



Glokomun Renkli Fundus Fotoğrafı ile Teşhisinde Evrişimsel Yapay Sinir Ağları Yapılarının Rolünün Araştırılması

Investigation of the Role of Convolutional Neural Network Architectures in the Diagnosis of Glaucoma using Color Fundus Photography

✉ Eray Atalay*, ✉ Onur Özalp*, ✉ Özer Can Devocioğlu**, ✉ Hakika Erdoğan***, ✉ Türker İnce**, ✉ Nilgün Yıldırım*

*Eskişehir Osmangazi Üniversitesi Tıp Fakültesi, Göz Hastalıkları Anabilim Dalı, Eskişehir, Türkiye

**İzmir Ekonomi Üniversitesi Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği Anabilim Dalı, İzmir, Türkiye

***Maltepe Üniversitesi Tıp Fakültesi, Göz Hastalıkları Anabilim Dalı, İstanbul, Türkiye

Öz

Amaç: Evrişimsel yapay sinir ağları (EYSA) yapılarının glokomlu gözleri normal gözlerden ayırt etme performansının değerlendirilmesidir.

Gereç ve Yöntem: Eskişehir Osmangazi Üniversitesi Tıp Fakültesi Göz Hastalıkları Kliniği Arşivi'nde bulunan 5.388 hastaya ait 9.950 fundus fotoğrafı, farklı 3 deneyimli göz hekimi tarafından glokom, glokom şüpheli ve normal olarak değerlendirildi. Sınıflandırılan fundus fotoğrafları; son teknoloji yapay sinir ağı mimarileri olan 2-B evrişimli yapay sinir ağları ile değerlendirilerek derin artık ağlar (ResNet) ve çok derin sinir ağları (VGG) ile karşılaştırıldı. İki yüz otuz sekiz normal ve 320 glokomlu fundus fotoğraflarıyla oluşturulan veri setinde farklı algoritmalarla EYSA'nın glokomu saptamada doğruluğu, duyarlılığı ve özgüllüğü belirlendi. Glokom şüphesi tespiti için 170 normal, 170 glokomlu ve 167 glokom şüpheli fundus fotoğrafından veri seti oluşturularak ResNet-101 yapıları ile test edildi.

Bulgular: Glokomu saptamada doğruluk oranı, duyarlılık ve özgüllük; ResNet-50 ile %96,2, %99,5 ve %93,7, ResNet-101 ile %97,4, %97,8 ve %97,1, VGG-19 ile %98,9, %100 ve %98,1, 2-B EYSA ile %99,4, %100 ve %99 olarak saptandı. Glokom şüphelileri; normallerden ayırmada doğruluk oranı %62, duyarlılık %68 ve özgüllük %56; glokomlulardan ayırmada doğruluk oranı ise %92, duyarlılık %81 ve özgüllük %97 idi. Glokom EYSA ile 55 fotoğraf 2 sn'de test edilirken, klinisyenler tarafından bir fotoğrafın ortalama 24,2 sn'de değerlendirildiği görüldü.

Sonuç: Çalışmada en uygun şekilde tasarlanan ve eğitilen EYSA yöntemi az sayıda fundus fotoğrafıyla dahi yüksek doğrulukla glokomlu fundus fotoğraflarını normallerden ayırt edebilmiştir.

Anahtar Kelimeler: Glokom, evrişimsel yapay sinir ağları, yapay zeka, tele-tıp

Abstract

Objectives: To evaluate the performance of convolutional neural network (CNN) architectures to distinguish eyes with glaucoma from normal eyes.

Materials and Methods: A total of 9,950 fundus photographs of 5,388 patients from the database of Eskişehir Osmangazi University Faculty of Medicine Ophthalmology Clinic were labelled as glaucoma, glaucoma suspect, or normal by three different experienced ophthalmologists. The categorized fundus photographs were evaluated using a state-of-the-art two-dimensional CNN and compared

Yazışma Adresi/Address for Correspondence: Nilgün Yıldırım, Eskişehir Osmangazi Üniversitesi Tıp Fakültesi, Göz Hastalıkları Anabilim Dalı, Eskişehir, Türkiye

E-posta: nyildirim@yahoo.com **ORCID-ID:** orcid.org/0000-0001-6266-4951

Geliş Tarihi/Received: 10.09.2020 **Kabul Tarihi/Accepted:** 22.06.2021

Cite this article as: Atalay E, Özalp O, Devocioğlu ÖC, Erdoğan H, İnce T, Yıldırım N. Investigation of the Role of Convolutional Neural Network Architectures in the Diagnosis of Glaucoma using Color Fundus Photography. Turk J Ophthalmol 2022;52:193-200

with deep residual networks (ResNet) and very deep neural networks (VGG). The accuracy, sensitivity, and specificity of glaucoma detection with the different algorithms were evaluated using a dataset of 238 normal and 320 glaucomatous fundus photographs. For the detection of suspected glaucoma, ResNet-101 architectures were tested with a data set of 170 normal, 170 glaucoma, and 167 glaucoma-suspect fundus photographs.

Results: Accuracy, sensitivity, and specificity in detecting glaucoma were 96.2%, 99.5%, and 93.7% with ResNet-50; 97.4%, 97.8%, and 97.1% with ResNet-101; 98.9%, 100%, and 98.1% with VGG-19, and 99.4%, 100%, and 99% with the 2D CNN, respectively. Accuracy, sensitivity, and specificity values in distinguishing glaucoma suspects from normal eyes were 62%, 68%, and 56% and those for differentiating glaucoma from suspected glaucoma were 92%, 81%, and 97%, respectively. While 55 photographs could be evaluated in 2 seconds with CNN, a clinician spent an average of 24.2 seconds to evaluate a single photograph.

Conclusion: An appropriately designed and trained CNN was able to distinguish glaucoma with high accuracy even with a small number of fundus photographs.

Keywords: Glaucoma, convolutional neural network, artificial intelligence, telemedicine

Giriş

Glokom geri dönüşümsüz körlüğün en sık sebebidir.¹ Tüm dünyada 2010 yılında 60 milyon insanı etkileyen glokom hastalığının 2020 yılında 80 milyon insanı, 2040 yılında ise yaklaşık 112 milyon insanı etkileyeceği öngörülmektedir.^{2,3} Dünyadaki prevalansı 40-80 yaş aralığındaki bireylerde ortalama %3,54 (güven aralığı: %2,09-%5,82) olarak bildirilen glokom hastalığının ülkemizdeki prevalansı farklı çalışmalarda %1,29 ve %2 olarak rapor edilmiştir (Yıldırım, N., Başmak, H., Kalyoncu, C., Özer, A., Aslantaş, D. Metintaş, S. 2008: Eskişehir bölgesinde 40 yaş üzeri popülasyonda glokom prevalansı TOD 42. Ulusal Kongresi, Antalya).^{3,4} Tanısal cihazların teknolojik olarak gelişmesiyle birlikte glokom hastalığının teşhisi her ne kadar daha etkin hale gelmiş olsa da hastaların %50-90'ının glokomu olduğunu bilmediği rapor edilmiştir.^{2,5,6,7,8}

Glokomun teşhisinde göz içi basıncı ölçümü, optik sinir muayenesi, görme alanı muayenesi, retina sinir lifi analizi gibi birtakım yöntemler kullanılmaktadır. Dijital fundus görüntülemenin non-invazif, maliyet-etkin ve hızlı bir yöntem olması, optik ortam opasitelerinden daha az etkilenmesi ayrıca tetatıp uygulamaları açısından pratik bir yaklaşım olmasından dolayı fundus görüntüleme sıkça tercih edilmektedir. Son yıllarda popüler hale gelen yapay zeka uygulamaları ile fundus görüntülemenin kullanım alanı daha da genişlemiş ve birçok göz hastalığının yapay zeka yardımıyla teşhisinde fundus fotoğrafları yaygın olarak kullanılmaya başlanmıştır.

Bu çalışmada, glokomlu fundus fotoğraflarının normal fundus fotoğraflarından ayırt edilmesinde evrimsel yapay sinir ağları yapısının performansı değerlendirilmiştir.

Gereç ve Yöntem

Eskişehir Osmangazi Üniversitesi Tıp Fakültesi Göz Hastalıkları Anabilim Dalı Arşivi'nde bulunan ve glokom prevalansı araştırması sonucu elde edilen 5.388 hastaya ait 9.950 optik sinir fotoğrafı değerlendirmeye alındı. Çalışma protokolü Helsinki Bildirgesi'ne uygun olarak hazırlandı ve Eskişehir Osmangazi Üniversitesi Etik Kurul Komitesi'nden etik kurul onayı alındı. Algoritmada değerlendirilecek tüm fotoğraflar kliniğimizde bulunan non-midriatik fundus kamera (Kowa nonmyd alpha-DIII, Kowa Company Ltd., Tokyo, Japonya) sabit diskinden elde edilmiş olup cihaz arşivindeki tüm 20°'lik arka

segment fotoğrafları kimlik ve cinsiyet bilgilerinden arındırılarak dış ortama aktarılmıştır.

Fotoğraf kaliteleri ilgili hekimler (EA, HA, OÖ) tarafından yönergelere göre sınıflandırılıp veritabanına kaydedilmiştir.⁹ Düşük kalite, optik disk sınırlarının %50'sinden fazlasının görülebilmesi veya optik disk ve/veya çukurluk sınırlarını bloke eden bulanıklık; orta kalite, optik disk sınırlarının 0-%50 aralığında tam seçilememesi fakat çukurluk ve sınırların izlenebilmesi ile optik disk sınırlarını veya çukurluğunu bloke eden bulanıklık olmaması; yüksek kalite, optik disk ve çukurluk sınırlarının tamamının seçilmesi +/- retina sinir liflerinin görülebilmesi olarak sınıflandırılmıştır.

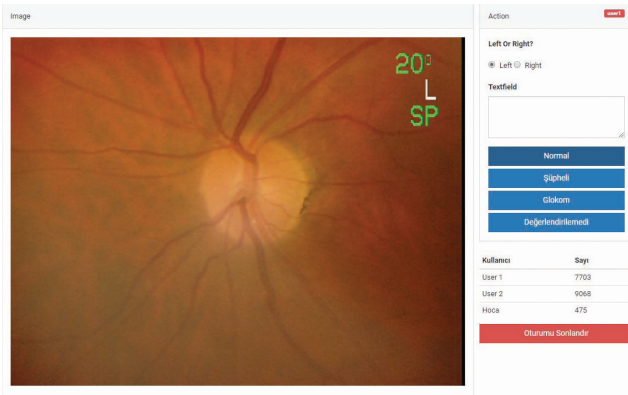
Toplam 9.950 fotoğraftan, 2.587 adet orta ve 5.970 adet yüksek kalite fotoğraf veri seti havuzunu oluşturmuş olup düşük kalitede olan 1.393 fotoğraf veri havuzuna dahil edilmemiştir. Çalışmada kullanılan tüm veri setleri, orta ve yüksek kalite fotoğraflardan oluşturulmuştur.

Fundus fotoğrafları klinisyenler tarafından kör ve rastgele bir şekilde bir yazılım yardımıyla normal, glokom şüphesi ve glokom olarak etiketlendi (Şekil 1). Bu etiketlemede daha önce yapılmış çalışmalarda kullanılan kriterler kullanıldı.^{9,10,11,12,13} Glokom; vertikal çukurluk/disk oranı $\geq 0,9$, rim genişliğinin disk çapına oranı $\leq 0,05$ olması veya lokalize çentiklenme, nöroretinal rimin incelendiği alanla veya lokalize çentiklenme bölgesi ile uyumlu retina sinir lifi defektinden herhangi birisinin bulunması olarak değerlendirildi. Glokom şüphesi ise; vertikal çukurluk/disk oranı $\geq 0,7$ ile $< 0,9$ aralığında olması, rim genişliğinin disk çapına oranının $\leq 0,1$ ile $> 0,05$ aralığında olması, retina sinir lifi defekti olması, disk hemorajisi olmasından herhangi birisinin bulunması olarak değerlendirildi. Bu özellikleri barındırmayan olgular normal olarak sınıflandırıldı.

Etiketleme iki aşamalı olarak gerçekleştirildi. İlk aşamada; 2 glokom uzmanının (EA ve HA) yaptığı değerlendirme neticesi elde edilen tanı kodları eğer 2 uzman arasında görüş birliği varsa kesin kanaat olarak veri tabanına aktarıldı. İhtilaf olan fotoğraflar işaretlenerek ikinci aşamaya geçildi. İkinci aşamada, ilk aşamada ihtilaf bildirilen fotoğraflar daha kıdemli ve deneyimli olan üçüncü uzmana (NY) gösterildi ve oy çokluğu ile kesin kanaata varıldı. Oy çokluğu sağlanmadığı durumda (örneğin; EA: şüpheli, HA: normal) üçüncü uzmanın kanaati (örneğin; NY: glokom), kesin kanaat olarak kayıt altına alındı. Ayrıca yazılımda fotoğrafın açılması ile uzmanın bir seçeneği işaretlemesi arasında geçen süre yazılım tarafından otomatik olarak hesaplandı.

Tüm fundus fotoğraflarının etiketlenmesi tamamlandıktan sonra evrişimsel yapay sinir ağları aşamasına geçildi.

Bu çalışmada, öznetelik çıkarma ve sınıflandırmayı tek bir öğrenme organında birleştirmek için 2-boyutlu (2-B) evrişimsel sinir ağları'nın (EYSA) uyarlanabilir bir uygulaması kullanılmıştır. EYSA, memeli görsel korteksinin basit hesaplama modelleri olarak kabul edilen beyin yapısından esinlenilen bir ileri beslemeli yapay sinir ağlarıdır.^{14,15,16} Bu nedenle, EYSA'lar çoğunlukla resimler ve videolar gibi 2-B sinyaller için kullanılır. Genellikle makine, görme toplulukları tarafından birçok görüntü ve video tanıma problemini hem anlarken hem de çözerken fiili standart olarak kullanılır. Bir evrişimi en basit şekilde anlamak için, bir matriste 2-B'de uygulanan bir sürgülü pencere işlevi olarak düşünülebilir. EYSA temel fikrini çok katmanlı algılayıcıların sınırlı bağlantı ve ağırlık paylaşımını kısıtlayarak alır. EYSA sadece alt örnekleme katmanlarının eklenen çok katmanlı algılayıcıların "kısıtlı" versiyonudur. EYSA'nın tamamen bağlı gizli ve çıkış katmanları çok katmanlı algılayıcıların katmanlarıyla tamamen aynıdır. Bu nedenle, EYSA karmaşık bir görevi öğrenirken çok katmanlı algılayıcıların



Şekil 1. Fundus fotoğraflarının etiketlendiği yazılım görüntüsü

karakteristik sınırlamalarına sahiptir. Çok katmanlı algılayıcılar gerçekten evrensel yaklaşımlardır; bununla birlikte, kompakt konfigürasyonları olan öğrenme görevleri için "daha iyi" evrensel bir yaklaşıma ihtiyaç vardır.

Konvansiyonel bir EYSA yapısının evrişimsel katmanları, ham veya ön-işlemeden geçmiş verilerden otomatik öznetelik öğrenimi ve çıkarımı için kullanır. Ardından sınıflandırma için çok katmanlı algılayıcı katmanlarını içerir ve bir boyutlu (1-B) öznetelik haritası çıktısı çok katmanlı bir algılayıcıya beslenir. Bu iki farklı katman, gelişmiş sınıflandırma performansı için en uygun öznetelikleri otomatik olarak öğrenmek üzere birleştirir. Girdi verileri, önceki katmanların işlemcileri (nöronları) tarafından filtrelenmiş ve alt örneklenmiş öznetelik haritalarını yayan dönüşümlü katmanlar ve aralarındaki alt örnekleme katmanları kullanılarak işlenir. EYSA'nın 2-B filtre çekirdekleri, hata geri-besleme algoritması (back-propagation) kullanılarak en iyilenir ve öğrenilir. Bu çalışmada glokom hastalığını saptamak için 2-B EYSA sınıflandırıcısı tasarlanmış ve uygulanmıştır. Önerilen sistem Şekil 2'de gösterilmektedir.

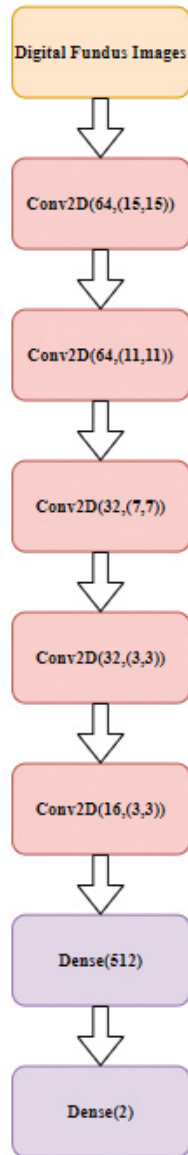
Önerilen 2-B EYSA sınıflandırıcısı, araştırma ve geliştiriciler için derin öğrenme modellerinde mümkün olduğunca hızlı ve kolay uygulama yapmak üzere geliştirilen Keras kütüphanesi ile Python kullanılarak uygulanmaktadır (Web sitesi: Keras Deep Learning Library Web Site: <https://keras.io/>). Önerilen 2-B EYSA sınıflandırıcısının yapısı, rekabetçi performans ile yüksek hesaplama verimliliği elde etmek için tüm deneylerde 5 evrişimsel ve iki skaler çok katmanlı sınıflandırıcı katmanından oluşur. Göreceli olarak sığ bir yapıda tasarlanan 2-B EYSA sınıflandırıcısı, gerçek zamanlı glokom tanısı için kullanılabilir. Evrişimsel katmanlardaki nöron sayısı [64 64 32 32 16] olarak ayarlanmıştır ve süzgeç boyutları (15,15), (11,11), (7,7), (3,3), (3,3) olarak belirlenmiştir. Tam olarak bağlanmış çok katmanlı algılayıcı katmanı 512 nörona ve çıkış katmanı glokomu tespit etmek için iki nörona sahiptir. 2-B sınıflandırıcısının mimarisi Şekil 3'te gösterilmektedir. Doğrusal olmayan bir aktivasyon fonksiyonu olarak relu, tüm EYSA katmanlarında



Şekil 2. Önerilen glokom tanı sistemi

kullanılmıştır ve maksimum örnekleme ile alt örnekleme katmanlarına uygulanmıştır. 2-B EYSA sınıflandırıcısının eğitimi için, genelleştirmeyi iyileştirmek için 10 katlı çapraz doğrulama tekniği uygulanmış ve bu sayede yapının fazla eğitimi engellenmiştir. Aynı amaçla ResNet yapısında Dropout (bazı bağlantıların belirli bir olasılıkla kaldırılması) tekniği uygulanmıştır. Tasarlanan yapının parametrelerinin optimizasyonu için eğitim içinde RMSprop algoritması kullanılmıştır.

Sınıflandırma sonrası veri setinden 238 sağlıklı - 320 glokom ile 170 sağlıklı - 170 glokom - 167 glokom şüpheli olarak



Şekil 3. Önerilen 2-B EYSA yapısı mimarisi [Conv2D (nöron sayısı, filtre büyüklüğü) and Dense (nöron sayısı)]

EYSA: Evrişimsel yapay sinir ağları

etiketlenmiş dijital fundus fotoğraflarıyla alt veri kümeleri oluşturuldu. Görüntülerin boyutu 512x512 piksele düşürülüp 10 katlı çapraz doğrulama ile eğitilip test edildi. Çalışmamızda kendimize ait veriler kullanılarak en uygun şekilde tasarlanmış ve eğitilmiş göreceli sığ (shallow) yapıda bir EYSA (CNN) ağ modeli (5 evrişimsel ve 1 saklı skaler katmana sahip) performansını karşılaştırmak amacıyla literatürdeki çalışmalarda sıklıkla kullanılan ResNet-50, ResNet-101 ve VGG-19 derin yapıları seçilerek elimizdeki aynı veriler ile transfer öğrenmesi kullanılarak eğitilmiştir.^{17,18,19} Burada amacımız büyük veriye (big data) sahip olunmadığı durumlarda (literatürdeki çalışmaların çoğunluğu bu durumdadır) kendi tasarlayıp eğittiğimiz göreceli sığ yapıdaki EYSA (CNN) modelinin glokom tanı performansının diğer çalışmalarda önerilmiş transfer öğrenmesi yöntemi ile eğitilmiş derin yapıların performanslarına benzer veya daha iyi olabileceğini göstermektir. Kendi verilerimizi kullanarak transfer öğrenme yöntemleri ile son 3 katmanı eğitilen derin yapılar, öncesinde yaklaşık 14 milyon resim içeren Imagenet veritabanı kullanılarak eğitilmiştir.²⁰ Tasarlanan yapıların eğitimi esnasında tüm veriler eğitim/doğrulama/genelleme amaçlı olarak rastgele 10 ayrı kümelere ayrılmıştır. Bu şekilde hem yapının doğru bir şekilde eğitilmesi hem de genelleme performansının en iyilenmesi amaçlanmıştır. Yapıların performans karşılaştırılması için standart performans ölçüm kriterleri (doğruluk, duyarlılık ve özgüllük) hesaplanmıştır. Isı haritaları ise son evrişimsel katmanın çıktılarının resim boyutuna üst örneklemeyle elde edilmiştir.

Glokom şüphesi tespiti için, 170 normal, 170 glokomlu ve 167 glokom şüpheli fundus fotoğrafından oluşan iki ayrı veri seti (glokom şüphesi-normal ve glokom şüphesi-glokom) oluşturuldu. Oluşturulan veri seti %90 eğitim ve %10 test olarak ayrılarak eğitilip test edildi. ImageNet veri kümesi ile daha önceden eğitilmiş olan 101 katmanlı ResNet yapıları ile transfer öğrenmesi metodu uygulandı.²⁰

Bu makalede bildirilen tüm deneyler, 8 GB RAM ve NVIDIA GeForce GTX 1050Ti grafik kartı ile 2.2 GHz Intel Core i7-8750H'de gerçekleştirilmiştir. Hem eğitim hem de test veri kümeleri toplam 768 CUDA çekirdeği tarafından paralel olarak işlenmiştir.

Bulgular

Çalışmaya dahil edilen toplam 5.388 olgunun 3.825'i kadın (%71), 1.563'ü erkek (%29) idi. Kadınların yaş ortalaması 54,88±10,31 yıl, erkeklerin 58,35±10,82 yıl olarak bulundu. Yaş bakımından iki grup arasında fark yoktu (p>0,05). Katılımcılardan 643'ünde (%11,9) diabetes mellitus, 1.734'ünde (%32,2) hipertansiyon, 505'inde (%9,4) koroner kalp hastalığı mevcut idi.

Orta ve yüksek kalitedeki tüm fotoğrafların glokom uzmanları tarafından etiketlenmesi sonucu 416 glokomlu ve 342 glokom şüpheli fotoğraf saptandı. Fotoğrafları etiketleyen EA ve HA arasındaki tanı uyumu %92 idi.

Glokomu saptamada doğruluk oranı, duyarlılık ve özgüllük; ResNet-50 ile %96,2, %99,5 ve %93,7, ResNet-101 ile %97,4,

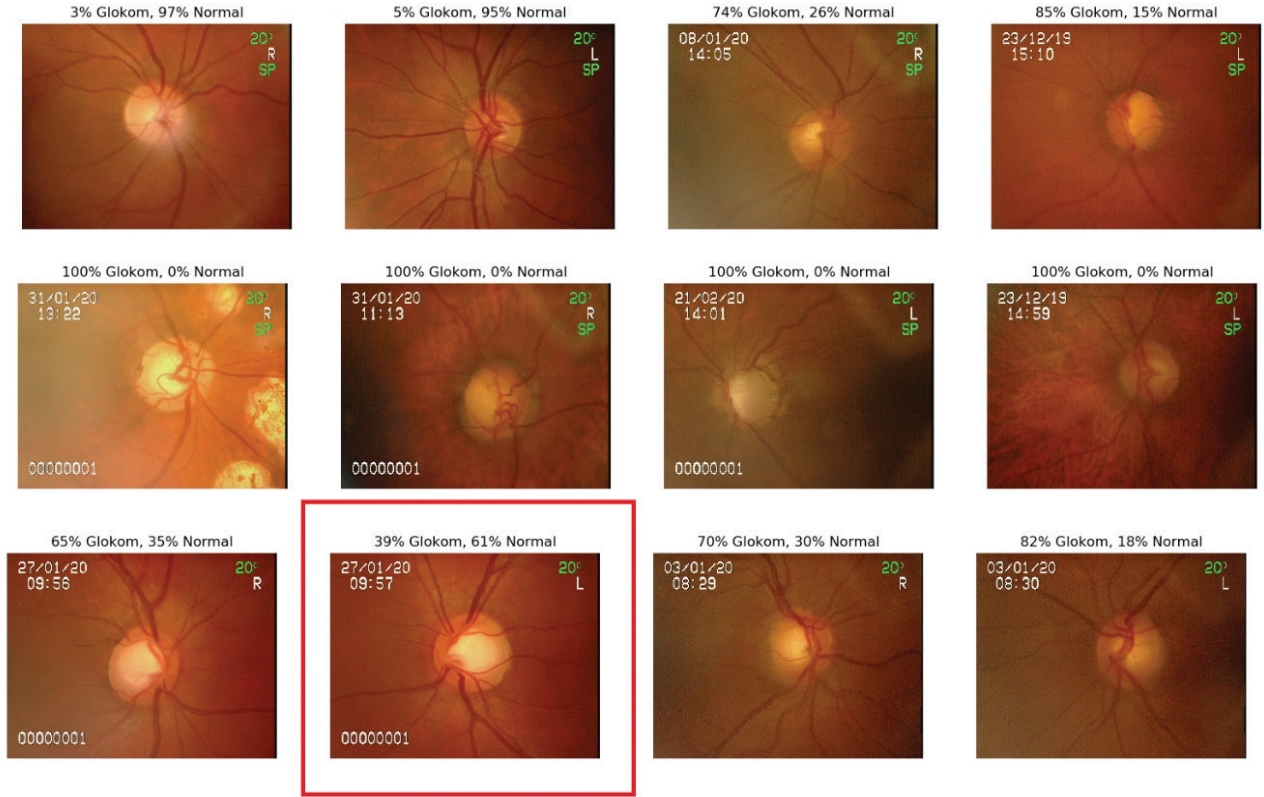
%97,8 ve %97,1, VGG-19 ile %98,9, %100 ve %98,1, 2-B EYSA ile %99,4, %100 ve %99 olarak saptandı. Şekil 4'te 2-B EYSA sınıflandırıcısının örnek verileri görülürken Şekil 5'te seçilmiş sonuçların ısı haritaları yer almaktadır.

ResNet-101 ile glokom şüpheli fundus fotoğraflarını normallerden ayırmada doğruluk oranı %62, duyarlılık %68 ve özgüllük %56 iken; glokomlulardan ayırmada doğruluk oranı %92, duyarlılık %81 ve özgüllük %97 idi.

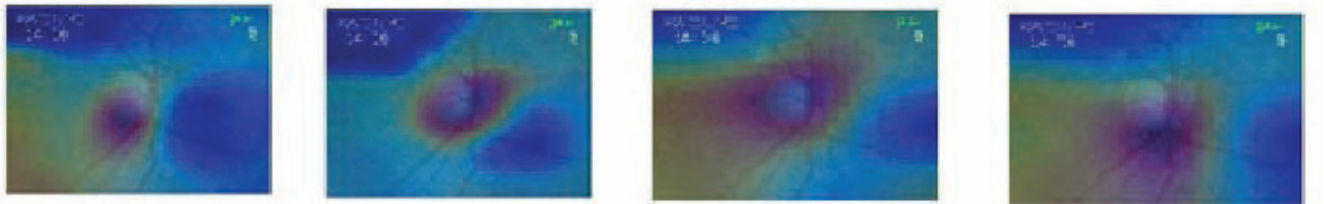
Glokom teşhisinde 10 epoch ile eğitilen sistemlerde, epoch başına geçen süre 2-B EYSA için 23,6 sn, VGG-19 için 23,2 sn, ResNet-50 için 35 sn ve ResNet-101 için 57 sn olarak saptandı. Ortalama test süreleri evrişimsel yapay sinir ağları ile 55 fotoğraf için 2 sn iken, klinisyenler tarafından bir fotoğrafın ortalama 24,2 sn'de değerlendirildiği görüldü.

Tartışma

Çalışmamızda glokomu saptamada doğruluk oranı, duyarlılık ve özgüllük; ResNet-50 ile %96,2, %99,5 ve %93,7, ResNet-101 ile %97,4, %97,8 ve %97,1, VGG-19 ile %98,9, %100 ve %98,1, 2-B EYSA ile %99,4, %100 ve %99 olarak saptandı. ResNet-101 ile glokom şüphelileri, normallerden ayırmada doğruluk oranı %62, duyarlılık %68 ve özgüllük %56; glokomlulardan ayırmada ise doğruluk oranı %92, duyarlılık %81 ve özgüllük %97 olarak saptandı. Glokom teşhisinde evrişimsel yapay sinir ağları ile 55 fundus fotoğrafı 2 sn'de test edilirken, klinisyenler tarafından bir fundus fotoğrafının ortalama 24,2 sn'de değerlendirildiği tespit edildi.



Şekil 4. 2-B evrişimli yapay sinir ağları sınıflandırıcısının örnek verileri (kırmızı işaret hatalı sonucu göstermektedir.)



Şekil 5. Seçilmiş sonuçların ısı haritaları

Fundus fotoğrafı ile görüntüleme optik sinirin yapısal değerlendirilmesinde en ucuz yöntemdir ancak glokom veya glokom şüphesini saptamada duyarlılık ve özgüllüğü optik koherens tomografi (OKT) gibi gelişmiş yöntemlerle karşılaştırılabilir değildir. Diyabetik retinopati tarama programlarından toplanan çok sayıda fundus görüntüsü, optik diskteki glokomatöz değişikliklerin derin öğrenme algoritmasıyla değerlendirilmesi için kaynak oluşturmuştur. Diyabetik retinopati taraması için çekilmiş olan fundus fotoğrafları kullanılarak glokom tespiti için derin öğrenme algoritmaları oluşturulan iki farklı çalışmada; duyarlılık %95,6 ve %96,4, özgüllük %92 ve %87,2 (eğri altında kalan alan [EAA]: 0,986 ve 0,942) gibi yüksek oranlar elde edilmiştir.^{9, 21} Fundus fotoğrafı ile glokomun ResNet yapısıyla değerlendirildiği başka bir çalışmada da EAA değeri 0,965 olarak bulunmuştur.²² Glokom tanısında kullanılan beş farklı ImageNet modelinin karşılaştırıldığı çalışmada duyarlılık ve özgüllükler VGG-16 ile %90,6 ve %88,2 (EAA: 0,96), VGG-19 ile %92,4 ve %88,5 (EAA: 0,97), InceptionV3 ile %92,2 ve %87,5 (EAA: 0,97), ResNet50 ile %91,1 ve %89,4 (EAA: 0,96) Xception ile %93,5 ve %85,8 (EAA: 0,96) olarak saptanmıştır.¹⁸ Literatürdeki sonuçlar ve çalışmamız göz önüne alındığında fundus fotoğrafları ile evrimsel yapay sinir ağları algoritmalarının oluşturulmasının glokom tespitinde etkili olduğu görülmektedir. Farklı evrimsel yapay sinir ağları modellerinin karşılaştırıldığı çalışmada glokom tespitinde duyarlılık ve özgüllükler yaklaşık olarak benzerken, bizim çalışmamızda VGG-19 ve 2-B EYSA ile daha yüksek duyarlılık ve özgüllük oranları saptandı.¹⁸

Glokomun derin öğrenme ile değerlendirilmesinde OKT ve ultra-geniş tarayıcı lazer oftalmoskopi (ultra-wide scanning laser ophthalmoscopy) ile çekilen görüntülerin kullanımı popülerlik kazanmaktadır. 1.399Optosfundusfotoğrafının glokom tespitinde kullanıldığı çalışmada ise duyarlılığın %81,3 ve özgüllüğün %80,2 (EAA: 0,872) olduğu, ileri düzeydeki glokomlu olguların tespitinde ise bu değerlerin arttığı bildirilmiştir.²³ 2.132 OKT görüntüsüyle yapılan derin öğrenme çalışmasında erken dönem glokom tespiti için duyarlılık %82,5 ve özgüllük %93,9 (EAA: 0,937) olarak saptanmıştır.²⁴ Farklı OKT parametrelerinin kullanıldığı diğer çalışmalarda da EAA değerlerinin 0,877 ile 0,981 aralığında değiştiği tespit edilmiştir.²⁵⁻²⁸ Glokom tanısında yapısal değerlendirmeler ile yapılan yapay zeka çalışmalarında konfokal tarayıcı lazer oftalmoskopi ("confocal scanning laser ophthalmoscopy", [CSLO]) parametreleri ve tarayıcı lazer polarimetri ("scanning laser polarimetry", [SLP]) parametreleri de çeşitli çalışmalarda kullanılmıştır. CSLO parametreleriyle yapılan taramalarda^{29,30,31,32,33} duyarlılık %83 ile %92 arasında, özgüllük %80 ile %91 arasında değişirken; SLP parametreleriyle yapılan taramalarda^{29,34} duyarlılık %74-%77, özgüllük %90-%92 olarak bulunmuştur. Makine öğrenmesiyle glokom tespitinde, optik diskin yapısal değerlendirilmesinde fundus fotoğrafı, OKT, CSLO ve SLP parametreleri gibi çeşitli veri kaynakları mevcut olup fundus fotoğraflarının kullanımının en ucuz ve kolay yöntem olması ayrıca yüksek duyarlılık ve

özgüllük oranlarına ulaşılması sebebiyle fundus fotoğrafları daha çok tercih edilebilir.

Literatürde yapısal testlerin birlikte değerlendirildiği çalışmalar da mevcuttur. Fundus fotoğrafları ve spektral domain-OKT'den elde edilen retinal sinir lifi kalınlık verilerinin birlikte kullanıldığı yapay zeka çalışmasında glokom teşhisinin, %80 özgüllük ve %90 duyarlılık ile yapıldığı bildirilmiştir.³⁵ Ayrıca glokomun tanısında; yapısal ve fonksiyonel test verilerinin kombinasyonuyla yapay zeka algoritmalarının etkinliğini artırmaya yönelik çalışmalar yapılmıştır. Fundus fotoğrafı ve görme alanı sonuçlarının ayrı ayrı değerlendirmelerine kıyasla birlikte değerlendirilmesinin glokom tanısında daha yüksek oranda (%88) doğruluk sağladığı rapor edilmiştir.³⁶

Son evrimsel katmanın çıktılarının resim boyutlarına üst örneklenmesiyle oluşan ısı haritaları, bir giriş görüntüsünün hangi bölgelerinin EYSA'nın çıktı tahminini etkilediğinin anlaşılmasına yardımcı olabilmektedir. Giriş görüntüsünün üzerine eklenen bu aktivasyon haritaları EYSA'nın en çok etkileşime girdiği bölgeleri vurgular. Glokom sınıfına ait örnek ısı haritaları Şekil 5'te verilmiştir. Glokomun tıbbi tanısında incelenen optik disk çevresindeki bozulmaların, EYSA'nın en çok etkileşime girdiği bölgeler olduğu görülmektedir. Hekimlerin glokomun veya başka hastalıkların teşhisinde, makine öğrenmesi çıktılarını yardımcı olarak kullanabilecekleri düşünülmektedir.

Derin öğrenme yöntemi, az sayıda fundus fotoğrafıyla dahi yüksek doğrulukla glokomlu gözleri normal gözlerden ayırt edebilmiştir. Daha fazla sayıda fotoğraf ve optimize edilmiş farklı öğrenme algoritmalarıyla daha yüksek duyarlılık ve özgüllük değerlerine ulaşılacağı gibi, fundus fotoğrafları ile OKT ya da görme alanı verilerinin kombinasyonu ile oluşturulan multi modelleme yöntemiyle de yapay zekanın glokom tanısında daha başarılı olabileceği öngörülmektedir. Çalışmamızda kullanılan fundus fotoğrafları, Eskişehir Osmangazi Üniversitesi Tıp Fakültesi Göz Hastalıkları Anabilim Dalı Arşivi'nde bulunan ve glokom prevalans araştırması sonucu elde edilen, tanıları etiketlenmemiş fotoğraflardır. Bu nedenle oluşturulan yapay zeka algoritması, farklı ırk ve etnik gruplarda kullanıldığında aynı doğruluğu göstermeyebilir. Ayrıca çalışmamızda glokom tanısında yalancı negatif ve pozitifliğin değerlendirilmemesi çalışmanın eksik yönlerini oluşturmaktadır.

Sonuç

Geri dönüşümsüz körlüğe yol açan glokomun, erken tanısı ve tedavisi önemlidir. Erken dönemlerde asemptomatik seyredabilen glokomda retinal ganglion hücrelerinde %20 ile %50 arasında kayıp olmadıkça görme alanı etkilenmemekte ve fonksiyonel testlerden ziyade yapısal değerlendirme daha çok önem kazanmaktadır. Glokomu erken dönemde tespit etmek, optik disk görüntülerinin makine öğrenmesi algoritmalarıyla değerlendirilmesiyle gerçekleştirilebilmektedir. Bu yöntemlerin geliştirilmesiyle birinci basamak sağlık merkezlerinde tele-tıp glokom tarama programları oluşturulabilecek ve kişilerin etkili biçimde triyajı sağlanabilecektir. Yapay zeka temelli algoritmalar hekimlerin yerine geçebilecek bir sistem olarak

değil hekimlere tanı ve takipte yardımcı birer araç olarak görülmelidir.

Etik

Etik Kurul Onayı: Çalışma protokolü Helsinki Bildirgesi'ne uygun olarak hazırlandı ve Eskişehir Osmangazi Üniversitesi Etik Kurul Komitesi'nden etik kurul onayı alındı.

Hasta Onayı: Alınmıştır.

Hakem Değerlendirmesi: Editörler kurulu ve editörler kurulu dışında olan kişiler tarafından değerlendirilmiştir.

Yazarlık Katkıları

Konsept: N.Y., E.A., T.İ., **Dizayn:** N.Y., E.A., T.İ., **Veri Toplama veya İşleme:** H.E., E.A., O.Ö.,

Analiz veya Yorumlama: E.A., N.Y., T.İ., Ö.D., **Literatür Arama:** E.A., O.Ö., Ö.D., **Yazan:** E.A., N.Y., O.Ö., T.İ., Ö.D.

Çıkar Çatışması: Yazarlar tarafından çıkar çatışması bildirilmemiştir.

Finansal Destek: Bu çalışma TÜBİTAK tarafından desteklenen 218E066 numaralı 1002 kodlu araştırma projesinden elde edilen sonuçlar ile yapılmıştır.

Kaynaklar

- Bourne RR, Stevens GA, White RA, Smith JL, Flaxman SR, Price H, Jonas JB, Keeffe J, Leasher J, Naidoo K, Pesudovs K, Resnikoff S, Taylor HR; Vision Loss Expert Group. Causes of vision loss worldwide, 1990-2010: a systematic analysis. *Lancet Glob Health*. 2013;1:e339-e349.
- Quigley HA, Broman AT. The number of people with glaucoma worldwide in 2010 and 2020. *Br J Ophthalmol*. 2006;90:262-267.
- Tham YC, Li X, Wong TY, Quigley HA, Aung T, Cheng CY. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology*. 2014;121:2081-2090.
- Yavaş GE, Küsbeci T, Şanlı M, Toprak D, Ermiş SS, İnan ÜÜ, Öztürk F. Risk Factors for Primary Open-Angle Glaucoma in Western Turkey. *Turk J Ophthalmol*. 2013;43:87-90.
- Bourne RR, Jonas JB, Flaxman SR, Keeffe J, Leasher J, Naidoo K, Parodi MB, Pesudovs K, Price H, White RA, Wong TY, Resnikoff S, Taylor HR; Vision Loss Expert Group of the Global Burden of Disease Study. Prevalence and causes of vision loss in high-income countries and in Eastern and Central Europe: 1990-2010. *Br J Ophthalmol*. 2014;98:629-638.
- Varma R, Ying-Lai M, Francis BA, Nguyen BB, Deneen J, Wilson MR, Azen SP; Los Angeles Latino Eye Study Group. Prevalence of open-angle glaucoma and ocular hypertension in Latinos: the Los Angeles Latino Eye Study. *Ophthalmology*. 2004;111:1439-1448.
- Baskaran M, Foo RC, Cheng CY, Narayanaswamy AK, Zheng YF, Wu R, Saw SM, Foster PJ, Wong TY, Aung T. The Prevalence and Types of Glaucoma in an Urban Chinese Population: The Singapore Chinese Eye Study. *JAMA Ophthalmol*. 2015;133:874-880.
- Gupta P, Zhao D, Guallar E, Ko E, Boland MV, Friedman DS. Prevalence of Glaucoma in the United States: The 2005-2008 National Health and Nutrition Examination Survey. *Invest Ophthalmol Vis Sci*. 2016;57:2905-2913.
- Li Z, He Y, Keel S, Meng W, Chang RT, He M. Efficacy of a Deep Learning System for Detecting Glaucomatous Optic Neuropathy Based on Color Fundus Photographs. *Ophthalmology*. 2018;125:1199-1206.
- Foster PJ, Buhrmann R, Quigley HA, Johnson GJ. The definition and classification of glaucoma in prevalence surveys. *Br J Ophthalmol*. 2002;86:238-242.
- Iwase A, Suzuki Y, Araie M, Yamamoto T, Abe H, Shirato S, Kuwayama Y, Mishima HK, Shimizu H, Tomita G, Inoue Y, Kitazawa Y, Tajimi Study Group, Japan Glaucoma Society. The prevalence of primary open-angle glaucoma in Japanese: the Tajimi Study. *Ophthalmology*. 2004;111:1641-1648.
- He M, Foster PJ, Ge J, Huang W, Zheng Y, Friedman DS, Lee PS, Khaw PT. Prevalence and clinical characteristics of glaucoma in adult Chinese: a population-based study in Liwan District, Guangzhou. *Invest Ophthalmol Vis Sci*. 2006;47(7):2782-2788.
- Topouzis F, Wilson MR, Harris A, Anastasopoulos E, Yu F, Mavroudis L, Pappas T, Koskosas A, Coleman AL. Prevalence of open-angle glaucoma in Greece: the Thessaloniki Eye Study. *Am J Ophthalmol*. 2007;144:511-519.
- Cireşan DC, Meier U, Gambardella LM, Schmidhuber J. Deep, big, simple neural nets for handwritten digit recognition. *Neural Comput*. 2010;22:3207-3220.
- Scherer D, Müller A, Behnke S. Evaluation of pooling operations in convolutional architectures for object recognition. *International conference on artificial neural networks*: Springer, 2010.
- Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*. 2017;60:84-90.
- Gómez-Valverde JJ, Antón A, Fatti G, Liefers B, Herranz A, Santos A, Sánchez CI, Ledesma-Carbayo MJ. Automatic glaucoma classification using color fundus images based on convolutional neural networks and transfer learning. *Biomed Opt Express*. 2019;10:892-913.
- Diaz-Pinto A, Morales S, Naranjo V, Köhler T, Mossi JM, Navea A. CNNs for automatic glaucoma assessment using fundus images: an extensive validation. *Biomed Eng Online*. 2019;18:29.
- Serener A, Serte S. Transfer learning for early and advanced glaucoma detection with convolutional neural networks. 2019 Medical Technologies Congress (TIPTeKNO): IEEE, 2019.
- Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L. Imagenet: A large-scale hierarchical image database. 2009 IEEE conference on computer vision and pattern recognition: Ieee, 2009.
- Ting DSW, Cheung CY, Lim G, Tan GSW, Quang ND, Gan A, Hamzah H, Garcia-Franco R, San Yeo IY, Lee SY, Wong EYM, Sabanayagam C, Baskaran M, Ibrahim F, Tan NC, Finkelstein EA, Lamoureux EL, Wong IY, Bressler NM, Sivaprasad S, Varma R, Jonas JB, He MG, Cheng CY, Cheung GCM, Aung T, Hsu W, Lee ML, Wong TY. Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes. *JAMA*. 2017;318:2211-2223.
- Shibata N, Tanito M, Mitsuhashi K, Fujino Y, Matsuura M, Murata H, Asaoka R. Development of a deep residual learning algorithm to screen for glaucoma from fundus photography. *Sci Rep*. 2018;8:14665.
- Masumoto H, Tabuchi H, Nakakura S, Ishitobi N, Miki M, Enno H. Deep-learning Classifier With an Ultrawide-field Scanning Laser Ophthalmoscope Detects Glaucoma Visual Field Severity. *J Glaucoma*. 2018;27:647-652.
- Asaoka R, Murata H, Hirasawa K, Fujino Y, Matsuura M, Miki A, Kanamoto T, Ikeda Y, Mori K, Iwase A, Shoji N, Inoue K, Yamagami J, Araie M. Using Deep Learning and Transfer Learning to Accurately Diagnose Early-Onset Glaucoma From Macular Optical Coherence Tomography Images. *Am J Ophthalmol*. 2019;198:136-145.
- Burgansky-Eliash Z, Wollstein G, Chu T, Ramsey JD, Glymour C, Noecker RJ, Ishikawa H, Schuman JS. Optical coherence tomography machine learning classifiers for glaucoma detection: a preliminary study. *Invest Ophthalmol Vis Sci*. 2005;46:4147-4152.
- Kim SJ, Cho KJ, Oh S. Development of machine learning models for diagnosis of glaucoma. *PLoS One*. 2017;12:e0177726.
- Barella KA, Costa VP, Gonçalves Vidotti V, Silva FR, Dias M, Gomi ES. Glaucoma Diagnostic Accuracy of Machine Learning Classifiers Using Retinal Nerve Fiber Layer and Optic Nerve Data from SD-OCT. *J Ophthalmol*. 2013;2013:789129.
- Christopher M, Belghith A, Weinreb RN, Bowd C, Goldbaum MH, Saunders LJ, Medeiros FA, Zangwill LM. Retinal Nerve Fiber Layer Features Identified by Unsupervised Machine Learning on Optical Coherence Tomography Scans Predict Glaucoma Progression. *Invest Ophthalmol Vis Sci*. 2018;59:2748-2756.
- Bowd C, Zangwill LM, Medeiros FA, Hao J, Chan K, Lee TW, Sejnowski TJ, Goldbaum MH, Sample PA, Crowston JG, Weinreb RN. Confocal

- scanning laser ophthalmoscopy classifiers and stereophotograph evaluation for prediction of visual field abnormalities in glaucoma-suspect eyes. *Invest Ophthalmol Vis Sci.* 2004;45:2255-2262.
30. Townsend KA, Wollstein G, Danks D, Sung KR, Ishikawa H, Kagemann L, Gabriele ML, Schuman JS. Heidelberg Retina Tomograph 3 machine learning classifiers for glaucoma detection. *Br J Ophthalmol.* 2008;92:814-818.
31. Zangwill LM, Chan K, Bowd C, Hao J, Lee TW, Weinreb RN, Sejnowski TJ, Goldbaum MH. Heidelberg retina tomograph measurements of the optic disc and parapapillary retina for detecting glaucoma analyzed by machine learning classifiers. *Invest Ophthalmol Vis Sci.* 2004;45:3144-3151.
32. Uchida H, Brigatti L, Caprioli J. Detection of structural damage from glaucoma with confocal laser image analysis. *Invest Ophthalmol Vis Sci.* 1996;37:2393-2401.
33. Adler W, Peters A, Lausen B. Comparison of classifiers applied to confocal scanning laser ophthalmoscopy data. *Methods Inf Med.* 2008;47:38-46.
34. Weinreb RN, Zangwill L, Berry CC, Bathija R, Sample PA. Detection of glaucoma with scanning laser polarimetry. *Arch Ophthalmol.* 1998;116:1583-1589.
35. Medeiros FA, Jammal AA, Thompson AC. From Machine to Machine: An OCT-Trained Deep Learning Algorithm for Objective Quantification of Glaucomatous Damage in Fundus Photographs. *Ophthalmology.* 2019;126:513-521.
36. Brigatti L, Hoffman D, Caprioli J. Neural networks to identify glaucoma with structural and functional measurements. *Am J Ophthalmol.* 1996;121:511-521.