



T.C.
KONYA TEKNİK ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ



**ERKEN VE ORTAK FÜZYON STRATEJİLERİ
İLE DERİN VE GELENEKSEL GÖRÜNTÜ
ÖZELLİKLERİNİN ENTEGRASYONUNA
DAYALI DİZ OSTEOARTRİTİ
DERECELENDİRMESİ**

USAME YILMAZ

YÜKSEK LİSANS TEZİ

Bilgisayar Mühendisliği Anabilim Dalı

Nisan-2025
KONYA
Her Hakkı Saklıdır

TEZ KABUL VE ONAYI

Usame YILMAZ tarafından hazırlanan “ERKEN VE ORTAK FÜZYON STRATEJİLERİ İLE DERİN VE GELENEKSEL GÖRÜNTÜ ÖZELLİKLERİNİN ENTEGRASYONUNA DAYALI DİZ OSTEOARTRİTİ DERECELENDİRMESİ” adlı tez çalışması 07.04.2025 tarihinde aşağıdaki jüri tarafından oy birliği ile Konya Teknik Üniversitesi Lisansüstü Eğitim Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı’nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Jüri Üyeleri

Başkan

Prof. Dr. Gülay TEZEL
(Konya Teknik Üniversitesi)

Danışman

Dr. Öğr. Üyesi Fatma Zehra SOLAK
(Konya Teknik Üniversitesi)

Üye

Dr. Öğr. Üyesi Güzin ÖZMEN
(Selçuk Üniversitesi)

İmza

.....

.....

.....

Yukarıdaki sonucu onaylarım.

Prof. Dr. Mevlüt UYAN
Enstitü Müdürü

TEZ BİLDİRİMİ

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

DECLARATION PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

İmza

Usame YILMAZ

Tarih:

ÖZET

YÜKSEK LİSANS

ERKEN VE ORTAK FÜZYON STRATEJİLERİ İLE DERİN VE GELENEKSEL GÖRÜNTÜ ÖZELLİKLERİNİN ENTEGRASYONUNA DAYALI DİZ OSTEOARTRİTİ DERECELENDİRMESİ

Usame YILMAZ

Konya Teknik Üniversitesi
Lisansüstü Eğitim Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Fatma Zehra SOLAK

2025, 116 Sayfa

Jüri

Dr. Öğr. Üyesi Fatma Zehra SOLAK
Prof. Dr. Gülay TEZEL
Dr. Öğr. Üyesi Güzin ÖZMEN

Diz Osteoartriti (Knee Osteoarthritis - KOA), özellikle ileri yaş grubunda sık görülen ve yaşam kalitesini düşüren kronik bir eklem hastalığıdır. Geleneksel teşhis yöntemlerinin uzman görüşüne dayanması, sübjektif hatalara ve tedavi sürecinde gecikmelere yol açabilmektedir. Bu nedenle, daha güvenilir ve objektif kararlar sunabilen otomatik teşhis sistemlerinin geliştirilmesi büyük önem taşımaktadır. KOA'nın yalnızca varlığını değil, derecesini de doğru şekilde belirlemek, uygun tedavinin zamanında uygulanabilmesi açısından klinik açıdan kritik bir gerekliliktir.

Bu tezde, diz osteoartritinin (KOA) derecelerini daha doğru ve kararlı biçimde belirlemek amacıyla, geleneksel ve derin görüntü özelliklerinin erken ve ortak füzyon stratejileriyle birleştirildiği çok bileşenli bir yaklaşım önerilmiştir. X-ışını görüntülerinden elde edilen morfolojik, istatistiksel, GLCM, HOG, LBP, dalgacık, çevre, alan ve Zernike gibi geleneksel özellikler; ResNet-50, VGG16, EfficientNetB0, Xception ve DenseNet201 modellerinden çıkarılan derin temsillerle entegre edilmiştir. Bu özellikler, erken füzyon stratejisinde doğrudan birleştirilmiş, ortak füzyon stratejilerinde ise farklı düzeydeki temsiller dikkate alınarak bütünleştirilmiştir. Böylece KOA dereceleri, KL sistemine göre “sağlıklı”, “şüpheli”, “hafif”, “orta” ve “ağır” olmak üzere beş sınıfa ayrılmış; bu birleşik özellik temsilleri yapay sinir ağı (Artificial Neural Network- ANN) temelli sınıflandırma yapılarında değerlendirilmiştir. Ayrıca, sınıflar arası dengesizliği azaltmak amacıyla veri artırımı ve Sentetik Azınlık Aşırı Örnekleme Tekniği (Synthetic Minority Over-sampling Technique - SMOTE) yöntemi uygulanarak her KOA derecesinin dengeli biçimde temsil edilmesi sağlanmıştır.

Önerilen füzyon stratejileri, KOA derecelendirmesinde bilgi temsilini zenginleştirerek sınıflar arasındaki ayırım gücünü artırmıştır. Her füzyon stratejisinde elde edilen en yüksek doğruluk ve F1-skoru değerleri sırasıyla; Erken Füzyon Tip II için %73,56 ve %71,46, Ortak Füzyon Tip I için %83,50 ve %83,15, Ortak Füzyon Tip II için ise %85,39 ve %84,96 olarak kaydedilmiştir. En yüksek başarıya ulaşılan Ortak Füzyon Tip II stratejisinde, doku ve morfoloji temelli geleneksel özelliklerin yanı sıra, VGG16 modelinden elde edilen derin temsillerin KOA derecelerini ayırt etmede etkin rol oynadığı görülmüştür. Bu çok düzeyli bilgi entegrasyonu, özellikle erken ve ileri evre KOA derecelerinin güvenilir şekilde belirlenmesini sağlamış ve klinik kullanıma uygun güçlü bir yapay zeka temelli yaklaşım sunmuştur.

Anahtar Kelimeler: Diz Osteoartriti Derecelendirmesi, Geleneksel Görüntü Özellikleri, Derin Öğrenme Tabanlı Özellikler, Özellik Füzyon Stratejileri, Yapay Sinir Ağı Sınıflandırması

ABSTRACT

MS THESIS

GRADING OF KNEE OSTEOARTHRITIS BASED ON THE INTEGRATION OF DEEP AND TRADITIONAL IMAGE FEATURES VIA EARLY AND JOINT FUSION STRATEGIES

Usame YILMAZ

**Konya Technical University
Institute of Graduate Studies
Department of Computer Engineering**

Advisor: Assist Prof Dr. Fatma Zehra SOLAK

2025, 116 Pages

**Jury
Assist Prof Dr. Fatma Zehra SOLAK
Prof. Dr. Gülay Tezel
Assist Prof Dr. Güzin ÖZMEN**

Knee Osteoarthritis (KOA) is a chronic joint disease commonly observed in the elderly population and significantly reduces quality of life. Since conventional diagnostic methods often rely on expert interpretation, they are prone to subjective errors and may lead to delays in treatment. Therefore, developing automated diagnostic systems that can provide more reliable and objective decisions is of great importance. In particular, determining not only the presence of KOA but also its severity accurately is a critical requirement for ensuring timely and effective clinical interventions.

In this thesis, a multi-component approach is proposed to improve the accuracy and reliability of KOA grading by integrating handcrafted and deep image features through early and joint fusion strategies. Handcrafted features such as morphological, statistical, GLCM, HOG, LBP, wavelet, perimeter, area, and Zernike moments were extracted from X-ray images and combined with deep representations obtained from ResNet-50, VGG16, EfficientNetB0, Xception, and DenseNet201 models. These features were directly fused in the early fusion strategy, while in joint fusion strategies, representations from different levels were integrated. Accordingly, KOA grades were categorized into five classes “normal,” “doubtful,” “mild,” “moderate,” and “severe” based on the Kellgren-Lawrence (KL) grading system. The combined feature representations were evaluated using Artificial Neural Network (ANN)-based classification architectures. Additionally, data augmentation and the Synthetic Minority Over-sampling Technique (SMOTE) were applied to mitigate class imbalance and ensure balanced representation of each KOA grade.

The proposed fusion strategies enriched information representation and enhanced inter-class separability in KOA grading. The highest accuracy and F1-score values obtained for each fusion strategy were recorded as follows: 73.56% and 71.46% for Early Fusion Type II, 83.50% and 83.15% for Joint Fusion Type I, and 85.39% and 84.96% for Joint Fusion Type II, respectively. In the most successful configuration, Joint Fusion Type II, both handcrafted features related to tissue and morphology and deep representations derived from the VGG16 model played a significant role in distinguishing KOA grades. This multi-level feature integration enabled the reliable identification of both early and advanced KOA stages, providing a robust AI-based framework suitable for clinical use.

Keywords: Knee Osteoarthritis Grading, Traditional Image Features, Deep Learning-Based Features, Feature Fusion Strategies, Artificial Neural Network Classification

ÖNSÖZ

Bu çalışma, diz osteoartriti derecelendirmesi sürecinde bilgisayar destekli yöntemlerin kullanımını ele almak ve bu alanda yeni bir yaklaşım sunmak amacıyla hazırlanmıştır. Çalışmanın her aşamasında, doğru bilimsel yöntemleri uygulama ve literatüre katkı sağlayacak bir yöntem geliştirme hedefi güdülmüştür.

Çalışmanın yürütülmesi sürecinde bana rehberlik eden ve yol gösteren, bilgi ve tecrübeleriyle bu çalışmanın şekillenmesine büyük katkı sağlayan değerli danışman hocam Dr. Öğr. Üyesi Fatma Zehra SOLAK'a içten teşekkürlerimi sunarım. Kendisinin akademik birikimi, yapıcı geri bildirimleri ve sürekli desteği, bu tezin başarıyla tamamlanmasında kritik bir rol oynamıştır.

Bu çalışmanın gerçekleştirilmesi sırasında bana destek olan aileme ve yakınlarıma, verdikleri moral ve anlayışları için teşekkür ederim. Ayrıca, çalışmanın her aşamasında beni teşvik eden ve destekleyen tüm arkadaşlarıma minnettarım.

Bu tezin, diz osteoartriti derecelendirmesi alanında bilimsel bir katkı sağlayacağını ve ilgili alandaki diğer çalışmalara ilham vereceğini ümit ediyorum.

USAME YILMAZ
KONYA-2025

İÇİNDEKİLER

ÖZET	vii
ABSTRACT.....	viii
ÖNSÖZ	ix
İÇİNDEKİLER	x
SİMGELER VE KISALTMALAR	xii
1. GİRİŞ	1
2. KAYNAK ARAŞTIRMASI	4
2.1. İkili Sınıflandırma	4
2.2. Çok Sınıflı Sınıflandırma	9
2.3. Füzyon Stratejileri (Fusion Strategies)	21
3. MATERYAL VE YÖNTEM.....	21
3.1. Veri Seti.....	21
3.2. Ön İşleme	21
3.2.1 Kontrast Sınırlama Adaptif Histogram Eşitleme (Contrast Limited Adaptive Histogram Equalization - CLAHE):	23
3.2.2 Sentetik Azınlık Aşırı Örnekleme Tekniği (Synthetic Minority Over-sampling Technique - SMOTE)	24
3.3. Geleneksel Görüntü Özelliklerinin Çıkarılması.....	25
3.3.1. Morfolojik Özellikler	25
3.3.2. Gri Seviye Eş Oluşum Matrisi (Gray Level Co-occurrence Matrix-GLCM) Özellikleri	28
3.3.3. İstatistiksel Özellikler.....	32
3.3.4. Yönlendirilmiş Gradyanların Histogramı (Histogram of Oriented Gradients - HOG)	36
3.3.5. Yerel İkili Örüntü (Local Binary Pattern - LBP).....	38
3.3.6. Dalgacık Dönüşümü	39
3.3.7. Zernike Moment	40

3.3.8. Diğer Özellikler.....	41
3.4. Boyut Azaltma Yöntemleri.....	41
3.4.1. ANOVA	41
3.4.1.1. ANOVA ile Özellik Seçimi.....	42
3.4.2. Doğrusal Diskriminant Analizi (Linear Discriminant Analysis – LDA)	45
3.4.3. Temel Bileşenler Analizi (Principal Component Analysis - PCA).....	47
3.5. Derin Öğrenme Tabanlı Görüntü Özelliklerinin Çıkarılması.....	48
3.5.1. ResNet-50	49
3.5.2. VGG16.....	50
3.5.3. EfficientNetB0.....	51
3.5.4. Xception.....	52
3.5.5. Densenet201	53
3.6. Füzyon Stratejileri	54
3.6.1. Erken Füzyon Stratejilesinin Uygulanması	55
3.6.2. Ortak Füzyon.....	58
3.6.2.1. Ortak Füzyon Tip I Stratejisinin Uygulanması.....	59
3.6.2.2. Ortak Füzyon Tip II Stratejisinin Uygulanması	62
3.7. Performans Değerlendirmesi	64
4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA.....	67
4.1. Erken Füzyon Tip II Sonuçları.....	68
4.2. Ortak Füzyon Tip I Sonuçları.....	75
4.3 Ortak Füzyon Tip II Sonuçları	83
5. SONUÇLAR VE ÖNERİLER	97
KAYNAKLAR	99

SİMGELER VE KISALTMALAR

Simgeler

- I_p : Komşu piksel yoğunluğu,
 I_c : Merkez piksel yoğunluğu,
 C_i : ilgili renk kanalındaki her bir pikselin değeri,
 $\bar{C}$: Renk kanalının ortalama değeri.
 $\bar{x}$: Piksel Yoğunluklarının ortalaması
 x_i : Her bir pikselin yoğunluğu
 μ_i, μ_j : Yoğunluk ortalamaları,
 σ_i, σ_j : Yoğunluk standart sapmaları
 N : İlgili renk kanalındaki toplam piksel sayısı.
 E_k : ROI deki kenar piksellerin toplamı
 $s[[n, m]]$: Görüntünün piksel yoğunluk değerleri.
 $\emptyset_{j,k}[[n, m]]$: Ölçekleme (scaling) fonksiyonu (seviyeye bağlı).
 $\psi_{j,k}^H[[n, m]]$: Yatay dalgacık fonksiyonu.
 $\psi_{j,k}^V[[n, m]]$: Dikey dalgacık fonksiyonu.
 $\psi_{j,k}^D[[n, m]]$: Diyagonal dalgacık fonksiyonu.

Kısaltmalar

- CNN: Evrimsel Sinir Ağı (Convolutional Neural Network)
JSN: Eklem Aralığının Daralması (Joint Space Narrowing)
HOG: Yönlendirilmiş Gradyanların Histogramı (Histogram of Oriented Gradients)
GLCM: Gri Seviye Eş Oluşum Matrisi (Gray Level Co-occurrence Matrix)
KL: Kellgren-Lawrence
KOA: Diz Osteoartriti (Knee Osteoarthritis)
LBP: Yerel İkili Örüntü (Local Binary Pattern)
LDA: Doğrusal Diskriminant Analizi (Linear Discriminant Analysis)
OA: Osteoartrit
PCA: Temel Bileşenler Analizi (Principal Component Analysis)
ROI: İlgi Alanı Çıkarma (Region of Interest)
YSA: Yapay Sinir Ağları

1. GİRİŞ

Diz osteoartriti (Knee Osteoarthritis - KOA), dünya genelinde yaşlanan nüfusun karşı karşıya olduğu en yaygın kas-iskelet sistemi hastalıklarından biridir (Martel-Pelletier, 1999; Anderson ve Loeser, 2010). Diz eklemindeki kıkırdak dokusunun aşınmasıyla gelişen bu kronik hastalık, yalnızca fiziksel ağrı ve hareket kısıtlılığına yol açmakla kalmayıp, bireylerin yaşam kalitesinde ciddi düşüslere de neden olmaktadır. KOA'nın doğru teşhisi ve hastalık derecesinin belirlenmesi, erken dönemde müdahale edebilmek ve tedavi başarısını artırmak açısından kritik öneme sahiptir.

KOA teşhisinde sıklıkla kullanılan Kellgren-Lawrence (KL) derecelendirme sistemi (Kellgren ve Bier, 1956; Kohn ve diğ., 2016), hastalığın radyografik görüntülerdeki yapısal değişikliklere göre değerlendirilmesini sağlar. Bu sistem, osteofit oluşumu, eklem aralığının daralması (Joint Space Narrowing – JSN), subkondral skleroz ve deformasyon gibi radyolojik bulgulara dayanır. Ancak KL derecelendirmesi uzman yorumu gerektirdiğinden, değerlendirme süreci subjektif farklılıklara açık olup, teşhis ve tedavi kararlarında tutarsızlıklara yol açabilmektedir. Bu nedenle, bilgisayar destekli sistemler kullanılarak KOA derecelerinin daha nesnel, doğru ve kararlı şekilde belirlenmesi gerekmektedir.

Son yıllarda görüntü işleme ve derin öğrenme alanında yaşanan gelişmeler, KOA gibi karmaşık hastalıkların otomatik analizi için önemli fırsatlar sunmuştur. Derin öğrenme modelleri, güçlü temsil öğrenme yetenekleri sayesinde umut verici sonuçlar vermektedir. Ancak bu modellerin sıfırdan eğitilmesi, büyük ve dengeli veri setleri gerektirdiğinden, özellikle tıbbi görüntüleme alanında sınırlı veri sorunu önemli bir engel oluşturmaktadır. Bu bağlamda, önceden eğitilmiş ağların yeniden kullanıldığı transfer öğrenme yöntemleri, veri kısıtlılığı problemini aşmak için etkili bir çözüm olarak öne çıkmıştır.

Literatür çalışmaları incelendiğinde, KOA teşhisine yönelik mevcut yaklaşımlarda bazı önemli eksiklikler olduğu görülmektedir. Birçok çalışmada yalnızca derin öğrenme veya transfer öğrenme algoritmaları doğrudan sınıflandırma amacıyla kullanılmıştır. Ancak bu yöntemler, özellikle küçük ve dengesiz veri setlerinde sınırlı genelleme başarısı göstermiştir. Diğer bazı çalışmalarda ise yalnızca sınırlı sayıda geleneksel özellik çıkarım tekniğine (morfolojik, istatistiksel, doku tabanlı vb.) dayalı klasik makine öğrenmesi yöntemleri kullanılmış; fakat bu yaklaşımlar da zengin ve çok

katmanlı bilgi temsili oluşturmamışları için sınıflandırma doğrulukları istenen düzeye ulaşamamıştır.

Öte yandan, farklı bilgi kaynaklarının sentezlenmesi yoluyla sınıflandırma performansını artırmayı hedefleyen füzyon stratejileri literatürde yeterince araştırılmamıştır. Özellikle erken ve ortak füzyon stratejilerinin etkinliğini sistematik biçimde inceleyen çalışmaların eksikliği dikkat çekmektedir. Bunun yanı sıra, sınıf dengesizliği sorunlarının yeterince ele alınmadığı ve düşük temsil edilen sınıflarda başarı oranlarının yetersiz kaldığı da önemli bir diğer eksiklik olarak öne çıkmaktadır.

Bu çalışmada, literatürdeki bu eksiklikleri gidermek amacıyla, geleneksel görüntü özellikleri (morfolojik, istatistiksel, GLCM, HOG, LBP, dalgacık dönüşümü, alan, çevre ve Zernike momentleri) ile transfer öğrenme modellerinden elde edilen derin özelliklerin bütüncül bir yapıda birleştirilmesi önerilmektedir. Derin özellik çıkarımı için, farklı mimari yapıları ve derinlik seviyeleriyle literatürde geniş kullanım alanı bulan ResNet-50, VGG16, EfficientNetB0, Xception ve DenseNet201 modelleri tercih edilmiştir (Usman ve diğ., 2023; Wani ve diğ., 2023; Chen ve diğ., 2019; Lau ve diğ., 2022; Solak, 2024; Mahfus ve diğ. 2025). Tez kapsamında transfer öğrenme modelleri doğrudan sınıflandırma amacıyla değil, daha çok füzyon stratejilerinde kullanılacak zengin ve çok katmanlı derin özelliklerin elde edilmesi amacıyla kullanılmıştır. Elde edilen çok boyutlu özellik kümesi, ANOVA, Doğrusal Diskriminant Analizi (Linear Discriminant Analysis – LDA) ve Temel Bileşenler Analizi (Principal Component Analysis – PCA) gibi istatistiksel yöntemlerle işlenerek, hem seçici hem de ayrıştırıcı nitelik taşıyan düşük boyutlu temsillere indirgenmiştir.

Özelliklerin birleştirilmesi sürecinde, çalışmanın temel katkı unsurlarından biri olan füzyon stratejileri uygulanmıştır. Geleneksel ve derin görüntü temsilleri, Erken Füzyon stratejisinde doğrudan birleştirilmiş; Ortak Füzyon stratejisinde ise farklı seviyelerde özetlenen temsiller sistematik biçimde entegre edilmiştir. Böylece, heterojen bilgi kaynakları çok düzeyli bir yapı altında bütünleştirilerek KOA derecelendirmesinde daha güçlü ve ayrıştırıcı özellik temsilleri oluşturulmuştur. Elde edilen bu birleşik vektörler, yapay sinir ağı (Artificial Neural Network – ANN) tabanlı sınıflandırma mimarisinde değerlendirilmiş ve her bir görüntü için KOA derecesinin tahmini gerçekleştirilmiştir.

Önerilen yaklaşım doğrultusunda, diz osteoartriti KL derecelendirme sistemine göre “sağlıklı”, “şüpheli”, “hafif”, “orta” ve “ağır” olmak üzere beş farklı kategoriye ayrılmıştır. Sınıflar arasındaki dengesiz örnek dağılımını dengelemek amacıyla veri

artırımı uygulanmış, azınlık grupların temsiliyetini güçlendirmek üzere Sentetik Azınlık Aşırı Örnekleme Tekniğı (Synthetic Minority Over-sampling Technique – SMOTE) kullanılmıştır. Böylece, sınıflandırma sürecinde her derecenin dengeli biçimde öğrenilmesi sağlanmış ve modelin genel kararlılığı artırılmıştır.

Bu tez çalışması sonucunda, çoklu bilgi kaynaklarının etkin biçimde bütünleştirilmesiyle KOA derecelendirmesinde bilgi temsiline zenginleştirildiğı, dereceler arası ayırım yeteneğinin artırıldığı ve doğruluk oranlarının literatürdeki pek çok çalışmaya kıyasla anlamlı şekilde iyileştirildiğı gösterilmiştir. Elde edilen sonuçlar, yalnızca KOA teşhisi için değil, klinik teşhis süreçlerinde genel kullanım potansiyeline sahip daha objektif ve güvenilir bir yaklaşım sunulduğunu ortaya koymaktadır.



2. KAYNAK ARAŞTIRMASI

Son yıllarda, diz osteoartriti (KOA) teşhisinde bilgisayar destekli yöntemlere yönelik ilgi ve araştırmalar belirgin şekilde artmıştır. Bu alanda gerçekleştirilen araştırmalar, temel olarak görüntü işleme, özellik çıkarımı, klasik makine öğrenmesi ve derin öğrenme tabanlı sınıflandırma yöntemleri etrafında yoğunlaşmaktadır. Literatürde, farklı sınıflandırma problemleri için çeşitli yaklaşımlar sunulmuş olup, bu yaklaşımlar genellikle sınıflandırma türlerine ve kullanılan stratejilere (örneğin, özellik füzyonu gibi) göre kategorize edilmektedir.

2.1. İkili Sınıflandırma (Binary Classification)

İkili sınıflandırma yöntemleri, KOA teşhisinde normal ve anormal durumların ayrımını yapmak için kullanılan en temel sınıflandırma yöntemlerinden biridir. Bu çalışmalar, genellikle sağlıklı ve osteoartritli dizlerin ayrımına odaklanmaktadır.

Kokkotis ve diğ. (2020), makine öğrenmesi tekniklerinin KOA teşhisi ve tahmini için nasıl kullanılabileceğine odaklanan bir derleme makalesi sunmaktadır. 2006-2019 yılları arasında yayımlanan makalelere dayanarak, derleme dört bölüme ayrılmıştır: segmentasyon, optimum sonrası tedavi planlama teknikleri, sınıflandırma ve tahmin/regresyon. Sonuçları, KOA'yı tahmin etmek için kullanılan çoğu tanı modellerinin doğruluk oranlarının %76,1 ile %92 arasında değiştiğini göstermektedir. Öğrenme teknikleri, doğrulama, sınıflandırma, segmentasyon vb. gibi konuların hepsi gözden geçirilmiş ve özetlenmiştir. Çalışma, makine öğrenmesinin KOA için yeni otomatik önce veya sonra tedavi çözümlerinin oluşturulmasında önemli bir rol oynadığını vurgulamaktadır.

Saleem ve diğ. (2020), osteoartriti tanılamak için eklem aralığı genişliğine dayalı bir model önermişlerdir. Yöntemleri, görüntü ön işleme, ilgi bölgesini çıkarma, kenarları hesaplama ve eklem aralığı genişliğini hesaplama aşamalarını içermektedir. Veri seti, iki radyolog ve iki ortopedi cerrahının danışmasıyla belirlenen ciddiyet düzeyine dayalı olarak etiketlenmiş 140 görüntü içermektedir. Önerilen yöntemler, KOA sınıflandırması için %98,4 F1 skoru ve %97,14 doğruluk elde etmiştir. Algılama yaklaşımlarının yüksek doğruluğuna rağmen, bu kadar küçük bir veri setinin gerçek dünya durumlarında modelin doğruluğunu etkileyebileceği unutulmamalıdır.

Brahim ve diğ. (2019), makine öğrenmesi algoritmalarını kullanarak KOA tespiti için bir bilgisayar destekli teşhis sistemini sunmaktadır. Yaklaşımları, X-ışını görüntülerini ön işlemden geçirme ve ardından çok değişkenli lineer regresyon kullanarak normelleştirme yöntemi uygulama aşamasıyla başlamaktadır. Özellik çıkarma aşamasında bağımsız bileşen analizi kullanılmakta ve son olarak Rastgele Orman ve Naive Bayes modelleri sınıflandırma sürecinde kullanılmaktadır. Yöntemleri için, Osteoartrit İnisiyatifi (Osteoarthritis Initiative - OAI)'den alınan 1024 diz X-ışını görüntüsünü kullanmışlardır. Denemelerinde %82,98 doğruluk, %80,65 özgüllük ve %87,15 duyarlılık elde edilmiştir. Bu yaklaşımda, özellik çıkarma yalnızca piksel yoğunluklarına dayanmaktadır ve bu da gerçek dünya koşullarında yüksek doğruluklara yol açabilir.

Nasser ve diğ. (2020), ayırt edici düzenlenmiş bir otomatik kodlayıcı önermektedir ve bu, tespit sürecini hassaslaştırmak için hem ayırt edici hem de anlamlı özellikleri anlamamıza olanak tanımaktadır. Eğitim kriterlerine ayırt edici kayıp eklenmesi, öğrenilen modelin aynı zamanda ayırt edici verileri de içermeye zorlamayı amaçlamaktadır. Araştırmacılar, çalışmalarında, OAI tarafından oluşturulan genel bir veritabanından elde edilen toplam 3900 diz radyografisini kullanmışlardır. Sonuçlarını diğer derin öğrenme yöntemleriyle karşılaştırmışlar ve doğruluklarını %82,53 olarak göstermişlerdir. Sonuçlar, yaklaşımlarının diğer güncel derin öğrenme tekniklerinden daha yüksek doğruluklar sunduğunu göstermiştir.

Lim ve diğ. (2019), osteoartriti tespit etmek için hasta sağlık davranışları ve tıbbi kullanım verilerinin istatistiksel verilerini kullanan bir derin sinir ağı kullanmışlardır. Bu çalışma, 2015-2016 KNHANES'ten alınan 5749 öznenin verilerini kullanmıştır. Önerilen yöntemde, KOA'nın risk faktörleri, özellikleri otomatik olarak çıkarmak için ölçekli bir temel bileşen analizi ve derin sinir ağları kullanılarak tespit edilmiştir. Sınıflandırma modeli %71,97 doğruluk ve %66,67 duyarlılık sağlamıştır. Bu makale, X-ışını görüntüleri kullanmadığından, osteoartriti tespit eden bir model oluşturmak için sağlık davranışları ve tıbbi kullanımın kullanılması umut vadeden alternatiflerdir ve muhtemelen diğer yöntemlerle işbirliği içinde kullanılabilir.

Guan ve diğ. (2022), erken KOA tespiti performansını artırmak için geleneksel ve derin öğrenme modelinin bir kombinasyonunu önermişlerdir. Geleneksel model, yapay sinir ağları, rastgele orman ve lojistik regresyon kullanılarak oluşturulmuştur. Derin öğrenme modeli, ağrı ilerleme olasılıklarını belirlemek için Evrişimli Sinir Ağlarını (Convolutional Neural Network - CNN) içermektedir. Veri seti, OAI'den elde

edilen 1389 KOA'lı ve 3285 KOA'sız bireyi içermektedir. Kombine modelin rapor edilen ROC eğrisi altında kalan alanı 0,807, duyarlılığı %72,3 ve özgüllüğü %80,9 olarak belirlenmiştir. Çalışma, değişkenlik gösteren veri setlerine uygulanan geleneksel ve derin öğrenme modellerinin birleştirilmesinin, KOA tespitinde daha yüksek doğruluklara yol açabileceğini göstermektedir.

Kokkotis ve diğ. (2020), radyologların hastalardaki KOA varlığını belirlemede yardımcı olabilecek bir özellik çıkarma yaklaşımı sunmaya çalışmışlardır. Buna ek olarak, çıkartılan özelliklere dayalı otomatik KOA tespiti için bir makine öğrenmesi modeli önermişlerdir. Destek Vektör Makinesi ve K-En Yakın Komşu algoritmalarını diz radyografilerini sınıflandırmak için kullanmışlardır. Çalışma, OAI'den alınan verileri kullanarak diz radyograflarını ilerlemeyenler ve KOA ilerleyenler olarak iki gruba ayırmıştır. Destek Vektör Makinesi yaklaşımıyla %74,07 sınıflandırma doğruluğu elde etmişlerdir.

Lau ve diğ. (2022), ImageNet, Xception modeli ve toplam diz artroplastisi hastalarının X-ışını görüntülerinden oluşan veri seti üzerinde bir yöntem geliştirmiş ve görüntü tabanlı makine öğrenmesi modeli oluşturmuşlardır. Bir klinik bilgi tabanı oluşturmak amacıyla, bu hastaların klinik parametrelerini içeren bir veri seti üzerinde rastgele orman sınıflandırıcısı kullanılarak bir makine öğrenmesi modeli geliştirilmiş ve daha sonra bu model farklı bir sistemle entegre edilmiştir. Modelin yaptığı tahmin seçimini yorumlamak için sınıf aktivasyon haritaları da kullanılmıştır. Makine öğrenmesinin görüntüler üzerindeki model gevşemesi için doğruluk oranı ve hatırlatma oranı sonuçları sırasıyla 0,92 ve 0,96'ya ulaşmıştır, görüntü sınıflandırması için %96,3'lük bir doğruluk oranı gözlemlenmiştir. Ancak, klinik bilgi tabanlı bir modelin eklenmesi, %71'lik bir doğruluk oranı ve %20'lik bir hatırlatma oranı ile doğrulukta bir gelişme göstermemiştir.

Schiratti ve diğ. (2021), KOA ile ilgili bir çalışma için diz MR görüntüleri üzerinde derin öğrenme kullanmışlardır. Araştırmacılar, MR görüntüleri ve klinik değişkenlere dayanarak daha fazla kıkırdak bozulmasını tahmin etmişlerdir. Sınıflandırma modeli için %65'lik bir ROC eğrisi altında kalan alan değeri başarıyla elde etmişlerdir.

Rehman de diğ. (2023), osteoartriti diz röntgen görüntülerinden etkili bir şekilde teşhis etmek için CNN-Rastgele Orman K-komşular adı verilen yenilikçi bir transfer öğrenme tabanlı özellik mühendisliği tekniği uygulamışlardır. Bu model, CNN'nin eğitilmesi için transfer öğrenme tabanlı bir yaklaşım sunar ve uzamsal özellikler

çıkarmak için modeli kullanarak tahmin yapmasını sağlar. Ana fikir, görüntülerin 2 evrişim katmanı ve 2 maksimum havuzlama katmanından oluşan bir CNN'e beslenmesidir. Çıkarılan uzamsal özellikler, aynı görüntülerden olasılıksal özellikler çıkarmak için bir Rastgele Orman ve bir K-En Yakın Komşu modelini eğitmek için kullanılır. Böylece model, diz görüntüsünün sınıfını ve güven düzeyini tahmin etmek için kullanılır. Çıkarılan olasılıksal özellikler, modelin daha doğru olmasını sağlar. Çalışmada %99 doğruluk, duyarlılık, f1-skoru ve kesinlik değerleri elde edilmiştir. Yani, önerilen yöntem osteoartrit teşhisinde oldukça yüksek bir başarı göstermiştir.

Rasool ve diğ. (2023), diz osteoporozu röntgen görüntülerinin ikili sınıflandırmasını ele almışlardır. Burada kullanılan veri seti, diğer çalışmalara kıyasla nispeten daha küçüktür. Bu çalışma, bireysel modellerin hatalarından kaynaklanan önyargılardan ve yanlılıklardan kaçınmak için tek bir yönteme veya modele bağlı kalmamayı hedefler. Bu, daha iyi performans gösteren modellere daha yüksek ağırlıklar vererek ve düşük performans gösteren modellere daha düşük ağırlıklar atayarak sağlanmıştır. Modeller, en iyi performansı elde etmek için birleştirilmiştir. Çoğu durumda, birden fazla modelin birleştirilmesi, bireysel modellerden daha iyi sonuçlar vermiştir. Bu çalışmada, Kaggle'dan alınan halka açık bir diz röntgen görüntü veri seti kullanılmıştır. Bu veri seti, her biri 2 sınıftan birine ait olan (Osteoporotik ve Osteoporotik Olmayan diz) toplam 372 görüntü içermektedir. %97 doğruluk olarak elde edilmiştir.

Usman ve diğ. (2023), diz röntgen görüntülerinde osteoporozu tanımlamak ve sınıflandırmak için üç transfer öğrenme modelini karşılaştırılmaya çalışmıştır. Uygulanan transfer öğrenme modelleri GoogLeNet, VGG-16 ve ResNet-50'dir. Ayrıca, modellerin performanslarını daha iyi karşılaştırmak için birçok son teknoloji metrik kullanılmıştır. Bu çalışmada kullanılan veri seti, Kaggle'ın Osteoporoz Diz Röntgeni Veri Seti-versiyon 1 (2021)'dir. Görüntü sayısı, Python'un veri artırma işlevleri kullanılarak istatistiksel veri artırma teknikleriyle artırılmıştır. Sonuç olarak, veri seti 323 normal diz görüntüsü ve 323 osteoporotik diz görüntüsü içermektedir. Bu görüntülerden 207'si eğitim için, 52'si doğrulama için ve 65'i test için kullanılmıştır. Görüntüler, Python'un ImageDataGenerator yöntemi kullanılarak, yeniden ölçeklendirme parametresi 1/255 olacak şekilde yeniden boyutlandırılmıştır. Keras'ın ImageDataGenerator sınıfı, gerçek zamanlı görüntü artırma teknikleri için kullanılmış ve standartlaştırma, döndürme, kaydırmalar, parlaklık değişiklikleri ve birkaç başka yöntem uygulanmıştır. Kullanılan model mimarileri GoogLeNet, VGG-16 ve ResNet-

50'dir. Kaybı ölçmek için kullanılan metrik ikili çapraz entropi olmuştur. Optimizasyon için RMSProp kullanılmış ve öğrenme oranı 0,001 olarak ayarlanmıştır. Her model 30 iterasyon boyunca eğitilmiştir. Tüm transfer öğrenme modellerinde, veri seti %80 eğitim ve %20 test olacak şekilde bölünmüştür. Model doğruluk oranları şu şekildedir: GoogLeNet model doğruluğu: %90 VGG-16 model doğruluğu: %87 ResNet-50 model doğruluğu: %83.

Wang ve diğ. (2023), OAI veri setinden elde edilen görüntüler kullanarak, VGG16, VGG19, ResNet101, MobileNetV2, InceptionResNetV2 ve DenseNet121 olmak üzere altı farklı önceden eğitilmiş derin sinir ağı modelinin KOA teşhisinde uygulanmasını incelemiştir. Çalışmada, KOA'nın varlığını veya yokluğunu belirlemeyi amaçlayan ikili sınıflandırma yöntemi kullanılmıştır. Bu amaçla özel olarak yapılandırılmış bir veri seti olan Veri Seti II üzerinde deneyler gerçekleştirilmiştir. Test edilen modeller arasında ResNet101, %83 doğruluk oranı ile en yüksek performansı göstermiştir. Çalışma sonuçları, literatürdeki önceki çalışmalara kıyasla daha iyi bir performans ortaya koyarak, derin öğrenme tabanlı otomatik KOA teşhisinin etkinliğini desteklemektedir.

Wei ve diğ. (2023), hastanın KAO olup olmadığını ikili sınıflandırma ile tespit etmek için yaş ve radyomik özelliklerin kullanımına odaklanmaktadır. Diz röntgen görüntülerinden radyomik özellikler çıkarılmış ve seçilmiş, ardından hasta KL-grade aşamalarına göre sınıflandırılmıştır. Çalışmada kullanılan veri seti, Çin'deki bir özel hastaneden, diz osteoartriti yaşayan hastalardan toplanmıştır. Röntgen görüntüleri ön işleme tabi tutulmuş, 2270x2270 piksel (yükseklik * genişlik) boyutlarına yeniden boyutlandırılmış ve tüm görüntüler normal dağılıma standartlaştırılmıştır. Veri seti, %80 eğitim ve %20 test olacak şekilde bölünmüştür. Lojistik regresyon en yüksek AUC puanını 0,843 ile elde etmiştir.

Solak (2024), KOA tanısında ikili sınıflama yaklaşımını benimsemiştir. Çalışmada, KL derecelendirme sistemine göre, derecesi 0 (sağlıklı) ve 1 (şüpheli) olan bireyler "KOA yok" grubunda; derecesi 2 (hafif), 3 (orta) ve 4 (ağır) olan bireyler ise "KOA var" grubunda toplanmıştır. Böylece, yapısal değişikliklerin henüz belirgin olmadığı erken evre bireyler ile klinik müdahale gerektirebilecek seviyede KOA bulguları taşıyan bireyler arasında ayırım yapılması hedeflenmiştir. Sınıflandırma sürecinde, ResNet-50, Xception, VGG16, EfficientNetB0 ve DenseNet201 gibi farklı transfer öğrenme modelleri kullanılmış; her model rmsprop, sgdm ve adam optimizasyon algoritmaları ile ayrı ayrı eğitilerek test edilmiştir. Elde edilen sonuçlar,

özellikle DenseNet201 modelinin rmsprop optimizasyon algoritması ile birlikte %87,7 doğruluk, %87,2 F1-skoru ve 0,75 Cohen's Kappa değeri elde ederek en yüksek başarıyı sağladığını göstermiştir. Bu bulgular, Solak'ın (2024) çalışmasında ikili sınıflama kapsamında yüksek doğruluk ve güvenilirlik seviyelerine ulaşıldığını ve KOA varlığının etkin bir şekilde tespit edilebildiğini ortaya koymaktadır.

2.2. Çok Sınıflı Sınıflandırma (Multiclass Classification)

Çok sınıflı sınıflandırma yöntemleri, KOA teşhisinde farklı hastalık derecelerini veya gruplarını ayırt etmek amacıyla geliştirilmiştir. Bu çalışmalar, genellikle üçlü (3 sınıf), dördü (4 sınıf) ve beşli (5 sınıf) sınıflandırma problemlerini kapsar. Çeşitli sınıflandırma sistemleri, özellikle Kellgren-Lawrence (KL) derecelendirme ölçeği gibi standart ölçütlere dayanarak, hastalığın farklı ilerleme düzeylerini doğru bir şekilde belirlemeyi hedefler. Bu yöntemler, KOA'nın daha gelişmiş seviyelerde analizini mümkün kılar ve tanı süreçlerini iyileştirmeye katkı sağlar.

Üçlü (3- Sınıf) Sınıflandırma

Von Tycowicz ve diğ. (2021), aşırı uyum sorununu ve şekil uzayının non-Euclidean modelini dikkate alarak ve intrinsik boyutu azaltma konseptini birleştirmeye çalışmışlardır. Sınıflandırıcının performansını alternatif bir ekstrinsik yaklaşım ile karşılaştırmışlardır. Kullanılan veri seti, OAI tarafından oluşturulan ve 201 rastgele seçilmiş örneği içeren bir veri setidir. Görüntüler sıfır, bir ve iki olarak derecelendirilmiş ve sıfır, dakika ve belirgin osteofit varlığı olarak belirlenmiştir. İçsel modelin başarı oranı %64,64 iken, Euclidean yöntemi %58,62 doğruluk oranı sağlamıştır. Modelin nispeten düşük tahmin doğruluğu, kullanılan derecelendirme sisteminin standart KL derecelendirme sisteminden farklı olmasından kaynaklanıyor olabilir.

Wani ve diğ. (2023), diz röntgen görüntülerinden osteoporoz tespiti için yeni bir derin öğrenme tabanlı yaklaşım sunmuşlardır. Bu, omurga veya kalça gibi diğer vücut bölgelerine kıyasla daha az araştırılmış bir uygulama alanıdır. Araştırmacılar, AlexNet, VGGNet ve ResNet gibi popüler CNN mimarileriyle deneyler yapmış ve önceden eğitilmiş AlexNet modelinin, röntgenleri normal, osteopenik ve osteoporotik kategorilere sınıflandırmada en yüksek doğruluğu olan %91'e ulaştığını tespit

etmişlerdir. Bu otomatik ve maliyet etkin çözüm, klinisyenlere osteoporozun erken tespitinde yardımcı olabilir ve kırık riskini azaltabilir. Ancak, özellikle normal ve osteoporotik sınıflar için küçük veri seti boyutu, modelin performansını olumsuz etkilemiş olabilecek bir sınırlamadır. Çalışma, diz osteoporozu teşhisi için derin öğrenmenin potansiyelini göstermektedir. Bununla birlikte, veri setinin genişletilmesi, klinik faktörlerin entegrasyonu ve modelin harici veri setlerinde genelleştirme yeteneğinin doğrulanması gibi konularda daha fazla araştırma yapılması gerektiği belirtilmektedir.

Wang ve ark. (2023), OAI veri setinden elde edilen görüntüleri kullanarak ikili sınıflandırmaya ek olarak 3 sınıflı sınıflandırma işlemi de gerçekleştirmişlerdir. Bunun için veri setini 3 sınıflı olacak şekilde yeniden düzenlemişler (Veri Seti 3) ve VGG16, VGG19, ResNet101, MobileNetV2, InceptionResNetV2, DenseNet121 modelleri ile KOA şiddeti teşhisini gerçekleştirmişlerdir. ResNet101 modeli, üç sınıf için %89 doğruluk oranı ile en iyi sonucu sağlamıştır.

Solak (2024), KOA derecelendirmesinde üçlü sınıflama yaklaşımını da değerlendirmiştir. Çalışmada, KL sistemine göre, derecesi 0 ve 1 olan bireyler “sağlıklı/şüpheli” grubunda, derecesi 2 ve 3 olan bireyler “erken-orta evre KOA” grubunda, derecesi 4 olan bireyler ise “ileri KOA” grubunda toplanmıştır. Böylece, yapısal değişikliklerin henüz sınırlı olduğu bireyler ile hafif düzeyde etkilenmiş ve ileri düzey KOA gelişimi göstermiş bireyler arasındaki ayrımın daha net yapılması amaçlanmıştır. Bu ayrım, hem hastalığın ilerleme düzeyine göre tanı koymak hem de uygulanacak tedavi sürecini doğru şekilde yönlendirmek açısından önemli bir rol oynamaktadır. Sağlıklı/şüpheli grupta koruyucu önlemler, hafif-orta KOA grubunda erken müdahaleler, ileri KOA grubunda ise daha yoğun tedavi yöntemleri uygulanmaktadır. Sınıflandırma sürecinde ResNet-50, Xception, VGG16, EfficientNetB0 ve DenseNet201 gibi transfer öğrenme modelleri kullanılarak farklı optimizasyon algoritmaları (rmsprop, sgd ve adam) ile eğitimler gerçekleştirilmiştir. Elde edilen sonuçlar, DenseNet201 modelinin rmsprop algoritması ile %85,6 doğruluk, %82,4 F1-skoru ve 0,71 Cohen’s Kappa değeri ile en yüksek performansı gösterdiğini ortaya koymuştur. Bu bulgular, Solak’ın (2024) çalışmasında üç sınıflı KOA derecelendirmesinde başarılı bir ayrım yapıldığını ve özellikle ileri evre KOA tespitinde modelin güçlü bir performans sergilediğini göstermektedir.

Beşli (5-Sınıf) Sınıflandırma

Thomas ve diğ. (2020) , radyograflardan KOA'nın şiddetini otomatik olarak tespit etmek için bir model geliştirmişlerdir. Modelin performansını test etmek için, muskuloskeletal radyologların görüşleriyle karşılaştırma yapmışlar. Radyografi görüntüleri önce bir CNN modeline giriş olarak kullanılmadan önce otomatik olarak artırılmıştır. Bu çalışmada kullanılan veri kümesi, 40000'den fazla görüntü içeren OAI veri setinden alınmıştır. Yazarlar test veri seti için ortalama doğruluk oranını %71 ve F1 skorunu %70 olarak raporlamışlardır. Büyük bir veri kümesi kullansalar da, modelin performansını eğitimli radyologların görüşleriyle değil, alanla ilgili diğer çalışmalarla karşılaştırmaları gereklidir.

Yong ve diğ. (2021), KL ölçek sistemini ordinal regresyon problemi olarak ele almak amacıyla nöral ağlar için ordinal regresyon modülü oluşturmayı amaçlamışlardır. Modelin performansını çeşitli mevcut nöral ağ çözümleriyle değerlendirmişlerdir. Kullanılan veri seti, OAI'den elde edilen ve toplamda 8260 dizi radyografisini içeren bir veri kümesidir. Bu radyografiler, sırasıyla doğrulama, test ve eğitim veri kümelerine 1:2:7 oranında bölünmüştür. Önerilen yaklaşım ile eğitilmiş DenseNet-161'in %88,09 doğruluk oranı ve 0,8609 Quadratic Weighted Kappa skoru elde ettiği bildirilmiştir. Yaklaşımları bazı iyi sonuçlar vermiş olsa da, modelin KL dereceleri sıfır ve bir olan KOA görüntülerini sınıflandırırken birkaç yanlış sınıflandırma yaptığını belirtmişlerdir.

Kwon ve diğ. (2020), daha önce gerçekleştirdikleri çalışmayı genişleterek KOA için otomatik bir sınıflandırma modeli geliştirmeyi amaçlamışlardır. Destek Vektör Makinesini, Inception-Net-V2'yi kullanan önceki çalışmalarındaki modele sınıflandırma için kullanılmıştır. Seul'deki bir hastanenin Kurumsal İnceleme Kurulu'nun onayı ile toplam 364 hastanın 728 diz görüntüsünden oluşan bir veri seti toplanmıştır. Çalışma, F1 skoru, duyarlılık ve hassasiyet açısından sırasıyla 0,71, 0,70 ve 0,76 başarı elde etmiştir. Bu çalışmada, radyografi KOA görüntülerinin şiddeti ile çıkarılan yürüyüş özellikleri arasında doğrudan bir ilişki kanıtlanmıştır. Ayrıca, çıkarılan özelliklerden yararlanarak daha hassas bir otomatik sınıflandırma modeli tasarlanabileceği belirtilmiştir.

Gornale ve diğ. (2020), diz X-ışını görüntülerinin geometrik değişikliklerini incelemek için Hu'nun invariant momentlerini kullanarak KOA'yı erken aşamalarda tespit etmişlerdir. Önerilen yöntem, istenmeyen bozulmaları gideren, kıkırdak bölgesini tanımlayan ve çıkararak, Hu'nun invariant momentlerini hesaplayan ve son olarak k en

yakın komşu ve karar ağacı modellerini kullanarak sınıflandıran bir görüntü ön işleme aşamasını içermektedir. 2000 görüntü içeren kendi veri setleri, KL derecelendirme çerçevesine göre iki tıp uzmanının manuel işaretleme ile oluşturulmuştur. Her iki uzmanın verilerinden elde edilen genel doğruluk oranı yaklaşık %99,23 olmuştur. Yüksek doğrulukları, yoğunluk temelinde belirlenen ilgi bölgelerini segmente etmelerinden kaynaklanmıştır.

Chen ve diğ. (2019), iki derin evrişimli sinir ağını başarıyla kullanarak otomatik olarak KOA'nın ve şiddetinin tespitini gerçekleştirmişlerdir. Bu çalışmada kullanılan temel X-ışını görüntüleri OAI'den alınmıştır. Önerilen yaklaşım ilk olarak özel bir YOLOv2 ağı kullanarak görüntülerdeki diz eklemelerini tanımıştır. DenseNet, VGG, ResNet ve InceptionV3 varyasyonlarını ayarlayarak, diz X-ışını görüntülerini KL derecelendirme sistemine dayanarak şiddet sınıflarına ayırmayı başarmışlardır. Diz eklem tespit yöntemleri 0,858 ortalama Jaccard endeksini ve %92,2 hatırlamayı, kalibre edilmiş VGG-19 modeli ise diz osteoartritin şiddetini %69,7 doğrulukla tespit etmiştir.

Yunus ve diğ. (2022), çalışmalarında, KOA şiddetinin otomatik olarak tanımlanması ve sınıflandırılması için yeni bir teknik geliştirmiştir. Geleneksel KL derecelendirme sistemi zaman alıcı olup yanlış sınıflandırmalara yol açabilirken, otomatik derecelendirme kısa sürede güvenilir sonuçlar sağlayabilir. Önerilen model şu adımları içermektedir: (a) ön işlenmiş veri seti, (b) el yapımı özelliklerin çıkarılması, (c) önceden eğitilmiş CNN modellerinden derin özelliklerin çıkarılması, (d) en iyi özelliklerin seçimi için temel bileşenler analizi, (e) özellik birleştirme, (f) sınıflandırma ve (g) YOLO-v2ONNX modeli kullanılarak sınıflandırılmış görüntülerin lokalizasyonu. Önerilen teknik, topluluk k en aykırı komşu sınıflandırıcısı ile 0. Derece için 0,95, 1. Derece için 0,85, 2. Derece için 0,82, 3. Derece için 0,85 ve 4. Derece için 0,81 hassasiyet oranına ulaşmıştır. KOA'nın lokalizasyonu için, YOLO-v2'nin ONNX modeli kullanılarak geliştirilen YOLO-v2ONNX modeli, sınıflandırılmış görüntülerde 0,96 IOU ve 0,98 mAP değerlerine ulaşmıştır.

Leung ve diğ. (2020), total diz replasmanı geçiren hastaların diz görüntülerini kullanarak bir derin öğrenme sınıflandırma modeli önermişlerdir ve bunu replasman yapmayan hastalarla karşılaştırmışlardır. KL tabanlı sınıf seviyeleri arasında ayırım yapmak için ResNet34 ve çapraz doğrulama kullanılmıştır. Çalışma, OAI'den toplam 4796 görüntü içeren bir veri seti kullanmıştır. Önerilen modelin doğruluğu %72,7 olarak

rapor edilmiştir. Sınırlı veri seti boyutu ve transfer öğrenme kullanımı, modelin daha yüksek doğrulukla performans göstermesini engellemiş olabilir.

Kondal ve diğ. (2022), Diz radyografilerini KL ölçeğinde otomatik olarak derecelendirmek için bir CNN tabanlı yaklaşım sunmuştur. Bireysel dizleri izole etmek için bir nesne tespit modeli ve her bir dize KL ölçeğini atamak için bir regresyon modeli kullanmışlardır. İnce ayar sonrasında ortalama mutlak hata düşüşü elde etmişlerdir. (1,09'dan 0,28'e).

Tolpadi ve diğ. (2020), OA lezyonlarını tespit etmek için derin Siyam CNN ve geliştirilmiş ResNet-34 tabanlı yarı otomatik bir model tasarlamıştır. Modelin ortalama çok sınıflı doğruluğu %61 olarak belirlenmiş ve KL-1 ile KL-2 seviyelerinin, KL-0, KL-3 ve KL-4'e kıyasla daha düşük başarı gösterdiği tespit edilmiştir.

Wang ve diğ. (2022), DNN'leri kullanarak otomatik olarak diz osteoartriti görüntülerini sınıflandırmayı önermişlerdir. İlk olarak görüntüler ön işlemden geçirilmiş, ardından VGG ağı kullanılarak dizin orta noktası çıkarılmıştır. Görüntüler daha sonra ResNet-50 ağı kullanılarak sınıflandırılmıştır. Yazarlar ayrıca veri setini eklemeli bir arama yöntemi uygulayarak dengelemişlerdir. Yukarıda bahsedilen tüm teknikleri uygulayarak %81,41 sınıflandırma doğruluğuna ulaşmışlardır.

Tiulpin ve diğ. (2018), KL derecelendirme ölçeğine dayalı bir diz osteoartriti teşhis sistemi sunmuşlardır. Yöntemleri, derin Siyami evrişimli sinir ağını kullanmaktadır. Modelleri, çok merkezli Osteoartrit çalışmasından elde edilen veri seti üzerinde eğitilmiş ve OAI'den alınan 5960 diz görüntüsü ile doğrulanmıştır. Modelleri, ortalama sınıflandırma doğruluğunu %66,71 ve kare Kappa değerini 0,83 olarak elde etmiştir. Ayrıca, ağ kararlarını etkileyen görüntü özelliklerine vurgu yapmışlardır. Bu detaylar, bu otomatik yöntemleri daha ileri geliştirmeyi amaçlayan diğer kişiler için süreci daha verimli hale getirmektedir.

Jakaite ve diğ. (2021), makine öğrenmesi tekniklerini kullanarak küçük bir veri setinin nasıl etkili bir şekilde erken aşamada KOA tespiti için kullanılabileceğini araştırmışlardır. Yapılan çalışma, yapay sinir ağları, destek vektör makinesi ve rastgele orman gibi makine öğrenmesi tekniklerini, kendi geliştirdikleri grup yöntemini içeren önerdikleri yaklaşımla karşılaştırarak gerçekleştirmiştir. Kullanılan veri seti yalnızca 31 görüntü içermektedir ve ciddiye etiketleri KL derecelendirme sistemine dayanmaktadır. Zernike tabanlı doku özellikleri kullanarak, önerdikleri çerçeve, %85,0 doğruluk oranı ile radyoloji testlerinde kesinlikle bir artış sağlamıştır. Ancak, büyük sayıda hasta durumu içeren durumlarda yeni çalışmaların yapılması gerekebilir.

Tiulpin ve diğ. (2019) tarafından yapısal OA ilerlemesini öngörmek için ham radyografik veri ve klinik bilgi kullanarak makine öğrenmesi tabanlı bir yaklaşım önerilmiştir. OA ilerlemesinin olasılığını değerlendirmek için bir CNN kullanılmış ve tahminleri klinik verilerle birleştirmek için bir Gradyan Artırmalı Makine Öğrenmesi kullanılmıştır. %62'lik bir hassasiyet ve %82'lik bir ROC eğrisinin altında kalan alan değeri elde edilmiştir.

Pi ve diğ. (2023), diz röntgen görüntülerini 5 sınıfa ayırmak için bir topluluk tabanlı model mimarisi sunmaktadır. DenseNet-161, EfficientNet-b5, EfficientNet-V2-s, RegNetY-8GF, ResNet-101, ResNext-50-32x4d, Wide-ResNet-50-2 ve QShuffleNet-V2-x2-0 gibi gelişmiş görüntü sınıflandırma modelleri birleştirilerek bir topluluk ağı oluşturulmuştur. Görüntüler, yeniden boyutlandırılmış ve 4x4 komşu pikseller kullanılarak işlenmiştir. Eğitim sürecini geliştirmek için görüntüler her epoch sırasında yatay olarak çevrilmiş ve kaynak görüntülerden minimum 1, maksimum 3 görüntü türetilmiştir. Her model belirlenen parametrelere göre eğitilmiş ve performansı test etmek için Test Zamanı Artırımı tekniği uygulanmıştır. Temel model ve optimize model olmak üzere iki model oluşturulmuştur. Optimize model, doğruluk oranı %70'in üzerinde olan bireysel modellerin birleştirilmesiyle oluşturulmuştur. Topluluk oyu için en çok tahmin edilen sınıfı seçme, modellerin tahminlerine ağırlık vererek seçim yapma ve karışık oylama yöntemleri kullanılmıştır. Sonuçlara göre temel model %74,21, optimize model ise %76,33 doğruluk oranı elde etmiştir. Bu yöntem, birden fazla modelin entegrasyonu ile diz röntgeni sınıflandırmasında daha yüksek doğruluk oranlarına ulaşmayı amaçlamaktadır.

Zhang ve diğ. (2020), diz osteoartritinin şiddetini sınıflandırmak için otomatik çok sınıflı bir CNN yaklaşımını tanıtmaktadır. Çalışma, daha önceki derin öğrenme yöntemlerini geliştirmeyi amaçlamaktadır ve iki ana yaklaşım içerir: Res-Nets ve Konvolüsyonel Blok Dikkat Modülü. Çalışmada, OAI veri seti kullanılmış olup, bu veri seti, bilateral diz radyografileri ve karşılık gelen KL dereceleri (0'dan 4'e kadar) içermektedir. Önerilen model, %74,81 çoklu sınıf ortalama doğruluk oranı, 0,36 ortalama karesel hata ve 0,88 kuadratik kappa skoru elde etmiştir. Bu doğruluk oranı, yazarların önceki çalışmalarında elde edilen %64,3 seviyesinden %67,49 seviyesine yükseltilmiştir.

Chen ve diğ. (2019), KOA şiddetini sınıflandırmak için tamamen otomatik bir yaklaşım sunmaktadır. Erken tespit, KOA'nın ilerlemesini yavaşlatmada kritik bir rol oynar ve KL derecelendirme sistemi, KOA şiddetinin değerlendirilmesinde yaygın

olarak kullanılır. Bu yöntem, OAI veri seti ile eğitilmiş, doğrulama ve test süreçlerinde uygulanmıştır. İlk olarak, diz eklem bölgelerini X-ray görüntülerinde tespit etmek için özelleştirilmiş bir YOLOv2 nesne tespit ağı kullanılmıştır. Ardından, ResNet, VGG, DenseNet ve InceptionV3 gibi CNN mimarileri, tespit edilen diz eklem görüntülerini 0-4 arasındaki beş KL derecesine sınıflandırmak için ince ayar yapılmıştır. Önerilen YOLOv2 modeli, 0,922'lik en yüksek duyarlılık ve 0,858'lik ortalama Intersection over Union ile 0,0105 saniye işlem süresi elde etmiştir. Bu model, daha önceki yöntemlerle kıyaslandığında üstün doğruluk ve verimlilik göstermektedir.

Gupta ve diğ. (2023), Osteoartritli diz X-ray görüntülerini, hastalığın şiddet seviyelerine göre 0'dan 4'e kadar olan 5 sınıfa ayırmayı amaçlamaktadır. X-ray görüntüleri, diz eklemine ait olup, her bir görüntü hastalığın şiddetine göre 0 (Osteoartrit yok) ile 4 (şiddetli Osteoartrit) arasında sınıflandırılır. Veri seti, OAI veri setinden alınmıştır ve 45-79 yaş arasındaki 4796 hastaya ait diz görüntülerini içermektedir. Görüntüler, tıbbi uzmanlar tarafından doğru bir şekilde sınıflandırılmıştır. Model için veri seti, eğitim, test ve doğrulama için %70, %10, %20 oranında rastgele bölünmüştür. Veri artırma teknikleri kullanılarak görüntülerdeki dengesizlikler giderilmiştir; bunlar arasında yatay çevirme, küçük açıyla döndürme, doygunluk ve parlaklık ayarlamaları bulunmaktadır. Tüm görüntülerin final çözünürlüğü 224x224 piksel olarak belirlenmiştir. Çeşitli CNN mimarileri, doğru sınıflandırma için en iyi doğruluğu elde etmek amacıyla uygulanmıştır. ResNet-34, DenseNet-121, VGG-19 ve Inception-V3 gibi modeller denenmiştir. Her bir mimarinin katmanına, evrişim blok dikkat modülü adı verilen özel bir modül eklenmiştir. Tüm modeller arasında, DenseNet-121 mimarisi bu katmanlarıyla en yüksek doğruluğu elde etmiştir. Bu mimari, 3 geçiş bloğu ve her bloğunda konvolüsyonel katman, batch normalizasyon, ReLU aktivasyon fonksiyonu, ortalama havuzlama ve filtre azaltma içerir. DenseNet-121, evrişim blok dikkat modülü katmanlarıyla 5 çıkış düğümü olan bir dense katmanına sahip olup, sınıflandırma için softmax aktivasyon fonksiyonu kullanmaktadır. 0 ve 1 numaralı sınıflar birleştirilerek doğruluk artırılmıştır. Sonuç olarak, DenseNet-121 mimarisi %86.714 doğrulukla en iyi sonucu vermiştir. Evrişim blok dikkat modülü katmanlarını eklemek doğruluğu yaklaşık %3, 0 ve 1'in birleştirilmesi ise doğruluğu yaklaşık %11,34 artırmıştır.

Cueva ve diğ. (2022), diz osteoartritinin tespiti ve sınıflandırılması için bir bilgisayarla yardımcı tanı sistemi sunmaktadır. Önerilen model, paralel çalışan ve birbirine bağlı iki aynı sinir ağından oluşan Deep Siamese konvolüsyonel sinir ağlarıdır.

Bu ağlar, her biri üzerine yığılan ince ayar yapılmış ResNet-34 katmanları ile iki dizin lezyonlarını aynı anda tespit etmeye çalışmaktadır. Çalışmanın amacı, derin öğrenme modeline dengeli bir veri seti sağlamaktır ve bu sorun transfer öğrenme ile çözülmüştür. Eğitim, kamuya açık veri setleri üzerinde yapılmış ve doğrulama özel veri setleriyle gerçekleştirilmiştir. Model, çoklu sınıf doğruluğu olarak %61 başarı elde etmiştir. KL derecelendirme sınıflarındaki dengesizliği iyileştirmek amacıyla, azınlık sınıflarının örnekleri çoğaltılmış ve veri artırma teknikleri ile yeni örnekler oluşturulmuştur. Bu dönüşümler arasında, -7 ile +7 arasında rastgele döndürme, renk değişimleri, parlaklık, kontrast ve doygunluk varyasyonları yer almaktadır. Kullanılan konvolüsyonel sinir ağı, önceden eğitilmiş ResNet-34 modelidir ve ardından hiperparametre ayarları yapılmıştır. Orijinal model, 1000 sınıf içeren ImageNet veritabanı üzerinde önceden eğitilmiş olup, çıkış katmanı 5 KL derecelendirme sınıfına uyacak şekilde değiştirilmiştir. Her bir görüntü, medial bölgeden sol ve sağ eklem olarak kesilir ve biri diğerine simetrik olacak şekilde çevrilir, her iki görüntü Siamese ağlarına paralel işlenmek üzere beslenir. Her Siamese ağında, bir konvolüsyon katmanı, ardından dört ResNet katmanı, normalizasyon katmanı, maksimum havuzlama katmanı ve üç tam bağlantılı katman bulunur. Son olarak, her KL derecelendirme sınıfı için olasılıkları çıktısı olarak veren SoftMax katmanı bulunur. Yazarlar, modeli özel klinik veri seti üzerinde test etmiş ve değerlendirmiştir. Model, ortalama çoklu sınıf sınıflandırmasında %61,71 başarı elde etmiştir ve özellikle KL grade 3 ve KL grade 4 sınıflarını doğru sınıflandırmada iyi performans gösterirken, KL grade 1 ve KL grade 2 sınıflarını sınıflandırmada zayıf sonuçlar elde etmiştir.

Yong ve diğ. (2022), X-ray görüntülerinin KL derecelendirme sınıflarına sınıflandırılması için yenilikçi bir ordinal regresyon modülü fonksiyonu kullanmışlardır. Bu yaklaşımda, her etiketin nominal bir tür olarak ele alınmasının yerine, KL derecelendirmesinin sıralaması da dikkate alınmıştır. Bu fonksiyon, doğru etikete daha yakın olan yanlış sınıflandırmaları diğerlerinden daha fazla ceza ile değerlendirir. Ayrıca, bu çalışmada bir kümülatif bağlantı fonksiyonu da kullanılmıştır. Yazarlar, 6 önemli sınıflandırma sinir ağı mimarisinden 23 farklı sinir ağı ile karşılaştırmalı bir çalışma yapmışlardır. Çalışma, OAI veri seti kullanılarak yapılmıştır ve bu veri seti 4796 katılımcıyı içermektedir. Tüm radyografiler, sinir ağına beslenmeden önce ön işleme tabii tutulmuştur. Ön işleme teknikleri arasında görüntülerin 0,14 mm/piksellik boyutlandırılması ve X-ray görüntülerinin diz eklemi kısmının kırılması yer alır. İşlenmiş görüntüler daha sonra popüler önceden eğitilmiş sınıflandırma modellerine

gönderilmiştir ve son katmanın çıktısı, etiketlerin sırasını dikkate alarak yanlış sınıflandırmaları cezalandıran ordinal regresyon modülü katmanına gönderilmiştir. Yazarlar, tüm önceden eğitilmiş modeller için standart hiperparametreler kullanmış ve 30 epoch boyunca SGD ile öğrenme oranı 0,001 ve momentum 0,9 olacak şekilde eğitilmiştir. Kullanılan batch boyutu ise 32'dir. Modeller arasında, DenseNet-161 katmanlı model, en düşük hata değeri olan 0,330 ile birinci sırada yer alırken, VGG-16bn 0,338 hata ile ikinci ve VGG-19 0,345 MAE ile üçüncü sıradadır. Doğruluk, Makro ve Kuadratik Ağırlıklı Kappa açısından ise DenseNet-161, sırasıyla %88,09 ve 0,8609 ile en yüksek değerlere ulaşmıştır.

Fatema ve diğ. (2023), çalışmalarında, diz röntgen görüntülerini kullanarak Diz Osteoartriti (KOA) sınıflandırması için optimal özellikleri ve bu özelliklere dayalı bir karar verme sistemi belirlemeyi amaçlamışlardır. Derin öğrenme ve makine öğrenimi yöntemleri kullanılmadan, yalnızca tıbbi özelliklere dayalı olarak KOA derecelerini sınıflandırmayı hedeflemiştir. Çalışmada, optimal segmentasyon yöntemi önerilmiş, belirgin tıbbi özellikler belirlenmiş ve bu özelliklere dayalı bir karar sistemi geliştirilmiştir. İki farklı veri seti birleştirilmiş, 711 aykırı değer çıkarılmış ve önerilen sistemin test edilmesi için 30 görüntü ayrılmıştır. Altı farklı özellik seti oluşturulmuş ve XGB sınıflandırıcısı, önerilen 12 özellik ile %96,64 doğruluk elde etmiştir. Özellik seçimi teknikleri uygulanarak, altı optimal mesafe özelliği ile model doğruluğu %99,46'ya yükseltilmiştir. Bu özelliklere dayalı karar sistemi, KOA derecelerini etkili bir şekilde belirlemiş ve tıbbi alanda önemli bir etki yaratacak potansiyele sahip olduğu görülmüştür. Gelecekte, veri seti genişletilerek gerçek zamanlı veri ile yöntemlerin test edilmesi ve KOA ilerlemesini anlamak için geometrik derin öğrenme ve grafik sinir ağlarının araştırılması planlanmaktadır.

Wang ve diğ. (2023), daha ayrıntılı bir değerlendirme yapmak amacıyla, hastalığı beş farklı evreye ayıran sınıflandırma modeli oluşturulmuştur. Beş sınıftan oluşan Veri Seti I üzerinde yapılan deneyler sonucunda, ResNet101 modeli %69 doğruluk oranı ile en iyi sonucu vermiştir. Bu çalışma, KOA'nın evrelerinin ayrıntılı bir şekilde belirlenmesinde derin öğrenme modellerinin potansiyelini ortaya koyarak, otomatik teşhis sistemlerinin geliştirilmesine önemli katkılar sunmaktadır.

Solak (2024), KOA derecelendirmesinde beş sınıflı ayrımı da uygulamıştır. Çalışmada, KL sistemine göre bireyler, derecelerine bağlı olarak "0: sağlıklı", "1: şüpheli", "2: hafif KOA", "3: orta KOA" ve "4: ağır KOA" olmak üzere beş ayrı sınıfa ayrılmıştır. Bu detaylı ayırım, KOA'nın ilerleyici doğasına uygun şekilde hastalık

evrelerinin daha hassas biçimde tespit edilmesini ve böylece bireylere uygun izlem ve tedavi stratejilerinin geliştirilmesini amaçlamaktadır. Sınıflandırma sürecinde, ResNet-50, Xception, VGG16, EfficientNetB0 ve DenseNet201 gibi transfer öğrenme modelleri kullanılmış; her bir model rmsprop, sgdm ve adam optimizasyon algoritmalarıyla eğitilerek performansları karşılaştırılmıştır. Elde edilen sonuçlar, Xception modelinin rmsprop algoritması ile %67,8 doğruluk, %68,8 F1-skoru ve 0,55 Cohen's Kappa değeri elde ederek beşli sınıflamada en yüksek başarıyı gösterdiğini ortaya koymuştur. Bu bulgular, Solak'ın (2024) çalışmasında KOA'nın farklı evrelerinin otomatik olarak ayırtmada orta düzey bir performans sergilediğini göstermektedir. Özellikle erken evreler arasındaki sınırların belirgin olmamasından kaynaklanan sınıflandırma zorlukları yaşandığı belirtilmiştir. Bu durum, KOA derecelendirmesinde çok sınıflı ayrımların daha karmaşık bir problem olduğunu ve bu alanda doğruluk oranlarını artırmak için ileri tekniklerin uygulanabileceği yeni çalışmalara ihtiyaç duyulduğunu göstermektedir.

Mahfus ve diğ. (2025) çalışmalarında, KOA derecelendirmesi için X-ışını görüntülerinde geliştirilmiş bir yaklaşım önermiştir. Çalışmada, görüntülerin kontrastını artırmak amacıyla CLAHE (Contrast Limited Adaptive Histogram Equalization) ve gürültüyü azaltmak için Bilateral Filtreleme yöntemleri uygulanmıştır. Özellik çıkarımı sürecinde, ResNet101 tabanlı bir transfer öğrenme modeli kullanılarak derin özellikler elde edilmiştir. Elde edilen özellikler üzerinde temel bileşenler analizi ile boyut indirgeme yapılmış ve nihai sınıflandırma aşamasında tam bağlantılı katmanlar içeren bir sinir ağı kullanılmıştır. 5 sınıflı KOA derecelendirmesinde önerilen yöntem, test verisinde %97,85 doğruluk oranı ile literatürdeki birçok çalışmanın üzerinde performans göstermiştir. Elde edilen yüksek doğruluk, kullanılan gelişmiş ön işleme adımları ve derin özellik çıkarımının etkili bir şekilde bütünleştirilmesiyle ilişkilendirilmiştir.

2.3. Füzyon Stratejileri (Fusion Strategies)

Füzyon stratejileri, çeşitli veri kaynaklarından veya farklı özellik çıkarma teknikleriyle elde edilen bilgilerin birleştirilmesiyle sınıflandırma ve tanı modellerinin performansını artırmayı amaçlayan bir yaklaşımdır. Bu stratejiler, yalnızca KOA teşhisinde değil, görüntü işleme ve sınıflandırma alanındaki pek çok çalışmada başarıyla uygulanmıştır. Farklı özelliklerin veya farklı veri türlerinin (örneğin, görüntü verisi ve

sinyal verisi) kombinasyonu, daha güçlü ve güvenilir sonuçlar elde edilmesini sağlamaktadır. Bu başlık altında, füzyon stratejilerinin uygulanışını ve bu yaklaşımların sınıflandırma performansına olan etkisini inceleyen çalışmalara yer verilmektedir.

Füzyonun önemi, karmaşık problemlerin çözümünde daha kapsamlı bilgi setleri sunarak hem genel doğruluk oranlarını artırması hem de modelin genellenebilirliğini geliştirmesinden kaynaklanmaktadır.

Chaugule ve arkadaşları (2023), KOA'nın erken ve doğru şekilde sınıflandırılmasının röntgen görüntülerinden otomatik olarak gerçekleştirilmesi gerektiğini vurgulamaktadır. Bu doğrultuda, KOA'nın şiddetini KL derecelendirme sistemine göre belirlemek amacıyla, özellik çıkarımı ve füzyonuna dayalı bir Derin Konvolüsyonel Sinir Ağı modeli önermektedir. Modelin işleyişi dört aşamadan oluşmaktadır. İlk olarak, görüntülerin kalitesini artırmak ve gürültüyü gidermek için bir otomatik kodlayıcı kullanılmıştır. Ardından, diz bölgesinin daha iyi analiz edilebilmesi için segmentasyon işlemi uygulanmıştır. Özellik çıkarımı aşamasında, bölge tabanlı, Zernike momentleri, dalgacık dönüşümleri ve Haralick metrikleri olmak üzere dört farklı teknik kullanılarak özellikler çıkarılmış ve bu özellikler birleştirilerek daha güçlü bir temsil oluşturulmuştur. Son aşamada ise, bu birleştirilmiş özellik kümesi, KOA şiddet seviyelerini belirlemek üzere eğitilmiş modele giriş olarak verilmiştir. Model, Adam optimizasyon algoritması ile eğitilmiş ve farklı değerlendirme metrikleri ile test edilmiştir. Sonuçlar, önerilen modelin doğrulama aşamasında %95,70 ve test aşamasında %96,31 doğruluk elde ettiğini göstermektedir. Özellikle, farklı özellik çıkarım tekniklerinin bir araya getirilmesiyle oluşturulan füzyon mekanizması, KOA'nın şiddetinin belirlenmesinde daha kapsamlı bir veri temsili sağlayarak önceki modellerden daha başarılı bir performans ortaya koymuştur.

Guida ve diğ. (2023) KOA sınıflandırması için X-ray'e dayalı çok modlu ara füzyon kullanan derin öğrenmeye dayalı bir yöntem sunmuşlardır. Bu çalışmada uygulanan derin öğrenme yaklaşımını oluşturmak için MRI, X-ray ve hastanın klinik bilgilerine dayalı bir veri seti kullanılmıştır. KOA'yı sınıflandırmak için önerilen yöntem CNN olmuştur. Çalışma sonuçları, uygulanan tekniğin %76 performans doğruluk oranı elde ettiğini göstermiştir, ancak bu oran düşüktür ve daha fazla iyileştirme gereklidir.

Alammar ve diğ. (2023), tıbbi görüntü sınıflandırmasında etiketli veri eksikliğinin büyük bir engel oluşturduğunu ve transfer öğrenmenin bu sorunu hafifletmek için kritik bir rol oynadığını belirtmektedir. Bu çalışmada, geleneksel

transfer öğrenme yöntemlerinin sınırlamalarını aşmak için, hedef veri setine benzer tıbbi görüntüler üzerinde önceden eğitilmiş modeller kullanılmış ve daha sonra küçük etiketli veri setleri ile ince ayar yapılmıştır. MURA veri setinden humerus ve bilek X-ray görüntüleri üzerinde test edilen yöntem, özellik çıkarımı için transfer öğrenme modellerini kullanarak çeşitli makine öğrenmesi sınıflandırıcılarını eğitmiştir. Özellikle, özellik füzyonu sürecinde, farklı transfer öğrenme modellerinden çıkarılan özellikler birleştirilerek daha kapsamlı bir veri temsili elde edilmiştir. Bu yaklaşım, humerus sınıflandırmasında %87,85 doğruluk ve %87,63 F1 puanı, bilek sınıflandırmasında ise %85,58 doğruluk ve %82,70 F1 puanı sağlayarak ImageNet transfer öğrenme ile eğitilen modellerden üstün performans göstermiştir. Ayrıca, Grad-CAM ve LIME görselleştirme teknikleri, önerilen yöntemin daha güçlü ve güvenilir bir karar mekanizmasına sahip olduğunu doğrulamıştır. Çalışma, önerilen transfer öğrenme ve özellik füzyon yönteminin sadece belirli görevlerle sınırlı olmadığını, farklı tıbbi görüntüleme senaryolarında da uygulanabileceğini göstermektedir.

Alzubaidi ve diğ. (2023), derin öğrenme modellerinin sınırlı eğitim verisi nedeniyle kötü genelleme yapabildiğini ve aşırı uyum riski taşıdığını belirtmektedir. Bu sorun, özellikle X-ray görüntülerinde önkol anormalliklerini tespit etmek için büyük bir engel oluşturmaktadır. Çalışmada, geleneksel ImageNet transfer öğrenme yöntemlerinin gri tonlamalı tıbbi görüntüler için yetersiz kaldığı vurgulanmış ve bu sorunu çözmek amacıyla, alan adaptasyonu tabanlı yeni bir transfer öğrenme yaklaşımı önerilmiştir. Bu yaklaşımda, farklı vücut bölgelerinden alınan X-ray görüntüleri üzerinde önceden eğitilmiş altı farklı model, hedef veri setindeki önkol X-ray görüntülerine ince ayar yapılarak adapte edilmiştir. Buna ek olarak, özellik füzyonu tekniği kullanılarak, farklı derin öğrenme modellerinden çıkarılan özellikler birleştirilmiş ve makine öğrenmesi sınıflandırıcıları eğitmek için daha kapsamlı bir veri temsili elde edilmiştir. Model kararlarının görselleştirilmesi için Grad-CAM yöntemi kullanılmış ve modelin hangi bölgeleri dikkate aldığı analiz edilmiştir. Çalışmada önerilen transfer öğrenme ve özellik füzyon yaklaşımı, %90,7 doğruluk, %90,6 F1 puanı ve %81,3 Cohen's Kappa katsayısı ile üstün performans göstermiştir. Sonuçlar, önerilen yöntemin alan uyumsuzluğunu azaltarak transfer öğrenme tabanlı modellere kıyasla daha iyi karar verme mekanizmaları sunduğunu ortaya koymuştur.

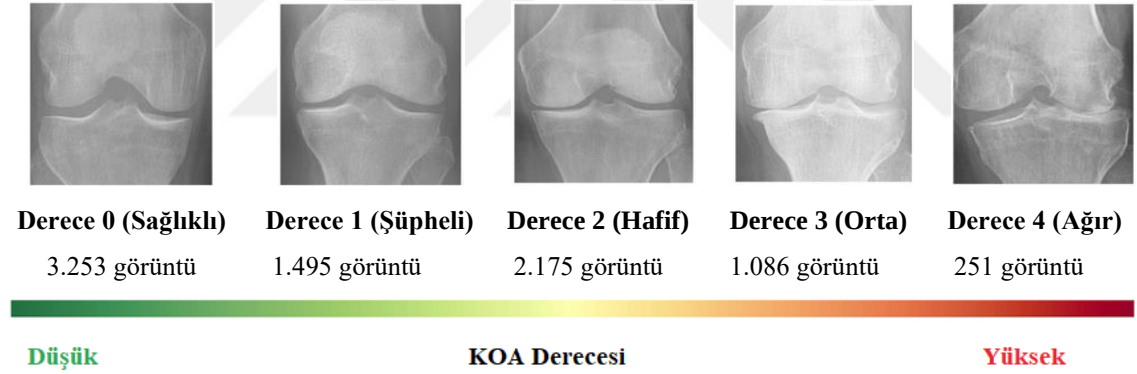
3. MATERYAL VE YÖNTEM

3.1. Veri Seti

Bu çalışmada, KOA derecelendirme süreçlerinde literatürde yaygın olarak kullanılan OAI veri setinden türetilmiş açık erişimli bir diz X-ışını görüntüleme veri setinden faydalanılmıştır. OAI veri seti, KOA'nın ilerleyişini izlemek ve teşhis süreçlerini desteklemek amacıyla geliştirilmiş büyük ölçekli ve uzunlamasına bir klinik çalışmadır. Veri seti, Chen ve arkadaşları (2018) tarafından derlenmiş ve hem “Hugging Face” hem de “Kaggle” platformlarında araştırmacıların kullanımına sunulmuştur.

Söz konusu veri kümesi, 4796 katılımcıya ait sol ve sağ diz X-ışını görüntülerinden oluşmakta olup, toplamda 8260 adet görüntü içermektedir

Görüntüler, diz osteoartritin şiddetini temsil eden KL derecelendirme sistemine göre beş sınıfa ayrılmıştır. Şekil 3.1., KL'nin tüm derecelerine ait bir örneği ve her bir dereceye ait veri setinde bulunan görüntü sayısını göstermektedir (Pi ve diğ. 2023).



Şekil 3.1. Tüm KL derecelerine ait diz eklemi örnekleri (Pi ve diğ. 2023)

3.2. Ön İşleme

Ön işleme aşamasında, diz osteoartriti (KOA) derecelendirmesi amacıyla kullanılan X-ışını görüntülerinin analiz edilebilirliğini standardizasyonu sağlamak için çeşitli işlemler uygulanmıştır.

İlk adımda, tüm görüntüler literatürde yaygın olarak benimsenen yöntemlere uygun şekilde yeniden boyutlandırılmıştır. Muthazhagu ve Surendiran (2024) ile Phan ve diğ. (2025) çalışmalarında da vurgulandığı gibi, KOA derecelendirme çalışmalarında giriş görüntülerinin standart bir boyuta getirilmesi, sınıflandırma sürecinin kararlılığı ve

başarımı açısından kritik öneme sahiptir. Bu çalışmada, tüm diz X-ışını görüntüleri 224×224 piksel boyutuna yeniden ölçeklendirilmiştir.

Ardından, literatürde Mohammed ve diğ. (2023) ile Solak (2024) tarafından önerilen yöntemlerden esinlenilerek, görüntülerin daha anlamlı bir analiz yapılabilmesi amacıyla kırpma işlemi uygulanmıştır. Mohammed ve ark. (2023) çalışmalarında, diz X-ışını görüntülerinin yalnızca ilgili bölgesini (İlgi Bölgesi/Region of Interest - ROI) içerecek şekilde kırılması gerektiği, aksi halde çevresel yapıların modele gereksiz bilgi yükleyebileceği vurgulanmıştır. Benzer şekilde, Solak (2024) çalışmasında da, diz ekleminin merkez bölgesine odaklanmak için kırpma işleminin uygulandığı ve böylece modelin KOA belirtilerinin yer aldığı temel anatomik yapıların öğrenimine odaklanmasının sağlandığı belirtilmiştir. Bu çalışmada da, diz ekleminin bulunduğu ROI bölgesi görüntüden manuel olarak çıkarılmıştır. SOLAK (2024) çalışmasında belirtilen değerler doğrultusunda, her yönden (üst, alt, sağ ve sol) 20 piksel kırpma işlemi uygulanmıştır. Çok fazla kırpma yapılmamasının nedeni, KOA'nın yalnızca eklem aralığıyla sınırlı kalmayıp, çevredeki osteofit oluşumu, subkondral skleroz ve kistler gibi yapısal değişikliklerle birlikte seyretmesidir. Uygulanan kırpma stratejisi, diz eklemi ve çevresindeki patolojik bulguların korunmasını sağlayarak, hastalığın ilerleme düzeyinin daha kapsamlı bir şekilde değerlendirilmesine imkân tanımıştır. Ayrıca, üst ve alt kenarlardan yapılan kırpma ile femur ve tibia üzerindeki deformasyonlar korunmuş, sağ ve sol kenarlardan yapılan kırpma ise görüntüdeki gereksiz siyah alanlar temizlenerek analiz süreci iyileştirilmiştir.

Kırpma işlemini takiben, düşük kontrastlı bölgelerdeki yapısal detayların daha belirgin hale getirilmesi amacıyla CLAHE (Contrast Limited Adaptive Histogram Equalization) yöntemi uygulanmıştır. Ahmed ve Mstafa (2022), Solak (2024) ve Raza ve diğ. (2025) çalışmalarında belirtildiği üzere, CLAHE işlemi sırasında 'NumTiles' parametresi (8,8) olarak belirlenmiştir. Ayrıca 'ClipLimit' değeri de yine literatür çalışmalarında kullanıldığı gibi 0,005 olarak ayarlanmıştır (Ahmed ve Mstafa, 2022; Solak, 2024). Bu sayede, görüntü kontrastı artırılarak KOA'ya özgü yapısal farklılıkların daha iyi görünür hale getirilmesi sağlanmıştır.

Tüm bu ön işleme adımlarının ardından, veri setindeki doğal sınıf dengesizliğini gidermek amacıyla veri artırımı teknikleri ve SMOTE yöntemi uygulanmıştır.

Bu kapsamda, derece 0, 1, 2, 3 ve 4 olarak etiketlenen KOA derecelerinin her biri için sınıf örneklerinin dengelenmesi hedeflenmiştir. Öncelikle, veri artırımı işlemleri ile mevcut görüntüler üzerinde rastgele döndürme (0–30 derece arası), yatay

ve dikey ekseninde kaydırma (yaklaşık %20 oranında) ve simetrik yansıtma (yatay ekseninde) gibi temel dönüşümler uygulanarak sınıflar arası örnek çeşitliliği artırılmıştır. Bu sayede görüntülerin sadece belirli görsel kalıplara göre değil, konum, yön ve perspektif gibi daha geniş bir örneklem yelpazesinde değerlendirilmesi mümkün olmuştur.

Daha sonra, SMOTE algoritması ile az temsil edilen KOA derecelerine ait sentetik örnekler üretilmiş ve her sınıfın örnek sayısı dengelenmiştir. Böylece, hem veri setindeki genel dengesizlik giderilmiş hem de modellerin tüm KOA derecelerini öğrenme kapasitesi artırılmıştır.

Bu adımlar, sınıf dengesizliğinden kaynaklanabilecek model önyargılarını azaltarak, KOA derecelendirmesi sürecinde daha adil ve dengeli bir sınıflandırma performansı elde edilmesine katkı sağlamıştır. Çizelge 3.1, veri artırımı ve dengesizlik giderme işlemleri öncesinde ve sonrasında her KOA derecesi için elde edilen güncellenmiş örnek sayılarını ve toplam veri sayısını özetlemektedir.

Çizelge 3.1. KOA derecelerine göre veri seti: artırım ve dengesizlik giderimi öncesi ve sonrası örnek sayıları

Veri Kümesi (Durum)	Toplam Veri Sayısı	Derece 0 (Sağlıklı)	Derece 1 (Şüpheli)	Derece 2 (Hafif)	Derece 3 (Orta)	Derece 4 (Ağır)
Orijinal Veri (Öncesi)	8260	3253	1495	2175	1086	251
Artırılmış Veri (Sonrası)	13010	2602	2602	2602	2602	2602

3.2.1 Kontrast Sınırlama Adaptif Histogram Eşitleme (Contrast Limited Adaptive Histogram Equalization - CLAHE)

CLAHE, görüntü kontrastını artırmak için kullanılan bir histogram eşitleme tekniğidir. Görüntü üzerindeki kontrastı artırmak ve ön plan/arka plan detaylarını daha belirgin hale getirmek amacıyla kullanılır. Standart histogram eşitleme yöntemlerinde, tüm görüntü üzerinde uygulanan eşitleme, parlaklık artışlarına neden olabilir. CLAHE yöntemi, görüntüyü küçük bölgelere ayırarak lokal kontrastı artırır ve aşırı gürültü oluşumunu önler. Özellikle X-ışını görüntüleri gibi düşük kontrastlı tıbbi verilerde, detayların vurgulanması ve görsel analiz başarısının artırılması amacıyla sıklıkla kullanılmaktadır (Nia ve Shih, 2024).

3.2.2 Sentetik Azınlık Aşırı Örneklemme Tekniği (Synthetic Minority Over-sampling Technique - SMOTE)

Veri setlerinde yer alan sınıflar arasında örnek sayısı bakımından dengesizlik, makine öğrenimi modellerinin performansını olumsuz yönde etkileyebilecek önemli bir sorundur. Bu dengesizlik, özellikle azınlık sınıfların model tarafından yeterince öğrenilememesine yol açabilir. Bu sorunu aşmak için yaygın olarak kullanılan yöntemlerden biri SMOTE olup, bu çalışmada da SMOTE kullanılarak veri setindeki sınıf dengesizliği giderilmiştir.

SMOTE, azınlık sınıflarının örnek sayısını artırmak amacıyla, mevcut örnekler arasındaki benzerlikleri kullanarak sentetik veri örnekleri üretir. Bu yöntem, azınlık sınıfındaki her bir örnek için en yakın komşularını belirleyerek, bu komşular arasındaki farklara dayalı olarak yeni sentetik veriler oluşturur. Böylece, veri setindeki azınlık sınıflarının örnek sayıları artırılarak, sınıf dağılımı daha dengeli hale getirilir (Chawla et al., 2002).

Bu çalışmada, veri setindeki sınıflar arasında belirgin bir dengesizlik gözlemlenmiştir. Örneğin, bazı sınıflar yeterince temsil edilmemekteydi. SMOTE yöntemi ile bu sınıflara ait sentetik veriler üretilmiş ve sınıf dağılımı dengelenmiştir. Özellikle 4 numaralı sınıfın örnek sayısının artırılması, modelin genel doğruluğunu ve genelleme yeteneğini iyileştirmiştir. Ayrıca, bazı sınıfların örnek sayısının oldukça düşük olduğu gözlemlenmiş ve SMOTE uygulaması ile bu sınıflar daha dengeli bir şekilde temsil edilmiştir.

SMOTE yönteminin etkin bir biçimde uygulanabilmesi için bazı önemli parametrelerin dikkatli bir şekilde belirlenmesi gerekmektedir. Bu parametrelerden biri olan `sampling_strategy`, hangi sınıfların ne oranda artırılacağını belirlemektedir. Bu çalışmada kullanılan 'auto' seçeneği, azınlık sınıflarının örnek sayısını, çoğunluk sınıfın örnek sayısına eşit olacak şekilde artırmakta ve böylece sınıflar arasındaki dengesizlik ortadan kaldırılmaktadır. Bu strateji, genellikle tüm sınıfların eşit şekilde temsil edilmesini sağlamak amacıyla tercih edilmektedir.

Bir diğer önemli parametre ise `random_state`'tir. Bu parametre, SMOTE algoritmasının yeniden çalıştırıldığında aynı sonuçları vermesini sağlayarak sürecin tekrarlanabilirliğini güvence altına alır. Belirli bir `random_state` değeri verilmesi, özellikle bilimsel çalışmalarda deneysel sonuçların tutarlılığı açısından büyük önem taşımaktadır. Bu bağlamda, kullanılan parametreler sayesinde SMOTE yöntemiyle

sentetik örnekler istikrarlı ve kontrollü bir şekilde üretilmiş, veri setindeki sınıf dağılımı başarılı bir şekilde dengelenmiştir.

3.3. Geleneksel Görüntü Özelliklerinin Çıkarılması

Füzyon stratejilerinde kullanılmak üzere, X-ışını görüntülerinden çeşitli geleneksel özellikler manuel olarak çıkarılmıştır. Bu tez çalışmasında, aşağıda belirtilen kategoriler altında gruplandırılan toplam 35 özellik değerlendirmeye alınmıştır.

3.3.1. Morfolojik Özellikler

Morfolojik özellikler, bir nesnenin veya yapının şekil, boyut ve geometrik yapısını analiz etmek için kullanılan ölçümler ve metriklerdir. Görüntü işleme ve bilgisayarlı görü alanında, özellikle tıbbi görüntüleme, biyomedikal mühendislik ve nesne tanıma uygulamalarında yaygın olarak kullanılır. Bu özellikler, ROI'nin yapısal özelliklerini sayısal olarak ifade eder ve analiz edilen bölgenin karakteristik yapısını belirlemek için kullanılır.

- **Basıklık (Kurtosis):**

Basıklık, bir veri kümesinin dağılımındaki uç değerlerin yoğunluğunu ve dağılımın tepe noktasının sivrilikliğini ölçen bir istatistiksel özelliktir. Bir dağılımın uç değerlerinin (aşırı büyük veya aşırı küçük değerlerin) ortalama çevresinde ne kadar yoğunlaştığını belirlemek için kullanılır.

Bu özellik, piksel yoğunluklarının uç değerler etrafındaki dağılımını inceleyerek, anormal bölgelerin tespit edilmesine yardımcı olur. Örneğin, tıbbi görüntülerde tümör gibi yapılar, normal dokulara kıyasla farklı basıklık değerlerine sahip olabilir. Denklem 3.1 ile hesaplanır. Bu özellik piksel dağılımının tepe yoğunluğunu tanımlayarak anormal bölgelerin tespitinde kullanılır.

$$\text{Basıklık} = \frac{n \cdot \sum (x_i - \bar{x})^4}{(\sum (x_i - \bar{x})^2)^2} - 3 \quad (3.1)$$

Denklemde;

n : Veri noktalarının toplam sayısını,

x_i : Her bir piksel yoğunluğunu,

$\bar{x}$:Piksel yoğunluklarının ortalamasını ifade etmektedir.

- **Çarpıklık (Skewness):**

Çarpıklık, bir veri kümesinin simetri derecesini ölçen bir istatistiksel metriktir. Pozitif veya negatif olabilir:

Pozitif çarpıklık: Sağ kuyruk uzun, yani büyük yoğunluk değerleri daha sık görülmektedir.

Negatif çarpıklık: Sol kuyruk uzun, yani düşük yoğunluk değerleri daha sık görülmektedir.

Bu özellik, özellikle tıbbi görüntülerde doku farklılıklarını anlamak için kullanılır. Asimetrik yoğunluk dağılımı, hastalıklı veya anormal dokuların varlığına işaret edebilir. Denklem 3.2’de görülen formülle hesaplanan çarpıklık, piksel yoğunluğunun simetrik dağılımından ne kadar saptığını ölçer. Pozitif veya negatif çarpıklık olabilir.

$$\text{Çarpıklık} = \frac{n \cdot \sum (x_i - \bar{x})^3}{(\sum (x_i - \bar{x})^2)^{3/2}} \quad (3.2)$$

Denklem de görülen;

n : Piksel sayısı

x_i : Her bir piksel yoğunluğu,

$\bar{x}$:Piksel yoğunluklarının ortalamasıdır.

Bu özellikle yoğunluk dağılımının asimetri derecesi ölçülür; hasarlı dokularda yoğunluk asimetrik olabilir.

- **Kapsama Oranı (Extent):**

ROI'nin toplam alanının, ROI'yi çevreleyen eksensel hizalanmış en küçük dikdörtgenin (Sınırlayıcı Kutu Alanı - Bounding Box) alanına oranıdır. Bu oran, nesnenin dikdörtgeni ne kadar verimli doldurduğunu gösterir. Nesnenin bu dikdörtgen içerisindeki yer kaplama oranını ifade eder. Yüksek extent değeri, şeklin bounding box'ı iyi doldurduğunu, düşük değer ise nesnenin uzun, ince ya da girintili olduğunu gösterebilir.

$$\text{Kapsama Oranı} = \frac{\text{ROI Alanı}}{\text{Sınırlayıcı Kutu Alanı}} = \frac{\text{ROI Alanı}}{(\max(X) - \min(X)) * (\max(Y) - \min(Y))} \quad (3.3)$$

Denklemden görülen $\max(X)$, ROI'deki en sağdaki noktanın x koordinatını, $\min(X)$, en soldaki noktanın x koordinatını, $\max(Y)$, en üstteki noktanın y koordinatını ve $\min(Y)$, en alttaki noktanın y koordinatını belirtir.

- **Katılık (Solidity):**

Denklemler 3.4'te verilen formülle hesaplanan bu metrik, ROI'nin konveks kabuğa kıyasla ne kadar sıkı ve dolu olduğunu ölçer. Yüksek solidity değerleri, daha kompakt ve sıkışık yapıları ifade ederken; düşük değerler, düzensiz ve girintili-çıkıntılı yapıları işaret eder. Örneğin, tümörler veya diğer anormal yapılar genellikle düşük solidity değerlerine sahiptir, çünkü şekilleri düzensizdir.

$$\text{Katılık} = \frac{\text{ROI Alanı}}{\text{Konveks Kabuğun Alanı}} = \frac{\text{ROI Alanı}}{\frac{1}{2} \sum_{i=1}^n (X_i * Y_{i+1} - X_{i+1} * Y_i)} \quad (3.4)$$

- **Dairesellik (Circularity):**

Dairesellik, bir nesnenin yuvarlaklık derecesini ölçen bir metriktir. Dairesellik değeri 1'e yaklaştıkça nesne ideal bir çembere benzemektedir. Düşük dairesellik değeri, nesnenin daha düzensiz bir yapıya sahip olduğunu gösterir ve Denklem 3.5 ile hesaplanır.

$$\text{Dairesellik} = \frac{4\pi * \text{ROI Alanı}}{(\text{ROI Çevresi})^2} \quad (3.5)$$

Bu özellik yuvarlak yapılar ve düzensiz dokuların ayrımında kullanılır.

- **Büyük Eksen Uzunluğu (Major Axis Length)**

ROI'nin en uzun ekseninin uzunluğunu verir. ROI'nin elipsle eşdeğer modeli kullanılarak hesaplanır. Elipsin ana eksen uzunluğudur. Tümörün uzunluğu veya diğer yapıların boyutlarının belirlenmesinde kullanılır.

- **Küçük Eksen Uzunluğu (Minor Axis Length):**

ROI'nin en kısa ekseninin uzunluğunu verir. ROI'nin elipsle eşdeğer modeli kullanılarak hesaplanır. Elipsin kısa eksen uzunluğudur. Şekil oranlarının (örneğin, tümör en-boy oranı) değerlendirilmesinde kullanılır.

- **Eşdeğer Çap (Equivalent Diameter):**

ROI'nin çevresine eşdeğer bir çemberin çapını tanımlar. Denklem 3.6 ile hesaplanan bu özellik ROI'nin boyutunun sadeleştirilmiş bir göstergesi olarak kullanılır; büyük ölçekli anormallikleri tanımlar.

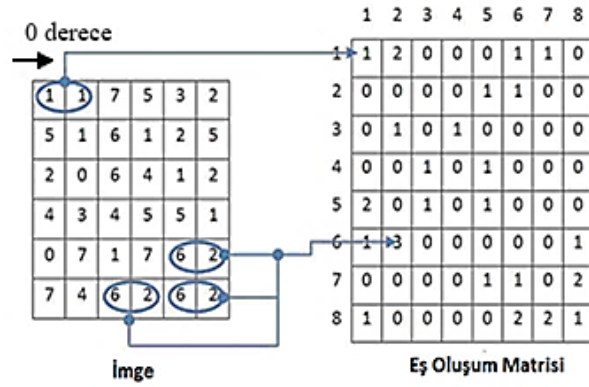
$$\text{Eşdeğer Çap} = \sqrt{\frac{4 * ROI \text{ Alanı}}{\pi}} \quad (3.6)$$

3.3.2 Gri Seviye Eş Oluşum Matrisi (Gray Level Co-occurrence Matrix- GLCM) Özellikleri

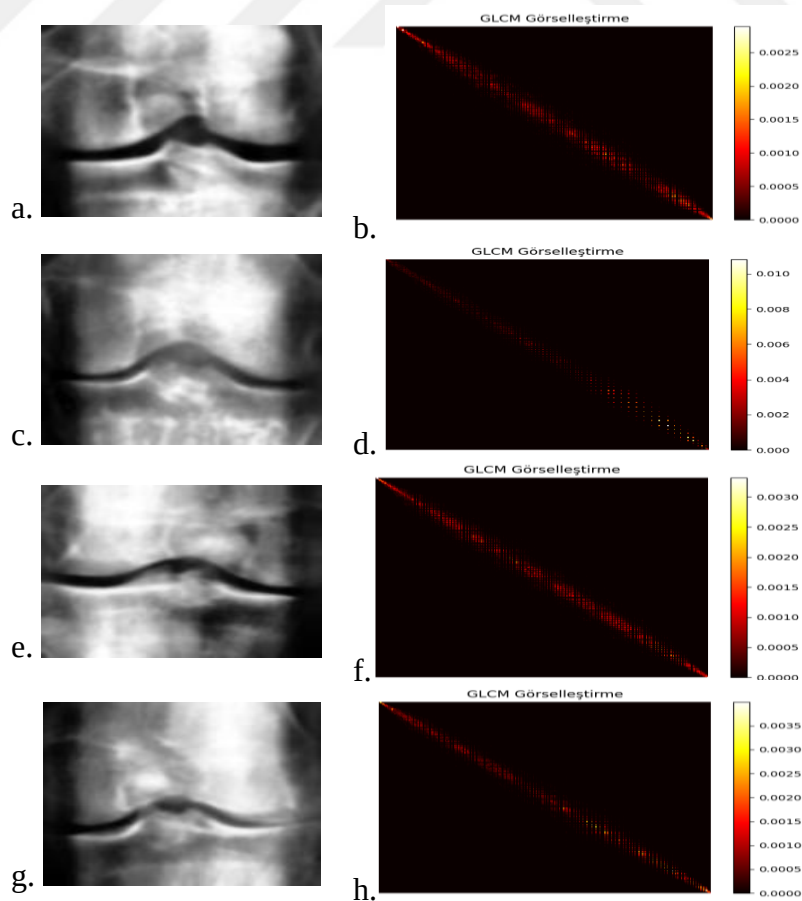
GLCM (Gray Level Co-occurrence Matrix), bir görüntüdeki piksel çiftlerinin yoğunluk değerleri arasındaki mekânsal ilişkileri modelleyen bir matristir ve görüntünün dokusal özelliklerinin çıkarılmasında yaygın olarak kullanılmaktadır. Bu yöntem, M. Haralick (1974) tarafından önerilmiş olup, gri tonlu bir görüntünün dokusal özniteliklerini çıkarmak amacıyla geliştirilmiştir. GLCM, bir görüntüde yan yana bulunan piksel çiftlerinin yoğunluk değerlerinin belirli bir mesafe ve açı doğrultusunda (örneğin, 1 piksel mesafe, 0° açı) birlikte ne sıklıkla tekrarlandığını hesaplar. Matrisin satır ve sütunları, görüntüdeki gri seviyelerini (örneğin, 0-255 arası) temsil ederken, her hücredeki değer, ilgili yoğunluk çiftinin görüntüde kaç kez gözlemlendiğini göstermektedir.

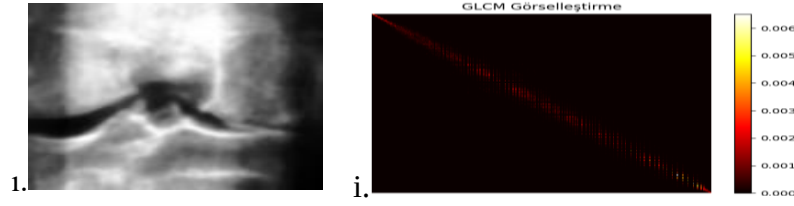
GLCM'in hesaplanmasında, pikseller arası uzaklık (d) ve yön (θ) gibi parametreler dikkate alınır. En yaygın kullanılan yönler 0°, 45°, 90° ve 135° olup, bu yönlerin simetrik benzerleri de analizlerde tercih edilebilir. Örneğin, bir görüntüde d=1 ve θ=0° parametreleriyle hesaplama yapıldığında, (1,1) piksel çifti yalnızca bir kez tekrarlandığı için GLCM'deki (1,1) hücresinin değeri 1 olur. Benzer şekilde, (6,2) piksel çifti üç kez tekrarlandığından, bu çiftin karşılık geldiği hücredeki değer 3 olarak belirlenir. Bu süreç tüm piksel çiftleri için tekrarlanarak, görüntünün dokusal özelliklerini modelleyen bir eş-oluşum matrisi elde edilir.(Ajmera, 2021).

Şekil 3.2, bir görüntüye ait eş oluşum matrisinin (GLCM) elde edilme sürecini göstermektedir (Acar ve Özerdem, 2007). Bu bağlamda, Şekil 3.3'te çalışmada kullanılan veri setinden seçilen beş farklı KOA derecesine ait örnek X-ışını görüntüleri (sol) ve bu görüntülere uygulanan GLCM çıktıları (sağ) birlikte sunulmuştur. Sol sütunlarda yer alan görüntüler, “sağlıklı”, “şüpheli”, “hafif”, “orta” ve “ağır” KOA derecelerini temsil eden ham X-ışını verilerini göstermektedir. Sağ sütunlarda ise bu görüntülerden 1 piksel mesafe ve 0° açı kullanılarak elde edilen GLCM temsilleri yer almaktadır.



Şekil 3.2. Eş oluşum matrisinin elde edilmesi (Acar ve Özerdem, 2007).





Şekil 3.3. Farklı KOA derecelerine ait örnek X-ışını görüntüleri ve bu görüntülere karşılık gelen GLCM çıktıları: (a) Sınıf 0 – Sağlıklı, (b) GLCM çıktısı; (c) Sınıf 1 – Şüpheli, (d) GLCM çıktısı; (e) Sınıf 2 – Hafif, (f) GLCM çıktısı; (g) Sınıf 3 – Orta, (h) GLCM çıktısı; (i) Sınıf 4 – Ağır, (i) GLCM çıktısı.

GLCM matrislerinde X eksenini komşu pikselin, Y eksenini ise merkez (ilk) pikselin gri seviye değerini ifade etmektedir. Matristeki yüksek değerler (sarı tonlar), ilgili gri seviye çiftlerinin daha sık tekrarlandığı konumları göstererek dokusal yoğunluk dağılımını görsel olarak analiz etmeyi mümkün kılmaktadır. GLCM, görüntülerin dokusal yapısını sayısal olarak temsil eden güçlü bir yöntemdir. Bu matrislerden enerji, kontrast, farklılık, homojenlik, korelasyon ve entropi gibi çeşitli dokusal özellikler hesaplanmış ve elde edilen değerler sonraki sınıflandırma aşamalarında temel giriş parametreleri olarak kullanılmıştır.

GLCM ile çıkarılan özellikler;

- **Enerji:**

GLCM'deki elemanların toplam enerjisini ölçer ve görüntüdeki tekdüzelik seviyesini temsil eder (Denklem 3.7). Bu özellik değeri yüksek enerji, daha tekdüze bir doku anlamına gelir.

$$\text{Enerji} = \sum_{i,j} P(i,j)^2 \quad (3.7)$$

Denklemdaki; $P(i,j)$ GLCM' nin i,j konumundaki normalleştirilmiş olasılık değerini göstermektedir.

- **Kontrast:**

Görüntüdeki yoğunluk farklarını ve kenar bölgelerinin belirginliğini ölçer. Denklem 3.8'deki formül ile hesaplanır.

$$\text{Kontrast} = \sum_{i,j} (i - j)^2 P(i,j) \quad (3.8)$$

Burada; $i - j$ piksel çiftleri arasındaki yoğunluk farkını ifade eder. $P(i,j)$ ise GLCM' nin i,j konumundaki normalleştirilmiş olasılık değerini göstermektedir. Yüksek kontrast değeri, yoğunluk farklılıklarının fazla olduğu dokuları ifade etmektedir.

- **Farklılık (Dissimilarity):**

Denklem 3.9'da görüldüğü gibi hesaplanan bu özellik, piksel çiftleri arasındaki yoğunluk farklarının bir ölçüsüdür. Görüntüdeki homojen olmayan bölgeleri belirlemek için kullanılır.

$$\text{Farklılık} = \sum_{i,j} |i - j| P(i, j) \quad (3.9)$$

$|i - j|$: Piksel çiftleri arasındaki gri seviye farkı,

$P(i, j)$: Bu farkın görüntüde ne sıklıkla gerçekleştiğini belirten normalize (olasılık) değeri.

- **Homojenlik:**

Piksel çiftlerinin yoğunluklarının ne kadar benzer olduğunu ölçer. Denklem 3.10'da bu özelliğin nasıl hesaplandığı verilmiştir. Yüksek homojenlik, yoğunlukların birbirine yakın olduğu düzenli dokulara işaret eder.

$$\text{Homojenlik} = \sum_{i,j} \frac{P(i,j)}{1+|i-j|} \quad (3.10)$$

- **Korelasyon:**

Piksel çiftleri arasındaki yoğunluk ilişkisinin derecesini ölçer (Denklem 3.11). Bu özellik dokular arasındaki yoğunluk benzerliğini ölçmek için kullanılır.

$$\text{Korelasyon} = \frac{\sum_{i,j} (i - \mu_i)(j - \mu_j) P(i, j)}{\sigma_i \sigma_j} \quad (3.11)$$

Denklemden görülen;

μ_i, μ_j : Yoğunluk ortalamaları,

σ_i, σ_j : Yoğunluk standart sapmaları

ifade etmektedir.

- **Entropi:**

GLCM entropisi, bir görüntünün gri seviyeler arasındaki ilişkilerinin karmaşıklığını ve rastgeleliğini ölçer. GLCM entropisi, gri seviyelerin birlikte dağılımını ifade eden GLCM matrisinin elemanları üzerinden hesaplanır ve Denklem 3.12 ile gösterilir. Daha yüksek bir GLCM entropisi, daha düzensiz ve karmaşık dokulara işaret eder.

$$\text{Entropi} = - \sum_{i,j} P(i,j) \log(P(i,j)) \quad (3.12)$$

$P(i,j) > 0$ olduğunda entropi hesaplanır.

3.3.3. İstatistiksel Özellikler

İstatistiksel özellikler, bir görüntünün veya veri kümesinin temel yapısını ve dağılımını anlamamıza yardımcı olan matematiksel ölçütlerdir. Görüntü işleme ve veri analizi alanlarında, bir görüntünün içeriğini daha iyi tanımlamak, sınıflandırmak veya karşılaştırmak için bu özellikler kullanılır.

Veri özetleme, büyük veri setleri veya görüntülerin ham halde karmaşık olmasından dolayı istatistiksel özellikler kullanılarak bu verilerin anlamlı sayısal değerlere indirgenmesini ve daha kompakt bir şekilde temsil edilmesini sağlar. Özellik seçimi ve tanımlama, görüntüdeki dokusal veya yapısal farklılıkları belirlemek için kullanılır; örneğin, bir yüzeyin pürüzlü mü yoksa düzgün mü olduğunu belirlemek için farklılık ve homojenlik gibi özellikler değerlendirilir. Makine öğrenimi ve sınıflandırma süreçlerinde, görüntü tanıma, nesne algılama ve tıbbi görüntüleme gibi birçok yapay zeka uygulamasında, görüntülerden elde edilen istatistiksel özellikler model eğitiminde giriş verisi olarak kullanılır. Görüntü karşılaştırma, iki görüntü arasındaki benzerlikleri veya farklılıkları ölçmek için istatistiksel özellikler yardımıyla mesafelerin hesaplanmasını sağlar. Bozukluk veya gürültü analizi ise görüntülerdeki anormallikleri veya bozulmaları tespit etmek için kontrast, entropi ve varyans gibi özelliklerden yararlanır. İstatistiksel özellikler, görüntü analizi, makine öğrenimi, tıbbi teşhis, kalite kontrol ve birçok farklı alanda kritik verilerdir. Görüntülerin temel yapısını anlamak,

desenleri belirlemek ve veri madenciliği yapmak için istatistiksel analiz büyük önem taşır.

İstatistiksel Çıkarılan Özellikler:

- **Enerji:**

İstatistiksel enerji, histogramın karelerinin toplamı olarak hesaplanır. Yani, her piksel değerinin olasılığının karesi alınarak toplam enerji değeri elde edilir. (Denklem 3.13).

$$\text{Enerji} = \sum_{i=0}^{255} P(i)^2 \quad (3.13)$$

Burada $P(i)$, i . gri seviye değerinin normalize edilmiş histogram olasılığıdır. Daha yüksek enerji değeri, görüntüde düşük kontrast ve daha fazla tekrarlanan desenlerin bulunduğunu gösterir.

- **Kontrast:**

İstatistiksel kontrast, görüntünün piksel değerlerinin varyansıdır. Bu, her pikselin yoğunluğunun ne kadar değiştiğini ölçer. Denklem 3.14'deki formül ile hesaplanır.

$$\text{Kontrast} = \text{Var}(I) = \frac{1}{N} \sum_{i=0}^N (I_i - \mu)^2 \quad (3.14)$$

Burada, I_i görüntüdeki her pikselin değeri, μ ise görüntüdeki ortalama piksel değeridir. Yüksek kontrast değeri, görüntüdeki piksel yoğunluklarının birbirinden çok farklı olduğu dokuları ifade eder.

- **Farklılık (Dissimilarity):**

Denklem 3.15'da görüldüğü gibi hesaplanan bu özellik, İstatistiksel farklılık, görüntüdeki her pikselin ortalama değeri ile olan farklarının mutlak değerinin ortalamasıdır. Bu, görüntüdeki homojen olmayan bölgeleri belirlemek için kullanılır.

$$\text{Farklılık} = \frac{1}{N} \sum_{i=0}^N |I_i - \mu| \quad (3.15)$$

Burada, I_i her pikselin yoğunluğunu, μ ise ortalama piksel yoğunluğunu ifade eder. Bu metrik, görüntüdeki homojen olmayan bölgeleri belirlemek için kullanılır.

- **Homojenlik:**

İstatistiksel homojenlik, her gri seviye için histogramın değerinin, o gri seviyenin konumuyla olan ters orantılı bir şekilde normalize edilmesidir. Bu, piksel çiftlerinin ne kadar benzer olduğunu ölçer. Denklem 3.16'da bu özelliğin nasıl hesaplandığı verilmiştir.

$$\text{Homojenlik} = \sum_{i=0}^{255} \frac{P(i)}{1+i} \quad (3.16)$$

Burada $P(i)$, gri seviye histogram olasılık değeridir. Yüksek homojenlik değeri, görüntüde benzer yoğunluk seviyelerinin daha fazla olduğunu gösterir.

- **Korelasyon:**

İstatistiksel korelasyon, görüntüdeki piksel değerlerinin ortalamadan ne kadar saptığını ve bu sapmaların birbirleriyle nasıl ilişkili olduğunu ölçer (Denklem 3.17).

$$\text{Korelasyon} = \frac{\sum_{i=0}^N (I_i - \mu)(I_j - \mu)}{\sigma_i \sigma_j} \quad (3.17)$$

Denklemden görülen;

I_i, I_j : İlgili piksellerin yoğunluk değerleri,

μ : Piksel yoğunluklarının ortalaması,

σ_i, σ_j : Piksel yoğunluklarının standart sapmasını

ifade etmektedir.

Pozitif korelasyon, yoğunluk değerlerinin birbirine benzer değiştiğini gösterirken, negatif korelasyon büyük değişimlerin olduğunu gösterir.

- **Entropi**

Entropi, histogramın elemanlarının logaritması alınarak hesaplanır. Görüntüdeki gri seviyeler arasındaki ilişkilerin rastgeleliğini ölçer (Denklem 3.18).

$$\text{Entorpi} = - \sum_{i=0}^{255} P(i) \log_2(P(i) + \epsilon) \quad (3.18)$$

Denklemde görülen;

P(i), gri seviye histogram olasılığıdır.

ϵ çok küçük bir sayıdır (örneğin, 10^{-10}), sıfır bölme hatasını önlemek için eklenmiştir

ifade etmektedir.

Yüksek entropi değeri, daha karmaşık ve düzensiz dokulara sahip görüntüleri ifade eder.

- **Ortalama (Mean):**

Denklem 3.19'te görülen formül ile hesaplanan ortalama, piksel yoğunluklarının aritmetik ortalamasıdır.

$$\text{Ortalama} = \frac{1}{N} \sum_{i=0}^N I_i \quad (3.19)$$

Denklemde I_i : görüntüdeki her pikselin değeri ifade etmektedir.

- **Varyans:**

Varyans, görüntüdeki her pikselin değerinin ortalamadan sapmalarının karesinin ortalamasını alır (Denklem 3.20). Görüntünün ne kadar çeşitlilik içerdiğini ve homojenlik seviyesini ölçer.

$$\text{Varyans} = \frac{1}{N} \sum_{i=0}^N (I_i - \mu)^2 \quad (3.20)$$

Denklemde, I_i görüntüdeki her pikselin değeri, μ ise görüntüdeki ortalama piksel değeridir

3.3.4. Yönlendirilmiş Gradyanların Histogramı (Histogram of Oriented Gradients - HOG):

HOG, bir görüntünün kenar ve yön bilgilerini analiz eden bir özellik çıkarma yöntemidir. HOG, özellikle nesne tespiti ve sınıflandırma için yaygın olarak kullanılır. Temel prensibi, görüntünün gradyan bilgilerini hesaplayarak bu bilgileri histogramlarla temsil etmektir (Dalal, 2005).

Histogram, verilerin belirli aralıklara (binlere) bölünerek frekanslarını gösteren bir grafikdir. Bir görüntü histogramı, piksellerin parlaklık veya renk değerlerine göre dağılımını ifade eder.

Gradyan ise bir görüntünün yoğunluk değişimlerini belirlemek için kullanılan bir operatördür. Gradyan, şu iki bileşeni kapsar:

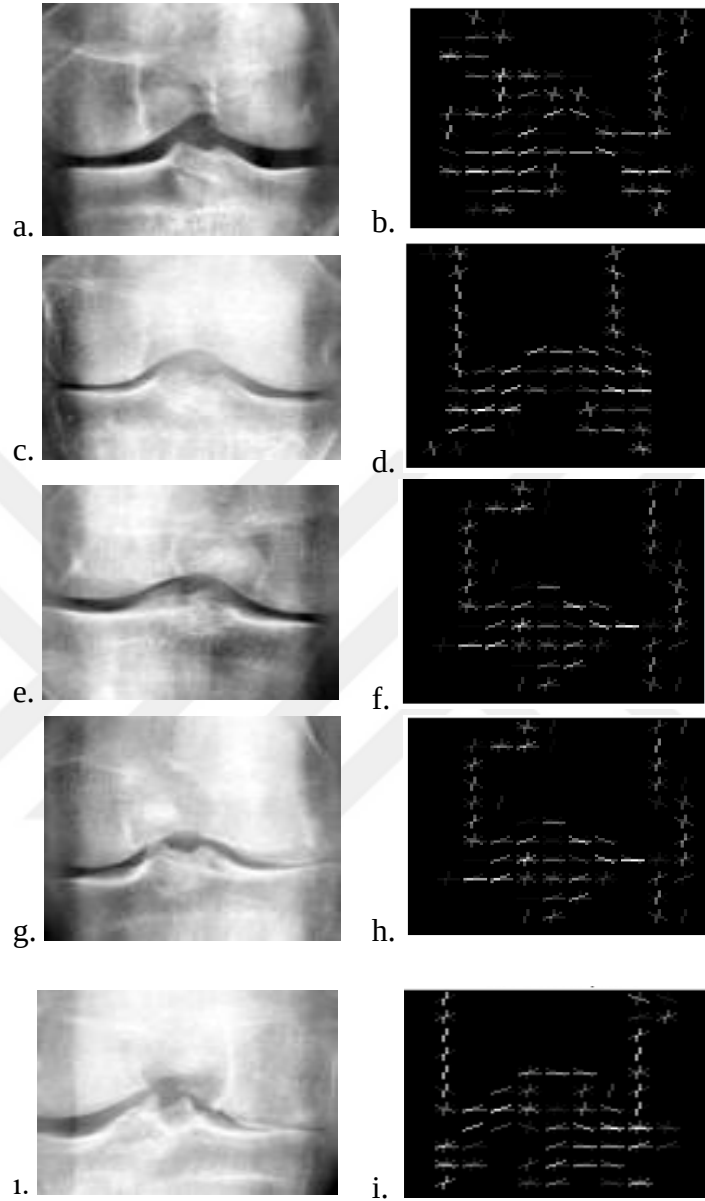
- Gradyan Büyüklüğü (Magnitude): Kenarın ne kadar belirgin olduğunu gösterir.
- Gradyan Yönü (Orientation): Kenarın hangi yönde olduğunu belirtir.

HOG algoritması aşağıdaki adımlardan oluşur:

- i. Girdi görüntüyü gri tonlamaya dönüştürme
- ii. Gradyan büyüklüğü ve yönü hesaplama
- iii. Görüntü bölgelere (hücrelere) bölme
- iv. Her hücre için yön histogramı oluşturma
- v. Histogramları normalize etme
- vi. Elde edilen özellik vektörünü nesne tanıma için kullanma

Görüntüdeki kenarları ve yön bilgilerini analiz eden bir yöntemdir. Gradyan yönlerinin dağılımını histogramlarla temsil eder (Kumar, 2020). Kenar ve yön bilgilerinin önemli olduğu nesne tespitinde ve sınıflandırmada kullanılır. Şekil 3.4, çalışmada kullanılan veri setinden seçilen farklı KOA derecelerine ait örnek X-ışını görüntüleri ile bu görüntülere uygulanan HOG temsillerini birlikte sunmaktadır. Görsel çiftleri sırasıyla “sağlıklı”, “şüpheli”, “hafif”, “orta” ve “ağır” KOA derecelerini temsil etmektedir. Sol sütunlarda yer alan görüntüler ham X-ışını verilerini, sağ sütunlarda ise bu görüntülerden çıkarılan HOG temsillerini göstermektedir. HOG temsilleri, görüntüdeki kenar, köşe ve yönelime dayalı yapısal bilgileri vurgulayan görsellerdir. Bu temsiller, özellikle morfolojik ayrıntıların belirginleştirilmesini sağlayarak KOA derecelerinin ayırımında önemli katkı sunmaktadır. Her bir HOG temsili, X ekseninde görüntünün yatay (soldan sağa), Y ekseninde ise dikey (yukarıdan aşağıya) piksel

konumlarını göstermektedir. Bu görselleştirme sayesinde, her KOA derecesine ait yapısal yönelim farklılıkları görsel olarak karşılaştırılabilir hale getirilmiştir.



Şekil 3.4. (a), (c), (e), (g), (i) X-ışını görüntülerini; (b), (d), (f), (h), (i) ise ilgili HOG (Histogram of Oriented Gradients) çıktıların görselleştirmesini göstermektedir. Görseller sırasıyla “sağlıklı”, “şüpheli”, “hafif”, “orta” ve “ağır” KOA derecelerine karşılık gelmektedir. HOG temsilleri, kenar bilgilerini ve yapısal yönelimleri vurgulayarak dokusal ayrımı güçlendirmektedir.

HOG Çıkarılan Özellikler:

- **Ortalama:**

HOG histogramındaki değerlerin aritmetik ortalamasıdır (Denklem 3.21).

$$\text{Ortalama} = \frac{1}{N} \sum_{i=1}^N HOG_i \quad (3.21)$$

- **Standart Sapma:**

HOG histogramındaki değerlerin dağılımını ölçer (Denklem 3.22).

$$\text{Standart Sapma} = \sqrt{\frac{1}{N} \sum_{i=1}^N (HOG_i - \overline{HOG})^2} \quad (3.22)$$

- **Maksimum:**

HOG histogramındaki en yüksek değerdir (Denklem 3.23).

$$\text{Maksimum} = \max (HOG_i) \quad (3.23)$$

- **Minimum:**

HOG histogramındaki en düşük değerdir (Denklem 3.24).

$$\text{Minimum} = \min (HOG_i) \quad (3.24)$$

3.3.5.Yerel İkili Örüntü (Local Binary Pattern - LBP):

Görüntünün her bir pikseline, çevresindeki piksellerin ilişkisine dayalı olarak bir doku deseni atar.

LBP Çıkarılan Özellikler:

- **Uniform LBP:**

Tekdüze (düzensiz olmayan) örüntüler.

- **Non-uniform LBP:**

Düzensiz örüntüler.

Denklem 3.25’de görülen formülle hesaplanan LBP değeri, merkez pikselin çevresindeki piksel yoğunluklarıyla karşılaştırılmasıyla hesaplanır:

$$LBP = \sum_{p=0}^{P-1} s(I_p - I_c) * 2^p \quad (3.25)$$

Denklemde;

I_p : Komşu piksel yoğunluğunu,

I_c : Merkez piksel yoğunluğunu

belirtmektedir.

Dokusal desenlerin analizinde, özellikle yüzey ve cilt dokusu gibi alanlarda kullanılır.

3.3.6. Dalgacık Dönüşümü:

Zamansal ve frekans analizini sağlayarak görüntüyü farklı çözünürlük seviyelerinde inceleyen bir yöntemdir. Farklı alt bantlarda özellik çıkarır. Çoklu çözünürlük analizi gerektiren uygulamalarda (örneğin, sinyal işleme ve dokusal analiz) kullanılır. Dalgacık dönüşümüyle analiz edilen alt bantlardan toplamda aşağıda belirtilen 4 özellik çıkarılmıştır.

Dalgacık Dönüşümü Çıkarılan Özellikler:

- **Düşük Frekanslı Bileşenlerin Varyansı (CA Katsayıları):**

Görüntünün temel yapısını ve düşük frekanslı bileşenlerini temsil eder. Bu varyans, görüntünün genel yapısının ne kadar sabit veya değişken olduğunu sayısal olarak ifade eder ve genellikle büyük ölçekli yapılar ile homojen bölgeleri analiz etmek için kullanılır. Denklem 3.26'da elde edilen özellik formülüne edilmiştir.

$$Var_{CA} = \frac{1}{N} \sum_{n,m} (s[n,m] * \phi[n,m] - \mu)^2 \quad (3.26)$$

Denklemde; $s[n,m]$: Görüntünün piksel yoğunluk değerlerini, $\phi_{j,k}[n,m]$: Ölçekleme (scaling) fonksiyonu (seviyeye bağlı) ifade eder. μ ise ortalama değeri temsil eder.

- **Yatay Detay Katsayılarının Varyansı (CH Katsayıları):**

Görüntünün yatay detaylarını içeren yüksek frekanslı bileşenleri temsil eder. Bu varyans, görüntüde yatay düzlemdeki kenarların, ince yapıların ve hızlı değişimlerin ne kadar yoğun olduğunu ölçer (Denklem 3.27).

$$Var_{CH} = \frac{1}{N} \sum_{n,m} (s[n, m] * \psi_H[n, m] - \mu)^2 \quad (3.27)$$

Denklemdeki, $\psi_V[n, m]$: Yatay dalgacık fonksiyonudur.

- ***Dikey Detay Katsayılarının Varyansı (CV Katsayıları):***

Görüntünün dikey yönlerdeki detaylarını temsil eder. Bu varyans, dikey düzlemdeki kenarları, ince detayları ve bölgesel değişimleri analiz etmek için kullanılır. CV katsayıları da yatay bileşenler gibi yüksek frekanslıdır. Bu nedenle varyansları da yüksek olacaktır. Dikey yapıların ne kadar belirgin olduğunu ve ne kadar değiştiğini yansıtır. Hesaplanması Denklem 3.28'de verilmiştir.

$$Var_{CV} = \frac{1}{N} \sum_{n,m} (s[n, m] * \psi_V[n, m] - \mu)^2 \quad (3.28)$$

Denklemdeki $\psi_V[n, m]$, dikey dalgacık fonksiyonunu ifade eder.

- ***Köşegen Detay Katsayılarının Varyansı (CD Katsayıları):***

Görüntünün köşegen yönlerdeki detaylarını ifade eder. Bu varyans, görüntüde çapraz yapıların ve yönlere göre farklılaşan özelliklerin yoğunluğunu değerlendirmek için önemlidir (Denklem 3.29). Denklemdeki $\psi_D[n, m]$, Diyagonal dalgacık fonksiyonudur.

$$Var_{CD} = \frac{1}{N} \sum_{n,m} (s[n, m] * \psi_D[n, m] - \mu)^2 \quad (3.29)$$

3.3.7. Zernike Moment

Zernike Momentleri, bir görüntünün dairesel bölgesinde tanımlanan ortogonal fonksiyonlar kullanılarak hesaplanan momentlerdir. Görüntü işleme ve desen tanıma

alanlarında yaygın olarak kullanılır. Döndürme değişmezliği sağladıkları için özellikle nesne tanıma uygulamalarında tercih edilirler (Tahmasbi, 2011).

3.3.8. Diğer Özellikler

- **Alan Özelliği:**

Görüntünün veya ROI'nin (İlgi Bölgesinin) kapladığı toplam piksel sayısını ifade eder. Denklem 3.30'daki gibi hesaplanır. Nesnelerin büyüklüklerini karşılaştırmada kullanılır.

$$\text{Alan} = \sum_{i,j} ROI(i,j) \quad (3.30)$$

- **Çevre Özelliği:**

Görüntüdeki nesnenin kenarlarının toplam uzunluğunu ifade eder (Denklem 3.31).

$$\text{Çevre} = \sum_k E_k \quad (3.31)$$

E_k : ROI deki kenar piksellerin toplamıdır.

Kenarların uzunluğunu ölçerek düzensiz yapıların tespiti için kullanılır.

3.4. Boyut Azaltma Yöntemleri

Boyut azaltma yöntemleri, yüksek boyutlu verilerdeki boyut sayısını düşürerek, analiz ve modelleme süreçlerini daha etkili ve yönetilebilir hale getirmeyi amaçlar. Bu yöntemler, verideki önemli bilgiyi korurken, gereksiz gürültü ve karmaşıklığı azaltarak modellerin daha hızlı ve doğru çalışmasını sağlar.

3.4.1. ANOVA

ANOVA (Analysis of Variance - Varyans Analizi), farklı gruplar arasında istatistiksel olarak anlamlı bir fark olup olmadığını analiz eden parametrik bir testtir. ANOVA, özellikle birden fazla grup arasında ortalama farklarını test etmek için kullanılan güçlü bir yöntemdir. Bu yöntem, farklı gruplardaki varyansların

karşılaştırılması yoluyla çalışır ve bağımlı değişkenin gruplara göre anlamlı farklılık gösterip göstermediğini belirlemek için kullanılır (Büyüköztürk, 2012).

Özellik seçimi süreçlerinde, hangi değişkenlerin sınıflar arasında ayırt edici olduğunu belirlemek amacıyla ANOVA kullanılabilir. Örneğin, X-ışını görüntülerinden elde edilen belirli bir özelliğin, sağlıklı bireyler ile KOA'lı bireyler arasında anlamlı bir farklılık gösterip göstermediğini test etmek için ANOVA uygulanabilir. ANOVA sonucunda istatistiksel olarak anlamlı bulunan özellikler seçilerek, makine öğrenmesi ve veri analizi süreçlerinin doğruluğu artırılabilir.

3.4.1.1. ANOVA ile Özellik Seçimi

Tek yönlü ANOVA, bağımsız değişkenlerin tek bir faktöre göre sınıflandırıldığı durumlarda kullanılır ve şu hipotezleri test eder:

- **H₀ (Null Hipotezi):** Tüm grupların ortalamaları eşittir, yani özelliğin farklı sınıflar arasında anlamlı bir farkı yoktur.
- **H₁ (Alternatif Hipotez):** En az bir grup ortalaması diğerlerinden farklıdır, yani ilgili özellik sınıflar arasında farklılık göstermektedir.

ANOVA testi sonucunda p-değeri (p-value) elde edilir. Eğer bu p-değeri belirli bir eşik değerinin (genellikle 0,05) altındaysa, ilgili özelliğin gruplar arasında anlamlı bir ayrım sağladığı kabul edilir.

ANOVA testi aşağıdaki adımlarla gerçekleştirilir:

i. Verinin Hazırlanması:

- Özelliklerin (feature) bağımsız değişken olarak belirlenmesi
- Hedef değişkenin (class label) bağımlı değişken olarak belirlenmesi

ii. F-İstatistiğinin Hesaplanması:

- F-Testi, grupların varyanslarının oranı üzerinden hesaplanır. F-değeri ne kadar büyükse, gruplar arasındaki farklılık o kadar belirgin olur.
- `f_oneway()` fonksiyonu ile her bir özelliğin sınıflara göre farklılığı analiz edilir.

iii. **p-Değerlerinin Analizi:**

- Eğer p-değeri $< 0,05$ ise, ilgili özelliğin sınıflar arasında anlamlı bir farklılık gösterdiği kabul edilir.
- Eğer p-değeri $\geq 0,05$ ise, ilgili özellik sınıflar arasında anlamlı bir fark yaratmadığı için elenir.

iv. **En Anlamlı Özelliklerin Seçilmesi:**

- En düşük p-değerine ve yüksek F değerine sahip en iyi n adet özellik seçilir.

Gerçekleştirilen tez çalışmasında, manuel olarak çıkarılan 35 farklı geleneksel görüntü özelliği arasından en anlamlı olanların seçilmesi amacıyla ANOVA testi uygulanmıştır. ANOVA, her bir özelliğin hedef değişken ile ilişkisini istatistiksel olarak analiz etmek için kullanılmıştır. Özellik seçim süreci şu şekilde gerçekleştirilmiştir:

i. **Özelliklerin p-değerleri hesaplandı:**

- Python `scipy.stats.f_oneway()` fonksiyonu kullanılarak her bir özellik için F-testi ve p-değeri hesaplandı.
- Bu işlem, özelliklerin farklı sınıflara göre değişkenlik gösterip göstermediğini anlamak için yapıldı.

ii. **p-değeri 0,05'ten küçük olanlar seçildi:**

- Eğer bir özelliğin p-değeri $< 0,05$ ise, bu özelliğin sınıflar arasında anlamlı bir ayrım yapabildiği kabul edildi.
- p-değeri 0,05'ten büyük olan özellikler elendi çünkü bunlar sınıflar arasında istatistiksel olarak anlamlı bir fark yaratmıyordu.

iii. **ANOVA F yöntemi kullanılarak özellikler seçildi:**

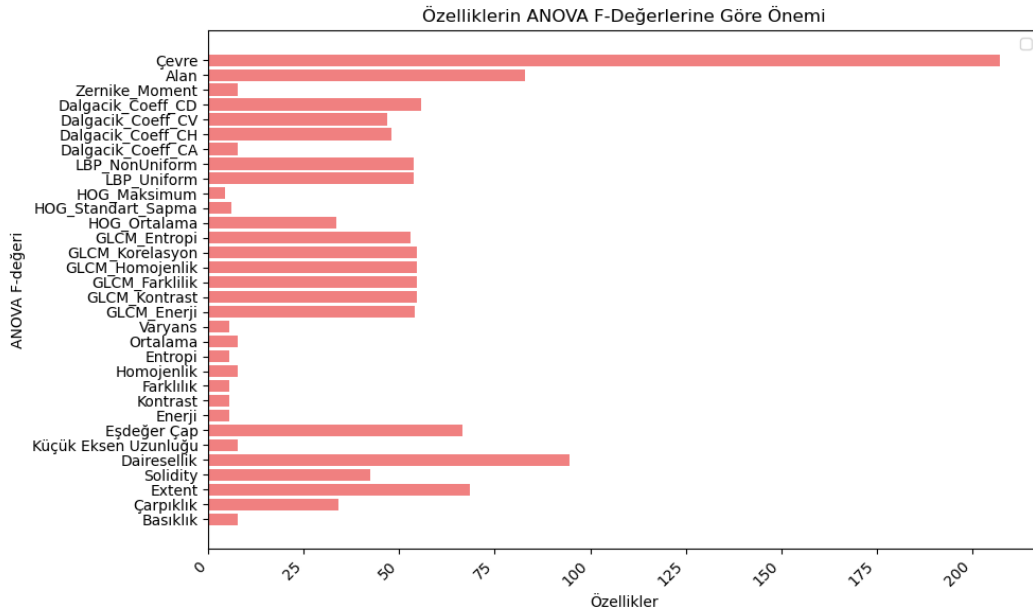
- Öncelikle, tüm veri kümesindeki her bir özellik için sınıflar arasındaki farkları ölçmek amacıyla F-değeri hesaplanmıştır. Ardından, bu özelliklerin ayırt edicilik gücünü değerlendirmek için F-değerlerinin ortalaması ve standart sapması temel alınarak (ortalama $\pm k * \text{Standart sapma}$) formülü ile bir eşik değeri (threshold) belirlenmiştir.

- $k=1$ temel alınarak hesaplanan eşik değerinin üzerinde F- değerine sahip 16 özellik, en anlamlı özellikler olarak seçilmiştir.

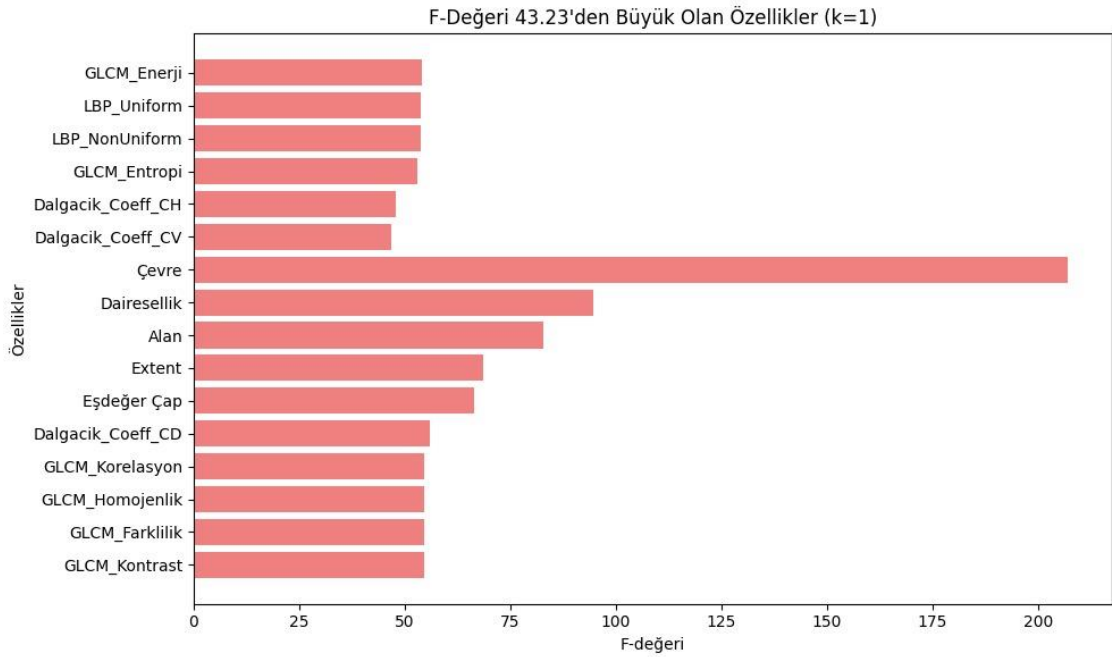
ANOVA F yöntemiyle özellik seçimi yapılırken, sınıflar arasındaki ayırt ediciliği belirlemek amacıyla her bir özellik için F-değeri hesaplanmıştır. Bu değerlerin ortalaması ve standart sapması temel alınarak belirlenen eşik değeri sayesinde, yalnızca anlamlı özelliklerin seçilmesi hedeflenmiştir. Ancak, veri setimizde ilk 10 özelliğin F-değerleri aşırı yüksek olduğundan, standart sapmayı da büyük ölçüde artırmıştır. Bu durum, $k=1$ seçildiğinde bile çok az sayıda özelliğin eşik değeri üzerinde kalmasına neden olmuştur. Kısıtlı sayıda özellik seçilmesi, modelin önemli bilgileri kaybetmesine yol açabileceğinden, bu yüksek değerli 10 özellik hariç tutularak standart sapma yeniden hesaplanmıştır.

Bu düzeltme sonrasında, standart sapma daha adil bir ölçüt haline gelmiş ve $k=1$ seçimiyle makul sayıda önemli özellik belirlenebilmiştir. Böylece, yalnızca en uç değerli birkaç özellik yerine, daha geniş bir özellik kümesi değerlendirilerek, modelin öğrenme kapasitesi artırılmıştır. $k=2$ ve $k=3$ değerleri ise eşik değerini çok katı hale getirdiğinden çok az özellik seçilmesine neden olmuştur. Sonuç olarak, $k=1$ seçimi, anlamlı 16 özelliğin belirlenmesini sağlamış ve modelin güçlü, dengeli ve yeterli çeşitliliğe sahip bir özellik kümesiyle çalışmasını mümkün kılmıştır.

Şekil 3.5, özelliklerin ANOVA F-Değerlerine göre önemini göstermektedir. Şekil 3.6 ise ANOVA ile seçilen en anlamlı 16 özelliği göstermektedir.



Şekil 3.5. Özelliklerin ANOVA F-Değerine göre önemi



Şekil 3.6. ANOVA ile en sonda işlemler sonucunda seçilen özellikler

ANOVA analizi sonucunda, Şekil 3.6'da gösterildiği üzere en anlamlı olarak değerlendirilen 16 özellik; çevre, dairesellik, alan, extent, eşdeğer çap, Dalgacik_Coeff_CD, GLCM_Korelasyon, GLCM_Homojenlik, GLCM_Farklılık, GLCM_Kontrast, GLCM_Enerji, LBP_NonUniform, LBP_Uniform, GLCM_Entropi, Dalgacik_Coeff_CH ve Dalgacik_Coeff_CV olarak belirlenmiştir. Bu özellikler, ANOVA testi ile elde edilen F-değerleri doğrultusunda seçilmiş ve sınıflar arası ayrımı en iyi sağlayan öznelilikler olarak değerlendirilmiştir. Görüldüğü üzere, özellikle GLCM ile çıkarılan doku tabanlı özellikler, LBP tabanlı özellikler ve çevre, dairesellik, alan gibi bazı morfolojik ölçümler yüksek F-değerleri ile öne çıkmıştır. En yüksek F-değerine sahip olan çevre özelliği, KOA derecelerinin ayrımında belirleyici bir unsur olmuştur. Bu sonuçlar, hem doku yapısının hem de morfolojik özelliklerin KOA derecelendirmesinde önemli bir rol oynadığını ortaya koymaktadır.

3.4.2. Doğrusal Diskriminant Analizi (Linear Discriminant Analysis – LDA)

Linear Discriminant Analysis (LDA), sınıflandırma problemlerinde kullanılan etkili bir boyut azaltma yöntemidir. Temel amacı, veriyi sınıflar arasında en iyi ayrımı sağlayacak şekilde dönüştürerek daha az boyutla daha anlamlı bir temsil elde etmektir.

Böylece, sınıflar arasındaki karışıklık azalırken, verinin daha iyi sınıflandırılmasına olanak tanınır.

Bu çalışmada, sağlıklı bireyler ile KOA bulunan bireyleri ayırt etmek için LDA kullanılmıştır. Veri kümesi toplamda manuel olarak elde edilen 35 farklı geleneksel özelliğe sahiptir ve LDA bu özellikler üzerinde uygulanmıştır. LDA'nın temel prensibi, sınıflar arası farkı artırırken, aynı sınıftaki veriler arasındaki değişkenliği en aza indirmektir. Bu sayede, sınıfların daha net ayrışması sağlanır.

LDA, verideki sınıf sayısına bağlı olarak en fazla (sınıf sayısı - 1) kadar yeni bileşen oluşturabilir. Bu çalışmada veri 5 farklı sınıftan oluştuğu için, elde edilen doğrusal bileşen sayısı 4 olmuştur.

LDA'nın etkili bir şekilde uygulanabilmesi için öncelikle veri üzerinde belirli işlemler gerçekleştirilmiştir:

i. **Özellik Vektörlerinin Hazırlanması**

Her bir veri örneği, 35 farklı geleneksel sayısal özelliğe sahiptir ve bu özellikler analiz için kullanılmıştır.

ii. **Sınıf Ortalamalarının Hesaplanması**

Her sınıfa ait özelliklerin ortalamaları belirlenerek, sınıflar arasındaki farklılıkların ortaya çıkması sağlanmıştır.

iii. **Kovaryans Matrislerinin Hesaplanması**

Her sınıfın içsel değişkenliği (kovaryans) hesaplanarak, sınıf içindeki dağılımın nasıl olduğu anlaşılmıştır.

iv. **Ayrımcı Bileşenlerin Belirlenmesi**

Sınıflar arasındaki farklar ve içsel değişkenlikler dikkate alınarak, veriyi en iyi şekilde ayıran doğrusal bileşenler elde edilmiştir.

v. **Boyut Azaltma İşlemi**

LDA, en iyi ayrımı sağlayan 4 doğrusal bileşeni seçerek veriyi daha düşük boyuta indirgemıştır.

LDA'nın uygulanması sonucunda:

Manuel olarak elde edilen 35 farklı geleneksel görüntü özelliği, 4 bileşene indirgenmiştir, bu da veri setinin daha kompakt hale gelmesini sağlamıştır. Yeni bileşenler, KOA derece sınıfların birbiriyle daha az örtüşmesini sağlayarak daha net bir

ayırım sunmuştur. LDA'nın sınıflar arasındaki ayrımı güçlendiren doğası, bu çalışmada daha düşük boyutlu ancak daha anlamlı bir veridir.

3.4.3. Temel Bileşenler Analizi (Principal Component Analysis - PCA)

Principal Component Analysis (PCA), boyut indirgeme amacıyla kullanılan istatistiksel bir yöntemdir. Yüksek boyutlu verileri daha düşük boyutlu bir uzaya projeksiyon yaparak, veri setindeki varyansı en iyi temsil eden yeni değişkenler (bileşenler) oluşturur.

Bu yöntem, veri içindeki korelasyonları dikkate alarak, değişkenler arasındaki ilişkileri özetler ve en fazla bilgi içeren doğrultuları belirler.

Temel amacı:

- Veriyi daha basit bir yapıda temsil etmek,
- Gürültüyü azaltmak,
- Görselleştirme ve işlem maliyetini düşürmek.

Özellikle yüksek boyutlu görüntü verilerinde, PCA önemli ölçüde veri sıkıştırması sağlar ve modelin genel performansını artırır.

PCA'nın çalışma prensibi, veri setindeki en yüksek varyansa sahip doğrultuları bulmaktır. Bu doğrultular, ana bileşenler (principal components) olarak adlandırılır. Ana bileşenler, aşağıdaki özelliklere sahiptir:

- I. Birinci bileşen: Verideki en büyük varyansı temsil eden doğrultudur.
- II. İkinci bileşen: Birinci bileşene dik (ortogonal) olacak şekilde, ikinci en yüksek varyansı temsil eden doğrultudur.
- III. Devam eden bileşenler de aynı mantıkla oluşturulur.
- IV. Her yeni bileşen, öncekilerle bağımsızdır ve veri setindeki farklı bir varyans boyutunu temsil eder.

Temel bileşenler hesaplanırken sırasıyla aşağıdaki adımlar takip edilir:

i. **Her boyut için ortalama vektör hesaplanır.**

Verinin her boyuttaki ortalama değeri bulunur ve tüm veriler bu ortalamadan çıkarılarak merkezleme işlemi yapılır.

ii. **Kovaryans matrisleri hesaplanır.** *(Değişkenler arası ilişkiyi anlatan matris)*

- Kovaryans matrisi, değişkenler arasındaki ilişkileri gösterir.
- Eğer değişkenler arasında yüksek kovaryans varsa, bunlar benzer bilgi taşıyor demektir.

- iii. **Her boyut için özvektörler ve karşılık gelen özdeğerler hesaplanır.**
 - **Özvektörler (eigenvectors):** Verinin en yüksek değişkenliğe sahip yönlerini belirler. (İdeal projeksiyonu vermesi beklenir.)
 - **Özdeğerler (eigenvalues):** Özvektörlere karşılık gelen ilişkilerin gücünü gösterir. Yani her temel bileşenin verideki varyansı ne kadar açıkladığını belirler.
- iv. **Her bir özdeğerin özdeğerler toplamına bölünmesi ile temel bileşenlerin toplam varyansı açıklama yüzdeleri elde edilir.**
 - Büyük özdeğerlere sahip olan bileşenler daha fazla bilgi içerir.
 - Özdeğerler sıralanarak en fazla varyansı açıklayan bileşenler seçilir.

Bu süreç, başlangıçta p boyutlu bir uzayda tanımlı veriyi, toplam varyansın büyük bir kısmını koruyarak $m < p$ boyutlu bir alt uzaya dönüştürmeyi amaçlar. Böylece, verinin temel yapısını en iyi temsil eden doğrusal bileşenler seçilir ve daha düşük boyutlu bir gösterime ulaşılır (Kara, 2019).

Boyut indirgeme, veri içerisindeki gereksiz veya yüksek korelasyonlu değişkenleri elimine ederek daha kompakt bir gösterim sağlar. Bu, hem hesaplama maliyetini azaltır hem de veri analizi ve modelleme süreçlerinde daha anlamlı desenlerin ortaya çıkmasını sağlar. Özellikle yüksek boyutlu veri kümelerinde, temel bileşenlerin seçilmesiyle bilgi kaybı minimize edilirken gürültü ve aşırı uyum riski de azaltılmış olur.

3.5. Derin Öğrenme Tabanlı Görüntü Özelliklerinin Çıkarılması

Bu çalışmada, geleneksel görüntü özelliklerine ek olarak derin öğrenme tabanlı özellikler de kullanılarak KOA derecelendirmesi sürecinde bilgi temsili güçlendirilmiştir. Derin özellikler, önceden büyük veri kümeleri üzerinde eğitilmiş başarılı CNN modelleri kullanılarak elde edilmiştir. Özellikle derin öğrenme tabanlı modeller, çok katmanlı yapıları sayesinde görüntülerdeki karmaşık yapısal örüntüleri otomatik olarak öğrenebilme kapasitesine sahiptir. Bu özellikleri ile geleneksel elle çıkarılan özniteliklere kıyasla daha zengin ve soyut temsiller sunarak sınıflandırma performansını artırmaktadır.

Ancak, derin öğrenme modellerinin sıfırdan eğitilmesi büyük miktarda etiketli veri gerektirdiğinden, tıbbi görüntüleme gibi sınırlı veri içeren alanlarda doğrudan

sıfırdan eğitim yapılması pratik olmamaktadır. Bu nedenle, bu çalışmada transfer öğrenme yaklaşımı tercih edilmiştir. Transfer öğrenme sayesinde, daha önce geniş veri kümeleri üzerinde eğitilmiş derin ağların, mevcut KOA görüntüleri üzerinde yeniden kullanılmasına olanak sağlanmış ve veri gereksinimi azaltılmıştır. Ayrıca, bu yöntem modelin eğitim süresini kısaltmış ve aşırı öğrenme riskini azaltmıştır.

Derin özelliklerin çıkarılması için ResNet-50, VGG16, EfficientNetB0, Xception ve DenseNet201 modelleri kullanılmıştır. Bu modeller, literatürde KOA teşhisi başta olmak üzere birçok tıbbi görüntü işleme uygulamasında yaygın şekilde başarıyla kullanılmıştır (Usman ve diğ., 2023; Wani ve diğ., 2023; Chen ve diğ., 2019; Lau ve diğ., 2022; Solak, 2024; Mahfus ve diğ. 2025). Farklı mimari yapıları ve derinlik seviyeleri sayesinde çeşitli görüntü özelliklerini öğrenme kabiliyetine sahip bu modeller, KOA derecelerinin doğru şekilde ayrılmasına katkı sağlaması amacıyla tercih edilmiştir.

3.5.1. ResNet-50

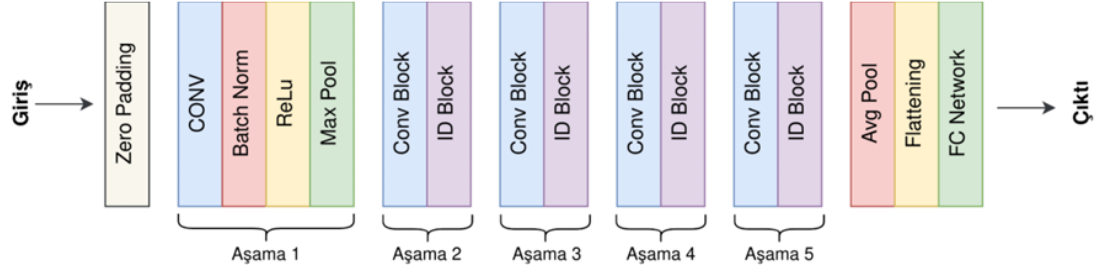
ResNet-50, derin yapısı ve residual bağlantıları sayesinde gradyan kaybını önleyerek çok katmanlı ağların etkin bir şekilde eğitilmesine olanak tanıyan bir CNN mimarisidir (He ve diğ. 2016). Bu yapı, özellikle morfolojik, tekstürel ve dokusal farklılıkları yakalamada yüksek temsil kapasitesi sunmaktadır. ResNet-50'nin genel mimarisi Şekil 3.7'de gösterilmiştir (Gençtürk ve diğ. 2023).

Model, geniş ve çeşitli görüntü veri kümelerinden (ImageNet) önceden eğitilmiş olması sayesinde, KOA derecelendirme gibi sınırlı veri içeren tıbbi görüntüleme problemlerinde transfer öğrenme ile etkili bir şekilde kullanılmaktadır. Bu çalışmada, ResNet-50 doğrudan sınıflandırma amacıyla kullanılmamış; bunun yerine, önceden eğitilmiş ağırlıklar korunarak, diz X-ışını görüntülerinden derin özellik vektörleri çıkarılmıştır.

Özellikle son sınıflandırma katmanları çıkarılarak, ağın ara katmanlarından elde edilen derin temsiller, geleneksel özelliklerle füzyon stratejileri kapsamında birleştirilmek üzere kullanılmıştır. Bu süreç, KOA'nın farklı derecelerini daha iyi temsil eden zengin ve ayrıştırıcı özelliklerin elde edilmesine katkı sağlamıştır.

Bu seçim, ResNet-50'nin derinliği sayesinde küçük yapısal değişiklikleri algılayabilme yeteneği ve daha önce elde edilmiş geniş genelleme kapasitesinden

faaydalanılarak gerekleřtirilmiřtir. Bylece, KOA derecelerinin daha gvenilir ve doėru řekilde ayrılması hedeflenmiřtir.

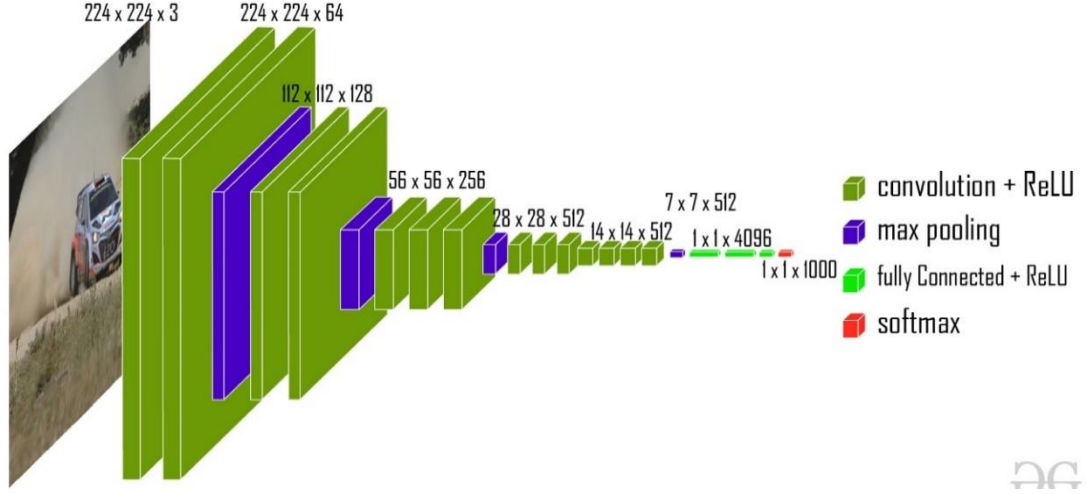


řekil 3.7. ResNet-50 Mimarisi

3.5.2. VGG16

Simonyan ve Zisserman (2014) tarafından nerilen VGG16, 13 evriřim katmanını ve 3 tam baėlantılı katmandan oluřan toplam 16 katmanlı bir CNN mimarisidir. Bu yapı, ardışık evriřim katmanlarının ardından uygulanan maksimum havuzlama katmanları ile uzamsal boyutların azaltılmasını saėlar (Li ve ark., 2023). Aė, tm evriřim iřlemlerinde 3×3 filtre boyutu kullanarak, daha az parametreyle daha derin bir aė oluřturulmasına olanak tanımakta ve bylece hiyerarřik grnt temsilleri ėrenmeyi hedeflemektedir (Yang ve ark., 2023).

Bu alıřmada, VGG16 modeli doėrudan sınıflandırma amacıyla deėil, nceden eėitilmiş aėırlıkları kullanılarak KOA X-ışını grntlerinden derin zellik ıkarımı amacıyla kullanılmıştır. zellikle ardışık evriřim katmanlarının saėladığı detaylı uzamsal temsiller sayesinde, KOA'ya zg morfolojik ve yapısal varyasyonların etkili bir řekilde modellenmesi hedeflenmiştir. VGG16'nın tutarlı ve basit mimarisi, farklı KOA derecelerine iliřkin karmařık yapısal bilgilerin ıkarımında avantaj saėlamaktadır. Bununla birlikte, aėın yksek parametre sayısı nedeniyle hesaplama maliyetinin diėer daha modern modellere kıyasla daha fazla olduėu da gz nnde bulundurulmuřtur. řekil 3.8 VGG16'nın genel mimarisini gstermektedir.



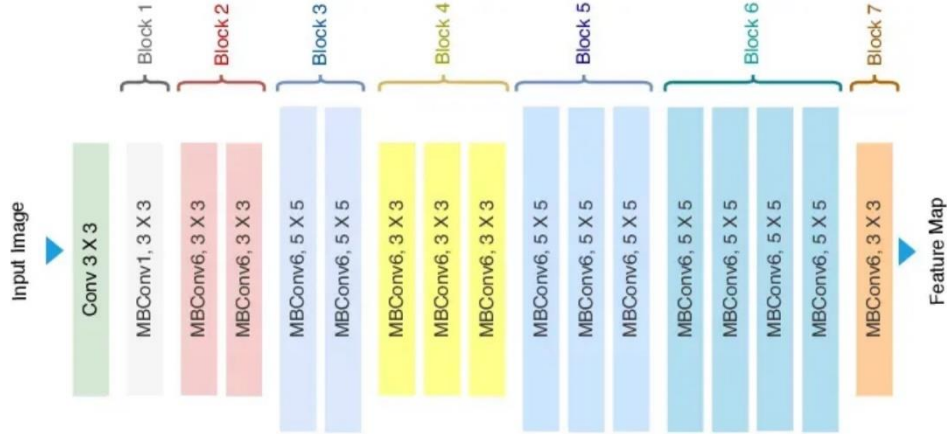
Şekil 3.8. VGG16 Mimarisi

3.5.3. EfficientNetB0

EfficientNetB0, Tan ve Le (2019) tarafından geliştirilen ve toplam 237 katmandan oluşan optimize edilmiş bir derin öğrenme modelidir. Bu modelde, mobil tersine çevrilmiş şişe boynu (Mobile Inverted Bottleneck) blokları ve Sıkıştırma ve Uyarım (Squeeze and Excitation - SE) mekanizmaları entegre edilmiştir. Ayrıca, Swish aktivasyon fonksiyonunun kullanımıyla ağırlık doğruluk ve verimlilik dengesinin artırılması hedeflenmiştir (Mou ve Razzak, 2023).

Model mimarisi, 3×3 alıcı alanlar kullanan çok katmanlı evrişim bloklarından oluşur ve AutoML MnasNet arama algoritması ile tasarlanmıştır. Bu yapı, parametre sayısını azaltırken, doğruluk oranını yüksek seviyede tutmaya olanak sağlamaktadır (Mou ve Razzak, 2023). Özellikle, kaynak kısıtlı cihazlar için geliştirilen EfficientNetB0, düşük hesaplama maliyeti ile dikkat çekmektedir.

Bu çalışmada EfficientNetB0, KOA derecelendirmesine yönelik doğrudan sınıflandırma amacıyla değil, X-ışını görüntülerinden derin özellik çıkarımı yapmak amacıyla kullanılmıştır. Böylece, KOA'nın farklı evrelerine ilişkin yapısal ve morfolojik detaylar daha etkili bir şekilde temsil edilmiştir. Modelin kompakt yapısı sayesinde, hesaplama yükü azaltılmış ve eğitim sürecinde daha verimli bir derin özellik temsili elde edilmiştir. Şekil 3.9'da EfficientNetB0'ın yapısı sunulmaktadır.



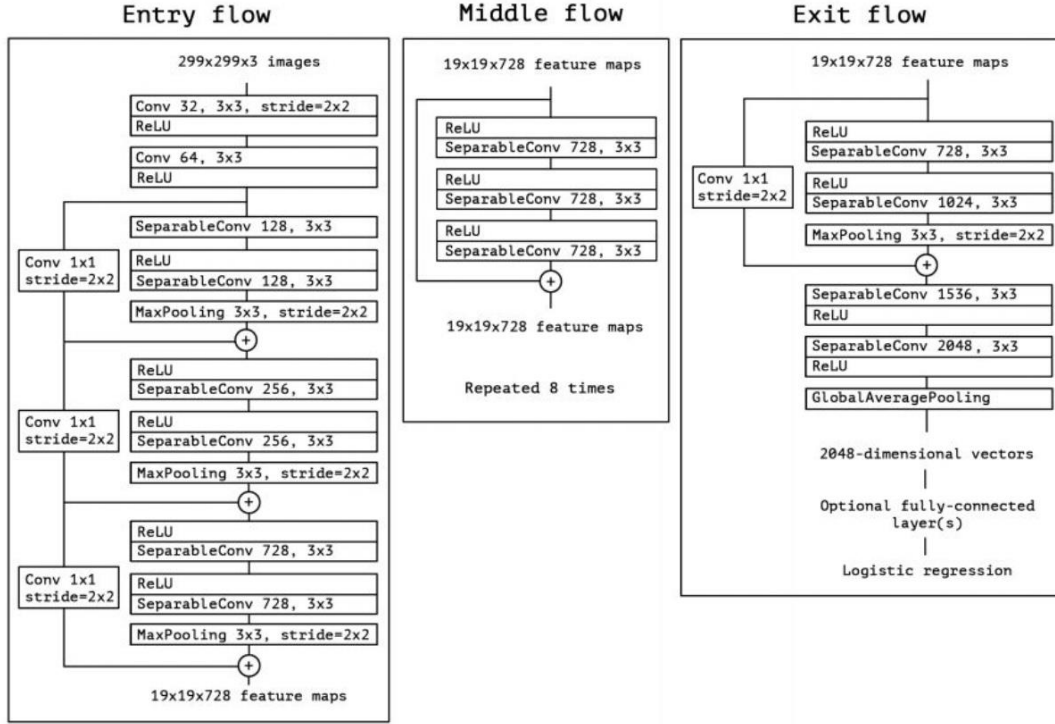
Şekil 3.9. EfficientNetB0 Mimarisi

3.5.4. Xception

Xception modeli, Chollet (2017) tarafından geliştirilmiş olup, yalnızca derinlik ayrılabilir evrişimlerden (depthwise separable convolutions) oluşan bir mimari üzerine inşa edilmiştir. Inception mimarisinin (Szegedy ve ark., 2016) bir uzantısı niteliğinde olan bu yapı, 71 katmanlı derin bir CNN modelidir. Xception, standart evrişim katmanları yerine derinlik ayrılabilir evrişimler ve artık bağlantılar (residual connections) kullanarak, hem hesaplama yükünü azaltmayı hem de ağıın öğrenme kapasitesini artırmayı hedeflemektedir (Shaheed ve ark. 2022).

Bu yaklaşım sayesinde, geleneksel evrişimsel işlemlere göre daha hafif ve verimli bir hesaplama süreci sağlanmaktadır. Özellikle, yüksek boyutlu verilerde daha az parametre kullanarak etkili özellik çıkarımı yapılmasına imkân tanır. Bu çalışmada, Xception modeli doğrudan sınıflandırma amacıyla değil, KOA derecelendirmesinde kullanılacak derin özelliklerin elde edilmesi amacıyla tercih edilmiştir. Modelin hesaplama verimliliği, sınırlı veri koşullarında bile detaylı morfolojik ve dokusal bilgilerin etkili bir şekilde çıkarılmasını sağlamıştır.

Bununla birlikte, çok büyük veri setlerinde zaman zaman ResNet gibi diğer derin modellerle karşılaştırıldığında doğruluk oranlarının daha düşük olabileceği bilinmektedir. Xception mimarisinin genel yapısı Şekil 3.10'da sunulmaktadır.



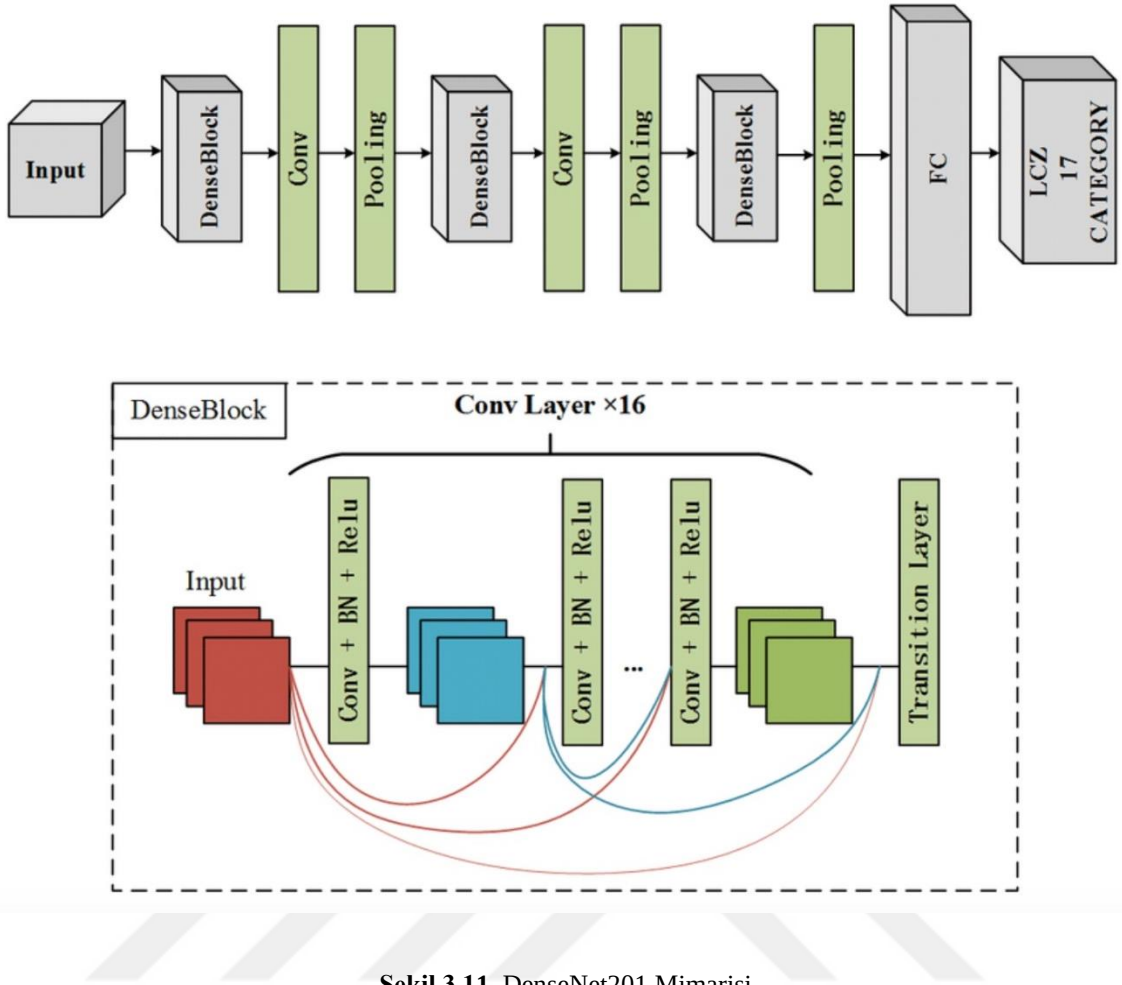
Şekil 3.10. Xception Mimarisi

3.5.5. Densenet201

DenseNet201 çerçevesi, Huang ve ark. (2017) tarafından önerilen yoğun bağlantılı bir mimariye dayanmaktadır. Bu modelde, her katman, önceki tüm katmanların çıktısını doğrudan girdi olarak alır. Böylece, özelliklerin ağ boyunca tekrar tekrar kullanılmasına olanak tanınarak bilgi akışı güçlendirilir ve öğrenme verimliliği artırılır. Ayrıca, havuzlama katmanları ile kompakt bir yapı sunarak, parametre sayısını ve genel model karmaşıklığını azaltmayı hedeflemektedir (Turkoğlu, 2021).

Bu yapısal yaklaşım, hem bellek hem de hesaplama açısından verimli bir derin öğrenme mimarisi ortaya koyar. Bu çalışmada DenseNet201, KOA derecelendirme sürecinde kullanılacak derin özelliklerin çıkarılması amacıyla tercih edilmiştir. Özellikle, farklı katmanlardan gelen özelliklerin paylaşılması sayesinde, diz eklem yapılarındaki ince varyasyonların daha etkili bir şekilde yakalanması sağlanmıştır.

Bununla birlikte, ağıın çok derin olması bazı durumlarda hesaplama maliyetini artırabilmekte ve eğitim süresini uzatabilmektedir. DenseNet201'in genel mimari yapısı Şekil 3.11'de gösterilmektedir.

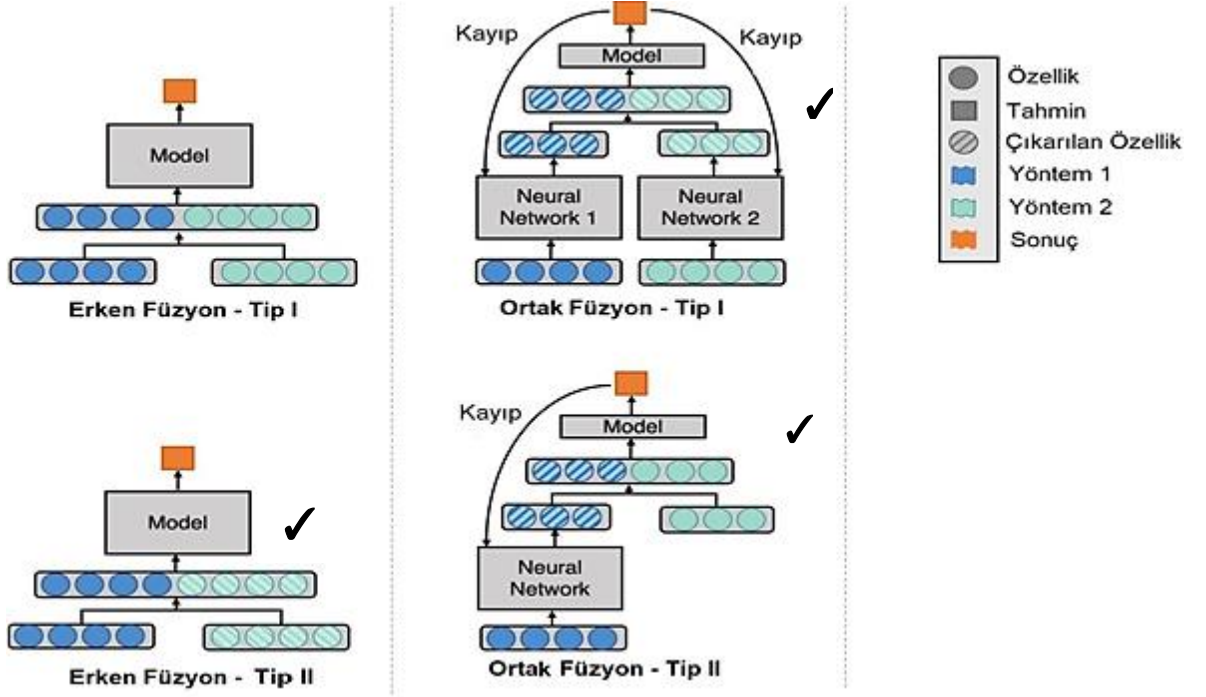


Şekil 3.11. DenseNet201 Mimarisi

3.6. Füzyon Stratejileri

Tez çalışmasında, Erken Füzyon Tip II, Ortak Füzyon Tip I ve Ortak Füzyon Tip II olmak üzere farklı füzyon stratejileri uygulanmıştır. Bu stratejiler, KOA derecelendirmesinde, aynı görüntülerin farklı temsil türlerinden (geleneksel ve derin özellikler) elde edilen bilgilerin entegrasyonu yoluyla model başarımını artırmayı hedeflemektedir. Erken füzyon yaklaşımı, farklı temsil seviyelerinden çıkarılan özelliklerin doğrudan birleştirilmesiyle bilgi bütünlüğü sağlamayı amaçlarken; ortak füzyon stratejisi ise, sinir ağlarının farklı katmanlarından öğrenilen temsilleri bir araya getirerek daha zengin ve soyut özellik vektörleri oluşturmayı hedefler. Her iki strateji de, bilgi çeşitliliğini artırarak sınıflar arası ayrımı güçlendirmekte ve KOA derecelendirme performansına anlamlı katkılar sunmaktadır.

Şekil 3.12. bu tez çalışmasında faydalanan füzyon stratejilerini göstermektedir.



Şekil 3.12. Derin öğrenmeyi kullanan füzyon stratejileri. Farklı füzyon stratejileri için model mimarisi. Erken füzyon (soldaki şekil), orijinal veya çıkarılan özellikleri giriş seviyesinde birleştirir. Ortak füzyon (ortadaki şekil) da özellikleri giriş seviyesinde birleştirir, ancak kayıp, özellik çıkarma modeline geri yayılır (Huang ve diğ., 2020).

3.6.1. Erken Füzyon Stratejisinin Uygulanması

Erken Füzyon (Early Fusion), çoklu veri kaynağından veya farklı özellik çıkarma tekniklerinden elde edilen bilgilerin, sınıflandırma veya karar verme aşamasından önce birleştirilmesiyle oluşturulan bir veri entegrasyon stratejisidir. Bu yöntem, verinin en erken aşamada birleştirilmesini sağladığı için "erken" olarak adlandırılır. Erken füzyonun temel amacı, farklı bilgi türlerini tek bir birleşik temsil (özellik vektörü) haline getirerek, sınıflandırma modeline daha zengin ve anlamlı bir giriş sağlamaktır. Bu sayede, sınıflandırma modelinin karmaşık ilişkileri daha iyi öğrenmesi ve genel doğruluğun artması hedeflenir. İki türü vardır Tip I ve Tip 2. Bu tez çalışmasında Erken Füzyon Tip II stratejisinden faydalanılmıştır.

Erken Füzyon Tip I farklı kaynaklardan gelen ham verilerin doğrudan sınıflandırma aşamasından önce birleştirilmesi ve sonrasında bir sınıflandırılma modeline verilmesi stratejisidir.

Erken Füzyon Tip II ise özelliklerin çıkarılmasından sonra, ancak sınıflandırma veya tahmin sürecinden önce gerçekleştirilen gelişmiş bir entegrasyon stratejisidir. Bu yöntemde, farklı kaynaklardan veya farklı özellik çıkarma tekniklerinden elde edilen

öznitelikler birleştirilerek daha zengin ve kapsamlı bir özellik vektörü oluşturulur. Özellikle, bu yaklaşım, modelin daha anlamlı ve bilgi açısından zengin bir giriş ile beslenmesini sağlar. Erken füzyonun bu türü, sınıflandırıcıya tek bir birleşik temsil sunarak, sınıf ayrımı için daha güçlü bir temel oluşturur (Huang ve ark., 2020).

Tez çalışmasında kullanılan tek bir veri kaynağı olduğu için Erken Füzyon Tip I stratejisi çalışma kapsamında gerçekleştirilememiştir. Erken Füzyon Tip 2 ise çalışmaya dahil olan veri kaynağından farklı özellik çıkarım teknikleri ile elde edilen özellikleri (geleneksel ve derin) birleştirilmesi yolu ile tez çalışmasında uygulanabilmektedir.

Bu kapsamda, KOA teşhisine yönelik olarak ilk aşamada X-ışını diz görüntülerinden morfolojik (8 adet), GLCM (6 adet), istatistiksel (8 adet), HOG (4 adet), LBP (2 adet), dalgacık dönüşümü (4 adet), çevre (1 adet), alan (1 adet) ve Zernike momenti (1 adet) olmak üzere toplamda 35 adet geleneksel görüntü özelliği manuel olarak çıkarılmıştır. Morfolojik özellikler görüntüdeki şekil, boyut ve yapısal detayları; dokusal özellikler yüzey desenlerini; istatistiksel özellikler ise piksel yoğunluğu ve histogram tabanlı bilgileri temsil etmektedir. Bu özellikler, KOA'nın derecelerini belirlemede önemli biyobelirteçler olarak değerlendirilmiştir.

Çıkarılan özelliklerin daha etkili ve anlamlı hale getirilmesi amacıyla LDA ve ANOVA testleri uygulanmıştır. LDA ile sınıf ayrımını en iyi yansıtan 4 yeni bileşen oluşturulurken, ANOVA testi sonucunda en anlamlı 16 özellik seçilmiştir. Böylece geleneksel özellikler, toplam 20 boyutlu bir vektör halinde temsil edilmiştir.

Erken Füzyon Tip II stratejisinin uygulanabilmesi için X-ışını görüntülerinin farklı bir temsili daha oluşturulmuştur. İlk olarak, her bir 92x92 piksel boyutundaki görüntü, sayısal olarak temsil edilebilmesi için tek boyutlu bir vektöre dönüştürülmüştür. Bu işlem sonucunda 8464 boyutlu bir veri uzayı elde edilmiştir. Ancak, bu kadar yüksek boyutlu bir veriyle çalışmak hesaplama maliyetini artırabileceği ve analiz sürecini zorlaştırabileceği için boyut azaltma işlemi gerekliliği ortaya çıkmıştır. Bu amaçla, PCA yöntemi uygulanarak, verinin bilgisini maksimum düzeyde koruyacak şekilde 97 bileşen seçilmiştir. Böylece, yüksek boyutlu verinin gereksiz ve tekrarlayan bilgileri elenmiş, %85'lik varyansı temsil eden daha kompakt bir veri yapısı elde edilmiştir. Bu süreç, görüntülerin temel yapısını kaybetmeden daha verimli bir şekilde işlenmesini sağlamış ve analiz sürecini hızlandırmıştır.

KOA derecelerinin “sağlıklı”, “şüpheli”, “hafif”, “orta” ve “ağır” olmak üzere belirlenebilmesi için sınıflandırma aşamasına geçilmeden önce, Erken Füzyon Tip II stratejisi kapsamında, manuel olarak elde edilen ve LDA ile ANOVA testleriyle

optimize edilen 20 geleneksel özellik ile PCA kullanılarak boyutu 97'ye indirgenmiş görüntü tabanlı özellikler birleştirilmiş ve her bir görüntü için toplamda 117 boyutlu birleşik bir özellik vektörü oluşturulmuştur.

Bu 117 boyutlu özellik vektörleri kullanılarak KOA derecelendirmesi gerçekleştirilmiştir. Sınıflandırma sürecinde, tüm deneylerde ortak bir yapay sinir ağı (ANN) mimarisi kullanılmıştır. Böylece, farklı model yapılarına bağlı kalmaksızın sabit bir sınıflayıcı tasarımı korunmuştur. Ancak, sinir ağı katmanlarının başlangıç ağırlıkları, farklı transfer öğrenme modellerinden (ResNet-50, VGG16, EfficientNetB0, Xception ve DenseNet201) alınarak belirlenmiştir. Bu sayede transfer öğrenme modellerinin taşıdığı ön bilgi, sınıflandırma sürecine aktarılmıştır.

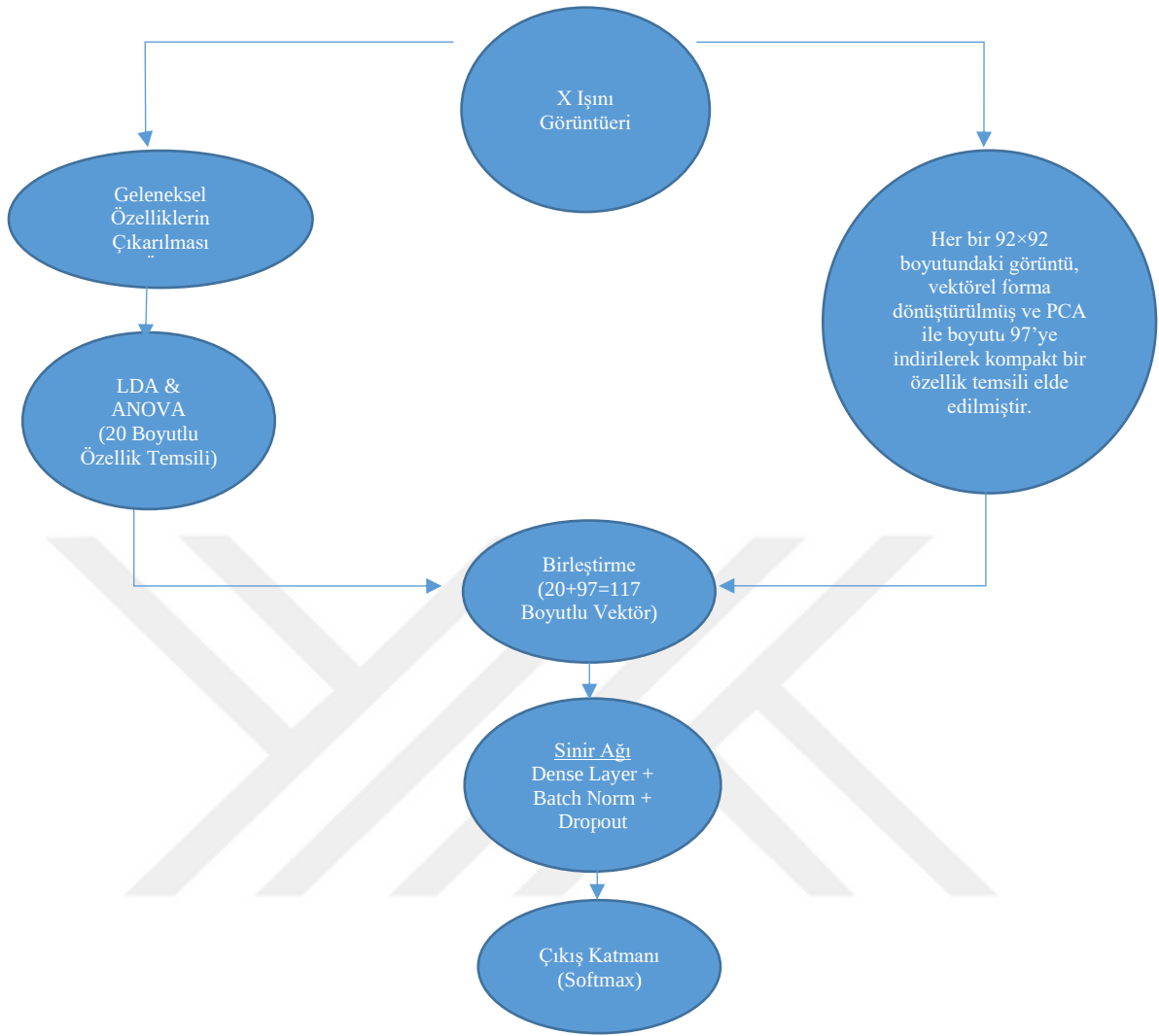
Kurulan sinir ağı mimarisi, KOA derecelendirmesi için şu şekilde yapılandırılmıştır. Giriş katmanı, 117 nöron içermekte olup, 117 boyutlu birleşik özellik vektörünü doğrudan almaktadır. Bu yapı, hem geleneksel hem de derin özelliklerin aynı anda modele aktarılmasına imkân sağlamaktadır.

İlk gizli katman, 128 nörondan oluşmakta ve ReLU aktivasyon fonksiyonu ile çalışmaktadır. Bu katmanda, ağı daha dengeli ve hızlı öğrenbilmesi için Batch Normalization uygulanmış ve aşırı öğrenmeyi önlemek amacıyla %30 oranında Dropout mekanizması eklenmiştir. İkinci gizli katman ise 64 nörondan oluşmakta olup, yine ReLU aktivasyonu, Batch Normalization ve %30 Dropout teknikleri ile desteklenmiştir. Böylece modelin öğrenme kapasitesi artırılırken aşırı uyum (overfitting) riski azaltılmıştır.

Çıkış katmanı, diz osteoartriti derecelerini temsil eden beş sınıflı modellemek amacıyla 5 nörondan oluşturulmuştur. Çıkış katmanında softmax aktivasyon fonksiyonu kullanılarak, her bir örnek için sınıflara ait olasılık değerleri hesaplanmıştır.

Modelin eğitimi sırasında Adam optimizasyon algoritması tercih edilmiş ve dengeli bir öğrenme süreci sağlamak amacıyla öğrenme oranı 0,0001 olarak belirlenmiştir. Kayıpların minimize edilmesi için, çok sınıflı sınıflandırma problemleri için uygun olan “Categorical Crossentropy” kayıp fonksiyonu kullanılmıştır. Ayrıca, eğitim süreci boyunca toplu işlem büyüklüğü (batch size) 32 olarak seçilmiş ve eğitim 70 epoch boyunca sürdürülmüştür. Aşırı öğrenmenin önüne geçebilmek amacıyla erken durdurma (early stopping) stratejisi uygulanarak modelin en iyi ağırlıklarının korunması sağlanmıştır.

Şekil 3.13 Erken Füzyon Tip II stratejisinin çalışmada uygulanmasına yönelik akışı göstermektedir.



Şekil 3.13. Erken Füzyon Tip II Süreci

3.6.2. Ortak Füzyon

Ortak füzyon, sinir ağlarının ara katmanlarından öğrenilen özellik temsillerini, nihai bir modele girdi olarak diğer yöntemlerden gelen özelliklerle birleştirme işlemidir. Bu füzyon yöntemi, özellikle sinir ağları ile uygulanır çünkü ağların ara katmanlarından elde edilen özellik temsilleri, daha gelişmiş ve genellikle daha bilgilendirici özellikler sağlar. Ortak füzyonun temel farkı, eğitim sırasında kaybın özellik çıkaran sinir ağlarına geri yayılmasıdır. Bu geri yayılım, her eğitim yinelemesi için daha iyi özellik temsilleri yaratılmasını sağlar ve bu da modelin performansını artırır. Ortak füzyon da iki temel türde uygulanabilir:

- **Ortak Füzyon Tip I:** Bu yöntem, farklı veri kaynaklarından veya özellik çıkarım yöntemlerinden elde edilen özelliklerin, bağımsız ağlar tarafından işlenip daha sonra üst katmanda birleştirilmesi prensibine dayanır. Her iki yöntemden gelen özellikler, iki ayrı sinir ağı üzerinden geçirilerek bağımsız tahminler üretilir. Bu tahminler, son aşamada bir araya getirilerek nihai karar veya tahmin yapılır. Bu yapı, farklı özellik setlerinin birbirinden bağımsız olarak işlenmesine olanak tanır, böylece her yöntemin kendi katkısını daha net bir şekilde analiz etmek mümkündür. Özellikle karmaşık veri setlerinde, her yöntemden elde edilen bilgilerin farklı açılardan değerlendirilmesini sağlar ve sonuç olarak daha kapsamlı bir model çıktısı sunar.
- **Ortak Füzyon Tip II:** Bu yöntemde, farklı veri kaynaklarından veya yöntemlerden elde edilen özellikler, tek bir sinir ağına verilerek doğrudan birleştirilir ve işlenir. Burada özellik füzyonu, ağı içinde gerçekleştiği için yöntemler arası etkileşim daha sıkı ve bütünlüklü olur. Tek bir ağ kullanıldığı için hesaplama açısından daha verimli olabilir ve iki özellik kümesi arasındaki örtüşme ya da tamamlayıcı bilgi doğrudan işlenir. Bu yapı, özellikle derin öğrenme modellerinde, verinin tek bir ağ üzerinden öğrenilmesini sağladığından parametre optimizasyonunu kolaylaştırabilir ve modelin genel performansını artırabilir. Ancak, iki veri kümesinin birbirini doğrudan etkileyebileceği durumlarda, yöntemler arası bağımlılık iyi yönetilmelidir (Huang ve ark., 2020).

Bu iki yaklaşım, füzyon stratejilerinde farklı avantajlar sunar: Tip I daha modüler ve bağımsız analiz imkânı verirken, Tip II daha bütünlüklü ve etkileşimli bir çözüm sağlar.

3.6.2.1. Ortak Füzyon Tip I Stratejisinin Uygulanması

Ortak Füzyon Tip I kapsamında, manuel olarak çıkarılan geleneksel 35 adet özelliğin yanı sıra X-ışını görüntülerinden transfer öğrenme modelleri (ResNet-50, VGG16, EfficientNetb0, Xception ve DenseNet201) ile elde edilen derin özellikler de kullanılarak füzyon işlemleri gerçekleştirilmiştir.

X-ışını görüntülerinden sırasıyla ResNet-50, VGG16, EfficientNetb0, Xception ve DenseNet201 modelleri ile derin özellikler çıkarılırken, 224x224 boyutunda hazırlanan X-ışını görüntüleri modellere verilmiş ve modeller tarafından önemli özellikler çıkarılmıştır. Bu işlem için her modelin en son evrimsel katmanından önceki

global ortalama havuzlama (Global Average Pooling - GAP) katmanı kullanılmıştır. Ancak, elde edilen bu özelliklerin boyutunun yüksek olması nedeniyle, boyut azaltma işlemi PCA yöntemiyle gerçekleştirilmiştir. PCA uygulanarak, vektörel halde bulunan derin özelliklerin boyutu başlangıçta birkaç bin boyuttan 103'e indirgenmiştir. PCA, en fazla bilgiyi koruyacak şekilde bileşenleri seçmek amacıyla uygulanmış ve toplamda %85'lik varyansı temsil eden 103 bileşen seçilmiştir. Bu sayede, yüksek boyutlu verinin gereksiz ve tekrarlayan bilgileri elenmiş, daha kompakt bir veri yapısı elde edilmiştir.

Bu stratejide de, manuel olarak çıkarılan geleneksel 35 özellik yerine, ANOVA ve LDA ile elde edilen 20 boyutlu özellik vektörü kullanılmıştır. Böylece, veri boyutu azaltılarak hem derin öğrenme tabanlı hem de istatistiksel yöntemlerle çıkarılan özellikler daha etkili bir şekilde birleştirilmiştir.

Gerçekleştirilen işlemlerle, derin ve geleneksel özellikler ayrı sinir ağlarında işlenmek üzere hazırlanmıştır.

- **Derin özellikler**, Sinir Ağları 2 olarak adlandırılan bir sinir ağına verilmiştir. Bu sinir ağı şu bileşenlerden oluşmaktadır:
 - Giriş katmanı: 103 **nöron** (PCA ile indirgenmiş derin özellikler)
 - Tam bağlı (Dense) katman: **256 nöron**
 - Batch Normalization ve ReLU aktivasyonu
 - Dropout (0,5 oranında)
 - Tam bağlı (Dense) katman: **128 nöron**
 - Batch Normalization ve ReLU aktivasyonu
 - Dropout (0,5 oranında)
- **Geleneksel özellikler**, Sinir Ağları 1 adı verilen ayrı bir sinir ağına verilmiştir. Bu sinir ağı şu bileşenlerden oluşmaktadır:
 - Giriş katmanı: **20 nöron** (LDA ve ANOVA ile optimize edilmiş geleneksel özellikler)
 - Tam bağlı (Dense) katman: **64 nöron**
 - Batch Normalization ve ReLU aktivasyonu
 - Dropout (0,5 oranında)
 - Tam bağlı (Dense) katman: **32 nöron**
 - Batch Normalization ve ReLU aktivasyonu
 - Dropout (0,5 oranında)

Her iki sinir ağına verilen özelliklerin çıktıları özetlendikten sonra birleştirilmiştir. Başlangıçta, Sinir Ağları 1'e LDA ve ANOVA ile optimize edilmiş 20

özellik verilmiş, bu ağ bu özellikleri özetlemiştir. Aynı şekilde, Sinir Ağları 2'ye PCA ile indirgenmiş 103 özellik verilmiş, bu ağ da özellikleri özetleyerek çıktılar üretmiştir.

Bu durumda, her iki ağın çıktıları şu şekilde olmuştur:

- Sinir Ağları 1'in çıktısı: 50 özetlenmiş özellik (Ağın özetleme işleminden sonra 50 boyutlu bir çıktı elde edilmiştir.)
- Sinir Ağları 2'nin çıktısı: 50 özetlenmiş özellik (Benzer şekilde, PCA ile indirgenmiş görsel özelliklerden sonra bu ağ da 50 boyutlu bir çıktı üretmiştir.)

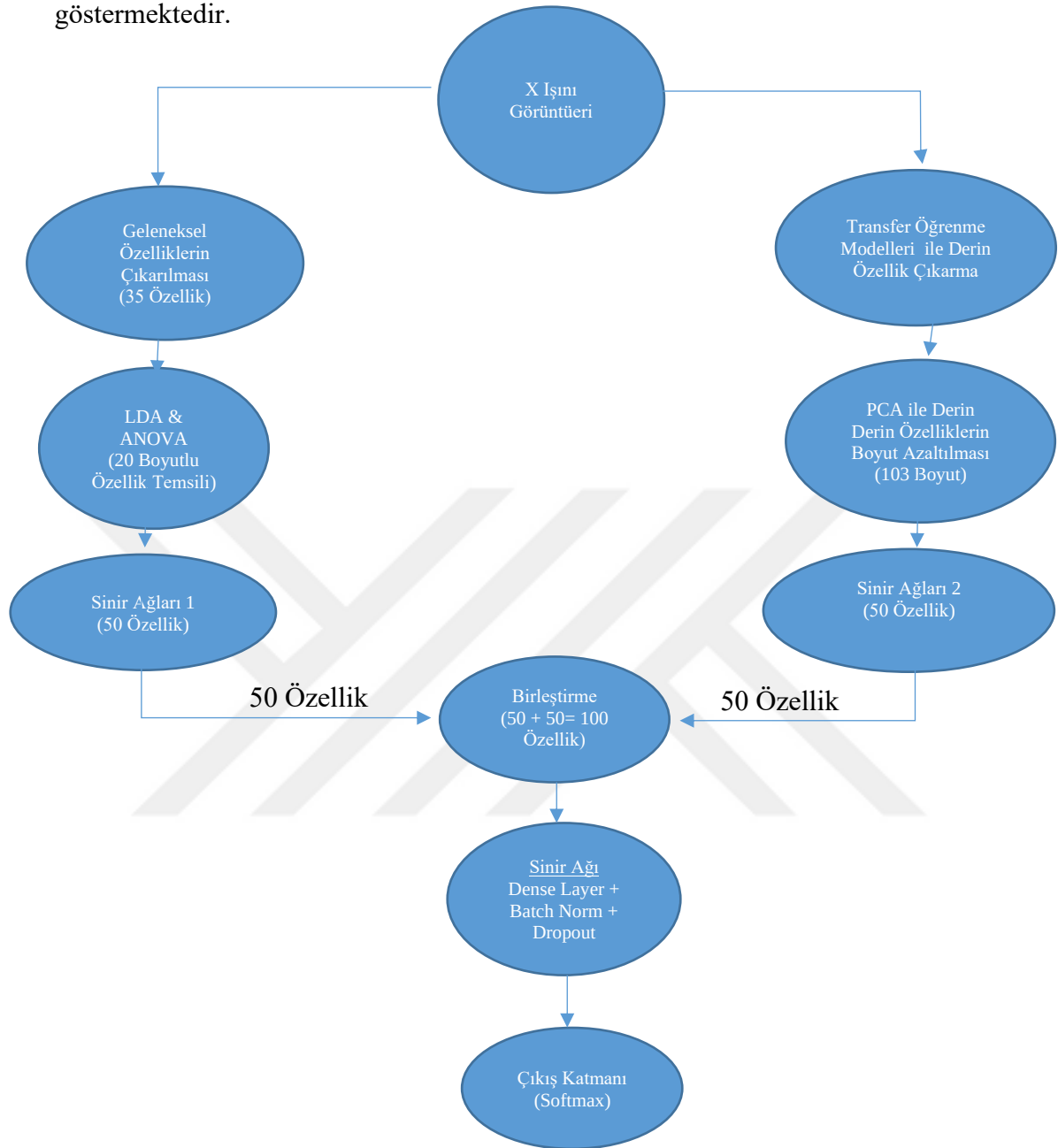
50 özellik seçiminde amaç, modelin gereksiz karmaşıklıktan kaçınarak, sadece veriyi en iyi şekilde temsil eden sayıda özelliği öğrenmesini sağlamaktır. Aşırı öğrenmeyi engellemek için modelin öğrendiği özelliklerin sayısının sınırlı tutulması ve genel doğruluğun artırılması sağlanmıştır. Bu sayede, modelin hem doğruluğu artırılmış hem de gereksiz öğrenmelerin önüne geçilmiştir. Bu çıktılar daha sonra birleştirme işlemi kullanılarak birleştirilmiştir. Sonuç olarak, her iki ağın çıktılarının birleştirilmesiyle toplamda 100 özetlenmiş özellik elde edilmiştir (50 özellik Sinir Ağları 1'den, 50 özellik Sinir Ağları 2'den).

Birleştirilen 100 boyutlu özellik vektörü, nihai sınıflandırma aşamasında kullanılmıştır. Doğrudan bu vektör üzerinden sınıflandırma yapmak mümkün olsa da, modelin performansı bu durumda sınırlı kalabilirdi. 100 özellik yalnızca başlangıç seviyesindeki bilgilere dayanıyordu ve modelin bu bilgiyi daha derinlemesine işlemeye ihtiyacı vardı. Bu nedenle, özelliklerin daha anlamlı ve soyut hale getirilmesi için ilk olarak 128 nöronlu bir dense katmanı kullanılmıştır. 100 nöronlu bir katman yerine 128 nöronun tercih edilme sebebi, modelin daha yüksek kapasiteye sahip olmasını sağlayarak daha karmaşık ilişkileri öğrenmesine yardımcı olmaktır.

Bu aşamadan sonra, modelin daha stabil çalışmasını sağlamak ve aşırı öğrenmeyi önlemek amacıyla batch normalization ve dropout işlemleri uygulanmıştır. Batch normalization, her katmandan sonra verilerin dağılımını normalize ederek ağın daha hızlı ve stabil öğrenmesini sağlarken, dropout rastgele nöronları geçici olarak devre dışı bırakarak modelin genelleme yeteneğini artırır. Kullanılan dropout oranı 0,3 olup, her eğitim adımında nöronların %30'u rastgele devre dışı bırakılmıştır.

Son aşamada, softmax aktivasyon fonksiyonu kullanılarak nihai sınıf tahmini gerçekleştirilmiştir. Bu yöntemle, X-ışını görüntülerinden ve manuel olarak çıkarılan özelliklerden faydalanılarak sağlıklı, şüpheli veya farklı derecelerdeki KOA gibi sınıfların belirlenmesi mümkün hale getirilmiştir.

Şekil 3.14 Ortak Tip I stratejisinin çalışmada uygulanmasına yönelik akışı göstermektedir.



Şekil 3.14. Ortak Füzyon Tip I Süreci

Ortak füzyon tip I, her bir transfer öğrenme modelinden elde edilen derin özellikler kullanılarak beş kez gerçekleştirilmiştir.

3.6.2.2. Ortak Füzyon Tip II Stratejisinin Uygulanması

Ortak Füzyon Tip II stratejisi uygulanırken, X-Işını görüntüleri transfer öğrenme modellerine (ResNet-50, VGG16, EfficientNetB0, Xception ve DenseNet201) verilerek

görsellerden yüksek seviye derin özellikler çıkarılmıştır. Boyutun yüksek olması nedeniyle PCA yöntemi her bir modelden elde edilen vektör halindeki özelliklere ayrı ayrı uygulanmış ve nihai olarak her bir modelden 103 boyutlu bir özellik vektörü elde edilmiştir. Manuel olarak elde edilen geleneksel özellikler yerine, LDA ve ANOVA ile optimize edilmiş 20 özellik kullanılmıştır. Sınıflandırma aşamasına geçilmeden önce, transfer öğrenme modelleri ile elde edilen ve boyutu azaltılmış 103 özellik ile optimize edilmiş 20 geleneksel özellik birleştirilerek 123 boyutlu tek bir özellik vektörü oluşturulmuştur. Ardından, bu birleşik özellik vektörleri, tek bir sinir ağı modeline girdi olarak verilmiştir.

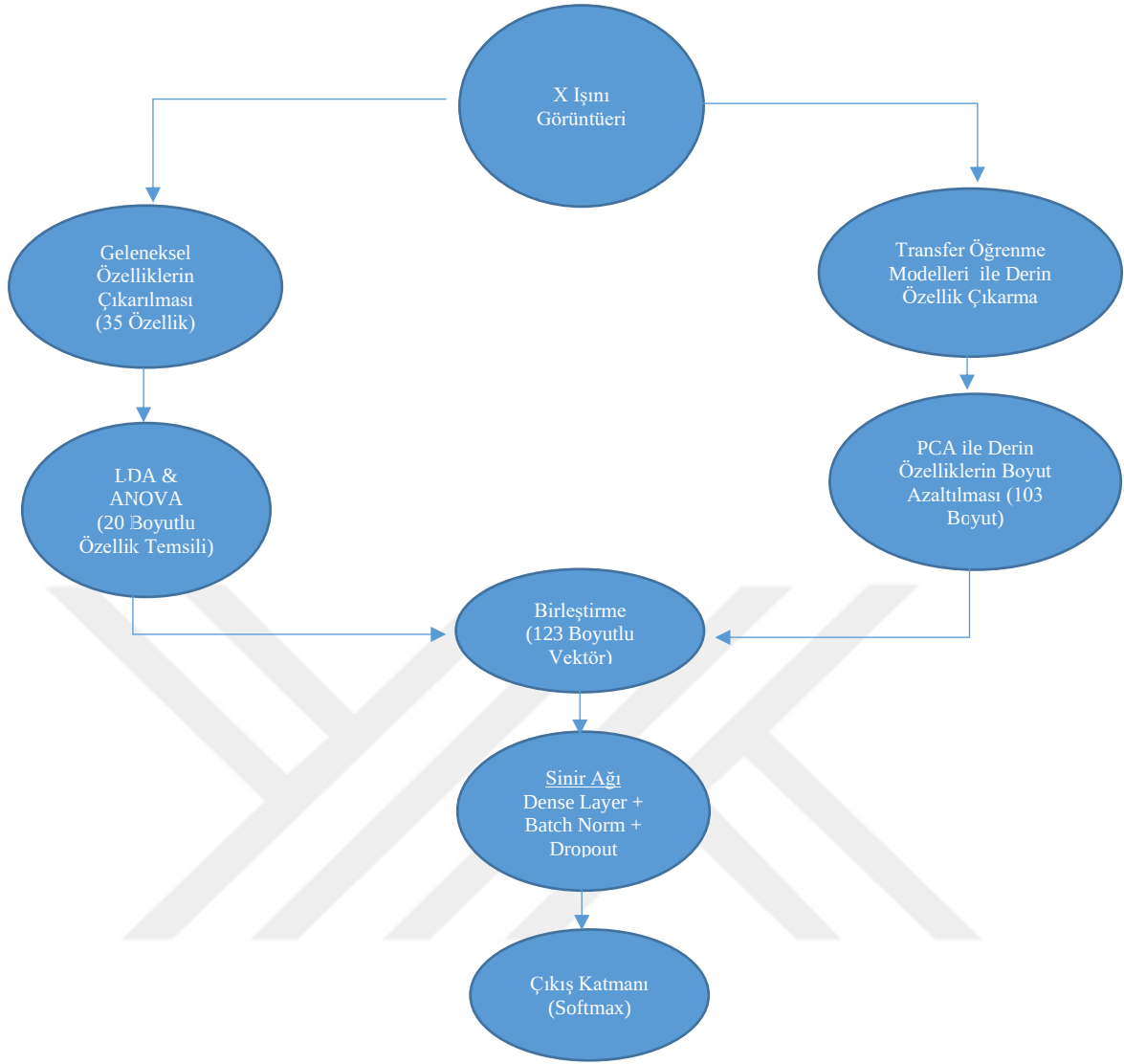
Sinir ağı mimarisinde, giriş katmanında her bir modelden elde edilen 103 boyutlu özellik vektörlerinin birleşiminden oluşan 123 boyutlu tek bir özellik vektörü kullanılmaktadır.

Sinir ağı şu bileşenlerden oluşmaktadır: Giriş katmanında, 123 boyutlu özellik vektörü modele giriş olarak verilmiştir. İlk gizli katmanda 512 nöron bulunmakta ve modelin yüksek seviyeli ilişkileri öğrenmesine yardımcı olmaktadır. İkinci gizli katmanda 256 nöron yer almakta olup, modelin daha derin katmanlarda özellikleri işleyerek daha kompakt bir temsil oluşturmasını sağlamaktadır. Çıkış katmanı ise 5 nörondan oluşmakta ve softmax aktivasyon fonksiyonu kullanılarak modelin her sınıfa ait olasılık değerlerini üretmesi sağlanmaktadır. Böylece model, girişte verilen özellik vektörünün hangi sınıfa ait olduğunu tahmin edebilmektedir.

Batch Normalization, her katmandan sonra uygulanarak modelin öğrenme sürecini hızlandırmakta ve daha stabil hale getirmektedir. Bu işlem, katmanların çıktısını normalize ederek verilerin belirli bir dağılım aralığında tutulmasını sağlar. Böylece modelin daha hızlı öğrenmesi ve aşırı öğrenmenin (overfitting) önlenmesi sağlanmaktadır. Batch Normalization için belirlenen öğrenme oranı 0,001 olup, bu değer veri kümesine ve uygulamaya göre optimize edilebilir.

Dropout yöntemi, modelin aşırı öğrenmesini engellemek için kullanılmıştır. Bu teknikte, her iterasyonda rastgele nöronlar devre dışı bırakılarak modelin daha genellenebilir hale gelmesi sağlanmaktadır. Kullanılan dropout oranı 0,3 olup, her seferinde nöronların %30'u rastgele devre dışı bırakılmaktadır. Böylece modelin belirli özelliklere fazla bağımlı olması engellenerek, daha sağlam bir öğrenme süreci elde edilmiştir.

Şekil 3.15 Ortak Füzyon Tip II sürecini göstermektedir.



Şekil 3.15. Ortak Füzyon Tip II Süreci

Ortak Füzyon Tip II, her bir transfer öğrenme modelinden çıkarılan derin özelliklerin geleneksel özelliklerle birleştirilmesi suretiyle beş kez uygulanmıştır.

3.7. Performans Değerlendirmesi

Performans değerlendirme, bir modelin doğruluk ve güvenilirliğini ölçmek için kullanılır. Çok sınıflı sınıflandırma problemlerinde, performans metrikleri her bir sınıf için ayrı ayrı hesaplanabilir ve ardından genel değerlendirme yapılır. Aşağıdaki metrikler, 5 sınıflı konfüzyon matrisi üzerinden hesaplanmaktadır.

Çok sınıflı bir konfüzyon matrisi, her bir satırın gerçek sınıfı ve her bir sütunun modelin tahmin ettiği sınıfı gösterdiği bir tablodur.

Örneğin, Şekil 3.16’da yer alan 5 sınıflı bir problemde matris aşağıdaki gibi ele alınır.

Gerçek \ Tahmin	Sınıf 0	Sınıf 1	Sınıf 2	Sınıf 3	Sınıf 4
Sınıf 0	TP0	FP0_1	FP0_2	FP0_3	FP0_4
Sınıf 1	FP1_0	TP1	FP1_2	FP1_3	FP1_4
Sınıf 2	FP2_0	FP2_1	TP2	FP2_3	FP2_4
Sınıf 3	FP3_0	FP3_1	FP3_2	TP3	FP3_4
Sınıf 4	FP4_0	FP4_1	FP4_2	FP4_3	TP4

Şekil 3.16. 5 sınıflı bir problem için örnek konfüzyon matrisi

Bu matriste:

- **Doğru Pozitif (TPi - True Positive):** Gerçek sınıfı i olan örneklerin doğru bir şekilde yine i olarak tahmin edilmesi.
- **Yanlış Negatif (FNi - False Negative):** Gerçek sınıfı i olan örneklerin, i dışındaki bir sınıf olarak yanlış tahmin edilmesi.
- **Yanlış Pozitif (FPi - False Positive):** Gerçek sınıfı i olmayan örneklerin, yanlış bir şekilde i olarak tahmin edilmesi.
- **Doğru Negatif (TNi - True Negative):** Gerçek sınıfı i olmayan örneklerin doğru bir şekilde i dışında bir sınıfa tahmin edilmesi.

Çok sınıflı durumda, her bir sınıf için ayrı ayrı metrikler hesaplanabilir ve bu metriklerin ortalamaları alınarak genel sonuçlara ulaşılır. İki yaygın ortalama yöntemi vardır:

1. **Makro Ortalama (Macro Averaging):** Her sınıf için metrikler ayrı ayrı hesaplanıp, sınıfların aritmetik ortalaması alınır. Tüm sınıflara eşit önem verir.
2. **Ağırlıklı Ortalama (Weighted Averaging):** Sınıfların örnek sayısına göre ağırlıklı ortalaması alınır. Daha fazla örnek içeren sınıflar daha fazla etkili olur. Özellikle sınıf dengesizliği olan ve büyük sınıfların daha kritik olduğu durumlarda kullanılır.

Doğruluk (Accuracy):

Doğru tahmin edilen örneklerin, toplam örnek sayısına oranıdır. Genel bir başarı ölçütüdür. Denklem 3.32 ile hesaplanır.

$$\text{Doğruluk} = \frac{\sum_{i=0}^4 TP_i}{\text{Toplam örnek sayısı}} \quad (3.32)$$

Duyarlılık (Recall):

Denklem 3.33’de görüldüğü gibi gerçek pozitiflerin, toplam pozitif örnek sayısına oranıdır. Modelin gerçek pozitifleri ne kadar doğru yakaladığını ölçer.

$$\text{Duyarlılık}_i = \frac{TP_i}{TP_i + FN_i} \quad (3.33)$$

Kesinlik (Precision):

Doğru pozitif tahminlerin, toplam pozitif tahminlere oranıdır (Denklem 3.34). Modelin pozitif tahminlerinin güvenilirliğini ifade eder.

$$\text{Kesinlik}_i = \frac{TP_i}{TP_i + FP_i} \quad (3.34)$$

F1-Skoru (F1-Score):

Kesinlik ve duyarlılık arasındaki dengeyi ifade eden bir ölçüttür. Harmonik ortalama kullanılarak Denklem 3.35’deki gibi hesaplanır.

$$F1_i = 2 * \frac{\text{Kesinlik}_i * \text{Duyarlılık}_i}{\text{Kesinlik}_i + \text{Duyarlılık}_i} \quad (3.35)$$

Bu metrikler, modelin her açıdan performansını anlamamızı ve modelin tüm sınıflar üzerindeki genel başarısını değerlendirmenizi sağlar. Doğruluk genel başarıyı ölçerken, duyarlılık modelin her sınıfı ne kadar doğru tanıdığını gösterir. Kesinlik, tahmin edilen pozitif sonuçların ne kadar doğru olduğunu belirtir. F1-Skoru ise duyarlılık ve kesinlik arasındaki dengeyi ölçer.

Bu metrikler sayesinde, modelin hangi sınıflarda daha başarılı olduğu ve hangi sınıflarda hata yaptığı detaylı bir şekilde analiz edilebilir.

4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA

Bu tez çalışması, X-ışını görüntülerini kullanarak KOA derecelendirmesini doğru ve güvenilir bir şekilde gerçekleştirmeye odaklanmıştır. Bu amaç doğrultusunda, X-ışını görüntülerine yönelik ön işleme adımları, geleneksel ve derin öğrenme tabanlı özellik çıkarma süreçleri, füzyon stratejileri entegre bir yaklaşımla bir araya getirilmiştir. İlk olarak, X-ışını görüntüleri üzerinde, yeniden boyutlandırma, ROI belirleme, kontrast iyileştirme, veri atırma ve veri dengesizliği giderme gibi ön işleme adımları gerçekleştirilmiştir. Daha sonra, görüntülerden morfolojik, GLCM, istatistiksel, HOG, LBP, Dalgacık Dönüşümü, zernike momenti, çevre ve alan özellikleri olmak üzere toplamda geleneksel 35 farklı görüntü özelliği manuel olarak çıkarılmış ve bu özelliklerden kapsamlı bir özellik vektörü oluşturulmuştur. Manuel olarak çıkarılan bu özelliklerin optimize edilmesi için LDA ve ANOVA tekniklerinden yararlanılmıştır.

Füzyon stratejileri kapsamında, X-ışını görüntülerinden farklı veri temsilleri oluşturulmuş ve boyut azaltımı amacıyla PCA ile işlenmiştir. Erken füzyon Tip II kapsamında, yalnızca X-ışını görüntüleri vektör haline getirilmiş ve bu vektörler PCA uygulanarak temsil edilmiştir. Bununla birlikte, ortak füzyon Tip I ve Tip II kapsamında, X-ışını görüntülerinden transfer öğrenme modelleri aracılığıyla derin yüksek seviyeli özellikler çıkarılmış, ardından bu özellikler PCA ile boyut azaltma işlemine tabi tutulmuştur. Son aşamada, manuel olarak çıkarılan ve optimize edilen geleneksel özellikler ile transfer öğrenme modellerinden elde edilen bu derin özellikler birleştirilerek daha zengin ve anlamlı veri temsilleri oluşturulmuştur.

KOA derecelerinin füzyon stratejileri ile belirlenebilmesi amacıyla, elde edilen birleşik özellik vektörleri ANN mimarisi kullanılarak sınıflandırılmıştır. Model eğitimi için dengelenmiş veri seti, %70'i eğitim, %15'i doğrulama ve %15'i test olmak üzere üç gruba rastgele ayrılmıştır. Tüm füzyon stratejileri, bu veri ayrımı doğrultusunda ANN tabanlı sınıflandırma yapıları üzerinden değerlendirilmiş ve karşılaştırılmıştır.

Çizelge 4.1, eğitim, doğrulama ve test grubundaki veri miktarlarını göstermektedir.

Çizelge 4.1. Eğitim, doğrulama ve test grupları veri sayıları

Veri Kümesi	Toplam Veri Sayısı	Derece 0 (Sağlıklı)	Derece 1 (Şüpheli)	Derece 2 (Hafif)	Derece 3 (Orta)	Derece 4 (Ağır)
Eğitim Seti	9107	1821	1821	1821	1822	1822
Doğrulama Seti	1951	390	390	391	390	390
Test Seti	1952	391	391	390	390	390

Füzyon stratejilerinde kullanılan transfer öğrenme modellerinin yapılandırma detayları ve hiperparametreleri ise Çizelge 4.2’de sunulmaktadır.

Çizelge 4.2. Füzyon Stratejilerinde Kullanılan Transfer Öğrenme Modellerinin Yapılandırma Detayları ve Hiperparametreleri

Model	Katman Sayısı	Öğrenme Oranı	Optimizasyon Yöntemi	Diğer Parametreler
EfficientNetb0	16	0,001	Adam	Batch Size: 32, Epochs: 50
DenseNet201	201	0,0005	SGD	Batch Size: 64, Epochs: 100
Xception	71	0,0001	Adam	Batch Size: 32, Epochs: 50
VGG16	16	0,0001	Adam	Batch Size: 64, Epochs: 100
Resnet 50	50	0,0001	Adam	Batch Size: 32, Epochs: 50

4.1. Erken Füzyon Tip II Sonuçları

Çalışmada, KOA derecelerinin sağlıklı, şüpheli, hafif, orta, ağır olarak belirlenebilmesi için ilk olarak Erken Füzyon Tip II stratejisinden yararlanılmıştır. Bu stratejide, manuel olarak çıkarılan ve LDA ile ANOVA testleri sonucunda optimize edilen 20 özellik, PCA yöntemi ile boyutu küçültülen X-ışını görüntü temsilleri ile birleştirilmiştir. Böylece, geleneksel ve görüntü tabanlı özelliklerin bütünleştirilmesiyle daha zengin ve anlamlı bir veri temsili oluşturulmuştur.

Oluşturulan 117 boyutlu birleşik özellik vektörleri, sabit bir ANN mimarisi kullanılarak KOA derecelerini belirlemek üzere sınıflandırılmıştır. Bu süreçte, ANN yapısının başlangıç ağırlıkları farklı transfer öğrenme modellerinden (ResNet-50, VGG16, EfficientNetB0, Xception ve DenseNet201) alınarak belirlenmiş, böylece farklı model mimarilerinden gelen bilgi çeşitliliği sınıflandırma sürecine entegre edilmiştir.

Çizelge 4.3, Çizelge 4.4, Çizelge 4.5, Çizelge 4.6 ve Çizelge 4.7, sırasıyla Erken Füzyon Tip II stratejisinde oluşturulan birleşik özellik vektörlerinin, başlangıç ağırlıkları farklı transfer öğrenme modelleri (ResNet-50, VGG16, EfficientNetB0,

Xception ve DenseNet201) kullanılarak başlatılan sabit ANN mimarisi ile sınıflandırılması sonucunda elde edilen performans metriklerini göstermektedir.

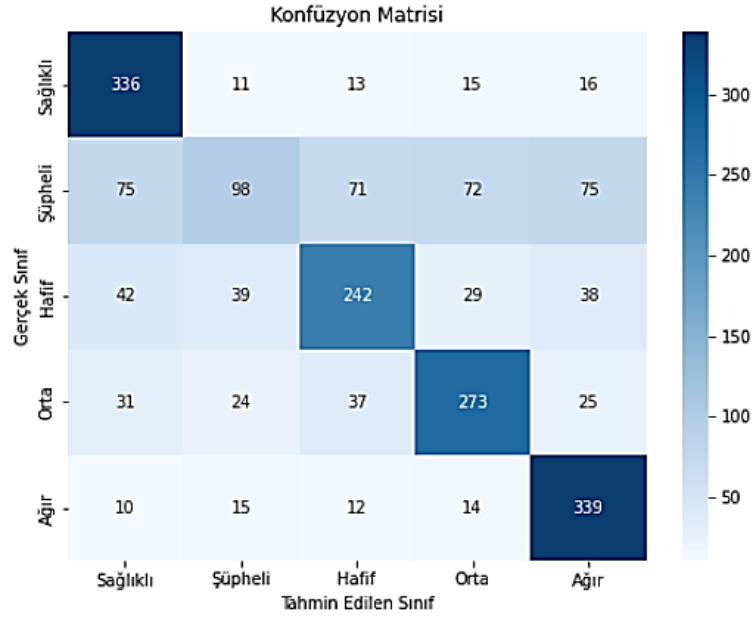
Çizelge 4.3. Erken Füzyon Tip II: ResNet-50 Tabanlı Başlangıç Ağırlıkları ile Sınıflandırma Sonuçları

Sınıf	Kesinlik	Duyarlılık	F1-Skoru	Sınıflandırma Doğruluğu
Sağlıklı	0,6801	0,8593	0,7593	0,6598
Şüpheli	0,5240	0,2506	0,3391	
Hafif	0,6453	0,6205	0,6326	
Orta	0,6774	0,7000	0,6885	
Ağır	0,6876	0,8692	0,7678	
Ortalama	0,6429	0,6599	0,6374	

Çizelge 4.3'te sunulan sonuçlara göre, ResNet-50 modelinden elde edilen başlangıç ağırlıkları ile eğitilen ANN kullanılarak gerçekleştirilen sınıflandırma performansı KOA dereceleri arasında farklılık göstermiştir. Ağır KOA derecesinde tüm metriklerde en yüksek başarı sağlanmış (F1-skoru: 0,7678, duyarlılık: 0,8692) ve model bu dereceleri ayırt etmede etkili olmuştur. Hafif ve orta dereceler için sırasıyla 0,6326 ve 0,6885 F1-skorları elde edilmiş olup, bu dereceler için makul bir sınıflandırma başarısı gözlemlenmiştir.

Buna karşın, şüpheli KOA derecesinde modelin performansı belirgin şekilde düşük kalmıştır (F1-skoru: 0,3391; duyarlılık: 0,2506). Özellikle şüpheli derecedeki düşük duyarlılık değeri, modelin bu vakaları doğru sınıflandıramadığını ortaya koymaktadır.

Genel ortalama değerlere bakıldığında, kesinlik 0,6429, duyarlılık 0,6599, F1-skoru 0,6374 ve doğruluk 0,6598 olarak hesaplanmıştır. Elde edilen sonuçlar, ağır KOA derecesinde başarılı bir performans sergilendiğini, ancak şüpheli derecede önemli iyileştirmelere ihtiyaç duyulduğunu göstermektedir. Şekil 4.1, ResNet-50 başlangıç ağırlıklarıyla eğitilen yapay sinir ağının test verileri üzerinde KOA derecelendirmesine yönelik sınıflama başarımını konfüzyon matrisi olarak özetlemektedir.



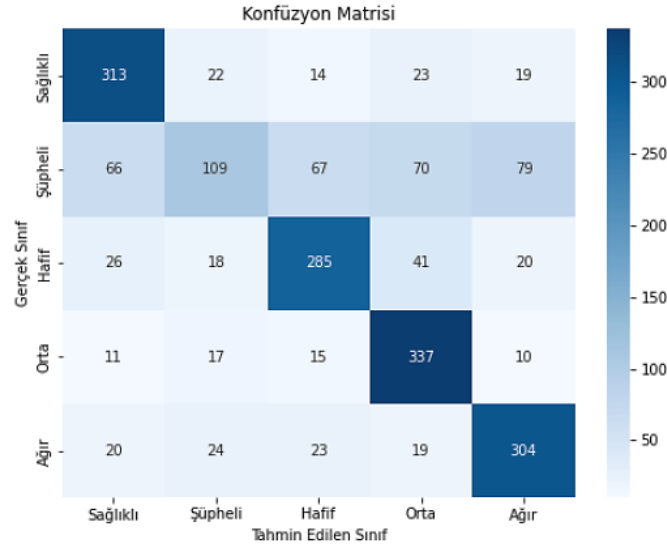
Şekil 4.1. Erken Füzyon Tip II stratejisi ile ResNet-50 başlangıç ağırlıkları kullanılarak test verilerinde gerçekleştirilen KOA derecelendirmesi sonuçlarının konfüzyon matrisi.

Çizelge 4.4. Erken Füzyon Tip II: VGG16 Tabanlı Başlangıç Ağırlıkları ile Sınıflandırma Sonuçları

Sınıf	Kesinlik	Duyarlılık	F1-Skoru	Sınıflandırma Doğruluğu
Sağlıklı	0,7178	0,8005	0,7569	0,6905
Şüpheli	0,5736	0,2787	0,3752	
Hafif	0,7054	0,7307	0,7178	
Orta	0,6877	0,8641	0,7659	
Ağır	0,7037	0,7794	0,7396	
Ortalama	0,6776	0,6907	0,6711	

Çizelge 4.4'te sunulan sonuçlar incelendiğinde, VGG16'dan elde edilen başlangıç ağırlıklarıyla eğitilen ANN ile gerçekleştirilen KOA derecelendirmesinde sınıflar arasında belirgin performans farklılıkları gözlemlenmiştir. Özellikle orta ve ağır derecelerde sırasıyla 0,7659 ve 0,7396 F1-skorları ile yüksek başarı elde edilmiştir. Orta derecede duyarlılık değeri 0,8641 ile en yüksek seviyeye ulaşırken, şüpheli derecede duyarlılık 0,2787 gibi oldukça düşük bir seviyede kalmıştır. Şüpheli derece için hesaplanan F1-skoru 0,3752 ve duyarlılık değeri 0,2787 olup, bu grubun doğru şekilde ayırt edilmesinde güçlük yaşandığı anlaşılmaktadır. Sağlıklı, hafif ve orta dereceler için ise daha dengeli sonuçlar elde edilmiştir. Genel ortalamalar dikkate alındığında, kesinlik 0,6776, duyarlılık 0,6907, F1-skoru 0,6711 ve doğruluk 0,6905 olarak hesaplanmıştır. Bu sonuçlar, şüpheli derece için performansın artırılmasına yönelik iyileştirmelere ihtiyaç olduğunu göstermektedir. Şekil 4.2, Erken Füzyon Tip II stratejisi ve VGG16

başlangıç ağırlıkları kullanılarak test verileri üzerinde gerçekleştirilen KOA derecelendirmesine ait konfüzyon matrisini sunmaktadır.



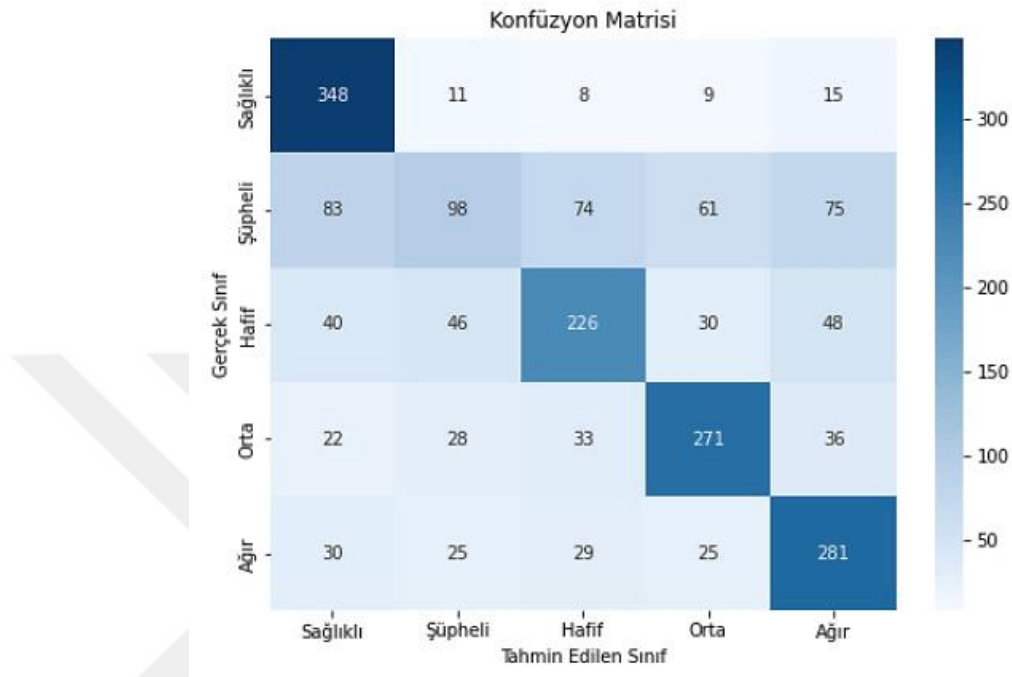
Şekil 4.2. Erken Füzyon Tip II stratejisi ile VGG16 başlangıç ağırlıkları kullanılarak test verilerinde gerçekleştirilen KOA derecelendirmesi sonuçlarının konfüzyon matrisi

Çizelge 4.5. Erken Füzyon Tip II: EfficientNetB0Tabanlı Başlangıç Ağırlıkları ile Sınıflandırma

Sınıf	Kesinlik	Duyarlılık	F1-Skoru	Sınıflandırma Doğruluğu
Sağlıklı	0,6653	0,8900	0,7614	0,6270
Şüpheli	0,4711	0,2506	0,3272	
Hafif	0,6108	0,5794	0,5947	
Orta	0,6843	0,6948	0,6895	
Ağır	0,9175	0,7205	0,6650	
Ortalama	0,6098	0,6271	0,6076	

Çizelge 4.5'te sunulan sonuçlar incelendiğinde, EfficientNetB0'dan alınan başlangıç ağırlıkları ile ANN kullanılarak gerçekleştirilen KOA derecelendirmesinde sınıflar arasında belirgin performans farklılıkları gözlemlenmiştir. Ağır ve orta derecelerde sırasıyla 0,6650 ve 0,6895 F1-skorları ile diğer derecelere kıyasla daha yüksek başarı elde edilmiştir. Sağlıklı derecede duyarlılık değeri 0,8900 gibi yüksek bir seviyeye ulaşırken, şüpheli derecede duyarlılık 0,2506 ve F1-skoru 0,3272 gibi düşük değerler elde edilmiştir. Bu sonuçlar, şüpheli vakaların doğru sınıflandırılmasında güçlük yaşandığını göstermektedir. Hafif derecede de duyarlılık 0,5794 seviyesinde olup, iyileştirilmesi gereken alanlar arasında yer almaktadır. Genel performans değerlendirildiğinde, ortalama kesinlik 0,6098, duyarlılık 0,6271, F1-skoru 0,6076 ve doğruluk 0,6270 olarak hesaplanmıştır. Bulgular, özellikle şüpheli ve hafif dereceler

için sınıflandırma başarısının artırılması gerektiğini ortaya koymaktadır. Şekil 4.3, Erken Füzyon Tip II stratejisi kapsamında EfficientNetB0 başlangıç ağırlıkları kullanılarak test verileri üzerinde gerçekleştirilen KOA derecelendirmesine ait konfüzyon matrisini göstermektedir.



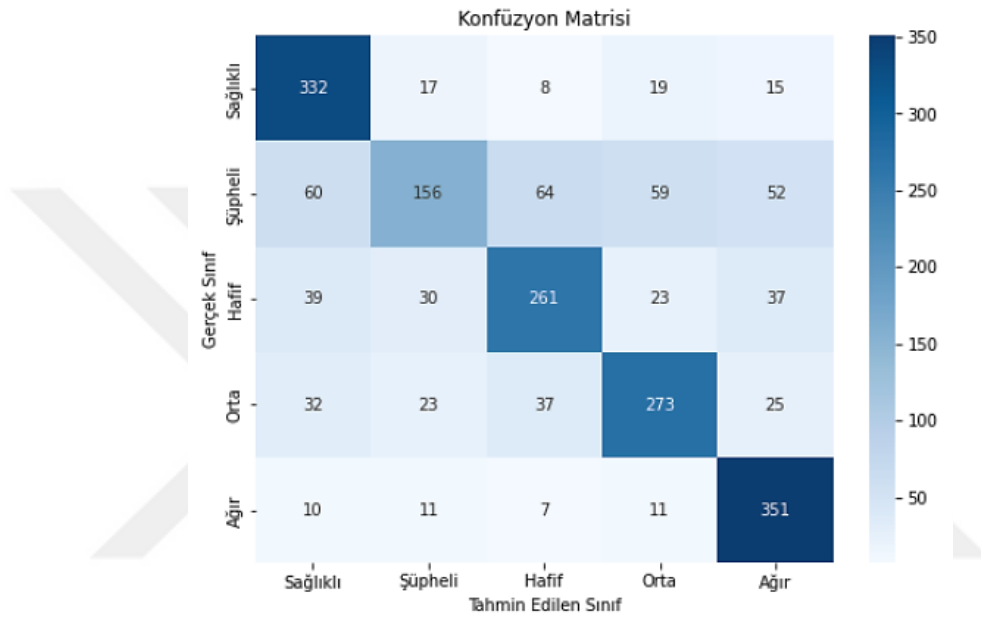
Şekil 4.3. Erken Füzyon Tip II stratejisi ile EfficientNetb0 başlangıç ağırlıkları kullanılarak test verilerinde gerçekleştirilen KOA derecelendirmesi sonuçlarının konfüzyon matrisi.

Çizelge 4.6. Erken Füzyon Tip II: Xception Tabanlı Başlangıç Ağırlıkları ile Sınıflandırma

Sınıf	Kesinlik	Duyarlılık	F1-Skoru	Sınıflandırma Doğruluğu
Sağlıklı	0,7019	0,8491	0,7685	0,7033
Şüpheli	0,6582	0,3989	0,4968	
Hafif	0,6923	0,6692	0,6805	
Orta	0,7090	0,7000	0,7045	
Ağır	0,7312	0,9000	0,8068	
Ortalama	0,6985	0,7034	0,6914	

Çizelge 4.6’da sunulan sonuçlar incelendiğinde, Erken Füzyon Tip II stratejisi kapsamında Xception’den alınan başlangıç ağırlıkları ile eğitilen ANN kullanılarak gerçekleştirilen KOA derecelendirmesinde, ağır ve orta dereceler en yüksek başarıyı göstermiştir. Ağır derecede F1-skoru 0,8068, orta derecede ise 0,7045 olarak hesaplanmış olup, her iki derecede de başarılı bir ayırım yapılabildiği görülmektedir. Sağlıklı ve hafif derecelerde F1-skorları sırasıyla 0,7685 ve 0,6805 düzeyinde bulunurken, şüpheli derecede 0,4968 gibi daha düşük bir F1-skoru elde edilmiştir. Bu

durum, şüpheli derecenin doğru tanımlanmasında güçlük yaşandığını göstermektedir. Genel ortalamalar değerlendirildiğinde, kesinlik 0,6985, duyarlılık 0,7034, F1-skoru 0,6914 ve doğruluk 0,7033 olarak hesaplanmıştır. Sonuçlar, genel performansın iyi olduğunu ancak şüpheli derecede ayırım yeteneğinin artırılması gerektiğini göstermektedir. Şekil 4.4, Erken Füzyon Tip II stratejisi kapsamında Xception başlangıç ağırlıkları kullanılarak test verileri üzerinde gerçekleştirilen KOA derecelendirmesine ait konfüzyon matrisini göstermektedir.



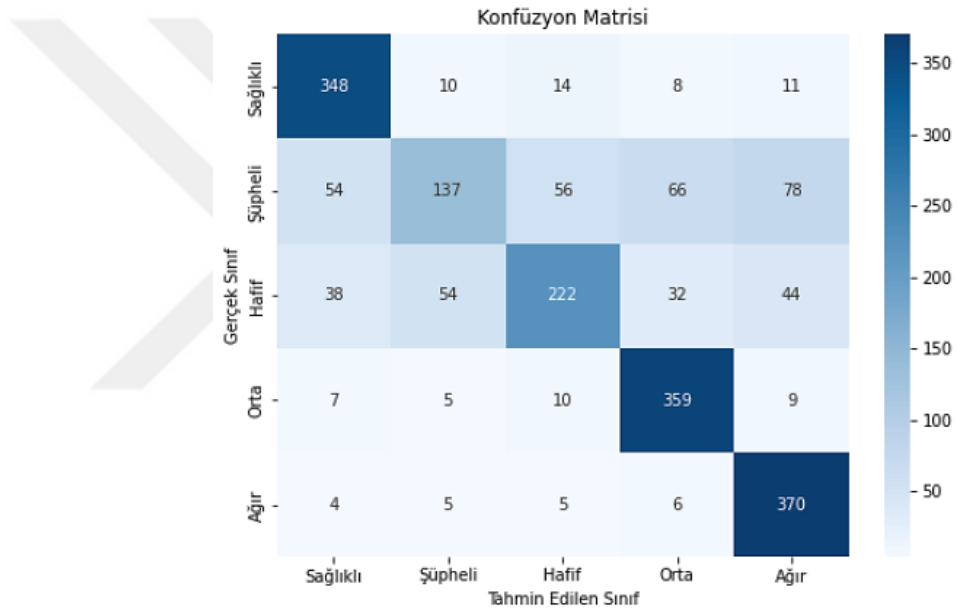
Şekil 4.4. Erken Füzyon Tip II stratejisi ile Xception başlangıç ağırlıkları kullanılarak test verilerinde gerçekleştirilen KOA derecelendirmesi sonuçlarının konfüzyon matrisi

Çizelge 4.7. Erken Füzyon Tip II: DenseNet201 Tabanlı Başlangıç Ağırlıkları ile Sınıflandırma Sonuçları

Sınıf	Kesinlik	Duyarlılık	F1-Skoru	Sınıflandırma Doğruluğu
Sağlıklı	0,7716	0,8900	0,8266	0,7356
Şüpheli	0,6492	0,3503	0,4551	
Hafif	0,7231	0,5692	0,6370	
Orta	0,7622	0,9205	0,8339	
Ağır	0,7226	0,9487	0,8203	
Ortalama	0,7257	0,7357	0,7146	

Çizelge 4.7’de sunulan sonuçlar incelendiğinde, Erken Füzyon Tip II stratejisi kapsamında DenseNet201’den alınan başlangıç ağırlıkları ile eğitilen ANN kullanılarak gerçekleştirilen KOA derecelendirmesinde, ağır ve orta dereceler en yüksek başarıyı göstermiştir. Ağır derecede F1-skoru 0,8203, orta derecede ise 0,8339 olarak

hesaplanmış olup, her iki derecede de başarılı bir ayırım yapılabildiği görülmektedir. Sağlıklı ve hafif derecelerde F1-skorları sırasıyla 0,8266 ve 0,6370 düzeyinde bulunurken, şüpheli derecede 0,4551 ile en düşük F1-skoru elde edilmiştir. Bu durum, şüpheli derecenin doğru tanımlanmasında güçlük yaşandığını göstermektedir. Genel ortalamalar değerlendirildiğinde, kesinlik 0,7257, duyarlılık 0,7357, F1-skoru 0,7146 ve doğruluk 0,7356 olarak hesaplanmıştır. Sonuçlar, genel performansın iyi olduğunu ancak şüpheli derecede ayırım yeteneğinin artırılması gerektiğini göstermektedir. Şekil 4.5, Erken Füzyon Tip II stratejisi kapsamında DenseNet201 başlangıç ağırlıkları kullanılarak test verileri üzerinde gerçekleştirilen KOA derecelendirmesine ait konfüzyon matrisini göstermektedir.



Şekil 4.5. Erken Füzyon Tip II stratejisi ile DenseNet201 başlangıç ağırlıkları kullanılarak test verilerinde gerçekleştirilen KOA derecelendirmesi sonuçlarının konfüzyon matrisi

Erken Füzyon Tip II stratejisi kullanılarak gerçekleştirilen analizde, ResNet-50, VGG16, EfficientNetB0, Xception ve DenseNet201 modellerinden elde edilen başlangıç ağırlıkları ile başlatılan ANN yapıları değerlendirilmiştir. Elde edilen sonuçlar incelendiğinde, her bir modelin KOA derecelendirmesinde farklı güçlü ve zayıf yönler sergilediği görülmektedir.

- **Ağır ve orta** KOA derecelerinde tüm modeller yüksek başarı göstermiştir. Özellikle DenseNet201 ve Xception tabanlı ANN yapıları, bu sınıflarda en yüksek F1-skorlarına ulaşarak öne çıkmıştır.

- **Şüpheli** derecede, tüm modellerin performansı görece düşük kalmıştır. Bu durum, şüpheli vakaların diğer derecelerle yapısal olarak benzerlik göstermesinden kaynaklanmakta olup, modelin bu sınıfları ayırt etmede zorlandığını göstermektedir. Özellikle EfficientNetB0 ve ResNet-50 tabanlı ANN yapılarının şüpheli sınıftaki duyarlılık ve F1-skorları diğer modellere kıyasla daha düşük bulunmuştur.
- **Hafif ve sağlıklı** sınıflarında ise DenseNet201 ve Xception modellerinden alınan başlangıç ağırlıkları ile başlatılan ANN yapıları, en yüksek F1-skorlarına ulaşmıştır. Bu sonuçlar, ilgili modellerin bu sınıflar arasında ayırım yapmada daha başarılı olduğunu göstermektedir.
- Genel ortalama değerlere bakıldığında, DenseNet201 tabanlı ANN yapısı en dengeli ve yüksek performansı sunmuştur. Bu model, %73,56 doğruluk, %71,46 F1-skoru, %72,57 kesinlik ve %73,57 duyarlılık değerleri ile diğer modellere göre daha üstün bir başarı sergilemiştir.
- **En başarılı model:** DenseNet201 tabanlı ANN yapısı, genel doğruluk ve F1-skoru açısından en başarılı sonuçları elde etmiştir. Xception tabanlı ANN yapısı da benzer şekilde güçlü bir performans göstererek hafif ve sağlıklı sınıflarda dikkat çekici sonuçlar vermiştir.
- **İyileştirme gereken alanlar:** Şüpheli sınıfta duyarlılık ve F1-skorlarının görece düşük kalması, tüm modeller için geliştirilmesi gereken önemli bir alan olarak değerlendirilmektedir.

Sonuç olarak, DenseNet201 modelinden elde edilen başlangıç ağırlıkları ile oluşturulan ANN yapısı, Erken Füzyon Tip II kapsamında, KOA derecelendirmesinde en başarılı sonuçları vermiştir. Bununla birlikte, şüpheli sınıfın doğruluk ve duyarlılık oranlarının artırılması için ileri düzey optimizasyon tekniklerinin uygulanması gerektiği değerlendirilmektedir.

4.2. Ortak Füzyon Tip I Sonuçları

Erken Füzyon Tip II stratejisinin ardından, X-ışını görüntülerinden manuel olarak çıkarılan geleneksel görüntü özellikleri ile transfer öğrenme modellerinin ara katmanlarından elde edilen derin özellikler, Ortak Füzyon Tip I yaklaşımı kapsamında entegre edilmiştir. Bu aşamada, sinir ağlarının orta düzey temsillerinden elde edilen özelliklerin geleneksel özelliklerle bütünleştirilmesi hedeflenmiş ve böylece KOA

derecelerinin “sağlıklı”, “şüpheli”, “hafif”, “orta” ve “ağır” olmak üzere beş kategoriye ayrılması sağlanmıştır.

Sonuçlar, sırasıyla ResNet-50, VGG16, EfficientNetB0, Xception ve DenseNet201 modellerinden elde edilen derin özelliklerin Ortak Füzyon Tip I stratejisi ile entegrasyonu sonrası Çizelge 4.8, Çizelge 4.9, Çizelge 4.10, Çizelge 4.11 ve Çizelge 4.12’de sunulmuştur.

Çizelge 4.8. Ortak Füzyon Tip I: Geleneksel ve ResNet-50 Tabanlı Derin Özelliklerle Sınıflandırma Sonuçları

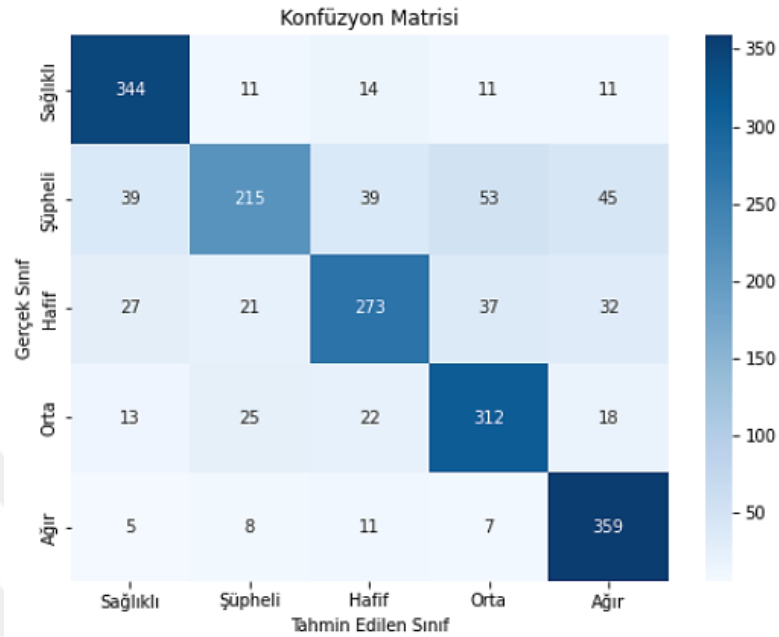
Sınıf	Kesinlik	Duyarlılık	F1-Skoru	Sınıflandırma Doğruluğu
Sağlıklı	0,8037	0,8797	0,8400	0,7699
Şüpheli	0,7678	0,5498	0,6408	
Hafif	0,7604	0,7000	0,7289	
Orta	0,7428	0,8000	0,7703	
Ağır	0,7720	0,9205	0,8397	
Ortalama	0,7693	0,7700	0,7639	

Çizelge 4.8’de sunulan sonuçlar incelendiğinde, ağır ve orta KOA derecelerinde yüksek performans elde edildiği görülmektedir. Ağır derece sınıfında 0,8397 ve orta derece sınıfında 0,7703 F1-skoruna ulaşılmış olup, bu dereceler doğru bir şekilde ayırt edilebildiği anlaşılmaktadır. Sağlıklı sınıf için elde edilen 0,8400 F1-skoru da bu kategoride güçlü bir performans sergilendiğini göstermektedir. Hafif KOA derecesinde 0,7289 F1-skoru ile orta düzeyde bir ayırım başarısı sağlanırken, şüpheli sınıfta performansın düştüğü gözlemlenmiştir. Şüpheli sınıf için 0,6408 F1-skoru ve 0,5498 duyarlılık değeri hesaplanmış, bu da şüpheli vakaların doğru şekilde sınıflandırılmasında zorluk yaşandığını ortaya koymuştur.

Genel performans değerlendirildiğinde, ResNet-50 modelinden elde edilen derin özelliklerin geleneksel özelliklerle Ortak Füzyon Tip I kapsamında birleştirilmesi sonucunda, ortalama kesinlik 0,7693, duyarlılık 0,7700, F1-skoru 0,7639 ve doğruluk 0,7699 olarak elde edilmiştir. Bu sonuçlar, KOA derese sınıfları arasında dengeli bir sınıflandırma başarısı sağlandığını göstermektedir. Ancak, şüpheli KOA derecesinde sınırlı duyarlılık ve F1-skoru değerleri, bu sınıfın daha etkin şekilde ayırt edilebilmesi için iyileştirme çalışmaları yapılması gerektiğini ortaya koymaktadır. Öte yandan, ağır ve orta derecelerdeki yüksek başarı oranları, genel sistem performansını olumlu yönde etkileyen başlıca unsurlar olmuştur.

Şekil 4.6’da, test verisi üzerinde, ResNet-50 modeliyle elde edilen derin özellikler ile geleneksel özelliklerin Ortak Füzyon Tip I stratejisi kapsamında

birleştirilmesi sonucu gerçekleştirilen KOA derecelendirmesine ait konfüzyon matrisi sunulmuştur.



Şekil 4.6. Ortak Füzyon Tip I ile ResNet-50 derin ve geleneksel özelliklerin birleşimi sonucu test verisinde elde edilen KOA derecelendirme konfüzyon matrisi

Çizelge 4.9. Ortak Füzyon Tip I: Geleneksel ve VGG16 Tabanlı Derin Özelliklerle Sınıflandırma Sonuçları

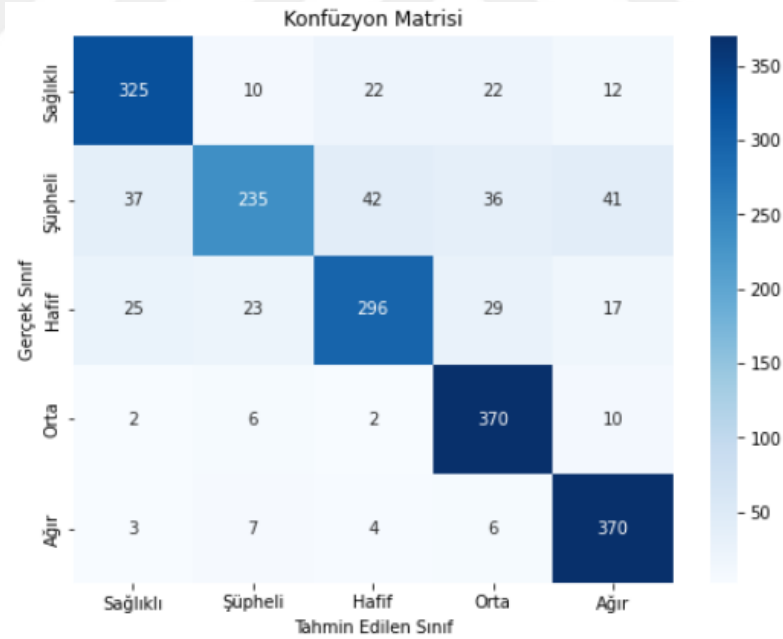
Sınıf	Kesinlik	Duyarlılık	F1-Skoru	Sınıflandırma Doğruluğu
Sağlıklı	0,8290	0,8312	0,8301	0,8176
Şüpheli	0,8362	0,6010	0,6994	
Hafif	0,8087	0,7589	0,7830	
Orta	0,7991	0,9484	0,8675	
Ağır	0,8222	0,9487	0,8809	
Ortalama	0,8190	0,8177	0,8122	

Çizelge 4.9’da sunulan sonuçlara göre, Ortak Füzyon Tip I stratejisi kapsamında, VGG16 modelinin ara katmanlarından çıkarılan derin özellikler ile geleneksel özelliklerin birleştirilmesiyle gerçekleştirilen sınıflandırmada ağır ve orta KOA derecelerinde yüksek performans elde edilmiştir. Ağır KOA derecesinde 0,8809 ve orta derecede 0,8675 F1-skoruna ulaşılmış olup, bu derecelerin başarılı bir şekilde ayırt edilebildiği gözlemlenmektedir. Sağlıklı bireyleri temsil eden derece grubunda da F1-skoru 0,8301 olup, kesinlik (0,8290) ve duyarlılık (0,8312) değerleriyle doğru ayırım yapılabildiği anlaşılmaktadır.

Şüpheli KOA derecesinde ise performans diğer derecelere kıyasla daha düşük kalmıştır. Bu grupta F1-skoru 0,6994 ve duyarlılık 0,6010 olarak hesaplanmış, şüpheli durumların tespitinde güçlük yaşandığı belirlenmiştir. Hafif KOA derecesinde ise 0,7830 F1-skoru elde edilmiş olup, ayırım başarısının orta seviyede olduğu görülmektedir.

Genel değerlendirmede, kesinlik 0,8190, duyarlılık 0,8177, F1-skoru 0,8122 ve doğruluk oranı 0,8176 olarak bulunmuştur. Bu sonuçlar, Ortak Füzyon Tip I stratejisi ile VGG16 derin özelliklerinin ve geleneksel özelliklerin entegrasyonu sonucunda KOA derecelerinin dengeli ve başarılı bir şekilde ayrılabilmesini göstermektedir. Özellikle ağır, orta ve sağlıklı derecelerde yüksek doğruluk sağlanırken, şüpheli derecelerde iyileştirme gereksinimi devam etmektedir.

Şekil 4.7’de, test verisi üzerinde, VGG16 modelinden çıkarılan derin özellikler ile geleneksel özelliklerin Ortak Füzyon Tip I stratejisi kapsamında birleştirilmesi sonucunda gerçekleştirilen KOA derecelendirmesine ait konfüzyon matrisi sunulmaktadır.



Şekil 4.7. Ortak Füzyon Tip I ile VGG16 derin ve geleneksel özelliklerin birleşimi sonucu test verisinde elde edilen KOA derecelendirme konfüzyon matrisi.

Çizelge 4.10. Ortak Füzyon Tip I: Geleneksel ve EfficientNetb0 Tabanlı Derin Özelliklerle Sınıflandırma Sonuçları

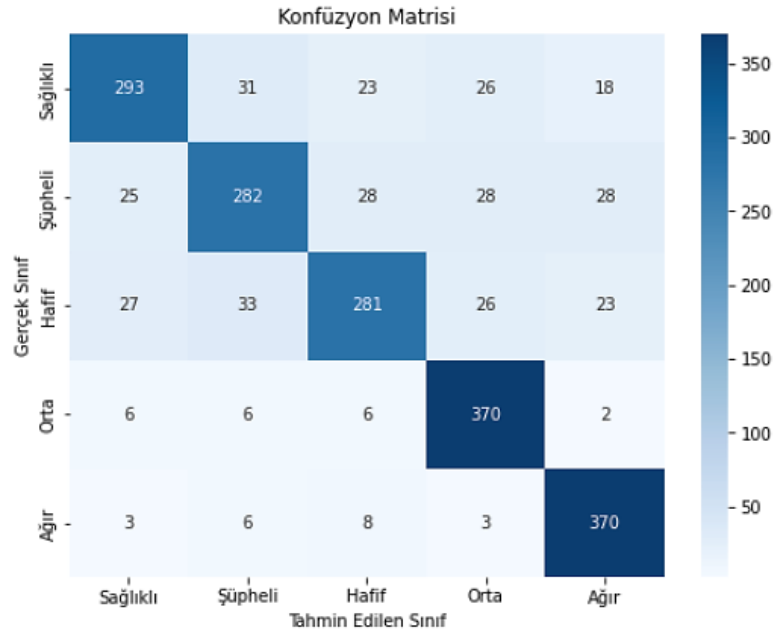
Sınıf	Kesinlik	Duyarlılık	F1-Skoru	Sınıflandırma Doğruluğu
Sağlıklı	0,8276	0,7493	0,7865	0,8176
Şüpheli	0,7877	0,7212	0,7530	
Hafif	0,8121	0,7205	0,7635	
Orta	0,8167	0,9487	0,8778	
Ağır	0,8390	0,9487	0,8904	
Ortalama	0,8166	0,8177	0,8142	

Çizelge 4.10’da sunulan sonuçlara göre, Ortak Füzyon Tip I stratejisi kapsamında EfficientNetB0 modelinin ara katmanlarından çıkarılan derin özellikler ile geleneksel özelliklerin birleştirilmesiyle gerçekleştirilen sınıflandırmada, ağır ve orta KOA derecelerinde en yüksek performans elde edilmiştir. Ağır KOA derecesinde F1-skoru 0,8904, orta derecede ise 0,8778 olarak hesaplanmış, duyarlılık değerlerinin de yüksek olması bu vakaların başarılı bir şekilde ayırt edilebildiğini göstermiştir.

Sağlıklı derece grubunda F1-skoru 0,7865 olup, modelin bu kategoride orta seviyede bir başarı sağladığı gözlemlenmiştir. Şüpheli KOA derecesinde ise modelin performansı diğer derecelere kıyasla daha düşük kalmıştır; bu grupta F1-skoru 0,7530 ve duyarlılık 0,7212 bulunmuştur. Hafif KOA derecesinde ise 0,7635 F1-skoru elde edilmiş ve orta düzeyde bir ayırım başarısı sağlanmıştır.

Genel ortalama değerlere bakıldığında, kesinlik 0,8166, duyarlılık 0,8177, F1-skoru 0,8142 ve doğruluk oranı 0,8176 olarak hesaplanmıştır. Bu sonuçlar, Ortak Füzyon Tip I yaklaşımıyla EfficientNetB0 modelinden elde edilen derin özellikler ile geleneksel özelliklerin entegrasyonunun KOA derecelerinin dengeli bir şekilde ayırımında etkili olduğunu göstermektedir. Ancak, şüpheli KOA derecesinde duyarlılık ve doğruluk değerlerinin artırılması gerektiği görülmektedir.

Şekil 4.8’de, test verisi üzerinde, EfficientNetB0 derin özellikleri ile geleneksel özelliklerin Ortak Füzyon Tip I stratejisi kapsamında birleştirilmesi sonucunda gerçekleştirilen KOA derecelendirmesine ait konfüzyon matrisi sunulmaktadır.



Şekil 4.8. Ortak Füzyon Tip I ile EfficientNetb0 derin ve geleneksel özelliklerin birleşimi sonucu test verisinde elde edilen KOA derecelendirme konfüzyon matrisi.

Çizelge 4.11. Ortak Füzyon Tip I: Geleneksel ve Xception Tabanlı Derin Özelliklerle Sınıflandırma Sonuçları

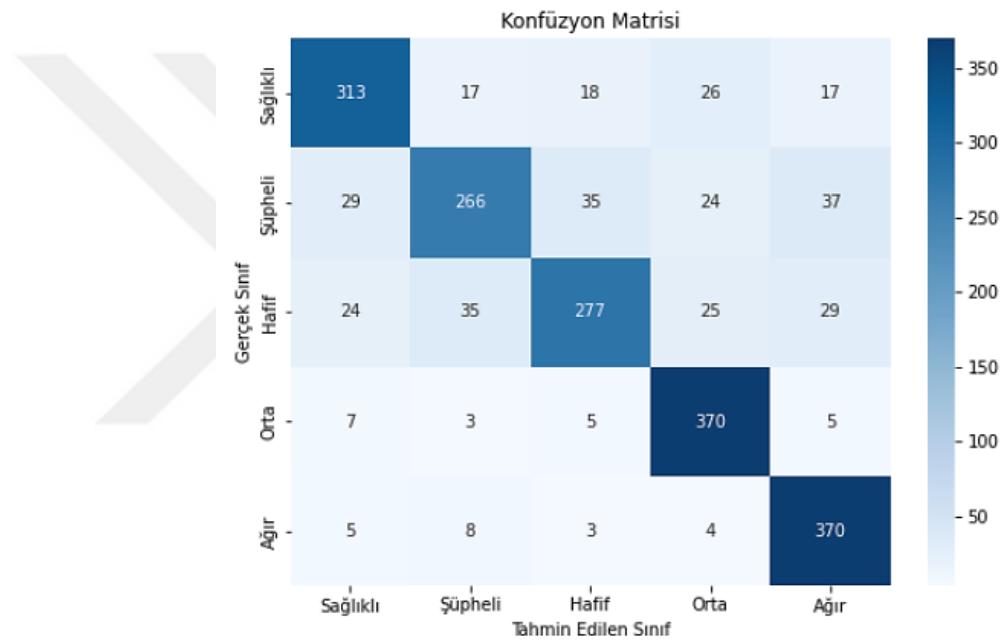
Sınıf	Kesinlik	Duyarlılık	F1-Skoru	Sınıflandırma Doğruluğu
Sağlıklı	0,8280	0,8005	0,8140	0,8176
Şüpheli	0,8085	0,6803	0,7388	
Hafif	0,8195	0,7102	0,7609	
Orta	0,8240	0,9487	0,8820	
Ağır	0,8078	0,9487	0,8726	
Ortalama	0,8175	0,8177	0,8137	

Çizelge 4.11’de sunulan sonuçlara göre, Ortak Füzyon Tip I stratejisi kapsamında Xception modelinin ara katmanlarından çıkarılan derin özellikler ile geleneksel özelliklerin birleştirilmesiyle gerçekleştirilen sınıflandırmada, ağır ve orta KOA derecelerinde yüksek başarı elde edilmiştir. Ağır derecede F1-skoru 0,8726, orta derecede ise 0,8820 olarak hesaplanmış olup, her iki derecede de kesinlik ve duyarlılık değerlerinin yüksek olduğu görülmüştür. Bu sonuçlar, modelin ileri KOA derecelerini doğru şekilde tespit etmede güçlü bir performans sergilediğini göstermektedir.

Sağlıklı derecede F1-skoru 0,8140 olup, modelin bu grubu ayırt etmede iyi bir başarı sağladığı söylenebilir. Buna karşılık, hafif KOA derecesinde (F1-skoru: 0,7609) ve özellikle şüpheli KOA derecesinde (F1-skoru: 0,7388, duyarlılık: 0,6803) model performansında düşüş gözlemlenmiştir. Şüpheli derecenin düşük duyarlılık değeri, modelin bu grubu doğru şekilde sınıflandırmada güçlük yaşadığını göstermektedir.

Genel ortalama değerlere bakıldığında, kesinlik 0,8175, duyarlılık 0,8177, F1-skoru 0,8137 ve sınıflandırma doğruluğu 0,8176 olarak elde edilmiştir. Bu sonuçlar, modelin genel olarak dengeli ve başarılı bir performans sunduğunu ortaya koymaktadır. Ancak, erken evre KOA derecelerinin tespitinde iyileştirme yapılması gerektiği söylenebilir. Model, ağır ve orta KOA derecelerinde yüksek doğruluk sağlarken, şüpheli ve hafif derecelerde daha düşük başarı sergilemiştir.

Şekil 4.9'da, test verisi üzerinde Xception derin özellikleri ve geleneksel özelliklerin Ortak Füzyon Tip I stratejisiyle entegrasyonu sonucunda gerçekleştirilen KOA derecelendirmesine ait konfüzyon matrisi sunulmaktadır.



Şekil 4.9. Ortak Füzyon Tip I ile Xception derin ve geleneksel özelliklerin birleşimi sonucu test verisinde elde edilen KOA derecelendirme konfüzyon matrisi.

Çizelge 4.12. Ortak Füzyon Tip I: Geleneksel ve DenseNet201 Tabanlı Derin Özelliklerle Sınıflandırma Sonuçları

Sınıf	Kesinlik	Duyarlılık	F1-Skoru	Sınıflandırma Doğruluğu
Sağlıklı	0,8320	0,8491	0,8405	0,8350
Şüpheli	0,8312	0,6803	0,7482	
Hafif	0,8414	0,7487	0,7924	
Orta	0,8390	0,9487	0,8904	
Ağır	0,8314	0,9487	0,8862	
Ortalama	0,8350	0,8351	0,8315	

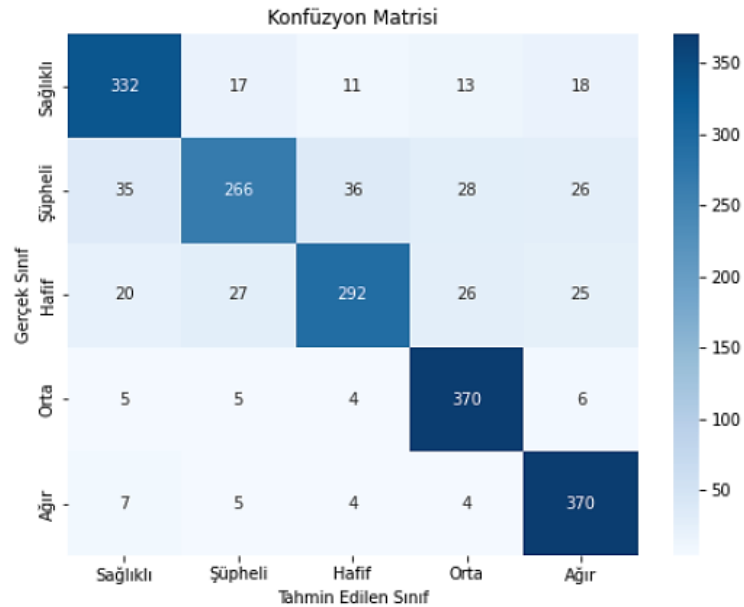
Çizelge 4.12'de sunulan sonuçlara göre, Ortak Füzyon Tip I stratejisi kapsamında DenseNet201 modelinin ara katmanlarından çıkarılan derin özellikler ile

geleneksel özelliklerin birleştirilmesiyle gerçekleştirilen sınıflandırmada, ağır ve orta KOA derecelerinde yüksek başarı elde edilmiştir. Ağır derecede F1-skoru 0,8862, orta derecede ise 0,8904 olarak hesaplanmış olup, her iki derecede de kesinlik ve duyarlılık değerlerinin yüksek olduğu görülmüştür. Bu bulgular, modelin ileri KOA derecelerini doğru şekilde tespit etmede güçlü bir performans sergilediğini göstermektedir.

Sağlıklı KOA derecesinde de model, F1-skoru 0,8405 ile iyi bir sınıflandırma performansı sunmuştur. Ancak, hafif ve şüpheli KOA derecelerinde daha düşük sonuçlar elde edilmiştir. Hafif derecede F1-skoru 0,7924, şüpheli derecede ise 0,7482 olarak hesaplanmıştır. Şüpheli derecedeki düşük F1-skoru, modelin bu grubu ayırt etmede zorlandığını ve bu alanda geliştirme ihtiyacı bulunduğunu göstermektedir.

Genel ortalama değerlere bakıldığında, kesinlik 0,8350, duyarlılık 0,8351, F1-skoru 0,8315 ve sınıflandırma doğruluğu 0,8350 olarak elde edilmiştir. Bu sonuçlar, modelin genel olarak başarılı ve dengeli bir performans sunduğunu ortaya koymaktadır. DenseNet201 modeli özellikle ağır ve orta KOA derecelerinde güçlü bir ayırım sağlarken, erken evre KOA derecelerinin (hafif ve şüpheli) sınıflandırılmasında iyileştirmelere ihtiyaç duyulmaktadır.

Şekil 4.10'da, test verisi üzerinde DenseNet201 derin özellikleri ve geleneksel özelliklerin Ortak Füzyon Tip I stratejisiyle entegrasyonu sonucunda gerçekleştirilen KOA derecelendirmesine ait konfüzyon matrisi sunulmaktadır.



Şekil 4.10. Ortak Füzyon Tip I ile DenseNet201 derin ve geleneksel özelliklerin birleşimi sonucu test verisinde elde edilen KOA derecelendirme konfüzyon matrisi.

Ortak Füzyon Tip I stratejisi kullanılarak gerçekleştirilen KOA derecelendirmesinde, ResNet-50, VGG16, EfficientNetB0, Xception ve DenseNet201 transfer öğrenme modellerinden elde edilen derin özellikler ile geleneksel özelliklerin entegrasyonu değerlendirilmiştir. Elde edilen sonuçlar incelendiğinde, tüm modellerin özellikle ağır ve orta KOA derecelerinde yüksek kesinlik, duyarlılık ve F1-skoru değerleri ile başarılı bir performans sergilediği görülmüştür. Bu derecelerde modellerin doğruluk oranları da oldukça yüksek seviyelere ulaşmıştır.

Sağlıklı derecede modeller genel olarak orta düzeyde performans göstermiş, her modelde farklı seviyelerde başarılar elde edilmiştir. Hafif KOA derecesinde ise tüm modellerde belirgin bir performans düşüşü gözlemlenmiştir. Şüpheli KOA derecesi, tüm modellerde en zayıf sonuçların elde edildiği grup olmuştur. Özellikle şüpheli derecede düşük duyarlılık değerleri, modellerin bu grubu doğru tespit etmede zorlandığını göstermektedir.

Genel performans değerlendirmesi yapıldığında, DenseNet201 modeli, şüpheli derecedeki nispeten düşük başarısına rağmen, ortalama sınıflandırma doğruluğu ve F1-skoru açısından en iyi sonuçları sunmuştur. Ağır ve orta KOA derecelerinde gösterdiği yüksek başarı, DenseNet201 modelini klinik uygulamalar için güçlü bir aday haline getirmektedir. Bununla birlikte, şüpheli derece için model performansının artırılmasına yönelik geliştirme çalışmalarına ihtiyaç duyulduğu sonucuna varılabilir.

4.3 Ortak Füzyon Tip II Sonuçları

Tez çalışmasında son olarak Ortak Füzyon Tip II stratejisi uygulanmıştır. Bu strateji kapsamında, derin özellik çıkarımı amacıyla transfer öğrenme modellerinden yararlanılmıştır. Önceden eğitilmiş modellerin belirli katmanlarından elde edilen derin özellikler, manuel olarak çıkarılan geleneksel görüntü özellikleri ile birleştirilerek KOA derecelendirmesi için sınıflandırma sürecine girdi olarak sunulmuştur.

ResNet-50, VGG16, EfficientNetB0, Xception ve DenseNet201 modellerinden çıkarılan derin özellikler ile manuel olarak elde edilen geleneksel özelliklerin birleştirilmesiyle oluşturulan Ortak Füzyon Tip II uygulamasına ait sınıflandırma sonuçları sırasıyla Çizelge 4.13, 4.14, 4.15, 4.16 ve 4.17’de sunulmuştur.

Çizelge 4.13. Ortak Füzyon Tip II: Geleneksel ve ResNet-50 Tabanlı Derin Özelliklerle Sınıflandırma Sonuçları

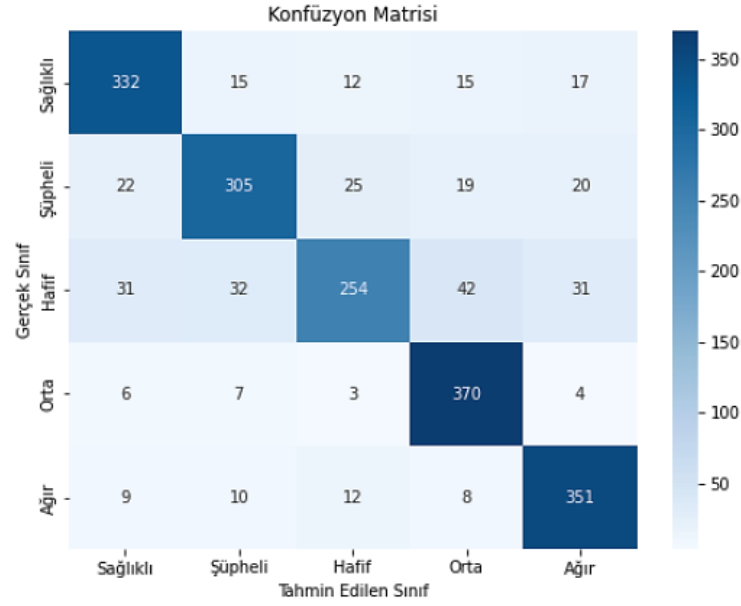
Sınıf	Kesinlik	Duyarlılık	F1-Skoru	Sınıflandırma Doğruluğu
Sağlıklı	0,8300	0,8491	0,8394	0,8258
Şüpheli	0,8265	0,7800	0,8026	
Hafif	0,8300	0,6512	0,7298	
Orta	0,8149	0,9487	0,8767	
Ağır	0,8297	0,9000	0,8634	
Ortalama	0,8262	0,8440	0,8224	

Çizelge 4.13'te sunulan sonuçlara göre, Ortak Füzyon Tip II stratejisinde ResNet-50 modelinden çıkarılan derin özellikler ile geleneksel görüntü özelliklerinin entegrasyonu sonucunda gerçekleştirilen KOA derecelendirmesinde dengeli ve yüksek doğruluk değerleri elde edilmiştir. Sağlıklı derece grubunda kesinlik (0,8300) ve duyarlılık (0,8491) değerleri yüksek bulunmuş, F1-skoru ise 0,8394 olarak hesaplanmıştır. Ağır KOA derecesinde de yüksek doğruluk sağlanmış, kesinlik (0,8297) ve duyarlılık (0,9000) değerleriyle başarılı bir ayırım yapılmıştır.

Şüpheli derece grubunda önceki yaklaşımlara göre bir iyileşme gözlenmiş; kesinlik 0,8265 ve duyarlılık 0,7800 seviyelerine ulaşılmıştır. Buna karşın, hafif KOA derecesinde doğruluk oranlarının bir miktar düştüğü ve F1-skorunun 0,7298 seviyesinde kaldığı görülmüştür. Orta KOA derecesinde ise 0,8767 F1-skoru ile yüksek bir ayırım başarısı sağlanmıştır.

Genel performans incelendiğinde, ortalama doğruluk oranı 0,8258 olarak elde edilmiş ve KOA derecelerinin ayırımında başarılı bir sonuç ortaya konmuştur. Ancak hafif KOA derecesinde ayırt edici gücün artırılması gerektiği dikkat çekmektedir.

Şekil 4.11, test verileri üzerinde gerçekleştirilen bu KOA derecelendirme sürecine ilişkin konfüzyon matrisini sunmaktadır.



Şekil 4.11. Test verisi üzerinde, Ortak Füzyon Tip II stratejisiyle ResNet-50 derin ve geleneksel özelliklerinin birleşimiyle yapılan KOA derecelendirme sonuçlarına ait konfüzyon matrisi

Çizelge 4.14. Ortak Füzyon Tip II: Geleneksel ve VGG16 Tabanlı Derin Özelliklerle Sınıflandırma Sonuçları

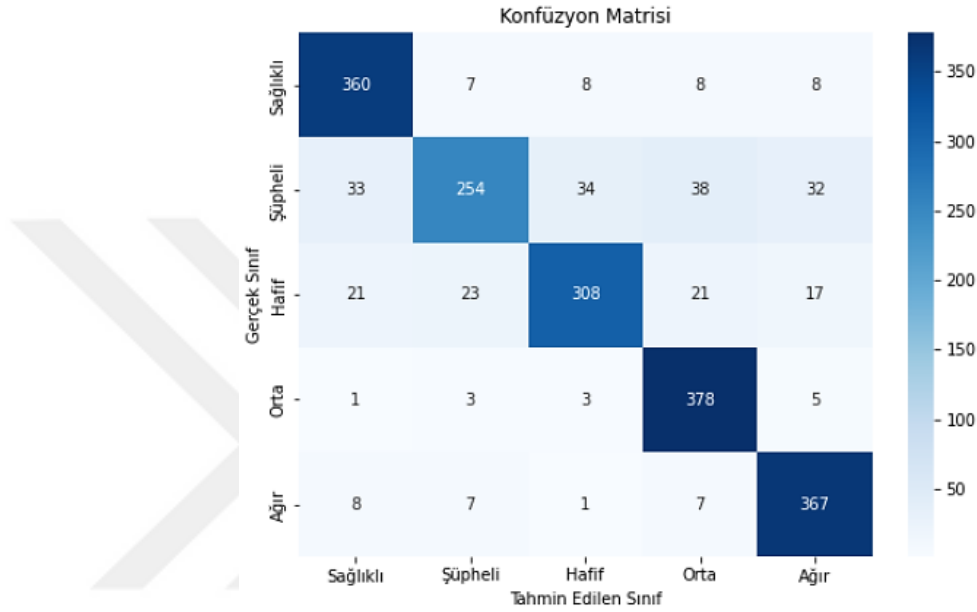
Sınıf	Kesinlik	Duyarlılık	F1-Skoru	Sınıflandırma Doğruluğu
Sağlıklı	0,8510	0,9207	0,8845	0,8539
Şüpheli	0,8639	0,6496	0,7416	
Hafif	0,8700	0,7897	0,8279	
Orta	0,8362	0,9692	0,8978	
Ağır	0,8554	0,9410	0,8962	
Ortalama	0,8553	0,8540	0,8496	

Çizelge 4.14'te sunulan sonuçlara göre, Ortak Füzyon Tip II stratejisinde VGG16 ile elde edilen derin ve geleneksel özelliklerin birleşimiyle gerçekleştirilen KOA derecelendirmesi genel olarak güçlü bir performans göstermiştir. Sağlıklı derecelerde kesinlik (0,8510) ve duyarlılık (0,9207) değerleri ile yüksek doğruluk sağlanmıştır. Ağır derecede de kesinlik (0,8554) ve duyarlılık (0,9410) değerleri yüksek olup, F1-skoru (0,8962) bu grubun başarılı şekilde ayırt edilebildiğini göstermektedir. Orta derecelendirme grubunda da F1-skoru (0,8978) ile oldukça başarılı sonuçlar elde edilmiştir.

Buna karşın, şüpheli derecelendirmede kesinlik (0,8639) ve duyarlılık (0,6496) değerlerinin düşük olması, ayırt etme başarısının diğer derecelere kıyasla zayıf olduğunu göstermektedir. Hafif KOA derecesinde ise kesinlik (0,8700) ve duyarlılık

(0,7897) ile orta düzeyde bir başarı sağlanmış ve F1-skoru 0,8279 olarak hesaplanmıştır.

Genel ortalamalar incelendiğinde, doğruluk (0,8539) ve F1-skoru (0,8496) değerleriyle güçlü ve dengeli bir derecelendirme performansı elde edilmiştir. Ancak şüpheli derecelerde ayırım başarısının artırılması gerektiği görülmektedir. Şekil 4.12, test verisi üzerinde gerçekleştirilen bu KOA derecelendirme sonuçlarına ait konfüzyon matrisini göstermektedir.



Şekil 4.12. Test verisi üzerinde, Ortak Füzyon Tip II stratejisi ile VGG16 ve geleneksel özelliklerin birleşimi sonucu elde edilen KOA derecelendirme konfüzyon matrisi

Çizelge 4.15. Ortak Füzyon Tip II: Geleneksel ve EfficientNetb0 Tabanlı Derin Özelliklerle Sınıflandırma Sonuçları

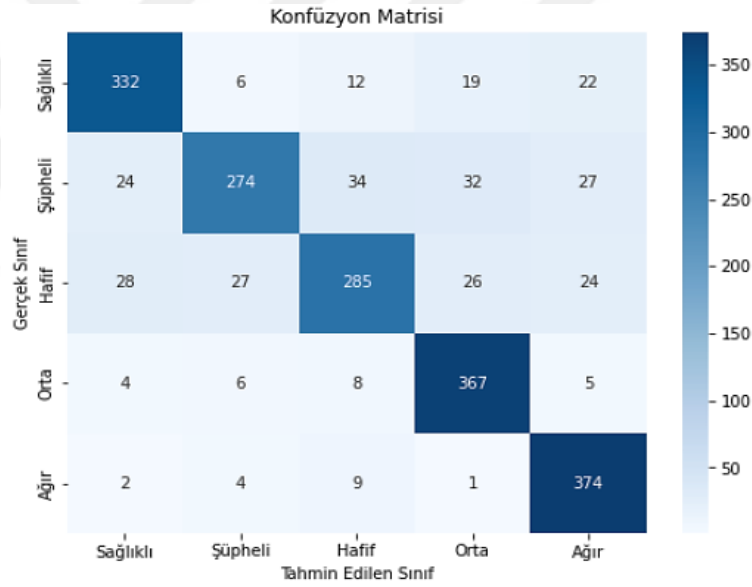
Sınıf	Kesinlik	Duyarlılık	F1-Skoru	Sınıflandırma Doğruluğu
Sağlıklı	0,8512	0,8491	0,8501	0,8360
Şüpheli	0,8643	0,7007	0,7740	
Hafif	0,8189	0,7307	0,7723	
Orta	0,8247	0,9410	0,8790	
Ağır	0,8274	0,9589	0,8883	
Ortalama	0,8373	0,8361	0,8327	

Çizelge 4.15'te sunulan sonuçlara göre, Ortak Füzyon Tip II stratejisinde EfficientNetB0 ile elde edilen derin ve geleneksel özelliklerin birleşimi sonucunda gerçekleştirilen KOA derecelendirmesi genel olarak dengeli ve yüksek doğruluklu bir performans sergilemiştir. Ağır derecelendirme grubunda kesinlik (0,8274) ve duyarlılık (0,9589) değerleri ile F1-skoru (0,8883) oldukça yüksek bulunmuştur. Orta

derecelendirmede ise duyarlılık (0,9410) ve F1-skoru (0,8790) değerleriyle güçlü bir başarı elde edilmiştir.

Sağlıklı derece grubunda kesinlik (0,8512) ve F1-skoru (0,8501) ile tatmin edici bir performans sergilenmiş, hafif derecede kesinlik (0,8189) ve duyarlılık (0,7307) ile dengeli sonuçlar elde edilmiştir. Buna karşılık, şüpheli derecelendirmede kesinlik (0,8643) ve duyarlılık (0,7007) değerleri diğer derecelere göre daha düşük bulunmuş, bu da bu derecenin belirlenmesinde iyileştirme gereksinimini ortaya koymuştur.

Genel olarak, ortalama doğruluk değeri 0,8360 olarak hesaplanmış ve EfficientNetB0 temelli yaklaşımın, özellikle ağır ve orta dereceler için oldukça başarılı sonuçlar ürettiği görülmüştür. Şüpheli derecenin belirlenmesinde ise performansın artırılması gerektiği değerlendirilmektedir. Şekil 4.13, test verisi üzerinde gerçekleştirilen bu KOA derecelendirme sonuçlarına ait konfüzyon matrisini sunmaktadır.



Şekil 4.13. Test verisi üzerinde, Ortak Füzyon Tip II stratejisi ile EfficientNetB0 ve geleneksel özelliklerin birleşimi sonucu elde edilen KOA derecelendirme konfüzyon matrisi

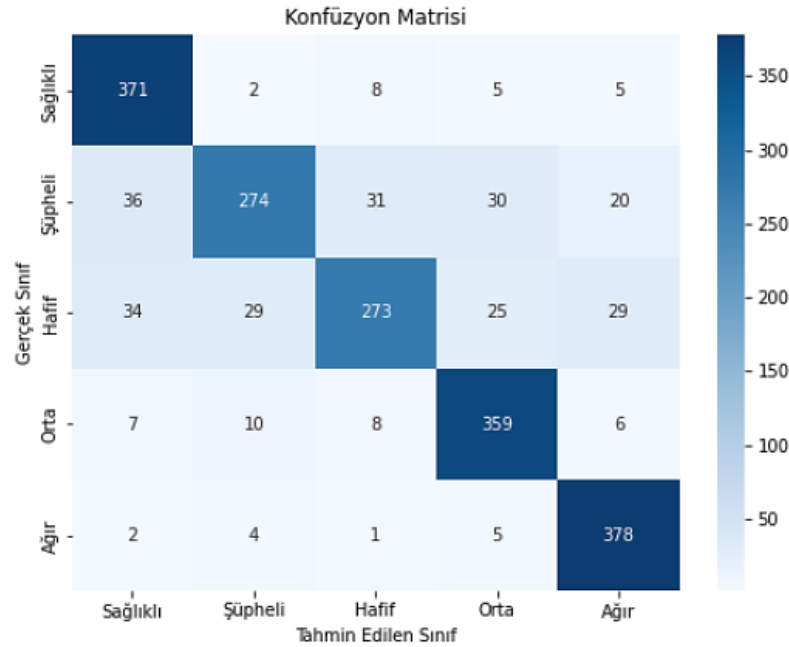
Çizelge 4.16. Ortak Füzyon Tip II: Geleneksel ve Xception Tabanlı Derin Özelliklerle Sınıflandırma Sonuçları

Sınıf	Kesinlik	Duyarlılık	F1-Skoru	Sınıflandırma Doğruluğu
Sağlıklı	0,8244	0,9488	0,8822	0,8478
Şüpheli	0,8589	0,7007	0,7718	
Hafif	0,8504	0,7000	0,7679	
Orta	0,8466	0,9205	0,8820	
Ağır	0,8630	0,9692	0,9130	
Ortalama	0,8448	0,8478	0,8434	

Çizelge 4.16'da sunulan sonuçlara göre, Ortak Füzyon Tip II stratejisinde Xception ile elde edilen derin ve geleneksel özelliklerin birleşimiyle gerçekleştirilen KOA derecelendirmesi genel olarak başarılı sonuçlar vermiştir. Ağır derecelendirme grubunda kesinlik (0,8630) ve duyarlılık (0,9692) değerleriyle F1-skoru (0,9130) oldukça yüksek bulunmuştur. Orta derecelendirmede de kesinlik (0,8466) ve F1-skoru (0,8820) ile güçlü bir performans elde edilmiştir.

Sağlıklı derece grubunda kesinlik (0,8244) ve F1-skoru (0,8822) değerleri ile iyi bir başarı sağlanmıştır. Ancak, şüpheli derecelendirme grubunda kesinlik (0,8589) ve duyarlılık (0,7007) değerlerinin diğer derecelere kıyasla düşük kaldığı gözlemlenmiştir. Bu durum, şüpheli vakaların doğru şekilde ayırt edilmesinde iyileştirme ihtiyacını göstermektedir.

Genel değerlendirmede, ortalama doğruluk 0,8478 ve ortalama F1-skoru 0,8434 olarak hesaplanmıştır. Xception modeli özellikle ağır, orta ve sağlıklı dereceler için güçlü bir performans sergilerken, şüpheli derecede iyileştirmelere ihtiyaç duyduğu anlaşılmaktadır. Şekil 4.14, test verisi üzerinde gerçekleştirilen bu KOA derecelendirme sonuçlarına ait konfüzyon matrisini sunmaktadır.



Şekil 4.14. Test verisi üzerinde, Ortak Füzyon Tip II stratejisi ile Xception ve geleneksel özelliklerin birleşimi sonucu elde edilen KOA derecelendirme konfüzyon matrisi

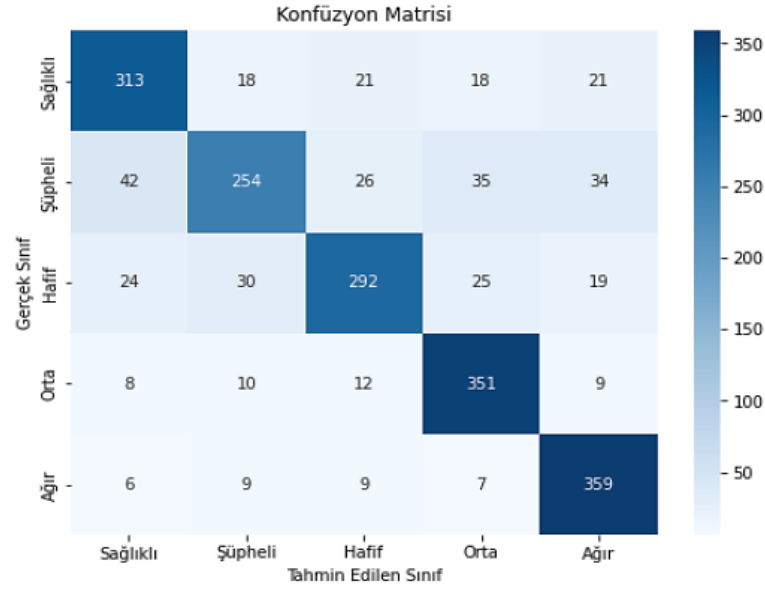
Çizelge 4.17. Ortak Füzyon Tip II: Geleneksel ve DenseNet201 Tabanlı Derin Özelliklerle Sınıflandırma Sonuçları

Sınıf	Kesinlik	Duyarlılık	F1-Skoru	Sınıflandırma Doğruluğu
Sağlıklı	0,7964	0,8005	0,7987	0,8037
Şüpheli	0,7912	0,6496	0,7134	
Hafif	0,8111	0,7487	0,7786	
Orta	0,8050	0,9000	0,8498	
Ağır	0,8122	0,9205	0,8629	
Ortalama	0,8032	0,8038	0,8006	

Çizelge 4.17 incelendiğinde, Ortak Füzyon Tip II stratejisinde DenseNet201 ile elde edilen derin ve geleneksel özelliklerin birleşimi sonucunda gerçekleştirilen KOA derecelendirmesinde, özellikle ağır ve orta derecelerinde başarılı sonuçlar elde edilmiştir. Ağır derecede kesinlik (0,8122) ve duyarlılık (0,9205) değerleriyle F1-skoru (0,8629) yüksek bulunmuştur. Orta derecelendirme grubunda ise kesinlik (0,8050) ve F1-skoru (0,8498) ile etkili bir performans sergilenmiştir.

Buna karşın, sağlıklı ve şüpheli derecelendirme gruplarında modelin daha düşük performans gösterdiği görülmektedir. Sağlıklı derecede kesinlik (0,7964) ve duyarlılık (0,8005) değerleri ortalama seviyede kalırken, şüpheli derecede kesinlik (0,7912) ve duyarlılık (0,6496) ile performansın daha da düştüğü gözlemlenmiştir. Bu durum, özellikle şüpheli vakaların doğru sınıflandırılmasında iyileştirmeye ihtiyaç olduğunu göstermektedir.

Genel değerlendirildiğinde, ortalama doğruluk 0,8037 olarak hesaplanmış olup, DenseNet201 ile elde edilen sonuçların, diğer modellere kıyasla bir miktar daha düşük bir genel performans sunduğu anlaşılmaktadır. Ancak, ağır ve orta derecelerinde gösterilen yüksek doğruluk ve duyarlılık, modelin bu alanlarda etkili bir ayırım yapabildiğini ortaya koymaktadır. Şekil 4.15, test verisi üzerinde gerçekleştirilen bu KOA derecelendirme sonuçlarına ait konfüzyon matrisini sunmaktadır.



Şekil 4.15. Test verisi üzerinde, Ortak Füzyon Tip II stratejisi ile DenseNet201 ve geleneksel özelliklerin birleşimi sonucu elde edilen KOA derecelendirme konfüzyon matrisi

Ortak Füzyon Tip II stratejisi kapsamında yapılan genel değerlendirme incelendiğinde, her bir transfer öğrenme modelinden elde edilen derin özelliklerin belirli güçlü yönleri ve iyileştirilmesi gereken alanları olduğu gözlemlenmiştir. ResNet-50 ile gerçekleştirilen derecelendirme sürecinde, özellikle ağır KOA derecesinde yüksek başarı sağlanmış ve genel doğruluk oranı 0,8258 olarak elde edilmiştir. VGG16 ile yapılan değerlendirmede ise ağır ve orta derecelerde çok yüksek doğruluk değerleri elde edilmiş, genel doğruluk oranı 0,8539 ile en iyi performanslardan biri olarak kaydedilmiştir; ancak şüpheli derece grubunda başarı oranının diğerlerine göre düşük olduğu dikkat çekmiştir. EfficientNetB0 ile ağır derecede yüksek performans sağlanmasına rağmen, şüpheli derecede daha düşük kesinlik ve duyarlılık değerleri gözlemlenmiş ve genel doğruluk oranı 0,8360 olarak belirlenmiştir. Xception ile de ağır derecede etkili sonuçlar alınmış, şüpheli derecede ise performansın zayıf olduğu görülmüştür; genel doğruluk oranı 0,8478'dir. DenseNet201 ile yapılan değerlendirmede ağır ve orta derecelerinde başarılı sonuçlar elde edilmiş ancak sağlıklı ve şüpheli derecelerinde daha düşük doğruluk ve duyarlılık oranları kaydedilmiştir; bu çalışmada genel doğruluk oranı 0,8037 olarak bulunmuştur.

VGG16 ile gerçekleştirilen Ortak Füzyon Tip II uygulamasında, elde edilen sonuçlar diğer modellerle karşılaştırıldığında en yüksek performansı göstermiştir. Özellikle orta KOA derecesinde 0,9692 duyarlılık ve 0,8978 F1-skoru ile üstün bir başarı sağlanmıştır. Ağır derecede ise 0,8962 F1-skoru ile öne çıkmıştır. Sağlıklı

derecelendirme grubunda da yüksek kesinlik ve duyarlılık değerleri elde edilerek güçlü bir ayrım başarısı gösterilmiştir. Ortalama doğruluk değeri 0,8539 olan VGG16 tabanlı derecelendirme yaklaşımı, genel doğruluk açısından en iyi sonucu sunmuştur. Bununla birlikte, şüpheli derecelendirme grubunda duyarlılık oranlarının diğer sınıflara göre düşük olması, bu grupta ek veri artırımı ya da model iyileştirme stratejilerinin uygulanması gerekliliğini ortaya koymaktadır. Genel sonuçlar değerlendirildiğinde, VGG16 ile elde edilen yaklaşımın KOA derecelerinin belirlenmesinde en dengeli ve güvenilir performansı sunduğu söylenebilir.

Çalışmada kullanılan üç farklı füzyon stratejisi (Erken Füzyon Tip II, Ortak Füzyon Tip I ve Ortak Füzyon Tip II) karşılaştırıldığında, her bir yaklaşımın belirli avantajlar sunduğu, ancak genel performans açısından farklılıklar gösterdiği belirlenmiştir.

Çizelge 4.18’de her bir füzyon stratejisinin elde ettiği en iyi performanslar ve karşılık gelen metrikler karşılaştırmalı olarak sunulmaktadır.

Çizelge 4.18. Füzyon stratejilerinin karşılaştırılması

Füzyon Tipi	Özellik Kombinasyonu	Doğruluk	Kesinlik	Duyarlılık	F1-Skoru
Erken Füzyon Tip II	Geleneksel + Görüntü Tabanlı	0,7356	0,7257	0,7357	0,7146
	Özellikler (PCA)				
	DenseNet201 başlangıç ağırlıklı model				
Ortak Füzyon Tip I	Geleneksel (ANOVA + LDA) + DenseNet201’den Elde Edilen Derin Özellikler	0,8350	0,8350	0,8351	0,8315
Ortak Füzyon Tip II	Geleneksel (ANOVA + LDA) + VGG16’dan Elde Edilen Derin Özellikler	0,8539	0,8553	0,8540	0,8496

Erken Füzyon Tip II stratejisinde, manuel olarak çıkarılan ve LDA ile ANOVA yöntemleriyle optimize edilen geleneksel özellikler ile PCA kullanılarak boyutu küçültülmüş görüntü temsilleri birleştirilmiştir. Bu strateji, farklı bilgi kaynaklarının bütünleştirilmesini amaçlamıştır. Ancak, elde edilen sonuçlar incelendiğinde, özellikle şüpheli KOA derecesinde duyarlılık değerlerinin düşük olduğu, bu nedenle erken evre KOA vakalarının sınıflandırılmasında zorluk yaşandığı görülmektedir. DenseNet201 başlangıç ağırlıkları ile eğitilen model ile en iyi sonuçlar elde edilmesine rağmen, bu

stratejinin genel doğruluk ve F1-skoru açısından diğer füzyon stratejilerine kıyasla daha düşük performans sergilediği belirlenmiştir.

Ortak Füzyon Tip I stratejisinde, ANOVA ve LDA ile optimize edilmiş geleneksel özellikler ve transfer öğrenme modellerinin ara katmanlarından çıkarılan derin özellikler ayrı ayrı işlenmiş ve daha sonra bu iki kaynaktan elde edilen özet temsiller birleştirilerek sınıflandırmaya sunulmuştur. Bu yapı, özellikle ağır ve orta KOA derecelerinin doğru sınıflandırılmasında yüksek doğruluk ve F1-skorları sağlamıştır. Ancak, hafif ve şüpheli derecelerde duyarlılık değerlerinin görece düşük kalması, erken evre KOA vakalarının ayırt edilmesinde iyileştirme yapılması gerektiğini göstermektedir. DenseNet201 modelinden elde edilen derin özelliklerin kullanıldığı bu yaklaşım, doğruluk (%83,50) ve F1-skoru (%83,15) değerleri ile Erken Füzyon Tip II'ye kıyasla anlamlı bir gelişim sağlamıştır.

Ortak Füzyon Tip II stratejisi ise çalışmada en başarılı yöntem olarak öne çıkmıştır. Bu stratejide, transfer öğrenme modellerinin belirlenen katmanlarından çıkarılan derin özellikler ile ANOVA ve LDA ile optimize edilmiş geleneksel özellikler doğrudan birleştirilmiş ve ortak bir sinir ağı yapısında sınıflandırma gerçekleştirilmiştir. Özellikle VGG16 modelinden elde edilen derin özellikler ile yapılan birleşim sonucunda, %85,39 doğruluk ve %84,96 F1-skoru elde edilerek en yüksek performans sağlanmıştır. Ağır ve orta KOA derecelerinde yüksek kesinlik ve duyarlılık değerleri elde edilmesi, modelin ileri evre vakaları başarıyla ayırt edebildiğini göstermektedir. Ayrıca, diğer füzyon stratejilerine kıyasla şüpheli ve hafif dereceler için de daha dengeli sonuçlar elde edilmiştir.

Genel değerlendirme sonucunda, Ortak Füzyon Tip II stratejisinin KOA derecelendirmesi için en etkili yaklaşım olduğu belirlenmiştir. Özellikle VGG16 modelinden çıkarılan derin temsillerin, ANOVA ve LDA ile optimize edilmiş geleneksel görüntü özellikleri ile doğrudan birleştirilmesi sonucu oluşturulan birleşik özellik vektörleriyle yapılan sınıflandırma sürecinde, en yüksek doğruluk ve F1-skoru değerlerine ulaşılmıştır. Ağır ve orta KOA derecelerinde elde edilen güçlü başarı, bu yöntemin ileri evre vakaların doğru şekilde tespit edilmesinde etkili olduğunu göstermektedir. Ayrıca, yapılan ANOVA analizi sonuçlarına göre özellikle GLCM tabanlı doku özellikleri, LBP desenleri ile çevre ve dairesellik gibi morfolojik ölçütlerin belirleyici olduğu görülmüştür. Bu geleneksel ölçütlerin, VGG16'dan elde edilen derin temsillerle birlikte entegre edilmesi, bilgi çeşitliliğini artırarak sınıflandırma başarısını önemli ölçüde yükseltmiştir.

Gerçekleştirilen çalışma sonuçları ayrıca literatürdeki mevcut çalışmalar ile karşılaştırılmıştır. Çizelge 4.19, beş sınıflı KOA derecelendirmesi amacıyla yapılan literatür çalışmaları ile bu tez çalışmasının sonuçlarını karşılaştırmalı olarak sunmaktadır.

Çizelge 4.19. Literatür çalışmaları ile karşılaştırma

Yazarlar (Yıl)	Yöntemler	Veri Seti	Sonuçlar
Tiulpin ve diğ. (2018)	Siamese CNN, Kellgren–Lawrence derecelendirme ölçeği	OAI veri seti (5960 diz görüntüsü)	%66,71 doğruluk, 0,83 Kappa skoru
Chen ve diğ. (2019)	DenseNet, VGG, ResNet ve InceptionV3	OAI	VGG16 doğruluk:%69,7
Tiulpin ve diğ. (2019)	CNN + GBM ile OA ilerleme tahmini	Ham radyografik veri ve klinik bilgiler	%62 hassasiyet, %82 AUC
Thomas ve diğ. (2020)	CNN, veri artırma, radyografi analizi	Osteoartrit İnisiyatifi (OAI) veri seti, 40,000+ görüntü	Ortalama doğruluk %71, F1 skoru %70
Zhang ve diğ. (2020)	ResNet-18 ile lokalizasyon, ResNet-34 + CBAM ile KL derecelendirme sınıflandırması	OAI	%74,81 doğruluk, 0,88 Quadratic Kappa, 0,36 hata
Tolpadi ve diğ. (2020)	Derin Siyam CNN, geliştirilmiş ResNet-34, CADx	OAI	OA lezyonları tespiti
Yong ve diğ. (2021)	Ordinal regresyon modülü (ORM), DenseNet-161	OAI, 8260 dizi radyografi	Doğruluk: %88,09, Quadratic Weighted Kappa: 0,8609
Kwon ve diğ. (2020)	SVM, Inception-Net-V2	364 hasta, 728 diz görüntüsü, Seul hastanesi	F1 skoru: 0,71, duyarlılık: 0,70, hassasiyet: 0,76
Gornale ve diğ. (2020)	Hu'nun invariant momentleri, K-NN, karar ağacı	2000 görüntü, KL derecelendirme	Genel doğruluk: %99,23
Leung ve diğ. (2020)	Resnet34	OAI 4796 görüntü	Doğruluk:%72,7
Jakaite ve diğ.	ANN, SVM, RF, Zernike	31 görüntü,	Doğruluk: %85,0

(2021)	tabanlı doku özellikleri	KL derecelendirme	
Chen ve diğ. (2019)	YOLOv2, DenseNet, VGG, ResNet, InceptionV3	OAI, diz X-ışını görüntüleri	Jaccard endeksi: 0,858 doğruluk: %92,2
Yong ve diğ. (2022)	Ordinal regresyon modülü, kümülatif bağlantı fonksiyonu, 23 farklı sinir ağı karşılaştırması	OAI veri seti (4796 katılımcı)	%88,09 doğruluk, 0,8609 QWK 0,330 Hata
Kondal ve diğ. (2022)	Nesne tespiti, CNN, regresyon modeli	Diz radyografileri (KL ölçeği)	Ortalama mutlak hata düşüşü: 1,09 → 0,28
Yunus ve diğ. (2022)	2D-3D dönüşümü, LBP özellik çıkarma, YOLOv2, Dark-net-53, Alex-Net	OAI, 3:1 eğitim/test verisi	Doğruluk: %90,6, mAP: 0,98
Cueva ve diğ. (2022)	Siamese CNN, transfer öğrenme (ResNet-34), veri artırma	Kamuya açık + özel veri setleri	%61 doğruluk
Gupta ve diğ. (2023)	DenseNet-121 + CBAM, veri artırma, 0 ve 1 sınıflarının birleştirilmesi	OAI veri seti (4796 hasta, 45-79 yaş)	%86,714 doğruluk
Pi ve diğ. (2023)	Topluluk tabanlı model: DenseNet-161, EfficientNet, ResNet, vb., TTA, Hard/Soft/Mix voting	CCI Interpolation yöntemiyle işlenmiş veri seti	Temel model: %74,21, Optimize model: %76,33 doğruluk
Solak (2024)	ResNet-50, VGG16, EfficientNetB0, Xception ve DenseNet201	OAI	Xception-rmsprop %67,8 doğruluk, %68,8 F1-Skoru 0,55 Cohenin kappa değeri
Mahfus ve diğ. (2025)	AlexNet, ResNet101, DenseNet201 ve VGG19	9786 X-ray images	ResNet101 %97,85 doğruluk
Bu Çalışma (2024)	Erken Füzyon Tip II (Geleneksel ve PCA Tabanlı Özellik Füzyonu)	OAI, 8260 diz X-ışını görüntüleri	0,7356 Doğruluk 0,7257 Kesinlik 0,7357 Duyarlılık 0,7146 F1-Skoru
	Ortak Füzyon Tip I (Geleneksel ve DenseNet201 Kaynaklı Derin Özellik Füzyonu)		0,8350 Doğruluk 0,8350 Kesinlik 0,8351 Duyarlılık 0,8315 F1-Skoru

	Ortak Füzyon Tip II (Geleneksel ve VGG16 Kaynaklı Derin Özellik Füzyonu)		0,8539 Doğruluk 0,8553 Kesinlik 0,8540 Duyarlılık 0,8496 F1-Skoru
--	---	--	--

Literatürde diz osteoartriti (KOA) derecelendirmesi üzerine yapılan çalışmalar incelendiğinde, kullanılan yöntemlerin doğruluk, kesinlik, duyarlılık ve F1-skoru açısından farklı seviyelerde performans sergilediği görülmektedir. Özellikle klasik CNN tabanlı yöntemlerin genellikle %70–75 doğruluk aralığında kaldığı ve sınıflandırma başarısında önemli iyileştirmelere ihtiyaç duyulduğu anlaşılmaktadır. Tiulpin ve ark. (2018) Siamese CNN ile %66,71 doğruluk, Thomas ve ark. (2020) CNN tabanlı analiz ile %71 doğruluk ve %70 F1-skoru, Kwon ve ark. (2020) SVM ve Inception-Net-V2 kullanarak %71 F1-skoru ve %76 hassasiyet, Leung ve ark. (2020) ResNet34 ile %72,7 doğruluk, Zhang ve ark. (2020) ResNet-18 ile %74,81 doğruluk, Chen ve ark. (2019) VGG16 ile %69,7, Solak (2024) Xception ile % 67,8 doğruluk elde etmiştir. Bu çalışmaların büyük bir kısmı geleneksel CNN modellerine dayalı olup, sınıflandırma performansı açısından yeterince yüksek doğruluk sunamamıştır.

Bu tez çalışmasında önerilen Erken Füzyon Tip II, Ortak Füzyon Tip I ve Ortak Füzyon Tip II stratejileri, literatürdeki yöntemlere kıyasla anlamlı bir performans artışı sağlamıştır. Erken Füzyon Tip II stratejisinde geleneksel ve PCA ile indirgenmiş görüntü tabanlı özelliklerin kombinasyonu kullanılarak %73,56 doğruluk, %72,57 kesinlik, %73,57 duyarlılık ve %71,46 F1-skoru elde edilmiştir. Bu sonuçlar, literatürdeki klasik CNN tabanlı yaklaşımlarla kıyaslanabilir düzeyde bir performans sunmaktadır. Ortak Füzyon Tip I stratejisinde geleneksel özellikler ile DenseNet201 modelinden çıkarılan derin özelliklerin birleştirilmesi sonucunda %83,50 doğruluk, %83,50 kesinlik, %83,51 duyarlılık ve %83,15 F1-skoru elde edilmiştir. Ortak Füzyon Tip II stratejisinde ise geleneksel özellikler ile VGG16 modelinden çıkarılan derin özellikler doğrudan birleştirilmiş ve %85,39 doğruluk, %85,53 kesinlik, %85,40 duyarlılık ve %84,96 F1-skoru ile çalışmadaki en yüksek başarı sağlanmıştır.

Elde edilen sonuçlar değerlendirildiğinde, özellikle Ortak Füzyon Tip I ve Ortak Füzyon Tip II stratejilerinin klasik CNN tabanlı yöntemlere kıyasla belirgin bir üstünlük sağladığı görülmektedir. Örneğin, Pi ve ark. (2023) tarafından topluluk tabanlı modeller

kullanılarak %76,33 doğruluk elde edilirken, bu çalışmada %85,39 doğruluk oranına ulaşılmıştır. Gornale ve ark. (2020) tarafından bildirilen %99,23 doğruluk oranı, yalnızca 2000 görüntü içeren küçük bir veri seti ile ve geleneksel makine öğrenmesi yöntemleri kullanılarak elde edildiğinden, geniş veri setleri üzerindeki genelleme performansı açısından doğrudan karşılaştırılabilir nitelikte değildir. Bu tez çalışmasında ise OAI veri setinden 8260 diz X-ışını görüntüsü kullanılarak, büyük ölçekli ve daha güvenilir bir değerlendirme yapılmıştır.

Genel değerlendirme sonucunda, Erken Füzyon Tip II yöntemi literatürdeki birçok çalışma ile kıyaslanabilir düzeyde bir performans sunarken, Ortak Füzyon Tip I ve özellikle Ortak Füzyon Tip II yöntemleri doğruluk oranı açısından literatürdeki pek çok çalışmayı geride bırakmıştır. Özellikle Ortak Füzyon Tip II stratejisi kapsamında geleneksel görüntü özelliklerinin ve VGG16 kaynaklı derin özelliklerin kullanımı ile elde edilen %85'in üzerindeki doğruluk oranı, KOA derecelendirmesi için en etkili yaklaşım olarak öne çıkmaktadır. Bu sonuçlar, önerilen füzyon stratejilerinin KOA derecelerinin daha doğru bir şekilde belirlenmesine olanak tanıdığını ve klinik teşhis süreçlerinde güvenilir bir çözüm sunduğunu göstermektedir.

5. SONUÇLAR VE ÖNERİLER

Bu tez çalışmasında, diz osteoartritinin (KOA) doğru tanısını gerçekleştirmek ve KOA derecelerini (“Sağlıklı”, “Şüpheli”, “Hafif”, “Orta” ve “Ağır”) doğru bir şekilde belirlemek amacıyla bütünleşik bir yaklaşım geliştirilmiştir. Çalışmada hem geleneksel yöntemlerle manuel olarak çıkarılan görüntü özellikleri hem de transfer öğrenme modellerinden elde edilen derin özellikler kullanılmış ve bu iki farklı bilgi türü, Erken Füzyon Tip II, Ortak Füzyon Tip I ve Ortak Füzyon Tip II stratejileri kapsamında entegre edilmiştir. Füzyon stratejileri sonucunda elde edilen birleşik özellik temsilleri, ANN temelli sınıflandırma yapılarında değerlendirilmiş ve KOA derecelerinin belirlenmesi amacıyla sınıflandırma işlemi gerçekleştirilmiştir.

Manuel çıkarılan geleneksel özelliklerin, derin sinir ağı modellerinden elde edilen derin özelliklerle birleştirilmesi, KOA derecelerinin belirlenmesinde önemli bir performans artışı sağlamıştır. Özellikle Ortak Füzyon Tip II stratejisinde, %85,53 kesinlik, %85,40 duyarlılık ve %84,96 F1-skoru elde edilerek literatürde bildirilen birçok çalışmadan daha yüksek doğruluk seviyelerine ulaşılmıştır. Bu başarıda, VGG16 modelinden çıkarılan derin temsillerle birlikte özellikle GLCM tabanlı doku özellikleri, LBP desenleri ve çevre, dairesellik gibi morfolojik ölçümlerin ayırt edici etkisi öne çıkmıştır.

Bu çalışmada yalnızca özellik çıkarımıyla yetinilmemiş; seçilen özelliklerin kalitesini artırmak için veri artırımı ve sınıf dengesizliğini giderme teknikleri de uygulanmıştır. Görüntü verileri üzerinde yapılan veri artırımı ve SMOTE uygulamaları sayesinde az temsil edilen KOA derecelerinin veri setindeki ağırlığı artırılmış ve modelin dengeli bir öğrenme gerçekleştirmesi sağlanmıştır. Ayrıca, çıkarılan geleneksel özelliklerin en anlamlı olanlarının seçilmesi amacıyla LDA ve ANOVA yöntemleri uygulanmış, derin özelliklerin boyutunun düşürülmesi için ise PCA kullanılmıştır. Bu optimizasyon süreçleri sonucunda, KOA derecelerinin tahmini için daha kompakt ve bilgi açısından zengin özellik vektörleri oluşturulmuştur.

Klinik açıdan değerlendirildiğinde, bu çalışmanın en önemli katkısı, Orta ve Ağır KOA derecelerinin %94’ün üzerinde doğruluk oranlarıyla doğru bir şekilde belirlenebilmesidir. Bu başarı, hastalığın ilerlemiş evrelerinin zamanında teşhisi ve erken tedavi sürecinin planlanması için büyük önem taşımaktadır. Ayrıca, sağlıklı bireylerin yüksek doğrulukla tanımlanması, gereksiz tıbbi müdahalelerin önlenmesine

yardımcı olmuştur. Ancak, Şüpheli ve Hafif KOA derecelerinde elde edilen sonuçlar, bu erken evrelerde daha fazla iyileştirmeye ihtiyaç duyulduğunu göstermektedir.

Çalışmanın literatüre katkısı, manuel olarak çıkarılan ve optimize edilen geleneksel özelliklerin, transfer öğrenme modellerinden elde edilen derin özelliklerle doğrudan bütünleştirilmesinin KOA derecelerinin belirlenmesinde doğruluk, kesinlik, duyarlılık ve F1-skoru değerlerini anlamlı şekilde artırdığını göstermesidir. Bu bütünleşik yaklaşım, yalnızca derin öğrenme tabanlı özelliklere dayanan yöntemlere kıyasla daha güvenilir sonuçlar sunmuş ve modelin genelleme kapasitesini artırmıştır. Özellikle Ortak Füzyon Tip II stratejisiyle sağlanan başarı, geleneksel ve derin bilgilerin etkili bir entegrasyonu ile KOA derecelerinin daha doğru tahmin edilebileceğini ortaya koymuştur.

Bu çalışmanın bazı sınırlamaları da bulunmaktadır. Şüpheli ve Hafif KOA derecelerinin belirlenmesinde, bu dereceler arasındaki görsel benzerliklerden kaynaklanan örtüşmelerin, modelin performansını olumsuz etkilediği gözlemlenmiştir. Ayrıca, yöntemin farklı veri setleri üzerinde test edilmemiş olması, modelin genelleme yeteneğinin tam olarak değerlendirilmesini sınırlamıştır.

Gelecek çalışmalar için öneriler şunlardır:

- Şüpheli ve Hafif KOA derecelerinde belirleme başarısını artırmak amacıyla daha gelişmiş özellik mühendisliği yöntemlerinin ve farklı ağ yapılandırmalarının denenmesi,
- Sağlıklı bireylerin yanlış derecelendirilmesini azaltmak için transfer öğrenme çıktılarının daha etkili şekilde optimize edilmesi,
- Ortak Füzyon Tip II stratejisinde, VGG16 modelinden elde edilen derin özelliklerin daha ileri seviyede işlenerek doğruluğun artırılması,
- Geliştirilen yaklaşımın farklı ve genişletilmiş KOA veri setlerinde denenerek genelleme kapasitesinin artırılması,
- Şüpheli ve Hafif dereceler için doğruluk oranlarının yükseltilmesi amacıyla sınıf ağırlıklandırma yöntemleri veya meta-öğrenme tekniklerinin uygulanması.

Bu öneriler doğrultusunda geliştirilecek çalışmalar ile KOA derecelerinin belirlenmesinde daha yüksek başarı oranlarına ulaşılması ve klinik uygulamalarda daha güvenilir sonuçlar elde edilmesi mümkün olacaktır.

KAYNAKLAR

Acar, Emrullah, ve Mehmet Sıraç Özerdem., 2012, Kızıltepe tarımsal Kızıltepe tarımsal alan imgelerinin ekinin ürün gelişimine göre sınıflandırılması, *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, 5(1).

Ajmera, G., 2021, Feature Extraction of Images using GLCM (Gray Level Co-occurrence Matrix) [online], Medium, <https://medium.com/@girishajmera/feature-extraction-of-images-using-glcm-gray-level-cooccurrence-matrix-e4bda8729498> [Ziyaret Tarihi: 17 Kasım 2024].

Alammar, Z., Alzubaidi, L., Zhang, J., Li, Y., Lafta, W. ve Gu, Y., 2023, Deep transfer learning with enhanced feature fusion for detection of abnormalities in x-ray images, *Cancers*, 15 (15), 4007.

Alzubaidi, L., Fadhel, M. A., Albahri, A., Salhi, A., Gupta, A. ve Gu, Y., 2023, Domain adaptation and feature fusion for the detection of abnormalities in X-Ray forearm images, *45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, IEEE, 1–5.

Anderson, A. S. ve Loeser, R. F., 2010, Why is osteoarthritis an age-related disease?, *Best Practice & Research Clinical Rheumatology*, 24 (1), 15–26. doi:10.1016/j.berh.2009.08.006.

Ahmed, S. M. ve Mustafa, R. J., 2022, A comprehensive survey on bone segmentation techniques in knee osteoarthritis research: From conventional methods to deep learning, *Diagnostics*, 12, 611.

Ahmed, S. M. ve Mstafa, R. J., 2022, Identifying severity grading of knee osteoarthritis from X-ray images using an efficient mixture of deep learning and machine learning models, *Diagnostics*, 12, 2939.

Brahim, A., Jennane, R., Riad, R., Janvier, T., Khedher, L., Toumi, H. ve Lespessailles, E., 2019, A decision support tool for early detection of knee osteoarthritis using X-ray imaging and machine learning: Data from the Osteoarthritis Initiative, *Computerized Medical Imaging and Graphics*, 73, 11–18.

Büyüköztürk, Ş., 2012, *Sosyal bilimler için veri analizi el kitabı: İstatistik, araştırma deseni, SPSS uygulamaları ve yorum* (17. Baskı), Ankara: Pegem Akademi Yayıncılık.

Chaugule, S. V. ve Malemath, V. S., 2023, Towards identifying key features in the classification of knee osteoarthritis – An enhanced feature fusion based deep network model. [Çevrimiçi] <https://doi.org/10.21203/rs.3.rs-3166084/v1>

Chawla, N. V., Bowyer, K. W., Hall, L. O. ve Kegelmeyer, W. P., 2002, SMOTE: Synthetic minority over-sampling technique, *Journal of Artificial Intelligence Research*, 16, 321–357.

Chen, P., 2018, Knee osteoarthritis severity grading dataset, *Mendeley Data*, 1. doi:10.17632/56rmx5bjcr.1

Chen, P., Gao, L., Shi, X., Allen, K. ve Yang, L., 2019, Fully automatic knee osteoarthritis severity grading using deep neural networks with a novel ordinal loss, *Computerized Medical Imaging and Graphics*, 75, 84–92.

Chollet, F., 2017, Xception: Deep learning with depthwise separable convolutions, *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1800–1807, USA.

Cueva, J. H., Castillo, D., Espinós-Morató, H., Durán, D., Díaz, P. ve Lakshminarayanan, V., 2022, Detection and classification of knee osteoarthritis, *Diagnostics*, 12 (10), 2362.

Dalal, N. ve Triggs, B., 2005, Histograms of Oriented Gradients for Human Detection, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1(1), 886–893.

Gençtürk, T. H., Gülağız, F. K. ve Kaya, İ., 2023, Derin öğrenme yöntemleri kullanılarak BT taramalarında beyin kanaması teşhisinin karşılaştırmalı bir analizi, *Zeki Sistemler Teori ve Uygulamaları Dergisi*, 6(1), 75–84.

Gornale, S. S., Patravali, P. U. ve Hiremath, P. S., 2020, Automatic detection and classification of knee osteoarthritis using Hu's invariant moments, *Frontiers in Robotics and AI*, 7, 591827.

Guan, B., Liu, F., Mizaian, A. H., Demehri, S., Samsonov, A., Guermazi, A. ve Kijowski, R., 2022, Deep learning approach to predict pain progression in knee osteoarthritis, *Skeletal Radiology*, 51, 363–373.

Guida, C., Zhang, M. ve Shan, J., 2023, Improving knee osteoarthritis classification using multimodal intermediate fusion of X-ray, MRI and clinical information, *Neural Computing and Applications*, 35, 9763–9772.

Haralick, R. M., Shanmugam, K. ve Dinstein, I., 1973, Textural features for image classification, *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6), 610–621. <https://doi.org/10.1109/TSMC.1973.4309314>

He, K., Zhang, X., Ren, S. ve Sun, J., 2016, Deep residual learning for image recognition, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778.

Huang, G., Liu, Z., Van Der Maaten, L. ve Weinberger, K. Q., 2017, Densely connected convolutional networks, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4700–4708. doi:10.48550/arXiv.1608.06993

Huang, S.-C., Pareek, A., Seyyedi, S., Banerjee, I. ve Lungren, M. P., 2020, Fusion of medical imaging and electronic health records using deep learning: A systematic review and implementation guidelines, *NPJ Digital Medicine*, 3, 136.

Jakaite, L., Schetinin, V., Hladůvka, J., Minaev, S., Ambia, A. ve Krzanowski, W., 2021, Deep learning for early detection of pathological changes in X-ray bone microstructures: Case of osteoarthritis, *Scientific Reports*, 11, 1–9.

Kara, M. ve Yıldız, S., 2019, Boyut indirgeme yöntemlerinden PCA'nın çalışma prensibi ve uygulama örnekleri, *İstatistiksel Araştırmalar Dergisi*, 15(2), 45–58.

Kellgren, J. ve Bier, F., 1956, Radiological signs of rheumatoid arthritis: A study of observer differences in the reading of hand films, *Annals of the Rheumatic Diseases*, 15(1), 55–60. doi:10.1136/ard.15.1.55

Kohn, M. D., Sassoon, A. A. ve Fernando, N. D., 2016, Classifications in brief: Kellgren-Lawrence classification of osteoarthritis, *Clinical Orthopaedics and Related Research*, 474(8), 1886–1893. doi:10.1007/s11999-016-4732-4

Kokkotis, C., Moustakidis, S., Giakas, G. ve Tsaopoulos, D., 2020, Identification of risk factors and machine learning-based prediction models for knee osteoarthritis patients, *Applied Sciences*, 10, 6797.

Kokkotis, C., Moustakidis, S., Papageorgiou, E., Giakas, G. ve Tsaopoulos, D. E., 2020, Machine learning in knee osteoarthritis: A review, *Osteoarthritis and Cartilage Open*, 2, 100069.

Kondal, S., Kulkarni, V., Gaikwad, A., Kharat, A., Pant, A., vd., 2022, Automatic grading of knee osteoarthritis on the Kellgren-Lawrence scale from radiographs using convolutional neural networks, *Advances in Deep Learning, Artificial Intelligence and Robotics*, Lecture Notes in Networks and Systems, 249, 163–173.

Kumar, N. ve Garg, R. K., 2020, Human detection using Histogram of Oriented Gradients and SVM, *Materials Today: Proceedings*, 33(1), 3803–3807.

Kwon, S. B., Han, H. S., Lee, M. C., Kim, H. C. ve Ku, Y., 2020, Machine learning-based automatic classification of knee osteoarthritis severity using gait data and radiographic images, *IEEE Access*, 8, 120597.

Lau, L. C. M., Chui, E. C. S., Man, G. C. W., Xin, Y., Ho, K. K. W., Mak, K. K. K., Ong, M. T. Y., Law, S. W., Cheung, W. H. ve Yung, P. S. H., 2022, A novel image-based machine learning model with superior accuracy and predictability for knee arthroplasty loosening detection and clinical decision making, *Journal of Orthopaedic Translation*.

Leung, K., Zhang, B., Tan, J., Shen, Y., Geras, K. J., Babb, J. S. ve Deniz, C. M., 2020, Prediction of total knee replacement and diagnosis of osteoarthritis by using deep learning on knee radiographs: Data from the osteoarthritis initiative, *Radiology*, 296, 584.

Li, W., Yu, S., Yang, R., Tian, Y., Zhu, T., Liu, H., Jiao, D., Zhang, F., Liu, X. ve Tao, L., 2023, Machine learning model of ResNet50-Ensemble voting for malignant–benign small pulmonary nodule classification on computed tomography images, *Cancers*, 15(22), 5417. doi:10.3390/cancers15225417

Lim, J., Kim, J. ve Cheon, S., 2019, A deep neural network-based method for early detection of osteoarthritis using statistical data, *International Journal of Environmental Research and Public Health*, 16, 1281.

Liu, M., Lan, J., Chen, X., Yu, G. ve Yang, X., 2018, Bone age assessment model based on multi-dimensional feature fusion using deep learning, *Academic Journal of Second Military Medical University*, 39, 909–916.

Mahfus, M., Özdemir, M., & Polat, H., 2025, Diz osteoartritinin derin öğrenme yöntemleri ile sınıflandırılması. *Journal of Engineering Sciences and Design (JESD)*, 13(2), 609–618.

Martel-Pelletier, J., 1999, Pathophysiology of osteoarthritis. *Osteoarthritis and Cartilage*, 7(4), 371–373. <https://doi.org/10.1053/joca.1998.0214>

Mohammed, A. S., Hasanaath, A. A., Latif, G., & Bashar, A., 2023, Knee osteoarthritis detection and severity classification using residual neural networks on preprocessed X-ray images. *Diagnostics*, 13, 1380.

Mou, S. F., & Razzak, S. A., 2023, Brain disease classification from MRI scans using EfficientNetB0 feature extraction. In *2023 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)* (pp. 336–340). IEEE.

Muthazhagu, V. H., & Surendiran, B., 2024, Deep learning-based osteoarthritis severity classification from radiographic images (X-ray). *2nd International Conference on Recent Advances in Information Technology for Sustainable Development (ICRAIS)*.

Nasser, Y., Jennane, R., Chetouani, A., Lespessailles, E., & El Hassouni, M., 2020, Discriminative regularized auto-encoder for early detection of knee osteoarthritis: Data from the osteoarthritis initiative. *IEEE Transactions on Medical Imaging*, 39, 2976–2984.

Nia, S. N., & Shih, F. Y., 2024, Medical X-ray image enhancement using global contrast-limited adaptive histogram equalization. *arXiv preprint arXiv:2411.01373*.

Phan, T. H., Nguyen, T. T., Nguyen, T. D., Pham, H. H., Ta, G. K., Tran, M. T., & Quan, T. T., 2025, DIKOApp: An AI-based diagnostic system for knee osteoarthritis. *Journal of Imaging Informatics in Medicine*.

Pi, S. W., Lee, B. D., Lee, M. S., & Lee, H. J., 2023, Ensemble deep-learning networks for automated osteoarthritis grading in knee X-ray images. *Scientific Reports*, 13, 22887.

Raza, A., Ching, C. T. S., Liu, P. T., Phan, T. L., Tran, N. L., Nguyen, H. A. T., Rameez, M., Li, H. C., Linh, D. N. P., Van Hieu, N., & Tran, T. N., 2025, Enhancing

diagnostic accuracy for knee osteoarthritis using a deep learning approach. *Journal of Physics: Conference Series*, 2949(1), 012015.

Rehman, A., Raza, A., Alamri, F. S., Alghofaily, B. ve Saba, T., 2023, Transfer learning-based smart features engineering for osteoarthritis diagnosis from knee X-ray images, *IEEE Access*.

Saleem, M., Farid, M. S., Saleem, S., & Khan, M. H., 2020, X-ray image analysis for automated knee osteoarthritis detection. *Signal, Image and Video Processing*, 14, 1079–1087.

Schiratti, J. B., Dubois, R., Herent, P., Cahané, D., & Dachary, J., et al., 2021, A deep learning method for predicting knee osteoarthritis radiographic progression from MRI. *Arthritis Research & Therapy*, 23(1), 262.

Shaheed, K., Mao, A., Qureshi, I., Kumar, M., Hussain, S., Ullah, I., & Zhang, X., 2022, DS-CNN: A pre-trained Xception model based on depth-wise separable convolutional neural network for finger vein recognition. *Expert Systems with Applications*, 191, 116288. <https://doi.org/10.1016/j.eswa.2021.116288>

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

Solak, F. Z., 2024, Classification of knee osteoarthritis severity by transfer learning from X-ray images. *Karaelmas Science and Engineering Journal*, 14(2), 119–133.

Tahmasbi, A., Saki, F., & Shokouhi, S. B., 2011, Classification of benign and malignant masses based on Zernike moments. *Computers in Biology and Medicine*, 41(8), 726–735.

Tan, M., & Le, Q., 2019, EfficientNet: Rethinking model scaling for convolutional neural networks. In *International Conference on Machine Learning* (pp. 6105–6114). California.

Thomas, K. A., Kidziński, Ł., Halilaj, E., Fleming, S. L., Venkataraman, G. R., Oei, E. H., & Delp, S. L., 2020, Automated classification of radiographic knee osteoarthritis severity using deep neural networks. *Radiology: Artificial Intelligence*, 2, e190065.

Tiulpin, A., Thevenot, J., Rahtu, E., Lehenkari, P., & Saarakkala, S., 2018, Automatic knee osteoarthritis diagnosis from plain radiographs: A deep learning-based approach. *Scientific Reports*, 8, 1727.

Tiulpin, S., Klein, S., Bierma-Zeinstra, S. M. A., Jérôme, T., Esa, R., et al., 2019, Multimodal machine learning-based knee osteoarthritis progression prediction from plain radiographs and clinical data. *Scientific Reports*, 9(1), 20038.

Tolpadi, A. A., Lee, J., Pedoia, V., & Majumdar, S., 2020, Deep learning predicts total knee replacement from magnetic resonance images. *Scientific Reports*, 10(1), 6371.

Turkoglu, M., 2021, COVID-19 detection system using chest CT images and multiple kernels-extreme learning machine based on deep neural network. *IRBM*, 42(4), 207–214. <https://doi.org/10.1016/j.irbm.2021.01.004>

von Tyrowicz, C., 2020, Towards shape-based knee osteoarthritis classification using graph convolutional networks. In *Proceedings of the 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)* (pp. 750–753). Piscataway, NJ, USA: IEEE.

Wang, Y., Li, S., Zhao, B., Zhang, J., Yang, Y., & Li, B., 2022, A ResNet-based approach for accurate radiographic diagnosis of knee osteoarthritis. *CAAI Transactions on Intelligent Technologies*, 7, 512–521.

Wang, Y., Wang, X., Gao, T., Du, L., & Liu, W., 2021, An automatic knee osteoarthritis diagnosis method based on deep learning: Data from the Osteoarthritis Initiative. *Journal of Healthcare Engineering*, 2021, 5586529.

Yang, L., Xu, S., Yu, X., Long, H., Zhang, H., & Zhu, Y., 2023, A new model based on improved VGG16 for corn weed identification. *Frontiers in Plant Science*, 14, 1205151. <https://doi.org/10.3389/fpls.2023.1205151>

Yong, C. W., Teo, K., Murphy, B. P., Hum, Y. C., Tee, Y. K., Xia, K., Lai, K. W., 2022, Knee osteoarthritis severity classification with ordinal regression module. *Multimedia Tools and Applications*, 1–13.

Zhang, B., Tan, J., Cho, K., Chang, G., & Deniz, C. M., 2020, Attention-based CNN for KL grade classification: Data from the osteoarthritis initiative. In *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, IEEE, 731–735.

Zhang, W., Li, X., Liu, Y., Wang, L., Zhao, Y., & Chen, X., 2023, Early detection and classification of knee osteoarthritis based on convolutional neural networks. *Journal of Healthcare Engineering*, 2023, 9296387. <https://doi.org/10.1155/2023/9296387>