

Dermoskopi Görüntülerinin İleri Beslemeli Yapay Sinir Ağı, Karar Ağaçları ve Rastgele Orman ile Sınıflandırılması

Esra Öztürk^{1*}, Semra İçer²

¹Biyomedikal Mühendisliği/Fen Bilimleri Enstitüsü, Erciyes Üniversitesi, Kayseri, Türkiye

²Biyomedikal Mühendisliği, Erciyes Üniversitesi, Kayseri, Türkiye

*esrgly.edu@gmail.com, ksemra@erciyes.edu.tr

Özet – Günümüzde kanser hastalıkları hızla artmaktadır. Cilt kanseri toplumlarda diğer kanser türlerine göre daha az görülmesine rağmen geç teşhisinde öldürücülüğü yüksek bir kanser türüdür. Cilt kanserini güneşten gelen zararlı ışınların tetikleyebileceği gibi genetik faktörler de cilt kanserinin oluşmasında büyük etkindir. Cilt kanserinde erken teşhiste ölüm oranı düşüken geç teşhisinde ise hayatta kalma oranı düşüktür. Yapay sinir ağları kullanılarak dermoskopi görüntülerinden malign (kötü huylu) ve benign (iyi huylu) lezyonların sınıflandırılması erken teşhisi kolaylaştırılacağı düşünülmektedir. Bu çalışmada veriler hazır veri seti kullanılarak Uluslararası Cilt Görüntüleme İşbirliği (ISIC) veri setinden alınmıştır. Alınan dermoskopi görüntülerine ön işlem uygulanmış ardından MATLAB’ de entropi, standart sapma, alan, homojenite, kontrast, korelasyon, enerji, çarpıklık ve basıklık özellikleri çıkarılmıştır. Bu özellikler ile birlikte demografik bilgiler olan yaş ve cinsiyet bilgisi de sınıflandırma için kullanılacak özelliklere eklenmiştir. Bu özellikler MATLAB ve WEKA programları kullanılarak sınıflandırma işlemleri gerçekleştirilmektedir. MATLAB’ de ileri beslemeli sinir ağı ve karar ağaçları algoritması ile sınıflandırılmıştır, rastgele orman algoritması için WEKA programı kullanılarak sınıflandırılmıştır. Bu üç yöntem ile ağlar eğitilerek sonuçlar elde edilmiştir.

Anahtar Kelimeler – Cilt kanseri (melanoma), melanostik nevüs, Yapay Sinir Ağı, sınıflandırma, karar ağaçları, rastgele orman

I. GİRİŞ

Deriye rengini veren melanin adlı renk pigmentleri, melanosit denilen cilt hücreleri tarafından üretilir. Bu hücrelerin kontrolsüz bir şekilde bölünüp çoğalması sonucunda meydana gelen cilt kanseri melanom veya melanoma olarak tanımlanır [1].

Melanositik nevus (mole), pigment üreten melanosit isimli hücrelerden oluşan benign poliferasyon/hamartom’dur. Melanositik nevuslar genellikle konjenital ve edinsel olarak iki gruba ayrılır. Konjenital nevuslar doğumda veya yaşamın ilk bir yılı içinde görülür, boyutları birkaç milimetreden vücut yüzeyinin büyük kısmını kaplayabilecek büyüklükte olabilir. Benign melanositik nevuslar koyu tenli bireylerde açık ve sarı tenli bireylere göre daha az sayıda görülür. Genelde benign nevuslar; küçük (6 milimetreden az), simetrik ve iyi konumlanmıştır [2].

Melanomun görülme sıklığı %4 olmakla birlikte cilt kanserinden dolayı olan ölümlerin %75’ine neden olmaktadır. Melanom insidansı son 30 yılda hızlı bir artış göstermiştir. Avrupa’da her yıl 62,000 yeni vaka saptanmaktadır. Amerikan Kanser Topluluğunun 2016 raporuna göre Birleşik Devletler’de, 2016 yılında 76,380 vaka teşhis edileceği ve 10,130 kişinin melanomadan öleceği tahmin edilmektedir. Ancak melanom erken teşhis edildiğinde sağ kalım oranı yaklaşık % 90 olan bir hastalıktır [3,4].

Codella ve arkadaşları, derin öğrenme, seyrek kodlama ve destek vektör makinesi (SVM) öğrenme algoritmalarını birleştiren dermoskopi görüntülerinde melanom tanıma için bir yaklaşım sunmaktadır. Metodolojiyi değerlendirmek için performans, 2624 klinik melanom vakası (334), atipik nevüs (144) ve iyi huylu lezyonlar (2146) içeren International Skin

Imaging Collaboration’dan elde edilen bir veri setinde ölçülür. Yaklaşım, bu veri kümesindeki önceki son teknoloji yöntemle karşılaştırılır. Değerlendirme için 20 kez iki kat çapraz doğrulama yapılır (toplam 40 deney) ve iki ayırım görevi incelenir: melanom ve melanom dışı tüm lezyonlar, melanom ve sadece atipik lezyonlar. Sunulan yaklaşım, ilk görev için %93,1’lik bir doğruluk ve ikinci görev için %73.9 doğruluk elde etmektedir. Son teknoloji ile karşılaştırıldığında anlamlı farklılıklar elde edilmiştir [5].

Milton, melanom ve cilt lezyonu kanserlerini tespit etmek için farklı derin öğrenme tabanlı yöntemler üzerinde çalışmıştır. Çalışmada, PNASNet-5-Large, InceptionResNetV2, SENet154, InceptionV4 gibi yeni derin öğrenme tabanlı modelleri kullanan çeşitli sinir ağlarını kullanılmıştır. Dermoskopik görüntüler, ağa beslenmeden önce uygun şekilde işlenir ve büyütülür. Yöntemler International Skin Imaging Collaboration (ISIC) 2018 veri seti üzerinde test edilmiştir. Sistem, PNASNet-5-Large modeli için 0,76 ile en iyi doğrulama puanını elde etmiştir [6].

Dubal ve arkadaşları, genel kameralardan alınan görüntülere dayanarak cilt lezyonlarını iyi huylu veya kötü huylu olarak tespit eden ve tanımlayan bir metodoloji önermektedir. Görüntüler bölümlere ayrılır, öznitelikler ABCD kuralı uygulanarak çıkarılır ve lezyonları yüksek doğruluk derecesinde sınıflandırmak için bir Sinir Ağı eğitilir. Eğitimli Sinir Ağı, altı farklı sınıfa bölünmüş 463 görüntüden oluşan bir veri setinde %76.9’luk bir genel sınıflandırma doğruluğu elde etmiştir [7].

Daskalu ve arkadaşları, polarize ışıklı (SMP) ilkel bir cilt büyüteci kullanarak sınıflandırma ile görüntü kalitesinin tanı

doğruluğu üzerindeki etkisini belirlemek istemişlerdir. Dermoskopik görüntülerin DL işlenmesi ve ardından sınıflandırma, SMP için doğru bir tanı çıktısı ile sonuçlanır, bu da dermoskopun kalitesinin cilt kanserinin DL tanısını etkileyen ana faktör olmadığı anlamına gelir. Mevcut sistem, uygulanabilir bir bilgisayar destekli tespit sistemi olarak tüm sağlık hizmeti sağlayıcılarına yardımcı olabilir [8].

Estava ve arkadaşları, girdi olarak yalnızca pikselleri ve hastalık etiketlerini kullanarak, doğrudan görüntülerden uçtan uca eğitilmiş tek bir CNN kullanarak cilt lezyonlarının sınıflandırılmasını göstermektedir. Bir CNN'yi 2.032 farklı hastalıktan oluşan 129.450 klinik görüntüden (önceki veri kümelerinden 12 iki kat daha büyük) oluşan bir veri kümesi kullanarak eğitilir. CNN, dermatologlarla karşılaştırılabilir bir yeterlilik düzeyi ile cilt kanserini sınıflandırabilen bir yapay zeka göstererek, her iki görevde de test edilen tüm uzmanlarla eşit performans elde eder [9].

Dorj ve arkadaşları, ECOC SVM ve derin evrişimli sinir ağı kullanarak cilt kanserini sınıflandırmışlardır. Deri kanserlerinin RGB görüntüleri internetten toplanmıştır. Özniteliklerin çıkarılmasında mevcut ve önceden eğitilmiş bir AlexNet evrişimsel sinir ağı modeli kullanılmıştır. Sonuçlar, dört çeşit cilt kanseri görüntüsü içeren toplam 3753 görüntü ile önerilen bir algoritma yürütülerek elde edilir. Uygulama sonucu, ortalama doğruluk, duyarlılık ve özgüllüğün maksimum değerlerinin sırasıyla 95.1 skuamöz hücreli karsinom), 98.9 (aktinik keratoz), 94.17 (skuamöz hücreli karsinom) olduğunu göstermektedir. Bu ölçümlerde ortalamanın minimum değerleri sırasıyla 91.8 (bazal hücreli karsinom), 96.9 (Skuamöz hücreli karsinom) ve 90.74'tür (melanoma) [10].

Bu çalışmada melanoma (malign) ve melanostik nevüs (benign) dermoskopi görüntüleri kullanılarak, benign ve malign lezyonların sınıflandırılması ileri beslemeli yapay sinir ağı, karar ağaçları ve rastgele orman algoritması kullanılarak sınıflandırılması gerçekleştirilmiştir.

II. MATERYAL VE METOTLAR

A. Veri Seti

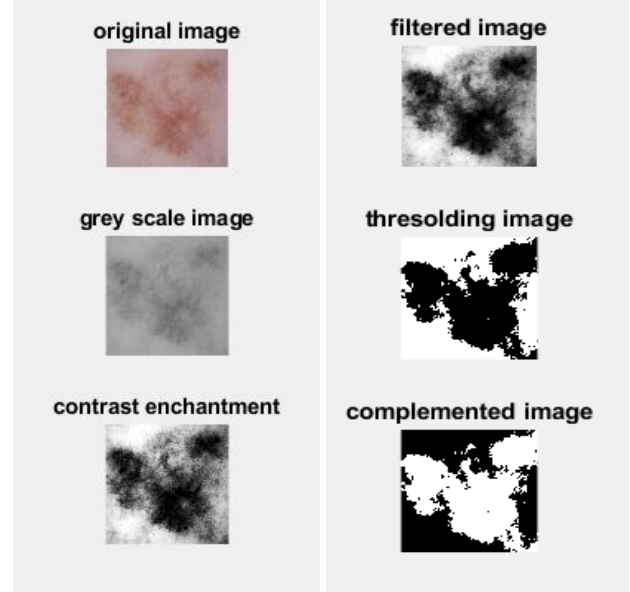
Veriler Uluslararası Cilt Görüntüleme İşbirliği (ISIC) veri setinden alınmıştır. Veri setinde farklı modaliteler tarafından elde edilen dermoskopik görüntüler mevcuttur. Veri seti, 10015 dermoskopik görüntüden oluşmaktadır. Veri seti farklı lezyonlar içeren görüntülerden oluşmaktadır. Veri setinde bulunan hastalıklar şunlardır: Aktinik keratoz ve intraepitelial karsinom / Bowen hastalığı (akiec), bazal hücreli karsinom (bcc), iyi huylu keratoz benzeri lezyonlar (bkl), dermatofibrom (df), melanom (mel), melanositik nevüs (nv) ve vasküler lezyonlar (vasc) [11].

Veri seçimi yapılırken malign olan melanoma görüntüleri ve benign olan melanostik nevüs görüntüleri seçilmiştir. Bu veriler içerisinde histolojik olarak kanıtlanmış, yüzde bulunan lezyonlar seçilmiştir. Bu görüntüler içerisinde de görüntü işleme de gürültü olmaması için kıl olmayan görüntüler seçilmiştir. Veri ayıklama işlemi yapıldıktan sonra 30 melanoma ve 30 melanostik nevüs görüntülerinden toplam 60 görüntü olacak şekilde veriler toplanmıştır.

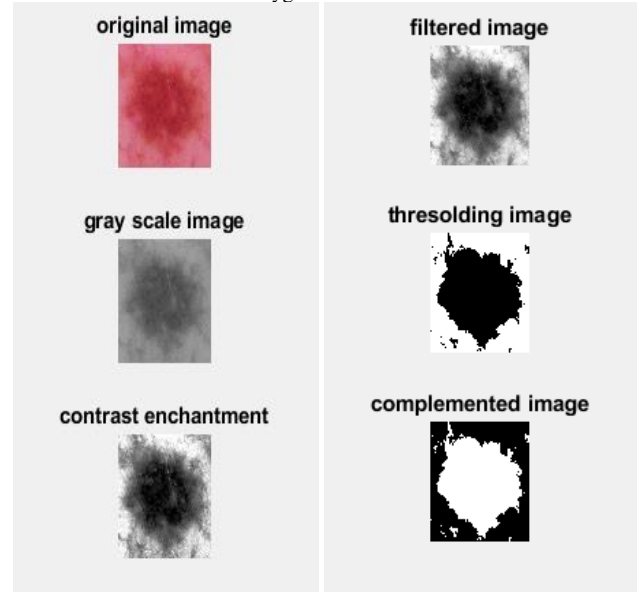
B. Ön İşlem

Dermoskopi görüntülerine ön işlem MATLAB2017b programı kullanılarak gerçekleştirilmiştir. Görüntüler

MATLAB ortamına alındıktan sonra renkli olan dermoskopi görüntüleri griye çevrilmiştir. Ardından sırasıyla kontrast artırılıp, medyan filtreleme ile filtreleme yapıp sonrasında otomatik eşik ile eşikleme yapılarak eşleniği alınmıştır. Yapılan ön işleme adımları Şekil 1 ve Şekil 2' de melanoma ve melanostik nevüs görüntüleri üzerinde gösterilmektedir.



Şekil 1. Melanoma dermoskopi görüntüsüne ön işlem adımlarının uygulanması



Şekil 2. Melanostik nevüs dermoskopi görüntüsüne ön işlem adımlarının uygulanması

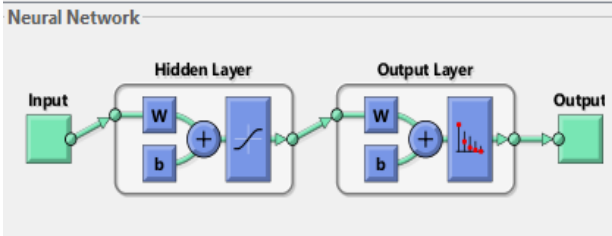
C. Özellik Çıkarımı

Verilere ön işleme yapıldıktan sonra özellik çıkarım aşamasına geçilmektedir. Özellik çıkarım işlemi MATLAB ortamında gerçekleştirilmektedir. Görüntülerde bulunan lezyonların alanları hesaplanıp, görüntülerin entropisi ve standart sapması hesaplanmıştır. Ardından görüntülerden GLCM matrisleri oluşturularak görüntülerin istatistiksel özellikleri olan, homojenite, enerji, kontrast ve korelasyon değerleri bulunmuştur. Görüntülerin çarpıklık ve basıklık değerleri hesaplanarak özelliklere eklenmiştir. Son olarak wavelet dönüşümü alınarak yaklaşıklık katsayısı özellik olarak eklenmiştir. Metadada bulunan demografik bilgiler olan yaş ve cinsiyet bilgileri de özelliklere eklenmiştir. Her

bir veri için toplamda biri sınıf değişkeni olmak üzere 13 özellik bulunmaktadır. Ek 1 TABLO 1’ de çıkarılan özellikler verilmiştir.

D. Çıkarılan Özelliklere Yapay Sinir Ağlarının Uygulanması
D.1. Neural Pattern Recognition (nprtool)

Örüntü tanıma problemlerinde, bir sinir ağının girdileri bir dizi hedef kategoriye sınıflandırmasını sağlar. Nöral Örüntü Tanıma uygulaması, verileri seçmenize, bir ağ oluşturup eğitmenize ve çapraz entropi ve karışıklık matrislerini kullanarak performansını değerlendirmenize yardımcı olur. Sigmoid gizli ve softmax çıkış nöronları (patternnet) ile iki katmanlı bir ileri beslemeli ağ, gizli katmanında yeterli nöron verildiğinde vektörleri iyi sınıflandırabilir. Ağ, ölçeklendirilmiş eşlenik gradyan geri yayılımı (trainscg) ile eğitilecektir [12]. Ağ yapısı Şekil 3’ te gösterilmektedir.



Şekil 3. Ağ yapısı [12]

İleri beslemeli yapay sinir ağı uygulaması MATLAB’ in nprtoolu kullanılarak gerçekleştirilmiştir. Veriler girilmeden önce yaş, alan, çarpıklık ve basıklık bilgileri diğer bilgilerden çok aykırı olduğu için normalizasyon işlemine tabi tutulmuştur. Ardından ileri beslemeli 15 hidden katmanı bulunan ağa 12 bağımsız değişken girdi olarak, 1 bağımsız sınıf değişkeni de çıktı olarak girilmiştir. Verilerin % 70’ i eğitime, %15’ i validasyona ve % 15’ i de test için ayrılmıştır. Bu bilgiler toola girildikten sonra ağın eğitimi gerçekleştirilmiştir.

D.2. Karar Ağaçları

Karar ağaçları (KA) kök düğüm, dallar ve yapraklardan oluşmaktadır. Karar ağaçlarındaki yapraklar sınıflandırmanın olduğu yerler olup dallar ise sonucu ifade etmektedir. Kök düğümünden yaprak düğümlere doğru tüme varım yöntemiyle ağaç oluşturulur. Karar ağacı iki temel işlem adımı içerir. Bu işlemler bölme ve budama işlemleridir. KA oluştururken en önemli adım hangi öznelik değerlerini baz alıp dallanmanın oluşturulacağına karar vermektir. Entropi kurallarını içeren bilgi kazancı ve bilgi kazancı oranı yaklaşımında eldeki tüm öznelikler özyineli olarak test edilir ve bilgi kazancı en yüksek olan öznelik dallanma için seçilir.

Formüller ile açılacak olursak:

Bir veri setinin A1, A2, ..., An şeklinde birkaç tane sınıftan oluştuğu ve n’ nin sınıf değerini gösterdiği düşünülürse, bir sınıfa ait olasılık $P_i=(A_i/T)$ olur ve sınıflara ait Entropi;

$$Entropi(T) = - \sum_{i=0}^n P_i \log_2(P_i) \quad (1)$$

Şeklinde hesaplanır. Veri setindeki B özneliğine göre T sınıf değerleri T1,T2,...Tn şeklinde alt kümelere ayrıldığı göz önünde bulundurulsun. B öznelik değeri kullanılarak T sınıf değerlerinin bölünmesi sonucunda elde edilecek kazanç,

$$Kazanç(B,T) = Entropi(T) - \sum_{i=0}^n \frac{|T_i|}{|T|} Entropi(T_i) \quad (2)$$

Şeklinde hesaplanır. T kümesi için B özneliğinin değerini belirlemede bölüme bilgisi kullanılır(Formül (3)).

$$Bilgi(B) = - \sum_{i=0}^k \frac{|T_i|}{|T|} \log_2\left(\frac{|T_i|}{|T|}\right) \quad (3)$$

Kazanç oranı,

$$Kazanç\ oranu = \frac{Kazanç(B,T)}{Bilgi(B)} \quad (4)$$

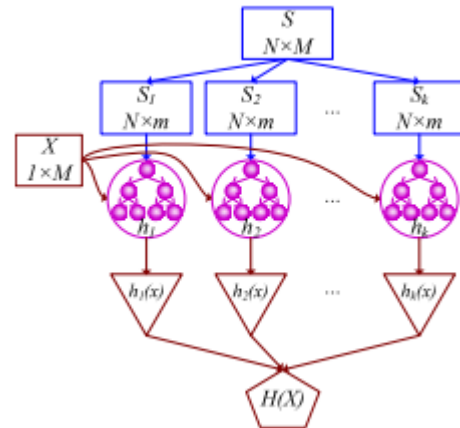
Şeklinde hesaplanır [13]. Bu formüllere göre kazancın en yüksek olduğu dal seçilmektedir.

Karar ağaçları algoritması MATLAB’ in Classification Learner toolu kullanılarak gerçekleştirilmiştir. Giriş verileri girildikten sonra farklı cross validation değerleri girilerek eğitim gerçekleştirilmiştir. En iyi sonuç 25 cross validation değerinde bulunmuştur.

D.3.Rastgele Orman

Rastgele orman algoritması, karar ağacı modeline dayalı bir topluluk sınıflandırıcı algoritmasıdır. Bir önyükleme örnekleme yaklaşımı kullanarak orijinal bir veri kümesinden k farklı eğitim verisi alt kümesi oluşturur ve ardından bu alt kümeleri eğiterek k adet karar ağacı oluşturulur. Sonunda bu karar ağaçlarından rastgele bir orman oluşturulur. Test veri setinin her bir örneği, tüm karar ağaçları tarafından tahmin edilir ve bu ağaçların oylarına bağlı olarak nihai sınıflandırma sonucu döndürülür.

Orijinal eğitim veri seti $S=\{(x_i,y_j),i=1,2,\dots,N;j=1,2,\dots,M\}$ şeklinde formüle edilmiştir, burada x bir örnektir ve y, S'nin bir özellik değişkenidir. Yani, orijinal eğitim veri kümesi N örnek içerir ve her örnekte M özellik değişkeni vardır. RF algoritmasının oluşturulmasının ana süreci Şekil 4’te verilmiştir.



Şekil 4. Rastgele orman algoritmasının ana süreci [14]

Rastgele orman algoritmasının yapım aşamaları aşağıdaki gibidir.

Adım 1. Örnekleme k eğitim alt kümesi.

Bu adımda, k eğitim alt kümesi, orijinal eğitim veri kümesinden (S) bir önyükleme örnekleme şeklinde örneklenir. Yani, her örnekleme zamanında rastgele örnekleme ve değiştirme yöntemi ile S'den N kayıt seçilir. Geçerli adımdan sonra, k eğitim alt kümesi, S_{Train} eğitim alt kümelerinin bir koleksiyonu olarak oluşturulur ($S_{Train} = \{S_1, S_2, \dots, S_k\}$).

Aynı zamanda her bir örnekleme periyodunda seçilmeyecek olan kayıtlar Out-Of-Bag (OOB) veri seti olarak oluşturulmaktadır. Bu şekilde, k OOB seti, S_{OOB} koleksiyonu olarak oluşturulur ($S_{OOB} = \{OOB_1, OOB_2, \dots, OOB_k\}$).

Burada $k \ll N$, $S_i \cap OOB_i = \emptyset$ ve $S_i \cup OOB_i = S$. Her bir ağaç modelinin sınıflandırma doğruluğunu elde etmek için bu OOB setleri eğitim sürecinden sonra test seti olarak kullanılır.

Adım 2. Her bir karar ağacı modelinin oluşturulması.

Bir RF modelinde, her meta karar ağacı, her S_i eğitim alt kümesinden bir C4.5 veya CART algoritması tarafından oluşturulur. Her ağacın büyüme sürecinde, S_i veri kümesinin m özellik değişkeni, M değişkenden rastgele seçilir. Her ağaç düğümünün bölme işleminde, her bir özellik değişkeninin kazanç oranı hesaplanır ve en iyi olanı bölme düğümü olarak seçilir. Bu bölme işlemi, bir yaprak düğüm oluşturulana kadar tekrarlanır. Son olarak, k tane karar ağacı aynı şekilde k tane eğitim alt kümesinden eğitilir.

Adım 3. k ağacı bir RF modeline toplamak. k eğitilmiş ağaç, şu şekilde tanımlanan bir RF modelinde toplanır:

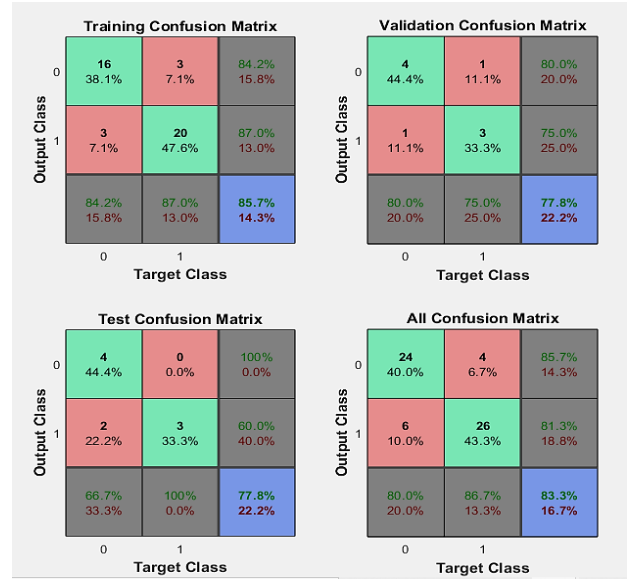
$$H(X, Q_j) = \sum_{i=1}^k h_i(x, Q_j), (j = 1, 2, \dots, m) \quad (5)$$

burada $h_i(x, Q_j)$ bir meta karar ağacı sınıflandırıcıdır, X, eğitim veri kümesinin girdi özellik vektörleridir ve Q_j , ağacın büyüme sürecini belirleyen bağımsız ve özdeş olarak dağıtılmış bir rastgele vektördür [14].

Rastgele orman algoritması WEKA programı kullanılarak gerçekleştirilmiştir. Programa giriş verileri girildikten sonra eğitim farklı cross validation değerleri ile gerçekleştirilmiştir. En iyi sonuç cross validation 10 olduğunda elde edilmiştir.

III. BULGULAR

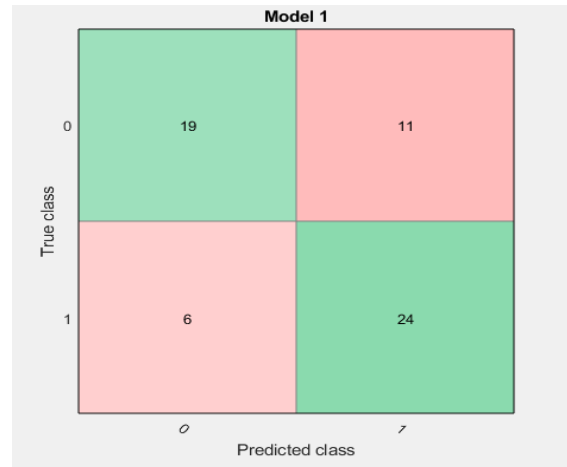
İleri beslemeli yapay sinir ağından elde edilen sonuç farklı hidden değeri, test, validasyon, eğitim değerleri için denenmiştir. En iyi sonucu 15 hidden katmanda, verilerin % 70' i eğitime, %15' i validasyona ve % 15' i de test için ayrıldığında gerçekleşmiştir. Elde edilen karmaşıklık matrisine baktığımızda eğitim için %85.7, validasyon için %77.8, test için %77.8 ve tüm ağ için %83.3' lük bir doğruluk elde edilmiştir. Şekil 4' de karmaşıklık matrisi verilmiştir.



Şekil 4. İleri beslemeli YSA ile elde edilen confusion (karmaşıklık) matrisi

Karar ağaçları algoritması ile elde edilen karmaşıklık matrisi Şekil 5' de verilmiştir. Program çıktısında %71.7' lik doğruluk elde edilmiştir. Karmaşıklık matrisine göre aşağıda verilen formül (6) ile de hesaplanırsa aynı sonuç elde edilmektedir.

$$\text{Doğruluk} = \frac{TP+TN}{TP+FP+TN+FN} \quad (6)$$



Şekil 5. Karar ağaçları ile elde edilen confusion (karmaşıklık) matrisi

WEKA program çıktısı Şekil 6' da verilmiştir. Rastgele orman algoritması 10 cross validation değeri ile 100 iterasyonda çalıştırıldığında %78.33' lük doğruluk elde edilmiştir.

Correctly Classified Instances	47	78.3333 %							
Kappa statistic	0.5667								
Mean absolute error	0.4645								
Root mean squared error	0.4674								
Relative absolute error	92.903 %								
Root relative squared error	93.4791 %								
Total Number of Instances	60								
== Detailed Accuracy By Class ==									
	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0,733	0,167	0,815	0,733	0,772	0,570	0,813	0,834	M
	0,833	0,267	0,758	0,833	0,794	0,570	0,813	0,749	B
Weighted Avg.	0,783	0,217	0,786	0,783	0,783	0,570	0,813	0,792	
== Confusion Matrix ==									
a	b	← classified as							
22	8	a = M							
5	25	b = B							

Şekil 6. Rastgele orman ile elde edilen sonuç

IV. SONUÇ

Elde edilen sonuçlar karşılaştırıldığında en iyi sonuç %83.3 ile ileri beslemeli yapay sinir ağında elde edilmiştir. Rastgele orman algoritmasının sonucu %78.33 ile karar ağaçları algoritmasına %71.7 göre daha iyi sonuç vermiştir. Karar ağacı bir kararın tüm olası sonuçlarını gösterirken, rastgele orman algoritması tüm karar ağaçlarının çıktılarını dayanarak nihai bir sonuç verir. Bundan dolayı çalışmada da görüldüğü gibi rastgele orman algoritması karar ağacı algoritmasına göre daha iyi bir sonuç vermiştir.

Milton yaptığı çalışmasında ISIC veri seti kullanarak farklı derin öğrenme metodları kullanmıştır [6]. Fakat kullandığı farklı yöntemlere rağmen en yüksek % 76 doğruluk elde etmiştir.

Dubal ve arkadaşlarının yaptığı ileri besleme yapay sinir ağı çalışmasında %76.9' luk bir doğruluk elde edilmiştir [7]. Yapılan çalışma ile karşılaştırıldığında daha iyi bir sonuç elde edilmiştir. Bunun sebebi ise Dubal ve arkadaşlarının internet ortamından rastgele görüntüler aldığından kaynaklandığı düşünülmektedir.

Daskula ve arkadaşları ISIC veri setini kullanarak derin öğrenme ile sınıflandırma gerçekleştirmişlerdir. Sonuçta %81 lik bir doğruluk elde edilmiştir [8].

Janney ve arkadaşları ISIC veri setini kullanarak dört yöntem üzerinde sınıflandırma yapmışlardır. k en yakın komşuluk (k- NN), destek vektör makineleri(SVM), çok katmanlı algılayıcı (MLP), rastgele orman ve karar ağacı yöntemlerini kullanmışlardır. Sonuçta k-NN, SVM, Karar Ağacı, MLP ve Rastgele Orman sınıflandırıcısının doğruluğu %71.5, %69, %86.9, %75 ve %93'tür [15]. Bu çalışmada yüksek başarımla elde edilmesinin sebebi çok fazla (900) görüntü üzerinde çalışılmasıdır.

Literatürdeki çalışmalar göz önüne alındığında iyi bir sonuç elde edilmiştir. Başarımın artırılması için kullanılan dermoskopi görüntü sayısının artırılması düşünülmektedir. Ayrıca çıkarılan özellikler artırılarak ve diğer makine öğrenmesi ve derin öğrenme metodları kullanılarak başarımın artırılması ileri çalışmaları için hedeflenmektedir.

KAYNAKLAR

- [1] (2020). Medicalpark. [Online]. Available: <https://www.medicalpark.com.tr/cilt-kanseri/hg-1808>
- [2] Yavuz, G. Ö., & Yavuz, İ. H. (2014). Melanositik Nevusler. *Van Tıp Dergisi*, 21(4), 259-268.

- [3] "Cancer facts and figures 2016," American Cancer Society
- [4] ÖZTÜRK, Banu, et al. "Kutanöz malign melanomda adjuvan medikal tedavi yaklaşımları." *Türk Onkoloji Dergisi* 25 (2010): 170-80
- [5] Codella, N., Cai, J., Abedini, M., Garnavi, R., Halpern, A., & Smith, J. R. (2015, October). Deep learning, sparse coding, and SVM for melanoma recognition in dermoscopy images. In *International workshop on machine learning in medical imaging* (pp. 118-126). Springer, Cham.
- [6] Milton, M. A. A. (2019). Automated skin lesion classification using ensemble of deep neural networks in ISIC 2018: Skin lesion analysis towards melanoma detection challenge. *arXiv preprint arXiv:1901.10802*.
- [7] Dubal, P., Bhatt, S., Joglekar, C., & Patil, S. (2017, November). Skin cancer detection and classification. In *2017 6th international conference on electrical engineering and informatics (ICEEI)* (pp. 1-6). IEEE.
- [8] Dascalu, A., & David, E. O. (2019). Skin cancer detection by deep learning and sound analysis algorithms: A prospective clinical study of an elementary dermoscope. *EBioMedicine*, 43, 107-113.
- [9] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *nature*, 542(7639), 115-118.
- [10] 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.
- [11] Noel Codella, Veronica Rotemberg, Philipp Tschandl, M. Emre Celebi, Stephen Dusza, David Gutman, Brian Helba, Aadi Kallou, Konstantinos Liopyris, Michael Marchetti, Harald Kittler, Allan Halpern: "Skin Lesion Analysis Toward Melanoma Detection 2018: A Challenge Hosted by the International Skin Imaging Collaboration (ISIC)", 2018; <https://arxiv.org/abs/1902.03368>
- [12] (2021). Mathworks website [Online]. Available: <https://www.mathworks.com/help/deeplearning/gs/classify-patterns-with-a-neural-network.html>
- [13] DUYGU, B., KOCAOĞLU, M., & COŞKUN, A. Karar Ağaçları ile Otistik Spektrum Bozukluğu Tanısı Koyma.
- [14] Chen, J., Li, K., Tang, Z., Bilal, K., Yu, S., Weng, C., & Li, K. (2016). A parallel random forest algorithm for big data in a spark cloud computing environment. *IEEE Transactions on Parallel and Distributed Systems*, 28(4), 919-933.
- [15] Roslin, S. E. (2020). Classification of melanoma from Dermoscopic data using machine learning techniques. *Multimedia tools and applications*, 79(5), 3713-3728.

EK 1. TABLO 1. ÖZELLİK ÇIKARIMININ SONUÇLARI

id	yaş	cinsiyet	entropi	stndrt sap	alan	homjnt	kontrast	enerji	korelasyon	skewness	kurtosis	wavelet	mal/ben
ISIC_417	70	0	1	0.5	9.09E+04	0.0293	0.9414	0.4716	0.9853	3.5797	92.3962	1.0067	M
ISIC_422	85	0	1	0.5	106312	0.01	0.98	0.4901	0.995	2.4489	45.3124	1.0001	M
ISIC_425	65	1	0.9996	0.4999	7.27E+04	0.0221	0.9557	0.4786	0.9889	3.899	83.6358	0.9767	M
ISIC_430	85	0	0.9994	0.4998	3.15E+04	0.0342	0.9315	0.4675	0.9829	0.9322	10.8409	1.0281	M
ISIC_432	70	0	1	0.5	7.54E+04	0.0096	0.9809	0.4905	0.9952	2.4304	64.2215	0.9949	M
ISIC_434	75	0	1	0.5	2.96E+04	0.0059	0.9881	0.4941	0.997	4.3474	62.4525	0.999	M
ISIC_436	75	0	1	0.5	7.06E+04	0.0154	0.9692	0.4848	0.9923	3.587	60.1044	0.9964	M
ISIC_444	70	0	1	0.5	5.91E+04	0.0154	0.9692	0.4848	0.9923	3.3884	40.7705	0.9938	M
ISIC_445	80	1	1	0.5	6.63E+04	0.0209	0.9581	0.4795	0.9895	-0.3833	43.2426	1.0014	M
ISIC_451	70	0	1	0.5	7.53E+04	0.019	0.96	0.4814	0.9905	-0.6691	97.5783	1.0009	M
ISIC_456	80	0	0.9999	0.5	3.42E+04	0.0093	0.9815	0.4908	0.9954	3.9616	67.9272	0.998	M
ISIC_459	65	0	0.9999	0.5	1.04E+05	0.0355	0.9289	0.4658	0.9822	0.5934	17.3544	1.0093	M
ISIC_462	60	1	1	0.5	5.44E+04	0.006	0.988	0.4941	0.997	3.5163	50.112	0.9904	M
ISIC_465	80	1	1	0.5	1.11E+05	0.02209	0.9582	0.4796	0.9896	0.6612	9.6139	0.9992	M
ISIC_468	70	1	1	0.5	3.55E+04	0.0133	0.9735	0.4869	0.9934	3.8403	35.2615	0.9996	M
ISIC_476	50	0	1	0.5	38172	0.0099	0.9803	0.4902	0.9951	4.449	95.2319	0.9937	M
ISIC_481	50	1	1	0.5	7.08E+04	0.0059	0.9882	0.4941	0.997	5.547	82.6401	1.0003	M
ISIC_485	70	0	0.9996	0.4999	1.16E+05	0.0236	0.9527	0.4772	0.9882	3.9382	181.3678	0.9747	M
ISIC_486	70	0	0.9998	0.4999	6.92E+04	0.0364	0.9272	0.4651	0.9818	-0.7652	19.8594	1.0166	M
ISIC_487	80	0	1	0.5	1.34E+05	0.0233	0.9534	0.4773	0.9884	1.6322	98.0939	0.9972	M
ISIC_491	40	0	1	0.5	4.61E+04	0.0206	0.9587	0.4798	0.9897	2.7194	226.2765	0.993	M
ISIC_492	40	0	1	0.5	118221	0.006	0.9881	0.4941	0.997	4.5357	192.7496	0.9943	M
ISIC_494	45	1	1	0.5	5.71E+04	0.014	0.9719	0.4862	0.993	4.1009	49.5041	1.0032	M
ISIC_501	85	1	1	0.5	8.52E+04	0.0212	0.9575	0.4792	0.9894	0.5538	56.5379	1.0051	M
ISIC_502	85	1	1	0.5	1.08E+05	0.0109	0.9781	0.4892	0.9945	6.8678	177.2046	0.992	M
ISIC_504	70	1	1	0.5	1.15E+05	0.0297	0.9407	0.4712	0.9852	1.5819	20.1511	0.994	M
ISIC_507	80	0	0.9997	0.4999	1.10E+05	0.0318	0.9364	0.4694	0.9841	0.4463	11.8532	1.0192	M
ISIC_508	55	0	0.9994	0.4998	3.24E+04	0.0448	0.9104	0.4576	0.9776	1.5515	216.729	0.9708	M
ISIC_510	55	0	0.9999	0.5	7.60E+04	0.0327	0.9346	0.4685	0.9837	1.0373	27.341	1.0084	M
ISIC_511	65	0	0.9999	0.5	1.34E+05	0.0365	0.9269	0.4649	0.9817	-4.3432	53.2293	0.9889	M

EK 1. TABLO 1. ÖZELLİK ÇIKARIMININ SONUÇLARI

id	yaş	cinsiyet	entropi	stndrt sap	alan	homjnt	kontrast	enerji	korelasyon	skewness	kurtosis	wavelet	mal/ben
ISIC_514	45	0	0.9989	0.4996	1.29E+05	0.0336	0.9327	0.4684	0.9832	-0.6935	41.1232	0.9602	B
ISIC_518	55	0	1	0.5	5.74E+04	0.2555	0.949	0.4752	0.9873	1.3677	14.3295	1.0014	B
ISIC_520	30	1	0.9949	0.4982	7308	0.0887	0.8212	0.4228	0.9556	0.8551	14.3624	0.9221	B
ISIC_521	30	1	0.998	0.4993	1.80E+04	0.0138	0.9724	0.4876	0.9931	3.3312	48.0997	0.9377	B
ISIC_525	75	0	1	0.5	5.83E+04	0.0263	0.9474	0.4744	0.9869	0.6402	42.2092	0.9993	B
ISIC_530	75	0	1	0.5	3.93E+04	0.0238	0.9523	0.4767	0.9881	3.1384	55.6061	0.9988	B
ISIC_532	50	0	1	0.5	1.18E+05	0.0238	0.9524	0.4768	0.9881	0.6393	6.8851	0.9952	B
ISIC_533	15	0	0.9999	0.5	9.77E+04	0.0231	0.9538	0.4775	0.9885	-3.8915	400.4259	0.9867	B
ISIC_534	35	0	1	0.5	9.57E+04	0.0171	0.9659	0.4832	0.9915	3.6638	72.9932	0.9945	B
ISIC_538	40	1	1	0.5	3.53E+04	0.0119	0.9763	0.4883	0.9941	4.3712	64.0998	0.9943	B
ISIC_541	40	1	1	0.5	4.96E+04	0.0076	0.9847	0.4924	0.9962	3.4395	84.8536	0.9951	B
ISIC_542	45	0	1	0.5	4.04E+04	0.0071	0.9857	0.4929	0.9964	4.9308	71.2999	0.9953	B
ISIC_543	45	0	1	0.5	4.91E+04	0.0159	0.9681	0.4843	0.992	2.7131	34.8624	0.9979	B
ISIC_545	20	1	1	0.5	1.02E+05	0.0062	0.9877	0.4939	0.9969	3.9594	70.2063	0.9948	B
ISIC_546	20	1	0.9999	0.5	5.40E+04	0.0202	0.9596	0.4803	0.9899	3.7586	54.4969	1.0095	B
ISIC_547	25	1	1	0.5	5.94E+04	0.0088	0.9824	0.4913	0.9956	4.5068	61.3985	0.9945	B
ISIC_549	25	1	0.9997	0.4999	7.63E+04	0.0373	0.9253	0.4643	0.9813	-0.1854	7.0744	0.9817	B
ISIC_550	25	1	1	0.5	4.73E+04	0.008	0.984	0.4921	0.996	4.4331	64.7683	0.99	B
ISIC_551	75	0	0.9998	0.4999	1.13E+05	0.0225	0.9549	0.4781	0.9887	3.6849	69.7613	1.0161	B
ISIC_554	35	1	0.9999	0.5	1.15E+04	0.0089	0.9821	0.4912	0.9955	2.9747	55.6825	0.9874	B
ISIC_555	15	1	1	0.5	7.95E+04	0.0314	0.9372	0.4696	0.9843	3.622	164.6249	1.0035	B
ISIC_557	65	1	0.9998	0.4999	3.46E+04	0.0135	0.973	0.4868	0.9933	3.0651	30.7206	0.9786	B
ISIC_559	60	1	1	0.5	6.32E+04	0.0091	0.9818	0.491	0.9955	4.72	204.5696	0.9982	B
ISIC_560	60	1	0.9999	0.5	7.20E+04	0.0112	0.9776	0.489	0.9944	2.6616	25.76	1.0063	B
ISIC_561	50	0	1	0.5	1.03E+05	0.0134	0.9732	0.4868	0.9933	0.2585	2.8263	0.9915	B
ISIC_562	50	0	1	0.5	3.67E+04	0.0128	0.9743	0.4873	0.9936	4.3033	91.4501	0.9951	B
ISIC_565	60	1	1	0.5	1.14E+05	0.0296	0.9408	0.4713	0.9852	0.5198	8.4912	0.9918	B
ISIC_567	15	0	1	0.5	2.58E+04	0.044	0.9119	0.4579	0.978	2.408	46.9422	1.0054	B
ISIC_571	60	1	1	0.5	1.35E+05	0.0291	0.9419	0.4718	0.9855	2.3016	105.5782	1.0041	B
ISIC_574	75	0	1	0.5	7.38E+04	0.0661	0.8678	0.4383	0.976	-2.2791	170.8791	0.9919	B