T.C. SAĞLIK BİLİMLERİ ÜNİVERSİTESİ
ANKARA BİLKENT ŞEHİR
SAĞLIK UYGULAMA VE ARAŞTIRMA MERKEZİ

RADYOLOJİ KLİNİĞİ

DERİN ÖĞRENME MİMARİLERİ İLE KELLGREN-LAWRENCE EVRELEME YAPAN MODELİN RETROSPEKTİF ÖN-ARKA DİZ GRAFİLERİ ÜZERİNDE DEĞERLENDİRİLMESİ VE RADYOLOG KARAR SÜREÇLERİNE ETKİSİNİN İNCELENMESİ

Dr. Mehmet Numan ÇOLAKOĞLU

TIPTA UZMANLIK TEZİ

ANKARA/2025

T.C. SAĞLIK BİLİMLERİ ÜNİVERSİTESİ
ANKARA BİLKENT ŞEHİR
SAĞLIK UYGULAMA VE ARAŞTIRMA MERKEZİ

RADYOLOJİ KLİNİĞİ

DERİN ÖĞRENME MİMARİLERİ İLE KELLGREN-LAWRENCE EVRELEME YAPAN MODELİN RETROSPEKTİF ÖN-ARKA DİZ GRAFİLERİ ÜZERİNDE DEĞERLENDİRİLMESİ VE RADYOLOG KARAR SÜREÇLERİNE ETKİSİNİN İNCELENMESİ

Dr. Mehmet Numan ÇOLAKOĞLU

Tez Danışmanı
Prof. Dr. Semra DURAN

TIPTA UZMANLIK TEZİ

ANKARA/2025

TEŞEKKÜR

Tezimin planlama, uygulama ve tamamlama aşamalarında deneyimiyle yönlendiren, her aşamada desteğini esirgemeyerek önemli katkılar sağlayan değerli tez danışmanım Prof. Dr. Semra Duran'a içtenlikle teşekkür ederim.

Asistanlığım süresince deneyim ve bilgilerini bizlerle fedakarlıkla paylaşarak mesleki gelişimime katkı sağlayan Ankara Bilkent Şehir Hastanesi Radyoloji Kliniği'ndeki değerli hocalarıma, uzman abi ve ablalarıma teşekkür ederim.

Uzmanlık eğitimim süresince birlikte çalışmaktan büyük memnuniyet duyduğum, destekleriyle süreci daha verimli ve keyifli kılan, birlikte birçok güzel anı paylaştığımız değerli arkadaşlarım — başta Faruk Dikici, Mehmet Kutlu ve Rıdvan Arda Demirci olmak üzere — tüm asistan arkadaşlarıma içtenlikle teşekkür ederim.

Tez sürecim ve eğitimim boyunca koşulsuz sevgi, sabır ve destekleriyle her zaman yanımda olan aileme; doktorluğu seçmemde büyük rol oynayan değerli annem Sibel Çolakoğlu'na, asistanlığım süresince radyoloji ile mühendislik alanları arasındaki yapay zeka ve diğer interdisipliner yaklaşımları keşfetmeme katkı sağlayan değerli babam Ali Çolakoğlu'na ve yazılım ile programlama öğrenme sürecimde desteğini esirgemeyen sevgili kardeşim Ahmet Kerem Çolakoğlu'na en derin şükranlarımı sunarım.

Ayrıca bu süreç boyunca anlayışı ve desteğiyle her zaman yanımda olan müstakbel eşim Nefise Garabli'ya özel olarak teşekkür ederim.

Dr. Mehmet Numan Çolakoğlu

Ankara/2025

İÇİNDEKİLER

TEŞEKKÜR i
İÇİNDEKİLER ii
KISALTMALAR v
TABLO LİSTESİ vi
ŞEKİL LİSTESİ vii
ÖZET viii
ABSTRACT x
1. GİRİŞ VE AMAÇ 1
2. GENEL BİLGİLER 3
2.1. OSTEOARTRİT 3
2.1.1. Tanım 3
2.1.2. Epidemiyoloji 3
2.1.2.1. Global Yük ve Prevalans 3
2.1.2.2. Yaş ve Cinsiyet Dağılımı 3
2.1.2.3. Ekonomik Yük 4
2.1.3. OA Sınıflandırması 4
2.1.4. Risk Faktörleri 5
2.1.4.1. Değiştirilemeyen Risk Faktörleri 5
2.1.4.2. Değiştirilebilir Risk Faktörleri 5
2.1.5. Patogenez 6
2.1.5.1. Moleküler Kaskadlar 6
2.1.5.2. Subkondral Kemik ve Osteofit Değişiklikleri 7
2.1.6. OA Kliniği 7
2.1.7. OA Tedavisi 8
2.2. OA'NIN RADYOLOJİK TANISI 10
2.2.1. Konvansiyonel Radyografi 10
2.2.2. Manyetik Rezonans Görüntüleme (MRG) 10
2.2.3. Bilgisayarlı Tomografi (BT) 10
2.2.4. Ultrasonografi (US) 11
2.3. KELLGREN–LAWRENCE EVRELEMESİ 12
2.3.1. Tarihi ve Güncel Durum 12

2.3.2. Evreleme Kriterleri 12
2.4. TIBBİ GÖRÜNTÜ ANALİZİNDE YAPAY ZEKA 12
2.4.1. Yapay Zeka, Makine Öğrenmesi ve Derin Öğrenme Temelleri 12
2.4.2. Yapay Sinir Ağları Temelleri 13
2.4.3. Yapay Sinir Ağlarının Öğrenme Türleri 14
2.4.4. Geri Yayılım Algoritması 14
2.4.5. Evrişim İşlemi 15
2.4.6. Evrişimli Sinir Ağları (ESA) 15
2.4.7. Evrişimli Sinir Ağları (ESA) Katmanları 15
2.4.8. Transfer Öğrenimi (Transfer Learning) 17
2.5. DENSENET MİMARİSİ 18
2.6. VISION TRANSFORMER 18
2.6.1. Vision Transformer'ın Gelişimi 18
2.6.2. ViT'nin Transfer Öğrenimiyle Kullanımı 19
2.7. GRAD-CAM YÖNTEMİ 19
2.7.1. Grad-CAM Yöntemi Kullanımı 19
2.7.2. Klinik Açıklanabilirlik 19
2.8. PERFORMANS DEĞERLENDİRME METRİKLERİ 19
2.9. GÖZLEMCİLER ARASI VE GÖZLEMCİ İÇİ UYUM METRİKLERİ 20
2.9.1. İkinci Dereceden Ağırlıklı Kappa(İDAK) 20
2.9.2. Krippendorff's Alpha 21
2.10. RECEIVER OPERATING CHARACTERISTIC CURVE (ROC) – ALICI-İŞLETİM KARAKTERİSTİĞİ EĞRİSİ 21
2.10.1. Eğri Altındaki Alan- Area Under the Curve (AUC) 21
2.11. İSTATİSTİKSEL ANLAMLILIK TESTLERİ 22
2.11.1. Wilcoxon İşaretli Sıra Testi 23
2.11.2. İşaret Testi 23
3. GEREÇ VE YÖNTEM 24
3.1. ÇALIŞMANIN PLANI 24
3.2. VERİ SETLERİ 24
3.2.1. Eğitim ve İç Doğrulama Veri Seti 24
3.2.2. Harici Doğrulama Veri Seti 26
3.3. GÖRÜNTÜ ÖN İŞLEME 28
3.3.1. Manuel Kırpma İşlemi 28

3.3.2. Boyutlandırma ve Normalizasyon 28
3.4. VERİ ARTIRMA 29
3.5. MIXUP VERİ KARIŞTIRMA STRATEJİSİ 29
3.6. MODEL MİMARİLERİ 30
3.6.1. Evrişimsel Sinir Ağı Tabanlı Modeller (n=10) 30
3.6.2. Vision Transformer Modelleri (n=4) 31
3.7. TRANSFER ÖĞRENME STRATEJİSİ 32
3.8. GRAD-CAM GÖRSELLEŞTİRME 33
3.9. RADYOLOG PERFORMANS DEĞERLENDİRMESİ 33
3.10. PERFORMANS METRİKLERİ 34
3.10.1. Radyologlar Arası Uyum Analizi 34
3.10.2. Gözlemci İçi Tutarlılık Analizi 34
3.11. İSTATİSTİKSEL ANALİZ 35
4. BULGULAR 36
4.1. MODEL PERFORMANSLARI - MENDELEY TEST SETİ 36
4.2. MODEL PERFORMANSI - HARİCİ HASTANE SETİ 37
4.2.1. Yapay Zeka Karışıklık Matrisi 38
4.3. DETAYLI MODEL PERFORMANSI DEĞERLENDİRME 42
4.4. GRAD-CAM GÖRSELLEŞTİRME SONUÇLARI 43
4.5. RADYOLOG PERFORMANS DEĞERLENDİRMESİ VE GÖZLEMCİLER ARASI UYUM ANALİZİ 46
4.5.1. Yapay Zeka Desteği Öncesi ve Sonrası Performans 46
4.5.2. Gözlemci İçi Tutarlılık Analizi 49
4.5.3. Gözlemciler Arası Tutarlılık Analizi 49
4.5.4. Krippendorff’s Alpha Sonuçları 50
4.5.6. Özet Bulgular 52
5. TARTIŞMA 53
6. SONUÇLAR 57
7. KAYNAKLAR 58
8. ÖZGEÇMİŞ 68
9. EKLER 70
EK 1: ETİK KURUL ONAYI 70

KISALTMALAR

3B	: 3 Boyutlu
ADAMTS	: A Disintegrin and Metalloproteinase with Thrombospondin Motifs
AUC	: Area Under the Curve (Eğri Altında Kalan Alan)
BT	: Bilgisayarlı Tomografi
DICOM	: Digital Imaging and Communications in Medicine
DPO	: Doğru Pozitif Oranı
ESA	: Evrişimsel Sinir Ağı
EYY	: Engellilikle Yaşanan Yıllar
Grad- CAM	: Gradient-weighted Class Activation Mapping
IL-1β	: İnterlökin-1 Beta
İDAK	: İkinci Dereceden Ağırlıklı Kappa
KL	: Kellgren-Lawrence
KOA-SG	: Knee Osteoarthritis Severity Grading
MMP	: Matriks Metalloproteinaz
MRG	: Manyetik Rezonans Görüntüleme
NSAİİ	: Non-Steroid Anti-İnflamatuar İlaç
OA	: Osteoartrit
PNG	: Portable Network Graphics
ReLU	: Rectified Linear Unit
ROC	: Receiver Operating Characteristic Curve- Alıcı-İşletim Karakteristiği Eğrisi
TNF-α	: Tümör Nekroz Faktörü Alfa
US	: Ultrasonografi
ViT	: Vision Transformer
VKİ	: Vücut Kitle İndeksi
YPO	: Yanlış Pozitif Oranı
YSA	: Yapay Sinir Ağı
YZ	: Yapay Zeka
YZÖ	: Yapay Zeka Öncesi
YZS	: Yapay Zeka Sonrası

TABLO LİSTESİ

Tablo 3.1. Mendeley Veri Seti Sınıf Dağılımı 26
Tablo 3.2. Harici Doğrulama Setinin Demografik Özellikleri 27
Tablo 4.1. Mendeley Test Setinde Model Performansları 36
Tablo 4.2. Mendeley Test Setinde Model Ailesi Performansları 36
Tablo 4.3. DenseNet-169 Harici Test Seti Performansları 37
Tablo 4.4. DenseNet-169 Evre Bazlı Performans (Harici Set) 38
Tablo 4.5. ConvNeXt ve DenseNet Modelleri 42
Tablo 4.6. EfficientNet, MobileNetV2 ve ResNet Modelleri 43
Tablo 4.7. Swin ve ViT (Vision Transformer) Modelleri 43
Tablo 4.8. Radyolog Performansı Karşılaştırması 47
Tablo 4.9. İşaret Testi İstatistikleri 47
Tablo 4.10. Raporlama Süresi ve Güven Skoru Değişimleri 48
Tablo 4.11. Wilcoxon İşaretli Sıra Testi Sonuçları 49
Tablo 4.12. Gözlemci İçi Tutarlılık 49
Tablo 4.13. Eşli İDAK Wilcoxon Testi 50
Tablo 4.14. Radyologlar Arası Uyumun Krippendorff's Alpha Sonuçları 50

ŞEKİL LİSTESİ

Şekil 2.1. Kellgren-Lawrence Evreleme Sistemi 12
Şekil 2.2. Yapay Sinir Ağı Örneği 14
Şekil 2.3. DenseNet Mimarisi 18
Şekil 3.1. Veri Arttırma ve Karıştırma Tekniği Örnekleri 30
Şekil 3.2. ImageNet Veri Seti Örneği 32
Şekil 4.1. DenseNet-169 Karışıklık Matrisi 39
Şekil 4.2. DenseNet-169 ROC Eğrileri Eğitim ve Doğrulama Seti 40
Şekil 4.3. DenseNet-169 ROC Eğrileri Mendeley ve Harici Hastane Test Verisi 41
Şekil 4.4. KL Evre 0-1 Grad-CAM Isı Haritası Örnekleri 45
Şekil 4.5. KL Evre 2-4 Grad-CAM Isı Haritası Örnekleri 45
Şekil 4.6. Özellik Haritası Örnekleri 46
Şekil 4.7. Yapay Zeka Öncesi ve Sonrası Gözlemciler Arası Kappa Değerleri 50
Şekil 4.8. Değerlendiriciler ve YZ için Gerçek Değerle İDAK Skorları 51

ÖZET

Amaç: Bu çalışmanın amacı, diz osteoartritinin Kellgren-Lawrence evrelemesini otomatik olarak gerçekleştiren derin öğrenme tabanlı bir sistem geliştirmek, farklı mimari yaklaşımları karşılaştırmak ve sistemin radyolog karar süreçlerine etkisini değerlendirmektir.

Gereç ve Yöntem: PyTorch kütüphanesi kullanılarak on dört farklı derin öğrenme mimarisi geliştirilmiş ve karşılaştırılmıştır. Evrişimli sinir ağı kategorisinde ResNet-18, ResNet-50, ResNet-101, DenseNet-121, DenseNet-169, DenseNet-201, EfficientNet-B0, MobileNetV2, ConvNeXt-Tiny ve ConvNeXt-Small; Vision Transformer kategorisinde Swin-T, Swin-S, ViT-B/16 ve ViT-L/16 mimarileri kullanılmıştır. Modeller, Mendeley Data platformundan elde edilen 8.260 görüntüden oluşan KOA-SG v1.0 veri seti ile eğitilmiş ve test edilmiştir. Harici doğrulama için Ankara Bilkent Şehir Hastanesi'nden toplanan 100 görüntülük bağımsız bir veri seti kullanılmıştır. Altı radyolog, yapay zeka desteği öncesi ve sonrası olmak üzere iki aşamada değerlendirilmiştir.

Bulgular: DenseNet-169 mimarisi en yüksek performansı göstermiş olup, Mendeley test setinde %70,17 genel doğruluk, 0,9079 Macro AUC ve 0,9340 Micro AUC değerleri elde etmiştir. Harici hastane setinde %65 genel doğruluk, 0,8575 Macro AUC ve 0,8863 Micro AUC değerleri kaydedilmiştir. Evrişimli sinir ağı tabanlı modeller açık erişim veri seti ve harici hastane veri setinde Vision Transformer modellerine göre anlamlı olarak üstün performans sergilemiştir.

Yapay zeka desteği, radyolog doğruluğunda önemli iyileşmeler sağlamıştır: R1 %51'den %61'e, R2 %52'den %63'e, R3 %65'ten %71'e, R4 %64'ten %72'ye, R5 %77'den %81'e ve R6 %78'den %83'e yükselmiştir. Grad-CAM analizi, modelin %78 oranında klinik olarak anlamlı anatomik bölgelere odaklandığını göstermiştir.

Gözlemciler arası uyum yapay zeka desteği ile ortalama eşli ikinci dereceden ağırlıklı kappa 0,708'den 0,836'ya yükselmiş, %18,1'lik artış bulunmuştur ($p<0,001$). Krippendorff's Alpha değerleri 0,726'dan 0,845'e yükselmiştir (%16,4 artış).

Gözlemci içi tutarlılık analizi tüm radyologlar için iyi uyum seviyesinde bulunmuştur (ortalama ikinci dereceden ağırlıklı kappa: 0,791). Yapay zekanın gerçek değerle uyumu mükemmel seviyede bulunmuştur (ikinci dereceden ağırlıklı kappa: 0,864).

Sonuç: DenseNet-169 mimarisi, diz osteoartrit tanısında klinik karar destek sistemi olarak kullanılabilecek yüksek potansiyel göstermiştir. Evrişimli sinir ağı tabanlı modeller genel olarak Vision Transformer modellerine kıyasla daha yüksek performans göstermiştir. Yapay zeka, bireysel performansı arttırmanın yanı sıra tanı standardizasyonunu sağlamaya da yardımcıdır.

Anahtar Kelimeler: Derin Öğrenme, Evrişimli Sinir Ağları, Osteoartrit, Radyografi, Tanısal Görüntüleme, Yapay Zeka

ABSTRACT

Objective: The aim of this study is to develop a deep learning-based system that automatically performs Kellgren-Lawrence staging of knee osteoarthritis, to compare different architectural approaches, and to evaluate the impact of this system on radiologist decision-making processes.

Materials and Methods: Fourteen different deep learning architectures were developed and compared using the PyTorch library. Among these architectures, in the convolutional neural network category, ResNet-18, ResNet-50, ResNet-101, DenseNet-121, DenseNet-169, DenseNet-201, EfficientNet-B0, MobileNetV2, ConvNeXt-Tiny, and ConvNeXt-Small were included; in the Vision Transformer category, Swin-T, Swin-S, ViT-B/16, and ViT-L/16 were included. The models were trained and tested with the KOA-SG v1.0 dataset consisting of 8,260 images obtained from the Mendeley Data platform. For external validation, an independent dataset of 100 images collected from Ankara Bilkent City Hospital was used. Six radiologists were evaluated in two stages: before and after artificial intelligence support.

Results: The DenseNet-169 architecture showed the highest performance, achieving 70.17% overall accuracy, 0.9079 Macro AUC, and 0.9340 Micro AUC values on the Mendeley test set. On the external hospital set, 65% overall accuracy, 0.8575 Macro AUC and 0.8863 Micro AUC values were recorded. Convolutional Neural Network-based models demonstrated significantly superior performance compared to Vision Transformer models on both the open-access dataset and the external hospital dataset.

Artificial intelligence support provided significant improvements in radiologist accuracy: R1 increased from 51% to 61%, R2 from 52% to 63%, R3 from 65% to 71%, R4 from 64% to 72%, R5 from 77% to 81%, and R6 from 78% to 83%. Grad-CAM analysis showed that the model focused on clinically significant anatomical regions at a rate of 78%.

Interobserver agreement increased from an average pairwise quadratic weighted kappa of 0.708 to 0.836 with artificial intelligence support, showing an 18.1% increase ($p<0.001$). Krippendorff's Alpha values increased from 0.726 to 0.845 (16.4% increase).

Intraobserver agreement analysis was found to be at a good agreement level for all radiologists (average quadratic weighted kappa: 0.791). The agreement of artificial intelligence with true values was found to be at an excellent level (quadratic weighted kappa: 0.864).

Conclusion: The DenseNet-169 architecture has shown high potential to be used as a clinical decision support system in knee osteoarthritis diagnosis. Convolutional Neural Network-based models generally demonstrated higher performance compared to Vision Transformer models. Artificial intelligence helps provide diagnostic standardization in addition to improving individual performance.

Keywords: Artificial Intelligence, Convolutional Neural Networks, Diagnostic Imaging, Deep Learning, Osteoarthritis, Radiography

1. GİRİŞ VE AMAÇ

Osteoartrit (OA), kapsül-bursa iltihabını, sinovyal sıvıdaki değişimleri, kıkırdak yıkımını ve erozyonlarını, osteokondral inflamatuvar hasara bağlı kemik erozyonu ve şekil bozukluklarını içeren bir kısır döngü içinde tüm eklem yapısını etkileyen ilerleyici bir dejeneratif hastalıktır (1). Global Burden of Disease Study 2021 verilerine göre, 2020 yılında dünya çapında tahminen 595 milyon kişi OA ile yaşamaktadır ve bu, küresel nüfusun yaklaşık %7,6'sına tekabül etmektedir. Bu rakam 1990 yılına göre önemli bir artış göstermektedir; 1990'da küresel nüfusun %4,8'i OA ile yaşarken, bu 256 milyon kişiye denk gelmekteydi (2).

OA ile yaşayan bireylerin yaklaşık %60'ı kadın, %73'ü 55 yaş üzerindedir (3). Diz, en sık etkilenen eklem olup, 2050 yılında yaklaşık 642 milyon kişinin diz OA'dan etkileneceği öngörülmektedir (2). OA'ya bağlı engelliliğe en çok neden olan semptom ağrı olup, klinik karar alma süreçlerini büyük oranda etkilemektedir (4).

Hastalığın küresel yükü önemlidir. 2020 yılında total osteoartrit için global yaş-standardize edilmiş Engellilikle Yaşanan Yıllar oranı 100.000 kişide 255,0 olup, bu 1990'dan %9,5 artış göstermektedir (2). ABD'de OA'ya bağlı yıllık harcama maliyetinin 185,5 milyar dolara ulaştığı tahmin edilmektedir (5).

Kellgren-Lawrence(KL) evrelemesi, 1957 yılında Kellgren ve Lawrence tarafından geliştirilmiş olup, diz OA'nın radyolojik evrelendirilmesinde yaygın olarak kullanılmaktadır. Bu sistemde KL evre 0 normal eklemi KL evre 4 ise şiddetli OA'yı tanımlar (6,7). KL sisteminde yorumcu bağımlılığı olsa da en yüksek uyum diz eklemindedir (r = 0,83) (7).

Günümüzde tanı amacı ile tıbbi görüntülemenin giderek artan kullanımı klinik tanı ve tedavi planlamasında uzmanlara yardımcı olmak için yapay zeka kullanımını neredeyse zorunlu hale getirmiştir. Bilgisayar destekli tanı kliniklerde yaygın olarak kullanılmaktadır. Derin öğrenme yöntemleriyle klinik başarılar giderek artmaktadır.

Bu çalışmanın temel amacı, diz anterior-posterior grafilerinde KL evrelemesi yapabilecek, yüksek doğruluklu ve güvenilir derin öğrenme modelleri geliştirmek ve farklı mimari yaklaşımları sistematik olarak karşılaştırmaktır.

Klinik hedefler ise radyolog karar süreçlerine destek olabilecek güvenilir bir araç sunulması, yorumcu bağımlılığının azaltılması ve standardizasyonun sağlanması, gözlemciler arası ve gözlemci içi uyumun iyileştirilmesi, özellikle deneyimsiz radyologlar için tanı desteği sağlanmasıdır.

2. GENEL BİLGİLER

2.1. OSTEOARTRİT

2.1.1. Tanım

Osteoartrit (OA), ağrı ve işlev kaybı ile seyreden yaygın görülen dejeneratif bir eklem hastalığıdır (8). OA en yaygın eklem hastalığı olup 60 yaş üstü erkeklerin %10'unu ve kadınların %18'ini etkilediği düşünülmektedir (9). Diz eklemi OA'dan en sık etkilenen eklemdir (4). OA, sağlık sisteminde ciddi kaynak tüketimine yol açar ve hastaların yaşam kalitesini belirgin şekilde olumsuz etkiler (8).

OA sadece mekanik kıkırdak hasarı ile sınırlı kalmayıp, kıkırdak, subkondral kemikler ve sinovyumun da rol aldığı ve sistemik inflamasyonun eşlik edebildiği kompleks bir durumdur (9).

2.1.2. Epidemiyoloji

2.1.2.1. Global Yük ve Prevalans

Global Burden of Disease Study 2021 verilerine göre, dünya çapında 595 milyon kişi OA'dan etkilenmektedir. Bu rakam toplam dünya popülasyonunun %7,6'sına tekabül etmekte olup, 1990 yılına göre %132,2 artmıştır (2). OA ile yaşayan kişilerin yaklaşık %60'ı kadın, %73'ü 55 yaş üzerindedir (3).

Güncel global prevalans verilerine göre, 365 milyon prevalansla diz en sık etkilenen eklem olup, bunu kalça ve el takip etmektedir (3). Araştırmacıların tahminlerine göre, 2050'ye kadar her OA lokalizasyonu için global prevalans %48,6-95,1 oranında artış gösterecektir (2).

2.1.2.2. Yaş ve Cinsiyet Dağılımı

OA için en büyük risk faktörü yaştır (10). Yetmişbeş yaş üzerindeki kişilerin %80'inde radyolojik olarak OA mevcuttur (11). OA ile yaşayan kişilerin yaklaşık %60'ı kadındır (3).

OA insidansı ve prevalansı açısından kadınlar genel olarak erkeklere göre daha yüksek risk altında olup, bu fark 55 yaş üstü kişilerde daha belirgindir. Erkeklerde diz ve el OA riski kadınlara oranla belirgin düşüktür (12).

2.1.2.3. Ekonomik Yük

2020 yılında total OA için global yaş-standardize edilmiş Engellilikle Yaşanan Yıllar oranı 100.000'de 255,0 olup, bu 1990'dan %9,5 artış göstermektedir.

Yetmiş yaş ve üzeri yetişkinlerde OA, EYY'nin yedinci sıradaki nedenidir (2). ABD'de OA'ya bağlı yıllık harcama maliyetinin 185,5 milyar dolara ulaşmakta olduğu düşünülmektedir (5).

2.1.3. OA Sınıflandırması

OA sınıflandırması etiyolojik faktörlere göre primer OA ve sekonder OA olarak ikiye ayrılmaktadır.

Primer (İdiopatik) OA: Altta yatan nedeni bulunmayan, yaşlanma süreciyle birlikte gelişen OA türüdür (13).

Sekonder OA: Belirli bir nedenle gelişen OA türü olup şu kategorilerde incelenmektedir:

- Metabolik nedenlere bağlı: Akromegali, okranozis, raşitizm, hemakromatozis (13), kristal depo hastalığı (14)
- Anatomik nedenlere bağlı: Femur başı epifiz kayması, gelişimsel kalça displazisi, Perthes hastalığı (15)
- Travmatik nedenlere bağlı: Cerrahi ve travma sonrası (13)
- İnflamatuar nedenlere bağlı: Romatoid artrit, enfeksiyöz artrit, psöriatik artrit (13)

2.1.4. Risk Faktörleri

OA risk faktörleri değiştirilemeyen ve değiştirilebilen risk faktörleri olarak ikiye ayrılır (13).

2.1.4.1. Değiştirilemeyen Risk Faktörleri

Yaş: Yaşın OA ile ilişkisi iyi biliniyor olsa da yaşlanmanın eklemde OA gelişimine yol açma mekanizması tam olarak anlaşılamamıştır (16). Semptomatik OA prevalansı yaş ile artmaktadır (17). Kondrositler yaş ve OA ile hücresel yaşlanmaya gitmektedir. Yaşlanan kondrositler ekstrasellüler matriks gen ekspresyonunu ve sentezini azaltır. Yaşlanmayla birlikte kas kütlesinde azalma, yağ kütlesi artışı ekleme binen yükte bir değişime yol açar. Bu durum adipokin ve sitokin salınımında artışa yol açar, bu durum düşük dereceli bir sistemik inflamasyonla sonuçlanır. Kondrositlerdeki mitokondrial disfonksiyon, artmış oksidatif stres ve otofajide azalma fonksiyon değişimine yol açar, katabolizma ve hücre ölümünü kolaylaştırır (18).

Cinsiyet: OA ile yaşayan kişilerin yaklaşık %60'ı kadındır (3). OA sıklığı kadınlarda erkeklere göre daha yüksek olup, erkeklerde diz ve el OA riski kadınlara oranla belirgin düşüktür (12). Orta yaş kadınlarda serum estradiol ve idrar 2-hydroxyestrone düzeylerinin düşüklüğü diz OA'sı gelişimiyle ilişkilidir. Bu durumun 2-hydroxyestrone'un araşidonik asit yolağı, lökotrien sentezi, lipooksijenaz modülasyonu üzerine olan etkilerinden kaynaklanabileceği düşünülmektedir (19).

Genetik Faktörler: OA'nın kalıtsal komponenti güçlü şekilde desteklenmektedir (20). Genome-wide association studies çalışmaları 100 adet OA ile ilişkili genetik varyantı tanımlamıştır (21). COL2A1, COL9A1, COL9A2, COL9A3, COL11A1, COL11A2, ACAN, COMP, MATN3, SMAD3, TRAPPC2, TNFRSF11B genlerindeki mutasyonlar erken başlangıçlı OA ile ilişkilidir (22).

2.1.4.2. Değiştirilebilir Risk Faktörleri

Obezite: Vücut kitle indeksi (VKİ) yüksekliği, osteoartrit (OA) gelişimi açısından dikkate alınması gereken değiştirilebilir bir risk faktörüdür. Her 5 kg/m² VKİ artışı için diz OA riski %35 artmaktadır (23).

Beslenme Faktörleri:

- D vitamini: Vitamin D takviyesi diz OA hastalarında fonksiyonelliğin iyileşmesinde etkilidir (24)
- K vitamini: Düşük seviyeleri el ve diz OA bulgularıyla ilişkilidir (25)
- C vitamini: C vitamini gibi antioksidanların yüksek miktarda alımı ile, kartilaj kaybı ve hastalığın ilerleyişinin önlenebileceği düşünülmektedir (26)
- Omega-3 yağ asitleri: Anti-inflamatuar etkileri nedeniyle OA'daki proinflammatuar sitokinler ve kartilaj kaybı açısından faydalı olabileceği düşünülmektedir (27)

Meslek ve Fiziksel Aktivite: Diz çökme ve çömelme gerektiren işlerde çalışanlarda OA gelişme riski artmaktadır (28). Diz OA tanılı hastalarda kara-su egzersiz programları ağrı, eklem hareket açıklığı ve hayat kalitesinde iyileşme sağlamaktadır (29).

2.1.5. Patogenez

Tüm ilerlemelere rağmen OA etiyolojisi ve patolojik mekanizmaları halen tam olarak anlaşılamamıştır (30).

2.1.5.1. Moleküler Kaskadlar

- IL-1β ve TNF-α: Katabolik-inflammatuar süreçleri başlatan ana sitokinler (31)
- Matriks metalloproteinazlar(MMP): MMP'ler kalsiyum ve çinko bağımlı yıkıcı proteinazlar olup sinoviyal sıvıdaki MMP-1, MMP-3, MMP-9 düzeylerinin yüksekliğinin OA şiddetiyle ilişkili olduğu gösterilmiştir (32)
- ADAMTS Ailesi: ADAMTS-4 ve ADAMTS-5 aggrekan yıkımından sorumlu olup, ADAMTS'ler ekstraselüler matriks kaybı, kartilaj dejenerasyonu ve sinovial inflamasyonda rol oynar (33)

Anabolik-Katabolik Denge Bozukluğu: Normal şartlarda dengede olan anabolik ve katabolik süreçler arasındaki dengenin katabolik yönde bozulması hastalığın temelini oluşturur (31).

2.1.5.2. Subkondral Kemik ve Osteofit Değişiklikleri

Erken evre OA'da eklem yüzeylerinde vaskülarizasyon artışı ve kemik dansitesinde azalma izlenmekte olup, geç dönem OA'da kemik rezorpsiyonunda azalma ve subkondral skleroz dikkati çekmektedir (34). Diz eklemlerinde, ön çapraz bağın yırtılmasından sonra öne ve arkaya doğru osteofitlerin geliştiği ve bunların, femurun tibia üzerindeki yer değiştirmesini sınırlayarak eklemi stabilize ettiği gösterilmiştir.

Primer diz OA'da, kenar osteofitlerinin çıkarılması eklem hareketliliğini önemli ölçüde artırmıştır; bu da osteofitlerin OA'lı dizlerin hareket kabiliyetini sınırladığını göstermektedir (35).

2.1.6. OA Kliniği

OA genellikle yavaş ve sinsi bir şekilde başlar; çoğu zaman bir veya birkaç eklemde ortaya çıkar. En erken belirti ağrıdır ve genellikle derin, künt bir ağrı şeklinde tanımlanır.

Ağrı, yük binmesiyle artar, dinlenmeyle azalır; hastalık ilerledikçe sürekli hale gelebilir. Tutukluluk hissi genellikle sabah uyanınca veya uzun süre hareketsiz kaldıktan sonra hissedilir, ancak genellikle 30 dakikadan kısa sürer ve hareketle birlikte azalır. Hastalık ilerledikçe eklem hareketleri kısıtlanır, hassasiyet, krepitasyon gelişir.

Büyük eklemlerde, özellikle dizlerde, eklem efüzyonu sık görülür. Zamanla fleksiyon kontraktürleri gelişebilir. Palpasyonla hassasiyet ve pasif hareket sırasında ağrı genellikle geç dönem bulgularıdır. Kas spazmı ve kontraktür ağrıyı artırabilir.

Eklem içinde serbest cisimler veya yırtılmış menisküs parçaları kilitlenme veya takılma hissi oluşturabilir. İleri evrelerde deformiteler ve subluksasyonlar gelişebilir.

En sık etkilenen eklemler şunlardır:

- Distal interfalangeal ve proksimal interfalangeal eklemler
- Başparmak karpometakarpal eklemi
- Servikal ve lomber intervertebral diskler ve zigapofiziyal eklemler
- Birinci metatarsofalangeal eklem
- Kalça eklemi
- Diz eklemi (36)

2.1.7. OA Tedavisi

Tedavi yaklaşımları farmakolojik ve non-farmakolojik yöntemleri içermektedir (37).

Non-Farmakolojik Tedavi Yaklaşımları

OA non-farmakolojik tedavisi, eklem çeşidi, komorbiditeler gibi hasta faktörlerine göre bireyselleştirilmelidir (38). Hastalık yönetimi, başvuran bireye özel olarak uyarlanmalı; öz-yönetim, eğitim, egzersiz ve gerekli durumlarda kilo kaybı gibi temel tedavilere odaklanmalıdır (4). Kilo kaybı obez hastalarda önemli bir tedavi seçeneğidir. Her 5 kg/m² VKİ artışı diz OA riskini %35 oranında arttırır (23). Cerrahi müdahale, daha az invaziv yöntemlere uygun yanıt vermeyen bireyler için saklanmalıdır. Önleme ve hastalık modifikasyonu, çeşitli araştırma girişimleriyle hedeflenen alanlar olup şu ana kadar büyük potansiyel göstermiştir (4). Kas güçlendirme, aerobik, esneklik ve zihin-beden egzersizlerini içeren egzersiz tedavileri, diz OA'da ağrıların hafifletilmesi ve işlevsel iyileşme açısından etkili olduğu düşünülmektedir (39). Diz OA tanılı hastalarda kara ve su egzersiz programları ağrı, diz hareket aralığı ve hayat kalitesinde iyileşme sağlamaktadır (29).

Diz OA olan hastalarda planlanmış davranış teorisine dayalı olarak uygulanan eğitim müdahalesi, öz bakım davranışlarını artırmıştır (40).

Farmakolojik Tedavi Yaklaşımları

Non-steroid anti-inflamatuar ilaçlar(NSAİİ) OA'da inflamatuvar ağrıyı hafifletmek ve fiziksel fonksiyonu iyileştirmek için ilk basamak tedavidir (41).

Analjezik tedaviye ihtiyaç duyan ancak topikal NSAİİ gibi birinci basamak tedavilere yanıt vermeyen ve oral NSAİİ kullanımı için kontrendikasyonu olan OA'lı hastalara, opioid veya parasetamol tedavisi verilebilir (42).

Glukokortikoidler ve hyaluronik asit, en yaygın kullanılan eklem içi tedaviler olup, farmakolojik olmayan tedavilere, NSAİİ'lere veya analjeziklere yanıt vermeyen OA kaynaklı diz ağrısının yönetiminde standart eklem içi tedavi seçeneği olarak kabul edilmektedirler (43).

Kondroprotektif ilaçlar (hastalığı modifiye edici osteoartrit ilaçları), genellikle hastalığın yapısal değişikliklerini hedef alarak, kıkırdak yıkımına neden olan faktörleri engellemeyi veya kıkırdak üretimini teşvik etmeyi amaçlamışlardır (44).

Cerrahi Tedavi Yaklaşımları

Artroskopik diz debridmanı, erken evre OA olan ve konservatif tedaviye yanıt vermemiş belirgin semptomlu hastalarda iyi semptomatik rahatlama ve kalıcı faydalar sağlayabilir (45). Görece daha genç ve dizin korunmasına ihtiyaç duyan hastalarda, eklem koruma yaklaşımını temel alan diz çevresi osteotomisi, alternatif bir cerrahi tedavi seçeneği olarak uygulanmaktadır. Kemiklerin kesilmesi ve yeniden hizalanması, alt ekstremitenin mekanik yük taşıma hattını düzeltir. Bu sayede diz çevresi osteotomisi, normal anatomik yapıyı korur ve diz ekleminin işlevsel olarak iyi bir şekilde iyileşmesini sağlar. Diz çevresi osteotomisi teknikleri arasında, diz eklem kompartmanlarının yük taşıma dengesini yeniden düzenlemeyi amaçlayan anti-varus deformite ve anti-valgus deformite osteotomileri yer alır (46). Diz OA'nın son evresinde, diz protezi kanıtlanmış ve maliyet etkin bir tedavi yöntemidir (47).

2.2. OA'NIN RADYOLOJİK TANISI

2.2.1. Konvansiyonel Radyografi

OA'nın tanısında konvansiyonel radyografi en kolay erişilir tetkik kabul edilmektedir (48). Radyolojik değerlendirmede beş temel bulgu aranmaktadır:

- Eklemi oluşturan kemik yapılarda osteofit oluşumu
- Periartiküler ossiküller
- Subkondral kemikte skleroz ile ilişkili eklem kıkırdağında incelme
- Genellikle subkondral kemikte yerleşmiş, sklerotik duvarlı küçük psödokistik alanlar
- Kemik uçlarında şekil değişikliği (7)

2.2.2. Manyetik Rezonans Görüntüleme (MRG)

MRG, OA ve kıkırdak değerlendirmesinde en önemli görüntüleme yöntemidir (49). Klinik uygulamada, MRG kıkırdak ve menisküs görüntülemesinde mükemmel yumuşak doku kontrastı ve tüm eklemi görselleştirme yeteneği sayesinde, OA gibi tüm eklemi etkileyen hastalıklarda referans yöntem olarak kabul edilmektedir (50).

Ayrıntılı kıkırdak yapı değerlendirmesinde T2 mapping, dGEMRIC, Sodyum 23 MRG ve T1ρ mapping gibi teknikler kullanılmaktadır (49).

2.2.3. Bilgisayarlı Tomografi (BT)

OA'nın erken evrelerinde dizin üç kompartmanında da küçük osteofitlerin değerlendirilmesinde MRG'den daha faydalı olabilir (51). BT görüntülemesindeki son gelişmeler—BT artrografi, spektral görüntüleme ve yük taşıyan tarayıcılar gibi—eklem değerlendirmesinde bu tekniğin kullanımına olan ilgiyi yeniden canlandırmıştır.

Klinik uygulamalarda BT, özellikle kortikal ve trabeküler kemik gibi mineralize dokuların yanı sıra kristal birikimlerinin değerlendirilmesindeki yeteneği sayesinde tercih edilmektedir.

BT'nin kontrast ve uzamsal çözünürlüğü son derece yüksektir; bu özellik, genellikle 3 mm kesit kalınlığına sahip geleneksel iki boyutlu MRG sekanslarına kıyasla düzlem içi çözünürlükte 10 kata kadar üstünlük sağlayabilir. BT, OA bağlamında genellikle sorun çözme aracı olarak kullanılır ve kemik yapıların görselleştirilmesindeki başarısıyla öne çıkar. Artrografi ile birleştirildiğinde, BT kıkırdak ve menisküs yüzey lezyonlarını etkili bir şekilde ayırt edebilir (50).

3B BT görüntüleme MRG ile karşılaştırıldığında, osteofitlerin ve subkondral kistlerin tespitinde sabit fleksiyon radyografisine göre daha hassas ve daha doğru bulunmuştur (52).

2.2.4. Ultrasonografi (US)

Ultrasonografi (US), invaziv olmayan, hızlı, pratik ve güvenli bir görüntüleme yöntemi olup; kıkırdak, menisküs ve tendon gibi yapılarının incelenmesinin yanı sıra eklem efüzyonu, bursit ve sinovit değerlendirmesinde de kullanılmaktadır. Ayrıca US, OA ile ilişkili osteofitler ile subkondral kemik yapıların görüntülenmesinde de etkili bir yöntemdir (53). US, yalnızca kemik yapıların değil, aynı zamanda yumuşak dokuları da ayrıntılı olarak değerlendirebilmesine imkân tanıyarak OA tanısında hem klinik muayeneyi hem de radyolojik görüntülemeyi tamamlayıcı bir rol üstlenir. US, eklem kıkırdağı, kortikal kemik ve sinovyal dokudaki en küçük anormallikleri dahi ortaya koyabilir. Ayrıca US, ilgili eklemlerde sinovyum proliferasyonu ve eklem içi sıvı miktarındaki değişiklikler gibi yumuşak doku patolojilerini saptamada son derece duyarlıdır. Kalça ve diz eklemlerindeki sıvı birikiminin US ile takibi, OA'lı hastaların klinik değerlendirilmesi ve tedaviye verdikleri yanıtın izlenmesi açısından yararlı bir belirteç olabilir. Hem renkli Doppler hem de power Doppler US teknikleri, artmış sinovyal damarlanmanın bir göstergesi olan sinovyal kan akımını ortaya koyabilmektedir. Doppler sinyalindeki artışın, OA'lı hastalarda histolojik incelemelerde gözlenen artmış sinovyal vaskülarite ile uyumlu olduğu gösterilmiştir (54).

2.3. KELLGREN–LAWRENCE EVRELEMESİ

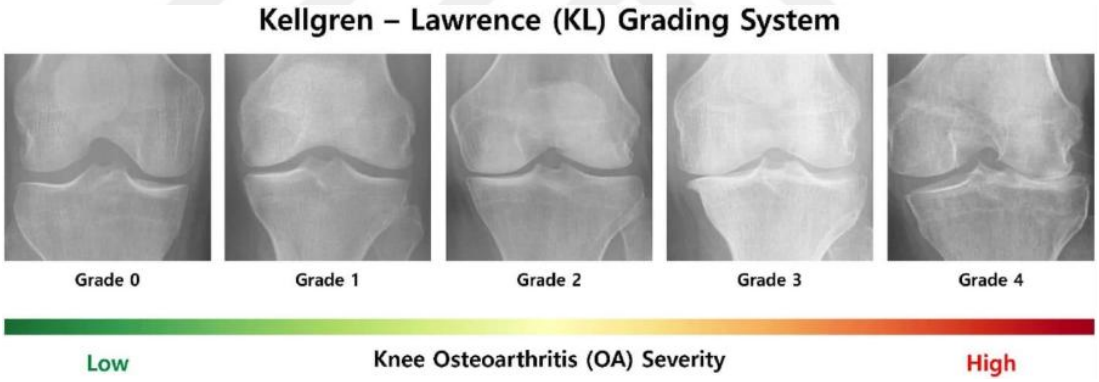
2.3.1. Tarihi ve Güncel Durum

KL evrelemesi, 1957 yılında Kellgren ve Lawrence tarafından geliştirilmiş olup, diz OA'nın radyolojik evrelendirilmesinde yaygın olarak kullanılmaktadır (6,7).

2.3.2. Evreleme Kriterleri

- Evre 0 (Normal): Radyolojik OA bulgusu yok
- Evre 1 (Şüpheli): Eklem aralığında şüpheli daralma, olası osteofit oluşumu
- Evre 2 (Hafif): Kesin osteofitler, eklem aralığında hafif daralma
- Evre 3 (Orta): Orta büyüklükte osteofitler, belirgin eklem aralığında daralma
- Evre 4 (Şiddetli): Büyük osteofitler, şiddetli eklem aralığında daralma (6,7)

KL sınıflamasına göre OA şiddeti Şekil 2.1'de gösterilmiştir.



Şekil 2.1. Kellgren-Lawrence Evreleme Sistemi (55)

2.4. TIBBİ GÖRÜNTÜ ANALİZİNDE YAPAY ZEKA

2.4.1. Yapay Zeka, Makine Öğrenmesi ve Derin Öğrenme Temelleri

Yapay zeka (YZ), bilgisayar sistemlerinde insanların öğrenme, problem çözme ve karar verme gibi becerilerini simüle etme teknolojisidir (56).

Makine öğrenmesi, YZ'nin bir alt dalı olup, algoritmaların eğitim verilerindeki örüntüleri öğrenerek yeni veriler hakkında doğru çıkarımlar yapabilmesine odaklanır. Bu örüntü tanıma yeteneği sayesinde makine öğrenimi modelleri, önceden kesin

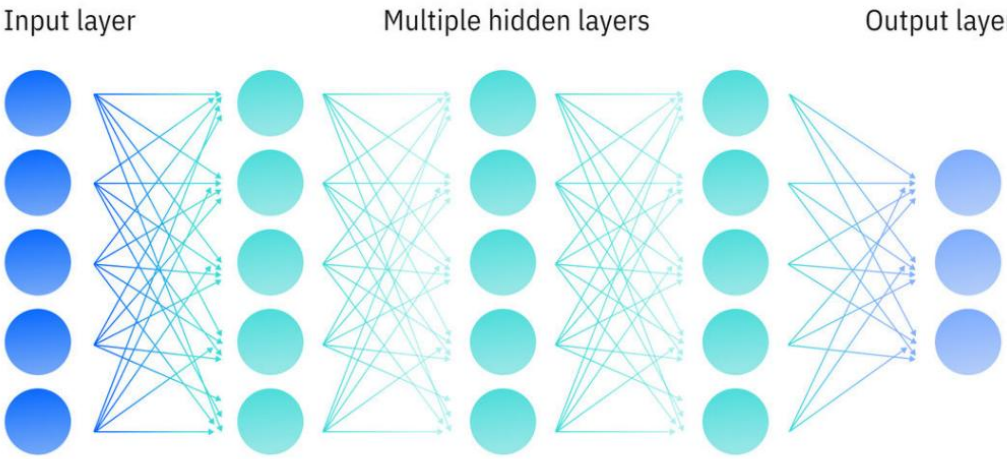
kurallarla programlanmış talimatlara ihtiyaç duymadan karar verebilir veya tahminlerde bulunabilir (57). Günümüzde bilgisayarlı görüden üretken YZ'ye, otonom araçlardan robotik sistemlere kadar birçok ileri düzey YZ uygulamasının temelini derin öğrenme modelleri oluşturmaktadır (58). Derin öğrenme algoritmaları, özellikle evrişim sinir ağları, tıbbi görüntü analizinde için en çok tercih edilen yöntem haline gelmiştir (59).

2.4.2. Yapay Sinir Ağları Temelleri

Yapay sinir ağı(YSA), sinir sisteminden esinlenen bir hesaplama modelidir. Bir sinir ağı, beyindeki nöronları modelleyen, "yapay nöron" olarak adlandırılan bağlı birimlerden oluşur. Yapay nöronlar arasındaki bağlantılar, beyindeki sinapsları temsil eder. Her nöron bağlı olduğu nöronlardan sinyaller alır, bu sinyalleri işler ve diğer nöronlara gönderir. Bu sinyal bir reel sayıdan oluşur ve her nöronun çıktısı, girdilerin toplamına uygulanan "aktivasyon fonksiyonu" adı verilen doğrusal olmayan bir fonksiyonla hesaplanır. Nöronlar arası bağlantılardaki sinyalin gücü ise öğrenme sürecinde ayarlanan ağırlıklar tarafından belirlenir.

Genellikle nöronlar katmanlar hâlinde düzenlenir. Farklı katmanlar, girdilere farklı işlemler uygular. Sinyaller, girdi katmanından, çıktı katmanına doğru ilerler ve bu yolculuk sırasında birden fazla ara katmandan (gizli katman) geçebilir. Bir ağda iki veya daha fazla gizli katman bulunuyorsa, bu yapı "derin sinir ağı" olarak adlandırılır.

Yapay sinir ağları; öngörüsel modelleme, uyarlanabilir kontrol ve yapay zekâ alanındaki çeşitli problemlerin çözümünde kullanılır. Deneyimlerden öğrenebilir, karmaşık ve birbiriyle ilgisiz görünen bilgilerden anlamlı sonuçlar çıkarabilirler (60). YSA örneği aşağıdaki şekilde gösterilmiştir (58).



Şekil 2.2. Yapay Sinir Ağı Örneği

2.4.3. Yapay Sinir Ağlarının Öğrenme Türleri

YSA'lar, onlara öğretim örnekleri verilerek ve belirli bir öğrenme kuralına göre ağırlıklarını değiştirmelerine izin verilerek eğitilebilir. Öğrenme türleri genel olarak üç başlıkta toplanabilir:

1. **Denetimli Öğrenme:** Ağa girdiler ve karşılık gelen çıktılar birlikte verilir. Bu girdi-çıktı çiftleri bir dış sistem tarafından sağlanır.
2. **Denetimsiz Öğrenme:** Ağ, verilen etiketlenmemiş veriyi içindeki örüntülere göre analiz ederek kümelere ayırır.
3. **Pekiştirmeli Öğrenme:** Bu öğrenme türü, yukarıda bahsedilen iki öğrenme türünün bir ara formu olarak düşünülebilir ve modelin, bir problemi çözmek için kendi başına ardışık kararlar vererek en uygun çözümü bulmasını sağlar (61).

2.4.4. Geri Yayılım Algoritması

Geri yayılım algoritması, YSA eğitimindeki temel bir yöntemlerden biridir. Bu algoritma ileri beslemeli sinir ağlarında uygulanır ve amacı, ağın çıktısı ile hedef çıktı arasındaki hatayı azaltarak ağın eğitim verilerini etkili öğrenmesini sağlamaktır. Algoritma süreci, eğitim örneğinin sinir ağına verilmesi ile başlar.

Daha sonra ağın çıktısı, veri setindeki hedef çıktı ile karşılaştırılır ve her çıkış nöronundaki hata hesaplanır. Her nöron için ölçekleme faktörü belirlenir ve çıktının hedefe ulaşması için ne kadar artırılması veya azaltılması gerektiği hesaplanır (61).

2.4.5. Evrişim İşlemi

Evrişim, derin öğrenmede, özellikle bilgisayarlı görü alanında temel bir işlem olup, evrişimli sinir ağlarının (ESA) ana yapı taşı olarak görev yapar. Modellerin, görüntülerdeki kenar ve doku gibi temel özelliklerden, karmaşık nesne biçimlerine kadar tüm özellikleri otomatik bir şekilde öğrenmesini sağlar.

Bu işlem, çekirdek olarak adlandırılan küçük bir filtreyi giriş görüntüsü üzerinde kaydırarak, kenarlar, dokular veya şekiller gibi belirli örüntüleri öne çıkaran özellik haritaları oluşturmayı içerir. Bu yöntem, görsel korteksin organizasyonundan esinlenmiştir (62).

2.4.6. Evrişimli Sinir Ağları (ESA)

ESA, YSA'ların gelişmiş versiyonu olup özellikle ızgara benzeri matris veri kümelerinden öznitelik çıkarmak amacıyla tasarlanmıştır (63). ESA'lar uzamsal paternleri algılayabildikleri için sıkça tercih edilir. Tıbbi görüntüleri yorumlayabilir, yüz tanıyabilir, görselleri sınıflandırabilir ve nesne tespiti yapabilirler. Bu tür işlemler yerel ayrıntılara dayanır ve ESA'lar bu ayrıntıları etkili bir şekilde yakalar (64). ESA'lar, görüntü tanıma, işleme ve sınıflandırma için kullanılan bir derin öğrenme modelidir. Karar ağaçları veya destek vektör makinesi gibi geleneksel modellerin aksine, ESA'lar kenarlar veya şekiller gibi örüntüleri otomatik olarak algılamak için filtreler kullanır (65).

2.4.7. Evrişimli Sinir Ağları (ESA) Katmanları

ESA, dört ana bileşenden oluşur:

- Evrişim Katmanları
- ReLU Aktivasyon Fonksiyonu
- Pooling Katmanları
- Tam Bağlı Katmanlar

Bu bileşenler, insan beyninin görsel verilerde örüntüleri ve özellikleri tanıma biçimini taklit etmeye yardımcı olur.

Evrişim Katmanları: Evrişim katmanı, ESA'nın temel yapı taşıdır. Burada görüntü matrisi üzerinde kayan filtre uygulanır. Her filtre, görüntüden çizgiler, eğriler, şekiller gibi belirli bir örüntüyü tanımak için kullanılır.

ReLU Aktivasyon Fonksiyonu: ReLU aktivasyon fonksiyonu her evrişim adımının ardından devreye girer. Bu fonksiyon, görüntüde yer alan özellikler arasındaki doğrusal olmayan bağıntıların öğrenilmesini sağlayarak ağın çeşitli desen ve örüntüleri daha etkili bir şekilde ayırt etme kapasitesini güçlendirir.

Havuzlama Katmanı: Havuzlama katmanının amacı, evrişim sonucu elde edilen matristen en önemli özellikleri seçmektir. Bu işlem, boyut indirgeme ve hafıza kullanımında azalma sağlar ve aşırı uyum riskini düşürür. Son havuzlama katmanı, özellik haritasını düzleştirir ve tam bağlı katmana aktarır.

Tam Bağlı Katmanlar: Tam bağlı katmanlar, son havuzlama katmanından gelen düzleştirilmiş matrisle çalışır. Son olarak, softmax aktivasyon fonksiyonu ile olasılık değerleri üretilir ve en yüksek olasılığa sahip etiket, modelin tahmini olur.

ESA'da Aşırı Uyum ve Düzenleme: Aşırı uyum modelin eğitim verilerini ezberlemesi ve gürültüleri de öğrenmesi durumudur. Bu durumda model eğitim verisinde iyi performans gösterirken, yeni verilerde başarısız olur.

ESA'lar, yüksek karmaşıklıkları ve büyük veri kümelerindeki detayları öğrenme kapasiteleri nedeniyle aşırı uyuma karşı hassastır. Bunun önüne geçmek için çeşitli düzenleme teknikleri uygulanabilir (65).

Hiyerarşik Öğrenme: ESA’lar; evrişim, havuzlama ve tam bağlı katmanlar gibi katmanlarla geri yayılım mekanizması sayesinde uzamsal özelliklerin hiyerarşik düzeylerini otomatik ve uyarlanabilir biçimde öğrenir. Evrişim katmanının temel

amacı, kenarlar, çizgiler ve diğer görsel unsurlar gibi ayırt edici yerel motifleri tespit etmektir.

Havuzlama katmanı, özellik haritalarının boyutlarını etkili bir şekilde küçültmek ve görüntü içindeki tespit edilen anlamsal özelliklerin şekil ve konum değişimlerine karşı dayanıklı kalması amacıyla kullanılır. Bir katmanın çıktısı bir sonraki katmana aktarıldıkça, çıkarılan özellikle aşamalı biçimde daha karmaşık hâle gelir. Derin ESA mimarisi, doğal nöromorfik çok katmanlı ağı taklit ederek, düşük seviyeden yüksek seviyeye kadar desenlerin hiyerarşik temsilini otomatik olarak öğrenir ve belirli bir görev için en önemli özellikleri belirler. ESA'lar, el yapımı özellik çıkarma ve uzman tarafından segmentasyon gerektirmez (66,67).

2.4.8. Transfer Öğrenimi (Transfer Learning)

Transfer öğrenimi, bir problem üzerinde öğrenilmiş özelliklerin alınarak benzer bir başka problemde yeniden kullanılması prensibine dayanır. Transfer öğrenimi genellikle, sıfırdan bir model eğitmek için yeterli veriye sahip olunmadığında tercih edilir.

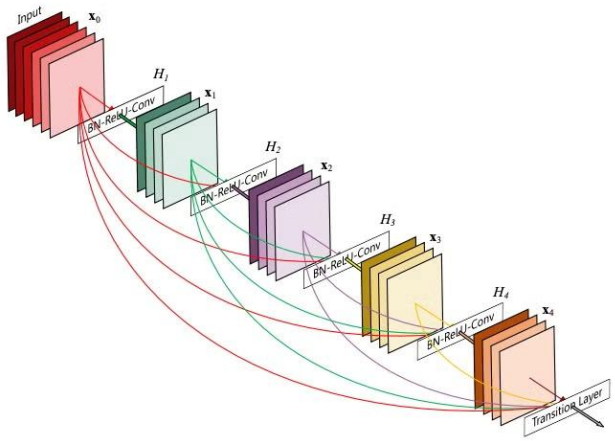
Derin öğrenme bağlamında transfer öğrenme genellikle şu adımlarla uygulanır:

1. Daha önce eğitilmiş bir modelin katmanları alınır.
2. Bu katmanlar dondurulur, yani içerdiği bilgilerin sonraki eğitimlerde bozulmaması sağlanır.
3. Dondurulmuş katmanların üzerine yeni, eğitilebilir katmanlar eklenir. Bu yeni katmanlar, eski modelin öğrendiği özellikleri kullanarak yeni veriler üzerinde tahmin yapmayı öğrenir.
4. Yeni katmanlar, mevcut veri kümesiyle eğitilir.
5. Son olarak, isteğe bağlı bir aşama olan ince ayar yapılabilir. Bu aşamada, modelin tamamı veya bir kısmı yeniden eğitime açılır ve düşük bir öğrenme oranıyla yeni veri üzerinde eğitilmeye devam edilir. Bu işlem, önceden öğrenilmiş özelliklerin yeni veriyle daha uyumlu hale getirilmesini sağlayarak performansta anlamlı iyileşmeler elde edilmesine katkı sağlayabilir (68).

2.5. DENSENET MİMARİSİ

DenseNet, ESA yapısında önerdiği yoğun bağlantı düzeniyle, kaybolan gradyan sorununa çözüm getirmiş, özellik tekrar kullanımı ve parametre verimliliği sağlayarak bilgisayarlı görü alanında önemli bir gelişme sağlamıştır. Klasik ESA yapılarında her katman yalnızca bir sonraki katmana bağlanırken, DenseNet mimarisinde bir blok içindeki tüm katmanlar birbirine doğrudan bağlanır. Böylelikle her katman, kendisinden önceki tüm katmanların çıktılarını girdi olarak alır.

Bu yoğun bağlantı yapısı, ağ boyunca bilgi akışını güçlendirir ve öğrenmeyi daha verimli hale getirir. DenseNet mimarisi aşağıdaki şekilde gösterilmiştir (69).



Şekil 2.3. DenseNet Mimarisi

2.6. VISION TRANSFORMER

2.6.1. Vision Transformer'ın Gelişimi

Transformer mimarisi, doğal dil işleme için standart haline gelmiştir, ancak bilgisayarlı görüdeki uygulamaları daha sınırlıdır. Vision Transformer(ViT), görüntü yamalarına doğrudan uygulanan saf bir transformer'dır ve görüntü sınıflandırma görevlerinde çok iyi performans gösterir (70).

2.6.2. ViT'nin Transfer Öğrenimiyle Kullanımı

Büyük miktarda veriyle ön eğitim yapıldığında ViT mimarisi en başarılı evrişimli sinir ağlarına göre mükemmel sonuçlar elde eder ve eğitim için önemli ölçüde daha az hesaplama kaynağı gerektirir (70).

2.7. GRAD-CAM YÖNTEMİ

2.7.1. Grad-CAM Yöntemi Kullanımı

Gradient-ağırlıklı Class Activation Mapping(Grad-CAM), diz OA sınıflandırmasında model yorumlanabilirliği için kullanılan bir tekniktir ve görüntü içerisinde sınıf belirleyici bölgeleri vurgulamak için görsel ısı haritaları üretir. Model eğitimi sonrasında uygulanır ve sınıflandırılan diz OA görüntülerinde modelin odaklandığı bölgeleri görselleştirir (71).

2.7.2. Klinik Açıklanabilirlik

Grad-CAM, modelin şeffaflığını ve güvenilirliğini artırır, klinik beklentilerle uyumunu sağlayarak diz OA teşhisinde yorumlanabilirliği geliştirir. Bu teknik, doktorların sınıflandırma sırasında modelin eklem boşluğu daralması veya kemik çıkıntısı oluşumu gibi tıbbi olarak ilgili yapılara odaklanıp odaklanmadığını değerlendirmesine olanak tanır (71).

Bu görselleştirme, modelin tanısal sürecini anlamayı kolaylaştırmanın yanı sıra, modelin klinik açıdan önemli özellikleri tespit edebildiğini de ortaya koyar. Böylece YZ destekli tanı araçlarında kullanılabilirliğini desteklerken, klinik karar verme süreçlerine de katkı sağlar (72).

2.8. PERFORMANS DEĞERLENDİRME METRİKLERİ

Bu çalışmada YZ modellerinin performansını ölçmek amacıyla birden fazla değerlendirme ölçütünden yararlanılmıştır. Bu ölçütler, modelin sınıflandırma yeteneğini farklı perspektiflerden inceleme olanağı sağlamıştır.

Karmaşıklık matrisi, sınıflandırıcının performansını her sınıf için doğru ve hatalı tahmin sayılarını göstererek resmeder. Bu matris, dikey eksende gerçek sınıfı ve yatay eksende model tarafından öngörülen sınıfı temsil eder. Bu temel sınıflandırma üzerinden aşağıdaki performans metrikleri hesaplanmaktadır:

- **Doğruluk:** Toplam doğru tahminlerin veri kümesindeki toplam örnek sayısına oranıdır.
- **Duyarlılık:** Gerçek pozitif tahminlerin, veri kümesindeki toplam pozitif örnek sayısına oranıdır ve modelin pozitif örnekleri doğru sınıflandırma yeteneğini gösterir.
- **Kesinlik:** Doğru sınıflandırılan gerçek pozitiflerin, model tarafından pozitif olarak sınıflandırılan toplam örnek sayısına oranıdır ve modelin pozitif sınıflandırmalarda güvenilirliğini vurgular.
- **F1-skoru:** Kesinlik ve duyarlılık puanlarının harmonik ortalamasıdır (73).

2.9. GÖZLEMCİLER ARASI VE GÖZLEMCİ İÇİ UYUM METRİKLERİ

Bu çalışmada hem gözlemcilerin aynı vakalar üzerinde yaptıkları değerlendirmelerin tutarlılığı, hem de aynı gözlemcinin 3 ay arayla aynı vakalar üzerinde yaptığı tahminlerin tutarlılığı ölçülmüştür. Bu çalışmada, gözlemciler arası ve gözlemci içi uyumu ölçmek için iki temel metrik kullanılmıştır:

2.9.1. İkinci Dereceden Ağırlıklı Kappa(İDAK)

İDAK, iki değerlendiricinin öğeleri sıralı bir ölçekte sınıflandırmalarındaki uyum düzeyini belirleyen istatistiksel bir göstergedir. Değerleri -1 (anlaşmazlık) ile 1 (anlaşma) arasında değişir ve 0 değeri, rastgele bir süreçten öte bir uyum olmadığını ifade eder. Bu, kategorik unsurlar için uyumu hesaplayan Cohen's Kappa'nın bir türevidir. İDAK'ın öne çıkan yanı, kategoriler arasındaki sıralı bağlantıyı dikkate almasıdır. Daha büyük uzaklıktaki yanlış tahminlere daha fazla ağırlık tanıyarak, bunları yakın olan yanlış tahminlere göre daha fazla cezalandırır (74).

2.9.2. Krippendorff's Alpha

Krippendorff's alfa, veri kümesi düzeyinde bir ölçüt olup değerlendiriciler arası güvenilirliği nicel olarak belirlemek için kullanılır.

Diğer uyum ölçütlerinden farklı olarak şu özelliklere sahiptir:

- Herhangi sayıda değerlendiriciyle çalışır.
- Eksik verileri yönetir.
- Nominal, ordinal, aralıklı ve oranlı veri türlerini destekler (75,76).

2.10. RECEIVER OPERATING CHARACTERISTIC CURVE (ROC) – ALICI-İŞLETİM KARAKTERİSTİĞİ EĞRİSİ

ROC eğrisi, bir modelin farklı eşik değerleri üzerindeki performansını görsel olarak gösteren bir grafiktir. ROC eğrisi, her olası eşik değeri için doğru pozitif oranı(DPO) ve yanlış pozitif oranı(YPO) hesaplanarak, DPO'nun YPO'ya karşı grafiğe dökülmesiyle elde edilir. Kusursuz bir model, belirli bir eşikte DPO = 1,0 ve YPO = 0,0 değerine ulaşır.

2.10.1. Eğri Altındaki Alan - Area Under the Curve (AUC)

ROC eğrisinin altındaki alan olan AUC, modelin rastgele seçilmiş bir pozitif örneğe negatif bir örnekten daha yüksek bir skor verme olasılığını gösterir. Tam başarılı bir modelin AUC değeri 1,0'dır; bu, modelin %100 ihtimalle pozitif örneği negatif örnekten daha yüksek olasılıkla sınıflandırdığını gösterir. Başka bir deyişle, AUC değeri, eşik değeri nerede olursa olsun, modelin rastgele seçilen bir "pozitif" örneği "negatif" örnekten daha doğru şekilde sıralama olasılığını belirtir. Daha somut bir örnekle açıklamak gerekirse, AUC değeri 1.0 olan bir diz OA ikili sınıflandırma modeli, her zaman rastgele seçilmiş bir OA'lı (pozitif) diz grafisini, rastgele seçilmiş bir normal (negatif) diz grafisinden daha yüksek OA varlığı olasılığıyla puanlar. Ancak nihai sınıflandırma, seçilen eşik değerine bağlıdır.

İkili sınıflandırma yapan bir modelde, eğer model yazı–tura atar gibi tamamen rastgele tahmin yapıyorsa, ROC eğrisi (0,0) noktasından (1,1) noktasına uzanan bir

köşegen olur. Bu modelin AUC değeri 0.5 olup bu modelin seçilen bir OA'lı veya bir normal görüntüyü doğru sırayla derecelendirme olasılığının %50 olduğu anlamına gelir.

Aynı mantıkla, AUC değeri 0.5 olan bir diz OA sınıflandırma modeli, rastgele seçilen bir OA'lı diz grafisine normal bir grafiden yalnızca %50 ihtimalle daha yüksek OA varlığı olasılık puanı verir. Bu da modelin rastgele tahmin eden bir sistemden daha iyi performans göstermediğini ifade eder (77).

Bu çalışmada çok sınıflı sınıflandırmada, her sınıfın AUC'si diğer tüm sınıflara karşı hesaplanmıştır. Bu yaklaşım, çok sınıflı durumu çok etiketli sınıflandırma gibi ele alır ve sınıf dengesizliğine karşı hassastır; dolayısıyla Macro ortalama alındığında da küçük sınıfların etkisi görünür (78).

- **Macro AUC:** Her sınıf için metriği hesaplar ve sınıf sayısına göre ağırlıksız ortalamasını alır.
- **Micro AUC:** Metriği, tüm sınıflardan bağımsız olarak toplam doğru pozitif, yanlış negatif ve yanlış pozitif sayıları üzerinden küresel şekilde hesaplar.
- **Ağırlıklı AUC:** Her sınıf için metriği hesaplar ve örnek sayısına göre ağırlıklı ortalamasını alır.

Her ortalama AUC hesap yöntemi kendi avantajlarına sahip olsa da uygun yöntemi seçerken sınıf dengesizliği önemli bir kriterdir. Eğer sınıfların örnek sayıları farklıysa, azınlık sınıflara da çoğunluk sınıflarla eşit ağırlık verecek şekilde Macro AUC kullanmak daha bilgilendirici olabilir (79).

2.11. İSTATİSTİKSEL ANLAMLILIK TESTLERİ

Çalışmada radyolog tahminlerindeki YZ öncesi-sonrası değişimin istatistiksel anlamlılığını değerlendirmek için Wilcoxon işaretli sıra testi (80) ve işaret testi (81) kullanılmıştır.

2.11.1. Wilcoxon İşaretli Sıra Testi

Wilcoxon işaretli sıra testi, bağımlı verilerde t-testine alternatif olarak kullanılan parametrik olmayan bir testtir. Bu test, özellikle t-testinin varsayımlarının (örneğin verilerin metrik ölçekte olması ve normal dağılması) sağlanmadığı durumlarda kullanılır. Sıralı veya normal dağılmayan veriler için oldukça uygundur.

Normal dağılım gerektirmediğinden, tekrarlanan ölçümler veya eşleştirilmiş gözlemler üzerinde yapılan analizlerde etkili bir alternatiftir.

Wilcoxon İşaretli Sıra Testinin Temel Özellikleri:

- **Parametrik olmayan yapısı:** Test, sıralı ölçekli veya normal dağılmayan verilere uygundur.
- **W-istatistiği:** Test, W-istatistiğini hesaplar. Bu istatistik, 10'dan fazla eşleştirilmiş gözlem olduğunda yaklaşık olarak normal dağılım gösterir (82).

2.11.2. İşaret Testi

İşaret testi, eşleştirilmiş gözlemler arasındaki tutarlı farkları değerlendirmek için kullanılan non-parametrik bir istatistiksel testtir. Örneğin, deneklerin tedavi öncesi ve sonrası ölçülen ağırlığı gibi ikili gözlemler verildiğinde, işaret testi bu ikililerdeki değerlerden birinin (örneğin tedavi öncesi) diğerine (örneğin tedavi sonrası) göre sistematik olarak daha büyük ya da daha küçük olup olmadığını belirler. (83).

3. GEREÇ VE YÖNTEM

“Derin Öğrenme Mimarileri ile Kellgren-Lawrence Evreleme Yapan Modelin Retrospektif Ön-Arka Diz Grafileri Üzerinde Değerlendirilmesi ve Radyolog Karar Süreçlerine Etkisinin İncelenmesi” konulu tez, Sağlık Bilimleri Üniversitesi Gülhane Tıp Fakültesi Dekanlığı’nın 17.06.2025 tarihli ve E-86241737-100-489113 sayılı yazısı ile verilmiştir. Çalışma, Ankara Bilkent Şehir Hastanesi Tıbbi Araştırmalar Bilimsel ve Etik Değerlendirme Kurulu’nun 02.07.2025 tarihli, TABED 1-25-1475 sayılı yazısıyla (Ek 1) onay almıştır. Bu çalışma, STARD-AI (The Standards for Reporting Diagnostic Accuracy -Artificial Intelligence) kılavuzuna uygun olarak tasarlanmış ve raporlanmıştır (84). Bu çalışmada yazarlar arasında herhangi bir çıkar çatışması bulunmamaktadır.

3.1. ÇALIŞMANIN PLANI

Bu çalışma, açık erişim ve hastane verileriyle yapay zeka geliştirme ve doğrulama çalışmasıdır. Çalışma iki ana aşamadan oluşmaktadır: Birinci aşamada açık kaynak veri seti (Mendeley Data KOA-SG v1.0) (85) kullanılarak on dört farklı derin öğrenme modelinin eğitimi ve iç doğrulaması yapılmıştır. İkinci aşamada ise Ankara Bilkent Şehir Hastanesi'nden elde edilen harici veri seti ile en iyi performans gösteren modelin gerçek dünya performansının test edilmesi ve radyolog performansına etkilerinin karşılaştırılması gerçekleştirilmiştir.

3.2. VERİ SETLERİ

3.2.1. Eğitim ve İç Doğrulama Veri Seti

Çalışmada kullanılan ana veri seti, Chen tarafından 2018 yılında Mendeley Data platformunda yayınlanan "Knee Osteoarthritis Severity Grading Dataset" (KOA-SG v1.0) olup, açık erişimlidir (85). Bu veri seti, çok merkezli Osteoarthritis Initiative’den elde edilen toplam 8.260 ön-arka diz grafisinden oluşmaktadır (86).

Veri Seti Dağılımı:

Eğitim Seti:

- Toplam: 5.778 örnek
- Evre 0: 2.286 örnek
- Evre 1: 1.046 örnek
- Evre 2: 1.516 örnek
- Evre 3: 757 örnek
- Evre 4: 173 örnek

Doğrulama Seti:

- Toplam: 826 örnek
- Evre 0: 328 örnek
- Evre 1: 153 örnek
- Evre 2: 212 örnek
- Evre 3: 106 örnek
- Evre 4: 27 örnek

Test Seti:

- Toplam: 1.656 örnek
- Evre 0: 639 örnek
- Evre 1: 296 örnek
- Evre 2: 447 örnek
- Evre 3: 223 örnek
- Evre 4: 51 örnek

Veri seti, %70 eğitim (5.778 görüntü), %10 doğrulama (826 görüntü) ve %20 test (1.656 görüntü) olacak şekilde rastgele bölünmüştür. Bölünme sırasında evre dengesi korunmuştur.

Mendeley veri seti sınıf dağılımı aşağıdaki tabloda gösterilmiştir.

Tablo 3.1. Mendeley Veri Seti Sınıf Dağılımı

KL	Eğitim	Doğrulama	Test	Toplam
Evre 0	2.286	328	639	3.253
Evre 1	1.046	153	296	1.495
Evre 2	1.516	212	447	2.175
Evre 3	757	106	223	1.086
Evre 4	173	27	51	251
Toplam	5.778	826	1.656	8.260

3.2.2. Harici Doğrulama Veri Seti

Katılımcı Seçimi ve Örneklem Büyüklüğü:

Harici doğrulama setine dahil edilecek 100 hasta, 1 Nisan-30 Mayıs tarihleri arasında Ankara Bilkent Şehir Hastanesinde ön-arka diz grafisi çekimi yapılan 18-80 yaş aralığındaki bireyler arasından seçilmiştir. Cinsiyet dengesi gözetilerek kadın ve erkek hastaların temsil edilmesi sağlanmıştır.

Dahil edilme kriterleri:

- 18-80 yaş arası bireyler
- Yüksek kaliteli standart ön-arka diz radyografisi bulunanlar
- Diz protezi veya implant geçirmemiş hastalar
- Yeterli klinik veri kaydı olan hastalar
- DICOM formatında kayıtlı görüntüsü olan hastalar

Dışlama kriterleri:

- Diz protezi veya implant varlığı
- Fraktür varlığı
- Yetersiz görüntü kalitesi (pozisyonlama hatası, hareket artefaktı, düşük kontrast)
- Diğer eklem hastalıkları (tümör, enfeksiyon, enflamatuar artropati)

- Daha önce diz cerrahisi geçirmiş hastalar (artroskopi, osteotomi, ligament rekonstrüksiyonu)
- Klinik verisi eksik olan hastalar

Harici veri seti demografik özellikleri:

100 hastanın demografik dağılımı aşağıdaki tabloda gösterilmiştir.

Tablo 3.2. Harici Doğrulama Setinin Demografik Özellikleri

Özellik	Değer
Toplam hasta sayısı	100
Cinsiyet dağılımı	52 kadın (%52), 48 erkek (%48)
Genel yaş ortalaması	52,11 ± 17,67 yıl (aralık: 18-80)
Kadınlarda yaş ortalaması	54,21 ± 18,02 yıl (medyan: 56,5)
Erkeklerde yaş ortalaması	49,02 ± 17,18 yıl (medyan: 50,5)
Görüntüleme cihazı	GE Optima XR646 ve GE Discovery XR656 (GE Healthcare, USA)

Yaş gruplarına göre dağılım:

- 18-30 yaş: 13 hasta (%13)
- 31-40 yaş: 14 hasta (%14)
- 41-50 yaş: 16 hasta (%16)
- 51-60 yaş: 19 hasta (%19)
- 61-70 yaş: 17 hasta (%17)
- 71-80 yaş: 21 hasta (%21)

KL Evre dağılımı (Referans Radyolog Konsensüsü):

- Evre 0: 24 hasta (%24) - Normal eklem
- Evre 1: 18 hasta (%18) - Şüpheli osteoartrit
- Evre 2: 22 hasta (%22) - Hafif osteoartrit

- Evre 3: 17 hasta (%17) - Orta osteoartrit
- Evre 4: 19 hasta (%19) - Şiddetli osteoartrit

Harici set görüntüleri, son sene asistanı ve 10 yıldan fazla kas-iskelet sistemi radyolojisi deneyimi olan radyolog tarafından ayrı ayrı değerlendirilmiş ve konsensüs skorları referans standart olarak kabul edilmiştir. Görüş ayrılığı durumlarında, üçüncü bir deneyimli radyolog görüşü alınmıştır.

3.3. GÖRÜNTÜ ÖN İŞLEME

Tüm görüntüler anonimize edildikten sonra standardize edilmiş bir ön işleme protokolünden geçirilmiştir. İlk olarak, DICOM formatındaki ham görüntüler Python pydicom kütüphanesi kullanılarak PNG formatına dönüştürülmüştür (87).

3.3.1. Manuel Kırpma İşlemi

Derin öğrenme modelinin diz eklemine odaklanmasını sağlamak ve gereksiz anatomik yapıların (femur proksimal kısmı, tibia distal kısmı, yumuşak dokular) model performansını olumsuz etkilemesini önlemek amacıyla, tüm görüntüler manuel olarak kırpılmıştır. Kırpma işlemi şu şekilde gerçekleştirilmiştir:

- **Kırpma bölgesi:** Diz eklemi merkez alınarak, femoral kondiller, tibial plato, patella ve proksimal fibula dahil edilmiştir
- **Standart kırpma protokolü:** Her görüntüde eklem aralığı merkez noktası olarak belirlenerek simetrik şekilde tüm eklem aralığını ve eklem komşuluğundaki kemik yapıları dahil edecek şekilde standart bir kırpma protokolü uygulanmıştır.

3.3.2. Boyutlandırma ve Normalizasyon

Kırpılmış görüntüler ardından 224×224 piksel boyutuna yeniden boyutlandırılmıştır. Piksel değerleri [0,1] aralığına normalize edilmiştir. Transfer öğrenme için ImageNet ortalama ve standart sapma değerleri kullanılarak standardizasyon yapılmıştır.

Ön işleme özeti:

1. DICOM → PNG dönüşümü
2. Manuel kırpma (diz eklemi odaklı)
3. 224×224 piksel yeniden boyutlandırma
4. [0,1] normalizasyon
5. ImageNet standardizasyonu (88)

3.4. VERİ ARTIRMA

Eğitim setinde çok sayıda veri artırma tekniği uygulanmıştır. Eğitim seti veri artırma teknikleri şunları içermektedir:

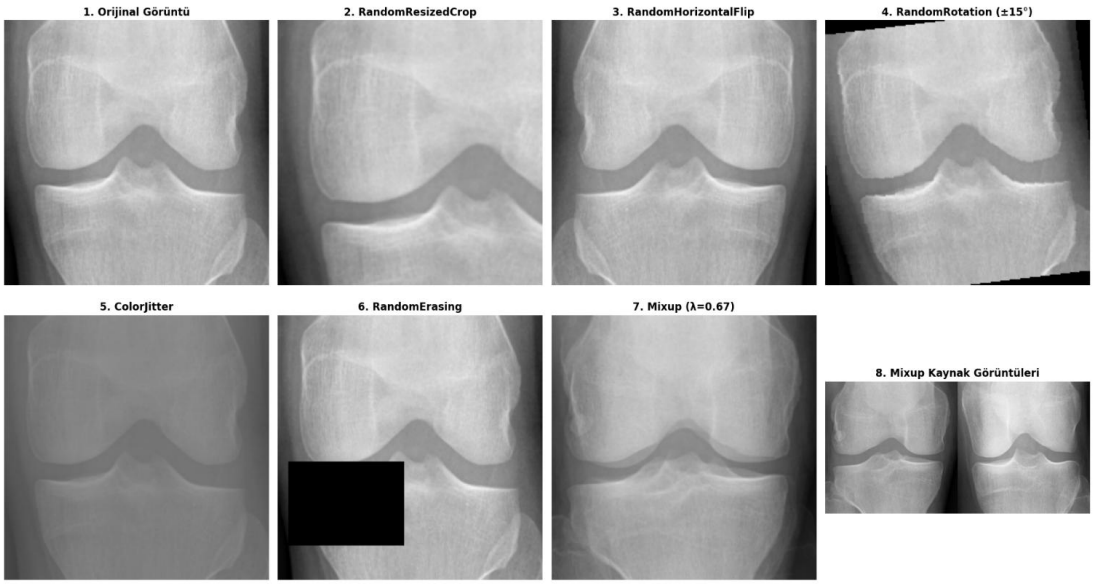
- **RandomResizedCrop():** Rastgele yeniden kırpma ile görüntünün belirli bir kesimini kırpar ve istenilen boyuta getirir (89)
- **RandomHorizontalFlip():** Rastgele bir olasılıkla yatay çevirme varyasyonlarını üretir (90)
- **RandomRotation():** Belirlenen açı aralığında rastgele döndürme ile görüntüyü rastgele olarak çevirir (91)
- **ColorJitter():** Bir görüntünün parlaklığını, kontrastını, doygunluğunu ve tonunu rastgele değiştirir (92)
- **RandomErasing():** Görüntü içerisinden rastgele bölgeleri silerek görüntü sınıflandırma başarısında iyileşme sağlar (93)

Veri artırma yöntemleri, görüntülerde ortaya çıkabilen parlaklık farklılıkları, örtülme, ölçek değişimleri, arka plan varyasyonları ve benzeri birçok önyargı kaynağını belirgin ölçüde azaltabilir (94). Doğrulama ve test setlerine veri artırma yapılmamıştır.

3.5. MIXUP VERİ KARIŞTIRMA STRATEJİSİ

Eğitim sürecinde Mixup veri karıştırma stratejisi uygulanmıştır. Mixup, iki görüntü x_1, x_2 ve iki etiket y_1, y_2 alarak sanal eğitim örnekleri üretir. Bu şekilde sinir ağlarının genelleme becerisini arttırır (95).

Yapılan veri arttırma ve karıştırma tekniği örnekleri aşağıdaki şekilde gösterilmiştir.



Şekil 3.1. Veri Arttırma ve Karıştırma Tekniği Örnekleri

3.6. MODEL MİMARİLERİ

Çalışmada on dört farklı derin öğrenme mimarisi kullanılmıştır.

3.6.1. Evrişimsel Sinir Ağı Tabanlı Modeller (n=10)

1. **ResNet-18:** 11,7 milyon parametre ve 18 katman derinliğine sahiptir; temel özelliği residual (atlama) bağlantılarıdır ve bu çalışmada ImageNet ön eğitimli ağırlıklar kullanılmıştır (96,97).
2. **ResNet-50:** 25,6 milyon parametre ve 50 katman derinliğine sahiptir; bottleneck mimarisi ile daha derin mimari sunmaktadır ve ImageNet ön eğitimli ağırlıklar kullanılmıştır (96,98).
3. **ResNet-101:** 44,5 milyon parametre ve 101 katman derinliğine sahiptir; daha derin bottleneck mimarisi ile gelişmiş özellik öğrenme kapasitesi sunmaktadır ve ImageNet ön eğitimli ağırlıklar kullanılmıştır (96,99).
4. **DenseNet-121:** 8,0 milyon parametre ve 121 katman derinliğine sahiptir; temel özelliği yoğun bağlantılardır (her katman önceki tüm katmanlara bağlı) ve ImageNet ön eğitimli ağırlıklar kullanılmıştır (69,100).

5. **DenseNet-169:** 14,1 milyon parametre ve 169 katman derinliğine sahiptir; daha derin yoğun bağlantı yapısı ile gelişmiş özellik yeniden kullanımı sunmaktadır ve ImageNet ön eğitimli ağırlıklar kullanılmıştır (69,101).
6. **DenseNet-201:** 20 milyon parametre ve 201 katman derinliğine sahiptir; çalışmada kullanılan DenseNet mimarileri içerisinde en yüksek derinliğe ve parametre sayısına sahiptir. ImageNet ön eğitimli ağırlıklar kullanılmıştır (69,102).
7. **EfficientNet-B0:** 5,3 milyon parametre ile compound scaling yöntemi ve verimlilik odaklı tasarım sunmaktadır; ImageNet ön eğitimli ağırlıklar kullanılmıştır (103,104).
8. **MobileNetV2:** 3,5 milyon parametre ile inverted residual bloklar içermekte ve mobil cihazlar için optimize edilmiştir; ImageNet ön eğitimli ağırlıklar kullanılmıştır (105,106).
9. **ConvNeXt-Tiny:** 28,6 milyon parametre ile modernize edilmiş ESA mimarisi, depthwise convolution ve LayerNorm kullanımı içermektedir; ImageNet-1K ön eğitimli ağırlıklar kullanılmıştır (107,108).
10. **ConvNeXt-Small:** 50,2 milyon parametre ile daha derin ConvNeXt mimarisi, gelişmiş özellik çıkarma kapasitesi sunmaktadır; ImageNet-1K ön eğitimli ağırlıklar kullanılmıştır (107,109).

3.6.2. Vision Transformer Modelleri (n=4)

1. **Swin-T (Swin Transformer Tiny):** 28,3 milyon parametre ile shifted window mekanizması, hiyerarşik özellik öğrenme içermektedir; ImageNet-1K ön eğitimli ağırlıklar kullanılmıştır (110,111).
2. **Swin-S (Swin Transformer Small):** 49,6 milyon parametre ile daha derin Swin Transformer mimarisi, gelişmiş global-local özellik kombinasyonu sunmaktadır; ImageNet-1K ön eğitimli ağırlıklar kullanılmıştır (110,112).
3. **ViT-B/16:** 86,6 milyon parametre ile 16×16 piksel patch boyutu kullanmaktadır. 12 transformer encoder blok ve multi-head self-attention mekanizması içermekte olup ImageNet-1K ön eğitimli ağırlıklar kullanılmıştır (70,113).

4. **ViT-L/16:** 304,3 milyon parametre ile 16×16 piksel patch boyutu kullanmaktadır. 24 transformer encoder blok ve daha geniş mimari ile daha fazla öğrenme kapasitesi sunmakta olup ImageNet-1K ön eğitimli ağırlıklar kullanılmıştır (70,114).

3.7. TRANSFER ÖĞRENME STRATEJİSİ

Tüm modeller için transfer öğrenme yaklaşımı uygulanmıştır. İlk olarak, ImageNet-1K veri setinde (115) (1,4 milyon görüntü, 1000 sınıf) eğitilmiş ağırlıklar başlangıç noktası olarak kullanılmıştır. Ardından tüm katmanlar eğitilebilir duruma getirilmiş ve tüm ağ diz grafisi veri seti üzerinde ince ayar yapılmıştır. Orijinal 1000 sınıflı sınıflandırma katmanı, 5 sınıflı (KL 0-4) yeni bir tam bağlı katman ile değiştirilmiştir. Her modelin çıkış katmanı ilgili mimariye uygun olarak değiştirilip, dropout (116) eklenmiştir. Eğer doğrulama başarısı 8 epoch boyunca artmazsa eğitim erken durdurulmuş ve en iyi kontrol noktası yüklenmiştir. Bu uygulamada amaç aşırı uyumun önüne geçmek ve eğitimi zamanında sonlandırmaktır (117).

ImageNet veri seti örneği aşağıdaki şekilde gösterilmiştir (118).



Şekil 3.2. ImageNet Veri Seti Örneği

3.8. GRAD-CAM GÖRSELLEŞTİRME

Model açıklanabilirliği için Grad-CAM yöntemi uygulanmıştır (119). Grad-CAM haritaları, radyologlara modelin hangi anatomik bölgelere (osteofit, eklem aralığı, subkondral kemik) odaklandığını ısı haritalarıyla göstermek için kullanılmıştır.

3.9. RADYOLOG PERFORMANS DEĞERLENDİRMESİ

Altı radyolog çalışmaya dahil edilmiştir.

Deneyim grupları:

- Az deneyimli (R1, R2): 2-3 yıl radyoloji deneyimi
- Orta deneyimli (R3, R4): 5-6 yıl radyoloji deneyimi
- Deneyimli (R5, R6): 15 yıl ve üzeri radyoloji deneyimi

Değerlendirme protokolü iki aşamadan oluşmuştur:

Aşama 1 (Yapay Zeka Öncesi):

Her radyolog harici veri setindeki 100 görüntüyü bağımsız olarak değerlendirmiş:

- Yapay zeka desteği olmadan KL evre atamış ve güven skoru 1-5 Likert ölçeğiyle (120) belirtilmiştir
- Grafileri değerlendirme süresi ölçülmüştür

Aşama 2 (Yapay Zeka Sonrası):

- Üç ay sonra aynı görüntüler tekrar değerlendirilmiştir
- Bu kez en iyi performans gösteren model tahminleri ve Grad-CAM görselleştirmeleri radyologlara sunulmuştur
- Radyologlar yapay zeka önerilerini göz önünde bulundurarak tekrar KL evre atamıştır
- Raporlama süresi ve güven skoru tekrar kaydedilmiştir

İki değerlendirme arasında 3 aylık süre (washing period) bırakılarak hafıza etkisi minimize edilmiştir.

Likert Ölçeği'nde Radyologlara KL evre kararlarına güvenme düzeyin nasıl? diye sorulmuş:

1. Hiç güvenmiyorum
2. Az güveniyorum
3. Orta düzeyde güveniyorum
4. Oldukça güveniyorum
5. Tamamen güveniyorum

Cevaplarından birini vermesi beklenmiştir.

3.10. PERFORMANS METRİKLERİ

Modeller değerlendirmesinde doğruluk, duyarlılık, kesinlik, F1-skoru metrikleri ve ROC eğrileri, Eğri Altındaki Alan kullanılmıştır. Her model için 5×5 karmaşıklık matrisi oluşturulmuş, gerçek ve tahmin edilen sınıfların dağılımı görselleştirilmiştir.

3.10.1. Radyologlar Arası Uyum Analizi

Radyologlar arası uyumu değerlendirmek için iki ölçüt kullanılmıştır:

- İkinci Dereceden Ağırlıklı Kappa (İDAK): Her radyolog çifti arasındaki uyum hesaplanmış ve İDAK değeri raporlanmıştır (121).
- Krippendorff's Alpha: Tüm radyologlar için genel uyum metriği olarak Krippendorff's Alpha hesaplanmıştır (75).

3.10.2. Gözlemci İçi Tutarlılık Analizi

Her radyologun iki zamanlı değerlendirmesi (YZ öncesi ve YZ sonrası) arasındaki tutarlılık İDAK ile ölçülmüştür (74). Bu analiz, radyologların kendi içlerindeki tutarlılığını ve YZ desteğinin bu tutarlılığa etkisini gösterir.

3.11. İSTATİSTİKSEL ANALİZ

Modelin geliştirme, istatistiksel analizi, sonuçların görselleştirilmesi için aşağıdaki yazılımlar kullanılmıştır.

- Model geliştirme için Python 3.12 (122), PyTorch 2.8 (123), torchvision (124) ve scikit-learn (125) kullanılmıştır.
- İstatistiksel analiz için numpy (126), scipy (127) ve scikit-learn (125) kullanılmıştır.
- Görselleştirme için matplotlib (128) ve seaborn (129) kullanılmıştır.
- Grad-CAM için pytorch-grad-cam kütüphanesi (130) kullanılmıştır.

4. BULGULAR

4.1. MODEL PERFORMANSLARI - MENDELEY TEST SETİ

On dört farklı derin öğrenme modelinin Mendeley test setindeki (n=1.656) performans sonuçları aşağıdaki tabloda özetlenmiştir.

Tablo 4.1. Mendeley Test Setinde Model Performansları

Model	Kategori	Doğruluk	Macro AUC	Eğitim-Test Farkı (%)
DenseNet-169	ESA	70,17	0,9079	0,93
ResNet-101	ESA	69,93	0,9072	1,48
Swin-S	ViT	68,84	0,8969	-1,64
ResNet-50	ESA	68,48	0,9051	1,70
DenseNet-121	ESA	68,24	0,8995	-0,47
ConvNeXt-Small	ESA	67,69	0,8941	-1,34
EfficientNet-B0	ESA	67,57	0,8986	5,36
DenseNet-201	ESA	67,33	0,9036	1,53
ResNet-18	ESA	67,27	0,8924	-1,35
MobileNetV2	ESA	66,61	0,9038	2,41
ConvNeXt-Tiny	ESA	65,16	0,8880	-1,73
ViT-L/16	ViT	60,57	0,8401	-5,55
Swin-T	ViT	59,42	0,8404	-3,88
ViT-B/16	ViT	52,36	0,7962	-4,13

Tablo 4.2. Mendeley Test Setinde Model Ailesi Performansları

Mimari Ailesi	Model Sayısı	Ortalama Test Doğruluğu (%)	Standart Sapma (Doğruluk)
ESA (Tümü)	10	67,85	1,43
ViT (Tümü)	4	60,30	6,75
Swin Transformer	2	64,13	6,66
Saf ViT	2	56,47	5,81
DenseNet Ailesi	3	68,58	1,46
ResNet Ailesi	3	68,56	1,35

Önemli bulgular:

- DenseNet-169, %70,17 test doğruluğu ile en yüksek test performansını göstermiştir.
- ResNet-101 %69,93 doğruluk ile ikinci sırada yer almıştır.
- ESA tabanlı modeller (n=10) ortalama %67,85 doğruluk elde etmiştir.
- ViT modelleri (n=4) ortalama %60,30 doğruluk ile ESA tabanlı modellerden belirgin şekilde düşük performans göstermiştir.
- Swin Transformer modelleri (n=2) ortalama %64,13 doğruluk ile saf ViT modellerinden (n=2, ortalama %56,47 doğruluk) daha iyi performans göstermiştir.
- ViT-L/16 (%60,57) ve ViT-B/16 (%52,36) ViT modelleri, ESA tabanlı modellerden belirgin şekilde düşük performans sergilemiştir.
- En iyi beş ESA modeli arasında doğruluk açısından 2,48 puanlık fark, tüm ESA modelleri arasında doğruluk açısından 5,01 puanlık fark vardır.

4.2. MODEL PERFORMANSI - HARİCİ HASTANE SETİ

Harici doğrulama setinde (Ankara Bilkent Şehir Hastanesi, n=100) en iyi performans gösteren model olan DenseNet-169 test edilmiştir.

Tablo 4.3. DenseNet-169 Harici Test Seti Performansları

Metrik	Değer
Genel Doğruluk	%65,00
Makro F1-Skoru	0,62
Ağırlıklı F1-Skoru	0,62
İkinci Dereceden Ağırlıklı Kappa	0,864
Micro AUC	0,8863
Macro AUC	0,8575

DenseNet-169'un harici sette evre bazlı performansı aşağıdaki tabloda gösterilmiştir.

Tablo 4.4. DenseNet-169 Evre Bazlı Performans (Harici Set)

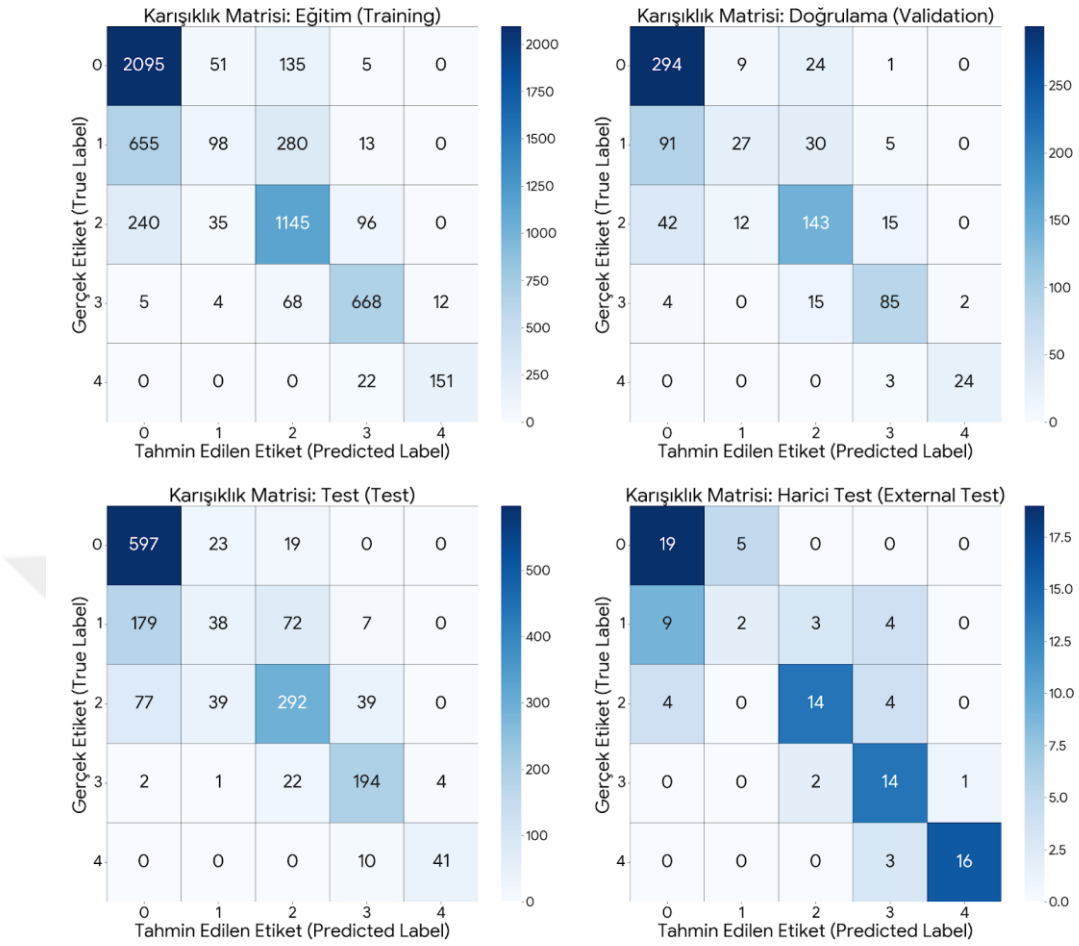
KL Evre	Kesinlik	Duyarlılık	F1-Skoru
Evre 0	0,5938	0,7917	0,6786
Evre 1	0,2857	0,1111	0,1600
Evre 2	0,7368	0,6364	0,6829
Evre 3	0,5600	0,8235	0,6667
Evre 4	0,9412	0,8421	0,8889

Gözlemler:

- Harici sette doğruluk Mendeley'e göre 5,17 puan düşüktür.
- Harici sette Macro AUC 0,8575 olup, Mendeley test setine göre (Macro: 0,9079) düşüş gözlenmiştir ancak halen yüksek performans sergilemektedir.
- KL evre 1 de model çok düşük performans göstermiştir (duyarlılık %11,1).
- İleri evrelerde (evre 3 ve 4) yüksek performans elde edilmiştir (duyarlılık %82-84).
- Normal eklem tespitinde de model başarılıdır (duyarlılık %79,2).

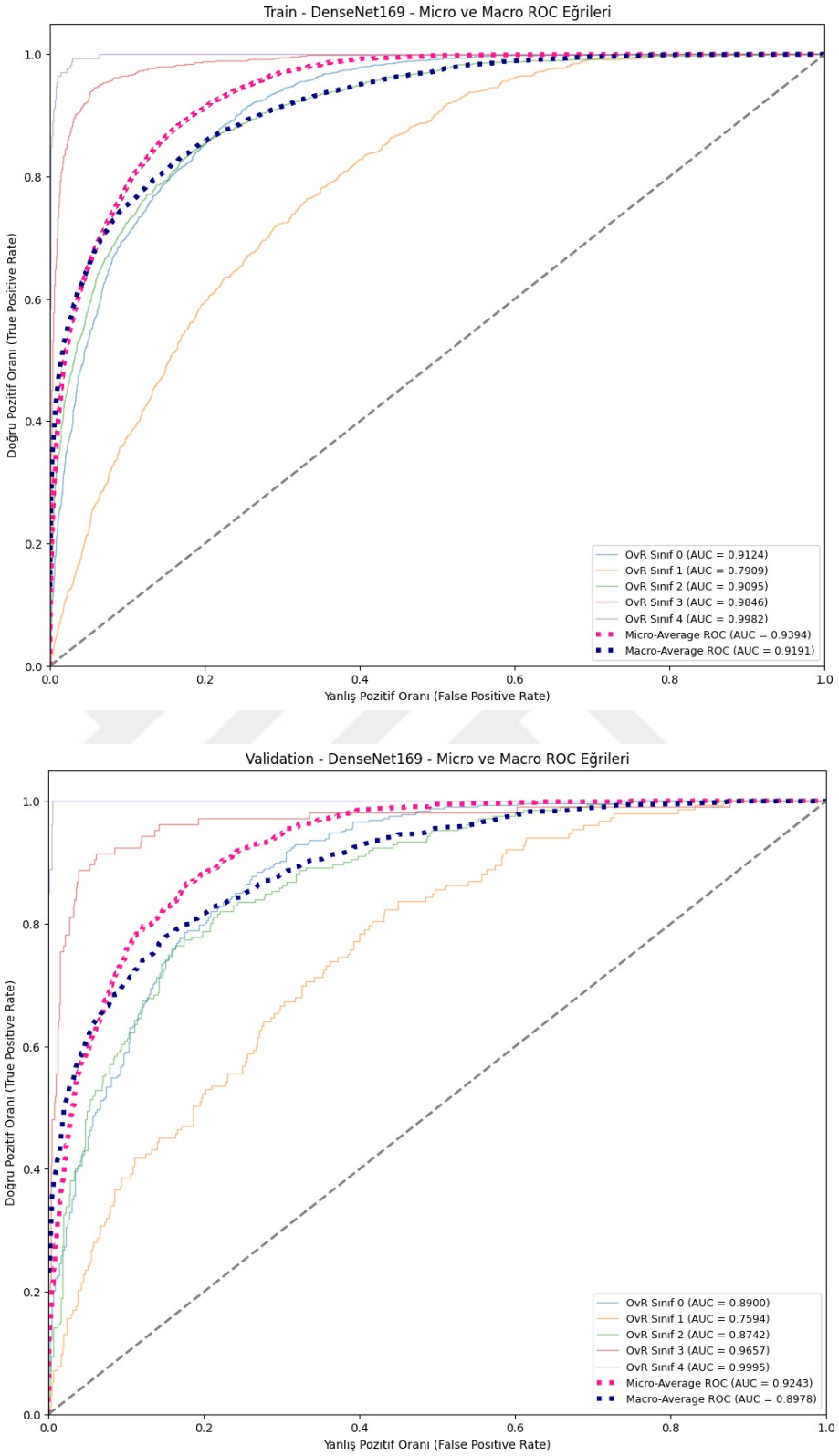
4.2.1. Yapay Zeka Karışıklık Matrisi

DenseNet-169 modelinin Mendeley ve harici sette evre bazlı performansı detaylı olarak incelenmiştir. DenseNet-169 Mendeley ve harici hastane verisi için karışıklık matrisleri aşağıdaki şekilde gösterilmiştir.

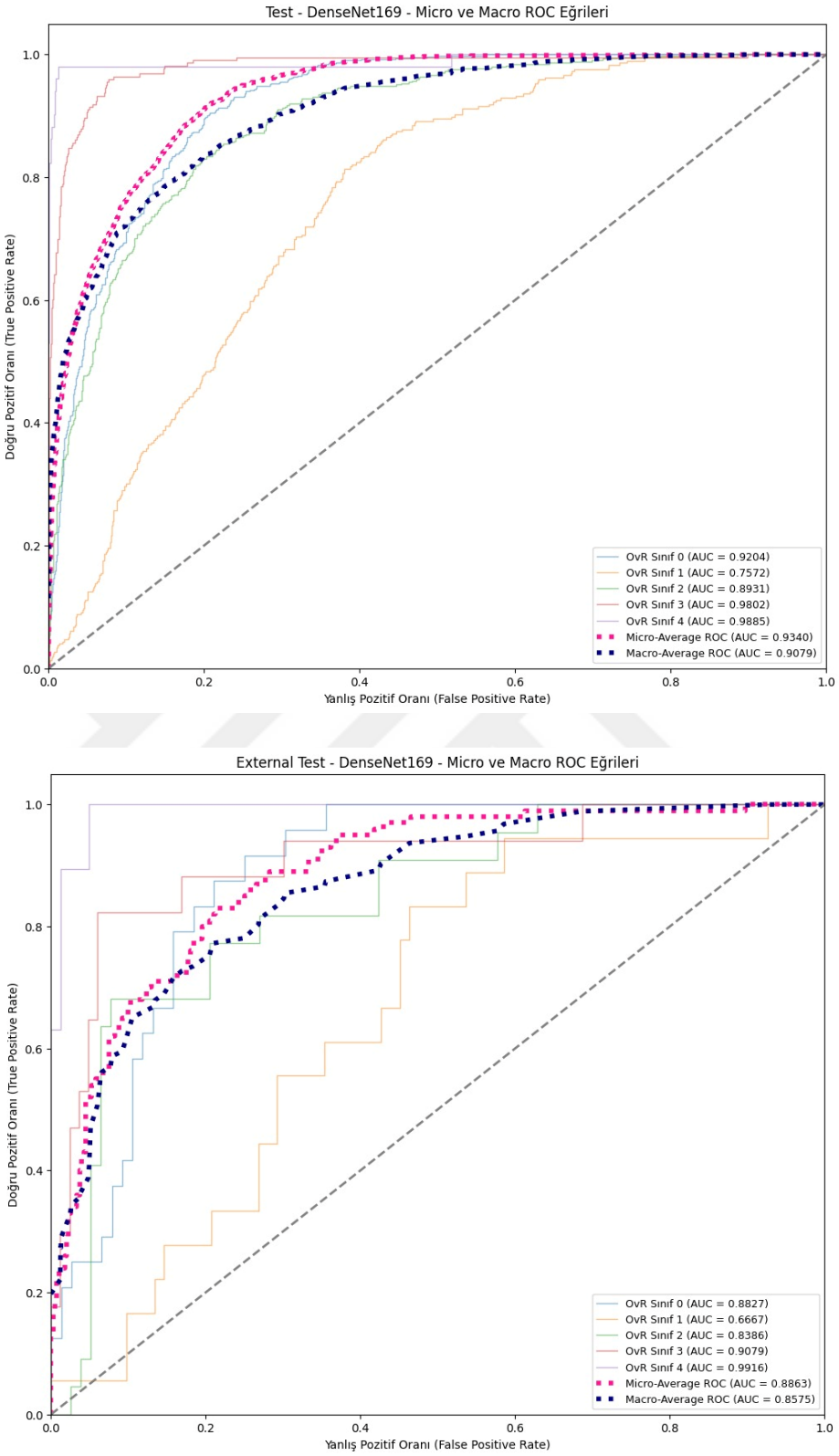


Şekil 4.1. DenseNet-169 Karışıklık Matrisi

DenseNet-169 ROC eğrileri eğitim ve doğrulama seti için Şekil 4.2'de, Mendeley ve harici hastane test verisi için Şekil 4.3'te gösterilmiştir.



Şekil 4.2. DenseNet-169 ROC Eğrileri Eğitim ve Doğrulama Seti



Şekil 4.3. DenseNet-169 ROC Eğrileri Mendeley ve Harici Hastane Test Verisi

Harici Hastane Verisinde Model Başarısının Değerlendirilmesi:

- Karışıklık matrisinde köşegendeki değerler doğru sınıflandırılan vakaları gösterir: 65/100 = %65 doğruluk
- En yüksek karışıklık Evre 0-Evre 1 ve Evre 1-Evre 2 arasında görülmektedir (toplam 17 vaka)
- Evre 4'te en yüksek doğruluk (%84,2, 16/19) elde edilmiştir
- Evre 1'de model en düşük performansı göstermektedir (%11,1 doğruluk, 2/18)
- Uzak evrelere hata çok nadir görülmektedir (örn: Evre 0→Evre 4 veya Evre 4→ Evre 0 yok)
- Model harici hastane verisinde 0,8575 Macro AUC ile yüksek başarı göstermiştir.

4.3. DETAYLI MODEL PERFORMANSI DEĞERLENDİRME

Tüm modellerin detaylı performansları aşağıdaki tablolarda gösterilmiştir.

Tablo 4.5. ConvNeXt ve DenseNet Modelleri

Model	Veri Seti	Doğruluk	Kesinlik	Duyarlılık	F1-Skoru	Macro AUC
ConvNeXt-S	Eğitim	0.6636	0.63	0.66	0.61	0.8794
ConvNeXt-S	Doğrulama	0.6550	0.61	0.66	0.61	0.8887
ConvNeXt-S	Test	0.6769	0.62	0.68	0.63	0.8941
ConvNeXt-T	Eğitim	0.6343	0.52	0.63	0.57	0.8668
ConvNeXt-T	Doğrulama	0.6489	0.53	0.65	0.58	0.8809
ConvNeXt-T	Test	0.6516	0.54	0.65	0.58	0.8880
DenseNet-121	Eğitim	0.6777	0.65	0.68	0.65	0.8904
DenseNet-121	Doğrulama	0.6731	0.65	0.67	0.65	0.8920
DenseNet-121	Test	0.6824	0.66	0.68	0.66	0.8995
DenseNet-169	Eğitim	0.7110	0.68	0.71	0.66	0.9191
DenseNet-169	Doğrulama	0.6937	0.68	0.69	0.66	0.8978
DenseNet-169	Test	0.7017	0.66	0.70	0.66	0.9079
DenseNet-169	Harici Hastane	0.6500	0.63	0.65	0.62	0.8575
DenseNet-201	Eğitim	0.6886	0.67	0.69	0.67	0.9055
DenseNet-201	Doğrulama	0.6780	0.68	0.68	0.67	0.8946
DenseNet-201	Test	0.6733	0.67	0.67	0.67	0.9036

Tablo 4.6. EfficientNet, MobileNetV2 ve ResNet Modelleri

Model	Veri Seti	Doğruluk	Kesinlik	Duyarlılık	F1-Skoru	Macro AUC
EfficientNet B0	Eğitim	0.7293	0.71	0.73	0.70	0.9264
EfficientNet B0	Doğrulama	0.6683	0.65	0.67	0.65	0.8944
EfficientNet B0	Test	0.6757	0.65	0.68	0.66	0.8986
MobileNetV2	Eğitim	0.6902	0.67	0.69	0.63	0.9126
MobileNetV2	Doğrulama	0.6719	0.65	0.67	0.63	0.8886
MobileNetV2	Test	0.6661	0.63	0.67	0.62	0.9038
ResNet-101	Eğitim	0.7141	0.69	0.71	0.68	0.9208
ResNet-101	Doğrulama	0.6792	0.66	0.68	0.66	0.8903
ResNet-101	Test	0.6993	0.68	0.70	0.68	0.9072
ResNet-18	Eğitim	0.6592	0.64	0.66	0.62	0.8833
ResNet-18	Doğrulama	0.6659	0.64	0.67	0.64	0.8832
ResNet-18	Test	0.6727	0.64	0.67	0.65	0.8924
ResNet-50	Eğitim	0.7018	0.67	0.70	0.66	0.9103
ResNet-50	Doğrulama	0.6768	0.65	0.68	0.65	0.8951
ResNet-50	Test	0.6848	0.65	0.68	0.65	0.9051

Tablo 4.7. Swin ve ViT (Vision Transformer) Modelleri

Model	Veri Seti	Doğruluk	Kesinlik	Duyarlılık	F1-Skoru	Macro AUC
Swin-S	Eğitim	0.6720	0.63	0.67	0.63	0.8884
Swin-S	Doğrulama	0.6683	0.63	0.67	0.63	0.8899
Swin-S	Test	0.6884	0.65	0.69	0.65	0.8969
Swin-T	Eğitim	0.5554	0.45	0.56	0.48	0.8106
Swin-T	Doğrulama	0.6138	0.51	0.61	0.54	0.8464
Swin-T	Test	0.5942	0.50	0.59	0.51	0.8404
ViT-B/16	Eğitim	0.4823	0.41	0.48	0.44	0.7670
ViT-B/16	Doğrulama	0.5351	0.44	0.54	0.48	0.7991
ViT-B/16	Test	0.5236	0.43	0.52	0.47	0.7962
ViT-L/16	Eğitim	0.5502	0.47	0.55	0.47	0.8291
ViT-L/16	Doğrulama	0.5799	0.47	0.58	0.51	0.8376
ViT-L/16	Test	0.6057	0.50	0.61	0.54	0.8401

4.4. GRAD-CAM GÖRSELLEŞTİRME SONUÇLARI

DenseNet-169 modelinin Grad-CAM analizleri, modelin klinik olarak anlamlı anatomik bölgelere odaklandığını göstermiştir.

KL Evre 0 (Normal) vakalarda:

- Eklem aralığının geniş olduğu bölgeye odaklanma
- Kemik kenarlarında osteofit olmadığını doğrulamak için eklem kenarlarını inceleme
- Subkondral kemik yoğunluğunun normal olduğu bölgelere dikkat gösterme

KL Evre 1 (Şüpheli) vakalarda:

- Eklem aralığının hafif daraldığı bölgelere odaklanma
- Tibial plato ve femoral kondile yüksek aktivasyon

KL Evre 2 (Hafif OA) vakalarda:

- Tibial plato medial kenarında osteofit formasyonuna yüksek aktivasyon
- Eklem aralığında hafif daralma gösteren bölgelere odaklanma
- Femoral kondil kenarlarındaki küçük osteofitlere dikkat

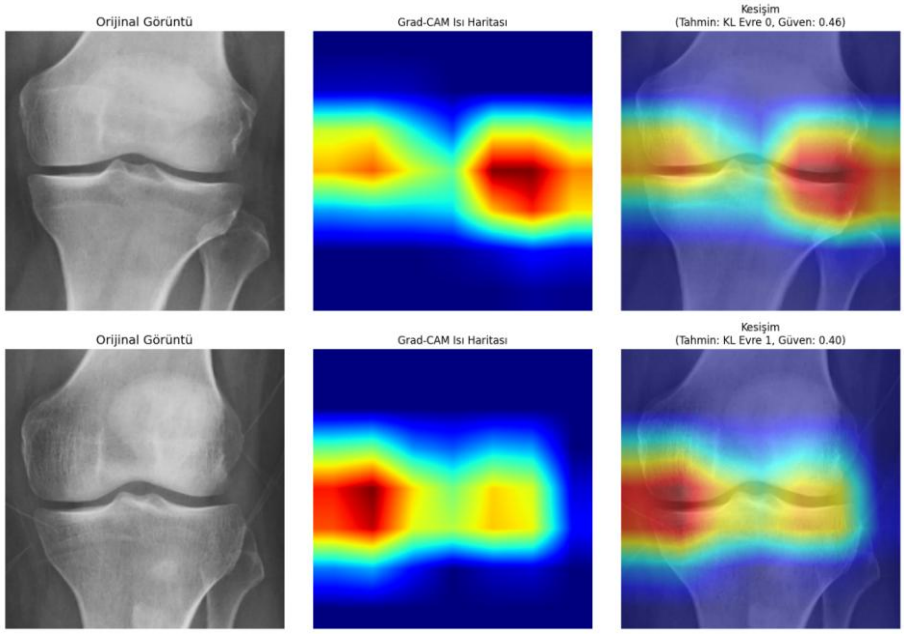
KL Evre 3 ve Evre 4 (İleri OA) vakalarda:

- Belirgin eklem aralığı daralması veya kaybolması bölgesine çok yüksek aktivasyon
- Osteofit bölgelerine güçlü odaklanma
- Subkondral skleroz alanlarına dikkat
- Kemik deformitesi bölgelerine aktivasyon

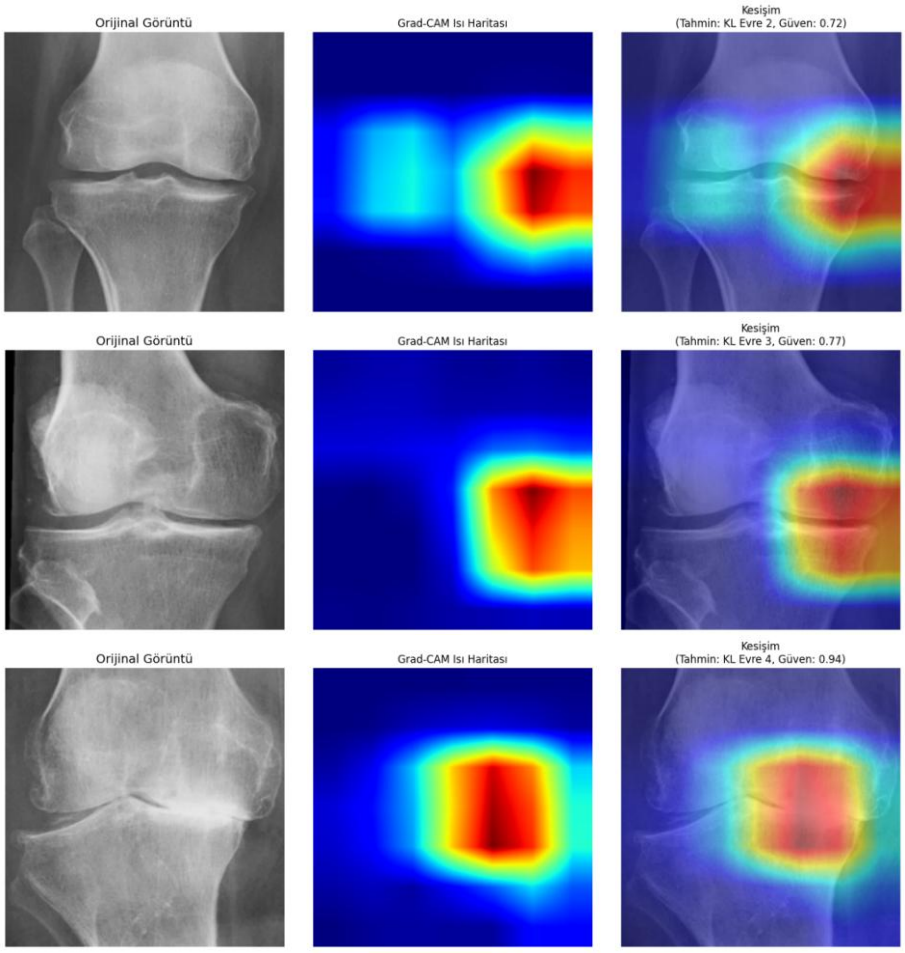
Klinik doğrulama: İki radyolog harici veri setinin Grad-CAM haritasını değerlendirmiştir:

- %78'inde model klinik olarak anlamlı bölgelere odaklanmış.
- %16'sında kısmen doğru bölgeler tespit edilmiş.
- %6'sında klinik olarak anlamsız bölgeler görülmüştür.

KL Evre 0-1 ve KL Evre 2-4 Grad-CAM ısı haritası örnekleri aşağıdaki şekillerde gösterilmiştir.

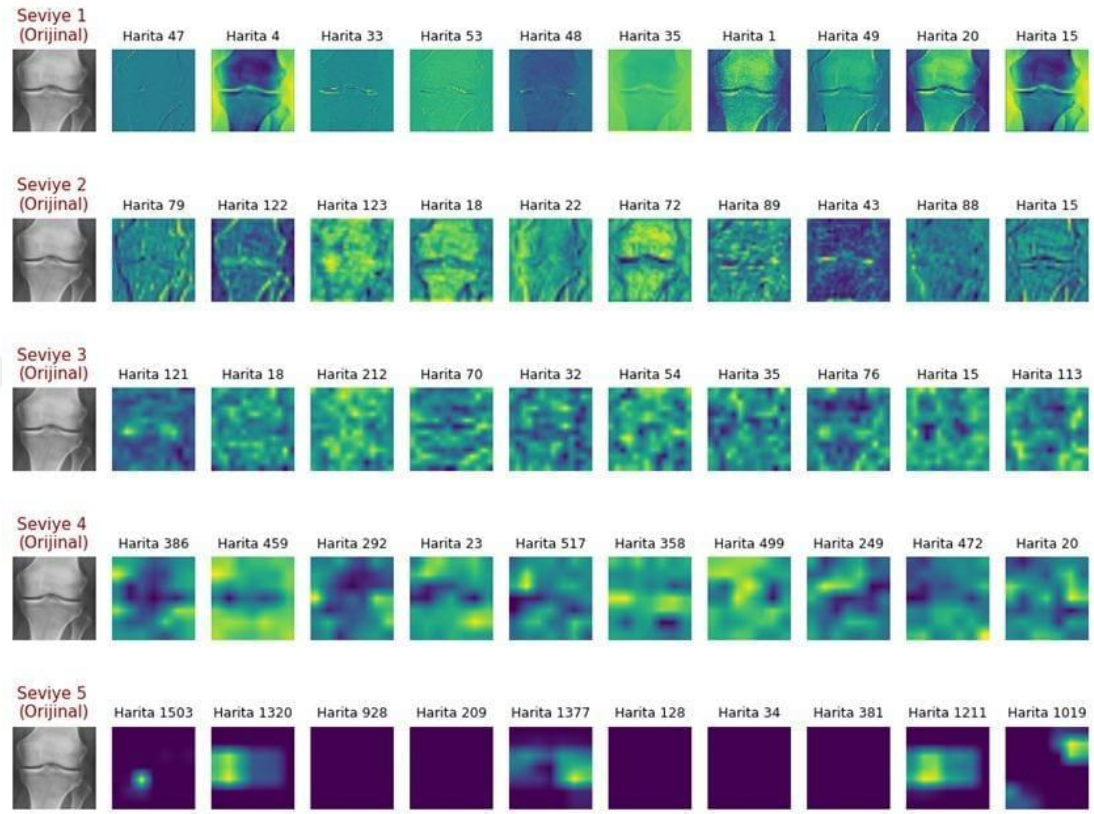


Şekil 4.4. KL Evre 0-1 Grad-CAM Isı Haritası Örnekleri



Şekil 4.5. KL Evre 2-4 Grad-CAM Isı Haritası Örnekleri

Modelin farklı katmanlarında Diz Grafisi görüntüsünün temsillerini görselleştirmek amacıyla, beş ayrı seviyeden özellik haritası örnekleri oluşturulmuştur. Oluşturulan bu özellik haritaları aşağıdaki şekilde gösterilmektedir.



Şekil 4.6. Özellik Haritası Örnekleri

4.5. RADYOLOG PERFORMANS DEĞERLENDİRMESİ VE GÖZLEMCİLER ARASI UYUM ANALİZİ

4.5.1. Yapay Zeka Desteği Öncesi ve Sonrası Performans

Radyologların YZ öncesi ve sonrası evre tahmin doğrulukları aşağıdaki tabloda gösterilmiştir.

Tablo 4.8. Radyolog Performansı Karşılaştırması

Radyolog	Deneyim	YZÖ Doğruluk (%)	YZS Doğruluk (%)	Mutlak Artış (%)	Relatif İyileşme (%)
R1	2-3 yıl	51,0	61,0	+10,0	19,6
R2	2-3 yıl	52,0	63,0	+11,0	21,2
R3	5-6 yıl	65,0	71,0	+6,0	9,2
R4	5-6 yıl	64,0	72,0	+8,0	12,5
R5	15+ yıl	77,0	81,0	+4,0	5,2
R6	15+ yıl	78,0	83,0	+5,0	6,4

Radyologların YZ öncesi ve sonrası evre tahminlerinin İşaret Testi ile değerlendirmesinin sonuçları aşağıdaki tabloda gösterilmiştir.

Tablo 4.9. İşaret Testi İstatistikleri

Metrik	Değer
Toplam Gözlem Sayısı	600
YZ Nedeniyle Tahmini İyileşen Vaka Sayısı (N+)	105
YZ Nedeniyle Tahmini Kötüleşen Vaka Sayısı (N−)	61
Toplam Net Doğruluk İyileşmesi (600 Vaka Üzerinden)	+%7,33
p-değeri (Tek Yönlü)	0,0004

Bu çalışmada, YZ desteğinin doğruluk üzerindeki etkisi, işaret testi kullanılarak değerlendirildi.

Toplam 166 vakanın YZ öncesi ve sonrası tahmininin değiştiği durumda, YZ sonrasında tahminler 105 vakada iyileşme yönünde (N+), 61 vakada ise kötüleşme yönünde (N-) gerçekleşmiştir.

Elde edilen p değeri ($p = 0,0004$), hem $\alpha = 0,05$ hem de $\alpha = 0,001$ eşik değerlerinden çok daha küçüktür. Tüm radyologlar genelinde yapay zeka desteği sayesinde gerçekleşen toplam doğruluk artışı istatistiksel olarak anlamlıdır ($p < 0,001$).

Wilcoxon İşaretli Sıra Testi Sonuçları

R1'den R6'ya kadar olan değerlendiricilerin ortalama evre tahmin doğruluğu hesaplanmıştır.

R1-R6 Ortalama Doğruluk:

- Önce (YZ Öncesi): 0.6450 (≈ 64.50%)
- Sonra (YZ Sonrası): 0.7183 (≈ 71.83%)

Wilcoxon İşaretli Sıra Testi Sonuçları:

- Test istatistiği (W): 0
- p-değeri: 0.0156

YZ'nın ortalama %7.33'lük doğruluk arttırıcı etkisi istatistiksel olarak anlamlıdır ($p<0.05$).

Radyologların YZ öncesi ve sonrası raporlama süreleri ve güven skoru değişimleri ve istatistiksel analizi aşağıdaki tablolarda gösterilmiştir.

Tablo 4.10. Raporlama Süresi ve Güven Skoru Değişimleri

Radyolog	YZÖ Süre (sn)	YZS Süre (sn)	YZÖ Güven	YZS Güven	Süre Azalması (%)	Güven Artışı (%)
R1	54,27	41,17	3,21	4,18	24,14	30,22
R2	56,96	45,95	3,90	4,40	19,33	12,82
R3	42,06	37,15	3,66	4,28	11,67	16,94
R4	46,34	39,62	4,12	4,70	14,51	14,08
R5	35,78	31,19	4,43	4,66	12,83	5,19
R6	36,40	33,07	4,56	4,74	9,15	3,95
Ortalama	45,30	38,03	3,98	4,49	16,07	12,81

Tablo 4.11. Wilcoxon İşaretli Sıra Testi Sonuçları

Ölçüt	Medyan Farkı	Tek Yönlü p-değeri	Sonuç (α=0,05)
Süre (Azalma)	5,82 sn	0,0156	İstatistiksel olarak anlamlı
Güven (Artış)	0,54	0,0156	İstatistiksel olarak anlamlı

4.5.2. Gözlemci İçi Tutarlılık Analizi

Aynı radyologun iki ayrı zamandaki değerlendirmesi arasındaki tutarlılık İDAK ile ölçülmüştür (74).

Tablo 4.12. Gözlemci İçi Tutarlılık

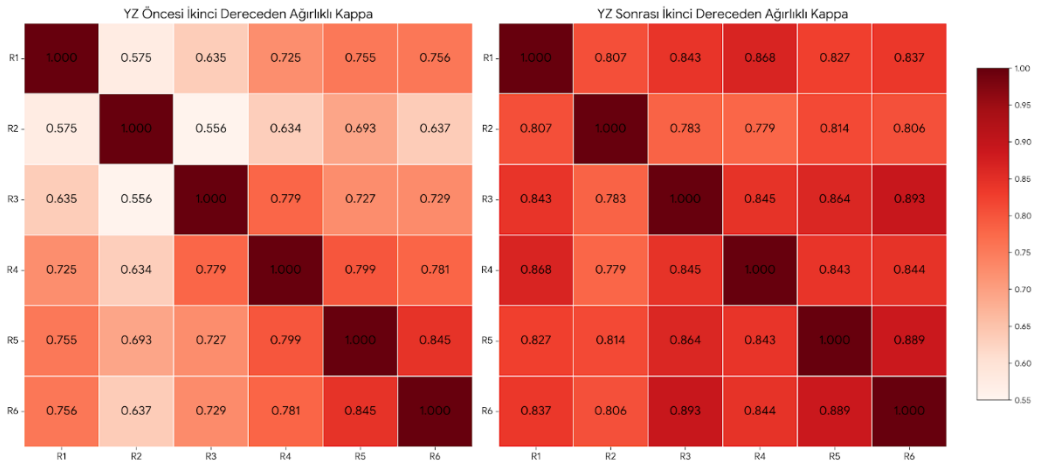
Radyolog	Deneyim	Gözlemci İçi İDAK	Yorum
R1	2-3 yıl	0,701	İyi uyum
R2	2-3 yıl	0,706	İyi uyum
R3	5-6 yıl	0,828	Çok iyi uyum
R4	5-6 yıl	0,795	İyi uyum
R5	15+ yıl	0,844	Çok iyi uyum
R6	15+ yıl	0,873	Çok iyi uyum
Ortalama	-	0,791	İyi uyum

Ağırlıklı Kappa Yorumlama Skalası (131):

- 0,00-0,20: Çok zayıf uyum
- 0,21-0,40: Zayıf uyum
- 0,41-0,60: Orta uyum
- 0,61-0,80: İyi uyum
- 0,81-1,00: Çok iyi uyum

4.5.3. Gözlemciler Arası Tutarlılık Analizi

YZ öncesi ve sonrası gözlemciler arası kappa değerleri aşağıdaki şekilde gösterilmiştir.



Şekil 4.7. Yapay Zeka Öncesi ve Sonrası Gözlemciler Arası Kappa Değerleri

YZ sonrası İDAK değişimi aşağıdaki tabloda gösterilmiştir.

Tablo 4.13. Eşli İDAK Wilcoxon Testi

Metrik	YZÖ Ortalama İDAK	YZS Ortalama İDAK	Mutlak Δ	Relatif Δ	Wilcoxon T / Fark	P değeri
Eşli İDAK (n=15)	0,708	0,836	+0,128	+18,1%	Medyan Fark> 0 (T=0)	0,0001

R1-R6 radyologlarında oluşturulan tüm 15 çift için, yapay zeka öncesi-sonrası İDAK artışı Wilcoxon t-testiyle anlamlı bulunmuştur ($p<0,001$).

4.5.4. Krippendorff's Alpha Sonuçları

Tüm radyologlar için genel uyum metriği olarak Krippendorff's Alpha hesaplanmıştır.

Tablo 4.14. Radyologlar Arası Uyumun Krippendorff's Alpha Sonuçları

Koşul	Krippendorff's Alpha	Mutlak İyileşme	Relatif İyileşme (%)	Yorumlama
YZÖ R1-R6	0,726	-	-	Orta düzeyde uyum
YZS R1-R6	0,845	+0,119	+16,4	İyi uyum

Krippendorff's Alpha Yorumlama (132):

- Alfa = 1: Kusursuz uzlaşı
- Alfa ≥ 0,80: Yeterli uzlaşı (güvenilir puanlama)
- Alfa [0,67 – 0,79]: Orta düzeyde uzlaşı (geçici çıkarımlar için minimum eşik)
- Alfa < 0,67: Zayıf uzlaşı (güvenilmez veri)
- Alfa = 0: Rastgele süreçle aynı
- Alfa < 0: Sistematik anlaşmazlık

R1-R6 (radyologlar arası): uyum 0,726'dan → 0,845'e yükselerek %16,4 artarak yeterli uzlaşı düzeyine gelmiştir.

4.5.5. Radyologların ve Yapay Zekanın Gerçek Değer ile Uyumu

YZ modelinin ve radyologların YZ öncesi ve sonrası tahminlerinin gerçek değerle uyumuna ilişkin İDAK değerleri aşağıdaki şekilde sunulmuştur.



Şekil 4.8. Değerlendiriciler ve YZ için Gerçek Değerle İDAK Skorları

Önemli bulgular:

- YZ modelinin gerçek değer ile uyumu çok iyi seviyededir (ikinci dereceden ağırlıklı kappa: 0,864)
- YZ desteği sonrası radyologların gerçek değer ile uyumu önemli ölçüde artmıştır (ortalama 0,827 → 0,884, +0,057 mutlak, +%6,9 relatif) ($p<0,05$)
- Az deneyimli radyolog R2 YZ desteği sonrası gerçek değer ile uyumları en büyük artışı göstermiştir (R2: +%20,4)
- Tüm radyologlar YZ sonrası 0,80'in üzerinde ağırlıklı kappa değerine ulaşmıştır, bu "çok iyi uyum" seviyesidir.

4.5.6. Özet Bulgular

- Model genellikle komşu evrelere hata yapmaktadır (örn: Evre1→Evre 0, Evre 2→Evre 3), bu klinik olarak kabul edilebilir bir hata tipidir
- En fazla doğruluk artışı az deneyimli radyologlarda görülmüştür (+10,5 puan mutlak ortalama, %20,4 ortalama relatif artış)
- Deneyimli olanlar da dahil olmak üzere tüm radyologlar YZ modelinden faydalanmıştır.
- Güven skorlarında da belirgin artış gözlenmiştir (ortalama +0,51 puan, %12,81 ortalama artış)
- Raporlama süreleri tüm gruplarda azalmıştır (ortalama-%16,07, en fazla az deneyimli R1'de-%24,14)
- Gözlemciler arası uyum "iyi uyum" seviyesinden "çok iyi uyum" seviyesine yükselmiştir (Ortalama eşli ikinci dereceden ağırlıklı kappa: 0,708 → 0,836, +%18,1, $p<0,001$)
- YZ sonrası radyologların uyumu Krippendorff's Alpha 0,726'dan 0,845'e yükselmiştir (+%16,4)
- En büyük uyum artışı az-orta deneyimli radyologlar arasında görülmüştür (R1-R2: +%40,3, R2-R3: +%40,8)
- YZ'nin gerçek değer ile mükemmel uyumu (ikinci dereceden ağırlıklı kappa: 0,864) sistemin güvenilirliğini doğrulamaktadır
- Radyologların YZ sonrası gerçek değer ile uyumları belirgin artmıştır (ortalama ikinci dereceden ağırlıklı kappa: 0,827 → 0,884, +%6,9)
- YZ sonrası tüm radyologlar referans değerlerle 0,80'in üzerinde ağırlıklı kappa'ya ulaşmış, "çok iyi uyum" seviyesine gelmiştir.

5. TARTIŞMA

Radyoloji, ortaya çıkışından bu yana modern tıbbı derinden etkileyen devrim niteliğinde bir gelişim süreci geçirmiştir. X-ışınlarının keşfinden başlayarak yapay zeka ve makine öğrenmesi teknolojilerinin entegrasyonuna kadar uzanan bu çok yönlü disiplin, kendisini ve temelini oluşturduğu sağlık ekosistemini sürekli olarak dönüştürmektedir (133).

Tıbbi görüntü analizi, genellikle yüksek boyutlu ve gürültülü veri kümelerinden ince desenleri ayırt edebilen güçlü algoritmalar gerektirir. Direkt grafi, BT, MRG ve pozitron emisyon tomografisi gibi görüntüleme yöntemlerinden karmaşık özellikleri doğrudan öğrenebilme yetenekleri sayesinde, ESA'lar otomatik tanı, tümör tespiti ve kişiselleştirilmiş tıp alanlarında önemli ilerlemeler sağlamıştır. Günümüzde ESA'lar, radyologların tanısal doğruluğu artırmak ve yorumlama süresini kısaltmak amacıyla klinik iş akışlarına giderek daha fazla entegre edilmektedir (134).

OA tanısı için literatürdeki çalışmalarda çeşitli ESA mimarileri kullanılmıştır. Nasef ve arkadaşlarının yürüttükleri çalışmada, Kaggle ve Mendeley veri setlerinden aldıkları görüntülerle eğittikleri KL evrelemesinin otomatik olarak belirlenmesi amacıyla geliştirilen modellerde, Derin Öğrenme modeli lojistik regresyon, destek vektör makineleri, Naive Bayes, karar ağacı, rastgele orman gibi farklı makine öğretme algoritmalarla karşılaştırılmış, derin öğrenme modeli, %70 doğruluk elde etmiş olup, diğer tüm yöntemlerden üstün performans sergilemiştir (135).

Tiulpin ve arkadaşları 2018 yılında Multicenter Osteoarthritis Study'den toplanan 18,376 diz grafisiyle görüntüsünde derin Siamese ESA eğiterek KL evrelemesi yapmışlardır. Model Osteoarthritis Initiative veri setinde 0,83 ikinci dereceden ağırlıklı kappa ve %66.71 çok sınıflı doğruluk elde etmiştir (136). Swiecicki ve arkadaşları 2021 yılında Multicenter Osteoarthritis Study veri setindeki 2802 hasta görüntüsü ile yaptıkları çalışmada çok girişli ESA ile KL evrelemesinde 0,9066 ikinci dereceden ağırlıklı kappa ve %71,9 ortalama çok sınıflı doğruluk elde etmiş, kas iskelet radyologlarına benzer performans göstermiştir (137).

Bizim çalışmamızda, DenseNet-169 ile Mendeley test setinde %70,17 doğruluk ve harici sette %65 doğruluk elde edilmiştir. Literatürdeki çalışmalarla karşılaştırıldığında, bu sonuçlar benzer performans seviyelerindedir.

Bu çalışmada 14 farklı mimari sistematik olarak karşılaştırılmıştır. Çalışmamızda, ESA tabanlı modeller (n=10) ViT modellerine (n=4) göre belirgin şekilde üstün performans göstermiştir. ESA modelleri ortalama %67,75 doğruluk elde ederken, Vision Transformer modelleri ortalama %60,30 doğruluk ile 7,45 puan doğruluk farkı göstermiştir.

ViT'nin orijinal çalışması, modelin etkili olması için çok büyük ölçekli veri setleri gerektirdiğini vurgulamaktadır. Dosovitskiy ve arkadaşlarına göre "ViT'ler yetersiz veride genelleme problemi yaşamaktadır" (70). Bu çalışmada kullanılan 8.260 görüntülük veri seti, ViT'nin optimal performansı için gereken yaklaşık 100M+ görüntünün sadece %0,0083'üne karşılık gelmektedir.

DenseNet-169 mimarisi, Mendeley veri setinde %70,17 test doğruluğu ve 0,9079 Macro AUC ile en yüksek performansı göstermiştir. Harici hastane setinde ise %65 genel doğruluk ve 0,8575 Macro AUC değerleri elde edilmiştir.

Harici hastane veri setimiz ile açık erişim veri seti arasındaki %5,17'lik farkın iki ayrı veri setinin cihaz farklılıkları, hasta demografisindeki değişim ve farklı etiketleyicilerin bulunması ile açıklanabilir.

KL evreleme sistemi, OA radyolojik değerlendirmesinde en yaygın kullanılan sınıflamadır (138). Ancak sistem, yorumcu bağımlılığı ve özellikle Evre 1'de düşük gözlemciler arası uyum gibi bazı limitasyonlara sahiptir.

Çalışmamızda KL evrelemesinde test setinde ve harici hastane veri setinde en yüksek performans normal ve şiddetli OA'da saptanmıştır. En düşük performans Evre 1 OA'da saptandı. Bu şüpheli osteofit gibi subjektif bulgulara normal ve hafif OA arasındaki ayrımın kişisel farklılıklarıyla açıklanabilir. Hastane veri setinde en düşük hasta sayısı Evre 1'de idi.

Şiddetli OA'da yüksek duyarlılık (%82-84) elde edilmesi, modelin belirgin osteoartrit bulgularını başarıyla tespit edebildiğini göstermektedir. Bu bulgular, sistemin özellikle ileri evre hastalıkların tespitinde güvenilir olduğunu desteklemektedir.

Grad-CAM analizi, modelin %78 oranında klinik olarak anlamlı anatomik bölgelere odaklandığını göstermiştir. Bu bulgu, yapay zeka sisteminin kara kutu algısını azaltması ve klinik güvenilirliğinin artırılması açısından kritik önem taşımaktadır.

Bu çalışmanın en önemli klinik bulgusu, yapay zeka desteğinin radyolog performansını istatistiksel olarak anlamlı şekilde artırmasıdır.

Yapay zeka modelinin gerçek değer ile ikinci dereceden ağırlıklı kappa uyumu 0,864 olarak bulunmuştur. Bu değer, modelin referans standart ile mükemmel düzeyde uyumlu olduğunu göstermektedir. Ayrıca, yapay zeka desteği sonrası radyologların gerçek değer ile uyumları belirgin şekilde artmıştır (ortalama ikinci dereceden ağırlıklı kappa: 0,827 → 0,884, +%6,9).

Vaattovaara ve arkadaşları gözlemciler arası uyum katsayısını radyoloji asistanları için 0,737, kas-iskelet sistemi radyolojisi fellow'u için 0,761, deneyimli kas-iskelet sistemi radyologu için 0,802 ve ortopedistler için 0,818 olarak bulmuştur (138).

Brejnebol ve arkadaşları, yapay zeka desteğiyle KL evreleme performansında anlamlı bir artış yakaladığını göstermiştir. Tüm gözlemciler arasında KL evrelemesinde gözlemciler arası uyum, yapay zeka desteğiyle anlamlı biçimde artmıştır (κ = 0,77'den 0,85'e). Yapay zeka desteği alan board sertifikalı radyologlar, KL evrelemesinde neredeyse mükemmel uyum (κ = 0,90) elde etmiştir (139). Bulgularımız literatürle uyumludur.

Gözlemciler arası uyumdaki %18,1'lik iyileşme (ortalama eşli ağırlıklı kappa: 0,708 → 0,836, $p<0,001$) ve Krippendorff's Alpha'daki %16,4'lük artış (0,726 →

0,845), yapay zeka destekli sistemlerin sadece bireysel performansı değil, aynı zamanda tanı standardizasyonunu da artırdığını göstermektedir.

Bunun klinik yararları;

- Farklı radyologların aynı hastayı benzer şekilde değerlendirmesi, hastalık takibi ve tedavi planlamasında tutarlılık sağlar.
- Yüksek gözlemciler arası uyum, çok merkezli klinik çalışmalarda veri etiketi kalitesini artırır.
- Yapay zeka desteği, az deneyimli radyologların performansını deneyimli radyologlara yaklaştırarak radyoloji eğitim sürecini hızlandırır.
- Yüksek standardizasyon, tanı uyuşmazlıklarını ve medikolegal riskleri azaltır.

Çalışmamızın diğer bir çarpıcı bulgusu yapay zekanın tüm radyologlara deneyim farkı gözetmeksizin fayda sağlaması ve radyologlar arası deneyim farkını azaltmasıdır.

Yapay zeka modelinin gerçek değer ile mükemmel uyumu (ağırlıklı kappa: 0,864) ve radyologların yapay zeka sonrası gerçek değer ile artan uyumları (0,827 → 0,884), sistemin güvenilir bir referans noktası olarak kullanılabileceğini göstermektedir.

Çalışmamızın limitasyonları mevcuttur. Sadece ön-arka grafiler kullanılmış olup, lateral grafiler dahil edilmemiştir. Çalışmamız tek merkez çalışması olup 100 hasta ile sınırlıdır.

6. SONUÇLAR

Sonuç olarak bu çalışma, derin öğrenme tabanlı OA tanı sistemlerinin klinik kullanıma hazır hale geldiğini göstermektedir. DenseNet-169 mimarisi, yoğun bağlantı yapısının tıbbi görüntüleme için optimal olduğunu kanıtlamış, özellikle doğruluk ve AUC metriklerinde üstün performans sergilemiştir.

Çalışmanın en önemli bulgusu, yapay zeka desteğinin sadece bireysel tanı doğruluğunu değil, aynı zamanda gözlemciler arası uyumu da önemli ölçüde artırmasıdır. Yapay zekanın gerçek değer ile mükemmel uyumu ve radyologların yapay zeka sonrası gerçek değer ile artan uyumları, sistemin güvenilir bir klinik karar destek aracı olarak kullanılabileceğini desteklemektedir.

Gelecekte, daha büyük çok merkezli veri setleri, çok modaliteli yaklaşımlar ve prospektif doğrulama çalışmaları ile bu sistemlerin klinik rutinde yaygınlaşması beklenmektedir. OA tanısında yapay zeka destekli sistemler, radyolog iş yükünü azaltırken, tanı doğruluğunu ve standardizasyonu artırma potansiyeline sahiptir.

Son söz olarak Dr. Curtis Langlotz'un dediği gibi "Yapay zeka, radyologların yerini almayacak ancak yapay zeka kullanan radyologlar, kullanmayanların yerini alacaktır" (140). Bu çalışma, bu vizyonun gerçekleştirilmesi yönünde önemli bir adımdır ve diz OA tanısında yapay zeka destekli klinik karar destek sistemlerinin pratik uygulanabilirliğini kanıtlamaktadır.

7. KAYNAKLAR

1. Di Nicola V. Degenerative osteoarthritis a reversible chronic disease. Regen Ther. 2020;15:149-160.
2. GBD 2021 Osteoarthritis Collaborators. Global, regional, and national burden of osteoarthritis, 1990-2020 and projections to 2050: a systematic analysis for the Global Burden of Disease Study 2021. Lancet Rheumatol. 2023;5(9):e508-e522.
3. World Health Organization. Osteoarthritis Fact Sheet [Internet]. 2023 Jul 14 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://www.who.int/news-room/fact-sheets/detail/osteoarthritis
4. Hunter DJ, Bierma-Zeinstra S. Osteoarthritis. Lancet. 2019;393(10182):1745-1759.
5. Kotlarz H, Gunnarsson CL, Fang H, Rizzo JA. Insurer and out-of-pocket costs of osteoarthritis in the US: Evidence from national survey data. Arthritis Rheum. 2009;60:3546-3553.
6. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis. 1957;16(4):494-502.
7. Kohn MD, Sassoon AA, Fernando ND. Classifications in Brief: Kellgren-Lawrence Classification of Osteoarthritis. Clin Orthop Relat Res. 2016;474(8):1886-1893.
8. Xie F, Kovic B, Jin X, et al. Economic and Humanistic Burden of Osteoarthritis: A Systematic Review of Large Sample Studies. PharmacoEconomics. 2016;34:1087–1100.
9. Glyn-Jones S, Palmer AJR, Agricola R, et al. Osteoarthritis. Lancet. 2015;386(9991):376-387.
10. Anderson AS, Loeser RF. Why is osteoarthritis an age-related disease? Best Pract Res Clin Rheumatol. 2010;24(1):15-26.
11. Arden N, Nevitt MC. Osteoarthritis: epidemiology. Best Pract Res Clin Rheumatol. 2006;20(1):3-25.
12. Srikanth VK, Fryer JL, Zhai G, Winzenberg TM, Hosmer D, Jones G. A meta-analysis of sex differences prevalence, incidence and severity of osteoarthritis. Osteoarthritis Cartilage. 2005;13(9):769-781.
13. Hsu H, Siwiec RM. Knee Osteoarthritis. In: StatPearls. Treasure Island (FL): StatPearls Publishing; 2023.
14. Abhishek A, Doherty M. Pathophysiology of articular chondrocalcinosis--role of ANKH. Nat Rev Rheumatol. 2011;7(2):96-104.
15. Subramanian P, Soundararajan R, Spalkit S, Sinha A, Verma N. A comprehensive review of the common developmental disorders of hip – Developmental dysplasia of the hip, slipped capital femoral epiphysis, and Perthes disease. Indian J Musculoskelet Radiol. 2025;7:13-25.

16. Loeser RF. Age-related changes in the musculoskeletal system and the development of osteoarthritis. Clin Geriatr Med. 2010;26(3):371-386.

17. Tang X, Wang S, Zhan S, et al. The Prevalence of Symptomatic Knee Osteoarthritis in China: Results From the China Health and Retirement Longitudinal Study. Arthritis Rheumatol. 2016;68(3):648-653.

18. Loeser RF, Collins JA, Diekman BO. Ageing and the pathogenesis of osteoarthritis. Nat Rev Rheumatol. 2016;12(7):412-420.

19. Sowers MR, McConnell D, Jannausch M, et al. Estradiol and its metabolites and their association with knee osteoarthritis. Arthritis Rheum. 2006;54(8):2481-2487.

20. Spector TD, MacGregor AJ. Risk factors for osteoarthritis: genetics. Osteoarthritis Cartilage. 2004;12 Suppl A:S39-44.

21. Boer CG, Hatzikotoulas K, Southam L, et al. Deciphering osteoarthritis genetics across 826,690 individuals from 9 populations. Cell. 2021;184(18):4784-4818.e17.

22. Aury-Landas J, Marcelli C, Leclercq S, Boumediene K, Bauge C. Genetic Determinism of Primary Early-Onset Osteoarthritis. Trends Mol Med. 2016;22(1):38-52.

23. Zheng H, Chen C. Body mass index and risk of knee osteoarthritis: systematic review and meta-analysis of prospective studies. BMJ Open. 2015;5(12):e007568.

24. Gao XR, Chen YS, Deng W. The effect of vitamin D supplementation on knee osteoarthritis: A meta-analysis of randomized controlled trials. Int J Surg. 2017;46:14-20.

25. Neogi T, Booth SL, Zhang YQ, et al. Low vitamin K status is associated with osteoarthritis in the hand and knee. Arthritis Rheum. 2006;54(4):1255-1261.

26. McAlindon TE, Jacques P, Zhang Y, et al. Do antioxidant micronutrients protect against the development and progression of knee osteoarthritis? Arthritis Rheum. 1996;39(4):648-656.

27. Cordingley DM, Cornish SM. Omega-3 Fatty Acids for the Management of Osteoarthritis: A Narrative Review. Nutrients. 2022;14(16):3362.

28. Verbeek J, Mischke C, Robinson R, et al. Occupational Exposure to Knee Loading and the Risk of Osteoarthritis of the Knee: A Systematic Review and a Dose-Response Meta-Analysis. Saf Health Work. 2017;8(2):130-142.

29. Wang TJ, Lee SC, Liang SY, et al. Comparing the efficacy of aquatic exercises and land-based exercises for patients with knee osteoarthritis. J Clin Nurs. 2011;20(17-18):2609-2622.

30. Tong L, Yu H, Huang X, et al. Current understanding of osteoarthritis pathogenesis and relevant new approaches. Bone Res. 2022;10:60.

31. Mueller MB, Tuan RS. Anabolic/Catabolic Balance in Pathogenesis of Osteoarthritis: Identifying Molecular Targets. PM&R. 2011;3(6):S3-S11.

32. Kumar S, Kumar H, Mittal A, Singh PP, Yadav V, Kumar D, et al. Correlation Between Synovial Fluid Levels of Matrix Metalloproteinase's (MMP-1, MMP-3, and MMP-9) and TNF-α with the Severity of Osteoarthritis Knee in Rural Indian Population. Indian J Orthop. 2023;57(10):1659-1666.

33. Li T, Peng J, Li Q, et al. The Mechanism and Role of ADAMTS Protein Family in Osteoarthritis. Biomolecules. 2022;12(7):959.

34. Burr DB, Gallant MA. Bone remodelling in osteoarthritis. Nat Rev Rheumatol. 2012;8(11):665-673.

35. van der Kraan PM, van den Berg WB. Osteophytes: relevance and biology. Osteoarthritis Cartilage. 2007;15(3):237-244.

36. Yaseen K. Osteoarthritis (OA). In MSD Manual Professional Edition. MSD Manuals; 2024. [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://www.msdmanuals.com/professional/musculoskeletal-and-connective-tissue-disorders/joint-disorders/osteoarthritis-oa

37. Osteoarthritis: care and management. London: National Institute for Health and Care Excellence (NICE); 2020 Dec 11. (NICE Clinical Guidelines, No. 177.)

38. Bannuru RR, Osani MC, Vaysbrot EE, et al. OARSI guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis. Osteoarthritis Cartilage. 2019;27(11):1578-1589.

39. Kitagawa T, Denda T, Okuyama W, Miyachi R, Nakamura K. Effectiveness of Rehabilitation for Osteoarthritis of the Knee: A Scoping Review of Network Meta-Analyses. Cureus. 2024;16(4):e57661.

40. Jormand H, Mohammadi N, Khani Jeihooni A, Afzali Harsini P. Self-care behaviors in older adults suffering from knee osteoarthritis: Application of theory of planned behavior. Front Public Health. 2022;10:958614.

41. Magni A, Agostoni P, Bonezzi C, et al. Management of Osteoarthritis: Expert Opinion on NSAIDs. Pain Ther. 2021;10(2):783-808.

42. da Costa BR, Pereira TV, Saadat P, et al. Effectiveness and safety of non-steroidal anti-inflammatory drugs and opioid treatment for knee and hip osteoarthritis: network meta-analysis. BMJ. 2021;375:n2321.

43. Jones IA, Togashi R, Wilson ML, Heckmann N, Vangsness CT Jr. Intra-articular treatment options for knee osteoarthritis. Nat Rev Rheumatol. 2019;15(2):77-90.

44. Mimpen JY, Snelling SJB. Chondroprotective Factors in Osteoarthritis a Joint Affair. Curr Rheumatol Rep. 2019;21(8):41.

45. Law GW, Lee JK, Soong J, Lim JWS, Zhang KT, Tan AHC. Arthroscopic debridement of the degenerative knee - Is there still a role? Asia Pac J Sports Med Arthrosc Rehabil Technol. 2018;15:23-28.

46. Peng H, Ou A, Huang X, et al. Osteotomy Around the Knee: The Surgical Treatment of Osteoarthritis. Orthop Surg. 2021;13(5):1465-1473.

47. Liddle AD, Pegg EC, Pandit H. Knee replacement for osteoarthritis. Maturitas. 2013;75(2):131-136.

48. Braun HJ, Gold GE. Diagnosis of osteoarthritis: imaging. Bone. 2012;51(2):278-288.

49. Roemer FW, Crema MD, Trattnig S, Guermazi A. Advances in imaging of osteoarthritis and cartilage. Radiology. 2011;260(2):332-354.

50. Omoumi P. Imaging of cartilage, meniscus, and beyond: Role of Magnetic Resonance Imaging (MRI) and Computed Tomography (CT). Osteoarthritis Imaging. 2025;5(2):100268.

51. Roemer FW, Engelke K, Li L, Laredo J-D, Guermazi A. MRI underestimates presence and size of knee osteophytes using CT as a reference standard. Osteoarthritis and Cartilage. 2023;31(5):656-668.

52. Segal NA, Nevitt MC, Lynch JA, et al. Diagnostic performance of 3D standing CT imaging for detection of knee osteoarthritis features. Phys Sportsmed. 2015;43(3):213-220.

53. Kumahashi N, Kuwata S, Takuwa H, Uchio Y. Ultrasonographic evaluation of early knee osteoarthritis. Annals of Joint. 2018;3:4583.

54. Möller I, Bong D, Naredo E, et al. Ultrasound in the study and monitoring of osteoarthritis. Osteoarthritis Cartilage. 2008;16 Suppl 3:S4-7.

55. Pi SW, Lee BD, Lee MS, et al. Ensemble deep-learning networks for automated osteoarthritis grading in knee X-ray images. Sci Rep. 2023;13:22887.

56. Stryker C, Kavlakoglu E. IBM. What is artificial intelligence (AI)? [Internet]. [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://www.ibm.com/topics/artificial-intelligence

57. Bergmann D. IBM. What is machine learning ? [Internet]. [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://www.ibm.com/think/topics/machine-learning

58. IBM. What is deep learning ? [Internet]. [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://www.ibm.com/think/topics/deep-learning

59. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. Med Image Anal. 2017;42:60-88.

60. Wikipedia. Neural Network [Internet]. [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://en.wikipedia.org/wiki/Neural_network_(machine_learning)

61. Borse D. Training of ANN in Data Mining. GeeksforGeeks [Internet]. [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://www.geeksforgeeks.org/data-science/training-of-ann-in-data-mining/

62. Ultralytics. Convolution [Internet]. [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://www.ultralytics.com/glossary/convolution

63. Mishra A. GeeksforGeeks. CNN Intro [Internet]. [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://www.geeksforgeeks.org/introduction-convolution-neural-network/

64. Gurucharan MK. UpGrad. Basic CNN Architecture [Internet]. [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://www.upgrad.com/blog/basic-cnn-architecture/

65. Keita Z. DataCamp. An Introduction to Convolutional Neural Networks (CNNs) [Internet]. [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://www.datacamp.com/tutorial/introduction-to-convolutional-neural-networks-cnns

66. Lee JG, Jun S, Cho YW, Lee H, Kim GB, Seo JB, et al. Deep Learning in Medical Imaging: General Overview. Korean J Radiol. 2017;18(4):570-584.

67. Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. Insights into Imaging. 2018;9(4):611-629.

68. Chollet F. Keras. Transfer Learning Guide [Internet]. 2024 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://keras.io/guides/transfer_learning/

69. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely Connected Convolutional Networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE; 2017:2261-2269.

70. Dosovitskiy A, Beyer L, Kolesnikov A, et al. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. arXiv preprint arXiv:2010.11929. 2020.

71. Amjad J, Sajid MZ, Amjad A, Hamid MF, Youssef A, Sharif MI. Knee Osteoarthritis Detection and Classification Using Autoencoders and Extreme Learning Machines. AI. 2025;6:151.

72. Abdusalomov A, Mirzakhalilov S, Umirzakova S, et al. Lightweight early detection of knee osteoarthritis in athletes. Sci Rep. 2025;15:31413.

73. Mohammed AS, Hasanaath AA, Latif G, Bashar A. Knee Osteoarthritis Detection and Severity Classification Using Residual Neural Networks on Preprocessed X-ray Images. Diagnostics. 2023;13:1380.

74. Ozturk A. Quadratic Weighted Kappa. Medium [Internet]. 2024 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://medium.com/@nlztrk/quadratic-weighted-kappa-%C4%B0DAK-metric-and-how-to-optimize-it-062cc9121baa

75. Krippendorff K. Content Analysis: An Introduction to Its Methodology. 3rd ed. Thousand Oaks, CA: Sage Publications; 2013.

76. Kaplan M. How to Use Krippendorff's Alpha to Measure Annotation Agreement. Label Studio Blog [Internet]. 2025 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://labelstud.io/blog/how-to-use-krippendorff-s-alpha-to-measure-annotation-agreement/

77. Classification: ROC and AUC. Google Developers [Internet]. 2024 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc

78. ROC AUC Score. Scikit-learn [Internet]. 2024 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.roc_auc_score.html

79. Azure Machine Learning Metrics. Microsoft Learn [Internet]. 2024 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://learn.microsoft.com/en-us/azure/machine-learning/how-to-understand-automated-ml

80. Wilcoxon F. Individual comparisons by ranking methods. Biometrics Bulletin. 1945;1(6):80–83.

81. Dixon, W. J., & Mood, A. M. (1946). The Statistical Sign Test. Journal of the American Statistical Association, 41(236), 557–566. https://doi.org/10.2307/2280577

82. Understanding the Wilcoxon Signed Rank Test. Statistics Solutions [Internet]. 2025 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/how-to-conduct-the-wilcox-sign-test/

83. Wikipedia. Sign Test [Internet]. 2025 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://en.wikipedia.org/wiki/Sign_test

84. Sounderajah, V., Guni, A., Liu, X. et al. The STARD-AI reporting guideline for diagnostic accuracy studies using artificial intelligence. Nat Med 31, 3283–3289 (2025). https://doi.org/10.1038/s41591-025-03953-8

85. Chen P. Knee Osteoarthritis Severity Grading Dataset. Mendeley Data. V1. 2018 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://data.mendeley.com/datasets/56rmx5bjcr/1

86. Chen P, Gao L, Shi X, Allen K, Yang L. Fully automatic knee osteoarthritis severity grading using deep neural networks with a novel ordinal loss. Comput Med Imaging Graph. 2019;75:84-92. doi:10.1016/j.compmedimag.2019.06.002

87. Mason D. pydicom: An open source DICOM library [Internet]. 2011 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://pydicom.github.io/

88. Khan SA. GeeksforGeeks. How to normalize images in PyTorch? [Internet]. [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://www.geeksforgeeks.org/python/how-to-normalize-images-in-pytorch/

89. PyTorch. RandomResizedCrop [Internet]. 2025 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://docs.pytorch.org/vision/stable/generated/torchvision.transforms.RandomResizedCrop.html

90. PyTorch. RandomHorizontalFlip [Internet]. 2025 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://docs.pytorch.org/vision/stable/generated/torchvision.transforms.RandomHorizontalFlip.html

91. PyTorch. RandomRotation [Internet]. 2025 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://docs.pytorch.org/vision/stable/generated/torchvision.transforms.RandomRotation.html

92. PyTorch. ColorJitter [Internet]. 2025 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://docs.pytorch.org/vision/stable/generated/torchvision.transforms.ColorJitter.html

93. Zhong Z, Zheng L, Kang G, Li S, Yang Y. Random Erasing Data Augmentation. arXiv preprint arXiv:1708.04896. 2017.

94. Shorten C, Khoshgoftaar TM. A survey on Image Data Augmentation for Deep Learning. J Big Data. 2019;6:60.

95. Zhang H, Cisse M, Dauphin YN, Lopez-Paz D. mixup: Beyond Empirical Risk Minimization. arXiv preprint arXiv:1710.09412. 2018.

96. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE; 2016:770-778.

97. PyTorch. ResNet18 [Internet]. 2024 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://docs.pytorch.org/vision/main/models/generated/torchvision.models.resnet18.html

98. PyTorch. ResNet50 [Internet]. 2024 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://docs.pytorch.org/vision/main/models/generated/torchvision.models.resnet50.html

99. PyTorch. ResNet101 [Internet]. 2024 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://docs.pytorch.org/vision/main/models/generated/torchvision.models.resnet101.html

100. PyTorch. DenseNet121 [Internet]. 2024 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://docs.pytorch.org/vision/main/models/generated/torchvision.models.densenet121.html

101. PyTorch. DenseNet169 [Internet]. 2024 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://docs.pytorch.org/vision/main/models/generated/torchvision.models.densenet169.html

102. PyTorch. DenseNet201 [Internet]. 2024 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://docs.pytorch.org/vision/main/models/generated/torchvision.models.densenet201.html

103. Tan M, Le Q. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. ICML. 2019;97:6105-6114.

104. PyTorch. EfficientNet-B0 [Internet]. 2024 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://docs.pytorch.org/vision/main/models/generated/torchvision.models.efficientnet_b0.html

105. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. MobileNetV2: Inverted Residuals and Linear Bottlenecks. CVPR. 2018:4510-4520.

106. PyTorch. MobileNetV2 [Internet]. 2024 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://docs.pytorch.org/vision/main/models/generated/torchvision.models.mobilenet_v2.html

107. Liu Z, Mao H, Wu CY, Feichtenhofer C, Darrell T, Xie S. A ConvNet for the 2020s. arXiv preprint arXiv:2201.03545. 2022.

108. PyTorch. ConvNeXt-Tiny [Internet]. 2024 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://docs.pytorch.org/vision/main/models/generated/torchvision.models.convnext_tiny.html

109. PyTorch. ConvNeXt-Small [Internet]. 2024 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://docs.pytorch.org/vision/main/models/generated/torchvision.models.convnext_small.html

110. Liu Z, Lin Y, Cao Y, Hu H, Wei Y, Zhang Z, et al. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. arXiv:2103.14030. 2021.

111. PyTorch. Swin-T [Internet]. 2024 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://docs.pytorch.org/vision/main/models/generated/torchvision.models.swin_t.html

112. PyTorch. Swin-S [Internet]. 2024 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://docs.pytorch.org/vision/main/models/generated/torchvision.models.swin_s.html

113. PyTorch. ViT-B/16 [Internet]. 2024 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://docs.pytorch.org/vision/main/models/generated/torchvision.models.vit_b_16.html

114. PyTorch. ViT-L/16 [Internet]. 2024 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://docs.pytorch.org/vision/main/models/generated/torchvision.models.vit_l_16.html

115. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. ImageNet: A large-scale hierarchical image database. CVPR. 2009:248-255.

116. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: A simple way to prevent neural networks from overfitting. J Mach Learn Res. 2014;15(1):1929-1958.

117. Prechelt L. Early stopping — but when? In: Neural Networks: Tricks of the Trade. Springer; 1998:55–69.

118. Karpathy A. t-SNE visualization of CNN codes. Stanford University [Internet]. [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://cs.stanford.edu/people/karpathy/cnnembed/

119. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. In: 2017 IEEE International Conference on Computer Vision (ICCV). IEEE; 2017:618-626.

120. Likert R. A Technique for the Measurement of Attitudes. Archives of Psychology. 1932;140:1–55.

121. Cohen J. A coefficient of agreement for nominal scales. Educational and Psychological Measurement. 1960;20(1):37–46.

122. Python Software Foundation. Python (Version 3.12) [Internet]. [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://www.python.org

123. Paszke A, Gross S, Massa F, et al. PyTorch: An imperative style, high-performance deep learning library. NeurIPS. 2019;32:8024-8035.

124. Marcel S, Rodriguez Y. Torchvision the machine-vision package of torch. ACM Multimedia. 2010:1485-1488.

125. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine learning in Python. JMLR. 2011;12:2825-2830.

126. Harris CR, Millman KJ, van der Walt SJ, et al. Array programming with NumPy. Nature. 2020;585:357–362.

127. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: Fundamental algorithms for scientific computing in Python. Nat Methods. 2020;17(3):261-272.

128. Hunter JD. Matplotlib: A 2D graphics environment. Comput Sci Eng. 2007;9(3):90-95.

129. Waskom ML. seaborn: Statistical data visualization. JOSS. 2021;6(60):3021.

130. Gildenblat J, contributors. PyTorch library for CAM methods. GitHub [Internet]. 2021 [Erişim tarihi: 5 Kasım 2025]. Erişim bağlantısı: https://github.com/jacobgil/pytorch-grad-cam

131. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159-174.

132. Marzi G, Balzano M, Marchiori D. K-Alpha Calculator —Krippendorff's Alpha Calculator: A User-Friendly Tool for Computing Krippendorff's Alpha Inter-Rater Reliability Coefficient. MethodsX. 2024;12:102545.

133. Najjar R. Redefining Radiology: A Review of Artificial Intelligence Integration in Medical Imaging. Diagnostics (Basel). 2023;13(17):2760.

134. Mienye ID, Swart TG, Obaido G, Jordan M, Ilono P. Deep Convolutional Neural Networks in Medical Image Analysis: A Review. Information. 2025;16:195.

135. Nasef D, Nasef D, Sawiris V, Girgis P, Toma M. Deep Learning for Automated Kellgren–Lawrence Grading in Knee Osteoarthritis Severity Assessment. Surgeries. 2025;6:3.

136. Tiulpin A, Thevenot J, Rahtu E, Lehenkari P, Saarakkala S. Automatic Knee Osteoarthritis Diagnosis from Plain Radiographs: A Deep Learning-Based Approach. Sci Rep. 2018;8(1):1727.

137. Swiecicki A, Li N, O'Donnell J, et al. Deep learning-based algorithm for assessment of knee osteoarthritis severity in radiographs matches performance of radiologists. Comput Biol Med. 2021;133:104334.

138. Vaattovaara E, Panfilov E, Tiulpin A, et al. KL grading of knee osteoarthritis using deep learning: Diagnostic performance with external dataset and comparison with four readers. Osteoarthr Cartil Open. 2025;7(2):100580.

139. Brejnebøl MW, Lenskjold A, Ziegeler K, et al. Interobserver Agreement and Performance of Concurrent AI Assistance for Radiographic Evaluation of Knee Osteoarthritis. Radiology. 2024;312:1.

140. Langlotz CP. Will Artificial Intelligence Replace Radiologists? Radiol Artif Intell. 2019;1(3):e190058.

8. ÖZGEÇMİŞ

I. Bireysel Bilgiler:

- Ad Soyad: Dr. Mehmet Numan Çolakoğlu
- Doğum yeri ve tarihi:
- Uyruğu: TC
- İletişim adresi ve telefonu:

- Yabancı dili: İngilizce

II. Eğitimi:

- Tıpta Uzmanlık Eğitimi, Radyoloji, Ankara Bilkent Şehir Hastanesi (2021-2026)
- Tıp Fakültesi, Hacettepe Üniversitesi (İngilizce) (2014-2020)
- Ankara Atatürk Anadolu Lisesi (2010-2014)

III. Ünvanları:

- Asistan Doktor (2021- Halen)

IV. Mesleki Deneyim:

- Ankara Bilkent Şehir Hastanesi Radyoloji Kliniği, Asistan Doktor (2021- Halen)
- Ayaş Şehit Mehmet Çifci Devlet Hastanesi, Pratisyen Doktor (2020-2021)

V. Üyelikler:

- Türk Radyoloji Derneği (TRD)
- Radiological Society of North America (RSNA)
- European Society of Radiology (ESR)
- American Roentgen Ray Society (ARRS)

VI. Yayınlar ve Sunumlar:

- Keles, E., Colakoglu, M. N., & Bengtsson, M. (2025). Comment on Dalboni da Rocha et al. Artificial Intelligence for Neuroimaging in Pediatric Cancer. Cancers (Basel), 17(11), 1776. https://doi.org/10.3390/cancers17111776
- Gunes, Y. C., Cesur, T., Camur, T., Cifci, B. E., Kaya, T., Colakoglu, M. N., et al. (2025). Textual Proficiency and Visual Deficiency: A Comparative Study of Large Language Models and Radiologists in MRI Artifact Detection and Correction. Acad Radiol.
- Büyükçapar, E., Görgülü, Ü., Güven, E. Z., Çopur, Ç., Giray, S., Ateş, S., Çolakoğlu, M. N., Günbey Karabekmez, L., & Bektaş, H. (2025). Predictors associated with hematoma volume expansion in spontaneous intracerebral hemorrhage. 20th Uludağ Neurology Days, Bursa, Türkiye.
- Colakoglu, M. N., & Demirci, R. A. (2024). Glioma, Meningioma and Pituitary Mass Diagnosis Using ResNet Architecture. Turkish National Congress of Radiology 2024.
- Temizhan, A. Y., Colakoglu, M. N., Kara, M., Köprücü, E., Korkmaz, A., Topaloğlu, S., Altay, F. P., Köroğlu, E. Y., Erbahçeci Timur, İ. E., Uğurlu, N., & Temizhan, A. (2024). Mulibrey Nanism: A Case with Heart Failure. Arch Turk Soc Cardiol.
- Colakoglu, M. N. (2023). Brain Tumor Diagnosis Using Convolutional Neural Network Architecture. Turkish National Congress of Radiology 2023.

9. EKLER

EK 1: ETİK KURUL ONAYI

T.C.
SAĞLIK BAKANLIĞI
İL SAĞLIK MÜDÜRLÜĞÜ
Ankara Bilkent Şehir Hastanesi

1 NOLU
TIBBİ ARAŞTIRMALAR BİLİMSEL VE ETİK DEĞERLENDİRME KURULU
(TABED) BAŞKANLIĞI

Sayı : TABED 1-25-1475

Ankara Bilkent Şehir Hastanesi Radyoloji Kliniği'nde yapılması planlanan "Derin Öğrenme Mimarileri ile Kellgren-Lawrence (KL) Evreleme Yapan Modelin Retrospektif Ön-Arka (AP) Diz Grafileri Üzerinde Değerlendirilmesi ve Radyolog Karar Süreçlerine Etkisinin İncelenmesi" konulu çalışma incelenmiş olup, Etik açıdan oy birliğiyle uygun görülmüştür.

Prof. Dr. Hurrem BODUR
Etik Kurul Başkanı

Etik Kurul Sekreterliği Üniversiteler Mah. Bilkent Cad. No:1 Çankaya/Ankara

Tel: 0 (312) 552 66 00 Dahili : 772 998 / 772 999