



MARMARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ



GÖRÜNTÜ İŞLEME VE DERİN ÖĞRENME YÖNTEMLERİ KULLANILARAK ANKİLOZAN SPONDİLİT ERKEN TEŞHİSİNE YÖNELİK KARAR DESTEK SİSTEMİNİN GELİŞTİRİLMESİ

EMRE ÇANAYAZ

DOKTORA TEZİ

Elektrik-Elektronik Mühendisliği Anabilim Dalı
Elektrik-Elektronik Mühendisliği Programı

DANIŞMAN

Dr. Öğr. Üyesi Zehra Aysun ALTIKARDEŞ

EŞ-DANIŞMAN

Prof. Dr. Alparslan ÜNSAL

İSTANBUL, 2024



MARMARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ



GÖRÜNTÜ İŞLEME VE DERİN ÖĞRENME YÖNTEMLERİ KULLANILARAK ANKİLOZAN SPONDİLİT ERKEN TEŞHİSİNE YÖNELİK KARAR DESTEK SİSTEMİNİN GELİŞTİRİLMESİ

EMRE ÇANAYAZ
(723116004)

DOKTORA TEZİ

Elektrik-Elektronik Mühendisliği Anabilim Dalı
Elektrik-Elektronik Mühendisliği Programı

DANIŞMAN

Dr. Öğr. Üyesi Zehra Aysun ALTIKARDEŞ

EŞ-DANIŞMAN

Prof. Dr. Alparslan ÜNSAL

İSTANBUL, 2024

ÖNSÖZ

Tez çalışmamı hazırlama sürecinde bana her konuda yardımcı olan, her soruma sabırla yanıt veren ve tez çalışmamın eksik yönlerinin ortaya konulup giderilmesi noktasında değerli katkılarda bulunan tez danışmamalarım Dr. Öğr. Üyesi Zehra Aysun ALTIKARDEŞ ve Prof. Dr. Alparslan ÜNSAL'a , çalışmam süresince fedakârlıklarıyla her zaman yanımda olan Prof. Dr. Hayriye KORKMAZ ve Doç. Dr. Mustafa GÖK'e, sonsuz söylenmelerimi büyük bir sabır ile dinleyen ve bana devamlı destek olan sevgili hayat arkadaşım Elif KÜÇÜKBÜCE'ye ve annem Ateş ÇANAYAZ, babam Uğur ÇANAYAZ'a, kayınvalidem ve manevi annem Canan TOKSÖZ'e teşekkürlerimi bir borç bilir, sonsuz minnettarlığımı sunarım. Pro tou, oloi oi theoi tremoun, so do I.

Kasım,2024

Emre ÇANAYAZ

ÖNSÖZ.....	iii
ÖZET.....	vii
ABSTRACT	viii
YENİLİK BEYANI.....	ix
SEMBOLLER.....	xi
KISALTMALAR	xii
ŞEKİL LİSTESİ.....	xiv
TABLO LİSTESİ.....	xv
1. GİRİŞ.....	1
1.1. Amaç ve Gerekçe	3
1.1.1. Materyal ve Yöntem	5
1.2. Çalışmanın Literatürdeki Yeri	6
2. GENEL BİLGİLER	15
2.1. Ankilozan Spondilit (AS)	15
2.1.1. AS Hastalığının Tarihçesi	15
2.1.2. AS Hastalığının Tanımı, Tanısı, İlerleyişi ve Vücutta Oluşturduğu Hasarlar ...	16
2.1.3. AS'nin Dünya ve Türkiye'deki Yeri	18
2.1.4. Türkiye'de Ankilozan Spondilit ile İlişkili Sağlık Bakım Maliyetleri.....	19
2.2. Manyetik Rezonans Görüntüleme (MRI)	20
2.2.1. Short Tau Inversion Recovery (STIR) Sekanslı MRI görüntüleri	21
2.2.2. STIR Sekansı ve AS	22
2.3. Yapay Zekâ (YZ) Tanımı ve Yaklaşımları	22
2.3.1. Makine Öğrenmesi (ML)	23
2.3.1.1. Destek Vektör Makinesi (SVM).....	24
2.3.1.2. k-En Yakın Komşu Algoritması (k-NN).....	27
2.3.1.3. Toplu Öğrenme Algoritması (TOA).....	28
2.3.1.4. Çoğunluk Oylama Algoritması (ÇOA).....	29
2.3.1.5. Yapay Sinir Ağları (YSA).....	30
2.4. Derin Öğrenme (DL)	32

2.4.1. Evrişimsel Sinir Ağları (ESA)	33
2.4.2. Transfer Öğrenme ve ESA İlişkisi	34
2.5. Makine Öğrenmesi ve Derin Öğrenme Arasındaki Farklar.....	35
2.5.1. Eğitim Kümesi Gereksinimi.....	35
2.5.2. Özellik Çıkarımı	36
2.5.3. Model Kompleksliği	36
2.5.4. Açıklanabilirlik.....	37
2.5.4.1. Polonyi Paradoksu ve GRAD-CAM Yöntemi	37
2.5.5. Performans	38
2.5.6. Donanım Gereksinimi.....	38
2.5.7. Hesaplama Maliyeti.....	39
2.5.8. Gerçek Zamanlı Uygulamalarda Kullanılabilirlik.....	39
3. ANKİLOZAN SPONDİLİT TESPİTİNE YÖNELİK YAPAY ZEKÂ TABANLI	
KARAR DESTEK SİSTEMLERİN TASARLANMASI.....	40
3.1. Görüntü Toplama ve Veri Seti	40
3.2. Yöntem 1: Klasik ML Algoritmaları Kullanılarak Geliştirilen AS Tespit Sistemi	
.....	43
3.1.1. Özellik Çıkarımı: Gri-Seviye Eş-Olay Matrisi ve Haralick Doku Özellikleri...45	
3.1.3. Veri Hazırlama ve Model Geliştirme	48
3.3. Yöntem 2: Haralick Özellikleri ile Eğitilen Derin Öğrenme Modeli.....	51
3.3. Yöntem 3: Derin Öğrenme Tabanlı Modeller ile Oluşturulan Oylama Sistemi ...	54
3.3.1. Görüntü Toplama ve Model Geliştirme.....	55
4. DENEYSEL ÇALIŞMALAR VE BULGULAR	59
4.1. Yöntem 1 ile Geliştirilen Modelinin Performans Analizi	59
4.2. Yöntem 2 Geliştirilen DL Modelinin Performans Analizi	61
4.3. DL Tabanlı Modeller ile Oluşturulan Çoğunluk Oylama Algoritmasının	
Performans Analizi.....	63
5. SONUÇLAR	73
5.1. Önerilen Modellerin Yöntem olarak benzeyen Çalışmalar ile Karşılaştırılması ve	
Literatüre Katkıları.....	73
5.2. Önerilen Sistemlerden Hangisi AS Tespiti için Daha Uygunur?.....	78
5.2.1. Genel Performans: Doğruluk ve AUC.....	79
5.2.2. Duyarlılık ve Geri Çağırma.....	79

5.2.4. Hatalı Pozitif ve Negatif Oranları	80
5.2.5. F1 Skoru.....	80
5.3. Grad-CAM Özellik Haritaları: DL Modelleri Nereden Öğrendi?.....	80
5.4. Yöntem 3'ün Erken Teşhis İçin Sınıflandırma Performans Analizi ve 223787 No'lu Denek.....	81
5.5. Yapay Zekâ Tabanlı Görüntü Sınıflandırma yönteminin Tıbbi alana Uygulanmasında Potansiyel Zorluklar	83
5.5.1. YZ Tabanlı Tıbbi Görüntüleme Çözümlerinin Benimsenmesinde Sağlık Profesyonellerinin Rolü ve Güçlendirilmesi	84
5.6. Çalışmanın Kısıtları ve Gelecek Çalışmalar.....	85
KAYNAKLAR.....	88
ÖZGEÇMİŞ	104

ÖZET

GÖRÜNTÜ İŞLEME VE DERİN ÖĞRENME YÖNTEMLERİ KULLANILARAK ANKİLOZAN SPONDİLİT ERKEN TEŞHİSİNE YÖNELİK KARAR DESTEK SİSTEMİNİN GELİŞTİRİLMESİ

Bu tez çalışmasında, ankilozan spondilitin erken teşhisinde kullanılmak üzere açıklanabilirliği yüksek, güç tüketimi düşük ve %90'ın üzerinde doğruluk oranına sahip yapay zekâ tabanlı bir karar destek sistemi geliştirilmiştir. Çalışma, tıp ve mühendislik disiplinlerinin iş birliği ile gerçekleştirilmiş ve STIR MRI görüntüleri üzerine odaklanmıştır. Bu bağlamda, üç farklı yöntemle modeller geliştirilmiştir.

Birinci yöntemde, klasik makine öğrenmesi algoritmalarına dayalı bir model oluşturulmuş, görüntülerden elde edilen gri düzey eş oluşum matrislerinden türetilen Haralick özellikleri kullanılmıştır. İkinci yöntemde, bu özellikler derin öğrenme modelleri ile birleştirilerek hibrit bir yaklaşım sunulmuştur. Üçüncü yöntemde ise, önceden eğitilmiş evrimsel sinir ağları kullanılarak bir topluluk modeli geliştirilmiş ve çoğunluk oylama mekanizması ile sınıflandırma gerçekleştirilmiştir. Bu sistemler hem bireysel hem de topluluk performansı açısından analiz edilmiş ve sınıflandırma performansları değerlendirilmiştir.

Veri seti, Adnan Menderes Üniversitesi Tıp Fakültesi'nden retrospektif olarak elde edilen 696 STIR MRI görüntüsünden oluşmuştur. Bu görüntüler, radyologlar tarafından global AS derecelendirme sistemi kullanılarak etiketlenmiştir. Her bir model, veri setinin özelliklerine göre özelleştirilmiş ve çeşitli performans metrikleri ile değerlendirilmiştir. Modeller, özellikle duyarlılık, özgünlük ve F1 skoru gibi metriklerde başarılı sonuçlar göstermiştir. Özellikle topluluk modeli, klinik uygulamalar için açıklanabilirlik ve güvenilirlik açısından üstün bir performans sunmuştur.

Geliştirilen sistemlerin, tıbbi tanı süreçlerini destekleyerek erken teşhis oranlarını artırması ve sağlık sisteminin yükünü hafifletmesi hedeflenmiştir. Bu tez çalışması hem tıp hem de mühendislik alanlarında literatüre önemli katkılar sağlamakta ve disiplinler arası bir model geliştirme örneği sunmaktadır. Ayrıca, sonuçlar, geliştirilen modellerin klinik ortamlarda uygulanabilir olduğunu göstermektedir. Bu tez, ankilozan spondilit teşhisinde yenilikçi yaklaşımlar sunarak tıp ve mühendislik disiplinlerinin iş birliğine dayalı çözümler üretmesine olanak sağlamaktadır.

ABSTRACT

DEVELOPMENT OF DECISION SUPPORT SYSTEM FOR EARLY DETECTION OF ANKYLOSING SPONDYLITIS USING IMAGE PROCESSING AND DEEP LEARNING METHODS

This thesis focuses on the development of an artificial intelligence-based decision support system for the early diagnosis of ankylosing spondylitis. The proposed system aims to achieve high explainability, low power consumption, and an accuracy rate exceeding 90%. To meet these objectives, three distinct AI-based models were developed.

The first model is based on classical machine learning algorithms, utilizing Haralick features derived from gray-level co-occurrence matrices generated from MRI images. The second model combines these Haralick features with deep learning methods, offering a hybrid approach. The third model leverages pre-trained convolutional neural networks and employs an ensemble learning mechanism with majority voting for classification.

The dataset used in this study consists of 696 short tau inversion recovery magnetic resonance images, retrospectively collected from Adnan Menderes University Medical Faculty. These images were labeled by medical experts using a global AS grading system. Following the preparation of the dataset, the models were tailored to the dataset's specific features and evaluated using metrics such as sensitivity, specificity, and F1 score. Among the three models, the ensemble learning model demonstrated superior performance, providing a balance of accuracy, reliability, and clinical interpretability.

This thesis aims to support medical diagnostic processes, improve early diagnosis rates, and reduce the healthcare system's burden. By combining expertise from the fields of medicine and engineering, the study offers a novel interdisciplinary approach to addressing the challenges of AS diagnosis. The findings highlight the potential of the developed models for clinical implementation, contributing significantly to the advancement of diagnostic technologies.

YENİLİK BEYANI

Ankilozan Spondilit (AS) tanısında yapay zekâ (AI) tabanlı yenilikçi yaklaşımlar geliştirilerek, bu alan için literatürde eksikliği görülen önemli teknik çözümler sunulmuştur. Bu çalışma hem mühendislik hem de tıp alanlarında ortak bir dil oluşturma amacıyla gerçekleştirilmiş, AS tanısı için en uygun yöntemlerin geliştirilmesine odaklanılmıştır. Çalışma kapsamında, Kısa Tau İnversiyon Kurtarma (Short-TI Inversion Recovery-STIR) sekansı kullanılarak elde edilen MRI görüntüleri üzerinden bir veri seti oluşturulmuş olup, tüm analizler bu görüntüler ya da bu görüntülerden türetilmiş özellikler üzerinde yürütülmüştür. Literatürde STIR sekanslı MRI görüntülerini AS erken tanısı için kullanan sadece bir çalışma bulunduğunda, bu tez literatüre önemli bir katkı sağlamaktadır.

AS tanısına yönelik geliştirilmiş olan bu sistemde, görüntülerin gri düzey eş oluşturma matrisleri (GLCM) oluşturulmuş ve bu matrisler üzerinden hesaplanan 21 Haralick özelliği ile derin öğrenme modeli eğitilmiştir. AS tanısı için Haralick özelliklerinin derin öğrenme sistemlerinde sayısal giriş verisi olarak kullanıldığı bir çalışmaya literatür taramasında rastlanmamıştır. Bu yenilikçi yaklaşım, hesaplama gücünden ve zamandan tasarruf sağlamanın yanı sıra daha hızlı ve verimli sonuçlar elde edilmesine olanak tanımaktadır.

Buna ek olarak, AS tanısı için önceden eğitilmiş evrimsel sinir ağları (ESA) bir araya getirilerek topluluk oylama sistemi kullanılması da literatürde yer almayan bir yaklaşımdır. Farklı ESA modellerini birleştiren bu topluluk modeli, bireysel modellerin tek başına sağladığı sınıflandırma doğruluğunu aşarak daha yüksek doğruluk oranına ve genelleme kapasitesine ulaşmıştır. Bu gelişim, AS teşhisinde klinik uygulama açısından oldukça önemli bir potansiyele işaret etmektedir.

Ayrıca, bu tez çalışmasında geliştirilen modelin klinik kabul edilebilirliğini ve açıklanabilirliğini artırmak amacıyla, yüksek enerji tüketimi ve sınırlı enerji kaynaklarının israfına yol açan derin öğrenme gibi karmaşık yapay zekâ teknikleri yerine, tıbbi uzmanlar tarafından kolayca yorumlanabilir bir sistem tasarlanmıştır. Derin öğrenme sistemlerindeki artan katman sayısı ve model karmaşıklığı, enerji verimliliğini düşürmenin yanı sıra, tıbbi açıklanabilirliği zorlaştırarak klinik uygulamalarda güvenilirliği olumsuz etkileyebilmektedir. Bu tezde sunulan yaklaşım, enerji tasarrufu ve açıklanabilirlik açısından mevcut yöntemlere kıyasla yenilikçi bir çözüm sunmaktadır. Tıp ve mühendislik alanları arasında ortak bir dil oluşturmaya hedefleyen bu yöntem, özellikle AS teşhisinde erken tanıya olanak sağlama potansiyeli taşımaktadır.

Sonu olarak, bu alıřmada geliřtirilen AI tabanlı model, AS teřhisinde yksek doęruluk ve hız sunarak klinik srelerin etkinlięini artırmakta; ayrıca, tıp ve mhendislik alanlarındaki uzmanların iř birlięine katkı saęlamaktadır. Tez kapsamında nerilen yeniliki yntemlerin AS'nin erken teřhisine ynelik disiplinler arası zm sunacaęı ve toplum saęlıęına nemli faydalar saęlayacaęı dřnlmektedir.

Kasım, 2024

Dr. ęr. yesi Zehra Aysun ALTIKARDEř

Emre ANAYAZ



SEMBOLLER

β^T	Ağırlık Vektörü
β_o	Uzamsal Doğru Denklemi
Δ	Fark Faktörü
$d(x, y)$	Öklit Uzaklık
o	Yapay Sinir Ağı Çıkışı
w	Sinir Hücresi Ağırlığı
s	Örnek Sayısı
f	Özellik Sayısı
C	Etiket Kategorisi Sayısı
N_g	Benzersiz Gri Seviye Toplam Sayısı
I	Gri Tonlama Görüntüsü
$p(i, j)$	Gri-Seviye Eş-Olay Matrisi
$p_x(i)$	Marjinal Satır Olasılıkları
$p_y(j)$	Marjinal Sütun Olasılıkları
μ	Gri Seviye Yoğunluğu
HX	x'in Entropisi
HY	y'nin Entropisi
HXY	Matris Entropisi

KISALTMALAR

adam	Uyarlanabilir Moment Tahmini
YZ	Yapay Zekâ
YSA	Yapay Sinir Ağları
AS	Ankilozan Spondilit
ASAS	Spondiloartrit Değerlendirme Uluslararası Topluluğu
AUC	Eğri Altında Kalan Alan
autoc	Oto Korelasyon
CM	Karışıklık Matrisleri
ESA	Evrişimsel Sinir Ağları
contr	Kontrast
cprom	Küme Önemi
CPU	Merkezi Hesaplama Ünitesi
cshad	Küme Gölgesi
BT	Bilgisayarlı Tomografi
dissi	Benzersizlik
DICOM	Tıpta Dijital Görüntüleme ve İletişim
DL	Derin Öğrenme
DSA	Derin Sinir Ağları
dvarh	Fark Entropisi
energ	Enerji
entro	Entropi
HN	Hatalı Negatif
HP	Hatalı Pozitif
GLCM	Gri Düzey Eş Oluşum Matrisleri
GPU	Grafik İşlemcileri
homom	Homojenlik Versiyon 1
homop	Homojenlik Versiyon 2
idmnc	Normalize Ters Fark Momenti
indnc	Normalize Ters Fark

inf1h	Korelasyon Bilgi Ölçütü Versiyon 1
inf2h	Korelasyon Bilgi Ölçütü Versiyon 2
k-NN	k-En Yakın Komşu
maxpr	Maksimum Olasılık
ML	Makine Öğrenmesi
MRI	Manyetik Rezonans Görüntüleme
NSAİİ	Non-Steroid Antiinflamatuvar İlaçlar
ResNet	Residual Networks
RO	Rastgele Orman
rmsprop	Kök Ortalama Kare Yayılımı
ROI	İlgi Bölgeleri
savgh	Toplam Ortalama
senth	Toplam Enerji
sgdm	Momentumlu Stokastik Gradyan İnişi
sosvh	Kareler Toplamı
STIR	Kısa Tau İnversiyon Kurtarma
svarsh	Toplam Varyans
SVM	Destek Vektör Makinesi
GN	Gerçek Negatif
TOA	Topluluk Öğrenmesi Algoritması
ÇOA	Çoğunluk Oylama Algoritması
GP	Gerçek Pozitif
TPU	Tensör İşlemcileri

ŞEKİL LİSTESİ

Şekil 1.1	Klinik karar destek sisteminin kısmı mimari yapısı	2
Şekil 1.2	Literatür taramasında takip edilen PRISMA metodolojisi.....	8
Şekil 1.3	Ulusal Tez Merkezi AS araştırma sayıları.....	8
Şekil 2.1	Ankilozan spondilit hastasının sakroiliak MRI bulguları	18
Şekil 2.2	Ankilozan spondilit: Deformitelerin ilerleyişi	18
Şekil 2.3	MRI cihazının temel çalışma prensibinin şematik gösterimi.....	21
Şekil 2.4	$\beta T x + \beta o = 0$ doğrusunun konumu	25
Şekil 2.5	Düzlem, üst ve alt destek doğru denklemleri	26
Şekil 2.6	$k=5$ 'e göre hesaplanmış bir k-NN algoritması uygulama örneği.	28
Şekil 2.7	Yapay bir nöronun ve çok katmanlı bir sinir ağının mimarisi	31
Şekil 2.8	Bir evrişimsel sinir ağının katmanları.....	33
Şekil 2.9	GRAD_CAM yönteminin genel çalışma prensibi.....	38
Şekil 3.1	Veri seti oluşturma adımları iş akış diyagramı.....	44
Şekil 3.2	Veri setindeki katılımcıların yaş bilgilerinin istatistiki olarak dağılımı.....	44
Şekil 3.3	Yöntem 1'in iş akış diyagramı.	44
Şekil 3.4	4x4'lük bir görüntüde, üç gri seviye 1'den 3'e kadar sayısal değerlerle gösterilir..	46
Şekil 3.5	Orijinal görüntüden ROI'lerin seçilmesi ve GLCM'nin $\theta=0^\circ$ and $d=2$ parametreleri için hesaplanması.	47
Şekil 3.6	Yöntem 2'in iş akış diyagramı	53
Şekil 3.7	Yöntem 2'de kullanılan DL mimarisi	53
Şekil 3.8	Yöntem 3'ün iş akış diyagramı.....	57
Şekil 5.1	Önceden eğitilmiş ESA tabanlı DL modellerinin doğruluğunun, hesaplama süresinin ve sınıflandırma süresinin karşılaştırılması.....	77
Şekil 5.2	Tez çalışması kapsamında, üç farklı yapay zeka yaklaşımıyla eğitilen modellerin sınıflandırma Performansları karşılaştırılması.....	78
Şekil 5.3	Tez çalışması kapsamında, üç farklı yapay zekâ yaklaşımıyla eğitilen modellerin ROC analizlerinin karşılaştırılması.....	57
Şekil 5.4	IncV3_5 modeli ile modeli ile doğru bir şekilde AS sınıflandırılması yapılan görüntülerin Grad-CAM yöntemi ile çıkarılmış özellik haritaları.	81
Şekil 5.5	223787 No'lu deneğe ait bir görüntüsünün ROI'si ve görüntünün Grad-CAM analizi.....	57

TABLO LİSTESİ

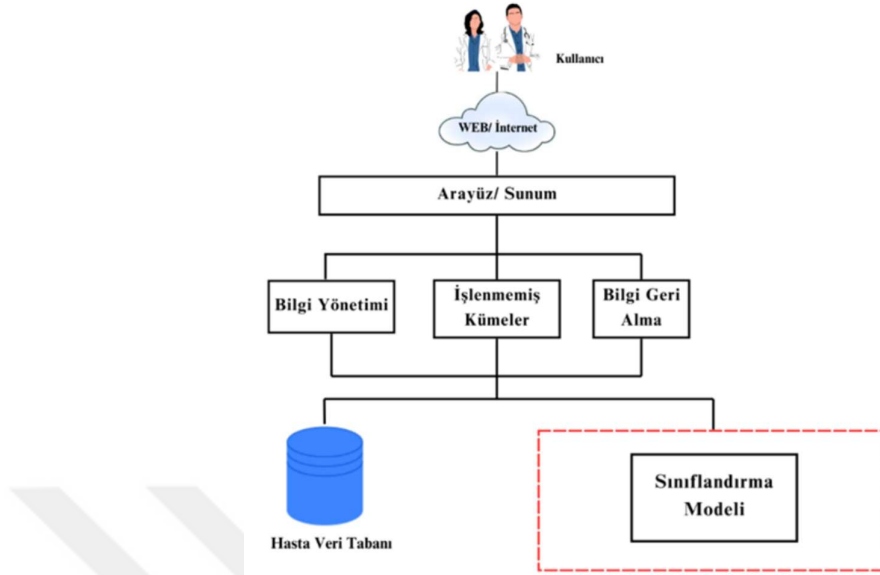
Tablo 1.1	Literatür tarama özeti.....	9
Tablo 2.1	Makine öğrenmesi ve Derin Öğrenme arasındaki farkların sistem özellikleri ve parametrelerine göre karşılaştırılması.....	35
Tablo 3.1	AS değerlendirmesi için incelenen görüntülerin global AS dereceleri.....	41
Tablo 3.2	GLCM'den türetilen Haralick yüzey özellikleri.. ..	49
Tablo 3.3	Haralick özellik tabanlı DL modelinin katman analizi.	54
Tablo 4.1	AS tespiti için Yöntem 1 ile geliştirilen modellerin geçerlilik denetim veri seti ve test seti üzerindeki performans analizleri.....	60
Tablo 4.2	Doğrulama veri setinde %85'in üzerinde Ankilozan Spondilit tespit doğruluğuna sahip modellerin eğitim hiper parametreleri.....	61
Tablo 4.3	AS tespiti için Yöntem 2 ile geliştirilen modellerin hiper parametreleri ve performans analizleri	62
Tablo 4.4	Resnet-18 temsilci seçimi için farklı hiperparametrelere sahip ön öğrenmeli ESA modelinin başarı Performansları	64
Tablo 4.5	Resnet-50 temsilci seçimi için farklı hiperparametrelere sahip ön öğrenmeli ESA modelinin başarı Performansları	65
Tablo 4.6	Resnet-101 temsilci seçimi için farklı hiperparametrelere sahip ön öğrenmeli ESA modelinin başarı Performansları	67
Tablo 4.7	InceptionV3 temsilci seçimi için farklı hiperparametrelere sahip ön öğrenmeli ESA modelinin başarı Performansları	68
Tablo 4.8	ShuffleNet temsilci seçimi için farklı hiperparametrelere sahip ön öğrenmeli ESA modelinin başarı Performansları	69
Tablo 4.9	MobileNet temsilci seçimi için farklı hiperparametrelere sahip ön öğrenmeli ESA modelinin başarı Performansları	69
Tablo 4.10	EfficientNet temsilci seçimi için farklı hiperparametrelere sahip ön öğrenmeli ESA modelinin başarı Performansları	71
Tablo 4.11	Yöntem 3 ile oluşturulan ÇOA modeli ve temsilci sınıflandırma performansları.	72
Tablo 4.13	Yöntem 3 ile oluşturulan ÇOA sistemi için oylama örnekleri.....	72
Tablo 5.1	Önerilen modellerin, yakın metotlara sahip çalışmalarla karşılaştırmalı avantaj ve dezavantaj analizi.	76
Tablo 5.2	Önerilen klasik ML tabanlı modelin tespit sonucu performansının en benzer araştırmayla karşılaştırılması.....	72
Tablo 5.3	Önerilen ML tabalı modelin tespit sonucu performansının en benzer araştırmayla karşılaştırılması.....	72

1. GİRİŞ

Ankilozan spondilit (AS), omurga ve eklemleri etkileyen kronik ve zamanla ilerleyen bir hastalıktır. Erken teşhis hastalığın seyrini belirlemek ve uygun tedaviyi başlatmak açısından büyük önem taşımaktadır. Ancak hastalığın erken dönem belirtilerinin genellikle diğer rahatsızlıklarla karıştırılması, tanıda gecikmelere yol açmakta ve bu durum hastaların tedavi şansını olumsuz etkilemektedir [1]. Gecikmiş teşhis AS hastalarının ağrılarının artmasına, yaşam kalitelerinin düşmesine ve omurga şekil bozukluklarının daha belirgin hale gelmesine neden olabilir [2]. Ayrıca, tedavi sürecinin karmaşıklaşması ve maliyetin artması, sağlık çalışanları ile sistemi iş yükünü de artırmaktadır. Araştırmalar AS'nin erken teşhis edilmesi durumunda, hastalığın daha iyi yönetilebileceğini ve sağlık hizmetlerinin maliyet etkinliğinin artabileceğini göstermektedir [3]. Bu nedenle AS'nin erken teşhisine yönelik yeni yaklaşımların geliştirilmesi gereklidir.

Bu tez çalışmasında, tıp ve mühendislik alanlarını bir araya getiren disiplinler arası bir yaklaşım benimsenmiştir. Radyoloji ile mühendislik bilimlerini birleştirerek ortak bir dil oluşturma çabası, bu çalışmanın en büyük yeniliklerinden biridir. Özellikle, Kısa Tau İnversiyon Kurtarma (Short Tau Inversion Recovery-STIR) echo sekanslı manyetik rezonans (MRI) görüntüleri kullanılarak yapılan bu çalışmada, disiplinler arasında radyoloji uzmanlarının ve mühendislerin hızlı, doğru ve verimli çalışmaları için ortak dilin geliştirilmesine katkı sağlaması amaçlanmaktadır.

Bu çalışma kapsamında klasik bir karar destek mimarisinin sınıflandırma işleminin gerçekleştiği ve karar verme sürecinin yönetildiği AS sınıflandırma modeli geliştirilmesi hedeflenmiştir (Şekil 1.1). İlgili probleme bir mühendislik çözümü geliştirmek amacıyla yapay zekâ (YZ) tekniklerinden geniş bir şekilde yararlanılmıştır. Bu çalışmada, klasik makine öğrenmesi (ML) yöntemlerinin yanı sıra, derin öğrenme (DL) modellerinin gücünden faydalanılmış ve bu iki yaklaşımın bir arada kullanıldığı hibrid bir sistem ortaya konmuştur. Geliştirilen tüm sistemler, hastalık belirtilerinin daha net ve doğru bir şekilde analiz edilmesini sağlamak için, uzmanlar eşliğinde çıkarılan sayısal görüntü özellikleri ile eğitilmiştir. Böylece hem mühendislik hem de tıp alanındaki uzmanların bilgi birikimi bir araya getirilerek, AS teşhisinin daha etkin bir şekilde yapılmasına yönelik özgün bir yöntem geliştirilmiştir.



Şekil 1.1: Klinik karar destek sisteminin kısmi mimari yapısı [4]

Bu çalışmada kullanılan veri seti, STIR echo MRI görüntüleme metodu görüntülerinden oluşmaktadır. Bu görüntü tipi, AS'nin erken teşhisi için altın standart olarak kabul edilmekte olup, hastalığın ilk evrelerinde ortaya çıkan karakteristik değişiklikleri belirlemede önemli avantajlar sunmaktadır. Bu doğrultuda ihtiyaç duyulan veriler, radyologlarla iş birliği yapılarak oluşturulmuş ve derlenmiştir. Tez çalışması, Adnan Menderes Üniversitesi Tıp Fakültesi Dahili Tıp Bilimleri Bölümü Radyoloji Anabilim Dalı ile Marmara Üniversitesi Fen Bilimleri Enstitüsü Teknoloji Fakültesi Elektrik ve Elektronik Mühendisliği Ana Bilim Dalı tarafından ortaklaşa gerçekleştirilmiştir. Disiplinler arası bir yaklaşım sergileyen bu çalışma, retrospektif ve gözlemsel bir tasarıma sahiptir. Çalışma kapsamında tedavi veya girişim üzerine herhangi bir müdahale planlanmamıştır.

Çalışmanın yürütülebilmesi için Adnan Menderes Üniversitesi Tıp Fakültesi Dekanlığı Girişimsel Olmayan Klinik Araştırmalar Etik Kurulu'ndan, 08.02.2021 tarihinde 2021/24 Protokol Numarası ile onay alınmış ve verilerin toplanmasına, ardından radyologlar tarafından etiketlenmesine başlanmıştır.

Yapılan kapsamlı literatür taramasında, MRI görüntüleme metodu kullanılarak gerçekleştirilen çeşitli çalışmalar bulunmakla birlikte, STIR sekanslarının kullanımına odaklanan bu çalışmanın literatürde özgün olduğu tespit edilmiştir. Bu bağlamda, çalışma,

AS'nin erken tanısına yönelik ilk veri sağlayan arařtırmalardan biri olarak bu alandaki bilgi boşluęunu doldurma potansiyeline sahiptir.

Geliřtirilen modeller, özellikle erken teřhis potansiyeline sahip bu tez kapsamında, klinik rutinelere entegre edilebilecek nitelikte bir yapı sunmayı hedeflemiřtir. Karmařık ve açıklanması güç YZ teknikleri yerine, süreçlerin ve sonuçlarının tıbbi uzmanlar tarafından daha kolay anlaşılabilir olduęu bir model tercih edilmiřtir. Böylelikle, geliřtirilen modellerin klinik kabul edilebilirlięi ve uygulanabilirlięi artırılmaya çalışılmıřtır.

Çalışmada, geliřtirilen YZ tabanlı modellerin istatistiksel başarı oranlarının %90'ın üzerine çıkarılması hedeflenmiř, bu doęrultuda geliřtirilen modellerin performans analizi ayrıntılı bir řekilde yapılmıřtır. Özellikle, modelin doęru bir řekilde eęitilmesi ve test edilmesi için veri sızıntısını engelleyecek yöntemler kullanılmıř, aşırı öğrenmenin önüne geçmek amacıyla çeřitli önlemler alınmıřtır. Bu sayede geliřtirilen modelin genelleme yeteneęinin korunması ve gerçek dünya verilerine uyarlanabilirlięinin sağlanması amaçlanmıřtır.

Tezin içerięi, okuyucuya konuyu geniş bir perspektifle sunmak üzere oluřturulmuřtur. İkinci bölümde, PRISMA metodolojisi kullanılarak AS ve YZ ile ilgili literatürdeki çalışmalar derlenmiřtir. Üçüncü bölümün ilk kısmında, AS'nin temel özellikleri, hastalıęın seyri ve Türkiye'deki mali yükü konuları ele alınmıřtır. Bu kısımda, disiplinler arası ortak dilin geliřtirilmesine katkıda bulunmak hedeflenmiřtir. Üçüncü bölümün devamında, YZ, makine öğrenmesi ve derin öğrenme kavramları açıklanmıř, bu kavramların arasındaki farklara değinilmiřtir. Ayrıca, çalışmada kullanılan YZ algoritmaları teorik olarak detaylandırılmıřtır. Dördüncü bölümde, veri toplama ve veri önileme işlemleri detaylı bir řekilde anlatılmıřtır. Buna ek olarak geliřtirilen modellerin eęitim aşamaları ve çalışma prensipleri sunulmuř, beřinci bölümde ise bu modellerin performanslarının istatistiksel analizi yapılmıřtır. Son olarak, sonuçların değeriendirilmesi ve önerilen modellerin klinik uygulamaya geçiři esnasında karşılaşılabilecek zorluklar ile çözüm önerilerine dair bir tartışma yer almaktadır.

1.1. Amaç ve Gerekçe

AS, inflamatuvar hastalık grubuna dahil olup, omurga eklemleri ve çevresindeki yapılarda belirgin inflamasyon ile karakterizedir. Hastalık omurgada ilerleyici ve yukarı doęru (asendan) kemik füzyonuna yol açmakta olup, etiyolojisi henüz tam olarak aydınlatılamamıřtır. İnflamasyon (enfamasyon, yangı, veya iltihaplanma olarak da bilinir),

canlı dokunun her türlü canlı, cansız yabancı etkene veya içsel/dışsal doku hasarına verdiği sellüler (hücresel), humoral (sıvısal) ve vasküler (damarsal) bir seri vital fizyolojik kökenli bir yanıttır [5]. Kalça ve omuz eklemleri olguların 1/3'ünde tutulur [6].

Erken teşhis hastalığın seyrini belirlemek ve uygun tedaviyi başlatmak için kritik öneme sahiptir. Hastalığın ilk zamanına ait belirtilerin genelliğinden dolayı, sorunun tespiti için yapılan çeşitli çalışmalar teşhisin gecikmesine neden olup, hastanın tamamen iyileşme şansını kaybetmesine neden olabilmektedir [1]. Teşhis gecikmesi, AS hastalarının çektiği acıyı artırabilir ve yaşam kalitelerini olumsuz yönde etkileyebilir [2]. Ayrıca, teşhis gecikmesi hastaların işlevselliklerini kaybetme riskini artırabilir ve omurga deformitelerinin daha belirgin hale gelmesine yol açabilir [2]. Geç teşhis edilen hastalarda, tedavi daha karmaşık hale gelebilir ve maliyeti artırabilir. Bu durumun gerek sağlık çalışanları gerekse hastane iş yükü üzerine olumsuz etkisi olduğu aşikârdır. Özellikle görüntüleme cihazlarından gelen görüntülerin raporlanması ve yorumlanması tamamen öznel ve radyologların iş yükünün büyük bir kısmını oluşturmaktadır. Yakın tarihli bir araştırmaya göre, bu prosedürdeki ana sorun, ağır iş yükleri nedeniyle tıp personelinin tükenmişlik yaşayarak yanlış tanılama yapabilmesi ve personelin üretkenliğinin negatif olarak etkilenmesine yol açabilmesidir [7]. AS'nin erken teşhisi, hastalığın ilerlemesini yavaşlatabilir, sakroiliak eklem hasarının ilerlemesini engelleyebilir ve hastaların yaşam kalitesini artırabilir [2].

Araştırmalar, AS hastalarında tedavinin erken başlanması, hastalığın seyrini olumlu yönde etkileyebileceğini göstermektedir [3]. Hastalığın erken teşhisi aynı zamanda sağlık hizmetlerinde maliyet etkinliğini artırabilir. Ancak, erken teşhis edilen hastalara erken müdahalelerle daha iyi sonuçlar elde edebilir ve sağlık sisteminin kaynaklarını daha etkin bir şekilde kullanabilir [8]. AS'nin erken teşhisi ya da hangi bireylerin ilerde AS'ye yakalanma ihtimalinin daha yüksek olduğunun tespit edilmesi için yapılan çeşitli çalışmalardan literatür taraması kısmında bahsedilmiştir. Literatürde görüntü işleme temelli çalışmalarda Sakroiliak eklemlerdeki subkondral kemik ödemi ve subkondral osteit, AS hastalarında erken teşhis açısından önemli bir bulgu olduğu kabul edilmiştir [9]. Bu bulgular, hastalığın erken evrelerinde sıklıkla görülür ve AS'nin diğer belirtileriyle birlikte değerlendirildiğinde, hastalığın teşhisi için kritik önem taşımaktadır [10]. Subkondral kemik ödemi ve osteit, manyetik rezonans görüntüleme (MRI) teknikleriyle tespit edilebilir. MRI, yumuşak doku kontrastını yüksek çözünürlükte sunarak bu lezyonların erken dönemde saptanmasına olanak tanır. Özellikle STIR (Short Tau Inversion Recovery) ve T2 ağırlıklı

sekanslar, kemik iliği ödemi belirlemede etkilidir. Alternatif olarak, bilgisayarlı tomografi (BT) ve sintigrafi gibi görüntüleme yöntemleri de kullanılabilir. Ancak, BT'nin yumuşak doku kontrastı MRI kadar yüksek değildir ve sintigrafi spesifik olmayan bulgular verebilir. Bu nedenle, subkondral kemik ödemi ve osteitin erken tanısında MRI, diğer yöntemlere göre daha üstün kabul edilmektedir [2, 11]. Spondiloartrit Değerlendirme Uluslararası Topluluğu (ASAS) kriteri, yarı koronal STIR sekanslarında sakroiliitisi güçlü bir şekilde düşündüren belirgin subkondral BME varlığını “pozitif MRI” olarak tanımlar [12].

Sonuç olarak, bilişim teknolojileri açısından tıbbi olarak AS'nin erken teşhisi, hastalığın ilerlemesini yavaşlatmak, komplikasyonları önlemek, yaşam kalitesini artırmak ve sağlık hizmetlerinde maliyet etkinliğini artırmak için STIR echo sekanslı MRI görüntüleri kullanılarak YZ yöntemlerinden olan ML ve transfer öğrenme ile ön eğitilmiş DL modelleri kullanılması hedeflenmiştir. Böylelikle tıp doktorunun teşhisine en yakın sınıflandırma performansına sahip olan modellerin ortaya konması ve bu modellerin doğruluklarının test edilmesi amaçlanmıştır.

1.1.1. Materyal ve Yöntem

Adnan Menderes Üniversitesi Tıp Fakültesi Dahili Tıp Bilimleri Bölümü Radyoloji Anabilim Dalı polikliniklerinde, 2015-2020 yılları arasında AS teşhisi konulan 87 hasta ve AS teşhisi bulunmayan 87 bireyden oluşan toplam 174 deneye ait STIR echo sekansı ile elde edilen MRI görüntüleri, cinsiyet ve yaş ayrımı yapılmaksızın seçilmiştir. AS teşhisi konulan deneklere ait tıbbi görüntüler, biri 22 yıllık, diğeri 13 yıllık deneyime sahip iki radyolog tarafından, sağ ve sol eklemler ayrı ayrı değerlendirilerek AS inflamasyon dereceleri belirlenmiştir. Puanlama işleminin ardından deneklerin kişisel bilgileri kimliksizleştirilmiştir.

AS'nin erken teşhis ve sınıflandırılması için subkondral kemik ödemi ve subkondral osteite odaklanmak büyük önem taşımaktadır. Bu doğrultuda, erken teşhis aşamasında AS tanısında kullanılmak üzere, orijinal MRI taramalarından, AS ile ilişkili patolojik değişikliklerin yüksek olasılıkla bulunduğu bölgeler seçilerek spesifik ilgi bölgeleri (ROI) belirlenmiş ve veri seti oluşturulmuştur.

Tezin amacını gerçekleştirmek için üç farklı yöntem uygulanmış ve sonuçları karşılaştırılmıştır:

- Birinci yöntemde, gri düzey eş oluşum matrisleri (GLCM) oluşturulmuş ve bu matrislerden elde edilen 21 farklı Haralick özelliği hesaplanmıştır. Daha sonra, bu özellikler kullanılarak farklı klasik makine öğrenmesi modelleri geliştirilmiştir.
- İkinci yöntemde, GLCM'ler kullanılarak hesaplanan 21 Haralick özelliği, derin öğrenme sisteminin eğitimi için kullanılmıştır. Bu yaklaşımda, derin öğrenme sistemi görsel veri yerine sayısal verilerle eğitilerek AS sınıflandırması yapılması hedeflenmiştir.
- Üçüncü yöntemde, ön eğitilmiş farklı evrimsel sinir ağlarının (ESA) eğitilmesi temel alınmıştır. Modelin eğitilmesi için STIR sekansı ile alınan MRI görüntüleri kullanılmıştır. Farklı hiper parametrelerle eğitilen ESA modelleri arasından en iyi sınıflandırma performansına sahip olanlar seçilmiş ve bu modellerin sağladığı sınıflandırma sonuçlarına dayalı bir çoğunluk oylama sistemi kullanılarak sınıflandırma işlemi gerçekleştirilmiştir.

Geliştirilen modellerin sınıflandırma performansları karışıklık matrisleri (CM) oluşturularak duyarlılık, özgünlük, kesinlik, geri çağırma, F1 Skoru hesaplanması ve alıcı işletim karakteristik eğrisi (ROC) çizilerek altında kalan alanın (AUC) hesaplanması ile gerçekleştirilmiştir.

1.2. Çalışmanın Literatürdeki Yeri

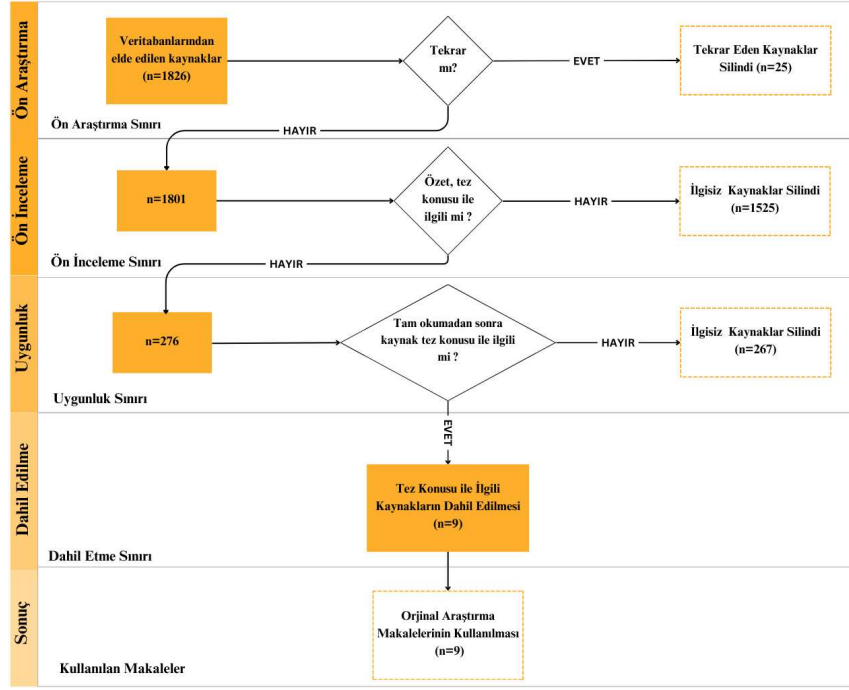
Bu literatür taraması, 10.07.2024 tarihini kapsayacak şekilde, PRISMA metodolojisine [13] uygun olarak gerçekleştirilmiştir. Bu sistematik literatür taraması için kaynakların belirlenmesinde, the National Library of Medicine, the Cochrane Library, Scopus, Web of Science, IEEE Explore, ELSEVIER, Springer, Science Direct (Şekil 1.2) ve YÖK Ulusal Tez Merkezi (Şekil 1.3) gibi çeşitli büyük atıf veri tabanları kullanılmıştır. Öncelikle “*Yapay Zekâ*”, “*Makine Öğrenmesi*”, “*Derin Öğrenme*”, “*Romatoloji ve Görüntü İşleme*” terimleri, farklı kombinasyonlarla AND komutlarıyla birleştirilmiştir. Sunulan incelemede kullanılan kaynaklar, en güncel yayın tarihine sahip olmaları için, 2019 yılından daha eski olmayan orijinal araştırma makalelerini içermektedir.

Şekil 1.2’de sunulan çalışma için PRISMA metodolojisinin nasıl kullanıldığını gösteren bir akış şeması sunmaktadır. Burada erişilen 1826 yayının analiz ve filtreleme adımları detaylı olarak gösterilmiştir. Buna ek olarak Şekil 1.2’de Ulusal Tez Merkezinde AS ile ilgili olan çalışmalar yıllara göre sıralanmıştır. İnceleme sonucunda filtrelenen yayınlar Tablo 1.1’de, özetlenmiştir. İncelemenin uygunluk kriterleri şu şekilde belirlenmiştir:

- Arařtırma, İngilizce yazılmıř olmalı ve 2004 yılından sonra yayımlanmıř olmalıdır.
- AS alanında makine öğrenmesi algoritmaları ve/veya derin öğrenme metodolojilerinin uygulanmıř olması gerekmektedir.
- Arařtırma, anatomik yapıları tespit etmek ve/veya AS hastalığını hedeflemek amacıyla tıbbi görüntüleri kullanmalıdır.
- Arařtırma, literatür taramasının birincil amacına katkıda bulunmalı ve ilgili çalışma alanında etkili bir içeriğe sahip olmalıdır.
- Arařtırma, ilgili otoriteler tarafından hakem deęerlendirmesinden gemiř olmalıdır.

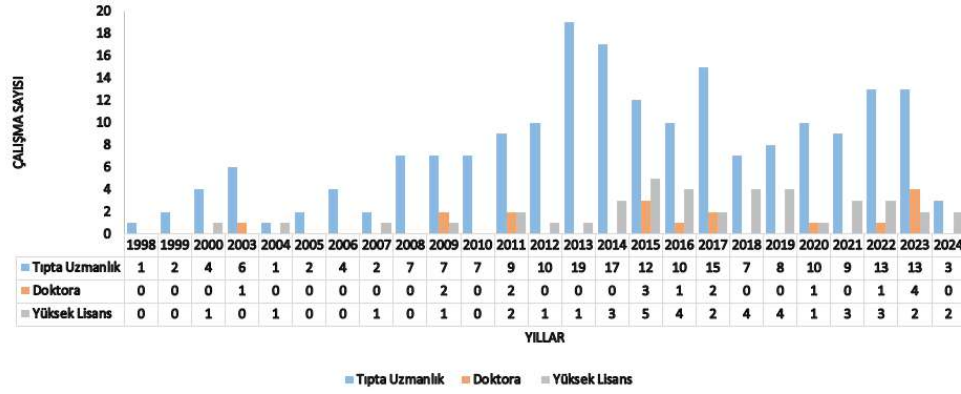
Dıřlanma nedenleri ise řu řekilde belirlenmiřtir:

- Tam metin elde edilemeyen alıřmalar.
- Medikal görüntüleme cihazlarından elde edilen görüntülerin, sınıflandırma için kullanılmadıęı arařtırmalar.
- AS sınıflandırma alıřması olmaması.
- Arařtırmanın yalnızca özet olarak sunulmuř olması.
- Arařtırmanın ana odak noktasının, incelemenin amaçlarına uygun olmaması.
- Yalnızca teorik kavramları ele alan alıřmalar.
- Omurgaya ilerlemiş olan ileri safha AS görüntülerini kullanan alıřmalar.
- Uygulanan yöntem hakkında detaylı bilgi saęlamayan arařtırmalar.



Şekil 1.2: Literatür taramasında takip edilen PRISMA metodolojisi

ANKYLOSING SPONDYLITIS ANAHTAR KELİMELİ ARAMA SONUÇLARI



Şekil 1.3: Ulusal Tez Merkezi AS araştırma sayıları

Tablo 1.1 Literatür tarama özeti

Yazar/ Yıl	Amaç	Veri	YZ Metodu	Sonuçlar
Shenkman ve Ark., 2019 [14]	Pelvis BT taramalarından sakroiliit tanı ve derecelendirmesini otomatik olarak yapacak bir sistem geliştirmek, erken tanı koymayı sağlamak (SIJ-grade).	Sheba Medical Center'dan 242 anonimleştirilmiş aksiyel BT taramaları	Rastgele Orman (RO) ROI tespiti için U-Net sınıflandırıcısı, ve ESA sınıflandırıcısı	- Önerilen algoritma iki aşamadan oluşmaktadır: Çevrimdışı eğitim aşaması ve çevrimiçi sınıflandırma aşaması. - Sırasıyla %91.9 ve %86 sınıflandırma doğruluğu, %95 ve %82 duyarlılık ve 0.97 ile 0.57 AUC değerleri elde edilmiştir. - Araştırmacılar, önerilen sistemin sakroiliiti derecelendirme ve tanı koyma potansiyeline sahip olduğunu iddia etmişlerdir.
Zunti ve Ark., 2020 [15]	Bilgisayarlı tomografi (BT) verilerini analiz ederek, erken dönem Ankilozan Spondilit belirtisi olan erozyonu tahmin etmek için istatistiksel makine öğrenimi (ML) ve derin öğrenme (DL) tabanlı bir sınıflandırıcı önerilmiştir.	681 gri tonlama JPG görüntüsü, her biri tek bir sakroiliak eklemi içermektedir.	ML (k-NN, RO) ve InceptionV3 ESA tabanlı DL modeline giriş özellikleri oluşturmak için GLCM ve yerel ikili desenler (LBP) kullanılmıştır.	- Random forest (RO) sınıflandırıcıları, 8 katlı çapraz doğrulama sonrasında k-NN sınıflandırıcılarını geçmekte olup, erozyon ile genç kontrol hastaları arasında ortalama doğruluk, geri çağırma ve alıcı işletim karakteristik eğrisi (ROC AUC) değerleri sırasıyla %96.0, %92.9 ve 0.97, erozyon ile genç kontrol hastaları için ise %82.4, %80.6 ve 0.91 olarak hesaplanmıştır. - Doğrulama kaybını sınırlamadan eğitilen bir DL sınıflandırıcısı, erozyon açısından tüm (karışık genç ve yaşlı) kontrol hastalarını geçerek, çapraz doğrulama doğruluğu, geri çağırma ve ROC AUC değerleri sırasıyla %99.0, %97.5 ve 0.97 ile üstünlük sağlamıştır.

Tablo 1.1 Literatür tarama özeti (devamı)

Yazar/ Yıl	Amaç	Veri	YZ Metodu	Sonuçlar
Faleiros ve Ark. , 2020 [16]	MRI aktif inflamatuvar sakroiliiti sınıflandırmak amacıyla, araştırmacılar geleneksel makine öğrenimi modellerinin ve özellik seçimi yöntemlerinin uygulanabilirliğini test etmişlerdir.	56 sakroiliak eklem MRI görüntüsü.	Support VeBTor Machine (SVM), the Multilayer Perceptron (MLP), ve the Instance-Based algoritmaları	1. ReliefF and Wrapper özellik çıkarma algoritmaları önemli özellikleri çıkarmak için kullanılmıştır. 2. Segmentasyon ve ROI seçim süreçleri Adobe Photoshop'ta gerçekleştirilmiştir. Eğitim veri setini kullanarak 10 katlı çapraz doğrulama esas alınarak, Wrapper özellik seçim yaklaşımı ile seçilen altı özellik içeren MLP sınıflayıcısı, %100 duyarlılık, %95,6 özgüllük ve %84,7 doğruluk ile en iyi Performansı göstermiştir.
Han ve Ark., 2021 [17]	Kalça eklemi MRI görüntülerinde BME'yi tespit etmek için derin öğrenmeye dayalı yeni bir segmentasyon ağı oluşturulmuştur.	141 vaka (101 eğitim için ve 40 doğrulama için).	ESA and Resnet-50	1. 40 test vakasından 31'i doğru bir şekilde sınıflandırılmıştır. 2. Önerilen sistemin doğruluk oranı %85,7'dir. MRI görüntülerini analiz eden otomatik bilgisayar tabanlı sistemler, AS'nin derecelendirilmesi ve erken tespiti için faydalı olup potansiyele sahiptir.
Li ve Ark. ,2023 [11]	Çalışmanın amacı, insan uzmanlarını geride bırakan çok merkezli araştırmalara dayanan AS teşhisi için kapsamlı bir YZ destekli araç geliştirmektir.	356 hastanın, 5389 dengesiz pelvik radyografi görüntüsü.	ResNet50, MobileNetV3 DenseNet201 ResNet101 MobileNetV3_large MobileNetV3_small DarkNet53 Xception	1. DL modeli hem iç hem de dış doğrulama setlerinde güçlü bir performans sergilemekte olup, %0,91 doğruluk, %0,90 hassasiyet, %0,89 geri çağırma ve %0,96 AUC değerleri elde edilmiştir. 2. DL modelinin tanısallık performansı, iki insan uzmanınınkini aşmış ve model tanımları ile insan uzman değerlendirmeleri birleştirildiğinde geri çağırma önemli bir iyileşme sağlanmıştır.

Tablo 1.1 Literatür tarama özeti (devamı)

Yazar/ Yıl	Amaç	Veri	YZ Metodu	Sonuçlar
Tas ve Ark., 2023 [18]	MRI görüntülerinden AS tanısı için doğru bir AI çerçevesi geliştirilerek, AS'nin erken tespitinin iyileştirilmesi ve hastaların yaşam kalitesinin artırılması hedeflenmiştir.	527 katılımcının aksiyel MRI görüntüleri, 821 katılımcının kontrastlı MRI görüntüleri ve 668 katılımcının koronal görüntüleri.	DenseNet201, ResNet50, ShuffleNet sonrasında farklı özellik seçim algoritmalarından elde edilen özellikler ML sınıflandırma için kullanılmıştır	1. Farklı manyetik rezonans görüntü veri setleri için doğruluk sonuçları %95,62 ile %100 arasında değişiklik göstermiştir. 2. NCA özellik seçici ve kNN sınıflayıcı ile birlikte DenseNet201, üstün Performans sergilemiştir
Lin ve Ark., 2023 [19]	Sakroiliit tespiti için DL modeli geliştirmek	10 adet hastaneden alınan 389 katılımcıya ait 3944 STIR MRI görüntüsü	Bölütleme için Attention Unet ve sınıflandırma için ESA	1. Sınıflandırma 0,86 hassasiyet, 0,92 özgünlük ile gerçekleştirilmiştir. 2. ROC hesaplanıp AUC 0,92 olarak hesaplanmıştır.
Zhang ve Ark.,2023 [20]	AS ile ilişkili sakroiliit tanısı ve derecelendirmesini BT görüntüleri kullanarak tamamen otomatik bir yöntem geliştirmek	İki farklı hastaneden AS tanısı almış hastalar ve kontrol gruplarına ait 435 BT incelemesi	Bölütleme için nnU-Net ve sınıflandırma için ESA	1. AS grade 0,1 ve 2 için hesaplanan AUC'ler sırası ile: 0,91, 0,80, 0,96'dır. 2. ESA, grade 1 görüntülerinde genç ve deneyimli radyologlardan daha üstün performans sergilemiştir. 3. Uzman radyologlar sistemin sonuçlarından daha iyi sınıflandırma yapmışlardır.
Prasad ve Ark.,2024 [21]	MRI taramalarından AS'nin doğru ve yorumlanabilir tanısı için derin öğrenme ve Açıklanabilir YZ tekniklerini entegre eden modeli	MRI görüntüleri, toplam sayı belirtilmemiş	ESA (Resnet, DenseNet) DL ön öğrenimli Modelleri	1. DenseNet önceden eğitilmiş DL modeli, Ankilozan Spondilit tespitinde %91.2 doğruluk, %90.5 hassasiyet ve %90.6 F1-Skoru ile en yüksek Performansı göstermiştir. 2. Grad-CAM metodu kullanılarak DL model sonuçlarının daha açıklanabilir hale getirilmesi hedeflenmiştir.

Shenkman ve Ark.'nın [14]yaptığı çalışmada, manyetik rezonans (MR) görüntüleri yerine 242 adet bilgisayarlı tomografi (BT) görüntüsü kullanılmıştır. Çalışmada, sakroiliit oluşumu görülen hastalarda teşhis koymak ve sakroiliitin sınıflandırılmasını sağlamak amaçlanmıştır. Bu hedef doğrultusunda, gözetimli makine öğrenimi (ML) yöntemleri ve derin öğrenme (DL) metotları uygulanmıştır. Geliştirilen algoritma, öncelikle ilgi alanlarını (ROI) U-Net algoritmaları ile otomatik olarak belirlemiş, ardından rastgele orman (RO) algoritması kullanarak sakroiliit oluşumunu tespit etmiştir. Daha sonra, evrişimsel sinir ağı (ESA) sınıflandırıcı algoritması ile geliştirilmiş New York kriterlerine [19] göre sakroiliitin ilerleme düzeyi belirlenerek uygun sınıfa atanması hedeflenmiştir. ROI'nin otomatik olarak tespit edilmesi, çalışmada %82 doğruluk oranıyla başarılı bir şekilde gerçekleştirilmiştir. Bunun ardından, DL tabanlı farklı sınıflandırma modelleri kullanılarak %80,6 ile %87,1 arasında değişen doğruluk oranlarıyla sınıflandırma yapılmıştır. Ancak bu çalışma, mevcut vakalar üzerinde gerçekleştirildiğinden, erken teşhis konusunda bir katkı sunmamış, yalnızca sakroiliit alanlarının otomatik tespitine odaklanmıştır.

AS'nin erken teşhisi üzere Zunti ve Ark. [15], BT görüntüleri üzerinde hem klasik ML algoritmaları hem de DL algoritmaları kullanarak AS'yi erken teşhis etmeyi hedeflemiştir. Klasik ML algoritmaları olarak RO ve k-NN yöntemlerini, DL araştırmaları için ise ön eğitilmiş InceptionV3 mimarisini kullanarak transfer öğrenme yöntemiyle sınıflandırma işlemleri gerçekleştirilmiştir. Çalışmada, toplamda 618 adet JPEG formatında bilgisayarlı tomografi görüntüsü kullanılmış ve 8 katlı çapraz doğrulama yöntemi uygulanmıştır. K-En yakın komşu algoritmasıyla yapılan analizlerde, doku erozyonu bulunan hastalar ile kontrol grubundaki bireyler %96 doğruluk oranıyla ayrılmıştır. Derin öğrenme mimarisi kullanılarak tasarlanan transfer öğrenme yöntemi ise doku erozyonu olan hastaları kontrol grubundan %99 doğrulukla ayırmayı başarmıştır. Klasik ML algoritmalarında, belirlenen ilgi alanlarından oluşturulan gri düzey eş oluşum matrislerinden yalnızca enerji, kontrast, korelasyon, homojenlik ve benzerlik parametreleri girdi olarak kullanılmıştır. Bu parametrelerin seçilmesindeki temel amaç, yaş farklılıklarından kaynaklanabilecek doku hasarını AS'den bağımsız tutmaktır. Ayrıca, yalnızca k-NN ve RO algoritmaları kullanılmış, GLCM oluşturulmadan önce herhangi bir veri ön işleme işlemi uygulanmamıştır.

Faleiros ve Ark. [16], SVM ve MLP gibi klasik makine öğrenimi tekniklerini kullanarak, MR taramalarından aktif enflamatuar sakroiliiti sınıflandırmak için bir sistem geliştirmişlerdir. Sistemde, girdinin boyutunu azaltmak için bir özellik seçme yaklaşımı uygulanmıştır. En iyi performans, 10-katlı çapraz doğrulama ile eğitilen veri setinde,

Wrapper özellik seçme yöntemi kullanılarak elde edilmiştir (duyarlılık = %100, özgüllük = %95,6, doğruluk = %84,7).

Diğer bir araştırmada [17] MRI görüntülerini kullanarak DL tabanlı bir yöntem geliştirmişlerdir. Bu yöntem, görüntüleri analiz ederek kalça kemik iliği ödemi otomatik olarak tespit etmekte ve derecelendirmektedir. Çalışma, AS hastalarında erken teşhis ve tedavi sürecini iyileştirmek amacıyla, derin öğrenmenin sağladığı yüksek doğruluk ve hassasiyetten yararlanarak, doktorlara yardımcı olacak bir araç sunduklarını iddia etmişlerdir. Önerilen sistem 40 test vakasının 31'ini başarılı sınıflandırmıştır ve doğruluk oranı %85,7'dir. Araştırmacılar MRI görüntülerini analiz eden otomatik bilgisayar tabanlı sistemlerin, AS derecelendirilmesi ve erken tespitinde büyük bir potansiyele sahip olduğunu belirtmişlerdir.

Li ve Ark. [11] tarafından gerçekleştirilen başka bir DL tabanlı AS sınıflandırma çalışmasında, 5389 pelvik radyografi verisi kullanılarak retrospektif bir çalışma yapılmış ve AS teşhisi için bir ensemble DL modeli geliştirilmiştir. Model, üç farklı tıp merkezinden alınan 583 görüntü üzerinde test edilmiş ve ROC analizi AUC (0.90), doğruluk (%94), kesinlik, duyarlılık (%91) ve F1 skorları (0.98) kullanılarak değerlendirilmiştir. Sonuçlar, modelin insan uzmanlardan daha yüksek performans gösterdiğini ve uzmanların teşhis doğruluğunu önemli ölçüde artırdığını göstermiştir. Geliştirilen klinik tahmin modeli, AS hastalarını yüksek ve düşük risk gruplarına doğru bir şekilde kategorize ederek bireyselleştirilmiş bakım için sağlam bir temel sunduğu iddia edilmiştir.

Tas ve Ark. [18] tarafından yapılan bir başka araştırmada, özgün bir MRI veri seti oluşturulmuş ve özellik çıkarımı, özellik seçimi ile sınıflandırma süreçlerinde k-NN sınıflandırıcısının kullanılması amacıyla önceden oluşturulmuş evrimsel sinir ağı (ESA) modellerinden faydalanılmıştır. Çalışmada, DL ve klasik ML algoritmaları tabanlı bir hibrit model geliştirilmiştir. Deneysel test sonuçları, önerilen modelin AS'nin erken tespitinde oldukça yüksek bir performans sergilediğini göstermektedir. Model, %98,7 doğruluk, %97,5 duyarlılık, %98,3 kesinlik ve %97,9 F-ölçüm değerlerine ulaşmıştır. Bu sonuçlar, önerilen hibrit yaklaşımın AS'nin erken teşhisinde etkili bir yöntem sunduğunu ortaya koymaktadır.

Literatür çalışmaları incelendiğinde, AS sınıflandırma için görüntü sınıflandırma temelli YZ kullanılarak yapılan çalışmaların sayısında ve popüleritesinde bir artış olduğu gözlemlenmiştir. Ancak hala klinik uygulamalarda aktif olarak kullanılan bir yaklaşıma

rastlanmamıştır. Buna ek olarak hiçbir çalışmada AS sınıflandırma ve teşhis için altın standart kabul edilen STIR echo sekanslı MRI görüntüsü [22-24] kullanılmamıştır. Diğer çalışmalardan farklı olarak tezimizde AS teşhisi için üç farklı yöntem kullanılarak performansları birbirleri ile karşılaştırılmıştır. Bunlardan ilki, GLCM özellikleri kullanılarak ML sınıflandırma modellerinin geliştirilmesi; ikincisi, GLCM özellikleri ve sayısal veriler kullanılarak bir DL modelinin geliştirilmesini kapsayan bir hibrid model oluşturulması; üçüncüsü ise deneklere ait görüntülerle eğitilerek, AS sınıflandırması yapan ve çoğunluk oylamasına göre çalışan keskin oylama modelinin geliştirilmesidir. Literatür taraması sonunda gerek makale veri tabanlarında ve de gerekse Ulusal Tez Merkezinde yukarıda sıralanan ikinci ve üçüncü yaklaşımı kullanarak gerçekleştirilen herhangi bir çalışmaya rastlanmamıştır (Şekil 1.3). 2023 yılına kadar romatoloji alanında kullanılan görüntü temelli Makine Öğrenmesi (ML) ve Derin Öğrenme (DL) ile ilgili makalelerden derlenen literatür taraması çalışması Dergipark, Erzincan Üniversitesi Fen Bilimleri Enstitüsü dergisinde yayınlanmıştır [25]. Sonrasında PRISMA kriterlerimize uyan makaleler de taramamıza eklenmiştir [11, 18-21].

2. GENEL BİLGİLER

Bu bölüm, tıp ve mühendislik alanlarında ortak bir dil geliştirme amacıyla hazırlanmıştır. Çalışmanın ana konusu olan AS hastalığının daha iyi anlaşılabilmesi için tarihçesi, tanımı, sınıflandırılması ve dünyadaki ve Türkiye'deki prevalansları (belirli bir nüfusta, belirli bir zaman diliminde belirli bir hastalığa sahip kişilerin oranı) ayrıntılı olarak incelenmiştir. Romatizmal hastalıklar çerçevesinde AS'in tanımları yapılmış; risk faktörleri, morbidite (bir toplulukta herhangi bir hastalığın görülme sıklığı) ve mortalite (belirli bir nüfusta, yaşa göre yıllık hayatta kalma ve ölüm oranlarını gösteren istatistikler) oranları ele alınmıştır. Hedef organ hasarı ve hastalığa bağlı diğer semptomların tespitinin önemi vurgulanmıştır. Bu bağlamda, çalışmanın halk sağlığı ve tıbbi sistem açısından taşıdığı önem açıklanmaya çalışılmıştır. Ardından, MRI görüntüleme yöntemi genel özellikleri ve STIR çekim teknikleri hakkında bilgi verilmiştir. Tezin tıp ve mühendislik alanlarında ortak bir terminoloji geliştirme amacına uygun olarak, hastalığa dair tıbbi bilgilerin yanı sıra YZ, ML ve DL konularında temel bilgiler ve tez kapsamında kullanılan yöntem ve teknikler sunulmuştur.

2.1. Ankilozan Spondilit (AS)

Kökene çok eski tarihlere dayanan AS hastalığını daha iyi tanımak adına tarihçesi, tanımı, dünyadaki ve ülkemizdeki prevalansı bu bölümde incelenmiştir.

2.1.1. AS Hastalığının Tarihçesi

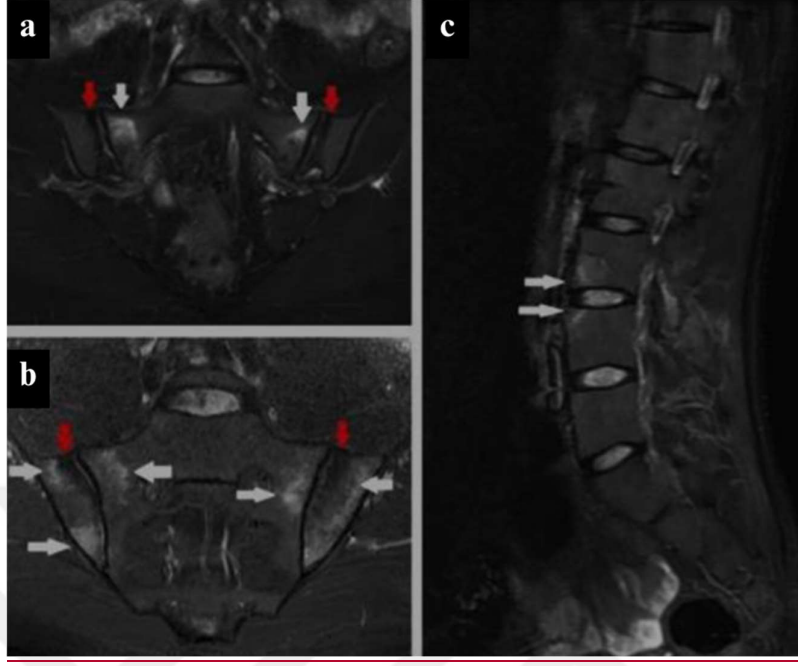
Son kanıtlar, M.Ö. 1500 civarına tarihlenen eski Mısır mumyalarında AS'nin mevcut olduğunu göstermektedir [26]. Bu kalıntılar, karakteristik ossifiye paraspinoz ligamanlar ve AS tanımıyla tutarlı belirgin SI eklem tutulumunu ortaya koymuştur. Omurga esnekliği ve ağrı ile ilgili ilk yazılı açıklamalar Hipokrat (MÖ 460 – MÖ 370) ve nörolog Caelius Aurelianus (1400'ler) tarafından yapılmış olsa da, bu tanımlamaların AS'yi güvenilir bir şekilde belirlediği söylenemez [27]. İrlandalı doktor Bernard Connor'ın omurgadaki bağların kalsifikasyonunu detaylandırarak sabit, kaynaşmış bir omurga ile sonuçlandığını belirten ilk kişi olduğuna inanılmaktadır [27]. Bununla birlikte, semptomatolojinin romatoid artrit ilerleyişi olduğu düşüncesi, önde gelen Rus nörolog Vladimir Bekhterev'in 1893'te AS'yi ayrı bir inflamatuvar hastalık olarak tanımasına kadar devam etmiştir [28]. Bu nedenle hastalık klasik olarak Bekhterev hastalığı olarak bilinir. AS'nin tarihçesini yazılı kayıtlardan izlemek, bu duruma verilen birçok isim nedeniyle zorlaşmaktadır [27]. AS, 1897'de Adolph

Strümpell ve 1898'de Pierre Marie tarafından resmi olarak tanımlanmış ve teşhis edilmiştir [29-31].

1800'lü yıllarda radyografi teknolojisinin gelişmesiyle birlikte klinik semptomlar görüntü bulgularıyla ilişkilendirilebilir hale gelmiştir. Bu tür bulgular, daha sonra lomber omurganın kareleşmesi, sindesmofit oluşumu, sakroileit ve patognomik bambu omurgasını içerecek şekilde açıklanmıştır [32]. 1980'lerde AS'yi incelemek için BT görüntüleme kullanılmaya başlanmıştır [33, 34]. Hem BT hem de manyetik rezonans görüntüleme (MRI) AS hastalığının tespitini artırmış ve 1990 yılına kadar erken hastalığın belirlenmesinde önemli bir rol oynamıştır [35, 36]. Günümüzde, bu teknolojiler AS'nin tanımlanmasında hayati öneme sahiptir ve tanı kriterlerinde temel teşkil etmektedir.

2.1.2. AS Hastalığının Tanımı, Tanısı, İlerleyişi ve Vücutta Oluşturduğu Hasarlar

AS, kronik, inflamatuvar bir romatizmal hastalık olup genellikle omurga ve sakroiliak eklemleri etkiler. AS'nin gelişiminde genetik yatkınlık, özellikle HLA-B27 geninin varlığı önemli bir rol oynar [37]. Hastalığın klinik semptomları erken dönemde bel ve kalça ağrısı gibi non-spesifik belirtilerle başlar ve ilerleyen aşamalarda omurgada hareket kısıtlılığı ve deformasyonlara yol açabilir [38]. AS'nin tanısı, klinik bulguların yanı sıra radyolojik görüntüleme ve laboratuvar testleri ile doğrulanır. Fiziksel muayeneye ek olarak farklı görüntüleme metodlarından elde edilen görüntülerin kullanılması X-Işını radyografi (X-Ray), ultrason (US), BT, MRI kullanılsa da AS'nin tanısında altın standart olarak MRI kabul edilmektedir [12, 22-24] (Şekil 2.1). Sakroiliak eklemlerde gözlemlenen subkondral kemik ödemi ve subkondral osteit, AS erken teşhis açısından önemli bulgular olarak değerlendirilir [9]. Bu belirtiler, hastalığın erken dönemlerinde sıkça görülür ve AS'nin diğer semptomlarıyla birlikte incelendiğinde teşhis için kritik bir rol oynar [10]. Ayrıca, subkondral kemik ödemi ve osteit, MRI teknikleri ile tespit edilebilir. Bu durum, hastalığın erken evrede tanı koyma sürecini kolaylaştırmaktadır [2, 11, 12, 22-24].

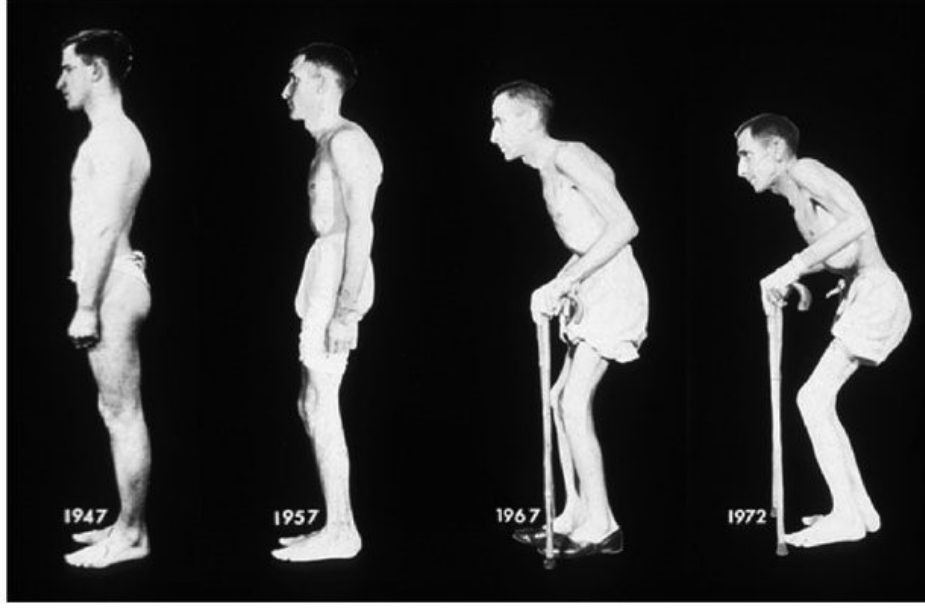


Şekil 2.1 Ankilozan Spondilit hastasının sakroiliak MRI bulguları [39] (a) İliak kanatlarda skleroz ve sakral kanatlarda kemik iliği ödemi (STIR sekansı), (b) Her iki taraf sakral ve iliak kanatlarda geniş alana yayılan kemik iliği ödemi (beyaz oklar) ve her iki eklem iliak kanatlarda skleroz (kırmızı oklar) (T1 ağırlıklı kontrastlı kesit), (c) lomber vertebra köşelerinde kemik iliği ödemi (beyaz oklar) görülmektedir (STIR sekansı)

Tedavi edilmediği takdirde hastalık ilerleyerek postural bozukluklar, fonksiyon kaybı ve yaşam kalitesinde belirgin düşüşe neden olabilir [40]. Şekil 2.2’de Ankilozan spondilitli bir hastanın fotoğraf serisi sunulmaktadır. Görüntüler 26 yıllık bir süre boyunca alınmıştır. 1957’de, torasik kifoz belirgin hale gelmiş ve normal lomber lordoz kaybı meydana gelmiştir. Kalçalar ve dizlerde erken fleksiyon kontraktürleri bulunmaktadır. İlerleyen iki seride, kalçalar ve dizlerdeki fleksiyon kontraktürlerinin arttığı ve omurganın ilerleyici ankilozisinin meydana geldiği gözlenmiştir.

AS, yalnızca omurga ve eklemleri etkilemekle kalmaz, aynı zamanda ilerleyen dönemlerde birçok sistemik rahatsızlığa yol açabilir. En yaygın ekstra-artiküler komplikasyonlar arasında üveit, inflamatuvar bağırsak hastalıkları ve psoriasis bulunur. Üveit, AS hastalarının %20-40’ında görülen, gözün ön kısmında inflamasyonla seyreden bir göz rahatsızlığıdır [41]. İnflamatuvar bağırsak hastalıkları, özellikle Crohn hastalığı ve ülseratif

kolit, AS'li bireylerin %5-10'unda gelişebilir ve genellikle kronik ishal ve karın ağrısı gibi belirtilerle kendini gösterir [42]. Psoriasis ise ciltte kızarıklık ve pullanma ile karakterize olup AS ile ilişkili olabilir [43]. Bu rahatsızlıkların yanı sıra, ilerleyen dönemlerde kardiyovasküler hastalıklar, osteoporoz, pulmoner fibrozis ve renal amiloidoz gibi ciddi sistemik komplikasyonlar da görülebilir [44].



Şekil 2.2 Ankilozan spondilit: Deformitelerin ilerleyişi [45].

AS'nin yönetiminde, non-steroid antiinflatuvar ilaçlar (NSAİİ), biyolojik ajanlar ve fiziksel rehabilitasyon önemli rol oynar. Erken tanı ve uygun tedavi ile hastalığın ilerleyişi kontrol altına alınabilir de, AS'nin uzun vadeli etkileri ve ilişkili komplikasyonlar hastaların yaşam kalitesini ciddi şekilde etkileyebilir [46].

2.1.3. AS'nin Dünya ve Türkiye'deki Yeri

AS, dünya genelinde farklı prevalans oranlarına sahip olup coğrafi, genetik ve çevresel faktörler bu dağılım üzerinde belirleyici rol oynamaktadır. Dünya genelinde AS prevalansı %0,1 ila %1 arasında değişirken, Kuzey Avrupa ve Amerika'da daha yüksek oranlar gözlemlenmektedir. Örneğin, Kuzey Avrupa'da bu oran %1'e kadar çıkabilmektedir, çünkü bu bölgelerde HLA-B27 pozitifliği daha yaygındır [38]. Asya ve Afrika gibi bölgelerde ise AS daha az sıklıkta görülür, bu da genetik farklılıklar ve çevresel faktörlerle ilişkilendirilmiştir. Ayrıca, bazı toplumlarda AS'nin geç tanı alması, farkındalık eksikliği ve

sağlık hizmetlerine sınırlı erişim de yaygınlığın doğru bir şekilde raporlanmasını zorlaştırmaktadır [32].

Türkiye'de ise AS prevalansı dünya ortalamalarına oldukça yakın seyretmektedir ve genel nüfusta yaklaşık %0,5 ila %1 oranında görülmektedir [47]. Türkiye'deki AS hastalarının genetik profili, HLA-B27 taşıyıcılığının daha düşük oranlarda bulunmasına rağmen, klinik seyrin dünya genelindeki diğer ülkelerle benzer olduğunu göstermektedir. Ancak Türkiye'deki AS hastalarının tanı alma süreci dünya genelindeki hastalara kıyasla daha geç gerçekleşmektedir. Bu durum, farkındalık eksikliği ve sağlık hizmetlerine erişimdeki farklılıklar nedeniyle, hastalığın genellikle ilerlemiş evrelerde teşhis edilmesine yol açmaktadır [47]. Buna karşılık, Avrupa ve Kuzey Amerika'da AS daha erken teşhis edilebilmekte ve bu da tedavi sürecini olumlu etkilemektedir. Türkiye'de yapılan bir çalışmada, AS hastalarının büyük bir kısmının geç teşhis aldığı ve bu nedenle hastalığın ilerleyişinin kontrol altına alınmasının zorlaştığı belirtilmiştir [48].

Bu veriler, AS'nin dünya çapında ve Türkiye'de önemli bir halk sağlığı sorunu olduğunu ve bölgesel farklılıkların tanı ve tedavi süreçlerini büyük ölçüde etkilediğini göstermektedir. Erken teşhis ve tedavi için farkındalık oluşturma çalışmaları, hastalığın uzun vadeli etkilerini hafifletmede kritik bir rol oynamaktadır. Türkiye'de, AS ile ilgili farkındalık kampanyalarının artırılması ve sağlık hizmetlerine daha kolay erişim sağlanması, hastalığın daha erken evrede tanınmasına ve yönetilmesine katkı sağlayacaktır [48].

2.1.4. Türkiye'de Ankilozan Spondilit ile İlişkili Sağlık Bakım Maliyetleri

Türkiye'de AS hastalığı ile ilişkili sağlık bakım maliyetleri, erken teşhis ve etkili tedavi yöntemlerine erişimde yaşanan zorluklar nedeniyle önemli bir yük oluşturmakta ve Avrupa'da AS hastalığının sağlık bakım maliyetleri ile benzerlik göstermektedir [48]. Türkiye'de yaygın vakalar için ortalama yıllık maliyet ortalama 4.233,15 €, acil vakalar için ise ayakta tedavi eczane maliyetleri ve ortak ödemeler hariç ortalama 2.253,63 € olduğu çalışmalarda kanıtlanmıştır [48]. Bu maliyetlerde özellikle geç teşhis edilen vakalarda biyolojik ilaç maliyetleri toplam harcamanın %77'sini oluşturmaktadır [48]. Özellikle hastalığın ilerleyişine bağlı olarak, uzun vadede sağlık sistemine ciddi maliyetler getiren komplikasyonlar ve ek tedavi gereksinimleri ortaya çıkmakta ve Türkiye'de AS hastalarının büyük bir kısmı geç teşhis aldığı için, tedavi süreci uzamakta ve daha ileri aşamalarda daha

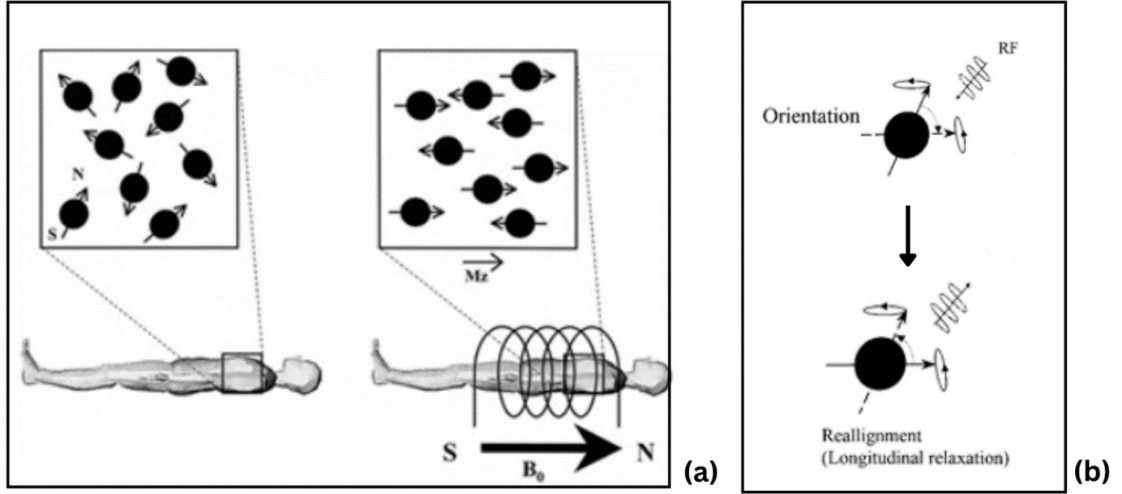
pahalı tedavi yöntemlerine başvurulması gerekmektedir. Bu durum hem doğrudan sağlık hizmetleri maliyetini hem de dolaylı iş gücü kayıplarını artırmaktadır [48].

Bu veriler, AS'nin Türkiye'deki ekonomik etkilerini anlamak için kritik öneme sahiptir ve sağlık politikalarında bu maliyetlerin azaltılması için farkındalık kampanyalarının artırılmasına, erken teşhis imkanlarının genişletilmesine ve biyolojik tedavilere daha kolay erişim sağlanmasına yönelik düzenlemelerin yapılması gerektiğini göstermektedir.

2.2. Manyetik Rezonans Görüntüleme

Manyetik rezonans görüntüleme (MRI), son yıllarda yumuşak dokuların görüntülenmesinde sıkça tercih edilen bir teknolojidir. Bu yöntemin temel çalışma prensibi, güçlü manyetik alanlar ve radyo dalgaları ile vücut dokularındaki protonların hizalanmasını ve uyarılmasını sağlamaktır [49-51]. Özellikle vücut dokularında yüksek oranda bulunan hidrojen atomları, bu süreçte önemli bir rol oynar. MRI görüntülemede kullanılan ana bileşenlerden biri, güçlü bir manyetik alan oluşturan büyük bir mıknatıstır. Bu manyetik alan, protonların belirli bir hizalanma yönüne girmesini sağlar [49-51]. Radyo frekans (RO) dalgaları ise bu hizalanmayı geçici olarak bozarak, protonların farklı bir enerji seviyesine geçmesine neden olur. Protonlar eski hallerine geri dönerken, RO sinyallerini tekrar yayarlar (Şekil 2.3). Bu sinyaller, vücuttaki dokuların yoğunluğu ve diğer özelliklerine bağlı olarak değişkenlik gösterir. MR cihazı bu enerjiyi algılayarak vücudun iç yapısının detaylı görüntülerini oluşturur [51, 52]. Özellikle beyin, omurga, kas-iskelet sistemi gibi yumuşak dokularda MR'ın üstün görüntüleme kabiliyetleri dikkat çeker [49-54].

Farklı dokuların MR sinyallerine verdiği tepki süresi, T1 ve T2 adı verilen zaman sabitleriyle ölçülür. T1 gevşeme zamanı, protonların manyetik alanla tekrar hizalanma süresini, T2 gevşeme zamanı ise protonların enerjilerini kaybedip eski hallerine dönme sürecini ifade eder. Bu süreler dokular arasında farklılık gösterdiğinden, MR görüntülemede çeşitli doku tiplerini ayırt etmek mümkündür. Genel olarak T1 ağırlıklı görüntüler anatomik veriyi, T2 ağırlıklı görüntüler ise patolojik veriyi daha iyi gösterir. Ayrıca, farklı MR sekanslarının kullanımıyla (T1-ağırlıklı, T2-ağırlıklı, FLAIR gibi) spesifik anatomik veya patolojik durumların daha net şekilde görüntülenmesi sağlanabilir [49-55].



Şekil 2.3 MRI cihazının temel çalışma prensibinin şematik gösterimi [51].(a) Manyetik alanın çekirdeklere etkisi.(b) Çekirdeklerin kendi eksenleri etrafında dönmesi ve harici manyetik alan yönü etrafında presesyon hareketi yapması.

2.2.1. Short Tau Inversion Recovery (STIR) Sekanslı MRI görüntüleri

Short Tau Inversion Recovery (STIR), MRI yaygın olarak kullanılan ve özellikle anormal doku kontrastı içeren çeşitli patolojik durumların tespitinde ve karakterizasyonunda büyük öneme sahip bir tekniktir. Bu teknik, farklı doku tiplerinin uzunlamasına gevşeme süreleri (T_1) arasındaki doğal farklılıkları kullanarak belirli dokuların sinyalini seçici olarak baskılar ve böylece klinik olarak ilgi duyulan bölgelerin görüntülenmesini artırır.

Short Tau Inversion Recovery'nin temel prensibi, 180 derecelik bir radyo frekans darbesinin uygulanması ve ardından nispeten kısa bir inversiyon zamanının kullanılmasıdır; bu, belirli bir T_1 gevşeme süresine sahip dokulardan gelen sinyalin seçici olarak yok edilmesine olanak tanır [56]. Yağ veya beyin omurilik sıvısı gibi belirli dokulardan gelen sinyalin bu şekilde hedeflenerek baskılanması, ilgi duyulan doku ile çevresindeki yapılar arasındaki kontrastın belirgin bir şekilde artmasını sağlayarak doğru tanı ve tedavi planlamasını kolaylaştırır.

Short Tau Inversion Recovery'nin kullanım alanı, kardiyovasküler görüntüleme ve nörogörüntüleme gibi çeşitli klinik uygulamaları kapsamaktadır. Kardiyovasküler manyetik rezonans görüntülemede, bu teknik, miyokard ödemi, enfarktüs ve fibrozisin görüntülenmesini iyileştirmek için kullanılarak çeşitli kardiyovasküler patolojilerin tanı ve

yönetimine katkı sağlar [57]. Benzer şekilde, nörogörüntülemeye, Short Tau Inversion Recovery, normal beyin parankimini baskılayarak ve patolojik değişikliklerin olduğu bölgeleri öne çıkararak, tümörler ve demiyelinizan hastalıklar gibi beyin lezyonlarının tespitinde ve karakterizasyonunda etkili olduğunu göstermiştir [57-60]

2.2.2. STIR Sekansı ve AS

STIR sekansı, AS tanısında özellikle erken inflamatuvar değişikliklerin tespitinde kritik bir rol oynamaktadır. Kemik iliği ödemi gibi AS'nin tipik bir özelliği olan inflamasyonu hassas bir şekilde saptayabilen bu sekans, yağ sinyalini baskılayarak iltihaplı dokuların kontrastını artırır ve özellikle sakroiliak eklemler ile omurgadaki inflamasyon odaklarının net bir şekilde görüntülenmesine olanak tanır. Bu özelliği sayesinde, AS gibi aksiyel iskeleti etkileyen inflamatuvar hastalıkların erken tanısı ve hastalığın ilerlemesinin izlenmesi açısından güvenilir bir yöntemdir [61]. Radyografik değişiklikler gözlenmeden önce inflamasyon belirtilerini tespit edebilen STIR, patolojik değişikliklerin olduğu bölgeleri seçici sinyal baskılama yoluyla daha net görüntüler. Bu, doğru tanıyı kolaylaştırırken, uygun tedavi stratejilerinin belirlenmesine de rehberlik eder [62].

2.3. Yapay Zekâ Tanımı ve Yaklaşımları

Evrensel olarak kabul görmüş kesin bir tanımı olmayan AI kavramı hem disiplin içinde hem de disiplin ötesinde çeşitli bağlamlarda kullanılmıştır. Örneğin, Turing [63] makine zekası kavramı üzerinden YZ'i tanımlamış ve insan yanıtlarını ikna edici şekilde taklit edebilen bir sistem üzerinde durmuştur. Benzer şekilde McCarthy [64] insansı bilişsel işlevlerini taklit eden zeki makineler yapma bilimi olarak kavrama yeni bir boyut kazandırmıştır. 90'lı yıllara geldiğimizde Kirsh [65] öğrenme , algılama ve dil anlama gibi insan zekasına ait özellikleri problem çözme amaçlı kullanılan bir disiplin olarak YZ'i tanımlamaya çalışmıştır. Russell ve Norving [66] ise AI kavramını dört farklı yaklaşımla ele almaya çalışmışlardır: insan gibi düşünme, insan gibi hareket etme, rasyonel düşünme ve rasyonel hareket etme. 2000'li yıllarda Brachman [67] YZ'i verilerden öğrenen ve uyum sağlayan sistemlerin geliştirilmesi olarak tanımlayarak özellikle DL ve YSA'ye vurgu yapmıştır. Buna ek olarak Monett ve Lewis [68] YZ'nin zeki davranışları simüle etmek için tasarlandığını vurgulayarak, günlük yaşam ve karar verme süreçlerine gitgide daha fazla entegre olacağını belirtmişlerdir. Zekanın karmaşıklığı göz önüne alındığında, yaygın olarak kabul edilen bir YZ tanımını tahmin etmek imkansız olduğu görülmektedir [69].

2.3.1. Makine Öğrenmesi

ML, bilgisayar biliminde mevcut verilerin, tahmin etmek veya bunlara yanıt vermek için kullanıldığı bir alandır. YZ'nin alt dallarından biridir. Tanım olarak ilk olarak Arthur Samuel tarafından kullanılmıştır [70]. Örüntü tanıma, hesaplamalı istatistik ve YZ alanları ile yakından ilgilidir. Yüz tanıma, spam filtreleme vs. gibi alanlarda makine öğrenimi önemlidir [71].

Farklı kullanım alanları olsa da algoritma yapıları ile çözümün zor hatta mümkün olmadığı durumlarda tahmine ya da sınıflamaya yardımcı olur. Makine öğrenimi verilerden anlam kazanmak için makinelerin (bilgisayarlar ve yazılımlar) kullanılması anlamına gelebileceği gibi aynı zamanda makinelere çevrelerinden öğrenme vasıtası ile yetenek kazandırmak anlamına da gelebilmektedir [72].

Öğrenme, bir konu hakkında bilgi veya beceri edinme sürecidir. İnsan öğrenimi ile karşılaştırıldığında, makine öğrenimi daha hızlı öğrenir, bunun sonucunda ise bilgi birikimini daha kolay bir şekilde ilişkilendirir. Yani, insanların ML alanında yaptıkları gelişim, bilgisayarların yapabileceklerini arttıracak ve dolayısıyla insan toplumu üzerinde olumlu bir etkiye sahip olacaktır [73]. Böylelikle tanımımıza ML'nin insan zekasını taklit edebilen bilgisayar algoritmaları geliştirme teknolojisi olduğunu da ekleyebiliriz. Bunu yaparken YZ, olasılık ve istatistik, bilgisayar bilimi, bilgi teorisi, psikoloji, kontrol teorisi ve felsefe gibi farklı disiplinlerden fikirlerden yararlanmaktadır [74-77].

Makine öğrenmesi farklı amaçlar için kullanılabilir. Makine öğrenmesi algoritmaları temel olarak birbirinden farklılık gösteren amaçlarına göre üç kategoriye ayrılır. Denetimli öğrenme, denetimsiz öğrenme ve pekiştirmeli öğrenmeyi içerir [78-80]. Denetimli öğrenme, modelin etiketlenmiş veriler üzerinde eğitilmesini içerir ve bu eğitilmiş modeli yeni veriler üzerinde tahminler yapmak için kullanır. Verilerin eğitim seti ve test seti olmak üzere iki gruba ayrılmasını içerir. Denetimli sınıflandırmada, etiketlenen değer münferit bir değerdir. Buradaki algoritmalar, problemin hangi sınıfa veya kategoriye ait olduğunu sınıflandırmak için kullanılır.

Denetimsiz öğrenmede, etiketli değerlerin veya hedef değerlerin bilinmediği gerçeği dışında, verilerin eğitimini de içerir. Bu tip öğrenmede, makine gizli deseni bularak benzer türdeki

verileri kümelemeye çalışır. Denetimsiz öğrenmenin temel amacı tahmin yapmaktan ziyade kalıpları keşfetmektir. Denetimli öğrenme etiketli veriler üzerinde ve denetimsiz öğrenme etiketlenmemiş veriler üzerinde çalıştığından, etiketlenmemiş verilerden elde edilebilen etiketli verilerden birçok bilgi kaybolur. Dolayısıyla bu durumda yarı denetimli öğrenme akla gelmektedir. Hem etiketlenmemiş hem de etiketlenmiş verileri aldığı denetimli ve denetimsiz öğrenmenin bir karışımıdır. Pekiştirmeli öğrenme, çevreden geri bildirim alarak Performansını artıran bir sistem geliştirerek ve bunları geliştirmek için olası adımlar atarak çalışır. İnsanlardan herhangi bir yardım almadan çevre ile etkileşime girerek ondan öğrenme eyleminin gerçekleştiği yenilemeli bir süreçtir [72, 75, 76, 79]. Bir sonraki alt başlıklarda AS teşhisi için bu tez kapsamında kullanılan YZ teknikleri ile ilgili teorik bilgiler verilecektir. Önerilen ML tabanlı ve DL tabanlı modeller için kullanılan teknikler sırası ile: Destek Vektör Makinesi algoritması (SVM), k-En yakın komşu algoritması (k-NN), Topluluk Öğrenmesi Algoritması (TOA), Çoğunluk Oylama Algoritması (ÇOA), Yapay Sinir Ağları (YSA).

2.3.1.1. Destek Vektör Makinesi

SVM, sınıflandırma, regresyon ve aykırı değer tespit görevleri için yaygın olarak kullanılan denetimli bir öğrenme modelidir. 1963 yılında Vladimir Vapnik ve Alexey Chervonenkis tarafından tanıtıldı ve o zamandan beri sağlamlığı ve yüksek boyutlu verileri işleme yeteneği nedeniyle makine öğreniminde en popüler yöntemlerden biri haline gelmiştir [81, 82]. SVM, veri noktaları ile hiperdüzlem arasındaki marjı maksimuma çıkararak farklı veri sınıflarını en iyi şekilde ayıran bir hiperdüzlem bulmayı amaçlamaktadır [81].

SVM'nin temel konsepti, aşağıdakilerle tanımlanan en uygun hiperdüzlemi bulmaktır. Bu hiperdüzlem denklem 2.1'de ki gibi ifade edilebilir [83].

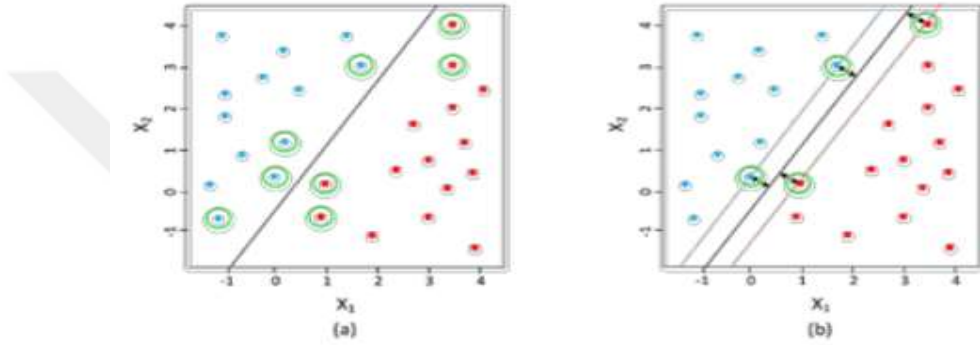
$$\beta^T x + \beta_o = 0 \quad (2.1)$$

Bu denklemde β^T ağırlık vektörü, x giriş özellik vektörü ve b önyargı terimidir. Amaç β^T ve β_o yi kullanarak hiperdüzleme en yakın veri noktaları olan destek vektörleri arasındaki marjı maksimuma çıkaracak hiperdüzlemi tespit etmektir.

Bu denklemin belirttiği doğrunun eğimi denklem 2.2'de ki gibi ifade edilmektedir [83].

$$Eğim = \beta^T = \frac{\sin \theta}{\cos \theta} = \frac{\sin \theta}{\beta_o} \quad (2.2)$$

Böylelikle ilgili doğru denkleminin oluşturduğu hizaya göre yeni eklenen değerin hangi grupta olduğu tespit edilebilmektedir (Şekil 2.4).



Şekil 2.4 $\beta^T x + \beta_o = 0$ doğrusunun konumu [84]. (a) İlgili doğruya yakın olan farklı sınıf değerlerinin gösterimi, (b) Destek vektörlerinin gösterimi

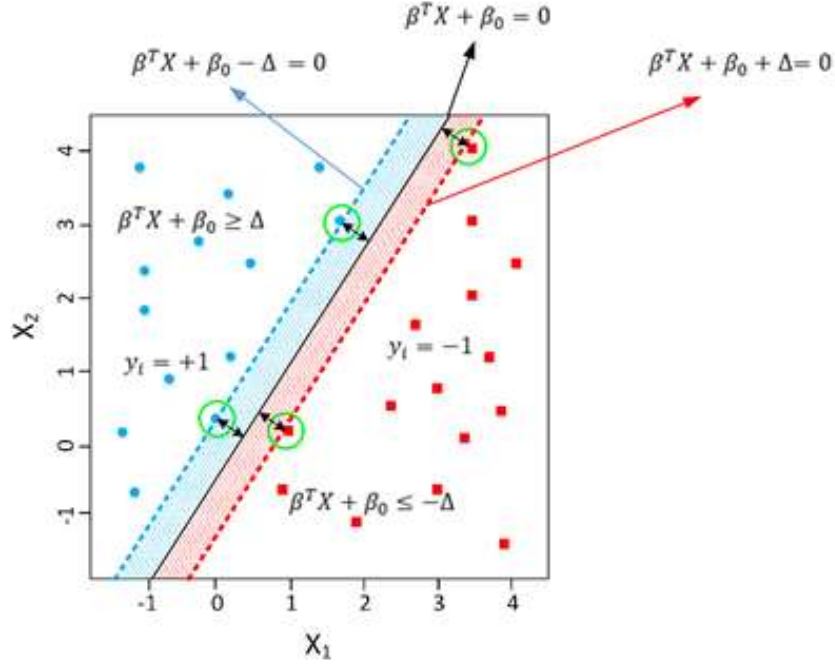
Şekil 2.4(a)'da görülen $\beta^T x + \beta_o = 0$ doğrusunun konumunu, farklı sınıflara ait en yakın noktalar belirler. Örneğin, Şekil 2.4(a)'da bu denkleme yakın olan pek çok farklı sınıfa ait nokta bulunmaktadır. Ancak doğrunun tam yerini, Şekil 2.4(b)'de gösterilen en yakın noktalar belirler. Farklı sınıflarda bulunan ve $\beta^T x + \beta_o = 0$ doğrusuna en yakın noktalar üzerinden geçen, bu doğruya paralel başka doğrular da vardır. Şekil 2.4(b)'de kırmızı ve mavi noktalarla gösterilen bu doğrulara destek vektörleri denir. Ayrıca Şekil 2.4 (b)'deki doğruya en yakın olan farklı sınıflardan geçen başka paralel doğrular da mevcuttur.

Şekil 2.5'de ana doğru denklemine Δ kadar yakından geçen alt ve üst destek vektörleri sunulmaktadır. Buna göre üst destek vektör denklemi (Denklem 2.3), Düzlem Doğru denklemi (Denklem 2.4) ve Alt destek vektör denklemi (Denklem 2.5) şu şekilde ifade edilebilir:

$$\beta^T x + \beta_o + \Delta = 0 \quad (2.3)$$

$$\beta^T x + \beta_o = 0 \quad (2.4)$$

$$\beta^T x + \beta_o - \Delta = 0 \quad (2.5)$$



Şekil 2.5 Düzlem, üst ve alt destek doğru denklemleri [84].

SVM, özellik sayısının örnek sayısından daha fazla olduğu durumları yönetebildiğinden, yüksek boyutlu verilerle uğraşırken özellikle etkilidir. Sağlam genelleme yeteneği, onu metin sınıflandırma, biyoinformatik ve görüntü tanıma görevleri için iyi bir seçim haline getirir [85]. Ek olarak, marjı maksimize eden karar sınırı aracılığıyla net bir geometrik yorum sağlayarak sinir ağları gibi daha karmaşık modellere kıyasla anlaşılmasını kolaylaştırır [86, 87].

Ancak SVM'nin uygun çekirdek fonksiyonlarının ve hiper parametrelerin seçimine güvenmesi, ayarlamayı zorlaştırabilir. Ek olarak, doğrusal olmayan çekirdekler kullanıldığında, orijinal özellik uzayında karar sınırı daha az sezgisel hale geldiğinden modelin yorumlanabilirliği azalır[87].

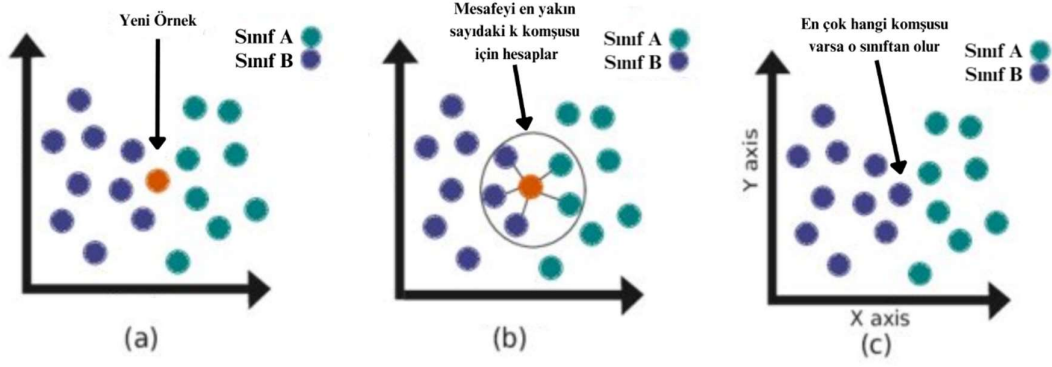
2.3.1.2. k-En Yakın Komşu Algoritması

Makine öğreniminde hem sınıflandırma hem de regresyon görevleri için kullanılan basit ama etkili, parametrik olmayan bir algoritma olan k-NN, özellikle bir sorgu noktasının özellik alanındaki en yakın veri noktalarına olan benzerliğine dayalı tahminler yapan sezgisel yaklaşımıyla tanınır. Bu yöntem açık bir eğitim aşaması gerektirmez, bu da onu bir tür örnek tabanlı öğrenme haline getirir [88, 89]. k-NN, görüntü tanıma, öneri sistemleri ve tıbbi teşhis dahil olmak üzere çeşitli alanlarda yaygın olarak kullanılmaktadır. Esnekliği ve basitliği, onu girdi veri noktaları arasındaki ilişkilerin karmaşık olduğu ve daha katı modellerle kolayca yakalanamadığı problemler için uygun kılar [74].

k-NN algoritması, özellik uzayındaki en yakın k komşusu arasındaki çoğunluk sınıfına dayalı olarak yeni bir veri noktasını sınıflandırır. Regresyon için çıktıyı bu komşuların hedef değerlerinin ortalaması olarak tahmin eder. Komşuları belirlemek için kullanılan mesafe ölçüsü tipik olarak Öklid mesafesidir, ancak Manhattan veya Minkowski mesafeleri gibi diğer ölçümler de kullanılabilir [90]. Sınıflarda bulunan örnekleri $x_i = (x_{i1}, x_{i2}, \dots, x_{in})$ ve sınıflandırılmak istenen verileri de $y_i = (y_{i1}, y_{i2}, \dots, y_{in})$ 'nin temsil ettiği düşünülürse, iki mesafe arasındaki öklit uzaklığı denklem 2.6'daki gibi hesaplanır.

$$d(x, y) = \sqrt{\sum_{i=1}^k (x_i - y_i)^2} \quad (2.6)$$

Mesafeler hesaplandıktan sonra algoritma en yakın k komşuyu, yani sorgu noktasına en yakın eğitim örneklerini tanımlar. Sınıflandırma için k-NN, sorgu noktasını bu k komşu arasında en sık meydana gelen sınıfa atar. Örneğin, bir noktanın en yakın beş komşusu varsa (Şekil 2.6) A Sınıfından üç ve B Sınıfından iki örnek içerirse, algoritma noktayı A Sınıfına ait olarak sınıflandırır.



Şekil 2.6 $k=5$ 'e göre hesaplanmış bir k-NN algoritması uygulama örneği. (a) Yeni örnek ve bilinmeyen bir mesafe (Turuncu daire). (b) $k=5$ için yakın komşuların bulunması. (c) Turuncu dairenin hangi sınıfa ait olduğunun fazla olan yakın komşusuna göre belirlenmesi

Denklem 2.6'da ifade edilen k parametresi, farklı gözlem birimlerini temsil etmekte olup k-NN algoritmasının performansı üzerinde doğrudan bir etkiye sahiptir. Küçük bir k değeri, modelin gürültüye karşı daha hassas olmasına yol açarak aşırı öğrenme riskini artırabilir. Öte yandan, çok büyük bir k değeri, gürültüye karşı dayanıklılık sağlasa da sınıflar arasındaki ince ayrımların gözden kaçmasına ve eksik öğrenmeye neden olabilir [76, 91]. k değerinin uygun şekilde belirlenmesi için k-katlı çapraz doğrulama yöntemleri sıkça kullanılmakta, ayrıca Bayesian en iyileyicisi gibi ileri algoritmalar da bu süreçte etkili bir şekilde tercih edilebilmektedir. Bununla birlikte, algoritma, sorgu noktası ile eğitim kümesindeki tüm noktalar arasındaki mesafelerin hesaplanmasını gerektirdiği için büyük veri kümelerinde hesaplama maliyetlerini önemli ölçüde artırabilir [92].

Ayrıca, k-NN algoritmasının performansı, büyük ölçüde seçilen mesafe ölçütüne ve veri özelliklerinin ölçeğine bağlıdır. Özelliklerden birinin daha geniş bir aralığa sahip olması durumunda, bu özelliğin mesafe hesaplamalarında baskın hale gelmesi olasıdır. Bu tür ölçeklenme sorunlarının önüne geçmek için özelliklerin normalleştirilmesi veya standartlaştırılması önerilmektedir [90].

2.3.1.3. Toplu Öğrenme Algoritması

TOA tahmin performansını artırmak, aşırı öğrenmeyi azaltmak ve bireysel modellere kıyasla genelleme yeteneklerini geliştirmek için birden fazla modelin birleştirildiği, makine öğreniminde güçlü bir tekniktir. TOA'nın temel amacı, daha doğru ve sağlam bir model

üretmek için farklı modellerin güçlü yönlerini bir araya getirmektir. Çoklu algoritmalarından gelen tahminleri entegre ederek topluluk yöntemleri, özellikle karmaşık ve yüksek boyutlu veri kümelerinde tek modelli yaklaşımlardan daha iyi performans gösterebilir [93]. Topluluk halinde öğrenmenin arkasındaki teori, farklı model setlerinin, bireysel olarak zayıf olsalar bile, uygun şekilde birleştirildiğinde kolektif olarak daha doğru tahminler yapabileceği önermesine dayanmaktadır [94].

Rastgele Orman (RO) daha sağlam bir sınıflandırıcı oluşturmak için birden fazla karar ağacını birleştiren bir topluluk tekniğidir. Ormandaki her ağaç, önyükleme toplama veya torbalama adı verilen bir teknik kullanılarak oluşturulan eğitim verilerinin farklı bir alt kümesiyle eğitilir. Bu süreç, veri kümesinin değiştirilerek rastgele örneklenmesini içerir ve bu da her ağaç için eğitim örneklerinde farklılıklara yol açar. Eğitim sırasında, Rastgele Ormanlar ayrıca her bir bölünmedeki özelliklerin seçimini rastgele hale getirerek bireysel ağaçları daha da çeşitlendirir ve aralarındaki korelasyonu azaltır [95]. Nihai tahmin, tüm ağaçlar arasında çoğunluk oyu (sınıflandırma için) alınarak veya ortalama alınarak elde edilir [96].

TOA'nın gücü, yüksek boyutlu verileri ve özellik etkileşimlerini ele alma yeteneklerinde yatarken, ağaçlar arasında ortalama alma işlemi nedeniyle fazla uyum sağlamaya daha az eğilimli olmalarında yatmaktadır. Farklı görüntüleme yöntemlerinden veya veri kaynaklarından elde edilen sonuçları birleştirerek tanı doğruluğunu artırmaya yardımcı olan biyomedikal araştırmalar gibi çeşitli alanlarda yaygın olarak uygulanmaktadır [97].

2.3.1.4. Çoğunluk Oylama Algoritması

Oylama algoritmaları, makine öğrenimi ve derin öğrenme modellerinde kullanılan topluluk öğrenme tekniklerinin ayrılmaz bir parçasıdır. Bu algoritmalar birden fazla modelin tahminlerini bir araya getirerek daha güçlü ve kesin bir nihai tahmine ulaşmak için kullanılır [98].

Oylama algoritmalarının iki ana türü vardır: sert oylama ve yumuşak oylama. Sert oylama, çoğunluk oylaması olarak da bilinir ve nihai tahminin, bireysel modellerden en fazla oyu alan sınıfa göre belirlendiği basit bir yaklaşımdır. Öte yandan, yumuşak oylama veya ağırlıklı ortalama oylama, her bir model tarafından atanan olasılıkları veya güven skorlarını dikkate alarak ağırlıklı bir ortalama hesaplar ve bu şekilde nihai tahmini belirler [98].

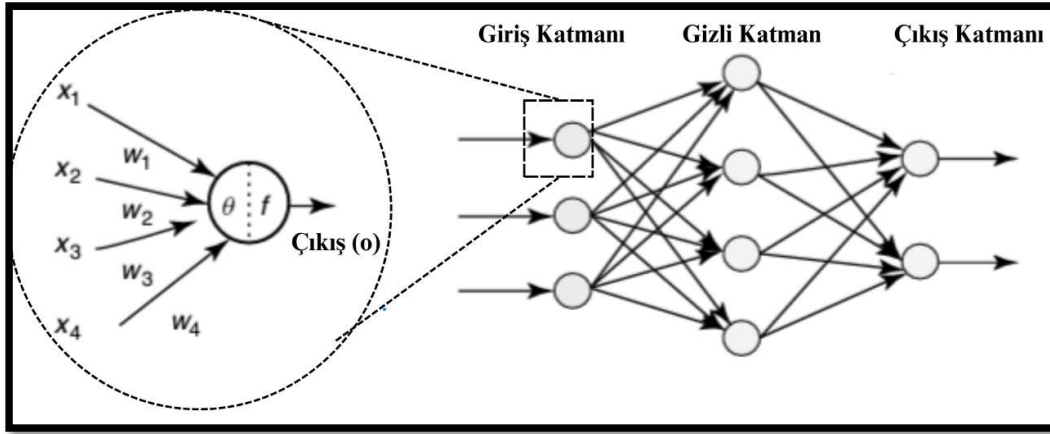
Oylama algoritmalarını kullanan topluluk öğrenme tekniklerinin, birden fazla algoritmanın güçlü yönlerinden yararlanarak ve zayıflıklarını azaltarak bireysel modelleri geride bıraktığı gösterilmiştir [99-101]. Ayrıca, oylama yöntemlerini kullanarak birden fazla algoritmayı birleştiren hibrit topluluk sınıflandırıcılarının, daha yüksek tahmin doğruluğu elde etmek amacıyla önerildiği belirtilmiştir [99].

2.3.1.5. Yapay Sinir Ağları

YSA, makine öğrenimi alanında, karmaşık problem çözme görevlerine yaklaşımımızda devrim yaratan güçlü bir araç olarak ortaya çıkmıştır. İnsan beyninin biyolojik yapısı ve işlevinden ilham alan bu hesaplamalı modeller, doğal zekamızın gelişmesini sağlayan sinir bağlantılarını ve öğrenme süreçlerini taklit edecek şekilde tasarlanmıştır [102]. Bu yüzden YSA kavramının net bir şekilde anlaşılabilmesi için bir sinir hücresinin (nöron) yapısını iyi anlaşılması gerekmektedir. İnsan serebral korteksi yaklaşık 10^{10} nöron içerir. Sinapsları dallandıran ve sonlandıran sinir şeritleri (aksonlar) ile birbirine bağlanırlar. Bu sinapslar diğer nöronlarla olan bağlantılardır. Sinapslar, diğer nöronlardan elektrik sinyalleri biçiminde girdi toplamak üzere tasarlanmış, sinir hücresi gövdesinden uzanan dallar olan dendritlere bağlanır. İnsan beynindeki bir nöronun diğer nöronlarla binlerce sinaptik bağlantısı olabilir. Sonuçta serebral kortekste birbirine bağlı nöronlardan oluşan ağ, görsel, işitsel ve duyuşsal verilerin işlenmesinden sorumludur.

YSA, girdi verilerini işlemek ve anlamlı çıktılara dönüştürmek için birlikte çalışan, beyindeki nöronlara benzer şekilde birbirine bağlı düğümler bulunur. Bu ağların, bağlantılarının güçlü yönlerini ayarlayarak öğrenme ve uyum sağlama yeteneği, onları geleneksel kural tabanlı algoritmalarından ayırmakta ve çözülmesi zor olan sorunların üstesinden gelmelerine olanak tanımaktadır [103].

Bir YSA'nın temel yapısı bir giriş katmanı, bir veya daha fazla gizli katman ve bir çıkış katmanından oluşur (Şekil 2.7). Her katman, temel hesaplama birimleri olan çok sayıda yapay nöronlardan oluşur. Bu nöronlar girdileri alır, ağırlıklı toplamı ve doğrusal olmayan aktivasyon fonksiyonunu uygular [103-105].



Şekil 2.7 Yapay bir nöronun ve çok katmanlı bir sinir ağının mimarisi [105].

Girişler olarak tanımlanan $x_1, \dots, x_n$ 'den gelen sinyal akışı, çıkışa (o) doğru Şekil 2.7'da gösterildiği üzere tek yönlü akış yaptığı kabul edilmektedir. Çıkış matematiksel olarak şu şekilde ifade edilebilir (Denklemler 2.7) [105]:

$$o = f(net) = f\left(\sum_{j=1}^n w_j x_j\right) \quad (2.7)$$

w_j ağırlık vektörünü simgelerken $f(net)$ aktivasyon fonksiyonunu temsil etmektedir. Dolayısı ile giriş ve çıkış sinyalleri arasındaki ilişki, girişlerin ağırlıklı toplamını çıkışa dönüştüren doğrusal olmayan bir fonksiyon olan aktivasyon fonksiyonu tarafından belirlenir. Aktivasyon fonksiyonu, diğerlerinin yanı sıra sigmoid fonksiyonu, hiperbolik tanjant fonksiyonu veya düzeltilmiş doğrusal birim fonksiyonu gibi çeşitli biçimlerde olabilir. Bu aktivasyon fonksiyonları, girdilerin ağırlıklı toplamını çıktıya dönüştüren doğrusal olmayan işlevlerdir.

YSA ile bir sorunu çözmek için ağın, eğitim verileri olarak bilinen bir dizi giriş-çıkış çifti üzerinde eğitilmesi gerekir. Eğitim süreci, tahmin edilen çıktılar ile gerçek çıktılar arasındaki farkı en aza indirmek için ağın ağırlıklarının ve sapmalarının ayarlanmasını içerir.

2.4. Derin Öğrenme

DL, bilgisayarlara insanlara doğal olarak gelen şeyi yapmayı öğreten bir makine öğrenimi dalıdır. Deneyimden öğrenme felsefesi ile çalışır. DL'nin detaylı tanımının yapılmasına geldiğinde şu anda hala kesin bir birlikteliğe varılan bir tanımlı bulunmamaktadır [106]. Dört adet temel DL tanımı yapılmıştır:

Tanım 1: DL, daha düşük seviyedeki bilgilerden (örneğin, ses, resim vb.) oluşan bilgi yığını, çoklu düzeylerde temsil ve soyutlama yöntemlerini kullanarak öğrenme işlemini gerçekleştirmektedir [106].

Tanım 2: Bengio'ya göre [107], "DL, daha yüksek seviyelerden özelliklere sahip öğrenme özelliği hiyerarşilerini, alt düzeyin bileşiminin oluşturduğu hiyerarşinin özelliklerinden oluşturmaktır."

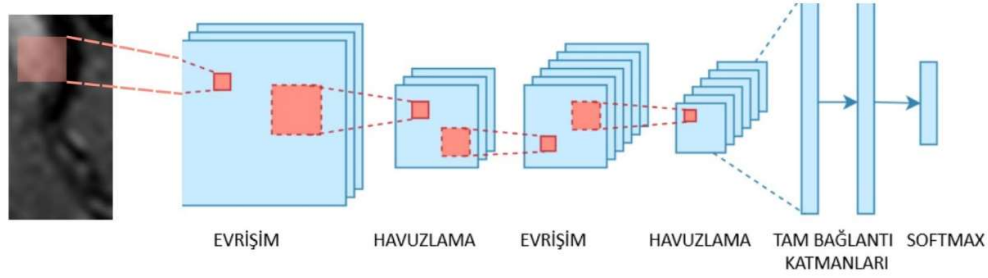
Tanım 3: DL (yapılandırılmış öğrenme veya hiyerarşik öğrenme olarak da bilinir) göreve özgü algoritmaların aksine bir veri temsildir. Öğrenme denetlenebilir, yarı denetimli veya denetimsiz olarak gerçekleştirilebilir [108, 109].

Tanım 4: Deng'ye göre [110], "DL bir çeşit ML'dir. Şu özellikleri taşır: (1) kademeli kullanım özelliği için birden fazla doğrusal olmayan işlem birimi katmanı kullanır. Her ardışık katmanın girdisi, bir önceki katmanın çıktısıdır. (2) farklı soyutlama seviyelerde ki hiyerarşik kavramlardan temsil seviyeleri vasıtası ile öğrenme gerçekleşir.

DL, özelliklerin yararlı temsillerini doğrudan verilerden öğrenmek için sinir ağlarını kullanır [111, 112]. Sinir ağları, paralel olarak çalışan ve biyolojik sinir sistemlerinden ilham alan basit öğeleri kullanarak birden fazla doğrusal olmayan işleme katmanını birleştirir. DL modelleri, nesne sınıflandırmasında bazen insan düzeyindeki Performansı aşan en son teknoloji doğruluğu sağlayabilir. Modelleri, genellikle bazı evrimsel katmanlar da dahil olmak üzere birçok katman içeren çok sayıda etiketli veri ve sinir ağı mimarileri kullanarak eğitilebilir. Bu modellerin eğitimi, hesaplama açısından yoğundur ve genellikle yüksek Performanslı bir grafik işlemci ünitesi (GPU) kullanarak eğitimi hızlandırılması gerekmektedir.

2.4.1. Evrişimsel Sinir Ağları

Romatoloji alanında DL kullanan görüntü işleme yöntemleri incelendiğinde karşımıza çıkan en önemli ifade evrişimsel sinir ağları (ESA) kavramıdır. ESA, hayvan görsel korteksinden biyolojik olarak ilham alan çok katmanlı bir sinir ağıdır [103]. Mimari özellikle görüntü işleme uygulamalarında kullanışlıdır. Şekil 2.8, ESA'nın yeni görüntüleri sınıflandırmak için birçok görüntüden özellikleri otomatik olarak öğrenen katmanları nasıl birleştirdiğini göstermektedir ve farklı katmanlardan oluşmaktadır. Her dikdörtgen görüntü, görüntü konumlarının her birinde saptanan öğrenilen özelliklerden birinin çıktısına karşılık gelen bir özellik haritasıdır. Yönlendirilmiş kenar detektörleri olarak hareket eden daha düşük seviyeli özellikler ile bilgi soldan sağa doğru akar ve en son aşamada sınıflandırma işlemi gerçekleştirilir.



Şekil 2.8: Bir evrişimsel sinir ağının temel katmanları [34].

ESA mimarisi görüntü işleme uygulamalarında kullanışlıdır. İlk ESA, YYS LeCun [34] tarafından oluşturulmuştur; O zamanlar mimari, posta kodu yorumlama gibi el yazısı karakter tanımaya odaklanmıştır. Görüntünün sınıflandırma katmanına gelinceye kadarki özellikleri evrişim katmanlarında çıkarılmaktadır. Erken katmanlar şeklin temel özelliklerini (kenarlar, renk özellikleri gibi), tanı ve sonraki katmanlar bu özellikleri girdinin daha yüksek düzey niteliklerinde yeniden birleştirir. Görüntü, daha sonra girdi görüntüsünden öznelilikler çıkaran bir evrişim katmanına beslenen alıcı alanlara bölünür. Bir sonraki adım en önemli bilgileri korurken (tipik olarak, maksimum havuzlama yoluyla) çıkarılan özelliklerin boyutsallığını azaltan (aşağı örnekleme yoluyla) havuzlamadır. Daha sonra, tamamen bağlantılı çok katmanlı bir algılayıcıyı besleyen başka bir evrişim ve havuzlama aşaması gerçekleştirilir. Bu ağın son çıktı katmanı, görüntünün özelliklerini tanımlayan bir dizi düğümdür. Ağın eğitimi geri yayılım yöntemi kullanılarak yapılmaktadır [103] (Şekil 2.8).

2.4.2. Transfer Öğrenme ve ESA İlişkisi

Transfer öğrenimi, bir problemi çözerek elde edilen bilgiyi ilgili ancak farklı bir problemi çözmek için kullanan bir makine öğrenme tekniğidir. ESA transfer öğreniminin temel fikri, farklı uygulama alanlarında geniş bir görüntü veri seti üzerinde eğitilmiş ağların üstünlüğünden yararlanmaktır [113]. Transfer öğrenme teknikleri mevcut ESA modellerinin eğitebileceği büyüklükte bir veri kümesinin mevcut olmadığı durumlarda başvurulacak bir tekniktir [113, 114].

Tıbbi görüntü çalışmaları için ESA transfer öğrenimini gerçekleştirmek için iki ana yaklaşım bulunmaktadır. Bunlardan ilki, önceden eğitilmiş bir ESA modeli kullanmak ve modelin tamamını tıbbi görüntü veri seti üzerinde yeniden eğitmektir. Bu, tüm ağırlıkların rastgele başlatılmasını ve ardından tam modelin yeni veriler üzerinde uçtan uca eğitilmesini içerir [113]. İkinci yaklaşım, görüntü ile ilgili temel özelliklerin çıkarımının yapıldığı ilk katmanları dondurarak ve yalnızca tıbbi görüntü verilerinin daha yüksek katmanlarına ince ayar yaparak, önceden eğitilmiş bir ESA modelinin özellik çıkarma yeteneklerinden yararlanmaktır [113]. Bu tekniklerle, geliştirilen modelin büyük veri kümelerinde öğrenilen genel görüntü özelliklerinden faydalanmasına ve aynı zamanda üst düzey göreve özgü özelliklerin tıbbi alana uyarlanmasına olanak tanınmaktadır.

Tıbbi görüntü sınıflandırması için transfer öğrenmede kullanılan en yaygın ESA algoritmalarından bazıları Very Deep Convolutinal Networks (VDS), Residual Networks (ResNet), Inception ve U-Net'tir. VDS ve ResNet, tıbbi görüntüleme de dahil olmak üzere çok çeşitli görüntü sınıflandırma görevlerinde etkileyici performans sergileyen, önceden eğitilmiş popüler modellerdir [115, 116]. Inception-V3 ve Inception-ResNet-V2 gibi Inception modelleri, çok ölçekli özellikleri etkili bir şekilde yakalama yeteneklerinden dolayı tıbbi uygulamalarda transfer öğrenimi için de yaygın olarak kullanılmaktadır [117]. Ek olarak, orijinal olarak biyomedikal görüntü segmentasyonu için geliştirilen U-Net mimarisi, tıbbi görüntü sınıflandırma görevlerinde transfer öğrenmeye başarıyla uyarlanmıştır [118].

Araştırmacılar, önceden eğitilmiş bu ESA modellerinin özellik çıkarma yeteneklerinden yararlanarak, sınırlı eğitim verileriyle bile tıbbi görüntü sınıflandırmasında umut verici sonuçlar elde edebilirler. Bu ESA algoritmaları, mevcut tıbbi görüntü veri kümesinin nispeten küçük olduğu senaryolarda özellikle kullanışlıdır çünkü transfer öğrenme

yaklaşımı, modelin büyük ölçekli genel görüntü veri kümeleri üzerinde eğitimden elde edilen bilgilerden faydalanmasına olanak tanır [119, 120]. Bu, modelin Performansını ve genelleme yeteneğini önemli ölçüde geliştirerek onu tıbbi görüntü analizi ve teşhisi için değerli bir araç haline getirebilir.

2.5. Makine Öğrenmesi ve Derin Öğrenme Arasındaki Farklar

ML ve DL modelleri özellikle yapay zekâ alanında sıkça kullanılan iki ana kavramdır. Bu bölümde, bu iki yöntemin çeşitli başlıklarda karşılaştırılmasını yaparak, avantaj ve dezavantajların incelenecektir. İncelememizde, eğitim kümesi gereksinimleri, özellik çıkarımı süreçleri, model kompleksliği, açıklanabilirlik, performans, donanım gereksinimleri, hesaplama maliyeti ve gerçek zamanlı uygulamalardaki kullanımları detaylandırılacaktır (Tablo 2.1).

Tablo 2.1 Makine öğrenmesi ve Derin Öğrenme arasındaki farkların sistem özellikleri ve parametrelerine göre karşılaştırılması

Karşılaştırma Konusu	Makine Öğrenmesi	Derin Öğrenme
Eğitim Kümesi Gereksinimi	Geniş	Çok Geniş
Özellik Çıkarımı	Uzman eşliğinde manuel olarak	Otomatik (Özellikleri kendi çıkarmaktadır)
Model Kompleksliği	Az kompleks	Kompleks
Açıklanabilirliği	Yüksek (İstatistik temelli)	Düşük (Kapalı-Kutu özellikli)
Performans	Ortalama	Yüksek
Donanım Gereksinimi	Düşük (CPU yeterli olmakta)	Yüksek (Çoğunlukla GPU gerekmekte)
Hesaplama Maliyeti	Düşük	Yüksek
Gerçek Zamanlı Uygulamalarda kullanılabilirliği	Genellikle uygun	Kompleksliğinin fazla olması nedeni ile daha az uygun

2.5.1. Eğitim Kümesi Gereksinimi

ML algoritmaları, genellikle daha küçük veri kümeleri ile yeterince verimli çalışabilir. Klasik algoritmalar, veri üzerinde belirli özellikler seçerek eğitilir ve büyük veri kümelerine ihtiyaç duymaz [100, 102]. Lojistik regresyon, karar ağaçları ve destek vektör makineleri (SVM) gibi algoritmalar, daha az sayıda veriden anlamlı sonuçlar çıkartabilir[112, 121]..

DL, büyük miktarda veriye dayalı öğrenme süreci ile tanınır. Sinir ağlarının büyük veri kümelerinden anlamlı özellikler çıkartabilmesi için büyük hacimli, çeşitli ve zengin veri setlerine ihtiyaç vardır [103, 122]. Örneğin, görüntü ve ses tanıma gibi karmaşık görevlerde DL'nin etkinliği, sadece büyük veri kümeleri kullanıldığında ortaya çıkar. Bu durum, özellikle internet devleri tarafından kullanılmakta olup, büyük çaplı verilerin toplanabileceği endüstrilerde avantaj sağlamaktadır. Özellikle ESA, büyük veri kümeleri olmadan yeterince iyi sonuçlar üretmez ve genellikle milyonlarca veri noktasına ihtiyaç duyar [103, 122]. Bu yüzden DL, büyük ölçekli projeler ve büyük veri kümeleri ile ilişkiliyken, ML daha küçük veri kümelerinde etkili olabilir.

2.5.2. Özellik Çıkarımı

Makine öğrenmesi yöntemleri, özellik çıkarımı sürecinde insan müdahalesine dayanmaktadır. Özellik mühendisliği olarak bilinen bu süreç, verinin manuel olarak incelenmesini ve anlamlı özelliklerin seçilmesini içerir [123, 124]. Örneğin, bir kredi risk modelinde, gelirin veya daha önceki ödeme geçmişinin manuel olarak modele dahil edilmesi, modelin başarısını doğrudan etkiler. Bu süreç uzman bilgi birikimi gerektirdiği gibi, aynı zamanda zaman alıcı olabilir.

Buna karşın derin öğrenme modelleri, veriden özellik çıkartma işlemini otomatik hale getirmiştir. Özellikle sinir ağları, farklı katmanlar aracılığıyla ham veriden soyut düzeyde anlamlı bilgiler çıkarabilir [125]. Bu, DL'nin en güçlü yönlerinden biri olup, özellik mühendisliği gerektirmeden verinin doğrudan kullanılmasına olanak tanır. DL modelleri, görüntü işleme ve doğal dil işleme gibi görevlerde başarılıdır çünkü ham veriden yüksek düzeyde özellikleri otomatik olarak çıkarabilir [116].

2.5.3. Model Kompleksliği

ML algoritmaları, daha basit ve anlaşılır modellere sahiptir. Lojistik regresyon, karar ağaçları gibi klasik makine öğrenmesi yöntemleri, genellikle az sayıda parametre içerir ve bu nedenle modellenmesi ve anlaşılması daha kolaydır [126]. Bu basitlik, hesaplama açısından daha az maliyetli olduğu gibi daha küçük ölçekli problemler için de uygundur.

DL ise, çok katmanlı sinir ağları ile daha kompleks yapılara sahiptir. Özellikle derin sinir ağları (DSA) ve ESA, yüzlerce hatta binlerce katmana ve milyonlarca parametreye sahip olabilir [127]. DL, doğrusal olmayan ilişkileri modelleme yeteneği ile karmaşık

problemlerde daha güçlüdür, ancak bu yüksek model karmaşıklığı, eğitimi zorlaştırır ve açıklanabilirliği kısıtlar.

2.5.4. Açıklanabilirlik

ML algoritmaları, genellikle daha şeffaf ve anlaşılır oldukları için açıklanabilirlik açısından avantajlıdır [128]. Karar ağaçları ve lojistik regresyon gibi yöntemler, hangi özelliklerin model kararlarını nasıl etkilediğini net bir şekilde gösterir. Bu durum, tıp ve hukuk gibi güven gerektiren alanlarda önemli bir tercih sebebidir. Kararları anlaşılır modeller, düzenleyici kurumlar tarafından da daha kolay kabul edilir.

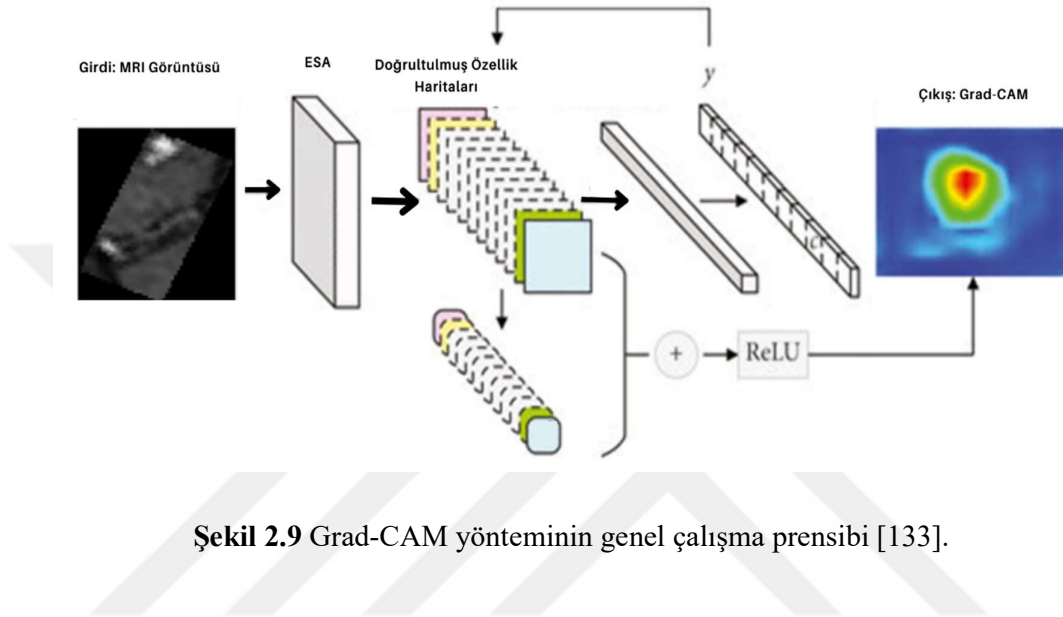
Derin sinir ağlarının karmaşık yapısı, bu modellerin yorumlanabilirliği ve şeffaflığı konusunda endişelerin artmasına yol açmış ve bu sorun genellikle "kara kutu" problemi olarak adlandırılmıştır. Milyonlarca parametreye sahip olan bu modellerin nasıl çalıştığını tam olarak anlamak zordur [129]. Özellikle sinir ağlarının iç dinamikleri ve çıktılarının nedenleri hakkında net bir açıklama yapılamaması, güvenlik açısından risk oluşturabilir. Dolayısıyla, DL'nin açıklanabilirliği üzerine yapılan çalışmalar hızla artmaktadır ancak hala ML kadar açıklanabilir değildir.

2.5.4.1. Polonyi Paradoksu ve GRAD-CAM Yöntemi

YZ ve ML'nin sürekli gelişen alanında, Polonyi paradoksu ve DL'nin kara kutu yapısı, hem araştırmacıların hem de uygulayıcıların dikkatini çeken, birbiriyle ilişkili ve ilgi çekici iki olgu olarak öne çıkmaktadır [103, 130]. Polonyi paradoksu, Macar-İngiliz fizikçi ve kimyager George Polonyi'nin adını taşıyan ve makine öğrenimi modellerinin öngörü gücü ile yorumlanabilirliği arasındaki içsel gerilimi vurgulayan bir kavramsal çerçevedir [126]. Özellikle derin öğrenme modelleri, yüksek öngörü Performansı sağlamada başarılı olmuş; ancak bu başarı, genellikle yorumlanabilirlikten ödün verilmesi pahasına elde edilmiştir [126]. Bölüm 3.5'de de bahsedildiği üzere DL modellerinin kara kutu niteliği, bu modellerin tahminlerini nasıl yaptığını anlamayı zorlaştıran önemli bir sorundur. Bu modellerin şeffaf olmaması, karar verme sürecinin anlaşılmasının kritik olduğu tıbbi teşhis veya risk değerlendirmesi gibi alanlarda önemli bir dezavantaj oluşturabilmektedir [131].

Bu sorunun üstesinden gelmek için, yorumlanabilir ve açıklanabilir YZ yöntemlerinin geliştirilmesine yönelik artan bir araştırma literatürü mevcuttur. Bu yöntemlerden biri olan Grad-CAM, hedef sınıfın son evrişim katmanına akan gradyanları kullanarak, sınıfın tahmini

için önemli bölgeleri vurgulayan kaba bir lokalizasyon haritası sunmaktadır (Şekil 2.9) [132, 133]. Bu ve benzeri yaklaşımlar, kara kutunun içini aydınlatarak, özellikle kararların gerekçesinin büyük önem taşıdığı kritik uygulamalarda derin öğrenme modellerinin güvenilirliğini ve kullanıcı güvenini artırma potansiyeline sahiptir.



Şekil 2.9 Grad-CAM yönteminin genel çalışma prensibi [133].

2.5.5. Performans

Büyük ve karmaşık veri setlerinde DL'nin performansı, genellikle ML'ye kıyasla daha üstündür. Derin sinir ağıları, veriden otomatik olarak öğrenerek daha yüksek doğruluk oranlarına ulaşabilir [129]. Bu avantaj, özellikle görüntü ve dil işleme gibi veri yoğun alanlarda DL'nin kullanılmasını teşvik eder. ML modelleri ise küçük ve yapılandırılmış veri kümelerinde daha verimli çalışır ve genellikle düşük veri miktarında DL'ye kıyasla daha iyi Performans gösterebilir [134].

2.5.6. Donanım Gereksinimi

ML algoritmaları, daha basit hesaplama yapıları ve genellikle merkezi hesaplama ünitesi (CPU) tabanlı donanımlarla çalıştırılabilir [135]. Düşük maliyetli bilgisayar sistemleri veya standart donanımlar, ML modellerinin eğitimi için yeterli olabilir.

DL modelleri, yüksek hesaplama gücü gerektirir ve özellikle grafik işlemcileri (GPU) veya tensör işlemcileri (TPU) gibi donanımlar kullanılarak eğitilir [136]. Bu, derin sinir ağlarının

eđitimi sırasında b y k veri k melerinin iřlenebilmesi ve milyonlarca parametrenin en iyileme iřlemine tabii tutulması i in gereklidir. B y k řirketler veya arařtırma laboratuvarları, bu t r donanımlara sahip oldukları i in DL'yi tercih ederken, k  k  l ekli iřletmeler i in bu t r donanım gereksinimleri bir sınırlayıcı fakt r olabilir.

2.5.7. Hesaplama Maliyeti

ML algoritmaları, d ř k hesaplama maliyeti ile bilinir. Bu modellerin eđitimi ve  ıkarımı, DL'ye kıyasla daha hızlıdır ve genellikle daha az kaynak gerektirir [137]. K  k  l ekli projelerde ve veri analizinde bu durum ML'nin tercih edilmesine neden olur.

DL'nin hesaplama maliyeti y ksektir. B y k veri k melerinin iřlenmesi,  ok katmanlı sinir ađlarının eđitimi i in gereken GPU ve TPU donanımlarının maliyeti y ksektir [138]. Ayrıca, eđitimin tamamlanması uzun s reabilir ve daha fazla enerji harcayabilir. Bu nedenle, DL'nin kullanımı y ksek b t eli projeler ve b y k verilerle sınırlıdır.

2.5.8. Ger ek Zamanlı Uygulamalarda Kullanılabilirlik

ML algoritmaları, hızlı sonu  verebilme yeteneđi ve d ř k donanım gereksinimleri sayesinde ger ek zamanlı uygulamalarda yaygın olarak kullanılır [139].  zellikle d ř k gecikmeli ve y ksek hız gerektiren g revlerde ML'nin hızlı tepki verme yeteneđi olduk a deđerlidir.

DL ise daha karmařık yapıları ve donanım gereksinimleri nedeniyle ger ek zamanlı uygulamalarda daha sınırlıdır [140]. Ancak, son yıllarda DL'nin ger ek zamanlı uygulamalarda kullanılabilirliđini artırmak i in GPU hızlandırma ve optimizasyon teknikleri geliřtirilmiřtir.  rneđin,  nceden eđitilmiř modellerin kullanımıyla DL, bazı ger ek zamanlı uygulamalarda bařarılı olabilir [141].

3. ANKİLOZAN SPONDİLİT TESPİTİNE YÖNELİK YAPAY ZEKÂ TABANLI KARAR DESTEK SİSTEMLERİN TASARLANMASI

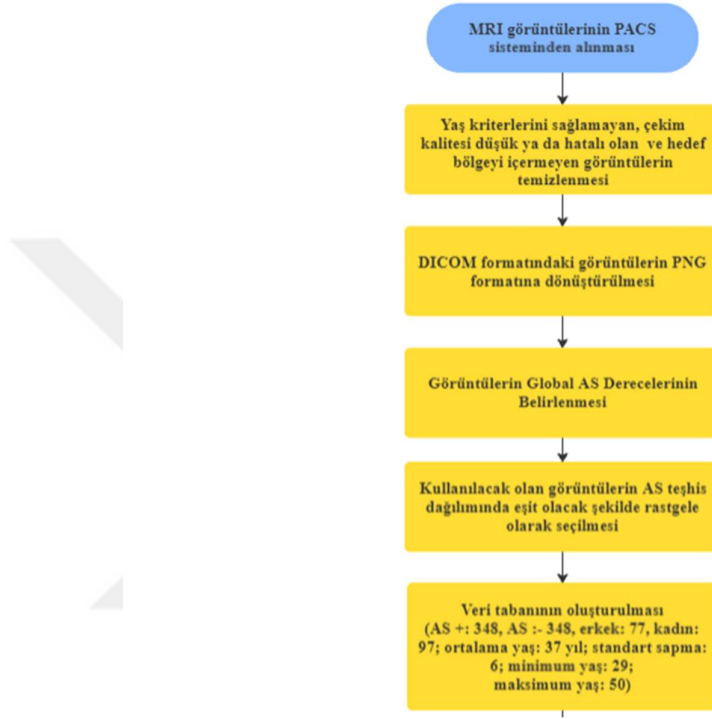
Bu bölümde, öncelikle veri seti ile ilgili bilgilere ve ön hazırlık aşamalarına yer verilmiştir. Ardından AS teşhisine yönelik geliştirilen üç YZ tabanlı modelin özellikleri ve tasarım süreçleri detaylandırılmıştır. İlk model GLCM oluşturulması ve bu matrislerden elde edilen 21 farklı Haralick özelliğinin kullanılmasıyla makine öğrenmesi tabanlı bir sistem olarak geliştirilmiştir. İkinci modelde, GLCM’lerden hesaplanan sayısal Haralick özellikleri, DL modeli ile kullanılarak görsel veri yerine sayısal veriler üzerinden AS sınıflandırması gerçekleştirilmiştir. Üçüncü modelde, önceden eğitilmiş farklı ESA modelleri kullanılmış ve bu modeller, farklı hiper parametreler denenerek eğitilmiş, böylece en iyi sınıflandırma performansına ulaşılması sağlanmıştır. En iyi sınıflandırma performansına sahip ESA modelleri seçilmiş ve çoğunluk oylama sistemi kullanılarak sınıflandırma işlemi gerçekleştirilmiştir.

3.1. Görüntü Toplama ve Veri Seti

Bu çalışmada, *Adnan Menderes Üniversitesi Tıp Fakültesi Girişimsel Olmayan Klinik Araştırmalar Etik Kurulu*’nun 08.02.2021 tarihli ve 2021/24 numaralı onayı ile elde edilen STIR MRI görüntüleri kullanılmıştır. 2015-2020 yılları arasında, *Adnan Menderes Üniversitesi Tıp Fakültesi Romatoloji ve Fiziksel Tıp ve Rehabilitasyon Anabilim Dalları*’na bel ağrısı şikayetiyle başvuran ve inflamatuvar hastalık ön tanısıyla Radyoloji Anabilim Dalı’na yönlendirilen olgulardan elde edilen Sakroiliak MR tetkikine ait arşiv görüntüleri incelenmiştir. Görüntüler, hastanenin PACS (Görüntü Saklama ve İletişim Sistemleri) sisteminden, kas-iskelet sistemi görüntüleme alanında sırasıyla 22 ve 13 yıllık deneyime sahip iki radyolog gözetiminde retrospektif olarak alınmıştır. Her deneğe ait görüntüler, STIR sekansı ile çekilen sakroiliak eklemleri içeren DICOM (Tıpta Dijital Görüntüleme ve İletişim) formatında elde edilmiş ve hasta bilgileri anonimleştirilerek global AS derecelerine göre sınıflandırılarak saklanmıştır (Şekil 3.1).

Verilerin kalite kontrol sürecinde, STIR sekansı ile çekim yapılmamış, PACS sisteminde hatalı başlıklar altında saklanmış, hedef bölgeyi içermeyen veya çekim hatası bulunan görüntüler çalışma kapsamından çıkarılmıştır. Ayrıca, AS’nin genellikle 50 yaş altındaki bireylerde görüldüğü göz önünde bulundurularak, 50 yaş üzerindeki deneklere ait görüntüler de analiz dışında bırakılmıştır. Kalite kontrol işlemi sonrasında görüntüler MATLAB

ortamında DICOM formatından gri tonlamalı PNG (Taşınabilir Ağ Grafiği- Portable Network Graphics) formatına dönüştürülmüştür.



Şekil 3.1: Veri seti oluşturma adımları iş akış diyagramı

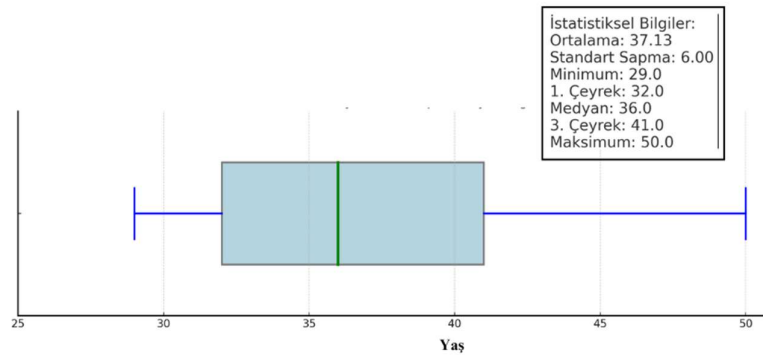
Deneyimli iki hekim tarafından bağımsız olarak incelenen STIR sekans MRI görüntülerinde, AS tespiti için kritik olan subkondral kemik ödemi ve subkondral osteit bulguları değerlendirilmiş ve görüntüler global AS derecelendirme yöntemine göre sınıflandırılmıştır. Derecelendirme işlemi gerçekleştirilmeden kişisel bilgiler anonim hale getirilmiştir. Bu sistem kapsamında:

- Global AS Derecesi 0: Hedef doku bozukluğu olmayan görüntüler (90 hasta).
- Global AS Derecesi 1: Subkondral kemik ödemi ve subkondral osteit oluşumlarının başlangıç aşamasında olduğu görüntüler (35 hasta)
- Global AS Derecesi 2: Doku bozukluğunun daha belirgin olduğu görüntüler (31 hasta).

- Global AS Derecesi 3: Hastalığın ilerlemiş olduğu görüntüler (10 hasta).
- Global AS Derecesi 4: Geniş bozulmaların görüldüğü görüntüler (8 hasta).
- Global AS Derecesi 5: Ciddi bozulukların olduğu görüntüler (2 hasta).
- Global AS Derecesi 6: En ileri düzeyde bozulmalara sahip görüntüler (1 hasta).

Her deneğe ait görüntü sayısı, modelin doğru sınıflandırma yapabilmesi için incelenmiştir. Çalışma sırasında, birçok hastaya ait birden fazla görüntü diliminin bulunduğu tespit edilmiştir. YZ modellerinin doğru bir şekilde sınıflandırma yapabilmesi için kullanılabilir en düşük dilim sayısı esas alınarak, her hastadan en fazla dört görüntü kullanılmıştır. Dört görüntüden fazla dilime sahip olan denekler için görüntüler rastgele seçilmiş ve fazla olanlar analiz dışında bırakılmıştır. Ayrıca, AS'nin var ya da yok durumunu dengelemek adına, her iki durum için eşit sayıda görüntü seçilmesine dikkat edilmiş ve bu görüntüler de rastgele bir şekilde belirlenmiştir. Bu yaklaşım, veri setinde tutarlılık sağlamak ve modelin gerçek dünya senaryolarına uyumunu artırmak amacıyla uygulanmıştır.

Veri temizliği işlemleri sonrasında çalışmaya toplam 174 katılımcı dahil edilmiştir (AS +: 87, AS -: 87; erkek: 77 kişi, kadın: 97 kişi; ortalama yaş: 37 yıl; standart sapma: 6; minimum yaş: 29; maksimum yaş: 50) (Şekil 3.2). Her bir katılımcıya ait 4 adet görüntü kullanılarak toplamda 696 STIR MR görüntüsü (AS +: 348, AS -: 348) elde edilmiş ve 3 farklı YZ tabanlı modelin eğitimi için kullanılmıştır.



Şekil 3.2: Veri setindeki katılımcıların yaş bilgilerinin istatistiki olarak dağılımı

AS tanısı global derecelendirme yöntemine göre değerlendirilen 174 katılımcının sağ ve sol eklemlerine ait AS global inflamasyon dereceleri Tablo 3.1’de sunulmuştur.

Tablo 3.1 AS değ erlendirmesi i in incelenen g r nt lerin global AS dereceleri

Sakroiliak Eklem (Sol/Sağ)	Global AS Derecesi	Toplam Denek Sayısı	ML Modeli i�in Elde Edilen STIR MR G�r�nt�lerinin Sayısı	Toplam STIR MRI G�r�nt�leri
Sol	0	44	176	696
Sağ	0	43	172	
Sol	1	17	68	
Sağ	1	17	68	
Sol	2	21	84	
Sağ	2	10	40	
Sol	3+	12	48	
Sağ	3+	10	40	

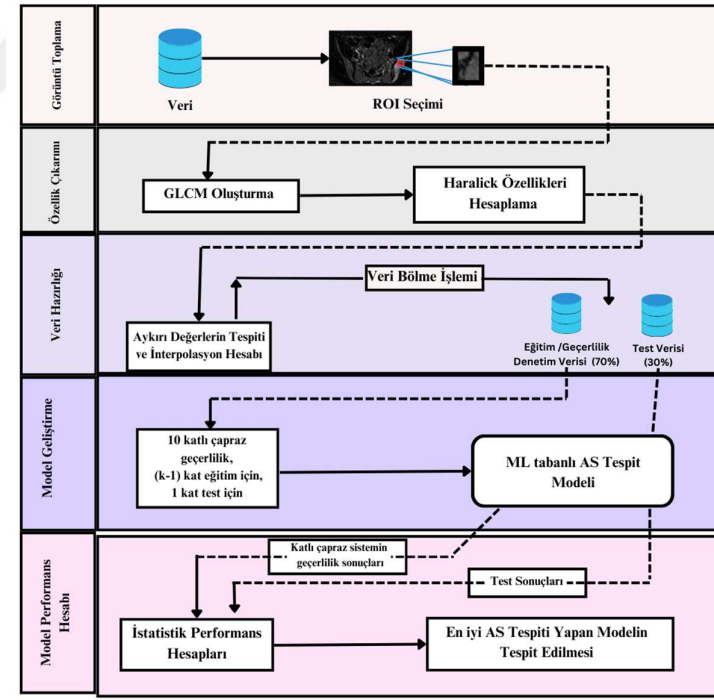
Yukarıda anlatılan adımlar ile bu tez kapsamında  nerilen YZ tabanlı sistemlerin eđitim, dođrulama ve test ařamaları i in kullanılacak olan ana veri seti oluřturulmuřtur. Modellerin eđitim stratejileri ve girdilere y nelik ihtiya larına g re, ana veri setinden ilgili modelin eđitimi i in yapılan d zenlemeler ve hesaplamalar, her y nteme ait bařlıklar altında detaylı bir řekilde a ıklanmıřtır.

AS derecelendirme s recinin ardından, kiřisel bilgiler anonim hale getirilmiřtir. AS tespiti yapabilmek i in, subkondral kemik  demi ve subkondral osteit  zerine odaklanmak kritik  nem tařımaktadır. Bu nedenle, AS'nin erken evredeki teřhisinde, orijinal MR taramalarından, AS ile iliřkili patolojik deđiřikliklerin y ksek olasılıkla bulunduđu belirli ROI'lerin tanımlanması gereklidir (řekil 3.3). Bu  alıřmada, ROI'ler 40*50 (en*boy) piksel geniřliđinde ve 50 piksel y ksekliđinde se ilip, bu g r nt ler herhangi bir  n g r nt  iyileřtirmesi iřleme tabi tutulmamıřtır.

3.2. Y ntem 1: Klasik ML Algoritmaları Kullanılarak Geliřtirilen AS Tespit Sistemi

Bu kısımda ML algoritmaları ile oluřturulmuř AS tespit sisteminin hangi ařamalar ile geliřtirildiđi ayrıntılı bir řekilde sunulacaktır. AS teřhisinde daha hassas, objektif ve hızlı

sonuçlar sunarak klinik süreçlere önemli katkılar sağlaması için, SVM, k-NN, TOA ve YSA yöntemleri bu çalışmanın kapsamında kullanılmıştır. Her bir farklı ML tekniği veri setinin farklı yöntemler ile yorumlanmasını sağlar. Örneğin SVM, veri kümelerini yüksek boyutlu bir uzaya haritalayarak sınıflar arasındaki en uygun ayırım sınırını belirlemekte ve karmaşık veri kümelerinde sınıflandırma doğruluğunu artırmaktadır. k-NN algoritması ise sınıflandırma işlemini, sınıflandırılacak örneğin en yakın komşularıyla olan benzerliklerine dayanarak gerçekleştirir; bu da, AS'nin farklı evrelerindeki hastaların doğru bir şekilde sınıflandırılmasına olanak tanır. TOA ise farklı yöntemlerin bir arada kullanılmasıyla modelin genelleme yeteneğini geliştirir ve teşhis sürecindeki hata payını azaltır. YSA insan beyninin bilgi işleme süreçlerinden ilham alarak, öğrenme yeteneğine sahip güçlü bir model sunar ve AS teşhisinde karmaşık ilişkileri öğrenerek hastalığın farklı semptomlarını ve klinik göstergelerini doğru bir şekilde tanımlamaktadır. Şekil 3.3'de yöntem 1 için tasarlama aşamaları sırasıyla sunulmuştur.



Şekil 3.3: Yöntem 1 iş akış diyagramı.

Veri seti temizliği tamamlandıktan sonra, klasik makine öğrenimi algoritmaları ile oluşturulacak modeller için yeterli örneklem sayısının olup olmadığı, “faktör 10 kuralı” yönergesine göre değerlendirilmiştir [142]. Bu kurala göre, ML modellerinin tespit

yapabilmesi için, her etiket kategorisi (C) için gerekli örnek sayısı (s), özellik sayısının (f) on katı olmalıdır (Denklem 3.1) [142]:

$$s \geq f \times 10 \times C \quad (3.1)$$

Bu nedenle, 21 Haralick özelliği ve iki kategori (AS ve AS olmayan) ile sınıflandırma yapacak bir makine öğrenimi modelinin, faktör 10 kuralına göre en az 420 örneğe ihtiyaç duyduğu belirlenmiştir. Önerilen çalışmada, aynı MRI'dan elde edilen 696 MRI görüntüsünün kullanılmasıyla bu gereklilik karşılanmış ve faktör 10 kuralı sağlanmıştır (Tablo 3.1).

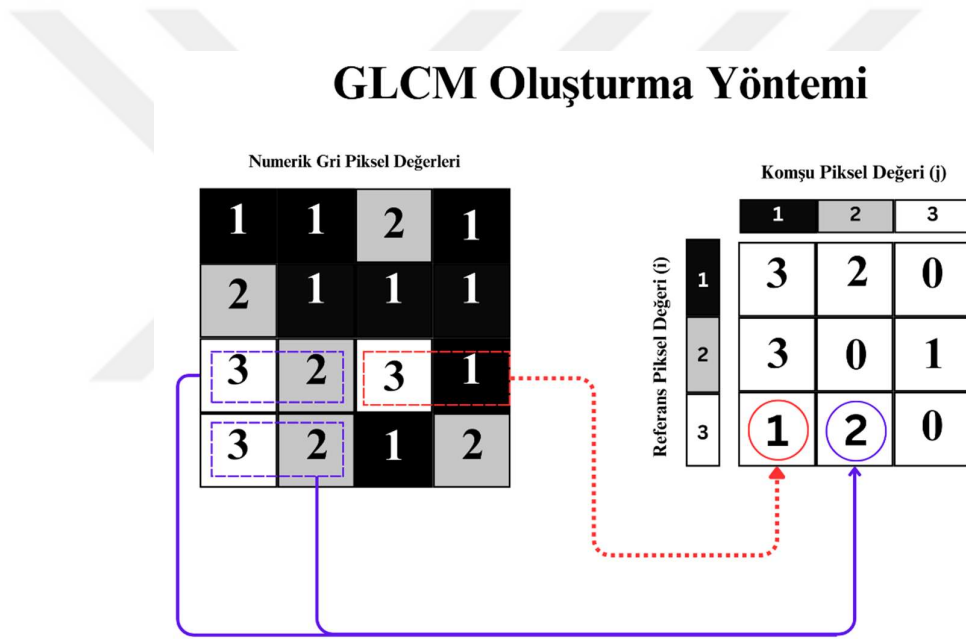
3.1.1. Özellik Çıkarımı: Gri-Seviye Eş-Olay Matrisi ve Haralick Doku Özellikleri

Gri Seviye Eş Oluşum Matrisi (GLCM), bir görüntüde özdeş gri seviyelerine sahip piksel çiftlerinin sıklığını ölçmek için kullanılan ve aynı zamanda bunların mekansal ilişkilerini de hesaba katan istatistiksel bir yöntemdir. GLCM ve doku özellikleri, basit ve sezgisel yorumlamaları nedeniyle sıklıkla kullanılır ve örneğin tıbbi görüntü analizinde, örneğin tümör sınıflandırması [143], X-ışını mamografisi [144], kanser için karaciğer MR görüntü analizi [145], [146], [147], [148], aritmi tanısı [149], osteosarkom tanısı [150], akciğer nodüllerinin malignite seviyesi sınıflandırması [151], Alzheimer hastalığı sınıflandırması [152], Lewy cisimcik hastalığı [153], pankreatit [154], servikal kanser tanısı [155], mamogram sınıflandırması [156] glokom evre tespiti [157] ve radyomiklerde [158] etkili bir şekilde uygulanmıştır.

GLCM, bir görüntünün tamamında veya belirli bölgelerde çeşitli gri seviyelerinin oluşumunu temsil eder [159]. Her yakın piksel bağlantısının benzersiz bir GLCM'si vardır. $n \times m$ boyutlarındaki bir gri tonlamalı görüntüde, I , görüntüdeki her piksel $g_{i,j}$, $[0, 1, \dots, N_g - 1]$ aralığında bir değere sahiptir, burada N_g , benzersiz gri seviyelerinin toplam sayısını temsil eder (Şekil 3.3). Görüntüyü I harfi ile temsil ettiğimizde, I boyunca birlikte görünen piksel değerlerinin orantılı oluşumlarının temsili şu şekilde adlandırılır (Denklem 3.2):

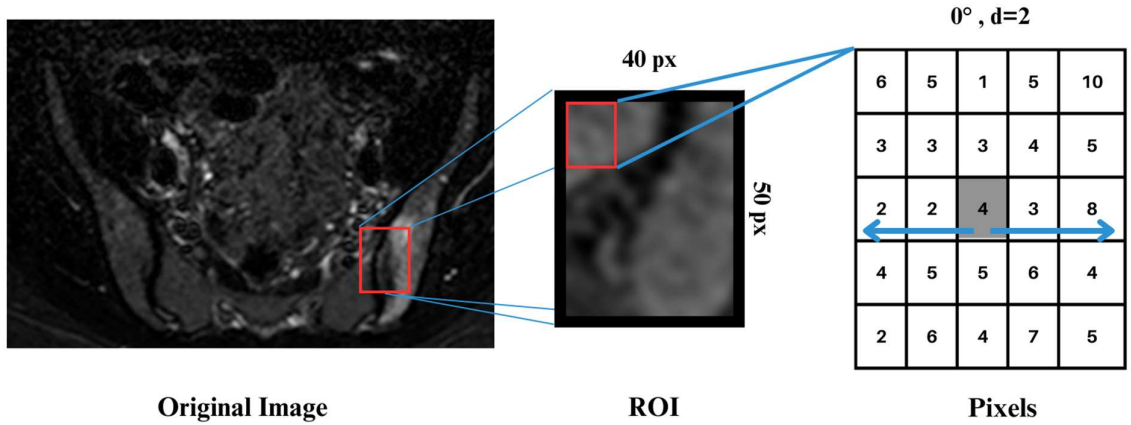
$$X(i,j) = \sum_{x=1}^n \sum_{y=1}^m \begin{cases} 1, & \text{if } I(n,m) = i \text{ ve } I(m+d_x, (k+d_y)) = j \text{ ise} \\ 0, & \text{diğer durumlar için} \end{cases} \quad (3.2)$$

$X(i,j)$ elemanı, referans pikselden d uzaklıkta, gri seviyesi i 'nin belirli bir yönde olduğu gri seviyesi j 'nin frekansını temsil eder [158], [159]. Tez çalışmasında GLCM, Haralick özelliklerini hesaplamak, doku özelliklerini sayısal değerlere dönüştürmek ve ML modellerinin eğitimi için kullanılmıştır (Şekil 3.4).



Şekil 3.4: 4x4'lük bir görüntüde GLCM oluşturulurken her piksel ile onun komşusu arasındaki ilişki

$X(i,j)$ elemanı, referans pikselden d uzaklıkta, gri seviyesi i 'nin belirli bir yönde olduğu gri seviyesi j 'nin frekansını temsil eder [158], [159]. Tez çalışmasında GLCM, Haralick özelliklerini hesaplamak, doku özelliklerini sayısal değerlere dönüştürmek ve ML modellerinin eğitimi için kullanılmıştır (Şekil 3.5).



Şekil 3.5: Orijinal görüntüden ROI'lerin seçilmesi ve GLCM'nin $\theta = 0^\circ$ and $d=2$ parametreleri için hesaplanması.

Görüntülenen bir nesnenin ölçülebilir bir özelliği olarak doku, onun temel niteliklerini kapsar [157], [158], [159], [160]. Görüntü işleme ve makine görüşündeki dokusal görüntüler, piksel aydınlatmasının yoğunluğunun belirli bir dağılım ve dağılım modelinin görüntü boyunca sırayla tekrarlandığı görüntüleri ifade eder [161].

Yüzey özelliklerini temsil eden GLCM [159, 162, 163], $\theta = 0^\circ$ and $d=2$ parametreleri ile Haralick özelliklerinin hesaplanabilmesi için oluşturuldu (Şekil 3.5) (Denklem 3.2). GLCM hesaplamalarından elde edilen bu sayısal çıktılar ML algoritması için girdi görevi görecektir. Sunulan çalışmada kullanılan klasik ML algoritma tabanlı modelleri geliştirmede kullanılan 21 Haralick özelliği şunlardır: Oto korelasyon (autoc), Kontrast (contr), Korelasyon (corm), Küme Önemi (cprom), Küme Gölgesi (cshad), Benzersizlik (dissi), Enerji (energ), Entropi (entro), Homojenlik Versiyon 1 (homom), Homojenlik Versiyon 2 (homop), Maksimum Olasılık (maxpr), Kareler Toplamı (sosvh), Toplam Ortalama (savgh), Toplam Varyans (svarh), Toplam Entropi (senth), Fark Varyansı (dvarh), Fark Entropisi (denth), Korelasyon Bilgi Ölçütü Versiyon 1 (inf1h), Korelasyon Bilgi Ölçütü Versiyon 2 (inf2h), Normalize Ters Fark (indnc) ve Normalize Ters Fark Momenti (idmnc). Tablo 3.2'de, ilgili Haralick özellik formülleri ve açıklamaları özetlenmiştir.

Tablo 4, $p(i,j)$ 'nin keyfi bir θ ve d ($\theta = 0^\circ$ ve $d=2$) için eş-oluşum matrisi olduğunu göstermektedir. Marjinal satır olasılıkları ve marjinal sütun olasılıkları sırasıyla $p_x(i)$ ve $p_y(j)$ olarak temsil edilmektedir. p_x 'in gri seviye yoğunluğu μ_x olarak ve p_y 'nin gri seviye

yoğunluğu μ_y olarak tanımlanmaktadır. HX , HY ve HXY , p_x , p_y ve $p(i,j)$ 'nin entropisini temsil etmektedir.

3.1.3. Veri Hazırlama ve Model Geliştirme

Veri setindeki her MR görüntüsü için GLCM oluşturulmuş ve ardından modeli eğitmek için Haralick özellikleri hesaplanmıştır. Sakroiliak eklemdaki MR görüntülerindeki parlaklık varyasyonlarının neden olduğu aykırı değerleri ortadan kaldırmak amacıyla, tüm ROI'ler için hesaplanan Haralick özellikleri filtrelenmiştir. Aykırı değerler, veri hazırlık aşamasında her bir özelliğin medyanının dört katından fazla olan değerler belirlenerek tespit edilmiştir. Wahir [160], aykırı değerlerin eksik değerler olarak ele alınmasının tahmin sonuçlarını iyileştirdiğini önermiştir. Ayrıca, aykırı veri çıkarmak, bu verilerde bulunan değerli bilgilerin kaybolma riskini taşımaktadır. Bu nedenle, çalışmamızda aykırı değerler, değerli bilgilerin kaybolmasını önlemek amacıyla eksik değerler olarak değerlendirilmiştir. Bu değerleri silmek yerine, komşu veri noktalarının değerlerine dayanarak eksik değerleri tahmin eden doğrusal ara kestirim hesabı yöntemi kullanılarak doldurulmuştur. Böylece verinin genel eğilimi korunmuştur.

Sonrasında, AS tespit problemi için çeşitli ML algoritmaları (k-NN, SVM ve Ağaç tabanlı TOA) için görüntü etiketleme, veri ön işleme ve eğitim prosedürleri tamamlanmıştır. AS tespit modeli eğitimi ve test işlemleri için Windows 10 işletim sistemi, AMD Ryzen 7 2.90 GHz işlemci, 16 GB DDR4 RAM ve Nvidia GTX 1050 TI grafik kartı bulunan bir PC kullanılmıştır.

Tablo 3.2 GLCM'den türetilen Haralick yüzey özellikleri. Sol sütun özellik adını, orta sütun özelliğin kısa bir açıklamasını ve sağ sütun ilgili özelliği hesaplamak için gereken formülü gösterir.

Hesaplanmış Haralick Özelliği	Açıklama	Formül
Oto korelasyon (autoc) [161]	Yüzeyin inceliği ve kabalığının büyüklüğünün ölçüsü	$\sum_{i=1}^N \sum_{j=1}^N (i, j) p(i, j)$
Kontrast (contr) [162]	Yerel yoğunluk değişiminin ölçüsü	$\sum_{i=1}^N \sum_{j=1}^N (i - j)^2 p(i, j)$
Korelasyon (corm) [162]	GLCM'de gri seviye değerlerinin ilgili voksellerine doğrusal bağımlılığını gösteren 0 (ilişkisiz) ile 1 (mükemmel şekilde ilişkili) arasındaki bir değer	$\sum_{i=1}^N \sum_{j=1}^N \left(\frac{i - \mu_x}{\sigma_x} \right) \left(\frac{j - \mu_y}{\sigma_y} \right) p(i, j)$
Küme Önemi (cprom) [161]	GLCM'nin çarpıklık ve asimetrisinin ölçüsü	$\sum_{i=1}^N \sum_{j=1}^N (i + j - 2\mu)^3 p(i, j)$
Küme Gölgesi (cshad) [161]	GLCM'nin çarpıklığı ve tekdüzelığının ölçüsü	$\sum_{i=1}^N \sum_{j=1}^N (i + j - 2\mu)^4 p(i, j)$
Benzersizlik (dissi) [161]	Komşu pikseller arasındaki yoğunluk farkını ölçüsü	$\sum_{i=1}^N \sum_{j=1}^N i - j p(i, j)$
Enerji (energ) [162]	Görüntünün düzenliliği veya homojenliğinin ölçüsü	$\sum_{i=1}^N \sum_{j=1}^N p(i, j)^2$
Entropi (entro) [161]	Komşuluk yoğunluk değerlerindeki rastgeleliğin ölçüsü	$- \sum_{i=1}^N \sum_{j=1}^N p(i, j) \log p(i, j)$
Homojenlik Versiyon 1 (homom) [162]	GLCM'deki elemanların dağılımının GLCM köşegenine yakınlığının ölçüsü Versiyon 1	$\sum_{i=1}^N \sum_{j=1}^N \frac{p(i, j)^2}{1 + (i - j)^2}$
Homojenlik Versiyon 2 (homop) [161]	GLCM'deki elemanların dağılımının GLCM köşegenine yakınlığının ölçüsü Versiyon 1	$\sqrt{1 - e^{-2 \log_2 p(i, j)}}$
Maksimum Olasılık (maxpr) [161]	En belirgin komşu yoğunluk değerleri çiftinin oluşumları	$\max(p(i, j))$
Kareler Toplamı (sosvh) [162]	Komşu yoğunluk seviyesi çiftlerinin dağılım ölçüsü	$\sum_{i=1}^N \sum_{j=1}^N (i - \mu)^2 p(i, j)$
Toplam Ortalama (savgh) [162]	Daha düşük yoğunluk değerlerine sahip çiftlerin oluşumu ile daha yüksek yoğunluk değerlerine sahip çiftlerin oluşumu arasındaki ilişkinin ölçüsü	$\sum_{i=2}^{2N_g} i p_{x+y}(i)$

Tablo 3.2 GLCM'den türetilen Haralick yüzey özellikleri. Sol sütun özellik adını, orta sütun özelliğin kısa bir açıklamasını ve sağ sütun ilgili özelliği hesaplamak için gereken formülü gösterir. (Devamı)

Hesaplanmış Haralick Özelliği	Açıklama	Formül
Toplam Varyans (svrh) [162]	Ortalamadan daha fazla sapan komşu yoğunluk seviyesi çiftlerine daha yüksek ağırlıklar veren heterojenlik ölçüsü	$\sum_{i=2}^{2N} \left(i - \sum_{i=2}^{2N} i p_{x+y}(i) \right)^2 p_{x+y}(i)$
Toplam Entropi (senth) [162]	Komşuluk yoğunluk değer farklarının toplamı	$- \sum_{i=2}^{2N} p_{x+y}(i) \log\{p_{x+y}(i)\}$
Fark Varyansı (dvarh)[162]	Ortalamadan daha fazla sapan farklı yoğunluk seviyesi çiftlerine daha yüksek ağırlıklar veren heterojenlik ölçüsü	$\sum_{k=0}^{N-1} (k - \mu_{x-y})^2 p_{x-y}(k)$
Fark Entropisi (denth) [162]	Komşuluk yoğunluğu değer farklarındaki rastgelelik/değişkenliğin ölçüsü	$\sum_{k=0}^{N-1} \mu_{x-y} \log_2(\mu_{x-y}(k) + \epsilon)$
Korelasyon Bilgi Ölçütü Versiyon 1 (inf1h) [162]	i ve j'nin olasılık dağılımları arasındaki korelasyon Versiyon 1	$HXY = - \sum_i \sum_j p(i, j) \log(p(i, j))$ $HY = - \sum_{j=1}^N p_y(j) \log(p_y(j))$ $HXY1 = - \sum_i \sum_j p(i, j) \log\{p_x(i)p_y(j)\}$ $inf1h = \frac{HXY - HXY1}{\max\{HX, HY\}}$
Korelasyon Bilgi Ölçütü Versiyon 2 (inf2h) [162]	i ve j'nin olasılık dağılımları arasındaki korelasyon Versiyon 2	$HXY2 = - \sum_i \sum_j p_x(i)p_y(j) \log\{p_x(i)p_y(j)\}$ $inf2h = \sqrt{(1 - \exp[-2(HXY2 - HXY)])}$
Normalize Ters Fark (indnc) [163]	Bir görüntünün yerel homojenliğinin ölçütü.	$C_{ij}(\text{Co-occurrence probability}) = \frac{P_{i,j}}{\sum_{i,j=1}^N P_{i,j}}$ $indnc = \sum \frac{C_{ij}}{1 + i - j }$
Normalize Ters Fark Momenti (idmnc) [163]	Bir görüntünün yerel homojenliğinin başka bir ölçütü.	$\sum \frac{C_{ij}}{1 + (i - j)^2}$

Mevcut görüntü sayısının sınırlı olmasından dolayı, aşırı öğrenmeyi önlemek amacıyla 10 katlı çapraz doğrulama yaklaşımı benimsenmiştir. Bu önleme ek olarak, aynı hastanın farklı görüntülerinin hem eğitim hem de doğrulama setlerinde yer almadığından emin olunarak

olası veri sızıntıları önlenmiştir. Bu önlem alınmadığı takdirde, ML modelleri aynı hastadan gelen görüntüler üzerinde eğitilip test edilebilir ve bu da yapay olarak yüksek tespit performans metriklerine yol açabilir [164]. Buna göre, toplam verinin %70'i eğitim seti ve 10-katlı çapraz doğrulama için (AS +: 220, AS -: 268) ayrılırken, geri kalan %30'u rastgele olarak test setine (AS +: 128, AS -: 80) tahsis edilmiştir. 10-katlı çapraz doğrulama bağlamında, eğitim veri seti on daha küçük alt küme olarak bölünmüştür (Şekil 3.1). Daha sonra, ML modeli bu katların dokuzunda eğitim verisi olarak kullanılarak model eğitilmiş, onuncu kat ise modelin performansını test etmek için kullanılmıştır. Model eğitim aşamasında gerçekleştirilen bu değerlendirme süreci, önerilen çalışmada çapraz doğrulama sonucu olarak kaydedilmiştir. Her bir modelin, çapraz doğrulama sırasında oluşturulması ve doğrulanması sürecinde tespit performansı, en iyi performans gösteren ML modelini belirlemek için bir karışıklık matrisine (CM) kaydedilmiştir (Şekil 3.3).

3.3. Yöntem 2: Haralick Özellikleri ile Eğitilen Derin Öğrenme Modeli

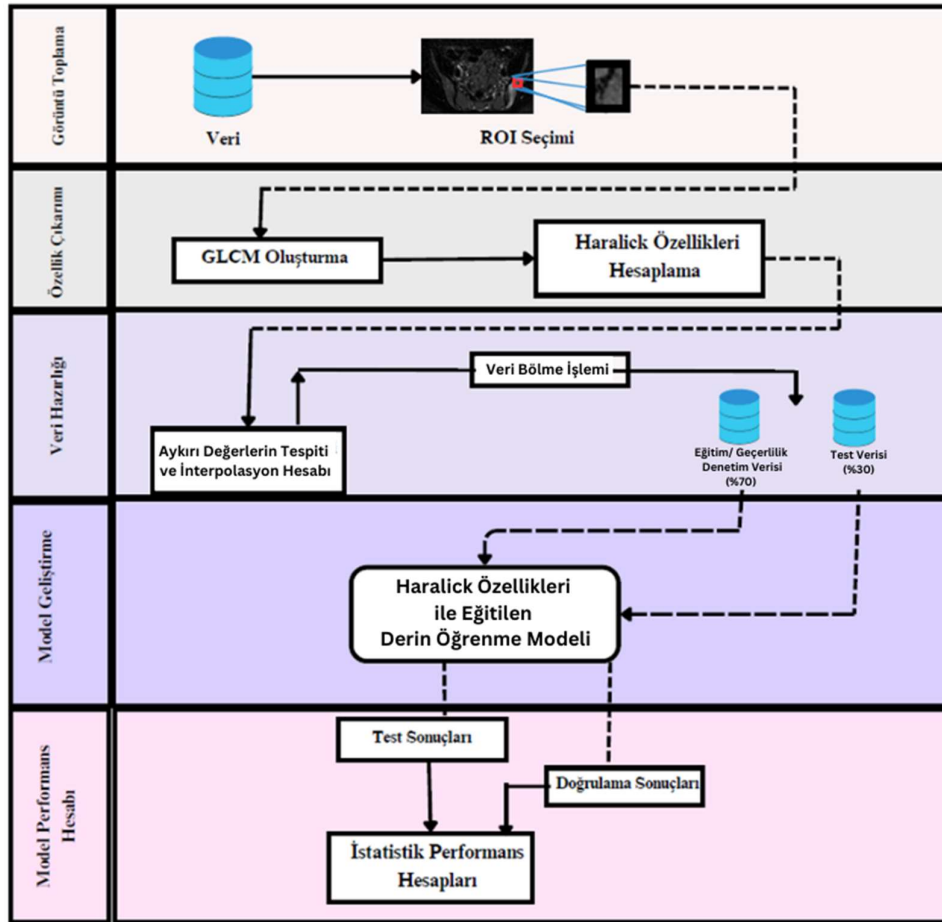
Önerilen bu model, her ne kadar DL tabanlı olsa da literatürde ilk defa AS tespiti için görüntü yerine, görüntüden çıkarılmış sayısal veriler üzerinden eğitilmesi temeline dayanan bir yöntem ile tasarlanmıştır (Şekil 3.6). DL modelleri, geleneksel istatistiksel yöntemlerle kolayca fark edilemeyen karmaşık kalıpları ve ilişkileri ortaya çıkarmak için bu sayısal veri kümeleri üzerinde eğitilebilir [165-167]. DL araştırmalarının büyük bir kısmı görüntü tabanlı girdilere odaklanmışken, sayısal verilere DL tekniklerinin uygulanması da son yıllardan ilgi gören bir yaklaşım halini almıştır [168].

Sayısal veriler için DL'nin kullanılmasının başlıca avantajlarından biri, geleneksel makine öğrenimi algoritmalarıyla hemen fark edilmeyen karmaşık, doğrusal olmayan ilişkileri ortaya çıkarma yeteneğidir. DL modelleri, verilerden hiyerarşik temsiller öğrenme kapasitesine sahip olup, bu sayede karmaşık desenleri yakalayabilir ve anlamlı özellikler çıkarabilirler [167-169]. Bu durum, sayısal verilerin bol olduğu finans, sağlık hizmetleri ve bilimsel araştırma gibi alanlardaki çalışmalara olumlu katkılar sağlamıştır.

Sayısal girdiler için DL tekniklerinin benimsenmesi, genellikle karmaşık yapılar sergileyen büyük ölçekli, yüksek boyutlu veri setlerinin işlenmesi ve analiz edilmesi ihtiyacından kaynaklanmıştır. Çok katmanlı mimarilere sahip DL modelleri, bu tür karmaşık verileri etkili bir şekilde ele alma yeteneklerini kanıtlayarak anomali tespiti, zaman serisi tahmini ve karar verme gibi çeşitli görevlerde geleneksel yöntemleri geride bıraktığı iddia edilmiştir. [167-

169]. Bölüm 3.1’de belirtilen görüntü toplama ve veri kümesi oluşturma adımları, bu yöntemde de uygulanmıştır. Bölüm 3.1.1’de bahsedilen STIR MRI görüntülerinden elde edilen GLCM’ler kullanılarak Haralick özellikleri hesaplanmış ve bu özellikler ile veri seti oluşturulmuştur (Şekil 3.6).

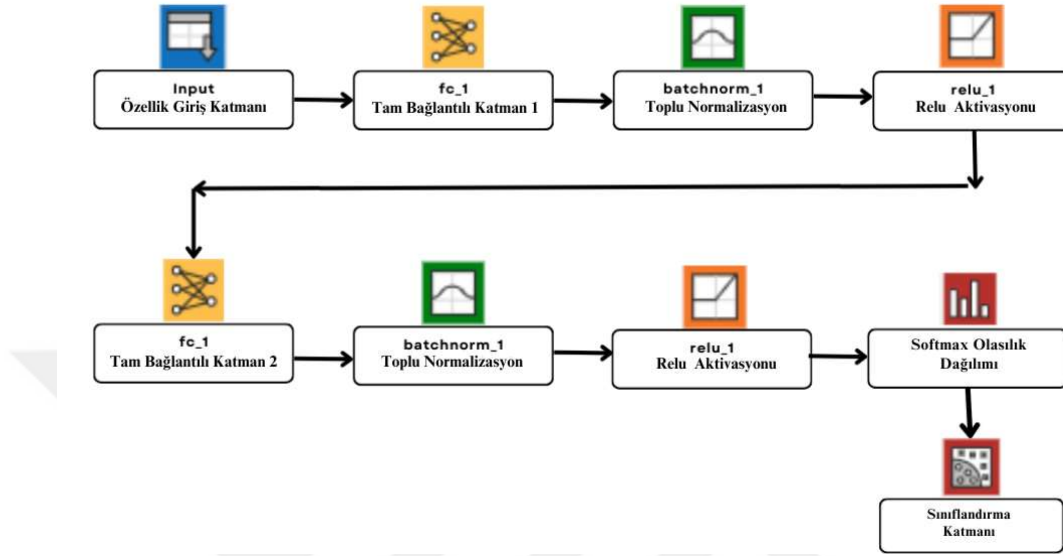
Daha sonra, aykırı veriler tespit edilerek lineer ara kestirim yöntemiyle hesaplanmış ve veri sızıntısını önleyecek şekilde düzenlenmiştir [160]. Bu düzenlemeye göre, toplam verinin %70’i eğitim seti olarak ayrılmıştır (AS +: 220, AS -: 268), geri kalan %30’luk kısmı ise rastgele doğrulama setine (AS +: 128, AS -: 80) tahsis edilmiştir.



Şekil 3.6 Yöntem 2 iş akış diyagramı.

10 katmanlı derin öğrenme ağı, sayısal verileri girdi olarak işleyerek ikili sınıflandırma görevi yürütmektedir (Şekil 3.7). Bu ağın işleyişini daha iyi anlamak için katmanlarını sırasıyla inceleyelim: İlk olarak, "Özellik Giriş Katmanı" ile ağ, girdiyi doğrudan alır. Bunu takiben, "Tam Bağlantılı Katman 1" ve "Tam Bağlantılı Katman 2" girdi verisi ve nöronlar

arasında bağlantılar kurar ve ağırlıklı hesaplamalar yoluyla çıktılar üretir. Bu çıktılar, ağırlıklı kararlılığını artırmak amacıyla "Toplu Normalizasyon" katmanında normalleştirilir.



Şekil 3.7 Sayısal Haralick özelliklerini girdi olarak kullanan AS sınıflandırması için eğitilen DL modelinin mimarisi

Sonrasında, "Relu" aktivasyon fonksiyonu, negatif değerleri sıfıra ayarlayarak seçici olarak nöronları etkinleştirir ve ağırlıklı esnekliğini artırır. Bu aktivasyon işlemi her tam bağlantılı katmandan sonra gerçekleşir. Son olarak, "Tam Bağlantılı Katman 3" ağı çıktılarını iki belirli sınıfa, AS pozitif ve AS negatif olarak sınıflandırır. Bu sınıflandırmalar, "Softmax" katmanında olasılık dağılımına dönüştürülerek sınıflandırma süreci tamamlanır. Bu sistematik yaklaşım, özellik çıkarımı ve doğru sınıflandırmada yetkin bir derin öğrenme modeli oluşturur (Tablo 3.3).

Tablo 3.3 Haralick özellik tabanlı DL modelinin katman analizi.

No	Katman Adı	Katman Tipi	Aktivasyonlar
1	Input	Giriş	21X1
2	fc_1	Tam Bağlantılı Katmanı 1	50X1
3	batchnorm_1	Toplu Normalizasyon Katmanı 1	50X1
4	relu_1	Relu Aktivasyon Fonksiyonu Katmanı 1	50X1
5	fc_2	Tam Bağlantılı Katmanı 2	50X1
6	batchnorm_2	Toplu Normalizasyon Katmanı 2	50X1
7	relu_2	Relu Aktivasyon Fonksiyonu Katmanı 1	50X1
8	fc_3	Tam Bağlantılı Katmanı 3	2X1
9	Softmax	Olasılık Dağılım Katmanı	2X1
10	Classification	Çıkış	2X1

Verilerin derlenmesinin ardından, önerilen DL sistemi çeşitli hiper parametreler kullanılarak eğitilmiştir. Çözümleyici olarak uyarlanabilir moment tahmini (adam), kök ortalama kare yayılımı (rmsprop) ve momentumlu stokastik gradyan inişi (sgdm) kullanılmıştır. Her çözümleyici için üç farklı başlangıç öğrenme oranı (0,001, 0,01 ve 0,1) uygulanmıştır. Ardından, eğitilen modellerden doğrulama verilerini sınıflandırmaları istenmiştir. Bu modellerin sınıflandırma performansları istatistiksel olarak değerlendirilmiş, CM oluşturulmuş, ROC eğrisi ve AUC hesaplanmıştır. En iyi sonuçları veren model performans istatistiklerinin değerlendirilmesi sonucunda seçilmiştir.

3.3. Yöntem 3: Derin Öğrenme Tabanlı Modeller ile Oluşturulan Oylama Sistemi

Tez kapsamında AS tespiti için geliştirilen üçüncü YZ modelimiz, transfer öğrenme ile ESA yapısı kullanılarak geliştirilen DL modelleri ile oluşturulan ÇOA modelidir. Bu sistemin geliştirilmesine ihtiyaç duyulmasının temel nedeni, karmaşık yapılar içeren verilerde tek bir modelin performansının yetersiz kalabileceği ve farklı modellerin güçlü yönlerini birleştirerek daha doğru ve genelleme yeteneği yüksek bir model elde edilebileceği düşüncesidir.

ML'nin sürekli gelişen dünyasında, topluluk öğrenmesi ve derin öğrenmenin entegrasyonu, karmaşık problemleri çözmek ve yeni performans seviyelerine ulaşmak için umut vadeden bir yaklaşım olarak öne çıkmıştır. ÇOA, birden fazla modelin birleştirilerek öngörü gücünün

artırılmasını sağlayan bir tekniktir ve çeşitli alanlarda zayıf öğrencilerin sonuçlarını birleştirerek genelleştirilmiş bir model sunmasıyla etkili bir çözüm olduğunu kanıtlamıştır [170]. Diğer yandan, derin öğrenme, verilerden hiyerarşik özellikler çıkartma ve öğrenme yeteneği ile bilgisayarla görmeden doğal dil işlemeye kadar birçok uygulamayı dönüştürmüştür [171, 172].

Bu iki güçlü paradigmanın sinerjik birleşimi, her ikisinin de güçlü yönlerini kullanarak bireysel sınırlılıklarının üstesinden gelme potansiyeline sahiptir. Özellikle AS tespiti gibi klinik kararların kritik olduğu durumlarda, tek bir modelin karmaşık ve yüksek boyutlu verilerde yeterince iyi performans gösterememesi, toplu öğrenme yaklaşımlarına olan gereksinimi artırmaktadır.

ÇOA temel avantajlarından biri, model çeşitliliğini artırma kabiliyetidir ve bu, genel performansın iyileştirilmesi için kritik bir öneme sahiptir. Farklı model mimarileri ve eğitim stratejilerinden yararlanılarak, her biri kendine özgü güçlü ve zayıf yönleri sahip heterojen bir temel öğrenci kümesi oluşturulabilir. Bu çeşitlilik, modelin genelleme yeteneğini, sağlamlığını ve tek bir modelin gözden kaçırabileceği karmaşık örüntüleri yakalama kapasitesini artırabilir.

DL ise, büyük veri setleri üzerinde öğrenme yapabilme kabiliyetiyle bu sürece önemli bir katkı sağlamaktadır. Transfer öğrenme ve ESA tabanlı derin öğrenme modelleri, verilerdeki karmaşık ve hiyerarşik özellikleri öğrenme kapasitesine sahiptir. Bu güçlü özellik çıkarma yeteneği, topluluk çerçevesi içinde temel öğrenci olarak kullanıldığında, AS tespiti gibi zorlu görevlerde model performansının artmasına olanak tanır [170-173] .

Önerilen ÇOA modeli, AS tanısında daha doğru ve güvenilir sonuçlar elde edebilmek için derin öğrenmenin güçlü yönleri ile toplu öğrenmenin çeşitliliği bir araya getirilerek oluşturulmuştur. Bu yaklaşımla, tek bir modelin sınırlamaları aşılmış ve farklı algoritmaların avantajları bir araya getirilerek daha üstün bir performans elde edilmesi hedeflenmiştir

3.3.1. Görüntü Toplama ve Model Geliştirme

Geliştirilen üçüncü sistemin eğitilmesi ve test edilmesi için, toplamda 174 katılımcıya ait görüntüler kullanılmıştır (AS +: 87, AS -: 87). Veri sızıntısını önlemek ve aynı hastaya ait görüntülerin birden fazla veri setinde bulunmasını engellemek amacıyla, hastalar %80'i eğitim setinde, %12'u geçerlilik denetim verisi olarak ve %8'i test setinde yer alacak şekilde

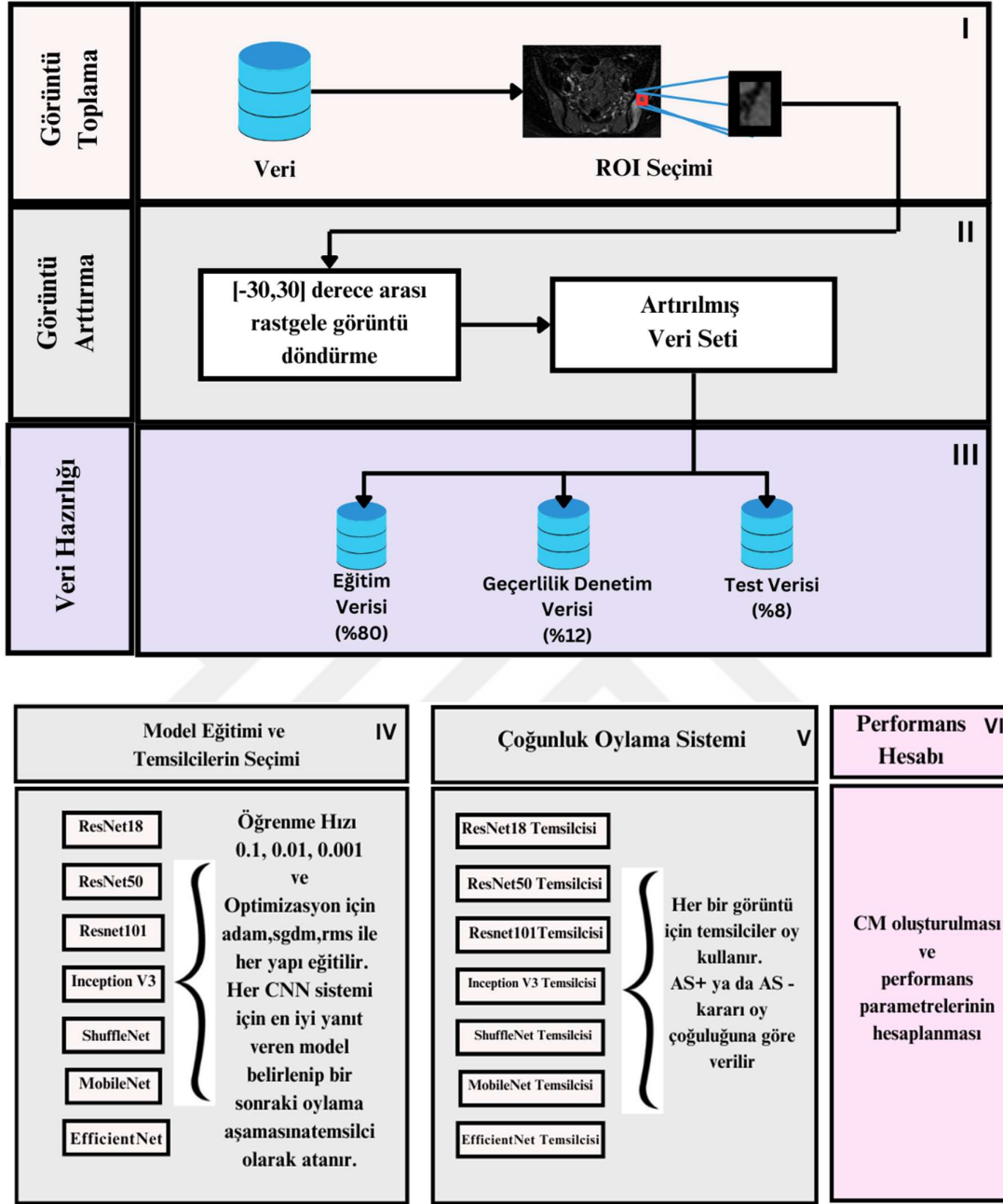
ayrılmıştır (Şekil 3.8). Bu yöntemle, modelin genelleme kabiliyetini artırmak ve farklı hastalardan elde edilen görüntülerle doğru bir eğitim süreci sağlamak amaçlanmıştır. Bu adımları takiben, hedef dokuyu içerecek şekilde 40*50 piksel boyutlarında ROI belirlenmesi yapılmıştır. ROI seçimi, modelin belirli bölgelere odaklanarak daha doğru sonuçlar elde etmesini sağlamak üzere planlanmıştır.

Görüntü sayısının model eğitimi için yeterli olup olmadığı değerlendirildiğinde, baş parmak kuralına göre [112] her bir sınıf için öğrenmenin yetersiz kalmaması ve aşırı öğrenmeyi engellemek adına 1000 adet görüntüye ihtiyaç vardır. Dolayısı ile sistemimizin çıktısını da iki adet sınıf olduğu için toplamda her sınıf için minimum 1000 görüntü olacak şekilde minimum 2000 adet görüntü kullanılması gerektiği hesaplanmıştır.

Eğitim, test ve doğrulama süreçleri için her sette yer alan, farklı hastalara ait olduğu doğrulanan görüntüler, veri artırma teknikleri kullanılarak çoğaltılmıştır. DL modellerini kullanan araştırmacılar, mevcut eğitim verilerini genişletmek için sıklıkla veri artırma tekniklerini kullanmaktadır. Yaygın olarak kullanılan veri artırma yöntemlerinden biri, veri setinin çeşitliliğini artırarak modelin genelleme yeteneğini geliştirebilen görüntü döndürme işlemidir [174-176]. Veri artırma, derin öğrenme modellerinin performansını artırmak için yaygın olarak kabul edilen bir yaklaşımdır; zira daha büyük veri setleri genellikle daha iyi performans sağlar [174]. Ancak, geniş veri setleri oluşturmak, veri toplama ve etiketleme için gerekli olan manuel çaba nedeniyle zorlu bir süreç olabilir [174]. Bu sorunu aşmak için araştırmacılar, çevirme, ölçeklendirme ve döndürme gibi basit düzenlemeler de içeren çeşitli veri artırma teknikleri geliştirmiştir [175].

Veri artırma işlemi kapsamında, yukarıda verilen bilgiler doğrultusunda orijinal MRI görüntüleri +30 ve -30 derece arasında sağa ve sola rastgele döndürülerek gerçekleştirilmiştir. Bu yöntem, sınırlı sayıdaki görüntülerden modelin öğrenmesi sırasında aşırı öğrenme riskini azaltmış ve görüntü sayısı, eğitim için gereken minimum veri miktarının üzerine çıkarılarak yetersiz öğrenme riski ortadan kaldırılmıştır.

Sonuç olarak, toplamda 2498 adet arttırılmış MRI görüntüsü elde edilmiştir (AS +: 1364, AS -: 1134) (Şekil 3.8). Elde edilen bu görüntüler, eğitim setinde (AS +: 1090, AS -: 902), test setinde (AS +: 146, AS -: 152) ve doğrulama setinde (AS +: 128, AS -: 80) kullanılarak modelin performansı kapsamlı bir şekilde değerlendirilmiştir.



Şekil 3.8 Yöntem 3 iş akış diyagramı.

Transfer öğrenme sürecinde kullanılacak olan önceden eğitilmiş ESA modelleri sırasıyla: ResNet-18, ResNet-50, ResNet-101, Inceptionv3, ShuffleNet, MobileNet ve EfficientNet olarak belirlenmiştir. Bu modeller, daha önce geniş veri setleri üzerinde eğitildikleri için AS sınıflandırması için de uygun bir başlangıç noktası sunmaktadır.

ResNet-18, ResNet-50 ve ResNet-101, güçlü özellik çıkarma yetenekleri ve önceden eğitilmiş ağırlıkların bulunabilirliği sayesinde transfer öğrenimi için uygun olan popüler ESA modelleridir. Bu modellerin, geniş bir görev ve veri seti yelpazesinde etkileyici performans sergilediği gösterilmiş olup [116], transfer öğrenimi uygulamaları için pratik bir seçenek sunmaktadır. Inception-v3 ise transfer öğrenimi için etkili bir şekilde kullanılacak bir diğer ESA modelidir [117]. Özel konvolüsyon modüllerini içeren benzersiz mimarisi, diğer görevlere aktarılabilecek son derece ayırtıcı özellikler öğrenebildiğini kanıtlamıştır. ShuffleNet, MobileNet ve EfficientNet ise kaynakları sınırlı cihazlarda verimlilik ve konuşlandırma için tasarlanmış hafif ESA modelleridir. Bu modeller, performans ve verimlilik arasında iyi bir denge sağlayabildiklerinden dolayı özellikle hesaplama gücü veya bellek sınırlı olduğunda transfer öğrenimi için avantajlı olabilir [177, 178].

Oylama aşamasına geçilmeden önce, her bir önceden eğitilmiş modelin AS sınıflandırmasında en iyi performansı gösteren temsilcilerinin belirlenmesi gerekmektedir. Bu adım, modellerin optimal performans göstermesini sağlayacak ayarların belirlenmesini ve en iyi sonuç veren model versiyonlarının seçilmesini hedeflemektedir.

Her bir ESA modelini temsil edecek en iyi performansa sahip modelin belirlenmesi için, modeller farklı hiper parametre kombinasyonlarıyla eğitilmiştir. Özellikle, her model için 3 farklı öğrenme oranı (0.1, 0.01, 0.001) ve 3 farklı optimizasyon algoritması (momentumlu stokastik gradyan inişi (sgdm), adaptif momentum AdaDelta (adam), Ortalama Karekök Yayılımı (rmsprop)) kullanılarak eğitimler gerçekleştirilmiştir. AS tespiti koymada en başarılı olan hiper parametre kombinasyonuna sahip model, ilgili ESA modelini temsil etmek üzere seçilmiştir. Bu yaklaşım sayesinde, her bir modelin performansını en üst düzeye çıkaran parametrelerle temsil edilmesi sağlanmış ve ÇOA sürecinde daha başarılı bir sonuç elde edilmesi hedeflenmiştir.

Bu yedi önceden eğitilmiş ESA modelinin, farklı öğrenme hızları ve optimizasyon algoritmaları ile eğitilip en iyi yanıt veren temsilcileri seçildikten sonra, her birinin aynı görüntüye AS + ya da AS - olarak oylaması istenmiştir. Bu sürecin sonucunda, oy çokluğu ile belirlenen karar, geliştirilen ÇOA sisteminin kararı olarak sunulmuştur. Ardından bu karar, gerçek durum ile karşılaştırılarak bir CM oluşturulmuş ve istatistiksel performans analizleri gerçekleştirilmiştir. Böylece, sistemin AS tespiti konusundaki performansı ayrıntılı bir şekilde değerlendirilmiştir.

4. DENEYSEL ÇALIŞMALAR VE BULGULAR

Bu kısımda bir önceki bölümde anlatılan AS tespiti için geliştirilen 3 farklı YZ modelinin performansları ile ilgili sonuçlar sunulmuş ve en sonunda en iyi performansı veren modeller listelenmiştir. Bahsedilen çalışmalar şöyle sıralanmıştır: ML Tabanlı geliştirilen modelinin performans analizi, Haralick özellikleri kullanılarak geliştirilen DL modelinin performans analizi, DL tabanlı modeller ile oluşturulan ÇOA modelinin performans analizi ve en iyi performans gösteren sistemlerin sonuçlarının ve hiper parametrelerinin sunulması.

4.1. Yöntem 1 ile Geliştirilen Modelinin Performans Analizi

Bu bölümde daha önceki 3.1 bölümünde anlatılan klasik ML algoritmaları kullanılarak elde edilen modelin sonuçlar paylaşılmıştır.

En iyi tanı performansına sahip modeller belirlemek için farklı hiperparametrelerin kullanıldığı 330 deney yapıldıktan sonra, test setinde ya da doğrulama setinde %70 doğrulukta sınıflandırma yapan modeller elenmiştir. Sonrasında ilgili kısıtımızın üzerinde başarı gösterebilen ML tabanlı modeller Tablo 4.1’de sunulmuştur. Geliştirilen modellerin performansını değerlendirmek amacıyla şu metrikler kullanılmıştır: Gerçek Pozitif (GP), hastanın AS'ye sahip olduğu durumda AS'yi doğru bir şekilde sınıflandırmayı ifade eder; Gerçek Negatif (GN), AS ile ilgili durumların yokluğunu doğru bir şekilde tanımayı ifade eder; Hatalı Pozitif (HP), AS olmayan bir hastayı yanlış bir şekilde AS olarak sınıflandırmayı; Hatalı Negatif (HN) ise, hastada AS ile ilişkili yapılar mevcutken yanlış bir şekilde AS negatif olarak sınıflandırmayı ifade eder (Tablo 4.1).

Önerilen ML modellerinin sınıflandırma performansları ve ilgili doğrulama seti üzerindeki AS tespit performansları Tablo 4.1’de sunulmuştur. AS tespiti için eğitilen modeller arasında yalnızca üçü %85’in üzerinde doğruluğa ulaşabilmiştir. Bu modeller; YSA tabanlı model (doğruluk: 0.885), Toplu Öğrenme tabanlı model (doğruluk: 0.870) ve SVM tabanlı modeldir (doğruluk: 0.856). Bu modeller arasında, YSA tabanlı model doğrulama verisinde en iyi performansı göstermiştir (doğruluk: 0.885, AUC: 0.941, duyarlılık: 0.883). Test verisinde en iyi AS tespit performansına sahip modellerin eğitiminde kullanılan hiper parametreler ise Tablo 4.2’te sunulmuştur.

Tablo 4.1 AS tespiti için Yöntem 1 ile geliştirilen modellerin geçerlilik denetim veri seti ve test seti üzerindeki performans analizleri

Model Adı	Geçerlilik Denetim Veri Seti Doğruluk	Test Seti Doğruluk	AUC	GP	GN	HP	HN	Duyarlılık	Özgünlük	Kesinlik	Geri Çağırma	F1 Skoru
Karar Ağaçları	0,960	0,841	0,833	103	72	8	25	0,805	0,900	0,928	0,805	0,862
SVM	0,959	0,856	0,921	115	63	17	13	0,898	0,788	0,871	0,899	0,885
Ağaç Tabanlı TOA	0,949	0,870	0,923	102	79	1	26	0,797	0,988	0,990	0,797	0,883
k-NN	0,957	0,813	0,858	96	73	7	32	0,750	0,913	0,932	0,750	0,831
Naive Bayes	0,746	0,678	0,764	78	63	17	50	0,609	0,788	0,821	0,609	0,700
YSA	0,961	0,885	0,941	113	71	9	15	0,883	0,888	0,926	0,883	0,904

Tablo 4.2 Test veri setinde %85'in üzerinde Ankilozan Spondilit tespit doğruluğuna sahip modellerin eğitim hiper parametreleri.

Model Adı	Eğitim Hiper parametreleri
YSA	Tam Bağlantılı Katman Sayısı:1 Aktivasyon Fonksiyonu: ReLU Düzenleştirme Gücü: 0.00012, İlk Katman Boyutu:280
Ağaç Tabanlı TOA	TOA Method: Gentle Boost Öğrenici Sayısı:88 Öğrenme Hızı:0.0108, Maksimum Bölünme Sayısı:21
SVM	Kutu Kısıtı Seviyesi: 969.394, Kernel Skalası:4.0944 Kernel Fonsiyonu: Gaussian

4.2. Haralick Özellikleri Kullanılarak Geliştirilen DL Modelinin Performans Analizi

Bu bölümde daha önceki 3.2 başlık altında oluşturulmuş Haralick özellikleri kullanılarak geliştirilen DL modelinin AS tespit performansının analizi yapılacaktır.

Veriler, Bölüm 3.2.1 başlığında bahsedildiği yöntem ile derlendikten bu başlık altında önerilen DL sistemi farklı hiper parametreler kullanılarak eğitilmiştir. Adam, rmsprop ve sgdm çözücüleri kullanılmış, her çözücü için 0,001, 0,01 ve 0,1 olmak üzere üç farklı başlangıç öğrenme hızı ile deneyler gerçekleştirilmiştir. Ardından, eğitilen modellerden doğrulama verilerini sınıflandırmaları istenmiştir. Modellerin sınıflandırma performansı istatistiksel olarak değerlendirilmiş ve ROC ve AUC değerleri hesaplanmıştır. Doğrulama setini %85 ve üzeri doğrulukla sınıflandıran modellerin hiper parametreleri ve performansları Tablo 4.3'de sunulmuştur.

Tablo 4.3 AS tespiti için Yöntem 2 ile geliştirilen modellerin hiper parametreleri ve performans analizleri

En İyileyici	Öğrenme Oranı	Test Doğruluğu	AUC	GP	GN	HP	HN	Duyarlılık	Özgünlük	Keskinlik	Geri Çağırma	F1 Skoru
adam	0,001	0,889	0,945	118	67	13	10	0,922	0,838	0,901	0,922	0,911
adam	0,01	0,875	0,933	118	64	16	10	0,922	0,800	0,881	0,922	0,901
adam	0,1	0,784	0,853	83	80	0	45	0,648	1,000	1,000	0,648	0,787
rms	0,001	0,918	0,941	117	74	6	11	0,914	0,925	0,951	0,914	0,932
rms	0,01	0,928	0,942	116	77	3	12	0,906	0,963	0,975	0,906	0,939
rms	0,1	0,764	0,904	79	80	0	49	0,617	1,000	1,000	0,617	0,763
sgdm	0,001	0,894	0,948	118	68	12	10	0,922	0,850	0,908	0,922	0,915
sgdm	0,01	0,904	0,944	114	74	6	14	0,891	0,925	0,950	0,891	0,919
sgdm	0,1	0,880	0,9418	115	68	12	13	0,898	0,850	0,906	0,898	0,902

4.3. DL Tabanlı Modeller ile Oluşturulan Çoğunluk Oylama Algoritmasının Performans Analizi

Bu bölümde, Bölüm 3.3’de önerilen ÇOA sisteminin performans analizi incelenmiştir.

Öncelikle ilgili modelin geliştirilmesi için oluşturulmuş olan 7 adet ön eğitimli ESA modeli farklı öğrenme hızları ve çözücüler ile oluşturulan hiper parametreler ile eğitilip, kendi içinde performans analizine tabii tutulmuştur. Böylelikle istatistiksel olarak en iyi performansa sahip olan sistem, kendi grubunun temsilcisi olarak atanmıştır. Temsilcilerin oy çoğunluğu sınıflandırmanın sonucunu belirlemiştir. Model isimleri, temsilcilerin hangi ESA modeline ait olduğunun anlaşılabilmesi için uygun bir şekilde isimlendirip, performansları ResNet-18 için Tablo 4.4’de , ResNet-50 için Tablo 4.5’de , ResNet-101 için Tablo 4.6’da, Inceptionv3 için Tablo 4.7’de, ShuffleNet için Tablo 4.8’de, MobileNet için Tablo 4.9’da ve EfficientNet için Tablo 4.10’da sunulmuştur. İlgili tablolarda temsilci koyu stilde yazılmıştır. Modellerin eğitiminde kullanılan temel hiper parametreler performans istatistik analizlerinde yer aldığı ilgili tablolarda sunulmuştur. Daha sonra bu temsilcilerin aynı doğrulama setinde gösterdikleri ve her bir görüntüye verdikleri yanıtlar kaydedilerek oylama sonuçları oluşturulmuştur. Örneğin doğruluk setinden aldığımız 240888, 1072819, 41074 ve 382668 numaralı verilere verdikleri oylama sonuçlarına ve görüntülerin uzmanlar tarafından incelenerek konulan gerçek AS teşhisleri karşılaştırıldığında, 240888 ve 1072819 temsilcilerin verdiği farklı oylara rağmen çoğunluğun oyları ile doğru olarak sınıflandırılmıştır. Buna karşın 41074 ve 382668 ise bazı sistemler gerçek teşhis lehinde oy verseler de ÇOA modeli hatalı karar vermiştir. Yukarıda anlatılan oylama sistemine göre modelin performansı, temsilcilerin performansları ile beraber Tablo 4.12’de sunulmuştur.

Tablo 4.4 Resnet_18 temsilci seçimi için farklı hiperparemetrelere sahip ön öğrenmeli ESA modellerinin sınıflandırma performansları

Evrişimsel Ağ	En İyileyici	Öğrenme Oranı	Doğruluk	AUC	GP	GN	HP	HN	Duyarlılık	Özgünlük	Keskinlik	Geri Çağırma	F1 Skoru
ResNet-18	sgdm	0,001	0,851	0,944	103	74	6	25	0,805	0,925	0,945	0,805	0,869
ResNet-18	sgdm	0,01	0,856	0,968	101	77	3	27	0,789	0,963	0,971	0,789	0,871
ResNet-18	adam	0,001	0,880	0,960	106	77	3	22	0,828	0,963	0,972	0,828	0,895
ResNet-18	adam	0,01	0,894	0,975	108	78	2	20	0,844	0,975	0,982	0,844	0,908
ResNet-18	rms	0,001	0,851	0,954	99	78	2	29	0,773	0,975	0,980	0,773	0,865
ResNet-18	rms	0,01	0,904	0,974	111	77	3	17	0,867	0,963	0,974	0,867	0,917

Tablo 4.5: ResNet-50 temsilci seçimi için farklı hiper parametrelere sahip ön öğrenmeli ESA modellerinin sınıflandırma performansları

Evrişimsel Ağ	En İyileyici	Öğrenme Oranı	Doğruluk	AUC	GP	GN	HP	HN	Duyarlılık	Özgünlük	Keskinlik	Geri Çağırma	F1 Skoru
ResNet-50	sgdm	0,001	0,913	0,961	114	76	4	14	0,891	0,950	0,966	0,891	0,731
ResNet-50	sgdm	0,01	0,856	0,968	101	77	3	27	0,789	0,963	0,971	0,789	0,770
ResNet-50	adam	0,001	0,880	0,960	106	77	3	22	0,828	0,963	0,972	0,828	0,798
ResNet-50	adam	0,01	0,894	0,975	108	78	2	20	0,844	0,975	0,982	0,844	0,729
ResNet-50	rms	0,001	0,851	0,954	99	78	2	29	0,773	0,975	0,980	0,773	0,811
ResNet-50	rms	0,01	0,904	0,974	111	77	3	17	0,867	0,963	0,974	0,867	0,711

Tablo 4.6: ResNet-101 temsilci seçimi için farklı hiper parametrelere sahip ön öğrenmeli ESA modellerinin sınıflandırma performansları

Evrişimsel Ağ	En İyileyici	Öğrenme Oranı	Doğruluk	AUC	GP	GN	HP	HN	Duyarlılık	Özgünlük	Keskinlik	Geri Çağırma	F1 Skoru
ResNet-101	sgdm	0,001	0,894	0,956	115	71	9	13	0,898	0,888	0,927	0,898	0,913
ResNet-101	sgdm	0,01	0,875	0,966	106	76	4	22	0,828	0,950	0,964	0,828	0,891
ResNet-101	adam	0,001	0,894	0,965	110	76	4	18	0,859	0,950	0,965	0,859	0,909
ResNet-101	adam	0,01	0,851	0,927	108	69	11	20	0,844	0,863	0,908	0,844	0,874
ResNet-101	rms	0,001	0,923	0,973	115	71	9	13	0,891	0,975	0,983	0,891	0,934
ResNet-101	rms	0,01	0,928	0,971	106	76	4	22	0,914	0,950	0,967	0,914	0,940

Tablo 4.7: Inceptionv3 temsilci seçimi için farklı hiper parametrelere sahip ön öğrenmeli ESA modellerinin sınıflandırma performansları

Evrişimsel Ağ	En İyileyici	Öğrenme Oranı	Doğruluk	AUC	GP	GN	HP	HN	Duyarlılık	Özgünlük	Keskinlik	Geri Çağırma	F1 Skoru
Inceptionv3	sgdm	0,001	0,904	0,967	113	75	5	15	0,883	0,938	0,958	0,883	0,919
Inceptionv3	sgdm	0,01	0,913	0,965	116	74	6	12	0,906	0,925	0,951	0,906	0,928
Inceptionv3	adam	0,001	0,928	0,979	118	75	5	10	0,922	0,938	0,959	0,922	0,940
Inceptionv3	adam	0,01	0,889	0,950	110	75	5	18	0,859	0,938	0,957	0,859	0,905
Inceptionv3	rms	0,001	0,923	0,975	120	72	8	8	0,938	0,900	0,938	0,938	0,938
Inceptionv3	rms	0,01	0,909	0,965	112	77	3	16	0,875	0,963	0,974	0,875	0,922

Tablo 4.8: ShuffleNet temsilci seçimi için farklı hiper parametrelere sahip ön öğrenmeli ESA modellerinin başarı performansları

Evrişimsel Ağ	En İyileyici	Öğrenme Oranı	Doğruluk	AUC	GP	GN	HP	HN	Duyarlılık	Özgünlük	Keskinlik	Geri Çağırma	F1 Skoru
ShuffleNet	sgdm	0,001	0,894	0,977	111	75	5	17	0,867	0,938	0,957	0,867	0,910
ShuffleNet	sgdm	0,01	0,894	0,977	111	75	5	17	0,867	0,938	0,957	0,867	0,910
ShuffleNet	adam	0,001	0,894	0,967	113	73	7	15	0,883	0,913	0,942	0,883	0,911
ShuffleNet	adam	0,01	0,904	0,967	113	75	5	15	0,883	0,938	0,958	0,883	0,919
ShuffleNet	rms	0,001	0,923	0,979	118	74	6	10	0,922	0,925	0,952	0,922	0,937
ShuffleNet	rms	0,01	0,913	0,961	114	76	4	14	0,891	0,950	0,966	0,891	0,927

Tablo 4.9: MobileNet temsilci seçimi için farklı hiper parametrelere sahip ön öğrenmeli ESA modelinin başarı performansları

Evrişimsel Ağ	En İyileyici	Öğrenme Oranı	Doğruluk	AUC	GP	GN	HP	HN	Duyarlılık	Özgünlük	Keskinlik	Geri Çağırma	F1 Skoru
MobileNet	sgdm	0,001	0,889	0,954	112	73	7	16	0,875	0,913	0,941	0,875	0,907
MobileNet	sgdm	0,01	0,875	0,961	105	77	3	23	0,820	0,963	0,972	0,820	0,890
MobileNet	adam	0,001	0,909	0,976	112	77	3	16	0,875	0,963	0,974	0,875	0,922
MobileNet	adam	0,01	0,822	0,947	94	77	3	34	0,734	0,963	0,969	0,734	0,836
MobileNet	rms	0,001	0,875	0,958	106	76	4	22	0,828	0,950	0,964	0,828	0,891
MobileNet	rms	0,01	0,923	0,959	115	77	3	13	0,898	0,963	0,975	0,898	0,935

Tablo 4.10: EfficientNet temsilci seçimi için farklı hiper parametrelere sahip ön öğrenmeli ESA modelinin başarı performansları

Evrişimsel Ağ	En İyileyici	Öğrenme Oranı	Doğruluk	AUC	GP	GN	HP	HN	Duyarlılık	Özgünlük	Keskinlik	Geri Çağırma	F1 Skoru
EfficientNet	sgdm	0,001	0,889	0,959	111	74	6	17	0,8672	0,9250	0,9487	0,867	0,9061
EfficientNet	sgdm	0,01	0,913	0,974	114	76	4	14	0,8906	0,9500	0,9661	0,891	0,9268
EfficientNet	adam	0,001	0,889	0,959	111	74	6	17	0,8672	0,9250	0,9487	0,867	0,9061
EfficientNet	adam	0,01	0,909	0,979	112	77	3	16	0,8750	0,9625	0,9739	0,875	0,9218
EfficientNet	rms	0,001	0,889	0,959	111	74	6	17	0,8672	0,9250	0,9487	0,867	0,9061
EfficientNet	rms	0,01	0,913	0,962	114	76	4	14	0,8906	0,9500	0,9661	0,891	0,9268

Tablo 4.11 Yöntem 3 ile oluşturulan ÇOA modeli ve temsilci sınıflandırma performansları

Evrişimsel Ağ	En İyileyici	Öğrenme Oranı	Doğruluk	AUC	GP	GN	HP	HN	Duyarlılık	Özgünlük	Keskinlik	Geri Çağırma	F1 Skoru
ResNet-18	rms	0,001	0,851	0,954	99	78	2	29	0,773	0,975	0,980	0,773	0,865
ResNet-50	rms	0,001	0,851	0,954	99	78	2	29	0,773	0,975	0,980	0,773	0,811
ResNet-101	rms	0,001	0,923	0,973	114	78	2	14	0,891	0,975	0,983	0,891	0,848
Inceptionv3	rms	0,001	0,923	0,975	120	72	8	8	0,9375	0,900	0,938	0,100	0,938
ShuffleNet	rms	0,001	0,923	0,979	118	74	6	10	0,922	0,925	0,952	0,938	0,937
MobileNet	rms	0,01	0,923	0,959	115	77	3	13	0,898	0,963	0,975	0,898	0,935
EfficientNet	sgdm	0,01	0,913	0,974	114	76	4	14	0,8906	0,950	0,966	0,891	0,927
TOS	-		0,938	0,942	118	77	3	10	0,922	0,963	0,975	0,922	0,948

Tablo 4.12: Yöntem 3 ile geliştirilen ÇOA sistemi için oylama örnekleri.

Data ID	ResNet-18	ResNet-50	ResNet-101	Inceptionv3	ShuffleNet	MobileNet	EfficientNet	Toplu Oylama Sonucu	Gerçek Teşhis Sınıfı
240888	Negatif	Pozitif	Pozitif	Pozitif	Pozitif	Pozitif	Negatif	Pozitif	Pozitif
1072819	Negatif	Negatif	Pozitif	Pozitif	Negatif	Negatif	Negatif	Negatif	Negatif
41074	Pozitif	Pozitif	Negatif	Negatif	Negatif	Pozitif	Pozitif	Negatif	Pozitif
382668	Pozitif	Negatif	Negatif	Negatif	Pozitif	Negatif	Pozitif	Negatif	Pozitif

5. SONUÇLAR

AS teşhisi, benzer semptomlara sahip diğer romatolojik hastalıklarla karıştırılabilmesi ve hastalığın klinik seyrinin değişkenliği nedeniyle önemli zorluklar barındırmaktadır. Erken ve doğru teşhis, tedavi sürecinin gecikmesini veya gereksiz tedavilerin önüne geçilmesini sağladığı için büyük önem taşımaktadır. Bu zorlukların üstesinden gelmek amacıyla, tezimiz kapsamında AS teşhisine yönelik üç farklı YZ modeli geliştirilmiştir. Bu modeller, %90 doğruluk oranına ulaşma hedefiyle tasarlanmış olup, geleneksel yöntemlere kıyasla daha yüksek doğruluk ve güvenilirlik sunmaktadır. Çalışmamız, özellikle STIR sekans MR görüntülerinin kullanılmasıyla önceki yaklaşımlardan ayrılmakta ve bu görüntülerin yüksek hassasiyetinden faydalanarak, subkondral kemik ödemi ve osteit gibi AS'nin erken belirtilerini daha isabetli şekilde, hedeflendiği gibi 90% doğruluk oranının üzerinde tespit edebilmektedir.

Ayrıca, farklı yaş gruplarını kapsayan büyük ölçekli bir veri seti üzerinde eğitim alan modellerin sağlamlığı artırılmış ve gerçek dünya uygulamalarına uygun hale getirilmiştir. Çeşitli Haralick özelliklerinin entegrasyonu ile AS ile ilişkili yapılar, MR görüntülerinde daha ayrıntılı sınıflandırılabilir. Bunun yanı sıra, çalışmamızda veri sızıntısına karşı alınan önlemler ile tahmin performansının yapay olarak artmasının önüne geçilmiş, böylece geliştirilen modellerin güvenilirliği sağlanmıştır. Bu yenilikler ışığında, çalışmamızın literatüre önemli katkılar sunduğu ve klinik ortamda uygulanabilirlik potansiyelinin yüksek olduğu düşünülmektedir. Bu bölümde, elde edilen sonuçlar detaylı olarak tartışılarak, geliştirilen sistemlerin klinik kullanım potansiyelleri ve bu alanda sağladıkları yenilikler değerlendirilecektir. Ayrıca, modellerin sınırlılıkları ve gelecekteki performanslarının iyileştirilmesi için öneriler de sunulacaktır.

5.1. Önerilen Modellerin Yöntem Olarak Benzeyen Çalışmalar ile Karşılaştırılması ve Literatüre Katkıları

Bu bölümde tez kapsamında önerilen modeller, literatürde yer alan diğer görüntü verisi kullanan sistemler ile karşılaştırılarak ve farklılıkları ortaya konmuştur.

Bu çalışmada, önerilen modelin performansı ve uygulama kolaylığı diğer çalışmalarla karşılaştırılmış ve önemli avantajlar sağladığı görülmüştür. Zunti ve Ark [15] tarafından yapılan çalışmada, yaşlanma nedeniyle eklemlerde oluşan erozyonları genç bireylerin

medikal görüntülerinden ayırmak için k-NN algoritması kullanılmıştır. Bu çalışmada, ön işleme yapılmaksızın yüksek doğruluk oranı (%93,6) elde edilmiştir. Ancak, veri setinin yaşa göre kategorize edilmesi, geliştirilebilirlik ve AS'nin başlangıcındaki tipik semptomları yakalamada eksikliklere neden olmuştur. Buna karşılık önerilen model, yaş farklılıklarını dikkate almadan daha geniş bir geliştirilebilirlik sağlamış ve gerçek dünya senaryolarına daha uygun bir yapı sunmuştur (Tablo 5.1).

Tablo 5.1 Önerilen modellerin, yakın metotlara sahip çalışmalarla karşılaştırmalı avantaj ve dezavantaj analizi

Çalışma	Yaklaşım	Avantajlar	Dezavantajlar	Tez kapsamında önerilen Model ile Karşılaştırma
Zunti ve Ark. [15]	Ön işleme yapılmadan klasik ML; k-NN ile sınıflandırma	Yaşlanmaya bağlı erozyonları başarıyla ayırır; %93,6 doğruluk oranı	Veri seti yaşa göre kategorize edilmiş, bu nedenle AS başlangıcı için geliştirilebilirliği düşüktür; ön işleme yoktur.	Tez kapsamında önerilen model yaş farklarını dikkate almayarak geliştirilebilirlik ve gerçek dünya uygulamaları için daha uygundur.
Tas ve Ark. [18]	Ön eğitilmiş ESA modelleriyle özellik çıkarımı ve hibrit ML-DL sınıflandırma	Yoğun özellik çıkarımı ile hibrit sınıflandırma; topluluk oylama mekanizması sunar	Karmaşık sistem; açıklanabilirliği azaltır, klinik uygulamada sorunlar ele alınmamıştır.	Tez kapsamında önerilen model, daha basit bir sistem sunarak açıklanabilirliği artırır ve klinik uygulanabilirlik açısından daha uygundur.
Faleiros ve Ark. [16]	SPAIR T2-ağırlıklı MR görüntülerinden Haralick özellikleri ile k-NN tabanlı sınıflandırma	13 Haralick özelliği kullanımıyla duyarlılık ve AUC değerlerinde başarılı sonuçlar	Eğitim ve doğrulama setlerinde aynı görüntülerin yer alması veri sızıntısı riskini artırır; entegre bir iş akışı yoktur.	Tez kapsamında önerilen model veri sızıntısını önler, MATLAB platformunda tüm süreci entegre eder ve AUC ile duyarlılığı iyileştirir.
Lin ve Ark. [19]	Sakroiliit tespiti için DL modeli geliştirmek	Attention Unet bölütme ve ESA ile sınıflandırmada yüksek hassasiyet (%86) ve özgünlük (%92)	Çalışmanın uygulanabilirliği sınırlı; veri setindeki görüntü çeşitliliği kısıtlı	Önerilen model, daha küçük ve klinik açıdan pratik bir veri seti ile benzer doğruluk ve geliştirilebilirlik sağlamaktadır.
Zhang ve Ark. [20]	BT görüntülerini kullanarak AS ile ilişkili sakroiliit tanısı ve derecelendirmesi	Yüksek AUC değerleri (AS Seviye 0: 0,91; AS Seviye 1: 0,80; AS Seviye 2: 0,96); genç radyologlardan üstün performans	AS seviye 1 sınıflandırmada deneyimli uzmanlardan düşük performans; klinik sonuçlara sınırlı katkı	Tez kapsamında önerilen model, STIR MRI kullanarak radyolojik sınıflamaya daha fazla klinik odaklı ve açıklanabilir bir alternatif sunar.

Tas ve Ark. [18] tarafından geliştirilen hibrit ML-DL modeli ise, yoğun özellik çıkarımı ve topluluk oylama mekanizması ile dikkat çekmektedir. Bu sistemin karmaşıklığı,

açıklanabilirliği ve klinik uygulanabilirliği azalmıştır. Ayrıca, modelin eğitim süresi ve klinik kullanımdaki olası sorunlar üzerinde yeterince durulmamıştır. Önerilen model, daha basit ve anlaşılabilir bir sistem sunarak her bir ESA modelinin performansını ayrı ayrı ortaya koymuş ve klinik uygulamalar için daha açıklanabilir bir yapı sağlamıştır (Tablo 5.1).

Faleiros ve Ark. [16] tarafından yapılan çalışmada ise 13 Haralick özelliği ile k-NN algoritması kullanılarak duyarlılık ve AUC değerlerinde (%91) başarılı sonuçlar elde edilmiştir. Aynı hastaya ait görüntülerin hem eğitim hem de doğrulama setlerinde yer alması olasılığı, veri sızıntısı riski yaratmıştır. Ayrıca, süreçler farklı platformlarda yürütülmüş ve entegre bir iş akışı sağlanamamıştır. Buna karşılık, bu çalışmada önerilen model, veri sızıntısını önlemiş, tüm süreci MATLAB platformunda entegre etmiş ve daha yüksek AUC (%94,1) ve duyarlılık değerlerine ulaşmıştır (Tablo 5.2). Faleiros ve Ark. [15] çalışması ile ortak 10 Haralick özelliği kullanılarak eğitilen modellerin çapraz doğrulama sonuçları yalnızca %86,1 doğruluk, 0.9194 AUC, %85,9 duyarlılık ve %85,9 özgüllük değerlerine ulaşmıştır. Sonuçlar, yaygın olarak kullanılan özelliklerin kullanılmasıyla daha yüksek doğruluk değerleri elde edildiğini göstermiştir. Ancak, önerilen AS tespit modelimizde kullanılan Haralick özellikleri, doğruluğu artırmış ve özellikle doğrulama verilerinde modellerin AS duyarlılığını önemli ölçüde iyileştirmiştir. Bu durum, AS tespitinde entegre edilen Haralick özelliklerinin sunduğu ek bilgilere bağlanabilir (Tablo 5.2).

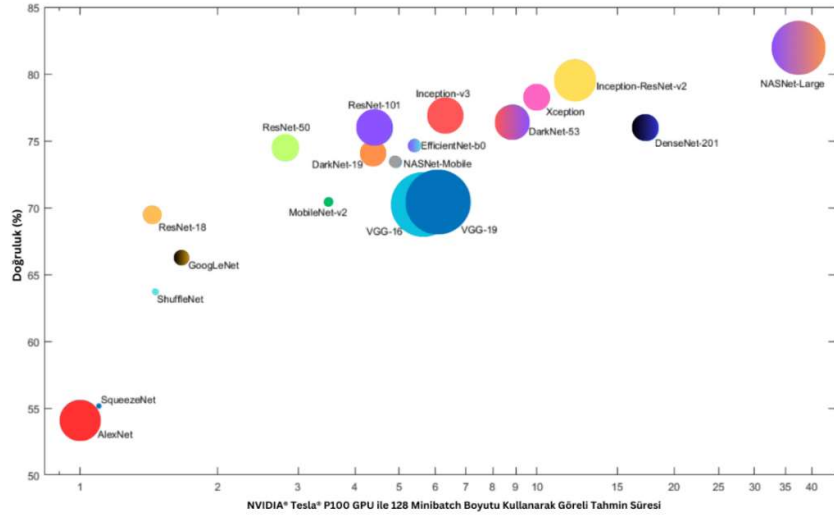
Bu sonuçlar önerilen modelin geliştirilebilirlik, açıklanabilirlik ve klinik uygulanabilirlik açısından daha üstün olduğunu ve literatürdeki diğer çalışmalardan önemli avantajlar sunduğunu göstermektedir.

Haralick özelliklerinin Yöntem 2’de ki gibi modellenmesi, özellikle DL çalışmaları için önemli bir avantaj sunmaktadır. DL modelleri genellikle yüksek hesaplama gücü gerektirir. Bu nedenle sayısal verilerin kullanılması, gerekli hesaplama gücünü önemli ölçüde düşürebilmektedir. Bu durum hesaplama kaynaklarının sınırlı olduğu senaryolarda DL modellerinin uygulanabilirliğini artırmaktadır. DL modellerinin sayısal veri ile işlenmesi durumunda daha düşük hesaplama gücüne ihtiyacı olduğuna dair yapılan çalışmalarda bulunmaktadır [179].

Tablo 5.2 Önerilen klasik ML tabanlı modelin tespit sonucu performansının en benzer araştırmayla karşılaştırılması.

	Faleiros et.al [12]	Önerilen Model
Veri Seti, Görüntü Sayısı	51	696
MR Görüntüsü	SPAIR T2-ağırlıklı görüntüler	STIR echo görüntüleri
Çalışmalardaki Ortak Haralick Özellikleri	1.contr 2.corm 3.idmnc 4.senth 5.svarh 6.dvarh 7.denth 8.inflh 9.inf2h 10.savgh	
Çalışmalardaki Farklı Haralick Özellikleri	1.Second Angular Momentum 2.Maximum Correlation Coef. 3. Variance	1.autoc 2.cprom 3.cshad 4.dissi 5.energ 6.entro 7.homom 8.homop 9.maxpr 10.sosvh 11.indnc
En iyi AS Tespit performansına sahip modellerin performans karşılaştırması	Sadece ROC parametreleri sunulmuştur AUC:0.91 Sensitivity:0.91 Specificity:0.89	Test Sonuçları Doğruluk:0.885 AUC:0.941 Hassasiyet:0.883 Specificity:0.888
İki çalışmadaki ortak Haralick özellikler ile geliştirilen modelin performans analizi	Test Sonuçları Accuracy:0.861 AUC: 0.9194 Sensitivity:0.859 Specificity:0.863	

Konu ile ilgili benzer çalışmalarda, modellerin eğitimi için harcanan sürelerin belirtilmediği tespit edilmiştir. Tez çalışması kapsamında, doğrulama veri setini %85 doğruluk ve üzeri ile sınıflandıran modellerin toplam eğitim süreleri Tablo 5.3'de sunulmuştur. Ayrıca, Şekil 5.1'de, bu modellerin eğitim süresinin yaklaşık tahmini görsel olarak ifade edilmiştir.



Şekil 5.1 Önceden eğitilmiş CNN tabanlı DL modellerinin doğruluğunun, hesaplama ve sınıflandırma sürelerinin karşılaştırılması [80].

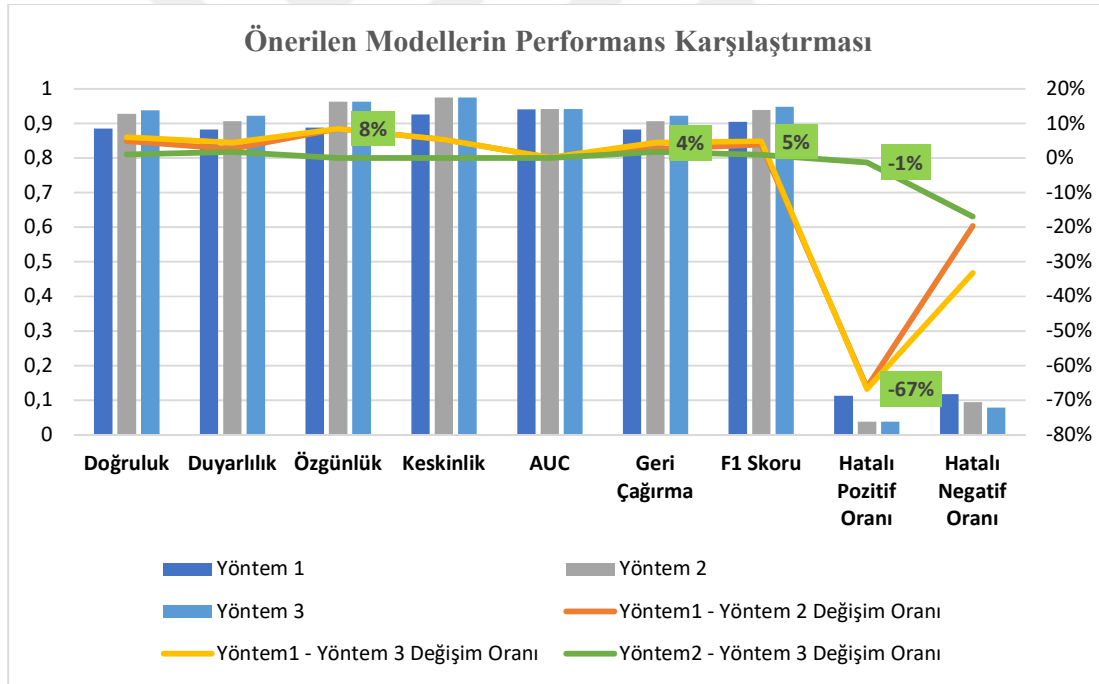
Tablo 5.3 Önerilen ML tabalı modelin tespit sonucu performansının en benzer araştırmayla karşılaştırılması.

	Model	Eğitim Süresi (saat)
Yöntem 1	YSA	0,43
	SVM	0,76
	Ağaç Tabanlı TOA	0,43
Yöntem 2	10 katmanlı DL modeli	0,1
Yöntem 3	ResNet-18	0,12
	ResNet-50	0,44
	ResNet-101	0,56
	Inceptionv3	0,95
	ShuffleNet	0,20
	MobileNet	0,35
	EfficientNet	0,86
		3,48

Buna göre Şekil 5.1 ve Tablo5.2 incelendiğinde önerilen modellerin, literatürde yapılan öneriler ile karşılaştırıldığında daha hızlı eğitilebildiği anlaşılmaktadır.

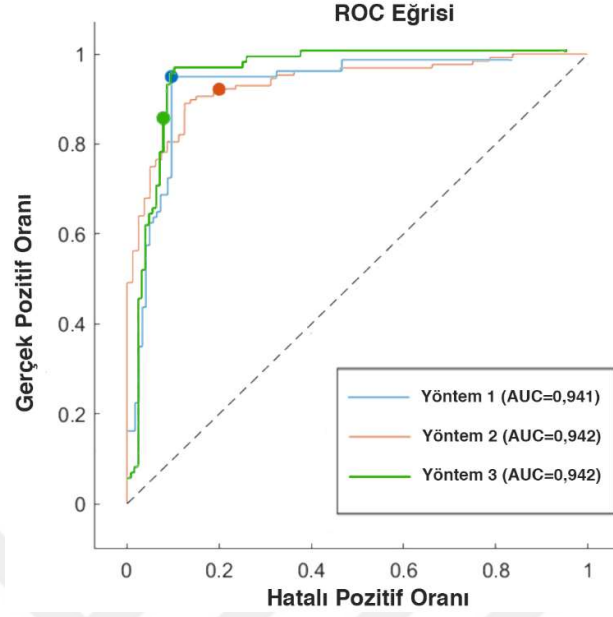
5.2. Önerilen Sistemlerden Hangisi AS Tespiti için Daha Uygunur?

Daha önceki bölümlerde belirtildiği gibi AS tespiti yapabilmek için üç farklı YZ modeli geliştirilmiştir. Modellerin nasıl oluşturulduğu Bölüm 3’de, modellerin sınıflandırma performanslarının istatistikleri ise Bölüm 4’de yer bulmuştur. Bu farklı yaklaşımlardan en iyi sonuç verenlerin performans ölçütlerinden bazıları Şekil 5.2’de karşılaştırılmıştır. Bu üç farklı yöntem ile geliştirilen modellerin ROC analizlerinin karşılaştırmaları Şekil 5.3’de sunulmuştur.



Şekil 5.2 Tez çalışması kapsamında, üç farklı yapay zekâ yaklaşımıyla eğitilen modellerin sınıflandırma performansları karşılaştırılması. Sunulan performans parametrelerinin sistemler arasındaki değişimleri, çizgi grafiklerle görselleştirilmiştir.

AS teşhisinde kullanılan üç farklı modelin sınıflandırma performansları, detaylı bir şekilde karşılaştırılarak tabloda ve grafikte analiz edilmiştir. OS modeli diğer yöntemlere kıyasla daha yüksek doğruluk, duyarlılık ve özgünlük oranlarıyla birlikte düşük hatalı negatif ve pozitif oranlarına sahiptir. Bu özellikler, modelin erken teşhis ve klinik uygulamalarda daha



Şekil 5.3 Tez çalışması kapsamında, üç farklı yapay zekâ yaklaşımıyla eğitilen modellerin ROC analizlerinin karşılaştırılması

güvenilir bir teşhis aracı olabileceğini göstermektedir. Özellikle, F1 skorundaki üstün performansı, modelin dengeli bir sınıflandırma sağladığını ve klinik ihtiyaçları karşılamada etkili bir çözüm sunduğunu ortaya koymaktadır. ÇOA modelinin, AS teşhisinde uygulanabilirlik ve güvenilirlik açısından en iyi seçenek olduğu değerlendirilmektedir. Bu önerinin nedenleri aşağıda detaylandırılmaktadır:

5.2.1. Genel Performans: Doğruluk ve AUC

ÇOA modeli, doğruluk açısından %93,8 ile en yüksek performansı sergilemiştir. AUC değeri (%94,2) ise pozitif ve negatif sınıfları ayırt etmede diğer modellerle eşdeğer düzeydedir. Yüksek doğruluk oranı, modelin örnekleri doğru bir şekilde sınıflandırma yeteneğini; yüksek AUC değeri ise modelin tüm karar eşiklerinde dengeli bir performans sunduğunu göstermektedir.

5.2.2. Duyarlılık ve Geri Çağırma

Duyarlılık oranı açısından ÇOA modeli %92,2 ile en iyi sonucu verirken, geri çağırma oranı da aynı şekilde %92,2 olarak hesaplanmıştır. Bu metrikler, modelin AS hastalarını doğru bir

şekilde tespit etme kapasitesini ölçmektedir. Yüksek duyarlılık ve geri çağırma oranları, yanlış negatiflerin sayısını azaltarak hastalığın erken teşhisine önemli bir katkı sağlamaktadır.

5.2.4. Hatalı Pozitif ve Negatif Oranları

ÇOA modeli, hatalı pozitif oranında %3,75, hatalı negatif oranında ise %7,81 ile en düşük değerlere ulaşmıştır. Bu durum, modelin yanlış pozitif ve yanlış negatif teşhis oranlarını en aza indirdiğini göstermektedir. Özellikle hatalı negatif oranında %33'lük bir iyileşme sağlanmıştır, bu da modelin hastaları gözden kaçırma olasılığını önemli ölçüde azaltmaktadır.

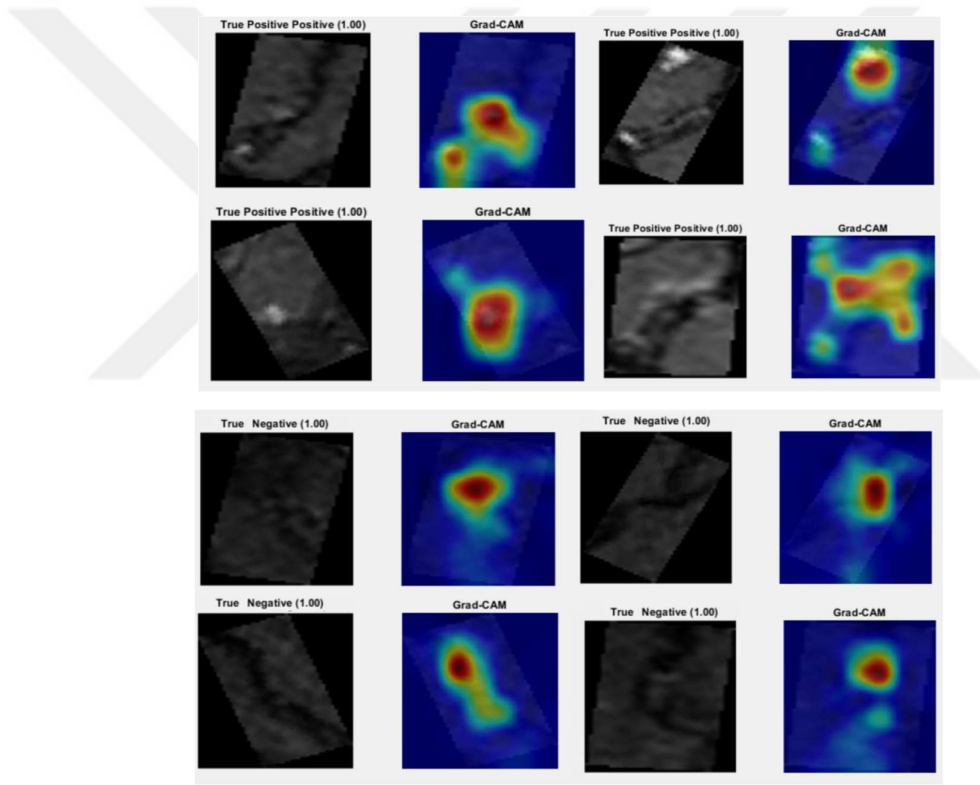
5.2.5. F1 Skoru

ÇOA modeli, %94,78'lik F1 skoru ile diğer modeller arasında en iyi değere sahiptir. F1 skoru, hem duyarlılığı hem de pozitif tahmin doğruluğunu dengeleyerek, modelin genel performansını ölçer. Bu yüksek skor, ÇOA modelinin pozitif ve negatif sınıfları başarılı bir şekilde ayırt ettiğini ve dengeli bir sınıflandırma sağladığını ortaya koymaktadır.

5.3. Grad-CAM Özellik Haritaları: DL Modelleri Nereden Öğrendi?

Bölüm 2.6.4.1'de açıklandığı üzere sinir ağlarının DL uygulanması, farklı türde problemlerde diğer ML algoritmalarından daha iyi performans gösterme yetenekleri nedeniyle hızla artmaktadır. Ancak bu yüksek performans kapasitesi özellikle tıbbi tanı gibi hassas araştırmalarda sistemlerin açıklanabilirliğini azaltmakta ve karşımıza “kara kutu” şeklinde davranan sistemler çıkmaktadır. Kara kutu sistemleri, girdiler ve çıktılar arasında bir ilişki kuran ancak bu ilişkinin nasıl oluşturulduğunu insan tarafından kolayca anlaşılabilir şekilde açıklayamayan modeller olarak tanımlanabilir. Özellikle derin öğrenme modellerinde, sınıflandırma işlemlerinin hangi özelliklere dayanarak yapıldığını ve bu özelliklerden nasıl bir sonuca varıldığını takip etmek oldukça zordur. Bu çalışma kapsamında önerdiğimiz çalışmalar incelendiğinde en iyi performans gösteren sistemin DL tabanlı ÇOA olduğu gözlemlenmişti. Ancak performans “kara kutu” probleminde de beraberinde getirmiş oldu. Bu sorunun, çalışmamız için minimize edilmesi ve de DL modellerinin karar verirken görüntünün hangi kısımlarını kullandıklarının belirlenebilmesi için Grad-CAM yöntemi kullanılmıştır. Önerilen ÇOA sistemindeki en yüksek duyarlılığa sahip olan sistem olan Inceptionv3 (Tablo 4.12) tarafından doğru olarak sınıflandırılmış

görsellerin bazıları ve ilgili görüntülerin Grad-CAM tekniği kullanılarak özellik haritasının çıkarılmış halleri Şekil 5.4’de sunulmuştur. Böylelikle Grad-CAM uygulaması ile görsel olarak da DL modellerinin nereden öğrendiğinin vurgulanması özellikle karar destek için başvuru yapan hekimlerin daha doğru karar vermesine yardımcı olabileceği ve Polonyi Paradox’unu aşılması için potansiyel bir çözüm olabileceği ön görülmektedir. Şekil 5.3’de farklı derecelerden dolayı oluşan siyah bölgelerin Grad-CAM uygulamasında da gözlemlendiği üzere karar verme üzerinde herhangi bir etkisi olmamıştır. Geliştirilen DL temelli AS sınıflandırma modellerimizin tamamen dokuda bulunan anormalliklerin bulunması ya da bulunmamasına göre karar verdiği gözlemlenmiştir.



Şekil 5.4 IncV3_5 modeli ile modeli ile doğru bir şekilde AS sınıflandırılması yapılan görüntülerin Grad-CAM yöntemi ile çıkarılmış özellik haritaları.

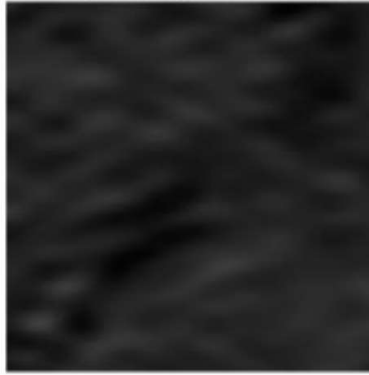
5.4. Yöntem 3'ün Erken Teşhis İçin Sınıflandırma Performans Analizi ve 223787 No'lu Denek

Yöntem 3 ile geliştirilen model, hasta olmayan bireyler ile AS'nin erken evrelerinde olan bireyleri ayırt etme kapasitesine odaklanmıştır. Test veri setinden global AS 0 derecesine

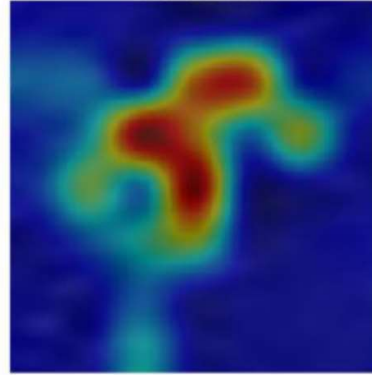
sahip bireyler (hasta olmayan) ile global AS 1 derecesine sahip bireyler (yeni AS teşhisi almış) arasında yapılan analizde model, %85 duyarlılık ve %96,39 özgünlük ve %85 geri çağırma değerlerine ulaşmıştır. Ayrıca, pozitif öngörü değeri %91,89, negatif öngörü değeri %93,02, doğruluk oranı %92,68 ve F1 skoru %88,31 olarak hesaplanmıştır. Hatalı pozitif oranı %3,61 ve hatalı negatif oranı %15,00 gibi düşük değerler, modelin sınıflandırma hatalarını minimize ettiğini göstermektedir.

Denek numarası 223787 bağlamında değerlendirildiğinde, modelin potansiyeli daha da belirginleşmekte ve geliştirilen modelin gerçek dünya vakalarındaki pratik uygulanabilirliğini göstermektedir. Genellikle, AS'nin ilerleyişini takip etmek, hastaların hastaneye tekrar başvurmadağı durumlarda oldukça zordur. Bu özel durumda, denek 2021 yılında bir MRI görüntülemesine girmiş ve radyologlar AS ile ilgili herhangi bir kanıt olmadığı sonucuna varmıştır. Ancak, denek 2023 yılında ısrar eden semptomlarla hastaneye tekrar başvurduğunda, fizik tedavi uzmanlarınca gerçekleştirilen fizik muayene sonucunda hastaya AS teşhisi konmuş ve ilaç tedavisine başlamıştır. Bu hastayı makine öğrenimi tabanlı modelimiz AS-pozitif olarak sınıflandırmıştır (Şekil 5.5).

223787 No'lu Deneğin Görüntüsü



Grad-Cam Analizi



Şekil 5.5 223787 No'lu deneğe ait bir görüntüsünün ROI'si ve görüntünün Grad-CAM analizi

Görüntülemelerde radyolojik belirtiler bulunmamasına rağmen, hastanede fizik tedavi departmanı tarafından yapılan fiziksel muayene bulgularına dayanarak AS tedavisi başlatılmıştır. Bu vaka, önerilen makine öğrenimi tabanlı algoritmaların AS'nin erken tespiti için potansiyelini vurgulamaktadır. Zira hastanın durumu, yalnızca hastaneye dönüşü ve araştırma ekibimiz tarafından yapılan yeniden değerlendirme sonrasında doğru bir şekilde ele alınabilmiştir. Bu vaka, önerilen modellerin AS'nin erken tespiti için sahip olduğu

potansiyeli ortaya koymaktadır. Ayrıca modelin gelecekte klinik ortamlarda erken teşhis ve tedavi sürecinde önemli bir yardımcı araç olarak kullanılabilme potansiyelini ortaya koymaktadır.

5.5. Yapay Zekâ Tabanlı Görüntü Sınıflandırma Yönteminin Tıbbi Alana Uygulanmasında Potansiyel Zorluklar

Tıbbi görüntü analiz görevlerinin otomatikleştirilmesinde YZ büyük bir potansiyel gösterse de, geniş çaplı klinik uygulama öncesinde ele alınması gereken önemli zorluklar bulunmaktadır. Bu zorluklardan biri, doğru derin öğrenme modellerini eğitmek için gerekli olan büyük, yüksek kaliteli ve etiketlenmiş tıbbi veri setlerinin eksikliğidir. Sağlık kurumları arasındaki gizlilik kaygıları ve farklı veri standartları, gerekli tıbbi görüntülerin bir araya getirilmesini ve paylaşılmasını zorlaştırmaktadır [180, 181]. Bu veri yetersizliği sorunu, tıbbi görüntülerin uzman klinisyenler tarafından manuel olarak etiketlenmesinin karmaşık ve maliyetli süreci ile daha da kötüleşmektedir [180]. Ayrıca, tıbbi veri setleri genellikle belirli hastalıkların diğerlerine göre çok daha nadir görüldüğü sınıf dengesizliklerinden muzdariptir[180]. Bu durum, yeterince temsil edilmeyen sınıflarda zayıf performans gösteren modellerin oluşmasına neden olabilir [180, 182]. Yeterli eğitim verisi bulunsa dahi, birçok yüksek performanslı derin öğrenme modelinin "kara kutu" doğası, doktorların sonuçlara güvenmesini ve bu sonuçları yorumlamasını zorlaştırarak klinik kabulü engelleyebilir [183]. Ek olarak tıbbi veri setleri, demografik veya altta yatan hastalık etkileri, korelasyonlar ve beklenmedik veya hatalı sonuçlar üretme gibi önemli kafa karıştırıcı faktörler içerebilir. Bu modellerin sahte öğrenmesine yol açabilir [184].

Bu problemlere ek olarak örüntü sınıflandırmanın tıbbi uygulamalarda kullanımıyla etik sorunlar da ortaya çıkmaktadır [185]. YZ modellerinin, eğitim verilerinde bulunan toplumsal önyargıları öğrenmesi ve sürdürmesi nedeniyle önyargı ve ayrımcılık potansiyeline dair endişeler bulunmaktadır. Bu durum belirli topluluklar için sağlık hizmetlerine erişim ve sağlık sonuçlarında eşitsizliklere yol açabilir [186]. Bununla birlikte YZ sistemlerinin hassas tıbbi görüntüleri ve verileri kullanması nedeniyle hasta gizliliği ve veri güvenliği konularında da sorunlar ortaya çıkmaktadır. YZ tabanlı görüntü sınıflandırmanın rutin klinik uygulamalarda yaygın olarak benimsenmesinden önce, bu etik etkilerin dikkatlice değerlendirilmesi gereklidir.

5.5.1. YZ Tabanlı Tıbbi Görüntüleme Çözümlerinin Benimsenmesinde Sağlık Profesyonellerinin Rolü ve Güçlendirilmesi

YZ'nin tıbbi görüntüleme alanına entegrasyonu, tanı doğruluğunu artırma, kişiye özel tedavi planlaması ve hasta sonuçlarını iyileştirme potansiyeline sahip olup, klinik uygulamalarda devrim yaratabilir [187]. Ancak, bu YZ tabanlı çözümlerin gerçek dünya uygulamalarında başarıyla kullanılması, sağlık profesyonellerinin eğitilmesi ve güçlendirilmesi için yoğun bir çaba gerektiren önemli bir zorluk olarak karşımıza çıkmaktadır.

YZ tabanlı tıbbi görüntüleme çözümlerinin klinik olarak benimsenmesindeki temel engellerden biri, sağlık hizmeti sunucuları arasında yeterli bilgi ve güven eksikliğidir. Tıbbın uygulaması klinik uzmanlık ve özerkliğe derinden dayandığı için, sağlık profesyonelleri YZ algoritmalarına karar verme yetkisini devretme konusunda isteksiz olabilirler [188]. Bu sorunun çözümü için, teknolojik gelişmeler ile sağlık çalışanlarının pratik ihtiyaçları arasında köprü kuracak eğitim girişimlerine yatırım yapmak hayati önem taşımaktadır [189].

Yakın zamanda yapılan bir çalışmada, radyologların ve tıbbi personelin yapay zekâya yönelik algıları incelenmiş ve radyoloji hizmetlerinin ön saflarında yer alan radyolojik tekniklerin kullanımına ilişkin bilgi eksikliği olduğu tespit edilmiştir. Bu durum, YZ eğitiminin yalnızca klinisyenlerle sınırlı kalmaması gerektiğini ve sağlık ekibinin tüm üyelerinin bu dönüşümsel teknolojilerin entegrasyonuna dair bilgi sahibi olması gerektiğini göstermektedir [190].

Bunun yanı sıra, YZ tabanlı tıbbi görüntüleme çözümlerinin başarılı bir şekilde benimsenmesi, kullanılan algoritmaların temel yapısını, sınırlamalarını ve bu algoritmaların eğitiminde kullanılan verilere özgü önyargı potansiyelini kavramayı gerektirmektedir. Sağlık profesyonellerinin YZ'nin temel prensipleri, düzenleyici çerçeve ve etik boyutları hakkında eğitilmesi, onların YZ destekli çıkarımları değerlendirmelerine, bu teknolojilerin yeteneklerini ve sınırlılıklarını hastalara doğru bir şekilde aktarabilmelerine olanak tanıyacaktır [191].

YZ tabanlı tıbbi görüntüleme konusunda sağlık profesyonellerini eğitme çabaları birkaç temel alana odaklanmalıdır. İlk olarak, YZ'nin temel prensipleri, yetenekleri ve sınırlamaları hakkında kapsamlı bir bilgi sağlanmalıdır. Bu temel bilgi, sağlık hizmeti sunucularının YZ tabanlı araçların uygun kullanımına dair bilinçli kararlar almalarını sağlayarak, teknolojinin aşırı güven ya da yanlış yorumlanma riskini azaltacaktır [192].

İkinci olarak, tıbbi eğitim müfredatı, sağlık bilişimi, bilgisayar bilimi ve istatistik konularını içerecek şekilde güncellenmelidir. Bu, sağlık profesyonellerine YZ algoritmalarının çıktısını yorumlama ve eleştirel analiz etme becerisi kazandırarak, bu çözümlerin klinik iş akışlarına daha etkin bir şekilde entegrasyonunu sağlayacaktır [193]. Ayrıca, müfredatın, YZ nin klinik uygulamalarda kullanımı ile ilgili düzenleyici ve etik hususları da kapsamı gerekmektedir. Sağlık profesyonelleri, veri gizliliği, algoritmaya karşı önyargı ve YZ destekli karar verme süreçlerinin etik sonuçlarına dair değişen yasal düzenlemelerde yol alabilecek bilgiye sahip olmalıdır [194]. Sağlık profesyonellerine gerekli bilgi ve becerileri kazandırarak, YZ tabanlı tıbbi görüntüleme çözümlerinin uygulanmasındaki zorluklar aşılabılır. Sağlık sektörü, YZ 'nin dönüştürücü potansiyeliyle araştırmalarına devam ederken, sağlık profesyonellerinin sürece aktif katılımcı olmaları ve hasta odaklı bakımın geleceğini şekillendirmeleri hayati önem taşımaktadır.

5.6. Çalışmanın Kısıtları ve Gelecek Çalışmalar

Bu çalışmada, AS tespiti için üç farklı YZ yöntemi kullanılarak yenilikçi bir yaklaşım sunulmuş ve bu yöntemlerin performansları karşılaştırılmıştır. Ancak, tüm modellerde kullanılan veri setinin boyutu önemli bir kısıt olarak öne çıkmaktadır. Örneklem büyüklüğünü artırmak amacıyla Haralick özelliklerini kullanan geleneksel makine öğrenimi (ML) teknikleri ile derin öğrenme (DL) tabanlı sistemlerde yapay görüntü iyileştirme veya artırma yöntemleri tercih edilmemiştir. İlk bulgular cesaret verici olsa da gelecekteki araştırmalarda daha büyük veri setlerine ihtiyaç duyulmaktadır. Örneklem büyüklüğünün artırılması, modellerin dayanıklılığını ve genellenebilirliğini artırarak klinik uygulamalarda daha güvenilir sonuçlar elde edilmesine katkı sağlayabilir. Bu çıkarım, YZ tabanlı tanı araçlarının kararlılığının artırılması ve aşırı öğrenmenin önlenmesi için yeterli veri miktarının önemini vurgulayan literatürle de uyumludur.

ML ve Haralick özellikleri ile eğitilen DL sistemi, ÇOA sistemine kıyasla daha açıklanabilir bir yapı sunmaktadır. Bununla birlikte, ÇOA sistemi, hızlı eğitilebilirlik ve önceden eğitilmiş modellerden faydalanma açısından avantajlıdır. Kara kutu niteliğinin azaltılması amacıyla Grad-CAM yöntemi, en hassas sonuçlar veren temsilci model üzerinde kullanılarak modelin hangi bölgeler üzerinden karar verdiği görselleştirilmiştir. YZ tabanlı sistemlerin açıklanabilirliğini artırmak, uzmanların modelin çalışma prensiplerini daha iyi anlamalarını sağlamak ve bu sistemlerin klinik ortamlarda yaygın kullanımını kolaylaştırmak önem arz etmektedir.

DL modelleri genellikle daha yüksek doğruluk oranları sağlarken, ML modelleri daha yüksek yorumlanabilirlik sunmaktadır. Bu durum, klinik uzmanların YZ sistemlerinden elde edilen sonuçları daha kolay anlamalarına ve bu sonuçlara güvenmelerine yardımcı olur. Çalışmamızın hedefi, AS tespiti için düşük hesaplama gücü gereksinimi ile yüksek yorumlanabilirlikle gerçekleştirmektir. Bu nedenle, ML sistemlerinin bu doğrultuda optimize edilmesi ve farklı yaklaşımların incelenmesi önemlidir. Özellikle kaynakların sınırlı olduğu durumlarda, ML tabanlı sistemler tercih edilebilir hale gelmektedir. Bu nedenlerle, en düşük performansı gösteren ML tabanlı sistemin geliştirilmesi için alternatif yaklaşımlara ihtiyaç duyulmaktadır.

Gelecek çalışmalarda, AS tespiti için önerilen çözüm stratejileri kapsamında öncelikle veri setinin büyütülmesi amacıyla görüntü iyileştirme ve artırma tekniklerinin uygulanması planlanmaktadır. Bu sayede mevcut veri setinin boyutu artırılarak modellerin eğitimi için daha fazla örnek sağlanacaktır. İkinci olarak, ML ve DL yaklaşımlarının kombinasyonunu içeren hibrit modeller geliştirilerek her iki yöntemin avantajlarından yararlanılması hedeflenmektedir; bu hem performansı artıracak hem de açıklanabilirliği güçlendirecektir. Üçüncü olarak, Grad-CAM gibi görselleştirme tekniklerinin yanı sıra diğer açıklanabilir YZ yöntemleri kullanılarak, ML ve Haralick özellikleri ile eğitilen DL sisteminin açıklanabilirliği artırılacak ve uzmanların modelin karar verme süreçlerini daha iyi anlamaları sağlanacaktır. Son olarak, kaynakların sınırlı olduğu durumlar göz önünde bulundurularak, ML sistemlerinin düşük hesaplama gücü gereksinimlerinin optimize edilmesi ve farklı yaklaşımların incelenmesi önem taşımaktadır. Bu stratejiler, YZ tabanlı sistemlerin klinik uygulamalarda daha etkin bir şekilde kullanılmasını sağlamak için bütüncül bir yaklaşım sunacaktır.

Bir diğer gelecekteki çalışma, modellerin farklı MRI modlarından elde edilen görüntüleri kullandıklarında performanslarının nasıl değişeceğini araştırmak olacaktır. Bu çerçevede, AS'yi sınıflandırmak için STIR echo MRI görüntüsünün kalitesi, STIR echo ile elde edilmeyen görüntüler kullanılarak değerlendirilecektir. Ayrıca, erişimi kolay bir ara yüz kullanarak geliştirilen modelin STIR MR görüntüleri üretemeyen hastanelerde veya farklı özelliklere sahip cihazlar kullanan hastanelerde uygulanabilirliği araştırılacaktır. Bu çalışma fikrine ek olarak yüksek çözünürlüklü görüntüleme sistemlerine sahip olmayan sağlık birimleri için, elektronik sağlık bilgilerinden derlenecek ve sayısal veriler ile eğitilebilecek denetimli ya da denetimsiz ML tekniklerinin kullanılması da hedeflenmektedir. Böylelikle

ilk deęerlendirme bu sistemlerde yapılıp, hastaların riskli olduęu tespit edilirse görüntüleme cihazlarının olduęu dięer saęlık kuruluşlarına yönlendirilmesi tavsiyesi edilebilecektir.

YZ tabanlı AS alışmalarının klinik pratięe entegrasyonunu artırmak ve kullanılabilirlik ile güvenilirliğini geliřtirmek amacıyla, kısa bir YZ eęitimi verilecek ve ardından AS'yi sınıflandıran yorumlanabilir bir YZ modeli kullanarak deneyimli radyologlarla denemeler gerekleřtirilecektir. Elde edilen sonuçlar belgelenecek ve AS ile YZ literatürüne katkıda bulunmak üzere yayımlanacaktır.

Son olarak, DL'nin güçlü yönlerini dengeleyerek hesaplama gereksinimlerini minimize eden modellerin geliřtirilmesi hedeflenmektedir. Bu bağlamda, derin öğrenme modellerini eęitmek için makine öğreniminde geleneksel olarak kullanılan özellik çıkarma teknikleri, sayısal özellik setleri gibi yöntemlerle uygulanabilir. Yapılandırılmış girdi verilerine dayanmak, büyük ve yüksek çözünürlüklü görüntüler yerine eęitim süresini ve hesaplama yükünü önemli ölçüde azaltabilir. Bu tür hibrit yaklaşımların, DL tabanlı sistemlerinin öngörücü doęruluęunu korurken gerek dünya uygulamaları için daha pratik hale gelen modellere yol açabileceęi düşünölmektedir.

KAYNAKLAR

1. Khan MA. (2002). Update on spondyloarthropathies. *Ann Intern Med*, 136(12), 896-907.
2. Sieper J, van der Heijde D, Dougados M, Mease PJ, Maksymowych WP, Brown MA, et al. (2013). Efficacy and safety of adalimumab in patients with non-radiographic axial spondyloarthritis: results of a randomised placebo-controlled trial (ABILITY-1). *Annals of the rheumatic diseases*, 72(6), 815-22.
3. Deodhar A, van der Heijde D, Gensler LS, Kim T-H, Maksymowych WP, Østergaard M, et al. (2020). Ixekizumab for patients with non-radiographic axial spondyloarthritis (COAST-X): a randomised, placebo-controlled trial. *The Lancet*, 395(10217), 53-64.
4. Yao Y, Yao J, Butz C, Lingras P, Jutla D, editors. *Web-based Support Systems (WSS): A Report of the WIC Canada Research Centre*. IEEE/WIC/ACM International Conference on Web Intelligence (WI'04); 2004: IEEE.
5. Moreland LW, Baumgartner SW, Schiff MH, Tindall EA, Fleischmann RM, Weaver AL, et al. (1997). Treatment of rheumatoid arthritis with a recombinant human tumor necrosis factor receptor (p75)-Fc fusion protein. *New England Journal of Medicine*, 337(3), 141-7.
6. Çeliker R. (2000). Ankilozan Spondilit: Klinik özellikleri. *Romatizma*, 15(1), 15-21.
7. Chen M, Gholamrezanezhad A. (2023). Burnout in radiology. *Academic Radiology*, 30(6), 1031-2.
8. Baraliakos X, Gossec L, Pournara E, Jeka S, Mera-Varela A, d'Angelo S, et al. (2021). Secukinumab in patients with psoriatic arthritis and axial manifestations: results from the double-blind, randomised, phase 3 MAXIMISE trial. *Annals of the rheumatic diseases*, 80(5), 582-90.
9. Weber U, Lambert RG, Østergaard M, Hodler J, Pedersen SJ, Maksymowych WP. (2010). The diagnostic utility of magnetic resonance imaging in spondylarthritis: an international multicenter evaluation of one hundred eighty-seven subjects. *Arthritis & Rheumatism*, 62(10), 3048-58.
10. van Hooft L, Luime J, Han H, Vergouwe Y, Weel A. (2014). Identifying axial spondyloarthritis in Dutch primary care patients, ages 20–45 years, with chronic low back pain. *Arthritis care & research*, 66(3), 446-53.

11. Li H, Tao X, Liang T, Jiang J, Zhu J, Wu S, et al. (2023). Comprehensive AI-assisted tool for ankylosing spondylitis based on multicenter research outperforms human experts. *Frontiers in Public Health*, 11, 1063633.
12. Sieper J, Rudwaleit M, Baraliakos X, Brandt J, Braun J, Burgos-Vargas R, et al. (2009). The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. *Annals of the Rheumatic Diseases*, 68(Suppl 2), ii1-ii44.
13. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Annals of internal medicine*, 169(7), 467-73.
14. Shenkman Y, Qutteineh B, Joskowicz L, Szeskin A, Yusef A, Mayer A, et al. (2019). Automatic detection and diagnosis of sacroiliitis in CT scans as incidental findings. *Medical Image Analysis*, 57, 165-75.
15. Castro-Zunti R, Park EH, Choi Y, Jin GY, Ko S-b. (2020). Early detection of ankylosing spondylitis using texture features and statistical machine learning, and deep learning, with some patient age analysis. *Computerized Medical Imaging and Graphics*, 82, 101718.
16. Faleiros MC, Nogueira-Barbosa MH, Dalto VF, Júnior JRF, Tenório APM, Luppino-Assad R, et al. (2020). Machine learning techniques for computer-aided classification of active inflammatory sacroiliitis in magnetic resonance imaging. *Advances in Rheumatology*, 60, 1-10.
17. Han Q, Lu Y, Han J, Luo A, Huang L, Ding J, et al. (2021). Automatic quantification and grading of hip bone marrow oedema in ankylosing spondylitis based on deep learning. *Modern Rheumatology*.
18. Tas NP, Kaya O, Macin G, Tasci B, Dogan S, Tuncer T. (2023). ASNET: A Novel AI Framework for Accurate Ankylosing Spondylitis Diagnosis from MRI. *Biomedicines*, 11(9), 2441.
19. Lin KYY, Peng C, Lee KH, Chan SCW, Chung HY. (2022). Deep learning algorithms for magnetic resonance imaging of inflammatory sacroiliitis in axial spondyloarthritis. *Rheumatology*, 61(10), 4198-206.
20. Zhang K, Luo G, Li W, Zhu Y, Pan J, Li X, et al. (2023). Automatic image segmentation and grading diagnosis of sacroiliitis associated with AS using a deep convolutional neural network on CT images. *Journal of Digital Imaging*, 36(5), 2025-34.
21. Prasad KS, Jana S, editors. Survey of Ankylosing Spondylitis for Biomedical Imaging with Deep Neural Networks. 2024 International Conference on Signal Processing,

Computation, Electronics, Power and Telecommunication (IConSCEPT); 2024 4-5 July 2024.

22. Diekhoff T, Hermann KG, Greese J, Schwenke C, Poddubnyy D, Hamm B, et al. (2017). Comparison of MRI with radiography for detecting structural lesions of the sacroiliac joint using CT as standard of reference: results from the SIMACT study. *Ann Rheum Dis*, 76(9), 1502-8.
23. Chen M, Elawad A, Herregonds N, Jans L, Rennie W. (2020). Cutting edge technologies in the imaging of spondyloarthritis. *Indian Journal of Rheumatology*, 15(5), 27-33.
24. Tsoi C, Griffith JF, Lee RKL, Wong PCH, Tam LS. (2019). Imaging of sacroiliitis: Current status, limitations and pitfalls. *Quant Imaging Med Surg*, 9(2), 318-35.
25. Altıkardeş ZA, Canayaz E, Ünsal A. (2023). Machine LearningDeep Learning in Rheumatological Screening A Systematic Review. *Erzincan University Journal of Science and Technology*, 16(3), 940-69.
26. Feldtkeller E, Lemmel E-M, Russell AS. (2003). Ankylosing spondylitis in the pharaohs of ancient Egypt. *Rheumatology international*, 23, 1-5.
27. Spencer DG, Sturrock RD, Buchanan WW. (1980). Ankylosing spondylitis: yesterday and today. *Medical History*, 24(1), 60-9.
28. Bekhterev VM. (2017). *Collective reflexology: The complete edition*, ed.: Routledge.
29. Bywaters E. (1979). Symposium on the spondylarthritideshistorical aspects of ankylosing spondylitis. *Rheumatology*, 18(4), 197-203.
30. Marie P. (1898). *Sur la spondylose rhizomélisque: publ. par les Laboratoires Boots-Dacour*, ed.: L. Pariente.
31. Strümpell A. (1897). Bemerkung über die chronische ankylosirende Entzündung der Wirbelsäule und der Hüftgelenke. *Deutsche Zeitschrift für Nervenheilkunde*, 11, 338-42.
32. Dean LE, Jones GT, MacDonald AG, Downham C, Sturrock RD, Macfarlane GJ. (2014). Global prevalence of ankylosing spondylitis. *Rheumatology (Oxford)*, 53(4), 650-7.
33. Carrera G, Foley W, Kozin F, Ryan L, Lawson T. (1981). CT of sacroiliitis. *American Journal of Roentgenology*, 136(1), 41-6.

34. Gordon T, Sage M, Bertouch J, Brooks P. (1984). Computed tomography of paraspinal musculature in ankylosing spondylitis. *The Journal of Rheumatology*, 11(6), 794-7.
35. Kenny J, Hughes P, Whitehouse G. (1990). Discovertebral destruction in ankylosing spondylitis: the role of computed tomography and magnetic resonance imaging. *The British Journal of Radiology*, 63(750), 448-55.
36. Ahlström H, Feltelius N, Nyman R, Hällgren R. (1990). Magnetic resonance imaging of sacroiliac joint inflammation. *Arthritis Rheum*, 33(12), 1763-9.
37. Brown MA, Kenna T, Wordsworth BP. (2016). Genetics of ankylosing spondylitis—insights into pathogenesis. *Nature Reviews Rheumatology*, 12(2), 81-91.
38. Braun J, Sieper J. (2007). Ankylosing spondylitis. *The Lancet*, 369(9570), 1379-90.
39. Karataş A, Öz B, Akar Za, Dönder E, Koca Ss. (2019). Aksiyel Spondilartritlerin Tanı ve Tedavisi. *Fırat Üniversitesi Sağlık Bilimleri Tıp Dergisi*, 33(2), 119-24.
40. Van Der Heijde D, Ramiro S, Landewé R, Baraliakos X, Van den Bosch F, Sepriano A, et al. (2017). 2016 update of the ASAS-EULAR management recommendations for axial spondyloarthritis. *Annals of the rheumatic diseases*, 76(6), 978-91.
41. Haroon M, O'Rourke M, Ramasamy P, Murphy CC, FitzGerald O. (2015). A novel evidence-based detection of undiagnosed spondyloarthritis in patients presenting with acute anterior uveitis: the DUET (Dublin Uveitis Evaluation Tool). *Ann Rheum Dis*, 74(11), 1990-5.
42. Fragoulis GE, Liava C, Daoussis D, Akriviadis E, Garyfallos A, Dimitroulas T. (2019). Inflammatory bowel diseases and spondyloarthropathies: from pathogenesis to treatment. *World journal of gastroenterology*, 25(18), 2162.
43. de Winter JJ, van Mens LJ, van der Heijde D, Landewé R, Baeten DL. (2016). Prevalence of peripheral and extra-articular disease in ankylosing spondylitis versus non-radiographic axial spondyloarthritis: a meta-analysis. *Arthritis research & therapy*, 18, 1-11.
44. Mathieu S, Pereira B, Soubrier M. (2015). Cardiovascular events in ankylosing spondylitis: an updated meta-analysis. *Semin Arthritis Rheum*, 44(5), 551-5.
45. Robinson Y. (2017). Spinal fractures related to ankylosing spondylitis: Epidemiology, clinical outcome and biomechanics. Institution, Pages.

46. Ward MM, Deodhar A, Gensler LS, Dubreuil M, Yu D, Khan MA, et al. (2019). 2019 update of the American College of Rheumatology/Spondylitis Association of America/Spondyloarthritis Research and Treatment Network recommendations for the treatment of ankylosing spondylitis and nonradiographic axial spondyloarthritis. *Arthritis & rheumatology*, 71(10), 1599-613.
47. Tuncer T, Gilgil E, Kaçar C, Kurtaiş Y, Kutlay Ş, Bütün B, et al. (2018). Prevalence of Rheumatoid Arthritis and Spondyloarthritis in Turkey: A Nationwide Study. *Arch Rheumatol*, 33(2), 128-36.
48. Baser O, Burkan A, Baser E, Koselerli R, Ertugay E, Altinbas A. (2013). Health care costs associated with ankylosing spondylitis in Turkey: an analysis from nationwide real-world data. *Int J Rheumatol*, 2013, 139608.
49. Sekino M, Ueno S. *Magnetic Resonance Imaging: Principles and Applications*. Bioimaging: CRC Press; 2020. p. 75-100.
50. Berger A. (2002). Magnetic resonance imaging. *Bmj*, 324(7328), 35.
51. van Geuns R-JM, Wielopolski PA, de Bruin HG, Rensing BJ, van Ooijen PMA, Hulshoff M, et al. (1999). Basic principles of magnetic resonance imaging. *Progress in Cardiovascular Diseases*, 42(2), 149-56.
52. Pai A, Shetty R, Chowdhury Y. (2020). Magnetic resonance imaging physics.
53. Boonen A, Brinkhuizen T, Landewé R, van der Heijde D, Severens JL. (2010). Impact of ankylosing spondylitis on sick leave, presenteeism and unpaid productivity, and estimation of the societal cost. *Annals of the Rheumatic Diseases*, 69(6), 1123-8.
54. Brown RW, Cheng Y-CN, Haacke EM, Thompson MR, Venkatesan R. (2014). *Magnetic resonance imaging: physical principles and sequence design*, ed.: John Wiley & Sons.
55. Vlaardingerbroek MT, Boer JA. (2013). *Magnetic resonance imaging: theory and practice*, ed.: Springer Science & Business Media.
56. Grover VP, Tognarelli JM, Crossey MM, Cox IJ, Taylor-Robinson SD, McPhail MJ. (2015). Magnetic resonance imaging: principles and techniques: lessons for clinicians. *Journal of clinical and experimental hepatology*, 5(3), 246-55.
57. Marcu CB, Beek AM, Van Rossum AC. (2006). Clinical applications of cardiovascular magnetic resonance imaging. *Cmaj*, 175(8), 911-7.

58. Xue Y, Du Y, Carloni G, Pachetti E, Jordan C, Tsiftaris SA, editors. Cine cardiac MRI reconstruction using a convolutional recurrent network with refinement. International Workshop on Statistical Atlases and Computational Models of the Heart; 2023: Springer.
59. Bradley Jr WG, editor Current And Future Indications For Magnetic Resonance In Medicine. Diagnostic Imaging Applications; 1985: SPIE.
60. Siddique MAB, Sakib S, Khan MMR, Tanzeem AK, Chowdhury M, Yasmin N, editors. Deep convolutional neural networks model-based brain tumor detection in brain MRI images. 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC); 2020: IEEE.
61. Dahlqvist JR, Salim R, Thomsen C, Vissing J. (2020). A quantitative method to assess muscle edema using short TI inversion recovery MRI. Scientific Reports, 10(1), 7246.
62. Raghavan M, Palestro CJ, editors. Imaging of spondylodiscitis: an update. Seminars in Nuclear Medicine; 2023: Elsevier.
63. Turing AM. (2009). Computing machinery and intelligence, ed.: Springer.
64. McCarthy J, Minsky ML, Rochester N, Shannon CE. (2006). A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. AI magazine, 27(4), 12-.
65. Kirsh D. (1991). Foundations of AI: the big issues. Artificial intelligence, 47(1-3), 3-30.
66. Russell SJ, Norvig P. (2016). Artificial intelligence: a modern approach, ed.: Pearson.
67. Brachman RJ. (2006). AI more than the sum of its parts. AI Magazine, 27(4), 19-.
68. Monett D, Lewis CW, editors. Getting clarity by defining artificial intelligence—A survey. Philosophy and theory of artificial intelligence 2017; 2018: Springer.
69. Wang P. (2019). On defining artificial intelligence. Journal of Artificial General Intelligence, 10(2), 1-37.
70. Samuel AL. (2000). Some studies in machine learning using the game of checkers. IBM Journal of research and development, 44(1.2), 206-26.
71. Park C, Took CC, Seong J-K. (2018). Machine learning in biomedical engineering. Biomedical Engineering Letters, 8, 1-3.

72. El Naqa I, Murphy MJ. What is machine learning? machine learning in radiation oncology: Springer; 2015. p. 3-11.
73. Wang H, Ma C, Zhou L, editors. A brief review of machine learning and its application. 2009 international conference on information engineering and computer science; 2009: IEEE.
74. Mitchell T. (1997). Machine Learning. New York: McGraw-Hill Companies Inc.
75. Alpaydin E. Introduction to machine learning, 3rd editio. ed. The MIT Press; 2014.
76. Bishop CM, Nasrabadi NM. (2006). Pattern recognition and machine learning, 4 ed.: Springer.
77. Jordan MI, Mitchell TM. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255-60.
78. Deodhar A, Rozycki M, Garges C, Shukla O, Arndt T, Grabowsky T, et al. (2020). Use of machine learning techniques in the development and refinement of a predictive model for early diagnosis of ankylosing spondylitis. Clinical Rheumatology, 39(4), 975-82.
79. Dhillon A, Singh A. (2019). Machine learning in healthcare data analysis: a survey. Journal of Biology and Today's World, 8(6), 1-10.
80. Paluszczek M, Thomas S. (2016). MATLAB machine learning, ed.: Apress.
81. Cortes C, Vapnik V. (1995). Support-vector networks. Machine Learning, 20(3), 273-97.
82. Burges CJ. (1998). A tutorial on support vector machines for pattern recognition. Data mining and knowledge discovery, 2(2), 121-67.
83. Osuna EE. (1998). Support vector machines: Training and applications. Institution, Pages.
84. Kavzoğlu T, Çölkesen İ. (2010). Destek vektör makineleri ile uydu görüntülerinin sınıflandırılmasında kernel fonksiyonlarının etkilerinin incelenmesi. Harita Dergisi, 144(7), 73-82.
85. Scholkopf B, Smola AJ. (2018). Learning with kernels: support vector machines, regularization, optimization, and beyond, ed.: MIT press.

86. Cristianini N. (2000). An Introduction to Support Vector Machines and other kernel-based learning methods, ed.: Cambridge University Press.
87. Campbell C, Ying Y. (2022). Learning with support vector machines, ed.: Springer Nature.
88. Cover T, Hart P. (1967). Nearest neighbor pattern classification. IEEE transactions on information theory, 13(1), 21-7.
89. Altman NS. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. The American Statistician, 46(3), 175-85.
90. Aha DW, Kibler D, Albert MK. (1991). Instance-based learning algorithms. Machine learning, 6, 37-66.
91. Ortiz-Bejar J, Téllez ES, Graff M, Moctezuma D, Miranda-Jiménez S. (2020). Improving k Nearest Neighbors and Naïve Bayes Classifiers Through Space Transformations and Model Selection. IEEE Access, 8, 221669-88.
92. Hastie T, Tibshirani R, Friedman JH, Friedman JH. (2009). The elements of statistical learning: data mining, inference, and prediction, 2 ed.: Springer.
93. Dietterich TG, editor Ensemble methods in machine learning. International workshop on multiple classifier systems; 2000: Springer.
94. Zhou Z-H. (2012). Ensemble methods: foundations and algorithms, ed.: CRC press.
95. Breiman L. (2001). Random forests. Machine learning, 45, 5-32.
96. Ho TK, editor Random decision forests. Proceedings of 3rd international conference on document analysis and recognition; 1995: IEEE.
97. Golden CE, Rothrock Jr MJ, Mishra A. (2019). Comparison between random forest and gradient boosting machine methods for predicting *Listeria* spp. prevalence in the environment of pastured poultry farms. Food research international, 122, 47-55.
98. Lin X, Yacoub S, Burns J, Simske S. (2003). Performance analysis of pattern classifier combination by plurality voting. Pattern Recognition Letters, 24(12), 1959-69.
99. Gandhi I, Pandey M, editors. Hybrid ensemble of classifiers using voting. 2015 international conference on green computing and Internet of Things (ICGCIoT); 2015: IEEE.

100. Muñoz-Aseguinolaza U, Sierra B, Aginako N. (2023). Rotational augmentation techniques: a new perspective on ensemble learning for image classification. arXiv preprint arXiv:230607027.
101. Mienye ID, Sun Y. (2022). A survey of ensemble learning: Concepts, algorithms, applications, and prospects. IEEE Access, 10, 99129-49.
102. Lowry OH, Rosebrough NJ, Farr AL, Randall RJ. (1951). Protein measurement with the Folin phenol reagent. J biol Chem, 193(1), 265-75.
103. LeCun Y, Bengio Y, Hinton G. (2015). Deep learning. nature, 521(7553), 436-44.
104. Raturi T, Memoria M, Kumar T, editors. Survey on Statistical Prediction Models Based on Intelligent Learning. International Conference on Advances in Engineering Science Management & Technology (ICAESMT)-2019, Uttarakhand University, Dehradun, India; 2019.
105. Abraham A. (2005). Artificial neural networks. Handbook of measuring system design.
106. Zhang W, Yang G, Lin Y, Ji C, Gupta MM, editors. On definition of deep learning. 2018 World automation congress (WAC); 2018: IEEE.
107. Bengio Y. Learning Deep Architectures for AI. Now Publishers Inc; 2009.
108. Bengio Y, Courville A, Vincent P. (2013). Representation learning: A review and new perspectives. IEEE transactions on pattern analysis and machine intelligence, 35(8), 1798-828.
109. Schmidhuber J. (2015). Deep learning in neural networks: An overview. Neural networks, 61, 85-117.
110. Deng L, Yu D. (2014). Deep learning: methods and applications. Foundations and trends® in signal processing, 7(3-4), 197-387.
111. Beale MH, Hagan MT, Demuth HB. (2018). Deep learning toolbox. R2018b User's Guide; The MathWorks, Inc: Natick, MA, USA.
112. Goodfellow I, Bengio Y, Courville A, Bengio Y. (2016). Deep learning, 1 ed.: MIT press Cambridge.
113. Kim HE, Cosa-Linan A, Santhanam N, Jannesari M, Maros ME, Ganslandt T. (2022). Transfer learning for medical image classification: a literature review. BMC medical imaging, 22(1), 69.

114. Aneja N, Aneja S, editors. Transfer learning using CNN for handwritten devanagari character recognition. 2019 1st International Conference on Advances in Information Technology (ICAIT); 2019: IEEE.
115. Simonyan K, Zisserman A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
116. He K, Zhang X, Ren S, Sun J, editors. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition; 2016.
117. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z, editors. Rethinking the inception architecture for computer vision. Proceedings of the IEEE conference on computer vision and pattern recognition; 2016.
118. Ronneberger O, Fischer P, Brox T, editors. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention; 2015: Springer.
119. Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, et al. (2016). Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE transactions on medical imaging, 35(5), 1299-312.
120. Yamashita R, Nishio M, Do RKG, Togashi K. (2018). Convolutional neural networks: an overview and application in radiology. Insights into imaging, 9, 611-29.
121. Breiman L. (2001). Statistical modeling: The two cultures (with comments and a rejoinder by the author). Statistical science, 16(3), 199-231.
122. Zhang Q, Wu YN, Zhu S-C, editors. Interpretable convolutional neural networks. Proceedings of the IEEE conference on computer vision and pattern recognition; 2018.
123. Zheng A. (2018). Feature engineering for machine learning: principles and techniques for data scientists, ed.: O'Reilly Media, Inc.
124. Murdoch WJ, Singh C, Kumbier K, Abbasi-Asl R, Yu B. (2019). Interpretable machine learning: definitions, methods, and applications. arXiv preprint arXiv:1901.04592.
125. Zeiler MD, Fergus R, editors. Visualizing and understanding convolutional networks. Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part I 13; 2014: Springer.
126. Rudin C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature machine intelligence, 1(5), 206-15.

127. Doshi-Velez F, Kim B. (2017). Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:170208608.
128. Lipton ZC. (2018). The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue*, 16(3), 31-57.
129. Tjoa E, Guan C. (2020). A survey on explainable artificial intelligence (xai): Toward medical xai. *IEEE transactions on neural networks and learning systems*, 32(11), 4793-813.
130. Van der Velden BH, Kuijf HJ, Gilhuijs KG, Viergever MA. (2022). Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Medical Image Analysis*, 79, 102470.
131. Lambert B, Forbes F, Doyle S, Dehaene H, Dojat M. (2024). Trustworthy clinical AI solutions: a unified review of uncertainty quantification in deep learning models for medical image analysis. *Artificial Intelligence in Medicine*, 102830.
132. Qiu Z, Rivaz H, Xiao Y, editors. Is visual explanation with Grad-CAM more reliable for deeper neural networks? A case study with automatic pneumothorax diagnosis. *International Workshop on Machine Learning in Medical Imaging*; 2023: Springer.
133. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D, editors. Grad-cam: Visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE international conference on computer vision*; 2017.
134. Rahimi A, Recht B. (2007). Random features for large-scale kernel machines. *Advances in neural information processing systems*, 20.
135. Domingos P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, 55(10), 78-87.
136. Krizhevsky A, Sutskever I, Hinton GE. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84-90.
137. Shwartz-Ziv R, Tishby N. (2017). Opening the black box of deep neural networks via information. arXiv preprint arXiv:170300810.
138. Kaplan J, McCandlish S, Henighan T, Brown TB, Chess B, Child R, et al. (2020). Scaling laws for neural language models. arXiv preprint arXiv:200108361.
139. Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, et al. (2015). Human-level control through deep reinforcement learning. *nature*, 518(7540), 529-33.

140. Han S, Mao H, Dally WJ. (2015). Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding. arXiv preprint arXiv:151000149.
141. Khan A, Sohail A, Zahoora U, Qureshi AS. (2020). A survey of the recent architectures of deep convolutional neural networks. *Artificial intelligence review*, 53, 5455-516.
142. Baeza-Delgado C, Cerdá Alberich L, Carot-Sierra JM, Veiga-Canuto D, Martínez de las Heras B, Raza B, et al. (2022). A practical solution to estimate the sample size required for clinical prediction models generated from observational research on data. *European Radiology Experimental*, 6(1), 22.
143. Hwan-Ho C, Hyunjin P. (2017). Classification of low-grade and high-grade glioma using multi-modal image radiomics features. *Annu Int Conf IEEE Eng Med Biol Soc*, 2017, 3081-4.
144. Hernández-Vázquez MA, Hernández-Rodríguez YM, Cortes-Rojas FD, Bayareh-Mancilla R, Cigarroa-Mayorga OE. (2024). Hybrid Feature Mammogram Analysis: Detecting and Localizing Microcalcifications Combining Gabor, Prewitt, GLCM Features, and Top Hat Filtering Enhanced with CNN Architecture. *Diagnostics*, 14(15), 1691.
145. Ranjitha K, Pushphavathi T. (2024). Analysis on Improved Gaussian-Wiener filtering technique and GLCM based Feature Extraction for Breast Cancer Diagnosis. *Procedia Computer Science*, 235, 2857-66.
146. Tarek Y, Elgohary R, Deif M. (2024). Brain Tumor Detection Using GLCM and Machine learning Techniques. *International Integrated Intelligent Systems*, 1(2).
147. Pagadala PK, D Yaso O, Lakshmi PS, Dewi C, Sutresno SA, Christanto HJ. (2024). Automated Brain Tumor Detection With GLCM-Based Feature Extraction and PCA For Dimension Reduction and Classification Using Machine Learning. *Journal of Theoretical and Applied Information Technology*, 102(12).
148. Ahmed A, Gibbs P, Pickles M, Turnbull L. (2013). Texture analysis in assessment and prediction of chemotherapy response in breast cancer. *Journal of Magnetic Resonance Imaging*, 38(1), 89-101.
149. Sun W, Zeng N, He Y. (2019). Morphological arrhythmia automated diagnosis method using gray-level co-occurrence matrix enhanced convolutional neural network. *IEEE Access*, 7, 67123-9.

150. Hu S, Xu C, Guan W, Tang Y, Liu Y. (2014). Texture feature extraction based on wavelet transform and gray-level co-occurrence matrices applied to osteosarcoma diagnosis. *Bio-medical materials and engineering*, 24(1), 129-43.
151. Rodrigues MB, Da Nobrega RVM, Alves SSA, Reboucas Filho PP, Duarte JBF, Sangaiah AK, et al. (2018). Health of things algorithms for malignancy level classification of lung nodules. *IEEE Access*, 6, 18592-601.
152. Zaina HS, Belhaouari SB, Stanko T, Gorovoy V. (2022). An exemplar pyramid feature extraction based Alzheimer disease classification method. *IEEE Access*, 10, 66511-21.
153. Ryman SG, Yutsis M, Tian L, Henderson VW, Montine TJ, Salmon DP, et al. (2021). Cognition at each stage of Lewy body disease with co-occurring Alzheimer's disease pathology. *Journal of Alzheimer's Disease*, 80(3), 1243-56.
154. Bing W, Zhang X, Wang D, Gu X. (2023). Clinical value of CT imaging features in the diagnosis of acute and chronic pancreatitis: A retrospective study. *Technology and Health Care*, (Preprint), 1-9.
155. Sagayam KM. (2024). An automated cervical cancer diagnosis using genetic algorithm and CANFIS approaches. *Technology and Health Care*, (Preprint), 1-17.
156. Uyun S, Barkah NM, Putri IE, Faridah N. (2024). A Systematic Literature Review on the Methods of Breast Cancer Classification. *Engineering Headway*, 6, 115-23.
157. Patel RK, Kashyap M. (2023). Automated screening of glaucoma stages from retinal fundus images using BPS and LBP based GLCM features. *International Journal of Imaging Systems and Technology*, 33(1), 246-61.
158. Zaccaria GM, Berloco F, Buongiorno D, Brunetti A, Altini N, Bevilacqua V. (2024). A time-dependent explainable radiomic analysis from the multi-omic cohort of CPTAC-Pancreatic Ductal Adenocarcinoma. *Computer Methods and Programs in Biomedicine*, 108408.
159. Wu Q, Gan Y, Lin B, Zhang Q, Chang H. (2015). An active contour model based on fused texture features for image segmentation. *Neurocomputing*, 151, 1133-41.
160. Wahir N, Nor M, Rusiman M, Gopal K, editors. Treatment of outliers via interpolation method with neural network forecast performances. *Journal of Physics: Conference Series*; 2018: IOP Publishing.

161. Soh L-K, Tsatsoulis C. (1999). Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. *IEEE Transactions on geoscience and remote sensing*, 37(2), 780-95.
162. Haralick RM, Shanmugam K, Dinstein I. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6), 610-21.
163. Clausi D. (2002). An analysis of co-occurrence texture statistics using gray level co-occurrence matrices. *Can J Remote Sens/J Can de Teledetect*, 28(1), 45-62.
164. Nayak SK, Ojha AC. (2020). Data leakage detection and prevention: Review and research directions. *Machine Learning and Information Processing: Proceedings of ICMLIP 2019*, 203-12.
165. Bendre N, Marín HT, Najafirad P. (2020). Learning from few samples: A survey. *arXiv preprint arXiv:200715484*.
166. Coelho IM, Coelho VN, Luz EJdS, Ochi LS, Guimaraes FG, Rios E. (2017). A GPU deep learning metaheuristic based model for time series forecasting. *Applied Energy*, 201, 412-8.
167. Ruff L, Kauffmann JR, Vandermeulen RA, Montavon G, Samek W, Kloft M, et al. (2021). A unifying review of deep and shallow anomaly detection. *Proceedings of the IEEE*, 109(5), 756-95.
168. Chen C, Zhang P, Zhang H, Dai J, Yi Y, Zhang H, et al. (2020). Deep learning on computational-resource-limited platforms: A survey. *Mobile Information Systems*, 2020(1), 8454327.
169. Ahmed SF, Alam MSB, Kabir M, Afrin S, Rafa SJ, Mehjabin A, et al. (2023). Unveiling the frontiers of deep learning: innovations shaping diverse domains. *arXiv preprint arXiv:230902712*.
170. Gupta A, Singh A. (2023). Prediction framework on early urine infection in IoT–Fog environment using XGBoost ensemble model. *Wireless Personal Communications*, 131(2), 1013-31.
171. Abe T, Buchanan EK, Pleiss G, Cunningham JP. (2023). Pathologies of predictive diversity in deep ensembles. *arXiv preprint arXiv:230200704*.
172. Li W, Paffenroth RC, Berthiaume D. (2021). Neural network ensembles: theory, training, and the importance of explicit diversity. *arXiv preprint arXiv:210914117*.
173. Sagi O, Rokach L. (2018). Ensemble learning: A survey. *Wiley interdisciplinary reviews: data mining and knowledge discovery*, 8(4), e1249.

174. Khalifa NE, Loey M, Mirjalili S. (2022). A comprehensive survey of recent trends in deep learning for digital images augmentation. *Artificial Intelligence Review*, 55(3), 2351-77.
175. Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H. (2018). GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing*, 321, 321-31.
176. Taylor L, Nitschke G. (2017). Improving deep learning using generic data augmentation. *arXiv preprint arXiv:170806020*.
177. Howard AG. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:170404861*.
178. Tan M, Le Q, editors. Efficientnet: Rethinking model scaling for convolutional neural networks. *International conference on machine learning*; 2019: PMLR.
179. Gupta S, Agrawal A, Gopalakrishnan K, Narayanan P, editors. Deep learning with limited numerical precision. *International conference on machine learning*; 2015: PMLR.
180. Madani A, Ong JR, Tibrewal A, Mofrad MR. (2018). Deep echocardiography: data-efficient supervised and semi-supervised deep learning towards automated diagnosis of cardiac disease. *NPJ digital medicine*, 1(1), 1-11.
181. Rajpurkar P, Park A, Irvin J, Chute C, Bereket M, Mastrodicasa D, et al. (2020). AppendiXNet: deep learning for diagnosis of appendicitis from a small dataset of CT exams using video pretraining. *Scientific reports*, 10(1), 3958.
182. Willemink MJ, Koszek WA, Hardell C, Wu J, Fleischmann D, Harvey H, et al. (2020). Preparing medical imaging data for machine learning. *Radiology*, 295(1), 4-15.
183. Giuste F, Shi W, Zhu Y, Naren T, Isgut M, Sha Y, et al. (2022). Explainable artificial intelligence methods in combating pandemics: A systematic review. *IEEE Reviews in Biomedical Engineering*, 16, 5-21.
184. Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. (2018). Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data. *JAMA Internal Medicine*, 178(11), 1544-7.
185. Vayena E, Blasimme A, Cohen IG. (2018). Machine learning in medicine: addressing ethical challenges. *PLoS medicine*, 15(11), e1002689.
186. Panch T, Mattie H, Atun R. (2019). Artificial intelligence and algorithmic bias: implications for health systems. *Journal of global health*, 9(2).

187. Law M, Seah J, Shih G. (2021). Artificial intelligence and medical imaging: applications, challenges and solutions. *The Medical Journal of Australia*, 214(10), 450-2. e1.
188. Sivaraman V, Bukowski LA, Levin J, Kahn JM, Perer A, editors. *Ignore, trust, or negotiate: understanding clinician acceptance of AI-based treatment recommendations in health care. Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*; 2023.
189. Bajwa J, Munir U, Nori A, Williams B. (2021). Artificial intelligence in healthcare: transforming the practice of medicine. *Future healthcare journal*, 8(2), e188-e94.
190. Pesapane F, Codari M, Sardanelli F. (2018). Artificial intelligence in medical imaging: threat or opportunity? Radiologists again at the forefront of innovation in medicine. *European radiology experimental*, 2, 1-10.
191. Rainey C, O'Regan T, Matthew J, Skelton E, Woznitza N, Chu K-Y, et al. (2021). Beauty is in the AI of the beholder: are we ready for the clinical integration of artificial intelligence in radiography? An exploratory analysis of perceived AI knowledge, skills, confidence, and education perspectives of UK radiographers. *Frontiers in digital health*, 3, 739327.
192. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. (2019). The practical implementation of artificial intelligence technologies in medicine. *Nature medicine*, 25(1), 30-6.
193. Lin SY, Mahoney MR, Sinsky CA. (2019). Ten ways artificial intelligence will transform primary care. *Journal of general internal medicine*, 34, 1626-30.
194. Gerke S, Minssen T, Cohen G. Ethical and legal challenges of artificial intelligence-driven healthcare. *Artificial intelligence in healthcare: Elsevier*; 2020. p. 295-336.

ÖZGEÇMİŞ

Adı Soyadı :Emre ÇANAYAZ

Öğrenim Durumu

Derece	Bölüm/ Program	Üniversite	Yılı
Önlisans	Biyomedikal Cihaz Teknolojisi	Marmara Üniversitesi	2000-2002
Lisans	Elektrik- Elektronik Mühendisliği	Eskişehir Osmangazi Üniversitesi	2003-2008
Y.Lisans	Biyomedikal Mühendisliği	İstanbul Teknik Üniversitesi	2010-2013

İş Deneyimi

Görevi	Üniversite	Yılı
Araştırma Görevlisi	Marmara Üniversitesi/ Teknik Bilimler Meslek Yüksekokulu/ Elektronik ve Otomasyon Programı	2010-2022
Öğretim Görevlisi	Marmara Üniversitesi/ Teknik Bilimler Meslek Yüksekokulu/ Elektronik ve Otomasyon Programı	2022-

Bilimsel Eserler:

1. SCI, SSCI ve AHCI İndekslerine Giren Dergilerde Yayınlanan Makaleler

1. Korkmaz H., Çanayaz E., Birtane S., Altıkardes A. (2019). Fuzzy logic based risk assessment system giving individualized advice for metabolic syndrome and fatal cardiovascular diseases. TECHNOLOGY AND HEALTH CARE, Volume 27, (SCI-Expanded).
2. Güçlü B., Tanıdır C., Canayaz E., Guner B., Toz H. I., Uneri O. S., et al. (2015). TaBTile processing in children and adolescents with obsessive-compulsive disorder.

SOMATOSENSORY AND MOTOR RESEARCH, Volume 32, Issue 3, pp.163-171. (SCI-Expanded).

2. Diğer Dergilerde Yayımlanan Makaleler

1. Altıkardes Z. A., Canayaz E., Ünsal A. (2023). Machine Learning and Deep Learning in Rheumatological Screening: A Systematic Review. Erzincan Üniversitesi Fen Bilimleri Enstitüsü Dergisi, Volume 16, Issue 3, pp.940-969. (TRDizin)
2. Çanayaz E., Tektaş M. (2018). Evaluation of *Neoergasilus japonicus* (Ergasilidae) Occurrence on Rudd Scardinius erythrophthalmus Using Fuzzy Logic Approach. ERZİNCAN FBE DERGİSİ, Volume 11, Issue 2, pp.277-284. (Hakemli Dergi)

3. Hakemli Kongre / Sempozyum Bildiri Kitaplarında Yer Alan Yayınlar

1. Çanayaz E., Altıkardes Z. A., Ünsal A. (2024). Haralick Feature-Based Deep Learning Model for Ankylosing Spondylitis Classification Using Magnetic Resonance Images. 2024 International Conference on Innovations in Intelligent Systems and Applications (INISTA), Craiova, Romania, 4-6 September 2024, pp.1-6.
2. Çanayaz E., Altıkardes Z. A., Mert K. U., Dural M., Aydın V., Kocakaya O., et al. (2021). Machine Learning Algorithms for Detecting Risk of Atrial Fibrillation Using Baseline Holter ECG and Echocardiography Data. 37th International Turkish Cardiology Congress, Istanbul, Turkey, 18-21 November 2021.
3. Korkmaz H., Çanayaz E., Birtane Akar S., Altıkardes Z. A. (2018). Fuzzy Logic Based Risk Assessment System Giving Individualized Advice For Metabolic Syndrome and Fatal Cardiovascular Diseases. The 7th International Conference on Biomedical Engineering and Biotechnology (ICBEB 2018), Nanjing, China, 17-20 October 2018.
4. Çanayaz E., Böcekçi V. G. (2018). Comparison of Performance of Different Background Subtraction Methods for Detection of Heavy Vehicles. Signal

Processing - Algorithms, Architectures, Arrangements, and Applications (SPA), Poznan, Poland, 19-21 September 2018, pp.50-54.

5. Çanayaz E., Akgün Ö. (2017). EKG'den Wavelet Dönüşümü Yardımı ile Transmural Miyokard İskemisinin Belirlenmesi. EEMKON 2017, Turkey, 16-18 November 2017.
6. Birtane S., Çanayaz E., Altıkardeş Z. A., Korkmaz H. (2017). Development of Decision Support System Using Mamdani Type Fuzzy Logic Clusters for Metabolic Syndrome Risk Assessment. Medical Technologies National Congress (TIPTEKNO), Trabzon, Turkey, 12-14 October 2017.
7. Kesen U., Sarıkaş A., Çanayaz E. (2014). Generating an Artificial Intelligent System to Diagnose Thyroid Gland Related Diseases using Fuzzy Logic and Neural Network. 2nd International Symposium on Innovative Technologies in Engineering and Science (ISITES'2014), 18-20 June 2014.
8. Kesen U., Kakilli A., Sarıkaş A., Çanayaz E. (2014). Web Based GPS Control System Design and Implementation. World Symposium on Web Application and Networking (ICNWTN'2014), Toronto, Canada, 22-24 May 2014, pp.12-15.
9. Çanayaz E., Buran R., Sarıkaş A. (2013). Interpretation of Thyroid Blood Tests Using Mamdani Type Fuzzy Logic Sets. 8th International Conference on Electrical and Electronics Engineering (ELECO), Bursa, Turkey, 28-30 November 2013, pp.614-616.
10. Çanayaz E., Sarıkaş A., Buran R. (2012). Mamdani Tip Bulanık Mantık Kümeleri Kullanarak Tiroid Hastalarının Kan Testlerinin Yorumlanması. TIPTEKNO 2012, Bursa, Turkey, 1-3 November 2012.

4. Kitap ve Kitap Bölümleri

1. Çanayaz E., Altıkardeş Z. A. (2022) Geliştirilen Yapay Zeka Teknolojileri Klinik Uygulamalarda Ne Kadar Yer Buluyor? Enfeksiyon Hastalıkları Özel Konular Türkiye Klinikleri Yayınevi, Ankara, ss.30-34