

GAZİ

JOURNAL OF ENGINEERING SCIENCES

Predicting of Melanoma Skin Cancer Using Machine Learning Methods

Resul Bütüner^a, M. Hanefi Calp^b

Submitted: 23.06.2024 Revised: 12.02.2024 Accepted: 04.04.2024 doi:10.30855/gmbd.0705N12

ABSTRACT

Cancer disease is known as the second highest cause of death in the world. One of the most common types of this disease is skin cancer. As in almost all cancer diseases, early diagnosis and treatment of skin cancer is of great importance. In the process of diagnosing cancer, Machine Learning-based methods are also widely used in addition to traditional methods. The most important advantage of these methods is that they eliminate or minimize human errors that may arise during the cancer diagnosis process. In this study, skin cancer was diagnosed with K-Nearest Neighbor (KNN), Naive Bayes (NB), Random Forest (RO), Logistic Regression (LR), and Artificial Neural Networks (ANN) methods using images taken from patients. In these algorithms, the training and testing process was run, the results were analyzed and models were created. With these models, benign and malignant lesions were compared and skin cancerous lesions were detected and success percentages were revealed. As a result, the best results were obtained using the ANN method, with a training percentage of 99.1% and a testing percentage of 98.6%. When different inputs were given to the created and proposed ANN model, it was observed that the model predicted skin cancer with a high accuracy rate. The results obtained revealed the success of the study and that machine learning methods are a usable method in the cancer diagnosis process.

Keywords: Skin Cancer, Machine Learning, Diagnosis, Prediction

^a2Ankara Beypazarı Fatih Vocational and Technical Anatolian High School, Dept. of Computer, Orcid: 0000-0002-9778-2349 e mail: rbutuner@gmail.com

^bAnkara Hacı Bayram Veli University, Faculty of Economics and Administrative Sciences, Dept. of Management Information Systems 06560 - Ankara, Türkiye Orcid: 0000-0001-7991-438X e mail: hanefi.calp@hbv.edu.tr

^cCorresponding author: hanefi.calp@hbv.edu.tr

Makine Öğrenmesi Yöntemleri Kullanılarak Melanom Cilt Kanserinin Tahmin Edilmesi

ÖZ

Kanser hastalığı, dünyadaki en yüksek ikinci ölüm nedeni olarak bilinmektedir. Bu hastalığın yaygın türlerinden biri de cilt kanseridir. Hemen hemen tüm kanser hastalıklarında olduğu gibi cilt kanserinde de erken teşhis ve tedavi sürecinde oldukça büyük bir öneme sahiptir. Kanser hastalığının teşhisi sürecinde, geleneksel yöntemlerin yanı sıra Makine Öğrenmesi tabanlı yöntemler de yaygın olarak kullanılmaktadır. Söz konusu yöntemlerin en önemli avantajı, kanser teşhisi sürecinde insan kaynaklı doğabilecek hataları ortadan kaldırması veya minimize etmesidir. Bu çalışmada, hastalardan alınan görüntüler kullanılarak K-En Yakın Komşu (KNN), Naive Bayes (NB), Rastgele Orman (RO), Lojistik Regresyon (LR) ve Yapay Sinir Ağları (YSA) yöntemleri ile cilt kanseri hastalığı teşhis edilmiştir. Bu algoritmalarda eğitim ve test süreci işletilerek sonuçlar analiz edilmiş ve modeller oluşturulmuştur. Bu modeller ile iyi ve kötü huylu lezyonlar karşılaştırılmış ve cilt kanseri olan lezyonlar tespit edilerek başarı yüzdeleri ortaya konulmuştur. Sonuç olarak, en iyi sonuçlar eğitim yüzdesi %99.1, test yüzdesi ise %98.6'lık bir oran ile YSA yöntemi kullanılarak elde edilmiştir. Oluşturulan ve önerilen YSA modeline farklı girdiler verildiğinde, modelin cilt kanseri hastalığını yüksek bir doğruluk oranıyla tahmin ettiği görülmüştür. Elde edilen sonuçlar, yapılan çalışmanın başarısını ve makine öğrenmesi yöntemlerinin kanser teşhisi sürecinde kullanılabilir bir yöntem olduğunu ortaya koymuştur.

Anahtar Kelimeler: Cilt Kanseri, Makine Öğrenmesi, Teşhis, Tahmin

1. Giriş (Introduction)

Kanser, farklı organlarda bulunan hücrelerin kontrol dışında çoğalmasından dolayı meydana gelen bir hastalık olarak tanımlanır [1]. Kanser hastalığı, hastalıklar içerisinde dünya üzerinde en yüksek karşılaşılan ikinci ölüm sebebidir [2]. Yaygın görülen kanser türlerinden birisi de cilt kanseridir. Bunun içerisinde en ölümcül olanı Melanoma Cilt Kanseridir. Cilt kanseriyle ilgili ölümlerin yaklaşık %75'i ise melanomadan dolayıdır [3]. Melanoma türüne ait cilt kanseri metastaz yapmak suretiyle diğer dokulara yayılmaktadır. Dolayısıyla bu durum ölüm oranlarını artırmaktadır [4]. Melanin renk pigmentleri deriye renk vererek, melanosit adlı cilt hücreleri tarafından üretilir. Bu hücrelerin deride kontrolsüz olarak bölünüp çoğalması ile meydana gelen cilt kanserine melanom ya da melanoma olarak tanımlanmaktadır [5]. Cilt kanseri, hem Türkiye'de hem de dünyada diğer kanser türleri gibi artış göstererek büyük dikkat çeker. En çok melanom dışı deri kanseri beyaz ırkta görülen kanserlerdir. Burada hastanın melanomunda benign (iyi huylu) malignant (kötü huylu) olarak ifade edilmektedir. Yaşam boyunca oluşma oranı erkek bireylerde 33 kişide birinde iken kadın bireylerde ise 52 kişide birinde rastlandığı görülür [6]. Melanomun en büyük nedeni ise ultraviyoleten (UV) kaynaklanan ışıklardır. Herkeste cilt kanserine yakalanma riski bulunur. Fakat cildinde açık ten olan, cildi fazlaca ben ve lekeleri bulunan, uzunca süre güneşin altında bulunan, çocukluk sürecinde güneş yanığı olan bireylerin cilt kanseri olma riski daha fazladır [7]. Tedavi sürecinde, hastanın geçmiş hayatından itibaren, kırmızımsı renk, kabarmış, tomografik açıdan anormal, büyüme gösteren, kanamış, kabuklanmış veya farklılaşan lezyonlar tanımlanarak daha sonrasında görsel olarak incelenir. Klinik uzman kontrolünde, şüphe görülen lezyonlar biyopsiye göre tedavi sürecine karar verilmektedir. Şu anda ise klinikteki tanının doğruluğunu artırmak için yeni teknolojiler de kullanılmaktadır. Neoplazm doku kitlesi 200 yıllık bir geçmişe dayanarak bir hastalık varlığı olarak kabul edilmiştir. Bundan dolayı melanomun tanısı ve tedavisi büyük ölçüde gelişim sağlamıştır. Fakat hastalarda melanom cilt kanserinin erken teşhisinin önemi çokça anlaşılmamaktadır [8]. Melanom cilt kanserinde önemli bir problem de lezyon bölütlemesinin yanlış yapılması ve lezyondaki kıl temizliğidir ve, bu durum yanlış kanser tanısına sebebiyet vermektedir [9]. Herhangi bir vücutun deri yüzeyinde melanositlerin fazlaca bulunduğu zaman melanoma sebep olur. Dermoskopi muayene tekniği, derideki yüzeylerin mikroskobik olarak incelenerek deride bulunan koyu renkli lekeler tanı koyulması için kullanılmaktadır. Melanom kişilerin ölümüne yol açan bir cilt kanseri olarak bilinir. Melanom cilt kanserine yakalanmış hastaların sadece %4'ünü oluşturmaktadır. Fakat tüm cilt kanserindeki ölümlerin %75'ini oluşturur. Uzman dermatologlar melanom tanısı için dermoskopi kullansalar bile, doğruluk başarıları %75 ile %84 arasında tahmini olarak bulunur [10]. Melanom cilt kanseri erken teşhis edilip, tedavi sürecine geçilirse iyileşmenin olumlu yönde sonuçlanmasını mümkün kılar. Ama melanomun teşhisi geç olursa bu sefer derinin daha diplerine doğru büyüme göstererek vücudun diğer bölgelerine yayılma imkânı doğar. İşte cildin yüzeyindeki diğer bölgelerine yayıldığı zaman tedavi süreci daha tehlikeli olabilir. Dünya da en yaygın şekilde görülen kanser türlerinden biri deri kanseridir ve giderek de artmaktadır [11]. Kişilerin cilt kanserine yakalanma riskinin ultraviyolenin ışımına maruz kalma, genetik ve çevresel faktörlerin bir oluşumundan kaynaklanır [12].

Deri anatomisinde iki ana katman bulunmaktadır. Bunlar epidermis yani dış katman ve dermis yani iç katman'dır. Epidermis, düz, ölçek benzeri hücreler ve bazal hücreler diye adı konulan yuvarlak Squamous hücreleri diye ifade edilen hücrelerden oluşmaktadır. Epidermisin de alt kısmında ise melanositi içermektedir. Melanositler, epidermisinin tam alt tarafında bulunmaktadır. Cilt kısmına doğal renk vererek pigment olan melanin üreten pigment hücreleridir. Cildin güneşle fazlaca kaldığında, melanositler daha fazla pigment üretir. Böylece cilt yüzeyinin bronzlaşması ya da koyulaşması gerçekleşerek melanomlara sebep olmaktadır [13]. Melanom cilt kanseri ise melanosit malign diye ifade edilen tümör olarak bilinir. Bu tarz hücrelerin çoğunlukla deri kısmında bulunmaktadır. Fakat bağırsak ve gözde de bulunma durumu görülür. Melanom diğerlerine göre görülme sıklığı daha azdır. Fakat cilt kanserinden dolayı ölüm vakalarının tümüne sebep olmaktadır. Uzun yıllardır devam eden laboratuvar ve klinik araştırmalarında, bu tümörlere uygulanacak erken cerrahi rezeksiyonu şuan da bile en büyük tedavi şansını sunar. Melanom en tehlikeli cilt kanseri türüdür ve melanosit diye tanımlanan cilt hücrelerinde başlamaktadır [14]. Avrupa'da yapılan araştırmaya göre malign melanom insidansı, 100 bin kişi başına 4 ile 14 hasta arasında sonuç vermektedir [15].

Günümüzde cilt kanseri teşhisinde geleneksel yöntemlerle (dermaskopi, gözle muayene ve cerrahi tabanlı) birlikte görüntü işleme, derin öğrenme ve makine öğrenmesi algoritmaları da kullanılmaktadır. Görüntülerin temizlenme sürecinde görüntü işleme ve yapay sinir ağları gibi tekniklerinden daha yaygın yararlanılmaktadır [16,17]. Derin Öğrenme ve Görüntü İşleme algoritmaları geleneksel yöntemlere göre, insan faktörünü ortadan kaldırdığı için daha güvenilir sonuçlar vermektedir. Bu kullanılan algoritmalarla, bir uzman kadar başarılı sonuçlar alınabilmekte ve insan hatası ortadan kaldırılmaktadır [18]. Kıl temizliği üzerine çözüm getiren ilk çalışmalardan biri kıl gürültüsünün temizliği için geliştirilen DullRazor algoritmasıdır. Bu algoritma sayesinde thresholding (eşikleme) ve morfolojik işlemler ile lezyonda kıl temizliği yapar ve hızlı, basit olmasından dolayı bu alanda yapılacak bilimsel çalışmalara katkı sağlamaktadır [19]. Melanom tespitinde, lezyonun kenarlarını tespit etmek ve doğru bölütleme amacıyla denetimsiz hızlı ve başarılı bir algoritma geliştirilmiştir [20]. Ayrıca, YUV renk uzayının kullanılmasıyla kıl temizliği ve bölütleme için yeni bir yaklaşım geliştirmiştir. Bu yaklaşım eşikleme ve morfolojik işlemler ile YUV renk uzayı üzerine temellenmiştir. Çalışmada 235 görüntü kullanılmış ve %98'lik başarı elde edilmiştir [3]. Başka bir çalışmada ise, kıl temizleme ve bölütleme çalışmalarında sadece görüntü işleme algoritmalarının yeterli olmadığı görülmüştür. Vnet tabanlı derin öğrenme ile bölütleme yapılmış ve %89 başarı değeri elde edilmiştir [21].

Literatürde, otomatik kodlayıcı, konuşma ve CNN dayalı başka bir yeni model önerilen çalışmaya da rastlanmıştır. Kullanılan veri kümesi, 1800 iyi huylu ve 1497 kötü huylu tümör görüntüsünü içeren Uluslararası Cilt Görüntüleme İş Birliği (ISIC) cilt kanseri adı verilen halka açık erişimli bir veri kümesidir. Veri kümesi, önerilen yöntemde otomatik kodlayıcı modeli kullanılarak yeniden oluşturulmuştur. Orijinal veri setini ve yapılandırılmış veri setini eğitmek ve sınıflandırmak için artık bloklardan ve artış ağlarından oluşan Mobile Net V2 modeli kullanılmıştır. Çalışmanın sınıflandırma başarı oranı %95,27 olarak gerçekleşmiştir [22]. Melanomu doğru bir şekilde tespit edip sınıflandırabilen bir sinir ağı oluşturmayı amaçlayan başka bir çalışmada, prosedürleri, görüntü geliştirmenin yanı sıra Maksimum Gradyan Yoğunluğu tekniği kullanılarak kılları çıkarmak için dermoskopik görüntülerin önceden işlenmesiyle başlamıştır. Cilt lezyonlarını görüntülerden ayırmak için Otsu Eşik algoritmasına dayalı bir segmentasyon tekniği kullanılmıştır. Bölümlere ayrılmış görüntüler daha sonra bir sinir ağını eğitmek için kullanılan ABCD, GLCM ve LBP gibi birçok özelliği hesaplamak için kullanılmıştır. ISIC arşivinin birleşik veri kümesi ve PH2 dermoskopik görüntü veritabanında ağ %97,7'lik bir doğruluk elde edilmiştir [23]. Yine, üç sınıf cilt kanseri görüntülerini tanımak amacıyla önceden eğitilmiş GoogLeNet ve AlexNet modelleri kullanılmıştır. Kullanılan veri seti, bölütlemesiz veri seti ve bölütlemeli veri seti olmak üzere iki formata ayrılmıştır. Sınıflandırma başarı oranı bölütlemeli veri seti için %92.2 ve bölütlemesiz veri seti için % 89.8 olmuştur [24]. Başka bir çalışmada ise yazarlar cilt kanseri görüntülerini Melanoma ve Nevus olarak iki sınıfa ayırmıştır. Sınıflandırmak amacıyla ResNet-50 modeli seçilmiştir. ResNet-50 modeli kullanılırken, sabit bir öğrenme oranının aksine modelin her katmanı için farklı bir öğrenme oranı kullanılmıştır. Ayrıca, öğrenme hızlarını düşürmek için kosinüs işlevine dayalı yeni yöntemler denenmiş ve bu yöntem ile % 82.3'lük bir başarı oranı elde edilmiştir [25]. Son olarak, farklı bir çalışmada veri setindeki her bir görüntüyü veri artırma yöntemini kullanarak veri seti artırılmış ve transfer öğrenmesi yaklaşımı ile sınıflandırma işlemi gerçekleştirilmiştir. Bu yöntem ile %95.91'lik başarı oranına ulaşılmıştır [26].

Çalışmada, günümüzde sıkça görülen cilt kanseri teşhisinde hızlı ve doğru tahminde bulunan bir makine öğrenmesi modelinin tasarlanması ve uygulanması amaçlanmıştır. Ulaşılmak istenen hedeflerden birisi, biyoloji alanında makine öğrenmesi uygulamalarını hayata geçirmektir. Böylece hastalık teşhisinde zaman ve emek kaybını minimuma indirmek ve hastalığı en yüksek doğruluk oranıyla teşhis edebilmek amaçlanmıştır. Çalışmanın bir başka hedefi ise, oluşturulan modelin farklı ve çok sayıda verileri kullanarak kendi kendini eğitmesi ve teşhis sürecindeki doğruluk oranını veya performansını giderek arttırmasıdır. Bu çerçevede, ikinci bölümde çalışma kapsamında kullanılan Materyal ve Yöntem, üçüncü bölümde elde edilen Bulgular ve Tartışma, dördüncü bölümde ise Sonuç ve Öneriler yer almaktadır.

2. Materyal ve Yöntem (Material and Method)

Bu bölümde, çalışma sürecinde makine öğrenmesi yöntemleri kullanılarak modeller oluşturulmuş, en yüksek performansa sahip model önerilmiş ve gerçekleştirilen tüm adımlar ayrıntılarıyla birlikte sunulmuştur. Öncelikle çalışmada kullanılan algoritmalar kısaca açıklanmış, daha sonra veriler hakkında bilgi verilmiş (veri setinin temini ve organize edilmesi), son olarak bu veriler kullanılarak oluşturulan modelin ayrıntılarına yer verilmiştir. Girdi değeri olarak sisteme cilt lezyonu görüntüsü verilmiş ve daha sonra görüntü işleme yöntemleri kullanılarak cilt kanseri varlığının olup olmadığı hakkında bir sonuç çıkarmak amacıyla analizler yapılmıştır.

2.1. Veri seti ve ön işleme (Dataset and preprocessing)

Çalışmanın verileri herkese açık bir platform olan <https://www.kaggle.com/datasets/fanconic/skin-cancer-malignant-vs-benign> adresinden elde edilmiştir. Bu aşamada; veriler analiz edilmiş ve en yüksek performans üretecek modeller için ön işlemeden geçirilmiştir. Ayrıca, veri hazırlama, keşifsel veri analizi ve modelleme gibi süreçler de icra edilmiştir. Çalışmada kullanılan veriler iyi ve kötü huylu cilt lezyonlarına aittir. Bu veri setinde toplam 2637 tane cilt lezyonu görüntüleri vardır. Bu görüntülerin her biri 224×224 yükseklik ve genişlikle JPG formatında kaydedilmiştir. Verilerin ayrıntılarına Tablo 1'de yer verilmiştir.

Tablo 1. Çalışmada kullanılan veri sayıları ve yüzdeleri (Data numbers and percentages used in the study)

Kullanım Alanı	Veri Sayısı	Yüzde (%)	İyi Huylu	Kötü Huylu
Eğitim (training)	1977	75	1482	495
Test (testing)	660	25	360	300
Toplam	2637	100	1842	795

2.2. Algoritma seçimi (Algorithm selection)

Melenom cilt kanserini tespit etmek amacıyla gerçekleştirilen çalışmada, KNN, RO, NB, YSA ve LR algoritmalarından faydalanılmış ve modeller oluşturulmuştur. Literatürdeki çalışmalar incelendiğinde, görüntü sınıflandırma ve tahmine yönelik yaygın olarak kullanılan algoritmalar çalışmada yer almıştır. Sözkonusu algoritmaların kısaca açıklamalarına bu bölümde yer verilmiştir.

2.2.1. KNN

En basit makine öğrenmesi algoritması olarak kabul edilir ve yaygın olarak kullanılan bir sınıflandırma algoritmasıdır. Bu yöntemde k sayısı önemlidir ve öncelikle bu sayı belirlenir. k sayısı, bir nesnenin hangi sınıfta olacağına karar vermek için o nesneye en yakın olan kaç tane komşu kullanılacağıyla ilgili bir sayıdır. k sayısı içinde en fazla hangi sınıfa yakınlık varsa, bilinmeyen nesne o sınıfa dâhil edilir. Ayrıca KNN eğitim aşamasına sahip değildir [27-29].

2.2.2. RO

Leo Breiman ve Adele Cutler tarafından oylama metodu ile geliştirilen bir sınıflandırma yöntemidir. Birden çok karar ağacının birleşmesiyle meydana gelmektedir. Bireysel ağaçlar tarafından oylanmasıyla kazanan sınıf belirlenir ve karar ağaçlarının oluşumunda, birbirinden bağımsız şekilde ve veri setinden bootstrap tekniği ile çekilen örneklerden faydalanılmaktadır. Torbalama yönteminin gelişmiş bir şekli olarak kabul edilmektedir [30].

2.2.3. NB

Birçok alanda yüksek performans gösteren istatistik tabanlı etkili bir tahmin etme tekniğidir. Çalışma prensibi olarak verideki nitelikleri sadece belirli bir sınıfa bağlı olması dikkate alınarak en uygun etikete göre sınıflandırma yapma işlemini gerçekleştirir [31-33].

2.2.4. YSA

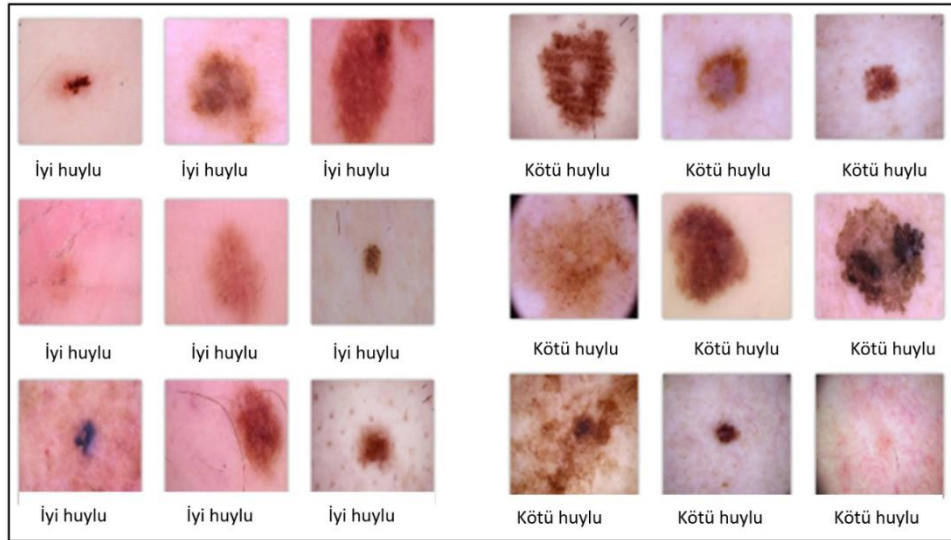
İnsan beyninin özelliklerinden biri olan öğrenmeyi dikkate alarak çalışır. Bir nöron ağı ve bu ağların ilişkilendirmiş ağırlık değerleri ve toplama fonksiyonlarından meydana gelmektedir. Yeni veriler oluşturabilme, tahmin etme ve keşfedebilme gibi yetenekleri, başka hiçbir yardım almadan otomatik bir şekilde gerçekleştirebilen bir yöntemdir [34-36].

2.2.5. LR

Bir veriye ait öğeyi gerçek değerli tahmin değişkenine eşleştirmek amacıyla kullanılmaktadır. Bu fonksiyonun öğrenilmesini kapsar. Temel amaç, verinin bilinen bir fonksiyon tipine uymasındır. Regresyon tekniğinde sürecinde verilen veriyi en iyi modelleyen fonksiyon bulunmaya çalışılmaktadır. Gerçek beklenen değerler ile tahmin değerleri arasındaki farka göre hangi fonksiyonun en başarılı performansa sahip olduğu belirlenir [37].

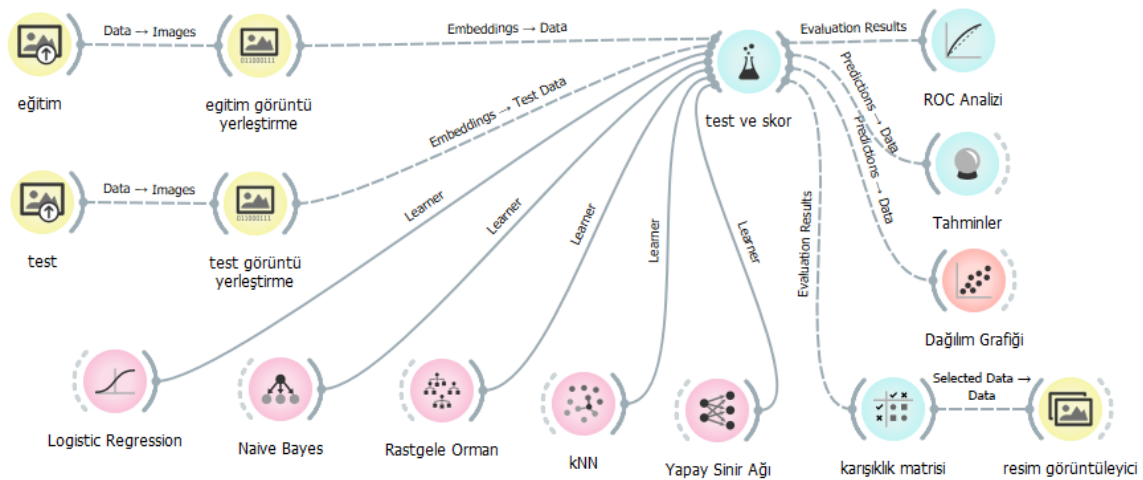
3. Bulgular ve Tartışma (Results and Discussion)

Melanom cilt kanserine ait görüntüler ile cilt beni fotoğrafları kullanılarak modelin eğitilmesi sağlanmıştır. Modele ait girdi değerlerinin tanımlanması yapılmıştır. Girdiye ait değerler iyi huylu (benign) ve kötü huylu (malignant) cilt lezyonlarıdır. Şekil 1'de çalışmada kullanılan iyi ve kötü huylu cilt lezyon görüntüleri verilmiştir.



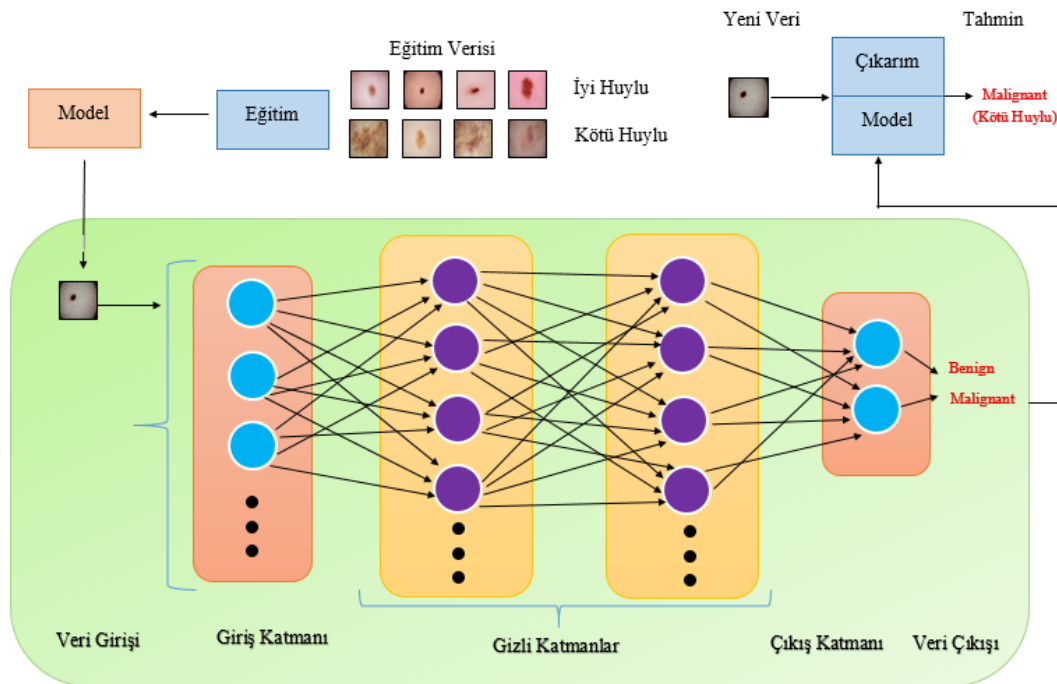
Şekil 1. İyi ve kötü huylu cilt lezyonları (Benign and malignant skin lesions)

Şekil 2'de hastaların cilt kanseri teşhisi için oluşturulan modele yer verilmiştir. Eğitim ve test olmak üzere iki bölümden oluşan görüntüler işlenmiştir. Görüntüler VGG-16 ağı modeli ile işlendikten sonra seçilen 5 adet algoritma ile modeller oluşturulmuştur. Bu modeller ile test ve skor (score) kısmında başarı değerleri tablolar ile verilmiştir. Daha sonraki kısımda ise test verilerinin algoritmalara göre görselleştirilmesi yapılmıştır. Her algoritmaya ait yüzdelik doğruluk değerleri, çıkış katmanına göre doğruluk ve kayıp (hata) grafiği vb. verilmiştir.



Şekil 2. Çalışmanın modeli (Model of the study)

Çalışmada, en iyi sonuç üreten veya en yüksek performansa sahip modele odaklanılmış olup sadece sözkonusu modelin genel tasarımına yer verilmiştir (Şekil 3). Bu bağlamda, yapay sinir ağına gelen veriler; giriş, gizli ve çıkış katmanlarından geçerek sınıflandırma yapılmış ve cilt lezyonu görüntüsünün melanom cilt kanseri olduğu tespit edilmiştir. Şekil 3'te, YSA mimarisine yer verilmiştir.



Şekil 3. YSA modeli ve sınıflandırma (ANN model and classification)

Bu algoritmalarda iyi huylu (benign) ve malignant (kötü huylu) kategorilerine göre Yapay Sinir Ağı'nın en başarılı sonuçlar verdiği görülmektedir. YSA tekniğinde, %98.8 iyi huylu ve %98.4 ile kötü huylu olarak elde edilmiştir. Tablo 2'de, çalışmada kullanılan algoritmalara yönelik Karışıklık Matrisi (Confusion Matrix) değerleri verilmiştir.

Tablo 2. Algoritma modellerinin karışıklık matrisi sonuçları (Confusion matrix results of algorithm models)

KNN				NB			
	İyi huylu	Kötü Huylu	Σ		İyi huylu	Kötü Huylu	Σ
İyi huylu	%81.4	%15.3	360	İyi huylu	%83.9	%24.1	360
Kötü huylu	%18.6	%84.7	300	Kötü huylu	%16.1	%75.9	300
Σ	385	275	660	Σ	336	324	660

RO				YSA			
	İyi huylu	Kötü Huylu	Σ		İyi huylu	Kötü Huylu	Σ
İyi huylu	%86.2	%17.7	360	İyi huylu	%98.8	%1.2	360
Kötü huylu	%13.8	%82.3	300	Kötü huylu	%1.6	%98.4	300
Σ	362	238	660	Σ	362	298	660

LR			
	İyi huylu	Kötü Huylu	Σ
İyi huylu	%84.3	%17.9	360
Kötü huylu	%15.7	%82.1	300
Σ	364	296	660

Tablo 3'te, çalışmada kullanılan algoritmaların eğitim ve test değerlerine göre yüzdelik değerleri verilmiştir. En başarılı sonuçlar, YSA yöntemi kullanılarak elde edilmiş olup Naive Bayes algoritması ise, eğitim %85.7 ve %80 test değeri ile en düşük performansa sahiptir.

Tablo 3. Modellerin karşılaştırılması (Comparison of models)

Model Adı	AUC	F1 Skor	Kesinlik	Duyarlılık
RO	0.920	0.844	0.844	0.844
LR	0.912	0.833	0.833	0.833
KNN	0.899	0.826	0.829	0.827
NB	0.857	0.800	0.803	0.800
YSA	0.991	0.987	0.986	0.986

Buna ilaveten, YSA algoritmasında bulunan farklı optimizasyonlar kullanılarak en başarılı eğitim ve test başarı yüzdeleri SGD optimizasyonu ile elde edilmiştir (Tablo 4).

Tablo 4. YSA'da kullanılan optimizasyon modellerinin karşılaştırılması (Comparison of optimization models used in ANN)

Optimizasyon Adı	Eğitim (%)	Test (%)
Adam	0.984	0.978
SGD	0.991	0.986
L-BFGS-B	0.988	0.981

Cilt kanseri teşhisinde ikili sınıflandırma algoritması ele alınırsa, bu model ile elde edilen veriler 1 ya da 0 olarak sınıflandırılır ve bu iki durumdan hangisine uyduğu belirlenir. 1 ya da 0 olma kararı, bir eşik (threshold) yardımıyla ele alınır. Eşik değerinden düşük olan veriler 0, eşik değerinden büyükler ise 1 olarak sınıflandırılır. Denetimli öğrenme tekniğinde modelin öngörülerini elde edilen verilerle test edilir ve modelin geçerliliği ve performansı bu şekilde denenmiş olur. İkili sınıflandırma modelinin tahminleri test verisi ile kıyas yapıldığında aşağıdaki durum oluşmaktadır.

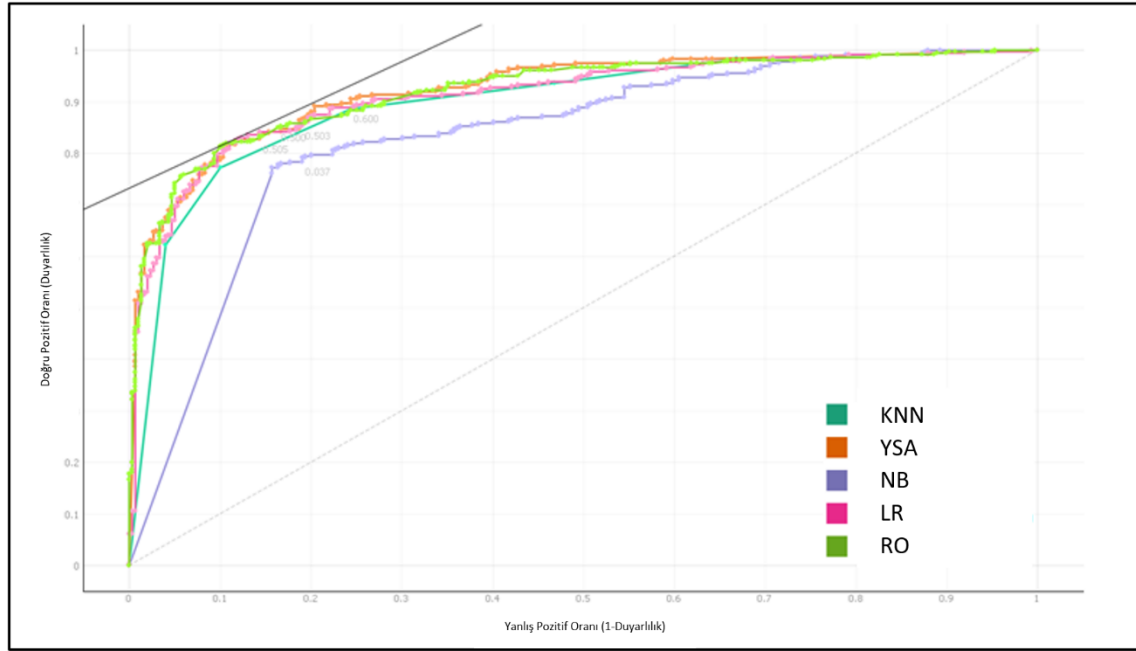
Model 1 olarak tahmin etmiş ve test verisi 1 ise: Doğru Pozitif (True Positive) (TP)

Model 0 olarak tahmin etmiş ve test verisi 0 ise: Doğru Negatif (True Negative) (TN)

Model 1 olarak tahmin etmiş ve test verisi 0 ise: Yanlış Pozitif (False Positive) (FP)

Model 0 olarak tahmin etmiş ve test verisi 1 ise: Yanlış Negatif (False Negative) (FN) olarak oluşturulur [38].

Buna göre Şekil 4'te hastaların cilt kanseri teşhisinde iyi huyluya göre TP ve FP grafiği ROC analiz grafiği verilmiştir. Çalışmada kullanılan algoritmalar için deneysel sonuçlar da Tablo 5'te verilmiştir. Duyarlılık oranı 0'dan başlamış, 0,986'ya kadar yükselmiştir. En yüksek doğruluk oranı YSA yöntemi ile elde edilmiştir.

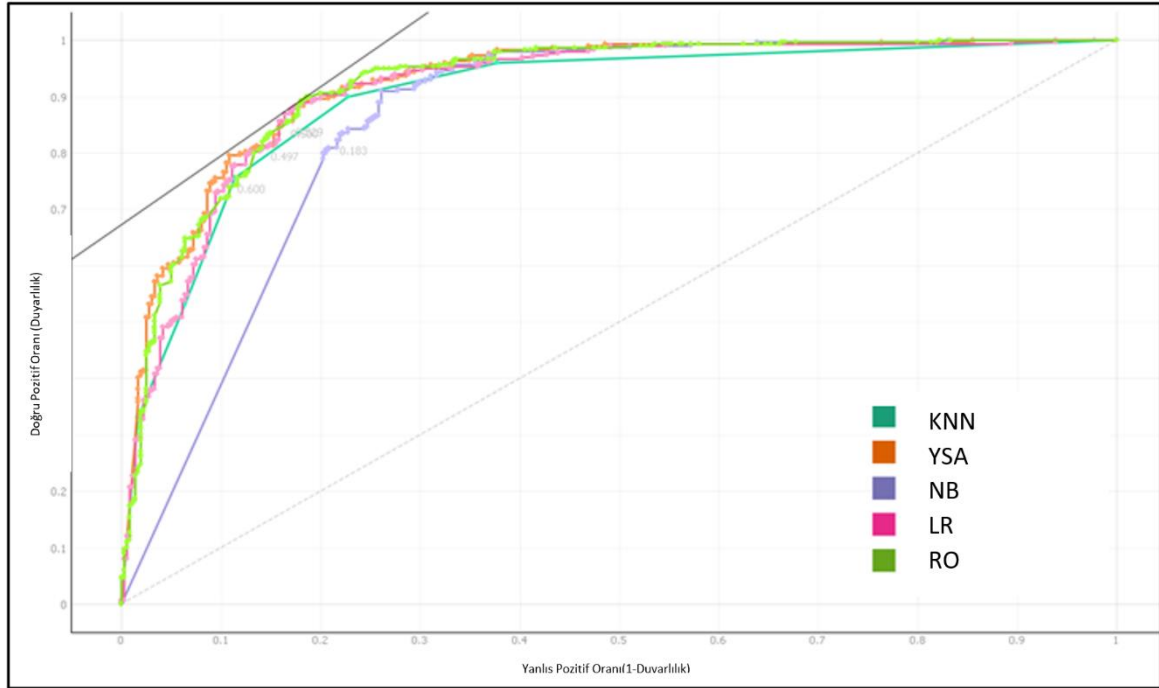


Şekil 4. İyi Huylu TP (Doğru Pozitif) ve FP (Yanlış Pozitif) ROC analiz grafiği (Benign TP (True Positive) and FP (False Positive) ROC analysis chart)

Tablo 5. Kullanılan modellerin performansları (iyi huylu) (Performances of the models used (well-behaved))

Model	Doğruluk Oranı (%)
KNN	81,4
RO	86,2
NB	83,9
LR	84,8
YSA	98,8

Şekil 5'te, cilt kanserinde kötü huylu olan hastaların çalışmada kullanılan algoritmalara göre ROC analiz grafiği verilmiştir. Deneysel sonuçlar da Tablo 6'da verilmiştir. Buna göre en yüksek doğruluk oranı YSA yöntemi ile elde edilmiştir.

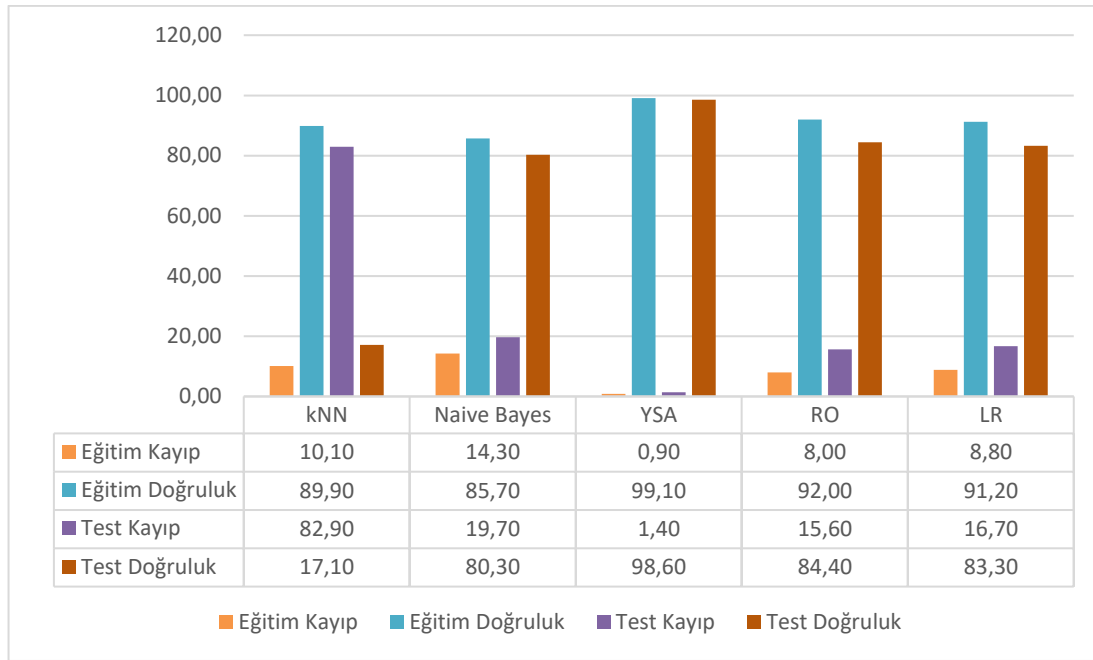


Şekil 5. Kötü huylu TP (Doğru Pozitif) ve FP (Yanlış Pozitif) ROC Analiz Grafiği (Malignant TP (True Positive) and FP (False Positive) ROC Analysis Graph)

Tablo 6. Kullanılan modellerin performansları (kötü huylu) (Performances of the models used (malignant))

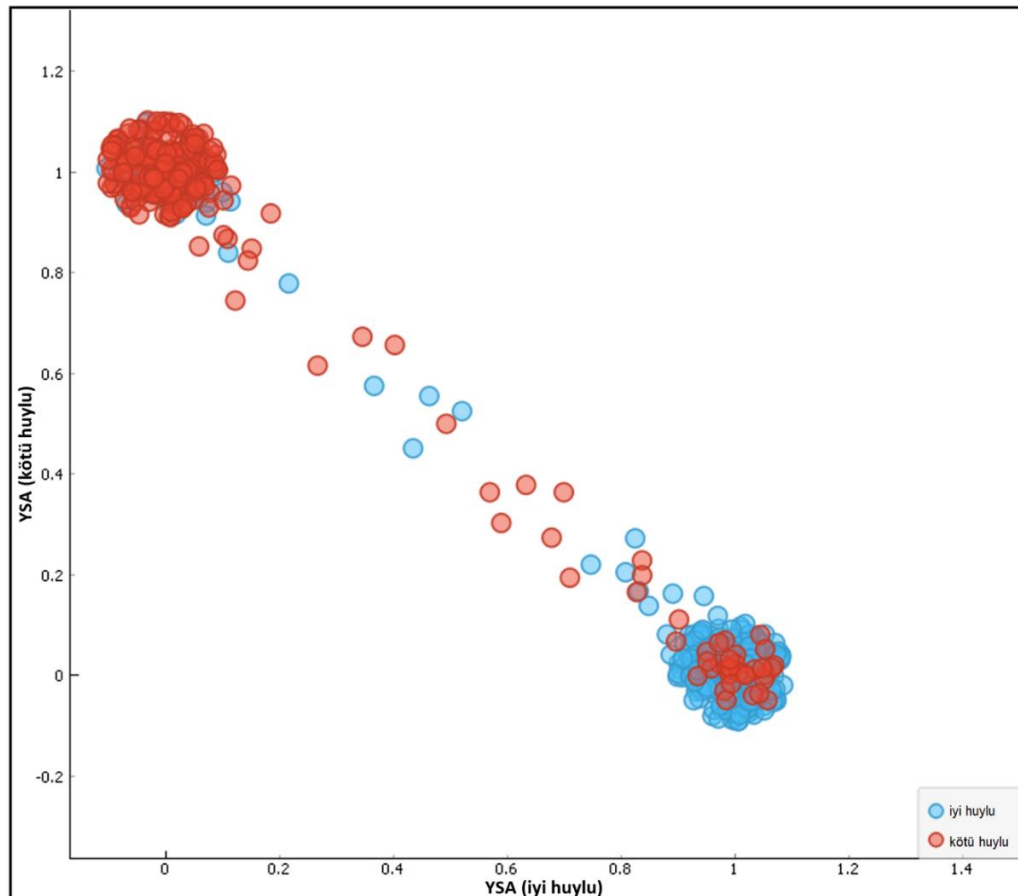
Model	Doğruluk Oranı (%)
KNN	84,7
RO	82,3
NB	83,9
LR	82,1
YSA	98,4

Şekil 6'da çalışmadaki algoritmalarla göre, doğruluk ve kayıp (hata) grafiğinde YSA algoritmasının hem eğitim hem de test sürecinde %99,1 ve %98,6 oranları ile daha başarılı olduğu görülmektedir.



Şekil 6. Kullanılan modellerin başarı oranları (Success rates of the models used)

Şekil 7’de test verisindeki YSA iyi huylu (benign) ve kötü huylu (malignant) çıktıları arasındaki dağılım diyagramı (Scatter Plot) verilmiştir. Bu diyagram aracılığıyla, değişkenler arasındaki ilişkinin sebebi görülmese de ilgili iki değişkenin arasında direkt olarak bir ilişki bulunup bulunmadığı ve bu ilişkinin ne derece güçlü olduğu görülmektedir.



Şekil 7. YSA dağılım grafiği (ANN scatter plot)

Son olarak, önerilen 5 adet algoritma modelleri ile gerçek veriler kullanılarak 15 (onbeş) farklı test gerçekleştirilmiştir. Bu testlerden elde edilen deneysel sonuçlar Tablo 7’de verilmiştir. Tablo 7’de elde edilen bulgular incelendiğinde, oluşturulan YSA modelinin melanom cilt kanserini tahmin etmede diğer kullanılan algoritmalara göre daha yüksek bir başarıya sahip olduğu görülmektedir.

Tablo 7. Modellerden elde edilen deneysel sonuçlar (Experimental results obtained from models)

No	KNN	Hata (KNN)	RO	Hata (RO)	NB	Hata (NB)	LR	Hata (LR)	YSA	Hata (YSA)
1	79,4	20,6	81,1	18,9	80,4	19,6	89,4	10,6	98,7	1,3
2	81,41	18,59	85,3	14,7	80,3	19,7	81,41	18,59	98,5	1,5
3	90,14	9,86	84,8	15,2	83,2	16,8	85,14	14,86	97,6	2,4
4	83,6	16,4	80,6	19,4	82,17	17,83	83,6	16,4	98,8	1,2
5	79,15	20,85	88,81	11,19	81,65	18,35	79,15	20,85	98,6	1,4
6	81,96	18,04	89,3	10,7	80,78	19,22	81,96	18,04	98,7	1,3
7	81,6	18,4	83,5	16,5	79,5	20,5	81,6	18,4	99	1
8	81,6	18,4	85,41	14,59	78,24	21,76	81,6	18,4	98,3	1,7
9	85,4	14,6	80,39	19,61	77,53	22,47	95,4	4,6	98,4	1,6
10	82,6	17,4	86,98	13,02	83,46	16,54	82,6	17,4	98,9	1,1
11	80,56	19,44	81,63	18,37	84,79	15,21	82,56	17,44	98,7	1,3
12	85,8	14,2	86,74	13,26	80,2	19,8	85,8	14,2	98,6	1,4
13	85,04	14,96	81,12	18,88	83,88	16,12	85,04	14,96	98,7	1,3
14	81,96	18,04	85,15	14,85	75,1	24,9	79,96	20,04	98,7	1,3
15	85,01	14,99	85,27	14,73	80,32	19,68	80,01	19,99	99	1
Ortalama	83,02	16,98	84,41	15,59	80,77	19,23	83,68	16,32	98,61	1,39

Tablo 8, önerilen YSA ile literatürde yer alan diğer çalışmaları karşılaştırmayı temsil etmektedir. Tablo 8 incelendiğinde, Namozov ve ark. tarafından gerçekleştirilen çalışma sonucu hariç, diğer üç modelin birbirine yakın doğruluk oranlarına sahip olduğu, önerilen modelin ise diğer tüm modellerden daha yüksek performansa sahip olduğu açıkça görülmektedir.

Tablo 8. Literatürde elde edilen diğer sonuçlarla karşılaştırma (Comparison with other results obtained in the literature)

Çalışma	Model	Doğruluk
Farooq ve ark., [39]	Inception-V3 ve MobileNetV1	86%
Demir [40]	MobileNetV2, L1Norm, RD ve DVM	88.35%
Namozov ve ark.,[41]	CNN model with keratosis/vascular LeNet approach	95.86%
Harangi ve ark., [42]	CNN as single neural-net architecture	83.8%
Bütüner ve Calp (Önerilen Metot)	YSA tekniği	98.4%

4. Sonuçlar ve Öneriler (Conclusions and Recommendations)

Bu çalışmada, hastalardan alınan gerçek görüntüler kullanılarak KNN, NB, RO, LR ve YSA yöntemleri ile cilt kanseri hastalığı tahmin edilmiştir. Bu çerçevede, her bir yöntem için birer model oluşturulmuş ve gerçek verilerle test edilmiştir. Oluşturulan modeller içerisinde en iyi model belirlenmiş ve önerilmiştir. Önerilen model, cilt kanseri tespiti için hem zamandan hem de maliyetten büyük tasarruf sağlamaktadır. Ayrıca oluşturulan model, çok daha fazla veri setine ulaşarak kapsamlı bir değerlendirme yapabilen bir model ile daha sağlıklı sonuçlara ulaşılması sağlanabilir. Bu verilerin kullanılabilmesi, küresel, ulusal veya bölgesel çapta çeşitli analizlerin ve yeni bilimsel çalışmaların yapılmasına şüphesiz katkı sunacaktır.

Bununla birlikte, bu tür çalışmalar sağlıkla ilgili tüm alanlarda uygulanarak önemli gelişmelere ve sonuçlara katkı sağlayacaktır. Çalışmada, önerilen makine öğrenmesi tekniklerinin erken melanom cilt kanseri tanı

koymada geleneksel biyopsi yöntemlerinden daha etkili olduğu görülmüştür. Bu noktada, özellikle YSA yöntemi ile sınıflandırma ve tanıma uygulamalarında karar verme sürecinde çok etkili olduğu kanaatine varılmıştır. Geliştirilen bu sistemin melanom cilt kanserinin otomatik tanılama amacına hizmet etmesi beklenmektedir.

Bu çalışmanın sonraki aşamasında, önceden eğitilmiş modellerle birlikte farklı özellik seçme yöntemleri kullanılarak diğer veri kümelerinde de benzer başarı oranlarının elde edilmesi amacıyla araştırmalar yapılarak geliştirilecektir. Bu geliştirilecek araştırma veriseti genişletilerek gerçekleştirilecek ve elde edilen sonuçlar analiz edilecektir. Sonuç olarak yapılan çalışma, düşük hata oranı minimum maliyetle tanı ve teşhisin sağlanabilmesi ve makine öğrenme tekniklerinin sağlık alanında kullanılmasında etkili olduğunu ortaya koymuştur.

Çıkar Çatışması Beyanı (Conflict of Interest Statement)

Yazarlar tarafından herhangi bir çıkar çatışması bildirilmemiştir.

Kaynaklar (References)

- [1] O. Baykara, "Kanser Tedavisinde Güncel Yaklaşımlar," *Balıkesir Sağlık Bilimleri Dergisi*, vol.5, no. 3, pp. 154-165, 2016.
- [2] WHO, "Cancer," *WHO-Home/Newsroom/Fact sheets/Detail/Cancer*, [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/cancer>. [Accessed: 3.02. 2023].
- [3] Z. Faisal and N. Abbadi, "New Segmentation Method for Skin Cancer Lesions," *Journal of Engineering and Applied Sciences*, vol. 12, no. 21, pp. 5598-5602, 2017.
- [4] K. H. Güngör, "Metastaz Yapmamış Melanoma ve Melanoma Dışı Deri Kanseri için Geliştirilmiş Olan Deri Kanseri İlişkili Yaşam Kalitesi Ölçeğinin (Dkykö) Türkçe Geçerlilik ve Güvenilirliğinin Araştırılması," *Tıpta Uzmanlık Tezi*, Ankara Üniversitesi Tıp Fakültesi, Ankara, 2016.
- [5] E. Acer ve H. K. Erdoğan, "Sık Görülen Deri Kanserlerinin Epidemiyolojisi," *Dermatoloji ve Halk Sağlığı Özel Sayısı*, vol.4, no. 1, ss. 52-60, 2019. doi:10.35232/estudamhsd.499257
- [6] A. Staff, "Cancer Facts and Figures Atlanta: American Cancer Society," *Cancer*, vol. 9, pp. 19-20, 2018.
- [7] A. Sümen ve S. Öncel, "Türkiye'de cilt kanseri ve güneşten korunmaya yönelik yapılan araştırmaların incelenmesi," *Türkiye Klinikleri Hemşirelik Bilimleri Dergisi*, vol. 10, no. 1, ss. 59-69, 2018. doi:10.5336/nurses.2017-56858
- [8] A. M. Glazer, D. S. Rigel, R. R. Winkelmann and A. S. Farberg "Clinical diagnosis of skin cancer: enhancing inspection and early recognition," *Dermatologic Clinics*, vol. 35, no. 4, pp. 409-416, 2017. doi:10.1016/j.det.2017.06.001
- [9] C. P. Akyel, ve N. Arıcı, "Cilt kanserinde kıl temizliği ve lezyon bölütlemesinde yeni bir yaklaşım," *Politeknik Dergisi*, vol. 23, no. 3, ss. 821-828, 2020.
- [10] Ü. Yılmaz, *Yapay Sinir Ağları ile Melanom Cilt Kanseri Tespiti, Geleceğin Dünyasında Bilimsel ve Mesleki Çalışmalar*, Ekin Basım Yayın Dağıtım, 2019, pp.104-114.
- [11] K. Lacy and W. Alwan, "Skin cancer," *Medicine*, vol. 41, no. 7, pp. 402-405, 2013. doi:10.1016/j.mpmed.2013.04.008
- [12] E. Craythorne and F. Al-Niami, "Skin cancer," *Medicine*, vol. 45, no. 7, pp. 431-434, 2017. doi:10.1016/j.mpmed.2017.04.003
- [13] A. H. Alasadi and B. M. Alsafty, "Early detection and classification of melanoma skin cancer," *Int. Journal of Information Technology and Computer Science*, vol. 7, no.12, pp. 67-74, 2017.
- [14] P. Das, N., Deshmukh, N., Badore, C., Ghulaxe, and P., Patel, "A review article on melanoma," *Journal of Pharmaceutical Sciences and Research*, vol. 8, no. 2, pp. 112-117, 2016.
- [15] F. Alendar, I., Drljević, K., Drljević, and T., Alendar, "Early detection of melanoma skin cancer," *Bosnian Journal of Basic Medical Sciences*, vol. 9, no. 1, pp. 77-80, 2009. doi:10.17305/bjbs.2009.2861

- [16] M. Demirci ve İ. Yabanova, "Model Tabanlı Tasarım Metotları Kullanılarak Gerçek Zamanlı Bir Görüntü İşleme Sisteminin Tasarımı ve Gerçeklemesi," *Politeknik Dergisi*, vol. 22, no. 4, pp. 827-838, 2019. doi:10.2339/politeknik.423603
- [17] Ş. Sağıroğlu ve E. Beşdok, "A Novel Approach for Image Denoising Based on Artificial Neural Networks," *Politeknik Dergisi*, vol. 15, no. 2, pp. 71-86, 2012. doi:10.2339/2012.15.2, 71-86
- [18] A. Kassam, "Segmentation of Skin Cancer by Using Image Processing Techniques," Master Thesis, Yıldız Technical University Department Of Computer Engineering, İstanbul, 2016.
- [19] T. Lee, V., Ng, R., Gallagher, A., Coldman, and D., McLean, "Dullrazor: A Software Approach to Hair Removal from Images," *Computers in biology and medicine*, vol. 27, no. 6, pp. 533-543, 1997. doi:10.1016/S0010-4825(97)00020-6
- [20] E.C. Celebi, Y., Alp Aslandogan, W. V., Stoecker, H., Iyatomi, H., Oka, and X. Chen., "Unsupervised Border Detection in Dermoscopy Images," *Skin Research and Technology*, vol. 13, no. 4, pp. 454- 462, 2007. doi:10.1111/j.1600-0846.2007.00251.x
- [21] M.Z. Alom, T. Aspiras, T.M. Taha and V. K. Asari, "Skin Cancer Segmentation and Classification with NABLA-N and Inception Recurrent Residual Convolutional Networks," 2019. doi:10.48550/arXiv.1904.11126
- [22] M. Toğaçar, Z. Cömert, and B. Ergen, "Intelligent skin cancer detection applying autoencoder, MobileNetV2 and spiking neural networks," *Chaos, Solitons and Fractals*, vol. 144, no. 1, pp.110714, 2021. doi:10.1016/J.CHAOS.2021.110714
- [23] P. P. Tumpa and M. A. Kabir, "An artificial neural network based detection and classification of melanoma skin cancer using hybrid texture features," *Sensors International*, vol. 2, no. 1, pp.100128, 2021. doi:10.1016/J.SINTL.2021.100128
- [24] A.M. Alqudah, H. Alquraan and I.A. Qasmieh, "Segmented and non-segmented skin lesions classification using transfer learning and adaptive moment learning rate technique using pretrained convolutional neural network," *Journal of Biomimetics, Biomaterials and Biomedical Engineering*, vol. 42, pp. 67-78, 2019. doi:10.4028/www.scientific.net/JBBBE.42.67
- [25] T.J. Brinker, A. Hekler, A.H. Enk, C. Berking, S. Haferkamp, A. Hauschild and J.S. Utikal, "Deep neural networks are superior to dermatologists in melanoma image classification," *European Journal of Cancer*, vol. 119, pp. 11-17, 2019. doi:10.1016/j.ejca.2019.05.023
- [26] K.M. Hosny, M.A. Kassem and M.M. Foad, "Classification of skin lesions using transfer learning and augmentation with AlexNet," *PloS one*, vol. 14, no. 5, pp. 217-293, 2019. doi:10.1371/journal.pone.0217293
- [27] J. Han, J. Pei and Kamber, M. *Data Mining: Concepts and Techniques*, Elsevier, 2011, pp. 2-10.
- [28] D. Kilinc, E. Borandag, F. Yucalar, V. Tunali, M. Simsek, and A. Ozcift, "Classification of Scientific Articles Using Text Mining with KNN Algorithm and R Language," *Marmara Journal of Pure and Applied Sciences*, vol.3, pp. 89-94, 2016. doi:10.7240/mufbed.69674
- [29] T. Savaş and S. Savaş, "Tekdüzen Kaynak Bulucu Yoluyla Kimlik Avı Tespiti için Makine Öğrenmesi Algoritmalarının Özellik Tabanlı Performans Karşılaştırması," *Politeknik Dergisi*, vol. 25, no. 3, ss. 1261-1270, 2022. doi:10.2339/politeknik.1035286
- [30] L. Breiman, *Random Forest*. Machine Learning, Statistics Department-University of California, vol. 45, pp. 5-32, 2001. doi:10.1023/A:1010933404324
- [31] D.D. Lewis, "Naive Bayes at forty: The independence assumption in information retrieval," *Proceedings of the Tenth European Conference on Machine Learning*, Part of the Lecture Notes in Computer Science book series (LNAI), vol.1398, pp. 4-15, 1998.
- [32] H.K. Yıldız, M. Gençtav, N. Usta, B. Diri and M. F. Amasyalı, "Metin Sınıflandırmada Yeni Özellik Çıkarımı," *15. Sinyal İşleme ve İletişim Uygulamaları Kurultayı*, 2007.
- [33] S. Savaş, "Naive Bayes Sınıflandırıcısı," *Teori ve Uygulamada Makine Öğrenmesi*, Nobel Akademik Yayıncılık Eğitim Danışmanlık Tic. Ltd. Şti, ss. 69-92, 2022.
- [34] C. Yerdelen, "Mevsimlik Kar Erimesinin Yapay Sinir Ağları Yöntemi ile Tahmin Edilmesi," *Sakarya Üniversitesi Mühendislik ve Mimarlık Fakültesi Dergisi*, vol. 21, pp. 3-4, 2006.
- [35] M.H. Calp and U. Kose, "Estimation of burned areas in forest fires using artificial neural networks," *Ingeniería Solidaria*, vol. 16, no. 3, pp. 1-22, 2022.
- [36] M. H. Calp and M. A. Akcayol, "Yazılım projeleri için yapay sinir ağlarına dayalı web tabanlı risk yönetim sisteminin tasarımı ve gerçekleştirilmesi: WEBRISKIT," *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, vol. 26, no. 5, ss. 993-1014, 2020. doi:10.5505/pajes.2019.29964

- [37] M.H. Dunham, "Data Mining introductory and advanced topics," New Jersey: Prentice Hall, 2003.
- [38] B. Ekici and H. Takcı, "Bilgisayar Ağlarında Anomali Tespiti Yaklaşımı ile Saldırı Tespiti," *Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, vol. 22, no. 5, ss. 1016-1027, 2022. doi:10.35414/akufemubid.1114906
- [39] M.A. Farooq, A., Khatoun, V., Varkarakis, and P. Corcoran, "Advanced Deep Learning Methodologies for Skin Cancer Classification in Prodromal Stages 2020," *arXiv preprint arXiv:2003.06356*.
- [40] F. Demir, "Derin öğrenme tabanlı yaklaşımla kötü huylu deri kanserinin dermatoskopik görüntülerden saptanması," *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, vol. 33, no. 2, ss. 617-624, 2021.
- [41] A. Namozov, D. Ergashev and Y.I. Cho, "Adaptive activation functions for skin lesion classification using deep neural networks," in *Proceedings- 2018 Joint 10th International Conference on Soft Computing and Intelligent Systems and 19th International Symposium on Advanced Intelligent Systems*, SCIS-ISIS 2018, ss. 232- 235, 2018. doi:10.1109/SCIS-ISIS.2018.00048
- [42] B. Harangi, A. Baran and A. Hajdu, "Classification of Skin Lesions Using An Ensemble of Deep Neural Networks," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, EMBS, 2018-July, pp. 2575-2578, 2018. doi:10.1109/EMBC.2018.8512800

This is an open access article under the CC-BY license

