

Derin Öğrenme ile Optik Koherens Tomografi Görüntülerinin Sınıflandırılması

Numan Halit GÜLDEMİR^{1*}, Ahmet ALKAN²

¹ Elektrik-Elektronik Mühendisliği Bölümü, Fen Bilimleri Enstitüsü, Kahramanmaraş Sütçü İmam Üniversitesi, Kahramanmaraş, Türkiye

² Elektrik-Elektronik Mühendisliği Bölümü, Mühendislik ve Mimarlık Fakültesi, Kahramanmaraş Sütçü İmam Üniversitesi, Kahramanmaraş, Türkiye

*¹ nhguldemir@gmail.com, ² aalkan@ksu.edu.tr

(Geliş/Received: 19/03/2021;

Kabul/Accepted: 18/04/2021)

Öz: Bu makalede, dünyada görme bozukluklarının %8.7'sinin sebebi olan yaşa bağlı makula dejenerasyonu hastalığının tespit edilmesi için otomatik bir sistem sunulmuştur. Hastalığın teşhisi için, son zamanlarda görüntüden özellik tanıma çıkarma işlemlerinde sıklıkla kullanılan Evrişimsel Sinir Ağı mimarisinden faydalanılmıştır. Çalışmada, yaşa bağlı makula dejenerasyonu hastalığının sınıflandırılması için Evrişimsel Sinir Ağı, etiketli 28031 adet Optik Koherens Tomografi (OKT) görüntüsü ile eğitilmiştir. Eğitilen Evrişimsel Sinir Ağı 6900 adet OKT görüntüsü ile test edilmiştir. Eğitim ve test için Xception, VGG16, InceptionV3 ve Resnet50 derin öğrenme ağları kullanılmıştır. Eğitilen ağlar karşılaştırılıp OKT veri seti için ağların performansları değerlendirilmiştir. Derin öğrenme ağlarının test edilmesi sonucunda Resnet50 ağı 0.95, VGG16 ağı 0.95, InceptionV3 ağı 0.97 ve Xception ağı 0.98 doğruluk oranı OKT veri setini sınıflandırmıştır. 0.96 değerindeki F1 skoruyla Xception ağı kullanılan OKT veri seti üzerindeki en başarılı sınıflandırmayı elde etmiştir.

Anahtar kelimeler: Yaşa bağlı makula dejenerasyonu, derin öğrenme, Evrişimsel Sinir Ağları.

Classification of Optical Coherence Tomography Images via Deep Learning

Abstract: In this article, an automatic system is presented to detect age-related macular degeneration, which is the cause of 8.7% of visual impairment in the world. For the diagnosis of the disease convolutional neural network, which has recently been used in feature extraction from the image, has been utilized. In the study, the convolutional neural network was trained with 28031 labeled Optical Coherence Tomography (OCT) images to classify age-related macular degeneration. The trained convolutional neural network was tested with 6900 OCT images. Xception, VGG16, Inception V3 and Resnet50 deep learning networks were used for training and testing. The trained networks were compared and the performance of the networks was evaluated for the OCT dataset. Classification of OCT images using Resnet50 network achieved 0.95, VGG16 network achieved 0.95, InceptionV3 network achieved 0.97 and Xception network achieved accuracy of 0.98. Xception network achieved F1 score of 0.96 which is highest rate of classification on the OCT dataset.

Key words: Age-related macular degeneration, deep learning, convolutional neural network.

1. Giriş

Retina ışığı yakalama özelliğine sahip bir dokudur ve 5 nöral hücre sınıfı ve 50'den fazla farklı hücre tipinden oluşur. Bu hücrelerin her biri, görüntü oluşturmak ve görsel algıyı elde etmek için beyin görme merkezlerinde gerekli olan bilgileri sağlayan benzersiz işlevlere sahiptir [1]. Makula, retinada bulunan 6 mm çapında, görüntü odağını ve renkli görmeyi sağlayan özel bir yapıdır [2]. Yaşa bağlı makula dejenerasyonu (YBMD) hastalığı, makulanın aşınmasıyla ilgili gelişen bir hastalıktır.

Dünyada 2020 yılında yaklaşık 196 milyon olarak tahmin edilen YBMD hastalığının 2050'li yıllarda 288 milyona kadar yükseleceği öngörülmektedir. Tedavi edilmediğinde genellikle okuma, araba kullanma ve nesnelere tanıma gibi günlük aktiviteleri etkileyen merkezi görme keskinliğinin kaybı oluşmaktadır ve hastalık ilerlediğinde ise körlüğe neden olmaktadır. Dünya çapındaki tüm körlüklerin %8.7'sini oluşturan YBMD hastalığı gelişmiş ülkelerde, 60 yaşın üzerindeki hastalarda yaşanan körlüklerin en büyük nedenlerinden biridir [3]. Hastalığın oluşumuna birçok risk faktörü etki eder. Literatürde ana risk faktörleri yaş, aile geçmişi ve sigara içme olarak belirtilmektedir. Vücut kitle indeksi, cinsiyet, kardiyovasküler hastalıklar, hipertansiyon, yeme alışkanlıkları ve etnik köken de risk faktörleri arasındadır [4].

Retinanın altında retina pigment epiteli (RPE) adı verilen bir membran bulunur. Bu yapı retinada bulunan ışığa duyarlı hücrelere gelen fazla ışığı emer ve o hücrelere besin ve oksijen taşınmasını sağlar [5]. Eğer bu zar görevini

* Sorumlu yazar: nhguldemir@gmail.com. Yazarların ORCID Numarası: ¹ 0000-0003-1202-6841, ² 0000-0003-0857-0764

yerine getiremezse besin aktarımı tam olarak gerçekleşemez ve artık besinler retina pigment epiteli altında birikir. Bu artıklar drusen olarak adlandırılır [6]. YBMD hastalığı bu drusenlerin varlığıyla ilişkilidir. Drusen yok ise veya çapı 63 mikrondan küçük ise YBMD hastalığı yoktur. YBMD'nin başlangıç aşamasında ise drusen çapı 63 ile 124 mikron arasında değişmektedir ve bu aşamada genellikle görme kaybı söz konusu değildir. Orta ve ileri evrelerde drusen çapı 125 mikron veya daha fazla olarak ölçülmektedir [7]. Drusen tespiti için farklı yöntemler bulunmaktadır. Başlıca yöntemler arasında fundus görüntülerinin, Optik Koherens Tomografi (OKT) görüntülerinin ve Optik Koherens Tomografi Anjiyografi (OKTA) görüntülerinin incelenmesi bulunmaktadır. Bu görüntüler aracılığı ile drusen tespiti yapıp hastalık hakkında bilgi edinilmektedir.

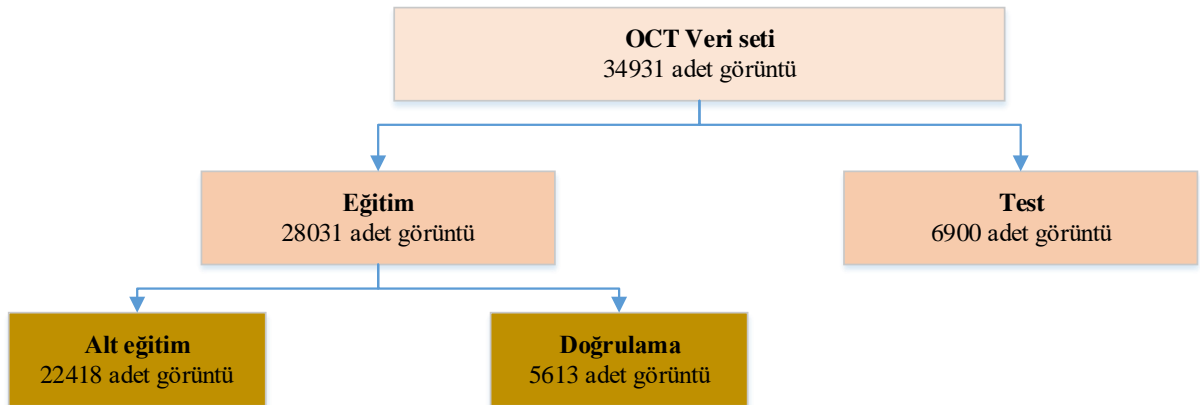
Drusen tespitinin manuel olarak yapılması öznel ve hataya açıktır, aynı zamanda vakit alıcıdır. Bu nedenle otomatize olarak drusen tespitinin yapılması uzman hekimlerin yüklerini azaltacak ve kararlarına destek olacaktır. Bu çalışmada drusen tespiti için makine öğrenmesi yaklaşımını kullanan derin öğrenmeden faydalanılmıştır. Çalışmada kullanılan mimari ise Evrişimsel Sinir Ağları (ESA) olarak isimlendirilmektedir.

Evrişimsel Sinir Ağları, görüntülerden özellik çıkartma işlemlerinde sıklıkla kullanılmaktadır. Günümüzde birçok farklı ESA mimarisi bulunmaktadır. Bu ağların başarıları ise genellikle ImageNet adı verilen ve içerisinde 14 milyona yakın görsel ve 1000 adet sınıf bulunan bir veri seti üzerinden test edilmektedir [8]. Bu çalışmada kullanılan model olan Xception ise ImageNet veri seti üzerinde %94.5 başarı oranı elde etmiştir [9].

Tsuji ve arkadaşları 2020 yılında yapmış oldukları bir çalışmada Kaggle OKT veri seti üzerinde "capsule network" ismi verilen ağ yapısını kullanmış ve veri setinde bulunan dört farklı sınıfla sistemi eğitmişlerdir. Eğitilen sistemin testi için kullanılan veri setinde 250'şer adet Yaşa Bağlı Makula Dejenerasyonu, Diyabetik Maküler Ödem, Koroidal Neovaskülarizasyon ve kontrol sınıfı görüntüsü bulunmaktadır. Test sonucunda %99,6'lık bir başarı yüzdesi elde etmişlerdir [10]. Tsuji ve arkadaşlarının yapmış oldukları çalışmanın aksine, tarafımızdan sunulan çalışmada eğitim setinden feragat edilerek test veri seti artırılmış ve böylelikle daha gerçekçi bir başarı oranı ortaya koymak amaçlanmıştır.

2. Materyal ve Metot

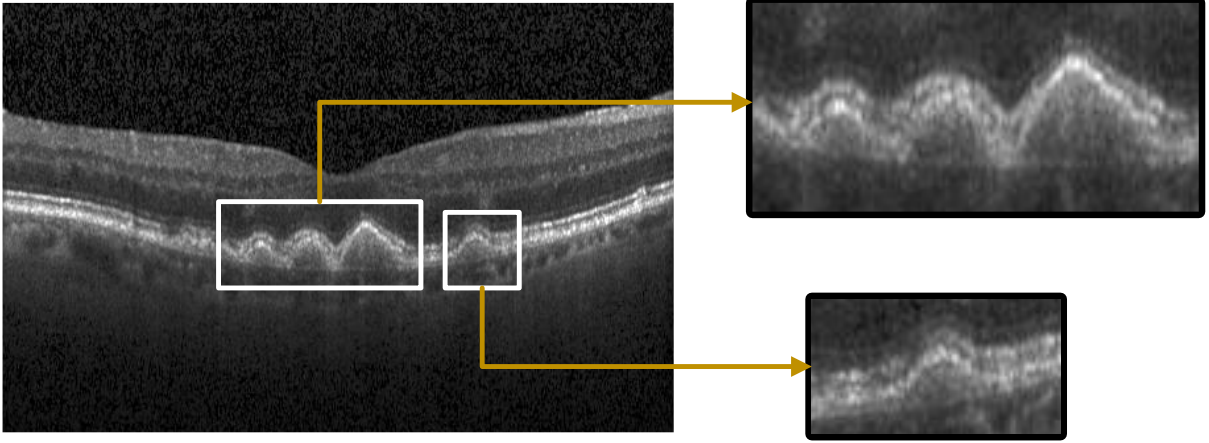
Çalışmada, drusenin tespiti için gerekli OKT veri seti, erişime açık Kaggle web sayfasından edinildi [11]. Bu veri setinde toplamda 34931 adet etiketli OKT görüntüsü bulunmaktadır. OKT görüntüleri California San Diego Üniversitesi, California Retinal Araştırma Vakfı, Shanghai First People's Hastanesi ve Beijing Tongren Göz Merkezi'nde 1 Haziran 2013 ve 1 Mart 2017 tarihleri arasında bulunan yetişkin hastalara aittir. Bu veri setinde bulunan OKT görüntülerinden eğitim için 22418 adet, doğrulama için 5613 adet ve modelin başarımını ölçmek için ise 6900 adet kullanılmıştır (Şekil 1.).



Şekil 1. Veri setinin kullanımı

2.1. OKT görüntülerinin sınıflandırılması

OKT görüntülerinde incelenerek retinada drusenlerin varlığı hakkında bilgi edinilebilmektedir. YBMD hastası olmayan bir insanda retina pigment epiteli düz bir yapıda iken YBMD hastalarında retina pigment epitelinde drusenlerin varlığından dolayı tepecikler oluşmaya başlar. YBMD hastalığı olan bir insanın OKT görüntüsü Şekil 2'deki gibidir.

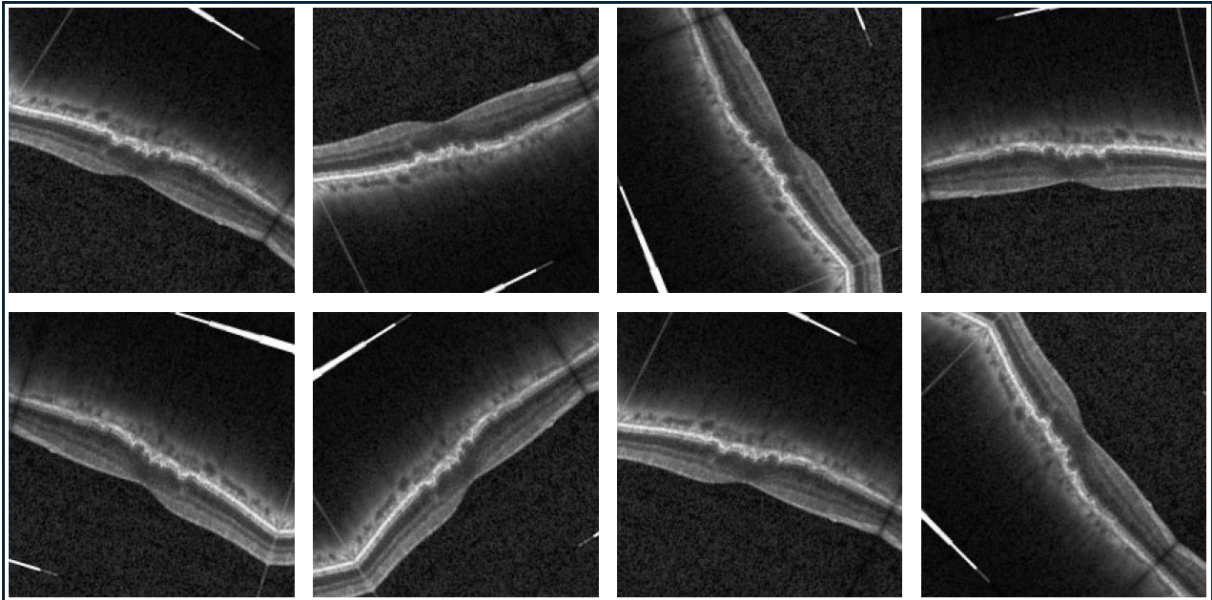


Şekil 2. Drusen görüntüsü

Drusenlerin tespiti için Evrişimsel Sinir Ağı mimarisinden faydalanılmıştır. Veri setinde bulunan “normal” ve “drusen” olarak etiketli OKT görüntüleri Evrişimsel Sinir Ağları’nın eğitilmesinde kullanılmıştır. Kullanılan ESA Xception derin öğrenme ağı olarak seçilmiştir. Xception derin öğrenme ağı dışında veri seti, en popüler ağlardan Resnet50, VGG16 ve InceptionV3 ağları ile de eğitilerek Xception ağının OKT veri seti üzerindeki sınıflandırma başarısı diğer ağlarla kıyaslanmıştır.

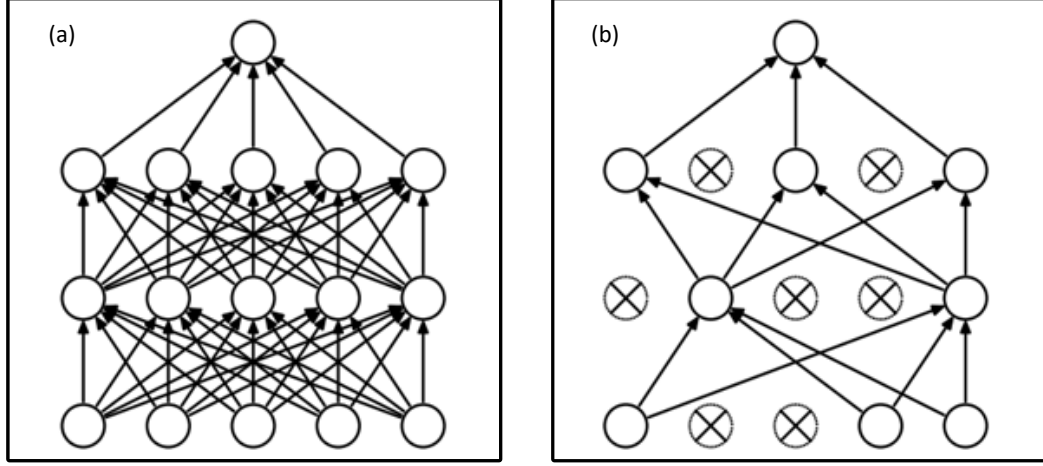
2.1.1. Deneysel çalışma

Kullanılan Evrişimsel Sinir Ağları’nda, veri setine uygun olması amacı ile veri setine göre giriş ağı giriş boyutunun ayarlanması, görüntü artırma işlemleri gibi değişiklikler yapılmıştır. Şekil 5’te gösterildiği gibi giriş görüntü boyutu 200x200x3 olarak ayarlanmıştır. Daha sonra görüntüler eğitim veri kümesinin artırılması ve OKT görüntü çekimlerindeki olası hata paylarının tolere edilebilmesi amacı ile rastgele açılarda sağa ve sola döndürülme, ayna görüntüsünü alma gibi yöntemler kullanılmıştır (Şekil 3.). Sonrasında görüntüdeki tüm pikseller 255’e bölünerek 0-1 aralığına indirgenerek normalizasyon yapılmıştır.

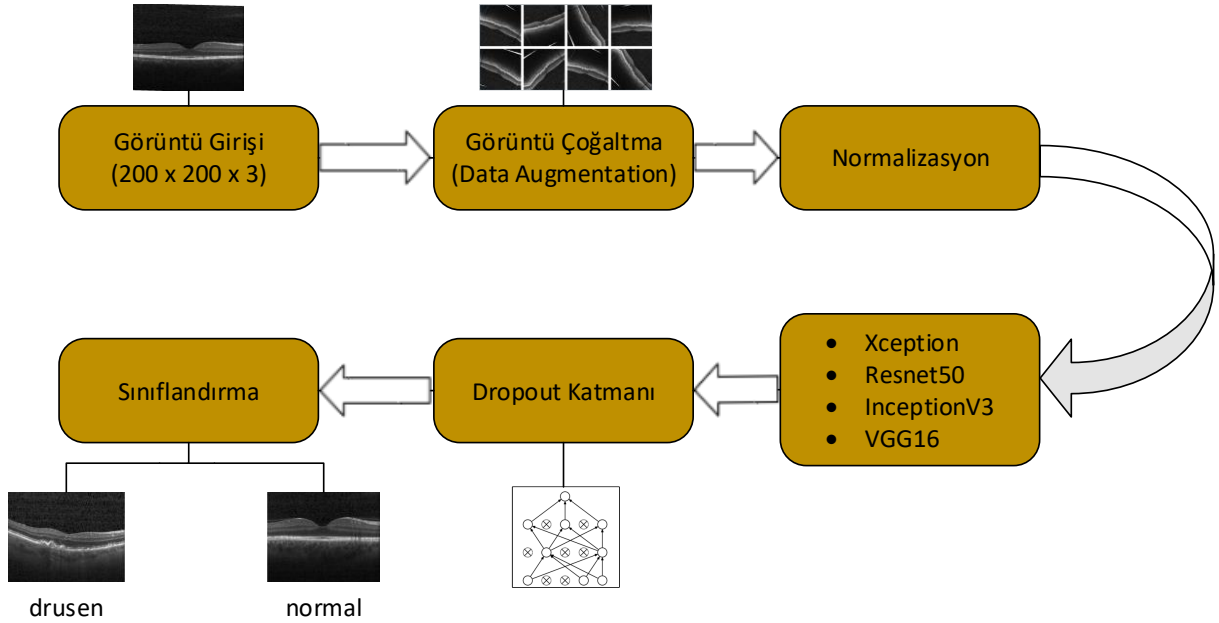


Şekil 3. Veri çoğaltma işlemi

Eğitimi uzun süren karmaşık ağlarda aşırı öğrenme (overfitting) olasıdır. Bunun önüne geçmek için ağıdaki sınıflandırma katmanından önce dropout katmanı kullanılmıştır. Bu katman ağıdaki bazı nöronların arasındaki ilişkiyi keserek o nöronları göz ardı eder. Bu sayede aşırı öğrenme önlenmiş olur. Şekil 4'te dropout katmanı uygulandıktan sonra ağıdaki nöron ilişkisi verilmiştir. Dropout katmanından sonra ağı çıkışında ikili sınıflandırma yapmak amacıyla sigmoid fonksiyonu kullanılmıştır ve birey hasta veya hasta değil olarak sınıflandırılmıştır.



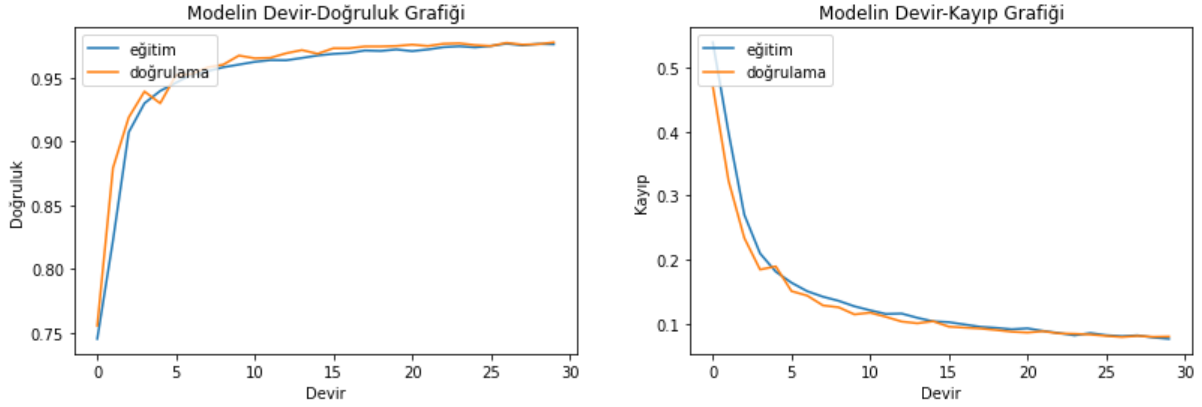
Şekil 4. a) Normal sinir ağı, b) Dropout katmanı uygulanmış sinir ağı



Şekil 5. Derin öğrenme akış şeması

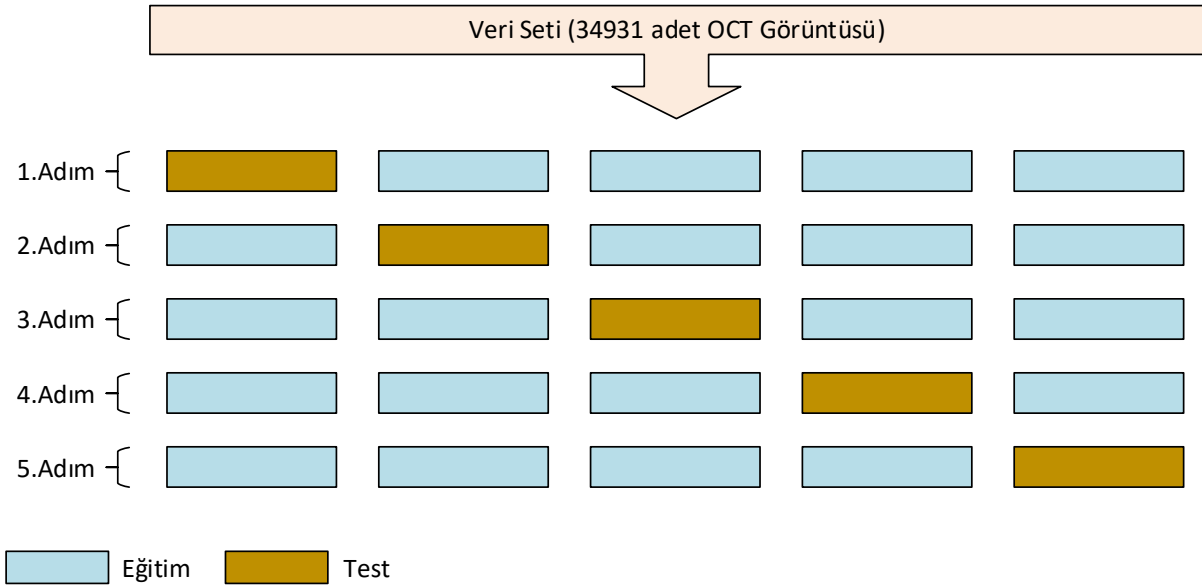
Ağın eğitilmesinde kullanılan bilgisayarda NVidia RTX 4000 Quadro GPU kartı ve 64 GB RAM donanımları bulunmaktadır. Ağ eğitilirken doğrulama veri setindeki kayıp fonksiyonları her adımda hesaplandı ve kayıp fonksiyonunun değerinin azalışı gözlemlendi. Doğrulama veri setindeki kayıp fonksiyonu değeri artışa geçtiğinde aşırı öğrenmeden kaçınmak amacıyla eğitim durduruldu.

Eğitimin bir adımının devir-doğruluk ve devir-kayıp grafikleri Şekil 6'da gösterilmiştir. Ağın eğitilmesi 29 devir boyunca devam etmiştir. Beklenildiği üzere doğruluk oranı her devir sonunda artmış ve hesaplanan kayıp her devir sonrası azalmıştır.



Şekil 6. Ağın eğitim grafikleri

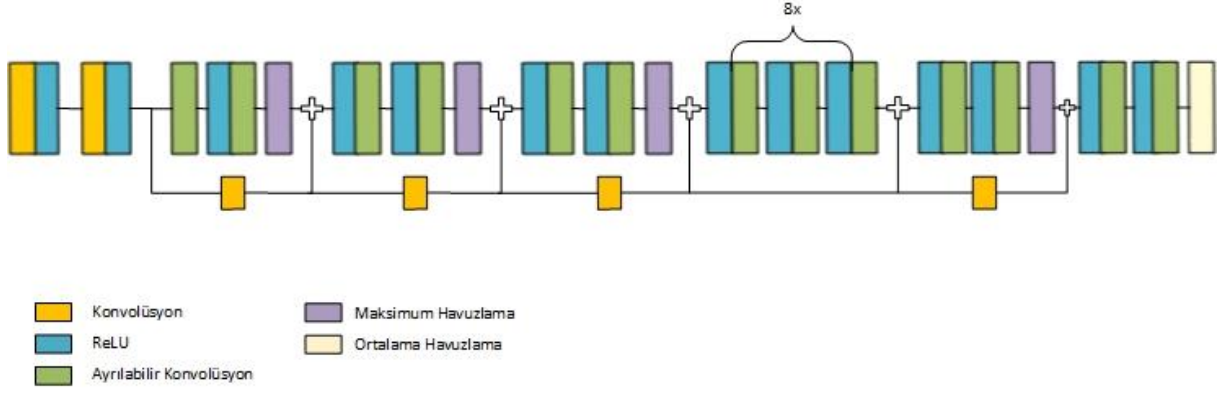
Ağların test parametrelerinin daha iyi bir şekilde değerlendirebilmesi için 5 katlamalı çapraz doğrulama yöntemi kullanılmıştır. Elde bulunan görüntülerden, her adımda farklı eğitim ve test grupları seçilerek ağ eğitilmiştir (Şekil 7). Her eğitim sonrası test veri seti kullanılarak parametreler hesaplanmıştır. Beş adımda da parametreler hesaplandıktan sonra ortalamaları alınarak 5 kat katlamalı çapraz doğrulama uygulanmıştır.



Şekil 7. 5 kat çapraz doğrulama gösterimi

2.1.2 Xception derin öğrenme ağı

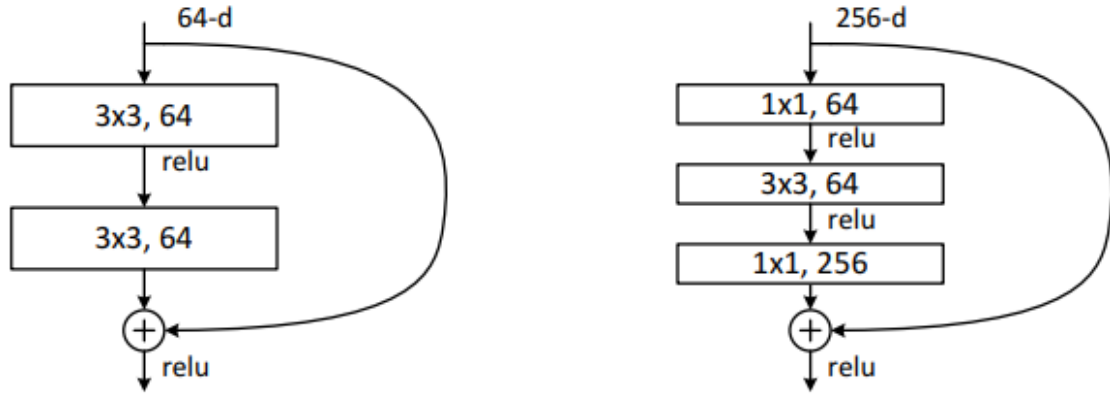
Çalışmada kullanılan Xception ağı François Chollet ve arkadaşları tarafından 2017 yılında ortaya atılmıştır [12]. Toplamda 170 adet katmandan oluşmaktadır. Toplamda 23 milyon civarında parametre bulunan Xception derin öğrenme ağı ImageNet veri seti üzerinde %94,5 doğruluk oranı elde etmiştir. Görüntü girişi 299x299x3 boyutundadır. Her konvolüsyon işlemi relu aktivasyonu takip etmektedir ve bu derin öğrenme ağında kullanılan havuzlama işlemi maksimum havuzlamadır. Ağın çıkışında ise ortalama havuzlama katmanı bulunmakta ve sonrasında sınıflandırma için softmax aktivasyonu bulunmaktadır. Ağın çıkışında ek olarak aşırı öğrenmeyi (ezberleme) engellemek adına dropout katmanı bulunmaktadır. Xception ağının akış şeması Şekil 8'de gösterildiği gibidir.



Şekil 8. Xception derin öğrenme ağı modeli

2.1.3. Resnet50 derin öğrenme ağı

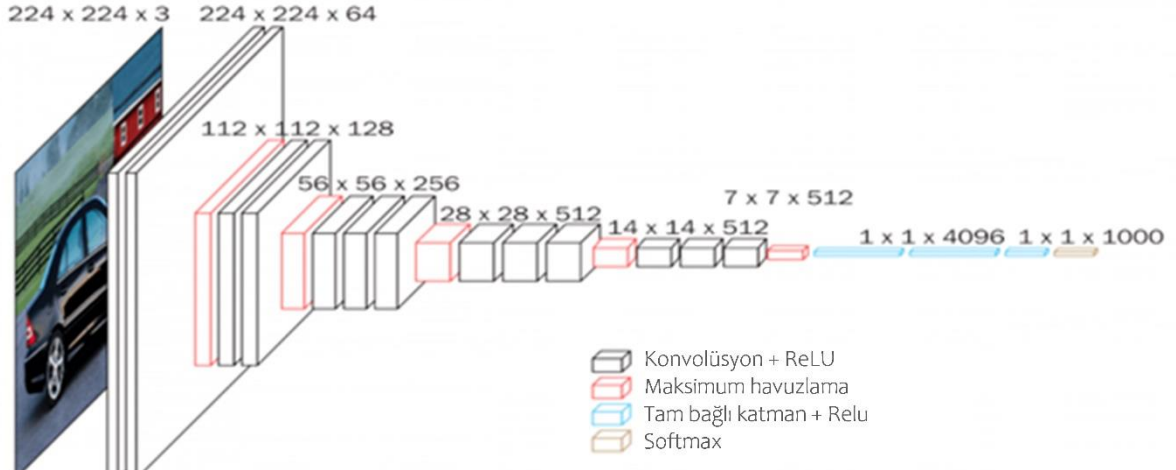
Kaiming He ve arkadaşları tarafından 2015 yılında ortaya atılan Resnet50 ağı 177 adet katmandan oluşmaktadır [13]. Görüntü girişi boyutu $224 \times 224 \times 3$ 'tür. Resnet50 ağı Şekil 9'da gösterildiği gibi daha hızlı eğitim yapabilmek için bazı katmanlarda atlama blokları kullanır. Ağda toplamda 25 milyon civarında parametre bulunmaktadır. ImageNet veri seti üzerinde %93,29 doğruluk elde etmiştir.



Şekil 9. Resnet50 derin öğrenme ağı bağlantı örneği

2.1.4. VGG16 derin öğrenme ağı

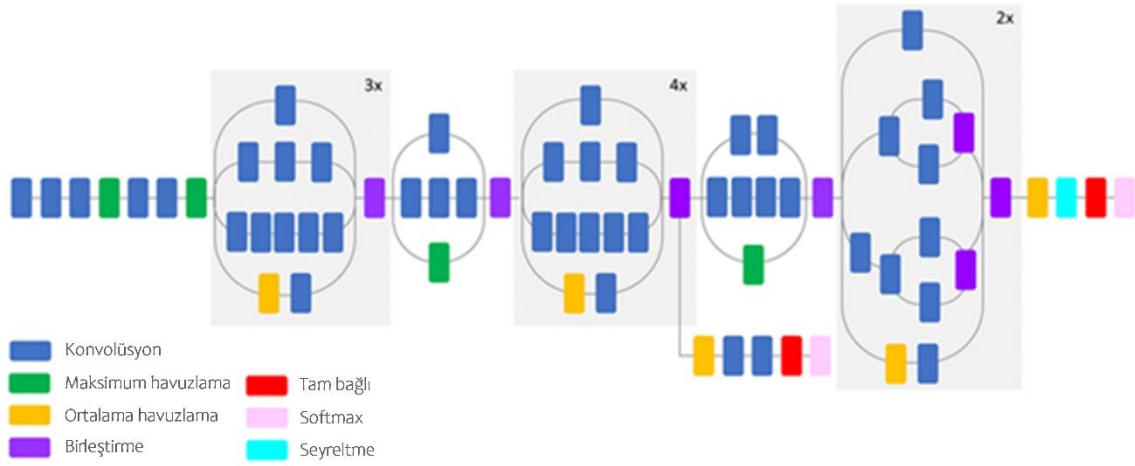
Karen Simonyan ve Andrew Zisserman tarafından 2014 yılında ortaya atılmıştır [14]. 41 katmandan oluşan bu ağın görüntü girişi $224 \times 224 \times 3$ boyutlarındadır. Toplamda 13 adet konvolüsyon katmanı, 15 adet relu aktivasyon katmanı, 5 adet maksimum havuzlama katmanı, 3 adet tam bağlantılı katmanı, 2 adet dropout katmanı bulunmaktadır. Ağın çıkışında ise softmax sınıflandırma katmanı bulunmaktadır. Ağdaki parametre sayısı 138 milyon civarındadır ve ImageNet veri seti üzerinde %91,6 doğruluk oranı elde etmiştir. Şekil 10'da VGG16 ağının katmanları gösterilmiştir.



Şekil 10. VGG16 derin öğrenme ağı modeli

2.1.5. Inception-V3 derin öğrenme ağı

Szegedy ve arkadaşları tarafından 2015 yılında sunulan bu ağ 315 katmandan oluşmaktadır [15]. Ağdaki görüntü girişi 299x299x3 boyutundadır. Ağdaki parametre sayısı 24 milyon civarında olup ImageNet veri seti üzerinde %93,7 doğruluk elde etmiştir. InceptionV3 ağının şeması Şekil 11’de verilmiştir.



Şekil 11. InceptionV3 derin öğrenme ağı modeli

Eğitilen modeller, 5200 adet sağlıklı ve 1700 adet YBMD hastası olan bireylerin OKT görüntüleri kullanılarak test edilmiştir. Eğitim sonucunda yorumlanması gereken hiper parametreler doğruluk, kesinlik, duyarlılık ve F1 skoru değerleri Tablo 1’de verilmiştir.

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

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

$$Duyarluluk = \frac{TP}{TP + FN} \quad (3)$$

$$F1 Skoru = \left(\frac{Duyarluluk^{-1} + Kesinlik^{-1}}{2} \right)^{-1} \quad (4)$$

TP (true positive) pozitif tahmin edilen verilerin kaç tanesinin pozitif olduğu, TN (true negative) negatif tahmin edilen verilerin kaç tanesinin negatif olduğu, FP (false positive) pozitif tahmin edilen verilerin kaç tanesinin negatif olduğu ve FN (false negative) negatif tahmin edilen verilerin kaç tanesinin pozitif olduğu bilgisini verir.

Doğruluk oranı değeri özellikle derin öğrenme ile hastalık sınıflandırma işlemlerinde tek başında yeterli değildir. Çünkü pozitif sınıflandırılması gereken bir verinin negatif sınıflandırılması ile negatif sınıflandırılması gereken bir verinin pozitif sınıflandırılmasının tıp alanındaki etkisi farklıdır. Hasta olan bir verinin hasta değil şeklinde sınıflandırılması büyük hatalara yol açabilir. Bu nedenle kesinlik, duyarlılık ve F1 skoru parametrelerinin de işleme alınması gerekir. Kesinlik değeri, pozitif veya hasta olarak tahmin edilen verilerin kaç tanesinin gerçek sınıfının pozitif veya hasta olduğu bilgisini verir. Duyarlılık değeri, pozitif veya hasta olarak tahmin edilmesi gereken verilerin kaç adedinin pozitif veya hasta olarak tahmin edildiği bilgisini verir. F1 skoru ise kesinlik ve duyarlılık değerlerinin harmonik ortalamasını verir. F1 skorunun hesaplanması genellikle FN ve FP değerlerinin önemli olduğu durumlarda gereklidir.

Resnet50 ve VGG16 ağlarının doğruluk oranına bakıldığında diğer ağlarla yakın sonuçlar aldığı görülmektedir. Ancak Resnet50 ve VGG16'ya ait duyarlılık değeri, bu ağların OKT veri seti üzerinde diğer ağlara kıyasla iyi bir sınıflama yapamadığını göstermektedir. Eğitilen tüm ağlarda hasta olarak sınıflandırılan görüntülerin tamamına yakınının gerçek etiketinin de hasta olduğu görülmektedir. F1 skoruna bakıldığında ise bu veri seti üzerindeki en başarılı sınıflandırma yapan ağın Xception derin öğrenme ağı olduğu görülmektedir.

Tablo 1. Diğer popüler ağların ve Xception ağının hesaplanan parametreleri.

Ağlar	Doğruluk	Kesinlik	Duyarlılık	F1 Skoru
Resnet50	0.95	1.00	0.78	0.87
VGG16	0.95	1.00	0.81	0.89
InceptionV3	0.97	0.98	0.89	0.94
Xception	0.98	0.99	0.92	0.96

3. Sonuç

Bu çalışma Evrişimsel Sinir Ağları kullanılarak OKT görüntülerinden drusen tespiti yapılmasının mümkün ve yüksek doğrulukta olduğunu göstermektedir. Doğruluk oranı olarak ağlar birbirine yakın olmasına karşın kesinlik, duyarlılık ve F1 skoru parametreleri ağlar arasındaki performans farkının daha net görülmesini sağlamaktadır. Duyarlılık değeri, incelendiğinde Resnet50 ve VGG16 ağlarının OKT veri setini sınıflandırma performansının nispeten düşük olduğu görülmektedir. F1 skoru dikkate alındığında ise InceptionV3 ve Xception ağlarının kullanılan OKT veri seti üzerindeki sınıflandırma başarısının yüksek olduğu görülmektedir. Xception ağı 0.96'lık F1 skoru ile en başarılı sınıflandırmayı yapmıştır. Farklı sağlık kurumlarından çeşitli insanlara ait OKT görüntüleri kullanılarak veri setinin artırılması ile bu başarının daha da artacağı öngörülmektedir.

Kaynaklar

- [1] Rodríguez, IR, Larrayoz IM. Cholesterol oxidation in the retina: implications of 7kch formation in chronic inflammation and age-related macular degeneration. Journal of Lipid Research 2010; 51(10): 2847-2862.
- [2] Fine SL, Berger JW, Maguire MG, Ho AC. Age-related macular degeneration. New England Journal of Medicine 2000; 342(7): 483-492.
- [3] Wong WL, Su X, Li X, Cheung CMG, Klein R, Cheng CY, Wong TY. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. The Lancet Global Health 2014; 2(2): 106-116.
- [4] Gwairi OA, Thach L, Zheng W, Osman N, Little PJ. Cellular and molecular pathology of age-related macular degeneration: potential role for proteoglycans. Journal of Ophthalmology 2016.
- [5] İnan S. Retina anatomisi. Kocatepe Tıp Dergisi 2014; 15(3): 355-359
- [6] Nowak JZ. Age-related macular degeneration (AMD): pathogenesis and therapy. Pharmacological Reports 2006; 58(3): 353.

- [7] Age-Related Eye Disease Study Research Group. The age-related eye disease study system for classifying age-related macular degeneration from stereoscopic color fundus photographs: the age-related eye disease study report number 6. *American Journal of Ophthalmology* 2001; 132(5): 668-681.
- [8] ImageNet. <https://image-net.org>, (17.03.2021).
- [9] Image Classification on ImageNet. <https://paperswithcode.com/sota/image-classification-on-imagenet>, (17.03.2021).
- [10] Tsuji T, Hirose Y, Fujimori K, Hirose T, Oyama A, Saikawa Y, Shiraishi K, Kobayashi T ve diğerleri. Classification of optical coherence tomography images using a capsule network. *BMC Ophthalmology* 2020; 20(1): 1-9.
- [11] Retinal OCT Images. <https://www.kaggle.com/paultimothymooney/kermany2018>, (12.02.2021).
- [12] Chollet F. Xception: deep learning with depthwise separable convolutions. 2017 IEEE Conference on Computer Vision and Pattern Recognition; 22-25 July 2017; Honolulu, Hawaii, USA. pp. 1251-1258.
- [13] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 27-30 June 2016; Las Vegas, Nevada, USA. pp. 770-778.
- [14] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. <https://arxiv.org/abs/1409.1556>, (18.01.2021).
- [15] Szegedy C, Vanhoucke V, Ioffe V, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 27-30 June 2016; Las Vegas, Nevada, USA. pp. 2818-2826.