T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

TRAVMA SONRASI KEMİK KIRIKLARININ TESPİTİNDE BİLGİSAYARLI GÖRÜ VE DERİN ÖĞRENME ALGORİTMALARI

DOKTORA TEZİ

Muhammed Taha ZEREN

Endüstri Mühendisliği Anabilim Dalı

Endüstri Mühendisliği Bilim Dalı

KASIM 2023

T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

TRAVMA SONRASI KEMİK KIRIKLARININ TESPİTİNDE BİLGİSAYARLI GÖRÜ VE DERİN ÖĞRENME ALGORİTMALARI

DOKTORA TEZİ

Muhammed Taha ZEREN

Endüstri Mühendisliği Anabilim Dalı

Endüstri Mühendisliği Bilim Dalı

Tez Danışmanı: Doç.Dr. Seher ARSLANKAYA

KASIM 2023

Muhammed Taha Zeren tarafından hazırlanan "Travma Sonrası Kemik Kırıklarının Tespitinde Bilgisayarlı Görü ve Derin Öğrenme Algoritmaları " adlı tez çalışması 10.10.2023 tarihinde aşağıdaki jüri tarafından oy birliği/oy çokluğu ile Sakarya Üniversitesi Fen Bilimleri Enstitüsü Endüstri Mühendisliği Anabilim Dalı Endüstri Mühendisliği Bilim Dalı'nda Doktora tezi olarak kabul edilmiştir.

Tez Jürisi

Jüri Başkanı : **Doç. Dr. Seher ARSLANKAYA** (Danışman) .....
Sakarya Üniversitesi

Jüri Üyesi : **Doç. Dr. Tijen ÖVER ÖZÇELİK** .....
Sakarya Üniversitesi

Jüri Üyesi : **Doç. Dr. Ayten YILMAZ YALÇINER** .....
Sakarya Üniversitesi

Jüri Üyesi : **Dr. Öğr. Üyesi S. Kerem AYTULUN** .....
İstanbul Arel Üniversitesi

Jüri Üyesi : **Dr. Öğr. Üyesi Pelin ALCAN GEZGİNCİ** .....
Okan Üniversitesi

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ

Sakarya Üniversitesi Fen Bilimleri Enstitüsü Lisansüstü Eğitim-Öğretim Yönetmeliğine ve Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesine uygun olarak hazırlamış olduğum "TRAVMA SONRASI KEMİK KIRIKLARININ TESPİTİNDE BİLGİSAYARLI GÖRÜ VE DERİN ÖĞRENME ALGORİTMALARI" başlıklı tezin bana ait, özgün bir çalışma olduğunu; çalışmamın tüm aşamalarında yukarıda belirtilen yönetmelik ve yönergeye uygun davrandığımı, tezin içerdiği yenilik ve sonuçları başka bir yerden almadığımı, tezde kullandığım eserleri usulüne göre kaynak olarak gösterdiğimi, bu tezi başka bir bilim kuruluna akademik amaç ve unvan almak amacıyla vermediğimi ve 20.04.2016 tarihli Resmi Gazete'de yayımlanan Lisansüstü Eğitim ve Öğretim Yönetmeliğinin 9/2 ve 22/2 maddeleri gereğince Sakarya Üniversitesi'nin abonesi olduğu intihal yazılım programı kullanılarak Enstitü tarafından belirlenmiş ölçütlere uygun rapor alındığını, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun ortaya çıkması halinde doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi beyan ederim.

(……/……/20…..).

(imza)

Muhammed Taha ZEREN

Eşime, kızıma ve aileme…

TEŞEKKÜR

Doktora çalışmam boyunca desteğini esirgemeyen, her daim kıymetli yönlendirmesi ile engebeli ve zorlu başarı patikasında, doğru yolu gösteren saygı değer danışman hocam Doç. Dr. Seher ARSLANKAYA'ya, çalışma sırasında gecesini gündüzüne katarak desteğini esirgemeyen Şişli Hamidiye Etfal Eğitim Araştırma Hastanesi Ortopedi ve Travmatoloji departmanı Doktor ve Asistanlarına, akademik yolculuğumun her adımında değerine paha biçemeyeceğim destekleri, fedakarlıkları ve sabırları için eşime, kızıma ve aileme sonsuz teşekkürler.

Muhammed Taha Zeren

İÇİNDEKİLER

Sayfa

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ v
TEŞEKKÜR ix
İÇİNDEKİLER xi
KISALTMALAR xv
SİMGELER xvii
TABLO LİSTESİ xix
ŞEKİL LİSTESİ xxi
ÖZET xxiii
SUMMARY xxvii
1. GİRİŞ 1
1.1. Tezin Kapsamı 2
1.2. Tezin Amacı ve Önemi 3
1.3. Tezin Yöntemi ve Özgün Katkısı 4
1.4. Tezin Çalışma Planı 6
1.5. İnsanda Kas İskelet Sistemi 7
1.6. İnsanda Kemik Kırıkları 9
1.7. İnsanda Femur Üst Bölge 12
1.8. İnsanda Kas İskelet Sistemi Görüntüleme Yöntemleri 15
1.8.1. X-ışını görüntüleme (radyografi) 15
1.8.2. Manyetik rezonans (MRI) 17
1.8.3. Bilgisayarlı tomografi (Computerized Tomography) 19
1.8.4. Ultrasonografi 21
2. KAYNAK ARAŞTIRMASI 25
3. MATERYAL VE YÖNTEM 41
3.1. Bilgisayarların Yükselişi ve Yapay Zekâ Kavramı 41
3.2. Yapay Zekâ Çeşitleri 43
3.3. Makine Öğrenmesi Kavramı ve Temelleri 45
3.3.1. Öğrenme algoritmaları 46
3.3.2. Kapasite, aşırı uyum (overfitting), eksik uyum (underfitting) 47
3.3.3. Hiper parametreler ve doğrulama kümeleri 48
3.3.4. Tahminleyiciler, bias ve varyans 49
3.3.5. Denetimli öğrenme algoritmaları 49
3.3.6. Gözetimsiz öğrenme 51
3.3.7. Stokastik gradyan azalma 52
3.4. Derin Öğrenme Kavramı 52
3.4.1. Parametre norm cezaları 53
3.4.2. Gürültü sağlamlığı 54
3.4.3. Yarı yardımlı öğrenme 54
3.4.4. Çok görevli öğrenme 54
3.4.5. Erken durdurma 55
3.4.6. Çekişmeli öğrenme 55

3.5. Yapay Sinir Ağları 55
3.5.1. Yapay sinir ağlarının yapısı 56
3.5.2. Aktivasyon fonksiyonları 58
3.5.2.1. Sigmoid fonksiyonu 58
3.5.2.2. Hiperbolik tanjant fonksiyonu (Tanh) 59
3.5.2.3. ReLU (Rectified linear unit) fonksiyonu 60
3.5.2.4. ELU (Exponential linear unit) aktivasyon fonksiyonu 61
3.6. Bilgisayarlı Görü ve Evrişimli Sinir Ağları 62
3.6.1. Evrişimli sinir ağları katmanları 63
3.6.1.1. Giriş katmanı (Input layer) 63
3.6.1.2. Konvolüsyon katmanı (Convulation layer) 64
3.6.1.3. Havuzlama katmanı (Pooling layer) 66
3.6.1.4. Bırakma katmanı (Dropout) 67
3.6.1.5. Tam bağlı katman (Fully connected layer) 68
3.6.1.6. Sınıflandırma katmanı 68
3.6.2. Yapay sinir ağı modelleri 69
3.6.2.1. MobileNet 69
3.6.2.2. Inception 70
3.6.2.3. LeNet-5 70
3.6.2.4. AlexNet 70
3.6.2.5. VGG (Visual Geometry Group) 71
3.6.2.6. ResNet 71
3.6.3. Görüntü sınıflandırma ve kritik bölge tespitimde kullanılan derin öğrenme algoritmaları 72
3.6.3.1. R-CNN 72
3.6.3.2. Fast R-CNN 74
3.6.3.3. Faster R-CNN 75
3.6.3.4. Single shot multibox detector (SSD) 76
3.6.3.5. You look only once (YOLO) 78
3.6.4. Görüntü sınıflandırma eğitimi çeşitleri 83
3.6.4.1. Sıfırdan eğitim (Training from scratch) 83
3.6.4.2. Aktarımlı öğrenme (Transfer learning) 83
3.6.5. Görüntü sınıflandırma eğitimi ve kullanılan veri setleri 84
3.6.5.1. ImageNet veri seti 84
3.6.5.2. MS COCO veri seti 85
3.6.5.3. MNIST veri seti 86
3.6.5.4. CIFAR-10 / CIFAR-100 86
3.6.6. Veri arttırma teknikleri 87
4. ARAŞTIRMA BULGULARI 89
4.1. Araştırma Çıktıları ve Performans Değerlendirme Kriterlerinin Belirlenmesi 89
4.1.1. Karmaşıklık matrisi ve çoklu performans değerlendirme tablosu 89
4.1.2. ROC eğrisi 91
4.1.3. Loss oranı grafikleri 92
4.2. Araştırmada Kullanılan Algoritmalar ve Hekim Gruplarının Seçilmesi 93
4.3. Araştırmada Kullanılan Veri Setlerinin Oluşturulması ve Özellikleri 94
4.4. Çalışmada Kullanılan Yazılım Dili ve Yapılar 99
4.5. Araştırmada Sonuçlarında Kullanılan İleri İstatistiksel Analizler 100
5. TARTIŞMA VE SONUÇ 103
5.1. Araştırmada Yer Alan Pratisyen Hekimlerin ve Ortopedi Uzmanlarının Ürettiği Sonuçlar 103

5.2. YOLO Algoritmasının Ürettiği Sonuçlar 106
5.3. Faster R-CNN Algoritmasının Ürettiği Sonuçlar 111
5.4. SSD Algoritmasının Ürettiği Sonuçlar 115
5.5. Bütün Derin Öğrenme Algoritmaları ile Hekim Gruplarının Ürettiği Sonuçların Kıyaslanması 120
5.6. Bütün Derin Öğrenme Algoritmaları ile Hekim Gruplarının Ürettiği Sonuçların İstatistiksel Olarak Yorumlanması 123
5.7. Sonuç ve Öneriler 129
KAYNAKLAR 133
ÖZGEÇMİŞ 145

KISALTMALAR

ACC : Accuracy – Doğruluk Oranı
ANN : Artificial Neural Network
ANOVA : Analysis of variance
AO : Arbeitsgemenienschaft für Osteosynthesefragen
AUC : Area Under the ROC Curve
BPNN : Back Propagation Neural Network
CIFAR : Canadian Institute for Advanced Research
CNN : Convolution Neural Network
CT : Computerized Tomograph
DCNN : Deep Convolutional Neural Networks
DRLSE : Distance Regularized Level Set Evolution
ERR : Error Rate – Hata Oranı
FPR : False Positive Rate – Yanlış Pozitif Oranı
FNR : False Negative Rate – Yanlış Negatif Oranı
FCM : Fuzzy C-Means
ILSVRC : ImageNet Large Scale Visual Recognition Challenge
MNIST : Modified National Institute of Standards and Technology
MRI : Magnetic Resonance Imaging
MS COCO : Microsoft Common Objects in Context
NMS : Non-Maximum Suppression
NPV : Negative Pred Value – Negatif Tahmin Oranı
OTA : Orthopedic Trauma Association
PMCT : Postmortem Computed Tomography
PPV : Positive Pred Value – Precision – Doğru Tahmin Oranı
R-CNN : Regional Convolution Neural Network
ReLU : Rectified Linear Unit
RF : Radyo Frekans
RGB : Red Green Blue
RPN : Region Proposal Network
ROC : Receiver Operating Characteristics

SaaS : Software as a Service

SGD : Stochastic Gradient Descent

SIFT : Scale-Invariant Feature Transform

SSD : Single Shot Multibox Detector

SVM : Support Vektor Mission

TPR : True Positive Rate – Doğru Pozitif Oranı

TNR : True Negative Rate – Doğru Negatif Oranı

VGG : Visual Geometry Group

WFP : Wavelet Fuzzy Phrases

YOLO : You Look Only Once

SİMGELER

$\mathbf{A_{ctr}}$	: Tahmin çerçevesinin merkez noktası koordinatlarını
$\boldsymbol{B_{ctr}}$	: Gerçek çerçevenin merkez noktası koordinatlarını
$\boldsymbol{\alpha}$	: Pozitif bir katsayı
$\boldsymbol{\lambda}$	: En-boy oranının tutarlılığı
w	: Tahmin çerçevenin genişliği
h	: Tahmin çerçevenin yüksekliği

TABLO LİSTESİ

Sayfa

Tablo 1.1. Diyafiz kırığını sınıflandırma ve kodlama süreci (Meinberg ve ark., 2018). 12
Tablo 4.1. Karmaşıklık matrisi. 90
Tablo 4.2. Çoklu değerlendirme kriterleri tablosu (Zeren ve ark., 2020). 91
Tablo 5.1. Çoklu değerlendirme kriterleri tablosu pratisyen ve uzman hekim sonuçları. 105
Tablo 5.2. Pratisyen ve uzman hekimlerin sonuç üretme süreleri (saniye). 105
Tablo 5.3. Çoklu değerlendirme kriterleri tablosu YOLO Algoritması sonuçları .. 108
Tablo 5.4. YOLO Algoritması sonuç üretme süreleri (saniye). 109
Tablo 5.5. Çoklu değerlendirme kriterleri tablosu Faster R-CNN Algoritması sonuçları. 112
Tablo 5.6. Faster R-CNN Algoritması sonuç üretme süreleri (saniye). 113
Tablo 5.7. Çoklu değerlendirme kriterleri tablosu SSD algoritması sonuçları 117
Tablo 5.8. SSD Algoritması sonuç üretme süreleri (saniye) 117
Tablo 5.9. Çoklu değerlendirme kriterleri tablosu test veri setine göre bütün sonuçlar 121
Tablo 5.10. Bütün çalışma grularının sonuç üretme süreleri (saniye)..... 122
Tablo 5.11. Femur üst bölge sonuç üretme süreleri çarpıklık ve basıklık analizi. .. 124
Tablo 5.12. Femur üst bölge kırık olan bölge doğruluğu çarpıklık ve basıklık analizi. 124
Tablo 5.13. Femur üst bölge kırık olmayan bölge doğruluğu çarpıklık ve basıklık analizi. 125

ŞEKİL LİSTESİ

Sayfa

Şekil 1.1. İnsan iskelet sistemi (Andrew, 2023). 8
Şekil 1.2. Travma sonrası çok parçalı kırık meydana gelen femur üst bölge görüntüsü. 10
Şekil 1.3. Kemik konumunun belirlenmesi (Meinberg ve ark., 2018). 11
Şekil 1.4. Femur kemiğinin anatomik yapısı (Waschke ve ark., 2011). 13
Şekil 1.5. Femur üst bölge anatomik yapısı. 14
Şekil 1.6. X-Işını Oluşumu (Ünal, 2008). 16
Şekil 1.7. MR görüntüleme cihazı (Ünal, 2008). 18
Şekil 1.8. MR görüntülemesi ile tanısı konulan beyin kanaması görüntüsü. 18
Şekil 1.9. Akçiğer grafisi ve bilgisayarlı tomografi cihazı görseli (Ünal, 2008). 20
Şekil 1.10. Yeni nesil bir ultrason cihazı 22
Şekil 3.1. Yapay Zeka'nın alt disiplinlerini gösteren Venn Şeması. 43
Şekil 3.2. Makine öğrenmesi işleyişi ile klasik programlama (Chollet, 2017). 46
Şekil 3.3. Makine öğrenmesi yetersiz uydurma, uygun kapasite ve aşırı uydurma örnekleri. 48
Şekil 3.4. 65 Yaş üstü insanların yaz aylarında sokağa çıkması ile ilgili karar ağacı. 51
Şekil 3.5. Kümeleme analizi örneği. 52
Şekil 3.6. Makine öğrenmesi ve derin öğrenme modellerinin çalışma sistematiği. .. 53
Şekil 3.7. İnsan vücudunda yer alan nöron hücresi (Zhang ve ark., 2018). 56
Şekil 3.8. Algılayıcıların şematik gösterimi (Shanmugamani, 2018). 57
Şekil 3.9. Tek ve çok katmanlı sinir ağları modelleri (Nielsen, 2015). 57
Şekil 3.10. Sigmoid fonksiyonu. 59
Şekil 3.11. Hiperbolik Tanjant (Tanh) fonksiyonu. 60
Şekil 3.12. ReLU fonksiyonu. 61
Şekil 3.13. ELU fonksiyonu. 62
Şekil 3.14. Konvolüsyonel sinir ağı katmanları. 63
Şekil 3.15. Kernel (çekirdek) matris önerileri (Zeren, 2020). 65
Şekil 3.16. Konvolüsyon işlemi örneği (Zeren, 2020). 66
Şekil 3.17. Maximum pooling layer örneği (Yani, 2019). 67
Şekil 3.18. Average pooling layer örneği (Yani, 2019). 67
Şekil 3.19. Bırakma katmanının sinir ağına uygulanması (Budhiraja, 2016). 68
Şekil 3.20. Flattening-düzleştirme işlemi. 68
Şekil 3.21. MobileNet V1 ve V2 mimarileri. 69
Şekil 3.22. LeNet Mimarisi (LeCun ve ark., 1998). 70
Şekil 3.23. AlexNet Mimarisi. 71
Şekil 3.24. VGG Mimarisi. 71
Şekil 3.25. ResNet Mimarisi (He ve ark., 2015). 72
Şekil 3.26. R-CNN çalışma yapısı (Girshick ve ark., 2014). 73
Şekil 3.27. R-CNN nesne tespiti örneği (Girshick ve ark., 2014). 73
Şekil 3.28. Fast R-CNN mimarisi (Girshick, 2015). 74

Şekil 3.29. Faster R-CNN mimarisi (Ren ve ark., 2015). 75
Şekil 3.30. Single shot multibox detector algoritması mimarisi (Liu ve ark., 2016). 77
Şekil 3.31. SSD Algoritmasında nesne çerçevelemesi (Liu ve ark., 2016). 78
Şekil 3.32. YOLO Mimarisi (Shafiee ve ark., 2017). 78
Şekil 3.33. YOLO darknet v4 detaylı yapısı (Gao ve ark., 2021). 79
Şekil 3.34. Öklitsel uzaklık diagramı (Gao ve ark., 2021). 81
Şekil 3.35. Hedef objenin resimdeki pozisyonu. 82
Şekil 3.36. ImageNet Veri Seti (Massouh ve ark., 2017). 84
Şekil 3.37. MS COCO Veri Seti (Lin ve ark., 2014). 85
Şekil 3.38. MNIST Veri Seti (Lecun ve ark., 1998). 86
Şekil 3.39. CIFAR-10 Veri Seti (ImageNet, 2016). 87
Şekil 4.1. ROC eğrisi analizi örneği (Lane ve Gantley, 2017). 92
Şekil 4.2. Femur Üst Bölge Kırıklarının Yaş Aralığı Dağılımı. 94
Şekil 4.3. İstanbul İli Nüfus Yaş Dağımı. 95
Şekil 4.4. Femur Üst Bölge Kırıklarının İstanbul İli Nüfusuna Göre Sıklığı. 96
Şekil 4.5. İstanbul İli Cinsiyet Dağılımı. 96
Şekil 4.6. Hasta Cinsiyet Dağılımı. 97
Şekil 4.7. Eğitim veri seti örnekleri. 98
Şekil 4.8. Doğrulama veri seti örnekleri. 98
Şekil 4.9. Test veri seti örnekleri. 99
Şekil 4.10. Text formatında kırık bölgelerin işaretlenmesi. 100
Şekil 5.1. ROC eğrisi analizinde ortopedi uzman ve pratisyen hekimlerin sonuçları. 106
Şekil 5.2. YOLO algoritması loss grafiği 107
Şekil 5.3. ROC eğrisi analizinde YOLO algoritması sonucu 109
Şekil 5.4. YOLO algoritması ile femur üst bölge kırığı tepiti örneği-1. 110
Şekil 5.5. YOLO algoritması ile femur üst bölge kırığı tepiti örneği-2. 110
Şekil 5.6. YOLO algoritması ile femur üst bölge kırığı tespiti örneği-3. 111
Şekil 5.7. Faster R-CNN algoritması loss grafiği. 112
Şekil 5.8. ROC eğrisi analizinde Faster R-CNN algoritması sonucu. 114
Şekil 5.9. Faster R-CNN algoritması ile femur üst bölge kırığı tepiti örneği-1. 114
Şekil 5.10. Faster R-CNN algoritması ile femur üst bölge kırığı tepiti örneği-2. 115
Şekil 5.11. Faster R-CNN algoritması ile femur üst bölge kırığı tepiti örneği-3. 115
Şekil 5.12. SSD algoritması loss grafiği 116
Şekil 5.13. ROC eğrisi analizinde SSD algoritması sonucu 118
Şekil 5.14. SSD algoritması ile femur üst bölge kırığı tepiti örneği-1 119
Şekil 5.15. SSD algoritması ile femur üst bölge kırığı tepiti örneği-2 119
Şekil 5.16. SSD algoritması ile femur üst bölge kırığı tepiti örneği-2 120
Şekil 5.17. ROC eğrisi analizinde bütün çalışma gruplarının sonuçları 123
Şekil 5.18. Femur üst bölge sonuç üretme süreleri tek yön anova analizi. 125
Şekil 5.19. Femur üst bölge sonuç üretme süreleri Tukey ikili karşılaştırma analizi. 126
Şekil 5.20. Femur üst bölge kırık olan bölge doğruluğu oranı tek yön anova analizi. 127
Şekil 5.21. Femur üst bölge kırık olan bölge doğruluğu oranı Tukey ikili karşılaştırma analizi. 127
Şekil 5.22. Femur üst bölge kırık olan bölge doğruluğu oranı tek yön anova analizi. 128
Şekil 5.23. Femur üst bölge kırık olmayan bölge doğruluğu oranı Tukey ikili karşılaştırma analizi. 129

TRAVMA SONRASI KEMİK KIRIKLARININ TESPİTİNDE BİLGİSAYARLI GÖRÜ VE DERİN ÖĞRENME ALGORİTMALARI

ÖZET

1950'li yıllarda makinelerin düşünebileceği ve öğrenebileceği görüşünün ortaya atılmasından hemen sonra yapay zekâ kavramı ortaya çıkmıştır. 21.yy. geldiğimizde GPU ve CPU'ların gelişimi ile birlikte makine öğrenmesi kavramının içerisinde yer alan derin öğrenme insanoğlunun hayatında vazgeçemeyeceği bir noktaya erişmiştir. Gündelik yaşantıda insanoğlu artık ulaşımdan, güvenliğe, üretimden sağlığa kadar birçok alanda bilgisayarlı görüde derin öğrenme algoritmalarının kullanılmasının faydasını görmektedir. Bu çalışmada özellikle yaşlı popülasyonun karşılaştığı en kritik ve ölüm/sakatlanma gibi sonuçların en çok ortaya çıktığı sağlık problemlerinden biri olan femur üst bölge (kalça) kırıkları odağa alınmıştır. Dünya genelinde yapılan araştırmalar yaşlı popülasyonun karşılaştığı en kritik sağlık problemlerinden biri kalça kırıkları olduğunu göstermektedir. Beklenen yaşam süresinin uzamasıyla birlikte yapılan araştırmalar 2040 yılına kadar kalça kırıklarının iki kat artması beklenmektedir. Özellikle ortopedi hekimlerinin çözümledikleri bu kalça kırığı vakalarının, dünya genelinde 1990 yılında 1,26 milyon gerçekleşmesine karşın, 2050 yılına gelindiğinde 4,5 Milyona yükseleceği yapılan araştırmalarca ortaya konulmaktadır. Bütün ortopedik kemik kırıkları değerlendirildiğinde sakat kalma ve örüm oranı en yüksek kırık olduğu görülmektedir.

Söz konusu femur üst bölge kalça kırıkları; derin öğrenme ve bilgisayarlı görünün en başarılı algoritmaları arasında yer alan "You Look Only Once - YOLO Darknet v4 algoritması, Faster R-CNN – Inception v2 Algoritması ve SSD - Mobilenet v2 algoritması ile ele alınmıştır. Söz konusu algoritmanın yeniden eğitilmesi adına nüfusu en kalabalık metropollerden biri olan İstanbul'da, kalça kırığı şikâyeti ile ortopedi servisine başvurmuş yaşları 22 ile 105 aralığında bulunan 500 üzeri hastanın X-ray görüntüsünde inceleme yapılmıştır. Bu incelemeler sonunda 410 hastanın X-ray görüntüsü veri çoklandırma tekniklerinden döndürme, ölçeklendirme kırpma ve yeniden boyutlandırma ile 820'ye çıkarılmıştır. Toplamda 1514 ((820-63)x2) adet femur üst bölge görüntüsüne yer verilmiştir. Söz konusu 820 X-ray görüntüsü %80 eğitim ve %20 doğrulama veri seti olarak ayrıştırılmıştır. Daha önce COCO – Common Object in Context veri setinde eğitilmiş olan YOLO Algoritması, dünyanın en güçlüsü olarak kabul edilmiş GPUlarından biri olan Tesla K80 24 GB GDDR5 GPUsu ile Google Colab üzerinden eğitim 5.000 iterasyon süresince eğitilmiştir. Bu eğitim 13 saat 6 dakikada tamamlanmıştır. Daha önce COCO veri setinde eğitilmiş olan Faster R-CNN algoritması ise Core i5-8300H CPU ve NVIDIA GeForce GTX 1050 GPU ya sahip iş istasyonu GPU'su ile eğitimi gerçekleştirilmiş, söz konusu 5000 adımdan oluşan eğitim 31,9 dakika sürmüştür. Yine daha önce COCO veri setinde eğitilmiş olan SSD Algoritması da yine Faster R-CNN Algoritması gibi Core i5-8300H CPU ve NVIDIA GeForce GTX 1050 GPU ya sahip iş istasyonu kullanılmış, toplamda 5000 iterasyona sahip eğitim 55,92 dk. da tamamlanmıştır.

Söz konusu algoritmanın başarısının performansının doğru ortaya koymak adına eğitim ve araştırma hastanesinden ortopedi uzman doktorları ve pratisyen hekimlerinden gönüllü olanları çalışmaya dahil edilmiştir. Ortopedi uzman hekimleri ve asistanlarından 12 kişilik grup oluşturulmuştur. Temel 6 yıllık tıp eğimi almış pratisyen hekimlerden oluşan 10 kişilik grup oluşturulmuştur. Bu iki farklı gruba X-ray görüntüleri üzerinden femur üst bölge kırıklarını tespit etmek üzere yeniden eğitilmiş olan algoritmaya sunulan test veri setindeki görüntüler karışık sıra ile gösterilmiştir. Ortopedi uzmanı hekimleri ve pratisyen hekimlerine tek tek gösterilen resimlerdeki hem tespit oranları hem de resim ilk karşılarında belirmesi ile tespit yapması arasındaki geçen süreler kayıt altına alınmıştır.

Eğitim ve doğrulama veri seti ve bu iki veri setinden bağımsız olarak değerlendirmeyi amaçlayan özgün toplamda 50 femur üst bölge kırığı içeren 50 sağlam femur üst bölgesini içeren toplamda 55 X-ray görüntüsünden oluşan test veri seti olmak üzere birbirinden bağımsız 3 veri seti oluşturulmuştur. Eğitim ve doğrulama veri setlerinden bağımsız ve objektif bir şekilde değerlendirmeyi yapılmasını sağlaması planlanan test veri seti femur üst bölge kırığı tespit etmek üzere yeniden eğitilmiş YOLO, Faster R-CNN ve SSD algoritmaları ve 2 hekim grubu olmak üzere 5 farklı grubun ürettiği sonuçlar değerlendirmek üzere kullanılmıştır.

Elde edilen sonuçlar, doğruluk oranı, duyarlılık, özgüllük, yanlış pozitif oranı, yanlış negatif oranı, doğru tahmin oranı, F1 puanı, hata oranı, sonuç üretme süresi ve ROC eğrileri gibi değerlendirme kriterleri kapsamında değerlendirilmiştir. Değerlendirme kriterlere ele alındığında doğruluk oranı modelin hem ilgili kritik bölge bulunan hem de bulunmayan görsellerde ne kadar bir %'lik başarı sağladığını göstermektedir. Duyarlılık olarak ifade edilen doğru pozitif oranı modelin kırık olan bölgeye sahip olan görüntülerde ne kadar doğru tespit ettiğini göstermektedir. Özgüllük olarak belirtilen doğru negatif oranı ise içerisinde kırık olan bölge barındırmayan görüntülerde ne kadar modelin başarılı olduğunu göstermektedir. Yanlış pozitif oranı 1- doğru negatif oranına eşit olduğu ifade edilebileceği gibi kırık olan bölge içermeyen görüntülerde modelin ne kadar yanılgıya düştüğü olarak da ifade edilebilmektedir. Yanlış negatif oranı ise 1-doğru pozitif oranı olarak düşünebileceği gibi kırık olan bölge içeren görüntülerde modelin ne kadar yanılgıya düştüğü olarak da gösterilebilmektedir. Doğru tahmin oranı modelin kırık olan bölge var dediği görüntülerde ne kadar başarılı olduğunu göstermektedir. F1 puanı doğru pozitif oranı ile doğru tahmin oranının harmonik ortalamasını göstermektedir. Hata oranı ise 1-doğruluk oranıdır.

Derin öğrenme algoritmalarının kendi içlerindeki eğitim sürecini kıyaslamak toplam kayıp (Total Loss) grafikleri kıyaslanıp incelendiğinde 5000 iterasyonda YOLO algoritmasının 18 seviyesinin üzerinden başlayarak 1000 iterasyona vardığında 1'in altına indiği görülmektedir. Hızlı bir öğrenme süreci geçirdiği tespit edilmiştir. Faster R-CNN algoritması eğitim sürecinde toplam kayıp grafiği incelendiğinde 5000 iterasyonda 0,6 seviyesinden başlayarak 1000 iterasyona vardığında 0,2'in altına indiği gözükmektedir. Faster R-CNN algoritmasının da YOLO Algoritması gibi hızlı bir öğrenme süreci geçirdiği gözükmektedir. SSD algoritması için eğitim sürecinde toplam kayıp grafiği incelendiğinde 5000 iterasyonda 11 seviyesinden başlayarak 1000 iterasyona vardığında 4'in altına indiği gözükmektedir. SSD Algoritmasının YOLO ve Faster R-CNN algoritmasından farklı olarak daha yavaş bir öğrenme süreci olduğu loss grafiğinden anlaşılmaktadır.

Ortopedi uzman doktorları ve asistanlarından oluşan 1. ekip %91,42 doğrulukta, yeniden eğitilerek ortaya çıkarılan YOLO algoritması doğrulukta %90,33 başarı

gösterirken, yeniden eğitilen Faster R-CNN algoritması %84,29, pratisyen hekimlerden oluşan 2. grup %81,30 doğrulukta, yeniden eğitilen SSD Algoritması %72,74 oranında doğrulukta, başarı göstermiştir. Doğru pozitif oranı olarak ele alındığında en yüksek performans yine %91,67 ile ortopedi uzman doktorları ve asistanlarının olurken sırası ile Faster R-CNN Algoritması %90,64, YOLO Algoritması %87,67, SSD algoritması %78,32, Pratisyen hekim grubu ise %74,80 oranında başarı göstermiştir. Doğru negatif oranında ise en yüksek başarı oranı %92,98 oranı ile YOLO algoritmasının olurken, %91,17 ile bu sonucu, ortopedi uzmanı ve asistanları izlemiş, %87,80 ile pratisyen hekim grubu takip ederken, Faster R-CNN algoritmasının başarısı %77,94 seviyesinde kalırken, SSD algoritması %67,16 seviyesinde başarı gösterebilmiştir.

Söz konusu beş farklı grubun ortalama tespit süreleri incelendiğinde ortalamada en hızlı sonucu 1,4 saniye ile YOLO Algoritması olmuş, Uzman doktorlardan ve pratisyen hekimlerden 3 kattan daha hızlı sonuç üretmiştir. Uzman doktorlar ortalamada 4,73 saniyede tespit yapabilirken pratisyen hekimler ortalama 5,10 saniyede tespit yapabilmişlerdir. Söz konusu tespit sürelerinin standart sapmaları kıyaslandığında 0,25 ile en düşük YOLO Algoritması olurken, 1,89 ile Pratisyen Hekimler sonrasında da 3,05 ile uzman hekimler gelmektedir. Faster R-CNN Algoritması ortalama tespit süresi 10,49 saniye olurken SSD algoritması ise 6,38 saniye ile sonuç üretebilmiştir. YOLO Algoritması her iki insan çalışma grubundan daha hızlı sonuç üretirken Faster R-CNN ve SSD algoritması her iki insan çalışma grubundan daha yavaş sonuç üretmiştir.

Bu yapılan çalışma ile Yapay Sinir Ağlarının kullanıldığı algoritmaların kabuklarının dışına çıkarak resimlerdeki; hayvan, bitki ve nesneler dışında konumları ve şekilleri oldukça farklılaşabilen femur üst bölge kırıklarını da tespit edebileceği ve bu alanda ne kadar başarı elde edebileceğini tarafsız bir test veri seti ile ortaya koymuştur. Yapılan çalışma ile hekimlerin gerçekleştireceği teşhislerde kendilerine saniyeler içerisinde destek olacak birden fazla yapı geliştirilmesi de sağlanmıştır.

Yapılan çalışma ile iskelet sistemi rahatsızlıklarının tespitinde, derin öğrenme algoritmalarından olan YOLO, Faster R-CNN ve SSD algoritmalarının doktorlar karşısında ne kadar başarılı olabildiği ortaya konulmuş, çıkan sonuçlar hem tek yönlü Anova Testleri hem de Tukey Pairwise testleri ile kıyaslanmıştır. Elde edilen bu sonuçlar doğrultusunda öğrenilebilen yapay zekâ programlarının geliştirilerek, yardımcı programlar olarak hekimlerin kullanımına sunulmasının teşhis aşamasında hekimlere önemli katkı sağlayacağı öngörülmektedir. Medikal alanda görüntü tespitinde sonraki çalışmalara ışık tutacak bu çalışma; oluşturulan ve genişletilen özgün veri setleri ile başarı oranının daha da yükseleceği görülmektedir.

COMPUTER VISION AND DEEP LEARNING ALGORITHMS ON POST-TRAUMATIC BONE FRACTURES DETECTION

SUMMARY

The concept of artificial intelligence emerged shortly after the proposition that machines can think and learn in the 1950s. The beginning of 21st century, with the improvement of GPU and CPU technologies, deep learning which is the concept under machine learning, has become an indispensable part of human life. In daily every daily activity, humans can now see the benefits of using deep learning algorithms in various areas such as transportation, security, production, and healthcare through computer vision. This study focuses on one of the most critical health problems faced by the elderly population, which is femoral proximal (hip) fractures, mostly caused to death or disability. Research conducted worldwide indicates that one of the most critical health problems faced by the elderly population is hip fractures. Studies suggest that with the increasing life expectancy, the incidence of hip fractures is expected to double by the end of 2040. Specifically, regarding these hip fracture cases treated by orthopedist, research reveals that while there were 1.26 million occurrences globally in 1990, it is projected to rise to 4.5 million by the year 2050. When evaluating all orthopedic bone fractures, it is evident that hip fractures have the highest rates of disability and mortality.

In this study, femoral proximal fractures were focused using the deep learning algorithms such as the "You Look Only Once - YOLO Darknet v4" algorithm, the "Faster R-CNN – Inception v2" algorithm, and the "SSD - Mobilenet v2" algorithm, which are among the most successful algorithms in computer vision using deep learning. For the retraining of the algorithm, X-ray images were examined which were taking over 500 patients in aged between 22 to 105 with complaints of hip fracture in the orthopedic service of Istanbul, one of the most populous metropolises. As a result of these examinations, the X-ray images of 410 patients were augmented using data augmentation techniques such as rotation, scaling, cropping, and resizing, increasing the number of images to 820. A total of 1514 (820-63*2) femoral proximal images were included. The 820 X-ray images were divided into 80% training and 20% validation datasets.

The YOLO algorithm which is considered one of the most powerful algorithms in the world, previously trained on the COCO (Common Object in Context). The new dataset was used for the retraining on the Tesla K80 24 GB GDDR5 GPU using Google Colab. The training duration is 5,000 iterations. This training was completed in 13 hours and 6 minutes. The Faster R-CNN algorithm which is also previously trained on the COCO dataset, was retrained using the Core i5-8300H CPU and NVIDIA GeForce GTX 1050 GPU. The retraining consisting of 5,000 steps and completed 31.9 minutes. Similarly, the SSD algorithm, previously trained on the COCO dataset, was trained using the Core i5-8300H CPU and NVIDIA GeForce GTX 1050 GPU on a workstation, and the training consisting of 5,000 iterations was completed in 55.92 minutes.

To accurately assess the performance of the algorithm, orthopedic specialist doctors and general practitioners from the training and research hospital were voluntarily included in the study. A group of 12 orthopedic specialist doctors and their assistants was formed. Another group of 10 general practitioners, who had received basic 6-year medical education, was also formed. The same images from the test dataset, which was presented to the retrained algorithm to detect femoral proximal fractures, were shown to the orthopedic specialist doctors and general practitioners in a mixed order. The duration of detection in seconds, the detection success of the fractures and the detection rates were recorded.

Three independent datasets have been created: a training set, a validation set, and a test set. The aim is evaluating the performance independently and objectively. The test dataset consists of a total of 55 X-ray images, with 50 images containing fractures in the femoral proximal region and 50 images with intact femoral proximal. The evaluation involves analyzing the results produced by five different groups, including retrained YOLO, Faster R-CNN, and SSD algorithms, as well as two groups of physicians.

The evaluation criteria includes accuracy rate, sensitivity, specificity, false positive rate, false negative rate, true prediction rate, F1 score, error rate, result generation time, and ROC curves. The accuracy rate indicates the percentage of successful detection in images with and without the critical region of interest. Sensitivity, also known as the true positive rate, measures the accuracy of fracture detection in images containing the fractured region. Specificity, expressed as the true negative rate, demonstrates the success of the model in images that do not contain the fractured region. The false positive rate can be calculated as 1 - the true negative rate, indicating the extent to which the model produces false positives in images without a fractured region. Similarly, the false negative rate can be calculated as 1 - the true positive rate, representing the extent to which the model commits errors in images containing the fractured region. The true prediction rate showcases the model's success in correctly identifying the presence of a fractured region in images. The F1 score represents the harmonic mean of the true positive rate and the true prediction rate. The error rate is equal to 1 - the accuracy rate.

When comparing the training processes of deep learning algorithms by examining the retrained algorithms' total loss graphs, it can be observed that the YOLO algorithm starts above level 18 and decreases below 1 after 1000 iterations, indicating a rapid learning process. Similarly, for the Faster R-CNN algorithm, the total loss graph starts at 0.6 and drops below 0.2 after 1000 iterations, indicating a fast-learning process comparable to YOLO. On the other hand, the SSD algorithm exhibits a slower learning process, as its total loss graph starts at 11 and decreases below 4 after 1000 iterations.

In terms of accuracy, the first group consisting of orthopedic specialist doctors and assistants achieves a 91.42% accuracy rate. The retrained YOLO algorithm achieves a 90.33% accuracy rate, while the retrained Faster R-CNN algorithm achieves an 84.29% accuracy rate. The second group, consisting of general practitioners, achieves an accuracy rate of 81.30%. The retrained SSD algorithm achieves a 72.74% accuracy rate. When considering the true positive rate, the orthopedic specialist doctors and assistants achieve the highest performance with 91.67%, followed by the Faster R-CNN algorithm with 90.64%, the YOLO algorithm with 87.67%, the SSD algorithm with 78.32%, and the general practitioner group with 74.80%. As for the true negative rate, the YOLO algorithm achieves the highest success rate of 92.98%, followed by

the orthopedic specialist doctors and assistants with 91.17%, the general practitioner group with 87.80%, the Faster R-CNN algorithm with 77.94%, and the SSD algorithm with 67.16%.

When examining the average detection durations of the five groups, the YOLO algorithm generates the fastest results with an average of 1.4 seconds, which is more than three times faster than both the specialist doctors (4.73 seconds) and the general practitioners (5.10 seconds). The standard deviations of these detection times are lowest for the YOLO algorithm (0.25), followed by the general practitioner group (1.89), and then the specialist doctors (3.05). The Faster R-CNN algorithm has an average detection time of 10.49 seconds, while the SSD algorithm produces results in 6.38 seconds. The YOLO algorithm outperforms both human working groups in terms of speed, while the Faster R-CNN and SSD algorithms are slower than both human working groups.

Receiver Operating Characteristic (ROC) curves are commonly used to measure prediction and classification performance. In this evaluation method, the True Positive Rate is plotted on the y-axis, and the False Positive Rate is plotted on the x-axis. Results that are closest to the top left corner of the ROC curve, which also lie above the diagonal curve, are considered to have better performance. When the True Positive Rate and False Positive Rate of the retrained YOLO algorithm, the general practitioners and the specialist doctors are compared using the ROC curve, the specialist doctors exhibit the most successful results, followed by the YOLO algorithm, and finally the general practitioners. According to the ROC curve, the Faster R-CNN algorithm does not outperform specialist doctors, and both the specialist doctors and general practitioners outperform the SSD algorithm.

This study has demonstrated that algorithms leveraging Artificial Neural Networks can extend beyond their conventional applications in recognizing animals, plants, and objects in images. They can also effectively detect femoral proximal fractures, which exhibit significant variations in both location and shape. The study substantiates the potential success rates achievable in this domain through the utilization of an impartial test dataset. Furthermore, the study has facilitated the development of multiple frameworks that can provide instantaneous support to healthcare practitioners during the diagnostic process.

During the training of the YOLO algorithm, the Tesla K80 24 GB GDDR5 GPU, considered one of the most powerful GPUs in the world, was used via Google Colab. It is evident that as more powerful processors, and potentially quantum computers, continue to advance and become directly available to humanity, the training duration of retrained algorithms will further decrease. When evaluating the test dataset using multiple retrained algorithms, it is expected that the use of a more powerful processor, particularly replacing the NVIDIA GeForce GTX 1050 GPU, will further decrease average processing times. Furthermore, the utilization of quantum computers in this field, particularly, will contribute to reducing these processing times even further.

The study successfully delineates the efficacy of deep learning algorithms, specifically YOLO, Faster R-CNN, and SSD, when compered by medical practitioners. The resulting outcomes were subjected to comparative analyses employing one-way ANOVA tests and Tukey Pairwise tests. In accordance with these findings, it is anticipated that the refinement of trainable artificial intelligence programs and their integration as auxiliary tools for healthcare professionals will substantially augment their diagnostic capabilities. Consequently, this study, illuminating subsequent

investigations in medical image detection, establishes that the attainment of higher success rates can be realized through the creation and expansion of original datasets.

1. GİRİŞ

Buhar gücüyle hesap yapabilen ilk otomatik hesap makinesinin 1823 yılında hayat bulması ile mekanik hesap makineleri bilgisayar olma yolunda insanlık adına ilk büyük adımı Charles Babbge tarafından atılmıştır. Bu adımın ardından geçen yaklaşık 200 yıl sonra bilinen insanlık tarihinin hiç sahip olmadığı bir birikim ile üreten, öğrenen ve özgün çıktılar sağlayan makineler insan hayatının artık vazgeçilemez bir konuma erişmiştir.

21. yüzyılın en yıkıcı etkisine sahip olacağı öngörülen ve bütün çalışmaların da bunu desteklediği yapay zekâ kavramı; makine öğrenmesi, derin öğrenme ve bilgisayarlı görü normal insan yaşantısının her alanında inanılamayacak boyutta etkin rol aldığı gözükmektedir (Torkul ve Tunacan, 2022). Günümüzde başta savunma sanayi, askeriye, güvenlik, medikal, otomotiv, havacılık, iletişim başta olmak üzere hem üretim süreçlerinde hem de elde edilen ürünlerin kullanımı sırasında yapay zekâ teknolojisi yer aldığı gibi, yapay zekâ destekli görüntü işleme ve bilgisayarlı görü kavramı önemi inkâr edilemeyecek boyuta erişmiş ve aktif şekilde kullanılmaya başlanmıştır. Buna en basit örneği dünya nüfusunun büyük çoğunluğunun kullanmakta olduğu akıllı telefonların tuş kilidi ve genel güvenliğinde yüz tanıma sistemlerinin kullanılmasıdır.

Karar verebilen ve düşünebilen makinelerin var olabileceği ve kullanılabileceği konusu 1950'li yılların hemen başında bilim insanlarının gündemine girmiştir. Gündeme gelmesinden birkaç yıl sonra 1956 yılında yapay zekâ kavramı ortaya atılmıştır. Bir insanoğlunun kendine özel zekasına sahip nitelikler bütünün yahut bir kısmının mekanik yahut fiziksel bir sistemde gözlenmesi hali yapay zekâ olarak nitelendirilmiştir. Bir nesnenin yapay zekaya sahip olup olmadığı günümüzde halen daha Turing testi ile ortaya konulmaktadır. Söz konusu test karşılıklı sorulara verilen yanıtlar ile bir hakemin iki farklı odadaki yanıtlandırıcıların insan mı yahut makine mi olduğunu anlayamaması halinde, karşısındaki makinenin yapay zekâya sahip olduğu kabulü üzerinedir (Turing, 1950; Russel ve Norvig, 2009).

Yapay zekanın alt başlığı olarak ifade edilen makine öğrenmesi kavramı, programlanmayan sonuçların üretilmesi yeteneği olarak nitelendirilmektedir. Klasik programlama mantığında kurallar ve veriler sisteme tanımlandırıldığında cevapların ortaya çıkarılması durumuyken, makine öğrenmesinde cevaplar ve verilerin sisteme tanımlanması halinde kural setleri ortaya çıkmaktadır. Bu kural setleri ile makine öğrenmesi yeni veriler ile yeni cevaplar üretebilir hale gelecektir (LeCun ve ark., 2017).

Derin öğrenme kavramı yapay zekanın derin uykusundan uyanışı ile ortaya çıkmış olup, özellik çıkarma işleminin insan tarafından yapıldığı bunun sonucunda sınıflandırma ve kural setlerinin ortaya çıktığı makine öğrenmesi kavramının bir alt dalıdır. Derin öğrenmede özellik çıkarma ve sınıflandırma hususlarının tamamını birçok katmanda işlem yaparak aynı anda sistemin kendisinin yapması sağlanmaktadır (LeCun ve ark., 2017).

Derin öğrenme algoritmalarının bilgisayarlı görü alanında kullanılması görüntü işlemede oldukça başarı sağladığı görülmektedir. 2013 yılına kadar başarı oranı %40 geçemeyen algoritmaların, derin öğrenmenin etkisi ile 2015 yılına vardığımızda %80 seviyelerine çıkarken, artık günümüzde insan ile hatta alanında uzman çalışmacılar ile kıyaslanabildiği bir seviyeye erişmiştir. Medikal alanda özellikle teşhis noktasında oldukça yol kat eden bu algoritmalar başarı noktasında bazı alanlarda başarılar elde ettiği hatta alanında uzmanları dahi geçmeye başladığı görülmüştür (Girshick ve ark., 2014; Girshick, 2015; Zeren ve ark., 2023).

1.1. Tezin Kapsamı

Yaklaşık iki buçuk yılı aşkın araştırmaların bir ürünü olarak söz konusu tezin kapsamı travma sonrası yaşlı popülasyonun ölüm ve sakat kalma olasılığının en yüksek olduğu femur üst bölge kırıkları odağa alınmıştır. Dünya genelinde yapılan araştırmalar değerlendirildiğinde yaşlı popülasyonunun karşılaştığı en kritik rahatsızlıklardan biri aynı zamanda mortalite oranın oldukça yüksek olduğu gözükmektedir. Ayrıca bütün ortopedik kemik kırıkları değerlendirildiğinde sakat kalma ve örüm oranı en yüksek olduğu da görülmektedir (Agar ve ark., 2017; Murat ve ark., 2020). Gelişen teknolojiler ile insan yaşam süresinin uzamasının bir sonucu olarak 2040 yılına kadar kalça kırıkları vakalarının 2 kat artması beklenmektedir (Zuckerman, 1996). 1990 yılında femur üst bölge kırığı (kalça kırığı) vakalarının 1,26 Milyon gerçekleşmesine

karşın, 2050 yılına vardığında 4,5 Milyon kayıtlı kalça kırığı vakası gözlemlenmesi yapılan araştırmalarca beklenmektedir (Veronese ve ark., 2018).

Femur üst bölge kırıkları Yapay Zekâ kavramının altında yer alan makine öğrenmesinin bir alt başlığı olan derin öğrenme ve bilgisayarlı görünün en güçlü ve başarı oranları en yüksek sayılabilecek üç algoritmasını olan YOLO – You Look Only Once, Faster – RCNN- Faster Region Based Convolutional Neural Network, SSD – Single Shot Multibox Detector algoritmaları odağa alınmıştır. Kapsama alınan söz konusu derin öğrenme algoritmalarının başarı seviyesini doğru anlayabilmek kıyaslamak ve performansını objektif bir şekilde ortaya koyabilmek adına, altı yıllık tıp eğitimini tamamlamış pratisyen hekim grupları ve altı yıllık tıp eğitimini tamamlaması sonrası üzerine ortopedi ve travmatoloji alanında uzmanlık yapma hakkı kazanmış 1 yıldan daha fazla tecrübeli asistan doktor ve uzman doktorlardan oluşan bir hekim grubu daha oluşturulmuştur.

Veri toplama sürecinde ülkemizin en yüksek popülasyonuna sahip ve Avrupa kıtasının da en yüksek nüfusa sahip metropollerinden biri olan İstanbul ili odağa alınmıştır. Söz konusu İstanbul ilinin en büyük hastanelerinden biri olan İstanbul Şişli Hamidiye Etfal Eğitim ve Araştırma Hastanesi Ortopedi ve Travmatoloji Anabilim Dalına travma sonrası kalça kırığı şüphesi veyahut kalça kırığı teşhisi ile müracaat eden hastalar odağa alınmıştır.

1.2. Tezin Amacı ve Önemi

Yapılan bütün araştırmalar bütün ortopedik kırık vakalarının kalıcı sakatlık ve ölüm oranın en yüksek olanının femur üst bölge yani kalça kırığı olduğu göstermektedir. Gelecekte de bu durumun değişmemesi ve rastlanan vaka sayısının önemli ölçüde artması beklenmektedir. Yine aynı araştırmaların sonuçları incelendiğinde yaşlı popülasyonun karşılaştığı sakat kalma ve ölüm riskinin en yüksek olduğu sağlık problemlerinin başında gelmektedir (Agar ve ark., 2017). 2050 yılına gelindiğinde 4,5 milyondan daha fazla femur üst bölge kırığı vakası beklendiği dünyamızda femur üst bölge kırıklarının doğru teşhis edilebilmesi bütün insanlığın ortak problemdir.

Yapay zekanın derin uykusundan uyanmasından sonra CPU, GPU ve hatta TPU ların gelişmesi ile hem nesne tespitinde oldukça başarılı hale gelmiştir (Karunakaran, 2019). Günümüzde çeşitli veri setleri ve çeşitli sektörlerdeki uzmanların başarısını geride bırakan derin öğrenme algoritmaları her geçen gün işlemcilerin gelişmesi ile daha

hızlı, mimarilerinin güncellenmesi ile daha güçlü bir şekilde nesne tespiti yapar hale gelmişledir.

Söz konusu çalışma, tezin kapsamı kısmında ifade edilen başarılarını kanıtlamış üç derin öğrenme algoritmasının kırık ve sağlam femur üst bölge X-ray görüntülerinin toplanarak yepyeni eğitim ve doğrulama veri setleri oluşturmak, akabinde transfer öğrenme yöntem ile yeniden eğitilmesi sağlayarak, kalça kırıklarının tespitinde özerkleşmiş derin öğrenme algoritmaları elde etmektir. Söz konusu alanda eğitim ve doğrulama veri setlerinden tamamen bağımsız olan test veri seti ile hem algoritmaların toplam başarılarını hem de birbiri içerisindeki başarılarını ortaya koymak en önemli amaçlarından biri olup, bunun yanında hem ortopedi asistanı ve uzmanları hem de tıp eğitimi almış pratisyen hekimlerle de kıyaslayarak tamamen objektif bir başarıyı göstermek amaçlanmıştır. Çalışmanın bir diğer sonucu olarak femur üst bölge yani kalça kırıklarının tespiti için hem medikal alanda çalışan radyologlara, acil hekimlerine ve ortopedi doktorlarına kalça kırıklarının teşhisinde destek olacak bir sistem sunulmuştur. Söz konusu çalışmanın ilave bir sonucu performansı nesnel bir şekilde ortaya koyacak şekilde bağımsız oluşturulmuş test veri setinin her bir çalışma grubu için kalça kırığı olan bölgedeki başarısı, kalça kırığı olmayan bölgedeki başarı ve sonuç üretme süresi varyans analizleri ile kıyaslanarak anlamlı bir fark olup olmadığı da ortaya konulması amaçlanmıştır.

1.3. Tezin Yöntemi ve Özgün Katkısı

Çalışmada yapay zekâ ve makine öğrenmesinin altında yer alan derin öğrenme kavramının bilgisayarlı görü ile bir araya gelmesi ile sağlanan nesne/kritik bölge tespiti işlemin en güçlü algoritmaları ele alınmıştır. Pyhton yazılım dilinin kullanılması ile YOLO Darknet v4 algoritması, Faster R-CNN Inception v2 Algoritması ve SSD-Mobilenet v2 algoritması odağa alınmıştır. Odağa alınan algoritmalar daha öncesinde MS COCO (Microsoft Common Objects in Context) veri seti ile eğitilmiş ve öğrenim transferi (transfer learning) yöntemi ile femur üst bölgesi X-ray görüntülerinden oluşturulan özgün veri seti ile yeniden 5.000 iterasyon boyunca eğitilerek femur üst bölge kırıklarını yani kalça kırıklarını tanıyabilir hale getirilmesi amaçlanmıştır. Söz konu eğitimin tamamlanması için kullanılan eğitim ve doğrulama veri seti oluşturulması sürecinde 410 hastadan alınan X-ray görüntüleri ile veri çoklandırma tekniklerinden döndürme, ölçeklendirme kırpma ve yeniden boyutlandırma ile 820'ye

çıkarılmıştır. Toplamda algoritmaların eğitiminde kullanılmak üzere 1514 adet femur üst bölge görüntüsü elde edilmiştir. Söz konusu eğitimin gerçekleştirilmesi için 820 X-ray görüntüsü %80 Eğitim ve %20 doğrulama veri seti olarak ayrıştırılmıştır.

Odağa alınan derin öğrenme ve bilgisayarlı görü algoritmalarının performansını objektif bir şekilde önyargı sapma ve yanılmaya düşmeden kıyaslamak üzere eğitim ve doğrulama veri setinden bağımsız 50 kırık femur görüntüsü 50 kırık olmayan femur bölgesi olmak üzere toplamda 55 hastadan alınarak test veri seti oluşturulmuştur. Odağa alınan derin öğrenme ve bilgisayarlı görü algoritmalarının gerçek başarısını göstermek adına 10 pratisyen hekimden oluşan pratisyen hekim grubu, 12 ortopedi ve travmatoloji asistanı ve uzmanlarından oluşan uzman hekim gurubu oluşturulmuştur. Test veri seti kendilerine önyargı yanılma ve sapmaya minimum düşmeleri için rastlantısal bir sıralama ile gösterilmiş verdikleri cevaplar kronometraj yöntemi ile kayıt altına alınmıştır.

Söz konusu elde edilen sonuçlar istatistiksel analizinin de gerçekleştirilmesi için MiniTab ve SPSS programları kullanılarak hem sonuç elde edilme süreleri, kırık olan görüntülerdeki doğru tespit oranı ve kırık olmayan görüntülerdeki doğru tespit oranları hem Tek Yönlü ANOVA testi hem de Tukey Kıyaslama testlerine tabi tutulmuştur.

Derin öğrenme ve bilgisayarlı görünün en güçlü kabul edilen YOLO algoritması, Faster R-CNN Algoritması ve SSD algoritması öğrenim transferi yönetimi kullanılarak yeniden eğitilmiş bunun sonucu olarak 3 farklı algoritma için başarı oranları performans değerlendirme tablolarına göre ölçümlenmiştir. Bunun sonucunda performansları farklılık gösteren ve doktorlara destek olabilecek 3 farklı sistem geliştirilmiştir. Femur üst bölge yani kalça kırıklarının tespitinde ilk defa hem ortopedi ve travmatoloji asistanı ve uzmanları, pratisyen hekimleri ve 3 farklı derin öğrenme algoritması bir araya getirilerek sonuçlar elde edilmiş, başarıları kıyaslanmış ve derin öğrenmenin kalça kırıklarını tespitindeki başarısı ortaya konulmuştur. Ayrıca söz konusu başarılar birbiri ile istatistiksel metodolojiler ile kıyaslanmış anlamlı fark olup olmayışı ortaya konulmuştur. Benzer medikal alanda görüntü işleme çalışmalarına kıyasla ilk defa bu sayıda hastanın özgün X-ray görüntülerinin işlendiği bir çalışma gerçekleştirilmiştir.

1.4. Tezin Çalışma Planı

Tez çalışma planında öncelikle olarak "Yapay Zekâ" kavramı araştırılmıştır. Sonrasında yapay zekâ kavramının bir alt araştırma alanı içerisinde yer alan "Makine Öğrenmesi" kavramı incelenmiştir. Sonrasında makine öğrenmesi kavramının altında yer alan gösterim öğrenme ve onun alt alanında yer alan derin öğrenme kavramı incelenmiştir. Derin Öğrenme ve Bilgisayarlı görü algoritmalarının incelenmesinin ardından uygulama için uygun olan algoritmaların belirlenmesi gerçekleştirilmiştir. Travma sonrası meydana gelen kemik kırıklarından hangi alanın odağa alınacağı ile ilgili araştırmalar yapılmış, alanında uzman hekimler, doçent doktorlar ve profesör doktorlar ile istişarelerde bulunulmuştur. Odağa alınacak kemik kırığı olarak ortopedi alanında en fazla ölüm ve sakat kalma oranı bulunan kalça kırıkları yani femur üst bölge kırıkları olarak belirlenmiştir. Sonrasında veri seti toplama aşamasına geçilmiştir. Veri toplama aşamasında femur üst bölgelerini içeren X-ray görüntüleri derlenmeye başlanmıştır. Algoritmaların yeniden eğitilmesi amacıyla eğitim, doğrulama ve test veri setleri oluşturulmuştur. Sonrasında son teknolojik gelişmeler de takip edilerek derin öğrenme ve bilgisayarlı görü algoritmaları seçilmiştir. Bunun ışında teknik yeterlilikleri karşılayacak teknoloji araştırması gerçekleştirilmiş konuyla ilgili hem Bulut (Cloud) servisleri araştırılmış hem de iş istasyonları belirlenmiştir. Sonrasında söz konusu algoritmaların eğitimlerinin gerçekleştirilmesi için yazılım dili ve Idle araştırması yapılarak seçimler gerçekleştirilmiştir.

Hem görüntü işleme ve derin öğrenme kütüphanelerine erişimi kolaylığı hem de son dönemde özellikle birçok derin öğrenme yazılımı geliştirme gerçekleştirilirken tercih edilmesi sebebiyle seçilen Python yazılım dili seçilmiştir. Sonrasında yazılım altyapısı hazırlanmış kod yazımı gerçekleştirilmiştir.

Sonrasında hem iş istasyonları hem de bulut tabanlı yazılım altyapıları kullanılarak ilgili algoritmalara hazırlanan veri setleri yardımı ile eğitimleri tamamlanmıştır. Femur üst bölge kırığının tespiti alanında yeniden eğitilmiş algoritmaların, eğitim ve doğrulama veri setlerinden tamamen bağımsız oluşturulan test veri seti ile karmaşıklık matrisi ve diğer belirlenen kriterler ışığında performansının ölçümü gerçekleştirilmiştir.

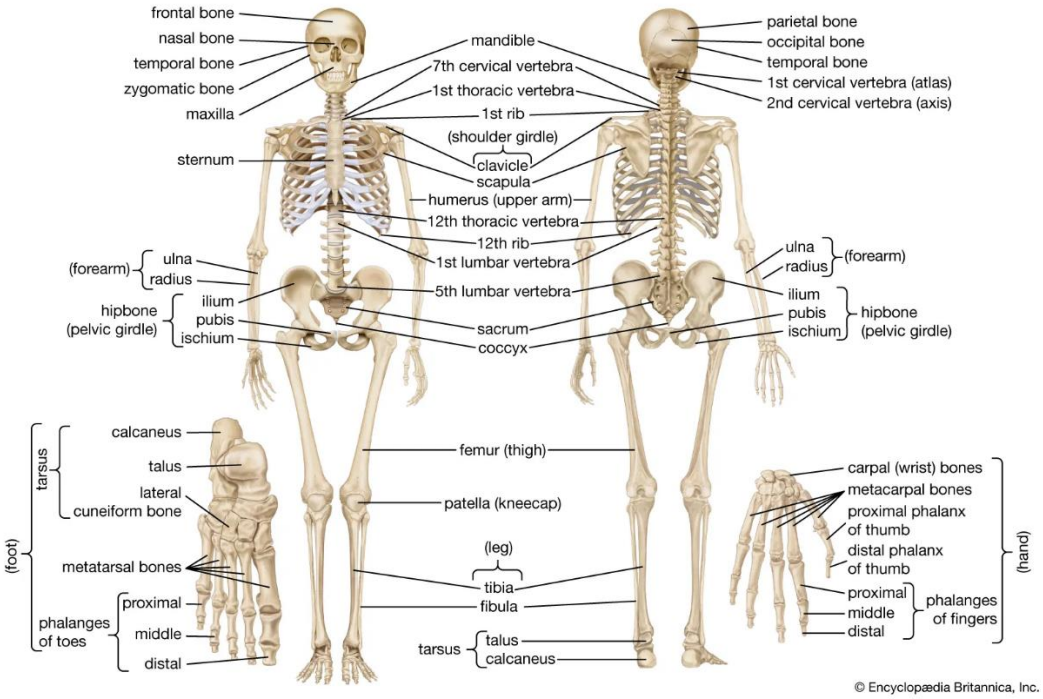
Bu arada söz konusu algoritmaların performansını daha doğru ortaya koymak adına pratisyen hekimlerden ve ortopedi ve travmatoloji alanı asistanları ve uzman doktorlarından oluşan iki grup belirlenmiş akabinde gönüllüler içerisinden grup

oluşturulmuştur. Bu iki grup yine test veri seti görüntüleri karışık olarak gösterilmiş ve performansı yine karmaşıklık matrisi ve diğer performans değerlendirme kriterleri kapsamında ölçülmüştür. Söz konusu çıktıların istatistiksel olarak analiz edilebilmesi için programlar araştırılmıştır. İstatistiksel inceleme için MiniTab ve SPSS programları odağa alınmış, istatistiksel yöntem olarak hem Tek Yönlü ANOVA testleri hem de Tukey karşılaştırma testleri kullanılmıştır. Çıktılar her aşamada kayıt altına alınmıştır.

1.5. İnsanda Kas İskelet Sistemi

İnsanda kas ve iskelet sistemi ele alındığında vücut organlarını dışarıdan gelecek mekanik etkilere karşı muhafaza eden, destek sağlayan ve haraketliliği sağlayan yapılardan meydana gelmektedir (Hall ve Hall, 2020).

Kemik dokusu vücuda ana desteği sağlayan yapıdır. Kas dokusu da ana haraketliliği sağlayan yapıdır. (Hall ve Hall, 2020). Erişkin bir insan vücudunda toplamda 206 kemik yer almaktadır. İnsan iskeleti vücudu çerçeve olarak saran çok sayıda kemik ve kıkırdaktan oluşan yapıdır. Bu iskelet diğer bütün omurgaya sahip canlılarda olduğu şekilde iki ana bölümden oluşmaktadır. Bunlardan ilki aksiyal omurga kemiği omurga ve kafatasıdır. İkincisi ise apendikülerdir (pelvik- kalça ve pektoral - omuz kuşakları ile eklem ve kıkırdaklar dahil olmak üzere eklem ve kıkırdakların olduğu kısımdır) (Andrew, 2023).



Şekil 1.1. İnsan iskelet sistemi (Andrew, 2023).

Söz konusu iskelet sisteminde yer alan bu kemikler dokuların yardımı olmadan hareketi sağlamaları mümkün değildir. Bir ya da biden fazla kemiğe bağlanmış kas dokuları kasılma ve gevşeme hareketleri ile vücut hareketini oluşturmaktadır (Sweeney ve Hammers, 2018).

Kemikler ele alındığında beş ana görevlerinin olduğunu söylemek mümkündür. Bunlardan ilki iskelet sistemini bir bütün olacak şekilde doku ve kaslara temel oluşturarak destek sağlamaktır. İkincisi dışarıdan gelecek bütün etkilere karşı iç organları korumaktadır. Üçüncüsü kemiklerde hareketi sağlayacak kas dokularına bağlanması için tendonlar yer almaktadır. Bu sayede hareket meydana gelir. Dördüncüsü vücuttaki hücresel hareketliliğin devamlılığı için Kalsiyum, Fosfor ve Magnezyum minerallerini depolamaktadır. Beşinci olarak da beyaz ve kırmızı kan hücrelerinin çoğunluğunun oluşumunu sağlamaktadır (Tuna, 2013).

Kemikler, uzun kemikler, kısa kemikler, yassı kemikler ve düzensiz kemikler olarak sınıflandırmak mümkündür. Uzun kemikler boyları enlerine göre daha uzun olan kemikler olarak izah edilebilir. Kol kemiği ya da femur kemiği uzun kemiklere örnek olarak verilebilir. Kısa kemikler enleri ile boylarının birbirlerine yakın olan kemikler olarak nitelendirilebilir. Karpal kemiği kısa kemiklere örnek olarak verilebilir. Görece ince olup kavisli bir eğriliğe sahip kemikler yassı kemiklerdir. Yassı kemiklerine örnek

olarak kafatası kemikleri verilebilmektedir. Düzensiz kemikler ise belirli bir formda nitelendirilmeyen karmaşık yapılı kemiklere denmektedir. Omurga kemikleri düzensiz kemiklere örnek olarak verilebilir (Yenigün, 2016).

Kimyasal yapıları bakımından kemikler ele alındığında %65'inin bileşenleri inorganikken kalan yaklaşık %35'lik kısmının bileşenlerinn organik olduğu ifade edilebilir (Tuna, 2013). Organik bileşenler içerisinde osteoid, osteojenik, osteosit, osteoblastlar ve osteoklast hücreleri yer almaktadır. Organik bileşenler içerisinde yer alan hücrelerce üretilen kolajen lifler kemik esnekliğinde önemli rol oynamaktadır. Bu liflerin yer almadığı kemik aşırı kırılganlaşmaktadır. İnorganik bileşenler içerisinde Kalsiyum Karbonat, Kalsiyum Fosfat, Magnezyum, Potasyum ve Florür minerallerinden oluşmaktadır. Minareller kolajen liflerin çeviresinde küçük kristaller formunda yer almakta bu form esasen kemiğin sıkılaştırma basıncına karşını direncini arttırmakta ve sertliğini sağlanmaktadır. Kemik yapısının inorganik ve organik bileşenlerinin ahenkli kombinasyonu ile kemikler daha basla hem dayanıklı hem de esnek olarak daha fazla travmalara karşı dayanıklı hale gelmektedir (Tuna, 2013; Yenigün, 2016).

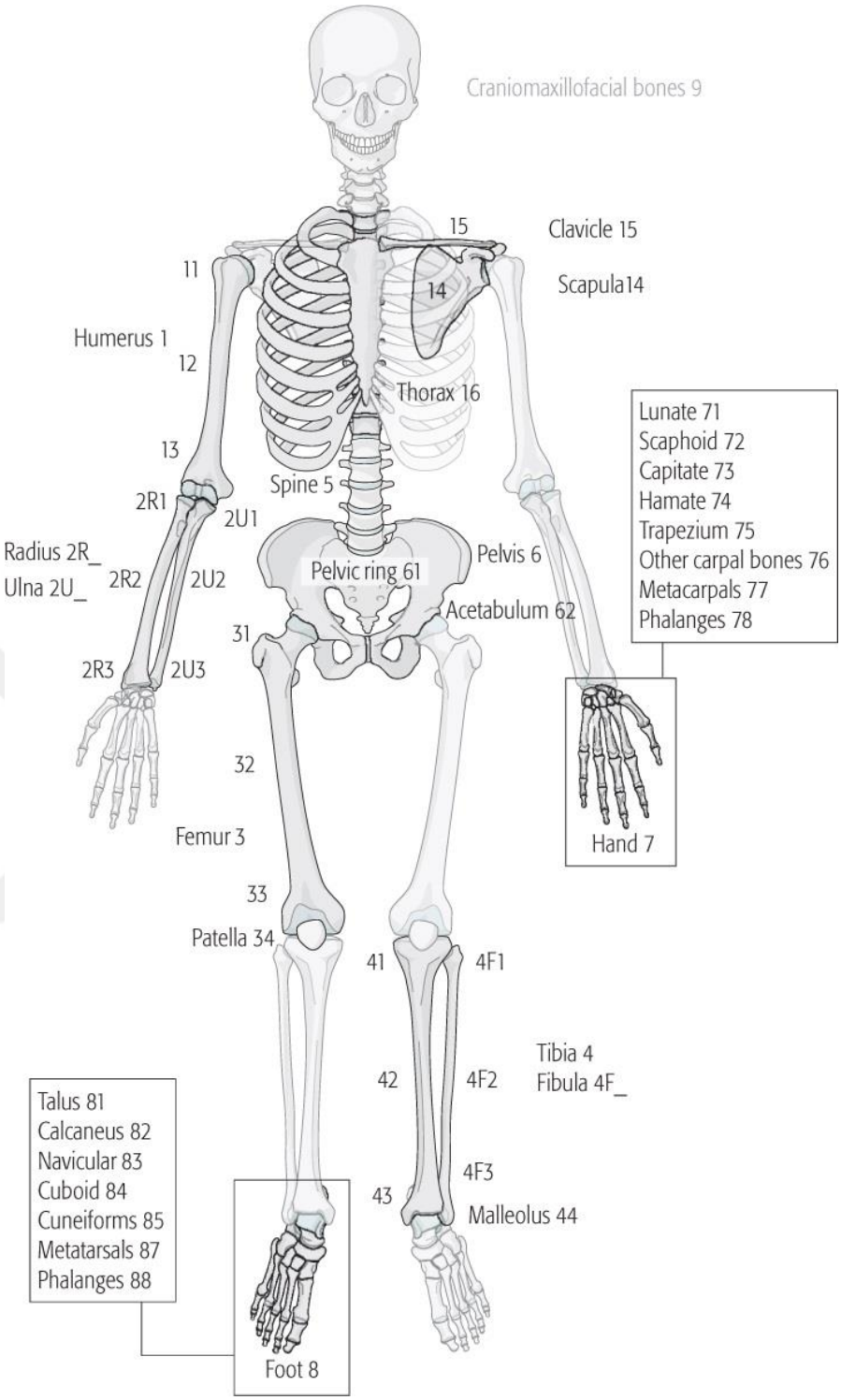
Sağlık bakanlığının yapmış olduğu araştırmalar ışığında temel hastalık grupları içerisinde üçüncü sırada kas iskelet sistemi hastalıklarının yer aldığı görülmektedir (Sağlık Bakanlığı, 2020).

1.6. İnsanda Kemik Kırıkları

Kas iskelet sisteminin haraketliliği ve statiğini sağlayan en önemli yapılar kemiklerdir. Kemikler dışarıdan gelecek kuvvetlerin yahut travmaların etkisi ile yapısal bütünlüğü bozulabilir. Kemik kırıklarının boyutu meydana gelen travmanın boyutuna göre oldukça değişkenlik göstermektedir. Basit bir çatlağın yanı sıra çoklu kırıkların olabileceği hatta kırığın vücut dışına çıkarak açık kırığa dönüşebileceği şeklinde gerçekleşebilmektedir (Ekşi ve Çakıroğlu, 2012; Upadhyay ve Tanwar, 2019).

Şekil 1.2. Travma sonrası çok parçalı kırık meydana gelen femur üst bölge görüntüsü.

Kırıkların farklı sınıflandırma sistemleri mevcuttur bunlar arasında en ön plana çıkan Orthopaedic Trauma Association (OTA) ve Arbeitsgemenienschaft für Osteosynthesefragen (AO)'dir. (Meinberg ve ark., 2018).



Şekil 1.3. Kemik konumunun belirlenmesi (Meinberg ve ark., 2018).

Diyafiz kırığı sınıflandırma sürecinde Tablo 1.1'deki sorulara cevap verilmek sureti ile sınıflandırma işlemi gerçekleştirilir.

Tablo 1.1. Diyafiz kırığını sınıflandırma ve kodlama süreci (Meinberg ve ark., 2018).

Adım	Soru	Cevap
1	Hangi kemik?	Belirli kemik (Şekil 1.3)
2	Kırık kemiğin neresinde?	Orta - diyafiz segmenti
3	Kırık hangi tip	Basit (A), Kama (B), Çok parçalı (C)
4a	Eğer basit ise kırık hangi gruptadır?	Sarmal (1), Eğik (2), Enine (3)
4b	Eğer kama ise kırık hangi gruptadır?	Bozulmamış (2) yada Parçalı (3)
4c	Eğer çok parçalı ise kırık hangi gruptadır?	Bozulmamış segmental (2) yada Parçalı segmental (3)
5	Nitelikler ve/veya evrensel değiştiriciler ekleyin	

Örnek olarak AO/OTA sınıflandırma tekniği 12-C2 şeklinde ifade edilen bir kırık Humerus (1) kemiği, Diyafiz bölgesi (2) , Çok parçalı (C) ve Bozulmamış segmental (2) olarak anlaşılmaktadır.

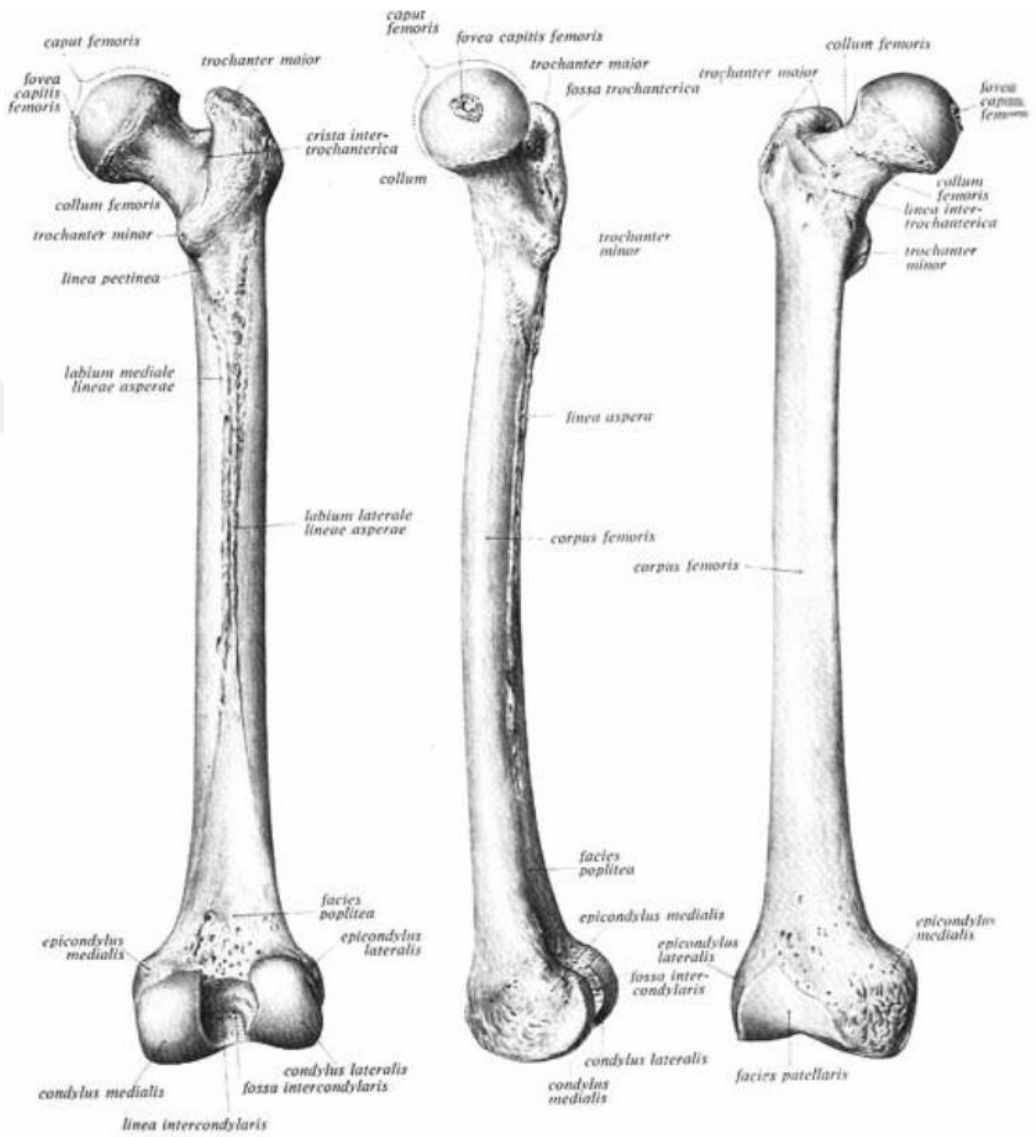
Kuvvet türlerine göre ortaya çıkması beklenen kırık türlerini ifade etmek istersek travma sırasında eğer kemiğe gerilme kuvveti etki etmiş ise kırık türünün enine kırık olması beklenmekte; eğer sıkıştırma kuvveti etki etmiş ise eğik kırık türünün ortaya çıkması beklenmekte; burulma kuvveti kemiğe etki etmiş ise sarmal kırık türünün ortaya çıkması beklenmekte, bükme kuvveti etki etmiş ise enine ve kelebek kırık türü beklenmekte; eğer aynı anda hem sıkıştırma ve bükme kuvvetleri etki etmiş ise enine ve eğik kırık türünün ortaya çıkması beklenmektedir (Tripathi ve ark., 2017).

1.7. İnsanda Femur Üst Bölge

Yapılan araştırmalar yaşlı hastalarda femur üst bölge kırıkları sonrasında ortaya çıkan en önemli komplikasyonun ölüm olduğunu göstermektedir. Kalça kırıkları sonrasında hastanede tedavileri sırasında %15, tedavi sonrasındaki bir yıl içerisinde ise ölüm oranının %30'lara çıktığı görülmektedir (Kreating, 2015). Söz konusu ölüm oranları aynı yaş gurubunda bulunan ve söz konusu kalça kırığına sahip olmayan insanlara göre dört ile beş kat fazla olduğu görülmektedir.

İlk femur boyun kırığının tariflemesi 1822 yılında Sir Astley Cooper tarafından gerçekleştirilmiştir. Söz konusu femur boyun kırığının tedavisini sağlayan ilk cerrah Von Lagenbeck'tir. Hasta bir süre sonra enfeksiyon nedeniyle vefat etmiştir (Freeman, 1904).

İnsan vücudundaki en sağlam ve en uzun olarak nitelendirilen kemik femur kemiğidir. Femur baş kısmı hiyalin adı verilen bir kıkırdak ile sarılıdır. Söz konusu kıkırdak perifere doğru incelir. Femur şaftını baş kısmına bağlayan femur boynudur (Hughes ve ark., 2002).



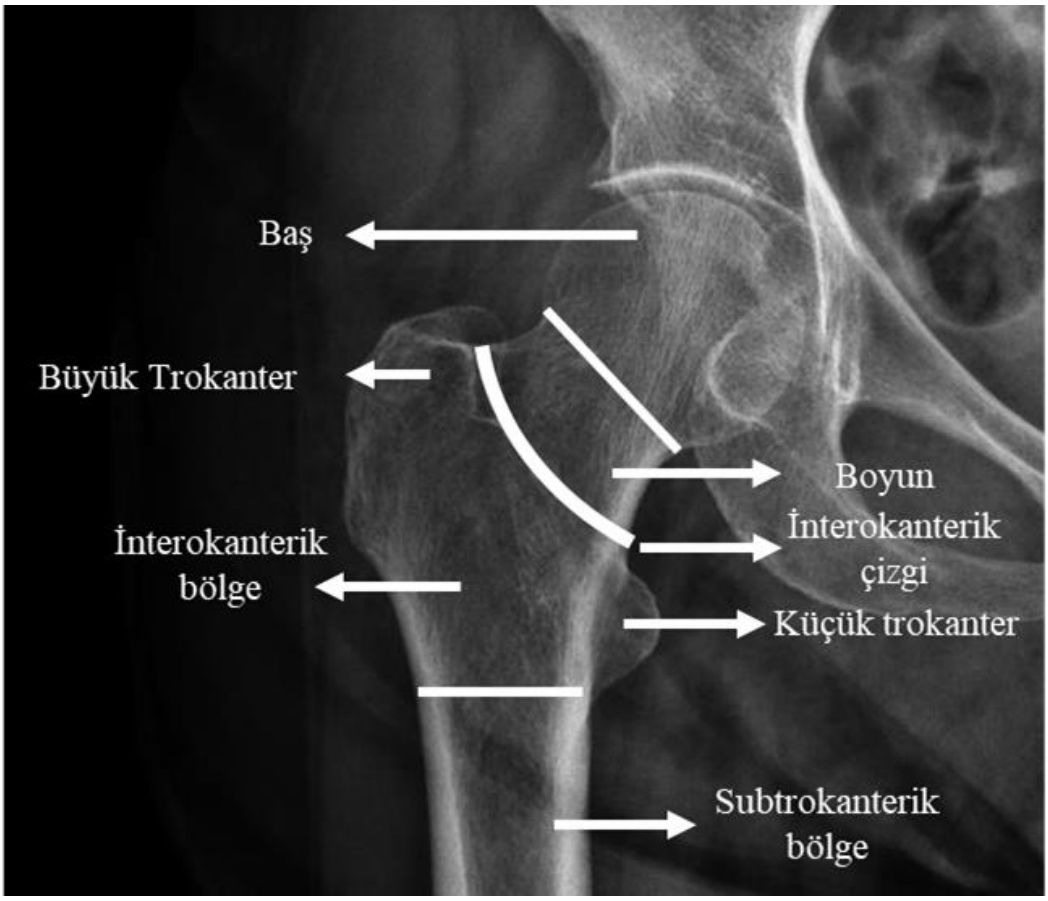
Şekil 1.4. Femur kemiğinin anatomik yapısı (Waschke ve ark., 2011).

Femur üst uç kısmında ağırlığı eşit dağıtan trabeküler sistemler yer almaktadır. Söz konusu trabeküler yapısı doğrudan kıkırdaklar ile bağlantılıdır. 1838 yılında Ward tarafından trabeküler sistemi tanımlanmıştır (Singh ve ark., 1970).

Bütün kırıklarda ilk olarak yaralanma hikayesinin araştırılması elzemdir. Sistematik bir muayene ile basit düşme ve travma hikayesinde eşlik edebilecek yaralanmaları bulmak ve tespit etmek elzemdir (Lowell, 1966; Vahvanen, 1971). Özellikle yüksek enerjili bir travma söz konusu ise genç yaştaki hastalarda kalça çevresi kırıkları

gözlemlenebilir. Bu nedenle öncelikli olarak hayati bölgelerde (kafatası, iç organlar ve boyun) yaralanmalarının tespitinin yapılması elzemdir. (Lowell, 1966; Vahvanen, 1971). Femur uç bölgede stres kırığına rastlanmışsa bu hastalarda genelde travma geçmişi yerine fiziksel aktivite değişikliğine yönelmiş olması olası olduğundan bu husus araştırılmalıdır (Lowell, 1966; Vahvanen, 1971).

Femur üst bölge kırıklarının sınıflandırmasında güncel olarak Garden ve AO sınıflandırmaları kullanılmasına karşın halen daha 1935 yılında adını literatüre kaydettirmiş Pauwels sınıflandırması kullanılmaktadır. Yapılan araştırmalar halen daha Pauwels sınıflandırmasının diğer sınıflandırmalara göre daha güvenilir olduğunu ifade etmektedir (DeLee, 1984; Turgut ve ark., 2016).



Şekil 1.5. Femur üst bölge anatomik yapısı.

Bütün proksimal femur kırıkları yani femur üst bölge kırıkları incelendiğinde söz konusu kırıkların %90 dan fazlası femur boyun kırıkları oluşturmakta geriye kalan kısmı subtrokanterik kırıklar oluşturmaktadır (Filipov, 2014; Gallagher ve ark., 1980; Karagas ve ark., 1996). Femur proximal yani femur üst bölge kırıklarının %98'i 50 yaşın üstündeki hastalarda gözlemlenmektedir (Berglund-Rödén ve ark., 1994).

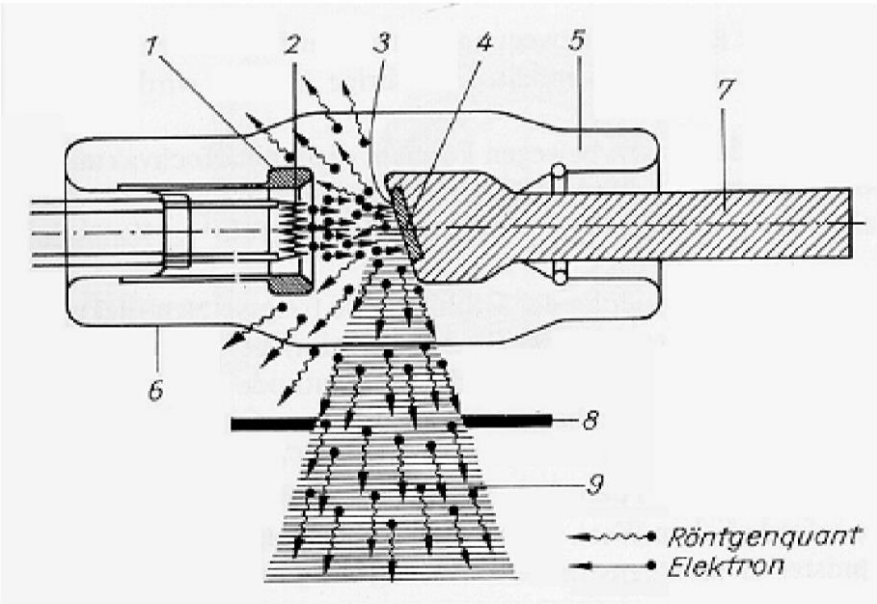
1.8. İnsanda Kas İskelet Sistemi Görüntüleme Yöntemleri

Tıp alanında kullanılan teknolojilere biyomedikal denmektedir. Biyomedikal sağlık alanında tedavi ve teşhiste kullanılan mekanik, elektronik ve megatronik sistemlerin tamamına denmektedir. Bu amaçla doktorlara kılavuz olabilecek onların işini kolaylaştıracak yeri geldiğinde rehber olacak yeri geldiğinde doğrudan veya dolaylı olarak ameliyat sırasında dahi kullanılabilecek, içerisinde birçok mühendislik dalının da geliştirilmesinde katkısı bulunan cihazlara biyomedikal cihazlar denmektedir. Günümüzde gelişen teknoloji ile birlikte bambaşka bir şehirdeki cerrahlar uzaktan bağlantı aracılığı ile Da Vinci Robotik Cerrahi Sistemini kullanarak ameliyat gerçekleştirebilmektedir.

Medikal alanda görüntüleme teknikleri herhangi bir cerrahi müdahale gerektirmeksizin ilgili alanda müdahale edilmeden önceden görülmesini sağlamaktadır. Tıbbi görüntüleme tekniklerine örnek olarak MRI olarak kısaltılan Manyetik Rezonans ile görüntüleme, CT olarak kısaltılan bilgisayarlı tomografi görüntüleme, ultrasonografi ve X-ray yahut X-ışını radyografi verilebilir.

1.8.1. X-ışını görüntüleme (radyografi)

Günümüzde dünya çapında en yaygın olarak kullanılan görüntüleme sistemi X-ray/X-ışınlı görüntüleme sistemidir. Medikal alanda görüntüleme sisteminin temeli olarak nitelendirilebilecek X-ray ışınları Wilhem Conrad Roentgen adındaki alman fizikçi tarafından 1895 yılında geliştirilmiştir. Roentgen bir tüpün içinden elektrik akımı geçirildiğinde ışınlar çıkardığı bu ışınların arkasından gölgeler oluşturabildiğini fark ederek devrim niteliğinde bir buluşu ortaya çıkarmıştır. Çıkan bu ışınların insan vücudu içerisinden geçerken yine aynı şekilde gölgeler oluşturduğunu keşfeden Roentgen ilk kez ilk ışınlarını eşi Anna Bertha Roentgen'in elinin X-ray görüntüsünü çekerek bir insan bedeninde kullanılmasını sağlamıştır. Bu durum insan bedeninde kullanımına öncülük etmiştir (Ünal, 2008). Roentgen'in bu keşfinin üzerinden altı hafta geçmesinin ardından farklı malzemelerin de X-ışınını soğurma özelliklerini içeren makalesi olan "Yeni bir ışın türü üzerine" ("Uber eine neue Art von Strahlen") başlığı ile yayınlamıştır. Bu keşfinin ardından geçen altı yıl sonra tarih 1901 yılına gelindiğinde ilk Nobel Fizik Ödülü'nü almaya hak kazanmıştır.



Şekil 1.6. X-Işını Oluşumu (Ünal, 2008).

X-Işınları daha önce üretilmiş olan sisteme verilen elektrik enerjinin kinetik enerjiye dönüştürülmesi ile elde edilmektedir. Öncelikle alınan alternatif akım transformatörler aracılığı ile yükseltilir. Sonrasında rektifiye edilir. Doğru akıma dönen yüksek voltajlı elektrik cam tüp içerisinde bir katot ile anot arasına uygulanırsa hem büyük ısı enerjisi hem de az miktarda X-ışınına dönüşmüş olur.

X-ray Röntgen görüntüleme tekniğinin en büyük avantajları (Kasban ve ark., 2015):

- Hızlı ve ağrısız uygulaması kolay bir süreç ile sonuçların hızlıca elde edilmesi,
- Özellikle travmatolojide tıbbı ve cerrahi müdahalelerde müdahale öncesi doğru bir kılavuz sunması,
- Diğer görüntüleme tekniklerine göre yaygın kullanımı ve kurulumu ile daha ucuz olması.

X-ray Röntgen görüntüleme tekniğinin dezavantajları;

- Sürekli radyoaktif seviyesi yüksek X-ray ışınları maruziyetinin kanser riskini yükseltmesi,
- Yüksek radyasyon ile ortaya çıkabilecek yan etkilerin nadiren de olsa görülmesi,
- Hamile ve hamilelik riski bulunan kişilerin maruziyetleri halinde embriyoda ortaya çıkabilecek olası yan etkiler.

1.8.2. Manyetik rezonans (MRI)

Diğer bir görüntüleme yöntemi olan Manyetik Rezonans görüntüleme tekniği (MRI) manyetik titreşim manasına gelmektedir. Yine tanı amacını gerçekleştirmek üzere kullanılan bu cihaz protonların manyetik alana maruz kaldıklarında gerçekleştirmiş olduğu titreşimden yola çıkılarak geliştirilmiştir. Söz konusu MR ilk prototip cihazlarının geliştirilmesi 1981 yılına dayanmaktadır. Geliştirilmesinden sonraki üç yıllık periyot sonrasında 1984 yılında gerçekten medikal alanda hizmete sunulabilmiştir. Sonrasında kullanımı ve üretimi oldukça yaygınlaşan MR görüntüleme sisteminin temeli ilk olarak 1946 yılında Nükleer Manyetik Rezonans olarak tanımlanmış ilk deneyler Purcell ve Bloch tarafından ihtiva edilmiştir. Purcell ve Bloch tarafından bulunan bu çalışma da 1952 yılında Nobel fizik ödülüne laik görülmüştür. 1973 senesinde Lauterbur ve Mansfield NMR tekniğini fiziksel yapıların incelenmesinde ve analiz edilmesinde kullanmışlardır.

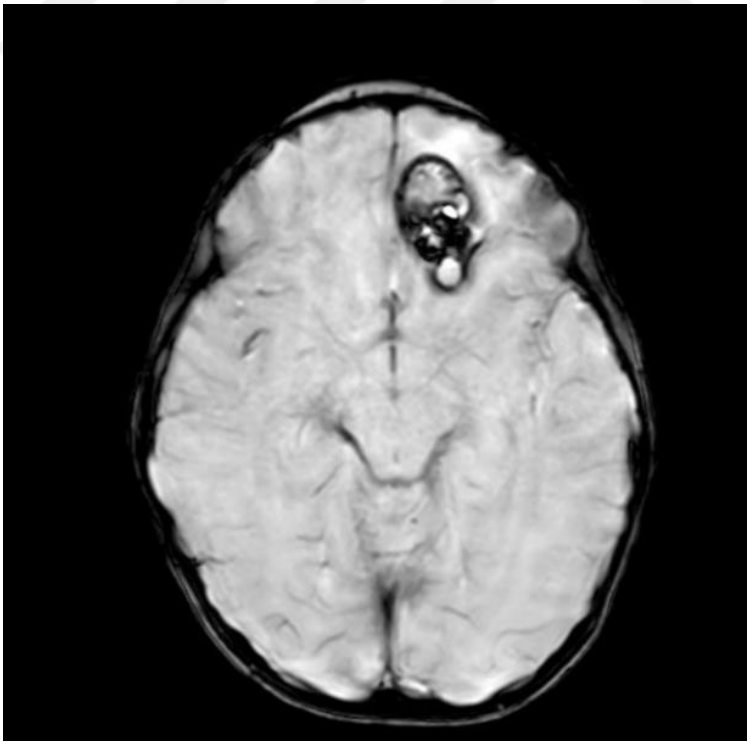
Günümüzde kullanım alanı ve imkanları oldukça genişlemiş olan MR cihazları 3 kısımdan oluştuğunu söylemek mümkündür. Bunlar, görüntü işlem ve kullanan operatörün bilgisayarları, kabinetler ve magnettidr. Magnet kısmının esas amacı görüntü alınacak alanla ilgili stabil ve manyetik açıdan steril bir alan yaratmaktır. İkinci kısmı olan kabinetler bilgisayar ve magnetler arasında veri aktarımı için bir bağlantı oluşturmaktadır. Bu alanda özellikle Helyum pompası kontrol kartları ve güç sağlayıcılar ve kontrol elementleri bulunur. Üçüncü kısmı olan görüntü işlem ve kullanan operatör bilgisayarları yer almaktadır. MR cihazının görüntülerini oluşturan bir bilgisayar ve görüntü işlem bilgisayarı yer almaktadır. Bu bilgisayarlar Manyetik alandan gelen görüntüleri proses ederler. Söz konusu bilgisayarlarda aynı zamanda görüntü ayarlamaları gerçekleştirilebilir (Ünal, 2008).

MR cihazının çalışma prensibinde H+ iyonları yani protonları normal ortamda titreşim hareketi gerçekleştirmektedirler. Bu protonlar manyetik alana giriş yaptıklarında doğal olarak oluşan manyetik alanın konumuna göre (pozitif atomlar kuzey yani N kutbuna göre) sıralanma eğilimi göstermektedir. Titreşimlerini bu yön doğrultusunda şekillendirirler. Söz konusu pozitif iyon olan protonların üzerine yüksek frekanslı olan radyo frekans (RF) dalgaları gönderilir. Bu durumda H+ iyonları söz konusu RF dalgasının bazılarını absorbe eder bazılarını ise yaymaktadırlar. Bu durum protonların dizilim, dağılım ve yoğunluklarına göre farklılık gösterecektir. Bu sayede özellikle

vücudunun büyük bir kısmı H2O atomlarından oluşan insan vücudunun rahatlıkla görüntülenmesinde kullanılır (Ünal, 2008).

Şekil 1.7. MR görüntüleme cihazı (Ünal, 2008).

İnsan vücudunda özellikle yumuşak dokuların yoğun olarak H+ iyonunu barındırması sayesinde özellikle yumuşak dokularda MR görüntüleme sisteminin başarısı daha büyüktür.



Şekil 1.8. MR görüntülemesi ile tanısı konulan beyin kanaması görüntüsü.

MRI cihazı ile gerçekleştirilen görüntülemenin avantajları olduğu gibi dezavantajları da söz konusudur. Avantaj ve dezavantajlar derlenerek listelenmiştir (Kasban ve ark., 2015; Oyar, 2008).

Avantajları arasında;

- Ağrısız ve acısız görüntü elde edilmesi,
- Radyasyon gibi zararlı ışınlar yerine manyetik titreşimin kullanılması,
- X-ray röntgen görüntülerinin aksine oldukça yüksek çözünürlüğe sahip görüntü elde edilmesi,
- İnsan etkisinden bağımsız olarak doğrudan bilgisayar ile görüntüleme yapıldığından hata oranı düşüktür.
- Yumuşak dokularda oldukça başarını görüntü sunmaktadır.
- Görüntüleme öncesi herhangi bir boyar maddeye ihtiyaç duymadan görüntüler elde edilir.

Dezavantajları arasında ise;

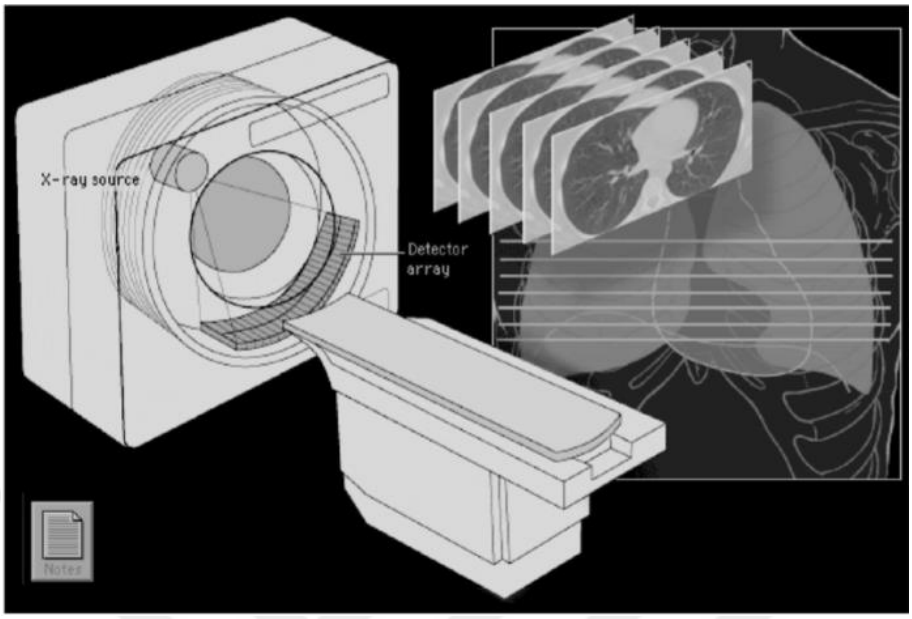
- Hastanın üzerine bulunacak ve manyetik alandan etkilenecek herşeye duyarlı olması,
- Görüntünün çekim sürecinin uzun sürmesi,
- Kurulumunun oldukça pahalı olması.

1.8.3. Bilgisayarlı tomografi (Computerized Tomography)

CT görüntüleme tekniği bilgisayarlı görüntüleme tekniği olarak nitelendirilmektedir. Bilgisayarlı tomografide X-ray ışınları kullanılmaktadır. Özellikle vücutta yer alan bir bölgenin görüntüsünü oluşturma yeteneğine sahip bir cihazdır. Söz konusu teknoloji bilgisayarların destekleri kullanılarak görüntülerin üretildiği ilk biyomedikal yöntemdir (Kumrular ve Polat, 2021; Wikipedia, 2022).

Sürekli hareket halinde bulunan X-ışını tüpüne benzeyen bir Gantry mevcuttur. Sürekli dönen tüpün karşısına konumlandırılmış olan dedektör vasıtası ile görüntülenmesi hedeflenen alanın her açıdan görüntüsünü alır ve bilgisayar aracılığı ile işlemektedir. CT cihazının bulunması görüntüleme alanında büyük bir devrimdir. Özellikle kanser alanında kullanılması hekimlerin teşhis ve tedavi süreçlerinde önemli bir avantaj sağlamıştır (Ünal, 2008; Molino ve ark., 2020).

CT cihazları MR cihazlarına benzer şekilde Gantry, kabinetler ve operatör ve görüntü işlem bilgisayarı olmak üzere üç kısımdan meydana gelmektedir.



Şekil 1.9. Akçiğer grafisi ve bilgisayarlı tomografi cihazı görseli (Ünal, 2008).

CT yani Computerized Tomography esasen eski yunanca kesit-tomo ve görüntü-graphy kelimelerinden meydana gelmiş manası ise bilgisayarlı kesitgörüntü anlamına gelmektedir. 1972 yılında Ambrose ve Hounsfield adı altındaki iki bilim insanı aracılığı ile tanıtımı gerçekleştirilmiştir. Kesitsel görüntü alınabilmesi bilgisayarlı tomografi için teşhis noktasında en avantajlı görüntüleme sistemidir. (Ünal, 2008; Molino ve ark., 2020).

Bilgisayarlı tomografi cihazları ele alandığında özellikle kemik kırıklarının detaylı ve etkili bir müdahale söz konusu olacaksa görüntülenmesinde, iç organlarda meydana gelecek zararın tespitinde ve iç kanamanın teşhislerinde, özellikle göğüs bölgesinde ve iç organlarda meydana gelecek rahatsızlıkların ve şüpheli durumların teşhisinde, kanser ve türevlerinde hastalığın ilerlemesi ve gerilemesi süreçlerinde aktif olarak kullanılmaktadır.

Söz konusu bilgisayarlı tomografi cihazlarının avantaj ve dezavantajları aşağıdaki şekilde listelenerek tasnif edilmiştir (Kasban ve ark., 2015).

Avantajları arasında:

- Ağrısız olması ve hızlı sonuç üretmesi,
- İyi ve yüksek çözünürlükte görüntü üretmesi,

- Her bir kesit bütün detayları ile irdelenebilinecek sonuç üretmesi,
- Doku hasarlarının ve söz konusu hasarlı dokuların sağlıklılarından ayrımının kolayca gerçekleştirilmesi.

Dezavantajları arasında:

- X-ışını ve radyasyon kullanımı radyoaktif maruziyeti arttırması,
- Anlık görüntü yerine çekim yapılırken ki sıradaki görüntüyü sunması,
- Eğer görüntüleme alanında yumuşak doku yoğunluğu azsa istenilen çözünürlükte görüntü sağlanamaması.

1.8.4. Ultrasonografi

Yüksek frekansa sahip ses dalgalarının gönderilerek geri gelmesi ile elde edilen bir görüntüleme yöntemidir. Denizaltılarında kullanılan yahut gece karanlığında yarasaların görmelerini sağlayan aynı metodun benchmarking edilerek sağlık görüntüleme sistemi ile oluşturulmuştur. Özellikle farklı doku yüzeylerinden yansıyan ses dalgaları görüntülemenin oluşmasında önemli rol oynar. Ultrason cihazlarında kullanılan ses dalgaları oldukça yüksek frekansa sahil olması nedeniyle insan kulağınca işitilmez. Söz konusu cihazların kullanım alanı oldukça yaygın kullanımı kolay ve X-ışınlarında olduğu gibi iyonlaştırılmış bir radyo aktif ışında kullanılmadığı için zararsızdır. Yumuşak dokuların gözlemlenmesinde çok kritik rol oynamaktadır (Ünal, 2008).

İnsan kulağının işitebileceği ses aralığı 20 Hertz ile 20 Kilohertz frekans arasında değişmektedir. Bu seslerin altı yahut üstünün insan kulağı tarafından işitilmesi mümkün değildir. Tıbbi alanda görüntülemede kullanılan ultrason cihazları ise 2 Megahertz ile 10 Megahertz arasında değişmektedir (Kasban ve ark., 2015).

Dr. Karl T. Dussik ve başarılı bir fizikçi olan Friederich Dussik tarafından daha öncesinde denizlerde derinlik taramasında kullanılan cihazın medikal alanda kullanılması amacıyla çalışmalarına başladılar. Ultrason dalgaları ile ilk olarak beyin üzerinde çalışsalar da günümüzde hala daha ultrason cihazları ile görüntülenememekte olup bu sebeple başarısız olduklarını düşündüler. 1952 yılı geldiğinde tanı amaçla kullanılan ilk araç Dussik kardeşlerin yapmış oldukları cihaz oldu. 1954 yılında ise kadın hastalıklarında kullanılması fikri Dr. Ian Donald'ın aklına geldi. İlerleyen yaptığı bu çalışmaların bir sonucu olarak çılgın doktor olarak anılacaktı.

Ultrasonografi cihazı ilk olarak makineyi besleyen elektrik enerjisini ses frekanslarına dönüştüren transdüser ve bağlı olduğu probu içerir. Söz konusu dalgaları görüntü olarak algılayan resmedici ve görüntünün şekillendiği katot ışık tüpünden oluşmaktadır (Ünal, 2008).

Şekil 1.10. Yeni nesil bir ultrason cihazı.

Ultrason cihazları yumuşak doku görüntülemelerindeki yetkinlik ve radyoaktif görüntüleme metodu kullanmaması sebebiyle özellikle hamilelik sırasında fetüs gelişiminin her aşamasında, karaciğer vb. abdomen bölgesinde yer almakta olan organların görüntülenmesinde ve kemik kırıkların önden olası eklem ve bağ sorunları ile tanımlanmasında kullanılmaktadır. (Kasban ve ark., 2015).

Söz konusu ultrason cihazlarının avantaj ve dezavantajları aşağıdaki şekilde listelenerek tasnif edilmiştir (Kasban ve ark., 2015).

Avantajları arasında:

- Ağrısız olması ve hızlı sonuç üretmesi,
- X-ışını kullanılmadığı için bir radyoaktif sorundan bahsedilemez,

- Çözünürlüğü yüksek görüntü alınmakta ve anlık görüntü alınmaktadır.

Dezavantajları arasında:

- Standart bir operasyonel prosedür olmadığı için operatörün yetkinliği doğrudan ortaya çıkacak görüntü kalitesini belirlemektedir.
- İstenilen alandan istenilen görüntünün elde edilebilmesi uzun sürebilmektedir.
- Dehidtasyon ve hidrasyon hususlarından etkilenmektedir.

2. KAYNAK ARAŞTIRMASI

Derin öğrenmede kullanılan algoritmaların analiz gücü değerlendirildiğinde ilk akla PASCAL VOC veri setinde yapılan çalışmalar gelmektedir. Yapay sinir ağlarının aktif olarak kullanılmaya başlanmadığı 2013 yılına kadar bu veri seti ile yapılan çalışmaların nesne tespiti başarısı %40 seviyesinin altında kalmaktadır (Girshick ve ark., 2014). Bu yıllardan sonra artan CPU ve GPU teknolojisi ve yapay sinir ağları ile derin öğrenme tekniklerinin algoritmalarda aktif şekilde kullanılması ile %80 seviyesinin üzerine nesne tespitinin yapılması sağlanmıştır (Girshick, 2015).

Medikal görüntü inceleme ile ilgili yapılan çalışmalarda 2022 yılında Samothai ve arkadaşların 1225 X-ray görüntüsünde yer alan küçük kemik kırıklarında yapmış oldukları çalışmanın sonucu olarak YOLO-X ve YOLO-R algoritmaları tekrar eğitilmiş, 122 test veri setinde YOLO-X algoritması %94 oranında kırıkların yerini doğru tespit etmiş, buna karşılık YOLO-R uygulaması ise %59 oranında kırık yerini doğru tespit edebilmiştir (Samothai ve ark., 2022).

2022 yılında Agrawal ve Choudhary tıbbi görüntülemede derin öğrenme üzerine çalışmalar gerçekleştirmiştir. Radyolojik görüntülemede derin öğrenme modelleri tanıtılmış ve klasik modeller ile tartışılmış farklı hastalıklarda için kullanılabilmek adına yeni bir model çizilmiştir (Agrawal ve Choudhary, 2022).

Archana ve Kalirajan'ın 2022 yılında yapmış olduğu çalışma ile 15 farklı hastalık için daha önce hangi bilimsel çalışmalarının yapıldığı ve hangi CNN mimarinin kullanıldığı ve nasıl bir yöntem izlendiği özetlemiştir. Söz konusu çalışmada diyabet hastalarının retina görsellerinden karaciğer hastalarının CT taramalarına kadar birçok görüntü şekli ile ilgili yorumlar derlenmiştir (Archana ve Kalirajan, 2022).

2020 yılında, endüstriyel robot kollarının konumlandırma hatasını belirlemek için bir görüntü işleme algoritması olan Levenberg-Marquardt (LM) Algoritması kullanılmıştır (Li ve ark., 2022).

2021'in başlarında, gerçek zamanlı gözlem temelli bir yangın tespit sistemi, SSD algoritmasının kullanıldığı insansız hava araçlarında uygulanmıştır. Çalışmada,

saniyede 26 kare hızında gerçek zamanlı görüntüler alınmış ve bu görüntülerde %92,7 başarı oranı elde edilmiştir (Nguyen ve ark., 2021).

2022 yılı sonuna gelindiğinde, pelvik X-ray görüntüleri kalça kırıklarının tespiti için YOLOv4 tiny yapay zekâ modeli kullanılmış ve küçük bir insan doktor grubuyla karşılaştırılmıştır. YOLOv4 tiny yapay zekâ modelinin tanısal performansta %94 doğruluk oranı elde edilmiştir (Twinprai ve ark., 2022).

Moharil ve Singh, X-ray görüntülerinde proksimal femur kırıklarının bölgelerini çıkarmak için CNN'ler ve Mask R-CNN'ler kullandılar. Bu çalışma, intrakapsüler ve ekstrakapsüler bölgelere odaklanmıştır (Moharil ve Singh,2021).

Mu ve ark. üç kurumdan 1.491 ön pelvik radyografisi bir araya getirmiş ve "Dijital Radyografi Kırık Tespit Sistemi" adını verdikleri bir model oluşturmuşlardır. Bu çalışma femur boyun kırıklarına odaklanmış ve 0,92 AUC değerleriyle umut verici sonuçlar elde etmiştir (Mu ve ark., 2021).

Tıbbi görüntü analiziyle ilgili çalışmalarda, 2022'de Samothai ve arkadaşları 1225 X-ray görüntüsünde bulunan küçük kemik kırıkları üzerine bir çalışma gerçekleştirmişlerdir. Çalışmalarının bir sonucu olarak, YOLO-X ve YOLO-R algoritmaları 122 görüntüden oluşan bir veri kümesi üzerinde yeniden eğitilmiş ve test edilmiştir. YOLO-X algoritması, kırık konumunu %94 doğruluk oranıyla doğru bir şekilde tespit etmiştir, YOLO-R uygulaması ise kırık konumunu yalnızca %59 doğruluk oranıyla doğru bir şekilde belirleyebilmiştir (Samothai ve ark., 2022).

2022'de Rakesh ve Akilandeswari'nin araştırması, bir Morfolojik algoritma kullanarak kemik kırıklarını tespit etmeyi ve bunun doğruluk oranını Genetik algoritma ile karşılaştırmayı amaçlamaktadır. MATLAB simülasyonundan elde edilen sonuçlara göre, Morfolojik algoritma %87,46 doğruluk oranıyla başarı elde etmişken, Genetik algoritmanın doğruluk oranı %83,25 olarak tespit edilmiştir (Rakesh ve Akilandeswari, 2022).

Peng 2003 yılında yapmış olduğu çalışmada femur üst bölge boyun ve sap arasındaki açının kullanılarak röntgen görüntülerinin ilk defa bilgisayar üzerindeki incelemesi gerçekleştirilmiştir. Kırık tespitinin gerçekleştirilmesi amacıyla eşik değeri belirlenmiş buna göre sınıflandırma yapılmış aktif konturlar yöntemi seçilmiştir. Olması gereken boyun ve sap açısının olağandışı bir aralığa çıkması sonucunda algoritma tarafından söz konusu kemiğin kırık olup olmama durumu belirlenmiştir.

Çalışmada kullanılan 64 X-ray görüntüsünün incelenmesi sonucunda 116 derecelik açı eşik değeri olarak belirlenmiştir. Yapılan çalışma sonucunda sağ femurda %93,8 ve sol femurda ise %95,3'lük bir başarı elde edilmiştir. Ortalama başarı sonucu her iki femur kemiği için %94,5'dır (Peng, 2003).

Mahendran 2012 yılında gerçekleştirmiş olduğu çalışmada kırık tespiti gerçekleştirmek için görüntü sınıflandırma tekniklerini kullanmışlardır. Sınıflandırıcı algoritmalar aracılığı ile kırık tespitini gerçekleştirmek için X-ray görüntüleri kullanılmıştır. Sınıflandırıcı algoritmalar için söz konusu çalışmada NB, DVM ve BPNN sınıflandırıcılarından yararlanılmıştır. Söz konusu füzyon tabanına sahip sınıflandırıcıların kullanımı neticesinde kırık algılama yetkinliğinin geliştiği BPNN ve DVM kombinasyonunun kullanılması sonucunda başarı oranının %97,97'ye geldiği ortaya çıkmıştır (Mahendran, 2012).

Bilgisayarlı tomografi görüntülerini kullanarak Chowdhury ve arkadaşları 2012 yılında otomatik kırık tespiti yapabilen bir sistem geliştirmişlerdir. Potansiyel kırık eğrilik ve yoğunluk çıkarımları ile gerçekleştirilmiştir. Gauss eğriliği ve ortalama ile yerelleştirme ikinci olarak gerçekleştirilmiş. Yanlış pozitif oranı minimize edilmesi söz konusu çalışma ile başarılmıştır. 3D Grafik kesimi, 3 boyutlu (3D) görüntülerde ve veri hacimlerinde kullanılan bir görüntü işleme ve segmentasyon tekniği, Uzamsal Olarak Tutarlı Vadi tekniği ve Ortak karar tekniği kullanılarak %90 üzerinde duyarlılık ve özüllük değerleri elde edilmiştir (Chowdhury, 2012).

Topuk kemiği olarak geçen kalkaneus kemiği kırığı genellikle yüksek enerjili bir travma sonucunda ortaya çıkmaktadır. Pranata ve arkadaşları 2019 yılında CT (bilgisayarlı tomografi) görüntülerinden kırık tepiti yapmak amacıyla çalışma başlatmışlardır. Söz konusu çalışmanın amacı bilgisayarlı tomografi görüntülerinde kırık ve kırık olmayan şekilde kalkaneus kemiği tespit etmek olarak belirlemişlerdir. Yapılan çalışma ile CNN kullanılarak %98 doğruluk oranı ile kalkaneus kemiği kırığını tespit edebildiklerini ortaya koymuşlardır. Söz konusu araştırmacılar transfer öğrenimi tekniklerini kullanarak ResNet modelini SURF ile birleştirerek bu başarı oranına ulaşmışlardır. Tayvanda bir hastaneden alınan 638 kırık ve 1248 normal olmak üzere toplamda 1931 bilgisayarlı tomografi görüntüsü elde edilerek veri seti oluşturulmuştur. Algoritmanın yeniden eğitilebilmesi için eğitim ve doğrulama veri seti için %80-%20 kuralına göre bölünmüştür. Yapılan çalışmada hızlı ve sağlam bir

şekilde özellik çıkarımı için SURF kullanılmış. Çerçevenin bulunması amacıyla Canny kenar seçimi ve OpenCV de çalışmada kullanılmıştır (Pranata ve ark., 2019).

Bin Guan ve arkadaşları 2019 yılında yapmış oldukları çalışma ile kol kemiği kırığı tespiti için röntgen resimlerine yeni bir derin öğrenme tekniği uygulamıştır. Çalışmanın esas amacı kırık ve kırık olmayan kol kemiği röntgenlerinin sınıflandırılması ve kol kırığının lokasyonunun tespiti amaçlanmaktadır. Çalışmada MURA veri seti kullanılarak ortalama hassasiyet (precision) değeri %62,04 olarak elde edilmiştir. Söz konusu çalışmada derin öğrenme yöntemlerinden biri olan iki kademeli R-CNN yöntemi kullanılmıştır. Ön işlem sırasında piksel transferi yöntemi kullandıktan sonra özellik çıkarımı adımı için özel bir piramit mimarisi yöntemi kullanılmıştır. Modelin eğitimi için 3392 adet resim kullanılırken test için 612 resim kullanılmıştır. Bu yöntemde çok küçük olan kırıkları, kırıktaki gri çizgileri tespit edemediği ve örtüşen kırıkları da yine tespit edemediği ortaya konulmuştur (Guan ve ark., 2019).

Thian ve arkadaşları 2019 yılında el bileği radyografilerini inceleyerek otomatik bir şekilde el bileği kırıklarını ortaya koymaya amaçlamışlardır. CNN algoritması kullanılarak el bileğindeki radius ve ulna kırıklarının tespit ve bölgesinin işaretlenmesindeki yeteneği ölçülmüştür. Yapılan çalışma ile %95 doğruluk oranında başarı elde edilmiştir. Yapılan çalışma araştırmacıların görev aldığı hastanenin acil servisinden elde edilen el bileği radyografileri kullanılmıştır. Veri kümesinde toplamda 7356 el bileği radyografisi derlenmiştir. Söz konusu veri kümesi %90-%10 metodu ile eğitim ve test veri seti olarak ayrıştırılmıştır. Görüntüler etiketlenirken uzman doktorlardan destek alınmıştır. Eğitim sırasında ön işleme esnasında resimlere yatay döndürme uygulanmıştır. Bu çalışmadaki veri setinin öne çıkan bir özelliği yetişkin ve pediatrik görüntülerin ayrı olarak hem eğitim hem test veri setinde yer verilmesidir. Söz konusu çalışmada analizin gerçekleştirilmesi için STATA istatistiksel yazılımı kullanılmıştır. Bu çalışma için elde edilen görüntüler sadece acil servisinden elde edilmiştir (Thian, 2019).

Lindsey ve arkadaşları 2018 yılında yaptıkları çalışmada el bileklerinin X-ray görüntülerini incelemiş ve söz konusu radyografileri CNN algoritmaları ile kırık ve kırık olmayan şeklinde gruplandırmaya çalışmışlardır. Söz konusu çalışmada gruplandırmanın yanı sıra kırık olan bölgenin de tespitinin yapılması amaçlanmıştır. DCNN (Deep convolutional neural networks) adı altındaki doğrusal olmayan bir

regresyon modelini kullanmışlardır. Veri setini oluştururken toplamda 2000-2016 olmak üzere 16 yıllık periyodu kapsayan sürede Özel Ameliyatlar Hastanesi'ne başvuruda bulunan hastaların X-ray görüntüleri kullanılmıştır. Söz konusu çalışmada toplamda insan vücudunun farklı kemiklerine ait 135.845 X-ray görüntüsü bir araya getirilmiş bunun neticesinde içerisinde seçim yapılarak 34.990 el bileği X-ray görüntüsü ayrıştırılmıştır. Modelin testinin gerçekleştirilmesi için 3.500 el bileği ve 1.400 anterior ve posteriyor görüntüleri kullanılmıştır. X-ray görüntülerinin kırpılması, yeniden boyutlandırılması işlemi en boy oranı korunarak gerçekleştirilmiş bu şekilde önişleme işlemi tamamlanmıştır. Sonuç olarak 1024x512 çözünürlüğüne sahip sabit çözünürlükte çıktılar elde edilmiştir. Bu çalışma yapay zekaya tıpkı ortopedi uzmanlarına öğretilebileceği şekilde el bileği kırıklarının öğretilebileceğini göstermiştir. Söz konusu çalışmanın sonucunda %96 seviyesinde AUC (Area Under the ROC Curve) başarısı elde etmiş aynı zamanda ortopedi el cerrahları ile model değerlendirilmiştir (Lindsey ve ark., 2018).

Bin Guan ve arkadaşları 2019 yılında yapmış oldukları çalışma ile uyluk kemiklerinden alınan görüntülere CNN tekniklerinin uygulanması sonucunda kırık tespitinde %82 seviyesinde bir başarı elde etmişlerdir. Çalışmayı gerçekleştiren araştırmacılar genişletilmiş konvolüsyonlu özellik piramidi ağı önerdiler ve bu uygulamanın gücünü klinik ortamda ortaya koydular. Linyi Hastanesinden elde edilen uyluk kemiği görüntüleri modelin eğitilmesi için kullanıldı ve test edildi. Dijital radyografi tekniği ile toplamda 3842 uyluk kırığı görüntüsü odağa alınmıştır. Söz konusu 3842 görüntü 3484 eğitim veri seti için 358 dijital radyografi görüntüsü ise test veri seti için rastgele olarak ayrıştırılmıştır. Veri arttırımı gerçekleştirmek için söz konusu çalışmada yatay döndürme kullanılmıştır. Araştırmacılar ResNet ağını R-CNN'de özellik çıkarımı için kullanmışlardır. Ağın ağırlığını güncellemek adına Stokastik Gradyan iniş yöntemi kullanılmıştır. Söz konusu modelin uyluk kemik kırıklarını tanıyacak şekilde yeniden eğitilmesi için Nvidia GeForce GTX 1080Ti GPU kullanılmıştır. Algoritmanın yeniden eğitilmesi 6 saat sürmüştür (Bin Guan ve ark., 2019).

2021 yılında Ma ve arkadaşları yapmış olduları özgün çalışma ile üst ve alt ekstremite, üst ve alt gövde kemiklerinin tamamını kırık ve kırık olmayan şeklinde sınıflandırmak için "cracknet" adı altında yeni bir sınıflandırma ağı önerdiler. Çalışmalarında öncelikli olarak Faster R-CNN algoritmasını kullanarak X-ray görüntülerinde 20 farklı

bölge önerisi buldular sonrasında bu 20 bölgede kırığın tespitini gerçekleştirmek adına cracknet'i uyguladılar. Çalışma sonucunda araştırmacılar %90,11 seviyesinde doğruluk ve %82 ortalama duyarlık değeri elde ettiler. Çalışma sırasında kafatası üst gövde, alt gövde alt ve üst ekstremite kemikleri üzerinde gerçekleştirilerek toplamda 3053 radyografi görüntüsünden oluşan bir veri seti kullanıldı. Araştırmacılar bu çalışma ile derin öğrenmenin tıbbi tanıda ne kadar hızlı ve doğru çözümler sunduğunu ortaya koydular (Ma, 2021).

Nguyen ve arkadaşlar 2021 yılında yapmış oldukları çalışma ile kol kırıkları X-ray görüntüleri üzerinden derin öğrenme algoritmalarının kırık tespitindeki başarısını ortaya koymayı amaçlamışlardır. Özellikle dünya genelindeki kol kırıklarının sayısına dikkat çeken araştırmada ilk olarak X-ray görüsünü incelemek için YOLACT++ ile görüntülerin segmentasyonunu ve görüntü kontrastlığının geliştirilmesi için algoritmalar kullanılmıştır. Veri arttırma teknikleri kullanılarak veri arttırımı yapılan çalışmada YOLOv4 algoritması X-ray görüntülerinde kemik kırıklarını hem tanımlanması hem de konumlarının belirlenmesi amacıyla kullanılmıştır. Kullanılan araştırmacıların geliştirmiş olduğu özgün yöntem Faster-RCNN algoritmasını da geride bırakarak test veri setinde ortalama duyarlılıkta %81,9 oranında başarı elde etmiştir (Nguyen ve ark., 2021).

Rao ve arkadaşlarının 2020 yılında yapmış oldukları çalışmada BPNN (Back Propagation Neroural Network) geri yayılım yapay sinir ağı ve evrişimli yapay sinir ağı kullanılmıştır. Söz konusu çalışmada el, ayak, göğüs ve kafatası kemiklerinin X-ray görüntüleri derlenmiştir. Bunun neticesinde veri seti için toplamda 400 resim elde edilmiştir. Her bir kemik grubundan elde edilen 100 resim 70 eğitim ve 30 test olacak şekilde ayrıştırılmıştır. Çalışmanın sonucunda değerlendirme matrisi ile BPNN ve CNN algoritmalarının sonuçları kıyaslanmış buna göre doğruluk oranı (accuracy) için %89 BPNN algoritması %89 oranında başarı elde edilirken CNN algoritmasında %91 oranında başarı elde edilmiş, duyarlılık oranı (sensitivity) için BPNN algoritması %87 oranında başarı elde edilirken CNN algoritmasında %90 oranında başarı elde edilmiş ve özüllük oranı (specificity) için ise BPNN algoritmasında %88 oranında başarı elde edilirken CNN algoritmasında %91 oranında başarı elde edilmiştir (Rao ve ark., 2020).

2019 yılında Basha ve arkadaşları radyologların ve ortopedi uzmanlarının kısa zamanda hem kemik kırıklarını tespit etmek hem de yorumlanmasına yardımcı olmak amacıyla bilgisayar destekli tanı sistemi geliştirmek amacıyla yola çıkmışlardır.

Çalışmada geliştirilmiş Haar Dalgacığı Dönüşümü (Haar Wavelet Transform), ölçekten bağımsız özellik dönüşümü (Scale-Invariant Feature Transform- SIFT) ve geri yayılım yapay sinir ağı (Back propagation neural network) kullanılarak etkili bir biçimde kemik kırığı tepiti hedeflenmiştir. İlk iki teknik özellik çıkarımı için kullanılmakta iken sonuncusu kırık görüntülerin sınıflandırılması için kullanılmıştır. Yapılan çalışmada SIFT tekniği ile yapılan tespitin daha iyi olduğu ortaya konulmuş 300 farklı kemik kırığının yer aldığı X-ray görüntüleri üzerinde gerçekleştirilen çalışmada sınıflandırma oranı %93,4 olarak ortaya çıkarılmıştır (Basha ve ark., 2019).

Yu ve çalışma arkadaşları 2020 yılında gerçekleştirdikleri çalışmada kalça kırıklarının tespiti ve yerinin işaretlenmesinde bir derin öğrenme algoritması geliştirmiş, geliştirilen bu algoritma test edilmiş ve doğruluğu değerlendirilmiştir. İkili sınıflandırma için söz konusu algoritmanın ürettiği çıktılarda %97,1 hassasiyet elde ederlerken, %96,7 oranında özüllük seviyesi elde etmiştirler. Bu çalışmada kırıkların lokasyonunun tespitinde hassasiyetin normal bulgularında %95,1, transservikal ve subkapital kırıkların olduğu X-ray görüntülerinde %84,1, interokanterik ve baziservikal kırıklarda %76 ve subtrokanterik kırıklarda başarı seviyesi %20'de kalmıştır (Yu ve ark., 2020).

2021 yılında Liu ve beraberindeki araştırmacılar TPF – Tibial plateu kırıklarının tespit edebilmesi için bir derin öğrenme algoritması geliştirdiler. Söz konusu algoritmanın performansını ortopedi doktorları ile kıyasladılar. Algoritma sonuç ürettiği sırada şüpheli tibial plateu kırık bölgesini otomatik olarak bir dikdörtgen içerisine almak sureti ile işaretledi. Ortopedi doktorlarının ürettiği sonuçlar ile algoritmanın sonuçları kıyaslanabilir seviyelerde olduğu görüldü. Algoritmanın elde ettiği sonuçlar %91 seviyelerindeyken insanın performansı %92 seviyelerinde ölçülmüştür. Sonuçların bu kadar yakın olmasına karşın asıl fark tespit sürelerinde ortaya çıkmakta ve yapay zekâ ortopedi doktorlarına kıyasla 16 kat daha hızlı sonuç üretmekteydi (Liu ve ark., 2021).

2012 yılında Ekşi ve Çakıroğlu yapmış oldukları araştırma ile uzun kemiklerde yer alan kırıkların tespiti için bir derin öğrenme modeli oluşturmuşlardır. Çalışmalarında popüler gruplama algoritmaları ele alınmıştır. Otsu, Bulanık C ortalama kümeleme (FCM), K-Ortalama olmak üzere üç kümeleme algoritması söz konusu çalışmada kullanılmıştır. Ön işlemenin bu algoritmalar ile gerçekleştirilmesi sonrasında yapay sinir ağları kullanılarak uzun kemiklerde kırıkların tespiti %89 oranında başarı ile gerçekleştirilmiştir (Ekşi ve Çakıroğlu, 2012).

2023 yılında Cohen ve araştırmacılar el bileği kırıklarının tespitinde yapay zekâ ile radyologların başarısını kıyaslamak adına çalışma başlatmışlardır. Söz konusu çalışmanın gerçekleştirilmesi adına 2017 yılı Ocak ayında başlayarak 2019 Aralık ayına kadar geçen süredeki el bileği travmasına maruz kalan 637 hastadan alınan görüntüler derlenmiş toplamda 1917 X-ray görüntüsü elde edilmiştir. Derin sinir ağı algoritması kullanılarak yapay zekâ tarafından sonuçlar elde edilmiştir. Çalışmada gerçek değerler alanında uzman kas iskelet alanında oldukça tecrübe sahibi radyolog tarafından oluşturulmuştur. Başlangıç radyoloji raporları ile yapay zekâ sonuçları ve her ikisinin kombinasyonu karşılaştırılmıştır. Yapılan çalışma ile yapay zekanın hassasiyeti %83, özüllük değeri %96 olarak ortaya çıkarılmıştır (Cohen ve ark., 2023).

Naderian ve Sanei 2022 yılında Nottingham Trent Üniversitesinde yaptıkları araştırmada sadece kemik kırıklarına değil aynı zaman kemik çıkıklarına da odaklanmışlardır. Yaptıkları çalışmada doğru tespitin önemi için kemik segmentasyonun önemini vurgulamışlardır. Geleneksel görüntü gruplamasında görüntünün homojenliği ön plana çıktığı ve piksel bazlı işaretlemenin uzun süreceğinden bahseden araştırmacılar bu yöntemin çıkık tespitinde ve kırık olan bölgesi algılamada yanılgıya sahip olabileceğini ifade etmişlerdir. Yaptıkları araştırmada Fast-RCNN algoritmasını öğrenme transferi metodu kullanılmıştır. Piksel bazlı etiketleme olmaksızın doğru kemik segmentasyonuna sahip olmak için uzaklık düzenlemeli seviye set evriminin (distance regularized level set evolution – DRLSE) bu çalışmada kullanılması hem kırık etrafı bölgenin hem de çıkık kemiklerin sınıflandırılmasını mümkün kılmıştır. Yapılan çalışmada söz konusu yöntem insan eli X-ray görüntülerine uygulandığında %95 oranında doğruluk, %96 oranında hassasiyet elde edilmiştir (Naderian ve Sanei, 2022).

Kim ve MacKinnon 2018 yılında yapmış oldukları çalışma ile Evrişimli sinir ağlarını kullanarak yapay zekâ ile el bileğindeki kemik kırıklarının tespiti üzerine çalışmışlardır. Inception v3 algoritması kullanılan çalışmada el bileğinin X-ray görüntüleri kullanılarak yeniden eğitilmiştir. Söz konusu çalışma için 695 kırk içeren, 694 kırık içermeyen olmak üzere toplamda 1.389 görüntü veri arttırma teknikleri de kullanılarak algoritma kırık veya kırık yok olacak şekilde yeniden eğitilmiştir. Eğitim veri seti %80 eğitim, %10 doğrulama ve %10 test olacak şekilde ayrıştırılmıştır. Çalışma sonucunda söz konusu yeniden eğitilerek oluşturulmuş olan yapay zekâ AUC

değeri %95,4, hassasiyet değeri (sensitivity) %90 ve özüllük değeri %88 olarak elde edilmiştir (Kim ve MacKinnon 2018).

Cheng ve araştırma grubu 2019 yılında düzlemsel ön pelvis radyografileri kullanarak kalça kırıklarının tespiti ve yerinin belirlenmesi için derin konvolüsyonel sinir ağlarını kullanmışlardır. 2012 yılının Ocak ayından başlayarak 2017 yılının Aralık ayına kadar geçen sürede toplamda 5,505 ekstremite radyografisi ile derin konvolüsyonel sinir ağı transfer öğrenim yöntemi ile yeniden eğitilmiştir. Bu veri setinden bağımsız olarak elde edilen 100 X-ray radyografi görüntüsü ile yapay zekâ algoritmasının performansı doğruluk, hassasiyet, yanlış negatif oran ve alıcı işletim karakteristiği eğrisi altında alan olan (AUC) değerleri ile ölçülmüştür. Yapılan çalışma sonucunda kalça kırıklarını tespit etme konusunda DCNN algoritması %91 doğruluk, %98 hassasiyet, %2 yanlış negatif oran ve %0,98 AUC sonuç üretmiştir (Cheng ve ark., 2019).

Kitamura ve arkadaşları 2019 yılında yapmış oldukları çalışma ile 596 kırık ve kırık olmayan ayak bileği X-ray görüntüsünü bir araya getirip bir veri seti oluşturmuştur. Söz konusu veri seti konvolüsyonel sinir ağları ile eğitilip eğitilemeyeceği denenmiştir. Eğitim öncesinde veri arttırma teknikleri kullanılarak veri seti çoklandırılmıştır. Çalışmada python yazılım dili kullanılırken TensorFlow kütüphanesi, Inception V3, ResNet ve Xception konvolüsyonel sinir ağları kullanılmıştır. Eğitimi gerçekleştirilen modelin performansının ölçülmesi için doğruluk, pozitif tahmin değeri (PPV), negatif tahmin değeri (NPV), hassasiyet ve özüllük değerleri seçilmiştir. Tek radyografik görüntüler için %76 doğruluk elde edilirken, üç görünüm kullanılarak %81 doğruluk oranı elde edilmiştir (Kitamura ve ark., 2019).

Gan ve arkadaşları 2019 yılında yapmış oldukları çalışma ile yapay zekâ algoritmalarından konvolüsyonel sinir ağları yardımı ile ön ve arka bilek radyografileri kullanılarak bilek distal radius kırıklarının tespitini ortaya koymayı amaçlamışlardır. 2.340 hastadan alınan X-ray görüntüleri çalışmada veri seti olarak dahil edilmiştir. Söz konusu konvolüsyonel sinir ağlarının eğitimi sonrasında performansını değerlendirmek adına alıcı işlem karakteristiği eğrisi altında kalan alan, doğruluk, hassasiyet ve özüllük değerleri kıyaslanmak adına odağa alınmıştır. CNN algoritmasının alıcı işlem karakteristiği eğrisi altında kalan alan değeri %96 olarak elde edilmiştir. Yeniden eğitilen yapay sinir ağı radyolog gruplarına kıyasla distal

radius kırıkları ve kırık olmayan görüntüleri ayırt etme açısından daha yüksek bir performans sergilediği kanıtlanmıştır (Gan ve ark., 2019).

2019 yılında Choi ve arkadaşları yapmış oldukları çalışma ile anteroposterior (AP) ve lateral dirsek radyografilerinin ikisini de kullanarak pediatrik supracondylar kırıkların otomatik tespiti için çift girişli konvolüsyonel sinir ağı (CNN) tabanlı bir derin öğrenme algoritması geliştirmeyi amaçlamışlardır. Derin öğrenme modelini yeniden eğitmek amacıyla 2013-2017 yılları arasında 1.266 çift anteroposterior ve lateral dirsek radyografisini derlenmiş ve veri seti oluşturulmuştur. Eğitimin gerçekleştirilebilmesi için söz konusu veri seti %80-%20 kuralına göre ayrıştırılmış buna göre 1.012 çift X-ray görüntüsü eğitim, 258 çift X-ray görüntüsü ise doğrulama veri seti olarak ayrıştırılmıştır. Bağımsız bir zaman test veri seti oluşturmak amacıyla aynı kuruluştan 2018 yılına ait 258 çift X-ray görüntüsü elde edilmiş, coğrafi test seti için farklı bir hastaneye ait 95 çift X-ray görüntüsü derlenmiştir. Modelin eğitiminin gerçekleştirilmesinin ardından karakteristik eğrisi altında alan (AUC), hassasiyet, özgüllük, pozitif prediktif değer (PPV) ve negatif prediktif değeri (NPV) hesaplanarak karşılaştırılmıştır. Eğitimi tamamlanan model doğrulama veri seti için %97,6, zaman test veri setinde %98,5 ve coğrafi test verisetinde %99,2 AUC sonucu üretmiştir (Choi ve ark., 2019).

2020 yılında Majkowska öncülüğündeki araştırma ekibi radyologlar tarafından karara bağlanmış referans standartları kullanarak göğüs röntgeni yorumlaması için derin öğrenme modelleri geliştirmeyi amaç edinmiştir. Ele aldıkları derin öğrenme modellerini kullanarak frontal göğüs röntgenlerinde yer alan pnömotoraks, opaklık, nodül veya kitle ve kırık olmak üzere dört bulguyu tespit etmeyi amaçlamışlardır. Bu çalışmada iki veri seti kullanılmıştır. İlk veri seti çoklu şehir hastane ağı 759.611 görüntü içerirken, ikinci veri seti ChestX-ray14 kamuya açık 112.120 görüntüden oluşmaktadır. Doğal dil işleme ve uzman incelemeleri neticesinde 657.954 eğitim görüntüsü için etiket sağlanmıştır. Test veri setleri için ise çoklu şehir hastane ağı veri setinden 1.818 ve ChestX-ray14'den 1.962 görüntü içerecek şekilde ayarlama yapılmıştır. Performans, alıcı işletim karakteristik eğrisi analizi, hassasiyet, özgüllük ve pozitif tahmin değeri ile çıktılar değerlendirilmiştir. Çoklu şehir hastane ağı veri setinde, pnömotoraks, nodül veya kitle, hava yolu opaklığı ve kırık için nüfus düzeyine ayarlanmış alıcı işletim karakteristik eğrisi altında alan sırasıyla %95, %72, %91 ve

%86 olarak bulunmuştur. ChestX-ray14 ile alıcı işletim karakteristik eğrisi altında alan yine sırasıyla %94, %91, %94 ve %81 olarak bulunmuştur (Majkowska ve ark., 2020).

2018 yılında Heimer ve araştırma grubu ölüm sonrası bilgisayarlı tomografi (PMCT) görüntüleri üzerine çalışmıştır. Ölüm sonrası bilgisayarlı tomografi genellikle büyük miktarlarda görüntüleme verisi üretmektedir. Söz konusu çalışmada kavisli maksimum yoğunluk projeksiyonlarında kafatası kırıklarına göre vakaların tasnif edilmesi için derin öğrenme modelleri kullanılmıştır. Veri seti oluşturulması için kavisli maksimum yoğunluk projeksiyonları 75 hasta için oluşturulmuştur. Çalışmada denetimli eğitim kullanılarak kırıklar manuel olarak işaretlenmiştir. Gerçekleştirilen eğitim sonucu olarak AUC değeri %96,5 olarak hesaplanmıştır. Çalışmada duyarlılık %91,4 özüllük ise %87,5 olarak hesaplanmıştır (Heimer ve ark., 2018).

2022 yılında Wang ve arkadaşları spiral bilgisayarlı tomografi üzerinde mandibur kırıklarının tespiti konvolüsyonel sinir ağlarının başarısını ortaya koymak adına çalışma gerçekleştirmiştir. 2013 yılının Ocak ayından başlayarak 2020 yılının Temmuz ayına kadar geçen 7,5 yıllık süre içerisinde bilgisayarlı tomografi taraması gerçekleştirilen 686 mandibular kırık hastasının verileri derlenmiştir. Söz konusu kırıklar deneyimli çene cerrahı uzmanlarınca etiketlendirilmiş ve sınıflandırılmıştır. Konvolüsyonel sinir ağı içeren U-Net ve ResNet algoritmaları yeniden eğitilmiştir. Konvolüsyonel sinir ağı içeren U-Net ve ResNet algoritmalarının kırıkları tespit edebilme performansları, yapılan çalışma ile Dice Skoru, doğruluk, duyarlılık, özgüllük ve ROC eğrisi altında alan (AUC) ile değerlendirilmiştir. Çalışmada veri seti olarak düzenlenen 686 hastanın bilgisayarlı tomografi görüntülerinden 1.506 adet mandibular kırık teşhis edilmiş Dice skoru %94,3 olarak hesaplanmıştır. Doğruluk oranı tespitlerinin tamamı %90 üzerinde kalırken, ortalama AUC değeri ise %95,6 olarak hesaplanmıştır (Wang ve ark., 2022).

2021 yılında Rayan ve arkadaşları akut pediatrik dirsek anormalliklerinin X-ray görüntülerinde derin öğrenmenin kullanılması sonucu ikilisel sınıflandırma yapılıp yapılamayacağı üzerine çalışmışlardır. Toplamda 58.817 dirsek görüntüsü ve ilgili radyoloji raporları 2014 yılı Ocak ayından, 2017 yılı Aralık ayına kadar geçen süredeki 4 yıllık dönem için geriye dönük olarak ele alınmıştır. Çocuk hastanesinden alınan verilerin yaş ortalaması 7,2 yıldır. Veri setinin %43 kız çocuklarına aitken %57'lik kısmı erkek çocuklarına aittir. Çalışmanın gerçekleştirilebilmesi için rastlantısal olarak 20.350 görüntü eğitim veri setinde kullanılırken 1.106 görüntü doğrulama veri seti için

ayrıştırılmıştır. Söz konusu anomalileri tespit etmek üzere özelleştirilecek modeli oluşturmak için hem konvolüsyonel sinir ağı hem de yinelenen sinir ağı bir araya getirilmiştir. Çalışmada performansı doğru ortaya koyabilmek adına Duyarlılık, özgüllük, pozitif öngörü değeri, negatif öngörü değeri, alıcı işletme karakteristik eğrisi altında alan değerleri hesaplanmıştır. Söz konusu modelin ürettiği sonuçlar ele alındığında AUC değeri %95, duyarlılık %91, özüllük %84 olarak hesaplanmıştır. Bu çalışma ile derin öğrenme ve bilgisayarlı görü teknikleri kullanılarak akut ve akut olmayan pediatrik dirsek anormalliklerini sınıflandırılmıştır (Rayan ve ark., 2021).

Adam ve arkadaşları 2019 yılında femur boyun kırıklarının radyografilerini odağa alıp derin konvolüsyonel sinir ağlarını kullanarak kırıkların tespitini sağlamayı amaçlamışlardır. AlexNet ve GoogLeNet mimarileri kullanılarak derin konvolüsyonel sinir ağları hem ön hem de arkadan gerçekleştirilen kalça radyografileri ile eğitilmişlerdir. Eğitimin gerçekleştirilmesi için farklı boyuttaki 200, 320 ve 640 radyografi olmak üzere üç veri seti %80-%20 olarak ayrıştırılmış ek olarak 160 radyografi test veri seti için ayrılmıştır. Doğruluk oranı ele alındığında AlexNet için sırası ile %81.9 ve %88.1'den %89.4'e yüksele gelen sonuçlar elde edilmiştir. GoogLeNet mimarisinde %94,4 oranında doğruluk oranı sonucu elde edilmiştir (Adam ve ark., 2019).

2019 yılında Balaji ve arkadaşları kaydırılabilir pencere yaklaşımı ve destek vektör makinesinin kombinasyonunu kullanarak X-ray görüntülerinde diyafiz femur kırığı bölgesini otomatik olarak tespit eden ve teşhis eden bir sistem sunmaktadır. Ayrıca geri yayılım (back probation) ve olasılık sinir ağı sınıflandırıcı diyafiz femur kırığının türünün ortaya konulması için kullanılmıştır. Yeniden eğitilerek ortaya konulan konvolüsyonel sinir ağının performansının objektif bir şekilde göstermek adına 175 gerçek hasta verisi de kırık ve kırık olmayan şekilde kaydedilerek ölçülmüştür. Yapılan çalışmanın sonucunda doğruluk oranı %90,7, duyarlılık oranı %92,3, özüllük oranı ise %86,6 olarak hesaplanmıştır (Balaji ve ark., 2019).

Rahmaniar ve Wang 2019 yılında yapmış oldukları çalışma ile daha hızlı ve daha ayrıntılı bir gözlem elde etmek için kalkaneus kırığı tespiti için bilgisayar destekli bir yöntem önermektedir. Kalkaneus kırıkları çoğunlukla egzersiz ve aktiviteler sırasında meydana gelmekte tespiti bilgisayarlı tomografi yöntemi ile manuel olarak gerçekleştirilmektedir. Gerçekleştirilen çalışma ile doğruluk oranı %86 olarak ortaya çıkarılmıştır. Test görüntüsündeki sonuçlar, referans veri kümelerindeki doktor

tarafından yapılan değerlendirme ve değerlendirme temel alınarak doğrulanmıştır (Rahmaniar ve Wang, 2019).

Tomita ve arkadaşları 2018 yılında yaptıkları çalışma ile radyologların uygulama seviyesinde rastlantısal osteoporotik vertebral kırıklarını karın bölgesi, pelvis bölgesi ve göğüs bölgesinde gerçekleştirilen bilgisayarlı tomografi taramaları sırasında yapay zekâ desteği ile tespit edebilen bir sistem kurulması amaçlanmıştır. Söz konusu sistem osteoporotik vertebral kırıklarını tespiti için her bir bilgisayarlı tomografi kesitinden radyolojik özelliklerin belirlenmesi için derin konvolüsyonel sinir ağı kullanmaktadır. Oluşturulan sistem toplamda 1.432 bilgisayarlı tomografi taramasında sagittal görünümdeki 10.546 iki boyutlu görüntülerden oluşan eğitim ve değerlendirme veri kümesi kullanılarak eğitilmiştir. 129 bilgisayarlı tomografi görüntüsünde bağımsız olarak test edilmiştir. Test veri kümesinde gerçekleştirilen çalışma sonucunda %89,2 oranında doğruluk ve %90,8 oranında F1 puanı elde edilmiştir. Sonrasında bu elde edilen test kümesi başarısı dünyada yer alan klinik koşullardaki radyologların performansı ile eşleştiği yine gerçekleştirilen bu çalışmanın bir sonucu olarak ortaya konulmuştur (Tomitta ve ark., 2018).

2018 yılında Muehlematter ve arkadaşları insüfiziyans kırığı riski taşıyan omurilik hastalarını tanımasının gerçekleştirilmesi için makine öğrenimi algoritmaları ile kemik doku analizinin birleştirilerek bilgisayarlı tomografi taramasında kullanılmasını sağlamıştır. 2006 ile 2013 yılları arasında yer alan 7 yıllık periyottaki 58 omurga insüfiziyans kırığı olan hastanın bilgisayarlı tomografi taramaları analiz edilmiştir. Denetimli makine öğrenmesi teknikleri kullanılarak omurga karşılaştırmasında, olguların omurgaları "istikrarsız" ve "istikrarlı" olarak sınıflandırılmıştır. Sonuçların elde edilmesi için 120 omurga kemiği kullanılmıştır. Buna göre Hounsfield birimi ölçümlerinin AUC değeri %83, makine öğrenmesi tekniklerinin kullanılması ile elde edilen AUC değeri %97 olarak elde edilmiştir (Muehlematter ve ark., 2018).

2018 yılında Urukawa ve arkadaşları kalça kırıklarının konvolüsyonel sinir ağlarının kullanılarak tespit edilmesi üzerine çalışmışlardır. X-ray görüntülerinden elde ettikleri veri seti ile intertrokanterik kalça kırıklarının tespiti amaçlanmıştır. Konvolüsyonel sinir ağı olarak VGG-16 kullanılmıştır. Konvolüsyonel sinir ağının eğitimi için toplamda 3.346 kalça X-ray görüntüsü derlenmiştir. Bunlardan 1.773 tanesinde kalça kırığı yer alırken 1.573 adet X-ray görüntüsünde sağlam kalça görüntüsü yer almaktadır. 2.678 görüntü eğitim veri setinde dahil edilmiş, 334 tanesi değerlendirme

veri setine dahil edilmiş, 334 tanesi ise test veri setine dahil edilmiştir. Modelin başarısını ortaya koymak adına bağımsız görüntülerden oluşan test veri seti sonuçları doğruluk oranı, duyarlık oranı, özüllük oranı ve AUC oranına göre ölçümleme yapılmıştır. Söz konusu çıktılar beş ortopedi uzmanı ile kıyaslanmıştır. Yapılan çalışma ile ortopedi doktorlarının performansları doğruluk oranında %92,2, duyarlık oranında %88,3, özüllük oranında %96,8 ve AUC oranında ise %96,9 olarak elde edilmesine karşın, VGG-16 mimarisine sahip yeniden eğitilen konvolüsyonel sinir ağı doğruluk oranında %95,5, duyarlık oranında %93,9 ; özüllük oranında %97,4 ve AUC oranında ise %98,4 olarak daha başarılı sonuçlar üretmiştir (Urukawa ve ark., 2018).

2020 yılında Yadav ve Rathor gerçekleştirmiş olduğu araştırmada kemik kırıklarının tespiti ve sınıflandırması için derin öğrenme yaklaşımını kullanmışlardır. Veri arttırma teknikleri kullanılarak X-ray görüntülerinden elde edilen veri seti çoklandırılmıştır. Veri çoklandırması sonrasında veri sayısı toplamda 4.000 adete erişmiştir. Modelin performansını değerlendirmek için softmax ve Adam optimizasyon yöntemlerini kullanarak üç deney gerçekleştirilmiştir. Modelin kırık ve sağlam kırık sınıflandırması 5 kat çapraz doğrulama kullanıldığında %92,44 oranında başarı elde edilmiştir (Yadav ve Rathor, 2020).

2019 yılında Vasilakakis ve beraberindeki çalışma ekibi hali hazırda zor olan kırık kemik tespiti sürecini gerçekleştirecek bir sistem üzerine çalışmışlardır. Çalışmanın gerçekleştirilmesi için üst ve alt ekstremiteye ait veri setini toplamda 300 X-ray kemik görüntüsünü toplayarak oluşturmuşlardır. Bahsi geçen kemik kırıklarının tespitinin sağlanması için yeni bir özellik çıkarma ve sınıflandırma yöntemi olan “Wavelet Fuzzy Phrases (WFP)” yöntemi kullanılmıştır. FP, 2B Kesirli Dalga Dönüşümü görüntülerinin farklı bantlarından dokusal bilgi çıkarır ve bir kelime kümesi tarafından ifade edilir. Her kelime, bir bulanık küme tarafından temsil edilmektedir. Kelimeler, görüntü içeriğini temsil eden bulanık küme birleştirmesi olan ifadeleri oluşturur. Yapılan çalışma ile kemik kırıklarının tespitinde elde edilen sınıflandırma doğruluğu %84 olarak elde edilmiştir (Vasilakakis ve ark., 2019).

Mukesh ve Dass 2022 yılında gerçekleştirdikleri çalışma ile kemik kırıklarının tespitinin gerçekleştirilmesi için yenilikçi bir sistem geliştirmeyi ve bunu da klinik uzmanlar ile kıyaslamayı amaçlamışlardır. X-ray görüntüleri ilk etapta açık kaynak veri seti sunan web sitesinden yararlanılarak veri seti derlenmiştir. Doğru kırık tespiti için MATLAB (2013), Derin Sinir Ağı ile Zekâ Kemik Kırığı Tespit Sistemi'ni

karşılaştırmak için kullanılmıştır. Birleşik Zekâ Kemik Kırığı Tespit Sistemi'nin derin sinir ağı üzerindeki karşılaştırması, SPSS yazılımı kullanılarak bağımsız örneklem t-testi ile yapılmıştır. Geliştirilen sistemin doğruluk oranı %91,67 olarak ölçülmüştür (Mukesh ve Dass, 2022).

2020 yılında Karimunnisa ve arkadaşları yapmış oldukları çalışma ile ortopedi uzmanlarının desteği olmaksızın X-ray görüntüleri üzerinden kemik kırıklarının tespitini sağlayan bir sistem geliştirmeyi amaçlamışlardır. Yapılan çalışmada geri yayılım sinir ağı (BPNN), Canny kenar segmentasyonu, koruyucu yumuşatma filtreleme tekniği önerilmiştir. Destek vektör makinesi (SVM) ve Yapay sinir ağı (ANN) ile kıyaslandığında geri yayılım sinir ağı (BPNN), Canny kenar segmentasyonu daha iyi doğruluk oranı elde edilmesine yardımcı olduğu görülmüştür. Bu teknik yapılan çalışma ile aynı zamanda Harris köşe tespit tekniği ile de kıyaslanmıştır. Geri yayılım sinir ağı (BPNN), Canny kenar segmentasyonu ve koruyucu yumuşatma filtreleme tekniklerinin bir araya getirilmesi ile oluşturulan yeni yapı %91 oranında doğruluk oranına sahip sonuçlar üretmiş ve diğer tekniklere kıyasla daha başarılı sonuçlar üretmiştir (Karimunnisa ve ark., 2020).

3. MATERYAL VE YÖNTEM

Travma sonrası kemik kırıklarında özellikle en fazla sakat kalma ve ölüm oranı bulunan kalça kırıklarını odağa alarak hem uzman hem pratisyen hekimler hem de derin öğrenme ve bilgisayarlı görünün en güçlü algoritmaları kullanılan bu çalışmanın materyal ve yöntem alanında; yapay zekâ kavramı, makine öğrenmesi kavramı derin öğrenme kavramı, yapay sinir ağları ve bilgisayarlı görü kavramı bunların uygulamaları ve yapıları ele alınarak irdelenmiştir.

3.1. Bilgisayarların Yükselişi ve Yapay Zekâ Kavramı

Yapılan çalışmalar, bilinen dünya tarihindeki en eski hesaplayıcının milattan önce 500'de abaküs olduğu kabul edilmektedir. 17 yy. başlarında ilk algoritma tablosunun kullanılması, ilerleyen yıllarda kayan cetvelin kullanılması sağlamış ve mekanik hesap makinesinin kullanılmaya başlanmasına öncülük etmiştir. Mekanik hesap makinesi kendisinin icadından üç yüzyıl sonra kullanılmaya başlanacak olan elektronik bilgisayarların atası niteliğindedir.

1821 yılında tekstil işçisi olan Ludd maliyetleri azaltan ve işgücüne ihtiyacı ortadan kaldıran makinelere karşı başkaldırmış ve günümüzde kullanılan Luddite teriminin ortaya çıkmasını sağlamıştır. 19. Yüzyıla gelindiğinde Babbage mekanik hesaplamaları gerçekleştiren Analitik Motor isimli bir mekanik hesaplayıcıyı tamamlamış, 1854 yılında Gorge Boole ise elektronik bilgisayarların bugünkü iletişimini sağlayan Boolean cebiri olarak bilinen iki sistemi oluşturmuştur (Çakır ve Eryılmaz, 2014).

20.yy. başlarında elektriğinin insan hayatında vazgeçilmez bir şekilde kullanılmaya başlanması ilk bilgisayarların kalbi olan mekanik sistemlere hayat vermiştir. Günümüzden tam 82 yıl önce 1941 yılında Kondrad Zuze, Z3 ismine sahip işlem yapma kabiliyetini elektrik motorlardan alan mekanik bir bilgisayar geliştirmiştir. Söz konusu bilgisayar bir programın kontrolünde hareket eden ilk bilgisayardır. Sonrasında 1943 yılında Alan Turing tarafından Colossus anında elektronik tüpleri kullanmakta olan bir bilgisayar geliştirmiştir. Takip eden yıllarda 23 haneli iki sayıyı

4,5 saniyede çarpabilen MARK I isimli bilgisayar Harvard Üniversitesinde geliştirilmiştir. Sonrasında 1946'da 30 tona yakın ağırlığı olan ENIAC ve yine aynı yıl bir programı bellekte tutabilme kapasitesine sahip EDVAC, 1951'de ilk ticari amaca sahip UNIVAC 1 geliştirilmiştir. 1968 yılında Intel firmasının kurulması bilgisayarlarda mikro işlemci çağını başlatmış, 1971 yılı geldiğinde sadece 4 bit işlem kapasitesine sahip ilk mikro işlemci üretilmiştir (Çakır ve Eryılmaz, 2014).

21. yüzyılın ilk senesine geldiğimizde bilinen insanlık tarihi boyunca üretilen veri toplamı kadar veri sadece bir yıl içerisinde üretilmiş; değişen, gelişen ve önlenemez bir yükselişe sahip, çok yüksek işlemcili bilgisayarlar sadece birkaç yıl içerisinde insanların cebine sığabilecek boyuta erişmiştir. Atalarına göre oldukça yüksek işlemcili bu bilgisayarlar yapay zekanın hiçbir zaman aralığında olmadığı kadar mükemmel bir gelişme ortamına sahip olmasına zemin hazırlamıştır (Özçelik ve ark., 2022).

1950'li yıllarda düşünebilen makineler yahut makinelerin düşünebileceği olgusu Alan Turing tarafından ortaya atılmasına karşın, bundan yüzyıllar önce Yunanlı düşünürler otomasyondan bahsetmiştir. Bunlara en güzel örnek İlyada destanında Homeros, kendi kendine hareket edebilen, ismine ise üçayaklı denen sandalyeleri kaleme almıştır. Milattan Önce 384-322 yılları arasında yaşayan Aristoteles de otomasyonu hayal etmiş, üzerine epeyi düşünmüş lakin bunun mümkün olmadığı sonucuna varmıştır.

Yapay zekâ, bir düşünce ve bir arayışla başlamış olsa da aslen insan zekasının öne çıkan ve insan zekasına özgü olan nitelikler, çıktılar ve davranış özelliklerinin bir sistemde gözlemlenmesi hali olarak nitelendirilebilir. Bir sistemin yapay zekaya sahip olup olmadığını anlamak için günümüzde halen Turing testlerine başvurulmaktadır. Turing testlerinde karşılıklı sorulara alınan yanıtlarda, karşımızdakinin makine yahut insan olduğu ayrımın yapılamadığı durumlarda, etkileşim kurulan taraf eğer bir bilgisayar yahut mekanik sistem ise yapay zekaya sahip olduğu söylenmektedir (Lattimore, 1962; Russel ve Norvig, 2009; Turing, 1950).

Şekil 3.1. Yapay Zeka'nın alt disiplinlerini gösteren Venn Şeması.

Yapay zekâ kavramı ve teriminin doğuşunun sıfır noktası aslında 1956 yılında bir grup bilim insanının Dartmouth College'de makinelerin insan düşünme süreçlerini simüle edebileceği fikrini tartışması ile ortaya çıkmıştır. Günümüze gelene kadar olan süreçte insan zekasını ile yapay zekayı kıyaslamak isteyen bilim insanları iki tarafında şampiyonları olarak nitelendilebilecekleri; santranç gibi, go oyunu gibi komplike oyunlarda rakipleri karşı karşıya getirmiştir (Russel ve Norvig, 2009). 2017 yılına geldiğimizde bu yarışmanın mutlak galibinin yapay zeka olacağı yıkıcı bir şekilde ortaya çıkmıştır.

3.2. Yapay Zekâ Çeşitleri

Günümüzde sadece tek bir tıklama ile erişebildiğimiz çok çeşitli alanlarda artık insanın karar vermesine ve insana destek olma noktasında önemli yer tutan yapay zekayı yetkinliğine göre sınıflandırmak mümkündür.

Zayıf yapay zekâ (Dar yapay zekâ): kendisinden istenilen görevleri gerçekleştirebilen lakin bu icra kapasitesi uzmanlaştığı tek bir görevlerle sınırlı kalan, ötesine geçemeyen yapay zeka türünü ifade etmektedir. Yetkinliğine göre bu kategoride sınıflandırılan

yapay zekalar gerçekleştireceği göreve göre tasarlanır ve öncelikli olarak programlandığı şekilde işlevselliğini sürdürebilmektedirler. Örnek olarak satrançta uzmanlaşmış dünya satranç şampiyonunu yenen yapay zeka gösterilerbilir (Strelkova, 2017).

Zayıf yapay zekaların temel özellikleri incelendiğinde görevleri belirli bir yetenek alanlarında sınıflıdır. Belli başlı bir görevi en iyi şekilde ortaya koymak üzere tasarlanmıştır. Bu alanda performansı düşünülenin çok üzerine çıkabilir. Buna karşın sınırlı bağlamda çalışabilmektedir. Veriye bağımlı olan bir yapıları vardır ve mutlak surette verinin bir yapı yahut kişi tarafından etiketlenmesine ihtiyaç duymaktadırlar. Etiketlenecek yahut sisteme eğitim için verilecek verinin eksik olması yahut yanlış olması zayıf yapay zekaların performansını doğrudan olumsuz yönde etkilemektedir. Esnek bir yapıya sahip olmadıklarından uzmanlaştığı alan dışında herhangi bir çıktı yahut sonuç üretemezler, farklılaşan problemlere cevap veremezler. Yapay dar zekalar kendilerinin varlıklarının yahut çevrelerindeki olguların farkında olmayıp herhangi bir düşünce duygu yahut bir bilince sahip değildirler. Dar yapay zekâ dediğimiz kavram esasen günümüzde aktif şekilde insanların kullanımına sunulmuş olan yapay zekâ çeşididir.

Güçlü yapay zekâ (Genel yapay zekâ): yetkinlikleri insan düşünme yetkinlikleri ile benzeşen, çok yönlü çıktı üretme kabiliyetine sahip yapay zekalar olarak adlandırılabilirler. Genel yapay zekâ sadece kendisine adreslenmiş ve bildiği görevler ile ilgili çıktı üretmek ile kalmaz, farklı görevlerin birbiri ile bağlantılarını anlamlandırma, ilişki kurma ve mantıksal yeni sonuçlar üretebilir (Özdemir ve ark., 2019; Strelkova, 2017).

Güçlü yapay zekâların temel özellikleri incelendiğinde, öncelikle çok yönlü ve genel yeteneklere sahiptir. Bu özelliği itibariyle belli başlı görevlere çözüm üretmez, farklılaşan görevleri anlayabilir, algılayabilir ve çözebilir. Çok yönlü problem çözme yeteneğine sahiptir. Öğrenme ve gelişme yeteneğine sahiptir. Her deneyimi yeni bir öğrenme olarak kayıt altına alınır. Özellikle daha önce hiç karşısına gelmeyen problemler ile geçmiş problemler arasında ilişki kurabildiğinden hiç olmayan çözümler üretir. Bağlamı dışında düşünme ve sonuç üretme yetkinliğine sahiptir. Çoklu senaryolara uyumlanabilir. Güçlü yapay zekaların öz farkındalıkları söz konusudur. Duygu ve düşünce gibi özelliklere sahip olabilir. Bu bağlamda tepki verebilirler. Şu an Dünya Ekonomik Forumunun her yıl yayınladığı geleceğin

meslekleri raporunda yer alan, insandan beklenen on yetkinlikte listenin başında olan yaratıcılık ve problem çözme yetkinliğine sahiptirler. Sadece öğrendikleri ile sonuç üretmezler aynı zamanda yeni bilgiler ile yeni ve farklı çözümler üretebilir. Söz konusu yapay zekaya örnek vermek henüz mümkün değildir. Araştırmacılar kuantum bilgisayarlarının gelişimi ile artık yapay genel zekaya hiç olmadığımız kadar yakın olduğumuzu belirtmektedir.

Süper yapay zekâ: insan zekasının sınırlarının çok ötesinde düşünme yeteneğine sahip bir yapay zekâ çeşidi olarak kabul edilmektedir. Karmaşık problemlere hızlı ve net bir çözüm üreten aynı zamanda duygu ve çok yönlülük alanlarında insan kapasitesinin çok üstüne çıkabilmektedir (Özdemir ve ark., 2019; Strelkova, 2017).

Süper yapay zekanın temel özellikleri arasında ilk olarak üstün zekâ kapasitesi yer almaktadır. Araştırmacı ve bilim insanlarının yıllar sürerek ortaya çıkardığı sonuçları hızlı ve kesin düşünme yapısı ile çok kısa sürede sonuç üretmesi beklenmektedir. Kapsamlı anlama ve öğrenme yeteneğine sahip olan bu yapay zeka çok çeşitli konuları derinlemesine anlar verileri hızla analiz ederek bilgiye dönüştürür. Yaratıcılık ve inovasyon alanında hiçbir yapay zekanın olmadığı bir seviyededir ve sanat eserleri ortaya çıkartabilir. Özgün tasarımlar ile kalıcı çözümler üretebilir. İletişim ve duygusal algılama alanında oldukça ileridir. İnsan tepki ve mimikleri değerlendirebilir duygusal ifadeleri yorumlayabilir. Yapay süper zekaların bir diğer ve en farklılaşan özelikleri kendi kendilerini geliştirme yeteneğine sahiptir ve bu şekilde kendi kendine yeni yetkinlikler kazanabilir.

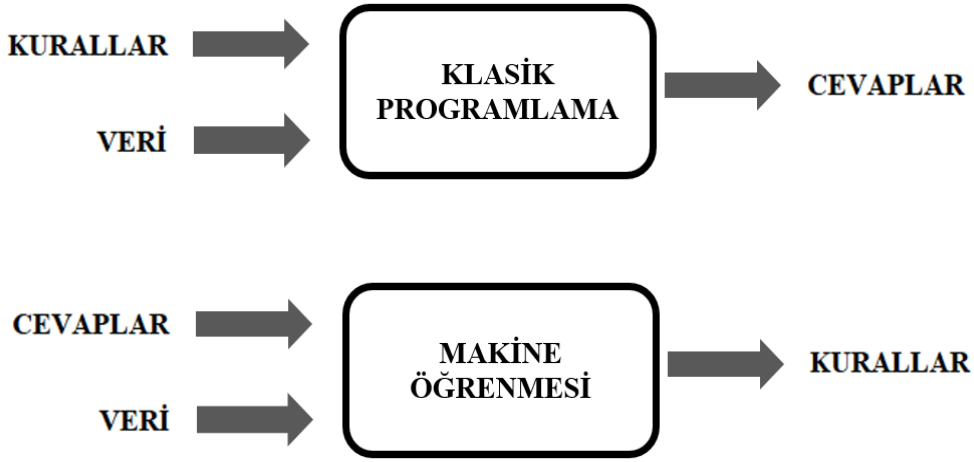
Şu an mevcut teknoloji seviyesi itibariyle insanoğlu dar ve genel yapay zekalara yoğunlaşmıştır. Gelecekte teknolojide yaşanılacak atılımlar ile ancak süper yapay zekaların oluşturulabileceğinden bahsedilebilir.

3.3. Makine Öğrenmesi Kavramı ve Temelleri

Makine öğrenmesi kavramı yapay zeka kavramını bir alt sınıfı içerisinde yer almaktadır. Makine öğrenmesi sistemlerin programlanmamış sonuçları öğrenmesi yeteneği olarak nitelendirilmektedir. İlk Arthur Samuel tarafından yapay zeka kavramından üç yıl sonra, 1959 yılında ortaya atılmıştır. Makine öğrenmesinin temelinde klasik programlamdan farklı olarak kod yazmak, kurallar ve veriyi sisteme vererek sonuçları almak yerine, sonuç ve veri sisteme verilerek makine öğrenmesi algoritması aracılığı ile kural setleri ortaya çıkmaktadır. Makine öğrenmesinin ortaya

çıkışında, makinelerin kendi kendilerine bir işi nasıl yapacaklarını öğrenmelerinin mümkün olup olmadığı sorusu etkili olmuştur (Karunakaran, 2018; Chollet, 2017).

Makine öğrenmesi yine insanoğlunun öğrenme süreçleri takip edilerek geliştirilmiştir. Ufak bir cocuk bir evin dışarıdan neye benzediğini anlamak için her bir duvarı ve evin dış yüzeyindeki her bir ayrıntıya odaklanmaz. Boyutuna, çatısına, pencere ve kapısına yani genel hatlarına odaklanır, bu şekilde bir binayı bir arabadan ayırt eder. Makine öğrenmesinde bu durum sınıflandırma olarak adlandırılmaktadır (Murphy, 2012).



Şekil 3.2. Makine öğrenmesi işleyişi ile klasik programlama (Chollet, 2017).

Klasik programlama ele alındığında, cevapların elde edilmesi için temelde sisteme kurallar ve verinin beslenmesi gerekliliği aşikardır. Makine öğrenmesinde ise eldeki veriden ve cevaplardan beslenen sistem, öğrenebilen algoritmalar ile kural setlerini üretmektedir.

3.3.1. Öğrenme algoritmaları

Eldeki veriler ile algoritmaların birleşimi sonrası öğrenme gerçekleştirerek daha iyi sonuçlar elde etmeyi başaran yapılar makine öğrenmesi algoritması olarak adlandırılmaktadır. Bir sistemde performans ölçütü, görevler ve deneyimleri düşünelim. Eğer sistemin performansı deneyimler arttıkça artıyorsa bu sistem öğrenebilen bir sistemdir. Sabit programlarca çözülemeyecek kadar komplike olan görevlere yanıt bulmada en etkili yöntemlerden biri makine öğrenmesi yöntemleridir. Bir örneğin nasıl işlenmesi gerektiğine göre makine öğrenmesi sistemlerinin görevleri tanımlanmaktadır. Makine öğrenmesi aracılığı ile çözülen en yaygın görevler; Sınıflandırma, eksik girdi ile sınıflandırma, bağlanım, transkripsiyon, makine tercümesi, yapılandırılmış çıktı, anomali saptama, sentezleme ve örneklendirme, eksik

ögelerin tamamlanması, gürültüsüzleştirme, yoğunluk kestirimi ve olasılık kütle fonksiyonu olarak gruplandırılabilmektedir.

Bir öğrenme algoritması tasarlanırken diğer bir önemli nokta performans ölçütünün belirlenmesidir. Görevlendirmelerin içerisinde sınıflandırma söz konusu ise genelde doğruluk modelin performans ölçütü olmaktadır. Makine öğrenmesi algoritmaların performans ölçütünde özellikle daha önce hiç karşılaşmadığı vakalar noktasında nasıl çalıştığını gözlemlemek önemlidir. Bu nedenle eğitim ve değerlendirme veri seti haricince bağımsız bir veri seti olan test veri seti kullanılmaktadır (Mitchell, 1997).

Yine makine öğrenmesi sisteminde yer alması gereken kısım deneyimdir. Makine öğrenmesi algoritmalarında öğrenme adımı gerçekleşmekteyken denetimli öğrenme ve denetimsiz öğrenme başlıklarının altına ikiye ayrıştırılabilir. Denetimli öğrenimde özniteliklerin tanımlandığı bir veri kümesi deneyimlenirken, denetimsiz öğrenmede bu durum söz konusu değildir.

Sonuç olarak bir makine öğrenmesi algoritması çalışılacaksa temel bileşenlerinde mutlak surette Görev (T), Performans Ölçütü (P) ve Denetim (E) parçaları yer almalıdır.

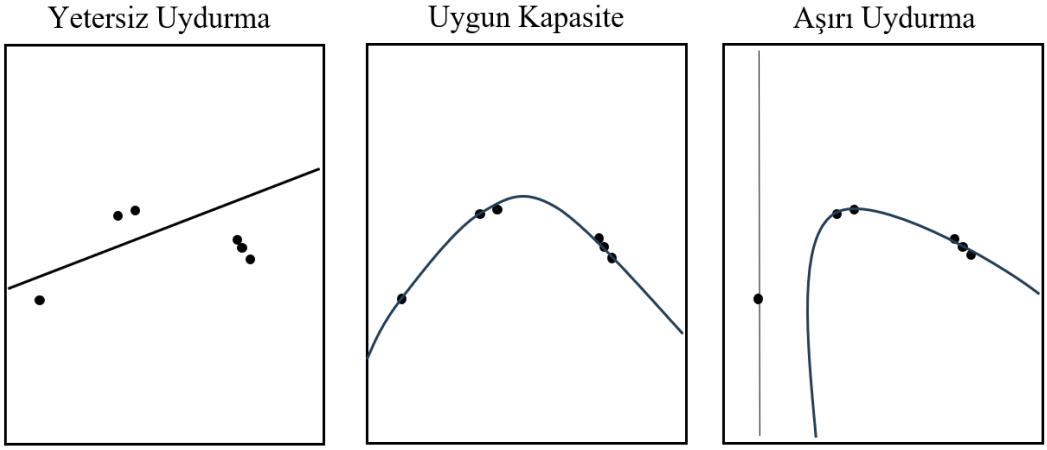
3.3.2. Kapasite, aşırı uyum (overfitting), eksik uyum (underfitting)

Makine öğrenmesinin temel zorluğu algoritmanın eğitimi sırasında kullanılan eğitim ve doğrulama veri setlerinin haricinde, özgün veri setlerinde ve özgün incelenecek yapılarda da beklenen performansı ortaya koymaları beklenir. Daha önce makine öğrenmesi algoritması tarafından incelemeye tabi tutulmamış girdilerdeki yüksek performans sergileme yeteneği genelleştirme olarak adlandırılmaktadır.

Makine öğrenmesinde söz konusu algoritmalardan esas beklenti, eğitim ve doğrulama veri setlerinde ortaya çıkan yanılgılardan oluşan eğitim hatası yerine, yeni girdilerde ortaya çıkan genelleştirme hatası veyahut diğer adlandırması ile bağımsız test veri setinde ortaya çıkan test hatasını minimize etmektir. Bir makine öğrenmesi algoritmasının esas performansı belirleyen faktörler; öncelikli olarak eğitim hatasının minimizasyonu akabinde eğitim hatası ve test hatası arasındaki farkın minimize edebilme kapasitesidir. Bu iki yetenek ölçütü iki temel engelleyiciye işaret etmektedir. Bu engelleyicilerden ilki yetersiz uydurma, modelin eğitimi sırasında eğitim girdilerine göre hedeflendiği kadar düşük hata oranı sunamadığında ortaya

çıkmaktadır. Aşırı uydurma ise eğitim hataları ve test hataları ele alındığında iki grup kıyaslandığında fark oldukça fazla olduğunda ortaya çıkmaktadır.

Model kapasitesi, eğitilen makine öğrenmesi modelinin çok farklı fonksiyonları modelleyebilme kapasitesine denilmektedir. Yukarıda ifade edilen iki temel zorluk olan aşırı uydurma ve yetersiz uydurma kavramlarından kurtulması için modelin kapasitesinde ayarlamaya gidilmesi en başarılı yöntem olacaktır. Bir modelin kapasitesi yetersiz ise karmaşık görevleri başarı ile ihtiva edemez (Goodfellow ve ark., 2016).



Şekil 3.3. Makine öğrenmesi yetersiz uydurma, uygun kapasite ve aşırı uydurma örnekleri.

3.3.3. Hiper parametreler ve doğrulama kümeleri

Hiper parametreler makine öğrenmesi algoritmalarının yönelim ve davranışlarının ayarlanmasını ve kontrol edilmesini sağlayan ayarlar bütünüdür. Eğer parametreler eğitimin gerçekleştirildiği veri grubu içerisinden öğrenilmeye çalışılırsa hiper parametreler olabilecek en yüksek kapasiteyi seçme eğilimi gösterir bu durumda model aşırı uydurma durumuna geçecektir (Goodfellow ve ark., 2016).

Bu gibi eğitim verisinden alınacak ve değerlendirme tabi tutulacak değerler probleme neden olacaktır. Bu sorunu çözmek için alınabilecek temel aksiyon, algoritmanın eğitim sürecinde içeriğinde yer almayan özgün örneklerden, bir araya gelen bir test kümesinin kullanılmasıdır. Test kümeleri model yahut hiper parametreler noktasında bir karara varılması noktasında kullanılmaz. Bu sebeptendir ki hiçbir test kümesi üyesi, doğrulama kümesinde yer alamaz. Dolayısı ile doğrulama kümeleri daima eğitim kümesinin içerisinden oluşturulmaktadır. Özetle eğitim kümesi her zaman iki alt kümeden oluşur bunlardan ilki parametrelerin oluşturulduğu küme diğeri ise

doğrulama kümesidir. Hiper parametrelerin seçimi doğrulama alt kümesi aracılığı ile gerçekleşmiş olur. Genellikle yapılan çalışmalarda eğitim verileri %80 eğitimin gerçekleştirilmesi %20'si de doğrulamanın gerçekleşmesi amacıyla kullanılmaktadır.

3.3.4. Tahminleyiciler, bias ve varyans

Makine öğrenmesi algoritmalarının esas amacı yalnızca eğitimde kullanılan küme üzerinde çözümleme yapmak değil aynı zamanda çözümün genele yayılması sureti ile çözüm üretmesidir. Nokta tahminleme ise odağa alınan bir çıktıda en başarılı öngörüyü yaparak çözüm üretmesi halidir.

Bias kavramı özellikle önyargı, yanılma ve sapma manasına gelmektedir. Bu kavramı herhangi biri ile ifade etmek hatalı olacaktır. Biri bias'e sahip biri de sahip olmayan iki adet tahminleyicimiz olduğunu varsayarsak, genellikle bias'e sahip olmaya tahminleyiciler daha çok tercih edilir lakin her zaman en iyi performansa sahip tahminleyiciler değildirler.

Tahminleyicilerin dikkate değer farklı bir özelliği ise tahminleyicinin veri örneğinin bir fonksiyon olarak varyansının ne kadar değişkenlik göstereceğine karşı olan beklentidir. Varyans kavramının karekökünе standart hata denilmektedir. Bir tahminleyicinin varyansı yahut standart hatası söz konusu algoritma hakkında önemli boyutta bilgi vermektedir. Algoritmaya yeni sonuçlar ürettirdiğimizde tahminlemenin nasıl bir değer alacağı yönünde varyans ölçütleri bilgi sunmaktadır.

Bias ve varyans bir tahminleyicideki farklılık gösteren iki hata kaynağını ölçmede kullanılırlar. Bias kavramı bir parametrenin veyahut bir fonksiyonun almış olduğu asıl değerden sapma miktarının beklentisini ölçerken, varyans ise tahminden herhangi bir veri örneklemesinin ortaya koyabileceği sapmayı ölçmektedir (Goodfellow ve ark., 2016).

3.3.5. Denetimli öğrenme algoritmaları

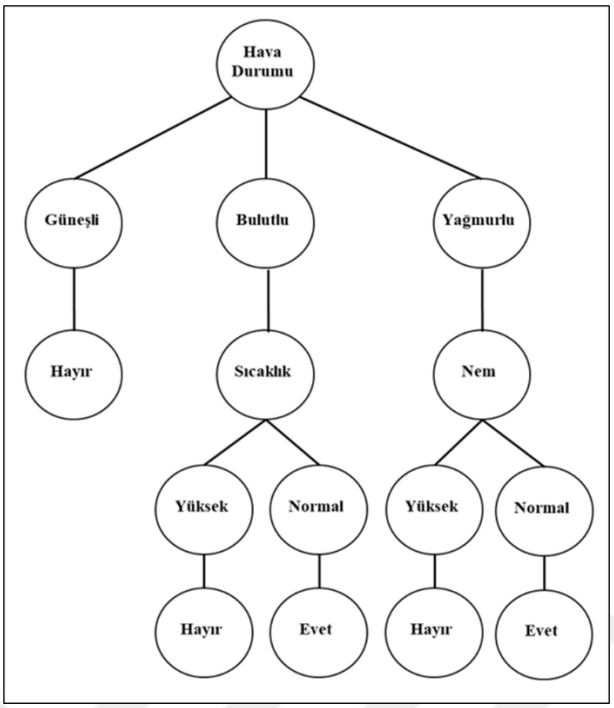
Eğitim örnekleri ve çıktıları verildiğinde girdi ile çıktı arasındaki ilişkiyi kurmayı öğrenen algoritmalara denetimli öğrenme algoritmaları denir. Birçok uygulamada çıktıların otomatik olarak bir araya getirilmesi güç olabilir. Çıktılar bu sebeple bir denetleyici aracılığı ile sağlanması gerekli olabilir. Otomatik toplanması halinde dahi denetimli öğrenme geçerlidir. Denetimli öğrenme algoritmalarına örmek verilmek istenirse; olasılıksal denetimli öğrenme, destek vektör makinesi, en yakın komşu bağlanımı, karar ağaçları verilebilir.

Olasılıksal denetimli öğrenmeye örneklendirmek istediğimizde farklı olasılıklara sahip dağlım ailelerini tanımlayarak doğrusal bir yapıya sahip sınıflandırma vakası oluşturabilmek mümkündür. Örneğin iki sınıfımızın olduğu varsayımı ile sınıflar 0 ve 1 olarak düzenlendiğinde sadece bir sınıfın olasılığının ifade edilmesi kâfi olacaktır. İki değerin toplamı 1 olacağı için 1 numaralı sınıfın olasılığı 0 numaralı sınıfın olasılığını belirlemektedir.

Denetimli öğrenme algoritmalarına destek vektör makinesi (SVM) örnek verilmek istenildiğinde en temel yaklaşımlardan biri olduğu ifade edilmektedir. (Boser ve ark., 1992; Cortes ve Vapnik, 1995). Destek vektör makinesi doğrusal fonksiyondan türemiştir. Destek vektör makinesinin bir diğer özelliği ise sadece sınıf kimliklerini ifade etmektedir olasılık değerlerini vermemektedir. $w^T x + b$ pozitif çıkması halinde destek vektör makinesi pozitif sınıfı öngörmekte iken aynı denklem negatif çıkması halinde negatif sınıfı öngörmektedir.

Denetimli öğrenme algoritmalarına bir diğer örnek K-en yakın komşu algoritmasıdır. Özellikle bağlanım ve sınıflandırma alanlarında kullanılan bu algoritma parametre içermemekte olup fonksiyon kullanan bir algoritmadır. K-en yakın komşu algoritmasında tam olarak bir eğitim safhası ya da öğrenme aşaması yoktur. Bu durumun aksine test sırasında her yeni test girdisi için bir çıktı üretmek istendiğinde eğitim verisinde, test girdisi için k-en yakın komşusunu buluruz. Daha sonrasında eğitim kümesinde yer alan ilgili çıktı değerlerinin ortalaması elde edilir.

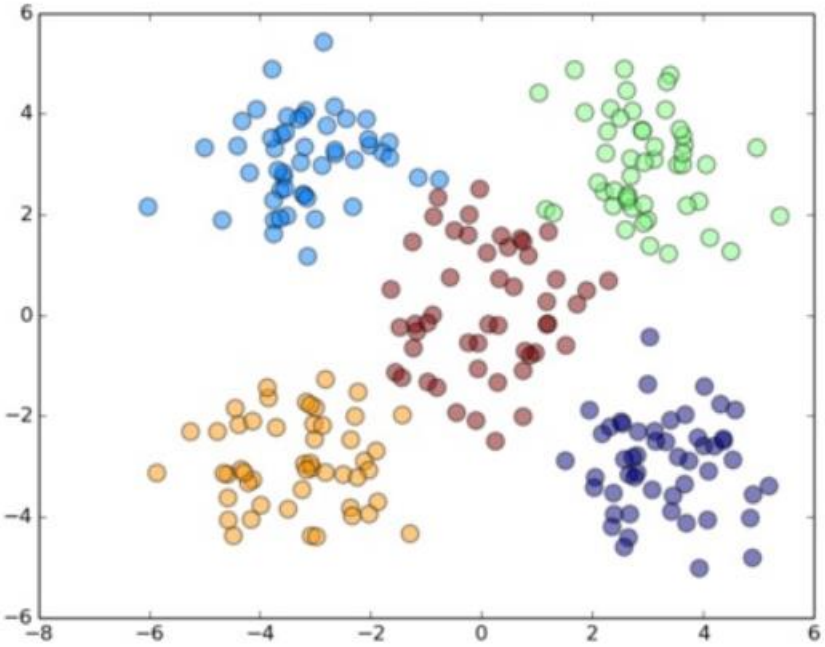
Bir diğer denetimli öğrenme algoritmalarına örnek karar ağaçlarıdır. Karar ağaçları girdileri boşluğunu bölgelere ayıran, sonrasında her bir bölge için ayrıma gidilecek parametreleri ayrıştıran bir öğrenme algoritmasıdır (Breiman ve ark., 1984). Karar ağacında yer alan her düğüm girdi boşluğunda bir bölge ile bağlantılıdır. İç düğüm ve bunun alt bileşenleri bu alanları alt alanlara bölmektedir.



Şekil 3.4. 65 Yaş üstü insanların yaz aylarında sokağa çıkması ile ilgili karar ağacı.

3.3.6. Gözetimsiz öğrenme

Gözetimsiz öğrenme verilerin bir etiket ile etiketlendirilmediği ya da belirli bir çıktı grubuna sahip olmaksızın kullanıldığı öğrenme şekli olarak nitelendirilebilir. Diğer bir değişle gözetimsiz öğrenme insan gücüne örneklerin etiketlenmesi noktasında bir desteğe ihtiyaç duymaz. Denetimsiz öğrenmenin ana görevi veriler için en iyi ifade etmenin yolunu bulmaktır. Gözetimsiz öğrenme yöntemlerine örnek olarak kümeleme, boyut indirgeme, anomali tespiti verilebilir. Kümele verileri aynı özellik gösteren gruplara ayıran bir yöntemdir buna verilebilecek en iyi örnek müşteri segmentasyonudur. Boyut indirgeme, veriler içerisinde yer alan boyut sayıları ya da fazla özellikleri azaltan bir yöntemdir. Çok yüksek hacme sahip verilerde veri görselleştirmenin yapılmasını sağlar. Anomali tespiti ise standart durumdan ayrışan hallerin ön plana çıkarılmasında kullanılmaktadır. Örnek olarak kredi kartı harcamalarındaki sahtekarlıklar verilebilir. Veri etiketlemek bir veri grubu için oldukça zor yahut maliyeti beklenenin oldukça üzerinde oluyorsa gözetimsiz öğrenme tercih edilebilir. Lakin bir denetimli öğrenme gibi net ve belirli sonuçlar üretmeyebilir çünkü net geri bilirim süreci olmaksızın çalışmaktadırlar (Goodfellow ve ark., 2016).



Şekil 3.5. Kümeleme analizi örneği.

3.3.7. Stokastik gradyan azalma

Stokastik gradyan azalımı bütün derin öğrenme yöntemlerde bulunan oldukça önemli bir algoritmadır. Makine öğrenmesi modelleri günümüzde gradyan azalma ile eğitildiğinde oldukça iyi çalıştığı bilinmektedir. Stokastik gradyan azalımı çok büyük veriye sahip kümelerde büyük doğrusal modellerin eğitiminin ana yoludur. Stokastik gradyan azalmada her bir eğitim iterasyonunda rasgele seçilen örnek verileri baz alarak bir eğim hesaplamakta, hesaplanan bu eğim ile ağırlık güncellemesi yapılmaktadır. Bu hesaplamanın bir diğer avantajı maliyeti düşürmesi ve eğitim süreleri optimize ederek aşağıya çekmesidir. Stokastik gradyan azalma işleyişi sırasında rastlantısal olarak bir veri seçilmekte, söz konusu veri baz alınarak eğimin hesaplanması ardından hesabı yapılan eğim kullanılarak ağırlıkların güncellenmesi söz konusudur (Goodfellow ve ark., 2016).

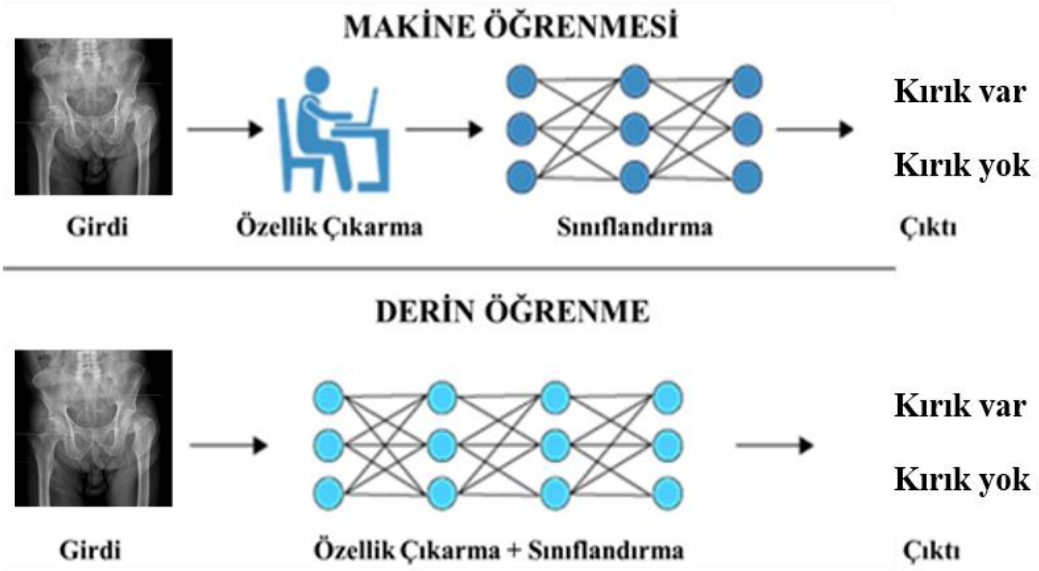
3.4. Derin Öğrenme Kavramı

Derin öğrenme verilen girdi üzerinden hem özellik çıkarma hem de sınıflandırma işlemlerinin kendisi tarafından yapıldığı, sonrasında ise yine kendisi tarafından bir çıktı üretilebilmekte olan bir makine öğrenmesi tekniğidir. Derin öğrenmede metinden başlamak üzere, ses ve görüntü formatında veriler sisteme tanımlatabilir. Derin öğrenmede yapay sinir ağları gibi büyük oldukça karmaşık yapılar kullanılarak veri

analizi gerçekleştirilmektedir. Derin öğrenme algoritmalarının karmaşık görevler ile ilgili sonuç üretme yetkinliği oldukça yüksektir.

Makine öğrenmesinden derin öğrenmeyi ayıran esas temel farklılık, derin öğrenme kavramı makine öğrenmesinin aksine tek katman yerine aynı anda birçok katmanda işlem yapma yeteneğine sahiptir. Makine öğrenmesinde özellik çıkarma adımı bir uzman tarafından yapılması gerekli iken, derin öğrenmede özellik çıkarımı yapay sinir ağları ile derin öğrenme modelinin kendisinin içerisinde çıkarılmaktadır (Karunakaran, 2018).

Derin öğrenme coğunlukla birden fazla katmana sahip yapay sinir ağları aracılığı ile çalışmaktadır. Bahsi geçen yapay sinir ağları girdi oluşturan verileri katmanlar halinde işlemekte özelliklerini ve desenlerini ortaya koymaktadır. Her bir yapay sinir ağı katmanı girdi olarak alınan veriyi yüksek seviyeye sahip özelliklere evirmektedir. Söz konusu derin öğrenme modelleri her aşamada ağırlık parametrelerini optimize eder ve bu sayede öğrenmeyi gerçekleştirir.



Şekil 3.6. Makine öğrenmesi ve derin öğrenme modellerinin çalışma sistematiği.

3.4.1. Parametre norm cezaları

Derin öğrenme modelleri içerisinde yer almakta olan parametre norm cezaları yapay sinir ağının ağırlıklarını ve parametrelerini regüle etmek için kullanılmakta olan bir düzenleme tekniğidir. Her bir işlem adımı sonrasında cezalar ile parametrelerin ve ağırlıkların yüksek miktarlarda artmasına engel olmakta aynı zamanda karmaşıklığı

baskılamaktadır. Modelin daha iyi sonuçlar üretmesinde önemli derecede yardım etmektedir.

En yaygın olarak kullanılan parametre norm cezaları: L1 Norm Cezası, L2 Norm Cezası ve Elastic Net'dir. L1 Norm Cezası, ağırlıkların mutlak değerinin toplamı üzerinden bir cezalandırma yöntemidir. Bu sayede ağırlıkları sıfıra yakınsayarak düzelmelerini sağlamaktadır. L2 Norm Cezası ise ağırlıkların karelerinin toplamı üzerinden cezalandırır bu da ağırlıkların birbirlerine yakınsamasını sağlamaktadır. Elasctic Net, hem L1 hem L2 norm cezalarının kombine edilmesi ile işletilen bir yöntemdir. Hem ağırlıklar büyümez hem de birbirlerinden uzaklaşmazlar (Goodfellow ve ark., 2016).

3.4.2. Gürültü sağlamlığı

Bishop 1995 yılında yapmış olduğu çalışmalarda çeşitli modellerde sonsuz küçüklükte bir değişikliğe sahip olan bir gürültü eğer modelin girdisine eklenirse aynı parametrelere norm cezası eklenmesi ile aynı özellik gösterdiğini ifade etmiştir. Bunun bir sonucu olarak eğer girdiye gürültü eklersek parametreleri küçültmekten daha fazla etkili olmaktadır (Bishop, 1995).

Modellerin düzenlileştirilmeleri ve regüle edilebilmeleri için gürültüyü ağırlıklandırmalarda da kullanılabiliriz. Özellikle yinelemeli sinir ağlarında kullanılabilirliği söz konusudur (Graves, 2011).

3.4.3. Yarı yardımlı öğrenme

Yarı yardımlı öğrenme (semi-supervised learning) hem etiketlenmemiş hem de etiketlenmiş verilerin kullanılması sonucu olarak modelin öğrenmesini sağlamayı amaçlar. Etiketlenmiş olan veriler giriş örneklerinin doğru çıktı değerleri ile eşleştiği verileri ifade etmekteyken, etiketlenmemiş olan veriler ise çıktı değerleri olmaksızın girdi verilerini içermektedir. Yarı yardımcı öğrenme az sayıda etiketlenmiş veri varsa ve etiketlenmesi zor ve maliyetli ise yarı yardımlı öğrenme tercih edilmektedir (Chapelle ve ark., 2003).

3.4.4. Çok görevli öğrenme

Çok görevli öğrenme, çeşitli görevlerde çıkan örneklerin biriktirilmek sureti ile genelleştirme yapmak için iyileştirmeler yapmanın yoludur. Çok görevli öğrenmede aynı modelin çok problemi çözmek için kullanılmasını amaçlamaktadır. Normalde her

bir görev için ayrı derin öğrenme algoritması eğitmek yerine tek model birden fazla işim yapabilmektedir. Çok görevli öğrenmede birbiri ile ilişkili yahut ortak özelliği bulunan bir görev söz konusu ile yapılabilir. Çeşitli görevleri için uygun miktarda veri yoksa bu görevlerin aynı modeller ile birleştirmesi daha başarılı sonuç elde etmeyi mümkün kılacaktır (Caruana, 1993).

3.4.5. Erken durdurma

Büyük modellerin eğitimi sırasında eğitim hata grafiğinin zamanla önce çok hızlı sonra yavaş yavaş azaldığını görürüz lakin bir süre sonra tekrar artmaya başlar. Demek ki öğrenim sırasında hatanın en düşük olduğu ağırlıkların olduğu ana geri dönersek en son iterasyondaki değere göre o andaki ağırlıkların daha iyi sonuç vermesini bekleriz. Bu nedenle sağlama kümesindeki hata oranı her iyileştiğinde bir kopya alırız eğitim bittiği anda son elde edilen değil kaydedilen en iyi parametreye geri çevrilir. Bu stratejiye erken durdurma denmektedir. Derin öğrenme modellerinde en çok kullanılan düzenlileştirme yöntemi erken durdurmadır (Goodfellow ve ark., 2016).

3.4.6. Çekişmeli öğrenme

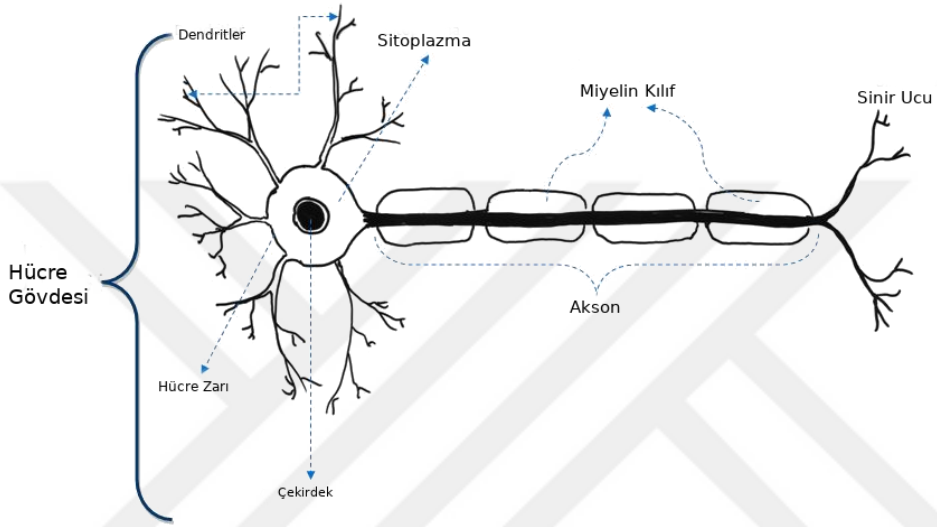
Çekişmeli öğrenme yine derin öğrenme modellerinde kullanılan bir yaklaşımdır. Bir modelin daha iyi özellikler göstermesi ve daha gerçekçi sonuçlar elde etmesini sağlamaktadır. Çekişmeli öğrenme iki ana bileşeni vardır bunlardan ilki üreteç ikincisi ise bir ayırt edici ya da eleştiricidir. Bu bileşenler arasında çekişme ve yarışma yaşanmaktadır. Üreteç rasgele olarak gürültü ya da başlangıç verilerinden gerçekçi veriler oluştururken, ayırt edici ise gerçek veriler ile üretilen verileri ayırt etme görevini gerçekleştirir. Üreteç, gerçekçi veriler üretmeye çalışırken, ayırt edici, gerçek ve üretilmiş verileri ayırt etme görevini yerine getirir bunun bir sonucu olarak üreteç ve ayırt edici arasında bir çekişme başlar. Üreteç, ayırt ediciyi aldatmak ve üretilmiş verileri gerçek gibi göstermek isterken; ayırt edici, gerçek ve üretilmiş verileri mümkün olduğunca iyi ayırt etmeye çalışır. Bu çekişme sonucunda, üreteç zamanla daha iyi veriler üretebilir hale gelmekte buna karşılık ayırt edici ise daha iyi ayırt etme yeteneğine sahip olmaktadır (Szegedy ve ark., 2014).

3.5. Yapay Sinir Ağları

Yapay sinir ağları temelde insanın beyninde bulunan sinir sistemi ve düşünme yapısının matematiksel olarak simüle edilmesi olarak ifade edilmektedir. Yapay sinir

ağları da aynı insan beyninde yer aldığı üzere nöron olarak adlandırımış olan temel işlem yapan küçük birimlerden bir araya gelir.

Biyolojik bir nöron hücresi incelendiğinde dentritlerden bilgi alınmakta bu alan giriş alanı olarak kabul edilmekte sonrasında hücre çekirdeğine gelen bilgi işlenmekte akabinde bir sonraki hücreye iletilmesi noktasında karar verilmekte, verilen karara göre akson uçları ile diğer nöron hücrelerine bilgi çıkış ile transfer edilmektedir (Zhang ve ark., 2018).



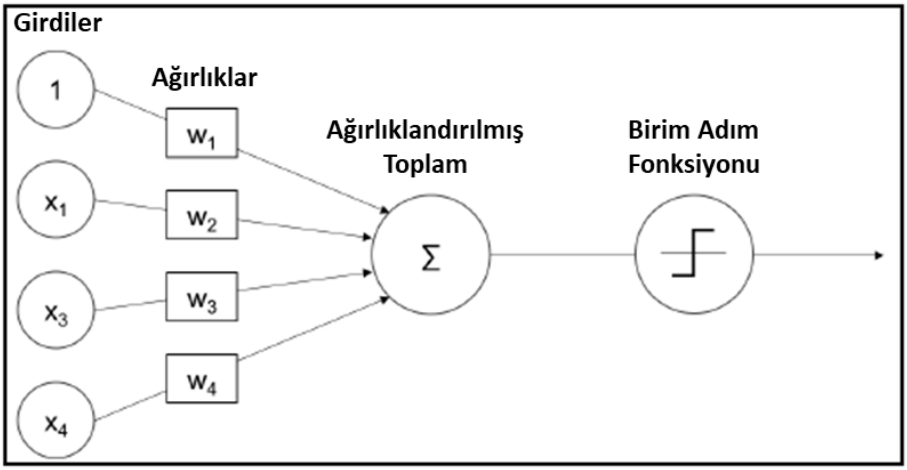
Şekil 3.7. İnsan vücudunda yer alan nöron hücresi (Zhang ve ark., 2018).

3.5.1. Yapay sinir ağlarının yapısı

İnsan sinir hücresinin bir modeli olan yapay sinir ağlarında yine benzer bir mantıkla hareket edilmiştir. Çevreden gelen girdiler x_i olarak alınmakta, tetiklenmesi noktasında ağırlık belirleme hususu için w_i değerleri ile ağırlıklandırılmakta, her biri sıralı olarak toplanıp sonrasında ise bias (yanılma, sapma, önyargı) değeri olan b değeri eklenmektedir. Buna göre bir algılaycının formülü aşağıdaki şekilde ifade edilmiştir.

$$z = b + \sum_{i=1}^{n} x_i w_i \tag{3.1}$$

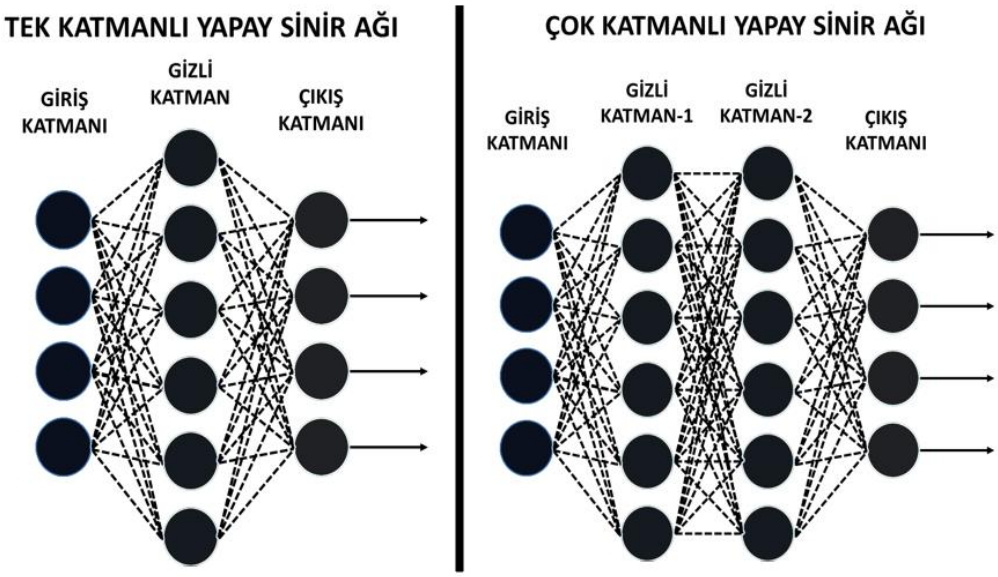
Söz konusu formül 3.1'de ifade edilen algılayıcının (perceptron) şema gösterimi aşağıda ifade edilmiştir.



Şekil 3.8. Algılayıcıların şematik gösterimi (Shanmugamani, 2018).

Tek başına bir algılayıcı sadece basit fonksiyonlar ile ilgili işlem yapabilir elde ettiği ağırlığı bu yönde tutabilir. Ağırlıkların hesaplanılması ve bir değer aklında tutması süreci eğitimdir. Ağırlıklandırılmış toplam yani algılayıcının çıktısı birim adım fonksiyonundan geçirilmektedir.

Yapay sinir ağları ise her bir algılayıcı nöronların bir araya gelmesi ve çoklu katmanların eklenmesi ile oluşmaktadır.



Şekil 3.9. Tek ve çok katmanlı sinir ağları modelleri (Nielsen, 2015).

Tek ve çok katmanlı sinir ağlarında gözüktüğü üzere her katmanın çıktıda doğrudan etkisi söz konusudur. Her bir nöron için doğrudan ya da dolaylı etkisi olduğu söylenebilir. Aynı hizada bulunan katmanlardaki algılayıcıların arasında şekil 3.8'de de görüldüğü üzere ilişki bulunmamaktadır. Değer akışı katmanlar arasındadır.

3.5.2. Aktivasyon fonksiyonları

Aktivasyon fonksiyonu, her bir nöronun ürettiği çıktıyı hesaplamada faaliyet gösteren matematiksel işlevdir. Bu fonksiyonlar ağırlıklı girdilerin toplamı ve bias değerini alarak bu toplamdan bir çıktıya dönüştürmektedirler. Aktivasyon fonksiyonlarının esas amaçları bir algılayıcının ateşlenme kararının verilmesinde rol oynamaktadır. Tıpkı nöron hücrelerindeki gibi aktivasyon fonksiyonları eşik değere göre çalışmakta, eşik değerlerin üzerinde bir çıktı algılamaları halinde nöron ateşlenmektedir. Bunun neticesinde çıktı değeri belirli bir değer olarak yeniden hesaplanır. Aksi durumda nörondan bir ateşleme olmaz çıktı değeri sıfır olur (Karpathy, 2018).

Yaygın olarak kullanılan aktivasyon fonksiyonlarının arasında: Sigmoid fonksiyonu, Hiperbolik tanjant fonksiyonu (Tanh), ReLU (Rectified linear unit) fonksiyonu ve Softmax aktivasyon fonksiyonu yer almaktadır.

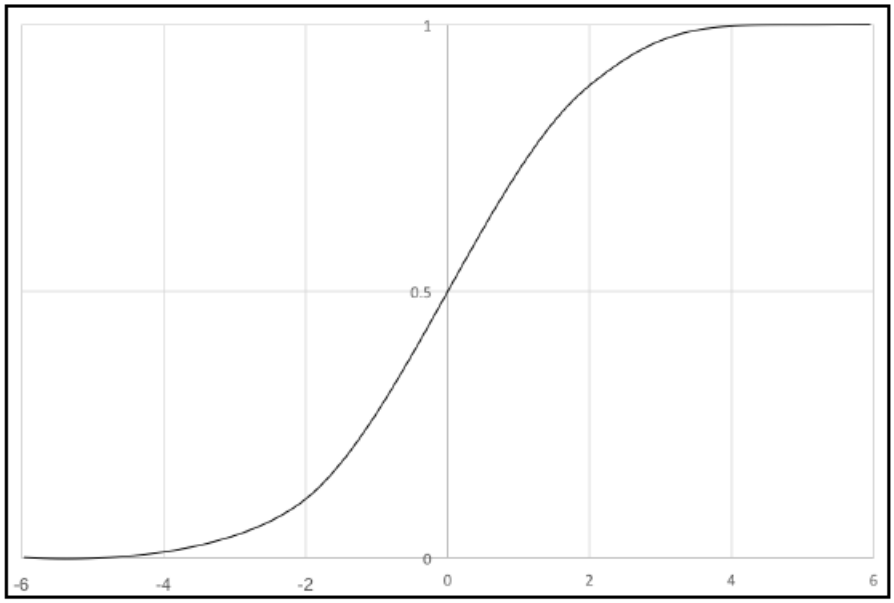
3.5.2.1. Sigmoid fonksiyonu

Sigmoid fonksiyonunda girdi değerleri 0 ile 1 arasında bir değer alabilmektedir. Bu fonksiyon özellikle gizli katman ve çıktı seviyesindeki katmanlarda kullanılmaktadır. Özellikle değerleri olasılık olarak ifade etme açısından başarılı olup sınırlandırması ikili olmaktadır.

Gradient kaybı sorunu nedeniyle bazı derin ağlarda kullanılmamaktadır. Çünkü her Y değerleri x değerinde değişikliklere göre oldukça az tepki vermektedir.

$$(0,1)\ aralığı\ için\ \ f(x) = 1/(1 + e^{-x}) \qquad (3.2)$$

Bu algoritmanın bir diğer dezavantajı olabilecek husus, öğrenmenin yavaş gerçekleştiği durumlarda hatayı minimum olarak gösteren yerel minimum değerlerine takılabilmektedir.



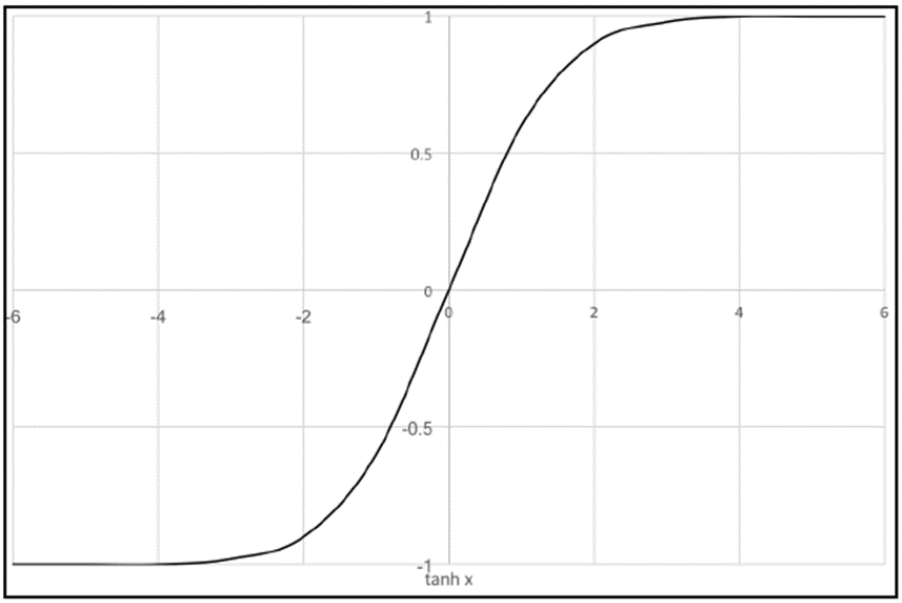
Şekil 3.10. Sigmoid fonksiyonu.

3.5.2.2. Hiperbolik tanjant fonksiyonu (Tanh)

Hiperbolik tanjant fonksiyonu aslında sigmoid fonksiyonunun ölçekli halidir. Sigmoid fonksiyonundan farklı olarak Tanh fonksiyonu -1 ile +1 arasında değerler almaktadır. Bu durumun bir sonucu olarak sigmoid fonksiyonunda farklı olarak daha açısı diktir ve fazla değer alabileceği için kararlı bir fonksiyondur. Gradient kaybı sorunu ile daha az yüzleşmesine karşın hala fonksiyonun uç kısımlarında problem gözlemlenir.

Tanh fonksiyonu, fonksiyon 3.3'de gösterilmiş buna karşılık gelen grafiğine şekil 3.10'da yer verilmiştir.

$$f(x) = \tanh(x) = \frac{(e^{x} - e^{-x})}{(e^{x} + e^{-x})} \tag{3.3}$$



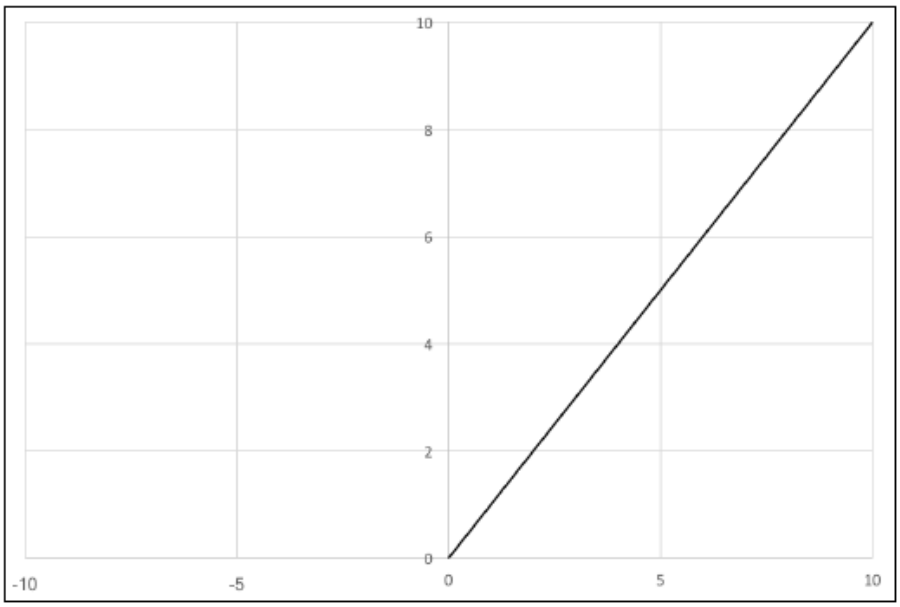
Şekil 3.11. Hiperbolik Tanjant (Tanh) fonksiyonu.

3.5.2.3. ReLU (Rectified linear unit) fonksiyonu

ReLU - Rectified linear unit fonksiyonu sigmoid ve Tanh fonksiyonundan farklı olarak büyük değerler alınmasına müsaade etmektedir. Buna karşın negatif değerler yerine de 0 değeri almaktadır. Bu durum algılayıcıların bir kısmını devre dışı kalmasını sağlayarak katmanlarda yer alan algılayıcıların seyreltilmesini sağlamaktadır. Rectified linear unit fonksiyonu pozitif girdi değerleri için büyüklüğüne bakmaksızın x değerlerinin kendisini alırken negatif değerler için "0" değeri almaktadır.

ReLU fonksiyonu, fonksiyon 3.4'de gösterilmiş buna karşılık gelen grafiğine şekil 3.11'da yer verilmiştir.

$$[0,\infty)\ aralığı\ için\ \ f(x) = \begin{cases} 0\ için\ x < 0 \\ x\ için\ x \geq 0 \end{cases} \tag{3.4}$$



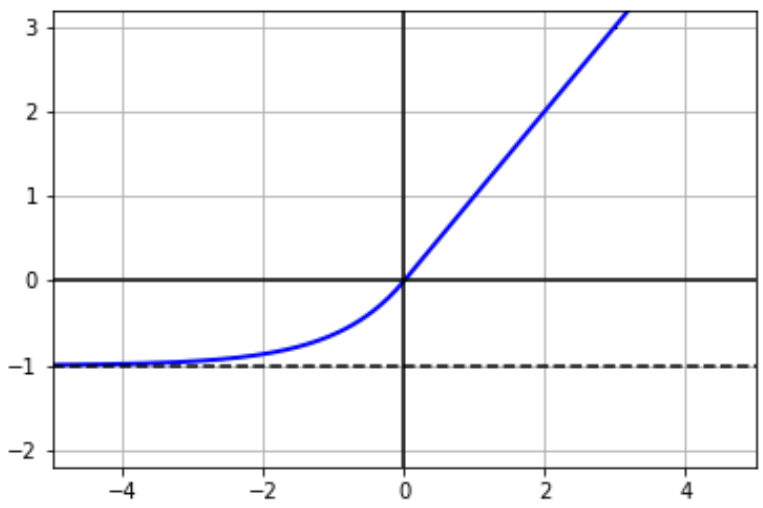
Şekil 3.12. ReLU fonksiyonu.

3.5.2.4. ELU (Exponential linear unit) aktivasyon fonksiyonu

Exponential linear unit aktivasyon fonksiyonu ReLU fonksiyonuna benzerlik göstermektedir. “x” pozitif değerleri için yine x değerini alırken negatif değeri için eksponansiyel bir değer alarak sonsuz uzayda “-1” değerine yaklaşmaktadır. ReLU aktivasyon fonksiyonundan önemli farklı algılayıcıların aldığı değer pozitiften negatife geçişte doğrudan kırılım göstermemeleri yavaş bir düzeltme ile “-1” değerine doğru gitmeleridir.

ELU fonksiyonu, fonksiyon 3.5’de gösterilmiştir.

$$[0,\infty)\ aralığı\ için\ \ f(x) = \begin{cases} \mathrm{a}(e^{x}-1)\ için\ x < 0 \\ x\ için\ x \geq 0 \end{cases} \tag{3.5}$$



Şekil 3.13. ELU fonksiyonu.

3.6. Bilgisayarlı Görü ve Evrişimli Sinir Ağları

Bilgisayarlı görü temelde bilgisayarların belir bir çalışma mantığından hareketle görüntü ve video verilerinin bilgisayarlar tarafından algılanması hadisesidir. Bilgisayarlı görü özellikle yapay zekâ ve makine öğrenmesi konuları ile eşlenik çalışır. Temelde her bir piksele atanan bir renk kodu ile verinin algılanması ve görülmesi hadisesi söz konusudur. Eğer ki görüntü siyah beyaz ise tek matriste 0 ile 255 arasında değer alarak bir piksel beyazdan siyaha doğru renk alır. Renkli bir görüntünün işlenmesi söz konusu ise RGB (Red Green Blue) kırmızı, yeşil ve mavi matrislerinden her biri için 0 ile 255 arasında değer alacak bu sayede ilgili pikselin rengi ortaya çıkacaktır.

Bilgisayarlı görü uygulama alanları özetlenecek olursa nesne algılama ve tanıma, yüz tanıma görüntü segmentasyonu, hareket analizi, sanal ve arttırılmış gerçeklik, tıbbi görüntüleme, endüstriyel denetim ve kalite kontrol verilebilir.

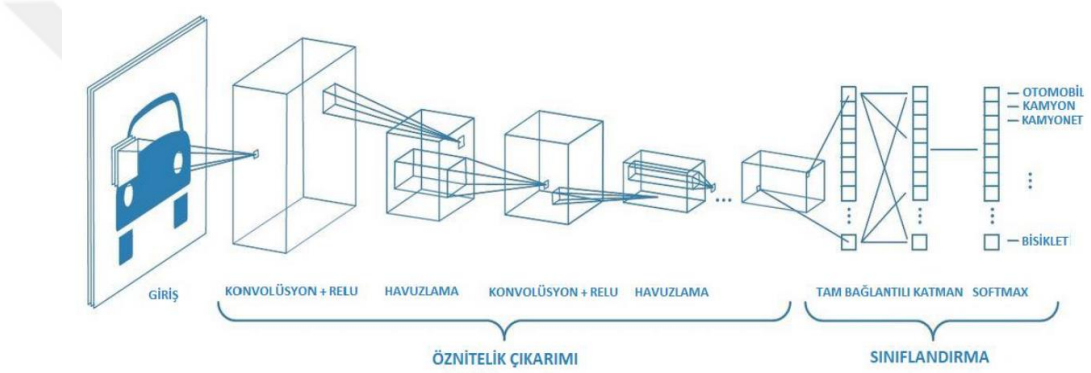
Konvolüsyonel sinir ağları özellikle bilgisayarlı görüde, nesne tanımada ve görsel verilerin analitik bir şekilde irdelenmesi, verilerin analiz edilmesi noktasında etki gösteren derin öğrenme modellerinden biridir. Konvolüsyonel sinir ağları, esasen evrişimli olarak geçmektedir. Bu kavramı kullanmasındaki temel sebep birçok katmandan oluşmasından kaynaklanmaktadır.

Yapay zekanın derin uykusundan uyanmasını sağlayan GPU ve CPU'ların gelişimi aynı zamanda çok katmanda işlem yapan derin öğrenme modellerinin ve

konvolüsyonel sinir ağlarının işlem yapmasına imkân sağlamıştır (İnik ve Ülker, 2017).

3.6.1. Evrişimli sinir ağları katmanları

Evrişimli sinir ağları çok çeşitli zorlukları bünyesinde barındıran görevleri, hem özellik çıkarma hem de sınıflandırmaları gerçekleştirerek tamamlayabilmektedirler. Bunu gerçekleştirebilmek için konvolüsyonel sinir ağlarının birçok katmanı bulunmaktadır. Konvolüsyonel sinir ağlarında sınıflandırma ve özellik çıkarımı adımlarının içerisinde de ilave katmanlar yer almaktadır. Bunlar, giriş katmanı, konvolüsyon katmanı, havuzlama katmanı, bırakma katmanı, tam bağlı katman ve sınıflandırma katmanıdır.



Şekil 3.14. Konvolüsyonel sinir ağı katmanları.

3.6.1.1. Giriş katmanı (Input layer)

Konvolüsyonel sinir ağında resmin ilk karşılandığı ve alındığı katmandır. Veri boyutu giriş katmanında ne kadar büyük ya da çözünürlüğü yüksek olarak alınırsa, modelin başarı ihtimali o kadar yükselecek ama veri sayısı artacağı için hızı düşecektir. 28x28 piksele sahip bir resimde resim eğer siyah beyaz ise 784 veri ile işlem yapılacakken renkli olduğunda 2.352 veri ile işlem yapacak eğer bu resim 1024*720 pikselden oluşuyorsa siyah beyaz olması halinde veri sayısı 737.280 olurken, bir de bu resim renkli olursa tek resim için 2.211.840 veri ile işlem yapılacaktır. Resmin çözünürlüğünün düşmesi halinde modelin başarısı bu sefer olumsuz etkileneceği için burada bir fayda zarar analizi yapılmalı gerekliliği ortaya çıkmaktadır (İnik ve Ülker, 2017).

3.6.1.2. Konvolüsyon katmanı (Convulation layer)

Konvolüsyon katmanında resim hem dikey hem yatay olarak incelenerek evrişim işlemi gerçekleştirilmektedir. Konvolüsyon işlemi neticesinde özellik haritaları elde edilmeye başlanmaktadır. Her bir ilgi bölgesi için girdiler bir ağırlık matrisi ile çapılarak eşik değer ilave edilmektedir. Evrişim işlemi kernel matrisin resim üzerinde her bir piksel ile kaydırarak çarpma işlemi gerçekleştirmesidir. Evrişim işleminin formülü, formül 3.6'da özetlenmiştir.

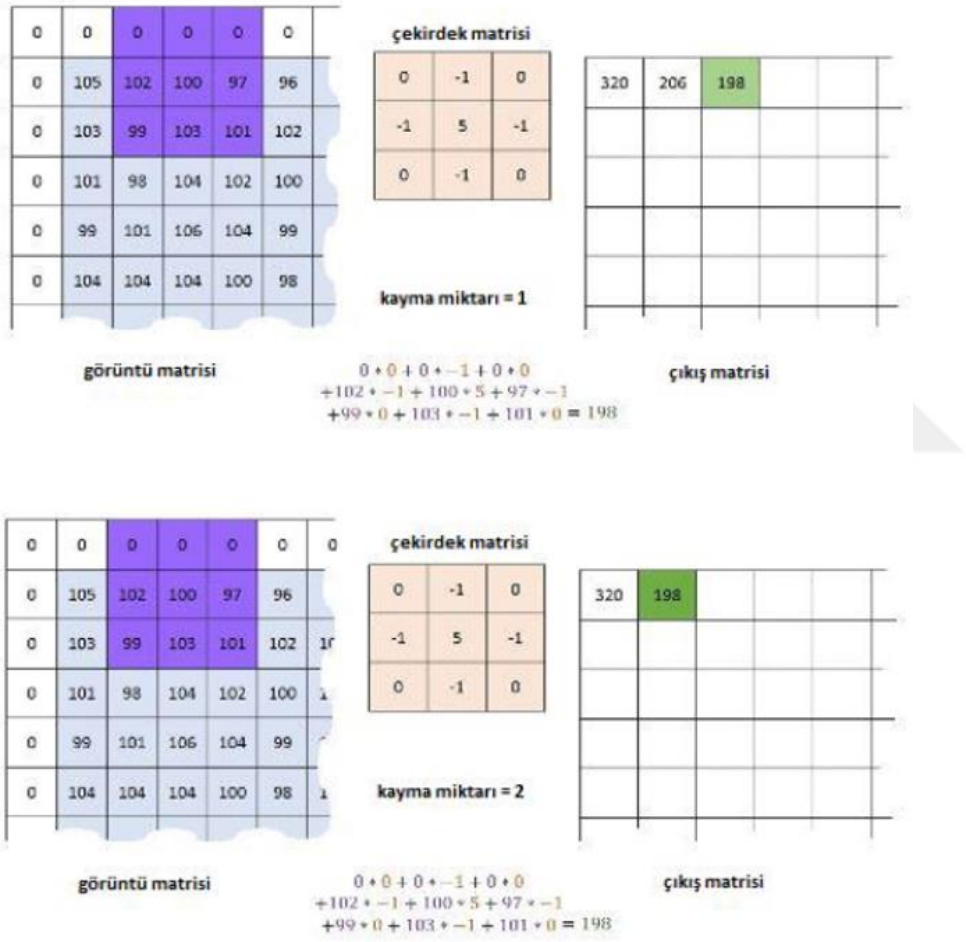
$$S(i,j) = (I * K)(i * j) = \sum_{m} \sum_{n} I(i+m, j+n)K(m+n) \tag{3.6}$$

Formül 3.6'da görüldüğü üzere I giriş görüntüsünün matrisi, K çekirdeğin matrisini (kernel); S ise yine konvolüsyon işlemi çıktısını ifade edilmektedir. Konvolüsyon işlemi '$*$' işareti ile ifade edilmektedir. Şekil 3.15 de evrişimli sinir ağları içerisinde kullanılan çekirdek matrisleri ve hangi işlemlerde kullanıldığı noktasında bilgiye yer verilmiştir her bir derin öğrenme algoritması mimarisinde yapılacak işleme göre evrişimli ağlarda aşağıda yer alan matrislerden tercihler kullanılmaktadır(Googfellow, Bengio ve Courville, 2016).

İşlem	Çekirdek Matrisi (Kernel -ω)	Görüntü Sonucu
Tanımlama	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	
Kenar Tespiti	$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$	
	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$	
	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	
Keskinleştirme	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	
Bulanıklaştırma	$\frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	
Gaussian Bulanıklaştırma 3 x 3	$\frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$	
Gaussian Bulanıklaştırma 5 x 5	$\frac{1}{256}\begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$	
Keskinlik Azaltım Maskesi 5 x 5	$\frac{-1}{256}\begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & -476 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$	

Şekil 3.15. Kernel (çekirdek) matris önerileri (Zeren, 2020).

Konvolüsyonel işlemler gerçekleştirilirken kernel çerçevesi alınarak modele ilk olarak girişi yapılan görünün üzerinde yatay ve düşey olarak kaydırmalı olarak hareket ettirilir hem de görüntünün piksel değeri ile çarptırılır. Her bir çıktının toplamı ele alınarak ilgili yeni piksele girişi yapılarak kaydı sağlanır. Yatay düzlemin tamamlanmasının ardından işlem için bir satır aşağı geçilerek yine aynı hesaplama işletilir. Eğer işlenecek görüntü renkli ise RGB matrisler dahilinde ayrı ayrı uygulaması gerekmektedir. Sonrasında hepsi toplanır aktivasyon haritası ortaya çıkarılır. Kayma miktarı parametresi her işlem sonrası ne kadar kayacağını söylemektedir.

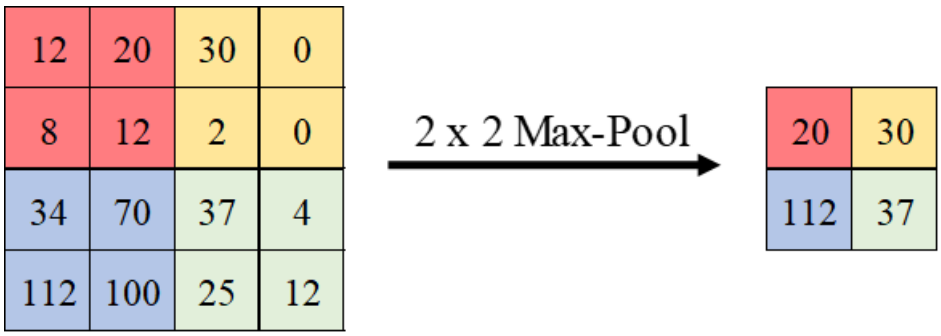


Şekil 3.16. Konvolüsyon işlemi örneği (Zeren, 2020).

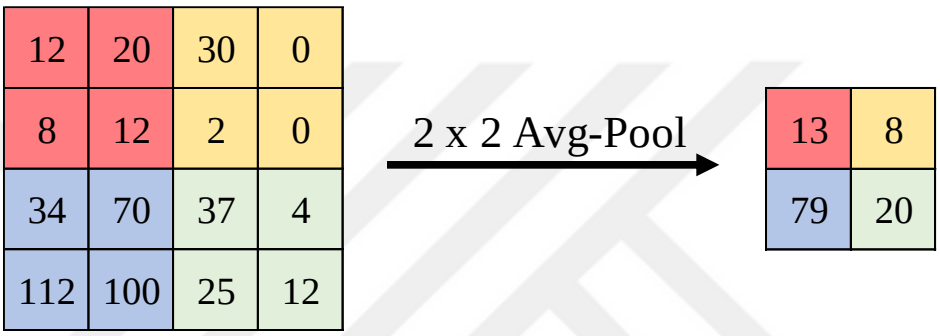
3.6.1.3. Havuzlama katmanı (Pooling layer)

Hesaplama alanı olarak kullanılan havuzlama katmanında modelin asıl amaç yeniden resmin boyutunu düzenlemek ve modelin başarısını yükseltmek esas amaçtır. Buradaki kullanılacak filtrenin kaça kaç olacağı önemlidir. Havuzlama işlemi

yapılırken maksimum havuzlama ya da ortalama havuzlama gibi yöntemler kullanılabilmektedir.



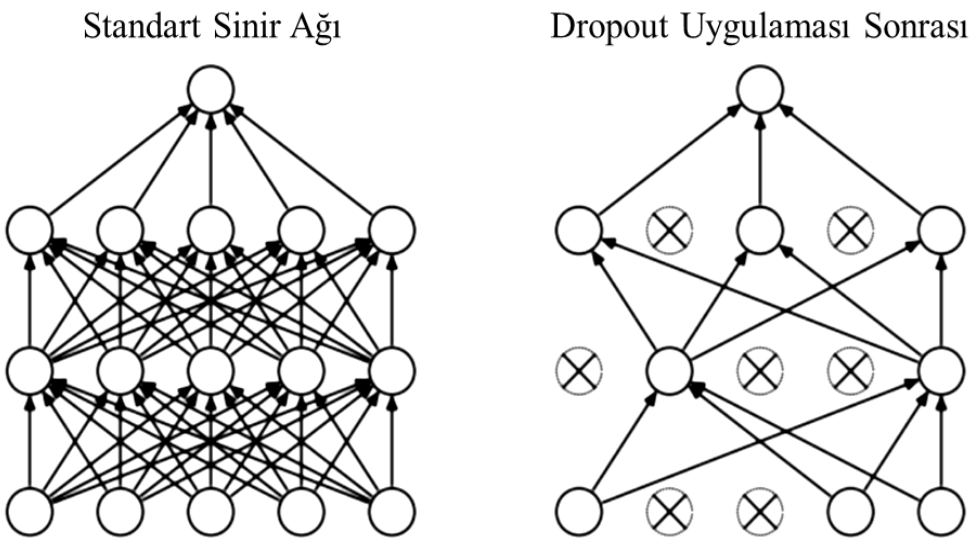
Şekil 3.17. Maximum pooling layer örneği (Yani, 2019).



Şekil 3.18. Average pooling layer örneği (Yani, 2019).

3.6.1.4. Bırakma katmanı (Dropout)

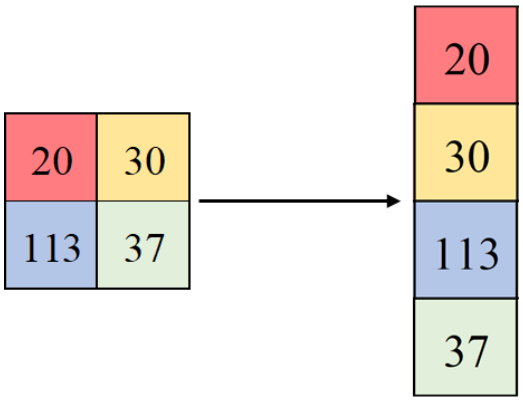
Modelin ezbere düşmemesi açısından evrişimli sinir ağları için en önemli katmanlardan biridir. Bırakma katmanı sinir ağının aşırı işlem yapmasının yahut sadece belli nöronlarda işlem yapmasının da önüne geçmektedir. Kapsam dahilinde bilinmesi gereken bilgilerin eksikliğinde dahi modelin doğru bir öğrenme sürecinden geçmesine yardımcı olmaktadır (Verma ve Pal, 2016).



Şekil 3.19. Bırakma katmanının sinir ağına uygulanması (Budhiraja, 2016).

3.6.1.5. Tam bağlı katman (Fully connected layer)

Tam bağlı katman genellikle evrişimli sinir ağlarının sonunda bulunmaktadır. Elde edilen matris şeklindeki verilerin düzleştirme işleminden geçmesini sağlamaktadır. Bu işlem sırasında sol üst satırdan başlayarak sağa doğru taranarak yukarıdan aşağıya yazılır sonrasında satır bitince bir alt satır için de aynı işlem yapılarak altına devam edilir.



Şekil 3.20. Flattening-düzleştirme işlemi.

3.6.1.6. Sınıflandırma katmanı

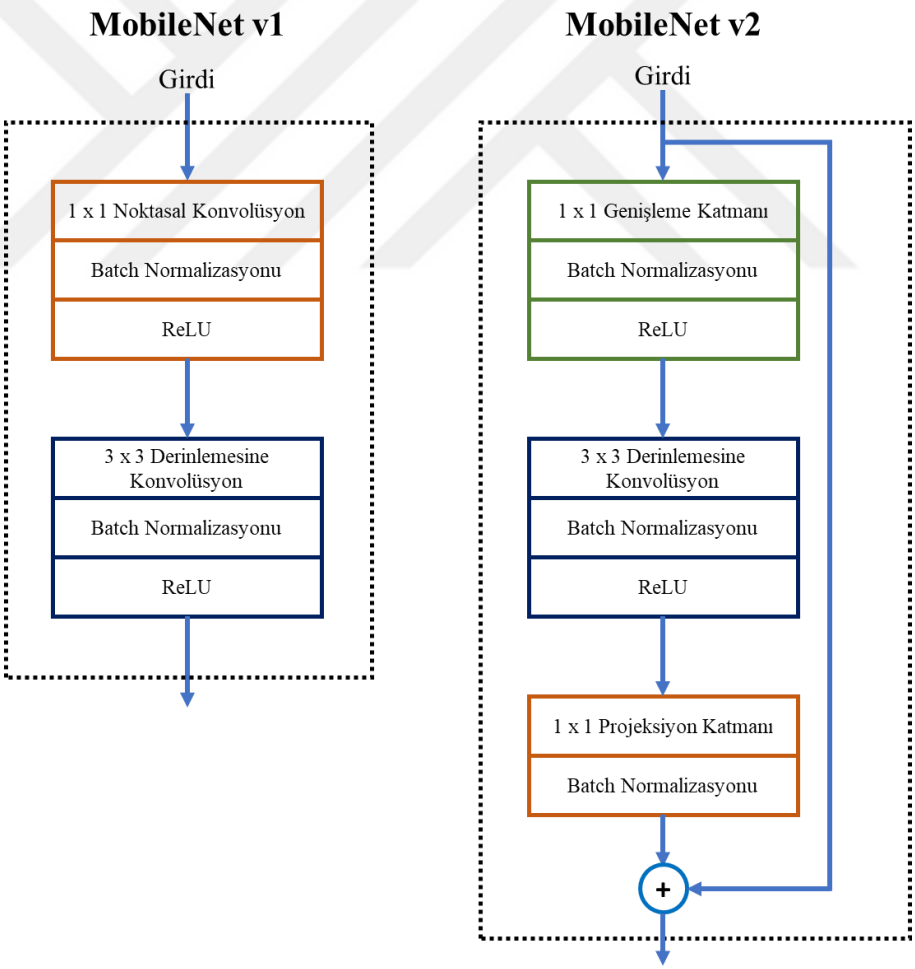
Sınıflandırma katmanı evrişimli sinir ağları için en önemli katmandır bu katmanda nesneler sınıflarına göre 0 ile 1 arasında bir değer alır. 1 değerine en yakın olarak sınıflandırılan ilgili tahmin edilen sınıf olarak işaretlenir. Modelde kaç adet sınıf tahmin edilmek isteniyorsa o kadar sayıda çıktı üretmektedir. Softmax sınıflandırıcısı yüksek oranda başarı göstermesi nedeniyle tercih edilmektedir (İnik ve Ülker, 2017).

3.6.2. Yapay sinir ağı modelleri

Derin öğrenme modellerinin mimarilerine dahil olan yapay sinir ağı modelleri doğru çıktı üretme noktasında büyük rol oynamaktadır. Çoğunlukla derin öğrenme mimarilerinde kullanılmakta olan modeller; MobileNet, GoogLeNet, Inception, Resnet, VGG (Visual Geometry Group), AlexNet, LeNet özetlenmiştir.

3.6.2.1. MobileNet

Diğer mimarilere göre boyutu oldukça küçük buna karşın hızı yüksek olan bir evrişimli sinir ağıdır. Her girdi kanalında tek filtre söz konusudur. Filtreleme ve birleştirme işlemleri ayrı ayrı katmanlarda gerçekleştirdiği için işlem hızı oldukça yüksektir. Çıktının sınıflandırılması amacıyla softmax katmanı kullanılmaktadır. Aktivasyon fonksiyonu olarak ReLU kullanılmaktadır. MobileNet mimarisi toplamda 28 katmandan oluşmaktadır (Howard ve ark., 2017).



Şekil 3.21. MobileNet V1 ve V2 mimarileri.

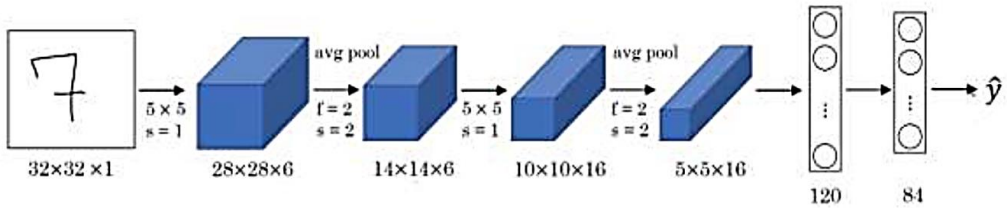
3.6.2.2. Inception

Inception modeli milyonlarca parametreye sahip daha önce oldukça güçlü işlemcili iş istasyonları ile eğitilmiş bir evrişimli sinir ağıdır. Bu modelin sıfırdan normal işlemcili bir bilgisayar ile eğitiminin gerçekleştirmesi saatlerce sürebilir. Bu mimarinin ilk versiyonu olan V1 2015 yılında tanıtılmıştır. 2016 yılında ise daha da geliştirilip topak normalleştirme tekniği ilave edilmiş ve V2 olarak tanıtılmıştır (Szegedy ve ark., 2015; Szegedy ve ark., 2016).

Aktivasyon fonksiyonu olarak ReLU aktivasyon fonksiyonu kullanılmaktadır. Inception mimarisi, sınıflandırma görevleri için tipik olarak tam bağlantılı (fully connected) katmanlar ile sonlandırılır. Bu tam bağlantılı katmanlar, öğrenilen özellikleri alır ve verilen girdiyi farklı sınıflara sınıflandırmak için kullanılır. Bu mimaride yine sınıflandırıcı olarak softmax sınıflandırıcı kullanılmıştır. Inception V1 modülü yaygın olarak GoogLeNet olarak da bilinmektedir.

3.6.2.3. LeNet-5

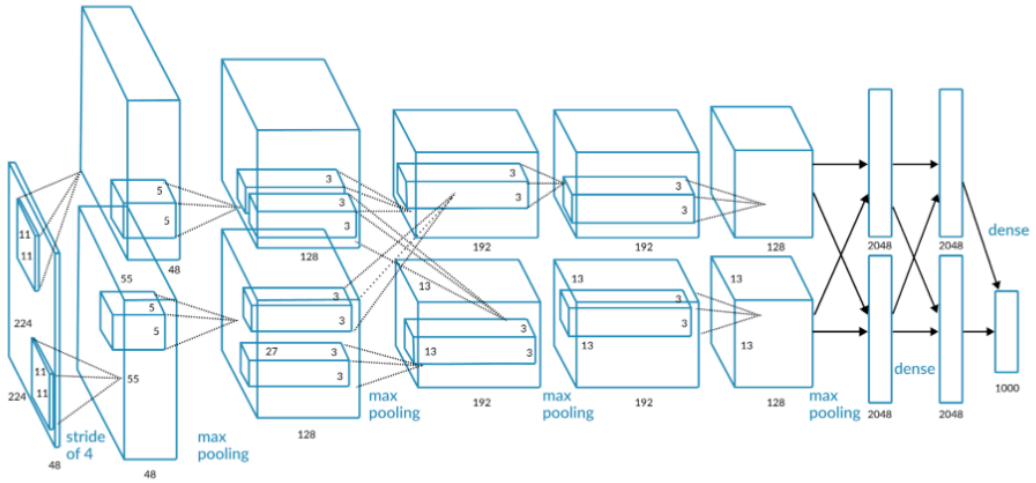
Tarihler 1998 yılını gösterdiğinde LeCun ve ekibi ilk başarılı sonuç veren evrişimli sinir ağı modeli olan LeNet Mimarisini tanıtmıştır. Bu mimaride aktivasyon fonksiyonu olarak Tanh (hiperbolik tanjant) ve sigmoid fonksiyonları kullanılmıştır. Bu mimarinin diğerlerinden ayıran önemli bir noktası ortalama havuzlama işleminin maksimum havuzlama işlemi yerine kullanılmasıdır 60.000 parametre hesaplama kapasitesine sahiptir (LeCun ve ark., 1998).



Şekil 3.22. LeNet Mimarisi (LeCun ve ark., 1998).

3.6.2.4. AlexNet

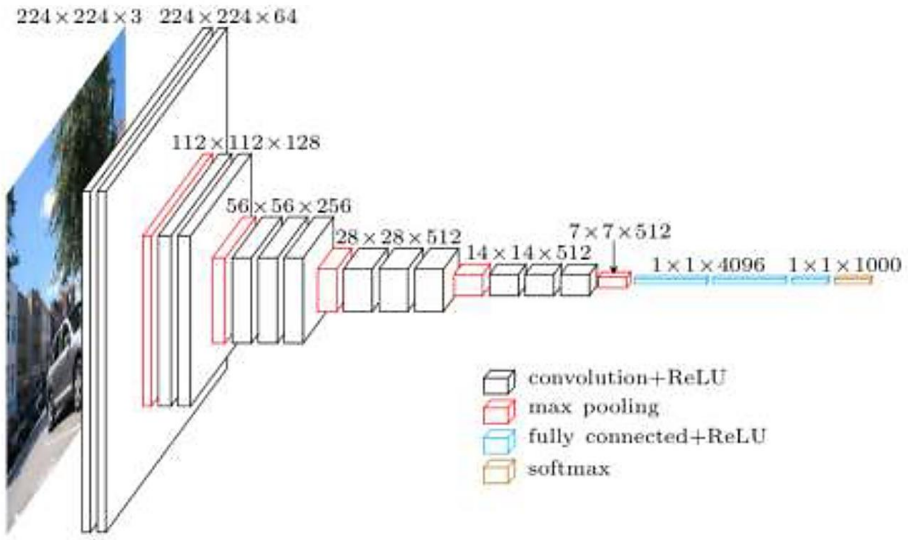
2009 yılından sonra görüntü tespiti noktasında algoritmaları zorlayan ve bu alanın önünü açan ImageNet yarışmasında 2012 yılının şampiyonu AlexNet modeli olmuştur. ReLu fonksiyonunu aktivasyon fonksiyonu olarak kullanan bu model, maksimum havuzlama katmanını kullanmıştır. AlexNet'i rakiplerinden ayıran önemli husus çift işlemci kullanması ve 60.000.000 parametre kullanma kapasitesine sahiptir. Şekil 3.22'de mimarisine yer verilmiştir (Krizhevsky ve ark 2012).



Şekil 3.23. AlexNet Mimarisi.

3.6.2.5. VGG (Visual Geometry Group)

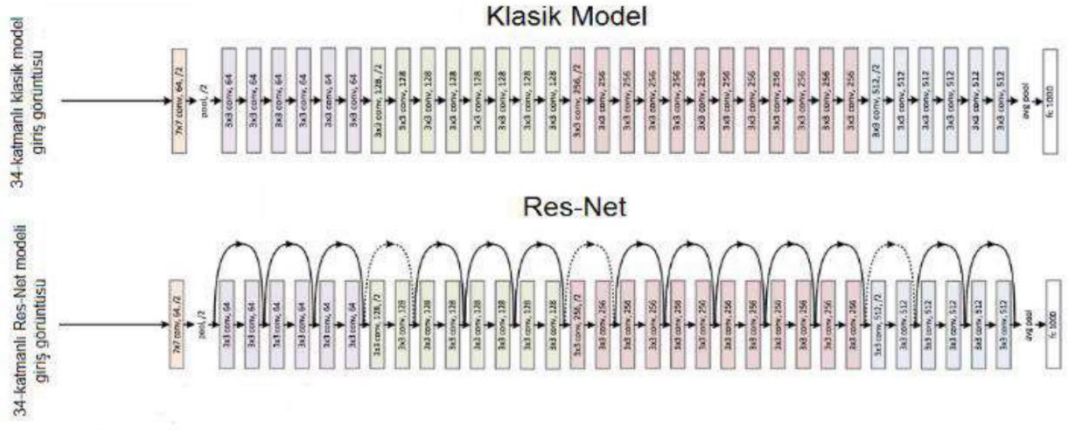
138 milyona yaklaşan parametre hesabı yapabilen VGG mimarisi, çoklu evrişim katmanlarını kullanır. Aktivasyon fonksiyonu olarak ReLu kullanan söz konusu mimari, maksimum havuzlama katmanını havuzlama katmanı olarak kullanmaktadır. Sınıflandırıcı olarak da yine softmax tercih edilmiştir (Simonyan ve Zisserman, 2014).



Şekil 3.24. VGG Mimarisi.

3.6.2.6. ResNet

Derin öğrenme mimarilerinde en fazla evrişimli katmanı bünyesinde barındıran mimari ResNet mimarisidir. Toplamda 152 katmanı söz konusudur. Evrişimli katmanlar atlamalı model ile sonraki katmanları beslemektedir. Bu modelde yine ReLU aktivasyon fonksiyonu çalışmaktadır.



Şekil 3.25. ResNet Mimarisi (He ve ark., 2015).

3.6.3. Görüntü sınıflandırma ve kritik bölge tespitimde kullanılan derin öğrenme algoritmaları

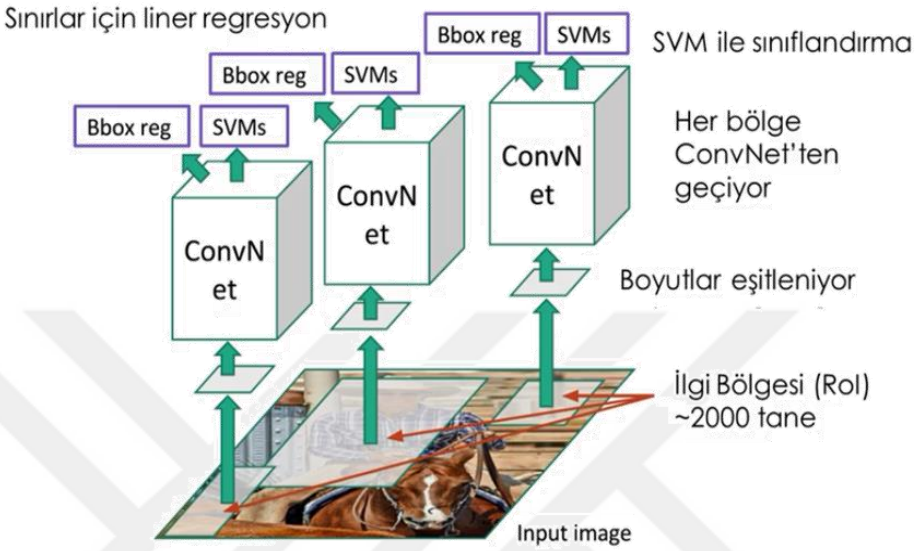
Derin öğrenmede kullanılan algoritmaların analiz gücü değerlendirildiğinde ilk olarak karşımıza PASCAL VOC veri setinde yapılan çalışmalar çıkmaktadır. Yapay sinir ağlarının aktif olarak kullanılmaya başlanmadığı 2013 yılına kadar bu veri seti ile yapılan çalışmaların nesne tespiti başarısı %40 seviyesinin altında kalmaktadır (Girshick ve ark., 2014). Bu yıllardan sonra artan CPU ve GPU teknolojisi ve yapay sinir ağları ile derin öğrenme tekniklerinin algoritmalarda aktif şekilde kullanılması ve algoritmalarını mimarilerinin optimize edilmesi ile %80 seviyesinin üzerine nesne tespitinin yapılması sağlanmıştır (Girshick, 2015).

Bu bölümde söz konusu çalışmaya konu olan ve derin öğrenmenin ne güçlü algoritmalarından olan R-CNN, Fast R-CNN, Faster R-CNN, SSD-Single shot multibox detector ve YOLO – You look only once algoritmalarına yer verilmiş detayları ile açıklanmıştır.

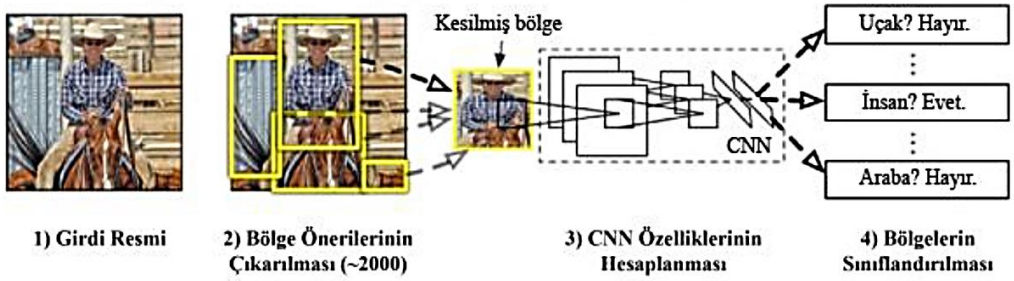
3.6.3.1. R-CNN

R-CNN kavramı evrişimli sinir ağlarının bölge tabanlı hale getirilmesi ile ortaya çıkmış ilk araştırmalardan biri, Girshick ve arkadaşları tarafından 2014 yılında ortaya konulmuştur. R-CNN temelinde bölge önerileri ve odak bölgesi yer almaktadır. Farklı ölçekteki pencerelerin CNN'den işlenmeden önce aynı boyutta olmaları önemlidir. Evrişimli sinir ağları aynı boyutta girdiyi kabul etmektedir. Her bir ilgi bölgesi öncelikle konvolüsyonel sinir ağından geçirilmekte ardından destek vektör makinesi ile sınıflandırılmaktadır. İlgi bölgesindeki nesnenin sınırlarının çizilmesi için lineer regresyon kullanılmaktadır. R-CNN'e beslenen resimde ilk olarak 2000 adedin

üzerinde ilgi bölgesi belirlenir. Her bir ilgi bölgesi öncelikle boyut eşitlenmesine tabi tutulur sonrasında ise konvolüsyonel ağ tarafından işlenilir. Destek vektör makinesi ile ilgi bölgesinde ne olduğu ortaya konulur yani sınıfı belirlenir. Lineer regresyon ile yerinin işaretlenmesi sınırlarının çizilmesi gerçekleştirilir. Çok katmanlı işlem görmesi R-CNN'in yavaş olmasına neden olmaktadır (Girshick ve ark., 2014).



Şekil 3.26. R-CNN çalışma yapısı (Girshick ve ark., 2014).



Şekil 3.27. R-CNN nesne tespiti örneği (Girshick ve ark., 2014).

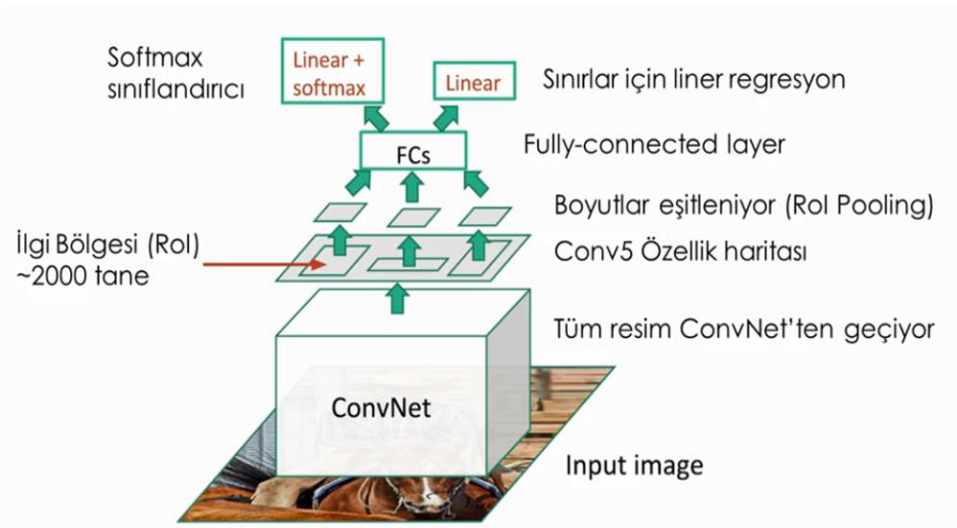
Yapılan çalışmalar konvolüsyonel yapay sinir ağlarını kullanması nedeniyle mükemmel seviyede nesne tespit doğruluğu elde etmesine karşın, R-CNN optimum bir yapıdan uzaktır. Öncelikle eğitim süreci çok aşamalı olmasına karşın buna cevap veremeyen bir tek şeritli yol gibidir. R-CNN log kaybı kullanarak nesne sınıflandırmayı konvolüsyonel ağ üzerinden ayarlar. Sonrasında konvolüsyonel ağ özelliklerini destek vektör makinesi aracılığı ile oluşturulur. Eğitimin son aşamasına gelindiğinde sınırlayıcı kutular regresyonlar aracılığı ile öğrenimi gerçekleştirilir. Bu

süreci oldukça uzatmaktadır. Eğitim zamanı açısından oldukça maliyetlidir. Ciddi seviye gigabayt veri depolama ihtiyacı gerektirir. Tespit aşamaları çoklu ve ayrı ayrı gerçekleştiği için oldukça yavaştır. Her bir ilgi bölgesi tek tek irdelenir.

3.6.3.2. Fast R-CNN

Faster R-CNN algoritmasında konvolüsyonel sinir ağı içeren aslında temelinde R-CNN Mimarisi ile aynı olup aslında nesne tespiti sırasındaki sürecin optimizasyonu ile daha hızlı hale gelmiş bir versiyonudur. Faster R-CNN algoritmasının R-CNN algoritmasından farklılaşan kısmı ilk olarak resim üzerinde çoklu bölge önerisinde bulunmaktansa resim bir bütün halinde konvolüsyonel sinir ağından geçirilmekte bunun sonucunda asıl resimle benzerlik gösteren yine yüksek çözünürlüğe sahip bir özellikler haritası elde edilmektedir. Özellik haritasının elde edilmesi sonrasında bölge önerilerini gerçekleştirilir. Fast R-CNN algoritmasında olduğu gibi her bir bölgenin tek tek konvolüsyonel sinir ağına verilerek incelenmesi yerine resim en baştan konvolüsyonel sinir ağından geçirilmektedir.

İki bin adet civarında bölge önerisinin öncelikle boyutları eşitlenir sonrasında, direk bağlı ağ katmanı kullanılmaktadır. Sınıflandırmanın gerçekleştirilmesi için yeni bir model ortaya konulması yerine lineer ve softmax sınıflandırıcılar kullanılmaktadır. Söz konusu nesnenin küçülen kutular ile sınırlarının işaretlenmesi için de lineer regresyon kullanılmaktadır (Girshick, 2015).



Şekil 3.28. Fast R-CNN mimarisi (Girshick, 2015).

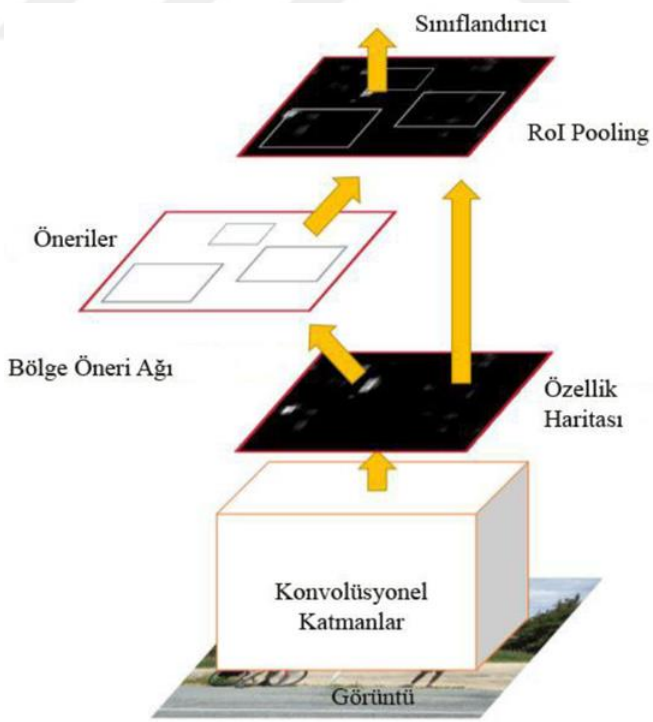
R-CNN algoritmasının eğitim süresi 84 saat olduğu bir örnekte, Fast R-CNN algoritması 8,75 saatte eğitimi gerçekleştirmekte 9,6 kat daha hızlı bir şekilde eğitim

sürecini tamamlamaktadır. Sonuç üretme noktasında ise R-CNN algoritması 49 saniyelik sürede bir çıktı ürettiği örnekte Faster R-CNN algoritması 2,3 saniyede sonuç üretmekte 21,3 kat daha hızlı sonuç üretebilmektedir. Her ne kadar Fast R-CNN hızlı görünse de henüz süreci tam olarak mükemmelleşmemiş olup özellikle bölge önerisinde ve sınıflandırma sürecinde vakit kaybetmektedir.

3.6.3.3. Faster R-CNN

Fast R-CNN'de yer alan bölge önerisi darboğazı ve sınıflandırma darboğazına çözüm olacak yeni bir ağ geliştirilmesi ile ortaya Faster R-CNN algoritması çıkmıştır. Faster R-CNN algoritması tıpkı Fast – RCNN algoritmasında olduğu gibi resmi alarak öncelikle konvolüsyonel sinir ağından geçirmekte ve yüksek çözünürlüğe sahip bir özellik haritası elde etmektedir. Faster R-CNN algoritması Fast R-CNN gibi seçimli arama yaparak ilgi bölgesi ya da bölge önerileri çıkarmak yerine bölge önerisi ağı oluşturmaktadır. Bölge önerileri konvolüsyonel ağ içerisinde yapılmakta bu durumda modelin çok hızlı sonuç üretmesini sağlamaktadır.

Ortaya çıkan öneri bölgeleri, doğrudan bağlı katmandan geçirilerek sınıflandırma yine konvolüsyonel sinir ağı tarafından yapılmaktadır.



Şekil 3.29. Faster R-CNN mimarisi (Ren ve ark., 2015).

Özetle Faster R-CNN'de Girdi Görüntü: İlk adım, nesneleri algılamak istediğiniz görüntüyü almak ve işlemeye hazırlamaktır. Evrişimli Sinir Ağı (Backbone Network): Önceden eğitilmiş bir evrişimli sinir ağı, girdi görüntüyü özellik haritalarına dönüştürür. Bu ağ, genellikle VGG, ResNet, ResNeXt gibi derin mimarileri içerebilir. Bölge Önerisi Ağı (Region Proposal Network- RPN): Özellik haritaları üzerinde çalışan bir ağdır. Bu ağ, olası nesne konumlarını önerir. Bu önerilen bölgeler daha sonra daha ayrıntılı bir analiz için seçilir. Bölgeyi Düzleştirme ve Özellik Çıkarma: RPN tarafından önerilen bölgeler, özellik haritalarından çıkarılır ve düzleştirilir. Bu adım, her bir bölgeyi daha küçük bir boyuta getirir ve aynı zamanda nesne algılama işlemi için gerekli olan özellikleri çıkarır. Bölgeye Özgü Evrişimler (RoI Pooling veya RoI Align): Bu adımda, her bölge için özellik haritasından özellikler alınır. Bu işlem, her bir bölgeyi sabit bir boyuta dönüştürmek için yapılır, böylece sonraki aşamalarda aynı boyutlu girdiye sahip olunur. Nesne Sınıflandırma ve Yer Tespiti: Bu aşamada, her bir bölge için nesne sınıflandırma ve yer tespiti yapılır. Bu işlem, her bölge içindeki nesnenin sınıfını tahmin etmeyi ve bölgenin nerede olduğunu tespit etmeyi içerir.

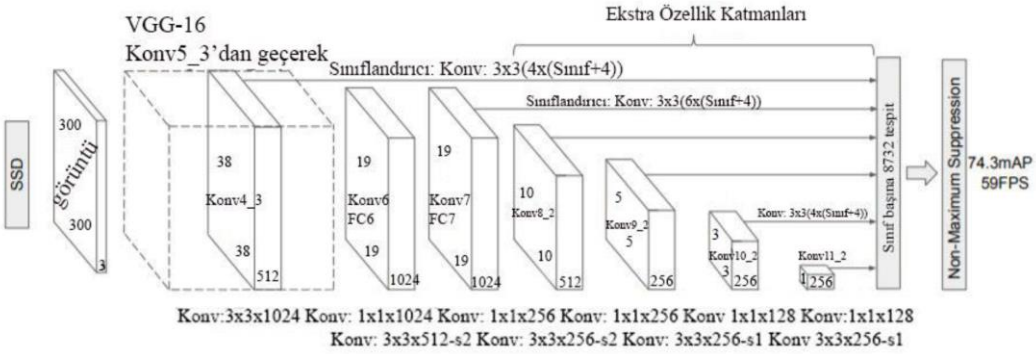
Faster R-CNN algoritmasının sonuç üretme hızı kıyaslandığında R-CNN algoritmasını 49 saniyede de ürettiği, Fast R-CNN algoritmasının 2,3 saniyede bir sonucu Faster R-CNN algoritması 0,2 saniyede üretmekte buna göre R-CNN'den 255 kat, Fast R-CNN den ise 11,5 kat daha hızlıdır. Bu hızı sahip olması gerçek zamanlı nesne tespitine imkân tanımaktadır.

3.6.3.4. Single shot multibox detector (SSD)

Faster R-CNN ile gerçek zamanlı nesne tespiti uygulamalarını dikkat çekmiş 2016 yılına gelindiğinde gerçek zamanlı nesne tanımaya bir imkân daha sağlayan SSD - Single Shot Multibox Detector derin öğrenme algoritması ortaya atılmıştır. SSD algoritmasının en önemli özelliği R-CNN algoritmaları gibi bir bölge ağı önerisi kullanmak yerine tek seferde nesne tespiti gerçekleştirmektedir.

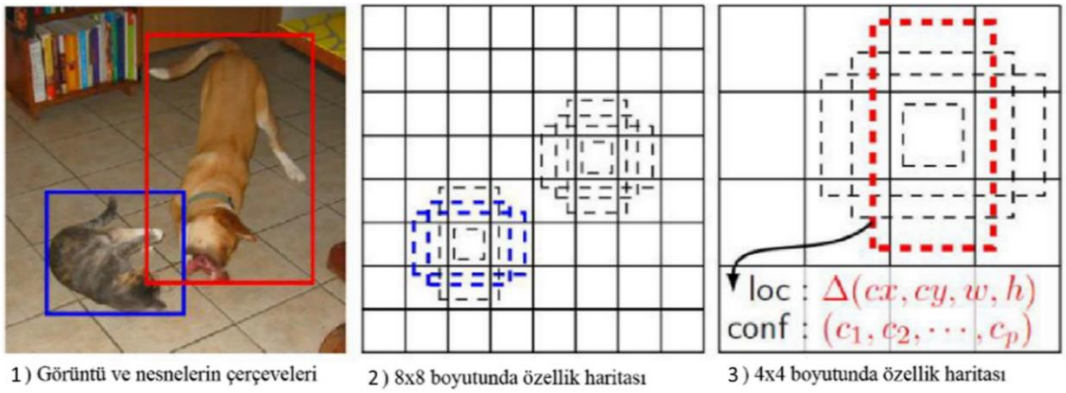
SSD'nin mimarisi özetle yedi temel adımları içermektedir. Bunlarda ilki girdi görüntü adımı olup nesneleri algılamak istediğiniz görüntüyü almak ve işlemeye hazırlamaktır. Sonrasında Evrişimli Sinir Ağı (Backbone Network) gelmekte, girdi görüntüyü işleyerek özellik haritalarını üretmektedir. Bu ağ, genellikle VGG, ResNet, MobileNet gibi evrişimli sinir ağlarından oluşabilmektedir. Sonraki adım çok ölçekli özellik haritaları adımıdır. SSD, farklı ölçeklerdeki özellik haritalarını kullanarak nesne tespiti yapmaktadır. Bu, farklı boyutlardaki nesneleri algılayabilme yeteneği sağlamaktadır.

Özellik haritaları, evrişimli sinir ağının farklı katmanlarından çıkarılan özellikleri temsil etmektedir. Konvolüsyonel Özellik Haritaları adımında her ölçek için çıkarılan özellik haritaları üzerinde konvolüsyon işlemleri yapılmaktadır. Bu konvolüsyonlar, her bir özellik haritasındaki her pikselin çeşitli nesne sınıfları ve ilgili özelliklere ait tahminleri üretmektedir. Çoklu Ölçekli Anchor Kutuları adımında ise her bir konvolüsyonel özellik haritası için önceden belirlenmiş boyutlara sahip kutular, "anchor" olarak adlandırılmaktadır. Bu anchor kutuları, farklı boyutlardaki nesneleri temsil etmektedir. Her bir anchor kutusu, konum ve boyut açısından çeşitli değişimlerle nesne önerileri üretmektedir. Sınıf sınıflandırma ve konum düzeltme adımında her anchor kutusu için, ilgili sınıfın olasılığını tahmin etmek için sınıf sınıflandırma ve kutunun gerçek nesne konumunu düzeltmek için konum regresyonu gerçekleştirilmektedir. Non-Maximum Suppression (NMS) adımında her bir ölçekten gelen nesne önerileri arasında gereksiz önerileri elemek ve aynı nesneyi birden fazla kez tespit etmeyi engellemek için NMS kullanılmaktadır (Liu ve ark., 2016).



Şekil 3.30. Single shot multibox detector algoritması mimarisi (Liu ve ark., 2016).

SSD algoritmasının çalışması sırasında şekil 3.11'de görüldüğü üzere görüntü ve nesnelerin çerçeveler, 8x8 boyutundaki özellik haritası oluşumu ve 4x4 boyutundaki özellik haritası çıkartma, nesnelerin çerçeveleri için pozisyon ve doğruluk değerleri hesaplanmasına yer verilmiştir.

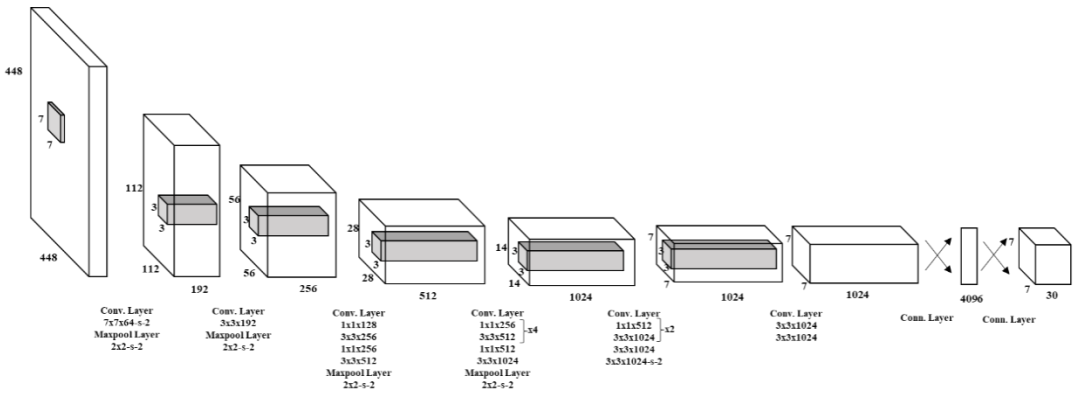


Şekil 3.31. SSD Algoritmasında nesne çerçevelemesi (Liu ve ark., 2016).

3.6.3.5. You look only once (YOLO)

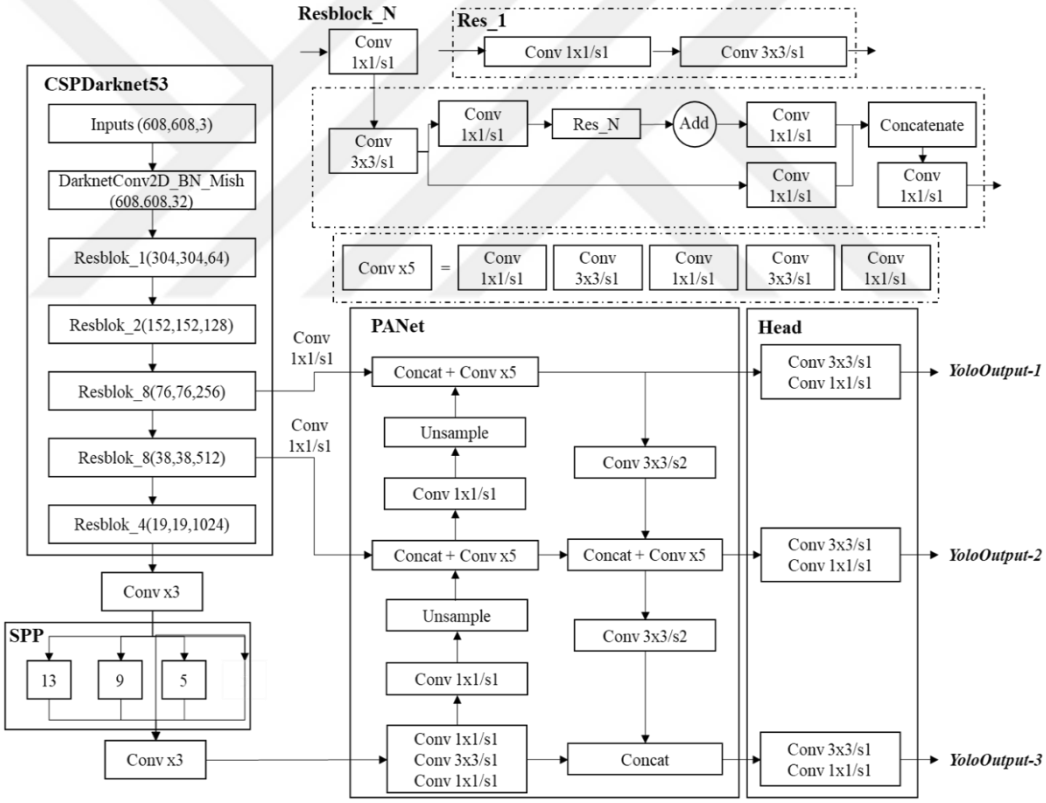
Nesne tespiti algoritmalarından beklenti çevik ve hızlı şekilde nesne tespitinin yapılmasıdır. YOLO algoritması Tek seferde bütün bir resmin verisini yapay sinir ağından geçirerek resmin içerisinde bulunan nesnelerin hem sınıfını hem de koordinatlarını tahmin eden bu algoritma bütün bir tahminleme işlemini tek bir regresyon problemi olarak ele almaktadır (Shinde ve ark., 2018). Söz konusu algoritma diğer görüntü tespitinde kullanılan derin öğrenme ve bilgisayarlı görü algoritmalarına nazaran tek seferde görüntü tespiti yaptığından YOLO Darknet v4 Mimarisi en güçlü algoritmalardan biri olarak ifade edilmektedir. Eş zamanlı nesne tespiti yapan YOLO Algoritmasının dördüncü versiyon Nisan 2020 de yayınlanmıştır. Yayınlanan bu model aynı zamanda COCO veri seti ile eğitilmiş halde yayınlamıştır.

Söz konusu algoritmanın mimarisi incelendiğinde Konvolüsyonel katmanlar ve Maxpool katman olarak adlandırılan en büyük değerin alınarak bir sonraki iterasyona iletildiği katmanlardan oluşmakta her bir katman modelin gücünü doğrudan etkilemektedir.



Şekil 3.32. YOLO Mimarisi (Shafiee ve ark., 2017).

YOLO V4 algoritması öncesinde geliştirilmiş olan YOLO v3. Algoritması Daknet-53 omurga ağına ve Fusion Feaute (FPN) ağına dayanan uçtan uca nesne tespiti yapan bir algoritmadır (Ju ve ark., 2019; Zhu ve ark., 2020). Darknet-53 omurga ağı nesne özelliklerinin çıkarılmasında ilk olarak kullanılmaktadır. Söz konusu ağ ana olarak beş ana bloktan oluşmaktadır. Her ana blok tekrar eden set halinde artık blok içermektedir. Batch Normalizasyon katmanı ve a Leaky ReLU aktivasyon fonksiyonu Con2d katmanı olarak her bir konvolüsyonel ağdan sonra gelmektedir. Ana YOLO nesne tanıma mimarisine göre YOLOv3 ün başlangıcını içermekte olup, çok daha güçlü bir omurga olan CSPDarknet53 içermektedir. İlave olarak YOLOv4 içerisinde spatial pyramid pooling (SPP) ve PANet boyun olarak kullanılmaktadır. Ayrıca Mish aktivasyon fonksiyonu, Mosaic veri güçlendirme ve DropBlock regülarizatörü ile iyileştirilmiş ve optimize edilmiştir.



Şekil 3.33. YOLO darknet v4 detaylı yapısı (Gao ve ark., 2021).

Ağın reseptive alanının genişletmek için, omurga sonrasına bir SPP ağı eklenmesi gerekmektedir. İlk olarak özellik çıkarım ağının çıktısı üç konvolüsyonel operasyona ve maksimize edilmiş 1*1, 5*5, 9*9 ve 13*13 havuzlama alanına bağlıdır. Sonrasında dört havuzlanan çıktı özellik haritasına sıralanmış ve son olarak konvolüsyonel boyutluluğunun azaltılması gerçekleştirilmiştir.

YOLOv3 MSE (ortalama karekök hatası) öngörülen kutunun merkezi ile gerçek kutunun Merkez koordinatları arasında genişlik ve yükseklik değerlerine göre kullanmaktadır. Bağımsız bir değişken olmasına karşın MSE kayıp fonksiyonu bilgiler arasındaki ilişkileri yansıtmada yanıltmaktadır. Bunu geliştirmek adına şekil, yön üst üste binme alanını göz önünde bulundurarak sınırlayıcı kutunun alanını dikkate alan IoU loss tavsiye edilmektedir (Huang ve ark., 2020; Kim ve ark., 2020). YOLOv4 algoritması MSE Loss yerine CloU Loss kullanmakta bu durumda şekil, yön ve aynı zamanda üst üstüne binen alanlar, merkezlere olan uzaklıklar ve öngörülen orantılar ortaya çıkmaktadır. Aşağıdaki denklemlerde ifade edilmiştir.

$$L_{CIoU} = 1 - IoU(A,B) + \frac{p^2(A_{ctr}, B_{ctr})}{c^2} + \propto * \lambda \tag{3.7}$$

$$IoU(A,B) = \frac{area\ (A \cap B)}{area\ (A \cup B)} \tag{3.8}$$

$$\propto = \frac{\lambda}{\big(1 - IoU(A,B)\big) + \lambda} \tag{3.9}$$

$$\lambda = \frac{4}{\pi^2}(arctan\frac{w^{gt}}{h^{gt}} - arctan\frac{w}{h})^2 \tag{3.10}$$

$$L_{ciou} = \Pr(object) * L_{CIoU} \tag{3.11}$$

$$L_{prob} = \Pr(object) * L_{SCE}(y,p) \tag{3.12}$$

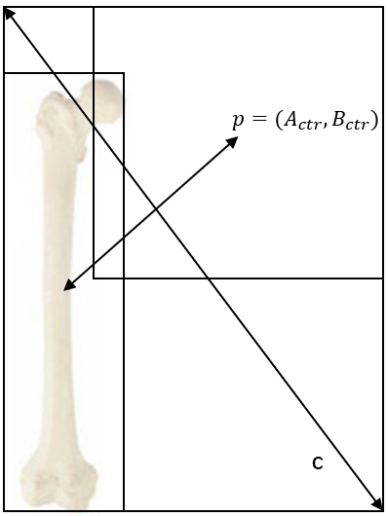
$$L_{conf} = \Pr(object) * L_{BCE}(y,p) + \big(1 - \Pr(object)\big) * L_{BCE}(y,p) * IgnoreMask \tag{3.13}$$

$$\lambda_{ciou} = 2 - (\frac{w^{gt} * h^{gt}}{w * h}) \tag{3.14}$$

$$Loss = \lambda_{ciou} * L_{ciou} + L_{prob} + L_{conf} \tag{3.15}$$

Denklem 3.7-3.15 arasındaki ifadelerde, "*B*" gerçek çerçeveyi temsil ederken, "*A*" tahmin çerçevesini temsil eder. Merkez noktaları arasındaki mesafe için ceza,

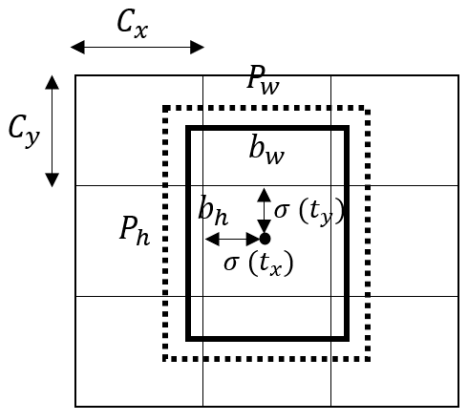
"$p^2(A_{ctr}, B_{ctr})$" olarak gösterilir. "A_{ctr}" tahmin çerçevesinin merkez noktası koordinatlarını, "B_{ctr}" gerçek çerçevenin merkez noktası koordinatlarını temsil eder. p() olarak gösterilen öklidyen mesafesi c'dir ve Şekil 3'te A ve B arasındaki mesafeyi temsil eder. En-boy oranı cezası "α * λ" olarak gösterilir, burada "α" pozitif bir katsayıyı temsil eder. "λ" ise en-boy oranının tutarlılığını ölçer. "gt" gerçek değerdir. "w^{gt}" ve "h^{gt}" gerçek çerçevenin yüksekliğini ve genişliğini göstermek için kullanılır. "w" ve "h" tahmin çerçevenin yüksekliğini ve genişliğini göstermek için kullanılır.



Şekil 3.34. Öklitsel uzaklık diagramı (Gao ve ark., 2021).

"λ_ciou", Loss değerinin dengelenmesinde önemli bir rol oynar. Böylece, bulunamayan nesnelerin sınırlayıcı kutularının gizliliğini sağlayarak konum kaybı ağırlığını artırır. "L_{SCE}" sigmoid çapraz entropiyi temsil eder. "L_{BCE}" ikili çapraz entropiyi temsil eder. y gerçek değeri temsil eder. p tahmin edilen değeri temsil eder. Gerçek çerçevenin merkez noktası tahmin çerçevesi içinde bulunuyorsa, tahmin çerçevesinin bir hedefi olduğu varsayılabilir. Bu durumda Pr(object) = 1, aksi halde Pr(object) = 0 olur. "$IgnoreMask$", hedef olmadığında tahmin çerçevesi ve gerçek çerçevenin "IoU" (Eşleşme üzerinde birlik) değerini hesaplar. En büyük "IoU" tahmin edilen ve gerçek değerin "IoU" olarak seçilir. Maksimum "IoU", önceden belirlenen eşik değerinin altına düştüğünde, denklem olarak kayıp fonksiyonu hesaplamasına dahil edilecektir. "L_{ciou}" location loss temsil eder. "L_{prob}" class loss temsil eder. "L_{conf}" confidence loss temsil eder (Gao ve ark., 2021).

Hedef pozisyonunun koordinatlarının hesaplanmasında denklem 3.16 - 3.23 nolu denklemlerden yararlanılmaktadır.



Şekil 3.35. Hedef objenin resimdeki pozisyonu.

YOLOv3 algoritmasında 3.16-3.19 denklemleri kullanılarak hedef nesnenin gerçek konumunu ve boyutunu belirlemek için kullanılır. Merkez noktanın bulunduğu alanın sol üst köşesi c_x ve c_y koordinatlarını gösterir. Ph ve Pw, referans noktasının yüksekliğini ve genişliğini gösterir. σ sigmoid fonksiyonunu temsil eder. $\sigma(t_x)$ ve $\sigma(t_y)$, tahmin edilen kutunun sol üst köşesi ile merkez arasındaki mesafeyi gösterir. "tw" ve "th", tahmin edilen genişlik ve yükseklik ofsetini ifade eder. "b_{x_YOLOv3}" ve "b_{y_YOLOv3}", konumu bulmak için kullanılırken, "b_{w_YOLOv3}" ve "b_{h_YOLOv3}" " hedef nesnenin boyutunu bulmak için kullanılır.

$$b_{x_YOLOv3} = \sigma(t_x) + c_x \tag{3.16}$$

$$\mathrm{b_{y_YOLOv3}} = \sigma(\mathrm{t_y}) + \mathrm{c_y} \tag{3.17}$$

$$b_{w_YOLOv3} = Pwe^{tw} \tag{3.18}$$

$$\mathrm{b_{h_YOLOv3}} = \mathrm{Phe^{th}} \tag{3.19}$$

Bu durum YOLOv4'te geliştirilmiş olup, 3.20 denklemi kullanılmaktadır. YOLOv3'te tahmin kutusunun merkez nokta koordinatlarını hesaplamak için kullanılan yöntem, 0 ile 1 arasında bir değer döndüren bir işlevi temsil etmektedir. Ancak, c_x veya $c_x + 1$ 'in kesin değerini elde etmek zor olabilir ve bu da tahmin çerçevesinin merkezinin ızgara sınırıyla hizalanmamasına neden olabilir. YOLOv4'te, bu sorunu çözmek için 1.0'dan büyük bir katsayı olan β tanıtılmıştır (Gao ve ark., 2021).

$$b_{x_YOLOv4} = \beta\sigma(t_x) - (\beta - 1)/2 + c_x \tag{3.20}$$

Sonuç olarak, YOLOv4 algoritmasında hedef nesnenin gerçek konumunu ve boyutunu belirlemede "b_{x_YOLOv3}", "b_{y_YOLOv3}", "b_{w_YOLOv4}" ve "b_{h_YOLOv4}" aracılığıyla 3.21-3.23 denklemleri kullanılmaktadır.

$$b_{x_YOLOv3} = \beta\sigma(t_x) - (\beta - 1)/2 + c_x \tag{3.21}$$

$$b_{y_YOLOv3} = \beta\sigma(t_y) - (\beta - 1)/2 + c_y \tag{3.22}$$

$$b_{w_YOLOv4} = Pwe^{tw} \tag{3.23}$$

3.6.4. Görüntü sınıflandırma eğitimi çeşitleri

Görüntü sınıflandırma eğitimi çeşitleri derin öğrenme modellere beslenen resimlerin içerdikleri objeleri veya özellikleri doğru bir şekilde sınıflandırmasını hedeflemektedir. Görüntü sınıflandırma eğitimleri sıfırdan eğitim ve aktarımlı öğrenme olarak iki temelli yaklaşımda toplamak uygun olacaktır.

3.6.4.1. Sıfırdan eğitim (Training from scratch)

Sıfırdan eğitimde derin öğrenme modelinin daha önce herhangi bir eğitime tabi tutulmamış olması bu sebeple temel ağırlıklarının dahi bulunmadığı durumda kullanılmaktadır. Bu eğitimde model sıfırdan başlayacağı için öncelikle rastgele başlangıç değerleri kullanır ve her işlem adımında günceller. Sıfırdan eğitimin kullanılması için oldukça büyük veri setleri ve yeterli hesaplama imkânı olan durumlar kullanılır. Eğitim süreleri sıfırdan eğitimde oldukça uzun olması beklenmektedir. Sıfırdan eğitimde bütüncül bir şekilde yepyeni bir problem için oldukça özelleştirilmiş bir model üretilir. İlgili veri setinde sadece problem alanına özgü bilgiler öğrenebilir (MathWorks, 2018).

3.6.4.2. Aktarımlı öğrenme (Transfer learning)

Aktarımlı öğrenmede çoğunlukla daha öncesinde büyük veri setlerinin kullanılması ile eğitimi gerçekleştirmiş ve genel obje özelliklerini öğrenmiş bir modelin farklı bir nesne ya da özelleşmiş alanı tanıması ya da tespit edilmesi noktasında adapta edilmesini içermektedir. Örneğin daha öncesinde COCO – Common Object in Context

veri seti ile eğitilmiş bir modelin sonra hedeflenen nesne üzerinde yeniden eğilerek o nesne özelliklerini çıkarmak için öğrenmesi aşamalarını kapsamaktadır. Aktarımlı öğrenmede daha az veri seti ile başarılı sonuçlar elde etmek mümkün hale gelir. Eğitim süreci çoğunlukla kısalmaktadır. Lakin hedeflenen nesneye ait veri sayısı azsa öğrenim süreci sonrası elde edilen özellikler uygun olmayabilir. Daha önce eğitilmiş modelin ilgili nesne alanında ne kadar uyumlanabileceği kestirilemeyebilinir (MathWorks, 2018).

3.6.5. Görüntü sınıflandırma eğitimi ve kullanılan veri setleri

Görüntü sınıflandırma eğitimleri sırasında algoritmaların daha önce eğitildiği ya da algoritmaların başarısını test etmek amacıyla birçok veri seri kullanılmaktadır. Bunların başlıcaları aşağıda yer alan alt bölümlerde irdelenmiştir.

3.6.5.1. ImageNet veri seti

ImageNet veri seti derin öğrenmenin gelişmesinde oldukça katkı sağlamış ve ilham vermiş bir veri setidir. Bu veri setinde toplam 1,3 milyon veri eğitimde kullanılmak üzere bulunmaktadır. Toplamda 1000 sınıf bulunmaktadır. Söz konusu veri setinde 100.000 test verisi bulunmakta ve toplamda 50.000 doğrulama verisi yer almaktadır. Bu bağlamda ImageNet veri setinin oldukça geniş bir görüntü veri seti olduğunu söylemek mümkündür (Massouh ve ark., 2017).

Şekil 3.36. ImageNet Veri Seti (Massouh ve ark., 2017).

ImageNet veri setinin yaygın hale gelmesi 2009 yılı itibariyle düzenli olarak gerçekleştiren ImageNet Large Scale Visual Recognition Challenge (ILSVRC) yarışmasıdır. ImageNet'in bir diğer önemli özelliği çok çeşitli bakış açılarını, sahneleri ve nesne tiplerini içermektedir. Aynı zamanda çok çeşitli zorluk derecelerine sahip görüntüleri barındırmaktadır. Bu veri seti içerisinde yer alan görüntüler hali hazırda etiketlidir.

3.6.5.2. MS COCO veri seti

Microsoft Common Objects in Context veri seti derin öğrenme alanında derin öğrenme modellerinin eğitiminde en çok kullanılan veri setleri arasında başı çekmektedir. Söz konusu veri seti 328.000 resim ve 2 Milyon üstünde etikete sahip nesne içermektedir. MS COCO veri seti ImageNet ile kıyaslandığında 91 sınıftan oluştuğu için daha az sınıfa sahiptir lakin bir sınıf içerisinde daha fazla görüntü içermektedir.

Şekil 3.37. MS COCO Veri Seti (Lin ve ark., 2014).

MS COCO veri setinde sadece nesneler odakta tutulmamış çevresel bağlamlarına da yer verilmiştir. Veri çeşitliliğinde oldukça büyük olanaklar sağlayan MS COCO veri seti modelin farklı nesneleri ve sahneleri algılamasında olanak sağladığı gibi diğer veri setlerine göre de ön plana çıkmasını sağlar. Çok çeşitli zorluk seviyesine sahip

görüntüleri içermektedir. MS COCO veri setinin bir diğer farklılaşan özelliği faaliyet halinde bulunan insanlar da yer almaktadır (Lin ve ark., 2014).

3.6.5.3. MNIST veri seti

Modified National Institute of Standards and Technology database olarak adlandırılan MNIST veri seti aslında lise öğrenimi içerisindeki öğrenciler ve Amerikan sayım bürosu bünyesinde çalışan kişilerce el yazısı kullanılarak yazılmış rakamların bir araya getirilmesi ile oluşmaktadır. Söz konusu yazılmış rakamların görsellerinin bir diğer özelliği ise siyah-beyaz olmak üzere 28x28 pikselden oluşmaktadırlar. MNIST veri seti incelendiğinde 60.000 eğitim ve 10.000 test olmak üzere toplamda 70.000 veriyi bünyesinde barındırmaktadır. Söz konusu veri seti örnek sayısı ve sınıf sayısı olarak değerlendirildiğinde başlangıç seviyesinde eğitimlerde ve derin öğrenme modellerinde kullanılmaktadır. Veri seti içerisinde 0 ve 9 rakamları dahil olmak üzere 0 ile 9 arasında rakamlar bulunur. Veri seti toplamında 10 farklı sınıf yer almaktadır. Oldukça temel ve basit bir veri setidir (Lecun ve ark., 1998).

Şekil 3.38. MNIST Veri Seti (Lecun ve ark., 1998).

3.6.5.4. CIFAR-10 / CIFAR-100

Hem CIFAR 100 hem de CIFAR 10 veri setleri görüntü sınıflandırmada kullanılan popüler veri setleridir. Canadian Institute For Advanced Research olarak geçen CIFAR veri seti 10 ve 100 olarak ikiye ayrılmaktadır.

CIFAR-10 veri seti 50.000 eğitim ve 10.000 test verisi olmak üzere toplamda 60.000 görüntüden oluşmaktadır. Bütün görüntülerin boyutu 32x32 piksellerden oluşmaktadır. Toplamda 10 sınıf yer almaktadır.

CIFAR-100 veri seti yine 50.000 eğitim ve 10.000 test verisi olmak üzere toplamda 60.000 görüntüden oluşmaktadır. Bütün görüntülerin boyutu 32x32 piksellerden oluşmaktadır. Toplamda 100 sınıf yer almaktadır. İçerisinde özellikle özgün ve ayrıntılı konseptleri içermektedir. Görece daha basit olup araştırmacılar için güzel bir başlangıç noktasıdır.

Şekil 3.39. CIFAR-10 Veri Seti (ImageNet, 2016).

3.6.6. Veri arttırma teknikleri

Veri arttırma teknikleri derin öğrenme modellerinin nesne tespitini daha doğru yapabilmeleri ve daha yetkin olabilmeleri için veri setindeki örnekleri oldukça çoğaltılmasını sağlayan bir tekniktir. Modellerin farklılaşan örnekleri görebilmesi aynı zamanda kısıtlı veri setlerinin çoklandırılmasını sağlayabilir. Veri çoklandırma veri sayısını arttıracağı için model sıfırdan da öğrense, aktarımlı öğrenmeye de tabi tutulsa performansı daha yüksek olacaktır (Alimovski ve Erdemir, 2021).

Veri çoklandırmada birçok teknik kullanılmasına karşın en yaygın olanlar aşağıda yer verilmiştir:

- Dikey ve yatay döndürme: elde edilen görüntülerin dikey yahut yatay olarak döndürülmesi sonucunda yeni görüntüler elde edilerek verilerin çoklandırılmasıdır. Bu işlemde resimlerin simetrisi alındığı için nesnelerin de farklı simetrilerini de algoritmanın öğrenmesi sağlanmaktadır.
- Rastgele kırpma: görüntülerin rastgele bölümlendirilmesi ile veri çoklandırılması yöntemidir.
- Rastgele döndürme: görüntünün saat yönünde ya da saat yönünün tersinde belirli açılarda döndürülmesi ile gerçekleştirilen veri çoklandırma yöntemidir.
- Parlaklık ayarlama: görüntünün parlaklığının azaltılması ya da arttırılması ile ortaya çıkan veri arttırım tekniğidir. Bu sayede test verileri arasına karanlık ortamda ya da daha aydınlık ortamda görüntülenmiş nesnelerin tespit edilmesine olanak sağlanmaktadır.
- Renk düzenleme: görüntünün renk doygunluğunu, renk tonunu yahut kontrastının oynanması ile gerçekleştirilen veri arttırım tekniklerdir.
- Rastgele gürültü eklenimi: görüntüye belir bir düzen dahilinde olmayacak şekilde gürültü eklenmesi sayesinde elde edilen veri çoklandırması tekniklerdir.

4. ARAŞTIRMA BULGULARI

Kalça kırıklarının tespitinde hem derin öğrenme ve bilgisayarlı görünün en güçlü algoritmaları hem de uzman hekim ve pratisyen hekimleri bir araya getiren çalışmanın bu kısmında; araştırma çıktıları ve performans değerlendirme kriterlerinin belirlenmesi, araştırmada kullanılan derin öğrenme ve bilgisayarlı görü algoritmalarının ve hekim gruplarının belirlenmesi, veri setlerinin oluşturulması ve özellikleri, çalışmada kullanılan yazılım dili ve yapıların seçimi ve araştırma sonuçlarının yorumlanmasında kullanılan ileri istatistiksel yöntemlerden bahsedilmiştir.

4.1. Araştırma Çıktıları ve Performans Değerlendirme Kriterlerinin Belirlenmesi

Çalışmanın gerçekleştirileceği YOLO Darknet v4, Faster R-CNN Inception v2, SSD Mobilenet - v2 algoritmaları olmak üzere derin öğrenmenin en güçlü sayılabilecek üç algoritması, uzman doktorlar ve pratisyen hekimler olmak üzere iki hekim grubunu performanslarını objektif bir şekilde ortaya koymak, başarı seviyelerini belirlemek ve birbiri ile kıyaslamak adına üretilen çıktıların hangi performans değerlendirme kriterlerine göre değerlendirileceği hususu büyük önem arz etmektedir. Bu nedenle karmaşıklık matrisi ve çoklu performans değerlendirme tablosu, ROC Eğrisi, Loss oranı grafiklerinden bahsedilmiştir.

4.1.1. Karmaşıklık matrisi ve çoklu performans değerlendirme tablosu

Karmaşıklık matrisi bir sistemin ürettiği sonuçlar itibariyle tahmin değerleri ile gerçekleşen değerler arasında kıyaslamaya imkân sunan bir yapıdır. Karmaşıklık matrisinde veri setinde bulunan değerler gerçekteki durumuna karşılık, çalıştırılan algoritmanın çıktısına göre hangi bölümde yer alacağı kıyasını ortaya koymaktadır (Polat ve ark., 2017).

Tablo 4.1. Karmaşıklık matrisi.

	Gerçekte Pozitif	*Gerçekte Negatif*
Pozitif Tahminleme	TP Doğru Pozitif	FP Yanlış Pozitif
Negatif Tahminleme	FN Yanlış Negatif	TN Doğru Negatif

Doğru Pozitif (TP): Gerçekte yer alan nesnenin tahminde de doğru sınıflandırıldığını göstermektedir. Doğru Negatif (TN): Gerçekte var olmayan nesnenin tahminde de yine var olmadığı şeklinde sınıflandırıldığını göstermektedir. Yanlış Negatif (FN): Gerçekte yer alan nesnenin tahminde de doğru sınıflandırılmadığını göstermektedir. Yanlış Pozitif (FP): Gerçekte var olmayan nesnenin tahminde var olduğu şeklinde sınıflandırıldığını göstermektedir.

Doğruluk oranı hem eğitimi gerçekleştirilen modellerin hem de çalışma gruplarının hem ilgili kemik kırığı bulunan hem de bulunmayan görsellerde ne kadar bir %'lik başarı sağladığını göstermektedir. Aynı zamanda 1- Hata Oranı olarak düşünülebilir. (Shafiee ve ark., 2017; Polat, Mehr ve Cetin, 2017).

Duyarlılık olarak ifade edilen doğru pozitif oranı modelin X-ray görüntülerinde femur üst bölge kırığına sahip olan görüntülerde kırığın hem çalışma gruplarınca hem de algoritmalarca ne oranda doğru tespit ettiğini göstermektedir.

Özgüllük olarak belirtilen doğru negatif oranı ise içerisinde kırık femur üst bölge olmayan femur üst bölge görüntülerinde ne kadar modelin başarılı olduğunu göstermektedir. Yanlış pozitif oranı 1- doğru negatif oranına eşit olduğu ifade edilebileceği gibi femur üst bölge kırığı içermeyen görüntülerde modelin ne kadar yanılgıya düştüğü yahut ne oranda yanıldığı olarak da ifade edilebilmektedir. Yanlış negatif oranı ise 1-doğru pozitif oranı olarak düşünebileceği gibi femur üst bölge kırığı içeren görüntülerde modelin ne kadar yanılgıya düştüğü olarak da gösterilebilmektedir. Doğru tahmin oranı modelin femur üst bölge kırığı var dediği görüntülerde ne kadar başarılı olduğunu göstermektedir. F1 puanı ise doğru pozitif oranı ile doğru tahmin oranının harmonik ortalamasını göstermektedir. Hata oranı 1-doğruluk oranıdır. Yani modelin hem femur üst bölge kırığı olan görüntülerde hem de

olmayan görüntülerde ne kadar başarısız olduğunu ortaya koymaktadır (Polat, Mehr ve Cetin, 2017). Birçok değerlendirme kriteri tek bir tablo olarak oluşturularak çoklu değerlendirme kriteri tablosu oluşturulmuştur.

Tablo 4.2. Çoklu değerlendirme kriterleri tablosu (Zeren ve ark., 2020).

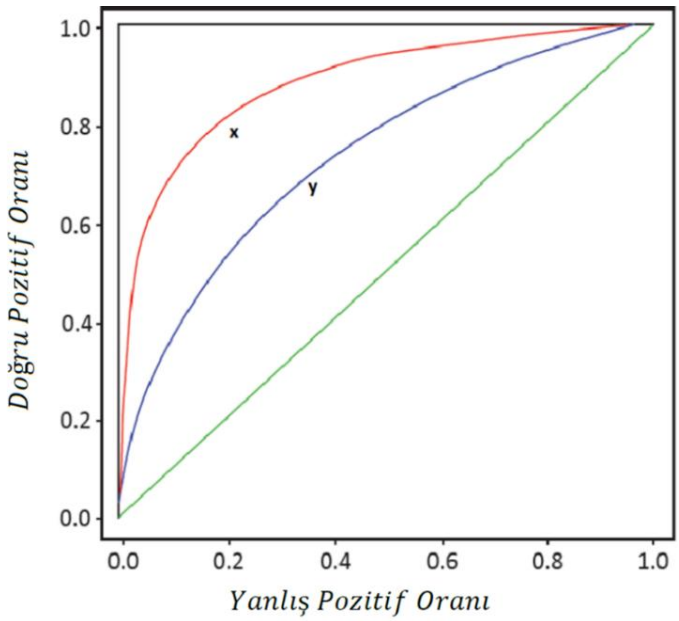
Kriterler	Formülleri
Doğruluk Oranı (ACC)	(TP+TN)/(N+P)
Duyarlılık - True Positive Rate (TPR)	TP/(TP+FN)
Özüllük- True Negative Rate (TNR)	TN/(TN+FP))
Yanlış Pozitif Oranı (FPR)	FP/(FP+TN)
Yanlış Negatif Oranı (FNR)	FN/(FN+TP)
Doğru Tahmin Oranı - Precision (PPV)	TP/(TP+FP)
F1 Puanı (F1)	2*((PPV*TPR)/(PPV+TPR))
Hata Oranı (ERR)	(FP+FN)/(N+P)

4.1.2. ROC eğrisi

Tahminlemelerin olduğu ve sınıflandırmanın olduğu sonuç üreten sistemler için kullanılmakta olan bir diğer performans değerlendirmede kullanılan yöntem "Alıcı İşlem Karakteristiği" olarak geçen ROC'dir (Lasko ve ark., 2006). Bu değerlendirme metodunda hem üretilen sonuçların doğru pozitif oranı hem de yanlış pozitif oranı kullanılmaktadır. Doğru Pozitif Oranı y eksenine, Yanlış Pozitif Oranı ise x eksenine konumlandırılmıştır. Örnek olması açısından aşağıda yer verilen ROC eğrisi örneğinde farklı eşik değerlerine sahip x ve y eğrileri yer almaktadır (Metz, 2006; Fawcett, 2006).

Elde edilen sonuçların Doğru Pozitif Oranı ve Yanlış Pozitif Oranı değerleri yerleştirildiğinde sol üst köşeye noktaya en yakın aynı zamanda diagonal eğrinin üstünde konumlanan sonuçlara sahip sistemlerin performansının daha başarılı olduğu kabul edilmektedir. Aşağıda ROC örneği yer verilen grafikte kırmızı ile çizilmiş olan x eğrisi sol üst köşeye mavi ile çizilen y eğrisine göre daha yakın konumlandığından

daha başarılı sonuçlar ürettiği kabul edilmektedir (Dirican, 2001; Lane ve Gantley, 2017).



Şekil 4.1. ROC eğrisi analizi örneği (Lane ve Gantley, 2017).

4.1.3. Loss oranı grafikleri

Kayıp (Loss) Oranı: Loss rate olarak ifade eden kayıp oranı her bir iterasyona sonucunda öğrenmeler ile ürettiği çıktılardan yola çıkarak düştüğü yanılgıların ortalamasıdır. Yani loss değeri geliştirilen modelin tahminlerinden ne kadar uzak olduğunu ölçmektedir. Kayıp oranına bakılarak veri beslemesi ve çıktı alınan modelin her bir işlem adımındaki davranışını ortaya koymaktadır. Beklenti ilgili algoritmanın her bir iterasyona sonucunda kayıp değerinin azalarak "0" erişmesi hedeflenmektedir (Chen, 2017). Yapılan iterasyonlar sonucunda elde edilen tahmin ile reel değer arasındaki fark loss fonksiyonu ile hesaplanmaktadır. Eğer modelimiz ya da algoritmamız güçlü değil yahut yanlış bir kurgu ile oluşturulmuşsa tahmin değer ile reel değer arasında farkın fazla olması beklenmektedir. Bunun aksine model güçlü ve doğru bir kurgu ile oluşturulmuşsa kayıp değerinin düşük çıkması ve gitgide düşmesi beklenmektedir. Loss fonksiyonları algoritmalarının ürettiği çıktıları normalize ederek sonuç üretmektedir. Günümüzde performans değerlendirmesinde çoğunlukla Multiclass Support Vector Machine (SVM), Sigmoid ve Softmax verilebilmektedir.

Örnek olarak sınıflandırma görevlerinde kullanılan loss fonksiyonlarından en yaygın olanlarından biri çapraz entropi kaybı olarak bilinmektedir. Gerçek sınıf etiketleri ile tahmin edilen sınıfların arasındaki olasılık farklılıklarını ölçmektedir. Bu fonksiyonda

“N” değeri veri noktalarının sayısı, “C” sınıf sayısı, "y_{ij}" i. veri noktasının j. sınıf etkiketi (0 veya 1 olmak üzere), "y'_{ij}" i. veri noktasının j. sınıfın tahmin edilen olasılığı olarak yer almak üzere formül 3.18 ile hesaplanmaktadır (Ho, 2019).

$$Cross - Entropy\ Loss = -\frac{1}{N}\sum_{i=1}^{N} y_{ij} Log(y'_{ij}) \tag{3.24}$$

4.2. Araştırmada Kullanılan Algoritmalar ve Hekim Gruplarının Seçilmesi

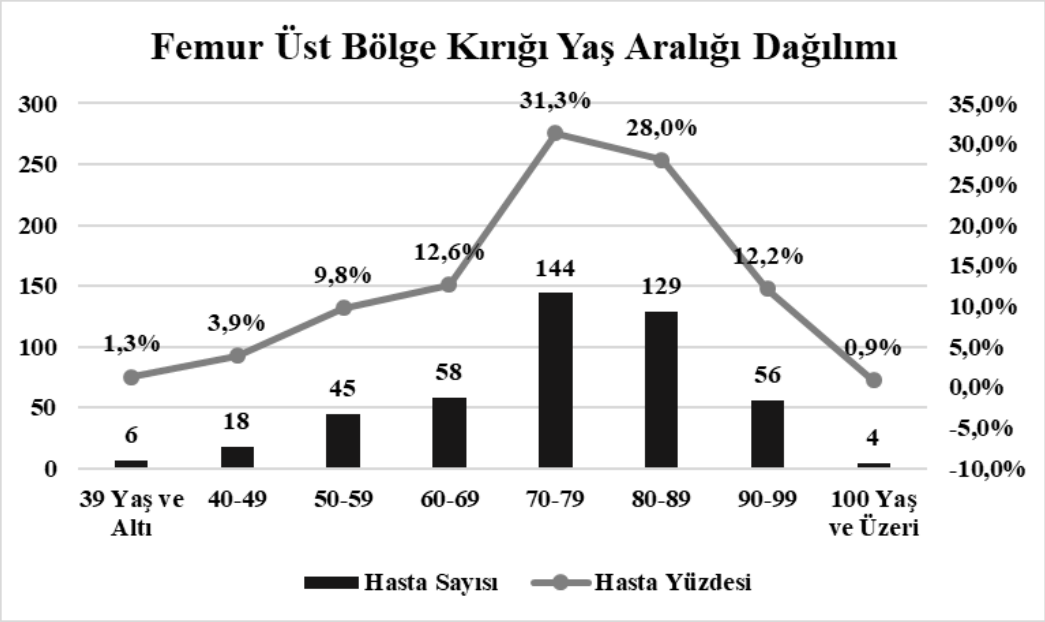
Araştırmada derin öğrenmenin en güçlü algoritmaları olan ve birçok farklı görüntü işleme veri seti tarafından performansı ortaya konulmuş güncel algoritmalar seçilmiştir. Bu algoritmalar önceki bölümlerde detaylı olarak hem mimarileri hem de görüntü işleme biçimleri açıklanmış olan tek seferde bütüncül olarak kritik bölge tespiti yapabilen YOLO – You look only once - Darknet V4 algoritması, SSD – Single shot multibox detector – Mobilenet v2 algoritması ve R-CNN algoritmalarının en güçlü olanı Faster R-CNN – Inception v2 algoritması kullanılmıştır. Söz konusu algoritmalar aktarımlı öğrenim tekniği kullanılarak femur üst bölge kırık tespit edebilir hale getirilmiştir. Aktarımlı öğrenim gerçekleştirilmeden önce her üç algoritmanın da daha öncesinde COCO (Common Object in Context) veri setindeki bütün görselleri tanıyabilecek halde eğitimden geçirilmiş versiyonları kullanılmıştır. Her bir algoritmanın en optimum şekilde birebir aynı veri setleri ile eğitiminin gerçekleştirilmesi için ortam özenle seçilmiştir.

Hekim grupları seçilirken gerçekten her bir derin öğrenme algoritmasının gerçek başarısını ortaya objektif bir şekilde koymak adına iki farklı hekim grubu belirlenmiştir. Bunlardan ilk grup Türkiye’nin en büyük metropolünde bulunan beş yıllık tıp eğitiminin tamamlamış içerisinde pratisyen hekimler, aile hekimleri ve acil hekimleri içerisinden seçilmiştir. Özellikle bu grup oluşturulurken tıp eğitimlerini son 1 yıl içerisinde bitirmiş olmalarına dikkat edilmiş. Teorik ve pratik bilgilerinin taze olmasına dikkat edilmiştir. Söz konusu çalışma grubu 10 kişilik hekim grubunun bir araya getirilmesi ile oluşturulmuştur. İkinci hekim grubu yine Türkiye’nin en büyük metropolünde yer alan en büyük eğitim ve araştırma hastanelerinden birinin ortopedi ve travmatoloji bölümü odağa alınmıştır. Söz konusu eğitim ve araştırma hastanesi ortopedi uzmanları ve ortopedi servisi asistanlarından oluşan toplamda 12 kişilik alanında Türkiye’nin sayılı ortopedi uzmanlarının da bulunduğu hekim grubu

oluşturulmuştur. Söz konusu hekimlerin tamamı gönüllü olarak bu çalışmada yer almışlardır.

4.3. Araştırmada Kullanılan Veri Setlerinin Oluşturulması ve Özellikleri

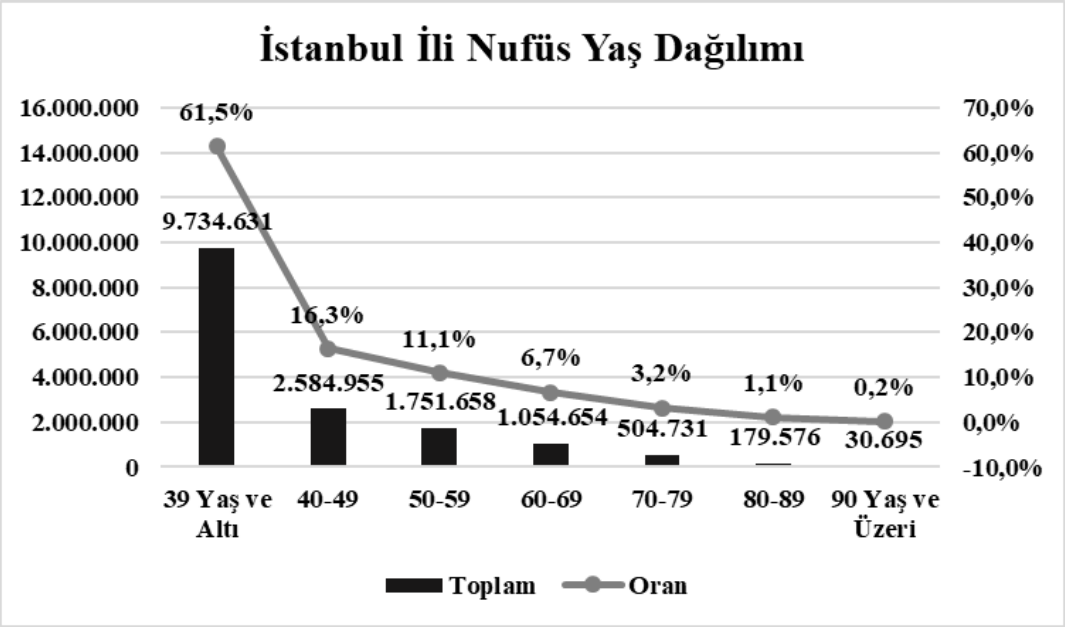
Eğitim ve doğrulama veri setinin oluşturulması adına İstanbul'da yer alan Eğitim ve Araştırma hastanesine yaşadığı travma sonrasında ortopedi yahut acil kliniğine başvurmuş 500 üzerinde hastanın X-raya görüntüleri incelenmiş. Bunlardan üzerinde bozulmalar ve problemler olmayan X-ray görüntüleri ayrıştırılmıştır. Toplamda 410 hastanın femur üst bölge x-ray görüntülerini içeren veri seti eğitim ve doğrulama veri setini oluşturmak amacıyla ele alınmıştır. 410 hastaya ilave olarak bağımsız olarak düzenlenen test veri seti için 50 adet femur üst bölge kırığına sahip X-ray görüntüsü ele alınmıştır. 460 hastadan oluşan bu küme incelendiğinde aşağıdaki şekil 4.2'de görüleceği üzere son 3 yılda başvuran hastalarının yaş dağılımı gözükmektedir. Buna göre 60 ile 100 üzeri yaş arası olan grup toplam femur üst bölge kırığı nedeniyle hastaneye kabul edilen kümenin %85'ine denk geldiği gözükmektedir. 39 yaş ve altı grup ise toplam popülasyonun sadece %1,3'sına denk geldiği gözlemlenmektedir. Özellikle söz konusu popülasyonun 70 yaş üzerine gelindiğinde agresif bir şekilde arttığı gözlemlenmektedir.



Şekil 4.2. Femur Üst Bölge Kırıklarının Yaş Aralığı Dağılımı.

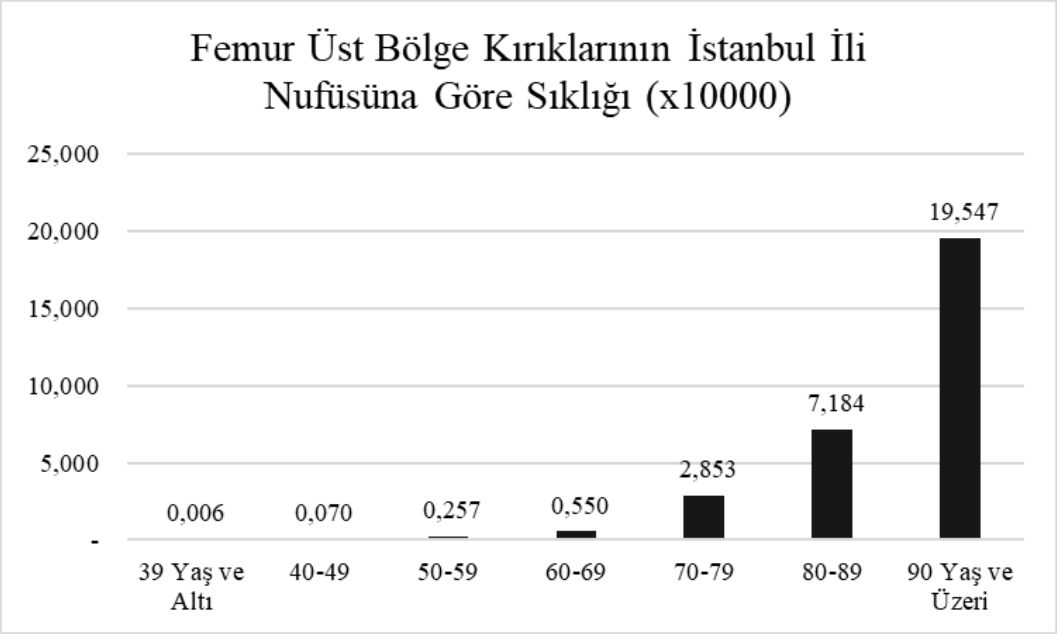
Kıyaslamak ve yorumlayabilmek adına İstanbul'da bulunan bu eğitim ve araştırma hastanesi için söz konusu metropolde yaşayan insanların yaş aralığı dağılımları

incelendiğinde 60- 100 yaş üzeri grup toplam popülasyonun %11,2 seviyesinde kaldığı, 39 yaş ve altı grubun toplam İstanbul nüfusunun %61,5'ine denk geldiği Türkiye İstatistik Kurumunun yapmış olduğu çalışma ile gözükmektedir (TUIK, 2022).



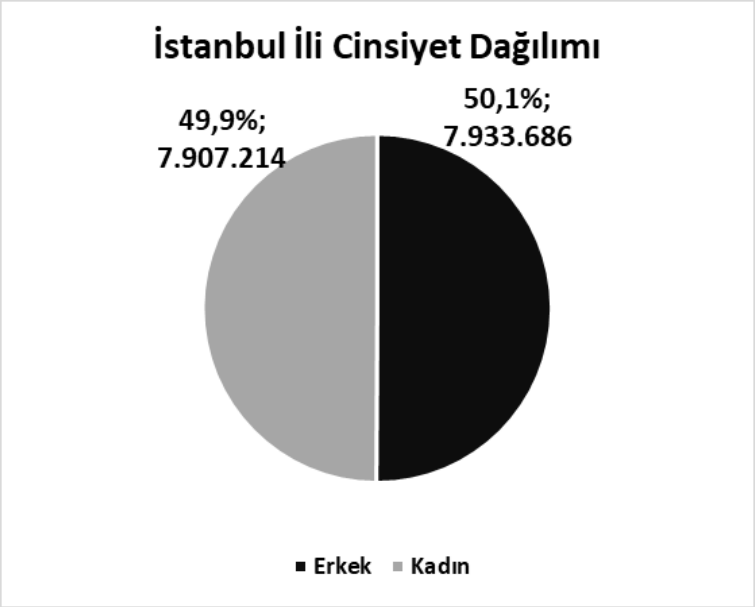
Şekil 4.3. İstanbul İli Nüfus Yaş Dağımı.

Söz konusu femur üst bölge kırığı rahatsızlığı gözlenen kümenin yaş grupları İstanbul İl nüfusuna göre sıklığına bakılacak olursa şekil 4.4'te görüleceği üzere 90 yaş üzeri insanlarda söz konusu rahatsızlığın ortaya çıkması 39 yaş altı aralığına göre 3257 kat daha fazla olduğu gözükmektedir. 39 yaş altı hastalarda sıkılık 10.000 kişide 0,006 iken 90 yaş üzeri hastalarda 10.000 kişide 19,5 kişidir. Yine 39 yaş altı hastalar ile 80-89 yaşa altı hastalar kıyaslandığında 1.183 kat daha fazla rastlanma sıklığına sahip olduğu; 70-79 yaş arası hastalarda ise 466 kat daha fazla olduğu; 60-69 yaş arasındaki hastalarda ise 91 kat daha fazla olduğu; 50-59 yaş arası hastalarda ise 42 kat daha fazla olduğu; 40-49 yaş arası hastalarda ise 11 kat daha fazla olduğu gözükmektedir. Yaş aralığının ilerlemesi hem söz konusu nüfusun azlığı hem de vaka sayısının artması ile sıklığın çok daha fazla arttığı gözükmektedir.



Şekil 4.4. Femur Üst Bölge Kırıklarının İstanbul İli Nüfusuna Göre Sıklığı.

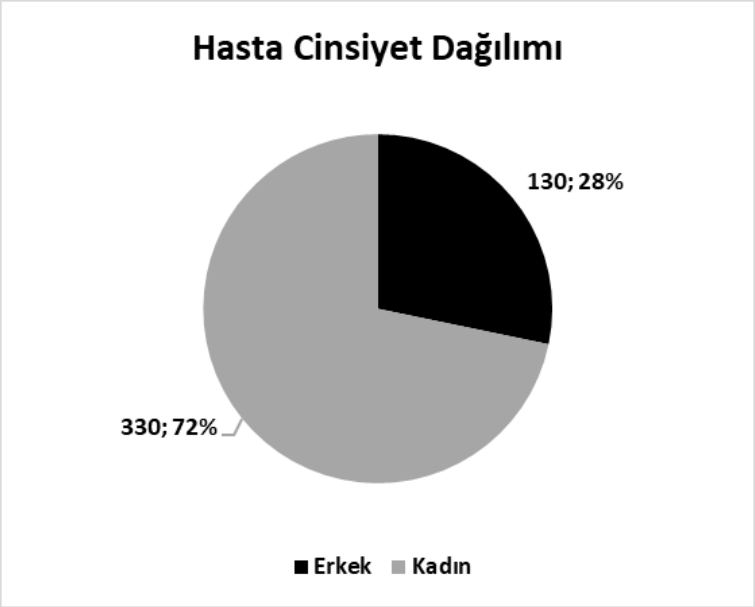
Femur üst bölge rahatsızlığı gözlemlenen kümenin cinsiyet açısından analiz edildiğinde ve İstanbul il nüfusu cinsiyeti kıyaslandığında şekil 4.5 ve şekil 4.6 ortaya çıkmaktadır. İstanbul ili cinsiyet dağılımının neredeyse aynı olduğu toplam nüfusun %50,1'inin kadın iken %49,9'unun erkek olduğu gözükmektedir.



Şekil 4.5. İstanbul İli Cinsiyet Dağılımı.

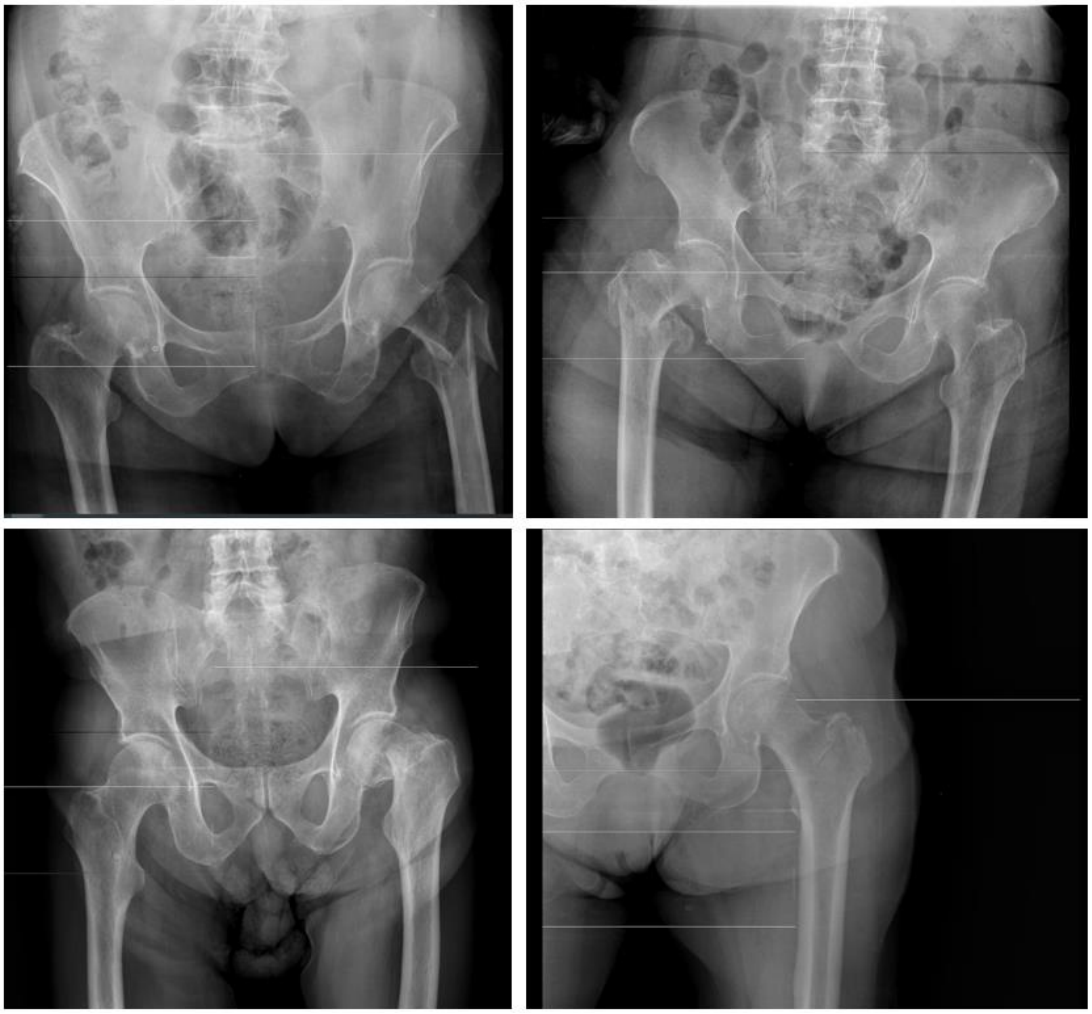
Son 3 yılda İstanbul'da yer alan Eğitim ve Araştırma hastanesine kalça kırığı şikâyeti ile başvuran hastaların cinsiyet dağılımı oranı incelendiğinde, incelenen hataların %28'ini erkek iken, %72'si kadın olduğu gözükmektedir. Buna göre kadınlarda femur

üst bölge kırıklarının erkeklere oranla 2,5 kat daha fazla gözlemlendiği yapılan araştırmanın bir bulgusu olarak görülmektedir.



Şekil 4.6. Hasta Cinsiyet Dağılımı.

Eğitim ve değerlendirme veri seti 410 hastadan alınan Femur üst bölge X-ray görüntülüleri veri çoklandırması tekniklerinden döndürme ve ölçeklendirme kullanılarak toplamda 820 adet görüntü elde edilecek şekilde iki katına çıkarılmıştır. Bunun sonucunda resimlerin çoğunluğunda 2 adet femur üst bölgesi olmak üzere algoritmanın transfer öğrenme ile tekrar eğitilmesi planlanan eğitim ve test veri setleri için 1514 femur üst bölge görüntüsü hazırlanmıştır. Söz konusu veri seti %80 eğitim ve %20 doğrulama olmak üzere ayrıştırılmıştır. Eğitim veri seti için 656 x-ray görüntüsü ayrılırken diğer bir yandan 164 adet X-ray görüntüsü doğrulama veri seti için ayrılmıştır. Söz konusu X-ray görüntüleri öncelikle uygun formatlara dönüştürülmesi sağlanmıştır.



Şekil 4.7. Eğitim veri seti örnekleri.

Şekil 4.8. Doğrulama veri seti örnekleri.

Modelin eğitim ve test veri setlerinde ezbere düşmemesi yahut önyargı oluşturmasına engel olmak objektif değerlendirmenin sağlanması adına eğitim ve doğrulama veri setinden tamamen bağımsız yeni bir veri seti oluşturulmuştur. Söz konusu test veri seti 55 farklı hastadan içerisinde 50 adet femur üst bölge kırığı içeren 50 adet femur üst bölge kırığı içermeyen X-ray görüntüleri derlenmiştir. Söz konusu hastalardan alınan X-ray sadece bir femur üst bölgesi ya da iki femur üst bölgesi içermektedir. Test veri setindeki görüntülerde en az 1 femur üst bölge kırığı içermesi beklenirken 5 hastada hiç kırık yer almayacak şekilde toplamda 100 femur üst bölgesi derlenmiştir.

Şekil 4.9. Test veri seti örnekleri.

4.4. Çalışmada Kullanılan Yazılım Dili ve Yapılar

Çalışmada hem YOLO – You look only once- Darknet V4 algoritması, hem SSD – Single shot multibox detector – Mobilenet v2 algoritması hem de R-CNN algoritmalarının en güçlü olanı Faster R-CNN – Inception v2 algoritması için eğitime

başlaması için her iki veri seti için de kırık femur üst bölgesinin koordinatları text formatında yeniden işaretlenmiştir.

Şekil 4.10. Text formatında kırık bölgelerin işaretlenmesi.

Söz konusu 3 algoritma olan YOLO Darknet V4 algoritması, SSD Mobilenet v2 Faster R-CNN Inception v2 algoritmalarının tamamı için hem kütüphane seçimi hem de yazılım dilinin yaygın kullanımı nedeniyle Python yazılım dili tercih edilmiştir. Idle yani kullanıcı web arayüzü seçimi sırasında kullanıcıya büyük kolaylık ve kısa yollar tanıyan Spider ve Pycharm arayüzleri seçilmiştir.

Bu çalışmada özellikle YOLO Darknet V4 algoritmasının herhangi bir iş istasyonu CPU ya da GPU'su ile eğitim sürecinin oldukça uzaması öngörüsü ile hizmet olarak yazılım (SaaS) bulut teknolojisi tercih edilmiştir. Bulut teknolojisi tercih edilirken çalışmada Google Colabratory'e ait dünyanın en güçlü GPU'larından biri olan Tesla K80 24 GB GDDR5 GPU ile modelin yeniden eğitilmesi hedeflenmiştir. SSD Mobilenet v2 Faster R-CNN Inception v2 algoritmalarının yeniden eğitimi için güçlü kabul edilebilecek bir iş istasyonu olan Intel(R) Core(TM) i5-8300H CPU @ 2.30GHz – 16 GB Ram – GTX1050 GPU tercih edilmiştir.

YOLO Darknet V4 algoritması, SSD Mobilenet v2 Faster R-CNN Inception v2 algoritmalarının tamamının yeniden eğitimi sonrası bağımsız veri seti olan test veri setinde sonuç üretmesi için yine güçlü kabul edilebilecek bir iş istasyonu olan Intel(R) Core (TM) i5-8300H CPU @ 2.30GHz – 16 GB Ram – GTX1050 GPU tercih edilmiştir.

4.5. Araştırmada Sonuçlarında Kullanılan İleri İstatistiksel Analizler

Ortopedi uzmanları ve pratisyen hekimler olmak üzere toplamda iki hekim grubu, YOLO Darknet V4 algoritması, SSD Mobilenet v2 Faster R-CNN Inception v2 algoritmaları olmak üzere üç derin öğrenme algoritması grubunun eğitim de

doğrulama veri setlerinden bağımsız test veri setinde ürettiği sonuç üretme süreleri, kırık kemikleri doğru tespit etme oranı, kırık olmayan kemiklerde doğru tespit etme oranı noktasında istatistiksel analiz yapılarak hem algoritmaların hem de insan çalışma gruplarının ürettiği sonuçlar noktasında anlamlı bir durum olup olmadığı tespit edilme ihtiyacı duyulmuştur.

Tek yönlü varyans analizi olarak geçen one way ANOVA testi, birbirinden bağımsız grupların ürettiği sonuçlarının ortalamalarının arasında istatistiksel anlamda anlamlı bir fark halinin var olup olmadığının kontrol edilmesinde kullanılır. ANOVA aslında "Analysis of variance" manasına gelmektedir. Tek yönlü ANOVA testi iki örneklem t testinden farklı olarak ikili ikili grupları kontrol etmek yerine bir seferde çoklu grupların ilişkisinin kontrol edilmesini sağlamaktadır. Üç veya daha fazla veri gruplarının aritmetik ortalamalarını kümülatif olarak kıyaslanmasını sağlamaktadır. Tek yönlü ANOVA testi parametrik bir testtir ve kullanılabilmesi için kullanılacak verilerin normal dağılması ve örneklem büyüklüğünün yeterli olması varsayımlarını sağlanması gerekmektedir. Tek yönlü varyans analizini gerçekleştirmede hızlı sonuç elde edilmesi noktasında MiniTab ya da SPSS programları kullanılmaktadır. Hangi program kullanılırsa kullanılsın tek yönlü varyans testinde P değeri 0,05'den küçük olduğu için ortalamaların arasındaki fark anlamlıdır. H0: Bütün serilerin ortalamaları eşittir; H1: En az bir serinin ortalaması diğerlerinden farklıdır; kabulü ile P değeri 0,05'den küçük olduğu H0 reddedilir, H1 kabul edilir. Tek yönlü varyans analizlerinde yapılan çalışmalar normalliğe karşı hassas olmayan bir test olduğunu göstermekte örneklem büyüklüğünün yeterli olması halinde analizler gerçekleştirilebilir (Montgomery ve Runger, 2010).

Tek yönlü varyans analizi bize grupların birbirlerinden farklı olup olmadığını göstermesine karşın hangi grupların birbirinden farklı olduğu bilgisini vermemektedir. Bu nedenle çoklu karşılaştırma testlerine devam edilir. Hem varyansların benzerlik gösterdiği hem de örneklemler eşit olduğu için Tukey karşılaştırma testlerine devam edilmektedir.

Tukey testi 1949 yılında J. W. Tukey tarafından birbiri ile ilişkili olmayan çoklu sayıda grup ortalamayı ortak bir hata yaklaşımı kullanımı ile ikişer olarak kıyaslamayı sağlayan çoklu kıyaslama testidir. Tukey testinin tercihinde önemli nedenler şu şekilde sınıflandırılabilir. Kesin Anlamlı Farklar: Tukey testi, çoklu karşılaştırmalar arasındaki gerçek anlamlı farkları belirlemek için kullanılır. Bu, türden bağımsız

olarak gruplar arasındaki anlamlı farkları belirlemek için güçlü bir yöntem olduğu anlamına gelir. Aşırı Tip I Hatasını Kontrol Eder: Çoklu karşılaştırmalar yaparken, birden fazla test gerçekleştirdiğinizde, tek bir hipotez testi yaparken belirlediğiniz anlamlılık düzeyi olan genellikle alfa (örneğin, 0.05) kullanmanız yaygındır. Bu durumda, aşırı tip I hata riski artar. Tukey testi, bu riski kontrol altında tutarak anlamlılık düzeyini ayarlar. Homojenlik Varsayımını Yumuşatır: Tukey testi, ANOVA gibi varyans analizlerinin sonucu olarak grupların homojen olmadığı durumlarda da kullanılabilir. Bu nedenle, varyansların homojen olmadığı durumlarda bile güvenilir sonuçlar elde etmek mümkündür. Grup Sayısını ve Büyüklüğünü İyi İdare Eder: Tukey testi, grup sayısı ve örneklem büyüklüğü konusunda esnek bir yapıya sahiptir. Yani, farklı grup sayıları ve örneklem büyüklükleri ile kullanılabilir. Basit ve Yapılandırılmış Sonuçlar: Tukey testi sonuçları genellikle basit ve anlaşılır bir şekildedir. Hangi grupların birbirinden anlamlı olarak farklı olduğunu net bir şekilde gösterir. Çeşitli Veri Dağılımları için Uygundur: Tukey testi, normal dağılım varsayımına bağımlı değildir ve değişik veri dağılımlarına karşı oldukça toleranslıdır. (Kayri, 2009; Montgomery ve Runger, 2010; Tukey, 1949).

5. TARTIŞMA VE SONUÇ

Bu kısımda aktarımlı öğrenim tekniği ile derin öğrenme ve bilgisayarlı görünün en güçlü sayılan YOLO Darknet V4 algoritması, SSD Mobilenet v2, Faster R-CNN Inception v2 algoritmalarının yeniden eğitilerek femur üst bölge kırıklarını tespit edecek şekilde yeniden eğitilmesi sonucu yeni algoritmaların üretilmesi, bunların çoklu performans kriterlerine göre performansının ortaya konulması aktarılmıştır. Söz konusu algoritmaların performansı daha doğru bir şekilde ortaya konulması için ortopedi uzman doktorları, pratisyen hekimler, özgün veri seti aracılığı ile sonuçların objektif bir şekilde ortaya konulması amaçlanmış, bunun sonucunda travma sonrası kemik kırıklarında derin öğrenme ve bilgisayarlı görü algoritmalarının performansı, ortopedi kırıklarında en fazla ölüm ve sakat kalma riski bulunan femur üst bölge yani kalça kırıklarında ortaya konulmuştur. Bunun yanında ileri istatistiksel metotlar ile de sonuçlar tekrar değerlendirilmiştir.

5.1. Araştırmada Yer Alan Pratisyen Hekimlerin ve Ortopedi Uzmanlarının Ürettiği Sonuçlar

Araştırmanın önemli aşamalarından biri pratisyen hekimler ile ortopedi uzmanların belirlenmesi ve gruplandırılmasıdır. Ortopedi uzmanı hekimler, Türkiye'nin en kalabalık nüfusuna sahip İstanbul ilinin en önemli eğitim ve araştırma hastanesi ortopedi ve travmatoloji servisinden biri seçilmiştir. Toplamda 12 uzman hekim ve ortopedi ve travmatoloji servisi kıdemli asistanları odağa alınmıştır. Kendilerine tamamen bağımsız bir veri seti olarak oluşturulan 55 X-ray görüntüsünden oluşan test veri setinden 50 adet femur üst bölge kırığı içeren 50 adet femur üst bölge kırığı içermeyen toplamda 100 femur üst bölge görüntüsü karışık olarak gösterilmiştir. X-ray görüntüleri kendilerine her gösterildiğinde sağ femur üst bölge kırığı, sol femur üst bölge kırığı, hem sağ hem sol femur üst bölge kırığı yahut kırık yok demeleri beklenmiştir. Her resim gösterilmesi sonrasında hekimin resmi gördüğü ilk andan yanıtlandırarak görüntü üzerinde kırığı göstermesine kadar geçen süreler sonuç üretme süresi olarak kronometraj yöntemi ile ölçümler gerçekleştirilmiştir. Ortaya çıkan

sonuçlar bu şekilde kayıt altına alınarak değerlendirme kriterleri tablosuna göre hesaplamalar ve ölçümlemeler gerçekleştirilmiştir.

Pratisyen hekimler belirlenirken yine Türkiye'nin en kalabalık nüfusuna sahip İstanbul ilinin en önemli eğitim ve araştırma hastanesine farklı branşlardan dahil olan uzmanlığında 1 aydan fazla geçirmemiş hekimler, acil hekimleri ve bu eğitim ve araştırma hastanesinden farklı lokasyonlardan aile hekimlerinden destek alınmıştır. Toplamda 10 pratisyen hekim ile çalışma gerçekleştirilmiştir. Kendilerine tamamen bağımsız bir veri seti olarak oluşturulan 55 X-ray görüntüsünden oluşan test veri setinden 50 adet femur üst bölge kırığı içeren 50 adet femur üst bölge kırığı içermeyen toplamda 100 femur üst bölge görüntüsü karışık olarak gösterilmiştir. Uzman hekimlerde olduğu şekilde X-ray görüntüleri kendilerine her gösterildiğinde sağ femur üst bölge kırığı, sol femur üst bölge kırığı, hem sağ hem sol femur üst bölge kırığı yahut kırık yok demeleri beklenmiştir. Yine kronometraj yöntemi ile resim gösterilmesi sonrasında hekimin resmi gördüğü ilk andan yanıtlandırarak görüntü üzerinde kırığı göstermesine kadar geçen süreler, sonuç üretme süresi olarak kayıt altına alınmıştır.

Bütün bu çalışmaların ardından aşağıda çoklu değerlendirme kriterleri tablosu pratisyen ve uzman hekim sonuçları tablosunda görüleceği üzere uzman hekimler doğruluk oranında %91,42 seviyesine başarı gösterirken pratisyen hekimler %81,30 seviyesinde başarı elde etmiştir. Duyarlılık oranında ise uzman hekimler %91,67 oranında başarı gösterirken, pratisyen hekimlerin oranı %74,80 seviyesinde kalmıştır. Özüllük oranı ölçüldüğünde uzman hekimlerin %91,17 olarak ölçülürken, pratisyen hekimlerin özüllük oranı %87,80 seviyesindedir.

Tablo 5.1. Çoklu değerlendirme kriterleri tablosu pratisyen ve uzman hekim sonuçları.

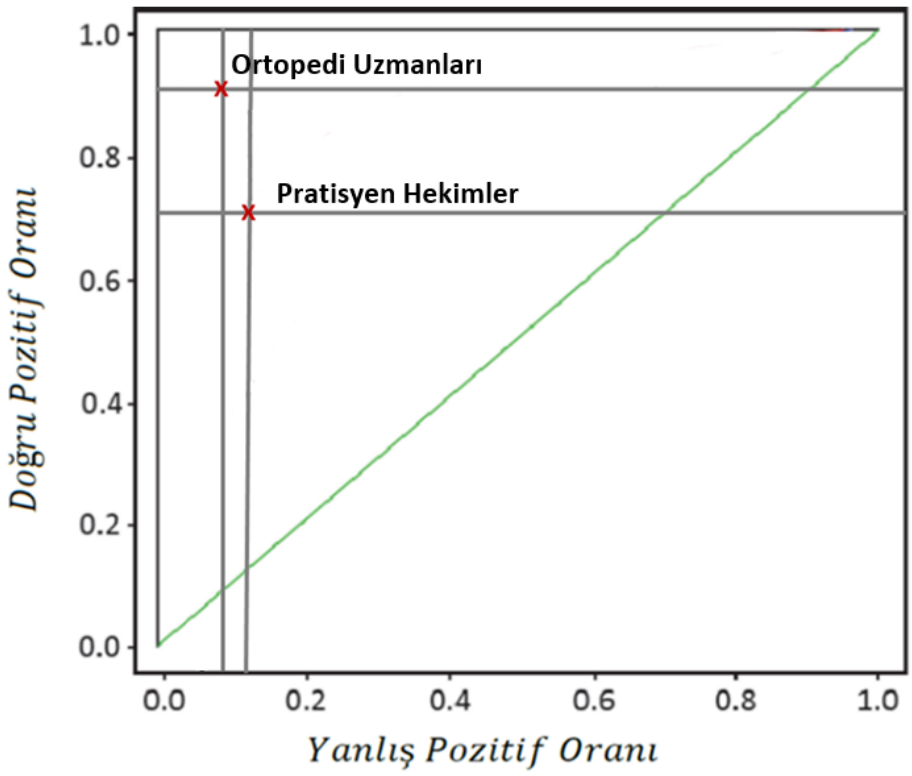
Kriterler	Hesaplama	Test veri setine göre Pratisyen Hekimlerin Sonuçları	Test veri setine göre Uzman Doktorların Sonuçları
Doğruluk Oranı	*(TP+TN)/(N+P)*	*81,30%*	*91,42%*
Duyarlılık	*TP/(TP+FN)*	*74,80%*	*91,67%*
Özüllük	*TN/(TN+FP))*	*87,80%*	*91,17%*
Yanlış Pozitif Oranı	*FP/(FP+TN)*	*12,20%*	*8,83%*
Yanlış Negatif Oranı	*FN/(FN+TP)*	*25,20%*	*8,33%*
Doğru Tahmin Oranı	*TP/(TP+FP)*	*85,98%*	*91,21%*
F1 Puanı	*2*((PPV*TPR)/ (PPV+TPR))*	*80,00%*	*91,44%*
Hata Oranı	*(FP+FN)/(N+P)*	*18,70%*	*8,58%*

Uzman hekimler ve pratisyen hekimlerin test veri setinde bulunan X-ray görüntüleri üzerindeki tespit süreleri ölçüldüğünde sonuç üretme süreleri tablosu ortaya çıkmaktadır buna göre ortalama tespit süreleri pratisyen hekim grubunun, uzman hekim grubuna göre 0,5 sn yavaş kaldığı görülmektedir. Buna karşın pratisyen hekimlerin tespit süreleri standart sapması uzman doktorlara göre 1,16 sn daha düşük olduğu görülmektedir. Uzman doktorların minimum tespit süresi pratisyen hekimlerin 0,51 sn altındayken, maksimum tespit süresi 3,54 sn daha uzun sürdüğü yapılan çalışma ile görülmüştür.

Tablo 5.2. Pratisyen ve uzman hekimlerin sonuç üretme süreleri (saniye).

Test Veri Seti Süreleri	Pratisyen Hekim Grubu Sonuçları	Uzman Doktor Grubu Sonuçları
Ortalama Tespit Süresi (sn)	*5,10*	*4,73*
Tespit Süresinin Standart Sapması (sn)	*1,89*	*3,05*
Maksimum Tespit Süresi (sn)	*10,18*	*13,72*
Minimum Tespit Süresi (sn)	*2,21*	*1,70*

Değerlendirme kriterleri içerisinde yer alan ROC grafiğinde hem ortopedi uzman hekimleri hem de pratisyen hekimler gösterildiğinde sol üst köşeye en yakın değer olağan şekilde ortopedi ve travmatoloji doktorlarının olmaktadır. Buna göre ortopedi ve travmatoloji uzman doktorlarının ürettiği sonuçlar ROC grafiğinin sol üst köşesine uzaklığı 0,121 birim olmaktadır. Pratisyen hekimler için ise bu uzaklık 0,279 birim olmaktadır.



Şekil 5.1. ROC eğrisi analizinde ortopedi uzman ve pratisyen hekimlerin sonuçları.

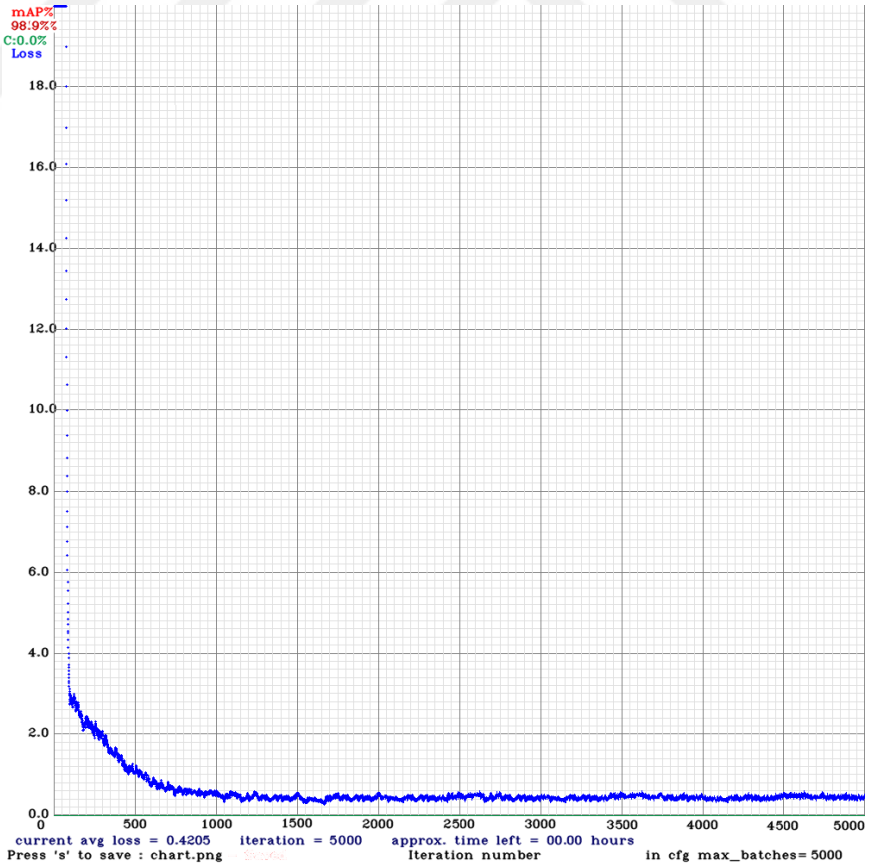
5.2. YOLO Algoritmasının Ürettiği Sonuçlar

Araştırmada kullanılan en güçlü derin öğrenme algoritmalarından biri YOLO Darknet V4 algoritmasıdır. MS COCO veri setinde daha öncesinde eğitilmiş YOLO Darknet v4 algoritması aktarımlı öğrenme kullanılmak sureti ile python yazılım dilinde yazılan kod blokları ile yeniden eğitilmiştir. YOLO algoritması 656 X-ray görüntüsü eğitim ve 164 X-ray görüntüsü doğrulama olmak üzere toplamda 820 X-ray görüntü ile toplamda 5.000 iterasyon boyunca eğitilmiştir. Bu eğitim öncesinde söz konusu veri setleri arayüz kullanılarak etiketlenmiştir. Söz konusu eğitim gerçekleştirilirken YOLO algoritması için dünyanın en güçlüsü olarak kabul edilmiş GPUlarından biri olan Tesla K80 24 GB GDDR5 GPUsu ile Google Colab üzerinden eğitim

gerçekleştirilmiştir. Bu 5.000 iterasyonluk eğitim toplamda 13 saat 6 dakikada tamamlanmıştır. Eğitimin çıktısı olan Weight dosyası iş istasyonu üzerine kullanılmak için kaydedilmiştir.

Eğitim sonuçlarının modelin ezberlemesinin önüne geçmek ve değerlendirme sonuçlarının tarafsız olması amacıyla eğitim ve doğrulama veri setinden tamamen tarafsız olarak oluşturulan test veri seti aracılığı ile performans değerlendirilmiştir. 50 femur üst bölge kırığı içermeyen, 50 femur üst bölge kırığı içeren toplamda 55 X-ray görüntüsünde sonuçlar algoritma aracılığı ile ortaya konulmuştur. Test veri setinde çıktılar üretilirken Intel Core i5-8300H CPU ve NVIDIA GeForce GTX 1050 GPU ya sahip iş istasyonu kullanılmıştır.

YOLO algoritması yeniden eğitilerek femur üst bölge kırıklarını tespit edebilecek yeni özgün modelin eğitimi sırasında elde edilen Loss (kayıp) grafiği aşağıdaki gibidir. Toplam kayıp (Total Loss) grafiği incelendiğinde 5000 iterasyonda YOLO algoritmasının 18 seviyesinin üzerinden başlayarak 1000 iterasyona vardığında 1'in altına indiği gözükmektedir. Hızlı bir öğrenme süreci geçirdiği gözükmektedir.



Şekil 5.2. YOLO algoritması loss grafiği.

YOLO algoritması çoklu değerlendirme kriterleri tablosunda performansı hesaplanırsa doğruluk oranı %90,33 seviyesinde duyarlılık oranı %87,67 seviyesinde, özüllük oranı ise %92,98 seviyesinde, doğru tahmin oranı ise %92,58 seviyesinde olduğu görülmüştür.

Tablo 5.3. Çoklu değerlendirme kriterleri tablosu YOLO Algoritması sonuçları.

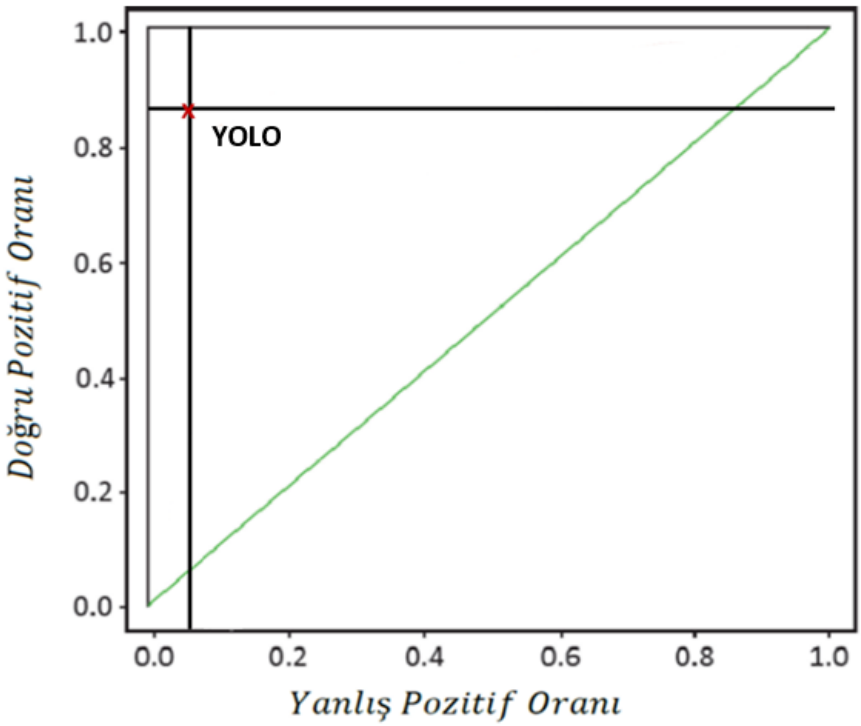
Kriterler	**Hesaplama**	**Test veri setine göre YOLO Algoritması Sonuçları**
Doğruluk Oranı	*(TP+TN)/(N+P)*	*90,33%*
Duyarlılık	*TP/(TP+FN)*	*87,67%*
Özüllük	*TN/(TN+FP))*	*92,98%*
Yanlış Pozitif Oranı	*FP/(FP+TN)*	*7,02%*
Yanlış Negatif Oranı	*FN/(FN+TP)*	*12,33%*
Doğru Tahmin Oranı	*TP/(TP+FP)*	*92,58%*
F1 Puanı	*2*((PPV*TPR)/(PPV+TPR))*	*90,06%*
Hata Oranı	*(FP+FN)/(N+P)*	*9,67%*

YOLO algoritmasının test veri seti üzerinde Intel Core i5-8300H CPU ve NVIDIA GeForce GTX 1050 GPU ya sahip iş istasyonu ile sonuç ürettirken sonuç üretme süreleri aşağıdaki şekilde kayıt alttına alınmıştır. Buna göre YOLO algoritmasının ortalama tespit süresi 1,40 sn olmuştur. Tespit sürelerinin standart sapması 0,25 sn maksimuum tespit süresi 2,27 sn sürerken minimum tespit süresi 1,17 sn olarak gerçekleşmiştir.

Tablo 5.4. YOLO Algoritması sonuç üretme süreleri (saniye).

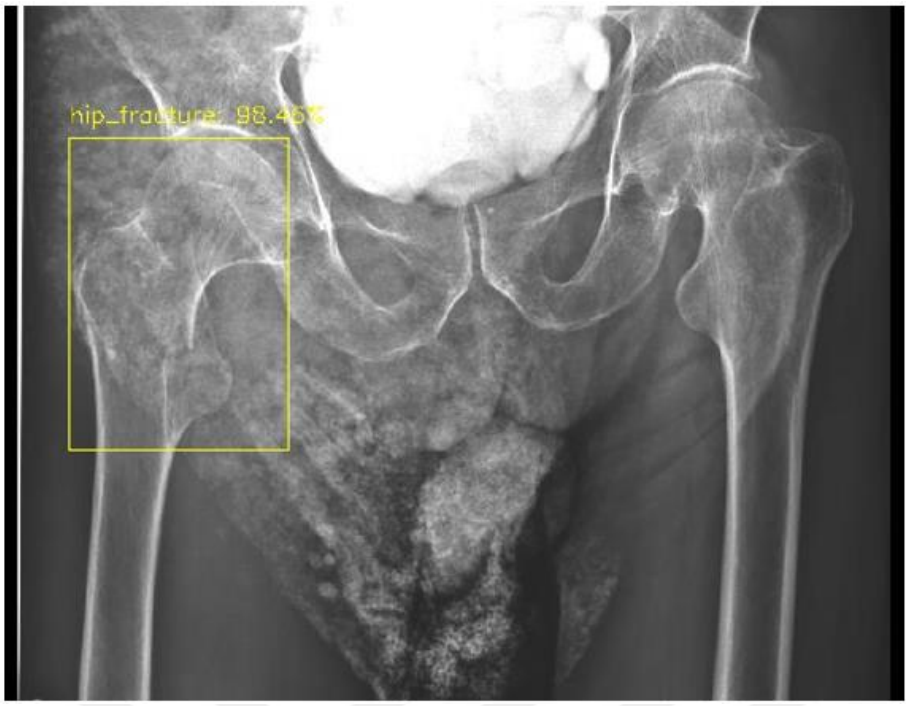
Test Veri Seti Süreleri	**YOLO Algoritması Sonuçları**
Ortalama Tespit Süresi (sn)	*1,40*
Tespit Süresinin Standart Sapması (sn)	*0,25*
Maksimum Tespit Süresi (sn)	*2,27*
Minimum Tespit Süresi (sn)	*1,17*

Değerlendirme kriterleri içerisinde yer alan ROC analizi için YOLO algoritmasının doğru pozitif oranı y eksenine, yanlış pozitif oranı ise x eksenine iliştirildiğinde YOLO algoritmasının sol üst köşeye olan uzaklığı 0,141 birim çıkmaktadır.

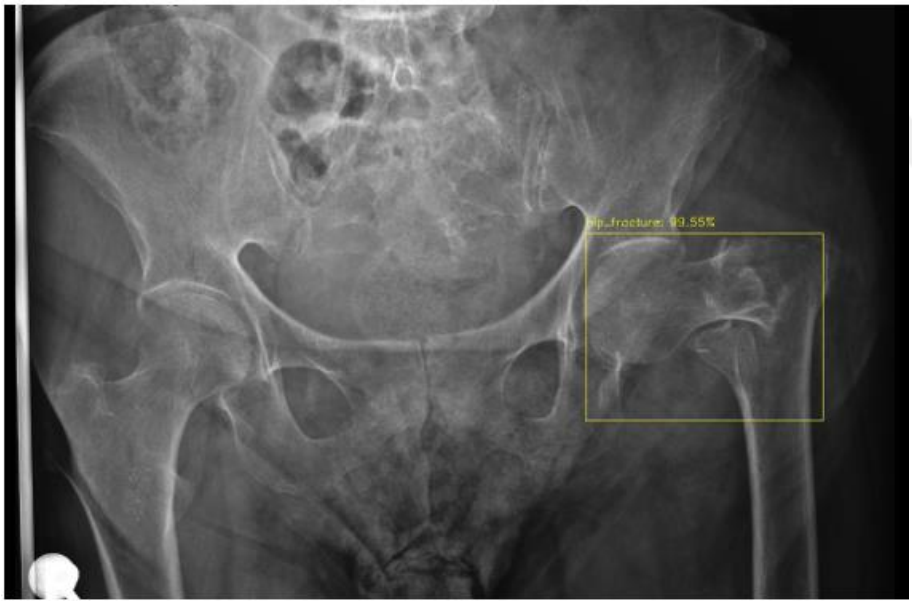


Şekil 5.3. ROC eğrisi analizinde YOLO algoritması sonucu.

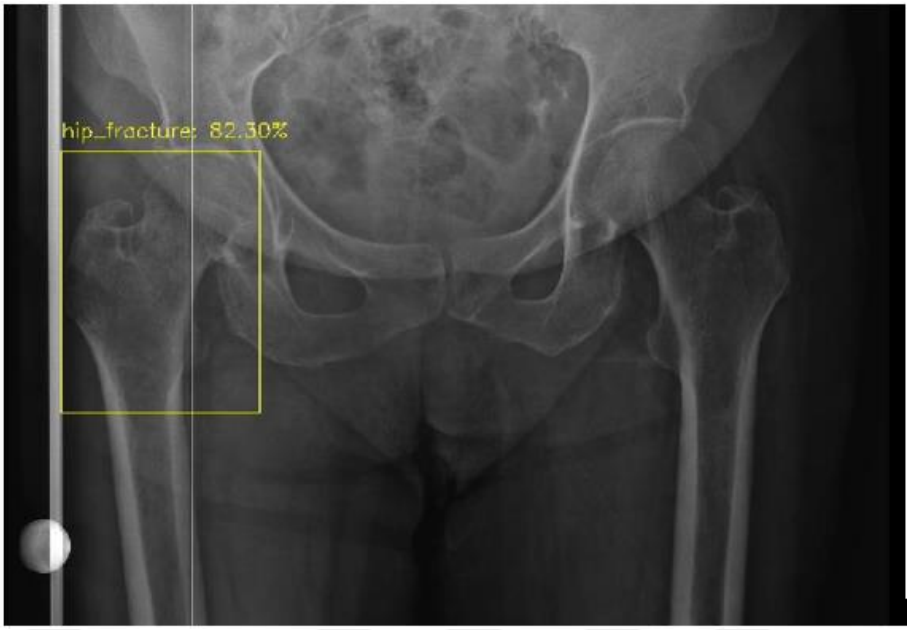
Araştırmada sonucunda yapılan çalışma ile femur üst bölge kırıklarını tespit etmek üzere geliştirilen YOLO algoritmasın üretmiş olduğu sonuçlardan örneklere aşağıda yer verilmiştir.



Şekil 5.4. YOLO algoritması ile femur üst bölge kırığı tepiti örneği-1.



Şekil 5.5. YOLO algoritması ile femur üst bölge kırığı tepiti örneği-2.



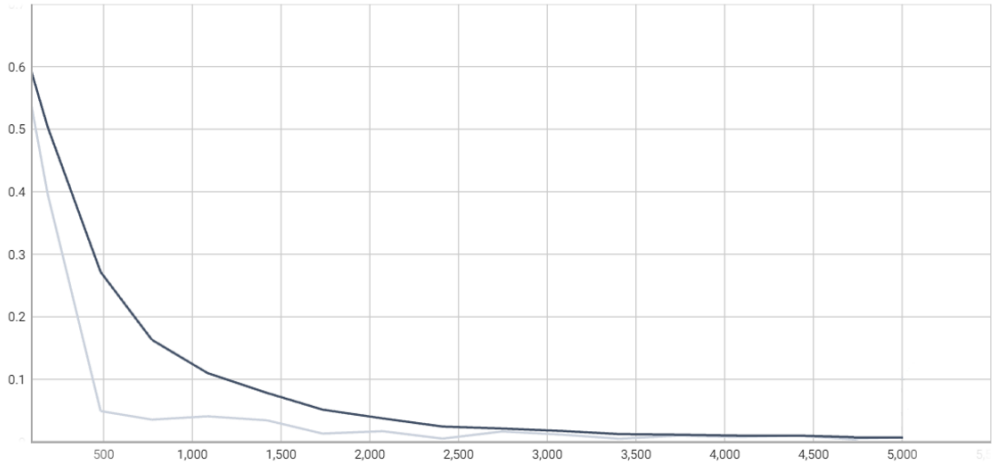
Şekil 5.6. YOLO algoritması ile femur üst bölge kırığı tespiti örneği-3.

5.3. Faster R-CNN Algoritmasının Ürettiği Sonuçlar

Araştırmada yine en güçlü derin öğrenme algoritmalarından biri olarak kabul edilebilecek Faster R-CNN Inception v2 algoritması kullanılmıştır. Tıpkı YOLO algoritması gibi Faster R-CNN algoritması da MS COCO veri setinde daha öncesinde eğitilmiştir. Faster R-CNN Inception v2 algoritması aktarımlı öğrenme kullanılmak sureti ile python yazılım dilinde yazılan kod blokları ile yeniden eğitimi gerçekleştirilmiştir. Faster R-CNN algoritması toplamda 820 X-ray görüntü ile toplamda YOLO ve SSD algoritmaları gibi 5.000 iterasyon boyunca eğitilmiştir. Bu eğitim öncesinde söz konusu veri setleri yine arayüz kullanılarak koordinatları etiketlenmiştir. Faster R-CNN Algoritmasının eğitimi gerçekleştirilirken Intel Core i5-8300H CPU ve NVIDIA GeForce GTX 1050 GPU ya sahip iş istasyonu kullanılmıştır. Bu 5.000 iterasyonluk eğitim toplamda 31,9 dakikada tamamlanmıştır. Eğitimin çıktısı olan Weight dosyası yine aynı iş istasyonu üzerine kullanılmak için kaydedilmiştir.

Eğitim sonuçlarının modelin ezberlemesinin önüne geçmek ve değerlendirme sonuçlarının tarafsız olması amacıyla, eğitim ve doğrulama veri setinden tamamen tarafsız olarak oluşturulan yine aynı test veri seti aracılığı ile performans değerlendirilmesi gerçekleştirilmiştir. 50 femur üst bölge kırığı içermeyen, 50 femur üst bölge kırığı içeren toplamda 55 X-ray görüntüsünde sonuçlar algoritma aracılığı ile ortaya konulmuştur. Test veri setinde çıktılar üretilirken yine Intel Core i5-8300H CPU ve NVIDIA GeForce GTX 1050 GPU ya sahip iş istasyonu kullanılmıştır.

Faster R-CNN algoritması yeniden eğitilerek femur üst bölge kırıklarını tespit edebilecek yeni özgün modelin eğitimi sırasında elde edilen Loss (kayıp) grafiği aşağıdaki gibidir. Toplam kayıp (Total Loss) grafiği Faster R-CNN algoritması için incelendiğinde 5000 iterasyonda 0,6 seviyesinden başlayarak 1000 iterasyona vardığında 0,2'in altına indiği görülmüştür. Faster R-CNN algoritmasının da YOLO Algoritması gibi hızlı bir öğrenme süreci geçirdiği görülmüştür.



Şekil 5.7. Faster R-CNN algoritması loss grafiği.

Tablo 5.5. Çoklu değerlendirme kriterleri tablosu Faster R-CNN Algoritması sonuçları.

Kriterler	**Hesaplama**	**Test veri setine göre Faster R-CNN Algoritması Sonuçları**
Doğruluk Oranı	*(TP+TN)/(N+P)*	*84,29%*
Duyarlılık	*TP/(TP+FN)*	*90,64%*
Özüllük	*TN/(TN+FP))*	*77,94%*
Yanlış Pozitif Oranı	*FP/(FP+TN)*	*22,06%*
Yanlış Negatif Oranı	*FN/(FN+TP)*	*9,36%*
Doğru Tahmin Oranı	*TP/(TP+FP)*	*80,43%*
F1 Puanı	*2*((PPV*TPR)/(PPV+TPR))*	*85,23%*
Hata Oranı	*(FP+FN)/(N+P)*	*15,71%*

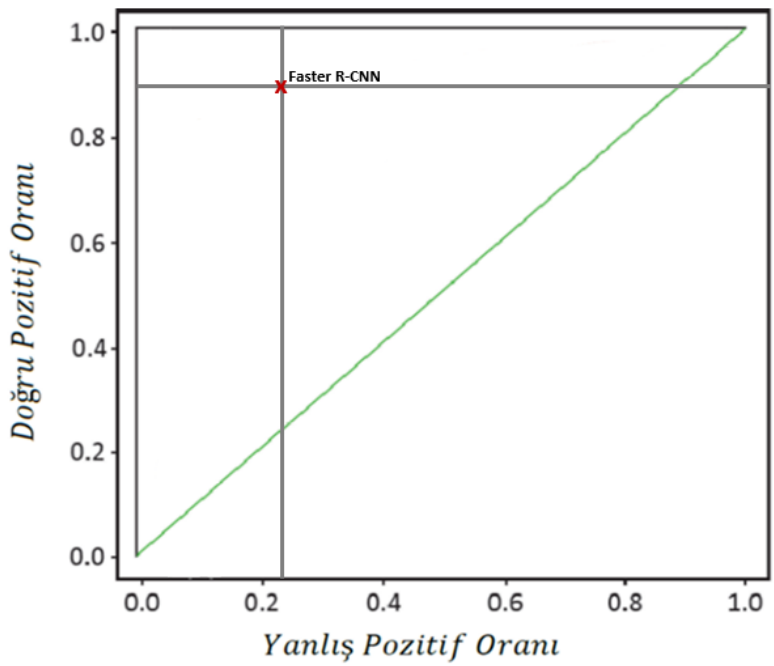
Faster R-CNN algoritması çoklu değerlendirme kriterleri tablosunda performansı hesaplanırsa doğruluk oranı %84,29 seviyesinde duyarlılık oranı %90,64 seviyesinde, özüllük oranı ise %77,94 seviyesinde, doğru tahmin oranı ise %80,43 seviyesinde olduğu görülmüştür.

Faster R-CNN algoritmasının test veri seti üzerinde Intel Core i5-8300H CPU ve NVIDIA GeForce GTX 1050 GPU ya sahip iş istasyonu ile sonuç ürettirken sonuç üretme süreleri aşağıdaki şekilde kayıt alttına alınmıştır. Buna göre Faster R-CNN algoritmasının ortalama tespit süresi 10,49 sn olmuştur. Tespit sürelerinin standart sapması 0,22 sn maksimuum tespit süresi 10,99 sn sürerken minimum tespit süresi 10,06 sn olarak gerçekleşmiştir.

Tablo 5.6. Faster R-CNN Algoritması sonuç üretme süreleri (saniye).

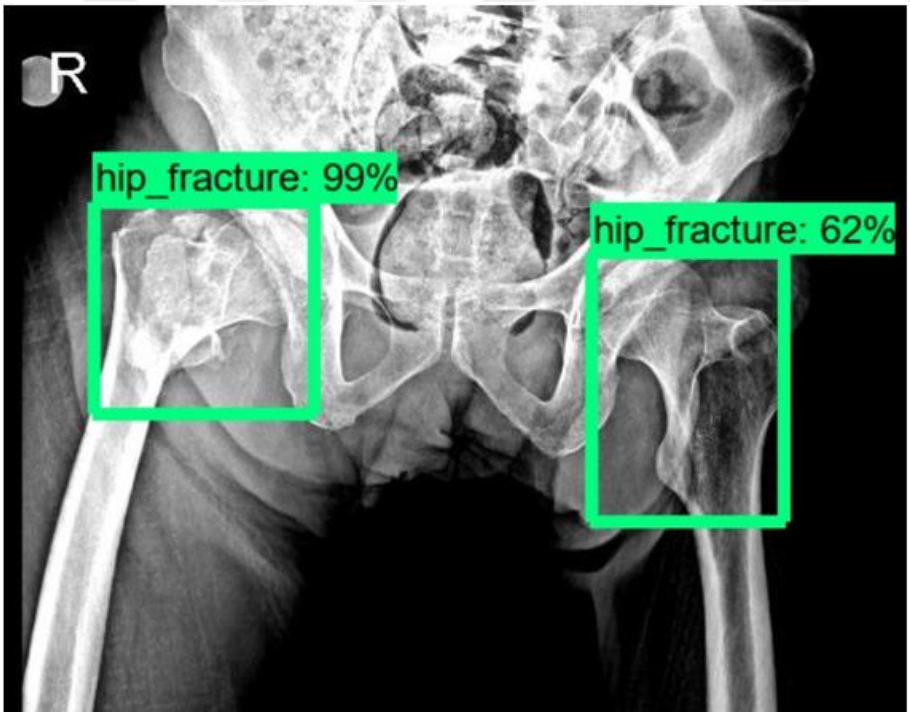
Test Veri Seti Süreleri	**Faster R-CNN Algoritması Sonuçları**
Ortalama Tespit Süresi (sn)	*10,49*
Tespit Süresinin Standart Sapması (sn)	*0,22*
Maksimum Tespit Süresi (sn)	*10,99*
Minimum Tespit Süresi (sn)	*10,06*

Değerlendirme kriterleri içerisinde yer alan ROC analizi için Faster R-CNN algoritmasının doğru pozitif oranı y eksenine, yanlış pozitif oranı ise x eksenine iliştirildiğinde Faster R-CNN algoritmasının sol üst köşeye olan uzaklığı 0,239 birim çıkmaktadır.

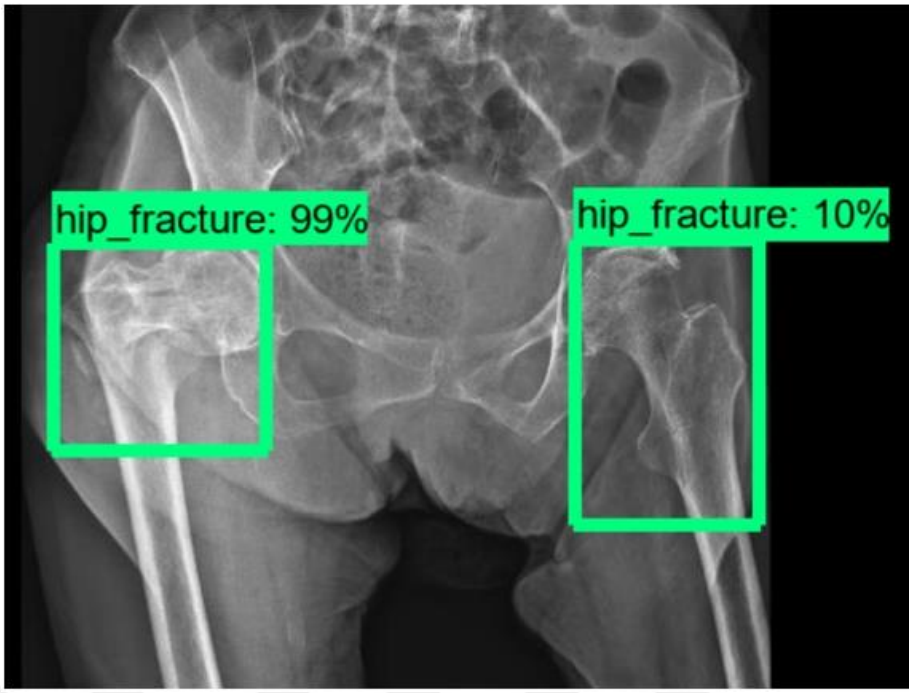


Şekil 5.8. ROC eğrisi analizinde Faster R-CNN algoritması sonucu.

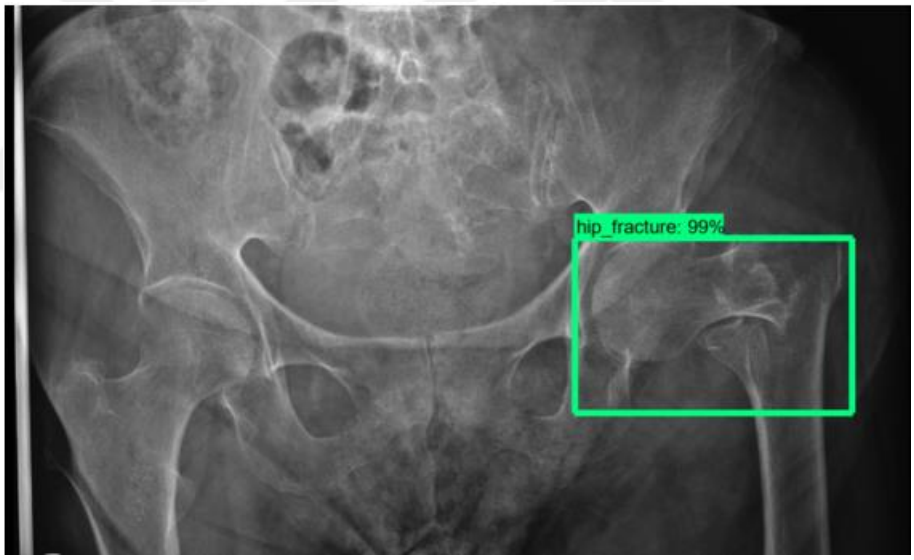
Araştırmada sonucunda yapılan çalışma ile femur üst bölge kırıklarını tespit etmek üzere geliştirilen Faster R-CNN algoritmasın üretmiş olduğu sonuçlardan örneklere aşağıda yer verilmiştir.



Şekil 5.9. Faster R-CNN algoritması ile femur üst bölge kırığı tepiti örneği-1.



Şekil 5.10. Faster R-CNN algoritması ile femur üst bölge kırığı tepiti örneği-2.



Şekil 5.11. Faster R-CNN algoritması ile femur üst bölge kırığı tepiti örneği-3.

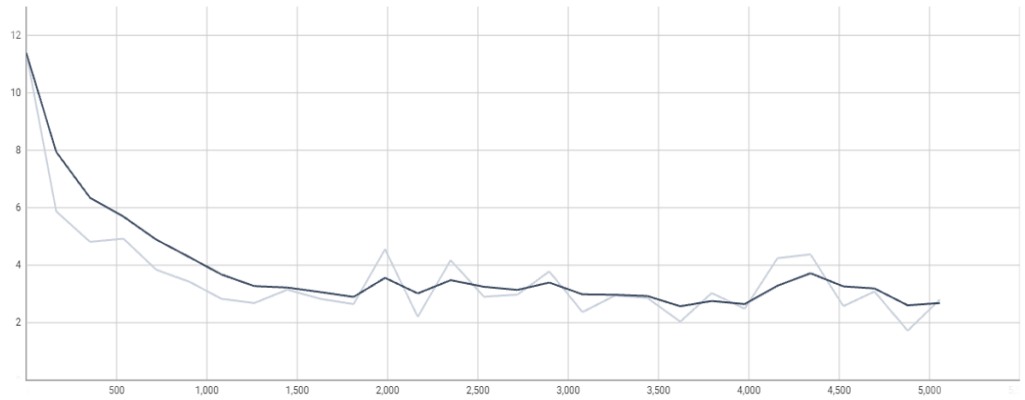
5.4. SSD Algoritmasının Ürettiği Sonuçlar

Araştırmada kullanılan yine bir diğer derin öğrenme algoritması SSD Mobilenet v2 algoritmasıdır. Bu çalışmada kullanılan YOLO algoritması ve Faster R-CNN algoritması gibi SSD algoritması da MS COCO veri setinde daha öncesinde eğitilmiştir. SSD Mobilenet v2 algoritması aktarımlı öğrenme kullanılmak sureti ile python yazılım dilinde yazılan kod blokları ile yeniden eğitimi gerçekleştirilmiştir. SSD Mobilenet v2 algoritması tıpkı diğer algoritmalarda birebir aynısı kullanılan

toplamda 820 X-ray görüntü ile eğitilmiştir. YOLO ve Faster R-CNN algoritmaları gibi SSD algoritması 5.000 iterasyon boyunca eğitilmiştir. Bu eğitim öncesinde söz konusu veri setleri yine farklı bir arayüz kullanılarak koordinatları etiketlenmiştir. SSD algoritmasının eğitimi gerçekleştirilirken diğer iki algoritmada olduğu gibi Intel Core i5-8300H CPU ve NVIDIA GeForce GTX 1050 GPU ya sahip iş istasyonu kullanılmıştır. Bu 5.000 iterasyonluk eğitim toplamda 55,92 dakikada tamamlanmıştır. Eğitimin çıktısı olan Weight dosyası yine aynı iş istasyonu üzerine kullanılmak üzere kaydedilmiştir.

Eğitim sonuçlarının değerlendirmek amacıyla bu çalışmada kullanılan diğer iki algoritma ve uzman hekimler ile pratisyen hekimlerde kullanılan yine aynı test veri seti aracılığı ile performans değerlendirilmesi gerçekleştirilmiştir. 50 femur üst bölge kırığı içermeyen, 50 femur üst bölge kırığı içeren toplamda 55 X-ray görüntüsünde sonuçlar algoritma aracılığı ile ortaya konulmuştur. Test veri setinde çıktılar üretilirken yine Intel Core i5-8300H CPU ve NVIDIA GeForce GTX 1050 GPU ya sahip iş istasyonu kullanılmıştır.

SSD algoritması yeniden eğitilerek femur üst bölge kırıklarını tespit edebilecek yeni özgün modelin eğitimi sırasında elde edilen Loss (kayıp) grafiği aşağıdaki gibidir. Toplam kayıp (Total Loss) grafiği SSD algoritması için incelendiğinde 5000 iterasyonda 11 seviyesinden başlayarak 1000 iterasyona vardığında 4'in altına indiği gözükmektedir. SSD Algoritmasının YOLO ve Faster R-CNN algoritmasından farklı olarak daha yavaş bir öğrenme süreci olduğu loss grafiğinden anlaşılmaktadır.



Şekil 5.12. SSD algoritması loss grafiği

SSD algoritması çoklu değerlendirme kriterleri tablosunda performansı hesaplanırsa doğruluk oranı %72,74 seviyesinde duyarlılık oranı %78,32 seviyesinde, özüllük oranı

ise %67,16 seviyesinde, doğru tahmin oranı ise %70,46 seviyesinde olduğu görülmüştür.

Tablo 5.7. Çoklu değerlendirme kriterleri tablosu SSD algoritması sonuçları.

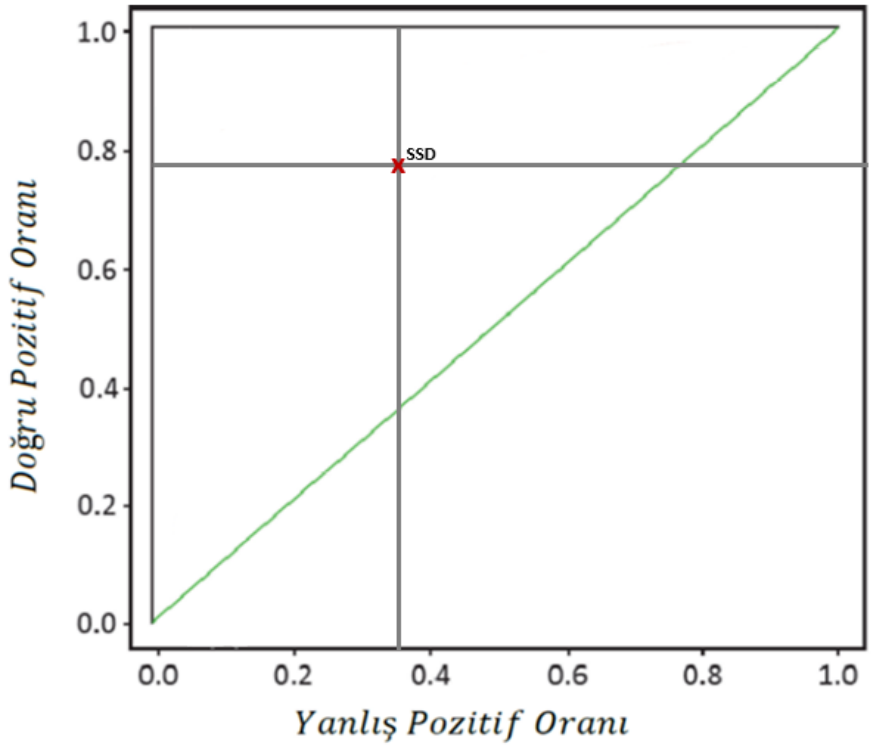
Kriterler	**Hesaplama**	**Test veri setine göre SSD Algoritması Sonuçları**
Doğruluk Oranı	*(TP+TN)/(N+P)*	*72,74%*
Duyarlılık	*TP/(TP+FN)*	*78,32%*
Özüllük	*TN/(TN+FP))*	*67,16%*
Yanlış Pozitif Oranı	*FP/(FP+TN)*	*32,84%*
Yanlış Negatif Oranı	*FN/(FN+TP)*	*21,68%*
Doğru Tahmin Oranı	*TP/(TP+FP)*	*70,46%*
F1 Puanı	*2*((PPV*TPR)/(PPV+TPR))*	*74,18%*
Hata Oranı	*(FP+FN)/(N+P)*	*27,26%*

SSD algoritmasının test veri seti üzerinde Intel Core i5-8300H CPU ve NVIDIA GeForce GTX 1050 GPU ya sahip iş istasyonu ile sonuç ürettirken sonuç üretme süreleri aşağıdaki şekilde kayıt alttına alınmıştır. Buna göre SSD algoritmasının ortalama tespit süresi 6,38 sn olmuştur. Tespit sürelerinin standart sapması 0,23 sn maksimuum tespit süresi 6,89 sn sürerken minimum tespit süresi 5,53 sn olarak gerçekleşmiştir.

Tablo 5.8. SSD Algoritması sonuç üretme süreleri (saniye).

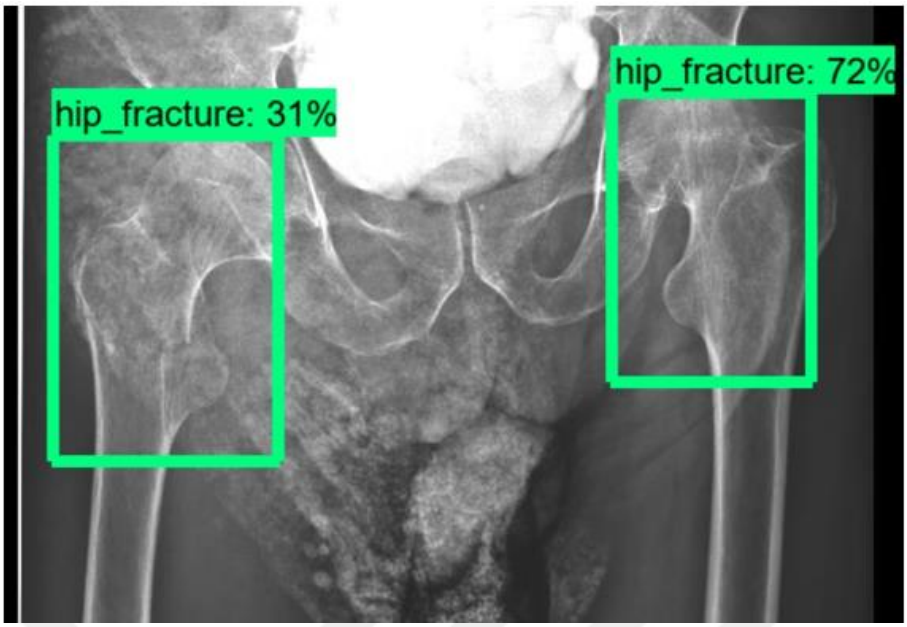
Test Veri Seti Süreleri	**SSD Algoritması Sonuçları**
Ortalama Tespit Süresi (sn)	*6,38*
Tespit Süresinin Standart Sapması (sn)	*0,23*
Maksimum Tespit Süresi (sn)	*6,89*
Minimum Tespit Süresi (sn)	*5,53*

Değerlendirme kriterleri içerisinde yer alan ROC analizi için Faster SSD algoritmasının doğru pozitif oranı y eksenine, yanlış pozitif oranı ise x eksenine iliştirildiğinde SSD algoritmasının sol üst köşeye olan uzaklığı 0,393 birim çıkmaktadır.

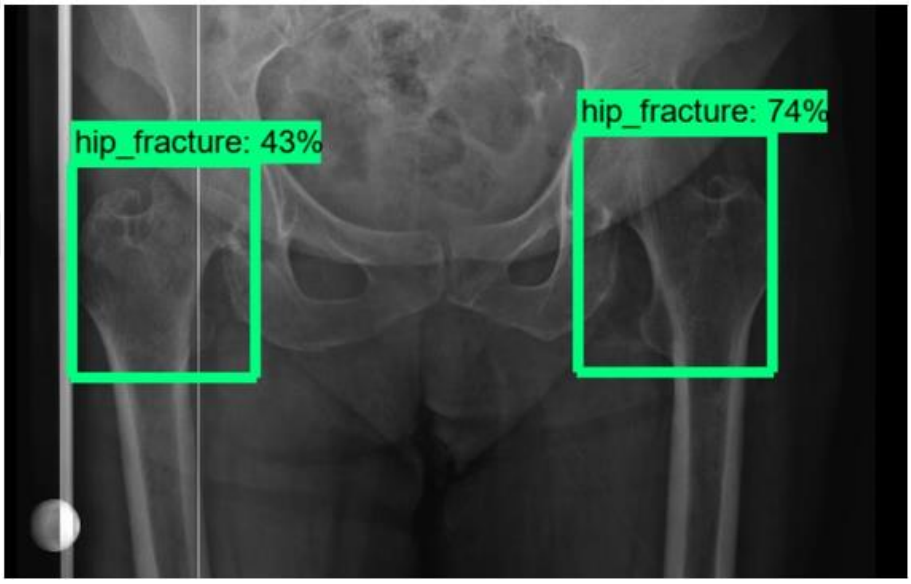


Şekil 5.13. ROC eğrisi analizinde SSD algoritması sonucu.

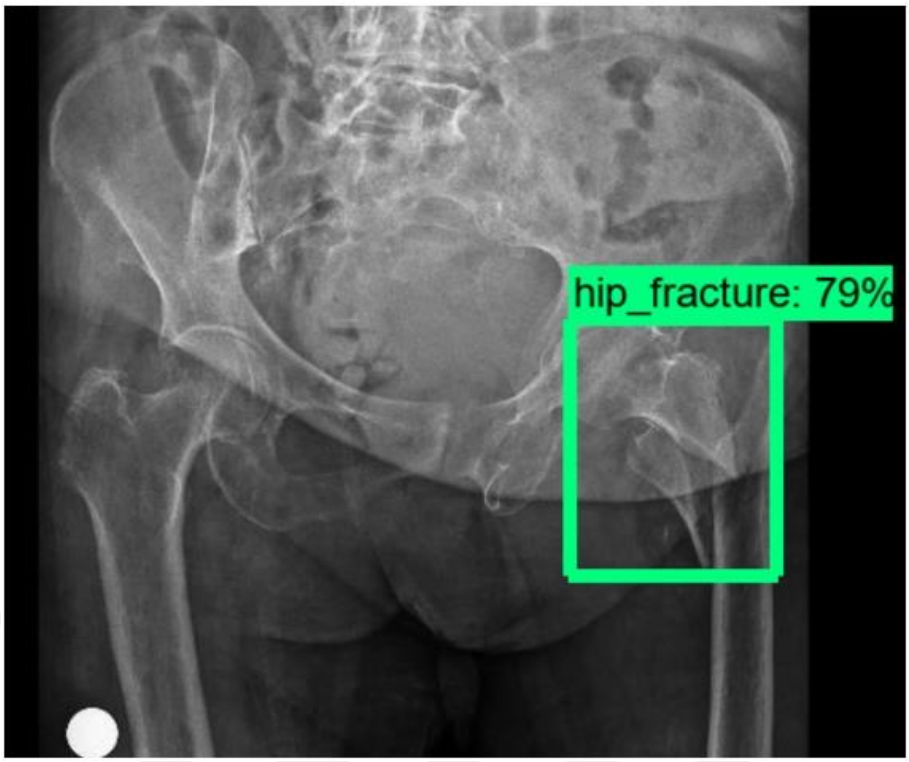
Araştırmada sonucunda yapılan çalışma ile femur üst bölge kırıklarını tespit etmek üzere geliştirilen SSD algoritmasının üretmiş olduğu sonuçlardan örneklere aşağıda yer verilmiştir.



Şekil 5.14. SSD algoritması ile femur üst bölge kırığı tepiti örneği-1.



Şekil 5.15. SSD algoritması ile femur üst bölge kırığı tepiti örneği-2.



Şekil 5.16. SSD algoritması ile femur üst bölge kırığı tepiti örneği-2.

5.5. Bütün Derin Öğrenme Algoritmaları ile Hekim Gruplarının Ürettiği Sonuçların Kıyaslanması

Alanında uzman ortopedi hekimleri ve toplamda altı yıl tıp eğitimi almış pratisyen hekimler olmak üzere iki medikal çalışma grubu, derin öğrenme ve bilgisayarlı görünün en güçlü algoritmalarından olan YOLO Darknet V4, SSD Mobilenet v2, Faster R-CNN Inception v2 algoritmaları olmak üzere üç algoritma kıyaslamanın gerçekleştirilmesi ve doğru performansın ortaya konulması adına bir araya getirilmiştir. Daha öncesinde COCO veri seti ile eğitilmiş üç algoritma; dünyada ortopedik kırıklar değerlendirildiğinde kalıcı sakatlık ve ölüm oranının en yüksek olduğu ve 2050 yılına kadar vaka sayının üç kattan fazla artması beklenen femur üst bölge kırıklarını tespit etmek üzere toplamda aynı 820 eğitim X-ray görüntüsü veri seti ile ile 5.000 iterasyon eğitilmiştir. Toplamda içi insan grubu ve üç yapay zekâ grubunun performansını doğru ortaya koymak adına eğitim ve doğrulama veri setinden bağımsız 50 kırık femur üst bölge içeren ve 50 sağlam femur üst bölge içeren olmak üzere test veri seti oluşturulmuş ve çıktı üretilmesi adına, gruplara ayrı ayrı verilmiştir.

Her grubun çıktılarının kıyaslanması için derlenmiş çoklu performans değerlendirme tablosunda yer verilmiştir. Buna göre doğruluk oranında (ACC) YOLO algoritması %90,33 değeri ile ortopedi uzmanlarının ortaya çıkardığı %91,42 seviyesi ile gerisinde

kalmaktayken, %81,30 doğruluk oranına sahip pratisyen hekimlerden daha iyi bir sonuç ortaya koymuştur. Faster R-CNN algoritması doğruluk oranında %84,29 seviyesinde bir sonuç üreterek yine pratisyen hekimlerden daha iyi bir sonuç ortaya koymuş lakin ortopedi uzman hekimlerinin gerisinde kalmıştır. SSD algoritması %72,74 ile hem uzman doktorların hem de pratisyen hekimlerin gerisinde kalmıştır. Duyarlık oranında (TPR) başarı sıralaması yapacak olursak %91,67 ile uzman hekimler en başarılı iken ardından %90,64 ile Faster R-CNN Algoritması gelmekte, sonrasında %87,67 ile YOLO algoritması, akabinde %78,32 ile SSD algoritması, sonuncu sırada ise pratisyen hekimler yer almaktadır. Özlülük oranında (TNR) başarı sıralamasında yapay zekâ uzman doktorları da geçmekte ve YOLO algoritması % 92,98 ile en yüksek performansı ortaya koymuştur. %91,17 ile ortopedi uzman doktorları takip etmiş, sonrasında %87,50 ile pratisyen hekimler gelmiştir. Faster R-CNN algoritması %77,94 ile her iki medikal grubun gerisinde kalırken %67,16 ile SSD algoritması sonuncu sırada başarı göstermiştir.

Tablo 5.9. Çoklu değerlendirme kriterleri tablosu test veri setine göre bütün sonuçlar.

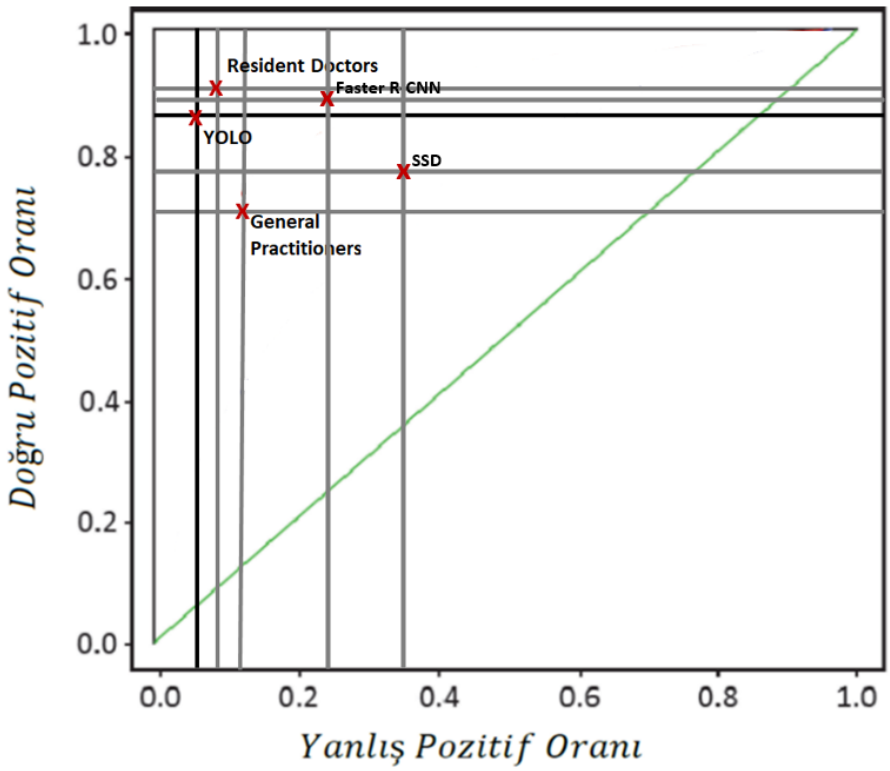
Kriterler	**Test veri setine göre YOLO Algoritması Sonuçları**	**Test veri setine göre Faster R-CNN Algoritması Sonuçları**	**Test veri setine göre SSD Algoritması Sonuçları**	**Test veri setine göre Pratisyen Hekimlerin Sonuçları**	**Test veri setine göre Uzman Doktorların Sonuçları**
Doğruluk Oranı	*90,33%*	*84,29%*	*72,74%*	*81,30%*	*91,42%*
Duyarlılık	*87,67%*	*90,64%*	*78,32%*	*74,80%*	*91,67%*
Özüllük	*92,98%*	*77,94%*	*67,16%*	*87,80%*	*91,17%*
Yanlış Pozitif Oranı	*7,02%*	*22,06%*	*32,84%*	*12,20%*	*8,83%*
Yanlış Negatif Oranı	*12,33%*	*9,36%*	*21,68%*	*25,20%*	*8,33%*
Doğru Tahmin Oranı	*92,58%*	*80,43%*	*70,46%*	*85,98%*	*91,21%*
F1 Puanı	*90,06%*	*85,23%*	*74,18%*	*80,00%*	*91,44%*
Hata Oranı	*9,67%*	*15,71%*	*27,26%*	*18,70%*	*8,58%*

İkisi insan çalışma grubu üçü derin öğrenme algoritması olmak üzere beş grubun en kritik performans sergileme alanlarından biri de sonuç üretme süreleridir. Sonuç üretme hızları kıyaslandığında (üç yapay zekâ birebir aynı iş istasyonu kullanılarak sonuç üretmiştir) beş grubun en hızlı sonuç üreteni insan çalışma gruplarından ortalama 3 kattan daha hızlı sonuç üreten YOLO algoritması olmuştur. YOLO algoritması ortalama 1,40 saniyede sonuç üretmekteyken uzman doktorlar ortalama 4,73 saniyede sonuç üretmiş, pratisyen hekimler 5,10 saniyede sonuç üretmiş, SSD algoritması 6,38 saniyede sonuç üretmiş, Faster R-CNN algoritması ise 10,49 saniyede sonuç üretmiştir. Bu araştırmada dikkat çeken noktalardan biri her üç algoritmanın da ortalama sonuç üretme sürelerinin standart sapması 0,25-0,23 saniye arasında değişirken, pratisyen hekimlerde bu süre 1,89 saniye ile 7,5 kat, uzman doktorlarda ise 3,05 saniye ile 12 kat fazla olmuştur.

Tablo 5.10. Bütün çalışma grularının sonuç üretme süreleri (saniye).

Test Veri Seti Süreleri	**YOLO Algoritması Sonuçları**	**Faster R-CNN Algoritması Sonuçları**	**SSD Algoritması Sonuçları**	**Pratisyen Hekim Grubu Sonuçları**	**Uzman Doktor Grubu Sonuçları**
Ortalama Tespit Süresi (sn)	*1,40*	*10,49*	*6,38*	*5,10*	*4,73*
Tespit Süresinin Standart Sapması (sn)	*0,25*	*0,22*	*0,23*	*1,89*	*3,05*
Maksimum Tespit Süresi (sn)	*2,27*	*10,99*	*6,89*	*10,18*	*13,72*
Minimum Tespit Süresi (sn)	*1,17*	*10,06*	*5,53*	*2,21*	*1,70*

Söz konusu 5 grubun performanslarını ortaya koymada bir diğer kritik analiz ROC analizidir. Aşağıda ROC grafiğinde görüldüğü üzere sol üst köşeye uzaklığı en az olan bunun sonucunda ROC grafiği analizine göre en başarılı kabul edilen 0,121 birim uzaklık ile uzman doktorlar gelmektedir sonrasında 0,141 birim uzaklıkla YOLO algoritması gelmekte, 0,239 birim uzaklık ile Faster R-CNN algoritması gelmekte, akabinde 0,279 birim uzaklık ile pratisyen hekimler, son olarak da be grup içerisinde en az başarılı kabul edilen ve sol üst köşeye en uzak olan 0,393 birim uzaklık ile SSD algoritması gelmektedir.



Şekil 5.17. ROC eğrisi analizinde bütün çalışma gruplarının sonuçları.

5.6. Bütün Derin Öğrenme Algoritmaları ile Hekim Gruplarının Ürettiği Sonuçların İstatistiksel Olarak Yorumlanması

Hem üç farklı yeniden eğitilen bilgisayarlı görü ve derin öğrenme algoritması hem de iki farklı hekim grubu ile yapılan çalışmanın neticesinde test veri seti ile sonuçlar elde edilmiş bu sonuçlar hem içerisinde Doğruluk Oranı, Duyarlılık, Özüllük, Yanlış Pozitif Oranı, Yanlış Negatif Oranı, Doğru Tahmin Oranı, F1 Puanı, Hata Oranını içeren çoklu değerlendirme kriterleri tablosu, hem sonuç üretme süreleri hem de ROC analizi eğrisi ile kıyaslanmış, derin öğrenme algoritmaları kendi içlerinde loss grafikleri ile kıyaslanmıştır.

Yapılan çalışma ile 50 femur üst bölge kırığı içeren 50 sağlam femur üst bölgesi içeren toplamda 55 farklı X-ray görüntüsü içeren veriler, Femur üst bölge sonuç üretme süreleri, femur üst bölge kırık olan bölge doğruluğu oranları ve femur üst bölge kırık olmayan bölge doğruluğu oranları ile seriler oluşturulmuştur. İlk seri olan Femur üst bölge sonuç üretme süreleri serisi her bir çalışma grubu için ayrı ayrı 55 olmak üzere toplamda 275 sonuç bulundurmakta, femur üst bölge kırık olan bölge doğruluğu oranları serisi her bir çalışma grubu için ayrı ayrı 50 toplamda 250 sonuç bulundurmakta, femur üst bölge kırık olmayan bölge doğruluğu oranları serisi de her

bir çalışma grubu için yine ayrı ayrı 50 toplamda ise toplamda 250 sonuç bulundurmaktadır.

Söz konusu sonuç serilerinin tek yön anova (varyans) anlamlılık testlerine tabi tutulmadan önce normal dağılımının kontrolü önemlidir. Normal dağılıma uygunluğunun kontrolü için her bir serinin ayrı olarak Skewness (çarpıklık) ve Kurtosis (basıklık) değerine bakılmıştır. Skewness (çarpıklık) değeri veri dağılımının sağa veya sola doğru ne kadar çarpık olduğunu ifade eden bir ölçüdür. Kurtosis (basıklık) değeri ise veri dağılımının tepesi (zirve) ne kadar sivri veya düz olduğunu belirten bir ölçüdür. Tabachnick ve Fidell'in 2013 yılında yapmış oldukları çalışmada serilerin çarpıklık ve basıklık değerlerinin -1,5 ile +1,5 arasında değişmesi halinde normal dağıldığının kabul edileceğini, George ve Mallery'in 2010 yılında yapmış oldukları çalışma ile çarpıklık ve basıklık değerlerinin -2,0 ile +2,0 arasında değişmesi halinde normal dağıldığının kabul edilmektedir (Tabachnick ve Fidell, 2013; George ve Mallery, 2010).

Aşağıda Femur üst bölge sonuç üretme süreleri serisi, femur üst bölge kırık olan bölge doğruluğu oranları serisi, femur üst bölge kırık olmayan bölge doğruluğu oranları serisi için her bir çalışma grubunun SPSS programı kullanılarak hesaplanan çarpıklık ve basıklık değerlerine yer verilmiştir.

Tablo 5.11. Femur üst bölge sonuç üretme süreleri çarpıklık ve basıklık analizi.

Femur Üst Bölge Sonuç Üretme Süresi		
	Skewness (Çarpıklık)	**Kurtosis (Basıklık)**
YOLO	1,291	1,853
Faster R-CNN	0,186	-0,377
SSD	-0,659	1,529
Pratisyen Hekim	0,872	0,351
Uzman Hekim	1,183	0,285

Tablo 5.12. Femur üst bölge kırık olan bölge doğruluğu çarpıklık ve basıklık analizi.

Femur Üst Bölge Kırık Olan Bölge Doğruluğu		
	Skewness (Çarpıklık)	**Kurtosis (Basıklık)**
YOLO	-1,778	1,567
Faster R-CNN	-1,056	1,617
SSD	-1,304	0,492
Pratisyen Hekim	-0,892	-0,280
Uzman Hekim	-1,921	1,969

Tablo 5.13. Femur üst bölge kırık olmayan bölge doğruluğu çarpıklık ve basıklık analizi.

Femur Üst Bölge Kırık Olmayan Bölge Doğruluğu		
	Skewness (Çarpıklık)	**Kurtosis (Basıklık)**
YOLO	-1,656	1,867
Faster R-CNN	-1,374	0,614
SSD	-0,570	-1,181
Pratisyen Hekim	-0,663	-0,603
Uzman Hekim	-1,810	1,201

Femur üst bölge sonuç üretme süreleri bakımından uzman doktorlar, pratisyen hekimler, YOLO algoritması, Faster R-CNN algoritması ve SSD algoritması tek yön anova testlerine tabi tutulur ve MiniTab programı aracılığı ile analiz edilirse aşağıda yer alan sonuç ortaya çıkmaktadır.

WORKSHEET 1

One-way ANOVA: Uzman; Pratisyen; YOLO; Faster-RCNN; SSD

Method

Null hypothesis	All means are equal
Alternative hypothesis	Not all means are equal
Significance level	α = 0,05

Equal variances were assumed for the analysis.

Factor Information

Factor	Levels	Values
Factor	5	Uzman; Pratisyen; YOLO; Faster-RCNN; SSD

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	4	2375,3	593,817	228,26	0,000
Error	270	702,4	2,602		
Total	274	3077,7			

Şekil 5.18. Femur üst bölge sonuç üretme süreleri tek yön anova analizi.

P değeri 0,05'den küçük olduğu için sonuç üretme süre ortalamalarının arasındaki fark anlamlıdır. Buna göre H0 reddedilir, H1 kabul edilir. H0: Bütün sonuç üretme süre ortalamaları eşittir.H1: En az bir ortalama diğerlerinden farklıdır.

Buna göre ikili karşılaştırma testlerinden olan Tukey ikili karşılaştırma testlerine devam edildiğinde, yine MiniTab programı ile şekil 5.19'da görülen sonuçlar elde edilmiştir.

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

Factor	N	Mean	Grouping			
Faster-RCNN	55	10,4884	A			
SSD	55	6,3825		B		
Pratisyen	55	5,100			C	
Uzman	55	4,734			C	
YOLO	55	1,3951				D

Means that do not share a letter are significantly different.

Şekil 5.19. Femur üst bölge sonuç üretme süreleri Tukey ikili karşılaştırma analizi.

Tukey pairwise analizinde göre aynı harf grubunu taşıyan ortalamalarının arasındaki fark anlamsızdır. Farklı harf grubu taşıyan ortalamalar arasındaki fark anlamlıdır. Buna göre Faster R-CNN Algoritmasının sonuç üretme süresi ortalaması; SSD Algoritması, Pratisyen hekim, Uzman hekim ve YOLO algoritmasına göre anlamlı bir şekilde yüksektir. Buna karşılık Pratisyen hekimlerin sonuç üretme süresi ortalaması Uzman Hekimlere göre anlamsızdır.

Femur üst bölge kırık olan bölge doğruluğu oranı bakımından uzman doktorlar, pratisyen hekimler, YOLO algoritması, Faster R-CNN algoritması ve SSD algoritması tek yön anova testlerine tabi tutulur ve MiniTab programı aracılığı ile analiz edilirse aşağıda yer alan sonuç ortaya çıkmaktadır.

WORKSHEET 1

One-way ANOVA: YOLO Tespit Kırık Bölge Doğrulu; Faster-RCNN

Method

Null hypothesis	All means are equal
Alternative hypothesis	Not all means are equal
Significance level	α = 0,05

Equal variances were assumed for the analysis.

Factor Information

Factor	Levels	Values
Factor	5	YOLO Tespit Kırık Bölge Doğrulu; Faster-RCNN Kırık Bölge Doğrulu; SSD Tespit Kırık Bölge Doğrulu; Pratisyen Hekim Kırık Bölge Doğ; Uzman Hekim Kırık Bölge Doğ

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	4	1,157	0,28917	4,84	0,001
Error	245	14,642	0,05976		
Total	249	15,799			

Şekil 5.20. Femur üst bölge kırık olan bölge doğruluğu oranı tek yön anova analizi.

P değeri 0,05'den küçük olduğu için kırık olan bölge doğruluğu oranlarının ortalamaları arasındaki fark anlamlıdır. Buna göre H0 reddedilir, H1 kabul edilir. H0: Bütün başarılı sonuç üretme oranlarının ortalamaları eşittir. H1: En az bir ortalama diğerlerinden farklıdır.

Tek yönlü varyans testleri sonrasında ikili karşılaştırma testlerinden olan Tukey ikili karşılaştırma testlerine devam edildiğinde, yine MiniTab programı ile Şekil 5.21'de yer alan sonuçlar elde edilmiştir.

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

Factor	N	Mean	Grouping		
Uzman Hekim Kırık Bölge Doğ	50	0,9167	A		
Faster-RCNN Kırık Bölge Doğrulu	50	0,9064	A	B	
YOLO Tespit Kırık Bölge Doğrulu	50	0,8767	A	B	C
SSD Tespit Kırık Bölge Doğrulu	50	0,7832		B	C
Pratisyen Hekim Kırık Bölge Doğ	50	0,7480			C

Means that do not share a letter are significantly different.

Şekil 5.21. Femur üst bölge kırık olan bölge doğruluğu oranı Tukey ikili karşılaştırma analizi.

Tukey pairwise analizinde göre aynı harf grubunu taşıyan ortalamalarının arasındaki fark anlamsızdır. Farklı harf grubu taşıyan ortalamalar arasındaki fark anlamlıdır. Buna göre Uzman Hekimlerin Kırık Bölge Doğruluğu başarı oranı SSD ve Pratisyen Hekimlerin başarı ortalamasına göre anlamlı bir şekilde yüksek iken Faster R-CNN ve YOLO Algoritmasına göre fark anlamsızdır.

Femur üst bölge kırık olmayan bölge doğruluğu oranı bakımından uzman doktorlar, pratisyen hekimler, YOLO algoritması, Faster R-CNN algoritması ve SSD algoritması tek yön anova testlerine tabi tutulur ve MiniTab programı aracılığı ile analiz edilirse aşağıda yer alan sonuç ortaya çıkmaktadır.

WORKSHEET 1

One-way ANOVA: YOLO Tespit Kırık Olmayan Bölge; Faster-RCNN

Method

Null hypothesis	All means are equal
Alternative hypothesis	Not all means are equal
Significance level	α = 0,05

Equal variances were assumed for the analysis.

Factor Information

Factor	Levels	Values
Factor	5	YOLO Tespit Kırık Olmayan Bölge; Faster-RCNN Kırık Olmayan Bölge; SSD Kırık Olmayan Bölge Doğrul; Pratisyen Hekim Kırık Olmayan B; Uzman Hekim Kırık Olmayan Bölge

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	4	2,325	0,58122	10,42	0,000
Error	245	13,666	0,05578		
Total	249	15,991			

Şekil 5.22. Femur üst bölge kırık olan bölge doğruluğu oranı tek yön anova analizi.

P değeri 0,05'den küçük olduğu için kırık olmayan bölge doğruluğu oranlarının ortalamaları arasındaki fark anlamlıdır. Buna göre H0 reddedilir, H1 kabul edilir. H0: Bütün başarılı sonuç üretme oranlarının ortalamaları eşittir. H1: En az bir ortalama diğerlerinden farklıdır.

Tek yönlü varyans testleri sonrasında ikili karşılaştırma testlerinden olan Tukey ikili karşılaştırma testlerine devam edildiğinde, yine MiniTab programı ile aşağıdaki sonuçlar elde edilmiştir.

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

Factor	N	Mean	Grouping		
YOLO Tespit Kırık Olmayan Bölge	50	0,9298	A		
Uzman Hekim Kırık Olmayan Bölge	50	0,9117	A		
Pratisyen Hekim Kırık Olmayan B	50	0,8780	A	B	
Faster-RCNN Kırık Olmayan Bölge	50	0,7794		B	C
SSD Kırık Olmayan Bölge Doğrul	50	0,6716			C

Means that do not share a letter are significantly different.

Şekil 5.23. Femur üst bölge kırık olmayan bölge doğruluğu oranı Tukey ikili karşılaştırma analizi.

Tukey pairwise analizinde göre aynı harf grubunu taşıyan ortalamalarının arasındaki fark anlamsızdır. Farklı harf grubu taşıyan ortalamalar arasındaki fark anlamlıdır. Buna göre YOLO Algoritmasının Kırık Olmayan Bölge Doğruluğu başarı oranı Faster R-CNN ve SSD Algoritmalarının başarı ortalamasına göre anlamlı bir şekilde yüksek iken Uzman ve Pratisyen Hekimlere göre fark anlamsızdır.

5.7. Sonuç ve Öneriler

Gerçekleştirilen araştırma dünyada 2050 yılına kadar femur üst bölge yani kalça kırıklarındaki vaka sayısının 3 kattan fazla artması beklenerek globalde 4,5 milyon vakaya ulaşması beklenmektedir. Çalışmada ortopedi kırıkları arasında en yüksek ölüm ve sakat kalma riski bulunan femur üst bölge kırıklarına odaklanmıştır. Toplamda 500'den fazla femur üst bölge kırığı vakası incelenmiş bu süreçte Türkiye'nin en büyük şehri olan İstanbul'da yer alan bir eğitim ve araştırma hastanesi uzman hekimleri ve pratisyen hekimlerinin katılımları ile yakinen çalışma gerçekleştirilmiştir.

Çalışmanın derin öğrenme ve bilgisayarlı görü algoritmalarının eğitimi aşamasında YOLO algoritmasında bir bulut teknolojisi kullanılmıştır. Söz konusu eğitimde aktarımlı öğrenme ile özgün algoritma geliştirilirken Tesla K80 24 GB GDDR5 GPUsu ile Google Colab bulut teknolojisi olarak kullanılmıştır. Dünyanın en güçlü GPU'larından biri yapılan çalışmanın eğitim aşamasında kullanılmasına rağmen bu eğitim 13 saat 6 dakikada tamamlanabilmiştir. Faster R-CNN algoritması ve SSD algoritması ise Core i5-8300H CPU ve NVIDIA GeForce GTX 1050 GPU ya sahip iş

istasyonu GPU'su ile yeniden eğitilme süreci gerçekleştirilmiş sırası ile eğitim Faster R-CNN algoritması için 31,9 dakika ve SSD algoritması için ise 55,92 sürmüştür.

Yapılan çalışmada veri çoklandırması ile eğitim ve doğrulama veri setinde toplamda 820 X-ray görüntüsü kullanılmıştır. Test veri setinde 55 X-ray görüntüsü olmak üzere toplamda her üç veri seti için toplamda 875 X-ray görüntüsü ele alınmıştır. Piksel çözünürlükleri ve renk matrisleri de hesaplandığında kullanılan ve işlenen verinin boyutu oldukça büyümekte toplamda 4,3 milyardan fazla veriye erişmektedir. Söz konusu çalışmada toplamda 4,3 milyardan fazla işlenerek sonuca erişilmiştir.

Hem ortopedi ve travmatoloji uzman hekimler hem pratisyen hekimler hem de derin öğrenmenin en güçlü olarak kabul edilebilecek algoritmaları ile bütün veri setlerinden bağımsız bir test veri seti ile inceme yapıldığında, uzman hekimler ile YOLO algoritması arasında %1,5'dan daha az bir doğruluk oranı görülmesine karşın, derin öğrenme 3 kat daha hızlı sonuç üretmiştir. Uzman hekimler %91,42 oranında doğruluk oranı başarısı elde ederken YOLO algoritması doğruluk oranında %90,33 oranında başarı elde etmiştir. Derin öğrenme algoritmalarından yeniden eğitilmiş YOLO algoritması özüllük oranında %92,98 ile uzman doktorlardan daha başarılı olmuştur. Uzman doktorların özüllük oranı %91,12 seviyesinde kalmıştır. Doğruluk oranında Faster R-CNN algoritması pratisyen hekimlerden %3'den fazla daha iyi sonuç üretmesine karşın 2 kat yavaş kalmıştır. Faster R-CNN doğruluk oranı %84,29 iken pratisyen hekimlerin doğruluk oranı %81,30'dur. Yapılan çalışma ile yapay zekâ ve derin öğrenme algoritmalarının çok çeşitli performansı ortaya konulmuş, neredeyse bir uzman doktor kadar başarı ortaya koyabileceği kanıtlanmıştır. Derin öğrenme algoritmalarının tamamında sonuç üretmesi için kullanılan iş istasyonu yerine daha güçlü bir GPU'ya sahip iş istasyonu kullanılması halinde bu sürenin daha da kısa süreceği yadsınamaz bir gerçektir. Daha çeşitli kırıklar ile veri setinin genişletilmesi hem çok geniş yelpazede kemik kırıklarının teşhisini hem de bu teşhisteki başarı oranını oldukça arttırması beklenmektedir. Özellikle daha güçlü GPU'lar ile işlem yapılması hem eğitim süresinin hem de tespit süresini aşağıya çekecektir.

Yapılan çalışmada ileri istatistiksel analiz kullanarak üretilen sonuçlar arasındaki ilişki çalışma gruplarının hem sonuç üretme süreleri hem femur üst bölge kırık olan bölge doğruluğu oranı hem de kırık olmayan bölge doğruluğu oranı noktasındaki benzerlikleri ve farklılıkları hem tek yönlü Anova Testleri hem de Tukey Pairwise testleri ile kıyaslanarak ortaya konulmuştur.

Bu çalışma ile hem ortopedi hekimlerine hem de pratisyen hekimlere femur üst bölge kırıklarının tespit ve teşhis edilmesinde birçok derin öğrenmesi algoritmasının görüş ortaya koyabildiği başarılı bir sistemler üretilmiştir. Üretilen sistemler yapay zekâ ve derin öğrenme tabanlı olması nedeniyle yorgunluk, dikkat dağınıklığı, gürültü, hastaların geliş sırası vs. ortam değişken şartlarından etkilenmeden her zaman objektif sonuçları benzer başarı ortalamaları ile üretecektir.

Yapılan çalışma ile derin öğrenmenin vardığı noktanın sadece günlük objeler ve rakamları tespit etmek, resimdeki suretleri karelerin içerisine almak yerine, kendi sınırlarını aşarak femur üst bölge gibi oldukça farklı vakaların tespiti noktasında işaretleyebileceği, tespit edebileceğini göstermiştir. Aynı zamanda objektif veri seti olan test veri setinde de ne kadar başarılı ve tarafsız olabileceği kanıtlanmıştır. Medikal alanda görüntü tespitinde sonraki çalışmalara ışık tutacak bu çalışma oluşturulan ve genişletilen özgün veri setleri ile başarı oranının daha da yükseleceğini ortaya koymaktadır.

1950 yılında henüz makinelerin düşünüp düşünemeyeceği tartışılırken; 1956 yılından sonra yapay zekâ kavramının ortaya atılmış bundan da 67 yıl sonra, artık yapay genel zekâ ve yapay süper zeka kavramları tartışılmaya başlanmıştır. Yapay zekanın derin uykusundan uyanışından bu yana GPU ve CPU'lar hiç bu kadar güçlü olmamış ve bu kadar başarılı algoritmaların geliştirilerek insanlığın hizmetine sunulmasına imkân vermemiştir. Günümüzde yapay dar zekâ kullanılıyor olsa da şu an yer almadığı sektör sayısı oldukça azdır. Özellikle üretici yapay zekalar birçok alanda farklarını ortaya koymakta özgün çıktılar üretebilmektedir.

Ortaya konulan çalışmanın sonrasında gelecek çalışmaların daha başarılı olmasında önemli bir husus çok daha fazla çeşitli kırıklar içeren x-ray görüntülerinin derin öğrenme modellerinin eğitiminde kullanılmasından geçmektedir. Bunun için gelişmiş ve entegre sağlık sistemine sahip ülkemizde çalışmalar bazında x-ray veri setlerinin anonim dahi olsa araştırmalar sırasında kullanıma açılmasından geçmektedir.

Günümüzde üzerine çalışılan kuantum bilgisayarları aracılığı ile insanlığın teknoloji ve yapay zekâ alanında büyük bir atılım yapması beklenmekte; veri işleme hızı noktasında ve veri işleme kapasitesi noktasındaki dar boğazların ortadan kalkması beklenmektedir. Olası kuantum bilgisayarların yaygın kullanımının genişletilmesi halinde oldukça yıkıcı teknolojilerin hayatımıza gireceği kaçınılmaz bir gerçektir.

KAYNAKLAR

Adams, M., Chen, W., Holcdorf, D., McCusker, M. W., Howe, P. D., & Gaillard, F. (2019). Computer vs human: deep learning versus perceptual training for the detection of neck of femur fractures. *Journal of medical imaging and radiation oncology*, *63*(1), 27-32. https://doi.org/10.1111/1754-9485.12828

Agar, A. & Sahin, A. (2022). Top 100 cited articles on geriatric hip fractures in orthopaedics: a bibliometric and visualised analysis. *Dicle Tıp Dergisi*, *49*(1), 102-110. https://doi.org/10.5798/dicletip.1086274

Agrawal, T., & Choudhary, P. (2023). Segmentation and classification on chest radiography: a systematic survey. *The Visual Computer*, *39*(3), 875-913. https://doi.org/10.1007/s00371-021-02352-7

Alimovski, E., & Erdemir, G. (2021). Veri artırma tekniklerinin derin öğrenmeye dayalı yüz tanıma sisteminde etkisi. *İstanbul Sabahattin Zaim Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, *3*(1), 76-80. https://doi.org/10.47769/izufbed.880581

Andrew, W. (2023, Mart 27). human skeleton. Encyclopedia Britannica. https://www.britannica.com/science/human-skeleton

Archana, B., & Kalirajan, K. (2022). A Survey of Medical Image Processing and its Applications. In *2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA)* (pp. 1541-1546). IEEE. http://dx.doi.org/10.1109/icirca54612.2022.9985621

Basha, C. Z., Padmaja, T. M., & Balaji, G. N. (2019). An Effective and Reliable Computer Automated Technique for Bone Fracture Detection. *EAI Endorsed Transactions on Pervasive Health and Technology*, *5*(18), e2-e2. https://doi.org/10.4108/eai.13-7-2018.162402

Balaji, G. N., Subashini, T. S., Madhavi, P., Bhavani, C. H., & Manikandarajan, A. (2020). Computer-aided detection and diagnosis of diaphyseal femur fracture. In *Smart Intelligent Computing and Applications: Proceedings of the Third International Conference on Smart Computing and Informatics, Volume 1* (pp. 549-559). Springer Singapore. https://doi.org/10.1007/978-981-13-9282-5_52

Berglund-Rödén, M., Swierstra, B. A., Wingstrand, H., & Thorngren, K. G. (1994). Prospective comparison of hip fracture treatment: 856 cases followed for 4 months in The Netherlands and Sweden. *Acta Orthopaedica Scandinavica*, *65*(3), 287-294. https://doi.org/10.3109/17453679408995455

Bishop, C. M. (1995). Regularization and complexity control in feed-forward networks. Unpublished. In *Proceedings International Conference on Artificial Neural Networks ICANN'95* (pp. 141-148). EC2 et Cie.

Breiman, L., Friedman, J., Olshen, R. A., & Stone, C. J. (1984). Classification and Regression Trees. Monterey, CA: Wadsworth & Brooks. *Cole Advanced Books*

Budhiraja, A. (2016). Dropout in (deep) machine learning. *Retrieved May*, *9*, 2020. https://medium.com/@amarbudhiraja/https-medium-com-amarbudhiraja-learning-less-to-learn-better-dropout-in-deep-machine-learning-74334da4bfc5

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992, July). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory* (pp. 144-152). http://dx.doi.org/10.1145/130385.130401

Caruana, R. A. (2014, March). Multitask connectionist learning. In *Proceedings of the 1993 connectionist models summer school* (pp. 372-379). Psychology Press.

Chapelle, O., Weston, J., & Schölkopf, B. (2002). Cluster kernels for semi-supervised learning. *Advances in neural information processing systems*, *15*.

Chen, K. (2017). How to interpret "loss" and "accuracy" for a machine learning model. *https://stackoverflow.com/questions/34518656/how-tointerpretloss-and-accuracy-for-a-machine-learning-model*.

Cheng, C. T., Ho, T. Y., Lee, T. Y., Chang, C. C., Chou, C. C., Chen, C. C., ... & Liao, C. H. (2019). Application of a deep learning algorithm for detection and visualization of hip fractures on plain pelvic radiographs. *European radiology*, *29*(10), 5469-5477. https://doi.org/10.1007/s00330-019-06167-y

Choi, J. W., Cho, Y. J., Lee, S., Lee, J., Lee, S., Choi, Y. H., ... & Ha, J. Y. (2020). Using a dual-input convolutional neural network for automated detection of pediatric supracondylar fracture on conventional radiography. *Investigative radiology*, *55*(2), 101-110. http://dx.doi.org/10.1097/RLI.0000000000000615

Chollet, F. (2021). *Deep learning with Python*. Simon and Schuster.

Chowdhury, A. S., Burns, J. E., Mukherjee, A., Sen, B., Yao, J., & Summers, R. M. (2012). Automated detection of pelvic fractures from volumetric CT images. In *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)* (pp. 1687-1690). IEEE.

Cohen, M., Puntonet, J., Sanchez, J., Kierszbaum, E., Crema, M., Soyer, P., & Dion, E. (2023). Artificial intelligence vs. radiologist: accuracy of wrist fracture detection on radiographs. European Radiology, 33(6), 3974-3983. http://dx.doi.org/10.1007/s00330-022-09349-3

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*, 273-297. http://dx.doi.org/10.1007/BF00994018

Çakır, H., & Eryılmaz, S., (2014). Bilgisayarın Tarihi Gelişimi. *Eğitimciler için Bilişim Teknolojileri* (pp.1-25), Ankara: Pegem Akademi Yayıncılık.

DeLee J.C. (1984): Fractures and dislocations of the hip. In Rockwood CA Jr (eds): Fractures in Adults, ed 2, JB Lippincott, Philadephia.

Dirican, A. (2001). Tani Testi Performanslarinin Değerlendirilmesi Ve Kiyaslanmasi. *Cerrahpaşa Tıp Dergisi*, *32*(1). https://dergipark.org.tr/tr/pub/iucerrahpasa/issue/665/6609

Eksi, Z., & Cakiroglu, M. (2012). Performance evaluation of the popular segmentation algorithms for bone fracture detection. *AWERProcedia Information Technology & Computer Science*, *1*, 1245-1249.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern recognition letters*, *27*(8), 861-874. https://doi.org/10.1016/j.patrec.2005.10.010

Filipov O. (2014), "Epidemiology and social burden of the femoral neck fractures," J IMAB., vol. 20, no.4, pp. 516-518, http://dx.doi.org/10.5272/jimab.2014204.516

Freeman, L. (1904). X. The Union of Ununited Fractures of the Neck of the Femur by Open Operation. *Annals of surgery*, *40*(4), 561. https://doi.org/10.1097/00000658-190410000-00011

Gallagher, J. C., Melton, L. J., Riggs, B. L., & Bergstrath, E. (1980). Epidemiology of fractures of the proximal femur in Rochester, Minnesota. *Clinical Orthopaedics and Related Research (1976-2007)*, *150*, 163-171.

Gan, K., Xu, D., Lin, Y., Shen, Y., Zhang, T., Hu, K., ... & Liu, Y. (2019). Artificial intelligence detection of distal radius fractures: a comparison between the convolutional neural network and professional assessments. *Acta orthopaedica*, *90*(4), 394-400. https://doi.org/10.1080/17453674.2019.1600125

Gao, M., Cai, Q., Zheng, B., Shi, J., Ni, Z., Wang, J., & Lin, H. (2021). A Hybrid YOLOv4 and Particle Filter Based Robotic Arm Grabbing System in Nonlinear and Non-Gaussian Environment. *Electronics*, *10*(10), 1140. https://doi.org/10.3390/electronics10101140

George, D. (2011). SPSS for windows step by step: A simple study guide and reference,. *Pearson Education India*.

Girshick, R., Donahue, J., Darrell, Malik, T. J., Berkeley, U. C. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. *Computer Vision and Pattern Recognition*, 1-21. https://doi.org/10.48550/arXiv.1311.2524

Girshick, R., (2015). Fast R-CNN. *Proceedings of the IEEE international conference on computer vision*, ss 1440-1448. https://doi.org/10.48550/arXiv.1504.08083

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.

Graves, A. (2011). Practical variational inference for neural networks. *Advances in neural information processing systems*, *24*.

Guan, B., Yao, J., Zhang, G., & Wang, X. (2019). Thigh fracture detection using deep learning method based on new dilated convolutional feature pyramid network. *Pattern Recognition Letters*, *125*, 521-526. https://doi.org/10.1016/j.patrec.2019.06.015

Guan, B., Zhang, G., Yao, J., Wang, X., & Wang, M. (2020). Arm fracture detection in X-rays based on improved deep convolutional neural network. *Computers & Electrical Engineering*, *81*, 106530.

Hall, J. E., & Hall, M. E. (2020). Guyton and Hall textbook of medical physiology e-Book. *Elsevier Health Sciences*.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778). https://doi.org/10.1109/CVPR.2016.90

Heimer, J., Thali, M. J., & Ebert, L. (2018). Classification based on the presence of skull fractures on curved maximum intensity skull projections by means of deep learning. *Journal of Forensic Radiology and Imaging*, *14*, 16-20. https://doi.org/10.1016/j.jofri.2018.08.001

Ho, Y., & Wookey, S. (2019). The real-world-weight cross-entropy loss function: Modeling the costs of mislabeling. *IEEE access*, *8*, 4806-4813. http://dx.doi.org/10.1109/ACCESS.2019.2962617

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M. & Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. https://doi.org/10.48550/arXiv.1704.04861

Huang, Y. Q., Zheng, J. C., Sun, S. D., Yang, C. F., & Liu, J. (2020). Optimized YOLOv3 algorithm and its application in traffic flow detections. *Applied Sciences*, *10*(9), 3079. http://dx.doi.org/10.3390/app10093079

Hughes, P. E., Hsu, J. C., & Matava, M. J. (2002). Hip anatomy and biomechanics in the athlete. *Sports medicine and arthroscopy review*, *10*(2), 103-114. http://dx.doi.org/10.1097/00132585-200210020-00002

ImageNet, ImageNet Large Scale Visual Recognition Challenge (2016)., Available: http://image-net.org/challenges/LSVRC/2016/results#team, 2016, Erişim Tarihi: 23.08.2023.

Özkan, İ. N. İ. K., & Ülker, E. (2017). Derin öğrenme ve görüntü analizinde kullanılan derin öğrenme modelleri. *Gaziosmanpaşa Bilimsel Araştırma Dergisi*, *6*(3), 85-104.

Ju, M., Luo, H., Wang, Z., Hui, B., & Chang, Z. (2019). The application of improved YOLO V3 in multi-scale target detection. *Applied Sciences*, *9*(18), 3775. https://doi.org/10.3390/app9183775

Karagas, M. R., Lu-Yao, G. L., Barrett, J. A., Beach, M. L., & Baron, J. A. (1996). Heterogeneity of hip fracture: age, race, sex, and geographic patterns of femoral neck and trochanteric fractures among the US elderly. *American journal of epidemiology*, *143*(7), 677-682. https://doi.org/10.1093/oxfordjournals.aje.a008800

Karimunnisa, S., Savarapu, P. R., Madupu, R. K., Basha, C. Z., & Neelakanteswara, P. (2020, July). Detection of Bone Fractures Automatically with Enhanced Performance with Better Combination of Filtering and Neural Networks. In *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)* (pp. 189-193). IEEE. https://doi.org/10.1109/ICIRCA48905.2020.9183085

Karpathy, A. (2018). Stanford university cs231n: Convolutional neural networks for visual recognition. *URL: http://cs231n. stanford. edu/syllabus. html*, 13-35.

Karunakaran, D. (2018). Deep learning series 1: Intro to deep learning. *linea}.{2018}. Disponible en:(https://medium.com/intro-toartificial-intelligence/deep-learning-series-1-intro-to-deep-learningabb1780ee20)*.

Kasban, H., El-Bendary, M. A. M., & Salama, D. H. (2015). A comparative study of medical imaging techniques. *International Journal of Information Science and Intelligent System*, *4*(2), 37-58.

Kayri, M. (2009). Araştirmalarda gruplar arasi farkin belirlenmesine yönelik çoklu karşilaştirma (post-hoc) teknikleri. *Journal of Social Science*, *55*, 22.

Keating JF, Femoral neck fractures. In: Court-Brown CM, Heckman, J. D., McKee, M., McQueen, M. M., Ricci, W., & Tornetta III, P. (2014). *Rockwood and Green's fractures in adults*. Lippincott Williams & Wilkins.

Kim, D. H., & MacKinnon, T. (2018). Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks. *Clinical radiology*, *73*(5), 439-445. https://doi.org/10.1016/j.crad.2017.11.015

Kim, W., Cho, H., Kim, J., Kim, B., & Lee, S. (2020). YOLO-based simultaneous target detection and classification in automotive FMCW radar systems. *Sensors*, *20*(10), 2897. http://dx.doi.org/10.3390/s20102897

Kitamura, G., Chung, C. Y., & Moore, B. E. (2019). Ankle fracture detection utilizing a convolutional neural network ensemble implemented with a small sample, de novo training, and multiview incorporation. *Journal of digital imaging*, *32*, 672-677. https://doi.org/10.1007/s10278-018-0167-7

Kumrular, R. K., & Polat, A. (2021). Klinik uygulamalarda ileri biyomedikal görüntüleme teknolojileri. *Avrupa Bilim ve Teknoloji Dergisi*, (23), 207-221. https://doi.org/10.31590/ejosat.840321

Lane, J. E., & Gantley, M. J. (2018). Utilizing complex systems statistics for historical and archaeological data. *Journal of Cognitive Historiography*, *3*(1-2), 68-92. http://dx.doi.org/10.1558/jch.31696

Lasko, T. A., Bhagwat, J. G., Zou, K. H., & Ohno-Machado, L. (2005). The use of receiver operating characteristic curves in biomedical informatics. *Journal of biomedical informatics*, 38(5), 404-415. http://dx.doi.org/10.1016/j.jbi.2005.02.008

Lattimore, R. (Ed.). (1962). *The Iliad of Homer*. CUP Archive.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444. https://doi.org/10.1038/nature14539

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278-2324. https://doi.org/10.1109/5.726791

Li Z., Li, S., Bamasag O. O., Alhothali A. and Luo X., (2022) "Diversified Regularization Enhanced Training for Effective Manipulator Calibration," in IEEE Transactions on Neural Networks and Learning Systems. https://doi.org/10.1109/TNNLS.2022.3153039

Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., & Zitnick, C. L. Microsoft coco: Common objects in context. InEuropean conference on computer vision 2014 Sep 6 (pp. 740-755). https://doi.org/10.1007/978-3-319-10602-1_48

Lindsey, R., Daluiski, A., Chopra, S., Lachapelle, A., Mozer, M., Sicular, S., ... & Potter, H. (2018). Deep neural network improves fracture detection by clinicians. *Proceedings of the National Academy of Sciences*, *115*(45), 11591-11596. https://doi.org/10.1073/pnas.1806905115

Liu, P. R., Zhang, J. Y., Xue, M. D., Duan, Y. Y., Hu, J. L., Liu, S. X., ... & Ye, Z. W. (2021). Artificial intelligence to diagnose tibial plateau fractures: An intelligent assistant for orthopedic physicians. *Current Medical Science*, *41*(6), 1158-1164. https://doi.org/10.1007/s11596-021-2501-4

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). Ssd: Single shot multibox detector. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14* (pp. 21-37). Springer International Publishing. http://dx.doi.org/10.1007/978-3-319-46448-0_2

Lowell, J. D. (1966). Fractures of the hip. *New England Journal of Medicine*, *274*(25), 1418-1425. https://doi.org/10.1056/nejm196606302742607

Ma, Y., & Luo, Y. (2021). Bone fracture detection through the two-stage system of crack-sensitive convolutional neural network. *Informatics in Medicine Unlocked*, *22*, 100452. https://doi.org/10.1016/j.imu.2020.100452

Mahendran, S. K., & Baboo, S. S. (2012). Ensemble systems for automatic fracture detection. *International Journal of Engineering and Technology*, *4*(1), 7. https://doi.org/10.7763/IJET.2012.V4.310

Majkowska, A., Mittal, S., Steiner, D. F., Reicher, J. J., McKinney, S. M., Duggan, G. E. & Shetty, S. (2020). Chest radiograph interpretation with deep learning models: assessment with radiologist-adjudicated reference standards and population-adjusted evaluation. *Radiology*, *294*(2), 421-431. https://doi.org/10.1148/radiol.2019191293

Massouh, N., Babiloni, F., Tommasi, T., Young, J., Hawes, N., & Caputo, B. (2017, September). Learning deep visual object models from noisy web data: How to make it work. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 5564-5571). IEEE. MathWorks, (2018). «What Is Deep Learning.», mathworks.com/discovery/deep-learning.html, [Erişim Tarihi: 22.08.2023].

Meinberg, E. G., Agel, J., Roberts, C. S., Karam, M. D., & Kellam, J. F. (2018). Fracture and dislocation classification compendium—2018. *Journal of orthopaedic trauma*, *32*, S1-S10. https://doi.org/10.1097/bot.0000000000001063

Metz, C. E. (2006). Receiver operating characteristic analysis: a tool for the quantitative evaluation of observer performance and imaging systems. *Journal of the American College of Radiology*, *3*(6), 413-422. http://dx.doi.org/10.1016/j.jacr.2006.02.021

Mitchell, T. M. (1997). *Machine Learning* . McGraw-Hill, Newyork. 96

Molino, G., Montalbano, G., Pontremoli, C., Fiorilli, S., & Vitale-Brovarone, C. (2020). Imaging techniques for the assessment of the bone osteoporosis-induced variations with particular focus on micro-ct potential. *Applied Sciences*, *10*(24), 8939. https://doi.org/10.3390/app10248939

Moharil, A., & Singh, S. (2021). Detection of Proximal Femoral Bone Fracture Using Mask R-CNN. In *Proceedings of International Conference on Intelligent Computing, Information and Control Systems: ICICCS 2020* (pp. 239-253). Springer Singapore. http://dx.doi.org/10.1007/978-981-15-8443-5_19

Mu L., Qu T. Dong D., Li X., Pei Y., Wang Y., Shi G., Li Y., He F., Zang H. (2021). Fine-Tuned Deep Convolutional Networks for the Detection of Femoral Neck Fractures on Pelvic Radiographs: A Multicenter Dataset Validation. *IEEE Access*, *9*, 78495-78503. http://dx.doi.org/10.1109/ACCESS.2021.3082952

Muehlematter, U. J., Mannil, M., Becker, A. S., Vokinger, K. N., Finkenstaedt, T., Osterhoff, G., ... & Guggenberger, R. (2019). Vertebral body insufficiency fractures: detection of vertebrae at risk on standard CT images using texture analysis and machine learning. *European radiology*, *29*, 2207-2217. https://doi.org/10.1007/s00330-018-5846-8

Mukesh, R., & Dass, P. (2022). Detection by Clinicians Comparison of Intelligence Bone Fracture Detection System with SIFT algorithm for Identification of Bone Fracture. In *2022 3rd International Conference on Intelligent Engineering and Management (ICIEM)* (pp. 533-537). IEEE.

Murat M; Ermutlu C. (2020). Opere Kalça Kırığı Olgularında D Vitamini ve Biokimyasal Parametrelerin Değerlendirilmesi. *Dicle Tıp Dergisi*, *47*(2), 446-454. https://doi.org/10.5798/dicletip.755774

Naderian, A., & Sanei, S. (2022). Automated Bone Segmentation Using a Hybrid Model of DRLSE and Faster R-CNN. In *2022 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)* (pp. 81-85). IEEE.

Nguyen, A. Q., Nguyen, H. T., Tran, V. C., Pham, H. X., & Pestana, J. (2021). A visual real-time fire detection using single shot multibox detector for uav-based fire surveillance. In *2020 IEEE Eighth International Conference on Communications and Electronics (ICCE)* (pp. 338-343). IEEE. http://dx.doi.org/10.1109/ICCE48956.2021.9352080

Nguyen, H. P., Hoang, T. P., & Nguyen, H. H. (2021, October). A deep learning based fracture detection in arm bone X-ray images. In *2021 international conference on multimedia analysis and pattern recognition (MAPR)* (pp. 1-6). IEEE. https://doi.org/10.1109/MAPR53640.2021.9585292

Nielsen, M. A. (2015). *Neural networks and deep learning* (Vol. 25, pp. 15-24). San Francisco, CA, USA: Determination press.

Oyar, O., (2008). Magnetik Rezonans Görüntüleme Mrg Nin Klinik Uygulamaları Ve Endikasyonları. *Harran Üniversitesi Tıp Fakültesi Dergisi*, *5*(2), 31-40.

Özçelik, T. Ö., Selvi, İ. H., Yalçıner, A. Y., & Zeren, M. T. (2022). Industry 4.0, Smart Factories And Effects On Business. *Knowledge Management and Digital Transformation Power*, 169.

Özdemir, Ş., & Kılınç, D. (2019). Geleceğin meslekleri. İstanbul: *Abaküs Yayınları*.

Polat, H., Mehr, D. H., & Cetin, A. (2017). Diagnosis of chronic kidney disease based on support vector machine by feature selection methods. *Journal of medical systems*, *41*, 1-11. https://doi.org/10.1007/s10916-017-0703-x

Pranata, Y. D., Wang, K. C., Wang, J. C., Idram, I., Lai, J. Y., Liu, J. W., & Hsieh, I. H. (2019). Deep learning and SURF for automated classification and detection of calcaneus fractures in CT images. *Computer methods and programs in biomedicine*, *171*, 27-37. https://doi.org/10.1016/j.cmpb.2019.02.006

Rahmaniar, W., & Wang, W. J. (2019). Real-time automated segmentation and classification of calcaneal fractures in CT images. *Applied Sciences*, *9*(15), 3011. https://doi.org/10.3390/app9153011

Rakesh, Y., & Akilandeswari, A. (2022). Bone Fracture Detection Using Morphological and Comparing the Accuracy with Genetic Algorithm. *Journal of Pharmaceutical Negative Results*, 270-276. https://doi.org/10.47750/pnr.2022.13.S04.030

Rao, L. J., Ramkumar, M., Kothapalli, C., Savarapu, P. R., & Basha, C. Z. (2020, August). Advanced computerized Classification of X-ray Images using CNN. In 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 1247-1251). IEEE. http://dx.doi.org/10.1109/ICSSIT48917.2020.9214136

Rayan, J. C., Reddy, N., Kan, J. H., Zhang, W., & Annapragada, A. (2019). Binomial classification of pediatric elbow fractures using a deep learning multiview approach emulating radiologist decision making. *Radiology: Artificial Intelligence*, *1*(1), e180015. https://doi.org/10.1148/ryai.2019180015

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, *28*. https://doi.org/10.48550/arXiv.1506.01497

Russel, S. ve Norvig, P., (2009). Artificial intelligence: A modern approach. *Prentice Hall*. ss 17.

Sağlık Bakanlığı. (Mart 10, 2022). *Sağlık İstatistikleri Yıllığı 2020*. https://dosyasb.saglik.gov.tr/Eklenti/43399,siy2020-tur-26052022pdf.pdf?0 adresinden alındı

Samothai, P., Sanguansat, P., Kheaksong, A., Srisomboon, K., & Lee, W. (2022, July). The Evaluation of Bone Fracture Detection of YOLO Series. In *2022 37th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC)* (pp. 1054-1057). IEEE. http://dx.doi.org/10.1109/ITC-CSCC55581.2022.9895016

Shinde, S., Kothari, A., & Gupta, V. (2018). YOLO based human action recognition and localization. *Procedia computer science*, *133*, 831-838. https://doi.org/10.1016/j.procs.2018.07.112

Shafiee, M. J., Chywl, B., Li, F., & Wong, A. (2017). Fast YOLO: A fast you only look once system for real-time embedded object detection in video. *arXiv preprint arXiv:1709.05943*. https://doi.org/10.48550/arXiv.1709.05943

Shanmugamani, R. (2018). *Deep Learning for Computer Vision: Expert techniques to train advanced neural networks using TensorFlow and Keras*. Packt Publishing Ltd.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*. https://doi.org/10.48550/arXiv.1409.1556

Singh, M., Nagrath, A., & Maini, P. S. (1970). Changes in trabecular pattern of the upper end of the femur as an index of osteoporosis. *JBJS*, *52*(3), 457-467. https://doi.org/10.2106/00004623-197052030-00005

Strelkova, O., Pasichnyk, O., (2017). Three types of artificial intelligence. *Khmelnitsky National University: Khmelnytskyi* (pp. 1-4).

Sweeney, H. L., & Hammers, D. W. (2018). Muscle Contraction. *Cold Spring Harbor perspectives in biology*, 10(2), a023200. https://doi.org/10.1101/cshperspect.a023200

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-9). https://doi.org/10.1109/CVPR.2015.7298594

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818-2826). https://doi.org/10.1109/CVPR.2016.308

Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I., & Fergus, R. (2013). Intriguing properties of neural networks. *arXiv preprint arXiv:1312.6199*. https://doi.org/10.48550/arXiv.1312.6199

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2013). *Using multivariate statistics* (Vol. 6, pp. 497-516). Boston, MA: pearson.

Thian, Y. L., Li, Y., Jagmohan, P., Sia, D., Chan, V. E. Y., & Tan, R. T. (2019). Convolutional neural networks for automated fracture detection and localization on wrist radiographs. *Radiology: Artificial Intelligence*, *1*(1), e180001. https://doi.org/10.1148%2Fryai.2019180001

Peng, T. T. (2002). Detection of Femur Fractures in X-ray images. *Master of Science Thesis, National University of Singapore*.

Tomita, N., Cheung, Y. Y., & Hassanpour, S. (2018). Deep neural networks for automatic detection of osteoporotic vertebral fractures on CT scans. *Computers in biology and medicine*, *98*, 8-15. http://dx.doi.org/10.1016/j.compbiomed.2018.05.011

Torkul O, Tunacan T, (2022). Knowledge Management and Digital Transformation Power. *Efe Akademik Publishing*, İstanbul, 170

Tripathi, A. M., Upadhyay, A., Rajput, A. S., Singh, A. P., & Kumar, B. (2017, March). Automatic detection of fracture in femur bones using image processing. In *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)* (pp. 1-5). IEEE. http://dx.doi.org/10.1109/ICIIECS.2017.8275843

Meral, T. (2013). Computer Simulation of Bone Remodeling. İstanbul Technical University,

Montgomery, D. C., & Runger, G. C. (2010). Applied statistics and probability for engineers. John wiley & sons.

Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.

Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics*, 99-114. https://doi.org/10.2307/3001913

Turgut, A., Kumbaracı, M., Kalenderer, Ö., İlyas, G., Bacaksız, T., & Karapınar, L. (2016). Is surgeons' experience important on intra-and inter-observer reliability of classifications used for adult femoral neck fracture?. *Acta orthopaedica et traumatologica turcica*, *50*(6), 601-605. https://doi.org/10.1016/j.aott.2015.11.004

Turing, A. M. (1950). Computing machinery and intelligence *The Essential Turing: the Ideas That Gave Birth to the Computer Age*, 433-464.

TUIK, (2022). Adrese Dayalı Nüfus Kayıt Sistemi (ADNKS) sonuçları. *Erişim adresi: http://www. tuik. gov. tr*

Twinprai, N., Boonrod, A., Boonrod, A., Chindaprasirt, J., Sirithanaphol, W., Chindaprasirt, P., & Twinprai, P. (2022). Artificial intelligence (AI) vs. human in hip fracture detection. *Heliyon*, *8*(11). https://doi.org/10.1016/j.heliyon.2022.e11266

Upadhyay, R. S., & Tanwar, P. (2019, May). A review on bone fracture detection techniques using image processing. In *2019 International Conference on Intelligent Computing and Control Systems (ICCS)* (pp. 287-292). IEEE. http://dx.doi.org/10.1109/ICCS45141.2019.9065874

Urakawa, T., Tanaka, Y., Goto, S., Matsuzawa, H., Watanabe, K., & Endo, N. (2019). Detecting intertrochanteric hip fractures with orthopedist-level accuracy using a deep convolutional neural network. *Skeletal radiology*, *48*, 239-244. https://doi.org/10.1007/s00256-018-3016-3

Ünal, D. (2008). Tıpta Kullanılan Görüntüleme Teknikleri. TC Gazi Üniversitesi. *Fen Bilimleri Enstitüsü, Doktora Tezi, Ankara*.

Vahvanen, V. (1971). Femoral neck fracture of the rheumatoid hip joint: a study of 20 operatively treated cases. *Acta Rheumatologica Scandinavica*, *17*(1-4), 125-136. https://doi.org/10.3109/rhe1.1971.17.issue-1-4.18

Vasilakakis, M., Iosifidou, V., Fragkaki, P., & Iakovidis, D. (2019, October). Bone fracture identification in x-ray images using fuzzy wavelet features. In *2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE)* (pp. 726-730). IEEE. http://dx.doi.org/10.1109/BIBE.2019.00136

Verma, H. K., & Pal, S. (2016). Modified sigmoid function based gray scale image contrast enhancement using particle swarm optimization. *Journal of The Institution of Engineers (India): Series B*, *97*, 243-251 http://dx.doi.org/10.1007/s40031-014-0175-z

Veronese, N., & Maggi, S. (2018). Epidemiology and social costs of hip fracture. *Injury*, *49*(8), 1458-1460. https://doi.org/10.1016/j.injury.2018.04.015

Wang, X., Xu, Z., Tong, Y., Xia, L., Jie, B., Ding, P., ... & He, Y. (2022). Detection and classification of mandibular fracture on CT scan using deep convolutional neural network. *Clinical Oral Investigations*, *26*(6), 4593-4601. https://doi.org/10.1007/s00784-022-04427-8

Waschke, Paulsen. "Sobotta Atlas of Human Anatomy Latin Nomenclature." (2011).

Wikipedia the free encyclopedia, "Computed Tomography", https://en.wikipedia.org/wiki/CT_scan (2022).

Yadav, D. P., & Rathor, S. (2020, February). Bone fracture detection and classification using deep learning approach. In *2020 International Conference on Power Electronics & IoT Applications in Renewable Energy and its Control (PARC)* (pp. 282-285). IEEE. http://dx.doi.org/10.1109/PARC49193.2020.236611

Yani, M., Budhi Irawan, S. S. M., & Casi Setiningsih, S. M. (2019, May). Application of transfer learning using convolutional neural network method for early detection of terry's nail. In *Journal of Physics: Conference Series* (Vol. 1201, No. 1, p. 012052). IOP Publishing. http://dx.doi.org/10.1088/1742-6596/1201/1/012052

Yenigün, Ç. (2016). *Mechanical design and analysis of a novel fixation device for human bone fractures* (Master's thesis, Middle East Technical University).

Yu, J. S., Yu, S. M., Erdal, B. S., Demirer, M., Gupta, V., Bigelow, M., ... & White, R. D. (2020). Detection and localisation of hip fractures on anteroposterior radiographs with artificial intelligence: proof of concept. *Clinical radiology*, *75*(3), 237-e1. https://doi.org/10.1016/j.crad.2019.10.022

Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J. (2021). Dive into deep learning. *arXiv preprint arXiv:2106.11342*. https://doi.org/10.48550/arXiv.2106.11342

Zeren, M. T. (2020). İnsansız hava araçları ve uydu görüntülerinden elde edilen veri seti ile havaalanlarının tespitinin yapılmasında SSD ve Faster R-CNN algoritmalarının karşılaştırılması (Master's thesis, Fen Bilimleri Enstitüsü).

Zeren, M. T., Arslankaya, S., Altuntaş, Y., Cam, N., Kırelli, Y., & Özdemir, M. H. (2023). Doctors Versus YOLO: Comparison Between YOLO Algorithm, Orthopedic and Traumatology Resident Doctors and General Practitioners on Detection of Proximal Femoral Fractures on X-ray Images with Multi Methods. *International Journal on Artificial Intelligence Tools* https://doi.org/10.1142/S0218213023500562

Zeren, M. T., Aytulun, S. K., & KIRELLİ, Y. (2020). Comparison of SSD and faster R-CNN algorithms to detect the airports with data set which obtained from unmanned aerial vehicles and satellite images. *European Journal of Science and Technology*, (19), 643-658. https://doi.org/10.31590/ejosat.742789

Zhu, Q., Zheng, H., Wang, Y., Cao, Y., & Guo, S. (2020). Study on the evaluation method of sound phase cloud maps based on an improved YOLOv4 algorithm. *Sensors*, *20*(15), 4314. http://dx.doi.org/10.3390/s20154314

Zuckerman JD., (1996) Hip Fracture. *N Engl J Med*. 1996; 334: 1519–25. https://doi.org/10.1056/nejm199606063342307

ÖZGEÇMİŞ

Ad-Soyad : Muhammed Taha ZEREN

ÖĞRENİM DURUMU:

- **Lisans :** 2013, Kocaeli Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği
- **Yüksek lisans :** 2020, Beykent Üniversitesi, Endüstri Mühendisliği Anabilim Dalı, Endüstri Mühendisliği
- **Doktora :** 2023, Sakarya Üniversitesi, Endüstri Mühendisliği Anabilim Dalı, Endüstri Mühendisliği

MESLEKİ DENEYİM VE ÖDÜLLER:

- 2013-2015 yılları arasında Ford Otosan A.Ş.'de Endüstri/Proje Mühendisi olarak çalıştı.
- 2016-2018 yılları arasında TRT World bünyesinde Proje Yöneticisi olarak çalıştı.
- 2018-2021 yılları arasında TAI – Türk Havacılık ve Uzay Sanayii A.Ş. bünyesinde Proje Yöneticisi olarak çalıştı.
- 2021-2023 yılları arasında LC Waikiki Mağazacılılık A.Ş.'de Direktör Yardımcısı olarak çalıştı.
- 2015 yılında Türk Silahlı Kuvvetleri bünyesinde Başarı Belgesi almaya hak kazandı.

TEZDEN TÜRETİLEN ESERLER:

- Zeren, M. T., Arslankaya, S., Altuntaş, Y., Cam, N., Kırelli, Y., & Özdemir, M. H. (2023). Doctors Versus YOLO: Comparison Between YOLO Algorithm, Orthopedic and Traumatology Resident Doctors and General Practitioners on Detection of Proximal Femoral Fractures on X-ray Images with Multi Methods. *International Journal on Artificial Intelligence Tools* https://doi.org/10.1142/S0218213023500562

DİĞER ESERLER:

- Zeren, M. T., Aytulun, S. K., & KIRELLİ, Y. (2020). Comparison of SSD and faster R-CNN algorithms to detect the airports with data set which obtained from unmanned aerial vehicles and satellite images. *European Journal of Science and Technology*, (19), 643-658. https://doi.org/10.31590/ejosat.742789
- Kırelli, Y., Arslankaya, S., & Zeren, M. T. (2020). Detection of credit card fraud in e-commerce using data mining. *Avrupa Bilim ve Teknoloji Dergisi*, (20), 522-529. https://doi.org/10.31590/ejosat.747399
- Özçelik, T. Ö., Selvi, İ. H., Yalçıner, A. Y., & Zeren, M. T. (2022). Industry 4.0, Smart Factories And Effects On Business. *Knowledge Management and Digital Transformation Power*, 169.