

## تشخیص رتینوپاتی دیابتی با تحلیل اگزودا در تصاویر شبکیه با استفاده از یادگیری عمیق

ملیحه شارعی نیا<sup>(۱)</sup> سید محمد حسین معطر<sup>(۲)\*</sup>

(۱) گروه مهندسی کامپیوتر، واحد فردوس، دانشگاه آزاد اسلامی، فردوس، ایران

(۲) گروه مهندسی کامپیوتر، دانشگاه آزاد اسلامی، واحد مشهد، مشهد، ایران\*

### چکیده

رتینوپاتی دیابتی، عارضه‌ای ناشی از دیابت است که بدلیل تغییرات ایجاد شده در رگ‌های خونی رخ می‌دهد. هدف اصلی این مقاله، ارتقاء دقت تشخیص رتینوپاتی دیابتی در تصاویر شبکیه نسبت به روش‌های معمول مبتنی بر استخراج بردار ویژگی می‌باشد. در این روش ابتدا بر اساس داده‌های آموزشی لایه به لایه شبکه عصبی آموزش دیده و در نهایت شبکه عصبی کانولوشن با کمک چهار لایه کانولوشن، چهار لایه ادغام و دو لایه تماماً متصل آموزش دیده و ساخته می‌شوند. شبکه عصبی کانولوشن ویژگی‌های موجود در تصویر شبکیه را آموخته و ویژگی‌های مناسب را برای دسته‌بندی تصاویر استخراج می‌کند. در این مقاله توانستیم در آزمایش‌های خود به بهبود قابل قبولی نسبت به کارهای انجام شده‌ی قبلی برسیم که با توجه به بررسی ۳۹۷ نمونه از پایگاه داده Stare، برای آزمودن روش پیشنهادی و بدست آوردن نرخ حساسیت ۹۰٪ و صحت ۹۶٪ موفقیت روش بر روی این پایگاه داده کاملاً مشهود هست. **واژه‌های کلیدی:** یادگیری عمیق، شبکه‌های عصبی کانولوشن، استخراج ویژگی، رتینوپاتی دیابتی.

\* عهده‌دار مکاتبات:

نشانی: گروه مهندسی کامپیوتر، واحد مشهد، دانشگاه آزاد اسلامی، مشهد، ایران

تلفن: ۰۹۱۵۵۰۳۰۵۴۰ پست الکترونیکی: [moattar@mshdiau.ac.ir](mailto:moattar@mshdiau.ac.ir)

## ۱- مقدمه

تأثیر مخرب بیماری دیابت بر روی شبکه را رتینوپاتی دیابتی<sup>۱</sup> (DR) می‌نامند، که یک عارضه جدی دیابت است و عامل اصلی نابینایی در جمعیت فعال در کشورهای جهان است. تصویربرداری به منظور تشخیص بیماری رتینوپاتی می‌تواند منجر به درمان موفقیت‌آمیز و پیش‌گیری از نابینایی شود. با پیشرفت بیشتر این بیماری، اختلالات جدیدی در سطح شبکه آشکار می‌شود. وجود جراحات‌ها در سطح شبکه می‌تواند به‌طور جدی بینایی را مختل نماید. تشخیص اختلالات در شمار زیادی از تصاویر شبکه با روش تصویربرداری، زمان‌بر، پرهزینه و توأم با خطای انسانی است. اگر بیماری در مراحل اولیه تشخیص داده شود، درمان می‌تواند پیشرفت آن را کند و یا متوقف کند.

در این بیماری ابتدا عروق کوچک در پرده شبکه صدمه دیده و مایع یا خون از آنها نشت می‌کند. مایع نشت کرده باعث تورم پرده شبکه شده و یا رسوباتی به نام آگزودا<sup>۲</sup> را ایجاد می‌نماید. آگزودا ضایعات زرد-سفید با حاشیه نسبتاً متمایز و شامل چربی هستند. این ترشحات تنوع سطح خاکستری بالاتری نسبت به پس زمینه اطراف دارند و از این رو به عنوان الگوهای روشن در تصاویر فوندوس<sup>۳</sup> رنگی بسیار قابل مشاهده هستند [۱] در مراحل پیشرفته‌تر بیماری، رگ‌های خونی جدید و غیرطبیعی بر روی سطح شبکه رشد می‌کنند. این عروق جدید دیواره ضعیف‌تری داشته و شکننده هستند و ممکن است منجر به خونریزی شوند.

تشخیص رگ‌های خونی در شبکه و لکه‌های زرد رنگ یکی از موارد مورد بحث روز در پردازش تصویر پزشکی می‌باشد، که بوسیله تصویر فوندوس از شبکه و استخراج رگ‌های خونی و ترشحات زرد رنگ (آگزودا)، نسبت به تشخیص اقدام می‌گردد.

این‌که برای تشخیص درست و سریع جهت شناسایی رتینوپاتی دیابتی از چه روشی استفاده کنیم، موضوعی است که در این مقاله با استفاده از روش‌های یادگیری عمیق و

استخراج خودکار ویژگی‌ها، به آن پرداخته‌ایم. روش پیشنهادی شناسایی آگزودا از سایر بخش‌های تصویر است. بنابراین در این تحقیق برای تشخیص سریع، ابتدا تصاویر را به فضای رنگی مناسب تبدیل کرده و سپس پیش پردازش مورد نیاز را بر روی آن‌ها انجام داده‌ایم. در نهایت با استفاده از شبکه عصبی کانولوشن<sup>۴</sup>، ویژگی به صورت خودکار استخراج و رتینوپاتی دیابتی مشخص می‌شود. به این ترتیب در روش پیشنهادی برخلاف روش‌های ارائه شده دیگر که ویژگی‌های طراحی شده دستی استخراج می‌شوند، ویژگی‌های خوب را مستقیماً از پیکسل‌های خام و به صورت خودکار استخراج می‌کنیم.

## ۲- پیشینه تحقیق

شبکه عصبی کانولوشن یک شبکه عصبی پیشرو<sup>۵</sup> است که الگوی اتصالات بین نرون‌های آن از ساختار قشر بینایی پستانداران الهام گرفته شده است. این شبکه نوع خاصی از شبکه‌های عصبی چندلایه است که می‌تواند ویژگی‌های توپولوژیکی یک تصویر را استخراج کند و تشخیص مستقیم الگوهای بصری از پیکسل‌های یک تصویر با حداقل پیش پردازش را انجام دهد.

در هر شبکه عصبی کانولوشن دو مرحله برای آموزش وجود دارد. مرحله روبه‌جلو و مرحله پس‌انتشار<sup>۶</sup>. در مرحله اول تصویر ورودی به شبکه تغذیه می‌شود، که به صورت ضرب نقطه‌ای بین ورودی و پارامترهای هر نرون و نهایتاً اعمال عملیات کانولوشن در هر لایه است. در ساختار ارائه شده یک شبکه کانولوشن بین کل پیکسل‌های تصویر ورودی به اشتراک گذاشته می‌شود. داشتن قابلیت اشتراک وزن به طور قابل توجهی سبب کاهش تعداد پارامترهای آزاد آموزش‌پذیر شبکه و در نتیجه افزایش تعمیم‌پذیری می‌شود. سپس خروجی شبکه محاسبه می‌شود.

<sup>4</sup> Convolutional Neural Networks (CNNs)

<sup>5</sup> feed-forward

<sup>6</sup> back propagation

<sup>1</sup> Diabetic retinopathy

<sup>2</sup> Exudate

<sup>3</sup> Fundus images

## ۲-۱ پیشینه در حوزه رتینوپاتی دیابتی

در مقاله [۲]، بخش‌بندی سلسله مراتبی رگ‌های خونی شبکه بر اساس روش ترکیبی شبکه‌های عصبی کانولوشن و جنگل تصادفی (RF)<sup>۷</sup> انجام شده است. در این روش، CNN برای استخراج ویژگی‌های سلسله مراتبی تصاویر و RF برای طبقه‌بندی رگ‌ها استفاده شده است. این روش شامل سه مرحله اساسی پیش‌پردازش، استخراج ویژگی سلسله مراتبی و طبقه‌بندی کلی است. ابتدا تصاویر با اعمال هیستوگرام و فیلتر گوسی پیش‌پردازش می‌شوند. سپس به منظور کاهش افزونگی با کمک روش مبتنی بر سوپر پیکسل، تنها یک پیکسل به صورت تصادفی انتخاب می‌شود. سپس برای هر پیکسل هدف بدست آمده، مقادیر پیکسل خام از مرکز پنجره مربع مستقیماً به عنوان ورودی به CNN تغذیه می‌شود. با فیله‌های موجود محلی، نرون‌ها در CNN می‌تواند ویژگی‌های بصری ابتدایی مانند تشخیص لبه، نقاط پایانی و گوشه‌ها را تشخیص دهد. این ویژگی‌ها توسط لایه‌های متوالی به منظور گرفتن ویژگی‌های سطح بالاتر ترکیب می‌شوند. ورودی CNN با توجه به متوسط عرض و ناحیه‌های روشن و تاریک ضایعات (اگزودا)، یک پنجره مربع به اندازه ۲۵×۲۵ در نظر گرفته می‌شود. ویژگی آزمایش‌های بدست‌آمده از لایه‌های CNN به دسته‌بند RF وارد و در نهایت برنده هر RF پیش‌بینی می‌کند که آیا پیکسل مرکزی پنجره به رگ خونی متعلق است یا نه؟ برای آزمایش، پنجره کشویی نمونه‌برداری به عنوان تصویر آزمون، مستقیم وارد CNN می‌شود. سپس طبقه‌بند برنده با این ویژگی‌ها برای پیش‌بینی نتایج استفاده می‌شود.

در مقاله [۳] یک روش مؤثر جدید برای تشخیص اگزودا در تصاویر شبکه با استفاده از شبکه عصبی مصنوعی پیشنهاد شد، که با استفاده از شبکه عصبی چندلایه پرسپترون به بررسی وجود بیماری کمک می‌کند. در ابتدا برای پیش‌پردازش، بر روی تصاویر شبکه حذف نویز انجام و تصویر از فضای رنگ به مقیاس خاکستری تبدیل و فیلتر میانگین روی آن اعمال و سپس جهت افزایش کنتراست تصاویر

شبکه، هیستوگرام تطبیق روی آن اعمال می‌شود. سپس بر اساس ویژگی‌های رنگ و بافت تصویر، بردارهایی با ابعاد ۱۵ ویژگی از رنگ و بافت تصویر بدست می‌آوریم. این بردار ویژگی بر اساس فضای رنگ YUV و خصوصیات آماری از لبه‌های تصویر به دست آمده با روش Canny می‌باشد که با ترکیب ویژگی‌های به دست آمده از اعمال فیلترهای گابور، روشی مناسب برای استخراج ویژگی‌های مختلف موجود در تصویر ارائه می‌دهد. با استفاده از این ویژگی‌ها تصاویر شبکه مربوط به افراد سالم و بیمار با دقت بالایی قابل شناسایی می‌باشند. این بردارهای ویژگی توسط شبکه پرسپترون دو لایه دسته‌بندی شده و نتایج برای تشخیص و تقسیم‌بندی ترشحات در تصاویر شبکه، استفاده شده است.

در مقاله [۴] یک روش ثبت کارآمد تصاویر شبکه با وضوح بالا ارائه شد. هسته این روش، یک تشخیص ناحیه مؤثر برای تعیین تشابهات بود. آشکارساز ناحیه یک ساختار عروقی استخراج می‌کند که در آن پس‌زمینه شبکه به صورت قابل توجهی، برای شناسایی مناطق تحت تغییرات محتوا و نور بین جفت تصاویر متفاوت است. الگوریتم قادر به مقابله بهتر با تغییرات کنتراست کم، ناپایداری شدت و محتوا و نواحی همپوشانی کوچک تصاویر شبکه با وضوح بالا است. در این مقاله، نقاط متناظر به منظور برآورد پارامترهای انتقال با استفاده از روش حداقل مربعات بدست می‌آید [۵]. توابع چند جمله‌ای عیب هندسی بین تصاویر شبکه را اصلاح می‌کنند.

در مقاله [۶] یک سیستم غربالگری دیابتی که قابلیت ارزیابی کیفیت تصاویر شبکه را داشت، توسط الگوریتم تحلیل اجزای مورفولوژیکی<sup>۸</sup> (MCA) بین ساختارهای شبکه نرمال و آسیب دیده تفاوت قائل می‌شود. همچنین اولین بار الگوریتم پیش‌پردازش برای ارزیابی کیفیت تصاویر شبکه استفاده شده است. در نهایت تصاویر طبیعی و غیرطبیعی شبکه توسط ویژگی‌های آماری از ضایعات شبکه مشخص می‌شود. در این الگوریتم، عروق شبکه و ضایعات با استفاده

<sup>۸</sup> Morphological Component Analysis

<sup>۷</sup> random forests (RF)

از الگوریتم MCA جدا می‌شوند. این الگوریتم قادر به شناسایی ضایعات مختلف مرتبط به DR مانند میکروآنوریزم، خونریزی و ترشحات می‌باشد.

در [7] یک روش مبتنی بر عکس‌های فوندس رنگی برای تشخیص رتینوپاتی دیابتی ارائه شده است. جهت پیش‌پردازش، برای تصحیح غیر یکپارچگی‌های موجود در تصویر، اعمال یکسان سازی هیستوگرام انجام می‌شود. سپس برای پردازش ساختار شیء‌های داخلی تصویر پیش‌پردازش شده، از پردازش مورفولوژیک استفاده می‌شود. بعد از اعمال عملگرهای مورفولوژیکال، تصویر اولیه از تصویر بدست آمده کم می‌شود. پس زمینه تصویر پردازش شده به اندازه تصویر اصلی دارای نویز نیست و رگ‌ها می‌توانند به صورت بهتر دیده شوند. در این صورت محیط و مکان ویژگی‌ها می‌توانند به آسانی با اعمال یک آستانه تعیین شوند که پس زمینه را سیاه و رگ‌ها را سفید می‌کند. در کانال سبز قدرت تمیز دادن اگزوداها بالا است، بنابراین رگ‌ها با عملگرهای بسته مورفولوژیکال با استفاده از عنصر ساختاری هشت ضلعی حذف می‌شوند.

در مقاله [8] مجموعه‌ای از ویژگی‌های مختلف از بافت و رنگ تصاویر شبکه استخراج شده است و در نهایت از آزمون فرضیه برای تعیین بیماری یا عدم بیماری استفاده شده است. هدف از این پژوهش ارائه ابزاری کارآمد جهت استفاده در کاربردهای واقعی بوده است. در مقاله [9] نیز علاوه بر مراحل به‌سازی تصویر و استخراج ویژگی‌ها، از یک رویکرد یادگیری ماشین تحت عنوان جنگل تصادفی تنظیم شده به منظور دسته‌بندی تصاویر شبکه استفاده شده است.

در مقاله [10] از ویژگی ضرایب ویولت بر روی دیسک نوری شبکه به منظور تشخیص استفاده شده است. همچنین پس از انتخاب ویژگی به کمک روش تحلیل مولفه‌های اصلی (PCA)، تعدادی از روش‌های یادگیری ماشین از جمله SVM، KNN، جنگل تصادفی و ... جهت دسته‌بندی تصاویر

مورد استفاده قرار گرفته است. در مطالعه دیگری [11]، از یادگیری دیکشنری و بازنمایی تنک به منظور تعیین محل رگهای در تصاویر شبکه و تشخیص بیماری رتینوپاتی دیابتی استفاده شده است. تفاوت این رویکرد با رویکردهای پیشین در این بوده است، که روش‌های پیشین از دیکشنری ثابت استفاده می‌کرده‌اند، اما در [11] دیکشنری از روی نمونه‌های موجود آموزش می‌بیند.

## ۲-۲ پیشینه در حوزه یادگیری عمیق

در مقاله [12] یکی از روش‌های جالب مدیریت لایه‌های کانولوشنی تحت عنوان روش شبکه در شبکه<sup>۹</sup> ارائه شده است، که در آن ایده اصلی جایگزینی لایه کانولوشنی با یک شبکه عصبی پرسپترون کوچک شامل چندین لایه تماماً متصل با توابع فعال سازی غیرخطی مطرح شده است. به این ترتیب فیلترهای خطی با شبکه‌های عصبی غیرخطی جایگزین می‌شوند. این روش باعث بدست آوردن نتایج خوبی در دسته‌بندی تصاویر می‌شود.

مقاله [13] مقایسه‌ای بین دو عملیات ادغام ماکزیمم<sup>۱۰</sup> و ادغام میانگین<sup>۱۱</sup> ارائه داد و بیان کرد که ادغام ماکزیمم می‌تواند باعث همگرایی سریع‌تر، تعمیم بهتر (بهبود تعمیم‌دهی) و انتخاب ویژگی‌های نامتغیر بسیار عالی<sup>۱۲</sup> شود. در مقاله [13] با استفاده از آموزش یک شبکه CNN بزرگ بر روی دادگان ImageNet و کسب بهترین نتیجه در رقابت ImageNet سال ۲۰۱۲، توانست اثبات کند که شبکه‌های عصبی کانولوشن علاوه بر تشخیص ارقام دست‌نویس [14] بخوبی می‌توانند در حوزه دسته‌بندی تصاویر هم عمل کنند. به منظور درک بهتر معماری‌های شبکه عصبی کانولوشن، مقاله [15] یک روش بصری‌سازی<sup>۱۳</sup> معروف برای معماری‌های شبکه عصبی کانولوشن ارائه کرد که باعث ارائه دیدی از آنچه در حال اتفاق افتادن در درون یک شبکه عصبی کانولوشن بود شد. این روش با مشخص کردن

<sup>12</sup> superior invariant features

<sup>13</sup> visualization technique

<sup>9</sup> Network in Network (NIN)

<sup>10</sup> max pooling

<sup>11</sup> average pooling

الگوهای قابل تفسیر توانست امکاناتی جهت بهتر سازی طراحی معماری فراهم سازد.

### ۳- چارچوب پیشنهادی

هدف اصلی این تحقیق بررسی این موضوع است که آیا شبکه عمیق کانولوشن می‌تواند دقت تشخیص بیماری رتینوپاتی دیابتی را نسبت به روش‌های معمول ارتقا دهد یا نه. در روش پیشنهادی ابتدا بر اساس داده‌های آموزشی لایه به لایه شبکه عصبی آموزش دیده و در نهایت تشکیل یک شبکه چند لایه از لایه‌های کانولوشن و لایه‌های متصل را می‌دهد. در نهایت شبکه عصبی آموزش دیده برای دسته‌بندی تصاویر شبکه‌ی متعلق به مجموعه آزمایش به کار می‌روند.

### ۳-۱ مرحله اول - پیش پردازش

در این قسمت بعد از بهبود کنتراست بر روی تصاویر رنگی فوندوس شبکه چشم به دلیل حجم بالای محاسبات در فرآیند آموزش و آزمایش، از کانال سبز تصویر رنگی به عنوان تصویر ورودی شبکه کانولوشن استفاده شده است. آزمایشات نشان دادند، که استفاده از کانال‌های دیگر تصویر رنگی یعنی کانال قرمز و آبی تاثیری در افزایش دقت شبکه کانولوشن نداشته و در عوض باعث افزایش تعداد پارامترهای شبکه و کند شدن فرآیند آموزش و اجرای شبکه کانولوشن خواهد شد.

### ۳-۲ مرحله دوم - یادگیری ویژگی‌ها با شبکه عصبی

#### کانولوشن

در این مقاله، یک شبکه عصبی کانولوشن با ساختار سلسله مراتبی استفاده شده است. شبکه‌های عصبی کانولوشن یکی از مهم‌ترین روش‌های یادگیری عمیق هستند، که در آن‌ها چندین لایه با روشی قدرتمند آموزش می‌بینند [۱۴]. این روش بسیار کارآمد بوده و یکی از رایج‌ترین روش‌ها در کاربردهای مختلف بینایی کامپیوتر است.

### ۳-۲-۱ الگوریتم آموزش شبکه عصبی کانولوشن

شبکه عصبی کانولوشن با توجه به این‌که بر اساس یادگیری با نظارت و در راستای کاهش خطای موجود میان خروجی شبکه و برچسب‌های تعیین شده فعالیت می‌کند، از الگوریتم‌های آموزش مبتنی بر پس‌انتشار خطا استفاده می‌نماید. در طول فرآیند آموزش، وزن‌های مشترک در لایه‌های کانولوشن و هم‌چنین وزن‌های موجود میان لایه‌های کاملاً متصل کانولوشن به طور قابل توجهی سبب کاهش تعداد پارامترهای آزاد آموزش‌پذیر شبکه و در نتیجه افزایش تعمیم‌پذیری می‌شود.

استفاده از شبکه‌های آموزش‌پذیر کوچک‌تر و بخش‌بندی شده برای حل یک مسئله بزرگ و پیچیده، یک معماری مقیاس‌پذیر برای پیاده‌سازی شبکه‌های بزرگ ارائه می‌دهد. ساختار ارائه شده سبب کاهش زمان آموزش، کاهش تعداد پارامترهای آموزش‌پذیر و افزایش دقت دسته‌بندی شده است.

بطور کلی، یک شبکه CNN از سه لایه اصلی تشکیل می‌شود که عبارتند از: لایه کانولوشن، لایه ادغام و لایه تماماً متصل. لایه‌های مختلف وظایف مختلفی را انجام می‌دهد.

### ۳-۲-۱-۱ لایه کانولوشن

لایه کانولوشن هسته اصلی تشکیل‌دهنده شبکه عصبی کانولوشن است که خروجی آن را می‌توان بصورت یک سه بعدی از نورون‌ها تفسیر کرد. پارامترهای لایه کانولوشن شامل مجموعه‌ای از فیلترهای قابل یادگیری هستند. هر فیلتر از لحاظ مکانی کوچک بوده و دارای سه بعد طول، عرض و عمق می‌باشد و در امتداد عمق ورودی ادامه پیدا می‌کند. یعنی در هر عمق در اصل یک ماتریس وجود دارد. بنابراین زمانی که ما می‌گوییم فیلتر مورد نظر در امتداد یا راستای عمق ادامه می‌یابد، به این معناست که این فیلتر بر روی تمامی ماتریس‌ها (برش‌های ورودی) اعمال می‌شود. در واقع در این لایه، فیلتر در تصویر کانوالو می‌شود. یعنی بر روی تصویر حرکت کرده و ضرب نقطه‌ای آن محاسبه می‌شود.

### ۳-۲-۱-۲ لایه ادغام

یک لایه ادغام معمولاً بعد از یک لایه کانولوشنی قرار می‌گیرد و از آن برای کاهش اندازه نقشه ویژگی‌ها و پارامترهای شبکه استفاده می‌شود. همان‌طور که می‌دانید ویژگی‌هایی که از یک قسمت از تصویر استخراج می‌شود، می‌تواند در قسمت‌های دیگر تصویر نیز به کار گرفته شود. برای بیان ویژگی‌های یک تصویر بزرگ یکی از روش‌ها جمع کردن آماره‌های این ویژگی‌ها در محل‌های مختلف تصویر است.

همانند لایه‌های کانولوشنی، لایه‌های ادغام بخاطر در نظر گرفتن پیکسل‌های همسایه در محاسبات خود، نسبت به تغییر مکان بی‌تغییر هستند. پیاده‌سازی لایه ادغام با استفاده از تابع ماکزیمم و تابع میانگین رایج‌ترین پیاده‌سازی‌ها هستند. بنابراین به وسیله لایه ادغام ابعاد خروجی لایه کانولوشن کم و حجم محاسبات پایین می‌آید.

### ۳-۲-۱-۳ لایه تماماً متصل

معمولاً لایه‌های آخر شبکه کانولوشن مشابه شبکه‌های پیشخور همچون پرسپترون تماماً متصل می‌باشند. در لایه‌های تماماً متصل، نرون‌ها یک اتصال کامل به تمام خروجی‌های لایه قبل دارند.

لایه‌های تماماً متصل، مشابه شبکه‌های عصبی مصنوعی معمولی هستند و نتیجه شبکه را در قالب یک بردار با اندازه مشخص ارائه می‌کند. از این بردار می‌توان برای دسته بندی تصاویر بهره برد به طوری که مشابه شبکه پرسپترون چند لایه، حداقل شامل دو لایه از نرون‌های تماماً هستند، که لایه آخر معمولاً به تعداد کلاس‌های موجود نرون دارد. هر بار یکی از این نرون‌ها نسبت به بیشترین سطح فعالیت فعال شده و در رقابت برنده می‌شود و کلاس برنده معین می‌شود.

### ۴- آزمایش‌ها و ارزیابی

سیستم مورد استفاده برای انجام آزمایشات و آموزش و ارزیابی شبکه کانولوشن یک لپ تاپ ASUS مجهز به پردازنده گرافیک Nvidia GeForce GTX 960M با ۴ گیگا

بایت حافظه داخلی و ۱۶ گیگا بایت حافظه RAM بوده است. در این تحقیق از بانک تصاویر Stare برای انجام تحقیقات استفاده شده است. این بانک تصاویر شامل ۴۰۲ تصویر بوده، که اخیراً به ۳۹۷ تصویر تقلیل پیدا کرده است. تعداد تصاویر برای آموزش شبکه کانولوشن کم بوده و نمی‌توان به دو بخش آموزش و آزمایش تقسیم کرد. بدین سبب با انجام چرخش حداکثر ۱۰ درجه در جهت عقربه‌های ساعت یا برعکس، که مقدار و جهت آن به طور تصادف تعیین می‌شود، تصاویر جدیدی ایجاد شده و به مجموعه تصاویر اولیه اضافه شد. از میان تصاویر موجود، به طور تصادف ۷۰٪ تصاویر را برای آموزش و ۳۰٪ را به عنوان مجموعه آزمایش در نظر گرفتیم. شکل ۱ یک نمونه از تصاویر موجود در بانک تصاویر به کاررفته را نمایش می‌دهد.



شکل ۱: تصویر رنگی مربوط به شبکه از دادگان STARE با محتوای آگزودا که در آموزش به کار رفته است [۱۶].

اندازه تصاویر برای ساخت شبکه کانولوشن بزرگ است و تعداد پارامترهای به کاررفته به حدی زیاد است که حافظه پردازنده گرافیک را پر کرده و فرآیند آموزش، که توسط پردازنده گرافیکی انجام می‌شود، با مشکل مواجه می‌شود. بنابراین اندازه تصاویر را به [۱۰۰ ۱۰۰] تغییر دادیم. شبکه کانولوشن مورد نظر بر اساس آزمون و خطا، با ساختار زیر در نظر گرفته شده است، که جدول ۱ جزئیات ساختاری

شبکه عصبی کانولوشن را نمایش می‌دهد. آموزش شبکه عصبی ۴۰۰ مرحله تکرار شده و با نرخ یادگیری ۰,۰۰۱ و روش بهینه‌سازی گرادیان تصادفی با ترم ممنت، فرآیند آموزش وزن‌های شبکه کانولوشن انجام شده است.

#### ۴-۱- معیارهای ارزیابی

به منظور ارزیابی نتایج معمولاً از معیارهای ارزیابی استاندارد استفاده می‌شود. مشهورترین و عمومی‌ترین معیار برای تعیین کارایی یک الگوریتم دسته‌بند صحت<sup>۱۴</sup> است.

در این معیار اصلاحات زیر را داریم: FP تعداد تصاویر مربوط به افراد سالم است که به اشتباه به عنوان بیمار تشخیص داده شده‌اند. FN تعداد تصاویر مربوط به افراد سالم است که به اشتباه، مربوط به افراد بیمار تشخیص داده شده‌اند. TP تعداد تصاویر شبکه‌ی مربوط به بیماری است که به درستی تشخیص داده شده است و TN تعداد تصاویر مربوط به افراد سالم است که به درستی تشخیص داده شده‌اند.

جدول ۱: نمایش جزئیات ساختار شبکه کانولوشن در روش پیشنهادی

تعداد پارامترها	گام حرکت	نوع لایه	لایه
8470		لایه کانولوشن با ۷۰ فیلتر ۱۱ در ۱۱	۱
		لایه تابع فعالیت خطی	۲
	[2 2]	ماکزیمم گیری	۳
1750		لایه کانولوشن با ۱۴۰ فیلتر ۵ در ۵	۴
		لایه تابع فعالیت خطی	۵
	[2 2]	ماکزیمم گیری	۶
2520		لایه کانولوشن با ۲۸۰ فیلتر ۳ در ۳	۷
		لایه تابع فعالیت خطی	۸
	[2 2]	ماکزیمم گیری	۹
2880		لایه کانولوشن با ۳۲۰ فیلتر ۳ در ۳	۱۰
		لایه تابع فعالیت خطی	۱۱
	[2 2]	ماکزیمم گیری	۱۲
57600		لایه از نرون‌های کاملاً متصل به تعداد ۲۰ عدد	۱۳
		لایه تابع فعالیت خطی	۱۴
40		لایه از نرون‌های کاملاً متصل به تعداد ۲ عدد	۱۵
		لایه تابع فعالیت خطی	۱۶
		لایه تابع فعالیت softmax	۱۷
		لایه دسته بندی کننده	۱۸

در نهایت درصد تصاویری که به درستی به عنوان بیمار تشخیص داده شده است<sup>۱۵</sup> به صورت رابطه (۱):

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (1)$$

و درصد تصاویری که مربوط به بیماری نمی‌باشند یا به عبارتی سالم بوده و به درستی تشخیص داده شده‌اند<sup>۱۶</sup> به صورت رابطه (۲):

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (2)$$

<sup>16</sup> specificity

<sup>14</sup> Accuracy

<sup>15</sup> Sensitivity

محاسبه می‌شوند. با استفاده از رابطه‌های بالا، رابطه (۳) میزان صحت دسته‌بندی را نشان می‌دهد.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

#### ۴-۲- نتایج به دست آمده با ساختارهای مختلف شبکه

##### کانولوشن

شبکه کانولوشن با توجه به تعداد کم تصاویر آموزشی عملکرد خوبی از خود نشان داده و نتایج بدست آمده قابل قبول می‌باشد. بدین جهت پروسه ارزیابی را ۱۰ بار تکرار کرده و به ازای داده‌های مختلف میانگین صحت شناسایی را در جدول ۲ و به ازای ساختارهای مختلف شبکه کانولوشن مورد ارزیابی قرار داده‌ایم. تصمیم در مورد ساختارهای مختلف با تعداد لایه‌ها و نرون‌های مختلف بر اساس پارامترهای مختلف، خود به نوعی مسئله بهینه‌سازی محسوب می‌شود. ولی به دلیل کندی و حجم محاسبات شبکه کانولوشن، ناگزیر با آزمون و خطا بر آن شدیم تا ساختار مناسب را ارزیابی و انتخاب کنیم. نتایج جدول ۲ بر اساس ساختارهای مختلف شبکه عصبی کانولوشن به دست آمده‌اند. همان‌طور که مشاهده می‌شود بهترین نتایج متعلق به شبکه‌ای با ساختار [70 140 280 320] است، در حالی که شبکه‌ای با ساختار بزرگ‌تر [70 140 280 320 400] نتایج بدتری به دست آورده است. این مشاهده اثبات‌کننده این موضوع است که تعیین ساختار مناسب شبکه عصبی برای حل مسائل مختلف، خود به نوعی یک مسئله بهینه‌سازی چالش برانگیز می‌باشد. حجم محاسباتی بالا و زمان زیاد مورد نیاز برای آموزش شبکه‌های عصبی کانولوشن مانع از استفاده از روش‌های بهینه‌سازی همچون الگوریتم ژنتیک برای تعیین ساختار مناسب می‌شود. در حال حاضر محققان بیشتر بر حسب تجربه و آزمون و خطا ساختار مناسب این

گونه شبکه‌های عصبی را تعیین می‌کنند. همان‌طور که در جدول ۲ ردیف ۴ می‌بینیم، ساختار این شبکه مشابه ساختار مناسب به دست آمده بر اساس جدول ۲ ردیف ۱ می‌باشد، ولی به دلیل تعداد نرون‌های کم‌تر در لایه متصل اول به دقت کمتری رسیده است؛ که این امر خود نشان‌دهنده اهمیت لایه‌های متصل در نتایج به دست آمده می‌باشد.

#### ۴-۳ مقایسه روش پیشنهادی با روش‌های پیشین

برای اثبات کارایی روش پیشنهادی، روش ارائه شده در مرجع [۳] را بر روی تصاویر به کاررفته در این تحقیق، مورد ارزیابی قرار دادیم. روش ارائه شده در این مرجع بر اساس ویژگی‌های رنگ و بافت تصویر بوده و ۱۵ ویژگی متفاوت از رنگ و بافت تصویر استخراج می‌کند. این بردار ویژگی بر اساس ویژگی‌های رنگ فرمت YUV و خصوصیات آماری لبه‌های تصویر استخراج می‌شوند. در واقع ویژگی‌های استخراج شده در این شبکه بر اساس اطلاعات لبه، مشتق و ویژگی‌های ساختاری موجود در تصویر بدست آمده‌اند و این بردارهای ویژگی، جهت دسته‌بندی، به شبکه پرسپترون دو لایه وارد می‌شوند. نتایج به دست آمده از این روش در جدول ۳ ارائه شده است. لایه‌های آخر شبکه کانولوشن نیز عملکردی مشابه شبکه پرسپترون چند لایه در دسته‌بندی بردارهای ویژگی دارند، و بردارهای ویژگی در لایه‌های اول شبکه کانولوشن با انجام پالایش در لایه‌های متوالی به دست می‌آید. این شبکه از اطلاعات رنگ تصویر استفاده نمی‌کند و تنها از کانال سبز تصویر رنگی برای انجام محاسبات خود بهره می‌برد. در واقع ویژگی‌های استخراج شده در شبکه کانولوشن بر اساس اطلاعات لبه و مشتق و ویژگی‌های ساختاری موجود در تصویر می‌باشد. با این حال می‌بینیم که بدون بهره بردن از اطلاعات رنگ موجود در تصویر، نتیجه به دست آمده از شبکه کانولوشن بسیار بهتر از روش مرجع ذکر شده می‌باشد.



جدول ۲: نتایج حاصل از طبقه‌بندی تصاویر با استفاده از شبکه کانولوشن با ساختار تعیین شده در دادگان stare

ردیف	ساختار لایه‌های کانولوشن	ساختار لایه‌های متصل	تعداد تصاویر ارزیابی	تعداد تصاویر آموزشی	Accuracy	Specificity	Sensitivity
۱	[70 140 280 320]	[202]	۲۳۹	۵۵۵	٪۹۶	٪۹۸	٪۹۰
۲	[70 140 280 320 400]	[20 2]	۲۳۹	۵۵۵	٪۸۲	٪۹۸	٪۵۵
۳	[70 140 280]	[20 2]	۲۳۹	۵۵۵	٪۹۲	٪۹۸	٪۶۴
۴	[70 140 280 320]	[10 2]	۲۳۹	۵۵۵	٪۹۴	٪۹۷	٪۷۷

ارزیابی‌های خود، از تصاویر به کار رفته در این تحقیق برای انجام طراحی روش [۱۷] و ارزیابی آن استفاده شده است. میانگین سه نتیجه از بهترین نتایج گزارش شده بر اساس تعداد خوشه‌های به کار رفته، در جدول ۳ گزارش شده است. یکی از مزیت‌های روش پیشنهادی ما این است که مجموعه ویژگی‌ها را به صورت خودکار از خود داده‌ها به جای استفاده از ترکیبات مختلف ویژگی‌های طراحی دستی استخراج می‌کند.

در مرجع [۱۷] از ویژگی‌های SIFT برای شناسایی نقاط مربوط به بیماری استفاده شده است. با استخراج این ویژگی‌ها و خوشه‌بندی بردارهای ویژگی و ایجاد مجموعه‌ای از خوشه‌ها به عنوان مجموعه منتخب، بانکی از ویژگی‌ها به وجود می‌آید. این ویژگی‌های منتخب به عنوان ستون‌های هیستوگرام در نظر گرفته شده و فراوانی ضرایب SIFT بر روی ویژگی‌های منتخب، تشکیل یک بردار ویژگی را می‌دهد. دسته‌بندی تصاویر بوسیله دسته‌بندی بردارهای ویژگی نهایی توسط ماشین بردار پشتیبان انجام می‌شود. در

جدول ۳: مقایسه نتایج به دست آمده از دسته‌بندی بردارهای ویژگی با روش‌های مرجع و روش پیشنهادی

روش مورد نظر	تعداد تصاویر ارزیابی	تعداد تصاویر آموزشی	Specificity	Sensitivity	Accuracy
روش پیشنهادی	۲۳۹	۵۵۵	۹۸٪	۹۰٪	۹۶٪
روش مرجع [۳]	۲۳۹	۵۵۵	۸۲٪	۳۳٪	۸۱٪
روش مرجع [۱۷]	۲۳۹	۵۵۵	۹۳٪	۸۸٪	۹۴٪

افزایش داده است، ولی روال آموزش شبکه‌های عمیق بخصوص کانولوشن زمان‌بر بوده و بدون سخت‌افزارهای قوی و گران‌قیمت امکان‌پذیر نمی‌باشد. از طرفی این شبکه‌ها نیازمند مقادیر زیادی از داده‌ها برای آموزش می‌باشند و تولید و یا دسترسی به این گونه داده‌ها به راحتی امکان‌پذیر نیست. به عنوان کارهای آتی می‌توان به روش‌های مختلف برای ارتقاء شبکه کانولوشن هم‌چون ساختارهای جدید و یا الگوریتم‌های آموزش بهتر اشاره کرد. هم‌چنین برای رفع مشکلاتی هم‌چون نیازمندی به داده‌های زیاد می‌توان بر روی روش‌های یادگیری بدون ناظر و یادگیری با داده‌های کم‌تر

## ۵- بحث و نتیجه‌گیری

در این تحقیق روشی برای تشخیص بیماری بر اساس شبکه عصبی کانولوشن ارائه شد. شبکه عصبی کانولوشن ویژگی‌های مناسب برای دسته‌بندی تصاویر را می‌آموزد. در این تحقیق از شبکه عصبی عمیق کانولوشن برای حل مسئله تشخیص بیماری رتینوپاتی دیابتی استفاده شد و به دقتی بالاتر از روش‌های معمول قبلی همچون شبکه‌های عصبی چندلایه رسید. در ساختار شبکه از چهار لایه کانولوشن و دو لایه کاملاً متصل استفاده شده است. یادگیری عمیق اگرچه دقت بسیاری از مسائل از جمله دسته‌بندی تصویر را

نظارت و دخالت کارشناس و متخصص انسانی بتوانند به طور آنلاین دانش موردنیاز از محیط خود را کسب کرده و هوشمندانه وظایف و فعالیت‌های خود را انجام دهند.

اشاره کرد. هم‌چنین با ترکیب ساختارها و روش‌های مختلف یادگیری عمیق می‌توان به شبکه‌های قوی‌تر و کاراتری رسید. در آینده نیاز به الگوریتم‌هایی خواهیم داشت که بدون

## ۶- مراجع

- [1] A. Tariq, M.U. Akram, A. Shaukat and S.A. Khan. "Automated detection and grading of diabetic maculopathy in digital retinal images". J. Digit. Imaging, vol. 26, pp. 803-812, 2013.
- [2] S. Wang, Y. Yin, G. Cao, B. Wei, Y. Zheng and G. Yang. "Hierarchical retinal blood vessel segmentation based on feature and ensemble learning." Neurocomputing, vol.149, pp.708-717, 2015.
- [3] S.W. Franklin and S.E. Rajan. "Diagnosis of diabetic retinopathy by employing image processing technique to detect exudates in retinal images." IET Image Processing, vol. 8, no. 10, pp. 601-609, 2014.
- [4] Z. Ghassabi, J. Shanbehzadeh and A. Mohammadzadeh. "A structure based region detector for high resolution retinal fundus image registration." Biomedical Signal Processing and Control, vol. 23, pp. 52-61, 2016.
- [5] Z. Ghassabi, B. Moaveni and A. Khaki-Sedigh. "Solving systems of linear equations and finding the inversion of a matrix by neural network using genetic algorithms." in 5th WSEAS Int. Conf. on Computational Intelligence, Man-Machine Systems and Cybernetics, 2006, pp. 129-132.
- [6] E. Imani, H.R. Pourreza and T. Banaee. "Fully automated diabetic retinopathy screening using morphological component analysis." Computerized medical imaging and graphics, vol. 43, pp. 78-88, 2015.
- [7] J. Nayak, P.S. Bhat, U.R. Acharya, C.M. Lim and M. Kagathi. "Automated identification of diabetic retinopathy stages using digital fundus images." Journal of Medical Systems, vol. 32, no. 2, pp. 107-115, 2007.
- [8] M. E. Gegundez-Arias et. al. "A tool for automated diabetic retinopathy pre-screening based on retinal image computer analysis". Computers in Biology and Medicine, vol. 88, pp. 100-109, 2017.
- [9] G. Leontidis. "A new unified framework for the early detection of the progression to diabetic retinopathy from fundus images". Computers in Biology and Medicine, vol. 90, pp. 98-115, 2017.
- [10] A. Singh, M. Kishore Dutta, M. Partha Sarathi, V. Uher and R. Burget. "Image processing based automatic diagnosis of glaucoma using wavelet features of segmented optic disc from fundus image". Computer Methods and Programs in Biomedicine, vol. 124, pp. 108-120, 2016.
- [11] M. Javidi, H.R. Pourreza and A. Harati. "Vessel segmentation and microaneurysm detection using discriminative dictionary learning and sparse representation". Computer Methods and Programs in Biomedicine, vol. 139, pp. 93-108, 2017.
- [12] H. Lee, C. Ekanadham, A. Y. Ng. "Sparse deep belief net model for visual area V2." in Proceeding of the 20th International conference on Neural Information Processing Systems, 2007, pp. 873-880.
- [13] D. Schere, A. Mulle and S. Behnk. "Evaluation of pooling operations in convolutional architectures for object recognition" in International Conference on Artificial Neural Networks, 2010, pp. 92-101.

- [14] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner. "Gradient-based learning applied to document recognition" in Proceedings of the IEEE, vol. 86, no. 11, 1998.
- [15] M.D. Zeiler and R. Fergus. "Visualizing and understanding convolutional networks." in European Conference on Computer Vision, 2014, pp. 818-833.
- [16] M. H. Goldbaum. "Structured analysis of the Retina." Internet: <http://cecas.clemson.edu/~ahoover/stare>. 1975
- [17] S.A.G. Naqvi, M.F. Zafar and I. ulHaq. "Referral system for hard exudates in eye fundus." Computers in Biology and Medicine, vol. 64, pp. 217-235, 2015.