

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ



VÜCUT BENLERİNİN KLİNİK GÖRÜNTÜLERİ ÜZERİNDE
DERİN ÖĞRENME TABANLI HASTALIK TAHMİNİ
SİSTEMİNİN GELİŞTİRİLMESİ

Zafer TOLAN

Yüksek Lisans Tezi

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
Donanım Bilim Dalı

OCAK, 2022

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Bilgisayar Mühendisliği Anabilim Dalı

Donanım Bilim Dalı

Yüksek Lisans Tezi

**VÜCUT BENLERİNİN KLİNİK GÖRÜNTÜLERİ ÜZERİNDE
DERİN ÖĞRENME TABANLI HASTALIK TAHMİNİ
SİSTEMİNİN GELİŞTİRİLMESİ**

Tez Yazarı
Zafer TOLAN

Danışman
Dr. Öğr. Üyesi Erkan DUMAN

OCAK, 2022
ELAZIĞ

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Bilgisayar Mühendisliği Anabilim Dalı

Yüksek Lisans Tezi

Başlığı: **Vücut Benlerinin Klinik Görüntüleri Üzerinde Derin Öğrenme Tabanlı Hastalık Tahmini Sisteminin Geliştirilmesi**

Yazarı: **Zafer TOLAN**

İlk Teslim Tarihi: **30.12.2021**

Savunma Tarihi: **30.01.2022**

TEZ ONAYI

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

Danışman:	Dr. Öğr. Üyesi Erkan DUMAN Fırat Üniversitesi, Mühendislik Fakültesi	<i>İmza</i> Onayladım
Başkan:	Doç. Dr. Ahmet ÇINAR Fırat Üniversitesi, Mühendislik Fakültesi	Onayladım
Üye:	Doç. Dr. Bülent Gürsel EMİROĞLU Kırıkkale Üniversitesi, Mühendislik Fakültesi	Onayladım

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

İmza

Prof. Dr. Kürşat Esat ALYAMAÇ
Enstitü Müdürü

BEYAN

Fırat Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygun olarak hazırladığım “Vücut Benlerinin Klinik Görüntüleri Üzerinde Derin Öğrenme Tabanlı Hastalık Tahmini Sisteminin Geliştirilmesi” isimli yüksek lisans tezimin içindeki bütün bilgilerin doğru olduğunu, bilgilerin üretilmesi ve sunulmasında bilimsel etik kurallarına uygun davrandığımı, kullandığım bütün kaynakları atıf yaparak belirttiğimi, maddi ve manevi desteği olan tüm kurum/kuruluş ve kişileri belirttiğimi, burada sunduğum veri ve bilgileri unvan almak amacıyla daha önce hiçbir şekilde kullanmadığımı beyan ederim.

30.01.2022

Zafer TOLAN



ÖNSÖZ

Bu çalışmada derin öğrenme teknikleriyle, hastaların vücudundaki ben görüntülerindeki örüntüler keşfedilerek melanoma cilt kanserinin tanısı yapılmaya çalışılmıştır. Günümüzde hastalığın erken teşhisi ve tedavisi her geçen gün önem kazanmaktadır. Melanoma tedavisinde erken teşhis, tedavi ve yaşam süreci açısından önemli bir rol oynamaktadır. Dermoskopik görüntülere ulaşmak ve bunlar üzerinde veri çoğaltma teknikleri uygulamak, çalışmanın en çok çaba sarfedilen kısmı olmuştur.

Bu çalışmada, bilgi ve deneyimleriyle beni yönlendiren, desteğini esirgemeyen ve karşılaştığım sıkıntılarda yardımcı olmaya çalışan, akademik tecrübesiyle bana yol gösteren değerli danışman hocam Dr. Öğr. Üyesi Üyesi Erkan DUMAN'a teşekkürlerimi sunarım. Ayrıca bu süreç boyunca bana yol gösteren, başarılı olacağıma inanan aileme, bana moral ve motivasyon kaynağı olan eşime teşekkür ederim.

Zafer TOLAN

ELAZIĞ, 2022

İÇİNDEKİLER

Sayfa

ÖNSÖZ.....	iv
İÇİNDEKİLER	v
ÖZET	vii
ABSTRACT	viii
ŞEKİLLER LİSTESİ	ix
TABLolar LİSTESİ	x
KISALTMALAR	xi
1. GİRİŞ	1
2. EVRİŞİMSEL SINIR AĞLARI İLE VÜCUT BENLERİNİN ANALİZİ	7
2.1 Önerilen Evrişimsel Mimari.....	8
2.2. İkili Sınıflandırma Uygulamalarında Başarı Ölçütleri – Metrik Çeşitleri	13
2.2.1. Başarı ölçütleri	14
2.2.2. Hata Matrisi (Confusion Matrix)	17
2.2.3. ROC Eğrisi (Receiver Operating Characteristic Curve)	18
3. ULUSLARARASI CİLT GÖRÜNTÜLEME BİRLİĞİ VERİ SETİ (ISIC ARŞİVİ)	21
3.1. ISIC Arşiv Görüntülerinin Elde Edilmesi ve Saklanması	26
3.1.1. PowerPoint dosyalarından görüntülerin ve meta verilerin çıkarılması	27
3.1.2. Dijital dermatoskopi sisteminden verilerin alınması	27
3.1.3. Dermatoskopik görüntülerin filtrelenmesi	27
3.2. ISIC Veri Setindeki Mevcut Hastalık Tanımları	28
3.2.1. Melanoma(Melanom)	28
3.2.2. Aktinik Keratoz(Aktinic Keratoses-Pigmented Bowen's).....	29
3.2.3. Seboreik Keratoz(Seborrheic Keratosis-Pigmented Benign Keratoses)	29
3.2.4. Bazal Hücreli Karsinoma(Basal Cell Carcinoma)	29
3.2.5. Nevüs (Nevus).....	30
3.2.6. Damarsal(Vasküler) Deri Lezyonları(Vascular Skin Lesions)	30
3.2.7. Dermatofibroma	30
4. ÖNERİLEN YÖNTEMLER VE DEĞERLENDİRME.....	33
4.1. Evrişimsel Modellerin Performans Karşılaştırması.....	33
4.2. Efficient Modellerinin Optimize Edilmesi	38
4.3. Toplu Öğrenme Yöntemiyle Evrişimsel Sinir Ağlarının Birleştirilmesi	48

5. SONUÇ	53
KAYNAKLAR.....	55
ÖZGEÇMİŞ	



ÖZET

Vücut Benlerinin Klinik Görüntüleri Üzerinde Derin Öğrenme Tabanlı Hastalık Tahmini Sisteminin Geliştirilmesi

Zafer TOLAN

Yüksek Lisans Tezi

FIRAT ÜNİVERSİTESİ
Fen Bilimleri Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı
Ocak 2022, Sayfa: xi+56

Bilimsel olarak Melanomlara diğer cilt kanserlerine göre daha az rastlanır. Ancak cilt kanserine bağlı ölümlerin yaklaşık olarak %75'i melanomlara bağlı olarak gelişir. Kanser hücreleri çok hızlı büyüyen ve sistemik olarak yayılan bir davranış sergiler. Tüm kanser çeşitlerinde olduğu gibi cilt kanserinin tedavisi için de erken tanı büyük önem taşır. Dolayısıyla ciltte meydana gelen kanser belirtilerinin erken fark edilmesi ve doğru değerlendirilmesi son derece önemlidir. Vücut benleri üzerinde bir değişiklik fark edildiğinde renk ve şekil açısından takip edilmeli, ten renginde veya pembe tonlarında görülen kabartılar da dikkate alınmalıdır. Dermatologlar sıklıkla Melanom'u erken aşamalarında tanılamak ve ortadan kaldırmak için düzenli cilt muayenesi önermektedir.

Bu çalışmada derin öğrenme teknikleriyle, hastanın vücudundaki renkli lezyon görüntüleri üzerindeki örüntüler incelenerek cilt kanseri tespiti yapılmıştır. Tıpta bu konuda erken teşhisin önemine defalarca vurgu yapılırken çalışmamız sayesinde ben üzerinde kanser riski varsa erken teşhis edilip, tedavi aşamasına geçilmesi hedeflenmektedir. Çalışmamızda, dermoskopik görüntüler kullanılarak patolojik doğrulamaları yapılmış açık kaynak bir veritabanı ile evrimsel sinir ağları başarılı bir şekilde eğitilmiştir. Dermoskopik görüntü cihazlarında; vücuttaki benlerin haritası oluşturularak renkli görüntüleri kayıt altına alınmaktadır. Böylece bir sonraki kontrolde elde edilecek görüntüyle karşılaştırma şansı sağlanmaktadır. Bu tez çalışmasında; Uluslararası Cilt Görüntüleme Birliği tarafından açık kaynak olarak erişime sunulan bir veri setindeki renkli lezyon görüntüleri, şekil ve boyut farklılığı üzerinden 7 farklı hastalığa ayrıştırılmıştır. Yapılan bu çoklu-sınıflandırma uygulamasında yüksek oranda doğruluk elde edilmiştir.

Anahtar Kelimeler: Derin öğrenme, Evrimsel Sinir Ağları, Melanoma tespiti, cilt hastalıklarının teşhisi, vücut benleri

ABSTRACT

Development of A Deep Learning Based Disease Prediction System on Clinical Images of Body Moles

Zafer TOLAN

Master's Thesis

FIRAT UNIVERSITY
Graduate School of Natural and Applied Sciences

Department of Computer Engineering
January 2022, Page: xi+56

Scientifically, melanomas are less common than other skin cancers. However, approximately 75% of skin cancer-related deaths are due to melanomas. Cancer cells show a behavior that grows very quickly metastasis systemically. As with all types of cancer, early diagnosis is of great importance for the treatment of skin cancer. Therefore, it is extremely important to recognize and accurately evaluate the signs of cancer occurring on the skin early. Therefore, when a color change occurring on the skin or a change in the body me is noticed, it should be followed in terms of color and shape, blisters in skin tones or pink tones should also be considered. Dermatologists often recommend regular skin examination to identify and eliminate Melanoma in its early stages.

In this study, skin cancer detection was made by examining the patterns on the colored lesion images on the patient's body with deep learning techniques. While the importance of early diagnosis in medicine has been emphasized many times, our study aims to diagnose early and move to the treatment stage if there is a risk of cancer. In our study, convolutional neural networks were successfully trained with an open source database with pathological validations with dermoscopic images. With this method, the moles in the body are created and the point areas are determined. In dermoscopic image devices; color images are recorded by creating a map of the moles in the body. Thus, the chance to compare with the image to be obtained in the next check is provided. In this multiclassification application, a high degree of success was achieved.

Keywords: Deep Learning, Convolutional Neural Networks-CNN, Melanoma Detection, Skin Diseases Diagnosis, Lesion Images

ŞEKİLLER LİSTESİ

Sayfa

Şekil 2.1.	CNN ile Lezyon Resimlerinin 7 Farklı Hastalığa Ayrıştırılması	8
Şekil 2.2.	K-fold Çapraz Doğrulama Yinelemesinde Kullanılan Veri Çoğaltma Süreci	11
Şekil 2.3.	Melanoma Üzerinde Data Augmentation ile Sentetik Resimler Üretme	12
Şekil 2.4.	Vgg16 Modelinin Adaptasyon İşlemlerinden Önceki Ve Sonraki F1-Score Değerlerinin Karşılaştırılması	13
Şekil.2.5.	VGG16 Modelinin Local-test Datası İçin Yaptığı Tahminlere ait ROC Eğrilerinin Adaptasyon İşlemlerinden Önceki ve Sonraki Hali	13
Şekil 2.6.	Karmaşıklık Matrisi	17
Şekil 2.7.	ROC Eğrisinde Metrik Sınırlar	19
Şekil 2.8.	FPR ve TPR Eşik Değerleri	19
Şekil 3.1.	ISIC Arşivinin Web Portal Arayüzünden Alınan Bir Ekran Görüntüsü	23
Şekil 3.2.	HAM10000 verisetindeki hastalıklara ait resim sayılarının yüzdesel dağılımı	25
Şekil 3.3.	HAM10000 Verisetindeki Her Hastalık İçin Etiketli Örnek Sayısı	25
Şekil 3.4.	ISIC Arşivinden Alınan 7 Farklı Lezyon Görüntü Örneği	31
Şekil 4.1.	CNN modellerin Off-line Test Datasına Ait ROC Eğrilerinin ve AUC Değerlerinin Karşılaştırılması	34
Şekil 4.2.	CNN Modellerinin Off-line Test Datası Tahminlerine Ait Confusion Matrislerinin Karşılaştırılması	36
Şekil 4.3.	Cnn Modellerinin Balanced Accuracy Score Değerleri ile Performans Kıyaslaması	38
Şekil 4.4.	Train.csv ve Test.csv Dosya İçerikleri	40
Şekil 4.5.	Cinsiyete Göre Veri Dağılımı	41
Şekil 4.6.	İyi-kötü Huylu Melanoma Veri Dağılımı	41
Şekil 4.7.	Melanoma Nitelikleri	42
Şekil 4.8.	Öğrenme Hızı Çizelgesi	43
Şekil 4.9.	EfficientNet Model Mimarisi – Kaynak Makale:” EfficientNet:Rethinking Model Scaling for Convolutional Neural Networks”	44
Şekil 4.10.	Parametre ve ImageNet Doğruluk Karşılaştırması- Kaynak Makale :” EfficientNet:Rethinking Model Scaling for Convolutional Neural Networks”.	45
Şekil 4.11.	Depth ve Point Wise Evrişimlerinin Temel Gösterimi – Kaynak: EfficientNet:Rethinking Model Scaling for Convolutional Neural Networks”.	46
Şekil 4.12.	ROC Eğrisi	47
Şekil 4.13.	Yeni ağırlıklı Birleştirme Yöntemi	50
Şekil 4.14.	Birleştirme yöntemleri ve CNN Modelleri Kullanılarak Çizilen ROC Eğrilerinin Karşılaştırılması	51
Şekil 4.15.	Birleştirme Yöntemleri Kullanılarak Elde Edilen Karışıklık Matrislerinin Bireysel CNN Modelleri İle Karşılaştırılması	52

TABLÖLAR LİSTESİ

Sayfa

Tablo 2.1. Özgüllük Ve Duyarlılık Metriklerinin Sayısal Bir Örneği	16
Tablo 3.1. Ham10000 Veri Setindeki Tüm Hastalıkların Açıklamaları ve Görüntü Sayıları	24
Tablo 3.2. HAM10000 ve Diğer Veri Setlerinin Cilt Hastalıkları Görüntü Sayıları Açısından Karşılaştırılması	31
Tablo 4.1. Cnn Modellerinin Online-Real Test Datası Üzerinde Gösterdikleri Performans Ve Parametre Sayılarının Karşılaştırılması	38
Tablo 4.2. Doğrulama Verileri İçin CNN Modellerinin Elde Edilen F1-skor Değerleri	49
Tablo 4.3. CNN Modellerinin Normalleştirilmesi Sonucu F1-skor Ağırlık Değerleri	49
Tablo 4.4. HAM10000'de Her Hastalık İçin CNN Modellerinin Duyarlılık Değerleri	51

KISALTMALAR

YSA	: Yapay Sinir Ağları
CNN	: Convolutional Neural Network – Evrişimsel Sinir Ağı
PCA	: Principal Component Analysis – Temel Bileşenler Analizi
EDCNN	: Encoder-Decoder Convolutional Neural Network – Kodlayıcı Evrişimli Sinir Ağı
ReLU	: Rectified Linear Unit – Doğrultulmuş lineer birim
K-NN	: K-Nearest Neighbours – K-En Yakın Komşu Algoritması
ISIC	: International Skin Imaging Collaboration – Uluslararası Cilt Görüntüleme İşbirliği



1. GİRİŞ

Yaşadığımız pandemi sürecinde hekim karar destek sistemlerinin önemi daha çok kendini göstermiştir. İnsanlar, bu dönemde hastaneye gitmekten çekinmekte ve olabildiğince uzaktan yardım almaya çalışmaktadırlar. Bu çalışmanın konusu olan renkli lezyon görüntülerinin kategorize edilmesinde kullanılan bilgisayar destekli sistemler, uzaktan teşhis için ideal örneklerdir [1, 2]

Renkli lezyon görüntülerinin, derin öğrenme tabanlı sınıflandırma uygulamalarında elde edilen başarı dermatologların dahi önüne geçmiş durumdadır. Almanya’da gerçekleştirilen çalışmada; dermatologların manuel olarak bir lezyona [0-1] aralığında melanoma olma olasılığı teşhisi koymaları istenmiştir. Buna karşın, klinik resimler ile eğitilmiş bir evrişimsel sinir ağının da aynı resim ile ilgili tahmini kaydedilmiştir [3]. Farklı enstitülerden katılan dermatolog sayısının oldukça fazla olduğu bu çalışmalarda, makine tahminlerinin katılımcılarının %85’inden daha doğru tahmin yapabiliyor olması çok dikkat çekicidir. Bu durum, evrişimsel sinir ağlarında gelinen başarı noktasının somut bir örneğidir [4, 5].

En çok rastlanılan kanser türü olan deri kanserlerinin erken teşhisinde, Convolution Neural Network(CNN) tabanlı görüntü işleme uygulamaları ön plandadır. Derin öğrenmenin her alanında olduğu gibi görüntü işleyen konvolüsyon modellerinde de her geçen gün popüler ve iddialı yeni mimariler ortaya çıkmaktadır [6, 7, 8]. Bu çalışmada performans karşılaştırması yapılan modellerin en yaşlısı bile beş yaşını doldurmamıştır. Bu çalışmada ele alınan popüler CNN modellerinin dermoskopik veri setleri üzerindeki sınıflandırma kabiliyetleri uygulamalı olarak kıyaslanmıştır. Hangi CNN modelinin ne kadar parametre ile ne denli başarı gösterdiği ortaya konulmuştur.

Dermatoskopi, iyi ve kötü huylu cilt kanserlerini belirlemede yaygın olarak kullanılan bir tanı tekniğidir. Çıplak gözle yapılan incelemeye kıyasla deri hastalıklarını otomatik olarak teşhis etmek için kullanılır. Dermaskopik görüntülerin alınmasında kullanılan çeşitli donanımsal aygıtlar mevcuttur. Derinin mikroskopla incelenmesi yöntemi olarak da tanımlanabilen dermoskopi, cilt görüntülerinin bilgisayar destekli analiz edilmesi yöntemidir [9]. Bu yöntemle vücuttaki benlerin haritası oluşturularak her noktanın lokalizasyonunda belirlenmektedir. Ardından her bir ben için dermaskopik görüntü alınıp kaydedilmektedir. Böylece bir sonraki kontrolde elde edilecek yeni görüntünün, uzman hekim tarafından önceki halleri ile karşılaştırılabilmesi mümkün olmaktadır.

Melanoma, cilt kanserleri içerisinde yer alan kanser türüdür. Cilt kanseri o kadar sıklıkla görülmektedir ki, Türkiye ve dünyada kanser sınıflaması yapılırken cilt kanseri ilk sıralardadır. Cilt kanseri, her yıl Amerika Birleşik Devletleri’nde 5.000.000’den fazla yeni tanı konulan vakayla büyük bir halk sağlığı sorunudur. Melanoma, cilt kanseri ölümlerinin ezici bir

çoğunluğundan sorumlu olan en ölümcül cilt kanseri şeklidir [10]. 2015 yılında küresel melanoma insidansının yaklaşık 60.000 ölüm ve 350.000'den fazla vaka olduğu tahmin edilmektedir. Mortalite önemli olmasına rağmen, erken saptandığında melanoma sağ kalımı %95'i geçmektedir. Deriye rengini veren melanin adlı renk pigmentleri, melanosit denilen cilt hücreleri tarafından üretilir. Bu hücrelerin kontrolsüz bir şekilde bölünüp çoğalması sonucunda meydana gelen cilt kanseri melanom veya melanoma olarak tanımlanır.

Melanomlara diğer cilt kanserlerine göre daha az rastlanır ancak cilt kanserine bağlı ölümlerin yaklaşık olarak %75'i melanomlara bağlı olarak gelişir. Kanser hücreleri çok hızlı büyüyen ve sistematik olarak yayılan bir davranış sergiler. Tüm kanser çeşitlerinde olduğu gibi cilt kanserinin tedavisi için de erken tanı büyük önem taşır. Dolayısıyla ciltte meydana gelen kanser belirtilerinin erken fark edilmesi ve doğru değerlendirilmesi son derece önemlidir. Dermatoloji uzmanları normal şartlar altında vücudunda çok sayıda beni olanların ayna karşısında belirli zamanlarda benlerini kontrol etmelerinin ve yıllık olarak dermatolojik muayeneden geçmelerinin son derece önemli olduğuna dikkat çekmektedirler. Çünkü en ölümcül cilt kanseri türü olan Melanoma geç teşhis edildiğinde tedavi şansı kalmamaktadır.

Bu çalışmada olduğu gibi; deri hastalıklarının teşhisinde yapay zekâ teknikleri kullanmak isteyen araştırmacıların ihtiyaç duyduğu dijital resimlerden oluşan en büyük uluslararası veri seti, ISIC(International Skin Imaging Collaboration) topluluğu tarafından açık kaynak olarak sunulmaktadır. Gerçekleştirdiğimiz lezyon sınıflandırma uygulamamız için dermatoskopi cihazlarından alınan renkli deri görüntülerinden oluşan ISIC arşivindeki HAM10000(Human Against Machine with 10.000 training images) veri setinden faydalanılmıştır. Bu çalışmada, veri seti olarak ISIC tarafından 2018 yılında düzenlenen uluslararası yarışmada katılımcılara modellerini eğitmeleri için sunulan 10.015 adet renkli lezyon resimlerinden oluşan HAM10000(Human Against Machine with 10.000 training images) veri seti kullanılmıştır [11]. "ISIC Arşivi" cilt görüntülerinin dijital resim formatında halka açık olarak sunulduğu en büyük global depo alanıdır. Bu veri setinde en yaygın olarak bilinen deri kanseri olan Melanoma ile birlikte yedi farklı cilt hastalığının da erken teşhisi için kullanılabilecek ön işleme yapılan dermoskopik resimler bulunmaktadır.

HAM10000 veri setinde olduğu gibi, gerçek dünya uygulamalarında veri setleri, genellikle dengeli dağılım göstermez. Veri setleri içerisindeki azınlık sınıfı, nadir rastlanan örneklerin oluşturduğu bir alt küme olduğundan bu kümedeki eleman sayısını arttırarak denge kurmaya çalışmak çok zordur. Çalışmamızın amacı olmamakla birlikte az da olsa veri çeşitlendirme tekniği ile HAM10000 veri setindeki minör sınıfların örnek sayısı kısmen arttırılmaya çalışılmıştır. Uygulamamızdaki temel amaç; popüler CNN modellerinin dengeli dağılım göstermeyen dermoskopik resim veri setleri üzerindeki doğal kabiliyetlerini analiz etmektir. Sentetik resim üretme teknikleri ile her bir modelin performans iyileştirmesi gerçekleştirilebilir [12, 13]. Tez

çalışmasında, HAM10000 veri setindeki minör sınıflara ait örneklerin popüler CNN modeller tarafından ne oranda entegre edilebildiği incelenmiştir. Çünkü tıbbi bir testin en kötü senaryosu, gerçekten hasta olan birine negatif olduğunu söylemektir. Dengeli dağılım göstermeyen veri setlerinde bu durumun gerçekleşme olasılığı çok yüksektir. Bu problemin önüne geçmek için temel karakteristiklerine değinilen CNN modellerinden yararlanılmıştır. Popüler CNN modellerinin dermoskopik görüntülere ne kadar yatkın olduğu uygulamalı olarak irdelenmiştir.

Szegedy ve arkadaşlarının 2014 yılındaki ILSVRC yarışmasında takım ismi olarak kullandıkları GoogleNet, Inception modül temelli mimarilerin atasıdır [14]. Network-in-Network düşüncesinin CNN katmanlarında uygulanması temel fikri ile doğmuş bir conv.net ailesidir. Bu çalışmada da Inception ailesinden V3 ve Chollet tarafından daha da geliştirilen sürümü olan Xception mimarilerinin performans denemeleri yapılmıştır [15, 16]. Inception network ailesi, hem Imagenet hem de JFT gibi devasa ölçekli resim veri tabanları üzerinde en iyi performansı gösterenlerdendir [17]. Her biri küçük bir ağ olan modüllerin paralel dallar şeklinde birleştirilmesi ile oluşan Inception modülünün çıkışı, dalların çıkışlarının birleştirilmesi şeklindedir. Her dalın başlangıcında 1x1 devamında ise 3x3 boyutunda Conv2D işlemleri uygulanmaktadır. Bu yapı sayesinde birlikte öğrenilmesi zor veya imkânsız olan uzaysal ve kanal bazlı niteliklerin ayrı ayrı keşfedilebilmesi mümkün hale gelmektedir.

Çok derin networklerin eğitimi oldukça zor bir süreçtir. Çünkü çok derinlere inildikçe eğim değerleri, kül olup kaybolan ateş misali sifıra yakınsamaktadır. Residual block diye isimlendirilen ve içerisinde skip-connections veya başka bir deyişle short-cuts barındıran katman mimarisi, eğri değerinin kaybolmadan derinliklere iletilmesi sağlanmaktadır [18]. Bu sayede çalışmamızda da kullandığımız ResNet152 gibi çok derin networkler eğitilmiştir. İsminden de tahmin edileceği üzere 152 katmandan oluşan bu model, eğri değeri ile başarılı bir şekilde eğitilmiştir [19]. CNN modellerin derinliği başka bir deyişle kapasitesi arttıkça ezber yapma ihtimali de artmaktadır. ResNet152 gibi çok derin networklerin inşası ile güdülen amaç, ezber yeteneğinin artırılması değil yeni örüntü keşiflerinin eklenen katmanlar ile ortaya çıkarılmasıdır. ResNet içerisindeki residual blocklar sayesinde overfit ihtimalinin önüne geçilmektedir. Residual bloklar içerisindeki skip-connectionlar sayesinde, ağın çıkış katmanındaki hatanın eğim değerinin çok derinlere kaybolmadan ulaşması sağlanmaktadır. Ayrıca herhangi bir ağ katmanında oluşabilecek darboğazın önüne geçilerek katmandan önceki öğrenilen özelliklerin bu darboğazın ötesine de aktarılması sağlanmaktadır.

Derin öğrenme teknikleriyle vücuttaki benler üzerinde cilt kanseri tespiti çalışması, bilgisayar sistemleriyle belli teknikler kullanılarak elimizdeki veri setleri ile karşılaştırılıp, herhangi bir bende cilt kanseri şüphesi olup olmadığının başarı yüzdesini ölçmek için kullanılan bir çalışmadır. Bu çalışmada tıpta tedavi konusunda temel ilke olan erken teşhis ön planda yer almaktadır. Öncelikle hastalık ihtimali taşıyan bir insanın vücut beni fotoğraflanmıştır. Daha

sonra derin öğrenme teknikleriyle elimizde bulunan veri seti kullanılarak vücut beni için şekil ve renk bakımından inceleme yapılmıştır. Karşılaştırmalar sonucunda varsa hastalık tespiti sağlanmaya çalışılmıştır. Hastanın vücudunda renk pigmentleri sonucu meydana gelen görüntünün ben veya kanser hücresi olup olmadığı tespit edilmiştir. Cilt kanseri belirtilerini gösteren ve melanoma şüphesi taşıyan lezyonlar önce fiziksel muayene ile detaylı şekilde değerlendirilir. Daha sonra doktor tarafından tercih edilen tarama yöntemine göre dermoskopik veya histopatolojik inceleme yapılır. Çalışma ile hastalık şüphesi taşıyan vücut ben görüntüsü alınarak sistemde eğitilmiş veriler ile karşılaştırılıp cilt kanseri tespiti yapılmıştır.

Inception modüllerindeki paralel network yapılarını, residual bloklardaki short-cut bağlantılar ile hızlı eğitilebilir hale getirebilme düşüncesi InceptionResNet mimarisini doğurmuştur. Bu hibrit yapı, aynı zamanda InceptionV4 olarak da isimlendirilmektedir. Klasik inception ağlarına göre eğitim süreci önemli oranda hızlanmıştır. Ayrıca inception ve residual blokların ensemble edilmesi ile oluşan mimarinin accuracy değeri artmaktadır. Res-Net ailesi derine inmeye çalışırken, inception ailesi yatayda paralellik sayısını arttırmaya çalışmaktadır. Çalışmamızda performans değerlendirmesi yaptığımız InceptionResNetV2 hibrit mimarisi ise hem derine hem de enine optimum hale getirilmiş mimari türevidir [20].

CNN tabanlı resim sınıflandırma uygulaması geliştirmek artık neredeyse bir mühendislik alanı haline gelmiştir. NASNet mimarisinin doğmasındaki temel düşünce bu tasarım uzmanlığı gereksinimini ortadan nasıl kaldırılabilceği düşüncesini doğurmuştur. Veri setine bakarak kendini optimum hale getirebilmeyi, bir ağa nasıl öğretebiliriz sorusu, özellikle büyük boyutlu gerçek dünya problemlerinde çok maliyetli cevaplara gebe dir. Network Architecture Search(Ağ mimarisi arama) düşüncesi, öncelikle küçük boyutlu veri seti üzerinde optimum bir temel blok inşa eder. Daha sonra bu temel yapıtaşının onlarcası kendi özgün parametreleri ile bir araya getirilerek büyük veri setine transfer edilir. Ancak temel yapıtaşının ve büyük binanın inşası bir arama uzayında global minimum bulma problemine benzemekte ve çok pahalıya patlamaktadır [21].

Son çalışmalar, CNN ağlarında giriş ve çıkışa yakın katmanlar arasında daha kısa bağlantılar içeriyorsa, eğitimin daha verimli olabileceğini göstermiştir. Dense Convolution Network(DenseNet), klasik feed-forward bağlantı şekline sahiptir ancak farklı olarak L tane layer için $L(L+1)/2$ tane ağırlık içermektedir [22]. Her katman kendinden sonraki katmanlara kısa devreler ile bağlanmıştır. Başka bir deyişle her katmanın girişi kendinden önceki tüm katmanların çıkışlarının bir araya gelmesi şeklinde oluşturulur. DenseNet içerisinde üç veya dört tane Dense Block mevcuttur. Bu bloklar arasındaki geçiş yani boyut küçültme transition layer ile gerçekleştirilir. Transition layer denilen geçiş conv+pooling işlemidir.

CNN tabanlı akıllı resim işleme uygulamaları, sadece masaüstü veya hosting sistemlerde değil cebimizde taşıyabildiğimiz akıllı telefonlar gibi kısıtlı kaynağa sahip donanımlarda da

ihtiyaç duyulan çözümlerdir [23, 24]. Bu kısıtlı kaynaklar üzerinde kořabilecek parametre sayısı az yani diđer bir deyiřle daha seyretilmiş mimarilere ihtiyaç MobileNet ailesini doğurmuřtur. Bu ailenin temel yapıtařı depthwise seperable convolution bloklardır. Depthwise seperable convolution blođu, klasik konvolüsyon işleminden farklı olarak her kanal için bir adet deđer hesaplar ve bunları arka arkaya yapıştırır. Daha sonra elde edilen bu çıktılarına 1x1 pointwise konvolüsyon uygulayarak modelin boyutunu çok düşük hale getirmiş olur. 1x1 pointwise konvolüsyon ise filtrelerdeki aynı konumda bulunan deđerlerin hepsinin alınıp tek bir piksel haline dönüřtürölmesi işlemidir. MobileNet ailesinde latency ve accuracy gibi ekstra 2 tane hiper parametre mevcuttur. Bu parametreler, tasarımcıya uygulaması için doğru boyutları belirlemesini sađlar.

NASNet ailesinden bir çözümler maliyet aısından yüksek olsada modelin mimarisini veri setine göre optimum hale getirmeye çalıřma fikri çok önemli bir gelişmeyi doğurmuřtur. EfficientNet olarak isimlendirilen bu yeni aile NASNet'ten farklı olarak mimariyi maliyetli bir şekilde optimum hale getirmek yerine, sadece bir tane katsayı ile mimariyi ölçekleme yoluna gitmiştir [25]. Örneđin; ölçeđin geçerli deđerı ikiye katlanınca, mimarideki derinlik, katman çıkıřlarındaki ara çıktıların geniřliđi ve çözünürlüğü de ikiye katlanmaktadır. Birleşik ölçekleme de denilebilecek bu yeni strateji çok basit ve sade olmasına karřın oldukça etkili sonuçlar vermektedir. Bu çalıřma içinde en etkili sonuçların EfficientNetB5 mimarisi ile elde edilmesi bu stratejinin verimliliđinin yeni bir örneđidir.

Derin öğrenme yöntemleri, tıbbi görüntüleme hastalıklarının teřhis ve tedavisinde giderek daha önemli bir seviyeye ulaşmıştır. Son yıllarda görüntü işleme teknikleri, birçok nedenden dolayı tıbbi görüntülerin çeřitli şekillerde analiz edilmesinde önemli bir faktör haline gelmiştir. Çođu dermoskopik görüntü, dijital kamerayla çekilenlere göre daha düşük bir gürültüye sahiptir. Benzer bir durum nükleer tıp görüntülerinde de görölmektedir. Bu nedenlerden ötürü, tıbbi görüntünün gürültüsünü azaltmak için kullanılan görüntü işleme teknikleri bu durumlarda vazgeçilmez olmaktadır [26].

Medikal alanda görüntü işleme tekniklerinin yanı sıra derin öğrenme yöntemleri kullanılmaya başlanması hem doğruluk hem de hız kazandırmıştır. Medikal alanda derin öğrenme yöntemlerinin kullanım alanları; görüntü sınıflandırma, nesne algılama, segmentasyon ve kayıttır. Tıptaki uygulama alanları genel olarak: nöroloji, retinal, pulmoner, dijital patoloji, meme, kalp, karın ve kas-iskelet sistemidir [27].

A.Kalinovsky ve arkadaşları [28], derin öğrenme yöntemleri ve Encoder-Decoder Konvolüsyonel Sinir Ağları (EDCNN) kullanarak X-Ray göđüs görüntülerinde (Göđüs Radyografileri) akciđerlerin segmentasyonu üzerine arařtırma yapmıştır. Elde edilen doğruluđun, Dice skorunu kullanarak manuel segmentasyon ile karřılařtırılması, ortalama doğruluđun 0.962,

minimum ve maksimum Dice skorlarının sırasıyla 0.926, 0.974 ve 0.008 standart sapması ile elde edildiğini ortaya çıkarmıştır.

Lo ve arkadaşları [29], göğüs röntgeni görüntülerinde akciğer nodüllerini saptamak için bir yöntem önermişlerdir. Bir bölgede bir akciğer nodülü olup olmadığını tespit etmek için, 55 X-ışını görüntüsü kullanılarak eğitilmiş iki gizli katmanı olan bir CNN kullanmışlardır. Sonuç olarak, bu görev için gelişmiş sınıflandırma oranını elde etmişlerdir.

Pratt ve arkadaşları [30], gözler üzerinde yapılan bu çalışmada göz problemini tespit etmek için 80.000 civarında dijital göz fotoğrafını eğitmiştir. Kullandıkları yöntem ile 5.000 doğrulama görüntüsü üzerinde % 95 hassasiyet ve % 75 doğruluk elde etmişlerdir.

N. Codella ve arkadaşları [31], derin öğrenme yöntemleri ve bununla birlikte klasik makine öğrenmesi yöntemleri ile dermoskopik görüntülerde melanom sınıflandırması gerçekleştirmiş bunu yaparken nitelik çıkarmasından faydalanmış ve %93 doğruluk değerine ulaşmışlardır.

Andre Esteva ve arkadaşları [32], deri lezyonlarının sınıflandırması üzerine yaptıkları çalışmalarında tek bir CNN kullanmışlardır. Doğrudan pikselleri ve hastalık etiketlerini girdi olarak almışlardır. 12932 dermoskopik görüntü ve 2.032 farklı hastalıktan oluşan bir veri seti kullanarak bir CNN eğitilmiştir. Bu çalışmada doğruluk değeri yaklaşık olarak %90'dır.

Bu noktaya kadar literatürdeki CNN tabanlı uygulamalardan bahsedilmiştir. İkinci bölümde evrimsel sinir ağları ile vücut benlerinin analizi için kurulan yapı anlatılmıştır. Daha sonraki bölümde ise kullanılan veri setinin içeriği ve nitelikleri verilmiştir. Tez çalışmasının 4. bölümünde önerilen yöntemler ve elde edilen sonuçlar şekillerle desteklenmiştir. Tezin son bölümünde, elde edilen sonuçlar ile varılan gerçekler ve yapılması düşünülen sonraki çalışmalara değinilmiştir.

2. EVRİŞİMSEL SİNİR AĞLARI İLE VÜCUT BENLERİNİN ANALİZİ

Derin öğrenme, yapı olarak insan gibi hareket eden ve karar veren sistemler elde etmek için kullanılan makine öğrenmesi algoritmalarının genel adıdır. Derin öğrenme insan beyin hücrelerinin birbiriyle olan etkileşimine benzer bir yapı modelleyip öğrenme ve karar verme işlemlerini gerçekleştirmeyi amaçlar. Derin öğrenmenin amacı, yüksek seviye özellikler ya da temsillerin alt seviyedeki temsillerden tanımlandığı verilerin aşamalı olarak özelliklerini hesaplamaktır. Bir görüntü için temsil denilince akla piksellerin yoğunluk değerleri veya kenar kümeleri gelebilir. Çoklu soyutlama seviyelerinde özelliklerin otomatik olarak öğrenilmesi, bir sistemin, insan eliyle çıkarılan özelliklere tamamen bağlı olmadan girişi doğrudan çıktıyla eşleştiren karmaşık işlevleri öğrenmesini sağlar. Derin öğrenme yöntemleri ailesi, sinir ağları, aşamalı olasılık modelleri, eğitici ve eğitici öğrenme algoritmalarını kapsamaktadır. Ayrıca giderek daha zengin hale gelmiştir. Derin öğrenme sinir ağlarının araştırma alanları, yapay zeka, grafik modelleme, optimizasyon, örüntü tanıma ve sinyal işleme arasındaki kesişme noktalarında [33]. Derin bir mimari kullanmanın avantajı, farklı bir soyutlama seviyesinin çok sayıda özellikten küçük bir alt kümeye odaklanmasıdır. Öğrenilecek bilgiler tek bir nöron tabakasında bulunmamakla birlikte, çok katmanlı olarak dağıtılmaktadır. Böyle dağıtılmış bir temsil, derin öğrenme ağlarının öğrenme için daha güçlü bir kapasiteye sahip olmasını sağlar. Geleneksel makine öğrenme yöntemlerine kıyasla çok daha iyi genellemeler üretebilir. Ayrıca, çok seviyeli ve verilerin dağıtılmış bir temsili üzerine kurulu bir mimari, derin öğrenme ağlarının farklı problem alanlarında paylaşılabilen ara temsilleri öğrenmesine olanak tanır. Bu, ara tasvirler olarak öğrendiği bilginin ortak bir havuzdaki alt düzey ara özelliklerin birleştirilmesiyle, yeni yüksek seviye özelliklerin öğrenilebileceği tekrar kullanılabilir olduğu anlamına gelir [34].

Derin öğrenme kullanmanın bazı avantajlarını şöyle sıralanabilir:

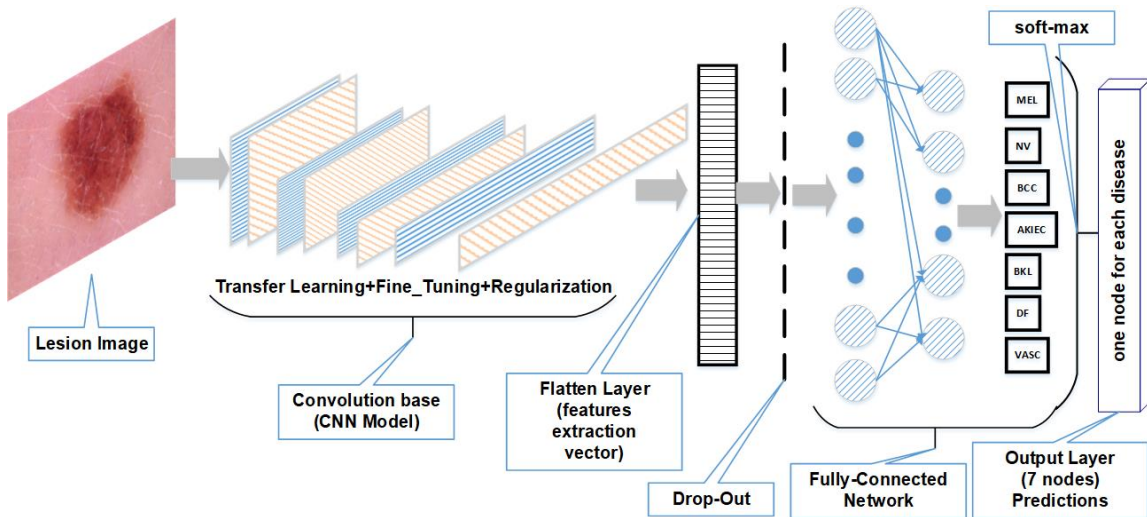
- Karmaşık, değişken fonksiyonları öğrenebilmek,
- Düşük düzeyde insan girişi ile düşük, orta ve yüksek seviye soyutlamaları analiz etmek,
- Çok geniş bir örnek kümesini işleyebilmek,
- Çoğunlukla etiketlenmemiş verileri değerlendirebilmektir.

Derin öğrenme teknikleriyle vücuttaki benler üzerinde cilt kanseri tespiti çalışması, görüntülerin bilgisayar sistemleriyle belli teknikler kullanılarak elimizdeki veri setleri ile karşılaştırıp, herhangi bir bende cilt kanseri şüphesi olup olmadığının başarı yüzdesini ölçmek için kullanılan bir çalışmadır. Bu çalışmada tıpta tedavi konusunda temel ilke olan erken teşhis ön planda yer almaktadır. Öncelikle hastalık ihtimali taşıyan bir insanın vücut beni fotoğraflanmıştır. Daha sonra derin öğrenme teknikleriyle elimizde bulunan veri seti kullanılarak vücut beni için

şekil ve renk bakımından inceleme yapılmıştır Karşılaştırmalar sonucunda var olan hastalık tespiti sağlanmıştır. Hastanın vücudunda renk pigmentleri sonucu meydana gelen görüntünün ben veya kanser hücresi olup olmadığını belirler. Cilt kanseri belirtilerini gösteren ve melanoma şüphesi taşıyan lezyonlar önce fiziksel muayene ile detaylı şekilde değerlendirilmiştir. Daha sonra doktor tarafından tercih edilen tarama yöntemine göre dermoskopik veya histopatolojik inceleme yapılır. Çalışma ile hastalık şüphesi taşıyan ben görüntüsü alınarak sistemde eğitilmiş veriler ile karşılaştırılıp cilt kanseri tespiti yapılmıştır.

2.1 Önerilen Evrişimsel Mimari

Evrişimsel Sinir Ağları, temel olarak görüntüleri sınıflandırmak için kullanılan çok katmanlı sinir ağları türüdür. CNN'ler daha çok görüntü problemlerinde sıklıkla kullanılan derin öğrenme mimarisidir. Konvolüsyonel sinir ağı mimarisi, insan beyindeki nöronların kendi aralarındaki haberleşme ve iletişim modeline benzemektedir. Konvolüsyonel Sinir Ağları oluşturulurken beyin fiziksel yapısı örnek alınmıştır. Beyindeki nöronlar, uyarıcılara sadece görsel alanın belirli bir bölgesinde yanıt verir. Öğrenilebilir ağırlıkları ve bias değerleri olan nöronlardan oluşurlar. Her bir nöron bazı girdiler alır, bir nokta ürünü yapar ve isteğe bağlı olarak onu doğrusal olmayan bir şekilde izler [35]. Şekil 2.1' de yer alan görüntüde bir Evrişimsel sinir ağının sınıflandırılmasını sağlayan, hangi katmanlardan oluştuğunu gösteren sinir ağı modeli yer almaktadır.



Şekil 2.1. CNN ile Lezyon Resimlerinin 7 Farklı Hastalığa Ayırılması.

Evrişimsel ağlar, sinir ağlarından katman yapıları ve işlevleri açısından farklı özellikler göstermektedir. Bilinen sinir ağlarının katmanları tek boyutludur. Ayrıca bu katmandaki nöronlar birbirine düzenli bir yapıda bağlıdır. Bir CNN, klasik neural networks yapısına benzer şekilde nöronlardan oluşur. Öğrenme için ağırlıklar ile bias değerine sahiptir. Her nöron girişleri alır, birleştirir ve genellikle non-lineer bir fonksiyon ile çıkış üretir. CNN uygulamaları girişleri görüntü olarak varsayar ve özellikleri mimariye kodlamamızı sağlar. Öte yandan, CNN'lerdeki katmanlar ise genişlik, yükseklik ve derinlik parametreleri olarak üç boyutludur.

CNN biyoloji ve bilgisayar bilimlerinin karışımı bir yapıya sahiptir. Görüntü tanıma uygulamalarında sıklıkla kullanılan çok etkili ağıdır. Bir konvolüsyonel sinir ağı, giriş görüntüsünü alabilen, görüntüdeki çeşitli yönler / nesnelere önem veren (öğrenilebilir ağırlıklar ve bias değerleri) ve birini diğerlerinden ayırabilen derin öğrenme algoritmasıdır. ConvNet'te gereken ön işleme, diğer sınıflandırma algoritmalarına kıyasla çok daha düşüktür. İlkel yöntemlerde, filtreler otomatik olmayıp manuel bir yapıda bulunsun da uygun görülen yeteri düzeyde bir eğitim ile ConvNets bu özellikleri öğrenebilir [36].

HAM10000 veri seti karşısında performans karşılaştırması yapılan bütün CNN modellerine aşağıda belirtilen adaptasyon aşamaları uygulanmıştır. Ele alınan her bir modelden olası en iyi performans yakalanmaya çalışılmıştır. Başarısı test edilen tüm CNN modellerin orijinal top_classifier kısımları elemine edilmiştir. En son konvolüsyon katmanından sonraki feature extraction diye isimlendirilen network çıkışı, flatten uygulanarak satır olarak bir vektör haline getirilmiştir. CNN modellerden elde edilen bu özellik vektörü 3 katmanlı bir fully-connected Neural Network sınıflandırıcıya Şekil 2.1'de gösterildiği gibi giriş olarak uygulanmıştır. Bu sinir ağının her biri bir hastalığa karşılık gelen 7 tane çıkışı bulunmaktadır. Dolayısıyla lezyon resmi Şekil 2.1'de gösterilen adapted CNN modellerinin girişine uygulanınca, belirtilen 7 tane çıkışta ilgili hastalıklara ait olma olasılıkları [0-1] aralığında elde edilmiştir.

Adım-1(Transfer Learning): 14 milyondan fazla resmin 21 binden fazla sınıfa ayrıştırma işleminden elde edilen filtre değerleri ile işe başlamak yani Imagenet veri seti ile eğitilmiş ağırlık değerlerini transfer etmek birçok çalışmada olduğu gibi bizim de temel başlangıç noktamız olmuştur. CNN modellerin ilk baştaki convolution katmanlarında algılanmaya çalışılan bileşenler, son katmanlara nazaran çok daha basit ve generic özelliklere sahiptir.

Adım-2(Fine-Tuning): CNN modellerin sadece en derin katmanlarındaki konvolüsyon filtreleri eğitim işlemine tabi tutulmuştur. En baştaki katmanlarda kullanılan filtreler, imagenet veri setinden transfer edilip dondurulmuştur. Hem eğitim süreci hızlandırılmış hem de overfit riski azaltılmıştır. Ne kadar çok parametre eğitilirse, overfit riski o denli artmaktadır. CNN modelinin son katmanındaki filtreler ve akabindeki fully-connected network fine-tuning edilerek, HAM10000 verisetindeki örnekler öğretilmeye çalışılmıştır.

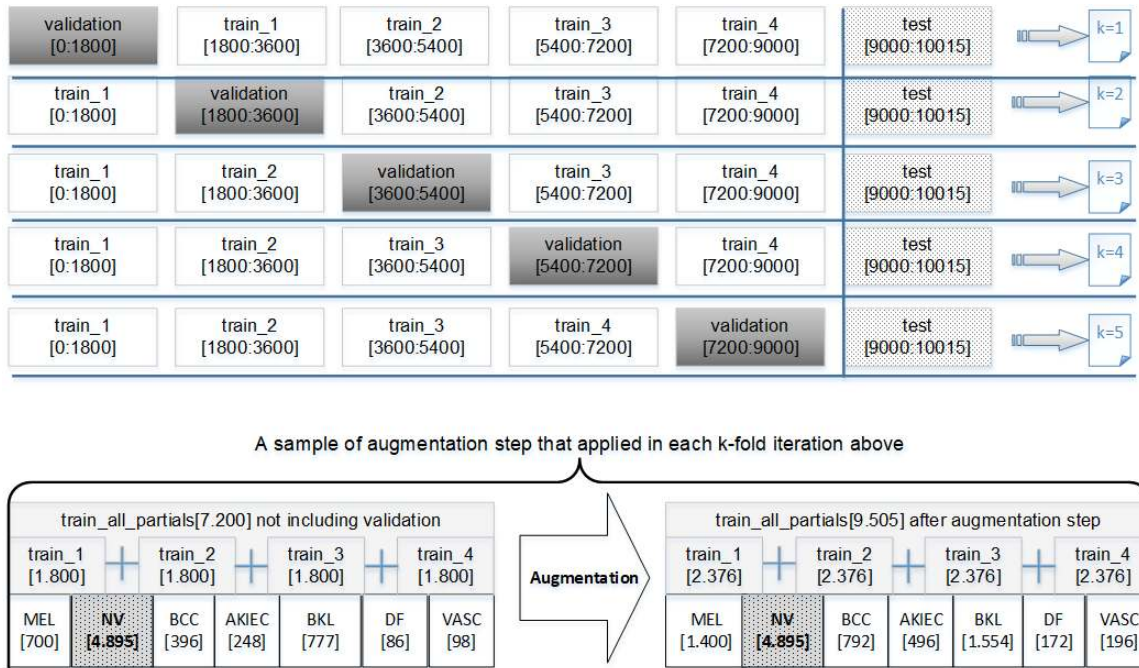
Adım-3(Regularization): CNN modellerin overfit yapmasını engellemek adına; drop-out, L2-regularization (weight-decay) ve kapasite indirgeme seçenekleri de göz önüne alınmıştır. Modellerin en son convolution katman çıkışları flatten edildikten sonra 50% oranında drop-out işlemine tabi tutulmuştur. Overfit problemi için ikinci bir adım olarak CNN modellerin feature extraction kısmına ve Fully-Connected sınıflandırıcının ağırlık güncelleme adımlarına weight decay(0.001) uygulanarak L2-regularization işlemi gerçekleştirilmiştir. Son olarak geliştirilen modellerin overfit yapmaması için Top-Classifier kısmındaki katman ve nöron sayısı da indirgenerek kapasite küçültülmüştür.

Adım-4(k-fold cross validation): Klasik hold-out doğrulama yerine, eğitim setinin beş farklı kombinasyonu ile CNN modeller eğitilerek ortalamaları alınmıştır. Bu sayede eğitim setinin train-validation ayrıştırmasında yapılabilecek olası bir hatanın önüne geçilmiştir. HAM10000 verisetinde bulunan 10.015 tane eğitim amaçlı lezyon resimlerinden 1.015 tanesi Şekil 2.2’de gösterildiği sadece off-line(local) test amacı ile kullanılmıştır. Local test kümesinin, eğitim ve doğrulama kümelerinden herhangi bir örnek içerme olasılığı bulunmamaktadır. Aynı şekilde her k. iterasyonda seçilen doğrulama alt kümesinin de eğitim alt kümelerinden aynı resmi içerme olasılığı bulunmamaktadır. Bilgi sızıntısı olmaması adına her iterasyon adımında 9.000 adet olan eğitim resmi 5 eşit parçaya ayrıştırılmakta ve bu parçalardan bir tanesi yani 1.800 tane resim doğrulama için rezerve edilmektedir. Geri kalan 4 parçanın oluşturduğu toplam 7.200 adetlik resim ise modelin eğitimi için bir araya getirilip kullanılmaktadır. Her k. iterasyonda seçilen doğrulama alt kümesi için herhangi bir veri çoğaltma işlemi yapılmamaktadır. Çünkü veri çoğaltma işlemi hem eğitim hem de doğrulama için kullanılırsa bilgi sızıntısı meydana gelmektedir. Bu neden ile oversampling işlemi her iterasyonda eğitim ve doğrulama kümeleri ayrıştırıldıktan sonra yapılmıştır. Bilgi sızıntısına dikkat edilmeden k-fold ve data augmentation işlemi birlikte gerçekleştirilirse, validation score değerini %100’e götürebilecek overfitting problemi gözlemlenmiştir.

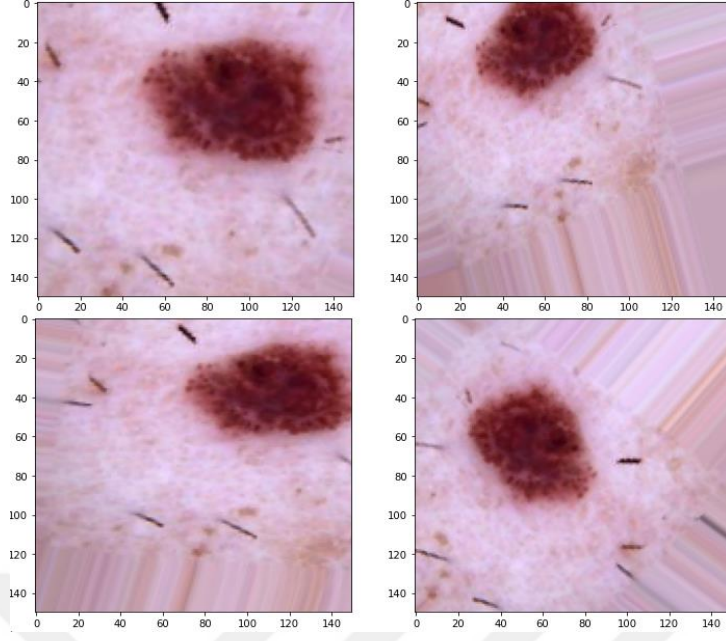
Adım-5(Data Augmentation): HAM10000 veri setindeki dengesizliği kısmen de olsa giderebilmek için, Şekil 2.2’de gösterildiği gibi her k. iterasyonda, eğitim kümelerindeki minör sınıfların pozitif örnekleri sentetik olarak çoğaltılarak ikiye katlanmıştır. Sadece NV majör sınıfına ait örnekler için oversampling yapılmamıştır. Şekil 2.2’de verilen sayısal bir örnekte bu çoğaltma detaylı olarak açıklanmıştır. Örneğin herhangi bir k.iterasyonda; 9.000 tane olan off-line eğitim resmi öncelikle 1.800 tane doğrulama ve 7.200 tane eğitim şeklinde ayrıştırılmaktadır. Daha sonra seçilen 7.200 tane eğitim resmi içerisindeki NV hastalık resimleri hariç geri kalan tüm lezyon resimlerinin sayısı, geometrik transformasyon teknikleri ile iki katına çıkarılmıştır. Şekil 2.2’deki sayısal örnekte MEL sınıfına ait 700 adet resim oversampling yapılarak 1.400 adet görüntüye çıkarılmıştır. Aynı şekilde diğer tüm minör sınıfların örnek sayıları iki katına

çıkarılarak toplam eğitim kümesi, 7.200 tane orijinal resimden 9.505 tane orijinal+sentetik resim haline getirilmiştir.

Data Augmentation işlemi sırasında; resimlerin rastgele döndürülmesi, resimlerin yatayda ve dikeyde kaydırma oranları, burkma işlemi, görüntü yakınlaştırma işlemi, dikeyde resmi döndürme ve oluşturulan yeni resimler üzerindeki olası boşlukların doldurulması seçenekleri aktifleştirilerek Şekil 2.2.'de görüldüğü gibi azınlık kümesine ait eğitim örnek sayıları çoğaltılmıştır. Şekil 2.2.'nin sol üst köşesindeki orijinal Melanoma resmine, yukarıda bahsedilen geometrik işlemlerin rastgele kombinasyonları uygulanarak Şekil 2.3'teki diğer 3 tane yeni resim elde edilmiştir.



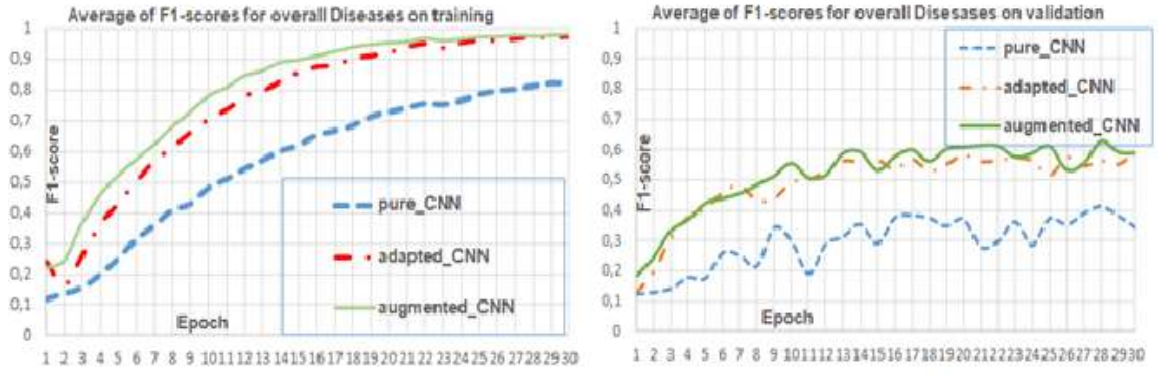
Şekil 2.2. K-fold Çapraz Doğrulama Yinelemesinde Kullanılan Veri Çoğaltma Süreci



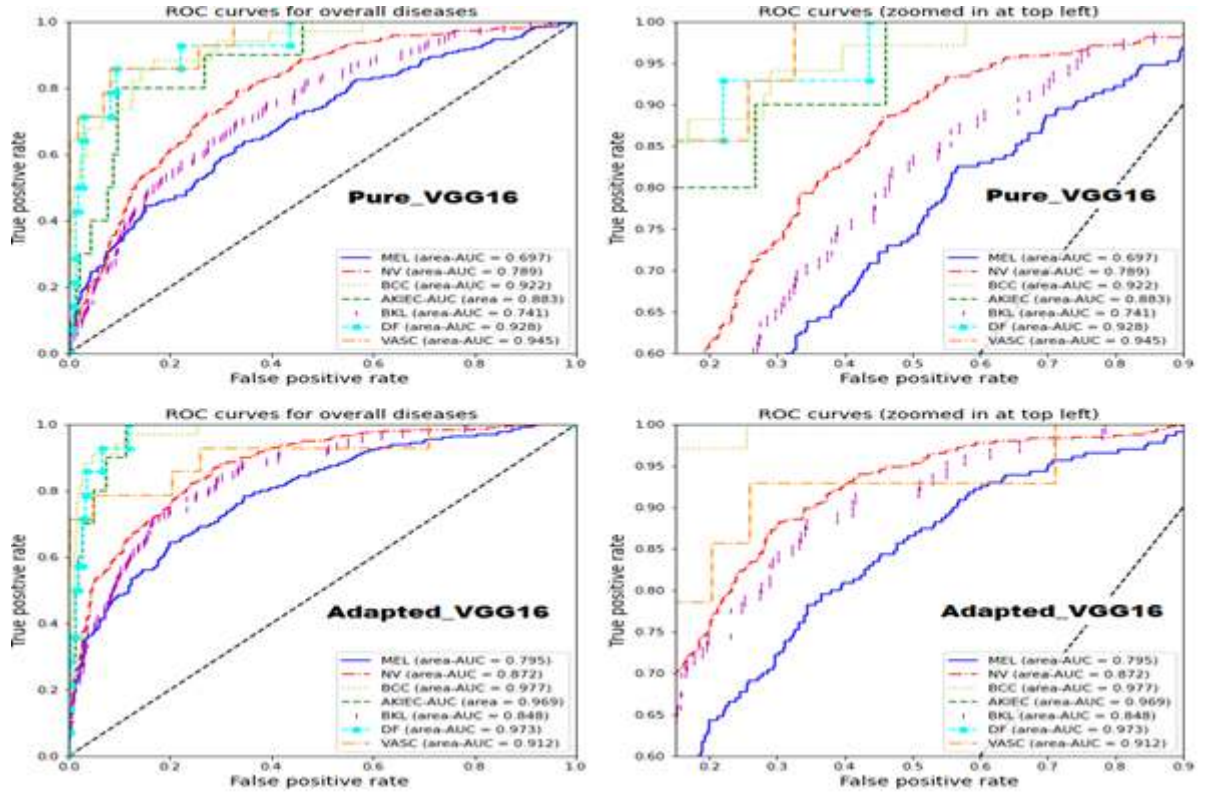
Şekil 2.3. Melanoma Üzerinde Data Augmentation ile Sentetik Resimler Üretme

Çalışmada uyguladığımız adaptasyon işlemlerinin bir CNN modelin başarısını nasıl iyileştirdiği Şekil 2.4. ve 2.5.'de gösterilmiştir. VGG16 olarak isimlendirilen kısmen eski sayılabilecek temel bir CNN modelinin eğitim ve doğrulama aşamalarına ait f1-score grafikleri Şekil 2.4'de verilmiştir. Her adaptasyon adımı, CNN modelin başarısına olumlu katkı sağlamıştır. Her ne kadar validation f1-score değeri eğitim sürecinden elde edilene yaklaşıp yaklaşmıyorsa da adaptasyon işlemlerinin burada da iyileştirme sağladığı görülmüştür. Şekil 2.4'teki pure_CNN versiyonu, VGG16 modelinin sadece transfer learning uygulanmış hali ile elde edilen başarıya ait eğridir. Adapted_CNN versiyonu ise VGG16 modeline sırası ile fine-tuning, regularization ve k-fold validation uygulanması ile elde edilen başarı eğrisidir. Son olarak augmented_CNN versiyonu ise sentetik veri çoğaltma da dâhil olmak üzere yukarıda anlatılan tüm adaptasyon işlemlerinin sonucunda elde edilen başarı eğrisidir.

Şekil 2.5'in üst kısmında verilen ROC eğrisi ve bu eğrinin zoom yapılmış hali, adaptasyon aşamalarından önceki VGG16 modellerine ait tahminlerin başarı durumunu göstermektedir. Adaptasyon işlemleri sayesinde VGG16 modelinin eğitim sonrası yaptığı tahminlere ait doğruluğun %15 - %20 aralığında iyileştiği Şekil 2.5'de alttaki ROC eğrisi ve zoom yapılmış halinde görülmektedir. Adapted VGG16 modelinin AUC değerlerinin, pure halinden ortalama %15 civarında daha iyi sonuçlar ürettiği görülmüştür.



Şekil 2.4. Vgg16 Modelinin Adaptasyon İşlemlerinden Önceki Ve Sonraki F1-Score Değerlerinin Karşılaştırılması



Şekil.2.5. VGG16 Modelinin Local-test Datası İçin Yaptığı Tahminlere ait ROC Eğrilerinin Adaptasyon İşlemlerinden Önceki ve Sonraki Hali

2.2. İkili Sınıflandırma Uygulamalarında Başarı Ölçütleri – Metrik Çeşitleri

Herhangi bir ikili sınıflandırıcının (binary classification) performans ölçütünü belirlemek, önemli bir tasarım parametresidir. Bu metrik seçimi ne kadar uygun yapılırsa modelin iyileştirme ve geliştirmesi de o denli doğru olacaktır. İkili sınıflandırıcının ne denli başarılı olduğunu sayısal bir değer ile ifade etmek gerekmektedir. Bu sayısal değeri hesaplamak için kullanılan

metrik seçiminde aşağıda irdelediğimiz belli başlı yaklaşımlar söz konusudur. Tez çalışmamızın bu kısmında genel kabul görmüş olan başarı ölçütleri incelenmiştir.

$$f1_score = 2 * \frac{precision*recall}{(precision+recall)}$$

$$precision = \frac{True\ Positive}{(True\ Positive+False\ Positive)}$$

$$recall = \frac{True\ Positive}{(True\ Positive+False\ Negative)}$$

Hastalık teşhisi yaparken; test sonucu pozitif yani test yapılan kişi hastalığa sahip ya da negatif yani test yapılan kişi hastalığı taşıyor olarak ayırım yapılmıştır. Bu durum, klasik ikili(binary classification) sınıflandırma örneğidir. Çalışmada her bir deneğin test sonucu, o kişinin melonoma hastası olup olmadığı şeklinde iki çıkışlı bir sınıflandırmadır. Çalışmada, melanomanın haricinde diğer sekiz deri hastalığı için de benzer şekilde olumlu veya olumsuz sonuç söylenmektedir.

Sınıflandırma algoritmalarını kullanarak yapılan çalışmalarda en büyük yanılgılardan biri başarı ölçütü olarak sadece doğruluk oranına bakmaktır. Özellikle dengeli dağılım göstermeyen veri setlerinde (imbalanced data sets) doğruluk oranı doğru bilgi vermez. Dengeli olmayan veri seti, sınıflara ait örnek sayılarının birbirine yakın olmadığını vurgulamaktadır. Örneğin; bir hastalığa ait 10.000 adet resim veri setinde mevcut iken diğer hastalık için bu değer sadece 1.000 tane olabilmektedir.

Herhangi bir ikili sınıflandırma uygulamasında performans ölçümü yapmaya çalışılıyorsa; doğruluk, duyarlılık, kesinlik, özgüllük, “F1 skoru”, “ROC eğrisi” ve “AUC alanı” gibi terimleri analiz etmek gerekmektedir. Bu terimlerin İngilizce kaynaklardaki karşılıkları ile Türkçe çevirileri, yaygın kullanımları göz önüne alınarak dile getirilmeye çalışılmıştır.

2.2.1. Başarı ölçütleri

a. Accuracy(Doğruluk)-Duyarlılık (Sensitivity): Sistemde doğru olarak yapılan tahminlerin tüm tahminlere oranıdır. (%90’lık tahmin sonucunun elde edildiği durumdur). Bu açıdan pek güvenilir bir karşılaştırma kriteri değildir.

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

Duyarlılık kimi zaman gerçek pozitif oranı (true pozitive rate) olarak da adlandırılır. Bir test ya da teşhis yönteminden bahsediliyorsa, duyarlılık testin hastalığı taşıyanları hangi oranla tespit edebildiği anlamına gelir. Burada altı çizilmesi gereken husus duyarlılığın kaç yanlış pozitif teşhis yapıldığıyla hiç bir ilgisi olmamasıdır. Eğer bir test yeteri kadar hassas yapılırsa o zaman duyarlılık derecesi çok rahatlıkla yukarıya çekilebilir. Duyarlılık sadece hasta bireylerin doğru

şekilde teşhis edilebilmesini ölçer. Bu açıdan kaç sağlıklı kişiye hasta teşhisi konduğu önemli değildir. Uç bir örnek olarak test edilen her kişiye hasta teşhisi koyan bir test mükemmel duyarlılığa sahip demektir. Elbette böyle bir test hiç bir işe yaramamaktadır. Test olan herkese melanoma olduğunu söyleyen bir test tüm melanoma pozitif vakalarını doğru teşhis etse bile hasta olmayan herkese de aynı teşhisi koyacaktır.

$$Sensitivity = \frac{TP}{(TP+FN)}$$

b. Precision (Kesinlik)-Recall (Duyarlılık): Pozitif durumların ne kadar başarılı tahmin edildiği bilgisidir. Çalışmaya uyarlandığında yani hasta olanları doğru tespit etme oranıdır. Recall kavramı bazı kaynaklarda Sensitivity-Duyarlılık olarak da isimlendirilmektedir.

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

Kanser vakalarını doğru tespit etmek yanlış tespitte bulunmaktan daha önemlidir. Diğer bir deyişle yanlış negatif, yanlış pozitiften daha kritiktir. Kanserli birini tespit edemeyip ölümüne neden olmaksızın kanser olmayan biri için yanlış tahmin yapıp onu hastaneye çağırarak daha kabul edilebilir. Herkesi hastaneye çağırma durumu (recall = 1) bütün kanserli vakalar bulunup fakat tespitlerin çoğu yanlış (FP yüksek) çıkmaktadır.

Precision ise, pozitif olarak tahmin edilen bir durumdaki başarıyı gösteren verilerdir.

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

Birine hasta demeden önce iyice düşünüp taşınmamızı sağlar. Herkese hasta denilirse precision = 10/10.000 = 0.001 çıkar. Bir kişiyi hasta olarak tespit edip ve o kişi gerçekten hasta çıkarsa precision = 1 olur. Fakat kalan 9 kişi ise FN değeri yüksek olduğundan tespit edilemez.

c. Özgüllük (Specificity): Özgüllük kimi zaman gerçek negatif oranı (true negative rate) olarak da adlandırılır. Bir test ya da teşhis yönteminin hasta olmayan bir hastaya doğru (yani negatif) teşhisi koyma oranını belirtir. Özgüllük hesaplanırken kaç hastaya doğru şekilde pozitif teşhis konduğu dâhil edilmez. Yeteri kadar muhafazakâr bir test yaratarak özgüllüğü keyfi bir şekilde arttırmak mümkündür. Duyarlılık gibi özgüllük de sadece teşhis hassasiyetinin tek bir yönüne bakar. Tüm test edilen hastalara negatif teşhisi koyan bir testin özgüllüğü mükemmel olacaktır [41].

$$Specificity = \frac{TN}{(TN+FP)}$$

d. F1-Score

F1 score sınıflandırma algoritmalarını karşılaştırırken sıklıkla kullanılan bir ölçüdür. Precision ve Recall değerlerinin harmonik ortalamasıdır. Bu sebeple başarının iyi bir göstergesidir. Aslında bir modelin F1 skoru ne kadar yüksekse çizginin altında kalan alan da o kadar yüksektir.

$$F1\ Score = \frac{2 * Precision * Recall}{Precision + Recall}$$

Tablo 2.1’de metrikler matematiksel bir örnekle açıklanmıştır.

Tablo 2.1 Özgüllük Ve Duyarlılık Metriklerinin Sayısal Bir Örneği

n:165	Tahmini Değer: HAYIR	Tahmini Değer: EVET	
Gerçek Değer: HAYIR	TN:50	FP:10	60
Gerçek Değer: EVET	FN:5	TP:100	105
	55	110	

$$Accuracy = \frac{TP+TN}{(TP+FP+TN+FN)} = \frac{100+50}{(100+10+50+5)} = 0.909$$

Sınıflandırıcının deneklere söylediği hastasın veya değilsin cümlelerinin doğruluk oranı 0.909’dur. Bu testin sonucu %90 doğrudur.

$$Recall = \frac{TP}{(TP+FN)} = \frac{100}{(100+5)} = 0.952$$

Sınıflandırıcının hasta olarak tahmin ettiği doğru sayısının tüm hastalara oranı 0.952 olarak çıkmaktadır. %95 başarı anlamına gelmektedir.

$$Sensitivity = \frac{TP}{(TP+FN)} = \frac{100}{(100+5)} = 0.952$$

Sınıflandırıcının hasta değil diye tahmin ettiği sayının tüm hasta olmayanlara oranı 0.952 çıkmaktadır.

$$Precision = \frac{TP}{(TP+FP)} = \frac{100}{(100+10)} = 0.909$$

Sınıflandırıcının hasta olarak tahmin ettiğinin hasta olup olmayanların toplamına oranı 0.909 yani %90 doğruluk oranıdır.

$$Specificity = \frac{TN}{(TN+FP)} = \frac{50}{(50+10)} = 0.833$$

Sınıflandırıcının hasta değil diye tahmin ettiklerinin tüm hasta olmayanlara oranı olarak 0.833 ya da % 83 başarı oranıdır

2.2.2. Hata Matrisi (Confusion Matrix)

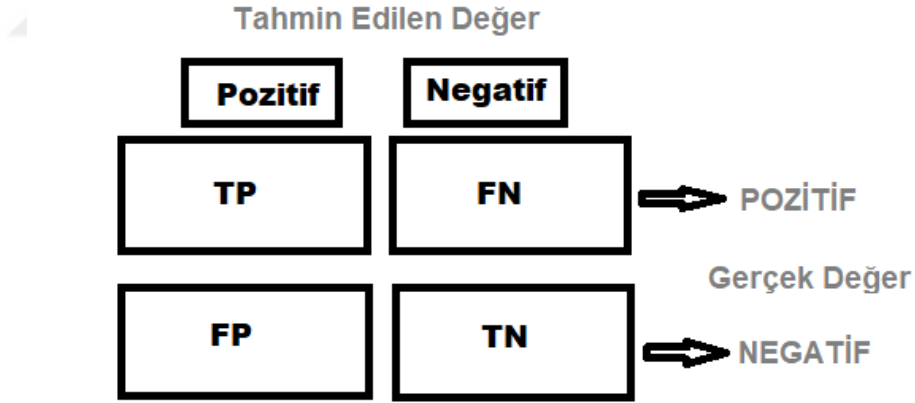
Bir karışıklık ya da hata matrisi, gerçek değerlerin bilinmekte olduğu bir dizi test verisi üzerinde, ikili sınıflandırma modelinin performansını tanımlamak için sıklıkla kullanılan bir şekildir. Şekil 2.6’da verilen sütunlar, ikili sınıflandırıcının çıkışları yani tahminleridir. Satırlar ise veri setindeki gerçek durumları göstermektedir. Pozitif sütunu; ikili sınıflandırıcının pozitif diye tahmin ettiği hastaların sayısını, negatif sütunu ise hasta değil diye tahmin edilen toplam denek sayısını göstermektedir. Hücrelere yazılan kısaltmaların açıklaması ise şu şekildedir:

TP(True Positive — Doğru Pozitif Tahminler): Hastaya hasta demek.

FP(False Positive — Yanlış Pozitif Tahminler): Hasta olmayana hasta demek.

TN(True Negative — Doğru Negatif Tahminler): Hasta olmayana hasta değil demek.

FN(False Negative — Yanlış Negatif Tahminler): Hasta olana hasta değil demek.



Şekil 2.6. Karmaşıklık Matrisi

Gerçek Pozitifler (TP): Bunlar gerçek değeri 1 ve tahmin edilen değerin de 1 olduğu örneklerdir.

Gerçek Negatifler (TN): Bunlar gerçek değeri 0 ve tahmin edilen değerin de 0 olduğu örneklerdir.

Yanlış Pozitifler (FP): Bunlar gerçek değeri 0 ancak tahmin edilen değerin 1 olduğu örneklerdir.

Yanlış Negatifler (FN): Bunlar gerçek değeri 1 ancak tahmin edilen değerin 0 olduğu örneklerdir.

$$\text{TOPLAM DENEK SAYISI} = \text{TP} + \text{TN} + \text{FP} + \text{FN}$$

$$\text{GERÇEK POZİTİFLER} = \text{TP} + \text{FN}$$

$$\text{GERÇEK NEGATİFLER} = \text{TN} + \text{FP}$$

2.2.3. ROC Eğrisi (Receiver Operating Characteristic Curve)

TPR (sensitivity) ile FPR (1-Specificity) eksenlerinden oluşan iki boyutlu bir grafikdir. TPR ve FPR açısından sınıflandırma sonucu, ROC grafiğindeki bir noktaya karşılık gelir. Tanı testlerinin sonuçları eğer var/yok, yüksek/normal veya düşük/normal şeklinde iki sonuçlu ise testlerin başarısı; sensitivite, spesifisite, pozitif prediktif değer, negatif prediktif değer ölçütleri ile değerlendirilir. Ancak test sonuçları sürekli (Hemoglobin değeri, WBC sayısı, plazma glukoz konsantrasyonu, PSA değeri gibi) ya da sıralı sayısal (BI-RADS gibi) sonuçlar şeklinde ise, bu testlerin tanı koymadaki başarılarını değerlendirmek biraz daha karmaşık olabilmektedir. ROC, farklı sınıflandırıcıların performanslarını karşılaştırmak için de kullanılabilir.

Eşik noktası, bir parametrenin eşik değeridir. Örneğin; kan testi sonucundaki bir hormonun değerine bakıp bir eşikten sonra karar vermektir. Eşik noktası değiştikçe tanı testinin sensitivite ve spesifitesi değişeceğinden dolayı, başarısının değerlendirmesi için belirlenecek eşik noktası da farklı olacaktır. Testin gerçek başarısını gösterecek bir yönteme olan ihtiyaç nedeniyle ilk kez 1971 yılında Lee Lusted'ın katkılarıyla, olası tüm eşik noktalarında tanı testinin her yanlış pozitif sonuç için ne kadar doğru pozitif sonuç vereceğini gösteren bir grafiğin de tanı testini değerlendirmede kullanılabileceği düşünülmüştür. Böylelikle farklı eşik noktalarında testin yanlış pozitif (1 – spesifisite) oranının x-ekseninde, doğru pozitif oranının ise (sensitivite) y-ekseninde olduğu ROC eğrisi (Receiver Operating Characteristic) elde edilmiştir.

ROC analizi istatistik paket programlarının yardımıyla yapılabilmektedir. Bu amaçla SPSS, Medcalc gibi istatistik paket programları sıklıkla kullanılan programlar arasındadır.

ROC analizi; bir testin tanı koyma başarısını (Recall ve Accuracy değerini) belirlemek, farklı eşik noktalarında elde edilen recall ve accuracy değerlerini karşılaştırmak amacıyla kullanılır. Ayrıca belirli recall değerlerine denk gelen accuracy değerlerini ya da belirli accuracy değerlerine denk gelen recall değerlerini belirlemekte de kullanılmaktadır. Bunların yanı sıra iki ya da daha fazla testin tanı koymadaki başarısını karşılaştırmak amacıyla da kullanılmaktadır. Sınıflandırma modellerinin başarı hesaplarında ROC eğrisi sık sık karşımıza çıkmaktadır. Temelde iki basit metrik üzerinden sınırlar belirlenip çizilmektedir. Bu iki metrik;

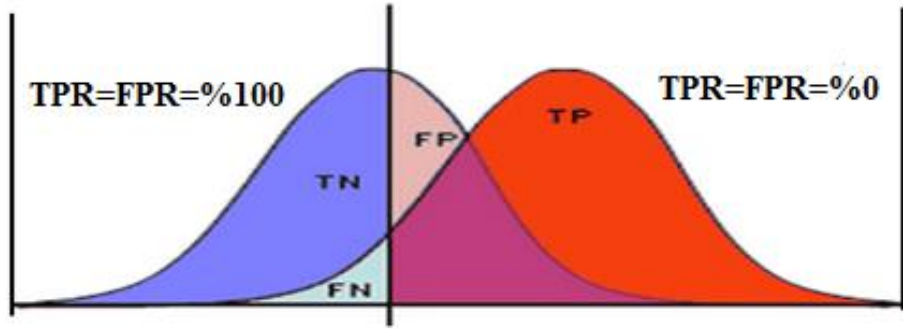
TPR, kanser olan insanların hangi olasılıkla doğru tespit edildiği bilgisidir. (Recall aynı zamanda Sensitivity)

$$\text{True Positive Rate} = \frac{TP}{(TP+FN)}$$

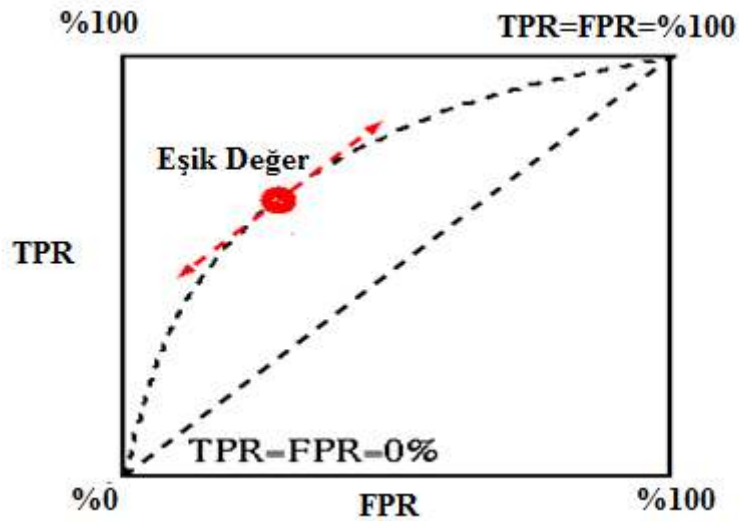
FPR(1-Specificity), Kanser olmayan insanları hangi olasılıkla yanlış tespit edildiği bilgisidir.

$$\text{False Positive Rate} = \frac{FP}{(TN + FP)}$$

Bu iki metrik x ve y eksenlerine yerleştirilerek çizginin altında kalan alan hesaplanır (AUC -Area Under Curve). Bu iki metriğe ait grafikler ve eğrinin altında kalan sınırlar Şekil 2.7’de detaylı bir şekilde gösterilmiştir.



Şekil 2.7. ROC Eğrisinde Metrik Sınırlar



Şekil 2.8. FPR ve TPR Eşik Değerleri

Şekil 2.8’de; iyi bir sınıflandırıcının ROC grafiği, $TPR = 1$ için çok hızlı bir şekilde üst kenara ulaşması gerektiği belirtilmiştir. Çünkü FPR 0’dan 1’e yükselir. Eğrinin altındaki alan sınıflandırıcının performansını ölçmek için kullanılır. ROC eğrisinin altındaki alan ne kadar büyükse sınıflandırma performansı o kadar iyi olur.

Örnekler;

- $TPR:0.6, FNR=1-TPR=0.4, FPR=0.3, TNR=1-FPR=0.7$

Bu durumda, ROC’deki karşılık gelen nokta (0.3; 0.6) üzerindedir, rastgele bir tahmine göre daha iyidir.

- $TPR:0.6, FNR=1-TPR=0.4, FPR=0.6, TNR=1-FPR=0.4$

Bu durumda, ROC’deki karşılık gelen nokta (0.6; 0.6) üzerindedir, rastgele bir tahminle aynıdır.

- $TPR:0.2, FNR=1-TPR=0.8, FPR=0.6, TNR=1-FPR=0.4$

Bu durumda, ROC’deki karşılık gelen nokta (0.4; 0.2) üzerindedir, rastgele bir tahminden daha kötüdür [42].

3. ULUSLARARASI CİLT GÖRÜNTÜLEME BİRLİĞİ VERİ SETİ (ISIC ARŞİVİ)

Bu tez çalışmasında veri seti olarak “International Skin Imaging Collaboration (ISIC)- Uluslararası Cilt Görüntüleme Birliği” tarafından hazırlanan resimler kullanılmıştır. “ISIC Arşivi” de denen bu veri seti, cilt görüntülerinin dijital resim formatında halka açık olarak sunulduğu en büyük global depo alanıdır. ISIC arşivi, deri ve zührevi hastalıklar alanında kullanılan, farklı niteliklere göre ayrılmış 24.000 cilt görüntüsünü, dijital resim formatında bulunduran bir veri setidir. Cinsiyet, patolojik büyüklük, cilt kanseri türü şeklinde değişik kriterlere göre resimler ayrıştırılmıştır. Bu veri setinde en yaygın olarak bilinen deri kanseri olan melanoma ile birlikte yedi farklı cilt hastalığının da erken teşhisi için özel olarak hazırlanmış resimler bulunmaktadır. ISIC uluslararası bir topluluk olup; melanoma gibi birçok cilt hastalığının tanı ve tedavi sürecinin iyileştirilmesine yardımcı olmaya çalışmaktadır. Cilt görüntülerinin dijital ortamda erişilebilirliğini kolaylaştırmak için tasarlanmış bir üniversite ve sanayi iş birliğidir.

ISIC arşivi, birçok farklı ülkedeki klinik merkezlerinden toplanan ve bu merkezlerdeki farklı cihazlardan elde edilen 24.000'den fazla dermoskopik görüntüden oluşmaktadır. ISIC, önerilen uluslararası standartları test etmek ve onaylamak için sahip olduğu gün geçtikçe artan cilt görüntülerini kamu erişimine açık bir arşiv haline getirmiştir. Bu uluslararası resim arşivi, otomatik cilt hastalıkları teşhis sistemlerinin geliştirilmesi ve test edilmesi için halka açık bir kaynak olarak hizmet vermektedir.

ISIC projesinin genel amacı; melanoma deri hastalığının erken tespitinin doğruluğunu arttırmak ve bu sayede melanoma ile ilişkili ölümleri ve gereksiz biyopsileri azaltmaktır. Bu amaçla ISIC, cilt lezyonlarının klinik ve dermoskopik resimlerinin dijital görüntüleme standartlarını geliştirmekte ve halka açık olarak arşivlemektedir. Dermatoskopi, iyi ve kötü huylu cilt kanserlerini belirlemede yaygın olarak kullanılan bir tanı tekniğidir. Çıplak gözle yapılan incelemeye kıyasla deri hastalıklarını otomatik olarak teşhis etmek için kullanılır.

Bu tez çalışmasında olduğu gibi; deri hastalıklarının teşhisinde yapay sinir ağını kullanmak isteyen araştırmacıların ihtiyaç duyduğu dijital resimlerden oluşan veri seti ISIC topluluğu tarafından açık kaynak olarak sunulmaktadır. Bu tez çalışmasında da, evrişimli sinir ağı(CNN) tabanlı lezyon sınıflandırma uygulaması için dermoskopi cihazlarından alınan renkli deri görüntülerinden faydalanılmıştır. Cilt hastalıklarının yapay sinir ağları ile teşhisinde; dermoskopik görüntülerden oluşan mevcut veri seti boyutlarının küçük olması nedeni ile sinir ağlarının eğitiminde sorunlar yaşanmaktaydı. Bu sorunu ortadan kaldırmak amacıyla kısa adı ISIC olan Uluslararası Cilt Görüntüleme İşbirliği ekibi HAM10000 veri setini yayınlayarak, sorunun çözümüne çok büyük katkı sağlamışlardır. HAM10000 veri seti en ölümcül cilt kanseri

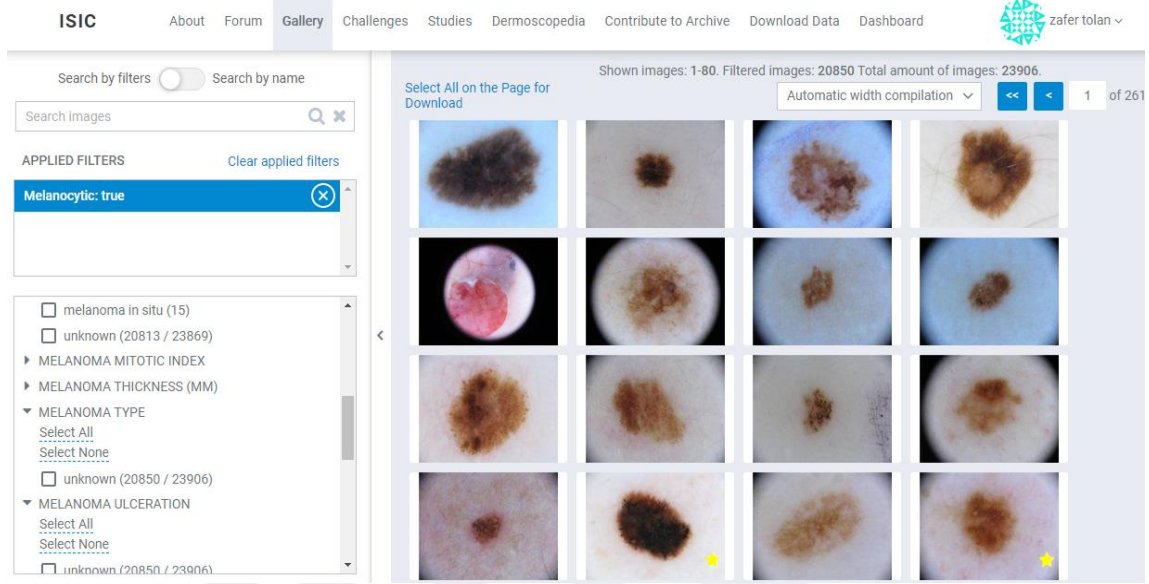
türü olan melanomayı diğer cilt hastalıklarından ayırmak için kullanılan cilt görüntülerinden oluşur.

Uluslararası Cilt Görüntüleme İşbirliği tarafından hazırlanan veri seti halka açık olduğundan ve kapsamlı veriler içerdiğinden, şimdiye kadar melanomanın yapay sinir ağları kullanılarak teşhis edilmesi için en büyük veri seti olarak hizmet etmiştir. Cilt kanserinden ölüm oranının azaltılmasına yardımcı olmak için tasarlanmış bir akademik arşivdir. Cilt kanserlerinde ilk aşamada tanı konulması, tedavide yüksek başarının temel şartıdır. Bu çalışmada kullandığımız veri seti, dijital cilt görüntüleri, klinik uzmanlarının karar desteği ve bilgisayar sistemleri aracılığıyla melanoma tanısına doğrudan yardımcı olmaktadır. Ayrıca ISIC veri seti, hem klinik hekimlerini hem de halkı eğitmek için kullanılmaktadır. ISIC topluluğu, cilt görüntüleme için kullanılan teknolojileri, teknikleri ve terminolojiyi oluşturmaktadır. Ayrıca gizlilik ve birlikte çalışabilirlik konularına özel standartlar geliştirmektedir.

Bu veri seti içinde farklı yöntemlerle ve farklı ülkelerden temin edilen dermoskopik görüntüler bulunmaktadır. Makine öğrenmesi alanında yapılan akademik araştırmalarda bir eğitim seti olarak kullanılabilmesi için ISIC tarafından makale olarak yayınlanmıştır. HAM10000 veri seti, ISIC Arşivi aracılığıyla herkese açık olan 10.015 tane dermoskopik görüntüden oluşmaktadır. 10.000 kişiden fazla hastadan alınan bu cilt görüntüleri, 2018'de İspanya'da düzenlenen uluslararası Tıbbi Görüntü Hesaplama ve Bilgisayar Destekli Müdahale (MICCAI) konferansında kullanılmıştır. Melanom Deri Lezyon Analizi Üzerine ISIC veri setinin ne amaçla oluşturulduğu, yöntemleri ve sonuçları bu konferansın sonuç makalesinde yayınlanmıştır. Bu çalışmada konferans kapsamında yapılan yarışmada araştırmacıların geldiği başarı noktası %90'a yakın doğru hastalık teşhisi olarak saptanmıştır. 2018 yılındaki ISIC yarışmasında HAM10000 veri seti, yarışmacıların algoritmalarını eğitmeleri için internet üzerinden sunulmuştur. Ancak test verisinin gerçek hastalık etiketleri gizlenmiş ve bu sayede eğitim verisinden izole edilmiştir. HAM10000 kısaltması, "Human Against Machine with 10.000 training images" İngilizce ifadesinden türetilmiştir. 10.015 hastadan alınan cilt görüntüsü dijital resim formatında ve gerçek hastalık etiketleri ile birlikte yarışmacılara sunulmuştur. Bunun yanında 1.512 tane hastadan alınan cilt görüntüsü ise etiketleri gizlenmiş olarak test verisi amacıyla yayınlanmıştır. Test verisinin gerçek dünya etiketleri bilinmediği ve eğitim verisinde yer almadığı için, yarışmacıların geliştirdikleri sınıflandırıcıların ezber yapmasının önüne geçilmiştir.

Dermatoskopik görüntülerin alınmasında kullanılan donanımsal aygıtlar mevcuttur. Derinin mikroskopla incelenmesi yöntemi olarak da tanımlanabilen dermatoskopi, cilt görüntülerinin bilgisayar destekli analiz edilmesi yöntemidir. Fotofinder videodermatoskop bu amaçla geliştirilmiş bir dermatoskopi cihazı olup, hastanın görüntüsünü ekrana aktararak benleri atlamadan tek tek inceleme olanağı sunmaktadır. Bu yöntemle vücuttaki benlerin haritası oluşturularak, her noktanın lokalizasyonu da belirlenmektedir. Ardından her bir ben için

dermatoskopik görüntü alınıp kaydedilmektedir. Böylece bir sonraki kontrolde elde edilecek yeni görüntünün, uzman hekim tarafından önceki halleri ile karşılaştırılabilmesi mümkün olmaktadır.



Şekil 3.1. ISIC Arşivinin Web Portal Arayüzünden Alınan Bir Ekran Görüntüsü

Şekil 3.1’de görüldüğü üzere ISIC arşivi içerisinde; cinsiyet, yaş aralığı, patolojik boyutlar ve hastalık tanısına göre görüntüler kategorize edilmiştir [37].

Bu semantik bilgiler, eğitim veri seti ile birlikte sunulan metadata dosyasında bulunmaktadır. Dijital resimlerin yanı sıra, .csv uzantılı bu text dosyası da arşivden birlikte indirilmekte ve içerisinde resimlerin dosya adları ile bağlantılı ve anlamlı metinler bulunmaktadır. Aslında ISIC Arşivinde, uluslararası alanda önde gelen klinik merkezlerindeki çeşitli cihazlardan elde edilen 24.000’den fazla dermoskopik görüntü mevcuttur. ISIC Arşivi’ne gelen tüm görüntüler hem gizlilik hem de kalite güvencesi açısından taranmaktadır. Görüntülerin çoğu tanınmış melanom uzmanları tarafından onaylanan klinik meta verilere sahiptir. Ayrıca bazı görüntüler üzerinde, bu uzman hekimler tarafından yapılan açıklama ve işaretlemeler de yer almaktadır. Bu işaretler, dermoskopik özellikleri vurgulamaktadır. ISIC Arşivi’nin yazılım altyapısı açık kaynaklı Girder platformu kullanılarak inşa edilmiştir. Arşive ait kaynak kodu GitHub’da ücretsiz olarak mevcuttur.

ISIC yarışmalarında test verisini eğitim verisinden izole etmek amacı ile test veri kümesi, gerçek hastalık etiketleri gizlenmiş olarak yarışmacılara sunulmaktadır. Bu amaç doğrultusunda HAM10000 veri setinde, 10.015 tane etiketli eğitim örneğinin yanı sıra 1.512 tane de etiketsiz cilt görüntüsü yayınlanmıştır. Test verisinin gerçek dünya etiketleri bilinmediği ve eğitim verisinde yer almadığı için yarışmacıların geliştirdikleri sınıflandırıcıların overfit yapmasının önüne geçilmiştir. 2016 yılından bu yana ISIC tarafından düzenlenen ödüllü yarışmaya, katılımcılar hem

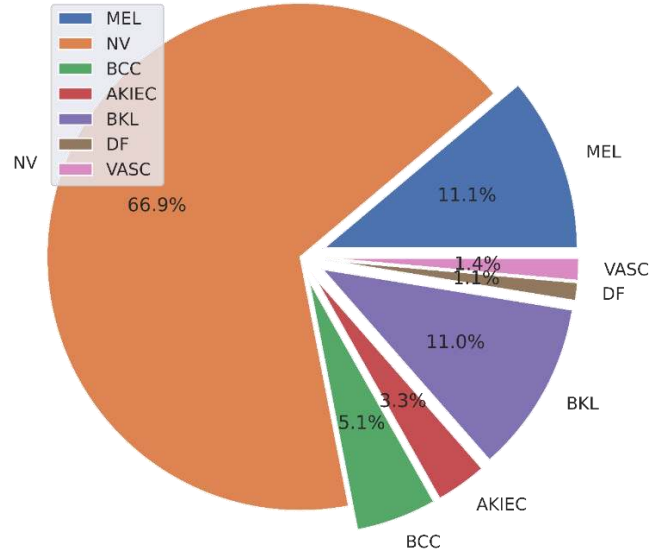
algoritmalarını tanımlayan kısa bildirileri hem de geliştirdikleri yöntemin test verisi üzerindeki performansları ile iştirak etmektedir. Ayrıca yeni araştırmacıların performanslarını test edebilmeleri için ISIC topluluğunun web sitesinde yıllara göre ayrıştırılmış veri setleri için online performans değerlendirme sistemi bulunmaktadır(challenge.isic-archive.com). Bu sayede çalışmamızda olduğu gibi geliştirilen her yeni algoritmanın, test verisi üzerindeki başarısı online olarak anında hesaplanabilmektedir. Tablo 3.1.’de veri setinin hastalık bazında niteliksel açıklamaları yer almaktadır.

Tablo 3.1. Ham10000 Veri Setindeki Tüm Hastalıkların Açıklamaları ve Görüntü Sayıları.

Disease	Name	Description	Count
MEL	Melanoma	Melanosit denilen cilt hücrelerinin kontrolsüz bir şekilde bölünüp çoğalması ile ortaya çıkan cilt kanseri türüdür.	1.113
NV	Nevus	melanositlerin iyi huylu halleridir.	6.705
BCC	Basal cell carcinoma	Deride bulunan eski hücrelerin ölmesiyle onların yerine yenilerini üreten, bazal hücrelerdeki kanser türü.	514
AKIEC	Actinic keratosis	uzun süreli güneşe maruz kalma Kansere dönüşme riski çok düşüktür.	327
BKL	Benign keratosis	iyi huylu keratoz olarak bilinir. genel itibariyle yaşlılarda görülür ve zararsız cilt büyümeleridir.	1.099
DF	Dermatofibroma	çoğalma açısından iyi huylu bir cilt hastalığıdır. sert, yüksek görünümlü ve farklı boyutlarda oluşan life benzer lezyonlardır.	115
VASC	Vascular lesion	Damarsal sorunlar nedeniyle, deri içindeki yüzeysel kılcal damarların genişlemesi veya normalden çok sayıda olmasıdır.	142

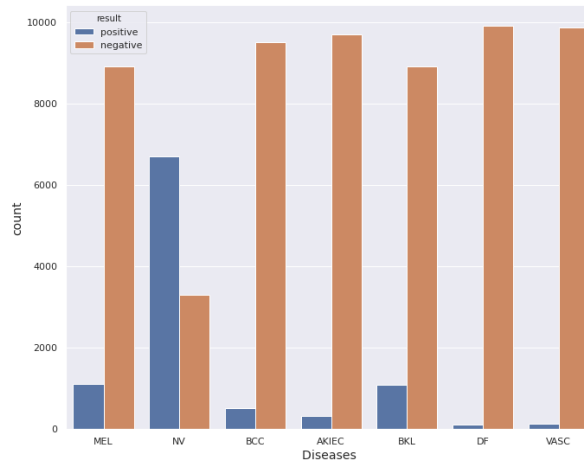
HAM10000 veri seti içerisinde, “Melanoma(MEL)” gibi ölümcül deri kanseri hastalarından alınan lezyon görüntülerinin yanı sıra cerrahi müdahaleye gerek kalmaksızın tedavi edilebilen “Nevus(NV)” veya çoğunlukla rahatsızlık vermediği için farkına bile varılmayacak olan “Aktinic Keratosis(AKIEC)” gibi iyi huylu lezyon görüntüleri de bulunmaktadır. Bu veri seti içerisindeki hastalık isimleri için yapılan kısaltmalar, hastalıkların kısaca tanımlamaları ve her bir hastalığa ait toplam görüntü sayıları Tablo 3.1’de açıklanmıştır. Çalışmanın genelinde hastalıklar, aynen HAM10000 veri setindeki sütun isimlerinde olduğu gibi kısaltmalar ile temsil edilmiştir.

Multi-classing uygulamalarda, baskın olan sınıfa ait pozitif örneklerin, eğitilen model tarafından yakalanma olasılığı yüksektir. Başka bir deyişle; eğitim sınıfındaki majör sınıfa ait örneklerle karşı, model daha hassas çalışmaktadır. Ancak Şekil 3.2’de vurgulandığı gibi HAM10000 veri seti pastası, hastalıklara ait örnek resim sayılarına göre kesilecek olursa; dilim büyüklüğü %5 ve daha düşük olan 4 tane azınlık sınıfı olduğu görülmektedir. Eğitilen modelin bu küçük dilimlere karşı önyargılı çıktı vermesi doğaldır. Sınıflandırıcılar, azınlık sınıflarına karşı daha az duyarlı olduğundan, dengeli dağılım göstermeyen bir veri seti ile eğitilen sınıflandırıcıların tahmin çıktıları önyargılıdır.



Şekil 3.2. HAM10000 verisetindeki hastalıklara ait resim sayılarının yüzdesel dağılımı

Binary bir sınıflandırma probleminden farklı olarak çoklu sınıflandırıcı uygulamalarda, negatif örnek sayılarının doğası gereği fazla olması kaçınılmazdır. Çünkü n tane çıkışı olan bir sınıflandırma probleminde herhangi bir sınıfın eğitim veri setindeki negatif örnek sayısı, geri kalan $(n-1)$ tane sınıfa ait pozitif örneklerin toplamıdır. Ancak Şekil 3.2’de gösterildiği gibi HAM10000 veri setinde minör sınıfların tamamı için negatif örnek sayısının aşırı derecede fazla olduğu göze çarpmaktadır. HAM10000 veri seti içerisinde sadece major sınıf olan “NV” hastalığına ait pozitif resim sayısı aynı hastalığın negative resim sayısından fazladır. Bu nedenle NV sınıfı haricinde kalan tüm sınıflarının pozitif örnek sayıları, sentetik olarak oversampling edilmiştir. NV sınıfı için örnekleme işlemi, veri kaybına ve kurban seçilecek olan örnek NV hastalık resimlerindeki özelliklerin sınıflandırıcıdan saklanması sebep olacağı için tercih edilmemiştir. Şekil 3.3’te her hastalık için ISIC veri setinde bulunan etiket sayısı gösterilmiştir.



Şekil 3.3. HAM10000 Verisetindeki Her Hastalık İçin Etiketli Örnek Sayısı

Sağlık hizmetlerinde, en kötü durum senaryosu yanlış negatif teşhistir (False Negative Diagnosis). Kullanılan bütün hekim karar destek sistemlerinde, ihtiyaç duydukları bakım ve tedaviyi sağlamak için mümkün olduğunca çok sayıda kanser hastasının belirlenmesi istenmektedir. Ayrıca bu makalede ele alınan deri kanseri hastalıklarında erken teşhis edebilme yeteneği, hayatta kalma olasılığını çok büyük oranda arttırmaktadır. Bu nedenle HAM10000 gibi negatif örnek sayısının çok fazla olduğu azınlık sınıfları içeren veri setleri ile uygulama geliştirirken accuracy değerini metrik olarak seçmek büyük bir hata olacaktır. Bunun yerine recall yani False Negative oranını hesaba katan bir ölçüm tercih edilmelidir. Bu doğrultuda bizim çalışmamızda, modellerin eğitim ve doğrulama aşamalarında f1-score puanı baz alınmıştır. Her evre sonunda veri setinin hem eğitim hem de doğrulama alt kümeleri için f1-score değerleri kayıt edilmiştir.

Denklem 1'deki bilinen eşitlikten de hatırlanacağı gibi f1-score puanı, precision ve recall metriklerinin harmonik ortalamasıdır. Precision metrik, eğitilen sınıflandırıcının çıktısında hasta diye tahmin edilen kişinin gerçekten hasta olma olasılığını, recall metrik ise gerçekten hasta olan örneklerin yüzde kaçının tespit edilebildiğini göstermektedir. Recall metrik aynı zamanda hit-rate yani isabet oranı olarak da isimlendirilmektedir. Modelimizi yanlış-negatif tahminler için olabildiğince cezalandırmak istediğimizden, modelimizi değerlendirmek için doğru ölçüt recall metriktir. Ayrıca çalışmamızda eğitilen modellerin değerlendirildiği off-line test veri setinde de balanced accuracy yani tüm sınıflara ait recall değerlerinin ortalaması göz önüne alınmıştır. Dolayısıyla False Negative oranı yüksek olan modellerin başarımı düşük kalmıştır. Accuracy gibi klasik optimizasyon ölçütleri, özellikle yanlış negatif veya yanlış pozitif tahminlerine ait risk arttığında, gerçek performansın göstergesi olarak kullanılamamaktadır.

3.1. ISIC Arşiv Görüntülerinin Elde Edilmesi ve Saklanması

HAM10000 eğitim setinin dermatoskopik cihazlardan alınan görüntüleri uzman bir ekip tarafından Viyana Üniversitesi Tıp Fakültesi Dermatoloji Departmanından ve Avustralya'daki Cliff Rosendahl cilt kanseri muayenehanesinden 20 yıllık bir süre içinde toplanmıştır. HAM10000 veri setinin tüm veri kayıtları Harvard Dataverse depo alanında saklanmaktadır. Avustralya üzerinden elde edilen görüntüler ve meta veriler PowerPoint dosyalarında ve Excel veritabanlarında saklanmıştır. Avusturya'dan elde edilen görüntülere gelince; dijital kameralardan önce görüntüler toplanmaya başlanmıştı ve görüntüleri ile meta verileri farklı zaman dilimlerinde farklı formatlarda depolanmaktaydı. Veri setini oluşturanlar görüntülerin elde edilmesi ve saklanması için aşağıda değinilen farklı yöntemlerden faydalanılmıştır.

3.1.1. PowerPoint dosyalarından görüntülerin ve meta verilerin çıkarılması

Her PowerPoint dosyası, bir takvim ayı içerisinde elde edilen klinik çalışmalardaki dermoskopik görüntüleri içermekteydi. Ayrıca her slayt, tek bir görüntü ve benzersiz bir cilt hastalığı kimlik numarasına sahip bir metin alanı içeriyordu. Büyük miktarda veri nedeniyle bu görüntüleri ayıklamak ve sıralamak için otomatik bir yaklaşım uygulanmıştır. PowerPoint dosyalarına erişmek ve içeriği elde etmek için Python dilinde python-pptx paketi kullanılmıştır. Her slayt tek tek ele alınmış, dosya adını oluşturmak için kaynak görüntü ismi, ilgili hastanın kimlik numarası ve dokümantasyon yılı otomatik olarak hesaplanıp kaydedilmiştir.

3.1.2. Dijital dermatoskopi sisteminden verilerin alınması

Viyana Üniversitesi Dermatoloji Bölümü dijital dermatoskopi sistemi MoleMax HD (Derma Medical Systems, Viyana, Avusturya) ile donatılmıştır. SQL tabloları kullanılarak özel bir araçla filtreleme yapıp bu sistemden cilt vakaları çekilmiştir. Aynı nevüsün bir dizi ardışık görüntüsünden sadece en yenisi göz önüne alınmıştır. Bu veri setinin bazı melanomları ayrıca bir DermLiteTM FOTO (3GenTM) kamera ile fotoğraflanmıştır. Bu ek görüntüler, aynı lezyonun farklı görüntülerinin ortak bir tanımlayıcı diziyle etiketlendiği ViDIR görüntü serisinin bir parçası haline getirmiştir. MoleMax HD sisteminin orijinal görüntüleri, kuadratik olmayan piksellerde 1872x1053px (MoleMax HD) çözünürlüğe sahiptir. Tüm HD görüntüler manuel olarak 800x600px (72DPI) olarak kırpılmıştır.

3.1.3. Dermatoskopik görüntülerin filtrelenmesi

Hem Avusturya'dan hem de Avustralya'dan alınan görüntü kaynakları sadece dermoskopik görüntüleri değil aynı zamanda klinik yakın çekimleri ve genel bakışları da içermekteydi. Görüntüleme türünün güvenilir bir açıklaması olmadığından, dermoskopik görüntüleri diğerlerinden ayırmak zorunda kalınmıştır. Büyük miktarda veriyle verimli bir şekilde başa çıkmak için, 30.000'den fazla görüntüyü taramak ve sınıflandırmak için otomatik bir yöntem geliştirmiştir. Avustralya koleksiyonunun 1.501 görüntü dosyası, "genel bakışlar", "yakın çekimler" ve "dermatoskopi" kategorileri şeklinde etiketlenmiştir. El ile etiketlenmiş görüntüleri bir eğitim seti olarak kullanıp, resimleri görüntü türüne göre sınıflandırmak için bir InceptionV3 ağ mimarisi tasarlanmıştır. Bu yapay sinir ağında Stokastik Gradient Descent eğitim modeli 20 döngülük ve öğrenme oranı 0.0003 olacak şekilde başlatılmıştır. 7. ve 13. döngülerde yavaşlama (Gama 0.1) saptanmıştır. Parti boyutu 64 resim olarak seçilmiştir. Bu eğitim ayarları sonucunda % 98.68'lik bir doğruluk elde edilmiştir. Bu doğruluk, dermoskopik görüntülerin seçim sürecini hızlandırmak için yeterli kabul edilmiştir. Kalan bir kaç yanlış sınıflandırılmış yakın plan ve genel bakış resimler ise ikinci bir revizyonda elle düzeltilmiştir.

3.2. ISIC Veri Setindeki Mevcut Hastalık Tanımları

ISIC veri setinde, melanoma cilt kanserinin resimlerinin yanı sıra altı farklı cilt hastalığının da görüntüleri mevcuttur. Bunların tamamı için cilt lezyonu ifadesi kullanılmakta olup farklı yapılarda karşımıza çıkmaktadırlar. Veri setinin Avustralya’da toplanan bölümü, yüksek risk gruplarında melanomun erken teşhisi için hazırlanmıştır. Bu hasta grubunda genellikle yüksek sayıda nevüs cilt hastalığı ve aile kaynaklı melanoma saptanmıştır. Avustralya’da meydana gelen cilt hastalıklarının temel sebebi güneş ışınlarına uzun süreli maruz kalmış olmasıdır.

Bu tez çalışmasında geliştirilen sınıflandırıcının 7 tane çıkışının her biri farklı bir cilt hastalığını işaret etmektedir. Bu nedenle öncelikle aşağıdaki paragraflarda kısaca bu hastalıklar ile ilgili tanımlayıcı bilgiler verilmiştir.

3.2.1. Melanoma(Melanom)

Deriye rengini veren melanin adlı renk pigmentleri, melanosit denilen cilt hücreleri tarafından üretilmektedir. Bu hücrelerin kontrolsüz bir şekilde bölünüp çoğalması sonucunda meydana gelen cilt kanserine melanoma denir. Bir başka deyişle melanom olarak da isimlendirilmektedir. Cildin yapısında mevcut olup cilde renk veren melanositlerin etkisiyle ciltte oluşan bir cilt kanseri türüdür. Bu hastalık için malign melanom ve kutanöz melanom gibi farklı isimlendirmeler de mevcuttur. Melanomlar cilt üzerinde farklı bölgelerde de gelişebilmektedir. Bilimsel olarak melanomlara diğer cilt kanserlerine göre daha az rastlanır. Ancak cilt kanserine bağlı ölümlerin yaklaşık olarak %75’i melanomlara bağlı olarak gerçekleşmektedir. Kanser hücreleri çok hızlı büyüyen ve sistematik olarak yayılan bir davranış sergiler. Tüm kanser çeşitlerinde olduğu gibi cilt kanserinin tedavisi için de erken tanı büyük önem taşımaktadır. Dolayısıyla ciltte meydana gelen kanser belirtilerinin erken fark edilmesi ve doğru değerlendirilmesi son derece önemlidir. Ciltte meydana gelen bir renk değişimi ya da vücut beni üzerinde bir değişiklik fark edildiğinde renk ve şekil açısından takip edilmeli, ten renginde veya pembe tonlarında görülen kabartılar da dikkate alınmalıdır.

Dermatologlar sıklıkla Melanomu erken aşamalarında tanılamak ve ortadan kaldırmak için düzenli cilt muayenesi önermektedir. Melanomların yaklaşık %25’i vücuttaki benlerden gelişir ve erkeklerde kadınlara göre daha sık görülmektedir [38].

Genellikle orta yaşlı insanlarda daha çok görülür. Erken teşhis çok önemlidir çünkü hayatta kalma oranları zamanla hızlı düşmektedir. Çıplak gözle erken evre melanom tanı şansı %50 iken dijital dermoskopik inceleme ile %90’lara çıkabilmektedir. Melanom çok erken evrede normal bir vücut beni gibi çıkıp sonradan yapı ve renk değiştirebilir. Dijital dermatoskoba kayıt edilen görüntülerle benlerde değişim ve ilerleme olup olmadığı takip edilebilmektedir. Önceki görüntüsüne göre ilerlediği tespit edilen vücut beni cerrahi olarak çıkartılıp, patolojik incelemeye

gönderilir. Çalışma bu kapsamda ele alınırsa, patolojik sonuçlardan gelen eğitim verileriyle oluşturulmuş hastalıklı benlerin bulunduğu resim arşivi kullanılarak, derin öğrenme teknikleri sayesinde yüksek doğruluk ile cilt kanseri tespiti gerçekleştirilmiştir. ISIC arşivindeki test verisi ile de eğitim resimlerindeki başarı seviyesine yaklaşılmıştır.

3.2.2. Aktinik Keratoz(Aktinic Keratoses-Pigmented Bowen's)

Aktinik keratoz, uzun süreli kontrolsüz bir şekilde güneşe maruz kalmaya bağlı olarak ortaya çıkan bir cilt hastalığıdır. Güneş ışığına daha fazla maruz kalan vücut bölgelerinde daha sık rastlanmaktadır. Deride anormal hücre gelişimi sonucu oluşan cilt değişiklikleridir. En sık görülen şekli ise yüzeyi pürüzlü yama şeklindedir. Aktinik keratozların düşük oranda deri kanserine dönüşme riski vardır. Herhangi bir rahatsızlık vermediğinden çoğu hasta tarafından farkına bile varılmaya bilinmektedir. Görünümleri, üzeri pürüklü, zaman zaman pullanan yama şeklindedir. Sıklıkla birden fazla olurlar. Kırmızı-pembe, ten rengi, renksiz olabilirler. Başlangıçta leke görünümündedir. Bazen gözle görmek zor olabilir, ancak elle üzerine dokunulduğunda zımpara kağıdı izlenimi verirler. Aktinik keratozlar dudakta olduklarında ‘aktinik keilit’ olarak adlandırılırlar. Bu cilt rahatsızlığının belirtileri arasında dudaklarda kuruluk, çatlaklar ve pullanma vardır.

3.2.3. Seboreik Keratoz(Seborrheic Keratosis-Pigmented Benign Keratoses)

Seboreik keratoz, iyi huylu keratoz olarak bilinir. Melanoma gibi gözle görünebilecek cilt kanseri olmayan bir lezyondur. Seboreik keratozlar genel itibariyle yaşlılarda görülür ve zararsız cilt büyümeleridir.

3.2.4. Bazal Hücreli Karsinoma(Basal Cell Carcinoma)

Bazal hücreli karsinom, deri kanserinin bir türüdür. Deride bulunan eski hücrelerin ölmesiyle onların yerine yenilerini üreten, bazal hücre olarak adlandırılan hücre tipinde başlar.

Bazal hücreli karsinom farklı şekillerde olabilmekle birlikte sıklıkla ciltte hafif şeffaf bir yumru olarak belirir. Baş ve boyun gibi güneşe maruz kalan alanlarda daha sık görülür. Bazal hücreli karsinomların birçoğuna uzun süreli ultraviyole (UV) ışına maruz kalmanın sebep olduğu düşünülmektedir. Güneşten korunmak ve güneş kremi kullanmak, bazal hücreli karsinomdan korunmaya yardımcı olabilir. Bazal hücreli karsinom, derideki bazal hücrelerin DNA'sında bir mutasyon meydana geldiğinde oluşur. Bazal hücreler, derinin en dış katmanı olan epidermisin altında bulunur ve yeni deri hücreleri üretirler. Yeni bir deri hücresi üretildiğinde eski hücreler cilt yüzeyine itilir ve orada ölüp dökülürler. Yeni deri hücrelerinin oluşması süreci bazal

hücrelerin DNA'sı tarafından kontrol edilir. DNA, hücreye ne yapması gerektiğini anlatan bilgileri içerir. Mutasyonlar ise bazal hücrelere hızlıca bölünmelerini ve normalde ölmeleri gereken zamanda yaşamaya devam etmelerini söyler. Sonuç olarak biriken anormal hücreler kanserli bir tümör haline gelebilir ve bu durum ciltte bir lezyon olarak görünür.

3.2.5. Nevüs (Nevus)

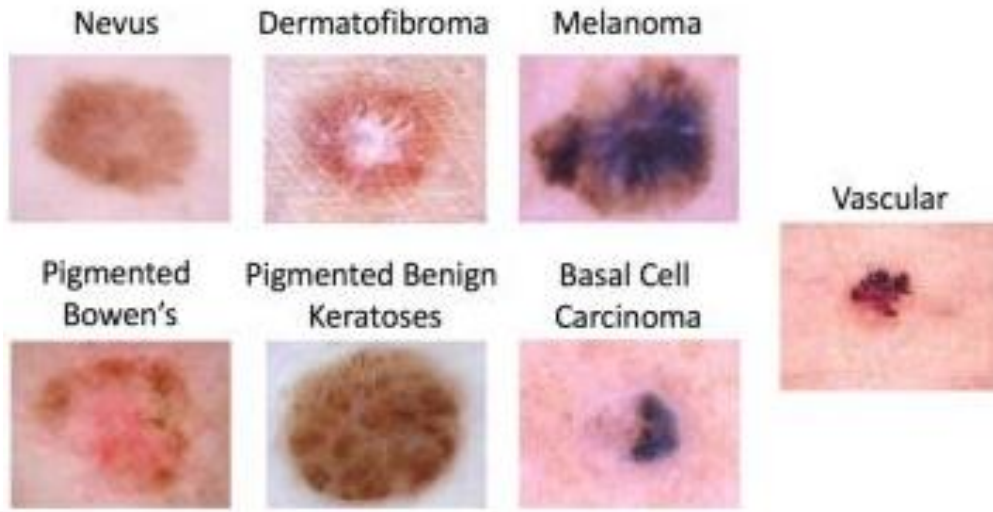
Melanositik nevuslar, melanositlerin iyi huylu halleridir ve sayısız varyantta görülebilmektedir. Varyantlar, dermatoskopik bir bakış açısı ile önemli ölçüde farklılıklar sergileyebilmişlerdir. Nevüsler melanomun aksine renk ve yapı dağılımı bakımından genellikle simetriklerdir[32]. Nevüs iyi huylu bir deri tümörüdür. Displastik nevi, nevusun çoğul halidir. Nevüs, çoğunlukla normal benler gibi görünür ve melanomun birçok özelliğini taşır. Genellikle diğer benlerden daha büyüktür. Şekli veya rengi alışılmışın dışındadır ve çoğunlukla kansere dönüşmez. Doğumda var olan benlere konjenital melanositik nevüs adı verilir. Doğumla birlikte var olan bu tür benlerde nevüsün büyüklüğüne bağlı olarak melanom gelişme riskinin en fazla %1 olduğu düşünülür. Bununla birlikte Konjenital(Doğumsal) melanositik nevüsü büyük olan kişilerde melanom gelişme riski daha fazladır. Örneğin; konjenital nevi avuç içinden daha küçükse, melanom riski daha düşüktür. Bunun aksine, sırtta veya kalçalarda büyük boyutlarda olan doğumsal nevi olan kişilerde melanom riski önemli ölçüde artmaktadır.

3.2.6. Damarsal(Vasküler) Deri Lezyonları(Vascular Skin Lesions)

Damarsal sorunlar deri içindeki yüzeysel kılcıl damarların genişlemesi veya normalden çok sayıda olmasıyla kendini göstermiştir. Klinik olarak damarlar, yüzeysel ise parlak kırmızı daha derinde veya geniş olması halinde ise bordodan maviye kadar değişen farklı tonlarda görüntü vermiştir.

3.2.7. Dermatofibroma

Dermatofibroma, çoğalma açısından ele alındığında iyi huylu bir cilt hastalığıdır. En sık görülen dermatoskopik görüntüsü, cilt üzerinde merkezi beyaz bir yama ile retiküler çizgilerin bulunduğu yapılardır [39]. Şekil 3.4'te eğitim ve test verileri olmak üzere 10015 görüntüden oluşan HAM10000 veri setinden alınan ve 7 farklı cilt hastalığını barındıran örnek görüntüler verilmiştir [40].



Şekil 3.4. ISIC Arşivinden Alınan 7 Farklı Lezyon Görüntü Örneği

Tablo 3.2. HAM10000 ve Diğer Veri Setlerinin Cilt Hastalıkları Görüntü Sayıları Açısından Karşılaştırılması

Veri Seti	Lisans	Toplam Görüntü	Patolojik Doğrulama (%)	akiec	bcc	bkl	df	mel	nv	vasc
PH2	Research&Education ^a	200	20.5%	-	-	-	-	40	160	-
Atlas	No license	1024	unknown	5	42	70	20	275	582	30
ISIC 2017 ^b	CC-0	13786	26.3%	2	33	575	7	1019	11861	15
Rosendahl	CC BY-NC 4.0	2259	100%	295	296	490	30	342	803	3
ViDIR Legacy	CC BY-NC 4.0	439	100%	0	5	10	4	67	350	3
ViDIR Current	CC BY-NC 4.0	3363	77.1%	32	211	475	51	680	1832	82
ViDIR MoleMax	CC BY-NC 4.0	3954	1.2%	0	2	124	30	24	3720	54
HAM10000	CC BY-NC 4.0	10015	53.3%	327	514	1099	115	1113	6705	142

Tablo 3.2’de HAM10000 ile diğer bilinen açık kaynak dermoskopik görüntü veri setlerinin hastalıklara ait resim sayıları karşılaştırmalı olarak verilmiştir. Veri setlerinin sayısal özetinin gösterildiği bu tabloda bulunan 7 farklı hastalık için etiketlendirmeler mevcuttur. Bu etiketler; akiec, bcc, bkl, df, mel, nv ve vasc kısaltmalarıyla verilmiştir. Tabloya göre en yüksek hastalık görüntü sayısı nevüs(nv) ve melanoma(mel) hastalıklarına aittir. Her bir veri seti için patolojik doğrulama oranları da ilgili sütunda gösterilmiştir. Ayrıca arşivlerin 2017 yılına ait barındırdığı toplam görüntü sayıları da mevcuttur. 2017 yılında yine nevüs(nv) ve melanoma(mel) görüntü sayıları daha ağır basmaktadır. 2018 yılındaki Uluslararası yarışmada ISIC HAM10000 veri setinin yanı sıra dijital ve analog kameralardan elde edilen ViDIR Molemax, ViDIR Current ve ViDIR legacy adlı veri setleri de kullanılmıştır.

ViDIR image seti Avusturya kaynağından elde edilen resimlerdir. ViDIR Group (Avusturya Viyana Üniversitesi Dermatoloji Bölümü) tarafından farklı zamanlarda temin edilmiştir. Tıbbi etik kurul onayından sonra işlenmiştir. Legacy Diapositives ise ViDIR grubunun en eski görüntü kaynağıdır. Analog kameralarla dermatoskopik görüntülerin çekildiği ve dijital kameraların kullanılabilirliğinden önceki zamanlara dayanmaktadır. Orijinal görüntüler Heine Dermaphot sistemi ile çekilmiştir. E-6 yöntemi ile eğitim ve arşivleme amacıyla üretilmiştir. MoleMax Serileri, Viyana'daki Dermatoloji Departmanı tarafından hazırlanmıştır. Nevüs çıkarımlarını azaltıp erken melanom sayısını artırmak için yüksek riskli hastalara dijital dermatoskopik takip uygulanmıştır. Takipler arasındaki zaman aralığı genellikle 6-12 ay arasında değişmektedir. Rinner ve arkadaşları yakın zamanda bu takip programının ayrıntılı bir açıklamasını yayınlamıştır. 2015 yılından beri MoleMax HD Sistemi, dermatoskopik görüntülerin alınması ve saklanması için kullanılmaktadır.

4. ÖNERİLEN YÖNTEMLER VE DEĞERLENDİRME

Çalışmamızda öncelikle güncel CNN modellerin HAM10000 veri seti karşısında sergiledikleri başarı performansları kıyaslanmıştır. 9 farklı CNN modelinin 7 farklı cilt hastalığına ait resimleri içeren bu veri seti için ne denli başarılı sınıflandırma yaptıkları analiz edilmiştir. Elde edilen sonuçlarda her ne kadar EfficientNet ailesi ön plana çıkmış olsa da, tüm hastalıklar için lider konumda olduğu söylenememiştir. Herbir modelin daha hassas olabildiği hastalıkların varlığı gözlemlenmiştir.

Bu bölümün ikinci alt başlığında ise en ölümcül cilt hastalığı olan Melanoma kanseri için EfficientNet mimarisi optimize edilmeye çalışılmıştır. Modelin hyper-parametreleri iyileştirilmeye çalışılmıştır. Diğer hastalıklar bir kenara bırakılarak sadece Melanoma tespiti için veritabanları birleştirilerek, EfficientNetB5 modeli ile başarılı sonuçlar elde edilmiştir.

Bu tez çalışmasının son aşamasında ise; bireysel CNN performanslarının, topluluk öğrenme yöntemleri ile birleştirilerek daha iyi sonuçlar elde edilebileceği düşüncesine yer verilmiştir. Topluluk öğrenme yöntemleri içerisindeki temel teknikler olan ortalama ve oylama yaklaşımlarından daha iyi sonuçlar elde etmemizi sağlayan ağırlıklandırılmış birleştirme tekniği geliştirilmiştir.

4.1. Evrimsel Modellerin Performans Karşılaştırması

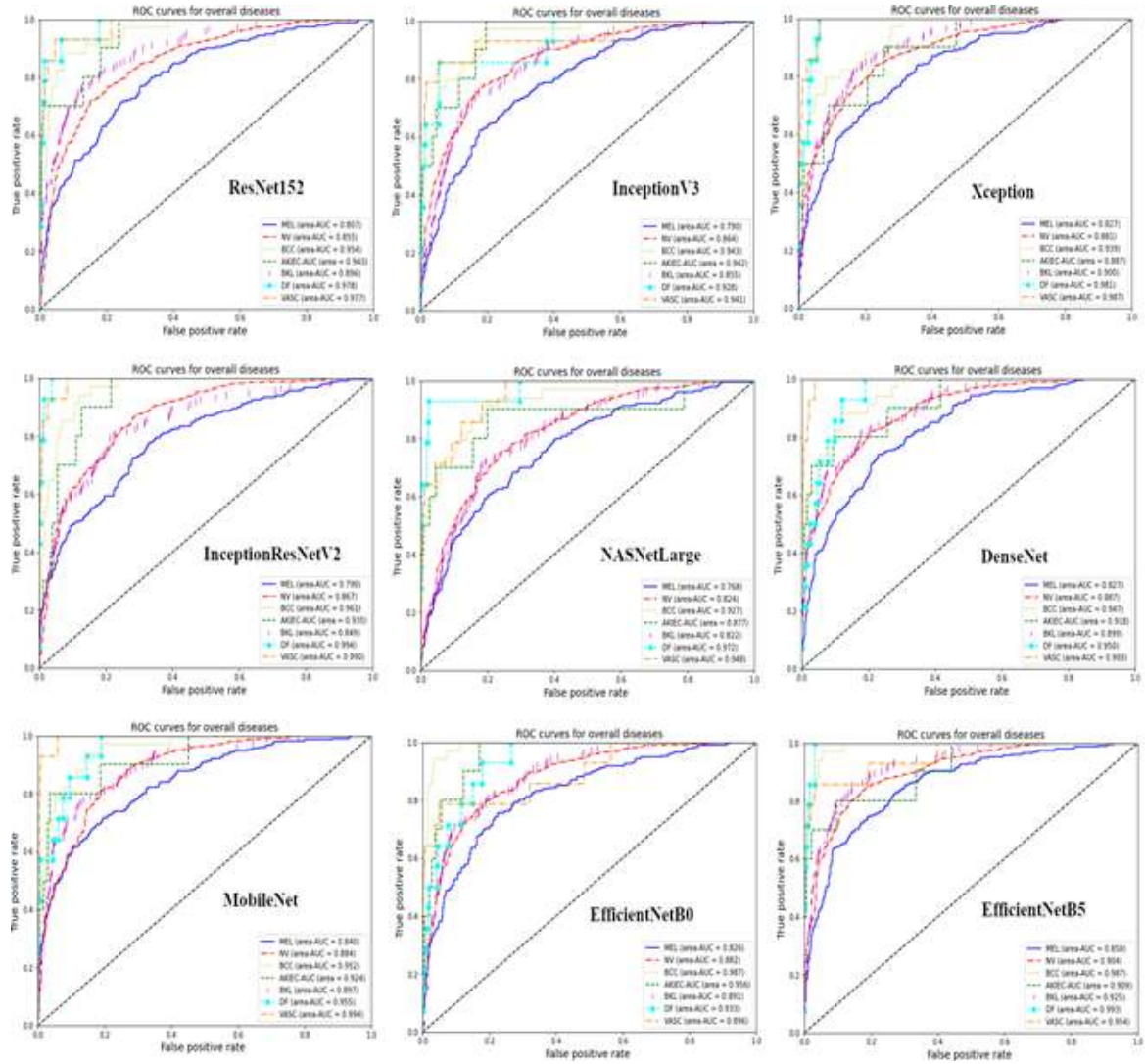
Eğitilen modellere off-line test yapabilmek için 1.015 tane eğitim resminin gizli tutulduğuna Şekil 2.2’de değinilmişti. Bu test verisine ait gerçek etiket değerleri yani hastalık bilgileri elimizde olduğundan eğitim ve doğrulama aşamasını tamamlayan modellerin hiç görmediği 1.015 tane resim için yaptığı tahminler yerelde değerlendirilebilmiştir. Bu sayede, eğitilen modelin bu off-line test kümesine ait tahminlerinin ne denli doğru olduğunu gösteren ROC eğrisi ve altında kalan AUC değerleri her bir hastalık sınıfı için ayrı ayrı çizdirilmiştir.

Tüm modellerin en çok zorlandıkları hastalık sınıfı MEL olduğu görülmektedir. Sadece “*Melanoma Detection*” başlığı ile araştırma çalışmaları yapılmasının temel nedeni bu zorluk derecesidir. MEL sınıfına ait AUC değerinin en yüksek olduğu CNN modeli, EfficientNetB5 mimarisi olmuştur. EfficientNetB5 mimarisi diğer hastalık sınıfları için de yüksek AUC değerleri ile ön plana çıkmaktadır.

MobileNet mimarisi, boyutunun küçük olmasına rağmen EfficientNet ailesine yakın doğrulukta tahminler yapabilmektedir. AUC değerlerinin en düşük olduğu mimarinin NASNet olması değinilmesi gereken diğer bir sonuçtur. Aslında maliyetine katlanıp NASNet çözümünün iyileşmesi için modelin optimizasyon süreci uzun tutulabilirdi ama bu diğer adaylar için bir haksızlık anlamına gelirdi.

ResNet ve Inception ailelerin hibrid hali olan InceptionResNet mimarisinin beklendiği gibi atalarından daha iyi sonuç verdiği görülmektedir. Bunun yanı sıra bu ailenin yakın dostu olan Xception ise her iki ailenin de önüne geçmiştir.

Model giriş ve çıkışlarındaki ağırlık sayılarının artması ile oluşturulan DenseNet mimarisinin de Inception ailesinden daha iyi sonuçlar doğurduğu görünmektedir.



Şekil 4.1. CNN modellerin Off-line Test Datasına Ait ROC Eğrilerinin ve AUC Değerlerinin Karşılaştırılması

CNN modellerinin off-line test datası için yaptıkları tahminlerin, gerçek etiket değerleri ile kıyaslanması Şekil 4.2’de sergilenmiştir. Şekil 4.2’de konfüzyon matrislerinin her birindeki hücrelerinin toplamı 1.015 sayıdır. HAM10000 veri setindeki hastalıklara ait resim sayılarının dağılımında olduğu gibi konfüzyon matrislerde de en koyu hücre NV majör sınıfına aittir. Tüm

hücrelerin toplamının %50'si civarındadır. En başarılı tahminlerin beklendiği gibi HAM10000 veri setindeki majör sınıf olan NV için yapıldığı gözlemlenmiştir.

Şekil 4.2'den bir satırı okuyacak olursak; örneğin ResNet152 ağıının ilk satırını göz önüne alındığında Offline test setindeki 230 adet MEL sınıfına ait dermoskopik resmin sadece 107 tanesi doğru tahmin edilmiştir. Gerçekte MEL sınıfına ait olan resimlerden 74 tanesi NV, 6 tanesi BCC, 1 tanesi AKIEC, 36 tanesi BKL, 4 tanesi DF ve 2 tanesi de VASC olarak tahmin edilmiştir. Konfüzyon matrisler sayesinde multi-classing bir uygulamada modelin hangi hastalığı ne ile en sıklıkla karıştırdığı veya hangisi yüksek doğrulukta tahmin edebildiği ortaya çıkmaktadır.

Multi-classing bir uygulamada CNN modelin Konfüzyon matrisinde, diyagonal çizgi üzerindeki hücre değerlerinin olabildiğince yüksek olması beklenir. Elde ettiğimiz çıktılarda bu duruma en yakın CNN modelinin EfficientNetB5 olduğu görülmüştür. Diyagonal hücrelerin haricindeki koyu hücrelerin sayısı ilgili satırdaki hastalığın, ilgili sütundaki hastalık ile çok fazla karıştırıldığını göstermektedir. Çalışmada bu kötü senaryoya en yakın örnek olarak NASNet mimarisidir. Ancak diğer modellerin en çok zorlandığı MEL sınıfına ait en yüksek doğru tahmini NASNet sınıfının yapması önemli bir avantajdır.

ROC eğrilerinde olduğu gibi, tüm CNN modellerin yanlış tahminlerinin en fazla olduğu satır melanoma hastalığıdır. En iyi tahminlerin sahibi EfficientNetB5 mimarisinin, melanoma tahminlerinin diğer modellerden düşük kalması da bir başka dikkat çekici noktadır.

Ayrıca tüm minör sınıflar için False Negative Rate hücrelerinin toplamının, True Positive Rate hücredeki sayıya yakın olduğu vurgulanması gereken diğer bir önemli zayıflıktır. Bunun anlamı yapılan tahminlerin hastalık bazında %50 doğruluk civarında kaldığıdır. Elbette bir yazı tura kadar kötü değildir ancak hata oranının yüksek olduğunu söylemek gerekir.



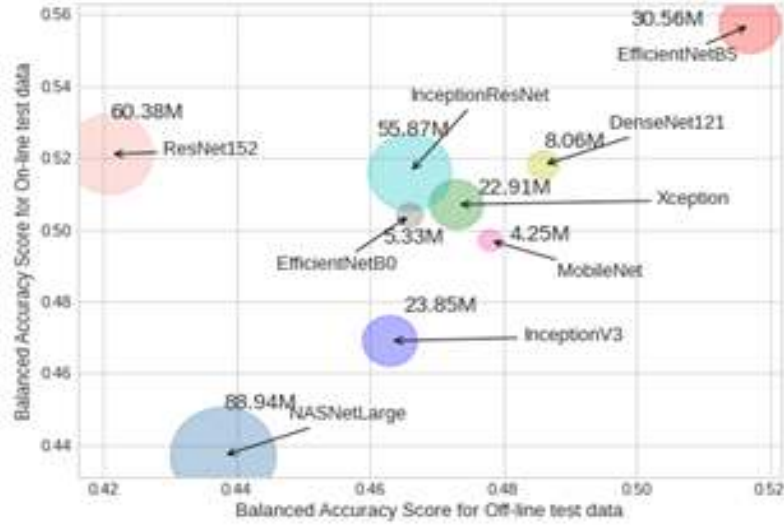
Şekil 4.2. CNN Modellerinin Off-line Test Datası Tahminlerine Ait Confusion Matrislerinin Karşılaştırılması

Melanoma tespitinde hangi mimarilerin seçileceği ve bunların performans kıyaslaması her ne kadar off-line test amacı ile 1.015 tane resmi test amaçlı izole etmiş olsakta aslında HAM10000 veri setinin zaten 1.512 tane resimden oluşan gerçek test kümesi bulunmaktadır. Ancak bu gerçek test kümesinin etiket değerleri gizli tutulduğu için model geliştirme aşamasında kullanmak mümkün değildir. Bu gerçek test kümesinin değerlendirmesi on-line platform üzerinden sadece Balanced Accuracy sonucu şeklinde Şekil 4.2’de gösterildiği gibi verildiğinden, çalışmamızda önce yerelde detaylı değerlendirme yapabilmek için eğitim seti içerisindeki 1.015 tane resim off-line test kümesi olarak kullanılmıştır. ISIC online değerlendirme sisteminin haftalık 10 adet submission limiti bulunmaktadır. Balanced Accuracy Score değeri, hastalıklara ait recall veya başka bir deyişle hit-rate oranlarının ortalamasıdır.

Off-line test işleminde başarılı bulunan model daha sonra HAM10000 verisetinin tamamı ile yani 10.015 tane lezyon resmi ile ön-eğitilmiş ağırlık değerlerinden başlanarak eğitim işlemine tabi tutulmuştur. Bu aşamada artık doğrulama ya da off-line test için eğitim verisinin parçalanmasına gerek kalmamıştır. Veri setinin tamamı ile eğitilen modellerin gerçek test kümesine ait performans değerleri ISIC topluluğunun on-line değerlendirme sisteminden elde edilmiştir. Şekil 4.1’de değerlendirmeye tabi tutulan CNN modellerin Off-line ve on-line test veri setleri için yaptıkları tahminlerin Balanced Accuracy Score değerleri, nokta büyüklükleri parametre sayılarını temsil edecek şekilde çizdirilmiştir. Şekil 4.1.’deki en büyük nokta NASNet ailesi olmasına karşın hem off-line hem de on-line tahminlerinin doğruluğu diğer ailelerden düşük kalmıştır. Aslında network mimari optimizasyonu süreci çok uzun tutularak daha iyi performans gösterebilecek bir hale getirilebilir. Ama tüm çalışma zamanının sadece bu amaca harcanması manasına gelirdi ki; bu durumda diğer CNN modellerine verilmemiş bir imtiyaz olurdu.

ResNet mimarisi temsil eden noktadan anlaşılacağı üzere; online ve offline tahminleri arasındaki mesafe, diğer konvolüsyon ağlarından daha fazladır. Online performans tahmini öncesi eğitim setinin tamamı ile yapılan komple eğitim en çok ResNet ailesinin performans artışına sebep olmuştur. Parametre sayısı yakın olmak ile birlikte Xception noktasının hem online hem de offline performansı, atası olan inception mimarisinden daha iyi konumdadır. Ebeveynleri Inception ve ResNet olan InceptionResNet ise off-line test datası üzerinde ResNet’den çok daha iyi performans göstermişken, online değerlendirmede çok yaklaşmıştır. InceptionResNet noktası, parametre sayısı olarak Inception mimarisinin iki katına çıkmasının getirisini online ve offline tahminlerinin doğruluğunun artması ile göstermiştir. Son zamanlarda küçük ağ yapıları ile büyük performanslar yakalama çabaları Şekil 4.2’de kendini MobileNet noktası ile ön plana çıkarmıştır. Nokta büyüklüğü diğer ailelere kıyasla çok düşük kalmasına rağmen tahmin doğruluğunda birçok köklü ailenin önüne geçmiştir.

DenseNet mimarisi de parametre sayısının düşüklüğüne karşın yüksek performans gösteren bir noktadır. 121 katmanlık derinliğine göre parametre sayısı nispeten az ve tahmin doğruluğu ise yüksektir. Tablo 4.1.’de Cnn Modellerinin Online-Real Test Datası Üzerinde Gösterdikleri Performans Ve Parametre Sayılarının Karşılaştırılması sayısal verilerle ifade edilmiştir. Son zamanlarda, sayısal olarak düşük veri yapılarıyla büyük performanslar elde etme çabaları Şekil 4.3’de ki MobileNet noktası ile kendini açıkça göstermiştir. MobileNet mimarisi veri boyutu diğer mimarilere kıyasla çok küçük olmasına rağmen, tahmin doğruluğunda birçok köklü mimariyi geride bırakmıştır.



Şekil 4.3. Cnn Modellerinin Balanced Accuracy Score Değerleri ile Performans Kıyaslaması

Tablo 4.1. Cnn Modellerinin Online-Real Test Datası Üzerinde Gösterdikleri Performans Ve Parametre Sayılarının Karşılaştırılması

CNN Model	Input Shape	Conv_base parameters	Required Memory	Trainable Parameters	Total parameters	Balanced Accuracy Score (offline test)	Balanced Accuracy Score (online test)
ResNet152V2	(224,224,3)	60,380,648	232 MB	117,9760,71	161,357,319	0.421	0.521
InceptionV3	(299,299,3)	23,851,784	92 MB	140,553,223	156,285,735	0.463	0.469
Xception	(299,299,3)	22,910,480	88 MB	216,756,087	230,841,903	0.473	0.507
InceptionResNetV2	(299,299,3)	55,873,736	215 MB	105,757,735	155,265,255	0.466	0.516
NASNetLarge	(331,331,3)	88,949,818	343 MB	135,892,039	216,189,529	0.438	0.437
DenseNet121	(224,224,3)	8,062,504	33 MB	53,775,367	58,682,951	0.486	0.518
MobileNet	(224,224,3)	4,253,864	16 MB	53,232,135	54,874,311	0.478	0.497
EfficientNetB0	(224,224,3)	5,330,571	29 MB	65,612,087	68,540,067	0.466	0.504
EfficientNetB5	(450,450,3)	30,562,527	118 MB	120,799,379	146,480,375	0.517	0.557

4.2. Efficient Modellerinin Optimize Edilmesi

Çalışmanın bir diğer aşamasında, günümüzün en güncel ve önde gelen CNN mimarisi olan EfficientNet ailesinden çözüm adaylarının Melanoma deri kanseri tespitindeki başarımlarını analiz edilmiştir. EfficientNet ailesinin temel modeli olan B0 ile işe başlayıp, daha yüksek doğruluk değerlerine ulaşmak için CNN model etkili bir şekilde ölçeklenerek B7'ye kadar ulaşılmıştır. CNN modelin ölçeği büyütülürken beklenenden daha az sayıda parametre artışı olması en temel kazanımdır. Uluslararası Cilt Görüntüleme Birliği (ISIC) tarafından sunulan renkli lezyon görüntüleri kullanılarak EfficientNet modellerinin eğitimi gerçekleştirilmiştir. ISIC arşivi, deri kanseri hastalarından alınan ve patolojik doğrulaması yapılmış dünya çapındaki en

büyük resim deposudur. CNN modellerin doğrulaması için ise 5-fold çapraz doğrulama tercih edilmiştir. Ayrıca CNN modellerinin performans iyileştirmesi için hyper-parametrelerin optimum değerleri yakalanmaya çalışılmıştır. Eğitim aşamasında başarımların ölçümü metriği olarak F1-score seçilmiş ve test verisi için ROC eğrileri çizdirilerek AUC değerleri ortaya konmuştur. Pandemi sürecinde hastalık teşhisinin erken tespiti ve tedavi yöntemlerinin önemi daha çok ortaya çıkmıştır. İnsanlar, bu dönemde hastaneye gitmekten çekinmekte ve olabildiğince uzaktan yardım almaya çalışmaktadırlar. Bu çalışmanın konusu olan renkli lezyon görüntülerinin kategorize edilmesinde kullanılan bilgisayar destekli sistemler, uzaktan teşhis için ideal örneklerdir.

Cilt lezyon görüntülerinden melanoma hastalığını tespit etmek için en yeni derin öğrenme modellerinden biri olan EfficientNet mimarilerinin başarısını ölçmek amaçlanmıştır. Bilimsel olarak Melanomlara diğer cilt kanserlerine göre daha az rastlanır. Ancak cilt kanserine bağlı ölümlerin yaklaşık olarak %75'i melanomlara bağlı olarak gelişir. Kanser hücreleri çok hızlı büyüyen ve sistemik olarak yayılan bir davranış sergiler. Tüm kanser çeşitlerinde olduğu gibi cilt kanserinin tedavisi için de erken tanı büyük önem taşır. Dolayısıyla ciltte meydana gelen kanser belirtilerinin erken fark edilmesi ve doğru değerlendirilmesi son derece önemlidir. Ayrıca ciltte meydana gelen bir renk değişimi ya da vücut beni üzerinde bir değişiklik fark edildiğinde renk ve şekil açısından takip edilmelidir. Ten renginde veya pembe tonlarında görülen kabartılar da dikkate alınmalıdır. Dermatologlar sıklıkla melanomayı erken aşamalarında tanımlamak ve ortadan kaldırmak için düzenli cilt muayenesi önermektedir.

En çok rastlanılan kanser türü olan deri kanserlerinin erken teşhisinde, Convolution Neural Network(CNN) tabanlı resim işleme uygulamaları ön planda gelmektedir. Derin öğrenmenin her alanında olduğu gibi resim işleyen konvolüsyon modellerinde de her geçen gün popüler ve iddialı yeni mimariler ortaya çıkmaktadır. Bu çalışmada performans karşılaştırması yapılan modellerin en yaşlısı bile beş yaşını doldurmamıştır. Bu çalışmada ele alınan popüler CNN modellerinin dermoskopik veri setleri üzerindeki sınıflandırma kabiliyetleri uygulamalı olarak kıyaslanmıştır. Hangi CNN modelinin ne kadar parametre ile ne denli başarı gösterdiği ortaya konmuştur.

Çalışmanın temel prensibi, her görüntü için ikili bir hedef tahmin ediyoruz. EfficientNET modelimiz, görüntüdeki lezyonun malign olma olasılığını 0.0 ile 1.0 arasında tahmin etmelidir. Eğitim verilerinde ise train.csv dosyası kullanılmaktadır. 0 değeri iyi huylu anlamına gelir.1 değeri ise kötü huylu anlamına gelir. Test verileri için test.csv dosyası kullanılmıştır.

Dermatoskopi, iyi ve kötü huylu cilt kanserlerini belirlemede yaygın olarak kullanılan bir tanı tekniğidir. Çıplak gözle yapılan incelemeye kıyasla deri hastalıklarını otomatik olarak teşhis etmek için kullanılır. Dermatoskopik görüntülerin alınmasında kullanılan çeşitli donanımsal aygıtlar mevcuttur. Derinin mikroskopla incelenmesi yöntemi olarak da tanımlanabilen dermatoskopi, cilt görüntülerinin bilgisayar destekli analiz edilmesi yöntemidir. Bu yöntemle vücuttaki benlerin haritası oluşturularak, her noktanın lokalizasyonu da belirlenmektedir.

Ardından her bir ben için dermatoskopik görüntü alınıp kaydedilmektedir. Böylece bir sonraki kontrolde elde edilecek yeni görüntünün, uzman hekim tarafından önceki halleri ile karşılaştırılabilmesi mümkün olmaktadır.

NASNet ailesinden bir çözüm çok tuzluya mal olsa da aslında temel düşüncesi yani modelin mimarisini datasetine göre optimum hale getirmeye çalışmak fikri çok önemli bir gelişmeyi doğurmuştur. EfficientNet olarak isimlendirilen bu yeni aile NASNet ten farklı olarak mimariyi maliyetli bir şekilde optimum hale getirmek yerine, sadece bir tane katsayı ile mimariyi ölçekleme yoluna gitmiştir. Örneğin; ölçeğin geçerli değeri ikiye katlanınca, mimarideki derinlik, katman çıkışlarındaki ara çıktılarının genişliği ve çözünürlüğü de ikiye katlanmaktadır. Birleşik ölçekleme de denilebilecek bu yeni strateji çok basit ve sade olmasına karşın oldukça etkili sonuçlar vermektedir. Çalışmamızda da EfficientNet mimarileri kullanarak iyi ya da kötü huylu Melanoma tespitinde hangisinin daha başarılı olduğunu tespit etmek temel amacımızdır.

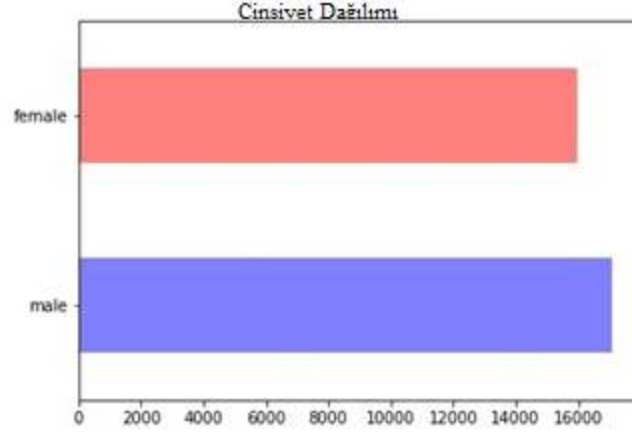
Bu çalışmada olduğu gibi; deri hastalıklarının teşhisinde yapay zekâ teknikleri kullanmak isteyen araştırmacıların ihtiyaç duyduğu dijital resimlerden oluşan en büyük uluslararası veri seti, ISIC(International Skin Imaging Collaboration) topluluğu tarafından açık kaynak olarak sunulmaktadır. Gerçekleştirdiğimiz lezyon sınıflandırma uygulamamız için dermoskopi cihazlarından alınan renkli deri görüntülerinden oluşan ISIC Arşivindeki HAM10000(Human Against Machine with 10.000 training images) veri setinden faydalanılmıştır.

Ham10000 veri setinde niteliklerine göre train.csv ve test.csv dosyaları başlığı altında görüntüler id,yaş, cinsiyet, iyi huylu ya da kötü huylu gibi kriterlere göre etiketlenmiştir. Bu verilerde TFRecord dosyaları içerisinde jpeg dosya türlerinde saklanmaktadır. Bir hastadan alınan tüm görüntüler tamamen tek bir Tensorflow kaydında üç katmanlı olarak TFRecord dosyası halinde bulunmaktadır. Tüm TFRecord'lar eşit veri sınıfı oranına sahiptir (yani her TFRecord artık %1,8 malign görüntü içerir). [43]

	image_name	patient_id	sex	age_approx	anatom_site_general_challenge	diagnosis	benign_malignant	target
0	ISIC_2637011	IP_7279968	male	45.0	head/neck	unknown	benign	0
1	ISIC_0015719	IP_3075186	female	45.0	upper extremity	unknown	benign	0
2	ISIC_0052212	IP_2842074	female	50.0	lower extremity	nevus	benign	0
3	ISIC_0068279	IP_6890425	female	45.0	head/neck	unknown	benign	0
4	ISIC_0074268	IP_8723313	female	55.0	upper extremity	unknown	benign	0

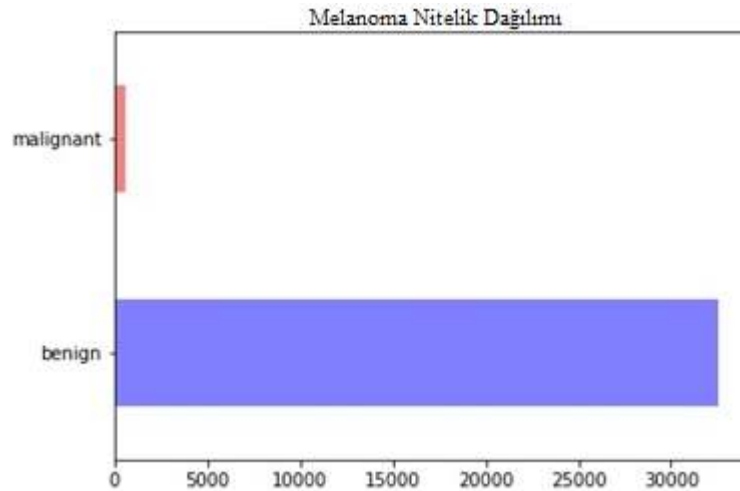
Şekil 4.4. Train.csv ve Test.csv Dosya İçerikleri

Daha sonra veri seti içerisinde cinsiyet, nitelik(iyi-kötü huylu) gibi özelliklere göre grafikler oluşturulmuştur. Bu grafikler Şekil 4.4, Şekil 4.5 ve Şekil 4.6’de gösterilmiştir.



Şekil 4.5.Cinsiyete Göre Veri Dağılımı

Eğitim veri setinde kötü huylu görüntülere kıyasla önemli ölçüde daha fazla iyi huylu görüntü olması dikkat çekici noktalardan biridir. Eğitim ve doğrulama setini bölerken ve çapraz doğrulama sırasında katmanlama yapmak açısından bu durum önemli olacaktır. Diğer bir avantaj ise, iyi huylu görüntülerin daha az ağırlıkta olması için kaybın önüne geçmekte faydalı olacaktır. Şekil.4.6’da görüldüğü gibi iyi huylu ve kötü huylu melanoma görüntüsü arasında sayıca farklılık mevcuttur. Bu durum kötü huylu melonomun tespitini kolaylaştırmaktadır.



Şekil 4.6.İyi-kötü Huylu Melanoma Veri Dağılımı

Veri seti ve csv dosyaları sisteme yüklenip eğitim ve doğrulama setleri sisteme tanıtıldıktan sonra daha hızlı ve verimli bir eğitim yapabilmek adına yapay zekâ işlemcisi TPU(Tensor Processing Unit) kurmakla işe başlanmıştır. Google'ın sunduğu bu teknoloji Yapay sinir ağları temelinden faydalanan ve bulut yapısı kullanılarak üretilen bu teknolojik yapılar yapay zeka uygulamalarını eğitmek için kullanılır.

TPU'lar yalnızca Tensorflowu desteklemektedirler. Matris hesaplamaları yoğun olan modellerdir. Çok büyük efektif kümeler üzerinde çok büyük modellerle çalışan yapılardır.

Eğitim veri setinde kötü huylu görüntülere kıyasla önemli ölçüde daha fazla iyi huylu görüntü olması dikkat çekici noktalardan biridir. Eğitim ve doğrulama setini bölerken ve çapraz doğrulama sırasında katmanlama yapmak açısından bu durum önemli olacaktır. Şekil.4.7'de veri setinde bulunan resimlerin gösterildiği melanoma özellikleri belirtilmiştir. Diğer bir avantaj ise, iyi huylu görüntülerin daha az ağırlıkta olması için kaybın önüne geçmekte faydalı olacaktır. Bazı hiper parametreler için birden fazla değer olması, her k-fold sırasında farklı modellerle eğitim ve denemeler yapmak içindir. Ayrıca son veri kümesinin eğitim aşamasında, hiper parametrelerin tümü için tüm k-foldlar aynı olmalıdır.

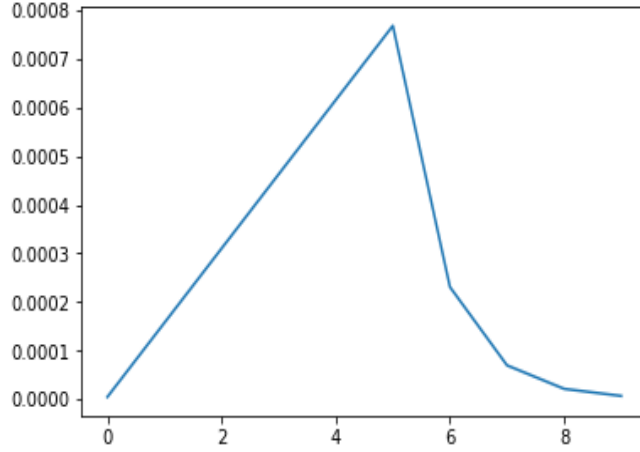


Şekil 4.7. Melanoma Nitelikleri

Makine öğrenimi için yaygın bir eğitim programı uygulamak gerekir. Öğrenme oranı sıfıra yakın başlar, sonra maksimuma çıkar, sonra zamanla azalır. Bu durum programın öğrenme oranlarını değiştirmektedir. Daha büyük veri boyutlarıyla öğrenme oranı daha fazla artmaktadır. Maksimum öğrenme oranının veri büyüklüğüyle doğru orantılı olarak arttığı tespit edilmiştir.

Get_lr_callback kullandığımız öğrenme oranı programı, bir döngü oluşturarak öğrenme oranı tespiti sağlamıştır. Öğrenme oranını bir başlangıç öğrenme oranından bir maksimum öğrenme hızına ve daha sonra bu maksimum öğrenme hızından başlangıç öğrenme hızından çok daha düşük bir minimum öğrenme hızına ayarlar. Bu politika ilk olarak Süper Yakınsama olarak literature geçmekte olup “Büyük Öğrenme Hızlarını Kullanarak Sinir Ağlarının Çok Hızlı

Eđitimi” olarak aıklanmaktadır. Őekil 4.8’de alıřmanın ğrenme zamanı gecikmesi grafiksel olarak gsterilmiřtir.



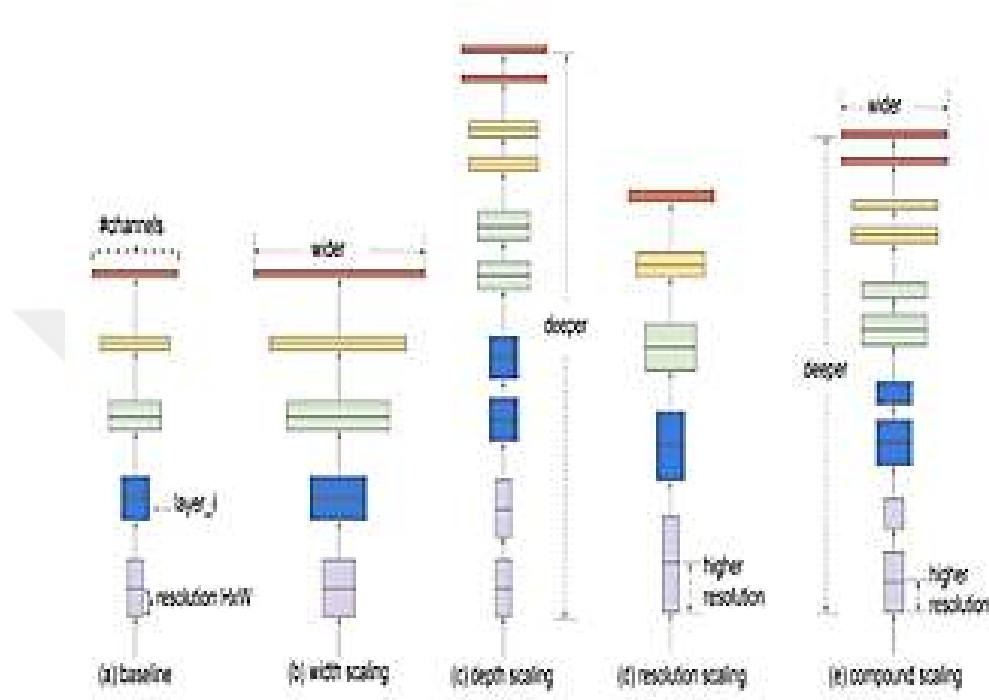
Őekil 4.8. Öğrenme Hızı Çizelgesi

Bu işlemlerden sonra daha ok veriyle daha ok bařarı elde etme amacıyla veri ođaltma işlemleri yapılmıřtır. Bu ařamada, tf.image komutu ile yapılamayan kırpma, yakınlılařtırma ve teleme gibi bazı manuel veri ođaltma işlemleri uygulanmıřtır. Resim dndürme yalnızca 90 derecelik faktrlerle yapılabilmektedir. Bu yzden bunun yerine işlem manuel olarak gerekleřtirilmiřtir. Ayrıca geometrik veri ođaltma işlemleri ile veri ođaltma gerekleřtirildi.

Evriřimsel sinir ađlarının eđitiminde, leklendirme genellikle modelin dođruluđunu artırmak iin yapıldıđı gibi aynı zamanda bir modelin verimliliđini artırmaya da yardımcı olmaktadır. Yani ne kadar hızlı eđitir ise o kadar verimlilik artmaktadır. Bir CNN'nin  leklendirme boyutu vardır; bunlar derinlik, geniřlik ve znrlktr. Derinlik, ađların ne kadar derin olduđunu ifade eder ve bu da iindeki katman sayısına eřittir. Geniřlik, ađın ne kadar geniř olduđunu ifade etmektedir. rneđin bir geniřlik ls Conv katmanındaki kanal sayısı iken znrlk sadece bir CNN'ye geirilen grnt znrlđdr.

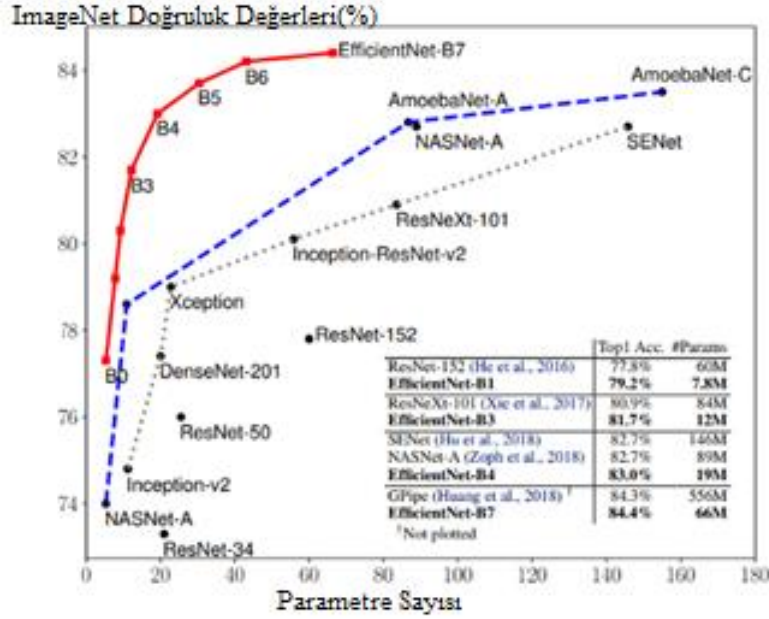
Derinlik leklendirme, bir ađın daha karmařık ve daha zengin zellikleri ğrenmesine izin verebilirken, ađ yapılarında atlamalı bađlantılar gibi teknikler kullanılır. Kaybolan gradyanlar nedeniyle derinlik ok bydke genellikle getiri az olur. te yandan geniřlik lekleme, modellerin kk olmasına izin verir. Daha geniř ađlar daha ince taneli zellikleri yakalama eđilimindedir. Daha kk oldukları iin daha hızlı eđitilirler. Bununla birlikte, performans yine daha geniř geniřlikte dřer. znrlk lekleme olduka sezgiseldir. Daha byk znrlkler kullanmak her zaman daha iyidir ancak iliřki dođrusal deđildir. Bu nedenle, ađın herhangi bir boyutunu (geniřlik, derinlik veya znrlk) bytmek dođruluđu artırır, ancak daha byk modeller iin dođruluk kazancı azalır [44].

EfficientNet, Bileşik Ölçekleme kullanma fikriyle oluşturulan bir mimaridir. Bu, ağ genişliğini, derinliğini ve çözünürlüğünü eşit olarak ölçeklendirmek için bir bileşik katsayı ϕ kullanır. ϕ , EfficientNets B1-B7'yi üretenler tarafından belirlenen bir katsayıdır. Şekil 4.9.'da Efficient mimarisi tüm detayları ile aktarılmıştır. [45].



Şekil 4.9. EfficientNet Model Mimarisi – Kaynak Makale:” EfficientNet:Rethinking Model Scaling for Convolutional Neural Networks” [46].

EfficientNet grubu B0-B7 arasında 8 tane modelden oluşur Sayı büyüdükçe hesaplanan parametre sayısı ve doğruluk artar. ImageNet veri setinde 2012 yılından sonra modeller karmaşıktıkça başarı yükselmiştir. Fakat işlem yükü bakımından bir çoğu efektif değildir. Son yıllarda daha küçük modellerle daha verimli yaklaşımlar benimsenmeye çalışılıyor. Öyle ki, modeli küçültmeye çalışırken derinlik, genişlik ve çözünürlükten yalnızca birinde değil her üçünde de ölçekleme yapmak çok daha efektif sonuçlar üretmeyi sağlıyor. EfficentNET B0-B7 ile diğer mimarilerin hesaplanan parametre ve doğruluk açısından ImageNET üzerindeki performans karşılaştırmaları Şekil 4.10’ da görülmektedir.



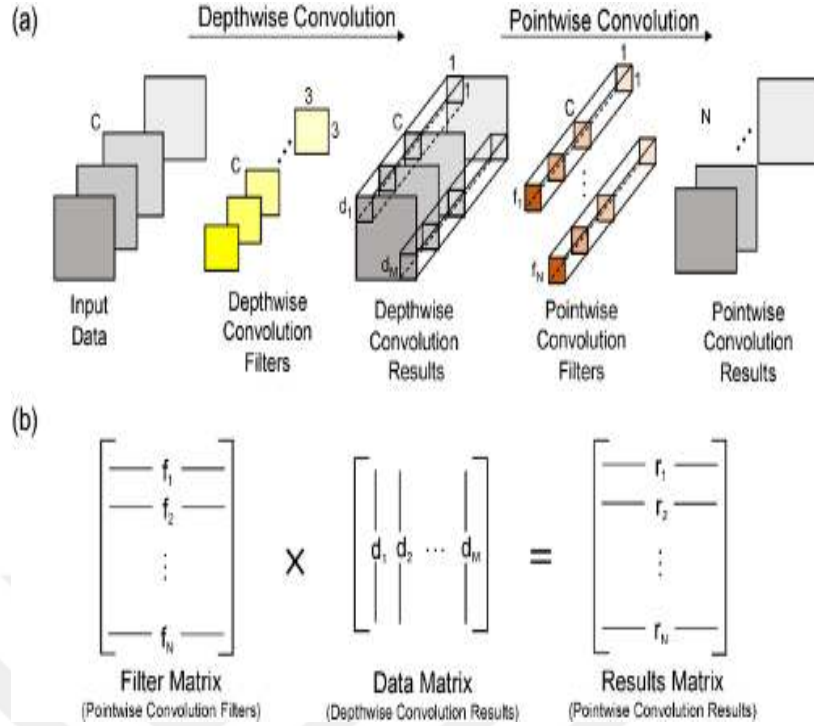
Şekil 4.10. Parametre ve ImageNet Doğruluk Karşılaştırması- Kaynak Makale :” EfficientNet:Rethinking Model Scaling for Convolutional Neural Networks” [46].

Depthwise Evrişim + Pointwise Evrişim: Hesaplama maliyetini asgari doğruluk kaybıyla önemli ölçüde azaltmak için orijinal evrişimi iki aşamaya bölmektedir.

Ters çevrilmiş Res: Orijinal ResNet blokları, kanalları sıkıştıran (squeeze) bir katmandan sonra kanalları genişleten (expand) bir katmandan oluşur. Böylece atlama bağlantıları (skip connections) ile zengin kanal katmanlarını bağlar. Ancak MBConv’da, bloklar önce kanalları genişleten, sonra bunları sıkıştıran bir katmandan oluşur, böylece daha az kanala sahip katmanlar atlanarak bağlanır.

Lineer darboğaz: ReLU’dan bilgi kaybını önlemek için her bloktaki son katmandaki lineer aktivasyonu kullanır.

EfficientNet için ana yapı taşı, ilk olarak MobileNetV2 olarak bilinen ters çevrilmiş darboğaz (bottleneck) MBConv’dur. Genişleme katmanlarına kıyasla çok daha az sayıda kanalı birbirine bağlayan darboğazlar arasındaki kısayolları kullanarak, hesaplamayı geleneksel katmanlara kıyasla neredeyse k^2 azaltan derinlemesine ayrılabilir evrişim ile birleştirilmiştir. k ’nın çekirdek boyutunu ifade ettiği için 2 boyutlu evrişim penceresinin yüksekliğini ve genişliğini belirtir. Şekil 4.11’de Depth ve Point Wise Evrişimlerinin gösterimleri şekilsel olarak ifade edilmiştir.



Şekil 4.11. Depth ve Point Wise Evrişimlerinin Temel Gösterimi – Kaynak: EfficientNet:Rethinking Model Scaling for Convolutional Neural Networks” [46].

Eğitim setinde yaptığımız veri büyütme benzer şekilde, Test Time Augmentation (TTA) gerçekleştirilmiştir. Bu işlemin amacı, test görüntülerinde rastgele değişiklikler yapmaktır. Gürültüden arındırılmış ve orijinal test görüntüleri üzerinde bir tahminde bulunmak için eğitilmiş modeli kullanmak yerine, görüntüler birkaç kez büyütülmüştür. Her biri üzerinde bir tahmin yapıldığına karşılık gelen her görüntünün ortalaması alınmıştır. Bu, o görüntü için elde ettiğimiz son tahmin olmaktadır. TTA'nın çalışmasının avantajı, tahminlerimizin ortalamasını aldığımız gibi hataların da ortalamasını almış olmasıdır. Doğrulama kümesinden daha fazla verim elde edebilmek için doğrulama veri kümesine Test Time Augmentation uygulanmıştır.

K-fold çapraz doğrulama gibi bir yeniden örnekleme tekniği kullanarak bir veri kümesinde bir makine öğrenimi algoritmasının performansını değerlendirmek yaygındır.

K-Fold Cross Validation’da veriler k farklı alt kümeye bölünmüştür. Veriyi eğitmek ve son alt kümeyi test verisi olarak bırakmak için k-1 adet alt küme kullanıldı. k adet deney sonucunda ortaya çıkan ortalama hata değeri modelimizin geçerliliğini belirtmektedir. K-fold çapraz doğrulama prosedürü, bir eğitim veri setini k gruba ayırmaktadır. Ardından kalan örnekler bir eğitim seti olarak kullanılırken bir doğrulama setinde k örnek grubunun her birini kullanmıştır. Bu, k farklı modelin eğitildiği ve değerlendirildiği anlamına gelir. Modelin performansı, tüm k-katlarında yapılan modellerin tahminleri kullanılarak tahmin edilir. k değeri genellikle 3 ya da 5 olarak seçilmektedir. Bu değer 10 ya da 15 de seçilebilir ancak bu oldukça maliyetli bir hesaplama ve zaman kaybına sebep olacaktır.

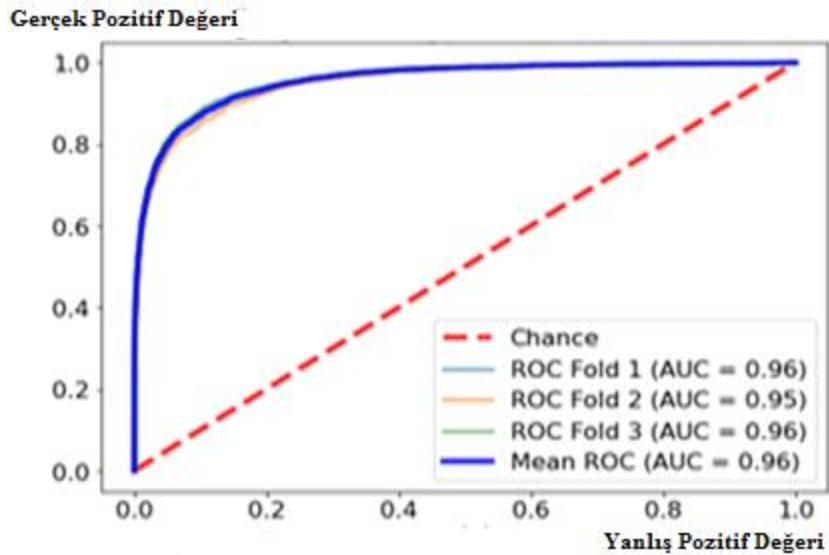
Son olarak, modeli eğitip değerlendirme aşamasında test setindeki en iyi tahmin elde edebilmek için her katın en iyi tahmin edildiği model kullanılmıştır. Tahminlerin daha sonra kat sayısı boyunca ortalaması alınmıştır.

İşlem sonunda doğrulama ve eğitim AUC'sinin birbirine oldukça yakın olduğu görülmüştür. Ancak bazı kısımlarda eğrilerin birbirinden ayrıldığı tespit edildi. Buda kaybın arttığını ve öğrenmenin yavaşladığını göstermektedir.

Sınıflar arasında dengeli dağılım olmadığı durumda iyi huylu görüntü sayısı, kötü huylu görüntülerin sayısını büyük ölçüde aşıyorsa, AUROC, modelin performansı hakkında aşırı iyimser bir görüş verebilir. Bu durum grafik açısından daha iyi olmaktadır.

Yani, yanlış pozitiflerin sayısındaki büyük bir gelişme, ROC kullanırken yalnızca yanlış pozitif oranında küçük bir değişikliğe yol açar. Hassasiyet ise, gerçek negatifler yerine yanlış pozitifleri gerçek pozitiflerle karşılaştırarak, çok sayıda negatif örneğin modelin performansı üzerindeki etkisini yakalar.

Büyük veri sınıfı dengesizliği olduğunda doğruluk iyi bir değerlendirme ölçüsü değildir. 100 örnek varsa: bunların 99'u melanoma ve 1'i ise normal olduğu varsayılırsa, o zaman her şeyi melanoma olarak tahmin eden bir model %99 doğruluk elde edecektir Bu durumda, kesinliğe, harmonik ortalamasına, F1 puanına bakmak daha iyi sonuçlar vermektedir. Şekil 4.12'de K-fold validation sonucunda elde edilen ROC eğrisi grafikleri verilmiştir.



Şekil 4.12. ROC Eğrisi

4.3. Toplu Öğrenme Yöntemiyle Evrimsel Sinir Ağlarının Birleştirilmesi

Sınıflandırma problemlerinde, çözüm kümesinin tahminlerinin birleştirilmesine topluluk öğrenmesi denir. Ayrıca içinde kullanılan toplama tekniğine topluluk yöntemi denir. Çalışmamızın bu bölümünde 3 yöntemin sonuçları, yani ortalama, oylama ve ağırlıklı birleştirme analiz edilmiştir. CNN modellerinin çıktılarının nasıl birleştirileceği sorusunun en basit cevabı ortalama almaktır. Ortalama, çözüm topluluklarında tercih edilebilecek basit ve etkili bir yöntemdir. Tüm modellerin çıktılarının aritmetik ortalaması basitçe hesaplanabilir olmuştur. Ortalama alma tekniği, yalnızca modellerin performansları iyiye işe yarayacaktır. Birbirine yakın olduğu durumlarda iyi aksi takdirde daha da kötü sonuçlar verir. Çalışmamızda başarı puanlarının CNN modelleri birbirinden çok uzak olmamakla birlikte, bu teknikle elde edilen birleştirme sonuçları da analiz edilmiştir.

Çoklu sınıflandırma problemlerinde (uygulamamızda yedi sınıf) daha verimli olması açısından birleştirme yöntemi tasarlanmalıdır. Bu karmaşık durumu ağırlıklı toplam tekniği ile çözmek mümkündür. Yüzlerce sınıflandırıcıdan oluşan bir aday topluluğu varsa, onların bir optimizasyon tekniği ile çıktılarını almak mümkündür. Örneğin; RandomForest tekniği ile optimum çıktılar almak için yüzlerce karar ağacından oluşan bir topluluk eğitebilir. Bu ağırlık değerlerini optimize etmeye çalışmak yerine her bir modelin nihai çıktı üzerindeki etkisini gösteren değerler, doğruluk puanlarına göre hesaplanabilmektedir. Doğrulama aşamasında ise; iyi tahminler yapan modellerin katsayıları daha yüksek olmaktadır. Sağlık hizmetlerinde en kötü durum senaryosu yanlış teşhistir (Yanlış Negatif). Hekimlerde kullanılan karar destek sistemleri, mümkün olduğunca çok sayıda kanser hastasının tespit edilerek tedavinin erken süreçte sağlanması istenmektedir. Bu çalışmada esas amacı ele alınan cilt kanseri hastalıklarını teşhis etme yeteneğini hızlıca sisteme kazandırmaktır. Bu durum hastanın erken bir aşamada hayatta kalma olasılığını büyük ölçüde artırır.

HAM10000 gibi dengeli dağılım göstermeyen veri kümelerine sahip uygulamalar geliştirirken doğruluk değerini metrik olarak seçmek gerekir. Bunun yerine, hatırlamayı hesaba katan bir ölçüm olan Yanlış Negatif oranı tercih edilmelidir. Her dönemin sonunda, hem eğitim hem de doğrulama alt kümesi için f1-skor değerleri kaydedildi. F1-skoru, kesinlik ve geri çağırma metriklerinin harmonik ortalamasıdır. Kesinlik metriği, eğitilmiş sınıflandırıcının çıktısında bir hasta olarak tahmin edilen kişinin gerçekten hasta olma olasılığını gösterirken, geri çağırma metriği, gerçekten hasta olan örneklerin yüzde kaçının tespit edilebileceğini gösterir. Modelimizi yanlış negatif tahminler için mümkün olduğunca cezalandırmak istediğimizden, değerlendirme için doğru kriter, isabet oranı olarak da adlandırılan hatırlama metriği kullanılmıştır.

Tablo 4.2. Doğrulama Verileri İçin CNN Modellerinin Elde Edilen F1-skor Değerleri

	MEL	NV	BCC	AKIEC	BKL	DF	VASC
ResNet152V2	0,53	0,91	0,67	0,59	0,55	0,48	0,77
InceptionV3	0,42	0,90	0,56	0,47	0,59	0,53	0,61
Xception	0,43	0,91	0,61	0,58	0,55	0,20	0,81
InceptionResNetV2	0,47	0,91	0,60	0,57	0,55	0,30	0,73
NASNetLarge	0,54	0,89	0,59	0,41	0,47	0,21	0,56
DenseNet	0,49	0,92	0,67	0,24	0,60	0,29	0,83
MobileNet	0,55	0,90	0,46	0,58	0,58	0,47	0,56
EfficientNetB0	0,54	0,90	0,77	0,56	0,53	0,29	0,87
EfficientNetB5	0,55	0,92	0,73	0,56	0,63	0,09	0,84
mean	0,50	0,91	0,63	0,51	0,56	0,32	0,73
std deviation	0,05	0,01	0,09	0,12	0,05	0,15	0,12

Tablo 4.2’de 7 hastalığa göre her satırın 7 boyutlu vektöre sahip olduğu bir f1-skor tablosu oluşturulmuştur. Tablo 4.2'deki her hücre, ilgili hastalık için doğrulama verileri üzerinde bir CNN modelinin f1-skorunu gösterir.

NV sütunu, HAM10000 veri kümesinde ana sınıf olması nedeniyle en büyük değerlere sahiptir. CNN'lerdeki çeşitlilik ve oluşturulan havuz bu tabloda rahatlıkla görülmektedir. Örneğin, InceptionV3 modeli, DF azınlık sınıfı için en iyi performansı gösterirken Efficient ailesi, NV hastalık etiketinde lider durumdadır.

Tablo 4.3. CNN Modellerinin Normalleştirilmesi Sonucu F1-skor Ağırlık Değerleri

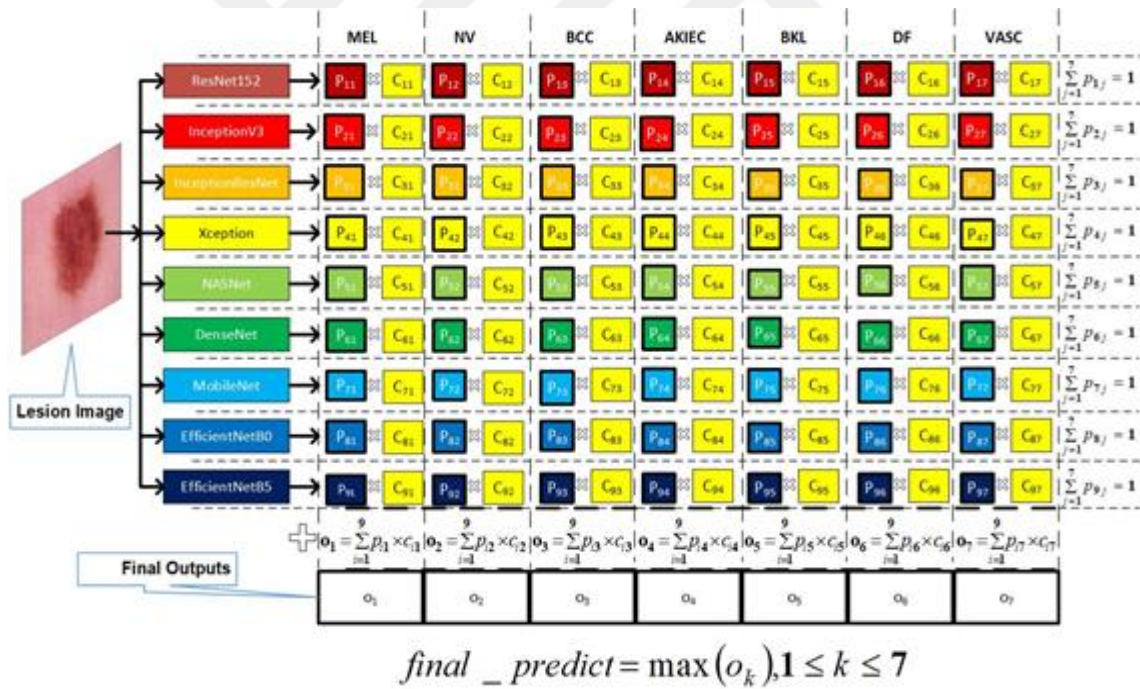
	MEL	NV	BCC	AKIEC	BKL	DF	VASC
ResNet152V2	$c_{11}=0,53$	$c_{12}=0,31$	$c_{13}=0,44$	$c_{14}=0,74$	$c_{15}=0$	$c_{16}=1,14$	$c_{17}=0,35$
InceptionV3	$c_{21}=0$	$c_{22}=0$	$c_{23}=0$	$c_{24}=0$	$c_{25}=0,72$	$c_{26}=1,42$	$c_{27}=0$
Xception	$c_{31}=0$	$c_{32}=0,09$	$c_{33}=0$	$c_{34}=0,63$	$c_{35}=0$	$c_{36}=0$	$c_{37}=0,61$
InceptionResNetV2	$c_{41}=0$	$c_{42}=0,43$	$c_{43}=0$	$c_{44}=0,53$	$c_{45}=0$	$c_{46}=0$	$c_{47}=0$
NASNetLarge	$c_{51}=0,88$	$c_{52}=0$	$c_{53}=0$	$c_{54}=0$	$c_{55}=0$	$c_{56}=0$	$c_{57}=0$
DenseNet	$c_{61}=0$	$c_{62}=1,11$	$c_{63}=0,47$	$c_{64}=0$	$c_{65}=0,76$	$c_{66}=0$	$c_{67}=0,83$
MobileNet	$c_{71}=0,89$	$c_{72}=0$	$c_{73}=0$	$c_{74}=0,60$	$c_{75}=0,32$	$c_{76}=1,02$	$c_{77}=0$
EfficientNetB0	$c_{81}=0,75$	$c_{82}=0$	$c_{83}=1,52$	$c_{84}=0,46$	$c_{85}=0$	$c_{86}=0$	$c_{87}=1,12$
EfficientNetB5	$c_{91}=0,81$	$c_{92}=1,46$	$c_{93}=1,06$	$c_{94}=0,45$	$c_{95}=1,53$	$c_{96}=0$	$c_{97}=0,88$

CNN modellerinin katsayı vektörlerini hesaplamak Tablo 4.3'deki her bir sütunun (hastalığın) ortalama ve standart sapması kullanılmıştır. Bu katsayılar birleştirme prosedüründe kullanılmaktadır. c_{ij} notasyonu, j. hastalık için i. CNN modelinin katsayı değeridir. i. CNN modeli j. hastalığın tahmininde söz sahibidir. c_{ij} değerinin 0'dan büyük olup olmaması göre nihai tahminde etkili olmuştur. NASNet'in doğrulama aşamasındaki f1-skor değeri tüm hastalıklar için ortalamanın altında sonuç vermiştir, Öte yandan, Tablo 4.2'te görüldüğü gibi, farklı CNN modelleri farklı hastalık tahminlerinde daha fazla söz sahibidir. Örneğin, MobileNet Melanoma cilt kanserinde en küçük etkiye sahip olmasına rağmen en büyük etkiye sahip model olarak

belirlenmiştir. EfficientNetB5 ise 6 hastalıkta da söz sahibi olmasına rağmen en az hastalığa sahip olan DF sınıfı için göz ardı edilmiştir.

$$c_{ij} = \frac{f1_score_{ij} - mean_j}{std_j} \quad | \quad c_{ij} = \begin{cases} c_{ij} & \text{if } c_{ij} > 0 \\ 0 & \text{if } c_{ij} \leq 0 \end{cases} \quad (2)$$

Şekil 4.13 önerilen topluluk yönteminin nasıl çalıştığını göstermektedir. İlk olarak, giriş lezyon görüntüsü uygulanmıştır. Denklem 2’den hareketle 7 hastalığa ait olma olasılığı [0-1] aralığındaki değerler bu modellerin çıktılarında hesaplanmıştır. p_{ij} notasyonu, i. modelinin görüntüyü j hastalığı olarak tahmin etme olasılığıdır. Softmax aktivasyon fonksiyonu son çıktı katmanında bulunduğundan CNN modellerinden herhangi bir satırdaki tüm olasılıkların toplamı 1'e eşittir. CNN tarafından yapılan hastalık tahminleri giriş lezyon görüntüsü için modeller, ilgili satırdaki en büyük değere sahip sütun başlığıdır.



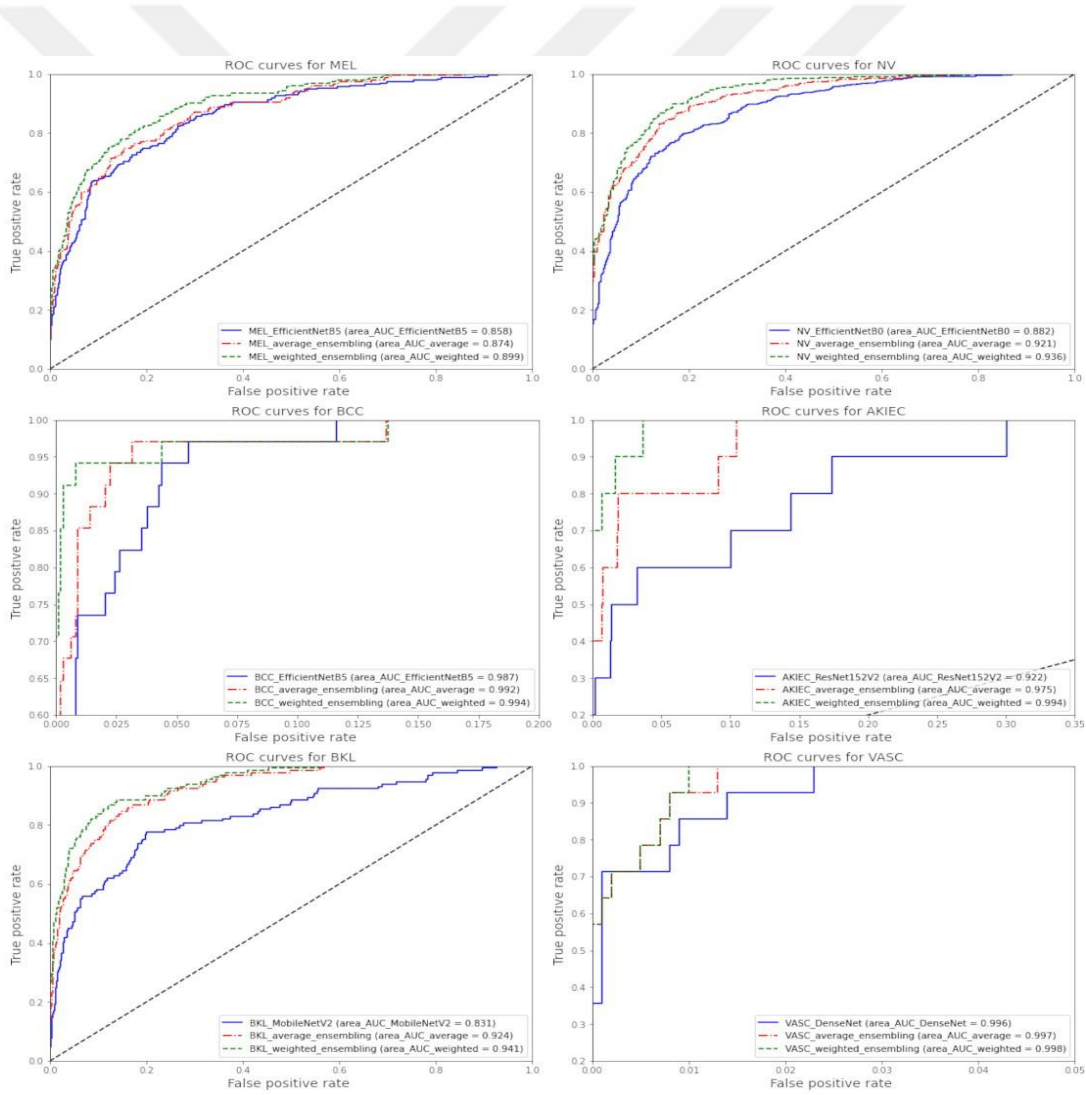
Şekil 4.13. Yeni ağırlıklı Birleştirme Yöntemi

Sonuç olarak, önerilen birleştirme yöntemini diğer birleştirme teknikleri ile karşılaştırmak için eğitim sürecinden HAM10000 veri setinde 1.015 lezyon görüntüsü izole edilmiştir. Öncelikle, CNN havuzundaki hangi modelin bireysel olarak en doğru tahminleri yaptığını belirlemek için bu yerel test verileri kullanıldı. Tablo 4.4'te gösterildiği gibi, farklı CNN

modelleri farklı hastalıkları teşhis etmede daha başarılı olmuştur. Tablo 4.4'te ölçüm metriği olarak duyarlılık (hit rate) veya geri çağırma olarak da adlandırılan metrik tercih edilmiştir.

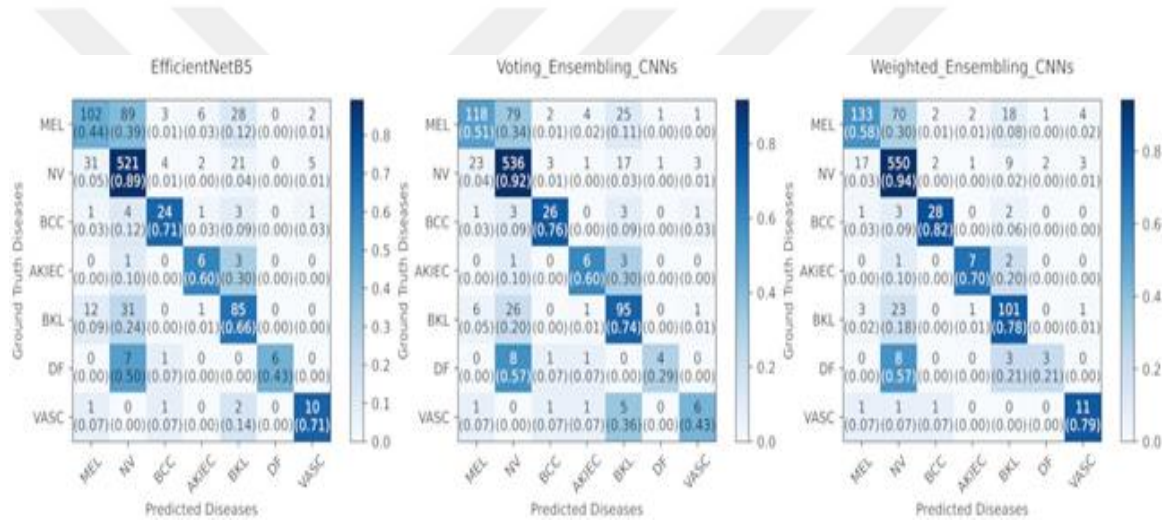
Tablo 4.4. HAM10000'de Her Hastalık İçin CNN Modellerinin Duyarlılık Değerleri

	MEL	NV	BCC	AKIEC	BKL	DF	VASC
ResNet152V2	0,6	0,6176	0,5891	0,8571	0,5391	0,9126	0,7857
InceptionV3	0,5	0,6470	0,6046	0,5	0,5	0,9212	0,7142
Xception	0,4	0,6176	0,5813	0,4285	0,5130	0,9332	0,7142
InceptionResNetV2	0,3	0,6176	0,6356	0,6428	0,5086	0,9674	0,6428
NASNetLarge	0,4	0,5588	0,2868	0,5714	0,6086	0,7294	0,5714
DenseNet	0,5	0,7352	0,6434	0,3571	0,5652	0,9229	0,8571
MobileNet	0,5	0,5882	0,6356	0,3571	0,6130	0,9743	0,7142
EfficientNetB0	0,5	0,8529	0,6511	0,6428	0,5608	0,8801	0,7142
EfficientNetB5	0,8	0,7352	0,6589	0,4285	0,5608	0,8921	0,7857



Şekil 4.14. Birleştirme yöntemleri ve CNN Modelleri Kullanılarak Çizilen ROC Eğrilerinin Karşılaştırılması

Şekil 4.15'te önerdiğimiz ağırlıklı birleştirme yöntemi ile ortalama olarak başarı düzeyinin üzerine çıkmıştır. Şekil 4.14'te grafiklerde görüldüğü üzere klasik yöntemle elde edilen sonuçların ROC eğrilerinde oluşan AUC değeri, önerdiğimiz birleştirme yönteminden ortalama %2-%5 arasında daha düşüktür. Karşılaştırdığımızda oylama tekniği ile elde edilen karışıklık matrislerinde köşegen hücrelerdeki değerler ve ağırlıklı yöntemde %10'un üzerinde bir iyileşme göze çarpmaktadır. CNN modellerinin bireysel performanslarına bakıldığında, EfficientNet ailesi en büyük 5. üyesi olan EfficientNetB5 modeli ile en üst sıradadır. HAM10000 veri setindeki görüntüler 450x450 boyutlarındadır. 6. ve 7. üyeler olan B0 ve B5 arasında %10 civarında önemli bir performans artışı meydana getirmiştir. Ancak, NV ana sınıfında B0 modeli, B5'ten daha iyi tahminler yapabilmektedir. Son zamanlarda yaygın olan küçük bir veri tabanı ile büyük performans elde etme çabaları çalışmada kendini göstermiştir.



Şekil 4.15. Birleştirme Yöntemleri Kullanılarak Elde Edilen Karışıklık Matrislerinin Bireysel CNN Modelleri İle Karşılaştırılması

5. SONUÇ

Bu tez çalışmasında son zamanların güncel çalışma alanlarından olan derin öğrenme yöntemleri kullanılarak evrişimsel sinir ağlarıyla hastalık teşhisinden bahsedilmiştir. Derin öğrenmenin yapısı, özellikleri ve mimarileri anlatılmıştır. Derin öğrenmenin sıklıkla kullanıldığı medikal alandaki çalışmalar incelenmiştir. Cilt kanseri tespiti ile ilgili farklı çalışmalar sonuçlarıyla beraber verilmiştir. Literatürde verilen çalışmalar incelendiğinde, konvolüsyonel sinir ağları kullanılarak geliştirilen yöntemlerin yüksek başarı oranı gösterdiği görülmektedir. Tez çalışması kapsamında, dermoskopi cihazlarından alınan renkli deri görüntüleri kullanılarak Evrişimsel sinir ağı tabanlı lezyon sınıflandırma yaklaşımı önerilmiştir. Önerilen yöntem için herkesin kullanımına açık en büyük dermatoloji veri setlerinden biri olan ISIC kullanılmıştır. ISIC veri seti üzerindeki çok sayıda görüntü üzerinden yapılan çalışmalar sonucu derin öğrenme modellerinin veri sayısı ile doğru orantılı olarak başarı gösterdiğini kanıtlamıştır. Derin öğrenme uygulamalarında yüksek başarı değerlerine ulaşmak için görüntü sayısının oldukça fazla olması gerekmektedir. Çalışma genelinde veri çoğaltma işlemi için geometrik dönüşümler kullanılmıştır. Tez çalışması sonucunda elde edilen bulgular derin öğrenme yapılarının melanomanın teşhisinde ciddi bir yardımcı kaynak olacağını göstermektedir.

Yarışın önde gelen mimarisi ise EfficientNetB5 modelidir. HAM10000 veri setindeki resimlerin boyutları 450x450 boyutlarında olup, bu mimarilerin 6 ve 7 numaralı Evrişimsel sinir ağlarının performans ölçümünü yapmaya gerek kalmamıştır. B0 modeli ile B5 modeli arasında %10 gibi önemli bir performans artışı gözlemlenmektedir. Modelimiz AUROC verisine göre bakıldığında oldukça iyi bir performans sergiliyor ancak AUPRC eğrisine göre ise iyileştirmeler yapılabileceği görülmektedir. Melanoma tanısında (lezyonları malign veya iyi huylu olarak sınıflandırmak için), tüm gerçek iyi huylu vakaları doğru bir şekilde tanımlamak istediğimiz için duyarlılık yüksek olmaktadır. Modelimiz için, yüksek bir geri çağırma elde etmek için düşük bir eşik ayarlandığında, hassasiyetimiz büyük ölçüde zarar görmektedir. Bu durum melanoma taraması sırasında birçok kişinin yanlış melanoma teşhisi alacağı anlamına gelmektedir. Verilerin dengeli dağılım göstermemiş olması işimizin zorluğunu arttırmıştır. Mevcut sorunu aşabilmek için yetersiz örneklenmiş sınıf için daha fazla veri toplamak ya da var olan çeşitli yöntemlerle artırma kanısına varmış bulunmaktayız.

Tüm modellerin en çok zorlandıkları hastalık sınıfı melanoma(MEL) olduğu görülmektedir. Sadece Melanoma Detection başlığı ile araştırma çalışmaları yapılmasının temel nedeni bu zorluk derecesidir. Melanoma sınıfına ait AUC değerinin en yüksek olduğu CNN modeli, EfficientNetB5 mimarisi olmuştur. EfficientNetB5 mimarisi diğer hastalık sınıfları için de yüksek AUC değerleri ile ön plana çıkmaktadır.

MobileNet mimarisi, boyutunun küçük olmasına rağmen EfficientNet ailesine yakın doğrulukta tahminler yapabilmektedir. AUC değerlerinin en düşük olduğu mimarinin NASNet olması değinilmesi gereken diğer bir sonuçtur. ResNet ve Inception ailelerin hibrid hali olan InceptionResNet mimarisinin beklendiği gibi atalarından daha iyi sonuç verdiği görülmektedir. Bunun yanı sıra bu ailenin yakın dostu olan Xception ise her iki mimarininde önüne geçmiştir. Model giriş ve çıkışlarındaki ağırlık sayılarının artması ile oluşturulan DenseNet mimarisinin de Inception ailesinden daha iyi sonuçlar doğurduğu görülmektedir.

Derin öğrenme uygulamalarında yüksek başarı değerlerine ulaşmak için veri tabanının oldukça yüksek sayıda olması gerekmektedir. Yani oldukça görüntüye ihtiyaç vardır. Ancak elimizdeki Uluslararası veri setindeki 10015 tane veri ile sistem başarıyla eğitilmiştir. Bu tez çalışması sonucunda elde edilen bulguların derin öğrenme yapılarının melanoma teşhisi ve tedavisinde önemli bir araç olabileceğini göstermektedir. Kullanılan veri setinin daha çok genişletilmesi, sınıf etiketlerinin artırılmasıyla sistemin performansının artacağı ortaya konulmuştur.

KAYNAKLAR

- [1] Do, T. T., Hoang, T., Pomponiu, V., Zhou, Y., Chen, Z., Cheung, N. M., . . . Tan, S. H. (2017, 11). Accessible Melanoma Detection using Smartphones and Mobile Image Analysis.
- [2] Okuboyejo, D. A., & Olugbara, O. O. (2018, 4). A Review of Prevalent Methods for Automatic Skin Lesion Diagnosis. *The Open Dermatology Journal*, 12, 14–53. doi:10.2174/187437220181201014
- [3] Brinker, T. J., Hekler, A., Enk, A. H., Klode, J., Hauschild, A., Berking, C., . . . Schrüfer, P. (2019, 4). A convolutional neural network trained with dermoscopic images performed on par with 145 dermatologists in a clinical melanoma image classification task. *European Journal of Cancer*, 111, 148–154. doi:10.1016/j.ejca.2019.02.005
- [4] Brinker, T. J., Hekler, A., Enk, A. H., Klode, J., Hauschild, A., Berking, C., . . . Schrüfer, P. (2019, 5). Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task. *European Journal of Cancer*, 113, 47–54. doi:10.1016/j.ejca.2019.04.001
- [5] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017, 1). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542, 115–118. doi:10.1038/nature21056
- [6] Li, Y., & Shen, L. (2018, 2). Skin Lesion Analysis towards Melanoma Detection Using Deep Learning Network. *Sensors*, 18, 556. doi:10.3390/s18020556
- [7] Majtner, T., Yildirim-Yayilgan, S., & Hardeberg, J. Y. (2018, 10). Optimised deep learning features for improved melanoma detection. *Multimedia Tools and Applications*, 78, 11883–11903. doi:10.1007/s11042-018-6734-6
- [8] Mishra, N. K., & Celebi, M. E. (2016, 1). An Overview of Melanoma Detection in Dermoscopy Images Using Image Processing and Machine Learning.
- [9] Barata, C., Celebi, M. E., & Marques, J. S. (2019, 5). A Survey of Feature Extraction in Dermoscopy Image Analysis of Skin Cancer. *IEEE Journal of Biomedical and Health Informatics*, 23, 1096–1109. doi:10.1109/jbhi.2018.2845939
- [10] Codella, N., Rotemberg, V., Tschandl, P., Celebi, M. E., Dusza, S., Gutman, D., . . . Halpern, A. (2019, 2). Skin Lesion Analysis Toward Melanoma Detection 2018: A Challenge Hosted by the International Skin Imaging Collaboration (ISIC).
- [11] Tschandl, P., Rosendahl, C., & Kittler, H. (2018, 3). The HAM10000 dataset, a large collection of multi-source dermoscopic images of common pigmented skin lesions. *Sci. Data* 5, 180161 (2018). doi:10.1038/sdata.2018.161.
- [12] Bisla, D., Choromanska, A., Berman, R. S., Stein, J. A., & Polsky, D. (2019). Towards Automated Melanoma Detection With Deep Learning: Data Purification and Augmentation. (pp. 2720–2728). Long Beach, CA, USA: IEEE. doi:10.1109/CVPRW.2019.00330
- [13] Vasconcelos, C. N., & Vasconcelos, B. N. (2017, 2). Convolutional Neural Network Committees for Melanoma Classification with Classical And Expert Knowledge Based Image Transforms Data Augmentation.
- [14] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., . . . Rabinovich, A. (2014, 9). Going Deeper with Convolutions.
- [15] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2015, 12). Rethinking the Inception Architecture for Computer Vision.
- [16] Chollet, F. (2016, 10). Xception:Deep Learning with Depthwise Separable Convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 1251–1258.
- [17] He, K., Zhang, X., Ren, S., & Sun, J. (2016, 2). Identity Mappings in Deep Residual Networks. Erişim 25.09.2021. <https://arxiv.org/abs/1905.11946>.
- [18] He, K., Zhang, X., Ren, S., & Sun, J. (2015, 12). Deep Residual Learning for Image Recognition. Erişim 11.09.2021. <https://arxiv.org/abs/1512.03385>.
- [19] He, K., Zhang, X., Ren, S., & Sun, J. (2016, 3). Identity Mappings in Deep Residual Networks. Erişim 12.09.2021. <https://arxiv.org/abs/1603.05027>.
- [20] Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. (2016, 2). Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning.
- [21] Zoph, B., Vasudevan, V., Shlens, J., & Le, Q. V. (2017, 7). Learning Transferable Architectures for Scalable Image Recognition.
- [22] Huang, G., Liu, Z., van der Maaten, L., & Weinberger, K. Q. (2016, 8). Densely Connected Convolutional Networks.

- [23] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018, 1). MobileNetV2: Inverted Residuals and Linear Bottlenecks. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 4510-4520.
- [24] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., . . . Adam, H. (2017, 4). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications.
- [25] Bengio, Y., Lamblin, P., Popovici, D., and Larochelle, H., 2007. Greedy Layer-Wise Training of Deep Networks, *Advances in Neural Information Processing Systems*
- [26] Zhenghao, S., Lifeng, H., Kenji S., Tsuyoshi N., and Hidenori I., 2009. Survey on Neural Networks Used for Medical Image Processing, *Int J Comput Sci.*, 3(1): 86–100.
- [27] Ying, W., and Rouzbeh R., 2015. An Introduction to Deep Learning. Examining the Advantages of Hierarchical Learning.
- [28] Kalinovskiy, A., and Kovalev, V., 2016. Lung image segmentation using Deep Learning methods and convolutional neural networks. In: XIII Int. Conf. on Pattern Recognition and Information Processing, Minsk, Belarus State University, 3-5 October.
- [29] Lo, S.C.B., Lou, S.L.A., Lin, J.S., Freedman, M.T., Chien, M.V., and Mun, S.K., 1995. Artificial convolution neural network techniques and applications for lung nodule detection, *IEEE Trans. Med. Imag.*, 14, no. 4, pp. 711-718.
- [30] Pratt, H., Coenen, F., Broadbent, D.M., Harding, S.P., and Zheng, Y., 2016. Convolutional neural networks for diabetic retinopathy, *Procedia Comput. Sci.*, 90, pp. 200-205.
- [31] Codella, N., Cai, J., Abedini, M., Garnavi, R., Halpern, A., and Smith, J.R., 2015. Deep learning sparse coding and svm for melanoma recognition in dermoscopy images, *International Workshop on Machine Learning in Medical Imaging*, pp. 118-126.
- [32] Esteva, A., Kuprel, B., Roberto, A., Novoa, Justin, K., Susan, M., Swetter, Helen M., Blau, and Sebastian, T., 2017. Dermatologist-level classification of skin cancer with deep neural Networks. *Research letter*, 542, pp.115-118.
- [33] Deng, L., and Dong Y., 2014. *Deep Learning Methods and Applications*, 198-349.
- [34] Bengio, Y., and LeCun, Y., 2007. In *Large Scale Kernel Machines*, MIT Press.
- [35] Karpathy, A., 2016. *Convolutional Neural Networks (CNNs / ConvNets)*. Retrieved CS231n Convolutional Neural Networks for Visual Recognition.
- [36] <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>. A Comprehensive Guide to Convolutional Neural Networks, 18 Mart 2021
- [37] ISIC, <https://www.isic-archive.com/#!/topWithHeader/wideContentTop/main>, Erişim Tarihi 25 Ekim 2020.
- [38] Kőszegi Balázs, Melanoma detection using deep learning technology , 2013, <https://pdfs.semanticscholar.org/ebf6/926101fbc704ed1f9e3d615eacd1138af433.pdf>
- [39] Zaballos, P., Puig, S., Llambrich, A. & Malvehy, J. Dermoscopy of dermatofibromas: a prospective morphological study of 412 cases. *Arch Dermatol* 144, 75–83 (2008).
- [40] Networks Noel Codella, Veronica Rotemberg, Philipp Tschandl, M. Emre Celebi, Stephen Dusza, David Gutman, Brian Helba, Aadi Kallou, Konstantinos Liopyris, Michael Marchetti, Harald Kittler, Allan Halpern: “Skin Lesion Analysis Toward Melanoma Detection 2018: A Challenge Hosted by the International Skin Imaging Collaboration (ISIC)”, 2018; <https://arxiv.org/abs/1902.03368>.
- [41] Dataİst Veri Analizi ,”Sınıflandırıcı Metrikleri”, Erişim 22 Ekim 2020, [https://tezverianaliz.com/biyoistatistik-dershanesi/ozgulluk-ve-duyarlilik-specificity-and-sensitivity1/#:~:text=%C3%96zg%C3%BCII%C3%BCK%20\(Specificity\)%20%3A%20%C4%B0ngilizcesi%20specificity,negatif\)%20te%C5%9Fhisi%20koyma%20oran%C4%B1n%C4%B1%20belirtir](https://tezverianaliz.com/biyoistatistik-dershanesi/ozgulluk-ve-duyarlilik-specificity-and-sensitivity1/#:~:text=%C3%96zg%C3%BCII%C3%BCK%20(Specificity)%20%3A%20%C4%B0ngilizcesi%20specificity,negatif)%20te%C5%9Fhisi%20koyma%20oran%C4%B1n%C4%B1%20belirtir).
- [42] Fourier Education, Ruye Wang 04.04.2016. “Performance Measurements”. Erişim 22 Ekim 2020. <http://fourier.eng.hmc.edu/e161/lectures/classification/node5.html>.
- [43] <https://www.kaggle.com/teyang/melanoma-detection-using-effnet-and-meta-data/data> , Erişim 22 Ekim 2020.
- [44] Li, Y., & Shen, L. (2018, 2). Skin Lesion Analysis towards Melanoma Detection Using Deep Learning Network. *Sensors*, 18, 556. doi:10.3390/s18020556
- [45] Majtner, T., Yildirim-Yayilgan, S., & Hardeberg, J. Y. (2018, 10). Optimised deep learning features for improved melanoma detection. *Multimedia Tools and Applications*, 78, 11883–11903.
- [46] Mingxing Tan, Quoc V. Le. *Machine Learning (cs.LG); Computer Vision and Pattern Recognition*. Erişim 25.08.2021. <https://arxiv.org/abs/1905.11946>.

YURT DIŐI DENEYİMLERİ VE SERTİFİKALAR

- ✓ Germany(16-23) Agust 2014 The Duke of Edinburgh's International Award Gold Member/GERMANY
- ✓ The Duke of Edinburgh's International Award / CERTIFICATE OF ATTENDANCE
- ✓ İş Analizi-Yazılım(Vb.net) Eğitimleri (NİLACCRA A.Ő)
- ✓ Zaman Yönetimi Eğitimi (Eğitim Sertifikası) (AKGEM)
- ✓ RYLA liderlik Eğitimleri / ROTARY
- ✓ Uygulamalı Araştırma ve Raporlama Teknikleri(SÜREKDER)
- ✓ Problem Çözme ve Çatışma yönetimi (AKADEMİKA)
- ✓ Doğru ve Etkili İfade ile Başarılı İletişim ve Beden Dili Eğitimi – PERYÖN
- ✓ Uygulamalı Münazara Teknikleri Eğitimi (TİKAV)
- ✓ Avrupa Birliği Proje Yazım Eğitimi(TİKAV)
- ✓ Menti-Mentör Menti Beceri Eğitimleri(TİKAV)
- ✓ Değişim Yönetimi ve İnovasyon Eğitimi (PERYÖN)
- ✓ Diksiyon ve Anlatım Bozuklukları Eğitimi (MEHPARE ÇELİK)
- ✓ Takım Çalışması Eğitimi(AKADEMİKA)
- ✓ Sunum Teknikleri Eğitimi (PERYÖN)
- ✓ Girişimcilik Eğitimi(TİKAV)
- ✓ Tiyatro, Konuşma Adabı, Hijyen ve Öz Bakım Eğitimleri (TİKAV)
- ✓ Protokol Kuralları Eğitimi (TDO-TİKAV)
- ✓ Kariyerimin Kontrolü Bende ve Mülakat Gelişim Eğitimleri (BİLİM İLAÇ)
- ✓ Herkes Eşit Herkes Farklı Eğitimi (TİKAV)
- ✓ İlk yardım Eğitimi (TİKAV)
- ✓ UGO Programı Altın Kampı Hazırlık Eğitimi(TİKAV)

SOSYAL SORUMLULUK PROJELERİ

- ✓ UGO (Uluslararası Gençlik Ödülü Programı) Altın Katılımcısı
- ✓ Gençlik Merkezi Gönüllü Eğitimci
- ✓ Fırat Üniversitesi GençTema Kulüp Başkanlığı(2011/-)
- ✓ TİKAV (Türkiye İnsan Kaynakları Vakfı) Bursiyerliği(2011-2015)
- ✓ The Duke of Edinburgh's International Award Altın Katılımcısı(2011-2015)
- ✓ Kızılay Gençlik Hizmetleri Gönüllülüğü(2012/-)
- ✓ Mersin Ortak Dilimiz Renkler Projesi Okul Boyama Faaliyeti