



Melanoma detection from dermoscopy images with deep learning methods: A comprehensive study

Oktay Yıldız

Department of Computer Engineering, Gazi University, Ankara, 06570, Turkey

Highlights:

- An automatic detection system for melanoma
- Comparative analysis for skin cancer detection
- Melanoma detection using convolutional neural network

Keywords:

- Melanoma detection
- Convolutional neural network
- Deep learning
- Machine learning
- Classification

Article Info:

Received: 20.06.2018

Accepted: 27.03.2019

DOI:

10.17341/gazimmfd.435217

Graphical/Tabular Abstract

Skin cancer is common and a serious disease, which can lead to death if not treated in time. Melanoma is the rarest and most dangerous type of skin cancer. It causes the most deaths. As in all diseases, early and correct detection of skin cancer are very important. Computer Aided Diagnosis systems can help physicians and patients make better decisions. Especially, machine learning and deep learning use effectively in Computer Aided Diagnosis systems. In this study, an automatic detection system for melanoma is suggested. To illustrate the advantage of the proposed CNN model C4Net, a comprehensive experimental study has been carried out. In addition, the proposed C4Net has been compared with not only the existing deep learning algorithms such as AlexNet, GoogLeNet, ResNet and VGGNet but also conventional machine learning algorithms such as Artificial neural networks, k-Nearest neighbor algorithm and Support vector machine. In experimental studies, C4Net, which is designed as deep neural network model for melanoma detection, has obtained more classification accuracy than other methods with 96.94%.

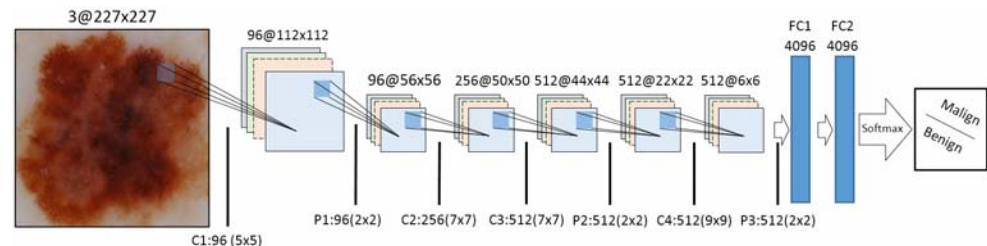


Figure A. Proposed CNN model for melanoma classification.

Purpose: In this study, an automatic detection system for melanoma is suggested. To illustrate the advantage of the proposed CNN model (C4Net), a comprehensive experimental study has been carried out.

Theory and Methods:

C4Net designed as deep neural network model for melanoma detection. In addition, the proposed model compared with the existing deep learning algorithms (i.e., AlexNet, GoogLeNet, ResNet and VGGNet) and conventional machine learning algorithms (i.e., ANN, kNN and SVM).

Results:

In experimental studies, C4Net, which is designed as deep neural network model for melanoma detection, has shown high classification accuracy than other methods.

Conclusion:

In this study, proposed CNN model has been obtained 96.94% classification accuracy. In future works, deep learning methods will be frequently used in the automatic detection of diseases such as melanoma.

Correspondence:

Author: Oktay Yıldız
e-mail: oyildiz@gazi.edu.tr
phone: +90 312 582 3139



Derin öğrenme yöntemleriyle dermoskopi görüntülerinden melanom tespiti: Kapsamlı bir çalışma

Oktay Yıldız*

Gazi Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Maltepe, Ankara, 06570, Türkiye

Ö N E Ç İ K A N L A R

- Melanom için otomatik tanı sistemi
- Cilt kanseri tespiti için karşılaştırmalı analiz
- Konvolüsyonel sinir ağı kullanarak melanom tespiti

Makale Bilgileri

Araştırma Makalesi
Geliş: 20.06.2018
Kabul: 27.03.2019

DOI:

10.17341/gazimmfd.435217

Anahtar Kelimeler:

Melanom tespiti,
konvolüsyonel sinir ağı,
derin öğrenme,
makine öğrenme,
sınıflandırma

ÖZET

Cilt kanseri yaygın görülen ve tedavi edilmemesi durumunda ölüme neden olan ciddi bir hastalıktır. Melanom ise nadir görülmesine rağmen ölüme en çok neden olan cilt kanseri türüdür. Tüm hastalıklarda olduğu gibi cilt kanserinin erken ve doğru tespit edilmesi hayati önem taşımaktadır. Bilgisayar destekli tanı sistemleri cilt kanseri tespitinde hekimlere ve hastalara yardımcı olabilir. Bilgisayar destekli tanı sistemlerinde özellikle makine öğrenmesi ve derin öğrenme modelleri etkin bir şekilde kullanılmaktadır. Gerçekleştirilen bu çalışmada cilt kanseri türü olan melanom için otomatik tanı koyabilecek bir sistem önerilmektedir. Melanom tanısı için tasarlanan C4Net derin sinir ağ modeli ile beraber literatürde ön plana çıkmış AlexNet, GoogLeNet, ResNet, VGGNet derin öğrenme modelleri ve Yapay sinir ağları, En yakın komşu algoritması ve Destek vektör makinesi gibi geleneksel makine öğrenme yöntemlerini de kapsayan detaylı bir deneysel çalışma yapılmıştır. Gerçekleştirilen deneysel çalışmalarda, tasarlanan C4Net derin sinir ağ modeli %96,94 doğrulukta melanom tanısı koyarak, diğer yöntemlere göre daha yüksek sınıflandırma başarısı göstermiştir.

Melanoma detection from dermoscopy images with deep learning methods: A comprehensive study

H I G H L I G H T S

- An automatic detection system for melanoma
- Comparative analysis for skin cancer detection
- Melanoma detection using convolutional neural network

Article Info

Research Article
Received: 20.06.2018
Accepted: 27.03.2019

DOI:

10.17341/gazimmfd.435217

Keywords:

Melanoma detection,
convolutional neural
network,
deep learning,
machine learning,
classification

ABSTRACT

Skin cancer is common and a serious disease, which can lead to death if not treated in time. Melanoma is the rarest and most dangerous type of skin cancer. It causes the most deaths. As in all diseases, early and correct detection of skin cancer are very important. Computer Aided Diagnosis systems can help physicians and patients make better decisions. Especially, machine learning and deep learning use effectively in Computer Aided Diagnosis systems. In this study, an automatic detection system for melanoma is suggested. To illustrate the advantage of the proposed CNN model C4Net, a comprehensive experimental study has been carried out. In addition, the proposed C4Net has been compared with not only the existing deep learning algorithms such as AlexNet, GoogLeNet, ResNet and VGGNet but also conventional machine learning algorithms such as Artificial neural networks, k-Nearest neighbor algorithm and Support vector machine. In experimental studies, C4Net, which is designed as deep neural network model for melanoma detection, has obtained more classification accuracy than other methods with 96.94%.

1. GİRİŞ (INTRODUCTION)

Malign melanom cilt dokusunda meydana gelen ve ölüme neden olabilen bir cilt kanseri türüdür [1]. Erken tanı ile hastalık tedavi edilebilir. Tıbbi hastalıkların tespitinde sıklıkla kullanılan Bilgisayar Destekli Tanı (Computer Aided Diagnosis - CAD) sistemleri cilt kanserinin tespitinde de hastalara ve hekimlere yardımcı olabilir.

CAD sistemlerinde makine öğrenmesi yöntemleri sıklıkla kullanılmaktadır [2-8]. Ye ve arkadaşları yapmış oldukları çalışmada [2], Hepatoselüler karsinom (Hepatocellular carcinoma - HCC) olarak adlandırılan, yaygın ve ölümcül olabilen karaciğer hastalığında, denetimli makine öğrenmesi yöntemlerini kullanarak, metastaz ve hasta sağ kalımı ile ilgili genleri tespit etmişlerdir. Kononenko yapmış olduğu çalışmada [3], tıpta tanı ve teşhiste kullanılan makine öğrenmesi yöntemlerini kapsamlı bir şekilde ele almıştır. Berikol ve arkadaşları [4] ise denetimli makine öğrenmesi yöntemlerini kullanarak Akut Koroner Sendromu (Acute Coronary Syndrome - ACS) tanısını yüksek doğrulukta gerçekleştirebilmişlerdir. Ancak son zamanlarda Derin Öğrenme (Deep Learning - DL) modelleri tıbbi görüntü sınıflandırma da dâhil olmak üzere birçok alanda başarılı bir şekilde uygulanmış ve makine öğrenme yöntemlerinin yerini almaya başlamıştır [5-8]. Litjens ve arkadaşları [5], prostat ve meme kanseri evrelemesinde derin öğrenme modellerinin başarılı bir şekilde kullanılabileceğini göstermişlerdir. Liu ve arkadaşları [6] ise Alzheimer hastalığı prodromal ve hafif bilişsel bozukluk aşamasını belirlemek için yeni bir derin öğrenme modeli tasarlamışlardır. Wang ve arkadaşları [7], otomatik medikal görüntü segmentasyonu için yeni bir Konvolüsyonel Sinir Ağ modeli önermişlerdir. Önerilen bu yeni model mevcut yöntemlere kıyasla daha az kullanıcı etkileşimli ve daha kısa sürede tıbbi görüntü segmentasyonu sağlayabilmektedir. Acharya ve arkadaşları [8], EEG sinyallerinin analizi için konvolüsyonel sinir ağı kullanılmışlar, 13 katmanlı bu ağ modeli yüksek doğrulukta sınıflandırma başarısı elde etmiştir.

Cilt doku görselleri gibi tıbbi görüntülerden hastalık tanı ve tespiti temel olarak bir görüntü sınıflandırma problemidir. Görüntü sınıflandırmada görüntüye ait temel özelliklerin yakalanması oldukça önemlidir. Görüntü sınıflandırmada en sık yaşanan problemler, görüntülerin açık bir tanıma sahip olmaması, farklı dokusal özellikler içermesi, görüntü özelliklerinden kaynaklanan gürültü ve bununla birlikte hesaplama karmaşıklığı ile yüksek zaman maliyeti sayılabilir. Bütün bunlar sınıflandırma başarısını olumsuz etkiler. Ayrıca tıbbi bir görüntü yüzlerce pikselden oluşabilir. Bir yapay sinir ağı (Artificial Neural Network-ANN) modeli düşünüldüğünde her pikselin bir nörona bağlanması hem yüksek maliyet hem de yüksek bellek gerektirir. ANN gibi klasik makine öğrenmesi yöntemleri görüntü sınıflandırma için girdi olarak görüntüleri tanımlayıcı özniteliklere ihtiyaç duyarlar. Son yıllarda yapılan araştırmalar göstermiştir ki DL modelleri, görüntü sınıflandırma da dâhil olmak üzere birçok alanda makine

öğrenmesi yöntemlerine göre daha iyi tahmin performansı göstermektedir. DL modelleri görüntü verisini temsil eden özellikleri çıkarmak için fazladan ön işleme ihtiyaç duymazlar. Ancak DL modelleri büyük miktarda veriye ihtiyaç duyar ve yüksek hesaplama maliyetine sahiptirler [9,10]. DL modelleri içinde en dikkat çeken Konvolüsyonel Sinir Ağlarıdır (Convolutional Neural Network-CNN). CNN, ANN'ye dayanan başarılı bir DL modelidir. Bazı çalışmalar görüntü sınıflandırmada CNN'nin geleneksel Makine Öğrenme (Machine Learning-ML) yöntemlerinden daha üstün olduğunu göstermektedir [11-14]. Görüntü tanıma problemlerinde ML ile karşılaştırıldığında DL, özellikle CNN, sınıflandırma performansı ile ön plana çıkmaktadır. Bu sebeple CNN görüntü tanıma problemlerinde yaygın olarak kullanılmaktadır. 2012 yılında Krizhevsky ve arkadaşları tarafından önerilen AlexNet, başarılı CNN modellerinden biridir [15]. Bu model, 60 milyon parametrelili, 650 bin nöron ve arkasından havuzlama katmanı bulunan beş konvolüsyon katmanından meydana gelmektedir. AlexNet, LSVRC-2010 ImageNet verisi içinde 1,2 milyon yüksek çözünürlüklü görüntüyü 1000 farklı sınıfa ayırmak için kullanılmış ve elde edilen %15,3 hata oranı ile daha önce aynı veri kümesinde yapılan çalışmalara göre yüksek bir sınıflandırma başarısı göstermiştir. Szegedy ve arkadaşları [16], "inception" olarak adlandırdıkları; farklı boyutlardaki filtreleri tek bir filtrede birleştiren dokuz modül ile iki konvolüsyon ve havuzlama katmanından meydana gelen GoogLeNet'i önermişlerdir. Bu mimari 2014 yılında ImageNet yarışmasında %6,67 hata oranı ile birinci olmuştur. Görüntü sınıflandırmada CNN'lerin başarısı sahip oldukları derinlik ile ilişkilidir. Ancak derinlik arttıkça eğitim için gerekli olan hesaplama maliyeti de artar. 2015 yılında Simonyan ve arkadaşları [17] mevcut CNN modellerinin derinliğini artırarak VGGNet'i geliştirmişlerdir. VGGNet 16 ila 19 katmandan ve 3x3 gibi çok küçük boyutlu filtrelerden meydana gelmektedir. Bu sebeple ağın eğitimi diğer CNN modellerine göre çok daha uzundur. 2015 yılında He ve arkadaşları [18] VGGNet'ten sekiz kat daha derin olan ancak daha düşük bir hesaplama maliyetine sahip ResNet adında yeni bir CNN modeli önermişlerdir. ResNet 2015 yılı ImageNet yarışmasında %3,57 hata oranı ile oldukça büyük bir başarı elde etmiştir.

Görüntü sınıflandırmada CNN'lerin göstermiş olduğu yüksek başarı, araştırmacıları bilgisayar destekli tanı sistemlerinde de CNN kullanmaya teşvik etmiştir [19]. Esteva ve arkadaşları tarafından 2017 yılında yapılmış olan çalışmada [20], cilt kanseri tanısı için CNN kullanılmışlar, 129.450 klinik görüntü ile eğitilen CNN'nin yüksek doğrulukta bir sınıflandırma yaptığını göstermişlerdir. Yazarlar 299x299 ölçeklendirdikleri görüntüler üzerinde Google'ın Inception v3 mimarisin kullanarak %72,1 ± 0,9 başarı elde etmişlerdir. Codella ve arkadaşları dermoskopi görüntülerinde melanom tanısı için DL-SVM tabanlı melez bir model sunmuşlardır [21]. Bu çalışmada Codella ve arkadaşları CNN için Berkeley tarafından geliştirilen Caffe CNN modelini tercih etmişlerdir. Bu modelin eğitimi için ISIC tarafından sağlanan 334 melanom, 144 atipik nevüs ve

2146 benign lezyonları barındıran toplam 2.624 görüntü kullanılmıştır. Melanom ve melanom dışı tüm lezyonlar ile melanom ve atipik lezyonlar ayrımı için iki farklı deney gerçekleştirilmiştir. Melanom ve melanom dışı tüm lezyonlar için yapılan deneysel çalışmada Caffe CNN %91,9 ve DL-SVM ile %93,1'lik sınıflandırma başarısı elde edilirken, melanom ve atipik lezyonlar ayrımı için yapılan çalışmada Caffe CNN %72,3 ve DL-SVM %73,9'luk sınıflandırma başarısı elde etmiştir. Önerilen melez yöntem melanom tanısı için az da olsa sınıflandırma başarısını arttırmıştır.

Sunulan bu çalışmada cilt kanseri tanısı için yüksek doğrulukta, otomatik tespit eden yeni bir CNN modeli önerilmiştir. Önerilen bu model, geleneksel makine öğrenmesi yöntemleri ile beraber dört farklı derin öğrenme modeli ile kıyaslanmıştır. Gerçekleştirilen çalışmada ISIC [22] tarafından derlenen 3920 dermoskopi görüntü kullanılmıştır. Etkin ve yüksek doğruluğa sahip CNN modeli için denemeler yapılmıştır. Tasarlanan CNN modeli (C4Net), literatürde sıklıkla kullanılan AlexNet, GoogLeNet, ResNet, VGGNet gibi CNN temsilcileri ve ANN, kNN ve SVM gibi geleneksel makine öğrenmesi yöntemleri ile karşılaştırılmıştır.

2. TEORİK ALTYAPI (THEORETICAL BACKGROUND)

2.1. Derin Öğrenme (Deep Learning - DL)

DL ham veriden özellik temsilini otomatik çıkarabilen, yapay sinir ağlarından türetilmiş yeni bir makine öğrenme yöntemidir. Çoğu durumda klasik makine öğrenme yöntemlerinin aksine veri ön işleme gerektirmez [23]. Bunun yerine DL, alt düzey özelliklerin bir bileşimi ile daha yüksek hiyerarşi özelliklerine sahip karakteristik hiyerarşileri öğrenir [24]. Böylece DL karmaşık, yüksek boyutlu problemlerin çözümünde başarılı bir şekilde kullanılmaktadır. Konvolüsyonel Sinir Ağları (Convolutional Neural Network-CNN) en başarılı derin öğrenme modellerinden biridir [10].

2.1.1. Konvolüsyonel sinir ağları (Convolutional neural network-CNN)

Literatürde çok sayıda CNN modeli bulunmaktadır; AlexNet [15], GoogLeNet [16], VGGNet [17] ve ResNet [18] dikkat çeken CNN modelleridir. Bununla birlikte bu modellerin temel bileşenleri birbirine çok benzer. 1998 yılında LeCun ve arkadaşları tarafından önerilen ünlü LeNet-5 [25] bugün ki CNN modellerinin temel yapısını oluşturmaktadır.

CNN, her biri bir veya daha fazla düzlem içeren çoklu bir katman yapısına sahiptir. Bu katmanlar; Konvolüsyon (Convolution), Havuzlama (Pooling) ve Tam bağlı (Fully-Connected) katman olarak adlandırılır. Konvolüsyon katmanında, her bir giriş için belirli özellikleri çıkaran bir dizi filtre uygulanır. Havuzlama, doğrusal olmayan alt örnekleme yaparken, ağırlık öğrenmesi gereken parametre sayısını azaltarak çıkışı basitleştirir. Böylece sadece aktive edilen özellikler bir sonraki katmana iletilir. Bu çoklu

katmanlı yapı öznelikleri öğrendikten sonra sınıflandırmaya geçer. Son katman c boyutlu tam bağlı bir katmandır. Her bir sınıfın olasılığı bu vektörde tutulur. CNN'nin başka bir özelliği paylaşılan ağırlıklardır. Bu, bir görüntünün farklı noktalarında yer alan piksel kümesinin özellik haritasında aynı ağırlık vektörüne sahip olacağı anlamına gelir. Her bir özellik haritası yerel alt örneklem çıkaran bir katman tarafından gerçekleştirilir [10].

Giriş katmanı normalleştirilmiş girdileri içerir. Bir katmandaki her düzlem, önceki katman düzleminin küçük bir girdisini alır. Her bir düzlem, bir önceki katmandaki düzlem üzerinden yerel bir pencere ile taranarak elde edilen sabit bir özellik haritası olarak düşünülebilir. Böylece çoklu özellikler tespit edilebilir. Bir özellik tespit edildikten sonra, kesin konumu daha az öneme sahiptir. Bu nedenle, konvolüsyon katmanları tipik olarak bir yerel ortalama ve alt örneklem işlemi yapan başka bir katman tarafından takip edilir. Sonunda, tüm piksel değerlerinin kombinasyonu bir vektöre dönüştürülür [25]. l katmanı ve k özellik haritasındaki (i, j) lokasyonunun değeri $Z_{i,j,k}^l$, Eş. 1'de gösterildiği gibi hesaplanabilir.

$$Z_{i,j,k}^l = w_k^{lT} x_{i,j}^l + b_k^l \quad (1)$$

Burada w_k^l ve b_k^l sırasıyla l katmanı k özellik haritasındaki ağırlık vektörü ve bias'dır. $x_{i,j}^l$ l katmanı (i, j) lokasyonunun ortalama giriş değeridir. w_k^l ağırlık vektörünün ağırlığında paylaşıldığını unutmamak gerekir. Böylece model karmaşıklığı azaltılırken, ağırlık daha kolay eğitilmesi gerçekleştirilir. Bütün bunlarla birlikte aktivasyon fonksiyonu ile CNN'nin doğrusal olmayan özellikleri tanınması sağlanabilir. Konvolüsyon özellik $Z_{i,j,k}^l$ için aktivasyon değeri $a_{i,j,k}^l$ Eş. 2'de gösterildiği gibi ifade edilebilir.

$$a_{i,j,k}^l = a(Z_{i,j,k}^l) \quad (2)$$

Sıklıkla kullanılan tipik aktivasyon fonksiyonları; Sigmoid, Tanh ve ReLU'dir [10].

CNN'lerin son katmanı çıkış katmanıdır. Çıkış katmanında sınıflandırma görevleri için softmax operatörü yaygın olarak kullanılmaktadır [10].

N giriş-çıkış ilişkisi $\{(x^n, y^n); n \in [1, \dots, N]\}$ şeklinde tanımlanır, x^n , n ci girişi, y^n karşılık gelen etiketi, o^n CNN'nin çıkışı temsil eder ve tüm parametreler Θ ile ifade edilirse, CNN kaybı Eş. 3 de gösterildiği gibi hesaplanabilir.

$$L = \frac{1}{N} \sum_{n=1}^N l(\Theta; y^n, o^n) \quad (3)$$

CNN'nin eğitimi bir hata-kayıp fonksiyonuna bağlı olarak gerçekleştirilir. Verilen q örnekleri için $e = (x, y_b)$ varsayımında, $Y^q = F(x, w)$ şeklinde bir kayıp fonksiyonu tanımlanabilir. Burada w ayarlanabilir ağırlığı temsil etmektedir.

$$E^q(w) = D(|Y^q - y_b|, F(x, w)) \quad (4)$$

Eş. 4 kullanılarak CNN tarafından üretilen çıktı ile olması gereken gerçek çıktı arasındaki fark bulunabilir. Bu çalışmada Minimum Ortalama Kareysel Hata (Minimum Mean Squared Error-MSE) (Eş. 5) tercih edilmiştir.

$$E^q(w) = \frac{1}{n} \sum_{q=1}^n (Y^q(F(x, w)) + \log(e^{-j} + \sum_i e^{-y^q(F(x, w))})) \quad (5)$$

Eş. 5'te y^q , q ncü katman çıktısını ifade etmektedir.

2.2. Makine Öğrenmesi (Machine Learning - ML)

2.2.1. Yapay sinir ağları (Artificial neural networks-ANN)

ANN mimarisi, Çok Katmanlı Perceptronun (Multilayer Perceptron - MLP) tipik olarak katman sayısı ve her katmana ait aktivasyon fonksiyonu ile farklı birimler arasındaki bağlantıların ağırlıklandırılmasından meydana gelir. Bir ANN eğitimi Eş. 6'da gösterildiği gibi ifade edilebilir.

$$f_{mtp}(x) = \theta \sum_{j=1}^n v_j \theta \left(\sum_{i=1}^m w_{ij} x_j + b_0 \right) + b_1 \quad (6)$$

Burada w_{ij} sinaptik ağırlıklar, b_0 ve b_1 bias ve θ genellikle sigmoid olan aktivasyon fonksiyonudur.

2.2.2. Destek vektör makinesi (Support vector machine-SVM)

Denetimli makine öğrenime yöntemi olan Destek Vektör Makinesi (Support Vector Machine-SVM), sınıflandırmak için bir hiper düzlemden faydalanır. Genellikle farklı kategorilere ait örnekleri sınıflandırmak için kullanılır. Hiper düzlem iki farklı sınıfı birbirinden ayırır ve iç bölgede örnek bulunmayan paralel iki çizginin maksimum genişliği sınır olarak tanımlanır. Destek vektörleri ise hiper düzleme en yakın veri noktalarıdır. Bir SVM için en iyi hiperdüzlem iki sınıf arasında en büyük mesafeye sahip olmalıdır.

2.2.3. k En yakın komşu (k-Nearest neighbor-kNN)

k-En Yakın Komşu (k-Nearest Neighbor - kNN) yaygın olarak kullanılan bir makine öğrenme yöntemidir. Denetimli öğrenme paradigmasına ait sınıflandırıcılar gibi kNN de $h(x)$ değerleri bilinen veri noktalarını $x \in X$ içeren bir $P \subseteq X$ eğitim verisine ihtiyaç duyar. P bilgisi kullanılarak, sınıflandırıcının yeni örneğin sınıf etiketini tahmin etmesi beklenir. kNN, iyi performansı yanı sıra basitliği sebebiyle yaygın olarak kullanılmaktadır. Ayrıca parametrik olmayan bir sınıflandırıcı olan kNN, veri dağılımı ile ilgili önceden yapılmış olan varsayımlara bağımlı değildir. Bununla birlikte, başarılı bir sınıflandırma için kNN k değeri,

komşuları belirlemek için kullanılan mesafe metriği ve örnek boyutu gibi üç önemli faktöre ihtiyaç duyar.

2.3. Doku Analizi (Texture Descriptor)

Görüntü sınıflandırmada doku analizi oldukça önemlidir. Doku analizi, bize görüntülerin yapısal özellikleri hakkında bilgi verir. Doku analizi için pek çok yöntem önerilmiştir. Bunlar içinde en yaygın kullanılanı Haralick ve arkadaşları [26] tarafından önerilen doku tanımıdır. Haralick dokuyu bir görüntüde gri seviye olasılık dağılımı için uzamsal etki alanından çıkarılan bir dizi özellik olarak tanımlar. Doku tanımlaması için yapısal ve istatistiksel olarak iki genel yaklaşım vardır. Yapısal yaklaşım için en temel örnek görüntünün Fourier dönüşümüdür. İstatistiksel yaklaşım için Haralick eş oluşum matrisi en iyi performansı vermektedir. Ayrıklaştırılmış bir görüntü $P^{m,n} = [i_{x,y}]$ Gauss rastgele alan şeklinde farz edilirse, $x, y \in \mathbb{Z}$ olmak üzere $i_{x,y}$ ilgili konumdaki pikselin gri düzeyini gösterir. Eş oluşum matrisi, belirli açı ve mesafede aynı parlaklığa sahip piksellerin sayılması ile elde edilir. $i_{x,y} < 2^8$ ve $i \in \mathbb{N}$ için n gri seviye Eş. 7'de gösterildiği gibidir.

$$b_i = \frac{2^8}{i}, i = 1, 2, \dots, n \quad (7)$$

Burada Uzaysal Gri Seviye Bağımlılık matrisi (Spatial Gray Level Dependence matrix - SGLD) SGLD = $[p_{i,j} | d, \theta]$, şeklinde ifade edilebilir. SGLD d uzaklık ve θ açısı için bileşik olasılık yoğunluk fonksiyonunun tahmini değeridir.

$$[p_{i,j}] = \sum_1^n \sum_1^n (i_{x,y} \in b_i) \cdot (i_{x,y} \in b_j) \quad (8)$$

Bu matris (Eş. 8) aralarındaki uzaklık d , açı θ iken piksellerin i gri seviyesinden j gri seviyesine geçme olasılığını gösterir. Genellikle θ açı değeri 0° , 45° , 90° ve 135° ve d ise 1 veya 2 olacak şekilde belirlenir. SGLD matrisinden görüntünün doku karakteristiği çıkarılabilir. Bu 5 karakteristik entropi (Eş. 9), enerji (Eş. 10), karışıklık (Eş. 11), korelasyon (Eş. 12) ve homojenliktir (Eş. 13).

$$\text{Entropi} = \sum_{i,j} p_{i,j} \log_2 p_{i,j} \quad (9)$$

$$\text{Enerji} = \sum_{i,j} p_{i,j}^2 \quad (10)$$

$$\text{Karışıklık} = \sum_{i,j} |i - j|^2 p_{i,j} \quad (11)$$

$$\text{Korelasyon} = \sum_{i,j} \frac{(1 - \mu_x)(1 - \mu_y)p_{i,j}}{\sigma_x \sigma_y} \quad (12)$$

$$\text{Homojenlik} = \sum_{i,j} \frac{p_{i,j}}{1 + |i - j|} \quad (13)$$

3. OTOMATİK MELANOM TESPİTİ (AUTOMATED MELANOMA DETECTION)

Bu bölümde, dermoskopi veri seti, kullanılan modellerin performans değerlendirme ölçütleri ve geleneksel makine

öğrenmesi yöntemleri için özellik çıkarımı ile sınıflandırıcılardan bahsedilmiştir.

3.1. Model Performans Değerlendirme (Model Performance Evaluation)

Model performans ölçütleri, Tablo 1’de verilen metriklere göre değerlendirilmiştir. Burada; Pozitif ve negatif etiketli iki sınıflı sınıma verisi için P pozitif etiketli örnekleri, N negatif etiketli örnekleri, DP sınıflandırıcı tarafından doğru bulunan pozitif etiketli örnek sayısını, DN sınıflandırıcı tarafından doğru bulunan negatif etiketli örnek sayısını, YP sınıflandırıcının yanlışlıkla pozitif bulduğu örnek sayısını, YN sınıflandırıcının yanlışlıkla negatif bulduğu örnek sayısını gösterir.

3.2. Dermoskopi veri kümesi (Dermoscopic Dataset)

Deneysel çalışmada International Skin Imaging Collaboration (ISIC) [22] tarafından sunulan dermoskopi görüntüleri kullanılmıştır. ISIC arşivi, uluslararası ve önde gelen klinik merkezlerden toplanan 13.000’den fazla dermoskopi görüntü içermektedir. Bu görüntüler 8 bit RGB ve 540×722 ila 4499×6748 pikselden oluşan farklı boyutlara sahiptir. ISIC arşivinden sunulan çalışma için 2000 malign (melanom) 1920 benign (iyi huylu) olmak üzere 3920 dermoskopi görüntü seçilmiştir. Şekil 1’de malign ve benign doku örnekleri görülmektedir.

3.3. Özellik Çıkarımı (Feature Extraction)

Görüntü sınıflandırmada, CNN’nin aksine ML yöntemleri görüntüyü tanımlayan özelliklere ihtiyaç duyarlar. Diğer bir ifadeyle CNN ham görüntü girdisinden özellik çıkarımı yaparken, geleneksel ML yöntemleri girdi olarak görüntüyü tanımlayan bir dizi özellik bekler. Bu çalışmada ML sınıflandırıcılarının eğitimi için literatürde yaygın bir şekilde kullanılan Haralick’in doku özellikleri kullanılmıştır. Özellik çıkarımı şu adımlardan oluşur:

1. *Adım* : Görüntünün gri seviye dönüştürülmesi.

2. *Adım* : Girdi bitlerinin belirlenmesi. Görüntünün gri seviye çözünürlüğüdür. Bu çalışmada 256 gri seviye değeri için 8 bit belirlenmiştir.
3. *Adım* : SGLD matrisini hesaplamak için pikseller arasındaki mesafe d değerinin belirlenmesi. Bu çalışmada $d=1$ kabul edilmiştir.
4. *Adım* : SGLD matrisini hesaplamak için θ açısı değerinin belirlenmesi. Bu çalışmada 0° , 45° , 90° ve 135° ile SGLD matrisi oluşturulmuştur.
5. *Adım* : SGLD matrisi oluşturulması.
6. *Adım* : Entropi, enerji, karışıklık, korelasyon ve homojenlik değerlerinin elde edilmesi.



Şekil 1. Malign ve benign doku görüntü örnekleri
(Malignant and benign skin lesions)

ML sınıflandırıcıları için girdi oluşturacak her bir görüntü örneği; entropi, enerji, karışıklık, korelasyon ve homojenlik değerlerinden oluşan beş Haralick özelliğinden oluşur.

3.4. Sınıflandırma (Classification)

Hekimler, melanom tanısı için beş özelliğin gözlemine dayalı ABCDE metodu olarak da bilinen bir yöntemi tavsiye

Tablo 1. Sınıflandırıcı performans ölçütleri (Classification Performance Metrics)

Ölçüt	Formül	Açıklama
Doğruluk (Accuracy)	$\text{Doğruluk} = \frac{DP + DN}{P + N}$	Bir sınıflandırıcının başarısını göstermek için kullanılan en yaygın ölçüttür. Doğru tahminlerin toplam örnekleme oranıdır.
Hassasiyet (Sensitivity)	$\text{Hassasiyet} = \frac{DP}{DP + YN}$	Doğru pozitif oranı. Sınıflandırıcı tarafından bulunan doğru pozitif örneklerin, tüm pozitif örneklere oranıdır.
Özgünlük (Specificity)	$\text{Özgünlük} = \frac{DN}{DN + YP}$	Doğru negatif oranı. Sınıflandırıcı tarafından bulunan doğru negatif örneklerin, tüm negatif örneklere oranıdır.
Kesinlik (Precision)	$\text{Kesinlik} = \frac{DP}{DP + YP}$	Sınıflandırıcı tarafından bulunan doğru pozitif sayısının, pozitif olarak sınıflandırılan örneklerin toplamına oranıdır.

etmektedirler. A (Asymmetric) asimetrik doku, B (Border) sınırlar, C (Color) renk, D (Diameter) çap, E (Evolving) değişim gösterme [27]. Gerçekleştirilen çalışmada, bu özellikler dikkate alınarak dermoskopi görüntüleri üzerinden melanom tespiti amaçlanmıştır.

3.4.1. Konvolüsyonel sinir ağı (Convolutional neural network-CNN)

Melanom tanısı için tasarlanan C4Net derin sinir ağı modeli ile beraber literatürde yaygın olarak kullanılan AlexNet, VGGNet, ResNet ve GoogLeNet gibi CNN modelleri içinde deneysel çalışma yapılmıştır. Gerçekleştirilen çalışmada C4Net için tipik CNN modeli doğrudan kullanılmamış, sınıflandırma başarısını arttırmak için bazı parametreler yeniden ayarlanmış, çeşitli denemeler yapılarak en uygun CNN topolojisi belirlenmiştir. Tablo 2’de tasarlanan C4Net topolojisi özetlenmiştir.

Tablo 2. Tasarlanan C4Net topolojisi
(Proposed C4Net topology)

Katman	Giriş	Filtre	Adım	Filtre
CV1	227x227	5x5	2	96
P1	112x112	2x2	2	96
CV2	56x56	7x7	1	256
CV3	50x50	7x7	1	512
P2	44x44	2x2	2	512
CV4	22x22	9x9	3	512
P3	6x6	2x2	2	512

Tablo 2’de *Katman*: Tasarlanan CNN modelinde kullanılan konvolüsyon, havuzlama ve tam bağlı katmanları ifade eder. *Giriş boyutu*: Görüntüler matris olarak ifade edilir. Her bir katman girdi olarak aldığı görüntünün boyutuna sahiptir. Diğer bir ifadeyle giriş boyutu, matris boyutudur. *Filtre*

boyutu: Özellik haritası çıkarımında filtreler önem arz etmektedir. Filtre, giriş matrisi üzerinde adım miktarınca kaydırılarak iki boyutlu özellik haritası çıkarımında kullanılır. *Adım (Stride)*: Filtre kaydırma miktarıdır. Özellik haritası çıkarımında filtrelerle birlikte kullanılır. Adım değeri daha küçük hacimde çıktı üretmeyi sağlar. *Filtre sayısı*: Özellik haritası olarak da ifade edilir. Her bir katman için kullanılan filtre sayısını gösterir. Farklı filtre sayısı görüntünün farklı özelliklerinin yakalanmasına yardımcı olur. Yapılan deneysel çalışmalarda konvolüsyon katmanı, havuzlama ve tam bağlı katman gibi temel CNN bileşenleri korunmuş, topoloji Python ve C ++ kullanılarak gerçekleştirilmiştir. Birinci konvolüsyon katmanı Python dilinde Şekil 2’de gösterildiği gibi kodlanmıştır. Bu katmanda; Filtre sayısı 96, Filtre boyutu 5x5, Adım 2x2 ve aktivasyon fonksiyonu Relu kullanılmaktadır. Çıkış ve tam bağlı katmanlar Şekil 3’te de gösterildiği gibi kodlanmıştır.

Melanom sınıflandırma için önerilen C4Net Şekil 4’te görüldüğü gibi dört konvolüsyon, üç havuzlama, iki tam bağlı katman ve çıkış olmak üzere on katmandan meydana gelmektedir.

Konvolüsyon katmanı

CV1 birinci konvolüsyon katmanı RGB (Red-Green-Blue) formatında 227x227 bir matris olan doku görüntülerini girdi olarak alır. Bu katmanda filtre boyutu 5x5, filtre sayısı 96 ve adım (stride) 2 olarak belirlenmiştir. Konvolüsyondan sonra, 112x112 boyutunda 96 özellik eşlemesi oluşturulur.

CV2 ikinci konvolüsyon katmanı kendinden önce yer alan P1 Havuzlama katmanından gelen 96@56x56 veriyi girdi olarak alır. Bu katmanda filtre boyutu 7x7, filtre sayısı 256 ve adım 1 olarak belirlenmiştir. Konvolüsyondan sonra, 50x50 boyutunda 256 özellik eşlemesi oluşturulur. CV3 üçüncü konvolüsyon katmanı kendinden önce yer alan CV2

```
model = Sequential()
# 1nci Konvolüsyon katmanı
model.add(Conv2D(filters=96, input_shape=(227,227,3),
kernel_size=(5,5), strides=(2,2), Activation('relu')))
```

Şekil 2. Birinci konvolüsyon katmanı (First convolution layer)

```
# 1nci Tam bağlı katman
model.add(Dense(4096))
model.add(Activation('relu'))
# 2nci Tam bağlı katman
model.add(Dense(4096))
model.add(Activation('relu'))
# Çıkış katmanı
model.add(Dense(2))
model.add(Activation('softmax'))
```

Şekil 3. Çıkış ve Tam bağlı katmanlar (Output and fully connected layer)

katmanından gelen $256@50 \times 50$ veriyi girdi olarak alır. Bu katmanda filtre boyutu 7×7 , filtre sayısı 512 ve adım 1 olarak belirlenmiştir. Konvolüsyondan sonra, 44×44 boyutunda 512 özellik eşlemesi oluşturulur.

CV4 dördüncü konvolüsyon katmanı kendinden önce yer alan P2 havuzlama katmanından gelen $512@22 \times 22$ veriyi girdi olarak alır. Bu katmanda filtre boyutu 9×9 , filtre sayısı 512 ve adım 3 olarak belirlenmiştir. Konvolüsyondan sonra, 6×6 boyutunda 512 özellik eşlemesi oluşturulur.

Havuzlama katmanı

Havuzlama katmanları, sınıflandırıcıdaki gizli katmanlar arasındaki bağlantı sayısını azaltarak hesaplama yükünü azaltır. Havuzlama katmanları P1, P2 ve P3, birinci, üçüncü ve dördüncü konvolüsyon katmanları arkasında yer alır. Havuzlama operatörü maksimum, filtre boyutu 2×2 ve adım 2 olarak belirlenmiştir. Konvolüsyon katmanı ve havuzlama katmanı ile beraber görüntü öznelikleri çıkarılır.

Tam bağlı katman ve çıkış katmanı

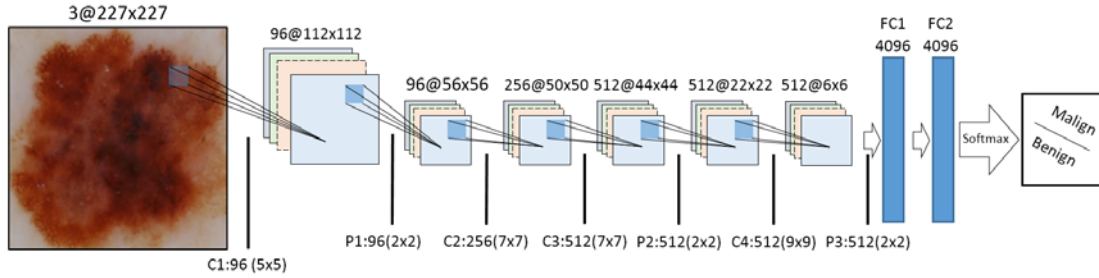
Dört konvolüsyon ve üç havuzlama katmanından sonra özellik gösterimi üzerinde yüksek düzeyde çıkarım

amaçlayan FC1 ve FC2, iki tam bağlı katman yer alır. FC1 ve FC2 4096 nörondan oluşmaktadır. İki tam bağlı katman arkasından çıkış katmanı gelmektedir. Çıkış katmanı malign ve benign kararının verildiği katmandır.

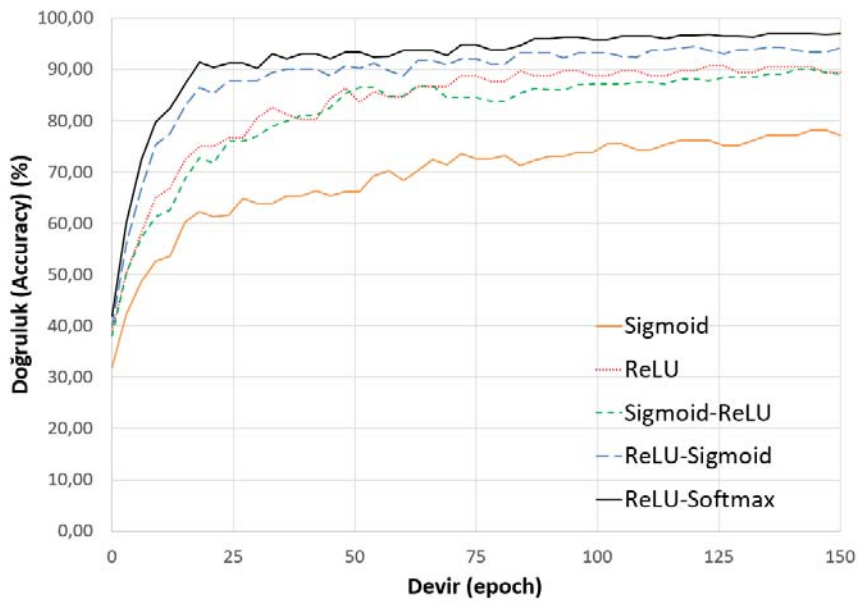
Aktivasyon fonksiyonu

Derin sinir ağlarında her nöron doğrusal veya doğrusal olmayan aktivasyon fonksiyonlarına sahip olabilir. Uygun aktivasyon fonksiyonlarının belirlenmesi ağıın genelleme performansında önemli rol oynar. Literatürde sıklıkla kullanılan aktivasyon fonksiyonları Sigmoid, Tanh, ReLU ve Softmax sayılabilir. En uygun aktivasyon fonksiyonu için beş muhtemel durum belirlenmiştir. Gerçekleştirilen çalışmada, uygun aktivasyon fonksiyonu seçimi için çeşitli denemeler yapılmıştır.

Gerçekleştirilen senaryolar için elde edilen aktivasyon fonksiyonu devir-doğruluk eğrisi Şekil 5'te gösterildiği gibidir. Birinci senaryo tüm gizli katmanlarda Sigmoid kullanıldığında en yüksek doğruluk oranı %77,77 elde edilmiştir. İkinci senaryo tüm gizli katmanlarda ReLU kullanıldığında en yüksek doğruluk oranı %88,30 elde edilmiştir. Üçüncü senaryo konvolüsyon katmanlarında Sigmoid, tam bağlı katmanda ReLU kullanıldığında en



Şekil 4. Melanom sınıflandırma için önerilen CNN modeli (Proposed CNN model for melanoma classification)



Şekil 5. Aktivasyon fonksiyonu devir-doğruluk eğrisi (Epoch-accuracy curve for activation functions)

yüksek doğruluk oranı %89,80 elde edilmiştir. Dördüncü senaryo konvolüsyon katmanlarında ReLU, tam bağlı katmanda Sigmoid kullanıldığında en yüksek doğruluk oranı %94,30 elde edilmiştir ve beşinci senaryo konvolüsyon katmanlarında ReLU, çıkış katmanında Softmax kullanıldığında en yüksek doğruluk %96,94 elde edilmiştir. Bu senaryolar içinde en yüksek doğruluk konvolüsyon katmanlarında ReLU ve çıkış katmanında Softmax kullanıldığında elde edildiği için önerilen CNN modeli buna göre tasarlanmıştır.

4. DENEYSEL ÇALIŞMA (EXPERIMENTAL STUDY)

Gerçekleştirilen deneysel çalışmalarda 10-kat çapraz geçişleme kullanılmıştır. Doğru ve yanlış sınıflandırılan Malign (M) ve Benign (B) doku örnekleri ile sınıflandırıcıların doğruluk, hassasiyet, özgünlük ve kesinlik değerleri ayrı ayrı tablolarda sunulmuştur.

10-kat çapraz geçişleme deneyleri için 3920 dermoskopi görüntü rastsal, 10 eşit parçaya bölünmüş, her bir sınıflandırıcı için aynı eğitim ve test veri kümesinin kullanılmasına dikkat edilmiştir. Böylece sınıflandırıcıların adil bir şekilde kıyaslanması sağlanmıştır.

4.1. Geleneksel Makine Öğrenmesi Sonuçları (Conventional Machine Learning Results)

4.1.1 MLP sonuçları (MLP results)

MLP katman ve nöron sayısı deneme yanılma ve en düşük hata ile beraber çalışma süresine göre belirlenmiştir. Buna göre 10 gizli katmandan oluşan geri yayımlı (Back Propagation Networks- BPN) ağ modeli ile %92,27 sınıflandırma doğruluğu elde edilmiştir. Tablo 3'te MLP için 10-kat çapraz geçişleme sonuçları, Tablo 4'de ise sınıflandırıcının doğruluk, hassasiyet, özgünlük ve kesinlik değerleri görülmektedir.

4.1.2 kNN sonuçları (kNN results)

kNN sınıflandırmada örnekler arasındaki mesafeyi hesaplamak için birçok farklı ölçüm kullanılabilir. Bu çalışmada Öklid mesafesi ve $k=5$ kullanarak en yüksek sınıflandırma başarısı elde edilmiştir. kNN ile %88,62 sınıflandırma doğruluğu elde edilmiştir. Tablo 5'te kNN için 10-kat çapraz geçişleme, Tablo 6'da ise sınıflandırıcının doğruluk, hassasiyet, özgünlük ve kesinlik değerleri görülmektedir.

Tablo 3. MLP 10-kat çapraz geçişleme sonuçları (MLP 10-fold cross validation results)

Test	Gerçek	Tahmin		Test	Gerçek	Tahmin	
		M	B			M	B
1	M	188	12	6	M	187	13
	B	16	176		B	18	174
2	M	187	13	7	M	188	12
	B	16	176		B	17	175
3	M	189	11	8	M	186	14
	B	15	177		B	19	173
4	M	187	13	9	M	185	15
	B	18	174		B	19	173
5	M	187	13	10	M	188	12
	B	19	173		B	18	174

Tablo 4. MLP sınıflandırma sonuçları (MLP classification results)

Test	Doğruluk %	Hassasiyet	Özgünlük	Kesinlik
1	92,86	0,940	0,917	0,922
2	92,60	0,935	0,917	0,921
3	93,37	0,945	0,922	0,926
4	92,09	0,935	0,906	0,912
5	91,84	0,935	0,901	0,908
6	92,09	0,935	0,906	0,912
7	92,60	0,940	0,911	0,917
8	91,58	0,930	0,901	0,907
9	91,33	0,925	0,901	0,907
10	92,35	0,940	0,906	0,913
Ortalama	92,27	0,936	0,909	0,915

4.1.3 SVM sonuçları (SVM results)

SVM ile %93,72 sınıflandırma doğruluğu elde edilmiştir. Tablo 7’de SVM için 10-kat çapraz geçişleme, Tablo 8’de ise sınıflandırıcının doğruluk, hassasiyet, özgünlük ve kesinlik değerleri görülmektedir.

4.2. CNN Sonuçları (CNNs Results)

4.2.1. C4Net Sonuçları (C4Net results)

C4Net ile 150 devir (epoch), 10-kat çapraz geçişleme ile ortalama sınıflandırma doğruluğu %96,94 elde edilmiştir.

Tablo 9’da C4Net için 10-kat çapraz geçişleme, Tablo 10’da ise sınıflandırıcının doğruluk, hassasiyet, özgünlük ve kesinlik değerleri görülmektedir.

Şekil 6’da C4Net için 10-kat çapraz geçişleme sonucunda elde edilen doğruluk grafik gösterimi, Şekil 7’de ise C4Net için 10-kat çapraz geçişleme sonucunda elde edilen hassasiyet grafik gösterimi yer almaktadır.

4.2.2. AlexNet sonuçları (AlexNet results)

AlexNet, 150 devir (epoch), 10-kat çapraz geçişleme ile ortalama sınıflandırma doğruluğu %93,93 elde edilmiştir.

Tablo 5. kNN 10-kat çapraz geçişleme sonuçları (kNN 10-fold cross validation results)

Test	Gerçek	Tahmin		Test	Gerçek	Tahmin	
		M	B			M	B
1	M	181	19	6	M	183	17
	B	26	166		B	28	164
2	M	183	17	7	M	184	16
	B	24	168		B	28	164
3	M	181	19	8	M	182	18
	B	25	167		B	30	162
4	M	184	16	9	M	183	17
	B	27	165		B	29	163
5	M	183	17	10	M	182	18
	B	26	166		B	29	163

Tablo 6. kNN sınıflandırma sonuçları (kNN classification results)

Test	Doğruluk %	Hassasiyet	Özgünlük	Kesinlik
1	88,52	0,905	0,865	0,874
2	89,54	0,915	0,875	0,884
3	88,78	0,905	0,870	0,879
4	89,03	0,920	0,859	0,872
5	89,03	0,915	0,865	0,876
6	88,52	0,915	0,854	0,867
7	88,78	0,920	0,854	0,868
8	87,76	0,910	0,844	0,858
9	88,27	0,915	0,849	0,863
10	88,01	0,910	0,849	0,863
Ortalama	88,62	0,913	0,858	0,870

Tablo 7. SVM 10-kat çapraz geçişleme sonuçları (SVM 10-fold cross validation results)

Test	Gerçek	Tahmin		Test	Gerçek	Tahmin	
		M	B			M	B
1	M	188	12	6	M	190	10
	B	15	177		B	13	179
2	M	190	10	7	M	190	10
	B	14	178		B	14	178
3	M	190	10	8	M	190	10
	B	15	177		B	15	177
4	M	190	10	9	M	190	10
	B	13	179		B	14	178
5	M	189	11	10	M	190	10
	B	15	177		B	15	177

Tablo 8. SVM sınıflandırma sonuçları (SVM classification results)

TEST	DOĞRULUK %	HASSASİYET	ÖZGÜNLÜK	KESİNLİK
1	93,11	0,940	0,922	0,926
2	93,88	0,950	0,927	0,931
3	93,62	0,950	0,922	0,927
4	94,13	0,950	0,932	0,936
5	93,37	0,945	0,922	0,926
6	94,13	0,950	0,932	0,936
7	93,88	0,950	0,927	0,931
8	93,62	0,950	0,922	0,927
9	93,88	0,950	0,927	0,931
10	93,62	0,950	0,922	0,927
Ortalama	93,72	0,949	0,926	0,930

Tablo 9. C4Net 10-kat çapraz geçirme sonuçları (C4Net 10-fold cross validation results)

TEST	GERÇEK	TAHMİN		TEST	GERÇEK	TAHMİN	
		M	B			M	B
1	M	195	5	6	M	196	4
	B	8	184		B	7	185
2	M	195	5	7	M	194	6
	B	7	185		B	7	185
3	M	196	4	8	M	195	5
	B	7	185		B	7	185
4	M	195	5	9	M	196	4
	B	6	186		B	7	185
5	M	195	5	10	M	194	6
	B	8	184		B	7	185

Tablo 10. C4Net sınıflandırma sonuçları (C4Net classification results)

TEST	DOĞRULUK %	HASSASİYET	ÖZGÜNLÜK	KESİNLİK
1	96,68	0,975	0,958	0,961
2	96,94	0,975	0,964	0,965
3	97,19	0,980	0,964	0,966
4	97,19	0,975	0,969	0,970
5	96,68	0,975	0,958	0,961
6	97,19	0,980	0,964	0,966
7	96,68	0,970	0,964	0,965
8	96,94	0,975	0,964	0,965
9	97,19	0,980	0,964	0,966
10	96,68	0,970	0,964	0,965
Ortalama	96,94	0,976	0,963	0,965

Tablo 11’de AlexNet için 10-kat çapraz geçişleme, Tablo 12’de ise sınıflandırıcının doğruluk, hassasiyet, özgünlük ve kesinlik değerleri görülmektedir.

Şekil 8’de AlexNet için 10-kat çapraz geçişleme sonucunda elde edilen doğruluk grafik gösterimi, Şekil 9’da ise AlexNet için 10-kat çapraz geçişleme sonucunda elde edilen hassasiyet grafik gösterimi yer almaktadır.

4.2.3. GoogLeNet sonuçları (GoogLeNet results)

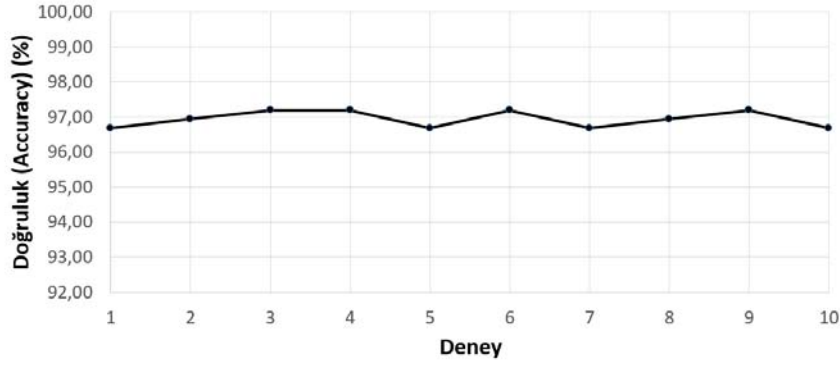
GoogLeNet, 150 devir (epoch), 10-kat çapraz geçişleme ile ortalama %94,95 sınıflandırma doğruluğu elde edilmiştir.

Tablo 13’de GoogLeNet için 10-kat çapraz geçişleme, Tablo 14’de ise sınıflandırıcının doğruluk, hassasiyet, özgünlük ve kesinlik değerleri görülmektedir.

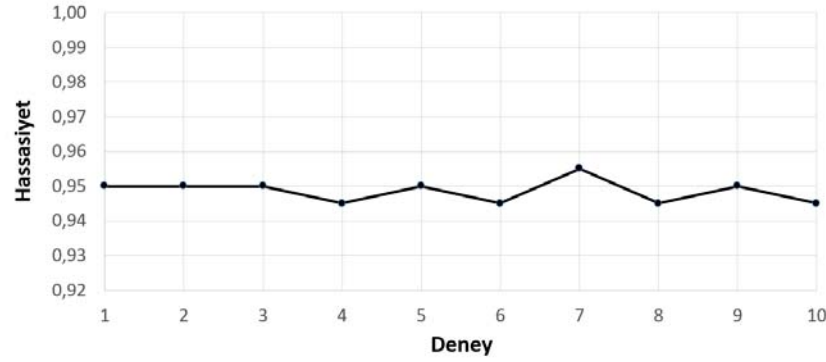
Şekil 10’da GoogLeNet için 10-kat çapraz geçişleme sonucunda elde edilen doğruluk grafik gösterimi, Şekil 11’de ise GoogLeNet için 10-kat çapraz geçişleme sonucunda elde edilen hassasiyet grafik gösterimi yer almaktadır.

4.2.4. ResNet sonuçları (ResNet results)

ResNet, 150 devir (epoch), 10-kat çapraz geçişleme ile ortalama %92,96 sınıflandırma doğruluğu elde edilmiştir.



Şekil 6. C4Net 10-kat çapraz geçişleme grafik gösterimi (C4Net 10-fold cross validation graphical representation)



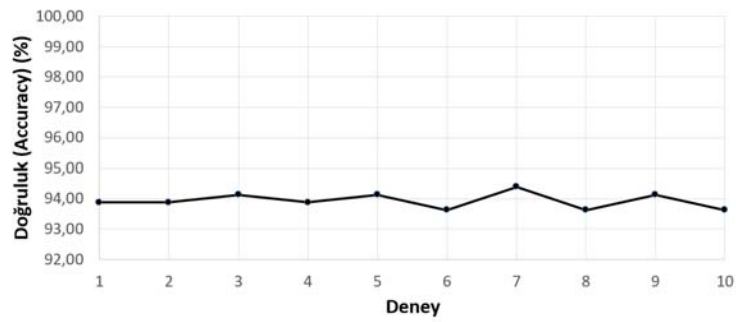
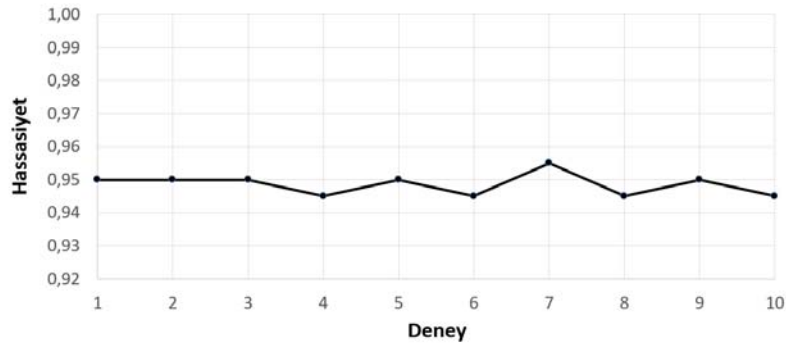
Şekil 7. C4Net 10-kat çapraz geçişleme hassasiyet grafik gösterimi (C4Net 10-fold cross validation graphical representation for sensitivity)

Tablo 11. AlexNet 10-kat çapraz geçişleme sonuçları (AlexNet 10-fold cross validation results)

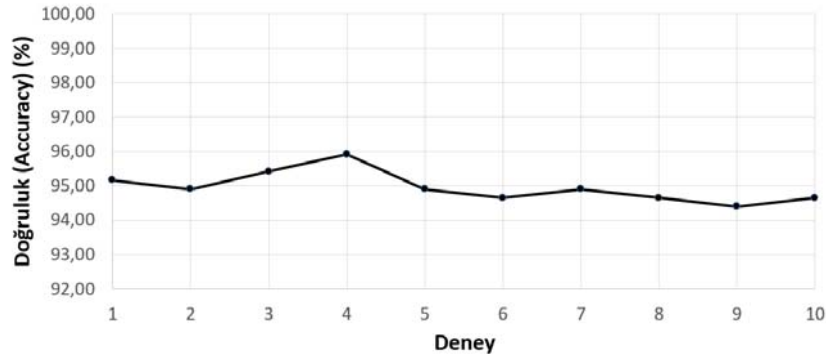
TEST	GERÇEK	TAHMİN		TEST	GERÇEK	TAHMİN	
		M	B			M	B
1	M	190	10	6	M	189	11
	B	14	178		B	14	178
2	M	190	10	7	M	191	9
	B	14	178		B	13	179
3	M	190	10	8	M	189	11
	B	13	179		B	14	178
4	M	189	11	9	M	190	10
	B	13	179		B	13	179
5	M	190	10	10	M	189	11
	B	13	179		B	14	178

Tablo 12. AlexNet sınıflandırma sonuçları (AlexNet classification results)

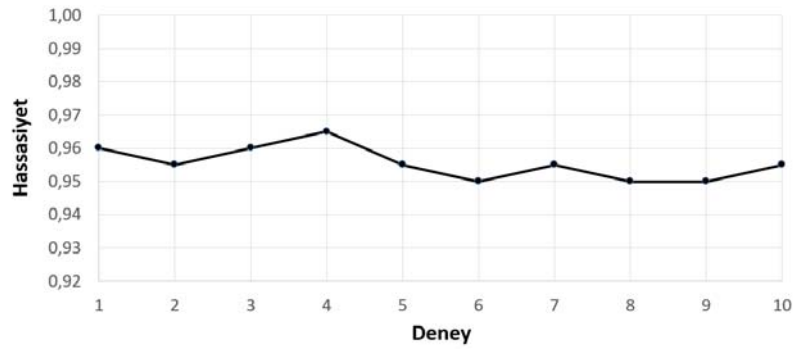
TEST	DOĞRULUK %	HASSASİYET	ÖZGÜNLÜK	KESİNLİK
1	93,88	0,950	0,927	0,931
2	93,88	0,950	0,927	0,931
3	94,13	0,950	0,932	0,936
4	93,88	0,945	0,932	0,936
5	94,13	0,950	0,932	0,936
6	93,62	0,945	0,927	0,931
7	94,39	0,955	0,932	0,936
8	93,62	0,945	0,927	0,931
9	94,13	0,950	0,932	0,936
10	93,62	0,945	0,927	0,931
Ortalama	93,93	0,949	0,930	0,934

**Şekil 8.** AlexNet 10-kat çapraz geçерleme grafik gösterimi (AlexNet 10-fold cross validation graphical representation)**Şekil 9.** AlexNet 10-kat çapraz geçерleme hassasiyet grafik gösterimi
(AlexNet 10-fold cross validation graphical representation for sensitivity)**Tablo 13.** GoogLeNet 10-kat çapraz geçерleme sonuçları (GoogLeNet 10-fold cross validation results)

TEST	GERÇEK	TAHMİN		TEST	GERÇEK	TAHMİN	
		M	B			M	B
1	M	192	8	6	M	190	10
	B	11	181		B	11	181
2	M	191	9	7	M	191	9
	B	11	181		B	11	181
3	M	192	8	8	M	190	10
	B	10	182		B	11	181
4	M	193	7	9	M	190	10
	B	9	183		B	12	180
5	M	191	9	10	M	191	9
	B	11	181		B	12	180



Şekil 10. GoogLeNet 10-kat çapraz geçerleme grafik gösterimi
(GoogLeNet 10-fold cross validation graphical representation)



Şekil 11. GoogLeNet 10-kat çapraz geçerleme hassasiyet grafik gösterimi
(GoogLeNet 10-fold cross validation graphical representation for sensitivity)

Tablo 14. GoogLeNet sınıflandırma sonuçları (GoogLeNet classification results)

TEST	DOĞRULUK %	HASSASİYET	ÖZGÜNLÜK	KESİNLİK
1	95,15	0,960	0,943	0,946
2	94,90	0,955	0,943	0,946
3	95,41	0,960	0,948	0,950
4	95,92	0,965	0,953	0,955
5	94,90	0,955	0,943	0,946
6	94,64	0,950	0,943	0,945
7	94,90	0,955	0,943	0,946
8	94,64	0,950	0,943	0,945
9	94,39	0,950	0,938	0,941
10	94,64	0,955	0,938	0,941
Ortalama	94,95	0,956	0,943	0,946

Tablo 15'te ResNet için 10-kat çapraz geçerleme, Tablo 16'da ise sınıflandırıcının doğruluk, hassasiyet, özgünlük ve kesinlik değerleri görülmektedir. Şekil 12'de ResNet için 10-kat çapraz geçerleme sonucunda elde edilen doğruluk grafik gösterimi, Şekil 13'te ise ResNet için 10-kat çapraz geçerleme sonucunda elde edilen hassasiyet grafik gösterimi yer almaktadır.

4.2.5. VGGNet sonuçları (VGGNet results)

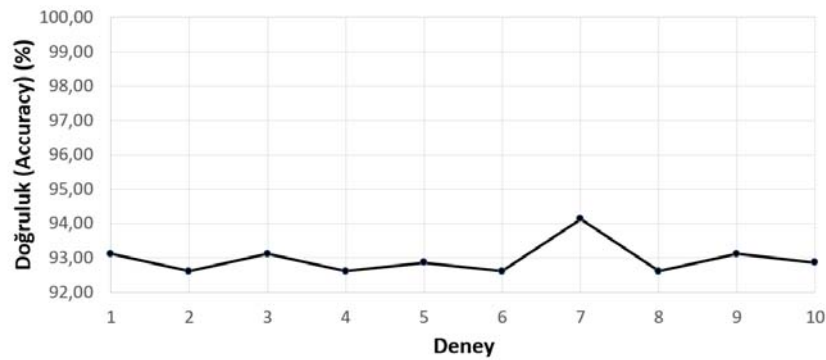
VGGNet, 150 devir (epoch), 10-kat çapraz geçerleme ile ortalama %94,85 sınıflandırma doğruluğu elde edilmiştir. Tablo 17'de VGGNet için 10-kat çapraz geçerleme, Tablo 18'de ise sınıflandırıcının doğruluk, hassasiyet, özgünlük ve kesinlik değerleri görülmektedir. Şekil 14'te VGGNet için

Tablo 15. ResNet 10-kat çapraz geçerieme sonuçları (ResNet 10-fold cross validation results)

TEST	GERÇEK	TAHMİN		TEST	GERÇEK	TAHMİN	
		M	B			M	B
1	M	188	12	6	M	186	14
	B	15	177		B	15	177
2	M	187	13	7	M	190	10
	B	16	176		B	13	179
3	M	189	11	8	M	186	14
	B	16	176		B	15	177
4	M	187	13	9	M	187	13
	B	16	176		B	14	178
5	M	189	11	10	M	188	12
	B	17	175		B	16	176

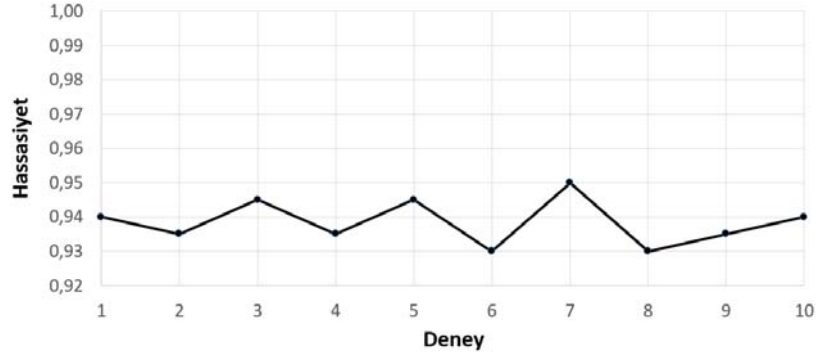
Tablo 16. ResNet sınıflandırma sonuçları (ResNet classification results)

TEST	DOĞRULUK %	HASSASİYET	ÖZGÜNLÜK	KESİNLİK
1	93,11	0,940	0,922	0,926
2	92,60	0,935	0,917	0,921
3	93,11	0,945	0,917	0,922
4	92,60	0,935	0,917	0,921
5	92,86	0,945	0,911	0,917
6	92,60	0,930	0,922	0,925
7	94,13	0,950	0,932	0,936
8	92,60	0,930	0,922	0,925
9	93,11	0,935	0,927	0,930
10	92,86	0,940	0,917	0,922
Ortalama	92,96	0,939	0,920	0,925

**Şekil 12.** ResNet 10-kat çapraz geçerieme grafik gösterimi
(ResNet 10-fold cross validation graphical representation)

10-kat çapraz geçerieme sonucunda elde edilen doğruluk grafik gösterimi, Şekil 15'te ise VGGNet için 10-kat çapraz

geçerieme sonucunda elde edilen hassasiyet grafik gösterimi yer almaktadır.



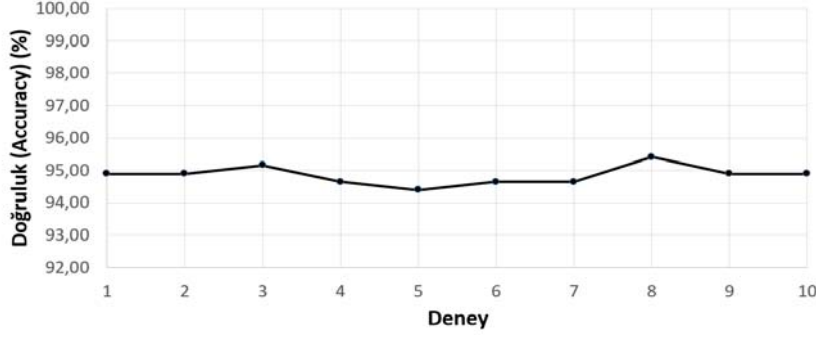
Şekil 13. ResNet 10-kat çapraz geçirme hassasiyet grafik gösterimi
(ResNet 10-fold cross validation graphical representation for sensitivity)

Tablo 17. VGGNet 10-kat çapraz geçirme sonuçları (VGGNet 10-fold cross validation results)

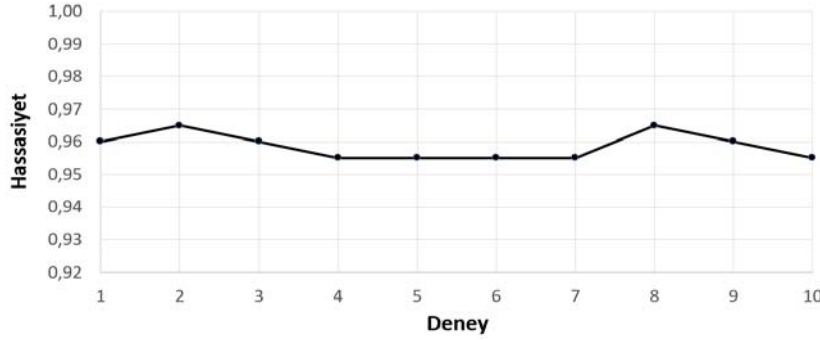
TEST	GERÇEK	TAHMİN		TEST	GERÇEK	TAHMİN	
		M	B			M	B
1	M	192	8	6	M	191	9
	B	12	180		B	12	180
2	M	193	7	7	M	191	9
	B	13	179		B	12	180
3	M	192	8	8	M	193	7
	B	11	181		B	11	181
4	M	191	9	9	M	192	8
	B	12	180		B	12	180
5	M	191	9	10	M	191	9
	B	13	179		B	11	181

Tablo 18. VGGNet sınıflandırma sonuçları (VGGNet classification results)

TEST	DOĞRULUK %	HASSASİYET	ÖZGÜNLÜK	KESİNLİK
1	94,90	0,960	0,938	0,941
2	94,90	0,965	0,932	0,937
3	95,15	0,960	0,943	0,946
4	94,64	0,955	0,938	0,941
5	94,39	0,955	0,932	0,936
6	94,64	0,955	0,938	0,941
7	94,64	0,955	0,938	0,941
8	95,41	0,965	0,943	0,946
9	94,90	0,960	0,938	0,941
10	94,90	0,955	0,943	0,946
Ortalama	94,85	0,959	0,938	0,942



Şekil 14. VGGNet 10-kat çapraz geçeleme grafik gösterimi (VGGNet 10-fold cross validation graphical representation)



Şekil 15. VGGNet 10-kat çapraz geçeleme grafik gösterimi (VGGNet 10-fold cross validation graphical representation)

5. SONUÇLAR VE TARTIŞMALAR (RESULTS AND DISCUSSIONS)

Deneysel çalışmalar ML ve CNN modelleri için aynı eğitim - test veri kümesi ile 10-kat çapraz geçeleme yöntemiyle gerçekleştirilmiştir. 10-kat çapraz geçeleme için 3920 dermoskopi görüntü rastsal 10 parçaya bölünmüş ve her bir deney için 200 malign (melanom) ve 192 benign (iyi huylu) olmak üzere 392 adet görüntü kullanılmıştır. ML yöntemleri için 3920 dermoskopi görüntüsünde öznel çıkarmı benzer şekilde gerçekleştirilmiş ve en yüksek başarıyı %93,72 doğruluk oranı ile SVM göstermiştir. Görüntü sınıflandırmada, model doğruluğunun yanı sıra hassasiyet yani melanom tespit etme gücünün de önemli olduğu bilinmektedir. Yapılan deneysel çalışmalarda geleneksel makine öğrenmesi yöntemleri içinde en yüksek hassasiyete 0,949 ile yine SVM'nin sahip olduğu görülmüştür. Yapılan çalışmada, otomatik melanom tespiti için önerilen C4Net CNN modeli yanı sıra literatürde ön plana çıkmış AlexNet, GoogLeNet, ResNet ve VGGNet de kullanılmıştır.

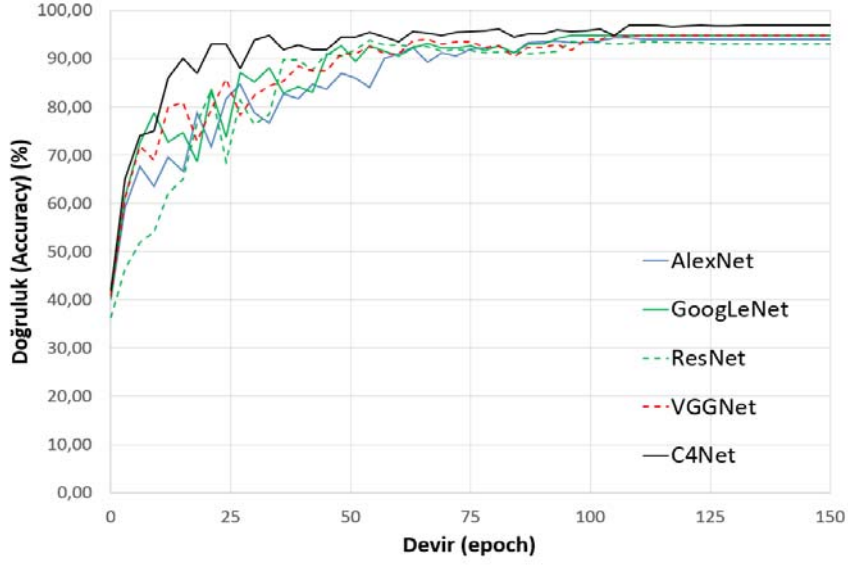
Elde edilen sonuçlar ışığında önerilen C4Net CNN modeli melanom tespitini %96,94 doğrulukta yaparak diğer tüm CNN modellerine göre daha yüksek bir başarı göstermiştir. Hassasiyet değeri kıyaslamasında yine C4Net, 0,976 ile melanom tespitini güçlü bir şekilde yapabileceğini göstermiştir. C4Net, 10-kat çapraz geçeleme deneylerinde Şekil 6'da da görüldüğü gibi en düşük %96,68 ve en yüksek %97,19 sınıflandırma başarıları göstermiştir. Şekil 7'de görüldüğü gibi C4Net için hassasiyet en düşük 0,970 ve en

yüksek 0,980 elde edilmiştir. AlexNet, %93,93 başarı elde etmiştir. 10-kat çapraz geçeleme deneylerinde Şekil 8'de de görüldüğü gibi en düşük %93,62 ve en yüksek %94,39 sınıflandırma başarıları göstermiştir. Şekil 9'da görüldüğü gibi AlexNet için hassasiyet en düşük 0,945 ve en yüksek 0,955 elde edilmiştir. GoogLeNet, %94,95 başarı elde etmiştir. 10-kat çapraz geçeleme deneylerinde Şekil 10'da görüldüğü gibi en düşük %94,39 ve en yüksek %95,92 sınıflandırma başarıları göstermiştir. Şekil 11'de görüldüğü gibi GoogLeNet için hassasiyet en düşük 0,950 ve en yüksek 0,965 elde edilmiştir.

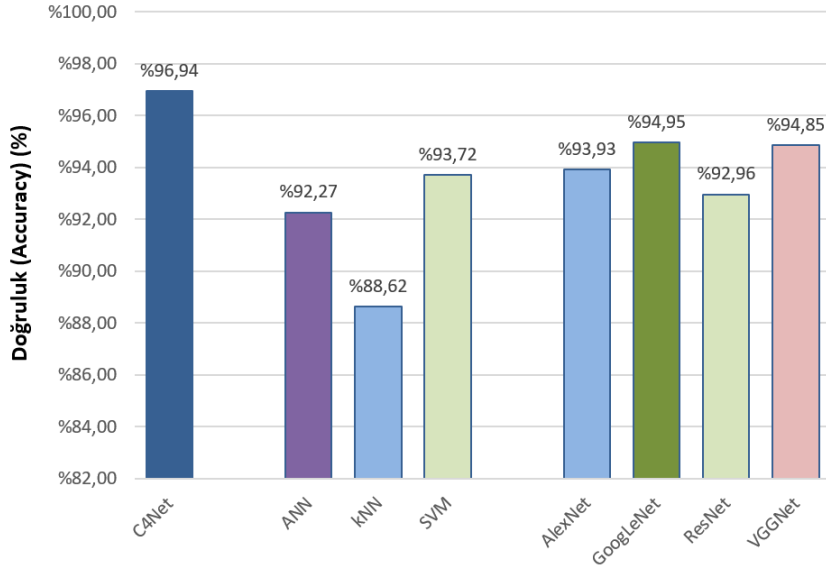
ResNet, %92,96 başarı elde etmiştir. 10-kat çapraz geçeleme deneylerinde Şekil 12'de görüldüğü gibi en düşük %92,60 ve en yüksek %94,13 sınıflandırma başarıları göstermiştir. Şekil 13'te görüldüğü gibi ResNet için hassasiyet en düşük 0,930 ve en yüksek 0,950 elde edilmiştir.

VGGNet, %94,85 başarı elde etmiştir. 10-kat çapraz geçeleme deneylerinde Şekil 14'te görüldüğü gibi en düşük %94,39 ve en yüksek %95,41 sınıflandırma başarıları göstermiştir. Şekil 15'te görüldüğü gibi VGGNet için hassasiyet en düşük 0,955 ve en yüksek 0,965 elde edilmiştir.

Şekil 16'de sınıflandırma başarıları açısından önerilen C4Net ile diğer CNN temsilcilerinin (AlexNet, GoogLeNet, ResNet, VGGNet) karşılaştırması gösterilmiştir. Şekil 17'de, C4Net, ML yöntemleri ve diğer CNN modelleri ile elde edilen sınıflandırma doğruluk oranlarının grafik gösterimi yer almaktadır.



Şekil 16. Önerilen C4Net ile diğer temsilcilerin karşılaştırılması
(Comparison of proposed C4Net with other representatives)



Şekil 17. Önerilen C4Net ile diğer temsilcilerin karşılaştırılması grafik gösterimi
(Graphical comparison of proposed C4Net with other representatives)

6. SONUÇLAR (CONCLUSIONS)

Melanom nadir görülmesine rağmen ölüme en çok neden olan cilt kanseri türüdür. Erken tanı hastalık tedavisinde büyük önem arz etmektedir. Melanom tanısı uzman hekimler tarafından gözle muayene sonucu gerçekleştirilebilir. Gerçekleştirilen bu çalışmada cilt kanseri türü olan melanom tanısı için geleneksel makine öğrenmesi yöntemleri ile derin öğrenme modellerini kıyaslayan kapsamlı bir değerlendirme yapılmış ve yüksek doğrulukta melanom tanısı koyabilecek bir CNN modeli tasarlanmıştır. Deneysel çalışmalarda ISIC (The International Skin Imaging Collaboration) tarafından

sağlanan 3920 gerçek doku görüntüleri kullanılmıştır. Önerilen CNN modeli, geleneksel makine öğrenmesi yöntemleri ve literatürde sıklıkla kullanılan AlexNet, GoogLeNet, ResNet, VGGNet ile karşılaştırılmıştır.

DL modellerinin en büyük avantajı sınıflandırma için kullanılacak özellikleri kendisinin belirlemesidir. Bilindiği gibi sınıflandırma yöntemlerinin performansı, özelliklerin nasıl belirlendiği ile yakından ilişkilidir. Özellikler pek çok kez uzman görüşü veya elle belirlenir ayrıca ML yöntemleri için kodlanması gerekir. Bu durum ML yöntemlerinin başarısını olumsuz etkiler. Diğer taraftan veri ölçeği arttıkça

DL modellerinin sınıflandırma performansı artmaktadır. Ancak bu durumda DL modelleri eğitim için yüksek işlem gücüne sahip donanıma ve zamana ihtiyaç duyar. ML yöntemleri bu açıdan avantajlıdır. Tasarlanan CNN modeli (C4Net) çeşitli denemeler sonucunda geliştirilmiş ve %96,94 doğruluk ile melanom tanısı koyarak, diğer yöntemlere göre daha yüksek başarı göstermiştir. Aynı zamanda yapılan deneysel çalışmalarda görülmüştür ki sınıflandırıcılara ait hassasiyet yani melanom tanısını doğru koyma gücü en yüksek 0,976 ile yine önerilen C4Net CNN modeli ile elde edilmiştir. Elde edilen bu sonuç veri kümesinde yer alan doku görüntülerinin doğasından kaynaklanabilir. Yüksek çözünürlüklü örneklem sayısının artırılması ile de bu başarı artırılabilir. Ancak büyük ölçekli ve dengeli bir dermoskopi görüntü verisi oluşturmak, bunların etiketlenmesi işlemleri ile CNN eğitimi için gereksinim duyulan yüksek performanslı donanım bu tür çalışmalar için önemli kısıtlardır. Bu kısıtlar giderildiğinde melanom vb. hastalıkların otomatik tespiti için daha başarılı modeller tasarlanabilir. Akıllı telefon, tablet vb. cihazlar için uygulamalar geliştirilebilir.

Gelecek çalışmalarda CNN eğitim sürecini hızlandıracak etkili ve ölçeklenebilir paralel algoritmalar geliştirilmesi amaçlanmaktadır. Ayrıca CNN model başarımını etkileyen parametreler farklı optimizasyon yöntemleri ile otomatik ve etkin bir şekilde belirlenebilir.

KAYNAKLAR (REFERENCES)

1. Halk Sağlığı Genel Müdürlüğü - <https://hsgm.saglik.gov.tr>. Erişim tarihi Nisan 10, 2018
2. Ye, Q. H., Qin, L. X., Forgues, M., He, P., Kim, J. W., Peng, A. C., & Ma, Z. C., Predicting hepatitis B virus-positive metastatic hepatocellular carcinomas using gene expression profiling and supervised machine learning. *Nature medicine*, 9 (4), 416, 2003.
3. Kononenko, I., Machine learning for medical diagnosis: history, state of the art and perspective. *Artificial Intelligence in medicine*, 23 (1), 89-109, 2001.
4. Berikol, G. B., Yıldız, O., & Özcan, İ. T., Diagnosis of acute coronary syndrome with a support vector machine. *Journal of medical systems*, 40 (4), 84, 2016.
5. Litjens, G., Sánchez, C. I., Timofeeva, N., Hermsen, M., Nagtegaal, I., Kovacs, I., & Van Der Laak, J., Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Scientific reports*, 6, 26286, 2016
6. Liu, S., Cai, W., Pujol, S., Kikinis, R., & Feng, D., Early diagnosis of Alzheimer's disease with deep learning. In *Biomedical Imaging (ISBI), 2014 IEEE 11th International Symposium on* (pp. 1015-1018), IEEE, 2014.
7. Wang, G., Li, W., Zuluaga, M. A., Pratt, R., Patel, P. A., Aertsen, M., Vercauteren, T., Interactive medical image segmentation using deep learning with image-specific fine-tuning. *IEEE Transactions on Medical Imaging*, 37 (7), 1562-1573, 2018
8. Acharya, U. R., Oh, S. L., Hagiwara, Y., Tan, J. H., & Adeli, H., Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. *Computers in biology and medicine*, 100, 270-278, 2018
9. Shin, H. C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., Summers, R. M., Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning, *IEEE transactions on medical imaging*, 35 (5), 1285-1298, 2016.
10. Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Chen, T., Recent advances in convolutional neural networks. *Pattern Recognition*, 2017.
11. Chen, X., Xiang, S., Liu, C. L., & Pan, C. H., Vehicle detection in satellite images by hybrid deep convolutional neural networks, *IEEE Geoscience and remote sensing letters*, 11 (10), 1797-1801, 2014.
12. Pal, A., Garain, U., Chandra, A., Chatterjee, R., & Senapati, S., Psoriasis skin biopsy image segmentation using Deep Convolutional Neural Network, *Computer methods and programs in biomedicine*, 159, 59-69, 2018.
13. Chan, T. H., Jia, K., Gao, S., Lu, J., Zeng, Z., & Ma, Y., PCANet: A simple deep learning baseline for image classification?, *IEEE Transactions on Image Processing*, 24 (12), 5017-5032, 2015
14. Chen, Y., Lin, Z., Zhao, X., Wang, G., & Gu, Y., Deep learning-based classification of hyperspectral data, *IEEE Journal of Selected topics in applied earth observations and remote sensing*, 7 (6), 2094-2107, 2014.
15. Krizhevsky, A., Sutskever, I., & Hinton, G. E., Imagenet classification with deep convolutional neural networks, In *Advances in neural information processing systems* (pp. 1097-1105), 2012.
16. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Rabinovich, A., Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-9), 2015.
17. Simonyan, K., & Zisserman, A., Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
18. He, K., Zhang, X., Ren, S., & Sun, J., Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778), 2016.
19. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., Sánchez, C. I., A survey on deep learning in medical image analysis, *Medical image analysis*, 42, 60-88, 2017
20. Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S., Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542 (7639), 115, 2017.
21. Codella, N., Cai, J., Abedini, M., Garnavi, R., Halpern, A., & Smith, J. R., Deep learning, sparse coding, and SVM for melanom recognition in dermoscopy images,

- In International Workshop on Machine Learning in Medical Imaging (pp. 118-126), Springer, Cham, 2015.
22. The International Society for Digital Imaging of the Skin (ISDIS) Archive - <https://isic-archive.com/> - Erişim tarihi Aralık 5, 2017.
 23. Li, Y., Nie, X., & Huang, R., Web spam classification method based on deep belief networks, *Expert Systems with Applications*, 96, 261-270, 2018.
 24. Bengio, Y., Learning deep architectures for AI. *Foundations and trends in Machine Learning*, 2 (1), 1-127, 2009.
 25. LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P., Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86 (11), 2278-2324, 1998.
 26. Haralick, R., Shanmugam, K., & Dinstein, I., Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions on*, SMC-3 (6), 610-621. doi: 10.1109/TSMC.1973.4309314, 1973.
 27. The Skin Cancer Foundation - SkinCancer.org - <https://www.skincancer.org/skin-cancer-information/melanoma#panel1-5> - Erişim tarihi Nisan 3, 2018.