

**KARADENİZ TEKNİK ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ**

**HARİTA MÜHENDİSLİĞİ ANABİLİM DALI**

**UZAKTAN ALGILANMIŞ GÖRÜNTÜLERDE MASK R-CNN KULLANIMI**

**YÜKSEK LİSANS TEZİ**

**Harita Müh. Betül SARALIOĞLU**

**Temmuz 2022  
TRABZON**



**KARADENİZ TEKNİK ÜNİVERSİTESİ**  
**FEN BİLİMLERİ ENSTİTÜSÜ**

**HARİTA MÜHENDİSLİĞİ ANABİLİM DALI**

**UZAKTAN ALGILANMIŞ GÖRÜNTÜLERDE MASK R-CNN KULLANIMI**

**Betül SARALIOĞLU**

**Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsünde**  
**"HARİTA YÜKSEK MÜHENDİSİ"**  
**Unvanı Verilmesi İçin Kabul Edilen Tezdir.**

**Tezin Enstitüye Verildiği Tarih : 06 / 06 / 2022**

**Tezin Savunma Tarihi : 04 / 07 / 2022**

**Tez Danışmanı : Prof. Dr. Oğuz GÜNGÖR**

**Trabzon 2022**

## ÖNSÖZ

“UZAKTAN ALGILANMIŞ GÖRÜNTÜLERDE MASK R-CNN KULLANIMI” adlı yüksek lisans çalışmamda danışmanım olarak araştırmalarımı yönlendiren, fikirleri ile vizyonumu geliştiren, çok değerli saygıdeğer hocam Prof. Dr. Oğuz GÜNGÖR’ e teşekkür ederim.

Yüksek lisans çalışmamda kullandığım Worldview-3 uydu görüntülerinin teminini sağlayan ve değerli fikirleriyle katkıda bulunan sayın hocam Prof. Vittorio CASELLA’ya teşekkür ederim.

Yüksek lisans çalışmam süresince bilgi, görüş, önerilerinden yararlandığım ve bana her zaman destek olan abim Arş. Gör. Dr. Ekrem SARALIOĞLU’na teşekkür ederim.

Karşılıksız destek ve sevgileriyle her zaman yanımda olan yoğun çalışma günlerimde moral ve destekleriyle bana güç veren ailem ve dostlarıma teşekkürü bir borç bilirim.

Betül SARALIOĞLU

Trabzon 2022

## TEZ ETİK BEYANNAMESİ

Yüksek Lisans Tezi olarak sunduğum “UZAKTAN ALGILANMIŞ GÖRÜNTÜLERDE MASK R-CNN KULLANIMI” başlıklı bu çalışmayı baştan sona kadar danışmanım Prof. Dr. Oğuz GÜNGÖR’ ün sorumluluğunda tamamladığımı, verileri/örnekleri kendim topladığımı, deneyleri/analizleri ilgili laboratuvarlarda yaptığımı/yaptırdığımı, başka kaynaklardan aldığım bilgileri metinde ve kaynakçada eksiksiz olarak gösterdiğimi, çalışma sürecinde bilimsel araştırma ve etik kurallara uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul ettiğimi beyan ederim. 04/07/2022

Betül SARALIOĞLU

## İÇİNDEKİLER

|                                                        | <u>Sayfa No</u> |
|--------------------------------------------------------|-----------------|
| ÖNSÖZ .....                                            | III             |
| TEZ ETİK BEYANNAMESİ.....                              | IV              |
| İÇİNDEKİLER .....                                      | V               |
| ÖZET .....                                             | VII             |
| SUMMARY .....                                          | VIII            |
| ŞEKİLLER DİZİNİ .....                                  | IX              |
| TABLolar DİZİNİ .....                                  | XII             |
| 1. GENEL BİLGİLER .....                                | 1               |
| 1.1. Giriş .....                                       | 1               |
| 1.2. Çalışmanın Amacı .....                            | 6               |
| 1.3. Uzaktan Algılamada Sınıflandırma .....            | 8               |
| 1.4. Yapay Sinir Ağları .....                          | 11              |
| 1.5. Derin Öğrenme .....                               | 13              |
| 1.5.1. Tekrarlayan Sinir Ağları (RNN).....             | 17              |
| 1.5.2. Uzun Kısa Vadeli Sinir Ağları (LSTM).....       | 19              |
| 1.5.3. Kısıtlı Bozaltman Makineleri (RBM) .....        | 20              |
| 1.5.4. Derin İnanç Ağları (DBN) .....                  | 21              |
| 1.5.5. Evrişimli Sinir Ağları (ESA).....               | 22              |
| 1.5.5.1. Evrişim Katmanı.....                          | 23              |
| 1.5.5.2. Doğrusal Olmayan Katman (Non-Linearity) ..... | 25              |
| 1.5.5.3. Havuzlama Katmanı .....                       | 26              |
| 1.5.5.4. Tamamen Bağlı Katman .....                    | 26              |
| 1.5.5.5. Sınıflandırma Katmanı.....                    | 27              |

|        |                                                      |    |
|--------|------------------------------------------------------|----|
| 1.6.   | Derin Ağ Mimarileri .....                            | 27 |
| 1.6.1. | AlexNet .....                                        | 27 |
| 1.6.2. | VGG-16 Net .....                                     | 28 |
| 1.6.3. | ResNet .....                                         | 29 |
| 1.7.   | Nesne Tespiti .....                                  | 31 |
| 1.7.1. | R-CNN .....                                          | 31 |
| 1.7.2. | SPP-net.....                                         | 33 |
| 1.7.3. | Fast R-CNN.....                                      | 34 |
| 1.7.4. | Faster R-CNN.....                                    | 34 |
| 1.7.5. | Mask R-CNN.....                                      | 36 |
| 1.8.   | Derin Öğrenmede Kullanılan Yaygın Veri Setleri ..... | 37 |
| 1.9.   | Transfer Öğrenme .....                               | 41 |
| 2.     | YAPILAN ÇALIŞMALAR .....                             | 42 |
| 2.1.   | Çalışma Alanı ve Kullanılan Veri .....               | 43 |
| 2.2.   | Çalışma Ortamı.....                                  | 48 |
| 2.3.   | Veri Setinin Oluşturulması.....                      | 49 |
| 2.4.   | Mask R-CNN Uygulaması .....                          | 54 |
| 2.4.1. | Kod Yapısı.....                                      | 54 |
| 2.4.2. | Mask R-CNN’de Sınıflandırma Adımları.....            | 58 |
| 2.5.   | Sınıflandırma Sonuçları .....                        | 62 |
| 2.5.1. | Bölge Tespiti Sınıflandırma Sonuçları.....           | 62 |
| 2.5.2. | Bina Tespiti Sınıflandırma Sonuçları .....           | 65 |
| 3.     | BULGULAR VE İRDELEMER .....                          | 66 |
| 3.1.   | Model Eğitimi ve Deneysel Sonuçlar .....             | 67 |
| 4.     | SONUÇLAR VE ÖNERİLER .....                           | 85 |
| 5.     | KAYNAKLAR.....                                       | 88 |

ÖZGEÇMİŞ

ÖZET

UZAKTAN ALGILANMIŞ GÖRÜNTÜLERDE MASK R-CNN KULLANIMI

Betül SARALIOĞLU

Karadeniz Teknik Üniversitesi  
Fen Bilimleri Enstitüsü  
Harita Mühendisliği Anabilim Dalı  
Danışman: Prof. Dr. Oğuz GÜNGÖR  
2022, 98 Sayfa

Çeşitli uzaktan algılama uyduları sayesinde günümüzde çok yüksek miktarda veri elde edilmektedir. Çok yüksek uzamsal çözünürlüklü uzaktan algılanmış verilerin yanında Sentinel gibi ücretsiz verilere erişmekte mümkün olmaktadır. Bu kadar büyük hacimli ve karmaşık yer yüzü detaylarına sahip uzaktan algılanmış görüntülerinin sınıflandırılması geleneksel yöntemlerle oldukça zordur. Mask R-CNN gibi hızlı derin öğrenme modelleri bir kere yüksek miktarda eğitim verisiyle eğitilerek farklı bölgelere ait görüntülerin saniyeler içerisinde sınıflandırılmasına imkân sağlayabilmektedir. Bu çalışmada en yeni evrişimli sinir ağı modellerinden biri olan Mask R-CNN kullanılarak Sentinel-2 ve Worldview-3 uydu görüntüleri üzerinde sınıflandırma çalışmaları yapılmıştır. Veri seti oluşturmak için toplamda 1515 görüntü, 18811 adet etiket oluşturulmuştur. Oluşturulan veri setleri üzerinde iki farklı senaryo gerçekleştirilmiştir. Bu senaryolardan ilkinde Sentinel-2 uydu görüntüleri kullanılarak bölge tespiti yapılmıştır. Sentinel-2 uydu görüntüsünde yapılaşma alanları tespit edilirken ResNet101 ve ResNet50 ağ mimarileri kullanılmıştır. Bu oluşturulan modellerden sırasıyla %81, %79 doğruluk elde edilmiştir. Sentinel-2 uydu görüntüsü üzerinden arazi örtüsü sınıflandırması amacıyla beş sınıf etiketli veri seti oluşturulmuş ve bunun sonucunda %74 doğruluk elde edilmiştir. İkinci senaryoda ise Worldview-3 uydu görüntüleri kullanılarak bina tespiti yapılmıştır ve %83 doğruluk elde edilmiştir. Elde edilen sonuçlar Mask R-CNN ile uydu görüntülerinin sınıflandırılmasında başarı ile kullanılabileceğini göstermektedir. Bu yöntemin uydu görüntülerinde kullanımı anlık sınıflandırmaya imkân sağladığından önem taşımaktadır.

**Anahtar Kelimeler:** Derin Öğrenme, Mask R-CNN, Sınıflandırma, Uzaktan Algılama

Master Thesis

## SUMMARY

### USING MASK R-CNN IN REMOTE SENSING IMAGES

Betül SARALIOĞLU

Karadeniz Technical University  
The Graduate School of Natural and Applied Sciences  
Geomatics Engineering Graduate Program  
Supervisor: Prof. Dr. Oğuz GÜNGÖR  
2022, 98 Pages

Thanks to various remote sensing satellites, a very large amount of data is obtained today. It is possible to access free data such as Sentinel and remotely sensed data with very high spatial resolution. It is challenging to classify remotely sensed images with such large volumes and complex earth details using traditional methods. Fast deep learning models such as Mask R-CNN can be trained once with a large amount of training data, allowing images of different regions to be classified within seconds. In this study, classification studies were carried out on Sentinel-2 and Worldview-3 satellite images using Mask R-CNN, one of the newest convolutional neural network models. A total of 1515 images and 18811 labels were created to create the data set. Two different scenarios were carried out on the created data sets. In the first of these scenarios, the region was determined by using Sentinel-2 satellite images. ResNet101 and ResNet50 network architectures were used while detecting built up area in the Sentinel-2 satellite image. 81% and 79% accuracy were obtained from these created models, respectively. Five class labeled data sets were created for land cover classification on the Sentinel-2 satellite image, and as a result, 74% accuracy was obtained. In the second scenario, the building was detected using Worldview-3 satellite images and 83% accuracy was obtained. Obtained results show that Mask R-CNN can be used successfully in classification of satellite images. The use of this method in satellite images is important because it allows instant classification.

**Keywords:** Deep Learning, Mask R-CNN, Classification, Remote Sensing

## ŞEKİLLER DİZİNİ

|                                                                       | <u>Sayfa No</u> |
|-----------------------------------------------------------------------|-----------------|
| Şekil 1. Yapay sinir hücresi.....                                     | 12              |
| Şekil 2. Yapay sinir ağı .....                                        | 12              |
| Şekil 3. Yapay zeka hiyerarşisi.....                                  | 15              |
| Şekil 4. (a) Makine öğrenimi, (b) Derin öğrenme .....                 | 16              |
| Şekil 5. Derin sinir ağ yapısı .....                                  | 17              |
| Şekil 6. Tekrarlayan sinir ağı yapısı .....                           | 18              |
| Şekil 7. RNN yapısı .....                                             | 19              |
| Şekil 8. LSTM yapısı .....                                            | 19              |
| Şekil 9. Kısıtlı bozaltman makinesi yapısı.....                       | 20              |
| Şekil 10. Derin inanç ağı yapısı.....                                 | 21              |
| Şekil 11. Evrişimli sinir ağı mimarisi .....                          | 23              |
| Şekil 12. Evrişim tanımlanması.....                                   | 23              |
| Şekil 13. Evrişim işlemi .....                                        | 24              |
| Şekil 14. ESA filtre örnekleri .....                                  | 25              |
| Şekil 15. Relu grafiği .....                                          | 25              |
| Şekil 16. Havuzlama işlemi.....                                       | 26              |
| Şekil 17. AlexNet ağ mimarisi yapısı (Krizesky vd., 2012). .....      | 28              |
| Şekil 18. VGG ağ mimari yapısı (Simonyan vd., 2014).....              | 29              |
| Şekil 19. ResNet atlama bağlantısı .....                              | 30              |
| Şekil 20. ResNet ağ mimarisi (He vd., 2016).....                      | 30              |
| Şekil 21. R-CNN mimari yapısı (Girshick vd., 2014) .....              | 32              |
| Şekil 22. Seçici arama (Uijlings vd., 2012).....                      | 32              |
| Şekil 23. SPP-net mimarisi (He vd., 2015) .....                       | 33              |
| Şekil 24. Fast R-CNN mimari ağ yapısı (Girshick, 2015) .....          | 34              |
| Şekil 25. Faster R-CNN yapısı (Girshick vd., 2017).....               | 35              |
| Şekil 26. Mask R-CNN ağ mimarisi yapısı.....                          | 36              |
| Şekil 27. ImageNet örnek veri seti (Deng vd., 2009) .....             | 37              |
| Şekil 28. PASCAL VOC örnek veri seti (Everingham ve Winn., 2012)..... | 38              |

|                                                                                                                                                             |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Şekil 29. Open Image örnek veri seti (Kuznetsova vd., 2020).....                                                                                            | 39 |
| Şekil 30. MS COCO örnek veri seti (Lin vd., 2014) .....                                                                                                     | 40 |
| Şekil 31. Tez iş akış şeması.....                                                                                                                           | 42 |
| Şekil 32. Türkiye haritası .....                                                                                                                            | 43 |
| Şekil 33. (a) İtalya haritası, (b) Pavia Şehri .....                                                                                                        | 44 |
| Şekil 34. USGS sitesi görüntü indirme modülü .....                                                                                                          | 45 |
| Şekil 35. Sentinel 2 uydu görüntülerinden bazı örnekler .....                                                                                               | 46 |
| Şekil 36. WV3 uydu görüntüsü (a)Pavia bölgesi, (b) Çalışma alanı .....                                                                                      | 48 |
| Şekil 37. Sentinel 2 veri seti örnekleri .....                                                                                                              | 49 |
| Şekil 38. WV3 veri seti örnekleri .....                                                                                                                     | 50 |
| Şekil 39. VGG Image Annotator’da görüntü etiketleme.....                                                                                                    | 51 |
| Şekil 40. Sentinel-2 görüntüsünde beş sınıf için etiketleme işlemi .....                                                                                    | 52 |
| Şekil 41. WV3 etiketleme işlemi .....                                                                                                                       | 53 |
| Şekil 42. Veri setine ait json dosyası .....                                                                                                                | 53 |
| Şekil 43. Veri seti örneği.....                                                                                                                             | 54 |
| Şekil 44. Veri seti tanımlatma .....                                                                                                                        | 55 |
| Şekil 45. Eğitim parametrelerinin belirlenmesi .....                                                                                                        | 56 |
| Şekil 46. RPN ile sınırlayıcı kutu belirleme a) orijinal görüntü, b) pozitif sınırlayıcı kutu<br>c) negatif sınırlayıcı kutu d) nötr sınırlayıcı kutu ..... | 59 |
| Şekil 47. Pozitif sınırlayıcı kutular (a) pozitif 100 sınırlayıcı kutu, (b) pozitif 50<br>sınırlayıcı kutu .....                                            | 59 |
| Şekil 48. En iyi sınırlayıcı kutular .....                                                                                                                  | 60 |
| Şekil 49. RoI üzerinde maske oluşturma.....                                                                                                                 | 61 |
| Şekil 50. Görüntü üzerinde maske oluşumu.....                                                                                                               | 61 |
| Şekil 51. ResNet101 ile oluşturulmuş modelin sınıflandırma sonuçları.....                                                                                   | 62 |
| Şekil 52. ResNet101 ile oluşturulmuş modelin sınıflandırma sonuçları.....                                                                                   | 63 |
| Şekil 53. Sentinel-2 görüntüsünde arazi örtüsü sınıflandırma sonuçları .....                                                                                | 64 |
| Şekil 54. WV3 uydu görüntüsünün sınıflandırma sonuçları .....                                                                                               | 65 |
| Şekil 55. Kutu boyutları .....                                                                                                                              | 68 |
| Şekil 56. ResNet101 için kayıp grafiği .....                                                                                                                | 69 |
| Şekil 57. ResNet50, ResNet101 kayıp grafikleri.....                                                                                                         | 70 |
| Şekil 58. Arazi örtüsü sınıf için kayıp grafiği .....                                                                                                       | 71 |
| Şekil 59. WV3 uydu görüntüleri için kayıp grafiği .....                                                                                                     | 72 |

|                                                                                                            |    |
|------------------------------------------------------------------------------------------------------------|----|
| Şekil 60. Tüm çalışma için kayıp grafiği .....                                                             | 72 |
| Şekil 61. Birleşim üzerinden kesişme .....                                                                 | 73 |
| Şekil 62. Sınıflandırma sonuçları (a) Orijinal görüntü, (b) Resnet101, (c) Resnet50,<br>(d) Beş sınıf..... | 77 |
| Şekil 63. Sınıflandırma sonuçları (a) Orijinal görüntü, (b) Resnet101, (c) Resnet50,<br>(d) Beş sınıf..... | 78 |
| Şekil 64. Sınıflandırma sonuçları (a) orijinal görüntü, (b) Resnet101, (c) Resnet50,<br>(d) Beş sınıf..... | 80 |
| Şekil 65. Sınıflandırma sonuçları (a) orijinal görüntü, (b) Resnet101, (c) Resnet50,<br>(d) Beş sınıf..... | 81 |
| Şekil 66. WV3 görüntüsünün sınıflandırma sonuçları (a) Orijinal görüntü, (b)<br>Sınıflandırma .....        | 82 |
| Şekil 67. WV3 görüntüsünün sınıflandırma sonuçları (a) Orijinal görüntü, (b)<br>Sınıflandırma .....        | 83 |
| Şekil 68. WV3 uydu görüntüsünün sınıflandırılması (a) çalışma alanı, (b) Sınıflandırma<br>sonuçları .....  | 84 |

## TABLÖLAR DİZİNİ

|                                                      | <b><u>Sayfa No</u></b> |
|------------------------------------------------------|------------------------|
| Tablo 1. Sentinel 2 uydu görüntüsü özellikleri.....  | 44                     |
| Tablo 2. WorldView-3 uydu görüntüsü özellikleri..... | 47                     |
| Tablo 3. Modeller için hiperparametreler .....       | 57                     |
| Tablo 4. Karışıklık matrisi.....                     | 74                     |
| Tablo 5. Değerlendirme sonuçları .....               | 75                     |



## 1. GENEL BİLGİLER

### 1.1. Giriş

Uzaktan algılanmış verilerin miktarı son yıllarda önemli derecede artmıştır. Bu artış günümüzde de hızlanan bir ivme ile devam etmektedir. Çeşitli amaçlara yönelik artan uydu sayısı, farklı karakteristiklerde (spektral, radyometrik, uzamsal, zamansal çözünürlük bakımında) gelişen sensör ve bilgisayar teknolojileri bu artışta önemli rol oynamaktadır. Bununla birlikte uzaktan algılanmış verilere erişim de giderek daha kolay hale gelmektedir. Ücretsiz veri imkânı sunan Landsat ve Sentinel takım uyduları sayısının artması, Copernicus, Sentinel, Hub, Google Earth ve EarthServer gibi çevrimiçi platformlar ile çeşitli uydu görüntülerine erişim sağlanabilmesi verilere ulaşımı kolaylaştırmaktadır. Amerika Birleşik Devletleri'nin Landsat serisi, Avrupa Birliği'nin Copernicus projesi kapsamında Sentinel uyduları ve Çin'in Gaofen serisi gibi sürekli güncellenerek devam eden büyük çaplı uydu programlarının yanında ticari firmaların da gözlem uydu sayıları sürekli artmaktadır. Örneğin Planet, her gün büyük miktarda uzaktan algılama verisi üreten yaklaşık 200 küçük uydu fırlatmıştır (Zhang vd., 2021). Günümüze kadar fırlatılmış 1000 üzeri uydunun bir tanesi tarafından üretilen veriler günde yaklaşık terabayt hacminde birikmektedir. Avrupa Uzak Uzay Ajansı (ESA)'nın ürettiği Dünya gözlem verileri sayısı 1.5 petabaytı aşmış durumdadır. Yalnızca Sentinel serisi günde yaklaşık 23 TB'lık veri üretme hacmine sahip olmaktadır. Ekvatorda 5 günlük bir zamansal çözünürlüğe (ve birden fazla yörünge tarafından kapsanan alanlar için daha da az) izin veren iki kutupsal yörüngeli uydu platformunun bir takımıyla Sentinel-2 ile Haziran 2017'den beri faaliyet gösteren Sentinel-2 görevi araştırma topluluğu içinde büyük ilgi görmüştür. 1 Aralık 2017'den 30 Kasım 2018'e kadar, Sentinel Veri Erişim Sistemi günde 26.500'den fazla ürün yayınlama oranına ve günlük ortalama 166 TB'lık veri indirme hacmine sahip olmuştur. Ayrıca, İHA'lar gibi diğer birçok yaklaşım, çeşitli uzaktan algılama veri toplama görevlerinde sıklıkla kullanılmaktadır (Bhardwaj vd., 2016; Prinz vd., 2010). Bu artan uzaktan algılanmış veriler büyük veri kriterlerini karşılamaktadır (Xu vd., 2020). Devlet projeleri, ticari uygulamalar veya akademik araştırmalarda büyük verilerin tanımlanmasında veri hacmi, veri hızı ve veri çeşitliliği şeklinde bu 3 kritere atıf yapılmaktadır (Liu vd., 2018).

Uzaktan algılanmış veriler için;

Veri Hacmi: Artan uydu ve sensör sayılarına bağlı üretilen çok sayıda veriyi,

Veri Çeşitliliği: Birden çok kaynaktan (uydu, İHA, yer vb.) çok modlu (LiDAR, radar, optik vb.), çok çözünürlüklü (1 cm'den 100 km'ye kadar) ve çok zamanlı (düzensiz tarih ve belirli periyotlarda toplanan) verileri oluşturmaktadır.

Veri Hızı: Dünya yüzeyinde Dünya gözleminin hedefi değişmesiyle Uzaktan Algılama (UA) verileri hızlı bir büyüme oranında elde edilebilmektedir. Ayrıca, artan sayıda görev, karar desteği için en son bilgileri üretmek (neredeyse) gerçek zamanlı işleme ve analiz gerektirir, örn. bir orman yangınında saatler içinde yüzlerce hektarlık orman kurtarılabilir.

Bunların dışında uzaktan algılanmış veriler kendilerine özgü bazı karakteristiklere de sahip olmaktadır.

- Çeşitli sensör yetenekleri ve algılama mesafeleri nedeniyle, UA verileri çok çözünürlüklü ve çok ölçeklidir.
- Görüntüleme geometrilerinin ve veri türünün tamamen farklı olduğu optik (çoklu ve hiperspektral), LIDAR ve sentetik açıklıklı radar (SAR) gibi UA verileri çok modlu bir yapıya sahiptir.
- Ayrıca uzaktan algılanmış veriler zaman değişkenine de sahiptir. Günümüzde zaman değişkeni giderek daha önemli hale gelmektedir. Geçen yüzyılda yüksek uzamsal ve spektral çözünürlük arayışına ek olarak, son zamanlarda yüksek zamansal çözünürlüğe olan talep artmıştır.

Bu büyük veriler uzaktan algılama uygulama alanlarının genişletilmesine yardımcı olmaktadır. Bunun yanında büyük verilerin işlenmesi yeni zorluklar oluşturmaktadır (Zhang vd., 2021). Buna bir örnek olarak arazi örtüsünün küresel boyutta doğru, hızlı ve otomatik olarak nasıl sınıflandırılabilirliği verilebilir. Bu büyüyen veri miktarına bağlı zorlukların çözümünde gelişen bilgisayar teknolojileri ve yapay zekâ kullanımındaki gelişmeler önemli rol oynamaktadır. UA uygulamalarının performansını artırmak için makine öğrenimi algoritmaları (yapay sinir ağları, destek vektör makineleri ve rastgele ormanlar gibi) sıklıkla kullanılmıştır. Son zamanlarda ise, birçok alanda geniş uygulamaları olan derin öğrenme (Deep Learning (DL)) yöntemleri kullanımı öne çıkmaktadır. Temsili ve ayırt edici özelliklerin hiyerarşik öğreniminin sağlanması ile derin öğrenme, uzaktan algılanmış büyük

verilerin işleme ve analizi konusunda umut verici yeni bir araştırma olanağı sunmuştur. (Ma vd., 2019).

Derin Öğrenme (DÖ) 1980'li yıllardan beri var olmasına rağmen hem bilgisayar teknolojileri (donanım ve yazılım) hem de yeterli veri seti olmamasından dolayı aktif bir şekilde kullanılamamıştır (Wani vd., 2020). Daha hızlı donanım kaynaklarının kullanılması ve bulut bilişim ortamlarına kolay erişim sayesinde DÖ algoritmaları yapay zekaya dayanan çok çeşitli bilimsel disiplinlerde artan bir popülerlik elde etmiştir. Uygulama alanları ilaç keşfinden otonom navigasyona, Dünya ve çevresel modellemeye kadar uzanmaktadır. DÖ'nün uzaktan algılama alanındaki başarıları sayesinde görüntü tanımlamada destek vektör makinesi (DVM) ve rastgele orman (RO) gibi geleneksel makine öğrenimi tekniklerini geride bırakmıştır. DÖ yöntemleri ile insan müdahalelerine ihtiyaç duymadan ses, video, resim ve metinlere ait verilerden (örneğin bir resim için piksel) otomatik olarak öğrenme yapabilmektedir. Bu sayede uzaktan algılamalı dijital görüntü işlemede farklı uygulamalarda en iyi sonuçlar elde edilmektedir (Ma vd., 2019). Yapılan uygulamalara kenar keskinleştirme (Zhang vd., 2020), nesne tespiti (Ammour vd., 2017), değişim analizi (De Bem vd., 2020), semantik segmentasyon (Diakogiannis vd., 2020), zaman serisi analizi (Crisóstomo De Castro Filho vd., 2020) örnek olarak verilebilmektedir (Carvalho vd., 2020).

Derin öğrenme yaklaşımı olan Evrişimli Sinir Ağları (ESA) İngilizcesi Convolutional Neural Networks (CNN), biyolojik görsel korteksten ilham alarak oluşturulmuştur. Görsel korteks, görme alanında yer alan belirli bölgelere duyarlı nöronlara sahiptir. Örneğin, bazı nöronlar dikey kenarlara bazıları ise yatay kenarlara maruz kaldığında aktive edilir (Hubel vd., 1968). Belirli bir göreve sahip bu nöronların fikri ESA'yı oluşturan temel olarak kabul edilmektedir (Wani vd., 2020). ESA, görüntü sınıflandırma, nesne algılama, konuşma tanıma, doğal dil işleme, tıbbi görüntü analizi gibi çeşitli uygulamalarda iyi sonuçlar göstermiştir. ESA'nın bilgisayarlı görme alanındaki bu yaygınlığı arazi kullanımı sınıflandırması, hava görüntülerinden nesne tespiti, uydu görüntülerinin analizi gibi birçok görevi yerine getirebildikleri için uzaktan algılama alanında da aktif bir şekilde kullanılmaktadır.

Literatürde ESA ile nesne tespiti, arazi kullanım alanlarının belirlenmesi gibi alanlarda çalışmalar bulunmaktadır:

(Kussul vd., 2017)'de Landsat-8 ve Sentinel-1A uydu görüntüleri kullanarak arazi örtüsü ve bitki türlerinin sınıflandırılması amaçlanmıştır. Bu çalışmada derin öğrenme tekniği olan 1D ve 2D ESA modelleri kullanılarak buğday, mısır, ayçiçeği, şeker pancarı

gibi ürünlerin sınıflandırılması %85 başarı puanı ile belirlenmiştir. (Zhao vd., 2018)'de Landsat uydu görüntüsü kullanılarak Çin'deki önemli bir yerel bölge olan Jinginji bölgesindeki arazi örtüsü sınıflandırılmıştır. Çalışmada ESA tabanlı PSPNet yöntemi kullanılmıştır. Çalışmanın sonucunda Landsat görüntüsündeki referans arazi örtüsü haritasının yeterince doğru olduğunda ve eğitim örneği yeterli olduğunda PSPNet'in iyi öğrenebildiği tespit edilmiştir. Genel doğruluk %83 olarak hesaplanmıştır. (Tong vd., 2018) farklı koşullar altında elde edilen uzaktan algılama görüntülerini etkili bir şekilde sınıflandırmak için ResNet-50 ağ mimarisine dayanan ESA kullanmıştır. Çalışma alanı için 150 adet Gaofen-2 uydu görüntüsünü kullanarak arazi örtüsü sınıflandırılmıştır. Yapılan çalışmada küçük ölçeğe sahip görüntülerin daha yüksek doğruluğa sahip olduğu görülmektedir ve geleneksel yöntemlere kıyasla üstün sınıflandırma performansı elde edilmiştir.

Son yıllarda ESA'nın nesne tespiti problemlerinde kullanımı için geliştirilen bölgesel ESA tabanlı modeller ile önemli bir ilerleme kaydedilmiştir. Bölge Tabanlı Evrişimli Sinir Ağı (R-CNN) (Girshick vd., 2014), R-CNN ailesindeki ilk makale olmuştur. Bu makalede CNN'lerin algılama performansını büyük ölçüde iyileştirmek için nasıl kullanılabileceği gösterilmiştir. R-CNN, nesne tespiti alanına yeni bir bakış açısı kazandırmıştır. Bunun yanında, R-CNN'de görüntü başına 2000 bölge önerisinin sınıflandırılması gerektiğinden bir ağı eğitmek hala çok fazla zaman almaktadır (Girshick vd., 2014; Zaidi vd., 2022). Her test görüntüsü için nesne tespit süresinin yaklaşık 47 saniye sürmesi ise R-CNN'in anlık nesne tespiti uygulamalarında kullanımına engel olmuştur. R-CNN'den kısa bir süre sonra He vd., (2015) tarafından yapılan çalışma ile Uzamsal Piramit Havuzlama (SPP) ağı geliştirilmiştir. SPP-Net R-CNN modeli ile benzer doğrulukta sonuçlar vermesine karşın, R-CNN modelinden önemli ölçüde daha hızlı olmuştur (Zaidi vd., 2022). SPPNet'in önemli özelliklerinden birisi ise, R-CNN'de giriş görüntüsünün sabit bir boyutta olma kısıtlamasının kaldırılmış olması olmuştur. Başka bir deyişle, SSPNet modeli girdi görüntü boyutundan bağımsız olarak çalışabilmektedir (Grauman vd., 2005). Bununla birlikte, SSPNet mimarisi R-CNN'ye benzer olduğundan, R-CNN'nin çok aşamalı eğitim, hesaplama açısından pahalılık ve eğitim süresi gibi dezavantajlarına da sahip olmuştur. Fast R-CNN, SSPNet ve R-CNN'deki bazı problemleri çözmüştür. SSPNet ve R-CNN'in aksine Fast R-CNN uçtan uca tek bir eğitilebilir sistem olarak tasarlanmıştır. Fast R-CNN'in, R-CNN'e göre doğrulukta bir miktar artışın yanında hız bakımından önemli derecede (146 kat) daha hızlı olduğu gösterilmiştir. Fast R-CNN ile eğitim prosedürü basitleştirilerek bölge önerisi 2000

farklı aday bölge olmadan tek bir aday bölge üzerinden çalıştığından gerçek zamanlıya yakın doğrulukta nesne tespiti gerçekleştirilmiştir (Girshick vd., 2014).

Fast R-CNN modelinin avantajlarına rağmen, bölge önerisi için zaman alıcı Seçici Arama (Selective Research) algoritması bir dezavantaj oluşturmaktadır. Bu sorunun çözümü için Faster R-CNN ağı sunulmuştur (Girshick, 2015) . Faster R-CNN’de Seçici Arama algoritması çıkartılmış ve bölge önerisi aşaması yeni önerilen Bölge Öneri Ağı (Region Proposal Network) üzerinden gerçekleştirilmiştir (Ren vd., 2015). Bölge önerileri artık algılama görevine göre eğitilebilen ve özelleştirilebilen bir ağ kullanılarak oluşturulduğundan uçtan uca eğitim yapılabilir. Böylece Selective Search ve EdgeBoxes gibi yöntemlere göre daha iyi bölge önerileri üretilebilmektedir. Faster R-CNN, Fast R-CNN ve Bölge Öneri Ağı’nın birleşiminden oluşturulmuştur. Bölge Öneri Ağı’nda, R-CNN ağı’nda kullanılan aynı evrişimli katmanlar kullanılarak görüntü işlendiğinden Seçici Arama gibi algoritmalarla kıyasla teklifleri üretmek için fazladan zaman harcanmamaktadır. Bu durum Fast R-CNN’deki yaklaşık 2 sn olan nesne tespit süresini yaklaşık 0.2 sn’ye indirmiştir.

Mask R-CNN, Faster R-CNN’in geliştirilmesi ile üretilmiştir. Faster R-CNN’nin her aday nesne için iki çıkışı varken (bir sınıf etiketi ve bir sınırlayıcı kutu) Mask R-CNN’de buna nesne maskesini çıkaran üçüncü bir kısım eklenmiştir. Üçüncü kısım her bir belirlenen nesneye maske oluşturmak için tam bağlantılı katman (Fully Convolutional Network, FCN) uygulanmaktadır. Bu sayede her bir piksel küçük hesaplama maliyetleri ile sınıflandırılmıştır. Mask R-CNN’in Faster R-CNN mimarisinden bir diğer farkı da özellik çıkartma katmanıdır. Bu katman evrişim katmanından geçtikten sonra uygulanmaktadır. Faster R-CNN’de RoIPool katmanı kullanırken Mask R-CNN’de RoIAlign katmanı kullanılmaktadır. RoIPool, maksimum havuzlama yöntemini kullanarak her bir ilgi bölgesindeki (Region of Interest, RoI) özellikleri küçük bir özellik haritasına dönüştürmektedir. Bu yöntem yuvarlama yaparak tam sayı aldığından büyük bir bilgi kaybına neden olmaktadır. RoIAlign, her bir ilgi bölgesini dört küçük bölgeye bölerek çift doğrusal enterpolasyon yöntemini kullanmaktadır. Bu sayede RoIPool’un neden olduğu yanlış hizalamayı önlemektedir (Potlapally vd., 2019). Eğitimin kolay ve esnek olması, genelleme yeteneğinin yüksek olması ve yüksek doğrulukta sınıflandırma yapabilme yeteneğine sahip olması Mask R-CNN yönteminin önemli özelliklerini oluşturmaktadır. Mask R-CNN literatürde tıp, biyoloji ve hayvancılık gibi alanlarda kullanıldığı gibi uzaktan algılama alanında da kullanılmaya başlamıştır.

Carvalho vd., (2020) tarafında yapılan çalışmada, Landsat 8 uydu görüntülerini kullanarak Brezilya bölgesindeki merkezi sulama sistemlerinin (Center Pivot Irrigation System, CPIS) otomatik olarak tespit edilmesi amaçlamıştır. CPIS, farklı dokulara, renklere ve spektral davranışlara sahip yuvarlak şekillerdir. Bu doğrultuda ilk olarak manuel yöntemler uygulanmıştır. Sonrasında otomatik olarak tespit etmek amacıyla örnek segmentasyon (instance segmentation) yöntemi olan Mask R-CNN kullanılmıştır. Mask R-CNN’de yedi farklı mimari yapı (ResNet50- Feature Pyramid Network (FPN), ResNet50-DC5, ResNet50-C4, ResNet101-FPN, ResNet101-DC5, ResNet101-FPN ve ResNeXt101) üzerinde çalışılmıştır. Modelleri değerlendirirken standart MS COCO metrikleri (ortalama hassasiyet AP, AP50, AP75) kullanılmıştır. Çalışma sonucunda ResNeXt101 mimari yapısının en iyi sonuçları verdiği tespit edilmiştir (Carvalho vd., 2020). Bhuiyan vd., (2020) tarafından yapılan çalışmada, farklı bölgelerden alınmış Worldview-2 uydu görüntüleri kullanarak tundra bölgelerindeki buz kamaları (ice-wedge, IWPs) Mask R-CNN ile tespit edilmiştir. Değişken tundra bölgelerinin tespitinde %89 - %96 doğruluk elde edilirken, sınıflandırma için %94 - %97 doğruluk elde etmiştir (Bhuiyan vd., 2020). Nie vd., (2018) tarafından yapılan çalışmada; kıyılarındaki gemilerin tespiti için Mask R-CNN yöntemine dayanan bir yöntem önermiştir. Bu yöntemde Mask R-CNN algoritmasına Soft-Non-Maximum Suppression (Soft-NMS) yöntemi eklenmiştir. Gerekli olan veri seti Google Earth görüntüleri kullanarak 1 m uzamsal çözünürlükteki kıyı bölgelerinden oluşturulmuştur. Bu veri seti Mask R-CNN ve Mask R-CNN Soft-NMS üzerinde çalıştırılmıştır. Modellerin doğruluklarının değerlendirilmesi aşamasında ortalama kesinlik puanı (AP) Mask R-CNN için 0.845, Soft-NMS için 0.852 olarak hesaplanmıştır. Sonuçlar incelendiğinde gemilerin çok yakın olduğu kısımlarda Soft- NMS sayesinde daha doğru sonuçlar elde edildiği tespit edilmiştir (Nie vd., 2018).

## 1.2. Çalışmanın Amacı

Uydu görüntüleri, geniş coğrafyaya sahip olması ve yeryüzü hakkında detaylı bilgi sağlamasından dolayı en önemli veri kaynaklarından biridir. Uydu görüntülerinin sınıflandırılması amacıyla araştırmacılar ve uygulayıcılar birçok yöntem denemiştir. Günümüzde büyük ölçüde artan uydu görüntüsü sayısı, spektral, konumsal, zamansal çözünürlüklerde yaşanan önemli artışlar uydu görüntülerinin işlenmesi aşamasında çeşitli zorluklar oluşturmaktadır. Örneğin piksel tabanlı bir sınıflandırıcı ile 5 m ve 30 cm uzamsal

özünürlüklü bir görüntüde binaların sınıflandırılması farklı sonuçlar ortaya çıkarabilmektedir. 30 cm'lik görüntüde bina çatısı 5 m'lik görüntüye göre çok daha fazla piksel ile temsil edileceğinden bina çatısı üzerindeki farklı renk tonları çatı üzerindeki piksellerin diğer sınıflar ile karışmasına neden olabilecektir. Bu tür sorunların daha aza indirilmesi amacıyla uzaktan algılama alanında nesne tabanlı sınıflandırma yöntemleri tercih edilmektedir. Nesne tabanlı sınıflandırma, görüntü segmentasyonu adı verilen ve homojen nesneler elde etmek için benzer pikselleri bir araya toplama işlemiyle başlamakta ve sonrasında elde edilen bu homojen nesne gruplarının sınıflandırılması ile tamamlanmaktadır. Nesne tabanlı sınıflandırma piksel tabanlı sınıflandırmaya göre genel olarak daha iyi sonuçlar vermesine rağmen sınıflandırma kalitesi görüntü segmentasyonuna bağlı olmaktadır. Görüntü segmentasyonu aşaması ise çok fazla parametreye bağlı olmakta ve görüntüden görüntüye bu parametreler değişmektedir. Derin öğrenme tabanlı yöntemlerde ise görüntü üzerinden yüksek seviye özellikler hiyerarşik olarak kullanıcı müdahalesi olmadan öğrenilmektedir. Pek çok alanda başarıyla uygulanan evrişimli sinir ağı tabanlı derin öğrenme yöntemleri geleneksel makine öğrenimi sınıflandırıcılarının (örneğin, RF ve SVM) yanı sıra nesne tabanlı görüntü analizi gibi daha yeni uygulamalardan da daha iyi performans göstermektedir (Bhuiyan vd., 2020).

Derin Öğrenme teknolojisindeki daha fazla ilerleme, nesnelerin tespit edilme sürecinin otomatikleştirilmesi ve anlık nesne tespitine imkân sağlaması üzerine gelişmiştir. Görüntü üzerindeki nesnelerin etraflarına bir sınırlayıcı kutu yerleştirilerek yapılan bu işlem R-CNN, Fast R-CNN, Faster R-CNN, SSD, YOLO gibi derin ağlar ile gerçekleştirilebilmiştir. Sınıf etiketiyle birlikte bir sınırlayıcı kutu, nesnelerin konumları ile birlikte tespit edilmesine imkân sağlamasına rağmen, her bir nesneyle ilişkili pikseller belirlenmemektedir. Faster R-CNN'in geliştirilmesi ile üretilen Mask R-CNN ile örnek segmentasyon mümkün olarak tüm görüntü üzerinde her bir nesneye ait piksellerde elde edilebilmiştir.

Bu çalışmanın amacı, bilgisayar görü alanında başarı ile uygulanan Mask R-CNN' i uzaktan algılanmış görüntüler üzerinde uygulamaktır.

Mask R-CNN'in uydu görüntüleri üzerinde sınıflandırmada kullanılması aşağıdaki avantajları:

- Eğitim süresi uzun olmasına rağmen tek bir eğitim ile farklı bölgelere ait görüntüler üzerinde sınıflandırma yapma imkanına sahip olunabilir. Bu, her bir görüntü için

yeniden eğitim yapma, eğitim verisi toplama gibi zorlu süreçleri ortadan kaldırmaktadır.

- Saniyeler içerisinde uydu görüntülerini sınıflandırma imkanına sahip olunabilir.
- Mask R-CNN'in örnek segmentasyon anlayışı renk, doku ve desen çeşitlerini göz önünde bulundurarak bir sınıflandırma yaptığından uzaktan algılama alanında çokça kullanılan sınıflandırma yöntemlerine göre yüksek doğruluk sağlayabilir.
- Örnek segmentasyon ile her bir sınıfa ait nesneler de ayrı ayrı tespit edilebilir. Örneğin bina sınıf içerisindeki her bir bina ayrı bir ID ile temsil edilebilir.

Mask R-CNN'in bu avantajlarından uydu görüntüleri üzerinde nasıl yararlanabileceğimizi test etmek için bu çalışmada 2 farklı uygulama gerçekleştirilmiştir. İlk uygulamada Türkiye'yi kapsan çeşitli yerlerden elde edilen büyük miktarda Sentinel-2 verileri ile Mask R-CNN ağının eğitimi ile bölge tespiti çalışması yapılmıştır. İkinci uygulamada ise Mask R-CNN' in çok yüksek uzamsal çözünürlüklü uydu görüntüsü üzerinden bina tespitinde kullanımı incelenmiştir.

### 1.3. Uzaktan Algılamada Sınıflandırma

Uzaktan algılama, bir nesne hakkında fiziksel temas olmadan bilgi edinme olarak tanımlanmaktadır. Bu bilgi, nesnenin çevresine uyguladığı elektromanyetik veya potansiyel değişiklikleri ölçerek tespit edilir. Uyduların ortaya çıkışı, Dünya ve diğer gezenlerin çevreleri hakkında küresel ve ayrıntılı bilgi edinilmesine olanak sağlamıştır. Dünya yörüngesinde bulunan uydulardaki sensörler; bulutların durumu, bitki örtüsü, mevsimler, değişimler, yüzeydeki morfolojik yapılar, okyanus yüzeyi sıcaklığı, kentsel yapılaşmayı inceleme gibi birçok konuda bilgi edinimine imkân sağlamaktadır (Elachi ve Van, 2006).

Uzaktan algılamada sınıflandırma işlemi karmaşık ve birçok faktörün dikkate alınmasının gerektiği bir süreçtir. Görüntü sınıflandırmanın temel adımları; uygun bir sınıflandırma algoritmasının belirlenmesi, eğitim örneklerinin seçimi, görüntüde ön işleme, özellik çıkarma, uygun sınıflandırma yaklaşımlarının seçilmesi ve sınıflandırma sonrası doğruluk değerlendirmesi gibi işlemlerden oluşmaktadır. Ayrıca kullanıcının ihtiyacı, çalışma alanının ölçeği, ekonomik durum, uzaktan algılanan verilerin seçimi, sınıflandırma algoritmasının tasarımı ve sınıflandırma sonuçlarının kalitesi sınıflandırma için önemli faktörlerdendir.

Son yıllarda teknolojinin gelişmesiyle beraber sınıflandırma, segmentasyon, gürültüden arındırma, nesne tespiti gibi alanlarda önemli gelişmeler gerçekleşmiştir. (Castelluccio vd., 2015).

Görüntü sınıflandırılmasının amacı bir görüntüdeki verilerinin özelliklerinin tanımlanmasını otomatikleştirmek ve görüntü verilerinin görsel analizinin yapılmasını sağlamaktır. Bu işlem aslında çok bantlı görüntü verilerinin analizini ve görüntüdeki her pikselin arazi örtüsü kimliğini belirlemek için istatistiksel karar kurallarının uygulanması sürecini içermektedir. Bu karar kuralları yalnızca spektral değerler üzerinden yapıldığında, bu sınıflandırma süreci piksel tabanlı sınıflandırma olarak adlandırılmaktadır (Lillesand vd., 2004). Karar kuralları görüntü üzerindeki geometrik şekillere, boyutlara ve modellere göre belirlendiğinde ise bu sınıflandırma süreci nesne tabanlı sınıflandırma olarak adlandırılmaktadır (Lillesand vd., 2004). Her iki durumda da sınıflandırma sürecinin amacı, dijital görüntüdeki tüm pikselleri çeşitli arazi örtüsü sınıfları olarak etiketlemektir. Piksel tabanlı sınıflandırma yöntemi; denetimli (kontrollü) sınıflandırma ve denetimsiz (kontrolsüz) sınıflandırma olarak alt başlıklarda incelenebilmektedir (Lillesand vd., 2004).

Denetimsiz sınıflandırma yönteminde sınıflandırılacak pikseller eğitim verileri kullanılarak sınıflandırılmamaktadır. Bunun yerine pikseller benzer spektral özelliklerine göre otomatik sınıflandırılmaktadır. Temel olarak, belirli bir sınıf içindeki piksel değerlerinin birbirine yakın olması gerekirken, farklı sınıflardaki piksel değerlerinin ayrı olması beklenmektedir. Oluşturulan spektral sınıflar yalnızca görüntü değerlerindeki doğal gruplamalara dayandıklarından, spektral sınıfların kimliği başlangıçta bilinmemektedir. Görüntü analisti, spektral sınıfların kimliğini ve bilgi değerini belirlemek için sınıflandırılmış verileri bazı referans verileriyle (daha büyük ölçekli görüntüler veya haritalar gibi) karşılaştırmalıdır. Genellikle denetimsiz sınıflandırma yöntemleri, denetimli sınıflandırma için güvenilir eğitim verilerinin elde edilmediği veya eğitim verilerinin oluşturulmasının çok zahmetli olduğu durumlarda kullanılmaktadır (Richards ve Jia, 2006).

Denetimli sınıflandırma, uzaktan algılama verilerinin analizi için en sık kullanılan yöntemlerden biridir. Denetimli sınıflandırma yapılırken ilk olarak görüntü analisti tarafından sınıflara karar verilerek her sınıf için örnek pikseller seçilmektedir. Görüntü analisti ayrıca pikselleri gruplarken diğer piksellere ne kadar benzer olması gerektiğine ilişkin sınırları belirlemektedir. Bu sınırlar genellikle eğitim alanının spektral özelliklerine göre belirlenir. Seçilen bu pikseller eğitim veri setini oluşturmaktadır. Eğitim veri seti referans alınarak tüm piksellerin sınıflandırması yapılmaktadır. Son olarak, etiketlenmiş

başka bir veri seti kullanılarak sınıflandırılmış görüntünün doğruluğu değerlendirilir (Knights vd., 2011) .

Son birkaç yılda derin öğrenme teknikleri uzaktan algılama alanında bilgisayarla görme ve nesne algılama problemlerinde sıklıkla kullanılan bir yöntem haline gelmiştir. DÖ tekniği olan ESA, orijinal görüntülerden çok seviyeli özellikleri çıkarmak için kullanılmaktadır. Bu durum uzaktan algılanmış görüntülerdeki karmaşık geometri, doku ve spektral özellikleri ayırt etmek için büyük bir avantaj sağlamaktadır. Buna dayanarak çoğu araştırmacı yüksek çözünürlüklü uzaktan algılama görüntülerinde ESA tabanlı çalışmalar kullanmıştır. Yapılan çalışmalardan bazılarına pan-keskinleştirme (pan-sharpening), değişim analizi (change detection), nesne tespiti, semantik segmentasyon, zaman serisi analizi ve otomatik bina çıkarımı örnek olarak verilebilir.

Nesne tespiti ile görüntüdeki nesnelerin konumları ve boyutları tanımlanmaktadır. Bu işlem yapılırken her nesne için bir sınırlayıcı kutu belirlenmekte ve nesne için sınıf etiketi tahmin edilmektedir. Önceleri nesne tespiti çalışmaları kenar algılama ve özellik çıkarma gibi teknikler ile yapılırken, günümüzde derin öğrenme yöntemleri de nesne tespitinde başarılı bir şekilde kullanılmaktadır (Fritz, 2020).

DÖ' de segmentasyon; semantik segmentasyon ve örnek segmentasyon olarak iki alt guruba ayrılmaktadır. Semantik segmentasyon klasik görüntü sınıflandırmasından farklı olarak her nesne parçasını oluşturan pikselleri semantik olarak gruplandırarak tanımlamaktadır. Bu işlem yapılırken konum, şekil ve sınıf bilgileri kullanılmaktadır (Carvalho vd., 2020). Semantik segmentasyon işleminde birbirine yakın olan aynı sınıfa ait nesneler tek bir sınıf altında toplanmaktadır. Bu durum karmaşık görüntülerdeki nesneleri ayırt etmede ve tek bina çıkarımına odaklanan uygulamalarda yetersiz kalmaktadır. Bu sorunu ortadan kaldırmak için örnek segmentasyon önerilmektedir (Carvalho vd., 2020). Örnek segmentasyon, tek tek binaları tanımlamakla kalmaz aynı zamanda her bir sınıflandırılacak pikselin semantik olarak sınıflandırılmasını sağlamaktadır. Başka bir deyişle örnek segmentasyon, sınıflandırılacak nesne için hem sınırlayıcı kutu oluşturmakta hem de nesneyi kapsayan bir maske üretmektedir.

#### 1.4. Yapay Sinir Ağları

İnsan beyninin örüntü tanıma problemlerindeki başarısı, araştırmacıları beynin basitleştirilmiş bir modeline dayalı sistemlerin standart istatistiksel sınıflandırma yöntemlerinden daha etkili olup olamayacağını düşünmeye yöneltmiştir. Bu tür araştırmalar, son on yılda uzaktan algılamada, özellikle sınıflandırma için giderek daha fazla kullanılan yapay sinir ağlarının (YSA) geliştirilmesine yol açmıştır (Wang vd., 2017).

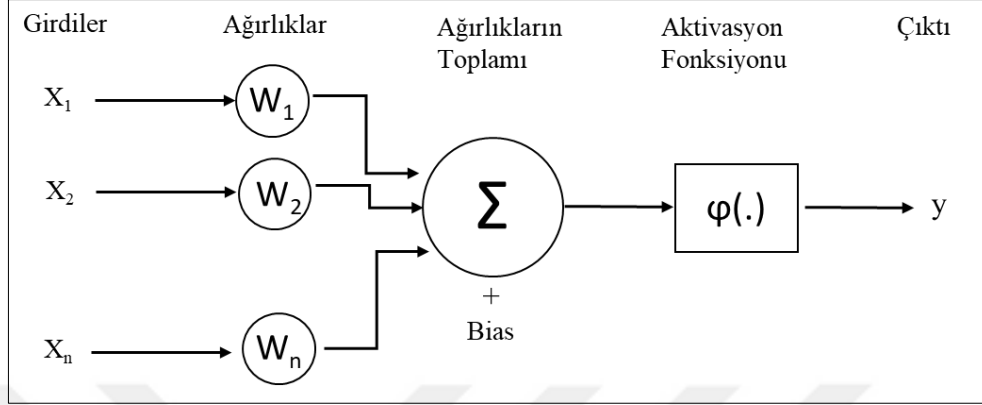
Yapay sinir ağları parametrik olmayan sınıflandırma grubu içine girmektedir. Bir sinir ağının kullanılması, verilerin istatistiksel dağılımı hakkında herhangi bir varsayım gerektirmemektedir. Bir sinir ağının performansı verilerin istatistiksel dağılımı ile ilgili varsayımların yeterliliğine değil, verilerin ne kadar iyi eğitildiğine bağlı olmaktadır. Eğitim aşamasında, sinir ağı eğitim verilerinde bulunan düzenlilikleri “öğrenir” ve bu düzenliliklere dayanarak bilinmeyen verilere uygulanabilecek kurallar oluşturmaktadır. Bu sistem sinir ağlarının özel yeteneğidir. Kullanıcı eğitim mimarisini, eğitim süresini ve bir sinir ağının öğrenme oranını etkileyen parametreleri belirlemelidir. Ağın tasarımına yardımcı olacak kesin kurallar yoktur (Mather ve Tso, 2009).

Yapay sinir ağları, geleneksel regresyon ve istatistiksel modellere göre daha başarılı sonuçlar vermektedir (Dave ve Dutta, 2014). Günümüzde yapay sinir ağları uygulamaları insan ihtiyaçlarını çeşitli alanlarda karşılayabildiği için popüler hale gelmiştir. Pek çok kuruluş ekonomik ve kendi alanlarındaki sorunları çözebilmek için sinir ağlarına yatırım yapmaktadır (Boyacioglu vd., 2009). Yapay sinir ağları hepatit teşhisi, konuşma tanıma, üç boyutlu nesne tanıma, doku analizi, yüz tanıma, deniz altı mayın tespiti gibi birçok alanda kullanılmaktadır (Abiodun vd., 2018). Yapay sinir ağları ayrıca optimizasyon (Haykin, 2004), saldırı tespiti (Fan vd., 2012) ve veri sınıflandırmasında (Saravanan ve Sasithra, 2014) kullanılmaktadır. Çoğu araştırmacı sınıflandırma problemi için makine öğrenimi tekniklerini kullanmaktadır (Das ve Behera, 2017).

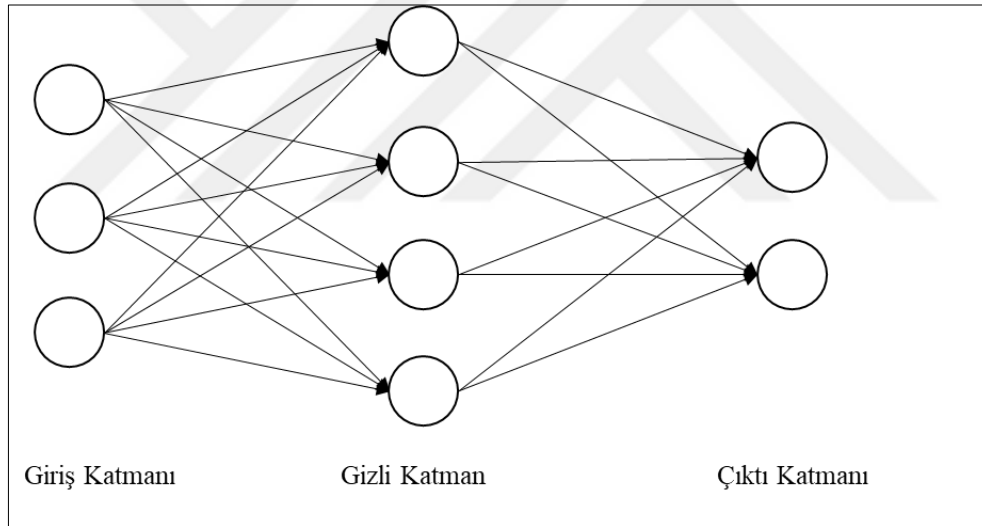
Biyolojik sinir ağlarından esinlenen yapay sinir ağları, nöronlar aracılığıyla işlemlerini gerçekleştirmektedir. Nöron, sinir ağının çalışması için temel olan bir bilgi işlem birimidir. Bir nöronda birden çok giriş katmanı olabilirken tek çıktı katmanı olabilmektedir.

Şekil 1’de bir yapay sinir hücresi gösterilmektedir. Girdiler, bir sinir ağına gelen verilerdir. Yapay sinir hücresine gelen verilerin her biri kendine ait ağırlıkla çarpılır. Toplama fonksiyonu ile ağırlıklarla çarpılmış olan veriler toplanır. Yapay sinir hücresinin çıktı değerini sınırlandırmak için aktivasyon fonksiyonu uygulanır. Bu fonksiyon  $[0, 1]$  veya

$[-1, 1]$  aralığında olmaktadır. Bunun sonucunda çıktı elde edilir (Csáji, 2001). Yapay sinir hücreleri yapay sinir ağlarını oluşturur (şekil 2).



Şekil 1. Yapay sinir hücresi



Şekil 2. Yapay sinir ağı

Yapay sinir ağları, belirli bir mimari yapıya göre ağırlıklı bağlantılarla bağlanan, düğüm adı verilen çok sayıda basit işlem birimi içermektedir. Temel olarak yapay sinir ağı modeli girdi katmanı, gizli katman ve çıktı katmanından oluşmaktadır. Öğrenme işlemi, çıkış düğümü ile çıktı arasındaki fark en aza indirilerek düğümlerdeki ağırlıkların ayarlanması ile gerçekleştirilmektedir. Yapay sinir ağlarının uygulanabilmesi için eğitim hızı, düğümler üzerindeki ağırlıklar, çıkış düğümlerinin sayısı, eğitim momentumu, çıkış kriterleri gibi

parametrelerin belirlenmesi gerekmektedir (Petroopoulos, 2010; Buduma, 2016). Aşağıda yapay sinir ağlarının avantajları ve dezavantajları listelenmektedir (Borra, 2019)

Yapay sinir ağlarının avantajları,

- Makine öğrenmesi yapabilir.
- Daha az eğitim verisi ile denetimli sınıflandırma yapabilir.
- Karmaşık bir alanda bile örüntüleri öğrenebilir.
- Hataları tolerans edebilir.
- Çeşitli veri türlerine uyarlanabilir.
- Girdiyi genelleştirebilir.
- Kendi kendine öğrenebilir.
- Öğrenme işlemi ağda saklanır.

Yapay sinir ağlarının dezavantajları,

- Bilgisayarın donanımına bağlıdır.
- Performansı yapay sinir ağının mimarisine bağlıdır.
- Çoklu ve karmaşık parametreler içerir.

### 1.5. Derin Öğrenme

Bir gizli katmana sahip yapay sinir ağlarının üzerinde çalışmalara 1960 yılında başlanmıştır (örneğin, Joseph, 1961; Viglione, 1970). 1965 yılında Ivakhnenko ve Lapa tarafından yapılan çalışmada Grup Veri İşleme Yöntemi (Group Method of Data Handling (GMDH)) kullanılmıştır. Bu yöntem ile eğitilen ağlar, çok katmanlı sinir ağlarından oluştuğu için ilk derin öğrenme algoritması olarak kabul edilmektedir (Ivakhnenko vd., 1965). GMDH yönteminde ağı eğitmek için en küçük kareler yöntemi kullanılmıştır (Ivakhnenko vd., 1965). 1971 yılında Ivakhnenko tarafından sekiz gizli katmandan oluşan derin sinir ağı tanıtılmıştır (Ivakhnenko, 1971). 1979 yılına gelindiğinde Fukushima tarafından ilk kez evrişimli sinir ağı kullanılmıştır. Fukushima, çoklu havuzlama katmanına ve evrişim katmanına sahip ağ tasarlayarak bunu 'Neocognitron' olarak adlandırmıştır. Bu tasarım ile bilgisayarın görüntüler üzerindeki örüntüleri öğrenmesi sağlanmıştır (Fukushima, 1979). 1989 yılında LeCun tarafından yapılan çalışmada geri yayılım evrişimli sinir ağları ilk kez kullanılmıştır. Ayrıca bu çalışmada makine öğrenimi alanında kullanılan en önemli veri

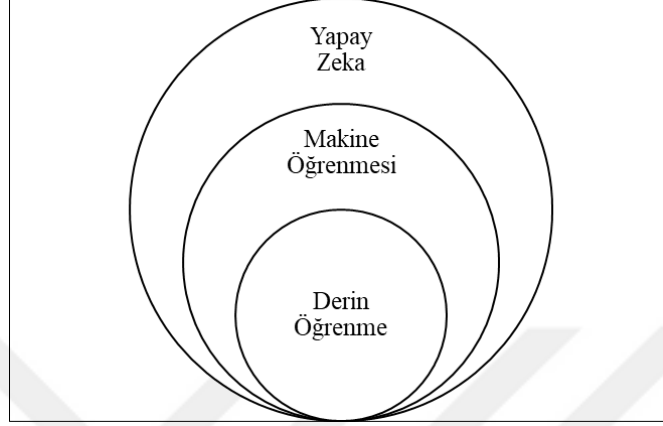
setlerinden biri olan MNIST veri seti tanıtılmıştır. MNIST, farklı insanlar tarafından el yazısı ile yazılmış rakamların resimlerinden oluşmaktadır. Geri yayılımın evrişimli sinir ağlarında kullanılmasıyla MNIST veri seti üzerinde iyi sonuçlar elde edilmiştir (LeCun vd., 1989). 1992 yılında Schmidhuber tarafından tekrarlayan sinir ağları (Recurrent Neural Network, RNN) önerilmiştir (Schmidhuber, 1992). 1997 yılında Hochreiter ve Schmidhuber tarafından uzun/kısa süreli hafıza (Long Short-Term Memory, LSTM) önerilmiştir (Hochreiter vd, 1997).

Derin öğrenme için bir diğer önemli adım çok işlemcili grafik kartlarının (GPU) bilgisayarlarda veri işleme için kullanılmasıyla 1999 yılında atılmıştır. GPU'lar yapay sinir ağlarındaki öğrenme işlemi için gerekli olan matris ve vektör çarpımlarını 50 kat daha hızlı yapabilmektedir. Bunun yanında GPU'lar 10 yıllık bir süre içinde hesaplama hızlarını 1000 kat daha artırmıştır. Bu süre zarfında yapay sinir ağları destek vektör makineleri ile rekabet etmeye başlamıştır. Bir sinir ağı destek vektör makinesine kıyasla daha yavaş olabilirken, aynı veri setinde daha iyi sonuçlar elde etmiştir (URL-1, 2022). 2000'li yıllarda derin sinir ağlarında kaybolan gradyan problemi görülmüştür. Bu problem alt katmandaki özelliklerin (features) üst katmanlara iletilmemesinden kaynaklanmıştır. Bu sorunu çözmek için LSTM geliştirilmiştir.

2010 yılında hızlanan GPU'lar sayesinde evrişimli sinir ağları ön-eğitim (pre-training) olmadan eğitilmiştir (Ciresan vd., 2010). 2012 yılında Krizhevsky ve arkadaşları, bilgisayarla görme topluluğunda popüler bir yarışma olan ILSVRC (ImageNet Large Scale Visual Recognition Challenge)'de Alex Krizhevsky'in adını taşıyan AlexNet evrişimli ağ mimarisini uygulamıştır. Bu ağ yarışmada en iyi sonuçları elde etmiştir (Krizhevsky vd., 2012). 2012'nin sonlarına doğru Ciresan ve arkadaşları AlexNet evrişimli sinir ağı mimarisinde ezberlemeyi önlemek için 'dropout' kullanarak ağı iyileştirmiştir (Cireşan vd., 2013). Böylelikle ilk kez 2012 yılında derin öğrenme algoritmaları kullanılarak bir yarışma kazanılmıştır. Derin öğrenmenin ImageNet yarışmasındaki başarılarından sonra Facebook, Amazon, Microsoft, Google ve Apple gibi teknoloji şirketleri son 10 yılda yapay zeka alanına yatırım yapmıştır.

Derin öğrenme, Şekil 3'te görüldüğü gibi hiyerarşik olarak yapay zekâ ve makine öğrenmenin altındadır. Derin öğrenme, her bir seviyedeki temsilleri veya özellikleri (ham girdiden başlayarak) daha yüksek seviyedeki temsillerine dönüştüren doğrusal olmayan işlem birimlerinden oluşmaktadır. Bu nedenle derin öğrenme yaklaşımı düşük seviyeli özelliklerden çoklu temsil seviyelerin öğrenilmesiyle sonuçlanan bir öğrenme yöntemidir

(Zhang vd., 2017). Başka bir deyişle derin öğrenme insan müdahalelerine bağımlı olmadan genelleme yeteneklerini geliştirebilmekte ve üst düzey özellik temsillerini modeller aracılığıyla kendileri öğrenebilmektedir (Lin vd., 2020).

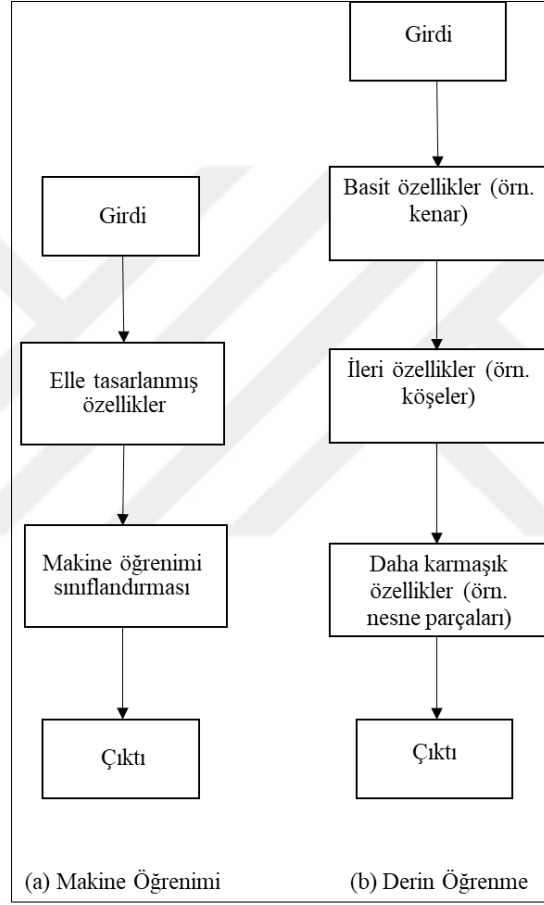


Şekil 3. Yapay zekâ hiyerarşisi

Derin öğrenme ve makine öğrenmesinin farklı olduğu nokta algoritmaların öğrenme şeklidir. Derin öğrenme, insan müdahalesine gerek kalmadan verilerin temsil değerlerinden öğrenebilmektedir. Klasik makine öğrenmesi veya derin olmayan makine öğrenmesi, öğrenme işlemi için insan müdahalelerine bağımlıdır. Verileri doğru şekilde işlemek, algoritmaları geliştirmek gibi durumlar için bilim adamları ve analistler kendi bilgilerini kullanmaktadır. Örneğin, duvar ölçülerini içeren verilere dayanarak bir evin yüzey alanının belirlenmesini isteyen bir problemde; makine öğrenimi algoritması, analist nasıl hesaplama yapılacağını belirtmediği sürece evin yüzeyini hesaplayamaz. Derin öğrenmede ise analistlerin herhangi bir tanımlama yapmasına ihtiyaç duyulmamaktadır. Derin sinir ağı yapısındaki birçok katman sayesinde kendisi en iyi temsilleri oluşturmaktadır. Bu sayede derin öğrenme, ses ve görüntü tanıma, metin anlama gibi çok zor olan görevlerde klasik makine öğrenimine göre daha iyi performans göstermektedir (Mueller ve Massaron, 2019).

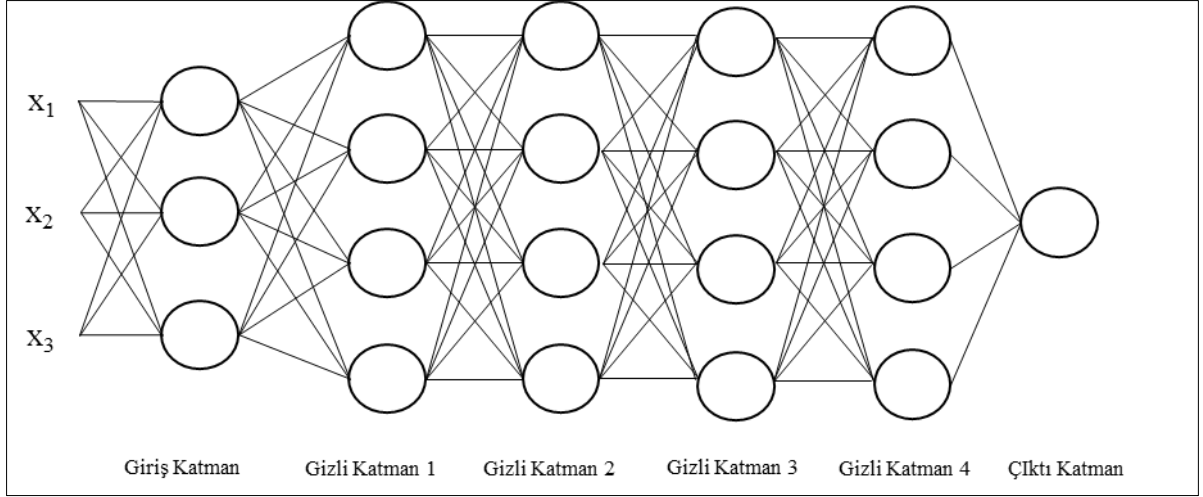
Derin öğrenme ile bir görüntü analiz edilirken gizli katmanlar giriş görüntüsünden özelliklerin çıkartılmasını sağlamaktadır. Bu gizli katmanlar hiyerarşik bir şekilde birbirlerinin üzerine inşa edilmiştir. İlk katmanlar kenarları algılamaktadır. Bu kenarlar daha sonra köşeleri ve konturları tanımlamak için kullanılmaktadır. Sonraki katmanlar, nesne parçalarına ulaşmak için köşeleri ve konturları birleştirmektedir. Daha önce de belirtildiği gibi derin öğrenmede bu özellik katmanları analistler tarafından tasarlanmamaktadır.

Öğrenme prosedürü kullanılarak verilerden yavaş yavaş öğrenilmektedir. Çıktı katmanında, görüntü sınıflandırılarak çıktı sınıfının etiketi atanmaktadır. Elde edilen çıktı, ağda bulunan diğer tüm düğümlerden doğrudan etkilenmektedir. Ağın içerisindeki her katman, daha önceki katmanların çıktısını kullandığından bu süreç hiyerarşik öğrenme olarak görülmektedir. Şekil 4'te elle tasarlanmış özelliklere dayalı makine öğrenmesi ile hiyerarşik temellere dayalı derin öğrenme yapısı gösterilmiştir (Wani vd., 2020).



Şekil 4. (a) Makine öğrenimi, (b) Derin öğrenme

Derin öğrenme birden fazla gizli katman içeren yapay sinir ağları ve makine öğrenimi algoritmalarından oluşmaktadır. Bu derin ağlar: özellik çıkarma ve dönüştürme işlemleri için doğrusal olmayan katmanlardan oluşmaktadır. Birbirini takip eden her katman bir önceki katmanın çıktısını girdi olarak kullanmaktadır. Şekil 5'te dört gizli katmandan oluşan derin sinir ağ yapısı gösterilmektedir.



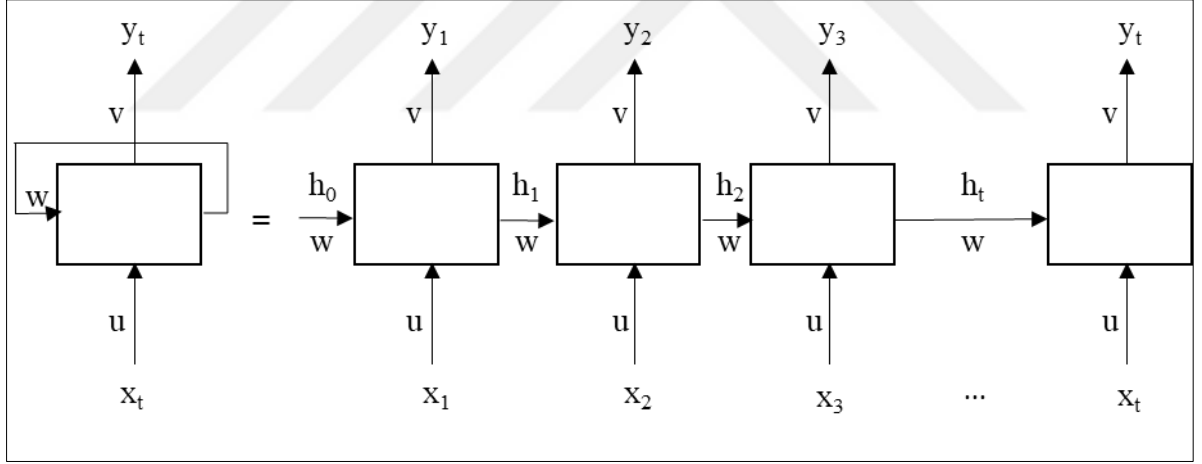
Şekil 5. Derin sinir ağı yapısı

Derin sinir ağlarından oluşan Tekrarlayan Sinir Ağları (Recurrent Neural Network (RNN)), Uzun Kısa Vadeli Sinir Ağları (Long Short-Term Memory (LSTM)), Kısıtlı Bozaltmanın Makineleri (Restricted Boltzmann Machines (RBM)), Derin İnanç Ağları (Deep Belief Network (DBN)), Evrişimli Sinir Ağları gibi derin öğrenme mimarileri; bilgisayarlı görme, doğal dil işleme ve ses tanıma alanlarında kullanılarak son teknolojik sonuçlar elde edilmiştir. Ayrıca derin sinir ağları nesne tanıma gibi görevlerde diğer makine öğrenimi algoritmalarından daha iyi sonuçlar elde etmiştir (Krizhevsky vd., 2012; Ongsulee, 2017).

### 1.5.1. Tekrarlayan Sinir Ağları (RNN)

Sıralı veri tahmini yapma birçok araştırmacı tarafından makine öğrenimi için bir sorun olarak kabul edilmektedir (Örneğin, Mahony, 1999). İstatiksel dil modellemesinin amacı, sıralı şekilde verilen metin içerisindeki verilerden bir sonraki sözcüğü tahmin edebilmektir. Bu nedenle dil modellerinin tahmini yapılırken sıralı verileri tahmin etme problemine çözüm bulmaya çalışılmaktadır. Bengio tarafından ileri beslemeli ağların istatiksel dil modellemesinde kullanılabileceği önerilmiştir (Bengio, 2001). Schwenk ise sinir ağı tabanlı modellerin konuşma tanıma önemli gelişmeler sağladığını göstermiştir (Schwenk, 2005). Tekrarlayan Sinir Ağları, konuşma tanıma, dil modelleme, duyarlılık analizi, altyazı tanımlama gibi problemleri çözmek için kullanılmaktadır (Gulli ve Pal, 2017).

Tekrarlayan Sinir Ağları, girdi verilerini sıralı olarak işleyebilmek için tekrarlayan hesaplamalar kullanan döngüsel bağlantılardan yararlanmaktadır (Min vd., 2017). Bu girdiler; konuşma, metin, zaman serileri ve dizideki bir ögenin oluşumunun kendisinden önceki öğelere bağı olan herhangi bir sıralı yapısı olabilmektedir. RNN, hücreleri öğrendikleri bilgileri içerisinde tutan gizli bir durum vektörüne (hidden state), veya belleğe sahip olarak işlemlerini gerçekleştirmektedir. Gizli durum vektörü; zamanın herhangi bir noktasındaki değerin, önceki zaman durumundaki değerin ve mevcut zaman içindeki değerin bir fonksiyonu olmaktadır. Şekil 6' RNN yapısı verilmektedir. Bu yapıda,  $t$  zamanındaki hücrenin  $x_t$  girdisi,  $y_t$  çıktısı vardır. Çıktının bir kısmı  $y_t$  (gizli durum vektörü  $h_t$ ) daha sonraki adım olan  $t+1$  de kullanılmak için hücreye geri beslenir. RNN'de giriş, çıkış ve gizli duruma karşılık gelen üç ağırlık vektörü;  $u$ ,  $v$  ve  $w$  tarafından oluşmaktadır. Ağırlık vektörleri her adımda farklı girdilere uygulanmaktadır. Bu ağırlık vektörlerinin tüm zaman diliminde paylaşılabilmesi RNN'in öğrenmesi gereken parametre sayısını azaltmaktadır (Gulli ve Pal, 2017).



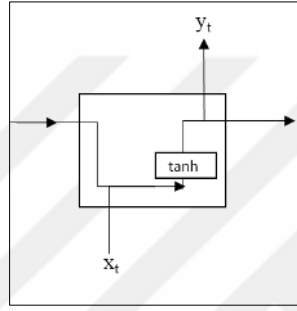
Şekil 6. Tekrarlayan Sinir Ağı yapısı

RNN'de, tekrarlayan bağlantılar kullandıklarından bilgiler ağı içinde uzun süre dolaşabilmektedir (Mikolov, 2010). Bununla birlikte stokastik gradyan inişiyle uzun süreli bağımlılıkları öğrenmenin zor olacağı sıklıkla iddia edilmektedir (Bengio, 1994)

### 1.5.2. Uzun Kısa Vadeli Sinir Ağları (LSTM)

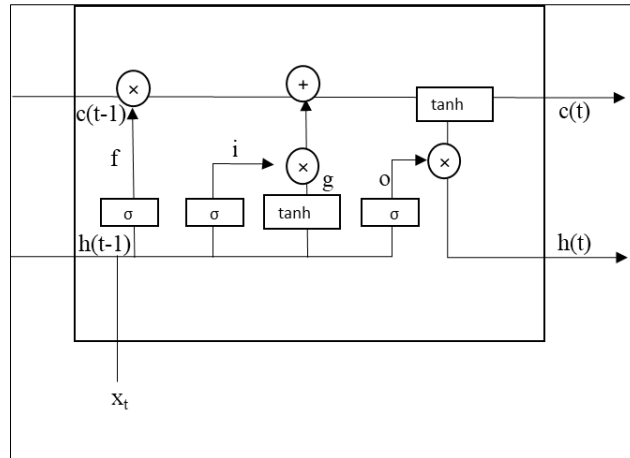
Uzun Kısa Vadeli Sinir Ağları, uzun vadeli bağımlılıkları öğrenebilen bir RNN çeşididir. 1997 yılında Hochreiter ve Schmidhuber tarafından önerilen RNN'ler diğer araştırmacılar tarafından geliştirilmiştir. LSTM, birçok problemin çözümünde kullanılan en yaygın RNN türüdür (Hochreiter ve Schmidhuber, 1997).

Tekrarlayan Sinir Ağları yinelemeyi uygulamak için önceki adımdaki gizli durum vektörünü ve bir tanh katmanını kullanmaktadır. RNN yapısı şekil 7'de gösterilmektedir.



Şekil 7. RNN yapısı

Uzun Kısa Vadeli Sinir Ağları da yinelemeyi uygulamaktadır. Fakat burada tek tanh katmanını yerine dört etkileşimli katman kullanılmaktadır. Şekil 8'de LSTM yapısı gösterilmektedir.



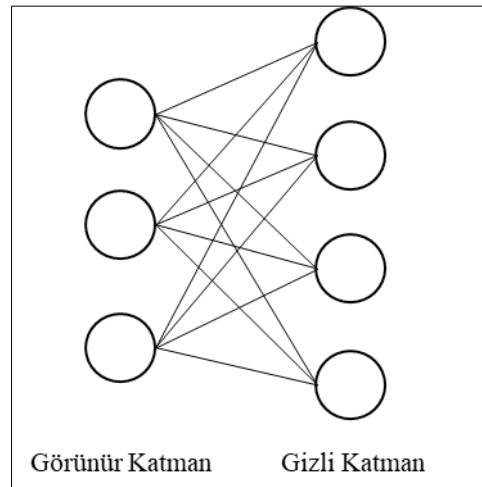
Şekil 8. LSTM yapısı

Şekil 8'deki üst çizgi ve c dahili belleği temsil etmektedir. Altındaki çizgi ise gizli durumu temsil etmektedir. LSTM'nin kaybolan gradyan problemini çözmek için i, f, o, ve g geçitleri kullanılmaktadır. Eğitim sırasında LSTM bu geçitler için parametreleri öğrenmektedir (Gulli ve Pal, 2017). Sigmoid katmanı unutma geçidi olarak adlandırılmaktadır. Unutma geçidi, birbirini takip eden katmanlar arasında önceki katmandaki bilgilerin bir sonraki katmana ne kadar geçilmesi gerektiğine karar veren bir katmandır (Kim vd., 2017).

### 1.5.3. Kısıtlı Bozaltman Makineleri (RBM)

Kısıtlı Bozaltman Makineleri, Geoffrey E. Hinton tarafından girdi verisi üzerinde olasılık değerlerini öğrenebilen stokastik bir sinir ağı olarak tanıtılmıştır (Ackley vd., 1985). Hesaplama gücündeki artış ve daha hızlı öğrenme algoritmalarının geliştirilmesi ile RBM, makine öğrenimi problemlerine uyarlanabilir hale getirilmiştir.

RBM'lerin yapısı Şekil 9'daki gibi iki katmandan oluşmaktadır. Bu katmanlar görünür ve gizli katmanlardır. Görünür katman, birinci katmanı oluşturur ve bir verinin bileşenlerine denk gelir (Örn. dijital bir görüntüdeki her bir piksel değeri). Gizli katman, ikinci katmanı oluşturur ve veriler arasındaki bağımlılıkları modeller (Örn. pikseller arasındaki ilişkiler) (Hinton vd, 2007). RBM ağ yapısında her nöron diğer katmandaki nöronlarla bağlı olmaktadır. Bunun yanında her nöronun kendi katmanındaki nöronlar ile bir bağlantısı bulunmamaktadır. Bu durum RBM'ye adını vermiştir (Fischer ve Igel, 2014).

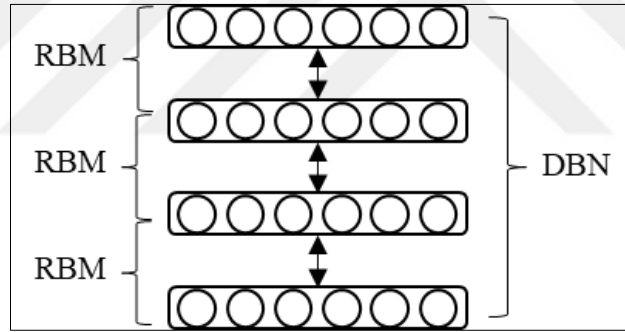


Şekil 9. Kısıtlı Bozaltman Makinesi yapısı

RBM görüntü sınıflandırma, görüntü işleme, film uygulamalarındaki film tavsiyeleri için filtre yapma (Salakhutdinov vd., 2007), örüntü hareketlerini öğrenme (Taylor vd., 2007) gibi problemlerin çözümünde kullanılmaktadır (Fischer ve Igel, 2014).

#### 1.5.4. Derin İnanç Ağları (DBN)

RBM iki katlı bir mimariye sahiptir. Derin İnanç Ağları, çok sayıda gizli katman içeren RBM'lerin katmanlar halinde sıralanmasıyla oluşmaktadır (Salakhutdinov vd., 2009). Şekil 10'da DBN yapısı gösterilmektedir. Şekilde görüldüğü gibi RBM'nin gizli katman çıktısından elde edilen veriler diğer bir RBN'nin giriş katmanı olarak kullanılmaktadır. Bu işlem ileri özellik çıkarma olarak adlandırılmaktadır. Bu özellikleri kullanarak Geoffrey E. Hinton (Hinton vd., 2006) tarafından 2006 yılında RBM'ye dayanan DBN oluşturulmuştur (Hua vd., 2015).



Şekil 10. Derin İnanç Ağı Yapısı

DBN, yüksek doğruluk ve hesaplama verimliliğinden dolayı en güvenilir yöntemlerden biri olmuştur. Bu nedenle uygulama alanları çok çeşitli olabilmektedir. DBN, birçok mühendislik alanında, duygu tespitinde, ekonomik tahminlerde, kanser teşhisi gibi alanlarda kullanılmaktadır (Mosavi vd., 2019).

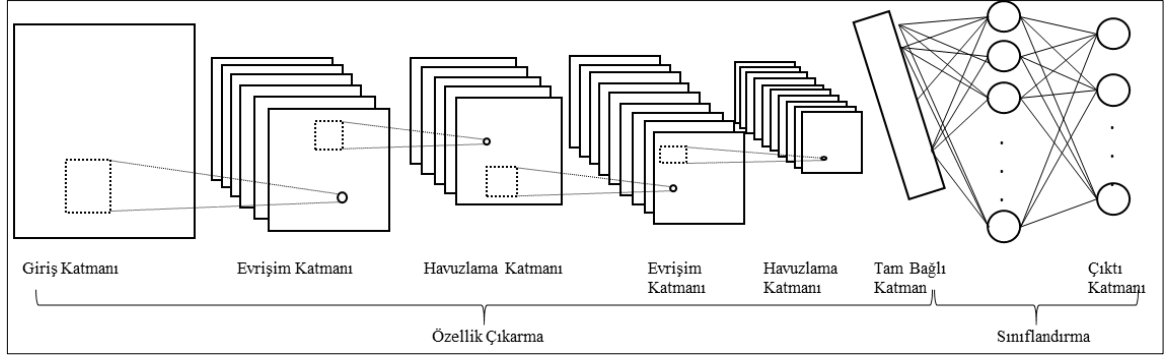
### 1.5.5. Evrişimli Sinir Ağları (ESA)

Evrişimli Sinir Ağı, nöronlar arasındaki bağlantı için hayvan görsel korteksinden ilham alınan bir tür ileri beslemeli yapay sinir ağıdır. Kortikal nöronlar, alıcı alan olarak bilinen sınırlı bir alan bölgesinde uyarılara yanıt vermektedir. Farklı nöronların alıcı alanları görsel alanı kapsayacak şekilde oluşturulmaktadır. Bir nöronun kendi alıcı alanı içerisindeki uyarılara tepkisi evrişim işlemi ile matematiksel olarak tahmin edilebilmektedir (Ongsulee, 2017). ESA, biyolojik süreçlerden esinlenmiştir ve minimum miktarda ön işleme kullanmak için tasarlanmış çok katmanlı algılayıcılardan oluşmaktadır (Ongsulee, 2017).

ESA'lar gradyan tabanlı bir öğrenme algoritmasına sahiptir. Bu sayede azalan gradyan probleminden daha az etkilenmektedir. Gradyan tabanlı algoritma hata değerini en aza indirmek için tüm ağı eğitir ve bu sayede ESA yüksek düzeyde optimize edilmiş ağırlıklar üretebilmektedir (Alom vd., 2019).

ESA, görüntü sınıflandırma, nesne algılama, konuşma tanıma, doğal dil işleme, tıbbi görüntü analizi gibi problemlerinin çözümünde yüksek performans sağlamıştır. ESA ile görsel işlemlerde tüm geleneksel yöntemlerden daha iyi sonuçlar elde edilmiştir (Wani vd., 2020). Örneğin, görüntü tanıma uygulamasında CNN'ler birden çok alıcı alan katmandan oluşur. Bunlar, girdi görüntüsünün bölümlerini işleyen nöron topluluklarıdır. Bu nöronların çıktıları daha sonra orijinal görüntünün daha yüksek çözünürlüklü bir temsili elde etmek için üst üste gelecek şekilde sıralanır. Bu durum her katman için tekrarlanmaktadır (Ongsulee, 2017).

ESA, girdi verilerinin ızgara benzeri bir topolojik yapıya sahip olduklarını varsaymaktadır. Örneğin; düzenli aralıklarla bilgi alan zaman serileri bir boyutlu ızgara tipine sahipken görüntü verileri iki boyutlu düşünülebilir (Goodfellow vd., 2016). ESA, Şekil 11'de görüldüğü üzere özellik çıkarma ve sınıflandırma adı altında iki ana bölümden oluşmaktadır. Özellik çıkarma bölümünde ağıdaki katmalar bir önceki katman çıktısını alarak girdi olarak kullanılmaktadır. Son katmanın çıktısı, tamamen bağlı katmanda (sınıflandırma katmanı) girdi olarak kullanılmaktadır. ESA mimarisi evrişim (convolution), havuzlama (pooling), tam bağlı (fully connected) katmanlarından oluşmaktadır (Zin, 2018).



Şekil 11. Evrişimli Sinir Ağı mimarisi

### 1.5.5.1 Evrişim Katmanı

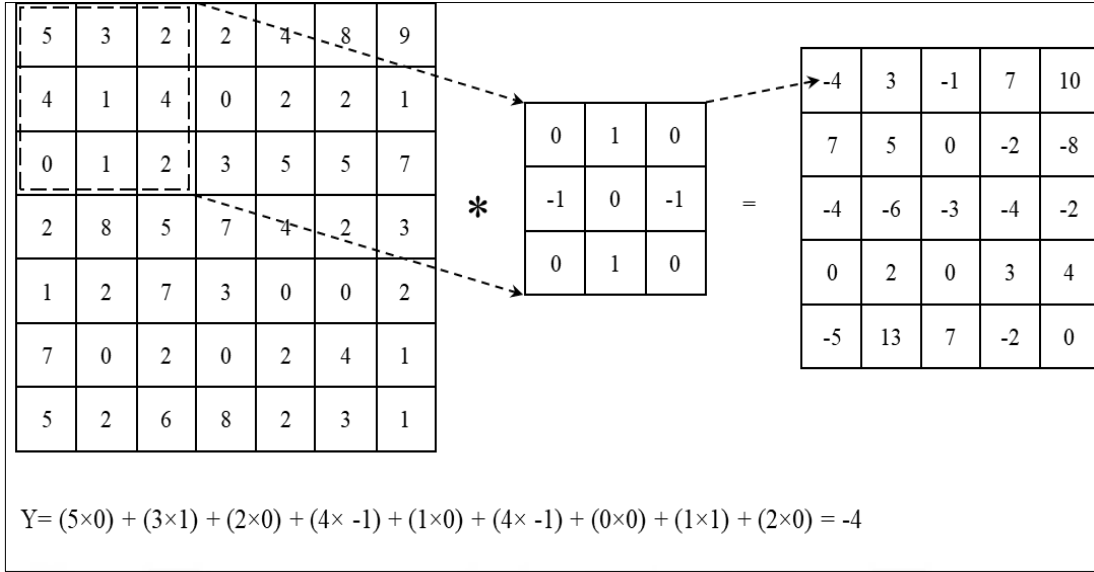
ESA'nın temel yapı taşı olan evrişim katmanı, giriş verisinin özelliklerini çıkarmak için kullanılan ilk katmandır. Evrişim katmanının amacı önceki katmanlardan elde edilen özelliklerin bağlantılarını tespit etmektir. Evrişim katmanı girdi olarak bir özellik haritası kullanmaktadır. Bu özellik haritası bir dizi yeni özellik haritası üretmek için filtrelenir. Filtre, giriş görüntüsü üzerinde adım adım kaydırılır ve Şekil 12'deki evrişim formülü uygulanarak özellik haritası çıkartılır. Formülde görüldüğü üzere m ve n terimleri evrişim sonucunda elde edilecek yeni matrisin konumlarını ifade etmektedir.

|     |      |     |     |
|-----|------|-----|-----|
| $m$ | $-1$ | $0$ | $1$ |
| $n$ | $-1$ | $0$ | $1$ |
|     | a    | b   | c   |
|     | d    | e   | f   |
|     | g    | h   | i   |

$$y[m, n] = \sum_{j=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} x[i, j] \cdot h[m - i, n - j]$$

Şekil 12. Evrişim tanımlanması

Şekil 13'de 7×7'lik bir girdiye 3×3'lük bir filtre uygulanarak evrişim işlemi yapılmıştır. Filtre sol üst tarafa denk gelecek şekilde yerleştirildikten sonra evrişim formülü uygulanarak çıktı değeri elde edilmektedir.



Şekil 13. Evrişim işlemi

Bu işlem bütün girdi görüntüsüne filtrenin adım adım kaydırılması ile uygulanmaktadır. Girdinin son sütununa gelindiğinde bir piksel aşağı inilerek tekrardan işlemlere devam edilir ve filtre girdinin tamamına uygulanana kadar süreç devam eder.

ESA tasarlanmaya başlarken katman sayısının ve her katmandaki filtre sayısının belirlenmesi gerekmektedir. Uygun filtreler ve ağırlıklar, ESA tarafından otomatik olarak bulunmaktadır (Gad vd., 2018). Kullanılan filtreler örnek olarak Şekil 14’te gösterilmiştir. Yatay kenar ve dikey kenar belirleyebilmek için a ve b’deki filtreler kullanılmaktadır. Sırasıyla köşeleri, görüntüdeki gri değerlerin keskinliğini, girdi verisinde dikdörtgen, çember ve üçgensel şekillerin tespiti için c, d, e, f, g’deki filtreler kullanılmaktadır (Albawi vd., 2017). Filtreler katmanlara eklenerek girdi verisinin farklı özellikleri çıkartılmaktadır.

|    |    |    |
|----|----|----|
| 1  | 1  | 1  |
| 0  | 0  | 0  |
| -1 | -1 | -1 |

a) Yatay kenar belirleme

|   |   |    |
|---|---|----|
| 1 | 0 | -1 |
| 1 | 0 | -1 |
| 1 | 0 | -1 |

b) Dikey kenar belirleme

|   |    |   |
|---|----|---|
| 0 | 1  | 0 |
| 1 | -4 | 1 |
| 0 | 1  | 0 |

c) Köşe belirleme

|    |    |    |
|----|----|----|
| 0  | -1 | 0  |
| -1 | 5  | -1 |
| 0  | -1 | 0  |

d) Keskinlik belirleme

|   |   |   |   |
|---|---|---|---|
| 1 | 1 | 1 | 1 |
| 1 | 0 | 0 | 1 |
| 1 | 0 | 0 | 1 |
| 1 | 1 | 1 | 1 |

e) Dikdörtgen belirleme

|   |   |   |   |
|---|---|---|---|
| 0 | 1 | 1 | 0 |
| 1 | 0 | 0 | 1 |
| 1 | 0 | 0 | 1 |
| 0 | 1 | 1 | 0 |

f) Çember belirleme

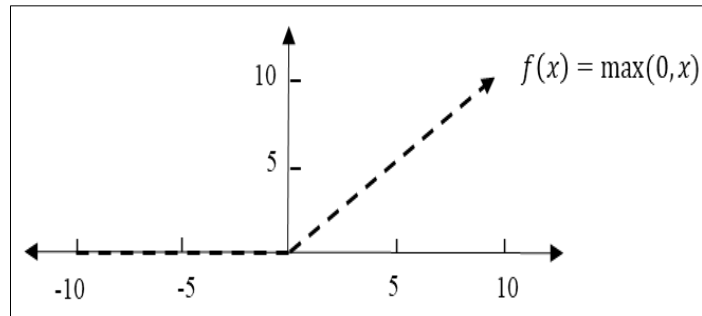
|   |   |   |   |
|---|---|---|---|
| 0 | 0 | 0 | 0 |
| 0 | 1 | 1 | 0 |
| 1 | 0 | 0 | 1 |
| 1 | 1 | 1 | 1 |

g) Üçgen belirleme

Şekil 14. ESA filtre örnekleri

#### 1.5.5.2. Doğrusal Olmayan Katman (Non-Linearity)

Evrişim katmanından sonra girdi verilerin doğrusal olmayan temsillerini geliştirmek için doğrusal olmayan aktivasyon fonksiyon katmanı (örneğin, ReLu, tanh ve sigmoid) kullanılır. Derin öğrenme algoritmalarında popüler olarak ReLu aktivasyon fonksiyonu kullanılmaktadır. Aktivasyon fonksiyonu olarak ReLu kullanılması durumunda tüm negatif değerler sıfır olarak ayarlanacaktır. Bu durum ağı daha hızlı çalışmasını sağlamaktadır (Venkatesan vd., 2017). Aktivasyon işlemi ağı doğrusal olmayan bir özellik getirerek hem ağı hem de gerçekleştireceği görevlerin karmaşıklığını artırmaktadır. Şekil 15'te Relu aktivasyon fonksiyonunun grafiği verilmiştir.

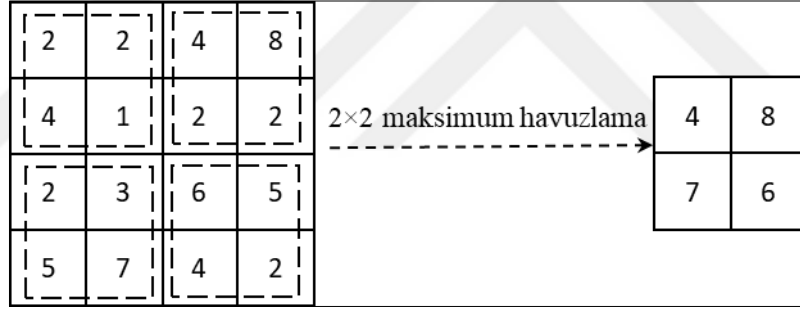


Şekil 15. Relu grafiği

### 1.5.5.3 Havuzlama Katmanı

Havuzlama katmanının rolü, girdideki anlamsal olarak benzer özellikleri bir araya getirmektir. Başka bir deyişle havuzlama katmanı, ağdaki parametre ve hesaplama miktarını azaltarak mekânsal boyutu küçültmektedir. Havuzlama işlemi her bir özellik haritası için ayrı ayrı gerçekleştirilir ve havuzlama biriminin boyutu her zaman özellik haritasından daha küçüktür. Birçok ESA algoritmasında maksimum havuzlama ve ortalama havuzlama katmanı kullanılmaktadır. Ağ mimarilerinde maksimum havuzlama katmanı daha iyi sonuç verdiği için yaygın olarak kullanılmaktadır. (Lee vd., 2016).

Şekil 16'da 4x4'lük bir girdiye 2x2'lik bir filtre 2 adım kaydırılacak şekilde uygulanarak maksimum havuzlama işlemi yapılmıştır. Bu işlem yapılırken her 2 x 2'lik matrisin içerisindeki maksimum değer alınarak boyutu küçültülmüş yeni bir matris oluşturulmaktadır.



Şekil 16. Havuzlama işlemi

### 1.5.5.4 Tamamen Bağlı Katman

Girdideki özelliklerin doğrusal olmayan ilişkilerini daha fazla modellemek için genellikle sondan bir önceki katman ve çıkış katmanı arasında tamamen bağlı bir katman kullanılır. Tamamen bağlı katmanda nöronlar önceki katmandaki tüm aktivasyonlara ve tam bağlantıya sahiptir. Tamamen bağlı katman sonucunda 2 boyutlu özellik haritaları bir boyutlu özellik vektörüne dönüştürülür (Zhang vd., 2021). Tamamen bağlı katman aşağıdaki formülle ifade edilmektedir (1).

$$y_i = f(W \cdot u_i) + b^h \quad (1)$$

Formülde;  $u_i$  : girdi,  $W$ ; nöron katmanı ile ilişkili ağırlık matrisi,  $b$ ; bias değeri,  $f$ ; aktivasyon fonksiyonunu ifade etmektedir (Bouvier, 2006).

#### 1.5.5.5. Sınıflandırma Katmanı

Yapısal olarak tam bağlantılı katman yapısına çok benzerdir. Sınıflandırma katmanı, sınıflandırılmak istenen nöron sayısı ile aynı sayıda nörona oluşmaktadır. Çıktı katmanındaki her nöron, belirli bir sınıfa karşılık gelen bir olasılık sonucu verir ve tahmin edilen sınıf maksimum olasılık sonucuna sahip olan sınıf olmaktadır (Morchhale, 2016).

### 1.6. Derin Ağ Mimarileri

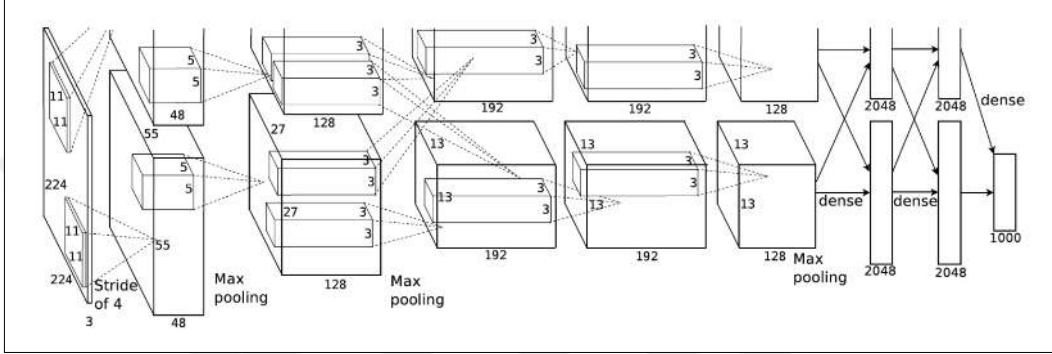
Ağ mimarileri, nesne tespiti için bir özellik oluşturma görevi görmektedir. Derin ağ mimarileri, girdi olarak görüntüyü alarak özellik haritasını oluşturmaktadır. Nesne tespiti için derin ağ mimarilerinin çoğu evrişim katmanında özellik haritası çıkarma işlemini yaparken, tam bağlı katmanda sınıflandırma işlemini yapmaktadır. Bu tür derin ağlara örnek olarak AlexNet, VGG16, ResNet ve Unet verilebilmektedir (Pal vd., 2021)

Daha derin ve yoğun bağlantılı ağ mimarileri, daha yüksek doğruluk elde etmek amacıyla daha sık ve seyrek olan ağ mimariler yerine kullanılmaktadır. Örneğin; Faster R-CNN'de zengin özellikleri tanımlayabilmek ve daha yüksek doğruluk elde etmek için VGG16 yerine yoğun bağlantılı ResNet kullanılmıştır. Dolayısıyla özelliklerin haritasının kalitesini ağ mimarileri belirlemektedir. Daha derin ve daha yoğun bağlantılı mimariler ile, sık ve seyrek bağlantılı mimarilere göre daha iyi sonuçlar elde edebilmektedir.

#### 1.6.1. AlexNet

2012 yılında (Krizesky vd., 2012) ImageNet Büyük Ölçekli Görsel Tanıma Yarışmasında (ImageNet Large Scale Visual Recognition Challenge, ILSVRC) evrişimli derin sinir ağ yapısı olan AlexNet'i tanıtmıştır. AlexNet, TOP-5 test doğruluğu puanında %84,6 olarak yarışmayı kazanmış ve öncü derin ağ mimarisi olarak kabul edilmiştir. Aynı yarışmada derin ağlar yerine geleneksel teknikleri kullanan en yakın rakibi %73,8 doğruluk elde etmiştir. AlexNet mimarisi oluştururken (Zhu vd., 2017);

- Geleneksel hiperbolik tanjant aktivasyon fonksiyonu yerine birkaç kat daha hızlı olan ve eğitim süresini azaltabilen doğrusal olmayan işlem birimleri (ReLU)'lar kullanılmıştır.
- Aşırı uyumu önlemek amacıyla bırakma katmanı (dropout) kullanılmıştır.
- Eğitim setinin boyutunu yapay olarak artırmak için veri artırma teknikleri kullanılmıştır. Şekil 17'de AlexNet mimarisi gösterilmektedir.

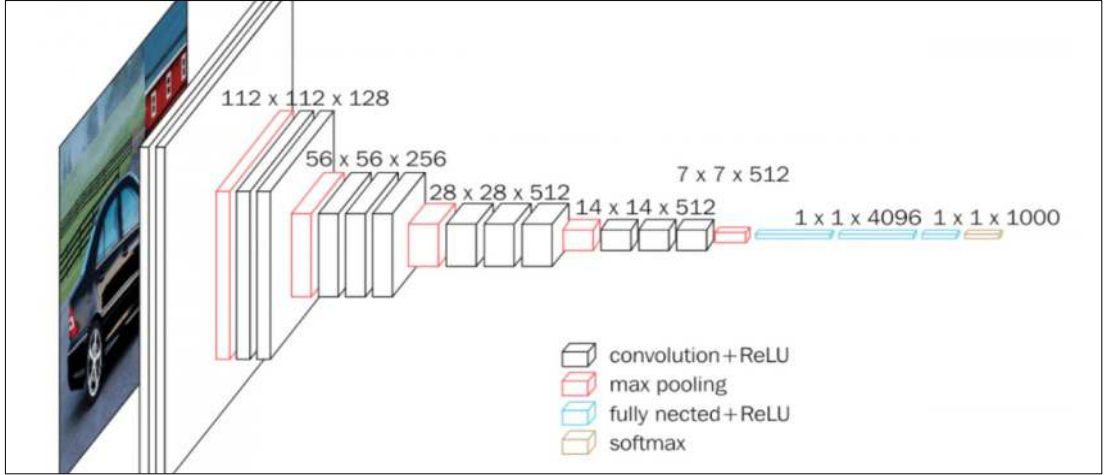


Şekil 17. AlexNet ağ mimarisi yapısı (Krizesky vd., 2012).

AlexNet'in başarılı olmasındaki bir diğer sebep ise GPU'lar kullanılarak eğitim işleminin yapılmasıdır. GPU'lardaki fazla çekirdek sayesinde büyük veri kümeleri ve büyük görüntülerin hızlı bir şekilde eğitilmesi sağlanmıştır.

### 1.6.2. VGG Net

VGG Oxford Üniversitesi görsel geometri gurubu tarafından üretilen bir ESA derin ağ modelidir. 2013 yılında düzenlenen ILSVRC yarışmasında TOP-5 doğruluk testinde %92,7 doğruluk elde ederek popüler hale gelmiştir. Şekil 18'de VGG ağ mimarisi gösterilmektedir. VGG diğer ağ mimarilerinden farklı olarak büyük alıcı alanlara sahip birkaç katman yerine küçük alıcı alanlara sahip evrişim katmanlarını kullanmaktadır. Bu durum daha az parametreye ve katmanlar arasında daha az doğrusal olmamaya yol açmaktadır, böylece karar fonksiyonu daha ayırt edici olur ve model daha kolay bir şekilde eğitilir. (Garcia-Garcia vd., 2017)

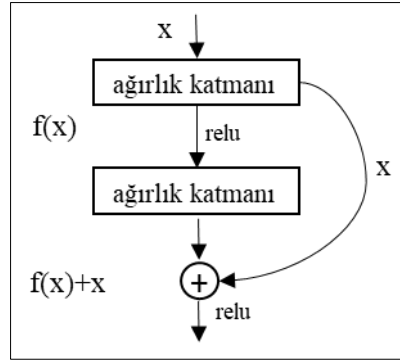


Şekil 18. VGG ağ mimari yapısı (Simonyan vd., 2014)

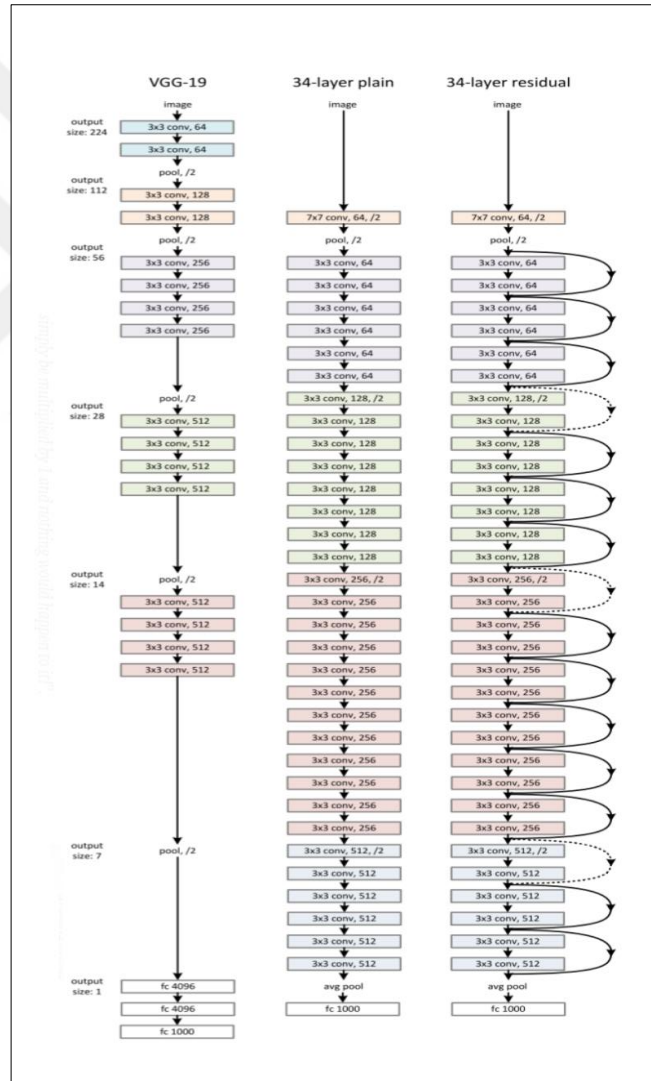
### 1.6.3. ResNet

He vd., (2016) evrişimli sinir ağları derinleştikçe doğruluk oranlarının en yüksek seviyeye ulaştığını sonrasında ise hızlı bir şekilde düşüşe geçtiğini göstermiştir. Performans düşüşünü azaltmak için evrişim katmanlarının arasına atlama bağlantılarının kullanılmasını önermiştir. Şekil 19’da ResNet atlama bağlantısı gösterilmektedir. Bu atlama bağlantısı, giriş ve çıkış blokları arasındaki eleman bazında bir eklemekten oluşmaktadır. Bu durum ağa fazladan parametre veya hesaplama karmaşıklığı eklememektedir. Standart bir 34 katmanlı ResNet yapısı temel olarak 7×7 evrişim katmanı, 16 darboğaz (aralarında atlama bağlantısı bulunan 3×3’lük filtre) ve tamamen bağlı katmandan oluşmaktadır. He ve arkadaşları ayrıca 16 katmanlı VGG ağının 101 ve 152 katmanlı ResNet mimarilerinden daha düşük doğruluğa sahip olduğunu göstermiştir.

ResNet ağ mimarisi 2016 yılında yapılan ILSVRC yarışmasında Top-5 testinde %96,4 doğruluk değerine ulaşarak yeni bir rekor kırmıştır. Bu durum sınıflandırma, algılama gibi alanlarda derin ağ mimarilerini daha ileriye taşımıştır. ResNet, sınıflandırma ve nesne tespitinde yaygınlıkla kullanılmaktadır ve temel ilkeleri birçok ağ mimarisine ilham vermiştir (Zaidi vd., 2022). Şekil 20’de VGG ve ResNet ağ mimarileri gösterilmektedir.



Şekil 19. ResNet atlama bağlantısı



Şekil 20. VGG ve ResNet ağ mimarisi (He vd., 2016)

## 1.7. Nesne Tespiti

Görüntüler üzerindeki nesne algılamaya yönelik derin öğrenme algoritmalarının (bölgesel tabanlı ESA'lar) başarısından esinlenerek, Uzaktan Algılama verileri için birçok derin öğrenme tabanlı nesne algılama yöntemi geliştirilmiştir. Bu yöntemler üst düzey anlamsal temsilleri daha iyi öğrenme yeteneklerine sahip olduğundan geleneksel yöntemlere göre avantaj göstermektedir (Zaidi vd., 2022).

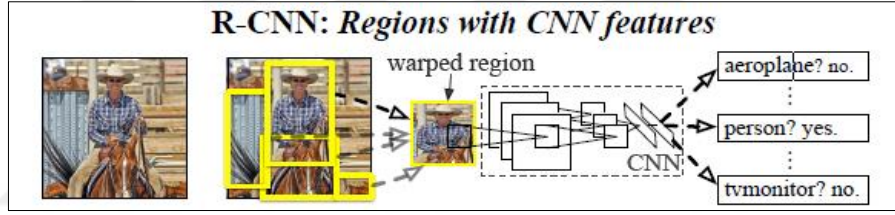
Genel nesne algılama, dikdörtgen sınırlayıcı kutu kullanarak nesnenin görüntüdeki yoğunluğunu belirtme ve nesneyi bir etiketle sınıflandırma işlemidir. Nesne algılayıcı, bölgesel tabanlı algılayıcı ve regresyon/sınıflandırma tabanlı algılayıcı olmak üzere iki kategoriye ayrılmıştır. Bölgesel tabanlı algılayıcı, geleneksel nesne algılama yöntemini takip etmektedir. Önce bölgenin önerilen nesnesi tespit edilmekte, ardından bölgelerin sınıflandırması yapılmaktadır. R-CNN (Girshick vd., 2014) , SPP-net (He vd., 2015), Fast R-CNN (Girshick, 2015), Faster R-CNN (Ren vd., 2015), R-FCN (Dai vd., 2016), FPN (Lin vd., 2017) ve Mask R-CNN (He vd., 2017) bölgesel tabanlı algılayıcılara örnek olarak verilebilir. Regresyon/sınıflandırma tabanlı algılayıcılar ise nesne algılamadaki sınırlayıcı kutuları ve nesne sınıflarının olasılıklarının bulunmasını bir regresyon problemi olarak ele alır. Tek bir sinir ağı, tek bir değerlendirme ile tüm görüntünün sınırlayıcı kutularını ve sınıf olasılıklarını tahmin etmektedir. Sınıflandırma ve regresyon tabanlı algılayıcılara Multibox (Erhan vd., 2014), AttentionNet (Yoo vd., 2015), G-CNN (Najibi vd., 2016), YOLO (Redmon vd., 2016), SSD (Liu vd., 2016), YOLOv2 (Redmon ve Farhadi, 2017), DSSD (Fu vd., 2017) ve DSOD (Shen vd., 2017) örnek olarak verilebilir (Aziz vd., 2020).

Bu tez kapsamında bölgesel tabanlı algılayıcılara değinilecektir. Bölgesel tabanlı algılayıcılar, insan beynindeki algılama işlevini taklit ederek tüm görüntüyü tarar ve ilgilenilen bölgeye odaklanır.

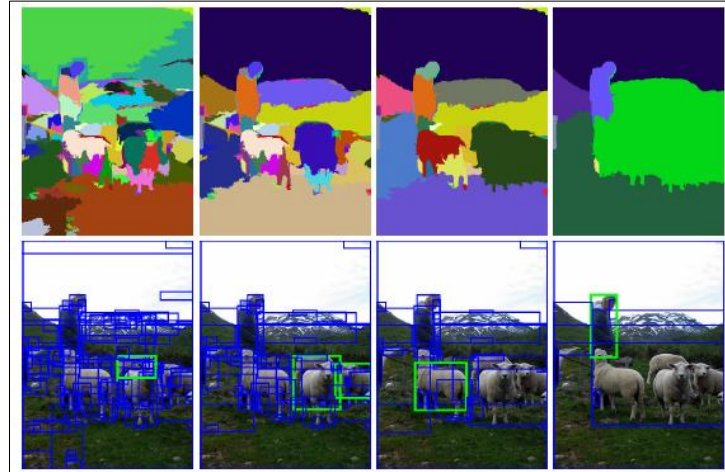
### 1.7.1. R-CNN

Bölgesel tabanlı evrişimli sinir ağı (R-CNN) 2014 yılında Girshick ve arkadaşları tarafından “Doğru nesne algılama ve anlamsal segmentasyon için zengin özellik hiyerarşileri Test raporu:5” başlığı altında tanıtılmıştır (Girshick vd., 2014). R-CNN, görüntüde bulunan nesneleri belirleyerek bunları çevreleyen sınırlayıcı kutular (bounding box) oluşturabilmektedir.

R-CNN, Şekil 21'deki gibi dört kısımdan oluşmaktadır. Birinci kısımda, giriş görüntüsü alınmaktadır. İkinci kısımda seçici arama (selective research) kullanarak 2000 bölge önerisi üretilmektedir. Seçici arama, görüntülerdeki nesneleri belirlemek için kullanılan bir algoritmadır (Uijlings vd., 2013). İlk olarak görüntü üzerindeki küçük bölgeler belirlenmektedir. Daha sonra birbirine benzer olan bölgeler birleştirilerek daha büyük bölge oluşturulmaktadır. Şekil 22'de seçici arama görseli bulunmaktadır. Önerilen bölgeler nesne içerme olasılığı yüksek bölgelerdir. Her bölge önerisindeki görüntü parçaları üzerinde CNN çalıştırılmaktadır. CNN, her bölge için özellik vektörü çıkarmaktadır. Bu özellik vektörleri sınıflandırma yapılması amacıyla destek vektör makinelerinin (SVM) girdisi olarak kullanılmakta ve nesne konumunun tespiti için sınırlayıcı kutu oluşturulmasını sağlamaktadır. R-CNN, nesne tespiti konusunda yüksek doğruluk göstermesine rağmen her nesne teklifi için ayrı bir CNN çalıştırdığı için uzun zaman almaktadır (Girshick vd., 2014).



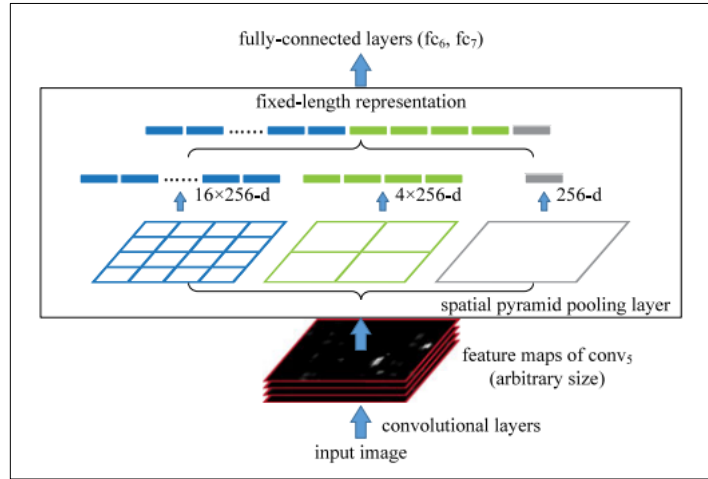
Şekil 21. R-CNN mimari yapısı (Girshick vd., 2014)



Şekil 22. Seçici arama (Uijlings vd., 2013)

### 1.7.2. SPP-net

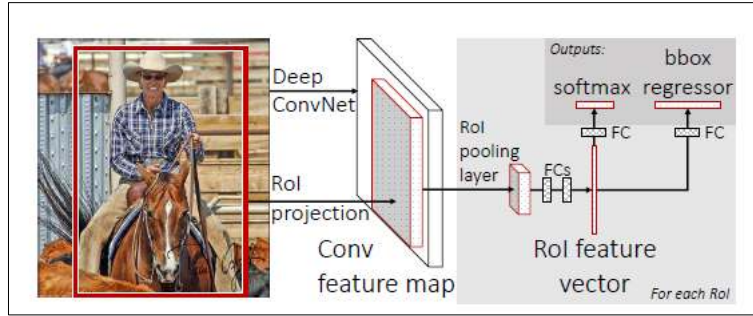
2015 yılında He vd., (2015) uzaysal piramit eşleştirme (Lazebnik vd., 2006) teorisine dayanarak oluşturulan SPP-net'i önermiştir. R-CNN, girdi görüntüsünü sabit boyutlu alırken SPP-net girdi boyutundan bağımsız şekilde çalışabilmektedir. R-CNN'de evrişim katmanından tam bağlantılı katmana geçişte tek bir havuzlama katmanı bulunmaktadır. SPP-net'te ise uzamsal piramit katmanı (spatial pyramid pooling layer) adıyla adlandırılan birden fazla havuzlama katmanı bulunmaktadır. Bu durum ağı girdi boyutundan bağımsız hale getirerek hesaplama miktarını azaltmaktadır. Nesnelerin tespitinde aday pencereleri oluşturmak için seçici arama algoritması kullanılmaktadır. Özellik haritalarını oluşturmak için ZF-5 ağ mimarisi kullanılmaktadır. Aday pencereler daha sonra havuzlama katmanı ile sabit uzunluktaki vektör temsillerine dönüştürülmektedir. Bu vektör, her bir aday nesnenin sınıfını tahmin etmek için tamamen bağılı katman ve SVM sınıflandırıcıları tarafından sınıflandırılmaktadır. SPP-net, R-CNN modeline kıyasla daha hızlıdır. SPP-net, herhangi en /boy oranına sahip görüntüleri işleyebilmekte ve böylece girdi farklılığından oluşan nesne deformasyonunu önleyebilmektedir (Aziz vd., 2020). Şekil 23'de SPP-net mimarisi gösterilmektedir.



Şekil 23. SPP-net mimarisi (He vd., 2015)

### 1.7.3. Fast R-CNN

Microsoft araştırma ekibinden olan Girshick, 2015 yılında nesne tespiti için R-CNN ve SPP-net'ten daha iyi bir model olan Fast R-CNN'i önermiştir (Girshick, 2015). R-CNN'de bölge önerisi olarak 2000 bölge kullanılmaktadır. Her bir bölge önerisi üzerinde CNN çalıştırıldığından dolayı zaman ve maliyet açısından önerilmemektedir. Fast R-CNN, girdi görüntüsü olarak tüm görüntüyü kullanmaktadır. Fast R-CNN, bir evrişim özellik haritası oluşturmak için tüm görüntüyü maksimum havuzlama katmanı ile işlemektedir. Ardından her nesne teklifi için bir ilgi bölgesi (Region of Interest (ROI)) havuzlama katmanı kullanılarak özellik haritasından sabit uzunlukta bir özellik vektörü çıkartılmaktadır. Fast R-CNN'de; CNN, sınıflandırıcı ve sınırlayıcı kutu regresörü tek bir modelde eğitilmektedir. Fast R-CNN, R-CNN ve SPP-net gibi destek vektör makineleri kullanmak yerine sınıflandırma için softmax katmanını kullanmaktadır. Ayrıca Fast R-CNN, sınırlayıcı kutuların çıktısını almak için bir doğrusal regresyon katmanı kullanmaktadır. Bu iyileştirmelerin yanı sıra nesne önerilerini oluşturmak için bölge önerisine ihtiyaç duyulmaktadır. Şekil 24'de Fast R-CNN yapısı gösterilmektedir.

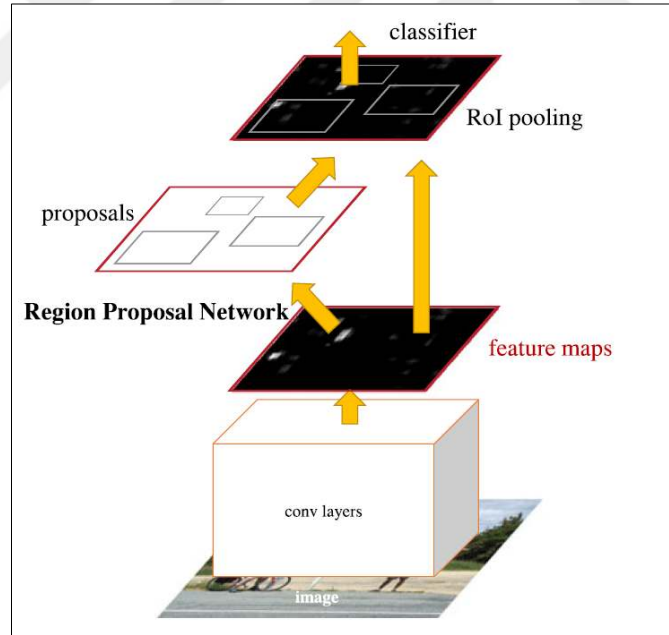


Şekil 24. Fast R-CNN mimari ağ yapısı (Girshick, 2015)

### 1.7.4. Faster R-CNN

Microsoft araştırma ekibinden olan Girshick ve arkadaşlarının yazmış oldukları “Faster R-CNN: Bölge Öneri Ağları ile Gerçek Zamanlı Nesne Algılamaya Doğru” başlıklı makalede, Fast R-CNN'deki bölge önerisi sorunu çözmek için bölge teklif ağını (RPN) önermiştir (Girshick, 2015). Bölge teklif ağı (RPN), tamamen evrişimli bir ağa benzemektedir. RPN'nin temel görevi nesneleri tanımlamak için algılama önerileri

oluşturmaktır. RPN'nin girdisi bir görüntüdür. Çıktısı, nesne önerileri için sınıflandırma puanları ve nesneleri çeviren dikkörtgenlerdir. Bölge önerilerini oluşturmak için, özellik haritası üzerinden kayan bir pencere (sliding window) çalıştırılmaktadır. Her sürgülü pencere konumunda birkaç bölge önerisi tahmin edilmektedir. Birleşim üzerinden kesişim (Intersection over Union (IoU)) kullanılarak önerilerin doğruluğu kontrol edilmektedir. IoU, iki sınırlayıcı kutunun kesişim alanının birleşim alanına oranıdır. IoU değeri 0,7'den büyükse nesne vardır veya 0,3'ten küçükse nesne yoktur şeklinde "nesnellik" puanları hesaplanarak nesne bölgesi önerilmektedir. Sonraki aşama Fast R-CNN ile benzer olmaktadır. Bölge önerileri ROI havuzlama katmanında ve tam bağlı katmanlarda işlenmektedir. Sınıflandırma katmanında softmax ve sınırlayıcı kutu regresörü kullanılmaktadır. Faster R-CNN ile yüksek doğrulukla bir görüntüdeki nesnenin tam konumu bulunabilmektedir. Şekil 25'de Faster R-CNN yapısı gösterilmektedir (Ren vd., 2015).

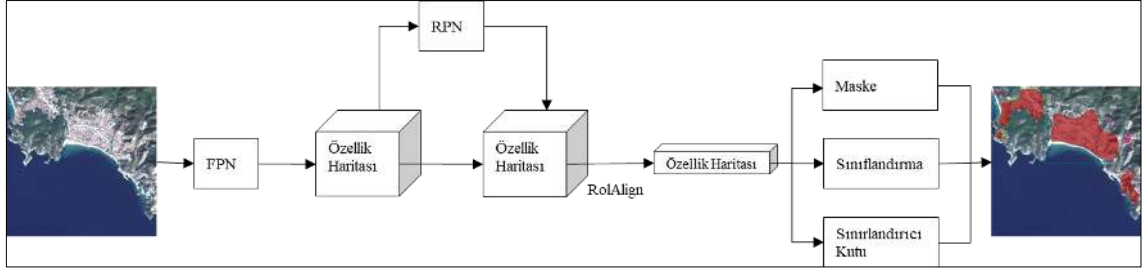


Şekil 25. Faster R-CNN yapısı (Ren vd., 2015).

### 1.7.5. Mask R-CNN

Bölgesel tabanlı CNN mimarileri kullanılarak nesne tespiti, semantik segmentasyon ve örnek segmentasyon alanlarında çalışmalar gerçekleştirilmiştir. Bu mimari yapılara; Fast R-CNN, Faster R-CNN, FCN (Fully Convolutional Network), Mask R-CNN örnek olarak verilebilmektedir (Bhuiyan vd., 2020). Nesne tespiti bir görüntüdeki nesneleri tespit edip sınıflandırmayı amaçlarken semantik segmentasyon görüntüdeki nesneleri tanımlayarak sınıf etiketlerini atamayı amaçlamaktadır. Örnek segmentasyon ise görüntüdeki nesneleri algılama ve her nesnenin maskesini oluşturma aşamasından oluşan zorlu bir süreçtir. Faster R-CNN'in geliştirilmesiyle oluşturulan Mask R-CNN örnek segmentasyon yapabilmektedir. Mask R-CNN, nesneler arasında net bir ayrım yaparak ve nesneler üst üste geldiğinde bile iyi sonuçlar vermektedir. Bu nedenle uzaktan algılama verilerinde örnek segmentasyon için Mask R-CNN algoritması kullanılmaktadır (Su vd., 2019).

Faster R-CNN'in geliştirilmesiyle Mask R-CNN oluşturulmuştur. Faster R-CNN, nesne tespiti yaparken her nesnenin sınırlayıcı kutusunu ve sınıf etiketini belirlemektedir. Mask R-CNN ise bunların yanı sıra her nesnenin maskesini oluşturmaktadır. Bu durumda Mask R-CNN iki aşamalı prosedür benimsemektedir. İlk aşamada bölge teklif ağı (RPN) kullanılarak nesnelerin sınırlayıcı kutularını belirlemektedir. İkinci aşamada RoIAlign kullanılarak her nesnenin maskesi belirlenmektedir. Başka bir deyişle; FPN (özellik piramit ağları), özellik çıkarmak amacıyla Faster R-CNN'e eklenmiştir. Böylece, ağın nesne tespit doğruluğu ve nesne tespit hızı artmıştır. Mask R-CNN'de sıklıkla ResNet101 ve ResNet50 ağ mimarileri kullanılmaktadır. FPN, girdi olarak görüntüyü alarak özellik piramitlerini oluşturmaktadır. Daha sonra çapalara (anchors) göre RPN farklı ölçek özelliklerinde bir RoI oluşturmaktadır. RoI'ler Faster R-CNN'de RoI havuzlama katmanı ile çıkarılırken nesne tespit doğruluğunu artırmak amacıyla Mask R-CNN'de RoIAlign katmanı kullanılmaktadır. Son aşamada nesne önerileri, softmax ve tam bağlantılı katman (FCN) kullanılarak sınıflandırılmaktadır (He vd., 2017). Şekil 26'da Mask R-CNN ağ mimarisi gösterilmektedir.

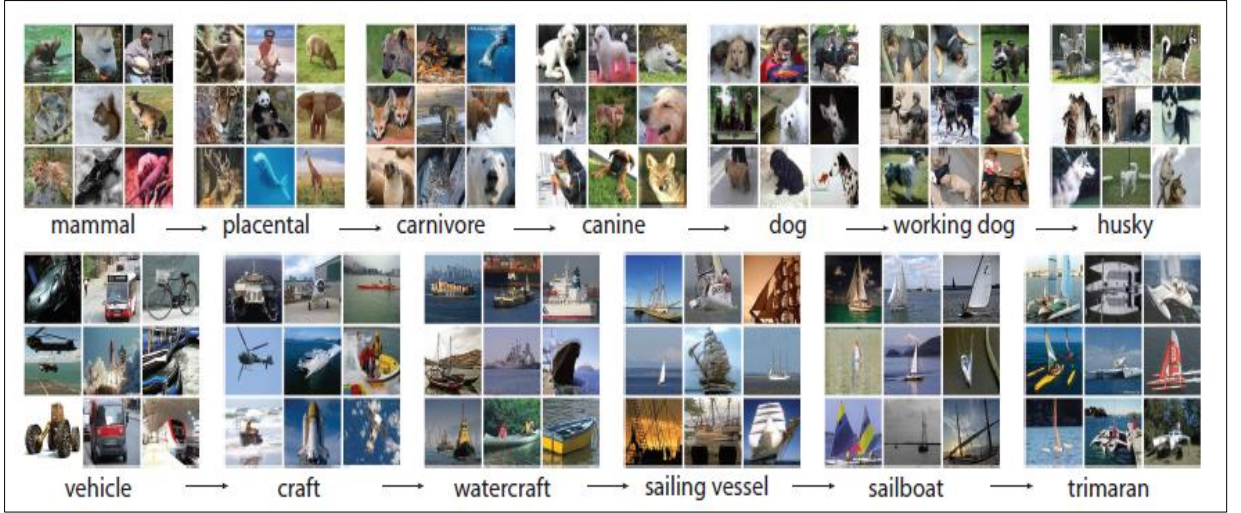


Şekil 26. Mask R-CNN ağ mimarisi yapısı

### 1.8. Derin Öğrenmede Kullanılan Yaygın Veri Setleri

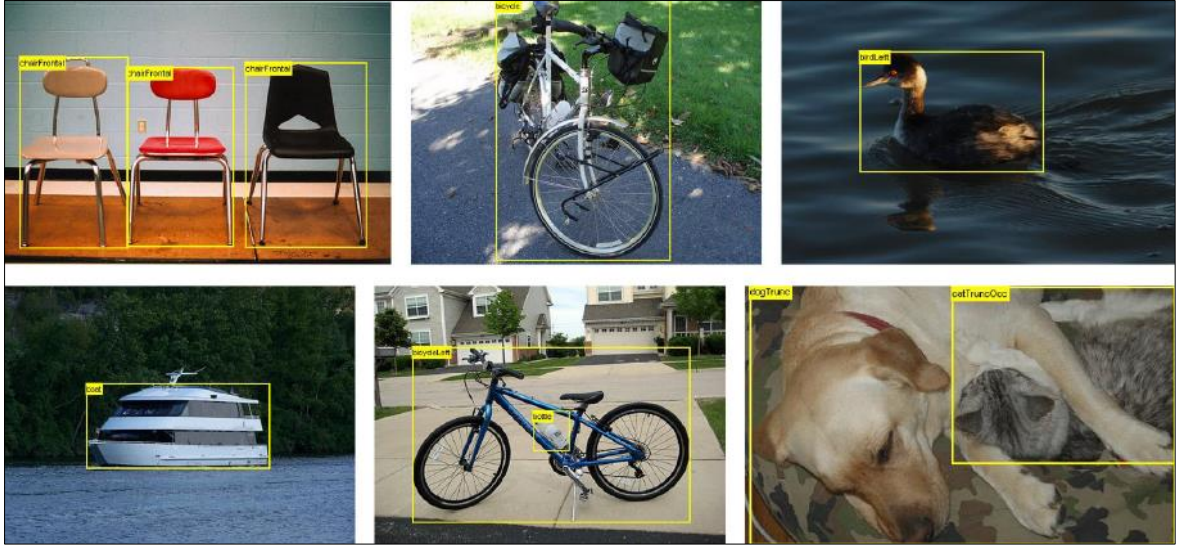
Derin öğrenme ile veri setleri arasında güçlü ve önemli bir ilişki bulunmaktadır. Derin öğrenme algoritmaları doğru bir şekilde hazırlanmış veri setleri olmadan anlamlı bilgiler çıkaramamaktadır. Veri setleri oluşturulurken görüntülere ve görüntü etiketlerine ihtiyaç bulunmaktadır. Görüntü etiketleri, görüntüdeki her bir nesne için sınırlayıcı kutulardan (nesneyi çevreleyen dikdörtgenin koordinatları) ve nesne maskelerinden oluşmaktadır. Bu etiketleme işlemi zahmetli ve karmaşık bilgiler içerdiğinden nitelikli kişiler tarafından yapılması gerekmektedir. Görüntü işlemede sıklıkla kullanılan veri setlerine ImageNet (Deng vd., 2009), İstatistiksel Modelleme ve Hesaplamalı Öğrenme Görsel Nesne Sınıfları (PASCAL VOC) (Everingham vd., 2015), Open Images (Kuznetsova vd., 2020), ve Microsoft Ortak Nesneleri (MS COCO) (Lin vd., 2014) örnek olarak verilebilmektedir.

ImageNet veri seti Deng vd., (2009) tarafından yapılan ‘ImageNet: A Large-Scale Hierarchical’ çalışması ile sunulmuştur. 2012 yılında ise ILSVRC yarışmasında ImageNet veri seti kullanılarak eğitilen derin sinir ağı birincilik elde etmiştir (Krizhevsky, vd. 2012). ImageNet; memeli, kuş, balık, sürüngen, amfibi, araç, mobilya, müzik aleti, jeolojik oluşum, alet, çiçek, meyve gibi kategorilerde binlerce resimden oluşmaktadır. ImageNet veri seti toplamda 14 milyondan fazla görüntü içermektedir. Şekil 27’de ImageNet veri seti örneği gösterilmektedir.



Şekil 27. ImageNet örnek veri seti (Deng vd., 2009)

Pascal VOC veri seti görüntü işleme alanındaki gelişimi hızlandırmak için oluşturulmuş başka bir veri setidir. PASCAL VOC 07 ve PASCAL VOC 12 olarak adlandırılan iki farklı seriden oluşmaktadır. PASCAL VOC 07 (Zaidi vd., 2022), 5000 eğitim ve 5000 test görüntüsünden oluşmaktadır. Bu görüntüler 12 binden fazla sınırlayıcı kutularla çevrilmiş nesnelerden oluşmaktadır. PASCAL VOC 12 (Everingham ve Winn., 2012) veri seti 5700 eğitim ve 5700 test görüntüsünden oluşmaktadır. Bu görüntüler 27 binden fazla sınırlayıcı kutularla çevrilmiş nesnelerden oluşmaktadır (Zaidi vd., 2022). Her bir PASCAL VOC serisi, araba, insan, bisiklet, otobüs, kedi, kuş, at, uçurtma, koyun, tekne, sandalye, yemek masası, kanepe gibi 20 nesne kategorisinden oluşmaktadır. Şekil 28’de PASCAL VOC veri setinden örnek görüntüler ve etiketleri gösterilmektedir.



Şekil 28. PASCAL VOC örnek veri seti (Everingham ve Winn., 2012)

Open Image (Kuznetsova vd., 2020) veri seti görüntü sınıflandırması, nesne algılama gibi görevlerde kullanılması amacıyla oluşturulmuş 9,2 milyon görüntüden oluşmaktadır. Bu görüntüler 600 kategoriden oluşmakta ve 16 milyon etiketlenmiş nesne içermektedir. Her bir görüntüde yaklaşık 8 etiketlenmiş nesne bulunmaktadır. Şekil 29’da Open Image veri setinden örnek görüntüler verilmiştir. Bu görüntüler üzerindeki nesneler sınırlayıcı kutular aracılığıyla belirlenmiş ve etiketlenmiştir.



Şekil 29. Open Image örnek veri seti (Kuznetsova vd., 2020)

Bilgisayarlı görme işleminde kullanılması amacıyla oluşturulmuş en popüler veri setlerinden birisi MS COCO veri setidir. MS COCO veri seti, görüntü sınıflandırmasının yanı sıra geniş sahne algılama sorunlarını gidermek için kullanılmaktadır (Carvalho vd., 2020). Veri setindeki çoğu görüntü ImageNet'ten farklı olarak aynı nesne gurubuna ait birden fazla nesne bulundurmaktadır. Toplamda 80 nesne kategorisi içeren yaklaşık 330.000 görüntü bulunmaktadır. Bu görüntülerden 200.000'den fazlası etiketli görüntüdür ve yaklaşık 1,5 milyon etiketli nesne bulunmaktadır (Lin vd., 2014). Şekil 30'da MS COCO veri setindeki nesne kategorilerindeki görüntüler ve her bir kategori için etiketli nesneler gösterilmektedir.



Şekil 30. MS COCO örnek veri seti (Lin vd., 2014)

### 1.9. Transfer Öğrenme

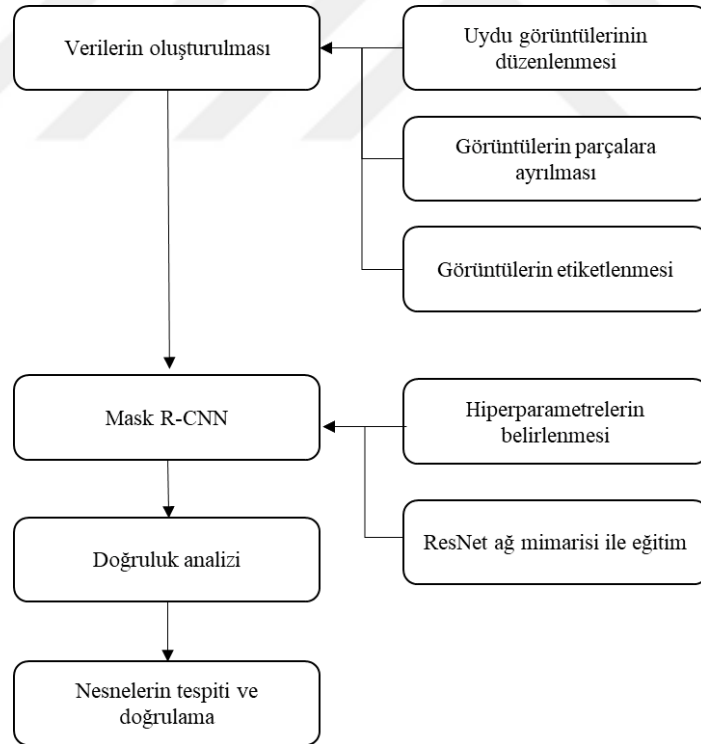
Klasik makine öğrenimi ve veri madenciliği algoritmalarındaki temel varsayım eğitim ve sonrasında kullanılacak verilerin aynı dağılıma sahip olması yönündedir (Zhu vd., 2005). Klasik makine öğrenmesi algoritmaları, uygulama yapılması istenilen alanda önceden oluşturulan etiketli veriler üzerinden matematiksel modeller kullanarak sonuçlar üretmektedir. Transfer öğrenmede ise eğitim ve test verisinde kullanılan veriler uygulama alanındaki verilerden farklı olabilmektedir (Fung vd., 2006). Örneğin, bir görüntü üzerinde sınıflandırma işlemi yapılacaksa bir veri setine ihtiyaç bulunmaktadır. Bunun öncesinde farklı bir uygulama alanındaki veri setine sahip olduğunda, bu veri setinin özelliklerini başarılı bir şekilde diğer algoritmaya transfer etmek mümkündür. Gerçek Dünya'ya bakıldığında birçok örneği görülmektedir. Örneğin; piyano çalmayı öğrenmek isteyen iki kişiyi ele alalım. Bunlardan birinin daha önce müzik çalma deneyimi olmadığını ve diğer kişinin ise gitar çalacak kapsamlı müzik bilgisi olduğunu kabul edelim. Geniş bir müzik altyapısına sahip olan kişi, daha önce öğrendiği müzik bilgilerini piyano çalmayı öğrenme görevine aktararak daha verimli bir şekilde piyano öğrenebilecektir (Pan vd., 2009). Ya da elmaların bir meyve olduğunu bilen bir kişi armutları daha kolay bir şekilde tanımlayabilecektir. Transfer öğrenme yönteminde insanların yeni fakat benzer sorunları çözerken önceden edindiği bilgileri kullanmasından ilham alınmıştır (Lu vd., 2015).

Başka bir deyişle transfer öğrenme, yeni ancak benzer sorunları çok daha hızlı ve etkili bir şekilde önceden edinilmiş bilgileri kullanarak çözebilmektedir. Yosinski vd., (2014) tarafından yapılan çalışma önceden eğitilmiş ağırlıklarını kullanarak ağı eğitmenin, derin sinir ağı performansını iyileştirmek için yararlı bir teknik olduğunu göstermektedir. Geniş uygulama olanakları sayesinde transfer öğrenme makine öğreniminde popüler ve gelecek vaat eden bir yöntem olarak kabul edilmektedir (Weiss vd., 2016). Bu sayede çok pahalı veri etiketleme görevlerinden kaçınılarak öğrenme performansı büyük ölçüde iyileştirilmektedir (Weiss vd., 2016).

## 2. YAPILAN ÇALIŞMALAR

Bu tez çalışmasında, Mask R-CNN kullanarak iki farklı senaryo gerçekleştirilmiştir. Bu senaryolardan ilkinde Sentinel-2 uydu görüntüleri kullanılarak bölge tespiti yapılmıştır. İkinci senaryoda ise WorldView-3 (WV3) uydu görüntüleri kullanılarak bina tespiti yapılmıştır. Bu senaryolar benzer adımları izlediğinden tezde bütün olarak anlatılmıştır.

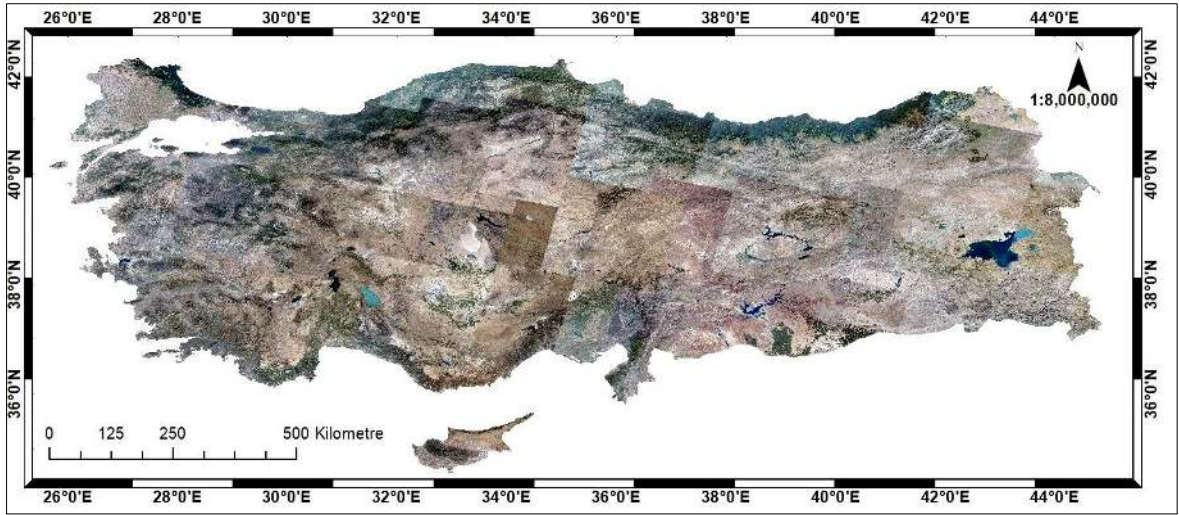
Tez çalışması aşağıda bulunan Şekil 31’deki metodolojik adımlara sahiptir. Sentinel-2 ve WV3 görüntülerinin temin edilerek; Sentinel-2 görüntüleri  $512 \times 512$ , WV3 görüntüleri  $1024 \times 1024$  boyutunda olacak şekilde görüntü parçalarına ayrılmıştır. Görüntülere etiketleme işlemi yapılarak eğitim ve test veri seti oluşturulmuştur. Hiperparametreler belirlenerek ResNet ağ mimarisi ile ağ eğitilmiştir. Doğrulama metrikleri kullanılarak değerlendirme işlemi yapılmıştır.



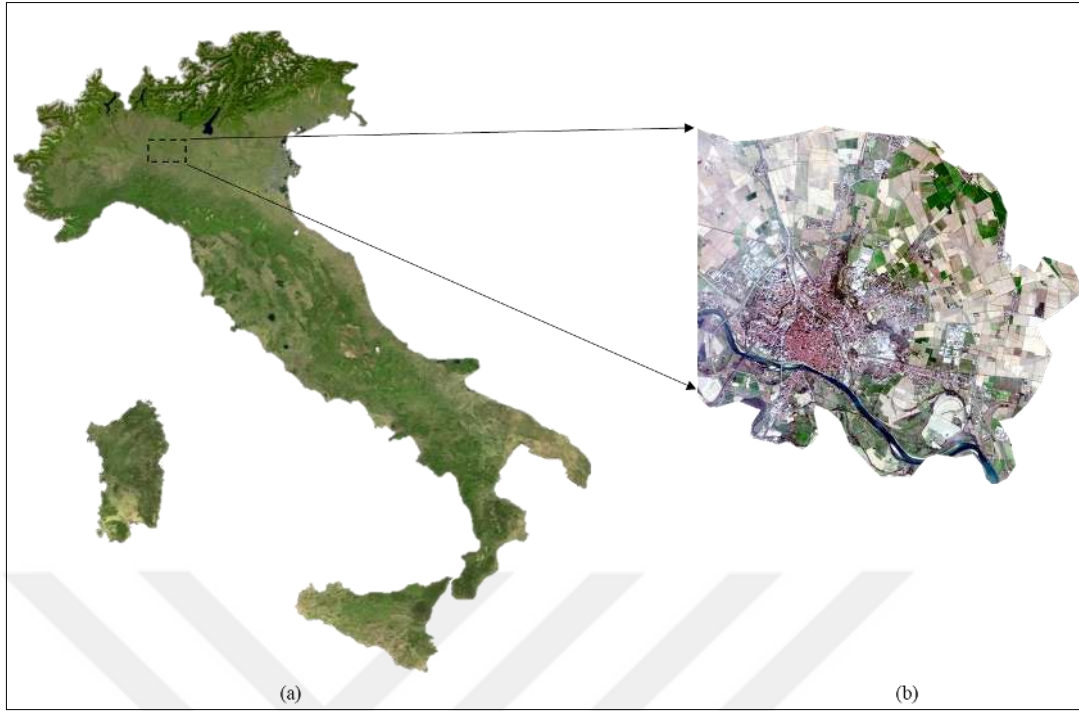
Şekil 31. Tez iş akış şeması

## 2.1. Çalışma Alanı ve Kullanılan Veri

Çalışma alanı olarak iki farklı bölge belirlenmiştir. Bunlardan ilki Türkiye’dir. Türkiye,  $36^{\circ}$  –  $42^{\circ}$  kuzey paralelleri,  $26^{\circ}$  –  $45^{\circ}$  doğu meridyenlerini arasında kalmaktadır. Şekil 32’de uydu görüntülerinden oluşturulmuş olan Türkiye haritası gösterilmektedir. Seçilen ikinci bölge ise İtalya’nın kuzeyinde kalan  $45^{\circ}$  kuzey paraleli  $09^{\circ}$  doğu meridyeni arasında bulunan Pavia şehridir. Şekil 33’te Pavia şehrinin İtalya’daki konumu gösterilmektedir.



Şekil 32. Türkiye haritası



Şekil 33. (a) İtalya haritası, (b) Pavia Şehri

Sentinel-2, Copernicus projesi kapsamında oluşturulan Dünya gözlem uydularından biridir. Sentinel-2 aynı yörüngede dolanan 2A ve 2B ikiz uydularına sahiptir. Bu sayede daha iyi kapsama alanına sahip olup daha kısa sürede döngülerini tamamlamaktadır. Sentinel-2 uydusu her beş günde bir aynı çekim şartlarında, aynı yere ait 12 spektral bant aralığında görüntü yakalamaktadır. ESA (European Space Agency) tarafından oluşturulan bu uydu 10m- 60m arasında uzamsal çözünürlüğe sahiptir. Tablo 1’de Sentinel-2 uydu görüntüsünün bant özellikleri gösterilmektedir (URL-2, 2021).

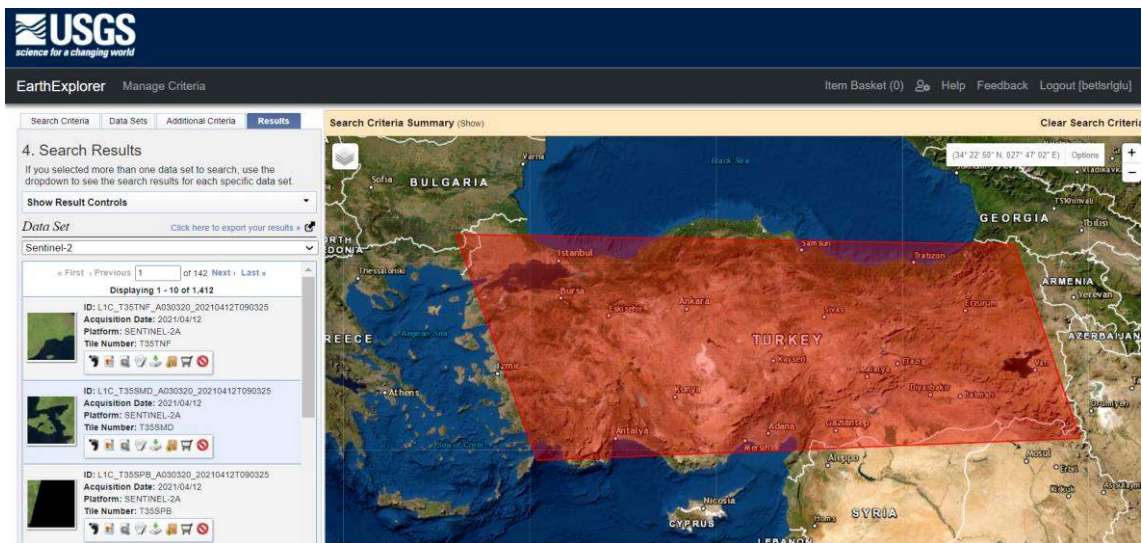
Tablo 1. Sentinel-2 uydu görüntüsü özellikleri

| Bant | Tanımlama                           | Uzamsal Çözünürlük (m) | Spektral Çözünürlük (nm) | Radyometrik Çözünürlük (bit) |
|------|-------------------------------------|------------------------|--------------------------|------------------------------|
| B1   | Ultra mavi (Kıyı ve Aerosol )       | 60                     | 443                      | 12                           |
| B2   | Mavi                                | 10                     | 490                      |                              |
| B3   | Yeşil                               | 10                     | 560                      |                              |
| B4   | Kırmızı                             | 10                     | 665                      |                              |
| B5   | Görünür ve Yakın kızıl Ötesi (VNIR) | 20                     | 705                      |                              |

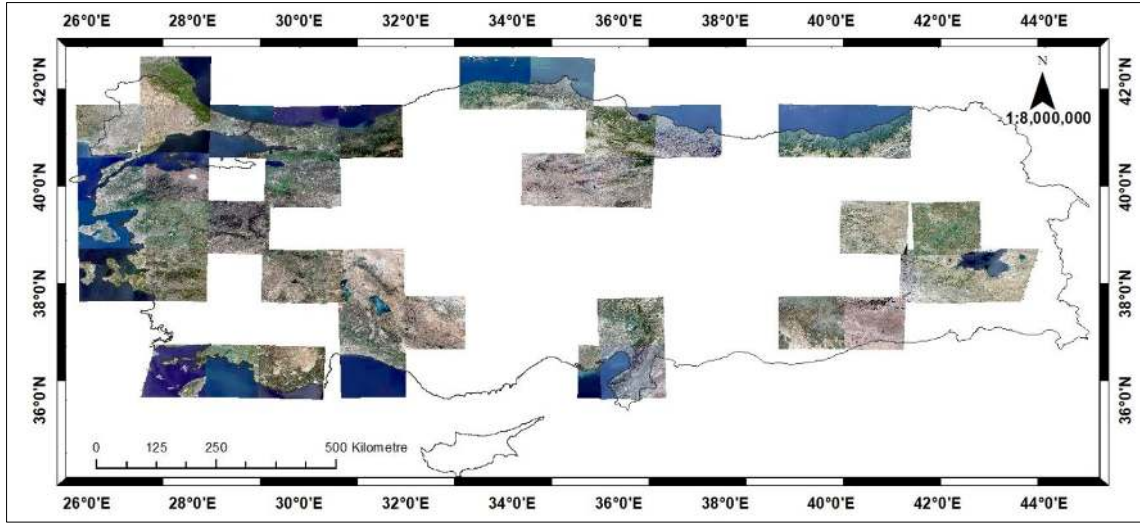
Tablo 1'in devamı

|     |                                     |    |      |    |
|-----|-------------------------------------|----|------|----|
| B6  | Görünür ve Yakın kızıl Ötesi (VNIR) | 20 | 740  | 12 |
| B7  | Görünür ve Yakın kızıl Ötesi (VNIR) | 20 | 783  |    |
| B8  | Görünür ve Yakın kızıl Ötesi (VNIR) | 10 | 842  |    |
| B8a | Görünür ve Yakın kızıl Ötesi (VNIR) | 20 | 865  |    |
| B9  | Kısa Dalga Kızıl Ötesi (SWIR)       | 60 | 940  |    |
| B10 | Kısa Dalga Kızıl Ötesi (SWIR)       | 60 | 1375 |    |
| B11 | Kısa Dalga Kızıl Ötesi (SWIR)       | 20 | 1610 |    |
| B12 | Kısa Dalga Kızıl Ötesi (SWIR)       | 20 | 2190 |    |

Çalışmada kullanılacak olan görüntüler, USGS (U.S. Geological Survey)'in resmî sitesi olan <https://earthexplorer.usgs.gov/> adresinden ücretsiz olarak indirilmiştir. İndirme işlemi yapılırken Türkiye sınır noktalarına işaretler koyulmuştur. Sonrasında zaman, bulutluluk oranı ve veri seti seçimi yapılarak arama yapılmıştır (Şekil 34). Çıkan sonuçlar incelenerek toplamda 50 görüntü indirilmiştir. İndirilen bu görüntüler sadece bir şehri ya da bir bölgeyi kapsamamaktadır. Çünkü tez çalışması Türkiye'nin genelini kapsadığından doğru veri seti oluşturmak amacıyla Türkiye'deki her bölgeyi en iyi temsil edecek şekilde görüntüler indirilmiştir. Şekil 35'te indirilen bazı Sentinel-2 uydu görüntüleri gösterilmektedir.



Şekil 34. USGS sitesi görüntü indirme modülü

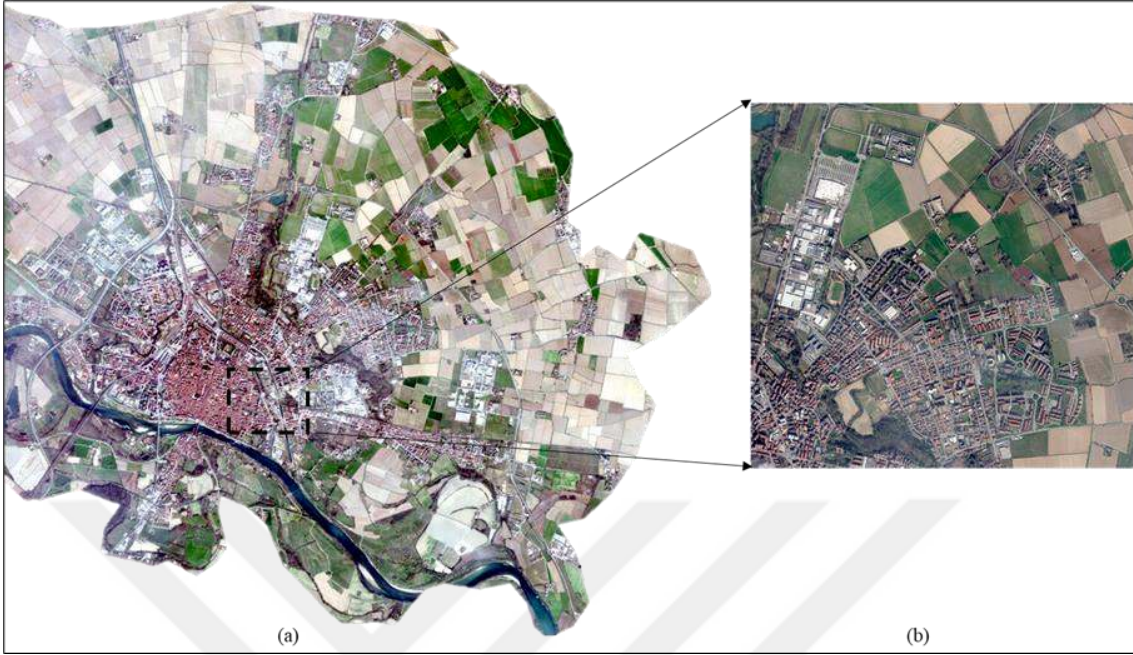


Şekil 35. Sentinel 2 uydu görüntülerinden bazı örnekler

Çalışmada kullanılan diğer bir uydu görüntüsü WorldView-3'tür. WV3 uydu görüntülerinde pankromatik, multispektral, 8 kısa dalga kızılötesi (SWIR), ve 12 CAVIS (Clouds, Aerosols, Vapours, Ice and Snow) bantları bulunmaktadır. WV3 uyduları 617 km yükseklikte bulunarak yüksek çözünürlüklü ticari uydu görüntüleri üretmektedir. WV3 uydu görüntüleri 0,31 m pankromatik, 1,24 m multispektral, 3,7 m SWIR, 30 m CAVIS uzamsal çözünürlüğüne sahiptir. WV3 uyduları Dünya üzerindeki bir tam turunu bir günde yaparak 680.000 km<sup>2</sup> yol almaktadır. Tablo 2'de WorldView-3 uydu görüntüsünün bant özellikleri gösterilmektedir (URL-3, 2022). Tez çalışmasında kullanılan WV3 uydu görüntüleri University of Pavia Laboratory of Geomatics tarafından satın alınmıştır. Şekil 36'da WV3 uydu görüntülerinden oluşturulmuş Pavia bölgesi ve çalışma alanı olarak seçilmiş olan bölge gösterilmektedir

Tablo 2. WorldView-3 uydu görüntüsü özellikleri

|                                               | Bant adı                            | Uzamsal Çözünürlük (m) | Spektral Çözünürlük (nm) | Radyo-metrik Çözünürlük (bit) |
|-----------------------------------------------|-------------------------------------|------------------------|--------------------------|-------------------------------|
| Pankromatik Bant                              |                                     | 0.31                   | 450- 800                 | 11                            |
| Multispektral Bant                            | Kıyı mavisi                         | 1.24                   | 400 - 450                |                               |
|                                               | Mavi                                |                        | 450 - 510                |                               |
|                                               | Yeşil                               |                        | 510 - 580                |                               |
|                                               | Sarı                                |                        | 585 - 625                |                               |
|                                               | Kırmızı                             |                        | 630 - 690                |                               |
|                                               | Görünür ve Yakın kızıl Ötesi (VNIR) |                        | 705 - 745                |                               |
|                                               | Görünür ve Yakın kızıl Ötesi (VNIR) |                        | 770 - 895                |                               |
|                                               | Görünür ve Yakın kızıl Ötesi (VNIR) |                        | 860 - 1040               |                               |
| Kısa Dalga Kızıl Ötesi (SWIR)                 | SWIR-1                              | 3.70                   | 1195 - 1225              | 14                            |
|                                               | SWIR-2                              |                        | 1550 - 1590              |                               |
|                                               | SWIR-3                              |                        | 1640 - 1680              |                               |
|                                               | SWIR-4                              |                        | 1710 - 1750              |                               |
|                                               | SWIR-5                              |                        | 2145 - 2185              |                               |
|                                               | SWIR-6                              |                        | 2185 - 2225              |                               |
|                                               | SWIR-7                              |                        | 2235 - 2285              |                               |
|                                               | SWIR-8                              |                        | 2295 - 2365              |                               |
| CAVIS (Clouds, Aerosols, Vapors, Ice, & Snow) | Çöl Bulutları                       | 30                     | 405 - 420                |                               |
|                                               | Aerosols-1                          |                        | 459 - 509                |                               |
|                                               | Yeşil                               |                        | 525 - 585                |                               |
|                                               | Aerosols-2                          |                        | 620 - 670                |                               |
|                                               | Su-1                                |                        | 845 - 885                |                               |
|                                               | Su-2                                |                        | 897 - 927                |                               |
|                                               | Su-3                                |                        | 930 - 965                |                               |
|                                               | NDVI-SWIR                           |                        | 1220 - 1252              |                               |
|                                               | Sirrus                              |                        | 1350 - 1410              |                               |
|                                               | Kar                                 |                        | 1620 - 1680              |                               |
|                                               | Aerosol-3                           |                        | 2105 - 2245              |                               |
|                                               | Aerosol-3                           |                        | 2105 - 2245              |                               |



Şekil 36. WV3 uydu görüntüsü (a) Pavia bölgesi, (b) Çalışma alanı

## 2.2. Çalışma Ortamı

Derin öğrenme çalışmaları büyük veri kümelerine ve yoğun hesaplama modellerine dayanmaktadır. Bu nedenle, bu tür yoğun hesaplamaları yapabilmek için grafik işlem birimleri (GPU) kullanılır. Bu tez kapsamında kendi çalışma istasyonum olan NVDI GeForce940MX üzerinden çalışma yapılmış fakat GPU yeterli gelmediği için algoritma çalışmamıştır. Bu yüzden çalışma zamanı ve ücretsiz GPU hizmeti sağlayan Google Colaboratory (Colab) kullanılmıştır. Çalışmada kullanılan Colab Tesla K80 GPU'yu hizmet olarak sunmaktadır. Colab ücretsiz çalışma zamanı en fazla 12 saat olarak tanımlanmıştır. Colab'ın çalışma saatlerine bağlı kalınarak algoritmanın çalıştırılmasına devam edilmiştir. Çalışma Colab üzerinden Jupyter Notebook kullanılarak Python dilinde yapılmış, TensorFlow, Keras, Numpy, Scipy, Pillow, Cython, Matplotlib, Scikit-image, Opencv-python, H5py, Imageaug ve Ipython kütüphaneleri kullanılmıştır.

### 2.3. Veri Setlerinin Oluřturulması

Yukarıda da belirtildiđi gibi veri setinde kullanmak amacıyla Sentinel-2 ve WV3 uydu görüntüleri seçilmiřtir. Sentinel-2 görüntüleri için ArcGIS ortamında gerçek renk değeri olacak řekilde bant-4, bant-3, bant-2 birleřtirilerek 3 bantlı görüntü oluřturulmuřtur. Birleřmiř görüntüler Matlab ortamında  $512 \times 512$  boyutunda kesilmiřtir. řekil 37’de eđitim veri setine ait bazı örnek görüntüler gösterilmektedir. WV3 uydu görüntüsünden seçilen çalıřma alanı yine Matlab ortamında  $1024 \times 1024$  boyutunda kesilmiřtir. řekil 38’de veri setine ait bazı örnekler sunulmaktadır.



řekil 37. Sentinel-2 veri seti örnekleri

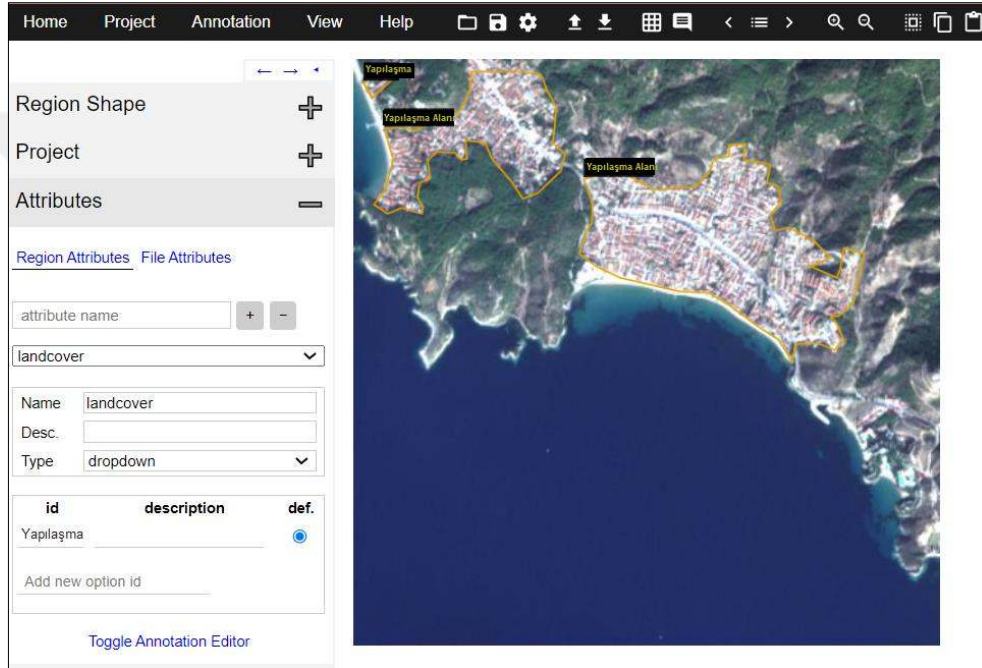


Şekil 38. WV3 veri seti örnekleri

Mask R-CNN algoritması denetimli sınıflandırmayı temel aldığından görüntülere ve görüntü etiketlerine ihtiyaç vardır. Görüntüdeki uzamsal bölgelerin tanımlanması, segmentlerin betimlenmesi için ücretsiz manuel açıklama aracı olan VGGImage Annotator (VIA) kullanılmıştır (URL-4, 2022). Herhangi bir kurulum gerektirmeden web tarayıcısında çevrimdışı bir uygulama olarak çalışmaktadır. Tanımlanan bölgeler daire, elips, çokgen, nokta ve poligon olabilmektedir. Çalışmadaki nesnelerin sınırlayıcı alanlarını belirlemek için poligon seçimi kullanılmıştır.

Sentinel-2 uydu görüntüsündeki yapılaşma alanlarının tespit edilmesi amacıyla veri seti oluşturulurken toplamda 600 görüntü kullanılmıştır. Bu görüntülerden %80'i eğitim işleminde kullanılması için 'train' klasörüne, %20'si test aşamasında kullanılması için 'val'

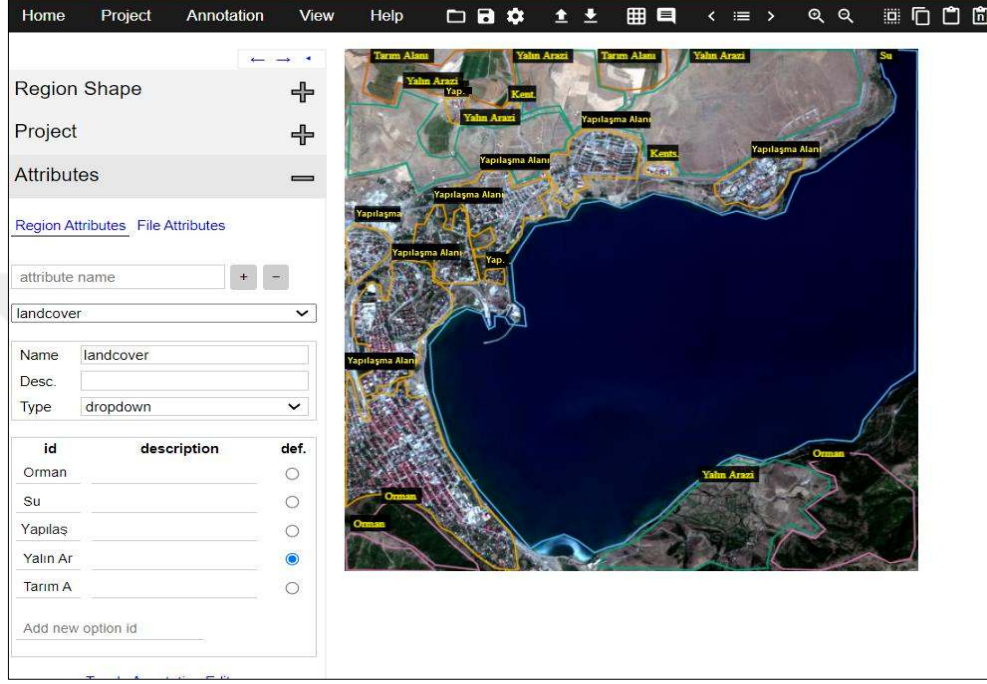
klasörüne rastgele olacak şekilde eklenmiştir. Sonrasında her bir klasördeki görüntülerin nesne tespitini yapabilmek amacıyla VIA kullanılmıştır. Etiketleme işlemi yapılırken etiket adı olarak ‘Yapılaşma Alanı’ seçilmiştir. Şekil 39’da VIA ara yüzü üzerinde görüntüde yapılan etiketleme işlemi gösterilmektedir. Etiketleme işlemi 60 saat sürmüştür. Toplamda 600 görüntü etiketlenmiştir. Eğitimde kullanılan görüntü sayısı 480, etiketlenen toplam nesne sayısı ise 5154’tür. Test için kullanılan toplam görüntü sayısı 120, bu görüntüler içerisindeki etiketlenen toplam nesne sayısı ise 663’tür.



Şekil 39. VGG Image Annotator’da görüntü etiketleme

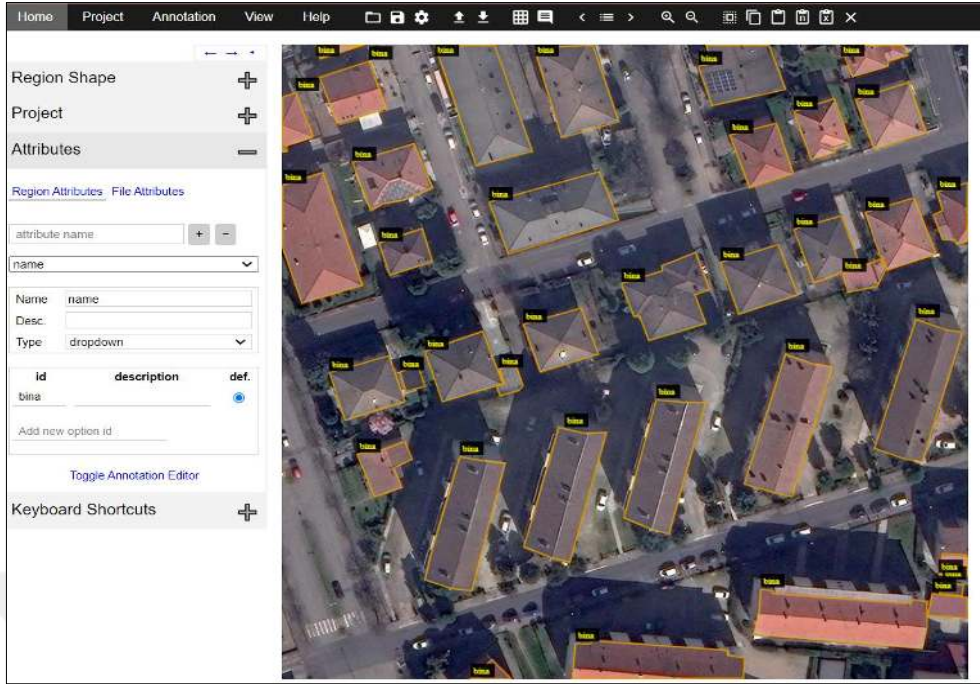
Sentinel-2 uydu görüntüsünde arazi örtüsünün sınıflandırılması amacıyla beş sınıf belirlenerek etiketleme yapılmıştır. Bu sınıflar; “Yapılaşma Alanı”, “Tarım Alanı”, “Yalın Arazi”, “Orman” ve “Su”dur. Veri seti içerisindeki tüm görüntüler etiketleme aracı olan VIA kullanılarak yapılmıştır. Etiketleme işlemi toplamda 120 saat sürmüştür. Şekil 40’da VIA üzerinden bir görüntüye ait yapılan etiketleme işlemi gösterilmektedir. Veri seti hazırlanırken toplamda 835 görüntü etiketlenmiştir. Eğitimde kullanılacak görüntü sayısı 700’dür ve içerdiği toplam nesne sayısı 9953’tür. Test işleminde kullanılan görüntü sayısı 135’tir ve içerdiği toplam nesne sayısı 1804’tür. Bu görüntülerin içerisinde beş sınıfı da temsil edecek etiketler bulunmaktadır. Eğitimde kullanılması amacıyla tarım alanını temsil eden 990 nesne, yalın araziyi temsil eden 1330 nesne, orman alanını temsil eden 1242 nesne,

yapılaşma alanlarını temsil eden 6115 nesne ve su alanını temsil eden 276 nesne etiketlenmiştir. Test aşamasında kullanılması amacıyla tarım alanını temsil eden 161 nesne, yalın araziyi temsil eden 129 nesne, orman alanını temsil eden 254 nesne, yapılaşma alanını temsil eden 1186 nesne ve su alanını temsil eden 74 nesne etiketlenmiştir.



Şekil 40. Sentinel-2 görüntüsünde beş sınıf için etiketleme işlemi

WV3 uydu görüntüsü üzerinden bina tespiti yapılması amacıyla VIA kullanılarak bina etiketleri oluşturulmuştur. Şekil 41’de VIA üzerinden yapılan etiketleme işlemi gösterilmektedir. Toplamda 80 görüntü etiketlenmiştir. Etiketleme işlemi 16 saat sürmüştür. Eğitimde kullanılan görüntü sayısı 56, eğitim için etiketlenen toplam nesne sayısı 870’tir. Test işleminde kullanılan görüntü sayısı 24, değerlendirme için etiketlenen nesne sayısı 367’dir.



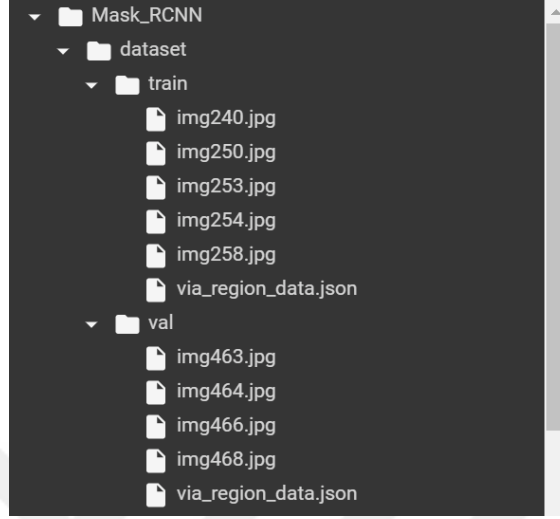
Şekil 41. WV3 görüntüsünde etiketleme işlemi

Veri集中的 tüm görüntülerin etiketlemesi yapıldıktan sonra veriler .json uzantılı olacak şekilde dışarıya aktarılmıştır. Şekil 42’de örnek bir .json uzantılı dosyanın içeriği görülmektedir. Bu dosya içerisinde her bir veriye ait görüntü ismi, boyutu; her poligonun ait olduğu sınıf ismi ve koordinat bilgileri bulunmaktadır.

```
{
  "img10.jpg39627": {
    "filename": "img10.jpg",
    "size": 39627,
    "regions": [
      {
        "shape_attributes": {
          "name": "polygon",
          "all_points_x": [
            50,
            76,
            78,
            53
          ],
          "all_points_y": [
            434,
            433,
            465,
            461
          ]
        },
        "region_attributes": {
          "landcover": "Yapılaşma Alanı"
        }
      }
    ]
  }
}
```

Şekil 42. Veri setine ait örnek bir .json dosyası

Şekil 43'te etiketleme işlemi sonrası oluşturulan veri seti klasör örneği gösterilmektedir.



Şekil 43. Veri seti klasör örneği

## 2.4. Mask R-CNN Uygulaması

Mask R-CNN uygulaması 2017 yılında Matterport mühendislik ekibi tarafından yayınlanmıştır (URL-5, 2020). Bu uygulama Python dili kullanılarak yazılmıştır. Açık kaynak kodlu derin öğrenme kütüphanesi olan Keras ve vektörleştirilmiş işlemleri hızlı gerçekleştiren, yüksek hızlı paralel hesaplamalar yapabilen açık kaynak kodlu Tensorflow kullanılmaktadır.

### 2.4.1 Kod Yapısı

Mask R-CNN uygulaması kendi içinde dört ana modülden oluşmaktadır.

1. Utilis: Özel görevler için gerekli olan tüm fonksiyonlar bu kısımdadır.
2. Config: Eğitim işlemi için gerekli olan çeşitli hiperparametreler bu kısımda yer almaktadır. Hiperpareterlerin optimum değerleri olmamakla birlikte veri seti değişikçe en iyi sonuç elde edilene kadar değiştirilmesi gerekmektedir. Tez kapsamında kullanılan hiperparametrelerin tanımları ve kullanım şekilleri aşağıdaki gibidir.

- NAME: Birden fazla işlem yapıldığı zaman ayırt etmeyi kolaylaştırılacak özel addır.
  - GPU\_COUNT: Kullanılacak GPU sayısı bu kısımda belirlenir.
  - IMAGES\_PER\_GPU: Her GPU için eğitilecek görüntü sayısıdır. En küçük grup boyutu GPU\_COUNT \* IMAGES\_PER\_GPU olarak hesaplanır.
  - STEPS\_PER\_EPOCH: Her epokta yapılacak eğitim adım sayısıdır.
  - BACKBONE: Ağ mimarisinin seçildiği bölümdür. ResNet50 ve ResNet101 ağ mimarisini desteklemektedir.
  - NUM\_CLASSES: Etiketlenen sınıf sayısının yazıldığı kısımdır.
  - USE\_MINI\_MASK: Maske boyutlarını daha küçük boyutlara indirmek için kullanılır. Yüksek çözünürlüklü görüntüler kullanırken bellek yükünü azaltmak için önerilir.
  - MINI\_MASK\_SHAPE: Yeniden boyutlandırılması istenilen maskenin boyutlarının girildiği kısımdır.
  - DETECTION\_MIN\_CONFIDENCE: Algılanan bölgenin kabul edilmesi için gereken güven eşiğidir.
  - VALIDATION\_STEPS: Doğrulama için istenilen adım sayısıdır.
  - RPN\_ANCHOR\_SCALES: Pikselleri tanımlamak için kullanılan çerçevenin kenar uzunluğudur.
  - LEARNING\_RATE: Algoritmanın öğrenme hızı bu kısımda belirlenir.
3. Model: Mask R-CNN ağ mimarisi bu kısımda yazılmıştır.
  4. Custom: Mask R-CNN'i kendi veri setimize göre kişiselleştirme işlemi bu kısımda yapılmıştır.
    - Veri Setinin Tanımlanması: Etiketleme işleminde nesneler için tanımlanan etiketler uygulamaya tanıtılmıştır. Şekil 44'te örnek veri seti tanımlatma işlemi gösterilmektedir.

```
class CustomDataset(utils.Dataset):
    def load_custom(self, dataset_dir, subset):
        self.add_class("object", 1, "Yapılaşma Alanı")
```

Şekil 44. Veri seti tanımlatma

- Eğitim İşlemi: Model için uygulanacak olan eğitim işlemi bu kısımda belirlenmiştir. Şekil 45'te örnek eğitim işlemi gösterilmektedir.

**dataset\_train:** Eğitim veri kümesi

**dataset\_val:** Doğrulama veri kümesi

**learning\_rate:** Öğrenme hızı

**epochs:** uygulanacak epok sayısı

**layers:** Eğitimde kullanılacak katmanlar burada belirlenir.

- Heads: Ağ yapısının başını temsil eder.
- 4+: ResNet ağını 4 katman ve sonrasında itibaren eğitir.
- all: Tüm katmanları eğitir.

```
from keras.callbacks import Callback

model.train(dataset_train, dataset_val,
            learning_rate=config.LEARNING_RATE,
            epochs=50,
            layers='heads')

model.train(dataset_train, dataset_val,
            learning_rate=config.LEARNING_RATE,
            epochs=110,
            layers='4+')

model.train(dataset_train, dataset_val,
            learning_rate=config.LEARNING_RATE/10,
            epochs=200,
            layers='all')
```

Şekil 45. Eğitim parametrelerinin belirlenmesi

Tez çalışması kapsamında örnek segmentasyon Mask R-CNN aracılığıyla yapılmıştır. Bölge tespiti ve bina tespiti olmak üzere iki farklı senaryo üzerinde çalışılmıştır. Bölge tespiti yapılırken Sentinel-2 uydu görüntüleri üzerinden 3 farklı işlem gerçekleştirilmiştir. İlk olarak ResNet50 ağ mimarisi kullanılarak ağ eğitilmiştir. İkinci olarak ResNet101 ağ mimarisi ile ağ eğitimi gerçekleştirilmiştir. Son olarak arazi örtüsünün sınıflandırılması amacıyla beş sınıf üzerinde sınıflandırma yapılmıştır.

Tez çalışmasının ikinci kısmında yüksek çözünürlüklü uydu görüntüsü olan WV3 uydu görüntüleri kullanılarak bina tespiti yapılmıştır.

Yukarıda bahsedilen çalışmalarda optimum sonuçlar elde etmek için birçok hiperparametre değeri denenmiştir. Bu işlemler toplamda 200 saat sürmüştür. Denemeler sonucunda elde edilen en iyi hiperparametre değerleri aşağıdaki tabloda gösterilmektedir.

Tablo 3. Modeller için hiperparametreler

|                                 | Sentinel-2 ile Bölge Tespiti |                  |                | WV3 ile Bina Tespiti |
|---------------------------------|------------------------------|------------------|----------------|----------------------|
|                                 | I.                           | II.              | III.           | I.                   |
| <b>Steps_per_epoch</b>          | 100                          | 100              | 100            | 100                  |
| <b>Backbone</b>                 | ResNet101                    | ResNet50         | ResNet101      | ResNet101            |
| <b>Num_classes</b>              | 1+1                          | 1+1              | 5+1            | 1+1                  |
| <b>Detection_min_confidence</b> | 0.6                          | 0.6              | 0.6            | 0.6                  |
| <b>Validation_steps</b>         | 50                           | 50               | 50             | 50                   |
| <b>Rpn_anchor_scales</b>        | (8,16,32,64,128)             | (8,16,32,64,128) | (4,8,16,32,64) | (8,16,32,64,128)     |
| <b>Learning_rate:</b>           | 0.001&0.0001                 | 0.001 & 0.0001   | 0.001&0.0001   | 0.001 & 0.0001       |

Hiperparametreler belirlendikten sonra eğitim parametreleri belirlenmiştir. Eğitim işlemi yapılırken MS COCO veri setinin önceden eğitilmiş ağırlıkları kullanılmıştır ve eğitim işlemi tüm modeller için üç aşamalı olacak şekilde gerçekleştirilmiştir. İlk olarak “learning\_rate” config dosyasında olduğu gibi 0.001 olarak alınmış ve “layers: heads” seçilerek sadece ağırlık baş kısmına ve her model için belirlenen epok sayısına göre eğitim yapılmıştır. İkinci olarak; “learning\_rate” yine 0.001 seçilmiş ve ResNet ağına 4’ncü ve diğer katmanlar işleme katılarak her model için belirlenen epok sayısına göre eğitim yapılmıştır. Son olarak “learning\_rate” 0.0001 seçilerek tüm katmanlarda belirlenen epok sayısına göre eğitim yapılmıştır.

### 2.4.2. Mask R-CNN’de Sınıflandırma Adımları

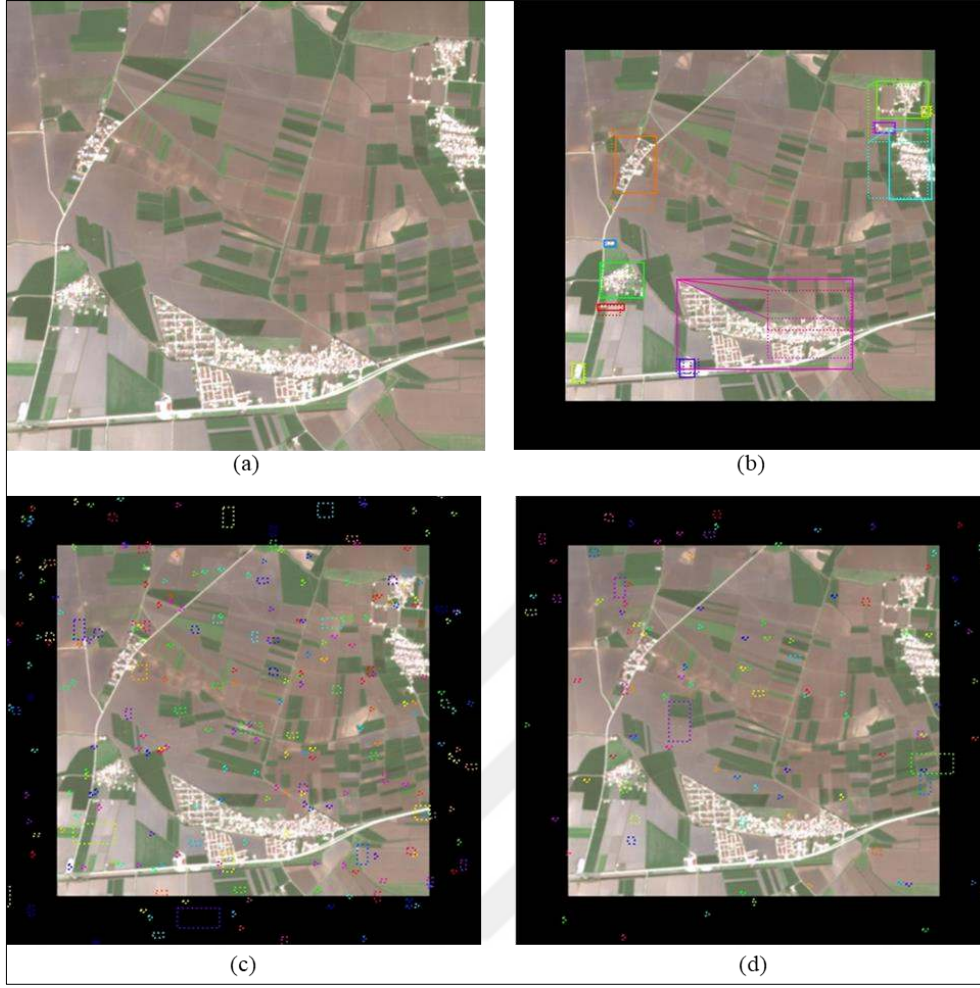
Mask R-CNN’de sınıflandırma iki aşamada gerçekleştirilmiştir. İlk aşamada nesne sınırlayıcı çapaları (*anchors*) önermek için bölge teklif ağı (RPN) kullanılmıştır. İkinci aşamada RoIAlign katmanı nesnenin maskesini belirlemek için tamamen bağlı katmana gönderilmiştir (He vd., 2017).

RPN, görüntü üzerinde birçok çapada ikili sınıflandırıcı çalıştırmaktadır. RPN, nesne/nesne yok puanlarını hesaplatır, yüksek nesne puanına sahip pozitif çapaları (*positive anchors*) sınıflandırılmak üzere ikinci aşamaya geçirir.

RPN hedefleri (*RPN Targets*), RPN için eğitim değerlerini ifade etmektedir. Hedefleri oluşturmak için farklı ölçeklerdeki çapalar ile tüm görüntü üzerinde tekliflerde bulunulur. Bu kutuların IoU değerleri yer gerçeği nesnesi (ground truth object) kullanılarak hesaplanmaktadır.  $\text{IoU} \geq 0,7$  değerine sahip olan nesneler pozitif,  $\text{IoU} \leq 0,3$  değerine sahip olan nesneler ise negatif olarak kabul etmiştir. Bu iki değerin arasında kalmış olan kutular ise tarafsız olarak kabul edilerek eğitim işlemine katılmamışlardır.

Şekil 46’da örnek bir görüntü üzerinde toplamda 2000 tane sınırlayıcı kutu ile teklifler gerçekleştirilmiştir.  $\text{Target\_rpn\_match} = 1$  olduğu kutular pozitif (b),  $\text{Target\_rpn\_match} = -1$  olduğu kutular negatif (c) ve  $\text{Target\_rpn\_match} = 0$  olduğu kutular nötr (d) olarak kabul edilmiştir.

2000 sınırlayıcı teklifinde çakışan bazı teklifleri filtrelemek için maksimum engelleme yapılmıştır. Maske dalı, en yüksek puana sahip 100 ve 50 algılama kutusunda çalışmıştır (Şekil 47).



Şekil 46. RPN ile sınırlayıcı kutu belirleme a) orijinal görüntü, b) pozitif sınırlayıcı kutu c) negatif sınırlayıcı kutu d) nötr sınırlayıcı kutu



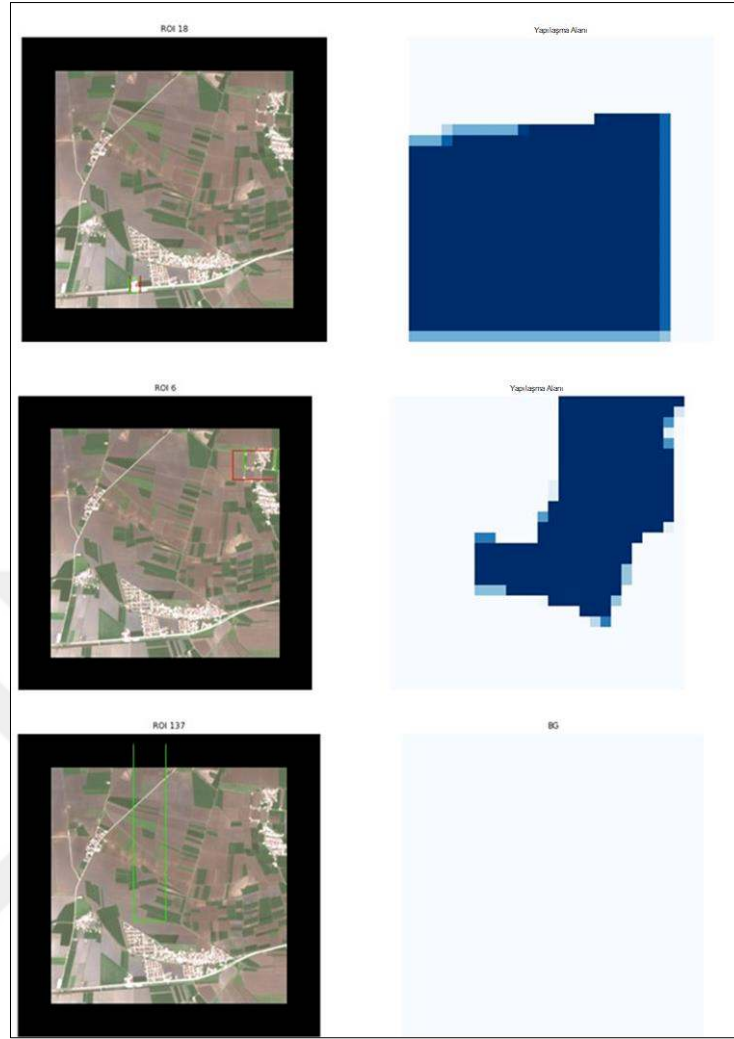
Şekil 47. Pozitif sınırlayıcı kutular (a) pozitif 100 sınırlayıcı kutu, (b) pozitif 50 sınırlayıcı kutu

Seçilen kutular arasında en yüksek puana sahip kutular tekrardan seçilerek ait oldukları sınıfa ait etiket bilgileri atanmıştır (Şekil 48).



Şekil 48. En iyi sınırlayıcı kutular

Maske R-CNN’de RoI’ye ve özellik haritasındaki özellik katmanının boyutuna bağlı olarak bir kutu belirlenmiştir. Bu kutulardaki nesnelerin tespit edilmesi ve maske oluşturulması için RoIAlign katmanı kullanılmıştır. Her örnekleme noktasının değeri özellik haritasının yakınındaki ızgara noktalarından çift doğrusal enterpolasyon yapılarak hesaplanmıştır. Şekil 49’da rastgele seçilen RoI’lerden RoIAlign kullanılarak nesne özelliklerinin çıkarılması gösterilmektedir. Şekil 50’de görüntü üzerinde maske oluşumu gösterilmektedir. Bu işlem tüm RoI’ler üzerinde yapılarak görüntünün tamamı sınıflandırılmıştır.



Şekil 49. RoI üzerinde maske oluşturma



Şekil 50. Görüntü üzerinde maske oluşumu

## 2.5. Sınıflandırma Sonuçları

### 2.5.1. Bölge Tespiti Sınıflandırma Sonuçları

Şekil 51’de Sentinel-2 Uydu görüntüsü ve ResNet101 ağ mimarisi ile oluşturulmuş model kullanılarak sadece yapılaşma alanlarını tespit etmek için yapılan sınıflandırma sonuçları görülmektedir. Kırmızı bölgeler yapılaşma alanlarını temsil etmektedir.



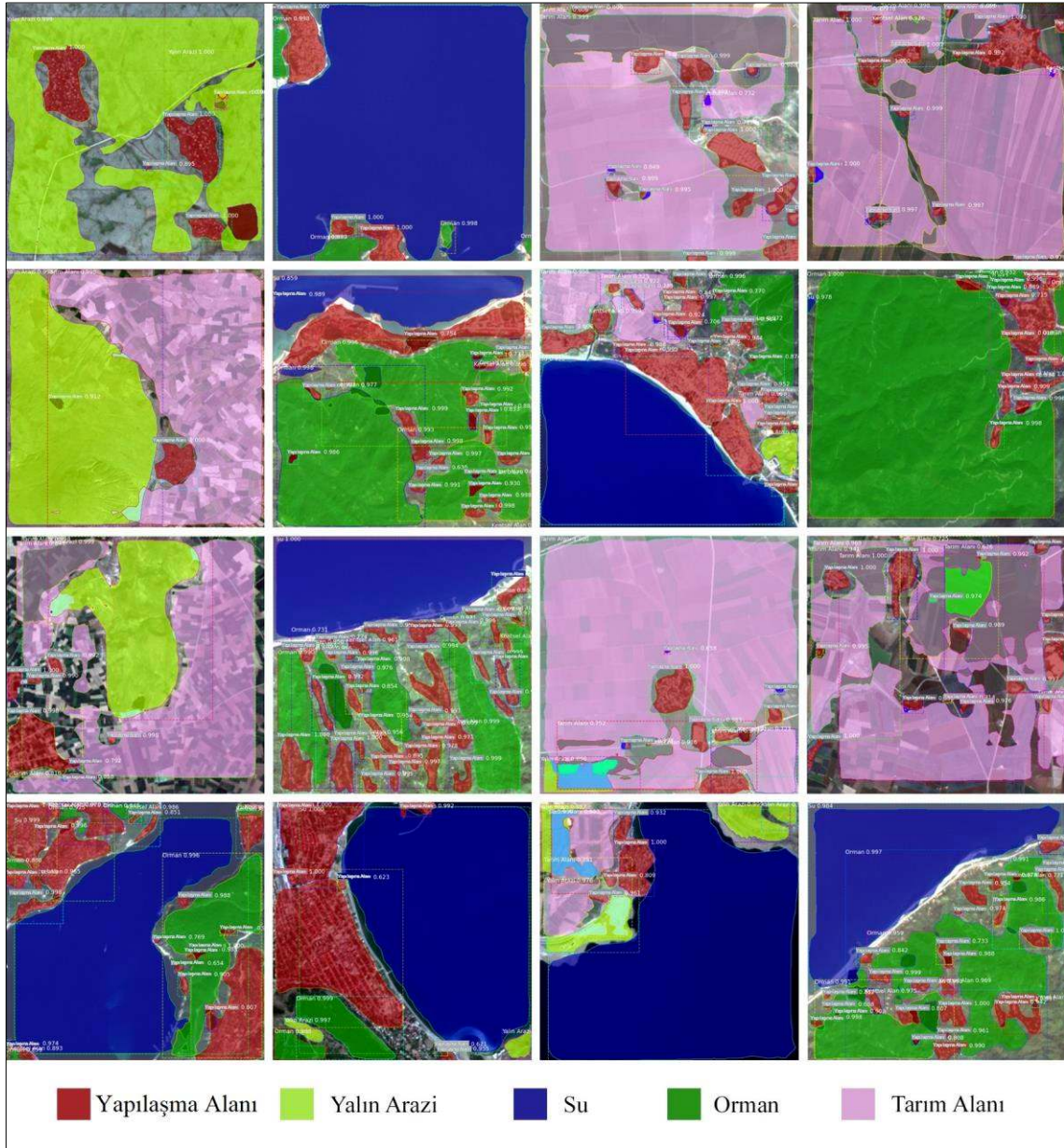
Şekil 51. ResNet101 ile oluşturulmuş modelin sınıflandırma sonuçları

Şekil 52’de Sentinel-2 Uydu görüntüsü ve ResNet50 ağ mimarisi ile oluşturulmuş model kullanılarak sadece yapılaşma alanlarını tespit etmek için yapılan sınıflandırma sonuçları gösterilmektedir. Kırmızı bölgeler yapılaşma alanlarını temsil etmektedir.



Şekil 52. ResNet50 ile oluşturulmuş modelin sınıflandırma sonuçları

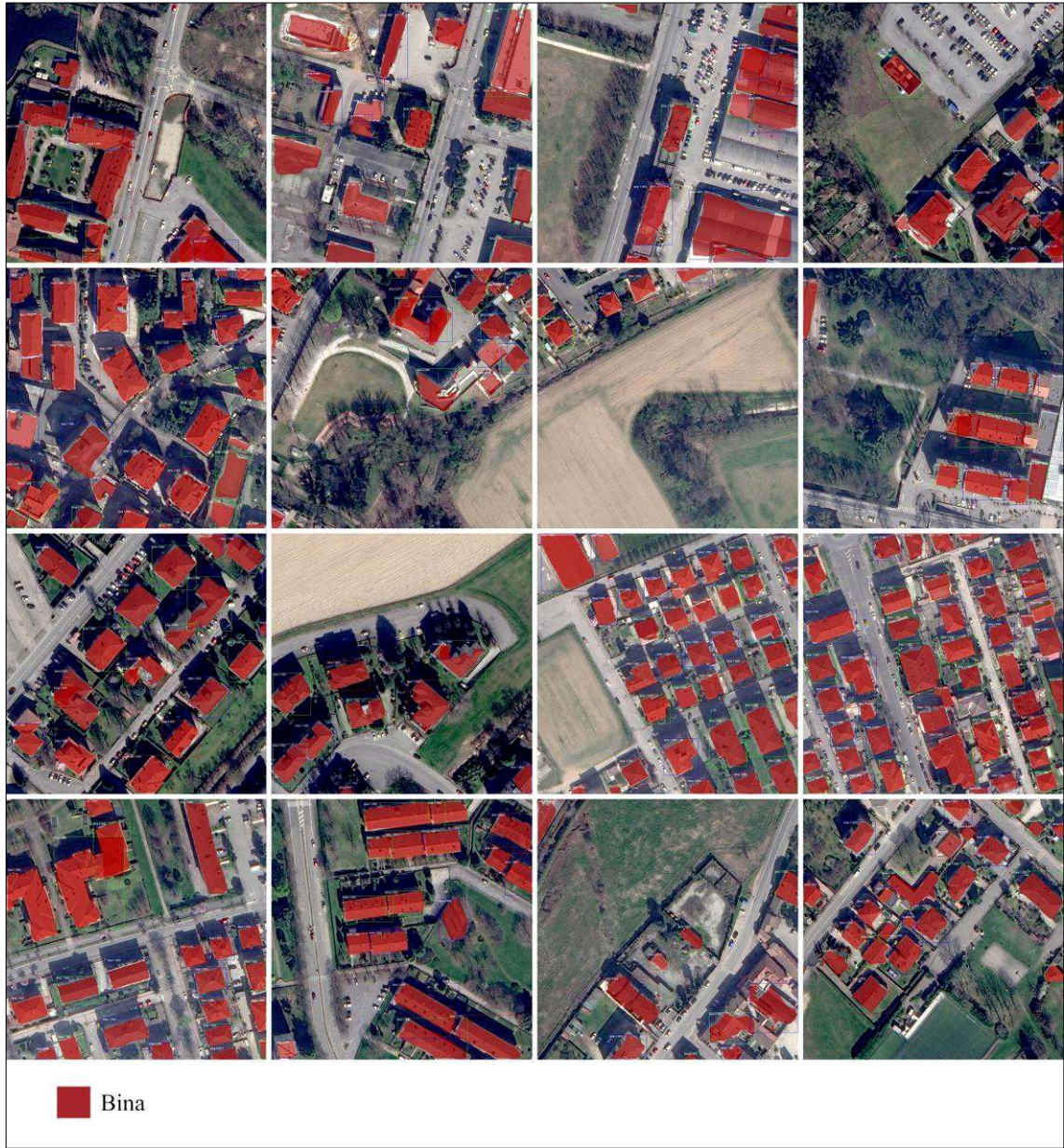
Şekil 53'te Sentinel-2 görüntüsü ile arazi örtüsünü sınıflandırmak amacıyla 5 sınıflı oluşturulan modelin sınıflandırma sonuçları gösterilmektedir. Kırmızı bölgeler yapılaşma alanı, sarı bölgeler yalın araziyi, mavi bölgeler suyu, yeşil bölgeler ormanı ve pembe bölgeler tarım alanlarını temsil etmektedir.



Şekil 53. Sentinel-2 görüntüsünde arazi örtüsü sınıflandırma sonuçları

### 2.5.2. Bina Tespiti Sınıflandırma Sonuçları

Şekil 54'te WV3 görüntüsü üzerinden bina tespiti için oluşturulmuş olan modelin sınıflandırma sonuçları gösterilmektedir. Kırmızı alanlar binayı temsil etmektedir.



Şekil 54. WV3 uydu görüntüsünün sınıflandırma sonuçları

### 3. BULGULAR VE İRDELEME

Derin öğrenme, artan başarı dereceleriyle beraber çok çeşitli problemleri çözmek için birçok alanda uygulanmıştır. Bu tez kapsamında oluşturulan model, Matterport tarafından üretilen Mask R-CNN uygulamasına dayanmaktadır. Mask R-CNN sitesinden indirilerek veri setine ve probleme uygun şekilde ayarlanarak kullanılmıştır.

Derin öğrenme modellerinin eğitimi zaman alıcıdır. Eğer bu durumdan kaçınmak mümkünse modellerin sıfırdan eğitilmemesi önerilmektedir. İlgili araştırmacılar tarafından sağlanan önceden eğitilmiş model ağırlıklarını kullanmak eğitim sırasında zamandan tasarruf ederek daha iyi sonuçlar elde edilmesini sağlamaktadır. Mask R-CNN, MS COCO veri setini ücretsiz bir şekilde temin ederek kullanabilmektedir. MS COCO veri seti 80 nesne kategorisi içeren yaklaşık 330.000 görüntüye sahip bir veri kümesidir. Bu görüntülerden 200.000'den fazlası etiketlidir ve yaklaşık 1,5 milyon nesne örneğine sahiptir. Tez çalışmasında zamandan tasarruf edebilmek ve daha iyi sonuçlara ulaşabilmek için MS COCO veri setinin ağırlıkları kullanılmıştır.

Maske R-CNN, ResNet50 ve ResNet101 ağ mimarilerini destekler. Bu tez kapsamında iki ağ mimarisi de denenmiş olup en iyi sonuçlar ResNet101'den elde edilmiştir. ResNet101 daha derin bir ağ yapısına sahip olduğu için GPU ihtiyacını artırmıştır. Bu tez çalışmasında sinir ağlarını eğitmek amacıyla NVDI GeoForce940MX ve NVDI GeoForceGTX1650 GPU'lara sahip olan bilgisayarlar kullanılmaya çalışılmıştır. NVDI GeoForce940MX yoğun sinir ağlarına sahip olan bu ağı çalıştıramamış ve "GPU' da bellek kalmadı (GPU- run out of memory)" hatası vermiştir. NVDI GeoForceGTX1650 ise ağı belli bir kısmını çalıştırmayı başaramamış ama belli bir süre sonra "GPU'da bellek kalmadı" hatası vermiştir. Yüksek GPU'lara sahip bilgisayarlar edinmek maliyetli olacağından Colab kullanılmıştır.

Colab, 12 GB NVIDIA Tesla K80 GPU'a sahip olup ücretsiz bir şekilde kullanılmıştır. Colab, 12 saat süreler ile çalışmaya izin vermektedir. Colab'ı uzun süreli hesaplamalar için kullanan kullanıcılar ile yakın zamanda Colab'da daha fazla kaynak kullanmış olan kullanıcıların GPU'lara erişimleri kısıtlanabilmektedir. Bu tez çalışmasında Colab üzerinden birçok model denemesi yapıldığından ilk başlarda 7, 8 saat kullanılabilirken Colab son zamanlarda 3 saat kadar kullanılabilmektedir. Colab, sanal makinelere bağlı olduğundan belli bir süre (yaklaşık yarım saat) herhangi bir işlem yapılmadığında otomatik olarak bağlantıyı kesebilmektedir. O yüzden tüm model denemelerinde bilgisayarın başında beklenip belli

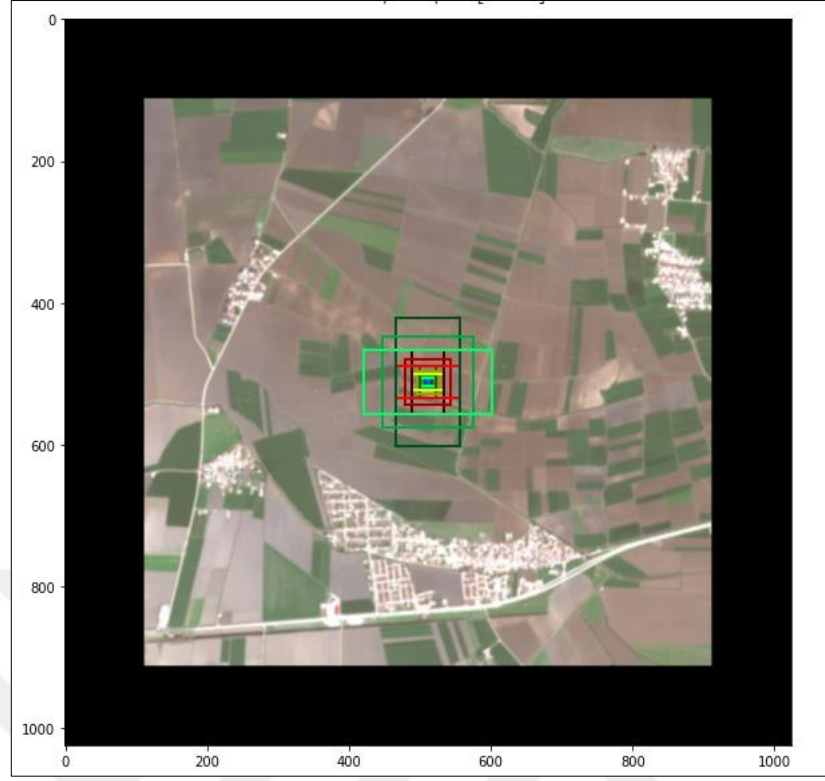
aralıklarla not defterine tıklama yapılmıştır. Colab aynı zamanda kullanıcının kontrolünü sağlamak adına zaman zaman “Robot musun?” sorusunu yöneltmekte ve eğer işaretlenme yapılmazsa GPU kullanımını sınırlandırmaktadır. Bu durumlar her ne kadar Colab kullanımında ücretsiz GPU sağlanması açısından işe yarasada, eğitim aşamasında sürekli bilgisayarın başında durmayı zorunlu kıldığından zaman kaybına neden olmuştur.

### 3.1. Model Eğitimi ve Deneysel Sonuçlar

Model eğitimi için etiketli veri oluşturma işleminde VGG Image Annotator (VIA) yazılımı kullanılmıştır. VIA, bilgisayar üzerinde herhangi bir kurulum yapılmasına gerek olmadan internet üzerinden çalışmaktadır. Sentinel-2 uydu görüntüleri 512×512, WV3 uydu görüntüleri 1024×1024 boyutlarında olacak şekilde kesilerek .jpg formatında kaydedilmiştir. Sonrasında VIA yazılımında açılarak her nesne için etiketleme işlemi yapılmıştır. Etiketleme işlemi toplamda 196 saat sürmüştür. Etiketleme işlemi sonucunda json dosyaları oluşturularak modele tanıtılmıştır.

Mask R-CNN kütüphanesi Colab üzerine eklenerek model oluşturulmaya başlanmıştır. Colab’da hafıza olarak 64 GB alan tanımladığından Google Drive yerine Colab Sanal bilgisayar kullanılmıştır. Colab sanal bilgisayarında bağlantı kesilmesi durumunda verileri kaydetmemekte ve işlemlerin tekrardan yapılması gerekmektedir.

Model’de en iyi eğitim parametrelerine ulaşabilmek için birçok deneme yapılmıştır. Mask R-CNN kendi kütüphanesinden otomatik olarak yüklenen “config” dosyasında RPN\_ANCHOR\_SCALES: (32, 64, 128, 256, 512) olarak tanımlanmıştır. İlk olarak bu ayar denendiğinde nispeten daha küçük olan nesnelerin tespit edilmediği fark edilmiştir. O yüzden RPN\_ANCHOR\_SCALES: (8, 16, 32, 64, 128) olarak seçilmiştir. Şekil 55’te rastgele bir görüntünün üzerinde çapalar (*anchors*) gösterilmiştir.



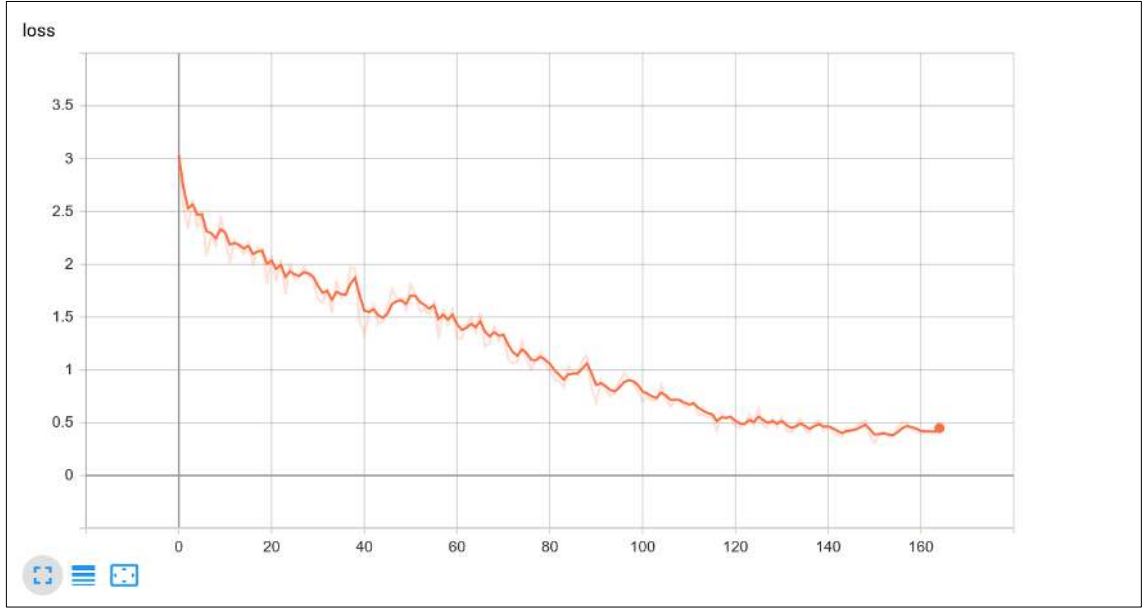
Şekil 55. Kutu boyutları

STEPS\_PER\_EPOCH her epokta yapılacak olan eğitim adım sayısını ifade etmektedir. Doğrulama istatistikleri her epok sonunda hesaplanıp TensorBoard güncellemeleri kaydedilmiştir. Bu nedenle epok sayısının küçük seçilmesi zaman ve saklama alanı açısından önerilmemektedir. Tez çalışmasında epok değeri olarak 1000 ve 100 değerleri seçilmiştir. Her epoktaki adım sayısı 1000 seçildiğinde TensorBoard’da oluşan grafiklerin açıklayıcı olmadığı fark edilmiştir. Eğitim işleminin bitirilmesi gerektiğinde 1000 adım beklemek gerekmiştir. Eğitim işlemindeki doğruluk değerlerinin daha iyi anlaşılması için tez çalışmasında her epoktaki adım sayısı (STEPS\_PER\_EPOCH) 100 seçilmiştir.

Bu tez çalışmasında iki farklı senaryo üzerinde çalışılmıştır. Bunlar; bölge tespiti ve bina tespittir. Bölge tespiti yapılırken üç farklı çalışma yapılmıştır. İlk çalışma Sentinel-2 uydu görüntüsü üzerinde ResNet101 ağ mimarisi kullanılarak yapılmıştır. Bu modelin eğitimi için; “learning\_rate” 0.001 ve “layers: heads” seçilerek 50 epok uygulanmıştır. İkinci olarak; “learning\_rate” yine 0.001 seçilmiş ve ResNet ağında 4’ncü ve diğer katmanlar işleme katılarak 60 epok uygulanmıştır. Son olarak “learning\_rate” 0.0001 seçilerek tüm katmanlarda 100 epok eğitim yapılmıştır. Toplamda 200 epok yapılması planlanırken eğitim işlemi anlık olarak TensorBoard’da izlenmiş ve değerlerde iyileşme olmadığı için 160.

epokta durdurulmuştur. En iyi sonuçlar 120. epokta elde edilmiştir. Sadece bu modelin eğitimi 18 saat sürmüştür.

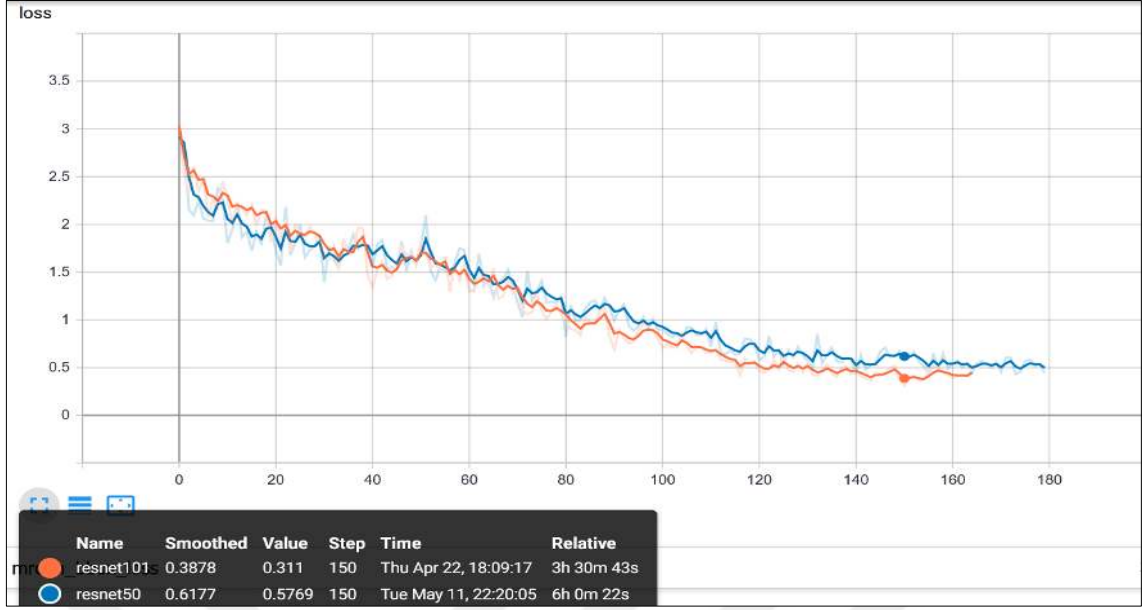
Şekil 56’da kayıp (loss) fonksiyon grafiği gösterilmektedir. Bu grafikte x eksenini epok sayısını y eksenini ise kayıp değerini göstermektedir. Grafik incelendiğinde kayıp değerinin giderek azaldığı ve sifıra yaklaştığı görülmektedir. Bu durum eğitim ilerledikçe doğruluğun arttığı anlamına gelmektedir.



Şekil 56. ResNet101 için kayıp grafiği

Sentinel-2 uydu görüntüsü üzerinden yapılaşma alanı tespiti için ResNet50 ağ mimarisi ile oluşturulmuş olan modelin eğitimi için; tüm hiperparametreler ve eğitim aşaması aynı seçilmiş sadece ResNet seçimi değiştirilmiştir.

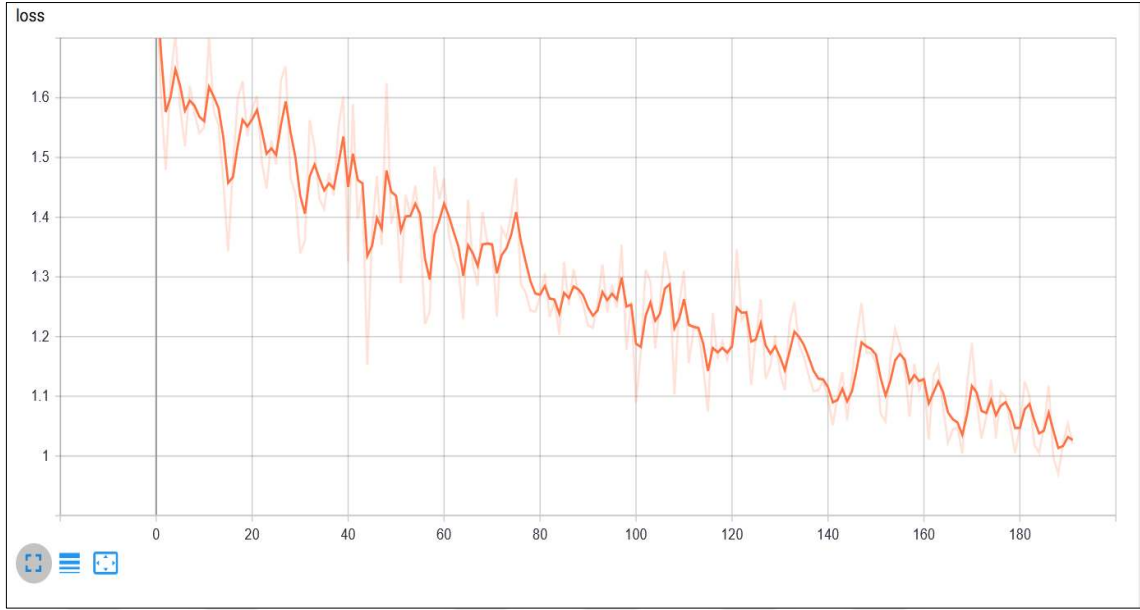
Şekil 57’de ResNet50 ile ResNet101’in kayıp (loss) grafiği gösterilmektedir. Bu grafikte x eksenini epok sayısını y eksenini ise kayıp değerini göstermektedir. Görüldüğü üzere 150. adımda ResNet101 ağ mimarisi ile elde edilen kayıp değeri 0.3878 iken ResNet50 ile 0.6177’dir, bu durum göstermektedir ki ResNet101 bu veri seti ve hiperparametre ayarlarında daha iyi sonuçlara ulaşmıştır.



Şekil 57. ResNet50, ResNet101 kayıp grafikleri

Sentinel-2 uydu görüntüsü üzerinden arazi örtüsünü sınıflandırmak için oluşturulan modelin eğitimi için birçok hiperparametre değeri denenmiştir. En iyi sonuçlar tekrardan tek sınıf için oluşturulmuş olan hiperparametre değerleriyle aynıdır.

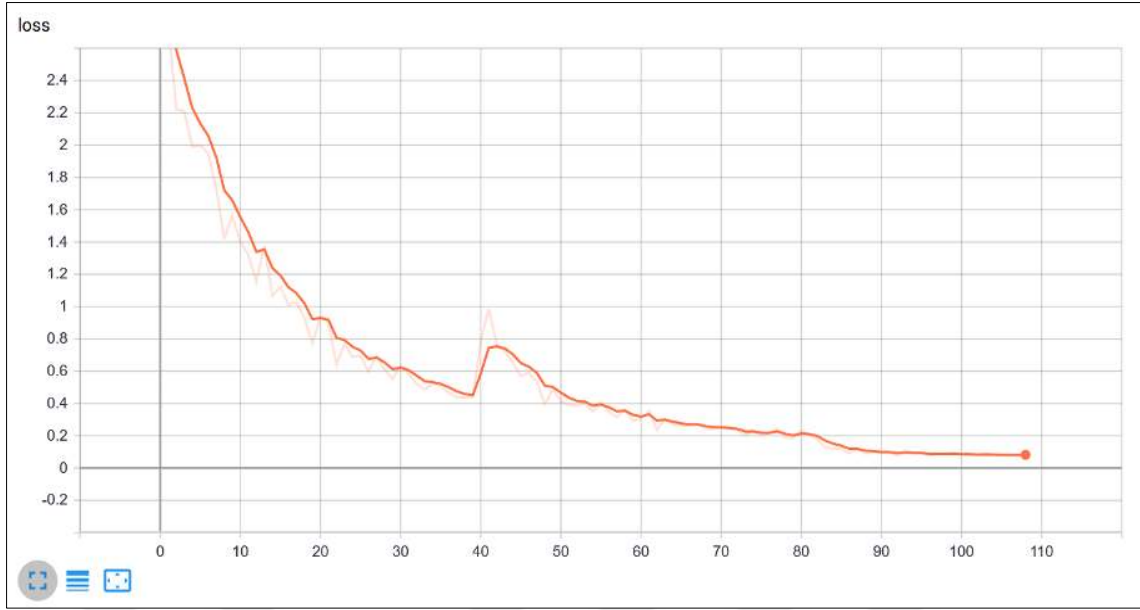
Bölge tespiti için oluşturulan son modelin eğitimi için ilk aşamada “learning\_rate” değeri 0.001 seçilmiş ve “layers: heads” seçilerek modelin sadece baş kısmına göre 45 epok uygulanmıştır. İkinci aşamada; “learning\_rate” yine 0.001 seçilmiş ve ResNet ağında 4’ncü ve diğer katmanlar işleme katılarak 55 epok uygulanmıştır. Son aşamada “learning\_rate” 0.0001 seçilerek tüm katmanlarda 190 epok eğitim yapılmıştır. Toplamda 400 epok yapılması planlanırken eğitim işlemi anlık olarak TensorBoard’da izlenmiş ve değerlerde iyileşme olmadığı için 290. epokta durdurulmuştur. Bu süreçte oluşan kayıp grafiğinin bir kısmı Şekil 58’de gösterilmektedir. Grafikteki x eksenini epok sayısını y eksenini ise kayıp değerini göstermektedir. Gösterilen bölüm eğitim aşamasının son aşamasındaki 190 epokluk kısımdır. Colab uzun süreler çalışmaya izin vermediğinden ilk iki aşama sonucunda oluşan eğitim verileri kullanılarak 3 aşama yeniden başlatılmıştır. Bu grafik incelendiğinde en iyi sonuçların 124. epokta olduğu görülmüş ve sınıflandırma aşamasında bu epoktaki eğitim verileri kullanılmıştır.



Şekil 58. Arazi örtüsü için kayıp grafiği

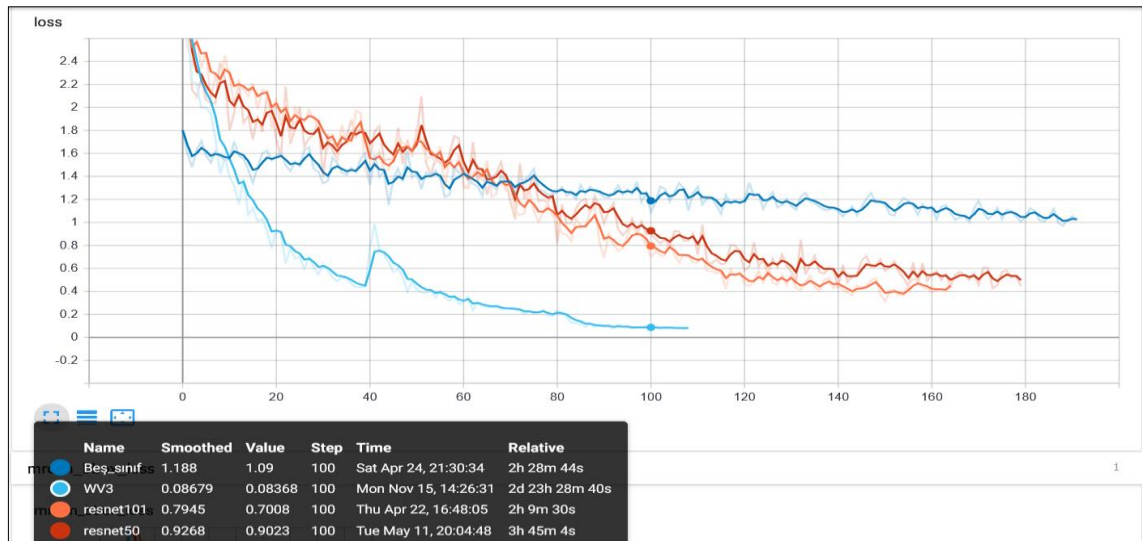
WV3 uydu görüntüleri Sentinel-2 uydu görüntülerinden daha yüksek uzamsal çözünürlüğe sahip olduğundan Mask R-CNN uygulaması için daha iyi sonuçlar elde edilmesi beklenmektedir. Bu çalışma kapsamında binaların tespit edilmesi amaçlanmış veri seti oluşturulurken ‘bina’ etiketleri oluşturulmuştur. Veri seti oluşturulurken kullanılan görüntü sayısı toplamda 80 olup bu sayı Sentinel-2 uydu görüntülerinden oluşturulan veri setinin neredeyse %8’dir. Bunun sebebi, kullanılan görüntüler daha yüksek çözünürlüğe sahip olduğundan daha az veri seti ile iyi sonuçlar verebilmesidir.

Bina tespiti için oluşturulan modelde eğitim süreci üç aşamalı olarak gerçekleşmiştir. İlk aşamada “learning\_rate” değeri 0.001 seçilmiş ve “layers: heads” seçilerek modelin sadece baş kısmına göre 40 epok uygulanmıştır. İkinci aşamada; “learning\_rate” yine 0.001 seçilmiş ve ResNet ağında 4’ncü ve diğer katmanlar işleme katılarak 40 epok uygulanmıştır. Son aşamada learning\_rate 0.0001 seçilerek tüm katmanlarda 29 epok eğitim yapılmıştır. Toplamda 200 epok yapılması planlanırken eğitim işlemi anlık olarak TensorBoard’da izlenmiş ve değerlerde iyileşme olmadığı için 109. epokta durdurulmuştur. Şekil 59’da TensorBoard grafiği gösterilmektedir. Grafikteki x eksenini epok sayısını y eksenini ise kayıp değerini göstermektedir. Grafik incelendiğinde en iyi sonucun 105. epokta olduğuna karar verilmiş ve sonuç sınıflandırma aşamasında kullanılmıştır.



Şekil 59. WV3 uydu görüntüleri için kayıp grafiği

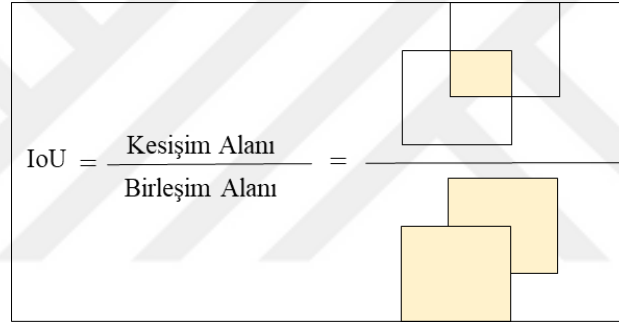
Şekil 60’da tüm çalışma sonucunda oluşan kayıp grafiği gösterilmektedir. Grafikteki x eksenini epok sayısını y eksenini ise kayıp değerini göstermektedir. Grafikteki koyu mavi çizgi beş sınıf ile yapılan modeli, mavi çizgi WV3 uydu görüntüsü için yapılan modeli, turuncu çizgi ResNet101 ağ mimari ile oluşturulan modeli, kırmızı çizgi ResNet50 ağ mimarisi ile oluşturulan modeli temsil etmektedir. Görüldüğü üzere en iyi değerler WV3 uydu görüntüsü ve Sentinel-2 uydu görüntüsünde ResNet101 kullanılarak eğitilen veri setinden elde edilmiştir.



Şekil 60. Tüm çalışma için kayıp grafiği

Nesne algılama modellerinin değerlendirilmesinde, eğitim sonucunda görüntüdeki nesne tespit performansının ölçülmesi gerekmektedir. Bu modelde değerlendirme yapabilmek için birleşim üzerinden kesişme (Intersection over union (IoU)), ortalama kesinlik puanı (Averaged Precision (AP)), ortalama geri çağırma puanı (Averaged Recall (AR)), ortalama ortalama hassasiyet (mean average precision (mAP)) gibi ölçütler kullanılmıştır.

Birleşim üzerinden kesişme (IoU), nesnelerin konumlarını tahmin etmek için çıktı olarak sınırlayıcı kutuları kullanmaktadır. IoU hesaplanırken yer gerçeği sınırlayıcı kutu ve model tarafından tahmin edilen kutu kullanılmaktadır. IoU'yu hesaplamak için, iki sınırlayıcı kutu arasındaki kesişme alanı birleşim alanına bölünmektedir (Şekil 61). IoU değeri 0-1 arasında bir değer alabilir.



Şekil 61. Birleşim üzerinden kesişme

Bir modelin performansını belirlemek için doğru pozitif (True positives (TP)), doğru negatif (True negatives (TN)), yanlış pozitif (False positives (FP)) ve yanlış negatif (False negatives (FN)) metrikleri kullanılarak kesinlik ve geri çağırma değerleri hesaplanmaktadır.

- Doğru pozitif (TP): Nesneyi oluşturan piksellerin doğru şekilde sınıflandırılması.
- Doğru negatif (TN): Arka plan piksellerinin doğru şekilde sınıflandırılması.
- Yanlış Pozitif (FP): Arka plan piksellerinin nesne pikseli olarak sınıflandırılması.
- Yanlış Negatif (FN): Nesne piksellerinin arka plan pikseli olarak sınıflandırılması.

Bu durum Tablo 4'deki gibi karışıklık matrisinde gösterilmektedir. Bir sınıflandırma algoritmasını değerlendirmek için kullanılan standart metrikler bu matrsten elde edilmektedir.

Tablo 4. Karışıklık matrisi

|                       | <b>Tahmin Pozitif</b> | <b>Tahmin Negatif</b> |
|-----------------------|-----------------------|-----------------------|
| <b>Gerçek Pozitif</b> | Doğru pozitif (TP)    | Yanlış Negatif (FN)   |
| <b>Gerçek Negatif</b> | Yanlış pozitif (FP)   | Doğru negatif (TN)    |

Kesinlik; doğru tahmin edilen pozitif gözlemlerin toplam tahmin edilen pozitif gözlemlere oranıdır. Modelin ne kadar kesin olduğunu göstermek için kullanılır.

$$\text{Kesinlik (Precision): } \frac{TP}{TP+FP} \quad (2)$$

Duyarlılık; pozitif olarak tahmin edilmesi gereken gözlemlerin ne kadarının pozitif tespit edildiğini gösteren orandır.

$$\text{Duyarlılık (Recall): } \frac{TP}{TP+FN} \quad (3)$$

AP; temel olarak duyarlılık-kesinlik eğrisine dayanır. AP'yi hesaplamadan önce çoklu geri çağırma seviyelerinde enterpolasyon yapılır. Enterpolasyonlu hassasiyet ( $P_{interp}$ ), çoklu geri çağırma değerlerinde ( $r_1, r_2, \dots$ ) hesaplanan en yüksek kesinlik değeri olarak tanımlanır. AP, enterpolasyonlu duyarlılık-kesinlik eğrisinin altındaki alan olarak tanımlanabilir (Zeng, 2021).

$$AP = \sum_n \{(r_{n+1} - r_n) P_{interp}(r_{n+1})\} \quad (4)$$

mAP; ortalama kesinliğin tüm sınıflardaki ( $K$ ) ortalamasıdır.

$$mAP = \frac{\sum_{i=1}^K AP_i}{K} \quad (5)$$

AR; farklı IoU değerlerindeki (0.5, 0.55, ...0.95) duyarlılığın ortalamasıdır. Duyarlılık- IoU eğrisinin altında kalan alanın 2 kere hesaplanmasıyla bulunur.

$$AR = 2 \int_{0.5}^1 recall(o) do \quad (6)$$

mAR; tüm sınıflardaki ortalama duyarlılığın ortalamasının hesaplanmasıdır.

$$mAR = \frac{\sum_{i=1}^K AR_i}{K} \quad (7)$$

F1 puanı; kesinlik ve duyarlılık değerlerinin harmonik ortalamasıdır.

$$F1 = 2 \times \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \quad (8)$$

Tez çalışmasında yukarıdaki değerlendirme metrikleri kullanılmıştır. Bunların sonucunda hesaplanan mAR, mAP ve F1 puanı değerleri Tablo 5’te verilmiştir.

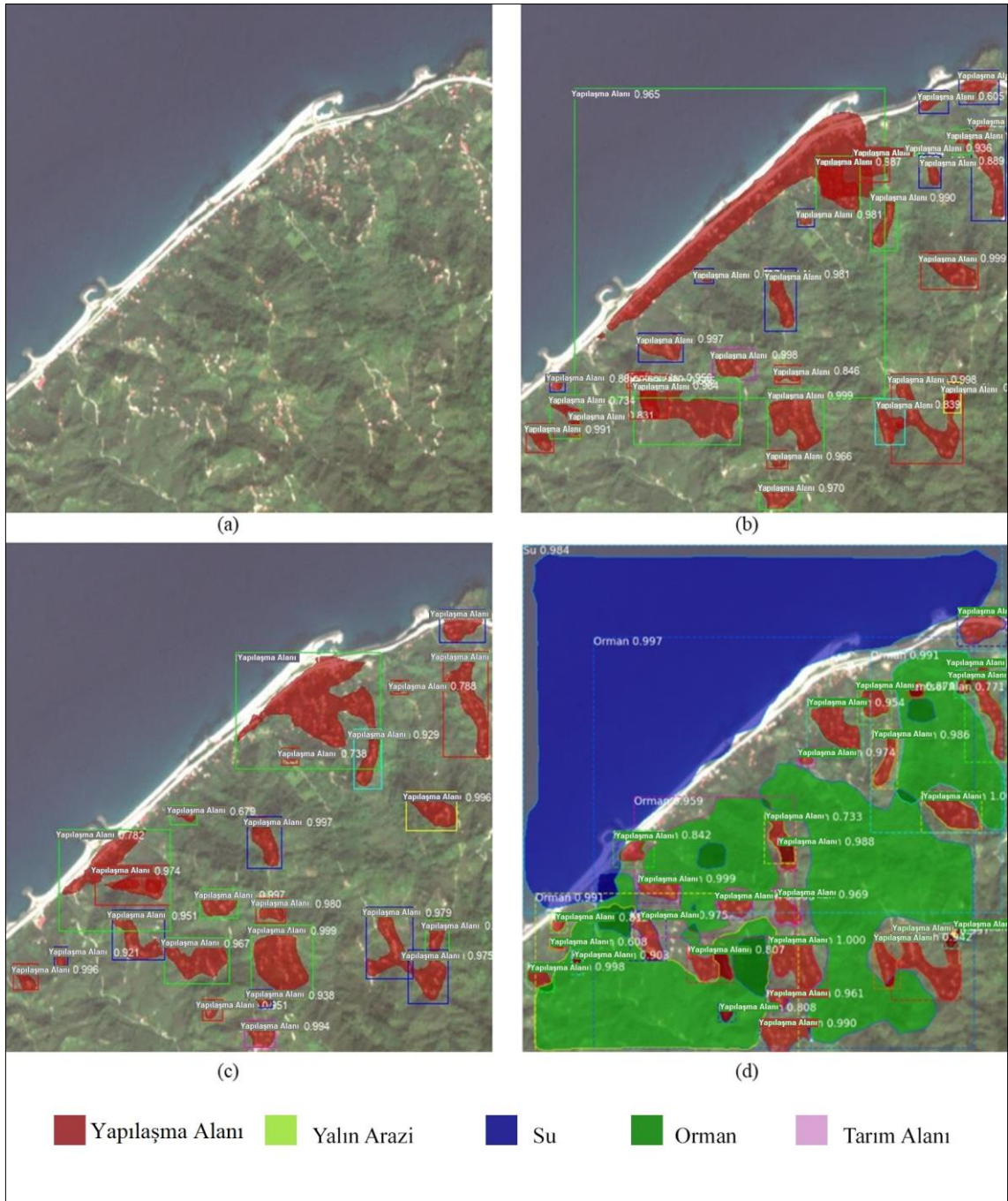
Tablo 5. Değerlendirme sonuçları

|                   |                      | <b>mAR</b> | <b>mAP</b> | <b>F1 Puanı</b> |
|-------------------|----------------------|------------|------------|-----------------|
| <b>Sentinel-2</b> | ResNet101            | 0.8914     | 0.7453     | 0.8118          |
|                   | ResNet50             | 0.8875     | 0.7144     | 0.7916          |
|                   | Arazi örtüsü tespiti | 0.8477     | 0.6649     | 0.7452          |
| <b>WV3</b>        | Bina tespiti         | 0.8669     | 0.8069     | 0.8358          |

Tablo 5 incelendiğinde; Sentinel-2 uydu görüntüleri üzerinden ResNet101 ağ mimarisi kullanılarak oluşturulan modelde mAR: 0,89, mAP: 0,74 ve F1 Puanı: 0,81 olarak bulunmuştur. ResNet50 ağ mimarisi kullanılarak oluşturulan modelde mAR: 0,88, mAP: 0,71 ve F1 Puanı: 0,89 olarak bulunmuştur. Bu sonuçlardan ResNet101’in ResNet50’ye göre daha iyi olduğu gözükmemektedir. Arazi örtüsü tespiti için Sentinel-2 uydu görüntüleri ve ResNet101 ağ mimari kullanılarak oluşturulan modelde mAR: 0,84, mAP: 0,66 ve F1 Puanı:

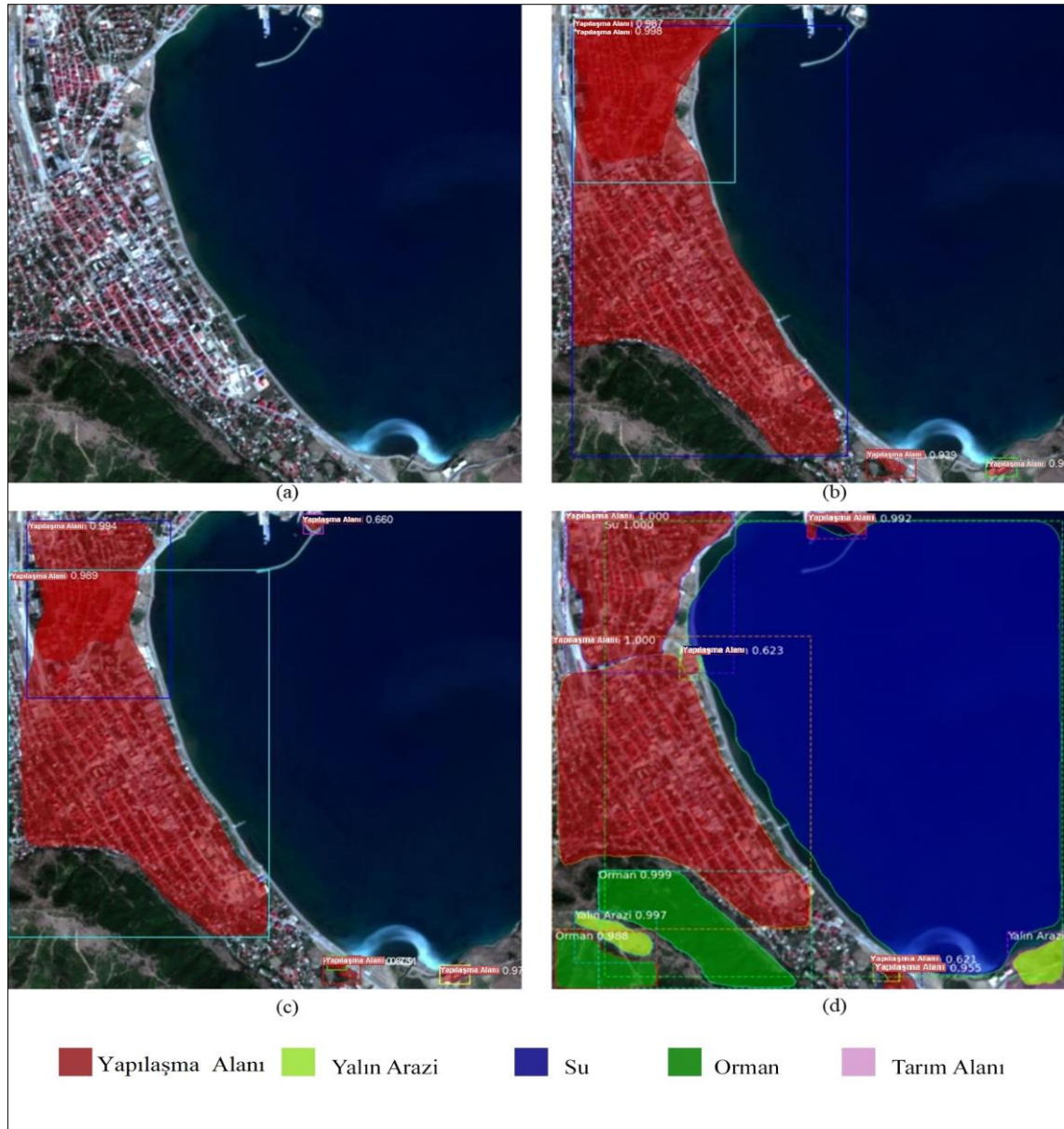
0,74 olarak bulunmuştur. WV3 uydu görüntüleri kullanılarak bina tespiti yapılması amacıyla oluşturulan modelde mAR: 0,86, mAP: 0,80 ve F1 Puanı: 0,83 sonuçları elde edilmiştir. Tüm modeller incelendiğinde en iyi sonuçların WV3 uydu görüntüsü ile oluşturulmuş olan modelden olduğu gözükmemektedir. Diğer sonuçlarda sınıflandırma için yeterli değerlere sahiptir. Aşağıdaki görseller ile sonuçlar incelendiğinde sonuçların tutarlı olduğu gözükmemektedir.

Şekil 62’de oluşturulan modellerden elde edilen sınıflandırma sonuçları gösterilmektedir. ResNet101 ağ mimarisi kullanılarak yapılan sınıflandırma işlemi (b) görüntüsünde, ResNet50 kullanılarak yapılan sınıflandırma işlemi (c) görüntüsünde gösterilmektedir. Bu iki görüntü kıyaslandığında (b) görüntüsünün (c) görüntüsüne göre daha iyi sonuçlar elde ettiği anlaşılmaktadır. ResNet101 kıyı bölgelerde kentsel arazi dokusunu daha iyi tespit etmiştir. Beş sınıf kullanılarak yapılan arazi örtüsü sınıflandırması (d) görüntüsünde gösterilmektedir. Görüntü incelendiğinde yanlış bir sınıf ataması olmayıp bazı piksellerin sınıflandırılmadığı gözükmemektedir.



Şekil 62. Sınıflandırma sonuçları (a) orijinal görüntü, (b) Resnet101, (c) ResNet50, (d) beş sınıf

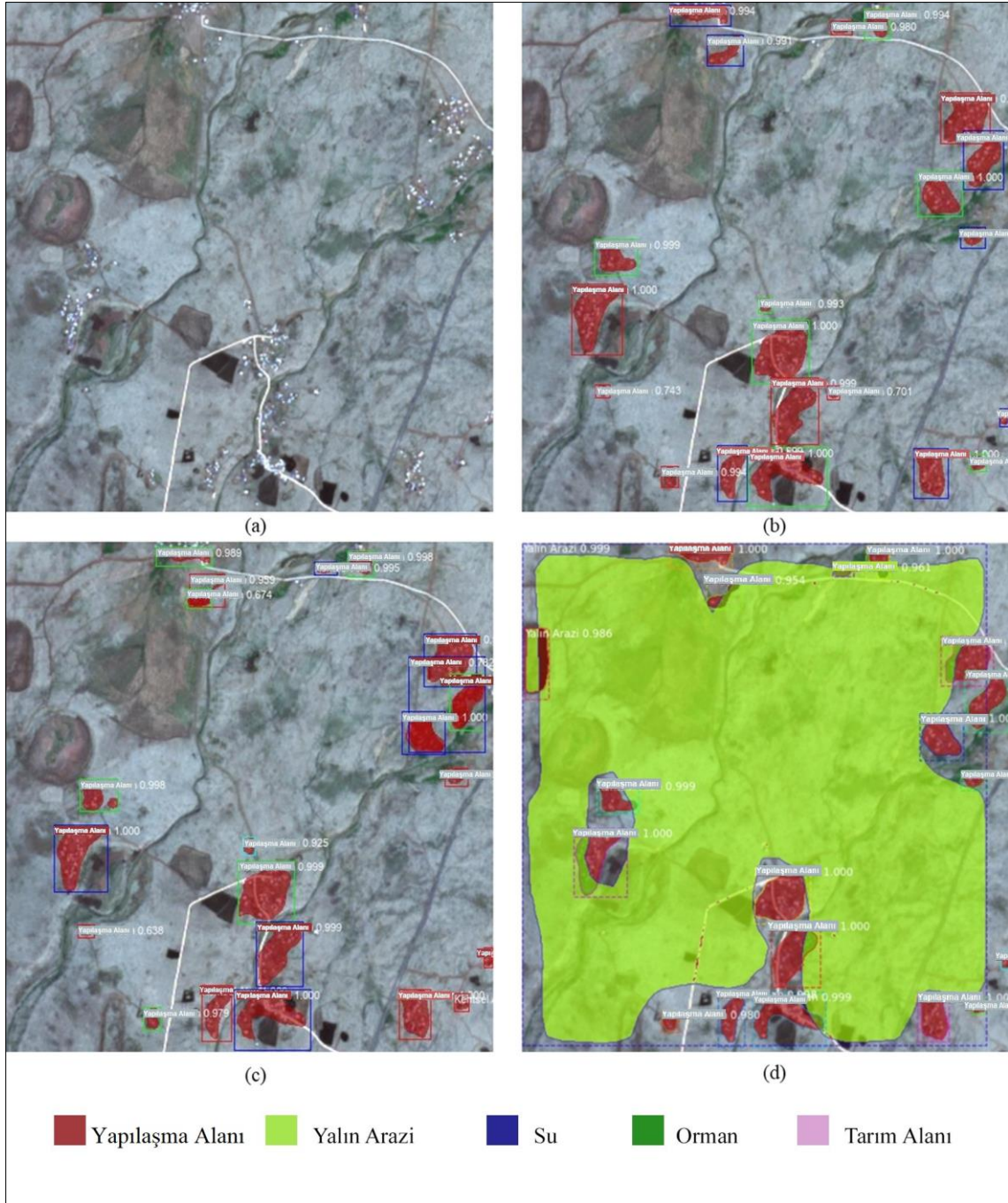
Şekil 63'te oluşturulan modellerden elde edilen sınıflandırma sonuçları gösterilmektedir. ResNet101 ağ mimarisi kullanılarak yapılan sınıflandırma işlemi (b) görüntüsünde, ResNet50 kullanılarak yapılan sınıflandırma işlemi (c) görüntüsünde gösterilmektedir. Yapılan sınıflandırma işleminde ResNet101 yapılaşma alanını ResNet50'ye göre daha iyi tespit ettiği gözükmemektedir. Beş sınıf kullanılarak yapılan arazi örtüsü sınıflandırması (d) görüntüsünde gösterilmektedir. (d) görüntüsü incelendiğinde bazı bölgelerin sınıflandırılmasının yapılmadığı gözükmemektedir.



Şekil 63. Sınıflandırma sonuçları (a) orijinal görüntü, (b) Resnet101, (c) Resnet50, (d) beş sınıf

Şekil 64'te oluşturulan modellerden elde edilen sınıflandırma sonuçları gösterilmektedir. ResNet101 ağ mimarisi kullanılarak yapılan sınıflandırma işlemi (b) görüntüsünde, ResNet50 kullanılarak yapılan sınıflandırma işlemi (c) görüntüsünde gösterilmektedir. ResNet101 ağ mimarisi ResNet50'ye göre daha az kutu kullanarak kentsel dokuyu tespit ettiği görülmüştür. Beş sınıf kullanılarak yapılan arazi örtüsü sınıflandırması (d) görüntüsünde gösterilmektedir. (d) görüntüsü incelendiğinde görüntünün kenarlarının iyi tespit edilmediği bazı bölümlerde yalın arazi ile kentsel dokunun iç içe geçtiği gözükmemektedir.

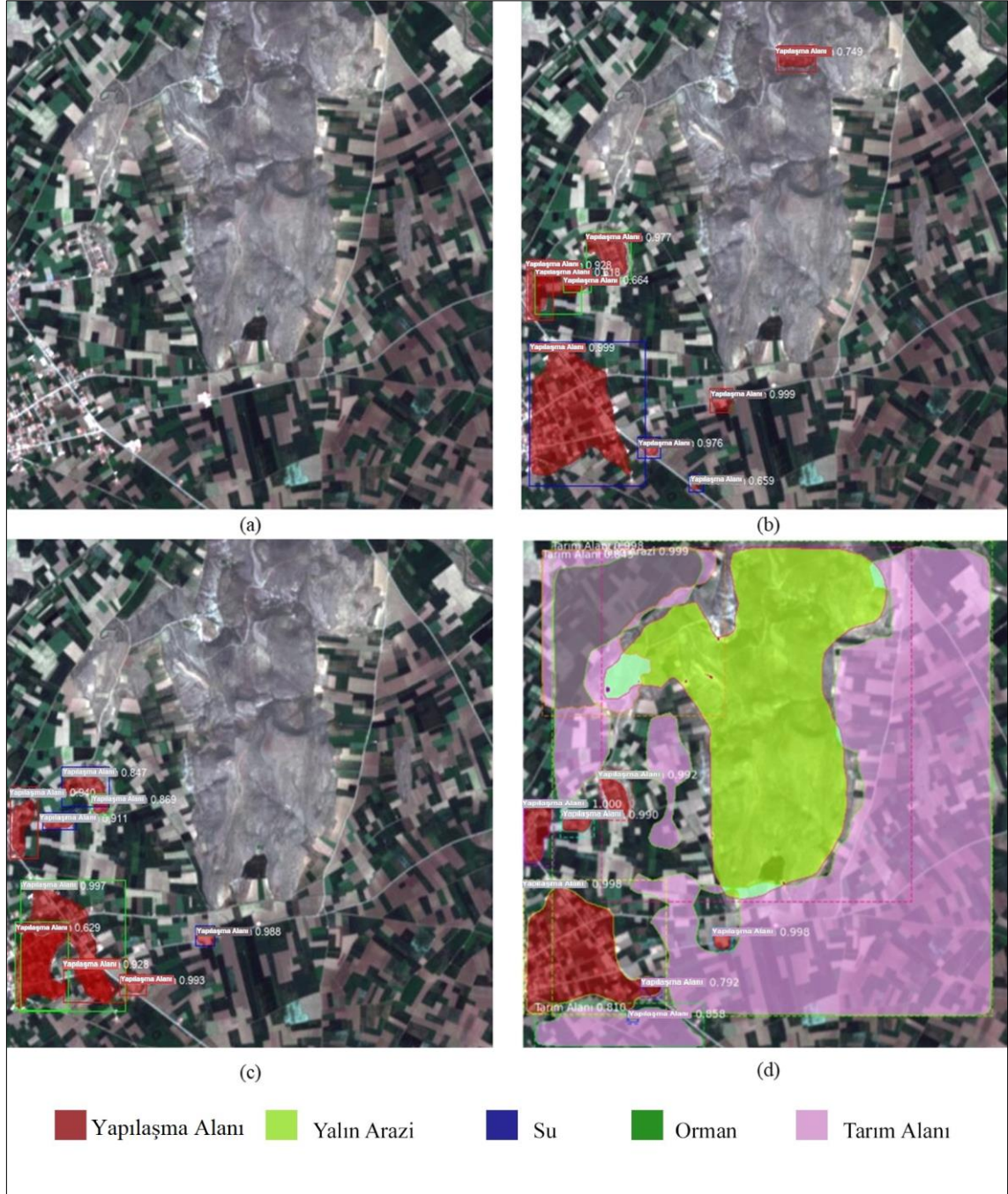




Şekil 64. Sınıflandırma sonuçları (a) orijinal görüntü, (b)Resnet101, (c)Resnet50, (d) beş sınıf

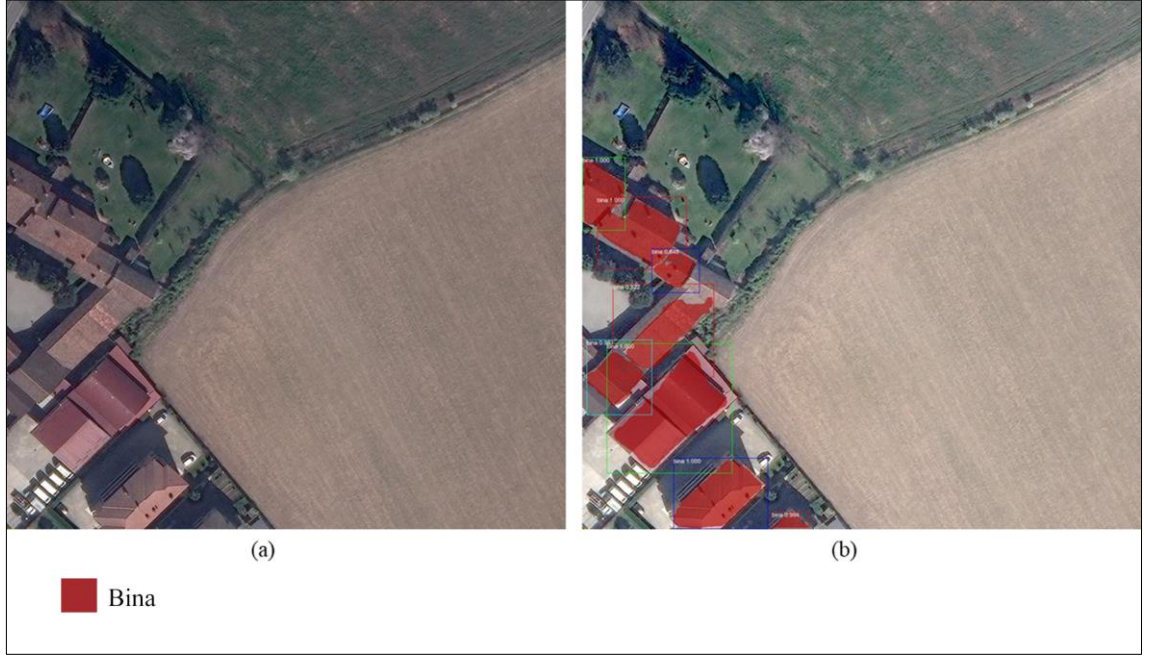
Şekil 65'te oluşturulan modellerden elde edilen sınıflandırma sonuçları gösterilmektedir. ResNet101 ağ mimarisi kullanılarak yapılan sınıflandırma işlemi (b) görüntüsünde, ResNet50 kullanılarak yapılan sınıflandırma işlemi (c) görüntüsünde gösterilmektedir. ResNet101 ile yapılan sınıflandırmada yapılaşma alanı olmamasına rağmen bölgenin yapılaşma alanı olarak sınıflandırıldığı gözükmemektedir. ResNet50 ile yapılan sınıflandırmada aynı nesne birden fazla kutu ile tespit edilmiş ama yapılaşma alanları

doğru bir şekilde tespit edilmiştir. Beş sınıf kullanılarak yapılan arazi örtüsü sınıflandırması (d) görüntüsünde gösterilmektedir. (d) görüntüsünde tarım alanlarının tespit edilemediği gözükmemektedir.



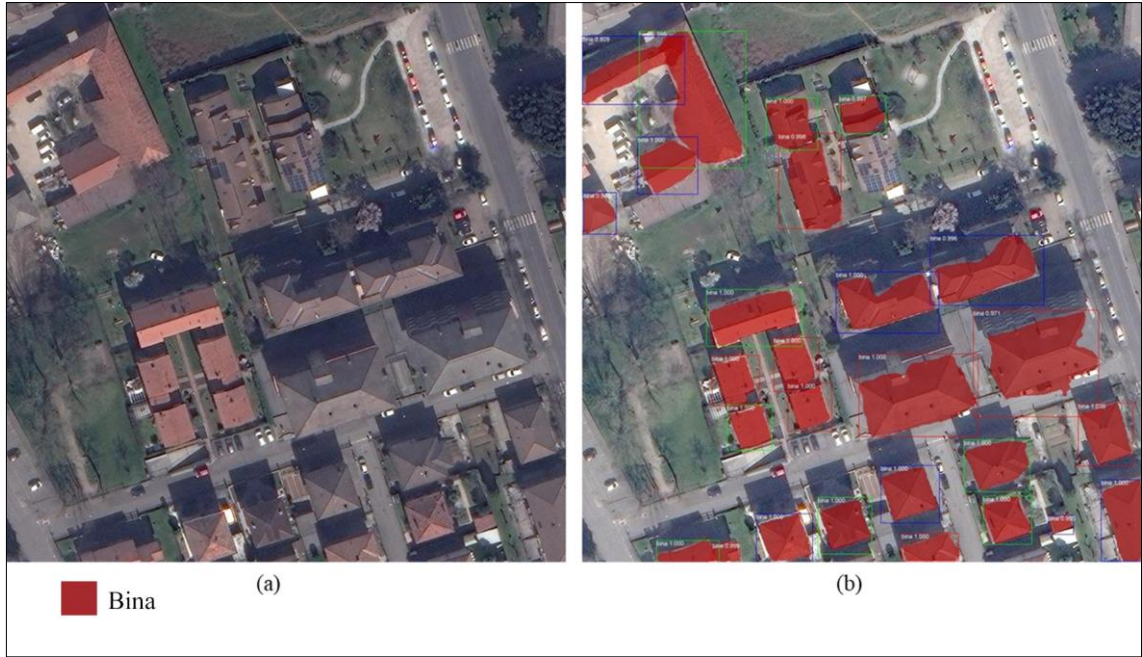
Şekil 65. Sınıflandırma sonuçları (a) orijinal görüntü, (b) ResNet101, (c) ResNet50, (d) beş sınıf

Şekil 66’da WV3 görüntüsü üzerinde yapılan sınıflandırma işlemi ile bina bulma performansı gösterilmektedir. Orijinal görüntü (a) görüntüsünde, sınıflandırma sonuç görüntüsü (b) görüntüsünde gösterilmektedir. (b) görüntüsü incelendiğinde bir bina haricinde tüm binaların tespit edilebildiği gözükmemektedir.



Şekil 66. WV3 görüntüsünün sınıflandırma sonuçları (a) orijinal görüntü, (b) sınıflandırma

Şekil 67’de WV3 uydu görüntüsü üzerinden yapılan bina tespiti gösterilmektedir. Orijinal görüntü (a) görüntüsünde, sınıflandırma sonuç görüntüsü (b) görüntüsünde gösterilmektedir. (b) görüntüsü incelendiğinde birkaç bina dışında tüm görüntüdeki binaları doğru şekilde tespit edilebildiği gözükmemektedir.



Şekil 67. WV3 görüntüsünde sınıflandırma sonuçları (a) orijinal görüntü, (b) sınıflandırma

WV3 görüntüsündeki tüm parçaların sınıflandırması yapıldıktan sonra MATLAB ortamında birleştirilmiştir. Çalışma alanı ve sınıflandırma sonuçları Şekil 68’de gösterilmektedir. Sınıflandırma sonuçları incelendiğinde kahverengi ve gri bina çatılarının tespitinin beyaz çatılı binalara göre daha iyi yaptığı gözükmemektedir. Bunun sebebi etiketli veri setinde beyaz çatılı bina örneklerinin yeterli sayıda olmamasıdır. Bazı tarım bölgelerine bina etiketi atanmıştır. Bunun sebebi tarım bölgelerinin renklerinin bina çatı renklerine benzemesidir. Bu tür problemlerin giderilmesi ve daha iyi sonuçlar elde edilmesi için daha veri seti kullanılarak eğitim işlemi yapılabilir. Bu sayede modelin renk ve doku bilgisi gelişerek daha iyi sonuçlar verebilir. Bunların dışında WV3 görüntüleri üzerinde yapılan sınıflandırma işleminin Sentinel-2 görüntülerine göre daha iyi sonuçlar verdiği gözükmemektedir.



Şekil 68. WV3 uydu görüntüsünün sınıflandırılması (a) çalışma alanı, (b) sınıflandırma sonuçları

#### 4. SONUÇLAR VE ÖNERİLER

Bu tez çalışmasında, yenilikçi bir yaklaşım olan örnek segmentasyonun (Mask R-CNN) yüksek/orta çözünürlüklü ve çok bantlı uydu görüntüleri üzerinde sonuçları sunulmaktadır. Örnek segmentasyon, semantik segmentasyondan daha fazla bilgi edinerek nesnelerin daha iyi anlaşılmasını sağlamaktadır. Mask R-CNN, maske tahminleri, nesneleri çevreleyen kutu sınırlarını, nesne koordinatlarını belirleme, farklı nesneleri daha iyi ayırt etme gibi zorlu uzaktan algılama görevleri için iyi sonuçlar vermektedir. Şu anda çok bantlı uydu görüntülerinde Mask R-CNN algoritmasını kullanan çalışmalar oldukça azdır.

Tez çalışması yapılırken Sentinel-2 ve WV3 uydu görüntülerinde örnek segmentasyon algoritması olan Mask R-CNN kullanarak sınıflandırma yapılmıştır. Veri seti oluşturulurken VGG Image Annotator kullanılmıştır. VGG Image Annotator ile veri seti hazırlanması aşamasının ileriki çalışmalarda doğrudan uygulanabilmesini mümkün kılmak için detaylı bir şekilde açıklanmıştır. Veri seti toplamda 1515 görüntü ve etiketli eğitim sınıflarını içeren 18811 adet poligondan oluşmaktadır. Bu işlem oldukça zaman alıcı bir işlem olup toplamda 196 saat sürmüştür. Ağ mimarisi oluşturulurken CNN'in iç katmanlarına ince ayarlar yapılarak hiperparametreler uzaktan algılanmış görüntülerdeki nesneleri ayırt edecek şekilde ayarlanmıştır. Mask R-CNN uygulaması ResNet50 ve ResNet101 ağ mimarilerini desteklemektedir. Bu çalışmada aynı veri seti ve hiperparametrelere sahip veri seti üzerinde sadece ağ mimarileri değiştirilerek eğitim işlemi yapılmıştır. Sınıflandırma sonuçları nispeten aynı olsa da görsel değerlendirme sonuçlarında ResNet101'in daha iyi sonuç aldığı gözükmemektedir. ResNet101 ağ mimarisi hesaplama maliyetleri yüksek olduğundan ileri teknoloji GPU sistemlerine ihtiyaç duyar. Bu sistemler donanımsal açıdan maliyetli olacağından Google Colab kullanılarak yüksek GPU'lu bilgisayarlar üzerinden eğitim işlemi yapılmıştır. Eğitim işleminde rastgele atanan ağırlıklarla başlamak yerine transfer öğrenme yöntemi kullanılmıştır. Bu yöntem ile MS COCO veri setinde eğitilen ağırlıklar başlangıç değeri olarak kullanılmıştır. Bu sayede modelin öğrenme performansı artarak daha iyi sonuçlar elde edilmiştir. Her veri seti için birçok eğitim parametresi denenerek optimum sonuç elde edilene kadar eğitim işlemi yapılmıştır. Sonuç olarak her veri seti için farklı eğitim parametreleri kullanılmıştır. Buradan da anlaşılacağı üzere her veri seti özgün olduğu için her modelde aynı eğitim parametrelerini kullanmak yerine en iyi sonucu veren parametrelerin bulunması gerekmektedir. Modeller oluşturulduktan sonra Sentinel-2 ve

WV3 uydu görüntüleri üzerinde sınıflandırma yapılmıştır. Uydu görüntüleri üzerindeki sınıflandırma işlemi 1-10sn arasında sürmüştür. Bu durum anlık bilgi üretmek için oldukça önemlidir.

Bu çalışmada kullanılan değerlendirme metrikleri doğru pozitif, yanlış pozitif, yanlış negatif ve doğru negatif verilerine dayanmaktadır. Doğru negatif tespit edilecek bir nesnenin olmadığını ve modelin 0.7'dan daha küçük bir güven puanı çıkardığı durumu ifade etmektedir. RPN nesneleri belirleyerek yalnızca mevcut bir nesneye sahip kutuları çıkararak bir sonraki işleme geçirir ve bir güven puanı verir. IoU 0,5 kabul edilerek hesaplanan değerlendirme ölçüsü ortalama kesinlik puanı (AP), ortalama hassasiyet puanı (mAP), ortalama duyarlılık puanı (mAR) ve F1 puanı modelleri değerlendirmek için kullanılmıştır. Bina tespiti için WV3 uydu görüntüsü kullanılarak yapılan Mask R-CNN çalışmasında %84'lük bir başarı puanı ile en iyi sonuçlar elde edilmiştir. Bölge tespiti yapılırken Sentinel-2 uydu görüntüsü üzerinde üç farklı model oluşturulmuştur. ResNet101 ağ mimarisi ile oluşturulan modelden %81 başarı puanı ile en iyi sonuca ulaşılmıştır. ResNet50 ağ mimarisi kullanarak oluşturulan modelden %79'luk bir başarıya ulaşılmıştır. Arazi örtüsünü sınıflandırmak amacıyla ResNet101 ağ mimarisi ile oluşturulan modelden %74'lük bir başarı puanı elde edilmiştir.

Derin öğrenme teknikleri, mevcut veri miktarının fazla olduğu durumlarda daha verimli çalışmaktadır. Bu çalışmada yaklaşık 600 görüntü kullanılırken gelecekteki çalışmalarda daha büyük veri setleri kullanılması amaçlanmaktadır. Bu şekilde sınıf sayısı artırılabilir ve nesnelerin tanımlanması kolaylaşabilir. Sunulan sonuçlar ile Mask R-CNN'nin uydu görüntülerinde sınıflandırma yapabildiği gösterilmiştir. Daha iyi sonuçlar elde etmek için daha yüksek çözünürlüklü verilere ihtiyaç olduğu anlaşılmaktadır. Mask R-CNN'in nesne tespit ederken belli bir şekle sahip olan nesnelerin tespitini daha kolay yaptığı çalışma sonucunda anlaşılmaktadır. Arazi örtüsü tespiti yapılırken yaşanan en büyük sorun her bir nesnenin birbirinden tamamen farklı olması ve belirli bir şekle sahip olmamasıdır. Buna rağmen Mask R-CNN anlaşılabilir ölçüde doğru sonuçlar elde etmiştir.

Uydu görüntülerinin sınıflandırılması sayesinde birçok alanda gelişim sağlanmaktadır. Örneğin bir bölgedeki kaçak yapılaşma tespit edilmek istendiğinde, iki farklı zamandaki uydu görüntüleri üzerinde bina tespiti yapılarak kaçak yapılaşmalar tespit edilebilmektedir. Son dönemlerde Türkiye'deki orman yaygınları artış göstermiştir. Uydu görüntülerindeki anlık sınıflandırmalar sayesinde orman yangınları ve orman yangınları sonucundaki hasar tespitleri yapılabilmektedir. Uydu görüntülerinin bir başka kullanım alanı ise coğrafi bilgi

sistemleri için veri katmanları oluřturmaektir. Coğrafi bilgi sistemleri pek çok problemin çözümlünde rol oynadığından uydu görüntülerin doğru ve hızlı bir şekilde sınıflandırılabilmesi bu alanlara katkı sağlayacaktır. Bu tez çalışmasında Mask R-CNN yönteminin uydu görüntüleri üzerindeki kullanımı anlatılmıştır. Mask R-CNN modeli bir kere oluřturulduktan sonra çok büyük alanların hızlı ve doğru bir şekilde sınıflandırılmasına olanak sağlamaktadır. Mask R-CNN değıřim analizi, kaçak yapılaşma, arazi örtüsünün tespiti, orman alanlarının tespiti gibi problemleri çözmekte kullanılabileceğinden önem kazanmaktadır.



## 5. KAYNAKLAR

- Abiodun, O., I., Jantan, A., Omolara, A. E., Dada, K., V., Mohamed, N., A. ve Arshad, H., 2018. State-of-the-art in Artificial Neural Network Applications: A survey, Heliyon, 4.
- Ackley, D., H., Hinton, G., E. ve Sejnowski, T., J., 1985. A Learning Algorithm for Boltzmann Machines, Cognitive Science, 9, 1, 147-169.
- Albawi, S., Mohammed, T., A. ve Al-Zawi, S., 2017. Understanding of a Convolutional Neural Network, In 2017 International Conference on Engineering and Technology (ICET), IEEE, 1-6.
- Ammour, N., Alhichri, H., Bazi, Y., Benjdira, B., Alajlan, N. ve Zuair, M., 2017. Deep Learning Approach for Car Detection in UAV Imagery, Remote Sensing, 9, 312.
- Alom, M., Z., Taha, T., M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M., S. ve Asari, V., K., 2019. A State-Of-The-Art Survey on Deep Learning Theory and Architectures, Electronics, 8,3, 292.
- Aziz, L., Salam, M. S. B. H., Sheikh, U. U. ve Ayub, S., 2020. Exploring Deep Learning-Based Architecture, Strategies, Applications and Current Trends in Generic Object Detection: A Comprehensive Review, IEEE Access, 8, 170461-170495.
- Bengio, Y., Simard, P. ve Frasconi, P., 1994. Learning Long-Term Dependencies with Gradient Descent Is Difficult, IEEE Transactions on Neural Networks, 5, 2, 157-166.
- Bengio, Y., Ducharme, R. ve Vincent, P., 2001. A Neural Probabilistic Language Model, In Advances in Neural Information Processing Systems, MIT Press, Canada, 932-938.
- Bhardwaj, A., Sam, L., Martín-Torres, F. J. ve Kumar, R., 2016. UAVs as Remote Sensing Platform in Glaciology: Present Applications and Future Prospects, Remote Sensing of Environment, 175, 196-204.
- Bhuiyan, M. A. E., Witharana, C. ve Liljedahl, A. K., 2020. Use of Very High Spatial Resolution Commercial Satellite Imagery and Deep Learning to Automatically Map Ice-Wedge Polygons Across Tundra Vegetation Types, Journal of Imaging, 6, 137.

- Borra S., Thanki R. ve Dey, N., 2019. Satellite Image Analysis: Clustering and Classification, Springer, 1st Edition, Singapore.
- Bouvier, J., 2006. Notes on Convolutional Neural Networks, Neural Nets, MIT CBCL Tech Report, 38–44.
- Boyacioglu, M., A., Kara, Y., ve Baykan, Ö., K., 2009. Predicting Bank Financial Failures Using Neural Networks, Support Vector Machines and Multivariate Statistical Methods: A Comparative Analysis in The Sample of Savings Deposit Insurance Fund (SDIF) Transferred Banks in Turkey, Expert Systems with Applications, 36, 3355-3366.
- Buduma, N., 2016. Fundamentals of Deep Learning, Designing Next-Generation Machine Intelligence Algorithms, 1st Edition, O'Reilly.
- Carvalho, O. L. F. D., De Carvalho Junior, O. A., Albuquerque, A. O. D., Bem, P. P. D., Silva, C. R., Ferreira, P. H. G. ve Borges, D. L., 2020. Instance Segmentation for Large, Multi-Channel Remote Sensing Imagery Using Mask-RCNN and A Mosaicking Approach, Remote Sensing, 13, 39.
- Castelluccio, M., Poggi, G., Sansone, C. ve Verdoliva, L., 2015. Land Use Classification in Remote Sensing Images by Convolutional Neural Networks. <https://arxiv.org/abs/1508.00092>, 21 Haziran 2017.
- Ciresan, D., C., Meier, U., Gambardella, L., M. ve Schmidhuber, J., 2010. Deep Big Simple Neural Nets for Handwritten Digit Recognition, Neural Computation, 22, 12, 3207–3220.
- Cireşan, D. C., Giusti, A., Gambardella, L., M. ve Schmidhuber, J., 2013. Mitosis Detection in Breast Cancer Histology Images with Deep Neural Networks, In International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 411-418, Berlin, Heidelberg.
- Crisóstomo De Castro Filho, H., Abílio de Carvalho Júnior, O., Ferreira de Carvalho, O. L., Pozzobon de Bem, P., Dos Santos De Moura, R., Olino De Albuquerque, A. ve Trancoso Gomes, R. A., 2020. Rice Crop Detection Using LSTM, Bi-LSTM, and Machine Learning Models from Sentinel-1 Time Series, Remote Sensing, 12, 2655.
- Csáji, B., C., 2001. Approximation with artificial neural networks, Faculty of Sciences, Eötvös Loránd University, Hungary, 24,48, 7.

- Dai, J., Li, Y., He, K. ve Sun, J., 2016. R-Fcn: Object Detection Via Region-Based Fully Convolutional Networks. Advances In Neural Information Processing Systems, 29.
- Das, K., Behera, R., N., 2017. A Survey on Machine Learning: Concept, Algorithms and Applications, International Journal of Innovative Research in Computer and Communication Engineering, 5.
- Dave, V., S. ve Dutta, K., 2014. Neural Network Based Models for Software Effort Estimation: A Review, Artificial Intelligence Review, 42, 295-307.
- De Bem, P. P., de Carvalho Júnior, O. A., de Carvalho, O. L. F., Gomes, R. A. T. ve Fontes Guimarães, R., 2020. Performance Analysis of Deep Convolutional Autoencoders with Different Patch Sizes for Change Detection from Burnt Areas, Remote Sensing, 12, 2576.
- Deng, J., Dong, W., Socher, R., Li, L. J., Li, K. ve Fei-Fei, L., 2009. Imagenet: A large-scale hierarchical image database, In 2009 IEEE conference on computer vision and pattern recognition, IEEE, 248-255.
- Diakogiannis, F. I., Waldner, F., Caccetta, P. ve Wu, C., 2020. ResUNet-a: A Deep Learning Framework for Semantic Segmentation of Remotely Sensed Data, ISPRS Journal of Photogrammetry and Remote Sensing, 162, 94-114.
- Elachi, C. ve Van, Z. J. J., 2006. Introduction to The Physics and Techniques of Remote Sensing, Second Edition, John Wiley & Sons, Inc, Hoboken, New Jersey.
- Erhan, D., Szegedy, C., Toshev, A. ve Anguelov, D., 2014. Scalable Object Detection Using Deep Neural Networks, In Proceedings of The IEEE Conference on Computer Vision and Pattern Recognition, 2147-2154.
- Everingham, M. ve Winn, J., 2011. The Pascal Visual Object Classes Challenge 2012 (Voc2012) Development Kit, Pattern Analysis, Statistical Modelling and Computational Learning, Tech. Rep, 8, 5.
- Everingham, M., Eslami, S. M., Van Gool, L., Williams, C. K., Winn, J. ve Zisserman, A., 2015. The Pascal Visual Object Classes Challenge: A Retrospective, International Journal of Computer Vision, 111, 98-136.
- Fan, W., Bouguila, N. ve Ziou, D., 2012. Variational Learning for Finite Dirichlet Mixture Models and Applications, IEEE Transactions on Neural Networks and Learning Systems, 23, 5, 762-774.

- Fischer, A. ve Igel, C., 2014. Training Restricted Boltzmann machines: An Introduction, Pattern Recognition, 47,1, 25-39.
- Fritz, K. 2020. Instance Segmentation of Buildings in Satellite Images, Master of Science Thesis in Electrical Engineering, Linköping University, Linköping, Sweden.
- Fu, C. Y., Liu, W., Ranga, A., Tyagi, A. ve Berg, A. C., 2017. Dssd: Deconvolutional Single Shot Detector, arXiv preprint arXiv:1701.06659.
- Fukushima, K., 1979. Neural Network Model for A Mechanism of Pattern Recognition Unaffected by Shift in Position-Neocognitron, IEICE Technical Report, A, 62, 10, 658-665.
- Fung, G. P. C., Yu, J. X., Lu, H. ve Yu, P. S., 2005. Text Classification without Negative Examples Revisit, IEEE Transactions on Knowledge and Data Engineering, 18, 6-20.
- Gad, A., F., Gad, A., F. ve John, S., 2018. Practical Computer Vision Applications Using Deep Learning with Cnns, Berkeley: Apress, Menoufia, Egypt.
- Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V. ve Garcia-Rodriguez, J., 2017. A Review on Deep Learning Techniques Applied to Semantic Segmentation. arXiv preprint arXiv:1704.06857.
- Girshick, R., Donahue, J., Darrell, T., ve Malik, J. 2014. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation, In Proceedings of The IEEE Conference on Computer Vision and Pattern Recognition, 580-587.
- Girshick, R., 2015. Fast R-Cnn, In Proceedings of The IEEE International Conference on Computer Vision, 1440-1448.
- Goodfellow, I., Bengio, Y. ve Courville, A., 2016. Deep learning, MIT Press, Cambridge, Massachusetts, London, England.
- Grauman, K. ve Darrell, T., 2005. The Pyramid Match Kernel: Discriminative Classification with Sets of Image Features. In Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, IEEE, 1458-1465.
- Gulli, A. ve Pal, S., 2017. Deep Learning with Keras. Packt Publishing Ltd., Birmingham, UK.
- Haykin, S., 2004. Kalman Filtering and Neural Networks, John Wiley & Sons Inc., 605 Third Avenue, New York, America.

- He, K., Zhang, X., Ren, S., ve Sun, J., 2015. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, 37, 1904-1916.
- He, K., Zhang, X., Ren, S. ve Sun, J., 2016. Deep Residual Learning for Image Recognition, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June, Las Vegas, Proceeding Books, 770-778.
- He, K., Gkioxari, G., Dollr, R. ve Girshick, R., 2017. Mask R-Cnn, In Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2961-2969.
- Hinton, G., E., Osindero, S. ve Teh, Y., W., 2006. A Fast Learning Algorithm for Deep Belief Nets, Neural Computation, 18,7, 1527-1554.
- Hinton, G., E., 2007. Boltzmann Machine, Scholarpedia, 2, 5, doi:10.4249/scholarpedia.1668.
- Hochreiter, S., ve Schmidhuber, J., 1997. Long Short-Term Memory, Neural Computation, 9,8, 1735-1780.
- Hubel, D., H. ve Wiesel, T., N., 1968. Receptive Fields and Functional Architecture of Monkey Striate Cortex, The Journal of Physiology, 195,1, 215-243.
- Ivakhnenko, A., G. ve Lapa, V., G., 1965. Cybernetic Predicting Devices, CCM Information Corporation.
- Ivakhnenko, A., G., 1971. Polynomial Theory of Complex Systems, IEEE Transactions on Systems, Man and Cybernetics, 4, 364-378.
- Joseph, R., D., 1961. Contributions to Perceptron Theory, PhD thesis, Cornell University, New York, America.
- Knights, D., Costello, E. K. ve Knight, R., 2011. Supervised Classification of Human Microbiota, FEMS Microbiology Reviews, 35, 343-359.
- Krizhevsky, A., Sutskever, I. ve Hinton, G., E., 2012. Imagenet Classification with Deep Convolutional Neural Networks, Advances in Neural Information Processing Systems, 25, 1097-1105.
- Kussul, N., Lavreniuk, M., Skakun, S. ve Shelestov, A. 2017. Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data, IEEE Geoscience and Remote Sensing Letters, 14, 5, 778-782.

- Kuznetsova, A., Rom, H., Alldrin, N., Uijlings, J., Krasin, I., Pont-Tuset, J. ve Ferrari, V., 2020. The Open Images Dataset V4, International Journal of Computer Vision, 128, 1956-1981.
- Lazebnik, S., Schmid, C., ve Ponce, J. 2006. Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories. In 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), IEEE, 2, 2169-2178.
- Lee, C., Y., Gallagher, P., W. ve Tu, Z. 2016. Generalizing Pooling Functions in Convolutional Neural Networks: Mixed, Gated, and Tree, In Artificial Intelligence and Statistics, PMLR, 464-472.
- LeCun, Y., Boser, B., Denker, J., S., Henderson, D., Howard, R., E., Hubbard, W. ve Jackel, L., D., 1989. Backpropagation Applied to Handwritten Zip Code Recognition, Neural Computation, 1,4, 541-551.
- Lillesand, T., M., Kiefer, R., W. ve Chipman, J., W., 2004. Remote Sensing and Image Interpretation, Fifth Edition, John Wiley & Sons, Inc., America.
- Lin, L., Zhang, D., Luo, P. ve Zuo, W., 2020. The Foundation and Advances of Deep Learning, In Human Centric Visual Analysis with Deep Learning, Springer, Volume 57, Singapore.
- Liu, P., Di, L., Du, Q. ve Wang, L. 2018. Remote Sensing Big Data: Theory, Methods and Applications. Remote Sensing, 10, 711.
- Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., ve Belongie, S., 2017. Feature Pyramid Networks for Object Detection. In Proceedings of The IEEE Conference on Computer Vision and Pattern Recognition, 2117-2125.
- Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D. ve Zitnick, C. L., 2014. Microsoft Coco: Common Objects in Context. In European Conference on Computer Vision, Springer, Cham, 740-755.
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y. ve Berg, A. C., 2016. Ssd: Single Shot Multibox Detector. In European Conference on Computer Vision, Springer, Cham 21-37.
- Lu, J., Behbood, V., Hao, P., Zuo, H., Xue, S. ve Zhang, G., 2015. Transfer Learning Using Computational Intelligence: A Survey. Knowledge-Based Systems, 80, 14-23.

- Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G. ve Johnson, B. A. 2019. Deep Learning in Remote Sensing Applications: A Meta-Analysis and Review. ISPRS Journal of Photogrammetry and Remote Sensing, 152, 166-177.
- Mahoney, M., V., 1999. Text Compression as A Test for Artificial Intelligence, AAAI/IAAI, 970.
- Mather, P. ve Tso, B. 2016. Classification Methods for Remotely Sensed Data, Second Edition, CRC Press, America.
- Min, S., Lee, B. ve Yoon, S. 2017. Deep Learning in Bioinformatics, Briefings in Bioinformatics, 18, 5, 851-869.
- Morchhale, S., 2016. Deep Convolutional Neural Networks for Classification of Fused Hyperspectral and Lidar Data. Wake Forest University, Winston-Salem, North Carolina.
- Mosavi, A., Ardabili, S. ve Varkonyi-Koczy, A., R., 2019. List Of Deep Learning Models, In International Conference on Global Research and Education, 202-214, Springer, Budapest, Hungary.
- Mueller, J., P. ve Massaron, L., 2019. Deep Learning for Dummies, John Wiley & Sons, Inc., Hoboken, New Jersey.
- Najibi, M., Rastegari, M. ve Davis, L. S., 2016. G-Cnn: An Iterative Grid Based Object Detector. In Proceedings of The IEEE Conference on Computer Vision and Pattern Recognition, 2369-2377.
- Nie, S., Jiang, Z., Zhang, H., Cai, B. ve Yao, Y., 2018. Inshore Ship Detection Based on Mask R-CNN. In IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, IEEE, 693-696.
- Ongsulee, P., 2017. Artificial Intelligence, Machine Learning and Deep Learning. In 2017 15th International Conference on ICT and Knowledge Engineering (ICT&KE), IEEE, 1-6.
- Pal, S. K., Pramanik, A., Maiti, J. ve Mitra, P., 2021. Deep Learning in Multi-Object Detection and Tracking: State of The Art. Applied Intelligence, 51, 6400-6429.
- Pan, S. J. Ve Yang, Q., 2009. A Survey on Transfer Learning, IEEE Transactions on Knowledge and Data Engineering, 22, 1345-1359.

- Petropoulos, G.P., Vadrevu, K., P., Xanthopoulos, G., Karantounias, G. ve Scholze, M., 2010. A Comparison of Spectral Angle Mapper and Artificial Neural Network Classifiers Combined with Landsat TM Imagery Analysis for Obtaining Burnt Area Mapping, Sensors, 10, 1967-1985.
- Potlapally, A., Chowdary, P. S. R., Shekhar, S. R., Mishra, N., Madhuri, C. S. V. D. ve Prasad, A. V. V., 2019. Instance Segmentation in Remote Sensing Imagery Using Deep Convolutional Neural Networks, In 2019 International Conference on Contemporary Computing and Informatics, IEEE, 117-120.
- Redmon, J., Divvala, S., Girshick, R. ve Farhadi, A., 2016. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of The IEEE Conference on Computer Vision and Pattern Recognition, 779-788.
- Redmon, J., ve Farhadi, A., 2017. YOLO9000: Better, Faster, Stronger. In Proceedings of The IEEE Conference on Computer Vision and Pattern Recognition, 7263-7271.
- Ren, S., He, K., Girshick, R. ve Sun, J., 2015. Faster R-Cnn: Towards Real-Time Object Detection with Region Proposal Networks, Advances in Neural Information Processing Systems, 28, 91-99.
- Prinz, T., Lasar, B. ve Krüger, K. P. 2010. High-Resolution Remote Sensing and GIS Techniques for Geobase Data Supporting Archaeological Surveys: A Case Study of Ancient Doliche, Southeast Turkey. Geoarchaeology, 25, 352-374.
- Richards, J. A. ve Jia, X., 2006. Remote Sensing Digital Image Analysis, Fourth Edition, Springer, Berlin, Heidelberg.
- Salakhutdinov, R., Mnih, A. ve Hinton, G., 2007. Restricted Boltzmann Machines for Collaborative Filtering, In Proceedings of the 24th International Conference on Machine Learning, 791-798.
- Salakhutdinov, R. ve Hinton, G., 2009. Deep Boltzmann Machines, Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics, PMLR, Florida USA, 448-455.
- Saravanan, K. ve Sasithra, S., 2014. Review on Classification Based on Artificial Neural Networks, International Journal Ambient Systems and Applications, 2, 4, 11-18.
- Schmidhuber, J., 1992. Learning Complex, Extended Sequences Using the Principle of History Compression, Neural Computation, 4,2, 234-242.

- Schwenk, H. ve Gauvain, J., L., 2005. Training Neural Network Language Models on Very Large Corpora, In Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing, 201-208.
- Shen, Z., Liu, Z., Li, J., Jiang, Y. G., Chen, Y. ve Xue, X., 2017. Dsod: Learning Deeply Supervised Object Detectors from Scratch. In Proceedings of The IEEE International Conference on Computer Vision, 1919-1927.
- Simonyan, K. ve Zisserman, A., 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition, arXiv:1409.1556 10 April 2015
- Su, H., Wei, S., Yan, M., Wang, C., Shi, J. ve Zhang, X., 2019. Object Detection and Instance Segmentation in Remote Sensing Imagery Based on Precise Mask R-CNN. In IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium, 1454-1457.
- Taylor, G., W., Hinton, G., E. ve Roweis, S., T., 2007. Modeling Human Motion Using Binary Latent Variables, In Advances in Neural Information Processing Systems, 1345-1352.
- Tong, X. Y., Lu, Q., Xia, G. S. ve Zhang, L. 2018. Large-Scale Land Cover Classification in Gaofen-2 Satellite Imagery. In IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium. 3599-3602.
- Uijlings, J., R., Van De Sande, K., E., Gevers, T. ve Smeulders, A., W., 2013. Selective Search for Object Recognition. International Journal of Computer Vision, 104,2, 154-171.
- Venkatesan, R. ve Li, B., 2017. Convolutional Neural Networks in Visual Computing: A Concise Guide, CRC Press, 6000 Broken Sound Parkway NW.
- Viglione, S., S. 1970. 4 Applications of Pattern Recognition Technology, In Mathematics in Science and Engineering, Elsevier, Volume 66, 115-162.
- Wang, D., He, H. ve Liu, D., 2017. Adaptive Critic Nonlinear Robust Control: A Survey, IEEE Transactions on Cybernetics, 47, 3429-3451.
- Wani, M., A., Bhat, F., A., Afzal, S. ve Khan, A., I., 2020. Advances in Deep Learning, Springer, Warsaw, Poland.
- Weiss, K., Khoshgoftaar, T. M. ve Wang, D., 2016. A Survey of Transfer Learning, Journal of Big Data, 3, 1-40.

- Xu, C., Du, X., Yan, Z. ve Fan, X. 2020. ScienceEarth: A Big Data Platform for Remote Sensing Data Processing. Remote Sensing, 12, 607.
- Yoo, D., Park, S., Lee, J. Y., Paek, A. S., ve So Kweon, I., 2015. Attentionnet: Aggregating Weak Directions for Accurate Object Detection. In Proceedings of The IEEE International Conference on Computer Vision, 2659-2667.
- Yosinski, J., Clune, J., Bengio, Y. ve Lipson, H., 2014. How Transferable are Features in Deep Neural Networks. Advances in Neural Information Processing Systems, 27.
- Zaidi, S. S. A., Ansari, M. S., Aslam, A., Kanwal, N., Asghar, M. ve Lee, B., 2022. A Survey of Modern Deep Learning Based Object Detection Models. Digital Signal Processing, 103514.
- Zeng, Nick., An Introduction to Evaluation Metrics for Object Detection, <https://blog.zenggyu.com/en/post/2018-12-16/an-introduction-to-evaluation-metrics-for-object-detection/>, 30 Haziran 2021
- Zhang, L., Tan, J., Han, D. ve Zhu, H., 2017. From Machine Learning to Deep Learning: Progress in Machine Intelligence for Rational Drug Discovery. Drug Discovery Today, 22, 1680-1685.
- Zhang, L., Li, W., Shen, L. ve Lei, D. 2020. Multilevel Dense Neural Network for Pan-Sharpening. International Journal of Remote Sensing, 41, 7217-7232.
- Zhang, X., Zhou, Y. N. ve Luo, J. 2021. Deep Learning for Processing and Analysis of Remote Sensing Big Data: A Technical Review. Big Earth Data, 1-34.
- Zhao, X., Gao, L., Chen, Z., Zhang, B. ve Liao, W., 2018. CNN-Based Large Scale Landsat Image Classification. In 2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC). IEEE. 611-617.
- Zhu, X. X., Tuia, D., Mou, L., Xia, G. S., Zhang, L., Xu, F. ve Fraundorfer, F., 2017. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources, IEEE Geoscience and Remote Sensing Magazine, 5, 8-36.
- Zhu, X. J., 2005. Semi-Supervised Learning Literature Survey, University of Wisconsin-Madison Department of Computer Sciences.
- Zin, T., T. ve Lin, J., C., W., 2018. Big Data Analysis and Deep Learning Applications: Proceedings of the First International Conference on Big Data Analysis and Deep Learning, Vol. 744, Springer, Warsaw, Poland.

URL-1 <https://www.dataiversity.net/brief-history-deep-learning/#>, 01.05.2022

URL-2, [http://nik.com.tr/content\\_sistem\\_uydu.asp?id=102](http://nik.com.tr/content_sistem_uydu.asp?id=102), Sentinel-2, 2 Nisan 2022

URL-3, <https://earth.esa.int/eogateway/missions/worldview-3>, 2 Nisan 2022

URL-4, <https://www.robots.ox.ac.uk/~vgg/software/via/via.html>, 3 Ocak 2022

URL-5, [https://github.com/matterport/Mask\\_RCNN](https://github.com/matterport/Mask_RCNN), 05 Nisan 2020



## ÖZGEÇMİŞ

İlk ve orta öğrenimini Kanuni İlköğretim Okulu'nda lise öğrenimini Tevfik Serdar Anadolu Lisesi'nde gördü. 2013 yılında Karadeniz Teknik Üniversitesi Harita Mühendisliği bölümünü kazandı. 1 yılı İngilizce hazırlık olmak üzere 5 yıllık lisans eğitimini 2018 yılında tamamladı. Aynı yıl Karadeniz Teknik Üniversitesinde yüksek lisans eğitimine başladı. 2021 yılında İtalya'daki Pavia Üniversitesi Geomatic laboratuvarında staj yaptı. 2022 Haziran ayından itibaren DSİ 8. Bölge Müdürlüğünde Harita Mühendisi olarak çalışmaktadır. B2 seviye İngilizce bilmektedir.

