

تفکیک عروق شبکیه در تصاویر فوندوس در شرایط افت کیفیت ناشی از آب‌مروارید با بهره‌گیری از شبیه‌سازی آگاه از میدان دید و یادگیری عمیق

آرش آقائی^۱، رضا حسن‌زاده پاک‌رضائی^۲ و اسدالله شاه‌بهرامی^۳

چکیده

تفکیک خودکار عروق شبکیه یکی از مؤلفه‌های اساسی در سامانه‌های کمک‌تشخیصی مبتنی بر تصاویر فوندوس شبکیه به‌شمار می‌رود. با این حال، وجود آب‌مروارید در بخش قابل توجهی از تصاویر بالینی موجب افت کیفیت تصویر شده و دقت روش‌های متداول تفکیک عروق را با چالش مواجه می‌سازد. در این پژوهش، تمرکز اصلی بر ارائه یک چارچوب مرحله‌ای، کنترل‌شده و آگاه از میدان دید برای بررسی پایداری الگوریتم‌های تفکیک عروق شبکیه در حضور افت کیفیت ناشی از آب‌مروارید است. در گام نخست، اثرات فیزیکی آب‌مروارید شامل تاری، تضعیف تمایز، نویز و پراکندگی نور به‌صورت کنترل‌شده و محدود به ناحیه میدان دید شبیه‌سازی شده و نسخه‌هایی از تصاویر با سطوح مختلف شدت تخریب تولید می‌گردند. سپس، چند معماری شاخص تفکیک عروق بر روی پایگاه داده مرجع FIVES آموزش داده شده و مدل‌های برتر بر اساس معیارهای IoU، Dice و Accuracy انتخاب می‌شوند. در ادامه، عملکرد این مدل‌ها در مواجهه با افزایش تدریجی شدت آب‌مروارید بر روی نسخه‌های شبیه‌سازی شده FIVES و همچنین به‌صورت مکمل بر روی نسخه‌های شبیه‌سازی شده پایگاه داده کلاسیک DRIVE ارزیابی می‌گردد. با توجه به نبود پایگاه داده عمومی استاندارد شامل تصاویر فوندوس مبتلا به آب‌مروارید همراه با ماسک پیکسلی عروق، استفاده از شبیه‌سازی کنترل‌شده امکان حفظ برجسب مرجع و انجام ارزیابی کمی قابل تکرار را فراهم می‌سازد. نتایج کمی و کیفی نشان می‌دهند که با افزایش شدت آب‌مروارید، معیارهای هم‌پوشانی نظیر IoU و Dice به‌صورت تدریجی کاهش می‌یابند؛ امری که موجب تضعیف رؤیت‌پذیری عروق باریک و کاهش تفکیک‌پذیری عروق از بافت زمینه شبکیه می‌گردد. با این حال، در چارچوب ارزیابی تعریف شده، ترکیب UNet با رمزگذار EfficientNet-B3 شیب افت عملکرد ملایم‌تری نسبت به سایر معماری‌ها نشان می‌دهد و در سطوح بالاتر تخریب نیز حفظ ساختار کلی شبکه عروقی را تا حد قابل قبولی ممکن می‌سازد.

کلیدواژه‌ها

تفکیک عروق شبکیه، تصاویر شبکیه (RF)، آب‌مروارید (CA)، یادگیری عمیق، UNet، شبیه‌سازی افت کیفیت تصویر.

۱ مقدمه

تصویر برداری فوندوس شبکیه (RF) یکی از مهم‌ترین روش‌های ارزیابی غیرتهاجمی در چشم‌پزشکی است و نقش اساسی در تشخیص، پایش و غربالگری طیف وسیعی از بیماری‌های شبکیه ایفا می‌کند. این تصاویر به‌طور گسترده برای شناسایی بیماری‌هایی

این مقاله در ۲۵ بهمن ۱۴۰۴ دریافت و در ۱۰ مرداد ۱۴۰۵ پذیرفته شد.

^۱ دانشجوی دکترای هوش مصنوعی، دانشگاه گیلان ایران

رایانامه: arashaghaci@webmail.guilan.ac.ir

^۲ دانشیار، دانشکده فنی، دانشگاه گیلان

رایانامه: hasanzadehpak@guilan.ac.ir

^۳ استاد، دانشکده فنی، دانشگاه گیلان

رایانامه: shahbahrani@guilan.ac.ir

نویسنده مسئول: رضا حسن‌زاده پاک‌رضائی

مانند رنجوری دیابتی^۱، آب‌سیاه^۲، لکه زرد شبکیه^۳ و سایر اختلالات شبکیه مورد استفاده قرار می‌گیرند [۱].

تحلیل ساختارهای شاخص شبکیه به‌ویژه شبکه عروق خونی، دیسک بینایی^۴ و ناحیه لکه‌زرد شبکیه^۵ امکان تشخیص زودهنگام بیماری‌ها و جلوگیری از پیشرفت آسیب‌های بینایی را فراهم می‌سازد [۱، ۲].

در میان ساختارهای قابل مشاهده در RF، شبکه عروقی شبکیه یک منبع اطلاعاتی بسیار غنی است؛ زیرا تغییرات مورفولوژی و هندسه عروق می‌تواند نشانه‌ای از بیماری‌های چشمی و وضعیت عمومی سلامت عروقی باشد. از این رو، بخش‌بندی^۶ خودکار عروق شبکیه به‌عنوان یکی از گام‌های پایه در بسیاری از سامانه‌های کمک تشخیصی رایانه‌ای^۷ و غربالگری مطرح است [۳، ۴].

در مطالعات گذشته، روش‌هایی مبتنی بر فیلترهای تطبیقی، آشکارسازهای شیاری^۸، پردازش‌های ریخت‌شناختی و آستانه‌گذاری های کلاسیک و تطبیقی برای استخراج عروق پیشنهاد شده‌اند [۵-۸]. این روش‌ها عمدتاً در شرایط تصویربرداری استاندارد و تصاویر با کیفیت مناسب، عملکرد قابل قبولی از خود نشان داده‌اند [۵]. اما با وجود کارایی مناسب این رویکردها در شرایط تصویربرداری استاندارد، وابستگی آن‌ها به کیفیت تصویر و حساسیت به افت تمایز، نویز و تاری باعث می‌شود در شرایط بالینی واقعی، پایداری کافی نداشته باشند.

با ظهور یادگیری عمیق^۹، شبکه‌های عصبی کانولوشنی^{۱۰} و به‌ویژه معماری UNet به‌عنوان یکی از موفق‌ترین چارچوب‌ها برای بخش‌بندی ساختارهای زیست‌پزشکی [۹] و استفاده از ساختار رمزگذار-رمزگشا^{۱۱} به همراه Skip Connections، امکان استخراج هم‌زمان اطلاعات سراسری و محلی را فراهم کرده و باعث بهبود دقت بخش‌بندی عروق شبکیه در تصاویر RF شده است [۱۰، ۱۱]. با این حال، بخش عمده‌ای از مطالعات بر روی تصاویر با کیفیت بالا انجام شده‌اند و پایداری این الگوریتم‌ها در مواجهه با افت کیفیت تصویر به‌صورت سیستماتیک مورد بررسی قرار نگرفته است.

یکی از مهم‌ترین عوامل بالینی مؤثر بر کیفیت تصاویر RF، وجود بیماری آب‌مروارید^{۱۲} (CA) در چشم است. این بیماری با ایجاد کدورت در عدسی چشم، موجب پراکندگی نور و ایجاد

نوعی هاله‌ی نوری در مسیر ثبت تصویر می‌شود که نتیجه‌ی آن افت محسوس تمایز تصویر است [۵]. تصاویری که از چشم مبتلا به CA گرفته می‌شوند معمولاً بصورت مات و مه‌آلود^{۱۳} دیده شده و لبه‌ها و جزئیات ریز ساختارهای شبکیه، تار و محو به‌نظر می‌رسند [۵، ۱۲]. علاوه بر این، کدورت عدسی و پراکندگی حاصل از آن توزیع روشنایی در تصویر را دستخوش تغییر می‌کند؛ بدین صورت که تصویر حالت یکنواخت‌تر و تیره‌تری پیدا کرده و تشخیص ساختارهای شبکیه دشوارتر می‌گردد [۱۳]. این کاهش کیفیت ناشی از CA می‌تواند تشخیص دقیق بیماری‌های چشمی از روی تصاویر RF را با چالش جدی مواجه کند [۱۲].

در برخی پژوهش‌ها، شدت CA در تصاویر RF به‌صورت کیفی و بر اساس ارزیابی بالینی در چند سطح گسسته مانند خفیف، متوسط و شدید دسته‌بندی شده است [۱۴]. این دسته‌بندی‌ها عمدتاً با هدف تشخیص خود CA انجام شده است. همچنین، در برخی مطالعات دیگر، شدت عملکردی CA بر اساس شاخص‌هایی نظیر حدت بینایی اصلاح‌شده^{۱۴} تعریف شده و تصاویر RF صرفاً به‌عنوان ورودی سامانه‌های طبقه‌بندی برای تشخیص وجود یا شدت CA مورد استفاده قرار گرفته‌اند [۱۵]. با این حال، این رویکردها وارد مرحله تحلیل ساختارهای شبکیه، از جمله تفکیک عروق، نمی‌شوند.

از دیدگاه تحلیل عروق شبکیه، هدف صرفاً تشخیص وجود CA نیست، بلکه بررسی میزان پایداری الگوریتم‌های بخش‌بندی در حضور درجات مختلف افت کیفیت ناشی از CA اهمیت دارد. در این سناریو، تعریف شدت CA به‌گونه‌ای که تصویر کاملاً غیرقابل تشخیص گردد، عملاً امکان تحلیل ساختارهای شبکیه را از بین می‌برد و با هدف مسئله هم‌راستا نیست. از این رو، در این پژوهش شدت‌های مختلف CA به‌صورت سطوح افزایشی تخریب تصویر تعریف شده‌اند که در آن‌ها، با وجود افت محسوس کیفیت، ساختارهای اصلی شبکیه همچنان قابل تشخیص باقی می‌مانند. نمونه‌هایی از تصاویر واقعی RF، صرفاً به‌منظور ارائه درک بصری از دامنه تغییرات کیفیت تصویر در حد بالای حالت تخریب در شکل (۱) نمایش داده شده‌اند [۱۴].

بر اساس گزارش‌های اپیدمیولوژیک^{۱۵}، CA یکی از اصلی‌ترین علل نابینایی قابل پیشگیری در جهان بوده و سهم قابل توجهی از اختلالات بینایی را به خود اختصاص می‌دهد [۴]. از این رو، بخش قابل توجهی از تصاویر RF مورد استفاده در محیط‌های بالینی، به‌طور بالقوه تحت تأثیر درجات مختلف CA قرار دارند.

Diabetic Retinopathy^۱
Glaucoma^۲
Age-Related Macular Degeneration^۳
Optic Disc^۴
Macula^۵
Segmentation^۶
Computer-Aided Diagnosis^۷
Ridge Detectors^۸
Deep Learning^۹
Convolutional Neural Networks^{۱۰}
Encoder-Decoder^{۱۱}
Cataract (CA)^{۱۲}

Haze^{۱۳}
Best Corrected Decimal Visual Acuity^{۱۴}
Epidemiological^{۱۵}

از سوی دیگر، برخی پژوهش‌ها تلاش کرده‌اند شدت تاری ناشی از CA را به‌صورت کمی ارزیابی کنند. در [۸] روشی مبتنی بر استخراج ویژگی‌هایی نظیر تمایز محلی، تعداد پیکسل‌های ساختارهای قابل مشاهده و انحراف معیار شدت روشنایی پیشنهاد داده‌شد که امکان طبقه‌بندی تصاویر RF به سطوح مختلف تاری را فراهم می‌سازد. با وجود این تلاش‌ها، هنوز یک چارچوب یکپارچه برای استخراج پایدار عروق شبکیه در حضور درجات مختلف CA به‌طور کامل توسعه نیافته است.

برای رفع این خلأ، در این مقاله یک چارچوب مرحله‌ای، کنترل‌شده و آگاه از میدان دید^۴ (FOV) برای ارزیابی پایداری الگوریتم‌های تفکیک عروق شبکیه در حضور افت کیفیت ناشی از CA ارائه می‌شود. مراحل اصلی این چارچوب به‌طور مشخص عبارت‌اند از:

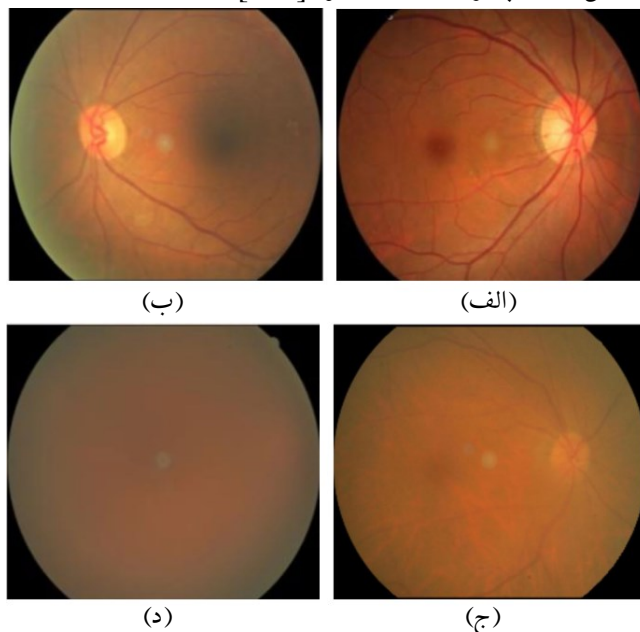
(۱) شبیه‌سازی چندمرحله‌ای اثرات اصلی CA شامل تاری اپتیکی، افت تمایز، نویز و پراکندگی نور به‌صورت کنترل‌شده و در سه سطح شدت، (۲) اعمال افت کیفیت به‌صورت آگاه و محدود به ناحیه FOV، به‌منظور جلوگیری از ایجاد نشانه‌های مصنوعی در نواحی خارج از میدان دید، (۳) ساخت نسخه‌های شبیه‌سازی‌شده و تخریب‌شده پایگاه‌داده‌های FIVES و DRIVE در سه سطح شدت CA، همراه با حفظ برجسته مرجع پیکسلی عروق برای انجام ارزیابی کمی قابل تکرار، (۴) بررسی انواع معماری‌های یادگیری عمیق کارآمد در تفکیک عروق شبکیه و آموزش، ارزیابی و مقایسه معماری‌های منتخب در چهارچوب کنترل شده پیشنهادی و انتخاب مدل‌های برتر و (۵) تحلیل روند افت عملکرد مدل‌ها با افزایش شدت CA و بررسی تعمیم‌پذیری بین پایگاه‌داده‌ای از طریق ارزیابی مکمل بر روی پایگاه‌داده مستقل DRIVE.

در ادامه مقاله، ابتدا در بخش ۲ فرآیند شبیه‌سازی تخریب ناشی از CA در تصاویر RF مورد بررسی قرار می‌گیرد. سپس در بخش ۳ پایگاه‌داده‌های مورد استفاده معرفی شده و در بخش ۴ روش پیشنهادی تفکیک عروق شبکیه به‌تفصیل تشریح می‌شود. نتایج تجربی و تحلیل عملکرد روش پیشنهادی در بخش ۵ ارائه شده و در نهایت، نتیجه‌گیری و مسیرهای پژوهشی آینده در بخش ۶ بیان می‌گردد.

۲ ملاحظات فیزیکی افت کیفیت تصاویر RF ناشی از CA

اثر CA با ایجاد کدورت در عدسی چشم، موجب پراکندگی نور ورودی و تغییر پاسخ نوری سیستم تصویربرداری چشم می‌شود. این پدیده به‌طور هم‌زمان باعث گسترش تابع پخش نقطه‌ای^۵،

مطالعات متعددی نشان داده‌اند که افت کیفیت تصویر ناشی از CA می‌تواند عملکرد سامانه‌های هوش مصنوعی در تشخیص بیماری‌های شبکیه را به‌طور چشمگیری کاهش دهد. به‌عنوان نمونه، در [۷] نشان داده شده که اعمال تاری شبیه‌سازی‌شده معادل CA بر تصاویر RF، دقت یک شبکه عمیق در تشخیص رنجوری دیابتی را از حدود ۹۷٪ در حالت تصویر شفاف به نزدیک ۵۳٪ در حالت تاری شدید کاهش می‌دهد [۶]. این نتایج نشان می‌دهد که در کاربردهای عملی بینایی‌ماشینی پزشکی^۱، نادیده گرفتن اثر CA می‌تواند منجر به ارزیابی‌های نادرست و کاهش اعتمادپذیری سامانه‌ها شود [۷، ۲].



شکل (۱): نمونه‌ای از تصاویر RF در سطوح مختلف شدت CA. (الف) تصویر RF فاقد CA، (ب) CA خفیف، (ج) CA متوسط، و (د) CA شدید [۱۴].

برای مقابله با این چالش، پژوهش‌های متعددی در زمینه بهبود کیفیت تصاویر RF مبتلا به CA یا افزایش استحکام الگوریتم‌ها در شرایط افت کیفیت انجام شده است. روش‌های کلاسیک مبتنی بر CLAHE^۲ در ترکیب با فیلترهای تقویتی، توانسته‌اند رویت‌پذیری عروق کم‌رنگ را تا حدی بهبود بخشند [۵]. علاوه بر این، روش‌های مبتنی بر یادگیری عمیق برای بازسازی یا بهبود تصاویر تاری RF نیز پیشنهاد شده‌اند. به عنوان مثال، در [۶] یک شبکه دومرحله‌ای مبتنی بر مکانیزم توجه چندمقیاسی^۳ معرفی شده است که بدون نیاز به تصاویر مرجع سالم متناظر، امکان بازیابی جزئیات تصاویر RF مبتلا به CA را فراهم می‌سازد. همچنین، در [۱۶] با بهره‌گیری از شبکه‌های یادگیری عمیق، مدلی برای تشخیص هم‌زمان چند بیماری چشمی در تصاویر RF دارای CA معرفی شد.

^۴ Field of View (FOV)
^۵ Point Spread Function

^۱ Medical Computer Vision
^۲ Contrast Limited Adaptive Histogram Equalization (CLAHE)
^۳ Multi-Scale Attention

۲-۳ افزودن نویز

اگرچه نویز ذاتاً ناشی از CA نیست، اما تصاویر واقعی ثبت‌شده در شرایط بالینی همواره حاوی درجاتی از نویز حسگر و اغتشاشات تصادفی هستند. مطالعات ارزیابی کیفیت تصاویر RF نشان داده‌اند که وجود نویز و ناهمگنی شدت، بخشی اجتناب‌ناپذیر از داده‌های واقعی بالینی است و باید در فرایند شبیه‌سازی مدنظر قرار گیرد [۵، ۲].

به‌منظور افزایش شباهت داده‌های شبیه‌سازی‌شده به تصاویر واقعی، نویز گاوسی دوبعدی با میانگین صفر به I_c افزوده می‌شود. لذا تصویر نویزدار I_n به‌صورت زیر تعریف می‌شود:

$$I_n(x, y) = I_c(x, y) + \eta(x, y), \quad \eta(x, y) \sim N(0, \sigma_n^2) \quad (3)$$

که در آن η نویز گاوسی افزوده‌شده، I_n تصویر نهایی پس از افزودن نویز، و σ_n^2 واریانس نویز است که میزان ناهمگنی شدت تصویر را کنترل می‌کند.

در این پژوهش مقدار σ_n به‌گونه‌ای انتخاب شده است که بدون تخریب ساختارهای اصلی تصویر، ناهمگنی‌های جزئی موجود در داده‌های واقعی را بازنمایی کند.

۲-۴ شبیه‌سازی پراکندگی نور^۳

یکی از تفاوت‌های اساسی تصاویر مبتلا به CA با تصاویر سالم، وجود هاله نوری ناشی از پراکندگی پیش‌رونده نور در عدسی است که باعث روشن‌شدن نواحی مجاور ساختارهای پرنور می‌شود [۱۲، ۱۸]. این پدیده از نظر فیزیکی با تاری اپتیکی ناشی از گسترش PSF متفاوت بوده و در مقیاسی گسترده‌تر رخ می‌دهد. برای شبیه‌سازی این اثر، ابتدا نسخه‌ای به‌شدت هموارشده از تصویر نویزدار I_n استخراج می‌شود. بدین منظور، تصویر هموارشده I_h از طریق کانولوشن دوبعدی I_n با G_{σ_s} ، که یک هسته گاوسی مدل‌کننده پراکندگی نور در عدسی کدر با انحراف معیار σ_s است، به صورت زیر بدست می‌آید:

$$I_h(x, y) = I_n(x, y) * G_{\sigma_s}(x, y), \quad \sigma_s \gg \sigma_b \quad (4)$$

در این رابطه، عملگر * کانولوشن دوبعدی را نشان می‌دهد و I_h تصویر هموارشده‌ای را نمایش می‌دهد که مؤلفه هاله نوری را بازنمایی می‌کند. شرط $\sigma_s \gg \sigma_b$ بیانگر آن است که مقیاس پراکندگی نور به‌طور معناداری بزرگ‌تر از مقیاس تاری اپتیکی بوده و این دو پدیده به‌صورت مجزا مدل‌سازی شده‌اند، به‌طوری که σ_b انحراف معیار هسته گاوسی تاری اپتیکی و مستقل از واریانس نویز در مرحله افزودن نویز است.

کاهش تمایز محلی، و ایجاد نوعی هاله یا مه‌نوری^۱ در تصویر ثبت‌شده از شبکیه می‌گردد [۷، ۱۷، ۱۸].

برخلاف مدل‌های ساده که افت کیفیت را صرفاً به‌صورت تاری یکنواخت در نظر می‌گیرند، مطالعات اپتیکی نشان داده‌اند که اثر CA ترکیبی از چند مؤلفه هم‌زمان است و مدل‌سازی واقع‌گرایانه آن مستلزم در نظر گرفتن این عوامل به‌صورت توأمان می‌باشد [۵، ۱۸]. بر این اساس، در این پژوهش یک چارچوب شبیه‌سازی چندمرحله‌ای برای بازتولید اثر CA در تصاویر RF ارائه می‌شود که شامل پنج مؤلفه اصلی تاری اپتیکی، تضعیف تمایز، نویز افزوده، پراکندگی نور و محدوده میدان دید است.

۲-۱ مدل‌سازی تاری اپتیکی^۲

اولین و بارزترین اثر CA، افزایش تاری تصویر به‌دلیل گسترش PSF عدسی است. این پدیده را می‌توان به‌صورت کانولوشن تصویر اصلی با یک هسته گاوسی مدل‌سازی کرد [۷، ۱۷]. اگر تصویر ورودی با وضوح کامل را در مختصات (x, y) با I نمایش دهیم، تصویر تار شده I_b به‌صورت زیر تعریف می‌شود:

$$I_b(x, y) = I(x, y) * G_{\sigma_b}(x, y) \quad (1)$$

که G_{σ_b} در آن هسته گاوسی تاری اپتیکی با انحراف معیار σ_b است. مقدار σ_b متناسب با شدت CA انتخاب می‌شود، به‌گونه‌ای که افزایش آن منجر به محوشدگی بیشتر لبه‌ها و ساختارهای ظریف عروقی گردد.

۲-۲ تضعیف تمایز تصویر

یکی از پیامدهای مستقیم پراکندگی نور در عدسی کدر، کاهش تمایز کلی و محلی تصویر شبکیه است؛ به‌طوری که اختلاف شدت بین ساختارهای عروقی و زمینه کاهش می‌یابد [۱۳، ۱۸]. اگر تصویر تار شده را در مختصات دوبعدی (x, y) با I_b نمایش دهیم و μ_b را به‌عنوان میانگین شدت I_b روی دامنه تصویر (یا ناحیه FOV) در نظر بگیریم، تصویر با تمایز کاهش‌یافته I_c به‌صورت زیر محاسبه می‌شود:

$$I_c(x, y) = \mu_b + c(I_b(x, y) - \mu_b) \quad (2)$$

که در آن $0 < c < 1$ ، ضریب تضعیف تمایز است. کاهش تمایز به‌عنوان یکی از اثرات فیزیکی اصلی CA در مطالعات ترمیم و تحلیل تصاویر RF گزارش شده است [۵، ۱۳]. بر این اساس، در این پژوهش مقدار c به‌صورت وابسته به شدت CA تنظیم شده و مقادیر آن در بازه [۰/۶۵، ۰/۸۵] انتخاب گردیده است تا روند تدریجی افت تمایز مشاهده‌شده در داده‌های بالینی بازنمایی شود.

این ماسک با اعمال عملگر اتساع^۱ ریکت شناختی بر مرز FOV و سپس هموارسازی گاوسی^۲ حاصل می شود تا گذار میان ناحیه شبکه و پس زمینه به صورت تدریجی انجام گیرد. لازم به ذکر است که هموارسازی در این مرحله با هدف مدل سازی تاری یا پراکندگی نور انجام نمی شود، بلکه صرفاً به منظور حذف ناپیوستگی های تیز در مرز FOV و جلوگیری از ایجاد هاله های مصنوعی در این ناحیه به کار می رود؛ در حالی که تاری اپتیکی و پراکندگی نور در مراحل پیشین به صورت مستقل مدل سازی شده اند.

۲-۶ تعریف سطوح شدت CA و سیاست انتخاب پارامترها

با توجه به ماهیت تدریجی و پیشرونده CA در شرایط بالینی، در این پژوهش سه سطح شدت شامل خفیف، متوسط و نسبتاً شدید تعریف شد. بر اساس مطالعات اپتیکی و شبیه سازی های پیشین، افزایش شدت CA معمولاً با تشدید هم زمان مؤلفه هایی مانند تاری اپتیکی، کاهش تمایز و پراکندگی نور همراه است [۵، ۷، ۱۳، ۱۸]. از این رو، مقادیر پارامترهای جدول (۱) به صورت افزایشی و در بازه های گزارش شده در مطالعات مرتبط انتخاب شدند تا امکان بررسی کنترل شده اثر افت کیفیت بر عملکرد الگوریتم های بخش بندی عروق فراهم شود. با این حال، هدف از انتخاب این مقادیر بازسازی کامل یا پیکسل به پیکسل تصاویر واقعی مبتلا به CA نبوده است، بلکه ایجاد سه سطح قابل تکرار از افت کیفیت تصویر بوده است. به همین دلیل، تنظیم نهایی پارامترها بر اساس کالیبراسیون کیفی و با توجه به میزان رؤیت پذیری ساختارهای عروقی انجام شد؛ به گونه ای که با وجود افت محسوس کیفیت تصویر RF، ساختارهای اصلی شبکه، به ویژه عروق بزرگ و متوسط، حتی در سطح نسبتاً شدید همچنان قابل مشاهده باقی بمانند.

نمونه هایی از خروجی شبیه سازی در شکل (۲) ارائه شده اند تا روند کاهش تدریجی وضوح، کاهش تمایز، محوشدگی لبه ها و یکنواخت شدن بافت شبکه در سطوح مختلف CA نشان داده شود. این ارزیابی ماهیتی کیفی دارد و صرفاً برای کنترل معنادار بودن شدت های تعریف شده از نظر حفظ اطلاعات ساختاری استفاده شده است. شدت هایی که باعث از بین رفتن کامل ساختارهای شبکه می شوند در این پژوهش در نظر گرفته نشده اند؛ زیرا در چنین شرایطی تحلیل و ارزیابی کمی تفکیک عروق از نظر عملی معنا نخواهد داشت.

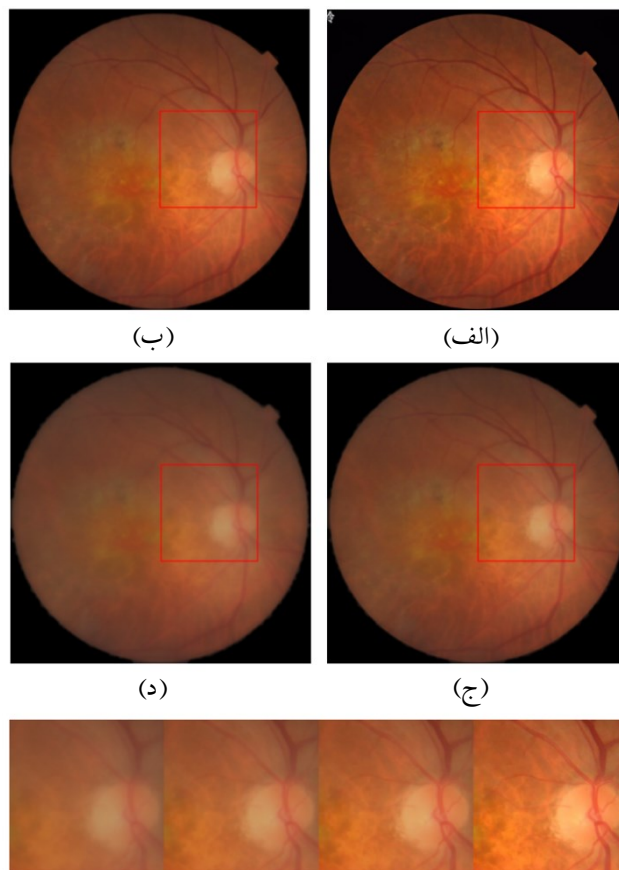
پس از آن، تصویر نهایی حاصل از پراکندگی نور I_s به صورت ترکیب خطی تصویر نویزدار I_n و تصویر هموار شده I_h محاسبه می شود:

$$I_s(x, y) = (1 - s)I_n(x, y) + sI_h(x, y) \quad (5)$$

که در آن s ضریب شدت پراکندگی و I_s تصویر حاصل از اعمال مراحل تاری، کاهش تمایز، نویز و پراکندگی نور است. افزایش مقدار s با افزایش شدت CA هم راستا با مدل های اپتیکی پراکندگی نور در عدسی کدر گزارش شده در [۱۸] است.

۲-۵ اعمال افت کیفیت محدود به ناحیه میدان دید

برای اعمال افت کیفیت درون FOV، ابتدا از ماسک دودویی FOV یک ماسک نرم دوبعدی $f(x, y)$ تولید می گردد.



(ه) تصویر بزرگ شده ROI (الف، ب، ج، د)

شکل (۲): نمونه ای از تصاویر RF در سطوح مختلف شدت CA شبیه سازی شده، (الف) فاقد CA، (ب) خفیف، (ج) متوسط، و (د) نسبتاً شدید که کاهش تدریجی وضوح، تمایز و رؤیت پذیری ساختارهای شبکه را نشان می دهند. کادر قرمز ناحیه مورد بررسی (ROI) را مشخص می کند، (ه) بزرگ نمایی نواحی ROI که محوشدگی جزئیات و یکنواخت شدن بافت تصویر در شدت های بالای CA را نشان می دهد.

جدول (۱): پارامترهای شبیه‌سازی CA در شدتهای مختلف

شدت	σ_B	c	σ_H	s
خفیف	۲	۰/۸۵	۴	۰/۲۵
متوسط	۳	۰/۷۵	۶	۰/۳۵
نسبتاً شدید	۵	۰/۶۵	۱۰	۰/۵۰

۳ پایگاه‌داده‌ها و روند آماده‌سازی داده‌ها

در این پژوهش، به‌منظور آموزش و ارزیابی الگوریتم‌های بخش‌بندی عروق شبکیه در شرایط مختلف کیفیت تصویر، از سه پایگاه‌داده استاندارد FIVES [۱۹]، DRIVE [۴] و RFMiD [۳۴] استفاده شده است.

ارزیابی کمی اصلی در این پژوهش بر پایه پایگاه‌داده‌های FIVES و DRIVE انجام شده است که شامل تصاویر واقعی RF همراه برجسب مرجع^۱ (GT) عروق شبکیه هستند، اما شامل تصاویر واقعی مبتلا به CA نیستند. اگرچه پایگاه‌داده‌هایی همچون RFMiD شامل تصاویر RF واقعی مبتلا به CA در مطالعات تشخیص بیماری یا طبقه‌بندی شدت CA مورد استفاده قرار گرفته‌اند، اما این داده‌ها معمولاً فاقد پیکسلی برای عروق شبکیه هستند. در مسئله تفکیک عروق، صرف وجود تصویر واقعی کافی نیست؛ زیرا ارزیابی کمی تفکیک عروق نیازمند وجود GT پیکسلی عروق است.

به همین دلیل، برای ایجاد شرایط کنترل‌شده CA، افت کیفیت ناشی از آب‌مروارید روی پایگاه‌داده‌های FIVES و DRIVE دارای GT شبیه‌سازی شد تا برجسب مرجع عروق حفظ شود و اثر شدت تخریب به‌صورت منظم بررسی گردد. افزون بر این، برای نزدیک‌تر شدن تحلیل به شرایط واقعی، تصاویر واقعی مبتلا به CA و فاقد GT پیکسلی از پایگاه‌داده عمومی RFMiD به‌صورت کیفی ارزیابی شده‌اند [۳۴].

۳-۱ معرفی پایگاه‌داده‌های مورد استفاده

پایگاه‌داده FIVES^۲ یک مجموعه‌داده مرجع و باکیفیت برای تفکیک عروق شبکیه است که شامل ۸۰۰ تصویر رنگی RF با وضوح بالا می‌باشد. تمامی تصاویر این پایگاه‌داده دارای GT عروق شبکیه هستند که فرآیند برجسب‌گذاری آن با مشارکت چندین برجسب‌گذار^۳ و بازبینی متخصصان انجام شده است. این پایگاه‌داده به‌صورت رسمی به ۶۰۰ تصویر آموزشی و ۲۰۰ تصویر آزمون تقسیم شده و به‌دلیل کیفیت بالای تصاویر و دقت GT، در این پژوهش به‌عنوان مرجع اصلی آموزش و ارزیابی کمی الگوریتم‌های بخش‌بندی عروق شبکیه مورد استفاده قرار گرفته است [۱۹].

در مقابل، پایگاه‌داده DRIVE^۴ یکی از مجموعه‌داده‌های کلاسیک و پرکاربرد در حوزه بخش‌بندی عروق شبکیه است که شامل ۴۰ تصویر رنگی RF می‌باشد. در این پایگاه‌داده، GT عروق شبکیه صرفاً برای تصاویر بخش آموزشی (۲۰ تصویر) در دسترس بوده و برای تصاویر بخش آزمون، GT عروق به‌صورت عمومی منتشر نشده است. از این‌رو، در این پژوهش ارزیابی کمی عملکرد مدل‌ها بر روی دیتاست DRIVE صرفاً بر اساس مجموعه آموزشی انجام شده است؛ رویکردی که امکان محاسبه معیارهای کمی نظیر Dice^۵ و IoU^۶ را فراهم می‌کند و با رویه رایج در ادبیات این حوزه هم‌خوانی دارد [۴]. مطابق با پیاده‌سازی‌های انجام‌شده، پایگاه‌داده DRIVE در این پژوهش صرفاً در نقش ارزیابی کمی و سنجش پایداری مدل‌ها مورد استفاده قرار گرفته است.

پایگاه‌داده RFMiD یک مجموعه‌داده عمومی از تصاویر RF است که با هدف بررسی بیماری‌ها و ناهنجاری‌های مختلف شبکیه در شرایط بالینی متنوع ارائه شده است [۳۴]. با توجه به گستردگی این پایگاه‌داده و ثبت تصاویر آن با سه دوربین فوندوس مختلف، تصاویر موجود در RFMiD دارای سه وضوح متفاوت هستند. در این پژوهش، از تصاویر دارای برجسب CA در این مجموعه، به‌عنوان نمونه‌هایی از تصاویر واقعی مبتلا به CA، برای ارزیابی کیفی رفتار مدل منتخب در شرایط نزدیک‌تر به داده‌های بالینی استفاده شده است. مشخصات پایگاه‌داده‌های FIVES، DRIVE و RFMiD و نقش هر یک در این پژوهش در جدول (۲) ارائه شده است.

جدول (۲): مشخصات آماری پایگاه‌داده‌های مورد استفاده در پژوهش

ویژگی	FIVES[۱۹]	DRIVE[۴]	RFMiD[۳۴]
تعداد کل تصاویر	۸۰۰	۴۰	۳۲۰۰
وضوح تصویر	۲۰۴۸×۲۰۴۸	۵۸۴×۵۶۵	۲۱۴۴×۱۴۲۴، ۴۲۸۸×۲۸۴۸، ۲۰۴۸×۱۵۳۶
تصاویر آموزش	۶۰۰	۲۰	-
تصاویر اعتبارسنجی	-	-	-
تصاویر آزمون	۲۰۰	۲۰	-
وجود GT عروق	✓	✓ (تصاویر آموزش)	ندارد
نقش در پژوهش	آموزش و ارزیابی کمی	ارزیابی کمی	ارزیابی کیفی

۳-۲ ساخت نسخه‌های شبیه‌سازی‌شده FIVES-CA و DRIVE-CA

بر اساس مدل افت کیفیت و مراحل شبیه‌سازی تشریح‌شده در بخش ۲، در این پژوهش دو مجموعه مشتق‌شده با نام‌های FIVES-CA و DRIVE-CA از پایگاه‌داده‌های اصلی FIVES و

^۴ Digital Retinal Images for Vessel Extraction^۵ Dice coefficient^۶ Intersection over Union^۱ Ground Truth^۲ Fundus Dataset for Artificial Intelligence-based Vessel Segmentation^۳ Annotator

بر این اساس، ابتدا نسخه‌های شبیه‌سازی‌شده پایگاه‌داده FIVES برای آموزش و انتخاب مدل‌های برتر ایجاد شده‌اند. سپس، به‌منظور ارزیابی مکمل و بررسی تعمیم‌پذیری، نسخه‌های شبیه‌سازی‌شده متناظر برای پایگاه‌داده DRIVE نیز تولید گردیده‌اند.

روند کلی انتخاب پایگاه‌داده‌ها و استفاده از نسخه‌های شبیه‌سازی شده مبتلا به CA در این پژوهش در شکل (۳) ارائه شده است. استفاده از نسخه‌های شبیه‌سازی‌شده، امکان انجام مقایسه‌های منصفانه و قابل تکرار را در سناریوهای دارای افت کیفیت تصویر فراهم می‌سازد.

۴ روش پیشنهادی تفکیک پایدار عروق شبکیه در حضور CA

در این بخش، چارچوب پیشنهادی پژوهش حاضر برای تفکیک عروق شبکیه از تصاویر RF در حضور افت کیفیت ناشی از CA ارائه می‌شود و تمرکز صرفاً بر جریان کامل روش، از تولید داده تا انتخاب مدل و ارزیابی پایداری خواهد بود.

ایده اصلی روش پیشنهادی بر این مبنا استوار است که بخش‌بندی عروق شبکیه در تصاویر RF به‌شدت به کیفیت تصویر حساس بوده و در حضور تاری، افت تمایز و پراکندگی نور، عملکرد مدل‌های یادگیری عمیق نیز دچار افت می‌شود. ازاین‌رو، به‌جای اتکا به تصاویر صرفاً شفاف، نسخه‌های شبیه‌سازی‌شده کنترل‌شده از داده‌ها در چند سطح شدت CA ساخته شده و سپس پایداری مدل‌های مختلف بخش‌بندی در این شرایط ارزیابی می‌گردد.

۴-۱ نمای کلی چارچوب پیشنهادی

نمای کلی چارچوب پیشنهادی شامل چهار گام اصلی زیر است: (الف) ساخت داده‌های CA: مطابق معادلات (۱) - (۶)، ملاحظات فیزیکی افت کیفیت ناشی از CA بر روی تصاویر RF اعمال و تخریب تصویر تنها در ناحیه FOV اعمال می‌گردد.

(ب) ارزیابی مدل‌های معیار روی پایگاه‌داده مرجع FIVES:

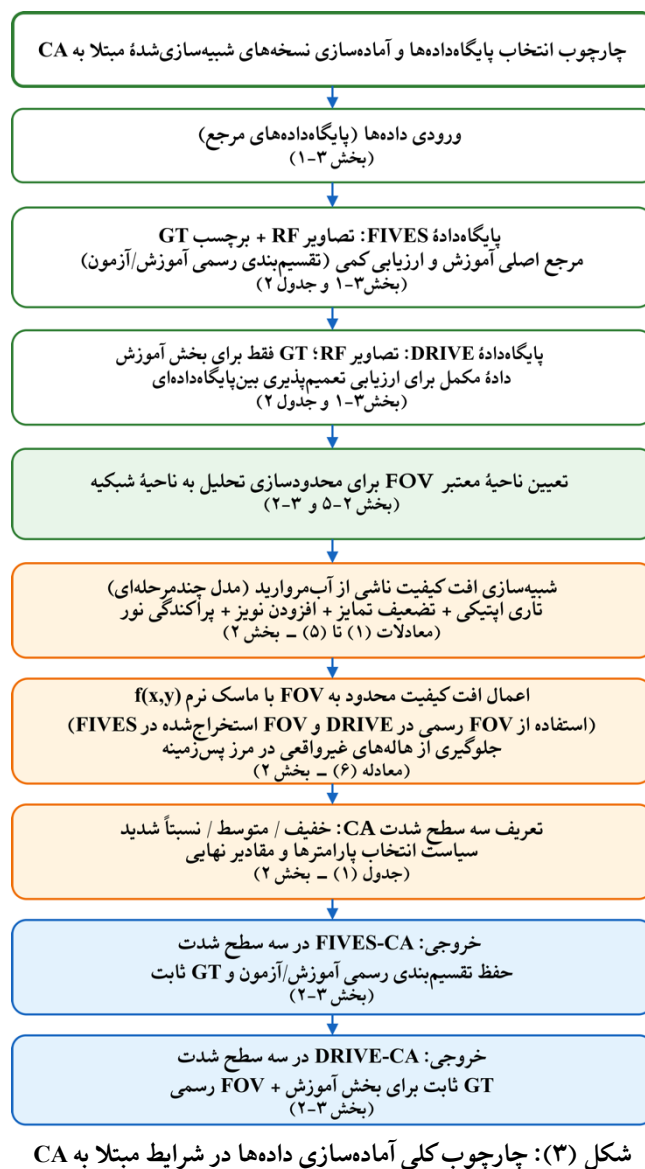
به‌منظور انتخاب الگوریتم پایه مناسب، چند معماری بخش‌بندی روی FIVES آموزش و ارزیابی شدند و مدل‌های برتر روی مجموعه آزمون انتخاب شدند. این انتخاب بر پایه این واقعیت است که FIVES به‌عنوان یک پایگاه‌داده با وضوح بالا و GT پیکسلی دقیق، مرجع مناسبی برای مقایسه کمی مدل‌ها فراهم می‌کند [۱۹]، و مشتقات آن نیز به‌عنوان چارچوب‌های رایج و مؤثر در بخش‌بندی زیست‌پزشکی شناخته می‌شوند [۹].

(ج) آموزش و ارزیابی مدل‌های منتخب روی FIVES-CA: سه مدل برتر در سه سطح شدت CA آموزش و ارزیابی می‌شوند تا

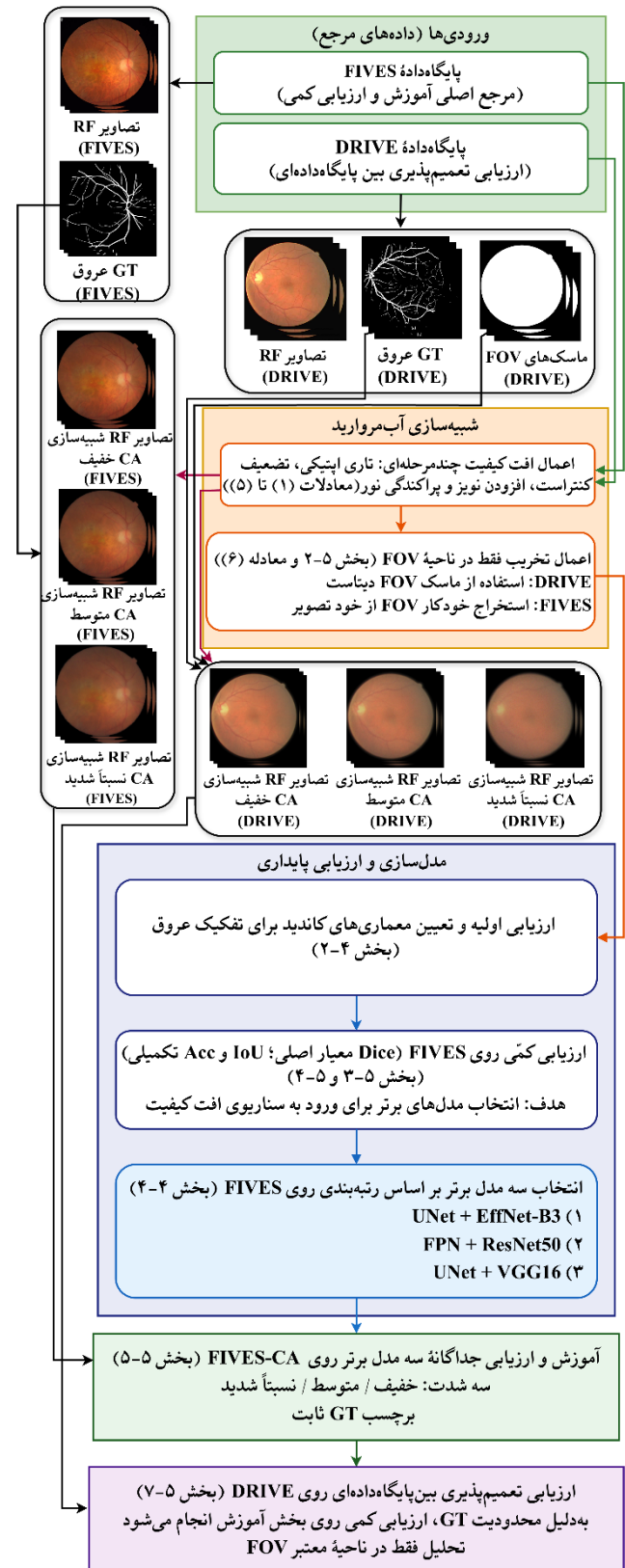
DRIVE ساخته شدند. این نام‌گذاری در ادامه مقاله برای تفکیک نسخه‌های اصلی از نسخه‌های شبیه‌سازی‌شده دارای افت کیفیت ناشی از CA به‌کار می‌رود.

برای پایگاه‌داده FIVES، نسخه‌های شبیه‌سازی‌شده در سه سطح شدت CA ایجاد شده و تقسیم‌بندی رسمی داده‌ها برای آموزش و آزمون بدون تغییر حفظ گردیده و مجموعه داده FIVES-CA برای پیش‌آموزش شبکه تشکیل گردیده است. از آنجا که این پایگاه‌داده فاقد ماسک صریح FOV است، ناحیه مؤثر شبکه برای هر تصویر به‌صورت خودکار تعیین شده و شبیه‌سازی تنها در این ناحیه اعمال گردیده است.

در مورد پایگاه‌داده DRIVE نیز نسخه‌های شبیه‌سازی‌شده متناظر تولید شده‌اند. نسخه‌های شبیه‌سازی‌شده DRIVE امکان بررسی رفتار الگوریتم در شرایط افت کیفیت کنترل‌شده را بدون تغییر در ساختار داده یا برجسب‌ها فراهم می‌کنند.



روند افت عملکرد با افزایش شدت تخریب به‌صورت کنترل‌شده بررسی شود.



شکل (۴): نمای کلی روش پیشنهادی تفکیک عروق شبکیه در حضور CA

(د) ارزیابی تعمیم‌پذیری بین پایگاه‌داده‌ای روی DRIVE-CA: به‌منظور بررسی رفتار مدل خارج از پایگاه داده مرجع، عملکرد همان مدل‌ها به‌صورت هم‌زمان روی DRIVE-CA نیز گزارش شدند. مطابق محدودیت ساختاری DRIVE، ارزیابی کمی روی قسمت آموزش این پایگاه‌داده انجام می‌گیرد تا امکان محاسبه معیارهای کمی با GT فراهم باشد [۴]. این ارزیابی مکمل، نشان می‌دهد آیا مدلی که روی FIVES-CA پایدارتر است، در یک پایگاه داده مستقل نیز رفتار سازگار نشان می‌دهد یا خیر. نمای شماتیک این فرآیند در شکل (۴) نمایش داده شده است.

۲-۴ معماری‌های مرجع مورد بررسی برای تفکیک عروق شبکیه

با توجه به ماهیت عروق شبکیه به‌عنوان ساختارهایی باریک، پیوسته و چندمقیاسی و نیز با در نظر گرفتن افت کیفیت ناشی از تاری، کاهش تمایز و پراکندگی نور در حضور CA، مجموعه‌ای متنوع از معماری‌های شاخص تفکیک تصویر پزشکی و تفکیک عروق شبکیه به‌عنوان کاندیدهای اولیه بررسی شدند. این مجموعه شامل خانواده‌های رمزگذار-رمزگشا، Residual، چندمقیاسی، زمینه محور و دارای رمزگذار از پیش آموزش‌دیده بود. در نهایت، معماری‌های لیست شده در جدول (۳)، به دلایل و رویکردهای رایج در تصاویر پزشکی و همچنین پردازش تصاویر RF انتخاب شدند. خانواده UNet و نسخه‌های توسعه‌یافته آن به‌دلیل کاربرد گسترده در تفکیک ساختارهای زیست‌پزشکی و عروق شبکیه، به‌عنوان مبنای اصلی مقایسه در نظر گرفته شدند [۲۶، ۱۱]. معماری‌های مبتنی بر اتصال‌های Residual، مدل‌های ترکیبی و روش‌های مبتنی بر رمزگذارهای از پیش آموزش‌دیده نیز به‌دلیل کاربرد آن‌ها در استخراج ویژگی‌های عمیق‌تر و پایدارتر از تصاویر RF وارد ارزیابی شدند [۲۷-۳۰]. علاوه بر این، مدل‌های زمینه‌محور و چندمقیاسی به دلیل قابلیت اعمال اثر اطلاعات بافتی و ترکیب ویژگی‌ها در مقیاس‌های مختلف، به‌ویژه در تفکیک عروق باریک و ضخیم، مورد بررسی قرار گرفتند [۳۰، ۳۲]. در نتیجه، جدول (۳) مجموعه‌ای متنوع از معماری‌های مرجع موفق را برای مقایسه تجربی تحت یک پروتکل آموزشی و ارزیابی یکسان فراهم می‌کند. جزئیات معماری‌های پیشنهادی به شرح ذیل می‌باشد:

(الف) معماری‌های مبتنی بر شبکه‌های UNet:

معماری‌های مبتنی بر رمزگذار-رمزگشا با اتصال‌های میان‌بُرا، به‌طور ویژه برای تفکیک زیست‌پزشکی طراحی شده‌اند و به دلیل حفظ اطلاعات مکانی در کنار ویژگی‌های عمیق، برای استخراج

جدول (۳): کاربرد معماری‌های مرجع بررسی شده در تفکیک عروق RF

نقش در پژوهش	روش
مدل پایه تفکیک زیست‌پزشکی و مبنای مقایسه؛ پرکاربرد در تفکیک عروق RF [۱۱].	UNet [۹]
بررسی اثر اتصال‌های تودرتو؛ استفاده شده در تفکیک عروق و پیش‌بینی بیماری از RF [۲۶].	UNet++ [۲۰]
بررسی اثر بلوک‌های Residual خانواده ResUNet در تفکیک عروق RF [۲۸, ۲۷].	ResUNet34 [۲۱]
بررسی اثر رمزگذار کلاسیک UNet-VGG در تفکیک عروق RF [۳۰, ۲۹].	UNet+VGG16 [۲۴]
بررسی اثر رمزگذار عمیق Residual خانواده ResUNet در تفکیک عروق RF [۳۰, ۲۹].	UNet+ResNet50 [۲۱]
بررسی اثر میدان دریافت بزرگ‌تر DeepLab در تفکیک عروق فوندوس [۳۱].	DeepLabV3+ResNet50 [۲۳]
بررسی اثر تجمیع چندمقیاسی FPN+ResNet در ارزیابی تفکیک عروق RF [۳۰].	FPN+ResNet50 [۲۲]
بررسی رمزگذار کارا و مقیاس‌پذیر UNet+EffNet در تفکیک عروق RF [۳۲].	UNet+EffNet-B3 [۲۵]

۳-۴ نحوه بررسی اولیه معماری‌ها و انتخاب سه مدل برتر برای تفکیک عروق شبکه

به‌منظور تعیین مدل‌های پایه مناسب برای بررسی اثر افت کیفیت ناشی از CA، تمامی معماری‌های کاندید پس از آموزش بر روی پایگاه‌داده مرجع FIVES، به‌صورت مدل‌های آموزش‌دیده مورد ارزیابی قرار گرفتند و عملکرد آن‌ها بر روی مجموعه آزمون این پایگاه‌داده سنجیده شد. هدف این مرحله، انتخاب مدل‌هایی بود که در ارزیابی اولیه، عملکرد مناسب‌تری در تفکیک ساختارهای عروقی نشان دهند و بتوانند مبنای قابل اتکاتری برای تحلیل رفتار مدل‌ها در شرایط افت کیفیت تدریجی تصویر باشند. لذا پس از انجام ارزیابی اولیه مدل‌ها، مدل‌های آموزش‌دیده در شرایط آزمون یکسان مورد تحلیل و به ترتیب عملکرد، مرتب می‌شوند. نهایتاً، سه مدل دارای بهترین عملکرد کلی به‌عنوان مدل‌های برتر انتخاب شدند تا در مراحل بعدی پژوهش، برای بررسی رفتار در حضور CA شبیه‌سازی شده مورد آموزش و ارزیابی قرار گیرند. نتایج تفصیلی‌تر مربوط به این مدل‌ها و مقایسه عملکرد آن‌ها در سناریوهای مختلف افت کیفیت، در بخش ۵ ارائه شده است.

لازم به ذکر است که اگرچه در بخش ۴-۲ مزایای ساختاری برخی معماری‌های تفکیک عروق از منظر تئوریک مورد بحث قرار گرفته است، انتخاب مدل‌های نهایی در این مرحله به‌صورت داده‌محور و بر اساس عملکرد تجربی آن‌ها بر روی پایگاه‌داده مرجع انجام شده است. بدین ترتیب، عدم قرارگیری برخی معماری‌های دارای مزایای نظری در میان سه مدل برتر، بیانگر تفاوت میان انتظار تئوریک و رفتار عملی مدل‌ها در شرایط واقعی داده است؛

ساختارهای ظریف مناسب‌اند [۹]. در این پژوهش، علاوه بر UNet پایه، نسخه‌های UNet با رمزگذارهای استاندارد و نیز نسخه توسعه‌یافته UNet++ نیز بررسی شده‌اند [۲۰].

(ب) معماری‌های مبتنی بر شبکه‌های Residual:

شبکه‌های Residual با تسهیل عبور گرادیان و یادگیری پایدارتر در عمق‌های بالاتر، به‌عنوان رمزگذارهای قدرتمند و پایدار در بسیاری از مسائل تفکیک به‌کار گرفته شده‌اند [۲۱]. از این منظر، استفاده از رمزگذارهایی نظیر ResNet می‌تواند در شرایط افت کیفیت، به پایداری استخراج ویژگی کمک کند.

(ج) معماری‌های چندمقیاسی مبتنی بر هرم ویژگی:

از آنجا که عروق شبکه شامل طیفی از قطر‌ها و الگوهای محلی/سراسری است، ترکیب ویژگی‌ها در مقیاس‌های مختلف می‌تواند به بهبود تشخیص هم‌زمان عروق نازک و ضخیم منجر شود. بر همین اساس، معماری شبکه عصبی هرم ویژگی (FPN) به‌عنوان یکی از چارچوب‌های شناخته‌شده برای تجمیع چندمقیاسی انتخاب شده است [۲۲].

(د) معماری‌های مبتنی بر زمینه با میدان دریافت مؤثر بزرگ‌تر:

روش‌هایی مانند DeepLabV3+ با بهره‌گیری از کانولوشن‌های حفره‌دار^۴ و طراحی رمزگذار-رمزگشا برای استخراج زمینه چندمقیاسی، معمولاً در تفکیک‌هایی که هم به ساختارهای ظریف و هم به زمینه نیاز دارند مورد توجه‌اند [۲۳]. این ویژگی در حضور تاری و افت تمایز می‌تواند اهمیت بیشتری پیدا کند.

(ه) رمزگذارهای کلاسیک و مدرن از پیش آموزش‌دیده:

برای بررسی اثر کیفیت رمزگذار در پایداری تفکیک، از رمزگذار کلاسیک VGG به‌عنوان مبنای ساده‌تر و قابل‌تفسیرتر [۲۴] و نیز رمزگذار EffNet^۵ به‌عنوان خانواده‌ای با مقیاس‌بندی اصولی و کارایی بالا [۲۵] استفاده شده است.

بر این اساس، مجموعه‌ای از مدل‌های کاندید برای پوشش رویکردهای مختلف تفکیک عروق شبکه انتخاب شد. فهرست مدل‌های بررسی‌شده و نقش هر یک در جدول (۳) ارائه شده است.

^۱ Feature Pyramid^۲ Feature Pyramid Neural Network^۳ Context-aware architectures with larger effective receptive fields^۴ Dilated ConvolutionsEfficientNet^۵

۵ ارزیابی و نتایج

در این بخش، عملکرد مدل‌های منتخب از چند منظر بررسی می‌شود. ابتدا نتایج کمی روی پایگاه داده مرجع FIVES و نسخه‌های شبیه‌سازی شده FIVES-CA گزارش می‌شود. سپس، اثر مؤلفه‌های شبیه‌سازی CA از طریق مطالعه فرسایشی تحلیل شده و تعمیم‌پذیری مدل‌ها روی DRIVE-CA ارزیابی می‌گردد. در ادامه، برای تکمیل تحلیل عملی، نتایج کیفی روی تصاویر واقعی فاقد GT از RFMiD و همچنین هزینه محاسباتی مدل‌های منتخب بررسی می‌شود.

۵-۱ هدف ارزیابی و سناریوی آزمون

هدف این بخش، ارزیابی پایداری^۱ الگوریتم‌های تفکیک عروق شبکیه از تصاویر RF در حضور افت کیفیت ناشی از CA و همچنین بررسی تعمیم‌پذیری بین پایگاه داده‌ای آن‌ها است. مطالعات پیشین نشان داده‌اند که این نوع افت کیفیت می‌تواند عملکرد سامانه‌های یادگیری عمیق را به‌صورت شدید و ناگهانی کاهش دهد [۷].

۵-۲ پروتکل ارزیابی و پیش‌پردازش

در فرآیند ارزیابی، هر شبکه یک نقشه احتمال پیکسل به پیکسل برای کلاس عروق تولید می‌کند. سپس با اعمال آستانه‌گذاری ثابت (آستانه ۰/۵)، ماسک دودویی نهایی استخراج می‌شود. برای یکسان‌سازی شرایط مقایسه، تمام مدل‌ها تحت یک پروتکل آموزشی ثابت آموزش داده شدند. تصاویر ورودی در مرحله آموزش به قطعات 512×512 برش داده شدند و افزایش داده شامل چرخش‌های ۹۰ درجه و قرینه‌سازی افقی و عمودی به‌کار رفت. برای آموزش، از تابع هزینه ترکیبی شامل Binary Cross-Entropy و Dice Loss با وزن یکسان استفاده شد. بهینه‌سازی با AdamW و نرخ یادگیری 3×10^{-4} انجام گرفت و مدل‌ها برای ۲۵ دوره آموزش داده شدند.

به‌منظور جلوگیری از تأثیر نواحی غیرمعتبر، محاسبه معیارهای کمی صرفاً درون ناحیه معتبر انجام شده است. در پایگاه داده DRIVE-CA، ماسک رسمی FOV برای محدودسازی ناحیه ارزیابی به کار گرفته شده است. در پایگاه داده FIVES و نسخه‌های FIVES-CA، از آنجا که ماسک رسمی FOV ارائه نشده است، ناحیه مؤثر شبکه برای هر تصویر به‌صورت خودکار استخراج شده و در محاسبه معیارها لحاظ گردیده است. علاوه بر این، به‌منظور سازگاری ابعادی با معماری‌های رمزگذار و جلوگیری از خطاهای ابعادی، تصویر و برچسب‌ها به نزدیک‌ترین مضرب ۳۲

موضوعی که خود بر اهمیت ارزیابی تجربی در مسائل تفکیک زیست‌پزشکی تأکید می‌کند. در ادامه، سه مدل منتخب که هر یک مصداق عملی یکی از راهبردهای معماری معرفی شده در بخش ۲-۴ هستند، به اختصار معرفی می‌شوند.

(الف) مدل UNet با رمزگذار EffNet-B3:

این مدل به‌عنوان نماینده ترکیب چارچوب تفکیک زیست‌پزشکی با یک رمزگذار مدرن و کارا انتخاب شد. با رویکرد مقیاس‌بندی هم‌زمان عمق، عرض و وضوح، به‌عنوان خانواده‌ای با توازن مناسب میان پیچیدگی و دقت شناخته می‌شود [۲۵]. از منظر تفکیک عروق، چنین رمزگذاری می‌تواند در استخراج الگوهای محلی کم‌تمایز (به‌ویژه در عروق نازک‌تر) نقش مؤثری داشته باشد، در حالی که ساختار UNet سازوکار حفظ جزئیات مکانی را فراهم می‌کند [۹].

(ب) مدل FPN با رمزگذار ResNet50:

این مدل به‌عنوان نماینده صریح ادغام چندمقیاسی ویژگی‌ها انتخاب شد. FPN به‌طور مستقیم با هدف ترکیب سلسله مراتبی ویژگی‌ها در چند مقیاس طراحی شده و برای مسائلی که در آن‌ها ساختار هدف در اندازه‌های متفاوت ظاهر می‌شود مناسب است [۲۲]. در مسئله عروق شبکیه، وجود هم‌زمان عروق بزرگ و شاخه‌های باریک، یک انگیزه روشن برای استفاده از چنین ساختاری است. همچنین به‌کارگیری رمزگذار Residual، امکان یادگیری پایداری و استخراج ویژگی‌های قوی‌تر را فراهم می‌کند [۲۱].

(ج) مدل UNet با رمزگذار VGG16:

این مدل به‌عنوان نماینده یک ترکیب کلاسیک و قابل مقایسه انتخاب شد. VGG به‌عنوان یک رمزگذار استاندارد و خوش‌ساخت در ادبیات شبکه‌های کانولوشنی شناخته می‌شود [۲۴]. حضور این مدل در سه مدل برتر این امکان را فراهم می‌کند که در فصل ۵، تفاوت پایداری میان رمزگذارهای مدرن‌تر و یک رمزگذار کلاسیک‌تر به‌صورت منصفانه تحلیل شود، بدون آنکه مقایسه صرفاً به معماری‌های پیچیده‌تر محدود گردد. علاوه بر این، سادگی نسبی چنین رمزگذاری معمولاً به رفتار قابل پیش‌بینی‌تر در شرایط تخریب منجر می‌شود که در بحث پایداری اهمیت دارد.

بر این اساس، سه مدل منتخب شامل UNet+EffNet-B3، UNet+VGG16 و FPN+ResNet50 به‌عنوان نمایندگان عملی سه راهبرد متفاوت معماری انتخاب شدند و در بخش ارزیابی، رفتار آن‌ها در سناریوهای مختلف افت کیفیت ناشی از CA بررسی می‌شود.

جدول (۴): نتایج ارزیابی اولیه تفکیک عروق روی کل تصاویر پایگاه داده FIVES

مدل	Dice	IoU	Acc
UNet	۰/۸۳۰۸	۰/۷۳۷۴	۰/۹۸۳۲
UNet + ResNet50	۰/۸۵۵۷	۰/۷۷۱۹	۰/۹۸۵۷
UNet++ + ResNet34	۰/۸۶۱۲	۰/۷۷۶۵	۰/۹۸۵۶
DeepLabV3+ + ResNet50	۰/۸۹۵۸	۰/۸۲۰۵	۰/۹۸۷۷
ResUNet34	۰/۸۹۸۳	۰/۸۲۴۹	۰/۹۸۸۱
UNet + VGG16	۰/۸۹۸۶	۰/۸۲۷۱	۰/۹۸۸۳
FPN + ResNet50	۰/۹۰۱۷	۰/۸۳۰۳	۰/۹۸۸۷
UNet + EffNet-B3	۰/۹۱۰۷	۰/۸۴۵۱	۰/۹۸۹۸

۵-۵ نتایج کمی روی پایگاه داده FIVES-CA (پایداری در حضور CA)

نتایج کمی سه مدل برتر انتخاب شده روی نسخه های شبیه سازی شده پایگاه داده FIVES در سه سطح شدت CA در جدول (۵) گزارش شده است.

مطابق جدول (۵)، با افزایش شدت CA از خفیف به متوسط و سپس نسبتاً شدید، کاهش تدریجی و یکنواختی در Dice و IoU مشاهده می شود که بیانگر دشوارتر شدن بازیابی عروق باریک در حضور تاری اپتیکی، افت تمایز و پراکندگی نور است. با این حال، حتی در شدیدترین سطح CA، مقادیر Dice همچنان در بازه های معنادار باقی مانده اند که نشان دهنده پایداری نسبی چارچوب انتخاب شده در شرایط افت کیفیت کنترل شده است.

جدول (۵): نتایج تفکیک عروق روی پایگاه داده FIVES-CA در شدت های مختلف CA

شدت CA	مدل	Dice	IoU	Acc
خفیف	FPN+ResNet50	۰/۸۷۹۶	۰/۷۹۴۲	۰/۹۸۵۶
	UNet + VGG16	۰/۸۹۱۸	۰/۸۱۳۶	۰/۹۸۷۱
	UNet+EffNet-B3	۰/۸۹۵۴	۰/۸۱۸۵	۰/۹۸۷۵
متوسط	FPN+ResNet50	۰/۸۶۳۴	۰/۷۶۸۰	۰/۹۸۳۲
	UNet+VGG16	۰/۸۶۹۴	۰/۷۷۸۳	۰/۹۸۴۴
	UNet+EffNet-B3	۰/۸۷۵۲	۰/۷۸۵۱	۰/۹۸۴۶
نسبتاً شدید	UNet+VGG16	۰/۷۹۹۴	۰/۶۷۴۷	۰/۹۷۵۲
	FPN+ResNet50	۰/۸۰۱۳	۰/۶۷۶۰	۰/۹۷۴۸
	UNet+EffNet-B3	۰/۸۱۰۶	۰/۶۸۸۸	۰/۹۷۵۷

هم اندازه سازی^۱ شده اند و ارزیابی کمی صرفاً بر روی ناحیه غیر هم اندازه شده^۲ انجام گرفته است. این راهبرد موجب می شود معیارهای گزارش شده تنها به نواحی دارای اطلاعات بالینی معنادار وابسته بوده و تحت تأثیر پاسخ های کاذب ناشی از نواحی هم اندازه سازی شده یا خارج از فوندوس قرار نگیرند.

۳-۵ معیارهای کمی ارزیابی

برای سنجش عملکرد تفکیک عروق، از معیارهای استاندارد Acc^۳، Dice و IoU استفاده شده است. در تمامی تعاریف، TP، TN، FN و FP به ترتیب نشان دهنده مثبت درست، منفی درست، مثبت غلط و منفی غلط هستند:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (۷)$$

$$Dice = \frac{2TP}{2TP + FP + FN} \quad (۸)$$

$$IoU = \frac{TP}{TP + FP + FN} \quad (۹)$$

معیارهای Dice و IoU نیز برای سنجش میزان هم پوشانی میان ماسک پیش بینی شده و GT به کار گرفته می شوند [۴، ۹]. همچنین، Dice در تفکیک پزشکی به عنوان معیاری سازگار با ارزیابی نواحی هدف کوچک تر و ساختارهای محدود به کار رفته است [۳۳].

۴-۵ نتایج کمی ارزیابی اولیه روی پایگاه داده FIVES

نتایج کمی معماری های کاندید روی پایگاه داده مرجع FIVES در جدول (۴) گزارش شده است. مطابق این جدول، بهترین عملکرد مربوط به UNet+EffNet-B3 بوده و پس از آن FPN+ResNet50 و UNet+VGG16 قرار گرفته اند. بر این اساس، این سه مدل به عنوان مدل های برتر انتخاب شده و در ادامه فصل برای ارزیابی پایداری در حضور CA مورد استفاده قرار گرفته اند.

به منظور تکمیل ارزیابی کمی، نمونه ای از خروجی تفکیک عروق برای هریک از مدل های منتخب روی تصاویر بدون CA در شکل (۵) ارائه می شود.

^۱ pad
^۲ non-pad
^۳ Accuracy

(ی) ResUNet34، (ح) UNet + VGG16، (ط) FPN + ResNet50، (ی) UNet + EffNet-B3.

۵-۶ مطالعه فرسایشی^۱ تحلیل مؤلفه‌ای شبیه‌سازی افت کیفیت ناشی از CA

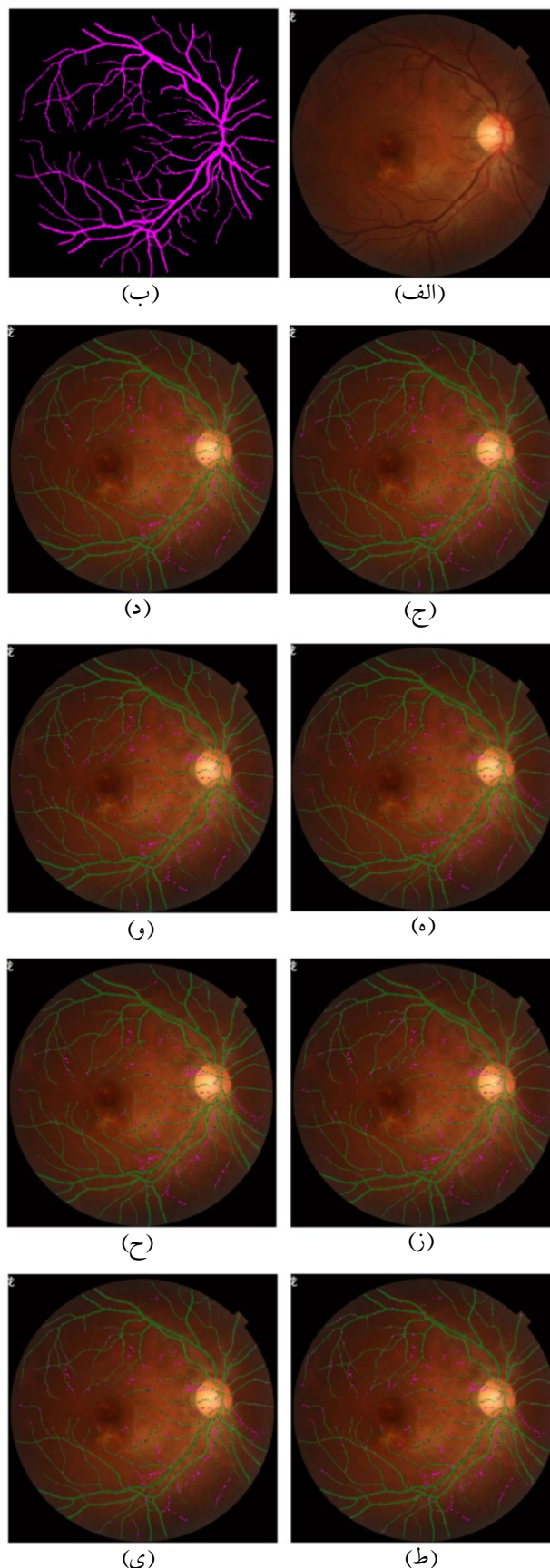
به منظور بررسی اثر مؤلفه‌های شبیه‌سازی CA در مرحله آموزش، یک مطالعه فرسایشی روی مدل منتخب UNet+EffNet-B3 انجام شد. در این مطالعه، شرایط آزمون برای همه حالت‌ها ثابت نگه داشته شد و تنها نوع داده‌های آموزشی تغییر یافت. برای هر وضعیت، مدل به صورت مستقل با یکی از ترکیب‌های مرحله‌ای معادلات معرفی شده در بخش ۲ آموزش داده شد و سپس روی تصاویر دارای تخریب کامل حاصل از معادلات (۱) تا (۶) ارزیابی گردید. نتایج این مطالعه در جدول (۶) گزارش شده است.

جدول (۶): مطالعه فرسایشی روی پایگاه داده FIVES-CA بر اساس ترکیب‌های مختلف معادلات شبیه‌سازی در داده‌های آموزشی

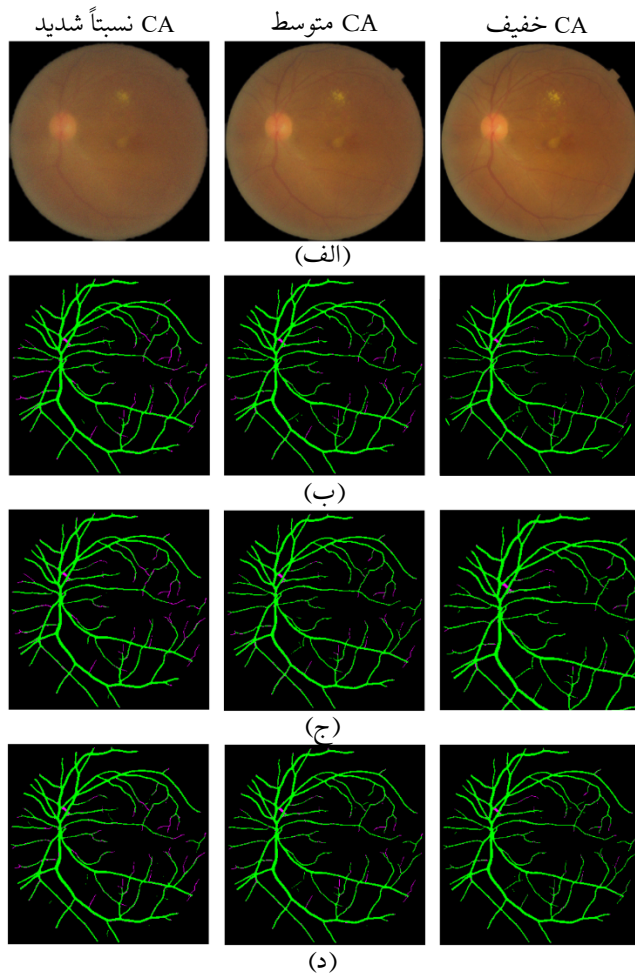
Acc	IoU	Dice	معادلات اعمال شده روی داده‌های آموزشی
۰/۹۷۹۸	۰/۷۳۶۳	۰/۸۴۸۱	فاقد تخریب
۰/۹۷۷۶	۰/۷۲۲۳	۰/۸۴۰۰	(۱)
۰/۹۷۹۵	۰/۷۳۵۵	۰/۸۴۷۶	(۱) و (۲)
۰/۹۸۴۱	۰/۷۷۹۱	۰/۸۷۲۰	(۱) الی (۳)
۰/۹۸۴۱	۰/۷۸۰۹	۰/۸۷۲۹	(۱) الی (۴)
۰/۹۸۳۹	۰/۷۷۸۶	۰/۸۷۱۷	(۱) الی (۵)
۰/۹۸۴۶	۰/۷۸۵۱	۰/۸۷۵۲	(۱) الی (۶)

مطابق جدول (۶)، آموزش با تصاویر اصلی (فاقد تخریب) در آزمون دارای تخریب کامل به مقادیر Acc، IoU، Dice به ترتیب برابر ۰/۸۴۸۱، ۰/۷۳۶۳ و ۰/۹۷۹۸ منجر شده اند که نشان دهنده افت مقادیر کمی نتایج تفکیک عروق شبکیه می باشد. با این حال، به دلیل عدم توازن ذاتی میان کلاس‌ها و غالب بودن تعداد پیکسل‌های زمینه نسبت به پیکسل‌های عروقی، حتی اگر بخش قابل ملاحظه‌ای از عروق به درستی شناسایی نشوند، تأثیر این خطا در محاسبه Acc به طور کامل قابل مشاهده نخواهد بود. از این رو، معیارهای Dice و IoU که تمرکز بیشتری بر میزان هم‌پوشانی نواحی عروقی تفکیک شده با مرجع واقعی دارند، ارزیابی دقیق‌تر و واقع‌بینانه‌تری از افت عملکرد مدل در شرایط تخریب ارائه می‌کنند.

همچنین به دلیل عدم اعمال کامل ملاحظات فیزیکی افت کیفیت تصاویر RF، اعمال معادله (۱) به تنهایی موجب کاهش عملکرد شده و افزودن معادله (۲) نیز تنها بخشی از این افت را جبران می‌کند. در مقابل، با ورود معادله (۳) به داده‌های آموزشی،



شکل (۵): تحلیل کیفی نمونه خروجی تفکیک عروق روی پایگاه داده FIVES (تصاویر RF فاقد CA) برای مدل‌های منتخب. (الف) تصویر اصلی، (ب) مرجع عروق، (ج) UNet، (د) UNet + ResNet50، (ه) UNet + ResNet50، (و) UNet++ + ResNet50، (ز) DeepLabV3+ + ResNet50، (ح) UNet + VGG16، (ی) FPN + ResNet50، (ط) UNet + EffNet-B3.



شکل (۶): تحلیل کیفی خروجی تفکیک عروق (GT پیش‌بینی روی واقعی) سه مدل برتر روی پایگاه‌داده FIVES-CA در سه سطح CA خفیف، متوسط و نسبتاً شدید. (الف) تصاویر با سه شدت CA (ب) UNet+VGG16 (ج) FPN+ResNet50 (د) UNet+EffNet-B3

جدول (۷): نتایج تفکیک عروق روی پایگاه‌داده DRIVE-CA (بخش دارای GT) برای مدل‌های آموزش‌دیده روی FIVES-CA

شدت CA	مدل	Dice	IoU	Acc
خفیف	FPN+ResNet50	۰/۶۲۶۵	۰/۴۵۷۸	۰/۹۳۳۴
	UNet+VGG16	۰/۶۶۲۲	۰/۴۹۷۷	۰/۹۴۲۱
	UNet+EffNet-B3	۰/۶۷۲۶	۰/۵۰۸۴	۰/۹۴۶۲
متوسط	FPN+ResNet50	۰/۶۰۹۵	۰/۴۴۰۹	۰/۹۳۱۷
	UNet+VGG16	۰/۶۲۳۸	۰/۴۵۵۴	۰/۹۴۰۱
	UNet+EffNet-B3	۰/۶۴۱۸	۰/۴۷۳۸	۰/۹۴۲۱
نسبتاً شدید	UNet+VGG16	۰/۴۹۷۵	۰/۳۳۲۸	۰/۹۲۰۱
	FPN+ResNet50	۰/۵۰۱۲	۰/۳۳۶۱	۰/۹۱۹۳
	UNet + EffNet-B3	۰/۵۳۶۴	۰/۳۶۷۴	۰/۹۲۸۲

مقدار Dice به ۰/۸۷۲۰ و مقدار IoU به ۰/۷۷۹۱ افزایش می‌یابد. این نتیجه نشان می‌دهد که ترکیب چند مؤلفه تخریب در داده‌های آموزشی، توان مدل را در مواجهه با تصاویر دارای افت کیفیت کامل افزایش می‌دهد.

در حالت‌های مربوط به اعمال معادلات (۱) تا (۴) و (۱) تا (۵)، مقادیر IoU و Dice در بازه‌ای نزدیک قرار دارند. این الگو نشان می‌دهد که پس از ورود مؤلفه‌های اصلی تخریب به داده‌های آموزشی، عملکرد مدل به سطح نسبتاً پایداری می‌رسد. در نهایت، اعمال کامل معادلات (۱) تا (۶)، که شامل مرحله آگاه از میدان دید است، بهترین نتیجه کلی را به دست می‌آورد. این نتیجه نشان می‌دهد که آموزش با داده‌های تولیدشده بر اساس زنجیره پیشنهادی، در مقایسه با آموزش بر تصاویر اصلی یا ترکیب‌های ناقص تخریب، موجب بهبود پایداری تفکیک عروق در حضور افت کیفیت ناشی از CA می‌شود.

برای تکمیل تحلیل نتایج، نمونه‌هایی از خروجی تفکیک عروق سه مدل برتر روی FIVES-CA در شکل (۶) ارائه می‌شود. این شکل شامل تصاویر ورودی، GT مرجع و ماسک‌های پیش‌بینی شده است و کاهش تدریجی رؤیت‌پذیری شاخه‌های باریک‌تر و تغییر رفتار مدل‌ها در سطوح بالاتر CA را نشان می‌دهد.

۵-۷ ارزیابی تعمیم‌پذیری بین پایگاه‌داده‌ای روی DRIVE-CA

به‌منظور بررسی اینکه پایداری مشاهده‌شده صرفاً محدود به پایگاه‌داده مرجع نباشد، عملکرد مدل‌های آموزش‌دیده روی FIVES-CA به‌صورت مکمل بر روی پایگاه‌داده DRIVE-CA نیز گزارش شده است. مطابق محدودیت ساختاری DRIVE و نبود GT عمومی در بخش آزمون، ارزیابی کمی صرفاً بر روی بخش آموزش دارای GT انجام شده است. نتایج این ارزیابی در جدول (۷) ارائه شده‌اند. از آنجا که پایگاه‌داده DRIVE از نظر وضوح تصویر، شرایط تصویربرداری، محدوده میدان دید، توزیع رنگ و ویژگی‌های برجسته‌گذاری با FIVES تفاوت دارد، مقادیر مطلق نتایج کمی روی DRIVE-CA نسبت به FIVES-CA کاهش قابل توجهی دارند. نتایج نشان می‌دهند که با افزایش شدت CA از خفیف به متوسط و نسبتاً شدید، عملکرد مدل‌ها در DRIVE-CA نیز کاهش تدریجی دارد. با این حال، روند افت عملکرد برای مدل UNet+EffNet-B3 همچنان پایین‌تر از سایر مدل‌ها است. نتایج حاصل بر روی DRIVE-CA نیز در سطوح خفیف تا نسبتاً شدید در شکل (۷) ارائه شده است.

تفکیک، پایداری در حضور CA و هزینه اجرایی است. مدل UNet+VGG16 نیز از نظر پیچیدگی در بازه میانی قرار می‌گیرد، اما در نتایج کمی، پایداری کمتری نسبت به UNet+EffNet-B3 نشان داده است.

جدول (۸): مقایسه پیچیدگی محاسباتی سه مدل منتخب

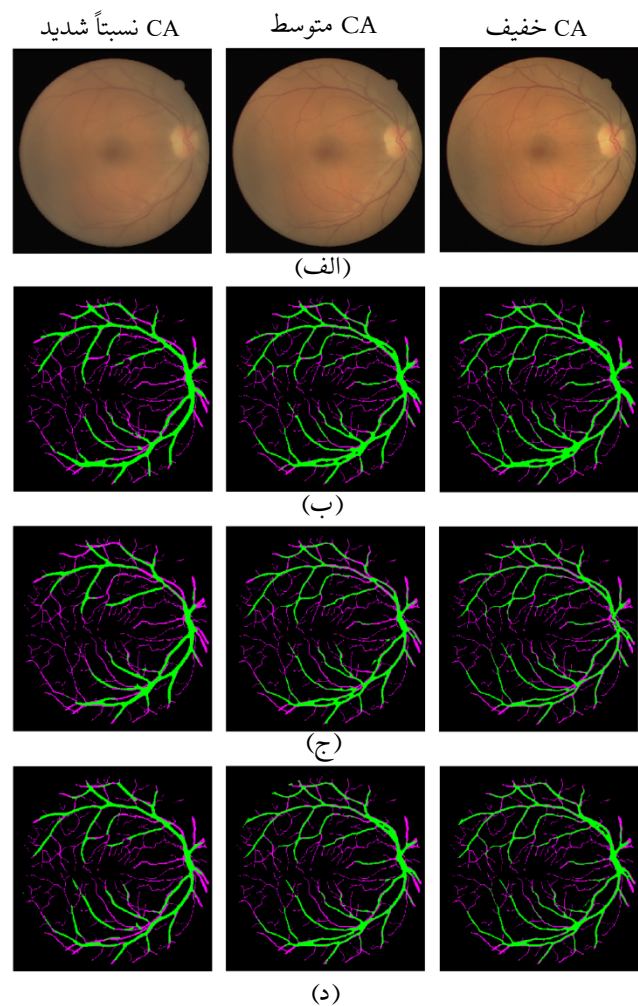
مدل	پارامترها (میلیون)	پیچیدگی محاسباتی (GFLOPs در 2048×2048)
UNet+VGG16	۲۰-۲۳	۵۲۰-۴۸۰
FPN+ResNet50	۲۵-۳۰	۵۶۰-۶۴۰
UNet+EffNet-B3	۱۲-۱۵	۳۲۰-۴۰۰

۵-۹ جمع‌بندی نتایج و تحلیل پایداری

بررسی نتایج کمی و کیفی نشان می‌دهد که انتخاب مدل برتر بر اساس ارزیابی روی پایگاه داده مرجع FIVES، منجر به انتخاب معماری‌هایی شده است که در حضور افت کیفیت ناشی از CA نیز کاهش عملکردی تدریجی و کنترل شده دارند. همچنین ارزیابی مکمل روی پایگاه داده DRIVE-CA نشان داد که اگرچه مقدار مطلق عملکرد نسبت به FIVES-CA کاهش می‌یابد، روند نسبی افت عملکرد با افزایش شدت CA تا حدی حفظ می‌شود و بهترین نتایج نیز متعلق به مدل UNet+EffNet-B3 است.

مطالعه فرسایشی نیز نشان داد که آموزش با داده‌های تولیدشده بر اساس ترکیب مرحله‌ای مؤلفه‌های تخریب، نسبت به آموزش بر تصاویر بدون تخریب، پایداری تفکیک عروق را در آزمون دارای تخریب کامل بهبود می‌دهد. همچنین بررسی هزینه محاسباتی نشان داد که UNet+EffNet-B3 در کنار عملکرد پایدارتر، پیچیدگی کمتری نسبت به FPN+ResNet50 دارد و از نظر توازن میان دقت، پایداری و هزینه اجرایی مناسب‌تر است.

اما به منظور بررسی رفتار مدل منتخب در مواجهه با تصاویر واقعی خارج از مجموعه‌های دارای برجسته مرجع، مجموعه‌ای از تصاویر RF از یک پایگاه داده عمومی مستقل RFMiD [۳۴] مورد بررسی قرار گرفت. با مشارکت متخصصان، تصاویر بر اساس شدت افت کیفیت ناشی از CA به سه سطح طبقه‌بندی شدند؛ به گونه‌ای که این سطوح از نظر شدت، با سطوح مورد استفاده در مرحله آموزش و ارزیابی مدل تا حدودی متناظر باشند. در ادامه، تصاویر انتخاب شده به مدل برتر UNet+EffNet-B3 اعمال شده و ماسک پیش‌بینی شده تفکیک عروق استخراج گردید. از آنجا که این تصاویر فاقد GT بودند، ارزیابی این بخش به صورت کیفی و توسط متخصصان انجام شد و به عنوان ارزیابی کمی در نظر گرفته



شکل (۷): تحلیل کیفی خروجی تفکیک عروق (GT پیش‌بینی روی GT واقعی) سه مدل برتر روی پایگاه داده DRIVE در سه سطح CA خفیف، متوسط و نسبتاً شدید. (الف) تصاویر با سه شدت CA (ب) UNet+VGG16 (ج) FPN+ResNet50 (د) UNet+EffNet-B3

۵-۸ بررسی هزینه محاسباتی مدل‌های منتخب

پس از ارزیابی کمی و کیفی مدل‌ها، هزینه محاسباتی سه مدل منتخب نیز بررسی شد تا توازن میان دقت، پایداری و پیچیدگی اجرایی آن‌ها مشخص شود. این مقایسه به سه مدلی محدود شده است که در مراحل اصلی ارزیابی FIVES-CA و DRIVE-CA مورد استفاده قرار گرفته‌اند. تعداد پارامترهای قابل آموزش و پیچیدگی محاسباتی تقریبی این مدل‌ها در جدول (۸) گزارش شده است.

مطابق جدول (۸)، مدل FPN+ResNet50 به دلیل استفاده از رمزگذار Residual و مسیرهای چندمقیاسی، بیشترین تعداد پارامتر و هزینه محاسباتی را دارد. در مقابل، UNet+EffNet-B3 با وجود دستیابی به بهترین عملکرد در اغلب سناریوهای ارزیابی، تعداد پارامتر و پیچیدگی محاسباتی کمتری نسبت به دو مدل دیگر نشان می‌دهد. این ویژگی بیانگر توازن مناسب این مدل میان دقت

حفظ شده‌اند. این یافته‌ها بیانگر رفتار ساختاری پایدار مدل در مواجهه با تصاویر واقعی است.

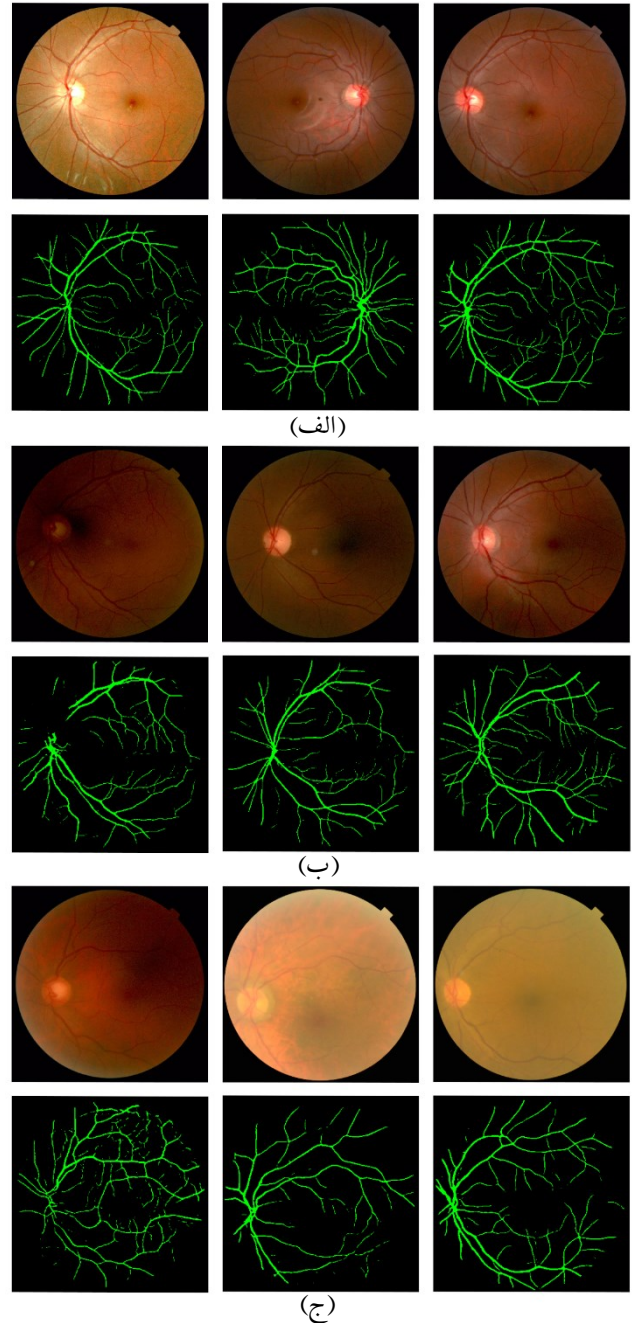
۶ نتیجه‌گیری

اکثر بیماری‌هایی که از طریق پردازش RF تشخیص داده می‌شوند، عمدتاً در سنن بالا رخ می‌دهد و در این سنن، بخش قابل ملاحظه‌ای از بیماران، مبتلا به CA نیز می‌باشند که تشخیص بیماری را از طریق بررسی RF با مشکل جدی روبرو می‌کند. لذا در این مقاله، برخلاف بسیاری از مطالعات پیشین که ارزیابی خود را به تصاویر سالم و با کیفیت بالا محدود کرده‌اند، اثر CA به‌عنوان یکی از شایع‌ترین و در عین حال نادیده‌گرفته‌شده‌ترین عوامل تخریب کیفیت تصاویر شبکه، به‌صورت سیستماتیک و کنترل‌شده مورد بررسی قرار گرفت. در این پژوهش، یک چارچوب مرحله‌ای و آگاه از افت کیفیت برای تفکیک عروق شبکه از تصاویر RF در حضور CA و مبتنی بر سه مؤلفه کلیدی (۱) شبیه‌سازی واقع‌گرایانه CA محدود به ناحیه FOV، (۲) آموزش شبکه‌های تفکیک عروق با داده‌های شبیه‌سازی‌شده حاوی CA در سطوح مختلف شدت و (۳) اعمال فرآیند استخراج نهایی عروق به‌گونه‌ای که محاسبه معیارهای کمی صرفاً در ناحیه معتبر شبکه انجام شود. این ترکیب امکان تحلیل منظم و قابل تکرار رفتار الگوریتم‌های تفکیک عروق را در حضور تاری اپتیکی، افت تمایز و پراکندگی نور فراهم می‌سازد، بدون آنکه نیاز به تنظیمات دستی وابسته به هر تصویر وجود داشته باشد.

مطالعه فرسایشی مؤلفه‌های شبیه‌سازی نشان داد که پایداری مشاهده‌شده صرفاً حاصل انتخاب مدل نیست، بلکه نوع داده‌های آموزشی و ترکیب مرحله‌ای مؤلفه‌های تخریب نیز در افزایش مقاومت مدل در برابر افت کیفیت نقش دارد.

نتایج ارزیابی اولیه روی پایگاه داده مرجع FIVES نشان داد که در میان معماری‌های متداول و مطرح تفکیک تصویر، ترکیب UNet با رمزگذار EffNet-B3 بالاترین مقادیر IoU، Dice و Accuracy را در تصاویر سالم به دست می‌آورد و نسبت به سایر معماری‌ها از توازن مناسب‌تری میان دقت و پایداری برخوردار است. بر این اساس، این مدل به‌عنوان گزینه اصلی برای ورود به سناریوی وجود CA انتخاب شد. ارزیابی‌های انجام‌شده روی نسخه‌های شبیه‌سازی شده FIVES-CA نشان داد که اگرچه با افزایش شدت CA، معیارهای هم‌پوشانی نظیر IoU و Dice به‌صورت تدریجی کاهش می‌یابند، اما این کاهش برای مدل UNet+EffNet-B3 یکنواخت و کنترل‌شده‌تر از بقیه مدل‌ها بوده و حتی در شدیدترین سطح CA نیز مقادیر غیرصفر و معناداری حفظ می‌شوند. این رفتار بیانگر توانایی مدل در حفظ ساختار کلی شبکه عروقی، به‌ویژه عروق اصلی و متوسط، در شرایط افت کیفیت قابل توجه است.

نمی‌شود. معیارهای ارزیابی کیفی شامل حفظ پیوستگی عروق اصلی، بازیابی شاخه‌های ثانویه، میزان پاسخ‌های کاذب در نواحی غیرعروقی و حفظ ساختار کلی شبکه عروقی بود.



شکل (۸): تفکیک عروق روی تصاویر واقعی فاقد GT از پایگاه داده RFMid در سه سطح شدت CA توسط مدل UNet+EffNet-B3. (الف) CA خفیف (ب) CA متوسط (ج) CA نسبتاً شدید.

همانطور که در شکل (۸) قابل مشاهده است، در سطوح خفیف و متوسط، ساختار کلی شبکه عروقی با انسجام مناسبی بازسازی شده است. در سطح نسبتاً شدید، اگرچه بازیابی عروق باریک کاهش یافته، اما عروق اصلی همچنان به‌صورت معنادار

مراجع

- [۱] M. D. Abràmoff, M. K. Garvin, and M. Sonka, "Retinal Imaging and Image Analysis," *IEEE Reviews in Biomedical Engineering*, vol. 3, pp. 169-208, 2010, doi: 10.1109/RBME.2010.2084567.
- [۲] A. D. Fleming, S. Philip, K. A. Goatman, J. A. Olson, and P. F. Sharp, "Automated Assessment of Diabetic Retinal Image Quality Based on Clarity and Field Definition," *Investigative Ophthalmology & Visual Science*, vol. 47, no. 3, pp. 1120-1125, 2006, doi: 10.1167/iov.05-1155.
- [۳] A. Hoover and M. Goldbaum, "Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels," *IEEE Transactions on Medical Imaging*, vol. 22, no. 8, pp. 951-958, 2003, doi: 10.1109/TMI.2003.815900.
- [۴] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. v. Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Transactions on Medical Imaging*, vol. 23, no. 4, pp. 501-509, 2004, doi: 10.1109/TMI.2004.825627.
- [۵] E. Gonzalez-Amador, J. Arines, P. Charlón, N. Garcia-Porta, M. J. Abralles, and E. Acosta, "Improvement of Retinal Images Affected by Cataracts," *Photonics*, vol. 9, no. 4, p. 251, 2022, doi: [10.3390/photonics9040251](https://doi.org/10.3390/photonics9040251).
- [۶] X. Fang, Y. Wang, X. Li, W. Fan, and D. Zhou, "A two-stage multi-scale attention-based network for weakly supervised cataract fundus image enhancement," *Scientific Reports*, vol. 15, no. 1, p. 27610, 2025/07/29 2025, doi: 10.1038/s41598-025-12157-6.
- [۷] A. B. Crane, H. S. Choudhry, and M. H. Dastjerdi, "Effect of simulated cataract on the accuracy of artificial intelligence in detecting diabetic retinopathy in color fundus photos," *Indian Journal of Ophthalmology*, vol. 72, no. Suppl 1, pp. S42-S45, 2024, doi: 10.4103/ijo.Ijo_1163_23.
- [۸] L. Xiong, H. Li, and L. Xu, "An Approach to Evaluate Blurriness in Retinal Images with Vitreous Opacity for Cataract Diagnosis," *Journal of Healthcare Engineering*, vol. 2017, no. 1, p. 5645498, 2017, doi: [10.1155/2017/5645498](https://doi.org/10.1155/2017/5645498).
- [۹] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Cham, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds., 2015// 2015: Springer International Publishing, pp. 234-241, doi: [10.1007/978-3-319-24574-4_28](https://doi.org/10.1007/978-3-319-24574-4_28).
- [۱۰] L. Jiao *et al.*, "A Survey of Deep Learning-Based Object Detection," *IEEE Access*, vol. 7, pp. 128837-128868, 2019, doi: 10.1109/ACCESS.2019.2939201.
- [۱۱] F. D. Hernandez-Gutierrez, E. G. Avina-Bravo, M. A. Ibarra-Manzano, J. Ruiz-Pinales, E. Ovalle-Magallanes, and J. G. Avina-Cervantes, "Retinal Vessel Segmentation Based on a Lightweight U-Net and Reverse Attention," *Mathematics*, vol. 13, no. 13, p. 2203, 2025. doi: 10.3390/math13132203
- [۱۲] Z. Gong *et al.*, "Versatile cataract fundus image restoration model utilizing unpaired cataract and high-

تحلیل‌های کیفی ارائه‌شده در فصل ۵ نیز این نتایج کمی را تأیید می‌کند. در تصاویر دارای CA خفیف و متوسط، خروجی مدل منتخب همچنان پیوستگی مناسبی در شاخه‌های اصلی شبکه عروقی حفظ می‌کند و در شدت‌های بالاتر، اگرچه عروق باریک‌تر به تدریج تضعیف می‌شوند، اما ساختار کلی شبکه عروق به طور کامل فرو نمی‌پاشد. این ویژگی برای کاربردهای بالینی، به‌ویژه در سناریوهای غربالگری که تصاویر ورودی اغلب دارای کیفیت غیرایده‌آل هستند، اهمیت ویژه‌ای دارد.

لازم به تأکید است که چارچوب پیشنهادی با هدف دستیابی به بیشترین دقت ممکن در تصاویر کاملاً سالم طراحی نشده است، بلکه تمرکز اصلی آن بر پایداری عملکرد در حضور افت کیفیت ناشی از CA بوده است. از این رو، کاهش نسبی عملکرد در تصاویر سالم در مقایسه با برخی روش‌های بهینه‌شده خاص برای تصاویر شفاف، قابل انتظار و از منظر هدف مسئله قابل توجیه است.

در مجموع، نتایج این پژوهش نشان می‌دهد که ترکیب شبیه‌سازی آگاه از CA، انتخاب معماری مناسب بر اساس ارزیابی مرجع، و ارزیابی سیستماتیک پایداری، می‌تواند به شناسایی مدل‌هایی منجر شود که برای کاربردهای عملی و بالینی واقع‌گرایانه مناسب‌تر هستند. بر اساس نتایج به‌دست‌آمده در چارچوب ارزیابی کنترل‌شده این پژوهش، UNet+EfficientNet-B3 در میان مدل‌های بررسی‌شده بهترین توازن را از نظر دقت و پایداری نسبی حتی در شرایط ارزیابی بر روی پایگاه‌های داده مختلف که دارای شرایط متفاوت وضوح و نحوه تصویربرداری هستند نیز نشان داد. با این حال، استفاده عملی از چنین مدلی در سامانه‌های کمک‌تشخیصی نیازمند ارزیابی بیشتر روی تصاویر واقعی مبتلا به آب مروارید همراه با GT پیکسلی است. علاوه بر ارزیابی‌های کمی مبتنی بر داده‌های دارای برجسب، ارزیابی کیفی روی تصاویر واقعی فاقد GT از یک پایگاه‌داده مستقل نیز انجام شد. نتایج این بررسی تخصصی نشان داد که مدل منتخب قادر است ساختار عروق اصلی و بخش قابل‌توجهی از شاخه‌های ثانویه را حتی در شرایط افت کیفیت واقعی بازسازی کند. این یافته بیانگر پتانسیل کاربردی چارچوب پیشنهادی در سناریوهای عملی و بالینی است که در آن‌ها دسترسی به GT پیکسلی معمولاً وجود ندارد.

در کارهای آینده، ساخت پایگاه‌داده‌ای از تصاویر واقعی RF مبتلا به CA که پس از درمان برجسب‌های پیکسلی عروق استخراج شوند، استفاده از راهبردهای یادگیری نیمه‌نظارتی یا خودنظارتی، و توسعه مدل‌های تطبیقی برای برآورد خودکار شدت افت کیفیت و تنظیم هوشمند سیاست تفکیک، می‌تواند به بهبود بیشتر عملکرد، به‌ویژه در شدت‌های بالای CA، منجر شود.

- [۲۵] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," *ArXiv*, vol. abs/1905.11946, 2019.
- [۲۶] M. S. Gargari, M. H. Seyedi, and M. Alilou, "Segmentation of Retinal Blood Vessels Using U-Net++ Architecture and Disease Prediction," *Electronics*, vol. 11, no. 21, p. 3516, 2022, doi: 10.3390/electronics11213516.
- [۲۷] W. Liu, Y. Jiang, J. Zhang, and Z. Ma, "RFARN: Retinal vessel segmentation based on reverse fusion attention residual network," *PLOS ONE*, vol. 16, no. 12, p. e0257256, 2021, doi: 10.1371/journal.pone.0257256.
- [۲۸] Y. Liu, J. Shen, L. Yang, G.-B. Bian, and H. Yu, "ResDO-UNet: A deep residual network for accurate retinal vessel segmentation from fundus images," *Biomedical Signal Processing and Control*, vol. 79, p. 104087, 2023, doi: 10.1016/j.bspc.2022.104087.
- [۲۹] N. Alsharabi, A. Echtioui, A. Alayba, G. Alshammari, and H. Hamam, "Advanced hybrid UNet architectures for retinal vessel segmentation," *Egyptian Informatics Journal*, vol. 32, p. 100823, 2025, doi: 10.1016/j.eij.2025.100823.
- [۳۰] T. Haklı and İ. Zaim Gökbay, "Evaluation of Image Segmentation Models in Retinal Vessel Segmentation," in *Medical Informatics IV*, N. Bozbuğa and S. Gülseçen Eds.: Istanbul University Press, 2024, pp. 105-112.
- [۳۱] M. C. S. Tang, S. S. Teoh, and H. Ibrahim, "Retinal Vessel Segmentation from Fundus Images Using DeepLabv3+," presented at the 2022 IEEE 18th International Colloquium on Signal Processing & Applications (CSPA), 2022, doi: [10.1109/CSPA55076.2022.9781891](https://doi.org/10.1109/CSPA55076.2022.9781891).
- [۳۲] M. R. Mathews, S. M. Anzar, R. Kalesh Krishnan, and A. Panthakkan, "EfficientNet for retinal blood vessel segmentation," presented at the 2020 3rd International Conference on Signal Processing and Information Security (ICSPIS), 2020, doi: [10.1109/ICSPIS51252.2020.9340135](https://doi.org/10.1109/ICSPIS51252.2020.9340135).
- [۳۳] F. Milletari, N. Navab, and S. A. Ahmadi, "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation," in *2016 Fourth International Conference on 3D Vision (3DV)*, 25-28 Oct. 2016, pp. 565-571, doi: 10.1109/3DV.2016.79.
- [۳۴] S. Pachade et al., "Retinal Fundus Multi-Disease Image Dataset (RFMiD): A Dataset for Multi-Disease Detection Research," *Data*, vol. 6, no. 2, p. 14, 2021. [Online]. Available: <https://www.mdpi.com/2306-5729/6/2/14>.
- [۱۳] E. Peli and T. Peli, "Restoration of retinal images obtained through cataracts," *IEEE Transactions on Medical Imaging*, vol. 8, no. 4, pp. 401-406, 1989, doi: 10.1109/42.41493.
- [۱۴] X. Xu, Y. Guan, J. Li, Z. Ma, L. Zhang, and L. Li, "Automatic glaucoma detection based on transfer induced attention network," *BioMedical Engineering OnLine*, vol. 20, no. 1, p. 39, 2021/04/23 2021, doi: 10.1186/s12938-021-00877-5.
- [۱۵] H. Xie et al., "Deep learning for detecting visually impaired cataracts using fundus images," (in English), *Frontiers in Cell and Developmental Biology*, Original Research vol. Volume 11 - 2023, 2023-July-28 2023, doi: 10.3389/fcell.2023.1197239.
- [۱۶] م. زمانی و ر. حسن‌زاده، «تشخیص بیماری‌های شایع چشمی در تصاویر شبکیه حاوی آب‌مروارید با استفاده از شبکه یادگیری عمیق»، مجله ماشین بینایی و پردازش تصویر، دوره ۱۰، شماره ۴، صص. ۷۵-۸۴، ۱۴۰۲.
- [۱۷] M. F. Marmor, "Ophthalmology and Art: Simulation of Monet's Cataracts and Degas' Retinal Disease," *Archives of Ophthalmology*, vol. 124, no. 12, pp. 1764-1769, 2006, doi: 10.1001/archophth.124.12.1764.
- [۱۸] V. F. Pamplona et al., "CATRA: interactive measuring and modeling of cataracts," *ACM Trans. Graph.*, vol. 30, no. 4, p. Article 47, 2011, doi: 10.1145/2010324.1964942.
- [۱۹] K. Jin et al., "FIVES: A Fundus Image Dataset for Artificial Intelligence based Vessel Segmentation," *Scientific Data*, vol. 9, no. 1, p. 475, 2022/08/04 2022, doi: 10.1038/s41597-022-01564-3.
- [۲۰] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A Nested U-Net Architecture for Medical Image Segmentation," Cham, 2018: Springer International Publishing, in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pp. 3-11.
- [۲۱] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.
- [۲۲] T. Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature Pyramid Networks for Object Detection," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 21-26 July 2017, pp. 936-944, doi: 10.1109/CVPR.2017.106.
- [۲۳] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation," Cham, 2018: Springer International Publishing, in *Computer Vision – ECCV 2018*, pp. 833-851.
- [۲۴] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *CoRR*, vol. abs/1409.1556, 2014.



آرش آقائی مدرک کارشناسی و کارشناسی ارشد خود را به ترتیب در گرایش نرم‌افزار در سال‌های ۱۳۹۹ و ۱۴۰۴ از دانشگاه گیلان اخذ نمود. وی هم‌اکنون دانشجوی دکتری مهندسی کامپیوتر در گرایش هوش مصنوعی می‌باشد. زمینه‌های پژوهشی ایشان

شامل اینترنت اشیا، امنیت اینترنت اشیا، یادگیری ماشین و شبکه‌های یادگیری عمیق است.



رضا حسن‌زاده پاک‌رضائی مدرک کارشناسی خود را در رشته مهندسی برق با گرایش الکترونیک در سال ۱۳۷۹ از دانشگاه گیلان دریافت کرد. ایشان مدرک کارشناسی‌ارشد و دکتری خود را به‌ترتیب در گرایش الکترونیک دیجیتال مهندسی برق در سال‌های ۱۳۸۱ و ۱۳۸۶ از دانشگاه صنعتی امیرکبیر (پلی‌تکنیک تهران) اخذ نمود. وی هم‌اکنون دانشیار گروه مهندسی برق دانشگاه گیلان است. زمینه‌های پژوهشی ایشان شامل پردازش سیگنال‌ها و تصاویر صنعتی و پزشکی، یادگیری ماشین، هوش محاسباتی و منطق فازی می‌باشد.



اسدالله شاه‌بهرامی مدرک کارشناسی مهندسی کامپیوتر را از دانشگاه علم و صنعت ایران، مدرک کارشناسی‌ارشد مهندسی کامپیوتر با گرایش هوش مصنوعی را از دانشگاه شیراز و مدرک دکتری مهندسی کامپیوتر را از دانشگاه صنعتی دلفت هلند دریافت نموده است. ایشان هم‌اکنون استاد تمام گروه مهندسی کامپیوتر دانشگاه گیلان می‌باشد. زمینه‌های پژوهشی ایشان شامل بینایی ماشین، پردازش تصویر دیجیتال، یادگیری ماشین، هوش مصنوعی، رایانش با کارایی بالا و کاربرد روش‌های هوشمند در مسائل مهندسی و پزشکی است.