

تقسیم‌بندی تصاویر فوندوس شبکیه چشم به کمک یادگیری عمیق

سید علیرضا خاتمی^۱، منصور فاتح^۲، محسن رضوانی^۳، امیررضا فاتح^۴

چکیده

تقسیم‌بندی دقیق رگ‌های خونی در تصاویر شبکیه به دلیل اهمیت آن در تشخیص زودهنگام بیماری‌های چشمی نظیر رتینوپاتی دیابتی و گلوکوم، از اهمیت بالایی برخوردار است. با توجه به چالش‌هایی مانند نویز، کنتراست پایین و عدم تعادل کلاسی در تصاویر شبکیه، تشخیص بیماری‌های چشمی سخت می‌باشد. این پژوهش با هدف توسعه یک روش خودکار و دقیق برای تقسیم‌بندی رگ‌های خونی، یک معماری جدید مبتنی بر U-Net ارائه می‌دهد. در مرحله پیش‌پردازش، از روش بهبود کنتراست CLAHE و تکنیک‌های افزایش داده مانند چرخش و بزرگ‌نمایی برای ارتقای کیفیت تصاویر و غنی‌سازی مجموعه داده استفاده شده است. در U-Net پیشنهادی، با جایگزینی بلوک‌های مرسوم با بلوک‌های توجه کانولوشنی تقویت شده است. این بلوک‌ها با توجه همزمان بر روی ویژگی‌های مکانی و کانالی، به مدل در استخراج مؤثرتر ویژگی‌های مرتبط با رگ‌ها کمک می‌کنند. روش پیشنهادی بر روی دادگان Retina Blood Vessel Segmentation دقت ۰/۹۵۲۳ را به دست آورد که نشان‌دهنده برتری عملکرد آن در مقایسه با روش‌های پایه است.

کلید واژه‌ها

تقسیم‌بندی رگ‌های خونی، یادگیری عمیق، شبکه U-Net، مکانیزم توجه، محدودسازی کنتراست، تابع زیان ترکیبی

۱- مقدمه

اطلاعات ساختاری عروق خونی، در تشخیص بیماری‌های چشمی مانند رتینوپاتی دیابتی، گلوکوم و دژنراسیون ماکولا و حتی بیماری‌های سیستمیک مانند فشار خون بالا و بیماری‌های قلبی‌عروقی از اهمیت ویژه‌ای برخوردارند [۲، ۳]. از آنجا که این بیماری‌ها در مراحل اولیه اغلب بدون علائم آشکار هستند، توسعه روش‌های خودکار برای تحلیل دقیق تغییرات شبکیه ضروری است [۴]. تقسیم‌بندی دقیق رگ‌های خونی در تصاویر فوندوس شبکیه به دلیل چالش‌هایی مانند نویز، کنتراست پایین، تنوع مقیاس رگ‌ها (از رگ‌های ضخیم تا مویرگ‌های نازک) و عدم تعادل شدید بین پیکسل‌های رگ و پس‌زمینه، یک چالش جدی است [۵]. روش‌های سنتی مانند فیلترهای گابور، روش Frangi و آستانه‌گذاری، به دلیل نیاز به تنظیم دستی و حساسیت به نویز، در تصاویر با کیفیت پایین عملکرد محدودی دارند [۶، ۷]. در مقابل، روش‌های مبتنی بر یادگیری عمیق، به‌ویژه شبکه‌های U-Net و نسخه‌های بهبودیافته آن با مکانیزم‌های توجه، دقت و تعمیم‌پذیری بهتری ارائه داده‌اند [۸].

پردازش تصاویر پزشکی در تشخیص سریع و دقیق بیماری‌ها نقش کلیدی ایفا می‌کند [۱]. تصاویر فوندوس شبکیه چشم به دلیل ارائه

مقاله در تاریخ ۲۳ مهر ماه ۱۴۰۴ دریافت شد

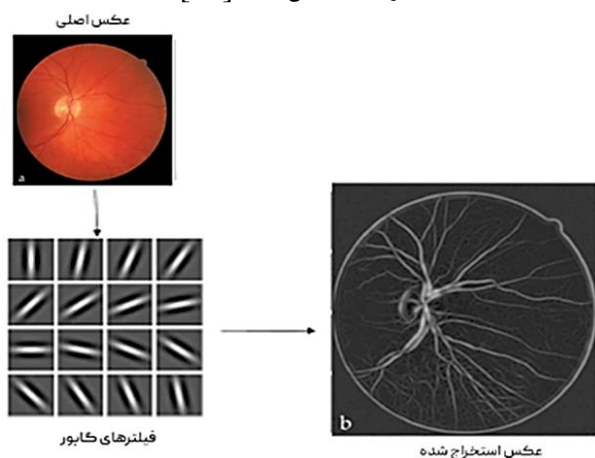
- دانشجوی کارشناسی ارشد دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شاهرود، شاهرود، رایانامه: alirezaaorg@gmail.com
 - دانشیار دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شاهرود، شاهرود، رایانامه: mansoor_fateh@shahroodut.ac.ir
 - دانشیار دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شاهرود، شاهرود، رایانامه: mrezvani@shahroodut.ac.ir
 - دانشجو دکتری دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران، تهران، رایانامه: amirreza_fateh@comp.iust.ac.ir
- نویسنده مسئول: منصور فاتح

۱-۲ روش‌های سنتی

روش‌های سنتی برای تقسیم‌بندی رگ‌های خونی از ویژگی‌های هندسی و آماری مانند فیلترهای همسان، گابور، Frangi و ردیابی استفاده می‌کنند. این روش‌ها به دلیل سادگی و سرعت در گذشته پرکاربرد بودند، اما در برابر نویز، کنتراست پایین و تنوع ساختار رگ‌ها محدودیت‌هایی دارند [۱۶].

فیلترهای همسان با استفاده از الگوهای شبیه به شکل رگ‌ها، ساختارهای عروقی را شناسایی می‌کنند. این روش‌ها نیازی به داده‌های آموزشی ندارند و سریع هستند، اما در تصاویر با کیفیت پایین یا پر نویز (مانند تصاویر بیماران مبتلا به رتینوپاتی دیابتی) اغلب رگ‌ها را به اشتباه تشخیص می‌دهند [۱۷]. برای مثال، در دادگان DRIVE، فیلترهای همسان رگ‌های اصلی را به خوبی شناسایی می‌کنند، اما در رگ‌های نازک‌تر دقت محدودی دارند [۱۸].

فیلترهای گابور با تقلید از عملکرد سیستم بینایی انسان، ساختارهای خطی و رگ مانند را تقویت می‌کنند. همان‌طور که در شکل (۱) مشاهده می‌شود، این روش برای شناسایی رگ‌ها با اندازه‌ها و جهت‌گیری‌های مختلف مناسب است، اما نیاز به تنظیم دستی پارامترها دارد که می‌تواند زمان‌بر باشد و در تصاویر متنوع (مانند دادگان STARE) عملکرد را کاهش دهد [۱۹].



شکل ۱: تصویر شبکه‌ای با فیلتر گابور پردازش شده و پس از شناسایی الگوهای عروقی، عروق به صورت مجزا از پس‌زمینه استخراج شده‌اند.

فیلترهای Frangi برای شناسایی ساختارهای لوله‌ای مانند رگ‌های خونی طراحی شده‌اند و با استفاده از تکنیک‌های اضافی مانند حذف نویز، نتایج دقیق‌تری تولید می‌کنند. این روش در تصاویر با وضوح بالا رگ‌های نازک را بهتر شناسایی می‌کند، اما در برابر تغییرات نوری حساس است [۲۰].

۲-۲ روش‌های ردیابی

روش‌های ردیابی از یک نقطه شروع آغاز شده و مسیر رگ‌ها را دنبال می‌کنند. که در شکل (۲) قابل نمایش است این روش برای شناسایی رگ‌های اصلی و شاخه‌دار مناسب است، اما اگر نقطه شروع به‌درستی

[۹]. با این حال، این روش‌ها همچنان با مشکلاتی مانند ناتوانی در شناسایی رگ‌های بسیار نازک، تأثیر منفی عدم تعادل کلاس‌ها و از دست دادن جزئیات در نواحی کوچک مواجه هستند [۱۰، ۱۱].

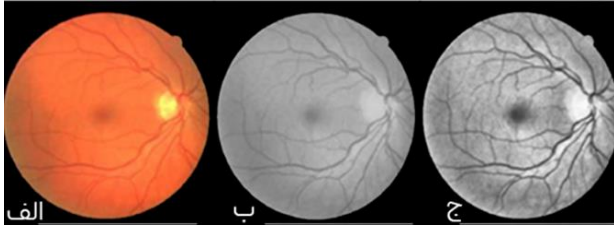
برای رفع این محدودیت‌ها، این پژوهش یک معماری جدید مبتنی بر U-Net، به نام ATEN-UNet پیشنهاد می‌دهد که با ادغام گره مرکزی چندمقیاسی، ماژول توجه تطبیقی ATEN و تابع زیان ترکیبی (Dice Loss + BCE)، دقت و تعمیم‌پذیری تقسیم‌بندی رگ‌های خونی را بهبود می‌بخشد. گره چندمقیاسی با کرنل‌های متنوع ویژگی‌های محلی و جهانی را به‌طور مؤثر استخراج می‌کند، ماژول ATEN ویژگی‌های مرتبط با رگ‌ها را تقویت کرده و نویز پس‌زمینه را سرکوب می‌کند و تابع زیان ترکیبی تعادل بین دقت پیکسلی و همپوشانی را برقرار می‌سازد [۱۲]. این نوآوری‌ها امکان شناسایی دقیق‌تر رگ‌های نازک و پیچیده را فراهم کرده و عملکرد مدل را در مقایسه با روش‌های پایه به‌طور قابل توجهی ارتقای می‌دهند.

ساختار مقاله به صورت زیر تنظیم شده است: در بخش مقدمه، اهمیت موضوع و چالش‌های موجود در تقسیم‌بندی رگ‌های خونی معرفی می‌شود. در بخش دوم، روش‌های پیشین سنتی و مبتنی بر یادگیری عمیق مرور و محدودیت‌های آن‌ها بیان می‌گردد. بخش سوم به تشریح معماری پیشنهادی ATEN-UNet، اجزای اصلی آن و تابع زیان ترکیبی اختصاص دارد. در بخش چهارم، جزئیات پیاده‌سازی، داده‌ها و معیارهای ارزیابی توضیح داده شده است. در نهایت، بخش نتایج و جمع‌بندی شامل تحلیل عملکرد مدل، مقایسه با روش‌های دیگر و پیشنهادهایی برای کارهای آینده است.

۲- کارهای پیشین

به طور کلی، تقسیم‌بندی خودکار در حوزه پزشکی برای بسیار حائز اهمیت است [۱۳، ۱۴]. تقسیم‌بندی خودکار رگ‌های خونی در تصاویر فوندوس شبکیه یکی از حوزه‌های کلیدی در پردازش تصویر پزشکی است که در تشخیص زودهنگام بیماری‌هایی مانند رتینوپاتی دیابتی، گلوکوم و دژنراسیون ماکولا نقش حیاتی دارد. تغییرات در الگوی عروق شبکیه همچنین می‌تواند نشانگر بیماری‌های سیستمیک مانند فشار خون بالا و بیماری‌های قلبی-عروقی باشد [۱۵]. با افزایش شیوع دیابت و پیری جمعیت، نیاز به سیستم‌های خودکار و مقیاس‌پذیر برای غربالگری چشمی به‌طور چشمگیری افزایش یافته است. این بخش به بررسی روش‌های پیشین در تقسیم‌بندی رگ‌های خونی پرداخته و محدودیت‌های آن‌ها را تحلیل می‌کند تا زمینه را برای نوآوری‌های روش پیشنهادی فراهم کند. به طور کلی، روش‌های موجود را می‌توان در سه دسته اصلی شامل روش‌های سنتی مبتنی بر پردازش تصویر، روش‌های یادگیری ماشین کلاسیک و روش‌های مبتنی بر یادگیری عمیق تقسیم‌بندی کرد که در ادامه هر یک به تفصیل بررسی می‌شوند.

داده‌های مصنوعی، این مشکل را برطرف می‌کنند. برای مثال، تابع زیان کانونی با تمرکز بر پیکسل‌های دشوار، دقت شناسایی رگ‌های نازک را بهبود می‌بخشد [۲۸]. همانطور که در شکل (۳) قابل مشاهده است روش‌های ترکیبی باعث شده خروجی نهایی ما بسیار قابل تشخیص و حتی بهتر از ماسک تصویری که ما به مدل دادیم شده باشد.



شکل ۳: روش‌های ترکیبی برای کاهش عدم تعادل کلاس. (الف) تصویر اصلی شبکه، (ب) تصویر پیش‌پردازش شده برای تقویت عروق، و (ج) تصویر نهایی بهبودیافته با تمایز بهتر کلاس‌ها.

۶-۲ مرور نهایی و شناسایی شکاف‌های پژوهشی

روش‌های سنتی به دلیل سادگی هنوز کاربرد دارند، اما در برابر نویز و تنوع ساختار رگ‌ها محدود هستند. روش‌های یادگیری عمیق مانند U-Net و نسخه‌های بهبودیافته آن با مکانیزم‌های توجه، دقت و حساسیت را بهبود بخشیده‌اند. با این حال، چالش‌هایی مانند شناسایی رگ‌های بسیار نازک، مدیریت نویز و عدم تعادل کلاسی همچنان باقی است. روش‌های ترکیبی و استفاده از داده‌های مصنوعی می‌توانند این محدودیت‌ها را کاهش دهند، اما نیاز به مدل‌هایی با تعمیم‌پذیری بالاتر و توانایی استخراج ویژگی‌های چندمقیاسی وجود دارد که روش پیشنهادی به آن پاسخ می‌دهد.

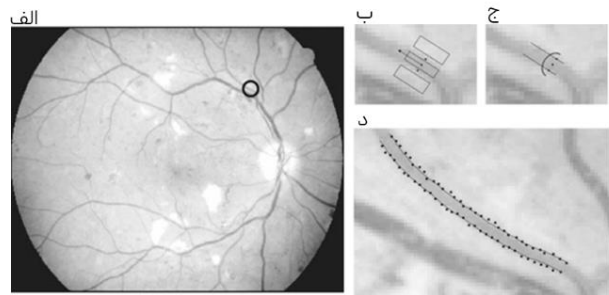
۳- روش پیشنهادی

در این بخش به توضیح معماری مدل و تابع زیان پرداخته می‌شود.

۱-۳ کلیات معماری مدل

شبکه U-Net به دلیل ساختار متقارن و توانایی استخراج ویژگی‌های چندمقیاسی، یکی از موفق‌ترین معماری‌ها در تقسیم‌بندی تصاویر پزشکی است [۲۵]. با این حال، نسخه کلاسیک U-Net در برابر چالش‌هایی مانند کنتراست پایین، نویز پس‌زمینه، تنوع مقیاس رگ‌های خونی و عدم تعادل کلاسی عملکرد محدودی دارد [۱۰]. این پژوهش معماری ATEN-UNet را پیشنهاد می‌دهد که با ادغام گره مرکزی چندمقیاسی، ماژول توجه تطبیقی ATEN^۴ و تابع زیان ترکیبی (Dice Loss + Binary Cross-Entropy) این محدودیت‌ها را برطرف می‌کند.

انتخاب نشود یا رگ‌ها در تصویر منقطع باشند، نتایج نادرستی تولید می‌کند [۲۱].



شکل ۲: روش‌های ردیابی

۳-۲ آستانه‌گذاری و پیش‌پردازش

پیش‌پردازش و آستانه‌گذاری به‌عنوان مراحل اولیه برای بهبود کیفیت تصاویر فوندوس و استخراج رگ‌های خونی استفاده می‌شوند. آستانه‌گذاری کلی (مانند روش اوتسو) برای بهینه‌سازی تفاوت بین کلاس‌های رگ و پس‌زمینه به کار می‌رود [۲۲]. در حالی که آستانه‌گذاری محلی در تصاویر با نور ناممکن عملکرد بهتری دارد [۲۳]. تکنیک‌های پیش‌پردازش مانند محدود سازی کنتراست^۱ با تنظیم کنتراست در نواحی کوچک تصویر، نویز را کاهش داده و رگ‌ها را برجسته می‌کند [۲۴]. سایر روش‌ها مانند فیلترهای گاوسی و عملیات مورفولوژیکی نیز برای حذف نویز و تقویت لبه‌ها استفاده می‌شوند.

۴-۲ یادگیری عمیق با شبکه U-Net

شبکه‌های عصبی کانولوشنی^۲، به‌ویژه معماری U-Net، با توانایی استخراج ویژگی‌های پیچیده و چندسطحی، تحول بزرگی در تقسیم‌بندی رگ‌های خونی ایجاد کرده‌اند. U-Net با ساختار متقارن شامل بخش رمزگذار (برای استخراج ویژگی‌ها) و رمزگشا (برای بازسازی تصویر) و اتصالات میانبر، دقت بالایی ارائه می‌دهد [۲۵]. مدل‌های مبتنی بر توجه مانند (DA-U)^۲Net با ترکیب توجه کانالی و فضایی، دقت شناسایی رگ‌های نازک را بهبود داده‌اند [۲۶]. همچنین، VGA-Net با مدل‌سازی ساختار گرافی رگ‌ها، تمرکز بر نواحی مرتبط را افزایش داده و عملکرد بهتری نسبت به U-Net پایه نشان داده است [۲۷].

۵-۲ روش‌های ترکیبی

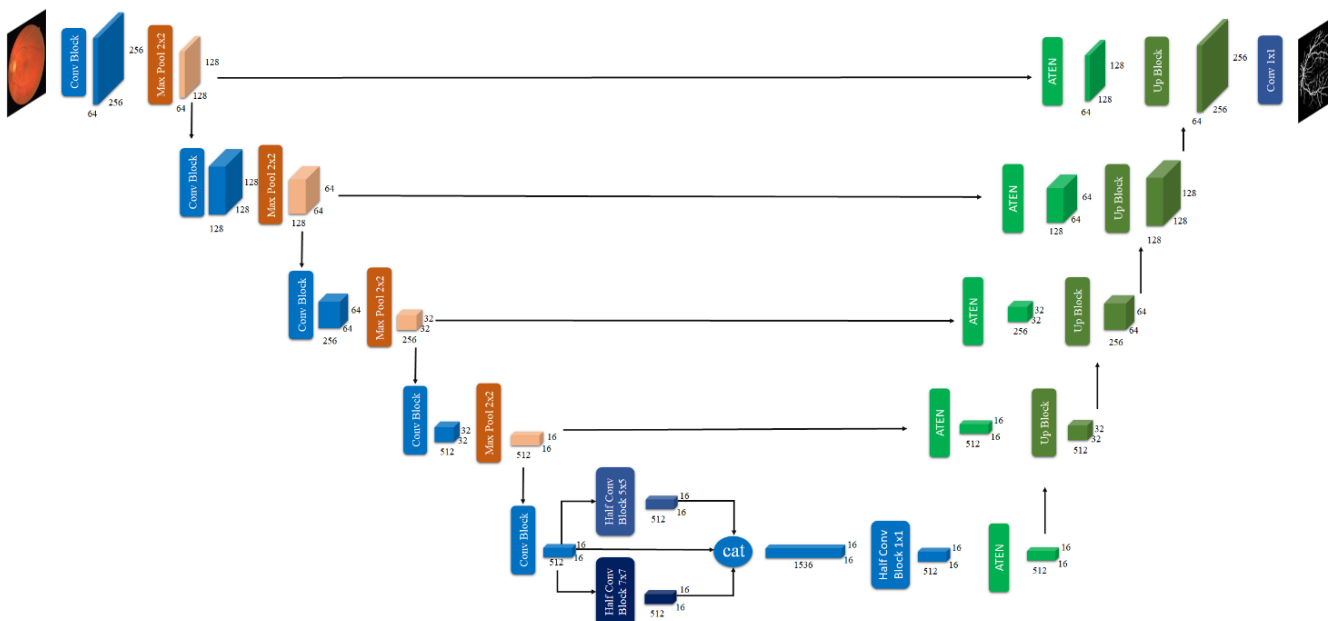
روش‌های ترکیبی برای حل عدم تعادل کلاسی، یکی از چالش‌های اصلی در تقسیم‌بندی رگ‌های خونی است. روش‌های ترکیبی با استفاده از توابع زیان پیشرفته مانند تابع زیان کانونی^۳ و تولید

3 Focal Loss

4Advanced TEchnology for Neural Attention

1 Contrast Limited Adaptive Histogram Equalization (CLAHE)

2 CNN



شکل ۴: مدل ATEN-UNet، بر پایه معماری U-Net، از بلوک‌های کانولوشنی در مسیر رمزگذاری برای استخراج ویژگی‌ها و از نمونه‌برداری افزایشی در مسیر رمزگشایی برای بازسازی نقشه‌ها استفاده می‌کند. ماژول‌های توجه ATEN در سطوح مختلف، تمرکز بر نواحی مهم را افزایش می‌دهند و اتصالات پرسی میان لایه‌های متناظر، اطلاعات مکانی و جزئیات ساختاری را حفظ می‌کنند.

روابط پیچیده می‌شود. رمزگذار از چهار مرحله‌ی اصلی تشکیل شده است. در مرحله‌ی نخست، دو بلوک کانولوشنی با 64×64 فیلتر برای استخراج لبه‌ها و بافت‌های محلی به‌کار گرفته می‌شوند و سپس عملیات تلفیق بیشینه^۱ با اندازه‌ی 2×2 ابعاد ویژگی‌ها را به 128×128 کاهش می‌دهد. در مرحله‌ی دوم، دو بلوک کانولوشنی با 128×128 فیلتر ویژگی‌های میانی را استخراج کرده و با انجام تلفیق بیشینه مشابه، ابعاد به 64×64 کاهش می‌یابد. در مرحله‌ی سوم، دو بلوک کانولوشنی با 256×256 فیلتر برای شناسایی ساختارهای شاخه‌ای و الگوهای پیچیده‌تر به‌کار می‌روند و مجدداً تلفیق بیشینه ابعاد را به 128×128 کاهش می‌دهد. در نهایت، در آخرین بلوک رمزگذار، دو بلوک کانولوشنی با 512×512 فیلتر جهت استخراج ویژگی‌های سطح بالا استفاده شده و با اعمال تلفیق بیشینه نهایی، اندازه‌ی ویژگی‌ها به 16×16 می‌رسد. اتصالات میانبر از هر مرحله به بخش رمزگشا منتقل می‌شوند تا اطلاعات مکانی حفظ شده و بازسازی دقیق‌تر ساختار رگ‌ها امکان‌پذیر شود. این طراحی موجب افزایش دقت در شناسایی رگ‌های خونی با مقیاس‌ها و الگوهای مختلف می‌شود. در طراحی معماری مدل، از بلوک‌های بلوک کانولوشنی استفاده شده است که ساختار آن در شکل (۵) نشان داده شده است. در این بلوک، ابتدا یک لایه‌ی کانولوشن 3×3 برای استخراج ویژگی‌های محلی از ورودی به‌کار گرفته می‌شود. سپس، یک نرمال‌ساز دسته‌ای جهت نرمال‌سازی مقادیر فعال‌سازی و تسریع فرایند یادگیری اعمال می‌گردد. پس از آن، از تابع فعال‌سازی Swish برای افزایش توانایی مدل در یادگیری روابط غیرخطی استفاده شده است. در ادامه، به‌منظور استخراج دقیق‌تر ویژگی‌ها و بهبود نمایش داده‌ها، مجدداً

این نوآوری‌ها دقت شناسایی رگ‌های نازک را افزایش داده و تعمیم‌پذیری مدل را بهبود می‌بخشند. شکل (۴) معماری کلی ATEN-UNet را نشان می‌دهد که در ادامه به بررسی دقیق‌تر هر بخش خواهیم پرداخت.

۳-۲ رمزگذار

رمزگذار ATEN-UNet وظیفه استخراج ویژگی‌های چندمقیاسی از تصاویر فوندوس را بر عهده دارد و شامل چهار بلوک کانولوشنی است. هر بلوک از دو لایه کانولوشن 3×3 ، نرمال‌سازی دسته‌ای^۱ و تابع فعال‌سازی Swish تشکیل شده است. تابع Swish، که در رابطه (۱) تعریف شده است.

$$f(x) = x \cdot \sigma(x) \quad (1)$$

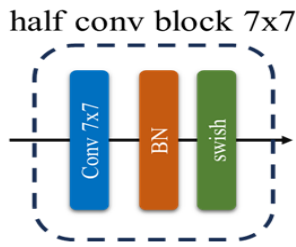
x ورودی تابع است که خروجی لایه کانولوشن پس از نرمال‌سازی دسته‌ای می‌باشد. $\sigma(x)$ تابع سیگموئید است که مقدار x را در بازه صفر و یک نگه می‌دارد.

این تابع به دلیل رفتار غیرخطی نرم و حفظ گرادیان‌های غیرصفر، جایگزین ReLU شده است [۲۹]. این ویژگی در تصاویر پزشکی با توزیع نامتوازن پیکسل‌ها (رگ‌های روشن در پس‌زمینه تاریک) از حذف ویژگی‌های ضعیف جلوگیری می‌کند.

در بخش رمزگذار شبکه، از بلوک‌های بلوک کانولوشنی برای استخراج ویژگی‌های تصویر در سطوح مختلف انتزاعی استفاده شده است. هر بلوک کانولوشنی شامل دو لایه‌ی متوالی از کانولوشن 3×3 ، نرمال‌سازی دسته‌ای و تابع فعال‌سازی Swish است. استفاده از Swish، به‌دلیل ماهیت پیوسته و غیرخطی آن، موجب بهبود جریان گرادیان، پایداری فرایند آموزش و افزایش توانایی مدل در یادگیری

² Max Pooling

¹ Batch Normalization

شکل ۷: بلوک 7×7 گره مرکزی

ادغام ویژگی‌ها: خروجی‌های سه مسیر با کانولوشن 1×1 (۵۱۲) فیلتر) فشرده شده و با ماژول ATEN تقویت می‌شوند. منظم‌سازی: لایه Dropout با نرخ 0.5 برای جلوگیری از بیش‌برازش. ماژول ATEN در این بخش ویژگی‌های کلیدی رگ‌ها را برجسته کرده و نویز پس‌زمینه را سرکوب می‌کند.

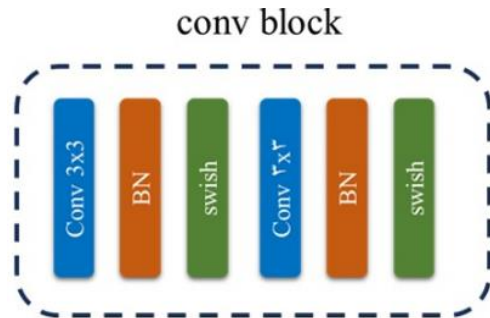
همان‌طور که در شکل‌های (۶) و (۷) مشخص است، در معماری مدل از بلوک‌هایی در گره مرکزی استفاده کرده‌ایم که شامل سه لایه در داخل بلاک‌های گره مرکزی هستند. محتوای این بلاک به این صورت است که ابتدا یک لایه کانولوشن برای استخراج ویژگی‌ها اعمال می‌شود. سپس یک نرمال‌ساز دسته‌ای برای نرمال‌سازی ویژگی‌ها به‌کارگرفته می‌شود. در نهایت، یک تابع فعال‌سازی Swish برای انتخاب تابع فعال‌سازی مناسب استفاده می‌شود. این ساختار باعث بهبود یادگیری و عملکرد مدل در استخراج ویژگی‌ها و بهینه‌سازی فرآیند آموزش دیده می‌شود.

۳-۴ رمزگشا

رمزگشا وظیفه بازسازی ویژگی‌های فشرده‌شده به ابعاد اصلی تصویر (256×256) را بر عهده دارد و از چهار بلوک متوالی تشکیل شده است. هر بلوک شامل بازنمونه‌برداری^۱ یک کانولوشن 3×3 ، ترکیب با ویژگی‌های میانبر متناظر از رمزگذار (که پیش‌تر توسط بلوک ATEN پردازش شده‌اند)، دو کانولوشن 3×3 و در نهایت یک بلوک ATEN است. در چهار بلوک رمزگشا، ابعاد ویژگی‌ها به ترتیب به 32×32 ، 64×64 ، 128×128 و 256×256 افزایش می‌یابد، و تعداد فیلترهای کانولوشن به ترتیب برابر با 512 ، 256 ، 128 و 64 است. ترکیب ویژگی‌های میانبر با خروجی هر بلوک موجب حفظ اطلاعات مکانی دقیق و بهبود شناسایی مرزها و الگوهای رگ‌ها می‌شود [۳۱].

توضیحات بلوک رمزگشا که در شکل (۸) قابل نمایش است. بدین صورت که این بلوک دارای دو ورودی اصلی است. اتصال عبوری و ویژگی‌های رمزگشا. اتصال عبوری به این صورت است که این ورودی از مرحله متناظر در رمزگذار (Encoder) آمده و معمولاً شامل اطلاعات مکانی با وضوح بالا و محلی است. ویژگی‌های رمزگشا به این صورت است که این ورودی از لایه قبلی مسیر رمزگشا آمده و معمولاً حاوی اطلاعات معنایی سطح بالا و وضوح پایین است.

یک لایه کانولوشن 3×3 دیگر به همراه نرمال‌ساز دسته‌ای و تابع فعال‌سازی Swish به صورت متوالی اعمال می‌شود. این ساختار با ترکیب مؤثر عملیات کانولوشن، نرمال‌سازی و فعال‌سازی، موجب بهبود پایداری آموزش و افزایش دقت مدل در استخراج ویژگی‌ها می‌گردد.



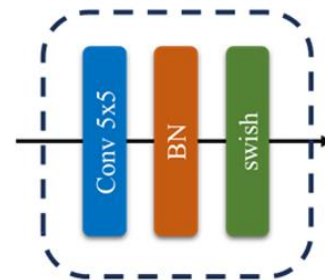
شکل ۵: بلوک کانولوشنی

۳-۳ گره مرکزی

گره مرکزی ATEN-UNet ویژگی‌های فشرده‌شده از رمزگذار را پردازش کرده و برای بازسازی دقیق در رمزگشا آماده می‌کند. برخلاف U-Net کلاسیک که تنها از کانولوشن‌های 3×3 استفاده می‌کند، گره مرکزی پیشنهادی از طراحی چندمقیاسی با کرنل‌های 3×3 ، 5×5 و 7×7 بهره می‌برد تا ویژگی‌های محلی و جهانی را به‌طور همزمان استخراج کند [۳۰]. این طراحی برای پوشش تنوع مقیاس رگ‌ها (از مویرگ‌های نازک تا رگ‌های ضخیم) حیاتی است.

جزئیات گره مرکزی به این صورت است یک لایه پایه با دو دو کانولوشن 3×3 با 512 فیلتر، نرمال‌سازی و Swish برای حفظ ویژگی‌های پایه و یک لایه متوسط با یک کانولوشن 5×5 با 512 فیلتر برای استخراج رگ‌های اصلی با میدان دید وسیع‌تر که در شکل (۶) آمده است.

half conv block 5x5

شکل ۶: بلوک 5×5 گره مرکزی

لایه پیشرفته: یک کانولوشن 7×7 با 512 فیلتر برای شناسایی مویرگ‌های نازک و شاخه‌دار. که جزئیات آن در شکل (۷) قابل نمایش است.

یکپارچه کرده و خروجی نهایی بلوک صعودی را تولید می‌کند که به مرحله بعدی رمزگشا منتقل می‌شود.

۳-۵ ماژول توجه ATEN

ماژول ATEN^۲ نوآوری کلیدی این پژوهش است که با ترکیب توجه کانالی و فضایی، ویژگی‌های مرتبط با رگ‌ها را تقویت کرده و نویز را سرکوب می‌کند. برخلاف مکانیزم‌های سنتی مانند CBAM یا Attention Gate از مسیرهای دوگانه (محلی و زمینه‌ای) و وزنه‌های تطبیقی استفاده می‌کند [۳۲].

توجه کانالی

در مرحله استخراج توصیف‌گرها، میانگین جهانی (GAP) و بیشینه جهانی (GMP) از ویژگی‌ها محاسبه می‌شوند تا به ترتیب توزیع کلی و ویژگی‌های برجسته نقشه‌های ویژگی ثبت شوند.

در مرحله تولید وزنه‌های تطبیقی، بردارهای حاصل از GAP و GMP از طریق یک لایه متراکم^۳ و تابع Softmax ترکیب می‌شوند تا ضرایب تطبیقی w_m و w_a تولید گردند. که وزن نسبی هر یک از توصیف‌گرها را مشخص می‌کنند و رابطه (۲) برقرار است.

$$w_a + w_m = 1 \quad (2)$$

سپس در مرحله تولید ماسک، توصیف‌گر ترکیبی مطابق با رابطه (۳) محاسبه می‌شود.

$$F_{fused} = GMP \cdot w_m + GAP \cdot w_a \quad (3)$$

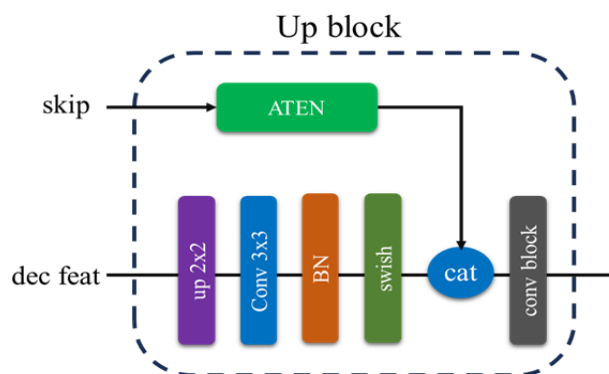
این توصیف‌گر ترکیبی توسط یک MLP دو لایه‌ای پردازش می‌شود؛ ابتدا ابعاد آن به $C/8$ کاهش یافته و تابع فعال‌سازی ReLU اعمال می‌گردد، سپس ابعاد به C بازگردانده شده و با استفاده از تابع Sigmoid نرمال‌سازی می‌شود. تا ماسک کانالی نهایی به دست آید. این ماسک اهمیت نسبی هر کانال ویژگی را مشخص می‌کند، به برجسته کردن ویژگی‌های کلیدی و سرکوب نویز پس‌زمینه کمک می‌کند. همچنین در شکل (۹) می‌توان تمامی کانال‌ها و مسیرهای ماژول توجه پیشنهادی را مشاهده کرد. در این شکل، نحوه‌ی جریان اطلاعات میان مسیرهای محلی و گسترده و نیز ادغام آن‌ها از طریق وزن تطبیقی به صورت بصری نمایش داده شده است، که درک بهتری از سازوکار درونی ماژول توجه فراهم می‌کند.

نقشه‌های فضایی

نقشه‌های فضایی: میانگین و بیشینه کانالی ویژگی‌ها به ماتریس [B, H, W, 2] تبدیل می‌شوند.

مسیرهای دوگانه

کانولوشن 3×3 برای ویژگی‌های محلی (مرزهای رگ‌های نازک) و کانولوشن 7×7 با نرخ گسترش^۴ برابر با ۲ برای ویژگی‌های گسترده (رگ‌های ضخیم).



شکل ۸: بلوک رمزگشا

مراحل پردازش درون بلوک صعودی در دو مسیر اصلی موازی رخ می‌دهد که سپس ادغام می‌شوند:

۱. مسیر ویژگی‌های رمزگشا

ورودی به صورت متوالی توسط یک سری لایه برای افزایش وضوح مکانی و بهبود ویژگی‌ها پردازش می‌شود. ابتدا لایه‌ی رمزگشا عملیات رمزگشایی را با استفاده از فیلترهای 2×2 انجام می‌دهد تا وضوح مکانی داده‌ها افزایش یابد. سپس لایه‌ی کانولوشن برای استخراج و بهبود ویژگی‌ها از داده‌های با وضوح بالاتر به کار گرفته می‌شود. پس از آن، لایه‌ی نرمال‌سازی دسته‌ای برای نرمال‌سازی فعال‌سازی‌ها اعمال می‌شود که موجب پایداری و تسریع فرآیند آموزش شبکه می‌گردد. در نهایت، از تابع فعال‌سازی Swish برای ایجاد ویژگی‌های غیرخطی و بهبود قابلیت نمایش شبکه استفاده می‌شود.

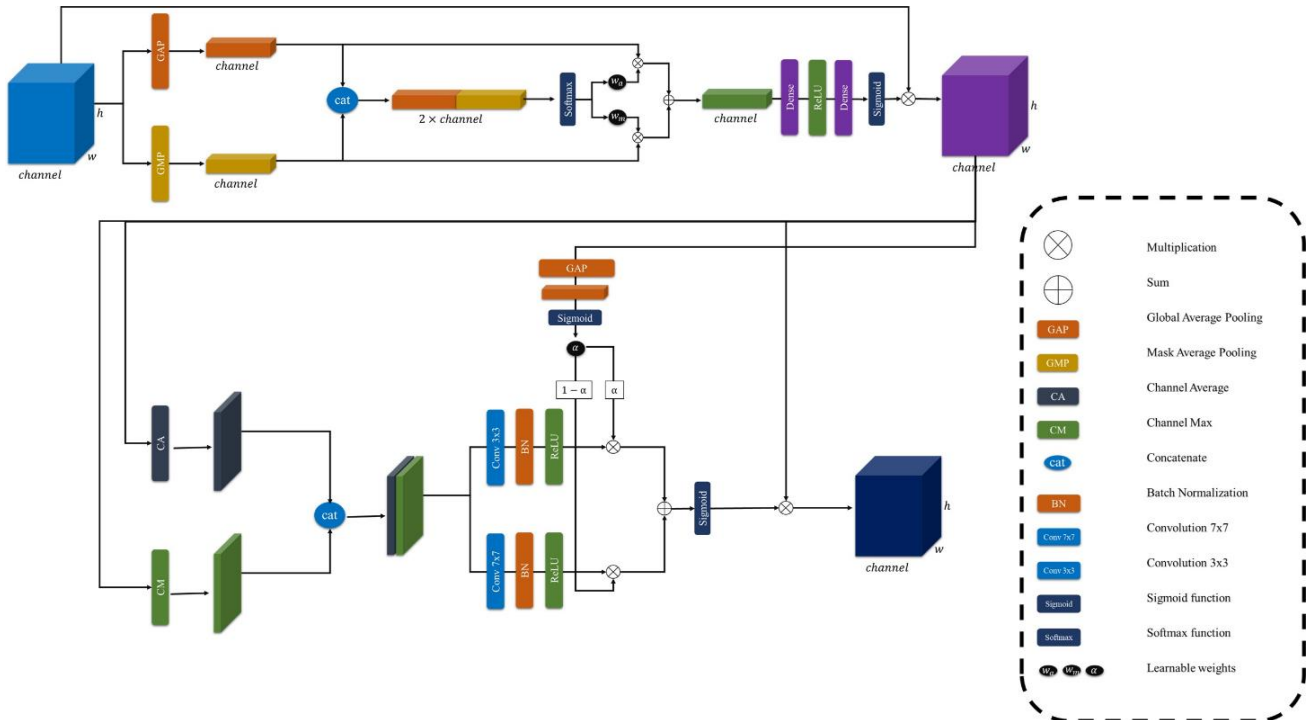
۲. مسیر اتصال عبوری

مکانیسم ATEN: ابتدا توسط مؤلفه‌ای با نام ATEN پردازش می‌شود. که مکانیزم توجه معرفی شده در این مقاله است. که برای وزن‌دهی انتخابی ویژگی‌ها از اتصال عبوری طراحی شده است تا قبل از ترکیب با مسیر رمزگشا، شبکه را در تمرکز بر ویژگی‌های مرتبط یاری دهد.

ترکیب ویژگی و خروجی نهایی

ادغام: خروجی لایه (ویژگی‌های رمزگشا) و خروجی بلوک ATEN (ویژگی‌های اتصال عبوری) با استفاده از عملیات ادغام^۱ با هم ترکیب می‌شوند. این عمل، ویژگی‌های معنایی سطح بالا را با ویژگی‌های با وضوح بالا ترکیب می‌کند.

بلوک کانولوشن نهایی: یک بلوک کانولوشن نهایی بر روی ویژگی‌های ادغام شده اعمال می‌شود. این بلوک اطلاعات ترکیبی را



شکل ۹: ساختار کلی ماژول توجه کانالی که با تجمع اطلاعات سراسری و اعمال مکانیزم توجه، اهمیت کانال‌های ویژگی را وزن‌دهی کرده و به تقویت ویژگی‌های معنادار و کاهش اطلاعات کم‌اهمیت کمک می‌کند

طراحی انعطاف‌پذیری لازم را برای تطبیق با ویژگی‌های خاص مجموعه داده‌های پزشکی فراهم می‌کند. زیان متقاطع دودویی یک معیار استاندارد برای مسائل طبقه‌بندی پیکسلی باینری است که در رابطه (۵) بیان شده است.

$$BCE = -\frac{1}{N} \sum_{i=1}^N [y_i \cdot \log(p_i) + (1 - y_i) \cdot \log(1 - p_i)] \quad (5)$$

در این رابطه، y_i برچسب واقعی پیکسل i -ام است که مقدار صفر برای پس‌زمینه و یک برای رگ دارد، و p_i بیانگر احتمال پیش‌بینی شده توسط مدل برای همان پیکسل است. N نیز تعداد کل پیکسل‌های تصویر را نشان می‌دهد.

برای اطمینان از پایداری محاسبات عددی، مقادیر پیش‌بینی شده پیش از اعمال تابع لگاریتم در یک بازه محدود (مانند $[\epsilon, 1 - \epsilon]$) قرار داده می‌شوند. این تابع زیان به مدل کمک می‌کند تا پیش‌بینی‌های پیکسلی دقیق‌تری انجام دهد؛ با این حال، به دلیل حساسیت کمتر آن به میزان هم‌پوشانی کلی بین نواحی پیش‌بینی و واقعی، به‌تنهایی برای تصاویر دارای عدم تعادل کلاس گزینه‌ی مناسبی نیست.

همچنین زیان دایس بر اساس ضریب دایس تعریف می‌شود که یک معیار هم‌پوشانی بین دو مجموعه است. این ضریب در رابطه (۶) بیان شده است.

$$Dice\ Coefficient = \frac{2 \cdot |Y \cap P| + smooth}{|Y| + |P| + smooth} \quad (6)$$

وزن تطبیقی: وزن $\alpha \in [0, 1]$ از طریق GAP و لایه متراکم با Sigmoid یادگیری می‌شود. اعمال ماسک: ماسک فضایی با ورودی ضرب شده و با اتصال باقی‌مانده ترکیب می‌شود. ATEN در اتصالات میانبر و گره مرکزی، شاخص‌های دایس و جاکارد را ۵٪ الی ۱۰٪ نسبت به U-Net پایه بهبود می‌بخشد.

۳-۶ تابع زیان

به‌منظور بهینه‌سازی معماری پیشنهادی در وظیفه‌ی تقسیم‌بندی رگ‌های خونی در تصاویر فوندوس شبکیه، یک تابع زیان ترکیبی طراحی شده است. که با ادغام زیان دودویی^۱ و زیان دایس^۲ به‌صورت وزن‌دار، عملکرد مدل را در برابر چالش‌های خاص این حوزه، نظیر عدم تعادل کلاس‌ها و نیاز به هم‌پوشانی دقیق، بهبود می‌بخشد. این تابع زیان به‌گونه‌ای تدوین شده است که تعادل مناسبی بین دقت پیکسلی و تطابق کلی بین ماسک پیش‌بینی شده و ماسک واقعی برقرار کند، که برای شناسایی رگ‌های نازک و ضخیم در تصاویر با کنتراست پایین حیاتی است.

تابع زیان ترکیبی به‌صورت یک ترکیب خطی از دو جزء اصلی تعریف می‌شود در رابطه (۴) بیان شده است.

$$Combo\ Loss = \alpha \cdot BCE + (1 - \alpha) \cdot Dice\ Loss \quad (4)$$

در این معادله، α یک ضریب وزن‌دهی است که به‌صورت تجربی به مقدار ۰/۳ تنظیم شده است تا تعادل بین دو جزء زیان برقرار شود. یک ثابت کوچک به نام "smooth" نیز در محاسبات گنجانده شده است تا از مشکلات عددی مانند تقسیم بر صفر جلوگیری شود. این

۴-۳ معیارهای ارزیابی

برای ارزیابی عملکرد مدل، از چند معیار رایج استفاده شده است که میزان تطابق خروجی مدل با برجسب‌های واقعی را نشان می‌دهند. این معیارها شامل دقت (Accuracy)، ضریب دایس (Dice) و شاخص همپوشانی (IoU) هستند.

دقت نشان می‌دهد چه نسبتی از کل پیکسل‌ها به درستی توسط مدل طبقه‌بندی شده‌اند. این معیار با در نظر گرفتن پیش‌بینی‌های صحیح برای هر دو کلاس عروق و پس‌زمینه محاسبه می‌شود و نحوه محاسبه آن در رابطه (۷) آورده شده است.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

ضریب دایس میزان همپوشانی بین ناحیه پیش‌بینی شده و ناحیه واقعی عروق را اندازه‌گیری می‌کند و نسبت به خطاهای ناشی از عدم تعادل کلاس حساسیت بالاتری دارد. این معیار به‌ویژه در مسئله بخش‌بندی عروق شبکیه اهمیت زیادی دارد. نحوه محاسبه دایس در رابطه (۸) بیان شده است.

$$Dice = \frac{|Prediction \cap GroundTruth| \times 2}{|Prediction| + |GroundTruth|} \quad (8)$$

این دو معیار میزان همپوشانی بین ناحیه پیش‌بینی شده و ناحیه واقعی را اندازه می‌گیرند.

شاخص همپوشانی یا نسبت اجتماع به اشتراک، معیار دیگری برای سنجش میزان همپوشانی بین خروجی مدل و برجسب واقعی است که سخت‌گیرانه‌تر از دایس عمل می‌کند. نحوه محاسبه IoU در رابطه (۹) ارائه شده است.

$$IoU = \frac{|Prediction \cap GroundTruth|}{|Prediction \cup GroundTruth|} \quad (9)$$

۴-۴ مقایسه کارهای قبلی

در این بخش، عملکرد مدل پیشنهادی ATEN-UNet با مدل‌های پایه و پیشرفته در حوزه تقسیم‌بندی رگ‌های خونی شبکیه مقایسه می‌شود. این مقایسه بر اساس معیارهای ارزیابی استاندارد شامل دقت، شاخص همپوشانی جاکارد و ضریب دایس انجام گرفته است. ارزیابی‌ها بر روی مجموعه داده Retina Blood Vessel صورت پذیرفته که شامل تصاویر شبکیه چشم برای تقسیم‌بندی عروق خونی است. برای دستیابی به یکنواختی در مقایسه، نتایج مدل‌های پایه از مقالات مرتبط استخراج شده و همه مدل‌ها بر روی دادگان مشابه ارزیابی شده‌اند. هدف اصلی این مقایسه، برجسته کردن برتری ATEN-UNet در شناسایی دقیق رگ‌های نازک، مدیریت نویز تصاویر و بهبود همپوشانی بین مناطق پیش‌بینی شده و واقعی است. این مدل با تمرکز بر معماری پیشرفته و تکنیک‌های بهینه‌سازی، توانسته است چالش‌های موجود در مدل‌های قبلی را کاهش دهد و عملکرد کلی را ارتقا بخشد.

که در آن $Y \cap P$ تعداد پیکسل‌های مشترک بین ماسک واقعی (Y) و ماسک پیش‌بینی شده (P) و $|Y|$ و $|P|$ به ترتیب تعداد پیکسل‌های مثبت در ماسک واقعی و پیش‌بینی شده هستند.

این معیار به‌طور خاص برای بهینه‌سازی همپوشانی بین پیش‌بینی و واقعیت طراحی شده و در شناسایی ساختارهای نازک مانند رگ‌های خونی، که بخش کوچکی از تصویر را اشغال می‌کنند، بسیار مؤثر است. استفاده از ثابت smooth ثبات محاسبات را در شرایط کم‌پیکسل تضمین می‌کند.

۴-۴ پیاده‌سازی و نتایج

۴-۱ دادگان

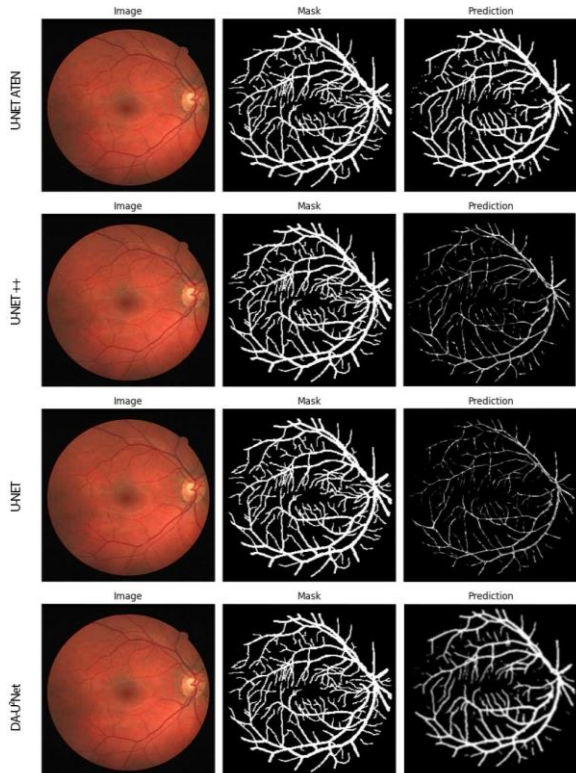
این پژوهش از مجموعه داده عمومی "Retina Blood Vessel Segmentation" موجود در پلتفرم Kaggle استفاده کرده است [۳۳]. این مجموعه شامل تصاویر فوندوس شبکیه و ماسک‌های باینری متناظر برای بخش‌بندی رگ‌های خونی است.

مطابق با رویه رایج در مطالعات پیشین، داده‌ها به دو بخش آموزش (۸۰٪) و آزمون (۲۰٪) تقسیم شده‌اند. آموزش مدل‌ها تنها بر روی مجموعه آموزش انجام شده و ارزیابی عملکرد بر اساس مجموعه آزمون صورت گرفته است. رزولوشن اصلی تصاویر متغیر است، اما در مرحله پیش‌پردازش به 256×256 پیکسل تغییر اندازه داده شده‌اند تا ورودی یکنواختی برای مدل ATEN-UNet فراهم شود.

مجموعه داده شامل تصاویر فوندوس از بیماران با شرایط متنوع (مانند افراد سالم و بیماران مبتلا به رتینوپاتی دیابتی) است که تنوع کافی در ساختار عروق، شرایط نوری و کیفیت تصویر را تضمین می‌کند. این ویژگی، تعمیم‌پذیری مدل در برابر چالش‌هایی مانند نویز و کنتراست پایین را تقویت می‌کند. برای مقایسه، دادگان استاندارد مانند DRIVE (شامل ۴۰ تصویر) و STARE (شامل ۲۰ تصویر) در حوزه تقسیم‌بندی رگ‌های خونی پرکاربرد هستند، اما دادگان Kaggle به دلیل حجم بیشتر و تنوع بالاتر انتخاب شد [۳۴].

۴-۲ جزئیات پیاده‌سازی

در تمامی آزمایش‌ها، تصاویر و ماسک‌ها به ابعاد 256×256 پیکسل تغییر اندازه داده شدند و تصاویر خاکستری پس از اعمال پیش‌پردازش همسان‌سازی تطبیقی هیستوگرام با محدودسازی کنتراست نرمال‌سازی شدند. مدل با اندازه دسته ۸، بهینه‌ساز Adam و نرخ یادگیری اولیه 10^{-4} حداکثر تا ۲۰۰ تکرار آموزش داده شد. تابع هزینه ترکیبی شامل آنتروپی متقاطع دودویی و تابع هزینه دایس مورد استفاده قرار گرفت. برای کنترل بیش‌برازش از توقف زودهنگام آموزش و کاهش نرخ یادگیری هنگام توقف بهبود استفاده شد.



شکل ۱۰: مقایسه کیفی نتایج بخش بندی عروق شبکیه برای مدل‌های مختلف. مدل پیشنهادی با حفظ بهتر ساختار عروق و کاهش نویز، عملکرد بهتری نسبت به سایر روش‌ها نشان می‌دهد.



شکل ۱۱: نتایج بزرگ‌نمایی شده، توانمندی مدل پیشنهادی در حفظ پیوستگی و تفکیک دقیق شاخه‌های باریک عروقی و شناسایی جزئیات ظریف را نشان می‌دهد.

۴-۵ مطالعات فرسایشی

در این مطالعه، برای بررسی اثر مکانیزم‌های توجه کانالی و فضایی، چهار مدل مختلف مبتنی بر U-Net طراحی و ارزیابی شدند. مدل اول (M1) نسخه پایه U-Net بدون استفاده از توجه کانالی و فضایی است. مدل دوم (M2) از توجه کانالی استفاده می‌کند، در حالی که مدل سوم (M3) به توجه فضایی مجهز شده است. در نهایت، مدل چهارم (M4) که به‌عنوان مدل پیشنهادی معرفی می‌شود، هر دو مکانیزم توجه کانالی و فضایی را به‌صورت هم‌زمان به‌کار می‌گیرد. نتایج کمی این مدل‌ها در جدول (۲) ارائه شده است.

جدول ۱: مقایسه عملکرد بخش‌بندی روش پیشنهادی با روش‌های پیشین بر روی دادگان Retina Blood Vessel

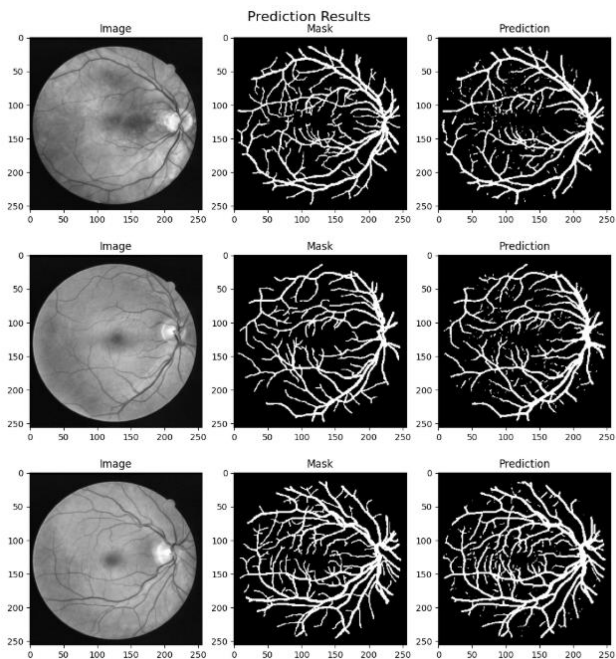
Accuracy	Dice	IoU	مدل
۰/۹۵۱۱	۰/۸۱۹۴	۰/۶۹۴۰	U-Net
۰/۹۲۶۱	۰/۶۵۳۳	۰/۴۸۵۱	U-Net با معماری VGG16
۰/۹۳۲۳	۰/۶۵۶۸	۰/۴۸۹۰	U-Net با معماری ResNet50
۰/۹۵۱۷	۰/۸۲۲۱	۰/۶۹۸۰	U-Net++
۰/۹۵۲۰	۰/۸۲۳۶	۰/۷۰۰۱	DA-U ² Net
۰/۹۵۱۸	۰/۸۲۱۳	۰/۶۹۶۸	VGA-Net
۰/۹۵۲۳	۰/۸۲۳۷	۰/۷۰۰۳	روش پیشنهادی

جدول (۱) مقایسه عملکرد بخش‌بندی روش پیشنهادی با روش‌های پیشین بر روی دادگان Retina Blood Vessel را نشان می‌دهد. در این جدول، مدل‌های مورد بررسی شامل U-Net پایه، معماری پیش‌آموزش‌دیده VGG16 و ResNet50، U-Net++ و همچنین معماری‌هایی بر پایه مکانیزم توجه شامل DA-U²Net و VGA-Net با روش پیشنهادی مقایسه شده‌اند. همان‌طور که در جدول (۱) نشان داده شده است، مدل پیشنهادی در مقایسه با سایر روش‌ها، بالاترین مقدار IoU و مقدار دایس را به‌دست آورده است که بیانگر دقت بالاتر در استخراج ساختار عروق شبکیه است. در مقابل، مدل‌های پایه از جمله U-Net و نسخه‌های مبتنی بر VGG16 و ResNet50 عملکرد ضعیف‌تری نسبت به مابقی روش‌ها داشته و مقادیر IoU آن‌ها کمتر از ۰/۷۰ گزارش شده است. اگرچه U-Net++ و DA-U²Net بهبود نسبی نسبت به U-Net را نشان می‌دهند، اما همچنان از نظر حفظ پیوستگی عروق و تشخیص شاخه‌های نازک، دقت پایین‌تری نسبت به مدل پیشنهادی دارند. در شکل (۱۰) نتایج عملکرد مدل‌های مورد بررسی به‌صورت بصری ارائه شده است. این شکل نشان می‌دهد که مدل پیشنهادی در مقایسه با سایر مدل‌ها، توانایی بالاتری در استخراج دقیق رگ‌های خونی شبکیه دارد. این برتری به‌ویژه در حفظ پیوستگی عروق و تشخیص شاخه‌های باریک‌تر بوده و در شکل (۱۰) و شکل (۱۱) به‌صورت واضح‌تری قابل مشاهده است.

تمامی نتایج بصری ارائه‌شده در این شکل‌ها مربوط به دادگان فوندوس شبکیه Retina Blood Vessel هستند. با توجه به اینکه روش‌های مختلف ممکن است از پیش‌پردازش‌های متفاوتی برای بهبود کیفیت تصاویر استفاده کنند، برای ایجاد یک مقایسه منصفانه، تمامی مدل‌ها تحت شرایط یکسان پیش‌پردازش ارزیابی شده‌اند. بدین ترتیب، تفاوت‌های مشاهده‌شده در خروجی‌ها صرفاً ناشی از ساختار و عملکرد مدل‌ها بوده و تأثیر روش‌های پیش‌پردازش در نتایج حذف شده است.

نتایج کیفی

تحلیل کیفی با مقایسه بصری ماسک‌های پیش‌بینی‌شده توسط ATEN-UNet با ماسک‌های واقعی^۱ از دادگان Kaggle انجام شد. شکل (۱۲) نمونه‌هایی از تصاویر فوندوس، ماسک‌های واقعی و ماسک‌های پیش‌بینی‌شده را نشان می‌دهد. در شکل (۱۲) مدل پیشنهادی توانایی بالایی در شناسایی رگ‌های اصلی و نازک به‌ویژه در نواحی با کنتراست نشان می‌دهد. با این حال، در نواحی با کنتراست پایین یا تصاویر پرنویز، ناپیوستگی‌های جزئی در رگ‌های نازک مشاهده شد که می‌تواند به دلیل محدودیت‌های دادگان یا پیچیدگی‌های نوری باشد. شکل (۱۲) مقایسه بصری ماسک‌های پیش‌بینی‌شده و واقعی اکتباس از دادگان می‌باشد.



شکل ۱۲: نتایج بصری

برای بهبود این محدودیت‌ها، پیشنهاد می‌شود از دادگان متنوع‌تر مانند DRIVE و STARE استفاده شود یا تکنیک‌های پیش‌پردازش پیشرفته‌تر مانند فیلترهای denoising اعمال گردد. این تحلیل کیفی تأیید می‌کند که ATEN-UNet در مقایسه با مدل‌های پایه، جزئیات ساختاری رگ‌ها را با دقت بیشتری بازسازی می‌کند، که برای تشخیص بیماری‌های چشمی حیاتی است.

۵- جمع‌بندی و نتیجه‌گیری

این پژوهش معماری ATEN-UNet را برای تقسیم‌بندی خودکار رگ‌های خونی در تصاویر فوندوس شبکیه ارائه کرد. این مدل با استفاده از گره چندمقیاسی، مازول توجه ATEN و تابع زیان ترکیبی، چالش‌هایی مانند شناسایی رگ‌های نازک و نویز پس‌زمینه را بهبود بخشید. همان‌طور که در جدول (۲) مشخص است ATEN-UNet به ضریب دایس ۰/۸۲۳۷، شاخص همپوشانی ۰/۷۰۰۳ و دقت ۰/۹۵۲۳ دست یافت، که نسبت به مدل‌های دیگر

جدول ۲: نتایج مطالعه فرسایشی برای بررسی تأثیر توجه کانالی و فضای در مدل پیشنهادی

Accuracy	Dice	IoU	توجه فضایی	توجه کانالی	U-Net	مدل
۰/۹۴۶۰	۰/۸۱۰۹	۰/۶۸۲۰	×	×	✓	M1
۰/۹۴۵۲	۰/۸۱۰۰	۰/۶۸۰۷	×	✓	✓	M2
۰/۹۴۸۵	۰/۸۲۲۴	۰/۶۹۸۴	✓	×	✓	M3
۰/۹۵۲۳	۰/۸۲۳۷	۰/۷۰۰۳	✓	✓	✓	M4

بر اساس نتایج، مدل M2 که تنها از توجه کانالی استفاده می‌کند، بهبود بسیار محدودی را نشان می‌دهد و عملکرد آن تفاوت چشمگیری با مدل پایه ندارد. در مقابل، مدل M3 با استفاده از توجه فضایی عملکرد بهتری ارائه می‌دهد، به‌ویژه در تشخیص نواحی با کنتراست پایین که نقش مهمی در تفکیک دقیق ساختارها دارند. بهترین عملکرد، مربوط به مدل M4 است که با ترکیب توجه کانالی و فضای توانسته است شناسایی رگ‌های نازک را بهبود داده و نویز را کاهش دهد. این مدل بالاترین مقادیر دقت، دایس و IoU را در میان همه مدل‌ها به دست آورده است. به‌طور کلی، نتایج این مطالعه فرسایشی نشان می‌دهد که استفاده هم‌زمان از توجه کانالی و فضای تأثیر قابل توجهی در بهبود عملکرد مدل U-Net در پردازش تصاویر شبکیه دارد.

۴-۶ نتایج کمی و کیفی

عملکرد مدل پیشنهادی ATEN-UNet بر روی مجموعه داده Retina Blood Vessel Segmentation با استفاده از معیارهای دقت، شاخص همپوشانی ضریب دایس و زیان ارزیابی شد. این معیارها به دلیل اهمیت آن‌ها در سنجش دقت پیکسلی و همپوشانی در مسائل تقسیم‌بندی تصاویر پزشکی انتخاب شدند. برای آموزش مدل، از بهینه‌ساز Adam با نرخ یادگیری ۰/۰۰۱ و ۱۰۰ epoch استفاده شد، و داده‌ها با نسبت ۸۰:۲۰ (آموزشی:آزمایشی) تقسیم شدند.

نتایج کمی

عملکرد مدل پیشنهادی ATEN-UNet با استفاده از معیارهای استاندارد دقت، ضریب دایس، شاخص جاکارد و تابع زیان بر روی مجموعه داده‌های آموزشی و اعتبارسنجی مجموعه Retina Blood Vessel Segmentation ارزیابی شده است. این معیارها توانایی مدل در تقسیم‌بندی دقیق رگ‌های خونی شبکیه، به‌ویژه در شناسایی لبه‌های نازک و مدیریت نویز، را به‌خوبی نشان می‌دهند.

همان‌طور که در شکل (۱۰) و شکل (۱۱) دیده می‌شود، لبه‌های نازک رگ‌ها را به‌خوبی تشخیص می‌دهد. برای بهبود، داده‌افزایی پیشرفته‌تر یا افزایش حجم داده پیشنهاد می‌شود. این نتایج، پتانسیل بالای مدل را برای کاربردهای پزشکی نشان می‌دهد.

- [11] Fateh, A., et al., *Providing a voting-based method for combining deep neural network outputs to layout analysis of printed documents*. Journal of Machine Vision and Image Processing, 2022. **9**(1): p. 47–64.
- [12] Oktay, O., et al., *Attention u-net: Learning where to look for the pancreas*. arXiv preprint arXiv:1804.03999, 2018.
- [13] Saber, A., et al., *Efficient and accurate pneumonia detection using a novel multi-scale transformer approach*. Sensors, 2025. **25**(23): p. 7233.
- [14] Fateh, A., M.R. Mohammadi, and M.R.J. Motlagh, *Msdnet: Multi-scale decoder for few-shot semantic segmentation via transformer-guided prototyping*. Image and Vision Computing, 2025: p. 105672.
- [15] Cheung, C.Y.-I., et al., *Retinal vascular geometry and 6 year incidence and progression of diabetic retinopathy*. Diabetologia, 2017. **60**:(9)p. 1770–1781.
- [16] Fraz, M.M., et al., *Blood vessel segmentation methodologies in retinal images—a survey*. Computer methods and programs in biomedicine, 2012. **108**(1): p. 407–433.
- [17] Hoover, A., V. Kouznetsova, and M. Goldbaum, *Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response*. IEEE Transactions on Medical imaging, 2000. **19**(3): p. 203–210.
- [18] Soares, J.V., et al., *Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification*. IEEE Transactions on medical Imaging, 2006. **25**(9): p. 1214–1222.
- [19] Rahmani, A.I., H. Akbari, and S.S. Esmaili, *Retinal blood vessel segmentation using gabor filter and morphological reconstruction*. Signal Processing and Renewable Energy (SPRE), 2020. **4**:(1)p. 77–88.
- [20] Nanmaran, R., et al. *Image Enhancement in Fundus Images using Modified Frangi Vesselness Filter in comparison with Hessian and Gaussian Filter*. in *2024 2nd International Conference on Networking and Communications (ICNWC)*. 2024. IEEE.
- [21] Yin, Y., M. Adel, and S. Bourennane, *Retinal vessel segmentation using a probabilistic tracking method*. Pattern Recognition, 2012. **45**(4): p. 1235–1244.
- [22] Lamrani, D., et al., *Enhanced Retinal Blood Vessels Segmentation using Deep Learning and Residual Network*. Intelligence-Based Medicine, 2025: p. 100263.
- [23] Ghislain, F., S.T. Beaudelaire, and T. Daniel, *An improved semi-supervised segmentation of the retinal vasculature using curvelet-based contrast adjustment and generalized linear model*. Heliyon, 2024.(18)10 .
- [24] Halder, N., et al. *Fusion-Based Segmentation and Classification of Novel Coronavirus or Pneumonia from Chest X-ray Images Using Machine Learning Techniques*. in *2022 International Interdisciplinary*

برتری دارد. تحلیل کیفی نیز توانایی مدل در شناسایی رگ‌های نازک را تأیید کرد، هرچند در نواحی کم‌کنتراست ناپیوستگی‌های جزئی مشاهده شد.

برای تحقیقات آینده، استفاده از مجموعه داده‌های متنوع‌تر و تکنیک‌های پیش‌پردازش پیشرفته‌تر پیشنهاد می‌شود تا دقت در نواحی چالش‌برانگیز افزایش یابد. ATEN-UNet پتانسیل بالایی برای تشخیص زودهنگام بیماری‌های چشمی مانند رتینوپاتی دیابتی دارد و گامی مهم در تحلیل خودکار تصاویر پزشکی است.

۶- مراجع

- [1] Fateh, A., et al., *BRISC: Annotated Dataset for Brain Tumor Segmentation and Classification with Swin-HAFNet*. arXiv preprint arXiv:2506.14318, 2025.
- [2] Radha, K., et al., *Unfolded deep kernel estimation-attention UNet-based retinal image segmentation*. Scientific Reports, 2023. **13**(1): p. 20712.
- [3] Guo, C., et al. *Sa-unet: Spatial attention u-net for retinal vessel segmentation*. in *2020 25th international conference on pattern recognition (ICPR)*. 2021. IEEE.
- [4] Gulshan, V., et al., *Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs*. jama, 2016. **316**(22): p. 2402–2410.
- [5] Abdulsahib, A.A., et al., *Comprehensive review of retinal blood vessel segmentation and classification techniques :intelligent solutions for green computing in medical images, current challenges, open issues, and knowledge gaps in fundus medical images*. Network Modeling Analysis in Health Informatics and Bioinformatics, 2021. **10**(1): p. 20.
- [6] Kumar, K., D. Samal, and Suraj, *Automated retinal vessel segmentation based on morphological preprocessing and 2D-Gabor wavelets*, in *Advanced Computing and Intelligent Engineering: Proceedings of ICACIE 2018, Volume 1*. 2020, Springer. p. 411–423.
- [7] Sidhu, R., J. Sachdeva, and D .Katoch, *Segmentation of retinal blood vessels by a novel hybrid technique-Principal Component Analysis (PCA) and Contrast Limited Adaptive Histogram Equalization (CLAHE)*. Microvascular Research, 2023. **148**: p. 104477.
- [8] Zhu, Y.-f., et al., *CCS-UNet: a cross-channel spatial attention model for accurate retinal vessel segmentation*. Biomedical Optics Express, 2023. **14**(9): p. 4739–4758.
- [9] Tong, L., et al., *LiViT-Net: A U-Net-like, lightweight Transformer network for retinal vessel segmentation*. Computational and Structural Biotechnology Journal, 2024. **24**: p. 213–224.
- [10] Siddique, N., et al., *U-Net and its variants for medical image segmentation: theory and applications*. arXiv preprint arXiv:2011.01118, 2020.



دکتر منصور فاتح مدرک کارشناسی خود را در رشته مهندسی برق از دانشگاه صنعتی شاهرود در سال ۱۳۸۶ دریافت کرد. سپس کارشناسی ارشد و دکتری خود را در رشته‌های مهندسی پزشکی و الکترونیک دیجیتال در سال‌های ۱۳۸۸ و ۱۳۹۳ از دانشگاه

تربیت مدرس دریافت کرد. پروژه کارشناسی ارشد خود را با عنوان "بررسی نقش و اثر نور پلاریزه در درماتوسکپی از بدن با استفاده از شبیه سازی" و پروژه دکتری خود را با عنوان "خواندن خودکار نقشه‌های دستی فرش" به انجام رسانید. از سال ۱۳۹۴ ایشان عضو هیئت علمی دانشگاه صنعتی شاهرود بوده و هم اکنون رئیس دانشکده کامپیوتر و برق دانشگاه صنعتی شاهرود است زمینه تحقیقاتی ایشان پردازش تصویر و یادگیری ماشین می‌باشد.



دکتر محسن رضوانی مدرک دکتری خود را از دانشگاه UNSW استرالیا در گرایش شبکه و امنیت در سال ۱۳۹۴ دریافت نموده‌اند. ایشان مدرک کارشناسی ارشد خود را در گرایش مهندسی نرم افزار از دانشگاه صنعتی شریف، و مدرک کارشناسی خود را در

همین گرایش از دانشگاه صنعتی امیرکبیر اخذ نموده اند. دکتر رضوانی پس از اخذ مدرک دکتری به مدت یک سال دوره Postdoc را در دانشکده مهندسی و علوم کامپیوتر دانشگاه UNSW استرالیا گذرانده و پس از آن در دانشگاه صنعتی شاهرود مشغول به کار شده اند. در حال حاضر ایشان رئیس مرکز تخصصی آپا دانشگاه صنعتی شاهرود و همچنین مدیر آزمایشگاه امنیت داده و شبکه در دانشکده مهندسی کامپیوتر است.



امیررضا فاتح مدرک کارشناسی و کارشناسی ارشد خود را در رشته مهندسی کامپیوتر از دانشگاه صنعتی شاهرود دریافت کرد. ایشان در حال حاضر دانشجوی دکتری رشته هوش مصنوعی و رباتیک در دانشگاه علم و صنعت ایران می‌باشند.

Conference on Mathematics, Engineering and Science (MESIICON). 2022. IEEE.

- [25] Ronneberger, O., P. Fischer, and T. Brox. *U-net: Convolutional networks for biomedical image segmentation*. in *International Conference on Medical image computing and computer-assisted intervention*. 2015. Springer.
- [26] Chu, B., et al., *(DA-U) 2Net: double attention U2Net for retinal vessel segmentation*. *BMC ophthalmology*, 2025. **25**(1): p. 86.
- [27] Jalali, Y., M. Fateh, and M. Rezvani, *VGA-Net: Vessel graph based attentional U-Net for retinal vessel segmentation*. *IET Image Processing*, 2024. **18**(8): p. 2191-2213.
- [28] Lin, T.-Y., et al. *Focal loss for dense object detection*. in *Proceedings of the IEEE international conference on computer vision*. 2017.
- [29] Ramachandran, P., B. Zoph, and Q.V. Le, *Searching for activation functions*. arXiv preprint arXiv:1710.05941, 2017.
- [30] Ni, Z.-L., et al., *SurgiNet: Pyramid attention aggregation and class-wise self-distillation for surgical instrument segmentation*. *Medical Image Analysis*, 2022. **76**: p. 102310.
- [31] Woo, S., et al. *Cbam: Convolutional block attention module*. in *Proceedings of the European conference on computer vision (ECCV)*. 2018.
- [32] Xie, Y., et al., *Attention mechanisms in medical image segmentation: A survey*. arXiv preprint arXiv:2305.17937, 2023.
- [33] Kumar, K.S. and N.P. Singh, *Retinal disease prediction through blood vessel segmentation and classification using ensemble-based deep learning approaches*. *Neural Computing and Applications*, 2023. **35**(17): p. 12495-12511.
- [34] Staal, J., et al., *Ridge-based vessel segmentation in color images of the retina*. *IEEE transactions on medical imaging*, 2004. **23**(4): p. 501-509.



سید علیرضا خاتمی مدرک کارشناسی خود را در رشته مهندسی نرم افزار در دانشگاه فنی و حرفه ای سمنان با معدل ۱۸ و کسب رتبه الف در سال ۱۴۰۰ دریافت کرد. ایشان در همان سال در مقطع کارشناسی ارشد در رشته مهندسی کامپیوتر و گرایش هوش مصنوعی و

رباتیک در دانشگاه صنعتی شاهرود مشغول به تحصیل شد. پروژه کارشناسی ارشد ایشان "بخش‌بندی تصاویر فوندوس شبکیه چشم به کمک یادگیری عمیق" است. زمینه تحقیقاتی ایشان پردازش تصویر، یادگیری ماشین و یادگیری عمیق است.