T.C.
SELÇUK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

CİLT KANSERİ TEŞHİSİNDE DERİN ÖĞRENME MODELİNİN GELİŞTİRİLMESİ VE AÇIKLANABİLİR YAPAY ZEKÂ TEKNİKLERİ İLE YORUMLANMASI

Havva Hazel ARAS

YÜKSEK LİSANS TEZİ

Bilgisayar Mühendisliği Anabilim Dalı

Nisan - 2025
KONYA
Her Hakkı Saklıdır

TEZ KABUL VE ONAYI

Havva Hazel ARAS tarafından hazırlanan “Cilt Kanseri Teşhisinde Derin Öğrenme Modelinin Geliştirilmesi ve Açıklanabilir Yapay Zekâ Teknikleri ile Yorumlanması” adlı tez çalışması 14/04/2025 tarihinde aşağıdaki jüri tarafından oy birliği ile Selçuk Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı’nda Yüksek Lisans Tezi olarak kabul edilmiştir.

Jüri Üyeleri	**İmza**
Başkan Doç. Dr. İsmail KOÇ	……..
Danışman Prof. Dr. Nurettin DOĞAN	……..
Üye Doç. Dr. Murat KÖKLÜ	……..

Yukarıdaki sonucu onaylarım.

Prof. Dr. Ömer Faruk YÜKSEL
FBE Müdürü

TEZ BİLDİRİMİ

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

DECLARATION PAGE

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

İmza

Havva Hazel ARAS
22.04.2025

ÖZET

YÜKSEK LİSANS TEZİ

CİLT KANSERİ TEŞHİSİNDE DERİN ÖĞRENME MODELİNİN GELİŞTİRİLMESİ VE AÇIKLANABİLİR YAPAY ZEKÂ TEKNİKLERİ İLE YORUMLANMASI

Havva Hazel ARAS

Selçuk Üniversitesi Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Prof. Dr. Nurettin DOĞAN

2025, 64 Sayfa

Jüri
Prof. Dr. Nurettin DOĞAN
Doç. Dr. İsmail KOÇ
Doç. Dr. Murat KÖKLÜ

Cilt kanseri, dünya genelinde en yaygın görülen kanser türlerinden biri olup erken teşhis edilmediğinde ciddi sağlık sorunlarına ve yüksek mortalite oranlarına yol açmaktadır. Geleneksel teşhis yöntemleri uzmanlık gerektiren zaman alıcı süreçlerdir ve bazı durumlarda doğruluk oranı düşebilmektedir. Bu çalışmada, cilt kanseri teşhisi için derin öğrenme modelleri eğitilmiş ve modellerin karar mekanizmalarının anlaşılabilir hale getirilmesi amacıyla Açıklanabilir Yapay Zekâ (XAI) yöntemleri kullanılmıştır. Araştırma kapsamında HAM10000 veri seti kullanılarak, deri lezyonlarının sınıflandırılması hedeflenmiştir. Model eğitimi sürecinde MobileNet, DenseNet ve Xception gibi derin öğrenme mimarileri kullanılmıştır. MobileNet için doğruluk oranı %84.04, DenseNet için %83.57 ve Xception için %81.52 olarak elde edilmiştir. Modelin tahminlerinin yorumlanabilirliğini artırmak amacıyla Grad-CAM ve LIME yöntemleri kullanılmıştır. Grad-CAM yöntemiyle, modelin sınıflandırma sırasında odaklandığı bölgeler görsel olarak vurgulanırken, LIME yöntemi ile lokal bazlı açıklamalar elde edilmiştir. Bu çalışma, yapay zekâ destekli sistemlerin cilt kanseri teşhisindeki doğruluğunu ve güvenilirliğini artırarak, sağlık profesyonellerinin teşhis süreçlerine önemli bir katkı sağladığını göstermektedir. Yapay zekâ teknolojilerinin tıbbi görüntüleme alanında etkinliği kanıtlanmış olup, erken teşhis oranlarını yükseltmek ve hastaların yaşam kalitesini iyileştirmek adına büyük bir potansiyel taşımaktadır. Bu bulgular, yapay zekâ uygulamalarının klinik karar destek sistemlerinde yaygın kullanımını teşvik ederek, sağlık hizmetlerinin etkinliğini artırmaya yönelik önemli bir adım niteliğindedir.

Anahtar Kelimeler: Açıklanabilir Yapay Zekâ, Cilt Kanseri, Derin Öğrenme, Grad-CAM, LIME

ABSTRACT

MS THESIS

DEVELOPMENT OF A DEEP LEARNING MODEL FOR SKIN CANCER DIAGNOSIS AND ITS INTERPRETATION USING EXPLAINABLE ARTIFICIAL INTELLIGENCE TECHNIQUES

Havva Hazel ARAS

The Graduate School Of Natural And Applied Science Of Selçuk University
The Degree Of Master Of Science

Advisor: Prof. Dr. Nurettin DOĞAN

2025, 64 Pages

Jury
Prof. Dr. Nurettin DOĞAN
Assoc. Prof. Dr. İsmail KOÇ
Assoc. Prof. Dr. Murat KÖKLÜ

Skin cancer is one of the most prevalent types of cancer worldwide, leading to severe health issues and high mortality rates if not diagnosed early. Traditional diagnostic methods are time-consuming processes requiring expertise and may exhibit reduced accuracy in some cases. In this study, deep learning models were trained for skin cancer diagnosis, and Explainable Artificial Intelligence (XAI) methods were employed to enhance the interpretability of the models' decision mechanisms. The HAM10000 dataset was utilized to classify skin lesions within the scope of the research. During the model training process, deep learning architectures such as MobileNet, DenseNet, and Xception were used. The accuracy rates were obtained as 84.04% for MobileNet, 83.57% for DenseNet, and 81.52% for Xception.To improve the interpretability of model predictions, Grad-CAM and LIME methods were employed. Grad-CAM highlighted the regions on which the model focused during classification through visual explanations, while LIME provided localized, instance-based explanations. This study demonstrates that AI-powered systems significantly enhance the accuracy and reliability of skin cancer diagnosis, providing substantial support to healthcare professionals during diagnostic processes. Artificial intelligence technologies have proven their effectiveness in the field of medical imaging, showcasing great potential for increasing early diagnosis rates and improving patients' quality of life. These findings represent a significant step toward promoting the widespread adoption of AI applications in clinical decision support systems, thereby enhancing the efficiency of healthcare services.

Keywords: Explainable Artificial Intelligence, Deep Learning, Grad-CAM, LIME, Skin Cancer

ÖNSÖZ

Bu tez çalışması, akademik yaşamımın önemli bir aşamasını temsil etmektedir. Bu süreçte yanımda olan, bana güç veren ve katkı sunan herkese içten teşekkürlerimi sunmak isterim.

Her şeyden önce, tüm eğitim hayatım boyunca maddi ve manevi desteğini esirgemeyen aileme en derin şükranlarımı sunarım. Özellikle her daim yanımda olan, desteği ve sevgisiyle bana güç veren sevgili kardeşim Lara'ya gönülden teşekkür ederim.

Tez sürecinde bilgi ve deneyimiyle bana yol gösteren, değerli katkılarını esirgemeyen danışmanım Sayın Prof. Dr. Nurettin Doğan'a teşekkür ederim.

Zor zamanlarda yanımda olan, desteğini ve emeğini eksik etmeyen değerli araştırma görevlisi arkadaşım Yusuf Eryeşil'e de ayrıca teşekkür ederim.

Bu süreçte doğrudan ya da dolaylı olarak katkı sunan herkese teşekkürü bir borç bilirim.

Havva Hazel ARAS
KONYA-2025

İÇİNDEKİLER

ÖZET ..... **iv**

ABSTRACT ..... **v**

ÖNSÖZ ..... **vi**

İÇİNDEKİLER ..... **vii**

TABLOLAR LİSTESİ ..... **ix**

ŞEKİLLER LİSTESİ ..... **ix**

ALGORİTMALAR LİSTESİ ..... **x**

SİMGELER VE KISALTMALAR ..... **xi**

1. GİRİŞ **1**

2. KAYNAK ARAŞTIRMASI **5**

3. MATERYAL VE YÖNTEM **15**

3.1. Veri Seti 15
3.2. Veri Önişleme 17
3.2.1. Veri Dengesizliği ve Veri Dengeleme 18
3.2.2. Undersampling 18
3.3. MobileNet 19
3.4. DenseNet 21
3.5. Xception 23
3.6. Açıklanabilir Yapay Zekâ 24
3.6.1. Grad-CAM 25
3.6.2. LIME 27
3.7. Değerlendirme Metrikleri 30
3.7.1. Karmaşıklık Matrisi 30
3.7.2. Performans Ölçütleri 32

4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA 35

4.1. Veri Önişleme 36
4.1.Model Eğitimi 38
4.2. Deney Sonuçları ve Performans Karşılaştırması 40
4.3. Grad-CAM Deneysel Sonuçlar 47
4.3.1. MobileNet Modeli 47
4.3.2. DenseNet Modeli 48
4.3.3. Xception Modeli 49
4.3.4. Modeller Arası Karşılaştırma 50
4.4. LIME Deneysel Sonuçlar 50
4.4.1. MobileNet Modeli 52
4.4.2. DenseNet Modeli 52
4.4.3. Xception Modeli 53
4.4.4. LIME'ın Sağladığı Lokal Açıklamalar 54
4.5. Grad-CAM ve LIME Karşılaştırması 54

5. SONUÇLAR VE ÖNERİLER 56

5.1 Sonuçlar 56
5.2 Öneriler 57

KAYNAKLAR 58

ÖZGEÇMİŞ 64

TABLOLAR LİSTESİ

Tablo 1: Malign melanom tipik belirtileri 2
Tablo 2: Cilt kanseri teşhisi konusunda yapılan çalışmalar ve özellikleri 14
Tablo 3: Kullanılan modellerin temel özellikleri 36
Tablo 4: Modellerinin doğrulama sonuçları ve toplam eğitim süresi karşılaştırması 38
Tablo 5: Modellerin değerlendirme metrikleri 40
Tablo 6: Karmaşıklık matrislerine göre değerlendirme 45
Tablo 7 : Çalışmamız ile literatürdeki çalışmaların karşılaştırılması 46
Tablo 8: tüm modellerin LIME maskeleri ve tahmin edilen sınıf 51
Tablo 9: Modellere göre LIME yorumları 54
Tablo 10: Grad-CAM ve LIME karşılaştırılması 55

ŞEKİLLER LİSTESİ

Şekil 1: Yöntemin genel akış şeması 15
Şekil 2:HAM10000 Veri Seti Sınıf Dağılımı 16
Şekil 3: HAM10000 veri setindeki cilt lezyonu Sınıfları 16
Şekil 4: HAM 10000 veri setindeki sınıf dağılımı 17
Şekil 5: İki sınıf için undersampling uygulama örneği 19
Şekil 6: MobileNetV2 modelinin genel mimarisi (Zuo et al., 2021) 20
Şekil 7:DenseNet121 model mimarisinin temel yapısı (Arora, August 2, 2020) 21
Şekil 8: DenseNet-121 mimarisinin katman yapısı (Tiwari et al., 2023) 22
Şekil 9: Xception'te Derinlik Ayrıştırmalı Konvolüsyonların Prensibi (Tsang, 2018) 23
Şekil 10: Xception'in genel mimarisi (Mohammed et al., 2023) 24
Şekil 11: Grad-CAM kullanılarak elde edilen aktivasyon haritası (Adadi & Berrada, 2018) 27
Şekil 12: LIME kullanılarak model tahmininin yorumlanması (Ribeiro, 2016). 29
Şekil 13: Karmaşıklık matrisi 31

Şekil 14: ROC eğrisi ve eşik belirleyici. Eğri altındaki alan (AUC), sınıflayıcının başarımını gösterir; 1'e yakın AUC değeri, en iyi performansı ifade eder (Liu et al., 2020). 34
Şekil 15: Veri arttırma sonrası örnek görüntüler 37
Şekil 16: Veri arttırma sonrası sınıf dağılımı 37
Şekil 17: Mobilenet modeline ait eğitim ve doğrulama kayıplarının grafiği 39
Şekil 18: DenseNet modeline ait eğitim ve doğrulama kayıplarının grafiği 40
Şekil 19: Xception modeline ait eğitim ve doğrulama kayıplarının grafiği 40
Şekil 20: MobileNet modeli için karmaşıklık matrisi 42
Şekil 21: DenseNet modeli için karmaşıklık matrisi 43
Şekil 22: Xception modeli için karmaşıklık matrisi 44
Şekil 23: Üç model için elde edilen Grad-CAM çıktıları ve model tahminleri .. 47
Şekil 24: MobileNet modeline ait Grad-CAM çıktıları ve model tahminleri 48
Şekil 25: DenseNet modeline ait Grad-CAM çıktıları ve model tahminleri 49
Şekil 26: Xception modeline ait Grad-CAM çıktıları ve model tahminleri 49
Şekil 27: MobileNet modeline ait LIME çıktıları ve model tahminleri 52
Şekil 28: DenseNet modeline ait LIME çıktıları ve model tahminleri 53
Şekil 29: Xception modeline ait LIME çıktıları ve model tahminleri 53

ALGORİTMALAR LİSTESİ

Algoritma 1: Grad-CAM algoritmasının temel adımları 26
Algoritma 2: LIME algoritmasının temel adımları 29
Algoritma 3: Yöntemin genel akışını gösteren Pseudo kod 35

SİMGELER VE KISALTMALAR

Kısaltmalar

CNN	: Convolutional Neural Networks (Evrişimli Sinir Ağları)
YSA	: Yapay Sinir Ağları
Grad-CAM	: Gradient-Weighted Class Activation Mapping (Gradyan Ağırlıklı Sınıf Aktivasyon Haritalaması)
LIME	: Local Interpretable Model-agnostic Explanations (Yerel Yorumlanabilir Modelden Bağımsız Açıklamalar)
TP	: True Positive (Doğru Pozitif)
FP	: False Positive (Yanlış Pozitif)
TN	: True Negative (Doğru Negatif)
FN	: False Negative (Yanlış Negatif)
RGB	: Red-Green-Blue (Kırmızı, Yeşil, Mavi)
XAI	: Explainable Artificial Intelligence (Açıklanabilir Yapay Zekâ)
ISIC	: International Skin Imaging Collaboration (Uluslararası Cilt Görüntüleme İş birliği)
FCN	: Fully Convolutional Network (Tam Konvolüsyonel Ağ)
GPU	: Graphics Processing Unit (Grafik İşlem Birimi)
AUC	: Area Under the Curve (Eğri Altındaki Alan)
Mel	: Melanom
Nv	: Benign Nevüs
Bcc	: Bazal Hücreli Karsinom
Ak	: Aktinik Keratoz
Bkl	: Benign Keratoz Lezyonları
Df	: Dermatofibroma
Vasc	: Vasküler Lezyonlar

1. GİRİŞ

Kanser, günümüzde küresel ölçekte önemli bir sağlık sorunu olarak kabul edilmektedir. Cilt kanseri, kanser türleri arasında en sık görülenlerden biri olarak öne çıkmaktadır. Cilt kanseri, deri hücrelerinin kontrolsüz büyümesi ile karakterize edilen ve dünya genelinde önemli bir sağlık sorunu olan bir hastalıktır (Uysal et al., 2004). Erken evrede tespit edildiğinde tedavi şansı yüksek olmasına rağmen, ileri aşamalarda teşhis edildiğinde tedavi seçenekleri sınırlı olabilir ve mortalite riski artabilir. Her yıl milyonlarca insan cilt kanserine yakalanmakta ve bu nedenle ciddi sağlık sorunlarıyla karşılaşmaktadır. Erken teşhis ve doğru sınıflandırma, cilt kanserinin tedavi başarısını ve hastaların yaşam kalitesini büyük ölçüde etkileyen önemli bir faktördür. Cilt kanserinin erken teşhisi, hastalığın tedavi edilebilirliğini artırabilir ve tedavi sonuçlarını iyileştirebilir (Akyel & Arıcı, 2022). Dermatologlar, cilt kanserini teşhis etmek için genellikle dermoskopi taraması yöntemini kullanır. Dermoskopi, cilt yüzeyindeki lezyonları daha ayrıntılı olarak incelemek için dermoskop adı verilen özel bir cihazın kullanıldığı bir görüntüleme tekniğidir (Alp et al., 2021). Bununla birlikte, dermoskopi görüntülerinin yorumlanması, uzmanlık gerektiren bir beceri gerektirebilir ve doğruluk oranını etkileyebilir.

Son yıllarda, derin öğrenme teknikleri tıp alanında önemli bir ilerleme kaydetmiş ve cilt kanseri teşhisi için bilgisayar destekli tanı (BDT) sistemleri geliştirilmeye başlanmıştır (Uslu et al., 2006). Derin öğrenme, karmaşık yapay sinir ağları kullanarak büyük veri kümelerini analiz etmeyi ve karmaşık desenleri algılamayı sağlayan bir yaklaşımdır. Cilt kanseri teşhisinde derin öğrenme algoritmaları, karmaşık yapıları ve büyük miktarda veriyi işleme yetenekleri sayesinde önemli bir araç haline gelmiştir. Bu teknik, cilt kanserinin erken teşhisi için yeni bir yaklaşım sunabilir ve dermatologların doğruluk oranını artırabilir (Öncel & Gündoğdu, 2017). Ancak, mevcut derin öğrenme modelleri ve yöntemleri hala bazı sınırlamalarla karşı karşıyadır ve cilt kanseri teşhisinde daha yüksek doğruluk oranlarına ulaşmak için yeni yaklaşımların geliştirilmesi gerekmektedir (Li et al., 2021). Bilgisayar mühendisliği alanında özellikle görüntü işleme, yapay zekâ ve makine öğrenmesi teknolojileri kullanılarak cilt kanseri teşhisinde önemli katkılar sağlanmaktadır. Geliştirilen algoritmalar, deri lezyonlarının dijital görüntülerinden kanserli hücreleri tespit edip sınıflandırarak, dermatologların teşhis süreçlerine destek olur ve bu süreçleri objektif verilerle güçlendirir. Bu bağlamda,

Bilgisayar Mühendisliği, cilt kanseri teşhisinde erken müdahalenin yanı sıra tedavi stratejilerinin geliştirilmesine yönelik çalışmalarda kritik bir rol üstlenmekte ve medikal görüntüleme ile analiz süreçlerini teknoloji ile entegre ederek zenginleştirmektedir. Cilt kanserinin belirtileri, kanser türüne ve gelişim aşamasına bağlı olarak değişkenlik gösterebilir. Ancak genel olarak, cilt kanseri teşhisinde dikkat edilmesi gereken bazı yaygın semptomlar bulunmaktadır. Malign melanom, cilt kanserlerinin en tehlikeli türü olup, erken teşhis edilmediğinde hızla ilerleyebilir ve metastaz yapabilir. Bu nedenle, malign melanomun erken belirtilerinin tanınması ve izlenmesi hayati öneme sahiptir (Rigel et al., 2010). Tablo 1'de Malign melanomun tipik belirtileri gösterilmiştir.

Tablo 1: Malign melanom tipik belirtileri

Kriter	Açıklama
Asimetri (A)	Benin veya lezyonun bir yarısının diğer yarısına benzememesi. Düzensiz şekillere sahip lezyonlar malign melanomun belirtisi olabilir.
Sınır (B)	Melanom lezyonlarının kenarları genellikle düzensiz, girintili çıkıntılı veya bulanıktır. Sağlıklı benler ise genellikle düzgün sınırlara sahiptir.
Renk (C)	Melanom lezyonları genellikle farklı renk tonları içerir; kahverengi, siyah, kırmızı, beyaz veya mavi renkte olabilir. Renk dağılımı düzensizdir.
Çap (D)	Melanomlar genellikle 6 mm'den büyük olur, ancak daha küçük çapta da olabilirler. Yeni veya mevcut bir benin büyüklüğünde değişiklikler gözlemlenmelidir.
Evrim (E)	Mevcut bir benin boyut, şekil veya renk açısından zamanla değişiklik göstermesi, malign melanomun önemli bir işaretidir. Ayrıca, kanama, kaşıntı veya kabuklanma gibi yeni belirtiler ortaya çıkabilir.

Diğer belirtiler arasında ciltte yeni oluşan pigmentli lezyonlar veya mevcut benlerde belirgin değişiklikler, lezyon bölgelerinde kaşıntı, hassasiyet, ağrı, lezyonların kanaması veya ülserleşmesi yer almaktadır. Bu belirtiler, malign melanomun erken tespiti için kritik olup, dermatolojik muayeneler sırasında dikkatle izlenmelidir (Crowson et al., 2014). Malign melanomun erken evrede tespit edilememesi durumunda, tedavi seçenekleri kısıtlı hale gelmekte ve hastaların yaşam kalitesi ciddi şekilde düşmektedir. Bu bağlamda, cilt kanserinin, özellikle malign melanomun erken teşhis edilebilmesi, hastalığın yayılmasını engellemek ve hastaların hayatta kalma oranlarını artırmak için önemli bir adımdır. Mevcut tanı yöntemleri, çoğunlukla klinik değerlendirmeler ve biyopsiler üzerine kuruludur. Ancak, bu yöntemler zaman alıcı olabilir ve bazı durumlarda kesin sonuçlar vermeyebilir. Bu nedenle, daha hızlı, hassas ve güvenilir teşhis

yöntemlerine ihtiyaç duyulmaktadır. Bilgisayar mühendisliği açısından, yapay zekâ ve makine öğrenmesi algoritmalarının geliştirilmesi, cilt lezyonlarının dijital görüntülerinden melanomun erken ve doğru teşhis edilmesini sağlamaktadır. Gelişmiş görüntü işleme teknikleri, bu belirtilerin otomatik olarak tespit edilmesine ve sınıflandırılmasına olanak tanır, bu da klinik tanı süreçlerini hızlandırır ve doğruluğunu artırır. Yapay zekâ destekli sistemler, cilt lezyonlarının erken evrede doğru bir şekilde tespit edilmesini sağlayarak dermatologlara yardımcı olabilir. Bu sistemler, geniş veri setleri üzerinde eğitilerek, melanom belirtilerini otomatik olarak tanıyabilir ve sınıflandırabilir. Bu sayede, manuel muayenelerde gözden kaçabilecek küçük ancak önemli detaylar yakalanabilir. Dolayısıyla, malign melanom teşhisinde teknolojik yaklaşımlar, cilt kanserinin erken evrede yakalanarak etkin bir şekilde tedavi edilmesine önemli katkılar sunar (Davis et al., 2019). Bu nedenle, yapay zekâ tabanlı çözümlerin geliştirilmesi ve uygulanması, malign melanomun yönetiminde büyük bir ihtiyaç ve fırsat olarak karşımıza çıkmaktadır. Bu bağlamda, teknolojik çözümlerin entegrasyonu, hem hastaların erken dönemde tedaviye başlamalarını sağlayarak yaşam sürelerini uzatmakta, hem de sağlık hizmetlerinin etkinliğini artırmaktadır.

Son yıllarda, yapay zekâ ve derin öğrenme alanındaki gelişmeler, tıp alanında devrim niteliğinde yenilikler sunmaktadır. Derin öğrenme modelleri, büyük veri setleri üzerinde eğitilerek, karmaşık desenleri ve anormallikleri insan gözünden çok daha hızlı ve doğru bir şekilde tanımlayabilmektedir. Bu teknolojiler tıbbi görüntüleme alanında büyük potansiyele sahiptir. Özellikle cilt kanseri teşhisinde bu potansiyel oldukça yüksektir (Nahata & Singh, 2020). Ancak yapay zekâ modellerinin kararları nasıl verdiğini anlamak, bu teknolojilerin kabul edilebilirliği ve güvenilirliği açısından son derece önemlidir. Açıklanabilir yapay zekâ (XAI) yöntemleri, bu modellerin iç işleyişini anlamamızı sağlayarak sağlık profesyonellerinin ve hastaların bu teknolojilere güvenini artırmaktadır (Roy et al., 2023). Bu durum, özellikle hassas ve hayati kararların alındığı sağlık alanında kritik bir rol oynamaktadır. Bu bağlamda, cilt kanseri teşhisinde kullanılacak derin öğrenme modellerinin geliştirilmesi ve bu modellerin açıklanabilir hale getirilmesi, tıp dünyasına önemli katkılar sunacaktır (Hauser et al., 2022). Bu çalışma, hem yapay zekânın sağlık alanında kullanımının yaygınlaşmasını destekleyecek hem de hastaların doğru ve zamanında tedaviye erişimini kolaylaştıracaktır. Böylece, sağlık sektöründeki profesyonellerin daha bilinçli ve güvenilir kararlar alması sağlanarak hem

hastaların yaşam kalitesi artırılacak hem de sağlık hizmetlerinin verimliliği önemli ölçüde iyileştirilecektir (Rahman, 2023).

2. KAYNAK ARAŞTIRMASI

Son yıllarda cilt kanseri teşhisinde özellikle malign melanom üzerine yapılan araştırmalar önemli ölçüde artış göstermiştir. Bu çalışmalar, erken teşhis ve tedavi yöntemlerinin geliştirilmesine odaklanmakta olup, biyomedikal ve bilgisayar mühendisliği disiplinlerinin kesişiminde yoğunlaşmaktadır (Jutzi et al., 2020).

Abdullah Al Kafee'nin 2013 yılında gerçekleştirdiği bir çalışma, görüntü işleme tekniklerinin kullanılarak deri kanserinin tespit edilmesini amaçlamaktadır. İstanbul Üniversitesi Cerrahpaşa Tıp Fakültesi Dermatoloji Bölümü Etik Kurul onayıyla gerçekleştirilen görüntü kayıtlarını temel almaktadır. Çalışmada, ön filtreleme, asimetrik puan algoritması, RGB renkleri yüzey puanları, ışık yansımaları ve görüntü işleme algoritmaları gibi yöntemler kullanılmaktadır. Ayrıca, ABCD puanlaması gibi bilgisayar tabanlı algoritmalarla standart risk analizi yapılmaktadır. Önerilen yöntemin doğruluğu ve faydası, belirli parametrelerle kanıtlanmıştır. Bilgisayar tabanlı bir teşhis olduğundan, yöntemlerin uygunluğu doğrulanmıştır (Al Kafee, 2013).

Muhammet Şahin'in 2016 yılında gerçekleştirdiği bir çalışma, görüntü işleme tekniklerinin kullanılarak malign melanom tipi deri kanserinin incelenmesini amaçlamaktadır. Çalışmada, sınır şekil ve renk öznitelikleri bölütlenmiş nevus (ben) görüntülerinden çıkarılmaktadır. Bu öznitelikler, nevusun iyi huylu mu yoksa maling melanom mu olduğuna karar vermek için kullanılmaktadır. Çalışmada, ön işleme, sabit eşikleme tabanlı bölütleme, öznitelik çıkarma ve Dull-Razor algoritması gibi görüntü işleme yöntemleri kullanılmaktadır. Bu yöntemler, dermaskopik (deri altı dokularının daha net görülmesini sağlayan bir mikroskobik yöntem) imgelerini işlemek için geliştirilmiş algoritmaları içermektedir (Şahin, 2016).

2018 yılında Abdelhafid Ali. I. Mohamed tarafından gerçekleştirilen bir araştırma, melanoma cilt kanseri segmentasyonunu bulanık kümeleme ortalaması ve görüntü bölgesi büyümesi kullanarak incelemektedir. Bu çalışma, cilt görüntülerinde melanoma kanserinin segmentasyonunu geliştirmek amacıyla yeni bir algoritma ve görüntü büyütme tekniğini kullanmaktadır. Ayrıca, bu yöntemin teşhis doğruluğunu ve tedaviye olan etkisini değerlendirmeyi hedeflemektedir. Veriler, Portekiz'deki Pedro Hastanesi veri tabanından elde edilmiştir. Çalışmada, bölge kümeleme yöntemi ve görüntü bölgesi büyütme yöntemi kullanılmaktadır. Özellik seçimi ve çıkarımı, görüntü bölütleme ve öznitelik sınıflandırması gibi adımlar da çalışmanın bir parçasıdır. Elde edilen sonuçlar, bu yöntemin doğruluk, hassasiyet ve duyarlılık açısından iyi performans gösterdiğini

göstermektedir. Doğruluk, hassasiyet ve duyarlılık değerleri sırasıyla 0.9685, 0.9542 ve 0.9829 arasında değişmektedir. Bu çalışma, melanoma cilt kanseri segmentasyonunda daha net ve doğru sonuçlar elde etmek için umut veren bir yöntem olarak değerlendirilmektedir (Ali & Mohamed, 2018).

Sarmad Mohsen Ali Hurmuzi'nin 2019 yılında gerçekleştirdiği bir çalışma, cilt kanseri riskinin değerlendirilmesinde evrişimsel sinir ağlarının kullanıldığı bir modelin geliştirilmesini amaçlamaktadır. Bu çalışma, ISIC (The International Skin Imaging Collaboration) veri seti üzerinde yapılmıştır. Çalışmada, evrişimsel sinir ağları kullanılarak bir model oluşturulmuştur. Bu sinir ağları, cilt kanseri riskini değerlendirmek için kullanılan bir modelin oluşturulmasında kullanılmıştır. Yapılan değerlendirmeler sonucunda, önerilen modelin başarısı, diğer çalışmalarla karşılaştırılabilir seviyede olduğu belirlenmiştir. Bu çalışma, evrişimsel sinir ağlarının kullanılarak etkili bir cilt kanseri risk değerlendirme modelinin oluşturulabileceğini göstermektedir (Hurmuzi, 2019).

2020 yılında Nuh Alpaslan tarafından yapılan bir çalışma, cilt lezyonlarının bölütlenmesinde SegNet mimarisi tabanlı bir sistem kullanarak performansı artırma amacını taşımaktadır. Aynı zamanda, bu çalışmada cilt lezyonları üzerinde veri büyütme, renk tutarlılığı ve kıl silme gibi ön işleme adımlarının bölütlenme performansı üzerindeki etkileri de araştırılmıştır. ISBI2016 veri kümesi kullanılarak geliştirilen SegNet mimarisi tabanlı sistem, cilt lezyonlarının bölütlenmesinde kullanılmıştır. Araştırma sonucunda, veri büyütme yönteminin, modellerin ezberlemesini önleyerek genelleme yeteneğini artırdığı tespit edilmiştir. Bu çalışma, SegNet mimarisi ve ön işleme adımlarının kullanılmasıyla cilt lezyonlarının bölütlenmesinde performansın iyileştirilebileceğini göstermektedir (Şahin & Alpaslan, 2020b).

Zhang ve arkadaşları 2020 yılında, deri kanserinin erken teşhisi için optimize edilmiş bir konvolüsyonel sinir ağı (CNN) yöntemi önerilmektedir. Deri kanseri teşhisinde, konvolüsyonel sinir ağları, görüntü işleme ve makine görmesi teknikleri kullanılarak yüksek doğruluk oranları elde edilmiştir. Önerilen yöntemde, CNN'nin verimliliğini artırmak için geliştirilmiş balina optimizasyon algoritması (WOA) kullanılmıştır. Bu algoritma, balina avlanma davranışını matematiksel olarak modelleyerek CNN'nin ağırlık ve önyargılarını optimize etmektedir. Çalışmada, yöntem iki farklı veri seti olan Dermquest ve DermIS üzerinde test edilmiştir. Simülasyon sonuçları, önerilen yöntemin mevcut diğer yöntemlerle karşılaştırıldığında üstün

olduğunu göstermektedir. Özellikle, doğruluk, duyarlılık ve özgüllük gibi performans metriklerinde yüksek değerler elde edilmiştir. Önerilen optimize edilmiş CNN'nin doğruluğu %97, duyarlılığı %95 ve özgüllüğü %98 olarak ifade edilmiştir (Zhang et al., 2020).

Nurullah Şahin'in 2020 yılında gerçekleştirdiği çalışma, cilt lezyonlarının doğru şekilde bölütlenmesi için metasezgisel temelli Otsu eşikleme yönteminin kullanılmasını amaçlamaktadır. Lezyonların doğru bir şekilde bölütlenmesi, özellik çıkarımı ve sınıflandırma performansı için son derece önemlidir. Bu çalışmada, lezyon bölütlenmesi için çeşitli ön işleme adımları, yani yeniden boyutlandırma, siyah çerçeve silme ve kıl silme gibi yöntemler uygulanmıştır. Çalışma, "Comparison of segmentation methods for automatic diagnosis of dermoscopy images" veri setine dayanmaktadır ve Otsu tabanlı HA (Harmony Search), SA (Simulated Annealing) ve PSO (Particle Swarm Optimization) yöntemlerini içermektedir. Bölütlemenin yanı sıra özellik çıkarımı ve sınıflandırma adımları da çalışmanın bir parçasıdır. Sonuçlar, çalışmanın hesaplanan ortalama iterasyon sayılarına göre çalışma süresini dikkate alındığında, PSO, HA ve SA yöntemlerinin geleneksel Otsu yöntemine göre sırasıyla %40, %29 ve %28.5 daha hızlı optimal eşik değerini bulduğunu göstermektedir. Bu bulgular, metasezgisel temelli Otsu eşikleme yönteminin cilt lezyonlarının bölütlenmesinde daha hızlı ve etkili bir yaklaşım sunulabileceğini göstermektedir (Şahin & Alpaslan, 2020a).

Kazım Kılıç'ın 2021 yılında yaptığı araştırma, cilt kanserinin tespiti için evrişimsel sinir ağı mimarilerinin karşılaştırılmasını amaçlamaktadır. Cilt kanseri erken teşhis edildiğinde tedavi şansı artmaktadır. Bu nedenle, çalışma, gereksiz biyopsi işlemlerini azaltmak ve cilt kanserinin erken teşhisini hızlandırmak amacıyla bir bilgisayar destekli tanı sistemi önermektedir. Araştırmada, ISIC 2018 yarışmasında kullanılan HAM10000 veri seti üzerinde çalışılmıştır. Lezyon bölgesinin tespiti ve ayırıcı özelliklerin elde edilmesi için Gauss bulanıklığı ve Otsu eşikleme yöntemleri kullanılmıştır. DenseNet-121 ve ResNet-152 gibi farklı evrişimsel sinir ağı mimarileri kullanılarak yapılan deneylerde elde edilen sonuçlar incelenmiştir. DenseNet-121 mimarisiyle %92 doğruluk, %91 AUC, %88 hassasiyet, %74 geri çağırma ve %80 F-Skoru elde edilmiştir. ResNet-152 mimarisiyle ise %89 doğruluk, %91 AUC, %85 hassasiyet, %81 geri çağırma ve %83 F-Skoru elde edilmiştir. Ayrıca, ResNet-152 mimarisiyle %84 doğruluk, %78 hassasiyet, %77 geri çağırma ve %78 F-Skoru elde edilmiştir. Bu sonuçlar, evrişimsel sinir ağı mimarilerinin cilt kanseri tespitinde etkili

olabileceğini göstermektedir. Araştırma, cilt kanseri teşhisinde farklı mimarilerin kullanılmasının önemli bir adım olduğunu vurgulamaktadır (Kılıç, 2021).

Gizem Poyraz'ın 2021 yılında yaptığı araştırma, kanser tespiti için görüntü işleme algoritmalarının performansını incelemeyi amaçlamaktadır. Araştırmanın ilk bölümünde, görüntü işleme, renk uzayı ve C#'da kullanılabilen fonksiyonlar ve çerçeveler üzerine bir değerlendirme yapılmıştır. Çalışmada, MatjazZwitter ve Milan Soklic tarafından geliştirilen meme kanseri veri seti kullanılmıştır. Görüntüleri gri seviyeye dönüştüren bir algoritma, Medyan filtreleme ve eşikleme gibi görüntü işleme yöntemleriyle birlikte Sobel Kenar Algılama algoritması da kullanılmıştır. Karar verme sürecinde ise makine öğrenimi yöntemleri kullanılmıştır. Bu proje, tümör bölgesini belirlemek için görüntü işleme yöntemlerini kullanmaktadır. Görüntü işleme yöntemleri sayesinde tümörün kanserli olup olmadığı hakkında bilgi elde edilmektedir. Ancak, makine öğrenimi yöntemleri bu çalışmada kullanılmamıştır (Poyraz et al., 2021).

Ameerah Mustafa Mohammed Alezabi'nin 2021 yılında gerçekleştirdiği araştırma, melanom cilt kanserini tespit etmek için görüntü işleme tekniklerini kullanmayı hedeflemektedir. Bu teknikler, ilgilenilen bölgenin belirlenmesine yardımcı olur ve görüntüyü daha anlaşılır hale getirme amacını taşır. Çalışmada, Pedro Hispano Hastanesi Dermatoloji Servisi'nden elde edilen görüntüler kullanılmıştır. Renk uzayı dönüşümü, kontrast iyileştirme, lezyon lokalizasyonu, artefakt giderimi, sınır genişletme, bölütleme, eşikleme, kümeleme tabanlı yöntemler, histogram tabanlı yöntemler, yapay zeka, makine öğrenimi ve derin öğrenme gibi çeşitli görüntü işleme yöntemleri kullanılmıştır. Yapılan başarım analizi sonuçlarına göre, veri tabanındaki görüntüler üzerinde %93'lük en yüksek sınıflandırma doğruluğu elde edilmiştir (Alezabı, 2021).

Emel Soylu'nun 2021 yılında yaptığı çalışma, cilt kanseri tanısı için bilgisayar analizi kullanmayı hedeflemektedir. Bu çalışmada, deri üzerindeki lekeli görüntüler, derin transfer öğrenme teknikleri kullanılarak iyi huylu veya kötü huylu olarak sınıflandırılmaktadır. Çalışmada, International Skin Imaging Collaboration (ISIC) tarafından sağlanan bir veri seti kullanılmıştır. DarkNet-53 adlı 53 katmanlı bir evrişimli sinir ağı ve YOLOv3 nesne algılama yöntemi gibi derin transfer öğrenme teknikleri, Matlab aracılığıyla uygulanmıştır. Elde edilen sonuçlara göre, DarkNet-19 mimarisi en yüksek performansı göstererek %89.89 doğruluk oranına ulaşmıştır. Diğer mimariler olan

DarkNet-53, ShuffleNet ve PinchNet ise sırasıyla %87.36, %86.15 ve %84.23 doğruluk oranlarına sahip olmuştur (Soylu & Demir, 2021).

Volkan Göreke tarafından gerçekleştirilen 2021 yılındaki araştırma, cilt lezyonlarının sınıflandırılmasında derin öğrenme temelli bir yöntemin kullanılmasını amaçlamaktadır. Bu çalışmada, öncelikle VGG16 modeli kullanılarak derin sinir ağı tarafından görüntülere ait öznitelikler elde edilmektedir. Çalışmada ISIC 2017 veri seti kullanılmış ve CNN (Convolutional Neural Network) ile RNN (Recurrent Neural Network) yöntemleri karşılaştırılmıştır. Veri artırma teknikleriyle önceden eğitilmiş VGG16 modeli kullanılarak öznitelikler çıkarılmıştır. Sonuçlara göre, CNN tabanlı derin sinir ağı %96 sınıflandırma doğruluğu ve %100 AUC başarımı elde etmiştir (Göreke, 2021).

Tolga Yalçın'ın 2021 yılında gerçekleştirdiği araştırma, cilt kanseri teşhisinde konvolüsyonel sinir ağları tabanlı bilgisayar destekli tanı (CNN-CAD) yöntemi kullanarak dijital görüntü kalitesini iyileştirmeyi hedeflemektedir. Çalışma, en yaygın beş cilt kanseri türünün teşhisinde kullanılacak görüntülerin önceden işlenerek daha iyi hale getirilmesiyle ilgili sonuçları sunmaktadır. Araştırmada, The International Skin Imaging Collaboration (ISIC) 2019 tarafından sağlanan veri seti kullanılmıştır. Görüntülerin kalitesini artırmak için uyarlanabilir medyan filtresi, medyan filtresi, Gauss yumuşatma ve ortalama filtresi gibi ön işleme adımları uygulanmıştır. Ayrıca, görüntüler filtrelenmiş ve bölütlenmiştir. Elde edilen sonuçlara göre, 30 epoch süresinde %89 doğruluk elde edilmiştir (Yalçın & Gürsel, 2021).

Junaid Iqbal'ın 2021 yılında gerçekleştirdiği araştırma, cilt kanserinin sınıflandırılmasında derin sinir ağı kullanımını hedeflemektedir. Bu çalışmada, el yapımı özellik çıkarıcılarının yerini alan derin ağlar ve evrişimli sinir ağları üzerinde odaklanılmıştır. Çalışmada, International Skin Imaging Collaboration (ISIC) veri seti kullanılmıştır. Monogenik ikili desen, yerel Gabor ikili model histogram dizisi gibi öznitelik çıkarıcıları kullanılmıştır. Derin öğrenme yaklaşımı için önceden eğitilmiş bir sinir ağı olan CNN (Convolutional Neural Network) kullanılmıştır. Birden fazla elle tasarlanmış öznitelik çıkarıcısı kullanılmış ve elde edilen sonuçlar %77.55 doğruluk değeri ile en yüksek performansı sergilemiştir (Iqbal et al., 2021).

Cihan Akyel'in 2022 yılında yaptığı araştırma, cilt kanseri görüntülerinde kıl temizliği ve lezyon bölütleme için FCN8-ResNetC ve görüntü işleme yöntemlerini kullanmaktadır. Bu çalışmanın amacı, lezyon görüntülerindeki kıl gürültüsünü doğru bir

şekilde temizleyerek doğru lezyon bölütlemesi yapmaktır. Bu sayede cilt kanseri teşhisinde başarı oranını artırmayı hedeflemektedir. Çalışmada, ISIC 2018 veri setinden iki farklı veri seti ve PH2 veri seti kullanılmıştır. FCN8 algoritması, ResNetC, U-Net ve LinkNet gibi yöntemler kullanılarak kıl temizliği ve lezyon bölütlemesi gerçekleştirilmiştir. Bu adımlar, görüntülerin optimize edilmesi ve bölütlenmesi sürecini içermektedir. Elde edilen sonuçlara göre, kıl temizliği eğitiminde %89.38 başarı oranı elde edilmiş, lezyon bölütlemesi ise %97.05 başarı oranına ulaşılmıştır. Bu sonuçlar, cilt kanseri teşhisinde daha doğru ve etkili bir yöntem geliştirilmesine katkı sağlamaktadır (Akyel & Arıcı, 2022).

Rukhsar ve arkadaşları 2022 yılında, RNA-Seq gen ekspresyon verilerinin derin öğrenme yaklaşımları kullanılarak kanser sınıflandırılması incelenmiştir. Çalışma, Mendeley veri deposundan alınan beş kanser türüne ait verileri analiz etmektedir: LUAD, BRCA, KIRC, LUSC ve UCEC. İlk adımda, RNA-Seq değerleri normalizasyon ve sıfır dolgu kullanılarak 2D görüntülere dönüştürülmüştür. İkinci adımda, derin öğrenme ile ilgili özellikler çıkarılmış ve seçilmiştir. Son aşamada, sekiz farklı derin öğrenme algoritması kullanılarak sınıflandırma yapılmıştır. En iyi sonuçlar %70 eğitim ve %30 test verisi kullanıldığında elde edilmiştir. En iyi performans Convolutional Neural Network (CNN) ile elde edilmiştir ve bu modelin doğruluk oranı %97'dir (Rukhsar et al., 2022).

Ogudo ve arkadaşları 2023 yılında, dermoskopik görüntülerdeki deri lezyonlarının tespiti ve sınıflandırılması için optimize edilmiş bir yapay zekâ modeli olan OSSAE-BPNN'yi (Optimize Edilmiş İstiflenmiş Seyrek Otokodlayıcı ve Geri Yayılım Sinir Ağı) sunmaktadır. Model, görüntülerin çok seviyeli eşikleme tabanlı segmentasyonunu, OSSAE tabanlı özellik çıkarımını ve SGO (Deniz Kuşu Optimizasyonu) algoritması ile parametre optimizasyonunu içermektedir. OSSAE-BPNN modeli, deri lezyonlarını otomatik olarak tespit etmek ve sınıflandırmak için dermoskopik görüntüler kullanmaktadır. Çalışma, yedi farklı sınıfa ait deri lezyonları üzerinde kapsamlı deneysel analizler gerçekleştirmiştir: Angioma, Nevus, Lentigo NOS, Solar Lentigo, Melanoma, Seboreik Keratoz ve Bazal Hücreli Karsinom. Deney sonuçları, OSSAE-BPNN modelinin diğer mevcut stratejilerden daha üstün performans gösterdiğini ortaya koymuştur. Modelin ortalama doğruluğu %96.2, duyarlılığı %85.3, özgüllüğü %97.8, kesinliği %85.4 ve F-skoru %84.9 olarak ifade edilmiştir (Ogudo et al., 2023).

Orhan ve Yavşan 2023 yılında, melanoma teşhisi için derin öğrenme teknikleri kullanan yapay zekâ destekli bir tespit modeli önerilmiştir. Çalışmada, iki farklı veri seti birleştirilerek toplamda 8598 görüntüden oluşan zengin bir veri seti oluşturulmuştur. Model geliştirilirken AlexNet, MobileNet, ResNet, VGG16 ve VGG19 gibi derin öğrenme algoritmaları kullanılmıştır. Eğitim, doğrulama ve test aşamalarında bu modellerin performansları değerlendirilmiştir. MobileNet modeli, %84.94 doğruluk oranıyla en yüksek performansı göstermiştir (Orhan & Yavşan, 2023).

Hussain ve arkadaşları, maymun çiçeği enfeksiyonlarının deri virolojik görüntülerinden tanı koymak için hiper-graf konvolüsyonel sinir ağları (HGCNs) tabanlı yeni bir derin öğrenme modelini önermektedir. Model, geleneksel CNN'lerin yerel özellik çıkarma gücünü, hiper-graf yapılarını kullanarak daha yüksek düzeydeki ilişkileri ve karmaşık özellik etkileşimlerini modelleme yeteneğiyle birleştirmektedir. Bu sayede, enfekte bölgelerin daha doğru tespit edilmesi ve sınıflandırılması sağlanmaktadır. HGCN, iki farklı doğrulama yöntemi (Holdout ve 3 katmanlı çapraz doğrulama) ile test edilmiş ve Holdout yöntemiyle %98,88 doğruluk, %98,13 hassasiyet ve %99,58 geri çağırma değerleri elde etmiştir. 3 katmanlı çapraz doğrulama yöntemi ise %99,17 doğruluk, %99,31 hassasiyet ve %99,12 geri çağırma sonuçlarıyla modelin performansını daha da ileri taşımıştır. Ayrıca, Grad-CAM yöntemi kullanılarak enfekte bölgelerin görselleştirilmesi, modelin açıklanabilirliğini ve klinik faydasını artırmıştır (Hussain et al., 2024).

Hosny ve arkadaşları, cilt lezyonlarını sınıflandırmak için açıklanabilir derin öğrenme yaklaşımına dayalı yeni bir model önermektedir. Çalışmada, "deep inherent learning" olarak adlandırılan bir teknik geliştirilmiş ve bu yöntemle yedi farklı cilt lezyonu sınıfı sınıflandırılmıştır. Model, sınırlı veri ve görüntü bozulmaları gibi sorunları aşmak için birden fazla evrişim filtresi ve açıklanabilir yapay zeka (XAI) tekniklerini kullanmaktadır. Bu sayede hem yerel hem de genel düzeyde karar verme süreçleri açıklanarak modelin güvenilirliği artırılmıştır. Model, ISIC 2018 (HAM10000) veri seti kullanılarak test edilmiş ve %92,89 doğruluk, %95,57 özgüllük ve %76,85 kesinlik gibi yüksek performans değerleri elde etmiştir. Ayrıca, occlusion sensitivity ve özellik görselleştirme teknikleri ile modelin hangi alanlara odaklandığı açıklanabilir hale getirilmiştir. Bu yaklaşım, doktorların yapay zekaya dayalı teşhis sistemlerine güvenini artırmayı ve lezyonları doğru bir şekilde tespit etmeyi amaçlamaktadır (Hosny et al., 2024).

Murmu ve arkadaşları, deri lezyonu sınıflandırmasında yüksek doğruluk elde etmek için UNet-MobileNet adlı bir derin öğrenme modeli önermektedir. Model, UNet'in detaylı özellik çıkarma yeteneğini MobileNet'in hafif ve hızlı sınıflandırma yetenekleriyle birleştirmektedir. Özellikle, UNet'in kodlayıcı ve kod çözücü blokları ile MobileNet tabanlı bir sınıflandırıcı entegrasyonu sayesinde model, deri lezyonlarının hem semantik hem de bağlamsal bilgilerini etkili bir şekilde işlemektedir. Eğitim sırasında aşırı öğrenmeyi önlemek için MobileNet, UNet kod çözücüsü yerine kullanılmıştır. HAM10000 veri seti kullanılarak test edilen model, yedi farklı deri lezyonu sınıfında %94,27 doğruluk ve %95,87 geri çağırma değeri elde etmiştir. Modelin hafif yapısı ve yüksek performansı, onu deri kanseri teşhisinde etkili bir araç haline getirmiştir (Murmu et al., 2024).

Ghazouani, deri lezyonu sınıflandırmasında yüksek doğruluk sağlamak için Multi-Residual Attention Network (MuRANet) adlı bir derin öğrenme modeli önermektedir. Model, Multi-Scale Dilated Attention kanal ve uzaysal dikkat mekanizmaları ile residual paths üzerinden entegre dikkat modülleri kullanarak lezyonların bağlamsal ve ince detaylarını etkili bir şekilde yakalamaktadır. HAM10000, ISIC-2019 ve PH2 veri setlerinde test edilen model, sırasıyla %90, %88 ve %97 F1 skorları ve %98 AUC değeri elde edilmiştir. Bu sonuçlar, MuRANet'in cilt kanseri teşhisinde doğruluk ve genelleme yeteneği açısından önemli bir potansiyele sahip olduğunu ortaya koymaktadır (Ghazouani, 2025).

Tablo 2'de bu alanda farklı derin öğrenme mimarileri ve farklı veri setleriyle ilgili yapılan çalışmalar ve bu çalışmalara ait özellikler yer almaktadır. Bu araştırmalar, derin öğrenme modellerinin cilt kanseri teşhisinde etkili bir şekilde kullanılabileceğini göstermektedir. Her bir modelin kendi avantajları ve dezavantajları olsa da genel olarak yüksek doğruluk oranları elde ederek cilt kanseri teşhisinin geliştirilmesine katkıda bulunmuşlardır. Bununla birlikte, her araştırma farklı yöntemler kullandığından, sonuçlar arasında doğrudan bir karşılaştırma yapmak zor olabilir (Bulam et al., 2010). Ancak genel olarak, derin öğrenme modellerinin cilt kanseri teşhisinde umut vadeden bir yol olduğu söylenebilir.

Cilt kanseri tespiti üzerine yapılan kaynak araştırmasının genel bir değerlendirmesi yapıldığında bu alanda gerçekleştirilen çalışmaların artış gösterdiğini ve bu çalışmaların elde ettiği sonuçların umut verici olduğunu ortaya koymaktadır. Derin öğrenme yöntemleri, cilt kanseri teşhisinde yüksek performans sergileyerek önemli bir

potansiyele sahiptir. Ancak, bu alanda hala geliştirmeye ihtiyaç duyulan noktalar bulunmaktadır. Özellikle daha geniş ve çeşitli veri setlerine erişimin sağlanması, daha etkili özellik çıkarımı tekniklerinin geliştirilmesi ve daha ileri model tasarımlarının kullanılması gibi alanlarda çalışmaların sürdürülmesi önemlidir. Bu sayede cilt kanseri teşhisinde daha hassas, doğru ve erken teşhis imkânı sağlanabilecektir.

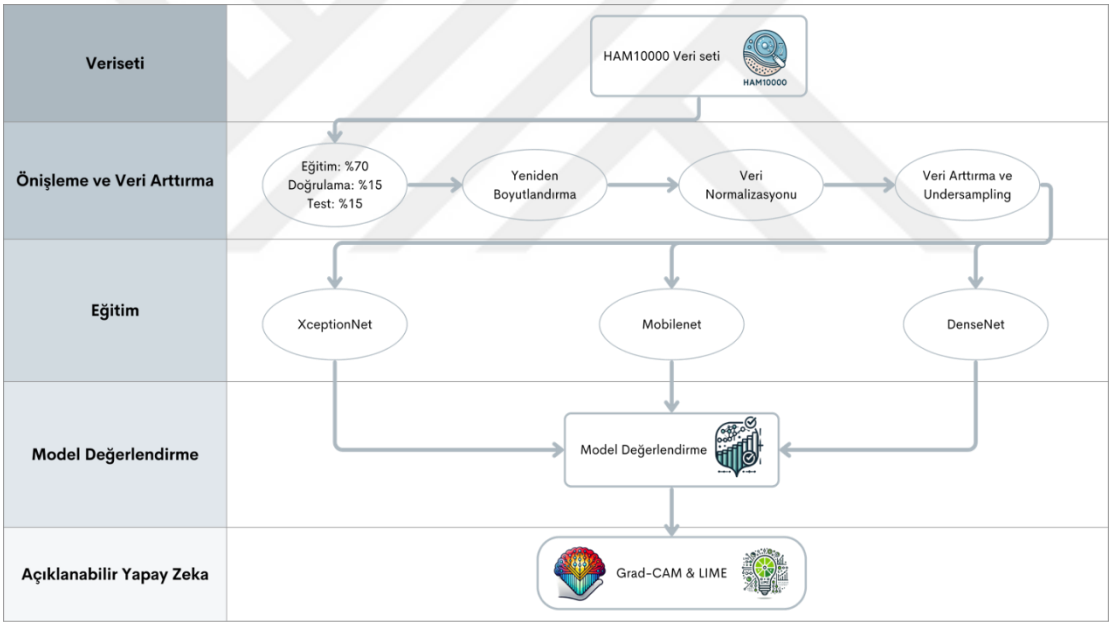
Son yıllarda yapay zekâ tabanlı sistemler, özellikle derin öğrenme yöntemleri sayesinde tıbbi görüntü analizinde oldukça başarılı sonuçlar elde etmektedir (Esteva et al., 2017). Ancak, bu modellerin "kara kutu" doğası, kullanıcıların ve özellikle sağlık profesyonellerinin sistemlerin karar mekanizmalarına duyduğu güveni sınırlandırmaktadır (Tjoa & Guan, 2020). Bu bağlamda, Açıklanabilir Yapay Zekâ (XAI), modelin verdiği kararların arkasındaki mantığın insan tarafından anlaşılır hale getirilmesini sağlayarak kritik bir rol oynamaktadır (Samek et al., 2017). XAI, yalnızca sonuca ulaşmakla kalmaz; aynı zamanda karar sürecine dair bilgi sunar. Bu, özellikle hayati kararların alındığı tıbbi teşhis süreçlerinde oldukça önemlidir (Amann et al., 2020). Bir cilt lezyonunun malign olup olmadığına dair verilen kararın açıklanabilirliği, hekimlerin sistem önerisini klinik bilgiyle birlikte yorumlamasını mümkün kılar. Böylece, modelin yalnızca bir tahmin mekanizması değil, aynı zamanda karar destek aracı olarak kullanılabilmesi sağlanır (Ghassemi et al., 2021). Bu nedenle, açıklanabilir yapay zekâ yöntemleri, hem araştırma hem de uygulama açısından büyük önem taşımaktadır.

Tablo 2: Cilt kanseri teşhisi konusunda yapılan çalışmalar ve özellikleri

YAZAR	ÇALIŞMA ADI	VERİ SETİ	DESEN	ÖRNEKLEM
Sarmad MOHSEN Ali HURMUZI (2019)	Evrişimsel sinir ağları kullanarak cilt kanseri risk değerlendirmesi	ISIC (The International Skin Imaging Collaboration)	Evrişimli Sinir Ağları	Evrişimli Sinir Ağları
Nuh ALPASLAN (2020)	SegNet Mimarisi Kullanılarak Cilt Lezyon Bölütleme Performansının İyileştirilmesi	ISBI2016 veri kümesi	SegNet mimarisi tabanlı bir sistem	
Zhang (2020)	Skin cancer diagnosis based on optimized convolutional neural network	Dermquest ve DermIS	İyileştirilmiş Balina Optimizasyon Algoritması (WOA) ve CNN kullanımı	WOA + CNN
Rukshar (2022)	Analyzing RNA-seq gene expression data using deep learning approaches for cancer classification	RNA-Seq gene expression data of BRCA, KIRC, LUAD, LUSC ve UCEC	1D RNA-Seq verilerini 2D görüntülere dönüştürmüştür. Sınıflandırma için sekiz farklı derin öğrenme modeli kullanılmıştır.	CNN, ResNet50, ResNet101, ResNet152, VGG16, VGG19, AlexNet, GoogleNet
Ogudo (2023)	Optimal Artificial Intelligence Based Automated Skin Lesion Detection and Classification Model	Dermoscopy Images Benchmark Dataset	Bu çalışma, OSSAE-BPNN tekniği olarak adlandırılan geri yayılımlı sinir ağı (BPNN) ile optimize edilmiş yığılmış seyrek otokodlayıcı (OSSAE) tabanlı özellik çıkarıcı kullanan otomatik bir cilt lezyonu algılama ve sınıflandırma tekniği sunmaktadır.	OSSAE-BPNN
Orhan ve Yavşan (2023)	Artificial intelligence-assisted detection model for melanoma diagnosis using deep learning techniques	Nevus classifier ve Skin lesions veri kümeleri	Melanom tespiti için veri çeşitliliğini ve gerçekçiliğini artırmaya yönelik birleşik veri kümesi yaklaşımı	MobileNet, AlexNet, ResNet, VGG16, VGG19
Anita Murmu ve arkadaşları (2024)	Deep Learning Based Encoder-Decoder Model for Skin Cancer Classification	HAM10000	UNet ve MobileNet tabanlı bir Encoder-Decoder yaklaşımı	UNet-MobileNet modeliyle 7 sınıflı sınıflandırma
Khalid M. Hosny ve arkadaşları (2024)	Explainable deep inherent learning for multi-classes skin lesion classification	ISIC 2018	Derin İçsel Öğrenme (Deep Inherent Learning), Açıklanabilir Yapay Zeka (XAI)	7 sınıf deri lezyonu, Occlusion Sensitivity ve Özellik Görselleştirme teknikleri
Haythem Ghazouani (2025)	Multi-residual attention network for skin lesion classification	HAM10000, ISIC-2019, PH2	Çoklu Artık Dikkat Ağı (MuRANet)	MuRANet, dikkat mekanizmaları, CNN

3. MATERYAL VE YÖNTEM

Bu bölüm, cilt kanseri tespiti için gerçekleştirilen işlemlerin teknik detaylarını, kullanılan yöntemleri ve açıklanabilir yapay zeka (XAI) tekniklerini içermektedir. Bu bilgiler, ilgili tezin temelini oluşturan veri seti, ön işlem adımları, sınıflandırma yöntemleri, açıklanabilirlik analizleri ve değerlendirme metriklerini kapsamaktadır. Çalışmada, elde edilen sonuçların yalnızca doğruluk ve geçerliliği değil, aynı zamanda modelin karar mekanizmalarının şeffaflığı ve yorumlanabilirliği de sağlanmaktadır. Bu sayede, tezin geri kalan bölümlerinde sunulacak olan bulguların güvenilirliği artırılmış ve cilt kanseri tespitinde açıklanabilir yapay zeka yaklaşımının katkıları vurgulanmıştır. Sürecin genel akışını ve kullanılan adımları gösteren Şekil 1, yöntemin özetini sunmaktadır. Bu şekil, sürecin aşamalarını sıralı ve anlaşılır bir şekilde ortaya koyarak yöntemsel akışın bilimsel bir temelde nasıl ilerlediğini detaylandırmaktadır.

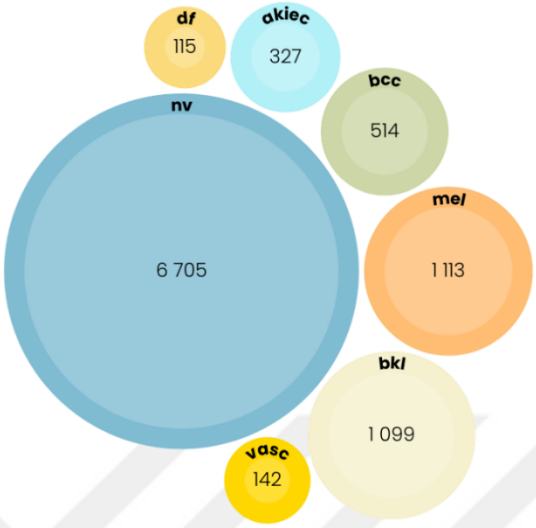


Şekil 1: Yöntemin genel akış şeması

3.1.Veri Seti

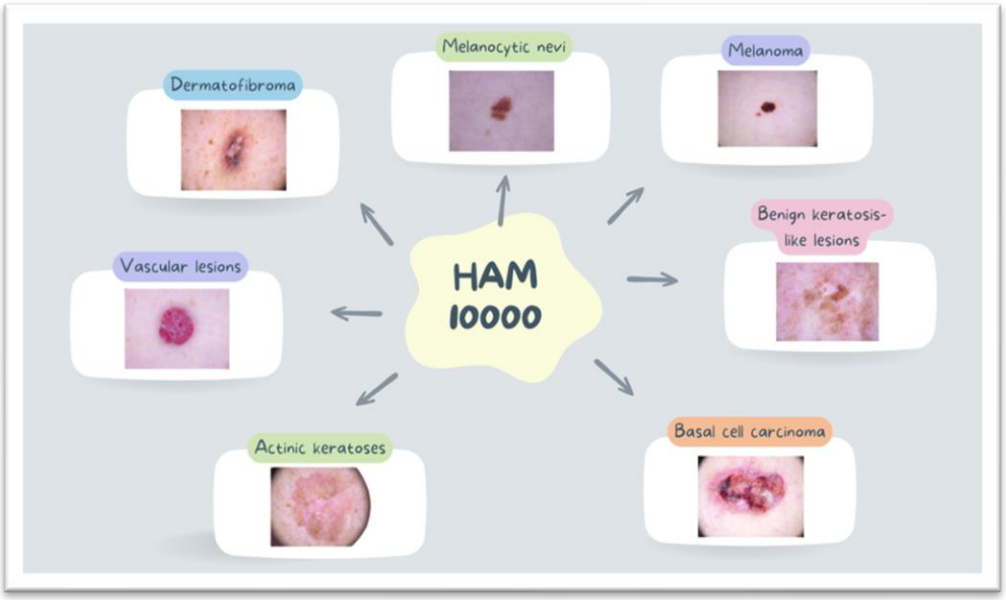
Bu çalışmada, deri lezyonlarının sınıflandırılması amacıyla açık erişimli HAM10000 veri seti kullanılmıştır. Veri seti, toplam 10015 adet dermoskopik görüntü içermekte olup, farklı deri lezyon tiplerine ait yedi sınıfa ayrılmıştır. Görüntüler, dermatologlar tarafından etiketlenmiş ve çeşitli kaynaklardan toplanmıştır. Veri setinde yer alan sınıflar ve bu sınıflara ait görüntü sayıları Şekil 2'de sunulmuştur. Sınıflar, melanom (mel), benign nevüs (nv), bazal hücreli karsinom (bcc), aktinik keratoz (ak),

benign keratoz lezyonları (bkl), dermatofibroma (df) ve vasküler lezyonlar (vasc) olmak üzere yedi farklı gruptan oluşmaktadır (Tschandl et al., 2018).



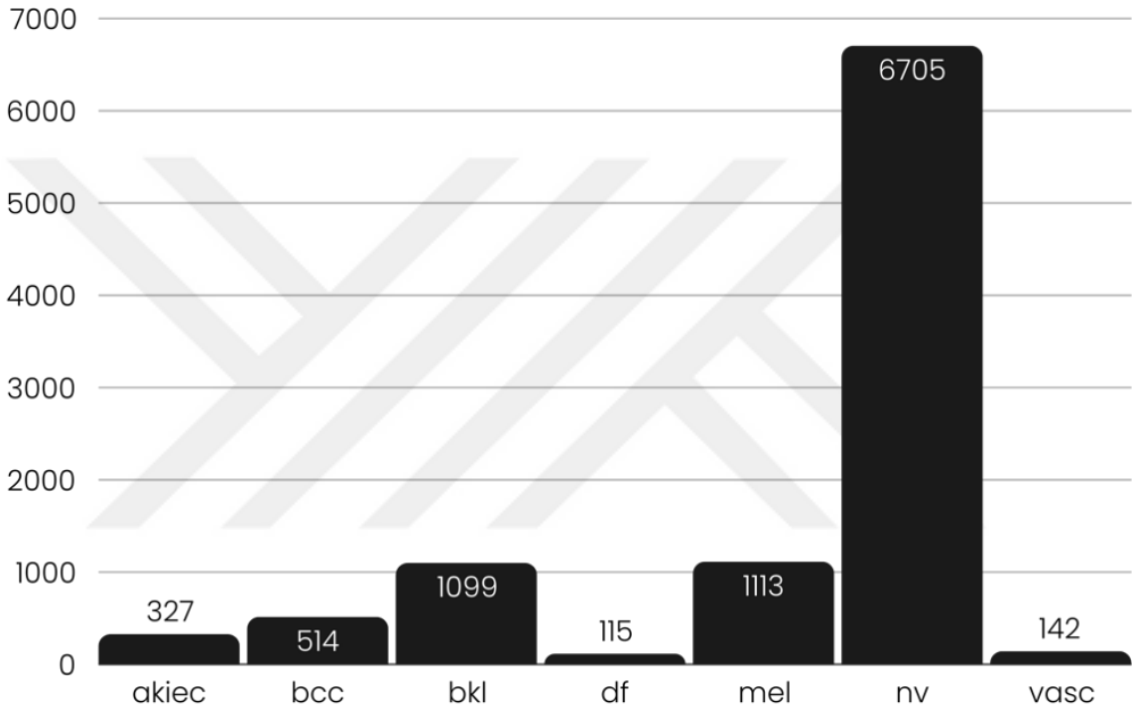
Şekil 2:HAM10000 Veri Seti Sınıf Dağılımı

HAM10000 veri setindeki tüm görüntüler standart JPG formatında kaydedilmiş olup, RGB renk formatında ve farklı cihazlardan elde edilmiştir. Görüntü çözünürlükleri değişkenlik göstermekle birlikte, çoğunlukla 450x600 piksel boyutundadır (Liu et al., 2024). Şekil 3'te veri setinde yer alan lezyon sınıflarından örnek görüntüler verilmiştir. Bu şekil, her bir sınıfı temsilen seçilmiş dermoskopik görüntüler içermektedir.



Şekil 3: HAM10000 veri setindeki cilt lezyonu Sınıfları

Şekil 4'te ise veri setindeki sınıfların dağılımını gösteren bir sütun grafiği yer almaktadır. Bu grafik, sınıflar arasındaki dengesiz veri dağılımını görselleştirerek veri setinin genel yapısını anlamayı kolaylaştırmaktadır. Veri seti, farklı kaynaklardan toplanmış olması nedeniyle lezyonların görsel karakteristiklerinde çeşitlilik içermektedir. Bu çeşitlilik, modelin genel performansını artırmaya yönelik olarak avantaj sağlamaktadır. Veri setinin etiketleri, uzman dermatologlar tarafından doğrulanmış olup çalışmanın güvenilirliğini artırmaktadır (Alam et al., 2022).



Şekil 4: HAM 10000 veri setindeki sınıf dağılımı

3.2.Veri Önişleme

Veri önişleme, ham verinin analiz ve modelleme süreçleri için uygun hale getirilmesini amaçlayan temel bir aşamadır. Bu süreç, veri kümesindeki gürültüyü azaltarak, modelin genelleştirme yeteneğini artırmak ve öğrenme sürecini iyileştirmek için uygulanır (Famili et al., 1997). Veri önişleme, eksik veya hatalı verilerin düzeltilmesi, veri ölçeklendirme, özellik mühendisliği ve veri dengesizliği gibi çeşitli aşamalardan oluşur (Alexandropoulos et al., 2019). Makine öğrenmesi ve derin öğrenme uygulamalarında veri kümesinin istatistiksel ve yapısal bütünlüğü, modelin başarımını doğrudan etkileyen en önemli faktörlerden biridir. Veri önişleme teknikleri, modelin

doğruluğunu artırırken, öğrenme sürecindeki önyargıları minimize etmek için de kritik bir rol oynar (Maharana et al., 2022).

3.2.1. Veri Dengesizliği ve Veri Dengeleme

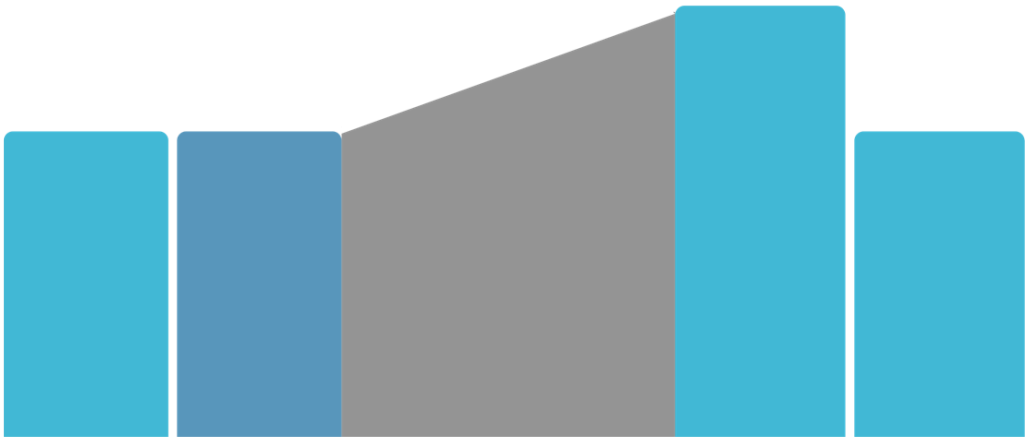
Makine öğrenmesi modellerinin başarıyla eğitilebilmesi için veri kümesindeki sınıf dağılımının dengeli olması büyük önem taşımaktadır. Veri dengesizliği, özellikle sınıflar arasında büyük farklılıklar olduğunda modelin az temsil edilen sınıfları öğrenmekte zorlanmasına ve hatalı tahmin yapmasına neden olabilir (Viloria et al., 2020). Bu durum, modelin yalnızca baskın sınıflara odaklanmasına ve azınlık sınıflarını görmezden gelmesine yol açarak tahmin performansını olumsuz etkiler. Bu bağlamda, veri dengesizliğini gidermek amacıyla iki temel yöntem uygulanmaktadır (Mohammed et al., 2020):

Oversampling: Azınlık sınıflarına ait veri noktalarının sayısını artırarak sınıf dengesizliğini gidermeyi amaçlar. Bu yöntem, orijinal veriye sadık kalarak, sentetik veri üretme teknikleriyle desteklenebilir. Yaygın aşırı örnekleme teknikleri arasında Veri Artırma ve SMOTE (Synthetic Minority Over-sampling Technique) gibi yöntemler bulunmaktadır (Won et al., 2020).

Undersampling: Baskın sınıfların örnek sayısını azaltarak veri dengesini sağlamayı hedefler. Bu yaklaşım, rastgele örnekleme veya kümeleme teknikleri kullanılarak gerçekleştirilebilir. Ancak, eksik örnekleme yöntemi veri kaybına neden olabileceğinden dikkatli uygulanmalıdır (Yang et al., 2024).

3.2.2. Undersampling

Çalışmada kullanılan veri setindeki sınıf dengesizliği, modelin performansını olumsuz etkileyebilecek önemli bir sorun olarak tespit edilmiştir. Bu dengesizlik, bazı sınıfların diğerlerine göre çok daha fazla sayıda örneğe sahip olması nedeniyle modelin bu sınıfları öncelikli olarak öğrenmesine ve az temsil edilen sınıflardaki örnekleri yetersiz öğrenmesine yol açmaktadır (Krawczyk et al., 2016). Bu problemi çözmek ve modelin tüm sınıflar için adil bir şekilde öğrenmesini sağlamak amacıyla undersampling yöntemi uygulanmıştır.



Şekil 5: İki sınıf için undersampling uygulama örneği

Şekil 5'te gösterildiği gibi undersampling, veri setindeki çoğunluk sınıfından rastgele örnekler çıkararak veri setinin boyutunu küçültmek ve sınıflar arasındaki dengesizliği azaltmak için kullanılan bir tekniktir (Hasanin & Khoshgoftaar, 2018). Bu çalışmada, çoğunluk sınıfındaki örneklerin belirli bir oranı rastgele seçilerek analiz dışında bırakılmış ve veri seti yeniden dengelenmiştir. Böylelikle, modelin azınlık sınıflarını da daha iyi öğrenmesi için uygun bir öğrenme ortamı oluşturulmuştur.

3.3.MobileNet

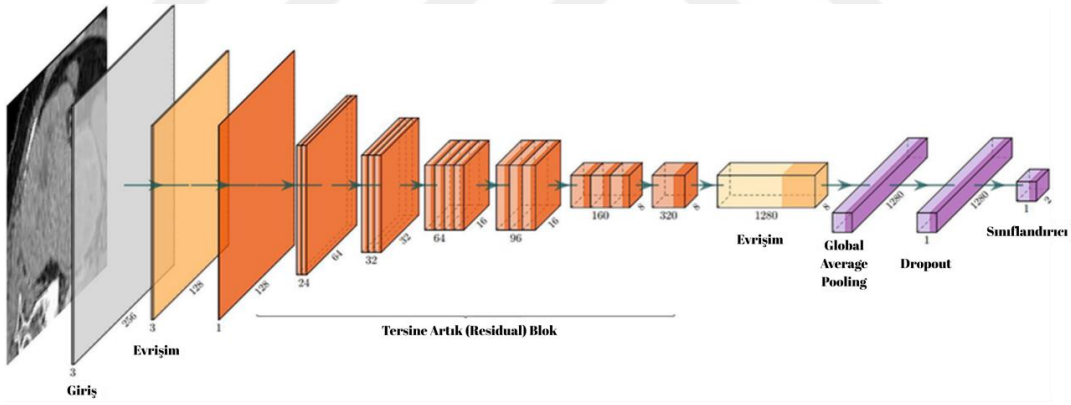
Bu çalışmada, sınıflandırma modelinin temelini oluşturmak için MobileNet mimarisi kullanılmıştır. MobileNet, derin sinir ağlarının hafif ve verimli bir varyantı olarak geliştirilmiştir ve özellikle sınırlı hesaplama kapasitesine sahip cihazlarda (örneğin, mobil cihazlar ve gömülü sistemler) kullanılmak üzere tasarlanmıştır. Bu mimari, geleneksel konvolüsyon işlemlerini optimize ederek hem parametre sayısını azaltmakta hem de hesaplama maliyetini düşürmektedir. MobileNet'in temelinde, derinlik ayrıştırmalı konvolüsyonlar (Depthwise Separable Convolutions) bulunmaktadır. Bu yaklaşım, standart konvolüsyonları iki ayrı aşamaya ayırarak işlem maliyetini büyük ölçüde azaltır (Howard, 2017).

1. Derinlik Ayrıştırmalı Konvolüsyon: Görüntüdeki her kanal üzerinde ayrı ayrı konvolüsyon işlemi gerçekleştirilir.
2. Noktasal Konvolüsyon: Ardından, 1x1 boyutundaki konvolüsyonlar kullanılarak kanallar arasındaki ilişkiyi öğrenir.

Bu iki aşama, geleneksel konvolüsyonlara kıyasla daha az sayıda parametre ve hesaplama gerektirir. MobileNet mimarisi, ağın esnekliğini artırmak için iki hiperparametre ile kontrol edilebilir:

1. Genişlik Katsayısı (Width Multiplier): Ağın kanal sayısını azaltarak veya artırarak modelin hesaplama maliyetini ayarlar.
2. Çözünürlük Katsayısı (Resolution Multiplier): Giriş görüntüsünün çözünürlüğünü azaltarak hesaplama maliyetini düşürür (Chen & Su, 2018).

MobileNet mimarisi, 28 katmanlı bir yapıdan oluşur. Ağ, sırasıyla derinlik ayrıştırmalı konvolüsyonlar, noktasal konvolüsyonlar ve doğrusal olmayan aktivasyon fonksiyonları (ReLU) ile donatılmıştır. Tüm ağın sonunda, ortalama havuzlama (Global Average Pooling) katmanı kullanılarak özellik haritaları bir vektöre indirgenir ve ardından Softmax katmanı ile sınıf tahmini gerçekleştirilir (Zhu et al., 2022). MobileNet'in genel mimarisi Şekil 6'da gösterilmiştir. Şekilde, derinlik ayrıştırmalı konvolüsyonların ve 1x1 konvolüsyonların ağı oluşturmadaki sıralı yapısı açıklanmaktadır.

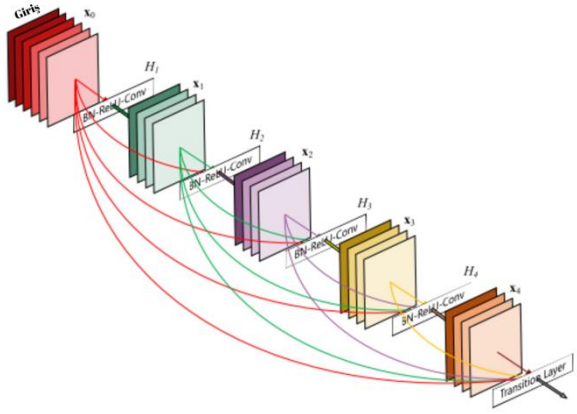


Şekil 6: MobileNetV2 modelinin genel mimarisi (Zuo et al., 2021)

MobileNet'in tercih edilme nedenleri arasında hesaplama verimliliği, küçük boyutlu modeller üretmesi ve sınıflandırma performansı yer almaktadır. Bu özellikler, özellikle HAM10000 veri seti gibi yüksek çözünürlüklü görüntülerin sınıflandırılmasında önemli bir avantaj sağlamaktadır (Zhu et al., 2022).

3.4.DenseNet

Bu çalışmada kullanılan bir diğer derin öğrenme mimarisi DenseNet (Densely Connected Convolutional Networks)'tir. DenseNet, özellikle derin sinir ağlarının öğrenme kapasitesini artırmak amacıyla geliştirilmiş yenilikçi bir mimaridir. DenseNet'in temel özelliği, ağ içindeki tüm katmanların birbirine yoğun şekilde bağlanmasıdır. Bu bağlantı yapısı, klasik konvolüsyonel sinir ağlarına (CNN) kıyasla öğrenmeyi iyileştirirken, parametre sayısını azaltarak hesaplama verimliliğini artırır (Albelwi, 2022). DenseNet, her bir katmanın kendisinden önceki tüm katmanlardan gelen çıktıları girdi olarak kullanması prensibiyle çalışır. Geleneksel ağlarda, bir katman yalnızca bir önceki katmandan veri alırken, DenseNet'te her katman, kendinden önceki tüm katmanların çıktılarıyla beslenir. Bu yapı, ağın gradyan yayılımını iyileştirerek öğrenme sürecini daha kararlı hale getirir ve derin katmanlı ağlarda görülen gradyan kaybolması problemini büyük ölçüde azaltır. DenseNet mimarisinin temel yapı taşı Dense Block olarak adlandırılır. Her Dense Block içerisinde, katmanlar yoğun bağlantılarla birbirine bağlanır ve bu bloklar arasında geçiş katmanları (Transition Layers) kullanılır. Geçiş katmanları, özellik haritalarının boyutunu küçültmek için 1x1 konvolüsyon ve ardından ortalama havuzlama işlemlerini uygular. Bu yapı, ağın parametre sayısını ve hesaplama maliyetini kontrol altında tutar (Huang et al., 2017).



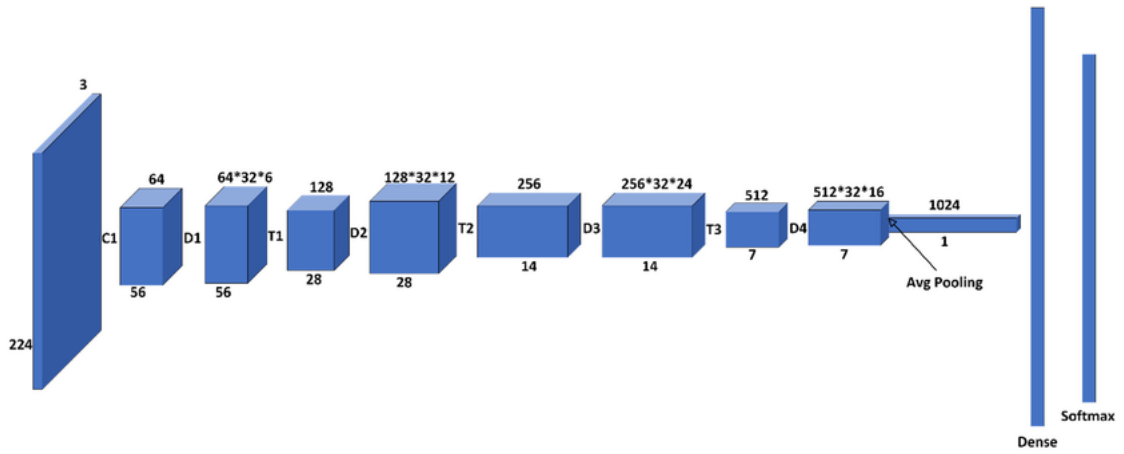
Şekil 7:DenseNet121 model mimarisinin temel yapısı (Arora, August 2, 2020)

Şekil 7'de DenseNet mimarisinin temel yapısı gösterilmektedir. Bu şekilde, bir Dense Block içerisindeki katmanlar arasındaki yoğun bağlantı yapısı açıklanmakta ve geçiş katmanlarının rolü vurgulanmaktadır. Her katmanın önceki katmanlardan gelen çıktıları birleştirme işlemi, özellik haritalarının sayısının giderek artmasına neden olur

(Zhu & Newsam, 2017). Ancak DenseNet, bu özellik haritalarını birleştirmek yerine "büyüme oranı" (growth rate) adı verilen bir parametre kullanarak özellik haritalarındaki artışı sınırlandırır. Büyüme oranı, her bir katmanın çıktısının ne kadarının bir sonraki katmana taşınacağını belirler (Huang et al., 2018). DenseNet, ağ içindeki bilgi akışını güçlendirirken gereksiz hesaplama maliyetlerinden kaçınır. Özellikle, ağın daha az parametreyle daha iyi performans göstermesini sağlayarak, derinlik arttıkça performans kaybı riskini ortadan kaldırır. DenseNet mimarisinde aşağıdaki avantajlar öne çıkmaktadır (Hou et al., 2024; Noh, 2021):

- **Gradyan Akışı:** Yoğun bağlantılar, gradyanların ağ boyunca daha verimli yayılmasını sağlar.
- **Parametre Verimliliği:** Yoğun bağlantı yapısı, daha az parametre ile öğrenmeyi sağlar.
- **Özellik Yeniden Kullanımı:** Önceki katmanların çıktıları, sonraki katmanlar tarafından yeniden kullanılarak daha zengin bir bilgi aktarımı sağlanır.

DenseNet mimarisi, derinliğine bağlı olarak DenseNet-121, DenseNet-169, DenseNet-201 gibi farklı varyantlara sahiptir (Khoei & Kaabouch, 2022). Bu çalışmada kullanılan varyant, parametre sayısı ve hesaplama verimliliği açısından dengeli olan DenseNet-121'dir. DenseNet-121 toplamda 121 katman içermekte olup, dört Dense Block ve aralarındaki geçiş katmanlarından oluşmaktadır. DenseNet'in genel yapısının sunduğu avantajlar sayesinde, sınırlı veri setleri üzerinde bile daha derin özellikler öğrenmek mümkün hale gelir (Chhabra & Kumar, 2022). Şekil 8'de DenseNet-121 mimarisinin katman yapısı ve katmanların bağlantı şeması gösterilmektedir. Bu yapı, ağın girişten çıkışa kadar olan akışını ve özellik haritalarının yoğun bağlantılarını görselleştirmektedir.



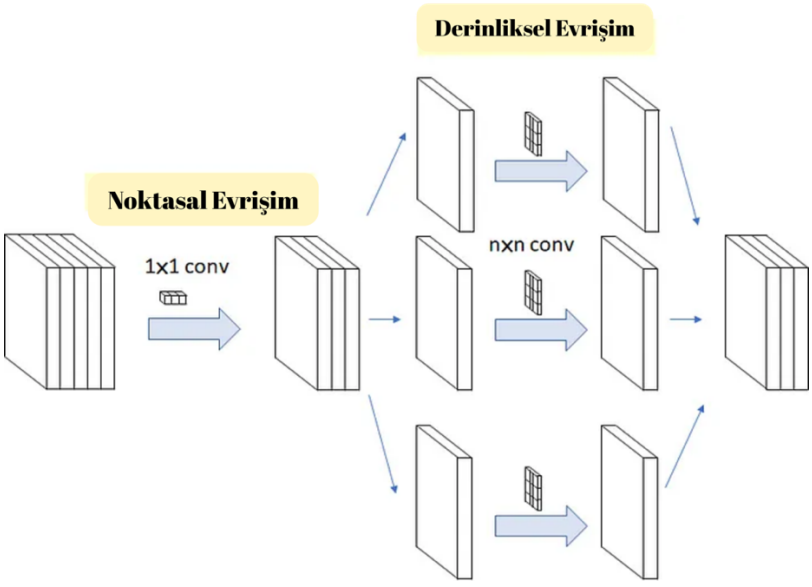
Şekil 8: DenseNet-121 mimarisinin katman yapısı (Tiwari et al., 2023)

3.5.Xception

Bu çalışmada kullanılan bir diğer derin öğrenme modeli Xception mimarisidir. Xception, Inception mimarisinin genişletilmiş ve optimize edilmiş bir versiyonu olarak geliştirilmiştir. Modelin temelinde, derinlik ayrıştırmalı konvolüsyonların daha verimli bir şekilde kullanılması yer alır (Chollet, 2017). Xception, klasik Inception mimarisindeki çoklu konvolüsyon bloklarının karmaşıklığını ortadan kaldırarak, daha sade ancak daha güçlü bir yapı sunar. Xception, temel olarak "derinlik ayrıştırmalı konvolüsyonlar" (Depthwise Separable Convolutions) ile donatılmış bir mimaridir. Bu yapı, konvansiyonel konvolüsyonları iki aşamada gerçekleştirmektedir (Muhammad et al., 2021):

1. Derinlik Ayrıştırmalı Konvolüsyon: Uzaysal konvolüsyonlar yalnızca görüntüdeki her kanal üzerinde ayrı ayrı uygulanır.
2. Noktasal Konvolüsyon (1x1 Convolution): Kanal boyutlarını yeniden birleştirerek özelliklerin kanal bazında ilişkisini öğrenir.

Bu iki adım, standart konvolüsyonlardaki hesaplama yükünü önemli ölçüde azaltırken, aynı zamanda ağın öğrenme kapasitesini artırır. Böylece Xception, parametre verimliliği ve hesaplama performansı açısından avantaj sağlar. Xception mimarisinin bir diğer önemli özelliği, lineer olmayan aktivasyon fonksiyonlarının derinlik ayrıştırmalı konvolüsyon ve noktasal konvolüsyon arasından kaldırılmasıdır. Bu düzenleme, verinin akışını daha verimli hale getirerek ağın doğruluğunu artırmaktadır (Muthu Subathra et al., 2023).

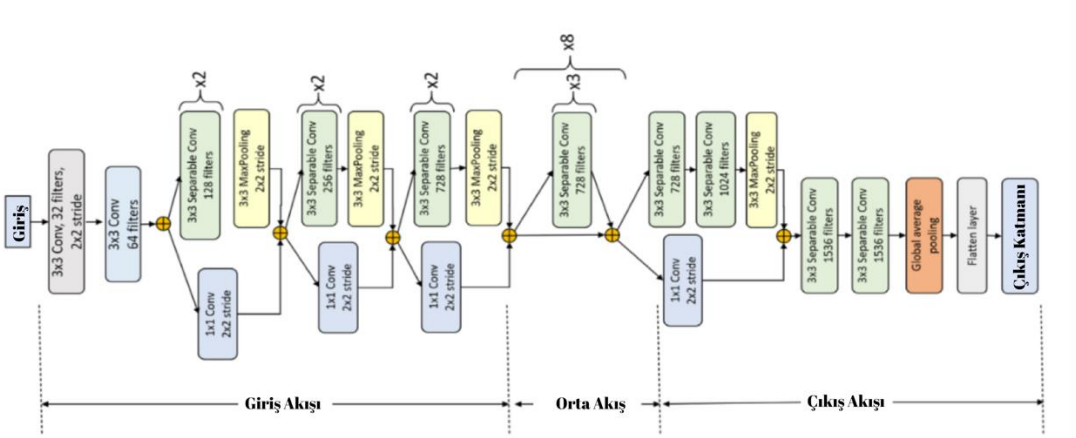


Şekil 9: Xception'te Derinlik Ayrıştırmalı Konvolüsyonların Prensibi (Tsang, 2018)

Şekil 9'da Xception'in temel yapı taşı olan derinlik ayrıştırmalı konvolüsyonların çalışma prensibi açıklanmaktadır. Şekilde, uzaysal filtrelerin bağımsız olarak her kanal üzerinde uygulandığı ve sonrasında noktasal konvolüsyon ile kanalların yeniden birleştirildiği görülmektedir. Xception mimarisi, toplamda 36 konvolüsyon katmanına sahiptir. Bu katmanlar, üç ana aşamada yapılandırılmıştır (Endah & Shiddiq, 2020):

- Giriş Akışı: İlk birkaç konvolüsyon katmanı görüntüde temel özellikleri çıkarır.
- Orta Akış (Middle Flow): Ağın derinlik ayrıştırmalı konvolüsyonlardan oluşan ana gövdesidir. Burada, özelliklerin detaylı bir şekilde öğrenilmesi sağlanır.
- Çıkış Akışı: Özellik haritaları global ortalama havuzlama (Global Average Pooling) ile özetlenir ve sınıflandırma işlemi gerçekleştirilir.

Bu yapı sayesinde, Xception hem hesaplama verimliliği sağlamakta hem de derin katmanlardan öğrenilen özelliklerle yüksek doğruluk performansı göstermektedir. Xception'in hesaplama maliyetini düşürmesine rağmen yüksek performans göstermesi, modelin derinlik ayrıştırmalı konvolüsyonların tam kapasiteyle kullanılması sayesinde mümkün olmuştur. Şekil 10'da Xception'in genel mimarisi görsel olarak gösterilmektedir. Şekilde, ağın girişten çıkışa doğru akışı, orta akıştaki derinlik ayrıştırmalı konvolüsyon bloklarının yapısı ve nihai sınıflandırma katmanı açıklanmaktadır.



Şekil 10: Xception'in genel mimarisi (Mohammed et al., 2023)

3.6.Açıklanabilir Yapay Zekâ

Son yıllarda yapay zekâ (YZ) modelleri, özellikle derin öğrenme mimarileri, çeşitli karmaşık problemlerde insan performansını aşarak önemli başarılar elde etmiştir. Ancak bu modellerin "kara kutu" olarak adlandırılan doğası, karar verme süreçlerinin insanlar tarafından anlaşılmasını zorlaştırmaktadır. Özellikle sağlık alanındaki kritik uygulamalarda, bir modelin verdiği kararın neden ve nasıl verildiğinin açıklanması büyük

bir gereklilik haline gelmiştir. Bu bağlamda, Açıklanabilir Yapay Zekâ yaklaşımları, yapay zeka modellerinin kararlarını daha anlaşılır ve şeffaf hale getirmeyi amaçlamaktadır (Angelov et al., 2021). Açıklanabilir yapay zekâ, bir modelin kararını açıklamak için neden-sonuç ilişkilerini ortaya koyan görsel veya sayısal açıklamalar sunar. Bu sayede, modelin doğruluğu ve güvenilirliği artırılarak kullanıcılar ve uzmanlar tarafından daha fazla kabul görmesi sağlanır. Özellikle tıbbi görüntüleme gibi hassas alanlarda, bir modelin hangi bölgelere odaklandığı veya hangi özellikleri dikkate aldığı gibi açıklamalar, doğru ve güvenilir teşhislerin yapılmasına yardımcı olmaktadır (Ghassemi et al., 2021).

Bu çalışmada, Grad-CAM (Gradient-weighted Class Activation Mapping) ve LIME (Local Interpretable Model-agnostic Explanations) olmak üzere iki farklı açıklanabilirlik yöntemi kullanılmıştır. Bu yöntemler, modelin tahminlerini görsel ve yerel bazda açıklayarak kararların yorumlanabilirliğini artırmıştır. Şekil 12'de bu iki yöntemden elde edilen örnek açıklamalar sunulmuştur. Şekilde, modelin bir görüntü üzerinde hangi bölgelere odaklandığı Grad-CAM ile vurgulanırken, LIME yöntemiyle kararın detaylı ve lokal açıklaması yapılmıştır. Açıklanabilir yapay zekâ yöntemlerinin kullanımı, özellikle yapay zekanın güvenilirliğini artırarak tıbbi karar destek sistemlerinin geliştirilmesinde önemli bir rol oynamaktadır (Yang, 2022). Bu bağlamda, Grad-CAM ve LIME yöntemlerinin teknik detayları aşağıda açıklanmıştır.

3.6.1. Grad-CAM

Grad-CAM, derin öğrenme modellerinin karar verme süreçlerini görsel olarak açıklamak için kullanılan popüler bir açıklanabilir yapay zekâ yöntemidir. Grad-CAM, bir görüntü üzerinde modelin tahminine katkıda bulunan bölgeleri vurgulayan bir ısı haritası oluşturarak, modelin odaklandığı alanları insan tarafından anlaşılır hale getirir. Bu yöntem, özellikle konvolüsyonel sinir ağı tabanlı modellerde kullanılmakta olup, görüntü sınıflandırma, obje tanıma ve tıbbi görüntü analizinde yaygın bir şekilde tercih edilmektedir (Selvaraju et al., 2016). Grad-CAM'in çalışma prensibi, modelin son konvolüsyon katmanının çıktılarına dayanmaktadır. Derin öğrenme modellerinde, son konvolüsyon katmanları genellikle görüntüdeki yüksek seviyeli uzamsal özellikleri temsil eder. Grad-CAM yöntemi, bu katmandaki özellik haritaları ile sınıf tahmini arasındaki gradyanları hesaplayarak sınıf tahminine en çok katkıda bulunan bölgeleri belirler (Selvaraju et al., 2017).

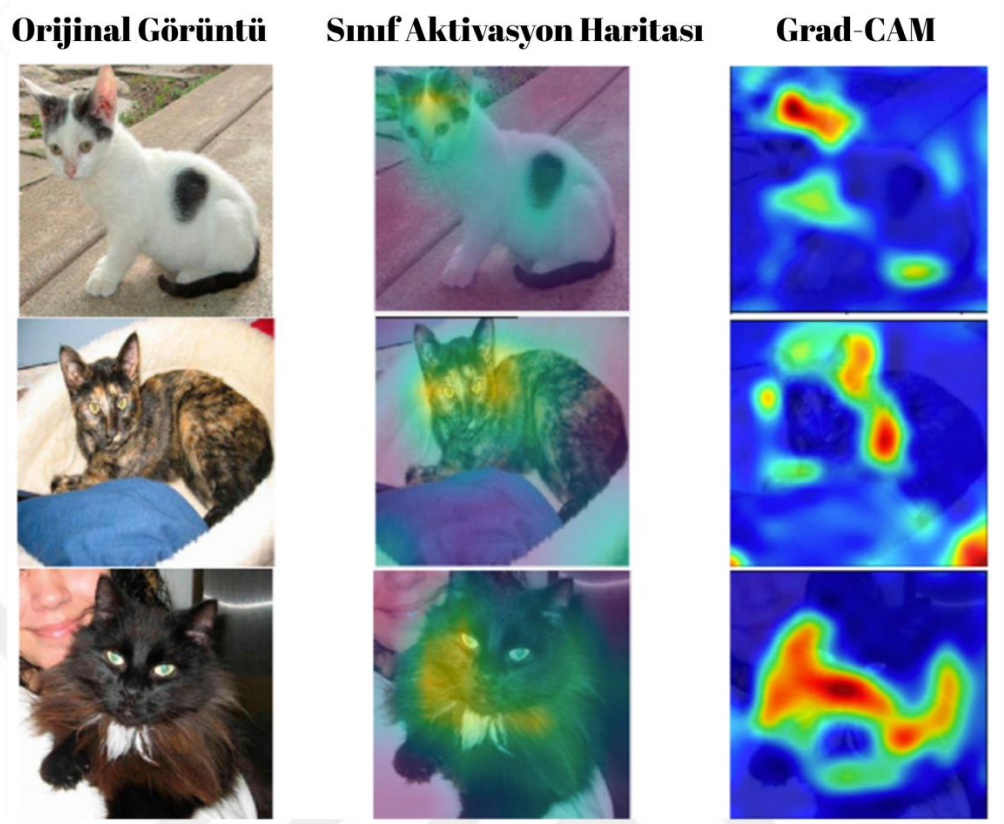
Grad-CAM yönteminin uygulanma süreci, Algoritma 1'de özetlenmiştir.

Algoritma 1: Grad-CAM algoritmasının temel adımları

Girdi: Görüntü G, Model M, Sınıf S
Çıktı: Sınıf Aktivasyon Haritası CAM

Son Konvolüsyon Katmanı:
$\{F_k\} \leftarrow M(G)$ modelinin son konvolüsyon katmanındaki K adet özellik haritası ▷ Her F_k, bir özellik haritasını temsil eder.
Gradyanların Hesaplanması:
$\frac{\partial y_S}{\partial F_k} \leftarrow S$ sınıfına ait tahmin değeri (y_S) ile F_k arasındaki gradyanlar ▷ Her özellik haritası için gradyanlar hesaplanır.
Ağırlıkların Hesaplanması:
for $k = 1$ **den** K **ye do**
    $w_k \leftarrow \frac{1}{Z} \sum_{i,j} \frac{\partial y_S}{\partial F_k(i,j)}$ ▷ Gradyanların ortalaması alınır (Global Ortalama Havuzlama).
end for
Sınıf Aktivasyon Haritasının Oluşturulması:
$CAM(i,j) \leftarrow \text{ReLU}\left(\sum_{k=1}^{K} w_k \cdot F_k(i,j)\right)$ ▷ Ağırlıklı özellik haritalarının kombinasyonu.
Görselleştirme:
$CAM_{\text{normalleştirilmiş}} \leftarrow \text{Normalleştir}(CAM)$
$Sonuç \leftarrow \text{Bindir}(CAM_{\text{normalleştirilmiş}}, G)$ ▷ Sınıf aktivasyon haritası, orijinal görüntü üzerine bindirilir.
Döndür: $Sonuç$

Grad-CAM yöntemi, özellikle modelin tahmininde dikkat ettiği bölgeleri görsel olarak ortaya koyduğu için tıbbi görüntü analizinde önemli bir araçtır. Şekil 11'de Grad-CAM kullanılarak elde edilen bir aktivasyon haritası örneği verilmiştir. Bu şekilde, bir deri lezyonu görüntüsü üzerinde modelin sınıflandırma yaparken odaklandığı bölgeler sıcak renk tonları (kırmızı) ile gösterilmiş, daha az önemli bölgeler ise soğuk renk tonları (mavi) ile ifade edilmiştir.



Şekil 11: Grad-CAM kullanılarak elde edilen aktivasyon haritası (Adadi & Berrada, 2018)

Grad-CAM'in en önemli avantajlarından biri, herhangi bir CNN tabanlı model ile uyumlu olmasıdır. Bu sayede, mevcut derin öğrenme modellerine herhangi bir yapısal değişiklik yapılmadan kolayca entegre edilebilir. Ayrıca Grad-CAM, sınıf tahminlerinin görsel açıklamalarıyla modelin güvenilirliğini artırır ve kullanıcıların modeli daha iyi yorumlamasına olanak tanır. Özetle, Grad-CAM yöntemi, modelin karar mekanizmasını anlamada etkili bir görsel açıklama sunarak, yapay zekâ sistemlerinin şeffaflığını artırmakta ve özellikle tıbbi teşhis gibi kritik alanlarda kullanıcı güvenini güçlendirmektedir (Zhang et al., 2021).

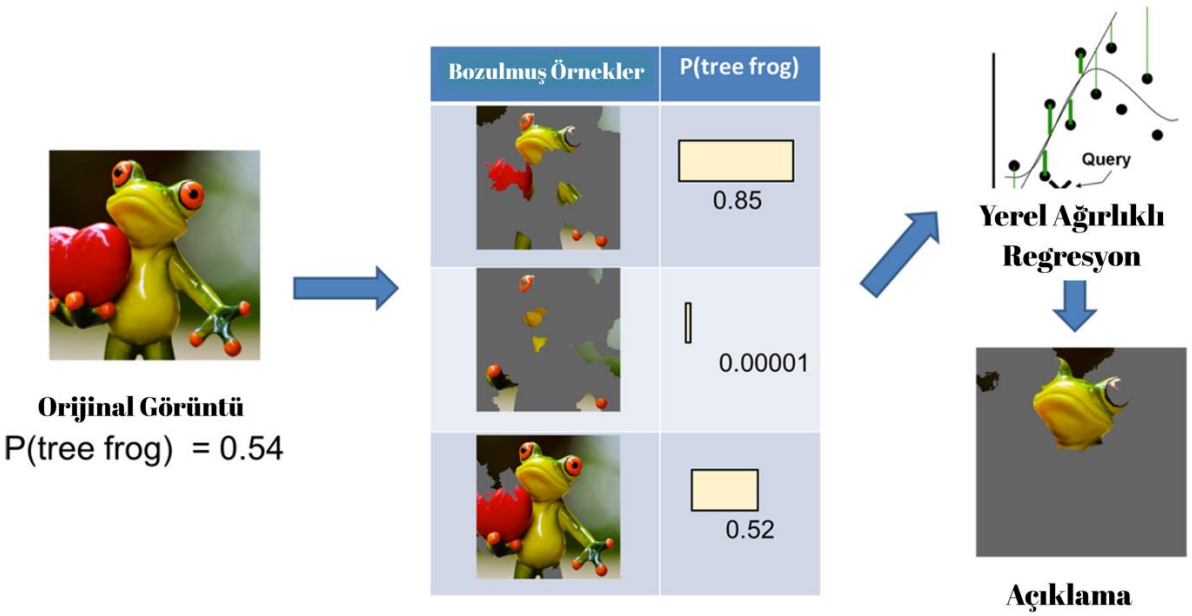
3.6.2. LIME

LIME, karmaşık yapay zekâ modellerinin kararlarını lokal düzeyde açıklayarak, modelin neden belirli bir tahmini yaptığına dair anlamlı bilgiler sunan bir yöntemdir. LIME, modeli doğrudan analiz etmez. Bunun yerine, modelin tahminlerini anlamak için girdilere yakın bir alanda yerel bir açıklayıcı model (genellikle basit bir lineer regresyon) oluşturur. Bu yerel model, karmaşık modelin çıktısını basitleştirir. Görüntü tabanlı uygulamalarda, LIME yöntemi, görüntüyü süperpiksel adı verilen küçük bölgelere

ayırarak her bir bölgenin tahmin üzerindeki katkısını analiz eder. Bu analiz sonucunda, modelin tahminini en çok etkileyen bölgeler belirlenir ve görsel olarak vurgulanır (Ribeiro et al., 2016). Örneğin, bir deri lezyonuna ait görüntüde modelin bu lezyonu "melanom" olarak sınıflandırdığını varsayalım. LIME bu görüntüyü süperpiksel adı verilen küçük parçalara böler ve ardından bu parçaların her birini ayrı ayrı modele sunarak karar üzerindeki etkilerini ölçer. Sonuç olarak, modelin kararını en çok etkileyen süperpiksel bölgeler (örneğin koyu renkli bir nokta, düzensiz kenar gibi) tespit edilerek görsel olarak vurgulanır. Böylece, modelin "melanom" kararını hangi görüntü parçacıklarına dayanarak verdiği kullanıcıya açıkça sunulur. Bu açıklayıcı harita hem modelin kararının anlaşılabilirliğini artırmakta hem de klinik uzmanların model çıktısını doğrulama sürecine katkı sağlamaktadır (Ribeiro et al., 2016; Saraswat et al., 2022).

LIME'ın temel amacı, modelin karmaşık karar verme mekanizmasını lokal düzeyde anlamlandırmak ve kullanıcıya açıklanabilir bir tahmin sunmaktır. Özellikle, görüntü tabanlı uygulamalarda hangi bölgelerin karar üzerinde etkili olduğu görsel olarak tanımlanabilir. Bu yöntem, modelin tahminlerini daha şeffaf hale getirerek, kullanıcının modelin kararlarına olan güvenini artırmaktadır. LIME'ın önemli bir avantajı, modelden bağımsız olarak çalışmasıdır. Yani, herhangi bir derin öğrenme modeli veya makine öğrenimi algoritması üzerinde uygulanabilir. Ayrıca, modelin genel yapısını değiştirmeden lokal düzeyde açıklamalar sağlayarak, kullanıcıların belirli bir tahmine yönelik çıkarımlar yapmasına olanak tanır. Bu durum, özellikle tıbbi görüntüleme gibi uygulamalarda önemli bölgelerin doğruluğunu değerlendirme açısından kritik bir rol oynar (Bhandari et al., 2023).

Şekil 12'de LIME yöntemi ile oluşturulmuş bir açıklama süreci örneği verilmiştir. Solda, modelin %54 olasılıkla "tree frog" olarak sınıflandırdığı orijinal görüntü yer almaktadır. Ortada, görüntü üzerinde farklı bölgelerin maskelendiği varyasyonlar ve bunlara ait tahmin sonuçları gösterilmiştir. Sağda ise modelin kararında etkili olan bölgeler öne çıkarılmış; kurbağanın yüzü ve gövdesi gibi ayırt edici alanların sınıflandırmada belirleyici olduğu görülmektedir.



Şekil 12: LIME kullanılarak model tahmininin yorumlanması (Ribeiro, 2016)

LIME yönteminin uygulanma süreci, Algoritma 2'de özetlenmiştir.

Algoritma 2: LIME algoritmasının temel adımları

Girdi: Görüntü x, sınıflandırıcı model f, hedef sınıf c, örnek sayısı N
Çıktı: Görsel açıklama maskesi A

Süper Piksel Bölütleme:
$S \leftarrow \text{segment}(x)$ ▷ Girdi görüntüsü süper piksellere ayrılır.
Bozulmuş Örneklerin Oluşturulması:
$Z \leftarrow \emptyset$ ▷ Örnek-tahmin çiftlerinin saklanacağı küme
for $i = 1$ **den** N **ye kadar do**
    $z_i \leftarrow$ Rastgele süper piksel kombinasyonu
    $\tilde{x}_i \leftarrow \text{boz}(x, z_i, S)$ ▷ Belirli bölgeler kapatılarak bozulmuş görüntü elde edilir.
    $y_i \leftarrow f(\tilde{x}_i)$
    $Z \leftarrow Z \cup \{(z_i, y_i)\}$
end for
Ağırlık Hesaplama:
$\pi_x(\tilde{x}_i) \leftarrow$ Her örneğin x görüntüsüne yakınlığına göre ağırlığını hesapla
Yerel Model Eğitimi:
$g \leftarrow Z$ ve π_x kullanılarak sınıf c için ağırlıklı doğrusal model öğren
Açıklamanın Oluşturulması:
$A \leftarrow$ Yerel modelde en büyük katsayıya sahip süper pikseller
return A ▷ Açıklama maskesi döndürülür

3.7.Değerlendirme Metrikleri

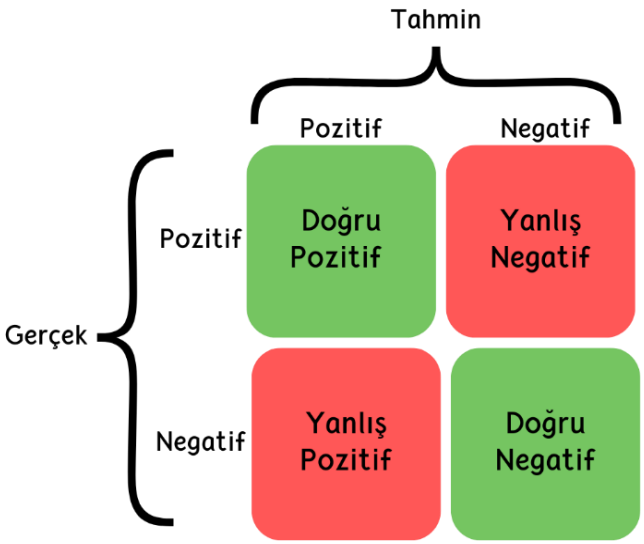
Bu çalışmada, sınıflandırma modellerinin performansını değerlendirmek için çeşitli performans ölçütleri kullanılmıştır. Kullanılan metrikler, modellerin genel doğruluğunu, sınıf tahminlerindeki tutarlılığı ve karar mekanizmalarının güvenilirliğini farklı açılardan analiz etmektedir. Her bir metrik, sınıflandırma performansını detaylı bir şekilde değerlendirme imkânı sunarak modelin başarısını kapsamlı bir şekilde ortaya koymaktadır (Sahoo et al., 2020).

3.7.1. Karmaşıklık Matrisi

Modellerin performansını kapsamlı bir şekilde analiz edebilmek için Karmaşıklık Matrisi (Confusion Matrix) kullanılmıştır. Karmaşıklık matrisi, sınıflandırma problemlerinde modelin tahmin performansını detaylı bir şekilde özetleyen bir araçtır. Modelin her bir sınıf için yaptığı doğru ve hatalı tahminlerin dağılımını görselleştirerek değerlendirme metriklerinin hesaplanmasına temel oluşturur (Townsend, 1971). Karmaşıklık matrisi, sınıflandırma problemlerinde dört temel bileşen üzerine kuruludur:

- True Positive (TP): Modelin, pozitif olarak tahmin ettiği ve doğru olan örneklerdir.
- False Positive (FP): Modelin pozitif olarak tahmin ettiği ancak yanlış olan örneklerdir (Tip I Hatası).
- True Negative (TN): Modelin negatif olarak tahmin ettiği ve doğru olan örneklerdir.
- False Negative (FN): Modelin negatif olarak tahmin ettiği ancak yanlış olan örneklerdir (Tip II Hatası).

Bu dört bileşen, modelin tahmin performansını sınıf bazında inceleyerek hangi sınıflarda hataların yoğunlaştığını anlamamıza yardımcı olur (Ting, 2010). Şekil 13'te bir karmaşıklık matrisinin örneği sunulmuştur. Şekil 13’te, matrisin satırları gerçek sınıfları (Actual), sütunları ise model tarafından tahmin edilen sınıfları (Predicted) temsil etmektedir. Diyagonal boyunca yer alan hücreler (TP ve TN) doğru tahmin edilen örnekleri, diyagonal dışındaki hücreler ise yanlış tahminleri göstermektedir.



Şekil 13: Karmaşıklık matrisi

Karmaşıklık matrisinin temel avantajı, modelin doğruluğunu sadece genel olarak değil, sınıf bazında da değerlendirmeye olanak sağlamasıdır. Örneğin, özellikle dengesiz veri setlerinde, nadir görülen sınıflardaki hataların analiz edilmesi ve modelin hangi sınıflarda zorlandığının belirlenmesi kritik öneme sahiptir. Karmaşıklık matrisinden elde edilen değerler, aşağıdaki performans metriklerinin hesaplanmasında kullanılır:

- Hassasiyet (Precision): FP hatalarını göz önünde bulundurarak modelin doğru pozitif tahmin yapma oranını ölçer.
- Duyarlılık (Recall/Sensitivity): FN hatalarını değerlendirerek modelin pozitif sınıfları ne kadar iyi tahmin ettiğini gösterir.
- F1-Skoru: Precision ve Recall'in harmonik ortalaması ile modelin dengeli performansını sunar.

Bu matris, modelin sınıf bazında performansını inceleyerek özellikle hassasiyetin önemli olduğu tıbbi tanı uygulamalarında büyük önem taşır. Yanlış negatif tahminlerin (FN) belirlenmesi, kritik durumların atlanmaması açısından hayati öneme sahiptir (Beauxis-Aussalet & Hardman, 2014).

3.7.2. Performans Ölçütleri

Doğruluk (Accuracy): Doğruluk, modelin doğru tahmin ettiği örneklerin toplam örnek sayısına oranını ifade eder. Matematiksel olarak Eşitlik 1'de tanımlanmıştır (Koklu & Ozkan, 2020).

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

Doğruluk, genel performansı ölçmek için yaygın olarak kullanılan bir metrik olmasına rağmen, dengesiz veri setlerinde yanıltıcı olabilir. Özellikle nadir görülen sınıfların hatalarını gizleyebilir.

Hassasiyet (Precision): Hassasiyet, bir sınıfa ait pozitif tahminlerin ne kadarının doğru olduğunu ölçer. Özellikle yanlış pozitif tahminlerin önemli olduğu durumlarda kullanılır. Eşitlik 2'de nasıl hesaplandığı gösterilmiştir.

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

Burada TP (True Positive) doğru pozitifleri, FP (False Positive) yanlış pozitifleri temsil eder. Hassasiyet, modelin belirli bir sınıfı yanlış tahmin etme oranını düşürmekte ne kadar başarılı olduğunu gösterir (Koklu et al., 2021).

Duyarlılık (Recall): Duyarlılık, modelin pozitif örnekleri ne kadar doğru bir şekilde tahmin ettiğini ölçer. Özellikle yanlış negatiflerin önemli olduğu durumlarda kullanılır. Matematiksel olarak Eşitlik 3'te gösterilmiştir.

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

Burada FN (False Negative) yanlış negatifleri ifade eder. Duyarlılık, modelin pozitif örnekleri kaçırma oranını ölçer ve tıbbi tanı gibi kritik uygulamalarda oldukça önemlidir (Koklu et al., 2021).

F1-Skoru: F1-Skoru, hassasiyet ve duyarlılığın harmonik ortalaması olarak tanımlanır. Modelin dengesiz veri setlerinde performansını ölçmek için kullanılır. F1-Skoru, Eşitlik 4'te gösterildiği gibi hesaplanmaktadır.

$$\text{F1-Skoru} = 2 \times \frac{\text{Hassasiyet} \times \text{Duyarlılık}}{\text{Hassasiyet} + \text{Duyarlılık}} \quad (4)$$

F1-Skoru, hem yanlış pozitif hem de yanlış negatif tahminlere karşı duyarlı bir ölçüttür (Koklu & Ozkan, 2020).

Mean Squared Error (MSE): MSE, modelin tahmin hatalarının ortalamasının karesi olarak hesaplanır ve Eşitlik 5'teki gibi ifade edilir:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2 \quad (5)$$

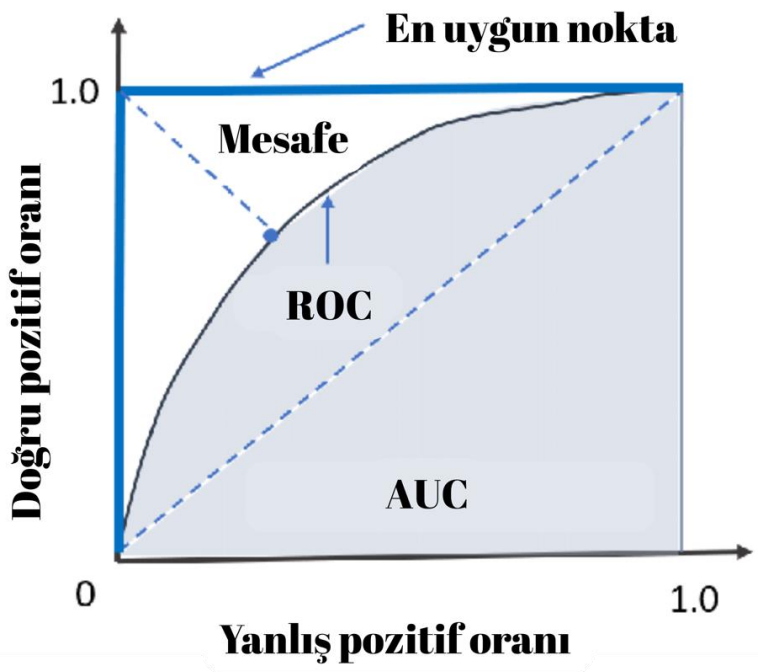
Burada y_i gerçek değerleri, x_i tahmin edilen değerleri ve n örnek sayısını temsil eder. MSE, tahmin hatalarının büyüklüğünü hesaplayarak modelin hatalarını sayısal olarak ölçer (Taspinar et al., 2022).

Bias: Bias (Önyargı), bir modelin gerçek değerlerden sistematik olarak sapma derecesini ölçen bir hata türüdür. Yüksek bias, modelin eğitim verisindeki temel desenleri yeterince öğrenemediğini ve bu nedenle tahminlerinde sürekli bir hata ürettiğini gösterir. Genellikle basit veya aşırı kısıtlanmış modellerde görülür ve underfitting olarak adlandırılan duruma yol açar. Underfitting, modelin hem eğitim hem de test verisi üzerinde düşük performans göstermesine neden olur, çünkü model karmaşık ilişkileri ve altta yatan veri yapısını yeterince yakalayamaz (Bullinaria, 2015).

Bias, makine öğrenimi modellerinin genelleme kapasitesini etkileyen en önemli faktörlerden biridir ve bias-varyans dengesi bağlamında değerlendirilmelidir. Bir modelin performansını optimize etmek için bias ve varyans arasında bir denge kurmak gerekir (Barzegar et al., 2020). Düşük bias genellikle daha esnek ve karmaşık modellerle sağlanabilir, ancak bu durum yüksek varyansa yol açarak modelin overfitting göstermesine neden olabilir. Overfitting durumunda, model eğitim verisine çok iyi uyum sağlarken yeni ve bilinmeyen verilere karşı düşük genelleme performansı gösterir. Dolayısıyla, bias seviyesinin dikkatlice analiz edilmesi, modelin hem eğitim hem de test verilerinde başarılı olması için kritik bir adımdır (Aliferis & Simon, 2024).

ROC/AUC: ROC eğrisi, sınıflandırıcı modelin tahmin ettiği olasılıklar için farklı karar eşikleri belirlenerek, her bir eşik noktasında recall ve false positive rate değerlerinin hesaplanmasıyla oluşturulan bir grafiksel gösterimdir. ROC eğrisinin yatay ekseni yanlış pozitif oranı (FPR), dikey ekseni ise doğru pozitif oranı (TPR, recall) değerlerini temsil eder. Farklı eşik değerlerinde modelin performansını gözlemleyerek ROC eğrisi

oluşturulur (Majnik & Bosnić, 2013). Eğrinin sol üst köşeye daha yakın olması, modelin daha iyi bir performans gösterdiğine işaret eder. Özellikle eğrinin yatay eksene yakın bir şekilde ilerlemesi modelin yüksek yanlış pozitif oranına sahip olduğunu, dik eksene yakın olması ise modelin daha iyi ayrım gücüne sahip olduğunu gösterir (Flach, 2016). ROC eğrisinin altında kalan alan (AUC) değeri, modelin genel sınıflandırma performansını tek bir skora indirger. AUC değeri 0.5 seviyesinde ise modelin rastgele tahmin yaptığı anlamına gelirken, 1.0 değerine yaklaşması modelin mükemmel bir sınıflandırıcı olduğunu gösterir. Genel olarak, AUC değeri ne kadar yüksekse, modelin pozitif ve negatif sınıfları ayırt etme gücü de o kadar yüksektir (Huang & Ling, 2005). Şekil 14'te ROC eğrisi ve sınıflayıcı performansını değerlendirmede kullanılan AUC metriği gösterilmiştir.



Şekil 14: ROC eğrisi ve eşik belirleyici. Eğri altındaki alan (AUC), sınıflayıcının başarımını gösterir; 1'e yakın AUC değeri, en iyi performansı ifade eder (Liu et al., 2020).

4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA

Bu bölümde, yapılan deneysel çalışmanın amacı, kullanılan yöntemlerin seçilme nedenleri ve araştırma sorularına nasıl yanıt arandığı açıklanmaktadır. Çalışmanın temel hedefi, Ham10000 veri seti kullanarak deri lezyonlarının sınıflandırılmasını gerçekleştirmek ve bu sınıflandırmada farklı derin öğrenme modellerinin performanslarını karşılaştırmaktır. Bunun yanı sıra, sınıflandırma modellerinin karar mekanizmalarını görselleştirerek model çıktılarının güvenilirliğini ve açıklanabilirliğini artırmak da çalışmanın önemli bir amacıdır.

Çalışmanın genel akışı, Algoritma 3'te özetlenmiştir.

Algoritma 3: Yöntemin genel akışını gösteren Pseudo kod

Girdi: HAM10000 veri seti V
Çıktı: Eğitilmiş modeller ve açıklanabilirlik görselleştirmeleri

Veri Önişleme:
$V \leftarrow$ RGB yeniden boyutlandırma ve normalizasyon
$V \leftarrow$ Gürültü azaltma ve sınıf etiketleme

Veri Dengeleme:
$V_{deng} \leftarrow$ Undersampling uygulanır

Model Eğitimi:
for $model \in \{MobileNet, DenseNet121, Xception\}$ **do**
    $model \leftarrow$ EğitimVerisi(V_{deng}) ile eğitilir
    $performans \leftarrow$ ModelDeğerlendir($model$)
    $modeller.append(model, performans)$
end for

Açıklanabilir Yapay Zekâ Analizi:
$gradcam_haritaları \leftarrow$ GradCAM(en_iyi, test_görüntüleri)
$lime_maskeleri \leftarrow$ LIME(en_iyi, test_görüntüleri)

Sonuçların Raporlanması:
Grad-CAM ve LIME çıktıları görselleştirilir ve yorumlanır

Çalışmada Mobilenet, Xception ve DenseNet modelleri kullanılmıştır. Mobilenet, hafif ve hızlı yapısı sayesinde özellikle mobil cihazlar gibi sınırlı donanımlara sahip sistemlerde bile etkin performans göstermesiyle öne çıkmaktadır. Xception ise derin öğrenme alanında yaygın olarak kullanılan, yüksek kapasiteli bir model olup büyük veri setlerinde elde ettiği başarılı sonuçlarla dikkat çekmektedir. DenseNet, yoğun bağlantıları sayesinde model öğrenmesini optimize ederek sınıflandırma performansını artıran bir yapıya sahiptir. Bu üç modelin farklı özellikleri, çalışmanın hipotezini test etmede ve elde

edilen sonuçların çeşitliliğini sağlamada önemli bir rol oynamaktadır. Tablo 3'te modellerin temel özellikleri gösterilmiştir. Modellerin performans analizlerinin yanı sıra, karar mekanizmalarının nasıl çalıştığını görselleştirmek ve anlaşılır hale getirmek için Grad-CAM ve LIME yöntemleri uygulanmıştır. Grad-CAM yöntemi, modellerin sınıflandırma yaparken hangi bölgeler üzerinde yoğunlaştığını görsel olarak açıklarken, LIME yöntemi modelin kararlarını yerel düzeyde analiz ederek açıklanabilirlik sağlar.

Tablo 3: Kullanılan modellerin temel özellikleri

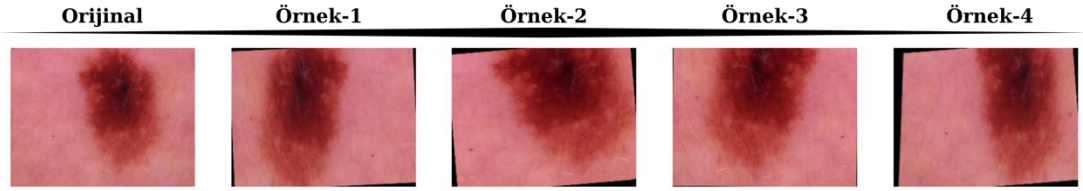
Model	Yayın Yılı	Parametre Sayısı	Katman Sayısı	Model Boyutu
MobileNet	2017	4.2M	28	16 MB
Xception	2017	22.8M	71	88 MB
DenseNet	2017	8M	121	32 MB

Bu deneysel çalışmada, farklı derin öğrenme modellerinin performansları karşılaştırılarak hangi modelin deri lezyonu sınıflandırmasında daha başarılı olduğu incelenmiş ve modellerin açıklanabilirliği artırılarak karar mekanizmalarının güvenilirliği değerlendirilmiştir. Elde edilen sonuçlar, derin öğrenme modellerinin klinik uygulamalarda kullanılabilirliğini ve kullanıcılar tarafından daha kolay yorumlanabilir hale gelmesini sağlamak için önemli bir katkı sunmaktadır.

4.1. Veri Önişleme

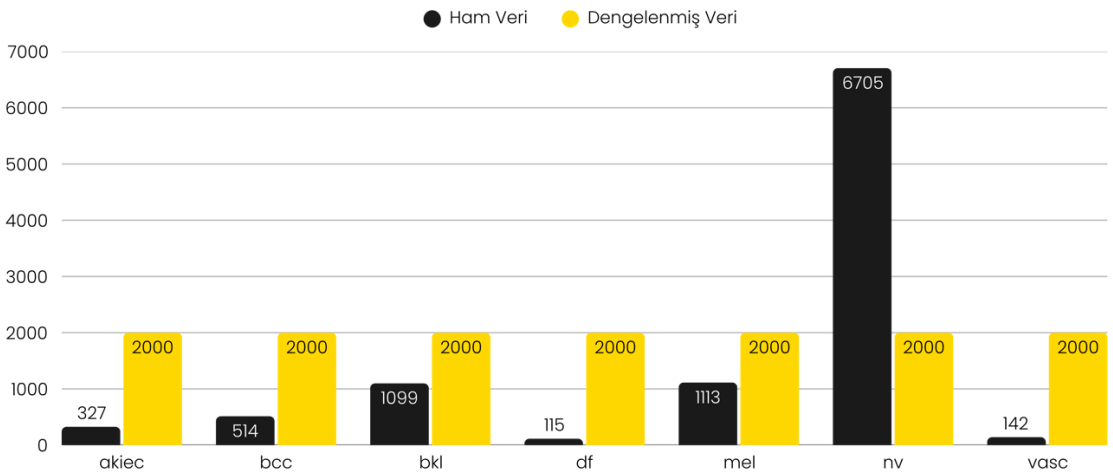
HAM10000 veri setinde sınıflar arasında belirgin bir dengesizlik gözlenmiştir. Örneğin, "benign nevüs" sınıfında 6.705 adet görüntü bulunurken, bazı sınıflardaki görüntü sayısı 200'ün altındadır. Bu dengesizlik, modelin eğitimi sırasında bazı sınıfların yeterince temsil edilememesine ve dolayısıyla model performansının düşmesine neden olabileceğinden, veri dengeleme işlemi uygulanmıştır. Her sınıf için hedeflenen görüntü sayısı 2000 olarak belirlenmiştir. Görüntü sayısı 2000'in üzerinde olan sınıflar rastgele örnekleme yöntemiyle 2000'e düşürülmüş, görüntü sayısı yetersiz olan sınıflarda ise veri artırma (augmentation) teknikleri uygulanarak yeni görüntüler oluşturulmuştur. Veri artırma işlemi sırasında kullanılan teknikler arasında yatay çevirme, rastgele döndürme (-10° ile 10°), ölçekleme (%90 ile %110), kırpma, kesme, hafif Gaussian bulanıklaştırma,

gürültü ekleme ve kontrast ayarlama yer almıştır. Ayrıca, tüm görüntüler sabit boyutta (224 x 224) yeniden boyutlandırılmıştır.



Şekil 15: Veri arttırma sonrası örnek görüntüler

Şekil 15'te veri artırma işleminin örnekleri sunulmuştur. Bu şekilde, bir orijinal görüntü ile veri artırma sonucunda oluşturulmuş çeşitli varyasyonlar yer almaktadır. Görüntülerde dönüşümler sonucunda yatay çevirme, hafif döndürme ve kontrast değişimi gibi farklılıklar gözlemlenebilir. Bu dönüşümler, modelin veriyi daha çeşitli açılardan öğrenmesini sağlayarak genelleme yeteneğini artırmayı amaçlamaktadır. Veri dengeleme işleminin ardından, her sınıf 2000 adet görüntü içerecek şekilde eşitlenmiştir. Bu durum, sınıflar arasında denge sağlanarak modelin tüm sınıfları eşit ağırlıkta öğrenmesine olanak tanımıştır. Şekil 16'da ham veri sayılarını ve veri dengeleme sonrası sınıf dağılımını gösteren sütun grafiği sunulmuştur. Grafik, her sınıfta görüntü sayısının eşitlendiğini ve dengesizliğin giderildiğini açıkça göstermektedir.



Şekil 16: Veri arttırma sonrası sınıf dağılımı

Elde edilen veri seti, modelin eğitimi, doğrulaması ve test edilmesi için %70, %15 ve %15 oranında üç alt kümeye ayrılmıştır. Buna göre, her sınıftan 1400 görüntü eğitim setinde, 300 görüntü doğrulama setinde ve 300 görüntü test setinde kullanılmıştır. Bu

işlem sonucunda oluşturulan eğitim, doğrulama ve test veri kümelerinin dengeli yapısı, modelin performans değerlendirmesinde tutarlılık sağlamıştır.

4.1.Model Eğitimi

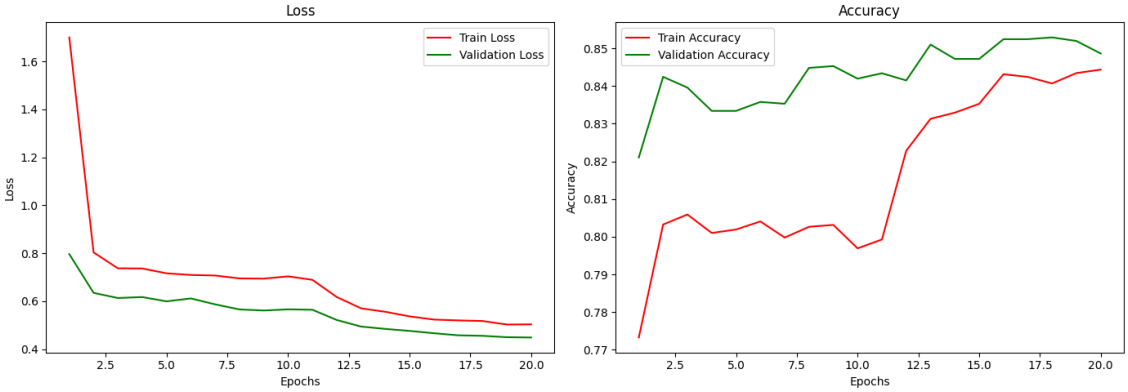
Model eğitiminde Intel i5 işlemci, 16 GB RAM ve NVIDIA RTX 2060 ekran kartına sahip bir bilgisayar kullanılmıştır. Eğitim süreci boyunca donanımın sağladığı GPU desteği, modellerin verimli ve hızlı bir şekilde eğitilmesine olanak sağlamıştır. Çalışmada üç farklı derin öğrenme modeli olan Mobilenet, Xception ve DenseNet eğitilmiştir. Bu süreçte her model, 20 epoch boyunca Adam optimizasyon algoritması ve 16 batch size ile çalıştırılmıştır. Öğrenme oranı başlangıçta 0.001 olarak belirlenmiş ve eğitim süresince ReduceLROnPlateau callback fonksiyonu kullanılarak, doğrulama kaybında iyileşme gözlenmediği durumlarda öğrenme oranı 0.1 oranında azaltılmıştır. Eğitim sürecinde EarlyStopping yöntemi uygulanarak doğrulama kaybının 5 epoch boyunca iyileşmemesi durumunda eğitim sonlandırılmış ve en iyi ağırlıklar modele geri yüklenmiştir. Ayrıca model ağırlıklarının kaydedilmesi için ModelCheckpoint callback'i kullanılmıştır. Tablo 4'te tüm modellerin eğitim süreleri gösterilmiştir.

Tablo 4: Modellerinin doğrulama sonuçları ve toplam eğitim süresi karşılaştırması

Model	Epoch	Eğitim Süresi (dk:sn)	Doğrulama Doğruluğu	Doğrulama Kaybı
MobileNet	20	23:35	0.8529	0.4489
DenseNet	20	50:13	0.8472	0.4820
Xception	20	62:45	0.8272	0.5730

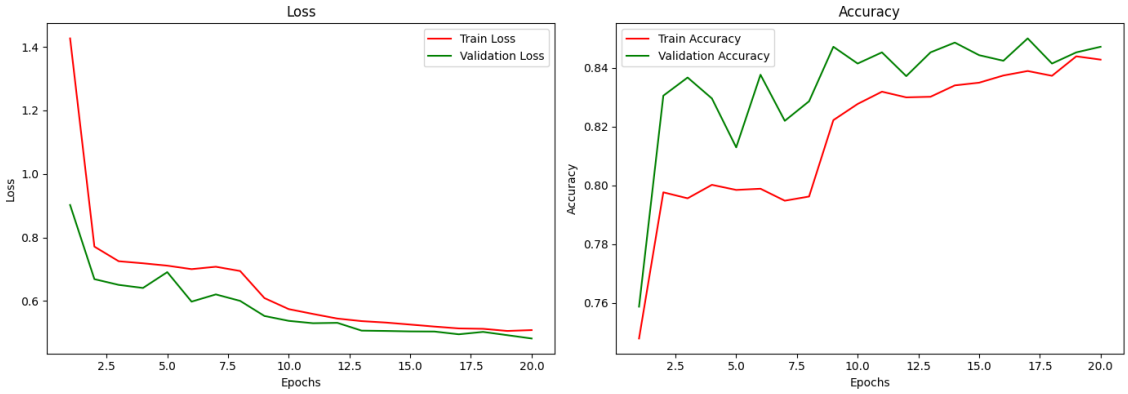
Mobilenet modeli eğitim süresince, doğrulama kaybında belirgin bir düşüş göstermiştir. İlk epoch'ta doğrulama kaybı 0.7966 seviyesinde başlarken, eğitim sürecinin sonlarına doğru bu değer 0.4489'a kadar gerilemiştir. Özellikle öğrenme oranının azaltılmasıyla kayıplardaki iyileşme hız kazanmıştır. Modelin doğrulama doğruluğu ise başlangıçta %82.1 seviyesinden başlayarak eğitim ilerledikçe %85.29 seviyesine ulaşmıştır. DenseNet modelinde ise daha uzun eğitim sürelerine rağmen, doğrulama kaybında benzer bir iyileşme gözlenmiştir. İlk epoch sonunda doğrulama kaybı 0.9024 seviyesinde iken, son epoch'ta bu değer 0.4820'ye kadar düşmüştür. Modelin doğrulama

doğruluğu ise eğitim sürecinde %75.87 seviyesinden %84.72 seviyesine kadar artış göstermiştir. Xception modeli ise doğrulama kaybında istikrarlı bir düşüş sergilemiş, ilk epoch'ta 0.9628 olan doğrulama kaybı, 15. epoch sonunda 0.5917 seviyesine kadar inmiştir. Modelin doğrulama doğruluğu ise eğitim boyunca %78.53 seviyesinden başlayarak %82.20 civarına yükselmiştir.

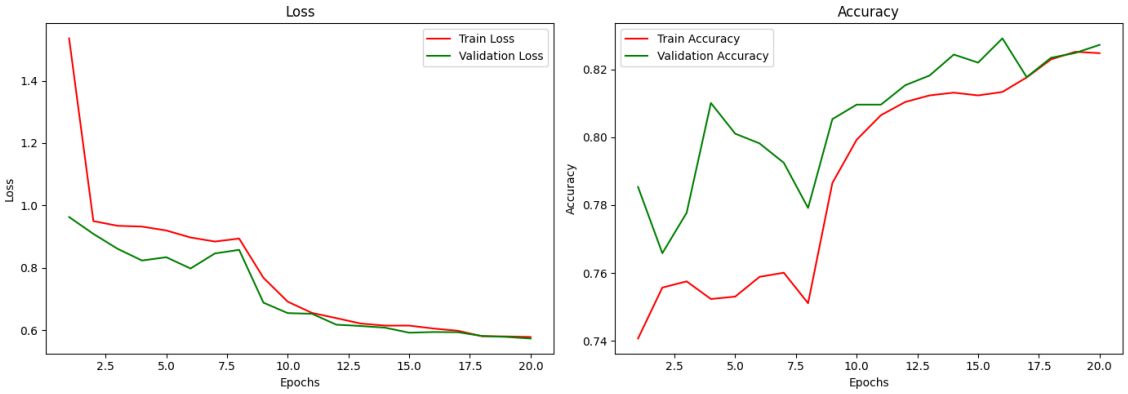


Şekil 17: Mobilenet modeline ait eğitim ve doğrulama kayıplarının grafiği

Şekil 17'de Mobilenet modeline ait eğitim ve doğrulama kayıp grafiği ile doğruluk grafiği sunulmuştur. Eğitim boyunca doğrulama kaybındaki düşüşün belirli noktalarda duraksadığı ancak öğrenme oranının azaltılması ile düşüşün devam ettiği gözlenmektedir. Şekil 18 ve Şekil 19'da sırasıyla DenseNet ve Xception modellerinin eğitim ve doğrulama kayıp grafiği ile doğruluk grafiği gösterilmiştir. DenseNet ve Xception modellerinde doğrulama kaybında benzer düşüş eğilimleri görülse de eğitim sürelerinin Mobilenet modeline kıyasla daha uzun olduğu dikkat çekmektedir. Elde edilen sonuçlar, Mobilenet modelinin hızlı ve hafif yapısıyla eğitim sürecinde avantaj sağladığını, DenseNet ve Xception modellerinin ise daha derin yapıları sayesinde doğrulama performansını iyileştirdiğini göstermektedir. Eğitim süresince kayıplardaki düşüşlerin ve doğrulama doğruluklarındaki artışların tutarlı bir şekilde ilerlemesi, modellerin veriye uyum sağladığını ve aşırı öğrenme durumunun önlendiğini ortaya koymaktadır.



Şekil 18: DenseNet modeline ait eğitim ve doğrulama kayıplarının grafiği



Şekil 19: Xception modeline ait eğitim ve doğrulama kayıplarının grafiği

4.2. Deney Sonuçları ve Performans Karşılaştırması

Deney sonuçlarında Mobilenet, DenseNet ve Xception modellerinin Ham10000 veri seti üzerinde elde ettikleri performans metrikleri incelenmiştir. Modellerin doğruluk (accuracy), hassasiyet (precision), duyarlılık (sensitivity/recall), F1-Skoru ve AUC değerleri Tablo 5'te sunulmuştur. Ayrıca, modellerin karar verme performansı karmaşıklık matrisi (confusion matrix) ve ROC eğrisi görselleriyle detaylandırılmıştır.

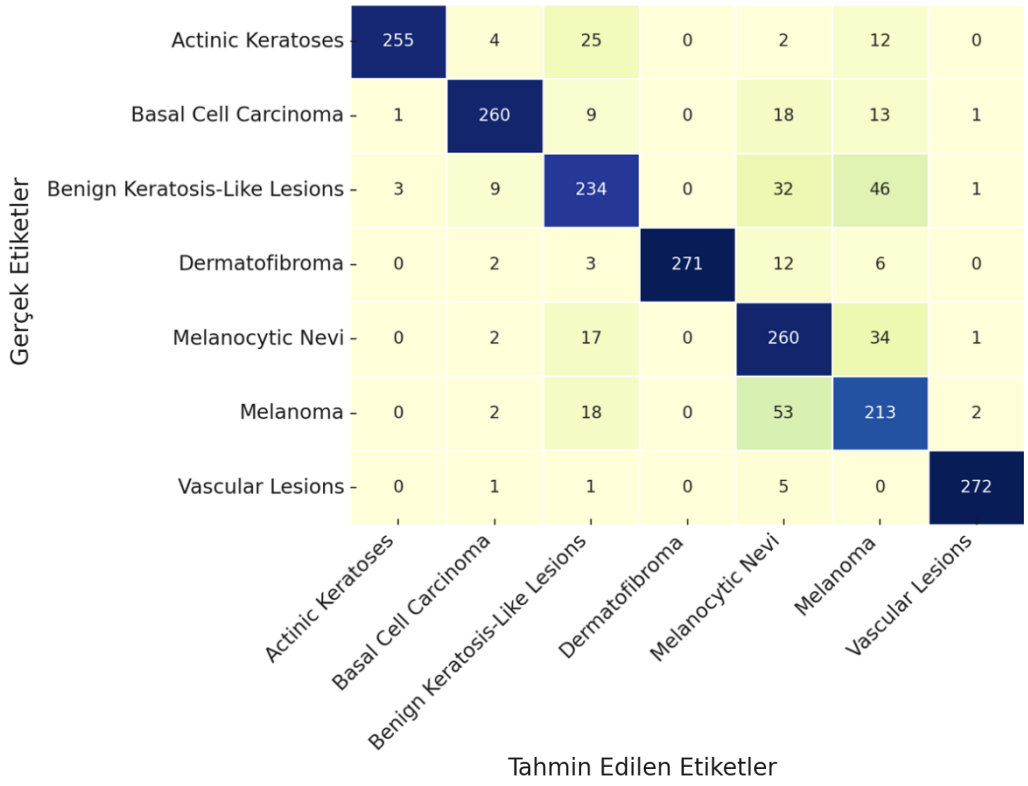
Tablo 5: Modellerin değerlendirme metrikleri

Model	Accuracy	Hassasiyet	Duyarlılık	F1-Score	AUC
MobileNet	0.8404	0.8564	0.8429	0.8472	0.9823
DenseNet	0.8357	0.8486	0.8387	0.8411	0.9799
Xception	0.8152	0.8416	0.8172	0.8236	0.9759

Mobilenet modeli, en yüksek accuracy ve F1-Score değerlerine sahip olup, %84.04 doğruluk oranı ile diğer modelleri geride bırakmıştır. DenseNet modeli, %83.57 doğruluk değeriyle Mobilenet'e yakın bir performans göstermiştir. Parametre sayısının daha fazla olması nedeniyle DenseNet'in eğitim süresi daha uzun sürmesine rağmen, doğruluk açısından tatmin edici sonuçlar elde edilmiştir. Xception modeli ise %81.52 doğruluk oranı ile diğer modellere kıyasla daha düşük bir performans sergilemiştir.8

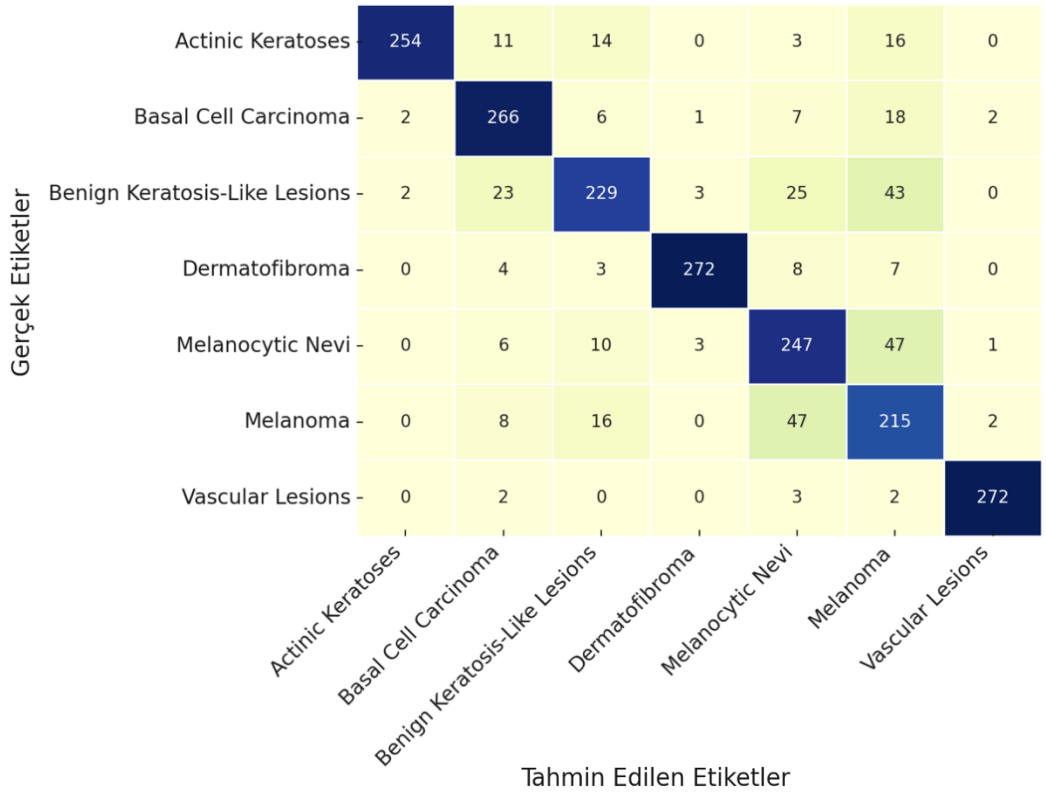
Elde edilen sonuçlar doğrultusunda Mobilenet modeli, sınırlı parametre sayısına sahip olmasına rağmen yüksek doğruluk ve AUC değerleriyle öne çıkmıştır. Modelin daha hafif ve hızlı yapısı, eğitim süresini önemli ölçüde kısaltmış ve kaynak kullanımını optimize etmiştir. DenseNet modeli, parametre sayısının artışına rağmen benzer performans sergileyerek özellikle büyük veri setlerinde derin yapıların faydasını göstermiştir. Xception modeli, daha derin katman yapısına sahip olmasına rağmen diğer modellerle kıyaslandığında doğruluk açısından geride kalmıştır. Bu durum, modelin aşırı öğrenme eğilimi veya veri setindeki dengesizliklerden etkilenmiş olabileceğini düşündürmektedir. Sonuç olarak, Mobilenet modeli, özellikle hız, kaynak kullanımı ve performans açısından çalışmanın hedeflerine en uygun yapı olarak değerlendirilmektedir. DenseNet modeli ise parametre sayısına bağlı olarak derin öğrenmenin faydalarını sergilemiştir.

HAM10000 veri seti üzerinde eğitilen ve test edilen MobileNet, DenseNet ve Xception modellerinin performansları karmaşıklık matrisleri üzerinden incelenmiştir. Karmaşıklık matrisleri, modellerin sınıf bazındaki tahmin doğruluklarını ve hata dağılımlarını göstererek, performans karşılaştırmasını bilimsel açıdan değerlendirilmesini sağlamaktadır.



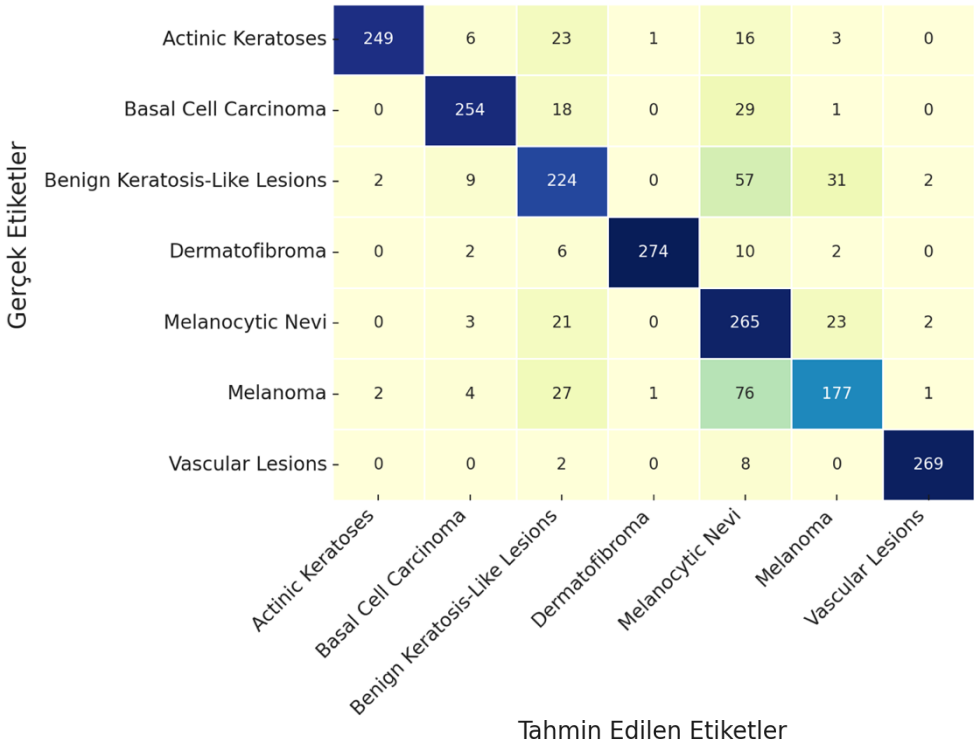
Şekil 20: MobileNet modeli için karmaşıklık matrisi

Şekil 20'de görüldüğü gibi MobileNet modeli, Dermatofibroma ve Vascular Lesions sınıflarında yüksek doğruluk oranına ulaşmıştır. Ancak, Benign Keratosis-Like Lesions ve Melanoma sınıflarında belirgin sınıflandırma hataları gözlenmiştir. Özellikle Benign Keratosis-Like Lesions sınıfı, sıklıkla Melanocytic Nevi olarak yanlış tahmin edilmiştir. Bu hata, görsel benzerliklerin MobileNet modeli tarafından ayırt edilemediğini ve daha gelişmiş özellik çıkarımı yöntemlerine ihtiyaç duyulduğunu göstermektedir. Benzer şekilde, Melanoma örneklerinin önemli bir kısmı Melanocytic Nevi sınıfına yanlış atandığı için, bu sınıfların arasındaki sınırların model tarafından yeterince öğrenilemediği anlaşılmaktadır.



Şekil 21: DenseNet modeli için karmaşıklık matrisi

Şekil 21'de gösterilen DenseNet modeli, genel sınıflandırma performansı açısından tutarlı sonuçlar üretmiştir. Özellikle Dermatofibroma ve Vascular Lesions sınıflarında yüksek doğruluk elde edilmiştir. Ancak Benign Keratosis-Like Lesions ve Melanocytic Nevi sınıfları arasında bir hata dağılımı dikkat çekmektedir. DenseNet, Melanoma sınıfında diğer modellere kıyasla daha iyi bir performans sergilese de, bu sınıfta hâlâ önemli oranda yanlış sınıflandırma meydana gelmiştir. Bu durum, DenseNet'in derin katmanlarının daha karmaşık özellikleri öğrenebildiğini ancak belirli sınıfların görsel benzerliklerini tamamen ortadan kaldıramadığını ortaya koymaktadır.



Şekil 22: Xception modeli için karmaşıklık matrisi

Şekil 22'de yer alan Xception modeli ise, bazı sınıflarda belirgin hata oranları göstermiştir. Xception, Dermatofibroma sınıfında en iyi performansı elde etmiştir; ancak Melanoma ve Benign Keratosis-Like Lesions sınıflarındaki hatalar dikkat çekicidir. Özellikle Melanoma sınıfındaki örneklerin büyük bir kısmının Melanocytic Nevi olarak tahmin edilmesi, bu iki sınıf arasındaki görsel benzerliğin Xception tarafından yeterince ayırt edilemediğini göstermektedir. Ayrıca Benign Keratosis-Like Lesions örneklerinin önemli bir kısmı yine Melanocytic Nevi sınıfına yanlış atanmıştır. Bu sonuçlar, Xception'in karmaşık sınıflar arasındaki ince ayrımları öğrenme kapasitesinin sınırlı kaldığını göstermektedir. Genel olarak, üç modelin de Dermatofibroma ve Vascular Lesions gibi ayırt edici özelliklere sahip sınıflarda tutarlı ve başarılı sonuçlar ürettiği gözlenmiştir. Ancak Melanocytic Nevi, Melanoma ve Benign Keratosis-Like Lesions sınıflarında yaşanan karışıklıklar, modellerin bu sınıfların ortak görsel özelliklerini ayırt etmekte zorlandığını göstermektedir. Tablo 6'da MobileNet, DenseNet ve Xception modellerinin en başarılı ve en başarısız sınıflarını göstermektedir. Her modelin en iyi performans gösterdiği sınıf, doğru sınıflandırma sayısına göre belirtilmiş, en başarısız sınıf ise en yüksek yanlış sınıflandırma sayısına göre tanımlanmıştır. MobileNet ve DenseNet için en başarılı sınıf "Vascular Lesions" (272 doğru sınıf), Xception için ise

"Dermatofibroma" (274 doğru sınıf) olmuştur. En başarısız sınıf tüm modellerde "Melanocytic Nevi" olarak tespit edilmiştir.

Tablo 6: Karmaşıklık matrislerine göre değerlendirme

Model	En Başarılı Olduğu Sınıf	Doğru Sayısı	En Başarısız Olduğu Sınıf	Yanlış Sayısı
MobileNet	Vascular Lesions	272	Melanocytic Nevi	53
DenseNet	Vascular Lesions	272	Melanocytic Nevi	47
Xception	Dermatofibroma	274	Melanocytic Nevi	76

Performans karşılaştırmasına bakıldığında DenseNet, genel doğruluk ve hata oranı açısından en başarılı sonuçları elde etmiştir. Xception, daha karmaşık sınıfları öğrenme potansiyeline sahip olsa da bazı sınıflarda yüksek hata oranı sergilemiştir. MobileNet ise verimliliğine rağmen sınıflar arası hassasiyeti yeterince sağlayamamıştır. Bu gözlemler, karmaşıklık matrislerinde net bir şekilde görülmekte ve sınıfların dağılımı ile görsel benzerliklerin modellerin performansında belirleyici rol oynadığını göstermektedir.

Tablo 7'de literatürdeki çalışmalar ile bu çalışma karşılaştırılmıştır. Kullanılan açıklanabilir yapay zeka yöntemleri ve doğruluk oranları tabloda gösterilmiştir. Tablo incelendiğinde, literatürdeki birçok çalışmanın Grad-CAM ve LIME gibi açıklanabilirlik tekniklerini kullanmadığı görülmektedir. Bu, derin öğrenme modellerinin karar alma süreçlerinin yorumlanmasında önemli bir boşluğa işaret etmektedir. Doğruluk oranları incelendiğinde, önerilen yöntemlerin doğruluk oranlarının literatürdeki bazı çalışmaların gerisinde kaldığı görülüyor. Ancak, doğruluk oranının tek başına bir yöntemin üstünlüğünü belirleyen bir faktör olmadığı unutulmamalıdır. Önerilen yöntemlerin Grad-CAM ve LIME gibi açıklanabilir yapay zeka (XAI) tekniklerini desteklemesi, çalışmanın en önemli avantajlarından biridir. Literatürdeki diğer çalışmaların çoğunda bu tür açıklanabilirlik mekanizmaları bulunmazken, önerilen yöntemler modelin karar verme sürecini daha anlaşılır hale getirmektedir. Bu durum, özellikle tıp, güvenlik veya kritik karar alma süreçlerinde modelin güvenilirliği açısından büyük bir avantaj sağlamaktadır.

Tablo 7 : Çalışmamız ile literatürdeki çalışmaların karşılaştırılması

Yazar	*Grad-CAM*	*LIME*	*Yöntem*	*Doğruluk*
Haythem Ghazouani	✗	✗	MuRANet	%91
Anita Murmu et al.	✗	✗	UNet-MobileNet	%94,27
Khalid M. Hosny et al.	✓	✗	Deep Inherent Learning	%92,89
Rishu Garg et al.	✗	✗	ResNet	%90,51
Wannipa Sae-Lim et al.	✗	✗	Modified MobileNet	%83,93
Abdulmateen Adebiyi	✗	✗	ALBEF	%94,11
Long Hoang et al.	✗	✗	Wide-ShuffleNet	%86,33
Önerilen Yöntem	✓	✓	MobileNet	%84,04
Önerilen Yöntem	✓	✓	DenseNet	%83,57
Önerilen Yöntem	✓	✓	Xception	%81,52

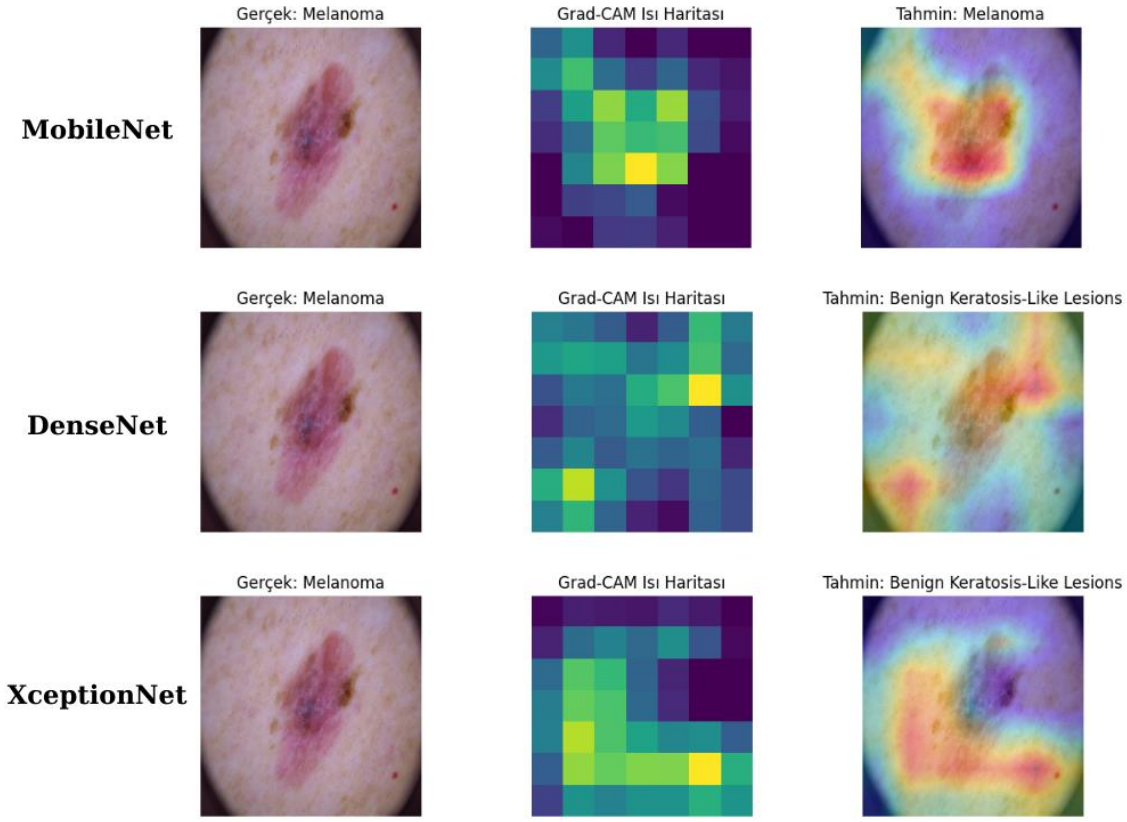
Yalnızca doğruluk oranları üzerinden yapılan değerlendirmelerin, derin öğrenme tabanlı tıbbi karar destek sistemlerinin bütüncül etkinliğini yansıtmada yetersiz kalabilir. Özellikle yüksek risk taşıyan alanlarda, model çıktılarının güvenilirliğini artırmak için kararların yorumlanabilirliğine olan ihtiyaç, son yıllarda literatürde vurgulanan temel gereksinimlerden biri hâline gelmiştir (Doshi-Velez & Kim, 2017; Tjoa & Guan, 2020). Bu bağlamda, çalışmada kullanılan açıklanabilir yapay zekâ yaklaşımlarının yalnızca görsel destek sunmakla kalmayıp, aynı zamanda modelin sınıflandırma sürecinde dikkate aldığı bölgesel örüntüleri kullanıcıya aktarabilme kapasitesi diğer çalışmalardan ayrışan önemli bir yöntemsel katkıdır. Bu katkı, modelin iç işleyişine dair sezgisel bir anlayış sunarak hem klinisyenlerin kararları sorgulayabilmesini hem de kullanıcı-makine güven ilişkisinin güçlendirilmesini mümkün kılar.

Ayrıca, model çıktılarının görsel ve lokal düzeyde açıklanabilmesi, yanlış sınıflandırmaların nedenlerini değerlendirme ve modelin eksik yönlerini hedefli şekilde iyileştirme imkânı da sağlamaktadır. Bu yönüyle çalışma, yalnızca tanı koyma sürecine değil aynı zamanda modelin sürekli geliştirilmesine olanak tanıyan bir geri bildirim

döngüsü oluşturmaktadır. Bu metodolojik yaklaşım, literatürdeki birçok çalışmada eksik kalan “karar süreçlerinin şeffaflığı” ihtiyacına doğrudan yanıt vermektedir.

4.3. Grad-CAM Deneysel Sonuçlar

Grad-CAM yöntemi, modellerin sınıflandırma yaparken dikkat ettiği bölgeleri görsel olarak vurgulamakta ve modelin karar mekanizmalarını anlaşılır hale getirmektedir. Şekil 23’te üç model için elde edilen Grad-CAM çıktıları örnek görseller üzerinden sunulmuştur.

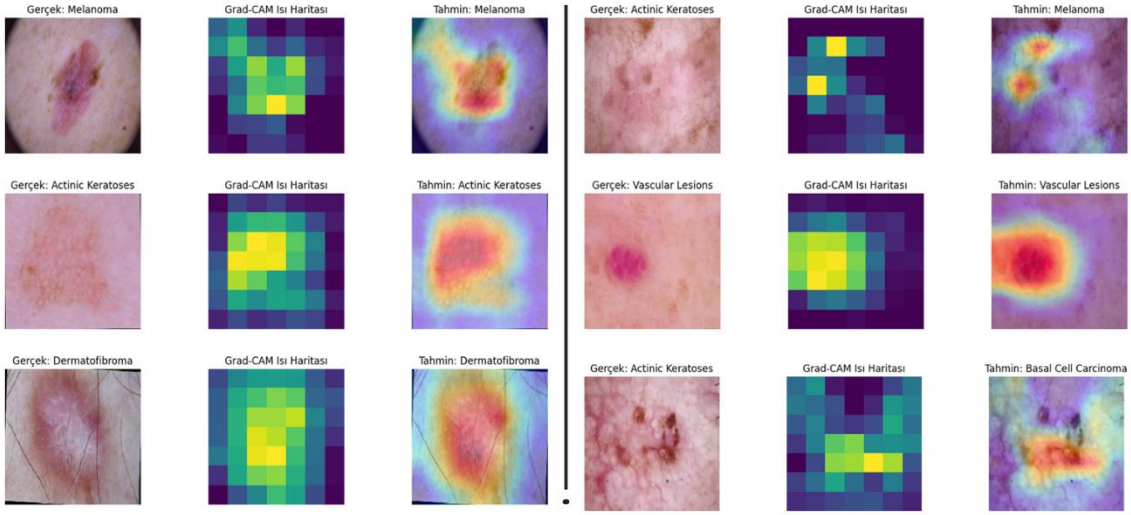


Şekil 23: Üç model için elde edilen Grad-CAM çıktıları ve model tahminleri

4.3.1. MobileNet Modeli

MobileNet modeli için üretilen Grad-CAM çıktıları incelendiğinde, modelin lezyonların genellikle merkez bölgelerine odaklandığı gözlemlenmiştir. Örneğin, "Melanoma" sınıfı için üretilen ısı haritasında, modelin lezyonun tam merkezinde yoğunlaştığı ve bu bölgede kararını şekillendirdiği dikkat çekmektedir. Bununla birlikte bazı örneklerde, özellikle "Actinic Keratoses" gibi sınıflarda, modelin kenar bölgelerine de odaklandığı görülmektedir. Bu durum, lezyonların sınırlarının belirgin olması ve sınıfın ayırt edici özelliklerinin bu bölgelerde yoğunlaşmasıyla açıklanabilir. Genel

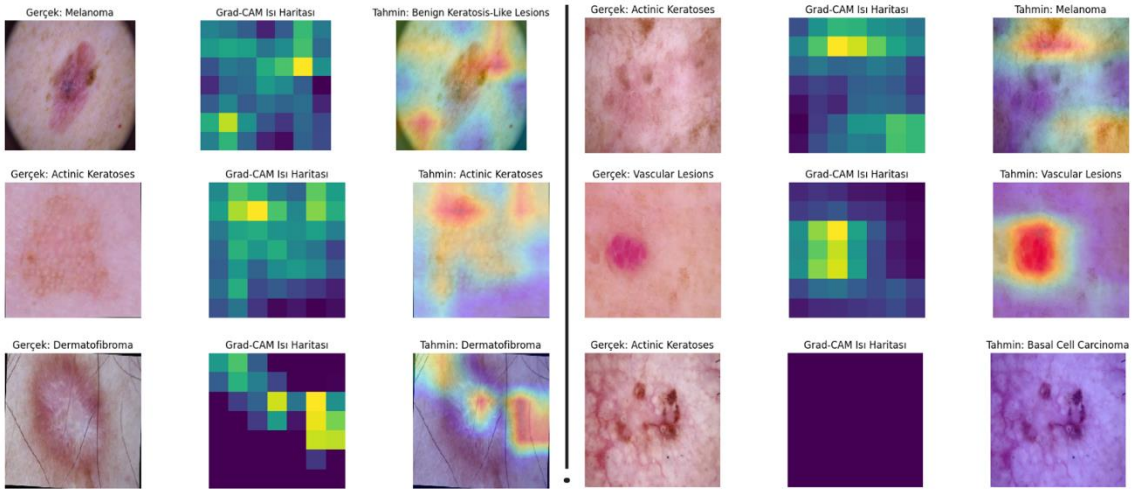
olarak, MobileNet modeli hızlı ve hafif yapısına rağmen karar mekanizmasını doğru bölgelerde yoğunlaştırarak başarılı bir görselleştirme sunmaktadır. Modelin Grad-CAM çıktıları Şekil 24'te gösterilmiştir.



Şekil 24: MobileNet modeline ait Grad-CAM çıktıları ve model tahminleri

4.3.2. DenseNet Modeli

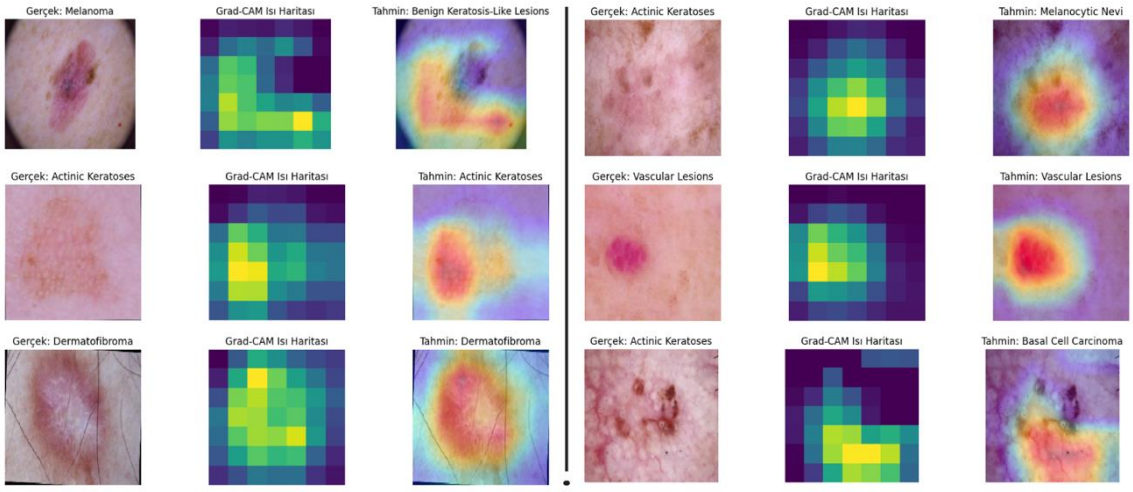
DenseNet modelinin Grad-CAM çıktıları, modelin lezyon bölgelerine daha geniş bir alanda odaklandığını göstermektedir. Örneğin, "Dermatofibroma" sınıfında, modelin sadece lezyon merkezine değil, lezyonun çevresel bölgelerine de dikkat ettiği gözlemlenmiştir. Bu durum, DenseNet'in yoğun bağlantılı katman yapısı sayesinde daha derin özellikleri öğrenme yeteneğiyle ilişkilendirilebilir. Bununla birlikte, bazı sınıflarda (örneğin, "Actinic Keratoses"), modelin dikkatinin lezyon merkezinden ziyade dağınık bir şekilde farklı bölgelere kaydığı görülmektedir. Bu durum, modelin daha fazla detay öğrenmesi sonucu karar mekanizmasında bir miktar belirsizliğe yol açabilir. Ancak genel olarak DenseNet, Grad-CAM çıktılarında kararını destekleyen geniş kapsamlı görsel kanıtlar sunmaktadır. Modelin Grad-CAM çıktıları Şekil 25'te gösterilmiştir.



Şekil 25: DenseNet modeline ait Grad-CAM çıktıları ve model tahminleri

4.3.3. Xception Modeli

Xception modelinin Grad-CAM çıktıları incelendiğinde, modelin hem lezyon merkezine hem de çevresel detaylara odaklandığı gözlemlenmiştir. Özellikle "Vascular Lesions" sınıfında, modelin lezyonun tam merkezini vurguladığı ve kararını bu bölgeye dayandırdığı net bir şekilde görülmektedir. Ancak, "Melanoma" sınıfında modelin odaklanma alanının merkezden kenarlara doğru yayıldığı gözlemlenmiştir. Xception'in daha derin katman yapısı, modelin karar verirken detaylı bilgiye dayandığını, ancak bazen dikkatin belirli bir noktada toplanmadığını göstermektedir. Bu durum, modelin genel hassasiyetini artırsa da odaklanma alanlarının tutarlılığını düşürebilir. Modelin Grad-CAM çıktıları Şekil 26'da gösterilmiştir.



Şekil 26: Xception modeline ait Grad-CAM çıktıları ve model tahminleri

4.3.4. Modeller Arası Karşılaştırma

Modellerin Grad-CAM çıktıları karşılaştırıldığında şu sonuçlar dikkat çekmektedir:

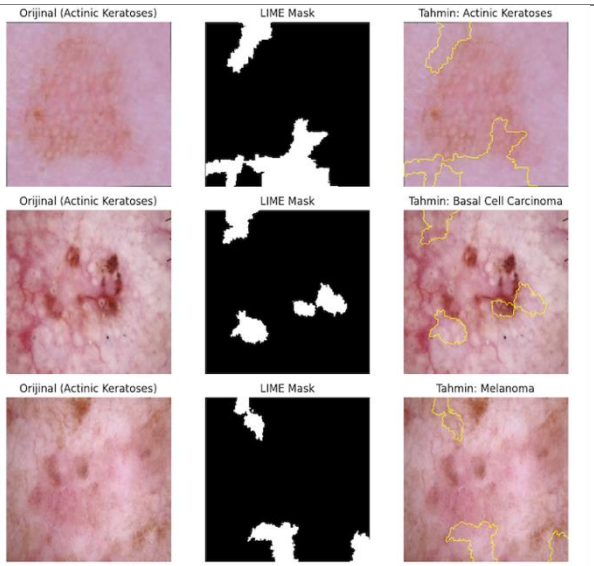
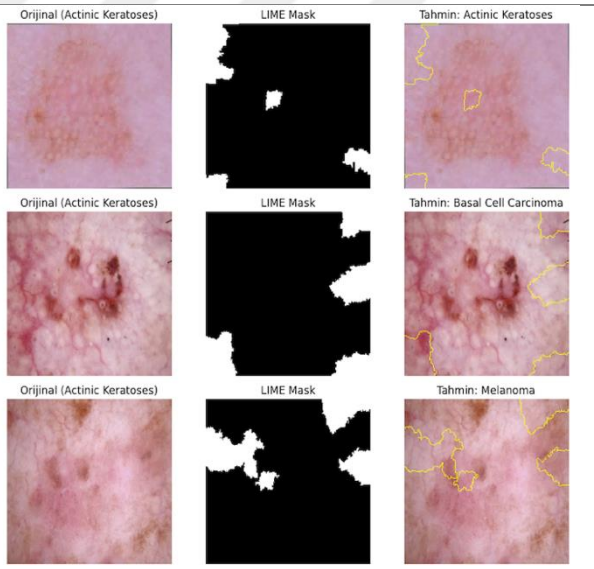
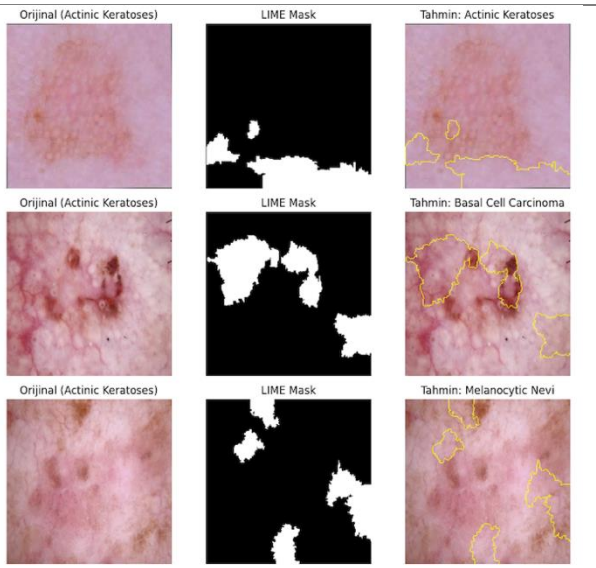
- MobileNet modeli, lezyonların daha çok merkezine odaklanmakta ve karar mekanizmasını belirgin bölgeler üzerinden şekillendirmektedir. Bu, modelin hızlı ve etkili bir görselleştirme sunduğunu göstermektedir.
- DenseNet modeli, geniş kapsamlı bölgelerde odaklanarak lezyonun hem merkezi hem de çevresel detaylarına dikkat etmektedir. Bu durum, karar mekanizmasını güçlendirse de odak noktasını dağıtabilir.
- Xception modeli ise lezyonun hem merkezini hem de kenarlarını dikkate alarak detaylı bir analiz yapmaktadır. Ancak bu durum, bazı örneklerde odaklanma alanının tutarlılığını azaltmaktadır.

Grad-CAM yönteminin sağladığı görsel açıklamalar, modellerin sınıflandırma kararlarını anlamada güçlü bir araç sunmaktadır. MobileNet, DenseNet ve Xception modellerinin dikkat mekanizmaları arasındaki farklılıklar, modellerin mimari yapıları ve öğrenme stratejileriyle doğrudan ilişkilidir. Lokal açıklamalar açısından Grad-CAM, modelin karar mekanizmasını şeffaf hale getirerek klinik uzmanların değerlendirmelerine destek sunar. Özellikle, modelin dikkate aldığı bölgelerin klinik anlamlılığı doğrulanabilir ve modelin potansiyel yanlılıkları tespit edilebilir. Ayrıca, Grad-CAM çıktıları doktorlar için ek bir referans sağlayarak, modelin güvenilirliğini artırmak ve tıbbi teşhis sürecini desteklemek adına önemli bir rol oynar.

4.4. LIME Deneysel Sonuçlar

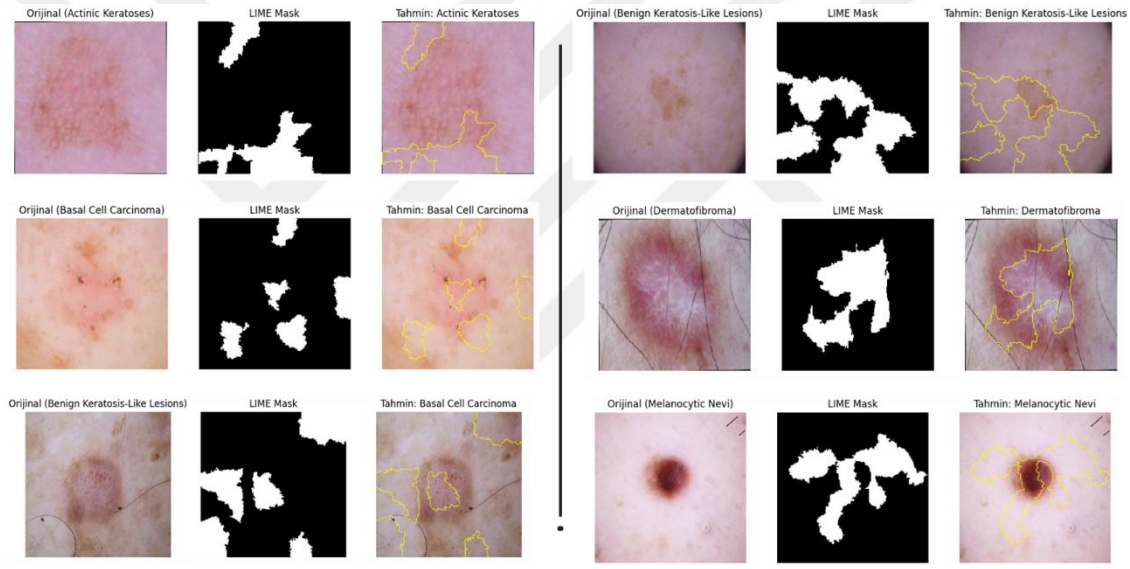
LIME yöntemi, modelin karar verme sürecinde hangi piksellerin veya alanların etkili olduğunu lokal düzeyde açıklamaktadır. Tablo 8’de her model için orijinal görüntüler, LIME maskeleri ve tahmin edilen sınıf sonuçları sunulmuştur.

Tablo 8: tüm modellerin LIME maskeleri ve tahmin edilen sınıf

MobileNet	
DenseNet	
Xception	

4.4.1. MobileNet Modeli

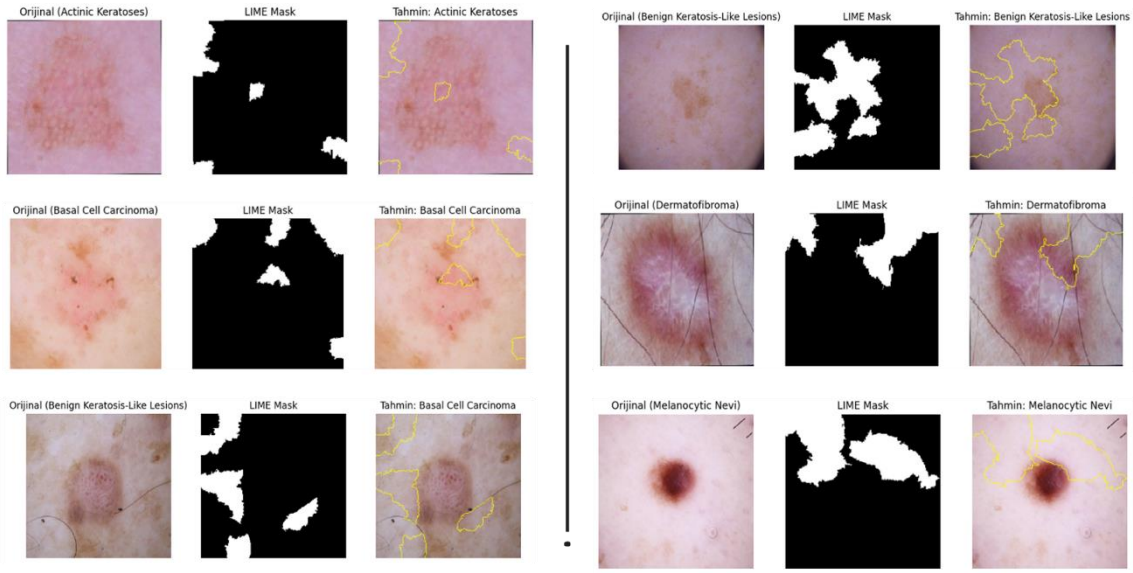
MobileNet modeli için LIME çıktıları incelendiğinde, modelin kararını genellikle lezyonun sınırlarına odaklanarak verdiği gözlemlenmiştir. Özellikle "Actinic Keratoses" ve "Basal Cell Carcinoma" sınıflarında, lezyonun çevresel kenarları belirgin bir şekilde vurgulanmıştır. Bu durum, MobileNet'in hızlı ve hafif yapısı nedeniyle sınıflandırmayı lezyon sınırlarını temel alarak gerçekleştirdiğini göstermektedir. Bazı örneklerde, "Melanoma" sınıfında modelin odaklandığı bölgelerin daha geniş ve dağınık olduğu görülmektedir. Bu durum, lezyon içindeki renk ve doku farklılıklarının modelin kararına katkı sağladığını göstermektedir. MobileNet modeli, LIME maskelerinde sınıflandırma için kritik bölgeleri başarılı bir şekilde işaretlemiştir. Modelin LIME çıktıları Şekil 27'de gösterilmiştir.



Şekil 27: MobileNet modeline ait LIME çıktıları ve model tahminleri

4.4.2. DenseNet Modeli

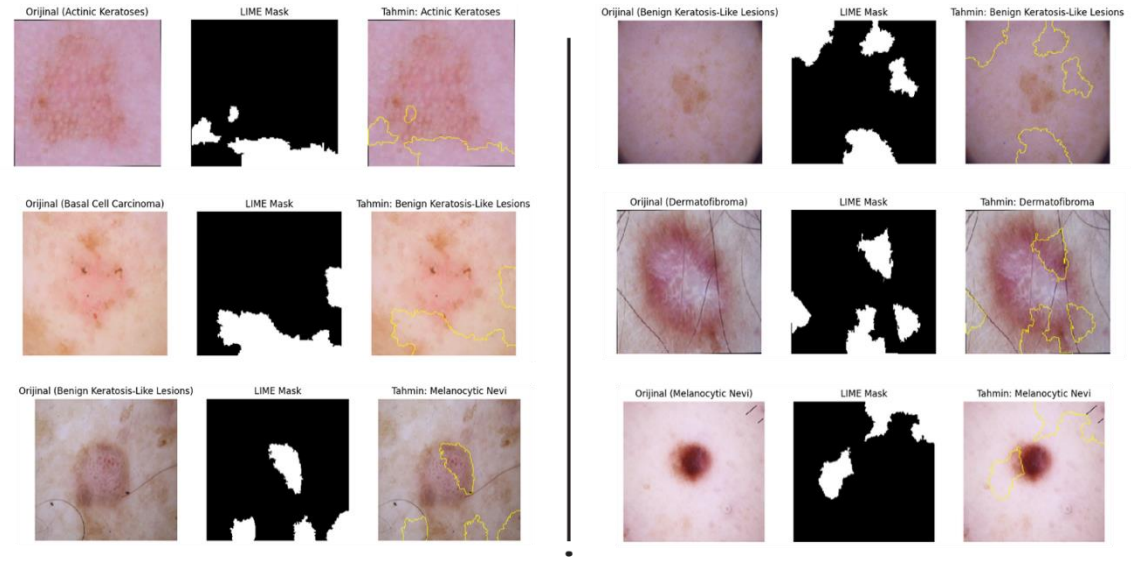
DenseNet modeli, LIME çıktılarında daha geniş kapsamlı bölgeleri açıklamaktadır. Örneğin, "Basal Cell Carcinoma" ve "Melanoma" sınıflarında modelin hem lezyonun merkezine hem de çevresel alanlara dikkat ettiği gözlemlenmiştir. Lezyonun içindeki farklı doku yapılarına ve renk tonlarına odaklanan maskeler, modelin derin bağlantılı yapısının detaylı özellikleri öğrenme yeteneğini göstermektedir. Bununla birlikte, "Actinic Keratoses" sınıfında DenseNet modelinin odaklandığı bölgelerin daha bütüncül olduğu, lezyonun büyük bir kısmını kapsadığı dikkat çekmektedir. Bu durum, modelin karar verirken sınıf özelliklerini daha geniş bir perspektifte değerlendirdiğini ortaya koymaktadır. Modelin LIME çıktıları Şekil 28'de gösterilmiştir.



Şekil 28: DenseNet modeline ait LIME çıktıları ve model tahminleri

4.4.3. Xception Modeli

Xception modelinin LIME çıktıları incelendiğinde, modelin lezyonun hem sınır bölgelerine hem de bazı örneklerde lezyon merkezindeki farklılıklara odaklandığı gözlemlenmiştir. Özellikle "Basal Cell Carcinoma" sınıfında, modelin dikkat ettiği alanların lezyonun sınır çizgisi boyunca yoğunlaştığı dikkat çekmektedir. Buna karşılık "Melanocytic Nevi" sınıfında, Xception modelinin odaklanma bölgelerinin daha dağınık ve bazı durumlarda lezyonun dışına taşarak genişlediği gözlenmiştir. Bu durum, modelin karar mekanizmasında detaylı analiz yapmasına karşın odak noktasının tutarlılığını kaybedebileceğini göstermektedir. Modelin LIME çıktıları Şekil 29'da gösterilmiştir.



Şekil 29: Xception modeline ait LIME çıktıları ve model tahminleri

4.4.4. LIME'ın Sağladığı Lokal Açıklamalar

LIME yöntemi, modellerin sınıflandırma kararlarını piksel düzeyinde açıklayarak, hangi alanların daha fazla katkı sağladığını ortaya koymaktadır. Görsellerde lezyonların kenarlarının ve iç doku farklılıklarının kritik bölgeler olarak vurgulanması, modellerin bu detayları kullanarak karar verdiğini göstermektedir. Tablo 9'da çalışmada kullanılan modellerin LIME'de aldığı sonuçlar yorumlanıştır.

Tablo 9: Modellere göre LIME yorumları

MobileNet	Lezyon sınırlarına odaklanarak hızlı ve doğru kararlar vermektedir.
DenseNet	Lezyonun geniş ve detaylı bölgelerini analiz ederek kapsamlı açıklamalar sunmaktadır.
Xception	Hem sınırları hem de merkezi analiz ederek lokal detayları dikkate almaktadır ancak bazı durumlarda odak noktası dağınık olabilmektedir.

LIME, modellerin karar mekanizmalarını detaylı ve lokal düzeyde açıklayarak, özellikle klinik uygulamalarda hangi bölgelerin model için önemli olduğunu anlamaya yardımcı olmaktadır. Bu sayede, model çıktılarının güvenilirliği ve yorumlanabilirliği önemli ölçüde artırılmıştır.

4.5. Grad-CAM ve LIME Karşılaştırması

Grad-CAM yöntemi, sınıflandırma sırasında modellerin dikkate aldığı genel ve bütüncül bölgeleri vurgularken, LIME yöntemi kararların lokal detaylarını açıklamakta ve piksel düzeyinde açıklamalar sunmaktadır.

Grad-CAM yöntemi ile elde edilen ısı haritaları, modellerin karar mekanizmalarını genel odaklanma alanları üzerinden açıklamaktadır. Görsellerde, özellikle lezyonların merkez bölgeleri Grad-CAM tarafından yoğun bir şekilde vurgulanmaktadır. Grad-CAM, karar mekanizmasını global seviyede açıklayarak modellerin genel davranışlarını anlamada etkili bir yöntemdir. Bununla birlikte, Grad-CAM yöntemi lokal piksel seviyesinde ayrıntı sunmamakta ve daha çok belirli bölgelerin bütüncül olarak önemli olduğunu vurgulamaktadır.

LIME yöntemi, model kararlarını lokal pikseller üzerinden açıklayarak karar mekanizmasının ince detaylarını ortaya çıkarmaktadır. Görsellerde, modellerin sınıflandırmayı özellikle lezyonların sınır bölgeleri ve iç yapısal detayları üzerinden gerçekleştirdiği gözlemlenmiştir. LIME yöntemi, karar mekanizmasını piksel düzeyinde

açıklayarak, modelin hangi detayları kullanarak sınıflandırma yaptığını daha net bir şekilde ortaya koymaktadır. Ancak LIME'ın lokal düzeyde çalışması, görsel bütünlük açısından bazı sınırlılıklar oluşturabilmektedir. Tablo 10'da Grad-CAM ve LIME yöntemleri arasında karşılaştırma yapılmıştır.

Tablo 10: Grad-CAM ve LIME karşılaştırılması

Özellik	**Grad-CAM**	**LIME**
Odaklanma	Bütüncül, geniş bölgelere odaklanma	Lokal, piksel bazlı detaylara odaklanma
Hız	Daha hızlı ve hesaplama açısından daha verimli	Daha fazla hesaplama gerektirir, nispeten yavaş
Açıklama	Anlaşılır, genel davranışı gösterir	Ayrıntılı, detaylı lokal açıklamalar sağlar
Kullanım Alanı	Global karar mekanizmasını anlamak için uygun	Lokal detayların incelenmesi için uygun
Dezavantajlar	Lokal ayrıntıları net göstermez	Bütüncül resmi görmekte zorlanabilir

Grad-CAM ve LIME yöntemleri, model karar mekanizmalarını açıklamada birbirini tamamlayıcı bir rol üstlenmektedir. Grad-CAM, modellerin genel olarak hangi bölgeleri dikkate aldığını görsel olarak ortaya koyarken, LIME bu bölgeleri detaylı bir şekilde analiz ederek hangi piksellerin model kararında etkili olduğunu açıklar.

- Grad-CAM, hızlı ve görsel olarak anlaşılır çıktılar sunarak, modellerin global davranışlarını anlamayı kolaylaştırmaktadır.
- LIME ise detaylı lokal açıklamalar sağlayarak, modellerin kararlarının daha ince ayrıntılar üzerinden nasıl şekillendiğini göstermektedir.

Sonuç olarak, Grad-CAM ve LIME yöntemleri, derin öğrenme modellerinin sınıflandırma kararlarını anlamada güçlü araçlar sunmaktadır. Grad-CAM, modellerin genel odak noktalarını vurgularken, LIME lokal detayları incelemeyi mümkün kılar.

5. SONUÇLAR VE ÖNERİLER

Çalışmada kullanılan yöntemler, deneysel çalışmanın temel hipotezine büyük ölçüde cevap vermektedir. Derin öğrenme modellerinin performanslarının karşılaştırılması, kullanılan farklı mimarilerin güçlü ve zayıf yönlerini ortaya koymuştur. MobileNet, hafif yapısı sayesinde performans ve verimlilik açısından öne çıkarken, DenseNet ve Xception modelleri derinlikleri sayesinde daha karmaşık özellikleri öğrenerek farklı avantajlar sağlamıştır. Görselleştirme yöntemleri olan Grad-CAM ve LIME, model kararlarının analizinde büyük katkı sağlamış ve modellerin hangi görsel detaylara dayalı olarak sınıflandırma yaptığı net bir şekilde açıklanmıştır. Bu yöntemler, çalışmanın güvenilirlik ve yorumlanabilirlik hedeflerine ulaşmasında kritik rol oynamıştır.

5.1 Sonuçlar

Bu çalışmada MobileNet, DenseNet ve Xception modellerinin Ham10000 veri seti üzerinde deri lezyonlarının sınıflandırılması performansları karşılaştırılmış ve görselleştirme yöntemleriyle analiz edilmiştir. Modellerin performans değerlendirmeleri doğrultusunda MobileNet, %84.04 doğruluk oranı ve yüksek F1-Score değeriyle en iyi performansı göstermiştir. MobileNet'in hızlı ve hafif yapısı, eğitim süresinin kısalmasına ve tutarlı sonuçlar elde edilmesine katkı sağlamıştır. DenseNet modeli %83.57 doğruluk oranıyla benzer performans gösterirken, daha derin yapısı sayesinde geniş kapsamlı bölgelerde analiz yaparak kararlarını şekillendirmiştir. Xception modeli ise %81.52 doğruluk oranıyla diğer modellere kıyasla daha düşük performans sergilemiş; ancak yüksek sensitivite değeriyle pozitif sınıfları doğru tahmin etmede güçlü bir performans göstermiştir.

Çalışmada modellerin karar mekanizmalarını analiz etmek için Grad-CAM ve LIME yöntemleri kullanılmıştır. Yapılan deneyler ve analizler bu yöntemlerin derin öğrenme tabanlı modellerin karar süreçlerini daha anlaşılır ve açıklanabilir hale getirdiğini göstermektedir. Grad-CAM, modellerin hangi bölgeleri dikkate alarak karar verdiğini görsel olarak sunarak, uzmanların modelin karar mekanizmalarını yorumlamasına olanak tanımıştır. LIME yöntemi ise her bir sınıflandırma kararını lokal düzeyde analiz ederek, hangi piksellerin veya görüntü bölgelerinin model kararında etkili olduğunu göstermiştir. Bu yöntemler sayesinde, modellerin doğru sınıflandırma yaptığı örneklerde hangi görsel detayların önemli olduğu belirlenmiş, aynı zamanda hatalı

tahminler yapan modellerde yanlış yönlendiren faktörler tespit edilmiştir. Grad-CAM, modelin sınıflandırma yaparken hangi bölgelere odaklandığını geniş ölçekli bir perspektifte gösterirken, LIME, kararın detaylarını piksel bazında açıklayarak daha spesifik bir içgörü sunmaktadır. Bu görselleştirme tekniklerinin kullanımı, tıbbi görüntüleme verilerinde kullanılan modellerin şeffaflığını artırarak klinik uygulamalarda AI tabanlı çözümlerin kabul edilebilirliğini yükseltmektedir. Sonuç olarak, Grad-CAM ve LIME yöntemleri, yapay zeka tabanlı deri lezyon sınıflandırma sistemlerinin güvenilirliğini artırarak klinik ortamlarda kullanılabilirliğini güçlendirmektedir.

5.2 Öneriler

Bu çalışmanın bulgularına dayanarak gelecekte yapılacak araştırmalar için bazı öneriler sunulabilir. İlk olarak, Ham10000 veri setinde gözlemlenen veri dengesizliği, özellikle azınlık sınıflarındaki performansı iyileştirmek adına veri çoğaltma (data augmentation) teknikleri veya sınıf ağırlıklandırma gibi yöntemlerle ele alınabilir. Veri setinin daha dengeli hale getirilmesi, modellerin bu sınıflardaki hassasiyetini artırabilir.İkinci olarak, DenseNet ve Xception gibi derin yapılı modellerin eğitimi zaman ve donanım açısından maliyetli olduğundan, bu tür sınırlı kaynaklara sahip çalışmalarda MobileNet gibi hafif ve hızlı modellerin kullanımı teşvik edilebilir. Bunun yanı sıra, model mimarilerinde optimizasyon tekniklerinin uygulanması da eğitim süresini ve donanım gereksinimlerini azaltmada faydalı olabilir. Üçüncü olarak, elde edilen sonuçların genel geçerliliğini artırmak için farklı veri setlerinde, cilt tiplerinde ve lezyon türlerinde model performanslarının test edilmesi önerilmektedir. Bu sayede, modellerin klinik ortamlarda geniş bir hasta grubuna uygulanabilirliği daha sağlıklı değerlendirilebilir. Son olarak, Grad-CAM ve LIME gibi açıklanabilirlik yöntemlerinin bazı durumlarda yetersiz kalabildiği göz önüne alındığında, daha gelişmiş veya alternatif açıklanabilirlik tekniklerinin kullanılması üzerinde durulabilir. Bu tür yöntemlerin entegrasyonu, modellerin karar mekanizmalarının daha tutarlı ve ayrıntılı bir şekilde analiz edilmesini sağlayarak güvenilirliği artırabilir. Bu öneriler doğrultusunda, deri lezyonlarının sınıflandırılmasında derin öğrenme modellerinin performansı ve klinik uygulamalardaki kullanılabilirliği daha ileri seviyelere taşınabilir.

KAYNAKLAR

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). *IEEE access*, *6*, 52138-52160.

Akyel, C., & Arıcı, N. (2022). Cilt Kanseri Görüntülerinde FCN8-ResNetC ve Görüntü İşleme ile Kıl Temizliği ve Lezyon Bölütleme. *International Journal of InformaticsTechnologies*, *15*(2).

Al Kafee, A. (2013). *Skin cancer detection using image processing techniques* Fatih Üniversitesi].

Alam, T. M., Shaukat, K., Khan, W. A., Hameed, I. A., Almuqren, L. A., Raza, M. A., Aslam, M., & Luo, S. (2022). An efficient deep learning-based skin cancer classifier for an imbalanced dataset. *Diagnostics*, *12*(9), 2115.

Albelwi, S. A. (2022). Deep architecture based on DenseNet-121 model for weather image recognition. *International Journal of Advanced Computer Science and Applications*, *13*(10).

Alexandropoulos, S.-A. N., Kotsiantis, S. B., & Vrahatis, M. N. (2019). Data preprocessing in predictive data mining. *The Knowledge Engineering Review*, *34*, e1.

Alezabı, A. M. M. (2021). *Görüntü işleme teknikleriyle melanom cilt kanseri bölütlemesi* Kastamonu üniversitesi].

Ali, A., & Mohamed, I. (2018). *Bulanık kümeleme ortalamasına dayalı görüntü bölgesi büyümesiyle melanoma cilt kanseri segmentasyonu (Melanoma skin cancer segmentation with image region growing based on fuzzy clustering mean)* Kastamonu University, Turkey].

Aliferis, C., & Simon, G. (2024). Overfitting, underfitting and general model overconfidence and under-performance pitfalls and best practices in machine learning and AI. *Artificial intelligence and machine learning in health care and medical sciences: Best practices and pitfalls*, 477-524.

Alp, H., Tutun, H., Altıntaş, L., Kaplan, H. M., Şingirik, E., & Kahraman, İ. (2021). Metforminin Melanom Cilt Kanseri Hücreleri Üzerindeki Etkisinin Araştırılması. *Genel Tıp Dergisi*, *31*(2), 111-115.

Amann, J., Blasimme, A., Vayena, E., Frey, D., Madai, V. I., & Consortium, P. Q. (2020). Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC medical informatics and decision making*, *20*, 1-9.

Angelov, P. P., Soares, E. A., Jiang, R., Arnold, N. I., & Atkinson, P. M. (2021). Explainable artificial intelligence: an analytical review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *11*(5), e1424.

Arora, A. (August 2, 2020). *DenseNet Architecture Explained with PyTorch Implementation from TorchVision*. Retrieved 12.12 from https://amaarora.github.io/posts/2020-08-02-densenets.html

Barzegar, R., Aalami, M. T., & Adamowski, J. (2020). Short-term water quality variable prediction using a hybrid CNN–LSTM deep learning model. *Stochastic Environmental Research and Risk Assessment*, *34*(2), 415-433.

Beauxis-Aussalet, E., & Hardman, L. (2014). Simplifying the visualization of confusion matrix. 26th Benelux conference on artificial intelligence (BNAIC),

Bhandari, M., Yogarajah, P., Kavitha, M. S., & Condell, J. (2023). Exploring the capabilities of a lightweight CNN model in accurately identifying renal abnormalities: Cysts, stones, and tumors, using LIME and SHAP. *Applied Sciences*, *13*(5), 3125.

Bulam, H. M., Kaya, B., Şencan, A., Tuncer, S., & Latifoğlu, O. (2010). Arsenikozis ve çoklu deri kanseri. *Türk Plastik Rekonstrüktif Ve Estetik Cerrahi Dergisi*, *18*(2), 81-84.

Bullinaria, J. A. (2015). Bias and variance, under-fitting and over-fitting. *Neural Computation: Lecture*, *9*, 16.

Chen, H.-Y., & Su, C.-Y. (2018). An enhanced hybrid MobileNet. 2018 9th International Conference on Awareness Science and Technology (iCAST),

Chhabra, M., & Kumar, R. (2022). A smart healthcare system based on classifier DenseNet 121 model to detect multiple diseases. In *Mobile Radio Communications and 5G Networks: Proceedings of Second MRCN 2021* (pp. 297-312). Springer.

Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. Proceedings of the IEEE conference on computer vision and pattern recognition,

Crowson, A. N., Magro, C. M., & Mihm Jr, M. C. (2014). *The melanocytic proliferations: a comprehensive textbook of pigmented lesions*. John Wiley & Sons.

Davis, L. E., Shalin, S. C., & Tackett, A. J. (2019). Current state of melanoma diagnosis and treatment. *Cancer biology & therapy*, *20*(11), 1366-1379.

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*.

Endah, S. N., & Shiddiq, I. N. (2020). Xception architecture transfer learning for garbage classification. 2020 4th International Conference on Informatics and Computational Sciences (ICICoS),

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *nature*, *542*(7639), 115-118.

Famili, A., Shen, W.-M., Weber, R., & Simoudis, E. (1997). Data preprocessing and intelligent data analysis. *Intelligent data analysis*, *1*(1), 3-23.

Flach, P. A. (2016). ROC analysis. In *Encyclopedia of machine learning and data mining* (pp. 1-8). Springer.

Ghassemi, M., Oakden-Rayner, L., & Beam, A. L. (2021). The false hope of current approaches to explainable artificial intelligence in health care. *The Lancet Digital Health*, *3*(11), e745-e750.

Ghazouani, H. (2025). Multi-residual attention network for skin lesion classification. *Biomedical Signal Processing and Control*, *103*, 107449.

Göreke, V. (2021). Cilt lezyonlarının sınıflandırılmasında derin öğrenme tabanlı bir yöntem. *Türk Doğa ve Fen Dergisi*, *10*(1), 30-36.

Hasanin, T., & Khoshgoftaar, T. (2018). The effects of random undersampling with simulated class imbalance for big data. 2018 IEEE international conference on information reuse and integration (IRI),

Hauser, K., Kurz, A., Haggenmüller, S., Maron, R. C., von Kalle, C., Utikal, J. S., Meier, F., Hobelsberger, S., Gellrich, F. F., & Sergon, M. (2022). Explainable artificial intelligence in skin cancer recognition: A systematic review. *European Journal of Cancer*, *167*, 54-69.

Hosny, K. M., Said, W., Elmezain, M., & Kassem, M. A. (2024). Explainable deep inherent learning for multi-classes skin lesion classification. *Applied Soft Computing*, *159*, 111624.

Hou, Y., Wu, Z., Cai, X., & Zhu, T. (2024). The application of improved densenet algorithm in accurate image recognition. *Scientific Reports*, *14*(1), 8645.

Howard, A. G. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.

Huang, G., Liu, S., Van der Maaten, L., & Weinberger, K. Q. (2018). Condensenet: An efficient densenet using learned group convolutions. Proceedings of the IEEE conference on computer vision and pattern recognition,

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. Proceedings of the IEEE conference on computer vision and pattern recognition,

Huang, J., & Ling, C. X. (2005). Using AUC and accuracy in evaluating learning algorithms. *IEEE Transactions on knowledge and Data Engineering*, *17*(3), 299-310.

Hurmuzi, S. M. A. (2019). Evrişimsel sinir ağları kullanarak cilt kanseri risk değerlendirmesi.

Hussain, S., Songhua, X., Aslam, M. U., Waqas, M., & Hussain, F. (2024). Hypergraph Convolutional Neural Networks for Clinical Diagnosis of Monkeypox Infections Using Skin Virological Images. *Applied Soft Computing*, 112673.

Iqbal, I., Younus, M., Walayat, K., Kakar, M. U., & Ma, J. (2021). Automated multi-class classification of skin lesions through deep convolutional neural network with dermoscopic images. *Computerized medical imaging and graphics*, *88*, 101843.

Jutzi, T. B., Krieghoff-Henning, E. I., Holland-Letz, T., Utikal, J. S., Hauschild, A., Schadendorf, D., Sondermann, W., Fröhling, S., Hekler, A., & Schmitt, M. (2020). Artificial intelligence in skin cancer diagnostics: the patients' perspective. *Frontiers in medicine*, *7*, 233.

Khoei, T. T., & Kaabouch, N. (2022). Densely connected neural networks for detecting denial of service attacks on smart grid network. 2022 IEEE 13th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON),

Kılıç, K. (2021). *Cilt kanseri tespitinde evrişimsel sinir ağları mimarilerinin karşılaştırılması* Ondokuz Mayıs Üniversitesi Lisansüstü Eğitim Enstitüsü].

Koklu, M., Cinar, I., & Taspinar, Y. S. (2021). Classification of rice varieties with deep learning methods. *Computers and Electronics in Agriculture*, *187*, 106285.

Koklu, M., & Ozkan, I. A. (2020). Multiclass classification of dry beans using computer vision and machine learning techniques. *Computers and Electronics in Agriculture*, *174*, 105507.

Krawczyk, B., Galar, M., Jeleń, Ł., & Herrera, F. (2016). Evolutionary undersampling boosting for imbalanced classification of breast cancer malignancy. *Applied Soft Computing*, *38*, 714-726.

Li, H., Pan, Y., Zhao, J., & Zhang, L. (2021). Skin disease diagnosis with deep learning: A review. *Neurocomputing*, *464*, 364-393.

Liu, X., Yu, Z., Tan, L., Yan, Y., & Shi, G. (2024). Enhancing skin lesion diagnosis with ensemble learning. 2024 4th International Conference on Computer Science and Blockchain (CCSB),

Liu, Y., Pang, Z., Karlsson, M., & Gong, S. (2020). Anomaly detection based on machine learning in IoT-based vertical plant wall for indoor climate control. *Building and Environment*, *183*, 107212.

Maharana, K., Mondal, S., & Nemade, B. (2022). A review: Data pre-processing and data augmentation techniques. *Global Transitions Proceedings*, *3*(1), 91-99.

Majnik, M., & Bosnić, Z. (2013). ROC analysis of classifiers in machine learning: A survey. *Intelligent data analysis*, *17*(3), 531-558.

Mohammed, M. A., Mukhlif, A. A., & Khateeb, B. A. (2023). Breast cancer images Classification using a new transfer learning technique. *Iraqi Journal for Computer Science and Mathematics*, *4*(1), 16.

Mohammed, R., Rawashdeh, J., & Abdullah, M. (2020). Machine learning with oversampling and undersampling techniques: overview study and experimental results. 2020 11th international conference on information and communication systems (ICICS),

Muhammad, W., Aramvith, S., & Onoye, T. (2021). Multi-scale Xception based depthwise separable convolution for single image super-resolution. *Plos one*, *16*(8), e0249278.

Murmu, A., Swati, S., Kumar, N., Murmu, R., & Dash, Y. (2024). Deep Learning Based Encoder-Decoder Model for Skin Cancer Classification. 2024 7th International Conference on Contemporary Computing and Informatics (IC3I),

Muthu Subathra, L., Krishnaveni, M., Paul Keins, B., & Joshua Vimal Raj, R. (2023). Breast Cancer Detection With Resnet50, Inception V3, And Xception Architecture. *Journal of Pharmaceutical Negative Results*, 60-68.

Nahata, H., & Singh, S. P. (2020). Deep learning solutions for skin cancer detection and diagnosis. *Machine learning with health care perspective: machine learning and healthcare*, 159-182.

Noh, S.-H. (2021). Performance comparison of CNN models using gradient flow analysis. Informatics,

Ogudo, K. A., Surendran, R., & Khalaf, O. I. (2023). Optimal Artificial Intelligence Based Automated Skin Lesion Detection and Classification Model. *Computer Systems Science & Engineering*, *44*(1).

Orhan, H., & Yavşan, E. (2023). Artificial intelligence-assisted detection model for melanoma diagnosis using deep learning techniques. *Mathematical Modelling and Numerical Simulation with Applications*, *3*(2), 159-169.

Öncel, S., & Gündoğdu, D. (2017). Deri Kanseri Risk Algısının Güneşten Korunma Davranışlarına Etkisi: Sistematik Derleme. *Turkiye Klinikleri Public Health Nursing-Special Topics*, *3*(1), 52-60.

Poyraz, G., Öztürk, F., & Kırcı, P. (2021). Cancer detection with an image processing application. *Avrupa Bilim ve Teknoloji Dergisi*(28), 1442-1146.

Rahman, I. (2023). AI-powered Personalized Treatment Recommendation Framework for Improved Healthcare Outcomes. *Journal of Computational Social Dynamics*, *8*(11), 42-51.

Ribeiro, M. T. (2016). *Local Interpretable Model-Agnostic Explanations (LIME): An Introduction*. https://www.oreilly.com/content/introduction-to-local-interpretable-model-agnostic-explanations-lime/

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). " Why should i trust you?" Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining,

Rigel, D. S., Russak, J., & Friedman, R. (2010). The evolution of melanoma diagnosis: 25 years beyond the ABCDs. *CA: a cancer journal for clinicians*, *60*(5), 301-316.

Roy, S., Pal, D., & Meena, T. (2023). Explainable artificial intelligence to increase transparency for revolutionizing healthcare ecosystem and the road ahead. *Network Modeling Analysis in Health Informatics and Bioinformatics*, *13*(1), 4.

Rukhsar, L., Bangyal, W. H., Ali Khan, M. S., Ag Ibrahim, A. A., Nisar, K., & Rawat, D. B. (2022). Analyzing RNA-seq gene expression data using deep learning approaches for cancer classification. *Applied Sciences*, *12*(4), 1850.

Sahoo, A. K., Pradhan, C., & Das, H. (2020). Performance evaluation of different machine learning methods and deep-learning based convolutional neural network for health decision making. *Nature inspired computing for data science*, 201-212.

Samek, W., Wiegand, T., & Müller, K.-R. (2017). Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. *arXiv preprint arXiv:1708.08296*.

Saraswat, D., Bhattacharya, P., Verma, A., Prasad, V. K., Tanwar, S., Sharma, G., Bokoro, P. N., & Sharma, R. (2022). Explainable AI for healthcare 5.0: opportunities and challenges. *IEEE access*, *10*, 84486-84517.

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. Proceedings of the IEEE international conference on computer vision,

Selvaraju, R. R., Das, A., Vedantam, R., Cogswell, M., Parikh, D., & Batra, D. (2016). Grad-CAM: Why did you say that? *arXiv preprint arXiv:1611.07450*.

Soylu, E., & Demir, R. (2021). Development and comparison of skin cancer diagnosis models. *Avrupa Bilim ve Teknoloji Dergisi*(28), 1217-1221.

Şahin, M. (2016). *Görüntü işleme tekniklerini kullanarak malign melanom tipi deri kanserinin incelenmesi* Altınbaş Üniversitesi].

Şahin, N., & Alpaslan, N. (2020a). Cilt Lezyon Bölütlemesi için Metasezgisel Temelli Otsu Eşikleme Yöntemi. *Türk Doğa ve Fen Dergisi*, *9*(1), 42-48.

Şahin, N., & Alpaslan, N. (2020b). SegNet Mimarisi Kullanılarak Cilt Lezyon Bölütleme Performansının İyileştirilmesi. *Avrupa Bilim ve Teknoloji Dergisi*, 40-45.

Taspinar, Y. S., Koklu, M., & Altin, M. (2022). Acoustic-driven airflow flame extinguishing system design and analysis of capabilities of low frequency in different fuels. *Fire technology*, *58*(3), 1579-1597.

Ting, K. M. (2010). Confusion Matrix. In C. Sammut & G. I. Webb (Eds.), *Encyclopedia of Machine Learning* (pp. 209-209). Springer US. https://doi.org/10.1007/978-0-387-30164-8_157

Tiwari, R. S., Dandabani, L., Das, T. K., Khan, S. B., Basheer, S., & Alqahtani, M. S. (2023). Cloud-Based Quad Deep Ensemble Framework for the Detection of COVID-19 Omicron and Delta Variants. *Diagnostics*, *13*(22), 3419.

Tjoa, E., & Guan, C. (2020). A survey on explainable artificial intelligence (xai): Toward medical xai. *IEEE transactions on neural networks and learning systems*, *32*(11), 4793-4813.

Townsend, J. T. (1971). Theoretical analysis of an alphabetic confusion matrix. *Perception & Psychophysics*, *9*, 40-50.

Tsang, S.-H. (2018). *Review: Xception — With Depthwise Separable Convolution, Better Than Inception-v3 (Image Classification)*. Retrieved 14.11.2024 from https://medium.com/data-science/review-xception-with-depthwise-separable-convolution-better-than-inception-v3-image-dc967dd42568

Tschandl, P., Rosendahl, C., & Kittler, H. (2018). The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific data*, *5*(1), 1-9.

Uslu, M., Karaman, G., Şavk, E., & Şendur, N. (2006). Adnan Menderes Üniversitesi hekimlerinin deri kanserleri ve güneşin etkileri konusundaki bilgi düzeyleri ile güneşten korunma davranışlarının değerlendirilmesi.

Uysal, A., Özsoy, S. A., & Ergül, Ş. (2004). ÖĞRENCİLERİN CİLT KANSERİ RİSKLERİNİN VE GÜNEŞ IŞINLARINDAN KORUNMAYA YÖNELİK UYGULAMALARININ DEĞERLENDİRİLMESİ. *Ege Tıp Dergisi*, *43*(2), 95-99.

Viloria, A., Lezama, O. B. P., & Mercado-Caruzo, N. (2020). Unbalanced data processing using oversampling: machine learning. *Procedia Computer Science*, *175*, 108-113.

Won, Y.-S., Jap, D., & Bhasin, S. (2020). Push for more: On comparison of data augmentation and SMOTE with optimised deep learning architecture for side-channel. Information Security Applications: 21st International Conference, WISA 2020, Jeju Island, South Korea, August 26–28, 2020, Revised Selected Papers 21,

Yalçın, T., & Gürsel, A. T. (2021). Improving Digital Image Quality for Convolution Neural Network Based Computer-Aided Diagnosis (CNN-CAD) of Skin Cancer. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, *36*(4), 1099-1110.

Yang, C., Fridgeirsson, E. A., Kors, J. A., Reps, J. M., & Rijnbeek, P. R. (2024). Impact of random oversampling and random undersampling on the performance of prediction models developed using observational health data. *Journal of big data*, *11*(1), 7.

Yang, C. C. (2022). Explainable artificial intelligence for predictive modeling in healthcare. *Journal of healthcare informatics research*, *6*(2), 228-239.

Zhang, N., Cai, Y.-X., Wang, Y.-Y., Tian, Y.-T., Wang, X.-L., & Badami, B. (2020). Skin cancer diagnosis based on optimized convolutional neural network. *Artificial intelligence in medicine*, *102*, 101756.

Zhang, Y., Hong, D., McClement, D., Oladosu, O., Pridham, G., & Slaney, G. (2021). Grad-CAM helps interpret the deep learning models trained to classify multiple sclerosis types using clinical brain magnetic resonance imaging. *Journal of Neuroscience Methods*, *353*, 109098.

Zhu, F., Liu, C., Yang, J., & Wang, S. (2022). An improved mobileNet network with wavelet energy and global average pooling for rotating machinery fault diagnosis. *Sensors*, *22*(12), 4427.

Zhu, Y., & Newsam, S. (2017). Densenet for dense flow. 2017 IEEE international conference on image processing (ICIP),

Zuo, T., Zheng, Y., He, L., Chen, T., Zheng, B., Zheng, S., You, J., Li, X., Liu, R., & Bai, J. (2021). Automated classification of papillary renal cell carcinoma and chromophobe renal cell carcinoma based on a small computed tomography imaging dataset using deep learning. *Frontiers in Oncology*, *11*, 746750.