

یک روش کارا برای غربالگری اولیه بیماری گلوکوم بر اساس محاسبه نسبت کاپ به دیسک نوری با استفاده از شبکه‌های عصبی کانولوشنی

فاطمه معادی^۱، ندا فرجی^۲ و محمدرضا حسن‌نژاد بی‌بالان^۳

چکیده

تشخیص و درمان به موقع بیماری گلوکوم می‌تواند از پیشروی این بیماری و از دست دادن ناگهانی بینایی جلوگیری کند. گلوکوم دیسک و کاپ نوری را که درون دیسک قرار دارد تحت تأثیر قرار می‌دهد. در این مقاله ابتدا به محلی‌سازی دیسک نوری پرداخته شده و سپس قطعه‌بندی دیسک و کاپ نوری به منظور تشخیص براساس نسبت کاپ به دیسک نوری (CDR) انجام می‌شود. برای مرحله محلی‌سازی دیسک از شبکه عصبی کانولوشنی مبتنی بر ناحیه سرعتر (Faster-RCNN) با شبکه پیش‌آموزش دیده ResNet50 استفاده می‌شود. مرحله قطعه‌بندی توسط معماری U-Net اصلاح شده با استفاده از شبکه SE-ResNet50 در قسمت کدگذار آن اجرا شده و در نهایت CDR مورد ارزیابی قرار می‌گیرد. پایگاه‌های داده DRISHTI-GS1 و RIM-ONE v3 برای آموزش و تست روش پیشنهادی و دسته‌داده MESSIDOR صرفاً در مرحله تست بکار رفته است. همچنین به منظور قطعه‌بندی کاپ و دیسک نوری دو رویکرد برای در نظر گرفتن حاشیه‌نویسی کاپ و دیسک نوری در حقیقت مبنای دسته‌داده DRISHTI-GS1 پیشنهاد می‌شود. در رویکرد پیشنهادی دوم و طبق معیار F_1 -score، نتیجه قطعه‌بندی کاپ و دیسک نوری برای دسته‌داده DRISHTI-GS1 به ترتیب ۰٫۹۳ و ۰٫۹۷، برای دسته‌داده RIM-ONE v3 به ترتیب ۰٫۷۹ و ۰٫۹۵ و برای دسته‌داده MESSIDOR به ترتیب ۰٫۸۴ و ۰٫۹۳ بدست آمده که قابل رقابت با سایر کارها است.

کلیدواژه‌ها

یادگیری عمیق، شبکه‌های عصبی کانولوشنی، گلوکوم، نسبت کاپ به دیسک نوری، قطعه‌بندی

۱ مقدمه

مطابق با اعلام سازمان بهداشت جهانی، گلوکوم یا آب سیاه دومین

این مقاله در اردیبهشت‌ماه سال ۱۳۹۹ دریافت، در آبان‌ماه بازنگری و در آذرماه پذیرفته شد.

^۱ دانش‌آموخته کارشناسی ارشد مهندسی برق گرایش مخابرات-سیستم، دانشگاه بین‌المللی امام خمینی (ره)، قزوین، ایران

رایانامه: f.maadi@edu.ikiu.ac.ir

^۲ گروه مهندسی برق، دانشکده فنی و مهندسی، دانشگاه بین‌المللی امام خمینی (ره)، قزوین، ایران

رایانامه: nfaraji@eng.ikiu.ac.ir

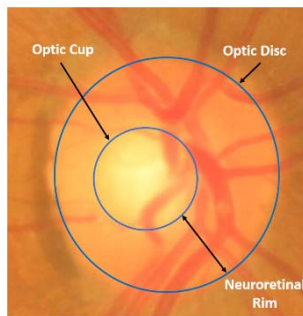
^۳ گروه مهندسی برق، دانشکده فنی و مهندسی، دانشگاه بین‌المللی امام خمینی (ره)، قزوین، ایران

رایانامه: mhbibalan@ikiu.ac.ir

نویسنده مسئول: ندا فرجی

علت نابینایی در جهان بعد از آب مروارید است [۱]. علاوه بر تحلیل بینایی، فشار بالای داخل چشم و فقدان خونرسانی به عصب بینایی و آسیب عصب بینایی، نسبت کاپ به دیسک^۱ (CDR) بزرگ از دیگر علائم بیماری گلوکوم است. حتی اگر درمانی برای بازگشت بینایی از دست‌رفته وجود نداشته باشد، زوال بینایی با درمان زودرس کنترل می‌شود. تشخیص گلوکوم در محیط بالینی شامل اندازه‌گیری فشار داخل چشم، آزمایش میدان دید و معاینه دیسک نوری در تصاویر فوندوس است. اندازه‌گیری فشار داخل چشم به تنهایی راه مؤثری برای تشخیص گلوکوم نیست؛ زیرا برخی از بیماران مبتلا به گلوکوم ممکن است فشار چشم طبیعی داشته باشند. از طرف دیگر آزمایش میدان دید به تجهیزات ویژه‌ای نیاز دارد که ممکن است برخی از کلینیک‌ها از آن برخوردار

^۱ Cup to Disc Ratio



شکل ۱. نمایش ناحیه دیسک و کاپ نوری در تصویر فوندوس

مرحله قطعه‌بندی است که بعنوان مهم‌ترین بخش تأثیر بسزایی در طبقه‌بندی گلوکوم در برابر نرمال دارد. ذکر این مطلب حائز اهمیت است که قطعه‌بندی کاپ نوری در اثر چگالی بالای رگ‌های خونی و تغییرات به شدت تدریجی رنگ بین کاپ و حاشیه‌هایش بسیار دشوار است [۵].

در این مقاله استخراج ویژگی بطور کامل برعهده ماشین است و با قطعه‌بندی معنایی بطور خودکار مرزهای دقیق کاپ و دیسک نوری استخراج می‌شوند که این ناحیه‌ها برای تشخیص گلوکوم مورد استفاده قرار می‌گیرند. برای انجام قطعه‌بندی کاپ و دیسک نوری از شبکه‌های عصبی کانولوشنی^۳ (CNN) استفاده می‌شود.

CNN از جمله معروف‌ترین و موفق‌ترین معماری‌های یادگیری عمیق در حوزه آنالیز تصاویر بوده و لایه‌های کانولوشنی آن، هسته اصلی تشکیل دهنده این شبکه است. وظیفه اصلی لایه‌های کانولوشنی استخراج سلسله مراتبی از ویژگی‌های غیرخطی تصاویر است. در هر لایه کانولوشن تعدادی فیلتر بر سطح تصویر اعمال می‌شود و یکسری نگاشت ویژگی حاصل می‌شود که در لایه‌های پایین این ویژگی‌ها، سطح پایین هستند و با افزایش تعداد لایه‌ها سطوح ویژگی افزایش می‌یابند. ویژگی‌های استخراج شده خاصیت سلسله مراتبی دارند به این معنی که در لایه‌های ابتدایی، بطور مثال گوشه‌ها، خط‌ها و لبه‌ها یاد گرفته می‌شوند و در لایه‌های عمیق‌تر به ترتیب ویژگی‌هایی با سطوح بالاتر، همانند اشیاء درون تصویر، استخراج می‌شوند [۶].

در این کار بعنوان یک نوآوری در بخش قطعه‌بندی، در شاخه کدگذار U-Net مرسوم بجای استفاده از مسیر انقباضی^۴ کلاسیک، از شبکه پیش آموزش دیده SE-ResNet50 جهت استخراج محتوای تصویر استفاده می‌شود که با کدگشای کلاسیک شاخه گسترشی^۵ ترکیب می‌گردد. برای آموزش و تست شبکه پیشنهادی قطعه‌بندی از پایگاه‌های داده Drishti-GS1 و RIM-MESSIDOR استفاده شده است و نیز از دسته‌داده صرفاً در فاز تست شبکه پیشنهادی استفاده گردیده است. نوآوری دوم ما در مقاله بررسی نحوه در نظرگرفتن حقیقت مبنا به گونه‌های مختلف در بخش آموزش شبکه برای دسته‌داده Drishti-GS1 است و اثر رویکردهای مختلف در مقاله مقایسه شده است.

نباشند [۲]. بنابراین، باتوجه به افزایش سربار گلوکوم در سیستم مراقبت‌های بهداشتی جمعیتی و نیز توسعه دستگاه‌های تصویربرداری چشم‌پزشکی، روش‌های خودکار آنالیز تصویر نقش مهمی در سیستم‌های تصمیم‌گیری پشتیبان برای مدیریت جمعیت و غربالگری به‌منظور مراقبت اولیه در بیماری گلوکوم دارد.

از آنجا که گلوکوم در مراحل اولیه علامتی ندارد، نیاز به غربالگری در جمعیت زیادی از افراد احساس می‌شود که بررسی جمعیت زیاد به تعداد متخصص زیادی نیاز دارد. از طرفی خستگی متخصصین در غربالگری جمعیت بالا کاهش دقت تشخیص را به دنبال دارد. به‌علاوه زمان زیادی برای بررسی این جمعیت مورد نیاز است و حتی ممکن است با کمبود متخصص روبرو شد. در اینصورت پردازش تصویر پزشکی و بینایی ماشین می‌تواند بعنوان دستیار متخصص عمل کند و افراد بیمار بدون علامت را شناسایی نماید تا برای معاینه بعدی به متخصص مراجعه کرده و تحت درمان مناسب قرار گیرند. این مقاله برای غربالگری و بعنوان تشخیص زودرس گلوکوم در مراحل اولیه ارائه شده است.

درحالی‌که دیسک نوری^۱ (OD) و کاپ نوری^۲ (OC) در همه افراد وجود دارد (شکل ۱)، نسبت کاپ به دیسک نوری بزرگ از نشانه‌های گلوکوم است. به همین دلیل، رویکردهای مختلفی برای قطعه‌بندی کاپ و دیسک نوری برای تشخیص گلوکوم در تصاویر رنگی فوندوس اتخاذ شده است. برخی از آثار موجود در این حوزه فقط بر قطعه‌بندی کاپ نوری و یا دیسک نوری متمرکزند و برخی دیگر به محاسبه نسبت کاپ به دیسک (CDR) پرداخته‌اند [۳].

اندازه‌گیری دقیق CDR به تلاش زیادی برای به دست آوردن قطعه‌بندی مناسب دیسک نوری و کاپ نوری نیاز دارد. در حال حاضر، قطعه‌بندی کاپ و دیسک نوری در تصاویر فوندوس به دو روش مبتنی بر پردازش تصویر مرسوم و روش‌های یادگیری عمیق صورت می‌گیرد [۴]. کارهای ارائه شده در روش‌های مبتنی بر پردازش تصویر مرسوم نیازمند مهندسی دستی ویژگی است که طراحی این ویژگی‌های دستی کار مشکل و زمانبری است و به دانش در رابطه با دسته‌داده مورد نظر نیازمند است. یکی از مزایای روش‌های یادگیری عمیق این است که عمل یادگیری، بازنمایی و قطعه‌بندی کاپ و دیسک نوری به‌صورت خودکار و همزمان صورت می‌پذیرد و همانند روش‌های سنتی، نیازی به مهندسی دستی ویژگی ندارد.

در این مقاله، دو مرحله در نظر گرفته می‌شود: در مرحله اول به محلی‌سازی دیسک نوری و استخراج آن بعنوان ناحیه مورد علاقه می‌پردازیم. سپس در مرحله دوم با استفاده از روش پیشنهادی به قطعه‌بندی کاپ و دیسک نوری برای ارزیابی نسبت این دو و نهایتاً طبقه‌بندی تصاویر به دو دسته دارای گلوکوم و نرمال خواهیم پرداخت. شایان ذکر است که عمده تمرکز این مقاله روی

³ Convolutional Neural Network

⁴ Contracting path

⁵ Expanding path

¹ Optic Disc

² Optic Cup

سواستوپولسکی U-Net را برای قطعه‌بندی دیسک و کاپ نوری در تشخیص گلوکوم اصلاح کرد [۱۵]. سپس سواستوپولسکی و همکارانش شبکه مطرح شده در [۱۵] را بهبود بخشیدند و ساختار شبکه U-Net انباشته شده^۸ را پیشنهاد کردند که مبتنی بر یک شبکه متوالی قطعه‌بندی تصویر U-Net است [۱۶]. هوژواف‌ئو و همکارانش یک معماری عمیق به نام M-Net را مطرح کردند که قطعه‌بندی دیسک نوری و کاپ نوری را بطور توأم و در سیستم چند برچسب^۹ یک مرحله‌ای انجام می‌دهد [۱۷]. آنها برای انجام این کار تصاویر ورودی را به سیستم مختصات قطبی بردند. در مقاله فریرا و همکارانش یک روش برای تشخیص خودکار گلوکوم در تصاویر شبکیه با استفاده از یک روش یادگیری عمیق و همچنین اکتشاف ویژگی‌های بافت^{۱۰} از طریق شاخص‌های تنوع فیلوژنتیک^{۱۱} ارائه شد [۱۸]. در کار البندر و همکارانش یک روش جدید مبتنی بر یادگیری عمیق برای قطعه‌بندی دیسک و کاپ نوری مطرح شد که هسته اصلی آن DenseNet با شبکه تمام کانولوشنی است که این روش یعنی FC-DenseNet فقط روی یک دسته‌داده آموزش دید و روی چهار دسته‌داده دیگر ارزیابی شد [۱۹]. در مقاله یوم‌اینگ‌جیانگ و همکارانش با فرض اینکه شکل دیسک و کاپ نوری بیضوی است، قطعه‌بندی توأم دیسک و کاپ نوری با CNN یعنی joint-RCNN^{۱۲} مطرح شد [۲۰]. در این مقاله DPN^{۱۳} و CPN^{۱۴} به ترتیب برای تولید پیشنهادهای جعبه مرزی دیسک و کاپ ارائه شدند. مازول اهمیت دیسک^{۱۵} برای اتصال DPN و CPN بکار رفت. ناحیه دیسک نوری و کاپ نوری بیضی‌های محاط شده درون این جعبه‌های مرزی بودند و سپس نسبت کاپ به دیسک محاسبه شد. شوانگ یو و همکارانش U-Net را با ResNet34 پیش آموزش دیده و U-Net کلاسیک به‌منظور قطعه‌بندی کاپ و دیسک نوری اصلاح کردند که در آن از مرحله آموزش از پایه اجتناب می‌شود و روند آموزش با تعداد کمتر دوره‌های آموزشی سریع‌تر صورت می‌گیرد [۵]. در مقاله مولانا و همکارانش یک روش تمام خودکار برای غربالگری گلوکوم مطرح شد که در آن در ابتدا با استفاده از معیار روشنایی و روش تطبیق قالب به تشخیص دیسک نوری پرداخته شد و سپس قطعه‌بندی کاپ و دیسک نوری با استفاده از رویکرد مبتنی بر بافت انجام شد و در نهایت CDR محاسبه شد [۲۱]. در کار تاکور و جونجا یک رویکرد ترکیبی مبتنی بر خوشه‌بندی و تنظیم سطح^{۱۶} برای قطعه‌بندی کاپ نوری و دیسک نوری در تصاویر شبکیه ارائه شده است [۲۲]. روش آنها با خوشه‌بندی تصویر با استفاده از C

بسیاری از تکنیک‌های پردازش تصویر و یادگیری ماشین برای قطعه‌بندی و طبقه‌بندی دیسک نوری و کاپ نوری استفاده می‌شود که چالش‌هایی فنی برای بهبود دقت روش‌های قطعه‌بندی و طبقه‌بندی در آنها به چشم می‌خورد [۷]. دامون و همکارانش روشی را برای تشخیص خودکار مرزهای کاپ از طریق استفاده از خمش رگ‌ها در مرز کاپ نوری مطرح کردند که خمش رگ‌ها در مرز کاپ نوری بدست آمده با اطلاعات قطبی برای تشخیص کاپ نوری ترکیب شدند [۸]. در کار چنگ و همکارانش قطعه‌بندی کاپ و دیسک نوری با استفاده از طبقه‌بندی سوپریکسل برای غربالگری گلوکوم مطرح شد [۹]. نور و همکارانش به پیاده‌سازی چند آستانه‌گذاری رنگی^۱ روی تصویر فوندوس برای قطعه‌بندی کاپ و دیسک نوری به‌منظور اندازه‌گیری CDR پرداخته‌اند [۱۰]. روش آنها شامل پیش‌پردازش، پردازش و پس‌پردازش است. مرحله پیش‌پردازش شامل برش و جداکردن کانال رنگ است. مرحله پردازش، روش چند آستانه‌گذاری رنگ است و سرانجام مرحله پس‌پردازش، شامل تجزیه و تحلیل قطعه‌بندی رنگ با استفاده از روش منحنی مشخصه عملکرد سیستم^۲ (ROC) و متعاقباً محاسبه و ارزیابی اندازه‌گیری CDR است. دوتا و همکارانش یک روش مبتنی بر قطعه‌بندی کاپ و دیسک نوری و محاسبه نسبت کاپ به دیسک را پیشنهاد کردند [۱۱]. در کار آنها برای قطعه‌بندی کاپ و دیسک نوری یک روش دو آستانه‌ای یکی برای پاک‌کردن عروق خونی و پس‌زمینه و دیگری برای قطعه‌بندی پیکسل‌های با شدت رنگ بالای موجود در کاپ و دیسک نوری استفاده شد. به‌علاوه تبدیل Hough برای محاسبه شعاع دیسک و کاپ نوری استفاده شد. الغمدی و همکارانش یک رویکرد موثر برای اندازه‌گیری نسبت کاپ به دیسک برای غربالگری گلوکوم ارائه کردند [۱۲]. در این رویکرد ویژگی‌ها از سوپریکسل‌ها استخراج شدند و از طبقه‌بندهای شبکه‌های عصبی برای تخمین مرز کاپ و دیسک نوری استفاده شد. میتاپالی و همکارانش دو روش قطعه‌بندی را یکی برای دیسک نوری و دیگری برای کاپ نوری مطرح کردند [۱۳]. در کار آنها یک مدل کانتور فعال مبتنی بر ناحیه ضمنی^۳ برای قطعه‌بندی دیسک نوری معرفی شد و یک روش جدید مبتنی بر خوشه‌بندی^۴ SWFCM برای قطعه‌بندی کاپ نوری ارائه شد. زیلی و همکارانش استفاده از معماری شبکه عصبی کانولوشنی مبتنی بر یادگیری گروهی^۵ را برای قطعه‌بندی دیسک و کاپ نوری مطرح کردند [۱۴]. در کار آنها تکنیک نمونه‌برداری آنتروپی^۶ برای نقاط حاوی اطلاعات مورد استفاده قرار گرفت. نقاط نمونه‌برداری شده برای طراحی یک چارچوب یادگیری جدید برای فیلترهای کانولوشنی مبتنی بر تقویت^۷ استفاده شدند.

⁸ Stack-U-Net network structure

⁹ Multi-label

¹⁰ Exploration of texture attributes

¹¹ Phylogenetic diversity indexes

¹² Joint Region-based Convolutional Neural Network

¹³ Disc Proposal Network

¹⁴ Cup Proposal Network

¹⁵ Disc attention module

¹⁶ Level set

¹ Color multi-thresholding

² Receiver Operating Characteristic

³ Implicit region based active contour model

⁴ Spatially Weighted Fuzzy C Means

⁵ Ensemble Learning

⁶ Entropy Sampling

⁷ Convolutional filters based on boosting

۲ معماری چند شبکه عصبی کانولوشنی

در گام نخست شبکه‌های عصبی کانولوشنی بکار رفته در مرحله محلی‌سازی و قطعه‌بندی معرفی می‌شوند.

۱-۲ بلوک باقیمانده

قبل از معرفی شبکه‌های باقیمانده^۸ (ResNet) محققان پی برده بودند که با افزایش عمق شبکه عصبی می‌توان به کارایی بهتری دست پیدا کرد. اما از نظر تجربی نشان داده شده که آستانه بیشینه‌ای برای عمق مدل مرسوم CNN وجود دارد. هنگامی که شبکه‌های عمیق‌تر شروع به همگرایی می‌کنند، مشکل تنزل^۹ به وجود می‌آید یعنی با افزایش عمق شبکه، دقت^{۱۰} اشباع می‌شود و سپس به سرعت کاهش می‌یابد. این تنزل ناشی از بیش‌برازش نیست و اضافه کردن لایه‌های بیشتر به یک مدل مناسب عمیق منجر به خطای آموزش بالاتر می‌شود [۲۵].

برای حل این مشکل، معماری ResNet معرفی شد و در آن به جای آموزش توابع غیر ارجاع داده شده، لایه‌های قابل یادگیری به عنوان توابع باقیمانده با ارجاع به ورودی لایه‌ها، مطرح شد. در کار هی و همکارانش شواهد تجربی جامعی ارائه شد که نشان می‌دهد این شبکه‌های باقیمانده برای بهینه‌سازی، ساده‌تر هستند و می‌توانند با استفاده از عمق افزایش یافته دقت خوبی را کسب کنند [۲۵].

شبکه‌های باقیمانده عمیق متشکل از بسیاری از واحدهای انباشته شده‌اند. هر واحد را می‌توان با معادلات زیر بیان کرد [۲۶].

(۱)

$$y_l = h(x_l) + F(x_l, W_l)$$

(۲)

$$x_{l+1} = f(y_l)$$

که در آن x_l ویژگی ورودی به واحد باقی‌مانده l ام است، $W_l = \{W_{l,k} | 1 \leq k \leq K\}$ مجموعه‌ای از وزن‌ها (و بایاس‌ها) است که با واحد باقی‌مانده l ام مرتبط است و K تعداد لایه‌ها در یک واحد باقی‌مانده است (K در [۲۵] ۲ یا ۳ است). F بیانگر تابع باقی‌مانده است، به عنوان مثال، پشته‌ای از دو لایه کانولوشنی 3×3 . در [۲۵] تابع f عملگری بعد از جمع المانی است و f ، ReLU^{11} می‌باشد. تابع h به عنوان نگاشت همانی تنظیم شده که برابر با $h(x_l) = x_l$ است. این جریانات انتشار مستقیم اطلاعات توسط پیکان‌های خاکستری در شکل ۲ نشان داده شده‌اند.

میانگین فازی شهودی مبتنی بر هسته تنظیم شده تطبیقی^۱ (ARKIFCM) شروع می‌شود. بعلاوه، مرز تصویری که با استفاده از خوشه‌بندی C میانگین فازی شهودی مبتنی بر هسته تنظیم شده تطبیقی قطعه‌بندی می‌شود به عنوان کانتور اولیه به رویکرد تنظیم سطح داده می‌شود. در کار کیم و همکارانش دو مرحله برای برآورد ناحیه دیسک و کاپ نوری وجود دارد [۲۳]. اول، تشخیص ناحیه موردعلاقه^۲ (ROI) با استفاده از یک شبکه عصبی کانولوشن معمولی (CNN) از تصاویر فوندوس که دیسک نوری در مرکز آن قرار دارد. دوم، قطعه‌بندی دیسک و کاپ نوری از ROI. برای مرحله دوم دو شبکه تمام کانولوشنی^۳ استفاده شده است: FCN2 (یک معماری U-Net برای قطعه‌بندی دو کلاسه) و FCNM (یک معماری U-Net برای قطعه‌بندی چندکلاسه). FCN2 برای آزمایش‌های قطعه‌بندی دو کلاسه و FCNM برای قطعه‌بندی همزمان دیسک و کاپ نوری استفاده می‌شود. در کار آنها سه راه مختلف برای قطعه‌بندی کاپ و دیسک نوری ارائه شده است: (۱) قطعه‌بندی دیسک نوری از ROI با آموزش یک FCN2 و قطعه‌بندی کاپ از ROI با آموزش یک FCN2 دیگر، (۲) قطعه‌بندی دیسک نوری از ROI با آموزش FCN2 و قطعه‌بندی کاپ از ROI ماسک شده با دیسک نوری از طریق آموزش FCN2 دیگر (۳) قطعه‌بندی دیسک و کاپ نوری از ROI به طور همزمان با استفاده از آموزش FCNM. کاریلو و همکارانش روشی برای قطعه‌بندی با استفاده از آستانه‌گذاری و اندازه‌گیری جدید برای نسبت کاپ و دیسک ارائه دادند [۲۴]. آنها ابتدا به قطعه‌بندی دیسک پرداختند، کانال رنگی که بهترین نتیجه را برای انجام این کار ارائه داد کانال قرمز بود. برای قطعه‌بندی کاپ، ابتدا رگ‌های خونی را قطعه‌بندی کردند، زیرا انحنای رگ‌ها به تشخیص لبه کاپ در ربع بینی^۴ (N)، فوقانی^۵ (S) و تحتانی^۶ (I) دیسک کمک می‌کند. در این مرحله از کانال سبز تصویر فوندوس و از کانال آبی برای قطعه‌بندی کاپ استفاده شده است. آنها همچنین برای یافتن یک آستانه خوب، الگوریتمی را پیشنهاد کردند که فقط روی ربع زمانی^۷ (T) کار می‌کند.

ساختار این مقاله بدین ترتیب است. در بخش دوم به توضیح بلوک‌ها و معماری‌های استفاده شده در این مقاله پرداخته می‌شود. در بخش سوم، روش پیشنهادی مطرح می‌گردد. سپس در بخش چهارم به توضیح آزمایش‌ها و نتایج مربوط به محلی‌سازی دیسک نوری و قطعه‌بندی کاپ و دیسک نوری و پس از آن به تشخیص گلوکوم با استفاده از نسبت این دو پرداخته شده است. در نهایت جمع‌بندی و نتیجه‌گیری در بخش پنجم بیان می‌شود.

¹ Adaptively Regularized Kernel-Based Intuitionistic Fuzzy C Means

² Region of Interest

³ Fully Convolutional Networks

⁴ Nasal

⁵ Superior

⁶ Inferior

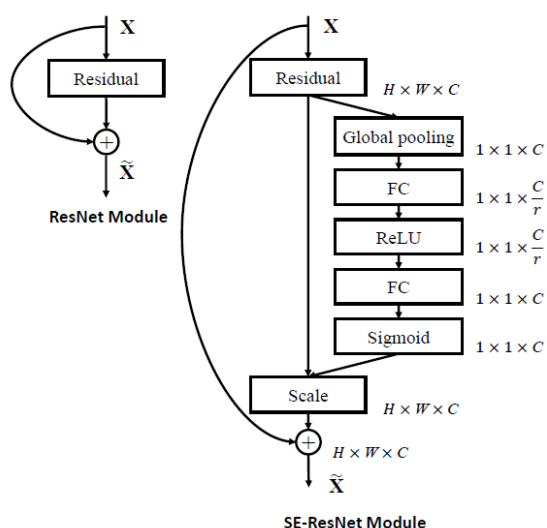
⁷ Temporal

⁸ Residual network

⁹ Degradation

¹⁰ Accuracy

¹¹ Rectified Linear Unit



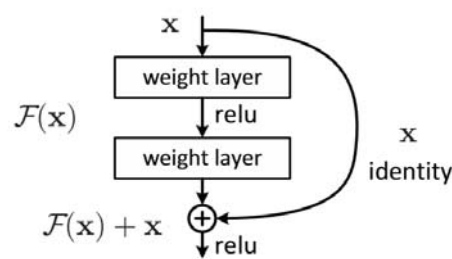
شکل ۳. طرح ماژول باقیمانده اصلی (چپ) و ماژول SE-ResNet (راست) [۲۷]

• وزن دهی: بعد از مرحله فشرده سازی و تحریک، وزن های خروجی نتیجه شده اهمیت هر کانال ویژگی را ارائه می کنند. همانطور که در شکل ۳ نشان داده شده است، بعد از اینکه کانولوشن به شبکه اعمال شد، مجموعه ای از ویژگی ها بطور مستقیم خارج می شوند و وزن های نتیجه شده، با مقیاس دهی کانال به ویژگی هایی که مستقیماً خارج شدند جهت تکمیل وزن دهی مجدد کانال های ویژگی اصلی، وزن دهی شده اند. بلوک های SE می توانند بطور مستقیم توسط شبکه های باقیمانده استفاده شوند. شکل ۳ طرح یک ماژول SE-ResNet را در برابر یک ماژول ResNet نشان می دهد. در اینجا تبدیل بلوک SE به عنوان شاخه نامشخص یک ماژول باقیمانده در نظر گرفته شده است. فشرده سازی و تحریک هر دو قبل از جمع شدن با شاخه معلوم اجرا می شوند. بلوک های SE نیز از نظر محاسباتی بسیار سبک بوده و تنها افزایش اندکی در پیچیدگی مدل و بار محاسباتی تحمیل می کنند [۲۷].

بطور خلاصه چهار مرحله برای معماری SE-ResNet50 و ResNet50 وجود دارد که در جدول ۱ آورده شده است. شکل ها و عملگرها با تنظیمات پارامتری خاص یک ساختمان بلوکی باقی مانده در هر مرحله در داخل براکت لیست شده اند و تعداد بلوک های انباشته شده در بیرون آن ارائه شده است. براکت هایی که با f_c دنبال می شوند نمایش دهنده بعد خروجی دو لایه تمام متصل در ماژول SE است.

۲-۳ RCNN سریعتز

در تمام روش های معرفی شده در برنامه های تشخیص اشیاء در سال های اخیر، یک گام مهم، تشخیص نواحی کاندید است. درواقع بدست آوردن نواحی که می توانند به عنوان شیء مطرح شوند، بخش مهمی در دسته بندی تصویر است [۲۸]. در مقاله رن



شکل ۲. ساختار یک بلوک یادگیری باقیمانده [۲۵]

بطور کلی در آموزش ResNet، یا لایه ها در بلوک های باقیمانده آموزش می بیند یا آموزش برای آن لایه ها با استفاده از اتصالات پرشی رد می شود. بنابراین برای نقاط مختلف آموزش داده، براساس چگونگی جریان پس انتشار خطا در شبکه، قسمت های مختلف شبکه با نرخ های مختلف آموزش داده می شوند.

۲-۲ معماری SE-ResNet

در این معماری جنبه متفاوت طراحی شبکه (رابطه بین کانال ها) نیز مورد بررسی قرار می گیرد. در آن یک واحد معماری جدید با هدف بهبود کیفیت بازنمایی های تولید شده توسط یک شبکه با مدل سازی صریح وابستگی های متقابل بین کانال های ویژگی های کانولوشنی آن معرفی می شود که آن را بلوک فشرده سازی و تحریک^۱ (SE) می نامند. این بلوک ابعاد نگاشت ویژگی را در نظر می گیرد، وابستگی هر کانال را پیدا می کند، اهمیت کانال های ویژگی متفاوت را با استفاده از تکنیک های یادگیری محاسبه می کند و بطور انتخابی ویژگی های اصلی را برای محو کردن ویژگی های غیر مهم ارتقا می دهد که در آن W عرض، H طول، C کانال ها و r ابرپارامتر است و FC تمام کانولوشنی است. شکل ۳ نشان می دهد که دسته ای از ویژگی ها در لایه بالاتر بطور مستقیم خارج می شوند درحالی که دسته دیگر ویژگی ها از عملگر وزن دهی مثل فشرده سازی و تحریک استفاده می کنند. بطور کلی در این معماری سه مرحله زیر اجرا می شود:

- فشرده سازی: یک پولینگ میانگین گیری کلی برای فشرده سازی نگاشت ویژگی دو بعدی استفاده می شود و یک ویژگی یک بعدی با استفاده از فشرده سازی بدست می آید که اطلاعات کلی میدان پذیرشی کانال ویژگی ورودی را بازتاب می کند.
- تحریک: برای استفاده از اطلاعات جمع شده در عملیات فشرده سازی، این عملیات با عملیات دوم دنبال می شود که هدف آن جمع آوری تمام وابستگی های سطح کانال است. همبستگی بین کانال های ویژگی می تواند بطور ضمنی بطور غیر خطی از طریق تحریک مدل شود.

¹ Squeeze-and-Excitation

یک روش کارا برای غربالگری اولیه بیماری گلوکوم بر اساس محاسبه نسبت کاپ به دیسک نوری با استفاده از شبکه‌های عصبی کانولوشنی

سافتمکس برای طبقه‌بندی نهایی و رگرسیون خطی استفاده می‌شود.

۲-۴ معماری U-Net

در قطعه‌بندی معنایی تصویر، بخش‌هایی از تصویر که مربوط به یک شیء از یک کلاس خاص است در یک دسته قرار می‌گیرند. به عبارت دیگر، قطعه‌بندی اجزای تصویر به کلاس‌های از پیش تعیین‌شده طوری که هر کلاس معنای یک بخش را مشخص کند، قطعه‌بندی معنایی تصویر نامیده می‌شود [۳۰].

موفقیت چشمگیر تکنیک‌های یادگیری عمیق در کارهای مختلف حوزه بینایی در سطح بالا به‌ویژه رویکردهای باناظر نظیر CNNها برای طبقه‌بندی تصویر، محققان را برای جستجوی توانایی‌های این شبکه‌ها به‌منظور برچسب‌گذاری سطح پیکسل مثل قطعه‌بندی معنایی تشویق می‌کند. مهم‌ترین مزیت تکنیک‌های یادگیری عمیق که آنها را در لبه بهترین روش‌های مرسوم قرار می‌دهد توانایی یادگیری بازنمایی‌های مناسب برای یک مساله موجود است، به‌عنوان مثال برچسب زدن پیکسل بر روی یک مجموعه داده خاص، به‌شیوه‌ای انتها به انتها به‌جای استفاده از ویژگی‌های دست‌ساز که به تخصص و تلاش نیاز دارد تا بتوان آنها را در یک بخش خاص بکار گرفت.

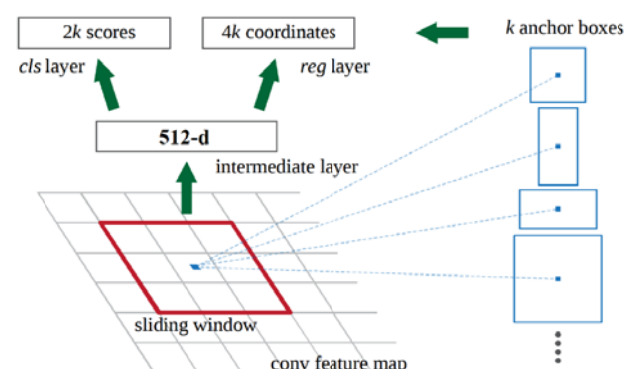
U-Net یکی از معماری‌های شبکه کانولوشنی برای قطعه‌بندی سریع و دقیق تصاویر است. این معماری در [۳۱] معرفی شد و بعنوان یک شبکه عصبی تمام کانولوشنی قادر به آموزش روی تعداد دسته‌داده به شدت کوچک است که به نتایج قابل رقابت با مدل‌های مبتنی بر پنجره لغزان دست یافته است. شبکه U-Net از یک مسیر منقبض شده و یک مسیر گسترشی تشکیل شده که به آن طرحی به شکل U می‌بخشد. در کار رانبرگر و همکارانش مسیر منقبض شده یک شبکه معمولی کانولوشنی است که از استعمال مکرر کانولوشن‌ها تشکیل شده است و هر یک با استفاده از یک ReLU و یک عملگر پولینینگ حداکثر دنبال می‌شود. در طول انقباض، ضمن افزایش اطلاعات ویژگی، اطلاعات مکانی کاهش می‌یابد. مسیر گسترده اطلاعات ویژگی و مکانی را از طریق دنباله‌ای از کانولوشن‌ها ترکیب می‌کند و با ویژگی‌های با وضوح بالا از مسیر انقباضی جمع می‌کند [۳۱].

۳ روش پیشنهادی در تشخیص گلوکوم

برای تشخیص گلوکوم الگوریتم نشان داده شده در شکل ۵ بکار می‌رود. در این الگوریتم از یک شبکه RCNN سریعتر با شبکه پیش‌آموزش‌دیده ResNet50 برای محلی‌سازی دیسک نوری استفاده شده است که این ناحیه مورد علاقه برای قطعه‌بندی کاپ نوری و دیسک نوری وارد شبکه قطعه‌بندی شده و بعد از بدست آوردن دیسک و کاپ قطعه‌بندی شده نسبت کاپ نوری به دیسک نوری که معیاری برای طبقه‌بندی گلوکوم است بدست می‌آید و

جدول ۱. معماری ResNet50 در سمت چپ و SE-ResNet50 در سمت راست [۲۷]

Output size	ResNet-50	SE-ResNet-50
112 × 112	conv, 7 × 7, 64, stride 2 max pool, 3 × 3, stride 2	
56 × 56	$\begin{bmatrix} \text{conv}, 1 \times 1, 64 \\ \text{conv}, 3 \times 3, 64 \\ \text{conv}, 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} \text{conv}, 1 \times 1, 64 \\ \text{conv}, 3 \times 3, 64 \\ \text{conv}, 1 \times 1, 256 \\ \text{fc}, [16, 256] \end{bmatrix} \times 3$
28 × 28	$\begin{bmatrix} \text{conv}, 1 \times 1, 128 \\ \text{conv}, 3 \times 3, 128 \\ \text{conv}, 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} \text{conv}, 1 \times 1, 128 \\ \text{conv}, 3 \times 3, 128 \\ \text{conv}, 1 \times 1, 512 \\ \text{fc}, [32, 512] \end{bmatrix} \times 4$
14 × 14	$\begin{bmatrix} \text{conv}, 1 \times 1, 256 \\ \text{conv}, 3 \times 3, 256 \\ \text{conv}, 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} \text{conv}, 1 \times 1, 256 \\ \text{conv}, 3 \times 3, 256 \\ \text{conv}, 1 \times 1, 1024 \\ \text{fc}, [64, 1024] \end{bmatrix} \times 6$
7 × 7	$\begin{bmatrix} \text{conv}, 1 \times 1, 512 \\ \text{conv}, 3 \times 3, 512 \\ \text{conv}, 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} \text{conv}, 1 \times 1, 512 \\ \text{conv}, 3 \times 3, 512 \\ \text{conv}, 1 \times 1, 2048 \\ \text{fc}, [128, 2048] \end{bmatrix} \times 3$
1 × 1	global average pool, 1000-d fc, softmax	

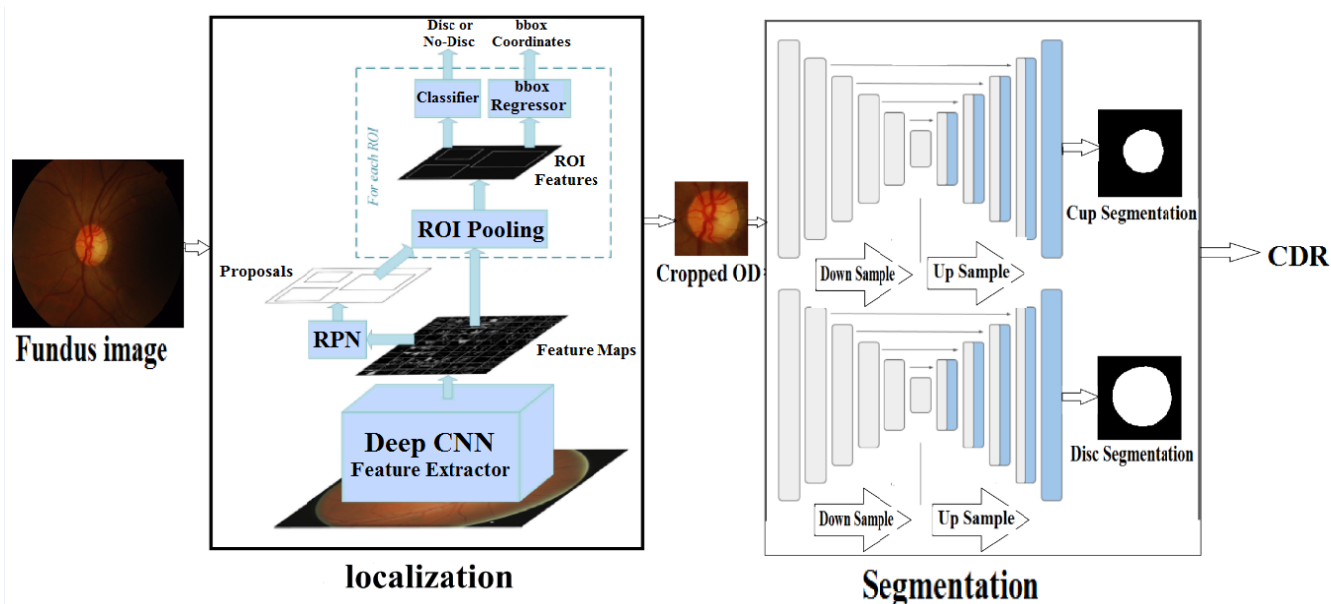


شکل ۴. RPN در R-CNN سریعتر. K جعبه لنگر از پیش تعریف شده با هر پنجره لغزان برای تولید بردارهایی با طول ثابت که توسط لایه رگرسیون و طبقه‌بندی برای بدست آوردن خروجی‌های متناظر گرفته می‌شوند، کانوال می‌شوند [۲۹].

و همکارانش یک شبکه پیشنهاد ناحیه کاندید^۱ (RPN) معرفی شد [۲۹].

RPN نداشت ویژگی بدست آمده از شبکه کانولوشنی پیش‌آموزش دیده را برای تولید مجموعه‌ای از پیشنهاد‌های شیء مستطیل-شکل می‌گیرد. معماری RPN در شکل ۴ نشان داده شده است. شبکه روی نگاشت ویژگی کانولوشنی می‌لغزد و بطور کامل به یک پنجره مکانی n × n متصل می‌شود. یک بردار کم بعد (۵۱۲ بعدی برای ResNet50) بدست می‌آید و به دو لایه کانولوشنی ۱ × ۱، یعنی لایه طبقه‌بندی (CLS) برای اینکه مشخص کند جعبه هدف است یا پس زمینه و لایه رگرسیون (reg) جعبه داده می‌شود. بعد از مرحله RPN، پولینینگ ROI برای نواحی پیشنهاد شده به‌منظور یکسان سازی اندازه آنها استفاده می‌شود و بعد از هموارسازی توسط لایه تمام متصل در مرحله نهایی از تابع

¹ Region Proposal Network



شکل ۵. چارچوب الگوریتم تشخیص گلوکوم

از RCNN سریعتر که برای تشخیص شی بکار رفته است با شبکه ResNet50 بعنوان استخراج کننده ویژگی استفاده می شود.

۲-۳ قطعه بندی

آزمایش عصب بینایی سر که اندازه نسبت کاپ به دیسک (CDR) را درگیر می کند، یکی از ارزشمندترین روش های تشخیص بیماری گلوکوم است. تخمین نسبت کاپ به دیسک نیازمند قطعه بندی دیسک و کاپ نوری در تصاویر فوندوس چشم است و می تواند با استفاده از الگوریتم های نوین بینایی ماشین انجام شود.

سه روش $VCDR^1$ ، $HCDR^2$ و $ACDR^3$ برای محاسبه CDR مطرح شده است. دو روش اول به ترتیب مبتنی بر نسبت عمودی قطر کاپ به قطر دیسک نوری و نسبت افقی قطر کاپ به قطر دیسک نوری و روش سوم مبتنی بر نسبت مساحت کاپ به دیسک نوری است. از آنجا که ممکن است کاپ در زوایای مختلف جهت دار شود، روش سوم بعنوان بهترین روش در نظر گرفته می شود.

۱-۲-۳ رویکرد مبتنی بر U-Net با شبکه پایه

SE - ResNet50

معماری پیشنهادی U-Net اصلاح شده در این پژوهش در شکل ۶ قابل مشاهده است. در این معماری بدلیل اینکه آموزش شبکه های عصبی از پایه، کاری زمانبر است در قسمت کدگذار از یک شبکه عصبی از پیش آموزش دیده استفاده شده است که این

براساس آن تصمیمات لازم برای تشخیص افراد دارای گلوکوم و نرمال اتخاذ می شود. نوآوری این مقاله در بخش قطعه بندی است که با استفاده از شبکه اصلاح شده U-Net به قطعه بندی کاپ و دیسک نوری پرداخته است.

۱-۳ محلی سازی دیسک نوری

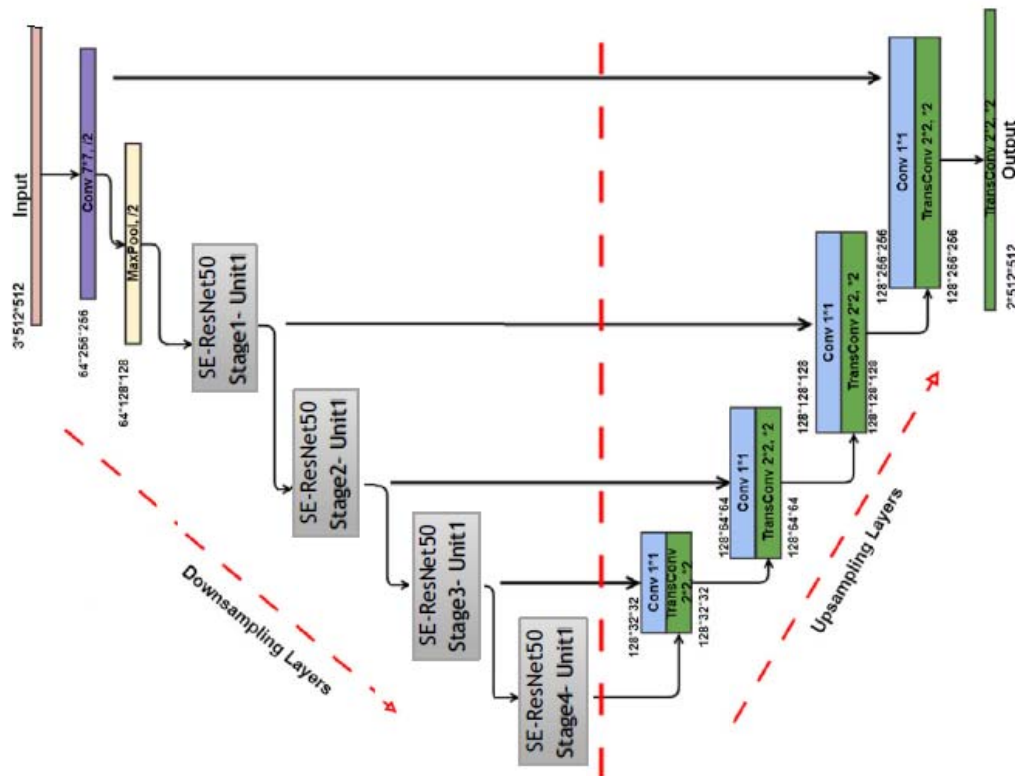
بدلیل اینکه بیماری گلوکوم دراصل دیسک نوری و اطراف آن را تحت تاثیر قرار می دهد بیشتر اطلاعات مربوط به گلوکوم در این ناحیه قرار دارد [۳۲] و [۳۳]. به علاوه اورلاندو و همکارانش نشان دادند هنگامی که از CNN برای ارزیابی گلوکوم استفاده می شود، برش تصویر در اطراف دیسک نوری و استخراج ناحیه مطلوب کارآمدترین راه نسبت به استفاده از تصویر کامل است [۳۴]. استخراج ناحیه مطلوب نه تنها تصویر آغازین کوچکتری را که از نظر محاسباتی کارآمد است تولید می کند، بلکه به CNN اجازه می دهد تا روی بخش مهم تصویر تمرکز کند.

بنابراین ناحیه مورد علاقه را باید استخراج کرده و به شبکه ارائه کرد. پس در مرحله اول محلی سازی دیسک نوری انجام می شود. روش های پردازش تصویر مرسوم برای یکی یا تعداد کمی از دسته داده ها جوابگو هستند و بخوبی قابل تعمیم به سایر دسته داده ها نیستند. بنابراین نیاز مبرم به یک روش مستقل از دسته داده و مقاوم و تمام خودکار برای محلی سازی دیسک وجود دارد که بتواند به تعداد زیادی از دسته داده های تصاویر فوندوس شبکه با اطمینان و دقت بالا اعمال شود؛ که این امر با رویکرد شبکه های عصبی بانظر تحقق می یابد. بنابراین در این مقاله یک رویکرد تمام خودکار محلی سازی دقیق دیسک نوری بدون استفاده از هیچ دانش تجربی درباره دسته داده اعمال می شود. برای این کار

¹ Vertical Cup to Disc Ratio

² Horizontal Cup to Disc Ratio

³ Area Cup to Disc Ratio



شکل ۶. معماری U-Net اصلاح شده پیشنهادی. (این شکل برای تصویری با ابعاد 512×512 نشان داده شده است.)

پایتون و کتابخانه Keras با پشتیبانی Tensorflow با قابلیت اجرای برنامه بر روی واحد پردازنده گرافیکی استفاده شده است.

۴-۱ پایگاه‌های داده مورد استفاده

نام هر پایگاه داده و اینکه آیا پایگاه‌های داده شامل حاشیه‌نویسی کاپ نوری و دیسک نوری است، در جدول ۲ ارائه شده است. در ادامه به تشریح مشخصات این دسته‌داده‌ها پرداخته می‌شود.

۴-۱-۱-۴ پایگاه داده Drishti-GS1

دسته‌داده Drishti-GS1 شامل ۱۰۱ تصویر شبکه‌ای است. تمام جدول ۲. پایگاه‌های داده مورد استفاده در ارزیابی گلوکوم

پایگاه داده	تعداد تصویر	حاشیه‌نویسی قطعه‌بندی کاپ	حاشیه‌نویسی قطعه‌بندی دیسک نوری
Drishti-GS1	۱۰۱ تصویر با برچسب گلوکوم و نرمال	✓	✓
RIM-ONE v3	۱۵۹ تصویر با برچسب گلوکوم و نرمال	✓	✓
MESSIDOR	۴۶۰ تصویر بدون برچسب	✓	✓
HRF	۳۰ تصویر با برچسب گلوکوم و نرمال	—	—
ORIGA-light	۶۵۰ تصویر با برچسب گلوکوم و نرمال	—	—
OCT & CFI	۱۰۰ تصویر با برچسب نرمال	—	—

امر موجب افزایش سرعت یادگیری ماشین شده و شبکه‌ی مقاومی بوجود می‌آید و نیز از بیش‌برازش جلوگیری می‌شود.

شبکه عصبی پیش آموزش داده شده برای انتقال یادگیری در این معماری SE-ResNet50 است. معماری U-Net ماژول SE-ResNet50 را بعنوان بدنه اصلی می‌گیرد تا وابستگی‌های بین کانال را با تکنیک یادگیری تعدیل کند. در اصل، تنظیم مجدد ویژگی‌های کانال، ویژگی‌های سطح پایین شبکه را با ویژگی‌های سطح بالا ترکیب می‌کند که منجر به نگاشت‌های ویژگی با رزولوشن بالا با اطلاعات معنایی قوی می‌شود. درمقایسه با ResNet معمولی که تنها اهمیت کانال ویژگی را در نظر می‌گیرد، مرحله تنظیم مجدد از ویژگی‌های کانال تضمین می‌کند که شبکه اصلی، میدان پذیرشی بزرگتری را برای بعضی از مقادیر بدست آورد. ترکیب ویژگی‌های سطح پایین و ویژگی‌های سطح بالا مزیت‌های اطلاعات هندسی و معنایی را برای بهبود مقاوم بودن شبکه ترکیب می‌کند.

۴ آزمایش‌ها و نتایج

در این بخش به بررسی و تحلیل نتایج حاصل از شبیه‌سازی در محلی‌سازی دیسک نوری، قطعه‌بندی کاپ و دیسک نوری و تشخیص حاصل از نسبت کاپ به دیسک پرداخته شده و نتایج با نتایج سایر مدل‌های اخیر در این حوزه مقایسه می‌شوند. تمام پیاده‌سازی‌ها بر روی پردازنده گرافیکی Tesla K80 با استفاده از Google Colab انجام گرفت. بدین منظور از زبان برنامه‌نویسی

۴-۱-۶- پایگاه داده OCT & CFI

این دسته‌داده شامل توموگرافی انسجام نوری^۵ (OCT) و تصاویر رنگی فوندوس از هر دو چشم ۵۰ فرد سالم جمع‌آوری شده در بخش چشم پزشکی بیمارستان فیض در شهر اصفهان است. از آنجا که تصاویر به‌عنوان بخشی از بررسی مقایسه OCT‌های ماکولای چشم راست و چپ افراد عادی گرفته شده است، هیچ حقیقت مبنایی را در رابطه با قطعه‌بندی دیسک یا کاپ نوری یا محلی‌سازی دیسک نوری ارائه نمی‌دهد [۴۰].

۴-۲- نحوه ارزیابی مدل پیشنهادی

مفاهیم اصلی مورد استفاده در ارزیابی الگوریتم‌های محلی‌سازی، قطعه‌بندی و طبقه‌بندی نهایی به شرح جدول ۳ است. معیارهایی که به‌منظور ارزیابی معماری مورد نظر و مقایسه آن با سایر روش‌ها مورد استفاده قرار می‌گیرد، در ادامه آمده است.

۴-۲-۱- منحنی مشخصه عملکرد سیستم

منحنی مشخصه عملکرد سیستم^۶ (ROC) نموداری است که عملکرد مدل‌های طبقه‌بند دو کلاسه را در آستانه‌های مختلف مورد بررسی قرار می‌دهد. در نمودار ROC مقدار نرخ مثبت صحیح (TPR) در محور عمودی و مقدار نرخ مثبت کاذب (FPR) در محور افقی بر حسب آستانه‌های مختلف رسم می‌گردد. مقادیر TPR و FPR به صورت زیر محاسبه می‌گردند:

$$TPR = \frac{TP}{TP + FN} \quad (۳)$$

$$FPR = \frac{FP}{FP + TN} \quad (۴)$$

جدول ۳. مفاهیم پارامترهای ارزیابی رایج

پارامتر	نماد	توضیح
مثبت صحیح ^۷	TP	داده بعنوان کلاس مثبت برچسب گذاری شده و حقیقتاً هم عضو کلاس مثبت است.
مثبت کاذب ^۸	FP	داده بعنوان کلاس مثبت برچسب گذاری شده و حقیقتاً عضو کلاس منفی است.
منفی صحیح ^۹	TN	داده بعنوان کلاس منفی برچسب گذاری شده و حقیقتاً هم عضو کلاس منفی است.
منفی کاذب ^{۱۰}	FN	داده بعنوان کلاس منفی برچسب گذاری شده و حقیقتاً عضو کلاس مثبت است.

تصاویر در بیمارستان چشم آراویند^۱ از بیماران بین ۴۰ تا ۸۰ سال با تعداد تقریباً مساوی مردان و زنان جمع‌آوری و توسط چهار متخصص بالینی برچسب‌گذاری شده‌اند. تمام تصاویر با محوریت دیسک نوری با میدان دید^۲ ۳۰ درجه و با ابعاد ۲۸۹۶ × ۱۹۴۴ پیکسل گرفته شده‌اند. کارشناسان بالینی با تجربه ۳، ۵، ۹ و ۲۰ سال با کمک یک ابزار ویژه نشانه‌گذاری که اجازه تعیین دقیق مرز را می‌دهد، حقیقت مبنای را ایجاد کردند [۳۵].

۴-۱-۲- پایگاه داده RIM-ONE v3

RIM-ONE مجموعه‌ای از پایگاه‌های داده شبکه برای ارزیابی عصب بینایی و تشخیص گلوکوم است. نسخه‌های مختلفی از این پایگاه داده وجود دارد که RIM-ONE v3 از ۱۵۹ تصویر تشکیل شده است. این پایگاه داده توسط چشم پزشکان از بخش چشم پزشکی بیمارستان دانشگاهی جزایر قناری^۳ در اسپانیا ایجاد شده است. علاوه بر تصاویر اصلی در RIM-ONE v3، این بانک اطلاعاتی حاوی دو حاشیه‌نویسی مجزا برای قطعه‌بندی دیسک نوری و کاپ نوری است که توسط دو متخصص انجام شده است [۳۶]. شایان ذکر است که در این دسته‌داده دیسک نوری برش‌خورده از تصویر فوندوس در پایگاه داده موجود است.

۴-۱-۳- دسته‌داده MESSIDOR

MESSIDOR بخشی از پایگاه داده RIGA است که از ۴۶۰ تصویر فوندوس با رزولوشن ۱۴۸۸ × ۲۲۴۰ پیکسل تشکیل شده است. هر تصویر فوندوس شامل شش حاشیه‌نویسی قطعه‌بندی کاپ و دیسک نوری است که توسط شش متخصص انجام شده است [۳۷].

۴-۱-۴- پایگاه داده HRF

HRF مخفف فوندوس با وضوح بالا^۴ است. این پایگاه داده حاوی ۴۵ تصویر است که ۱۵ تصویر مربوط به افراد سالم، ۱۵ تصویر بیماران مبتلا به رتینوپاتی دیابتی و ۱۵ تصویر مربوط به بیماران گلوکومی است. این تصاویر توسط یک دوربین فوندوس Canon CR-1 با میدان دید ۴۵ درجه با وضوح ۲۳۳۶ × ۳۵۰۴ پیکسل جمع‌آوری شدند [۳۸].

۴-۱-۵- پایگاه داده ORIGA-light

ORIGA-light شامل ۶۵۰ تصویر شبکه است که توسط متخصصان آموزش دیده از موسسه تحقیقاتی چشم سنگاپور حاشیه‌نویسی شده است. این دسته‌داده از ۱۶۸ تصویر گلوکومی و ۴۸۲ تصویر نرمال با رزولوشن ۲۰۴۸ × ۳۰۷۲ پیکسل تشکیل شده است [۳۹].

⁵ Optical Coherence Tomography

⁶ Receiver Operating Characteristic

⁷ True Positives

⁸ False Positives

⁹ True Negatives

¹⁰ False Negative

¹ Aravind

² Field of View (FOV)

³ Universitario de Canarias

⁴ High-Resolution Fundus

ضریب همبستگی برابر صفر است نشان می‌دهد که بین دو متغیر رابطه خطی وجود ندارد.

با تعریف $Z_x = \frac{x - \bar{x}}{s_x}$ و $Z_y = \frac{y - \bar{y}}{s_y}$ که در آن s_x و s_y

به ترتیب انحراف معیارهای x و y و نیز $\bar{x}$ و $\bar{y}$ به ترتیب میانگین‌های x و y می‌باشند ضریب همبستگی پیرسون بصورت معادله ۹ تعریف می‌شود.

$$r = \frac{\sum Z_x Z_y}{n} \quad (9)$$

که در آن n تعداد نمونه‌ها است.

۳-۴ آزمایشات تشخیص گلوکوم با معماری پیشنهادی

در این بخش به نتایج حاصل از محلی‌سازی دیسک نوری، قطعه‌بندی کاپ و دیسک نوری و در نهایت تشخیص بیماری گلوکوم طبق CDR پرداخته می‌شود.

۳-۴-۱- نتایج محلی‌سازی دیسک نوری

به‌منظور محلی‌سازی دیسک نوری از چهار دسته‌داده HRF، ORIGA-light، DRISHTI GS1 و CFI & OCT در فاز آموزش استفاده شده است. ترکیب این چهار دسته‌داده برای تغییرات سایر دسته‌داده‌ها کفایت می‌کند و نتیجه مقاومی بدست می‌آید. برای کاهش حجم محاسبات و یکسان‌سازی رزولوشن در تمام دسته‌داده‌ها، رزولوشن همه تصاویر برای فاز آموزش برابر با 128×128 پیکسل قرار داده شده است. تنظیمات ابرپارامترهای مورد استفاده در محلی‌سازی دیسک نوری در جدول ۴ ارائه شده است. نتایج محلی‌سازی دیسک نوری با توجه به اینکه در این مقاله محلی‌سازی درست به معنای IoU بیشتر از ۵۰٪ در نظر گرفته شده است، در تصاویری از پایگاه‌های داده Drishti-GS1 و Drions_DB و نیز دسته داده MESSIDOR در جدول ۵ آمده است. همچنین نمونه‌ای از محلی‌سازی دیسک نوری با استفاده از RCNN سریعتر در این تصاویر در شکل ۷ آمده است.

نتایج محلی‌سازی بطور میانگین در زمان ۰٫۴۱ ثانیه برای هر تصویر بدست آمد. به‌منظور ارزیابی کارایی روش محلی‌سازی در دسته داده‌ای متمایز از دسته داده‌های فاز آموزش، همانگونه که از

جدول ۴. ابرپارامترهای مورد استفاده در محلی‌سازی دیسک نوری

نام پارامتر	مقدار
اندازه تصویر ورودی	128×128
اندازه دسته آموزشی	۳۲
تعداد تکرار	۵۰۰
بهینه‌ساز	Adam
نرخ یادگیری	۰٫۰۰۰۱
شبکه پیش آموزش دیده	ResNet50

با استفاده از منحنی به‌دست آمده می‌توان با محاسبه مساحت زیر نمودار ROC مقدار AUC را بدست آورد.

۴-۲-۴-۲- معیار F1-score

فراخوانی^۱ یا میزان حساسیت، نسبتی از نمونه‌های مرتبطی است که که بازیابی شده‌اند. دقت^۲ یا ارزش پیش‌بینی‌شده مثبت به نسبت نمونه‌های بازیابی شده‌ای گفته می‌شود که بدرستی پیش‌بینی شده‌اند. F1-score معیاری است که با محاسبه میانگین هندسی دقت و فراخوانی بدست می‌آید. با توجه به جدول ۳ معیارهای ارزیابی به‌صورت زیر محاسبه می‌شوند:

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$F_1 - score = \frac{Pr \cdot Re}{Pr + Re} \times 2 \quad (7)$$

۴-۲-۴-۳- معیار اشتراک بر اجتماع (IoU)

در برنامه قطعه‌بندی بدلیل اینکه سطح پیش‌بینی‌شده برای هدف دقیقاً سطح حقیقت مبنا را نمی‌پوشاند و یا جعبه مرزی پیش‌بینی‌شده در برنامه محلی‌سازی دقیقاً جعبه مرزی حقیقت مبنا را پوشش نمی‌دهد از معیار اشتراک بر اجتماع^۳ استفاده می‌شود. این معیار برای گستره همپوشانی بین دو سطح استفاده می‌شود که مقداری بین ۰ و ۱ را به خود اختصاص می‌دهد و بصورت زیر محاسبه می‌شود.

$$IoU = \frac{GT \cap Pr}{GT \cup Pr} \quad (8)$$

که در آن GT و Pr به ترتیب سطح حقیقت مبنا و سطح پیش‌بینی شده در برنامه قطعه‌بندی یا سطح جعبه مرزی حقیقت مبنا و سطح جعبه مرزی پیش‌بینی‌شده در برنامه محلی‌سازی هستند. علائم $\cap$ و $\cup$ به ترتیب معادل اشتراک و اجتماع است.

۴-۲-۴-۴- ضریب همبستگی پیرسون

این ضریب به‌منظور تعیین میزان رابطه، نوع و جهت رابطه بین دو متغیر فاصله‌ای یا نسبی و یا یک متغیر فاصله‌ای و یک متغیر نسبی به‌کار برده می‌شود. ضریب همبستگی پیرسون بین ۱- و ۱ تغییر می‌کند. $r = 1$ بیانگر رابطه مستقیم کامل بین دو متغیر است، رابطه مستقیم یا مثبت به این معناست که اگر یکی از متغیرها افزایش (کاهش) یابد، دیگری نیز افزایش (کاهش) می‌یابد. $r = -1$ وجود یک رابطه معکوس کامل بین دو متغیر را نشان می‌دهد. رابطه معکوس یا منفی نشان می‌دهد که اگر یک متغیر افزایش یابد متغیر دیگر کاهش می‌یابد و بالعکس. زمانی که

¹ Recall

² Precision

³ Intersection over Union

⁴ Batch size

با ۱۹۱ بعنوان هدف در نظر گرفته شده و به سطح ۲۵۵ تبدیل می‌شوند و بقیه سطوح به سطح ۰ تبدیل می‌شوند. در دسته‌داده RIM-ONE v3، میانگین حاشیه‌نویسی‌های دو متخصص بعنوان حقیقت مبنا انتخاب می‌شود. در رویکرد پیشنهادی دوم در دسته‌داده Drishti-GS1 با هدف میانگین‌گیری از نظرات ۴ متخصص، سطوح خاکستری بالاتر و مساوی با ۱۲۸ بعنوان هدف در نظر گرفته شده و به سطح ۲۵۵ تبدیل می‌شوند و بقیه سطوح به سطح ۰ تبدیل می‌شوند. در دسته‌داده RIM-ONE v3 به روال رویکرد پیشنهادی اول عمل می‌شود.

همچنین کارایی شبکه تحت تأثیر تعداد داده‌های فاز آموزش است. به همین دلیل و برای جلوگیری از بیش‌برازش از تقویت داده برای آموزش شبکه استفاده می‌شود. این تقویت داده توسط تبدیلات پردازش تصویر نظیر دوران نسبت به محور افقی و محور عمودی، جابجایی مرکز تصویر، مقیاس‌دهی تصاویر، افزودن نویز گاوسی به تصاویر، تارکردن، همسان‌سازی هیستوگرام تطبیقی با کنتراست محدود^۱ به‌منظور بهبود کنتراست محلی دسته‌داده‌ها و تقویت لبه انجام می‌شود. همچنین بدلیل تطبیق با برخی دسته‌داده‌ها که دارای رنگ، سطح اشباع و مقادیر پیکسل متفاوت و کنتراست و روشنایی متفاوتی هستند، در روش اعمال‌شده بطور تصادفی تصاویر با رنگ‌ها و سطوح اشباع و کنتراست و روشنایی متفاوت تولید شد و نیز بصورت تصادفی تیزی^۲ بر روی تصاویر اعمال شد که کارایی شبکه را بطور چشمگیری افزایش داد.

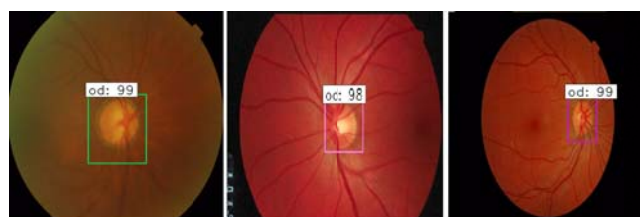
نتایج قطعه‌بندی معنایی برای دسته‌داده Drishti-GS1 در جدول ۷ و برای دسته‌داده RIM-ONE v3 در جدول ۸ آورده شده است. در شکل ۸ نمونه‌ای از قطعه‌بندی معنایی توسط الگوریتم پیشنهادی در دو رویکرد پیشنهادی و نیز حقیقت مبنا به همراه تصویر اصلی در دسته‌داده Drishti-GS1 قابل مشاهده است.

همچنین در شکل ۹ نمونه‌ای از قطعه‌بندی معنایی توسط الگوریتم پیشنهادی و در دو رویکرد پیشنهادی و نیز قطعه‌بندی حقیقت مبنا به همراه تصویر اصلی در دسته‌داده RIM-ONE v3 قابل مشاهده است. همانطور که از جدول‌های ۷ و ۸ برمی‌آید معیارهای اندازه‌گیری تحت تأثیر نحوه انتخاب حقیقت مبنا برای دسته‌داده Drishti-GS1 نیز است. این تأثیرگذاری شامل دسته‌داده RIM-ONE v3 نیز که در هر دو رویکرد پیشنهادی حقیقت مبنای یکسانی داشت می‌شود. بعنوان مثال برای قطعه‌بندی کاپ نوری معیارهای F_1 -score و IoU رویکرد پیشنهادی دوم در هر دو دسته‌داده بهبود پیدا کرده‌اند.

بمنظور بررسی کارایی شبکه در دسته‌داده متمایز نتایج قطعه‌بندی که روی ۹۰ تصویر از دسته‌داده MESSIDOR انجام شده است در جدول ۹ قابل مشاهده است که در شکل ۱۰ نمونه‌ای از قطعه‌بندی معنایی توسط الگوریتم پیشنهادی و در دو رویکرد

جدول ۵. نتایج محلی‌سازی دیسک نوری برای مقدار IoU بیشتر از ۵۰٪ (%)

دسته‌داده	MESSIDOR	Drions_DB	Drishi_GS1
معیار			
Precision	۹۹,۱۳	۱۰۰	۱۰۰
Recall	۹۹,۱۳	۹۴,۵۴	۱۰۰
F ₁ -score	۹۹,۱۳	۹۷,۱۹	۱۰۰



Drishti-GS1 Drions-DB MESSIDOR

شکل ۷. نمونه‌ای از محلی‌سازی دیسک نوری با استفاده از RCNN سریعتر

جدول ۶. ابرپارامترهای مورد استفاده در قطعه‌بندی

نام پارامتر	مقدار
اندازه تصویر ورودی	۱۲۸×۱۲۸
اندازه دسته آموزشی	۲
تعداد تکرار	۲۰۰
بهینه‌ساز	RMSprop
نرخ یادگیری	۰,۰۰۰۱
شبکه پیش‌آموزش دیده	SE- ResNet50

جدول ۵ مشاهده می‌شود، در دسته‌داده‌های Drions_DB و MESSIDOR که در فاز آموزش وجود نداشت، دقت قابل توجهی بدست آمد که مقاوم بودن شبکه را در برابر دسته‌داده‌های مختلف که دارای رزولوشن و سطح رنگی مختلف هستند، نشان می‌دهد.

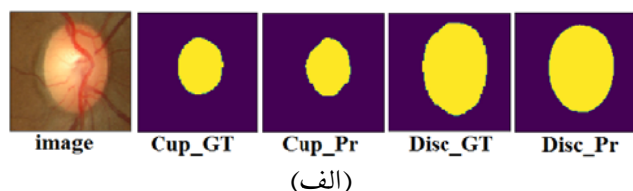
۴-۳-۲- نتایج قطعه‌بندی

نتایج قطعه‌بندی با تنظیم ابرپارامترهای جدول ۶ در ادامه آورده شده است. به‌منظور قطعه‌بندی کاپ و دیسک نوری، دو دسته‌داده RIM-ONE v3 و Drishti-GS1 برای هر دو فاز آموزش و تست مورد استفاده قرار گرفتند. از دسته‌داده RIM-ONE v3 به تعداد ۱۱۱ تصویر برای فاز آموزش، ۲۴ تصویر برای اعتبارسنجی و ۲۴ تصویر برای تست الگوریتم و از دسته‌داده Drishti-GS1 به تعداد ۷۱ تصویر برای آموزش، ۱۵ تصویر برای اعتبارسنجی و ۱۵ تصویر برای تست انتخاب شدند. در این مقاله، دو رویکرد به‌منظور انتخاب حقیقت مبنا در دسته‌داده Drishti-GS1 برای قطعه‌بندی معنایی باینری دیسک و کاپ نوری پیشنهاد می‌شود. در رویکرد پیشنهادی اول در دسته‌داده Drishti-GS1 با هدف پذیرفتن نظر حداقل ۳ متخصص از ۴ متخصص برای حاشیه‌گذاری کاپ و دیسک نوری، سطوح خاکستری بالاتر و برابر

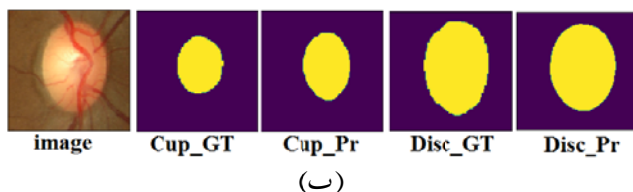
¹ Contrast-Limited Adaptive Histogram Equalization (CLAHE)

² Sharpening

یک روش کارا برای غربالگری اولیه بیماری گلوکوم بر اساس محاسبه نسبت کاپ به دیسک نوری با استفاده از شبکه‌های عصبی کانولوشنی



(الف)



(ب)

شکل ۱۰. نمونه‌ای از قطعه‌بندی معنایی روش پیشنهادی در دسته‌داده MESSIDOR در رویکرد پیشنهادی: (الف): اول و (ب): دوم

جدول ۱۰. میانگین نتایج قطعه‌بندی در رویکرد پیشنهادی اول

معیار	کاپ نوری		دیسک نوری		زمان آموزش (s)	شبکه پایه
	اول	دوم	اول	دوم		
IoU	۰,۸۶۵	۰,۸۷۷	۰,۹۵۶	۰,۹۵۳	۱۶۲۰	VGG16
F ₁ -score	۰,۹۲۶	۰,۹۳۴	۰,۹۷۷	۰,۹۷۶	۱۹۸۰	ResNet50
IoU	۰,۶۴۳	۰,۶۶۵	۰,۹۱۴	۰,۹۱۵	۴۲۵۵	ResNet101
F ₁ -score	۰,۷۵۸	۰,۷۸۷	۰,۹۵۵	۰,۹۵۵	۲۱۰۰	SE-ResNet50

جدول ۱۱. میانگین نتایج قطعه‌بندی در رویکرد پیشنهادی دوم

معیار	کاپ نوری		دیسک نوری		زمان آموزش (s)	شبکه پایه
	اول	دوم	اول	دوم		
IoU	۰,۸۱۱	۰,۷۱۰	۰,۹۳۲	۰,۸۷۸	۱۶۴۳	VGG16
F ₁ -score	۰,۸۲۳	۰,۷۲۸	۰,۹۶۳	۰,۹۳۰	۱۹۹۲	ResNet50
IoU	۰,۷۲۲	۰,۷۳۰	۰,۸۸۲	۰,۸۸۵	۴۲۵۵	ResNet101
F ₁ -score	۰,۸۳۰	۰,۸۳۶	۰,۹۳۲	۰,۹۳۳	۲۱۰۰	SE-ResNet50

پیشنهادی و نیز قطعه‌بندی حقیقت مبنا به همراه تصویر اصلی در دسته‌داده MESSIDOR قابل مشاهده است.

برای بررسی تأثیر شبکه پیش‌آموزش دیده در بخش کدگذار معماری U-Net، شبکه‌های VGG16، ResNet50 و ResNet101 در بخش کدگذار با همان تنظیمات جدول ۶ بکار رفته و نتایج قطعه‌بندی کاپ و دیسک نوری در رویکرد پیشنهادی اول و دوم به ترتیب در جدول ۱۰ و جدول ۱۱ آورده شده است. همانگونه که مشاهده می‌شود در هر دو رویکرد پیشنهادی، معماری U-Net که از SE-ResNet50 در بخش کدگذار خود استفاده می‌کند نتایج بهتری بدست آورده است. این درحالی است که زمان آموزش و حجم محاسباتی آن تفاوت اندکی با زمان آموزش ResNet50 دارد و خیلی سریعتر از ResNet101 است.

۴-۳-۳- نتایج تشخیص گلوکوم بر اساس CDR

برای ارزیابی نسبت کاپ به دیسک، به این دلیل که شکل هندسی کاپ و دیسک بطور کامل به دایره و بیضی شباهت ندارد، در این

جدول ۷. میانگین نتایج قطعه‌بندی برای دسته‌داده Drishti-GS1

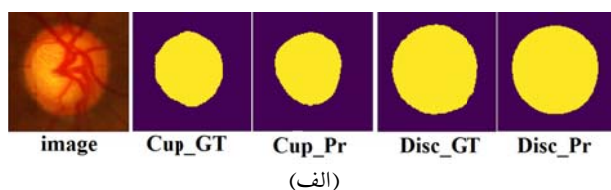
معیار	کاپ نوری		دیسک نوری	
	اول	دوم	اول	دوم
IoU	۰,۸۶۵	۰,۸۷۷	۰,۹۵۶	۰,۹۵۳
F ₁ -score	۰,۹۲۶	۰,۹۳۴	۰,۹۷۷	۰,۹۷۶

جدول ۸. میانگین نتایج قطعه‌بندی برای دسته‌داده RIM-ONE v3

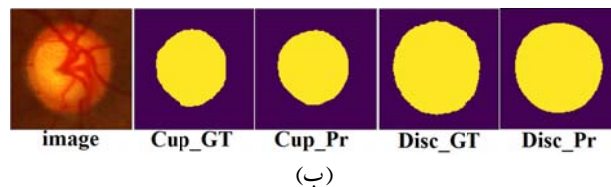
معیار	کاپ نوری		دیسک نوری	
	اول	دوم	اول	دوم
IoU	۰,۶۴۳	۰,۶۶۵	۰,۹۱۴	۰,۹۱۵
F ₁ -score	۰,۷۵۸	۰,۷۸۷	۰,۹۵۵	۰,۹۵۵

جدول ۹. میانگین نتایج قطعه‌بندی برای دسته‌داده MESSIDOR

معیار	کاپ نوری		دیسک نوری	
	اول	دوم	اول	دوم
IoU	۰,۷۲۲	۰,۷۳۰	۰,۸۸۲	۰,۸۸۵
F ₁ -score	۰,۸۳۰	۰,۸۳۶	۰,۹۳۲	۰,۹۳۳

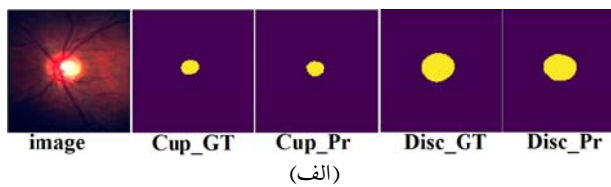


(الف)

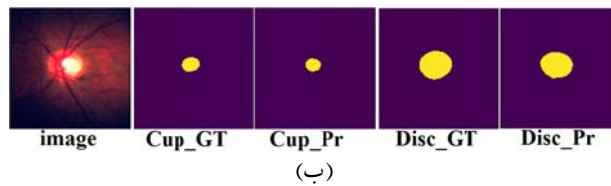


(ب)

شکل ۸. نمونه‌ای از قطعه‌بندی معنایی روش پیشنهادی برای دسته‌داده Drishti-GS1 در رویکرد پیشنهادی: (الف): اول و (ب): دوم



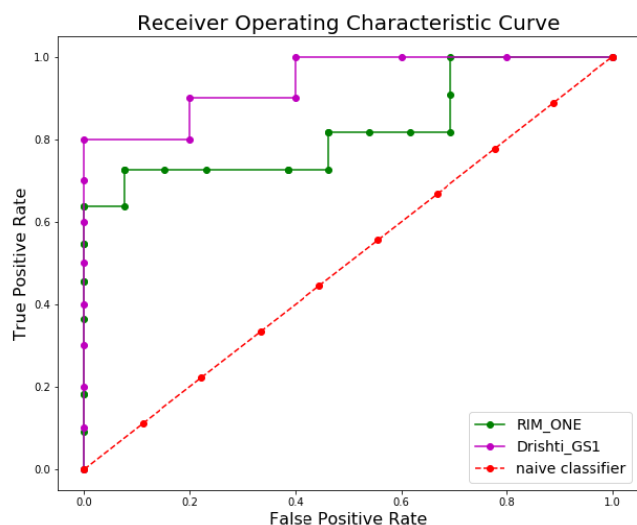
(الف)



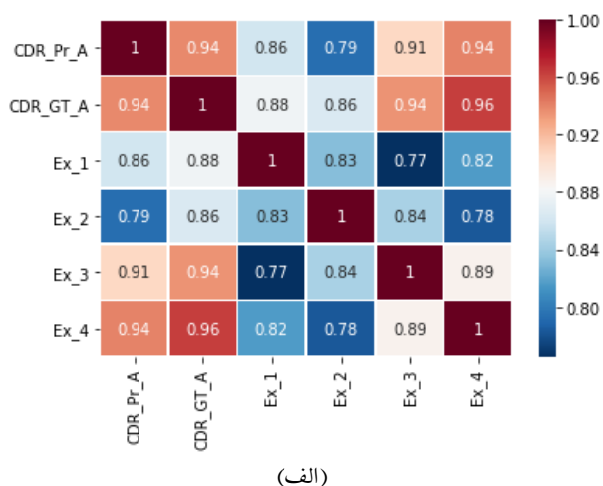
(ب)

شکل ۹. نمونه‌ای از قطعه‌بندی معنایی روش پیشنهادی در دسته‌داده RIM-ONE v3 در رویکرد پیشنهادی: (الف): اول و (ب): دوم

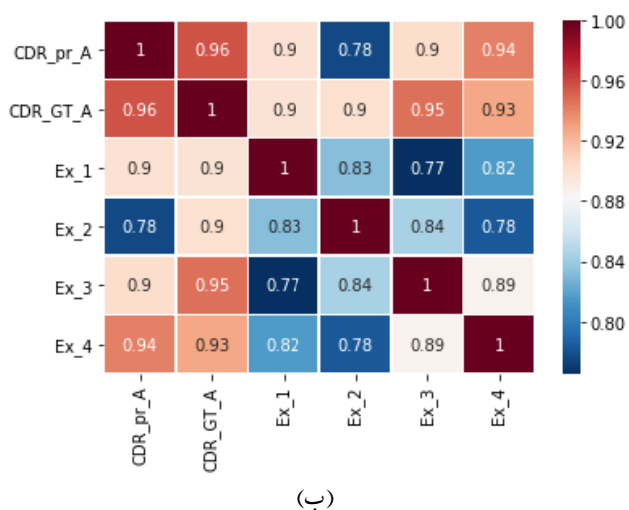
دسته‌داده Drishti_GS1 باعث آموزش بهتر شبکه شده و به نحوی معیار بهتری برای در نظر گرفتن بعنوان حقیقت مبنا برای این دسته‌داده است.



شکل ۱۱. منحنی ROC برای رویکرد پیشنهادی دوم



(الف)



(ب)

شکل ۱۲. نمودار دمانگاشت مربوط به دسته‌داده تست Drishti-GS1 برای رویکرد پیشنهادی (الف): اول و (ب): دوم

مقاله از نسبت مساحت کاپ به مساحت دیسک نوری (ACDR) استفاده می‌شود. بطور کلی در [۵] مقدار CDR بالاتر از ۰٫۵ در نسبت قطری کاپ به دیسک نوری که معادل ۰٫۲۵، در نسبت مساحت کاپ به مساحت دیسک نوری است بعنوان احتمال وجود بیماری گلوکوم در نظر گرفته شده است که در [۲۱] بهترین سطح آستانه برای طبقه‌بندی براساس CDR در دسته‌داده Drishti-GS1 برابر با ۰٫۴۰۷، یعنی ۰٫۶۳ در نسبت قطری کاپ به دیسک نوری بدست آمده است. به منظور تست دسته‌داده Drishti-GS1، ۱۵ تصویر از مجموعه تست این دسته‌داده شامل ۵ تصویر نرمال و ۱۰ تصویر گلوکومی انتخاب شده است. همچنین برای تست دسته‌داده RIM-ONE v3، ۲۴ تصویر شامل ۱۳ تصویر نرمال و ۱۱ تصویر گلوکومی بعنوان داده‌های تست انتخاب شده است. منحنی ROC برای دسته‌داده Drishti-GS1 و RIM-ONE v3 در شکل ۱۱ قابل مشاهده است. سطح زیر منحنی ROC برای این دو دسته‌داده به ترتیب برابر با ۰٫۹۴ و ۰٫۸۲۵ است.

نمودار دمانگاشت مربوط به ضریب همبستگی متقابل پیرسون در سنجش CDR برای روش پیشنهادی، متخصصان و حقیقت مبنا و در دسته‌داده Drishti_GS1 در شکل ۱۲ قابل مشاهده است. همانگونه که در نمودار شکل ۱۲ مشاهده می‌شود نتایج مربوط به پیش‌بینی CDR مدل پیشنهادی در دسته‌داده Drishti_GS1 در رویکرد اول از نظر همبستگی با نتایج CDR مربوط به حقیقت مبنا به میزان ۲٪ نسبت به متخصص چهارم (Ex_4) بدتر عملکرد است و از متخصص دوم (Ex_2) به میزان ۸٪ و از متخصص اول (Ex_1) به میزان ۶٪ بهتر عمل کرده است. نتایج مربوط به CDR مدل پیشنهادی در رویکرد دوم از نظر همبستگی با نتایج CDR مربوط به حقیقت مبنا نسبت به متخصص چهارم (Ex_4) به میزان ۳٪، نسبت به متخصص سوم (Ex_3) به میزان ۱٪، نسبت به متخصص دوم (Ex_2) به میزان ۶٪ و نسبت به متخصص اول (Ex_1) به میزان ۶٪ بهتر عمل کرده است.

نمودار دمانگاشت مربوط به ضریب همبستگی متقابل پیرسون برای دسته‌داده RIM-ONE v3 در سنجش CDR و در روش پیشنهادی، متخصصان و حقیقت مبنا در شکل ۱۳ قابل مشاهده است. همانگونه که مشاهده می‌شود نتایج مربوط به پیش‌بینی CDR در رویکرد پیشنهادی اول از نظر همبستگی با نتایج CDR مربوط به حقیقت مبنا به میزان ۷٪ نسبت به متخصص دوم (Ex_2) و ۶٪ نسبت به متخصص اول (Ex_1) بدتر عمل کرده است. همچنین نتایج مربوط به پیش‌بینی CDR در رویکرد پیشنهادی دوم از نظر همبستگی با نتایج CDR مربوط به حقیقت مبنا به میزان ۴٪ نسبت به متخصص دوم (Ex_2) و ۳٪ نسبت به متخصص اول (Ex_1) بدتر عمل کرده است.

باتوجه به نمودارهای دمانگاشت برای ارزیابی نسبت کاپ به دیسک می‌توان نتیجه گرفت بطور کلی رویکرد پیشنهادی دوم نسبت به رویکرد پیشنهادی اول در هر دو دسته‌داده بهتر عمل می‌نماید و حقیقت مبنای مورد نظر در رویکرد پیشنهادی دوم برای

یک روش کارا برای غربالگری اولیه بیماری گلوکوم بر اساس محاسبه نسبت کاپ به دیسک نوری با استفاده از شبکه‌های عصبی کانولوشنی

جدول ۱۲. نتایج قطعه‌بندی برای دسته‌داده Drishti-GS1 (%)

روش	کاپ نوری		دیسک نوری	
	IoU	F ₁ -score	IoU	F ₁ -score
سواستوپولسکی [۱۵]	۸۵,۰	۷۵,۰	-	-
سواستوپولسکی [۱۶]	۸۹,۰	۸۰,۰	۹۷,۰	۹۵,۰
زیلی [۱۴]	۸۷,۱	۸۵,۰	۹۷,۳	۹۱,۴
البندر [۱۹]	۸۲,۸	۷۱,۱	۹۰,۴	۹۴,۹
شوانگ یو [۵]	۸۸,۷	۸۰,۴	۹۷,۴	۹۴,۹
الگوریتم ارائه‌شده در رویکرد پیشنهادی اول	۹۲,۶	۸۶,۵	۹۷,۷	۹۵,۶
الگوریتم ارائه‌شده در رویکرد پیشنهادی دوم	۹۳,۴	۸۷,۷	۹۷,۶	۹۵,۳

جدول ۱۳. نتایج قطعه‌بندی برای دسته‌داده RIM-ONE v3 (%)

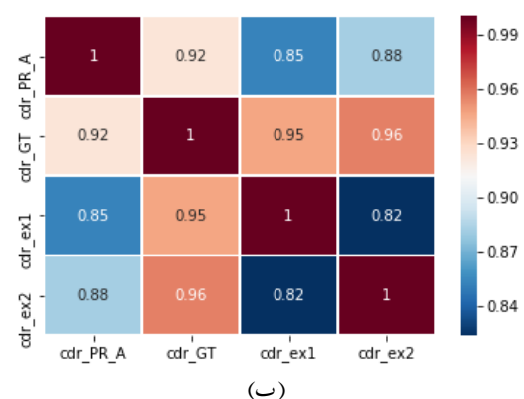
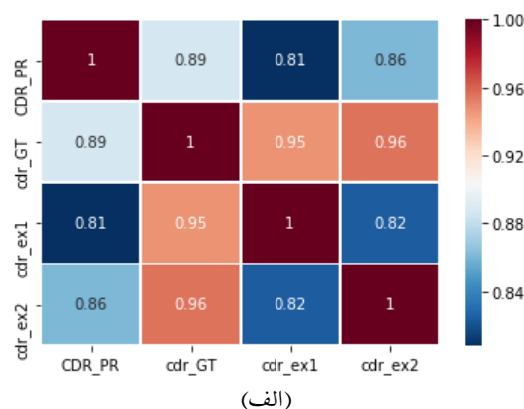
روش	کاپ نوری		دیسک نوری	
	IoU	F ₁ -score	IoU	F ₁ -score
سواستوپولسکی [۱۵]	۸۲,۰	۶۹,۰	۸۹,۰	۹۵,۰
سواستوپولسکی [۱۶]	۸۴,۰	۷۳,۰	۹۶,۰	۹۲,۰
زیلی [۱۴]	۸۲,۴	۸۰,۲	۹۴,۲	۸۹,۰
شوانگ یو [۵]	۸۴,۴	۷۴,۳	۹۶,۱	۹۲,۶
البندر [۱۹]	۶۹,۰	۵۵,۷	۹۰,۴	۸۲,۹
الگوریتم ارائه‌شده در رویکرد پیشنهادی اول	۷۵,۸	۶۴,۳	۹۵,۵	۹۱,۴
الگوریتم ارائه‌شده در رویکرد پیشنهادی دوم	۷۸,۷	۶۶,۵	۹۵,۵	۹۱,۵

جدول ۱۴. نتایج قطعه‌بندی برای دسته‌داده MESSIDOR (%)

روش	کاپ نوری		دیسک نوری	
	IoU	F ₁ -score	IoU	F ₁ -score
کیم [۲۳]	۹۰,۷	۸۳,۳	۹۷,۶	۹۵,۳
الگوریتم ارائه‌شده در رویکرد پیشنهادی اول	۸۳,۰	۷۲,۲	۹۳,۲	۸۸,۲
الگوریتم ارائه‌شده در رویکرد پیشنهادی دوم	۸۳,۶	۷۳,۰	۹۳,۳	۸۸,۵

۵ جمع‌بندی و نتیجه‌گیری

در این مقاله ابتدا با استفاده از الگوریتم خودکار RCNN سریعتر، دیسک نوری استخراج شد. با این فرض که IOU بیشتر از ۵۰٪ به معنای محلی‌سازی درست باشد، کارایی محلی‌سازی ۱۰۰٪ برای دسته‌داده Drishti_GS1، بدست آمد. سپس شبکه U-Net اصلاح‌شده که از SE-ResNet50 بعنوان لایه‌های نمونه‌بردار با نرخ پایین در قسمت کدگذار خود استفاده می‌کند، برای قطعه‌بندی کاپ و دیسک نوری بکار رفت. در الگوریتم پیشنهادی و دسته‌داده Drishti_GS1 به ترتیب مقدار ۹۲,۶٪ و ۹۳,۴٪ برای قطعه‌بندی کاپ نوری و به ترتیب مقدار ۹۷,۷٪ و ۹۷,۶٪ برای قطعه‌بندی دیسک نوری در دو رویکرد پیشنهادی اول و دوم برای معیار F₁-score بدست آمد که بهبود قابل توجهی را



شکل ۱۳. نمودار دمانگاشت مربوط به دسته‌داده تست RIM-ONE v3 برای رویکرد پیشنهادی (الف): اول و (ب): دوم

در ادامه به مقایسه نتایج بدست آمده از کار ما با سایر کارهای صورت‌گرفته در زمینه قطعه‌بندی تصاویر پرداخته می‌شود. به‌منظور مقایسه با سایر روش‌های موجود، نتایج حاصل از قطعه‌بندی در دسته‌داده Drishti-GS1 در جدول ۱۲ آمده است. باتوجه به این جدول مشاهده می‌شود که قطعه‌بندی کاپ و دیسک نوری حاصل از الگوریتم ارائه‌شده در دو رویکرد پیشنهادی برای دسته‌داده Drishti-GS1 قابل رقابت با پیشرفته‌ترین الگوریتم‌ها است و حتی از آنها پیشی می‌گیرد.

همچنین، نتایج حاصل از قطعه‌بندی بر روی دسته‌داده RIM-ONE v3 در جدول ۱۳ آورده شده است. با توجه به این نتایج مشاهده می‌شود که اگرچه در قطعه‌بندی کاپ نوری بهبودی صورت نگرفته ولی قطعه‌بندی دیسک نوری در روش پیشنهادی و در دو رویکرد پیشنهادی در دسته‌داده RIM-ONE v3 قابل رقابت با پیشرفته‌ترین الگوریتم‌ها است.

نتایج حاصل از قطعه‌بندی دسته‌داده MESSIDOR در جدول ۱۴ آمده است. باتوجه به این نتایج مشاهده می‌گردد که این نتایج قابل رقابت با نتایج کار کیم و همکارانش در پژوهش [۲۳] است و این نکته نیز حائزاهمیت است که در کار کیم و همکارانش از دسته‌داده MESSIDOR در مرحله آموزش نیز استفاده شده است در حالیکه در کار ما این دسته‌داده در مرحله آموزش وجود نداشت و فقط در مرحله تست استفاده شده است.

[۶] جوانمردی، شیما و زارع چاهوکی، محمدعلی، "پالایش شرح‌گذاری مجموعه تصاویر با مقیاس بزرگ با یادگیری انتقالی در شبکه عصبی کانولوشنال عمیق"، مجله ماشین بینایی و پردازش تصویر، شماره اول، ص ۳۹ تا ۵۲، ۱۳۹۷.

- [7] Veena, H. N., Muruganandham, A., and Kumaran, T.S., "A Review on the optic disc and optic cup segmentation and classification approaches over retinal fundus images for detection of glaucoma," SN Appl. Sci. 2, no. 1476, 2020.
- [8] Damon, W. W. K., Liu, J., Meng, T. N., Fengshou, Y., and Yin, W. T., "Automatic detection of the optic cup using vessel kinking in digital retinal fundus images," in 9th IEEE International Symposium on Biomedical Imaging (ISBI), Barcelona, 2012.
- [9] Cheng, J., Liu, J., Xu, Y., Yin, F.S., Wong, D.W.K., Tan, N.M., Tao, D.C., Cheng, C.-Y., and Aung, T., "Superpixel Classification Based Optic Disc and Optic Cup Segmentation for Glaucoma Screening," IEEE Transactions on Medical Imaging, vol. 32, no. 6, pp. 1019–1032, 2013.
- [10] Noor, N., Khalid, N., and Ariff, N., "Optic cup and disc color channel multi-thresholding segmentation," in 2013 IEEE International Conference on Control System, Computing and Engineering, Mindeb, 2013.
- [11] Dutta, M. K., Mourya, A. K., Singh, A., Parthasarathi, M., Burget, R. and Riha, K., "Glaucoma Detection by Segmenting the Super Pixels from Fundus Colour Retinal Images," in International Conference on Medical Imaging, M-Health and Emerging Communication Systems (MedCom), Greater Noida, 2014.
- [12] Alghmdy, H., Tang, H. L., Hansen, M., O'Shea, A., Al Turk, L., and Peto, T., "Measurement of optical cup-to-disc ratio in fundus images for glaucoma screening," in International Workshop on Computational Intelligence for Multimedia Understanding (IWCIM), Prague, 2015.
- [13] Mittapalli, P.S. and Kande, G.B., "Segmentation of optic disk and optic cup from digital fundus images for the assessment of glaucoma," Biomedical Signal Processing Control, vol. 24, p. 34–46, 2016.
- [14] Zilly, J., Buhmann, J. M., and Mahapatra, D., "Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation," Computerized Medical Imaging and Graphics, vol. 55, pp. 28–41, 2017.
- [15] Sevastopolsky, A., "Optic disc and cup segmentation methods for glaucoma detection with modification of U-Net convolutional neural network," Computer Vision and Pattern Recognition, vol. 27, p. 618–624, 2017.
- [16] Sevastopolsky, A., Drapak, S., Kiselev, K., Snyder, B.M., and Georgievskaya, A., Stack-U-Net: Refinement Network for Image Segmentation on the Example of

در قطعه‌بندی کاپ و دیسک نوری در این دسته‌داده نسبت به کارهای جدید ارائه داده است. همچنین در دسته‌داده RIM-ONE v3 به ترتیب مقدار ۷۵٫۸٪ و ۷۸٫۷٪ برای قطعه‌بندی کاپ نوری در رویکرد پیشنهادی اول و دوم و نیز به ترتیب مقدار ۹۵٫۵٪ و ۹۵٫۵٪ برای قطعه‌بندی دیسک نوری برای رویکرد پیشنهادی اول و دوم برای معیار F_1 -score بدست آمد. در دسته‌داده MESSIDOR که در فاز آموزش وجود نداشت به ترتیب مقدار ۸۳٫۰٪ و ۸۳٫۶٪ برای قطعه‌بندی کاپ نوری در رویکرد اول و دوم و نیز به ترتیب مقدار ۹۳٫۲٪ و ۹۳٫۳٪ برای قطعه‌بندی دیسک نوری در رویکرد اول و دوم برای معیار F_1 -score بدست آمد. از نظر همبستگی با حقیقت مبنا، ضریب همبستگی برای CDR بدست آمده از الگوریتم پیشنهادی برای دسته‌داده Drishti_GS1 به ترتیب ۰٫۹۴ و ۰٫۹۶، در رویکرد پیشنهادی اول و دوم است که این همبستگی در رویکرد پیشنهادی دوم از متخصصان پیشی می‌گیرد. در دسته‌داده RIM-ONE v3 این مقادیر به ترتیب ۰٫۸۹ و ۰٫۹۲، در رویکرد پیشنهادی اول و دوم است.

ذکر این نکته لازم است که محدود بودن دسته‌داده‌های آموزشی که دارای حاشیه‌نویسی‌های قطعه‌بندی دیسک و کاپ نوری باشند، الگوریتم ارائه‌شده را به چالش می‌کشد. لذا علیرغم تعداد داده زیاد که در دسترس عموم قرار دارد، تعداد کمی از آنها قابل استفاده برای آموزش شبکه است. علاوه بر این، عدم توازن بین تعداد داده‌های نرمال و گلوکومی در پایگاه‌های داده مورد استفاده به چشم می‌خورد. به‌عنوان یک پیشنهاد جهت کار در آینده، درجه‌بندی میزان پیشرفت بیماری گلوکوم مورد بررسی قرار خواهد گرفت.

مراجع

- [1] Barton, K. and Hitchings, R. A., *Medical Management of Glaucoma*, London, UK: Springer Healthcare Communications, 2013, pp. 71–100.
- [2] Serener A. and Serte S., "Transfer Learning for Early and Advanced Glaucoma Detection with Convolutional Neural Networks," in Medical Technologies Congress (TIPTTEKNO), Izmir, Turkey, 2019.
- [3] Morales, S., Naranjo, V., Angulo, J., and Alcañiz, M., "Automatic Detection of Optic Disc Based on PCA and Mathematical Morphology," IEEE Transactions on Medical Imaging, vol. 32, pp. 786–796, 2013.
- [4] Xiao, Z., Zhang, X., Geng, L., Zhang, F., Wu, J., and Liu, Y., "Research on the Method of Color Fundus Image Optic Cup Segmentation Based on Deep Learning," Symmetry, vol. 11, no. 7, 2019.
- [5] Yu, S., Xiao, D., Frost, S., and Kanagasigam, Y., "Robust optic disc and cup segmentation with deep learning for glaucoma detection," Computerized Medical Imaging and Graphics, vol. 74, pp. 61–71, 2019.

- Computer Vision and Pattern Recognition, 2018, pp. 7132–7141.
- [۲۸] قنبری سرخی، علی، حسن پور، حمید و فاتح، منصور، "بهبود شبکه عمیق R-FCN در آشکارسازی و برجسبزی اشیاء"، مجله ماشین بینایی و پردازش تصویر، شماره ۲، ص ۴۳ تا ۵۹، ۱۳۹۸.
- [29] Ren, S., He, K., Girshick, R., and Sun, J., "Faster R-CNN: Towards real-time object detection with region proposal networks," IEEE Transactions on Pattern Analysis & Machine Intelligence, vol. 39, pp. 1137–1149, 2017.
- [۳۰] نصیری، مجید، رشیدی کنعان، حمیدرضا و امیری، سید حمید، "ویژگی‌های آگاه به محتوا برای قطعه‌بندی معنایی تصویر"، مجله ماشین بینایی و پردازش تصویر، ارسال شده برای انتشار، ۱۳۹۸.
- [31] Ronneberger, O., Fischer, P., and Brox, T., "U-Net: Convolutional Networks for Biomedical Image Segmentation," in Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015.
- [32] Acharya, U. R., Ng, E. Y. K., Eugene, L. W. J., Noronha, K. P., Min, L. C., Nayak, K. P., and Bhandary, S. V., "Decision support system for the glaucoma using Gabor transformation," Biomedical Signal Processing and Control, vol. 15, pp. 18–26, 2015.
- [33] Gajbhiye, G. O. and Kamthane, A. N., "Automatic classification of glaucomatous images using wavelet and moment feature," in Annual IEEE India Conference (INDICON), New Delhi, 2015.
- [34] Orlando, J. I., Prokofyeva, E., Fresno, M. d., and Blaschko, M. B., "Convolutional neural network transfer for automated glaucoma identification," in 12th International Symposium on Medical Information Processing and Analysis, 2017, pp. 10160–10170.
- [35] Sivaswamy, J., Krishnadas, S. R., Chakravarty, A., Joshi, G. D., and Ujjwal, "A Comprehensive Retinal Image Dataset for the Assessment of Glaucoma from the Optic Nerve Head Analysis," JSM Biomed Imaging Data Papers, vol. 2, no. 1, 2015.
- [36] Fumero, F., Alayón, S., Sanchez, J.L., Sigut, J., and Hernandez, G., "RIM-ONE: An open retinal image database for optic nerve evaluation," in 24th International Symposium on Computer-Based Medical Systems (CBMS), 2011, pp. 1–6.
- [37] Almazroa, A., Alodhayb, S., Osman, E., Ramadan, E., Hummadi, M., Dlain, M., Alkatee, M., Raahemifar, K., and Lakshminarayanan, V., "Retinal fundus images for glaucoma analysis: the RIGA dataset," in Proc. SPIE 10579, Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications, 2018.
- [38] High-Resolution Fundus (HRF) Image Database, Technische Facultat, [Online]. Available: Optic Disc and Cup, arXiv preprint arXiv: 1804.11294, 2018.
- [17] Fu, H., Cheng, J., Xu, Y., Wong, D. W. K., Liu, J., and Cao, X., "Joint Optic Disc and Cup Segmentation Based on Multi-Label Deep Network and Polar Transformation," IEEE Transactions on Medical Imaging, vol. 37, no. 7, pp. 1597–1605, 2018.
- [18] Ferreira, M. V. d. S., Filho, A. O. d. C., Sousa, A. D. d., Silva, A. C., and Gattass, M., "Convolutional neural network and texture descriptor-based automatic detection and diagnosis of Glaucoma," Expert Systems With Applications, vol. 110, pp. 250–263, 2018.
- [19] Al-Bander, B., Williams, B. M., Al-Nuaimy, W., Al-Tae, M. A., Pratt, H., and Zheng, Y., "Dense Fully Convolutional Segmentation of the Optic Disc and Cup in Colour Fundus for Glaucoma Diagnosis," Symmetry, vol. 10, no. 4, pp. 87–103, 2018.
- [20] Jiang, Y., Duan, L., Cheng, J., Gu, Z., Xia, H., Fu, H., Li, C., and Liu, J., "JointRCNN: A Region-based Convolutional Neural Network for Optic Disc and Cup Segmentation," IEEE Transactions on Biomedical Engineering, vol. 67, pp. 335–343, 2019.
- [21] Mvoulana, A., Kachouri, R., and Akil, M., "Fully automated method for glaucoma screening using robust optic nerve head detection and unsupervised segmentation based cup-to-disc ratio computation in retinal fundus images," Computerized Medical Imaging and Graphics, vol. 77, pp. 1–19, 2019.
- [22] Thakur, N. and Juneja, M., "Optic Disc and Optic Cup Segmentation from Retinal Images Using Hybrid Approach," Expert Systems with Applications, vol. 127, pp. 308–322, 2019.
- [23] Kim, J., Tran, L., Chew, E. Y., and Antani, S., "Optic Disc and Cup Segmentation for Glaucoma Characterization Using Deep Learning," in IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS), Cordoba, Spain, 2019.
- [24] Carrillo, J., Bautista, L., Villamizar, J., Rueda, J., Sanchez, M., and Rueda, D., "Glaucoma Detection Using Fundus Images of The Eye," in 2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA), Bucaramanga, Colombia, 2019.
- [25] He, K., Zhang, X., Ren, S., and Sun, J., "Deep Residual Learning for Image Recognition," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.
- [26] He, K., Zhang, X., Ren, S., and Sun, J., "Identity Mappings in Deep Residual Networks," arXiv:1603.05027 [cs.CV], 2017.
- [27] Hu, J., Shen, L., Albanie, S., and Sun, G., "Squeeze-and-Excitation Networks," in IEEE/CVF Conference on

<https://www5.cs.fau.de/research/data/fundus-images>.

- [39] Zhang, Z., Yin, F. S., Liu, J., Wong, W. K., Tan, N. M., Lee, B. H., and Cheng, J., "ORIGA-light : An Online Retinal Fundus Image Database for Glaucoma Analysis and Research," in Annual International Conference of the IEEE Engineering in Medicine and Biology, 2010, pp. 3065-3068.
- [40] Mahmudi, T., Kafieh, R., and Rabbani, H., "Comparison of macular OCTs in right and left eyes of normal people," in SPIE 9038, Medical Imaging 2014: Biomedical Applications in Molecular, Structural, and Functional Imaging, San Diego, California, 2014.



فاطمه معادی مدرک کارشناسی خود را در رشته مهندسی برق گرایش مخابرات از دانشگاه بین‌المللی امام خمینی (ره) قزوین دریافت کرد. سپس مدرک کارشناسی ارشد خود را در گرایش مخابرات-سیستم از همان دانشگاه اخذ نمود. حوزه‌های پژوهشی و مورد علاقه ایشان پردازش سیگنال، پردازش تصویر، شناسایی آماری الگو، بینایی ماشین، یادگیری ماشین و یادگیری عمیق است.



ندا فرجی مدرک کارشناسی خود را در رشته مهندسی برق گرایش الکترونیک از دانشگاه علم و صنعت ایران دریافت نمود. سپس در رشته مهندسی برق گرایش الکترونیک دیجیتال و در مقاطع کارشناسی ارشد و دکترا از دانشگاه صنعتی امیرکبیر فارغ‌التحصیل گردید. ایشان در سال ۱۳۹۰ یک دوره فرصت مطالعاتی نه ماهه را در دانشگاه دلفت هلند گذرانده و از سال ۱۳۹۲ تاکنون استادیار گروه مهندسی برق-مخابرات سیستم در دانشگاه بین‌المللی امام خمینی (ره) هستند. زمینه کاری مورد علاقه ایشان، پردازش سیگنال گفتار، یادگیری ماشین، یادگیری عمیق و پردازش تصویر می‌باشد.



محمدرضا حسن نژاد بی بالان مدرک کارشناسی خود را در رشته مهندسی برق، گرایش الکترونیک از دانشگاه گیلان اخذ نمود. مقطع کارشناسی ارشد مهندسی برق، گرایش مخابرات سیستم را در دانشگاه صنعتی شریف به پایان رساند. او مدرک دکتری تخصصی مهندسی برق، گرایش مخابرات سیستم را در سال ۱۳۹۵ از دانشگاه صنعتی امیرکبیر دریافت نمود. وی هم‌اکنون به عنوان عضو هیأت علمی در دانشگاه بین‌المللی امام خمینی (ره) مشغول به فعالیت می‌باشد. از زمینه‌های پژوهشی مورد علاقه ایشان، می‌توان به پردازش سیگنال آماری، پردازش تصویر دیجیتال، تئوری تخمین و آشکارسازی اشاره نمود.