نوع مقاله: پژوهشی

تشخیص دیابت چشمی با استفاده از شبکه‌های عصبی کانولوشنال عمیق

علی کارساز، دانشیار[۱]، صبورا محمدیان روشن، کارشناسی ارشد[۲]

۱- گروه مهندسی پزشکی و برق، موسسه آموزش عالی خراسان، مشهد، ایران، karsaz@khorasan.ac.ir
۲- گروه مهندسی پزشکی، موسسه آموزش عالی خراسان، مشهد، ایران، sa.m.roshan@khorasan.ac.ir

چکیده: دیابت چشمی به عنوان یکی از عوارض مهم دیابت، باعث آسیب به شبکیه چشم بیمار شده و تشخیص دیرهنگام آن حتی می‌تواند موجب نابینایی گردد. برخی از روش‌های دسته‌بندی مبتنی بر یادگیری ماشین بر اساس استخراج داده‌های تصاویر شبکیه به صورت دستی بوده و توسط متخصصین پردازش تصویر صورت می‌پذیرد. در سال‌های اخیر روشی جدید برای تشخیص و طبقه‌بندی تصاویر شبکیه چشم بدون نیاز به استخراج ویژگی‌های آن به‌صورت دستی مبتنی بر شبکه‌های عصبی کانولوشنال *(CNN)* ارائه شده است. در زمینه تشخیص و تصویربرداری پزشکی، به علت کمبود داده‌های طبقه‌بندی شده و زمان‌بر بودن فرآیند آموزش تا یک همگرایی مناسب، آموزش یک شبکه *CNN* از ابتدا دشوار بوده بنابراین یک روش متداول برای آموزش شبکه‌های *CNN* در حوزه پزشکی، بر اساس تنظیم مجدد شبکه‌های از پیش آموزش یافته، می‌باشد. برای این منظور در این مقاله، شبکه از پیش آموزش داده شده گوگل‌نت (*GoogLeNet*) به عنوان یکی از قوی‌ترین شبکه‌های عصبی کانولوشنال بر روی تصاویر شبکیه چشم بانک اطلاعات چشم پزشکی کگل *(Kaggle Database)* جهت تشخیص رتینوپاتی دیابتی اعمال می‌شود. همچنین جهت ارزیابی کلینیکی ساختار پیشنهادی، شبکه آموزش دیده جهت تشخیص دیابت چشمی بر روی ۱۰۱ تصویر شبکیه از کلینیک تخصصی چشم‌پزشکی نوید دیدگان با موفقیت اعمال گردید.

واژه‌های کلیدی: دیابت شبکیه، شبکه‌های عصبی کانولوشنال، شبکه گوگل‌نت، بانک اطلاعات چشم پزشکی کگل.

Deep Convolutional Neural Networks for Diabetic Retinopathy Screening

Ali Karsaz, Associate Professor[1], Saboora Mohammadian, M.Sc.[2]

1- Electrical and Bioelectric Engineering Department, Khorasan Institute of Higher Education, Mashhad, Iran, karsaz@khorasan.ac.ir
2- Bioelectric Engineering Department, Khorasan Institute of Higher Education, Mashhad, Iran, sa.m.roshan@khorasan.ac.ir

Abstract: Diabetic Retinopathy (DR) is one of the major complications of Diabetes, which is the injury to the retina of the diabetic patient and causes blindness if not diagnosed in early stages. Various machine learning classification and clustering approaches have been studied in literature with the purpose of improving the accuracy of the screening methods. Some of machine learning classification and clustering approaches are based on manually feature extraction of fundus images by image processing experts. In recent years, a new approach for image classification and diagnosis without using any manual feature extraction is proposed based on convolutional neural network (CNN). In medical imaging and diagnosis, training a deep CNN from scratch is difficult because it requires a large amount of labeled training data and the training procedure is a time consuming task to ensure proper convergence. Therefore, a very common method to train CNNs for medical diagnosis is fine-tuning a pre-trained CNN. In this paper, the pre-trained GoogleNet as a powerful CNN is employed on the Kaggle database for DR diagnosis from retinal images. To assess the efficacy of the clinical results, the proposed CNN algorithm is performed to diagnose DR from the images that are gathered from the the Navid-Didegan ophthalmology clinic.

Keywords: Diabetic retinopathy; Convolutional neural network (CNN), GoogleNet, Kaggle retinopathy database.

نام نویسنده مسئول: علی کارساز
نشانی نویسنده مسئول : : ایران - مشهد - موسسه آموزش عالی خراسان - گروه مهندسی پزشکی و برق.

Digital Object Identifier 10.22034/JASP.2021.45134.1135

۱- مقدمه

دیابت چشمی یا رتینوپاتی دیابتی[1] (*DR*) یکی از علل اصلی نابینایی افراد بزرگسال در سنین بین ۲۰ تا ۷۴ سالگی به شمار می‌رود. این بیماری زمانی اتفاق می‌افتد که بیماری دیابت باعث تخریب رگ‌های خونی در شبکیه چشم بیمار شده باشد. رتینوپاتی دیابتی تا زمانی که به مرحله خطرناکی نرسیده بدون علامت است بنابراین تشخیص این بیماری در مراحل اولیه نقش اساسی در درمان آن دارد [۲،۱]. پیش‌بینی‌ها از افزایش تعداد مبتلایان به بیماری *DR*، از ۱۲۶/۶ میلیون نفر در سال ۲۰۱۰ به ۱۹۱ میلیون نفر در سال ۲۰۳۰ حکایت دارد. هم‌چنین دربازه زمانی مذکور تعداد افراد مبتلا به *DR* که در معرض نابینایی قرار دارند از ۳۷/۳ میلیون نفر به ۵۶/۳ میلیون نفر خواهند رسید [۴،۳]. تشخیص رتینوپاتی دیابتی توسط متخصص و با استفاده از عکس‌های فوندوس[2] (عکس‌های شبکیه) صورت می‌گیرد. عکس‌های فوندوس شامل اطلاعات مفیدی از جمله وجود یا عدم وجود اگزودا[3]، میکروآنوریسم[4]، خونریزی[5] و رگ‌سازی مجدد[6] برای تشخیص رتینوپاتی می‌باشند. بنابراین پزشک متخصص براساس عکس شبکیه به بررسی وجود یا عدم وجود این نواحی در شبکیه چشم بیمار و در نهایت تشخیص بیماری *DR* می‌پردازد [۵-۷]. همواره دسترسی به کادر درمانی حرفه‌ای با هزینه‌ی زیاد روبرو بوده و نیز احتمال خطای انسانی در تشخیص مراحل اولیه‌ی این بیماری وجود دارد [۸]. مطالعات نشان می‌دهد که با گسترش روش‌های تشخیص اتوماتیک همانند یادگیری ماشین، می‌توان رتینوپاتی دیابتی را با دقت بالا و در مراحل اولیه و بدون حضور کادر درمانی متخصص تشخیص داد [۴]. در حال حاضر با افزایش اهمیت استفاده از الگوریتم‌های یادگیری ماشین و شبکه‌های عصبی برای تشخیص اتوماتیک بیماری‌ها، مطالعات بسیاری برای تشخیص زودهنگام رتینوپاتی دیابتی صورت گرفته است [۹]. برای تشخیص بیماری رتینوپاتی دیابتی از روی تصاویر شبکیه به کمک روش‌های متفاوت یادگیری، بایستی ویژگی‌های عکس‌ها به صورت دستی استخراج شوند. این کار با استفاده از روش‌های متفاوت پردازش تصویر و ارتباط با افراد متخصص در این زمینه صورت می‌گیرد. سپس ویژگی‌های استخراج شده مربوط به هر عکس به عنوان ورودی برای آموزش یک شبکه عصبی متداول[7] (*TNN*) و یا یک روش دسته‌بندی هوشمند اعمال می‌گردد. برخی از مطالعات از شناسایی میکروآنوریسم‌ها [۱۰،۹]، اگزوداها [۱۱-۱۴]، خونریزی‌ها [۱۵] و رگ‌سازی مجدد [۱۶] در تصاویر فوندوس برای تشخیص رتینوپاتی دیابتی بهره برده‌اند.

شبکه عصبی پرسپترون چندلایه (*MLP*) بر روی بانک اطلاعاتی *DIARETDB0* که یکی از معروف‌ترین مجموعه‌های اطلاعات تصاویر شبکیه چشم محسوب می شود، برای تشخیص *DR* اعمال شده است [۱۷]. ماشین بردار پشتیبان[8] (*SVM*) به عنوان یک روش توانمند در موضوع دسته‌بندی داده‌ها، و روش نایو بیز[9] (*NB*) برای غربال‌گری رتینوپاتی دیابتی مورد مطالعه قرار گرفته است [۱۸]. در پروژه‌ای دیگر روش نزدیکترین همسایگی *k* (*KNN*)[10] برای داده‌های بانک اطلاعاتی مسیدور[11] و داده‌های *DRIVE* مورد استفاده قرار گرفته است [۱۹]. استخراج ویژگی‌های از عکس‌های شبکیه چشم، پروسه‌ای زمان‌بر بوده و به متخصصین حرفه‌ای برای مطالعه دقیق عکس‌ها و اعمال ابزار استخراج ویژگی‌ها مناسب از روی تصاویر جهت آموزش الگوریتم‌های یادگیری نیاز دارد.

شبکه‌های عصبی کانولوشنال[12] (*CNN*) یکی از جدیدترین روش‌ها برای استخراج ویژگی‌های شبکیه چشم بدون نیاز به استخراج دستی آنها محسوب می‌گردد. در حال حاضر مطالعات بسیاری برای گسترش این شیوه نوین در زمینه‌های آنالیز تصاویر پزشکی و تشخیص بیماری‌ها در حال انجام است [۲۰]. مهم‌ترین مزیت استفاده از *CNN* توانایی آن در استخراج ویژگی‌های تصاویر به صورت خودکار با استفاده از مفهوم یادگیری عمیق[13] می‌باشد [۲۱]. به دلیل همین مزیت بسیار مهم اخیراً استفاده از *CNN*ها برای تشخیص بیماری‌ها در کاربردهای متفاوت پزشکی مورد توجه محققین قرار گرفته است. برای مثال، در مطالعه‌ای از *CNN* برای تشخیص تومورهای مغز از روی عکس‌های ام. ار. آی استفاده شده است [۲۲]. شبکه‌ی کانولوشنال عمیق چندگانه برای قسمت‌بندی پانکراس از روی اسکن‌های توموگرافی با موفقیت خوبی همراه بوده است [۲۳]. هم‌چنین در زمینه‌ی تشخیص رتینوپاتی دیابتی به‌وسیله‌ی *CNN*، ویژگی‌های تصاویر به کمک شبکه‌ی کانولوشنال استخراج گردیده و سپس به روش یادگیری گروهی، رگ‌های خونی شبکیه بخش‌بندی شده‌اند [۲۴]. از *CNN* برای تشخیص مراحل پیشرفت بیماری *DR* نیز استفاده می‌گردد [۲۵]. همچنین در یک مطالعه دیگر، مقایسه‌ای بین دو ساختار *CNN* با ابعاد و تعداد لایه‌های چندگانه برای تشخیص هرچه دقیقتر *DR* انجام شده است [۲۶].

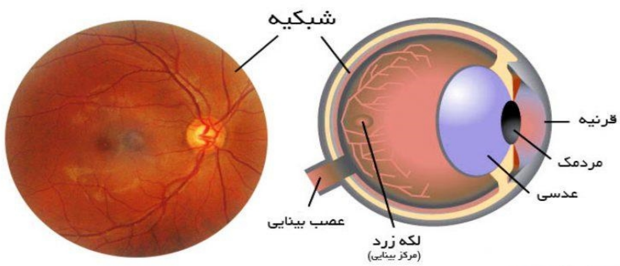
در این مقاله برای اولین بار یکی از معروف‌ترین و عمیق‌ترین شبکه‌های عصبی کانولوشنال را به شکل باز تنظیم لایه‌های انتهایی روی بزرگ‌ترین پایگاه اطلاعات تصاویر چشم پزشکی جهت غربال‌گری دیابت چشمی اعمال نموده و با استفاده از پیش‌پردازش مناسب تصاویر و نیز افزایش تعداد تصاویر افراد دارای دیابت چشمی و ایجاد توازن در نمونه‌های دارای بیماری با افراد سالم به بهبودهای قابل قبولی در معیارهای ارزیابی و نیز زمان پردازش و آموزش دست یافته است. انتخاب مناسب تعداد لایه‌های بازتنظیم شونده و اتخاذ روش مناسب بهینه‌سازی و نیز بکارگیری توابع فعالیت مناسب برای افزایش دقت همراه با حفظ سرعت آموزش، از جمله دیگر نوآوری‌های این مقاله جهت بهبود نتایج، محسوب می‌گردد.. همچنین جهت نمایش کارکرد بالینی روش پیشنهادی، شبیه سازی روی تصاویر مرکز چشم‌پزشکی نوید دیدگان نیز صورت پذیرفت که نتایج در

این بخش با نتایج یک شـــبکه گوگل‌نت پایه مقایســـه و بهبود قابل ملاحظه‌ای حاصل گردید. بر این اساس در این مقاله و در بخش دوم پس از بررســی کوتاه در خصـوص عوارض بیماری دیابت چشــمی، به تشــریح ساختار شبکه عصبی کانولو شنال در بخش سوم پرداخته می شود. در بخش چهارم، شـبکه عصـبی کانولوشـنال گوگل‌نت با اسـتفاده از تصـاویر شبکیه چشم باز تنظیم می‌گردد. دو بانک اطلاعات تصاویر چشم پزشکی برای بازتنظیم و ارزیابی شبکه مورد نظر در این مقاله استفاده شده است اولین بانک اطلاعات چشـــم پزشـــکی که یک بانک داده عمومی و در دســـترس محســـوب می‌شـــود بانک اطلاعات داده کگل[14] بوده پس از بازتنظیم و ارزیابی نتایج با تصـــاویر این پایگاه داده از یک بانک تصـــاویر خصـــوصـــی مربوط به مرکز چشـــم پزشـــکی نوید دیدگان برای ارزیابی توانمندی غربال‌گری شـبکه آموزش یافته اسـتفاده خواهد شـد. شـبکه عصـبی کانولوشـنال مورد اسـتفاده، مربوط به جدیدترین نسـخه شـبکه گوگل‌نت *Inception-V3* بوده که بر روی داده‌های بانک تصاویر *ImageNet* آموزش دیده اســت تمامی الگوریتم‌های مورد نظر در محیط نرم‌افزاری پایتون[15] و به‌وســـیله‌ی پکیج *Numpy* که برای محاســـبات آرایه‌های چندبعدی بسـیار قوی، پرسـرعت و مناسـب می‌باشـد اجرا گردیده اسـت. درنهایت در بخش پنجم نتیجه‌گیری و پیشنهادات ارائه می‌شود.

۲- دیابت چشمی و عوارض آن

از جمله عوارض مهم دیابت می‌توان به موارد کتواســـیدوز دیابتی[16]، هیپوگلیســـمی[17]، نفروپاتی دیابتی[18] و نوروپاتی دیابتی[19] اشـــاره نمود. هیپوگلیسمی هنگامی که غلظت قند خون به ۷۰ میلی‌گرم بر دسی‌لیتر یا کمتر بر سد، نامیده می شود. هیپوگلیسمی ممکن است به علت مصرف زیاد انسـولین یا قرص‌های پایین‌آورنده‌ی قند خون، کم خوردن وعده غذایی و فعالیت بدنی زیاد در بدن ایجاد شـود. کتواسـیدوز دیابتی دراثر کمبود نسـبی یا مطلق انسـولین و افزایش سـطح هورمون‌های مخالف آن (کتوکولامین‌ها، گلوکاگون، هورمون رشد و کورتیزول) ایجاد می شود. در سـیر بیماری دیابت و عدم کنترل مطلوب آن، احتمال آسـیب به کلیه‌ها زیاد اسـت به طوری که قادر به تصـفیه کامل خون نبوده و مواد زاید در بدن تجمع پیدا می‌کنند و در صورت عدم درمان مناسب، کلیه‌ها به سمت نار سایی پیش می‌روند که به آن نفروپاتی دیابتی گفته می شود. نوروپاتی دیابتی عارضه شایع دیگر دیابت به دلیل آسیب‌دیدگی‌های عصبی بوده که در آین حالت بیماری دچار فقدان حس در پاها می‌شـــود رشـــد این نوع بیماری منجر به آسیب دیدن پا شده از این رو به آن پای دیابتی نیز گفته می‌شود [۲۷].

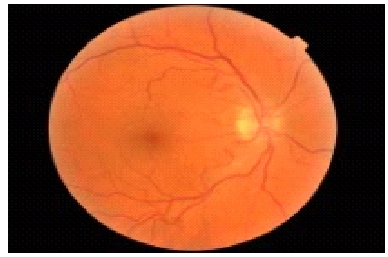
در بیماری دیابت با افزایش قند خون، رگ‌های بســـیار ریز داخل شبکیه تخریب می شوند و این خود باعث نشت و تجمع خون و توده‌های چربی ریز در شـــبکیه می‌گردد. همچنین به دلیل از بین رفتن رگ‌های خونی که وظیفه خون‌رسانی به محیط شبکیه را بر عهده داشتند، رگ‌های جدید و غیرطبیعی در شـبکیه تشـکیل می‌شـود. رتینوپاتی دیابتی در دو نوع کلی تکثیری[20] (*PDR*) و غیرتکثیری *(NoDR)* طبقه‌بندی می‌شـــود. رتینوپاتی غیرتکثیری یا زمینه‌ای اولین مرحله *DR* می‌باشـــد که در آن رگ‌های تخریب شـده شـروع به نشـت خون و چربی در شـبکیه می‌کنند [۲۹،۲۸]. *PDR* اثرات جدی‌تری نســـبت به *NoDR* برای بیمار ایجاد می‌نماید به دلیل اینکه به سرعت وارد مرحله‌ی پیشرفته شده و بینایی را کاهش می‌دهد [۲۹]. شکل (۱) ساختار چشم به همراه نمایشی از یک تصویر شبکیه چشم را نشان می‌دهد.



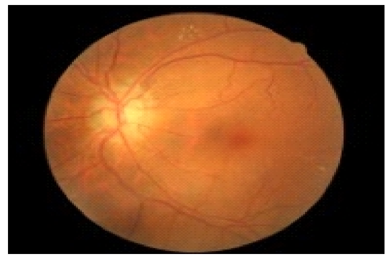
شکل ۱ ساختار چشم به همراه تصویر شبکیه چشم

شـــکل (۲) چند نمونه از تصـــاویر شـــبکیه چشـــم مربوط به بانک اطلاعات چشم پزشکی کگل، یک فرد مبتلا به بیماری دیابت چشمی را در مقایسه با شبکیه چشم یک فرد سالم نشان می‌دهد. همچنین در این شـــکل یک تصـــویر شـــبکیه مبتلا به *DR* از نوع غیرتکثیری *(NoPDR)* (شکل ۲ ب) در مقایسه با تصویر شبکیه یک شخص مبتلا به *DR* تکثیری یا *PDR* (شـکل ۲ ج) را نشـان می‌دهد. عوارض مختلفی از بیماری دیابت بر روی تصاویر گرفته شده از چشم شامل میکروآنوریسم‌ها، خونریزی‌ها، اگزوداها، ایسکیمی شبکیه و غیره قابل مشاهده است میکروآنوریسم‌ها به معنای ورم‌های کوچک در رگ‌های خونی بوده که گاهی اوقات باعث نشت چربی شده و به صورت نقاط تیره‌رنگ در تصاویر شبکیه قابل شناسایی می‌باشند. کاهش خونرسانی به سلول‌ها در شبکیه موجب مرگ سلول‌های محافظ شبکیه، ضعف دیواره مویرگ‌ها و مرگ سلول‌های اطراف مویرگ‌ها می‌شـــود. این تغییرات موجب می‌شـــوند که قســـمت‌هایی از دیواره مویرگ‌های شـــبکیه به صـــورت برآمده تغییر شـــکل دهند. این تغییرات ساختاری در دیواره عروق، میکروآنوریسم نام دارند. از سوی دیگر عروق ایجاد شده در پروسه رگ‌سازی مجدد، بسیار شکننده هستند و موجب خونریزی‌های متعدد می‌شوند. که به‌صورت نقاط کوچک خونی در شبکیه قابل روئیت هستند. اگزوداها یا ادم ماکولا[21] در هر مرحله ای از رتینوپاتی دیابتی می‌تواند ایجاد شـود و به معنای افزایش ضـخامت شـبکیه و تجمع آب میان بافتی در ماکولا است. آسیب به عروق کوچک در بیماران دیابتی منجر به افزایش نفوذپذیری مویرگ‌ها می‌شـــود. در نتیجه آب و مواد محلول یا نامحلول از مویرگ‌ها خارج و در شـــبکیه تجمع می‌یابند و ...

ضخامت شبکیه را بی شتر می‌کنند. پس از جذب دوباره بخش اعظم این مواد توسط مویرگ‌ها، قسمتی از نامحلول‌ترین آن‌ها در شبکیه باقی می‌مانند و اگزودا را تشکیل می‌دهند.



الف)

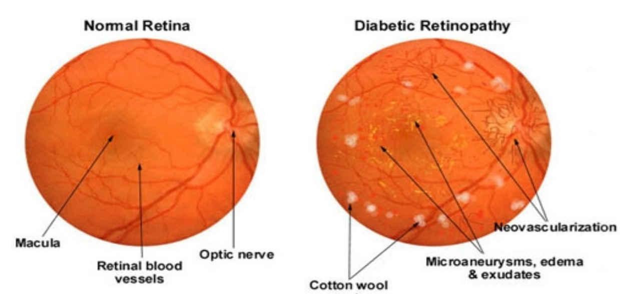


ب)

ج)

شکل ۲ نمونه تصاویر شبکیه بانک اطلاعات چشم پزشکی کگل الف) شبکیه چشم سالم، ب) شبکیه دیابت چشمی غیر تکثیری د) دیابت چشمی تکثیری

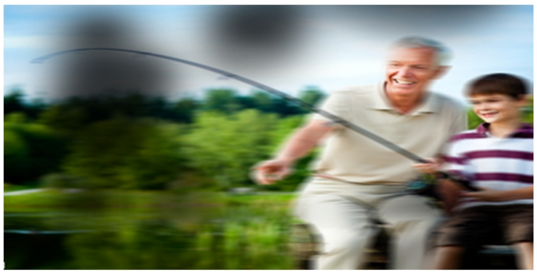
ادم ماکولا با معاینه تخصصی فوندوس، آنژیوگرافی[۲۲] تشخیص داده می شود. شکل (۳) (سمت راست) وجود میکروآنوریسم‌ها، خونریزی‌ها و اگزوداها را برروی تصویر شبکیه یک فرد بیمار نشان می‌دهد. در حالت ایسکیمی شبکیه، رگ‌های خونی بسته می‌شود و ماکولا به اندازه‌ی کافی خون دریافت نمی‌کند [۲۹].



شکل ۳ عوارض میکروآنوریسم‌ها، خونریزی‌ها، اگزوداها ناشی از دیابت شبکیه

شکل (۴) دید یک فرد سالم را در مقایسه با دید یک فرد مبتلا به رتینوپاتی دیابتی نشان می‌دهد.

الف)



ب)

شکل۴ الف) دید شخص با چشم سالم، ب) دید شخص مبتلا به بیماری دیابت چشمی

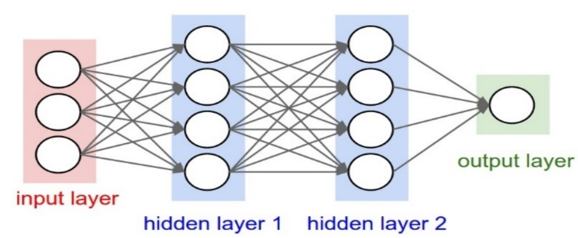
۳- شبکه‌های عصبی کانولوشنال

شبکه‌های ع صبی کانولو شنال *(CNN)* مدلی از شبکه‌های ع صبی م صنوعی بوده که همانند شبکه‌های ع صبی معمولی *(TNN)* از نورون‌ها، لایه‌ها و وزن‌ها تشکیل شده است. مهم‌ترین تفاوت *CNN*ها با *TNN*ها، در قابلیت ویژه آنها در یادگیری‌های عمیق مانند ت شخیص ت صاویر، صداها و دست‌خط‌های متفاوت نهفته است. لذا ورودی این شبکه‌ها می‌تواند شامل ماتریس‌هایی با ابعاد بالاتر باشد. مطالعه‌ای که در سال ۱۹۶۸ صورت گرفت نشان داد که قشر بینایی مغز برای پردازش اطلاعات تصاویر از الگوی پیچیده‌ای استفاده می‌نماید [۳۰]. نواحی ادراکی که قشر بینایی در آن قرار دارد، همانند فیلترهای محلی بر روی اطلاعات تصویر اعمال می شود سلول‌های ساده‌تر برای ت شخیص ویژگی‌های سطح پائین‌تر در نواحی ادراکی همچون لبه‌ها کاربرد دارند، همچنین سلول‌های پیچیده قادر به تشخیص ویژگی‌های مهم‌تر و اختصاصی‌تر و در سطح بالاتر می‌باشند. تشخیص ویژگی‌های اختصاصی‌تر نتیجه و ترکیبی از ویژگی‌های سطح پایین می‌باشند. این عملکرد مغز الهام بخش شبکه‌های عمیق عصبی امروزی شده است. مفهوم شبکه کانولوشنال برای اولین بار در سال ۱۹۸۰ تو سط فکو شیما[۲۳] مطرح گردید [۳۱]. در سال ۱۹۹۸ شبکه‌های کانولو شنال امروزی به منظور استفاده در شبکه‌های عصبی بررسی شدند [۳۲]. اما به دلیل نیاز به سخت‌افزارها و پردازش‌گرهای گرافیکی قوی ا ستفاده از این شبکه برای ت شخیص ت صاویر تا سال ۲۰۱۲ که به شکل

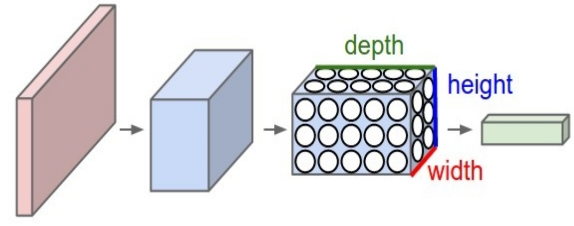
اختصـاصـی برای تشـخیص تصـاویر ارائه و معروف گردید به تعویق افتاد [۳۲].

۳-۱- ساختار شبکه عصبی کانولوشنال

شبکه عصبی کانولوشنال مدلی از شبکه‌های عصبی چندلایه هستند که معمولاً شـامل لایه‌های کانولوشـنال، تلفیق[۲۴] و تمام متصـل[۲۵] *(FC)* می‌باشند [۲۴]. در یک *TNN*، شبکه ورودی را دریافت و آن را به تعدادی لایه مخفی می‌فرستد. هر لایه مخفی از تعدادی نورون به شکل یک بردار تشـکیل شـده اسـت که هر نورون به طور کامل به تمامی نورون‌های لایه قبلی خود متصـل می‌باشـد نورون‌ها در هر لایه به‌طور مسـتقل عمل می‌کنند و در ارتباطات و وزن‌ها با یکدیگر شریک نمی‌شوند. لایه آخر این شبکه، لایه خروجی بوده که مسئول نهایی تعیین کلاس داده‌ها است. اما در *CNN*ها، ورودی یک ماتریس چند لایه یا دارای عمق که به آن تنسـور نیز گفته می شود، مانند یک تصویر است و لایه‌های این شبکه بر خلاف *TNN*ها، متشـکل از نورون‌هایی در سـه بعد هسـتند. در این شـبکه‌ها نورون‌های یک لایه تنها به ناحیه کوچکی از لایه قبلی متصل می‌باشند و به تمامی نورون‌های لایه قبلی متصـل نیسـتند. شـکل (۵) تفاوت شـبکه عصبی کانولوشنال با شبکه عصبی معمولی را نشان می‌دهد. همانطور که در شـکل مشـخص اسـت تمامی نورون‌ها در یک *TNN* به هم متصـل می‌باشند اما در یک *CNN* نورون‌ها در سه بعد قرار گرفته‌اند. در شکل (۵-ب)، لایه‌ی ورودی که با رنگ قرمز مشخص شده است عکس ورودی به شـبکه اسـت، و دارای طول، عرض و عمق بوده و برای تصـاویر با فرمت *RGB* معرف سه کانال قرمز، سبز و آبی تصویر می‌باشد [۳۳].



الف)



ب)

شکل۵ مقایسه شبکه عصبی معمولی (*TNN*) و شبکه عصبی کانولوشنال (*CNN*)، الف) شبکه عصبی معمولی، ب) شبکه عصبی کانولوشنال

۳-۱-۱- لایه کانولوشنال

این لایه به‌عنوان اصلی‌ترین لایه در این دسته از شبکه‌های عصبی محسـوب شـده به همین دلیل، آنها را به عنوان شـبکه‌های عصـبی کانولوشنال می‌شناسند. وظیفه این لایه استخراج ویژگی‌ها می‌باشد. این لایه عملیات کانولوشـنال را بر روی داده ورودی اعمال می‌کند و خروجی‌هایی به نام نقشـه ویژگی[۲۶] از این لایه به‌دسـت می‌آید. در نتیجه تمامی نورون‌ها در یک نقشـه ویژگی، مجموعه‌ای از وزن‌ها و بایاس‌های مشـابه و مشـترکی دارند که باعث می‌شـود که ویژگی‌های تصـویر در موقعیت‌های مختلف قابل شـناسـایی باشـند. در شـبکه‌های کانولوشـنال اتصالات به صورت نواحی کوچک و محلی صورت می‌گیرد. به بیان دیگر هر نورون در اولین لایه مخفی به ناحیه کوچکی از نورون‌های ورودی متصل می‌شود. این ناحیه کوچک به عنوان ناحیه ادراکی محلی[۲۷] یا کرنل کانولوشـنال[۲۸] نامیده می‌شـود [۳۴]. به بیان دیگر خروجی نورون لایه کانولوشـنال $y_{w,h,m}$ در طول و عرض w، h و عمق m به صـورت رابطه (۱) قابل بیان است.

$$y_{w,h,m} = f\left(\sum_{i=(w-1)S+1}^{(w-1)S+K}\sum_{j=(h-1)S+1}^{(h-1)S+K}\sum_{k=1}^{N} W_{k,m}x_{i,j,k}+b_m\right) \quad (۱)$$

که در این رابطه f تابع فعالیت[۲۹]، b_m بایاس مشـترک نورون‌ها، $W_{k,m}$ وزن‌های ۵×۵ مشترک نورون‌ها و نهایتا $x_{i,j,k}$ ورودی در موقعیت i، j و k می‌باشد. بکارگیری توابع فعالیت در لایه کانولوشنال، باعث افزایش خصـوصـیات غیرخطی در خروجی می‌شـوند. توابع فعالیت انواع متفاوتی دارند. چند نمونه از معروف‌ترین توابع فعالیت، تابع سـیگموئید[۳۰]، تابع تانژانت هایپربولیک *(tanh)*، تابع یکسو ساز خطی[۳۱] *(ReLU)*، تابع خطی نمای[۳۲] *(ELU)* و *Softmax* اسـت. بنابراین تمامی نورون‌های واقع در لایه مخفی اول دقیقاً ویژگی مشـابهی را در نواحی مختلف تصـویر شـناسـایی می‌نمایند. در نهایت خروجی لایه ورودی یا نورون‌های لایه مخفی به نام نقشـه ویژگی شـناخته می‌شـوند. ابعاد مربوط به ماتریس خروجی لایه کانولوشنال $W_2 \times H_2 \times D_2$ که از ماتریس ورودی با ابعاد $W_1 \times H_1 \times D_1$ به صورت روابط (۲)، (۳) و (۴) به‌دست می‌آید:

$$W_2 = \frac{W_1 - F + 2P}{S+1} \quad (۲)$$

$$H_2 = \frac{H_1 - F + 2P}{S+1} \quad (۳)$$

$$D_2 = K \quad (۴)$$

در این روابط، F، P، S و K به‌ترتیب نشـان‌دهنده اندازه کرنل، مقدار *zero padding* اندازه‌ی گام و تعداد فیلترها می‌باشـد. از این روابط مشـخص اسـت که به‌ازای هر فیلتر $F \times F \times D_1$ وزن داریم و با توجه به تعداد K فیلتر موجود، در مجموع $(F \times F \times D_1)K$ وزن و K بایاس ایجاد می‌شود. بنابراین همانطور که مشخص است تعداد پارامترهایی که شبکه در یک لایه کانولوشنال آموزش می‌بینند بسیار زیاد هستند.

۳-۱-۲-لایه تلفیق[۳۳]:

در یک شـــبکه عصـــبی کانولوشـــنال به‌طور معمول پس از هر لایه کانولوشـــنال یک لایه تلفیق قرار می‌گیرد. این لایه از آن جهت اهمیت دارد که باعث کاهش تعداد پارامترهایی می‌شـــود که بایســـتی آموزش ببینند. بنابراین با بکارگیری این لایه ضمن کاهش محاسبات مورد نیاز در بخش آموزش، باعث کنترل فرایادگیری[۳۴] احتمالی در شبکه نیز می‌گردد. این لایه بر روی هر برش از لایه‌های عمق ورودی اعمال می‌شـود و اندازه آن را تغییر می‌دهد. دو تابع عملکردی معروف این لایه، شـــامل تلفیق بیشـــینه[۳۵] و میانگین[۳۶] نام دارد، که مدل اول دارای کاربرد بیشـــتری در CNNها است. طریقه عمل تلفیق بیشـینه، بدین صـورت اسـت که در هر پنجره بزرگترین پیکســـل را به خروجی می‌فرســـتد. این پنجره بر روی تصویر همانند تابع کانولوشنال از چپ به راست و از بالا به پایین با اندازه گام‌های مشخص شده حرکت می‌کند و نتیجه را به خروجی می‌فرستد. به دلیل اینکه این عملیات بر روی تمامی برش‌ها اعمال می‌گردد، عمق خروجی همان عمق ورودی به لایه می‌باشـــد و تغییری نمی‌کند. رابطه مربوط به تنســـور خروجی با ابعاد $W_3 \times H_3 \times D_3$ که از تنســـور ورودی مرحله کانولوشـــنال با ابعاد $W_2 \times H_2 \times D_2$ به‌دســـت می‌آید به صـــورت روابط ذیل قابل بیان است.

$$W_3 = \frac{W_2 - F_p}{S_p + 1} \quad (۵)$$

$$H_3 = \frac{H_2 - F_p}{S_p + 1} \quad (۶)$$

$$D_3 = D_2 \quad (۷)$$

در روابط فوق، F_p و S_p به‌ترتیب اندازه کرنل و گام لایه تلفیق می‌باشند.

۳-۱-۳- لایه‌کاملاً متصل:

لایه کاملاً متصل، لایه آخر یک شبکه عصبی کانولوشنال محسوب می شود و مشابه شبکه عصبی معمولی اتصالات کاملی با خروجی لایه قبلی ایجاد می‌کند. این لایه ورودی را دریافت می‌کند و سپس خروجی را به‌صـــورت برداری با N مولفه تولید می‌کند. تعداد کلاس‌هایی که شـــبکه مورد نظر آن را طبقه‌بندی می‌کند برابر با N است. بنابراین خروجی شبکه طراحی شده برای تصویر ورودی به شبکه، برداری با N مولفه عددی بوده، که هر عدد درصـــد احتمال تعلق به کلاس مورد نظر را نشـــان می‌دهد. بالاترین درصـــد بیان‌گر نتیجه نهایی آموزش شـــبکه اســـت [۳۳]. در الگوریتم‌های یادگیری، بحث مهم در نظر گرفتن رابطه‌ای جهت ارزیابی میزان صحت یادگیری یا دسته‌بندی است این رابطه که با نام تابع تلفات نیز از آن یاد می شود بایستی با روش‌های مناسب، بهینه سازی یا کمینه شود. بنابراین ایده اساسی در آموزش یک شبکه و یا الگوریتم یادگیری کمینه‌کردن تابع تلفات با پیمودن گام‌های مناسب برای حصول به نتیجه بهتر اســـت. روش‌های مهم کمینه‌ســـازی شـــامل گرادیان نزولی، گرادیان نزولی دســـته‌ای یا ***batch***، گرادیان نزولی تصـادفی و گرادیان نزولی ***mini-batch***، الگوریتم ***Adagrad***[۳۷]، الگوریتم ***RMSprop***[۳۸] و ***Adam***[۳۹] می‌باشند.

در شـبکه‌های عصـبی کانولوشـنال به دلیل اینکه تعداد پارامترهایی که آموزش می‌بینند بســیار زیاد اســت و همچنین این شـــبکه‌ها دارای لایه‌های عمیق هستند لذا در طی زمان آموزش، شبکه مستعد فرایادگیری می‌باشند. این به این معنا است که شبکه برای داده‌های آموزش عملکرد خوبی از خود نشان می‌دهد اما این عملکرد با وارد شدن داده‌های تست دچار مشـــکل می‌شـــود. برای حل نمودن این مشـــکل راه‌های متفاوتی پیشـــنهاد شـــده اســـت. یکی از روش‌های معروف، روش افزایش تعداد داده‌های آموزش و روش دیگر، روش اضافه نمودن لایه *dropout* مناسب در لایه کاملاً متصل برای رفع این معضل است. روش افزایش داده‌ها در موضـــوع فرآیندهای یادگیری، یک روش مرســـوم بوده که در آن با اعمال تغییراتی در تصـــاویر مانند چرخاندن، بزرگ‌نمایی، جابجایی و غیره به افزایش تعداد تصاویر ورودی به شبکه می‌پردازند. در روش *dropout* قبل از هر مرحله آموزش، تعدادی از نورون‌ها به صـــورت رندم و با نســـبت تعیین‌شـــده در برنامه مورد نظر از فرآیند آموزش خارج می‌شـــوند و تنها نورون‌های باقی‌مانده در آموزش شرکت می‌کنند [۳۴].

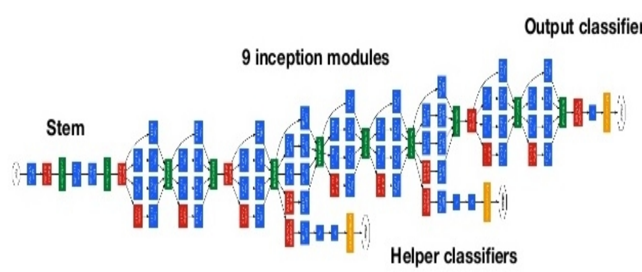
۲-۳- باز تنظیم شبکه‌های از قبل آموزش دیده

در زمینه‌های پزشکی به علت کمبود تعداد تصاویر دارای طبقه‌بندی مناسب، آموزش کامل یک شبکه عصبی کانولوشنال بسیار دشوار و یا غیر ممکن است. علاوه بر این آموزش شبکه از ابتدا به سخت افزارهای قوی همانند پردازش‌گرهای قوی گرافیکی نیاز دارد که عملاً با محدودیت سخت‌افزاری مذکور آموزش این شبکه‌ها از ابتدا ناممکن می‌باشد [۲۱]. بنابراین به دو روش اســتخراج ویژگی و یا اســتفاده از روش بازتنظیم یک شبکه از پیش آموزش دیده در لایه‌های انتهایی آن می‌توان عمل نمود. در روش اســتخراج ویژگی، به کمک یک *CNN*، ویژگی‌های تصــاویر پزشــکی بدســـت آمده و به کمک یک روش طبقه‌بندی، همانند ماشـــین بردار پشــتیبان می‌توان به دســته‌بندی مورد نظر دســت یافت. این شــیوه برای زمانی که تعداد نمونه‌ها بســیار کم اســت دارای کاربرد بوده در این حالت بایستی ابتدا ویژگی‌های تصاویر را به کمک لایه‌های کانولوشنال و تلفیق اســـتخراج نموده و در نهـایـت این ویژگی‌هـا را بـه‌عنوان ورودی بـه الگوریتم‌های طبقه‌بندی داد، تا طبقه‌بندی نهایی صـــورت گیرد. اما زمانی که تعداد نمونه‌ها بیشـــتر بوده ولیکن به اندازه‌ای نیســـت که بتوان کل پارامترهای شـــبکه را از ابتدا آموزش داد، می‌توان از روش دوباره تنظیم شبکه استفاده شود [۳۲]. در روش بازتنظیم یک *CNN* ابتدا یک شبکه

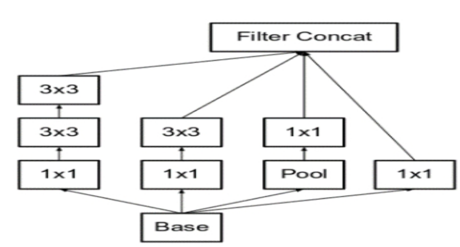
آموزش دیده که نتایج بسیار خوبی در بانک‌های تصاویر معروف به‌دست آورده انتخاب می‌گردد، سپس لایه تماماً متصل نهایی با لایه تماماً متصل جدید که قابلیت طبقه‌بندی به تعداد کلاس‌های مورد نظر را دارد، جایگزین می‌گردد. همچنین وزن‌های لایه‌های اولیه شبکه اصلی بدون تغییر وارد شبکه مورد نظر جدید می‌شوند و لایه‌های نهایی شبکه دوباره آموزش می‌بینند تا شبکه مورد نظر قابلیت طبقه‌بندی تصاویر جدید را داشته باشد.

۳-۳- شبکه CNN گوگل‌نت و مدل Inception-v3

شبکه عصبی کانولوشنال گوگل‌نت به عنوان عمیق‌ترین شبکه *CNN* در بین مدل‌های رایج بوده مهم‌ترین وجه تمایز این شبکه نسبت به مدل‌های قبلی مانند الکس‌نت[۴۰] یا سیفارنت[۴۱] استفاده از ماژول *Inception* در ساختار خود است [۳۵]. به شکل کلی یک شبکه کانولوشنال گوگل‌نت دارای دو لایه کانولوشنال، دو لایه تلفیق و ۹ لایه *Inception* بوده هر لایه *Inception* خود دارای ۶ لایه کانولوشنال و یک لایه تلفیق است. شکل (۶) ساختار یک گوگل‌نت را نشان می‌دهد. مدل *Inception-v3* به عنوان یک مدل پیشرفته از شبکه عصبی کانولوشنال گوگل‌نت، اولین بار توسط *Szegedy* و همکاران در سال ۲۰۱۵ معرفی گردید [۳۶]. در این مدل کانولوشنال‌های ۱×۱، ۳×۳ و ۵×۵ در یک ماژول یکسان و بر روی یک ورودی اعمال می‌شوند و خروجی تمامی این فرآیندها با یکدیگر الحاق شده تا به عنوان ورودی لایه بعدی در نظر گرفته شوند. شکل (۷) یک ماژول *Inception-v3* از شبکه کانولوشنال گوگل‌نت را نشان می‌دهد.



شکل ۶ مدل شبکه عصبی کانولوشنال گوگل‌نت با ماژول‌های *Inception*



شکل ۷ ماژول *Inception-V3*

۴- نتایج شبیه سازی

در این بخش جهت دست‌یابی به نتایج قابل تحلیل، ضمن معرفی معیارهای مختلف ارزیابی عملکرد روش‌های هوشمند دسته‌بندی، به معرفی بانک‌داده کگل و ویژگی‌های ارائه شده در این مجموعه از تصاویر شبکیه چشم جهت بازتنظیم شبکه کانولوشنال مورد نظر، پرداخته می‌شود [۳۸]. سپس جهت ارزیابی قدرت طبقه‌بندی شبکه بازتنظیم شده از تصاویر مرکز چشم پزشکی نوید دیدگان استفاده خواهد شد.

۴-۱- معیارهای ارزیابی عملکرد روش‌های دسته‌بندی

معیارهای ارزیابی عملکرد الگوریتم‌های دسته‌بندی معمولاً شامل دقت[۴۲]، قابلیت پیش‌بینی[۴۳] یا صحت[۴۴]، حساسیت[۴۵] و معیار امتیاز یا رتبه *F1* می‌باشند. با توجه به نتایج این معیارها می‌توان در مورد عملکرد الگوریتم‌های ارائه شده بحث نمود. برای توضیح این چهار معیار بایستی ابتدا اجزای مورد استفاده برای محاسبه آن‌ها بررسی شوند. این اجزا شامل چهار مورد مثبت درست[۴۶] *(TP)* یعنی آن دسته از نمونه‌هایی که دارای بیماری *DR* بوده و به درستی تشخیص داده شده و به عنوان بیمار دسته‌بندی شده‌اند، مثبت نادرست[۴۷] *(FP)* اشاره به آن دسته از تصاویر شبکیه چشم دارد که *NoDR* بوده و به اشتباه در دسته *DR* قرار گرفته‌اند، منفی درست[۴۸] *(TN)* اشاره به نمونه‌هایی دارد که *NoDR* بوده و به درستی در دسته *NoDR* طبقه‌بندی شده‌اند و نهایتا منفی نادرست[۴۹] *(FN)* اشاره به نمونه‌هایی دارد که DR بوده و اشتباها *NoDR* کلاسه‌بندی شده‌اند [۳۷]. با توجه به توضیحات مطرح شده روابط ارزیابی عملکرد سیستم در جدول (۱) تعریف می‌گردد.

جدول ۱: معیارهای ارزیابی عملکرد روش‌های دسته‌بندی داده

معیار	رابطه
دقت	$(TP+TN)/(P+N)$
قابلیت پیش‌بینی یا صحت	$TP/(TP+FP)$
حساسیت	$TP/(TP+FN)$
معیار رتبه *F1*	$2TP/(2TP+FP+FN)$

در جدول (۱)، $\boldsymbol{P+N}$ نشان‌دهنده تمام نمونه‌ها می‌باشد.

۴-۲- بازتنظیم شبکه کانولوشنال گوگت‌نت با استفاده از داده‌های بانک اطلاعاتی کگل

تصاویر مورد استفاده به‌عنوان ورودی شبکه کانولوشنال تصاویر مربوط به رقابت رتینوپاتی دیابتی کگل می‌باشند. تعداد این تصاویر ۳۵۱۲۶ عکس می‌باشد که در پنج گروه سالم، رتینوپاتی دیابتی خفیف، متوسط، شدید و تکثیری طبقه‌بندی شده‌اند. مشخصات این تصاویر در جدول (۲) آمده است.

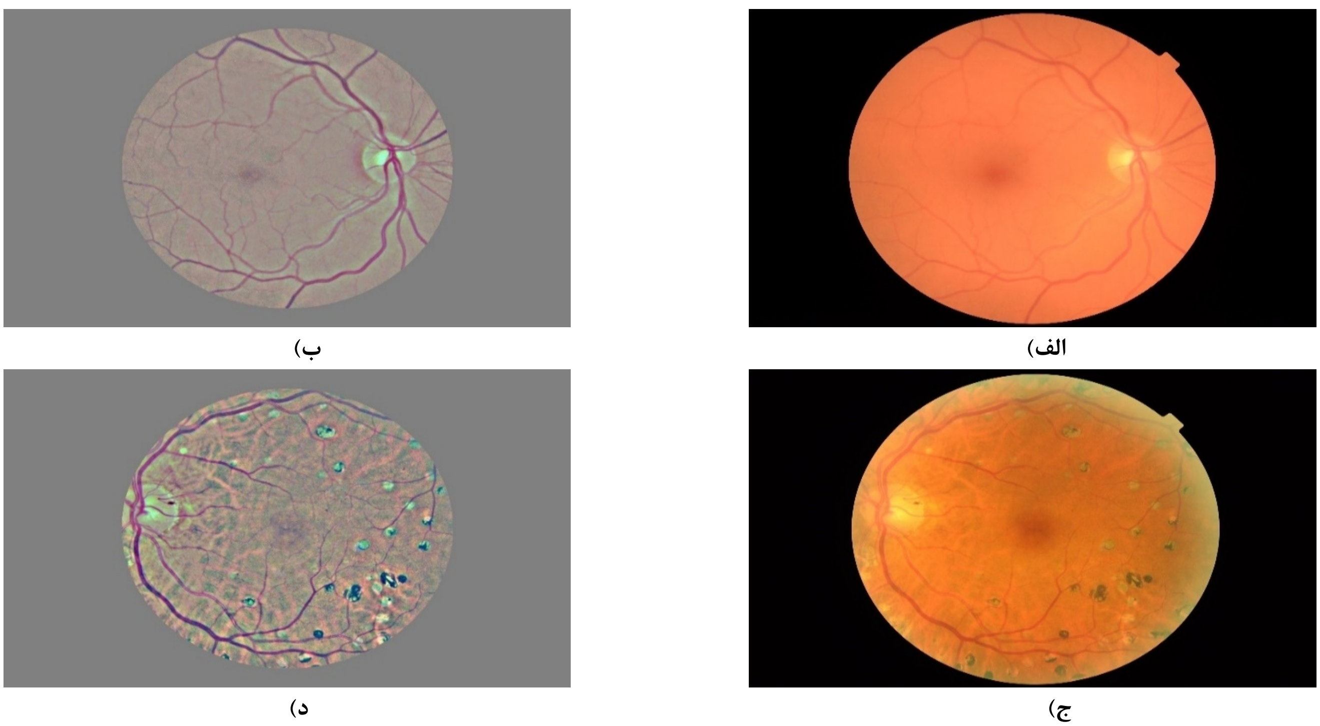
جدول ۲ تعداد و مشخصات تصاویر بانک اطلاعاتی کگل [۳۸]

شدت بیماری دیابت چشمی	تعداد تصاویر
تصاویر شبکیه سالم یا *(NoDR)*	۲۵۸۱۰
تصاویر شبکیه با رتینوپاتی دیابتی خفیف	۲۴۴۳
تصاویر شبکیه با رتینوپاتی دیابتی متوسط	۵۲۹۲
تصاویر شبکیه با رتینوپاتی دیابتی شدید	۸۷۳
تصاویر شبکیه با رتینوپاتی دیابتی تکثیری *(PDR)*	۷۰۸

به دلیل کم بودن تصاویر شبکیه در بخش‌های مربوط به رتینوپاتی دیابتی متوسط، شدید و تکثیری نسبت به تصاویر بدون رتینوپاتی دیابتی، طراحی شبکه‌ای که بتواند بدون فرایادگیری نتایج مطلوبی بدهد با دوباره تنظیم شبکه‌های از پیش آموزش دیده عملاً ممکن نیست. بنابراین با یک گروه در نظر گرفتن تمامی تصاویر مربوط به مبتلایان به رتینوپاتی دیابتی اعم از خفیف، متوسط، شدید و تکثیری می‌توان شبکه را برای دو گروه *DR* و *NoDR* آموزش داد. گروه اول شامل تصاویر شبکیه‌های سالم یا بدون رتینوپاتی دیابتی که طبق جدول (۲) دارای ۲۵۸۱۰ تصویر و گروه دوم شامل تصاویر شبکیه مبتلا به رتینوپاتی دیابتی با ۹۳۱۶ عکس شبکیه که از جمع تصاویر چهار دسته باقیمانده جدول (۲) بدست می‌آید.

۴-۲-۱- پیش پردازش تصاویر

تصاویر بانک اطلاعاتی کگل با دستگاه‌های مختلف و با رزولوشن‌های متفاوت تصویربرداری شده لذا دارای کیفیت‌های متفاوتی هستند. بنابراین اولین مرحله، اعمال الگوریتم مناسب پیش‌پردازش بر روی تصاویر می‌باشد. برای انجام پیش‌پردازش در این مقاله از بسته *OpenCV* استفاده شده است. در این بسته اولین مرحله پیش‌پردازش، تغییر مقیاس عکس‌ها است تا تصاویر ورودی اندازه کره چشم برابر داشته باشند. در مرحله بعد رنگ هر پیکسل از تصویر، از میانگین محلی شدت رنگ‌های آن کم می‌شود تا میانگین شدت رنگ‌ها به ۵۰ درصد خاکستری برسد. در نهایت لبه‌های تصاویر حذف می‌شوند تا اثر لبه‌ها در طبقه‌بندی اختلال ایجاد نکند [۳۷]. شکل (۸) دو تصویر از داده‌های بانک اطلاعاتی مورد نظر، یک تصویر مربوط به شخص بیمار و تصویر دیگر مربوط به شخص سالم را به همراه تصاویر پیش‌پردازش شده آنها نشان می‌دهد.



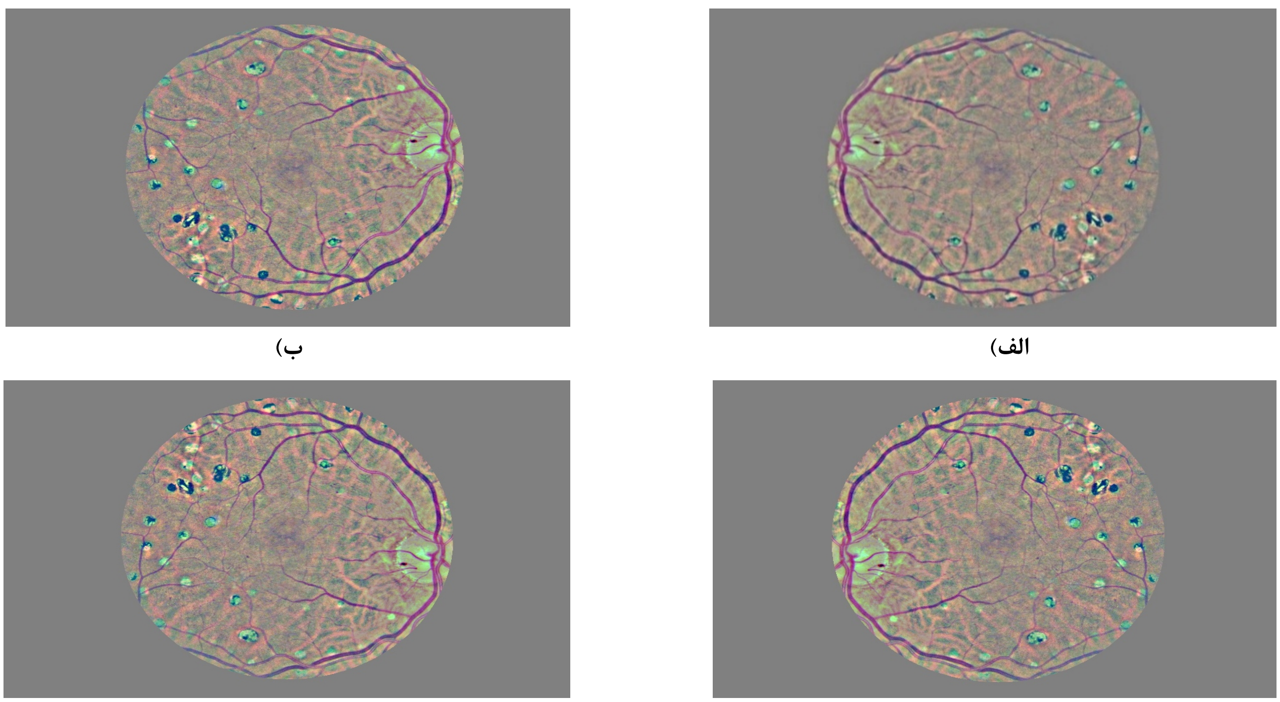
شکل ۸ پیش‌پردازش تصاویر شبکیه چشم بانک اطلاعاتی کگل به لحاظ اندازه، شدت روشنایی و لبه‌های تصویر الف) یک نمونه تصویر شبکیه *NoDR* ب) تصویر پیش‌پردازش‌شده *NoDR* ج) تصویر شبکیه شخص مبتلا به *PDR* د) تصویر پیش‌پردازش شده شبکیه *PDR*

۴-۲-۲- افزایش داده‌ها

چنانچه اشاره گردید تصاویر موجود در بانک اطلاعاتی کگل شامل ۳۵۱۲۶ عکس شبکیه می‌باشد که از این تعداد ۲۵۸۱۰ تصویر آن مربوط به افراد سالم یا *NoDR* بوده و بقیه ۹۳۱۶ تصویر دارای علائم *DR* می‌باشند. برای متوازن نمودن تعداد تصاویر در دو کلاس طبقه‌بندی، در این مقاله از روش افزایش داده‌ها استفاده شده است. به این صورت که در

ابتدا نیمی از تصاویر بیمار به‌صورت افقی و نیمه دیگر به‌طور عمودی و البته به صورت تصادفی برعکس شده‌اند. سپس ۷۵ درصد تصاویر بیمار هم عمودی و هم افقی و به‌طور تصادفی برعکس شده‌اند. در نهایت با این عمل تعداد تصاویر *DR* به ۲۵۶۱۹ عدد که مقداری نزدیک تعداد تصاویر *NoDR* دارد، رسید. شکل (۹) افزایش داده را برای یک تصویر شبکیه از بانک مورد نظر، نشان می‌دهد.

بسته‌های نرم افزاری مورد استفاده در این مقاله شامل *Tensorflow*، *Numpy*، *h5py*، *Sciki-learn*، *OpenCv* و *Keras* می‌با شد. پکیج *Keras* یک کتابخانه شبکه عصبی عمومی تحت محیط نرم افزاری پایتون می‌با شد که شامل مدل‌های از قبل آموزش‌دیده *CNN* بوده این شبکه‌ها بهترین شبکه‌های کانولوشنال می‌باشند که در چند سال اخیر روی داده‌های بانک اطلاعاتی *ImageNet* آموزش دیده‌اند [۳۲] و نتایج خوبی به‌دست آورده‌اند. بانک تصاویر *ImageNet*، یک بانک اطلاعاتی پرکاربرد بوده که تعداد ۱/۲ میلیون تصویر در ۱۰۰۰ کلاس متفاوت را گردآوری کرده است و هر ساله رقابتی برای طبقه‌بندی تصاویر این بانک اطلاعاتی برگزار می‌شود. برنامه‌های *Keras* و *Tensorflow* برای اجرای الگوریتم‌های *CNN* بایستی در سیستم عامل *Ubuntu* نصب شوند. سایر بسته‌های نرم افزاری نیز برای انجام محاسبات مورد نیاز و پیش‌پردازش تصاویر در محیط نرم افزاری پایتون اجرا می‌شوند. استفاده از شبکه‌های از پیش آموزش دیده‌ی *CNN* برای طبقه‌بندی تصاویر، باعث می شود تا نیازی به استفاده از واحد پردازش گرافیک و حافظه خارجی که برای آموزش این شبکه‌ها از ابتدا ضروری است، نبا شد. این قضیه یکی از دلایل اصلی استفاده از این شبکه‌ها در این مقاله است. در این مقاله شبیه سازی‌ها با لپ‌تاپی با پردازش‌گر هفت هسته‌ای شرکت اینتل و حافظه رم ۸ گیگا بایتی انجام شده است که در مقایسه با سخت‌افزارهای مورد نیاز برای آموزش شبکه مورد نظر از ابتدا، بسیار کم‌هزینه‌تر می‌باشد. مدت آموزش شبکه‌ها با توجه به تعداد تکرارها متفاوت می‌باشد.



الف) ب)

ج) د)

شکل ۹ افزایش تعداد تصاویر *DR* برای متوازن نمودن تعداد نمونه‌ها در دو دسته الف) تصویر پیش پردازش شده اصلی ب) تصویر قرینه نسبت به محور عمودی ج) تصویر قرینه نسبت به محور افقی د) تصویر قرینه نسبت به محور افقی و محور عمودی

۴-۲-۳- نتایج حاصل از آموزش شبکه

برای انجام شبیه‌سازی‌ها و همچنین تست شبکه مورد نظر، ۲۰ در صد داده‌ها به صورت تصادفی برای تست و ۸۰ در صد دیگر نیز برای آموزش انتخاب شده‌اند. برای آموزش شبکه، ابتدا وزن‌های شبکه از پیش آموزش داده شده، ذخیره می‌شوند. چنانچه اشاره گردید شبکه مورد استفاده در این مقاله، شبکه جدید گوگل‌نت *Inception-v3* می‌با شد. در ابتدا وزن‌های اولیه شبکه پیشنهادی به غیر از لایه تماماً متصل را وزن‌های اولیه شبکه *Inception-v3* در نظر گرفته و وزن‌های لایه تماماً متصل به صورت تصادفی مقداردهی اولیه می شوند، حال شبکه‌ای جدید

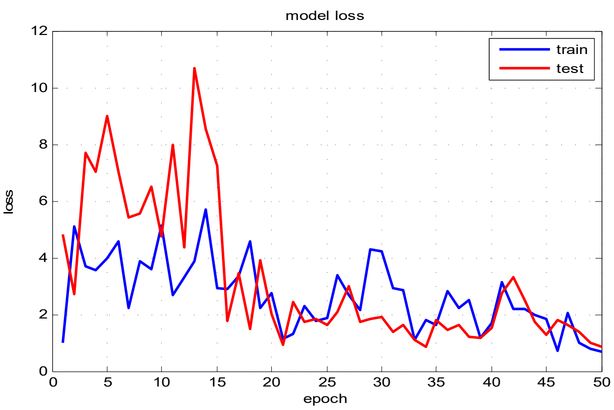
برای تشخیص رتینوپاتی دیابتی مدل شده است. ابتدا بایستی لایه اضافه شده به این شبکه را برای تعداد اندکی تکرار آموزش دهیم تا لایه تماماً متصل جدید با لایه‌های قبلی هماهنگ گردد. در این مرحله بایستی وزن‌های لایه‌های دیگر بدون تغییر باقی بماند.

در مرحله بعد بایستی به تنظیم دوباره لایه‌های آخر شبکه برای حصول به نتایج بهتر پرداخت. برای این کار لایه‌های اول تا لایه مورد نظر بدون تغییر باقی می‌ماند و یا به اصطلاح فریز می‌شوند و لایه‌های نهایی دوباره آموزش می‌بینند. بنابراین ابتدا دو بلوک آخر یا لایه‌ی ۱۷۲ تا آخر، برای دوباره آموزش دادن در نظر گرفته می شود و لایه‌های قبل لایه ۱۷۲ فریز شده به این معنا که وزن‌ها و بایاس‌های لایه اول تا لایه ۱۷۲ بدون تغییر باقی می‌ماند. هرچند برای بررسی تأثیر آموزش لایه‌های بیش‌تر می‌توان لایه‌های فریز شده را کمتر نمود اما از طرف دیگر هرچه تعداد لایه‌ها برای آموزش، بیش‌تر در نظر گرفته شوند زمان آموزش طولانی‌تر می‌گردد برای مثال زمان آموزش با توجه به سخت‌افزار مورد استفاده در این مقاله و بسته به تعداد لایه مورد نظر جهت آموزش، بین ۳ تا ۱۰ ساعت متغیر می‌باشد.

برای آموزش این شبکه از توابع فعالیت متفاوتی استفاده شده است. توابع استفاده شده برای لایه *dense* که یکی از لایه‌های بخش تماماً متصل می‌باشد، شامل *ReLu*، *ELU*، *tanh* و *Sigmoid* است. نتایج استخراج شده برای تابع *ELU* قابل قبول‌تر بود، بنابراین برای کلیه شبیه سازی‌ها از این تابع فعالیت استفاده شده است. از طرف دیگر برای انتخاب الگوریتم بهینه سازی مناسب نیز از بهینه‌کننده‌های متفاوتی نظیر *SGD*، *adam*، *adagrad*، *adamax* و *nadam* مورد آزمایش قرار گرفتند از بین تمامی این توابع بهینه‌کننده *adam* و *SGD* نتایج قابل قبول‌تری از خود نشان دادند. اما به دلیل اینکه *SGD* سرعت بسیار کمی دارد و به تعداد تکرارهای بسیار زیاد و همچنین زمان اجرای بالا نیاز دارد، *adam* الگوریتم بهینه سازی بهتری به نظر رسید. این تابع سرعت بسیار بالایی دارد و در تکرارهای کم، زودتر به نتیجه می‌رسد.

معیار ارزیابی و بررسی نتایج به‌دست آمده نمودار تلفات می‌باشد. طبیعتا الگوریتمی مناسب است که بتواند تلفات کمتری داشته باشد. شکل (۱۰) تلفات برنامه را به ازای بازتنظیم دو بلوک آخر به همراه لایه تماماً متصل نشان می‌دهد.

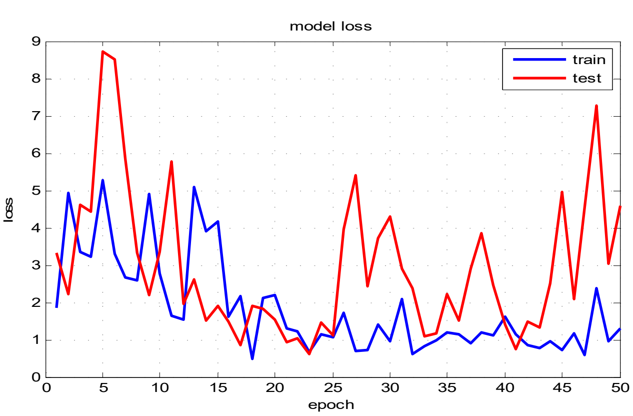
نوسانات موجود در نمودار تلفات به دلیل کم بودن اندازه بسته‌های داده در نظر گرفته شده در برنامه است که این اندازه به دلیل محدودیت سخت‌افزاری ۱۶ در نظر گرفته شده است. برای به‌دست آوردن نموداری با نرخ کاهش هموارتر بایستی این عدد بسیار بزرگ‌تر در نظر گرفته شود همچنین افزایش تعداد تکرارها، به بیش از صدهزار تکرار در هموار نمودن شکل کاملاً اثرگذار خواهد بود. که با توجه به توضیحات ارائه شده در مورد سخت‌افزار مورد استفاده جهت اجرای برنامه تعداد تکرار بیش‌تر عملاً



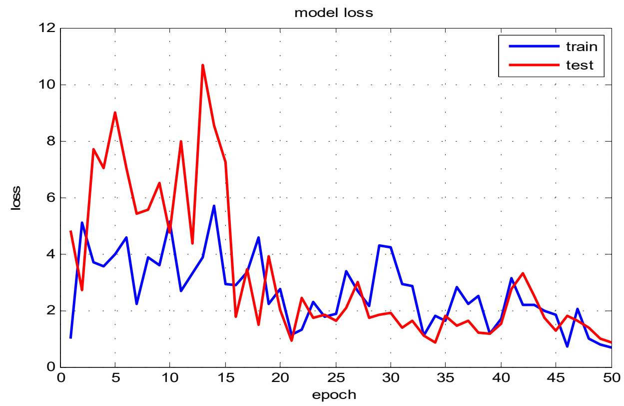
شکل ۱۰ نمودار تلفات آموزش دوباره لایه ۱۷۲ تا آخر شبکه گوگل‌نت *Inception-v3* بر اساس تصاویر بانک اطلاعاتی کگل

ناممکن می‌باشد. اما با توجه به این نمودار کاملاً مشخص است که با آموزش دو بلوک آخر شبکه از قبل آموزش داده شده تلفات مدل در حال کاهش است و در نتیجه عملکرد الگوریتم مورد نظر قابل توجیه می‌باشد. برای بررسی تأثیر آموزش لایه‌های بیش‌تر بر عملکرد شبکه پیشنهادی، این‌بار سه بلوک آخر را دوباره آموزش می‌دهیم. این به این معنا است که از لایه ۱۵۸ تا به آخر و همچنین لایه تماماً متصل را دوباره آموزش می‌دهیم و لایه‌های ماقبل لایه ۱۵۸ فریز شده و وزن‌ها و بایاس‌های آن‌ها بدون تغییر باقی می‌ماند. نمودار تلفات به صورت شکل (۱۱) خواهد بود.

در این شکل همانطور که مشخص است شبکه دچار پدیده فرایادگیری شده که نمودار تست و آموزش از تکرار ۲۴ به بعد از هم فاصله می‌گیرند. برای حل این مشکل از روش اضافه نمودن لایه *dropout* که در بخش سوم مورد بررسی قرار گرفت می‌توان استفاده نمود. با افزودن لایه *dropout* با نسبت ۰/۵ و با اجرای دوباره‌ی برنامه نمودار (۱۲) از تابع تلفات شبکه حاصل می‌گردد. در این نمودار همان‌طور که مشخص است هم نمودار آموزش و هم نمودار تست روند کاهشی مناسبی دارند.



شکل ۱۱ نمودار تلفات آموزش دوباره لایه ۱۵۸ تا آخر شبکه گوگل‌نت *Inception-v3* برای تشخیص دیابت چشمی (شبکه دچار فرایادگیری)



شکل ۱۲ نمودار تلفات آموزش دوباره لایه ۱۵۸ تا آخر شبکه گوگل‌نت *Inception-v3* با افزودن لایه *dropout* برای تشخیص دیابت چشمی

با توجه به شبیه‌سازی‌های متعدد صورت پذیرفته می‌توان به این نتیجه‌گیری مهم رسید که برای تشخیص رتینوپاتی دیابتی از روی تصاویر شبکیه با استفاده از دوباره آموزش دادن لایه‌های آخر شبکه گوگل‌نت، لایه‌های منا سب برای تنظیم دوباره لایه‌ی ۱۵۸ به آخر یا سه بلوک آخر شبکه *Inception-v3* می‌باشد. تنظیم لایه‌های بیش‌تر و کم‌تر نتایج قابل قبولی نخواهد داشت.

۴-۳- ارزیابی شبکه بازتنظیم شده با استفاده از تصاویر کلینیک چشم پزشکی نوید دیدگان

به منظور برر سی بیش‌تر نتایج به‌د ست آمده و همچنین تلاش برای کاربردی نمودن نتایج این مقاله، از داده‌های واقعی کلینیکی برای ارزیابی شبکه استفاده گردیده است. به این منظور ۱۰۱ داده از کلینیک نوید دیدگان دریافت شده است. این ۱۰۱ تصویر، تصاویر شبکیه چشم مراجعه‌کنندگان به این کلینیک بوده که توسط پزشک متخصص کلینیک عکس‌برداری و به دو گروه شبکیه‌های سالم و شبکیه‌های با علائم بیماری رتینوپاتی دیابتی طبقه‌بندی شده‌اند. به منظور بررسی عملکرد شبکه ارائه شده بر روی داده‌های کلینیک، از معیارهای برر سی عملکردی جدول (۱) ا ستفاده می‌گردد. بنابراین با اعمال این ت صاویر به شبکه بازتنظیم شده، معیارهای ارزیابی عملکرد به صورت جدول (۳) به‌دست می‌آید. همان‌طور که از جدول (۳) مشخص است، دوباره آموزش دادن لایه ۱۵۸ تا آخر و به عبارت دیگر سه بلوک آخر شبکه بهترین نتایج را برای داده‌های کلینیکی ارائه می‌کند. اگرچه معیار صحت در آموزش لایه ۱۷۲ به آخر بالا می‌باشد، اما این عدد بسیار بالا است و به‌عبارت دیگر می‌توان گفت حساسیت شبکه مورد نظر برای شناسایی بیماران بالا است بنابراین بسیاری از تصاویر سالم را نیز بیمار تشخیص دهد. به همین دلیل این شبکه، شبکه مناسبی برای فرآیند غربالگری نیست. از میان سه شبکه توضیح داده شده مناسب‌ترین نتایج برای داده‌های کلینیکی، شبکه دوم است که با دوباره

جدول ۳ عملکرد ساختار پیشنهادی بر روی داده‌های کلینیک خصوصی چشم پزشکی نوید دیدگان

معیار / روش	دقت	صحت	پیش‌بینی	امتیاز F1
لایه ۱۳۶ تا آخر	۰/۵۸۴۱	۰/۵۸۳۳	۰/۵۶	۰/۵۷۱۴
لایه ۱۵۸ تا آخر	۰/۷۰۳۰	۰/۸۵۷۱	۰/۴۸	۰/۶۱۵۴
لایه ۱۷۲ تا آخر	۰/۶۷۳۳	۰/۹۴۷۳	۰/۳۶	۰/۵۲۱۷
گوگل‌نت پایه	۰/۳۲۴۱	۰/۴۳۲۳	۰/۴۷۸۳	۰/۵۰۴۳
گوگل نت بدون پیش‌پردازش تصاویر	۰/۳۱۲۴	۰/۴۲۳۵	۰/۴۵۶۷	۰/۴۷۷۹

آموزش دادن سه بلوک آخر مدل *Inception-v3* حاصل شده است. همچنین در این جدول نتایج روش پیشنهادی، با یک شبکه گوگل‌نت که فرآیند آموزش روی تمام لایه‌های آن اجرا شده (گوگل‌نت پایه) مقایسه شده است مدت زمان این آموزش بسیار زیاد بوده و نتایج حاصل مطلوب نیستند. بر اساس دو معیار دقت و صحت روش پشینهادی با بازتنظیم لایه ۱۷۲ تا آخر، نتایجی با بیش از ۱۰۰ درصد بهبود را نسبت به نتایج گوگل‌نت پایه نشان می‌دهد. البته گوگل‌نت پایه که نتایج ضعیفی در دو معیار دقت و صحت دارد در معیار پیش‌بینی وضعیت بهتری را دارد. در معیار امتیاز F1 نتایج روش پیشنهادی و روش گوگل نت پایه مشابه همدیگر هستند. همچنین فرآیند آموزش روی شبکه گوگل‌نت پایه بدون فرآیند پیش پردازش تصاویر و بدون فرآیند افزایش داده که در بخش ۴ توضیح داده شد نیز حکایت از پائین بودن تمامی معیارهای ارزیابی دارد.

۵- جمع‌بندی و نتیجه‌گیری

در این مقاله، تصاویر شبکیه بانک اطلاعاتی کگل به شبکه عصبی کانولوشنال از پیش آموزش دیده گوگل‌نت *Inception-v3* به عنوان عمیق-ترین شبکه کانولوشنال، جهت بازتنظیم این شبکه برای تشخیص رتینوپاتی دیابتی پرداخته شد. تنظیم دوباره شبکه‌های از پیش آموزش دیده به صورت دوباره آموزش دادن تعدادی از لایه‌های آخر شبکه اصلی، معضل کمبود داده‌های آموزش در حوزه پزشکی را مرتفع می‌سازد. در این مقاله شبکه‌ی معروف *Inception-v3* به‌عنوان شبکه از پیش آموزش دیده معرفی گردید و لایه‌های ۱۳۶، ۱۵۸ و ۱۷۲ به آخر این شبکه دوباره آموزش داده شد و تنظیم گردید. آموزش لایه ۱۵۸ به آخر از بین روش‌های دیگر نتایج بهتری از خود نشان داده است.

در بخش پایانی این مقاله، به اعمال تصاویر شبکیه اتخاذ شده از کلینیک نوید دیدگان به شبکه پیشنهادی پرداخته شد. در روش آموزش دوباره لایه ۱۵۸ تا آخر، نتایج غربال‌گری تصاویر کلینیکی با دقت بیش از ۷۰ درصد نسبت به سایر روش‌ها حاصل گردید.

مراجع

[1] X. Zhang and e. al., "Exudate detection in color retinal images for mass screening of diabetic retinopathy," *Medical Image Analysis,* vol. 18, pp. 1026-1043, October 2014.

[2] K. S. Argade, K. A. Deshmukh, M. M. Narkhede, N. N. Sonawane, and S. Jore, "Automatic detection of diabetic retinopathy using image processing and data mining techniques," presented at the 2015 Int. Conference on Green Computing and Internet of Things, 2015.

[3] H. M. Zheng Y, Congdon N, "The worldwide epidemic of dneaiabetic retinopathy," *Indian J Ophthalmol,* vol. 60, pp. 428-431, 2012.

[4] B. Antal and A. Hajdu, "An ensemble-based system for Microaneurysm detection and diabetic retinopathy grading," *IEEE Trans. Biomeical Engineering,* vol. 59, pp. 1720–172 ,June 2012.

[5] P. P. Conde, J. d. l. Calleja, A. Benitez, and M. A. Medina, "Image-based classification of diabetic retinopathy using machine learning," Int. Conf. Intell. Syst. Design and Applications, Nov., 2012.

[6] M. U. Akram, S. Khalid, A. Tariq, and F. Azam, "Detection and classification of retinal lesions for grading of diabetic retinopathy," *Computers in Biology and Med.,* vol. 45, pp. 161–171, Feb. 2014.

[7] C. Sundhar and D. Archana, "Automatic screening of fundus images for detection of Diabetic Retinopathy," *Int. J. Communication and Computer Technologies,* vol. 2, April 2014.

[8] A. F. M. Hani and H. A. Nugroho, "Gaussian Bayes classifier for medical diagnosis and grading: Application to diabetic retinopathy," IEEE Conf. Biomedical Engineering and Sciences, Nov. 2010.

[9] B. v. G. M. Niemeijer, Michael J. Cree, "Retinopathy online challenge: Automatic detection of Microaneurysms in digital color Fundus photographs," *IEEE Trans. Medical Imaging,* vol. 29, pp. 185-195, Jan. 2010.

[10] S. S. Rahim, C. Jayne, V. Palade, and J. Shuttleworth, "Automatic detection of microaneurysms in colour fundus images for diabetic retinopathy screening," *Neural Computing and Applications,* 2015.

[11] L. G. L. Giancardo, F. Meriaudeaub, T. P. Karnowskic, Y. L. G. Kenneth W. Tobin, J. C. EdwardChaumd, "Exudate-based diabetic macular edema detection in fundus images using publicly available datasets," Medical Image Analysis, vol. 16, pp. 216–226, Jan. 2012.

[12] C. Jayakumari and T. Santhanam, "Detection of hard exudates for diabetic retinopathy using contextual clustering and fuzzy art neural network," *J. Information Technology,* vol. 6, pp. 842-846, 2012.

[13] A. Osareh, B. Shadgar, and R. Markham, "A computational-intelligence-based approach for detection of exudates in diabetic retinopathy images," *IEEE Trans. Information Technology in Biomedicine,* vol. 13, pp. 535–545, July 2009.

[14] S. Franklin and S. Rajan, "Diagnosis of diabetic retinopathy by employing image processing technique to detect exudates in retinal images," *IET Image Process.*, vol. 8, no. 10, pp. 601-609, 2014.

[15] G. B. Kande, T. S. Savithri, and P. V. Subbaiah, "Automatic detection of Microaneurysms and hemorrhages in digital Fundus images," *J. Digital Imaging,* vol. 23, pp. 430–437, November 2009.

[16] M. U. Akram, S. Khalid, A. Tariq, and M. Y. Javed, "Detection of neovascularization in retinal images using multivariate m-mediods based classifier," *Computerized Medical Imaging and Graphics,* vol. 37, pp. 346–357, July 2013.

[17] A. P. Bhatkar and G. U. Kharat, "Detection of diabetic Retinopathy in retinal images using MLP Classifier," IEEE Int. Symposium on Nanoelectronic and Information Syst., Dec. 2015.

[18] R. Priya and P. Aruna, "Diagnosis of diabetic retinopathy using machine learning techniques," *J. Soft Computing,* vol. 3, pp. 563–575, July 2013.

[19] K. Saranya, B. Ramasubramanian, and S. K. Mohideen, "A novel approach for the detection of new vessels in the retinal images for screening diabetic Retinopathy," Int. Conf. Communication and Signal Processing, April 2012.

[20] W. Zhang, R. Li, H. Deng, L. Wang, "Deep convolutional neural networks for multi-modality isointense infant brain image segmentation," *NeuroImage,* vol. 108, pp. 214-224, 2015.

[21] J. Y. Tajbakhsh, Suryakanth and R. Gurudu, "Convolutional neural networks for medical image analysis: fine tuning or full training?," *IEEE Trans. Medical Imag.,* vol. 35, pp. 1299–1312, May 2016.

[22] D. Nie, L. Wang, E. Adeli, C. Lao, W. Lin and D. Shen, "3-D fully convolutional networks for multimodal isointense infant brain image segmentation," *IEEE Trans. Cybern.*, vol. PP, no. 99, pp. 1-14, 2018.

[23] L. L. olger R. Roth, A. Farag, H. Shin, J. Liu, E. Turkbey and R. M. Summers, "Deeporgan: multi-level deep convolutional networks for automated pancreas segmentation," Medical Image Computing and Computer-Assisted. Lecture Notes in Computer Science, 2015.

[24] Y. Wanga, G. Caoa, B. Weib and G. Yang, "Hierarchical retinal blood vessel segmentation based on feature and ensemble learning," *J. Neurocomput.,* vol. 149, pp. 708–717, Feb. 2015.

[25] F. C. Harry Pratta, Deborah M. Broadbentc, P. Simon P. Hardingac and Y. Zhengac, "Convolutional neural networks for diabetic retinopathy," International Conference On Medical Imaging Understanding and Analysis, July 2016.

[26] H. H. Vo and A. Verma, "New deep neural nets for fine-grained diabetic retinopathy recognition on hybrid color space," IEEE Int. Symposium on Multimedia, Dec. 2016.

[27] Y. M. S. Reddy, R. E. Ravindran and K. H. Kishore, "Diabetic retinopathy through retinal image analysis: A review," *Int. J. of Engineering Technology,* vol. 7, no. 1-5, p. 19, 2017.

[28] B. a. Antal and A. a. Hajdu, "An ensemble-based system for automatic screening of diabetic retinopathy," *Knowledge-Based Systems, Elsevier,* vol. 60, pp. 20-27, April 2014.

[29] L. Ryan, T. Y. Wong, and C. Sabanayagam. "Epidemiology of diabetic retinopathy, diabetic macular edema and related vision loss," Eye and Vision 2 (2015): 17. PMC. Web. 5 Sept. 2017.

[30] D. H. Hubel and T. N. Wiesel, "Receptive fields and functional architecture of monkey striate cortex," *J. Physiol*, vol. 195, pp. 215–24, Jun. 1968.

[31] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biological Cybernetics*, vol. 36, no. 4, pp. 193–202, 1980.

[32] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Conf. Neural Inf. Processing Syst. (NIPS)*, 2012.

[33] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient based learning applied to document recognition," *Proc. IEEE,* vol. 86, pp. 2278–2324, 1998.

[34] [Online]. Available: https://arxiv.org/abs/ 1207.0580. Accessed: Nov. 10, 2016.

[35] C. Szegedy et al., "Going deeper with convolutions," *IEEE Conf. on computer Vision and Pattern Recognition (CVPR)*, 2015.

[36] C. Szegedy et al., "Rethinking the Inception Architecture for Computer Vision," *IEEE Conf. on Computer Vision*, Dec. 2015.

[37] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters,* vol. 27, no. 6, pp. 861-874, Jun. 2006.

[38] [Online]. Available: https://kaggle2.blob. core. windows.net/forummessage attachments/88655 /2795/ competitionreport.pdf. Accessed: Jul. 5, 2017.

زیرنویس‌ها

1- Diabetic Retinopathy
2- Fundus images
3- Exudate
4- Microaneurysm
5- Hemorrhages
6- Neovascularization
7- Traditional Neural Network
8- Support vector machine
9- Naïve Bayes
10- K-nearest neighbors
11- Messidor database
12- Convolutional neural network
13- Deep learning
14- Kaggle Database
15- Paython
16- Diabetic ketoacidosis
17- Hypoglycemia
18- Diabetic nephropathy
19- Diabetic neuropathy
20- Proliferative
21- Macular edema
22- Optical coherence tomography
23- Fukushima
24- Pooling layer
25- Fully connected layer
26- Feature map
27- Local receptive field
28- Convolutional kernel
29- Activation function
30- Sigmoid
31- Rectified linear unit
32- Exponential linear unit
33- Pooling
34- Overfittinh
35- Max-pooling
36- Average-pooling
39- Adaptive Gradient Algorithm
40- Root Mean Square Propagation
41- Adaptive Moment Estimation
40- AlexNet
41- CifarNet
42- Accuracy
43- Predictive value
44- Precision
45- Sensitivity
48- True Positive (TP)
49- False Positive (FP)
50- True Negative (TN)
51- False Negative (FN)