



Teknoloji Fakültesi

BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

**AKCİĞER BT GÖRÜNTÜLERİNİN MAKİNE
ÖĞRENMESİ ALGORİTMALARIYLA ANALİZİ
VE AKCİĞER KANSERİ TESPİTİ**

BİTİRME PROJESİ 2. ARA RAPORU

Bilgisayar Mühendisliği Bölümü

"

DANIŞMAN

Doç. Dr. AYŞE BERNA ALTINEL GİRGİN

İSTANBUL, 2025

"

MARMARA ÜNİVERSİTESİ
TEKNOLOJİ FAKÜLTESİ
BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

Marmara Üniversitesi Teknoloji Fakültesi Bilgisayar Mühendisliği Öğrencileri Sude Nur GÜLCAN, Karun ACAR, Elif TILKI tarafından “AKCİĞER BT GÖRÜNTÜLERİNİN MAKİNE ÖĞRENMESİ ALGORİTMALARIYLA ANALİZİ VE AKCİĞER KANSERİ TESPİTİ ” başlıklı proje çalışması, xxx tarihinde savunulmuş ve jüri üyeleri tarafından başarılı bulunmuştur.

Jüri Üyeleri

Dr. Öğr. Üyesi xxx xxx
Marmara Üniversitesi
Prof. Dr. Xxx xxx
Marmara Üniversitesi
Prof. Dr. Xxx xxx
Marmara Üniversitesi

(Danışman)

(Üye)

(Üye)

(İMZA).....

(İMZA).....

(İMZA).....

ÖNSÖZ

Proje çalışmamız süresince karşılaştığım bütün problemlerde, sabırla yardım ve bilgilerini esirgemeyen, tüm desteğini sonuna kadar yanımda hissettiğim değerli hocalarım, sayın Dr. Öğr. Üyesi Xxx xxx ve sayın Prof. Dr. Xxx xxx' a en içten teşekkürlerimi sunarım.

Bu proje çalışması fikrinin oluşması ve ortaya çıkmasındaki önerisi ve desteğinden dolayı değerli hocam Dr. Öğr. Üyesi Xxx xxx' a teşekkür ederim.

Proje çalışmam sırasında maddi ve manevi desteklerini esirgemeyen okul içerisinde ve okul dışında her zaman yanımda olan değerli çalışma arkadaşlarım ve hocalarım Doç. Dr. Xxx xxx ve Dr. Öğr. Üyesi ' xxx xxx a sonsuz teşekkürlerimi sunarım.

İÇİNDEKİLER

1. GİRİŞ.....	1
1.1 Proje Çalışmasının Amacı ve Önemi.....	1
2. LİTERATÜR TARAMASI.....	2
3. MATERYAL VE YÖNTEM.....	4
3.1 Veri Setleri.....	4
3.1.1. The IQ-OTH/NCCD Lung Cancer Dataset	4
3.1.2. SPIE-AAPM-NCI Lung Nodule Classification Challenge Dataset.....	5
3.1.3. CT-Scan Images Dataset	7
3.2 Metotlar	7
3.2.1. CNN.....	8
3.2.2. VGG16	8
3.2.3. ResNet50	9
3.2.4. EfficientNetV2-S	10
3.2.5. SwinV2-CR-Small-224.....	12
3.2.6. Beit-Base	12
4. BULGULAR VE TARTIŞMA.....	14
5. SONUÇLAR.....	34

ÖZET

AKCİĞER BT GÖRÜNTÜLERİNİN MAKİNE ÖĞRENMESİ ALGORİTMALARIYLA ANALİZİ VE AKCİĞER KANSERİ TESPİTİ

Bu çalışma, akciğer bilgisayarlı tomografi (BT) görüntülerinin derin öğrenme temelli yaklaşımlar ile analiz edilerek akciğer kanserinin doğru ve güvenilir bir şekilde tespit edilmesini amaçlamaktadır. Bu kapsamda, çalışmada üç farklı açık erişimli veri seti kullanılmıştır: IQ-OTH/NCCD Lung Cancer Dataset, SPIE-AAPM-NCI Lung Nodule Classification Challenge Dataset ve CT-Scan Images Dataset.

Sınıflandırma aşamasında çeşitli derin öğrenme temelli mimariler tercih edilmiştir. Bu doğrultuda, temel Convolutional Neural Network (CNN) mimarisine ek olarak, önceden eğitilmiş ve literatürde sıklıkla kullanılan gelişmiş derin öğrenme modelleri olan VGG16, ResNet, EfficientNetV2-S, Swin Transformer ve BEiT Base değerlendirmeye alınmıştır.

Model performansları; doğruluk (accuracy), kesinlik (precision), duyarlılık (recall) ve F1-skoru gibi temel sınıflandırma metrikleri üzerinden analiz edilmiştir. Elde edilen bulgular, farklı derin öğrenme yaklaşımlarının akciğer kanseri teşhisindeki etkinliğini karşılaştırmalı olarak ortaya koymakta ve en uygun model mimarilerinin belirlenmesine katkı sunmaktadır.

Sonuç olarak, bu çalışma; derin öğrenme tabanlı yöntemlerin tıbbi görüntü analizi alanındaki yüksek potansiyelini ortaya koymakta ve yapay zekâ destekli yaklaşımların, akciğer kanseri gibi kritik hastalıkların erken tanı süreçlerinde önemli bir rol oynayabileceğini göstermektedir.

Mayıs, 2025

Öğrenciler

Sude Nur GÜLCAN

Karun ACAR

Elif TİLKİ

ABSTRACT

ANALYSIS OF LUNG CT IMAGES USING MACHINE LEARNING ALGORITHMS AND LUNG CANCER DETECTION

This study aims to accurately and reliably detect lung cancer by analyzing lung computed tomography (CT) images using deep learning-based approaches. In this context, three publicly available datasets were utilized: IQ-OTH/NCCD Lung Cancer Dataset, SPIE-AAPM-NCI Lung Nodule Classification Challenge Dataset, and CT-Scan Images Dataset.

In the classification stage, various deep learning-based architectures were employed. Specifically, in addition to the basic Convolutional Neural Network (CNN) architecture, several advanced and widely used pre-trained models were evaluated, including VGG16, ResNet, EfficientNetV2-S, Swin Transformer, and BEiT Base.

The performance of these models was evaluated using standard classification metrics such as accuracy, precision, recall, and F1-score. The experimental results provided a comparative analysis of the effectiveness of different deep learning models in detecting lung cancer and contributed to identifying the most suitable model architectures.

In conclusion, this study highlights the high potential of deep learning-based methods in the field of medical image analysis and demonstrates that AI-powered approaches can play a significant role in the early diagnosis of critical diseases such as lung cancer.

May, 2025

Students

Sude Nur GÜLCAN

Karun ACAR

Elif TILKİ

ŞEKİL LİSTESİ

Şekil 1. CNN Modelinin IQ-OTH/NCCD Lung Cancer Datasetindeki Sonuçları.....	15
Şekil 2. CNN Modeli IQ-OTH/NCCD Lung Cancer Dataseti Confusion Matrix Sonuçları	15
Şekil 3. CNN Modelinin SPIE-AAPM-NCI Datasetindeki Sonuçları	15
Şekil 4. CNN Modeli SPIE-AAPM-NCI Dataseti Confusion Matrix Sonuçları.....	16
Şekil 5. CNN Modeli CT-Scan Images Datasetindeki Sonuçları.....	16
Şekil 6. CNN CT-Scan Images Dataseti Confusion Matrix Sonuçları.....	17
Şekil 7. VGG16 Modelinin IQ-OTH/NCCD Lung Cancer Datasetindeki Sonuçları	18
Şekil 8. VGG16 Modelinin IQ-OTH/NCCD Dataseti Confusion Matrix Sonuçları.....	18
Şekil 9. VGG-16 Modelinin SPIE-AAPM-NCI Datasetindeki Sonuçları.....	19
Şekil 10. VGG16 Modelinin SPIE-AAPM-NCI Dataseti Confusion Matrix Sonuçları.....	19
Şekil 11. VGG-16 Modelinin CT-Scan Images Datasetindeki Sonuçları	20
Şekil 12. VGG16 Modelinin CT-Scan Images Dataseti Confusion Matrix Sonuçları	20
Şekil 13. Resnet50 Modelinin IQ-OTH/NCCD Lung Cancer Datasetindeki Sonuçları.....	21
Şekil 14. ResNet50 Modelinin IQ-OTH/NCCD Dataseti Confusion Matrix Sonuçları.....	21
Şekil 15. Resnet50 Modelinin SPIE-AAPM-NCI Datasetindeki Sonuçları.....	22
Şekil 16. ResNet50 Modelinin SPIE-AAPM-NCI Datasetindeki Confusion Matrix Sonuçları.....	22
Şekil 17. Resnet50 Modelinin CT-Scan Images Datasetindeki Sonuçları	22
Şekil 18. ResNet50 Modelinin CT-Scan Images Datasetindeki Confusion Matrix Sonuçları.....	23
Şekil 19. EfficientNetV2-S Modelinin IQ-OTH/NCCD Dataseti Sonuçları.....	24
Şekil 20. EfficientNetV2-S Modelinin IQ-OTH/NCCD Datasetindeki Confusion Matrix Sonuçları.....	24
Şekil 21. EfficientNetV2-S Modelinin SPIE-AAPM-NCI Datasetindeki Sonuçları.....	25

Şekil 22. EfficientNetV2-S Modelinin SPIE-AAPM-NCI Dataseti Confusion Matrix Sonuçları.....	25
Şekil 23. EfficientNetV2-S Modelinin CT-Scan Images Datasetindeki Sonuçları	26
Şekil 24. EfficientNetV2-S Modelinin CT-Scan Images Dataseti Confusion Matrix Sonuçları.....	26
Şekil 25. Swin Transformer Modelinin IQ-OTH/NCCD Datasetindeki Sonuçları	27
Şekil 26. Swin Transformer Modelinin IQ-OTH/NCCD Dataseti Confusion Matrix Sonuçları.....	28
Şekil 27. Swin Transformer Modelinin SPIE-AAPM-NCI Datasetindeki Sonuçları	28
Şekil 28. Swin Transformer Modelinin SPIE-AAPM-NCI Dataseti Confusion Matrix Sonuçları.....	29
Şekil 29. Swin Transformer Modelinin CT-Scan Images Datasetindeki Sonuçları	29
Şekil 30. Swin Transformer Modelinin CT-Scan Images Dataseti Confusion Matrix Sonuçları.....	30
Şekil 31. BEiT Base Modelinin IQ-OTH/NCCD Datasetindeki Sonuçları.....	31
Şekil 32. BEiT Base IQ-OTH/NCCD Datasetindeki Confusion Matrix Sonuçları	31
Şekil 33. BEiT Base Modelinin SPIE-AAPM-NCI Datasetindeki Sonuçları	32
Şekil 34. BEiT Base Modelinin SPIE-AAPM-NCI Datasetindeki Confusion Matrix Sonuçları.....	32
Şekil 35. BEiT Base Modelinin CT-Scan Images Datasetindeki Sonuçları.....	33
Şekil 36. BEiT Base CT-Scan Images Datasetindeki Confusion Matrix Sonuçları	33

TABLO LİSTESİ

Tablo 1. IQ-OTH/NCCD Datasetinin Modellere Göre Sonuçları.....	33
Tablo 2. SPIE-AAPM-NCI Datasetinin Modellere Göre Sonuçları.....	34
Tablo 3. CT-Scan Images Datasetinin Modellere Göre Sonuçları	34

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 arttırmaktadır [4][5]. Geleneksel tanı yöntemleri, radyologların bilgisayarlı tomografi (BT) ve manyetik rezonans (MR) 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 makine öğrenmesi algoritmaları aracılığıyla analiz edilmesi ve sınıflandırılması amaçlanmıştır. Bu bağlamda, makine öğrenmesi 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, makine öğrenmesi 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 makine öğrenmesi 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 radyoloji raporu oluşturulmasında kullanılan derin öğrenme teknikleri ve tıbbi görüntüleme uygulamaları incelenmiştir.

R2GenGPT modeli, radyoloji raporu üretiminde büyük dil modellerinin (LLM) potansiyelini optimize etmeye yönelik yapılan önemli bir çalışmadır. Wang ve arkadaşları bu modeli kullanarak görsel özellikleri dil modeli ile hizalayarak, düşük hesaplama karmaşıklığı ve yüksek performans sunmayı başarmıştır. MIMIC-CXR ve IU-Xray veri setlerinde test edilen R2GenGPT, görsel özellikleri sığ, derin ve delta hizalama stratejileriyle entegre ederek, BLEU, ROUGE-L, METEOR ve CIDEr gibi metriklerde üstün performans sergilemiştir. Model, yalnızca hafif bir modül eğitirken LLM parametrelerini dondurarak hızlı yakınsama sağlamış ve klinik etkinlikte diğer yöntemleri geride bırakmıştır. Bu, radyoloji raporu oluşturma otomasyonunun geleceği için önemli bir adımdır [7].

COVID-19 pandemisi ile birlikte, bu hastalığın tanısında kullanılan X-ray ve bilgisayarlı tomografi (BT) görüntülerine dayalı derin öğrenme modelleri de araştırılmaya başlanmıştır. Özellikle CT görüntüleri üzerinde yapılan çalışmalar, virüsün akciğerlerdeki etkilerini tespit etmede önemli bir araç olarak öne çıkmıştır. Derin öğrenme tabanlı CNN modelleri, özellikle COVID-19 tanısında başarılı sonuçlar elde etmiştir. Haryanto ve arkadaşlarının SCOV CNN modeli kullanarak yapmış olduğu çalışma, CT görüntülerinde yüksek doğruluk oranları ile performans sergileyerek web tabanlı bir sistemle entegre edilmiştir [8]. Bu tür yaklaşımlar, sağlık profesyonellerine tanı koyma süreçlerinde hızlı ve doğru kararlar alabilme yeteneği sunmaktadır.

Ayrıca, akciğer kanseri tespitine yönelik yapılan çalışmalarda, 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 [9]. 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 [10].

BT görüntülerinde akciğer kanserinin tespiti için segmentasyon ve sınıflandırma yöntemleri de önemli bir yer tutmaktadır. Özdet ve arkadaşının RIDER Lung-CT veri seti üzerinde yaptığı çalışmada, geleneksel eşikleme yöntemleri ve sütunsal tarama algoritmaları kullanılarak yüksek doğruluk oranları elde edilmiştir. Segmentasyon süreci, tümör tespiti için kritik bir öneme sahiptir ve doğru ROI (Region of Interest) çıkarımı, sınıflandırma başarısını doğrudan etkilemektedir [11]. Bu çalışmalar, görüntü işleme tekniklerinin kanser teşhisindeki rolünü vurgulamaktadır.

Son olarak, akciğer kanseri tespitinde kullanılan yeni bir Evrişimsel Sinir Ağı (ESA) modeli, ResNet tabanlı yapısı ve optimize edilmiş parametreleri ile %99 doğruluk oranına ulaşmıştır[6]. Erdoğan ve arkadaşlarının yapmış olduğu çalışma, daha az parametreyle yüksek doğruluk elde etmenin mümkün olduğunu ve klinik uygulamalarda kullanılabilecek güçlü bir model sunduğunu ortaya koymaktadır.

Bu literatür taraması, derin öğrenme ve görüntü işleme yöntemlerinin akciğer kanseri teşhisi ve radyoloji raporu oluşturulmasında nasıl önemli bir rol oynadığını açıkça göstermektedir. Bu alandaki çalışmalar, hem tanı doğruluğunu artırmakta hem de klinik uygulamaların etkinliğini desteklemektedir. Gelecekte yapılacak araştırmalarla bu yöntemlerin daha geniş veri setleri ve daha gelişmiş algoritmalar kullanılarak daha da optimize edilmesi beklenmektedir.

3. MATERYAL VE YÖNTEM

3.1 Veri Setleri

Bu çalışmada, derin öğrenme tabanlı sınıflandırma modellerinin performanslarının güvenilir ve genellenebilir bir biçimde değerlendirilebilmesi amacıyla K-Fold çapraz doğrulama yöntemi uygulanmıştır. Bu yöntem, veri kümesini k adet alt kümeye ayırarak her bir alt kümenin test ve eğitim verisi olarak ayrılması esasına dayanır. Böylece model, farklı veri bölünmeleri üzerinde test edilerek daha sağlam ve önyargısız bir performans değerlendirmesi gerçekleştirilmiş olur. Çalışma kapsamında $k = 5$ olarak belirlenmiş ve her veri seti beş katmana ayrılmıştır.

Her veri seti içerisinde yer alan sınıflar ayrı ayrı ele alınarak tüm görüntü örnekleri rastgele karıştırılmış ve eşit sayıda olacak şekilde beş alt kümeye ayrılmıştır. Her bir fold, %80 oranında eğitim ve %20 oranında test alt kümeleri oluşturularak yapılandırılmıştır.

Ayrıca, deneysel tekrarların tutarlılığını sağlamak ve rastlantısal değişkenlikleri kontrol altına almak amacıyla veri dağıtımında sabit bir rastgelelik tohumu (`random.seed(42)`) kullanılmıştır. Bu uygulama, modelin farklı çalıştırmalarda aynı veri dağılımı ile eğitilmesini ve karşılaştırmalı analizlerin sağlıklı bir şekilde yapılabilmesini mümkün kılmıştır.

3.1.1. 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 (CT) görüntüleri ile sağlıklı bireylerin görüntülerini içermektedir. Veri setinde toplamda 110 vakayı temsil eden 1190 CT tarama dilimi bulunmaktadır. Her bir tarama, değişen 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. CT 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[12].

3.1.2. 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.

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 betiği, 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.

3.1.3. 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 (BT) görüntüsünden oluşan bu veri kümesi üzerinde çeşitli ön işleme adımları uygulanmış; bu sayede görüntülerin derin öğrenme tabanlı sınıflandırma modellerine daha uygun hâle getirilmesi amaçlanmıştır. Görüntü işleme sürecinde OpenCV (cv2) temel kütüphane olarak kullanılırken, dosya işlemleri için os ve Pathlib, sayısal hesaplamalar için ise NumPy kütüphanesinden faydalanılmıştır.

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 marjine 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. 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.

3.2 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.

3.2.1. CNN

Bu çalışmada, tıbbi görüntü sınıflandırma problemine yönelik özel olarak tasarlanmış ve optimize edilmiş bir Convolutional Neural Network (CNN) mimarisi kullanılmıştır. Model, 224x224 boyutlarında RGB görüntüleri giriş olarak alacak şekilde yapılandırılmıştır ve toplamda 13 konvolüsyonel, 3 tam bağlantılı katmandan oluşmaktadır. İlk katmanda 128 filtreli 8x8 kernel yapısı kullanılmış, bunu 256 filtreli 5x5 kernel içeren ikinci katman takip etmiştir. Sonraki katmanlar 256 ve 512 filtreli yapılarla derinleştirilmiş, çıkış katmanında ise sınıf sayısı kadar (3) nöron yer almıştır. Modelde her konvolüsyon katmanından sonra batch normalization uygulanarak öğrenme süreci stabilize edilmiştir. Tam bağlantılı katmanlarda overfitting riskini azaltmak amacıyla %50 oranında dropout kullanılmıştır. Optimizasyon işlemleri, öğrenme oranı $1e-4$ olarak belirlenmiş Adamax algoritması ile yürütülmüştür. Modelin eğitimi sırasında erken durdurma (early stopping) mekanizması aktif hale getirilmiş; böylece doğrulama kaybında 50 epoch boyunca iyileşme gözlemlenmezse eğitim süreci sonlandırılmıştır. Ayrıca eğitim sürecinin bellek verimliliği açısından optimize edilmesi için GPU üzerinde karma hassasiyetli hesaplama (mixed precision training) uygulanmıştır. Modelin değerlendirilmesinde accuracy, loss, precision, recall, F1-score gibi metrikler kullanılmış; ayrıca her fold için confusion matrix ve classification report oluşturulmuştur. Verilerin işlenmesinde MobileNetV2 normalizasyon değerleri esas alınmış, veri yükleme işlemleri ise ImageDataGenerator aracılığıyla gerçekleştirilmiştir. 5 katlı çapraz doğrulama yöntemi uygulanarak modelin genellenebilirlik performansı test edilmiş ve her bir fold için ayrı eğitim gerçekleştirilmiştir. Eğitim süreçleri sırasında GPU bellek yönetimi dikkatle yürütülmüş ve her fold sonrası bellek temizliği yapılmıştır. Sonuçlar zaman damgalı dosyalarla birlikte kayıt altına alınmış, grafiksel analizler ve metrik raporları ile desteklenmiştir.

3.2.2. VGG16

Bu çalışmada kullanılan bir diğer derin öğrenme yöntemi, 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 konvolüsyonel sinir ağı (CNN) modelidir. Mimari, 3×3 boyutundaki küçük konvolüsyon filtrelerini çok katmanlı bir şekilde istifleyerek daha karmaşık özelliklerin hiyerarşik biçimde öğrenilmesini sağlamaktadır [16]. 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.

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 (kenar, köşe, doku vb.) korunması sağlanırken, çalışma veri kümesine özgü yüksek seviyeli özelliklerin öğrenilmesi mümkün olmuştur.

Veri setinde yer alan tüm CT 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.

Modelin eğitiminde erken durdurma (early stopping) stratejisi uygulanmış, eğitim sürecinde doğrulama kaybı (validation loss) belirli bir süre 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 mini-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 başarımı 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/doğruluk eğrileri grafiksel olarak sunulmuştur.

3.2.3. ResNet50

ResNet-50, derin öğrenme alanında artık (residual) bağlantılar ve geleneksel bağlantılar kullanan gelişmiş bir sinir ağı mimarisidir. Artık bağlantılar, bir konvolüsyon katmanının çıktısını, daha sonraki evrişim katmanına doğrudan ileterek ağı derinliğine rağmen öğrenme sürecini kolaylaştırır ve özellik haritalarının daha etkili bir şekilde çıkarılmasını sağlar

Model, beş temel aşamadan oluşmaktadır. İlk aşamada, 7x7 boyutunda 64 filtreli bir konvolüsyon katmanını takiben 3x3 boyutunda ve 2 adımda uygulanan bir maksimum havuzlama işlemi yer alır. İkinci aşamada, sırasıyla 1x1 boyutlu 64 filtreli, 3x3 boyutlu 64 filtreli ve 1x1 boyutlu 256 filtreli üçlü konvolüsyon blokları bulunur. Üçüncü aşamada ise bu yapı 128 ve 512 filtreli konvolüsyon katmanları ile devam eder. Dördüncü aşamada 256 filtreli 1x1 ve 3x3 konvolüsyon katmanları ile birlikte 1024 filtreli 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. ResNet-50, özellikle görüntü sınıflandırma gibi görevlerde derin yapısına rağmen yüksek başarı oranları ile öne çıkmaktadır[13].

Bu çalışmada, beş katlı çapraz doğrulama yöntemi kullanılarak modelin genellenebilirlik performansı değerlendirilmiştir. Bu kapsamda, önceden ImageNet veri kümesi üzerinde eğitilmiş olan ResNet-50 mimarisi transfer öğrenme yöntemiyle yeniden yapılandırılmış; yalnızca tam bağlantılı (fully connected) katman güncellenerek sınıflandırma işlemi gerçekleştirilmiştir. Veri kümesindeki görüntüler standart boyutlara getirilerek tensör formatına dönüştürülmüştür.

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; karışıklık 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 ve etkinliği akademik ölçütlerle kapsamlı bir biçimde değerlendirilmiştir.

3.2.4. EfficientNetV2-S

EfficientNetV2 ailesi, parametrik verimlilik ile yüksek doğruluk oranlarını bir araya getiren mimari yapısıyla öne çıkmaktadır. 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.

Bu çalışmada, EfficientNetV2-S modeli kullanılarak veri setleri üzerinde sınıflandırma işlemi gerçekleştirilmiş ve model performansı, beş katlı çapraz doğrulama (5-Fold Cross Validation) yöntemiyle sistematik olarak değerlendirilmiştir. Veri seti, modele uygun hale getirilmek üzere EfficientNetV2-S mimarisi için optimize edilmiş ön işleme adımlarına tabi tutulmuştur. Bu kapsamda, EfficientNet_V2_S_Weights.DEFAULT ağırlıkları kullanılarak normalize edilmiş dönüşümler (transforms.Compose) uygulanmıştır.

Modelin mimarisinde, önceden eğitilmiş ağırlıklar ile başlatılan özellik çıkarıcı (feature extractor) katmanlar dondurulmuş, yalnızca sınıflandırma katmanı güncellenmiştir. Bu katmanda, veri setinde yer alan sınıf sayısına uygun olacak şekilde tam bağlantılı (fully connected) bir çıkış katmanı yapılandırılmıştır. Modelin eğitimi, CrossEntropyLoss kayıp fonksiyonu ve Adamax optimizasyon algoritması kullanılarak gerçekleştirilmiştir. Aşırı öğrenmeyi engellemek amacıyla erken durdurma (early stopping) kriteri uygulanmış, 50 epoch boyunca doğrulama kaybında iyileşme gözlemlenmemesi durumunda eğitim süreci sonlandırılmıştır.

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 (F1-score) gibi temel sınıflandırma metrikleriyle değerlendirilmiş; ayrıca sınıflandırma raporları ve karışıklık

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 karışıklık matrisi ile modelin genel genellenebilirlik performansı ortaya konmuştur. Ortalama doğruluk oranı hesaplanarak tüm eğitim sürecinin istatistiksel başarımı yorumlanmış, modelin tutarlı ve güvenilir bir sınıflandırma yetisi sunduğu gösterilmiştir.

3.2.5. 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. 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. "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 (CNN) 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 çıkmakta; bu yönüyle hem akademik araştırmalarda hem de endüstriyel uygulamalarda tercih edilen etkin bir model mimarisi olarak literatürde yer almaktadır.

Bu model kullanılarak 5 katlı çapraz doğrulama (K-Fold Cross Validation) yöntemi ile kapsamlı bir eğitim ve değerlendirme süreci gerçekleştirilmiştir. Eğitim ve test verileri, belirlenen klasör yapısına göre her bir fold için torchvision.datasets.ImageFolder sınıfı aracılığıyla yüklenmiş ve transforms.Compose ile normalize edilerek modele uygun hale getirilmiştir.

Model olarak, timm kütüphanesi üzerinden önceden eğitilmiş (pretrained) 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 (early stopping) mekanizması uygulanarak aşırı öğrenmenin (overfitting) önüne geçilmiştir. 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 (classification_report) ve karışıklık matrisi (confusion_matrix) görselleştirilerek kayıt altına alınmıştır. Eğitim sonunda her bir fold için eğitilen modeller .pt uzantılı dosyalar halinde kaydedilmiş ve tüm fold'ların çıktıları birleştirilerek ortalama performans metrikleri elde edilmiştir. Bu ortalama metrikler doğrultusunda modelin genellenabilirlik performansı değerlendirilmiş ve eğitim sürecinin genel kayıp ve doğruluk grafikleri görselleştirilmiştir.

3.2.6. Beit-Base

Bu çalışmada, sınıflandırma görevini gerçekleştirmek amacıyla BEiT (Bidirectional Encoder representation from Image Transformers) base modelinden de yararlanılmıştır. Modelin, ImageNet veri kümesi üzerinde önceden eğitilmiş (pretrained) versiyonu kullanılmış olup, bu sayede yüksek düzeyde genelleme yeteneği sağlanmıştır. Projedeki sınıf sayısına uyum sağlayabilmesi amacıyla modelin son katmanı yeniden yapılandırılmıştır.

Modelin eğitimi ve değerlendirilmesi sürecinde, 5-katlı çapraz doğrulama (5-fold cross-validation) yöntemi uygulanmıştır. Her bir fold için eğitim ve test veri kümeleri torchvision.datasets.ImageFolder yapısı kullanılarak tanımlanmış; görseller 224x224 boyutlarına yeniden ölçeklendirilmiş ve normalize edilmiştir. Normalize işlemi, BEiT modelinin giriş beklentilerine uygun şekilde $\text{mean}=[0.485, 0.456, 0.406]$ ve $\text{std}=[0.229, 0.224, 0.225]$ değerleriyle gerçekleştirilmiştir.

Eğitim sürecinde kayıp fonksiyonu olarak cross-entropy loss kullanılmış; optimizasyon işlemleri Adamax algoritması ile yürütülmüştür. Her bir eğitim süreci, 50 epoch üzerinden gerçekleştirilmiş olup, early stopping yöntemiyle gereksiz tekrarların önüne geçilmiştir.

Her epoch sonunda eğitim ve doğrulama kayıpları ile doğruluk değerleri kayıt altına alınmıştır.

Model performansı; accuracy, precision, recall ve F1-score metrikleri ile değerlendirilmiştir. Ayrıca, sınıflandırma başarısını görselleştirmek amacıyla her fold için confusion matrix ayrı ayrı oluşturulmuş ve kaydedilmiştir. Buna ek olarak, classification_report fonksiyonu ile her bir sınıfa ait ayrıntılı performans raporları metin dosyası olarak arşivlenmiştir.

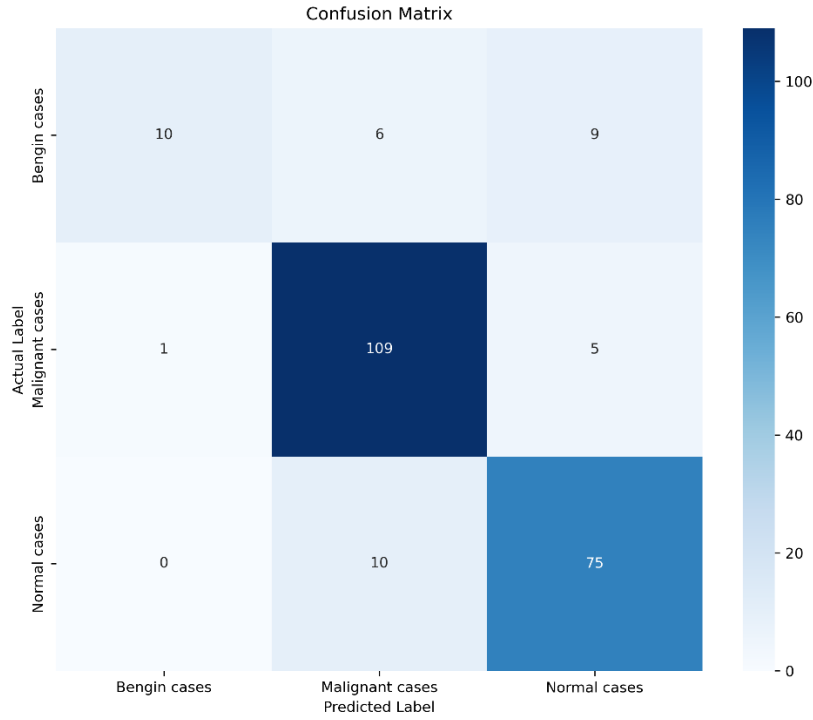
Her fold için eğitilen modeller .pt formatında kaydedilmiş; tüm fold'ların ortalama performans değerleri hesaplanarak genel başarı düzeyi ortaya konmuştur. Fold'ların birleştirilmesiyle genel bir confusion matrix görselleştirilmiş; ayrıca ortalama loss ve accuracy grafiklerinin analizi gerçekleştirilmiştir. Bu sayede, modelin genel performansı daha kapsamlı şekilde değerlendirilmiştir.

4. BULGULAR VE TARTIŞMA

Modellerin performans analizi incelendiğinde, farklı veri setlerinin model türleri üzerindeki sonuçları belirgin bir şekilde gözlemlenmiştir. CNN modeli, üç farklı veri seti üzerinde test edilmiş ve elde edilen yüzdesel sonuçlar doğrultusunda değerlendirildiğinde, modelin genel olarak yüksek doğrulukla çalıştığı görülmüştür. IQ-OTH/NCCD Lung Cancer Datasetinde modelin genel doğruluk oranı %86 olarak ölçülmüştür. Özellikle Malignant (kötü huylu) vakalarda %87 precision, %95 recall ve %91 f1-score ile yüksek başarı sağlanmıştır. Normal vakalarda %84 precision, %88 recall ve %86 f1-score elde edilirken, Benign (iyi huylu) vakalarda recall oranı %40'a kadar düşmüş ve bu sınıfta modelin örnekleri tanımada zorlandığı görülmüştür. SPIE-AAPM-NCI Datasetinde model %92 doğruluk oranına ulaşmış, Benign sınıfında %97 precision, %84 recall ve %90 f1-score; Malignant sınıfında ise %88 precision, %98 recall ve %93 f1-score elde etmiştir. Bu veri setinde model, her iki sınıfta da dengeli ve yüksek performans göstermiştir. CT-Scan Images Datasetinde ise genel doğruluk oranı %66'da kalmıştır. Cancerous (kanserli) sınıfında model %66 precision, %100 recall ve %79 f1-score ile yüksek başarı gösterirken, Non-Cancerous (kanser olmayan) sınıfında precision, recall ve f1-score değerleri %0 olarak kalmış ve model bu sınıfta hiçbir doğru tahmin gerçekleştirememiştir.

	precision	recall	f1-score	support
Bengin cases	0.91	0.40	0.56	25
Malignant cases	0.87	0.95	0.91	115
Normal cases	0.84	0.88	0.86	85
accuracy			0.86	225
macro avg	0.87	0.74	0.78	225
weighted avg	0.87	0.86	0.85	225

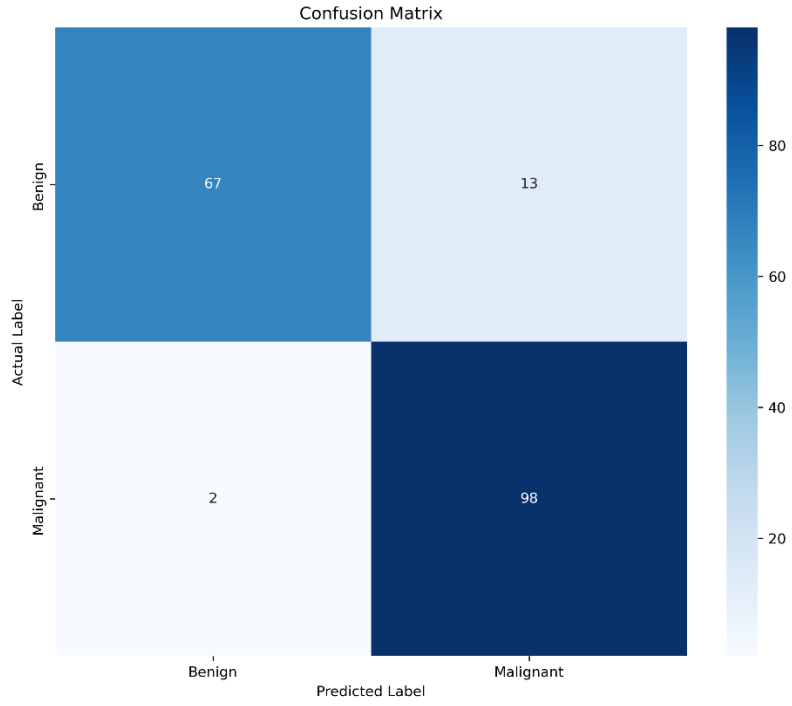
Şekil 1. CNN Modelinin IQ-OTH/NCCD Lung Cancer Datasetindeki Sonuçları



Şekil 2. CNN Modeli IQ-OTH/NCCD Lung Cancer Dataseti Confusion Matrix Sonuçları

	precision	recall	f1-score	support
Benign	0.97	0.84	0.90	80
Malignant	0.88	0.98	0.93	100
accuracy			0.92	180
macro avg	0.93	0.91	0.91	180
weighted avg	0.92	0.92	0.92	180

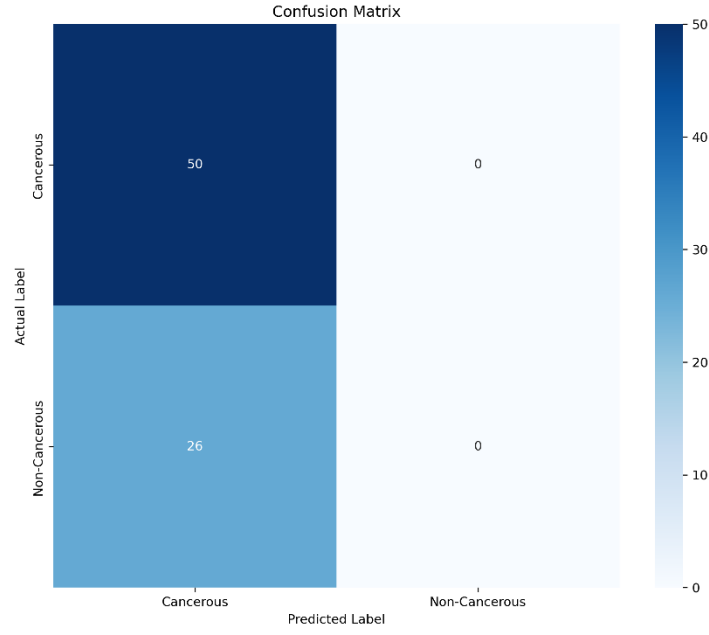
Şekil 3. CNN Modelinin SPIE-AAPM-NCI Datasetindeki Sonuçları



Şekil 4. CNN Modeli SPIE-AAPM-NCI Dataseti Confusion Matrix Sonuçları

	precision	recall	f1-score	support
Cancerous	0.66	1.00	0.79	50
Non-Cancerous	0.00	0.00	0.00	26
accuracy			0.66	76
macro avg	0.33	0.50	0.40	76
weighted avg	0.43	0.66	0.52	76

Şekil 5. CNN Modeli CT-Scan Images Datasetindeki Sonuçları



Şekil 6. CNN CT-Scan Images Dataseti Confusion Matrix Sonuçları

Bir diğer çalışmada, önceden eğitilmiş VGG16 derin öğrenme mimarisi kullanılarak geliştirilen akciğer görüntü sınıflandırma modeli, üç farklı veri seti üzerinde test edilmiş ve doğruluk (accuracy), duyarlılık (recall), kesinlik (precision) ve F1-skoru gibi temel sınıflandırma metrikleri aracılığıyla performansı değerlendirilmiştir. Elde edilen bulgular, modelin genel başarısının yanı sıra sınıf bazlı ayırım gücüne dair önemli ipuçları sunmuştur.

IQ-OTH/NCCD Lung Cancer Datasetinde model %89.78 doğruluk sağlamıştır. Malignant (kötü huylu) vakalarda %97.39 gibi yüksek oranlarda duyarlılık ve kesinlik elde edilmiştir. Normal sınıf için de benzer şekilde yüksek metrikler gözlenmiştir. Ancak benign (iyi huylu) sınıfta duyarlılık %28.00 seviyesinde kalmış ve modelin bu sınıfı tanımada yetersiz kaldığı görülmüştür. Bu durum, sınıf dengesizliği veya görsel benzerliklerden kaynaklanıyor olabilir.

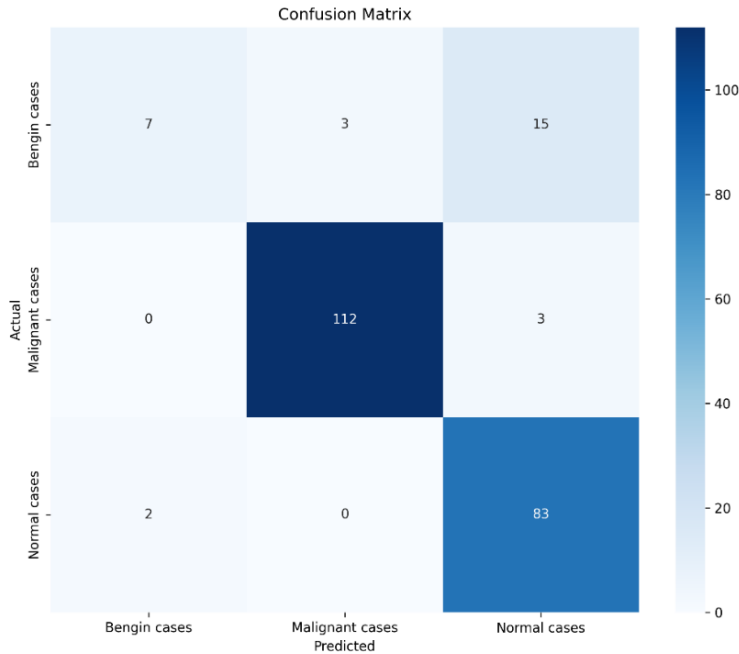
SPIE-AAPM-NCI Datasetinde model %90.00 doğruluk ile daha dengeli bir performans sergilemiştir. Benign ve malignant sınıflarında duyarlılık ve kesinlik değerleri sırasıyla %88.75 ve %91.00 seviyelerindedir. Model, her iki sınıfı da başarılı şekilde ayırt edebilmiştir.

CT-Scan Images Datasetinde doğruluk %96.05'e ulaşarak en yüksek başarı sağlanmıştır. Cancerous (kanserli) sınıfta %94.00 duyarlılık ve %100.00 kesinlik; non-cancerous (kanserli olmayan) sınıfta ise %100.00 duyarlılık ve %89.66 kesinlik elde edilmiştir. Bu sonuçlar, modelin özellikle yanlış negatifleri azaltmada etkili olduğunu ve klinik açıdan kritik vakaları doğru tanıyabildiğini göstermektedir.

Bu sonuçlar, VGG-16 mimarisinin özellikle dengeli ve yeterli veri ile desteklendiğinde yüksek doğrulukta sınıflandırma yapabildiğini; ancak bazı sınıflarda veri dengesizliğinin model performansını olumsuz etkileyebileceğini göstermektedir.

	precision	recall	f1-score	support
Bengin cases	0.7778	0.2800	0.4118	25
Malignant cases	0.9739	0.9739	0.9739	115
Normal cases	0.8218	0.9765	0.8925	85
accuracy			0.8978	225
macro avg	0.8578	0.7435	0.7594	225
weighted avg	0.8946	0.8978	0.8807	225

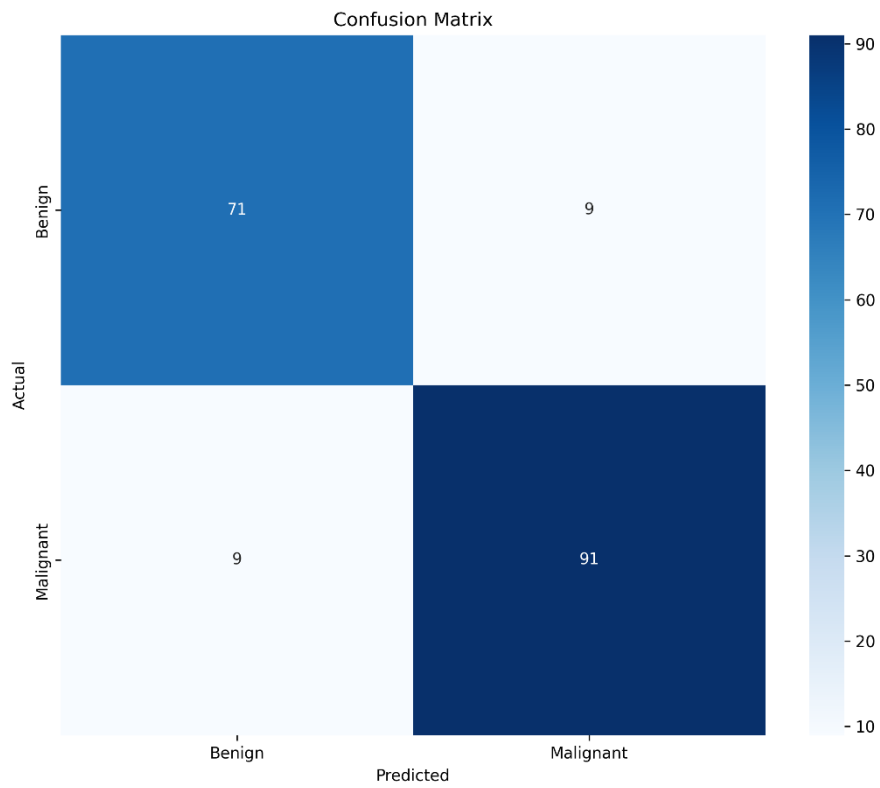
Şekil 7. VGG16 Modelinin IQ-OTH/NCCD Lung Cancer Datasetindeki Sonuçları



Şekil 8. VGG16 Modelinin IQ-OTH/NCCD Dataseti Confusion Matrix Sonuçları

	precision	recall	f1-score	support
Benign	0.8875	0.8875	0.8875	80
Malignant	0.9100	0.9100	0.9100	100
accuracy			0.9000	180
macro avg	0.8987	0.8987	0.8987	180
weighted avg	0.9000	0.9000	0.9000	180

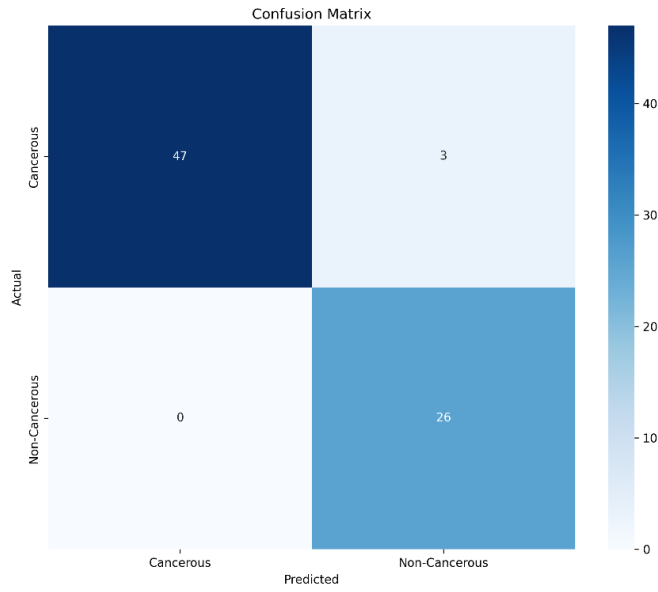
Şekil 9. VGG-16 Modelinin SPIE-AAPM-NCI Datasetindeki Sonuçları



Şekil 10. VGG16 Modelinin SPIE-AAPM-NCI Dataseti Confusion Matrix Sonuçları

	precision	recall	f1-score	support
Cancerous	1.0000	0.9400	0.9691	50
Non-Cancerous	0.8966	1.0000	0.9455	26
accuracy			0.9605	76
macro avg	0.9483	0.9700	0.9573	76
weighted avg	0.9646	0.9605	0.9610	76

Şekil 11. VGG-16 Modelinin CT-Scan Images Datasetindeki Sonuçları



Şekil 12. VGG16 Modelinin CT-Scan Images Dataseti Confusion Matrix Sonuçları

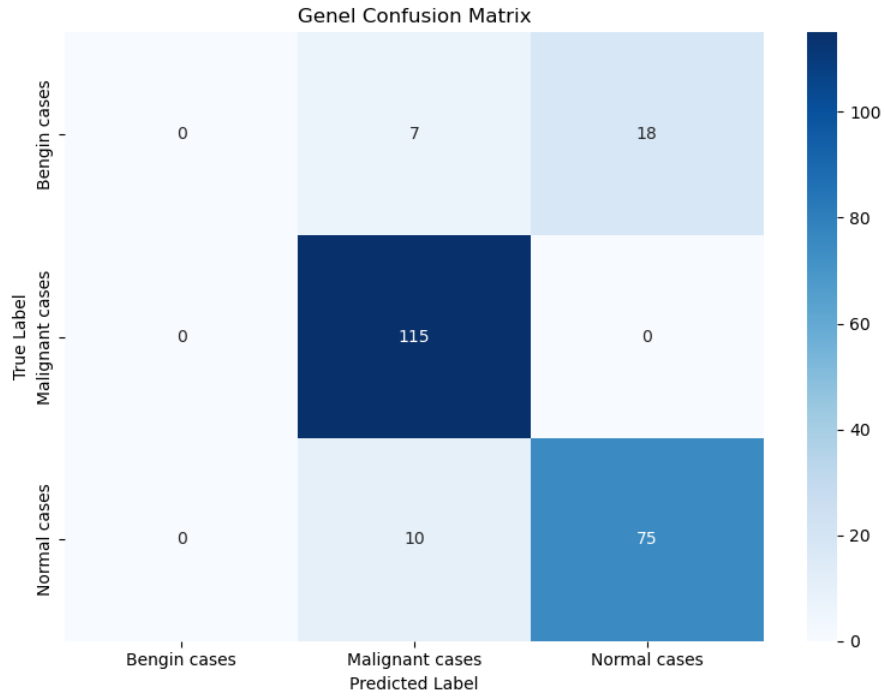
Bu ResNet mimarisinde ise diğer modellerde olduğu gibi üç farklı veri kümesi üzerinde beş katlı çapraz doğrulama yöntemi ile eğitilerek elde edilen performans metrikleri analiz edilmiştir. The IQ-OTH/NCCD lung cancer dataset veri kümesi üzerinde gerçekleştirilen eğitim sonucunda modelin doğruluk (accuracy) değeri %84,44, hassasiyet (precision) değeri 0.5595, duyarlılık (recall) değeri 0.6275 ve F1 skoru 0.5911 olarak elde edilmiştir. Bu sonuçlar, modelin genel olarak başarılı bir sınıflama performansı sergilediğini göstermektedir. Buna karşılık, SPIE-AAPM-NCI Lung Nodule Classification Challenge Dataset veri kümesi üzerinde elde edilen doğruluk değeri %54,44 ile oldukça düşüktür. Ayrıca, bu veri kümesinde precision, recall ve F1 skoru sırasıyla 0.4547, 0.5175 ve 0.4529

olarak kaydedilmiştir. Bu durum, ResNet mimarisinin bu veri kümesine karşı sınırlı genelleme yeteneğine sahip olduğunu göstermektedir. Öte yandan, CT-Scan Images Dataset veri kümesinde doğruluk %79,17, precision 0.6879, recall 0.7000 ve F1 skoru 0.6693 olarak elde edilmiştir. Genel olarak değerlendirildiğinde, ResNet mimarisinin veri kümesinin özelliklerine bağlı olarak sınıflama başarısında anlamlı farklılıklar gösterdiği söylenebilir.

Ortalama Metrikler (5 Fold):

Accuracy: 84.44%
Precision: 0.5595
Recall: 0.6275
F1 Score: 0.5911

Şekil 13. Resnet50 Modelinin IQ-OTH/NCCD Lung Cancer Datasetindeki Sonuçları

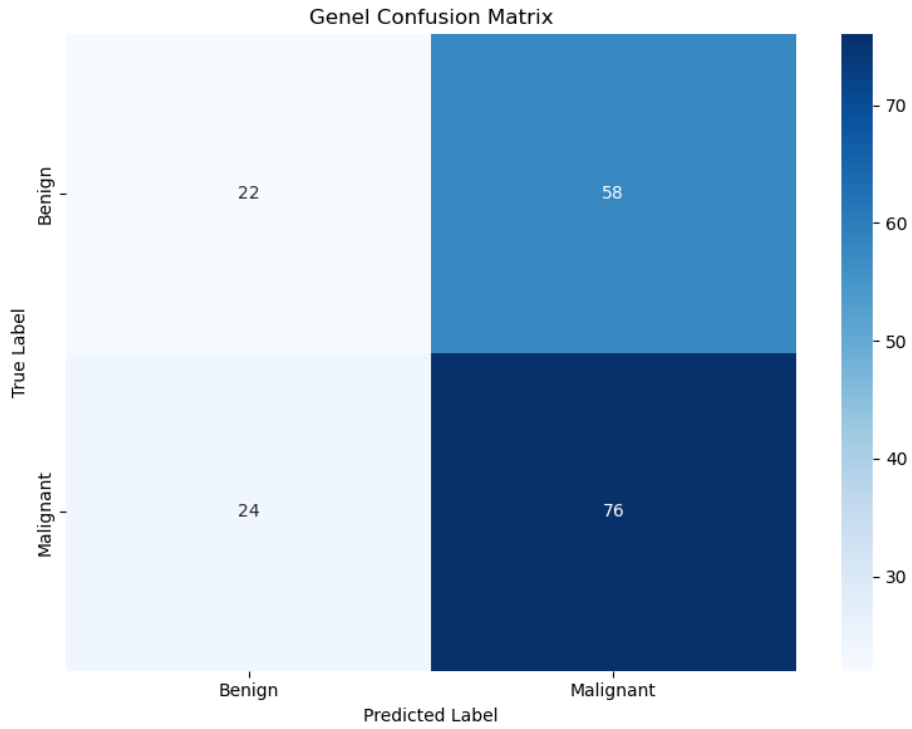


Şekil 14. ResNet50 Modelinin IQ-OTH/NCCD Dataseti Confusion Matrix Sonuçları

Ortalama Metrikler (5 Fold):

Accuracy: 54.44%
Precision: 0.4547
Recall: 0.5175
F1 Score: 0.4529

Şekil 15. Resnet50 Modelinin SPIE-AAPM-NCI Datasetindeki Sonuçları

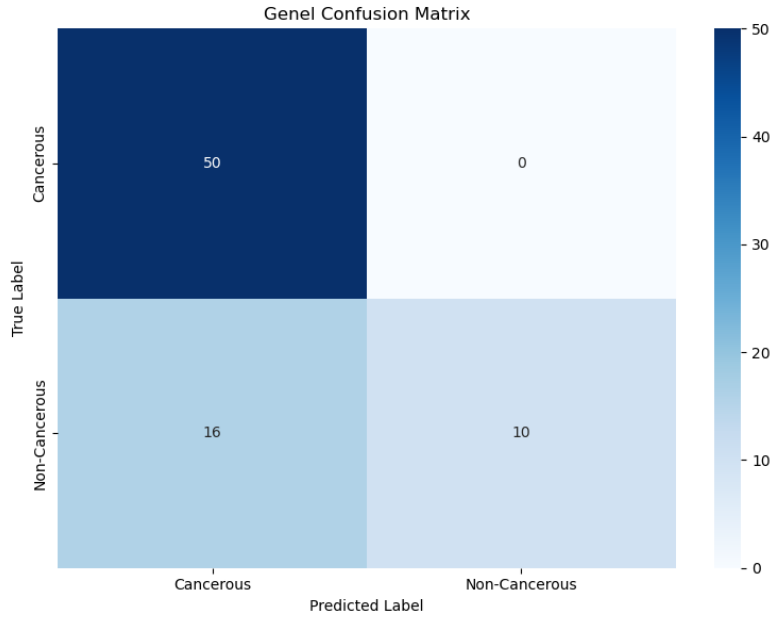


Şekil 16. ResNet50 Modelinin SPIE-AAPM-NCI Datasetindeki Confusion Matrix Sonuçları

Ortalama Metrikler (5 Fold):

Accuracy: 79.17%
Precision: 0.6879
Recall: 0.7000
F1 Score: 0.6693

Şekil 17. Resnet50 Modelinin CT-Scan Images Datasetindeki Sonuçları



Şekil 18. ResNet50 Modelinin CT-Scan Images Datasetindeki Confusion Matrix Sonuçları

Ayrıca, EfficientNetV2-S derin öğrenme mimarisi kullanılarak geliştirilen akciğer görüntü sınıflandırma modelinde ise yine üç farklı veri seti üzerindeki performansı değerlendirilmiştir. Modelin başarımı; doğruluk (accuracy), kesinlik (precision), duyarlılık (recall) ve F1-skoru gibi temel sınıflandırma metrikleriyle ölçülmüş, elde edilen sonuçlar sınıf ayırım gücü ve genel genellenebilirlik açısından analiz edilmiştir.

Modelin genel doğruluk oranları, sırasıyla %83.56, %56.11 ve %53.95 olarak hesaplanmıştır. Bu sonuçlar, EfficientNetV2-S'in The IQ-OTH/NCCD lung cancer dataset veri setinde makul düzeyde bir başarı gösterdiğini ancak diğer iki veri setinde özellikle sınıf dengesizliği veya örnek benzerliği gibi faktörler nedeniyle sınıflandırma performansının önemli ölçüde düştüğünü ortaya koymaktadır. Özellikle SPIE-AAPM-NCI Lung Nodule Classification Challenge Dataset ve CT-Scan Images Dataset veri setlerinde hem doğruluk hem de sınıf bazlı metriklerde gözlemlenen düşüşler, modelin sınıflar arası ayırım konusunda yetersiz kaldığını göstermektedir.

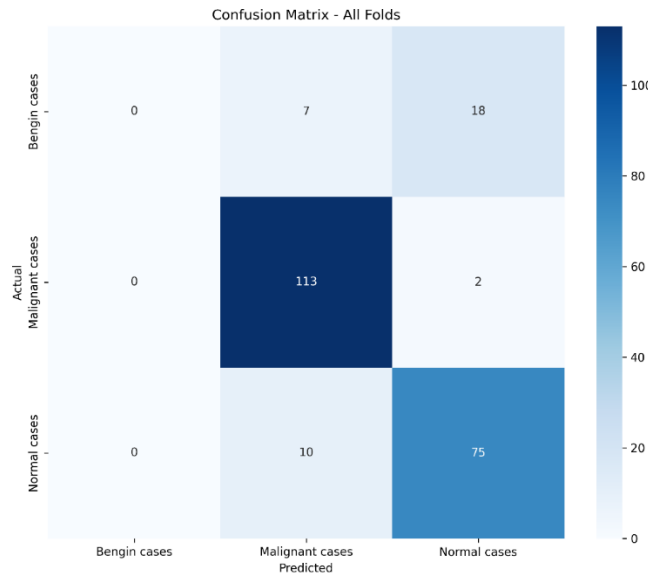
The IQ-OTH/NCCD lung cancer dataset veri setinde malign vakaların yüksek doğrulukla sınıflandırılması olumlu bir bulgu olarak öne çıkarken, benign sınıfındaki başarısızlık, veri setindeki örnek dengesizliğinin ve ayırt edici özellik eksikliğinin etkilerini göstermektedir.

Diğer veri setlerinde ise precision, recall ve F1-skoru gibi metriklerin çoğunlukla %50–60 aralığında seyretmesi, EfficientNetV2-S’in genelleme kapasitesinin belirli durumlarda sınırlı olduğunu işaret etmektedir.

Sonuç olarak, EfficientNetV2-S modeli sınıflandırma görevini genel anlamda yerine getirebilse de, özellikle karmaşık veya dengesiz veri setlerinde performans düşüşü gözlenmiştir. Bu durum, modelin daha etkin öğrenme stratejilerine, veri ön işleme tekniklerine ve dengeli veri kümelerine ihtiyaç duyduğunu göstermektedir. Gelecek çalışmalarda, transfer öğrenme süreçlerinin optimize edilmesi, sınıf dengesizliklerinin giderilmesi ve veriye özgü mimari uyarlamaların yapılması modelin başarımını artırmak adına önemli adımlar olacaktır.

	precision	recall	f1-score	support
Bengin cases	0.0000	0.0000	0.0000	25
Malignant cases	0.8692	0.9826	0.9224	115
Normal cases	0.7895	0.8824	0.8333	85
accuracy			0.8356	225
macro avg	0.5529	0.6217	0.5853	225
weighted avg	0.7425	0.8356	0.7863	225

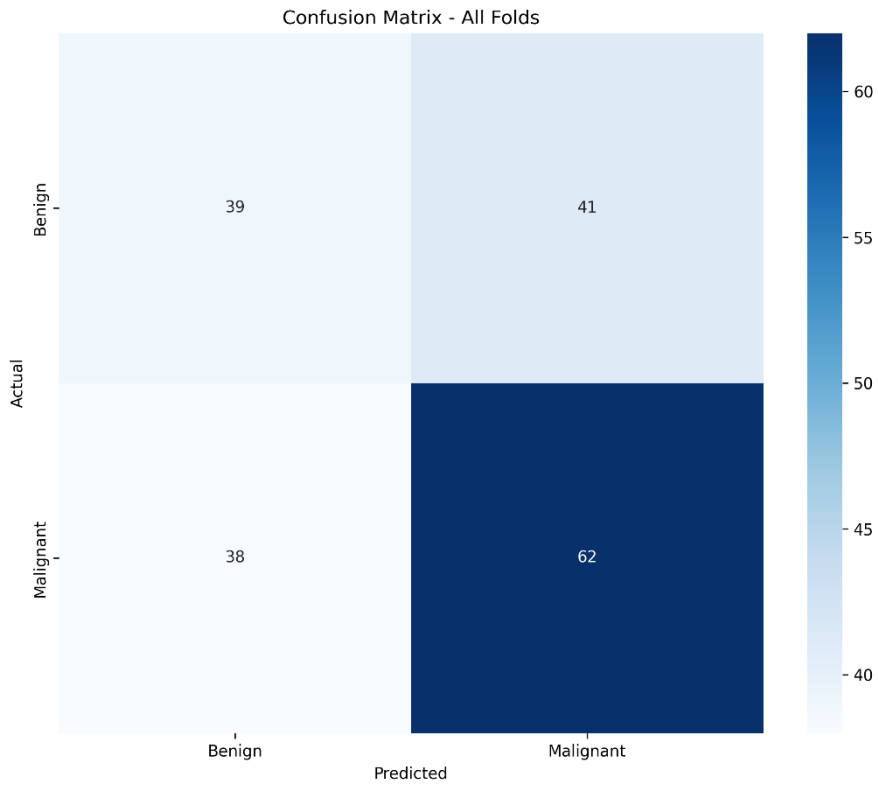
Şekil 19. EfficientNetV2-S Modelinin IQ-OTH/NCCD Dataseti Sonuçları



Şekil 20. EfficientNetV2-S Modelinin IQ-OTH/NCCD Datasetindeki Confusion Matrix Sonuçları

	precision	recall	f1-score	support
Benign	0.5065	0.4875	0.4968	80
Malignant	0.6019	0.6200	0.6108	100
accuracy			0.5611	180
macro avg	0.5542	0.5537	0.5538	180
weighted avg	0.5595	0.5611	0.5602	180

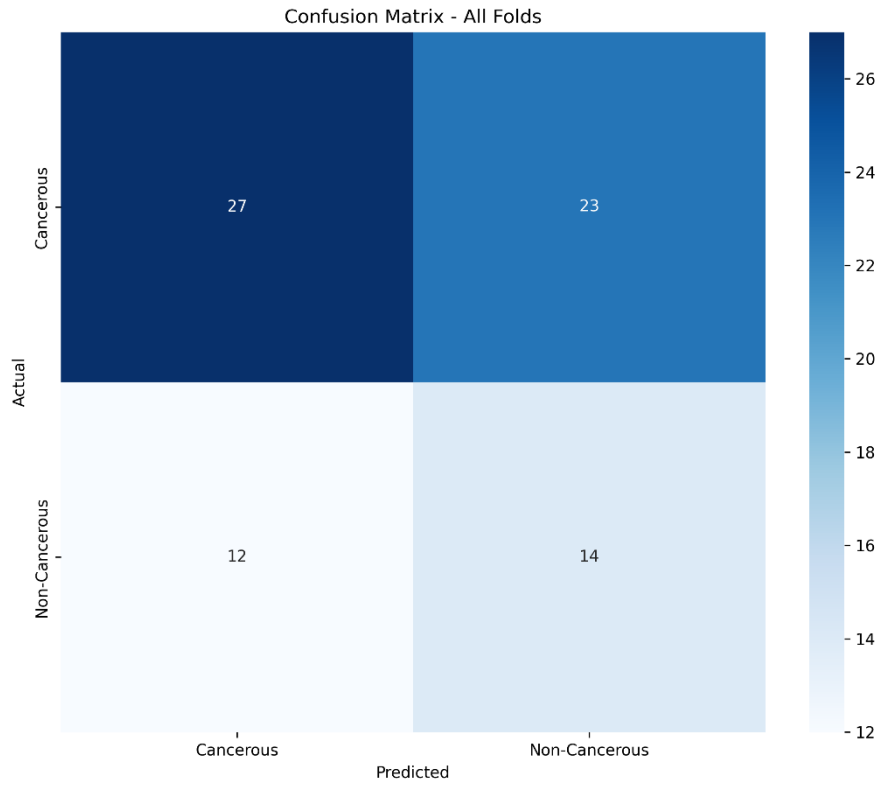
Şekil 21. EfficientNetV2-S Modelinin SPIE-AAPM-NCI Datasetindeki Sonuçları



Şekil 22. EfficientNetV2-S Modelinin SPIE-AAPM-NCI Dataseti Confusion Matrix Sonuçları

	precision	recall	f1-score	support
Cancerous	0.6923	0.5400	0.6067	50
Non-Cancerous	0.3784	0.5385	0.4444	26
accuracy			0.5395	76
macro avg	0.5353	0.5392	0.5256	76
weighted avg	0.5849	0.5395	0.5512	76

Şekil 23. EfficientNetV2-S Modelinin CT-Scan Images Datasetindeki Sonuçları



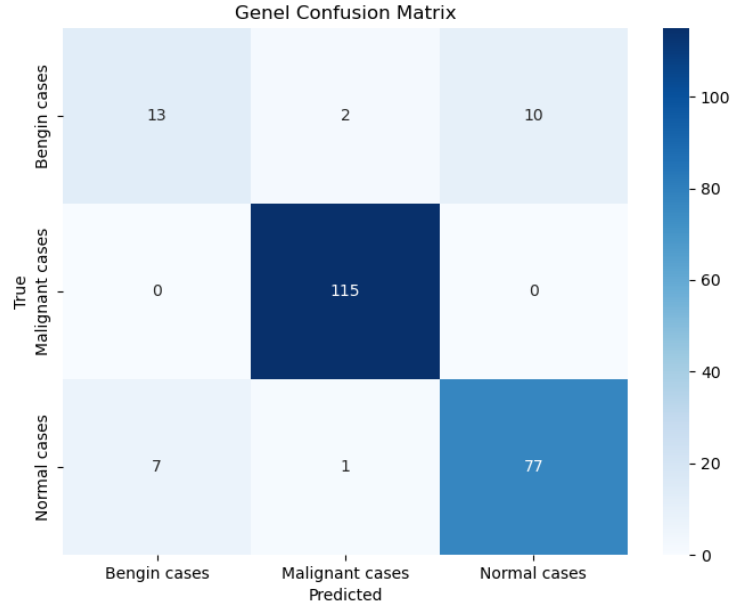
Şekil 24. EfficientNetV2-S Modelinin CT-Scan Images Dataseti Confusion Matrix Sonuçları

Bu çalışmada, Swin Transformer mimarisi üç farklı veri kümesi üzerinde beş katlı çapraz doğrulama yöntemi ile eğitilerek modelin sınıflama performansı değerlendirilmiştir. The IQ-OTH/NCCD lung cancer dataset veri kümesi üzerinde elde edilen ortalama doğruluk (accuracy) değeri %91,11 olup; precision, recall ve F1 skoru sırasıyla 0.9086, 0.8086 ve 0.8027 olarak hesaplanmıştır. Bu değerler, modelin yüksek doğruluk oranına karşın bazı sınıflarda hata yapma eğiliminin varlığına işaret etmektedir. SPIE-AAPM-NCI Lung Nodule Classification Challenge Dataset veri kümesinde ise doğruluk yine %91,11 olarak gerçekleşmiş, precision değeri 0.9196, recall 0.9050 ve F1 skoru 0.9086 olarak ölçülmüştür. Bu sonuçlar, modelin hem olumlu sınıfları doğru tespit etme yeteneğinin hem de genel başarı düzeyinin oldukça yüksek olduğunu göstermektedir. Öte yandan, CT-Scan Images Dataset veri kümesi üzerinde elde edilen sonuçlar dikkat çekici düzeydedir. Tüm metriklerde — doğruluk, precision, recall ve F1 skoru — %100 başarı sağlanmıştır. Bu durum, modelin söz konusu veri kümesi üzerinde hatasız bir sınıflama gerçekleştirdiğini ve öğrenme sürecinde ideal bir genelleme yeteneği sergilediğini ortaya koymaktadır. Bu bulgular, Swin mimarisinin görüntü sınıflandırma problemlerinde güçlü bir alternatif olarak öne çıktığını göstermektedir.

Ortalama Metrikler (5 Fold):

Accuracy: 91.11%
Precision: 0.9086
Recall: 0.8086
F1 Score: 0.8027

Şekil 25. Swin Transformer Modelinin IQ-OTH/NCCD Datasetindeki Sonuçları

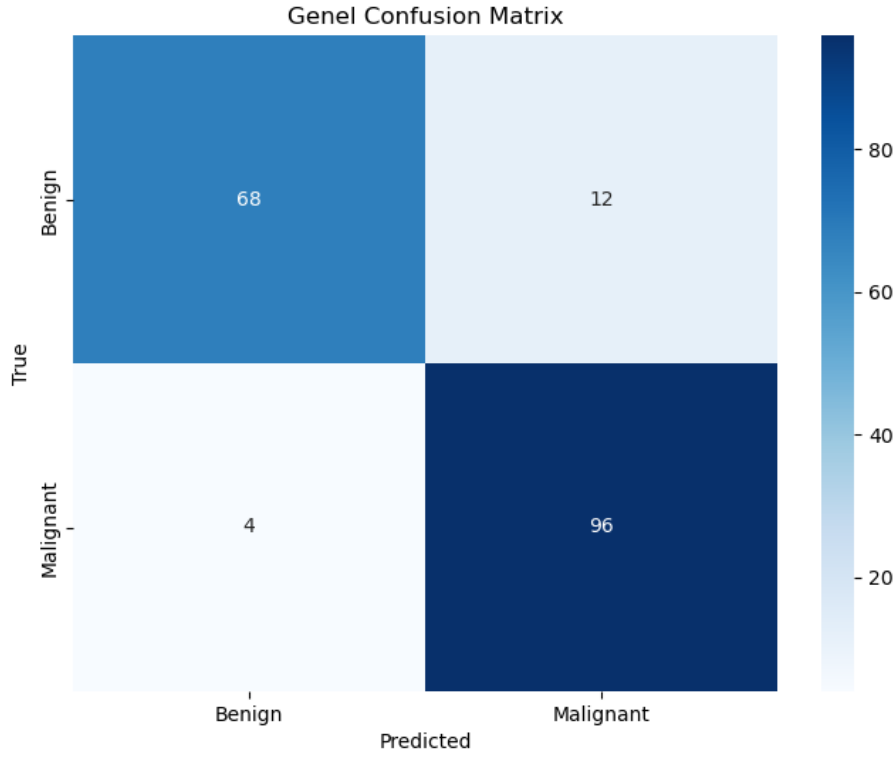


Şekil 26. Swin Transformer Modelinin IQ-OTH/NCCD Dataseti Confusion Matrix Sonuçları

Ortalama Metrikler (5 Fold):

Accuracy: 91.11%
 Precision: 0.9196
 Recall: 0.9050
 F1 Score: 0.9086

Şekil 27. Swin Transformer Modelinin SPIE-AAPM-NCI Datasetindeki Sonuçları

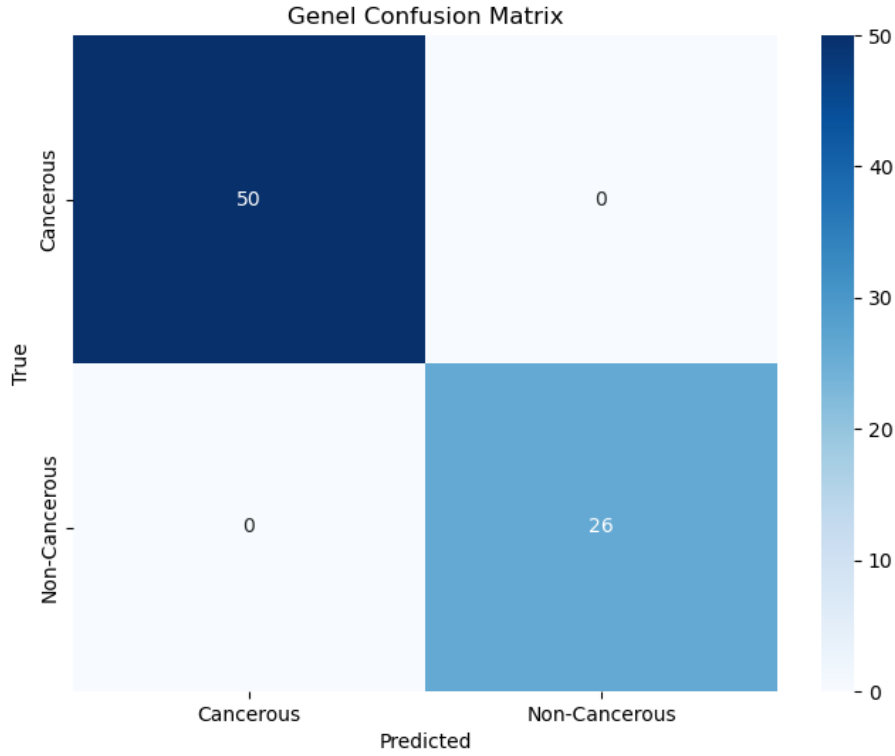


Şekil 28. Swin Transformer Modelinin SPIE-AAPM-NCI Dataseti Confusion Matrix Sonuçları

Ortalama Metrikler (5 Fold):

Accuracy: 100.00%
Precision: 1.0000
Recall: 1.0000
F1 Score: 1.0000

Şekil 29. Swin Transformer Modelinin CT-Scan Images Datasetindeki Sonuçları



Şekil 30. Swin Transformer Modelinin CT-Scan Images Dataseti Confusion Matrix Sonuçları

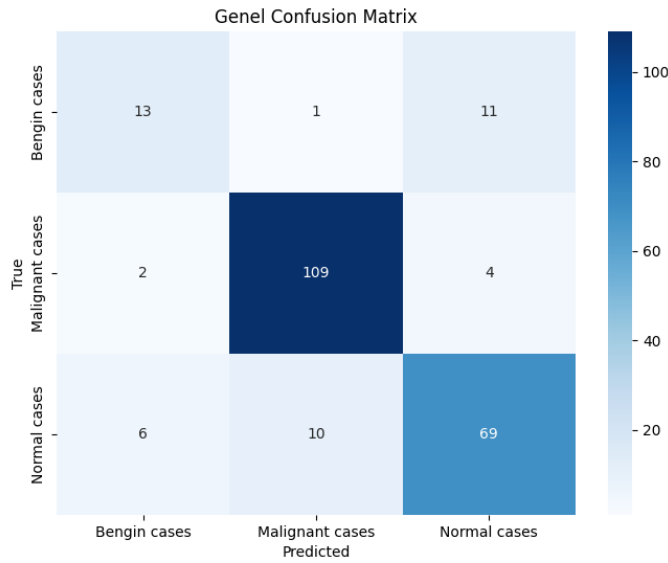
BEiT Base modeli, üç farklı veri kümesi üzerinde 5 katlı K-fold çapraz doğrulama yöntemi ile değerlendirilmiştir. IQ-OTH/NCCD lung cancer dataset veri kümesinde modelin ortalama doğruluğu %84.89 olarak hesaplanırken, precision ve recall değerleri sırasıyla 0.8066 ve 0.7599 olarak elde edilmiştir. Bu veri setinde genel başarı dengeli olmakla birlikte, özellikle recall değerinin precision'a kıyasla düşük kalması, modelin bazı sınıfları yeterince duyarlı şekilde tanımlayamadığını düşündürmektedir. SPIE-AAPM-NCI Lung Nodule Classification Challenge Dataset veri kümesine geçildiğinde model performansında gözle görülür bir artış yaşanmış; doğruluk oranı %88.89, precision değeri 0.9000, recall değeri 0.8875 ve F1 skoru 0.8868 olarak ölçülmüştür. Bu, modelin her iki sınıf için dengeli ve yüksek başarıyla çalıştığını göstermektedir. En yüksek performans ise CT-Scan Images Dataset veri kümesi üzerinde elde edilmiş; %93.33 doğruluk, 0.9250 precision, 0.9200 recall ve 0.9175 F1 skoru ile modelin güçlü bir sınıflandırma kabiliyetine sahip olduğu ortaya konmuştur. Bu sonuçlar, BEiT Base modelinin özellikle daha dengeli

ve iyi etiketlenmiş veri kümelerinde yüksek genelleme kapasitesine ulaştığını, sınıflar arası ayrımı başarılı bir şekilde gerçekleştirebildiğini ve görev odaklı görsel temsillerde etkin bir biçimde kullanılabileceğini göstermektedir. Genel olarak değerlendirildiğinde, modelin performansı kullanılan veri kümesinin kalitesi, sınıf dengesi ve örnek çeşitliliğine önemli ölçüde bağlıdır.

Ortalama Metrikler (5 Fold):

Accuracy: 84.89%
Precision: 0.8066
Recall: 0.7599
F1 Score: 0.7591

Şekil 31. BEiT Base Modelinin IQ-OTH/NCCD Datasetindeki Sonuçları

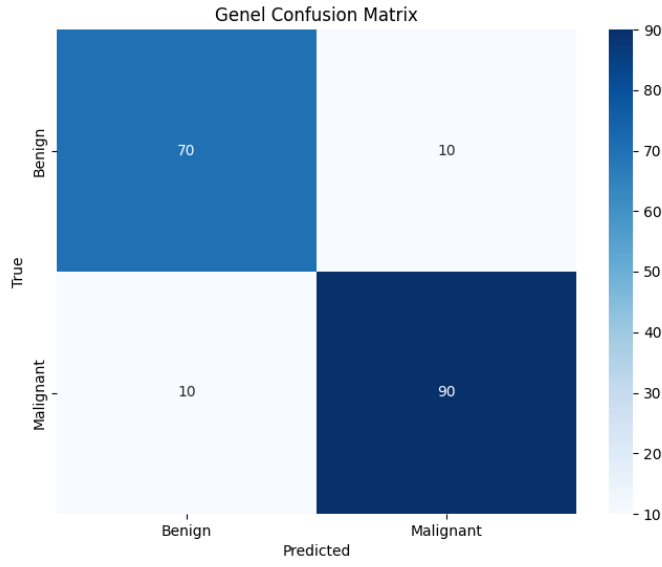


Şekil 32. BEiT Base IQ-OTH/NCCD Datasetindeki Confusion Matrix Sonuçları

Ortalama Metrikler (5 Fold):

Accuracy: 88.89%
Precision: 0.9000
Recall: 0.8875
F1 Score: 0.8868

Şekil 33. BEiT Base Modelinin SPIE-AAPM-NCI Datasetindeki Sonuçları

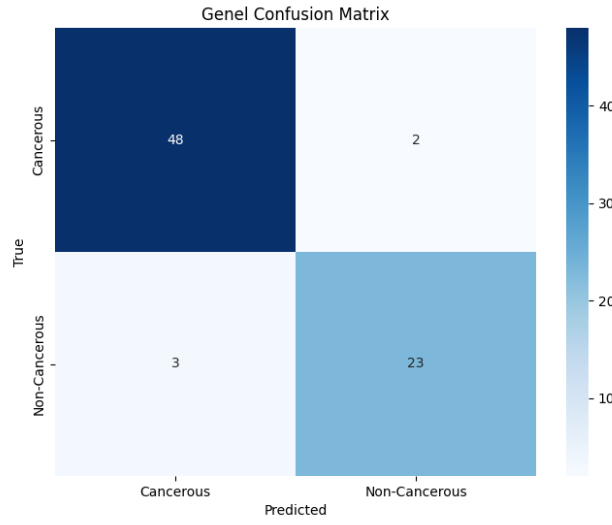


Şekil 34. BEiT Base Modelinin SPIE-AAPM-NCI Datasetindeki Confusion Matrix Sonuçları

Ortalama Metrikler (5 Fold):

Accuracy: 93.33%
Precision: 0.9250
Recall: 0.9200
F1 Score: 0.9175

Şekil 35. BEiT Base Modelinin CT-Scan Images Datasetindeki Sonuçları



Şekil 36. BEiT Base CT-Scan Images Datasetindeki Confusion Matrix Sonuçları

Tablo 1. IQ-OTH/NCCD Datasetinin Modellere Göre Sonuçları

IQ-OTH/NCCD Lung Cancer Dataset						
Model/Folder	Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	Genel
CNN	0.84	0.89	0.87	0.89	0.82	0.86
VGG-16	0.89	0.91	0.89	0.87	0.93	0.89
ResNet50	0.87	0.84	0.87	0.84	0.80	0.84
EfficientNetV2-S	0.80	0.87	0.82	0.84	0.84	0.84
SwinV2-CR-Small-224	0.91	0.93	0.91	0.84	0.96	0.91
BEiT-base-patch16-224	0.89	0.80	0.82	0.89	0.84	0.85

Tablo 2. SPIE-AAPM-NCI Datasetinin Modellere Göre Sonuçları

SPIE-AAPM-NCI Lung Nodule Classification Challenge Dataset						
Model/Folder	Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	Genel
CNN	0.94	0.92	0.92	0.92	0.89	0.92
VGG-16	0.89	0.97	0.89	0.86	0.89	0.90
ResNet50	0.47	0.53	0.56	0.50	0.67	0.55
EfficientNetV2-S	0.52	0.64	0.58	0.50	0.56	0.56
SwinV2-CR-Small-224	0.92	0.86	0.92	0.94	0.92	0.91
BEiT-base-patch16-224	0.94	0.86	0.89	0.92	0.83	0.89

Tablo 3. CT-Scan Images Datasetinin Modellere Göre Sonuçları

CT-Scan Images Dataset						
Model/Folder	Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	Genel
CNN	0.62	0.67	0.67	0.67	0.67	0.66
VGG-16	0.94	1.00	1.00	0.93	0.93	0.96
ResNet50	0.62	0.80	0.67	0.93	0.93	0.79
EfficientNetV2-S	0.56	0.73	0.87	0.33	0.20	0.54
SwinV2-CR-Small-224	1.00	1.00	1.00	1.00	1.00	1.00
BEiT-base-patch16-224	1.00	1.00	0.73	0.93	1.00	0.93

Çalışmanın ilerleyen aşamalarında, mevcut veri setlerindeki sınıf dengesizliklerini gidermek ve modelin genelleme yetisini artırmak amacıyla kapsamlı bir veri artırma (data augmentation) süreci uygulanacaktır. Bu süreçte, her bir sınıfa ait örneklerin sayısı analiz edilerek, hedeflenen örnek sayısına ulaşmak için eksik sınıflarda sentetik veriler üretilenektir. Bu amaç doğrultusunda, döndürme (rotation), kaydırma-ölçekleme-döndürme (shift-scale-rotate), parlaklık/kontrast değişimi (brightness/contrast adjustment), Gauss bulanıklığı (Gaussian blur), Gauss gürültüsü (Gaussian noise) ve elastik bozulma (elastic deformation) gibi çeşitli dönüşümler içeren bir augmentasyon bileşimi kullanılacaktır. Beş katlı çapraz doğrulama (5-fold cross-validation) yaklaşımı çerçevesinde her bir fold için orijinal ve artırılmış veriler ayrı dizinlerde organize edilerek veri yönetimi sistematik bir şekilde sürdürülecektir. Ayrıca, test veri setlerinde herhangi bir veri arttırımı yapılmayacaktır böylece orijinal biçimleri korunacaktır. Nihai olarak, augmentasyon sonrası her sınıfa ait örneklerin eşit ya da eşite yakın bir şekilde dağılımı sağlanarak, sınıf dengesizliğinin neden olabileceği önyargıların önüne geçilmesi hedeflenmektedir. Bu çalışmalar, modelin doğruluk, hassasiyet ve genelleme kabiliyeti açısından daha istikrarlı performans göstermesine katkı sağlayacaktır.

Gelecekte yapılması planlanan çalışmalar kapsamında, mevcut modellerin ötesine geçilerek daha gelişmiş ve modern derin öğrenme mimarilerinin uygulanması hedeflenmektedir. Bu doğrultuda EfficientNetV2-S, Swin Transformer, BEiT, VGG16, CNN ve ResNet gibi bireysel modellerin yanında, bu mimariler arasında hibrit yaklaşımlar da geliştirilecektir. Özellikle farklı mimarilerin güçlü yönlerini birleştiren hibrit modellerin, sınıflandırma doğruluğu ve sınıf bazlı ayırt edicilik açısından daha üstün performans sergilemesi beklenmektedir.

5. SONUÇLAR

Bu çalışma, akciğer kanseri tespiti gibi hayati öneme sahip bir alanda yapay zekâ tabanlı yöntemlerin performansını ortaya koymakta ve farklı derin öğrenme mimarilerinin çoklu veri kümeleri üzerindeki sınıflandırma başarımlarını kapsamlı bir şekilde analiz etmektedir. Kullanılan modeller arasında CNN, VGG16, ResNet, EfficientNetV2-S, Swin Transformer ve BEiT Base gibi çeşitli derin öğrenme yapıları yer almakta olup, bu modellerin farklı veri setleri üzerindeki doğruluk, kesinlik (precision), duyarlılık (recall) ve F1-skoru gibi metriklerle değerlendirilmeleri gerçekleştirilmiştir.

Modellerin performansları incelendiğinde, CNN mimarisinin özellikle ilk iki veri setinde yüksek doğruluk ve dengeli sınıf ayrımı sağladığı görülmüştür. Malign vakalarda başarı oranı oldukça yüksek olsa da, benign sınıfındaki performans düşüklüğü sınıf dengesizliğine işaret etmektedir.

VGG16 modeli, genel olarak tüm veri setlerinde istikrarlı bir başarı göstermiş, özellikle dengeli veri dağılımına sahip durumlarda %96'ya varan doğruluk oranları ile dikkat çekmiştir. Bu sonuçlar, VGG16'nın medikal görüntü sınıflandırma alanında etkili bir temel model olduğunu göstermektedir.

EfficientNetV2-S mimarisi ise ilk veri setinde tatmin edici sonuçlar üretmiş; ancak daha karmaşık ve dengesiz veri setlerinde başarı oranı belirgin şekilde azalmıştır. Bu durum, modelin genelleme kapasitesinin sınırlı olduğunu ve performansının veri setinin yapısına duyarlı olduğunu göstermektedir.

En dikkat çekici sonuçlar ise Swin Transformer ve BEiT Base mimarilerinden elde edilmiştir. Swin Transformer, özellikle CT görüntüleri üzerinde tüm metriklerde %100 başarıya ulaşarak modelin yüksek genelleme gücünü ortaya koymuştur. BEiT Base ise

daha dengeli veri setlerinde üstün sınıflandırma performansı göstermiştir.

Genel olarak değerlendirildiğinde, bu çalışma; derin öğrenme mimarilerinin tıbbi görüntü analizinde etkin bir şekilde kullanılabileceğini ve özellikle dengeli, çeşitlendirilmiş veri setleri ile model başarımının önemli ölçüde artırılabilirliğini ortaya koymaktadır. Model seçimi kadar, veri setinin kalitesi, sınıf dağılımı ve ön işleme teknikleri de sınıflandırma performansı üzerinde belirleyici rol oynamaktadır.

Gelecekteki çalışmalarda, sınıf dengesizliğini gidermeye yönelik stratejiler ve hibrit model yaklaşımlarının kullanılması, tıbbi tanı süreçlerinde derin öğrenmenin güvenilirliğini ve genellenebilirliğini artıracaktır. Bu bağlamda, elde edilen bulguların klinik karar destek sistemlerinin geliştirilmesine yönelik katkılar sunması beklenmektedir.

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] Z. Wang, L. Liu, L. Wang, and L. Zhou, “R2GenGPT: Radiology Report Generation with Frozen LLMs,” *Meta-radiology*, vol. 1, no. 3, pp. 100033–100033, Nov. 2023, doi: <https://doi.org/10.1016/j.metrad.2023.100033>.
- [8] T. Haryanto, Heru Suhartanto, Aniati Murni, Kusmardi Kusmardi, and J. M. Zain, “SCOV-CNN: A Simple CNN Architecture for COVID-19 Identification Based on the CT Images,” *JOIV International Journal on Informatics Visualization*, vol. 8, no. 1, Mar. 2024, doi: <https://doi.org/10.62527/joiv.8.1.1750>.

- [9] 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
- [10] D. Şeker, M. S. Kartal, A. Yıldız, and I. Öksüz, “Akciğer Kanseri tespitinde Dönüşüm ve Evrişim Tabanlı modeller İle Açıklanabilir Yapay Zeka Uygulaması,” *EMO Bilimsel Dergi*, <https://dergipark.org.tr/tr/pub/emobd/issue/86413/1484039> (accessed Mar. 23, 2025).
- [11] Berat ÖZDET and Semra İÇER, “AKCİĞER BİLGİSAYARLI TOMOGRAFİ GÖRÜNTÜLERİNDE GÖRÜNTÜ İŞLEME UYGULAMALARI İLE TUMÖRLERİNİN TESPİT EDİLMESİ,” *Deleted Journal*, pp. 135–150, Mar. 2022, doi: <https://doi.org/10.17482/uumfd.947619>.
- [12] “The IQ-OTH/NCCD lung cancer dataset,” *www.kaggle.com*. <https://www.kaggle.com/datasets/hamdallak/the-iqothnccd-lung-cancer-dataset>
- [13] S. Zhang, “Challenges in KNN Classification,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 10, pp. 1–1, 2021, doi: <https://doi.org/10.1109/tkde.2021.3049250>.
- [14] Yrd Doç, Emre Yakut, Bekir Elmas, S. Yavuz, and Destek Vektör Makineleri, “Yapay Sinir Ağları ve Destek Vektör Makineleri Yöntemleriyle Borsa Endeksi Tahmini,” Jan. 01, 2014. https://www.researchgate.net/publication/349027298_Yapay_Sinir_Aglari_ve_Destek_Vektor_Makineleri_Yontemleriyle_Borsa_Endeksi_Tahmini
- [15] Y. Özüpak, “Evrişimli Sinir Ağı (ESA) Mimarileri ile Hücre Görüntülerinden Sıtmanın Tespit Edilmesi,” *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, vol. 39, no. 1, pp. 197–210, Mar. 2024, doi: <https://doi.org/10.21605/cukurovaumfd.1460434>.