardışık yüzlerce dar bantta görüntülenmesiyle meydana gelirler. Örnek bir hiper-spektral görüntü küpüne Şekil 2.7’de yer verilmiştir.



Şekil 2.7. Hiper-spektral Görüntü

2.1.2. Uydu Sistemleri ile Uzaktan Algılama

Uzaktan algılama teknolojilerinin kullanıldığı en önemli alanlardan birisi tematik harita üretimidir. Uydu görüntüleri, arazi örtüsü ve arazinin kullanımı belirlenerek tematik haritaların üretilmesi amacıyla kullanılmaktadır. Bu haritalar yardımıyla planlama veya doğanın korunmasına yönelik çeşitli çalışmalar yapılmaktadır. Doğal afete maruz kalan yerleşim yerlerinde hasar tespiti yapılması, uydu görüntülerinin kullanıldığı diğer bir alandır. Doğal afet sonucu hasar tespiti, hasar durumuna göre müdahale planlarının yapılması ve afet hasarlarının en hızlı şekilde giderilmesi için kullanılan uydu görüntüleri, günümüzde afet anında hayat kurtaran bir yöntem olmuştur. Uydu görüntüleri ülkelerin ekonomik ve siyasi kalkınma politikalarının oluşturulmasında dahi kullanılmaktadır. Haritacılık, hidroloji, tarım, maden, ormancılık, çevre ve şehircilik başlıca kullanım alanlarıdır.

Farklı kullanım amaçları için dünya üzerinde çeşitli teknik özellik ve çözünürlükte uydu sistemleri bulunmaktadır. Bu uydu sistemleri aşağıdaki tabloda gösterilmiştir (Tablo 1). Bu tablo içerisinde gösterilen uydu sistemlerinin konumsal çözünürlükleri 1 m ile 1 km, radyometrik çözünürlükleri 3 ile 12 bit ve zamansal çözünürlükleri 1 ile 40 gün arasında değişmektedir.

Tablo 1. Uydu Sistemleri ve Özellikleri [19]

UYDU	TARİH*	ALGILAYICI TİPİ	ÇÖZÜNÜRLÜK			ŞERİT GENİŞLİĞİ** (km)
			Konumsal (metre)	Radyometrik (bit)	Zamansal (gün)	
ALOS (NASDA)	2004-	PALSAR PRISM AVNIR-2	10-100 2.5 10	3 bit/5 bit 8 bit 8 bit	- - -	20-350
ALSAT-1 (Cezayir)	2002-	Multispektral	32	8 bit	4	600
ASTER/ TERRA (METI& NASA)	1999-	ASTER	15 30 90	8 bit 8 bit 12 bit	48 (16) 48 (16) 48 (16)	60 60 60
MODIS/ TERRA (METI&NASA)	1999-	MODIS	250 500 1000	8 bit 8 bit 8 bit	1-2 1-2 1-2	2330 2330 2330
BİLSAT-1 (TÜBİTAK - BİLTEN)	2003-	Pankromatik Multispektral	12 26	8 bit 8 bit	(116) 4 (52)	25 55
ENVİSAT (ESA)	2002-	ASAR MERIS	12.5-75 300	8 bit	3	54-400
ED-1 (NASA)	2000-	ALI HYPERION	10-30 30	12 bit 12 bit	16 (7-9) 16 (7-9)	37 7.7
EROS-A1 (Image Sat)	2000-	Pankromatik	1.8/0.9	11bit	1.8-4	12.5/6
ERS-1/2 (ESA)	91-00/ 94-	SAR Image Mode	30	-	3-35	100
IKONOS (Space Imaging)	1999-	SAR Pankromatik Multispektral	18 1 4	3 bit 11 bit 11 bit	44 3.5-5 3.5-5	75 11 11
IRS-1C/D (ISRO Hint.)	1995- / 97-	Pankromatik LISS-III LISS-III WIFS	5.8 23 70 188	6 bit 7 bit 7 bit 7 bit	24/25 24/25 24/25 24/25	70.5 141 141 812
IRS-P6 (ISRO Hint.)	2003-	LISS-III LISS-IV WIFS	24 6 60	7bit 7bit 10bit	24 5 24	140 24/70 mono 740
JERS-1 (Japonya NASDA)	1992-99	OPS OPS	18 18	6 bit 6 bit	44 44	75 75
KOMSAT-1 (G.Korea KARI)	1999-	EOC OSMI (ocean)	6.6	8 bit	-	17 800
LANDSAT - 1/2/3 (NASA-EOSAT)	1972-78/ 1975-82/ 1978-83	MSS	80	8 bit	18	180

* Uydunun faaliyete geçtiği ve kullanım dışı kaldığı yıllar (Örnek; 1972-78: 1972 yılında faaliyete geçmiş ve 1978 yılında kullanım dışı kalmış; 1999- : 1999 yılında faaliyete geçmiş ve halen faal)

** Uydu görüntüsünün geçiş izine dik yönde yeryüzünde kapsadığı alan

UYDU	TARİH	ALGILAYICI	ÇÖZÜNÜRLÜK			ŞERİT GENİŞLİĞİ (km)
			Konumsal (metre)	Radyometrik (bit)	Zamansal (gün)	
LANDSAT-4/5 (Space Imaging)	1982-87/ 1984-	TM	30	8 bit	16	183
		TM	30	8 bit	16	183
		TM	120	8 bit	16	183
LANDSAT -7 (NASA)	1999-	Pankromatik	15	8 bit	16	185
		ETM	30	8 bit	16	185
		ETM	60	8 bit	16	185
NIGERIASAT-1 (NIGERIA)	2003-	Multispektral	32	-	-	600
ORBVİEW -3 (ORBIMAGE)	2003-	Pankromatik	1	11 bit	8	8
		Multispektral	4	11 bit	8	8
RADARSAT -1 (CSA)	1995-	SAR	8-100	-	3-35 (24)	50-500
RADARSAT-2 (CSA-MDA)	2003-	SAR	3-100	-	3-35 (24)	20-500
SAC-C (Arjantin)	2000-	MMRS	175	-	9/7	360
		HRTC	35	-	16	90
		HSTC	250	-	2	1000
SeaStar (ORB-IMAGE NASA)	1997-	SeaWifs	4500 GAC 1100 LAC	10 bit 10 bit	1	150200 GAC 280100 LAC
SPIN-2 ve Mekikler (Rusya)	88/95 Periyodik	MK-4	15,8,15	-	8	150
		KVR-1000/DKI	2-1	-	-	40
		TK-350	10	-	-	200
SPOT-1/2/3 (CNES/ SPOT)	1986- / 1990- / 1993-96	HRV-PAN	10	8 bit	1-4(26)	60
		HRV	20	8 bit	1-4(26)	60
SPOT-4 (CNES/ SPOT)	1998-	HRV-PAN	10	8 bit	1-4(26)	60
		HRVIR	20	8 bit	1-4(26)	60
		HRVIR	20	8 bit	1-4(26)	60
		Vegetation	1000	4/8 bit	1	2200
SPOT-5 (CNES/ SPOT)	2002-	HRS-PAN	10	8 bit	1-4(26)	120
		HRG-PAN	2.5-5	8 bit	1-4(26)	60
		HRG	10	8 bit	1-4(26)	60
		HRG	20	8 bit	1-4(26)	60
		Vegetation	1000	4/8 bit	1	2250
QUICKBIRD-2 (Digital Globe)	2001-	Pankromatik	0,61-0,73	11 bit	3,5	16,5
		Multispektral	2,5-2,9	11 bit	3 5	16,5

Uzaktan algılama için kullanılan algılayıcılar optik ve radar olmak üzere ikiye ayrılır.

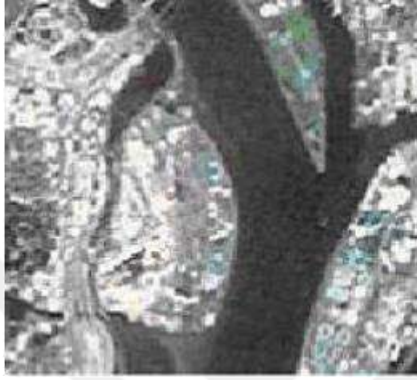
2.1.3. Optik Algılayıcılar

Optik algılayıcılar güneşten yayılan enerjinin yeryüzünden yansımaları algılayan sensörlerdir. Bu nedenle gece ve kötü hava şartlarında çalışamazlar. Uydu üzerine yerleştirilen algılayıcılar, elektromanyetik spektrumun optik dalga boylarında faaliyet gösteren algılayıcılardır.

Optik algılayıcılardan elde edilen görüntüler çözünürlüklerine göre 4'e ayrılır. Bunlar; Konumsal, Radyometrik, Zamansal ve Spektral çözünürlüklerdir.

2.1.3.1. Konumsal Çözünürlük

Konumsal çözünürlük görüntü içerisindeki piksellerin boyutlarını ifade eder. Konumsal çözünürlüğün düşmesi görüntü üzerindeki detayların artması ve görüntünün daha net olması anlamına gelmektedir. Farklı konumsal çözünürlüğe ait 4 adet görüntü Şekil 2.8’de gösterilmiştir.



LANDSAT TM (30 metre)



SPOT Pan (10 metre)



KVR1000 (2 metre)



IKONOS (1 metre)

Şekil 2.8 Farklı Konumsal Çözünürlükte Optik Uydu Görüntüleri

Konumsal çözünürlük üç gruba ayrılmaktadır. Bunlar; düşük, orta ve yüksek çözünürlüklü uydu görüntüleridir.

2.1.3.1.1. Düşük Çözünürlüklü Uydu Görüntüsü

Konumsal çözünürlüğü 100 metreden fazla olan görüntülerdir (OrbView2, MODIS). Piksel büyüklüğü 0,25 – 1,1 km arasında değişmektedir. 1/500.000 ve daha küçük ölçekli haritaların üretiminde kullanılır. Ormancılık açısından bu tür görüntüler genel bitki örtüsü hakkında bilgi sağlarlar. Düşük çözünürlüklü bir uydu görüntüsü örneği Şekil 2.9’da verilmiştir.



Şekil 2.9. Düşük Çözünürlüklü Sentinel-2 Görüntüsü

2.1.3.1.2. Orta Çözünürlüklü Uydu Görüntüleri

Konumsal çözünürlüğü 10 ile 100 metre arasında olan uydu görüntüleridir. Bu çözünürlükteki görüntüler çoğunlukla 1:50.000 ve 1:100.000 ölçekli veri tabanlarının hazırlanmasında kullanılır. Ulusal orman envanterinde arazi kullanım haritalarının oluşturulmasında kullanılmaktadırlar. Görüntü kaplama alanı olarak 30x30 km’den 180x180 km genişliğe kadar algılama yapılmaktadır. Landsat 7 ETM+, SPOT 5 ve ASTER bu tip çözünürlüğe sahip uydulardan bazılarıdır. Orta seviye çözünürlüklü bir uydu görüntüsü örneği Şekil 2.10’da verilmiştir.



Şekil 2.10. Orta Çözünürlüklü Sentinel-2 Görüntüsü

2.1.3.1.3. Yüksek Çözünürlüklü Uydu Görüntüleri

Konumsal çözünürlüğü 10 metreden daha düşük olan görüntülerdir. Bu görüntüler 1999 yılından itibaren kullanılmaktadırlar ve birçok alanda hava fotoğrafları yerine tercih edilmeye başlanmıştır. Yüksek çözünürlüğe sahip uydu görüntülerinin avantajları; sayısallık, yüksek geometrik doğruluk, topoğrafik kaymanın olmaması, yılın her zamanında veri alımı yapılması ve alanın büyüklüğüne bağlı kalmaksızın alım yapılabilmesi olarak sıralanabilir. Bu tür görüntüler veren uydulardan bazıları Ikonos ve Quickbird olarak sıralanabilir. Yüksek çözünürlüğe sahip iki farklı uydu görüntüsüne Şekil 2.11 ve Şekil 2.12’de yer verilmiştir.



Şekil 2.11. Yüksek Çözünürlüklü PlanetScope Görüntüsü

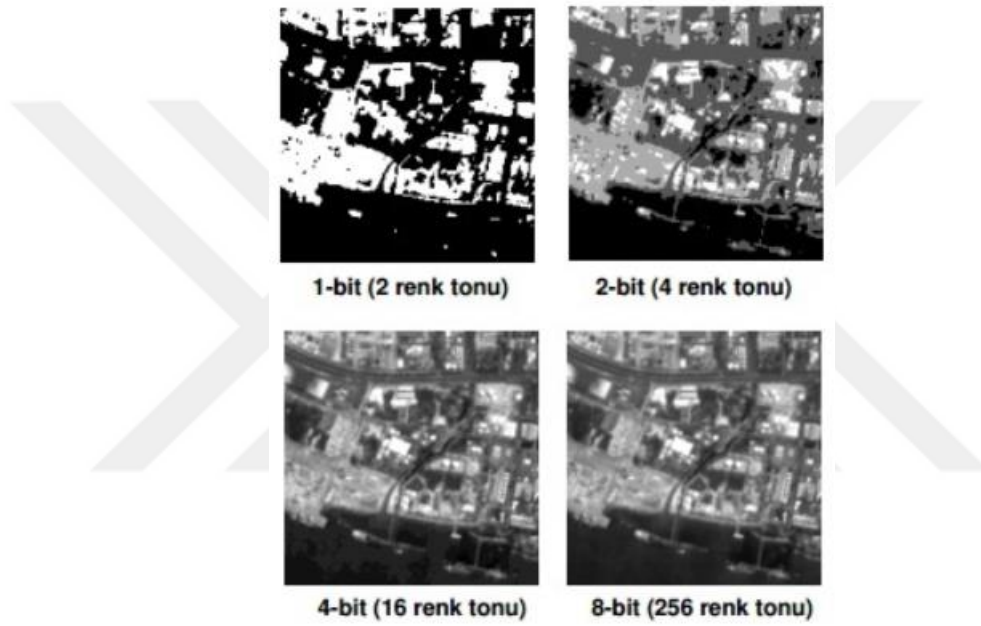


Şekil 2.12. Yüksek Çözünürlüklü SkySat Görüntüsü

2.1.3.2. Radyometrik Çözünürlük

Radyometrik çözünürlük, algılayıcının elektromanyetik enerjinin büyüklüğüne karşı duyarlılığını ifade eder. Görüntüleme sisteminin radyometrik çözünürlüğü enerjideki küçük farklılıkları ayırma yeteneği ile tanımlanır. Diğer bir anlatımla bu, kaydedilen enerjinin bölündüğü “bit” sayısıdır (<https://www.hvkk.tsk.tr/Custom/Hvkk/244>). Herhangi bir

görüntüye ait her bir pikselin alabileceği rakamsal değerleri kapsayan aralıktır. Örneğin, 8-bitlik bir radyometrik çözünürlüğe sahip görüntüde her bir piksel 0-255 arası bir değer alır. Bu değer, siyah ile beyaz arasındaki ton farklarına ait değerlerdir. Piksellerin görüntü değer aralığının artması görüntünün radyometrik çözünürlüğünü ve dolayısıyla ayırt etme hassasiyetini arttırmaktadır. Uyduların büyük çoğunluğunun radyometrik çözünürlüğü 8-bittir. Ikonos ve Quickbird gibi bazı uyduların radyometrik çözünürlükleri 11-bit olduğu için radyometrik çözünürlükleri diğerlerine göre daha yüksektir. Farklı radyometrik çözünürlüklere ait uydu görüntülerinden bazı örnekler Şekil 2.13'te yer almaktadır.



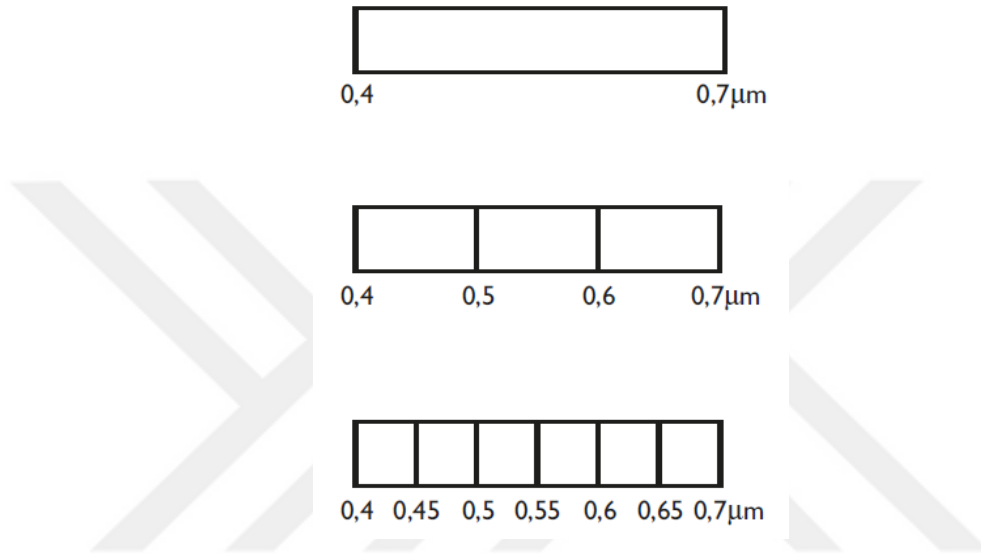
Şekil 2.13. Radyometrik çözünürlüğe ait görüntüler

2.1.3.3. Zamansal Çözünürlük

Uydular dünya etrafında belirli bir yörüngede hareket eder. Zamansal çözünürlük, uydunun bir noktadan görüntü aldıktan sonra aynı yörünge üzerinde tekrardan aynı noktadan aldığı görüntü arasında geçen süreye bağlı olan çözünürlüktür. Örneğin Landsat-TM algılayıcısı bir alanın görüntüsünü 16 günde bir alırken, NOAA-AVHRR algılayıcısı aynı alanı günde 2 defa algılamaktadır.

2.1.3.4. Spektral Çözünürlük

Spektral çözünürlük, algılayıcıların duyarlı olduğu dalga boyu aralığıdır. Bir bandın algıladığı dalga boyu aralığının küçük olması ve çok sayıda algılayıcı sistemin oluşması Spektral çözünürlüğün iyi olduğunu gösterir. Spektrumun küçük parçalara ayrılarak algılama yapması, spektral çözünürlüğün yüksek olduğunu ifade eder. Şekil 2.14’te farklı spektrum aralıklarına yer verilmiştir.



Şekil 2.14. Spektral Çözünürlüğe Ait Bant ve Kanal Değerleri

Yukarıdaki şekilde aynı dalga boyu aralığına ait farklı bant sayıları içeren algılayıcı bulunmaktadır. Bant sayısının artması diğer bir deyişle aralıkların çoğalarak daralması spektral çözünürlüğü arttırmaktadır. Aşağıdaki tabloda bazı uyduların algılayıcı sistemlerine ait dört farklı çözünürlükteki değerleri gösterilmektedir (Tablo 2).

Tablo 2. Çeşitli Algılayıcılar için Farklı Çözünürlüklerdeki Değerler

Algılayıcı	Piksel boyutu	Radyometrik çözünürlük (bit)	Spektral çözünürlük Bant sayısı	Spektral bölge	Zamansal çözünürlük (gün)
NOAA AVHRR	1 km (LAC) 4km (GAC)	10	5	Görünür YKö., Isıl Kö.	günde 2 kez
TERRA/ AQUA MODIS	0.25km 0.5km 1 km	12	1-2 3-7 8-36	Görünür/YKö./ Kısa Dalga Kö./ Isıl Kö.	1 gün
LANDSAT-7 ETM+	15m 30m 30m 60m	8	8	Pan Görünür, YKö. Kısa Dalga Kö. Isıl Kö.	16
SPOT-4 HRV	10m 20m	8	4	Pan Görünür, YKö.	26
SPOT-5 HRG	5m 10m 20m	8	5	Pan Görünür, YKö. Kısa Dalga Kö.	2-3 (Enleme bağlı)
RapidEye	5m	12	5	Görünür, Ykö.	1 gün (5 uydu ile)
QuickBird	0.6m 2.4m	11	5	Pan Görünür, YKö.	1-4 (Enleme bağlı)
IKONOS	1m 4m	11	5	Pan Görünür, YKö.	3
GeoEye-1	0.5m 1.7m	11	5	Pan Görünür, YKö.	2-8
Worldview-2	0.5m 1.8m	11	9	Pan Görünür, YKö.	1-4

2.1.4. Radar

Radar sistemleri, enerji kaynağını kendileri ürettikleri için aktif sistemler olarak anılırlar. Araziye ait bilgileri toplayabilmek için gerekli enerjiyi üretir ve bu enerjiyi cisimlerin üzerine gönderir. Gönderilen bu enerji, nesneden yansarak geri döner. Sistem yansıyan bu enerjiyi ölçer. Radar sistemleri, elektromanyetik spektrumun mikrodalga aralığında faaliyet göstermektedir. Dünyayı gözlemleyen radar uydularından başlıcaları şunlardır:

- RADARSAT-2 (SAR, Kanada, 2007)

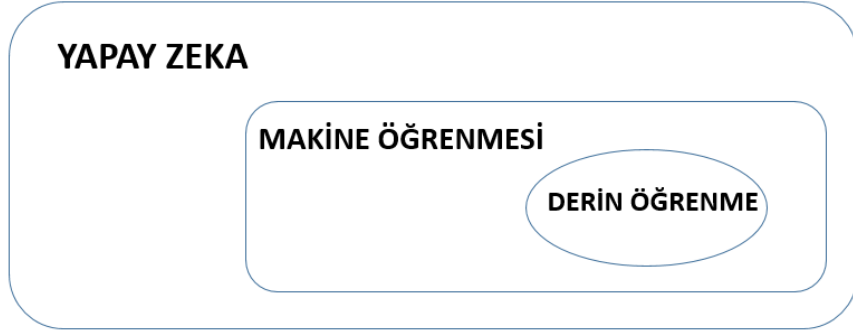
- ERS-1 & ERS-2 (Avrupa Uzaktan Algılama Uydusu) (altimetre, birleşik SAR/scatterometer)
- JERS-1 (SAR Japonya, 1992)
- TanDEM-X (SAR Almanya, 2010)
- RISAT-1 (SAR, ISRO Hindistan, 2012)

Radar görüntüleri, topoğrafik haritaların üretiminde, sivil ve askeri uygulama alanlarında en az kullanılan görüntülerdir. Radar sistemlerinin kullanım alanları ve geometrik düzeltmeleri üzerine çok sayıda araştırma bulunmaktadır. Özellikle radar görüntülerinin rektifikasyonunda, radarın algılama geometrisinden kaynaklı sorunlar bulunmaktadır. Bu sorunların çözümü için yeni algoritmalar geliştirilmiştir. Yapılan bir çalışmada polinomlarla rektifikasyon ve yükseklik düzeltmesine dayalı bir yöntem kullanılmış ve bu yöntem RADARSAT, ERS-2 ve uçağa takılı SAR sistemi görüntüleri üzerinde test edilmiştir. Sonuçlar görüntü türüne ve arazi yapısına bağlı olarak özellikle dağlık alanlarda 2 ila 5 piksel arasında doğruluklara ulaşılabileceğini göstermiştir [20].

2.2. Derin Öğrenme

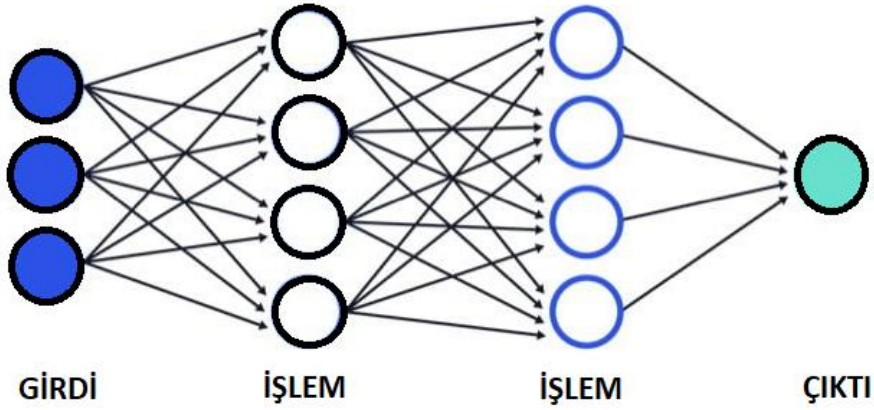
2.2.1. Derin Öğrenme (Deep Learning) Nedir?

Derin öğrenmenin bir makine öğrenmesi yöntemi olduğunu söylememiz yanlış bir tabir olmayacaktır. Derin öğrenme, yapay sinir ağları olarak adlandırılan, insan beyninin yapısından ve işlevinden ilham alan algoritmaları destekleyen bir yöntemdir. Derin öğrenme özellik çıkarma ve bu özellikleri işleme üzerine çalışmaktadır. Özellik çıkarımında birçok doğrusal olmayan işlem katmanını kullanılmaktadır. Bu yöntemde önceki katmanın çıktısı bir sonraki katmanın girdisi olacak şekilde ardışık katmanlar kullanılmaktadır. Derin öğrenmede, verilerin birden fazla özellik seviyesinin veya temsillerinin öğrenilmesine dayanan bir yapı söz konusudur. Üst düzey özellikler, alt düzey özelliklerden türetilerek hiyerarşik bir temsil oluşturur. Bu temsilde, soyutlamanın farklı seviyelerine karşılık gelen birden çok temsil seviyesini öğrenir [21]. Derin öğrenmenin yapay zekâdaki yerini aşağıdaki şekilde daha iyi görebiliriz (Şekil 2.15).



Şekil 2.15. Derin Öğrenme ve Yapay Zekâ İlişkisi

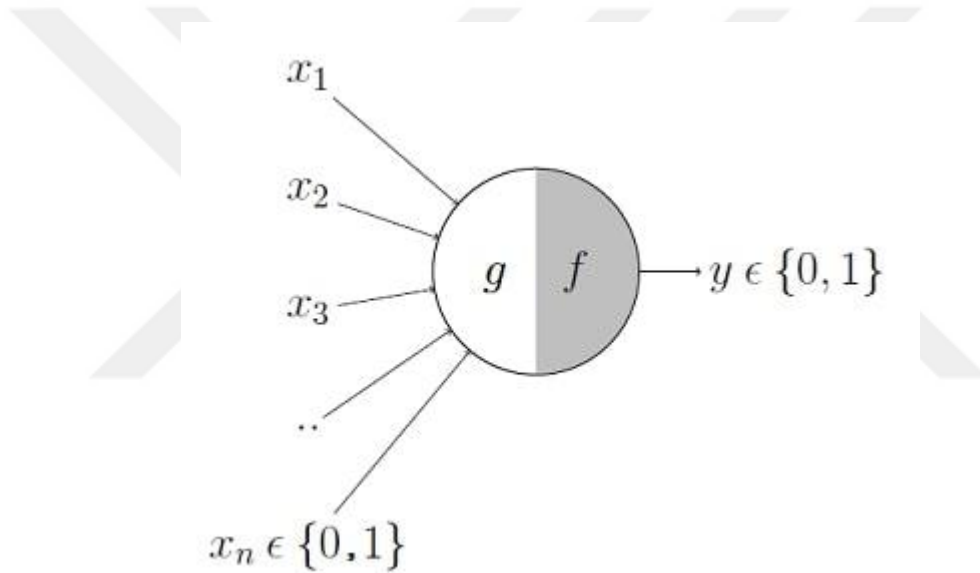
Derin öğrenmenin çalışma mimarisi üç temel adımdan meydana gelir. Bunlar; girdi, işlem ve çıktı adımlarıdır. İşlem adımlarının çok olması öğrenme modelinin derinliğini belirlemektedir. Girdi olarak alınan değer, işlem adımlarından geçirilerek öznitelik çıkarımları yapılmaktadır. Ortaya çıkan özniteliklerin hepsi ayrı bir katman olarak oluşturulur. Bu işlem adımlarında işlem sonrası elde edilen çıktı, bir sonraki işlemin girdisi olarak kullanılır. Mimari elde ettiği öznitelik katmanlarından çıkarımlar yaparak bir çıktı elde etmektedir. Şekil 2.16’da sistemin ana çalışma adamları sunulmuştur.



Şekil 2.16. Derin Öğrenme Sistemi

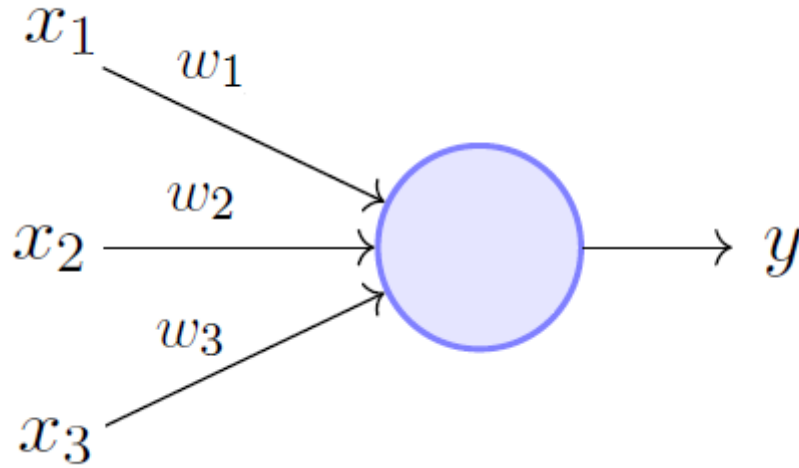
2.2.2. Derin Öğrenmenin Tarihi

Derin öğrenme 1940 senesinde basit bir yapay sinir ağı modeli ile başlamış, büyük gelişimler göstererek günümüze kadar gelmiştir. İnsan beyninin düşünme işlevini inceleyerek bir sinir ağının matematiksel modelini ortaya koyan mantıkçı Walter Pitts ve sinir bilimci Warren McCulloch “A Logical Calculus of Ideas Immanent in Nervous Activity” adlı ortak çalışmalarını ortaya çıkarmış ve derin öğrenmenin ilk adımını atmışlardır. Walter Pitts ve Warren McCulloch’un basit bir sinir ağı modeli, girişlerden ve girişlerden elde edilecek bir çıkıştan ibarettir. X girdileri Y ise çıktıyı belirtmektedir (Şekil 2.17). İlk yapay sinir ağı yapısı olarak çıkan bu model, diğer modellere de ışık tutmuştur.



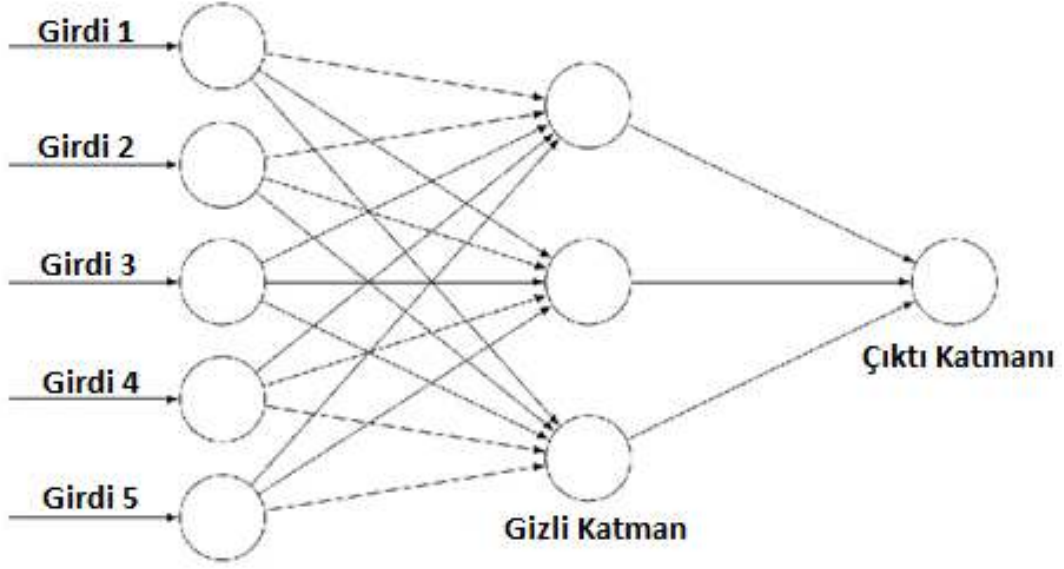
Şekil 2.17. Yapay Sinir Ağı Modeli (<https://www.elektrikport.com/haber-roportaj/derin-ogrenme-nedir-serisi-2-bolum/23407#ad-image-0>)

Yeni bir kavram olan “Perceptron” 1950-1960 tarihinde karşımıza çıkmaktadır. Cornell Üniversitesi’nden psikolog Rosenblatt tarafından bir yapay sinir ağı olarak önerilen Perceptron’lar ismini sadece girdi ve çıktı katmanları bulunan ağlardan yani tek katmanlı sinir ağlarından (Perceptron) alır. Basit tek katmanlı sinir ağı modeli olan Perceptron, temel olarak eğitilebilecek tek bir yapay sinir hücresinden oluşmaktadır. En eski sinir ağlarından olan Perceptron’lar son derece sınırlıdır ve bu sebeple karmaşık problemlerin çözümünde etkisizlerdir.



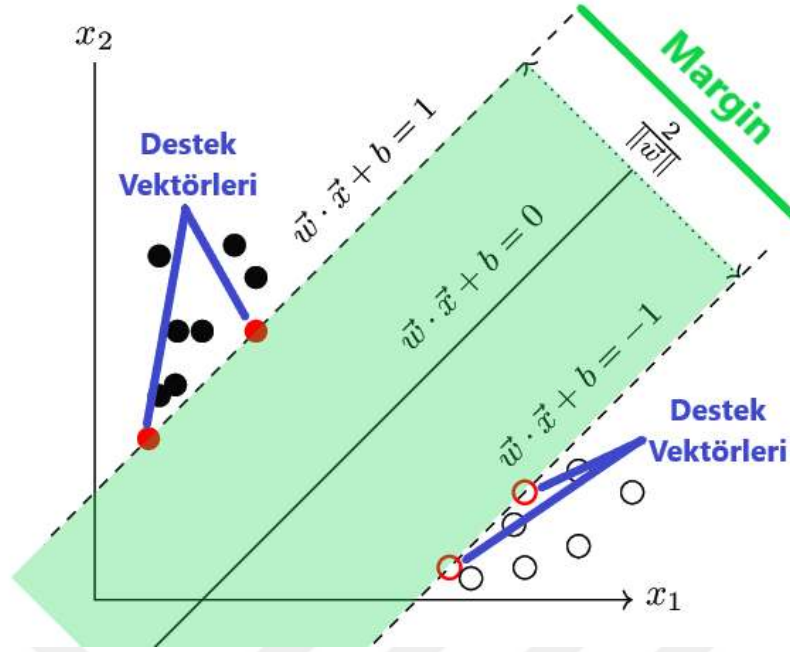
Şekil 2.18. Perceptron Model [22]

Şekil 2.18'de de görüldüğü gibi Perceptron yapısı tek katmanlı ve sadece giriş/çıkışa sahip olan bir işlem ünitesidir. Burada x değerleri girişleri, w değerleri ise ağırlıkları göstermektedir. 1970'li yıllarda tek katmanlı algılayıcıların doğrusal olmayan durumların çözümünde başarılı olmadığı anlaşılmıştır. Günlük hayatta karşılaşılan problemler her zaman doğrusal değildir, bu sebeple doğrusal olmayan problemlerin çözümü önem arz etmektedir. Perceptron yapısının XOR çözümlerinde başarısız olması sebebiyle doğrusal olmayan durumların çözülemeyeceği iddia edilmiştir. Bu sebeple yapay sinir ağı çalışmaları aniden durma noktasına kadar gelmiştir. Bu durum yapay zekâ çalışmalarının da önüne geçmiştir. Bu yüzden 1986 yılına kadar olan bu dönem “Karanlık Dönem” olarak adlandırılır. 1980'lerin ortalarında ise multi-layer perceptron'ların (çok katmanlı algılayıcılar) geliştirilmesi ile XOR problemi çözülmüş ve doğrusal olmayan problemlerin çözülebileceği anlaşılmıştır. Çok katmanlı bir ağ yapısına Şekil 2.19'te yer verilmiştir.



Şekil 2.19. Çok Katmanlı Ağ Modeli

Çok katmanlı algılayıcılar çok yüksek işlem gücüne gereksinim duydukları için farklı bir algoritma peşinde olan araştırmacılar yeni bir algoritma keşfetmişlerdir. 1995 yılında V. Vapnik ve C. Cortes yaptıkları araştırmalar sonucunda Destek Vektör Makineleri (Support Vector Machine) adlı bir öğrenme algoritması geliştirdiler (Şekil 2.20). Destek Vektör Makineleri, temelde iki farklı sınıfa ait verileri en uygun şekilde birbirinden ayırmak için kullanılır. Bunun için karar sınırları ya da diğer bir ifadeyle hiper düzlemler belirlenir. Veri setini ikiye ayıran bu doğru, karar doğrusu olarak adlandırılır. Sınırsız karar doğrusu çizebilme imkânı olsa da önemli olan en uygun karar doğrusunu belirlemektir.



Şekil 2.20. Destek Vektör Makineleri

2006 yılında Geoffrey Hinton tarafından “Çok Katmanlı Yapay Sinir Ağları”nın daha verimli bir şekilde eğitilebileceğinin öne sürülmesiyle “Derin Öğrenme” kavramı ortaya çıkmıştır. Derin sinir ağlarının ortaya çıkmasıyla yapay sinir ağlarının işlem gücünün artması sağlanmıştır. 2000 yılından sonra özellikle bilgisayar donanımlarının hızla gelişmesi yapay zekâ çalışmalarını daha da ateşlemiştir.

Stanford Üniversitesi Yapay Zekâ Laboratuvarı başı olan Profesör Fei-Fei Li, 2009’da ImageNet’i başlatmıştır. ImageNet ile derin öğrenme, evrimsel yapay sinir ağı modelleri kavramını sözlüğüne eklemiş ve derin öğrenmede hızlı ilerleyecek bir yol açılmıştır. Nesne tanımlama alanında en büyük yarışma olan Büyük Ölçekli Görsel Tanıma Yarışması (ImageNet Competition 2012) o yıl derin öğrenmede temel mimari kabul edilen Evrimsel Sinir Ağı (ESA) [23] ile kazanılmıştır. Bu gelişme, Derin Öğrenmenin inanılmaz bir yükselişi olmuştur. Çünkü yarışmada %26,1 olan Top-5 hata oranı %15,3 gibi bir orana düşürülmüştür [23]. Derin Öğrenmedeki ilerlemeler Top-5 hata oranını %3,6’ya kadar düşürmüştür [24]. Günümüzde evrimsel yapay sinir ağları yapay zekâ sektöründe kullanılan en popüler yöntemlerden biridir. Yıllar içinde yapılan çalışmalarda evrimsel yapay sinir ağları ciddi bir yol kat etmiş ve kendini ispatlamıştır.

2.2.3. Derin Öğrenmenin Kullanım Alanları

Derin öğrenme günümüze kadar önemli bir yol almıştır. Yapılan çalışmalarda elde edilen yüksek başarı oranları, derin öğrenmenin daha etkin bir kullanıma sahip olmasını sağlamıştır. Günümüzde derin öğrenme pek çok alanda başarılı bir şekilde kullanılmaktadır. Derin öğrenmenin başlıca kullanıldığı alanlardan bazıları aşağıda gösterilmektedir.

- 1- Görüntü İşleme
- 2- Sınıflandırma
- 3- Nesne Tespiti
- 4- Ses Verilerinin Sınıflandırılması
- 5- Medikal Alanda Teşhis ve Tanı
- 6- Endüstriyel Alanda Tespit ve Sınıflandırma

2.2.3.1. Görüntü İşleme

Basit nesnelere ait görsellerin işlenmesi ile başlayan derin öğrenme, bu alanda kendini geliştirilmiştir. Görsel verilere ait düşük çözünürlüklü görüntülerin çözünürlüklerini iyileştirme, nesne tespiti, medikal çalışmalar, parmak izi ve göz taraması gibi biyometrik verilerin işlenmesini konu alan birçok kullanım alanları vardır. Yüksek doğruluk ve düşük hata oranı derin öğrenmenin en dikkat çekici özelliklerinden biri olsa da, işlem süresinin kısalması diğer bir önemli özelliktir. Bu alanda yapılan bazı çalışmalardan bahsedecek olursak, birincisi Google Brain araştırmacılarından Dahl, Norouzi ve Shlens tarafından yapılan çalışmadır. Derin öğrenme yöntemi kullanılarak bir yüzün çok düşük çözünürlüklü bir görüntüsüne dayanarak yüzün nasıl görüldüğünü anlamaya çalışan uygulama gerçekleştirmişlerdir. Video görüntülerindeki yüzlerin çözünürlüğü çok düşük olduğu için yüzler görülemez ve net bir şekilde görüntülenemediğinden bunu ortadan kaldırmak için bir uygulama uygulamışlardır. Olasılıksal derin öğrenme mimarileri kullanılarak düşük çözünürlüklü yüz görüntülerini yüksek çözünürlüklü yüz görüntülerine dönüştürmeye çalışılmıştır [25]. Isola ve arkadaşları, yapmış oldukları çalışmada pix2pix yazılımında kullandıkları bir derin öğrenme yaklaşımını oluşturmuşlardır. Derin öğrenme yöntemini kullanarak, haritalanmış bir görüntüden renklendirilmiş ve tamamlanmış yeni bir görüntü elde etmişlerdir. Çalışmalarında bu derin öğrenme yaklaşımı pix2pix yazılımında kullanılmıştır [26]. Iizuka, Simo-Serra, Ishikawa tarafından yürütülen araştırmada, gri

tonlamalı siyah beyaz görüntülerin renklendirilmesi başarıyla sağlanmıştır [27]. Bu çalışmada, evrişimli sinir ağlarına dayalı derin öğrenme algoritmasına ek olarak bir füzyon katmanı kullanılmıştır. Genel ve bölgesel önceliklere sahip tüm çerçeveler uçtan uca eğitilir. Bu modeli eğitmek için geniş bir sahne sınıflandırma veritabanı kullanılmıştır. Bu çalışmada, yüzlerce yıl öncesine ait görüntüler otomatik olarak renklendirilmiştir. Bu araştırma, derin öğrenme ağları tarafından kendisine verilen kalıpları öğrenmiştir. Örneğin; denizin mavi olması, gökyüzünün mavi olması ve bulutların beyaz-gri olması gibi. Yapılan bu çalışmada gri seviyeli görüntüleri renklendirmek için farklı sonuçlar ortaya çıkabilmektedir. Ancak, aynı resim için farklı renklendirmeler sunulabileceği için bu durumun insan tarafından fark edilmesi de pek mümkün değildir [28].

2.2.3.2. Sınıflandırma

Görüntü sınıflandırılmasında görüntü içerisinde tanımladığımız nesneler, belirlenen bazı kriterler doğrultusunda etiketlenerek sınıflandırılır. Örnek verecek olursak, bir görüntü içerisinde farklı renklere ait bardakları renklerine göre sınıflandırmak diyebiliriz. Görüntü sınıflandırılması ile ilgili yapılmış pek çok çalışma mevcuttur. Derin öğrenme ile sınıflandırmada; uygulama, çalışma alanı ve yöntemlere göre farklı algoritmalar kullanılmaktadır. Bu sebeple algoritmalar her çalışmada farklı sonuçlar verebilmektedir. Sınıflandırma için kullanılacak en uygun derin öğrenme algoritmasını seçmek, çalışmanın istenilen başarıya ulaşmasında ilk adım olarak düşünülebilir. Sınıflandırma; uydu görüntülerinin sınıflandırılması, trafik işaret ve plakaların sınıflandırılması, araçların sınıflandırılması gibi birçok problemin çözümünde kullanılmaktadır. Sınıflandırmanın kullanıldığı diğer bir önemli alan da medikal sektördeki teşhis ve tanı koyma uygulamalarıdır. Trafik işaretlerinin sınıflandırılması [29], uydu görüntülerinin sınıflandırılması [30], [31], [32], yüz tanıma-sınıflandırma [33], [34], [35], [36], hiperspektral verilerin sınıflandırılması [37], tomografi görüntülerinden akciğer keseciklerinin sınıflandırılması [38], EEG verilerinin sınıflandırılması başlıca yapılan çalışmalardan sadece birkaçıdır [39]. Sınıflandırmada elde bulunan etiketlere göre sınıflandırma yapılırken nesne tespitinde görüntü içerisinde yer alan bir nesnenin konumu tespit edilir.

2.2.3.3. Nesne Tespiti

Nesne tespiti, görüntü içerisindeki nesnelerin yerlerinin tespit edilmesi anlamına gelmektedir. Nesne tespitinde derin öğrenme yaklaşımları başarılı sonuçlar elde etmişlerdir. Nesne tespiti ile sınıflandırma birbirine benzer görünse de birbirinden farklıdır. Sınıflandırmada eldeki etiketlere göre görüntülerin ait oldukları sınıflar belirlenir. Nesne tespitinde ise bir görüntü içerisinde bir nesnenin varlığı aranarak konumu belirlenmeye çalışılır [40]. Özellikle 2014 yılından sonra nesne tespiti çalışmalarında derin öğrenmenin kullanılmasıyla başarılı sonuçlar elde edilmeye başlanmıştır. Derin öğrenmede kullanılan ve geliştirilen birçok nesne tespiti algoritması bulunmaktadır. Elde edilen başarılı çalışmalar nesne tespitinde derin öğrenmenin kullanılmasıyla daha iyi sonuçlar alınabileceği konusunda umut olmuştur. Özellikle son zamanlarda bilinen algoritmalarda yapılan değişiklikler ile algoritmalar modifiye edilerek daha iyi sonuçlar bulma arayışları ortaya çıkmıştır. Bir çalışmada, AlexNet mimarisinin son katmanı değiştirilerek DetectorNet adlı bir algoritma önerilmiştir. Aynı çalışmada, bir görüntüdeki aynı nesnenin birden fazla tespitine izin veren bir derin öğrenme mimarisi olan DeepMultiBox önerilmiştir [41]. Elde edilen sonuçların oldukça başarılı olduğu raporlanmıştır.

2.2.3.4. Ses Verilerinin Sınıflandırılması

Derin öğrenme artık günümüzde ses verilerinin tanımlanması ve sınıflandırılması gibi çalışmalarda da kullanılmaktadır. Ses verilerinin eğitilmesi ile seslerin ait olduğu canlı ve nesnelerin tanımlanması, ses verilerinin işlenerek farklı alanlarda kullanılması başlıca çalışmalardan bazılarıdır. Salamon ve Bello ses verilerini kullanarak sesin ait olduğu nesne ve canlıların sınıflandırmasını yapmıştır. Bu çalışmada siren, araba kornası, silah sesi, sokak müziği, sondaj, köpek havlaması, klima, çekiç, oynayan çocuklar ve araç motoruna ait 10 adet sınıf belirlenmiştir. Kısa ses verileri ile oluşturulan bu sınıflara ait bir sınıflandırma yapılmıştır [42]. Fried ve Fiebrink ses haritaları oluşturmak için derin öğrenme metodu geliştirmişlerdir. Görüntü haritasını çıkarabilmek için sesin tanımlanması yapılmıştır. Sese karşılık gelen görüntüler eşleştirilmiştir. Bu çalışmada derin öğrenme mimarilerinden oto kodlayıcılar kullanılmıştır [43].

2.2.3.5. Medikal Alanda Teşhis ve Tanı

Derin öğrenmenin girdiği çalışma alanlarından biri de medikal alandır. Bilgisayar sistemlerinin insan gözünden daha hassas değerlendirme yapabildiği bilinmektedir. Sağlık alanında teşhis ve tanı konmasında insan gözünün yetersiz kalabildiği bazı durumlar söz konusu olmaktadır. Sağlıkta erken teşhisin hayat kurtaran bir durum olması ve bilgisayar sistemlerinin teşhis konusunda insan gözünden çok daha hassas noktaları yakalayabileceği düşünüldüğünde, derin öğrenmenin bu alanda hayati bir konuya el attığı aşikârdır. Röntgen, tomografi, laboratuvar örneklerinin derin öğrenme ile detaylı incelenmesi sayesinde tespit edilmesi zor olan hasarlı veya zararlı organizmalar tespit edilebilmektedir. Bir çalışmada diz kireçlenmesi oranını, radyolojik görüntülerden derin konvolüsyonel sinir ağlarını kullanarak ölçmeye çalışan bir yöntem hakkında çalışılmıştır. Bu çalışmada VGG16 ve AlexNet algoritmalarından faydalanmışlardır [44]. Bir diğer çalışmada ise 17 adet hastalık sınıfının tespitinin nasıl yapıldığı sunulmuştur. Burada göğüs röntgen görüntüleri değerlendirilmiş, daha sonra GoogleNet ve AlexNet ile kıyaslanmıştır [45].

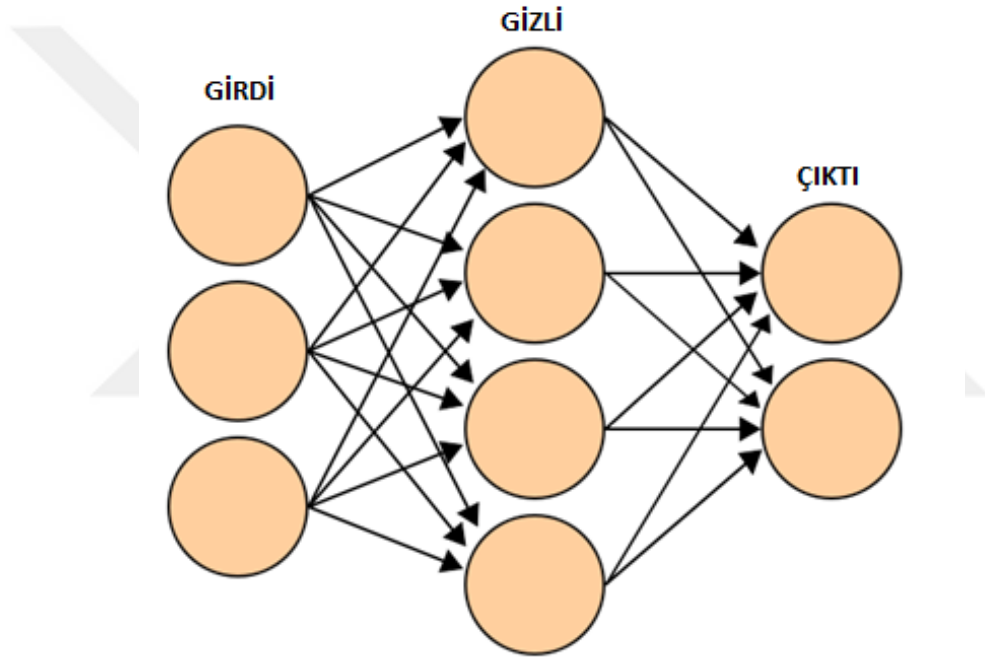
2.2.3.6. Endüstriyel Alanda Tespit ve Sınıflandırma

Derin öğrenmenin kullanıldığı diğer bir alan ise endüstridir. Üretim günümüzün en önemli konularından biridir. Üretimde minimum hata, düşük maliyet, hız ve kalite önemli etkenlerdir. Derin öğrenme algoritmalarını bu otomasyonlara entegre ederek daha hızlı çalışmanın yanında daha hatasız sonuçlar almak, derin öğrenmenin sektörde ileriye dönük büyük umutlar taşımasına neden olmaktadır. Üretimin iyileştirilmesinde, firenin daha az verilerek hammaddenin daha etkin kullanılması, hızlı üretim bantlarında hataların daha hassas tespit edilebilmesi gibi birçok amaçla kullanılmaktadır. Bir çalışmada, nesnelerin duruş-poz tahmini yapmak, onları kavramak ve seçilen hedefe doğru yönelen bir robot görme sistemi sunulmuştur [46].

2.3. Evrimsel Sinir Ağları (CNN)

Yapay Sinir Ağı, insan beyninden esinlenilerek oluşturulan bir yöntemdir. Bu yöntemde insan beyninin sinir yapısı kullanılmaktadır. Her nöron bilgi işlem birimini temsil

eder ve birbirine bağı yapay nöronlardan oluşur. Bu nöronlara düğüm de denilmektedir. Düğümler kendi aralarında etkileşimdedir ve her düğüm bir sonraki düğüme bilgi taşımaktadır. Her düğüm girdi (input) alır ve işlemlerden geçirerek bir sonraki düğüme çıktı (output) verisi aktarır. Bu düğümdeki çıktı verisi bir sonraki düğümde girdi olarak alınarak bu iletim devam eder. Doğrusal olmayan matematiksel bir fonksiyon olan aktivasyon fonksiyonu ile bu işlemler gerçekleştirilir. Her düğümün bir ağırlığı vardır ve bu ağırlıklar sürekli güncellenmektedirler. Doğruluk oranı yüksek olan ağırlıklar güncellenmezken düşük olan ağırlıklar öğrenmedeki her geri beslemede doğruluğu arttırmak için güncellenirler. Örnek bir yapay sinir ağı Şekil 2.21’de gösterilmektedir.



Şekil 2.21. Yapay Sinir Ağı

Şekilde de görüldüğü gibi en soldaki katman giriş (input) katmanıdır. Birçok gizli (hidden) katmandan geçirildikten sonra çıktı (output) katmanına ulaşarak çıktıyı oluşturur. Bu gizli katmanlar içerisinde birçok düğüm barındırır. Düğümlerin artması daha fazla veri çıkarımı yapılabilmesini sağlar. Bu sebeple temelde düğüm sayısı arttıkça doğruluk artar. Ancak, düğüm sayısının belli bir sayıdan daha fazla arttırılması tersine bir etki yaratarak doğruluğu azaltmaya başlar. Düğüm sayısındaki artışın diğer bir olumsuz etkisi ise her düğümde işlem yapılması sebebiyle donanımsal olarak sistemin zorlanmaya başlamasıdır.

Bu zorlanma sebebiyle işlem süresinin uzaması ve donanımın yetersiz kalması gibi olumsuz durumlar oluşabilmektedir.

Yapay sinir ağları öğrenme esasına dayalı, deneyimlerden yararlanan bir sistemdir. Öğrenme ile oluşturduğu deneyimi kullanarak karar veren bir mekanizmaya sahiptir. Cnn (Convolutional Neural Network) görüntüyü çeşitli katmanlarda işler. Bu katmanlar sırasıyla Convolutional layer, ReLu, Pooling, Flattening Layer ve Fully-Connected Layer'dır.

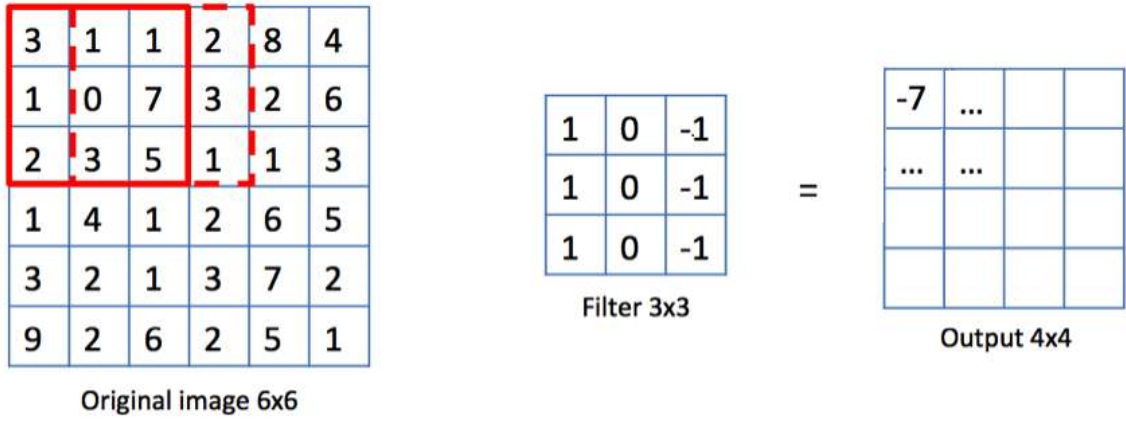
2.3.1. Convolutional Layer

Convolutional Layer CNN'nin en önemli katmanıdır. Bu katmanda görüntüye ait düşük ve yüksek seviyede özellikler çıkarılır. Bu çıkarım için bir filtre kullanılır. Aşağıdaki şekilde convolutional layer'a ait 3x3 bir filtre gösterilmektedir (Şekil 2.22).

1	0	-1
1	0	-1
1	0	-1

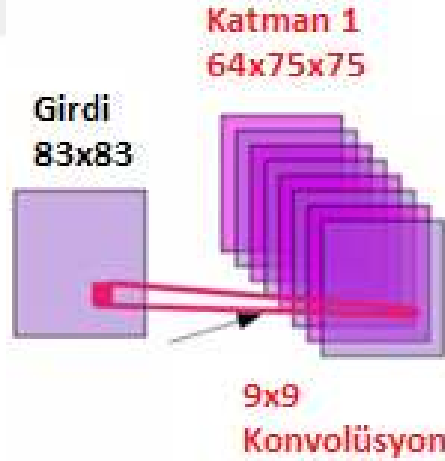
Şekil 2.22. 3 x 3 Convolutional Layer Filtresi

Bu filtre yardımıyla görüntü üzerinden çıkarımlar yapılır. Birden çok özelliği tespit etmek için birden fazla konvolüsyonel filtre kullanılır. Aşağıdaki şekilde konvolüsyonel filtrenin nasıl çalıştığı gösterilmektedir (Şekil 2.23).



Şekil 2.23. Konvolüsyonel Filtrenin Çalışma Prensibi

Konvolüsyonel katmanda seçilen filtreler ana görüntü üzerinde işlenerek bir çıktı oluşturulur. Her konvolüsyonel işlem sonucunda ana görüntü üzerinden bir özellik çıkarımı yapılır. Her özellik çıkarımı bir katmana ait olmaktadır. Şekil 2.24'te görüldüğü gibi, konvolüsyonel işlem sonrası görüntünün boyutu azalırken kayman sayısı artmaktadır.

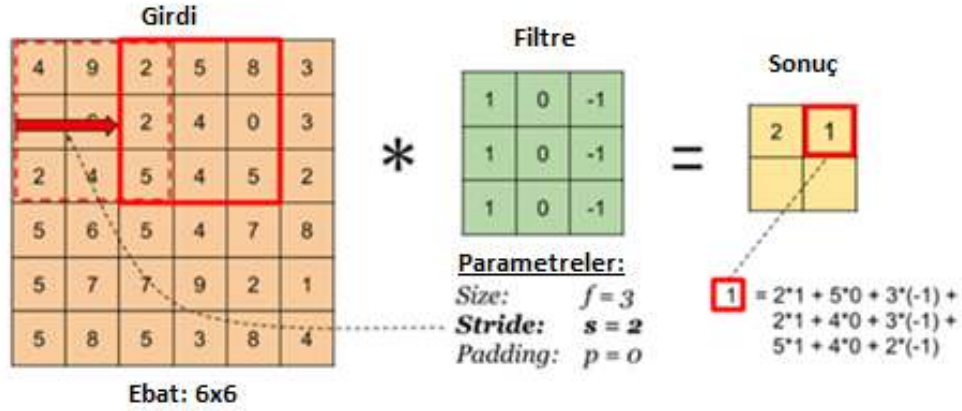


Şekil 2.24. Konvolüsyonel Çıktı

2.3.1.1. Stride

Stride, sinir ağındaki konvolüsyonel filtrenin bir parametresi olup görüntü veya video üzerindeki hareket miktarını değiştirir. Aşağıdaki görselde de görülebileceği gibi, bir sinir

ağında stride değeri 2 olarak ayarlanırsa, filtre bir seferde iki piksel veya iki birim hareket eder.



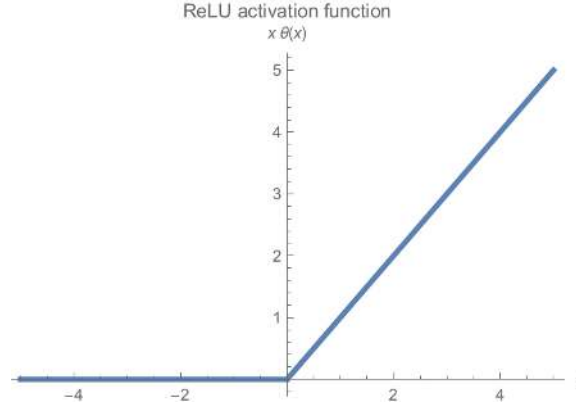
Şekil 2.25. Stride işleminin etkisi

2.3.1.2. Padding

Bir imgeye filtre uygulandığında boyutlarından dolayı çıktı orijinal görselden daha küçük hale gelir. Bunu önlemek için kullanabileceğimiz yöntem ise padding yani dolgulamadır. Dolgulama işleminde görüntüye dört taraftan da sıfırlar eklenir. Filtrenin boyutuna göre eklenen bu sıfır değerleri artırılabilir.

2.3.2. ReLu

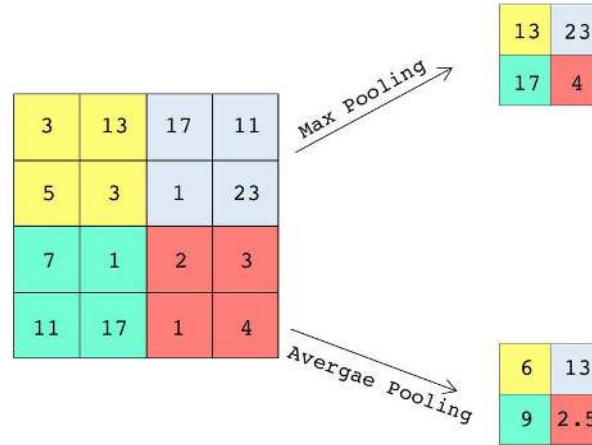
ReLu (Rectified Linear Unit) $f(x) = \max(0, x)$ şeklinde çalışan doğrusal olmayan bir fonksiyondur. Basit bir anlatımla negatif değerlerden kurtulmamızı sağlayan bir fonksiyondur. CNN işlemlerinde çok önemli olan bu fonksiyon, modeli ya da bazı özellikleri negatif değerlerden dolayı kavrayamamasının önüne geçmek için kullanılır. Örnek bir ReLu değer grafiği Şekil 2.26'da görülmektedir.



Şekil 2.26. ReLu Değer Grafiği

2.3.3. Pooling

Pooling yani havuzlama katmanı konvolüsyonel katmandaki gibi görüntünün boyutunu küçültür. Amacı görüntünün boyutunu küçülterek gereksiz özellikleri atmak ve görüntüyü daha düşük boyuta indirgemesi sayesinde işlem süresi ve hacmini azaltmaktır. Performansı doğrudan etkileyen bu katman, CNN mimarisi için önemli bir işlem adıdır. İki farklı pooling tekniği vardır. Bunlar; Max Pooling ve Average Pooling'dir. Her iki yönteme birer örnek Şekil 2.27'de verilmiştir.



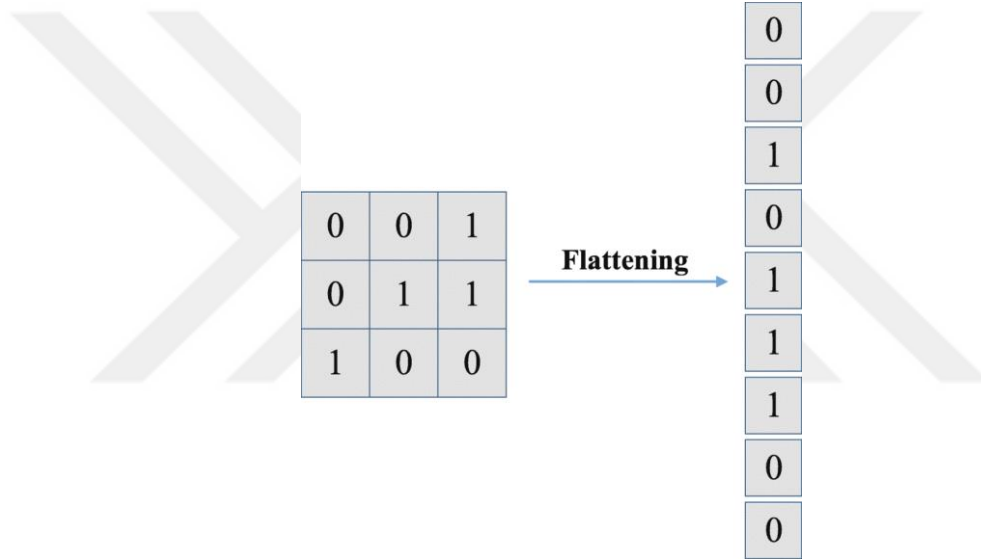
Şekil 2.27. Pooling İşlemi

Yukarıdaki şekilden de anlaşılacağı gibi evrimsel katmandaki gibi bir filtreye sahip olan pooling, bu filtre değeri kadar görüntüyü bölerek içerisindeki değerleri alır ve görüntü

boyutunu küçültür. İki yöntem arasındaki fark ise; birinde filtre boyutu kadar seçilen ana görüntü içerisindeki en büyük değer alınırken diğerinde bu değerlerin ortalaması alınır.

2.3.4. Flattening

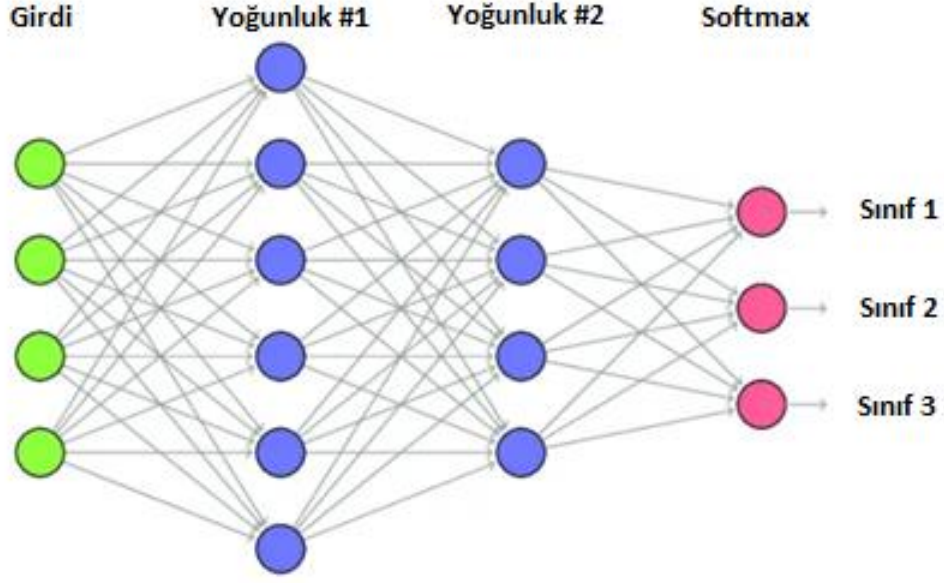
Konvolüsyonel ve havuzlama işlemleri matris üzerine işlem yapan ve çıktısını da matris olarak veren adımlardır. CNN'nin son işlem adımı olan Fully Connected Layer (tam bağlantılı katman) tek boyutlu diziyi girdi olarak kullanır. Flattening katmanının görevi matris olarak alınan çıktıyı tek boyutlu diziye çevirerek Fully Connected Layer için girdi oluşturmaktır. Örnek bir düzleştirme işlemine Şekil 2.28'de yer verilmiştir.



Şekil 2.28. Flattening İşlemi

2.3.5. Fully Connected Layer

Bu katman yapay sinir ağı ile öğrenme işleminin gerçekleştiği bölümdür. Bu katmanda tek sıra dizi haline getirilen özellik çıkarımları yapay sinir ağı ile değerlendirilerek öğrenme gerçekleştirilir (Şekil 2.29). Sonrasında elde edilen değerler softmax gibi bir sınıflandırma katmanına aktararak sınıflandırma işlemi tamamlanır.



Şekil 2.29. Fully Connected İşlemi

2.4. R-CNN, Fast R-CNN ve Faster R-CNN

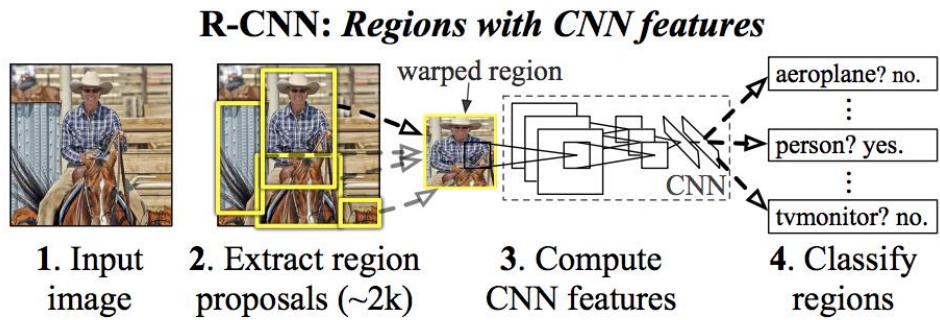
2.4.1. R-CNN

Region based Convolutional Neural Network yani kısaltması R-CNN, CNN üzerine kurulu bir yöntemdir. Adından da anlaşılacağı gibi bölge bazlı bir metottur. Bu yöntemin geliştirilmesinin ana sebeplerinden biri görüntü içerisindeki çoklu obje seçiminde yaşanan sorunlardır. R-CNN mimarisi, görüntülerde bulunan objelere ait sınıfları ve bu objelere ait bounding box'ları (sınırlayıcı kutular) tespit etmek için kullanılır. R-CNN temelde iki adımdan oluşmaktadır. Birinci adım olan selective search (seçici arama) ile görselde obje olmaya aday özellikler belirlenir. Selective search görsellerde yakalanması gereken bölgeleri belirlemek için kullanılır. Küçükten büyüğe doğru çalışan yöntemde öncelikle küçük bölgeler belirlenir. Sonra birbirine benzer olan bölgeler birleştirilir ve daha büyük bölgeler ortaya çıkar. Bu işlem sürekli olarak devam ederek daha büyük bölgeler ortaya çıkar ve görselde bulunan objeler bir nevi gruplandırılır. R-CNN ile gerçekleştirilen örnek bir obje tespiti uygulamasına ait görsel Şekil 2.30'da verilmiştir.



Şekil 2.30. R-CNN Bölgesel İşaretleme

Birinci adımda yaklaşık 2000 bölge belirlendikten sonra ikinci adımda her biri ayrı ayrı bir CNN modeline gönderilir ve sınırları yani bounding box'ları tahmin edilir. CNN(ConvNet) işlemi ile elde edilen özellik haritalarının SVM kullanılır. R-CNN yapısının işlem adımlarına Şekil 2.31'de yer verilmiştir.



Şekil 2.31. R-CNN İşlem Adımları

Alınan sonuç umut verici olmasına rağmen, bu yöntemin de bazı olumsuz yönleri bulunmaktadır. Bu sorunlardan biri eğitim ve test sürelerinin uzun sürmesi, diğeri ise işlem yükünün fazlalığı sebebiyle donanımsal olarak sistemin zorlanmasıdır. Bu sorunlar üzerine yapılan çalışmalar sonucunda Fast R-CNN yaklaşımı ortaya çıkmıştır.

2.4.2. Fast R-CNN

R-CNN yönteminde karşılaşılan işlem ve zaman sorunlarını çözmek amacıyla Fast R-CNN yöntemi geliştirilmiştir. Bu yöntem R-CNN'nin aksine önce ana görüntüye CNN yaklaşımını uyguladıktan sonra ortaya çıkan COV5 özellik haritası üzerinde bölge önerilerini ortaya çıkarır. Diğer bir farkı ise sınıflandırmada SVM yerine yapay sinir ağı katmanları içinde gerçekleştirebileceği bir derin öğrenme sınıflandırması olan softmax yöntemini kullanmasıdır. Bu yöntem sayesinde eğitim süresi yaklaşık 10 kat düşürülebilmüş ve tahmin süresi de yaklaşık 2-3 saniyeye kadar indirilmiştir.

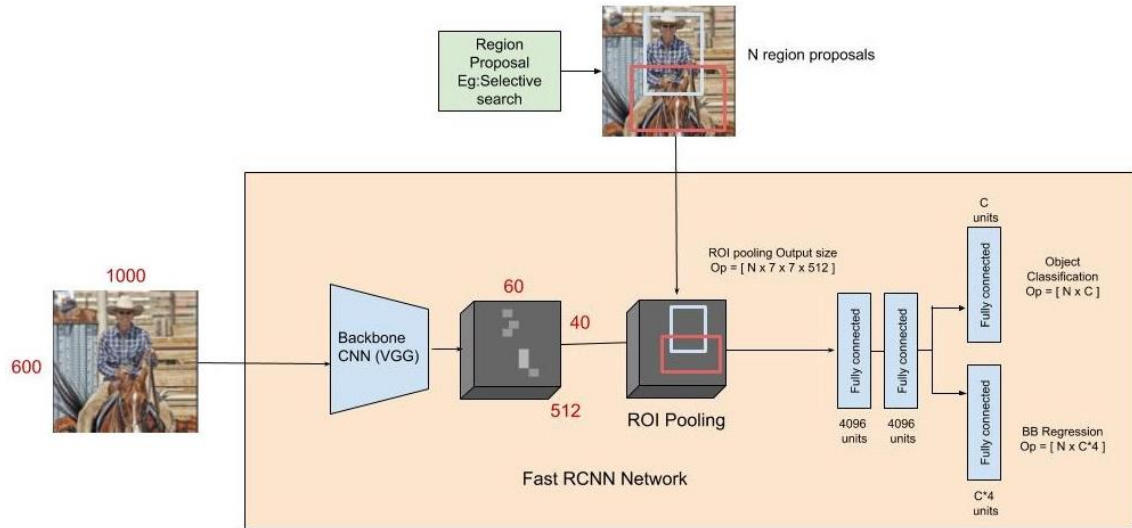
2.4.3. Faster R-CNN

Bu yöntem ile, Fast R-CNN yöntemine benzer şekilde öncelikle görüntüyü CNN işlemine sokarak özellik haritası çıkarılır. Bu noktadan sonra Fast R-CNN'e göre değişiklik başlar. Bölge önerileri kısmı selective search ile değil daha kullanışlı olan Region Proposal Network ile yapılmaktadır. Faster R-CNN yaklaşımı Region Proposal Network (RPN) ve Fast R-CNN şeklinde iki aşamada incelenebilir.

2.4.3.1. Region Proposal Network (RPN)

RPN bölge önermeye yarayan derin, evrişimli bir sinir ağıdır. İlk aşamada RPN, girdi olarak herhangi bir boyutta girdiyi alır ve obje skoruna göre bir dizi nesnelere ait olabilecek dikdörtgen tekliflerini ortaya çıkarır. Bu öneriyi, evrişimli katman tarafından oluşturulan öznitelik haritası üzerinde küçük bir ağı kaydırarak yapar.

RPN tarafından gerçekleştirilen bu hesaplamalar Fast R-CNN mimarisine sokulur ve bir sınıflandırıcı ile objenin sınıfı, regressor (değişken) ile de bounding box'u tahmin edilir. Örnek bir Faster R-CNN yapısına Şekil 2.32'de yer verilmiştir.



Şekil 2.32. **Faster R-CNN yapısı**

3. MATERİYAL VE METOT

Gerçekleştirilen çalışma üç başlık altında incelenebilir. Bunlar; model eğitimi, arayüz tasarımı ve test adımlarıdır.

3.1. Model Eğitimi

Eğitim aşaması, modelin işlevselliği ve yüksek doğruluk oranı için önemli bir adımdır.. Eğitim başarısını etkileyen çeşitli etmenler bulunmaktadır. Bunlar; eğitimde kullanılacak veri seti büyüklüğü ile çözünürlük kalitesi, evrimsel sinir ağı mimari yapısı ve eğitim için seçilen parametrelerdir. 1989 yılında LeNet ile başlayan CNN macerası zamanla daha iyi sonuçlar almamızı sağlayan farklı mimari yapılar oluşturularak günümüze kadar gelmiştir. LeNet, AlexNet, VGGNet, GoogleNet, Resnet, ZFNet gibi birçok farklı CNN mimarisi geliştirilmiş ve bu mimariler birçok problemin çözümünde kullanılmıştır. Daha önce yapılan çalışmaları incelediğimizde her mimarinin farklı alanda başarılı sonuçlar vererek ön plana çıktığını görmekteyiz. Bu tez çalışmasında bina bölge tespiti ve çatı sınıflandırılmasında AlexNet, GoogleNet ve DarkNet53 mimarileri kullanılarak başarı oranları incelenmiştir. Elde edilen sonuçlara göre, görece düşük çözünürlüklü görüntülerden oluşan veri setinin üç mimarideki başarı oran çıkarımları yapılmıştır.

3.1.1. AlexNet

ImageNet Büyük Ölçekli Görsel Tanıma Yarışması ile ilk kez ortaya çıkan AlexNet, CNN kullanarak yarışmayı %80 doğruluk oranıyla büyük bir farkla kazanmıştır. Bu ağ ilk kez, öğrenme yoluyla elde edilen özniteliklerin el ile tasarlanmış öznitelikleri aşabildiğini ve bilgisayarla görmede daha önceden alışlagelen kalıpların kırılabileceğini göstermiştir. AlexNet mimarisinde 5 adet konvolüsyon katmanı, 3 adet maxpool katmanı, 2 adet dropout katmanı, 3 adet tam bağlı katman, 7 adet relu katmanı, 2 adet normalizasyon katmanı, softmax katmanı, giriş ve sınıflandırma (çıkış) katmanı olmak üzere toplam 25 katman bulunmaktadır. Aşağıdaki şekil bu katmanların sıralamasını göstermektedir (Şekil 3.1). Giriş

katmanı 227x227x3 boyutlarında verilerden oluşmaktadır. Son katmanda ise sınıflandırma yapılarak giriş görüntüsündeki sınıflandırma sayısı değeri elde edilir.

1	1x1 ImageInputLayer
2	1x1 Convolution2DLayer
3	1x1 ReLULayer
4	1x1 CrossChannelNormalizationLayer
5	1x1 MaxPooling2DLayer
6	1x1 GroupedConvolution2DLayer
7	1x1 ReLULayer
8	1x1 CrossChannelNormalizationLayer
9	1x1 MaxPooling2DLayer
10	1x1 Convolution2DLayer
11	1x1 ReLULayer
12	1x1 GroupedConvolution2DLayer
13	1x1 ReLULayer
14	1x1 GroupedConvolution2DLayer
15	1x1 ReLULayer
16	1x1 MaxPooling2DLayer
17	1x1 FullyConnectedLayer
18	1x1 ReLULayer
19	1x1 DropoutLayer
20	1x1 FullyConnectedLayer
21	1x1 ReLULayer
22	1x1 DropoutLayer
23	1x1 FullyConnectedLayer
24	1x1 SoftmaxLayer
25	1x1 ClassificationOutputLayer

Şekil 3.1 AlexNet Katman Yapısı

3.1.2. GoogleNet

GoogleNet, ImageNet veri tabanında %93 doğruluk oranıyla yüksek bir başarı elde etmiştir. Alexnet'e göre çok daha fazla katmana sahiptir. GoogleNet algoritmasında konvolüsyon, maxpooling, softmax, fully connected, relu, giriş katmanı ve çıkış katmanı

gibi 144 adet katman bulunmaktadır. 224x224x3 boyutunda görüntüyü girdi verisi olarak almaktadır. Konvolüsyon katmanında kullanılan filtreler 1x1, 3x3 ve 5x5 boyutundadır. 3x3 boyutunda havuzlama kullanılmaktadır. Aktivasyon için doğrusal aktivasyon kullanılmaktadır [47]. GoogleNet mimarisine ait işlem adımları aşağıdaki şekilde gösterilmektedir (Şekil 3.2).



1	1x1 ImageInputLayer	49	1x1 Convolution2DLayer	97	1x1 Convolution2DLayer
2	1x1 Convolution2DLayer	50	1x1 ReLULayer	98	1x1 ReLULayer
3	1x1 ReLULayer	51	1x1 Convolution2DLayer	99	1x1 MaxPooling2DLayer
4	1x1 MaxPooling2DLayer	52	1x1 ReLULayer	100	1x1 Convolution2DLayer
5	1x1 CrossChannelNormalizationLayer	53	1x1 ReLULayer	101	1x1 ReLULayer
6	1x1 Convolution2DLayer	54	1x1 DepthConcatenationLayer	102	1x1 Convolution2DLayer
7	1x1 ReLULayer	55	1x1 Convolution2DLayer	103	1x1 ReLULayer
8	1x1 Convolution2DLayer	56	1x1 ReLULayer	104	1x1 Convolution2DLayer
9	1x1 ReLULayer	57	1x1 Convolution2DLayer	105	1x1 ReLULayer
10	1x1 CrossChannelNormalizationLayer	58	1x1 ReLULayer	106	1x1 Convolution2DLayer
11	1x1 MaxPooling2DLayer	59	1x1 Convolution2DLayer	107	1x1 ReLULayer
12	1x1 Convolution2DLayer	60	1x1 ReLULayer	108	1x1 Convolution2DLayer
13	1x1 ReLULayer	61	1x1 Convolution2DLayer	109	1x1 ReLULayer
14	1x1 Convolution2DLayer	62	1x1 ReLULayer	110	1x1 DepthConcatenationLayer
15	1x1 Convolution2DLayer	63	1x1 Convolution2DLayer	111	1x1 MaxPooling2DLayer
16	1x1 ReLULayer	64	1x1 ReLULayer	112	1x1 Convolution2DLayer
17	1x1 ReLULayer	65	1x1 MaxPooling2DLayer	113	1x1 Convolution2DLayer
18	1x1 Convolution2DLayer	66	1x1 Convolution2DLayer	114	1x1 MaxPooling2DLayer
19	1x1 ReLULayer	67	1x1 ReLULayer	115	1x1 Convolution2DLayer
20	1x1 MaxPooling2DLayer	68	1x1 DepthConcatenationLayer	116	1x1 Convolution2DLayer
21	1x1 Convolution2DLayer	69	1x1 Convolution2DLayer	117	1x1 ReLULayer
22	1x1 Convolution2DLayer	70	1x1 Convolution2DLayer	118	1x1 ReLULayer
23	1x1 ReLULayer	71	1x1 ReLULayer	119	1x1 Convolution2DLayer
24	1x1 ReLULayer	72	1x1 MaxPooling2DLayer	120	1x1 ReLULayer
25	1x1 DepthConcatenationLayer	73	1x1 ReLULayer	121	1x1 ReLULayer
26	1x1 Convolution2DLayer	74	1x1 Convolution2DLayer	122	1x1 ReLULayer
27	1x1 ReLULayer	75	1x1 Convolution2DLayer	123	1x1 Convolution2DLayer
28	1x1 Convolution2DLayer	76	1x1 ReLULayer	124	1x1 ReLULayer
29	1x1 ReLULayer	77	1x1 ReLULayer	125	1x1 DepthConcatenationLayer
30	1x1 Convolution2DLayer	78	1x1 Convolution2DLayer	126	1x1 Convolution2DLayer
31	1x1 ReLULayer	79	1x1 ReLULayer	127	1x1 MaxPooling2DLayer
32	1x1 Convolution2DLayer	80	1x1 Convolution2DLayer	128	1x1 Convolution2DLayer
33	1x1 ReLULayer	81	1x1 ReLULayer	129	1x1 Convolution2DLayer
34	1x1 Convolution2DLayer	82	1x1 DepthConcatenationLayer	130	1x1 ReLULayer
35	1x1 ReLULayer	83	1x1 Convolution2DLayer	131	1x1 ReLULayer
36	1x1 MaxPooling2DLayer	84	1x1 Convolution2DLayer	132	1x1 Convolution2DLayer
37	1x1 Convolution2DLayer	85	1x1 Convolution2DLayer	133	1x1 ReLULayer
38	1x1 ReLULayer	86	1x1 ReLULayer	134	1x1 Convolution2DLayer
39	1x1 DepthConcatenationLayer	87	1x1 Convolution2DLayer	135	1x1 ReLULayer
40	1x1 MaxPooling2DLayer	88	1x1 ReLULayer	136	1x1 Convolution2DLayer
41	1x1 Convolution2DLayer	89	1x1 MaxPooling2DLayer	137	1x1 ReLULayer
42	1x1 ReLULayer	90	1x1 Convolution2DLayer	138	1x1 ReLULayer
43	1x1 MaxPooling2DLayer	91	1x1 ReLULayer	139	1x1 DepthConcatenationLayer
44	1x1 Convolution2DLayer	92	1x1 ReLULayer	140	1x1 GlobalAveragePooling2DLayer
45	1x1 ReLULayer	93	1x1 ReLULayer	141	1x1 DropoutLayer
46	1x1 Convolution2DLayer	94	1x1 Convolution2DLayer	142	1x1 FullyConnectedLayer
47	1x1 Convolution2DLayer	95	1x1 ReLULayer	143	1x1 SoftmaxLayer
48	1x1 ReLULayer	96	1x1 DepthConcatenationLayer	144	1x1 ClassificationOutputLayer

Şekil 3.2. GoogleNet Katman Yapısı

3.1.3. DarkNet53

DarkNet53, 183 katmandan oluşan bir yapay sinir ağı modelidir. 256x256x3 görüntü verisini girdi olarak alır. DarkNet modelleri çoğunlukla tek başına kullanılmamaktadır. YOLO ile birlikte adı çokça duyulmaktadır. Bu algoritma çoğunlukla YOLO çalışmalarında özellik haritasının çıkarılmasında kullanılmaktadır. DarkNet53'e ait mimari katman yapısı Şekil 3.3'te gösterilmektedir.



1	1x1 ImageInputLayer	51	1x1 LeakyReLULayer	101	1x1 Convolution2DLayer	151	1x1 BatchNormalizationLayer
2	1x1 Convolution2DLayer	52	1x1 Convolution2DLayer	102	1x1 BatchNormalizationLayer	152	1x1 LeakyReLULayer
3	1x1 BatchNormalizationLayer	53	1x1 BatchNormalizationLayer	103	1x1 LeakyReLULayer	153	1x1 Convolution2DLayer
4	1x1 LeakyReLULayer	54	1x1 LeakyReLULayer	104	1x1 Convolution2DLayer	154	1x1 BatchNormalizationLayer
5	1x1 Convolution2DLayer	55	1x1 AdditionLayer	105	1x1 BatchNormalizationLayer	155	1x1 LeakyReLULayer
6	1x1 BatchNormalizationLayer	56	1x1 Convolution2DLayer	106	1x1 LeakyReLULayer	156	1x1 Convolution2DLayer
7	1x1 LeakyReLULayer	57	1x1 BatchNormalizationLayer	107	1x1 AdditionLayer	157	1x1 BatchNormalizationLayer
8	1x1 Convolution2DLayer	58	1x1 LeakyReLULayer	108	1x1 Convolution2DLayer	158	1x1 LeakyReLULayer
9	1x1 BatchNormalizationLayer	59	1x1 Convolution2DLayer	109	1x1 BatchNormalizationLayer	159	1x1 AdditionLayer
10	1x1 LeakyReLULayer	60	1x1 BatchNormalizationLayer	110	1x1 LeakyReLULayer	160	1x1 Convolution2DLayer
11	1x1 Convolution2DLayer	61	1x1 LeakyReLULayer	111	1x1 Convolution2DLayer	161	1x1 BatchNormalizationLayer
12	1x1 BatchNormalizationLayer	62	1x1 AdditionLayer	112	1x1 BatchNormalizationLayer	162	1x1 LeakyReLULayer
13	1x1 LeakyReLULayer	63	1x1 Convolution2DLayer	113	1x1 LeakyReLULayer	163	1x1 Convolution2DLayer
14	1x1 AdditionLayer	64	1x1 BatchNormalizationLayer	114	1x1 AdditionLayer	164	1x1 BatchNormalizationLayer
15	1x1 Convolution2DLayer	65	1x1 LeakyReLULayer	115	1x1 Convolution2DLayer	165	1x1 LeakyReLULayer
16	1x1 BatchNormalizationLayer	66	1x1 Convolution2DLayer	116	1x1 BatchNormalizationLayer	166	1x1 AdditionLayer
17	1x1 LeakyReLULayer	67	1x1 BatchNormalizationLayer	117	1x1 LeakyReLULayer	167	1x1 Convolution2DLayer
18	1x1 Convolution2DLayer	68	1x1 LeakyReLULayer	118	1x1 Convolution2DLayer	168	1x1 BatchNormalizationLayer
19	1x1 BatchNormalizationLayer	69	1x1 AdditionLayer	119	1x1 BatchNormalizationLayer	169	1x1 LeakyReLULayer
20	1x1 LeakyReLULayer	70	1x1 Convolution2DLayer	120	1x1 LeakyReLULayer	170	1x1 Convolution2DLayer
21	1x1 Convolution2DLayer	71	1x1 BatchNormalizationLayer	121	1x1 AdditionLayer	171	1x1 BatchNormalizationLayer
22	1x1 BatchNormalizationLayer	72	1x1 LeakyReLULayer	122	1x1 Convolution2DLayer	172	1x1 LeakyReLULayer
23	1x1 LeakyReLULayer	73	1x1 Convolution2DLayer	123	1x1 BatchNormalizationLayer	173	1x1 AdditionLayer
24	1x1 AdditionLayer	74	1x1 BatchNormalizationLayer	124	1x1 LeakyReLULayer	174	1x1 Convolution2DLayer
25	1x1 Convolution2DLayer	75	1x1 LeakyReLULayer	125	1x1 Convolution2DLayer	175	1x1 BatchNormalizationLayer
26	1x1 BatchNormalizationLayer	76	1x1 AdditionLayer	126	1x1 BatchNormalizationLayer	176	1x1 LeakyReLULayer
27	1x1 LeakyReLULayer	77	1x1 Convolution2DLayer	127	1x1 LeakyReLULayer	177	1x1 Convolution2DLayer
28	1x1 Convolution2DLayer	78	1x1 BatchNormalizationLayer	128	1x1 AdditionLayer	178	1x1 BatchNormalizationLayer
29	1x1 BatchNormalizationLayer	79	1x1 LeakyReLULayer	129	1x1 Convolution2DLayer	179	1x1 LeakyReLULayer
30	1x1 LeakyReLULayer	80	1x1 Convolution2DLayer	130	1x1 BatchNormalizationLayer	180	1x1 AdditionLayer
31	1x1 AdditionLayer	81	1x1 BatchNormalizationLayer	131	1x1 LeakyReLULayer	181	1x1 GlobalAveragePooling2DL...
32	1x1 Convolution2DLayer	82	1x1 LeakyReLULayer	132	1x1 Convolution2DLayer	182	1x1 Convolution2DLayer
33	1x1 BatchNormalizationLayer	83	1x1 AdditionLayer	133	1x1 BatchNormalizationLayer	183	1x1 SoftmaxLayer
34	1x1 LeakyReLULayer	84	1x1 Convolution2DLayer	134	1x1 LeakyReLULayer	184	1x1 ClassificationOutputLayer
35	1x1 Convolution2DLayer	85	1x1 BatchNormalizationLayer	135	1x1 AdditionLayer		
36	1x1 BatchNormalizationLayer	86	1x1 LeakyReLULayer	136	1x1 Convolution2DLayer		
37	1x1 LeakyReLULayer	87	1x1 Convolution2DLayer	137	1x1 BatchNormalizationLayer		
38	1x1 Convolution2DLayer	88	1x1 BatchNormalizationLayer	138	1x1 LeakyReLULayer		
39	1x1 BatchNormalizationLayer	89	1x1 LeakyReLULayer	139	1x1 Convolution2DLayer		
40	1x1 LeakyReLULayer	90	1x1 AdditionLayer	140	1x1 BatchNormalizationLayer		
41	1x1 AdditionLayer	91	1x1 Convolution2DLayer	141	1x1 LeakyReLULayer		
42	1x1 Convolution2DLayer	92	1x1 BatchNormalizationLayer	142	1x1 AdditionLayer		
43	1x1 BatchNormalizationLayer	93	1x1 LeakyReLULayer	143	1x1 Convolution2DLayer		
44	1x1 LeakyReLULayer	94	1x1 Convolution2DLayer	144	1x1 BatchNormalizationLayer		
45	1x1 Convolution2DLayer	95	1x1 BatchNormalizationLayer	145	1x1 LeakyReLULayer		
46	1x1 BatchNormalizationLayer	96	1x1 LeakyReLULayer	146	1x1 Convolution2DLayer		
47	1x1 LeakyReLULayer	97	1x1 Convolution2DLayer	147	1x1 BatchNormalizationLayer		
48	1x1 AdditionLayer	98	1x1 BatchNormalizationLayer	148	1x1 LeakyReLULayer		
49	1x1 Convolution2DLayer	99	1x1 LeakyReLULayer	149	1x1 AdditionLayer		
50	1x1 BatchNormalizationLayer	100	1x1 AdditionLayer	150	1x1 Convolution2DLayer		

Şekil 3.3. DarkNet53 Katman Yapısı

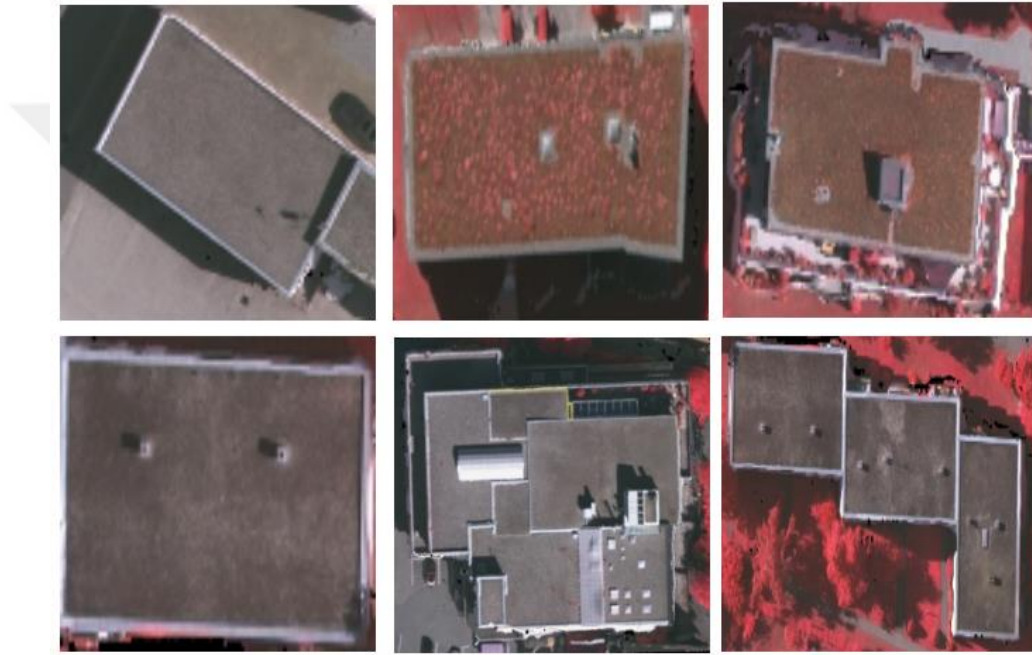
Eğitim safhasında kullanılan CNN mimarisinin parametre değerleri, eğitim süresini ve doğruluk değerini doğrudan etkiler. Eğitimde en çok kullanılan parametreler; "learning rate", "max epoch" ve "mini batch size" olarak görüntülenir. Bu parametrelere en ideal değerleri vermek kısa eğitim süresinde başarılı sonuçlara ulaşmamızı sağlar. Bu çalışma kapsamında karşılaştırma için üç farklı CNN mimarisi seçileceğinden, her üç mimarinin eğitim parametreleri aynı değerlere sahiptir. Eğitim için seçilen parametreler ve değerleri Şekil 3.4'te gösterilmiştir.

SOLVER	
Solver	sgdm
InitialLearnRate	1e-05
BASIC	
ValidationFrequency	50
MaxEpochs	20
MiniBatchSize	32
ExecutionEnvironment	auto
SEQUENCE	
SequenceLength	longest
SequencePaddingValue	0
SequencePaddingDirection	right
ADVANCED	
L2Regularization	0.0001

Şekil 3.4. Eğitim Parametre Değerleri

Veri setleri, eğitim aşamasının en temel bileşenlerinden biridir. Veri seti, eğitilecek nesnelere ait bir dizi görüntü kümeleridir. Veri seti ile ilgili iki önemli konu vardır. Bunlar veri setine ait görüntü sayısı ve görüntü kalitesidir. Eğitim aşamasında, model, veri kümesindeki nesnelerin görüntüleri kullanılarak eğitilir. Bu aşama beynin çalışması ile aynıdır. Bu eğitimde nesnelere ait öznitelikler çıkarılır ve model nesneleri öğrenir. Modelin eğitilmesinde kullanılan görüntülerin çeşitliliği ve sayısı arttıkça modelin eğitim düzeyi de artmakta ancak belirli bir düzeyden sonra görüntü sayısındaki artış eğitim kalitesini düşürmeye başlamaktadır. Bu nedenle başarılı bir eğitim için veri setine ait görüntü sayısı en uygun düzeyde seçilmelidir. Diğer bir kriter ise eğitim görüntülerinin uzamsal

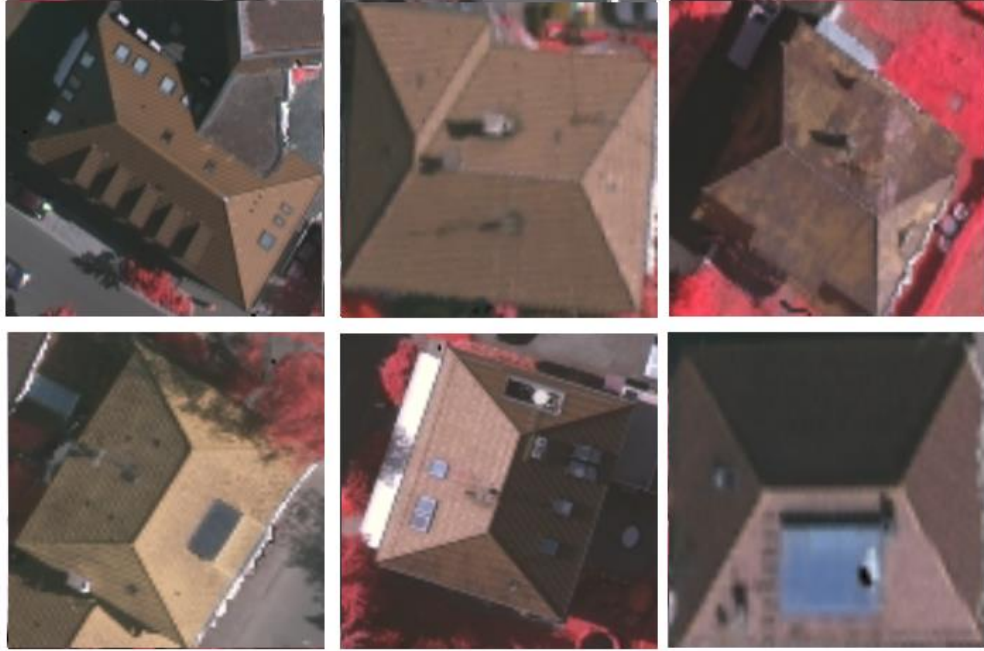
özünürlüğüdür. Görüntüdeki nesneler ne kadar kolay anlaşılırsa ve nesnelerin detayları ne kadar net olursa modelin eğitimi o kadar başarılı olacaktır. Bu çalışmada, veri setinin görüntüleri ve test görüntüleri özellikle orta düzey bir çözünürlükte seçilmiştir. Bunun amacı, orta seviye çözünürlüklü görüntüler üzerinde CNN mimarisinin başarı düzeyini ölçmektir. Veri setinde düz, üçgen ve sarkık çatı sınıflarına ait sınıf başına yaklaşık 4000 görüntü bulunmaktadır. Veri setindeki üç adet sınıfa ait görüntülerden bazıları aşağıda görülmektedir (Şekil 3.5 – 3.7).



Şekil 3.5. Düz Çatı Tipine Ait Veri Seti Örneği

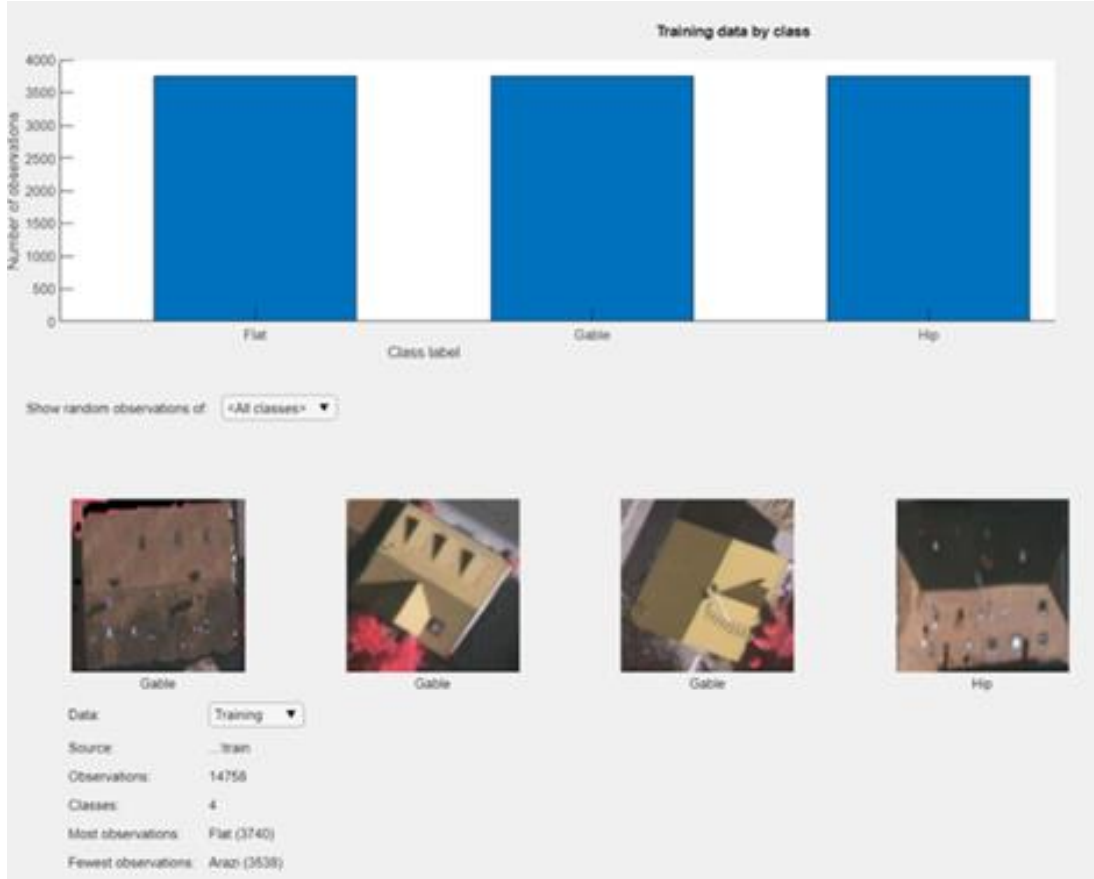


Şekil 3.6. Üçgen Çatı Tipine Ait Veri Seti Örneği



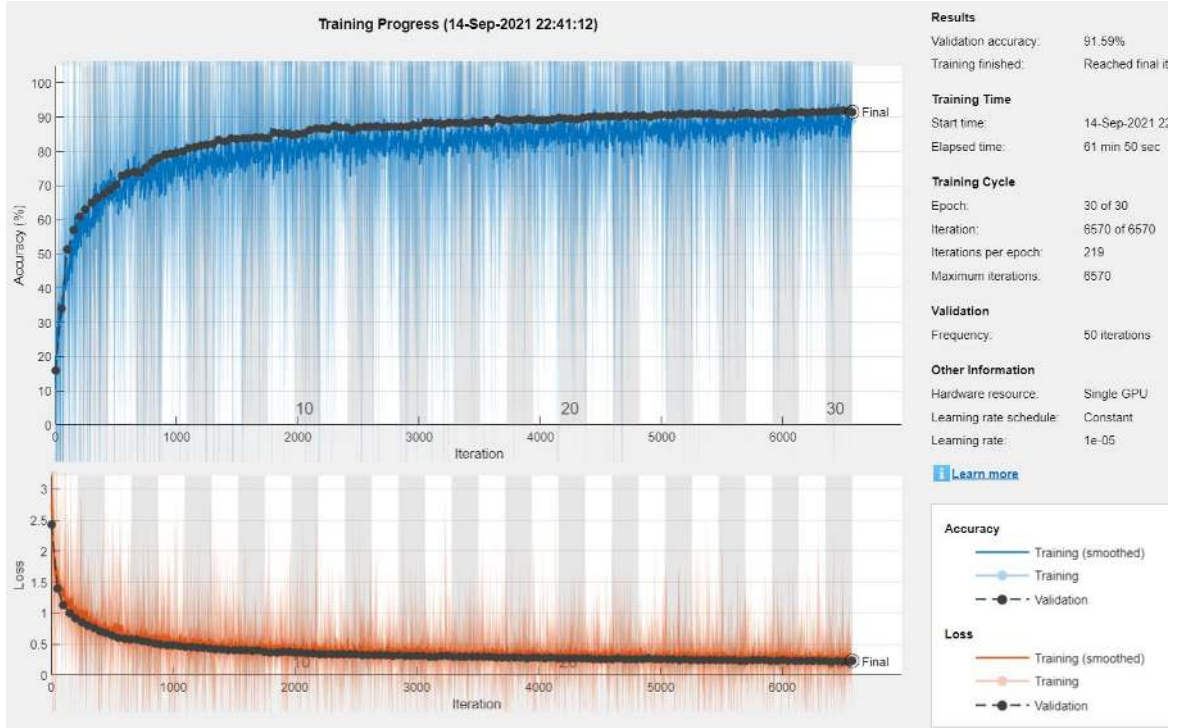
Şekil 3.7. Sarkık Çatı Tipine Ait Veri Seti Örneği

Bu görüntüler sadece çatı alanını içeren düşük çözünürlüklü verilerden oluşmaktadır. Veri setine ait örnek görseller ve çatı sınıfına ait eğitim örnek sayıları aşağıda gösterilmiştir (Şekil 3.8).

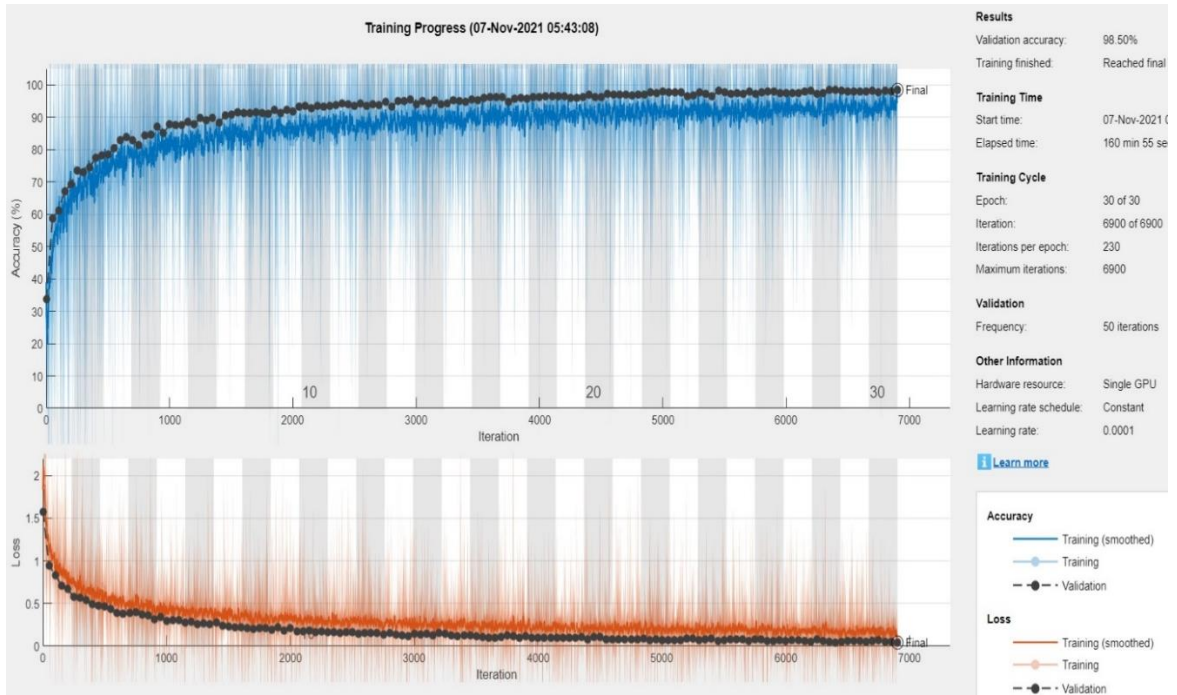


Şekil 3.8. Sınıflara Ait Eğitim Veri Seti Dağılımları

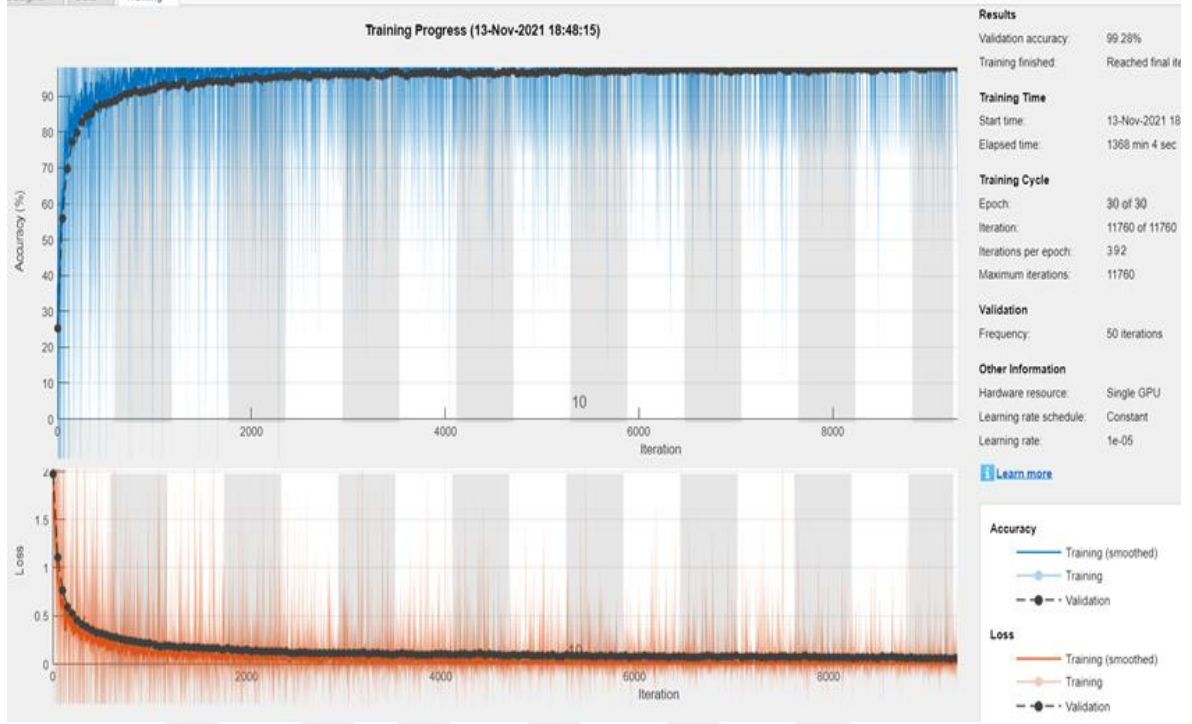
Eğitim için seçilen veri seti, Matlab'a ait "Derin Öğrenme Araç Kutusu" kullanılarak yukarıdaki parametreler seçilerek eğitilmiştir. Veri setinin %85'i eğitim ve %15'i test için kullanılmıştır. Eğitim aşamasındaki başarı oranları ve çalışma süreleri, mimarilerin bina çatı tiplerini sınıflandırmadaki başarı oranını yansıtmaktadır. GoogleNet, AlexNet ve DarkNet53 mimarileri kullanılarak eğitilen modellerin başarı oranları sırasıyla % 98.50, % 91.59 ve % 99.28 şeklinde elde edilmiştir. Eğitime ait sonuçlar ve eğitim sürelerini gösteren grafikler sırasıyla aşağıdaki şekillerde gösterilmektedir (Şekil 3.9 – Şekil 3.11).



Şekil 3.9. AlexNet Eğitim Sonuçları



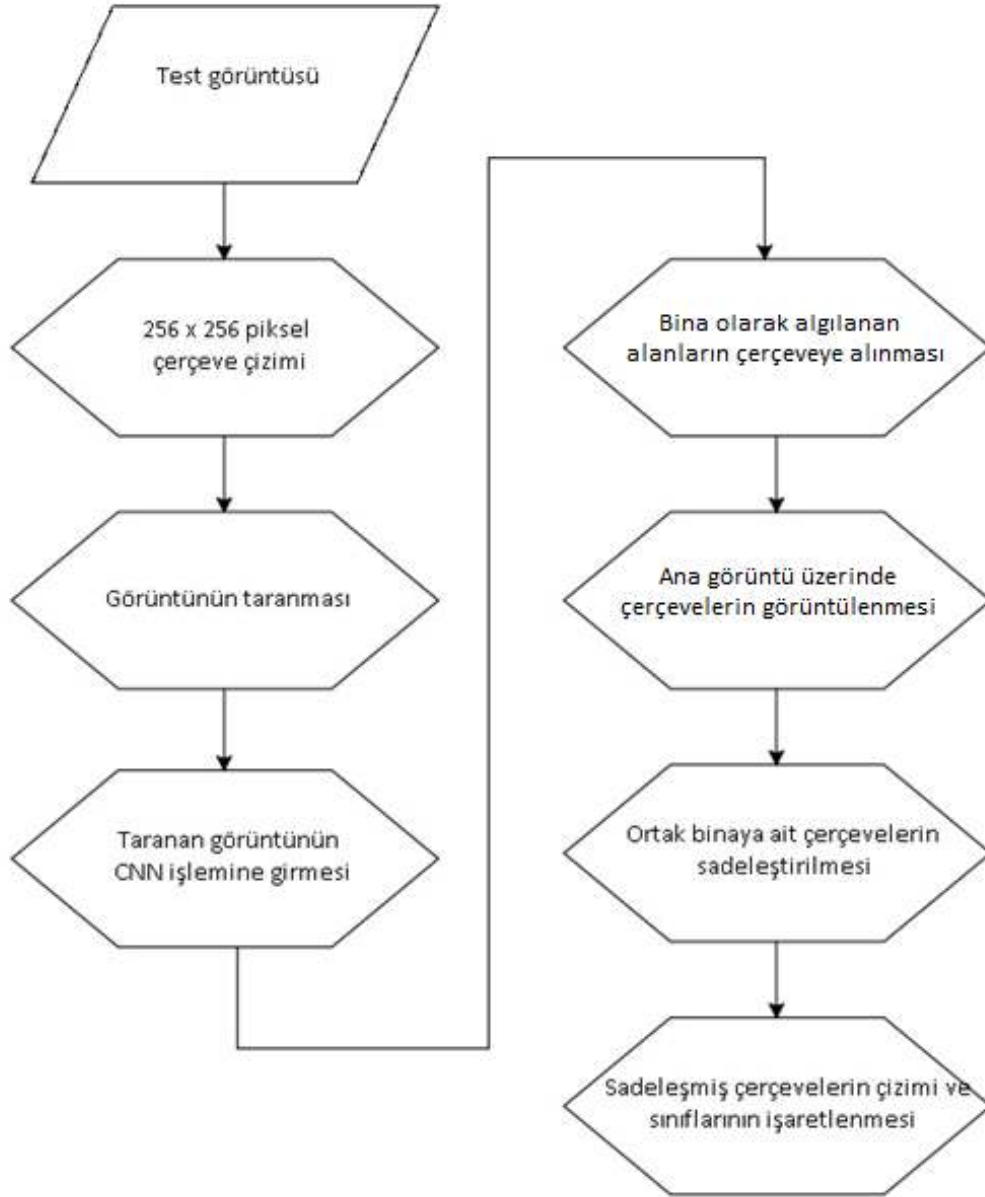
Şekil 3.10. GoogleNet Eğitim Sonuçları



Şekil 3.11. DarkNet53 Eğitim Sonuçları

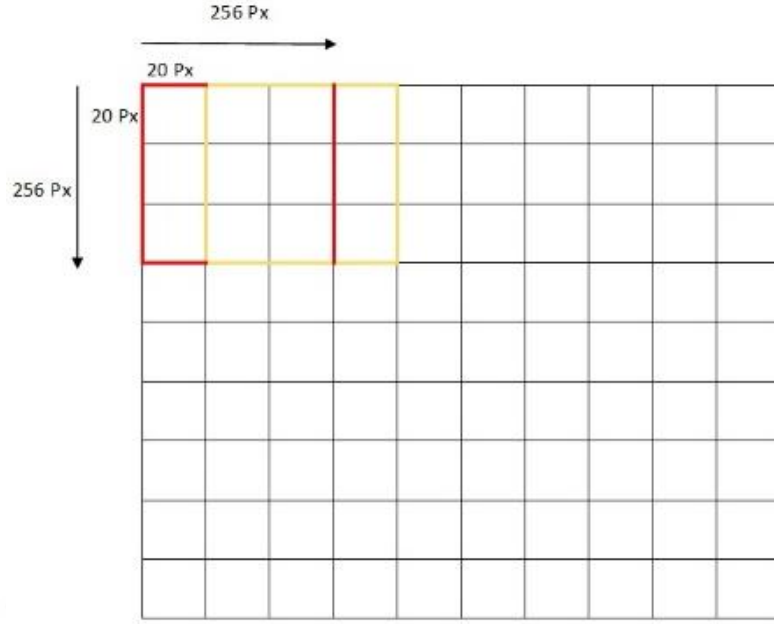
3.2. Arayüz Tasarımı

Arayüz çalışmaları için Matlab ortamı seçilmiştir. Burada tasarlanan arayüz kullanılarak; girdi, tanıma, sınıflandırma ve çıktı üretimi gibi işlemler tek bir ekranda gerçekleştirilmektedir. Girdi olarak bir test görüntüsü, binanın konumunu ve sınıfını gösteren bir çıktı verir. Arayüz tasarımında test verileri parçalara ayrılarak eğitilen modele dâhil edilir ve sınıflandırılır. Böylece bina olarak algılanan görüntü etiketlenir ve çıktı verisi hazır olur. Tasarımın yapısı birden fazla aşamadan meydana gelmektedir. Bunlar; “Tarama”, “Tespit”, “Sınıflandırma” ve “Birleştirme” aşamalarıdır. Bu aşamalardan “Tarama”, “Tespit” ve “Sınıflandırma” adımları birbiri ardı sıra gelen ardışık işlemlerdir. Her görüntü ayrı ayrı taranır, algılanır ve sınıflandırılır. Ana görüntünün bütün alanları tarandığında, son olarak birleştirme işlemi uygulanır. Arayüz tasarımının akış şeması Şekil 3.12’de gösterilmiştir.



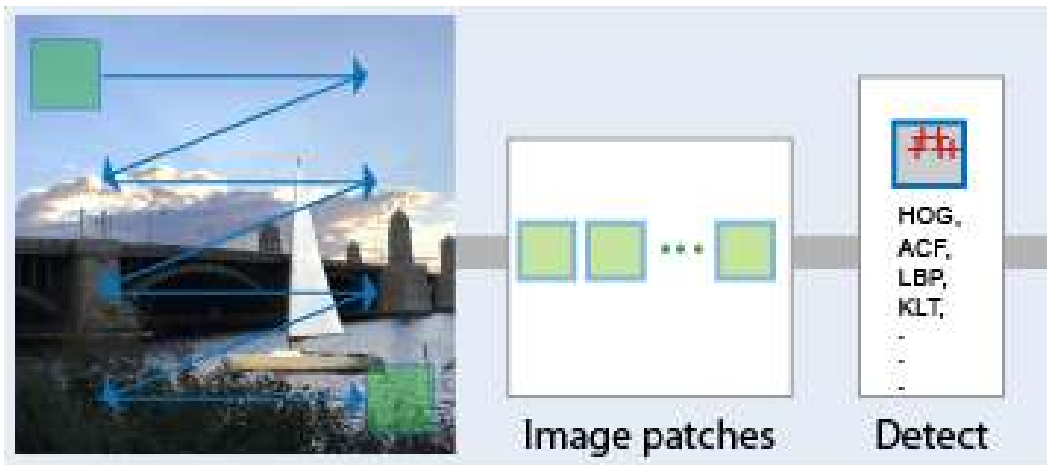
Şekil 3.12. İşlem Akış Diyagramı

Bu çalışmada, bir binayı içeren görüntü verilerinin yanı sıra tek görüntü içerisinde birden fazla binaya sahip görüntülerin analiz edilebilmesi amacıyla bir tarama algoritması oluşturulmuştur. Kısaca, algoritma test görüntüsünü giriş değeri olarak alır ve 256 x 256 piksel karelerde tarar. Tarama algoritması, çerçeveyi sırasıyla soldan sağa ve yukarıdan aşağıya hareket ettirme prensibine dayanmaktadır. Çerçeve, 20 piksellik bir adım değeriyle kaydırılır ve tüm görüntüyü taramasına izin verir. Tarama hareket yönünü gösteren görsel Şekil 3.13'te verilmiştir.



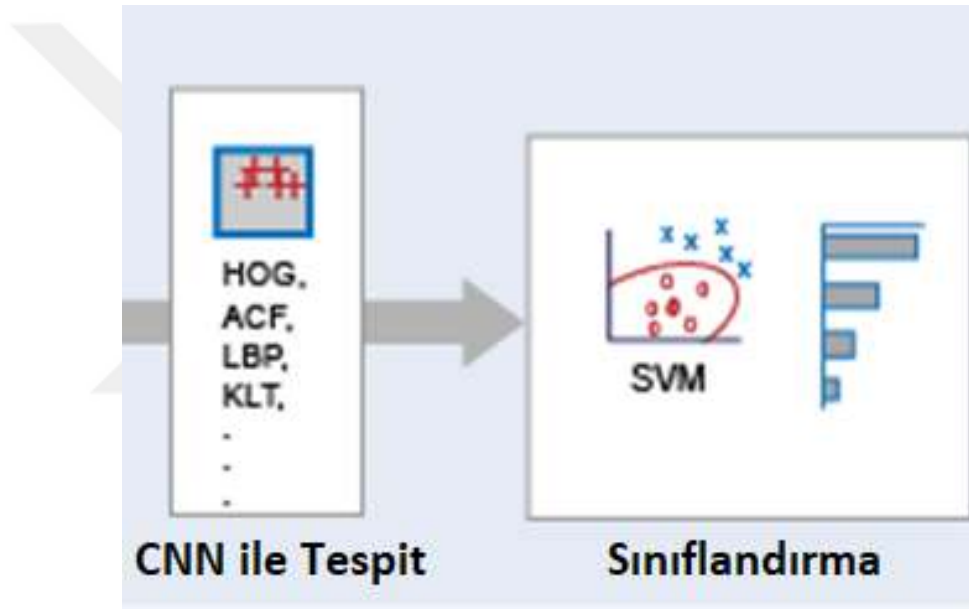
Şekil 3.13. Tarama Hareket Yönü

Tespit işlemi ile tarama işlem birleştirilir. Her tarama, görüntünün ilgili bölümünün bir bina alanı olup olmadığını belirlemek için önceden eğitilmiş bir CNN mimarisine sokulur. Bina olduğu kabul edilen çerçeveye ait konum bilgileri işaretlenir ve çerçevenin hareketini görüntü üzerinde ve sonraki görüntü parçası boyunca devam ettirir. Bina olarak algılanan ve işaretlenen görüntünün koordinatları farklı bir alana kaydedilir ve binaların koordinat değerleri bir liste halinde saklanır. Bu işlemi özetleyen örnek bir görüntü Şekil 3.14'te gösterilmiştir.



Şekil 3.14. Görüntü Tarama Algoritması

Sınıflandırma aşaması, bina alanı tespitinden elde edilen sonuçlara göre çalışır. Seçilen görüntülerin çatı tipi tespiti için eğitilmiş bir modele girdi olarak alınmasıyla önceden tanımlanmış üç adet sınıf için skor değerleri belirlenir. Bu değerler 0 ile 1 arasındaki olasılık değerleridir. Üç sınıf için puan değerlerinin toplamı 1'dir ve tahmin edilen değerler buna göre dağıtılır. En yüksek puana sahip sınıfı seçerek bir binanın çatı sınıfını tahmin yapılmıştır.



Şekil 3.15. Bina Bölge Tespiti ve Çatı Sınıflandırılmasına Ait İşlemler

Tüm test görüntülerini 256 x 256 kare ve 20 piksel kaydırma değeri ile taradıktan sonra son aşama olan "birleştirme" aşaması başlar. Birleştirme aşamasının görevi, bir önceki aşamada bina alanı olarak tanımlanan görüntüleri sadeleştirmektir. Tarama sırasında 256 x 256 boyutlarında seçilen tarama çerçeveleri, binaları tek bir çerçeveye dâhil etmeyebilir. Bunun başlıca sebebi, görüntü içerisindeki bina alanının çerçeve boyutundan daha büyük olmasıdır. Diğer bir olasılık ise binanın kaydırma piksel boyutu olan 20 pikselden daha büyük olmasıdır. Örneğin, bir bina kaydırma değeri olan 20 piksel ile çerçeve boyutu olan 256 x 256 piksel alanı arasında bir büyüklükte olursa tek bir görüntü

içerisinde yer alır. Ancak bu değer aralığı dışında bir değer olursa birden fazla tarama çerçevesi içerisinde görüntülenebilir. Böylelikle aynı binaya ait çerçeveler birleştirilerek sadeleştirme yapılmıştır. Bu nedenle, bir bina birden çok kez görüntüleniyorsa, sadeleştirme gereklidir. Aynı ya da farklı binalara ait olma durumları sınıfları ve aralarındaki mesafeler için seçilen eşik değer ile tespit edilmiştir. Yeni koordinat değerleri oluşturmak için aynı olduğu tahmin edilen görüntüler birleştirilmiştir. Böylelikle aynı binaya ait birden fazla oluşan çerçeveler sadeleştirilmiştir. Bu koordinat değerleri daha sonra ana test görüntüsü üzerinde sınıfları ile etiketlenir. Örnek bir tespit ve sadeleştirme durumunu örnekleyen görseller Şekil 3.16’da verilmiştir.



Şekil 3.16. Tarama Algoritması ile Test Görüntüsünde Bina Bölge Tespiti (sol) ve Sadeleştirme sonucu (sağ)

Bu sayede klasik yöntem olan bounding box ile nesnenin dörtgen içine alınması yerine asimetrik şekillere sahip binalara ait asimetrik çerçeveler oluşturulabilmiştir. Bu yöntem aynı zamanda kendi bounding box algoritmamız ile asimetrik bir bounding box meydana getirmemizi sağlamıştır. Şekil.3.16’te sağ taraftaki görüntü içerisinde sol üst kısımda görünmekte olan asimetrik binaya ait çerçeve, yöntemin başarısını göstermektedir. Bu başarıyı açıkça gösteren diğer bir görüntü örneği ise aşağıdaki şekilde gösterilmektedir (Şekil 3.17).



Şekil 3.17. DarkNet53 Sınıflandırma Örneği

DarkNet53 algoritması ile teste sokulan örnek görüntülerden birinde asimetrik boyutu çok yüksek olan bu bina için oluşturulan bounding box görülmektedir.

3.2.1. Test

Test aşaması, çalışmanın deneysel sonuçlara ait elde ettiğimiz çıktı aşamasıdır. Başka bir ifadeyle, tasarlanan arayüz kullanılarak yapılan eğitim modelini test ettiğimiz ve çalışmamızın bina bölge tespiti ve çatı tipi sınıflandırma doğruluğunu ortaya koyduğumuz aşama olarak tanımlanabilir. Test aşaması, tasarladığımız entegre arayüze test görüntüsünü girdi alarak test görüntüsündeki binaların konumlarını belirler ve çatı tiplerini çıkarır. Arayüzü ve eğitilmiş CNN modelini test etmek için 10 adet farklı özelliklere sahip görüntüden oluşan bir veri seti oluşturulmuştur. Bu veri seti, Google Earth uygulamasında manuel yöntemlerle üretilmiştir. Geniş alanlara ait Google Earth verileri kaydedilmiştir. Matlab ile gerçekleştirilen algoritma ile bu görüntüler parçalara ayrılarak tekrardan kaydedilmiştir. Manuel yöntemle elde edilen bu görüntüler incelenerek içlerinden

10 adet test görüntüsü seçilmiştir. Test görüntüleri seçilirken bazı kriterler göz önünde bulundurulmuştur. Bu kriterler çatı rengi, binaların asimetrik şekilleri ve binaların aralıklı olmasıdır. Düz çatı tipinin beton yapısı sebebiyle gri renkte olmasından dolayı özellikle gri renkli üçgen çatı tipinde bina görüntüsü aranmıştır. Bunun amacı aynı renkte farklı sınıfa ait çatı tipini doğru şekilde sınıflandırabileceğini sorgulamaktır. Diğer seçim kriteri olan asimetrik binaları seçme sebebi, görüntü içerisindeki asimetrik binanın çatı sınıfını doğru tahmin edebildiğini ve çerçevenin asimetrik şekilde binayı tamamen alabildiğini test etmektir. Binaların arasının aralıklı olması binalara ait oluşturulacak çerçevelerin sade ve belirgin şekilde görüntülenebilmesini sağlayacaktır. Çatılara ait renk çeşitliliğinin bulunduğu test verisi örneklerinden biri Şekil 3.18’ de gösterilmiştir.



Şekil 3.18. Çatı Renk Çeşitliliği İçeren Test Görüntü Örneği

Resimlerin çekim yükseklik değerleri farklı seçilmiştir. Bu, binaların ve diğer nesnelerin görüntü üzerinde farklı boyutlarda olduğu anlamına gelir. Çalışmamızda görüntü çözünürlük kalitesi, binanın geometrik şekli ve çatı rengi, bina boyutu gibi farklı etmenlerin seçtiğimiz CNN mimarilerindeki başarıya etkileri gözlemlenmiştir.

Test için görüntüler oluşturulurken seçilen binaların farklı şekillerde, farklı çatı renklerinde ve boyutlarında olmasına dikkat edilmiş ve çözünürlüğü bir miktar azaltılmıştır. Şekil 3.19, örnek bir test görüntüsünü göstermektedir.



Şekil 3.19. Test Görüntü Örneği

4. DENEYSEL SONUÇLAR

Üç farklı CNN mimarisi kullanılarak, bina alanı tespiti ve çatı sınıflandırmasındaki doğruluk değerlerinin tespiti için çoklu bina içeren 10 adet test görüntüsü işleme sokulmuştur. Bu çalışmada iki ana sonuç çıkarılması hedeflenmiştir. Bunlar, seçilen CNN mimarilerinin bina bölge tespiti ve çatı sınıflandırılmasındaki başarılarıdır. Bina bölge tespiti için elde edilen başarı değerleri DarkNet53, GoogleNet ve AlexNet mimarileri için sırasıyla %38.48, %87.18, %92.30 olarak elde edilmiştir. Diğer yandan, binaların çatı tiplerinin sınıflandırılmasında Top-1 ve Top-2 değerleri olmak üzere iki farklı sonuç hesaplanmıştır. ‘Top-1’ değeri, belirlediğimiz farklı üç adet çatı sınıfı için elde ettiğimiz en yüksek skor değerini ifade etmektedir. ‘Top-2’ ölçütü ise tespiti yapılan çatı sınıfının için en olası iki sınıftan bir tanesine ait olma durumudur. Sınıflandırmanın hesaplanmasında çok sınıflı hata matrisi kullanılmıştır. Bu hesaplama için öncelikle sınıflara ait hata matris tablosu oluşturulmuştur. Üç sınıfa ait hata matris tablosu aşağıda gösterilmektedir (Tablo 3).

Tablo 3. AlexNet Hata Matrisi

T O P 1				
AlexNet		Gerçek		
		Düz (Flat)	Üçgen (Gable)	Sarkık (Hip)
Tahmin	Flat (Düz)	0	0	3
	Gable (Üçgen)	0	14	9
	Hip (Sarkık)	0	6	12

T O P 2				
AlexNet		Gerçek		
		Düz (Flat)	Üçgen (Gable)	Sarkık (Hip)
Tahmin	Flat (Düz)	0	1	1
	Gable (Üçgen)	0	16	1
	Hip (Sarkık)	0	4	9

Tablo 4. GoogleNet Hata Matrisi

T O P 1				
GoogleNet		Gerçek		
		Düz (Flat)	Üçgen (Gable)	Sarkık (Hip)
Tahmin	Flat (Düz)	2	8	2
	Gable (Üçgen)	0	13	1
	Hip (Sarkık)	0	11	4

T O P 2				
GoogleNet		Gerçek		
		Düz (Flat)	Üçgen (Gable)	Sarkık (Hip)
Tahmin	Flat (Düz)	2	3	1
	Gable (Üçgen)	0	11	0
	Hip (Sarkık)	0	4	6

Tablo 5. DarkNet53 Hata Matrisi

T O P 1				
DarkNet53		Gerçek		
		Düz (Flat)	Üçgen (Gable)	Sarkık (Hip)
Tahmin	Flat (Düz)	0	6	4
	Gable (Üçgen)	0	1	1
	Hip (Sarkık)	0	1	3

T O P 2				
DarkNet53		Gerçek		
		Düz (Flat)	Üçgen (Gable)	Sarkık (Hip)
Tahmin	Flat (Düz)	0	4	4
	Gable (Üçgen)	0	1	0
	Hip (Sarkık)	0	1	3

Sınıflandırmanın değerlendirilmesinde iki farklı metrik kullanılmıştır. Bunlar, F1 skoru ve sınıflara ait hata matrisidir. F1 skoru her sınıfı tek tek dikkate almaz, metrikleri global olarak hesaplar. Her üç sınıfa ait F1 skoru değerleri aşağıdaki tablolarda sırasıyla gösterilmektedir.

Tablo 6. Mikro F1 Skor Tablosu

GoogleNet	TOP-1	TOP-2
Toplam TP	19	19
Toplam FP	22	8
Toplam FN	22	8
Tahmin	0,46	0,70
Hassasiyet	0,46	0,70
F1 Skoru	0,46	0,70

AlexNet	TOP-1	TOP-2
Toplam TP	26	25
Toplam FP	18	7
Toplam FN	18	7
Tahmin	0,59	0,78
Hassasiyet	0,59	0,78
F1 Skoru	0,59	0,78

DarkNet53	TOP-1	TOP-2
Toplam TP	4	4
Toplam FP	12	9
Toplam FN	12	9
Tahmin	0,25	0,31
Hassasiyet	0,25	0,31
F1 Skoru	0,25	0,31

Test aşamasında tespit edilen binalara ait çatı sınıflandırılması yapılırken, bazı binaların çatıları her üç sınıfa da ait olarak görülmüştür. Bu durum, skor tablosu içerisindeki üç mimariye ait Top-1 ve Top-2 sonuçlarında bina sayılarının birbirinden farklı olmasına sebep teşkil etmiştir. Hesaplama çıkan sonuçlara göre Tahmin, Hassasiyet ve F1 skorları birbirlerine eşit olarak çıkmıştır. Buradan anlaşılabileceği gibi F1 skoru bize aynı zamanda doğruluk (accuracy) değerini vermektedir. Bu sonuçlara göre Top-1 genel doğruluk oranları AlexNet, GoogleNet ve DarkNet mimarileri için sırasıyla %59, %46 ve %25 olarak

hesaplanmıştır. Top-2 doğruluk oranları ise aynı sırayla %78, %70 ve %31 olarak elde edilmiştir.

Yapılan diğer bir hesaplama ise hata matrisi tablosuna göre her sınıfa ait Tahmin, Hassasiyet, F1 Skoru ve Doğruluk değerlerinin bulunmasıdır. Aynı şekilde bu hesaplamalar Top-1 ve Top-2 için ayrı ayrı hesaplanmıştır. Elde ettiğimiz ‘Top-1’ ve ‘Top-2’ skorları Tablo 6’da sunulmuştur. Bu hesaplamada Top-1 ve Top-2 için her sınıfa ait doğruluk tahmin değerleri ayrıntılı olarak gösterilmektedir.

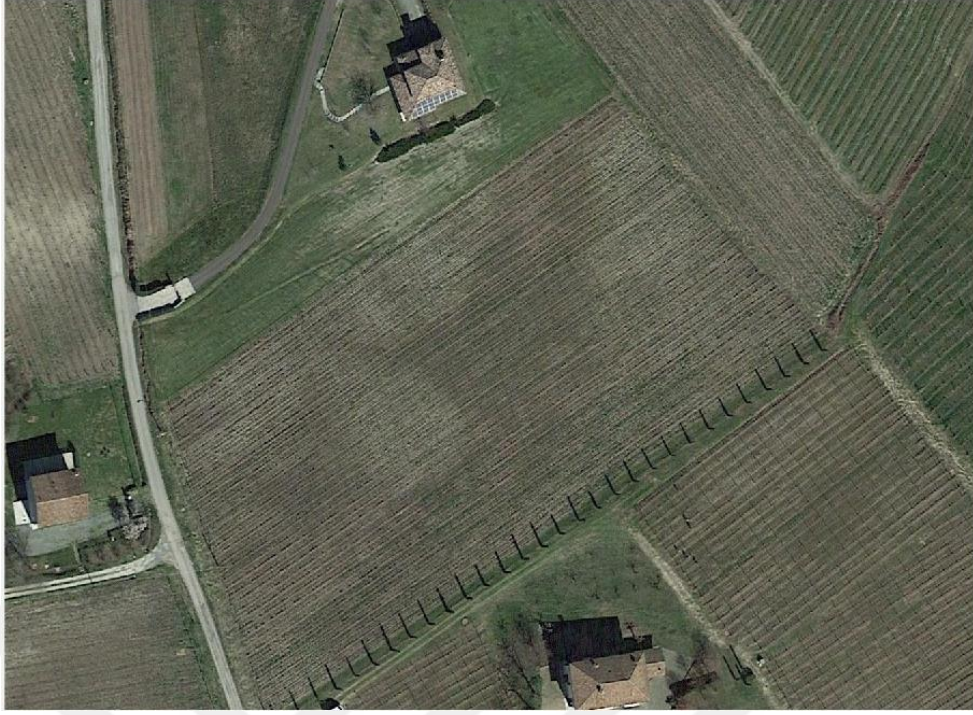
Tablo 7. Top-1 ve Top-2 Değerlerine Göre Sınıflandırma Değerleri

GoogleNet		Toplam TP	Toplam TN	Toplam FP	Toplam FN	Tahmin	Doğruluk	Hassasiyet	F1 Skoru
TOP-1	Flat (Düz)	2	29	10	0	0,17	0,76	1	0,29
	Gable (üçgen)	13	8	1	19	0,93	0,51	0,41	0,57
	Hip (Sarkık)	4	23	11	3	0,27	0,66	0,57	0,36
TOP-2	Flat (Düz)	2	21	4	0	0,33	0,85	1	0,5
	Gable (üçgen)	11	9	0	7	1	0,74	0,61	0,76
	Hip (Sarkık)	6	16	4	1	0,6	0,81	0,86	0,71

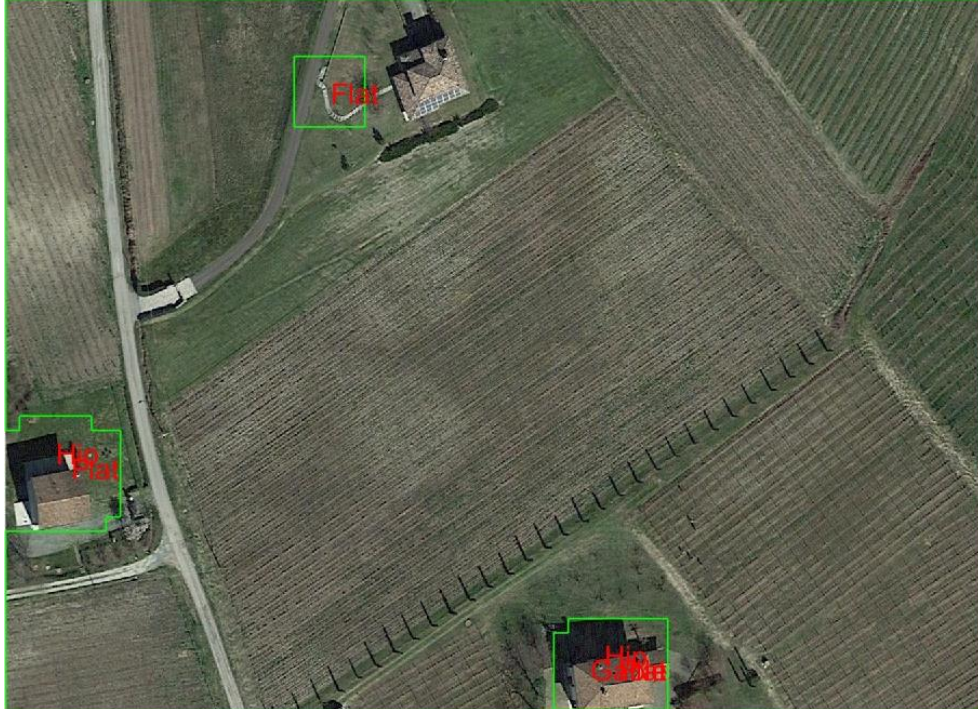
AlexNet		Toplam TP	Toplam TN	Toplam FP	Toplam FN	Tahmin	Doğruluk	Hassasiyet	F1 Skoru
TOP-1	Flat (Düz)	0	41	3	0	0	0,93	0	0
	Gable (üçgen)	14	15	9	6	0,61	0,66	0,7	0,65
	Hip (Sarkık)	12	14	6	12	0,67	0,59	0,5	0,57
TOP-2	Flat (Düz)	0	30	2	0	0	0,94	0	0
	Gable (üçgen)	16	10	1	5	0,94	0,81	0,76	0,84
	Hip (Sarkık)	9	17	4	2	0,69	0,81	0,82	0,75

DarkNet53		Toplam TP	Toplam TN	Toplam FP	Toplam FN	Tahmin	Doğruluk	Hassasiyet	F1 Skoru
TOP-1	Flat (Düz)	0	6	10	0	0	0,38	0	0
	Gable (üçgen)	1	7	1	7	0,5	0,50	0,13	0,20
	Hip (Sarkık)	3	7	1	5	0,75	0,63	0,38	0,50
TOP-2	Flat (Düz)	0	5	8	0	0	0,38	0	0,00
	Gable (üçgen)	1	7	0	5	1	0,62	0,17	0,29
	Hip (Sarkık)	3	5	1	4	0,75	0,62	0,43	0,55

DarkNet53, GoogleNet ve AlexNet mimarileri ile işleme alınan test görüntülerinden bir tanesine ait sonuçlar Şekil 4.1, Şekil 4.2 ve Şekil 4.3’te gösterilmiştir.



Şekil 4.1. DarkNet53 Test Sonucu



Şekil 4.2. GoogleNet Test Sonucu



Şekil 4.3. AlexNet Test Sonucu

Gösterilen test sonuçları incelendiğinde DarkNet53 mimarisi ile yapılan testte örnek görüntü içerisindeki üç adet bina bulunamamıştır. Binaların tespit edilememesi sebebiyle sınıflandırma da gerçekleştirilememiştir. Şekil 4.2 incelendiğinde ise GoogleNet mimarisi kullanılarak test edilen görüntü içerisindeki üç adet bina tespit edilebilmiş ancak bir tanesinin yeri tam olarak bulunamamıştır. Tespit edilen üç binaya ait sınıflandırma sonucu başarısız olmuştur. Şekil 4.3 incelendiğinde ise AlexNet mimarisi kullanılarak test edilen görüntünün diğer iki mimariye göre daha iyi sonuç verdiği görülmektedir. AlexNet mimarisi ile yapılan testte görüntü içerisindeki üç adet binanın tamamı tespit edilebilmiştir. Tespit edilen üç binanın ikisinin çatı tipi sınıflandırması doğru yapılırken, üçüncü binaya ait çatı sınıflandırmasında Üçgen ve Sarkık çatı skoru birbirine yakın değerlerde çıktı vermiştir. Bu görüntü içerisinde Top-1 sonucu ele alındığında üç görüntünün ikisi doğru sınıflandırılırken, Top-2 sonucu ele alındığında üç sonucun tamamında doğru sınıflandırma yapıldığı kabul edilmiştir.

5. SONUÇ VE DEĞERLENDİRME

Bu çalışma ile iki ana sonuca varmak amaçlamıştır. Bunlar, bina bölge tespiti ve çatı tipi sınıflandırması için seçilen CNN mimarilerinin başarılarıdır. GoogleNet, AlexNet ve DarkNet53 mimarileri için bina bölgesi tespiti doğruluk performans değerleri sırasıyla %87.18, %92.30 ve %38.48'dir. Bina çatı tiplerinin sınıflandırılmasında ise Top-1 ve Top-2 olmak üzere farklı iki sonuç değeri hesaplanmıştır. Çatı sınıfları için alınan en yüksek skor "Top-1" değeri olarak kabul edilmiştir. "Top-2", tespit edilen çatı sınıfının üç sınıf içinden en yüksek iki sınıf skorunu alarak, gerçek sınıfın bu seçilen iki sınıf içinde olma durumunu göstermektedir. Çalışma sonucunda bina alanı tespiti ve çatı sınıflandırması için üç farklı yapay sinir ağı modelinin sonuçları gözlemlenmiştir. DarkNet53'ün bina bölge tespiti ve sınıflandırmadaki doğruluk oranlarının sırasıyla %38,48 ve %25 olduğu gözlemlenmiştir. Bu sonuçlara dayanarak DarkNet53 mimarisinin binaların bölge tespiti ve çatı sınıflandırması için uygun olmayan bir mimariye sahip olduğu sonucuna varılmıştır. Elde edilen sonuçlar değerlendirildiğinde, bina bölge tespitinde en uygun mimari %92.30 ile AlexNet olmuştur. GoogleNet mimarisi de aldığı yüksek doğruluk oranıyla bina bölge tespiti konusunda kullanılabilir bir mimari olduğunu göstermiştir. CNN mimarilerinden AlexNet ve GoogleNet mimarilerinin bölge tespiti yapabilmek için elde ettiği yüksek doğruluk, görüntü çözünürlüğünün tespit konusunda çok az etkisi olduğu sonucunu çıkarmamızı sağlamıştır.

Üç mimarinin de çatı tipi sınıflandırmasında düşük performans gösterdiği ancak %59,09 ile en yüksek doğruluğa AlexNet mimarisinin sahip olduğu görülmüştür. Eğitim ve test için kullanılan veri setlerinin orta düzey çözünürlükte olması, sınıflandırmada belirleyici bir etken olduğunu göstermiştir. Özellikle düz çatı sınıfının betonarme yapıda olması ve gri renk tonuna sahip olması diğer gri çatı sınıfları ile karıştırılmasına neden olmuş ve eğitim verilerinin çözünürlüğünün düşük olması nedeniyle ayırt edici faktörlerin öğrenmeye dâhil olamamasına sebep olmuştur.

Bu tez çalışmasında, birçok çalışmadaki doğruluk oranını arttıran etkenlerin kullanılması ile yüksek doğruluk oranına sahip sonuçlar elde etmek yerine doğruluk oranını etkileyen olumsuz şartların etkileri gözlemlenmiştir. Çalışma sonucunda elde edilen sonuçlara göre, bina bölge tespitinde çözünürlüğün etkisinin belirli bir seviyede göz ardı edilebileceği ancak binalara ait çatı sınıflandırılmasında göz ardı edilemeyecek seviyede

önem arz ettiđi görölmüştür. Yapılan çalışma sonucunda düşük çözünürlüğün çatı tipi sınıflandırma başarımını çok düşürdüğü görünmüş olup, üç mimari üzerinden yapılan testler sonucunda mimariler arası doğruluk oranının farklı değerlerde olduđu gözlemlenmiştir. Farklı mimariler kullanıldığında farklı sonuçların alınabilmiş olması ileriki çalışmalarda daha uygun bir mimari ile sınıflandırma başarısının arttırılabileceğini göstermektedir. İleride yapılacak çalışmalar için doğruluk oranını arttırabilecek diğeri bir etkende eğitim veri setinin sayısının arttırılması ve eğitim parametrelerinde daha uygun değerlerin bulunmasıdır. Bu değişiklikler uygulanarak ilerideki çalışmalarda daha yüksek sınıflandırma doğruluđu elde edileceđi düşünülmektedir.



KAYNAKLAR

- [1] Mathworks. "Getting Started with YOLO v2." MATLAB. <https://ch.mathworks.com/help/vision/ug/getting-started-with-yolo-v2.html> 2022. (Accessed: Oca.1, 2022).
- [2] G. Chitturi, "Building Detection in Deformed Satellite Images Using Mask R-CNN," Master of Science in Computer Science, 2020.
- [3] H. Ma, Y. Liu, Y. Ren, J. Yu, "Detection of Collapsed Buildings in Post-Earthquake Remote Sensing Images Based On the Improved YOLOv3," *International DOI Foundation (IDF)*, 2019.
- [4] F. Özyurt, "Uzaktan Algılama Görüntülerinin Evrişimsel Sinir Ağları ve Komşuluk Bileşen Analizi Tabanlı Özniteliklerinin Sınıflandırılması," *Afyon Kocatepe University Journal of Science and Engineering*, vol. 19, pp. 669-675, 2019.
- [5] O. Yılmaz, H. Aydın, A. Çetinkaya, "Faster R-CNN Evrişimsel Sinir Ağı Üzerinde Geliştirilen Modelin Derin Öğrenme Yöntemleri ile Doğruluk Tahmini ve Analizi: Nesne Tespiti Uygulaması," *European Journal of Science and Technology*, vol. 20, pp. 783-795, 2020.
- [6] M. Toğaçar, B. Ergen, F. ÖZYURT, "Evrişimsel Sinir Ağı Modellerinde Özellik Seçim Yöntemlerini Kullanarak Çiçek Görüntülerinin Sınıflandırılması," *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, pp. 37-45, 2020.
- [7] F. Alidoost, H. Arefi, "A CNN-Based Approach for Automatic Building Detection and Recognition of Roof Types Using a Single Aerial Image," *Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, vol. 86, pp. 235–248, 2020.

- [8] L. Hang, G. Cai, "Cnn Based Detection Of Building Roofs From High Resolution Satellite Images," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42-3, pp. 187, 2020.
- [9] F. Alidoost, H. A. Arefi, "CNN-Based Approach for Automatic Building Detection and Recognition of Roof Types Using a Single Aerial Image," *PFG*, vol. 86, pp. 235, 2018, doi: 10.1007/s41064-018-0060-5.
- [10] J. Kim, H. Bae, H. Kang, S. G. Lee, "CNN Algorithm for Roof Detection and Material Classification in Satellite Images," *Electronics*, vol. 10-13, pp. 1592, 2021, doi: 10.3390/electronics10131592.
- [11] C. Ye, H. Li, C. Li, X. Liu, Y. Li, J. Li, W. N. Gonçalves, J. M. Junior, "A Building Roof Identification CNN Based on Interior-Edge-Adjacency Features Using Hyperspectral Imagery," *Remote Sens.*, vol. 13, pp. 2927, 2021, doi: 10.3390/rs13152927.
- [12] Y. Wang, S. Li, F. Teng, Y. Lin, M. Wang, H. Cai, "Improved Mask R-CNN for Rural Building Roof Type Recognition from UAV High-Resolution Images: A Case Study in Hunan Province, China," *Remote Sens.*, vol. 14, pp. 265, 2022, doi: 10.3390/rs14020265.
- [13] M. Sun, Y. Deng, M. Li, H. Jiang, H. Huang, W. Liao, Y. Liu, J. Yang, Y. Li, "Extraction and Analysis of Blue Steel Roofs Information Based on CNN Using Gaofen-2 Imagery," *Sensors*, vol. 20, pp. 4655, 2020, doi: 10.3390/s20164655.
- [14] N. Musaoğlu, "Temel Uzaktan Algılama." *Orman Genel Müdürlüğü*. <https://www.ogm.gov.tr/tr/e-kutuphane-sitesi/EgitimDokumanlari/BSD-CBS%20E%C4%9Fitimi/Uzaktan%20Alg%C4%B1lama.pdf>. (Accessed: Nis. 16, 2014).
- [15] Z. Ekram, J. Moreau, "Photogrammetry and Remote Sensing." *Science Aid*. https://scienceaid.net/Photogrammetry_and_Remote_Sensing.

- [16] M. Queisera, M. Burtona, R. Kazahaya, “Insights into geological processes with CO2 remote sensing – A review of technology and applications,” *Earth-Science Reviews*. Vol. 188, pp. 389–426, 2019.
- [17] A.M. Lechner, G. M. Foody, and D. S. Boyd, “Applications in Remote Sensing to Forest Ecology and Management,” *One Earth Primer*, pp.2, 2020, doi: 10.1016/j.oneear.2020.05.001.
- [18] S. Ateş, E. Demir, “Uzaktan Algılamada Çözünürlüğe Bağlı Veri Kazanımı Potansiyeli,” *TMMOB Harita ve Kadastro Mühendisleri Odası 12. Türkiye Harita Bilimsel ve Teknik Kurultayı*, pp. 4, 2009.
- [19] M. Erdoğan, H. Akdeniz, “Uzaktan Algılama Amaçlı Uydu Sistemlerindeki Son Gelişmeler,” *Harita Genel Komutanlığı Harita Dergisi*, vol. 132, pp. 12, 2017.
- [20] M. Erdoğan, “Uzaktan Algılama Amaçlı Uydu Sistemlerindeki Son Gelişmeler,” *Hrt.Gn.K.lığı, Harita Dergisi*, vol. 132, pp. 3-4.
- [21] Y. Bengio, “Learning Deep Architectures for AI,” *Found. Trends Mach. Learn*, vol. 2, no. 1, pp. 1–127, 2009.
- [22] M. Marvin, S. Papert, “Perceptrons: An Introduction to Computational Geometry,” Cambridge, MA, USA: *MIT Press*, 1969.
- [23] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *In Proceedings of the 25th International Conference on Neural Information Processing Systems*, vol. 1, pp. 1097–1105, 2012.
- [24] Ö. İnik, E. Ülker, “Derin Öğrenme ve Görüntü Analizinde Kullanılan Derin Öğrenme Modelleri,” *Gaziosmanpasa Journal of Scientific Research*, vol. 6, pp. 85-104, 2017.

- [25] R. Dahl, M. Norouzi, and J. Shlens, "Pixel recursive super resolution," *arXiv preprint*, arXiv: 1702.00783.
- [26] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," *arXiv preprint*, 2017.
- [27] S. Iizuka, E. Simo-Serra, and H. Ishikawa, "Let there be color!: joint end-to-end learning of global and local image priors for automatic image colorization with simultaneous classification," *ACM Transactions on Graphics (TOG)*, vol. 35(4), pp. 110, 2016.
- [28] F. Doğan, İ. Türkoğlu, "Derin Öğrenme Modelleri ve Uygulama Alanlarına İlişkin Bir Derleme," *DÜMF Mühendislik Dergisi*, vol. 10, pp. 409-445, 2018.
- [29] D. Ciregan, U. Meier, and J. Schmidhuber, "Multi-column deep neural networks for image classification," *In Computer vision and pattern recognition (CVPR)*, 2012 IEEE conference on, pp. 3642-3649, 2012.
- [30] F. Dogan, I. Turkoglu, "Classification Of Satellite Images By Deep Learning," *8th International Advanced Technologies Symposium*, 2017.
- [31] F. Hu, G. S. Xia, J. Hu, and L. Zhang, "Transferring deep convolutional neural networks for the scene classification of highresolution remote sensing imagery," *Remote Sensing*, vol. 7(11), pp. 14680-14707, 2015.
- [32] Q. Zou, L. Ni, T. Zhang, and Q. Wang, "Deep learning based feature selection for remote sensing scene classification," *Geoscience and Remote Sensing Letters*, vol. 12(11), pp. 2321-2325, 2015.
- [33] T.H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng, and Y. Ma, "PCANet: A simple deep learning baseline for image classification?," *IEEE Transactions on Image Processing*, vol. 24(12), pp. 5017-5032, 2015.

- [34] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep Face Recognition," *In BMVC*, Vol. 1, No. 3, pp. 6, 2015.
- [35] Y. Sun, Y. Chen, X. Wang, and X. Tang, "Deep learning face representation by joint identification-verification," *In Advances in neural information processing systems*, pp. 1988-1996, 2014.
- [36] Y. Sun, D. Liang, X. Wang, and X. Tang, "Deepid3: Face recognition with very deep neural networks," *arXiv preprint*, arXiv:1502.00873, 2015.
- [37] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep learning-based classification of hyperspectral data," *IEEE Journal of Selected topics in applied earth observations and remote sensing*, vol. 7(6), pp. 2094-2107, 2014.
- [38] K. L. Hua, C. H. Hsu, S. C. Hidayati, W. H. Cheng, and Y. J. Chen, "Computer-aided classification of lung nodules on computed tomography images via deep learning technique," *OncoTargets and therapy*, pp. 8, 2015.
- [39] X. An, D. Kuang, X. Guo, Y. Zhao, and L. He, "A deep learning method for classification of EEG data based on motor imagery," *In International Conference on Intelligent Computing*, pp. 203-210, 2014.
- [40] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, "Deep learning for visual understanding: A review," *Neurocomputing*, vol. 187, pp. 27-48, 2016.
- [41] D. Erhan, C. Szegedy, A. Toshev, and D. Anguelov, "Scalable object detection using deep neural networks," *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2147-2154, 2014.
- [42] J. Salamon, and J. P. Bello, "Deep convolutional neural networks and data augmentation for environmental sound classification," *IEEE Signal Processing Letters*, vol. 24(3), pp. 279-283, 2017.

- [43] O. Fried, and R. Fiebrink, "Cross-modal Sound Mapping Using Deep Learning," *In NIME*, pp. 531-534, 2013.
- [44] J. Antony, K. McGuinness, N. E. O'Connor, and K. Moran, "Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks," *In Pattern Recognition (ICPR), 2016 23rd International Conference on*, pp. 1195-1200, 2016.
- [45] H. C. Shin, K. Roberts, L. Lu, D. DemnerFushman, J. Yao, and R. M. Summers, "Learning to read chest x-rays: Recurrent neural cascade model for automated image annotation," *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2497-2506, 2016.
- [46] J. Yu, K. Weng, G. Liang, and G. Xie, "A vision-based robotic grasping system using deep learning for 3D object recognition and pose estimation," *In Robotics and Biomimetics (ROBIO), 2013 IEEE International Conference on*, pp. 1175-1180, 2015.
- [47] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, and A. Rabinovich, "Going deeper with convolutions," *Computer Vision and Pattern Recognition Conference*, 2015.