

## Cilt Kanseri Görüntü Sınıflandırması için Görüntü Ön İşlemenin Evrişimsel Sinir Ağları Performansı Üzerindeki Etkileri

Burcu BAKİR-GUNGOR<sup>ID\*</sup><sup>1</sup>, Beyhan ADANUR DEDETURK<sup>ID</sup><sup>2</sup>, Kasım TAŞDEMİR<sup>ID</sup><sup>3</sup>

\*<sup>1,2,3</sup>Abdullah Gül Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği, KAYSERİ

(Alınış / Received: 11.05.2022, Kabul / Accepted: 18.07.2022, Online Yayınlanma / Published Online: 23.08.2022)

### Anahtar Kelimeler

Cilt Kanseri,  
Sınıflandırma,  
Ön İşleme,  
Evrişimsel Sinir Ağları

**Öz:** Cilt kanseri, dünya çapında yaygın olarak karşılaşılan kanser türleri arasındadır. Günümüzde pek çok kanser vakasının yanlış ya da geç teşhisi sonucunda, hasta ölümleri de dahil olmak üzere ciddi problemler yaşanmaktadır. Bu çalışmada, evrişimli sinir ağları kullanılarak cilt kanseri sınıflandırması problemi incelenmiştir. Çalışmanın temel amacı, farklı öğrenme mimarilerini karşılaştırmak yerine, görüntüleri farklı ön işlemlere tabi tutup, bu işlemin kullanılan mimari performansına etkisini incelemektir. Bu amaç doğrultusunda, ISIC 2018 Cilt Görüntü Analizi Yarışması'na ait veri seti kullanılarak, iki farklı görüntü ön işleme yöntem dizisi ResNet50 mimarisi için uygulanmıştır. Bunlardan birincisinde sırasıyla ikili ve otsu eşikleme, CLAHE dönüşümü uygulanırken, ikincisinde morfolojik filtreleme, renk normalizasyonu ve dolgu işlemleri uygulanmıştır. F1 Puanı başta olmak üzere farklı performans metrikleri baz alındığında, cilt kanseri görüntüleri üzerinde ikinci ön işleme yöntem dizisinin performans iyileştirmesi yapabildiği gösterilmiştir.

## Impact of Image Preprocessing on Skin Cancer Image Classification Using Convolutional Neural Networks

### Keywords

Skin Cancer,  
Classification,  
Preprocessing,  
Convolutional Neural  
Networks

**Abstract:** Skin cancer is one of the most common types of cancer worldwide. Nowadays, as a result of wrong or late diagnosis of many cancer cases, serious problems including patient deaths are faced. This study deals with the skin cancer classification problem using convolutional neural networks. Instead of comparing different learning architectures, this research effort aims to preprocess the images and examine the effect of these processes on the performance of the architecture. For this purpose, using the ISIC 2018 Skin Image Analysis Competition dataset, two different sets of preprocessing methods have been applied on the ResNet50 architecture. In the first set up, binary and OTSU's thresholding and CLAHE transformation were applied, respectively. In the second set up, morphological filtering, color normalization and filling operations were applied. Based on different performance metrics including F1 Score, it has been shown that the second set of preprocessing methods can improve the performance for the skin cancer classification problem.

\*İlgili Yazar, email: bakirburcu@gmail.com

### 1. Giriş

Cilt kanseri, dünya çapında artış gösteren en yaygın kanser türleri arasındadır. Erken teşhisi ve doğru tedavi seçimi yapıldığı takdirde, sağ kalım oranlarının yüksek olduğu görülmektedir. Geleneksel olarak, doktorlar cilt kanseri tespiti için çıplak gözlerini kullanmaktadır. Ancak bu yöntem birçok durumda doğru tespit oranını azaltmaktadır

[1]. Bu nedenle günümüzde, makine öğrenmesi ve derin öğrenme yöntemlerinin bu alandaki kullanımı önem kazanmıştır, ve hastalık görüntülerini sınıflandırma tekniklerinin kullanımı artmıştır.

Görüntü sınıflandırma yaklaşımları iki ana başlık altında incelenebilir. Birinci kategori, Karar Ağacı, Destek Vektör Makineleri, Bulanık Ölçümler vb. gibi geleneksel makine öğrenme algoritmalarını içerir. İkinci kategori ise, evrimsel sinir ağları (Convolutional Neural Networks, CNN), otomatik kodlayıcılar vb. gibi derin öğrenmeye dayalıdır. Zaman içerisinde araştırmacılar derin öğrenme yöntemlerini kullanmayı tercih etmişlerdir. Bunun başlıca nedeni, derin öğrenme yöntemlerinin 2012 öncesindeki hiçbir makine öğrenimi yönteminde görülmemiş olan yüksek doğrulukta sonuçlar elde edebilmiş olmasıdır [2].

Derin bir sinir ağı, hem iyi huylu, hem de kötü huylu olmak üzere her iki hastalık kategorisine ait binlerce görüntü üzerinde eğitilebilir. Bu eğitim sonucunda elde edilen model, doğrusal olmayan etkileşimleri öğrenerek, yeni bir görüntünün iyi huylu veya kötü huylu bir sınıfa karşılık gelip gelmediğini tahmin edebilir. Bu otomasyon, yalnızca hem yanlış pozitifleri, hem de yanlış negatifleri önlemede daha yüksek bir verimliliğe sahip olmakla kalmaz, aynı zamanda tespit için harcanan zamanı da azaltır. Derin öğrenme ve bilgisayarlı görme teknikleri kullanılarak cilt kanseri sınıflandırması alanında yayınlanmış çok sayıda çalışma bulunmaktadır. Bu çalışmalar, sınıflandırma, segmentasyon ve algılama, farklı filtre türleri kullanarak görüntü işleme vb. dahil olmak üzere birçok farklı yaklaşımları kullanmaktadır [3].

Esteva ve diğerleri [4] cilt lezyonlarını sınıflandırmak için AdaBoost sınıflandırma yöntemini kullanmıştır. Xu ve diğerleri [5] sağlam bir tanı sistemi elde edebilmek için lezyon tipi, doku, renk gibi farklı özellik setleri ve sinir ağlarını kullanmıştır. Şimdiye kadar sunulan örnekler, yalnızca geleneksel makine öğrenimi tekniklerini kullanmıştır. Ancak son zamanlarda derin öğrenmenin daha doğru sonuçlar verdiği gösterilmiştir [6]. Bunun nedeni, derin öğrenme ile özellik çıkarma işleminin eğitimle kendiliğinden ortaya çıkması ve böylece tasarımcının etkisini azaltmasıdır.

Lopez ve diğerleri [7], 129.450 klinik cilt kanseri görüntüsünü sınıflandırmak için önceden eğitilmiş bir GoogleNet Inception v3 CNN modeli kullanarak, cilt kanseri sınıflandırmasında ön plana çıkmıştır. Dorj ve diğerleri [8], malign melanomun sınıflandırılması için ISBI 2016 meydan okuma veri setinde 50'den fazla katmana sahip bir evrimsel sinir ağı (ESA) geliştirmiştir. 2018'de Brinker ve diğerleri [9], dermoskopi görüntülerinin ikili sınıflandırma problemi için derin bir evrimsel sinir ağı kullanmıştır. Rezvantalab ve diğerleri [10], klinik cilt kanseri görüntülerinin 4 tanı kategorisinde sınıflandırılması için derin bir konvolüsyonel sinir ağı yaklaşımıyla birleştirilmiş Destek Vektör Makinelerini kullanan bir algoritma geliştirmiştir. Codella ve diğerleri [11], 12 cilt hastalığının klinik görüntülerini sınıflandırmak için derin bir evrimsel sinir ağı kullanmıştır.

Mevcut çalışmalar incelendiğinde şu gözlemler yapılabilir:

- i) Nesne algılama ve sınıflandırma problemlerinde, evrimsel sinir ağları yöntemi diğer yöntemlere kıyasla daha yüksek doğruluk değeri üretmesi nedeniyle çoğunlukla ilk tercih edilen yöntem olmuştur [12].
- ii) Karşılaştırma için farklı öğrenme mimarileri kullanılmaktadır.
- iii) Performansı artırmak için farklı yaklaşımları birleştirerek çalıştıran yeni algoritmalar kullanılmaktadır.
- iv) Değişik kanser türlerinin ve cilt hastalıklarının kategorik olarak sınıflandırılması aktif çalışma alanlarıdır.

Bu çalışmada, evrimsel sinir ağları kullanılarak cilt kanseri sınıflandırması sorunu ele alınmıştır. Ağı hem eğitmek, hem de test etmek için halka açık ISIC 2018 Cilt Görüntü Analizi Yarışması'na ait veri seti kullanılmıştır. Bu veri seti 7 farklı cilt hastalığına ait dermoskopik görüntüler içermektedir [13]. Bu çalışma, ikili tipte (iyi huylu-kötü huylu) olan cilt kanseri sınıflandırması problemini incelemektedir. Bu nedenle, veri seti yeniden düzenlenmiştir. Bu çalışmanın temel amacı, farklı öğrenme mimarilerini karşılaştırmak yerine, görüntüleri farklı ön işlemlere tabi tutup, bu işlemlerin kullanılan mimari performansına etkisini incelemektir. Bu amaç doğrultusunda veriler üzerinde iki farklı ön işleme dizisi uygulanmıştır. Bunlardan birincisinde sırasıyla; ikili ve otsu eşikleme, belirli oranda kırma ve CLAHE dönüşümü uygulanmıştır [14]. İkincisinde ise sırasıyla; yeniden boyutlandırma, görüntülerdeki kırımları yok etmek için morfolojik filtreleme, renk normalizasyonu, dolgu ve kırma işlemleri uygulanmıştır. Çalışmada kullanılmak üzere, ResNet50 mimarisi tercih edilmiştir. Sonuçların değerlendirilmesi için doğruluk, kesinlik, geri çağırma, F1 puanı ve ROC-AUC metrikleri kullanılmıştır. Bu çalışma ile, dermoskopi görüntüleri üzerinde eğitilmiş derin öğrenme mimarilerinin farklı ön işleme yöntemleri ile birlikte kullanıldığında

performans üzerindeki etkilerini incelemeyi, bulgularımız ile, ön işleme yöntem seçiminin de mimari tercihi kadar önemli olduğunu vurgulayabilmeyi hedeflemekteyiz. Çalışmanın temel katkıları şu şekilde özetlenebilir:

- i) Mevcut cilt kanseri sınıflandırma çalışmaları incelenmiş, en yaygın kullanılan yöntemin evrimsel sinir ağı olduğu tespit edilmiştir. Literatüre bakıldığında, hemen hemen her çalışmanın farklı CNN topluluklarını bir arada kullandığı gözlenmiştir. Bu tespitler doğrultusunda, bu çalışmada tek bir mimari seçilip veriler üzerinde iki farklı ön işleme dizisi uygulanmıştır. Böylelikle ideal bir mimarinin vazgeçilmez parçası olan ön işleme yöntemlerinin seçilen yöntemin performansı üzerindeki etkisi analiz edilmiştir.
- ii) Aynı veri setini kullanan çalışmalar, veri setinin tümünü kullanıp, kategorisel sınıflandırma yöntemini tercih etmişlerdir. Bildiğimiz kadarıyla kullanılan veri seti üzerinde ikili tipte sınıflandırma yönteminin performans analizi yapılmamıştır. Bu sebeple çalışmamızda 7 hastalık kategorisi içeren veri setinden ikili sınıflandırma yapabilmek adına 4 kategori seçilip veri dağılımlarına göre yeni bir veri topluluğu oluşturulmuştur. Bu deney sayesinde, ikili tipte sınıflandırma yönteminin, ilgili veri seti üzerindeki performansı da analiz edilmiştir.

## 2. Materyal ve Yöntem

### 2.1. Veri Kümesi

Bu çalışmada, ISIC 2018 Cilt Görüntü Analizi Yarışması'na ait veri seti kullanılmıştır. Bu veri seti, melanom, melanositik nevus, bazal hücreli karsinom, aktinik keratoz, iyi huylu keratoz, dermatofibroma ve vasküler lezyonu olmak üzere 7 farklı cilt hastalığına ait dermoskopik görüntüler içermektedir. Bu çalışmada, kategorisel olarak hastalık sınıflandırması yerine, ikili tipte (iyi huylu-kötü huylu) cilt kanseri sınıflandırması problemi ele alınmıştır. Bunun sebebi, bu veri setini kullanan çalışmaların hemen hemen hepsinin kategorisel sınıflandırma yapması, ve bizim bu çalışmada farklı olarak ikili tipte sınıflandırma yönteminin ilgili veri seti üzerindeki performansını değerlendirmek istememizdir. Bunun için veri seti yeniden düzenlenmiştir. Vasküler lezyon, aktinik keratoz ve melanositik nevus hastalık görüntülerinin bazıları iyi huylu, bazıları kötü huylu olduğu için bu görüntüler çalışmaya dahil edilmemiştir. Melanom ve bazal hücreli karsinom görüntüleri kötü huylu, dermatofibroma ve iyi huylu keratoz görüntüleri ise iyi huylu kategoriye dahil edilmiştir. Bunun sonucunda kötü huylu sınıfta 1627 görüntü, iyi huylu sınıfta ise 1216 görüntü yer almıştır. Her sınıfın verilerinin %80'i eğitim, %20'si test için kullanılmıştır ve veri seti numpy dizisi olarak işlenirken girdi boyutu 224 olarak düzenlenmiştir. Böylece görüntü boyutları 224 × 224 olarak düzenlenmiştir. Veri dağılımı Tablo 1'de özetlenmiştir.

**Tablo 1.** ISIC 2018 Cilt Görüntü Analizi Yarışması'na ait Veri Setinin Dağılımı

	Toplam	Eğitim (%80)	Test (%20)
<b>İyi Huylu</b>	1216	973	243
<b>Kötü Huylu</b>	1627	1302	325

### 2.2. Ön İşleme

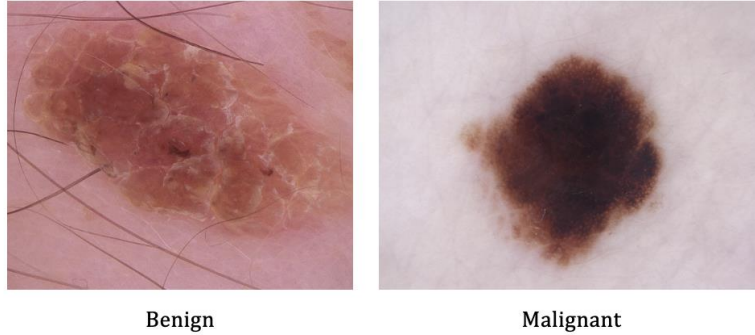
Öncelikle veri kümesi boyutunu orijinal veri kümesi boyutunun neredeyse iki katına çıkarmak için, kırpma, yakınlaştırma, çevirme ve parlaklık değişikliği gibi literatürde sıkça kullanılan [12] veri büyütme yöntemlerinden faydalanılmıştır. Bu çalışmanın temel amacı, görüntüleri farklı ön işlemlere tabi tutup, bu işlemlerin kullanılan mimari performansına etkisini incelemektir. Görüntüler incelediğinde, görüntülerin başlıca görüntü azaltılmasına, belirli oranlarda istenen nesnenin ortaya çıkarılmasına ve kontrast iyileştirilmesine ihtiyacının olduğu belirlenmiştir. Bu sebeple ve ihtiyaç duyulan gereksinimleri karşılayacağı düşüncesiyle, verilere iki farklı ön işleme dizisi uygulanmıştır. Bunlardan birincisinde sırasıyla; ikili ve otsu eşikleme, belirli oranda kırpma (ROI işlemleri için bölge sınırı +20 piksel tolerans değeri olarak belirlenmiştir. Bu sınır yardımıyla da orijinal görüntü kırpılmıştır) ve CLAHE dönüşümü uygulanmıştır [14]. CLAHE, kontrastın aşırı yükseltilmesiyle ilgilenen uyarlanabilir histogram eşitlemenin (AHE) bir çeşididir. CLAHE, görüntünün tamamı yerine, görüntüde karo adı verilen küçük bölgeler üzerinde çalışır. Komşu karolar daha sonra yapay sınırları kaldırmak için çift doğrusal enterpolasyon kullanılarak birleştirilir. İkincisinde ise sırasıyla; yeniden boyutlandırma, görüntülerdeki kılları yok etmek için morfolojik filtreleme, renk normalizasyonu, dolgu ve kırpma işlemleri uygulanmıştır.

Eşikleme, değeri belirli bir eşik üzerinde olan pikselleri tam olarak beyaza ayarlayıp, diğer pikselleri siyaha ayarlama ve gri ölçekli bir görüntüden siyah beyaz görüntü oluşturma işlemidir. Otsu yöntemi, kümeleme tabanlı

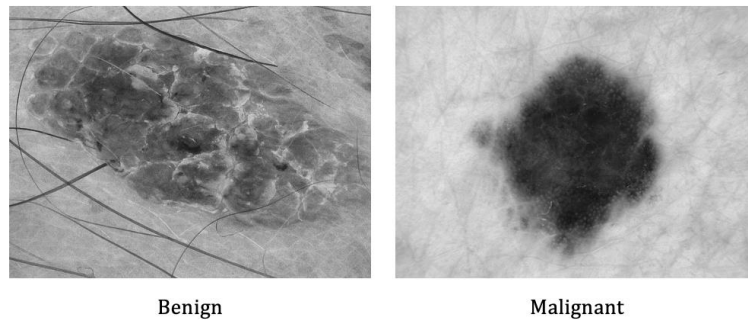
görüntü eşikleme örneğidir. Gri seviyeli bir görüntünün ikili bir görüntüye indirgenmesini otomatik olarak gerçekleştirmek için kullanılır. Algoritma, görüntünün iki modlu histogramı izleyen iki piksel sınıfı içerdiğini varsayar, daha sonra iki sınıfı ayıran optimum eşik hesaplar, böylece sınıf içi varyans minimum veya eşdeğer olurken, sınıflar arası varyanslar maksimum olur [15]. Görüntüleri yeniden boyutlandırma ve kırpma, iki temel geometri dönüşümüdür. Kırpma, görüntülerden istenmeyen dış alanların çıkarılmasıdır. İşlem genellikle, görüntülerden istenmeyen alanların çıkarılması için kullanılır. Yeniden boyutlandırma, görüntüleri küçültmeye veya büyütme olarak tanır. Genellikle dosya boyutunu ve görüntü kalitesini değiştirmek için kullanılır. Dolgu ise şekillerdeki boşlukların doldurulması için kullanılmaktadır.

Görüntünün piksel değerlerinin belli bir bölgede toplandığı ve kalitesinin düştüğü durumlarda histogram eşitleme yöntemi kullanılır. Geleneksel histogram eşitleme yöntemleri uygulandığında, görüntünün farklı bölgelerinde çok karanlık veya parlak bölgeler oluşabilir. Bu sorunu çözebilmek adına Kontrast Sınırlı Adaptif Histogram Eşitleme (CLAHE) kullanılır. CLAHE yönteminin geleneksel eşitlemeye göre farkı, görüntüyü parçalara ayırarak her bir kare içerisinde histogram eşitlemesi yapmasıdır, ayrıca kontrast limiti belirleme imkanı da tanır [16]. Bir görüntüdeki renk değerlerinin dağılımı, aydınlatma koşullarına, kameralara ve diğer faktörlere bağlı olarak değişebilir. Renk normalizasyonu, bu varyasyonları telafi etmek için piksel yoğunluğu değerlerinin aralığını değiştiren bir tekniktir. Morfolojik filtreleme iki temel işlemden türetilmiştir. Bunlar aşındırma ve genişletme işlemleridir. Aşındırma ikili bir görüntüde bulunan şeklin boyutunu, seçilen yapısal elemente bağlı olarak küçültürken, genişletme nesnenin alanını artırır. Bu işlemlerden aşındırma işlemi birbirine gürültü ile bağlanmış şekilleri birbirinden ayırmak için kullanılırken, genişletme işlemi ise şekillerin bir gürültü bölünerek ayrı iki nesne gibi görünmesini engellemek için kullanılır [17].

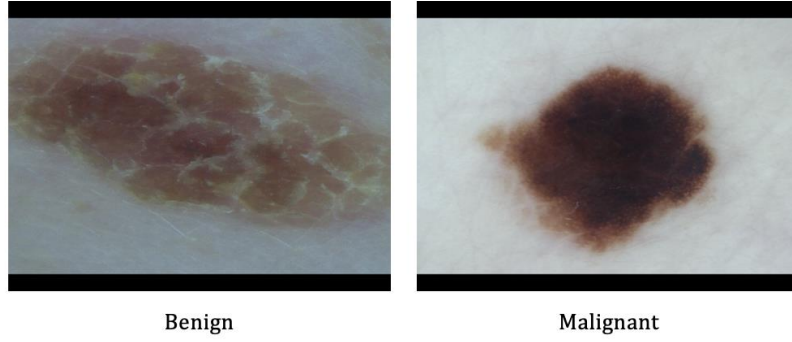
Şekil 1, Şekil 2, Şekil 3 sırasıyla; veri setinin ön işleme tabi tutulmamış haline, veri setinin birinci ön işleme dizisine tabi tutulmuş haline, ve son olarak veri setinin ikinci ön işleme dizisine tabi tutulmuş haline örnekler vermektedir. Şekil 3’de kılların görüntüden silindiği ve ayrıca yüzeydeki cilt dokusunun daha belirginleştiği görülmektedir. Makalenin Ek Kısımında, ön işleme yöntemindeki her bir adımın sonucunda oluşan resimlere örnekler sunulmuştur.



**Şekil 1.** Veri setinin ön işlem yöntem dizileri uygulanmamış hali, örnek resimler



**Şekil 2.** Veri setinin, birinci ön işlem yöntem dizisi uygulanmış hali, örnek resimler



Şekil 3. Veri setinin, ikinci ön işlem yöntem dizisi uygulanmış hali, örnek resimler

### 2.3. Sınıflandırma Yöntemi

Sınıflandırma için Sagar ve diğerleri [18] tarafından önerilen transfer öğrenme konsepti, çalışmamıza uygun olan parametre değişiklikleri ile kullanılmıştır. Transfer öğrenme ile, öğrenme sürecini sıfırdan başlatmak yerine, model, farklı bir problem çözülürken öğrenilen örüntülerden başlar. Görüntü sınıflandırmasında, transfer öğrenmesi genellikle önceden eğitilmiş modellerin kullanımı ile ifade edilir. Önceden eğitilmiş bir model, bizim çözmek istediğimize benzer bir sorunu çözmek için büyük bir kıyaslama veri kümesi üzerinde eğitilmiş bir modeldir [19]. Bu çalışmada karşılaştırmalarda genel geçer yöntem olarak kullanılan önceden eğitilmiş ResNet50 mimarisi tercih edilmiştir. ResNet50 mimarisi, tüm katmanlarda ince ayar yapılarak, ve üst katmanları bir ortalama havuzlama, bir tam bağlantılı ve son olarak 2 tanı kategorisini sınıflandırmaya izin veren softmax katmanı ile değiştirilerek veri kümesi üzerinde yeniden eğitilmiştir. Öğrenme oranı 0.0001 olarak ayarlanmış ve optimize edici olarak Adam kullanılmıştır.

Mimari, girdileri eşlemek için kimlik eşlemesini kullanır. Bu kimlik eşleme, herhangi bir parametreye sahip değildir ve yalnızca önceki katmandan gelen çıktıyı sonraki katmana ekler. Kimlik eşleme aşamasında, kısayol kanallarını kalanlarla eşleştirecek şekilde genişletmek için doğrusal bir izdüşüm ile çarpılır. Atlama bağlantıları (skip connections) önceki katmanlardan gelen çıktıları yığılmış katmanların çıktılarına ekler. Bu da, daha önce mümkün olandan çok daha derin ağları eğitme yeteneği ile sonuçlanır. Çalışmamızda fazla uyumu azaltmak için, arada yüzde 50 atlama (dropout) ve parça (batch) normalleştirme katmanları kullanılmıştır. Diğer mimari çalıştırma değerleri sırasıyla şu şekilde özetlenebilir. Sırasıyla 64 nöron ve 2 nöron içeren iki yoğun katman kullanılmıştır. Son katman, aktivasyon fonksiyonu olarak softmax ile sınıflandırma için kullanılmıştır. Kayıp fonksiyonu olarak ikili çapraz entropi kullanılmıştır. Model epok sayısı 10, parça boyutu (batch size) 32 olacak şekilde eğitilmiştir.

### 2.4. Performans Ölçütleri

Sonuçların değerlendirilmesi için doğruluk, kesinlik, geri çağırma, F1 puanı ve ROC-AUC metrikleri kullanılmıştır. Sınıflandırma performansındaki anahtar kavramlardan biri, kesin referans etiketlerine karşı model tahminlerinin tablo şeklinde görselleştirilmesi olan karışıklık matrisidir. Tablo 2’de gösterildiği üzere bu yöntemde, örnek pozitif (P) ve elde edilen sınıflandırma sonucu pozitif ise Gerçek Pozitif (TP); örnek pozitif ancak negatif (N) sınıflandırılmışsa Yanlış Negatif (FN); örnek negatif ve negatif sınıflandırılmışsa Gerçek Negatif (TN); örnek negatif ancak pozitif sınıflandırılmışsa da Yanlış Pozitif (FP) olarak gruplandırılır [20]-[21].

Tablo 2. Karışıklık Matrisi

Tahmin Edilen Değeri			
Pozitif	Negatif		
Gerçek Pozitif	Yanlış Negatif	<b>Pozitif</b>	<b>Gerçek Değeri</b>
Yanlış Pozitif	Gerçek Negatif	<b>Negatif</b>	

Doğruluk değeri, Denklem 1’de olduğu gibi hem olumlu hem de olumsuz kaç gözlemin doğru sınıflandırıldığını ölçerek modelin genel performansını gösterir. Kesinlik, Denklem 2’de olduğu gibi doğru tahminlerin ne kadar kesin olduğunu gösterir. Geri çağırma, Denklem 3’de olduğu gibi gerçek pozitif örneklerin oranını gösterir. F1 puanı, Denklem 4’de olduğu gibi dengesiz sınıflar için kullanılan hibrit bir metriktir. ROC (İşlem Karakteristik Eğrisi), çeşitli eşik değerleri için yanlış pozitif oranına (FPR) karşı gerçek pozitif oranı (TPR) gösteren eğridir.

Eğrinin ne kadar iyi olduğunu söyleyen bir sayı elde etmek için ise ROC Eğrisinin Altındaki Alan (AUC) veya ROC AUC puanı hesaplanır. Benzer şekilde, (AUPRC: The Area Under Precision-Recall Curve) metriği kesinlik-geri çağırma eğrisinin altındaki alanı ifade eder.

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

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

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

$$\text{F1 Puanı} = \frac{(2TP)}{(2TP + FP + FN)} \quad (4)$$

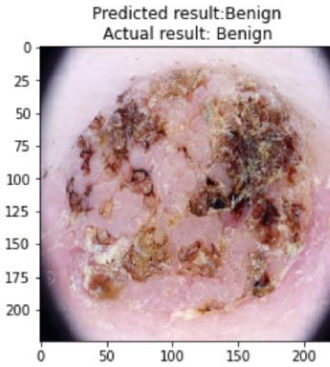
### 3. Testler ve Performans Analizi

Çalışmamızda dermoskopik görüntülerin sınıflandırılması amacıyla oluşturulan Resnet50 mimarisine dayanan ESA modeli; iyi huylu ve kötü huylu olmak üzere iki sınıftan oluşan ve toplamda 2841 dermoskopik görüntü içeren ISIC 2018 Cilt Görüntü Analizi Yarışması'na ait veri seti üzerinde test edilmiştir. Model epok sayısı 10, parça boyutu (batch size) 32 olacak şekilde eğitilmiştir. Çalışmanın temel amacı görüntüleri farklı ön işlemlere tabi tutup kullanılan mimari üzerindeki etkisini incelemektir. Bu sebeple iki farklı ön işlem yöntem dizisi denenmiştir. Bunlardan birincisinde sırasıyla; ikili ve otsu eşikleme, belirli oranda kırpma and CLAHE dönüşümü yapılmıştır. İkincisinde ise sırasıyla; yeniden boyutlandırma, görüntülerdeki kılları yok etmek için morfolojik filtreleme, renk normalizasyonu, dolgu ve kırpma işlemleri uygulanmıştır.

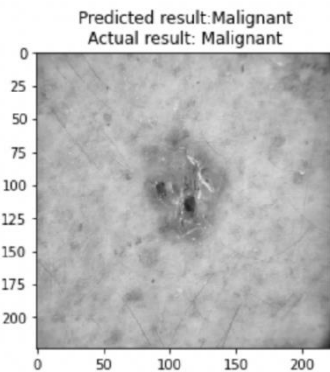
Bu amaç doğrultusunda geliştirilen model üç senaryo ile test edilmiştir. Sonuçların değerlendirilmesi için yapılan testler sırasıyla şu şekilde özetlenebilir:

- A= Ön işleme yöntemi kullanılmadan,
- B= Birinci ön işleme yöntem dizisi kullanılarak,
- C= İkinci ön işleme yöntem dizisi kullanılarak,

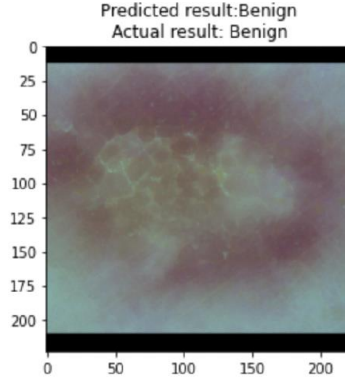
yapılan testleri temsil etmektedir. Şekil 4, Şekil 5, ve Şekil 6'da sırasıyla; A, B ve C test senaryolarının doğru sınıflandırma sonuçlarına birer örnek sunulmuştur.



Şekil 4. "A" test senaryosunun doğru sınıflandırma sonucuna bir örnek



Şekil 5. "B" test senaryosunun doğru sınıflandırma sonucuna bir örnek



Şekil 6. "C" test senaryosunun doğru sınıflandırma sonucuna bir örnek

Tüm test sonuçlarının değerlendirilmesi için Tablo 3'te gösterildiği üzere; doğruluk, kesinlik, geri çağırma, F1 puanı ve ROC-AUC metrikleri kullanılmıştır. Tablo 3'te görüldüğü üzere her bir senaryo, model üzerinde test edilmiş olup, belirtilen metriklere göre sonuçlar sunulmuştur. Sonuçlar incelendiğinde, aşağıdaki çıkarımlar elde edilmiştir:

- i) Doğruluk ve Geri Çağırma metrikleri açısından değerlendirildiğinde, B senaryosunda diğerlerinden yüksek değerli sonuç elde edilmiştir.
- ii) Kesinlik metriği açısından, A senaryosunda diğerlerinden yüksek değerli sonuç elde edilmiştir.
- iii) F1 Puanı açısından, C senaryosunda diğerlerinden yüksek değerli sonuç elde edilmiştir.
- iv) ROC-AUC metriği açısından, A ve C senaryolarında, B senaryosundan daha yüksek değerli sonuç elde edilmiştir.

Tablo 3. Senaryoların performans değerlendirmesi

Test Senaryosu	Doğruluk	Kesinlik	Geri Çağırma	F1 Puanı	ROC-AUC	AUPR
A	0.804	<b>0.83</b>	0.86	0.84	<b>0.80</b>	<b>0.976</b>
B	<b>0.824</b>	0.76	<b>0.93</b>	0.83	0.76	0.983
C	0.819	0.80	0.92	<b>0.85</b>	<b>0.80</b>	<b>0.984</b>

Testler sonucunda, ön işleme yöntemlerinin model ile beraber kullanıldığında ve Doğruluk, Geri Çağırma, F1 puanı metriklerine göre değerlendirildiğinde performansı arttırdığı; ROC-AUC değerine göre değerlendirildiğinde performans farkı oluşturmadığı; Kesinlik değeri için ise anlamlı bir performans artışı elde edilemediği gözlemlenmiştir.

Doğruluk, en sezgisel performans ölçüsüdür ve doğru tahmin edilen gözlemin toplam gözlemlere oranıdır. Doğruluk değeri yüksekse, modelin en iyisi olduğu düşünülebilir. Fakat bu durum yalnızca yanlış pozitif ve yanlış negatif değerlerinin neredeyse aynı olduğu simetrik veri kümelerinde geçerli olur. Bu nedenle, modelin performansını değerlendirmek için diğer performans metriklerine de bakmak gerekmektedir [22]. Kesinlik, doğru tahmin edilen pozitif gözlemlerin toplam tahmin edilen pozitif gözlemlere oranıdır. Bu metrik, kötü huylu olarak etiketlenen tüm resimlerin gerçekte kaç tanesinin kötü huylu olduğu sorusunun cevabıdır. Kesinlik sonuçları incelendiğinde, ön işleme yöntemlerinin kullanılmadığı durumda bu değer daha yüksek çıkmıştır. Geri Çağırma, doğru tahmin edilen pozitif gözlemlerin gerçek sınıftaki tüm gözlemlere oranıdır. Kanseri tespiti probleminde yanlış negatiflik yanlış pozitifliğe göre daha önemli bir hata olarak değerlendirilmektedir. Bunun nedeni, bir kanser çeşidinin tespitinin geciktiği veya yapılamadığı zaman verdiği zararın, yanlış kanser tespitinde verdiği zarardan daha fazla olmasıdır. Bu metrik yanlış negatifliği de ölçüme katarak, tüm kötü huylu lezyon resimlerinden kaç tanesi etiketlenmiştir sorusunu yanıtlar. Bu sebeple önemli bir metrik olarak değerlendirilir ve mümkün olduğunca yüksek olması beklenir. F1 Skoru, Kesinlik ve Geri Çağırmanın ağırlıklı ortalamasıdır. Bu nedenle, bu

puan hem yanlış pozitifleri, hem de yanlış negatifleri hesaba katar. Sezgisel olarak, doğruluk kadar anlaşılması kolay değildir, ancak F1 genellikle, özellikle eşit olmayan bir sınıf dağılımı söz konusuysa, doğruluktan daha kullanışlıdır. Doğruluk, yanlış pozitifler ve yanlış negatifler benzer maliyete sahipse en iyi sonucu verir. Yanlış pozitiflerin ve yanlış negatiflerin maliyeti çok farklıysa, hem Kesinlik hem de Geri Çağırma metriklerini değerlendirmek daha iyidir. Sonuçlara bakıldığında, F1 puanı, ön işleme yöntemlerinin uygulandığı durumlarda daha yüksektir. ROC bir olasılık eğrisidir ve altında kalan alan olan AUC ayrılabilirliğin derecesini veya ölçüsünü temsil eder. Eğrinin altında kalan alan arttıkça, sınıflar arasında ayırt etme performansı artmaktadır.

Başlıca F1 Puanı, ROC-AUC, AUPRC metrikleri baz alındığında, görüntüler üzerinde C) morfolojik filtreleme, renk normalizasyonu ve dolgu işlemlerinin, B) ikili ve otsu eşikleme, CLAHE dönüşümünden ve A) ön işleme yöntem dizisi uygulanmamış halinden daha fazla iyileştirme yaptığı sonucuna varılmaktadır. Bu bulgu sonucunda, görüntülerin başlıca gürültü azaltılmasına, belirli oranlarda istenen nesnenin ortaya çıkarılmasına ve kontrast iyileştirilmesine ihtiyacının olduğu anlaşılmaktadır.

#### 4. Tartışma ve Sonuç

Bu çalışmada, iyi huylu ve kötü huylu cilt kanseri sınıflandırmasında, derin evrimsel sinir ağlarının ve farklı ön işleme yöntemlerinin performansa etkileri araştırılmıştır. Ağı hem eğitmek, hem de test etmek için halka açık ISIC 2018 Cilt Görüntü Analizi Yarışması'na ait veri seti kullanılmıştır. Bu çalışmanın temel amacı, farklı öğrenme mimarilerini karşılaştırmak yerine, görüntüleri farklı ön işlemlere tabi tutup, bu işlemin kullanılan mimari performansına etkisini incelemektir. Bu amaç doğrultusunda iki farklı ön işleme yöntem dizisi ResNet50 mimarisi için uygulanmıştır. Bunlardan birincisinde sırasıyla ikili ve otsu eşikleme, CLAHE dönüşümü uygulanırken, ikincisinde morfolojik filtreleme, renk normalizasyonu ve dolgu işlemleri uygulanmıştır. Sonuçların değerlendirilmesi için doğruluk, kesinlik, geri çağırma, F1 puanı ve ROC-AUC metrikleri kullanılmıştır.

Başlıca F1 Puanı ve ROC-AUC metrikleri baz alındığında, görüntüler üzerinde yapılan morfolojik filtreleme, renk normalizasyonu ve dolgu işlemler dizisinin, diğer ön işleme yöntem dizisine göre daha fazla iyileştirme yaptığı sonucuna varılmıştır. Sonuçlarımız, dermoskopi görüntüleri üzerinde eğitilmiş derin öğrenme mimarilerinin farklı ön işleme yöntemleri ile birlikte kullanıldığında, daha iyi bir sınıflandırma yapabileceğini göstermiştir. Bu kombine model daha da geliştirilerek dermatologlara yardımcı olmak için kullanılan çeşitli dermoskopi sistemlerinde kolayca uygulanabilir. Daha fazla iyileştirme oranı için çok daha fazla dermoskopi görüntüsü gerekmektedir. Ayrıca görüntüler üzerindeki farklı gereksinimler tespit edilebilirse ve bunlar doğrultusunda farklı ön işleme yöntem dizileri denenirse, performans üzerindeki etkinin daha da artacağını düşünmekteyiz.

#### Kaynakça

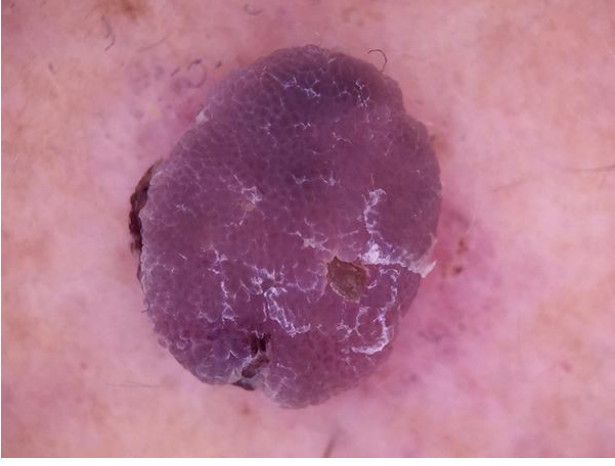

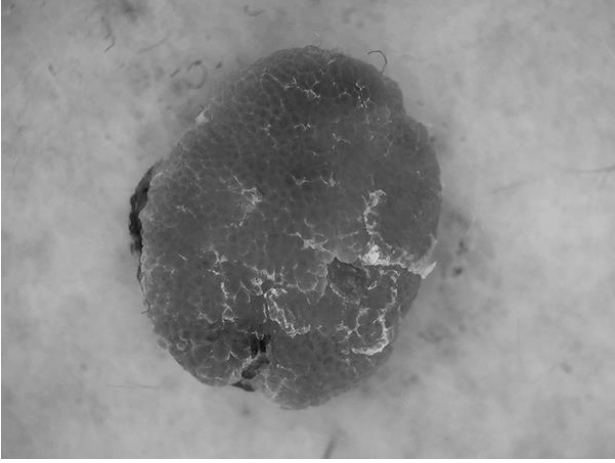
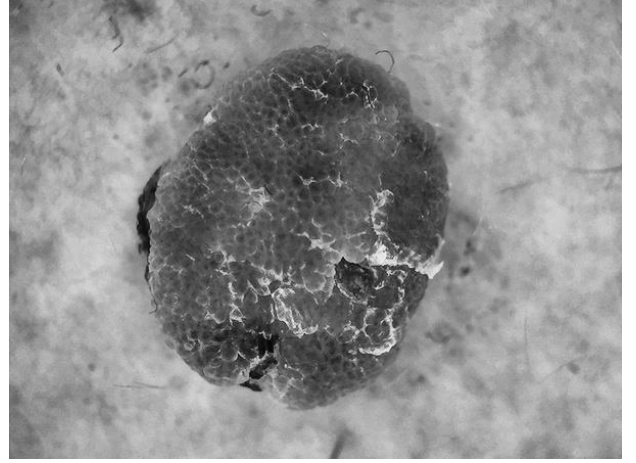
- [1] Chante Karimkhani, Adele C Green, Tamar Nijsten, MA Weinstock, ` Robert P Dellavalle, Morteza Naghavi, and C Fitzmaurice. The global burden of melanoma: results from the global burden of disease study 2015. *British Journal of Dermatology*, 177(1):134–140, 2017.
- [2] Andre Esteva, Brett Kuprel, Roberto A Novoa, Justin Ko, Susan M Swetter, Helen M Blau, and Sebastian Thrun. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639):115, 2017.
- [3] D. Gutman et al., "Skin lesion analysis toward melanoma detection: A challenge at the international symposium on biomedical imaging (ISBI) 2016, hosted by the international skin imaging collaboration (ISIC)," 2016.
- [4] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun. Dermatologistlevel classification of skin cancer with deep neural networks. *nature*, 542(7639):115–118, 2017.
- [5] L. Xu, J. S. Ren, C. Liu, and J. Jia. Deep convolutional neural network for image deconvolution. In *Advances in neural information processing systems*, pages 1790–1798, 2014.
- [6] Tran, K. A., Kondrashova, O., Bradley, A., Williams, E. D., Pearson, J. V., & Waddell, N. (2021). Deep learning in cancer diagnosis, prognosis and treatment selection. *Genome Medicine*, 13(1), 1-17.
- [7] A. R. Lopez, X. Giro-i Nieto, J. Burdick, and O. Marques. Skin lesion classification from dermoscopic images using deep learning techniques. In *2017 13th IASTED international conference on biomedical engineering (BioMed)*, pages 49–54. IEEE, 2017.



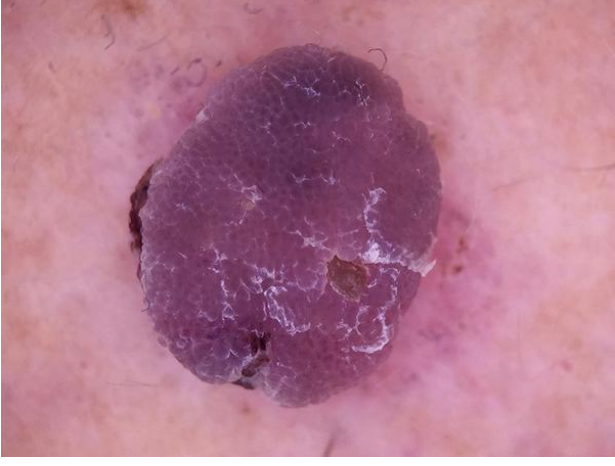
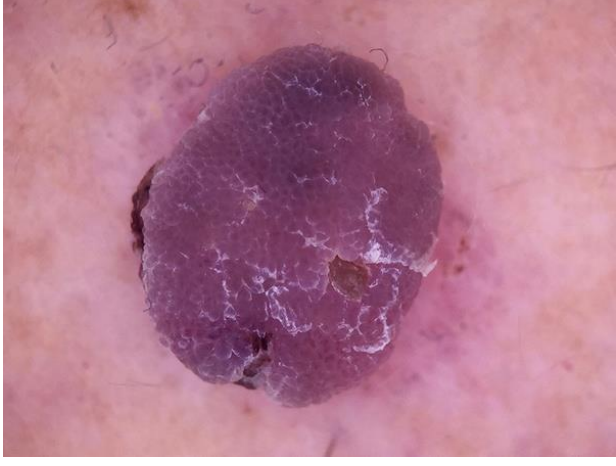
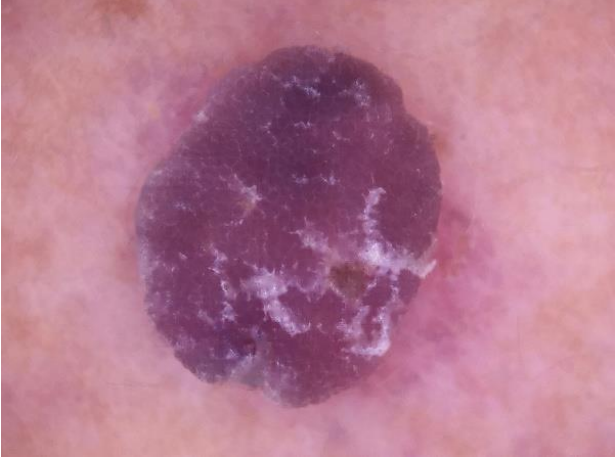
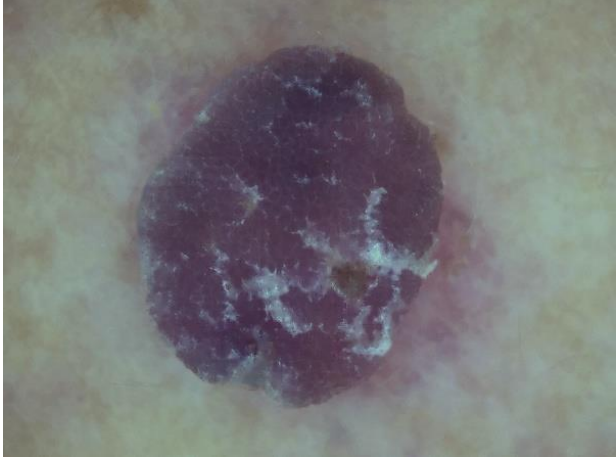
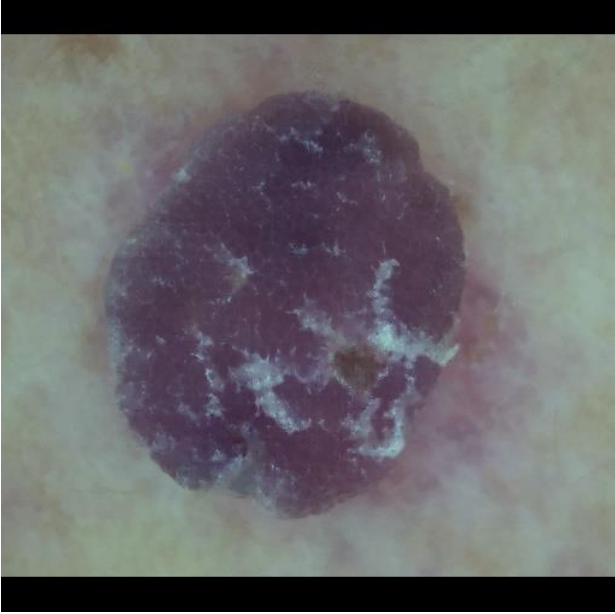
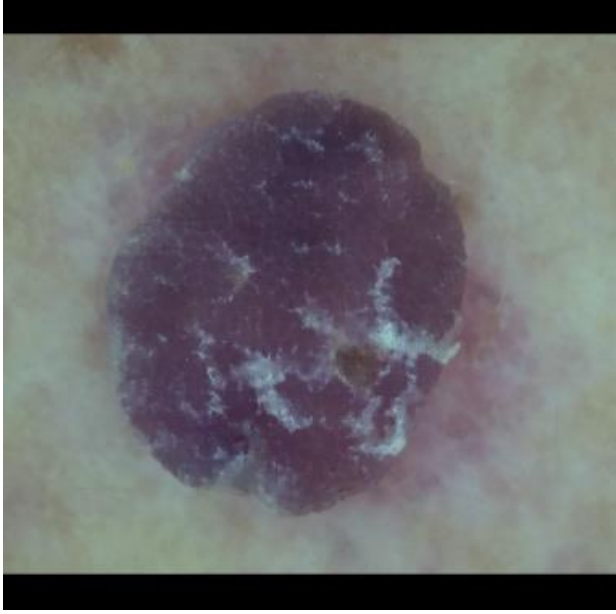
- [8] Dorj, U. O., Lee, K. K., Choi, J. Y., & Lee, M. (2018). The skin cancer classification using deep convolutional neural network. *Multimedia Tools and Applications*, 77(8), 9909-9924.
- [9] T. J. Brinker, A. Hekler, J. S. Utikal, N. Grabe, D. Schadendorf, J. Klode, C. Berking, T. Steeb, A. H. Enk, and C. von Kalle. Skin cancer classification using convolutional neural networks: systematic review. *Journal of medical Internet research*, 20(10):e11936, 2018.
- [10] A. Rezvantalab, H. Safigholi, and S. Karimijeshni. Dermatologist level dermoscopy skin cancer classification using different deep learning convolutional neural networks algorithms. *arXiv preprint arXiv:1810.10348*, 2018.
- [11] N. C. Codella, Q.-B. Nguyen, S. Pankanti, D. A. Gutman, B. Helba, A. C. Halpern, and J. R. Smith. Deep learning ensembles for melanoma recognition in dermoscopy images. *IBM Journal of Research and Development*, 61(4/5):5-1, 2017.
- [12] Hussain, M., Bird, J. J., & Faria, D. R. (2018, September). A study on cnn transfer learning for image classification. In *UK Workshop on computational Intelligence* (pp. 191-202). Springer, Cham.
- [13] Purnama IKE, Hernanda AK, Ratna AAP, Nurtanio I, Hidayati AN, Purnomo MH, Rachmadi RF. 2019. Disease classification based on dermoscopic skin images using convolutional neural network in teledermatology system. In: *Proceedings of International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM)*, November 19 - 20, 2019, Surabaya, Cava, pp. 1-5.
- [14] R. J. Hemalatha, B. Babu, A. J. A. Dhivya, T. R. Thamizhvan, J. E. Joseph and R. Chandrasekaran, "A comparison of filtering and enhancement methods in malignant melanoma images," 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI), 2017, pp. 2704-2710, doi: 10.1109/ICPCSI.2017.8392209.
- [15] Guruprasad, P. (2020, June). Overview of different thresholding methods in image processing. In *TEQIP Sponsored 3rd National Conference on ETACC*.
- [16] Sahu, S., Singh, A. K., Ghreera, S. P., & Elhoseny, M. (2019). An approach for de-noising and contrast enhancement of retinal fundus image using CLAHE. *Optics & Laser Technology*, 110, 87-98.
- [17] Ezhilarasan, K., Praveenkumar, S., Somasundaram, K., Kalaiselvi, T., Magesh, S., Kiruthika, S., & Jeevarekha, A. (2021). Automatic brain extraction from MRI of human head scans using Helmholtz free energy principle and morphological operations. *Biomedical Signal Processing and Control*, 64, 102270.
- [18] Sagar, A., & Dheeba, J. (2020). Convolutional neural networks for classifying melanoma images. *bioRxiv*.
- [19] Agarwal, N., Sondhi, A., Chopra, K., & Singh, G. (2021). Transfer learning: Survey and classification. In *Smart Innovations in Communication and Computational Sciences* (pp. 145-155). Springer, Singapore.
- [20] "cs-229, cheatsheet machine learning tips and tricks", Stanford University. <https://stanford.edu/~shervine/teaching/cs-229/cheatsheet-machine-learning-tips-and-tricks> .
- [21] "Understanding Confusion Matrix". Sarang Narkhede. May 9, 2018. <https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62> .
- [22] Hossin, M., & Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. *International journal of data mining & knowledge management process*, 5(2), 1.

## Appendix

### Birinci Ön İşleme Yönteminin Çıktı Örnekleri

	
1) Orjinal Görüntü	2) İkili ve OTSU Eşikleme Uygulanmış Hali
	
3) Boundary of ROI ile kırılmış hali	4) CLAHE uygulanmış hali

## İkinci Ön İşleme Yönteminin Çıktı Örnekleri

	
<b>1. Orjinal Görüntü</b>	<b>2. Boyut küçültme (% 55)</b>
	
<b>3. Renk Normalizasyonu</b>	<b>4. Morfolojik Uygulama sonrası</b>
	
<b>5. Dolgu ve Kırpma</b>	<b>6. Boyutu 380x380 yapma</b>