

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ



**SAYISAL VERİ VE GÖĞÜS RÖNTGEN GÖRÜNTÜLERİNDEN
DERİN ÖĞRENME YAKLAŞIMLARI İLE PNÖMONİ TESPİTİ**

Zehra KADİROĞLU

Doktora Tezi

ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

Elektrik Elektronik Mühendisliği Teknolojileri Programı

Elektrik Tesisleri Bilim Dalı

ŞUBAT 2024

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Elektrik Elektronik Mühendisliği Anabilim Dalı
Elektrik Elektronik Mühendisliği Teknolojileri Programı

Doktora Tezi

SAYISAL VERİ VE GÖĞÜS RÖNTGEN GÖRÜNTÜLERİNDEN DERİN ÖĞRENME YAKLAŞIMLARI İLE PNÖMONİ TESPİTİ

Tez Yazarı

Zehra KADİROĞLU

Danışman

Prof. Dr. Hanifi GÜLDEMİR

İkinci Danışman

Prof. Dr. Abdurrahman ŞENYİĞİT

ŞUBAT 2024

ELAZIĞ

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Elektrik Elektronik Mühendisliği Anabilim Dalı
Elektrik Elektronik Mühendisliği Teknolojileri Programı

Doktora Tezi

Başlığı:	Sayısal Veri ve Göğüs Röntgen Görüntülerinden Derin Öğrenme Yaklaşımları ile Pnömoni Tespiti
Yazarı:	Zehra KADİROĞLU
İlk Teslim Tarihi:	28.12.2023
Savunma Tarihi:	09.02.2024

TEZ ONAYI

Fırat Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına göre hazırlanan bu tez aşağıda imzaları bulunan jüri üyeleri tarafından değerlendirilmiş ve akademik dinleyicilere açık yapılan savunma sonucunda OYBİRLİĞİ ile kabul edilmiştir.

Danışman:	Prof. Dr. Hanifi GÜLDEMİR Fırat Üniversitesi, Teknoloji Fakültesi	İmza Onayladım
Başkan:	Prof. Dr. Abdulkadir ŞENGÜR Fırat Üniversitesi, Teknoloji Fakültesi	Onayladım
Üye:	Doç. Dr. Yaman AKBULUT Fırat Üniversitesi, Teknoloji Fakültesi	Onayladım
Üye:	Doç. Dr. Muzaffer ASLAN Bingöl Üniversitesi, Mühendislik-Mimarlık Fakültesi	Onayladım
Üye:	Doç. Dr. Ömer Faruk ALÇİN İnönü Üniversitesi, Mühendislik Fakültesi	Onayladım

Bu tez, Enstitü Yönetim Kurulunun/...../20..... tarihli toplantısında tescillenmiştir.

İmza

Prof. Dr. Burhan ERGEN
Enstitü Müdürü

BEYAN

Fırat Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygun olarak hazırladığım “Sayısal Veri ve Göğüs Röntgen Görüntülerinden Derin Öğrenme Yaklaşımları ile Pnömoni Tespiti” Başlıklı Doktora Tezimin içindeki bütün bilgilerin doğru olduğunu, bilgilerin üretilmesi ve sunulmasında bilimsel etik kurallarına uygun davrandığımı, kullandığım bütün kaynakları atıf yaparak belirttiğimi, maddi ve manevi desteği olan tüm kurum/kuruluş ve kişileri belirttiğimi, burada sunduğum veri ve bilgileri unvan almak amacıyla daha önce hiçbir şekilde kullanmadığımı beyan ederim.

09.02.2024

Zehra KADİROĞLU



ÖNSÖZ

Kliniklerde doktorlar pnömoniye ilişkin teşhisi, hastanın semptomlarına, fizik muayeneye, laboratuvar testleri, görüntü verileri (bilgisayarlı tomografi, göğüs röntgeni, MRI) vb. dahil olmak üzere çeşitli hastane verilerine dayanarak yapmaktadır. Hastaların semptomları, belirtileri, görüntü verileri (röntgen veya bilgisayarlı tomografi bulguları) ve biyosinyaller (kalp atış hızı, vücut sıcaklığı vb.) gibi çok sayıda yapılandırılmamış verinin eklenmesi daha iyi ve doğru tahmin modelleriyle sonuçlanabilir. Ancak bu yapılandırılmamış verileri kullanmak zordur ve bu verileri işlemek için çok fazla insan gücü ve bilgi işlem gücü gerekir. Derin öğrenme tabanlı yaklaşımlar, hızlı ve doğru tanı sağlayarak, ekipman ve uzman talebini azaltarak, doktorlara ikinci bir görüş sağlayarak sınırlı tıbbi kaynaklara ve personele sahip gelişmekte olan ülkelerde bu problemlerin çözümünde fayda sağlamaktadır. Bu tez çalışmasında pnömoninin çok yönlü teşhisini iyileştirmek için derin öğrenme tabanlı yaklaşımlar kullanılmıştır.

Bu tez çalışmasında ilgisini ve desteğini benden esirgemeyen değerli hocam ve tez danışmanım Prof. Dr. Hanifi GÜLDEMİR'e, verileri anlayıp yorumlamamda yardımcı olan ve çalışmaya olan katkılarından dolayı tez danışmanım Prof. Dr. Abdurrahman ŞENYİĞİT'e sonsuz teşekkürlerimi sunarım. Tez çalışmaları boyunca bilgilerini paylaşan, bu süreçte tecrübeleriyle yol gösterici olan ve değerli fikirleriyle çalışma sürecini zenginleştiren Prof. Dr. Abdulkadir ŞENGÜR'e teşekkürü bir borç bilirim.

Tez sürecinde beni sürekli destekleyen ve yol gösteren başta babam Prof. Dr. Ali Kemal KADİROĞLU'na, ablam Dr. Öğr. Üyesi Berfin KADİROĞLU'na ve her zaman ve her şartta yanımda olarak beni motive eden anneme ve aileme teşekkürlerimi sunarım.

Tez çalışmaları sürecinde bilgileriyle ufku açan ve yapacağım diğer çalışmalarda da desteğini esirgemeyeceğini düşündüğüm Doç. Dr. Erkan DENİZ'e ve Öğr. Gör. Mazhar KAYAOĞLU'na teşekkürlerimi sunarım.

Bu tez çalışması, Dicle Üniversitesi Tıp Fakültesi Girişimsel Olmayan Klinik Araştırmalar Etik Kurulunun 13/10/2021 tarih ve 421 sayılı izni ile gerçekleştirilmiştir. Dicle Üniversitesi Tıp Fakültesi Göğüs Hastalıkları ve Tüberküloz Ana Bilim Dalı'na çalışmaya veri katkılarından dolayı teşekkür ederim.

Zehra KADİROĞLU

ELAĞIĞ, 2024

İÇİNDEKİLER

	Sayfa
ÖNSÖZ.....	iv
İÇİNDEKİLER	v
ÖZET	vii
ABSTRACT	viii
ŞEKİLLER LİSTESİ	ix
TABLolar LİSTESİ	xi
EKLER LİSTESİ	xii
SİMGELER	xiii
KISALTMALAR	xiv
1. GİRİŞ	1
2. MAKİNE ÖĞRENMESİ.....	10
2.1. K En Yakın Komşu (KNN)	12
2.2. Destek Vektör Makineleri (DVM).....	13
2.3. Rastgele Orman Algoritması (RO)	14
3. DERİN ÖĞRENME	16
3.1. Özellik Çıkarımı	18
3.2. Transfer Öğrenimi	19
3.3. Uçtan Uca Öğrenme	20
3.4. Evrişimsel Sinir Ağları (ESA)	21
3.4.1. Giriş Katmanı	23
3.4.2. Evrişim Katmanı	23
3.4.3. Toplu Normalizasyon Katmanı	24
3.4.4. Etkinleştirme (Aktivasyon) katmanı.....	25
3.4.5. Havuzlama Katmanı	26
3.4.6. Tam Bağlantılı Katman	27
3.4.7. DropOut Katmanı.....	28
3.4.8. Sınıflandırma Katmanı	29
3.5. Kullanılan Derin Öğrenme Algoritmaları.....	29
3.5.1. AlexNet	29
3.5.2. SqueezeNet.....	30
3.5.3. VGG16 ve VGG19.....	32
3.5.4. ResNet50	32
3.5.5. DenseNet201	34
3.5.6. MobileNetV2.....	35
3.5.7. ShuffleNet	36
3.5.8. NasNetMobile	37
3.5.9. Darknet53.....	38
3.5.10. Uzun Kısa Süreli Bellek (UKSB).....	39
3.5.11. Otomatik Kodlayıcı	41
3.5.12. Görüntü Dönüştürücü (Vision Transformer / ViT)	43
3.5.13. MLP-mixer	44
3.5.14. FNet.....	45

3.5.15. gMLP	46
4. MATERYAL VE METOT	48
4.1. Özel Veri Tabanı	48
4.2. Genel Veri Tabanı	50
4.3. Geliştirilen Pnömoni Tespit Modeli	50
4.4. Model Değerlendirme/Doğrulama Yöntemleri.....	52
4.4.1. K-Katlı Çapraz Doğrulama (K-Fold Cross Validation):	52
4.4.2. Ayırarak çapraz doğrulama (Hold-out Cross Validation)	52
4.5. Değerlendirme Metrikleri	53
4.5.1. Kesinlik	54
4.5.2. Duyarlılık	54
4.5.3. Doğruluk	55
4.5.4. F1 skor.....	55
4.6. İstatistiksel Yöntemler	55
5. BULGULAR VE TARTIŞMA	56
5.1. Göğüs röntgen görüntülerinde pnömoni tespiti için derin öğrenme modellerinin incelenmesi	56
5.2. Göğüs röntgen görüntülerinde pnömoni tespiti için transformer modellerinin incelenmesi ve ESA mimarisi ile karşılaştırılması	69
5.3. Sayısal Verilerden Derin Öğrenme Yaklaşımları ile Pnömoni Tespiti.....	76
6. SONUÇLAR.....	83
ÖNERİLER	86
KAYNAKLAR.....	87
EKLER	98
ÖZGEÇMİŞ	

ÖZET

Sayısal Veri ve Göğüs Röntgen Görüntülerinden Derin Öğrenme Yaklaşımları ile Pnömoni Tespiti

Zehra KADİROĞLU

Doktora Tezi

FIRAT ÜNİVERSİTESİ

Fen Bilimleri Enstitüsü

Elektrik Elektronik Mühendisliği Anabilim Dalı
Elektrik Elektronik Mühendisliği Teknolojileri Programı
Şubat 2024, Sayfa: xiv + 97

Pnömoni, dünya çapında tüm yaş gruplarında yüksek morbidite ve mortaliteye sahip akciğer enfeksiyonudur. Pnömoniye bağlı ölümleri azaltmak ve iyileşme oranlarını artırmak için hastalığın erken evrelerde etkili bir şekilde tespit edilmesi çok önemlidir. Pnömoniyi tanımlamak için fizik muayene, laboratuvar testleri, klinik semptomlar ve çeşitli tıbbi görüntüleme yöntemleri gibi teknikler kullanılmaktadır. Diğer akciğer hastalıklarının pnömoniye benzer semptomlar göstermesi ve görüntüleme bulgularının spesifik olmaması nedeniyle pnömoni tanısının belirlenmesi zaman alıcı ve hataya açıktır. Hastalıkların erken teşhisinde uzmanlara yardımcı olmak için derin öğrenme tabanlı yöntemlerin benimsenmesi, pnömoni teşhisi de dahil tıbbi uygulamalarda doğru ve verimli tanı yöntemi olarak etkili sonuçlar elde etmektedir. Bu tez çalışmasının amacı derin öğrenme metotları kullanılarak sayısal veriler ve göğüs röntgen (CXR) görüntülerinden pnömoni hastalığını teşhis etmektir. 2000 kişiye ait (1000 pnömoni, 1000 sağlıklı) demografik özellikler, semptomlar ve laboratuvar test sonuç bilgilerinden oluşturulan sayısal tıbbi veri kümesine makine öğrenimi ve derin öğrenme algoritmaları uygulanmıştır. Aynı kişilere ait CXR görüntüleri de derin öğrenme tabanlı yöntemlerle sınıflandırılmıştır. Sayısal verilerin ve CXR görüntülerinin sınıflandırma sonuçları karşılaştırılmıştır. Çalışmanın verileri Dicle Üniversitesi Tıp Fakültesi Göğüs Hastalıkları ve Tüberküloz kliniği ile yoğun bakım ünitesi ve göğüs polikliniğine başvuran hastaların dosyalarının retrospektif olarak taranmasıyla oluşturulmuştur. Bu çalışmanın amaçlarından biri de pahalı ekipman ve yüksek eğitilmiş klinisyen eksikliğinin olduğu yerlerde uzmanlara yardımcı olacak derin öğrenme tabanlı bilgisayar destekli otomatik tespit sistemiyle pnömoninin çok yönlü teşhisini geliştirmek ve iyileştirmektir. Bu tez çalışması kapsamında elde edilen sonuçlar diğer çalışmalarla karşılaştırıldığında önerilen yöntemlerin mevcut literatüre katkısı olduğu ve iyi performans gösterdiği tespit edilmiştir.

Anahtar Kelimeler: Derin Öğrenme, Pnömoni Tespiti, Göğüs Röntgen Görüntüleme, Sayısal Veriler

ABSTRACT

Pneumonia Detection with Deep Learning Approaches from Numerical Data and Chest X-Ray Images

Zehra KADIROĞLU

Ph.D. Thesis

FIRAT UNIVERSITY

Graduate School of Natural and Applied Sciences

Department of Electrical and Electronics Engineering

Electrical and Electronics Engineering Technologies

February 2024, Pages: xiv + 97

Pneumonia is a lung infection with high morbidity and mortality in all age groups worldwide. Effective detection of the disease in the early stages is crucial to reduce pneumonia-related deaths and increase cure rates. Techniques such as physical examination, laboratory tests, clinical symptoms and various medical imaging methods are used to identify pneumonia. Because other lung diseases present similar symptoms to pneumonia and imaging findings are non-specific, the diagnosis of pneumonia is time-consuming and error-prone. The adoption of deep learning-based methods to assist experts in the early diagnosis of diseases is achieving effective results as an accurate and efficient diagnostic method in medical applications, including pneumonia diagnosis. The aim of this thesis is to diagnose pneumonia from numerical data and chest X-ray (CXR) images using deep learning methods. Machine learning and deep learning algorithms were applied to a numerical medical dataset of 2000 individuals (1000 pneumonia, 1000 healthy) with demographic characteristics, symptoms and laboratory test results. CXR images of the same individuals were also classified using deep learning-based methods. The classification results of the numerical data and CXR images were compared. The data of the study were collected by retrospectively reviewing the files of patients admitted to the Chest Diseases and Tuberculosis clinic, intensive care unit and chest outpatient clinic of Dicle University Faculty of Medicine. One of the aims of this study is to develop and improve the multidimensional diagnosis of pneumonia with a deep learning-based computer-aided automated detection system that will help experts where there is a lack of expensive equipment and highly trained clinicians. When the results obtained in this thesis are compared with other studies, it is found that the proposed methods contribute to the existing literature and perform well.

Keywords: Deep Learning, Pneumonia Detection, Chest X-ray Imaging, Numerical Data

ŞEKİLLER LİSTESİ

	Sayfa
Şekil 2.1. Alan verilerine dayalı bir modelin uygulanması [59].....	11
Şekil 2.2. KNN çalışma yapısı [67].....	13
Şekil 2.3. DVM hiper düzlem yapısı [72]	14
Şekil 2.4. Rastgele orman ağaçlarının gösterimi [76].....	15
Şekil 3.1. Derin Öğrenme kavramı ve onun Makine Öğrenimi ile ilişkisi [59].....	17
Şekil 3.2. Derin sinir ağlarının aşamalı öznetelik çıkarımı [91].....	18
Şekil 3.3. Pnömoni sınıflandırılmasına ait transfer öğrenme yöntemi gösterimi [95]	20
Şekil 3.4. Manuel özellik çıkarma ile karşılaştırıldığında uçtan uca öğrenme kavramının gösterimi [98]... 21	
Şekil 3.5. ESA'nın tipik mimarisi [100].....	22
Şekil 3.6. Evrişimli işleminin bir gösterimi [100]	24
Şekil 3.7. Aktivasyon fonksiyonu olarak ReLU kullanımına bir örnek [100].....	26
Şekil 3.8. Maksimum havuzlamaya örnek bir gösterim [100].....	27
Şekil 3.9. Tam bağlantılı katman [112]	28
Şekil 3.10. Dropout'a örnek [114].....	29
Şekil 3.11. AlexNet [115].....	30
Şekil 3.12. SqueezeNet'de kullanılan ateşleme modülü [117]	31
Şekil 3.13. SqueezeNet mimarisi [118]	31
Şekil 3.14. VGG16 [115]	32
Şekil 3.15. Vgg19 [115]	32
Şekil 3.16. Artık öğrenme: bir yapı taşı [120]	33
Şekil 3.17. ResNet50 mimarisi [122]	34
Şekil 3.18. DenseNet201 basit gösterimi [127]	35
Şekil 3.19. MobileNetV2 ağının mimarisi [130]	36
Şekil 3.20. ShuffleNet Birimleri [132]	37
Şekil 3.21. UKSB biriminin temel yapısı [138]	39
Şekil 3.22. Otomatik kodlayıcının şematik diyagramı [144]	42
Şekil 3.23. Pnömoni tespiti için ViT modelinin gösterimi	44
Şekil 3.24. MLP-mixer mimarisinin gösterimi.....	45
Şekil 3.25. Pnömoni tespiti için önerilen FNet mimarisinin gösterimi.....	46
Şekil 3.26. gMLP mimarisinin gösterimi	47
Şekil 4.1. Oluşturulan veri setine ait pnömoni ve sağlıklı CXR görüntüleri	48

Şekil 4.2. Halka açık CXR veri setine ait sağlıklı ve pnömoni CXR görüntüleri [22]	50
Şekil 4.3. Pnömoni tespiti için önerilen yeni 21 katmanlı ESA modeli.....	51
Şekil 4.4. 5 katlı çapraz doğrulama örneği [154].....	52
Şekil 4.5. Hold out çapraz doğrulama örneği [154].....	53
Şekil 4.6. Modelin performansının görselleştirilmesini kolaylaştıran bir karışıklık matrisi [157]	54
Şekil 5.1. CXR görüntüleri ile pnömoni tespiti için önerilen yöntemin akış diyagramı [161]	57
Şekil 5.2. Uygulanan yöntemlerin en yüksek doğruluk puanı açısından karşılaştırılmasının grafiksel analizi	68
Şekil 5.3. Pnömoni ve sağlıklı CXR görüntülerinin dönüştürücü tabanlı modeller ve 21 katmanlı ESA ile sınıflandırılması	69
Şekil 5.4. 21 katmanlı geliştirilen ESA modeli kullanılarak elde edilen karışıklık matrisi	71
Şekil 5.5. 21 katmanlı ESA modeli için Grad-CAM tekniği kullanılarak sağlıklı ve pnömoni sınıfları için elde edilen ısı haritaları.....	72
Şekil 5.6. CXR görüntülerinden pnömoni tespiti için önerilen yöntemin performans değerlendirme sonuçlarının grafiksel analizi	73
Şekil 5.7. Sayısal veriler ile pnömoni tespiti için önerilen yöntemin akış diyagramı.....	76
Şekil 5.8. Sayısal veriler ile pnömoni tespiti için önerilen yöntemin performans metriklerinin grafiksel analizi.....	80

TABLÖLAR LİSTESİ

	Sayfa
Tablo 4.1. Oluşturulan sayısal parametre veri seti öznetelikleri	49
Tablo 4.2. Geliştirilen model için kullanılan hiper parametreler	51
Tablo 5.1. Pnömoni/Sağlıklı sınıflandırmasında önceden eğitilmiş ESA modellerinin DVM başarıları	58
Tablo 5.2. Pnömoni/Sağlıklı sınıflandırmasında önceden eğitilmiş ESA modellerinin KNN başarıları	59
Tablo 5.3. Pnömoni/Sağlıklı sınıflandırmasında önceden eğitilmiş ESA modellerinin RO başarıları	60
Tablo 5.4. Pnömoni/Sağlıklı sınıflandırmasında önceden eğitilmiş ESA modellerinin DVM başarıları	61
Tablo 5.5. Pnömoni/Sağlıklı sınıflandırmasında önceden eğitilmiş ESA modellerinin KNN başarıları	62
Tablo 5.6. Pnömoni/Sağlıklı sınıflandırmasında önceden eğitilmiş ESA modellerinin RO başarıları	63
Tablo 5.7. Pnömoni/Sağlıklı sınıflandırmasında ESA modellerinin transfer öğrenme başarıları	64
Tablo 5.8. Pnömoni/Sağlıklı sınıflandırmasında ESA modellerinin transfer öğrenme başarıları	65
Tablo 5.9. Geliştirilen ESA modelinin uçtan uca eğitimi ile elde edilen sınıflandırma performansı	66
Tablo 5.10. Uçtan uca eğitilmiş ESA modelinin çapraz doğrulama performansları	66
Tablo 5.11. Önerilen modelin CXR görüntü veri setindeki performans metrikleri	66
Tablo 5.12. Uygulanan yöntemlerin performans karşılaştırması	67
Tablo 5.13. Önerilen yöntemin literatürdeki çalışmalarla performans karşılaştırması	67
Tablo 5.14. Kullanılan modeller için elde edilen sınıflandırma performanslarının özeti	70
Tablo 5.15. Özel veri tabanı için ESA modeli kullanılarak elde edilen performansın özeti	71
Tablo 5.16. CXR görüntüleri üzerinde 21 katmanlı model kullanılarak elde edilen performans matrislerinin özeti	73
Tablo 5.17. Pnömoni tespiti için önerilen yaklaşımın CXR görüntüleri kullanılarak geliştirilen en son tekniklerle karşılaştırılması	74
Tablo 5.18. Pnömoni hastası ve sağlıklı bireylerin demografik, klinik ve laboratuvar bulguları	77
Tablo 5.19. Modellere ait sınıflandırma performans sonuçları	79
Tablo 5.20. Önerilen yaklaşımın mevcut metodolojilerle karşılaştırılması	81

EKLER LİSTESİ

Sayfa

Ek- 1:	Tez Çalışması Verilerinin Kullanımına Dair Etik Kurul Onay Sayfası	98
--------	--	----



SİMGELER

b	: Sapma
c_{t-1}	: Bir önceki hücre durumu
$\hat{c}_t$	: Yeni hücre durumu
C	: Kanal sayısı
f_t	: Unut kapısı (sigmoid fonksiyonunun çıktısı)
h_{t-1}	: Önceki gizli durumdaki bilgiler ($t - 1$ zaman tekrarındaki gizli durum)
H, W	: Orijinal görüntünün çözünürlüğü (yükseklik ile genişlik değerleri)
i_t	: Güncelleme/Giriş kapıları
L	: Dönüştürücü kodlayıcı katmanı
LYM	: Lenfosit
n	: Sınıf sayısı
NEU	: Nötrofil
$\sigma(\cdot)$	: Sigmoid aktivasyon fonksiyonu
o_t	: Çıkış kapısı
$\mathcal{P}_i$	: Veri setinde gözlemlenen sınıfın göreceli sıklığı
PLT	: Trombosit
$s(\cdot)$	: Uzamsal etkileşimleri yakalayan katman
x_t	: Mevcut girişten gelen bilgiler (t . zaman tekrarındaki girdi vektörü)
x_p	: Düzleştirilmiş yamalar
W	: Ağırlık vektörü
WBC	: Lökosit

KISALTMALAR

BN	: Toplu Normalizasyon
BT	: Bilgisayarlı Tomografi
CAO	: CRP Albümin Oranı
CLAHE	: Kontrast Sınırlı Uyarlanabilir Histogram Eşitleme
CRP	: C Reaktif Protein
CXR	: Göğüs Röntgen Görüntüleme
DAA	: Doğrusal Ayırma Analizi
DNN	: Derin Sinir Ağı
DÖ	: Derin Öğrenme
DVM	: Destek Vektör Makinesi
ESA	: Evrişimsel Sinir Ağları
FN	: Yanlış Negatif
FP	: Yanlış Pozitif
GMLP	: Kapılı Çok Katmanlı Algılayıcı
GSHHO	: Gauss Mutasyon Mekanizmasına ve Benzetilmiş Tavlama Yöntemine Dayanan Geliştirilmiş Harris Şahini Optimizasyonu
HBYS	: Hastane Bilgi Yönetim Sistemi
HE	: Histogram Eşitleme
ILSVRC	: ImageNet Büyük Ölçekli Görsel Tanıma Mücadelesi
KA	: Karar Ağacı
KNN	: K En Yakın Komşu
LR	: Lojistik Regresyon
MLP	: Çok Katmanlı Algılayıcı
MSA	: Çok Başlı Öz-Dikkat
MRG	: Manyetik Rezonans Görüntüleme
MP-ViT	: Çok Seviyeli Yama Birleştirme Görüntü Dönüştürücüsü
NAS	: Nöral Mimari Arama
NB	: Naive Bayes
NLO	: Nötrofil Lenfosit Oranı
NLP	: Doğal Dil İşleme
PACS	: Görüntü Saklama ve İletişim Sistemleri
RO	: Rastgele Orman
ReLU	: Doğrultulmuş Doğrusal Birim
TLO	: Trombosit Lenfosit Oranı
TN	: Gerçek Negatif
TP	: Gerçek Pozitif
UKSB	: Uzun Kısa Süreli Bellek
VGG	: Görsel Geometri Grubu
ViT	: Görüntü Dönüştürücü
YSA	: Yapay Sinir Ağları

1. GİRİŞ

Pnömoni, akciğerlerdeki küçük hava keseleri olan alveolleri de etkileyen, sol akciğer, sağ akciğer veya her iki akciğerde meydana gelen akciğer parankiminin iltihaplanmasıdır [1]. Pnömoni, akut hastalıklar kategorisi altında yer alan bir alt solunum yolu enfeksiyonudur ve dünya çapında çocuklarda görülen en büyük bulaşıcı hastalık ve ölüm nedenidir [2]. 2019 yılında 5 yaşın altında 740.180 çocuğun ölümüne neden olarak beş yaşın altındaki tüm çocuk ölümlerinin %14'ünü oluşturan Pnömoni, 1 ile 5 yaş arasındaki çocuk ölümlerinin ise %22'sini oluşturmaktadır [3]. Pnömoni prevalansı yaşla birlikte artmakta ve 65 yaş üstü kişiler için risk oluşturmaktadır [4]. Göğüs ağrısı, nefes alma güçlüğü, ateş, kuru ve bazen balgamlı öksürük pnömoninin sık görülen semptomlarıdır ve bunların şiddeti hastalığın seyrine bağlı olarak değişebilir. Pnömoni 'ye virüsler, bakteriler, mantarlar, immünolojik bozukluklar ve hatta kimyasallar neden olabilir [4]. Klinikte bakteriyel pnömoni, mikoplazma pnömonisi, fungal pnömoni ve influenza gibi virüslere bağlı olarak viral pnömoni görülmektedir [5]. Pnömoni, özellikle yüksek düzeyde hava ve ortam kirliliğinin, aşırı kalabalığın, sağlıksız yaşam koşullarının yanı sıra tıbbi kaynak ve personel eksikliğinin yaygın olduğu gelişmekte olan veya az gelişmiş ülkelerde birçok kişiyi etkilemektedir [6]. Dolayısıyla gelişmekte olan veya az gelişmiş ülkelerde pnömoniyi doğru ve hızlı bir şekilde teşhis etmek son derece hayatidir.

Pnömoninin erken bir aşamada ve doğru bir şekilde teşhis edilmesi, küratif tedavi sağlamak, etkili hasta bakımı ve hayatta kalma oranlarını artırmak için çok önemlidir [7]. Bununla beraber pnömoninin erken teşhis ve tedavisi bu ülkeler için zaman ve finans açısından tasarruf sağlayabilir. Pnömoniyi teşhis etmek için kliniklerde çeşitli tıbbi görüntüleme yöntemleri ve teşhis araçları kullanılmaktadır. Risk altındaki hastaları belirlemek, hastalığın önemini ve pnömoni prognozunu belirlemek ve ayrıca uygun antibiyotik kullanımına karar vermek için kapsamlı bir fiziksel muayene, klinik semptomlar, laboratuvar bulguları ve çeşitli tıbbi görüntüleme yöntemleri bir bütün olarak incelenmektedir [8].

Görüntüleme, pnömonili hastaların saptanmasında ve yönetiminde çok önemli bir rol oynar. Pnömoniyi teşhis etmek için kullanılan tıbbi görüntüleme yöntemleri, göğüs röntgen görüntüleme (CXR), bilgisayarlı tomografi (BT), manyetik rezonans görüntüleme (MRG) ve ultrasondur. Düşük maliyeti, daha hızlı görüntüleme süresi, erişim kolaylığı ve invazif olmayan bir yaklaşım olması nedeniyle, CXR dünya çapında pnömoniyi tespit etmek için kullanılan en yaygın tıbbi görüntüleme tekniğidir [9]. CXR enfeksiyonun şiddeti hakkında önemli miktarda bilgi içerir, ancak son derece deneyimli radyologlar için bile benzer lezyonları ayırt etmek veya çok belirsiz nodülleri saptamak ve doğru yorumlamak zor olabilir [10]. Buna ek olarak, röntgen bulguları hastalığın erken evrelerinde her zaman mevcut olmayabilir, bu da geç tanıya neden olur; göğüs röntgenlerinin yorumlanması zorlaşır [11]. CXR görüntülerinin düşük çözünürlüğü, hastalık belirtileri arasındaki

benzerlikler ve bir CXR görüntüsünü incelerken deneyim ve odaklanma eksikliği, yaşamı tehdit etme potansiyeline sahip teşhis hatalarına yol açabileceğinden yorumlamayı radyologlar için zorlu bir görev haline getirmektedir [12]. Bu nedenle, CXR pnömoninin teşhis ve tespiti konusunda zaman alıcıdır ve doğruluğu kafa karıştırıcı olabildiğinden tanı ve tedavi süreci gecikebilir. Bununla birlikte düşük gelirli ülkelerde ve hastalık yükünün en yüksek olduğu bölgelerde ileri radyografi teknikleri genellikle mevcut değildir [13]. Benzer şekilde, tüm şüpheli pnömoni vakalarına CXR yapmak da zordur. Kaynaklara erişimin sınırlı olduğu az gelişmiş ülkelerdeki klinisyenler çoğunlukla klinik belirti, semptom ve basit laboratuvar testlerinin kullanımı gibi invaziv olmayan ölçümlere dayalı olarak pnömoni teşhisi yapmaktadır [14]. Bu bağlamda hem CXR'lerden hem de laboratuvar testleri, klinik belirti ve semptomlardan pnömoninin çok yönlü teşhisinin doğruluğunu ve hızını artırabilecek otomatik pnömoni tanı ve tespit yöntemlerinin geliştirilmesine ihtiyaç duyulmaktadır. Pnömoninin otomatik tespiti, uzman radyologlara ve klinisyenlere yardımcı olacaktır ve uygun tıbbi olanaklara sahip olmayan gelişmekte olan ülkeler için avantajlı olacaktır [15].

Bilgisayar destekli tanı sistemleri, hekimlerin gözlemsel hatalarını ve dolayısıyla yanlış negatif oranlarını azaltmak ve tanı performansını artırmak için tasarlanmış bir teknolojidir. Bu sistemler hastalıkların teşhisini ve tahmin görevlerini küçük hata paylarıyla daha kolay hale getirmektedir [16]. Derin öğrenme (DÖ), günümüzde tıp endüstrisinde kilit bir rol oynayacak kadar ilerleme kaydetmiştir. DÖ, beyin-bilgisayar arayüzü, bilgisayar destekli teşhis, sağlıkla ilgili elektronik verilerin analizi, tedavi planlaması ve ilaç alımı, ortam tanıma ve tıbbi görüntülerdeki tümör ve lezyonların tespiti gibi çeşitli uygulamalara uygulanmaktadır [17]. DÖ algoritmaları, CXR dahil olmak üzere çeşitli görüntüleme muayenelerinin yorumlanmasını hızlandırmak, otomatikleştirmek ve iyileştirmek için bir çözüm olarak önerilmiştir. DÖ tabanlı bilgisayar destekli tespit ve teşhis sistemleri ile hekimler tanıları daha hızlı ve doğru bir şekilde koyabilirler. DÖ algoritmaları, yoğun klinik uygulamalarda eğitilmiş bir radyologun bulunmadığı veya aşırı yüklendiği acil durumlarda görüntü yorumlamasını hızlandırabilir. Ayrıca, radyologların doğruluğunu artırmak için ikinci bir okuyucu olarak da kullanılabilir [18]. DÖ, insanların nasıl öğrendiğini ve düşündüğünü taklit ederek bilgisayarların ham verilerden ilgili bilgileri otomatik olarak çıkarmasına, analiz etmesine ve yorumlamasına olanak tanıyan daha gelişmiş bir yaklaşım sunan bir makine öğrenimi dalıdır. DÖ, otonom özellik mühendisliği süreçlerine dayanan bir dizi nöral veri odaklı yaklaşımdır; doğruluğu ve performansı, özellikleri girdilerden otomatik olarak öğrenmesinden kaynaklanmaktadır [19]. Evrimsel sinir ağları (ESA), tekrarlayan sinir ağları, derin inanç ağları, grafik sinir ağları ve üretken rakip ağlar gibi birkaç tür derin öğrenme mimarisi vardır. Ham veriden başlayarak modelin her katmanı, bir önceki katmandan öğrenerek daha yüksek boyutlu analizler gerçekleştirir ve bu da tanısal performansı artırmaktadır. Derin öğrenme algoritmalarının çalışma prensibi sayesinde pnömoni teşhisinin etkinliği ve doğruluğu büyük

ölçüde iyileştirilmiştir [20,21]. Literatürde hem CXR görüntüleri hem de laboratuvar testleri, klinik bulgu ve semptomlar gibi sayısal parametreler kullanılarak yapılan pnömoni tespit çalışmaları bulunmaktadır. Bu tez çalışması kapsamında derin öğrenme tabanlı bilgisayar destekli tespit ve teşhis sistemlerinin kullanımının pnömonin çok yönlü teşhisinde etkisi ve başarısı araştırılmıştır. Literatürde yer alan çalışmalar sınıflandırma performansına, sınıf sayısına, kullanılan veri setine, ön işleme aşamasına ve kullanılan derin öğrenme algoritmasının yapısına göre gözden geçirilmiştir. Bu amaçla ilk olarak CXR görüntülerinden pnömoni tespit çalışmaları incelenmiştir.

Ayan çalışmasında, Kaggle veri tabanındaki CXR görüntülerini [22], sınıflandırmak için ince ayarlı bir Xception modeli eğitmiştir. Ardından, ince ayarlı Xception modelini, çeşitli makine öğrenimi algoritmaları için bir özellik çıkarıcı olarak kullanmıştır. İnce ayarlı Xception modelinin evrişim katmanları tarafından çıkarılan özellikler destek vektör makinesi (DVM), k en yakın komşu (KNN), lojistik regresyon (LR), naive bayes (NB) ve karar ağaçları (KA) kullanılarak eğitilmiştir. Test sonuçlarına göre ince ayarlı Xception modeli %89,74 doğrulukla diğer sınıflandırıcılara göre daha iyi sınıflandırma sonuçları elde etmiştir. Makine öğrenimi sınıflandırıcıları arasında en iyi puanı %89,58 doğrulukla DVM algoritması almıştır [23].

Demir ve ark., CXR görüntüleri [22] aracılığıyla çocuklarda pnömoninin erken ve doğru teşhisinde uzmanlara yardımcı olabilecek ve sağlıklı ve hasta bireyleri sınıflandırabilecek transfer öğrenme tabanlı ESA yaklaşımı önermiştir. Çalışma sonucunda AlexNet mimarisi katmanlarına ek katmanlar eklenerek özgün bir ESA mimarisi önerilmiş ve %96,31 test doğruluğu elde edilmiştir [24].

Bakır ve arkadaşları, CXR veri setini [22] kullanarak çoklu ve ikili sınıflandırma görevlerinde çeşitli derin öğrenme teknikleri benimsemiştir. Öznitelik çıkarıcı olarak ResNet, Inception ve MobileNet gibi ESA mimarileri benimsenmiş ve yapay sinir ağları (YSA) algoritması sınıflandırma için kullanılmıştır. Çalışma sonucunda, önerilen ResNet öznitelik çıkarma aşamasına sahip YSA modeli, CXR veri seti [22] üzerinde çok sınıflı sınıflandırma (bakteriyel pnömoni, viral pnömoni ve sağlıklı) yapıldığında en yüksek sınıflandırma doğruluk oranını %81,67 olarak vermiştir. Öte yandan, önerilen MobileNet öznitelik çıkarma aşamasına sahip YSA modeli, CXR görüntü veri seti üzerinde ikili bir sınıflandırma yapıldığında %95,67 ile en yüksek doğruluk oranını vermiştir [25].

Ouledroun ve ark., pnömoniyi saptamak ve radyologlara karar verme süreçlerinde yardımcı olmak için yeni bir yöntem önermiştir. İlk olarak, göğüs röntgeni görüntüleri için histogram eşitleme (HE) ve Kontrast Sınırlı Uyarlanabilir Histogram Eşitleme (CLAHE) hesaplanmıştır. Ardından, çıkarılan görüntüler, [22] veri kümesi üzerinde eğitilmiş iki ESA akışından oluşan bir modele beslenmiştir. Son olarak, çıkarılan derin özelliklere dayalı algılama sürecini gerçekleştirmek için birkaç makine öğrenme sınıflandırıcısı (DAA (Doğrusal ayırma analizi),

DVM, KNN) kullanılmıştır. Önerilen sistem (ESA-KNN), [22] veri setinde doğruluk açısından %97,86'ya ulaşmıştır [26].

Aljawarneh ve ekibi çalışmalarında, geliştirilmiş bir ESA modeli, ResNet, VGG19 ve ince ayarlı ResNet modellerini geniş bir XRI veri seti ile kullanarak çok sayıda pnömoni saptama modelini değerlendirmiştir. Geliştirilmiş ESA modeli, ilk dördü 3×3 filtre boyutuna sahip olan yedi katmana sahiptir. Çalışma için veri seti Kaggle'dan temin edilmiştir [22]. Deneysel bulgulara göre ResNet50 modeli en düşük doğruluğu yani %82,80'i gösterirken, geliştirilmiş ESA modeli %92,40 ile en yüksek doğruluğu göstermiştir [27].

Kanawade ve ark., tarafından pnömoni tespitinde kullanılmak üzere VGG16, VGG19, DenseNet121, DenseNet201, DenseNet169, InceptionResNet, MobileNet, MobileNetV2 ve Xception gibi önceden eğitilmiş ESA modellerinin çeşitli varyantları test edilmiştir. Sonuçlara göre Xception, VGG16 ve DenseNet201 diğer modellerden daha iyi performans göstermiştir. VGG16, Xception ve Densenet201 birleştirildikten sonra sistemin genel doğruluğu %94,39'dur [28].

Iparraguirre-Villanueva ve arkadaşları, akciğer röntgeni görüntülerinden pnömoniyi saptamak için dört ESA modeli sunmaktadır. ESA'lar, X-ışını görüntülerini birkaç evrişimli katman kullanarak normal ve pnömoni olmak üzere iki türe ayırmak üzere eğitilmiştir. Çalışmada VGG16, VGG19, ResNet50 ve InceptionV3 kullanılmıştır. Sonuçlar, Inceptionv3 modelinin %72,90 doğruluk, %93,70 duyarlılık ve %82,00 F1-Skor ile en iyi performansı elde ettiğini göstermiştir [29].

Derin öğrenme teknikleri ve ESA'lar, görüntü sınıflandırma ve tıbbi görüntü işleme alanlarında üstün başarılar sağlamıştır. Bununla birlikte, ESA'ların bazı dezavantajları vardır. Daha yüksek hesaplama maliyetleri, kaybolan veya patlayan gradyan sorunları, aşırı uyum, daha fazla hiper parametre sayısı vb. bu dezavantajlardan bazılarıdır. Bir ESA, maksimum havuzlama gibi bir işlem nedeniyle önemli ölçüde yavaştır. Ayrıca sinir ağını işlemek ve eğitmek büyük bir veri kümesi gerektirir [30,31]. 2020 yılında Dosovitskiy ve arkadaşları, daha düşük hesaplama maliyeti ile ESA ile karşılaştırılabilir veya daha iyi sınıflandırma performansı elde eden Görüntü Dönüştürücü (Vision Transformer / ViT) modelini önermiş ve ViT büyük ölçekli bilgisayarlı görme veri kümeleri için ESA yerine tercih edilmeye başlanmıştır. Bilgisayarla görmede, ViT evrişimsel bloklar olmadan saf bir dönüştürücü modeli uygular. ESA'da her bir evrişim katmanı, neyin ön plan neyin arka plan olduğuna bakmaksızın tüm görüntüyü işler. Öz dikkat tabanlı dönüştürücü mimarisi, ViT 'in yalnızca görüntünün ana kısmına odaklanarak ve görüntünün geri kalan gereksiz kısmını göz ardı ederek yalnızca anlam özelliklerini öğrenmesini sağlar [32,33]. Dikkat tabanlı dönüştürücü mimarileri kullanılarak yapılan pnömoni tespit çalışmaları literatürde yer almaktadır.

Singh ve ark., CXR görüntülerini kullanarak pnömoni hastalığının tahmin etmek için derin sinir ağını (DNN), bir dikkat mekanizması ile birleştirmiştir. Dikkat farkındalığı olan özellikler üretmek için önerilen ağ, DNN mimarisinde kanal ve uzamsal dikkat modüllerinin

birleştirilmesiyle oluşturulmuştur. Ayrıca deneysel çalışmalarda transfer öğrenme ile önceden eğitilmiş sınıflandırıcılar kullanarak veri kümesi üzerindeki sınıflandırma performansını gözlemlemişlerdir. İki ESA mimarisi ile deneyler yapılmıştır, bunlardan ilki temel ESA, diğeri ise Resnet50'dir. Daha sonra dikkat modülünü bu mimariye entegre etmiş ve performans metriklerindeki iyileşmeyi gözlemlemişlerdir. Dikkat mekanizması tabanlı ResNet50 ağını CXR veri seti [22] ile test etmişlerdir. Önerilen ağ, %95,47'lik bir sınıflandırma doğruluğu ve %92,00'lik bir F1 skor puanı elde etmiştir [30].

Tyagi ve ark., pnömoninin erken saptanması için bir dikkat tabanlı dönüştürücü modeli önermiştir. Çalışmada ESA, VGG16 ve ViT olmak üzere üç model kullanılmıştır. Ek olarak, modellerini geliştirmek için halka açık bir CXR veri seti üzerinde çalışmışlardır [22]. Deneysel çalışmalarda kullanılan ViT yöntemi pnömoni tanımlamasında %96,45 doğruluk performansı göstermiştir [31].

Mabrouk ve ark., CXR görüntülerinden bilgisayar destekli pnömoni teşhisi için topluluk öğrenme modeli önermişlerdir. Önerilen model, üç iyi bilinen ESA modelini, yani DenseNet169, MobileNetV2 ve ViT'i içermektedir. Bu modeller, ince ayar kullanılarak CXR veri kümesi [22] üzerinde eğitilmiştir. Önerilen topluluk öğrenme modeli, %93,91 doğruluk ve %93,43 F1 skor puanı elde ederek diğer son teknoloji yöntemlerden daha iyi performans göstermiştir [34].

Ukwuoma ve ark., pnömoniyi teşhis etmek için hibrit bir derin öğrenme çerçevesi tasarlamıştır. Tasarlanan hibrit çerçeve hem evrişimli ağlar hem de dönüştürücü kodlayıcısı mekanizmaları birleştirilerek geliştirilmiştir. Yöntem, ikili ve çoklu sınıflandırma görevleri için Mendeley [22] ve göğüs röntgeni [35] veri setleri kullanılarak eğitilmiş ve değerlendirilmiştir. Hibrit çerçeveleri, sınıflandırma görevleri için %95,00'in üzerinde doğruluk ve F1 skor değerleri üretmiştir [36].

Cha ve ark., CXR görüntülerinde etkili pnömoni tespiti için dikkat mekanizmasına dayalı bir transfer öğrenme çerçevesi geliştirmiştir. İlk olarak, bir özellik çıkarıcı olarak, önceden eğitilmiş üç modelden, yani ResNet152, DenseNet121 ve ResNet18'den özellikler toplanmıştır. Daha sonra özellik seçme işlemi olarak dikkat mekanizması uygulanmıştır. Önerilen yaklaşım, %96,63 doğruluk, %97,30 F1 skor değeri, %96,23 kesinlik ve %98,46 duyarlılık elde etmiştir [37].

Jiang ve arkadaşları, CXR görüntülerinden otomatik pnömoni teşhisi için çok seviyeli yama birleştirme görüntü dönüştürücüsünü (MP-ViT) önermiştir. Deneylerini kamuya açık bir veri kümesi [22] üzerinde gerçekleştirmişlerdir. Önerdikleri model %91,00 doğruluk, %92,00 kesinlik, %89,00 duyarlılık ve %90,00 F1-skoru elde etmiştir [38].

Ma ve arkadaşları CXR görüntülerinde pnömoni tanıma modeli olarak Swin dönüştürücüsünü kullanmış ve bunları CXR görüntülerinin özelliklerine göre optimize etmiştir. İki farklı veri kümesi [22, 35] üzerinde yapılan karşılaştırmalı deneylerden sonra, deneysel sonuçlar

modelin doğruluğunun sırasıyla %76,30'den %87,30'e ve %92,80'den %97,20'ye yükseldiğini göstermiştir [39].

Belirti, semptom gibi klinik parametreler ve biyobelirteçler gibi klinik laboratuvar parametreleri, pnömoni hastalarının yönetiminde başarı oranlarını artırmak için yararlı araçlardır. Hastalık şiddetini değerlendirmek ve tanı sürecini basitleştirmek için biyobelirteçlere ihtiyaç duyulmaktadır. Normal veya patojenik bir biyolojik sürecin ya da bir maruziyete veya müdahaleye verilen yanıtın göstergesi olan ölçülebilir bir özellik olarak tanımlanan biyobelirteçler, solunum tıbbında hastalık temelli değil hasta temelli bir terapötik yaklaşımın geliştirilmesinde kilit bir rol kazanmıştır [40]. Biyobelirteçlerin değerlendirilecek hastalık veya olay için yüksek özgüllüğe sahip olmaları, ölçülmelerinin kolay ve ucuz olması, iyi bir ayırt etme kapasitesi göstermeleri ve şu anda geleneksel klinik uygulamada kullanılan göstergelerden daha uygun maliyetli olmaları gerekir [40,41]. Biyobelirteçlerin klinik uygulamalarda kullanılması yalnızca doğru teşhisi kolaylaştırmakla kalmayacak, aynı zamanda aşırı antibiyotik kullanımının azaltılmasına da yardımcı olacaktır [42]. Tam kan sayımı, lökosit, nötrofil, trombosit, lenfosit, Nötrofil Lenfosit Oranı (NLO) ve Trombosit Lenfosit Oranı (TLO) düzeyleri kan içeriği hakkında bilgi sağlayan kolay, ucuz, rutin bir inceleme teknikleridir [43]. Pnömoni hastalığının tanısı karaciğer tarafından üretilen bir akut faz proteini olan C-reaktif protein (CRP) gibi biyokimyasal parametrelerle desteklenir. CRP, birçok enfeksiyonda, otoimmün hastalıkta yükselen pozitif bir akut faz reaktanıdır. Albümin ise çeşitli hastalıklarda inflamasyonun şiddeti ile ters korelasyon gösteren negatif bir akut faz reaktanıdır [44]. Yeni bir parametre olarak CRP Albümin Oran'ının (CAO), belirli klinik durumların genel prognozunu tahmin etmede tek başına albümin ve CRP'den daha doğru olduğu gösterilmiştir [44]. Literatürdeki çalışmalarda, CRP, albümin veya CAO'nun enflamatuvar için yararlı bir prognostik faktör olarak kullanılabileceğini ve özellikle CAO'nun, enflamatuvar yanıtın güçlü bir göstergesi olabileceği gösterilmektedir [45, 46]. Nötrofil, lenfosit ve trombosit düzeylerinin sistemik inflamasyon ve enfeksiyonda önemli rol oynadığı bildirilmiştir [47]. NLO ve TLO gibi yeni inflamatuvar biyobelirteçlerin pnömonili hastalarda anlamlı olarak artmış olması pnömoni varlığı için öngörücü olarak kullanılabileceğini göstermektedir [48, 49]. Sonuç olarak son yıllarda yapılan çalışmalarda rutin laboratuvar ve/veya klinik verilere dayalı erken teşhis yöntemleri, tıbbi görüntüleme tekniklerine göre daha hızlı, kullanımı kolay, ulaşılabilir ve daha ucuz alternatifler olması nedeniyle tercih edilmektedir [50]. Teşhis doğruluğunu artırmak ve pnömonide çeşitli tedavi stratejileri geliştirmek için kullanılmaya başlanan makine öğrenmesi ve derin öğrenme tabanlı tahmin modelleri bu sorunlara çözüm sağlama potansiyeline sahip olup tıpta oldukça iyi ilerlemeler ve büyük faydalar sağlamıştır [51]. Literatürde çeşitli makine öğrenmesi ve derin öğrenme algoritmaları kullanılarak pnömoni hastalığına ait laboratuvar parametreleri ve klinik bulguların sınıflandırma çalışmaları yer almaktadır [52].

Effah ve ark., pnömoniye laboratuvar parametrelerine ve fiziksel özelliklere dayalı olarak tahmin etmek için her biri 45 özelliğe sahip 535 farklı hasta üzerinde sekiz makine öğrenimi modeli kullanmışlardır. Çalışmada Rastgele Orman (RO) (doğruluk = %92,00, kesinlik = %91,30, duyarlılık = %96,00, f1-Skor = %93,60) ve XGBoost (doğruluk = %90,80, kesinlik = %92,60, duyarlılık = %92,30, f1-skor = %92,40) modelleri orijinal veri setinde en yüksek performansı elde etmiştir [51].

Zhang ve ark., klinik, biyokimyasal ve göğüs bilgisayarlı tomografi özelliklerine dayalı bir veri kümesi oluşturarak, grip virüsü kaynaklı pnömoniye bakteriyel pnömoniden ayırmada makine öğrenimi tekniklerinin kapasitesini keşfetmeye odaklanmıştır. Gauss mutasyon mekanizmasına ve benzetilmiş tavlama yöntemine dayanan geliştirilmiş bir Harris Şahini optimizasyonu (GSHHO) önerilmiş ve ardından bir makine öğrenme modeli olan GSHHO-FKNN oluşturmak için GSHHO'yu Fuzzy k-en yakın komşu (FKNN) ile birleştirmişlerdir. Önerdikleri model ortalama %90,00 doğruluk, %82,64 duyarlılık, % 93,00 kesinlik ve %94,67 özgünlük puanları elde etmiştir [52].

Stokes ve ark., orta gelirli bir ülkedeki 4344 hastadan (1500 bronşit, 2844 pnömoni) elde edilen denek popülasyon özellikleri, semptomlar ve laboratuvar test sonuçları hakkında bilgi içeren bir veri kümesine makine öğrenimi algoritmaları uygulamışlardır. Uygulanan makine öğrenmesi algoritmaları LR, KA ve DVM'dir. Karar ağacı %84,00 doğruluk, %87,00 F1 skor, %80,00 duyarlılık ve %73,00 kesinlik ile diğer modellerden daha iyi performans göstermiştir [53].

Bu çalışmada, pnömoninin çok yönlü tespiti için sayısal tıbbi veriler ve göğüs röntgen görüntüleri derin öğrenme yaklaşımları kullanılarak sınıflandırılmıştır. Diele Üniversitesi Tıp Fakültesi Göğüs Hastalıkları ve Tüberküloz kliniği ile yoğun bakım ünitesine 2001 Ocak- 2021 Ocak ayları arasında pnömoni tanısıyla yatırılan hastalar ve herhangi bir solunum şikâyetiyle göğüs hastalıkları polikliniğine başvurup CXR çekilen ve normal olan hastaların dosyalarının retrospektif olarak taranmasıyla veri kümeleri oluşturulmuştur. Çalışmaya CXR'de pnömoni tanısı alan 1000 pnömoni hastası ve CXR'si normal olan 1000 sağlıklı birey olmak üzere toplam 2000 kişi dâhil edilmiştir. 2000 kişiye ait (1000 pnömoni, 1000 sağlıklı) demografik özellikler, semptomlar ve laboratuvar test sonuç bilgilerinden oluşturulan sayısal tıbbi veri kümesine makine öğrenimi ve derin öğrenme algoritmaları uygulanmıştır. Aynı kişilere ait CXR görüntüleri de derin öğrenme tabanlı hibrit yöntemlerle sınıflandırılmıştır. Sayısal tıbbi veri kümesi için yaş, cinsiyet, dispne, öksürük, yan ağrısı, nötrofil, lenfosit, trombosit, lökosit, CRP ve albümin gibi demografik, klinik ve laboratuvar bulguları seçilerek bu verilerin pnömoni hastalarının ve sağlıklı bireylerin ayırıcı tanısında etkisi ve rolü incelenmiştir. 2000 kişiye ait 15 özellikli bir sayısal veri seti oluşturulmuştur. Sayısal tıbbi veri kümesinin sınıflandırma aşamasında kullanılan sınıflandırıcılar KNN, DVM, RO, Otomatik Kodlayıcı ve Uzun Kısa Süreli Bellek (UKSB)'dan oluşturulmuştur. Oluşturulan sayısal parametre veri setinden pnömoni tespiti için de bir ESA modeli geliştirilmiş ve

uçtan uca eğitilmiştir. Oluşturulan sayısal parametre veri setinin sınıflandırılmasında en yüksek doğruluk %96,25 olarak elde edilmiştir. CXR görüntülerinin sınıflandırılması iki aşamada gerçekleştirilmiştir. İlk aşamada pnömoni ve sağlıklı CXR görüntülerini sınıflandırmak için derin öğrenmeye dayalı yaklaşımlar kullanılmıştır. Bu yaklaşımlar, derin öznitelik çıkarımı, önceden eğitilmiş ESA'ların ince ayarı ve geliştirilmiş bir ESA modelinin uçtan uca eğitimidir. Derin öznitelik çıkarımı ve transfer öğrenme için önceden eğitilmiş AlexNet, VGG16, GG19, ShuffleNet, NasNetMobile, MobileNetV2, DarkNet53, SqueezeNet, ResNet50 ve DenseNet201 modelleri kullanılmıştır. AlexNet'in fc6; VGG16 ve VGG19'un fc7; ResNet50 ve DenseNet201'in fc1000; DarkNet53'ün conv53; NasNetMobile'in predictions; MobileNetV2'nin Logits; ShuffleNet'in node_202 ve SqueezeNet'in pool10 katmanlarından öznitelikler çıkarılıp sınıflandırma algoritmalarına verilmiştir. Çıkarılan derin öznitelikleri sınıflandırmak için kullanılan algoritmalar KNN, DVM ve RO algoritmasıdır. Uçtan uca öğrenme yöntemi için 21 katmanlı yeni bir ESA mimarisi önerilmiş ve test edilmiştir. İnce ayarlı AlexNet modelinin başarısı, elde edilen tüm sonuçlar arasında en yüksek olan %98,50 doğruluk puanı üretmiştir. Geliştirilen 21 katmanlı ESA modelinin uçtan uca eğitimi %96,75 sonuç vermiştir. İkinci aşamada tez çalışması kapsamında önerilen ESA modelinin performansı yakın zamanda geliştirilen dönüştürücü tabanlı mimarilerle karşılaştırılmıştır. Bu dönüştürücü mimariler ViT, kapılı çok katmanlı algılayıcı (gMLP), MLP-mixer ve FNet'tir. ViT, FNet, MLP-mixer ve gMLP modellerinin sınıflandırma performans doğrulukları sırasıyla %96,43, %96,43, %95,41 ve %94,39'dur. Bu tez çalışması kapsamında gerçekleştirilen pnömoni tespit yöntemlerinin literatüre katkısını incelemek amacıyla mevcut veri setleri ile önerilen yöntemler test edilmiş ve yapılan çalışmalar ile sınıflandırma performansları karşılaştırılmıştır.

Bu tez çalışmasının katkıları şunlardır:

- 1- Pnömoninin otomatik tespiti için 1000'i pnömoni 1000'i sağlıklı bireylere ait olmak üzere toplam 2000 yeni CXR görüntüsü ve aynı kişilere ait demografik veri, klinik bulgu ve laboratuvar parametreleri toplanmış ve CXR görüntülerinden ve sayısal tıbbi verilerden 2 yeni veri seti oluşturulmuştur.
- 2- CXR görüntülerinden pnömoni tespiti için yeni bir 21 katmanlı ESA modeli önerilmiştir.
- 3- Pnömoni tespitinde derin öğrenme yöntemlerinin ve derin öğrenme modellerinin performansları karşılaştırılmıştır ve analiz edilmiştir.
- 4- Veri seti hem k kat çapraz doğrulama hem de ayırarak çapraz doğrulama yöntemleriyle ayrılarak önerilen yöntemlerin geçerliliği test edilmiştir.
- 5- Yeni ESA modeli, sağlıklı bireylerden zatürreyi doğru bir şekilde tespit etmede özel ve kamu veri tabanları için yüksek doğruluk oranları sağlamıştır.
- 6- Ayrıca önerilen model, pnömoni tespiti için yakın zamanda geliştirilen dönüştürücü tabanlı modellerden daha yüksek sınıflandırma performansı sağlamıştır.

- 7- İlgili alanlarını göstererek klinisyenlerde güven oluşturmak amacıyla sağlıklı ve zatürre sınıfları için tipik ısı haritaları gösterilmiştir.
- 8- Pnömoninin ayırıcı tanısında sayısal tıbbi verilerin etkisi ve rolü incelenmiştir.
- 9- Sayısal verilerden derin öğrenme yaklaşımıyla pnömoni tespiti için yeni bir ESA mimarisi geliştirilmiştir.
- 10- Yaygın teşhis yöntemi olan CXR'lere alternatif olarak daha düşük maliyetli, yüksek doğruluklu bir karar destek sistemi geliştirilmiştir.
- 11- Önerilen yaklaşımların mevcut modellerle karşılaştırmalı analizleri yapılmıştır.

Bu tez çalışması giriş, makine öğrenmesi, derin öğrenme, materyal ve metot, bulgular ve tartışma, sonuçlar ve öneriler olmak üzere altı bölümden oluşmaktadır. İlk bölüm olan Giriş bölümünde, pnömoninin tanımı yapılmış olup dünya çapında mevcut durumu ile ilgili istatistikler verilmiştir. Yaygın olarak kullanılan pnömoni tespit yöntemlerinden bahsedilmiştir. Bu tezin konusu olan derin öğrenme ve tekniklerinin pnömoni tespitindeki önemi vurgulanmıştır. Son yıllarda derin öğrenme algoritmaları kullanılarak yapılan pnömoni tespiti çalışmalarına ayrıntılı bir şekilde yer verilmiş olup yöntemin üstünlüklerinden bahsedilmiştir. Ayrıca bu tezin çalışmanın amacı, önemi ve tezin organizasyonundan bahsedilmiştir.

Bölüm 2'de makine öğrenmesinin teorik tanımı yapılmış ve algoritmalarından detaylı bahsedilmiştir. Bu tez kapsamında kullanılan makine öğrenmesi algoritmaları detaylı açıklanmıştır. Bölüm 3'de derin öğrenme, kullanıldığı alanlar, bu tez kapsamında kullanılan derin öğrenme yöntemleri, derin öğrenme modelleri ve derin öğrenme katmanları ayrıntılı bir şekilde anlatılmıştır.

Dördüncü bölüm olan Materyal ve Metot bölümünde bu tez için oluşturulan sayısal tıbbi veri kümesi ve göğüs röntgen görüntüleri veri kümesi açıklanmıştır. Pnömoni tespiti için literatürde mevcut kullanılan veri kümesinden bahsedilmiştir. Bu tez kapsamında geliştirilen pnömoni tespit modelinin yapısı, hiper parametreleri ayrıntılı olarak açıklanarak detaylandırılmıştır. Önerilen modelleri değerlendirme/doğrulama yöntemleri ve elde edilen sınıflandırma performanslarını değerlendirmek için kullanılan değerlendirme metrikleri detaylı açıklanmıştır. Ayrıca bu bölümün sonunda sayısal tıbbi verilerin istatistiksel analizinde kullanılan yöntemlere de değinilmiştir.

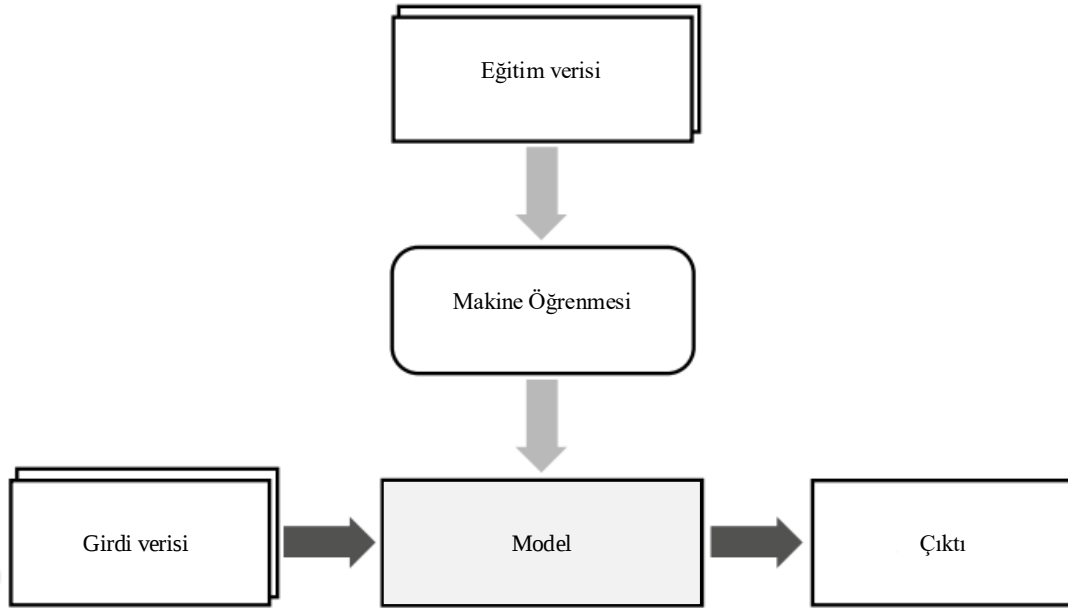
Beşinci bölüm olan Bulgular ve Tartışma bölümünde tez kapsamında yapılan deneysel çalışmalar tablolar halinde verilerek elde edilen sonuçlar ayrıntılı bir şekilde açıklanmıştır. Bununla birlikte tez kapsamında elde edilen sonuçlar literatürdeki diğer benzer çalışmalarla karşılaştırılarak önerilen yöntemlerin üstünlüğü tartışılmış ve literatüre katkısı incelenmiştir.

Altıncı ve son bölüm olan Sonuç bölümünde tez kapsamında yapılan işlemlerin kısa bir özeti anlatıldıktan sonra tez çalışmasının önerdiği modellerin sonuçları bulunmaktadır.

2. MAKİNE ÖĞRENMESİ

Makine öğrenimi, bilgisayarların veri biçimindeki örneklerden öğrenmesine ve verilere dayalı tahminler veya kararlar almasına olanak tanıyan algoritmaların ve modellerin geliştirilmesine odaklanan bir teknolojidir [54]. Programlamaya yönelik geleneksel yaklaşımlar, bir sorunun adım adım nasıl çözüleceğini belirleyen sabit kodlanmış kurallara dayanır. Buna karşılık, makine öğrenimi sistemlerine bir görev ve bu görevin nasıl gerçekleştirilebileceğine veya kalıpların saptanabileceğine dair kullanılmak üzere büyük miktarda örnek veri verilir. Sistem daha sonra istenen çıktıya en iyi nasıl ulaşılacağını öğrenir [55]. Makine öğrenimi, bir veri kümesini en iyi şekilde temsil eden algoritmalar geliştirerek yapay zekânın öğrenme yönüne odaklanan bir alandır [56]. Artan veri kullanılabilirliği, makine öğrenimi sistemlerinin geniş bir örnek havuzu üzerinde eğitilmesine olanak tanırken, bilgisayarların artan bilgi işleme gücü bu sistemlerin analitik yeteneklerini desteklemektedir [57]. Makine öğrenimi günümüzde farklı sektör ve uygulamalarda uygulanmaktadır ve kullanımı giderek artmaktadır. Görüntülerdeki nesneleri tanımlama, konuşmayı metne dönüştürme, haber öğelerini, gönderileri veya ürünleri kullanıcıların ilgi alanlarıyla eşleştirme ve ilgili arama sonuçlarını seçme makine öğrenimi uygulamalarından bazılarıdır [58].

Makine Öğrenimi, veriden bir model oluşturmak (veya öğrenmek) için kullanılan bir tekniktir. Burada veri; belgeler, sesler, görüntüler vb. gibi bilgiler anlamına gelir. Model, Makine Öğreniminin nihai ürünüdür. Makine öğreniminin modelleme sürecinde kullandığı verilere eğitim verileri denir. Makine öğrenimi süreci eğitim verilerinden modeli bulduktan sonra, model gerçek alan verilerine uygulanır. Bu süreç Şekil 2.1’de gösterilmektedir. Şeklin dikey akışı öğrenme sürecini gösterir ve eğitilen model, çıktı olarak adlandırılan yatay akış olarak tanımlanır [59].



Şekil 2.1. Alan verilerine dayalı bir modelin uygulanması [59]

Makine öğrenimi algoritmaları öğrenme türüne göre iki kategoriye ayrılmaktadır. Bunlar gelecekteki çıktıları tahmin edebilmesi için etiketlenmiş girdi ve çıktı verileri üzerinde bir model eğiten denetimli öğrenme ve girdi verilerinde gizli kalıpları veya içsel yapıları bulan denetimsiz öğrenmedir.

Denetimli öğrenme, makine öğrenimi sistemini eğitmek için örnek etiketli veriler sağlanarak bu temelde çıktının tahmin edildiği bir makine öğrenimi yöntemidir [60]. Sistem, veri kümelerini anlamak ve her bir veri hakkında bilgi edinmek için etiketli verileri kullanarak bir model oluşturur, eğitim ve işleme tamamlandıktan sonra, tam çıktıyı tahmin edip etmediğini kontrol etmek için bir örnek veri sağlanarak model test edilir. Denetimli öğrenmenin amacı, girdi verilerini çıktı verileriyle eşleştirmektir [61]. Bir tweet veya ürün incelemesi gibi bir metnin sınıf etiketini veya duyarlılığını tahmin etmek, yani metin sınıflandırması, denetimli öğrenmeye bir örnektir. Denetimli öğrenme algoritmaları sınıflandırma ve regresyon algoritmaları olarak kategorize edilmektedir. Sınıflandırma teknikleri, bir e-postanın gerçek mi yoksa istenmeyen posta mı olduğu veya bir tümörün kanserli mi yoksa iyi huylu mu olduğu gibi ayrık yanıtları tahmin eder. Sınıflandırma modelleri, verileri kategorilere ayırmak için bir eğitim veri kümesindeki önceden tanımlanmış sınıfları kullanır. Yaygın uygulamalar arasında tıbbi görüntüleme, konuşma tanıma ve kredi puanlama yer alır. Regresyon yönteminde, veriler analiz edilerek değişkenler arasındaki ilişki elde edilir ve bu daha sonra gelecekteki eğilimi tahmin etmek için kullanılır. Regresyon teknikleri pilin şarj durumu, şebekedeki elektrik yükü veya finansal varlıkların fiyatları gibi ölçülmesi zor fiziksel niceliklere sahip sürekli yanıtları tahmin eder. Tipik uygulamalar arasında sanal algılama, elektrik yükü tahmini ve algoritmik ticaret yer alır [62].

Denetimsiz öğrenme, denetimden yoksun bir öğrenme tekniğini ifade eder. Burada makine, etiketlenmemiş bir veri seti kullanılarak eğitilir ve herhangi bir denetim olmaksızın çıktıyı tahmin etmesi sağlanır. Denetimsiz bir öğrenme algoritması, sıralanmamış veri kümesini girdinin benzerliklerine, farklılıklarına ve örüntülerine göre gruplandırmayı amaçlar [63]. En yaygın denetimsiz öğrenme görevleri, kümeleme, yoğunluk tahmini, özellik öğrenme, boyutluluk azaltma, birliktelik kurallarını bulma, anormallik tespittir [60]. Kümeleme tekniği, adından da anlaşılacağı gibi, bir veri kümesindeki örnekleri, özelliklerinin belirli kombinasyonlarına göre ayrı kümeler halinde gruplandırır. Müşterileri satın aldıkları ürünlere göre gruplamak kümeleme uygulamasına bir örnektir. Birliktelik kuralı algoritmaları, büyük bir veri kümesinin değişkenleri arasındaki tipik ilişkileri tanımlamayı ifade eder. Çeşitli veri öğelerinin bağımlılığını belirler ve birleşmiş değişkenleri haritalar. Tipik uygulamalar arasında web kullanım madenciliği ve pazar veri analizi yer alır [58].

Bu çalışmada pnömoni tespiti için denetimli öğrenme algoritmalarından olan DVM, KNN ve RO algoritmaları kullanılmıştır. Kullanılan sınıflandırma algoritmalarının teorik altyapıları ve çalışma prensipleri sırasıyla verilmiştir.

2.1. K En Yakın Komşu (KNN)

K en yakın komşu algoritması (KNN), denetimli makine öğrenmesi algoritmaları içerisinde en basit ve en yaygın kullanılan sınıflandırma algoritmalarından biridir. KNN algoritması, çok sınıflı etiket sınıflandırma problemine uygundur ve iyi bir genelleme yeteneğine sahiptir [64]. KNN, sınıflandırılmamış örneğin en yakın komşularını bulmaya ve benzerliği yüksek sınıflara göre tahminler yapmaya dayalı bir sınıflandırma yöntemidir. En yakın komşuları bulmak için veri setini tek tek taramak algoritmanın performansını düşürdüğü için tembel öğrenme yöntemi veya vaka (case) tabanlı öğrenme yöntemi olarak adlandırılır [65]. Bu dezavantaj nedeniyle, KNN algoritması, özellikle büyük hacimli verilerde yavaş çalışma süresine sahiptir.

KNN, örneklem mesafesinin hesaplanmasına dayalı bir öğrenme algoritmasıdır, algoritma yeni bir veri örneğiyle her karşılaştığında, yeni verideki tüm örnekler üzerinden mesafe hesaplanır ve en kısa mesafeye sahip nokta en yakın komşu olarak adlandırılır. Bu hesaplamadan sonra, önceden bilinen veri örneklerinden k adet en yakın komşu bulunup yeni örnekteki eğitim verilerindeki örneklerle karşılaştırılarak ve aralarındaki benzerliklere bakılarak sınıf etiketleri seçilir [66].

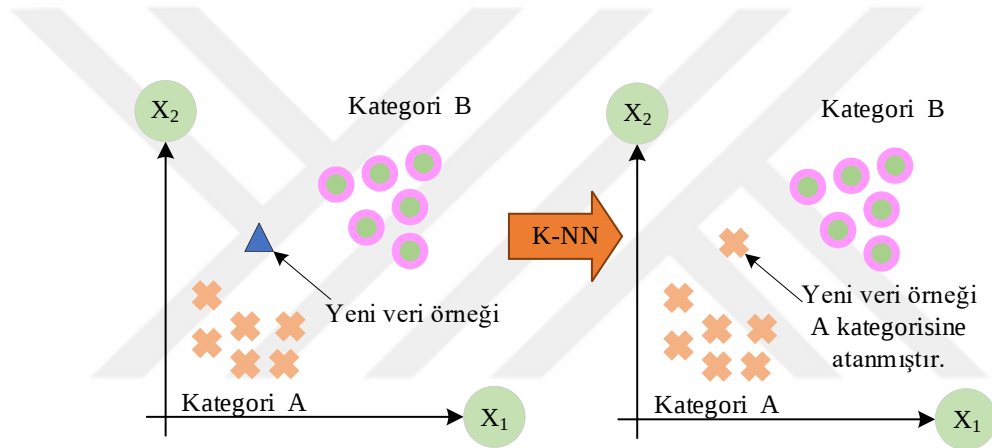
Algoritma basit ve uygulaması kolaydır. Bir model oluşturmaya, birkaç parametreyi ayarlamaya veya ek varsayımlar yapmaya gerek yoktur. KNN hiper parametre olarak diğer makine öğrenimi algoritmalarına kıyasla yalnızca bir k değeri ve bir mesafe metriği gerektirir. Yeni eğitim örnekleri eklendikçe, tüm eğitim verileri belleğe kaydedildiğinden, algoritma herhangi bir yeni veriyi hesaba katacak şekilde ayarlanır. Algoritma çok yönlüdür. Sınıflandırma, regresyon ve

arama için kullanılabilir [65]. KNN mesafe hesaplamasında hiyerarşik kümeleme yöntemlerinde genel olarak kullanılan Öklid, manhattan ve minkowski uzaklık ölçümleri ile işlem yapılmaktadır. Uzaklık formülleri aşağıdaki Denklem 2.1, Denklem 2.2 ve Denklem 2.3'e göre hesaplanır [65]. Şekil 2.2'de KNN algoritmasının çalışma prensibine örnek verilmiştir. Yeni gelen örnek ile etiketlenmiş örnek grupları arasındaki mesafe ve en yakın komşuluklar hesaplanarak yeni gelen örneğe sınıf etiketi atanarak sınıflandırma gerçekleştirilmiş olur [67].

$$\text{Öklid uzaklığı: } \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.1)$$

$$\text{Manhattan uzaklığı: } \sum_{i=1}^n |x_i - y_i| \quad (2.2)$$

$$\text{Minkowski uzaklığı: } (\sum_{i=1}^n |x_i - y_i|^p)^{1/p} \quad (2.3)$$



Şekil 2.2. KNN çalışma yapısı [67]

2.2. Destek Vektör Makineleri (DVM)

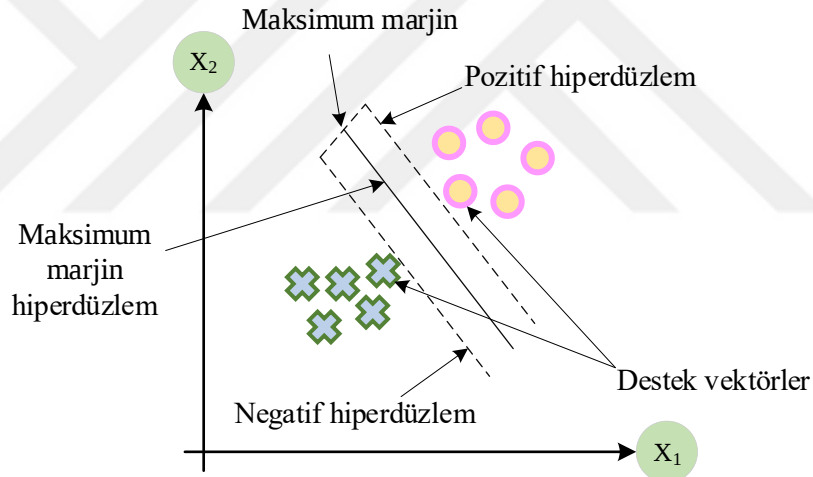
Destek vektör makinesi (DVM), denetimli sınıflandırma ve denetimsiz veri kümeleme veya regresyon uygulamaları için kullanılabilen önemli bir makine öğrenimi algoritmasıdır [68]. DVM sınıflandırıcısının arkasındaki fikir, farklı sınıflara ait veri noktalarını bölen N boyutlu bir uzayda en uygun hiper düzlemi (çizgi/alan sınırlayıcı) bulmaktır. Ancak, bu hiper bölme iki sınıf arasında maksimum marjı sağlayan hiper düzlem dikkate alınarak yapıldığından, marja göre seçilir. Bu marjlar, destek vektörleri olarak bilinen veri noktaları kullanılarak hesaplanır. Destek vektörleri, hiper düzleme daha yakın olan ve hiper düzlemin konumunu ve yönünü etkileyen veri noktalarıdır. Bu destek vektörlerini kullanarak, sınıflandırıcının marjı maksimize edilir [69].

DVM kolayca uygulanabilir ve mevcut veri örneğinin sınırlamaları dahilinde yüksek boyutlu problemlerin çözümünde kullanılan uygun yöntemlerden biridir. Temel olarak DVM, doğrusal ayrılabilir veriler için iyi çalışır, ancak doğrusal olarak ayrılamayan veriler için de DVM'lerin

çekirdek fonksiyonları kullanılarak ayarlanması gerekir. Çekirdekler, ayrılamayan eğitim verilerini farklı sınıfların ayrılabilir verilerine dönüştüren matematiksel fonksiyonlardır. Doğrusal olmayan problemleri çözmek için DVM Gauss, Sigmoid ve Polinom gibi doğrusal olmayan çekirdek fonksiyonlarını kullanır. Veri analizinde DVM uygularken uygun çekirdek ve parametrelerin seçilmesi iki önemli faktördür [70,71]. Hiper düzlemde her nokta Denklem 2.4’de olduğu gibi gösterilmektedir. W ; ağırlık vektörü, x ; giriş vektörü, b ; bir sapmadır. İki doğru denkleminin gösterimi Denklem 2.5’de gösterilmiştir [69]. Şekil 2.3’de DVM’nin temel çalışma yapısı gösterilmiştir.

$$wx + b = 0 \quad (2.4)$$

$$y_i = \begin{cases} -1, & wx + b \leq -1 \\ +1, & wx + b \geq +1 \end{cases} \quad (2.5)$$



Şekil 2.3. DVM hiper düzlem yapısı [72]

2.3. Rastgele Orman Algoritması (RO)

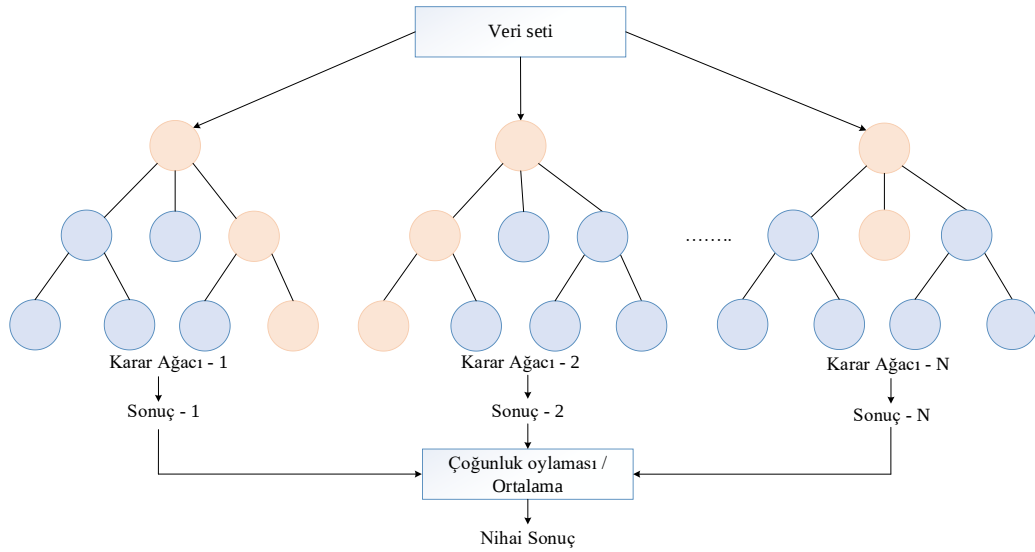
Rastgele orman sınıflandırıcı, basitliği ve kullanılabilirliği nedeniyle hem sınıflandırma hem de regresyon için en yaygın kullanılan makine öğrenimi algoritmalarından biridir. Çok sayıda karar ağacı oluşturularak rastgele orman sınıflandırıcısı güçlendirilmiştir. Birden fazla karar ağacından oluştuğu için bir düğümü kaldırmak, ağaç rastgele seçildiği için çıktıyı çok fazla etkilemez. Temel fikri, zayıf öğrenen gruplarının bir araya gelerek daha güçlü bir öğrenen oluşturması olan karar ağaçları, bir kökle başlar, dallarını büyötmeye devam eder ve sonunda yapraklar adı verilen uç düğüme ulaşır. Ağaca aktarılan dallar, özellikler veya bu özelliklere dayalı olarak işlenmiş

bilgilerdir. Ağaçlar daha sonra bir sınıf sonucu tahmin eder ve ağaçlar arasındaki çoğunluk oyu, modelin nihai sınıf tahmini olarak kullanılır. Diğer algoritmalarla karşılaştırıldığında Rastgele Orman Sınıflandırıcıları, aşırı öğrenme riskinin daha düşük olması nedeniyle nispeten yüksek doğrulukla büyük bir veri tabanı üzerinde verimli bir şekilde çalışır [73,74].

Bu sınıflandırıcının iyi performansının temel sebebi, ormandaki ağaçların birbiriyle nispeten bağlantısız olması ve bir bütün olarak verdikleri kararın, her birinin ayrı ayrı verdiği kararlardan daha iyi olmasını sağlamasıdır. Rastgele orman, kalabalığın bilgeliği adı verilen basit ve güçlü bir temel kavram kullanır. Ağaçlar arasındaki düşük korelasyon, modelin başarısı için çok önemlidir. Bu öncül altında, birkaç ağacın tahmin sonuçları doğru olmasa bile, diğer ağaçların çoğunun tahmin sonuçları doğru olduğu sürece, bu ağaçlar bir grup olarak sonunda doğru tahmin sonuçlarını alabilir. Başka bir deyişle, rastgele orman modeli iyi performans gösterir çünkü bir bütün olarak işleyen nispeten bağlantısız çok sayıda model, herhangi bir bileşen modelden daha iyi performans gösterir [74,75].

Sınıflandırma problemleri için Rastgele Orman kullanılırken genellikle Gini indeksinden yararlanır. Bu formül, bir düğümdeki her dalın Gini'sini belirlemek için sınıfı ve olasılığı kullanır ve dallardan hangisinin oluşma olasılığının daha yüksek olduğunu belirler. Denklem 2.6'da yer alan Gini indeks formülünde $\mathcal{P}_i$, veri setinde gözlemlenen sınıfın göreceli sıklığını temsil eder ve n , sınıf sayısını temsil eder [74]. Şekil 2.4'de rastgele orman sınıflandırıcının nasıl çalıştığını ve tüm karar ağaçlarından son bir sınıf çıkardığını gösterilmiştir.

$$Gini(tvs) = 1 - \sum_{i=1}^n (\mathcal{P}_i)^2 \quad (2.6)$$

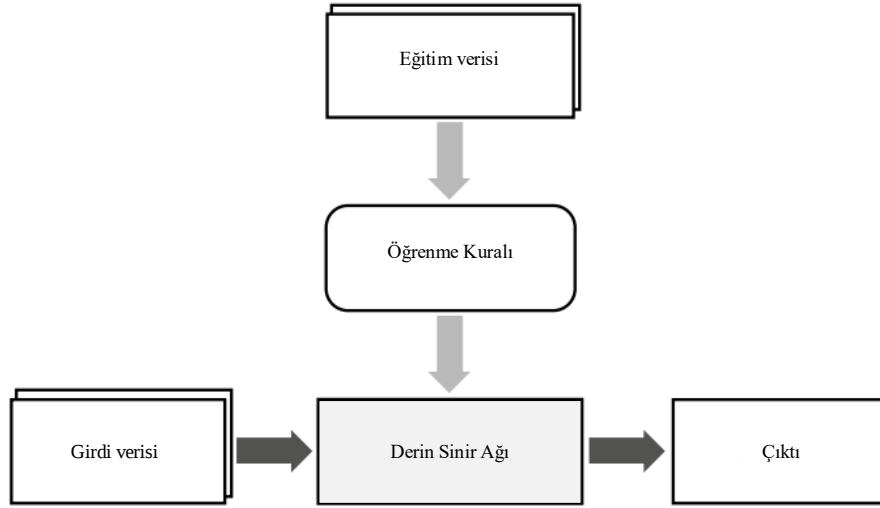


Şekil 2.4. Rastgele orman ağaçlarının gösterimi [76]

3. DERİN ÖĞRENME

Derin öğrenme ham girdi verilerine karşı daha yüksek seviyeli özellikleri çıkarmak için birden çok katmanın kullanımına dayanan bir tür makine öğrenmesi algoritmasıdır [77]. Derin öğrenme algoritmaları, YSA olarak adlandırılan katmanlı bir algoritma mimarisi kullanır. YSA'da tek bir algılayıcının (perceptron) yapısı, insan beynindeki nöron yapısına benzer şekilde tasarlanmıştır. YSA kendi kendine öğrenme ve karar verme yeteneğine sahiptir [78]. Derin öğrenme ağları bir girdi katmanı, bir çıktı katmanı ve her biri birbirine bağlı ağırlıklı düğümler veya nöronlar içeren bir veya daha fazla gizli katmandan oluşur. Derin terimi, bir sinir ağındaki gizli katmanların sayısını ifade eder. Geleneksel bir sinir ağında gizli katman sayısı iki ile üç iken, derin öğrenmede bir sinir ağında gizli katman sayısı ikiden daha fazladır. Bu miktar, derin öğrenmenin kapsamlı verileri eğitme yeteneğine sahip olduğunu gösterir [79]. Derin öğrenme algoritması, verileri birkaç katmandan geçirir; her katman, özellikleri aşamalı olarak çıkarma yeteneğine sahiptir ve onu bir sonraki katmana aktarır. Ardışık her katman, bir önceki katmanın çıktısını girdi olarak kullanır. İlk katmanlar, düşük seviyeli özellikleri çıkarır ve sonraki katmanlar, hiyerarşik bir temsil oluşturmak için özellikleri birleştirir [80].

Derin öğrenme, makine öğrenimi teorisinin temsili öğrenim (veya özellik öğrenimi) dalına dayanır. Bir gözlem (örneğin bir görüntü), piksel başına yoğunluk değerlerinin bir vektörü olarak veya daha soyut bir şekilde bir dizi kenar, belirli bir şekle sahip bölgeler vb. gibi birçok şekilde temsil edilebilir. Derin öğrenmenin sağladığı yararlarından biri, denetimsiz veya yarı denetimli özellik öğrenimi ve hiyerarşik özellik çıkarımı için manuel çıkarılan özellikleri verimli algoritmalarla değiştirmektir [81]. Bu açıdan derin öğrenme, yapılandırılmamış verilerle uğraşırken çok sayıda özelliği işleme yeteneği sayesinde mevcut makine öğrenimi tekniklerinden daha iyi performans göstermektedir [82]. Makine öğrenimi algoritmasında problem çözme, problemi farklı parçalara ayırmayı, bunları tek tek çözmeyi ve sonucu elde etmek için birleştirmeyi gerektirir. Buna karşılık derin öğrenme, sorunu uçtan uca çözer [58]. Şekil 3.1'de derin öğrenme kavramı ve makine öğrenimi ile ilişkisi gösterilmiştir. Derin sinir ağı, Makine Öğreniminin nihai ürünü (final product) yerine geçer ve öğrenme kuralı, eğitim verilerinden modeli (derin sinir ağı) üreten algoritma haline gelir. Derin sinir ağının önemi, bilginin hiyerarşik olarak işlenmesi için karmaşık doğrusal olmayan modele ve sistematik yaklaşıma kapı açmış olmasında yatmaktadır [59].



Şekil 3.1. Derin Öğrenme kavramı ve onun Makine Öğrenimi ile ilişkisi [59]

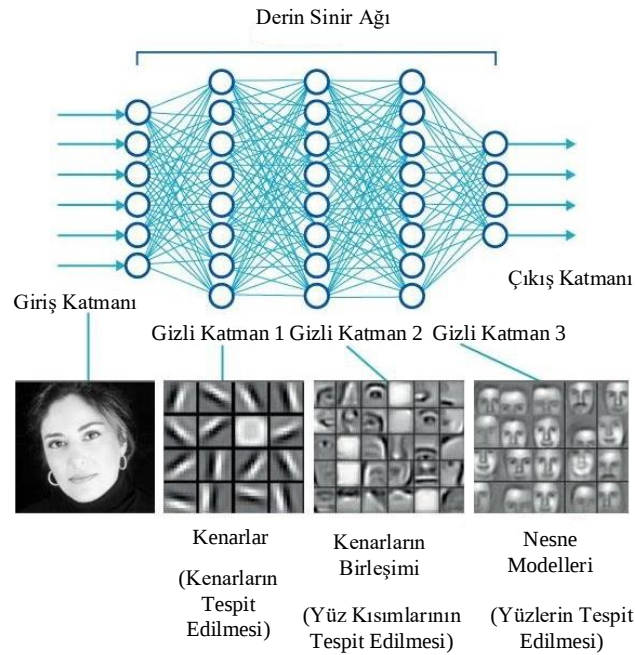
Derin sinir ağlarını uygulayan derin öğrenme teknikleri, yüksek performanslı bilgi işlem sağlayan donanımın artmasının yanı sıra veri miktarının artması ile son yıllarda popüler hale gelmiştir. Sağlık alanında tıbbi görüntü analizi, otonom araçlar, bankacılık, doğal dil işleme, e-ticaret uygulamaları için tavsiye sistemi, kişisel asistan olarak otomatik konuşma tanıma, ses ve video yakalamayı geliştirme dahil olmak üzere birçok alanda derin öğrenme teknikleri kullanılmaktadır [83].

Derin öğrenme yöntemleri, uçtan uca özelliklerin otomatik olarak çıkarılmasını ve sınıflandırılmasını sağlar. Son on yılda, bir tür derin sinir ağı olan ESA, görüntü tanıma ile ilgili farklı görevlerde çığır açan sonuçlar elde etmiştir. İnsan görsel korteksinden esinlenen ESA'lar, görüntü tanıma problemlerinde çeşitli sınıfları ayırt eder [84,85]. Bununla birlikte, ESA büyük miktarda eğitim verisi gerektirir. ESA kullanarak tıbbi görüntü analizi genellikle ESA'yı sıfırdan eğiterek, yeniden eğitmeden önceden eğitilmiş bir ağ kullanarak veya önceden eğitilmiş bir ağı bir hedef veri kümesinde ayarlayarak gerçekleştirilir [86]. Bu zorluğun üstesinden gelmek için ImageNet [87] gibi çok sayıda görüntü üzerinde eğitilmiş ESA modellerinden bilgi aktaran bir transfer öğrenme yaklaşımı geliştirilmiştir. Bu yaklaşımda modelin ilk katmanları kilitlenir veya ayarlanır, son katmanlar ise hedef veri kümesi üzerinde eğitilir. Kaynak ve hedef veri kümeleri farklılık gösterse de model, herhangi bir görüntüyü değerlendirirken genellikle aynı olan düşük seviyeli özellikleri çıkarır, böylece modeli eğitmek ve ek özellikler çıkarmak için büyük veri kümesi gereksinimini ortadan kaldırır. Bu teknik, eğitim süresini ve maliyetini azaltır [88]. Transfer öğreniminde iki yaklaşım kullanılır. Birincisi, aktarılan modelin çıktı katmanını yeni veri setinin sınıf sayısına göre özelleştirmek ve ardından ince ayar yapmaktır. Diğer yaklaşım, aktarılan modeli doğrudan bir özellik çıkarıcı olarak kullanır [89].

3.1. Özellik Çıkarımı

Özellik çıkarımı, verilerin karmaşıklığını azaltmak, algoritmaların çalışmasını öğrenmek ve kalıpları daha görünür hale getirmek için alan bilgisini özellik çıkarıcıların oluşturulmasına dahil etme sürecidir. Bu süreç zaman alıcıdır, uzmanlık açısından zor ve pahalıdır. Uygulanan özelliklerin çoğunun bir uzman tarafından tanımlanması ve ardından makine öğrenimindeki etki alanı ve veri türüne göre manuel kodlanması gerekir. Özellikler piksel değerleri, şekil, doku, konum ve yön olabilir. Makine öğrenimi algoritmasının çoğunun performansı, özelliklerin ne kadar doğru tanımlanıp çıkarıldığına bağlıdır. Derin öğrenme algoritmaları, verilerden yüksek seviyeli özellikler çıkarır. Bu, geleneksel makine öğrenimine göre büyük avantaj sağlayan derin öğrenmenin ana özelliğidir. Sonuç olarak derin öğrenme, her problem için yeni özellik çıkarıcı geliştirme görevini azaltır [58].

Derin özellik çıkarma, ImageNet [87] gibi büyük bir veri kümesinde önceden eğitilmiş bir ESA'dan elde edilen özelliklerin çıkarılmasına dayanır. Özellikler, evrişimli katmanlar kullanılarak çıkarılır ve bir özellik vektörü elde edilir. Veriler daha sonra, yeni bir görev için makine öğrenimi sınıflandırıcılarından (DVM, KNN, vb.) biri ile veya tamamen bağlantılı katmanlardan oluşan sınıflandırıcıya (softmax) iletilir. Tüm modeli yeniden eğitmek yerine yalnızca yeni sınıflandırıcı eğitilir [90]. Bu yaklaşımın temel avantajı, önceden eğitilmiş modeli her eğitim periyodunda bir kez çalıştırmak yerine yeni veriler üzerinde yalnızca bir kez çalıştırmaktır, bu nedenle çok daha hızlıdır. Şekil 3.2'de derin bir sinir ağının giriş katmanına gelen görüntüye ait öznelikleri düşük seviyeden yüksek seviyeye çıkararak görüntü tanımasına bir örnek verilmiştir.



Şekil 3.2. Derin sinir ağlarının aşamalı öznelilik çıkarımı [91]

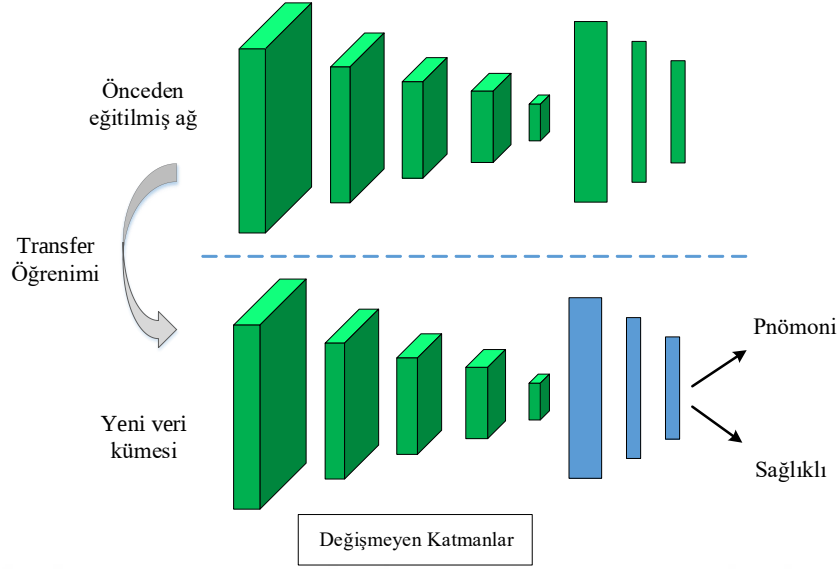
Yukarıda verilen örnekte, girdi katmanına görüntülerin ham verileri verilir. Daha sonra, ilk gizli katman yerel kontrast modellerini belirleyecek, yani renkler, parlaklık vb. temelinde ayırım yapacaktır. Ardından 2. gizli katman yüz özelliğini belirleyecek, yani gözlere, buruna ve dudaklara vb. sabitlenecektir. Ve sonra, bu yüz özelliklerini doğru yüz şablonuna sabitleyecektir. Dolayısıyla, 3. gizli katmanda, yukarıdaki şekilde de görülebileceği gibi, doğru yüzü belirleyecek ve ardından çıktı katmanına gönderilecektir.

3.2. Transfer Öğrenimi

Transfer öğrenimi, öğrenmenin bir modelden diğerine aktarılmasıdır. Transfer öğrenimi ile kapsamlı bir veri kümesi üzerinde eğitilmiş bir modelden öğrenmeyi alıp başka bir veri kümesine tam olarak uyarlamak mümkündür. Bu yaklaşım, modeli daha hızlı eğitmede etkilidir ve aynı zamanda hesaplama açısından daha ucuzdur. Transfer öğrenimi, bir modelin büyük bir veri kümesinden bilgi edinmesini sağlarken aynı zamanda daha az etiketli gözlem içeren daha küçük bir veri kümesindeki performansını artırır. Derin öğrenme modellerinin son birkaç katmanı, veri kümesine özgü bilgiler içerir; bu nedenle, onları başka bir veri kümesinde yeniden eğitmek, modeli yeni verilerle uyumlu hale getirir [92].

Sınıflandırma için kullanılabilecek hâlihazırda eğitilmiş çok sayıda transfer öğrenme modeli bulunmaktadır. Önceden eğitilmiş bir model, başka birinin benzer nitelikteki bir sorunu ele almak için geliştirdiği bir modeldir. Sınıflandırma sorunları için, onları eğitmek üzere çok büyük veri seti kullanılır. Benzer bir sorunu ele almak için sıfırdan bir model oluşturmak yerine, başka bir sorun üzerinde eğitilen model bir başlangıç noktası olarak kullanılabilir [93]. Hedef etki alanından veri kullanılabilirliğine bağlı olarak, tüm katmanlar yeniden eğitilebilir veya yalnızca son (tamamen bağlı) katman yeniden eğitilebilir. Bu yaklaşım, sinir ağlarının nispeten daha küçük veri kümeleri kullanılarak yeni görevler için eğitilmesine izin verir, çünkü yararlı düşük seviyeli özellikler kaynak etki alanı verilerinden öğrenilir [94].

Herhangi bir derin öğrenme modelinin performansı mevcut eğitim verilerine dayanır. Daha büyük veri kümeleri daha iyi öğrenme ve daha iyi performans sağlar. Ancak, gizlilik endişeleri nedeniyle sağlık uygulamalarında büyük eğitim verilerinin kullanılabilirliği her zaman mümkün değildir. Bu yetersiz veri sorununu çözmek için transfer öğrenme yöntemi uygulanmaktadır. Transfer öğrenme, daha büyük bir veri kümesi üzerinde zaten eğitilmiş olan ve daha sonra yeni hedef görev için değiştirilen ağırlıkları kullanır [92,95].



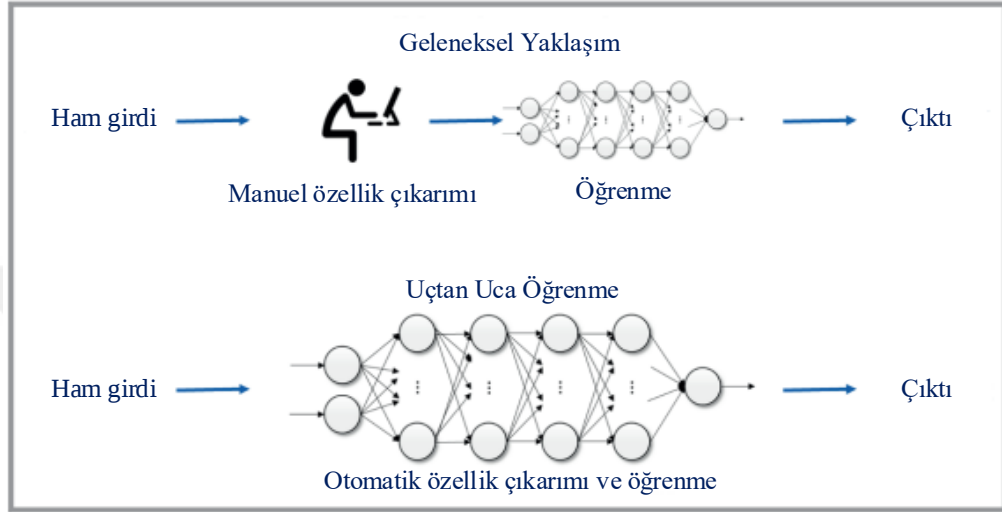
Şekil 3.3. Pnömoni sınıflandırılmasına ait transfer öğrenme yöntemi gösterimi [95]

Şekil 3.3’de pnömonili CXR’leri sağlıklı CXR’lerden ayırt ederken transfer öğrenme yönteminin uygulanmasına örnek gösterilmiştir. ImageNet veri tabanında önceden eğitilen ESA modellerinin son 3 katmanı yani sınıflandırmaya ait etiketlerin bulunduğu katmanların değiştirilmesiyle ESA yeni veri kümesine uyum sağlayarak başarılı bir sınıflandırma performansı göstermektedir.

3.3. Uçtan Uca Öğrenme

Uçtan uca derin öğrenme, herhangi bir manuel özellik çıkarımı olmaksızın doğrudan ham girdi verilerinin girdi olarak kullanılarak karmaşık görevler için tek bir sinir ağının eğitildiği bir makine öğrenimi tekniği olarak tanımlanmaktadır. Büyük ölçekli veri kümelerinin oluşturulması sayesinde, uçtan uca derin öğrenme konuşma tanıma, makine çevirisi, yüz algılama vb. gibi çeşitli alanlarda devrim yaratmıştır. Geleneksel makine öğreniminde, eğitim veri hattı en az 2 aşamadan oluşur: İlk aşamanın amacı, ham girdi verileri göz önüne alındığında ayırt edici özellikler üretmektir. Bu, her görev için en alakalı özellikleri belirlemek üzere alana özgü bilgi kullanılarak yapılır. Bir sonraki aşama, çıkarılan özellikleri alır ve bazı geleneksel makine öğrenimi algoritmalarını kullanarak tahminleri oluşturur. Başarılı olmasına rağmen, bu prosedür çok zaman alıcıdır ve alana özgü çok fazla bilgi gerektirir. Mevcut veri kümelerinin boyutundaki son artış, manuel çıkarılan özellikleri manuel olarak çıkarmadan girdi çıktı eşlemesini doğrudan verilerden öğrenmeyi amaçlayan uçtan uca derin öğrenmenin yükselişini sağlamıştır. Yani, uçtan uca öğrenme ile derin öğrenme algoritmaları girdi verilerini alır ve bunlar (genellikle büyük) bir sinir ağından geçirilir. Ardından, ağ girdi verilerini işler ve daha sonra tahminler oluşturmak için kullanılan ilgili özellikleri otomatik olarak çıkarır. Tüm prosedür, manuel mühendisliğe ihtiyaç duyulmadan

gerçekleştirilir ve gereken zaman ve çaba miktarını azaltır. Bu tekniğin başarısı, eğitim sırasında kullanılan muazzam miktarda veride ve sinir ağlarının sadece çok sayıda veri üzerinde eğitim yaparak yüksek seviyeli özellikleri öğrenme yeteneğinde yatmaktadır [96,97]. Şekil 3.4’de geleneksel makine öğrenimi ile derin öğrenme teknikleri arasındaki girdi çıktı ilişkisi karşılaştırılmıştır.



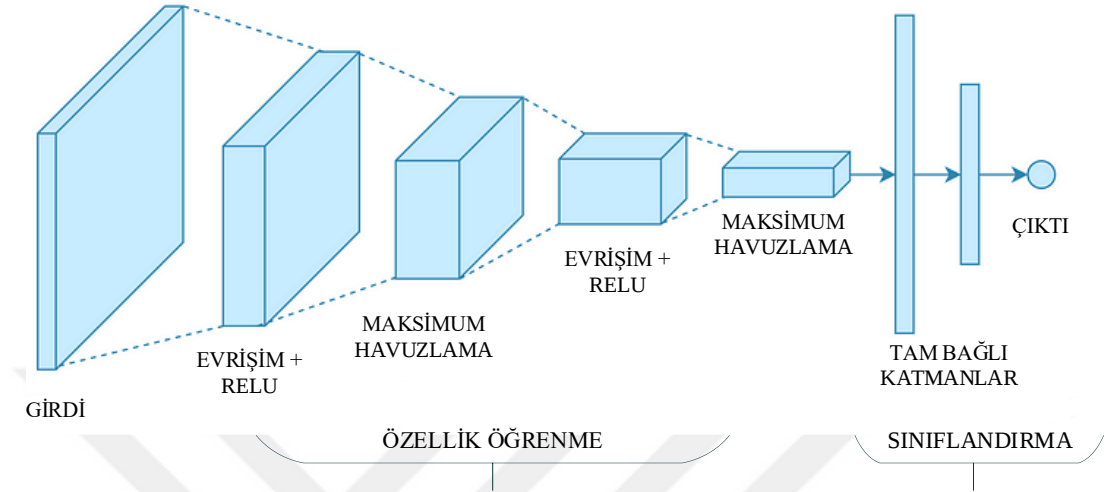
Şekil 3.4. Manuel özellik çıkarma ile karşılaştırıldığında uçtan uca öğrenme kavramının gösterimi [98]

Hem derin öğrenme hem de geleneksel makine öğrenimi, Şekil 3.4’de gösterildiği gibi girdi ve çıktı arasındaki karmaşık ilişkiyi modellemek için veriye dayalı yapay zekâ teknikleridir. Yüksek hiyerarşik yapıya ek olarak, derin öğrenme ayrıca özellik öğrenme, model oluşturma ve model eğitimi açısından geleneksel makine öğrenimine göre ayırt edici niteliklere sahiptir. Şekil 3.4’de gösterildiği gibi, uçtan uca derin öğrenme yaklaşımı, manuel özellik seçimi ihtiyacını ortadan kaldırır. Derin öğrenme, farklı çekirdekler seçerek veya uçtan uca optimizasyon yoluyla parametreleri ayarlayarak özellik öğrenmeyi ve model oluşturmaya tek bir modelde birleştirir. Geleneksel makine öğrenimi ise özellik çıkarma ve model oluşturma işlemlerini ayrı bir şekilde gerçekleştirir ve her modül adım adım oluşturulur [99].

3.4. Evrimsel Sinir Ağları (ESA)

Evrimsel Sinir Ağı ya da kısaca ESA, kenarlar, doku, renk ve şekiller gibi görüntülerdeki örüntüleri tanımlamak için öğrenilebilir filtreler veya çekirdekler kullanan bir derin öğrenme algoritmasıdır [28]. Derin bir sinir ağı, paralel olarak çalışan ve biyolojik sinir sistemlerinden esinlenen basit öğeler kullanarak birden fazla doğrusal olmayan işlem katmanını birleştirir. ESA’nın çıkış katmanı genellikle çok sınıflı sınıflandırma sinir ağını kullanır. ESA, özellik çıkarıcıyı manuel olarak tasarlamak yerine eğitim sürecine dahil eder. ESA’nın özellik çıkarıcısı, ağırlıkları eğitim süreci ile belirlenen özel tür sinir ağlarından oluşmaktadır. ESA’nın manuel

özellik çıkarma tasarımı otomatikleştirilmiş sürece dönüştürmesi, birincil özelliği ve avantajıdır. Şekil 3.5’de ESA’nın tipik mimarisi yer almaktadır. Bir ESA, bu grafik gösterime benzer bir mimari etrafında inşa edilmiştir [59].



Şekil 3.5. ESA’nın tipik mimarisi [100]

Şekil 3.5’den de görüleceği üzere bir ESA, özellik çıkarma ağı ve sınıflandırma ağının seri bağlantısından oluşur. Giriş görüntüsü özellik çıkarma ağına girer. Çıkarılan özellik sinyalleri sınıflandırma sinir ağına girer. Sınıflandırma sinir ağı daha sonra görüntünün özelliklerine göre çalışır ve çıktıyı üretir. Özellik çıkarma sinir ağı, evrişim katmanı, aktivasyon fonksiyonu ve havuzlama katmanı çiftlerinden oluşur. Evrişim katmanı, evrişim işlemini kullanarak görüntüyü dönüştürür. Bir dijital filtre koleksiyonu olarak düşünülebilir. Havuzlama katmanı komşu pikselleri tek bir pikselde birleştirir. Bu nedenle, havuzlama katmanı görüntünün boyutunu ve ağı öğrenmesi gereken parametre sayısını azaltır. Bu işlemler onlarca veya yüzlerce katman üzerinde tekrarlanır ve her katman farklı özellikleri tespit etmeyi öğrenir. Özellik tespitinden sonra, ESA mimarisi sınıflandırmaya geçer. Sonndan bir önceki katman, K ’nın ağı tahmin edebileceği sınıf sayısı olduğu K boyutlu bir vektör çıkaran tam bağlı bir katmandır. Bu vektör, sınıflandırılan herhangi bir görüntünün her bir sınıfı için olasılıkları içerir. ESA mimarisinin son katmanı, sınıflandırma çıktısını sağlamak için bir softmax işlevi kullanır [100,101].

ESA Katmanları

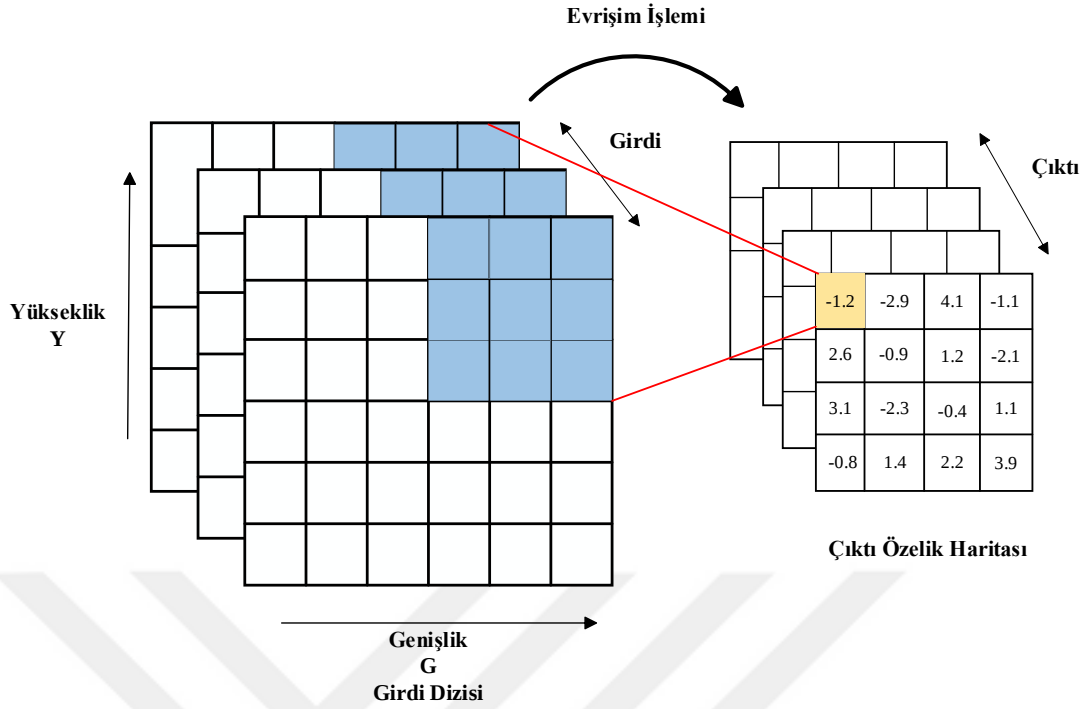
Bu kısımda genel bir ESA mimarisini oluşturan katmanlardan sırasıyla bahsedilmiştir. Ayrıca bu tez çalışması kapsamında geliştirilen pnömoni tespit modelinde kullanılan toplu normalizasyon katmanından da bahsedilmiştir.

3.4.1. Giriş Katmanı

Giriş katmanı ESA'nın ilk katmanını oluşturur. Giriş katmanında, giriş verileri ağa ham olarak verilir. Girdi bir görüntü ise eğer görüntüleri her bir elemanın bir pikselin renk yoğunluğunun değeri olduğu tensörler (çok boyutlu matrisler) olarak temsil eder [100]. Kullanılan bir ağda giriş katmanındaki görüntünün boyutu önemlidir. Büyük girdi boyutları çok fazla bellek kullanımına neden olur ve parametre sayısını artırır, bu da ağı yavaşlatabilir, ancak daha fazla öznetelik çıkarılabileceği için daha iyi sonuçlar verebilir. Ayrıca küçük girdi boyutları ile parametre sayısı azalacağından ağ daha hızlı sonuç verebilir ancak öznetelik çıkarımı daha az olduğu için ağın performansı düşebilir [102].

3.4.2. Evrişim Katmanı

Evrişim katmanları, evrişimsel sinir ağı yapısının ana yapı taşlarıdır. Evrişim katmanı özellik çıkarma katmanı olarak da bilinir çünkü ağa verilen görüntünün özellikleri bu katmanda çıkarılır. Evrişim katmanında çeşitli filtreler (kernel) kullanılarak görüntünün özellik haritası oluşturulur [103]. Kullanılan filtreler 2×2 , 3×3 , 5×5 gibi farklı boyutlarda olabilmektedir. Bu filtreler ile görüntüye evrişim işlemi uygulanır. Evrişim, bir matrisin (filtre) diğerinin (görüntü) üzerinde kaydırıldığı ve üst üste gelen değerleri arasında eleman bazında çarpımların gerçekleştirildiği, ardından tüm çarpımların toplandığı (üst üste binen her konumda, bir sonrakine geçmeden önce) ve bir çıktı elde edildiği doğrusal bir işlemdir [100]. ESA'ların eğitimi sırasında eğitim setindeki her öğrenme iterasyonu ile filtrelerin katsayıları değişerek ağı özneteliklerinin belirlenmesinde verinin hangi bölgelerinin önemli olduğu belirlenir [102]. Giriş verilerindeki özellikler, bir modelin bir göreve göre öğrenmesini ve genelleştirmesini sağlayan anlamlı bilgiler içeren köşeler, lekeler, kenarlar, eğriler, şekiller, dokular, yoğunluk örüntüleri, küresel bağlamdaki soyut ve karmaşık ayrıntılar vb. anlamına gelir [100]. Şekil 3.6'da girdi dizisine evrişim işlemi uygulandıktan sonra elde edilen özellik haritasının bir gösterimi bulunmaktadır.



Şekil 3.6. Evrişimli işleminin bir gösterimi [100]

Her filtre tarafından üretilen evrişimsel çıktılar (bir sapma değeri eklenmiş) sonraki katmanlara aktarılmak üzere bir matriste eşlenir. Bu matrise özellik haritası denir. Çünkü bir filtre tarafından çıkarılan özellikleri saklar. Her filtre kendi özellik haritasını üretir, bu nedenle bir katman tarafından üretilen tek kanallı özellik haritalarının sayısı, içinde bulunan filtrelerin sayısına eşittir. Bir özellik haritası, nöronların çıkışlarını (filtreler yerine) gösteren aktivasyon fonksiyonları uygulandıktan sonra alındığında aktivasyon haritası olarak adlandırılabilir [100].

3.4.3. Toplu Normalizasyon Katmanı

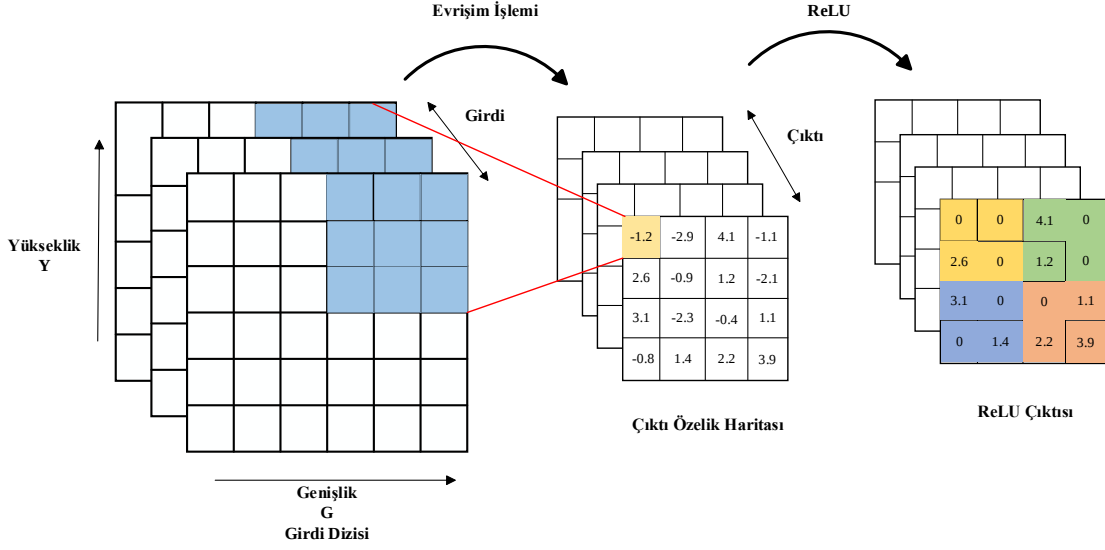
Bir toplu normalizasyon katmanı (Batch Normalization / BN), her kanal için bağımsız olarak tüm gözlemler genelinde mini bir veri grubunu normalleştirir. BN, derin öğrenme sinir ağındaki bir katmanın girdilerini otomatik olarak standartlaştırmak için tasarlanmış bir tekniktir. BN bir kez uygulandığında, bir sinir ağının eğitim sürecini önemli ölçüde hızlandırma etkisine sahiptir ve bazı durumlarda modelin performansını artırır. Teknik olarak, katman girdileri standartlaştırılacak şekilde dönüştürecek, yani sıfır ortalamaya ve bir standart sapmaya sahip olacaklardır, eğitim sırasında katman her girdi değişkeni için istatistikleri takip edecek ve bunları verileri standartlaştırmak için kullanacaktır. BN, bir modelin çoğu noktasında ve çoğu derin öğrenme sinir ağı türünde kullanılabilir. Yazarlar çalışmalarında [104] toplu normalleştirme katmanlarının eklenmesinin 14 kat daha az eğitim adımına ihtiyaç duyan ancak yine de aynı doğruluğu sağlayan bir ağ oluşturduğunu göstermiştir. Evrişimsel sinir ağının eğitimini hızlandırmak ve ağı başlatılma

performansını iyileştirmek için evrişim katmanları ile ReLU katmanları gibi doğrusal olmayan katmanlar arasında toplu normalleştirme katmanları kullanılabilir. [104,105].

3.4.4. Etkinleştirme (Aktivasyon) katmanı

Aktivasyon katmanları, evrişimsel sinir ağlarının bir diğer önemli yapı taşıdır. Aktivasyon katmanları sinir ağlarına doğrusal olmayan bir özellik kazandırmak için kullanılır. Aktivasyon katmanı, evrişim katmanından sonra gelir. Genellikle doğrusal olmayan bu katman üretilen çıktıyı ayarlamak ve sınırlamak için kullanılır [106]. Bu katmanın kullanım amacı evrişim katmanından sonra oluşan doğrusal ağdaki negatif değerleri sıfıra indirerek ağı doğrusal olmayan bir forma sokmak ve ağı daha hızlı öğrenmesini sağlamaktır. Aktivasyon fonksiyonu olarak da adlandırılır. Tanh, sigmoid, ELU gibi birçok farklı aktivasyon fonksiyonu vardır ancak ESA bağlamında en yaygın kullanılan fonksiyon Doğrultulmuş Doğrusal Birim (Rectified Linear Unit / ReLU)'dir. ReLU'nun matematiksel yorumu $f(x) = \max(0, x)$ şeklindedir. Girdinin tüm değerlerini pozitif sayılara dönüştürür. Ana avantajlarından biri, giriş negatif olduğunda sıfır çıkışa sahip olması ve giriş pozitif olduğunda girişle aynı değere sahip olmasıdır. Dolayısıyla aktivasyon fonksiyonlarının gradyanları her zaman ya 1 ya da 0'dır ve bu da daha derin sinir ağlarında kaybolan gradyan [107] sorunundan kaçınmaya yardımcı olur. Kaybolan gradyan sorunu, aktivasyon fonksiyonunun gradyanı sinir ağlarının kaldırabileceğinden daha küçük hale geldiğinde ortaya çıkar. ReLU'da fonksiyonun gradyanı ya sıfır olacaktır (girdi sıfırdan küçük olduğunda) ya da yeterince büyük bir değer olacaktır (girdi sıfırdan büyük olduğunda). Böylece ReLU, aktivasyon fonksiyonlarının neden olduğu kaybolan gradyan sorunundan kaçınmaya yardımcı olur [108]. Düşük hesaplama yükü, ReLU'nun diğerlerine göre bir diğer avantajıdır [109]. Girdi değerlerine ReLU aktivasyon fonksiyonunun uygulanması Denklem 3.1'e göre hesaplanır [102].

$$f(x) = \begin{cases} 0 & \text{eğer } x < 0 \\ x & \text{eğer } x \geq 0 \end{cases} \quad (3.1)$$

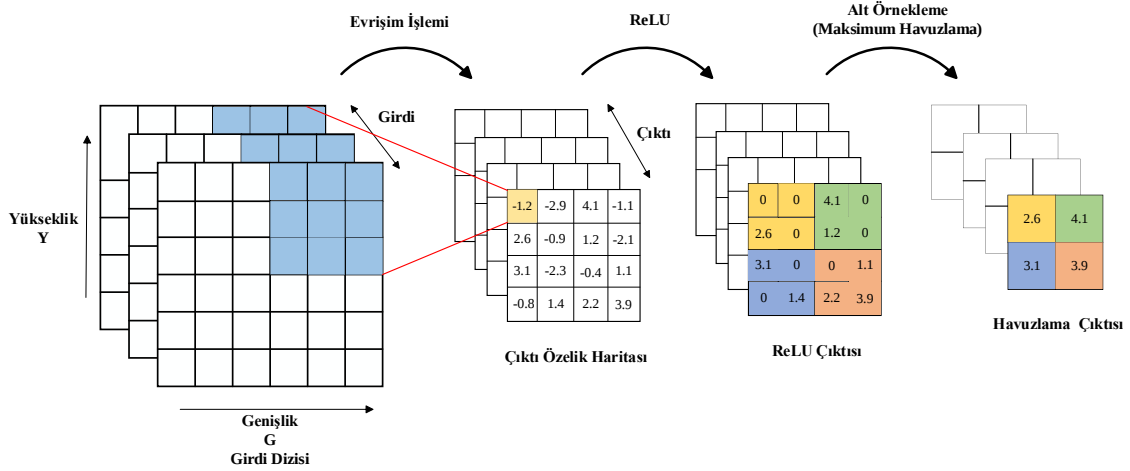


Şekil 3.7. Aktivasyon fonksiyonu olarak ReLU kullanımına bir örnek [100]

Şekil 3.7’de girdi dizisinin evrişim katmanından geçip bir öznetelik haritası oluştuktan sonra aktivasyon fonksiyonu olarak ReLU kullanımına bir örnek verilmiştir. Özellik haritasındaki düğümler, aktivasyon fonksiyonundan geçirilir. Pozitif değerler aynı değerle çıkarken, negatif değerler bir eşikte karşılandığını ve 0 değer olarak çıkar. Böylelikle doğrultulmuş bir öznetelik haritası elde edilir.

3.4.5. Havuzlama Katmanı

Havuzlama işlemi veya havuzlama katmanı genellikle ESA’larda parametre sayısını ve hesaplama yükünü azaltmak için kullanılır. Derinliği etkilemeden bir sonraki evrişim katmanı için girdi boyutunu azaltır. Yani havuzlama sonrasında giriş verisinin yükseklik ve genişlik değerleri azalırken derinlik yani kanal sayısı değişmez. Havuzlama katmanları, giriş dizilerinin aşağı örnekleme için kullanılır. Giriş matrisini alt matrislere bölerek ve her bir alt matrisi temsil edecek bir değer seçerek aşağı örnekleme gerçekleştirir. Genel olarak iki farklı havuzlama katmanından bahsedilebilir [102]. Bunlar maksimum havuzlama ve ortalama havuzlama katmanlarıdır. Maksimum havuzlamada alt matrisi temsil etmek için alt matristeki maksimum değer alınırken, ortalama havuzlamada alt matrisin ortalaması alınır. Scherer ve diğerleri [110]’da maksimum havuzlamanın nesne sınıflandırma görevlerinde ortalama havuzlamadan daha etkili olduğunu göstermiştir.

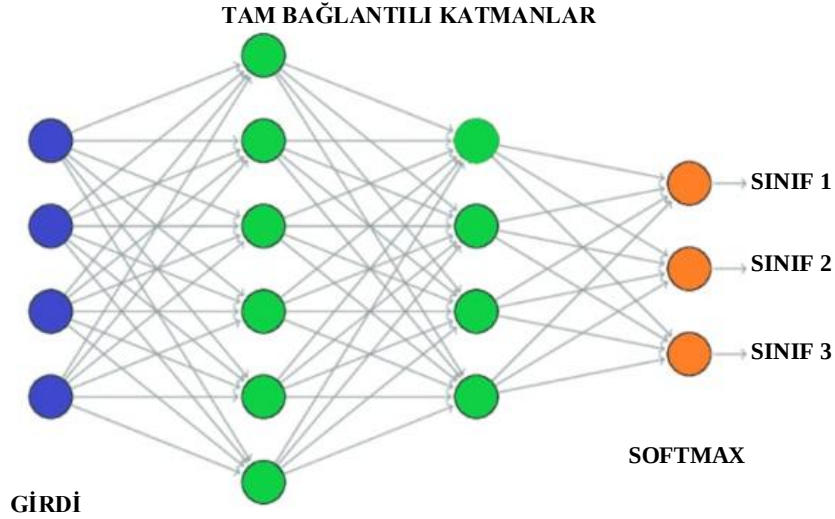


Şekil 3.8. Maksimum havuzlamaya örnek bir gösterim [100]

Şekil 3.8’de 2 boyutunda ve 2 adımlı bir maksimum havuzlama örneği yer almaktadır. Maksimum havuzlama kullanan bir ESA, özelliklerin uzamsal düzenini koruyarak boyutları azaltılmış özellik haritalarıyla sonuçlanacaktır. Bu şekilde, 2 boyutunda ve 2 adımlı bir havuzlama çekirdeği, orijinal boyutların yarısı kadar küçültülmüş bir özellik haritası elde eder.

3.4.6. Tam Bağlantılı Katman

Bir ESA’da birbirini izleyen evrişim, ReLU ve havuzlama katmanlarından sonra tam bağlantı katmanı gelir. Evrişim ve havuzlama katmanlarından gelen çıktı, girdi görüntüsünü sınıflandırmak veya segmentasyon gibi diğer görevleri gerçekleştirmek için standart bir sinir ağı mimarisi kullanan bir veya daha fazla tamamen bağlı katmandan (yoğun katmanlar olarak da bilinir) geçirilir. Tam bağlantı katmanı için özellik haritası 1 boyutlu bir özellik vektörüne dönüştürülür, buna düzleştirme denir [111]. Bu katman, bir önceki katmanın tüm alanlarına bağlı olduğu için tam bağlantı katmanı olarak adlandırılır. Tam bağlantı katmanı, görüntüleri farklı kategorilerde sınıflandırmak için kullanılır. Bir ESA’daki evrişimli ve havuzlama katmanlarının çıktısı, giriş görüntüsünün önemli özelliklerini temsil eden bir dizi üst düzey özelliktir. Tamamen bağlantılı katmanların rolü, girdi görüntüsünün sınıfı veya etiketi hakkında bir tahmin yapmak için bu öğrenilmiş özellikleri kullanmaktır. Tam bağlantı katmanları yüksek seviyeli (karmaşık) özellikleri yakalama ve öğrenme ve bunları bir şekilde anlamlı sonuçlara dönüştürme yeteneğine sahiptir. Yoğun bağlantılar nedeniyle çok sayıda parametreye ve yüksek hesaplama yüküne sahiptirler. Bir ağda tam bağlantı katmanı yoksa karmaşık özellikleri işleme yeteneğinden yoksun olacak ve genel performansı önemli ölçüde düşürecektir [100].

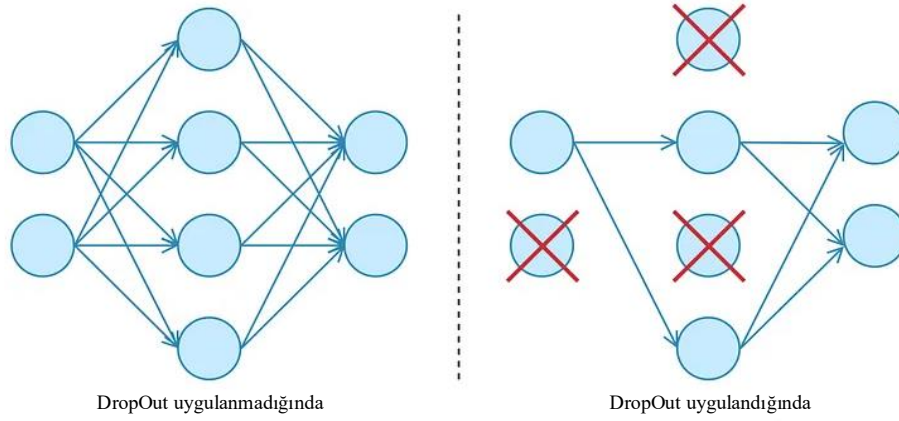


Şekil 3.9. Tam bağlantılı katman [112]

Şekil 3.9’da gösterildiği gibi komşu katmanlar arasındaki nöronlar tamamen bağlıdır.

3.4.7. DropOut Katmanı

DropOut, derin sinir ağları için en popüler düzenli hale getirme tekniğidir. Genellikle, tüm özellikler tam bağlantı katmanına bağlandığında, eğitim veri setinde aşırı öğrenmeye neden olabilir. Aşırı öğrenme, belirli bir model eğitim verileri üzerinde çok iyi çalıştığında ortaya çıkar ve yeni bir veri üzerinde kullanıldığında modelin performansında olumsuz bir etkiye neden olur. Bu sorunun üstesinden gelmek için, eğitim süreci sırasında birkaç nöronun sinir ağından çıkarıldığı ve modelin boyutunun küçültüldüğü bir bırakma katmanı kullanılır. Eğitim süresi boyunca, her iterasyonda, bir nöron geçici olarak bırakılır veya p olasılığı ile devre dışı bırakılır. Bu, bu nöronun tüm giriş ve çıkışlarının mevcut iterasyonda devre dışı bırakılacağı anlamına gelir. Devre dışı bırakılan nöronlar her eğitim adımında p olasılığı ile yeniden örneklenir, böylece bir adımda devre dışı bırakılan bir nöron bir sonraki adımda aktif olabilir. Bırakma oranı 0,3’ü geçtiğinde, düğümlerin %30’u sinir ağından rastgele çıkarılır. Bırakma, ağı daha basit hale getirerek aşırı öğrenmeyi önlediği için bir derin öğrenme modelinin performansını artırır. Eğitim sırasında nöronları sinir ağlarından düşürür [113]. Ağı daha iyi performans göstermesinin nedeni, devre dışı bırakmanın ağı az sayıda nörona çok fazla bağımlı olmasını engellemesi ve her nöronu bağımsız olarak çalışmaya zorlamasıdır. Şekil 3.10’da nöronlara DropOut uygulandığı ve uygulanmadığı zamanki durumları gösterilmektedir.



Şekil 3.10. Dropout'a örnek [114]

3.4.8. Sınıflandırma Katmanı

ESA'nın önemli bir bileşenidir ve işlevi, ağ tarafından çıkarılan ve öğrenilen özelliklere dayanarak kararlar vermektir. Çıkış katmanının tüm nöronları önceki tam bağlantı katmanı ile tamamen bağlantılıdır. ESA'nın çıkış katmanındaki nöron sayısı, sınıflandırma görevleri için sınıf sayısına ve regresyon görevleri için regresyon çıktılarına eşittir. Her bir nöron her bir sınıfa karşılık gelir ve bir nöron tarafından gösterilen en yüksek aktivasyon, ağ tarafından yapılan ilgili sınıfın tahminini gösterir [100]. Genel olarak, Softmax Aktivasyon Fonksiyonu çıkış katmanında kullanılır, çünkü görüntü sınıflandırması yapılacak tüm sınıflar için olasılık dağılımı verir. En yüksek olasılığa sahip olan sınıf, modelin tahmini olarak seçilir. Softmax işlevi, rastgele gerçek değerli puanlardan oluşan bir vektörü alır ve bunu sıfır ile bir arasında toplamı 1 olan değerlerden oluşan bir vektöre sıkıştırır [106].

3.5. Kullanılan Derin Öğrenme Algoritmaları

Bu tez çalışması kapsamında pnömoni tespiti için kullanılan derin öğrenme yaklaşımları için seçilen derin öğrenme mimarileri bu kısımda detaylı olarak anlatılmıştır.

3.5.1. AlexNet

Alex Krizhevsky ve ark., LeNet'e kıyasla daha derin ve daha geniş bir ESA modeli önerdiler ve 2012'de ImageNet Büyük Ölçekli Görsel Tanıma Mücadelesi (ILSVRC) olarak adlandırılan görsel nesne tanıma yarışmasını kazandılar. Görüntü giriş katmanı $227 \times 227 \times 3$ görüntü boyutu gerektiren ağ, 8 katman derinliğinde toplamda 25 katmandan oluşmaktadır. AlexNet, beş evrişim katmanı, 4096 çıktı ile iki tam bağlı katman ve görüntüleri 1000 nesne kategorisinde sınıflandırabilen Softmax çıkış katmanından oluşmaktadır. Her bir evrişim katmanından sonra, AlexNet ağ boyutunu azaltmak için maksimum havuzlamaya sahiptir. Doğrusal olmayan özellikleri

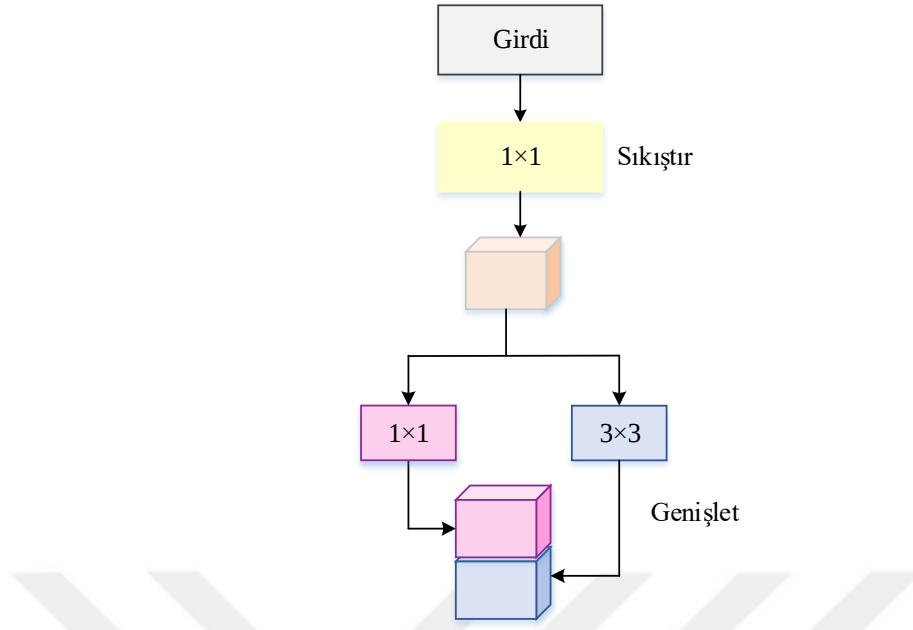
dahil etmek için, aktivasyon fonksiyonu olarak LeNet'teki *tanh* fonksiyonunun yerini ReLU almıştır. ReLU fonksiyonunun eğitim sürecini hızlandırmak, iterasyon sayısını azaltmak ve aşırı uyumu etkili bir şekilde önlemek gibi birçok avantajı vardır. AlexNet'te bırakma katmanı, modelin eğitim sürecini yani aşırı uyum derecesini azaltmak için kullanılır, nöronlar belirli bir olasılıkla durdurulur, böylece yerel düğümlere olan bağımlılık azalır ve modelin genelleme yeteneği iyileştirilir [84]. Şekil 3.11'de AlexNet'in temsili bir gösterimi yer almaktadır.



Şekil 3.11. AlexNet [115]

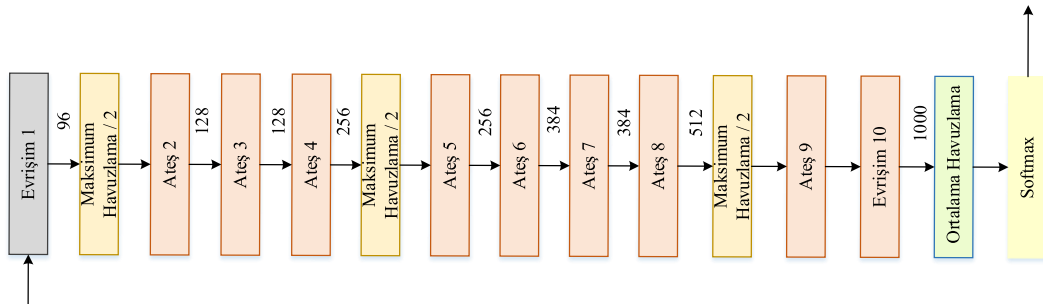
3.5.2. SqueezeNet

SqueezeNet, 18 katman derinliğindedir ve 68 katmanlıdır. Ağ mimarisi bağımsız bir evrişim katmanı (conv1) ile başlar, ardından 8 ateşleme modülü (fire2-9) takip eder ve son bir evrişim katmanı (conv10) ile biter. Bir ateşleme modülü, 1×1 ve 3×3 evrişim filtrelerinin bir karışımına sahip bir genişletme katmanına beslenen (yalnızca 1×1 filtreleri olan) bir sıkıştırılmış evrişim katmanından oluşur. Ateşleme modülü başına filtre sayısı, ağın başından sonuna kadar kademeli olarak artırılır. SqueezeNet, conv1, fire4, fire8 ve conv10'dan sonra 2'lik bir adımla maksimum havuzlama gerçekleştirir. Kısaca 2 evrişimsel katmana, 8 ateşleme modülüne ve 4 havuzlama katmanına sahiptir. $227 \times 227 \times 3$ boyutunda bir giriş görüntüsü alır. SqueezeNet mimarisi, parametre sayısını azaltmak ve yüksek doğruluk elde etmek için çeşitli iyileştirmelere sahiptir [116]. SqueezeNet mimarisi, verimlilik (daha az parametreye ve daha küçük model boyutuna sahip olmak) ile ilgili temel kaygılarla bilgisayarlı görme görevleri için önerilen derin bir ESA'dır [117].



Şekil 3.12. SqueezeNet’de kullanılan ateşleme modülü [117]

SqueezeNet mimarisinin temel yapı taşı Şekil 3.12’de gösterilen ateşleme modülüdür. Modül bir sıkıştırma aşaması ve bir genişletme aşaması içerir. Sıkıştırma aşaması, bir dizi 1×1 filtre ve ardından bir ReLU aktivasyonu uygular. Öğrenilen sıkıştırma filtrelerinin sayısı her zaman giriş hacminin boyutundan daha küçüktür. Sonuç olarak, sıkıştırma aşaması bir boyutsallık azaltma işlemi olarak düşünülebilir ve aynı zamanda giriş kanalları arasındaki piksel korelasyonlarını yakalar. Sıkıştırma aşamasının çıktısı, 1×1 ve 3×3 konvolüsyonların bir kombinasyonunun öğrenildiği genişletme aşamasına beslenir. Daha büyük 3×3 filtreler pikseller arasındaki uzamsal korelasyonları yakalamak için kullanılır. Genişletme aşamasının çıktıları kanal boyutu boyunca birleştirilir ve ardından bir ReLU aktivasyonu ile değerlendirilir. Orijinal makale, genişletme aşamasında n , 1×1 ; ve n , 3×3 filtre kullanılmasını önermiştir; burada n , sıkma aşamasında kullanılan filtre sayısından 4 kat daha büyüktür. Tüm SqueezeNet mimarisi, geleneksel evrişim katmanları, maksimum havuzlama, ateşleme modülleri ve sonunda bir ortalama havuzlama katmanı istifleterek oluşturulmuştur. Modelde tam bağlantılı katman bulunmamaktadır [117,118]. SqueezeNet mimarisi Şekil 3.13’de yer almaktadır.



Şekil 3.13. SqueezeNet mimarisi [118]

3.5.3. VGG16 ve VGG19

Simonyan ve Zisserman tarafından önerilen VGG (Visual Geometry Group), bir ESA mimarisidir ve 2014 yılında ILSVRC yarışmasını kazanmak için kullanılmıştır. VGG16, ImageNet veri setinde %92,7 ile ilk 5 test doğruluğuna ulaşmıştır. Toplamda ağ 16 katman derinliğindedir. VGG16'da, daha önceki modellerde kullanılan büyük çekirdek boyutlu filtrelerin yerine birbiri ardına çoklu 3×3 çekirdek boyutlu filtreler kullanılmıştır. Birden fazla çekirdek katmanı, sinir ağının derinliğinin artmasına neden olur. Bu, sinir ağının daha karmaşık özellikleri ve kalıpları anlamasını ve tanımasını sağlar. Bu mimarinin en önemli özelliği, çok sayıda hiper parametreye sahip olmak yerine, evrişim katmanlarında basit 3×3 boyutundaki çekirdeklere ve maksimum havuz katmanlarında 2×2 boyutuna sahip olmalarıdır. VGG19, kısaca 19 katmandan (16 evrişim katmanı, 3 tam bağlantılı katman, 5 maksimum havuzlama katmanı ve 1 softmax katmanı) oluşan VGG modelinin bir çeşididir ve esas olarak görüntü sınıflandırması için kullanılan bir ESA'dır. Temel mimarisi VGG16'ninkine benzemektedir. Bu ağlara girdi olarak sabit boyutta (224×224) RGB görüntü verilmiştir, bu da giriş matrisin ($224 \times 224 \times 3$) şeklinde olduğu anlamına gelmektedir [119]. Şekil 3.14'de VGG16 mimarisi yer alırken Şekil 3.15'de VGG19 mimarisi yer almaktadır.



Şekil 3.14. VGG16 [115]

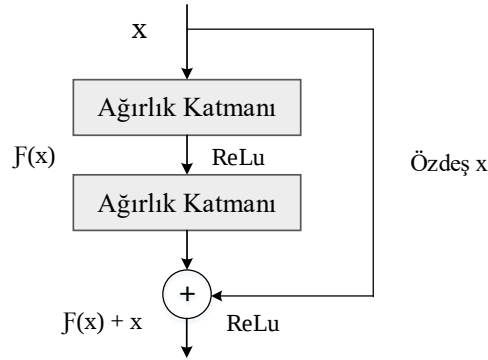


Şekil 3.15. Vgg19 [115]

3.5.4. ResNet50

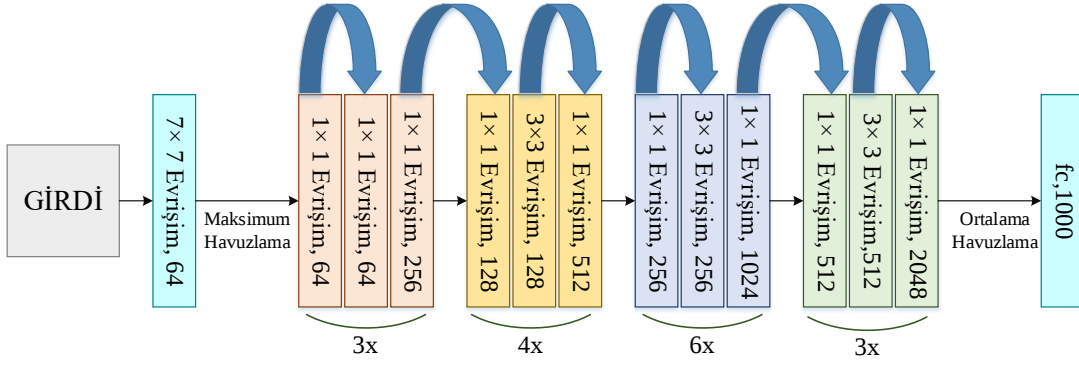
ResNet, Residual Network (Artık Ağ) anlamına gelmektedir. He ve ark., tarafından 'Görüntü Tanıma için Derin Artık Öğrenme' adlı makalesinde tanıtılan model, ImageNet veri setinde %3,57 ilk 5 hatası elde etmiş ve 2015'de ILSVRC yarışmasını kazanmıştır [120]. Ağın evrişimli katmanları 3×3 filtrelerle sahiptir ve alt örnekleme doğrudan 2 adımlı evrişimli katmanlar tarafından yapılır. Ağın son katmanı, sırasıyla ReLU ve softmax aktivasyon fonksiyonlarını kullanan 256 ve

iki kanallı tam bağlantılı bir katmandır. Bu mimari, derin öğrenme eğitimindeki ikilemleri ortadan kaldırmak için oluşturulmuştur, çünkü derin öğrenme eğitimi genel olarak oldukça fazla zaman alır ve belirli sayıda katmanla sınırlıdır. ResNet, hata yüzdesini en aza indirirken daha fazla sinir katmanına sahip derin sinir ağlarının verimliliğini artırır. Başka bir deyişle, atlama bağlantıları önceki katmanlardan gelen çıktıları yığılmış katmanların çıktılarına ekleyerek daha önce mümkün olandan çok daha derin ağların eğitilmesini mümkün kılar. Ayrıca, ağları eğitme yeteneği daha iyidir. ResNet’de, derin sinir ağlarında meydana gelen azalan doğruluk ve kaybolan gradyan sorunlarını düzeltmek için kısayol bağlantıları kullanılır. Bu bağlantılar, ağın eğitim için alakasız olduğunu düşündüğü katmanları atlamasına izin verir. Bu, eğitim hatasını azaltır ve ağın diğer ağlara kıyasla daha hızlı yakınsamasına yardımcı olur. ResNet’in ResNet50, ResNet18, ResNet34, ResNetV2 vb. gibi birden çok çeşidi vardır. ResNet50, 50 kat derinliğindedir ve artık öğrenme çerçevesine sahiptir. 48 evrişimsel katmana, maksimum havuzlama katmanına, ortalama havuzlama katmanına ve tam bağlantılı katmana sahiptir. 224×224 şeklinde girdi alır [121]. ResNet, eğitim aşamasında kaybı azaltmaya, bilgi kazanımını korumaya ve performansını artırmaya yardımcı olan katmanlar arasında artık bağlantıları tanıtmıştır. Bir katmandaki artık bağlantı, bir katmanın çıktısının, girdisi artı girdisinin bir evrişimi olduğu anlamına gelir. ResNet makalesi, Atlama Bağlantılarını kullanma yaklaşımını popüler hale getirmiştir [120]. Şekil 3.16’da ResNet mimarilerinin yapı taşı olan atlama bağlantıları yer almaktadır.



Şekil 3.16. Artık öğrenme: bir yapı taşı [120]

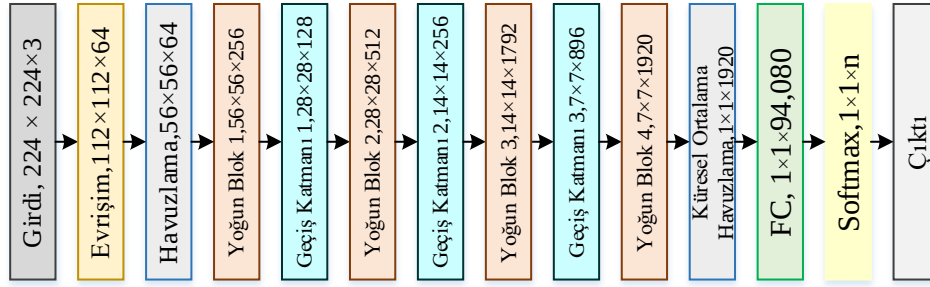
Matematiksel olarak, $y = x + F(x)$ anlamına gelir; burada y katmanın nihai çıktısıdır. Mimari açıdan ağdaki herhangi bir katman modelin performansına zarar verirse, atlama bağlantılarının varlığı nedeniyle atlanır. Bu tür bir atlama bağlantısının dahil edilmesinin faydası düzenlileştirmenin mimari performansını düşüren herhangi bir katmanı atlayacak olmasıdır. Sonuç olarak, kaybolan veya genişleyen gradyanlarla ilgili sorunlarla karşılaşmadan derin bir sinir ağını eğitmek mümkündür [121]. Şekil 3.17’de ResNet50 mimarisinin yapısı yer almaktadır.



Şekil 3.17. ResNet50 mimarisi [122]

3.5.5. DenseNet201

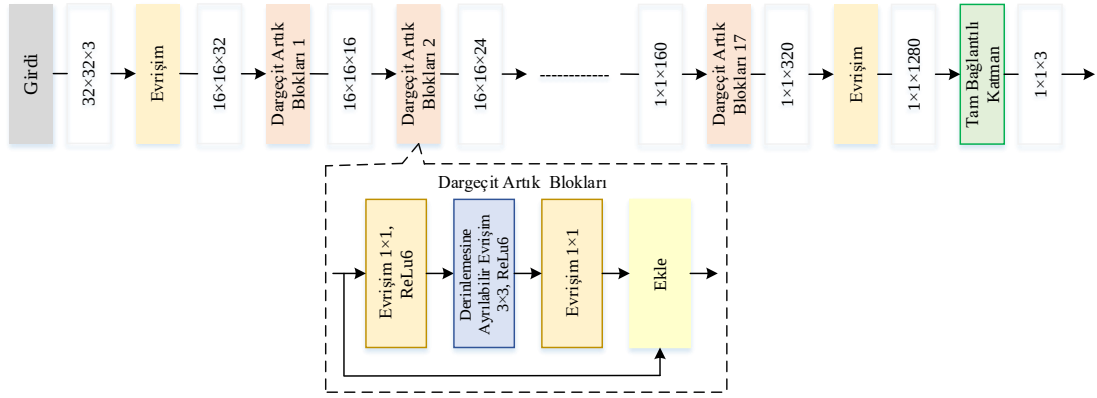
DenseNet201, 201 katman derinliğinde olan ve 224×224 görüntü giriş boyutunu kabul eden bir ESA'dır [123]. DenseNet, gradyan akışını optimize ederek ResNet mimarisi üzerinde geliştirilmiştir. DenseNet201 modeli parametre sayısını azaltmak ve katmanlar arasında daha verimli ve daha kısa bağlantılara sahip olmak üzere tasarlanmıştır. DenseNet, geleneksel ESA'lara göre daha az parametre ile ağırlı bilgi akışını ve gradyanını iyileştirebilmektedir. DenseNet yapısındaki her katman orijinal sinyal ve kayıp fonksiyonuna bağlanarak ağırlı eğitilmesini kolaylaştırır. Ayrıca, veri kümelerinin organizasyonunu etkileyen yoğun bağlantılara sahiptir. DenseNet201, katman girdilerindeki çeşitliliği artıran ve performansı iyileştiren farklı katmanlar tarafından özelliğin yeniden kullanılması olasılığı nedeniyle eğitilmesi kolay, parametrik olarak yüksek verimli modeller sağlayan yoğunlaştırılmış bir ağdan oluşur [124]. DenseNet201, ImageNet ve CIFAR-100 gibi önceden eğitilmiş çeşitli ağırlıklara sahip veri kümelerinde dikkate değer bir performans göstermiştir. DenseNet201 modelinde bağlanabilirliği artırmak için önceki tüm katmanlar sonraki tüm katmanlara ileri beslemeli bir şekilde bağlanır. DenseNet mimarilerinde önceki katmanlardan gelen özellik haritalarının mevcut katmandaki özellik haritasıyla birleştirilmesi, özellik temsilini geliştirir. Ayrıca evrişim katmanlarında daha az bağlantı yer aldığından, eğitilebilir parametrelerin sayısı azalır ve bu nedenle model hesaplama açısından verimlidir. DenseNet mimari modelleri üç geçiş katmanı ve dört yoğun blok içerir. Yoğun bloklar 1×1 ve 3×3 matris boyutlarında evrişim çekirdeklerine sahiptir. DenseNet içindeki yoğun bloklardaki evrişim çekirdekleri altı, on iki, yirmi dört ve altı kez tekrar eder. Bu mimari modelde yoğun katmanlar arasında bir geçiş katmanı bulunmaktadır. Yoğun bir bloktaki özellikleri çıkaran her bir evrişim katmanı, başka bir özellik çıkarıcı evrişim katmanına ileri beslemeli olarak bağlıdır. DenseNet mimarisindeki geçiş katmanı, 1×1 çekirdek boyutuna sahip evrişim, toplu normalizasyon ve havuzlama katmanlarından oluşur. Havuzlama katmanı 2×2 adımlıdır [125,126]. Şekil 3.18'de DenseNet201 mimarisinin basit bir gösterimi yer almaktadır.



Şekil 3.18. DenseNet201 basit gösterimi [127]

3.5.6. MobileNetV2

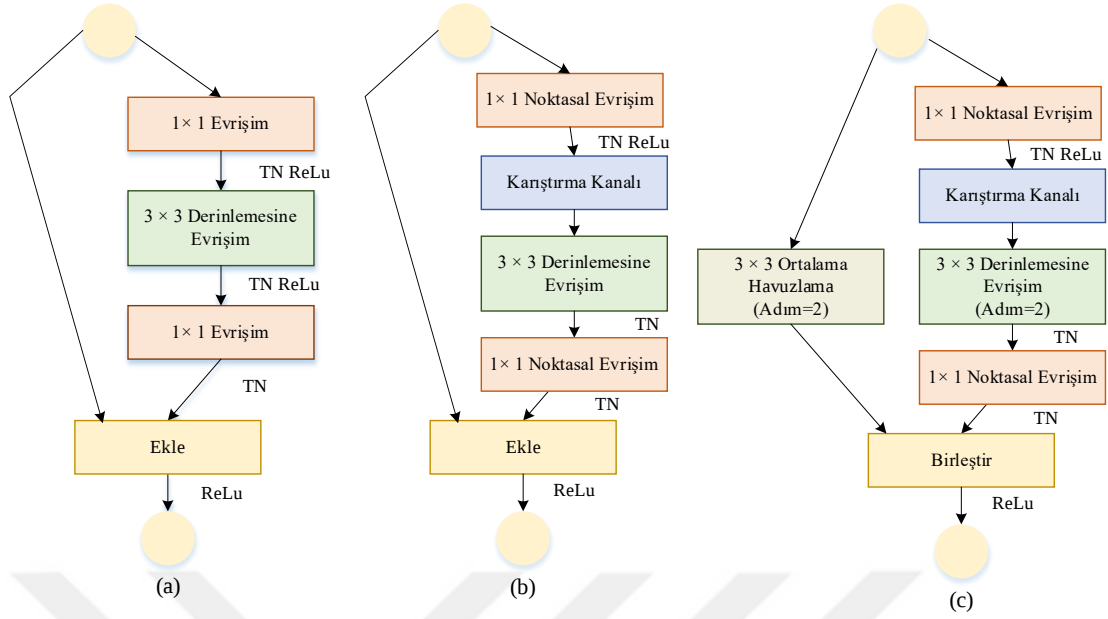
MobileNetV2, MobileNet mimarisine doğrusal dargeçit ve ters çevrilmiş artık modülleri (inverted residual with linear bottleneck modules) eklenerek mimarinin doğruluğunu önemli ölçüde artıran geliştirilmiş bir versiyonudur [128]. 53 katman derinliğindedir ve girdi katmanı 224×224 boyutunda bir görüntü ile beslenmektedir. ImageNet sınıflandırma, COCO nesne algılama ve VOC görüntü segmentasyonu kullanılarak başarıları değerlendirilen modelde daha az bellek kullanımı için bir evrişim modülü eklenmiştir. MobileNet tasarımı derinlemesine ayrılabilir evrişim katmanlarına dayanır ve iki bileşenden oluşur: (a) Derinlemesine evrişim; her giriş kanalına tek bir filtre uygulanır; (b) Noktasal evrişim; 1×1 evrişim, derinlemesine evrişimin sonuçlarını toplar [129]. Derinlemesine evrişim, giriş görüntüsünü ve filtreyi farklı kanallara ayırır ve ardından her giriş kanalını karşılık gelen filtre kanalıyla sarar. Filtrelenmiş çıkış kanalı üretildikten sonra, bu çıkış kanalları tekrar istiflenir. Ayrılabilir derinlikli evrişimde, kümelenmiş çıkış kanalları daha sonra kümelenmiş çıkış kanallarını bir kanalda birleştirmek için noktasal evrişim olarak adlandırılan 1×1 'lik bir evrişim kullanılarak filtrelenir. Derinlemesine evrişim, hesaplama süresini ve model boyutunu azaltmak için kullanılır. Derinlik ve noktasal evrişimi ayrı katmanlar olarak sayan MobileNetV1, $7 \times 7 \times 1280$ piksel boyutunda çıktı üreten 28 evrişim katmanına sahiptir. MobileNetV2, 32 filtreli ilk evrişim katmanından sonra 19 ters artık dargeçit katmanı ekler ve ardından $7 \times 7 \times 1280$ piksel boyutunda çıktı üreten noktasal bir evrişim ile sona erer. Artık blok, bilgileri ağına daha derin katmanına iletmek amacıyla bir atlamalı bağlantı ile bir evrişimli bloğun başlangıcını ve sonunu birbirine bağlar [128]. Şekil 3.19'da MobileNetV2 mimarisi yer almaktadır.



Şekil 3.19. MobileNetV2 ağının mimarisi [130]

3.5.7. ShuffleNet

ShuffleNet, 2018’de Zhang ve ark., tarafından önerilen, esas olarak kısıtlı hesaplama gücüne sahip mobil cihazlar için tasarlanmış, hesaplama açısından çok verimli bir ESA mimarisidir [131,132]. Mimari, doğruluğu korurken hesaplama maliyetini ve karmaşıklığını önemli ölçüde azaltmak için iki önemli işlem sunar. İlk işlem, 1×1 evrişimlerin hesaplama karmaşıklığını azaltabilen noktasal grup evrişimleridir. İkinci işlem, özellik kanalları arasında bilgi akışına yardımcı olan kanalların karıştırılmasından oluşur. Kanal karıştırmanın amacı, özellik ağırlıklarının kanalların birleşimi olacak şekilde kanallardan özellikleri karıştırmak ve çaprazlamaktır. Bu işlem, evrişimsel katmanlarda bulunan çoklu öznitelik ağırlıkları nedeniyle ESA mimarisinde üstün bir yapıya yol açar. Ayrıca, kanal karıştırma katmanları, daha güçlü bir ESA modeli elde etmek için grup evrişim katmanlarıyla birleştirilir [132]. Genel ShuffleNet ağı, diğer geleneksel evrişim ve havuzlama katmanlarıyla birlikte üç farklı aşamada gruplandırılmış bu birimlerin bir yığınından oluşur. Şekil 3.20’de ShuffleNet birimleri yer almaktadır.



Şekil 3.20. ShuffleNet Birimleri [132]

Şekil 3.20(a), derinlemesine evrişimli (3×3 DWConv) bir dargeçit birimidir [133,120,129]. Şekil 3.20(b), noktasal grup evrişimli (GConv) ve kanal karıştırmalı bir ShuffleNet birimidir. İkinci noktasal grup evrişiminin amacı, kanal boyutunu kestirme yolla eşleşecek şekilde geri kazanmaktır. Şekil 3.20(c), 2 adımlı bir ShuffleNet birimidir. Kanal karıştırma ile noktasal grup evrişimi sayesinde ShuffleNet birimindeki tüm bileşenler verimli bir şekilde hesaplanabilir.

3.5.8. NasNetMobile

Google Brain ekibi 2016 yılında sinir ağı yapılandırmaları alanında arama yapmak için NASNet mimarisini geliştirmiştir [134]. Nöral mimari arama (NAS), YSA alanındaki en yeni DÖ tekniğidir. Bu mimarinin ilk fikri, küçük veri kümelerinde en iyi evrişimsel mimariyi bulmak için bir arama yöntemi olarak NAS çerçevesinin kullanılmasından kaynaklanmaktadır. Daha sonra NASNet arama uzayı adı verilen yeni bir arama uzayı tasarımının katkısıyla mimari daha büyük bir veri kümesine aktarılmıştır. NASNet arama uzayında en iyi mimari bulunmuş ve bu mimariye NASNet adı verilmiştir. NAS, ESA'ları farklı boyutlar için optimize etmek için CIFAR10 ve ImageNet gibi standart veri kümeleriyle birlikte kullanılmıştır. İndirgenmiş sürümüne NasNetMobile adı verilir. NASNet mimarisinde Normal Hücreler ve Azaltma Hücreleri olarak adlandırılan iki evrişimsel hücre bulunmaktadır. Normal Hücre, tam boyutlara sahip bir özellik haritası döndürmek için işlev görürken; Azaltma Hücresi, özellik haritasının yüksekliğinin ve genişliğinin iki kat azaltıldığı bir özellik haritası üretmek için işlev görür. Mimari genel olarak ilk 3×3 evrişim katmanından ve ardından bir dizi azaltma hücresi ve dört normal hücreden oluşan üç kez tekrar eden bir diziden oluşur. Normal ve azaltma hücreleri hem önceki hücreden gelen girdiyi hem de

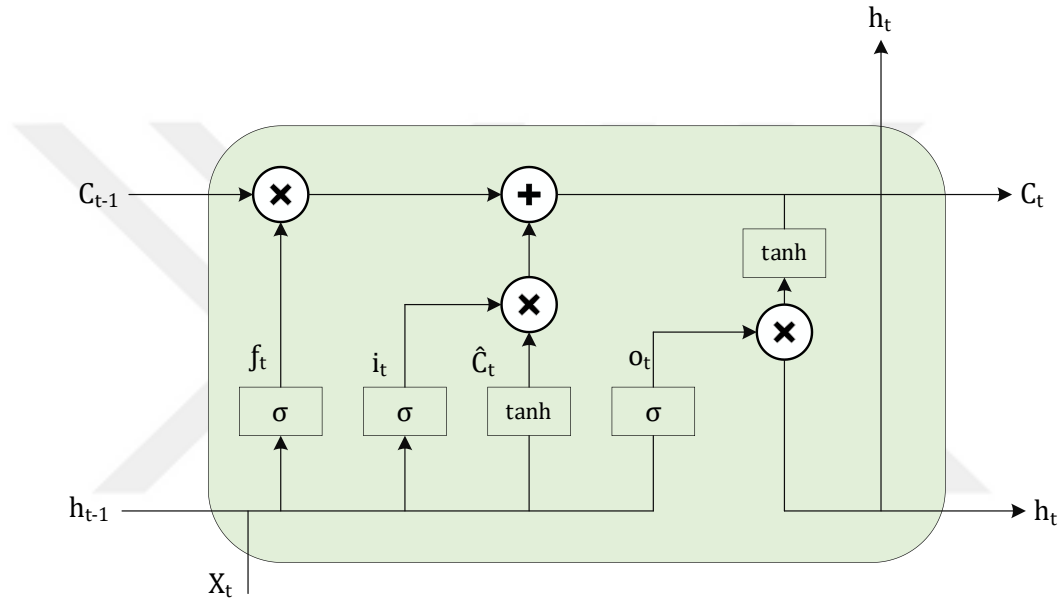
önceki hücreyi besler, artık bir ağ oluşturur. Normal hücrede bulunan tek evrişim katmanları, üç adet 3×3 ve iki adet 5×5 derinlikte ayrılabilir evrişim katmanlarıdır. Azaltma hücresi bir adet 3×3 , iki adet 5×5 ve iki adet 7×7 derinlikte ayrılabilir evrişim içerir. Katmanların geri kalanı ya ortalama ya da maksimum havuzlama katmanlarıdır [134,135].

3.5.9. Darknet53

Darknet53, eğitim görüntülerinin özelliklerini çıkarmak için kullanılan hedef tespit ağı YOLOv3'ün temel özellik çıkarma ağıdır. Darknet53 temel olarak 1×1 ve 3×3 boyutlarında, toplam 53 katmanlı (son tam bağlı katman dahil ancak artık katman hariç) bir dizi evrişim katmanından oluşur. Her bir evrişim katmanını bir BN katmanı ve LeakyReLU katmanı takip etmektedir. Darknet53'e bir dizi artık ağ modülü, yani ResNet'den türetilen artık katman eklenmiştir. Artık katmanın eklenmesinin amacı, ağıdaki gradyan kaybolması veya gradyan patlaması sorunlarını çözmektir, böylece gradyanın yayılmasını daha kolay kontrol edebilir ve ağ eğitimi gerçekleştirilebilir. Darknet53'de, çok sayıda 1×1 evrişim çekirdeği kullanılmasına ve maksimum havuzlama yerine adım boyutu 2 olan 3×3 evrişim çekirdeği kullanılmasına rağmen, parametre sayısı çok azaltılmıştır. Bu ağ için görüntünün giriş boyutu 256×256 'dır. Evrişim katmanı, BN katmanı ve LeakyRelu katmanı birlikte mimarinin en küçük bileşenini oluştururlar [136,137].

3.5.10. Uzun Kısa Süreli Bellek (UKSB)

Hochreiter ve Schmidhuber (1997) tarafından tanıtılan Uzun Kısa Süreli Bellek (UKSB), uzun süreli bağılıkları öğrenebilen ve bağımlılık sorunundan kaçınmak için tasarlanmış ağlardır. Bu ağlar özgün bir kapı yapısı kullanılarak inşa edilmiştir. Şekil 3.21’de bir UKSB’nin temel bileşenleri gösterilmektedir. Unut kapısı, Güncelleme/Giriş kapısı ve Çıkış kapısı bir UKSB biriminin işlevini yerine getirmek ve bilgi akışını kontrol etmek için birlikte çalışır. UKSB ağlarının çalışma prensibi; kapı olarak adlandırılan yapılarda hangi bilgilerin atılıp hangilerinin hafızada tutulacağına karar vermektir.



Şekil 3.21. UKSB biriminin temel yapısı [138]

UKSB’nin ilk adımında unut kapısı (f_t), önceki gizli durumdaki bilgiler (h_{t-1}) ve mevcut girişten gelen bilgilerden (x_t) hangilerinin atılıp atılmayacağı, hangilerinin hafızada tutulacağı kararını sigmoid aktivasyon fonksiyonu $\sigma(\cdot)$ kullanarak verir. Denklem 3.2’de bu sürecin matematiksel ifadesi yer almaktadır. Burada W_f ve b_f eğitimden sonra belirlenecek parametrelerdir. h_{t-1} , $t - 1$ zaman tekrarındaki gizli durumdur. x_t , t . zaman tekrarındaki girdi vektörüdür. f_t sigmoid fonksiyonunun çıktısıdır [138]. Fonksiyonun girişleri bir önceki UKSB biriminden gelen gizli durum ve giriş vektörüdür. Fonksiyonların çıkışları, bir önceki UKSB biriminden hücre durumundaki her bir sayıya karşılık gelen 0 ile 1 arasında değişen değerlerdir. Değerler, bir önceki hücre durumu c_{t-1} ’deki her bir sayının unutma derecesini temsil etmektedir. 1, “bunu geçir” i, 0 ise “geçmesine izin verme” i temsil eder.

$$f_t = \sigma(W_{f,x} * X_t + W_{f,h} * h_{t-1} + b_f) \quad (3.2)$$

UKSB'nin bir sonraki adımı, yeni bilginin hangi kısmının mevcut hücre durumunda saklanacağına karar vermektir. Bu işlem Güncelleme/Giriş (i_t) kapılarında gerçekleştirilir ve iki aşamadan oluşmaktadır. İlk olarak, $t-1$ anındaki gizli durum ve t anındaki girdi, bir sigmoid fonksiyonuna girişler olarak verilir (çıkış 1'e ne kadar yakınsa, bilgi o kadar önemlidir). Ayrıca, ağıın ayarını iyileştirmek için -1 ile 1 arasındaki değerleri sıkıştırmak için gizli durumu ve girdiyi bir $\tanh$ işlevinden geçirir. Sigmoid çıktısı, $\tanh$ çıktısından hangi bilgilerin korunmasının önemli olduğuna karar verir. Daha sonra, bir $\tanh$ katmanı, duruma eklenebilecek yeni aday değerlerin $\hat{C}_t$ vektörünü oluşturur. Denklem 3.3 ve 3.4'de güncelleme/girdi kapısı (i_t) ve yeni aday değerlerin matematiksel ifadesi yer almaktadır. Denklem 3.3, Denklem 3.2'e benzer bir sigmoid fonksiyondur. Bu fonksiyon, giriş vektöründeki her bir sayının güncelleme derecesine karar vermek için kullanılır. Denklem 3.4 yeni hücre durumu $\hat{C}_t$ çıktısı veren bir $\tanh$ katmanıdır.

$$i_t = \sigma(W_{i,x} * X_t + W_{i,h} * h_{t-1} + b_i) \quad (3.3)$$

$$\hat{C}_t = \tanh(W_{c,x} * X_t + W_{c,h} * h_{t-1} + b_c) \quad (3.4)$$

Giriş kapısının etkinleştirilmesinden sonra sıradaki adım, eski hücre durumu C_{t-1} 'i yeni hücre durumu C_t 'e güncellemektir. Unut kapısının çıkışı olan f_t önceki zaman adımının hücre durumu C_{t-1} ile çarpılır. Sonra bu çarpıma Girdi kapısının çıkışları olan i_t ve $\hat{C}_t$ çarpımı eklenir. Çıktı, her bir durum değerini ne kadar güncellemeye karar verdiğimizize göre ölçeklenen yeni aday değerlerdir. Denklem 3.5'de yeni hücre durumu C_t 'nin matematiksel ifadesi yer almaktadır.

$$C_t = C_{t-1} * f_t + i_t * \hat{C}_t \quad (3.5)$$

Son adım, önceki girişler hakkında bilgi içeren ve bir sonraki gizli durumun değerine karar veren çıkış kapısından (o_t) hücre çıktısına (h_t) karar vermektir. İlk olarak, önceki gizli durum ve girdi toplanır ve bir sigmoid fonksiyonuna aktarılır. Daha sonra yeni hücre durumu $\tanh$ fonksiyonuna geçer. Sonunda, gizli durumun hangi bilgileri içermesi gerektiğine karar vermek için sigmoid çıktısıyla $\tanh$ çıktısı çarpılır. Çıktı, yeni gizli durumdur. Yeni hücre durumu ve yeni gizli durum daha sonra bir sonraki zaman adımına taşınır. Bu işlemin matematiksel ifadesi Denklem 3.6 ve Denklem 3.7'da yer almaktadır. Bu denklemlerdeki W ağırlık vektör matrisini ve b sapma vektörünü ifade etmektedir [139-142].

$$o_t = \sigma(W_{o,x} * X_t + W_{o,h} * h_{t-1} + b_o) \quad (3.6)$$

$$h = o_t * \tanh(C_t) \quad (3.7)$$

3.5.11. Otomatik Kodlayıcı

Otomatik kodlayıcılar, veri sıkıştırma, görüntü renklendirme ve görüntüden gürültüyü arındırma problemlerinde yaygın olarak kullanılmaktadır [143]. Veri sıkıştırmanın amacı, boyutsallığı azaltarak giriş verisinin daha küçük bir gösterimini elde etmektir. Otomatik kodlayıcı modelinin mimarisi, boyut azaltma tekniklerini kullanarak giriş verilerine benzer verilerin öğrenilmesine olanak sağlamaktadır.

Bir otomatik kodlayıcı oluştururken 3 bileşene ihtiyaç vardır. Bunlar; kodlayıcı, kod çözücü ve girdinin sıkıştırılmış düşük boyutlu gösterimini gerçekleştirecek bir koddur. Otomatik kodlayıcı mimarisinin en temel halinde 3 katman bulunmaktadır. Bu 3 katman sinir ağı temelindeki giriş, gizli ve çıkış katmanlarıdır. Giriş ve çıkış katmanlarındaki nöron sayısı eşittir, sebebi ise otomatik kodlayıcıdaki temel çalışma prensibinin veriyi yeniden inşa etmesidir. Giriş ve gizli katman arasında kodlayıcı varken gizli katman ile çıkış katmanı arasında kod çözücü vardır. Kodlayıcı yapısı ile otomatik kodlayıcı ağı, gelen çok boyutlu verileri sıkıştırarak daha az boyutlu verilere düşürmeyi amaçlar. Kod çözücü ise kodlayıcının aksine gelen sıkıştırılmış veriyi yeniden inşa etmektedir. Kodlama aşamasında, kodlayıcı girişi ($x = [x_{(1)}, x_{(2)}, \dots, x_{(m)}]^T \in R^m$), f fonksiyonu aracılığıyla gizli katmana ($h = [h_{(1)}, h_{(2)}, \dots, h_{(d_h)}]^T \in R^{d_h}$) eşler. Bu işlemin matematiksel ifadesi Denklem 3.8'de yer almaktadır.

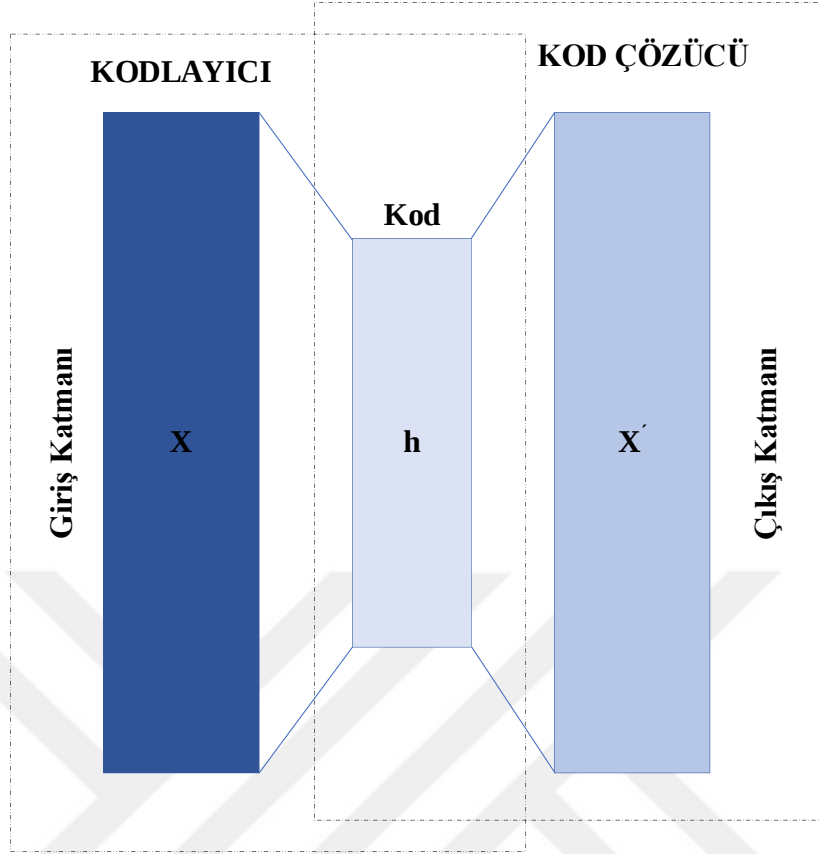
$$h = f(x) = \sigma_f(W_x + b) \quad (3.8)$$

Burada W , giriş katmanı ile gizli katmanı bağlamak için kullanılan bir $d_h \times m$ ağırlık matrisidir; $b \in R^{d_h}$ sapma vektörüdür; σ_f , sigmoid, tanh, relu, vb. olabilen aktivasyon fonksiyonudur. Kod çözücü aşamasında otomatik kodlayıcı, gizli katman özelliklerini (h) $\tilde{f}$ fonksiyonuyla çıktı katmanına ($\tilde{x} \in R^m$) eşler. Bu işlemin matematiksel ifadesi Denklem 3.9'de yer almaktadır.

$$\tilde{x} = \tilde{f}(h) = \tilde{\sigma}_f(W_h + \tilde{b}) \quad (3.9)$$

Burada $\tilde{W}$, bir $d_h \times m$ ağırlık matrisidir; $\tilde{b} \in R^m$ çıktı katmanındaki sapma vektörüdür; $\tilde{\sigma}_f$, çıktı katmanındaki aktivasyon fonksiyonudur. Otomatik kodlayıcının amacı, yeniden yapılandırılmış verileri ($\tilde{x} \in R^m$) öğrenerek girdi verilerine mümkün olduğunca benzer hale getirmektir. Bu nedenle, otomatik kodlayıcının kayıp fonksiyonu, eğitim verilerinin ortalama karesel hatası (MSE) olarak ifade edilebilir. N , eğitim örneklerinin sayısıdır; x_i , i 'inci eğitim örneğidir. Otomatik kodlayıcının kayıp fonksiyonu Denklem 3.10'da gösterildiği gibi hesaplanır [143,144]. Şekil 3.22'de otomatik kodlayıcıya ait şematik bir diyagram gösterilmektedir [144].

$$K(W, b, \tilde{W}, \tilde{b}) = \frac{1}{2N} \sum_{i=1}^N \|x_i - \tilde{x}_i\|^2 \quad (3.10)$$



Şekil 3.22. Otomatik kodlayıcının şematik diyagramı [144]

Otomatik kodlayıcı, yüksek boyutlu bir girdiyi gizli bir düşük boyutlu koda (kodlayıcı) dönüştüren ve ardından bu gizli kodla (kod çözücü) girdinin yeniden yapılandırılmasını gerçekleştiren bir darboğaz mimarisidir [144].

ESA, ağdaki düşük evrişim katmanlarından derin evrişim katmanlarına kadar etkili bir şekilde öznetelikler çıkardığı için görüntü sınıflandırma ve bilgisayarla görme problemlerinde en gelişmiş mimari olarak kabul edilir [145]. Ancak daha sonraki çalışmalarda, araştırmacılar, bilgisayarla görme uygulamaları için derin öğrenme modellerinde evrişim fonksiyonları olmadan başarılı tanıma ve performans iyileştirmesi aradılar [146]. Dikkat mekanizması tabanlı mimariler olan ve doğal dil işlemede (NLP) yaygın olarak kullanılan dönüştürücü modelleri bu sorunu çözmek için uyarlanmış ve dikkat çekici sonuçlar elde edilmiştir [147,32]. Dönüştürücü mimarisinin görüntü modeli olan Görüntü Dönüştürücü (ViT), bir girdi görüntüsünü sabit boyutlu bir görüntü yamaları dizisi olarak temsil eder ve NLP uygulamalarında kullanılan kelime gömme dizisine benzer şekilde görüntü sınıfı etiketlerini tahmin eder [33].

Dönüştürücüler, Vaswani ve arkadaşları tarafından "Attention Is All You Need" adlı makalede tanıtılan bir tür derin öğrenme modelidir. Dönüştürücülerdeki temel yenilik, modelin girdi verilerinin farklı kısımlarına farklı şekilde odaklanmasını sağlayan dikkat mekanizmasıdır [32]. Dönüştürücü ağları, girdi olarak bir kelime dizisini alır ve daha sonra sınıflandırma, çeviri

veya diğer NLP görevleri için kullanır. Kodlayıcı-kod çözücü mimarisinden oluşan dönüştürücü ağları, tekrarlamaya dayanmadan giriş ve çıkış dizileri arasındaki bağımlılıkları öğrenen, kelime konumlarını ve aynı cümle içindeki kelimeler arasındaki ilişkiyi dikkate alan bir mekanizma üzerine kuruludur. Bu, dönüştürücü uygulamalarının kolayca paralelleştirilmesine ve hesaplama açısından verimli olmasına olanak tanır. Dönüştürücülerin kullandığı matematiksel hesaplamalar paralel işlemeye elverişlidir ve böylece bu modeller hızlı çalışabilirler. Bu modeller, ağa giren ve çıkan veri öğelerini etiketlemek için konumsal kodlayıcıları kullanır. Dikkat birimleri bu etiketleri takip ederek her bir ögenin diğerleriyle nasıl ilişkili olduğuna dair bir tür cebirsel harita hesaplar [32]. Bir dönüştürücü modeli, bu cümledeki kelimeler gibi sıralı verilerdeki ilişkileri izleyerek bağlamı ve dolayısıyla anlamı öğrenen ileri beslemeli bir sinir ağıdır. Bir dizideki birbirini etkileyen ve birbirinden uzak veri öğelerinin bile ince bağlantılarını tespit etmek için, dikkat ve öz dikkat adı verilen ve hala gelişmekte olan bir dizi matematiksel teknik uygular.

3.5.12. Görüntü Dönüştürücü (Vision Transformer / ViT)

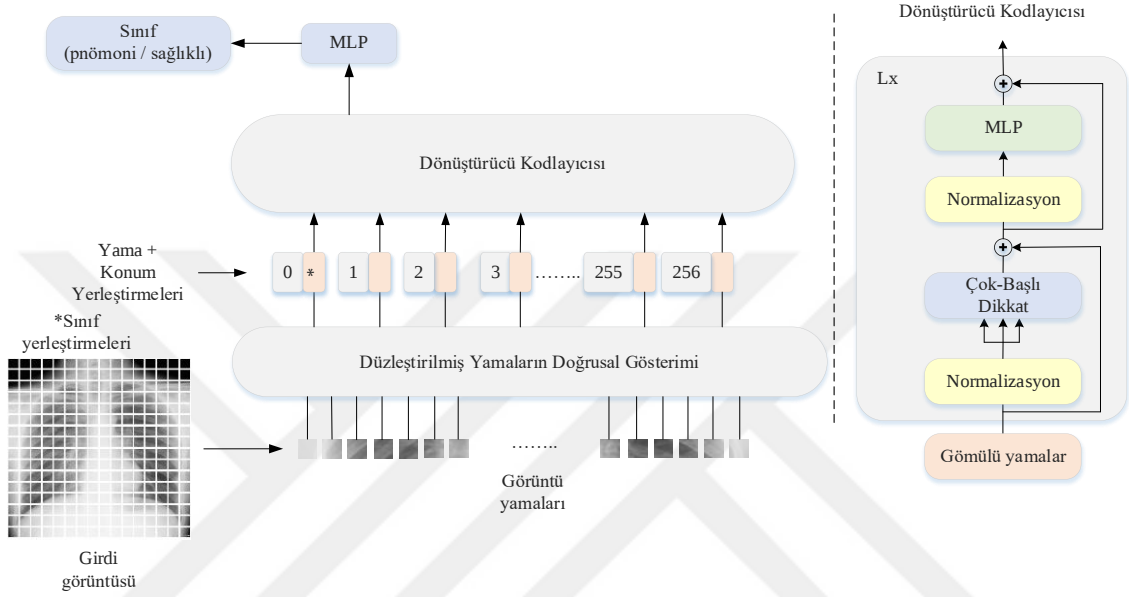
Görüntü Dönüştürücü (ViT) [33], NLP uygulamalarında başarılı bir şekilde gerçekleştirilmiş olan dikkat mekanizması tabanlı dönüştürücü modellerin bilgisayar görüşüne uyarlanmış halidir. Bir görüntüyü sınıflandırırken ViT onu, NLP dönüştürücüsü tarafından oluşturulan bir kelime gömme dizisine benzer şekilde bir yama dizisi olarak ele alır. ViT modelinin çalışma prensibi aşağıdaki adımlardan oluşmaktadır. ViT, bir giriş görüntüsünü yama dizilerine veya görsel belirteçlere böler. Her 2D görüntü yaması düzleştirilir. Ardından, düzleştirilmiş yamalar, yama gömme olarak adlandırılan bir işlemle doğrusal olarak gömülür. Her görüntü yamasına öğrenilebilir konum yerleştirmeleri eklenir. Gömülü bir görüntü yaması, yeni görüntünün sınıfını tahmin etmek için ekstra öğrenilebilir sınıf belirteçleriyle birleştirilir. Son olarak, elde edilen dizi dönüştürücü kodlayıcı bloğuna verilir. Standart dönüştürücü girdi olarak 1D belirteç gömme dizilerini alırken, ViT’de 2D görüntüleri işlemek için $x \in \mathbb{R}^{H \times W \times C}$ girdi görüntüsü; $x_p \in \mathbb{R}^{N \times (P^2 \cdot C)}$ düzleştirilmiş 2D yamalar dizisi olarak yeniden şekillendirilmiştir. Burada (H, W) orijinal görüntünün çözünürlüğüdür, yani yükseklik ile genişlik değerleridir, C kanal sayısıdır, (P, P) her görüntü yamasının çözünürlüğüdür, $N = HW/P^2$ yama sayısıdır. Aşağıdaki denklemlerde görüntü dönüştürücüsünün görüntüleri işleyişi matematiksel olarak açıklanmıştır. Denklem 3.11’de yamaların düzleştirilip, eğitilebilir doğrusal bir izdüşümle D boyutuna eşlenmesi yer almaktadır. Düzleştirilmiş yamaların (x_p) toplam boyutu $N \times (P^2 \cdot C)$ ‘dir. Denklem 3.12 ve 3.13’de dönüştürücü kodlayıcı katmanının (L) MLP (çok katmanlı algılayıcı) ve MSA (çok başlı öz-dikkat) blokları yer almaktadır. ℓ ’inci katmanın çıktıları denklemlerdeki gibi hesaplanır. Denklem 3.14’de kodlayıcının çıkışının (z_L^0) görüntü gösterimi y ile ifade edilmektedir [33]. Şekil 3.23’de pnömoni tespiti için ViT modelinin gösterimi yer almaktadır.

$$z_0 = [x_{class}; x_p^1 E; x_p^2 E; \dots; x_p^N E] + E_{pos}, \quad E \in \mathbb{R}^{(P^2 \cdot C) \times D}, E_{pos} \in \mathbb{R}^{(N+1) \times D} \quad (3.11)$$

$$z'_\ell = MSA(LN(z_{\ell-1})) + z_{\ell-1}, \quad \ell = 1 \dots L \quad (3.12)$$

$$z_\ell = MLP(LN(z'_\ell)) + z'_\ell, \quad \ell = 1 \dots L \quad (3.13)$$

$$y = LN(z_L^0) \quad (3.14)$$



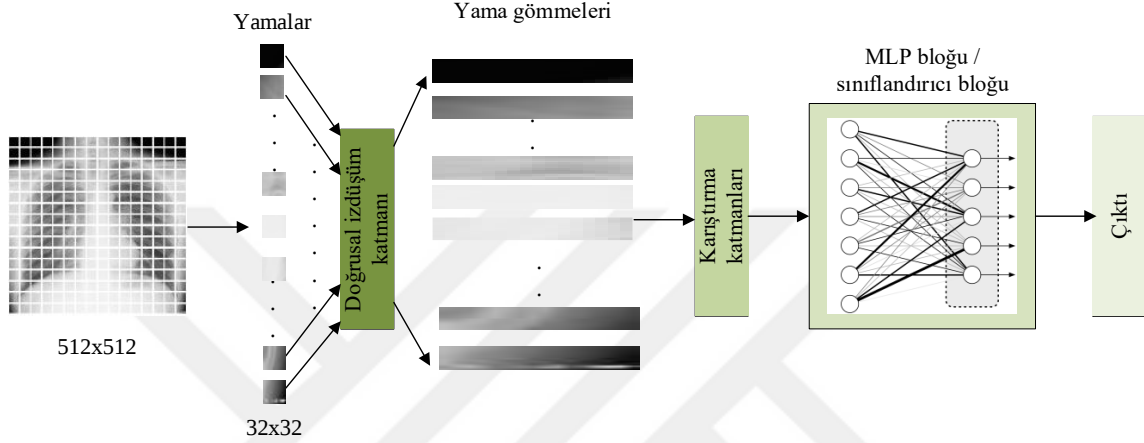
Şekil 3.23. Pnömoni tespiti için ViT modelinin gösterimi

3.5.13. MLP-mixer

MLP-mixer, yalnızca MLP'e dayanan ve iki tür MLP katmanı içeren bir mimaridir. Bunlardan biri, konum başına özellikleri karıştıran görüntü yamalarına bağımsız olarak uygulanır. Diğeri, uzamsal bilgileri karıştıran yamalar boyunca (kanallar boyunca) uygulanır. Mixer, "yamalar×kanallar" tablosu olarak şekillendirilmiş doğrusal olarak yansıtılmış görüntü yamaları (belirteçler) dizisini girdi olarak kabul eder ve bu boyutluluğu korur. Geliştirilen MLP-mixer'in temel mimarisi Şekil 3.24'de verilmiştir. Görüldüğü gibi, giriş renkli görüntüsünün boyutu 512×512'dir ve yama çıkartmak için 32×32 boyutunda örtüşmeyen bir pencere kullanılmaktadır. Sonuçta yamalar çıkarılır, her yama başlangıçta düzeltilir ve her düzeltilmiş yamadaki örnek sayısını azaltmak için doğrusal bir izdüşüm katmanı kullanılır. Doğrusal izdüşüm katmanı Denklem 3.15'de formüle edilmiştir.

$$y = x\omega^T + b \quad (3.15)$$

Burada x giriş vektörüdür ω ve b sırasıyla ağırlık vektörü ve sapma değeridir. Düzleştirilmiş bir yamanın uzunluğu 3072'dir; doğrusal izdüşümden sonra, her bir yama vektörünün uzunluğu 2048 olur. Benzer şekilde, doğrusal izdüşüm katmanından sonra, tüm yamalar için 256×1024 boyutunda bir yama gömme matrisi elde edilir. Daha sonra, bir giriş görüntüsünün 256×1024 şeklindeki bu yama yerleştirmesi, sınıflandırma için MLP bloğuna beslenmeden önce mixer katmanlarından geçer [148].

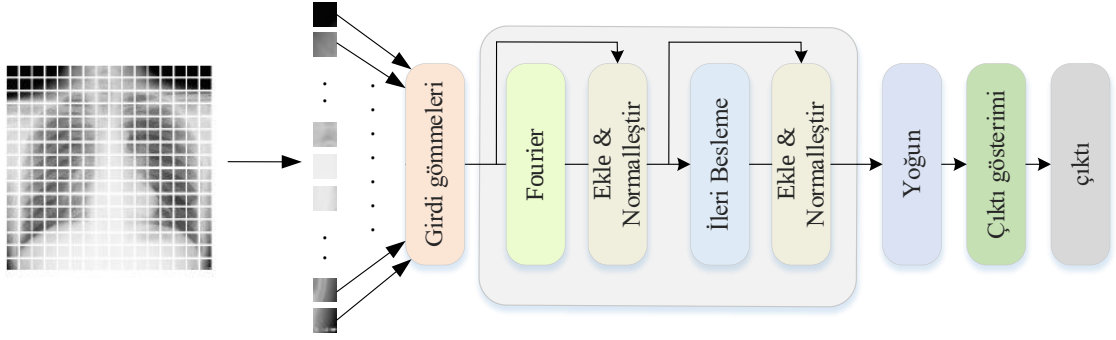


Şekil 3.24. MLP-mixer mimarisinin gösterimi

3.5.14. FNet

Lee-Thorp ve arkadaşları tarafından tanıtılan modelde, belirteç karıştırma mekanizması olarak Fourier dönüşümü kullanılmıştır [149]. Bu model ile Fourier dönüşümleri dahil olmak üzere basit doğrusal dönüşümlerin metin verilerindeki çeşitli ilişkileri modellemede yetkin olduğu gösterilmiştir. Yalnızca iki öz-dikkat alt katmanı içeren FNet hibrit modellerinin diğer dönüştürücü modellerine göre avantajı, daha uzun giriş uzunluklarında daha hızlı çalışması ve daha iyi performans göstermesidir. Ayrıca bu model, dikkat mekanizmasının artan doğruluğu etkilediğini, ancak her katmanda kullanılmasının gerekli olmadığını göstermiştir. FNet dikkat gerektirmeyen bir dönüştürücü mimarisidir (dönüştürücü benzeri mimari), her katman bir Fourier karıştırma alt katmanından ve ardından bir ileri besleme alt katmanından oluşur. Denklem 3.16'da gösterildiği gibi, FNet'in temelinde dikkat alt katmanı 2D DFT uygulayan Fourier alt katmanı ile değiştirilmiştir. Burada $\mathcal{F}_{seq}$ dizi uzunluğu, $\mathcal{F}_h$ gizli boyuttur. Bu çalışma için geliştirilen FNet mimarisi Şekil 3.25'de verilmiştir.

$$y = \Re(\mathcal{F}_{seq}(\mathcal{F}_h(x))) \quad (3.16)$$



Şekil 3.25. Pnömoni tespiti için önerilen FNet mimarisinin gösterimi

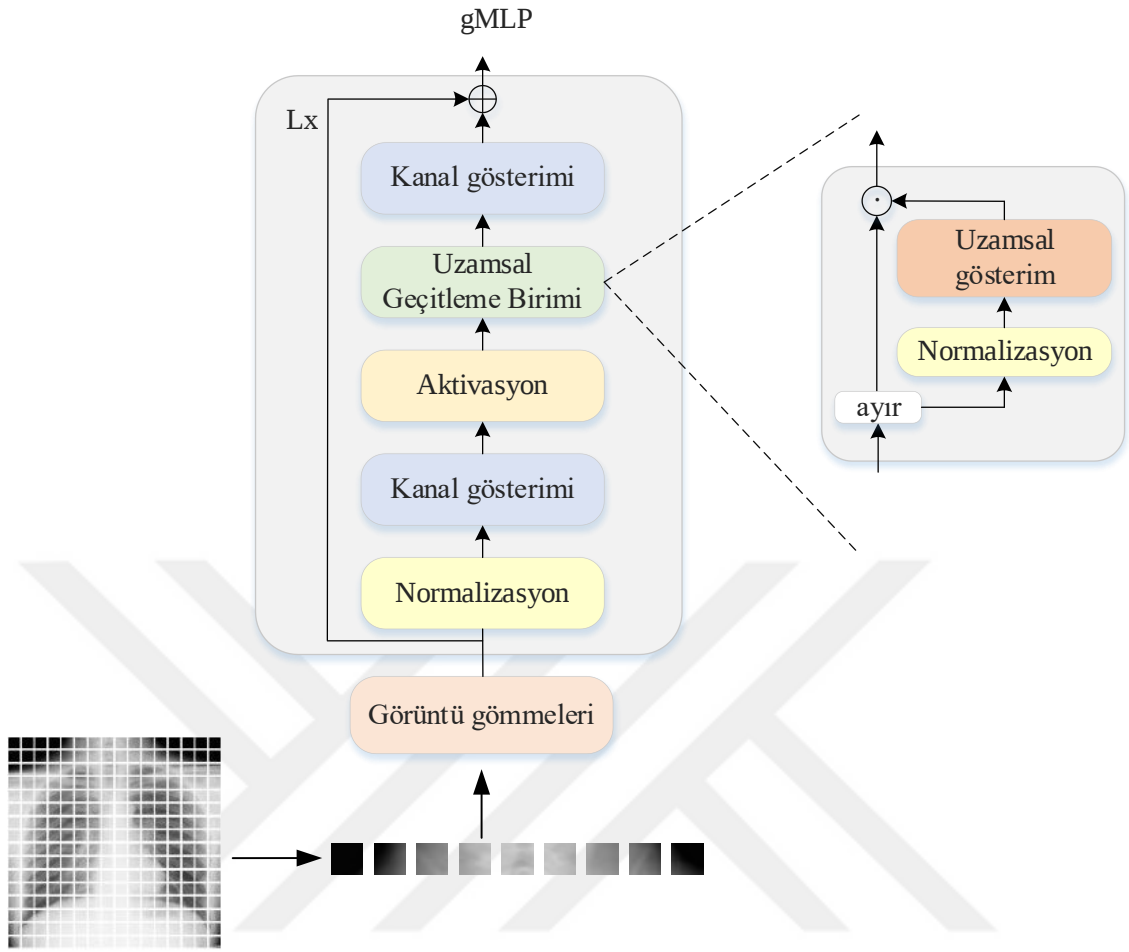
3.5.15. gMLP

Liu ve arkadaşları tarafından önerilen gMLP, dikkat mekanizmasız MLP tabanlı dönüştürücü modelidir. Mimari, kapı (geçit) eklenmiş bir MLP olduğu için bu şekilde adlandırılmıştır. Dönüştürücü mimarilerindeki dikkat mekanizmasının gerekliliğini araştırmak için temel MLP katmanları geçitleme ile birleştirilerek tasarlanmıştır. GMLP'deki yenilik, dizi öğeleri arasındaki uzamsal etkileşimleri öğrenen dikkat mekanizmalarına alternatif olarak Uzamsal Geçitleme Birimi (Spatial Gating Unit / (SGU)) kullanılmasıdır. SGU öğe konumları için kodlama gerektirmez çünkü bu tür bilgiler uzamsal etkileşimleri yakalayan katman $s(\cdot)$ 'de tutulur. gMLP, diğer dönüştürücü modellerine göre daha az eğitilebilir parametreye sahiptir. gMLP ağı, L adet aynı boyutta ve yapıda blok yığından oluşur. Belirteç dizisinin uzunluğu n ve belirtecin boyutu d olan $X \in \mathbb{R}^{n \times d}$ başlangıç matrisi girdisi aşağıdaki formüllerle temsil edilir. Her vektör (gömm) belirli bir kelimenin temsilidir. Burada σ , GeLU gibi bir aktivasyon fonksiyonudur, U ve V kanal boyutu boyunca doğrusal projeksiyonları tanımlayan matrisler, $s(\cdot)$ uzamsal etkileşimleri yakalayan katmandır. Y bloğun çıktısıdır [150]. Önerilen gMLP mimarisinin gösterimi Şekil 3.26'de verilmiştir. Denklem 3.17, Denklem 3.18 ve Denklem 3.19'da ağda gerçekleşen süreçlerin matematiksel ifadeleri yer almaktadır.

$$Z = \sigma(XU) \quad (3.17)$$

$$\check{Z} = s(Z) \quad (3.18)$$

$$Y = \check{Z}V \quad (3.19)$$



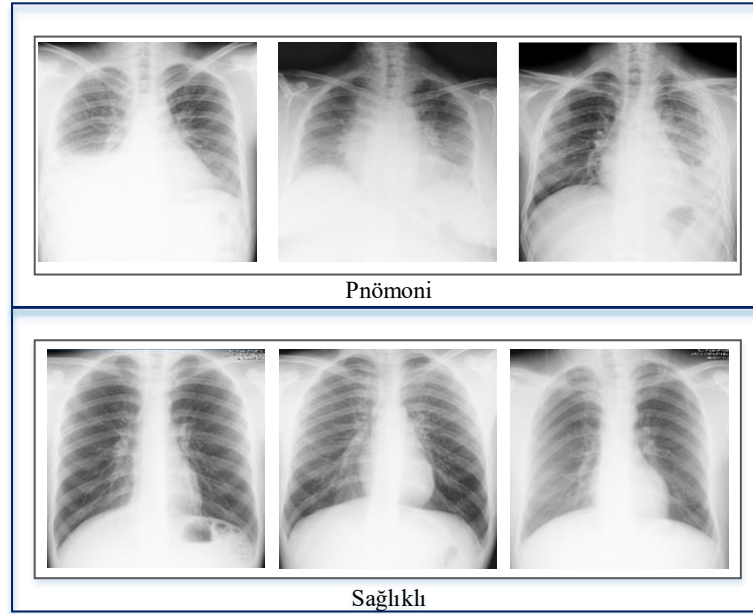
Şekil 3.26. gMLP mimarisinin gösterimi

4. MATERYAL VE METOT

Bu bölümde tez çalışması kapsamında oluşturulan CXR ve sayısal tıbbi parametre veri setleri, literatürde CXR görüntülerinden pnömoni tespiti için en yaygın kullanılan CXR veri seti [22], pnömoni tespiti için geliştirilen 21 katmanlı yeni ESA modeli ve hiper parametreleri, önerilen modelleri değerlendirmek için kullanılan çapraz doğrulama yöntemleri, performans metrikleri ve sayısal parametrelerin istatistiksel analizinde kullanılan istatistik testleri hakkında detaylı bilgi verilmiştir.

4.1. Özel Veri Tabanı

Bu çalışma Dicle Üniversitesi Tıp Fakültesi Göğüs Hastalıkları ve Tüberküloz kliniği ile yoğun bakım ünitesine 2001 Ocak- 2021 Ocak ayları arasında pnömoni tanısıyla yatırılan hastalar ve herhangi bir solunum şikâyetiyle göğüs hastalıkları polikliniğine başvurup CXR çekilen ve normal olan hastaların dosyalarının retrospektif olarak taranmasıyla gerçekleştirilmiştir. Çalışmaya CXR’de pnömoni tanısı alan 1000 pnömoni hastası ve CXR’si normal olan 1000 sağlıklı birey olmak üzere toplam 2000 kişi dâhil edilmiştir. Önerilen yöntemin etkinliğini ve performansını karşılaştırmak için aynı kişilere ait hem CXR görüntüleri hem de sayısal tıbbi verileri kullanılmıştır. Tüm görüntüler farklı boyutlarda ve RGB’dir. Şekil 4.1, özel veri tabanından elde edilen pnömoni ve sağlıklı CXR görüntülerini göstermektedir.



Şekil 4.1. Oluşturulan veri setine ait pnömoni ve sağlıklı CXR görüntüleri

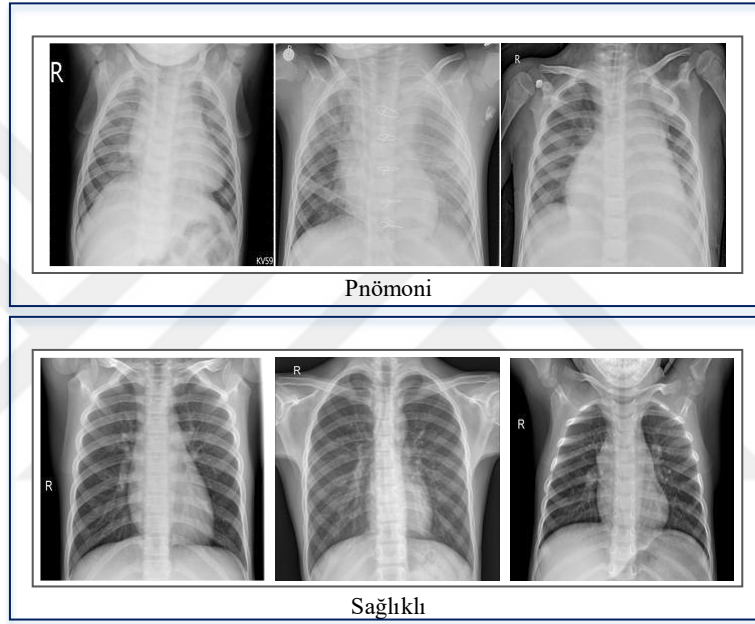
Çalışmaya CXR'si çekilen 1000 pnömoni hastası ve 1000 sağlıklı bireyin hastalığın ayırıcı tanısında etkisi olan sayısal tıbbi verileri de dahil edilmiştir. Sayısal tıbbi veriler hasta dosyasından elde edilmiştir ve sayısal verilere dayalı bir veri seti oluşturulmuştur. Demografik veriler olarak, yaş ve cinsiyet seçilmiştir. Hastalık şikâyeti olarak, yan ağrısı, nefes darlığı ile öksürük verileri seçilmiştir. Laboratuvar parametreleri olarak; lökosit sayısı (WBC, $10e^3\mu\text{L}$), nötrofil sayısı (NEU, $10e^3\mu\text{L}$), lenfosit sayısı (LYM, $10e^3\mu\text{L}$), trombosit sayısı (PLT, $10e^3\mu\text{L}$), C-reaktif protein (CRP (mg / dL)) ve Albümin (mg / dL) çalışılmıştır. Bu laboratuvar verilerinden hareketle pnömoni ve sağlıklı hastaların ayırıcı tanısında NLO, TLO ve CAO verileri kullanılmıştır. Tablo 4.1'de oluşturulan sayısal parametre veri seti öznitelikleri yer almaktadır. Pnömoni hastaları ve sağlıklı bireylere ait CXR görüntüleri ve aynı kişilere ait sayısal tıbbi veriler PACS (Görüntü Saklama ve İletişim Sistemleri) ve HBYS (Hastane Bilgi Yönetim Sistemi) veri tabanlarından sağlanmıştır. Bu çalışma, Dicle Üniversitesi Tıp Fakültesi Girişimsel Olmayan Klinik Araştırmalar Etik Kurulu onayı ile yapılmıştır. (Tarih: 13.10.2021 ve Karar No. 421).

Tablo 4.1. Oluşturulan sayısal parametre veri seti öznitelikleri

Öznitelik Sayısı	Öznitelik
1	Hasta/Sağlıklı Yaş
2	Cinsiyet
3	Öksürük
4	Dispne (nefes darlığı)
5	Yan ağrısı
6	CRP (mg / dL)
7	Albümin (g / dL)
8	WBC $10e^3\mu\text{L}$
9	LYM $10e^3\mu\text{L}$
10	NEU $10e^3\mu\text{L}$
11	PLT $10e^3\mu\text{L}$
12	CAO
13	NLO
14	TLO
15	Sınıf (0/1)

4.2. Genel Veri Tabanı

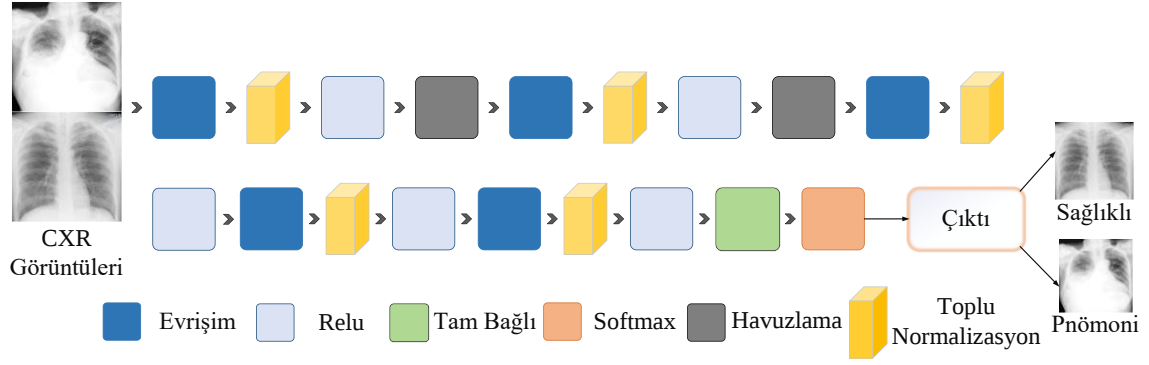
Tez çalışması kapsamında geliştirilen pnömoni tespit ağının etkinliğini test etmek amacıyla pnömoni tespitinde literatürde en çok kullanılan CXR görüntü veri setleri ile deneysel çalışmalar gerçekleştirilmiştir [22,35]. Sınıf dengesizliği problemlerini önlemek için veri setlerindeki her iki sınıftan eşit sayıda CXR görüntüsü rastgele seçilmiş ve kullanılmıştır. Bu çalışmada 1750 adedi sağlıklı sınıftan ve diğer 1750 örneği pnömoni sınıfından olmak üzere 3500 görüntü kullanılmıştır. Şekil 4.2’de ikili sınıflandırma görevlerinde literatürde en yaygın kullanılan Kaggle CXR veri setine ait sağlıklı ve pnömoni CXR görüntü örnekleri yer almaktadır.



Şekil 4.2. Halka açık CXR veri setine ait sağlıklı ve pnömoni CXR görüntüleri [22]

4.3. Geliştirilen Pnömoni Tespit Modeli

Geliştirilen ESA modeli Şekil 4.3’de gösterildiği gibi 21 katmandan oluşmaktadır. Yeni ESA mimarisi, bir giriş katmanı ile başlayıp, toplu normalleştirme ve daha sonra her bir evrişim katmanını takip eden ReLU katmanı ile devam etmektedir. Havuzlama işlemlerinden havuzlama katmanlarında ‘max’ operatör fonksiyonu kullanılmıştır. 21 katmanlı ESA mimarisinin evrişim1, evrişim2, evrişim3, evrişim4 ve evrişim5 olmak üzere beş evrişim katmanı bulunmaktadır. Ayrıca sırasıyla ReLu1 ve ReLu2 katmanlarından sonra gelen havuzlama1 ve havuzlama 2 olmak üzere iki havuz katmanı vardır. Sınıflandırma amacıyla tam bağlantılı katman, softmax katmanı ve sınıflandırma katmanı da kullanılmıştır. Evrişim1, evrişim2 ve evrişim3 katmanları, 3×3 piksel boyutunda 64, 32, 16, 8 ve 4 filtre içermektedir. Katmanların, aktivasyonların ve öğrenilebilir ağırlıkların açıklamalarını kapsayan yeni ESA mimarisinin detayları Şekil 4.3 ve Tablo 4.2’de sunulmuştur.



Şekil 4.3. Pnömoni tespitinde önerilen yeni 21 katmanlı ESA modeli

Tablo 4.2. Geliştirilen model için kullanılan hiper parametreler

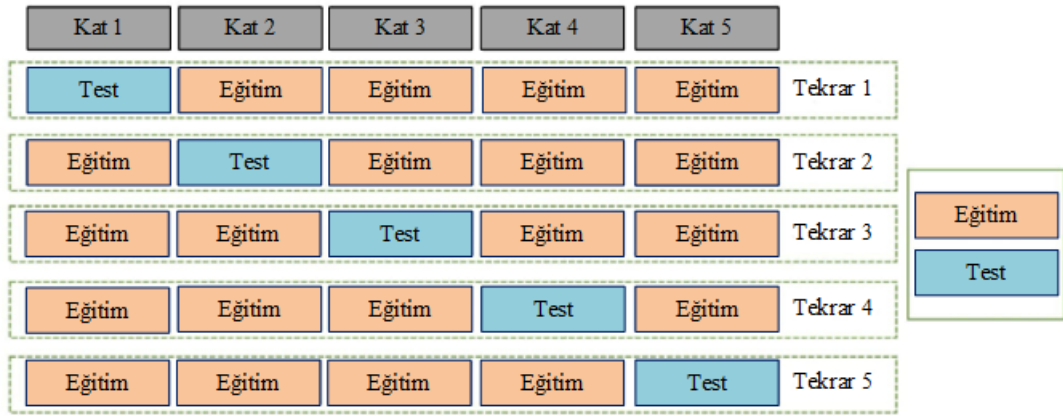
Katmanlar	Aktivasyonlar	Öğrenilebilir Ağırlıklar
Görüntü girişi	224×224×3	-
Evrişim Katmanı	224×224×64	Ağırlıklar 3×3×3×64 Sapma 1×1×64
Toplu Normalizasyon	224×224×64	Offset 1×1×64 Scale 1×1×64
ReLu	224×224×64	-
Maksimum Havuzlama	112×112×64	-
Evrişim Katmanı	112×112×32	Ağırlıklar 3×3×3×64 Sapma 1×1×64
Toplu Normalizasyon	112×112×32	Offset 1×1×32 Scale 1×1×32
ReLu	112×112×32	-
Maksimum Havuzlama	56×56×32	-
Evrişim Katmanı	56×56×16	Ağırlıklar 3×3×3×64 Sapma 1×1×64
Toplu Normalizasyon	56×56×16	Offset 1×1×16 Scale 1×1×16
ReLu	56×56×16	-
Evrişim Katmanı	56×56×8	Ağırlıklar 3×3×3×64 Sapma 1×1×64
Toplu Normalizasyon	56×56×8	Offset 1×1×8 Scale 1×1×8
ReLu	56×56×8	-
Evrişim Katmanı	56×56×4	Ağırlıklar 3×3×3×64 Sapma 1×1×64
Toplu Normalizasyon	56×56×4	Offset 1×1×4 Scale 1×1×4
ReLu	56×56×4	-
Tam Bağlı Katman	1×1×2	Ağırlıklar 2×12544 Sapma 2×1
Softmax	1×1×2	-
Sınıflandırma Çıktısı	-	-

4.4. Model Değerlendirme/Doğrulama Yöntemleri

Çapraz doğrulama, bir makine öğrenimi algoritmasının eğitilmediği yeni veri kümeleri üzerinde tahminlerde bulunma performansını değerlendirmek için kullanılan bir model değerlendirme tekniğidir. Bu, belirlenen veri kümesini bölümlere ayırarak, algoritmayı eğitmek için bir alt küme ve test için kalan verileri kullanarak yapılır. Her çapraz doğrulama turu, orijinal veri setinin rastgele bir şekilde bir eğitim setine ve bir test setine bölünmesini içerir. Eğitim seti daha sonra denetimli bir öğrenme algoritmasını eğitmek için kullanılır ve test seti de algoritmanın performansını değerlendirmek için kullanılır. Bu işlem birkaç kez tekrarlanır ve ortalama çapraz doğrulama hatası bir performans göstergesi olarak kullanılır. Çapraz doğrulama, model için en iyi algoritmayı bulmak amacıyla verileri farklı şekilde bölen çeşitli teknikler sunar [151,152].

4.4.1. K-Katlı Çapraz Doğrulama (K-Fold Cross Validation)

Çapraz doğrulama veya 'k-kat çapraz doğrulama' yönteminde tüm veri kümesi eşit büyüklükte k parçaya bölünür ve her bölüme bir kat adı verilir. K 3,4,5 vb. herhangi bir tam sayı olabilen k parça olduğu için k -katlı olarak bilinir. Gruplardan biri test seti olarak, diğer $k-1$ grup ise eğitim seti olarak kullanılır. Model eğitim seti üzerinde eğitilir ve test seti üzerinde değerlendirilir. Daha sonra bu işlem, her bir grup test kümesi olarak kullanılabileceği kadar k kez tekrarlanır. Şekil 4.4'de 5 katlı çapraz doğrulamayı ve dolayısıyla 5 iterasyonu göstermektedir. Her iterasyonda, bir kat test seti/doğrulama seti ve diğer $k-1$ setleri (4 set) eğitim setidir [153].

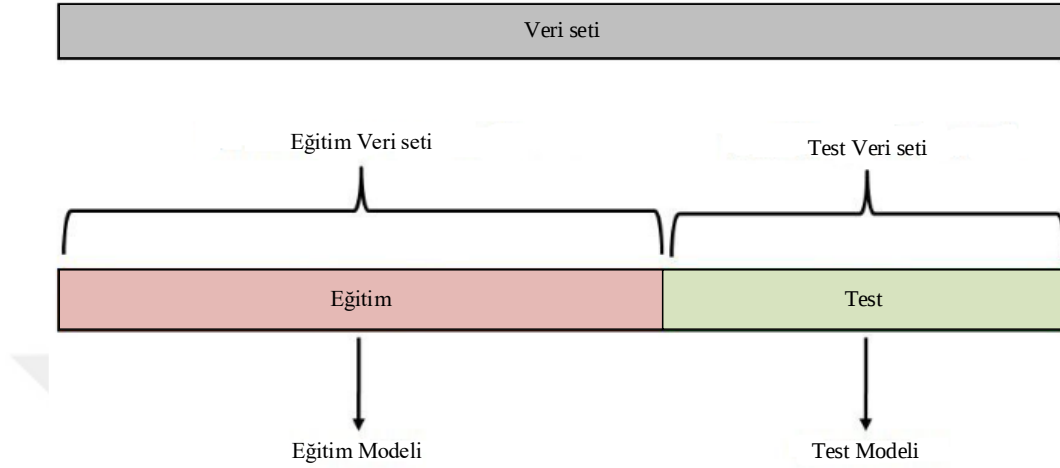


Şekil 4.4. 5 katlı çapraz doğrulama örneği [154]

4.4.2. Ayırarak çapraz doğrulama (Hold-out Cross Validation)

Ayırma, veri setinin bir eğitim ve test setine ayrılmasıdır. Tüm veri kümesinin rastgele bir şekilde bir eğitim kümesi ve bir test kümesi olarak bölünmesini sağlar. Eğitim seti, modelin üzerinde eğitildiği settir ve test seti, modelin görünmeyen veriler üzerinde ne kadar iyi performans

gösterdiğini görmek için kullanılır. Veri kümesi yalnızca iki sete bölündüğünden, model eğitim kümesi üzerinde yalnızca bir kez oluşturulur ve daha hızlı yürütülür. Şekil 4.5’de veri kümesi bir eğitim kümesi ve bir test kümesi olarak ikiye ayrılmıştır. Model eğitim kümesi üzerinde eğitilir ve test veri kümesi üzerinde test edilir [153].



Şekil 4.5. Hold out çapraz doğrulama örneği [154]

K katlı çapraz doğrulama daha çok tercih edilen yöntemdir çünkü modele birden fazla eğitim-test ayrımı üzerinde eğitim alma/test etme fırsatı verir. Bu, modelin görünmeyen veriler üzerinde ne kadar iyi performans göstereceğine dair daha iyi bir gösterge sağlar. Öte yandan hold out, yalnızca bir eğitim-test bölünmesine bağlıdır. Bu da hold-out yönteminin sınıflandırma performansının, verilerin eğitime ve test kümelerine nasıl bölündüğüne bağlı hale getirir. Bununla birlikte k katlı çapraz doğrulamada birden fazla eğitim-test bölmesi kullandığından, hold-out yöntemine göre daha fazla hesaplama gücü ve zaman gerektirir [153].

4.5. Değerlendirme Metrikleri

Eğitimden sonra oluşturulan modelin kalitesini değerlendirmek, performansını anlamak için çok önemlidir. Modelin performansını değerlendirmek için doğruluk, kesinlik, duyarlılık ve F1 skor gibi metrikler kullanılabilir. Bunlar, oluşturulan model tarafından verilen tahmine ve temel gerçek olarak hizmet eden etiketli verilere dayanan gerçek negatifler, gerçek pozitifler, yanlış negatifler ve yanlış pozitifler gibi parametreler aracılığıyla elde edilir. Gerçek Pozitif (TP), pozitif sınıfın doğru sınıflandırmasıdır, örneğin, pnömoni içeren bir görüntünün model tarafından pnömoni olarak sınıflandırılmasıdır. Gerçek Negatif (TN), negatif sınıfın doğru sınıflandırmasıdır, örneğin pnömoni içermeyen görüntünün yani sağlıklı görüntünün model tarafından pnömoni içermediği ve sağlıklı sınıfına ait olduğunun tespit edilmesidir. Yanlış Pozitif (FP), pozitiflerin yanlış tahminidir, örneğin pnömoni içeren bir görüntünün model tarafından yanlış etiketlenmesidir yani görüntünün

sağlıklı olarak sınıflandırılmasıdır. Yanlış Negatif (FN), negatiflerin yanlış tahmin edilmesidir, örneğin sağlıklı bir görüntünün model tarafından yanlış etiketlenmesidir yani görüntünün pnömoni olarak sınıflandırılmasıdır. Karışıklık matrisleri, tahmin edilen ve gerçek değerlerden elde edilen sonuçları bir tablo şeklinde temsil ederek modelin performansını görselleştirmenin bir yoludur [155-157]. Şekil 4.6’da modellerin performansını değerlendirmek için kullanılan karışıklık matrisinin bir gösterimi yer almaktadır.

		Tahmin etiketi	
		Var	Yok
Gerçek etiket	Var	Gerçek Pozitif (TP)	Yanlış Negatif (FN)
	Yok	Yanlış Pozitif (FP)	Gerçek Negatif (TN)

Şekil 4.6. Modelin performansının görselleştirilmesini kolaylaştıran bir karışıklık matrisi [157]

4.5.1. Kesinlik

Tahmin edilenlerden doğru gerçek pozitifleri kontrol ederek modelin ne kadar hassas çalıştığını kontrol eder. Kesinlik/pozitif öngörülen değer; Denklem 4.1’de gösterildiği gibi, doğru tespit edilen pozitif vakaların beklenen tüm pozitif vakalara oranıdır.

$$Kesinlik = \frac{TP}{TP+FP} \quad (4.1)$$

4.5.2. Duyarlılık

Modelin kaç tane gerçek (actual) TP yakaladığını hesaplar ve bunları pozitif olarak etiketler. Duyarlılık/hassasiyet/gerçek pozitif oranı; Denklem 4.2’de gösterildiği gibi, tüm gerçek pozitif vakalara kıyasla pozitif vaka olarak doğru şekilde tanımlanan örneklerdir.

$$Duyarlılık = \frac{TP}{TP+FN} \quad (4.2)$$

4.5.3. Doğruluk

Denklem 4.3’de yer aldığı üzere doğruluk, verilen toplam veri örneği sayısı üzerinden doğru olarak tanımlanan veri örneği sayısını ifade eder.

$$Doğruluk = \frac{TP+TN}{TP+TN+FP+FN} \quad (4.3)$$

4.5.4. F1 skor

Karşılaşılabilecek her durumu hesaba katmak için duyarlılık ve kesinliğin harmonik ortalaması olarak alınır ve yöntemin genel doğruluğunu gösterir.

$$F1 \text{ skor} = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (4.4)$$

4.6. İstatistiksel Yöntemler

Araştırma verilerinin istatistiksel değerlendirmesinde IBM SPSS İstatistik 26 programı kullanılmıştır. Verilerin değerlendirilmesinde nicel değişkenler ortalama $\pm$ standart sapma (SD) ile nitel değişkenler sayı ve yüzde (%) ile sunulmuştur. Nitel değişkenlerin karşılaştırılmasında Ki-kare (χ^2) testi kullanılmıştır. Nicel değişkenlerin normal dağılıma uygunluğu Kolmogorov-Smirnov testi ile değerlendirilmiştir. Grupların karşılaştırılmasında Mann-Whitney U testi kullanılmıştır. Hipotezler çift yönlü alınarak, $p \leq 0.05$ ise istatistiksel olarak anlamlı sonuç kabul edilmiştir.

Mann-Whitney U testi, normal dağılıma sahip olmayan sıralı veriler veya sürekli veriler üzerinde iki bağımsız grup arasındaki farklılıkları test etmek için kullanılan parametrik olmayan testlerdendir [158]. Ki-kare anlamlılık testi, kategorik ve nitel değişkenler arasındaki ilişkinin istatistiksel olarak anlamlı olup olmadığını belirlemek için kullanılan parametrik olmayan bir testtir [159]. Kolmogorov-Smirnov testi, örneklem büyüklüğüne bakılmaksızın normal dağılıma sahip olup olmadığını ve istatistiksel analizde anlamlı sonuçlar elde etmek için kullanılır [160].

5. BULGULAR VE TARTIŞMA

Bu tez çalışması kapsamında seçilen tüm ağların derin öğrenme yöntemlerindeki başarımlarını değerlendirmek için kapsamlı deneyler gerçekleştirilmiştir. Bu bölümde CXR görüntülerinden ve sayısal tıbbi parametrelerden derin öğrenme yaklaşımları ile pnömoni tespiti için deneysel çalışmaların sonuçları detaylı bir şekilde yer almaktadır. İlk önce CXR görüntülerinden elde edilen bulgular tablo şeklinde verilmiştir. Literatürde benzer çalışmaların sonuçları tez kapsamında kullanılan yöntemlerle karşılaştırılmıştır. Daha sonra sayısal parametre veri setinden elde edilen bulgular ve literatür karşılaştırması yer almaktadır.

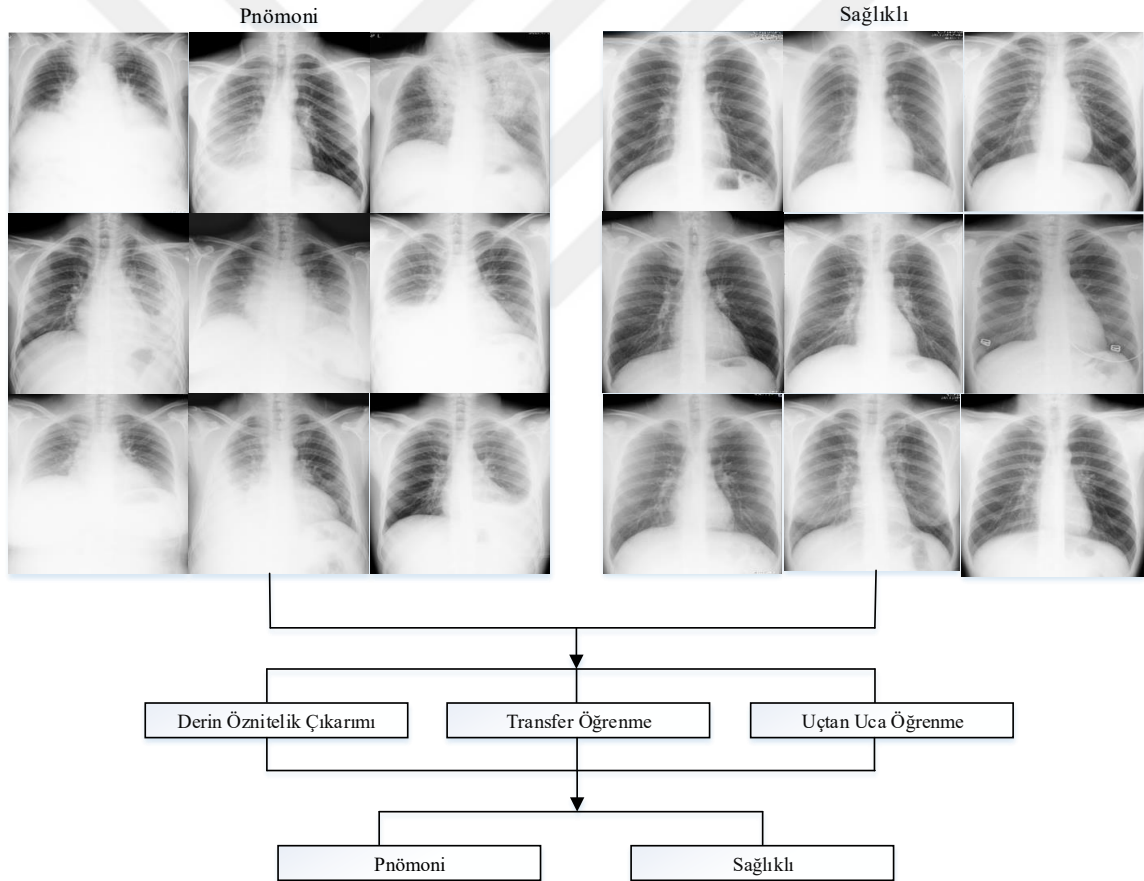
CXR görüntülerinden derin öğrenme algoritmalarını ve yöntemlerini kullanarak pnömoni tespitinde iki farklı çalışma yapılmıştır. İlk çalışmada pnömoni ve sağlıklı CXR görüntülerini sınıflandırmak için derin öğrenmeye dayalı yaklaşımlar kullanılmıştır. Bu yaklaşımlar, derin öznitelik çıkarımı, transfer öğrenme ve tez çalışması kapsamında geliştirilmiş bir ESA modelinin uçtan uca eğitimidir. Derin öznitelik çıkarımı ve transfer öğrenme için 10 farklı önceden eğitilmiş ESA modelleri kullanılmıştır. Bu çalışmada kullanılan önceden eğitilmiş ESA modelleri AlexNet, VGG16, VGG19, ResNet50, DenseNet201, DarkNet53, ShuffleNet, SqueezeNet, MobileNetV2 ve NasNetMobile'dir. İkinci çalışmada geliştirilen yeni ESA modeli kullanılarak elde edilen sonuçlar dönüştürücü tabanlı mimariler ile karşılaştırılmıştır. Bu dönüştürücü mimarileri, ViT, gMLP, MLP-mixer ve FNet'tir. Önerilen yöntemlerin etkinliğini değerlendirmek için doğruluk, kesinlik, duyarlılık ve F1-ölçümü gibi performans ölçütleri kullanılmıştır.

5.1. Göğüs röntgen görüntülerinde pnömoni tespiti için derin öğrenme modellerinin incelenmesi

Yapılan çalışmada önceden eğitilmiş AlexNet, VGG16, VGG19, ResNet50, DenseNet201, DarkNet53, ShuffleNet, SqueezeNet, MobileNetV2 ve NasNetMobile modelleri ile derin öznitelik çıkarımı ve transfer öğrenme çalışmaları gerçekleştirilmiştir. AlexNet'in fc6; VGG16 ve VGG19'un fc7; ResNet50 ve DenseNet201'in fc1000; DarkNet53'ün conv53; NasNetMobile'in predictions; MobileNetV2'nin Logits; ShuffleNet'in node_202 ve SqueezeNet'in pool10 katmanlarından öznitelikler çıkarılıp sınıflandırıcıya verilmiştir. Çıkarılan özellikler daha sonra RO, DVM ve KNN dahil olmak üzere çeşitli makine öğrenimi algoritmalarına girildikleri bir sonraki adımda kullanılmıştır. Deneysel çalışmalarda kullanılan DVM hiper parametreleri; çekirdek fonksiyonu 'linear' olarak belirlenmiştir. KNN hiper parametrelerinde komşu sayısı '10' olarak belirlenmiştir. Uzaklık ölçü birimi 'minkowski' olarak seçilmiştir. RO hiper parametrelerinde eğitim metodu 'bag', toplulukta kullanılacak ağaç sayısı 50 seçilmiştir. Uçtan uca öğrenme için ise 21 katmanlı model oluşturularak sıfırdan eğitilmiştir. 21 katmanlı yeni ESA hiper

parametreleri; 'trainingOptions', 'InitialLearnRate', 'MaxEpochs' ve 'MiniBatchSize' sırasıyla 'SGDM', 0.0001, 8 ve 10'dur.

Çalışmanın ön işleme aşaması olarak; kullanılan farklı algoritmaların görüntü girişleri farklı olduğu için CXR görüntüleri yeniden boyutlandırılmıştır. Hem derin öznetelik çıkarma hem de ince ayar prosedürlerinde önceden eğitilmiş ESA modellerini eşit şartlar altında karşılaştırmak için tüm modeller minimum girdi boyutlarını karşılayacak ortak boyut olan 256×256 yeniden boyutlandırılmıştır. Giriş CXR görüntüleri, tez çalışması kapsamında geliştirilen ESA modeli için 224×224 piksel olarak yeniden boyutlandırılmıştır. Deneysel çalışmalarda veri seti hem belirli oranlarda eğitim ve test veri setlerine ayrılmıştır, hem de k-katlı çapraz geçerlilik testi uygulanmıştır. Böylece veri seti %80 eğitim ve %20 test olarak ayrılmıştır. Ayrıca k-katlı çapraz geçerlilik testi için k değeri 5 seçilmiştir. Toplanan CXR görüntülerinden pnömoni tespiti için önerilen yöntemin akış diyagramı Şekil 5.1'de yer almaktadır.



Şekil 5.1. CXR görüntüleri ile pnömoni tespiti için önerilen yöntemin akış diyagramı [161]

İlk deneyler derin öznitelik çıkarımı ve makine öğrenmesi sınıflandırıcıları ile gerçekleştirilmiştir. Önceden eğitilmiş 10 adet ESA modelinin ilk derin öznitelik çıkarımından elde edilen sonuçlar sırasıyla Tablo 5.1, Tablo 5.2 ve Tablo 5.3’de sunulmuştur. Tablolarda her bir modelin en yüksek performans metrik puanları koyu ile gösterilmiştir. Tablo 5.1, derin özniteliklerin DVM ile sınıflandırılmasından elde edilen sonuçları gösterirken; Tablo 5.2, aynı özniteliklerin KNN sınıflandırıcısından üretilen doğruluk puanlarını göstermektedir. Tablo 5.3, derin özniteliklerin RO ile sınıflandırılmasından elde edilen sınıflandırma performans metriklerini göstermektedir.

Tablo 5.1. Pnömoni/Sağlıklı sınıflandırmasında önceden eğitilmiş ESA modellerinin DVM başarıları

Derin Öznitelikler + DVM (%80 Eğitim %20 Test)	Doğruluk (%)	Kesinlik (%)	Duyarlılık (%)	F1 skor (%)
AlexNet (fc6)	97,00	95,50	98,45	96,95
SqueezeNet(pool10)	97,25	96,00	98,46	97,22
ResNet50(fc1000)	96,50	96,00	96,97	96,48
ShuffleNet(node202)	96,75	96,00	97,46	96,73
NasNetMobile(predictions)	95,75	93,00	98,41	95,63
MobileNetV2(Logits)	96,00	95,50	96,46	95,98
VGG16 (fc7)	96,50	95,50	97,45	96,46
VGG19 (fc7)	96,00	95,00	96,94	95,96
DenseNet201(fc1000)	95,00	94,00	95,92	94,95
DarkNet53(conv53)	93,50	95,50	97,45	96,46

Önceden eğitilmiş ESA modellerine ait derin öznitelik çıkarımı deneysel çalışmalarının sonuçları Tablo 5.1’de yer almaktadır. Bu çalışma yürütülürken veri setinin %80’i eğitim için kullanılmış, önceden eğitilmiş modeller veri setinin %20’si ile test edilmiştir. Tablo 5.1’deki satırlar model türünü, sütunlar ise elde edilen sınıflandırma performansı parametrelerini göstermektedir. Tablo 5.1’den, SqueezeNet modelinin %97,25 ortalama doğruluk puanı ile en yüksek ortalama doğruluk puanını ürettiği, AlexNet, ShuffleNet, VGG16, ResNet50, MobilenetV2, VGG19 ve NasNetMobile modellerinin ise sırasıyla %97,00, %96,75, %96,50, %96,50, %96,00, %96,00 ve %95,75 ortalama doğruluk puanı ürettiği görülebilir. DenseNet201 modeli %95,00 ortalama doğruluk puanı üretirken, DarkNet53 modeli %93,50 ile en düşük doğruluk puanları üretmiştir. SqueezeNet, ResNet50 ve ShuffleNet %96,00 ile aynı kesinlik puanını elde ederken NasNetMobile %93,00 ile en düşük kesinlik başarımı göstermiştir. AlexNet, VGG16, MobileNetV2 ve DarkNet53 %95,50 kesinlik başarımı üretirken; VGG19 %95,00 ve DenseNet201 %94,00 kesinlik performansı göstermiştir. SqueezeNet, AlexNet ve NasNetMobile sırasıyla %98,46, %98,45 ve %98,41 ile en yüksek duyarlılık puanları elde ederken, DenseNet201 %95,92 ile en düşük duyarlılık başarımı göstermiştir. VGG16 ve DarkNet53 modelleri %97,45; ResNet50 %96,97; VGG19 %96,94 ve MobileNetV2 %96,46 duyarlılık performansı göstermiştir.

SqueezeNet, AlexNet, ShuffleNet ve ResNet50 sırasıyla %97,22, %96,95, %96,73, %96,48 F1 skor puanı üretirken, DenseNet201 %94,95 ile en düşük F1 skor puanı elde etmiştir. VGG16 ve DarkNet53 %96,46; VGG19 %95,96 F1 skoru elde etmiştir.

Tablo 5.2, her bir model için elde edilen sınıflandırmanın değerlendirme metriklerini vermektedir. Önceden eğitilmiş ESA modellerinden çıkarılan derin öznitelikler KNN algoritması ile sınıflandırılmıştır. Bu çalışma ile hibrit modellerin başarımları da sınıflandırma performansı açısından karşılaştırılmaktadır.

Tablo 5.2. Pnömoni/Sağlıklı sınıflandırmasında önceden eğitilmiş ESA modellerinin KNN başarımları

Derin Öznitelikler + KNN (%80 Eğitim %20 Test)	Doğruluk (%)	Kesinlik (%)	Duyarlılık (%)	F1 skor (%)
AlexNet	96,00	92,50	99,46	95,85
SqueezeNet	89,50	82,50	95,93	88,71
ResNet50	94,00	94,00	94,00	94,00
ShuffleNet	93,75	89,50	97,81	93,47
NasNetMobile	87,50	84,50	89,89	87,11
MobileNetV2	93,00	87,50	98,31	92,59
VGG16	91,50	91,00	91,92	91,46
VGG19	92,00	89,50	94,21	91,79
DenseNet201	93,25	89,00	97,27	92,95
DarkNet53	92,75	95,00	90,91	92,91

Tablo 5.2'deki sonuçlar göz önüne alındığında AlexNet'in fc6 katmanından çıkarılan özniteliklerin sınıflandırmada %96,00 doğruluk puanı ile en yüksek ortalama doğruluk puanını ürettiği, ResNet50'nin fc1000 katmanı kullanılarak çıkarılan özniteliklerin sınıflandırmada ikinci en iyi puan olarak %94,00 doğruluk puanı ürettiği görülebilir. ShuffleNet modeli %93,75 ortalama doğruluk puanı üretirken, NasNetMobile modeli %87,50 ile en düşük doğruluk oranına sahiptir. ShuffleNet %93,75, DenseNet201 %93,25, MobileNetV2 %93,00, DarkNet53 %92,75, VGG19 %92,00 ve VGG16 %91,50 doğruluk performansı göstermiştir. Darknet53 %95,00 ile en yüksek kesinlik puanı elde ederken; ResNet50 %94,00, AlexNet %92,50 ve VGG16 %91,00 kesinlik puanı elde etmiştir. SqueezeNet %82,50 ve NasNetMobile %84,00 ile en düşük kesinlik başarımları göstermiştir. ShuffleNet ve VGG19 % 89,50, DenseNet201 %89,00, MobileNetV2 %87,50 kesinlik elde etmiştir. AlexNet %99,46 ile en yüksek duyarlılık elde ederken; MobileNetV2 %98,31 ile en yüksek ikinci duyarlılık değeri elde etmiştir. ShuffleNet %97,81, Densenet201 %97,27, SqueezeNet %95,93, VGG19%94,21 ve ResNet50 %94,00 duyarlılık puanı elde etmiştir. %89,89 ile NasNetMobile en düşük duyarlılık puanı üretmiştir. AlexNet ve ResNet50 sırasıyla %95,85 ve %94,00 ile en yüksek F1 skor puanları elde ederken; NasNetMobile %87,11 ile en düşük F1 skor puanı elde etmiştir. ShuffleNet %93,47, DenseNet201 %92,95, Darknet53 %92,91 ve MobileNetV2 %92,59 F1 skor performansı göstermiştir.

Tablo 5.3. Pnömoni/Sağlıklı sınıflandırmasında önceden eğitilmiş ESA modellerinin RO başarıları

Derin Öznitelikler + RO (%80 Eğitim %20 Test)	Doğruluk (%)	Kesinlik (%)	Duyarlılık (%)	F1 skor (%)
AlexNet	95,75	96,00	95,52	95,76
SqueezeNet	91,50	90,50	92,35	91,41
ResNet50	93,50	93,00	93,94	93,47
ShuffleNet	95,25	94,50	95,94	95,21
NasNetMobile	90,75	90,00	91,37	90,68
MobileNetV2	94,00	95,00	93,14	94,06
VGG16	92,50	92,00	92,93	92,46
VGG19	92,25	92,00	92,46	92,23
DenseNet201	92,75	93,50	92,12	92,80
DarkNet53	93,25	93,00	93,47	93,23

Önceden eğitilmiş ESA modellerinden çıkarılan derin özniteliklerin RO sınıflandırma performansları Tablo 5.3’de yer almaktadır. AlexNet’in fc6 katmanından çıkarılan özniteliklerin sınıflandırmada %95,75 doğruluk puanı ile en yüksek ortalama doğruluk puanını ürettiği, ShuffleNet’in node202 katmanı kullanılarak çıkarılan özniteliklerin sınıflandırmada ikinci en iyi puan olarak %95,25 doğruluk puanı ürettiği görülebilir. MobileNetV2 modeli %94,00 ortalama doğruluk puanı üretirken, NasNetMobile modeli %90,75 ile en düşük doğruluk oranına sahiptir. ResNet50, DarkNet53, DenseNet201, VGG16 ve VGG19 modellerinin doğruluk oranları sırasıyla %93,50, %93,25, %92,75, %92,50 ve %92,25’dir. SqueezeNet modeli %91,50 doğruluğa ulaşmıştır. AlexNet %96,00 ile en yüksek kesinlik puanı elde ederken; MobileNetV2 %95,00, ShuffleNet %94,50 ve DenseNet201 %93,50 kesinlik puanı elde etmiştir. VGG modelleri %92,00 kesinlik başarımları gösterirken; SqueezeNet %90,50 ile NasNetMobile %90,00 ile en düşük kesinlik başarımları göstermiştir. ShuffleNet %95,94 ile en yüksek duyarlılık elde ederken; AlexNet %95,52 ile en yüksek ikinci duyarlılık değeri elde etmiştir. %91,37 ile NasNetMobile en düşük duyarlılık puanı üretmiştir. AlexNet ve ShuffleNet sırasıyla %95,76 ve %95,21 ile en yüksek F1 skor puanları elde ederken; NasNetMobile %90,68 ile en düşük F1 skor puanı elde etmiştir. MobileNetv2 %94,06 ile ortalama bir F1 skor başarımları gösterirken; ResNet50 %93,47, DarkNet53 %93,23, DenseNet201 %92,80, VGG16 %92,46 ve VGG19 %92,23 F1 skor performansı üretmiştir.

Tablo 5.4. Pnömoni/Sağlıklı sınıflandırmasında önceden eğitilmiş ESA modellerinin DVM başarıları

Derin Öznitelikler + DVM (5 kat çapraz doğrulama)	Doğruluk (%)	Kesinlik (%)	Duyarlılık (%)	F1 skor (%)
AlexNet	97,75	97,00	98,48	97,73
SqueezeNet	97,50	98,50	96,57	97,52
ResNet50	96,75	96,50	96,98	96,74
ShuffleNet	98,00	97,00	98,98	97,98
NasNetMobile	96,25	95,00	97,44	96,20
MobileNetV2	96,75	96,50	96,98	96,74
VGG16	96,00	96,00	96,00	96,00
VGG19	97,00	97,50	96,53	97,01
DenseNet201	97,25	96,50	97,97	97,23
DarkNet53	96,00	95,00	96,46	95,98

İkinci derin öznitelik çıkarma deneylerinde veri seti 5 kat çapraz doğrulama kullanılarak ayrılmıştır. Tablo 5.4’de önceden eğitilmiş ESA modellerinin DVM sınıflandırma başarımları yer almaktadır. Görüldüğü gibi, neredeyse tüm ağlar tüm ölçütlerde %95,00’in üzerinde bir puanla tatmin edici sonuçlar elde edebilmektedir, bu da derin öznitelik çıkarımı yönteminin pnömoni tespitinde etkili bir yaklaşım olduğunu göstermektedir. Dahası, tüm modeller yüksek kesinlik puanları elde etmektedir, bu da her sınıf için genellikle düşük sayıda yanlış pozitif olduğunu, yani pnömoni sınıfının tahmin sırasında karıştırılmama eğiliminde olduğunu göstermektedir. ShuffleNet %98,00 doğruluk puanı ile en yüksek doğruluğa ulaşırken; AlexNet %97,75, SqueezeNet %97,50, DenseNet201 %97,25 doğruluk elde etmiştir. VGG16 ve DarkNet53 %96,00 ile en düşük doğruluk puanı üretmiştir. VGG19 %97,00, ResNet50 ve MobileNetV2 %96,75 doğruluk başarımları elde etmiştir. SqueezeNet %98,50 ile en yüksek kesinlik puanı elde ederken; DarkNet53 ve NasNetMobile %95,00 ile diğer modellere göre daha düşük kesinlik puanı elde etmiştir. VGG19 %97,50; AlexNet ve ShuffleNet %97,00; ResNet50, MobileNetV2 ve DenseNet201 %96,50 kesinlik performansı sergilemiştir. ShuffleNet %98,98 ve AlexNet %98,48 ile en yüksek duyarlılık puanlarını üretirken; VGG16 %96,00 ile en düşük duyarlılık başarımları göstermiştir. DenseNet201 %97,97; NasNetMobile %97,44; MobileNetV2 ve ResNet50 %96,98 duyarlılık başarımları göstermiştir. ShuffleNet, AlexNet ve SqueezeNet sırasıyla %97,98, %97,73, %97,52 ile ortalama en yüksek F1 skor puanı elde etmiştir. DarkNet53 en düşük F1 skor puanını %95,98 ile göstermiştir. DenseNet201 %97,23; VGG19 %97,01; MobileNetV2 ve ResNet50 %96,74 F1 ölçüm sonucu elde etmiştir.

Tablo 5.5. Pnömoni/Sağlıklı sınıflandırmasında önceden eğitilmiş ESA modellerinin KNN başarıları

Derin Öznitelikler+KNN (5 kat çapraz doğrulama)	Doğruluk (%)	Kesinlik (%)	Duyarlılık (%)	F1 skor (%)
AlexNet	96,25	94,50	97,93	96,18
SqueezeNet	91,50	88,00	94,62	91,19
ResNet50	93,25	95,00	91,79	93,37
ShuffleNet	95,75	94,50	96,92	95,70
NasNetMobile	90,00	88,00	91,67	89,80
MobileNetV2	92,50	89,50	95,21	92,27
VGG16	93,50	93,00	93,94	93,47
VGG19	92,50	91,00	93,81	92,39
DenseNet201	93,25	89,00	97,27	92,95
DarkNet53	91,75	94,00	89,95	91,93

Tablo 5.5’de önceden eğitilmiş ESA modellerinin KNN sınıflandırma başarımları yer almaktadır. AlexNet %96,25 doğruluk puanı ile en yüksek doğruluğa ulaşırken; ShuffleNet %95,75, VGG16 %93,50, DenseNet201 ve ResNet50 %93,25 doğruluk elde etmiştir. NasNetMobile %90,00 ile en düşük doğruluk puanı üretmiştir. SqueezeNet %91,50 ve DarkNet53 %91,75 sınıflandırma doğruluğu üretmiştir. ResNet50 %95,00 ile en yüksek kesinlik puanı elde ederken; SqueezeNet ve NasNetMobile %88,00 ile diğer modellere göre daha düşük kesinlik puanı elde etmiştir. AlexNet ve ShuffleNet %94,50; DarkNet53 %94,00; VGG16 %93,00; VGG19 %91,00 ortalama kesinlik performansı göstermiştir. DenseNet201 %89,00 ve MobileNetV2 %89,50 ile düşük kesinlik başarımları gösteren modellerdendir. AlexNet %97,93 ve DenseNet201 %97,27 ile en yüksek duyarlılık puanlarını üretirken; DarkNet53 %89,95 ile en düşük duyarlılık başarımları göstermiştir. ShuffleNet %96,92 ve MobileNetv2 %95,21 ile ortalama yükseklikte bir duyarlılık başarımları elde etmiştir. ResNet50 %91,79 ile NasNetMobile %91,67 düşük duyarlılık puanı üreten mimarilerdendir. AlexNet ve ShuffleNet sırasıyla %96,18, %95,70 ile ortalama en yüksek F1 skor puanı elde etmiştir. NasNetMobile en düşük F1 skor puanını %89,80 ile göstermiştir. VGG16 %93,47; ResNet50 %93,37; DenseNet201 %92,95; VGG19 %92,39; MobileNetV2 %92,27 F1 skor puanı elde etmiştir.

Tablo 5.6. Pnömoni/Sağlıklı sınıflandırmasında önceden eğitilmiş ESA modellerinin RO başarıları

Derin Öznitelikler+RO (5 kat çapraz doğrulama)	Doğruluk (%)	Kesinlik (%)	Duyarlılık (%)	F1 skor (%)
AlexNet	96,75	97,50	96,06	96,77
SqueezeNet	93,00	94,00	92,16	93,07
ResNet50	95,00	94,00	95,92	94,95
ShuffleNet	97,00	97,00	97,00	97,00
NasNetMobile	91,25	91,00	91,46	91,23
MobileNetV2	93,00	93,00	93,00	93,00
VGG16	93,50	93,00	93,94	93,47
VGG19	93,25	93,50	93,03	93,27
DenseNet201	95,25	95,00	95,48	95,24
DarkNet53	90,50	90,00	90,91	90,45

Veri setinin 5 kat çapraz doğrulamayla ayrıldığı son derin öznitelik çıkarımı deneysel çalışması Tablo 5.6’da sunulmuştur. Tablodan görüleceği üzere, tüm ağlar tüm ölçütlerde %90,00’in üzerinde sınıflandırma performans sonuçları elde etmiştir, bu da RO algoritmasının CXR görüntülerine ait derin öznitelikleri yüksek bir oranla tanıyabildiğini göstermektedir. Önceden eğitilmiş ESA modellerinden çıkarılan özniteliklerin RO sınıflandırma performansları açısından ShuffleNet %97,00 ile en yüksek doğruluğu elde etmiştir. AlexNet %96,75, DenseNet201 %95,25 ve ResNet50 %95,00 ortalama doğruluk puanları üretmiştir. DarkNet53 %90,50 ile en düşük doğruluk puanına sahiptir. VGG16 %93,50; VGG19 %93,25; SqueezeNet ve MobileNetV2 %93,00 sınıflandırma doğruluğu elde etmiştir. Kesinlik başarımları açısından AlexNet %97,50 ile en yüksek; ShuffleNet %97,00 ile en yüksek ikinci puanı üretmiştir. DarkNet53 %90,00 ve NasNetMobile %91,00 ile en düşük kesinlik puanı elde etmiştir. DenseNet201 %95,00; SqueezeNet ve Resnet50 %94,00; VGG19 %93,50; VGG16 ve MobileNetV2 %93,00 kesinlik puanına ulaşmıştır. Önceden eğitilmiş ESA modelleri duyarlılık açısından karşılaştırıldığında DarkNet53 %90,91 ile en düşük; ShuffleNet %97,00 ile en yüksek puanları elde etmiştir. AlexNet %96,06; DenseNet201 %95,48; ResNet50 %95,92; VGG16 %93,94; VGG19 %93,03 başarımlarını sergilemişlerdir. F1 skor açısından da DarkNet53 %90,45 ile en düşük; ShuffleNet %97,00 ile en yüksek puanları üretmiştir. AlexNet %96,77; DenseNet201 %95,24; ResNet50 %94,95; VGG16 %93,47 ve VGG19 % 93,27 F1 skor puanına ulaşmıştır.

İki derin öznitelik çıkarımı deneylerinin sonuçları karşılaştırıldığında; 5 katlı çapraz doğrulama kullanılarak ayrılan veri seti ile yapılan test sonuçlarının daha yüksek doğruluk puanlarına ulaştığı görülmektedir. K-katlı çapraz doğrulama kullanılırken, verilerin tüm bölümleri test verilerinin bir parçası olarak kullanılabilir. Bu şekilde, veri kümesindeki tüm veriler hem eğitim hem de test için kullanılabilir ve modelin performansının daha iyi değerlendirilmesini sağlar [152,161].

Çapraz doğrulama kullanarak daha fazla metrik elde edilebilir ve hem algoritma hem de veriler hakkında önemli sonuçlar çıkarılabilmektedir. Seçilen algoritmanın veya verilerin (veya her ikisinin) tutarlı olup olmadığı, algoritmanın problemi iyi öğrenip öğrenemediği veya verilerin karmaşık olup olmadığı gibi bilgiler sınıflandırma performans metriklerinden anlaşılabilmektedir.

İlk deneysel çalışmalar değerlendirildiğinde önceden eğitilmiş ESA modellerinden çıkarılan derin özneteliklerin makine öğrenmesi algoritmalarıyla sınıflandırılmasında DVM algoritması hem k kat çapraz doğrulama hem de ayırarak çapraz doğrulama yöntemlerinde her değerlendirme metriğinde ortalama %90,00'nin üzerinde sınıflandırma performansı sergilemiştir. Böyle bir sonuç, DVM algoritmasının diğer makine öğrenmesi algoritmalarına kıyasla pnömoniden etkilenenleri sağlıklı bireylerden daha doğru bir şekilde ayırt edebildiğini göstermektedir. Sonuç olarak, önerilen derin öğrenme çerçeveleri, pnömoni vakalarının yüksek doğrulukta sınıflandırılmasını ve teşhisini sağlamak için çıkarılan özelliklerden ve makine öğrenimi algoritmalarından başarıyla yararlanmaktadır.

Tablo 5.7. Pnömoni/Sağlıklı sınıflandırmasında ESA modellerinin transfer öğrenme başarıları

Transfer Öğrenme (%80 Eğitim %20 Test)	Doğruluk (%)	Kesinlik (%)	Duyarlılık (%)	F1 skor (%)
AlexNet	98,50	97,00	100,00	98,48
SqueezeNet	95,00	91,50	98,39	94,82
ResNet50	97,25	94,50	100,00	97,17
ShuffleNet	97,00	98,50	95,63	97,04
NasNetMobile	98,00	98,00	98,00	98,00
MobileNetV2	97,75	96,00	99,48	97,71
VGG16	97,50	98,50	96,57	97,52
VGG19	95,75	99,50	92,56	95,90
DenseNet201	94,75	89,50	100,00	94,46
DarkNet53	96,50	94,50	98,44	96,43

Tablo 5.7'de önceden eğitilmiş modellerin ince ayarının doğruluk sonuçları gösterilmiştir. Veri seti %80 eğitim ve %20 test olarak ayrılmıştır. Sınıflandırma doğruluğu, kesinlik ve F1 puanı gibi performans ölçütlerinin sonuçları, AlexNet modelinin diğer mimarilerden daha iyi bir performans elde ettiğini göstermektedir. En yüksek doğruluk puanı %98,50 ile AlexNet modeli tarafından üretilirken; ikinci en iyi doğruluk puanı %98,00 ile NasNetMobile modelinden elde edilmiştir. MobileNetV2, VGG16, ResNet50 ve ShuffleNet modelleri sırasıyla %97,75, %97,50, %97,25 ve %97,00 doğruluk puanları elde edilmiştir. SqueezeNet %95,00 ve VGG19 %95,75 ile düşük doğruluk başarımı gösteren mimarilerdendir. Densenet201 %94.75 ile en düşük doğruluk puanı elde etmiştir. Kesinlik puanları açısından karşılaştırıldığında sonuçlar %89,50 ile %99,50 arasında değişmektedir. VGG19 %99,50 ile en yüksek; DenseNet201 %89,50 ile en düşük kesinlik puanlarını üretmiştir. ResNet50 ve DarkNet53 %94,50 ile ortalama kesinlik değeri elde etmiştir.

ShuffleNet ve VGG16 %98,50; NasNetMobile %98,00; AlexNet %97,00 ve MobileNetV2 %96,00 kesinlik skoru üretmiştir. AlexNet, ResNet50 ve DenseNet201 %100,00 duyarlılık değeri üretmiştir. %99,48 ile en yüksek ikinci duyarlılık değerinin MobileNetV2 üretmiştir. VGG19 %92,56; ShuffleNet %95,63 ile diğer modellere göre daha düşük duyarlılık elde etmiştir. DarkNet53 %98,44; SqueezeNet %98,39 ve NasNetMobile %98,00 duyarlılık değeri elde etmiştir. F1 skor değerleri %94,46 ile %98,48 arasında değişmektedir. Alexnet %98,48 ile en yüksek F1 skor puanı elde ederken; ResNet50 %97,17; ShuffleNet %97,04 puan elde etmiştir. VGG16, NasNetMobile, MobileNetV2 modelleri sırasıyla %98,00, %97,71, %97,52 elde etmiştir. DenseNet201'in ürettiği %94,46 ve SqueezeNet'in ürettiği %94,82 düşük F1 skor değerlerindendir.

Tablo 5.8. Pnömoni/Sağlıklı sınıflandırmasında ESA modellerinin transfer öğrenme başarıları

Transfer Öğrenme (5 kat çapraz doğrulama)	Doğruluk (%)	Kesinlik (%)	Duyarlılık (%)	F1 skor (%)
AlexNet	94,00	90,60	97,21	93,79
SqueezeNet	92,00	89,00	94,78	91,80
ResNet50	98,00	98,00	98,00	98,00
ShuffleNet	97,75	97,80	97,70	97,75
NasNetMobile	97,10	99,50	94,94	97,17
MobileNetV2	97,15	98,50	95,91	97,19
VGG16	98,25	99,00	97,54	98,26
VGG19	97,00	99,50	94,76	97,08
DenseNet201	98,10	98,50	97,72	98,11
DarkNet53	98,00	97,80	98,10	97,95

Tablo 5.8'de önerilen ESA modellerinin transfer öğrenme yöntemine ait 5 katlı çapraz doğrulama test sonuçları yer almaktadır. Modeller doğruluk değerleri açısından karşılaştırıldığında; VGG16 %98,25 ile en yüksek puana ulaşırken, DenseNet201 %98,10, Resnet50 ve DarkNet53 %98,00 başarımlarını göstermiştir. ShuffleNet %97,75, MobileNetV2 %97,15, NasNetMobile %97,10 doğruluk değeri üretmiştir. SqueezeNet %92,00 ile en düşük transfer öğrenme doğruluğunu göstermiştir. VGG19 ve NasNetMobile %99,50 ile en yüksek kesinlik değeri elde etmiştir. VGG16 %99,00; DenseNet201 ve MobileNetV2 %98,50; ResNet50 %98,00 kesinlik puanına erişmiştir. SqueezeNet %89,00 ve AlexNet %90,60 ile düşük kesinlik performansı gösteren modellerdir. DarkNet53 %98,10, ResNet50 %98,00 ile en yüksek duyarlılık değerlerini üretmiştir. NasNetMobile %94,94 ile en düşük duyarlılık performansı göstermiştir. DenseNet201, ShuffleNet, VGG16 ve AlexNet mimarilerinin duyarlılık puanları sırasıyla %97,72, %97,70, %97,54 ve %97,21'dir. VGG16'nin F1 skor değeri %98,26'dır, %91,80 ile en düşük F1 skor performansı %91,80 ile SqueezeNet'e aittir. DenseNet201, ResNet50, DarkNet53 ve ShuffleNet için F1 skor değerleri sırasıyla %98,11, %98,00, %97,95 ve %97,75'dir.

Veri setinin %80 eğitim, %20 test olarak ayrıldığı derin öğrenme deneylerinden elde edilen sonuçlar karşılaştırıldığında ince ayarlı AlexNet modeli %98,50 ile derin öznitelik çıkarımı ve uçtan uca öğrenme sonuçlarından daha iyi sonuç elde ettiği açıkça görülmektedir. Veri setinin 5 kat çapraz doğrulama ile ayrıldığı derin öğrenme deneylerinden elde edilen sonuçlar karşılaştırıldığında ince ayarlı VGG16 modeli %98,25 ile derin öznitelik çıkarımı ve uçtan uca öğrenme sonuçlarından daha iyi sonuç elde ettiği açıkça görülmektedir. Bu sonuç, transfer öğrenmenin etkinliğini vurgulamakta ve pnömoninin erken teşhisi için önceden var olan modellerin becerisi sayesinde pratik bir strateji olabileceğini göstermektedir.

Tablo 5.9. Geliştirilen ESA modelinin uçtan uca eğitimi ile elde edilen sınıflandırma performansı

ESA (%80eğitim %20test)	Doğruluk (%)	Kesinlik (%)	Duyarlılık (%)	F1 skor (%)
21 katman	96,75	96,50	97,00	96,75

Tablo 5.10. Uçtan uca eğitilmiş ESA modelinin çapraz doğrulama performansları

ESA (5 kat çapraz doğrulama)	Doğruluk (%)	Kesinlik (%)	Duyarlılık (%)	F1 skor (%)
21 katman	95,75	93,50	97,91	95,65

Tablo 5.9 ve Tablo 5.10’da geliştirilen ESA modelinin uçtan uca eğitiminden k kat çapraz doğrulama ve ayırarak çapraz doğrulama ile elde edilen sınıflandırma performansları gösterilmiştir. Tablo 5.9’de 21 katmanlı ESA %96,75 doğruluk değeri elde etmiştir. Ayırarak doğrulama ile yeni ESA modelinin kesinliği %96,50, duyarlılığı %97,00 ve F1 skor ölçütü %96,75’dir. Tablo 5.10’da görüleceği üzere k kat çapraz doğrulama ile yeni ESA modelinin doğruluğu %95,75, kesinliği %93,50, duyarlılığı %97,91 ve F1 skor ölçütü %95,65’dir.

Tablo 5.11. Önerilen modelin CXR görüntü veri setindeki performans metrikleri

ESA (5 kat çapraz doğrulama)	Doğruluk (%)	Kesinlik (%)	Duyarlılık (%)	F1 skor (%)
21 katman	94,50	90,50	98,37	94,27

Çalışmanın son aşamasında tez çalışması kapsamında geliştirilen ESA modelinin etkinliğini test etmek için CXR veri setleri kullanılmıştır [22,35]. Bu çalışma için önerilen 21 katmanlı yeni modelin Pnömoni CXR veri setindeki performans değerlendirmeleri Tablo 5.11’de yer almaktadır. Tablo 5.11’de 21 katmanlı yeni ESA modelinin CXR veri setindeki doğruluk performansı %94,50’dir.

Tez çalışması kapsamında gerçekleştirilen derin öğrenme yaklaşımlarından elde edilen sonuçlar arasında Tablo 5.12’de verildiği gibi bir karşılaştırma daha yapılmıştır. Karşılaştırmalar, her bir derin öğrenme yaklaşımından elde edilen en yüksek doğruluk puanına göre temsil edilmiştir. Tablo 5.12’de görüldüğü gibi, ince ayarlı AlexNet modeli %98,50 doğruluk ile en yüksek doğruluk skorunu gerçekleştirmiştir. Derin öznetelik çıkarma ve transfer öğrenme yöntemleri önceden eğitilmiş derin ESA modelleri kullanıldığından, uçtan uca eğitime göre daha yüksek doğruluk puanları üretmiştir.

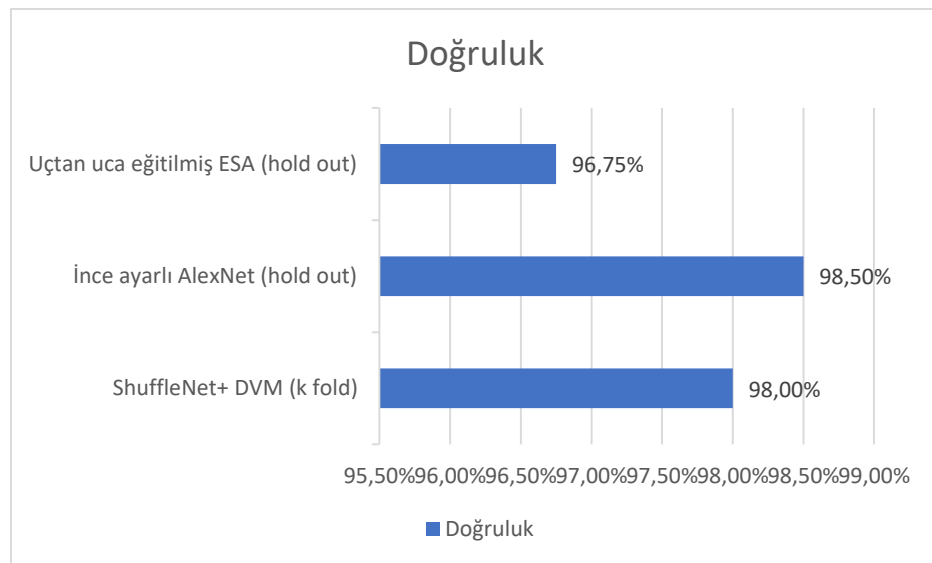
Tablo 5.12. Uygulanan yöntemlerin performans karşılaştırması

Yöntem	Doğruluk
ShuffleNet+ DVM (k fold)	%98,00
İnce ayarlı AlexNet (hold out)	%98,50
Uçtan uca eğitilmiş ESA (hold out)	%96,75

Tablo 5.13. Önerilen yöntemin literatürdeki çalışmalarla performans karşılaştırması

Yazarlar	Önerilen yöntem	Karşılaştırılan teknikler	Veri seti	Performans metrikleri (%)
Ayan vd. [23]	Çeşitli makine öğrenimi algoritmaları	Klasik makine öğrenimi algoritmaları İnce ayarlı Xception	[22]	Kesinlik:85,95 Duyarlılık:89,74 F1 skor:89,72 Doğruluk:89,74
Demir vd. [24]	Transfer öğrenme tabanlı ESA	AlexNet	[22]	Kesinlik:98,21 Duyarlılık:95,99 F1 skor:97,08 Doğruluk:96,31
Bakır vd. [25]	Derin öğrenme teknikleri	YSA ResNet-YSA İnce ayarlı ResNet MobileNet-YSA	[22]	Kesinlik:94,50 Duyarlılık:95,50 F1 skor:95,00 Doğruluk:95,67
Ouledroun vd. [26]	Derin öznetelikler	ESA-DAA ESA-DVM ESA-KNN	[22]	Kesinlik:97,05 Duyarlılık:95,20 F1 skor:96,12 Doğruluk:97,86
Aljawarneh vd. [27]	Geliştirilmiş ESA Modeli	Geliştirilmiş ESA ResNet İnce ayarlı ResNet VGG19	[22]	Kesinlik:90,0 Duyarlılık:96,0 F1 skor:92,90 Doğruluk:92,40
Kanawade vd. [28]	Derin öğrenme yaklaşımı	Xception VGG16 DenseNet201 Hibrit yaklaşım	[22]	Kesinlik:94,63 Duyarlılık:93,37 F1 skor:93,93 Doğruluk:94,39
Iparraguirre-Villanueva vd. [29]	Transfer öğrenme tabanlı ESA	VGG16 VGG19 ResNet50 InceptionV3	[22]	Kesinlik:- Duyarlılık:93,7 F1 skor:82,0 Doğruluk:72,9
(Bu tez çalışması kapsamında yayınlanmıştır) [161]	Derin öğrenme teknikleri	VGG16, VGG19, ResNet50, DenseNet201, AlexNet, NasNetMobile, MobileNetV2, DarkNet53, SqueezeNet, ShuffleNet, 21 katmanlı ESA	Özel veri seti	İnce ayarlı AlexNet Kesinlik:97,00 Duyarlılık:100,00 F1 skor:98,48 Doğruluk:98,50

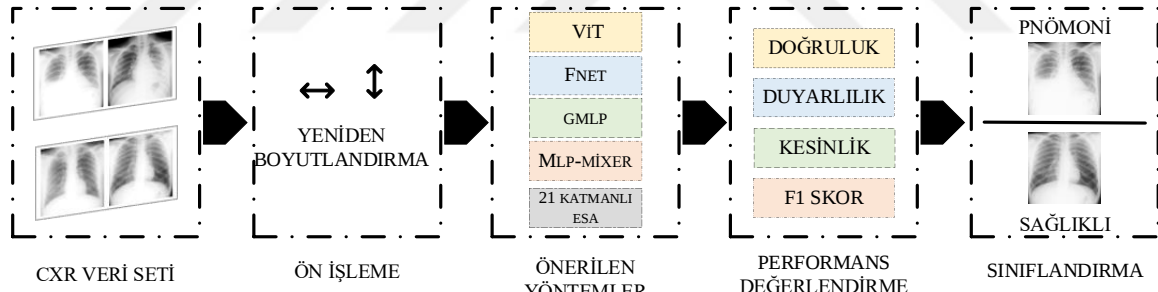
Bu çalışmanın temel amacı, ESA tabanlı derin öğrenme yaklaşımlarının pnömoni tespiti için oldukça etkili olduğunu göstermektir. Önceki çalışmalara benzer şekilde, bizim çalışmamız da derin öğrenmeye dayanmaktadır. Ancak metodolojik katkısı nedeniyle farklılık göstermektedir. Daha önce derin öğrenme kullanılarak yapılan farklı yöntemler öneren çalışmalar Tablo 5.13’de ilk çalışmada kullanılan yöntemlerle doğrulukları açısından karşılaştırılmıştır. Önerilen yöntemlerden olan ince ayarlı önceden eğitilmiş AlexNet ağının oluşturulan CXR veri kümesinin %80 eğitim %20 test ayrımındaki sınıflandırma performans metrikleri kesinlik haricinde diğer çalışmalardan daha iyi performans göstermiştir. İnce ayarlı AlexNet modelinin sınıflandırma metrikleri doğruluk, kesinlik, duyarlılık ve F1 skor açısından sırasıyla %98,50, %97,00, %100,00, %98,48’dir. [24]’de araştırmacılar CXR görüntülerinde pnömoni tespiti için transfer öğrenme yöntemini uygularken, [26]’da ise yazarlar CXR görüntülerinden çıkardıkları derin öznitelikleri sınıflandırmak için DAA, DVM ve KNN gibi çeşitli sınıflandırıcılar kullanmışlardır. [24] ve [26] kesinlik metrikleri sırasıyla %98,21 ve %97,05’dir. Buna göre, önerilen yöntemin doğruluk açısından önceki çalışmalarla karşılaştırılabilir olduğu görülmektedir. Tablo 5.13’de de görüleceği üzere bugüne kadar pnömoni tespitinde ESA mimarisini içeren birçok çalışma yapılmıştır. Bu mimarilerin en büyük avantajı uçtan uca öğrenme yapısı içermeleri, yani elle hazırlanmış bir özellik çıkarma adımının olmamasıdır. Transfer öğrenme yöntemi, büyük ölçekli veri kümelerinden öğrenilen önceden eğitilmiş modellerden elde edilen bilgilerden yararlanabildiği, performansı artırdığı, minimum eğitim sağladığı ve küçük veri kümelerinin dezavantajının üstesinden gelebildiği için pnömoni tespiti de dahil tıbbi görüntü sınıflandırması için uygun bir teknik olduğunu kanıtlamıştır. Tez çalışması kapsamında uygulanan her bir derin öğrenme yöntemi içinde en yüksek doğruluk performansı gösteren yaklaşımın karşılaştırmasının grafiksel analizi Şekil 5.2’de yer almaktadır.



Şekil 5.2. Uygulanan yöntemlerin en yüksek doğruluk puanı açısından karşılaştırılmasının grafiksel analizi

5.2. Göğüs röntgen görüntülerinde pnömoni tespiti için transformer modellerinin incelenmesi ve ESA mimarisi ile karşılaştırılması

CXR görüntülerinden pnömoni tespiti için gerçekleştirilen ikinci deneysel çalışmalarda tez çalışması kapsamında geliştirilen 21 katmanlı ESA modeli ile çeşitli dönüştürücü tabanlı yaklaşımların sınıflandırma performansları karşılaştırılmıştır. CXR görüntüleri kullanarak pnömoni tespiti için kullanılan dört görüntü dönüştürücü mimarileri, ViT, FNet, gMLP ve MLP-mixer'dir. Çalışmanın ön işleme aşamasında veri集中的 CXR görüntüleri farklı boyutlarda olduğu için dönüştürücü tabanlı modellerin minimum giriş boyutlarını karşılayacak şekilde 512×512 olarak yeniden boyutlandırılmıştır. Daha sonra her görüntü 32×32 piksel boyutunda toplam 256 parçaya bölünmüştür. Bahsedilen yöntemlerin eğitiminde ADAM optimizasyonu dikkate alınmıştır. Ayrıca, veri setini eğitim ve test verilerine bölmek için ayırma çapraz doğrulaması kullanılmıştır; burada bölme oranı eğitim için %90 ve test için %10'dır. CXR görüntüleri ayrıca 21 katmanlı ESA modeli için 224×224 olarak yeniden boyutlandırılmıştır. Geliştirilen 21 katmanlı ESA modeli için öğrenme oranı 0.0001, yığın boyutu 10 ve devir sayısı 8'dir. Geliştirilen 21 katmanlı ESA yönteminin eğitiminde momentumlu stokastik gradyan iniş (SGDM) optimizasyon algoritması dikkate alınmıştır. Şekil 5.3'de pnömoni tespiti için önerilen yöntemin akış diyagramı yer almaktadır.



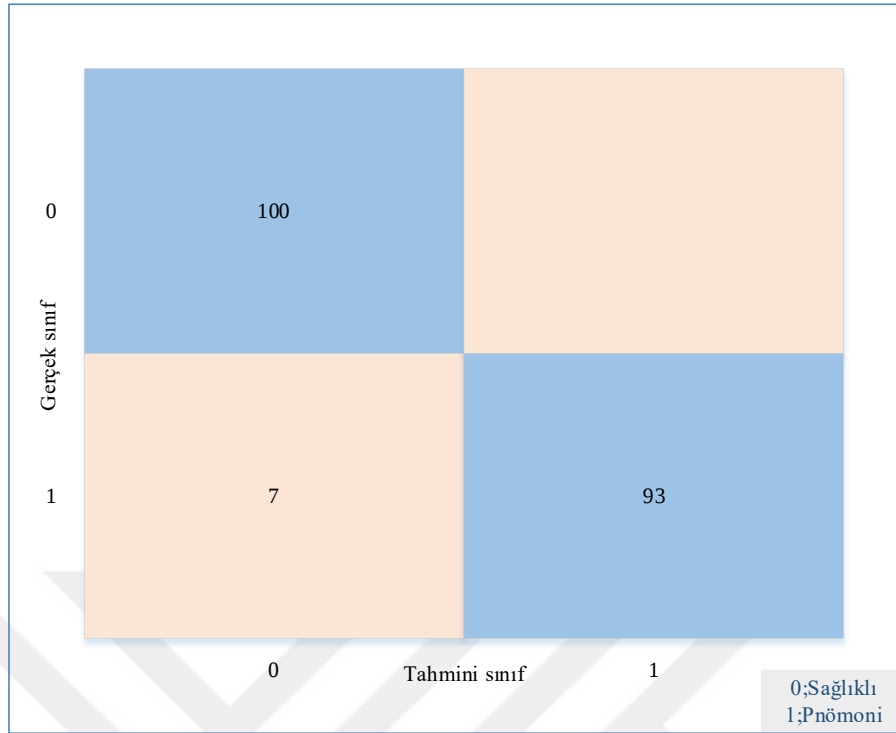
Şekil 5.3. Pnömoni ve sağlıklı CXR görüntülerinin dönüştürücü tabanlı modeller ve 21 katmanlı ESA ile sınıflandırılması

Dönüştürücü tabanlı modellerin parametreleri, her kod çalıştırılırken buluşsal olarak atanmıştır. ViT modeli için öğrenme oranı 0.001, ağırlık azalması 0.0001 ve yığın boyutu, eğitim tur sayısı, yansıtma boyutu, başlık sayısı ve dönüştürücü katman sayısı sırasıyla 64, 400, 64, 4 ve 8'dir. FNet, MLP-mixer ve gMLP modelleri için ağırlık azalması 0,0001 olarak ayarlanmıştır, yığın boyutu, eğitim tur sayısı, gömme boyutu, MLP- mixer katman sayısı ve bırakma oranı sırasıyla 64, 400, 256, 4 ve 0,2 olarak ayarlanmıştır. Öğrenme oranı, MLP mixer için 0,005, FNet için 0,001 ve gMLP için 0,003 olarak ayarlanmıştır.

Tablo 5.14. Kullanılan modeller için elde edilen sınıflandırma performanslarının özeti

Model	F1 skor (%)	Duyarlılık (%)	Kesinlik (%)	Doğruluk (%)
ViT	96,52	95,10	97,98	96,43
FNet	96,52	95,10	97,98	96,43
MLP-mixer	95,57	95,10	96,04	95,41
gMLP	94,42	91,18	97,89	94,39
21 katmanlı ESA	96,62	93,46	100,00	96,50

Tablo 5.14’de, her bir yöntem için elde edilen sınıflandırmanın değerlendirme ölçütlerini vermektedir. Tablo 5.14’deki satırlar modelin türünü, sütunlar ise elde edilen sınıflandırma performans parametrelerini göstermektedir. Tablo 5.14’te görüldüğü gibi, ViT ve FNet modelleri aynı doğruluk, kesinlik, duyarlılık ve F1 puanları değerleri üretmiştir. Üretilen değerlendirme puanları doğruluk, kesinlik, duyarlılık ve F1 puanları için sırasıyla %96,43, %97,98, %95,10 ve %96,52’dir. Ayrıca, MLP-mixer sırasıyla %95,41 doğruluk puanı, %96,04 kesinlik değeri, %95,10 duyarlılık değeri ve %95,57 F1 puanı kaydetmiştir. Son olarak gMLP yaklaşımı ile %94,39 doğruluk puanı, %97,89 kesinlik, %91,18 duyarlılık ve %94,42 F1-skor değerleri elde edilmiştir. Hesaplanan değerlendirme metrikleri incelendiğinde, önerilen dönüştürücü tabanlı yaklaşımlar arasında ViT ve FNet başarılarının en yüksek olduğu görülmüştür. MLP-mixer modeli, en iyi ikinci değerlendirme puanlarını (F1 puanı, duyarlılık ve doğruluk) üretmiştir. gMLP, hesaplanan kesinliğin %97,89 olduğu en iyi ikinci kesinlik skorunu elde etmiştir. gMLP kesinlik metriği dışındaki diğer performans metriklerinde en düşük başarıyı sergilemiştir. Bununla birlikte, önerilen ESA modeli, tüm dönüştürücü tabanlı modellerden daha iyi performans göstermiş ve tez çalışması kapsamında oluşturulan veri setinde ayırma çapraz doğrulama yöntemiyle %96,50 doğruluk, %100,00 kesinlik, %93,46 duyarlılık ve %96,62 F1 puanları kaydetmiştir.



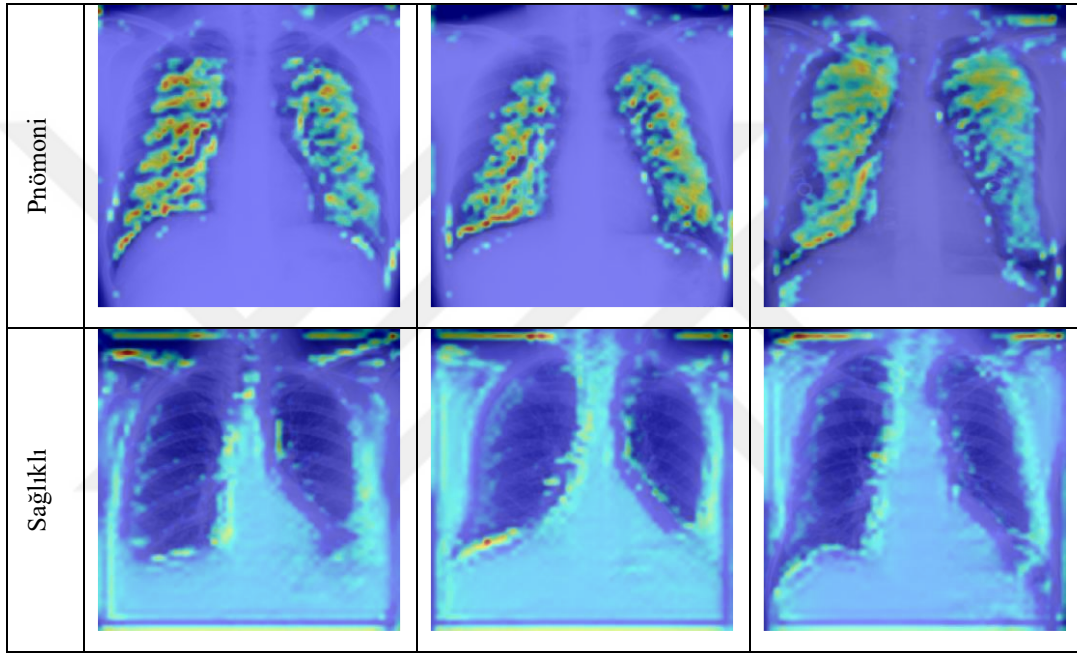
Şekil 5.4. 21 katmanlı geliştirilen ESA modeli kullanılarak elde edilen karışıklık matrisi

Geliştirilen ESA modeli tarafından elde edilen karışıklık matrisi Şekil 5.4’de gösterilmektedir. Karışıklık matrisinin satırları gerçek sınıfların örnek sayısını gösterirken, sütunlar tahmin edilen sınıfların örnek sayısını göstermektedir. Şekil 5.4’den de görülebileceği gibi, sağlıklı sınıf %100 doğru sınıflandırılmış ve 7 pnömoni örneği yanlış sınıflandırılmıştır. 21 katmanlı ESA modeli kullanılarak elde edilen performansın özeti Tablo 5.15’de yer almaktadır. Tablo 5.15’de belirtildiği gibi, geliştirilen 21 katmanlı ESA modeli, özel veri tabanı için hold-out doğrulama stratejisi ile %96,50 doğruluk puanı, %100,00 kesinlik, %93,46 duyarlılık ve %96,62 F1 puanı değerleri üretmiştir. Buna ek olarak, 21 katmanlı ESA modeli, on kat çapraz doğrulama stratejisi ile özel bir veri tabanı kullanarak sırasıyla %93,33, %92,15, %91,00 ve %92,25 kesinlik, F1 puanı, duyarlılık ve doğruluk değerleri vermiştir.

Tablo 5.15. Özel veri tabanı için ESA modeli kullanılarak elde edilen performansın özeti

Model	Veri Bölme	Doğruluk	Kesinlik	Duyarlılık	F1 skor
21 katmanlı ESA	%90 eğitim %10 test	96,50	100,00	93,46	96,62
21 katmanlı ESA	10 kat çapraz doğrulama	92,25	93,33	91,00	92,15

Şekil 5.5, 21 katmanlı ESA için Grad-CAM yöntemi kullanılarak sağlıklı ve pnömoni sınıfları için elde edilen ısı haritalarını göstermektedir. Grad-CAM tekniği, Zhou ve diğerleri tarafından geliştirilen ve yalnızca tam bağlantılı katmanları olmayan ESA modelleri için uygun olan orijinal Sınıf Aktivasyon Haritalama (CAM) tekniğinin geliştirilmiş bir versiyonudur [162]. Grad-CAM, girdi değerlerinin model sınıflandırmasını nasıl etkilediğini anlamak için gerekli olan her model katmanının görüntülenmesini ve her özellik haritası katmanının incelenmesini sağlar [163]. Grad-CAM, bir görüntünün belirli bir tahminde bulunmada en önemli olan bölgelerini vurgulayarak ESA modellerinin karar verme sürecine ilişkin önemli bilgiler sağlayabilir [164].



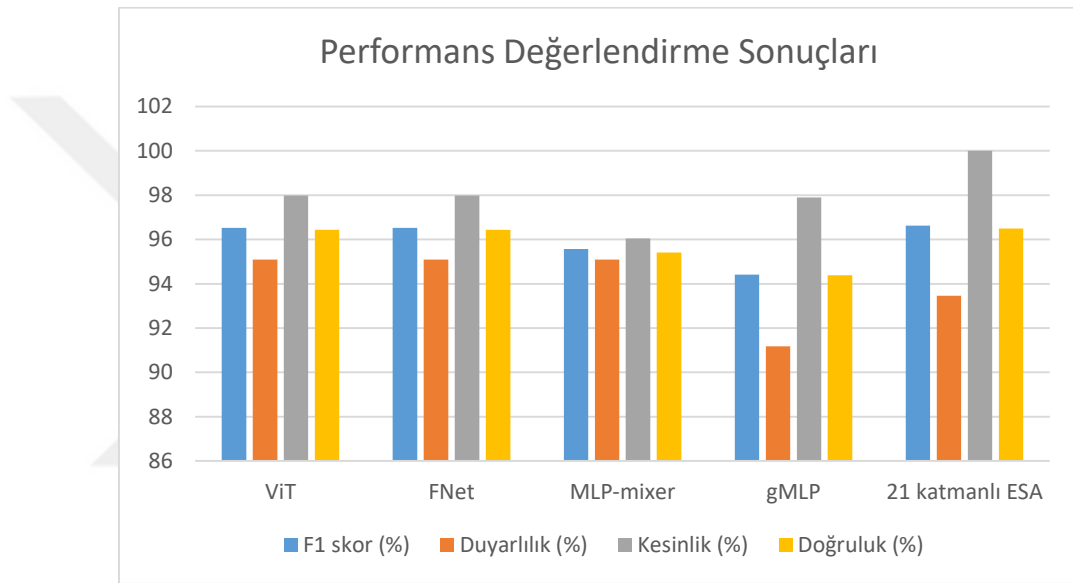
Şekil 5.5. 21 katmanlı ESA modeli için Grad-CAM tekniği kullanılarak sağlıklı ve pnömoni sınıfları için elde edilen ısı haritaları

Şekil 5.5’de görüldüğü gibi, birinci sıradaki resimler pnömonili bireylere ait CXR görüntüleri, ikinci sıradaki resimler ise sağlıklı deneklerin CXR görüntüleridir. Şekil 5.5’de gösterildiği gibi, pnömoni tespiti için 21 katmanlı ESA modeli pnömoni sınıfında, çoğunlukla sol akciğer blobunu dikkate almıştır ve sağlıklı sınıfında akciğer bloblarına odaklanmamıştır. Bu teknik, klinisyenler için giriş görüntüsünün sorunlu bölgesini belirlemeye yardımcı olur ve sağlıklı bireyler ile pnömoniden etkilenenleri ayırt etmede kolaylık sağlar.

Tablo 5.16. CXR görüntüleri üzerinde 21 katmanlı model kullanılarak elde edilen performans matrislerinin özeti

ESA (10 kat çapraz doğrulama)	Doğruluk (%)	Kesinlik (%)	Duyarlılık (%)	F1 skor (%)
21 katman	94,29	94,46	94,10	94,27

On kat çapraz doğrulama stratejisiyle CXR görüntüleri üzerinde 21 katmanlı model kullanılarak elde edilen performans değerlendirmesi Tablo 5.16’da verilmiştir [22,35]. %94,46 kesinlik değeri, %94,29 doğruluk, %94,27 F1 puanı ve %94,10 duyarlılık oranları elde edilmiştir.



Şekil 5.6. CXR görüntülerinden pnömoni tespiti için önerilen yöntemin performans değerlendirme sonuçlarının grafiksel analizi

Tablo 5.14’deki performans metriklerini daha net yorumlayabilmek için Şekil 5.6’daki grafik Tablo 5.14’deki değerlerden oluşturulmuştur. Şekil 5.6’da CXR görüntülerinden pnömoni tespiti için önerilen ikinci çalışmanın performans değerlendirme sonuçlarının grafiksel analizi yer almaktadır. Bu çalışma, CXR görüntülerinden pnömoniyi tespit etmek amacıyla yeni bir 21 katmanlı ESA modeli önermiş ve performansını çeşitli dönüştürücü tabanlı yaklaşımlarla karşılaştırmıştır. Şekil 5.6’da görüldüğü gibi, önerilen model doğruluk ve F1 skor açısından ViT ve FNet’e yakın bir performans göstermiştir. Bu çalışmada, yeni model özel ve genel veri tabanları kullanılarak geliştirilmiştir. Elde edilen sonuçlar, önerilen ESA tabanlı modelin, bu çalışma için toplanan CXR veri kümesini kullanarak pnömoniyi tespit etmede etkili ve doğru olduğunu göstermiştir. Tablo 5.15 ve tablo 5.16’da gösterildiği gibi, geliştirilen modelin sınıflandırma doğruluğu, pnömoni tespitinde özel ve genel veri tabanları için sırasıyla %96,50 ve %94,29’dur.

Tablo 5.17. Pnömoni tespiti için önerilen yaklaşımın CXR görüntüleri kullanılarak geliştirilen en son tekniklerle karşılaştırılması

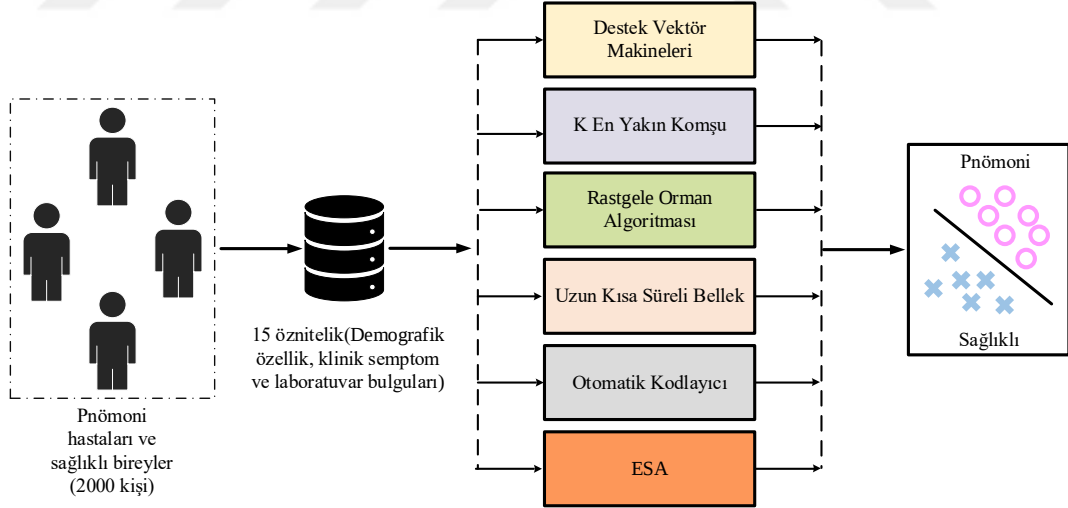
Yazarlar	Önerilen yöntem	Karşılaştırılan teknikler	Veri seti	Performans metrikleri (%)
Singh vd. [30]	Derin Dikkat Ağı	ResNet, Dikkat tabanlı ResNet	[22]	Kesinlik:- Duyarlılık:- F1 skor:92,0 Doğruluk:95,47
Tyagi vd. [31]	ViT ve diğer tekniklerle karşılaştırılması	ESA, VGG16, ViT	[22]	Kesinlik:- Duyarlılık:- F1 skor:- Doğruluk:96,45
Mabrouk vd. [34]	Derin ESA Topluluğu	DenseNet169, MobileNetV2, ViT	[22]	Kesinlik:93,96 Duyarlılık:92,99 F1 skor:93,43 Doğruluk:93,91
Ukwuoma vd. [36]	Hibrit Topluluk Dönüştürücü Kodlayıcısı	Hibrit A (DenseNet201, VGG16, GoogleNet) Hibrit B(DenseNet201, InceptionResNetV2, Xception)	Kermany [22] Chest X-ray [35]	Kesinlik:99,21 Duyarlılık: 99,21 F1 skor: 99,21 Doğruluk: 99,21
Cha vd. [37]	Dikkat Tabanlı Transfer Öğrenme Çerçevesi	ResNet152 ResNet18 DenseNet	[22]	Kesinlik:96,23 Duyarlılık:98,46 F1 skor:97,3 Doğruluk:96,63
Jiang vd. [38]	Çok Anlamlı Seviye Yama Birleştirme Görüntü Dönüştürücüsü	Temel (ResNet50), ViT, ViT + yama birleştirme	[22]	Kesinlik:91,82 Duyarlılık:89,36 F1 skor:90,34 Doğruluk:91,19
Ma vd. [39]	Dönüştürücü temel ağı	Swin dönüştürücü	[22] [35]	Kesinlik: - Duyarlılık: - F1 skor: - Doğruluk:97,2
Bu çalışma	Derin öğrenme teknikleri	21 katmanlı ESA ViT gMLP FNet MLP-mixer	Özel veri seti	21 katmanlı ESA Doğruluk:96,50 Kesinlik: 100,00 Duyarlılık: 93,46 F1 skor:96,62
Bu çalışma	Derin öğrenme teknikleri	21 katmanlı ESA	Kamu veri setleri [22,35]	Doğruluk:94,29 Kesinlik: 94,46 Duyarlılık: 94,10 F1 skor:94,27

Tablo 5.17, kamuya açık CXR veri kümelerini kullanarak pnömoniyi tespit etmek için yayınlanan önceki çalışmaları listelemektedir. Çalışmaların çoğunun pnömoni tespiti için farklı ESA teknikleri ve dönüştürücü tabanlı topluluk ESA çerçeveleri kullandığı belirtilebilir. Pnömoni tespiti için 5856 CXR görüntüsü içeren bir veri kümesi [22] literatürde ikili sınıflandırma görevlerinde yaygın olarak kullanılmıştır. Veri tabanında üç klasör bulunmaktadır: eğitim, test ve

doğrulama. Üç klasörün her biri normal ve pnömoni alt klasörlerini içerir. Veri kümesi, 3875'i eğitim, 390'ı test ve 8'i doğrulama için olmak üzere 4273 pnömoni görüntüsünden oluşmaktadır. Buna ek olarak, veri kümesi toplam 1583 sağlıklı görüntü içermektedir. Tablo 5.17'de görüldüğü gibi, ikili sınıflandırmada en iyi doğruluk puanı %99,21 ile Ukwuoma ve arkadaşları [36] tarafından hibrit dönüştürücü kodlayıcı tabanlı bir derin öğrenme modeli kullanılarak elde edilmiştir. [39]'deki yazarlar Swin dönüştürücü modeli kullanarak %97,20 doğruluk puanı elde ederken, Tyagi ve diğerleri [31] aynı CXR veri tabanı üzerinde ViT ile diğer ESA'ların performansını karşılaştırdığı çalışmada %96,45 doğruluk elde etmiştir. Çalışmamız, CXR görüntülerini doğru bir şekilde kullanarak pnömoni tespiti için özel ve genel veri tabanlarını kullanan ilk çalışmadır. Bu çalışmadaki veri sayısının azlığı ve deneysel çalışmalar sırasında dönüştürücü tabanlı modellerin temel mimari olarak kullanıldığı göz önünde bulundığında performanslarının nispeten düşük olması makul kabul edilmektedir.

5.3. Sayısal Verilerden Derin Öğrenme Yaklaşımları ile Pnömoni Tespiti

Retrospektif olarak yürütülen bu çalışmada Dicle Üniversitesi Tıp Fakültesi Göğüs Hastalıkları ve Tüberküloz kliniği ile yoğun bakım ünitesine 2001 Ocak- 2021 Ocak ayları arasında başvuran hastaların verileri HBYS veri tabanından sağlanmıştır. Sayısal tıbbi veriler hasta dosyasından elde edilmiştir ve sayısal verilere dayalı bir veri seti oluşturulmuştur. Demografik veriler olarak, yaş ve cinsiyet seçilmiştir. Hastalık şikâyeti olarak, yan ağrısı, nefes darlığı ile öksürük verileri seçilmiştir. Laboratuvar parametreleri olarak temel biyokimyasal parametrelerden CRP (mg / dL) ve Albümin (g / dL) çalışılmıştır. Tam kan sayımı parametrelerinden lökosit sayısı (WBC, $10e^3 \mu L$), nötrofil sayısı (NEU, $10e^3 \mu L$), lenfosit sayısı (LYM, $10e^3 \mu L$), trombosit sayısı (PLT, $10e^3 \mu L$), çalışılmıştır. Bu laboratuvar verilerinden hareketle pnömoni ve sağlıklı hastaların ayırıcı tanısında NLO, TLO ve CAO verileri kullanılmıştır. Sınıflandırma aşamasında kullanılan sınıflandırıcılar KNN, DVM, RO, otomatik kodlayıcı ağı, UKSB ve uçtan uca eğitilmiş yeni bir ESA'dan oluşturulmuştur. Verilerin %80'i eğitim, %20'si test için kullanılmıştır. Modellerin performansları doğruluk, duyarlılık, kesinlik ve F1 skor sınıflandırma metrikleri ile değerlendirilmiştir. Pnömoni hastası ve sağlıklı bireylerin demografik, klinik ve laboratuvar bulgularının dağılımı Tablo 5.18'de verilmiştir. Sayısal veriler ile pnömoni tespiti için önerilen yöntemin akış diyagramı Şekil 5.7'de gösterilmektedir.



Şekil 5.7. Sayısal veriler ile pnömoni tespiti için önerilen yöntemin akış diyagramı

Tablo 5.18. Pnömoni hastası ve sağlıklı bireylerin demografik, klinik ve laboratuvar bulguları

Parametreler	Hasta (N=1000) Mean±SD %	Sağlıklı (N=1000) Mean±SD %	Normal değer	P
CRP	5,73±7,27	0,31±0,48	0-0,5 mg / dL	< 0,001
Albümin	3,04±0,67	4,31±0,40	3,5-5,2 g / dL	< 0,001
WBC (Lökosit)	9,45±4,75	7,88±2,02	3,7-10,1 10e ³ µL	< 0,001
NEU (Nötrofil)	6,38±4,37	4,58±1,73	1,63-6,96 10e ³ µL	< 0,001
PLT (Trombosit)	304,32±130,89	261,42±68,53	155-366 10e ³ µL	< 0,001
LYM (Lenfosit)	1,98±1,32	2,42±0,68	1,09-2,99 10e ³ µL	< 0,001
CAO	2,26±3,14	0,07±0,13		< 0,001
NLO	4,64±5,23	2,05±1,13		< 0,001
TLO	220,72±263,29	115,83±42,15		< 0,001
Yaş	56,53±19,75	34,77±10,71		0,014
Cinsiyet	Kadın Erkek	440 (%44) 560 (%56)	237(%23,7) 763(%76,3)	0,002
Dispne	917(%91,7)	206(%20,6)		< 0,001
Öksürük	662(%66,2)	145(%14,5)		< 0,001
Yan ağrısı	666(%66,6)	29(%2,9)		< 0,001

Çalışma (hasta) ve kontrol (sağlıklı) gruplarının demografik, klinik ve laboratuvar bulgularının hastalık durumuna göre dağılımı Tablo 5.18’de verilmiştir. Pnömoni hastası ve sağlıklı bireyler için ortalama, standart sapma ve aralık belirtilmiştir. Çalışmaya PA AC grafisinde pnömoni tanısı alan n:1000 çalışma grubu ve PA AC grafisi normal olan n:1000 kontrol grubu olmak üzere toplam 2000 kişi dahil edilmiştir. Dicle Üniversitesi Tıp Fakültesi Göğüs Hastalıkları ve Tüberküloz kliniği ile yoğun bakım ünitesine 2001 Ocak- 2021 Ocak tarihleri arasında başvuran, tez çalışması kapsamında dahil edilme ve hariç tutulma kriterlerini karşılayan toplam 2000 kişinin HBYS kayıtları, arşivleri ve dosya kayıtları geriye dönük olarak incelenmiştir. Sayısal tıbbi verilerden pnömoni tespiti ile ilgili istatistiksel analize öncelikle toplanan verilerin normal dağılımına bakılarak başlanmıştır. Verilerin normal dağılıp dağılmadığını test etmek için Kolmogorov-Smirnov ve Shapiro-Wilk testleri uygulanmıştır. Sig. değeri 0.05 den küçük olduğu için veriler normal dağılım göstermediği kabul edilip veriler üzerinde nonparametrik testlerin uygulanmasına karar verilmiştir. Hasta özelliklerinin belirlenmesinde tanımlayıcı istatistiklerden

(ortalama, standart sapma), hasta gruplarının karşılaştırılmasında ise ki-kare ve Mann-Whitney U testlerinden yararlanılmıştır.

Bildirilen p değeri, pnömoni ve sağlıklı gruplarındaki nicel değişkenler arasındaki varyasyon için Mann-Whitney U testinin, nitel değişkenler için ki-kare testinin sonucuna karşılık gelir. Genellikle kategorik değişkenler olarak adlandırılan nitel değişkenler, farklı gruplar halinde kategorize edilebilen özellikleri barındırır. Ölçülebilir bir miktarı belirtmekten ziyade farklı kategorileri veya sınıfları temsil etmektedirler. Nitel veriler genellikle sayısal değildir ve bir veri kümesi veya denek grubunun doğasında bulunan nitelikleri veya özellikleri ifade eder [165]. Bu çalışmadaki nitel değişkenler cinsiyet, dispne, yan ağrısı, öksürük verileridir. Niceliksel değişkenler, sayısal bir değeri temsil eden, ölçülebilen veya sayılabilen miktarda özellikleri olan verilerdir [165]. Yaş verileri ve tez çalışması kapsamında toplanan biyobelirteçler nicel değişkenlerdir. Tablo 5.18'deki sonuçlar incelendiğinde seçilen sayısal parametreler açısından gruplar arasında istatistiksel olarak anlamlı fark saptanmıştır ($p \leq 0,05$). Oluşturulan veri setindeki biyobelirteçlere ait sonuçlar incelendiğinde CRP değerinin çalışma grubu için normal değer aralığından çok yüksek olduğu görülmektedir. CRP'nin yüksekliği vücutta pnömoniyeye ait inflamasyon ve enfeksiyon olabileceğini göstermekle birlikte spesifik bir belirteç değildir. Aynı şekilde Albümin değeri çalışma grubu için normal değer aralığından daha düşük çıkmıştır. Çalışma grubunda kontrol grubuna kıyasla CRP'nin yüksek ve Albümin 'in düşük çıkması literatürdeki çalışmalarla uyumludur ve istatistiksel olarak anlamlı farklılık bulunmaktadır [44-46,166]. Akut enfeksiyon sırasında azalan albümin düzeyleri de doğrudan altta yatan iltihaplanma sürecinden kaynaklanmaktadır ve bir dereceye kadar enfeksiyon/iltihap şiddetinin bir örneğini sağlayabilir. Akut inflamasyon sırasındaki bu özelliği nedeniyle albüminden sıklıkla negatif bir akut faz proteini olarak bahsedilir. Aynı zamanda, pnömoni dahil olmak üzere çeşitli akut enfeksiyon türlerinde tanınan ve geniş çapta incelenen pozitif bir akut faz proteini olarak CRP düzeyi de yükselir [166].

WBC, nötrofil, trombosit, lenfosit gibi kan parametreleri sistemik inflamasyon ve enfeksiyon hakkında bilgi sağlayan biyobelirteçler olarak çeşitli hastalıklarda kullanılmaktadırlar. Son zamanlarda, NLO ve TLO gibi çeşitli indeksler pnömoninin teşhisi ve şiddetinin değerlendirilmesinde benzer faydalar sağlamaktadır. NLO, daha şiddetli hastalık ile daha hafif hastalığı ayırt etmede yardımcıdır. NLO ucuz, basit, hızlı yanıt veren, yüksek duyarlılık ve düşük özgüllük ile stres ve inflamasyonun kolay erişilebilir bir parametresidir [47-49,167]. Bu çalışmada da literatürdeki çalışmalarla uyumlu olarak bu parametrelerde çalışma grubunda kontrol grubuna kıyasla istatistiksel olarak anlamlı fark bulunmuştur.

İkinci olarak aynı hastaların demografik özellik, klinik semptomlar ve laboratuvar bulguları analiz edilerek bu sayısal tıbbi veriler ile oluşturulan veri seti üzerinde makine öğrenmesi ve derin öğrenme yöntemleriyle sınıflandırma işlemi yapılmıştır. Sınıflandırma aşaması DVM, KNN, RO, Otomatik Kodlayıcı ve UKSB algoritmaları ile tamamlanmıştır. Ayrıca sayısal verileri

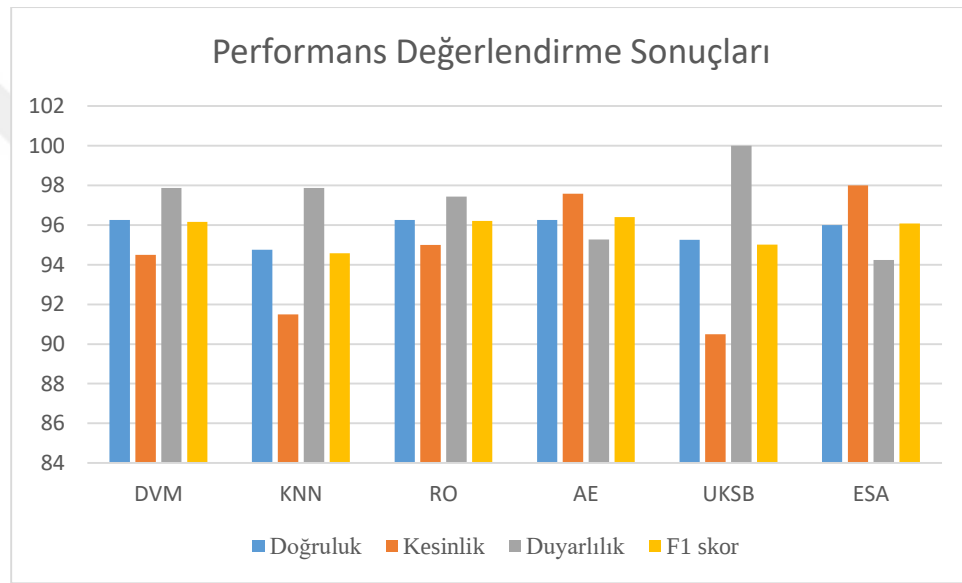
sınıflandırmak için de yeni bir ESA modeli uçtan uca eğitilmiştir. Sayısal özellikler veri kümesindeki yeni ESA mimarisi oluşturulurken özellik giriş katmanına sahip bir ağ tanımlanmış ve özelliklerin sayısı belirtilmiştir. Ayrıca, Z-skor normalizasyonu kullanılarak verileri normalleştirmek için giriş katmanı yapılandırılmıştır. Daha sonra, çıktı boyutu 50 olan tamamen bağlı bir katmanı ve ardından bir toplu normalleştirme katmanı ve bir ReLU katmanı eklenmiştir. Sınıflandırma için, sınıf sayısına karşılık gelen çıktı boyutuna sahip bir tam bağlantılı katman ve ardından bir softmax katmanı ve bir sınıflandırma katmanı eklenmiştir. Ağın eğitiminde ADAM optimizasyon algoritması seçilmiştir. Yığın boyutu 16, öğrenme oranı 0,01 ve devir sayısı 30'dur. Deneysel çalışmalarda kullanılan DVM hiper parametreleri; çekirdek fonksiyonu 'linear' olarak belirlenmiştir. KNN hiper parametrelerinde komşu sayısı '10' olarak belirlenmiştir. Uzaklık ölçü birimi 'euclidean' olarak seçilmiştir. RO hiper parametrelerinde eğitim metodu 'bag', toplulukta kullanılacak ağaç sayısı 30 seçilmiştir. Tablo 5.19'da modellerin sınıflandırma performans sonuçları yer almaktadır.

Tablo 5.19. Modellere ait sınıflandırma performans sonuçları

Model	Doğruluk	Kesinlik	Duyarlılık	F1 skor
DVM	96,25	94,50	97,86	96,15
KNN	94,75	91,50	97,86	94,57
RO	96,25	95,00	97,43	96,20
Otomatik Kodlayıcı	96,25	97,57	95,27	96,40
UKSB	95,25	90,50	100,0	95,01
ESA	96,00	98,00	94,23	96,08

Tablo 5.19'da 80-20% eğitim ve test oranları için DVM, KNN, RO, Otomatik Kodlayıcı, UKSB modellerinin ve uçtan uca eğitilmiş yeni ESA mimarisinin performansını değerlendirmek amacıyla karışıklık matrisi sonuçları verilmektedir. Performans ölçüm sonuçları, bu karışıklık matrislerindeki TP, TN, FP ve FN değerleri kullanılarak hesaplanmaktadır. Gösterildiği gibi neredeyse tüm ağlar, tüm ölçümlerde %90'nın üzerinde bir puanla tatmin edici sonuçlar elde edebiliyor; bu da sayısal parametreleri sınıflandırmak için seçilen mimarilerin yeni hastalıkların erken taranması için önemli bir araç olabileceğini gösteriyor. Üstelik tüm modeller yüksek kesinlik puanlarına ulaşıyor; bu da her sınıf için genellikle düşük sayıda yanlış pozitif bulunduğunu, yani pnömoni kaynaklarının tahminler sırasında karıştırılmadığını gösteriyor. Sınıflandırma algoritmalarına ait performans sonuçları incelendiğinde DVM, RO ve Otomatik Kodlayıcı %96,25 ile en yüksek doğruluk puanı üretirken; KNN %94,75 ile en düşük doğruluk puanı üretmiştir. Yeni ESA modeli %96,00 ve UKSB %95,25 doğruluk puanına ulaşmıştır. Kesinlik puanları karşılaştırıldığında %98,00 ile ESA en yüksek kesinlik değerine ulaşmıştır. Otomatik Kodlayıcı %97,57 en iyi ikinci kesinlik değerini üretmiştir. UKSB %90,50 ile en düşük kesinlik performansı göstermiştir. RO, DVM ve KNN algoritmalarının kesinlik değerleri sırasıyla %95,00, %94,50 ve

%91,50'dir. Duyarlılık performansları karşılaştırıldığında UKSB %100 ile en yüksek puanı üretirken, DVM ve KNN %97,86 duyarlılık göstermiştir. ESA %94,23 ile en düşük duyarlılığa sahiptir. Otomatik Kodlayıcı %95,27 duyarlılık performansı gösterirken; RO %97,43 puanı elde etmiştir. F1 skor açısından Otomatik Kodlayıcı %96,40 ile en yüksek performansı elde ederken; RO, DVM, UKSB ve KNN sırasıyla %96,20, %96,15, %95,01 ve %94,57 puanlarına ulaşmıştır. ESA %96,08 F1 skor performansı sağlamıştır. Şekil 5.8'de sayısal parametreler ile pnömoni tespiti için önerilen yöntemin performans metriklerinin grafiksel analizi yer almaktadır. Şekil 5.8'deki grafik, seçilen algoritmaların performans sonuçlarını daha iyi incelemek amacıyla Tablo 5.19'daki veriler kullanılarak oluşturulmuştur.



Şekil 5.8. Sayısal veriler ile pnömoni tespiti için önerilen yöntemin performans metriklerinin grafiksel analizi

Literatürdeki güncel çalışmalarda makine öğrenmesi ve derin öğrenme algoritmaları kullanılarak CXR görüntülerinden pnömoni tespiti mevcuttur [23-36]. Ancak aynı hastalarda laboratuvar verileri analiz edilmemiş ve iki veri türünden elde edilen tanımlar karşılaştırılmamıştır. Tablo 5.18'de oluşturulan sayısal parametre veri setinin istatistiksel analizinin sonuçları yer alırken; Tablo 5.19'da makine öğrenmesi ve derin öğrenme algoritmalarıyla sınıflandırma sonuçları yer almaktadır. Sonuçlar karşılaştırıldığında önerilen yöntemin daha düşük maliyetle ve daha kısa sürede doğru sonuçlar elde ettiği ve CXR görüntüleme için alternatif, tamamlayıcı bir test olduğu sonucuna varılmıştır. Tablo 5.20'de literatürdeki mevcut laboratuvar ve klinik bulgular kullanılarak yapılan pnömoni tespit çalışmaları yer almaktadır. Tablo 5.20'de önerilen yaklaşımın mevcut metodolojilerle karşılaştırılması yer almaktadır.

Tablo 5.20. Önerilen yaklaşımın mevcut metodolojilerle karşılaştırılması

Yazarlar	Seçilen öznitelikler	Karşılaştırılan teknikler	Performans değerlendirmeleri (%)
Effaf vd. [51]	Laboratuvar bulguları	LR, NB, DVM, ADT, KNN, RO, XGBoost, YSA	Kesinlik: 91,30 Duyarlılık: 96,00 F1 skor: 93,60 Doğruluk: 92,00
Zhang vd. [52]	Biyokimyasal ve klinik bulgular	GSHHO-FKNN	Doğruluk: 90,00 Duyarlılık: 82,64 Kesinlik: 93,00
Strokes vd. [53]	Denek popülasyon özellikleri, semptomlar ve laboratuvar bulguları	KA, DVM, LR	Doğruluk: 84,00 Duyarlılık: 80,00 Kesinlik: 73,00 F1 skor: 87,00
Önerilen Metot	Demografik özellik, klinik semptom ve laboratuvar bulguları	DVM, RO, KNN, Otomatik Kodlayıcı, UKSB, ESA	Doğruluk: 96,25 Duyarlılık: 95,27 Kesinlik: 97,57 F1 skor: 96,40

Düşük gelirli ortamlar gibi hastalık yükünün en yüksek olduğu bölgelerde radyografi genellikle mevcut değildir. Sonuç olarak, kliniklerde pnömoni dahil olmak üzere alt solunum yolu hastalıklarının tanı ve takibinde tam kan sayımı, temel biyokimyasal parametreler ve çeşitli enfeksiyon parametreleri kullanılmaktadır. Tanısal doğruluğu artırmak ve pnömoniye yönelik çeşitli tedavi stratejilerini geliştirmek için, invaziv olmayan ölçümlere dayalı bilgisayar tabanlı otomatik tespit modelleri önerilmiştir. Yapılan çalışmalarda hastalığın tanısı, şiddeti, yoğun bakım ihtiyacı gibi durumların değerlendirilmesinde temel biyokimya ve tam kan sayımı dahil çeşitli inflamasyon belirteçleri kullanılmıştır [51-53]. Enfeksiyon, enflamasyon ve travmaya yanıt olarak konakçıda meydana gelen spesifik olmayan dönem akut faz olarak bilinir. Günlük pratikte, akut faz yanıtları bakteriyel ve viral enfeksiyonların ayırt edilmesinde yaygın olarak kullanılmaktadır. Klinik uygulamalarda en sık kullanılan akut faz yanıtları lökosit sayısı (WBC), mutlak nötrofil sayısı ve CRP'dir [168]. Bu çalışmada da literatürdeki çalışmalara uyumlu olarak WBC, Nötrofil, Lenfosit, Trombosit, CRP, Albümin belirteçleri çalışılmıştır.

Effaf ve ark. her biri 45 özelliğe sahip 535 farklı hasta üzerinde pnömoniyi öngörmedeki etkinliğini değerlendirmek için 8 makine öğrenmesi modelini (LR, NB, DVM, AdaBoost KA,

KNN, RO, YSA ve XGBoost) karşılaştırmışlardır [51]. Bu 45 özellik içerisinde hastalara ait demografik bilgileri (yaş, cinsiyet ve komorbiditeler dahil), fiziksel parametreleri (taşikardi, trakeal sekresyon, plevral efüzyon, ortalama arteriyel basınç, kalp atım hızları, solunum hızları ve sistolik kan basıncı) ve hematolojik parametreleri (serum sodyum, serum potasyum, serum kreatin, hematokrit, WBC, trombosit, toplam bilirubin, hemoglobin CRP) yer almaktadır. RO ve XGBoost sırasıyla %92,00 ve %90,80 ile en yüksek doğruluk performansı gösteren ağlardır.

Zhang ve ark. 100 hastaya ait cinsiyet, yaş, hasta bölümünde yattığı gün sayısı ve hastalıkların prognozu, pıhtılaşma ve böbrek fonksiyon testi, WBC, PLT, Albümin g/L, CRP, mg/L) dahil olmak üzere inflamatuvar kan testi parametrelerini toplamışlardır [52].

Strokes ve ark. 4500 hastaya ait çeşitli klinik bilgileri toplamışlardır. Toplanan bilgiler arasında solunum yolu hastalıklarına özgü bir dizi semptom, laboratuvar test sonuçları ve hava kirliliğine maruz kalma veya yetersiz beslenme gibi çeşitli nüfus tanımlayıcı özellikleri yer almıştır. Laboratuvar test sonuçları arasında CRP (mg/ml), Lökositler, Nötrofiller, Lenfositler gibi biyobelirteçler yer almaktadır. Belirti ve semptomlar arasında öksürük, dispne, ateş, plevral ağrı, kas ağrısı, terleme, baş ağrısı gibi çeşitli ikili (evet ya da hayır şeklinde) özellikler yer akmaktadır. LR, DVM, KA arasında KA %84,00 en yüksek doğruluğa ulaşmıştır [53].

Bu çalışmada olduğu gibi literatür çalışmalarında da sayısal tıbbi verilerinden pnömoni tespiti yapılırken kullanılacak öznitelikler belirlenerek özel veri setleri oluşturulmuştur. Bu çalışmanın sonuçlarını diğer çalışmalarla karşılaştırdığımızda, bu çalışmayla elde edilen sonuçların mevcut diğer çalışmalarla karşılaştırılabilir olduğunu ve daha iyi performans gösterdiği tespit edilmiştir. Birçok çalışmada pnömonili hastalarına ait laboratuvar parametreleri araştırılmıştır. Ancak bu çalışmalar çoğunlukla pnömoninin prognozu ve şiddeti ile ilgilidir. Bu çalışmada, laboratuvar parametrelerinden pnömoni tespitine daha kısa sürede, daha ucuz ve daha doğru sonuçlar ile olanak sağlayacak bilgisayar destekli bir tanı sisteminin geliştirilmesi amaçlanmıştır.

6. SONUÇLAR

Pnömoni, bakteri, virüs ve mantar enfeksiyonlarının neden olduğu akciğer dokusunun iltihaplanmasıdır. Pnömoninin varlığının ve etkeninin tespit edilmesi, tedavi sürecini hızlandırmak ve hastanın hayatta kalma oranını artırmak için çok önemlidir. Pnömoninin doğru biyomedikal teşhisi, çeşitli teşhis araçlarının kullanılması ve çeşitli klinik özelliklerin değerlendirilmesini gerektirmektedir. Klinik semptomlar, kapsamlı bir fiziksel muayene ve göğüs röntgeni gibi görüntüleme çalışmaları uzun zamandır pnömoni teşhisinde aşamalı olarak kullanılmaktadır. Ancak bu teknikler genellikle öznel ve güvenilir sonuçlar vermeleri garanti değildir. Bilgisayar destekli teşhisteki ilerlemeyle birlikte teşhis doğruluğunda bir artış eğilimi görülebilir. Son yıllarda, derin öğrenme teknikleri kullanılarak sayısal tıbbi veriler ve görüntü verilerinin analizi artan bir popülerlik alanı haline gelmiştir. Derin öğrenme metodolojileri, pnömoni hastalığının türünü tanımlamak ve tanımak için kullanılırken, otomatik algılama, zamanı azaltmaya ve doğruluğu artırmaya yardımcı olur.

Bu tez çalışmasının amacı; derin öğrenme teknikleri kullanılarak pnömoni hastalığına ait hem sayısal veri hem de göğüs röntgen görüntülerinden hastalık ile ilgili özellikleri etkili bir şekilde çıkarıp hastalığın tespiti için otomatik pnömoni tespit sisteminin geliştirilmesi, sağlık hizmetlerinin kalitesinin daha düşük maliyetlerle ve daha hızlı yanıtla iyileştirilmesidir. Bu amaçla Dicle Üniversitesi Tıp Fakültesi Göğüs Hastalıkları ve Tüberküloz kliniği, yoğun bakım ünitesi ve göğüs polikliniğinden elde edilen pnömoni ve sağlıklı CXR görüntüleri ile aynı kişilere ait demografik özellik, klinik semptom ve laboratuvar bulgularından oluşan sayısal tıbbi verilerden 2 yeni veri kümesi oluşturulmuştur. 2000 denek üzerinde yapılan çalışmada, 1000 denek Poster Anterior (PA) akciğer grafisinde pnömoni tanısı almış, 1000 denek ise normal PA akciğer grafisine sahiptir. Sayısal tıbbi veriler hasta dosyasından elde edilmiştir. Demografik veriler olarak, yaş ve cinsiyet seçilmiştir. Hastalık şikâyeti olarak, yan ağrısı, nefes darlığı ile öksürük verileri seçilmiştir. Laboratuvar parametreleri olarak; lökosit sayısı (WBC, $10^3/\mu\text{L}$), nötrofil sayısı (NEU, $10^3/\mu\text{L}$), lenfosit sayısı (LYM, $10^3/\mu\text{L}$), trombosit sayısı (PLT, $10^3/\mu\text{L}$), CRP (mg / dL) ve Albümin (mg / dL) çalışılmıştır. Bu laboratuvar verilerinden hareketle pnömoni ve sağlıklı hastaların ayırıcı tanısında NLO, TLO ve CAO verileri kullanılmıştır. Sayısal tıbbi verilerin sınıflandırma aşamasında KNN, DVM, RO, otomatik kodlayıcı ve UKSB algoritmaları kullanılmıştır. Ayrıca sayısal parametre veri setinden pnömoni tespiti için yeni bir ESA modeli uçtan uca eğitilmiştir. Verilerin %80'i eğitim, %20'si test için kullanılmıştır. Sayısal tıbbi verilerin sınıflandırılmasında en yüksek doğruluk %96,25 olarak elde edilmiştir. Önerilen ESA modeli %96,00 doğruluk başarımı göstermiştir.

Aynı kişilere ait CXR görüntüleri de derin öğrenme tabanlı hibrit yaklaşımlarla sınıflandırılmıştır. CXR görüntülerinin sınıflandırılmasında iki farklı çalışma yapılmıştır. İlk çalışmada sınıflandırma aşamasında derin öğrenme tabanlı yaklaşımlar kullanılmıştır. CXR görüntülerinin sınıflandırılmasında kullanılan derin öğrenme yöntemleri, derin öznitelik çıkarımı, transfer öğrenme ve uçtan uca öğrenmedir. Derin öznitelik çıkarımı ile transfer öğrenme yöntemleri için önceden eğitilmiş 10 farklı ESA ağı kullanılmıştır. Uçtan uca öğrenme yöntemi için 21 katmanlı yeni bir ESA mimarisi önerilmiştir. Önerilen yöntemlerin performansını değerlendirmek amacıyla doğruluk, kesinlik, duyarlılık ve F1 skor metrikleri kullanılmıştır. CXR görüntüleri sınıflandırılırken veri seti hem k kat çapraz doğrulama hem de ayırarak çapraz doğrulama yöntemleri ile test edilmiştir. Verilerin %80'i eğitim, %20'si test için kullanılmıştır. Çapraz doğrulama için k değeri 5 seçilmiştir. Çalışmanın ön işleme aşaması olarak; kullanılan farklı algoritmaların görüntü girişleri farklı olduğu için CXR görüntüleri yeniden boyutlandırılmıştır. Önceden eğitilmiş ESA modellerini eşit şartlar altında karşılaştırmak için tüm modeller minimum girdi boyutlarını karşılayacak ortak boyut olan 256×256 yeniden boyutlandırılmıştır. Çıkarılan derin öznitelikler DVM, KNN ve RO algoritmaları ile sınıflandırılmışlardır. CXR görüntülerinin sınıflandırılmasında derin öğrenme yaklaşımlarının performans karşılaştırmaları sırasıyla şöyledir. Derin öznitelik çıkarımında 5 kat çapraz doğrulamalı ShuffleNet+DVM hibrit yaklaşımı %98,00 doğrulukla en yüksek performansı göstermiştir. Hold out yöntemiyle ayrılan veri setinde ince ayarlı AlexNet %98,50 ile en yüksek transfer öğrenme başarımı elde etmiştir. Hold out yöntemiyle ayrılan veri setinde 21 katmanlı ESA modeli %96,75 sınıflandırma performansı sağlamıştır. Önerilen 21 katmanlı ESA modelinin literatürde mevcut Kaggle veri setindeki başarımı %94,50'dur.

CXR görüntülerinin sınıflandırılmasında yapılan ikinci çalışmada tez çalışması kapsamında önerilen 21 katmanlı ESA modeli ile yakın zamanda geliştirilen dönüştürücü tabanlı modellerin sınıflandırma performansı karşılaştırılmıştır. Karşılaştırma için kullanılan ağlar ViT, gMLP, FNet ve MLP-mixer modelleridir. Verilerin %90'i eğitim, %10'si test için kullanılmıştır. Çalışmanın ön işleme aşamasında, veri kümesindeki CXR görüntüleri farklı boyutlarda olduğundan, dönüştürücü tabanlı modellerin minimum giriş boyutlarını karşılamak için 512×512 olarak yeniden boyutlandırılmıştır. Daha sonra her görüntü 32×32 piksel boyutunda toplam 256 yamaya bölünmüştür. CXR görüntüleri ayrıca 21 katmanlı ESA modeli için 224×224 olarak yeniden boyutlandırılmıştır. ViT, FNet, MLP-mixer ve gMLP modellerinin sınıflandırma performans doğrulukları sırasıyla %96,43, %96,43, %95,41 ve %94,39'dur. Geliştirilen 21 katmanlı yeni model, sağlıklı bireylerden pnömoniye doğru bir şekilde tespit etmede özel ve genel veri tabanları için sırasıyla %96,50 ve %94,29 doğruluk sağlamıştır. Önerilen model, pnömoni tespiti için yakın zamanda geliştirilen dönüştürücü tabanlı modellerden daha yüksek sınıflandırma performansı sağlamıştır.

Sayısal verilerin ve CXR görüntülerinin sınıflandırma sonuçları karşılaştırıldığında her iki yöntemle pnömoni tespiti yüksek doğruluk ile sonuçlanmaktadır. Bu tez çalışması kapsamında önerilen yöntemler, pnömoni tanısı için bir karar destek sistemi olarak kullanılabilir ve farklı bir bakış açısıyla alternatif bir çözüm olarak klinisyenlere yardımcı olabilir. Bu bulgular, gelişmiş ekipman ve yüksek eğitimli uzmanların az olduğu düşük gelirli ve kaynak sınırlı ülkelerde pnömoninin otomatik tespit edilmesine, doktorların hastaları için daha etkili bir tedavi planı geliştirmesine, aşırı antibiyotik kullanımının azaltılmasına ve iyileşme oranlarının artmasına yardımcı olacaktır.

Literatür çalışmaları incelendiğinde pnömoni tespitinin mevcut dezavantajlarından biri, yüksek açıklamalı veri setlerine ve hastanın tıbbi geçmişine sahip olmamasıdır. Bu tez çalışması kapsamında hem pnömonili hastalara hem de sağlıklı bireylere ait CXR görüntüleri toplanmış ve derin öğrenme yaklaşımları ile pnömoni tespit edilmeye çalışılmıştır. Aynı kişilere ait sayısal tıbbi veriler de toplanarak kapsamlı bir veri seti oluşturmak amaçlanmıştır ve sayısal tıbbi verilerin derin öğrenme algoritmaları ile pnömoni tespitindeki yüksek doğruluk oranları ve kayda değer sonuçlar elde ettiği gösterilmiştir.

ÖNERİLER

Çalışmanın sınırlılığı, bu çalışma için yalnızca iki veri kümesi kullanmış olmamızdır. Göğüs röntgen görüntüleri hasta bireylere ve sağlıklı kişilere aittir. Bu çalışmada pnömoninin erken, hafif ve şiddetli gibi farklı aşamalarına ve/veya bakteri ve virüs gibi farklı türlerine ait alt bir sınıflandırma yapılmamıştır. Literatürdeki veri kümelerine kıyasla daha küçük boyutludur ve sınıflandırma performansı açısından tek yönlüdür. Farklı pnömoni aşamalarına ve türlerine ait göğüs röntgeni görüntüleri eklenerek veri kümesi boyutu ve kapsamı genişletilebilir. Aynı zamanda diğer alt solunum yolu enfeksiyonları da eklenerek CXR görüntülerinin kalitesi bilgisayar destekli sistemlerle iyileştirilebilir. Daha fazla veri eklenerek ve renk, doku ve şekil gibi düşük seviyeli öznitelikleri çıkarmak için gelişmiş yöntemler kullanılarak performans daha da artırılabilir. Düşük seviyeli öznitelikler ile derin özniteliklerin sınıflandırma performansları ayrı ayrı ve/veya bir füzyon modeli oluşturularak karşılaştırılabilir. Bu çalışmanın kısıtlılıklarından biri de retrospektif bir çalışma olmasıdır. Gerçek zamanlı takip değerleri yerine pnömoni tespiti için öncül parametreler kullanılmıştır. Bu çalışmada CXR görüntülerini işlerken ön işleme olarak sadece yeniden boyutlandırma işlemi yapılmıştır. Literatürdeki çalışmalarda histogram eşitleme, gürültü giderme veya görüntü kümesinin boyutunu artırma gibi işlemler de yapılmıştır. Bu ve bunun gibi farklı ön işlemler denenerek derin öğrenme modellerinin sınıflandırma performansı artırılabilir. Ayrıca, önerilen derin ESA modellerinin performansı daha fazla sayıda katman ve parametre ile geliştirilebilir. Bu, klinisyenlerin göğüs röntgeni görüntülerinden akciğer hastalıklarını hastalığın daha erken bir aşamasında daha düşük prevalansla tanımasına olanak sağlayacaktır.

KAYNAKLAR

- [1] Torres, A., Serra-Batlles, J., Ferrer, A., Jiménez, P., Celis, R., Cobo, E., & Rodriguez-Roisin, R. (1991). Severe community-acquired pneumonia. *Am Rev Respir Dis*, 144(2), 312.
- [2] Meedeniya, D., Kumarasinghe, H., Kolonne, S., Fernando, C., De la Torre Díez, I., & Marques, G. (2022). Chest X-ray analysis empowered with deep learning: A systematic review. *Applied Soft Computing*, 109319.
- [3] WHO (2021) Pneumonia “Pneumonia in children” Available: <https://www.who.int/news-room/fact-sheets/detail/pneumonia> (Erişim: 23/06/2023)
- [4] Bhatt, H., & Shah, M. (2023). A Convolutional Neural Network ensemble model for Pneumonia Detection using chest X-ray images. *Healthcare Analytics*, 100176.
- [5] Yi, R., Tang, L., Tian, Y., Liu, J., & Wu, Z. (2023). Identification and classification of pneumonia disease using a deep learning-based intelligent computational framework. *Neural Computing and Applications*, 35(20), 14473-14486.
- [6] Sharma, S., & Guleria, K. (2023). A Deep Learning based model for the Detection of Pneumonia from Chest X-Ray Images using VGG-16 and Neural Networks. *Procedia Computer Science*, 218, 357-366.
- [7] Parveen, S., & Khan, K. B. (2020). Detection and classification of pneumonia in chest X-ray images by supervised learning. In *2020 IEEE 23rd International Multitopic Conference (INMIC)* (pp. 1-5). IEEE.
- [8] Lamia, A., & Fawaz, A. (2022). Detection of Pneumonia Infection by Using Deep Learning on a Mobile Platform. *Computational Intelligence and Neuroscience*, 2022.
- [9] Govindarajan A, Govindarajan A, Tanamala S, Chatteraj S, Reddy B, Agrawal R, Iyer D, Srivastava A, Kumar P, Putha P. (2022) Role of an Automated Deep Learning Algorithm for Reliable Screening of Abnormality in Chest Radiographs: A Prospective Multicenter Quality Improvement Study. *Diagnostics*. 2022; 12(11):2724. <https://doi.org/10.3390/diagnostics12112724>
- [10] Qin, C., Yao, D., Shi, Y., & Song, Z. (2018). Computer-aided detection in chest radiography based on artificial intelligence: a survey. *Biomedical engineering online*, 17(1), 1-23.
- [11] Ait Nasser, A., & Akhloufi, M. A. (2023). A Review of Recent Advances in Deep Learning Models for Chest Disease Detection Using Radiography. *Diagnostics*, 13(1), 159.
- [12] Kwon T, Lee SP, Kim D, Jang J, Lee M, Kang SU, Kim H, Oh K, On J, Kim YJ, Yun SJ, Jin KN, Kim EY, Kim KG. (2021) Diagnostic performance of artificial intelligence model for pneumonia from chest radiography. *PLoS One*. 2021 Apr 15;16(4):e0249399. doi: 10.1371/journal.pone.0249399. PMID: 33857181; PMCID: PMC8049482.
- [13] Elemraid MA, Muller M, Spencer, D. A., Rushton SP, Gorton R, Thomas MF, Eastham KM, Hampton F, Gennery AR, Clark JE; North East of England Paediatric Respiratory Infection Study Group. (2014). Accuracy of the interpretation of chest radiographs for the diagnosis of paediatric pneumonia. *PLoS One*. 2014 Aug 22;9(8):e106051. doi: 10.1371/journal.pone.0106051. PMID: 25148361; PMCID: PMC4141860.
- [14] Htun, T. P., Sun, Y., Chua, H. L., & Pang, J. (2019). Clinical features for diagnosis of pneumonia among adults in primary care setting: A systematic and meta-review. *Scientific reports*, 9(1), 7600.
- [15] Ahmadova, A., Huseynov, I., & Ibrahimov, Y. (2023). Improving Pneumonia Diagnosis with RadImageNet: A Deep Transfer Learning Approach. *Authorea Preprints*.

- [16] Arun Prakash, J., Asswin, C. R., Ravi, V., Sowmya, V., & Soman, K. P. (2023). Pediatric pneumonia diagnosis using stacked ensemble learning on multi-model deep CNN architectures. *Multimedia Tools and Applications*, 82(14), 21311-21351.
- [17] Chiwariro, R., & Julius, B. (2023). Comparative analysis of deep learning convolutional neural networks based on transfer learning for pneumonia detection. *Int. J. Res. Appl. Sci. Eng. Technol*, 11(1), 1161-1170.
- [18] Singh R, Kalra MK, Nitiwarangkul C, Patti JA, Homayounieh F, Padole A, Rao P, Putha P, Muse VV, Sharma A, Digumarthy SR. (2018) Deep learning in chest radiography: Detection of findings and presence of change. *PLoS One*. 2018 Oct 4;13(10):e0204155. doi: 10.1371/journal.pone.0204155. PMID: 30286097; PMCID: PMC6171827.
- [19] Rashed, B. M., & Popescu, N. (2023). Performance Investigation for Medical Image Evaluation and Diagnosis Using Machine-Learning and Deep-Learning Techniques. *Computation*, 11(3), 63. <https://doi.org/10.3390/computation11030063>
- [20] Van Ginneken, B., Romeny, B. T. H., & Viergever, M. A. (2001). Computer-aided diagnosis in chest radiography: a survey. *IEEE Transactions on medical imaging*, 20(12), 1228-1241. doi: 10.1109/42.974918. PMID: 11811823.
- [21] Li, Y., Zhang, Z., Dai, C., Dong, Q., & Badrigilan, S. (2020). Accuracy of deep learning for automated detection of pneumonia using chest X-Ray images: A systematic review and meta-analysis. *Computers in Biology and Medicine*, 123, 103898.
- [22] Kermany DS, Goldbaum M, Cai W, Valentim CCS, Liang H, Baxter SL, McKeown A, Yang G, Wu X, Yan F, Dong J, Prasadha MK, Pei J, Ting MYL, Zhu J, Li C, Hewett S, Dong J, Ziyar I, Shi A, Zhang R, Zheng L, Hou R, Shi W, Fu X, Duan Y, Huu VAN, Wen C, Zhang ED, Zhang CL, Li O, Wang X, Singer MA, Sun X, Xu J, Tafreshi A, Lewis MA, Xia H, Zhang K. (2018) Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell*. 2018 Feb 22;172(5):1122-1131.e9. doi: 10.1016/j.cell.2018.02.010. PMID: 29474911.
- [23] Ayan, E. (2022). Using a Convolutional Neural Network as Feature Extractor for Different Machine Learning Classifiers to Diagnose Pneumonia. *Sakarya University Journal of Computer and Information Sciences*, 5(1), 48-61.
- [24] Demir, Y., & Bingöl, Ö. (2023). Detection of pneumonia from pediatric chest X-ray images by transfer learning. *Sigma Journal of Engineering and Natural Sciences*, 41(6), 1264-1271.
- [25] Bakır, H., Oktay, S., & Tabaru, E. (2023). Detection Of Pneumonia From X-Ray Images Using Deep Learning Techniques. *Journal of Scientific Reports-A*, (052), 419-440.
- [26] Ouledroun, T., Ellahyani, A., & El Ansari, M. (2023). Automated Pneumonia Detection using deep features in chest X-ray images. In *2023 3rd International Conference on Intelligent Communication and Computational Techniques (ICCT)* (pp. 1-6). IEEE.
- [27] Aljawarneh, S. A., & Al-Quraan, R. (2023). Pneumonia Detection Using Enhanced Convolutional Neural Network Model on Chest X-Ray Images. *Big Data*.
- [28] Kanawade, B. R., Zaware, S. N., Nandre, J., Mahale, Y., & Dhake, K. (2023). A Deep Learning Approach for Pneumonia Detection from X ray Images. *International Journal of Intelligent Systems and Applications in Engineering*, 11(2), 262-266.
- [29] Iparraguirre-Villanueva, O., Guevara-Ponce, V., Roque Paredes, O., Sierra-Liñan, F., Zapata-Paulini, J., & Cabanillas-Carbonell, M. (2022). Convolutional Neural Networks with Transfer Learning for Pneumonia Detection. *International Journal of Advanced Computer Science and Applications*, 13(9), 544-551.

- [30] Singh, S., Rawat, S., Gupta, M., Tripathi, B., Alanzi, F., Majumdar, A., Khuwuthyakorn, P., Thinnukool, O. (2022) Deep Attention Network for Pneumonia Detection Using Chest X-ray Images. *Computers, Materials & Continua* 74(1):1673–1692.
- [31] Tyagi, K., Pathak, G., Nijhawan, R., & Mittal, A. (2021). Detecting Pneumonia using Vision Transformer and comparing with other techniques. In *2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA)* (pp. 12-16). IEEE.
- [32] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. Attention is all you need. In *Advances in neural information processing systems*, pp. 5998–6008, 2017.
- [33] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J. & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.
- [34] Mabrouk, A., Díaz Redondo, R. P., Dahou, A., Abd Elaziz, M., & Kayed, M. (2022). Pneumonia detection on chest X-ray images using ensemble of deep convolutional neural networks. *Applied Sciences*, 12(13), 6448.
- [35] Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM. ChestX-ray8: hospitalscale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: *Proc - 30th IEEE Conf Comput Vis Pattern Recognition, CVPR 2017*, vol. 2017-January; 2017. p. 3462–71. doi: 10.1109/CVPR.2017.369.
- [36] Ukwuoma, C. C., Qin, Z., Heyat, M. B. B., Akhtar, F., Bamisile, O., Muaad, A. Y., Addo, D., & Al-Antari, M. A. (2023). A hybrid explainable ensemble transformer encoder for pneumonia identification from chest X-ray images. *Journal of Advanced Research*, 48, 191-211.
- [37] Cha, S. M., Lee, S. S., & Ko, B. (2021). Attention-Based transfer learning for efficient pneumonia detection in chest X-ray images. *Applied Sciences*, 11(3), 1242.
- [38] Jiang, Z., & Chen, L. (2022). Multisemantic Level Patch Merger Vision Transformer for Diagnosis of Pneumonia. *Computational and Mathematical Methods in Medicine*, 2022.
- [39] Ma, Y., Lv, W. (2022). Identification of Pneumonia in Chest X-Ray Image Based on Transformer, *International Journal of Antennas and Propagation*, 1-8, 2022, 2022:5072666
- [40] García-Río, F., Alcázar-Navarrete, B., Castillo-Villegas, D., Cilloniz, C., García-Ortega, A., Leiro-Fernández, V., Irene Lojo-Rodríguez, Alicia Padilla-Galo, Carlos A. Quezada-Loaiza, Jose Antonio Rodríguez-Portal, Manuel Sánchez-de-la-Torre, Oriol Sibila, & Martínez-García, M. A. (2022). Biological Biomarkers in Respiratory Diseases. *Archivos de bronconeumologia*, 58(4), T323-T333.
- [41] Ölmez, H. & Tosun, M. (2023). Significance of laboratory biomarkers in monitoring patients with COVID-19 pneumonia. *Health Sciences Quarterly*, 3(1), 13-25.
- [42] Ito, A. & Ishida, T. (2020). Diagnostic markers for community-acquired pneumonia. *Annals of Translational Medicine*, 8(9).
- [43] Özdin, S., Sarisoy, G., & Böke, Ö. (2017). A comparison of the neutrophil-lymphocyte, platelet-lymphocyte and monocyte-lymphocyte ratios in schizophrenia and bipolar disorder patients—a retrospective file review. *Nordic journal of psychiatry*, 71(7), 509-512.
- [44] Bekis Bozkurt, H. (2021). Is there any relationship between C-reactive protein/albumin ratio and clinical severity of childhood community-acquired pneumonia. *Turkish Journal of Biochemistry*, 46(6), 647-653.

- [45] Hoşgün, D., Gülensoy, E. S., Akpınar, E., Ogan, N., Ateş, C., & Aydemir, S. (2022). Serum albumin and C-reactive protein/albumin ratio in community-acquired pneumonia. *Journal of Medicine and Palliative Care*, 3(2), 111-116.
- [46] Uzum, Y., & Turkkan, E. (2023). Predictivity of CRP, Albumin, and CRP to Albumin Ratio on the Development of Intensive Care Requirement, Mortality, and Disease Severity in COVID-19. *Cureus*, 15(1).
- [47] Huang, Y., Liu, A., Liang, L., Jiang, J., Luo, H., Deng, W., Lin, G., Wu, M., Li, T., Jiang, Y. (2018). Diagnostic value of blood parameters for community-acquired pneumonia. *Int Immunopharmacol*. 2018 Nov;64:10-15. doi: 10.1016/j.intimp.2018.08.022. Epub 2018 Aug 23. PMID: 30144639.
- [48] Kartal, O., & Kartal, A. T. (2017). Value of neutrophil to lymphocyte and platelet to lymphocyte ratios in pneumonia. *Bratislavske lekarske listy*, 118(9), 513-516.
- [49] Kaya, Y., Taş, N., Çanakçı, E., Cebeci, Z., Özbilen, M., Keskin, H., & Yıldırım, B. B. (2018). Relationship of neutrophil-to-lymphocyte ratio with presence and severity of pneumonia. *J Clin Anal Med*, 9(5), 452-7.
- [50] AlJame, M., Imtiaz, A., Ahmad, I., & Mohammed, A. (2021). Deep forest model for diagnosing COVID-19 from routine blood tests. *Scientific reports*, 11(1), 16682.
- [51] Effah, C. Y., Miao, R., Drokow, E. K., Agboyibor, C., Qiao, R., Wu, Y., Miao, L., Wang, Y. (2022). Machine learning-assisted prediction of pneumonia based on non-invasive measures. *Front Public Health*. 2022 Jul 28;10:938801. doi: 10.3389/fpubh.2022.938801. PMID: 35968461; PMCID: PMC9371749.
- [52] Zhang, Q., Huang, A., Shao, L., Wu, P., Heidari, A. A., Cai, Z., Mafarja, M., Alotaibi, F. S., Chen, H., Liang, G. & Ouyang, J. (2022). A machine learning framework for identifying influenza pneumonia from bacterial pneumonia for medical decision making. *Journal of Computational Science*, 65(5): 101871.
- [53] Stokes, K., Castaldo, R., Franzese, M., Salvatore, M., Fico, G., Pokvic, L. G., Badnjević, A. & Pecchia, L. (2021). A machine learning model for supporting symptom-based referral and diagnosis of bronchitis and pneumonia in limited resource settings. *Biocybernetics and Biomedical Engineering*, 41(4), 1288-1302.
- [54] Teng, X., & Gong, Y. (2018). Research on application of machine learning in data mining. In *IOP conference series: materials science and engineering* (Vol. 392, No. 6, p. 062202). IOP Publishing.
- [55] Vidhya, B., Nikhil Madhav, M., Suresh Kumar, M., & Kalanandini, S. (2022). AI Based Diagnosis of Pneumonia. *Wireless Personal Communications*, 126(4), 3677-3692.
- [56] Bellinger, C., Mohomed Jabbar, M. S., Zaiane, O., & Osornio-Vargas, A. (2017). A systematic review of data mining and machine learning for air pollution epidemiology. *BMC public health*, 17, 1-19
- [57] Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN computer science*, 2(3), 160.
- [58] Tiwari, T., Tiwari, T., & Tiwari, S. (2018). How Artificial Intelligence, Machine Learning and Deep Learning are Radically Different?. *International Journal of Advanced Research in Computer Science and Software Engineering*, 8(2), 1.
- [59] Kim, P. (2017). Matlab deep learning with machine learning, neural networks and artificial intelligence. 130(21). DOI 10.1007/978-1-4842-2845-6.
- [60] Han, J., Kamber, M., Pei, J. (2011). Data mining: Concepts and techniques, 3. izd. ISBN 978-0-12-381479-1, 225 Wyman Street, Waltham, MA 02451, US, 2012.

- [61] Sarker, I. H., Kayes, A. S. M., Badsha, S., Alqahtani, H., Watters, P., & Ng, A. (2020). Cybersecurity data science: an overview from machine learning perspective. *Journal of Big data*, 7, 1-29.
- [62] Benke, K., & Benke, G. (2018). Artificial intelligence and big data in public health. *International journal of environmental research and public health*, 15(12), 2796.
- [63] Deo, R. C. (2015). Machine learning in medicine. *Circulation*, 132(20), 1920-1930.
- [64] Güvenç, E., Çetin, G., & Koçak, H. (2021). Comparison of KNN and DNN classifiers performance in predicting mobile phone price ranges. *Advances in Artificial Intelligence Research*, 1(1), 19-28.
- [65] Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1), 21-27.
- [66] Nahzat, S., & Yağanoğlu, M. (2021). Classification of Epileptic Seizure Dataset Using Different Machine Learning Algorithms and PCA Feature Reduction Technique. *Journal of Investigations on Engineering and Technology*, 4(2), 47-60.
- [67] “K-Nearest Neighbor(KNN) Algorithm for Machine Learning” Available: <https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning> (Erişim: 23/05/2023).
- [68] Çakar, H., & Sengur, A. (2021). Machine learning based emotion classification in the COVID-19 real world worry dataset. *Computer Science*, 6(1), 24-31.
- [69] Setiawan, E. B., Widyantoro, D. H., & Surendro, K. (2016). Feature expansion using word embedding for tweet topic classification. In *2016 10th International Conference on Telecommunication Systems Services and Applications (TSSA)* (pp. 1-5). IEEE.
- [70] Demir, K., Arı, B., & Demir, F. (2023). Detection of brain tumor with a pre-trained deep learning model based on feature selection using MR images. *Firat University Journal of Experimental and Computational Engineering*, 2(1), 23-31.
- [71] Ayhan, S., & Erdoğan, Ş. (2014). Destek vektör makineleriyle sınıflandırma problemlerinin çözümü için çekirdek fonksiyonu seçimi. *Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 9(1), 175-201.
- [72] “Introduction to SVM – Support Vector Machine Algorithm of Machine Learning” Available: <https://www.analytixlabs.co.in/blog/introduction-support-vector-machine-algorithm/> (Erişim: 23/05/2023).
- [73] Fu, Y. (2017). Combination of random forests and neural networks in social lending. *Journal of Financial Risk Management*, 6(4), 418-426.
- [74] Breiman, L. (2001). Random forests. *Machine learning*, 45, 5-32. <http://dx.doi.org/10.1023/A:1010933404324>
- [75] Chang, V., Bailey, J., Xu, Q. A., & Sun, Z. (2022). Pima Indians diabetes mellitus classification based on machine learning (ML) algorithms. *Neural Computing and Applications*, 1-17.
- [76] Khan, M. Y., Qayoom, A., Nizami, M. S., Siddiqui, M. S., Wasi, S., & Raazi, S. M. K. U. R. (2021). Automated prediction of Good Dictionary EXamples (GDEX): a comprehensive experiment with distant supervision, machine learning, and word embedding-based deep learning techniques. *Complexity*, 2021, 1-18.
- [77] Richards BA, Lillicrap TP, Beaudoin P, Bengio Y, Bogacz R, Christensen A, Clopath C, Costa RP, de Berker A, Ganguli S, Gillon CJ, Hafner D, Kepecs A, Kriegeskorte N, Latham P, Lindsay GW, Miller KD, Naud R, Pack CC, Poirazi P, Roelfsema P, Sacramento J, Saxe A, Scellier B, Schapiro AC, Senn W, Wayne G, Yamins D, Zenke F, Zylberberg J, Therien D, Kording KP. (2019). A deep

- learning framework for neuroscience. *Nat Neurosci.* 2019 Nov;22(11):1761-1770. doi: 10.1038/s41593-019-0520-2. Epub 2019 Oct 28. PMID: 31659335; PMCID: PMC7115933.
- [78] Schaefer, M., Michelin, L., & Kepner, J. (2021). Naming Schema for a Human Brain-Scale Neural Network. *arXiv preprint arXiv:2109.10951*.
- [79] Aggarwal, K., Mijwil, M. M., Al-Mistarehi, A. H., Alomari, S., Gök, M., Alaabdin, A. M. Z., & Abdulrhman, S. H. (2022). Has the future started? The current growth of artificial intelligence, machine learning, and deep learning. *Iraqi Journal for Computer Science and Mathematics*, 3(1), 115-123.
- [80] Mathew, A., Amudha, P., & Sivakumari, S. (2021). Deep learning techniques: an overview. *Advanced Machine Learning Technologies and Applications: Proceedings of AMLTA 2020*, 599-608.
- [81] Ongsulee, P. (2017). Artificial intelligence, machine learning and deep learning. In *2017 15th international conference on ICT and knowledge engineering (ICT&KE)* (pp. 1-6). IEEE.
- [82] Sebastian, N., & Ankayarkanni, B. (2022). Deep Learning applications to detect pneumonia on chest X-ray: A systematic study. In *2022 International Conference on Advanced Computing Technologies and Applications (ICACTA)* (pp. 1-5). IEEE.
- [83] Shukla, S., Badal, N., & Thakur, B. K. (2023). Introduction to Deep Learning. In *Sustainable Computing: Transforming Industry 4.0 to Society 5.0* (pp. 253-267). Cham: Springer International Publishing.
- [84] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84-90.
- [85] Kim, J., Sangjun, O., Kim, Y., & Lee, M. (2016). Convolutional neural network with biologically inspired retinal structure. *Procedia Computer Science*, 88, 145-154.
- [86] Mehra, R. (2018). Breast cancer histology images classification: Training from scratch or transfer learning?. *ICT Express*, 4(4), 247-254.
- [87] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In: *2009 IEEE conference on computer vision and pattern recognition* (pp. 248–255). IEEE.
- [88] Morid, M. A., Borjali, A., & Del Fiol, G. (2021). A scoping review of transfer learning research on medical image analysis using ImageNet. *Computers in biology and medicine*, 128, 104115.
- [89] Yaman, S., Ünlü, E. I., Güler, H., Sengur, A., & Acharya, U. R. (2023). Application of novel DIRF feature selection algorithm for automated brain disease detection. *Biomedical Signal Processing and Control*, 85, 105006.
- [90] Al Hadhrami, E., Al Mufti, M., Taha, B., & Werghi, N. (2018). Transfer learning with convolutional neural networks for moving target classification with micro-Doppler radar spectrograms. In *2018 International Conference on Artificial Intelligence and Big Data (ICAIBD)* (pp. 148-154). IEEE.
- [91] “Verilerin temsillerinden öğrenmek” Available: <https://mdemlabs.github.io/AI101-7-DataRepresentations/> (Erişim tarihi:25/07/2023)
- [92] Nalluri, S., & Sasikala, R. (2022). A deep neural architecture for SOTA pneumonia detection from chest X-rays. *International Journal of System Assurance Engineering and Management*, 1-14.
- [93] Acharya, U. K., Ali, M. T., Ahmed, M. K., Siddiqui, M. T., Gupta, H., Kumar, S., & Mishra, A. S. (2023). Hybrid deep neural network for automatic detection of COVID-19 using chest x-ray images. *International Journal of Imaging Systems and Technology*, 33(4), 1129-1143.

- [94] Baltruschat, I.M., Steinmeister, L., Ittrich, H., Adam, G., Nickisch, H., Saalbach, A., von Berg, J., Grass, M., Knopp, T., (2019b). When Does Bone Suppression And Lung Field Segmentation Improve Chest X-Ray Disease Classification?, in: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), IEEE. pp. 1362–1366. Doi:10.1109/ISBI.2019. 8759510.
- [95] Deniz, E., Şengür, A., Kadiroğlu, Z., Guo, Y., Bajaj, V., & Budak, Ü. (2018). Transfer learning based histopathologic image classification for breast cancer detection. *Health information science and systems*, 6, 1-7.
- [96] Lee, W., Jo, O., & Kim, M. (2018). Application of end-to-end deep learning in wireless communications systems. *ArXiv preprint arXiv:1808.02394*.
- [97] Mahony, N.O., Campbell, S., Carvalho, A., Harapanahalli, S., Velasco-Hernández, G.A., Krpalkova, L., Riordan, D., & Walsh, J. (2020). Deep learning vs. traditional computer vision. In *Advances in Computer Vision: Proceedings of the 2019 Computer Vision Conference (CVC), Volume 1 1* (pp. 128-144). Springer International Publishing.
- [98] Schweidtmann, A. M., Esche, E., Fischer, A., Kloft, M., Repke, J. U., Sager, S., & Mitsos, A. (2021). Machine learning in chemical engineering: A perspective. *Chemie Ingenieur Technik*, 93(12), 2029-2039.
- [99] Wang, J., Ma, Y., Zhang, L., Gao, R. X., & Wu, D. (2018). Deep learning for smart manufacturing: Methods and applications. *Journal of manufacturing systems*, 48, 144-156.
- [100] Singh, N., & Sabrol, H. (2021). Convolutional neural networks-an extensive arena of deep learning. A comprehensive study. *Archives of Computational Methods in Engineering*, 28(7), 4755-4780.
- [101] Çallı, E., Sogancioglu, E., van Ginneken, B., van Leeuwen, K. G., & Murphy, K. (2021). Deep learning for chest X-ray analysis: A survey. *Medical Image Analysis*, 72, 102125.
- [102] İnik, Ö., Ülker, E. (2017). Derin öğrenme ve görüntü analizinde kullanılan derin öğrenme modelleri. *Gaziosmanpaşa Bilimsel Araştırma Dergisi*, 6(3), 85-104.
- [103] Jiang, S.; Qin, S.; Pulsipher, J.L.; Zavala, V.M. (2022) Convolutional Neural Networks: Basic Concepts and Applications in Manufacturing. In *Artificial Intelligence in Manufacturing* (pp. 63-102). Academic Press.
- [104] Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *Proceedings of the 32nd International Conference on Machine Learning*, PMLR 37:448-456, 2015.
- [105] Barhoom, A. M., & Abu-Naser, S. S. (2022). Diagnosis of pneumonia using deep learning. *International Journal of Academic Engineering Research (IJAER)* 6 (2):48-68 (2022)
- [106] Albawi, S., Mohammed, T. A., & Al-Zawi, S. (2017). Understanding of a convolutional neural network. In 2017 International Conference on Engineering and Technology (ICET) (pp. 1-6). Ieee.
- [107] Hochreiter, S. (1998). The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 6(02), 107-116.
- [108] Pauly, L., Hogg, D., Fuentes, R., & Peel, H. (2017). Deeper networks for pavement crack detection. In *Proceedings of the 34th ISARC* (pp. 479-485). IAARC.
- [109] Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M. & Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of big Data*, 8, 1-74.

- [110] Scherer, D., Müller, A., & Behnke, S. (2010). Evaluation of pooling operations in convolutional architectures for object recognition. In *International conference on artificial neural networks* (pp. 92-101). Berlin, Heidelberg: Springer Berlin Heidelberg.
- [111] Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., & Lew, M. S. (2016). Deep learning for visual understanding: A review. *Neurocomputing*, 187, 27-48.
- [112] Pelletier, C., Webb, G. I., & Petitjean, F. (2019). Temporal convolutional neural network for the classification of satellite image time series. *Remote Sensing*, 11(5), 523.
- [113] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1), 1929-1958.
- [114] Dertat, A., "Applied Deep Learning- Part 4: Convolutional Neural Networks" Available: <https://towardsdatascience.com/applied-deep-learning-part-4-convolutional-neural-networks-584bc134c1e2> (Erişim tarihi:25/07/2023)
- [115] Adaloglou, N., "Best deep CNN architectures and their principles: from AlexNet to EfficientNet" Available: <https://theaisummer.com/cnn-architectures/> (Erişim tarihi:26/07/2023)
- [116] Iandola, F.N.; Han, S.; Moskewicz, M.W.; Ashraf, K.; Dally, W.J.; and Keutzer, K. (2016). SqueezeNet: AlexNet-level accuracy with 50× fewer parameters and<0.5 MB model size. ArXiv preprint arXiv:1602.07360.
- [117] Alshazly, H., Linse, C., Barth, E., & Martinetz, T. (2021). Explainable COVID-19 Detection Using Chest CT Scans and Deep Learning. *Sensors* (Basel). 2021 Jan 11;21(2):455. doi: 10.3390/s21020455. PMID: 33440674; PMCID: PMC7828058.
- [118] Pothos, V., Kastaniotis, D., Theodorakopoulos, I., & Fragoulis, N. (2016). A fast, embedded implementation of a Convolutional Neural Network for Image Recognition. Technical Report, Aug. 2016,[Online].Available:https://www.researchgate.net/publication/306003694_A_fast_embedded_implementation_of_a_Convolutional_Neural_Network_for_Image_Recognition.
- [119] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *ArXiv preprint arXiv:1409.1556*.
- [120] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).
- [121] Sarwinda, D., Paradisa, R. H., Bustamam, A., & Anggia, P. (2021). Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer. *Procedia Computer Science*, 179, 423-431.
- [122] Ali, L., Alnajjar, F., Jassmi, H. A., Gocho, M., Khan, W., & Serhani, M. A. (2021). Performance evaluation of deep CNN-based crack detection and localization techniques for concrete structures. *Sensors*, 21(5), 1688.
- [123] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).
- [124] Pleiss, G., Chen, D., Huang, G., Li, T., Van Der Maaten, L., & Weinberger, K. Q. (2017). Memory-efficient implementation of densenets. *ArXiv preprint arXiv:1707.06990*.
- [125] Iandola, F.; Moskewicz, M.; Karayev, S.; Girshick, R.; Darrell, T.; Keutzer, K. (2014) Densenet: Implementing efficient convnet descriptor pyramids. ArXiv 2014, arXiv:1404.1869.
- [126] Çetiner, H., & Çetiner, İ. (2022). Classification of cataract disease with a DenseNet201 based deep learning model. *Journal of the Institute of Science and Technology*, 12(3), 1264-1276.

- [127] Sharma, N., Saba, L., Khanna, N. N., Kalra, M. K., Fouda, M. M., & Suri, J. S. (2022). Segmentation-based classification deep learning model embedded with explainable AI for COVID-19 detection in chest X-ray scans. *Diagnostics*, 12(9), 2132.
- [128] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4510-4520).
- [129] Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. (2017) Mobilenets: Efficient convolutional neural networks for mobile vision applications. ArXiv 2017, arXiv:1704.04861.
- [130] Eryılmaz, F., & Karacan, H. (2021). Akciğer X-Ray Görüntülerinden COVID-19 Tespitinde Hafif ve Geleneksel Evrişimsel Sinir Ağ Mimarilerinin Karşılaştırılması. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, 9(6), 26-39.
- [131] Ma, N.; Zhang, X.; Zheng, H.T.; Sun, J. (2018) ShuffleNet V2: Practical guidelines for efficient CNN architecture design. In *Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018*; pp. 116–131.
- [132] Zhang, X.; Zhou, X.; Lin, M.; Sun, J. (2018) Shufflenet: An extremely efficient convolutional neural network for mobile devices. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–22 June 2018*; pp. 6848–6856
- [133] Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1251-1258).
- [134] Zoph, B., Vasudevan, V., Shlens, J., & Le, Q. V. (2018). Learning transferable architectures for scalable image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 8697-8710).
- [135] Putra, M. D. A., Winanto, T. S., Hendrowati, R., Primajaya, A., & Adhinata, F. D. (2023). A Comparative Analysis of Transfer Learning Architecture Performance on Convolutional Neural Network Models with Diverse Datasets. *Komputika: Jurnal Sistem Komputer*, 12(1), 1-11.
- [136] Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. *ArXiv preprint arXiv:1804.02767*.
- [137] Ma, H., Liu, Y., Ren, Y., & Yu, J. (2019). Detection of collapsed buildings in post-earthquake remote sensing images based on the improved YOLOv3. *Remote Sensing*, 12(1), 44.
- [138] Jiang, C., Chen, Y., Chen, S., Bo, Y., Li, W., Tian, W., & Guo, J. (2019). A mixed deep recurrent neural network for MEMS gyroscope noise suppressing. *Electronics*, 8(2), 181.
- [139] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
- [140] Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. *Neural computation*, 12(10), 2451-2471.
- [141] Gers, F. A., Schraudolph, N. N., & Schmidhuber, J. (2002). Learning precise timing with LSTM recurrent networks. *Journal of machine learning research*, 3(Aug), 115-143.
- [142] Xiao, Y., & Yin, Y. (2019). Hybrid LSTM neural network for short-term traffic flow prediction. *Information*, 10(3), 105. <https://doi.org/10.3390/info10030105>.
- [143] Dong, J., Tu, Y., & Peng, K. (2023). Fault classification based on variable-weighted dynamic sparse stacked autoencoder for industrial processes. *The Canadian Journal of Chemical Engineering*, 101(1), 420-430.

- [144] Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, 313(5786), 504-507.
- [145] Al Mamlook, R. E., Chen, S., & Bzizi, H. F. (2020). Investigation of the performance of machine learning classifiers for pneumonia detection in chest X-ray images. In *2020 IEEE International Conference on Electro Information Technology (EIT)* (pp. 098-104). IEEE.
- [146] Bai, Y., Mei, J., Yuille, A. L., & Xie, C. (2021). Are Transformers more robust than CNNs?. *Advances in Neural Information Processing Systems*, 34, 26831-26843.
- [147] Usman, M., Zia, T., & Tariq, A. (2022). Analyzing Transfer Learning of Vision Transformers for Interpreting Chest Radiography. *J Digit Imaging*. 2022 Dec;35(6):1445-1462. Doi: 10.1007/s10278-022-00666-z. Epub 2022 Jul 11. PMID: 35819537; PMCID: PMC9274969.
- [148] Tolstikhin, I., Houlsby, N., Kolesnikov, A., Beyer, L., Zhai, X., Unterthiner, T., Yung, J., Keysers, D., Uszkoreit, J., Lucic, M., & Dosovitskiy, A. (2021). Mlp-mixer: An all-mlp architecture for vision. *Advances in Neural Information Processing Systems*, 34, 24261-24272.
- [149] Lee-Thorp, J., Ainslie, J., Eckstein, I., & Ontanon, S. (2021). Fnet: Mixing tokens with fourier transforms. *ArXiv preprint arXiv:2105.03824*.
- [150] Liu, H., Dai, Z., So, D., & Le, Q. V. (2021). Pay attention to mlps. *Advances in Neural Information Processing Systems*, 34, 9204-9215.
- [151] Zheng, A. (2015). Evaluating machine learning models: a beginner's guide to key concepts and pitfalls. O'Reilly Media.
- [152] Yadav, S., & Shukla, S. (2016). Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification. In *2016 IEEE 6th International conference on advanced computing (IACC)* (pp. 78-83). IEEE.
- [153] Blum, A., Kalai, A., & Langford, J. (1999). Beating the hold-out: Bounds for k-fold and progressive cross-validation. In *Proceedings of the twelfth annual conference on Computational learning theory (COLT '99)*. Association for Computing Machinery, New York, NY, USA, 203–208. <https://doi.org/10.1145/307400.307439>
- [154] “Different Types of Cross-Validations in Machine Learning and Their Explanations” <https://www.turing.com/kb/different-types-of-cross-validations-in-machine-learning-and-their-explanations> (Erişim tarihi:01/06/2023)
- [155] Al Mamlook, R. E., Chen, S., & Bzizi, H. F. (2020). Investigation of the performance of machine learning classifiers for pneumonia detection in chest X-ray images. In *2020 IEEE International Conference on Electro Information Technology (EIT)* (pp. 098-104). IEEE.
- [156] Güvenç, E., Çetin, G., & Koçak, H. (2021). Comparison of KNN and DNN classifiers performance in predicting mobile phone price ranges. *Advances in Artificial Intelligence Research*, 1(1), 19-28.
- [157] Pereira, A. I., Franco-Gonçalo, P., Leite, P., Ribeiro, A., Alves-Pimenta, M. S., Colaço, B., Loureiro, C., Gonçalves, L., Filipe, V., Ginja, M. (2023). Artificial Intelligence in Veterinary Imaging: An Overview. *Vet Sci*. 2023 Apr 28;10(5):320. doi: 10.3390/vetsci10050320. PMID: 37235403; PMCID: PMC10223052.
- [158] Nachar, N. (2008). The Mann-Whitney U: A test for assessing whether two independent samples come from the same distribution. *Tutorials in quantitative Methods for Psychology*, 4(1), 13-20.

- [159] Mindrila, D., & Balentyne, P. (2013). The Chi Square Test. *The Basic Practice of Statistics*. 6th ed. New York: WH Freeman.
- [160] Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C., & Keshri A. (2019). Descriptive statistics and normality tests for statistical data. *Ann Card Anaesth*. 2019 Jan-Mar;22(1):67-72. doi: 10.4103/aca.ACA_157_18. PMID: 30648682; PMCID: PMC6350423.
- [161] Kadiroğlu, Z., Deniz, E., & Şenyiğit, A. (2023). Göğüs röntgen görüntülerinde pnömoni tespiti için derin öğrenme modellerinin karşılaştırılması. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 39(2), 729-740. <https://doi.org/10.17341/gazimmfd.1204092>
- [162] Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., & Torralba, A. (2016). Learning Deep Features for Discriminative Localization," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 2921-2929, doi: 10.1109/CVPR.2016.319.
- [163] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization, *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, 2017, pp. 618-626, doi: 10.1109/ICCV.2017.74.
- [164] Loh, H. W., Ooi. C. P., Seoni, S., Barua, P. D., Molinari, F., & Acharya, U. R. (2022). Application of explainable artificial intelligence for healthcare: A systematic review of the last decade (2011-2022). *Comput Methods Programs Biomed*. 2022 Nov; 226:107161. Doi: 10.1016/j.cmpb.2022.107161. Epub 2022 Sep 27. PMID: 36228495.
- [165] Landau, S., Everitt, B.S. (2003). *A Handbook of Statistical Analyses Using SPSS* (1st ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9780203009765>
- [166] Harimurti, K., Setiati, S. (2007). C-reactive protein levels and decrease of albumin levels in hospitalized elderly patients with community-acquired pneumonia. *Acta Med Indones*. 2007 Jan-Mar;39(1):13-8. PMID: 17297204.
- [167] Ng, W. W., Lam, S. M., Yan, W. W., Shum, H. P. (2022). NLR, MLR, PLR and RDW to predict outcome and differentiate between viral and bacterial pneumonia in the intensive care unit. *Sci Rep*. 2022 Sep 24;12(1):15974. Doi: 10.1038/s41598-022-20385-3. PMID: 36153405; PMCID: PMC9509334.
- [168] Bekdas, M., Goksugur, S. B., Sarac, E. G., Erkocoglu, M., Demircioglu, F. (2014). Neutrophil/lymphocyte and C-reactive protein/mean platelet volume ratios in differentiating between viral and bacterial pneumonias and diagnosing early complications in children. *Saudi Med J*. 2014 May;35(5):442-7. PMID: 24825803.

EKLER

EK- 1: TEZ ÇALIŞMASI VERİLERİNİN KULLANIMINA DAİR ETİK KURUL ONAY SAYFASI

DİCLE ÜNİVERSİTESİ TIP FAKÜLTESİ GİRİŞİMSEL OLMAYAN KLİNİK ARAŞTIRMALAR ETİK KURULU DİCLE UNIVERSITY MEDICAL FACULTY ETHICS COMMITTEE FOR NONINTERVENTIONAL STUDIES					
KARAR					
Zehra KADIOĞLU isimli araştırmacılar tarafından planlanan "Sayısal veri ve göğüs röntgen görüntülerinden derin öğrenme yaklaşımları ile pnömoni teşhisi" başlıklı araştırmaya Dicle Üniversitesi Tıp Fakültesi Etik Kurul'u tarafından toplantıda hazır bulunan üyeler tarafından oy birliği ile onay verilmiştir. Klinik araştırma tamamlanıp yayın aşamasına geldiğinde, yayına sunulan bildiri veya makalenin bir örneğinin Etik Kurul'a verilmesi zorunludur.					
DECISION					
The project titled as "Diagnosis of pneumonia with deep learning approaches from numerical data and chest x-ray image" planned by Erkan DENİZ, Abdurrahman ŞENYİĞİT, Zehra KADIOĞLU has been approved by Ethics Committee of Dicle University Faculty of Medicine.					
Oturum No (Meeting number):		Tarih (Date): 13.10.2021		Saat (Hour): 10:00-12:00	
KURUL BAŞKANI (CHIEF)		Prof. Dr. [REDACTED]			
KURUL ÜYELERİ / MEMBERS					
	ÜNVANI	ADI-SOYADI	KURUMU	BRANŞI	İMZA
1	Prof. Dr.	[REDACTED]	Dicle Üniversitesi Tıp Fakültesi	Tıbbi Farmakoloji	[REDACTED]
2	Prof. Dr.	[REDACTED]	Dicle Üniversitesi Tıp Fakültesi	Anesteziyoloji ve Reanimasyon	[REDACTED]
3	Prof. Dr.	[REDACTED]	Dicle Üniversitesi Tıp Fakültesi	İç Hastalıkları	[REDACTED]
4	Doç. Dr.	[REDACTED]	Dicle Üniversitesi Tıp Fakültesi	Genel Cerrahi	[REDACTED]
5	Doç. Dr.	[REDACTED]	Dicle Üniversitesi Tıp Fakültesi	Tıbbi Biyokimya	[REDACTED]
6	Dr. Öğretim Üyesi	[REDACTED]	Dicle Üniversitesi Tıp Fakültesi	Biyoistatistik	[REDACTED]
7	Dr. Öğretim Üyesi	[REDACTED]	Dicle Üniversitesi Tıp Fakültesi	Tıbbi Biyoloji	[REDACTED]
8	Dr. Öğretim Üyesi	[REDACTED]	Dicle Üniversitesi Tıp Fakültesi	Tıbbi Patoloji	[REDACTED]
9	Dr. Öğretim Üyesi	[REDACTED]	Dicle Üniversitesi Tıp Fakültesi	Radyoloji	[REDACTED]
10	Avukat	[REDACTED]	Dicle Üniversitesi Hastaneleri	Hukuk İşleri Birimi	[REDACTED]

Dicle Üniversitesi Tıp Fakültesi Dekanlık Binası Zemin Kat 21280 Kampüs/DIYARBAKIR
Telefon:+90.412 . 248 80 01-16/4631 Faks:+90.412. 248 84 40 kuruletikdiyar@gmail.com

CS CamScanner ile tarandı

Şekil E1.1. Tez Çalışması Verilerinin Kullanımına Dair Etik Kurul Onay Sayfası

ÖZGEÇMİŞ

Zehra KADIROĞLU

[illegible]

AKADEMİK FAALİYETLER

Makaleler:

1. Deniz, E., Şengür, A., Kadiroğlu, Z., Guo, Y., Bajaj, V., & Budak, Ü. (2018). Transfer learning based histopathologic image classification for breast cancer detection. *Health information science and systems*, 6, 1-7.
2. Kadiroğlu, Z., Deniz, E., & Şenyiğit, A. (2023). Göğüs röntgen görüntülerinde pnömوني tespiti için derin öğrenme modellerinin karşılaştırılması. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 39(2), 729-740. <https://doi.org/10.17341/gazimmfd.1204092>

Bildiriler:

1. A. Demir, B. N. Akılotu, Z. Kadiroğlu and A. Şengür, "Bitcoin Price Prediction Using Machine Learning Methods," *2019 1st International Informatics and Software Engineering Conference (UBMYK)*, Ankara, Turkey, 2019, pp. 1-4, doi: 10.1109/UBMYK48245.2019.8965445.
2. B. N. Akılotu, Z. Kadiroğlu and A. Şengür, "Information Security and Related Machine Learning Applications," *2019 1st International Informatics and Software Engineering Conference (UBMYK)*, Ankara, Turkey, 2019, pp. 1-6, doi: 10.1109/UBMYK48245.2019.8965484.
3. Z. Kadiroğlu, B. N. Akılotu and A. Şengür, "Mechanism of Bitcoin and Investigation of the Studies in the Literature Related to Bitcoin," *2019 1st International Informatics and*

Software Engineering Conference (UBMYK), Ankara, Turkey, 2019, pp. 1-5, doi: 10.1109/UBMYK48245.2019.8965449.

4. Akılotu, B. N., Kadiroğlu, Z., Şengür, A., & Kayaoğlu, M. (2019). Evrişimsel Sinir Ağları ve Transfer Öğrenme Yöntemi Kullanılarak Sıtma Tespiti. In *International Engineering and Science Symposium, (IESS'19), June 2019. Siirt.*
5. Kadiroğlu, Z., Akılotu, B. N., Şengür, A., Kayaoğlu, M., Derin Öznelikler ve Renk Kullanılarak Dermoskopik Görüntülerin Sınıflandırılması. *International Engineering and Sciences Symposium (IESS'19)*, June 2019.
6. Kadiroğlu, Z., Şengür, A., & Deniz, E. (2018, November). Classification of histopathological breast cancer images with low level texture features. In *2018 International Engineering and Natural Sciences Conference (IENSC)* (pp. 1765-1772).
7. Şengür, A., Akılotu, B. N., Tuncer, S. A., Kadiroğlu, Z., Yavuzkılıç, S., Budak, Ü., & Deniz, E. (2018, May). Optic disc determination in retinal images with deep features. In *2018 26th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.

Projeler:

1. 122E215 no'lu "Nonstemi Hastalarının Tıkalı Koroner Arterlerinin Spesifik Derin Öğrenme Tabanlı Yaklaşımlarla Tespit Edilmesi" isimli Tübitak 1001 projesinde Birlikte Çalışıp Birlikte Başaracağız (BİÇABA) bursiyeri (05/04/23-devam ediyor.)