



Teknoloji Fakültesi

BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

**EVRIŞİMSEL SİNİR AĞLARI ve
TRANSFORMER MİMARİLERİNİN AKCİĞER
BT GÖRÜNTÜLERİNİ SINIFLANDIRMA
PERFORMANS ANALİZİ**

BİTİRME PROJESİ

Bilgisayar Mühendisliği Bölümü

"

DANIŞMAN

Doç. Dr. AYŞE BERNA ALTINEL

İSTANBUL, 2025

"

MARMARA ÜNİVERSİTESİ
TEKNOLOJİ FAKÜLTESİ
BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

Marmara Üniversitesi Teknoloji Fakültesi Bilgisayar Mühendisliği Öğrencileri Sude Nur GÜLCAN, Karun ACAR, Elif TİLKİ tarafından “EVİRİŞİMSEL SİNİR AĞLARI ve TRANSFORMER MİMARİLERİNİN AKCİĞER BT GÖRÜNTÜLERİNİ SINIFLANDIRMA PERFORMANS ANALİZİ” başlıklı proje çalışması, 19/06/2025 tarihinde savunulmuş ve jüri üyeleri tarafından başarılı bulunmuştur.

Jüri Üyeleri

Doç. Dr. AYŞE BERNA ALTINEL	(Danışman)	
Marmara Üniversitesi	
Dr. Öğr. Üyesi NEŞE ÖZDEMİR	(Üye)	
Marmara Üniversitesi	
Arş. Gör. NURSAÇ KURT	(Üye)	
Marmara Üniversitesi	

ÖNSÖZ

Proje çalışmamız süresince karşılaştığımız bütün problemlerde, sabırla yardım ve bilgilerini esirgemeyen, tüm desteğini sonuna kadar yanımızda hissettiğimiz değerli hocalarımız, sayın Doç. Dr. Ayşe Berna Altınel ve sayın Arş. Gör. Dr. Abdulsamet Aktaş'a en içten teşekkürlerimizi sunarız.

İÇİNDEKİLER

1. GİRİŞ.....	1
1.1 Proje Çalışmasının Amacı ve Önemi.....	2
2. LİTERATÜR TARAMASI.....	2
3. MATERYAL VE YÖNTEM.....	5
3.1 Veri Setleri.....	5
3.1.1 Veri Ön İşlemleri.....	5
3.1.2 The IQ-OTH/NCCD Lung Cancer Dataset.....	8
3.1.3 SPIE-AAPM-NCI Lung Nodule Classification Challenge Dataset.....	9
3.1.4 CT-Scan Images Dataset	12
3.2 Deney Ortamı Özellikleri ve Kullanılan Kütüphaneler.....	13
3.2.1 Deney Ortamı ve Makine Özellikleri.....	13
3.2.2 Kullanılan Kütüphaneler	14
3.2.2.1 NumPy	15
3.2.2.2 Matplotlib.....	15
3.2.2.3 Seaborn	15
3.2.2.4 Torch (PyTorch).....	15
3.2.2.5 Torchvision	15
3.2.2.6 Scikit-learn.....	15
3.2.2.7 Timm.....	16
3.2.2.8 OpenCV	16
3.2.2.9 Pandas	16
3.2.2.10 Glob	16
3.2.2.11 Tqdm.....	16
3.2.2.12 Transformers	17

3.2.2.13 Safetensors	17
3.2.2.14 Torch.cuda.amp.....	17
3.2.2.15 Torch.nn.functional.....	17
3.2.2.16 Torch.multiprocessing.....	17
3.2.2.17 JSON & Pickle.....	18
3.2.2.18 Albumentations	18
3.2.2.19 Time.....	18
3.3 Metotlar	18
3.3.1 VGG16	18
3.3.2 ResNet50	21
3.3.3 EfficientNetV2-S	23
3.3.4 SwinV2-CR-Small-224.....	24
3.3.5. Beit-Base	26
3.3.6 ConvNeXtV2	28
3.3.7 CoAtNet-V1	30
3.4. Performans Metrikleri.....	32
3.4.1 Doğruluk (Accuracy):.....	32
3.4.2 Kesinlik (Precision):.....	32
3.4.3 Duyarlılık (Recall):	33
3.4.4 F1 Skoru (F1 Score):.....	33
3.4.5 Confusion Matrix	33
3.4.6 T Testi	34
4. BULGULAR VE TARTIŞMA.....	36
4.1. Model Performans Analizi	36
4.2. Veri Artırımı Uygulanmamış ve Veri Arttırımı Uygulanmış Veri Setleri	

Üzerindeki Modellerin Sonuçlarının Karşılaştırılması.....	41
4.2.1 IQ-OTH/NCCD	42
4.2.2 SPIE-AAPM Dataset	43
4.2.3 CT-Scan Images Dataset	45
4.3. T- Testi Karşılaştırması.....	47
5. SONUÇLAR.....	50
[EK].....	58

ÖZET

EVRIŞİMSEL SİNİR AĞLARI ve TRANSFORMER MİMARİLERİNİN AKCİĞER BT GÖRÜNTÜLERİNİ SINIFLANDIRMA PERFORMANS ANALİZİ

Bu çalışma, derin öğrenme algoritmalarının tıbbi görüntüleme alanındaki uygulamalarını inceleyerek, özellikle akciğer kanseri gibi hayati öneme sahip hastalıkların erken teşhisine yönelik potansiyelini ortaya koymayı amaçlamaktadır. Akciğer kanseri, dünya genelinde en yaygın görülen ve ölüm oranı en yüksek olan kanser türlerinden biridir. Bu nedenle hastalığın erken ve doğru teşhisi, tedavi başarısı açısından kritik öneme sahiptir. Bu bağlamda, bilgisayarlı tomografi görüntülerinin otomatik olarak işlenerek analiz edilmesi, tanı sürecinde hekimlere destek sağlayabilecek yapay zekâ tabanlı sistemlerin geliştirilmesi oldukça önem taşımaktadır.

Çalışmada, IQ-OTH/NCCD, SPIE-AAPM-NCI Lung Nodule Classification Challenge ve CT-Scan olmak üzere üç farklı açık erişimli akciğer BT görüntüsü veri seti kullanılmıştır. Görüntüler, öncelikle veri ön işleme adımlarından geçirilmiş, ardından 5 katlı çapraz doğrulama yöntemiyle eğitim ve test alt kümelerine ayrılmıştır. Ayrıca bu veri setlerinde veri artırımı teknikleri uygulanmıştır. Sınıflandırma sürecinde VGG16, ResNet50, EfficientNetV2-S, SwinV2, BEiT-Base, ConvNeXtV2 ve CoAtNet gibi güncel CNN ve Transformer tabanlı derin öğrenme modelleri hem veri artırımı yapılmadan hem de veri artırımı yapıldıktan sonra karşılaştırmalı olarak değerlendirilmiştir. Model performansları; doğruluk, kesinlik, duyarlılık, F1 skoru gibi sınıflandırma metrikleri ile ölçülmüş, ayrıca sonuçların istatistiksel anlamlılığı t-testi ile analiz edilmiştir.

Elde edilen bulgular, transformer tabanlı gelişmiş modellerin yüksek başarı oranlarına ulaştığını, derin öğrenme algoritmalarının tıbbi görüntü analizi alanında güçlü bir potansiyele sahip olduğunu ve akciğer kanseri gibi ölümcül hastalıkların teşhisinde yapay zekâ destekli karar destek sistemlerinin etkili bir şekilde kullanılabileceğini ortaya koymaktadır.

Haziran, 2025

Öğrenciler

Sude Nur GÜLCAN

Karun ACAR

Elif TİLKİ

ABSTRACT

PERFORMANCE ANALYSIS OF CONVOLUTIONAL NEURAL NETWORKS AND TRANSFORMER ARCHITECTURES IN THE CLASSIFICATION OF LUNG CT IMAGES

This study aims to examine the applications of deep learning algorithms in the field of medical imaging and to reveal their potential for early diagnosis of vital diseases, especially lung cancer. Lung cancer is one of the most common cancers worldwide and has the highest mortality rate. Therefore, early and accurate diagnosis of the disease is of critical importance for the success of treatment. In this context, automatic processing and analysis of computed tomography images and the development of artificial intelligence-based systems that can support physicians in the diagnosis process are of great importance.

Three different open access lung CT image datasets, IQ-OTH/NCCD, SPIE-AAPM-NCI Lung Nodule Classification Challenge and CT-Scan, were used in the study. The images were first subjected to data preprocessing steps and then divided into training and test subsets with a 5-fold cross-validation method. In addition, data augmentation techniques were applied to these datasets. In the classification process, current CNN and Transformer-based deep learning models such as VGG16, ResNet50, EfficientNetV2-S, SwinV2, BEiT-Base, ConvNeXtV2 and CoAtNet were comparatively evaluated both without and after data augmentation. Model performances were measured with classification metrics such as accuracy, precision, sensitivity, F1 score, and the statistical significance of the results was analyzed with t-test.

The findings show that advanced transformer-based models achieve high success rates, deep learning algorithms have strong potential in the field of medical image analysis, and artificial intelligence-supported decision support systems can be used effectively in the diagnosis of fatal diseases such as lung cancer.

June, 2025

Students

Sude Nur GÜLCAN

Karun ACAR

Elif TİLKİ

SEMBOLLER

\bar{X} :	: örneklem ortalaması,
μ :	: popülasyon ortalaması (veya karşılaştırma grubunun ortalaması),
s:	: örneklem standart sapması,
n	: örneklem büyüklüğüdür.

KISALTMALAR

ACO	: Ant Colony Optimization
AOC	: Area Under the Curve
BEiT	: Bidirectional Encoder representation from Image Transformers
BT	: Bilgisayarlı Tomografi
CNN	: Convolutional Neural Network
Coatnet	: Convolutional Attention Network
ConvNextV2	: Convolutional Network Next Version 2
DICOM	: Digital Imaging and Communications in Medicine
EfficientNetV2	: Efficient Network Version 2
MR	: Manyetik rezonans görüntüleme
NSGA-II	: Non-dominated Sorting Genetic Algorithm II
PCO	: Particle Swarm Optimization
ReLU	: Rectified Linear Unit
ResNet	: Residual Network
SVM	: Support Vector Machine
SwinV2	: Shifted Window Transformer Version 2
VGG	: Visual Geometry Group

ŞEKİL LİSTESİ

Şekil 1	Veri Artırımı Teknikleri Kod Bloğu.....	7
Şekil 2	Veri Setlerinde Veri Artırımı Öncesi ve Sonrasında Toplam Sayıları.....	7
Şekil 3	CT-Scan Images Datasetinin Veri Artırımı Öncesi ve Sonrası Sınıflara Göre Dağılımı.....	7
Şekil 4	IQ-QTH/NCCD Lung Cancer Datasetinin Veri Artırımı Öncesi ve Sonrası Sınıflara Göre Dağılımı.....	7
Şekil 5	SPIE-AAPM Lung Nodule Datasetinin Veri Artırımı Öncesi ve Sonrası Sınıflara Göre Dağılımı	7
Şekil 6	Orijinal Görüntüye Brightness İşlemi Uygulanmış Hali.....	8
Şekil 7	Orijinal Görüntüye Döndürme İşlemi Uygulanmış Hali.....	8
Şekil 8	Orijinal Görüntü	8
Şekil 9	Benign Case (16) [14].....	9
Şekil 10	Malignant (165) [14]	9
Şekil 11	Normal Case (32) [14].....	9
Şekil 12	CT-048_1-117 (malignant) [15]	12
Şekil 13	CT-060_1-132 (benign) [15]	12
Şekil 14	Cancerous108 [16].....	13
Şekil 15	Non-Cancerous85 [16]	13
Şekil 16	VGG 16 Model Mimarisi [39].....	19
Şekil 17	VGG 16 Katmanları [40].....	19
Şekil 18	VGG16 Model Oluşturma Kod Bloğu.....	20
Şekil 19	ResNet50 Mimarisi [43]	21
Şekil 20	ResNet50 Model Oluşturma Kod Bloğu.....	22
Şekil 21	EfficientNetV2-S Mimarisi [46].....	23

Şekil 22 EfficientnetV2-S Model Oluşturma Kod Bloğu	24
Şekil 23 SwinV2-CR-Small-224 Mimarisi [49]	25
Şekil 24 SwinV2-CR-Small-224 Model Oluşturma Kod Bloğu.....	26
Şekil 25 Beit-Base Modelinin Mimarisi [52].....	27
Şekil 26 Beit-Base Model Oluşturma Kod Bloğu.....	28
Şekil 27 ConvNeXtV2 Modelinin Mimarisi [53]	29
Şekil 28 ConvNeXtV2 Modelinin Katmanları [54]	29
Şekil 29 ConvNeXtV2 Model Oluşturma Kod Bloğu	30
Şekil 30 CoAtNet Modelinin Katmanları [56].....	30
Şekil 31 CoAtNet Model Oluşturma Kod Bloğu	31
Şekil 32 a) Vgg 16 Modelinin Veri Artırımın Yapılmamış IQ-OTH/NCCD Veri Setindeki Confusion Matrix Sonuçları b) Vgg 16 Modelinin Veri Artırımın Yapılmamış SPIE-AAPM-NCI Veri Setindeki Confusion Matrix Sonuçları c) Vgg 16 Modelinin Veri Artırımın Yapılmamış CT-SCAN Images Veri Setindeki Confusion Matrix Sonuçları ..	36
Şekil 33 a) ResNet50 Modelinin Veri Artırımın Yapılmamış IQ-OTH/NCCD Veri Setindeki Confusion Matrix Sonuçları b) ResNet50 Modelinin Veri Artırımın Yapılmamış SPIE-AAPM-NCI Veri Setindeki Confusion Matrix Sonuçları c) ResNet50 Modelinin Veri Artırımın Yapılmamış CT-SCAN Images Veri Setindeki Confusion Matrix Sonuçları	37
Şekil 34 a) EfficientNetV2-S Modelinin Veri Artırımın Yapılmamış IQ-OTH/NCCD Veri Setindeki Confusion Matrix Sonuçları b) EfficientNetV2-S Modelinin Veri Artırımın Yapılmamış SPIE-AAPM-NCI Veri Setindeki Confusion Matrix Sonuçları c) EfficientNetV2-S Modelinin Veri Artırımın Yapılmamış CT-SCAN Images Veri Setindeki Confusion Matrix Sonuçları.....	38
Şekil 35 a) SwinV2-CR-Small-224 Modelinin Veri Artırımın Yapılmamış IQ-OTH/NCCD Veri Setindeki Confusion Matrix Sonuçları b) SwinV2-CR-Small-224 Modelinin Veri Artırımın Yapılmamış SPIE-AAPM-NCI Veri Setindeki Confusion Matrix Sonuçları c) SwinV2-CR-Small-224 Modelinin Veri Artırımın Yapılmamış CT-SCAN Images Veri Setindeki Confusion Matrix Sonuçları.....	39

Şekil 36 a) BEiT-base Modelinin Veri Artırımın Yapılmamış IQ-OTH/NCCD Veri Setindeki Confusion Matrix Sonuçları b) BEiT-base Modelinin Veri Artırımın Yapılmamış SPIE-AAPM-NCI Veri Setindeki Confusion Matrix Sonuçları c) BEiT-base Modelinin Veri Artırımın Yapılmamış CT-SCAN Images Veri Setindeki Confusion Matrix Sonuçları39

Şekil 37 a) ConvNeXtV2 Modelinin Veri Artırımın Yapılmamış IQ-OTH/NCCD Veri Setindeki Confusion Matrix Sonuçları b) ConvNeXtV2 Modelinin Veri Artırımın Yapılmamış SPIE-AAPM-NCI Veri Setindeki Confusion Matrix Sonuçları c) ConvNeXtV2 Modelinin Veri Artırımın Yapılmamış CT-SCAN Images Veri Setindeki Confusion Matrix Sonuçları.....40

Şekil 38 a) CoAtNet Modelinin Veri Artırımın Yapılmamış IQ-OTH/NCCD Veri Setindeki Confusion Matrix Sonuçları b) CoAtNet Modelinin Veri Artırımın Yapılmamış SPIE-AAPM-NCI Veri Setindeki Confusion Matrix Sonuçları c) CoAtNet Modelinin Veri Artırımın Yapılmamış CT-SCAN Images Veri Setindeki Confusion Matrix Sonuçları41

TABLO LİSTESİ

Tablo 1 Literatür Taraması Özet Tablosu.....	5
Tablo 2 Çalışma Ortamı Tablosu	14
Tablo 3 Confusion Matrix Değerleri Tablosu	34
Tablo 4 IQ-OTH/NCCD Lung Cancer Datasetin Modellere Göre Karşılaştırması.....	43
Tablo 5 IQ-OTH/NCCD Lung Cancer Datasetin Veri Artırımı Sonrasında Modellere Göre Karşılaştırması	43
Tablo 6 SPIE-AAPM Lung Nodule Classification Challenge Datasetin Modellere Göre Karşılaştırması	44
Tablo 7 SPIE-AAPM Lung Nodule Classification Challenge Datasetin Veri Artırımı Sonrasında Modellere Göre Karşılaştırması	45
Tablo 8 CT-Scan Images Datasetin Modellere Göre Karşılaştırması.....	46
Tablo 9 CT-Scan Images Datasetinin Veri Artırımı Sonrası Modellere Göre Karşılaştırması	46
Tablo 10 Modellerin T-Test Karşılaştırma Sonuçları.....	48

1. GİRİŞ

Kanserin küresel çapta görülme sıklığı ve kansere bağlı ölüm oranları her geçen yıl artış göstermektedir [1][2]. Dünya Sağlık Örgütü'nün (WHO) verilerine göre, en yaygın kanser türleri arasında ilk sırada meme kanseri, ikinci sırada ise akciğer kanseri yer almaktadır [3]. Akciğer kanserinin zamanında ve doğru şekilde teşhis edilmesi, hastaların hayatta kalma oranını belirgin şekilde artırmaktadır [4][5]. Geleneksel tanı yöntemleri, radyologların bilgisayarlı tomografi ve manyetik rezonans görüntülerini manuel olarak incelemesine dayanmakta olup, bu süreç hem zaman açısından maliyetli hem de insan hatalarına açık bir yaklaşım sunmaktadır.

Son yıllarda, yapay zekâ ve makine öğrenmesi tabanlı sistemlerin tıbbi görüntü analizinde kullanımının yaygınlaşması, hastalık teşhisinde hekimlere önemli bir destek sağlamaktadır. Derin öğrenme algoritmalarının ilerlemesiyle birlikte, özellikle Evrişimli Sinir Ağları (CNN) gibi modeller, BT ve MR gibi tıbbi görüntüleme yöntemlerinden elde edilen verileri işleyerek akciğer kanseri gibi hastalıkların teşhisinde yüksek doğruluk oranlarına ulaşmaktadır [6]. Dolayısıyla, tıbbi görüntü analizi ve yapay zekâ destekli sınıflandırma yöntemleri, akciğer kanserinin erken teşhisinde kritik bir rol oynamaktadır.

Bu çalışmada ise IQ-OTH/NCCD Lung Cancer Dataset, SPIE-AAPM-NCI Lung Nodule Classification Challenge Dataset ve CT-Scan Images Dataset veri setleri kullanılarak akciğer BT görüntülerinin derin öğrenme algoritmaları ve transformer mimarileri aracılığıyla sınıflandırılması ve performans analizinin yapılması amaçlanmıştır. Bu bağlamda, derin öğrenme temelli tanı sistemlerinin tıbbi görüntü analizi alanındaki etkinliğini ortaya koyarak, akciğer kanseri teşhisinde daha hızlı, güvenilir ve objektif karar destek sistemlerinin geliştirilmesine bilimsel katkı sunması hedeflenmektedir.

1.1 Proje Çalışmasının Amacı ve Önemi

Bu araştırma, derin öğrenme tabanlı tanı sistemlerinin tıbbi görüntü analizi alanındaki potansiyelini ortaya koyarak, akciğer kanserinin daha hızlı, güvenilir ve nesnel bir şekilde teşhis edilmesine katkı sunmayı hedeflemektedir. Çalışmanın temel amacı, farklı derin öğrenme algoritmalarının sınıflandırma performanslarını analiz ederek en etkili modelin belirlenmesini sağlamaktır. Böylece, elde edilen sonuçlar doğrultusunda geliştirilecek olan modellerin klinik uygulamalara entegre edilerek radyologlara karar destek mekanizması sunması ve akciğer kanseri teşhis sürecinde doğruluğun artırılması hedeflenmektedir.

2. LİTERATÜR TARAMASI

Son yıllarda, tıbbi görüntüleme alanında derin öğrenme ve yapay zeka teknolojilerinin kullanımı önemli bir gelişim göstermiştir. Özellikle akciğer kanseri gibi ciddi hastalıkların erken teşhisinde ve doğru sınıflandırılmasında, bilgisayarla desteklenen görüntü işleme yöntemleri büyük bir potansiyel taşımaktadır. Bu literatür taramasında, akciğer kanseri tespiti ve kullanılan derin öğrenme teknikleri incelenmiştir.

Siddiqui ve arkadaşları, akciğer kanseri sınıflandırmasında Vision Transformer (ViT) tabanlı modellerin etkinliğini değerlendirmişlerdir. Çalışmada CoAtNet, Nested-TNT, CrossViT ve HorNet gibi gelişmiş mimariler, IQ-OTH/NCCD (3 sınıf) ve LC25000 (5 sınıf) veri kümelerinde transfer öğrenme ile eğitilmiştir. Görüntülere ön işleme olarak histogram eşitleme, median filtre ve Gaussian yumuşatma uygulanmış; ardından veri artırma teknikleri ile sınıf dengesi sağlanmıştır. En başarılı sonuçları LeViT modeli elde etmiş; IQ-OTH/NCCD veri setinde %99.43 doğruluk, %98.44 F1 skoru ve %98.36 MCC değerine ulaşmıştır. CoAtNet modeli ise %99.10 doğruluk ve %97.27 PR AUC ile güçlü bir performans sergilemiştir. CoAtNet'in konvolüsyon ve self-attention yapısını birleştiren hibrit mimarisi, medikal görüntülerde hem yerel hem de küresel özelliklerin etkili şekilde öğrenilmesini sağlamıştır. Bu tür ViT tabanlı modellerin, medikal görüntü sınıflandırmalarında yüksek doğruluk ve genel performans sunduğu gösterilmiştir [7].

Hassan ve arkadaşları (2025), akciğer kanserinin erken teşhisine yönelik olarak, bilgisayarlı tomografi (BT) görüntüleri üzerinde EfficientNet-B0 mimarisi tabanlı bir derin öğrenme modeli geliştirmiştir. Çalışmada, IQ-OTH/NCCD veri seti ile ek bir veri kümesinin birleştirilmesiyle toplamda yaklaşık 1.415 görüntü kullanılarak sınıflandırma

işlemi gerçekleştirilmiştir. Modelin eğitiminde, dönme, yansıtma ve affine dönüşümler gibi ileri düzey veri artırma teknikleri uygulanmış; ayrıca görüntü kalitesini artırmak amacıyla medyan filtreleme ile gürültü azaltımı sağlanmıştır. Performans değerlendirmesi için kullanılan 5 katlı çapraz doğrulama sonucunda, model en iyi senaryoda %99.32 doğruluk, %98.20 makro ortalama kesinlik (precision) ve %97.80 makro ortalama duyarlılık (recall) değerlerine ulaşmıştır. Tüm doğrulama katlarında elde edilen ortalama doğruluk değeri ise %97.43 olarak rapor edilmiştir. Ek olarak, veri setindeki sınıf dengesizliğini gidermek amacıyla ağırlıklı çapraz entropi kaybı fonksiyonu tercih edilmiştir. Bu çalışma, EfficientNet-B0 mimarisinin tıbbi görüntü sınıflandırmasında yüksek doğrulukla çalışabildiğini ve k-kat çapraz doğrulamanın modelin genellenebilirliğini artırmadaki etkili rolünü ortaya koymaktadır [8].

Adams ve arkadaşları (2018), Haralick doku özellikleri ve Destek Vektör Makineleri kullanarak akciğer BT görüntülerinde tümör tespiti ve malign-benign sınıflandırması yapan tam otomatik bir sistem geliştirmiştir. SPIE-AAPM Lung CT Challenge veri setiyle yapılan değerlendirmede, tümör tespiti için iki farklı strateji uygulanmıştır. Strateji 1'de, tüm tespit edilen pencereler tümör olarak kabul edilmiş ve bu yöntemle %80.82 duyarlılık sağlanmış ancak düşük doğruluk (Precision: %34.50) nedeniyle F-measure değeri 0.4835'te kalmıştır. Strateji 2'de ise yalnızca en yüksek SVM skoruna sahip pencere değerlendirilmiş; yanlış pozitif oranı önemli ölçüde azalmış (Precision: %74.58), ancak duyarlılık %60.67'ye düşmüştür. Bu strateji ile F-measure 0.6666'ya yükselmiştir. Ayrıca, malign-benign ayrımı için sadece istatistiksel olarak anlamlı bulunan doku özellikleriyle eğitilen ikinci bir SVM modeli, sınırlı eğitim verisiyle 0.61 AUC değerine ulaşarak, benzer yöntemler arasında ikinci en iyi sonucu vermiştir [9].

Thambi ve arkadaşları bu çalışmada, akciğer kanseri teşhisi için BT görüntülerinden oluşan veri kümesi üzerinde derin öğrenme tabanlı bir sınıflandırma modeli geliştirmiş ve çeşitli optimizasyon algoritmalarıyla modelin performansını artırmaya çalışmışlardır. 364 adet 512x512 boyutundaki görüntü, Python Augmentor ile 2000'e çıkarılmıştır. Özellik seçimi için ACO, PSO ve Knapsack; hiperparametre ayarı için PSO, NSGA-II ve Nelder-Mead Simplex algoritmaları kullanılmıştır. Sınıflandırmada transfer öğrenmeye dayalı ResNet50 mimarisi tercih edilmiş, farklı yapılandırmalar karşılaştırılmıştır. Temel model %90.45 doğruluk elde ederken, ACO ve PSO ile optimize edilen model %98.83 doğruluk ve %98.84 F1 skoru gibi yüksek performans sergilemiştir. Bulgular, bu optimizasyon

tekniklerinin küçük veri kümelerinde dahi sınıflandırma başarımını anlamlı şekilde artırdığını göstermektedir [10].

Ayrıca, akciğer kanseri tespitine yönelik yapılan bir diğer çalışmada, derin öğrenme modellerinin başarılı sonuçlar verdiği gözlemlenmiştir. Ayayna ve arkadaşları, IQOTH/NCCD veri setini kullanarak VGG16 sinir ağı modelini uygulamış ve modelin doğruluğunu %94.19'a çıkarmıştır [11]. Akciğer kanserinin erken evrelerinde doğru tanı koymak için kullanılan transfer öğrenme ve CNN mimarileri, bu tür hastalıkların tedavisinde önemli bir yer tutmaktadır. Aynı şekilde, EfficientNet-B1 ve Vision Transformer (ViT) gibi dönüşüm tabanlı derin öğrenme modelleri, LC25000 veri seti üzerinde yapılan çalışmalarda sırasıyla %96.97 ve %98.4 doğruluk oranlarına ulaşmış ve klinik uygulamalarda güvenilirliklerini artırmıştır [12].

Yapılan literatür incelemesi, derin öğrenme tabanlı modellerin akciğer kanseri tespitinde yüksek doğruluk ve güvenilirlik sunduğunu göstermektedir. Özellikle ViT tabanlı ve hibrit mimariler, geleneksel yöntemlere kıyasla daha başarılı sonuçlar elde etmiştir. Ayrıca, veri artırımı ve optimizasyon tekniklerinin model performansını belirgin biçimde artırdığı görülmektedir. Tablo 1'de de görüldüğü üzere, farklı mimariler çeşitli veri setlerinde %98'in üzerinde doğruluk oranlarına ulaşarak klinik uygulamalar açısından umut verici sonuçlar ortaya koymuştur.

Tablo 1 Literatür Taraması Özet Tablosu

Yazar	Yıl	Yöntem / Yaklaşım	Veri Seti	Performans Metrikleri	Sonuçlar
Siddiqui ve ark.	2025	ViT tabanlı modeller (CoAtNet, LeViT, Nested-TNT, CrossViT, HorNet)	IQ-OTH/NCCD, LC25000	LeViT accuracy	%89.43
				F1 score	%88.44
				CoAtNet accuracy	%89.10
				PR AUC	%87.27
Hassan ve ark. (2025)	2025	EfficientNet-B0	IQ-OTH/NCCD + ek veri seti	Accuracy	%89.32
				Precision	%88.20
				Recall	%87.80
Adams ve ark. (2018)	2018	SVM + Haralick doku özellikleri	SPIE-AAPM Lung CT Challenge	Strateji 1 recall	%80.82
				Strateji 1 precision	%34.50
				Strateji 2 recall	%60.67
				Strateji 2 precision	%74.58
				F-measure	0.6666
Thambi ve ark.	2025	ResNet50 + ACO, PSO, NSGA-II, Nelder-Mead	CT scan images (364 → 2000)	Temel model accuracy	%80.45
				ACO-PSO optimizasyonlu model accuracy	%88.83
				F1 score	%88.84
Ayayna ve ark.	2024	VGG16, EfficientNet-B1, Vision Transformer (ViT)	IQOTH/NCCD, LC25000	VGG16 accuracy	%84.19
				EfficientNet-B1 accuracy	%86.97
				ViT accuracy	%88.4

3. MATERYAL VE YÖNTEM

3.1 Veri Setleri

3.1.1 Veri Ön İşlemleri

Bu çalışmada, derin öğrenme tabanlı sınıflandırma modellerinin performanslarını güvenilir ve genellenebilir bir şekilde değerlendirebilmek amacıyla Stratified K-Fold çapraz doğrulama yöntemi uygulanmıştır. Bu yöntem, veri kümesini k adet alt kümeye ayırarak her bir alt kümenin hem eğitim hem de test verisi olarak dönüşümlü kullanılması esasına dayanır. Bu sayede modelin, farklı veri bölünmeleri üzerinde test edilerek daha sağlam ve ön yargısız bir performans değerlendirmesi yapılması sağlanır [13]. Çalışmada $k = 5$ olarak belirlenmiş ve veri kümesi, sınıf dağılımı korunacak şekilde beş katmana ayrılmıştır.

Ayrıca, rastlantısal veri dağılımındaki tutarlılığı sağlamak ve deneysel sonuçları tekrarlanabilir kılmak adına “random_state=42” kullanılarak sabit bir rastgelelik tohumu belirlenmiştir. Bu yaklaşım, modelin her çalıştırmada aynı veri dağılımı ile eğitilmesine olanak tanıyarak karşılaştırmalı analizlerin güvenilirliğini artırmıştır.

Bu çalışmada veriler üzerinde gerçekleştirilen önemli işlemlerden biri de veri artırımı (data augmentation) sürecidir. Eğitim verilerinin çeşitliliğini artırmak ve sınıflar arası dengesizlik problemini azaltmak amacıyla, görüntü işleme alanında yaygın olarak kullanılan Albumentations kütüphanesi aracılığıyla çeşitli dönüşüm teknikleri uygulanmıştır. Geliştirilen veri artırımı stratejisi, özellikle sınıf dağılımı dengesiz ve örnek sayısı sınırlı olan tıbbi görüntü veri kümelerinde, her sınıfın temsil oranını artırmaya ve modelin genelleme performansını iyileştirmeye yönelik olarak tasarlanmıştır.

Veri kümesi, 5 katlı çapraz doğrulama (5-fold cross-validation) yöntemiyle organize edilmiş olup, her bir fold için eğitim ve test alt klasörleri ayrı ayrı işlenmiştir. Test verileri, model değerlendirmesinin nesnelliğini ve bütünlüğünü korumak amacıyla hiçbir dönüşüm uygulanmadan yalnızca kopyalanarak sonuç klasörlerine aktarılmıştır. Böylece modelin sadece eğitim verisi üzerinde gerçekleştirilen veri artırımı etkilerini öğrenmesi ve test verilerinde yapay ön yargılardan etkilenmemesi sağlanmıştır.

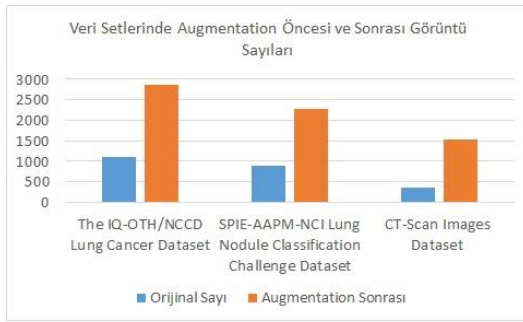
Eğitim verilerine uygulanan veri artırımı sürecinde, sentetik örnekler oluşturularak modelin karşılaşılabileceği çeşitli senaryolar test edilmiştir.

Bu kapsamda, Şekil 1’de görüldüğü gibi görüntülere $\pm 15^\circ$ aralığında rastgele döndürme, %5 oranında kaydırma, %10 oranında ölçekleme ve sınırlı açılarda dönme işlemleri uygulanmıştır. Ayrıca, görüntülerin parlaklık ve kontrast seviyeleri rastgele değiştirilerek farklı ışık koşulları altında modelin öğrenme kapasitesi artırılmıştır. Modelin gürültüye karşı dayanıklılığını artırmak amacıyla Gauss bulanıklığı ve Gauss gürültüsü eklenmiş; görüntülerin geometrik yapısına elastik bozulmalar (elastic deformations) uygulanarak örnek çeşitliliği artırılmıştır.

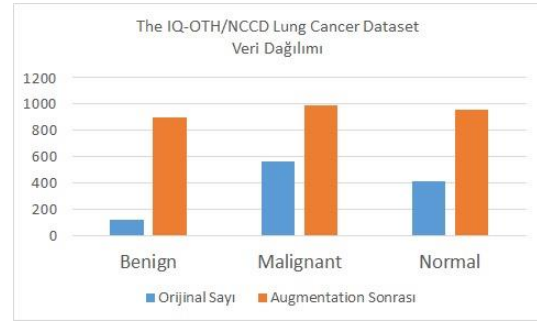
```
transform = Compose([
    Rotate(limit=15, p=0.5),
    ShiftScaleRotate(shift_limit=0.05, scale_limit=0.1, rotate_limit=0, p=0.5),
    RandomBrightnessContrast(p=0.3),
    GaussianBlur(blur_limit=(3, 5), p=0.2),
    GaussNoise(std_range=(0.04, 0.05), mean_range=(0.0, 0.0), p=0.2),
    ElasticTransform(alpha=1, sigma=50, p=0.1),
])
```

Şekil 1 Veri Artırımı Teknikleri Kod Bloğu

Uygulanan dönüşüm teknikleri titizlikle seçilmiş, özellikle insan anatomisine aykırı düşebilecek aşırı deformasyonlardan kaçınılmıştır. Bu sayede, oluşturulan sentetik verilerin gerçek biyolojik yapılara uygunluğu korunmuş ve modelin anlamsız yapay örnekleri öğrenmesi önlenmiştir.

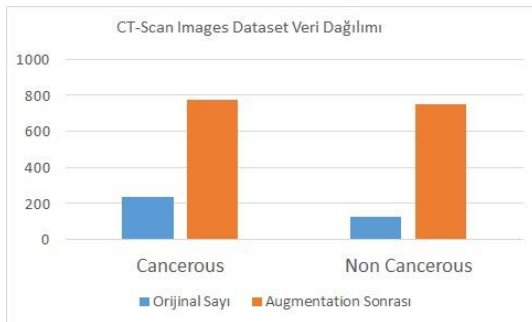


Şekil 2 Veri Setlerinde Veri Artırımı

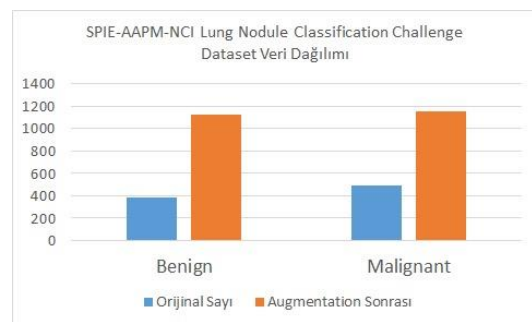


Şekil 3 IQ-QTH/NCCD Lung Cancer

Öncesi ve Sonrasında Toplam Sayıları Datasetinin Veri Artırımı Öncesi ve Sonrası Sınıflara Göre Dağılımı



Şekil 4 CT-Scan Images Datasetinin Veri Artırımı Öncesi ve Sonrası Sınıflara Göre Dağılımı

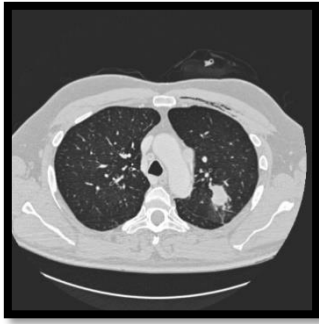


Şekil 5 SPIE-AAPM Lung Nodule Datasetinin Veri Artırımı Öncesi ve Sonrası Sınıflara Göre Dağılımı

Her bir sınıf için hedeflenen örnek sayısı, veri setinin orijinal boyutu dikkate alınarak belirlenmiştir. Şekil 2, Şekil 3 ve Şekil 4'te görüldüğü gibi IQOTH/NCCD ve SPIE-AAPM Lung Challenge veri setlerinde bu hedef, toplam orijinal görüntü sayısının yaklaşık üç katı; CT Scan Images veri setinde ise yaklaşık beş katı olarak tanımlanmıştır. Her sınıftaki eksik örnekler, bu hedef değerlere ulaşılan kadar sentetik olarak üretilmiştir.

Gerekli veri artırımı miktarı, her fold içindeki mevcut sınıf dağılımına orantılı olarak hesaplanmış ve veri artırımı bu oranlara uygun biçimde gerçekleştirilmiştir. Böylece, hem sınıflar arası dengeli bir örnek dağılımı sağlanmış hem de modelin sınıf ön yargısı geliştirmesinin önüne geçilmiştir.

Sürecin sonunda, her bir fold için sınıf bazında üretilen toplam örnek sayıları ayrı ayrı raporlanmış ve genel dağılım istatistikleri sunulmuştur. Bu işlem, veri artırımı sürecinin şeffaflığını ve tutarlılığını belgeleyen bir kontrol mekanizması olarak işlev görmektedir.



Şekil 6 Orijinal
Görüntüye Brightness
İşlemi Uygulanmış Hali



Şekil 7 Orijinal Görüntüye
Döndürme İşlemi
Uygulanmış Hali



Şekil 8 Orijinal Görüntü

3.1.2 The IQ-OTH/NCCD Lung Cancer Dataset

Bu çalışmada, Hamdalla F. Al-Yasriy tarafından Kaggle'da yayımlanan IQ-OTH/NCCD akciğer kanseri veri seti kullanılmıştır. Bu veri seti, 2019 yılının sonbahar döneminde, Irak'taki Iraq-Oncology Teaching Hospital ve National Center for Cancer Diseases (IQ-OTH/NCCD) hastanelerinde toplanmış olup farklı evrelerde akciğer kanseri tanısı konmuş hastalara ait bilgisayarlı tomografi görüntüleri ile sağlıklı bireylerin görüntülerini içermektedir. Veri setinde toplamda 110 vakayı temsil eden 1097 BT tarama dilimi

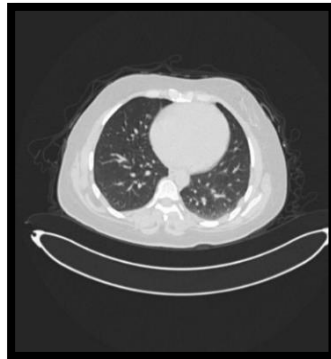
bulunmaktadır. Her bir tarama, deęişen sayılarda dilim içermekte olup, her bir dilim, insan göğsünün farklı yön ve açılardan alınan bir kesitini temsil etmektedir.

Çalışmaya dahil edilen 110 vakada, cinsiyet, yaş, eğitim durumu, ikamet edilen bölge ve yaşam durumu gibi demografik özellikler açısından çeşitlilik bulunmaktadır. Veri setinde bulunan vakalar üç sınıfa ayrılmaktadır: Normal, Benign (iyi huylu) ve Malign (kötü huylu). Veri seti vakaların sınıfına göre 3 farklı klasörde toplanmıştır. Bu klasörlerde 40 adet Malign vakasına ait 561, 15 adet Benign vakasına ait 120 ve 55 adet Normal vakaya ait 416 görüntü yer almaktadır. İlgili sınıfların görsel özelliklerini temsil edecek şekilde Şekil 9, Şekil 10 ve Şekil 11'de sırasıyla benign, normal ve malign vakalarına ait örnek görüntüler sunulmuştur. BT tarama görüntüleri, Siemens markalı SOMATOM cihazı kullanılarak elde edilmiştir ve DICOM formatında saklanmıştır. Taramalar, nefes tutarak tam inspirasyon anında yapılmıştır. Tüm görüntüler, kimlik bilgisi gizliliğini koruyacak şekilde anonimleştirilmiştir ve bu süreç, katılımcı kurumların etik kurulları tarafından onaylanmıştır [14].



Şekil 9 Benign Case (16)

[14]



Şekil 10 Malignant (165)

[14]



Şekil 11 Normal Case (32)

[14]

3.1.3 SPIE-AAPM-NCI Lung Nodule Classification Challenge Dataset

Çalışma kapsamında kullanılan bu veri seti, 2015 yılında SPIE Medical Imaging Konferansı bünyesinde düzenlenen LUNGx Challenge için oluşturulan SPIE-AAPM-NCI Lung Nodule Classification Challenge Dataset'idir. Yarışma, Amerikan Tıp Fizikçileri Derneği (AAPM) ve Ulusal Kanser Enstitüsü (NCI) desteğiyle gerçekleştirilmiştir. Temel amacı, akciğer nodüllerini benign (iyi huylu) ve malignant (kötü huylu) olarak sınıflandırmaya yönelik geliştirilen görüntü analiz algoritmalarının karşılaştırmalı olarak

değerlendirilmesidir.

Veri seti, “CT-Training” ile adlandırılan 10 etiketli eğitim hastası ve “LUNGx” ile başlayan 60 etiketsiz test hastasından oluşmaktadır. Test verilerine ait sınıf etiketleri başlangıçta paylaşılmamış, ancak yarışma sonrasında nodül konumlarını ve tanı bilgilerini içeren bir Excel dosyası yayımlanarak araştırmacıların erişimine sunulmuştur [15].

Bu çalışmada, bu etiket dosyaları kullanılarak hem eğitim hem de test verileri yeniden yapılandırılmış, nodüller benign ve malignant olmak üzere iki sınıfa ayrılmıştır. Bu doğrultuda, 35 benign ve 31 malignant hastaya ait klasörler oluşturulmuş; her hastaya ait farklı BT kesitlerinden elde edilen yaklaşık 21.178 DICOM görüntüsü ile dengeli ve sınıflandırmaya uygun bir veri kümesi hazırlanmıştır.

Veri işlemenin bir sonraki aşamasında, DICOM formatındaki bu görüntülerin daha yaygın analiz edilebilir ve görselleştirilebilir hale gelmesi amacıyla tüm görüntüler .png formatına dönüştürülmüştür. Bu işlem sırasında geliştirilen Python kod bloğu, kaynak klasörlerdeki dosya hiyerarşisini koruyarak DICOM dosyalarını taramış, piksel değerlerini normalize ederek 0–255 aralığına dönüştürmüş ve görüntüleri gri tonlamalı .png formatında hedef klasöre kaydetmiştir. Böylece, orijinal klasör yapısını koruyarak dicom_png_output klasöründe görsel bir veri seti elde edilmiştir.

Görüntü işleme aşamasında, DICOM’dan dönüştürülen .png görüntüler üzerinden yalnızca nodül içeren kesitlerin otomatik olarak seçilmesini amaçlayan bir yöntem uygulanmıştır. Nodüller her zaman aynı dilimde belirgin şekilde görünmeyebileceğinden, yalnızca merkez nodül dilimi değil, aynı zamanda bu merkezin çevresindeki benzer dilimlerin de seçilmesine olanak tanıyan benzerlik tabanlı bir dilim seçme algoritması geliştirilmiştir.

Etiket dosyasında yer alan merkez dilim bilgileri kullanılarak her hasta için merkez nodül görüntüsü tespit edilmiş, bu merkez dilimin hem öncesindeki hem de sonrasındaki dilimler incelenmiştir. Her bir çevre dilim, merkez görüntüyle karşılaştırılmış ve benzerlik ölçümleri kullanılarak uygun dilimler seçilmiştir. Bu benzerlik değerlendirmesinde iki metrik kullanılmıştır:

1. Structural Similarity Index (SSIM): Görüntülerin yapısal benzerliğini değerlendirerek detay ve kontrast açısından yakın olan dilimlerin tespit edilmesini sağlar.
2. Histogram Bhattacharyya Mesafesi: Görüntü histogramları arasındaki benzerliği ölçerek yoğunluk dağılımındaki farkları analiz eder.

Bir çevre dilimin SSIM değeri belirlenen eşik olan 0.6'nın altındaysa veya histogram farkı 0.3'ün üzerindeyse, ilgili yönde seçim işlemi sonlandırılmıştır.

Böylece yalnızca merkeze yapısal olarak benzer olan dilimler seçilerek, nodülün net biçimde yer aldığı bir alt görüntü kümesi oluşturulmuştur.

Bu süreçte, dosya adlarında yer alan sayısal indekslerden yararlanılarak her görüntüye benzersiz bir sıra numarası atanmış ve merkez-çevre dilim ilişkisi düzenli biçimde korunmuştur. Seçilen dilimler, orijinal hasta klasör yapısı korunarak ayrı bir hedef klasöre kopyalanmış ve yalnızca nodül içeren kesitlerden oluşan bir alt veri kümesi klasörü oluşturulmuştur.

Tüm bu işlemler sonucunda oluşturulan alt görüntü kümesinde, benign (iyi huylu) sınıfa ait 384 ve malignant (kötü huylu) sınıfa ait 492 olmak üzere toplam 876 adet .png formatında görüntü yer almaktadır. İlgili sınıfların görsel özelliklerini temsil edecek şekilde Şekil 12 ve Şekil 13'te sırasıyla benign ve malignant vakalarına ait örnek görüntüler sunulmuştur.

3.1.4 CT-Scan Images Dataset

Bu çalışmada, Kaggle üzerinden temin edilen ve 238 adet cancerous ile 126 adet non-cancerous olmak üzere iki sınıfa ayrılan toplam 364 adet bilgisayarlı tomografi görüntüsünden oluşan CT-Scan Images veri seti kullanılmıştır [16]. İlgili sınıfların görsel özelliklerini temsil edecek şekilde Şekil 14 ve Şekil 15'te sırasıyla cancerous ve non-cancerous vakalarına ait örnek görüntüler sunulmuştur.



Şekil 12 CT-048_1-117 (malignant) [15]



Şekil 13 CT-060_1-132 (benign) [15]

Bu veri kümesi üzerinde çeşitli ön işleme adımları uygulanmıştır. Bu sayede görüntülerin derin öğrenme tabanlı sınıflandırma modellerine daha uygun hâle getirilmesi amaçlanmıştır.

İlk adımda, göğüs kafesinin merkezini belirlemek amacıyla görüntülere binary threshold uygulanarak konturlar tespit edilmiş; bu konturlar arasından en büyük olanın moment hesaplamaları yapılmış ve böylece göğüs merkezinin koordinatları elde edilmiştir.

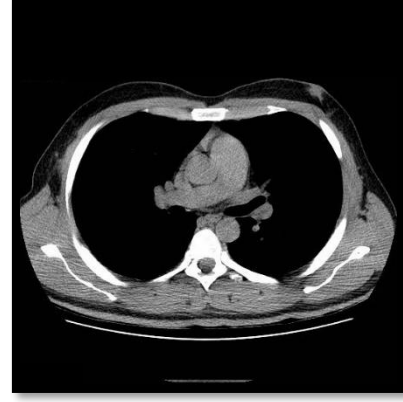
Ardından, bu merkez noktası referans alınarak her bir görüntü, 512x512 piksel boyutlarında yeni bir tuval üzerine ortalanarak yerleştirilmiş ve normalize edilmiş bir görsel elde edilmiştir. Devamında, yeniden threshold uygulanarak en büyük kontur belirlenmiş ve sabit koordinatlara (üst sol, üst sağ ve alt sol köşeler) göre siyah dikdörtgenlerle manuel maskeleme yapılmıştır. Bu maskeleme sonrası, belirli bir marjinle birlikte ilgili bölge kırılmış ve kırılan alan tekrar 512x512 boyutundaki bir tuval üzerine ortalanarak yerleştirilmiştir. Tüm bu işlemler, görüntülerin hizalanmış, normalize edilmiş ve istenmeyen alanlardan arındırılmış şekilde elde edilmesini sağlamıştır.

Çalışmanın bir diğer aşamasında ise, özellikle kanserli doku içeren BT görüntülerinin kontrast seviyelerinin iyileştirilmesi ve görüntüler arasında tutarlılığın artırılması hedeflenmiştir. Bu amaçla, işlenecek görüntülerin bulunduğu dizin belirlenmiş ve çıktıların kaydedileceği bir hedef klasör oluşturulmuştur. Görsellerin iyileştirilmesi aşamasında, lokal kontrastı artırmak için Contrast Limited Adaptive Histogram Equalization (CLAHE) yöntemi uygulanmıştır. CLAHE, medikal görüntülerde detayların daha belirgin hâle getirilmesinde etkili bir yöntemdir [17]. Süreçte, siyah-beyaz bir görüntü referans olarak seçilmiş ve diğer tüm görüntüler, bu referans görüntüye histogram eşleme (histogram matching) tekniği ile normalize edilmiştir. Bu adım, görüntüler arasındaki aydınlatma ve kontrast farklarını azaltarak sınıflandırma modellerinin başarımını artırmayı amaçlamaktadır. Histogram eşlemesinin ardından, görüntüler uint8 formatına dönüştürülmüş ve CLAHE yöntemiyle nihai kontrast iyileştirmesi uygulanmıştır.

Sonuç olarak, her bir BT görüntüsü hem global histogram özellikleri hem de lokal kontrast açısından optimize edilerek çıktı klasörüne kaydedilmiştir.



Şekil 14 Cancerous108 [16]



Şekil 15 Non-Cancerous85 [16]

3.2 Deney Ortamı Özellikleri ve Kullanılan Kütüphaneler

3.2.1 Deney Ortamı ve Makine Özellikleri

Bu çalışmada kullanılan bilgisayar altyapısı, yüksek işlem gücü ve geniş bellek kapasitesi sayesinde veri işleme, model eğitimi ve yazılım geliştirme süreçlerinin verimli bir şekilde yürütülmesine olanak tanımıştır. Tablo 2’de gözüktüğü gibi sistem, 4 fiziksel çekirdek ve 8 iş parçacığına sahip Intel Core i7-7700K @ 4.20GHz işlemci ile donatılmış olup, çoklu

görevlerin etkin bir biçimde gerçekleştirilmesini sağlamıştır. Toplam 32 GB DDR4 RAM, yüksek bellek gerektiren işlemler sırasında sistemin kararlı ve kesintisiz çalışmasına katkıda bulunmuştur.

Grafik işlem birimi (GPU) olarak kullanılan NVIDIA GeForce GTX 1080 Ti, 11 GB GDDR5X bellek kapasitesi ve güçlü paralel işlem yetenekleri ile derin öğrenme tabanlı modellerin eğitim sürecini önemli ölçüde hızlandırmıştır. Sistemde yer alan SSD (Solid State Drive), yüksek veri okuma/yazma hızları sayesinde genel sistem performansını artırmıştır.

Geliştirme ve izleme süreçlerinde uzak erişim için AnyDesk aracı tercih edilmiş, bu sayede sistem kaynaklarına farklı konumlardan güvenli erişim sağlanabilmiştir. İşletim sistemi olarak Windows 10 kullanılmış; sistem kaynakları, görev yöneticisi üzerinden düzenli aralıklarla izlenmiş ve gerek görüldüğünde optimizasyon işlemleri uygulanmıştır.

Tablo 2 Çalışma Ortamı Tablosu

İşlemci	Intel Core i7-7700K @ 4.20GHz (4 çekirdek, 8 thread)
Bellek	32 GB DDR4 RAM
Ekran Kartı	NVIDIA GeForce GTX 1080 Ti (11 GB)
Depolama	SSD
İşletim Sistemi	Windows 10
Uzaktan Erişim	AnyDesk ile bağlantı

3.2.2 Kullanılan Kütüphaneler

Bu çalışmada, yedi farklı derin öğrenme tabanlı görüntü sınıflandırma modeli kullanılarak deneysel analizler gerçekleştirilmiştir. Bu analizler kapsamında veri ön işleme, model eğitimi, değerlendirme ve görselleştirme işlemleri için Python programlama dilinde çeşitli açık kaynak kütüphaneler kullanılmıştır. Kullanılan bu kütüphaneler aşağıda açıklanmıştır.

3.2.2.1 NumPy

NumPy, Python programlama dili için geliştirilmiş temel bir dizi programlama kütüphanesidir. Çok boyutlu diziler, matrisler ve daha yüksek boyutlu veri yapıları üzerinde hızlı ve verimli işlemler yapmaya olanak tanır. Bilimsel verilerin düzenlenmesi, keşfedilmesi ve analiz edilmesi süreçlerinde temel bir araç olarak kullanılmaktadır [18].

3.2.2.2 Matplotlib

Matplotlib, Python programlama dili için geliştirilmiş, iki boyutlu grafiklerin oluşturulmasında yaygın olarak kullanılan kapsamlı bir görselleştirme kütüphanesidir. Etkileşimli betik yazımı, uygulama geliştirme ve yüksek kalitede grafik üretimi gibi çeşitli amaçlarla kullanılabilir [19].

3.2.2.3 Seaborn

Seaborn, Python programlama dili için geliştirilmiş istatistiksel veri görselleştirme kütüphanesidir. Matplotlib üzerine inşa edilmiş olan Seaborn, daha estetik, okunabilir ve anlamlı grafikler oluşturmayı kolaylaştırır. Veri kümeleri üzerinde istatistiksel ilişkileri keşfetmek amacıyla çizgi grafik, çubuk grafik, kutu grafik (boxplot), dağılım grafiği (scatter plot) ve ısı haritası (heatmap) gibi çeşitli görsel araçlar sunmaktadır [20].

3.2.2.4 Torch (PyTorch)

PyTorch, Python tabanlı açık kaynaklı bir derin öğrenme kütüphanesidir. Kullanıcıların model geliştirme sürecini hızlandırmak için sezgisel ve etkileşimli bir programlama ortamı sunar. Aynı zamanda GPU gibi donanım hızlandırıcılarını destekleyerek yüksek performanslı hesaplamalara olanak tanımaktadır [21].

3.2.2.5 Torchvision

Torchvision, PyTorch ile birlikte kullanılan, veri seti yönetimi, görüntü dönüşümleri ve hazır modeller sunan bir yardımcı kütüphanedir [22].

3.2.2.6 Scikit-learn

Scikit-learn, Python programlama dili için geliştirilmiş, orta ölçekli denetimli ve denetimsiz öğrenme problemlerini çözmeye yönelik kapsamlı bir makine öğrenimi kütüphanesidir. Sınıflandırma, regresyon, kümeleme, boyut indirgeme ve model seçimi gibi birçok temel makine öğrenimi algoritmasını bünyesinde barındırır [23].

3.2.2.7 Timm

Timm (PyTorch Image Models), Ross Wightman tarafından geliştirilmiş açık kaynaklı bir kütüphane olup, EfficientNet ve ConvNeXt gibi güncel görsel modellerin kolayca kullanılmasını sağlamakta ve ImageNet sonuçlarının yeniden üretilebilmesine olanak tanımaktadır [24].

3.2.2.8 OpenCV

OpenCV (Open Source Computer Vision Library), bilgisayarla görme uygulamaları için geliştirilmiş açık kaynaklı bir kütüphanedir. OpenCV, görüntü işleme ve bilgisayarla görme alanlarında geniş bir fonksiyon yelpazesi sunar. Bu fonksiyonlar arasında görüntü filtreleme, kenar bulma, nesne takibi, yüz tanıma, kamera kalibrasyonu, hareket takibi ve üç boyutlu görme gibi uygulamalar yer alır. Kütüphane, kullanıcıların görüntü ve video verileriyle çalışmasını kolaylaştırmak amacıyla tasarlanmış olup, akademik ve endüstriyel projelerde sıklıkla tercih edilmektedir [25].

3.2.2.9 Pandas

Pandas, Python programlama dili için geliştirilmiş, özellikle veri analizi ve istatistiksel işlemler için güçlü araçlar sunan bir açık kaynak kütüphanedir. Bu kütüphane, veri yapılarını ve veri manipülasyonu işlevlerini yüksek düzeyde bir soyutlamayla sunarak Python’u istatistiksel hesaplamalar ve bilimsel çalışmalar için daha pratik hale getirmeyi amaçlar [26].

3.2.2.10 Glob

Glob modülü, dosya sistemi içerisinde belirli desenlere uyan dosya ve klasör yollarını listelemek için kullanılan bir Python kütüphanesidir. Bu modül, özellikle dosya adlarını joker karakterler örneğin *, ?, ve [] kullanarak aramak isteyenler için oldukça faydalıdır [27].

3.2.2.11 Tqdm

Tqdm, Python programlarında özellikle uzun süren yinelemeli işlemlerde ilerlemenin görsel olarak takip edilmesini sağlayan bir ilerleme çubuğu kütüphanesidir. İşlem sırasında kaçınıcı adıma gelindiğini, saniyede kaç adım tamamlandığını, geçen süreyi ve tahmini bitiş zamanını tek bir satırda dinamik olarak güncelleyerek kullanıcıya sunar [28].

3.2.2.12 Transformers

Transformers kütüphanesi, Hugging Face tarafından geliştirilmiş olup, Transformer tabanlı derin öğrenme modellerinin uygulanması ve eğitilmesi için tasarlanmış açık kaynaklı bir yazılım kütüphanesidir. Kütüphane, BERT, BEiT gibi önceden eğitilmiş modern derin öğrenme modellerinin kolayca erişilebilir, yeniden kullanılabilir ve dağıtılabilir olmasını sağlamaktadır [29].

3.2.2.13 Safetensors

Safetensors, derin öğrenme modellerinin ağırlıklarını güvenli ve hızlı bir şekilde saklamak ve yüklemek amacıyla geliştirilmiş bir dosya formatı ve kütüphanedir. Geleneksel olarak PyTorch gibi kütüphanelerde kullanılan pickle tabanlı model kaydetme yöntemlerinin aksine, safetensors formatı dışarıdan gelebilecek zararlı kodların çalıştırılmasını engeller ve böylece dosya güvenliğini artırır [30].

3.2.2.14 Torch.cuda.amp

Torch.cuda.amp, PyTorch tarafından sağlanan ve karışık hassasiyetli (mixed precision) eğitim ve çıkarım işlemlerini kolaylaştıran bir modüldür. Bu modül sayesinde, model eğitimi sırasında bazı işlemler daha düşük hassasiyetli veri tiplerinde gerçekleştirilirken, diğer işlemler ise daha yüksek dinamik aralığa sahip veri tipi ile yapılabilir [31].

3.2.2.15 Torch.nn.functional

Torch.nn.functional modülü, sinir ağı katmanlarının işlevsel biçimde uygulanmasına olanak tanımaktadır. Bu modülde yer alan fonksiyonlar, herhangi bir öğrenilebilir parametre (örneğin ağırlık veya bias) içermeyen stateless yapılar olup, modelin eğitim sürecinde güncellenen bir duruma sahip değildir. Konvolüsyon, aktivasyon ve havuzlama gibi temel işlemler, bu modül aracılığıyla katman örneği oluşturmaya gerek kalmaksızın doğrudan gerçekleştirilebilmektedir [32].

3.2.2.16 Torch.multiprocessing

Torch.multiprocessing, Python'un standart çoklu işlem modülünün üzerine inşa edilmiş bir PyTorch sarmalayıcısıdır. Bu modül, çoklu işlem desteği sağlayarak derin öğrenme modellerinin özellikle veri paralelleştirmesiyle daha verimli bir şekilde çalıştırılmasına olanak tanır. Temel farkı, PyTorch tensörlerini ve verilerini işlemler arasında kopyalama yapmadan paylaşabilmesidir [33].

3.2.2.17 JSON & Pickle

JSON ve Pickle, Python'da veri serileştirme ve kalıcı depolama işlemleri için yaygın olarak kullanılan iki farklı kütüphanedir. JSON (JavaScript Object Notation), özellikle insanlar tarafından okunabilir yapısıyla veri değişimi için tercih edilen hafif bir formattır [34].

Pickle ise Python'a özgü bir serileştirme yöntemidir ve nesne yapılarının ikili formata dönüştürülerek diske yazılmasını sağlar. Bu süreçte Python nesneleri bayt dizilerine çevrilir ve gerektiğinde bu bayt dizilerinden tekrar orijinal nesneler elde edilebilir [35].

3.2.2.18 Albumentations

Albumentations, görüntü artırma işlemleri için geliştirilmiş güçlü ve esnek bir kütüphanedir. Görüntü sınıflandırma, nesne tespiti ve segmentasyon gibi birçok bilgisayarlı görü görevinde kullanılabilir. Açık kaynaklı olan bu kütüphane, GPU üzerinde hızlı çalışacak şekilde optimize edilmiştir ve PyTorch ile kolayca entegre edilebilir [36].

3.2.2.19 Time

Time modülü, Python programlama dilinde zamanla ilgili işlemleri gerçekleştirmek için kullanılan yerleşik bir kütüphanedir. Bu modül, özellikle zaman damgaları (timestamp), süre ölçümleri ve gecikme işlemleri gibi zaman bazlı görevlerde yaygın olarak kullanılmaktadır [37].

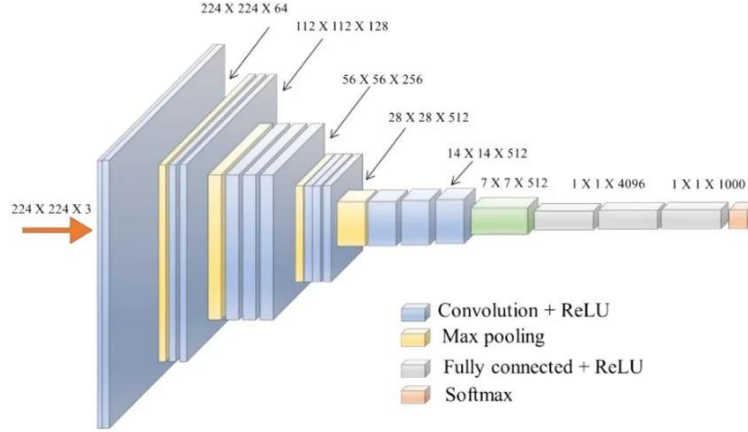
3.3 Metotlar

Bu çalışmada, akciğer kanseri tespiti için farklı derin öğrenme yöntemleri ve Transformer tabanlı modeller kullanılmıştır. Kullanılan veri setindeki görüntüler belirli ön işleme adımlarından geçirilerek sınıflandırma modellerine uygun hale getirilmiştir.

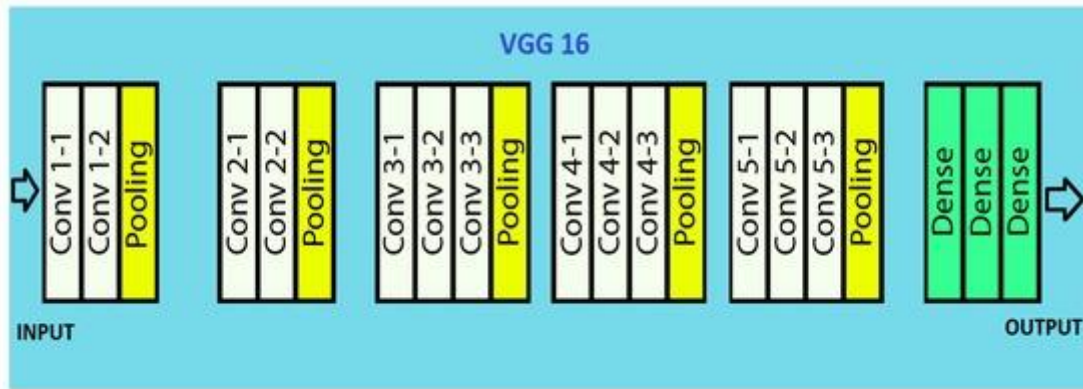
3.3.1 VGG16

Bu çalışmada kullanılan derin öğrenme yöntemlerinden biri, Oxford Visual Geometry Group tarafından geliştirilen ve görsel tanıma görevlerinde yüksek başarı gösteren VGG16 mimarisidir. VGG16, derinliği artırılmış ve katmanlar arası ağırlık paylaşımı ile optimize edilmiş bir evrimsel sinir ağı modelidir. Mimari, 3×3 boyutundaki küçük konvolüsyon filtrelerini çok katmanlı bir şekilde istifleyerek daha karmaşık özelliklerin hiyerarşik biçimde öğrenilmesini sağlamaktadır. Her konvolüsyon katmanından sonra ReLU (Rectified Linear Unit) aktivasyon fonksiyonu kullanılmakta olup, belirli katmanlarda

maksimum havuzlama (MaxPooling) işlemleri ile boyutsal azalma sağlanmaktadır. Modelin son kısmında yer alan tam bağlantılı (fully connected) katmanlar, sınıflandırma görevini gerçekleştirmektedir [38]. Söz konusu mimarinin katmanları ve genel yapısı Şekil 16 ve Şekil 17’de gösterilmiştir.



Şekil 16 VGG 16 Model Mimarisi [39]



Şekil 17 VGG 16 Katmanları [40]

Çalışma kapsamında, VGG16 modeli ImageNet veri kümesi üzerinde önceden eğitilmiş ağırlıklarla kullanılmış ve transfer öğrenme yaklaşımı benimsenmiştir. Özellikle modelin "feature extraction" katmanları (özellik çıkarımı) dondurularak yalnızca sınıflandırma katmanları eğitime uygun hale getirilmiştir. Böylece, önceden öğrenilmiş düşük seviyeli özelliklerin korunması sağlanırken, veri kümesine özgü yüksek seviyeli özelliklerin öğrenilmesi sağlanılmıştır. Bu kapsamda oluşturulan model mimarisi, Şekil 18’de gösterilmiştir.

Veri setinde yer alan tüm BT görüntüleri, modelle uyumlu hale getirilmek üzere öncelikle 224×224 piksel boyutuna yeniden ölçeklendirilmiştir. Ardından, önceden eğitilmiş

modellerin beklediği giriş formatına uygun biçimde normalize edilerek tensör yapısına dönüştürülmüştür. Görüntüler RGB formatında olup, kanal başına ortalama ve standart sapma değerleri sırasıyla [0.485, 0.456, 0.406] ve [0.229, 0.224, 0.225] olacak şekilde normalize edilmiştir. Veri seti, beş katlı çapraz doğrulama (5-Fold Cross Validation) yöntemiyle değerlendirilmiş, her kat için ayrı eğitim ve test işlemleri gerçekleştirilmiştir.

```
model = models.vgg16(pretrained=True)
for param in model.features.parameters():
    param.requires_grad = False

num_features = model.classifier[6].in_features
model.classifier[6] = nn.Linear(num_features, len(class_names))
model = model.to(device)

criterion = nn.CrossEntropyLoss()
optimizer = optim.Adamax(model.classifier.parameters(), lr=1e-4)

model, history, total_train_time, epoch_durations = train_model(model, dataloaders, criterion, optimizer,
                                                                num_epochs=EPOCHS, patience=PATIENCE)
total_all_training_time += total_train_time
```

Şekil 18 VGG16 Model Oluşturma Kod Bloğu

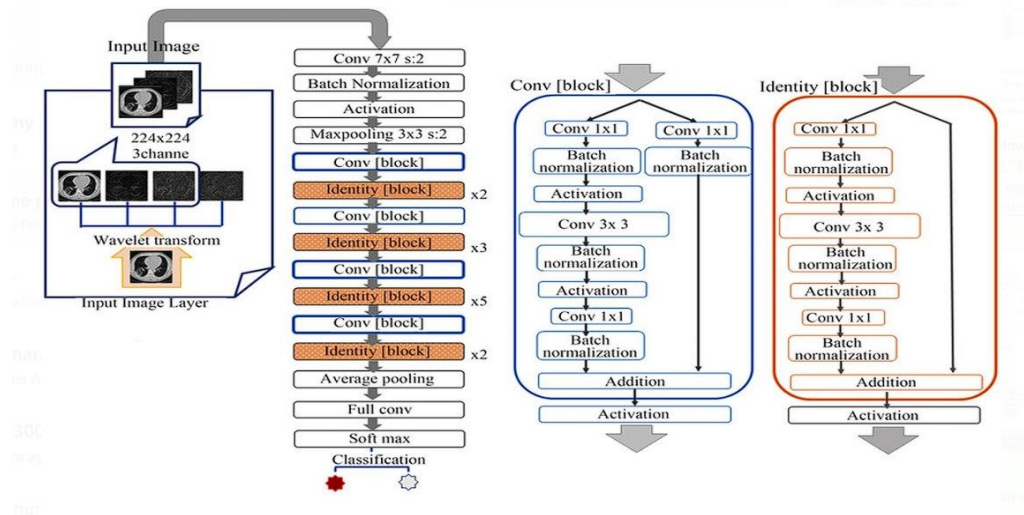
Modelin eğitiminde erken durdurma (early stopping) stratejisi uygulanmış, eğitim sürecinde doğrulama kaybı (validation loss) 5 epoch boyunca iyileşme göstermediğinde eğitim sonlandırılmıştır. Optimizasyon süreci için Adamax algoritması tercih edilmiş ve yalnızca sınıflandırma katmanlarının ağırlıkları güncellenmiştir. Kayıp fonksiyonu olarak çok sınıflı çapraz entropi (Cross Entropy Loss) kullanılmıştır. Eğitim süreci 50 epoch boyunca gerçekleştirilmiş ve batch boyutu 32 olarak belirlenmiştir.

Eğitim ve doğrulama verilerinin performans göstergeleri, epoch bazlı olarak doğruluk (accuracy) ve kayıp (loss) değerleri üzerinden takip edilmiştir. Ayrıca, her folder sonunda modelin test seti üzerindeki sınıflandırma başarımları değerlendirilmek üzere saklanmış ve model ağırlıkları her fold için ayrı ayrı kaydedilmiştir. Görselleştirme amaçlı olarak her fold için confusion matrisi oluşturulmuş ve eğitim sürecine ilişkin kayıp ve doğruluk eğrileri grafiksel olarak sunulmuştur.

3.3.2 ResNet50

ResNet50, derin öğrenme alanında hem artık (residual) bağlantıları hem de geleneksel bağlantıları bir arada kullanan gelişmiş bir sinir ağı mimarisidir. Bu mimaride kullanılan artık bağlantılar, katmanların doğrudan karmaşık fonksiyonlar öğrenmesi yerine, giriş verilerine göre artık fonksiyonlar öğrenmesini sağlayarak, derin sinir ağlarının eğitim sürecini büyük ölçüde kolaylaştırmaktadır [41].

Model, beş temel aşamadan oluşmaktadır. İlk aşamada, 7x7 boyutunda 64 filtrelili bir konvolüsyon katmanını takiben 3x3 boyutunda ve 2 adımda uygulanan bir maksimum havuzlama işlemi yer alır. İkinci aşamada, sırasıyla 1x1 boyutlu 64 filtrelili, 3x3 boyutlu 64 filtrelili ve 1x1 boyutlu 256 filtrelili üçlü konvolüsyon blokları bulunur. Üçüncü aşamada ise bu yapı 128 ve 512 filtrelili konvolüsyon katmanları ile devam eder. Dördüncü aşamada 256 filtrelili 1x1 ve 3x3 konvolüsyon katmanları ile birlikte 1024 filtrelili bir 1x1 katman yer alır. Beşinci ve son aşamada ise modelde sırasıyla 512, 512 ve 2048 filtre içeren 1x1, 3x3 ve tekrar 1x1 boyutlu konvolüsyon katmanları yer almakta olup, ardından bir maksimum havuzlama işlemi uygulanır. Son olarak, 1000 nöronlu oluşan tam bağlantılı (fully connected) katman ile sınıflandırma gerçekleştirilir. Bu mimarinin örnek bir yapısı Şekil 19’da verilmiştir. ResNet50, özellikle görüntü sınıflandırma gibi görevlerde derin yapısına rağmen yüksek başarı oranları ile öne çıkmaktadır [42].



Şekil 19 ResNet50 Mimarisi [43]

Bu çalışmada, beş katlı çapraz doğrulama yöntemi kullanılarak modelin genellenebilirlik performansı değerlendirilmiştir. Bu kapsamda, önceden ImageNet veri kümesi üzerinde eğitilmiş olan ResNet50 mimarisi transfer öğrenme yöntemiyle yeniden yapılandırılmış; yalnızca tam bağlantılı (fully connected) katman güncellenerek sınıflandırma işlemi gerçekleştirilmiştir. Bu kapsamda oluşturulan model mimarisi, Şekil 20’de gösterilmiştir. Veri kümesindeki görüntüler standart boyutlara getirilerek tensör formatına dönüştürülmüştür.

```
model = models.resnet50(pretrained=True)
for param in model.parameters():
    param.requires_grad = False
model.fc = nn.Linear(model.fc.in_features, num_classes)
model = model.to(device)

criterion = nn.CrossEntropyLoss()
optimizer = optim.Adamax(model.fc.parameters(), lr=0.0001)

num_epochs = 50
train_losses, test_losses = [], []
train_accuracies, test_accuracies = [], []
epoch_times = []

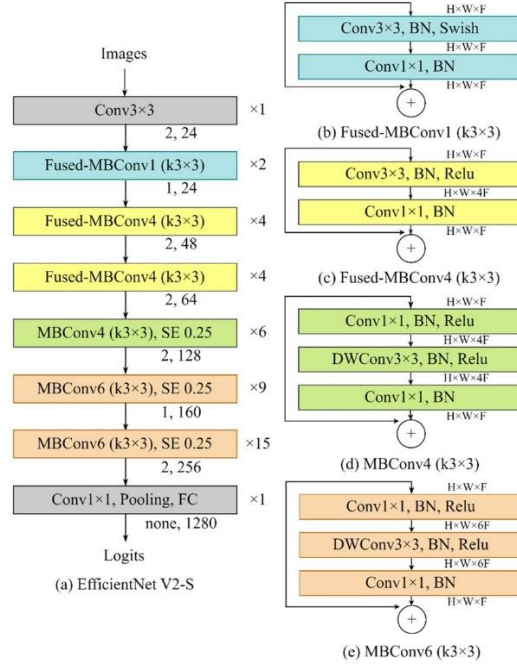
patience = 5
best_loss = float('inf')
epochs_no_improve = 0
```

Şekil 20 ResNet50 Model Oluşturma Kod Bloğu

Model eğitimi sürecinde kayıp fonksiyonu olarak çok sınıflı çapraz entropi, optimizasyon algoritması olarak ise Adamax tercih edilmiştir. Her bir folder’ın sonunda doğruluk, hassasiyet, duyarlılık ve F1 skoru gibi performans metrikleri hesaplanmış; confusion matrisi ve eğitim süreçlerine ilişkin grafiksel analizler görselleştirilmiştir. Ayrıca, her folder için elde edilen sınıflandırma sonuçları ve model ağırlıkları kayıt altına alınmıştır. Sonuç olarak, tüm folderlardan elde edilen metriklerin ortalaması alınarak modelin genel başarı düzeyi ortaya konmuş; geliştirilen sistemin farklı veri alt kümeleri üzerindeki tutarlılığı değerlendirilmiştir.

3.3.3 EfficientNetV2-S

EfficientNetV2 ailesi, parametrik verimlilik ile yüksek doğruluk oranlarını bir araya getiren mimari yapısıyla öne çıkmaktadır [44]. EfficientNetV2-S ise bu mimarinin daha kompakt ve hızlı eğitilebilir versiyonu olup, özellikle sınırlı veri ve hesaplama kapasitesine sahip tıbbi görüntüleme uygulamaları için ideal bir çözüm olarak değerlendirilmektedir [45].



Şekil 21 EfficientNetV2-S Mimarisi [46]

Bu çalışmada, EfficientNetV2-S modeli kullanılarak veri setleri üzerinde sınıflandırma işlemi gerçekleştirilmiş ve model performansı, beş katlı çapraz doğrulama yöntemiyle sistematik olarak değerlendirilmiştir. Veri seti, modele uygun hale getirilmek üzere EfficientNetV2-S mimarisi için optimize edilmiş ön işleme adımları uygulanmıştır. Bu kapsamda, EfficientNet_V2_S_Weights.DEFAULT ağırlıkları kullanılarak normalize edilmiş dönüşümler (transforms.Compose) uygulanmıştır.

Modelin mimarisinde, önceden eğitilmiş ağırlıklar ile başlatılan özellik çıkarıcı katmanlar dondurulmuş, yalnızca sınıflandırma katmanı güncellenmiştir. Bu katmanda, veri setinde yer alan sınıf sayısına uygun olacak şekilde tam bağlantılı bir çıkış katmanı yapılandırılmıştır. Modelin eğitimi, CrossEntropyLoss kayıp fonksiyonu ve Adamax optimizasyon algoritması kullanılarak gerçekleştirilmiştir. Aşırı öğrenmeyi engellemek amacıyla erken durdurma kriteri uygulanmıştır. Bu kapsamda oluşturulan model mimarisi, Şekil 22’de gösterilmiştir.

```
model = efficientnet_v2_s(weights=weights)
for param in model.features.parameters():
    param.requires_grad = False

in_features = model.classifier[1].in_features
model.classifier[1] = nn.Linear(in_features, len(class_names))
model = model.to(device)

criterion = nn.CrossEntropyLoss()
optimizer = optim.Adamax(model.classifier.parameters(), lr=1e-4)

model, history = train_model(model, dataloaders, criterion, optimizer, num_epochs=EPOCHS, patience=PATIENCE)
```

Şekil 22 EfficientnetV2-S Model Oluşturma Kod Bloğu

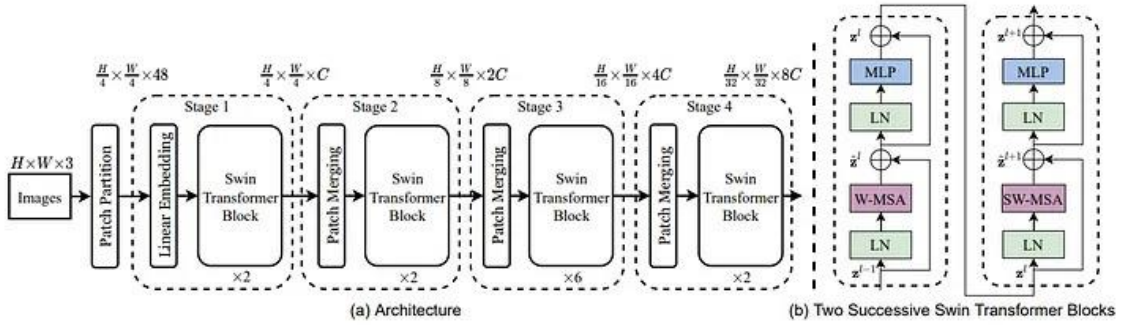
Her bir epoch sonunda eğitim ve doğrulama veri kümeleri üzerinden kayıp (loss) ve doğruluk (accuracy) değerleri kayıt altına alınmış; eğitim süreci boyunca elde edilen bu metrikler grafiksel olarak görselleştirilmiştir. Modelin test verisi üzerindeki performansı, doğruluk, hassasiyet (precision), geri çağırma (recall) ve F1-skoru gibi temel sınıflandırma metrikleriyle değerlendirilmiş; ayrıca sınıflandırma raporları ve confusion matrisleri her bir fold için ayrı ayrı hesaplanarak analiz edilmiştir.

Eğitim sonunda, her bir folder’da eğitilen modeller pth uzantılı dosyalar halinde kaydedilmiş, ilgili görselleştirmeler ve metrik çıktıları sistematik biçimde klasörlenmiştir. Beş farklı folder’a ait sınıflandırma çıktıları birleştirilerek genel bir değerlendirme yapılmış; tüm test sonuçları üzerinden hesaplanan birleşik confusion matrisi ile modelin genel genellenebilirlik performansı ortaya konmuştur.

3.3.4 SwinV2-CR-Small-224

SwinV2-CR-Small-224 modeli, Swin Transformer mimarisinin gelişmiş bir versiyonu olarak sunulmuş, görsel tanıma görevlerinde yüksek performans sağlamak üzere tasarlanmış bir derin öğrenme modelidir [47]. Bu model, "Shifted Window" (kaydırılmış pencere) mekanizmasını temel alarak, küresel dikkat (global attention) hesaplamalarının

karmaşıklığını azaltmakta ve görsel verilerdeki yerel ve küresel bağlamları etkili biçimde modelleyebilmektedir [48]. "CR" (Cosine Rescaled) ifadesi, modelin dikkat mekanizmasında kullanılan ağırlıkların yeniden ölçeklendirilmesi sürecini tanımlarken, 224 ibaresi modelin giriş görüntülerinin boyutunu (224x224 piksel) ifade etmektedir. SwinV2 mimarisi, geleneksel evrişimsel sinir ağlarının sınırlılıklarını aşmak amacıyla, çok ölçekli özellik çıkarımı ve hiyerarşik yapısı sayesinde, görsel öğeleri farklı çözünürlüklerde temsil ederek daha zengin bir bilgi öğrenme yetisi sunmaktadır. SwinV2-CR-Small-224, daha az parametreye sahip olmasına karşın, sınıflandırma, nesne tespiti ve segmentasyon gibi görevlerde yüksek doğruluk oranlarıyla öne çıkmaktadır. Bu mimarinin örnek bir yapısı Şekil 23'te verilmiştir.



Şekil 23 SwinV2-CR-Small-224 Mimarisi [49]

Bu model kullanılarak 5 katlı çapraz doğrulama yöntemi ile kapsamlı bir eğitim ve değerlendirme süreci gerçekleştirilmiştir. Eğitim ve test verileri, belirlenen klasör yapısına göre her bir fold için yüklenmiş ve normalize edilerek modele uygun hale getirilmiştir.

Model olarak, timm kütüphanesi üzerinden önceden eğitilmiş swinv2_cr_small_224 mimarisi seçilmiş ve çıktı katmanı, veri setindeki sınıf sayısına göre yeniden yapılandırılmıştır. Eğitim süreci boyunca optimizasyon için Adamax algoritması ve kayıp fonksiyonu olarak CrossEntropyLoss kullanılmış; erken durdurma mekanizması uygulanarak aşırı öğrenmenin önüne geçilmiştir. Bu kapsamda oluşturulan model mimarisi, Şekil 24'te gösterilmiştir. Her bir epoch sonunda hem eğitim hem de doğrulama verisi üzerindeki doğruluk ve kayıp değerleri kaydedilmiş; test verisi üzerinden ise precision, recall, F1-score ve accuracy gibi performans metrikleri hesaplanmıştır. Ayrıca, her fold için sınıflandırma raporu ve confusion matrisi görselleştirilerek kayıt altına alınmıştır. Eğitim sonunda her bir fold için eğitilen modeller pt uzantılı dosyalar halinde

kaydedilmiş ve tüm foldların çıktıları birleştirilerek ortalama performans metrikleri elde edilmiştir. Bu ortalama metrikler doğrultusunda modelin genellenebilirlik performansı değerlendirilmiş ve eğitim sürecinin genel kayıp ve doğruluk grafikleri görselleştirilmiştir.

```
# SwinV2 modelini oluşturun
model = timm.create_model('swinv2_cr_small_224', pretrained=True, num_classes=num_classes)
model.to(device)

criterion = nn.CrossEntropyLoss()
optimizer = optim.Adamax(model.parameters(), lr=1e-4)
num_epochs = 50
patience = 5

train_losses, test_losses = [], []
train_accuracies, test_accuracies = [], []
best_loss = float('inf')
epochs_no_improve = 0

epoch_times = [] # Bu fold için epoch süreleri

for epoch in range(num_epochs):
    epoch_start_time = time.time()

    model.train()
    train_loss, correct = 0, 0
    for images, labels in train_loader:
        images, labels = images.to(device), labels.to(device)
        optimizer.zero_grad()
        outputs = model(images)
        loss = criterion(outputs, labels)
        loss.backward()
        optimizer.step()
        train_loss += loss.item()
        correct += (outputs.argmax(1) == labels).sum().item()

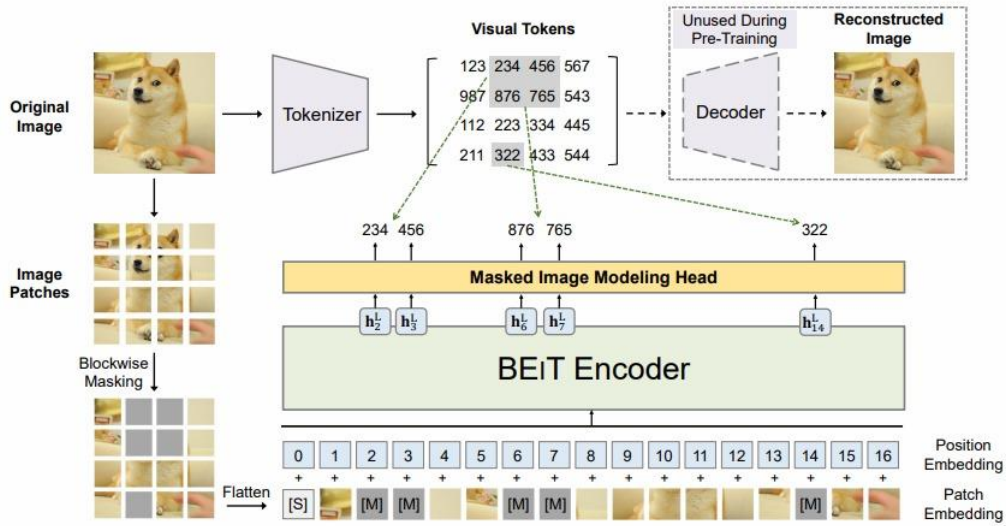
    epoch_end_time = time.time()
    epoch_times.append(epoch_end_time - epoch_start_time)
```

Şekil 24 SwinV2-CR-Small-224 Model Oluşturma Kod Bloğu

3.3.5. Beit-Base

Bu çalışmada, sınıflandırma görevini gerçekleştirmek amacıyla BEiT (Bidirectional Encoder representation from Image Transformers) tabanlı modelden de yararlanılmıştır. BEiT (Bidirectional Encoder representation from Image Transformers), doğal dil işleme alanında kullanılan BERT modelinden esinlenerek geliştirilmiş, kendi kendine denetimli bir görsel temsil modelidir [50]. Çift yönlü transformer kodlayıcı mimarisi sayesinde, görüntü yamalarının bağlamsal temsillerini etkili bir şekilde öğrenebilmektedir. Modelin ön eğitim sürecinde, BERT modelinin dil işleme alanındaki maskelenmiş dil modelleme

yaklaşımına benzer şekilde, bağlamsal bilgilere dayanarak maskelenmiş görsel sembollerin tahmin edilmesi hedeflenmiştir. Bu sayede, model etiketli veri ihtiyacını azaltarak kendi kendine denetimli (self-supervised) öğrenme yoluyla güçlü görsel temsiller edinme yeteneği kazanmıştır [51]. Modelin genel katman yapısı ve mimarisi Şekil 25’te sunulmuştur.



Şekil 25 BeiT-Base Modelinin Mimarisi [52]

Çalışmada, BEiT modelinin ImageNet veri kümesi üzerinde önceden eğitilmiş versiyonu kullanılmış ve böylece yüksek düzeyde genelleme yeteneği elde edilmiştir. Projeye özgü sınıf sayısına uyum sağlayabilmesi amacıyla modelin son katmanını yeniden yapılandırılmıştır. Bu kapsamda oluşturulan model mimarisi, Şekil 26’te gösterilmiştir.

Modelin eğitimi ve değerlendirilmesi aşamalarında 5 katlı çapraz doğrulama yöntemi kullanılmıştır. Her bir fold için eğitim ve test veri kümeleri tanımlanmış; görseller 224x224 piksel boyutlarına yeniden ölçeklendirilmiş ve BEiT modelinin giriş beklentilerine uygun olarak mean= [0.485, 0.456, 0.406] ve std= [0.229, 0.224, 0.225] değerleriyle normalize edilmiştir.

Eğitim sürecinde kayıp fonksiyonu olarak çapraz entropi kullanılmış; optimizasyon işlemleri ise Adamax algoritması ile gerçekleştirilmiştir. Her bir eğitim süreci 50 epoch üzerinden yürütülmüş olup, erken durdurma yöntemiyle gereksiz yinelemelerin önüne geçilmiştir. Her epoch sonunda eğitim ve doğrulama kayıpları ile doğruluk değerleri kaydedilmiştir.


```

# BEiT base modelini oluştur
model = timm.create_model('beit_base_patch16_224', pretrained=True, num_classes=num_classes)
model.to(device)

criterion = nn.CrossEntropyLoss()
optimizer = optim.Adamax(model.parameters(), lr=0.0001)
num_epochs = 1
patience = 5

train_losses, test_losses = [], []
train_accuracies, test_accuracies = [], []
best_loss = float('inf')
epochs_no_improve = 0
epoch_times = []

training_start_time = time.time() # Toplam eğitim süresi başlangıcı

```

Şekil 26 Beit-Base Model Oluşturma Kod Bloğu

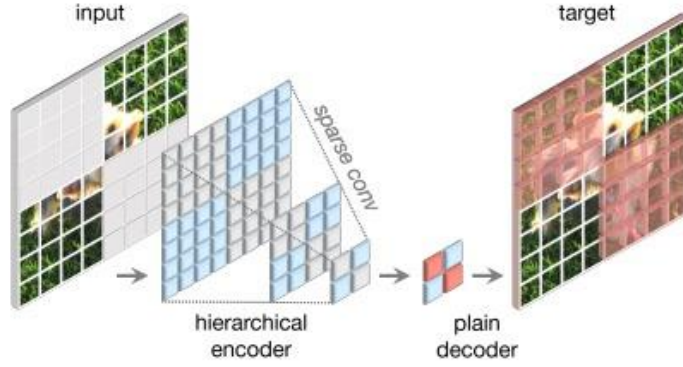
Model performansı; accuracy, precision, recall ve F1-score metrikleri üzerinden değerlendirilmiştir. Ayrıca, sınıflandırma başarısını görselleştirmek amacıyla her fold için confusion matrisi ayrı ayrı oluşturulmuş ve kaydedilmiştir.

Her fold için eğitilen modeller pt formatında saklanmış; tüm foldlara ait ortalama performans değerleri hesaplanarak genel başarı düzeyi belirlenmiştir. Foldların birleştirilmesiyle genel bir confusion matrisi oluşturulmuş, ayrıca ortalama kayıp (loss) ve doğruluk (accuracy) grafiklerinin analizi gerçekleştirilmiştir.

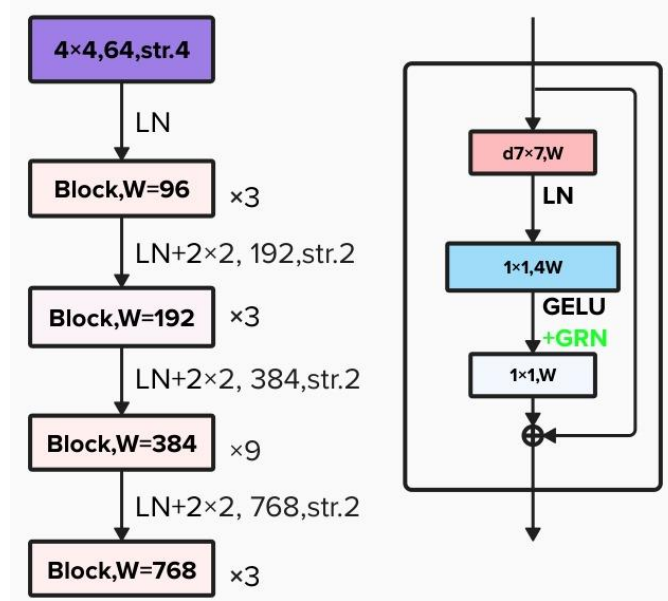
3.3.6 ConvNeXtV2

Bu çalışmada kullanılan derin öğrenme tabanlı yöntemlerden biri, Meta AI Research tarafından geliştirilen ve modern CNN mimarileri arasında yer alan ConvNeXtV2 modelidir. ConvNeXtV2, orijinal ConvNeXt mimarisinin daha gelişmiş bir versiyonudur ve görsel tanıma görevlerinde hem doğruluk hem de verimlilik açısından Transformer tabanlı modellere rekabetçi bir alternatif sunmaktadır [53]. Bu mimari, evrimsel sinir ağlarının geleneksel yapısını korurken, Transformerlardan ilham alan yapısal iyileştirmeler içermektedir. Özellikle derinliği artırılmış katman yapısı, geniş kernel boyutları ve normalize edilmiş aktivasyon fonksiyonları ile yüksek seviyeli görsel özellikleri etkili bir şekilde öğrenmektedir. Söz konusu mimarinin katmanları ve genel yapısı Şekil 27 ve Şekil 28’de gösterilmiştir.

Modelin backbone kısmı olarak kullanılan ConvNeXtV2-Base, geniş çaplı görüntü veri kümelerinde (özellikle ImageNet-22k ve MAE ön-eğitimli) önceden eğitilmiş ağırlıklarla kullanılmıştır. Bu sayede, düşük seviyeli evrimsel özelliklerin önceden öğrenilmesi sağlanmış ve transfer öğrenme yaklaşımı benimsenerek, yalnızca sınıflandırma katmanları yeniden eğitilmiştir. Bu kapsamda oluşturulan model mimarisi, Şekil 26’te gösterilmiştir.



Şekil 27 ConvNeXtV2 Modelinin Mimarisi [53]



Şekil 28 ConvNeXtV2 Modelinin Katmanları [54]

Veri setinde yer alan tüm BT görüntüleri, ConvNeXt mimarisiyle uyumlu olacak şekilde 224×224 piksel boyutuna ölçeklendirilmiştir. Görüntüler RGB formatında normalize edilerek $[0.485, 0.456, 0.406]$ ortalama ve $[0.229, 0.224, 0.225]$ standart sapma değerleri kullanılarak tensör formuna dönüştürülmüştür. Modelin eğitimi sırasında batch boyutu 32 olarak belirlenmiş ve eğitim süreci toplam 50 epoch ile sınırlandırılmıştır.

Model, 5 katlı çapraz doğrulama yöntemiyle değerlendirilmiştir. Her fold için eğitim ve test işlemleri ayrı ayrı gerçekleştirilmiş, eğitim sürecinde Adamax optimizasyon algoritması kullanılmıştır. Kayıp fonksiyonu olarak çok sınıflı çapraz entropi tercih edilmiş ve erken durdurma stratejisi uygulanarak doğrulama kaybı 5epoch boyunca iyileşme göstermediğinde eğitim sonlandırılmıştır.

```
# Model
model = timm.create_model('convnextv2_base.fcmae_ft_in22k_in1k', pretrained=True, num_classes=len(class_names))
model = model.to(device)

criterion = nn.CrossEntropyLoss()
optimizer = optim.Adamax(model.parameters(), lr=1e-4)

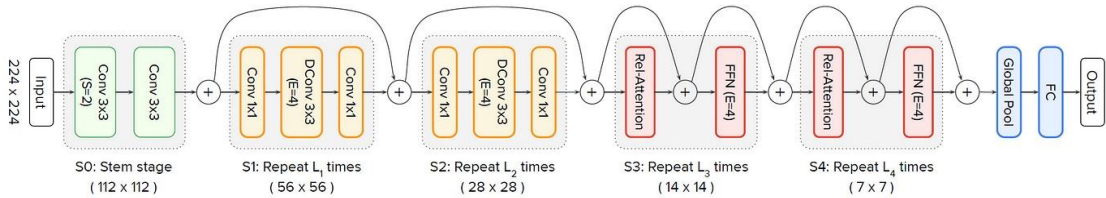
model, history, total_train_time, epoch_durations = train_model(model, dataloaders, criterion, optimizer,
                                                                num_epochs=EPOCHS, patience=PATIENCE)
total_all_training_time += total_train_time
```

Şekil 29 ConvNeXtV2 Model Oluşturma Kod Bloğu

Eğitim süreci boyunca doğruluk ve kayıp değerleri epoch bazlı olarak izlenmiş, performans takibi görselleştirme amaçlı grafiklerle desteklenmiştir. Her fold sonunda modelin test seti üzerindeki sınıflama başarımı değerlendirilmiş, confusion matrix görselleştirilmiş ve model ağırlıkları kaydedilmiştir.

3.3.7 CoAtNet-V1

Bu çalışmada, akciğer hastalıklarının tespiti ve sınıflandırılması amacıyla CoAtNet-V1 (Convolution-and-Attention Network) adlı derin öğrenme mimarisi tercih edilmiştir. CoAtNet yapısı, evrimsel sinir ağlarının (CNN) güçlü genelleme özelliklerini, Transformer tabanlı modellerin yüksek ölçeklenebilirlik yeteneğiyle bir araya getiren hibrit bir yaklaşıma sahiptir. Bu mimari, öz dikkat mekanizmalarını evrimsel işlemlerle harmanlayarak hem yerel hem de küresel düzeyde anlamlı özellikler öğrenmeyi mümkün kılmaktadır [55].



Şekil 30 CoAtNet Modelinin Katmanları [56]

Görseller, her fold için ayrı ayrı tanımlanmış veri kümeleri üzerinde 224x224 piksel boyutuna yeniden ölçeklendirilmiş ve ImageNet istatistiklerine uygun şekilde normalize

edilmiştir. Model olarak, timm kütüphanesi aracılığıyla erişilen ve görüntü sınıflandırma problemlerinde yüksek başarı elde eden coatnet_1_224 modeli tercih edilmiştir.

Modelin ImageNet veri kümesi üzerinde önceden eğitilmiş versiyonu kullanılmıştır. Bu işlem kapsamında, pretrained=False olarak belirtilmiş olsa da modelin ağırlıkları manuel olarak indirilip modele yüklenmiştir. Bu sayede, modelin genel özellikleri korunduğu halde sınıflandırma görevine özgü esneklik sağlanmıştır. CoAtNet-V1 modelinin çıkış katmanı, veri kümesindeki sınıf sayısına göre özelleştirilerek yeniden yapılandırılmıştır.

Eğitim sürecinde kayıp fonksiyonu olarak CrossEntropyLoss, optimizasyon algoritması olarak ise Adamax tercih edilmiş ve öğrenme oranı 1e-4 olarak belirlenmiştir. Eğitim boyunca erken durdurma stratejisi uygulanmış; doğrulama kaybı üst üste beş epoch boyunca iyileşme göstermediğinde eğitim süreci sonlandırılmıştır. Her epoch sonunda eğitim ve doğrulama kayıpları ile doğruluk değerleri takip edilmiş; en düşük doğrulama kaybının elde edildiği noktadaki model ağırlıkları kaydedilmiştir. Ayrıca, her bir fold için eğitim süreleri ayrı ayrı hesaplanarak toplam eğitim süresi de kayıt altına alınmıştır.

```
model = create_model('coatnet_1_224', pretrained=False, num_classes=len(class_names))
model = model.to(device)

criterion = nn.CrossEntropyLoss()
optimizer = optim.Adamax(model.parameters(), lr=1e-4)

model, history, total_train_time, epoch_durations = train_model(model, dataloaders, criterion, optimizer, num_epochs=EPOCHS, patience=PATIENCE)
total_all_training_time += total_train_time

model.eval()
```

Şekil 31 CoAtNet Model Oluşturma Kod Bloğu

Test aşamasında, modelin tahminleri ile gerçek etiketler karşılaştırılmış; elde edilen sonuçlar doğrultusunda confusion matrisi ve sınıflandırma raporu oluşturulmuştur. Bu çıktılar, her bir fold için ayrı ayrı görselleştirilmiş; model ağırlıkları pth uzantılı olarak, doğruluk-kayıp grafikleri, confusion matrisleri ve sınıflandırma raporları düzenli biçimde klasörlerde kaydedilmiştir.

Beş fold tamamlandıktan sonra, tüm test sonuçları birleştirilerek genel bir confusion matrisi ve ortalama sınıflandırma raporu oluşturulmuş; modelin genel başarı oranı hesaplanmıştır. Bu kapsamda, tüm sınıflar için ortalama doğruluk, hassasiyet, duyarlılık ve F1 skorları elde edilmiştir.

3.4. Performans Metrikleri

Makine öğrenmesi projelerinde modelin başarısını objektif bir biçimde değerlendirebilmek için performans metriklerinin doğru şekilde seçilmesi ve titizlikle uygulanması büyük önem taşımaktadır. Performans metrikleri, geliştirilen modellerin doğruluğunu, genelleme yeteneklerini ve veri ile ne ölçüde uyum sağladıklarını belirlemek amacıyla kullanılan ölçütlerdir. Bu metrikler sayesinde modellerin gerçek dünya verileri üzerindeki etkinliği sayısal olarak ifade edilebilmekte, böylece modelin güçlü ve zayıf yönleri analiz edilerek gerekli iyileştirmeler planlanabilmektedir.

Bu çalışmada, dört temel performans metriği kullanılmıştır: doğruluk (Accuracy), duyarlılık (Recall), kesinlik (Precision) ve F1 skoru (F1-Score).

3.4.1 Doğruluk (Accuracy):

Doğruluk, sınıflandırma algoritmalarının genel performansını değerlendirmede kullanılan temel metriklerden biridir ve toplam örnekler içerisindeki doğru tahminlerin oranını ifade etmektedir. Bu metrik, Denklem 1’de görüldüğü gibi hesaplanmaktadır:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Sınıf dağılımının dengeli olduğu durumlarda doğruluk metriği anlamlı sonuçlar sunmaktadır. Ancak sınıf dağılımının dengesiz olduğu senaryolarda, model baskın sınıfa yönelerek yüksek doğruluk değeri elde edebilir; bu ise azınlık sınıfın doğru tahmin edilmemesi riskini beraberinde getirmektedir. Bu tür durumlarda, modelin pozitif ve negatif sınıfları ayırt etme kapasitesini daha sağlıklı değerlendirebilmek için doğruluk metriği yerine kesinlik ve duyarlılık gibi daha ayrıntılı performans ölçütlerinin dikkate alınması gerekmektedir.

3.4.2 Kesinlik (Precision):

Kesinlik, modelin pozitif sınıf olarak tahmin ettiği örnekler içerisinde, gerçekten pozitif olanların oranını göstermektedir. Bir başka ifadeyle, pozitif sınıf etiketine sahip olduğu tahmin edilen örneklerin ne kadarının gerçekte de pozitif olduğunu ölçmektedir. Bu metrik, özellikle Tip-I hata (False Positive) üzerinde odaklanmakta ve Denklem 2’de görüldüğü gibi hesaplanmaktadır:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

Kesinlik, modelin yanlış bir şekilde pozitif tahmin yapmasının yüksek maliyetli olduğu durumlarda önem arz etmektedir.

3.4.3 Duyarlılık (Recall):

Duyarlılık, modelin gerçek pozitif sınıfa ait örnekleri ne ölçüde doğru şekilde tespit ettiğini ifade etmektedir. Başka bir deyişle, tüm pozitif örnekler içerisinde doğru biçimde sınıflandırılanların oranını ortaya koymaktadır. Bu metrik Tip-II hata (False Negative) ile ilgilenmekte olup Denklem 3'te görüldüğü gibi hesaplanmaktadır:

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

Duyarlılık, gözden kaçan pozitif örneklerin ciddi sonuçlara yol açabileceği bağlamlarda kritik bir rol üstlenmektedir.

3.4.4 F1 Skoru (F1 Score):

F1 skoru, kesinlik ve duyarlılık metriklerinin harmonik ortalaması olarak tanımlanmakta olup, modelin hem pozitif tahminlerinin doğruluğunu hem de gerçek pozitif örnekleri tespit etme başarısını aynı anda değerlendirmeye olanak tanımaktadır. Denklem 4'teki gibi hesaplanmaktadır:

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

F1 skoru, özellikle sınıf dağılımının dengesiz olduğu veri kümelerinde, model performansının daha dengeli ve anlamlı bir şekilde ölçülmesini sağlamaktadır [57].

3.4.5 Confusion Matrix

Confusion matris, sınıflandırma modellerinin performans değerlendirmesinde kullanılan temel ve sistematik bir araçtır. Bu matris, modelin öngörülleri ile gerçek sınıf etiketleri arasındaki ilişkileri detaylı bir şekilde karşılaştırarak, doğru ve hatalı sınıflandırmaları dört ayrı kategori halinde sınıflandırmaktadır. Bunun sonucunda modelin hangi tür hatalara eğilimli olduğu belirlenebilir ve performans iyileştirme stratejileri geliştirilebilir.

Confusion matris dört bileşenden oluşmaktadır:

True Positive (TP): Modelin pozitif olarak sınıflandırdığı örneklerin gerçekte de pozitif olduğu durumu ifade etmektedir.

True Negative (TN): Modelin negatif olarak sınıflandırdığı örneklerin gerçekte de negatif olduğu durumu belirtmektedir.

False Positive (FP): Modelin pozitif sınıf tahmininde bulunduğu ancak gerçek etiketin negatif olduğu durumdur; literatürde Tip I hata olarak tanımlanmaktadır.

False Negative (FN): Modelin negatif sınıf tahmininde bulunduğu ancak gerçek etiketin pozitif olduğu durumu ifade eder ve Tip II hata olarak adlandırılmaktadır.

Bu dört kategorinin ayrıntılı analizi, yalnızca genel sınıflandırma doğruluğunun ötesinde, modelin yanlış pozitif ve yanlış negatif tahminlere karşı hassasiyetini de ortaya koymaktadır. Dolayısıyla, confusion matris, özellikle tıbbi teşhis, güvenlik ve kritik karar destek sistemleri gibi alanlarda, model performansının derinlemesine incelenmesi ve güvenilirlik değerlendirmesi için vazgeçilmez bir araçtır [58].

Tablo 3 Confusion Matrix Değerleri Tablosu

	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TP)	False Negative(FN)
Actual Negative	False Positive(FP)	True Negative(TN)

3.4.6 T Testi

T-testi, iki grup ortalaması veya bir örneklem ortalaması ile önceden belirlenmiş bir referans değeri arasında istatistiksel açıdan anlamlı farklılıkların varlığını araştırmak amacıyla kullanılan parametrik bir hipotez testidir. Özellikle küçük örneklem büyüklüklerinde ve popülasyon standart sapmasının bilinmediği durumlarda tercih edilen bu test, örneklem ortalamasının popülasyon ortalamasından veya iki bağımsız örneklem ortalamaları arasındaki farkın tesadüfi olup olmadığını değerlendirmektedir.

Tek örneklem t-testinin test istatistiği aşağıdaki formül ile hesaplanmaktadır:

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

T-testinin başlıca üç türü bulunmaktadır:

Tek Örneklem T-testi: Tek bir örneklem ortalamasının, belirli bir sabit parametre veya popülasyon ortalaması ile anlamlı derecede farklılık gösterip göstermediğini tespit etmek için uygulanmaktadır. Örneğin, belirli bir sınıfın başarı ortalamasının ulusal ortalamadan anlamlı biçimde sapıp saptığını incelemesi bu teste örnektir.

Bağımsız İki Örneklem T-testi: İki bağımsız örneklemin ortalamaları arasındaki farkın istatistiksel olarak anlamlı olup olmadığını sınamak için kullanılmaktadır. Bu test, farklı gruplara uygulanan tedavi veya koşulların etkilerini karşılaştırmada yaygın şekilde tercih edilmektedir. Örneğin, farklı öğretim yöntemlerinin öğrenci başarısı üzerindeki etkilerinin karşılaştırılması bu test kapsamında değerlendirilmektedir.

Eşleştirilmiş (İlişkili) T-testi: Aynı birey veya birim üzerinde farklı zamanlarda veya farklı koşullar altında ölçülen ortalamaların karşılaştırılması için uygulanmaktadır. Bu test, müdahale öncesi ve sonrası ölçümler arasındaki değişimin anlamlılığını ortaya koymaktadır. Örneğin, yeni bir eğitim programının uygulanmasından önce ve sonra öğrencilerin akademik performanslarının karşılaştırılması bu teste örnek teşkil eder.

T-testlerinin geçerliliği aşağıdaki temel varsayımların sağlanmasına bağlıdır:

Normallik Varsayımı: Verilerin, özellikle küçük örneklem büyüklüklerinde, yaklaşık olarak normal dağılım göstermesi gerekmektedir. Normallikten önemli sapmalar, test sonuçlarının güvenilirliğini azaltabilmektedir.

Varyans Homojenliği: Bağımsız iki örneklem t-testinde, karşılaştırılan grupların varyanslarının eşit olduğu varsayılmaktadır. Varyanslar arasında belirgin farklılıkların bulunması, testin sonuçlarının doğruluğunu olumsuz etkileyebilir.

Gözlemler Arası Bağımsızlık: Her bir gözlemin birbirinden bağımsız olduğu kabul edilmektedir; bir gözlemin sonucu diğerlerini etkilememelidir. Bu varsayımın ihlali, testin istatistiksel geçerliliğini zedeleyebilir.

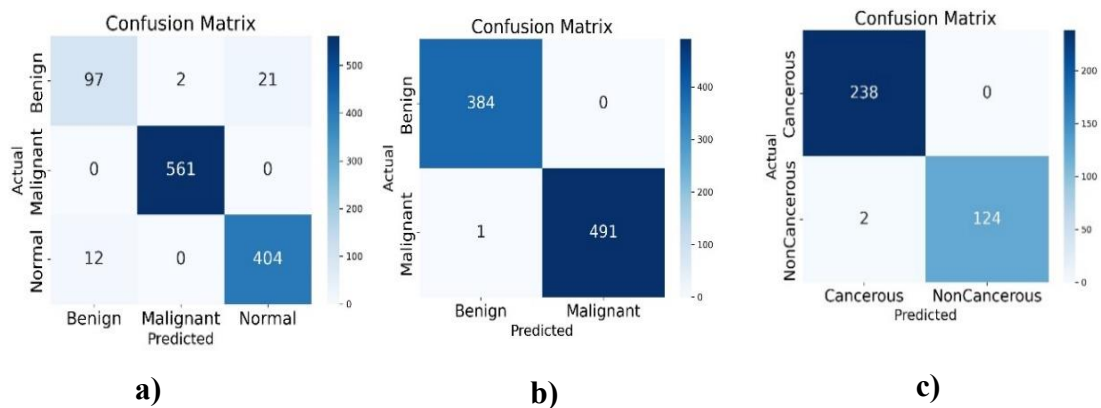
Bu varsayımların analize başlamadan önce dikkatle değerlendirilmesi, t-testinin sonuçlarının tutarlılığı ve geçerliliği açısından kritik önem taşımaktadır [59].

4. BULGULAR VE TARTIŞMA

4.1. Model Performans Analizi

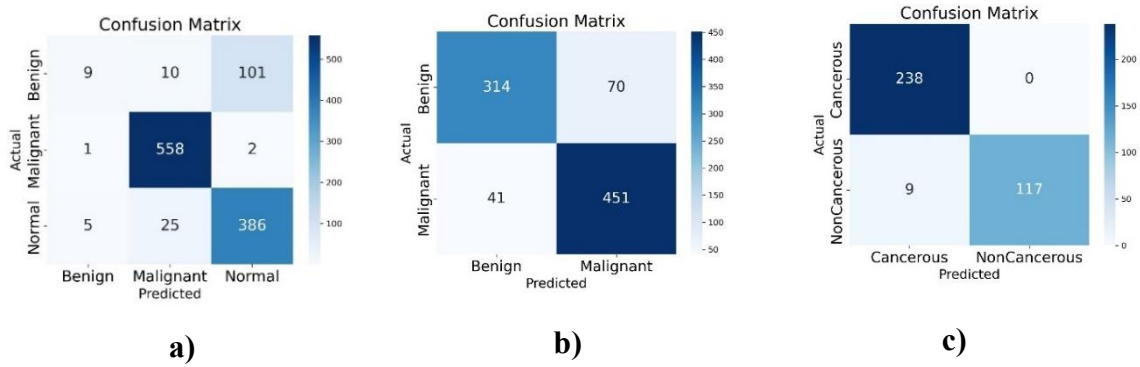
Bu çalışmada, derin öğrenme tabanlı görüntü sınıflandırma modellerinin performansları, üç farklı akciğer görüntüleme veri seti (IQ-OTH/NCCD, SPIE-AAPM-NCI ve CT-Scan Images) üzerinde 5 katlı çapraz doğrulama yöntemiyle değerlendirilmiştir. Model karşılaştırmaları, katman derinliği, mimari yapısı, dikkat mekanizmaları ve parametrik karmaşıklıkları açısından anlamlı farklılıklar göstermektedir.

VGG-16 modeli, basit ve sabit filtre boyutlarından oluşan yapısı sayesinde üç veri setinde dengeli ve güçlü performans sergilemiştir. IQ-OTH/NCCD veri seti üzerinde gerçekleştirilen değerlendirmede %96,81 doğruluk ve %93,54 F1 skoru elde edilmiştir. Şekil 32a'da sunulan confusion matrix incelendiğinde, modelin bu veri setinde oldukça düşük sayıda yanlış negatif tespit gerçekleştirdiği gözlemlenmektedir. Bu durum, modelin recall değerinin precision'a göre biraz daha düşük olmasına neden olmuştur. SPIE-AAPM-NCI ve CT-Scan Images veri setlerinde ise doğruluk ve F1-skorları %99'a yaklaşmış, Şekil 32b ve Şekil 32c'de sunulan confusion matrix'ler ise yanlış sınıflandırmaların minimum düzeyde ve dengeli olduğunu ortaya koymuştur. Bu sonuçlar, VGG-16'nın küçük ve orta ölçekli medikal görüntü verilerinde aşırı öğrenme riskini azaltarak istikrarlı sonuçlar verdiğini göstermektedir.



Şekil 32 a) Vgg 16 Modelinin Veri Artırımın Yapılmamış IQ-OTH/NCCD Veri Setindeki Confusion Matrix Sonuçları b) Vgg 16 Modelinin Veri Artırımın Yapılmamış SPIE-AAPM-NCI Veri Setindeki Confusion Matrix Sonuçları c) Vgg 16 Modelinin Veri Artırımın Yapılmamış CT-SCAN Images Veri Setindeki Confusion Matrix Sonuçları

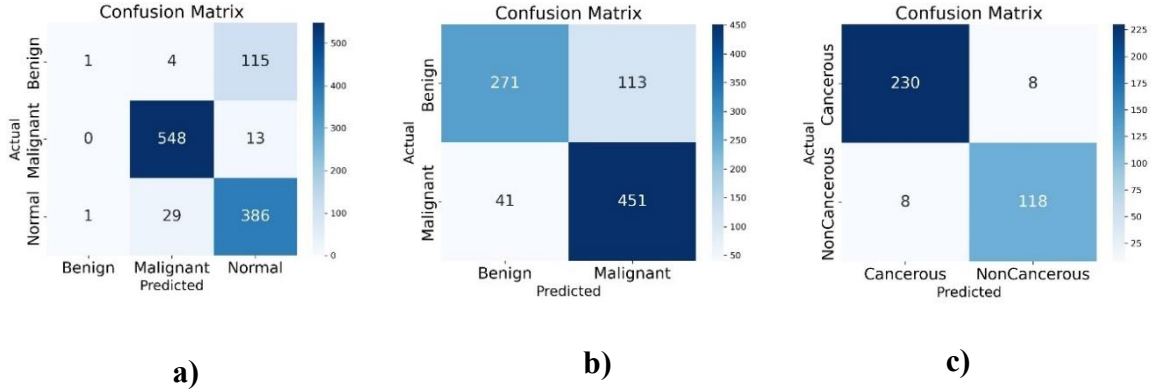
ResNet50, derin yapısı ve residual bağlantıları sayesinde teoride karmaşık özellikleri öğrenme potansiyeline sahiptir. Ancak IQ-OTH/NCCD veri setinde %86.87 doğruluk ve %64.97 F1-skoru ile diğer modellere kıyasla düşük performans göstermiştir. Bu durum, özellikle benign sınıfta sıkça yanlış negatif üretilmesiyle desteklenmektedir; Şekil 33a’da sunulan confusion matris, modelin bu sınıftaki zayıflığını net bir şekilde ortaya koymaktadır. Buna ek olarak, SPIE-AAPM-NCI ve CT-Scan Images veri setlerindeki performans ölçümlerinde belirli bir iyileşme gözlenmiştir. Şekil 33b ve Şekil 33c’de sunulan confusion matrisler incelendiğinde, modelin bu veri setlerinde daha dengeli ve tutarlı tahminler yaptığı anlaşılmaktadır. Yine de precision ve recall arasındaki farklar modelin heterojen veri setlerine adaptasyonunun kısıtlı olduğunu işaret etmektedir.



Şekil 33 a) ResNet50 Modelinin Veri Artırımın Yapılmamış IQ-OTH/NCCD Veri Setindeki Confusion Matrix Sonuçları b) ResNet50 Modelinin Veri Artırımın Yapılmamış SPIE-AAPM-NCI Veri Setindeki Confusion Matrix Sonuçları c) ResNet50 Modelinin Veri Artırımın Yapılmamış CT-SCAN Images Veri Setindeki Confusion Matrix Sonuçları

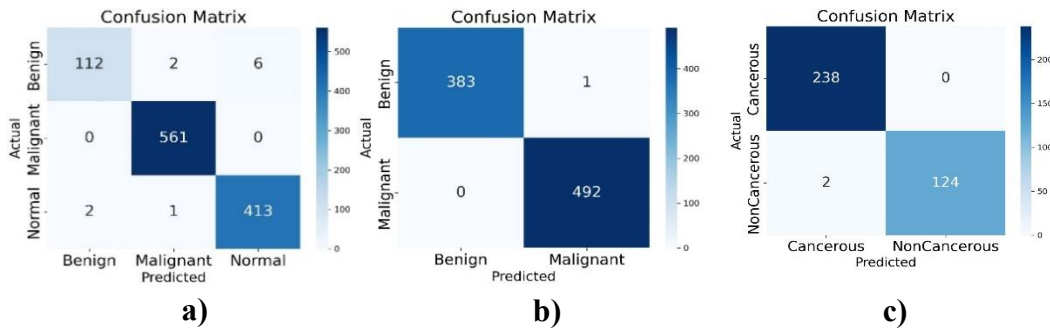
EfficientNetV2-S, parametre verimliliği yüksek modern bir model olup, parametre optimizasyonu ve ölçeklendirme avantajlarıyla genellikle dengeli sonuçlar vermesine karşın, IQ-OTH/NCCD veri setinde %85.23 doğruluk ve yalnızca %60.21 F1-skoru ile diğer modellere göre daha düşük başarı göstermiştir. Şekil 34a’da sunulan confusion matris’de benign sınıfını ayırt etmede zorlandığı ve yüksek sayıda yanlış negatif ürettiği açıkça görülmektedir. CT-Scan Images veri setinde %95.60 doğrulukla görece bir iyileşme yaşanmış, fakat Şekil 34b ve Şekil 34c’de görülen confusion matris’ler yine modelin sınıf dengesizliklerine karşı duyarlı olduğunu ve daha fazla optimizasyona ihtiyaç duyduğunu

işaret etmektedir.



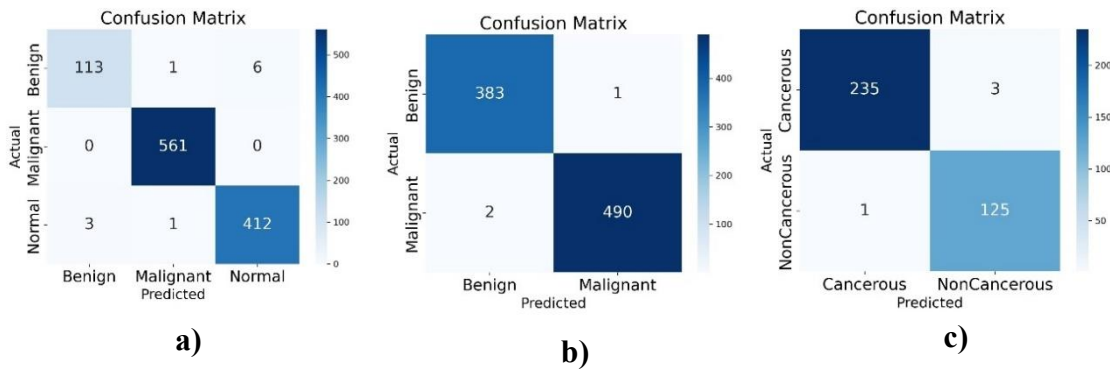
Şekil 34 a) EfficientNetV2-S Modelinin Veri Artırımın Yapılmamış IQ-OTH/NCCD Veri Setindeki Confusion Matrix Sonuçları b) EfficientNetV2-S Modelinin Veri Artırımın Yapılmamış SPIE-AAPM-NCI Veri Setindeki Confusion Matrix Sonuçları c) EfficientNetV2-S Modelinin Veri Artırımın Yapılmamış CT-SCAN Images Veri Setindeki Confusion Matrix Sonuçları

Transformer tabanlı SwinV2-CR-Small-224, üç veri setinde yüksek doğruluk, precision ve recall değerleriyle ön plana çıkmıştır. IQ-OTH/NCCD veri setinde %99'un üzerinde F1 skoru elde etmiş olup, Şekil 35a'da sunulan confusion matris bu veri setinde sadece birkaç hata yaptığını ve sınıflar arasında dengeli tahminler sunduğunu göstermektedir. SPIE-AAPM-NCI veri setinde ise %99.89 doğruluk ve %99.45 F1-skoru ile neredeyse hatasız sınıflandırma gerçekleştirmiştir. Bu başarı, modelin yerel ve global özellik çıkarımı yapabilme kabiliyetinin medikal görüntülerdeki karmaşık yapıları yakalamadaki etkinliğini yansıtmaktadır. Şekil 35c'de görülen CT-Scan Images veri setinin confusion matris'i de benzer şekilde yüksek doğruluk ve dengeli sınıflandırma başarısını doğrulamaktadır.



Şekil 35 a) SwinV2-CR-Small-224 Modelinin Veri Artırımın Yapılmamış IQ-OTH/NCCD Veri Setindeki Confusion Matrix Sonuçları b) SwinV2-CR-Small-224 Modelinin Veri Artırımın Yapılmamış SPIE-AAPM-NCI Veri Setindeki Confusion Matrix Sonuçları c) SwinV2-CR-Small-224 Modelinin Veri Artırımın Yapılmamış CT-SCAN Images Veri Setindeki Confusion Matrix Sonuçları

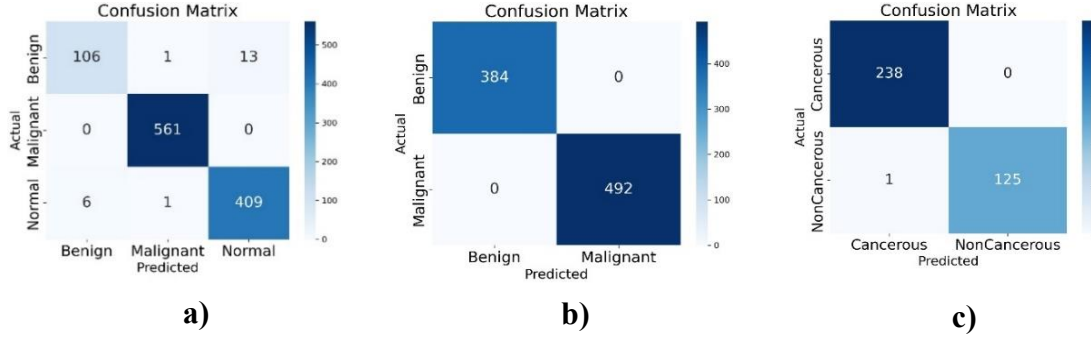
BEiT-base modeli, küçük parça tabanlı dikkat mekanizmaları ile üç veri setinde tutarlı ve yüksek performans sergilemiştir. Şekil 36a ve Şekil 36b’de görülen confusion matrisler, modelin özellikle sınıflar arası ayırmada başarılı olduğunu, yanlış sınıflandırmaların minimumda tutulduğunu göstermektedir. Precision ve recall değerlerinin birbirine çok yakın olması, modelin sınıflandırmada dengeyi koruduğunu ve hem pozitif hem negatif sınıflarda etkin olduğunu göstermektedir. Şekil 36c’de sunulan CT-Scan Images veri setinin confusion matris’i de yüksek metriklerle birlikte bu durumu desteklemektedir.



Şekil 36 a) BEiT-base Modelinin Veri Artırımın Yapılmamış IQ-OTH/NCCD Veri Setindeki Confusion Matrix Sonuçları b) BEiT-base Modelinin Veri Artırımın Yapılmamış SPIE-AAPM-NCI Veri Setindeki Confusion Matrix Sonuçları c) BEiT-base Modelinin Veri Artırımın Yapılmamış CT-SCAN Images Veri Setindeki Confusion Matrix Sonuçları

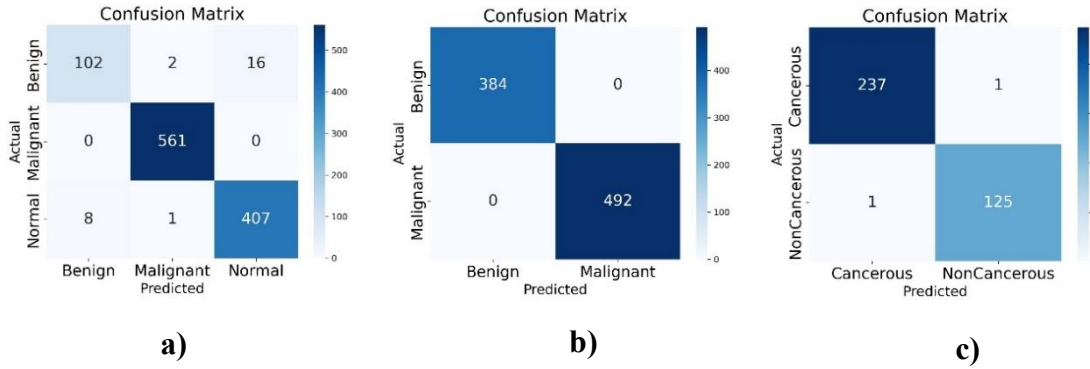
ConvNeXtV2, CNN ve Transformer unsurlarını birleştiren hibrit mimarisiyle üç veri setinde yüksek başarılar elde etmiştir. SPIE-AAPM-NCI veri seti için tüm metriklerde yüksek başarı gösterirken, Şekil 36a ve Şekil 36c’de verilen IQ-OTH/NCCD ve CT-Scan Images veri setlerinin confusion matris’leri sadece bir veya iki yanlış sınıflandırma içererek modelin yüksek kararlılık ve genelleme kabiliyetini ortaya koymuştur. Precision

ve recall oranlarının neredeyse eşit olması, overfitting riskinin düşük olduğunu ve modelin dengesiz veri setlerinde bile tutarlı tahminler yaptığını göstermektedir.



Şekil 37 a) ConvNeXtV2 Modelinin Veri Artırımın Yapılmamış IQ-OTH/NCCD Veri Setindeki Confusion Matrix Sonuçları b) ConvNeXtV2 Modelinin Veri Artırımın Yapılmamış SPIE-AAPM-NCI Veri Setindeki Confusion Matrix Sonuçları c) ConvNeXtV2 Modelinin Veri Artırımın Yapılmamış CT-SCAN Images Veri Setindeki Confusion Matrix Sonuçları

CoAtNet modeli, dikkat mekanizmaları ile konvolüsyonel katmanların hibrit yapısını kullanarak üç veri setinde üstün performans sergilemiştir. SPIE-AAPM-NCI veri setinde yüksek doğruluk ve F1-skoru ile iyi sınıflandırma sağlarken, Şekil 38b’de sunulan confusion matris’te herhangi bir hata bulunmamaktadır. Şekil 38a ve Şekil 38c’de görülen IQ-OTH/NCCD ve CT-Scan Images veri setlerinin confusion matris’leri ise sınıf dengesizliklerine rağmen yüksek precision ve recall oranlarıyla modelin tutarlı ve dengeli tahminlerde bulunduğunu göstermektedir. Bu özellikleri sayesinde CoAtNet, klinik uygulamalar için oldukça uygun bir model olarak öne çıkmaktadır.



Şekil 38 a) CoAtNet Modelinin Veri Artırımın Yapılmamış IQ-OTH/NCCD Veri Setindeki Confusion Matrix Sonuçları b) CoAtNet Modelinin Veri Artırımın Yapılmamış SPIE-AAPM-NCI Veri Setindeki Confusion Matrix Sonuçları c) CoAtNet Modelinin Veri Artırımın Yapılmamış CT-SCAN Images Veri Setindeki Confusion Matrix Sonuçları

Her bir modelin confusion matrislerine dayanarak yapılan değerlendirmelerde, Transformer tabanlı modeller (SwinV2, BEiT) ile modern CNN ve hibrit mimarilerin (ConvNeXt v2, CoAtNet), geleneksel CNN mimarilerine kıyasla daha yüksek doğruluk ve daha dengeli precision-recall değerleri sunduğu gözlemlenmiştir. Özellikle SwinV2, ConvNeXt v2 ve CoAtNet modelleri, tüm veri kümelerinde yüksek tutarlılık ve minimum hata oranı ile öne çıkmıştır. Bu sonuçlar, bu tür modellerin klinik uygulamalarda kullanıma daha uygun olduğunu göstermektedir.

4.2. Veri Artırımı Uygulanmamış ve Veri Artırımı Uygulanmış Veri Setleri Üzerindeki Modellerin Sonuçlarının Karşılaştırılması

Bu çalışma kapsamında, IQ-OTH/NCCD, SPIE-AAPM ve CT-Scan olmak üzere üç farklı akciğer görüntüleme veri seti üzerinde çeşitli derin öğrenme modelleri test edilmiştir. Her modelin, veri artırımı öncesi ve sonrası performans metrikleri karşılaştırmalı olarak incelenmiştir. Temel amaç, veri artırımı işleminin farklı mimariler üzerindeki etkisini ortaya koymak ve modellerin genelleme kapasitelerinin sentetik verilerle nasıl değiştiğini analiz etmektir. Karşılaştırmalarda F1-Skor başta olmak üzere Accuracy, Precision ve Recall gibi metrikler dikkate alınmıştır.

4.2.1 IQ-OTH/NCCD

Aşağıdaki tabloda, IQ-OTH/NCCD akciğer kanseri veri seti üzerinde eğitilen farklı derin öğrenme modellerinin veri artırımı öncesi ve sonrası elde ettikleri performans metrikleri karşılaştırmalı olarak sunulmuştur. Bu karşılaştırmalar doğrultusunda, veri artırımı uygulamasının model performanslarına olan etkisi açıkça görülmektedir:

Tablo 4 ve Tablo 5 incelendiğinde, VGG-16 modeli veri artırımı öncesinde %93.54 F1-Skor elde ederken; veri artırımı sonrası bu değeri %96.33'e çıkarmıştır. Bu artış, modelin sentetik verilerle daha güçlü genelleme yapabildiğini göstermektedir.

ResNet50, başlangıçta %64.97 F1-Skor ile en düşük performansı gösterirken, veri artırımı uygulamasıyla %79.64 seviyesine ulaşarak en yüksek performans artışını elde etmiştir. Bu durum, ResNet50'nin sınıflar arası dengesizlik nedeniyle zayıf genelleme yaptığı ancak veri artırımı ile bu eksikliğin giderilebildiğini göstermektedir.

EfficientNetV2-S, veri artırımı öncesi %60.21 F1-Skor seviyesindeyken, veri artırımı sonrası %77.10'a çıkmıştır.

SwinV2-CR-Small-224, her iki durumda da %98 üzeri F1-Skor elde ederek en iyi performansı sergileyen modellerden biri olmuştur. Veri artırımı sonrası çok az bir düşüş (%98.12 → %97.96) yaşanmış olsa da modelin istikrarlı performansı dikkat çekmektedir.

BEiT-base-patch16-224, veri artırımı öncesi %98.13 ile en yüksek F1-Skoru elde eden model olurken, veri artırımı sonrası bu skor %95.30'a düşmüştür. Bu durum, bazı modellerin sentetik veri karşısında genelleme performansında dalgalanma yaşayabileceğini göstermektedir.

ConvNeXt v2, veri artırımı öncesinde %96.27 F1-Skor ile başarılı bir performans sunarken, veri artırımı ile bu skoru %97.16'ya yükselterek istikrarlı ve güçlü bir genel başarı göstermiştir.

CoAtNet, veri artırımı öncesi %95.15 F1-Skor ile dikkat çekerken, veri artırımı sonrası skorunun %94.45'e gerilemesi, modelin belirli sentetik varyasyonlara karşı hassas olabileceğine işaret etmektedir.

Tablo 4 IQ-OTH/NCCD Lung Cancer Datasetin Modellere Göre Karşılaştırması

IQ-OTH/NCCD Lung Cancer Dataset				
Model/Folder	Accuracy	Presicion	Recall	F1-Skor
VGG-16	0.9681	0.9456	0.9265	0.9354
ResNet50	0.8687	0.7547	0.6659	0.6497
EfficientNetV2-S	0.8523	0.7314	0.6377	0.6021
SwinV2-CR-Small-224	0.9900	0.9885	0.9754	0.9812
BEiT-base-patch16-224	0.9900	0.9861	0.9773	0.9813
Convnext v2	0.9809	0.9707	0.9555	0.9627
Coatnet	0.9754	0.9614	0.9428	0.9515

Tablo 5 IQ-OTH/NCCD Lung Cancer Datasetin Veri Artırımı Sonrasında Modellere Göre Karşılaştırması

IQ-OTH/NCCD Lung Cancer Dataset(Augmentation)				
Model/Folder	Accuracy	Presicion	Recall	F1-Skor
VGG-16	0.9818	0.9734	0.9543	0.9633
ResNet50	0.8842	0.8034	0.7929	0.7964
EfficientNetV2-S	0.8496	0.7635	0.7938	0.7710
SwinV2-CR-Small-224	0.9900	0.9892	0.9714	0.9796
BEiT-base-patch16-224	0.9699	0.9491	0.9614	0.9530
Convnext v2	0.9854	0.9700	0.9733	0.9716
Coatnet	0.9654	0.9362	0.9544	0.9445

4.2.2 SPIE-AAPM Dataset

SPIE-AAPM akciğer nodülü sınıflandırma veri seti üzerinde test edilen modellerin, veri artırımı öncesi ve sonrası performanslarını karşılaştırmalı olarak incelemek amacıyla Tablo 6 ve Tablo 7’den yararlanılmıştır. Elde edilen sonuçlar, veri artırımı işleminin model başarımına olan etkisini model bazlı olarak değerlendirme imkânı sunmaktadır:

VGG-16, veri artırımı öncesinde zaten oldukça yüksek olan %99.89 F1-Skorunu veri artırımı sonrası %100’e çıkararak veri artırımının bu mimari üzerinde olumlu bir katkı sağladığını göstermiştir. Model, tüm metriklerde maksimum başarıya ulaşmıştır.

ResNet50, veri artırımı öncesinde %87.01 F1-Skor elde etmişken, veri artırımı uygulaması sonrası bu skoru %89.57’ye yükseltmiştir. Bu artış, modelin daha dengeli sınıflandırma yapabildiğini ve veri artırımı sayesinde genelleme yetisinin geliştiğini göstermektedir.

EfficientNetV2-S için, %81.65 olan F1-Skor, veri artırımı sonrası %83.39’a çıkmıştır.

Bu artış görece sınırlı olsa da modelin veri artırımı ile bir miktar daha sağlam sonuçlar ürettiği gözlemlenmiştir.

SwinV2-CR-Small-224, veri artırımı öncesinde %99.88 F1-Skor ile mükemmel yakın bir performans göstermiştir. Veri artırımı sonrası model, tüm metriklerde %100 başarıya ulaşarak en üst seviyede sınıflandırma başarısı sergilemiştir.

BEiT-base-patch16-224, veri artırımı öncesinde %99.65 F1-Skor elde etmiş, veri artırımı ile birlikte bu değer %99.88'e yükselmiştir. Yüksek başarı düzeyi korunmuş ve model, sentetik veriyle birlikte daha da istikrarlı hale gelmiştir.

ConvNeXt v2, her iki durumda da %100 başarıya ulaşarak, veri artırımı öncesi ve sonrası fark göstermeden en üst düzeyde performans sergilemiştir. Bu modelin güçlü yapısı, sınırlı veriyle bile üst düzey genelleme yapabildiğini ortaya koymaktadır.

CoAtNet, veri artırımı öncesinde %100 F1-Skor elde etmişken, veri artırımı sonrası bu skor %99.30'a düşmüştür. Bu durum, bazı modellerde veri artırımı sonrası küçük performans dalgalanmalarının yaşanabileceğini göstermektedir.

Tablo 6 SPIE-AAPM Lung Nodule Classification Challenge Datasetin Modellere Göre Karşılaştırması

SPIE-AAPM Lung Nodule Classification Challenge Dataset				
Model/Folder	Accuracy	Presicion	Recall	F1-Skor
VGG-16	0.9989	0.9987	0.9990	0.9989
ResNet50	0.8733	0.8758	0.8672	0.8701
EfficientNetV2-S	0.8242	0.8341	0.8112	0.8165
SwinV2-CR-Small-224	0.9989	0.9990	0.9987	0.9988
BEiT-base-patch16-224	0.9966	0.9964	0.9967	0.9965
Convnext v2	1.000	1.000	1.000	1.000
Coatnet	1.000	1.000	1.000	1.000

Tablo 7 SPIE-AAPM Lung Nodule Classification Challenge Datasetin Veri Artırımı Sonrasında Modellere Göre Karşılaştırması

SPIE-AAPMLung Nodule Classification Challenge Dataset (Augmentation)				
Model/Folder	Accuracy	Presicion	Recall	F1-Skor
VGG-16	1.000	1.000	1.000	1.000
ResNet50	0.8973	0.8963	0.8965	0.8957
EfficientNetV2-S	0.8425	0.8421	0.8396	0.8339
SwinV2-CR-Small-224	1.000	1.000	1.000	1.000
BEiT-base-patch16-224	0.9989	0.9987	0.9990	0.9988
Convnext v2	1.000	1.000	1.000	1.000
Coatnet	0.9932	0.9928	0.9933	0.9930

4.2.3 CT-Scan Images Dataset

Tablo 8 ve Tablo 9’a göre, CT-Scan akciğer görüntüleri veri seti üzerinde yapılan testlerde, veri artırımı uygulaması öncesi ve sonrası elde edilen sonuçlar detaylı bir şekilde incelenmiştir. Yapılan bu değerlendirme, farklı model mimarilerinin veri artırımı işlemine nasıl tepki verdiğini ortaya koymaktadır:

VGG-16, veri artırımı öncesinde %99.39 F1-Skoruna ulaşmışken, veri artırımı sonrası bu değer %99.09’a düşmüştür. Her ne kadar düşüş minimal olsa da, modelin sentetik veriyle birlikte performansında küçük bir azalma yaşandığı gözlemlenmiştir.

ResNet50 modeli, veri artırımı işlemiyle belirgin bir performans artışı göstermiştir. F1-Skoru %97.20’den %99.08’e yükselmiş; özellikle Recall değerinde önemli bir artış yaşanmıştır. Bu durum, veri artırmanın ResNet50 mimarisi için olumlu etkiler yarattığını göstermektedir.

EfficientNetV2-S, veri artırımı öncesinde %95.14 F1-Skoruna sahipken, veri artırımı sonrası bu değer %97.60’a ulaşmıştır. Tüm metriklerde görülen iyileşmeler, bu modelin veri artırımıyla daha dengeli sonuçlar verebildiğini ortaya koymuştur.

SwinV2-CR-Small-224, zaten oldukça yüksek bir başarı düzeyine sahipken, veri artırımı sonrası %99.70 F1-Skoruna ulaşarak performansını daha da artırmıştır. Bu modelin güçlü genel performansı, veri artırımıyla daha da pekiştirilmiştir.

BEiT-base-patch16-224 modeli de benzer şekilde veri artırımı sonrası %99.70 F1-Skoruna ulaşmış, veri artırımı öncesine kıyasla belirgin bir performans artışı sağlamıştır.

ConvNeXt v2, veri artırımı öncesinde zaten çok yüksek olan %99.70 F1-Skorunu veri artırımı sonrası %100'e çıkarmıştır. Model, tüm metriklerde tam başarı sağlayarak sınıflandırma görevinde en yüksek verimi göstermiştir.

CoAtNet, augmentation öncesinde %99.39 F1-Skoruna sahipken veri artırımı sonrası %99.40 ile hemen hemen aynı seviyede kalmıştır. Bu küçük artış, modelin zaten yüksek olan performansını koruduğunu ve veri artırımı etkisinin sınırlı olduğunu göstermektedir.

Tablo 8 CT-Scan Images Datasetin Modellere Göre Karşılaştırması

CT-Scan Images Dataset				
Model/Folder	Accuracy	Presicion	Recall	F1-Skor
VGG-16	0.9945	0.9958	0.9921	0.9939
ResNet50	0.9753	0.9821	0.9643	0.9720
EfficientNetV2-S	0.9560	0.9514	0.9514	0.9514
SwinV2-CR-Small-224	0.9945	0.9958	0.9922	0.9939
BEiT-base-patch16-224	0.9890	0.9864	0.9897	0.9879
Convnext v2	0.9973	0.9979	0.9962	0.9970
Coatnet	0.9945	0.9939	0.9939	0.9939

Tablo 9 CT-Scan Images Datasetinin Veri Artırımı Sonrası Modellere Göre Karşılaştırması

CT-Scan Images Dataset (Augmentation)				
Model/Folder	Accuracy	Presicion	Recall	F1-Skor
VGG-16	0.9918	0.9918	0.9900	0.9909
ResNet50	0.9918	0.9818	0.9899	0.9908
EfficientNetV2-S	0.9780	0.9713	0.9813	0.9760
SwinV2-CR-Small-224	0.9973	0.9963	0.9979	0.9970
BEiT-base-patch16-224	0.9973	0.9963	0.9979	0.9970
Convnext v2	1.000	1.000	1.000	1.000
Coatnet	0.9945	0.9922	0.9958	0.9940

Tüm veri setleri ve modeller göz önüne alındığında, veri artırımı tekniklerinin genel olarak model performanslarını stabilize ettiği ve bazı durumlarda belirgin iyileştirmelere yol açtığı söylenebilir. Özellikle IQ-OTH/NCCD gibi karmaşık ve dengesiz sınıf dağılımına sahip veri setlerinde veri artırımı, modellerin düşük temsilli sınıfları daha iyi öğrenmesini sağlamış ve bu sayede F1-Skor gibi kritik metriklerde anlamlı artışlar sağlanmıştır.

Bununla birlikte, BEiT ve CoAtNet gibi bazı ileri düzey mimariler veri artırımı sonrası sentetik varyasyonlara karşı hassasiyet göstererek performans düşüşlerine maruz kalabilmektedir. Bu durum, veri artırımı stratejilerinin model mimarisi ve veri setinin özelliklerine göre dikkatli biçimde optimize edilmesi gerektiğini ortaya koymaktadır.

Yüksek performanslı modellerde veri artırımı, çoğunlukla stabilite ve küçük iyileştirmeler sağlarken, daha düşük performanslı modellerde (örneğin EfficientNetV2-S, ResNet50) veri artırımı etkisi daha az olmaktadır. Bu nedenle, veri artırımının etkinliği hem model kapasitesine hem de veri setinin yapısına bağlı olarak değişkenlik göstermekte, dolayısıyla uygulama aşamasında model-veri seti uyumu göz önünde bulundurulmalıdır.

Sonuç olarak, veri artırımı medikal görüntü sınıflandırma problemlerinde önemli bir destekleyici teknik olarak öne çıkmakta, ancak stratejilerin her modelin karakteristiğine göre uyarlanması başarının anahtarı olmaktadır.

4.3. T- Testi Karşılaştırması

Bu çalışmada, yedi farklı derin öğrenme mimarisinin (VGG16, ResNet50, EfficientNetV2S, SwinV2-CR-Small-224, BEiT-base-patch16-224, ConvNeXt V2 ve CoAtNet-1) tıbbi görüntü sınıflandırmadaki başarımları, hem ham veri hem de veri artırımı uygulanmış versiyonlarıyla karşılaştırmalı olarak incelenmiştir. Analizler, IQ-OTH/NCCD, SPIE-AAPM-NCI ve CT-Scan olmak üzere üç farklı akciğer görüntüleme veri seti üzerinde gerçekleştirilmiştir. Tüm modeller, her veri setinde 5-fold çapraz doğrulama yöntemi ile test edilmiş ve doğruluk değerleri elde edilmiştir.

Her iki senaryo (veri artırımı öncesi ve sonrası) için elde edilen doğruluk değerleri, bağımsız örneklem t-testi ile analiz edilmiştir. Bu testler sırasında model çiftleri arasında eşit varyans varsayımı yapılmaksızın p-değerleri hesaplanmış ve 0.1'in altındaki değerler istatistiksel olarak anlamlı fark olarak değerlendirilmiştir. Her bir karşılaştırmada, doğruluk ortalaması daha yüksek olan model "kazanan", daha düşük olan "kaybeden", fark anlamlı değilse "beraberlik" olarak sınıflandırılmıştır.

Tüm bu karşılaştırmalar sonucunda, her modelin diğerlerine karşı kaç kez kazandığı, kaybettiği veya berabere kaldığı belirlenmiş ve bu bilgiler K/B/Kyp (Kazanma/Beraberlik/Kaybetme) formatında özet tablolar halinde sunulmuştur. Ayrıca, her model için toplam başarı durumu da hesaplanarak tabloya dahil edilmiştir.

Orijinal veri setleri kullanılarak gerçekleştirilen bağımsız örneklem t-testi sonuçları Tablo 10'da sunulmaktadır. Bu tablo incelendiğinde, en fazla istatistiksel galibiyet elde eden modelin SwinV2-CR-Small-224 olduğu ve toplam 8 galibiyet ile diğer modellerden ayrıştığı görülmektedir. Bu modeli sırasıyla BEiT-base-patch16-224 (7 galibiyet) ve ConvNeXt V2 ile CoAtNet-1 (6'şar galibiyet) takip etmektedir. Bu bulgu, Transformer tabanlı (SwinV2, BEiT) ve modern CNN/hibrit mimarilerin (ConvNeXt, CoAtNet), geleneksel CNN mimarilerine kıyasla daha rekabetçi bir performans sergilediğini göstermektedir. Geleneksel yapılardan VGG16 yalnızca 6 galibiyet elde ederken, ResNet50 ve EfficientNetV2S modelleri sırasıyla yalnızca 1 ve 0 galibiyet ile listenin sonunda yer almıştır.

Tablo 10 Modellerin T-Test Karşılaştırma Sonuçları

	VGG16	ResNet50	EfficientNetV2S	SwinV2-CR-Small-224	BEiT-base-patch16-224	Convnext v2	Coatnet-1
VGG16	-	0/0/3	0/0/3	1/2/0	1/2/0	0/3/0	0/3/0
ResNet50	3/0/0	-	0/2/1	3/0/0	2/1/0	3/0/0	3/0/0
EfficientNetV2S	3/0/0	1/2/0	-	3/0/0	3/0/0	3/0/0	3/0/0
SwinV2-CR-Small-224	0/2/1	0/0/3	0/0/3	-	0/3/0	0/3/0	0/2/1
BEiT-base-patch16-224	0/2/1	0/1/2	0/0/3	0/3/0	-	0/3/0	0/2/1
Convnext v2	0/3/0	0/0/3	0/0/3	0/3/0	0/3/0	-	0/3/0
Coatnet-1	0/3/0	0/0/3	0/0/3	1/2/0	1/2/0	0/3/0	-
Toplam Kazanma	6	1	0	8	7	6	6
Toplam Beraberlik	10	3	2	10	11	12	10
Toplam Kaybetme	2	14	16	0	0	0	2

Veri artırımı stratejileri uygulandıktan sonra yapılan t-testi analiz sonuçları Tablo 11'de sunulmuştur. Bu aşamada, ConvNeXt V2 modeli 9 galibiyet ile istatistiksel olarak en başarılı model konumuna yükselmiştir. Ardından SwinV2-CR-Small-224 (8 galibiyet) ve VGG16 (6 galibiyet) gelmektedir. Bu durum, ConvNeXt V2'nin veri artırımı sonrasında gösterdiği istatistiksel üstünlüğe işaret etmektedir. Öte yandan, EfficientNetV2S, veri artırımı sonrası değerlendirmelerde de herhangi bir galibiyet elde edememiştir. Ayrıca, bazı modellerin beraberlik sayılarında gözlenen artış, veri artırımının model kararlılığını artırıcı etkisine işaret etmektedir.

Tablo 11 Veri Artırımı Sonrası Modellerin T-Test Karşılaştırma Sonuçları

	VGG16	ResNet50	EfficientNetV2S	SwinV2-CR-Small-224	BEiT-base-patch16-224	Convnext v2	Coatnet-1
VGG16	-	2/1/0	3/0/0	0/3/0	0/3/0	0/2/1	1/2/0
ResNet50	0/1/2	-	2/1/0	0/1/2	0/1/2	0/1/2	0/1/2
EfficientNetV2S	0/0/3	0/1/2	-	0/0/3	0/0/3	0/0/3	0/0/3
SwinV2-CR-Small-224	0/3/0	2/1/0	3/0/0	-	1/2/0	0/3/0	2/1/0
BEiT-base-patch16-224	0/3/0	2/1/0	3/0/0	0/2/1	-	0/2/1	0/3/0
Convnext v2	1/2/0	2/1/0	3/0/0	0/3/0	1/2/0	-	2/1/0
Coatnet-1	0/2/1	2/1/0	3/0/0	0/1/2	0/3/0	0/1/2	-
Toplam Kazanma	6	2	0	8	5	9	5
Toplam Beraberlik	11	6	1	10	11	9	8
Toplam Kaybetme	1	10	17	0	2	0	5

Her iki aşamada yapılan t-testi analizlerinin sonuçları birlikte değerlendirildiğinde, hem orijinal hem de veri artırımı uygulanmış veri setlerinde Transformer tabanlı (SwinV2-CR-Small-224, BEiT-base-patch16-224) ve modern CNN/hibrit mimarilerin (ConvNeXt V2, CoAtNet-1) istatistiksel açıdan geleneksel CNN modellerine kıyasla daha başarılı ve daha kararlı performans sergilediği anlaşılmaktadır.

Veri artırımı stratejileri, bazı modellerin doğruluk performanslarını artırmakla kalmamış, aynı zamanda istatistiksel testlerde anlamlı fark oluşturarak modeller arası kıyaslamalarda galibiyet sayılarının yeniden dağılımına neden olmuştur. Özellikle ConvNeXt V2 modeli, veri artırımı sonrasında en yüksek galibiyet sayısına ulaşarak dikkat çekmiştir.

Sonuç olarak, bu bulgular yalnızca model seçiminin değil, aynı zamanda veri ön işleme ve artırma stratejilerinin de derin öğrenme tabanlı medikal görüntüleme sistemlerinde kritik bir rol oynadığını ortaya koymaktadır.

5. SONUÇLAR

Bu çalışma kapsamında, evrişimsel sinir ağları (CNN) ve Transformer tabanlı modern derin öğrenme mimarilerinin, akciğer BT (bilgisayarlı tomografi) görüntülerini sınıflandırmadaki başarı performansları karşılaştırmalı olarak incelenmiştir. Akciğer kanseri gibi ölümcül hastalıkların erken teşhisi, hastalığın ilerlemesini önlemek ve tedavi sürecini optimize etmek açısından kritik öneme sahiptir. Bu doğrultuda, geliştirilen yapay zekâ destekli sınıflandırma sistemleri, klinik karar destek mekanizmaları için umut verici çözümler sunmaktadır.

Çalışmada kullanılan IQ-OTH/NCCD, SPIE-AAPM-NCI ve CT-Scan olmak üzere üç farklı akciğer BT veri seti, veri artırımı ve özenli ön işleme adımlarıyla hazırlanmış ve 5 katlı çapraz doğrulama yöntemiyle değerlendirilmiştir. VGG16, ResNet50, EfficientNetV2-S gibi klasik CNN mimarileri ile SwinV2, BEiT-Base, ConvNeXtV2 ve CoAtNet gibi modern Transformer ve hibrit tabanlı modeller eğitilmiş ve test edilmiştir.

Deneysel sonuçlar, Transformer tabanlı modellerin özellikle BEiT-Base ve SwinV2 mimarilerinin, doğruluk, kesinlik, duyarlılık ve F1 skoru gibi metriklerde üstün performans gösterdiğini ortaya koymuştur. Geleneksel CNN mimarileri, daha düşük hesaplama gereksinimleri sayesinde bazı durumlarda avantaj sağlasa da, yüksek doğruluk oranları açısından Transformer mimarileriyle rekabet etmekte zorlanmıştır. Ayrıca, veri artırımı tekniklerinin model başarımını anlamlı düzeyde artırdığı ve özellikle dengesiz sınıf dağılımı olan veri kümelerinde genelleme kabiliyetini yükselttiği gözlemlenmiştir. T-testi ile yapılan istatistiksel analizler, elde edilen başarıların rastlantısal olmadığını ve modeller arasındaki farklılıkların anlamlı olduğunu doğrulamıştır.

Sonuç olarak, bu çalışmada elde edilen bulgular, derin öğrenme ve özellikle Transformer tabanlı mimarilerin, tıbbi görüntü analizi ve akciğer kanseri teşhisinde etkin ve güvenilir araçlar olarak kullanılabileceğini göstermektedir. Yapay zekâ destekli bu sistemler, radyologlara nesnel ve hızlı analiz imkânı sunarak klinik karar süreçlerine değerli katkılar sağlayabilir. Gelecekte, daha geniş ve dengeli veri setleri ile yapılacak çalışmalar ve çok yönlü multimodal yaklaşımlar sayesinde, bu sistemlerin hastane ortamlarına entegre edilerek daha yaygın ve etkili şekilde kullanılabileceği öngörülmektedir.

KAYNAKLAR

- [1] A. PAVLOPOULOU, D. A. SPANDIDOS, and I. MICHALOPOULOS, “Human cancer databases (Review),” *Oncology Reports*, vol. 33, no. 1, pp. 3–18, Oct. 2014, doi: <https://doi.org/10.3892/or.2014.3579>.
- [2] R. L. Siegel, K. D. Miller, and A. Jemal, “Cancer statistics, 2016,” *CA: A Cancer Journal for Clinicians*, vol. 66, no. 1, pp. 7–30, Jan. 2016.
- [3] World Health Organization, “Cancer,” *World Health Organization*, Feb. 03, 2025. <https://www.who.int/news-room/fact-sheets/detail/cancer>
- [4] V. Gant and R. Dybowski, *Clinical Applications of Artificial Neural Networks*. Cambridge University Press, 2001. doi: <https://doi.org/10.1017/cbo9780511543494>.
- [5] Azzam Taktak, “The use of artificial neural networks in decision support in cancer: A systematic review,” *Neural Networks*, 2006. https://www.academia.edu/13719857/The_use_of_artificial_neural_networks_in_decision_support_in_cancer_A_systematic_review (accessed Mar. 22, 2025).
- [6] M. Nur, Tülin ÖZTÜRK, and M. TALO, “Yeni bir Evrişimsel Sinir Ağı Modeli Kullanarak Bilgisayarlı Tomografi Görüntülerinden Akciğer Kanseri Tespiti,” *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, vol. 34, no. 2, pp. 795–802, Sep. 2022, doi: <https://doi.org/10.35234/fumbd.1135691>.
- [7] M. Rasel, None Hasib Fardin, M. Ismail, A. Haque, and A. Sakib, “Hybrid deep learning for interpretable lung cancer recognition across computed tomography and histopathological imaging modalities,” *International Journal of Science and Research Archive*, vol. 15, no. 1, pp. 1798–1810, Apr. 2025, doi: <https://doi.org/10.30574/ijrsra.2025.15.1.1161>.
- [8] S. Ameer, “DETECTION OF LUNG CANCER THROUGH COMPUTED TOMOGRAPHIC IMAGES USING DEEP LEARNING MODELS,” *Journal of Population Therapeutics & Clinical Pharmacology*, pp. 4188–4198, Sep. 2024, doi: <https://doi.org/10.53555/j7tj1t43>.

- [9] T. Adams, J. Dörpinghaus, M. Jacobs, and V. Steinhage, “Automated lung tumor detection and diagnosis in CT Scans using texture feature analysis and SVM,” Communication Papers of the 2018 Federated Conference on Computer Science and Information Systems, Sep. 2018, doi: <https://doi.org/10.15439/2018f176>.
- [10] S. V. Thambi, I. G. K. S. Kammari, and A. Sahay, “Hybrid Metaheuristic Optimisation for Lung Cancer Image Classification: Leveraging MOEA, PSO, and ACO Algorithms,” *Procedia Computer Science*, vol. 258, pp. 3781–3793, May 2025, doi: <https://doi.org/10.1016/j.procs.2025.04.633>.
- [11] Ferhat Ayayna, Abidin Çalışkan, and A. Aslan, “Akciğer Kanserinin Derin Öğrenme Yaklaşımları Kullanılarak Tespit Edilmesi,” *International Conference on Scientific and Academic Research (ICSAR’22)*, Dec. 2022, Available: https://www.researchgate.net/publication/366464730_Akciger_Kanserinin_Derin_Ogrenme_Yaklasimlari_Kullanilarak_Tespit_Edilmesi
- [12] D. Şeker, M. S. Kartal, A. Yıldız, and I. Öksüz, “Akciğer Kanseri tespitinde Dönüşüm ve Evrişim Tabanlı modeller Ile Açıklanabilir Yapay Zeka Uygulaması,” *EMO Bilimsel Dergi*, <https://dergipark.org.tr/tr/pub/emobd/issue/86413/1484039> (accessed Mar. 23, 2025).
- [13] R. Kohavi, “A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection,” 1995. Available: <https://www.ijcai.org/Proceedings/95-2/Papers/016.pdf>
- [14] “The IQ-OTH/NCCD lung cancer dataset,” *www.kaggle.com*. <https://www.kaggle.com/datasets/hamdallak/the-iqothnccd-lung-cancer-data>
- [15] The, “SPIE-AAPM-NCI Lung Nodule Classification Challenge Dataset,” *The Cancer Imaging Archive (TCIA)*, 2023. <https://www.cancerimagingarchive.net/collection/spie-aapm-lung-ct-challenge/> (accessed Jun. 14, 2025).
- [16] N. Maleki, “CT-Scan images,” vol. 3, Jun. 2020, doi: <https://doi.org/10.17632/p2r42nm2ty.3>.

- [17] Y. R. Haddadi, B. Mansouri, and F. zohra D. Khoudja, “A Novel Medical Image Enhancement Algorithm Based on CLAHE and Pelican Optimization,” Jan. 2023, doi: <https://doi.org/10.21203/rs.3.rs-2443705/v1>.
- [18] C. R. Harris et al., “Array Programming with NumPy,” *Nature*, vol. 585, no. 7825, pp. 357–362, Sep. 2020, doi: <https://doi.org/10.1038/s41586-020-2649-2>.
- [19] J. D. Hunter, “Matplotlib: A 2D Graphics Environment,” *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90–95, 2007, Available: <https://ieeexplore.ieee.org/document/4160265>
- [20] M. Waskom, “seaborn: statistical data visualization,” *ResearchGate*, Apr. 06, 2021. https://www.researchgate.net/publication/350701622_seaborn_statistical_data_visualization
- [21] A. Paszke et al., “PyTorch: An Imperative Style, High-Performance Deep Learning Library,” *arXiv.org*, 2019. <https://arxiv.org/abs/1912.01703>
- [22] “torchvision — Torchvision 0.22 documentation,” *Pytorch.org*, 2024. <https://docs.pytorch.org/vision/stable/index.html>
- [23] F. Pedregosa et al., “Scikit-learn: Machine Learning in Python,” *arXiv:1201.0490 [cs]*, Jun. 2018, Available: <https://arxiv.org/abs/1201.0490>
- [24] “Pytorch Image Models (timm) | timmdocs,” *timm.fast.ai*. <https://timm.fast.ai/>
- [25] G. Bradski, “The OpenCV library,” *Dr. Dobb’s Journal*, vol. 25, no. 11, pp. 120–125, 2000, Available: <https://www.proquest.com/docview/202684726?sourcetype=Trade%20Journals>
- [26] W. Mckinney, “pandas: a Foundational Python Library for Data Analysis and Statistics,” *ResearchGate*, 2011. https://www.researchgate.net/publication/265194455_pandas_a_Foundational_Python_Library_for_Data_Analysis_and_Statistics
- [27] “glob — Unix style pathname pattern expansion — Python 3.8.3rc1 documentation,” *docs.python.org*. <https://docs.python.org/3/library/glob.html>

- [28] C. O. da Costa-Luis, “tqdm: A Fast, Extensible Progress Meter for Python and CLI,” *Journal of Open Source Software*, vol. 4, no. 37, p. 1277, May 2019, doi: <https://doi.org/10.21105/joss.01277>.
- [29] T. Wolf et al., “HuggingFace’s Transformers: State-of-the-art Natural Language Processing,” *arXiv:1910.03771 [cs]*, Feb. 2020, Available: <https://arxiv.org/abs/1910.03771>
- [30] Hugging Face, "safetensors: Simple, safe way to store and distribute tensors," GitHub, [Çevrimiçi]. Mevcut: <https://github.com/huggingface/safetensors>. Erişim: 14 Haz. 2025.
- [31] “Automatic Mixed Precision package - torch.amp — PyTorch 2.7 documentation,” Pytorch.org, 2024. <https://docs.pytorch.org/docs/stable/amp.html>
- [32] GeeksforGeeks, “Differences between torch.nn and torch.nn.functional,” GeeksforGeeks, Feb. 20, 2024. <https://www.geeksforgeeks.org/deep-learning/differences-between-torch-nn-and-torch-nn-functional/#what-is-torchnnfunctional> (accessed Jun. 14, 2025).
- [33] “Multiprocessing package - torch.multiprocessing — PyTorch 2.7 documentation,” Pytorch.org, 2024. <https://docs.pytorch.org/docs/stable/multiprocessing.html> (accessed Jun. 14, 2025).
- [34] Python Software Foundation, “json — JSON encoder and decoder — Python 3.8.3rc1 documentation,” [docs.python.org](https://docs.python.org/3/library/json.html), <https://docs.python.org/3/library/json.html>
- [35] Python Software Foundation, “pickle — Python object serialization — Python 3.7.3 documentation,” Python.org, 2019. <https://docs.python.org/3/library/pickle.html>
- [36] A. Buslaev, V. I. Iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin, and A. A. Kalinin, “Albumentations: Fast and Flexible Image Augmentations,” *Information*, vol. 11, no. 2, p. 125, Feb. 2020, doi: <https://doi.org/10.3390/info11020125>.
- [37] Python Software Foundation, “time — Time access and conversions — Python 3.7.2 documentation,” Python.org, 2000. <https://docs.python.org/3/library/time.html>

- [38] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv.org, Apr. 10, 2015. <https://arxiv.org/abs/1409.1556>
- [39] mrgrhn, "VGGNet with TensorFlow (Transfer Learning with VGG16 Included)," Medium, Jan. 23, 2021. <https://ai.plainenglish.io/vggnet-with-tensorflow-transfer-learning-with-vgg16-included-7e5f6fa9479a>
- [40] M. Humayun, R. Sujatha, S. N. Almuayqil, and N. Z. Jhanjhi, "A Transfer Learning Approach with a Convolutional Neural Network for the Classification of Lung Carcinoma," *Healthcare*, vol. 10, no. 6, p. 1058, Jun. 2022, doi: <https://doi.org/10.3390/healthcare10061058>.
- [41] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," arXiv.org, Dec. 10, 2015. <https://arxiv.org/abs/1512.03385>
- [42] Ç. ERÇELİK and K. HANBAY, "Gauss Filtreleme ve ResNET50 Modeli Kullanılarak Beyin Tümörlerinin Sınıflandırılması," *Computer Science*, Oct. 2023, doi: <https://doi.org/10.53070/bbd.1345848>.
- [43] "Understanding ResNet-50 in Depth: Architecture, Skip Connections, and Advantages Over Other Networks - Wisdom ML," Mar. 30, 2023. <https://wisdomml.in/understanding-resnet-50-in-depth-architecture-skip-connections-and-advantages-over-other-networks/>
- [44] M. Tan and Q. V. Le, "EfficientNetV2: Smaller Models and Faster Training," arXiv:2104.00298 [cs], Jun. 2021, Available: <https://arxiv.org/abs/2104.00298>
- [45] A. A. Abd El-Aziz, M. A. Mahmood, and S. Abd El-Ghany, "A Robust EfficientNetV2-S Classifier for Predicting Acute Lymphoblastic Leukemia Based on Cross Validation," *Symmetry*, vol. 17, no. 1, p. 24, Dec. 2024, doi: <https://doi.org/10.3390/sym17010024>.
- [46] J. Seo, H. Kim, and D. Lee, "Identification of Individual Hanwoo Cattle by Muzzle Pattern Images through Deep Learning," ResearchGate, [Online]. Available: https://www.researchgate.net/figure/Efficientnet-v2-s-Architecture-a-represents-the-schematic-of-the-model-from-image-input_fig4_373782480. [Accessed: Jun. 12, 2025].

- [47] Z. Liu et al., “Swin Transformer V2: Scaling Up Capacity and Resolution,” arXiv:2111.09883 [cs], Apr. 2022, Available: <https://arxiv.org/abs/2111.09883>
- [48] Z. Liu et al., “Swin Transformer: Hierarchical Vision Transformer using Shifted Windows,” arXiv:2103.14030 [cs], Mar. 2021, Available: <https://arxiv.org/abs/2103.14030>
- [49] N. DiSalvo, “Swin-Transformer from Scratch in PyTorch - Python in Plain English,” Medium, Sep. 12, 2023. <https://python.plainenglish.io/swin-transformer-from-scratch-in-pytorch-31275152bf03> (accessed Jun. 14, 2025).
- [50] H. Bao, L. Dong, and F. Wei, “BEiT: BERT Pre-Training of Image Transformers,” arXiv:2106.08254 [cs], Jun. 2021, Available: <https://arxiv.org/abs/2106.08254>
- [51] Muhammad Waseem Sabir, M. Farhan, Nabil Sharaf Almalki, M. M. Alnfai, and Gabriel Avelino Sampedro, “FibroVit—Vision transformer-based framework for detection and classification of pulmonary fibrosis from chest CT images,” *Frontiers in Medicine*, vol. 10, Nov. 2023, doi: <https://doi.org/10.3389/fmed.2023.1282200>.
- [52] S.-H. Tsang, “Review — BEiT: BERT Pre-Training of Image Transformers,” Medium, Sep. 2022. <https://sh-tsang.medium.com/review-beit-bert-pre-training-of-image-transformers-c14a7ef7e295>
- [53] S. Woo et al., “ConvNeXt V2: Co-designing and Scaling ConvNets with Masked Autoencoders,” arXiv:2301.00808 [cs], Jan. 2023, Available: <https://arxiv.org/abs/2301.00808>
- [54] M. A. Rahman, A. Joy, A. T. Abir, and T. Shimamura, “Unleashing the Power of Open-Source Transformers in Medical Imaging: Insights from a Brain,” *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 7, p. 2024, Jul. 2024, doi: <https://doi.org/10.14569/IJACSA.2024.01507126>.
- [55] Z. Dai, H. Liu, Q. V. Le, and M. Tan, “CoAtNet: Marrying Convolution and Attention for All Data Sizes,” arXiv:2106.04803 [cs], Sep. 2021, Available: <https://arxiv.org/abs/2106.04803>

- [56] S.-H. Tsang, “Review — CoAtNet: Marrying Convolution and Attention for All Data Sizes,” Medium, May 17, 2022. <https://sh-tsang.medium.com/review-coatnet-marrying-convolution-and-attention-for-all-data-sizes-1462b9bc25ac>
- [57] “Lojistik Regresyon ve Sınıflandırma Metrikleri,” Miuul.com, 2022. <https://miuul.com/blog/lojistik-regresyon-ve-siniflandirma-metrikleri> (accessed Jun. 11, 2025).
- [58] GeeksforGeeks, “Understanding the Confusion Matrix in Machine Learning,” GeeksforGeeks, Oct. 15, 2017. https://www-geeksforgeeks-org.translate.goog/confusion-matrix-machine-learning/?_x_tr_sl=en&_x_tr_tl=tr&_x_tr_hl=tr&_x_tr_pto=tc (accessed Jun. 11, 2025).
- [59] A. Thevapalan, “T-test vs. Z-test: When to Use Each,” Datacamp.com, Aug. 15, 2024. <https://www.datacamp.com/tutorial/t-test-vs-z-test>

[EK]

Hafta	Faaliyet	Açıklama
1-2. Hafta	Literatür Taraması	İlgili alandaki güncel akademik yayınlar ve çalışmalar detaylı biçimde incelenerek proje kapsamında kullanılacak yöntemler, modelleme yaklaşımları ve veri kaynakları belirlenmiştir.
3-5. Hafta	Veri Setlerinin Temini, İncelenmesi ve Veri Ön İşleme	Çalışmanın amacına uygun, güvenilir ve yeterli büyüklüğe sahip veri setleri belirlenmiş; veri setlerinin içerik, format ve kalite açısından analizi gerçekleştirilmiştir. Ayrıca veri seti üzerinde gerekli ön işlemler yapılmıştır.
6-8. Hafta	Modelin Eğitilmesi, Optimizasyonu ve Model Seçimi	Literatürde başarı göstermiş çeşitli derin öğrenme ve Transformer mimarileri değerlendirilerek problem yapısına en uygun model seçilmiştir. Seçilen model, eğitim verisi üzerinde eğitilerek doğruluk, kayıp ve diğer performans metrikleri izlenmiştir.
9. Hafta	Model Performansının Değerlendirilmesi	Model, test verisi üzerinde değerlendirilmiş ve elde edilen performans sonuçları yorumlanmıştır.
10-12. Hafta	Bitirme Raporunun Hazırlanması	Bitirme raporu, proje çıktıları doğrultusunda uygun şekilde hazırlanmıştır.
13-14. Hafta	Sunum Materyallerinin Hazırlanması	Sunum materyalleri hazırlanmış ve bitirme raporu son şekline getirilmiştir.
15. Hafta	Proje Sunumu ve Son Teslimler	