

**VISION TRANSFORMER TABANLI ÖZ DENETİMLİ
ÖĞRENME İLE HİSTOPATOLOJİK GÖRÜNTÜLERDE MİTOZ
HÜCRELERİNİN TESPİTİ**

MUHAMMED CİHAD ARSLANOĞLU

HAZİRAN 2024

DİYARBAKIR

DİCLE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**VİSİON TRANSFORMER TABANLI ÖZ DENETİMLİ
ÖĞRENME İLE HİSTOPATOLOJİK GÖRÜNTÜLERDE MİTOZ
HÜCRELERİNİN TESPİTİ**

MUHAMMED CİHAD ARSLANOĞLU

DİCLE ÜNİVERSİTESİ LİSANSÜSTÜ EĞİTİM-ÖĞRETİM VE SINAV
YÖNETMELİĞİNİN BİR PARÇASI OLARAK
ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANA BİLİM DALINDA
YÜKSEK LİSANS TEZİ
OLARAK HAZIRLANMIŞTIR

HAZİRAN 2024

DİYARBAKIR

VİSİON TRANSFORMER TABANLI ÖZ DENETİMLİ ÖĞRENME İLE HİSTOPATOLOJİK GÖRÜNTÜLERDE MİTOZ HÜCRELERİNİN TESPİTİ

Muhammed Cihad ARSLANOĞLU tarafından Dicle Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliği'nin bir parçası olarak hazırlanan bu çalışma, aşağıda bilgileri yazılı jüri üyeleri tarafından değerlendirilerek **Elektrik-Elektronik Mühendisliği Ana Bilim Dalı'nda Yüksek Lisans Tezi** olarak kabul edilmiştir.

Prof. Dr. Neslihan DALKILIÇ
Müdür, **Fen Bilimleri Enstitüsü**

Dr. Öğr. Üyesi Hüseyin ACAR
Danışman, **Elektrik-Elektronik Mühendisliği Bölümü,**
Dicle Üniversitesi

Doç. Dr. Abdülkadir ALBAYRAK
İkinci Danışman, **Bilgisayar Mühendisliği Bölümü,**
Dicle Üniversitesi

Sınav Jürisi:

Prof. Dr. Mehmet Sıraç ÖZERDEM (*)
Elektrik-Elektronik Mühendisliği Bölümü, Dicle Üniversitesi

Dr. Öğr. Üyesi Hüseyin ACAR (**)
Elektrik-Elektronik Mühendisliği Bölümü, Dicle Üniversitesi

Dr. Öğr. Üyesi Ömer TÜRK
Bilgisayar Mühendisliği Bölümü, Mardin Artuklu Üniversitesi

ONAY

Savunma Tarihi: 25/06/2024

(*) Jüri Başkanı.

(**)Tez Danışmanı.

Aileme...



Dicle Üniversitesi Fen Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak hazırlanan bu tez çalışmasında yer alan tüm bilgilerin akademik kurallara ve etik ilkelere uygun olarak elde edildiğini ve sunulduğunu beyan ederim. Ayrıca, bahse konu bu kural ve ilkelerin gerektirdiği üzere, bu çalışmada özgün olmayan tüm bilimsel içerikleri kurallara uygun biçimde alıntılıyıp kaynak gösterdiğimi beyan ederim. Beyanımınla çelişen herhangi bir delil bulunduğu takdirde tüm sorumluluğu üstleneceğimi kabul ederim.

Ad, Soyad: Muhammed Cihad ARSLANOĞLU

İmza:

TEŞEKKÜR

Öncelikle danışmanım Dr. Öğr. Üyesi Hüseyin ACAR'a ve ikinci danışmanım Doç. Dr. Abdülkadir ALBAYRAK'a çalışmanın daha fikir aşamasından, örnekleme, deney setlerini oluşturma, analizlerin yorumlanması ve tez taslağındaki düzeltmeler konularına kadar gösterdikleri destek ve sabırdan dolayı teşekkür ederim. Jüri üyelerime de bu tezin daha nitelikli bir hale gelmesi için yaptıkları eleştiri ve önerileri için teşekkür ederim.

En önemlisi, beni araştırma hevesiyle coşturdukları ve kendime bazı zamanlar inanmasam da bana inandıkları için anneme, babama, abime ve ablama teşekkür ederim.

İÇİNDEKİLER

TEŞEKKÜR.....	v
İÇİNDEKİLER	vi
ŞEKİLLER LİSTESİ	viii
TABLolar LİSTESİ.....	ix
SİMGELER VE KISALTMALAR LİSTESİ	x
ÖZET.....	xiii
ABSTRACT.....	xiv
1. GİRİŞ	1
2. ÖNCEKİ ÇALIŞMALAR.....	5
3. MALZEME VE YÖNTEM	11
3.1 Veri Toplama.....	11
3.1.1 MIDOG2021	13
3.1.2 MIDOG2022	13
3.1.3 ICPR2014.....	14
3.2 Bölütleme	14
3.2.1 Ortanca filtresi.....	14
3.2.2 K-ortalamalar	15
3.2.3 Bölge analizi (Blob analysis)	15
3.3 Veri Ön İşleme	15
3.3.1 Yeniden boyutlandırma.....	19
3.3.2 Normalizasyon	19
3.4 Öznitelik Çıkarma	20
3.4.1 Temel modeller	21
3.4.2 Özdenetimli öğrenme	22
3.4.3 Self-distillation with no labels (DINO).....	22
3.4.4 Vision transformer	26

3.4.5	Çapraz kovaryans imge dönüşümü (XCiT).....	34
3.4.6	ResNet.....	35
3.4.7	Yerel ikilik desen (LBP)	37
3.5	Sınıflandırma	38
3.5.1	SVM.....	39
4.	BULGULAR VE TARTIŞMA	42
4.1	Kurulum.....	42
4.2	Değerlendirme Yöntemi	42
4.2.1	Duyarlılık	42
4.2.2	Kesinlik	43
4.2.3	F-skoru	43
4.3	DeneySEL Sonuçlar	43
5.	SONUÇLAR VE ÖNERİLER	59
5.1	Sonuçlar.....	59
5.2	Öneriler.....	62
	KAYNAKLAR	64
	ÖZGEÇMİŞ	71

ŞEKİLLER LİSTESİ

Şekil 3.1 Bilgisayarlı görü işlem adımları	11
Şekil 3.2 MIDOG2021 veri setinden örnek görüntüler.....	16
Şekil 3.3 MIDOG2022 veri setinden örnek görüntüler.....	17
Şekil 3.4 ICPR2014 veri setinden örnek görüntüler	18
Şekil 3.5 DINO algoritmasının genel mimarisi.....	24
Şekil 3.6 ViT algoritmasının genel mimarisi	27
Şekil 3.7 XCIT algoritmasının mimarisi.....	35
Şekil 3.8 ResNet ve normal evrişim katmanları	36
Şekil 3.9 ResNet50 mimarisi.....	37
Şekil 3.10 LBP'nin çalışma prensibinin akış diyagramı	38
Şekil 3.11 Doğrusal olarak ayrılan ve ayrılmayan veri kümeleri.....	39
Şekil 3.12 verinin boyutunun artırılmış ve artırılmamış hali	40
Şekil 3.13 $z = x^2 + y^2$ dönüşümü.....	40
Şekil 4.1 MIDOG2021 ve MIDOG2022 elde edilen sonuçların grafiği.....	48
Şekil 4.2 MIDOG2021 ve MIDOG2022 hata matrisleri.....	49
Şekil 4.3 ICPR2014 denetimli öğrenme doğruluk ve hata grafiği.....	55
Şekil 4.4 ICPR2014 veri kümesi için hata matrisleri.....	56

TABLÖLAR LİSTESİ

Tablo 4.1 Kullanılan modellerin bazı özellikleri	44
Tablo 4.2 MIDOG2021 veri kümesi için farklı modellerin performansları.....	45
Tablo 4.3 MIDOG2022 veri kümesi için farklı modellerin performansları.....	45
Tablo 4.4 MIDOG2021 yarışması için takımların test veri kümesindeki model performansları.	46
Tablo 4.5 MIDOG2022 yarışması için takımların test veri kümesindeki model performansları.	46
Tablo 4.6 MIDOG2021 yarışması için literatürde bulunan çalışmaların performansları.	47
Tablo 4.7 MIDOG2022 veri kümesi için literatürde bulunan çalışmaların performansları.	48
Tablo 4.8 ICPR2014 1 kat dengesizlik durumunda farklı modellerin performansları	50
Tablo 4.9 ICPR2014 2 kat dengesizlik durumunda farklı modellerin performansları	50
Tablo 4.10 ICPR2014 3 kat dengesizlik durumunda farklı modellerin performansları	52
Tablo 4.11 ICPR2014 4 kat dengesizlik durumunda farklı modellerin performansları	53
Tablo 4.12 ICPR2014 5 kat dengesizlik durumunda farklı modellerin performansları	54
Tablo 4.13 ICPR2014 dengesiz veri kümeleri örnek sayısı.	54
Tablo 4.14 ICPR2014 veri kümesi için literatürde yapılan çalışmaların sonuçları. ..	54

SİMGELER VE KISALTMALAR LİSTESİ

Simge	Açıklama
σ	Standart Sapma
I	İmge
m	İmgenin piksel yoğunluklarının ortalaması
I_n	Normalize edilmiş imge
t	DINO modelinin son katmandaki gömme değerleri
t_o	DINO öğretmen modelinin çıktısının merkezleme ve softmax fonksiyonunun uygulanmış hali
s_o	DINO öğrenci modelinin çıktısının merkezleme ve softmax fonksiyonunun uygulanmış hali
C	Merkezleme değeri
τ_s	Öğrenci ısınma katsayısı
τ_t	Öğretmen ısınma katsayısı
t_1^g	x_1^g 'nin DINO modeli ile çıkarılmış öz nitelikleri
t_2^g	x_2^g 'nin DINO modeli ile çıkarılmış öz nitelikleri
x_1^g	DINO modeli için 1. genel parça
x_2^g	DINO modeli için 2. genel parça
d	DINO çıktı gömme değerleri boyutu
$H(.)$	Çapraz Entropi Hatası
$P_t(.)$	Öğretmen modelinin çıktısı
$P_s(.)$	Öğrenci modelinin çıktısı
V	DINO genel ve yerel parçalarını bulunduran küme

x	DINO bütün genel ve yerel parçalar
x'	DINO her bir genel ve yerel parçalar
W_t	DINO öğretmen modelinin ağırlıkları
W_s	DINO öğrenci modelinin ağırlıkları
l	DINO modelinin ağırlıklarının güncellenmesi için momentum katsayısı
QK	ViT dikkat mekanizması ilişki matrisi
Q	ViT algoritması Query gömme değerleri
K	ViT algoritması Key gömme değerleri
V	ViT Algoritması Value gömme değerleri

Kısaltma

Açıklama

CPU	Merkezi İşlem Birimi
SVC	Destek Vektör Sınıflandırıcısı
RF	Rastgele Orman
CNN	Evrişimsel Sinir Ağları
ViT	Vision Transformer
ResNet	Residual Network
DINO	self-distillation with no labels
WSI	Whole Slide Image
YOLO	You Only Look Once
HPF	Yüksek Güçlü Alan
HOG	Histogram of Oriented Gradient
MIDOG	Mitosis Domain Generalization
ICPR	International Conference on Pattern Recognition
LBP	Local Binary Pattern
SURF	Speedup Robust Features
XCIT	Cross Variational Image Transform

NLP	Natural Language Processing
MLP	Multi Layer Perceptron
XCA	Çapraz Kovaryans Dikkat
LPI	Local Patch Interaction
SVM	Support Vector Macine
RBF	Radial Based Function
GPU	Graphical Processing Unit
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
R	Recall
P	Precision
f	F-skoru

ÖZET

VİSİON TRANSFORMER TABANLI ÖZ DENETİMLİ ÖĞRENME İLE HİSTOPATOLOJİK GÖRÜNTÜLERDE MİTOZ HÜCRELERİNİN TESPİTİ

ARSLANOĞLU, M. Cihad

Yüksek Lisans, Elektrik-Elektronik Mühendisliği Bölümü

Danışman: Dr. Öğr. Üyesi Hüseyin ACAR

İkinci Danışman: Doç. Dr. Abdülkadir ALBAYRAK

Haziran 2024, 86 sayfa

Doku içerisindeki mitozlu hücrelerin belirlenmesi, kanser, parkinson, viral enfeksiyonlar gibi hastalıkların teşhis edilmesi ve uygulanacak tedavinin belirlenmesi açısından büyük önem taşımaktadır. Doku bilimciler, hastadan alınan doku örneklerini manuel olarak inceleyerek mitozlu hücreleri tespit ederler. Ancak bu işlem, doktorlar için zaman alıcı ve yorucudur. Ayrıca, çok sayıda hücre inceleme sürecine maruz kalan doktorların hata yapma olasılığı da artmaktadır. Bilgisayarlı görü teknikleri kullanılarak bu işlemlerin otomatik bir şekilde yapılması ve mitozlu hücrelerin tespit edilmesi mümkündür. Mitozlu hücrelerin bilgisayarlı görü ile tespit edilmesi, harcanan zamanı ve hata oranını azaltacaktır. Bu çalışmada, geleneksel makine öğrenimi ve derin öğrenme yöntemlerinin biyomedikal görüntülerde hücre tespiti performansları değerlendirilmiştir. Geleneksel makine öğrenmesi yöntemi olarak Yerel İkili Desen algoritması, derin öğrenme yöntemi olarak ise Transformer ve CNN tabanlı olmak üzere 3 farklı yöntem kullanılmıştır. Transformer tabanlı yöntemler Vision Transformer, XCIT algoritmalarıdır. CNN tabanlı yöntem olarak ise ResNet50 algoritması kullanılmıştır. Derin öğrenme tabanlı yöntemler sadece denetimli olarak değil özdenetimli öğrenme yaklaşımı kullanılarak da eğitilmiş ve özellikle temel modellerin biyomedikal görüntüler üzerindeki performansları değerlendirilmiştir. Elde edilen sonuçların güvenilirliğini artırmak ve en iyi performansı elde etmek için yapılan deneyler derin öğrenme modellerinin farklı parametreleri ile tekrarlanmış ve sonuçlar paylaşılmıştır. Dengesiz veri kümesi problemi geleneksel makine öğrenmesi veya derin öğrenme algoritmaları için aşılması gereken bir problemdir. Dengesiz veri kümesi problemi kullanılan algoritmaların performansını düşürmekte ve istenilen sonucun elde edilmesini engellemektedir. Kullanılan yöntemlerin biyomedikal görüntülerde oldukça sık karşılaşılan dengesiz veri kümesi problemine karşı ne kadar dayanıklı oldukları da deneyler vasıtasıyla gözlemlenmiş ve sonuçları paylaşılmıştır. Çalışmada, algoritmaların performanslarının ölçülebilmesi için literatürde eski bir veri kümesi olan ICPR2014 ile son yıllarda yayımlanan MIDOG2021 ve MIDOG2022 veri kümeleri kullanılmıştır.

Anahtar Kelimeler: Vision transformer, CNN, Sınıflandırma, Temel modeller, Özdenetimli öğrenme

ABSTRACT

DETECTION OF MITOSIS CELL IN HISTOPATHOLOGICAL IMAGES USING ViT-BASED SELF-SUPERVISED LEARNING

ARSLANOĞLU, M. Cihad

Master of Science in Department of Electrical and Electronics Engineering

Supervisor: Assist. Prof. Dr. Hüseyin ACAR

Co-Supervisor: Assoc. Prof. Dr. Abdülkadir ALBAYRAK

June 2024, 86 pages

The identification of mitotic cells within tissue is of great importance for the diagnosis of diseases such as cancer, parkinson's, and viral infections, as well as for determining the appropriate treatment. Pathologist manually examine tissue samples taken from patients to detect mitotic cells. However, this process is time-consuming and tiring for doctors. Additionally, the likelihood of errors increases as doctors are exposed to examining a large number of cells. It is possible to automate these processes using computer vision techniques to detect mitotic cells. The detection of mitotic cells through computer vision will reduce the time spent and the error rate. In this study, the performance of traditional machine learning and deep learning methods in cell detection in biomedical images was evaluated. The Local Binary Pattern algorithm was used as the traditional machine learning method, while three different methods were used as deep learning methods: Transformer-based and CNN-based approaches. The Transformer-based methods are Vision Transformer and XCIT algorithms, and the CNN-based method is the ResNet50 algorithm. Deep learning-based methods were trained not only using supervised learning but also using a self-supervised learning approach to evaluate and the performance of foundational models on biomedical images. To increase the reliability of the results and achieve the best performance, experiments were repeated with different parameters of deep learning models, and the results were shared. The imbalanced dataset problem is a significant issue for both traditional machine learning and deep learning algorithms. It reduces the performance of the algorithms and hinders the achievement of the desired results. The robustness of the methods against the imbalanced dataset problem, which is frequently encountered in biomedical images, was also observed through experiments and the results were shared. In the study, to measure the performance of the algorithms, an older dataset from the literature, ICPR2014, as well as more recent datasets, MIDOG2021 and MIDOG2022, were used.

Keywords: Vision Transformer, CNN, Classification, Foundation models, self-supervised learning

1. GİRİŞ

Histopatoloji, temel ve genel dokuları mikroskopik düzeyde inceleyerek kanser, parkinson ve viral enfeksiyon gibi birçok hastalığın daha kesin tanımlarının konulmasında ve yapılacak olan tedavinin seçilmesinde oldukça önemli bir role sahip olan tıbbi disiplin veya bilim dalıdır. Histopatoloji, aynı zamanda patolojik histoloji veya daha sade bir dil ile doku bilimi olarak da bilinmektedir. Tıp eğitimi almış ve doku bilimi alanında uzmanlaşmış kişiler ise “Histopatolog” veya “Doku Bilimci” olarak adlandırılırlar. Doku bilimciler bir hastalığa ait tanının bırakılması için dokuların şekilsel ve işlevsel açıdan bütünlüğünü inceler ve durum hakkında karar verir. Canlılara ait dokuların doku bilimciler tarafından incelenmesi 5 aşamadan oluşmaktadır. Bunlar hastalıklı bölgeye ait doku örneğinin alınması (biyopsi örneği), doku örneklerinin “mikrotom” denilen cihazlar ile uygun boyutlara kesilmesi, dokuların ve patolojik özelliklerin daha ayırt ediciliğini ön plana çıkarmak için “Hematoksilen-eozin” gibi boyalar ile boyanması, uygun hale getirilen dokuların doku bilimciler tarafından mikroskoplar altında incelenmesi ve son olarak elde edilen bulguların raporlanması aşamalarıdır. Bu süreçlerin uzun sürmesi ve özellikle incelemelerin insanlar tarafından yapılması oldukça fazla zaman kaybına sebep olmakta ve doku bilimcilerde “dikkatsizlik körlüğü” durumunun meydana gelmesi sebebiyle hata payı artabilmektedir [1]. Doku bilimcilerin işlerini kolaylaştırmak adına “QuPath” gibi bazı açık kaynaklı veya lisanslı uygulamalar vardır fakat bunlar da manuel ilerletilen sürecin otomatikleştirilmesi noktasında yetersiz kalmaktadır [2].

Mitoz, bir hücrenin ikiye bölünme suretiyle çoğalma evresine verilen isimdir. Kanser ve parkinson gibi birçok hastalığın tespit edilmesinde hücrelerin mitoz evresini incelemek önemli bilgiler sağlamaktadır. Bir bölgedeki mitozlu hücre sayısı belirlenerek kanser tespiti yapmak mümkünken kanserin türünün saptanması noktasında ise bu hücrelerin boyutları etkin bir rol almaktadır. Bunun yanı sıra hastalığa uygulanacak olan tedavi yönteminin seçilmesinde ve hastanın tedaviye karşı tepkisinin belirlenmesi noktasında da mitozlu hücrelerin durumları dikkate alınmaktadır.

Bilgisayar donanımlarının gelişmesiyle beraber bilgisayarlı görü ve derin öğrenme algoritmalarının da gelişim hızı artmıştır. Güçlü donanımlar derin öğrenme

algoritmalarının ihtiyaç duyduğu yüksek işlem gücüne karşılık vermiş ve performanslarının artmasında büyük bir katkı sağlamıştır. Bilgisayarlı görü ve derin öğrenme tekniklerinin bu gelişimi ile beraber manuel olarak yapılan inceleme süreçlerinin en aza indirgenmesi mümkündür. Doku içerisindeki normal olmayan hücrelerin belirlenmesi ve sayılması gibi uygulamalar bilgisayarlar tarafından oldukça kısa bir sürede yapılabilmektedir [3]. Klasik görüntü işleme algoritmaları manuel olarak uzmanlar tarafından geliştirilmesi sebebiyle bazı kısıtlara sahiptir. Klasik görüntü işleme algoritmalarının ortam parlaklığı, cihaz çözünürlüğü ve doku kalitesi gibi birçok dış etmenlere karşı hassas olması sebebiyle her zaman iyi sonuç elde etmek mümkün değildir [4]. Öte yandan Evrimsel Sinir Ağları (CNN) ve Vision Transformer (ViT) gibi derin öğrenme algoritmaları çok daha fazla öğrenme kapasitesine sahip olması ve otonom bir şekilde probleme özgü çözümler geliştirmesi sebebiyle histopatolojik görüntülerin işlenmesinde daha yüksek başarımlar elde etmek mümkündür [5]. Bu tür yaklaşımlar kullanılarak yüksek çözünürlüklü histopatolojik görüntülerin incelenmesi ve doku bilimciler tarafından harcanan sürenin en aza indirgenmesiyle hastanın tedavisine daha erken başlanması mümkündür.

Derin öğrenme modellerinin performansının artırılmasında verinin kalitesi ve sayısı oldukça büyük bir önem arz etmektedir. Veri çeşitliliğinin fazla ve kalitesinin iyi olması modelin problemi daha iyi ifade edebilmesine olanak sağlamaktadır [6], [7]. Her ne kadar internet dünyasında oldukça fazla veri bulunsun da medikal görüntüler gibi belirli bir alan için veriler yetersiz olabilmekte veya veri gizliliği sebebiyle herkese açık bir şekilde paylaşılması mümkün olmamaktadır. Veri kalitesinin ve sayısının az olması sebebiyle derin öğrenme modellerinin bazı alanlar için kullanılması mümkün olmamakta ya da beklenen performans elde edilememektedir. Bu tür problemlerin üstesinden gelmek için derin öğrenme modelleri öncelikle çok büyük veri setleri ile eğitilmekte ve öğrenme aktarımı (transfer learning) veya ince ayar (fine tuning) yaklaşımları kullanılarak asıl çözülmek istenen probleme odaklanılmaktadır. Böylece derin öğrenme modelleri büyük veri kümelerinden kıymetli bilgileri çıkarmayı öğrenir ve bu bilgileri kullanarak daha küçük veri kümelerini öğrenmeye çalışır.

Derin öğrenme modellerinin çok büyük veri kümeleri ile eğitilerek elde edilen modellere temel modeller (foundation models) denilmektedir. Temel modellerin fazla veriye ihtiyaç duyması sebebiyle eğitilmesi oldukça yüksek maliyetlere ihtiyaç duymaktadır. Fakat bu modeller büyük veri kümeleri kullandığı için yerel minimum noktalarına takılma olasılıkları çok daha düşüktür. Bu modeller eğitilmeden direkt olarak çözülmek istenen problemler üzerinde kullanılabileceği gibi ince ayar ve bilgi aktarımı gibi yaklaşımlar ile eğitilmesi de mümkündür. Temel modellerin ince ayar ve bilgi aktarımı ile eğitilmesi modelin çözülmek istenen problem için performansı artabilmektedir. Bunun yanı sıra temel modelleri ince ayar ve bilgi aktarımı yaklaşımları ile eğitmek modelleri sıfırdan eğitmekten çok daha kısa bir süreye ve maliyete ihtiyaç duymaktadır.

Derin öğrenme modellerinin eğitilmesi aşamasında kullanılan farklı öğrenme yöntemleri mevcuttur. Bunlardan birkaçı denetimli, öz denetimli veya denetimsiz öğrenme yöntemleridir. Derin öğrenme modellerinin denetimli öğrenme yöntemleri ile eğitilebilmesi için uzman bilgisine ihtiyaç duyulurken denetimsiz ve öz denetimli öğrenme yöntemleri verinin kendi içerisindeki ilişkiyi kullanarak öğrenme işlemini gerçekleştirir. Temel modellerin eğitim aşamasında büyük veri kümelerine ihtiyaç olması sebebiyle denetimli öğrenme yaklaşımı, denetimsiz ve öz denetimli öğrenme yöntemlerine göre maliyet açısından daha pahalı olabilmektedir. Bu sebeple temel modellerin eğitilmesinde çoğunlukla uzman bilgisine ihtiyaç duyulmayan öz denetimli ve denetimsiz öğrenme gibi yöntemler kullanılmaktadır.

Tespit veya sınıflandırma gibi bilgisayarlı görü uygulamalarında kullanılan görüntülerde kullanılan veri kümesinin sınıflar arası örnek sayısı geliştirilen algoritmanın performansına etki eden önemli bir parametredir. Veri kümesinde bir sınıfın örnek sayısının diğer sınıfın örnek sayısına karşılık ezici bir şekilde fazla olması modelin performansını düşürmekte ve tespit edilmek istenen sınıfın tespit edilme işlemini güçleştirmektedir. Veri kümesi içerisinde bulunan her bir sınıfa ait örnek sayısının yakın olmaması veri dengesizliği olarak adlandırılmaktadır. Bu tür veri kümeleri aynı zamanda uzun kuyruklu dağılıma (long-tailed distribution) sahip veri kümeleri olarak da adlandırılır.

Bu çalışmada mitozlu hücre tespiti yapılabilmesi için 3 farklı veri setinde 3 farklı yaklaşım, 4 farklı algoritma ve bunların farklı parametreler ile oluşturulmuş mimarileri kullanılmıştır. Geleneksel görüntü işleme tekniklerinden Dairesel Yerel İkilik Desen (Circular Local Binary Pattern, LBP), derin öğrenme tekniklerinden CNN tabanlı ResNet-50 [8] ve Transformer tabanlı Vision Transformer (ViT) [9] ve Çapraz Kovaryans İmge Dönüşümü (XCIT) [10] algoritması kullanılmıştır. Bu modellerin bir kısmı denetimli öğrenme algoritması ile eğitilirken bir kısmı ise özdenetimli öğrenme algoritması olan self-distillation with **no** labels (DINO) [11] yaklaşımı ile eğitilmiştir. Kullanılan her bir modelin çıkardığı gömme değerleri ve öznitelikler Destek Vektör Sınıflandırıcısı (Support Vector Classifier, SVC) ile eğitilmiş ve nihai sonuçlar elde edilmiştir. Bunun yanı sıra ICPR veri kümesi için veri dengesizliğinin kullanılan algoritmanın performansına olan etkisini gözlemek için farklı deneyler yapılmıştır.

2. ÖNCEKİ ÇALIŞMALAR

Histopatolojik görüntülerden mitoz hücre tespiti ve sınıflandırılması için literatürde birçok yöntem kullanılmıştır. Kullanılan yöntemlerin genel olarak geleneksel görüntü işleme yöntemleri ve derin öğrenme tabanlı yöntemler olarak iki başlıkta toplanması mümkündür. Donanım kapasitesinin yetersiz olduğu dönemlerde yapılan çalışmalar temelinde ağırlıklı olarak geleneksel görüntü işleme yöntemlerini barındırmaktadır. Donanım kapasitelerinin artmasının ardından derin öğrenme doku bilimi de dahil olmak üzere birçok alana yön vermeye başlamıştır. Her iki başlık altındaki yöntemlerin birbirine karşı üstünlükleri ve zayıflıkları bulunmaktadır.

Geleneksel görüntü işleme algoritmalarının geliştirilmesinde insan sezgileri ve matematiksel işlemlerin oldukça fazla ön planda olması sebebiyle nasıl çalıştıklarının anlaşılması kolaydır. Bu tür yöntemlerin geliştirilmesinde problem ile ilgili uzman bilgisine ihtiyaç duyulmaktadır. Geleneksel görüntü işleme yöntemleri çok fazla işlem gücüne ihtiyaç duymaması sebebiyle Merkezi İşlem Birimi (CPU) gibi seri işlem yapabilen donanımlar üzerinde de çalıştırılabilmektedir. Fakat geleneksel görüntü işleme yöntemleri ortam ışığı, gürültü veya görüntü kalitesi gibi birçok dış etmene karşı oldukça hassastır. Görüntülerin alınması için kullanılan mikroskobun değişmesi durumunda bile bu tür yöntemler yeni veriler için başarısız olabilmektedir.

Derin öğrenme yöntemlerinin temelleri istatistiksel hesaplamalara dayanmaktadır. Geleneksel görüntü işleme yöntemlerinin aksine derin öğrenme yöntemleri bir öğrenme sürecine ihtiyaç duymaktadır. Bu süreci ise problemi ifade edecek veri kümelerini kullanarak gerçekleştirmektedir. Kullanılan veri kümesinin kalitesi ve boyutu ise kullanılan yöntemin performansını doğrudan etkilemektedir. Derin öğrenme yöntemlerinin genelleyebilme yeteneği geleneksel görüntü işleme yöntemlerine kıyasla çok daha fazladır ve algoritmanın performansı ortam koşullarından çok daha az etkilenmektedir. Derin öğrenme yöntemlerinin nasıl çalıştığı sezgisel olarak anlaşılabilmesi geleneksel görüntü işleme yöntemlerine göre çok daha fazla karmaşıktır. Bunun yanı sıra derin öğrenme yöntemleri oldukça fazla paralel işlem gücüne ihtiyaç duymaktadır.

Faridi vd. histopatolojik görüntülere yumuşatma ve morfolojik filtreler uygulayarak kanserli çekirdeklerin merkezlerini tespit eder. Daha sonra her bir çekirdeğin sınırlarını belirlemek için seviye kümesi (level set) algoritmasını kullanır ve kanserli çekirdekleri tespit eder [12].

Irshad vd. histopatolojik görüntülerdeki mitoz hücrelerinin daha belirgin hale gelebilmesi için tespit aşamasından önce imgenin HSV, Lab, Luv ve Mavi Oranı renk uzaylarını hesaplar. Görüntüde bulunan aday çekirdek bölgelerinin tespiti için Mavi Oran renk kanalına birtakım kenar tespiti, eşikleme ve morfolojik filtreler uygular. Tespit edilen olası çekirdek adaylarının toplam 143 olmak üzere alan ve çevre gibi bazı istatistiksel özniteliklerini çıkartır. Çıkarılan özniteliklerden sonuca etkisinin en fazla olanlarını seçmek için altküme değerlendirme (subset evaluation) yöntemini kullanır. Son olarak bu öznitelikler rastgele orman (random forest RF) sınıflandırma algoritmasına girdi olarak verilir ve bölgenin mitoz hücresi olup olmadığı tespit eder [13].

Maroof vd. imgenin Mavi Oran renk uzayını hesaplar ve olası çekirdek bölgelerin bölge analizi (blob analysis) yaparak seçer. Tespit edilen bölgelerden morfolojik, istatistiksel ve renksel öznitelikler çıkarır. Çıkarılan öznitelikler kullanılarak SVC ve RF sınıflandırma algoritmaları eğitilir ve seçilen bölgenin mitoz hücresi olup olmadığı kararını verir [14].

Zakariapour vd. histopatolojik görüntülerdeki gürültü ve bulanıklık gibi problemleri en aza indirmek için ön işlem adımımda imgelere uyarlanabilir histogram eşitlemesi (adaptive histogram equalization) uygular. Ön işleminden geçirilmiş imgelerin Mavi Oran renk uzayı hesaplanır ve uyarlanabilir eşikleme (adaptive thresholding) yöntemi kullanır. Eşikleme yöntemi sonucu elde edilen görüntüden gürültü ve istenmeyen detayların temizlenmesi için morfolojik filtreler ve yumuşatma algoritmaları uygular. Tespit edilen çekirdeklere ait toplam 112 öznitelik oluşturulur. Elde edilen öznitelikler 11 farklı sınıflandırıcı ve topluluk öğrenmesi (ensemble learning) yaklaşımı ile eğitilir ve mitoz hücre tespiti yapılır [15].

Rehman vd. mitoz hücrelerinin tespit edilebilmesi için doku öznitelikleri belirleyici olduğunu söyler ve 3 farklı öznitelik kümesi olmak üzere toplam 89 tane öznitelik çıkarır. Çıkarılan öznitelikler Genişletilmiş Yerel İkili Desen (Extended Local Binary Pattern), Gri Seviye Birliktelik Matrisi (Grey Level Co-occurrence Matrix) ve istatistiksel yöntemler ile elde edilmiştir. Çıkarılan her bir özneliğin çıktıya olan katkısı aynı olmaması sebebiyle daha sonra bu öznitelikler ağırlıklandırılmıştır. Son aşamada ise ağırlıklandırılmış öznitelikler kullanılarak farklı sınıflandırıcılar eğitilerek mitoz hücre tespiti yapılmıştır [16].

Li vd. veri kümesinde sadece zayıf etiketlerin bulunması durumunda mitoz hücre tespiti yapabilecek bir derin öğrenme modeli önermiştir. Önerilen model mitoz hücre tespitini yaparken hücrelerin çevresel bilgilerini kullanmaktadır. Modelin eğitim aşamasında hücreleri çevreleyen etiketler yerine zayıf etiketleri kullanması sebebiyle veri işaretleme için harcanan eforun da azaltıldığı belirtilmiştir. Bunun yanı sıra önerilen modelin başarı oranını artırmak ve yanlış tespitleri azaltmak için bir doğrulama yöntemi önermişlerdir [17].

Chen ve vd. mitoz hücre tespiti için CNN tabanlı bir model önermiştir. Önerilen model girdi olarak verilen imgeyi evrişim katmanlarından geçirerek mitoz hücrelerine ait bilgileri çıkarır. Daha sonra elde edilen öznitelik matrisleri üst örnekleme (upsampling) kullanılarak imgenin orijinal boyutuna getirilir. Elde edilen çıktı herhangi bir bölgenin mitoz hücre bulundurması durumuna göre yüksek veya düşük puanlar içermektedir [18].

Mahmood vd. yanlış tespit edilen mitoz hücreleri en aza indirmek için iki aşamalı bir algoritma önermiştir. Birinci adımda Daha Hızlı Bölgesel Evrişim Sinir Ağları (Faster R-CNN) obje tespiti algoritması kullanılan ICPR-2012 ve ICPR-2014 eğitim veri kümeleri ile eğitilmiştir. İkinci adımda ise ResNet50 ve DenseNet201 CNN mimarileri aynı eğitim veri setleri ile eğitilmiştir. Kullanılan test veri kümesindeki olası mitoz hücreler Faster R-CNN modeli ile çıkarılmış, ResNet50 ve DenseNet201 mimarileri ile sınıflandırılmıştır [19].

Sohail vd. çok aşamalı bir mitoz hücre tespiti algoritması önermiştir. İlk aşamada zayıf etiketlere sahip olan mitoz hücrelerinin piksel seviyesindeki etiketleri Maske Bölgesel Evrişimsel Sinir Ağları (Mask R-CNN) yöntemi ile oluşturulur. Mitoz hücrelerini barındıran bölgelerin seçilmesi için yine aynı şekilde Mask R-CNN ve bunun yanı sıra çok amaçlı hata fonksiyonu kullanılır. Elde edilen sonuçlara bölge analizi (blob analysis) uygulanarak olası mitoz hücreleri tespit edilir. Tespit edilen hücreler makalede önerilen bir CNN mimarisi ile sınıflandırılır ve mitoz hücre tespiti yapılır. [20]

Sebai vd. mitoz hücre tespiti için denetimli, denetimsiz veya zayıf denetimli öğrenme yaklaşımlarının hepsinde kullanılabilecek bir algoritma önermiştir. Önerilen algoritmanın tabanında Mask R-CNN ve ResNet gibi CNN tabanlı derin öğrenme modelleri bulunmaktadır. Model ayrıca bölütleme ve nesne tespiti gibi iki farklı görevi de gerçekleştirebilmektedir [21].

Yücel vd. obje tespiti algoritması olan YOLO algoritması ile Dönüştürücü (Transformer) yapısının birleştirilmesiyle elde edilen bir model önermiştir. Dönüştürücü modelin YOLO algoritmasına eklenmesiyle beraber mitoz hücrelerinin genel özelliklerinin ve yerel özelliklerin arasındaki ilişkinin çıkarılması hedeflenmiştir [22].

Raj vd. histopatolojik görüntülerden mitoz hücre tespiti yapmak için CNN ve geleneksel öznitelik çıkarma algoritmalarını birleştiren bir teknik önermiştir. Histopatolojik imgelerde bulunan çekirdek gibi önem arz eden bölgelerin ön plana çıkarılması için Yüksek Güçlü Alan (HPF) renk uzayına sahip imgeler Mavi Oran renk uzayına dönüştürülmüştür. İmgelerde bulunan olası çekirdek bölgelerini çıkarmak için U-Net tabanlı olan StarDist bölütleme algoritması kullanılmıştır. Çekirdek adayları olan bölgelerden şekil ve doku gibi bilgileri içeren 144 farklı öznitelik geleneksel yöntemler ile çıkarılmıştır. Bunun yanı sıra çekirdek bölgelerinden önerilen CNN algoritması kullanılarak daha derin öznitelikler çıkarılmıştır. Elde edilen öznitelikler birleştirilerek topluluk öğrenmesi yöntemiyle eğitilmiş ve MITOS-ATYPIA-14 veri seti için %96 f-skoru elde edilmiştir [23].

Albayrak ve Bilgin histopatolojik görüntülerde mitozlu hücrelerin tespiti için bölütleme ve yönlü gradyanlar histogramı (Histogram of Gradient, HOG) yöntemini kullanmıştır. Daha iyi sonuçlar elde etmek için RGB renk uzayındaki histopatolojik görüntüler Lab renk uzayına dönüştürülmüştür. İmgeler üzerinde bulunan hücrelerin tespit edilebilmesi için bulanık C-ortalamları (Fuzzy C-means) algoritması kullanılmıştır. Ardından gürültü gibi istenmeyen bilgilerin temizlenmesi için bağlı bileşen analizi (connected component analysis) yöntemi kullanılmıştır. Son olarak damla analizi yapılarak olası hücreler belirlenmiş ve yönlü gradyanlar histogramı yöntemi kullanılarak bazı öznelilikler çıkarılmıştır. Bu öznelilikleri ise SVC ve RF gibi sınıflandırıcı algoritmaları ile eğitilmiş ve sonuçlar elde edilmiştir [24].

Albayrak ve Bilgin mitozlu hücrelerin bölütlenmesi için geleneksel yöntemlerden süperpiksel ve derin öğrenme yöntemlerinden CNN kullanarak hibrit bir yöntem önermiştir. Histopatolojik imgelere medyan filtresi uygulanarak yumuşatılmış ve daha sonra süperpiksel algoritması ile hücreler bölütlenmiştir. Daha sonra olası mitozlu hücreler ESA algoritması ile sınıflandırılmış ve sonuçlar elde edilmiştir [25].

Wang vd. geleneksel yöntemleri ve derin öğrenme yöntemlerini kullanan iki aşamalı bir mitozlu hücre tespiti algoritması önermiştir. Öncelikle histopatolojik görüntüler RGB renk uzayından Mavi Oran renk uzayına dönüştürülür. Daha sonra hücrelerin tespit edilmesi için Gauss Laplasyanı (Laplacian of Gaussian) yöntemi kullanılır. Tespit edilen hücrelere ait derin öğrenme öznelilikleri ve geleneksel yöntemler ile elde edilmiş öznelilikler çıkarılır ve iki ayrı sınıflandırıcı eğitilir. Test aşamasında ise her bir hücre bu iki ayrı sınıflandırıcı ile tahminlenir. Eğer her iki sınıflandırıcı da birbirinin aksi kararlar verir ise derin öğrenme ve geleneksel yöntemler ile çıkarılan öznelilikler üçüncü bir sınıflandırıcıya verilir ve yeni bir tahmin yapılır. Son olarak ise üç sınıflandırıcının tahminin ağırlıklı ortalaması alınır ve son tahmin olarak kabul edilir [26].

Saha vd. CNN ve geleneksel yöntemlerden elde edilen öznelilikleri birleştiren bir yöntem önermektedir. Görüntülere ait boyut ve doku gibi 55 farklı geleneksel öznelilikler çıkarılır. Ayrıca görüntüler önerilen 5 katmanlı CNN ağına girdi olarak verilir ve derin öznelilikler elde edilir. Geleneksel yöntemler ve derin öğrenme ile

elde edilen öznitelikler birleştirilir ve bir doğrusal sınıflandırıcı eğitilir. Önerilen yöntem MITOS-ATYPIA-14 veri seti için %90 f-skoru elde etmiştir [27].



3. MALZEME VE YÖNTEM

Histopatolojik görüntülerden hücrelerin veya mitoz hücrelerinin tespit edilmesi için birkaç farklı yaklaşım vardır. Bunlar temel olarak sınıflandırma (classification), bölütleme (segmentation) ve obje tespiti (object detection) yaklaşımlarıdır. Bu yöntemlerin hepsinin bir arada kullanıldığı çalışmalar olduğu gibi sadece bir tanesinin kullanıldığı çalışmalar da mevcuttur. Kullanılan her bir yaklaşıma göre izlenmesi gereken yollar ve algoritmalar farklılık göstermektedir. Bu çalışmada bahsedilen yaklaşımlardan bölütleme ve sınıflandırma yaklaşımları kullanılmıştır. Bilgisayarlı görü ile bölütleme ve sınıflandırma uygulamalarının gerçekleştirilmesinde probleme göre değişiklik göstermekle beraber bu çalışmada genel olarak Şekil 3.1’de gösterildiği gibi beş adımdan oluşan bir yol izlenmektedir. Bunlar veri toplama, bölütleme, veri ön işleme, öznetelik çıkarma ve son olarak sınıflandırma aşamalarıdır. Bu aşamalar bölüm 3’ün altında alt başlıklar halinde daha detaylı olarak incelenmiştir.



Şekil 3.1 Bilgisayarlı görü işlem adımları

3.1 Veri Toplama

Sınıflandırma veya bölütleme yöntemleri ile bir problemi çözmek için problemi ifade edecek verilere ihtiyaç duyulmaktadır. Problemi temsil edecek verilerin elde edilmesi için daha önceden oluşturulan veri kümelerinin kullanılması mümkün olduğu gibi sıfırdan bir veri kümesi oluşturmak da mümkündür. Sıfırdan bir veri kümesi oluşturmak maliyetli ve uzun bir süreç olduğu için genellikle eğer varsa hazır veri kümeleri kullanılmaktadır.

Kullanılması planlanan veri verilerin çözülmek istenen problemi en iyi şekilde ifade etmesi ve sayıca yeterli olması gerekmektedir. Veri kümesindeki veriler tekrarlı olmamalı, birbirine çok benzememeli ve doğrulanmış olmalıdırlar. Veri kümesini problemi temsil edememesi durumunda aşırı uyum (overfitting) veya yetersiz uyum

(underfitting) gibi sorunlar ortaya çıkmaktadır [28], [29]. Uygun olmayan veri kümeleri ile eğitilen modeller yanlış, güven skoru düşük ve ön yargılı sonuçlar üretmektedir.

Kullanılan veri kümesindeki örnek sayısının belirlenmesi için nicel bir yöntem bulunmamakla beraber karmaşık problemler için basit problemlere göre çok daha fazla veriye ihtiyaç duyulmaktadır. Yeterli sayıda örnek bulunmayan veri kümelerini veri artırma (data augmentation) yöntemleri kullanılarak zenginleştirmek ve örnek sayısını artırmak mümkündür. Her ne kadar veri artırma teknikleri ile üretilen yeni veriler orijinal verilere benzese dahi yapılan çalışmalar bu tür yaklaşımların modelin performansını olumlu etkilediğini göstermiştir [30].

Veri kümesinde bulunan örnek sayısı önemli olduğu gibi sınıflar arasındaki örnek sayısının dengesi de önemlidir. İki sınıflı bir veri kümesinde bir sınıfın örnek sayısının diğer sınıfın örnek sayısına karşı ezici bir üstünlük sağladığı takdirde dengesiz veri kümesi problemi ortaya çıkmaktadır. Örneğin veri kümesinde 1000 tane mitozlu hücre bulunurken 10000 tane mitozlu olmayan hücrenin bulunması veri kümesini dengesiz olduğu anlamına gelmektedir. Dengesiz veri kümeleri modelin performansını düşürmekte ve az örnek sayısına sahip sınıfın öğrenilmesi engellemektedir. Dengesiz veri kümesi problemini çözmek için literatürde alt örnekleme (downsampling), üst örnekleme (upsampling), yapay veri üretme (synthetic data generation) ya da bunların birleşimi gibi çözümler kullanılmaktadır.

Bu çalışmada önerilen yaklaşımların değerlendirilebilmesi için MIDOG2021, MIDOG2022 ve ICPR2014 veri kümeleri kullanılmıştır. Kullanılan veri kümelerinin detaylı açıklamaları alt başlıklar altında verilmiştir.

Mitozlu hücre tespiti insan veya hayvanların farklı organlarında yapılabilmektedir. İnsan veya hayvanlarda bulunan hücrelerinin karakteristik özellikleri farklı olmakla beraber aynı türün farklı organlarındaki hücrelerinin karakteristik özellikleri de değişiklik gösterebilmektedir. Aynı hücrelerin farklı tarayıcılar kullanılarak dijitalleştirilmesi de nihai olarak elde edilen histopatolojik görüntülerdeki hücreler de birbirinden farklı karakteristikler sergilemektedir. Veri kümesinde bulunan

hücrelerin karakteristik özelliklerinin değişiklik göstermesi geliştirilen kanser tespiti algoritmalarının istenen performansını sergilemesinin önüne geçmektedir [31]. Bu durum literatürde alan kayması (domain shifting) olarak da bilinmektedir. Bu çalışmada kullanılan MIDOG2021, MIDOG2022 ve ICPR2014 veri kümeleri geliştirilen mitozlu hücre tespiti algoritmalarının alan kayması problemine karşı gürbüz olabilmesi için verilerin toplanması aşamasında farklı tarayıcılar kullanılmıştır. Bunun yanı sıra toplanan mitozlu hücreler farklı kanser türlerinden elde edilmiştir.

3.1.1 MIDOG2021

MIDOG2021 (Mitosis Domain Generalization Challenge 2021) veri seti 2021 yılında düzenlenen bir yarışmada [32] kullanılmak için yayınlanan ve içerisinde insana ait meme kanseri hücrelerini barındıran bir histopatolojik veri setidir [33]. Veri setinin eğitim kümesi herkese açık bir şekilde yayınlanırken test kümesinin yayınlanmaması sebebiyle bu çalışmada sadece eğitim veri kümesi kullanılmıştır. Veri seti 1721 mitozlu ve 2714 mitozlu olmayan hücre görüntüsüne sahiptir. Test veri kümesini oluşturmak için toplam veri kümesinin %20'lik kısmı kullanılmıştır. Veri kümesinde bulunan tüm slayt hücreler, 4 farklı Tüm Slayt Görüntü (Whole Slide Image, WSI) tarayıcıları (Hamamatsu XR nanoscooper 2.0, Hamamatsu S360 (0.5 NA), Aperio ScanScope CS2, Leica GT450) ile elde edilmiştir. Tarayıcıların dijitalleştirme yöntemlerinin sonuca olan etkisini de veri kümesine yansıtmak için meme kanserinin 50 farklı tanısına ait görüntüler farklı tarayıcılar kullanılarak elde edilmiştir. MIDOG2021 veri setine ait birkaç örnek görüntü Şekil 3.2'deki gibidir.

3.1.2 MIDOG2022

MIDOG2022 (Mitosis Domain Generalization Challenge 2022) veri kümesi farklı laboratuvarlardan elde edilmiş 370 farklı tümör tanısına ait histopatolojik görüntüleri içermektedir [34]. Toplanan veriler ilk olarak MIDOG2022 yarışmasında kullanılmıştır [35]. Veri kümesinin gerçek dünya problemlerini daha iyi temsil edebilmesi için elde edilen görüntüler birkaç farklı türe (insan, kedi köpek) ve tarayıcıya aittir. MIDOG2022 eğitim veri kümesi her bir tümör tipi için prostat, lenfoma melanoma ve meme kanseri gibi en az 50 farklı tanıya sahiptir. Test veri kümesinin herkese açık bir şekilde paylaşılmaması sebebiyle paylaşılan veri

kümesinin %20'lik kısmı bu çalışmada test için ayrılmıştır. Toplam veri kümesinde 9501 mitozlu ve 11051 tane mitozlu olmayan hücre görüntüsü vardır. MIDOG2022 veri setine ait birkaç örnek görüntü Şekil 3.3'deki gibidir.

3.1.3 ICPR2014

ICPR2014 (International Conference on Pattern Recognition) veri kümesi 2014 yılında İsveç'in Stockholm kentinde düzenlenen bir konferans esnasında yayınlamıştır [36]. Veri kümesinin yayınlanmasındaki temel amaç meme kanserinin tespit edilmesinde sıkça başvurulmuş ve büyük bir öneme sahip olan mitozlu hücrelerin sayısının belirlenmesidir. Paylaşılan veri kümesi içerisinde toplam 749 tane mitozlu hücre bulunurken yaklaşık olarak 180 bin tane de mitozlu olmayan hücre bulunmaktadır. Paylaşılan görüntüler 2 farklı tarayıcı kullanılmış olup bunlar Aperio Scanscope XT ve Hamamatsu Nanozoomer 2.0-HT tarayıcılarıdır. ICPR2014 veri kümesine ait bazı örnek görüntüler Şekil 3.4'teki gibidir.

3.2 Bölütleme

Kullanılan MIDOG2021 ve MIDOG2022 veri kümeleri sınıflandırmaya uygun olarak kaynağından temin edilmiştir. Fakat ICPR2014 veri kümesindeki mitozlu hücrelerin sınıflandırılabilmesi için öncelikle hücrelerin tespit işleminin gerçekleştirilmesi gerekmektedir. ICPR2014 veri kümesindeki hücrelerin tespit edilebilmesi için 3 işlem gerçekleştirilmiştir. Bunlar görüntülere ortanca (medyan) filtresinin uygulanması, k-ortalamlar yöntemi ile bölütlenmesi ve bölge analizi ile hücrelerin seçilmesi aşamalarıdır.

3.2.1 Ortanca filtresi

Ortanca filtresi, $n \times n$ 'lik bir pencerede bulunan piksel değerlerini sıralar. Sıralanan piksel değerlerinden $\frac{n \cdot n}{2}$ konumunda olan piksel değerini seçer ve bulunduğu pikselin konumuna bu yeni değeri atamasını yapar.

3.2.2 K-ortalamalar

K-ortalamalar algoritması girdi olarak verilen bütün görüntüyü n adet renk kümesi ile temsil etmeyi amaçlar. Girdi olarak verilen imgede rastgele n tane piksel değeri seçer. Bunlar küme merkezleri olarak da adlandırılmaktadır. Görüntü üzerinde bulunan diğer piksel değerleri ile bu küme merkezleri arasında bir mesafe hesabı yapılır ve piksel en yakın olduğu kümeye bırakılır. Bu işlem bütün piksel değerleri için yapıldıktan sonra kümelerde bulunan piksel değerlerinin ortalaması hesaplanır ve küme merkezlerinin değerleri olarak atanır. Bu işlem belirli bir iterasyon kadar devam eder ve iterasyonların bitmesinin ardından görüntü n tane farklı renk kümesinden oluşmaktadır.

3.2.3 Bölge analizi (Blob analysis)

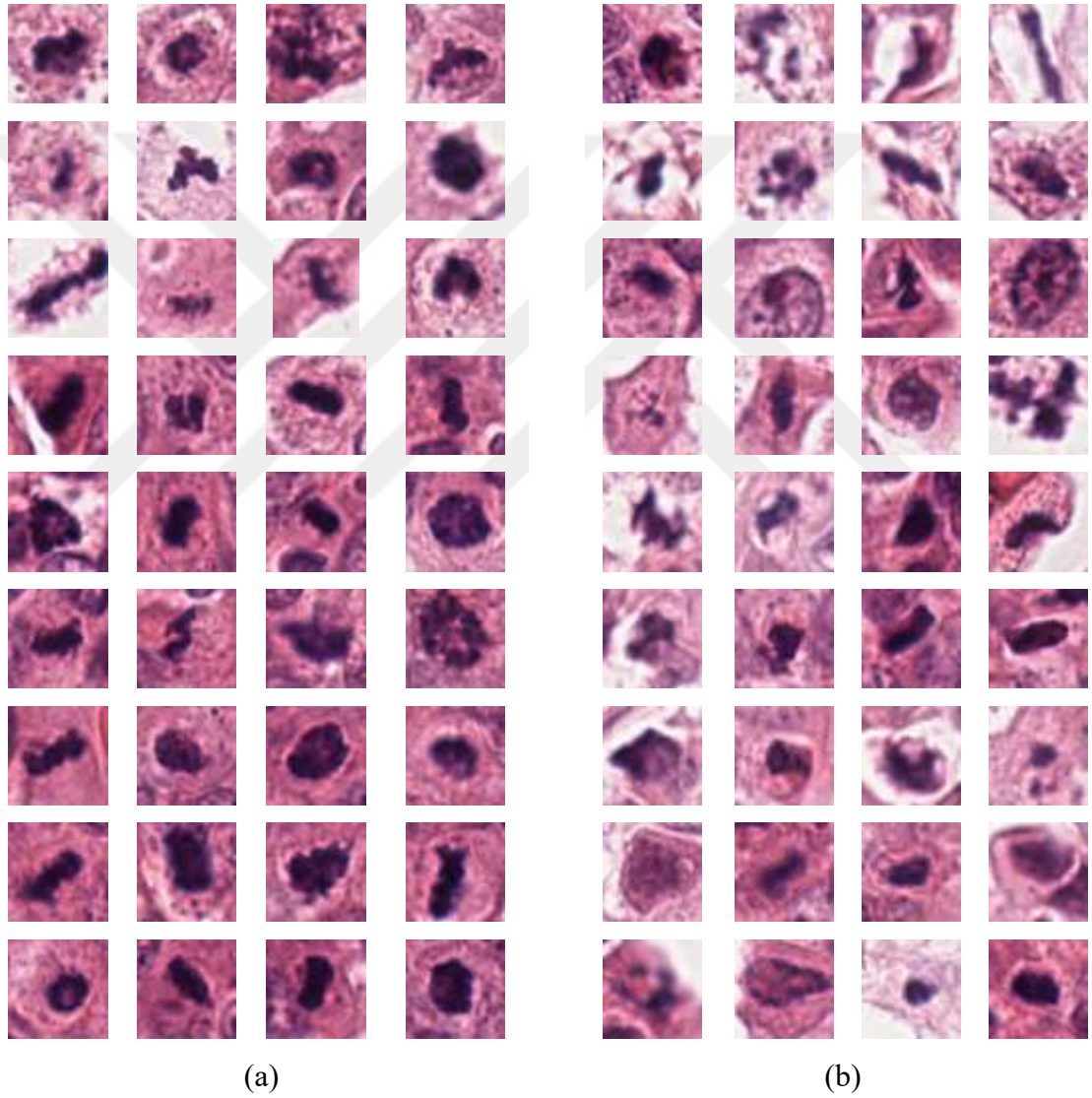
Bölge analizi aşamasında bölütlenmiş görüntüdeki bulunan bütün hücreler tespit edilmiştir. Hücre olması çok mümkün olmayan çok küçük veya büyük bölgeler bölge analizi yapılarak elimine edilmiştir. Kalan hücreler ise asıl görüntüden kırpılarak ayrı ayrı dosyalarda kayıt edilmiştir.

3.3 Veri Ön İşleme

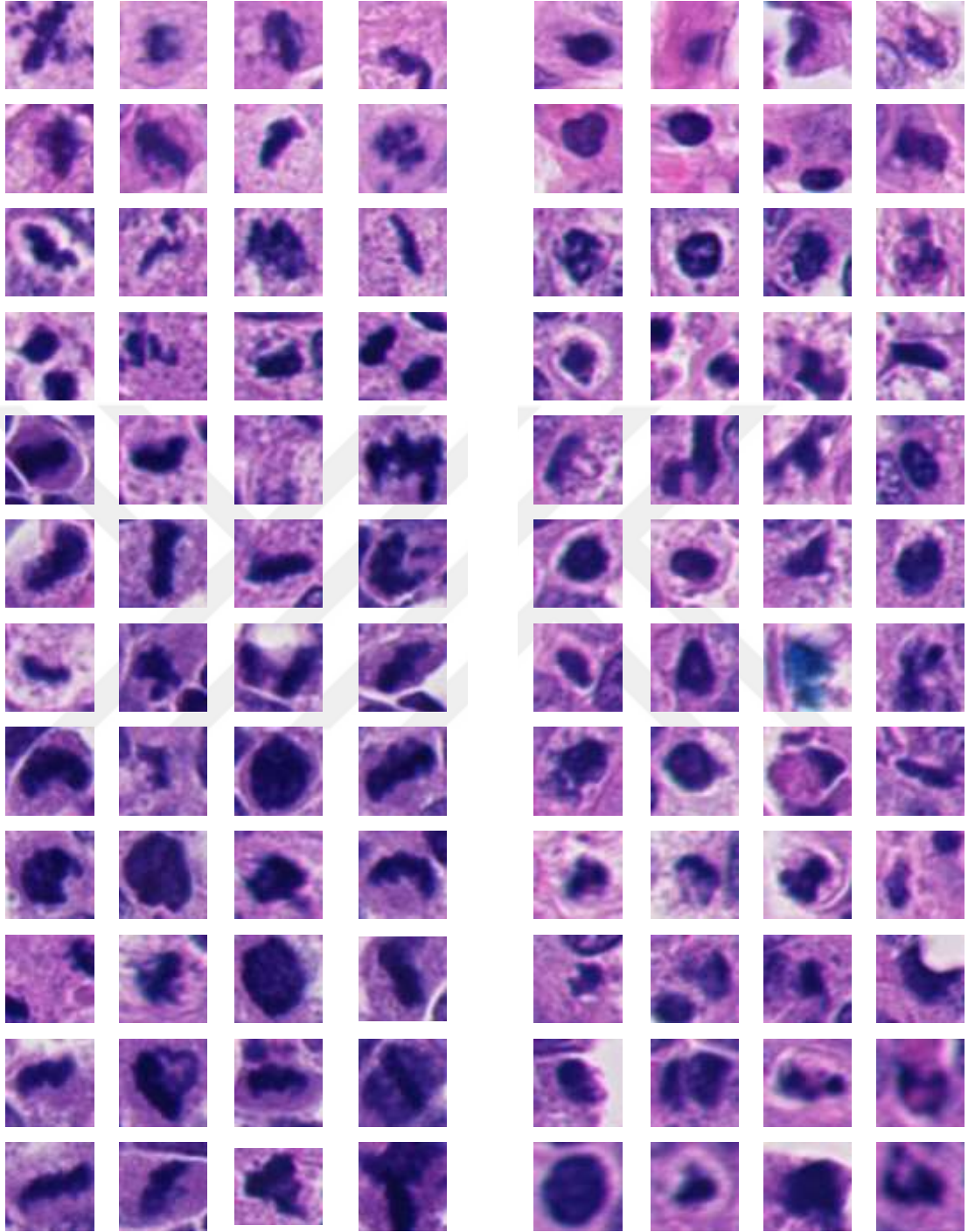
Problemi ifade etmek için oluşturulan veri kümesi temsiliyeti yetersiz özelliklere sahip olabilmektedir. Bunun yanı sıra ham veriler çoğunlukla derin öğrenme modeline girdi olarak verilebilecek formatta değildir. Bu tür problemlerin giderilmesi veya veri kümesindeki örneklerin zenginleştirilmesi için yapılan işlemlere veri ön işleme ismi verilmektedir. Veri ön işleme adımları derin öğrenme modelinin eğitim aşamasından önce gerçekleştirilmektedir. Örneğin veri kümesi içerisinde bulunan örneklerin karanlık ortamda alınmış olması ya da gürültü içermesi olası bir durumdur. Bu tür istenmeyen etkilerin azaltılması için veri kümesinde bulunan imgelere kontrast artırma veya gürültü giderme gibi ön işlemler uygulanmaktadır.

Literatürde kullanılan 10'dan fazla ön işleme adımı bulunmaktadır. Normalizasyon, kontrast artırma, yeniden boyutlandırma veya imge döndürme literatürde sıkça kullanılan ön işlem adımlarından birkaçıdır. Bunların her biri tek tek kullanılabileceği gibi birleştirilmesi de mümkündür. Kullanılacak ön işlem

yöntemlerinin seçilmesinde çözülmek istenen problem ve kullanılacak olan derin öğrenme algoritmaları önemli bir rol oynamaktadır. Bu çalışmada ön işlem aşamasında yeniden boyutlandırma ve normalizasyon olmak üzere toplam iki ön işleme algoritması kullanılmıştır. Bu aşamada bahsedilen bütün işlemler sadece derin öğrenme tabanlı algoritmalar için uygulanmış olup geleneksel görüntü işleme yöntemlerine uygulanmamıştır. Kullanılan ön işleme algoritmaları daha detaylı olarak alt başlıklar altında anlatılmıştır.



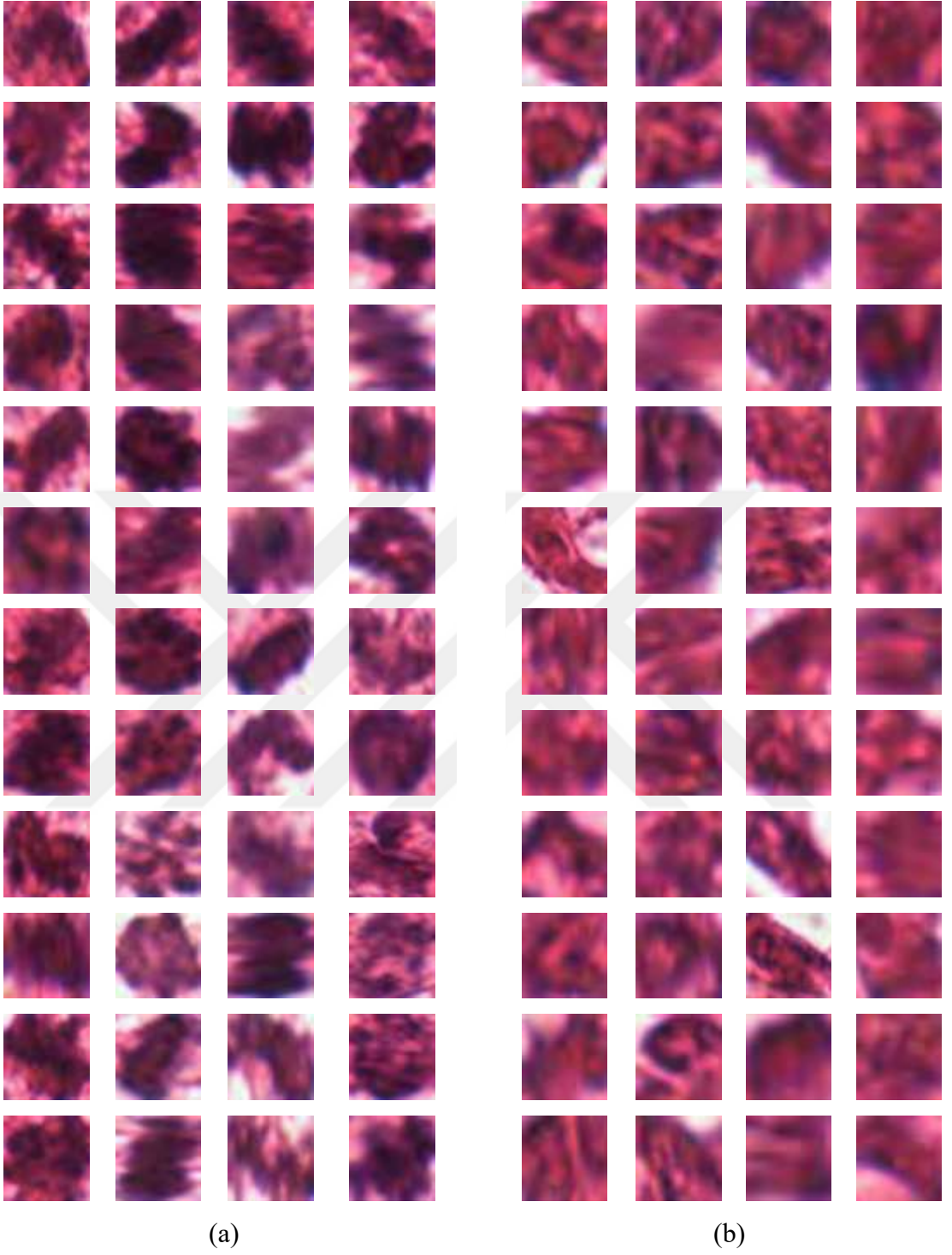
Şekil 3.2 MIDOG2021 veri setinden örnek görüntüler: (a) mitozlu hücreleri (b) ise mitozlu olmayan hücreleri temsil etmektedir



(a)

(b)

Şekil 3.3 MIDOG2022 veri setinden örnek görüntüler: (a) mitozlu hücreleri (b) ise mitozlu olmayan hücreleri temsil etmektedir



Şekil 3.4 ICPR2014 veri setinden örnek görüntüler: (a) mitozlu hücreleri (b) ise mitozlu olmayan hücreleri temsil etmektedir

3.3.1 Yeniden boyutlandırma

Gerek CNN gerek Dönüştürücü tabanlı birçok algoritma sabit boyutlu veriler ile işlem yapmaktadır. Bunun temel sebebi girdi olarak verilen her bir görüntü ve görüntü pikseli için bir parametre ataması yapmasıdır. Literatür incelendiğinde bilgisayarlı görü alanında kullanılan derin öğrenme yöntemlerinin çoğunlukla 224x224 veya 384x384 boyutundaki görüntüleri girdi olarak kullandığı görülmektedir. Bu çalışmada, kullanılan veri kümelerindeki histopatolojik görüntüler 224x224 olacak şekilde yeniden boyutlandırılmıştır.

3.3.2 Normalizasyon

Gündelik hayatımızda karşılaştığımız birçok görüntü her kanal için 8 bit büyüklüğünde bir veriye sahiptir. Görüntünün her bir kanalının 8 bitlik bir derinliğe sahip olması, her bir piksel değerinin 0-255 arasında değerler alabileceği anlamına gelir. Bu değerlerin sıfıra yakın olması ortamın karanlık veya cismin renginin siyaha yakın olacağını, 255 değerine yakın olması ise ortamın aydınlık veya cismin renginin beyaza daha yakın olacağını ifade etmektedir. Görüntüdeki aynı nesnelerin parlaklık gibi farklı ortam koşullarına göre değerinin değişmesi modelin performansını doğrudan etkileyebilmekte ve düşürebilmektedir. Aynı nesne görüntülerinin farklı değerlerle ifade edilmesi görüntüde bulunan pikseller arasındaki oransal bilginin kayıp olmasına neden olmaktadır. Özellikle dağılımsal bilgiden çıkarım yapan derin öğrenme modelleri bu durumdan çok fazla etkilenmektedir. Veri kümesinde bulunan görüntülerine normalizasyon tekniklerinin uygulanması oransal bilginin saklanması ve kullanılan modelin ortam koşullarından çok daha az etkilenmesini sağlamaktadır.

Derin öğrenme algoritmalarının yüksek işlem gücü ile öğrenmeleri eğitim aşamasının uzun sürmesine ve eğitim maliyetinin artmasına sebep olmaktadır. Veri kümesindeki benzer örneklerin birbirinden uzak olması da modelin optimizasyon sürecini güçleştirmekte ve eğitim maliyetini artırmaktadır. Normalizasyon teknikleri ile veri kümesinde bulunan örnekler birbirine yaklaştırılmakta ve modelin en az hata yapması için gerçekleştirilen optimizasyonu işlemleri daha kısa sürmekte ve böylece modelin eğitilmesi için gereken maliyet azalmaktadır [37].

Derin öğrenme modellerinin eğitilmesi esnasında farklı problemler meydana gelmektedir. Bu tür problemler modelin tamamen öğrenmesini engellemekte veya performansını düşürmektedir. Sıkça karşılaşılan problemlerden biri ise gradyan patlaması (gradient exploding) veya gradyan sönümlenmesidir (gradient vanishing). Verilere derin öğrenme modellerinden önce normalizasyon tekniklerinin uygulanması ile kullanılan modelin gradyan patlaması ve sönümlenmesi problemleri ile daha az karşı karşıya gelmesini sağlamak mümkündür.

Literatürde kullanılan farklı normalizasyon teknikleri bulunmaktadır. Bu çalışmada standart normalizasyon tekniği kullanılmıştır ve hesaplanma şekli Denklem 3.1’de verilmiştir.

$$I_n = \frac{I - m}{\sigma} \quad (3.1)$$

Denklem 3.1’de bulunan “ I ”, “ m ” ve “ σ ” ifadeleri sırasıyla imgeyi, imgenin piksel yoğunluklarının ortalamasını ve piksel yoğunluklarının standart sapmasını ifade etmektedir. “ I_n ” ise normalizasyon işlemi sonucu elde edilmiş görüntüyü ifade etmektedir.

3.4 Öznitelik Çıkarma

Öznitelik çıkarma aşaması bilgisayarlı görü uygulamalarında oldukça kritik bir öneme sahiptir. Kameralar, mikroskoplar veya görüntü elde etmek için kullanılan diğer cihazlar yüksek çözünürlükte veriler toplamaktadır. Yüksek çözünürlüklü verilerin işlenmesi oldukça fazla zamana ve işlem gücüne ihtiyaç duymaktadır. Düşük çözünürlüklü elde edilen görüntüler ise yeteri kadar bilgi içermemekte ve geliştirilen algoritmanın istenen performansı vermesini engellemektedir. İmgelerin yüksek veri hacmini en aza indirmek ve bu indirgemeyi yaparken imgeden kıymetli bilgileri elde etmek için öznitelik çıkarma işlemleri uygulanır. Yüksek veri hacminin azalmasıyla beraber geliştirilen algoritmanın çok daha az kaynak ile çok daha hızlı çalışması mümkündür. Literatürde bulunan öznitelik çıkarma algoritmalarının 2 başlık altında toplanması mümkündür. Bunlar derin öğrenme yaklaşımları ve geleneksel görüntü işleme algoritmalarıdır.

Geleneksel görüntü işleme algoritmalarına Gradyanların Histogramı (Histogram of Oriented Gradients, HOG), Yerel İkilik Desen (Local Binary Pattern, LBP) ve Hızlandırılmış Gürbüz Öznitelikler (Speedup Robust Features, SURF) örnek olarak vermek mümkündür. CNN, Vision Transformer (ViT) veya MLP-Mixer ise öznitelik çıkarmak için literatürde sıkça kullanılan derin öğrenme yaklaşımlarıdır. Bu çalışmada veri kümelerinde bulunan görüntülerden öznitelik çıkarmak için derin öğrenme tabanlı DINO özdenetimli öğrenme yaklaşımı ve temel model olma özelliği taşıyan dönüştürücü tabanlı ViT, XCiT ve CNN tabanlı ResNet mimarileri kullanılmıştır. Geleneksel görüntü işleme yöntemi olarak ise LBP algoritması kullanılmıştır.

3.4.1 Temel modeller

Farklı veya aynı çok fazla veri ile eğitilen ve daha sonra farklı görevler için kullanılan modellere temel model ismi verilmektedir [38]. Temel modeller kavramı daha çok doğal dil işleme alanında kullanılırken bilgisayarlı görü veya diğer alanlarda da yaygınlaşmaya başlamıştır. Temel modeller çok büyük veri setleri ile eğitilmesi sebebiyle yüksek kaynak ve paralel işlem gücüne ihtiyaç duymaktadır. Özellikle Dönüştürücü modellerinin gelişimi ile beraber modeller daha çok paralel çalışma kapasitesine erişmiştir. Derin öğrenme modellerinin denetimli, denetimsiz, yarı denetimli veya özdenetimli gibi birkaç farklı öğrenme yaklaşımları ile eğitilmesi mümkündür. Temel modeller genellikle özdenetimli öğrenme yaklaşımları ile eğitilirler. Çok büyük veri kümeleri ile eğitilen modellerin doğrudan kullanılması mümkün olduğu gibi daha küçük veri kümeleri ile bilgi aktarımı veya ince ayar ile eğitildikten sonra da kullanılması mümkündür.

Derin öğrenme modellerinin eğitilmesi gerek veri gerek kaynak açısından yüksek bütçelere ihtiyaç duymaktadır. Özellikle küçük ölçekteki yapılanmaların derin öğrenme modellerinin eğitilmesi için bir kaynak ayırması çok mümkün olmamaktadır. Temel modellerin ortaya çıkışı eğitim aşamasının maliyetini azaltması sebebiyle daha küçük ölçekteki organizasyonların yapay zekâ alanında çalışabilmesi için büyük bir fırsat niteliğindedir.

3.4.2 Özdenetimli öğrenme

Özdenetimli öğrenme yaklaşımı denetimsiz öğrenme yaklaşımında olduğu gibi verideki bilginin çıkarılabilmesi için herhangi bir dış müdahaleye ihtiyaç duymamaktadır. Bunun yerine verinin kendisini veya kısmen değiştirilmiş halini kullanmaktadır. Verinin kendisini ve biraz değiştirilmiş halini kullanarak veride bulunan öznitelikleri çıkarmaya çalışan öğrenme algoritmaları karşılaştırmalı (contrastive learning) olarak da bilinmektedir. Öz denetimli öğrenme yaklaşımı etiket gibi dış yönlendirmeye ihtiyaç duymaması sebebiyle veri etiketleme veya hazırlama için gereken eforu düşürmekte ve çok daha fazla veri ile eğitim aşamasının tamamlanmasına katkı sunmaktadır. Fakat öğrenme işleminde fazla veri kullanılması sebebiyle çok daha yüksek işlem gücüne ihtiyaç duyulmaktadır.

Noroozi vd. Jigsaw Puzzle problemini çözmek için öz denetimli bir öğrenme modeli geliştirmiştir. Model, bir imgeyi 3x3'lük eşit parçalara böler ve bu parçaların yerlerini karıştırarak doğru yerleri bulmaya çalışır [39]. Özellikle modellerin denetimli öğrenme yaklaşımı ile küçük veri setlerini öğrenmesi noktasında öz denetimli öğrenme yöntemleri modelin performansın artmasına ve eğitim süresinin azalmasına katkı sunmaktadır [40].

Dengesiz veri setleri derin öğrenme modellerinin öğrenme aşamasında büyük bir problemdir. Eğitim veri setinde bir sınıftan az veri olması modelin onu aykırı değer olarak görmesine sebep olabilmektedir. Bunu yanı sıra yetersiz sayıdaki veriler model tarafından öğrenilmek yerine ezberlenebilmektedir. Liu vd. yaptıkları çalışmada dengesiz veri kümeleri için özdenetimli öğrenme yaklaşımı ile ön eğitimi yapılmış modeller denetimli öğrenme ile eğitilmiş modellere göre daha iyi ve gürbüz sonuçlar verdikleri görülmüştür. Bunun yanı sıra Liu vd. özdenetimli öğrenme modelinin dengesiz veri kümeleri için çok daha zengin öznitelikler çıkardığını göstermiştir [41].

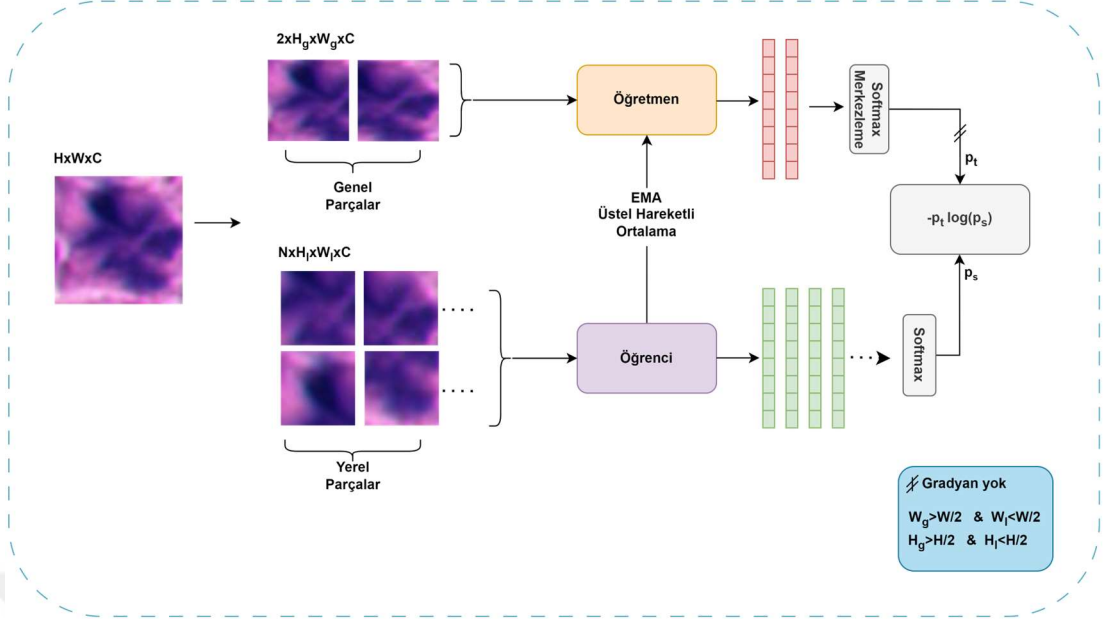
3.4.3 Self-distillation with no labels (DINO)

İnsan beyni gözün gördüğü objenin çevre şartlarından bağımsız olarak bir nesneyi kolayca tanıyabilmektedir. İdeal bir öznitelik çıkama algoritmasından da beklenen

performans bu şekildedir. Algoritma alınan görüntüdeki nesnenin hangi ışık koşulları altında veya hangi perspektif bakış açısından alındığından bağımsız aynı veya benzer öznitelik vektörleri çıkarabilmelidir. DINO algoritması ise bu prensibi benimseyen bir öğrenme yaklaşımı olup farklı derin öğrenme modelleri ile kullanılması mümkündür. DINO algoritması aynı görüntünün farklı koşullar altında çekilmesi durumunda bile aynı öznitelik vektörünü elde etmeye çalışır. DINO öğrenme modeli aynı zamanda bir karşılaştırmalı öğrenme yöntemidir. DINO öğrenme algoritmasının genel mimarisi Şekil 3.5’deki gibidir.

DINO girdi olarak verilen bir imgeyi 2 tane genel n tane yerel olmak üzere parçalara böler. Genel parçalar imgenin boyutunun %50’sinden daha büyük iken yerel parçalar daha küçüktür. Bu parçaların her biri aynı görüntünün farklı şartlar altında alınması durumunu simgeler. Gerçek dünya senaryolarının elde edilebilmesi için yerel ve genel parçalara, kontrast artırma ve görüntü yumuşatma gibi bazı ön işlem adımları uygulanır.

Elde edilen yeni yerel ve genel parçalar öğrenci ve öğretmen olmak üzere farklı ağırlıklara fakat aynı mimariye sahip iki modele girdi olarak verilir ve bu parçalara ait gömme değerleri elde edilir. Eğitilen modelin daha iyi bilgi damıtması (knowledge distillation) yapabilmesi için çıkarılan öznitelik vektörlerine sıcaklık sabiti ile öğretmen modeli için Eşitlik 3.2’deki, öğrenci modeli için ise Eşitlik 3.3’deki gibi softmax fonksiyonu uygulanır [42], [43]. Bilgi damıtması yaklaşımı modele çıkarılan özniteliklerin yetersiz olduğu izlenimini verir ve onu temsiliyeti daha yüksek öznitelikler çıkarması için teşvik etmeyi amaçlamaktadır. Eşitlik 3.2 ve 3.3’de bulunan t , τ_t , τ_s ve C sabitleri sırasıyla modelin çıktısını, öğretmen ısınma katsayısını, öğrenci ısınma katsayısını ve merkezleme değerini ifade etmektedir. Öğretmen modeli için kullanılan ısınma katsayısı genel olarak öğrenci ısınma katsayısından daha küçük seçilir.



Şekil 3.5 DINO algoritmasının genel mimarisi. DINO algoritması imgeden farklı perspektifleri temsil eden genel ve yerel parçalar çıkarır. Gerçek hayat koşullarını taklit edebilmek için veri artırma yöntemlerini içeren iki farklı işlem bloğu çıkarılan genel ve yerel parçalara uygulanır. Parçalar öğrenci ve öğretmen modellerine verilir. Modellerin çıktısı olarak verdiği öznitelik vektörleri ısınma katsayısı ile softmax fonksiyonundan geçirilir ve her iki modelin de ağırlıkları güncellenir

$$t_o = softmax\left(\frac{t-c}{\tau_t}\right) \quad (3.2)$$

$$s_o = softmax\left(\frac{t}{\tau_s}\right) \quad (3.3)$$

Eşitlik 3.2’de bulunan c ve merkezleme değeri olan C , Eşitlik 3.4’teki gibi hesaplanmaktadır. Eşitlik 3.4’te bulunan m momentum oranını, d gömme değeri boyutunu, t_1^g ve t_2^g ise x_1^g ve x_2^g genel parçalarına ait öznitelikleri ifade etmektedir. Özdenetimli öğrenme modelleri boyutsal çöküş (dimensional collapse) ve tamamen çöküş (complete collapse) gibi bazı problemler ile karşılaşabilmektedir [44], [45]. Tamamen çöküş, modelin verilerden hiçbir şey öğrenememesine ve bütün girdiler için hemen hemen aynı çıktıları üretmesine sebep olmaktadır. Boyutsal çöküş ise modelin n boyutlu kapasitesinin kullanılamamasına ve eğitilen verileri çok daha az bir boyut ile temsil etmeye çalışması problemidir. Caron vd. bu tür problemleri aşmak için çıkarılan gömme değerlerine merkezleme değerini kullanmışlardır.

Modele girdi olarak verilen imgeler her zaman kendi içerisinde bir ilişkiye sahip olması mümkün değildir. Yerel ve genel parçalar arasındaki ilişkinin az olması

sebebiyle modelin çıktı olarak vereceği gömme değerleri birbirinden farklı olması mümkün olabilmektedir. Bu tür girdiler optimizasyon yönteminin modeli yanlış yönlendirmesine veya yerel minimum noktalarda sıkışıp kalmasına sebep olabilmektedir. Eşitlik 3.4'te bulunan m momentum katsayısı, C merkezleme değerinin bir önceki değerlerini ağırlıklandırarak kullanır ve ani olarak ortaya çıkabilecek bu tür problemleri en aza indirmeyi amaçlar. Sutskever vd. yaptıkları çalışmada momentum katsayısının derin öğrenmede ne kadar önemli olduğunu göstermişlerdir [46].

$$C_{1xd} = m \cdot C_{1xd} + (1 - m) \cdot \frac{\text{konkat}(t_1^g, t_2^g)_{1xd}}{2} \quad (3.4)$$

Model çıktılarının elde edilmesinin ve softmax fonksiyonundan geçirilmesinin ardından model ağırlıklarının güncellenebilmesi için çapraz entropi hatası (Cross Entropy Loss) Eşitlik 3.5'teki gibi hesaplanır. Buradaki P_t ve P_s , V , x_1^g ve x_2^g , x ve x' ifadeleri sırasıyla öğretmen ve öğrenci modellerinin verilen imge parçası için çıktısını, genel ve yerel imge parçalarını bulunduran seti, bütün genel imge parçalarını, her bir genel ve yerel imge parçalarını ifade etmektedir. Çapraz entropi hata fonksiyonu ise Eşitlik 3.6 'daki gibi tanımlanmıştır.

$$hata = \sum_{x \in [x_1^g, x_2^g]} \sum_{\substack{x' \in V \\ x' \neq x}} H(P_t(x), P_s(x')) \quad (3.5)$$

$$H(t, s) = -t_o \cdot \log(s_o) \quad (3.6)$$

Hatanın hesaplanmasının ardından öğretmen ve öğrenci modellerinin ağırlıkları güncellenir. Öğrenci modelinin ağırlıkları geriye yayılım fonksiyonu ile hesaplanırken öğretmen modelinin ağırlıkları, Eşitlik 7'deki gibi öğrenci modelinin bir önceki ağırlıklarının üstel hareketli ortalaması ile güncellenir [47]. Buradaki W_t öğretmen modelinin ağırlıklarını, l momentum oranını ve W_s öğrenci modelinin ağırlıklarını ifade etmektedir.

$$W_t = l \cdot W_t + (l - 1) \cdot W_s \quad (3.7)$$

Eğitimin tamamlanmasının ardından öğrenci ve öğretmen modellerinin her ikisi de öznitelik çıkarmak için kullanılabilir. Caron vd. DINO öğretmen modelinin

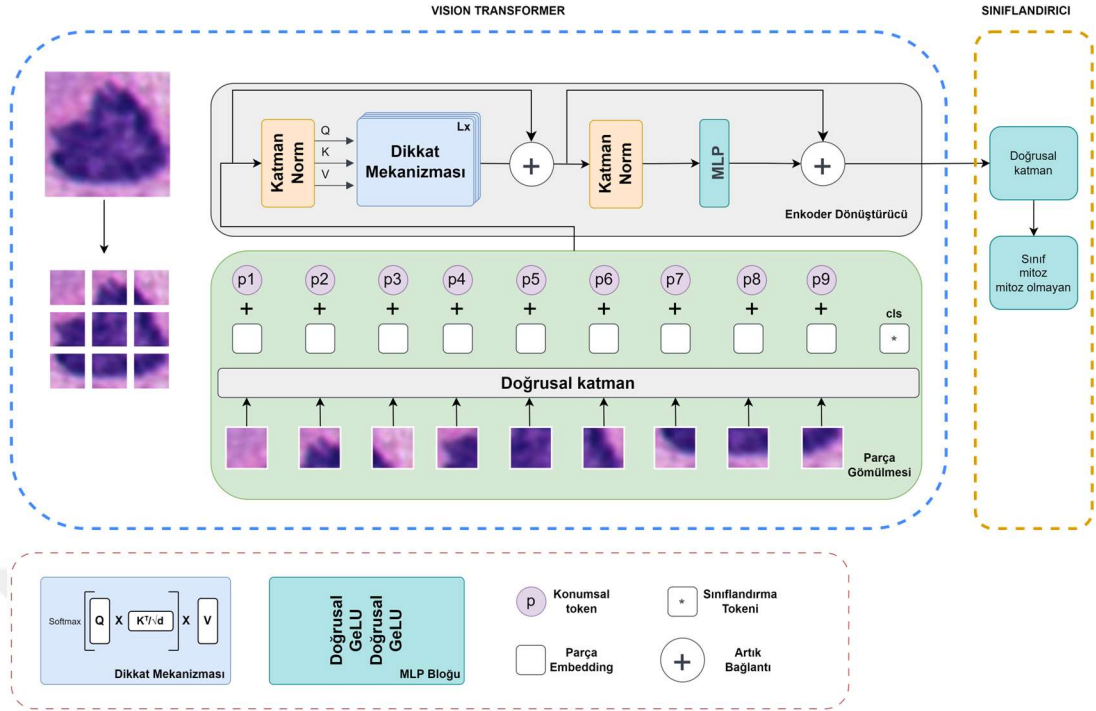
performansı, öğrenci modelinin performansından daha iyi olduğunu bildirmiştir [11]. Bu yüzden çalışmanın devamında öğretmen modeli kullanılmıştır.

3.4.4 Vision transformer

Vision Transformer (ViT), ilk olarak Doğal Dil İşleme (NLP) problemleri için Vaswani vd. tarafından önerilen Dönüştürücü (transformer) tabanlı bir öznitelik çıkarma ve görüntü sınıflandırma algoritmasıdır [48]. ViT, temelde bir sınıflandırma algoritması olmasına rağmen farklı formdaki halleri bölütleme [49], obje tespiti [50] ve imge restorasyonu [51] gibi diğer birçok alanda da kullanılmaktadır. ViT sınıflandırma algoritması üç ana aşamadan meydana gelmektedir: parça gömülmesi (patch embedding), enkoder dönüştürücü (transformer encoder) ve son olarak sınıflandırma aşamalarıdır.

ViT, CNN tabanlı derin öğrenme yaklaşımlarından farklı olarak imgeleri bütün şeklinde değil parça parça işlemektedir. Girdi olarak verilen imge $n \times n$ 'lik eşit parçalara bölünür. Buradaki n sayısı parça boyutu olarak adlandırılmaktadır. Parçalara ayrılan imgelerden kıymetli özniteliklerin çıkarılması ve her parçanın boyutsal olarak indirgenmesi için bir doğrusal katman uygulanır. Bunun yanı sıra her bir parçanın uzamsal bilgisinin kayıp edilmemesi için doğrusal katman çıktısına konumsal bilgiyi temsil edebilecek eğitilebilir parametreler eklenir. Bütün bu işlemlerin tamamı parça gömülmesi aşamasında meydana gelmektedir.

Parça gömülmesi aşamasında çıkarılan gömme değerleri enkoder dönüştürücü yapısına girdi olarak verilir. Enkoder dönüştürücü bu gömme değerleri arasında ilişki veya benzerlik matrisi olarak adlandırılan bir matris hesabı yapar. Bu matris parça gömülmesi aşamasında elde edilen gömme değerlerinin ağırlıklandırılması için kullanılmaktadır. Son olarak ise elde edilen nihai öznitelik vektörleri kullanılarak çoğunlukla MLP tipindeki bir sınıflandırıcı eğitilir. ViT algoritması CNN tabanlı



Şekil 3.6 ViT algoritmasının genel mimarisi. ViT algoritması girdi olarak verilen imgeyi parçalara böler. Her bir parça çok düğümlü algılayıcılara verilerek boyut indirgeme ve öznitelik çıkarma işlemleri yapılır. Çıkarılan öznitelikler enkoder dönüştürücü yapısına girdi olarak verilir ve her bir parça arasındaki ilişki bulunur. Çıkarılan ilişki kullanılarak çıkarılan öznitelikler ağırlıklandırılır. Son katmanda ise genellikle çok katmanlı bir algılayıcı bu öznitelikler ile eğitilerek sınıflandırma işlemi tamamlanır.

algoritmaların aksine yerel öznitelikleri değil imgeyi parçalara bölerek ve enkoder dönüştürücü yapısını kullanarak temsil ediciliği yüksek genel öznitelikler çıkarmayı amaçlar. ViT algoritmasının genel öznitelikleri çıkarması onu ışık seviyesi gibi ortam koşullarına ve el titremesi gibi insandan kaynaklı hatalara karşı çok daha fazla gürbüz yapar. ViT algoritmasının genel mimarisi Şekil 3.6’da verilmiştir.

Her ne kadar ViT algoritması ResNet gibi CNN algoritmalarından çok daha iyi performans sergilese dahi bazı zayıflıkları vardır. Bu tür zayıflıkların giderilebilmesi için literatürde farklı ViT mimarisine sahip modeller önerilmiştir.

ViT algoritması imgenin genel özniteliklerine odaklanması ve göreceli olarak az bir önyargıya (bias) sahip olması sebebiyle CNN algoritmalarına göre çok daha yavaş öğrenmektedir. Bunun yanı sıra çoğunlukla CNN algoritmalarından daha iyi sonuç verebilmesi için çok fazla veriye ihtiyaç duymaktadır. Bu durum ViT algoritmasının özellikle küçük ölçekte altyapılara sahip projelerde kullanılmasının önüne bir engel

olarak çıkmaktadır. Steiner vd. ViT algoritmasının fazla veriye ihtiyaç duyması problemini aşmak için düzenleme (regularization) ve veri artırma (data augmentation) yöntemlerinin kullanılmasını önermiştir [52]. Ayrıca yapılan çalışmada önerilen yöntemin kullanılması toplam eğitim maliyetini de azalttığı vurgulanmaktadır. Lee vd. eksik önyargı probleminin çözülmesi ve ViT modelinin küçük veri setlerinde daha iyi performans sergileyebilmesinin sağlanması için yerel dikkat mekanizması (locality self-attention) ve kayan parça tokenleştirme (shifted patch tokenization) yaklaşımlarını önermiştir [53].

Derin öğrenme modellerinin işlem karmaşıklığı sonuç elde etme hızı ve eğitim maliyeti açısından büyük bir öneme sahiptir. Kullanılan modelin başarılı bir performans sergilemesinin yanında bunu en az kaynak ile yapması da beklenmektedir. ViT algoritması CNN tabanlı derin öğrenme modellerinin aksine $O(d^2)$ işlem karmaşıklığına sahiptir. Buradaki d , parça gömülmesi aşamasında çıkarılan gömme değerlerinin boyutunu ifade etmektedir. Öğrenme aşamasında çok fazla veriye ihtiyaç duyması ve algoritmanın karesel bir işlem karmaşıklığına sahip olması sebebiyle eğitim aşamasını zorlaştırmakta ve güçlü donanım kapasiteleri gerektirmektedir. Modelin karesel karmaşıklığının giderilebilmesi için farklı mimariler ve hesaplama yöntemleri önerilmiştir. Vasu vd. modelin parametre karmaşıklığını azaltmak için bazı normalizasyon ve aktivasyon katmanlarını makalede önerilen RepMixer yapısı ile değiştirmeyi ve çarpanlara ayrılmış (factorized) evrişim katmanlarını kullanmayı önermiştir [54]. Bunun yanı sıra dikkat mekanizmasının hesaplama maliyetinin azaltılabilmesi için matematiksel ve donanımsal en iyileme yaklaşımları da önerilmiştir [55], [56].

CNN algoritmaları çıkarılan özneteliklerin birbiri ile etkileşime girebilmesi için hiyerarşik bir yapı üzerine kurulmuştur. Bir katmanın çıktısı diğer katmanın girdisi olması sebebiyle düşük seviyeli öznetelikler ile yüksek seviyeli öznetelikler aralarında bir etkileşime girebilir ve daha karmaşık olan desen öznetelikleri ortaya çıkabilir. Öte yandan ViT algoritması çıkarılan özneteliklerin birbiri ile hiyerarşik bir şekilde etkileşime girmesine olanak sağlayacak bir mimariye sahip değildir. Liu vd. çok aşamalı ve çok ölçekli bir dikkat mekanizması önererek ViT algoritmasının da hiyerarşik bir şekilde öznetelik çıkarmasına olanak tanımıştır fakat işlenen bilgiler

farklı katmanlara girdi olarak verilmesi sebebiyle algoritmanın paralel işlem yapma kapasitesi ve saniye başına işlenen veri sayısı azalmıştır [57].

Derin öğrenme modellerinin performansları kadar gürbüzlükleri de önemlidir. Dış etmenler tarafından yanıltıcı sonuçlar üretmeye yönelik girişimlere karşı dayanıklı olmaları gerekir. Shao vd. ViT modelinin CNN modeline göre dış saldırılara karşı çok daha fazla gürbüz olduğunu göstermiştir [58]. Bunun sebebi olarak ise ViT algoritmasının verileri yüksek frekanslı öznitelikler ile değil düşük frekanslı öznitelikler ile öğrenmesi olarak göstermişlerdir. İşlenecek olan verilerin her zaman aynı koşullar altında toplanması mümkün değildir. Gerçek hayat problemleri için elde edilen imgeler bazen döndürülmüş veya çok parlak bir ortamdan elde edilebilmektedir. Liu vd. ImageNet-C veri seti üzerinde yaptıkları deneyler sonucu ViT modelinin birçok CNN modeline göre bu tür dış faktörlere karşı çok daha gürbüz olduğunu ortaya koymuştur [59].

Birçok gerçek hayat verisi ile oluşturulan veri kümeleri sınıflar arası veri dengesizliğine sahiptir. Veri kümesinin dengesiz olmasının birçok sebebi olmasına rağmen sık karşılaşılanlardan iki tanesi verinin toplanma aşamasının meşakkatli olması veya toplanması planlanan verinin diğer durumlara göre çok daha az bir gerçekleşme olasılığına sahip olmasıdır. Veri kümesinin dengesiz olması literatürde uzun kuyruklu dağılım (long-tailed distribution) olarak da adlandırılmaktadır. Derin öğrenme yöntemleri ise bu tür veri dengesizliklerinden negatif olarak etkilenmekte ve performansları düşmektedir. Xu vd. yaptığı araştırmaya göre ViT modeli CNN algoritmalarına göre veri dengesizliği probleminden daha fazla etkilenmektedir [60]. Bu problemi çözmek için özdenetimli bir öğrenme yönteminden esinlenilerek farklı bir yaklaşım önermişlerdir.

Her ne kadar ViT algoritması görüntü içerisinde bulunan uzamsal bilginin korunması için konum tokenlerini korusa da bu bilgi CNN algoritmalarındaki kadar net değildir. Bunun en büyük sebeplerinden biri ViT algoritmasının permütasyondan bağımsız öznitelik çıkarmasıdır. Permütasyondan etkilenmemesi uzamsal bilginin azalmasına sebep olurken diğer bir taraftan döndürme gibi uzamsal gürültülere karşı daha gürbüz sonuçlar vermesinde önemli bir rol oynamaktadır [61]. ViT algoritması CNN tabanlı

algoritmaların aksine hiyerarşik bir öznitelik çıkarma işlemi yapmamaktadır. Bu durum çıkarılan özniteliklerin birbirleri ile daha az etkileşime girmelerine sebep olmaktadır. Fakat bu durum aynı zamanda ViT algoritmasının daha fazla paralel işlem yapabilmesine olanak sağlamaktadır.

3.4.4.1 Parça gömülmesi

ViT algoritması girdi olarak verilen A imgesini parçalara böler ve $H \times W \times C$ boyutlarından $L \times (P \times P) \times C$ boyutlarına çevirir. Buradaki H , W , C , L ve P değişkenleri sırasıyla imgenin yüksekliğini, genişliğini, kanal sayısını, parça sayısını ve her bir parçanın boyutlarını ifade etmektedir. Çıkarılan bu parçalar ayrıca NLP'deki kelimeler adlandırmasına benzer olarak görsel kelimeler (visual words) olarak isimlendirilir. Görsel kelimeler bir doğrusal katmana (linear layer) girdi olarak verilir. Doğrusal katmanlar her bir parçanın boyutlarını indirgemesinin yanı sıra aynı zamanda parçalar içerisinde bulunan kıymetli öznitelikleri Eşitlik 3.8'deki gibi çıkarır. ViT algoritmasının görüntüleri parçalara bölmesi ve bu parçalardan gömme değeri (embedding) çıkarması ile elde edilen özniteliklerin gürültü gibi istenmeyen ve ufak değişimlere karşı daha gürbüz olmasını sağlamaktadır. Bunun yanı sıra çıkarılan gömme değeri döndürme gibi uzamsal değişikliklerden çok daha az etkilenmektedir. Eşitlik 3.8'de bulunan n_{km} , k . parçanın m . pikselini ifade etmektedir. Elde edilen bu öznitelikler ayrıca akademik çalışmalarda görsel token (visual token) olarak da adlandırılmaktadır. Görsel tokenler doğrusal katmanlardan geçirilmesi sebebiyle $L \times d$ boyutuna indirgenmiş ve gürültü gibi gereksiz bilgiler elimine edilmiştir. Buradaki L toplam token sayısını ve d ise her bir tokenin eleman sayısını belirtmektedir.

Eşitlik 3.8'de de görüldüğü üzere her bir token aynı doğrusal katmanlar vasıtasıyla elde edilmiştir. Bu durum tokenlerin her birinin aynı uzamsal bilgiye sahip olmasına yani uzamsal bilginin kaybolmasına sebep olmaktadır. Uzamsal bilginin kaybolması imgede bulunan objenin yönü, belirli bir düzleme göre açısı ve bunun gibi bilgilerin kaybolmasına sebep olmaktadır. Bu problemin aşılabilmesi için Eşitlik 3.9'daki gibi gömme değeri boyutu ile aynı boyutta eğitilebilir parametreler gömme değerleri ile toplanır. Böylece 2 farklı parçada aynı objenin bulunmasına rağmen çıkarılan gömme değerleri büyük bir değişime uğramadan uzamsal bilgileri de taşıyabilecektir. Elde

edilen gömme değerlerine sınıflandırma tokeni (classification token) olarak adlandırılan token yeni bir parçaymış gibi Eşitlik 3.10'daki gibi eklenir ve $(L + 1)xd$ boyutunda yeni bir gömme değeri elde edilir. Eklenen bu sınıflandırma tokeni eğitilebilir olması ve bütün imge ve parçalar ile etkileşim içerisinde olması sebebiyle veri kümesini temsil etme niteliğine sahiptir.

$$V = \begin{bmatrix} n_{11} \cdot w_{11} + n_{12} \cdot w_{12} + n_{13} \cdot w_{13} + \dots + n_{1x(p \times p \times c)} \cdot w_{1x(p \times p \times c)} \\ n_{21} \cdot w_{21} + n_{22} \cdot w_{22} + n_{23} \cdot w_{23} + \dots + n_{2x(p \times p \times c)} \cdot w_{2x(p \times p \times c)} \\ \vdots \\ n_{L1} \cdot w_{L1} + n_{L2} \cdot w_{L2} + n_{L3} \cdot w_{L3} + \dots + n_{Lx(p \times p \times c)} \cdot w_{Lx(p \times p \times c)} \end{bmatrix}_{L \times d} \quad (3.8)$$

$$V' = V + P, \quad V, P \in R^{L \times d} \quad (3.9)$$

$$V' = ekle(V', cl), \quad V' \in R_{\square}^{(L+1) \times d}, \quad cl \in R^{1 \times d} \quad (3.10)$$

3.4.4.2 Enkoder dönüştürücü

Enkoder dönüştürücü mimarisi ViT algoritması için çok kritik olmakla beraber dikkat mekanizmasının kullanıldığı bölümdür. Literatürde sıkça kullanılan CNN gibi diğer derin öğrenme algoritmalarından ayrıştığı nokta olma niteliğini taşımaktadır. Enkoder Dönüştürücü bloğu dikkat mekanizmasını (attention mechanism) kullanarak imgelere ait her bir parça veya temsiliyetleri olan gömme değerleri arasındaki ilişkiyi ortaya çıkarmayı amaçlamaktadır. Çıkarılan bu ilişkiye göre imgenin her bir parçasından çıkarılan gömme değerlerinin sonuca olan etkisinin azaltılmasında ve artırılmasında kullanılmaktadır. Enkoder dönüştürücü yapısı katman normalizasyonu (layer normalization) [62], çoklu dikkat mekanizması bloğu (multi-head self attention block), MLP bloğu ve artık bağlantılar olmak üzere toplamda 4 temel aşamadan meydana gelir.

İmgeden çıkarılan gömme değerleri ışık seviyesi gibi dış faktörler sebebiyle çok fazla büyüyebilir veya küçülebilmektedir. Gömme değerlerinin belirli bir aralığın içerisinde olmaması modelin, dağılımsal bilgiden çok verinin genliğine odaklanmasına sebep olmaktadır. Verinin genlik bilgisi dış ortama bağlı olarak değişmesi modelin yanlış karar vermesini sağlamaktadır. Bunun yanı sıra genlik bilgisinin çok fazla veya az olması doğrusal katmanların ağırlıklarının güncellenmesi aşamasında gradyan patlaması veya sönümlenmesi problemine yol açmaktadır. Gradyan patlaması ve sönümlenmesi ise doğrusal katmanların eğitilmesini

engellemektedir. Genlik bilgisin kısıtlanması ve verinin dağılımının daha fazla ön plana çıkarılabilmesi için gömme değerleri enkoder dönüştürücü ile işlenmeden önce katman normalizasyonuna [62] tabi tutulur.

Bilgisayarlı görüde ve derin öğrenmede temsiliyeti yüksek öznitelik çıkarmak modelin performansını doğrudan etkilemektedir. Bu sebeple derin öğrenme modellerinde birçok farklı aşamada öznitelik çıkarma işlemleri uygulanmaktadır. Normalize edilmiş gömme değerleri 3 farklı doğrusal katmana girdi olarak verilerek farklı özelliklere sahip Query(Q), Key(K) ve Value(V) olarak adlandırılan 3 farklı gömme değerleri vektörleri elde edilir. ViT algoritmasının kalbi olan dikkat mekanizması bu vektörleri kullanarak Eşitlik 3.12'deki gibi bir ilişki veya benzerlik matrisi hesaplanır ve Eşitlik 3.11'de tanımlanan softmax fonksiyonundan geçirilir. Öklid, Minkowski veya Manhatttan gibi birçok farklı benzerlik hesaplama yöntemi olmasına rağmen dikkat mekanizması ilişki matrisinin hesaplanmasında kosinüs benzerliğinden faydalanmaktadır. Bunun sebebi kosinüs benzerlik yönteminin diğer yöntemlerden farklı olarak verinin büyüklüğüne değil açısal ve dağılımsal bilgisine odaklanmasıdır. Bunun yanı sıra kosinüs benzerlik algoritması verinin ölçeklenmesinden bağımsız olarak aynı dağılım için aynı sonuçları vermektedir. Çıktı olarak -1 ile 1 aralığında bir değer üretmekte olup -1 değeri iki vektör arasında benzerliğin hiç olmamasını, 1 değeri ise iki vektörün özdeş olduğu anlamına gelmektedir. Elde edilen ilişki matrisinin dağılımsal veya olasılıksal bilgi içerebilmesi için içerisindeki verilerin 0 ile 1 arasında olması gerekmektedir. Verilerin bu aralığa çekilebilmesi için ilişki matrisi Eşitlik 3.11'de tanımlanan softmax fonksiyonundan satır bazında geçirilir. Verilerin 0 ile 1 aralığına çekilebilmesi için ölçekleme veya sigmoid fonksiyonunun kullanılması da mümkündür. Fakat sigmoid fonksiyonu vektörün dağılım bilgisini dikkate almaksızın sadece tek bir verinin genlik bilgisi üzerinden ölçekleme yapmaktadır. Bu durum elde edilen vektör değerlerinin toplamının birden büyük olmasına sebep olmaktadır ve vektörün olasılık dağılımı bilgisi kaybolmaktadır. En az ve en çok ölçekleme (min-max scaling) gibi yöntemler de aynı şekilde vektörün olasılık dağılımı bilgisinin korunmasında başarısız olmaktadır.

Elde edilen yeni matrisin deęerleri her bir paranın dięerleri ile arasındaki iliřkiyi veya benzerlięi ifade eder. Yksek iliřkiye sahip paralar iin benzerlik matrisinin deęeri 1'e yakın iken dřk iliřkiye sahip olanlar iin ise deęerler 0'a daha yakındır. Daha sonra benzerlik matrisi ile V gmme deęeri vektr Eřitlik 3.13'deki gibi matris arpımı yapılır ve aęırlıklandırılmış gmme deęerleri elde edilir. Bylelikle nem arz eden gmme deęerleri artarken grlt veya nemsiz bilgilerin deęeri azalacaktır.

Aęırlıklandırılmış olan gmme deęerleri dikkat mekanizmasının son ařamasında Eřitlik 3.14 gibi bir MLP bloęuna girdi olarak verilir. MLP bloęunda ardışık 2 tane doęrusal ve GeLU aktivasyon fonksiyonu bulunmaktadır. Derin ęrenme modellerinin ęrenme kapasitesinin artırılabilmesi iin doęrusal olmayan verileri de temsil edebilmesi gerekir. Modelin doęrusallıęının azaltılması iin ise aktivasyon fonksiyonları sıka kullanılmaktadır. Literatrde ReLU, Tanh ve GeLU gibi birok farklı aktivasyon fonksiyonu kullanılmaktadır. Kullanılan aktivasyon fonksiyonlarının birbirlerine gre gl ve zayıf ynleri bulunmaktadır. Dikkat mekanizmasında her yerde trevinin olması, srekli olması ve modelin performansını artırması sebebiyle GeLU aktivasyon fonksiyonu kullanılmıştır [63].

ViT modelinde dikkat mekanizmasının ncesi ile sonrasını, MLP bloęunun ncesini ve sonrasını baęlayan 2 tane artık baęlantı bulunmaktadır. Artık baęlantılar ilk olarak ResNet mimarisinde kullanılmıştır [8]. Artık baęlantılar modellerin gradyan patlaması ve snmlenmesi gibi problemlerin engellenmesinde nemli bir rol oynamaktadır. Katmanlar arası bilginin aktarılması sebebiyle nemli bilgilerin kayıp edilmesi gibi durumlar da en aza indirgenmektedir. Bunların yanı sıra artık baęlantılar modelin ezberlemesini engellerken modelin optimizasyonu da kolaylařtırmaktadır.

$$\text{softmax}(x_i) = \frac{e^{x_i}}{\sum_j^n e^{x_j}} \quad (3.11)$$

$$QK = \text{softmax}\left(\frac{KxQ^T}{\sqrt{d}}\right) \quad (3.12)$$

$$x = QKxV \quad (3.13)$$

$$x = \begin{bmatrix} Doğrusal(x) \\ GeLU(x) \\ Doğrusal(x) \\ GeLU(x) \end{bmatrix} \quad (3.14)$$

3.4.4.3 Sınıflandırma

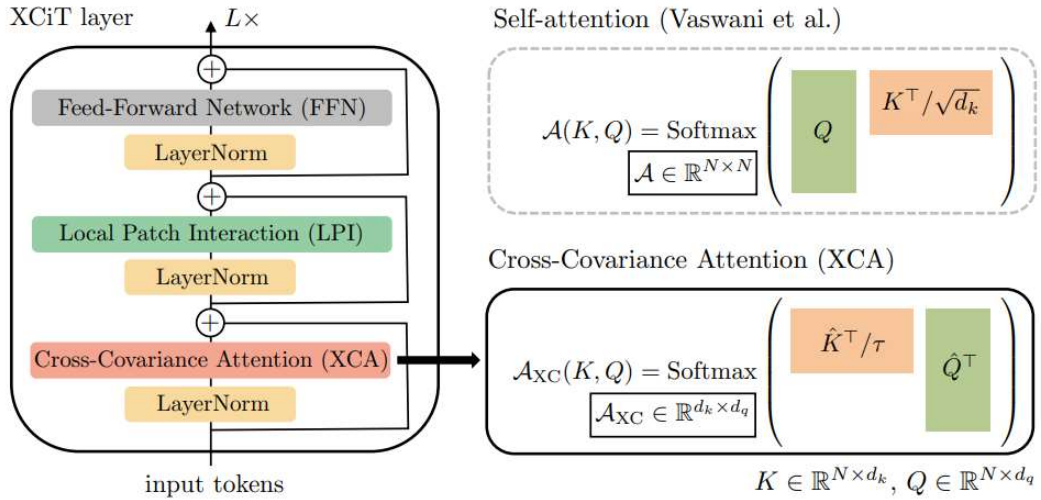
Parça gömülmesi ve dönüştürücü enkoder katmanlarında çıkarılan özniteliklerin işlenerek hangi verinin mitozlu hücre hangi verinin mitozsuz hücre olduğu kararına sınıflandırma aşamasında varılmaktadır. Literatürde Xgboost, karar ağaçları (decision tree), doğrusal katmanlar veya SVM gibi farklı sınıflandırma algoritmaları bulunmaktadır. Özgün ViT algoritmasının sınıflandırma aşamasında ise doğrusal katman kullanılmaktadır. SVM algoritması doğrusal olmayan bir sınıflandırıcı olması sebebiyle doğrusal katmanlara göre daha iyi performans verebilmektedir. Bu çalışmada sınıflandırma performansını artırmak için bütün yaklaşımlar için SVM sınıflandırıcı kullanılmıştır. DINO eğitim yaklaşımı ile eğitilen ViT modelinin son katmandan bir önceki katmanında bulunan gömme değerleri alınır ve SVM algoritması eğitilir. Test aşamasında ise aynı şekilde imgeler ViT algoritmasına girdi olarak verilir ve gömme değerleri elde edilir. Bu gömme değerleri SVM algoritması ile sınıflandırılır ve nihai sonuç elde edilir.

3.4.5 Çapraz kovaryans imge dönüşümü (XCiT)

ViT algoritmasında kullanılan dikkat mekanizması N parça sayısı olmak üzere $O(N^2)$ işlem karmaşıklığına sahiptir. Bu işlem karmaşıklığı ViT algoritmasını yavaş yapmakla beraber yüksek bellek tüketmesine sebep olmaktadır. Dikkat mekanizmasının bu dezavantajı, ViT algoritması ile yüksek çözünürlüklü imgelerin işlenmesini güçleştirmektedir. El-Nouby vd. dikkat mekanizmasını Eşitlik 3.15'deki gibi güncelleyerek karmaşıklık seviyesini $O(d^2)$ 'den $O(d)$ 'ye düşürmüşlerdir [10]. Eşitlik 3.15'de bulunan K ve Q değişkenleri özgün ViT algoritmasında bulunan K ve Q değişkenleri ile aynı şeyi ifade etmektedir. Paydada bulunan τ değişkeni ise öğrenilebilir bir normalize etme değişkenidir. El-Nouby vd. yaptıkları çalışmada dikkat mekanizmasında elde edilen sonuçların -1 ve 1 aralığına ölçeklenmesinin modelin öğrenme kabiliyetini düşüreceğini söylemiştir. Elde edilen çıktıların hem normalize edilebilmesi hem de öğrenme kabiliyetinin kısıtlanmaması için

öğrenilebilir bir ölçekleme katsayısı önermiştir. Önerilen yeni dikkat mekanizmasına ise çapraz-kovaryans dikkat (XCA) ismi verilmiştir. Özgün dikkat mekanizmasında parçalar birbirleri ile etkileşime girerken Eşitlik 3.15 de verilen XCA dikkat mekanizmasında girdi imgelerden çıkarılan öznitelikler birbirleri ile etkileşime girmektedir. XCA dikkat mekanizması öznitelikler arası bir ilişki matrisi hesaplaması sebebiyle parçalar arası ilişki kaybolmaktadır. Bu ilişkinin de hesaplanabilmesi için içerisinde kaybolmaması için XCA katmanından sonra Yerel Parça Etkileşimi (Local Patch Interaction LPI) katmanını kullanılmaktadır. Her bir LPI katmanını içerisinde iki tane derinlik bazlı evrişim (depthwise convolution), grup normalizasyonu (batch normalization) [64] ve GeLU [65] aktivasyon fonksiyonu bulunmaktadır. XCIT sınıflandırmasına ait mimari Şekil 3.7’deki gibidir.

$$QK = \frac{K^T x Q}{\tau} \quad (3.15)$$

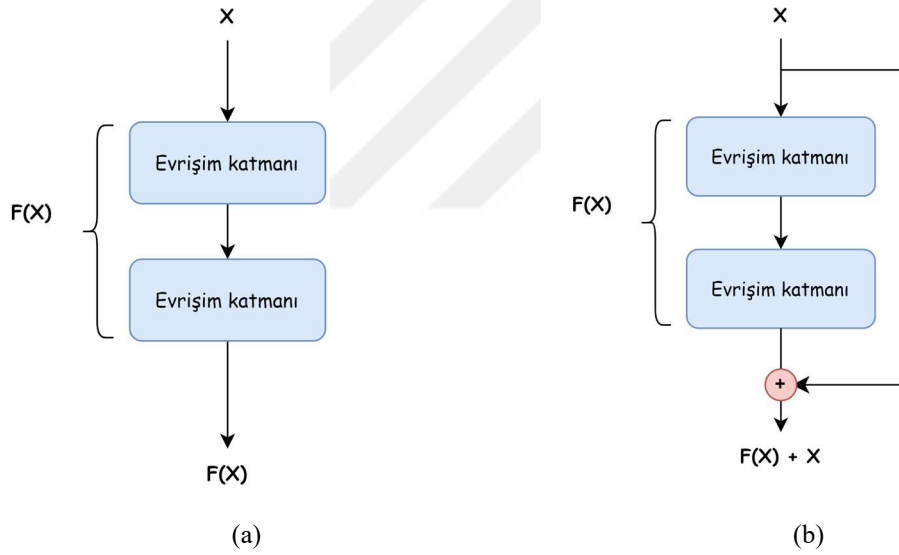


Şekil 3.7 XCIT algoritmasının mimarisi. XCIT algoritması ViT algoritmasının hesaplama karmaşıklığını (d^2)’den $O(d)$ ’e düşürmek için dikkat mekanizmasını Eşitlik 3.15’deki gibi modifiye etmiştir. Yapılan değişiklik sebebiyle parçalar değil çıkarılan gömme değerleri arasında bir ilişki matrisi hesaplanmıştır. Parçalar arasındaki ilişkinin de hesaplanabilmesi için XCA katmanından sonra içerisinde derinlik bazlı evrişim katmanını bulunan LPI katmanını kullanılmıştır [10].

3.4.6 ResNet50

CNN tabanlı berin öğrenme modellerinde kullanılan katman sayısı arttıkça öznitelik çıkarma kapasitesi artar ve genellikle modelin performansı artmaktadır. Fakat modelin katman sayısının artması belirli bir süre sonra problem olmaya başlamaktadır. Kullanılan mimarinin çok fazla parametreye sahip olması modelin

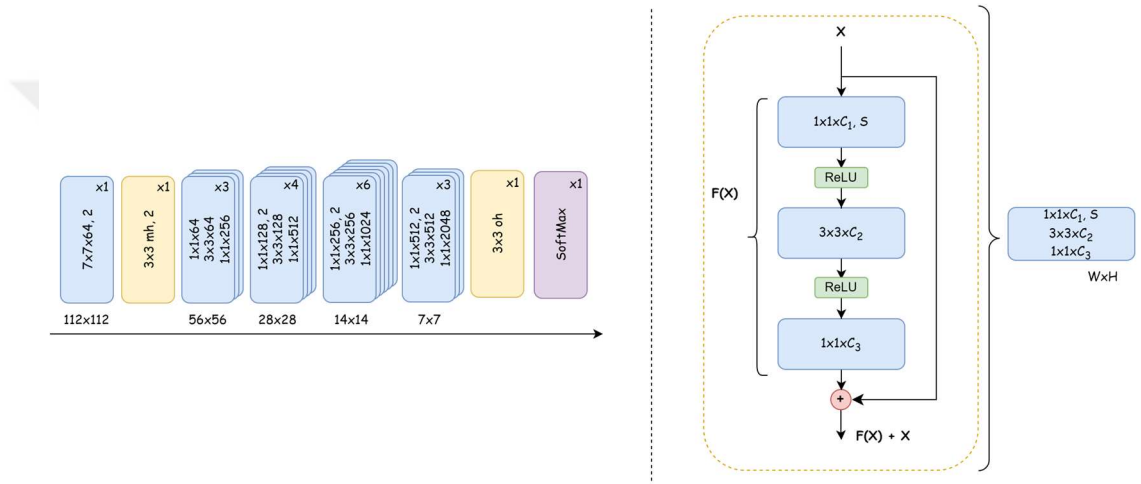
veri kümesini ezberlemeye başlamasına sebep olmaktadır. Katman sayısı derin öğrenme modellerinde sıkça karşılaşılan gradyan sönümlenmesi veya gradyan patlaması gibi problemlerin de sıkça meydana gelmesine neden olmaktadır. He vd. artık bağlantıları önererek bu tür problemlerin aza indirgenmesini ve modelin, katman sayısından bağımsız olarak veri kümesini öğrenebilmesini amaçlamıştır [8]. Artık bağlantılar iki katman arasında paralel bir bağlantı olup birinci katmanın çıktılarını ikinci katmana ekleme işlemi yapmaktadır. Şekil 3.8’de normal ve artık bağlantısı arasındaki farkı gösteren 2 basit evrişim katmanı paylaşılmıştır. He vd. tarafından paylaşılan sonuçlarda modelin katman sayısının 100 üzerine çıkması durumunda bile eğitim işlemi başarıyla tamamlanabilmekte ve diğer modellerden daha iyi performans sergileyebilmekte olduğu görülmektedir. ResNet mimarisinde kullanılan katman sayısına göre ismi ResNetX olarak tanımlanmaktadır.



Şekil 3.8 ResNet ve normal evrişim katmanları. Şekil (a) normal bir CNN mimarisinde bulunan evrişim katmanını ifade etmektedir. Şekil (b) ResNet mimarisinde bulunan ve artık bağlantı içeren bir evrişim katmanını ifade etmektedir. Normal evrişim katmanlarında bir katmanın çıktısı diğer bir katmanın girdisi şeklinde bir ilişki varken ResNet mimarilerinde bir katmanın girdisi kendinden önce gelen katmanların çıktılarının toplamı şeklinde bir ilişki vardır

Buradaki X değişkeni model içerisinde bulunan toplam evrişim katman sayısını belirtmektedir. ResNet50 modelinin mimarisi Şekil 3.9’daki sol bölmedeki gibi verilmiştir. Şekil 3.9’da bulunan sağ taraftaki blok ResNet mimarilerinde bulunan bir bloğun içeriğini ifade etmektedir. C_n değişkeni evrişim katmanında bulunan filtre sayısını ifade etmektedir. S değişkeni evrişim ve havuzlama katmanlarında bulunan filtrelerin kaç piksel kayacağını ifade etmektedir (stride). $W \times H$ ifadesi katmandan

elde edilen çıktılarının genişlik ve yükseklik olarak hangi boyutta olduğunu göstermektedir. mh ve oh kısaltmaları ise sırasıyla maximum havuzlama (maximum pooling) ve ortalama havuzlama (average pooling) anlamlarını taşımaktadır. Resnet mimarisinde bulunan evrişim bloklarının her birinde 3 tane evrişim katmanı bulunmaktadır. Bu blokların ardışık olarak kaç defa kullanıldığı bilgisi ise bloğun sağ üst tarafında bulunan xn ifadesi ile temsil edilmektedir. Ayrıca blokların içerisindeki her bir evrişim katmanı arasında ReLU aktivasyon fonksiyonu bulunmaktadır [66].

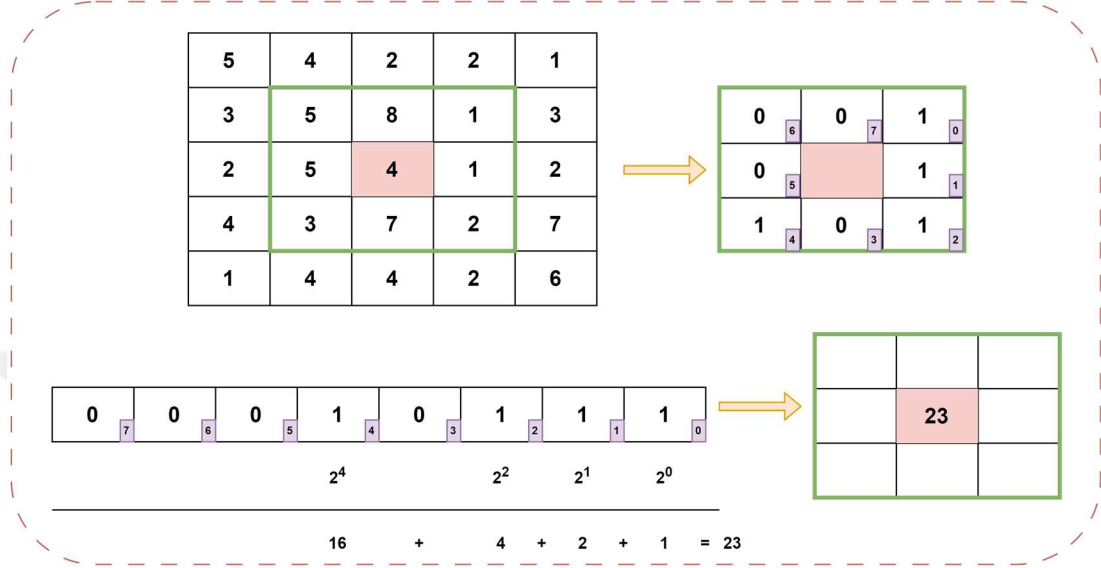


Şekil 3.9 ResNet50 mimarisi. Sağ taraftaki görsel ResNet mimarisi içerisinde bulunan her bir evrişim bloğunu ifade etmektedir. Sol taraftaki görsel ResNet50 mimarisini temsil etmektedir

3.4.7 Yerel ikilik desen (LBP)

Çembersel Yerel İkilik Desen (LBP) algoritması görüntü üzerindeki piksellerin komşu pikseller arasındaki ilişkisine odaklanarak desen tanımlayıcıları ortaya çıkarmayı amaçlamaktadır. p_{xy} görüntü üzerindeki (x, y) koordinatında bulunan bir pikselin değerini ifade etsin. (x, y) koordinatlarında bulunan pikselin etrafındaki diğer 8 pikselin değeri p_{xy} ile karşılaştırılır, p_{xy} piksel değerinden küçük olanlara 1 büyük olanlara ise 0 değeri atanır. $(x, y + 1)$ koordinatı referans alınarak değeri 1 olan her piksel değeri için 1, 2, 4, 8, 16, 32, 64 ve 128 değerleri atanır ve bütün piksellerin değerleri toplanarak (x, y) koordinatında bulunan pikselin yeni değeri olarak atanır. Örneğin $(x - 1, y - 1)$ koordinatındaki pikselin değerinin 1 olması durumunda 1 değeri atanırken $(x + 1, y + 1)$ koordinatları için 128 değeri atanır. LBP işlemi görüntü üzerindeki bütün piksel değerleri için tekrarlandıktan sonra elde

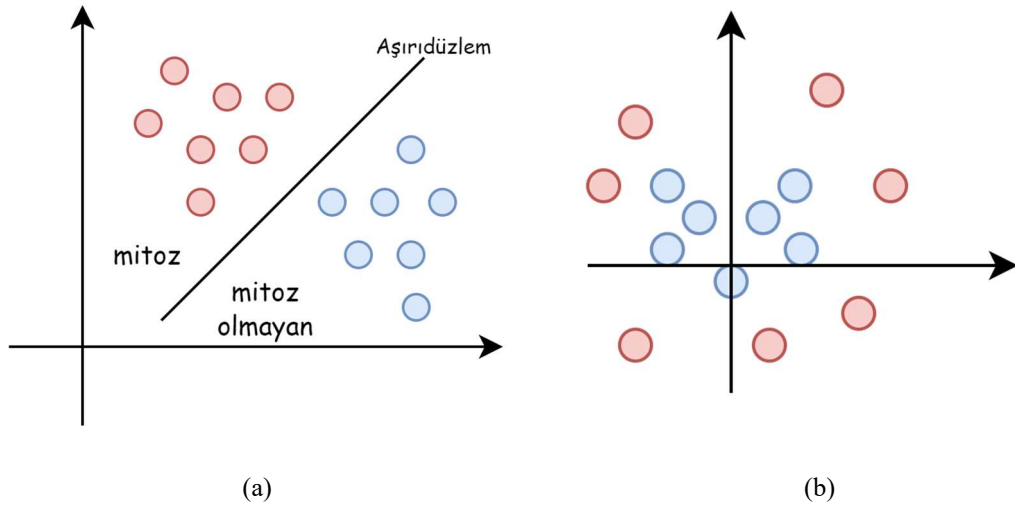
edilen LBP sonuçları üzerinden histogram hesabı yapılır ve bu histogram öznitelik vektörü olarak kullanılır. LBP algoritmasının çalışma prensibini gösteren akış diyagramı Şekil 3.10’da paylaşılmıştır.



Şekil 3.10 LBP'nin çalışma prensibinin akış diyagramı

3.5 Sınıflandırma

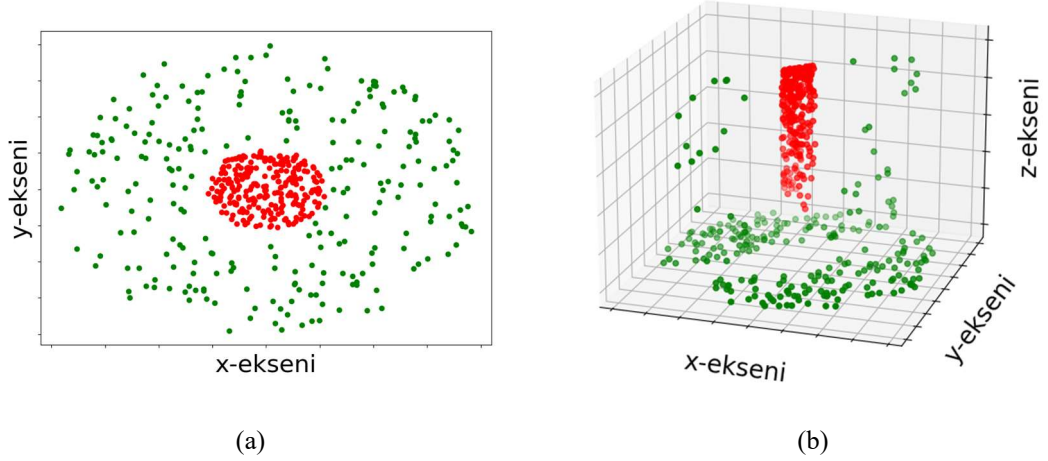
ViT, CNN veya diğer modeller ile çıkarılan gömme değerlerinin sezgisel çıktılara dönüştürülmesinde sınıflandırıcılar kullanılmaktadır. Sınıflandırıcılar eğitildikten sonra girdi olarak verilen gömme değerlerinin bizim anlayabileceğimiz bir boyuta indirir. Literatürde doğrusal, XGBoost [67] ve SVM [68] gibi birçok farklı sınıflandırma algoritması bulunmaktadır. Bu çalışmada çıkarılan gömme değerlerinin sınıflandırılması ve mitozlu hücrelerin tespit edilebilmesi için SVM sınıflandırıcı algoritması kullanılmıştır.



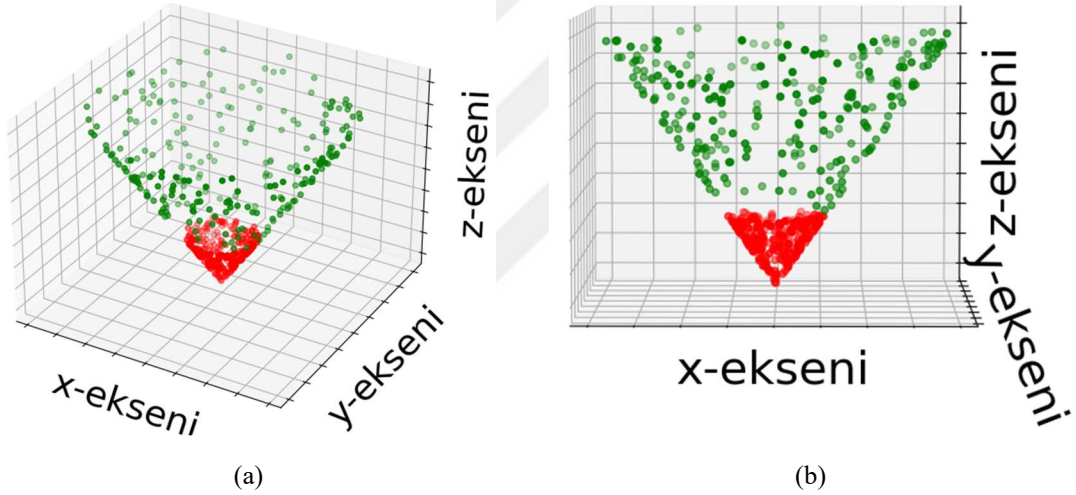
Şekil 3.11 Doğrusal olarak ayrılan ve ayrılmayan veri kümeleri. (a) 2 boyutta doğrusal bir düzlem ile ayrışabilen veriyi temsil etmektedir. (b) 2 boyutta doğrusal olarak ayrışamayan veriyi göstermektedir

3.5.1 SVM

Cortes ve Vapnik tarafından önerilen SVM sınıflandırma [68] algoritması 2 sınıflı ve 2 boyutlu bir sınıflandırma problemi için her bir sınıfa ait gömme değerleri arasında Şekil 3.11a'daki gibi bir aşırıdüzlem (hyperplane) veya diğer bir adıyla karar düzlemi bulmayı hedeflemektedir. Hesaplanmak istenen karar düzlemi iki sınıfa ait veriyi en az hata payıyla ayıracak şekilde optimize edilmektedir. Fakat sınıflandırılmak istenen gömme değerleri Şekil 3.11'deki n . boyutta her zaman doğrusal olmamaktadır. Özellikle gerçek dünyadan toplanan veriler çoğunlukla doğrusal değildir. Bu tür verilerin doğrusal bir sınıflandırıcı ile ayrıştırılması mümkün olmamaktadır. Doğrusal olarak sınıflandırılmaya çalışıldığında veriler tam anlamıyla ayrışmaz ve istenen sonuç elde edilemez. Doğrusal olmayan verilerin sınıflandırılabilmesi için ya kullanılacak olan veriler doğrusallaştırılmalı ya da kullanılacak olan sınıflandırma algoritması da doğrusal olmamalıdır. SVM algoritması verileri doğrusallaştırmak için bir boyut dönüşümü (Kernel trick) yapmaktadır. Doğrusal olmayan ve n boyutlu olan bir veriyi kernel olarak adlandırılan bir fonksiyondan geçirir ve $n+1$ boyutuna yükseltir. Literatürde sıkça kullanılan fonksiyonlar polinom fonksiyonları, doğrusal fonksiyonlar, Gauss fonksiyonları veya Radyal tabanlı Gauss fonksiyonlarıdır (RBF). Bu çalışmada veri kümesinin veri kümesinin $n+1$ boyutunda doğrusallaştırılabilmesi için RBF kullanılmıştır. Şekil 3.12a ve 3.12b'de sırasıyla ham veri ve ham verinin RBF fonksiyonu ile boyutunun artırılmış halinin grafiği verilmiştir.



Şekil 3.12 verinin boyutunun artırılmış ve artırılmamış hali.(a) rastgele oluşturulmuş ham veriyi temsil etmektedir. (b) rastgele oluşturulan ham verinin Radyal tabanlı Gauss fonksiyonu ile boyut artırılmış halini göstermektedir



Şekil 3.13 (a) $z = x^2 + y^2$ ile boyut artırımı yapılmış verinin x, y ve z eksenlerinden görünümü. (b) $z = x^2 + y^2$ ile boyut artırımı yapılmış verinin x ve z eksenlerinden görünümü

Şekil 3.13’de ise durumun daha anlaşılabilir olması için verinin $z = x^2 + y^2$ dönüşümü ile boyutunun artırılması durumunda elde edilen yeni verinin dağılımını gösteren grafikler verilmiştir. Şekil 3.12 ve 3.13’de de görüldüğü üzere 2 boyutta doğrusal olarak ayrışması mümkün olmayan verinin RBF veya $z = x^2 + y^2$ dönüşümü ile 3. boyutta doğrusal bir düzlem ile ayrışabilir bir hale geldiği anlaşılmaktadır.

SVM algoritmasının veriyi farklı bir düzleme taşıyarak doğrusallaştırabilmesinin yanı sıra diğer bir özelliği minimize etmeye çalıştığı hata fonksiyonunun dışbükey olmasıdır. Dışbükey bir hata fonksiyonuna sahip olması SVM algoritmasının

doğrusal katmanlar gibi yerel minimum noktalara takılmayacağını ve mutlak suretle en optimum hata değerine yakınsayacağını garanti etmektedir.



4. BULGULAR VE TARTIŞMA

4.1 Kurulum

MIDOG2021, MIDOG2022 ve ICPR2014 veri kümelerinin yarışma esnasında paylaşılan toplam veri, eğitim ve test olmak üzere sırasıyla %80 ve %20 olmak üzere bölünmüştür. Modellerin eğitilmesi, test edilmesi, başarımlarının hesaplanması ve grafiklerin çizilmesi gibi bütün işlemler Google Colab ortamında gerçekleştirilmiştir [69]. Google Colab 2-6 saat arası kesintisiz ve ücretsiz olarak 16GB, Intel(R) Xeon(R) CPU @ 2.20GHz işlemci ve 16GB Tesla T4 GPU özelliklerine sahip sanal bir makine sunmaktadır. Kullanılan derin öğrenme algoritmaları Python programlama dili ve PyTorch derin öğrenme kütüphanesi [70] ile geliştirilmiştir. Modellerin ImageNet1k veri kümesi ile ön eğitim yapılmış olup ağırlıkları Python'da bulunan Timm kütüphanesi [71] ile temin edilmiştir.

4.2 Değerlendirme Yöntemi

Literatürde çok farklı değerlendirme yöntemleri bulunmaktadır. Kullanılan değerlendirme yöntemi çözülmek istenen probleme göre de değişebilmektedir. Örneğin bölütleme problemleri için jaccard veya dice skoru kullanılırken sınıflandırma için kesinlik veya doğruluk gibi değerlendirme yöntemleri kullanılmaktadır. Bu çalışmada yapılan deneylerin objektif olarak değerlendirilebilmesi için üç farklı değerlendirme yöntemi kullanılmıştır. Bunlar f-skoru, kesinlik ve duyarlılıktır. Kullanılan her bir değerlendirme yönteminin farklı avantajları ve dezavantajları bulunmaktadır. Bu değerlendirme yöntemleri daha detaylı olarak alt başlıklar altında incelenmiştir.

4.2.1 Duyarlılık

Duyarlılık (Recall) değeri modelin mitoz hücrelerini ne kadar iyi tespit edebildiğine odaklanırken mitoz olmayan hücrelerin tespitini göz ardı eder. Asıl tespit edilmek istenmeyen sınıfın başarımlarını dikkate almaması sebebiyle tıbbi görüntüler gibi özellikle veri dengesizliğinin çok fazla olduğu problemlerde sıklıkla kullanılmaktadır. Duyarlılık değeri Eşitlik 4.1'de verildiği gibi Gerçek Pozitif (TP) değeri Gerçek Pozitif ve Yanlış Negatif (FN) değerlerine bölünerek elde edilir.

$$R = \frac{TP}{TP+FN} \quad (4.1)$$

4.2.2 Kesinlik

Kesinlik (Precision) değeri modelin mitozlu ve mitozlu olmayan hücreleri ne kadar iyi tespit edebildiğine odaklanmaktadır. Duyarlılık değerinin aksine mitozlu olmayan hücreleri göz ardı etmez. Kesinlik değeri Eşitlik 4.2’de olduğu gibi TP değerinin TP ve Yanlış Pozitif (FP) değerlerinin toplanmasına bölünmesi ile elde edilir.

$$P = \frac{TP}{TP+FP} \quad (4.2)$$

4.2.3 F-skoru

F-skoru, kesinlik ve duyarlılık değerlerinin harmonik ortalamasını alarak bir denge kurmayı amaçlamaktadır. Değeri ise Eşitlik 4.3’deki gibi hesaplanmaktadır. Modellerin dengeli veya dengesiz veri kümelerindeki performanslarının değerlendirilmesi için kullanılan etkili bir yöntemdir.

$$f = 2 \cdot \frac{P \cdot R}{P+R} \quad (4.3)$$

4.3 Deneysel Sonuçlar

DINO bir öğrenme yaklaşımı olması sebebiyle CNN gibi farklı mimarideki modeller ile de kullanmak mümkündür. DINO öz denetimli öğrenme yönteminin farklı modeller ile sergilediği performansı değerlendirmek için ResNet50, XCIT ve ViT modelleri kullanılmıştır. Bunun yanı sıra modelin büyüklüğünün etkisini görebilmek için yapılan deneyler ViT modelleri için farklı parametreler ile tekrarlanmıştır.

MIDOG2021 ve MIDOG2022 veri kümeleri için sırasıyla Tablo 4.2 ve 4.3’te yapılan deneylerin sonuçları paylaşılmıştır. Model sütunu kullanılan modelin ismini ve mimarisini temsil etmektedir. ViT B/16, ViT mimarisinin temel (base) modelinin 16 parça boyutu (patch size) ile kullanıldığını ifade etmektedir. Parametre kolonu ise bahsi geçen modelin toplam kaç milyon eğitilebilir parametreye sahip olduğunu gösterir. Kullanılan dönüştürücü tabanlı modellerin parametre bilgileri daha detaylı

olarak Tablo 4.1’de verilmiştir. Tablo 4.2 ve 4.3 incelendiğinde kullanılan modelin toplam parametre sayısının sonuca olan etkisinin en fazla %2.5 civarında olduğu gözlemlenmektedir.

Tablo 4.1 Kullanılan modellerin bazı özellikleri

Model	Girdi	Parça	Gömme değeri boyutu	Baş	Katman	Parametre (M)
ViT S/16	224x224	16	384	6	12	21M
ViT B/16	224x224	16	768	12	12	86M
ViT B/8	224x224	8	768	12	12	86M
XCIT M/16	224x224	16	512	24	8	25M
XCIT M/8	224x224	8	512	24	8	84M
XCIT S/16	224x224	16	384	12	8	84M
XCIT S/8	224x224	8	384	12	8	26M

Bunun yanı sıra aynı modeller için parça boyutunun (patch size) değiştirilmesi de XCIT M/16 ve M/8 modelleri için performansa en fazla %2-3 civarında etki etmektedir. CNN mimarisine sahip olan ResNet-50 modeli de dönüştürücü tabanlı ViT modelleri ile benzer sonuçlar vermektedir. En düşük performansı sergileyen algoritma ise geleneksel bir yöntem olan LBP yaklaşımıdır.

Tablo 4.4’te MIDOG2021 veri kümesi için düzenlenen yarışmada elde edilen başarımlar oranları verilmiştir. Tablo kolonları sırasıyla takım ismi, kullanılan algoritmanın ismi, f-skoru, kesinlik ve duyarlılık değerlerini ifade etmektedir. Takım kolonunda bulunan veriler köprü bağlantıya sahip olup önerilen çözümlere gitmek mümkündür. Düzenlenen yarışma toplamda 2 bölümde gerçekleşmiştir: eğitim ve test aşamaları. Eğitim aşamasında yarışmaya katılan takımlardan verilen eğitim kümesi kullanılarak modellerinin eğitilmesi beklenmiştir.

Test aşamasında ise eğitilen modellerin performansları daha önce kullanılmayan görüntüleri içeren test kümesi üzerinde değerlendirilmiştir. Yarışma esnasında sunulan test veri kümesi herkese açık olarak paylaşılmaması sebebiyle bu çalışmada kullanılamamıştır. Bu çalışmada sunulan test performans değerleri yarışmada

paylaşılan veri kümesinin %80 eğitim %20 test olarak ayrılması ile elde edilen test kümesi üzerinden hesaplanmıştır.

Tablo 4.2 MIDOG2021 veri kümesi için farklı modellerin performansları

Model	Parametre	F1-skoru	Kesinlik	Duyarlılık
ViT S/16 (dino)	21M	0.7778	0.7758	0.7802
ViT S/8 (dino)	21M	0.7773	0.7731	0.7839
ViT B/16 (dino)	86M	0.7811	0.7784	0.7843
ViT B/8 (dino)	86M	0.7818	0.7862	0.7784
ResNet-50 (dino)	25M	0.7715	0.7677	0.7764
XCIT M/16 (dino)	84M	0.7628	0.7605	0.7654
XCIT M/8 (dino)	84M	0.7833	0.7815	0.7853
XCIT S/16 (dino)	26M	0.7809	0.7771	0.7859
XCIT S/8 (dino)	26M	0.7689	0.7662	0.7721
ResNet-50 (denetimli)	25.6M	0.7650	0.7613	0.7699
LBP		0.5739	0.6128	0.5770

Tablo 4.3 MIDOG2022 veri kümesi için farklı modellerin performansları

Model	Parametre	F1-skoru	Kesinlik	Duyarlılık
ViT S/16 (dino)	21M	0.7788	0.7808	0.7778
ViT S/8 (dino)	21M	0.7646	0.7747	0.7547
ViT B/16 (dino)	86M	0.7885	0.7909	0.7874
ViT B/8 (dino)	86M	0.7935	0.7949	0.7927
ResNet-50 (dino)	25M	0.7823	0.7840	0.7815
XCIT M/16 (dino)	84M	0.7789	0.7812	0.7779
XCIT M/8 (dino)	84M	0.7797	0.7818	0.7787
XCIT S/16 (dino)	26M	0.7820	0.7812	0.7838
XCIT S/8 (dino)	26M	0.7765	0.7792	0.7754
ResNet-50 (denetimli)	25.6M	0.8448	0.8448	0.8449
LBP		0.6399	0.6457	0.6348

Tablo 4.4 MIDOG2021 yarışması için takımların test veri kümesindeki model performansları

Takım	Algoritma	F-skoru	Kesinlik	Duyarlılık
adamshephard	HEUC-Rebuild	0.7657	0.7282	0.8072
n.zamanitajeddin	Mitosis-Old-Original	0.7656	0.7543	0.7771
mositemp	mitosis-Old	0.7656	0.7543	0.7771
n.zamanitajeddin	HEUC	0.7640	0.7157	0.8192
gauthierroy	V2_updated	0.7578	0.7820	0.7349
skylibration	MIDOG_jitt	0.7550	0.7237	0.7891
whwl0882	MIDOG baseline-mk2	0.7548	0.8125	0.7048
n.zamanitajeddin	HEUC	0.7544	0.7329	0.7771
marcaubreville	MIDOG Reference Test Algorithm	0.7514	0.6938	0.8192
zerostarcraft	release	0.7500	0.7792	0.7228
mositemp	mitosis-new	0.7500	0.7096	0.7951
salarra	MIDOG	0.7492	0.7707	0.7289
zerostarcraft	release	0.7484	0.7828	0.7168
AI_medical	mitos1	0.7465	0.6943	0.8072
hard-name	1-0.4	0.7462	0.7577	0.7349

Tablo 4.5 MIDOG2022 yarışması için takımların test veri kümesindeki model performansları

Takım	Algoritma	F-skoru	Kesinlik	Duyarlılık
adamshephard	Mitosis Detection, Fast and Slow (MDFS)	0.7844	0.7855	0.7833
mositemp	MDFS_T2_SC	0.7719	0.7982	0.7472
saipradeepvg	STG2	0.7704	0.7336	0.8111
sujaatha.kotte	STG_ENH	0.7607	0.7118	0.8166
AI_medical	midog2022task2	0.7543	0.7764	0.7333
kgicmd	CSN4-STG1	0.7323	0.7313	0.7333
SKJP	MIDOG2022_SK	0.7163	0.7396	0.6944
frauke.wilm	Baseline 2 for MIDOG 2022, RetinaNet-DA	0.7152	0.6955	0.7361
pz.suen	RetinaNet-All	0.6966	0.7867	0.6249
mlafarge	P4-Detector	0.6904	0.6566	0.7277
florent.cd	IN_i2i_aug (task_2)	0.6645	0.7819	0.5777
marin.scalbert	IN_i2i_aug (task_2)	0.6604	0.6437	0.6777
JonaGanz	Baseline 1 for MIDOG 2022, Mask-RCNN	0.6285	0.6320	0.6249
engin.bozaba	MitDet	0.6035	0.7148	0.5222
wangzy.gz.02	Retina -V3	0.5809	0.4729	0.7527

Tablo 4.6 MIDOG2021 yarışması için literatürde bulunan çalışmaların performansları

Çalışma	Test kümesi	F1-skoru	Kesinlik	Duyarlılık
[72]	Eğitim kümesi bölünmüş	0.7577	0.7820	0.7349
[73]	Ön çalışma kümesi	0.7550	0.7238	0.7892
[74]	Ön çalışma kümesi	0.7500	0.7792	0.7229
[75]	Ön çalışma kümesi	0.7492	0.7707	0.7289
[76]	Eğitim kümesi bölünmüş	0.747	0.76	0.75
[77]	Ön çalışma kümesi	0.7323	0.7313	0.7333
[78]	Ön çalışma kümesi	0.6541	0.6895	0.7289
[79]	Ön çalışma kümesi	0.6645	0.7319	0.6084
[80]	Belirtilmemiş	0.6575	0.6481	0.6673
[81]	Belirtilmemiş	0.52	0.71	0.41
Önerilen	Eğitim kümesi bölünmüş	0.7833	0.7815	0.7853

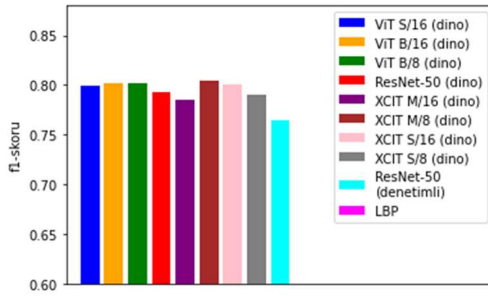
Tablo 4.4'te de görüldüğü üzere en iyi takım 0.7657 f-skoru, 0.7282 kesinlik ve 0.8072 duyarlılık değerlerini elde etmiştir. Bütün takımlar göz önünde bulundurulduğunda birinci ve sonuncu takımın f-skoru arasında sadece %2'lik bir oynama olduğu görülmektedir. Fakat kesinlik ve duyarlılık kriterleri için bu fark çok daha fazladır. Tablo 4.5'te MIDOG2022 yarışmasının sonuçları paylaşılmıştır. MIDOG2022 yarışması için de süreç MIDOG2021'de olduğu gibi gerçekleşmiştir. MIDOG2022 sonuçları için sonuçlar değerlendirildiğinde f-skoru için takımlar arası makas açılırken duyarlılık kriteri için aralık azalmıştır.

Tablo 4.6'da ve Tablo 4.7'de MIDOG2021 ve MIDOG2022 veri kümeleri için literatürde bulunan sonuçlar ve bu çalışmada elde edilen sonuçlar paylaşılmıştır. Tablo 4.7'de bulunan çalışma sayısının Tablo 4.6'dan oldukça az olmasının temel sebebi MIDOG2022 veri kümesinin yeni yayınlanmış olmasıdır. Şekil 4.1a ve 4.1b'de sırasıyla MIDOG2021 ve MIDOG2022 veri kümeleri için bütün modellerin f-skoru değerleri çubuk grafiği şeklinde verilmiştir. MIDOG2021 ve MIDOG2022 veri kümeleri için sırasıyla en iyi sonucu elde eden XCIT M/8 ve ResNet50 denetimli modellerin hata matrisleri Şekil 4.2a ve 4.2b'de verilmiştir. Tablo 4.2, 4.4 ve 4.6 incelendiğinde MIDOG2021 veri kümesi için en yüksek sonuçlar 0.7833 f-skoru, 0.7815 kesinlik değeri ve 0.7853 duyarlılık değeri olarak elde edilmiştir. Bunun yanı sıra MIDOG2022 için metrikler Tablo 4.3, 4.5 ve 4.7'de de görüldüğü üzere sırasıyla

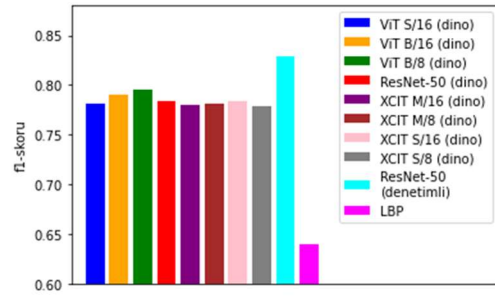
0.8448, 0.8448 ve 0.8449 şeklindedir. Elde edilen f-skoru, kesinlik ve duyarlılık değerlendirme metriklerinin birbirine yakın olduğu görülmektedir. Skorların bu denli birbirine yakın olmaları modelin aynı zamanda veri dengesizliği ile ilgili bir problem yaşamadığını da göstermektedir.

Tablo 4.7 MIDOG2022 veri kümesi için literatürde bulunan çalışmaların performansları

Çalışma	Test kümesi	F1-skoru	Kesinlik	Duyarlılık
[82]	Yarışma test veri kümesi	0.7569	-	-
[83]	Yarışma test veri kümesi	0.697	-	-
[84]	Yarışma test veri kümesi	0.6658	-	-
Önerilen	Eğitim kümesi bölünmüş	0.7807	0.7812	0.7813

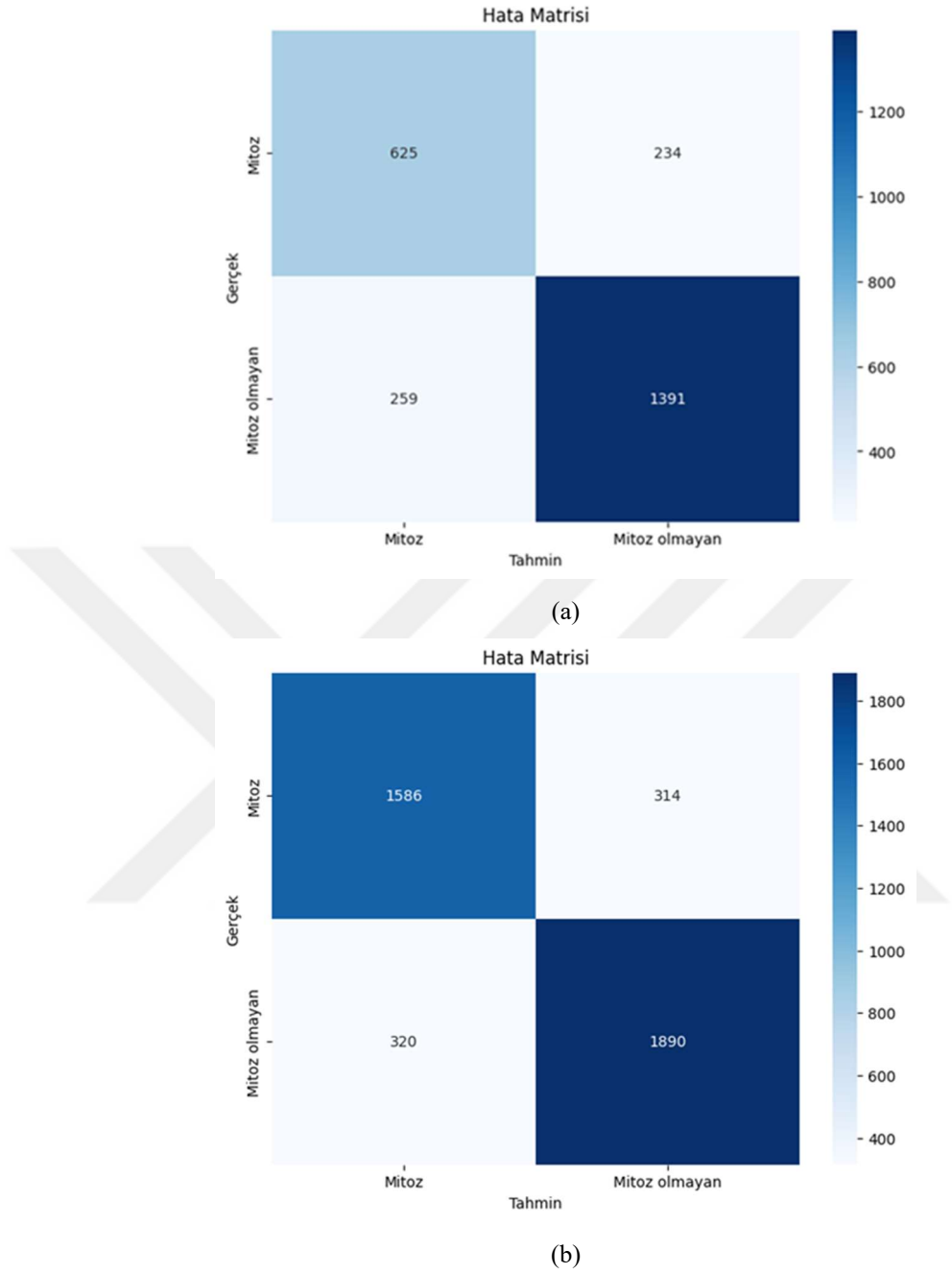


(a)



(b)

Şekil 4.1 MIDOG2021 ve MIDOG2022 elde edilen sonuçların grafiği. (a) MIDOG2021 test veri kümesi için kullanılan bütün modellerin f-skoru. (b) MIDOG2022 test veri kümesi için kullanılan bütün modellerin f-skoru



Şekil 4.2 Sırasıyla MIDO2021 ve MIDO2022 veri kümeleri için hata matrisi. (a) MIDO2021 test veri kümesi için hata matrisi. (b) MIDO2022 test veri kümesi için hata matrisi

ICPR2014 veri kümesinde bulunan 749 adet mitozlu hücrelerin sadece 690 tanesi doğru olarak tespit edilebilmiştir. Sınıflandırma aşamasında ise tespit edilen 690 tane mitozlu hücre ile işlemlere devam edilmiştir. Kullanılan algoritmaların sınıflandırma başarımının ölçülmesinin yanı sıra veri dengesizliği problemine karşı performanslarının incelenebilmesi için 5 farklı alt veri kümeleri elde edilmiştir. Veri dengesizliğinin bulunmaması kullanılan algoritmaların ICPR2014 veri kümesi için

sınıflandırma başarımını temsil etmektedir. Her bir veri kümesindeki toplam örnek sayısı, mitozlu ve mitozlu olmayan örnek sayısı Tablo 4.13'deki gibi paylaşılmıştır. Tablo 4.13'de bulunan dengesizlik sütunu mitozlu ve mitozlu olmayan hücreler arasında kaç katlık bir fark olduğunu, ETÖ kolonu eğitim toplam örnek sayısını, TTÖ kolonu toplam test örnek sayısını, EMÖ sütunu eğitim mitozlu örnek sayısını, EMOÖ sütunu eğitim mitozlu olmayan örnek sayısını, TMÖ test mitozlu örnek sayısını ve TMOÖ test mitozlu olmayan örnek sayısını ifade etmektedir. Veri dengesizliğinin artması durumunda modellerin performansları 1 ve 5 kat dengesizlikler için sırasıyla Tablo 4.8 ve Tablo 4.12 arasında paylaşılmıştır. Veri dengesizliğinin bulunmaması durumunda Tablo 4.8'de de görüleceği üzere özdenetimli olarak eğitilen ViT B/8 mimarisi en iyi performansı sergilemiştir.

Değerlendirme metriği olarak f1-skoru baz alındığında dengeli ICPR2014 veri kümesi için ViT B/8 mimarisi 0.9239 değerini elde etmiştir. Tablo 4.9, 4.10, 4.11 ve 4.12'de de paylaşıldığı gibi 2 kat veri dengesizliği durumunda XCIT M/8 mimarisi 0.8956 f1-skoru performansını sergilerken 3 kat, 4 kat ve 5 kat veri dengesizliği durumunda denetimli olarak ICPR2014 eğitim veri kümesi ile eğitilen ResNet50 mimarisi sırasıyla 0.8627, 0.8695 ve 0.8268 f1-skoru ile en yüksek değerlendirme metriğini elde etmiştir. Bunun yanı sıra bütün ICPR2014 veri kümeleri için geleneksel görüntü işleme yöntemi olan LBP yaklaşımı en düşük performansı sergilemiştir.

Tablo 4.8 ICPR2014 1 kat dengesizlik durumunda farklı modellerin performansları

Model	Parametre	F1-skoru	Kesinlik	Duyarlılık
ViT S/16 (dino)	21M	0.9130	0.9134	0.9130
ViT B/16 (dino)	86M	0.9203	0.9204	0.9203
ViT B/8 (dino)	86M	0.9239	0.9241	0.9239
ResNet-50 (dino)	25M	0.9130	0.9134	0.9130
XCIT M/16 (dino)	84M	0.9239	0.9239	0.9239
XCIT M/8 (dino)	84M	0.9239	0.9239	0.9239
XCIT S/16 (dino)	26M	0.8841	0.8841	0.8841
XCIT S/8 (dino)	26M	0.9203	0.9204	0.9203
ResNet-50 (denetimli)	25.6M	0.9056	0.9088	0.9057
LBP		0.7199	0.7244	0.7210

Veri dengesizliğinin 2 kattan fazla artması durumunda özdenetimli öğrenme yöntemi ile eğitilmiş modellerin performansının denetimli öğrenme yöntemi ile eğitilen ResNet-50'de aşağı düştüğü gözlemlenmiştir. Özdenetimli olarak eğitilen ResNet-50 mimarisi de bu düşüşün içerisinde olduğu açıkça görülmektedir. Veri dengesizliğinin artması göz önünde bulundurulduğunda özdenetimli olarak eğitilen ResNet-50 mimarisinin değerlendirme kriteri 0.9130'dan 0.7074'e gerileyerek yaklaşık olarak %22.5'luk bir düşüş sergilemiştir. Denetimli öğrenme yaklaşımı ile eğitilen ResNet-50 mimarisi için düşüş miktarı 0.9056'dan 0.8268'e olup yaklaşık olarak %8.7 civarındadır. Özdenetimli öğrenme yöntemi ile eğitilen modeller arasında ise en büyük ve en küçük düşüşü sırasıyla ViT B/8 modeli %19.8 ve XCIT S/16 modeli %19.3 olarak gerçekleştirir. Genel olarak değerlendirilmesi durumunda ise özdenetimli öğrenme yöntemi ile eğitilen modeller için düşüş ortalama %21.0 civarındadır.

Şekil 4.3a ve b ICPR2014 veri kümesi için denetimli öğrenme yaklaşımı ile eğitilen ResNet-50 modeli için sırasıyla doğruluk ve hata fonksiyonunun her bir eğitim döngüsüne (epoch) göre çizdirilmiş halini temsil etmektedir. Farklı renkler grafik üzerinde de belirtildiği üzere n kat veri dengesizliğini ifade etmektedir. Şekil 4.3a ve b'nin gösterdiği üzere veri dengesizliğinin artması halinde başarımlar çok daha kısa sürede yükselmektedir. Veri dengesizliği arttıkça aynı eğitim döngüsünde varılan hata da daha azalmaktadır. Tablo 4.8 – 4.12 arası dikkate alındığında hatanın azalmasına ve başarımların daha çabuk daha yüksek seviyelere ulaşması modelin test veri kümelerindeki başarımlarını artırmamakta aksine düşürmektedir. Bu durum veri dengesizliğinin modelin problemi öğrenmesine engel olduğunu ve problemi yani eğitim veri kümesini ezberlemesine sebep olduğunu göstermektedir.

Tablo 4.9 ICPR2014 2 kat dengesizlik durumunda farklı modellerin performansları

Model	Parametre	F1-skoru	Kesinlik	Duyarlılık
ViT S/16 (dino)	21M	0.8649	0.8814	0.8533
ViT B/16 (dino)	86M	0.8848	0.9009	0.8732
ViT B/8 (dino)	86M	0.8662	0.8893	0.8514
ResNet-50 (dino)	25M	0.8674	0.8852	0.8551
XCIT M/16 (dino)	84M	0.8956	0.9147	0.8822
XCIT M/8 (dino)	84M	0.8956	0.9147	0.8822
XCIT S/16 (dino)	26M	0.8453	0.8713	0.8297
XCIT S/8 (dino)	26M	0.8612	0.8812	0.8478
ResNet-50 (denetimli)	25.6M	0.8761	0.8931	0.8641
LBP		0.5688	0.7324	0.5851

Tablo 4.10 ICPR2014 3 kat dengesizlik durumunda farklı modellerin performansları

Model	Parametre	F1-skoru	Kesinlik	Duyarlılık
ViT S/16 (dino)	21M	0.8474	0.8882	0.8213
ViT B/16 (dino)	86M	0.8544	0.8889	0.8309
ViT B/8 (dino)	86M	0.8574	0.8905	0.8345
ResNet-50 (dino)	25M	0.8514	0.8872	0.8273
XCIT M/16 (dino)	84M	0.8336	0.8661	0.8116
XCIT M/8 (dino)	84M	0.8336	0.8661	0.8116
XCIT S/16 (dino)	26M	0.8179	0.8664	0.7899
XCIT S/8 (dino)	26M	0.8305	0.8644	0.8080
ResNet-50 (denetimli)	25.6M	0.8627	0.8761	0.8514
LBP		0.4285	0.375	0.5

Şekil 4.4’de ICPR2014 veri kümesi için en iyi modellerin hata matrisleri paylaşılmıştır. Şekil 4.4a, 4.4b, 4.4c, 4.4d ve 4.4e sırasıyla 1 kat, 2 kat, 3 kat, 4 kat ve 5 kat veri dengesizliğine aittir. Kullanılan modeller ise sırasıyla ViT B/8, XCIT M/8, denetimli ResNet50, denetimli ResNet50 ve denetimli ResNet50’dir. Yapılan analizler göz önünde bulundurulduğunda dengesiz veri kümesi problemi ile denetimli öğrenme yaklaşımının daha iyi başa çıkabileceği gözlemlenmiştir. Bunun yanı sıra her ne kadar dönüştürücü tabanlı ve CNN tabanlı yöntemler farklı çalışma prensiplerine sahip olsalar da denetimli öğrenme yöntemleri ile eğitildiklerinde

dengeşiz veri kümesi problemlerine karşı benzer gürbüzlük göstermektedirler. Parametre sayısı fazla olan modeller de keza aynı şekilde veri dengeşizliğı problemi için değerdendirilmesi durumunda benzer yaklaşımı sergilemektedirler. LBP geleneksel görüntü işleme yöntemi dikkate alındığında veri dengeşizliğinin bulunmaması durumunda Tablo 4.8’de de görüldüğü üzere 0.7199 f1-skoru metriğı elde edilmiştir. Veri dengeşizliğinin 2 kat olmasıyla beraber yaklaşık olarak %21’lik bir değerd kaybedilir ve f1-skoru metriğı hızlıca 0.5688’e düşer. Dengeşizliğin 3 kat ve daha fazla olması durumunda f1-skoru daha da geriler ve 0.4757 seviyesine kadar geriler. Ayrıca duyarlılık değerdinin 0.5’e yakın olması ve kesinlik değerdinin de 0.40 civarında olması uygulanan yöntemin artık rastgele sınıflandırmalar yaptığı anlaşılmaktadır. Tablo 4.14’te ICPR2014 veri kümesi için yapılan bazı çalışmaları ve bunların değerdendirme metrikleri bulunmaktadır.

Tablo 4.11 ICPR2014 4 kat dengeşizlik durumunda farklı modellerin performansları

Model	Parametre	F1-skoru	Kesinlik	Duyarlılık
ViT S/16 (dino)	21M	0,7746	0,8894	0,7254
ViT B/16 (dino)	86M	0,7817	0,8730	0,7370
ViT B/8 (dino)	86M	0.8082	0.8771	0.7681
ResNet-50 (dino)	25M	0.7566	0.8612	0.7116
XCIT M/16 (dino)	84M	0.7817	0.8730	0.7370
XCIT M/8 (dino)	84M	0.7817	0.8730	0.7370
XCIT S/16 (dino)	26M	0.7851	0.8745	0.7406
XCIT S/8 (dino)	26M	0.8076	0.8975	0.7609
ResNet-50 (denetimli)	25.6M	0.8695	0.8912	0.8514
LBP		0.4546	0.4167	0.5

Tablo 4.12 ICPR2014 5 kat dengesizlik durumunda farklı modellerin performansları

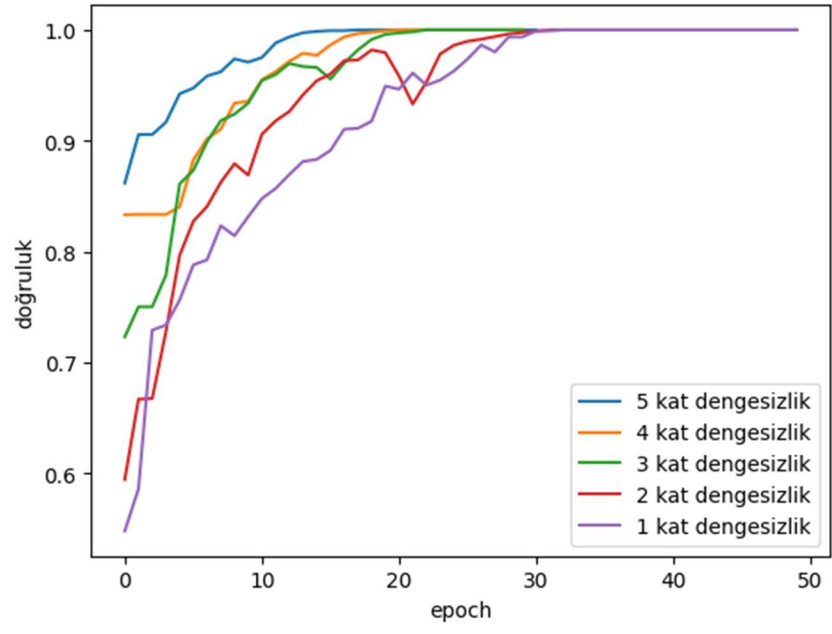
Model	Parametre	F1-skoru	Kesinlik	Duyarlılık
ViT S/16 (dino)	21M	0.7074	0.8418	0.6574
ViT B/16 (dino)	86M	0.7115	0.8442	0.6610
ViT B/8 (dino)	86M	0.7409	0.8819	0.6842
ResNet-50 (dino)	25M	0.7074	0.8594	0.6549
XCIT M/16 (dino)	84M	0.7385	0.8870	0.6810
XCIT M/8 (dino)	84M	0.7385	0.8870	0.6810
XCIT S/16 (dino)	26M	0.7129	0.8687	0.6589
XCIT S/8 (dino)	26M	0.7341	0.8673	0.6799
ResNet-50 (denetimli)	25.6M	0.8268	0.8689	0.7952
LBP		0.4752	0.4527	0.5

Tablo 4.13 ICPR2014 dengesiz veri kümeleri örnek sayısı. ETÖ eğitim toplam örnek, TTÖ, test toplam örnek, EMÖ eğitim mitozlu örnek, EMOÖ eğitim mitozlu olmayan örnek, TMÖ test mitozlu örnek, TMOÖ test mitozlu olmayan örnek

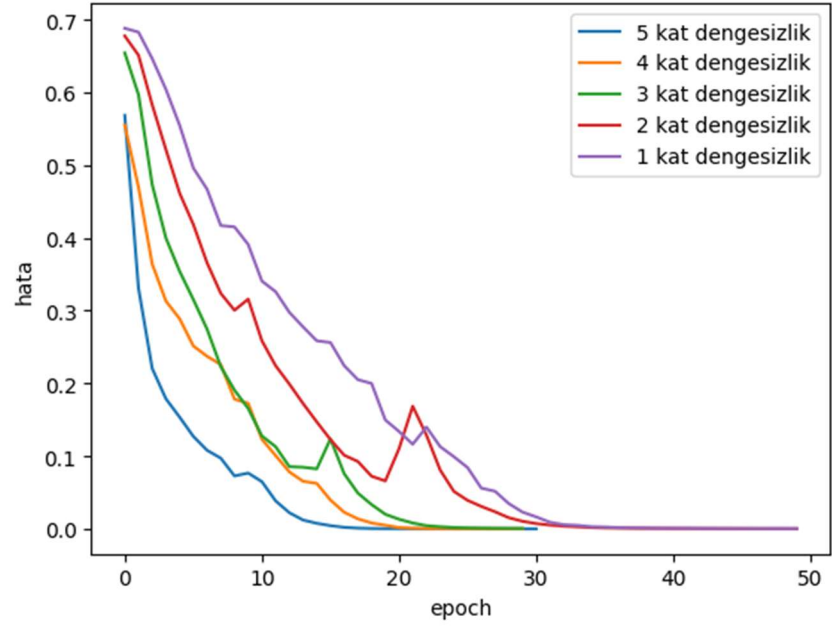
Dengesizlik	ETÖ	TTÖ	EMÖ	EMOÖ	TMÖ	TMOÖ
1 kat	1104	276	552	552	138	138
2 kat	1657	414	552	1105	138	276
3 kat	2210	552	552	1658	138	414
4 kat	3316	829	552	2764	138	691
5 kat	5846	1461	552	5294	138	1323

Tablo 4.14 ICPR2014 veri kümesi için literatürde yapılan çalışmaların sonuçları

Çalışma	F1-skoru	Kesinlik	Duyarlılık
[16]	0.8638	-	-
[85]	0.7800	-	-
[86]	0.780	-	-
[87]	0.8000	-	-
[5]	0.8697	0.9678	0.7942
[88]	0,969	1.0000	0.9400
Önerilen	0.9239	0.9241	0.9239

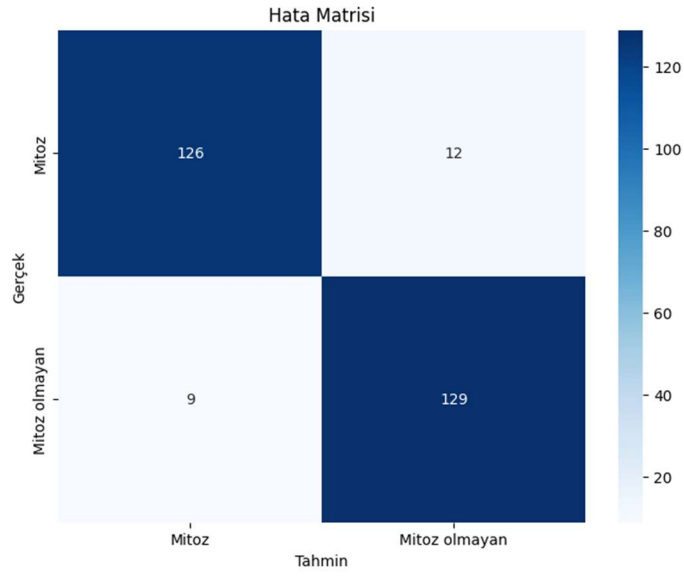


(a)

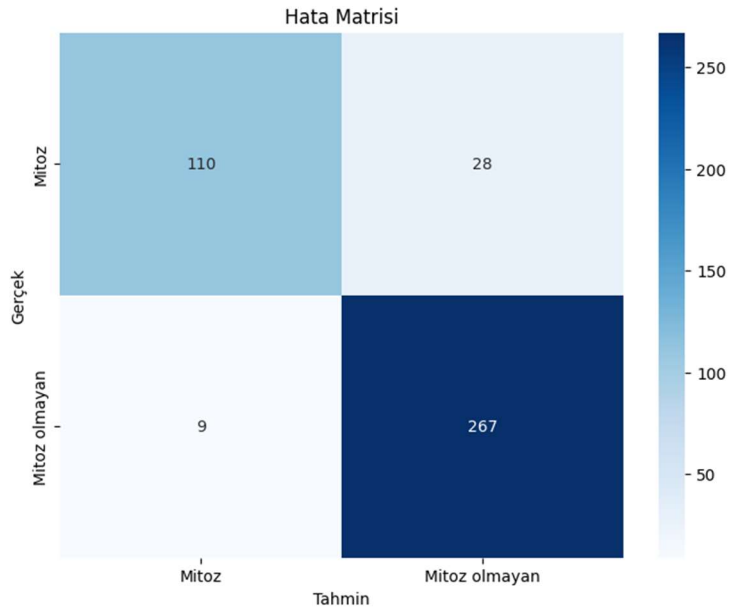


(b)

Şekil 4.3 ICPR2014 veri kümesi için denetimli öğrenme yaklaşımı ile eğitilen ResNet-50 mimarisi için sırasıyla doğruluk ve hata grafiği.

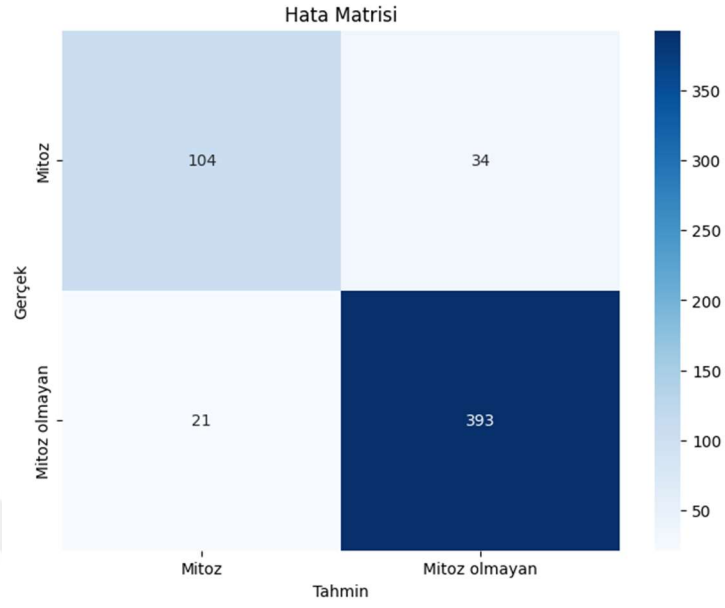


(a)

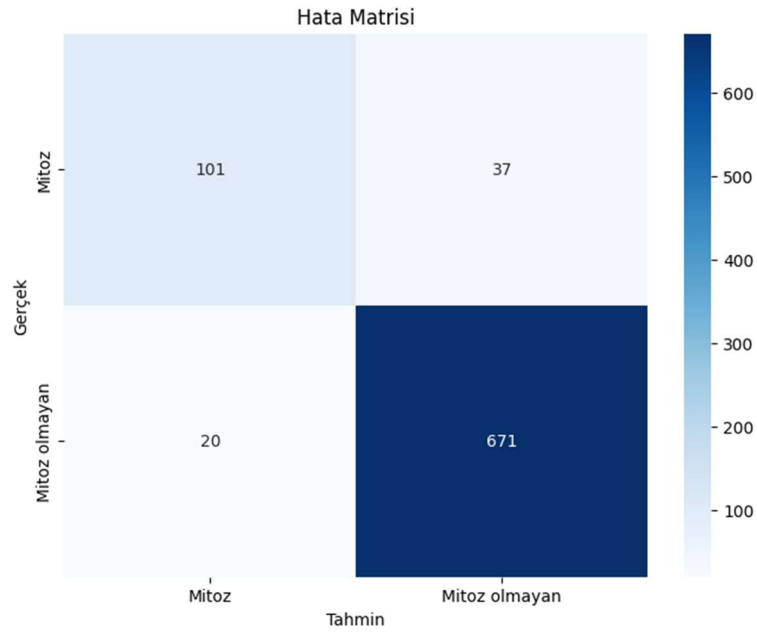


(b)

Şekil 4.4 ICPR2014 veri kümesi için en iyi modellerin hata matrisleri. (a) 1 kat dengesizlik durumu, kullanılan model ViT B/8. (b) 2 kat dengesizlik durumu, kullanılan model XCIT M/8. (c) 3 kat dengesizlik durumu, kullanılan model denetimli ResNet50. (d) 4 kat dengesizlik durumu, kullanılan model denetimli ResNet50. (e) 5 kat dengesizlik durumu, kullanılan model denetimli ResNet50.

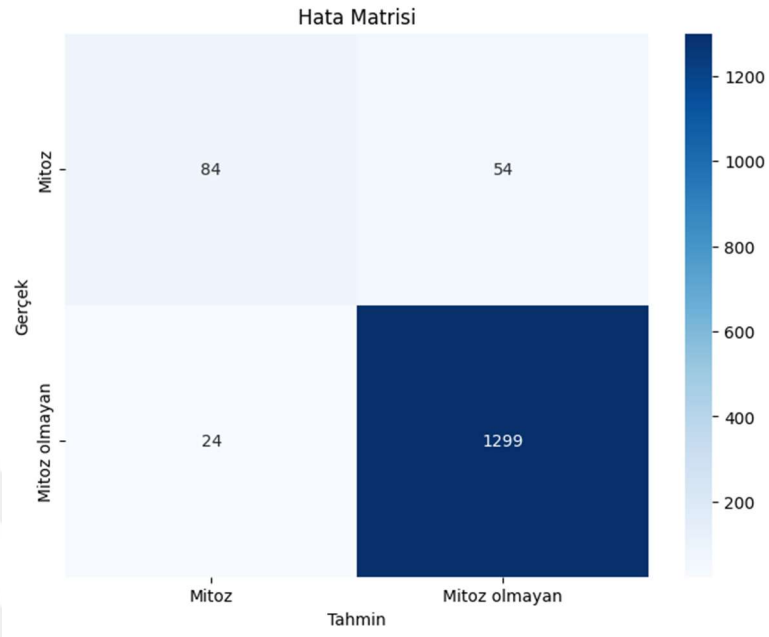


(c)



(d)

Şekil 4.4 (Devam)



(e)

Şekil 4.4 (Devam)

5. SONUÇLAR VE ÖNERİLER

5.1 Sonuçlar

Histopatolojik imgelerdeki mitozlu hücrelerin sayısı kanserin agresiflik seviyesinin belirlenmesinde önemli bir rol oynamaktadır. Bunun yanı sıra tespit edilen mitozlu hücreler, hastaya uygulanacak olan tedavi yönteminin seçilmesinde de etkin bir rol oynamaktadır. Histopatolojik görüntülerdeki mitozlu hücrelerin tespit edilmesi doktorlar açısından oldukça yorucu ve zaman alıcı bir işlemdir. Histopatolojik görüntülerdeki mitozlu hücrelerin tespit edilebilmesi için klasik görüntü işleme teknikleri veya derin öğrenme yöntemlerini kullanmak mümkündür. Derin öğrenme algoritmalarının daha gürbüz ve iyi sonuç vermesi sebebiyle mitozlu hücrelerin tespit edilebilmesi için literatürde daha çok derin öğrenme teknikleri kullanılmaktadır.

Bu çalışmada denetimli, öz denetimli öğrenme yöntemleri ve geleneksel görüntü işleme yöntemleri kullanılarak MIDOG2021, MIDOG2022 ve ICPR2014 veri kümelerinde mitozlu hücre tespiti yapılmıştır. Denetimli öğrenme yaklaşımı için ResNet-50 mimarisi kullanılırken özdenetimli öğrenme yaklaşımı için ise ViT, XCIT ve ResNet-50 modelleri kullanılmıştır. Geleneksel görüntü işleme tekniği olarak ise LBP algoritması kullanılmıştır.

Elde edilen sonuçların güvenilirliğinin sağlanması ve daha iyi hale getirilebilmesi için çalışmada farklı modeller ve bu modellerin farklı parametreler ile deneyler yapılmıştır. Yapılan deneyler sonucunda farklı modeller ve parametreler ile elde edilen başarımların birbirlerine çok yakın olduğu gözlemlenmiştir. Transfirmer ve evrişim gibi farklı mimarisinde farklı katmanlara sahip olan ViT ve ResNet gibi modellerin özdenetimli öğrenme yaklaşımı ile kullanılmaları durumunda yakın başarımlar elde edilmiştir. Bunu yanı sıra ViT ve XCIT modellerinin farklı parametreleri ile elde edilen sonuçlar da birbirine yakındır.

Yapılan bütün deneyler göz önünde bulundurulduğunda MIDOG2021 ve MIDOG2022 veri kümeleri için sırasıyla XCIT M/8 modeli 0.7833 ve denetimli ResNet-50 modeli 0.8448 en iyi f1-skoru performanslarını sağlamışlardır. ICPR2014 veri kümesi için ise veri dengesizliğinin olmadığı ve 2 kat veri dengesizliğinin olduğu durumda sırasıyla ViT B/8 modeli 0.9239 ve XCIT M/8 modeli ise 0.8956 f1-

skoru performansını sergilemiştir. Veri kümesinde bulunan dengesizliğin 5 kata kadar artması ile en iyi performansı sergileyen model denetimli öğrenme yöntemiyle eğitilen ResNet-50 mimarisi olmuştur.

Verilerin dengesiz olarak dağılması durumunda yapılan deneyler özdenetimli öğrenme yaklaşımlarının dengesizliğin 2 kattan fazla artması durumunda denetimli öğrenme modellerinden daha kötü sonuç verdiği gözlemlenmiştir. Bunun yanı sıra farklı mimarideki modellerin özdenetimli öğrenme yaklaşımları ile kullanılması sonucu elde edilen performansların verinin dengesizliğine göre düşüşü de birbirine yakın olmaktadır. Aynı deneyler denetimli öğrenme yaklaşımı ve ResNet-50 modeli kullanılarak yapıldığında denetimli öğrenme yaklaşımının veri dengesizliğine karşı çok daha gürbüz olduğu görülmektedir. Bunun yanı sıra verinin dengesizleşmesi sonucu elde edilen performans düşüşü de öz denetimli öğrenme yaklaşımına göre çok daha azdır. Geleneksel görüntü işleme yöntemi olan LBP dikkate alındığında ise performans düşüşünün ani ve derin öğrenme tabanlı yöntemlere göre çok daha fazla olduğu görülmektedir.

Denetimli ve öz denetimli öğrenme yaklaşımlarının performanslarının veri sayısı ile de ilişki içerisinde olduğu gözlemlenmektedir. MIDOG2022 ve MIDOG2021 veri kümeleri göz önünde bulundurulduğunda her ne kadar MIDOG2022 veri kümesinin içerisinde MIDOG2021'e ait veriler bulunsun da özdenetimli öğrenme yaklaşımların MIDOG2022 veri kümesi için daha kötü sonuçlar sergilemektedir. Buna sebep olan durumun MIDOG2021 veri kümesinde MIDOG2022 veri kümesine göre çok daha az örnek bulunmasının gösterilmesi mümkündür. ICPR2014 veri kümesi de göz önünde bulundurulduğunda bu çıkarım desteklenmektedir. ICPR2014 veri kümesi için veri dengesizliğinin az olması ve örnek sayısının da az olması durumunda özdenetimli öğrenme yöntemleri çok daha iyi sonuçlar verirken veri dengesizliğinin artması yani veri sayısının da artması durumunda denetimli öğrenme yaklaşımlarının daha iyi sonuç verdiği gözlemlenmiştir. Buradan yola çıkarak özellikle problemi ifade edebilecek kadar yeterli sayıda örnek bulunmaması durumunda temel ve özdenetimli öğrenme modellerinin kullanılması yapılacak olan çalışmanın başarı oranına pozitif yönde bir katkı sağlarken kullanılacak olan kaynakların daha verimli bir şekilde yönetilmesini mümkün kılacaktır.

ICPR2014 veri kümesi göz önünde bulundurulduğunda veri dengesizliğinin artışı ile beraber eğitim hata değeri daha çabuk indirgenir ve eğitim doğruluk değeri daha çabuk yükselirken test veri kümesindeki başarı oranı düşmektedir. Bu durum veri dengesizliğinin denetimli öğrenme ile eğitilen modellerin problemi öğrenmesini engellediği ve ezberlemesine sebep olduğunu göstermektedir.

MIDOG2021 ve MIDOG2022 yarışmalarında test veri kümeleri herkese açık bir şekilde yayınlanmaması sebebiyle bu çalışmada ve literatürde bulunan birçok çalışmada yayınlanan eğitim veri kümesinin bölünmesi ile test kümesi elde edilmiştir. Bölünerek elde edilen test veri kümeleri ile yarışmada kullanılan test veri kümelerinin elde ettiği sonuçlar gözlemlendiğinde yakın başarı performanslarının elde edildiği görülmüştür. Bu durum eğitim veri kümesinin bölünmesi ile elde edilen test kümesinin başarı performanslarının tutarlı olduğunu göstermektedir.

Bu çalışmada da kullanılan temel modeller çok büyük veri kümeleri ile eğitilmeleri sebebiyle tekrar bir eğitim gerçekleştirmeden kullanılmaları mümkündür. Kullanılan temel modeller tekrar MIDOG2021 ve MIDOG2022 veri kümeleri kullanılarak tekrar eğitilmedikleri için derin öğrenme modelinin eğitim aşamasında ihtiyaç duyduğu GPU gibi yüksek işlem gücü sağlayan donanımsal altyapı yatırımlarına olan ihtiyaç azaltılmıştır. Derin öğrenme modelleri yüksek işlem gücüne ihtiyaç duymalarının yanı sıra eğitim aşamasında uzun eğitim sürelerine de ihtiyaç duymaktadır. Derin öğrenme modelinin eğitim aşamasının ortadan kaldırılmasıyla beraber zamandan da tasarruf edilmiştir. Histopatolojik ve diğer birçok sağlık görüntülerinde veri gizliliği çok kritik bir noktadır. Bu sebeple histopatoloji alanında büyük ve dengeli veri kümeleri oluşturmak zordur. Veri kümesinin dengesiz ve az veriye sahip olması derin öğrenme modellerinin performansını olumsuz olarak etkilemektedir. Bu tür problemlerin çözülebilmesi için düzenleme (regularization) gibi ekstra özel tekniklerin uygulanmasına ihtiyaç duyulmaktadır. Kullanılan temel model herhangi bir eğitim aşamasına ihtiyaç duymaması sebebiyle bu tür problemlerin de önüne geçilmiştir.

5.2 Öneriler

Literatürde bu çalışmada kullanılan modeller dışında birçok farklı model bulunmaktadır. Bu modellerin birbirlerine göre farklı güçlü ve zayıf yönleri bulunmaktadır. İlerleyen çalışmalarda kullanılan modellerin yelpazesi genişletilerek performansın ne kadar yükseltilebileceğinin veya işlem maliyetinin ne kadar düşürülebileceğinin araştırılması mümkündür.

Temel modeller de dahil olmak üzere derin öğrenme modellerinin performansının artması noktasında verinin niceliği ve niteliği büyük önem arz etmektedir. Kullanılan verinin niceliğinin artırılması genel olarak modelin performansını olumlu etkilemektedir. Dosovitskiy vd. tarafından da belirtildiği üzere ViT modeli gerçek performansını ancak büyük veri kümeleri ile eğitilmesi durumunda ortaya çıkarabilmektedir. Bu çalışmada kullanılan modeller ImageNet1k veri kümesi kullanılarak ön eğitim aşması gerçekleştirilmiştir. İlerleyen çalışmalarda ImageNet21k veya daha fazla veri sayısına sahip nitelikli veri kümeleri kullanılarak kullanılan modeller DINO öğrenme yaklaşımıyla eğitilerek elde edilecek olan sonuçların değerlendirilmesi mümkündür. Bunun yanı sıra histopatoloji alanına özgü veri kümeleri de kullanılarak modelin görüntülerdeki istatistiksel özellikleri daha iyi çıkarması mümkün kılınabilir.

Özdenetimli öğrenme yöntemi temelde verinin kendisinden öğrense de literatürde MoCo [89], BYOL [90] ve VICReg [91] gibi farklı özdenetimli ve karşılaştırmalı öğrenme yaklaşımları bulunmaktadır. Farklı yaklaşımların farklı problemler için birbirlerine karşı üstünlükler sergilemesi mümkündür. Literatürde bulunan farklı özdenetimli öğrenme yaklaşımlarının histopatolojik görüntüler üzerindeki performansının değerlendirilmesi mümkündür.

Literatür çalışmasında da görüldüğü üzere derin öğrenme modelleri ile geleneksel yöntemler ile çıkarılan özneteliklerin birleştirilmesi elde edilen başarımların performanslarını artırdığı gözlemlenmiştir. Bunun yanı sıra farklı derin öğrenme modellerinden çıkarılan gömme değerlerinin de birleştirilmesi mümkündür. Farklı özdenetimli öğrenme yaklaşımları ile eğitilen aynı derin öğrenme modellerinin gömme değerlerinin veya farklı derin öğrenme modellerinin aynı özdenetimli

öğrenme yaklaşımları ile eğitilmesi sonucu elde edilen gömme değerlerini birleştirmek mümkündür. Yapılacak olan çalışmalar ile bu tür yaklaşımların histopatoloji görüntüleri üzerinde performansı nasıl etkileyeceği araştırılabilir.

Yapılan çalışmada veri kümesindeki örnek sayısı ile özdenetimli ve denetimli öğrenme modellerinin başarımı arasında bir ilişki olduğu öngörülmektedir. Bu ilişkinin kesinleştirilebilmesi ve teorik olarak sayısının belirlenebilmesi için daha fazla veri kümesi üzerinde deneylerin yapılması ve karşılaştırılması mümkündür.

Dokulardaki mitozlu olmayan hücrelerin sayısının mitozlu olan hücrelerden kat kat fazla olması sonuçları negatif yönde etkilediği gözlemlenmiştir. Bu problemin çözülebilmesi veya kullanılan modellerin veri dengesizliğine karşı olan gürbüzlüğünün artırılması noktasında araştırmaların yapılması mümkündür.

KAYNAKLAR

- [1] A. Mack, “Inattentional Blindness: Looking Without Seeing”, *Curr. Dir. Psychol. Sci.*, c. 12, sy 5, ss. 180-184, Eki. 2003, doi: 10.1111/1467-8721.01256.
- [2] P. Bankhead vd., “QuPath: Open source software for digital pathology image analysis”, *Sci. Rep.*, c. 7, sy 1, s. 16878, Ara. 2017, doi: 10.1038/s41598-017-17204-5.
- [3] M. Z. Alom, T. Aspiras, T. M. Taha, T. J. Bowen, ve V. K. Asari, “MitosisNet: End-to-End Mitotic Cell Detection by Multi-Task Learning”, *IEEE Access*, c. 8, ss. 68695-68710, 2020, doi: 10.1109/ACCESS.2020.2983995.
- [4] J. Breen, K. Zucker, N. M. Orsi, ve N. Ravikumar, “Assessing Domain Adaptation Techniques for Mitosis Detection in Multi-scanner Breast Cancer Histopathology Images”, içinde *Biomedical Image Registration, Domain Generalisation and Out-of-Distribution Analysis*, c. 13166, M. Aubreville, D. Zimmerer, ve M. Heinrich, Ed., içinde *Lecture Notes in Computer Science*, vol. 13166. , Cham: Springer International Publishing, 2022, ss. 14-22. doi: 10.1007/978-3-030-97281-3_2.
- [5] I. O. Sigirci, A. Albayrak, ve G. Bilgin, “Detection of mitotic cells in breast cancer histopathological images using deep versus handcrafted features”, *Multimed. Tools Appl.*, c. 81, sy 10, ss. 13179-13202, Nis. 2022, doi: 10.1007/s11042-021-10539-2.
- [6] A. M. Güneş, W. Van Rooij, S. Gulshad, B. Slotman, M. Dahele, ve W. Verbakel, “Impact of imperfection in medical imaging data on deep learning-based segmentation performance: An experimental study using synthesized data”, *Med. Phys.*, c. 50, sy 10, ss. 6421-6432, Eki. 2023, doi: 10.1002/mp.16437.
- [7] D. Hirahara, E. Takaya, T. Takahara, ve T. Ueda, “Effects of data count and image scaling on Deep Learning training”, *PeerJ Comput. Sci.*, c. 6, s. e312, Kas. 2020, doi: 10.7717/peerj-cs.312.
- [8] K. He, X. Zhang, S. Ren, ve J. Sun, “Deep Residual Learning for Image Recognition”. arXiv, 10 Aralık 2015. doi: 10.48550/arXiv.1512.03385.
- [9] A. Dosovitskiy vd., “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale”. arXiv, 03 Haziran 2021. doi: 10.48550/arXiv.2010.11929.
- [10] A. El-Nouby vd., “XCiT: Cross-Covariance Image Transformers”. arXiv, 18 Haziran 2021. doi: 10.48550/arXiv.2106.09681.
- [11] M. Caron vd., “Emerging Properties in Self-Supervised Vision Transformers”, program adı: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, ss. 9650-9660. Erişim: 01 Şubat 2024. [Çevrimiçi]. Erişim adresi: https://openaccess.thecvf.com/content/ICCV2021/html/Caron_Emerging_Properties_in_Self-Supervised_Vision_Transformers_ICCV_2021_paper.html
- [12] P. Faridi, H. Danyali, M. S. Helfroush, ve M. A. Jahromi, “An automatic system for cell nuclei pleomorphism segmentation in histopathological images of breast cancer”, içinde *2016 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, Philadelphia, PA, USA: IEEE, Ara. 2016, ss. 1-5. doi: 10.1109/SPMB.2016.7846861.

- [13] H. Irshad, L. Roux, ve D. Racocceanu, "Multi-channels statistical and morphological features based mitosis detection in breast cancer histopathology", içinde *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Osaka: IEEE, Tem. 2013, ss. 6091-6094. doi: 10.1109/EMBC.2013.6610942.
- [14] N. Maroof, A. Khan, S. A. Qureshi, A. U. Rehman, R. K. Khalil, ve S.-O. Shim, "Mitosis detection in breast cancer histopathology images using hybrid feature space", *Photodiagnosis Photodyn. Ther.*, c. 31, s. 101885, Eyl. 2020, doi: 10.1016/j.pdpdt.2020.101885.
- [15] S. Zakariapour, H. Jazayeri, ve M. Ezoji, "Mitosis detection in breast cancer histological images based on texture features using AdaBoost", *J. Inf. Syst. Telecommun. JIST*, c. 2, sy 18, s. 1, Tem. 2017, doi: 10.7508/jist.2017.18.003.
- [16] M. U. Rehman, S. Akhtar, M. Zakwan, ve M. H. Mahmood, "Novel architecture with selected feature vector for effective classification of mitotic and non-mitotic cells in breast cancer histology images", *Biomed. Signal Process. Control*, c. 71, s. 103212, Oca. 2022, doi: 10.1016/j.bspc.2021.103212.
- [17] C. Li, X. Wang, W. Liu, ve L. J. Latecki, "DeepMitosis: Mitosis detection via deep detection, verification and segmentation networks", *Med. Image Anal.*, c. 45, ss. 121-133, Nis. 2018, doi: 10.1016/j.media.2017.12.002.
- [18] H. Chen, X. Wang, ve P. A. Heng, "Automated mitosis detection with deep regression networks", içinde *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, Prague, Czech Republic: IEEE, Nis. 2016, ss. 1204-1207. doi: 10.1109/ISBI.2016.7493482.
- [19] T. Mahmood, M. Arsalan, M. Owais, M. B. Lee, ve K. R. Park, "Artificial Intelligence-Based Mitosis Detection in Breast Cancer Histopathology Images Using Faster R-CNN and Deep CNNs", *J. Clin. Med.*, c. 9, sy 3, s. 749, Mar. 2020, doi: 10.3390/jcm9030749.
- [20] A. Sohail, A. Khan, N. Wahab, A. Zameer, ve S. Khan, "A multi-phase deep CNN based mitosis detection framework for breast cancer histopathological images", *Sci. Rep.*, c. 11, sy 1, Art. sy 1, Mar. 2021, doi: 10.1038/s41598-021-85652-1.
- [21] M. Sebai, X. Wang, ve T. Wang, "MaskMitosis: a deep learning framework for fully supervised, weakly supervised, and unsupervised mitosis detection in histopathology images", *Med. Biol. Eng. Comput.*, c. 58, sy 7, ss. 1603-1623, Tem. 2020, doi: 10.1007/s11517-020-02175-z.
- [22] "Mitotic cell detection in histopathological images of neuroendocrine tumors using improved YOLOv5 by transformer mechanism | Signal, Image and Video Processing". Erişim: 03 Şubat 2024. [Çevrimiçi]. Erişim adresi: <https://link.springer.com/article/10.1007/s11760-023-02642-8>
- [23] A. M. Raj ve L. S. Nair, "Breast Cancer Histopathology Image Classification Using Ensemble Feature Set", içinde *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Delhi, India: IEEE, Tem. 2023, ss. 1-6. doi: 10.1109/ICCCNT56998.2023.10307803.
- [24] A. Albayrak ve G. Bilgin, "Detection of mitotic cells in histopathological images using textural features", içinde *2013 21st Signal Processing and Communications Applications Conference (SIU)*, Haspolat: IEEE, Nis. 2013, ss. 1-4. doi: 10.1109/SIU.2013.6531197.

- [25] A. Albayrak ve G. Bilgin, “Automatic cell segmentation in histopathological images via two-staged superpixel-based algorithms”, *Med. Biol. Eng. Comput.*, c. 57, sy 3, ss. 653-665, Mar. 2019, doi: 10.1007/s11517-018-1906-0.
- [26] H. Wang vd., “Cascaded ensemble of convolutional neural networks and handcrafted features for mitosis detection”, program adı: SPIE Medical Imaging, M. N. Gurcan ve A. Madabhushi, Ed., San Diego, California, USA, Mar. 2014, s. 90410B. doi: 10.1117/12.2043902.
- [27] M. Saha, C. Chakraborty, ve D. Racoceanu, “Efficient deep learning model for mitosis detection using breast histopathology images”, *Comput. Med. Imaging Graph.*, c. 64, ss. 29-40, Mar. 2018, doi: 10.1016/j.compmedimag.2017.12.001.
- [28] X. Ying, “An Overview of Overfitting and its Solutions”, *J. Phys. Conf. Ser.*, c. 1168, s. 022022, Şub. 2019, doi: 10.1088/1742-6596/1168/2/022022.
- [29] D. Bashir, G. D. Montanez, S. Sehra, P. S. Segura, ve J. Lauw, “An Information-Theoretic Perspective on Overfitting and Underfitting”, arXiv.org. Erişim: 18 Şubat 2024. [Çevrimiçi]. Erişim adresi: <https://arxiv.org/abs/2010.06076v2>
- [30] R. Wu, S. Yan, Y. Shan, Q. Dang, ve G. Sun, “Deep Image: Scaling up Image Recognition”. arXiv, 05 Temmuz 2015. doi: 10.48550/arXiv.1501.02876.
- [31] K. Stacked, G. Eilertsen, J. Unger, ve C. Lundstrom, “Measuring Domain Shift for Deep Learning in Histopathology”, *IEEE J. Biomed. Health Inform.*, c. 25, sy 2, ss. 325-336, Şub. 2021, doi: 10.1109/JBHI.2020.3032060.
- [32] “MIDOG Challenge 2021 - Grand Challenge”, grand-challenge.org. Erişim: 14 Mart 2024. [Çevrimiçi]. Erişim adresi: <https://midog2021.grand-challenge.org/>
- [33] M. Aubreville vd., “Abstract: the MIDOG Challenge 2021: Mitosis Domain Generalization in Histopathology Images”, içinde *Bildverarbeitung für die Medizin 2023*, T. M. Deserno, H. Handels, A. Maier, K. Maier-Hein, C. Palm, ve T. Tolxdorff, Ed., içinde Informatik aktuell. , Wiesbaden: Springer Fachmedien Wiesbaden, 2023, ss. 115-115. doi: 10.1007/978-3-658-41657-7_26.
- [34] M. Aubreville vd., “Mitosis domain generalization in histopathology images — The MIDOG challenge”, *Med. Image Anal.*, c. 84, s. 102699, Şub. 2023, doi: 10.1016/j.media.2022.102699.
- [35] “MITosis DOmain Generalization Challenge 2022 - Grand Challenge”, grand-challenge.org. Erişim: 14 Mart 2024. [Çevrimiçi]. Erişim adresi: <https://midog2022.grand-challenge.org/>
- [36] “MITOS-ATYPIA-14 - Grand Challenge”, grand-challenge.org. Erişim: 26 Mayıs 2024. [Çevrimiçi]. Erişim adresi: <https://mitos-atypia-14.grand-challenge.org/>
- [37] J. Watt, R. Borhani, ve A. K. Katsaggelos, “Feature Scaling via Standard Normalization”, içinde *Machine learning refined: foundations, algorithms, and applications*, Second Edition., New York: Cambridge University Press, 2020.
- [38] R. Bommasani vd., “On the Opportunities and Risks of Foundation Models”. arXiv, 12 Temmuz 2022. doi: 10.48550/arXiv.2108.07258.
- [39] M. Noroozi ve P. Favaro, “Unsupervised Learning of Visual Representations by Solving Jigsaw Puzzles”, içinde *Computer Vision – ECCV 2016*, c. 9910, B. Leibe, J. Matas, N. Sebe, ve M. Welling, Ed., içinde Lecture Notes in Computer Science, vol. 9910. , Cham: Springer International Publishing, 2016, ss. 69-84. doi: 10.1007/978-3-319-46466-4_5.

- [40] A. Newell ve J. Deng, “How Useful Is Self-Supervised Pretraining for Visual Tasks?”, program adı: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, ss. 7345-7354. Erişim: 17 Şubat 2024. [Çevrimiçi]. Erişim adresi: https://openaccess.thecvf.com/content_CVPR_2020/html/Newell_How_Useful_Is_Self-Supervised_Pretraining_for_Visual_Tasks_CVPR_2020_paper.html
- [41] H. Liu, J. Z. HaoChen, A. Gaidon, ve T. Ma, “Self-supervised Learning is More Robust to Dataset Imbalance”. arXiv, 22 Mayıs 2022. doi: 10.48550/arXiv.2110.05025.
- [42] C. Buciluă, R. Caruana, ve A. Niculescu-Mizil, “Model compression”, içinde *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, içinde KDD '06. New York, NY, USA: Association for Computing Machinery, Ağu. 2006, ss. 535-541. doi: 10.1145/1150402.1150464.
- [43] G. Hinton, O. Vinyals, ve J. Dean, “Distilling the Knowledge in a Neural Network”. arXiv, 09 Mart 2015. doi: 10.48550/arXiv.1503.02531.
- [44] T. Hua, W. Wang, Z. Xue, S. Ren, Y. Wang, ve H. Zhao, “On Feature Decorrelation in Self-Supervised Learning”. arXiv, 25 Ağustos 2021. doi: 10.48550/arXiv.2105.00470.
- [45] L. Jing, P. Vincent, Y. LeCun, ve Y. Tian, “Understanding Dimensional Collapse in Contrastive Self-supervised Learning”. arXiv, 23 Nisan 2022. doi: 10.48550/arXiv.2110.09348.
- [46] I. Sutskever, J. Martens, G. Dahl, ve G. Hinton, “On the importance of initialization and momentum in deep learning”, içinde *Proceedings of the 30th International Conference on Machine Learning*, PMLR, May. 2013, ss. 1139-1147. Erişim: 24 Mart 2024. [Çevrimiçi]. Erişim adresi: <https://proceedings.mlr.press/v28/sutskever13.html>
- [47] K. He, H. Fan, Y. Wu, S. Xie, ve R. Girshick, “Momentum Contrast for Unsupervised Visual Representation Learning”. arXiv, 23 Mart 2020. doi: 10.48550/arXiv.1911.05722.
- [48] A. Vaswani vd., “Attention Is All You Need”. arXiv, 01 Ağustos 2023. doi: 10.48550/arXiv.1706.03762.
- [49] A. Kirillov vd., “Segment Anything”. arXiv, 05 Nisan 2023. doi: 10.48550/arXiv.2304.02643.
- [50] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, ve S. Zagoruyko, “End-to-End Object Detection with Transformers”. arXiv, 28 Mayıs 2020. doi: 10.48550/arXiv.2005.12872.
- [51] K. He, X. Chen, S. Xie, Y. Li, P. Dollár, ve R. Girshick, “Masked Autoencoders Are Scalable Vision Learners”. arXiv, 19 Aralık 2021. doi: 10.48550/arXiv.2111.06377.
- [52] A. Steiner, A. Kolesnikov, X. Zhai, R. Wightman, J. Uszkoreit, ve L. Beyer, “How to train your ViT? Data, Augmentation, and Regularization in Vision Transformers”. arXiv, 23 Haziran 2022. doi: 10.48550/arXiv.2106.10270.
- [53] S. H. Lee, S. Lee, ve B. C. Song, “Vision Transformer for Small-Size Datasets”. arXiv, 26 Aralık 2021. Erişim: 21 Mart 2024. [Çevrimiçi]. Erişim adresi: <http://arxiv.org/abs/2112.13492>
- [54] P. K. A. Vasu, J. Gabriel, J. Zhu, O. Tuzel, ve A. Ranjan, “FastViT: A Fast Hybrid Vision Transformer using Structural Reparameterization”. arXiv, 17

- Ağustos 2023. Erişim: 25 Şubat 2024. [Çevrimiçi]. Erişim adresi: <http://arxiv.org/abs/2303.14189>
- [55] I. Han, R. Jayaram, A. Karbasi, V. Mirrokni, D. P. Woodruff, ve A. Zandieh, “HyperAttention: Long-context Attention in Near-Linear Time”. arXiv, 01 Aralık 2023. doi: 10.48550/arXiv.2310.05869.
 - [56] T. Dao, D. Y. Fu, S. Ermon, A. Rudra, ve C. Ré, “FlashAttention: Fast and Memory-Efficient Exact Attention with IO-Awareness”. arXiv, 23 Haziran 2022. doi: 10.48550/arXiv.2205.14135.
 - [57] Z. Liu *vd.*, “Swin Transformer: Hierarchical Vision Transformer using Shifted Windows”. arXiv, 17 Ağustos 2021. doi: 10.48550/arXiv.2103.14030.
 - [58] R. Shao, Z. Shi, J. Yi, P.-Y. Chen, ve C.-J. Hsieh, “On the Adversarial Robustness of Vision Transformers”, sy arXiv:2103.15670. arXiv, 02 Kasım 2022. Erişim: 26 Temmuz 2023. [Çevrimiçi]. Erişim adresi: <http://arxiv.org/abs/2103.15670>
 - [59] C. Liu *vd.*, “A Comprehensive Study on Robustness of Image Classification Models: Benchmarking and Rethinking”. arXiv, 27 Şubat 2023. doi: 10.48550/arXiv.2302.14301.
 - [60] Z. Xu, R. Liu, S. Yang, Z. Chai, ve C. Yuan, “Learning Imbalanced Data with Vision Transformers”. arXiv, 08 Mart 2023. doi: 10.48550/arXiv.2212.02015.
 - [61] J. Yung *vd.*, “SI-Score: An image dataset for fine-grained analysis of robustness to object location, rotation and size”, sy arXiv:2104.04191. arXiv, 09 Nisan 2021. Erişim: 26 Temmuz 2023. [Çevrimiçi]. Erişim adresi: <http://arxiv.org/abs/2104.04191>
 - [62] J. L. Ba, J. R. Kiros, ve G. E. Hinton, “Layer Normalization”. arXiv, 21 Temmuz 2016. doi: 10.48550/arXiv.1607.06450.
 - [63] M. Lee, “Mathematical Analysis and Performance Evaluation of the GELU Activation Function in Deep Learning”, *J. Math.*, c. 2023, s. 4229924, Ağu. 2023, doi: 10.1155/2023/4229924.
 - [64] S. Ioffe ve C. Szegedy, “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift”, içinde *Proceedings of the 32nd International Conference on Machine Learning*, PMLR, Haz. 2015, ss. 448-456. Erişim: 30 Mart 2024. [Çevrimiçi]. Erişim adresi: <https://proceedings.mlr.press/v37/ioffe15.html>
 - [65] D. Hendrycks ve K. Gimpel, “Gaussian Error Linear Units (GELUs)”. arXiv, 05 Haziran 2023. doi: 10.48550/arXiv.1606.08415.
 - [66] V. Nair ve G. E. Hinton, “Rectified linear units improve restricted boltzmann machines”, içinde *Proceedings of the 27th International Conference on International Conference on Machine Learning*, içinde ICML’10. Madison, WI, USA: Omnipress, Haz. 2010, ss. 807-814.
 - [67] T. Chen ve C. Guestrin, “XGBoost: A Scalable Tree Boosting System”, içinde *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Ağu. 2016, ss. 785-794. doi: 10.1145/2939672.2939785.
 - [68] C. Cortes ve V. Vapnik, “Support-vector networks”, *Mach. Learn.*, c. 20, sy 3, ss. 273-297, Eyl. 1995, doi: 10.1007/BF00994018.
 - [69] E. Bisong, “Google Colaboratory”, içinde *Building Machine Learning and Deep Learning Models on Google Cloud Platform*, Berkeley, CA: Apress, 2019, ss. 59-64. doi: 10.1007/978-1-4842-4470-8_7.

- [70] A. Paszke vd., “PyTorch: An Imperative Style, High-Performance Deep Learning Library”. arXiv, 03 Aralık 2019. doi: 10.48550/arXiv.1912.01703.
- [71] R. Wightman vd., “rwightman/pytorch-image-models: v0.8.10dev0 Release”. [object Object], 07 Şubat 2023. doi: 10.5281/ZENODO.4414861.
- [72] G. Roy vd., “Robust Mitosis Detection Using a Cascade Mask-RCNN Approach With Domain-Specific Residual Cycle-GAN Data Augmentation”. arXiv, 28 Eylül 2021. Erişim: 14 Mart 2024. [Çevrimiçi]. Erişim adresi: <http://arxiv.org/abs/2109.01878>
- [73] J. Liang vd., “Detecting Mitosis against Domain Shift using a Fused Detector and Deep Ensemble Classification Model for MIDOG Challenge”. arXiv, 31 Ağustos 2021. Erişim: 14 Mart 2024. [Çevrimiçi]. Erişim adresi: <http://arxiv.org/abs/2108.13983>
- [74] X. Long, Y. Cheng, X. Mu, L. Liu, ve J. Liu, “Domain Adaptive Cascade R-CNN for MITosis DOmain Generalization (MIDOG) Challenge”. arXiv, 29 Eylül 2021. Erişim: 14 Mart 2024. [Çevrimiçi]. Erişim adresi: <http://arxiv.org/abs/2109.00965>
- [75] S. Razavi, F. Dambandkhameneh, D. Androustos, S. Done, ve A. Khademi, “Cascade RCNN for MIDOG Challenge”. arXiv, 26 Eylül 2021. Erişim: 14 Mart 2024. [Çevrimiçi]. Erişim adresi: <http://arxiv.org/abs/2109.01085>
- [76] M. W. Lafarge ve V. H. Koelzer, “Rotation Invariance and Extensive Data Augmentation: A Strategy for the MITosis DOmain Generalization (MIDOG) Challenge”, içinde *Biomedical Image Registration, Domain Generalisation and Out-of-Distribution Analysis*, c. 13166, M. Aubreville, D. Zimmerer, ve M. Heinrich, Ed., içinde Lecture Notes in Computer Science, vol. 13166. , Cham: Springer International Publishing, 2022, ss. 62-67. doi: 10.1007/978-3-030-97281-3_9.
- [77] H. Gu vd., “Detecting Mitoses with a Convolutional Neural Network for MIDOG 2022 Challenge”, içinde *Mitosis Domain Generalization and Diabetic Retinopathy Analysis*, c. 13597, B. Sheng ve M. Aubreville, Ed., içinde Lecture Notes in Computer Science, vol. 13597. , Cham: Springer Nature Switzerland, 2023, ss. 211-216. doi: 10.1007/978-3-031-33658-4_21.
- [78] R. Nateghi ve F. Pourakpour, “Two-Step Domain Adaptation for Mitotic Cell Detection in Histopathology Images”, içinde *Biomedical Image Registration, Domain Generalisation and Out-of-Distribution Analysis*, c. 13166, M. Aubreville, D. Zimmerer, ve M. Heinrich, Ed., içinde Lecture Notes in Computer Science, vol. 13166. , Cham: Springer International Publishing, 2022, ss. 32-39. doi: 10.1007/978-3-030-97281-3_4.
- [79] M. Halmes, H. Heuberger, ve S. Berlemont, “Deep Learning-based mitosis detection in breast cancer histologic samples”. arXiv, 02 Eylül 2021. Erişim: 14 Mart 2024. [Çevrimiçi]. Erişim adresi: <http://arxiv.org/abs/2109.00816>
- [80] H. Anand, A. K. Rajan, G. Santhosh, ve L. S. Nair, “Mitotic Instance Detection in Stain Normalized Histopathological Images using Faster R-CNN”, içinde *2022 7th International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India: IEEE, Haz. 2022, ss. 1310-1313. doi: 10.1109/ICCES54183.2022.9835762.
- [81] M. Hussain, R. Gangnani, ve S. Kadiyala, “Robust Multi-Domain Mitosis Detection”. arXiv, 13 Eylül 2021. Erişim: 14 Mart 2024. [Çevrimiçi]. Erişim adresi: <http://arxiv.org/abs/2109.15092>

- [82] S. Kotte vd., “A Deep Learning Based Ensemble Model for Generalized Mitosis Detection in H &E Stained Whole Slide Images”, içinde *Mitosis Domain Generalization and Diabetic Retinopathy Analysis*, c. 13597, B. Sheng ve M. Aubreville, Ed., içinde Lecture Notes in Computer Science, vol. 13597. , Cham: Springer Nature Switzerland, 2023, ss. 221-225. doi: 10.1007/978-3-031-33658-4_23.
- [83] M. W. Lafarge ve V. H. Koelzer, “Fine-Grained Hard-Negative Mining: Generalizing Mitosis Detection with a Fifth of the MIDOG 2022 Dataset”, içinde *Mitosis Domain Generalization and Diabetic Retinopathy Analysis*, c. 13597, B. Sheng ve M. Aubreville, Ed., içinde Lecture Notes in Computer Science, vol. 13597. , Cham: Springer Nature Switzerland, 2023, ss. 226-233. doi: 10.1007/978-3-031-33658-4_24.
- [84] J. Annuscheit ve C. Krumnow, “Radial Prediction Domain Adaption Classifier for the MIDOG 2022 Challenge”, içinde *Mitosis Domain Generalization and Diabetic Retinopathy Analysis*, c. 13597, B. Sheng ve M. Aubreville, Ed., içinde Lecture Notes in Computer Science, vol. 13597. , Cham: Springer Nature Switzerland, 2023, ss. 206-210. doi: 10.1007/978-3-031-33658-4_20.
- [85] A. Paul, “Intravenous Regional Anaesthesia”, içinde *Regional Anaesthesia*, Jaypee Brothers Medical Publishers (P) Ltd., 2014, ss. 85-85. doi: 10.5005/jp/books/12372_8.
- [86] N. A. Aloraidi, K. Sirinukunwattana, A. M. Khan, ve N. M. Rajpoot, “On generating cell exemplars for detection of mitotic cells in breast cancer histopathology images”, içinde *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Chicago, IL: IEEE, Ağu. 2014, ss. 3370-3373. doi: 10.1109/EMBC.2014.6944345.
- [87] A. Paul, D. P. Mukherjee, P. Das, A. Gangopadhyay, A. R. Chintha, ve S. Kundu, “Improved Random Forest for Classification”, *IEEE Trans. Image Process.*, c. 27, sy 8, ss. 4012-4024, Ağu. 2018, doi: 10.1109/TIP.2018.2834830.
- [88] A. Albayrak ve G. Bilgin, “Mitosis detection using convolutional neural network based features”, içinde *2016 IEEE 17th International Symposium on Computational Intelligence and Informatics (CINTI)*, Budapest, Hungary: IEEE, Kas. 2016, ss. 000335-000340. doi: 10.1109/CINTI.2016.7846429.
- [89] X. Chen, H. Fan, R. Girshick, ve K. He, “Improved Baselines with Momentum Contrastive Learning”. arXiv, 09 Mart 2020. doi: 10.48550/arXiv.2003.04297.
- [90] J.-B. Grill vd., “Bootstrap your own latent: A new approach to self-supervised Learning”. arXiv, 10 Eylül 2020. doi: 10.48550/arXiv.2006.07733.
- [91] A. Bardes, J. Ponce, ve Y. LeCun, “VICReg: Variance-Invariance-Covariance Regularization for Self-Supervised Learning”. arXiv, 28 Ocak 2022. doi: 10.48550/arXiv.2105.04906.

ÖZGEÇMİŞ

Kişisel Bilgiler

Soyad, Ad

ARSLANOĞLU, Muhammed Cihad

Web sayfası

(Research Gate, Academia, vs.)

Eğitim Bilgileri

Derece	Kurum	Mezuniyet Yılı
Yüksek Lisans	Dicle Üniversitesi	2024
Lisans	Dicle Üniversitesi	2022
Lise	Nafiye Ömer Şevki Cizrelioğlu Meslek Lisesi	2017

İş Deneyimi

Dönem (Yıl)	Şirket, Kurum	Görev
2022 – Devam ediyor	Dicle Elektrik Dağıtım	Büyük Veri ve Yapay Zeka Uzmanı

Yabancı Dil

İngilizce

Yayınlar

1. M. C. ARSLANOĞLU, A. ALBAYRAK, and H. ACAR, “Facial Expression Recognition with Vision Transformer,” in Facial Expression Recognition with Vision Transformer Dicle University 1st International Symposium on Graduate Researches In Science, Jan. 2023.
2. M. C. ARSLANOĞLU, H. ACAR, and A. ALBAYRAK, “Analyzing Performance of Deep Learning Methods on Unbalanced Medical Dataset,” in Dicle University 2nd International Symposium on Graduate Researches In Science, May 2024.

Özel İlgiler

DİCLE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
TEZ BENZERLİK BİLDİRİMİ FORMU

Öğrencinin Adı, Soyadı	Muhammed Cihad ARSLANOĞLU
Öğrenci No	22805004
Ana Bilim Dalı	Elektrik-Elektronik Mühendisliği Anabilim Dalı
Program Türü	Proje ☐ Yüksek Lisans ☒ Doktora ☐
Tez Danışmanı (Ünvanı, Adı, Soyadı)	Dr. Öğr. Üyesi Hüseyin ACAR
(Varsa) II. Tez Danışmanı (Ünvanı, Adı, Soyadı)	Doç. Dr. Abdülkadir ALBAYRAK
Tez Başlığı	Vision Transformer Tabanlı Öz Denetimli Öğrenme ile Histopatolojik Görüntülerde Mitoz Hücrelerinin Tespiti
RAPOR BİLGİLERİ	
Raporlama Aşaması	Tez Savunma Sınavı Sonrası
Sayfa Sayısı	86
Raporlama Tarihi	30.06.2024
Benzerlik Oranı (%)	4

Yukarıda bilgileri verilen tez çalışmamın toplam 79 sayfalık kısmına ilişkin, 30/06/2024 tarihinde şahsım/tez danışmanım tarafından Turnitin isimli intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan intihal raporuna göre, tezimin benzerlik oranı %4 olarak tespit edilmiştir.

Uygulanan filtrelemeler:

- ☐Başlangıç Bölümleri (Kabul ve Onay sayfası, Teşekkür sayfası, Özet/Abstract) hariç
☒Kaynaklar hariç
☐Alıntılar hariç/dâhil
☐Diğer (Açıklayınız)

Tezimin benzerlik oranı, Dicle Üniversitesi Fen Bilimleri Enstitüsü İntihal Raporu Uygulama Esaslarında belirtilen üst sınır benzerlik oranını aşmamaktadır. Tez benzerlik oranı üst sınır benzerlik oranının altında olsa dahi aksinin tespit edilmesi durumunda her türlü yasal sorumluluğu kabul ettiğimi ve hukuki sonuçlarına razı olduğumu bildirir, gereğini arz ederim.

Öğrencinin Adı, Soyadı: Muhammed Cihad ARSLANOĞLU

Tarih: 30.06.2024

İmza:

1. Danışman Adı, Soyadı: Dr. Öğr. Üyesi Hüseyin ACAR

İmza:

Tarih: 03.07.2024

2. Danışman Adı, Soyadı: Doç. Dr. Abdülkadir ALBAYRAK

İmza:

Tarih: 03.07.2024

Ana Bilim Dalı Başkanı Adı, Soyadı: Doç. Dr. Cafer BUDAK

İmza:

Tarih: 03.07.2024
