

Research Article

Enhancing the Detection Accuracy of Mild and Proliferative Diabetic Retinopathy Using Augmented Retinal Images

Babak Esmaeilpour Quchani *¹ 

¹ Department of Computer and Artificial Intelligence, Quchan, Islamic Azad University, Quchan, Iran,

Babak.esmaeilpour@iau.ac.ir

Corresponding Author

* Babak Esmaeilpour Quchani
Department of Computer and Artificial Intelligence, Quchan, Islamic Azad University, Quchan, Iran,

Babak.esmaeilpour@iau.ac.ir

Main Subjects: Improving Diagnostic Accuracy Using Retinal Images

Received: 13 November 2025

Revised: 19 December 2025

Accepted: 20 December 2025

<https://doi.org/10.82195/ntds.2025.1224526>

Abstract

Diabetic retinopathy is one of the most prevalent and serious chronic complications of diabetes mellitus and may lead to progressive vision impairment and irreversible blindness if not diagnosed early. Accurate detection and grading of disease severity, particularly in mild and advanced stages, are essential for effective clinical management and timely intervention. Nevertheless, the performance of intelligent diagnostic systems is often limited by challenges such as data imbalance and insufficient samples in certain severity classes. This study proposes a machine learning-based framework to improve the detection and severity classification of diabetic retinopathy using retinal fundus images. Initially, preprocessing techniques including normalization and contrast limited adaptive histogram equalization applied to the green channel were employed to enhance diagnostically significant features. To address data imbalance, data augmentation methods were applied to the training dataset. Multiple

machine learning classifiers were then evaluated, and their performances were comparatively analyzed. The experimental results demonstrate that the combination of targeted preprocessing and data augmentation significantly improves classification accuracy, particularly for underrepresented classes. Additionally, the proposed approach enhances model robustness when handling heterogeneous retinal image data. Overall, the findings suggest that the presented framework can serve as an effective decision-support tool for clinical screening and treatment planning, contributing to the integration of intelligent technologies into practical ophthalmic healthcare applications.

Keywords: Diabetic Retinopathy, Convolutional Neural Network, Data Augmentation, Early Detection, Machine Learning, Ophthalmology

افزایش دقت تشخیص رتینوپاتی دیابتی خفیف و تکثیری با استفاده از تصاویر شبکه‌ای افزوده شده

بابک اسماعیل‌پور قوچانی* 

چکیده:

رتینوپاتی دیابتی یکی از شایع‌ترین و جدی‌ترین عوارض مزمن بیماری دیابت است که در صورت عدم تشخیص به‌موقع می‌تواند به کاهش تدریجی بینایی و در مراحل پیشرفته به نابینایی منجر شود. تشخیص دقیق مراحل مختلف این بیماری، به‌ویژه در طبقات خفیف و پیشرفته، نقش تعیین‌کننده‌ای در کنترل روند پیشرفت آسیب‌های شبکه‌ای و بهبود اثربخشی مداخلات درمانی دارد. با این حال، توسعه سامانه‌های هوشمند تشخیصی همواره با چالش‌هایی نظیر عدم توازن داده‌ها و کمبود نمونه‌های تصویری در برخی طبقات مواجه است که می‌تواند موجب کاهش کارایی مدل‌های تشخیصی شود. هدف این پژوهش ارائه یک چارچوب مبتنی بر یادگیری ماشین برای بهبود تشخیص و درجه‌بندی شدت رتینوپاتی دیابتی از تصاویر فوندوس شبکه‌ای است. در این مطالعه، تصاویر ابتدا تحت فرآیندهای پیش‌پردازشی شامل نرمال‌سازی و بهبود کنتراست تطبیقی محدود بر روی کانال سبز قرار گرفتند تا ویژگی‌های تشخیصی مهم تقویت شوند. سپس، به‌منظور کاهش اثر عدم توازن داده‌ها، از روش‌های تقویت داده بر روی مجموعه آموزش استفاده شد. مدل‌های مختلف یادگیری ماشین برای طبقه‌بندی تصاویر شبکه‌ای مورد ارزیابی قرار گرفتند و عملکرد آن‌ها به‌صورت مقایسه‌ای تحلیل شد. نتایج حاصل نشان داد که به‌کارگیری پیش‌پردازش هدفمند و تقویت داده می‌تواند موجب بهبود قابل توجه دقت تشخیص، به‌ویژه در طبقات کم‌نمونه، شود. این رویکرد همچنین پایداری مدل‌ها را در مواجهه با داده‌های ناهمگون افزایش داد. در مجموع، یافته‌های این پژوهش بیانگر آن است که استفاده از چارچوب پیشنهادی می‌تواند به‌عنوان یک ابزار پشتیبان مؤثر در فرآیند غربالگری بالینی و تصمیم‌گیری درمانی مورد استفاده قرار گیرد و گامی مؤثر در جهت پیوند فناوری‌های هوشمند با نیازهای عملی حوزه سلامت چشم باشد.

^۱ گروه کامپیوتر و هوش مصنوعی، قوچان، دانشگاه آزاد اسلامی، قوچان، ایران،

Babak.esmaeilpour@iaui.ac.ir

نویسنده مسئول

* بابک اسماعیل‌پور قوچانی، گروه کامپیوتر و هوش مصنوعی، قوچان، دانشگاه آزاد اسلامی، قوچان، ایران،

Babak.esmaeilpour@iaui.ac.ir

موضوع اصلی: افزایش دقت تشخیص با استفاده از تصاویر شبکه‌ای

تاریخ دریافت: ۱۴۰۴/۸/۲۲

تاریخ بازنگری: ۱۴۰۴/۹/۲۸

تاریخ پذیرش: ۱۴۰۴/۹/۲۹

<https://doi.org/10.82195/ntds.2025.1224526>

کلیدواژه‌ها: رتینوپاتی دیابتی، شبکه عصبی کانولوشنی، افزایش داده، تشخیص زودهنگام، یادگیری ماشین، بینایی‌سنجی.

۱- مقدمه

رتینوپاتی دیابتی^۱ به عنوان یکی از پیچیده ترین و شایع ترین پیامدهای بلندمدت دیابت، همواره مورد توجه جامعه پزشکی بوده است. این بیماری از طریق تأثیرات مخرب قند خون بالا بر ساختار عروقی شبکیه چشم^۲ شکل می گیرد و در صورت عدم تشخیص و درمان به موقع، می تواند منجر به افت شدید بینایی یا نابینایی کامل شود. رتینوپاتی دیابتی در سیر پیشرفت خود دارای مراحل مختلفی نظیر رتینوپاتی خفیف^۳، متوسط^۴، شدید^۵ و تکثیری^۶ است که شناسایی به موقع این مراحل به ویژه در مراحل ابتدایی نقش مهمی در کنترل و مداخله بالینی ایفا می کند. با پیشرفت فناوری های تصویربرداری پزشکی از جمله فوندوس فتوگرافی^۷، امکان ثبت داده های دقیق تر فراهم شده اما چالش های تحلیل حجم انبوه داده های تصویری همچنان پابرجاست. در این میان، ضرورت بهره گیری از روش های خودکار و هوشمند برای تفسیر داده های شبکیه، انگیزه ای برای ورود علوم داده و الگوریتم های یادگیری ماشین به حوزه تشخیص بالینی فراهم ساخته است.

با وجود گسترش کاربرد مدل های مبتنی بر بینایی ماشین^۸ و یادگیری عمیق^۹ در تحلیل تصاویر پزشکی، همچنان یک چالش بنیادین در داده های مربوط به رتینوپاتی دیابتی به چشم می خورد و آن هم عدم توازن تعداد نمونه ها در میان طبقات مختلف بیماری است. طبقات خفیف، شدید و تکثیری معمولاً با کمبود نمونه های تصویری مواجه اند که این مسئله باعث کاهش دقت و اطمینان مدل های پیش بینی می شود. از سوی دیگر، سیستم های تصمیم یار پزشکی^{۱۰} برای بهره مندی از این فناوری ها نیازمند داده هایی با کیفیت بالا، تنوع کافی و نمایندگی کامل از تمامی مراحل بیماری هستند. در این شرایط، توجه به فرایندهای پیش پردازشی^{۱۱} مانند نرمال سازی^{۱۲}، فیلترگذاری، و مهم تر از آن، افزایش داده^{۱۳}، به عنوان ابزارهایی برای مقابله با عدم توازن داده اهمیت فزاینده ای یافته است. استفاده از این تکنیک ها می تواند باعث ارتقای قابلیت تعمیم پذیری مدل های یادگیری عمیق و افزایش دقت در شناسایی مراحل کم نمونه شود. این مقاله تلاش دارد تا با تمرکز بر این رویکردها، مسیری مؤثر برای بهبود عملکرد تشخیصی مدل ها فراهم کند.

در ادامه این پژوهش، یک رویکرد بین رشته ای با محوریت استفاده از شبکه های عصبی کانولوشنی^{۱۴}، همراه با تکنیک های افزایش داده و ارزیابی جامع عملکرد بر اساس شاخص های متنوع پزشکی مانند دقت^{۱۵}، حساسیت، ویژگی و معیار طراحی شده است. داده های تصویری مورد استفاده از بیماران دارای رتینوپاتی دیابتی استخراج شده و در دسته بندی های کلینیکی استاندارد قرار گرفته اند. تمرکز ویژه ای بر تحلیل عملکرد مدل در طبقات خفیف و تکثیری صورت گرفته که در مقایسه با طبقات دیگر، دارای داده های کمتری هستند. نوآوری اصلی مقاله در آن است که با بهره گیری از مفاهیم پیشرفته علوم داده نظیر الگوریتم های افزایش داده، مشکل کمبود نمونه در کلاس های حیاتی بالینی را هدف قرار داده و در نهایت، مدلی پیشنهاد شده که بتواند در کاربردهای

¹ Diabetic Retinopathy

² Retinal Vessels

³ Mild

⁴ Moderate

⁵ Severe

⁶ Proliferative

⁷ Fundus Photography

⁸ Computer Vision

⁹ Deep Learning

¹⁰ Clinical Decision Support Systems

¹¹ Preprocessing

¹² Normalization

¹³ Data Augmentation

¹⁴ Convolutional Neural Networks

¹⁵ Accuracy

بالینی واقعی به عنوان یک ابزار مکمل و قابل اعتماد برای متخصصان چشم پزشکی مورد استفاده قرار گیرد. این تلاش می تواند گامی مهم در جهت همگرایی دانش مهندسی داده و نیازهای تشخیص پزشکی در حوزه بینایی سنجی باشد.

۲- پیشینه پژوهش:

۱-۲- اهمیت تشخیص زودهنگام رتینوپاتی دیابتی

تشخیص زودهنگام رتینوپاتی دیابتی^۱ به عنوان یکی از جدی ترین عوارض مزمن بیماری دیابت، نقش حیاتی در پیشگیری از کاهش بینایی و نابینایی دائمی ایفا می کند. مطالعات اخیر در حوزه چشم پزشکی نشان داده اند که شروع آسیب های میکروواسکولار شبکه^۲ ممکن است بدون علائم بالینی محسوس باشد، اما با استفاده از تکنولوژی های تصویربرداری دقیق مانند فوندوس فتوگرافی^۳ و توموگرافی همدوسی نوری^۴، امکان شناسایی زودهنگام تغییرات جزئی فراهم شده است [۱]. با این حال، تفسیر این تصاویر به صورت دستی توسط متخصصان چشم^۵ علاوه بر وقت گیر بودن، مستعد سوگیری انسانی است، به ویژه زمانی که تغییرات شبکه در مراحل اولیه خفیف^۶ بروز می کنند. در نتیجه، نیاز به توسعه سامانه های پشتیبان تصمیم گیری بالینی^۷ که بتوانند با استفاده از الگوریتم های هوشمند، تصاویر را تحلیل و طبقه بندی کنند، بیش از پیش احساس می شود [۲]. از منظر پزشکی، مداخله به موقع در مراحل اولیه رتینوپاتی می تواند از طریق کنترل دقیق قند خون، فشار خون و تجویز لیزرتراپی^۸ از پیشرفت بیماری جلوگیری کند [۳]. مطالعات بین المللی نظیر تحقیق های انجام شده در برنامه ی غربالگری NHS انگلستان، تأکید دارند که سیستم های خودکار تشخیص اولیه با دقت بالا می توانند نرخ تشخیص به موقع را تا ۲۰ درصد افزایش دهند. بنابراین، استفاده از روش های داده محور^۹ نه تنها از منظر کارایی سیستم سلامت، بلکه از دیدگاه کیفیت مراقبت از بیماران نیز دارای اهمیت بالایی است [۴]. تمرکز این مقاله بر طراحی مدلی است که بتواند در چنین شرایطی با تشخیص سریع و قابل اطمینان مراحل ابتدایی بیماری، نقش مکمل و مؤثری در فرآیند درمان و پیشگیری ایفا کند.

۲-۲- چالش های تشخیص طبقات خفیف و تکثیری رتینوپاتی دیابتی^{۱۰}

تشخیص دقیق مراحل خفیف و تکثیری^{۱۱} رتینوپاتی دیابتی^{۱۲} همواره یکی از چالش های جدی در حوزه تشخیص بالینی و تحلیل داده های تصویری پزشکی بوده است. مرحله خفیف اغلب با رگ برآمدگی های ریزی^{۱۳} کوچک و بسیار ظریف آغاز می شود که به راحتی در تصاویر فوندوس^{۱۴} پنهان می شوند یا با نویزهای تصویر اشتباه گرفته می شوند. در مقابل، مرحله تکثیری با رشد غیرعادی عروق خونی جدید^{۱۵} همراه است که ممکن است بسته به شرایط تصویربرداری یا پوشاندگی ناحیه ماکولا^۱ به درستی

¹ Diabetic Retinopathy

² Retinal Microvascular Damage

³ Fundus Photography

⁴ Optical Coherence Tomography

⁵ Ophthalmologists

⁶ Mild Stage

⁷ Clinical Decision Support Systems

⁸ Laser Photocoagulation

⁹ Data-Driven Approaches

¹⁰ Challenges in Diagnosing Mild and Proliferative Diabetic Retinopathy

¹¹ Proliferative

¹² Diabetic Retinopathy

¹³ Microaneurysms

¹⁴ Fundus Images

¹⁵ Neovascularization

شناسایی نشوند. این وضعیت باعث شده است که بسیاری از مدل‌های یادگیری ماشین شبکه‌های عصبی عمیق^۲ در طبقه‌بندی صحیح این دوطبقه دچار ضعف عملکرد شوند[۲]. از منظر پزشکی، این دو مرحله، به‌ویژه مرحله تکثیری، نقش بحرانی در تعیین نوع و شدت مداخلات درمانی ایفا می‌کنند؛ زیرا عدم تشخیص به‌موقع آن می‌تواند منجر به خونریزی زجاجیه^۳، جداسازی شبکیه^۴ و در نهایت نابینایی کامل گردد[۵]. در مطالعات اخیر، محققان سعی کرده‌اند با به‌کارگیری تکنیک‌های پیش‌پردازشی، استخراج ویژگی‌های پیچیده‌تر^۵ و تمرکز بر داده‌های نادر^۶ دقت مدل‌ها را در این دو طبقه افزایش دهند. در عین حال، عدم توازن آماری موجود در مجموعه داده‌ها، به‌ویژه کم‌نمونه بودن تصاویر مربوط به مراحل خفیف و تکثیری، مانعی جدی برای یادگیری دقیق مدل‌ها ایجاد کرده است[۶]. از این رو، استفاده از روش‌های افزایش داده و طراحی معماری‌های تخصصی برای تمرکز بیشتر بر جزئیات ناحیه آسیب‌دیده به‌عنوان راهکارهای نوین در مطالعات جدید مطرح شده‌اند[۷].

۲-۳- چالش پیگیری میزان پیشرفت بیماری:

پیش‌بینی پیشرفت رتینوپاتی دیابتی به عنوان یکی از چالش‌های مهم در مدیریت بالینی بیماران دیابتی، در سال‌های اخیر با بهره‌گیری از یادگیری عمیق و تحلیل تصاویر طولی فوندوس چشم، پیشرفت‌های قابل توجهی داشته است. مطالعات جدید نشان می‌دهند که استفاده از تصاویر متوالی یک بیمار در بازه‌های زمانی مختلف، امکان شناسایی تغییرات تدریجی در ساختار شبکیه را فراهم می‌سازد؛ تغییراتی که ممکن است در یک تصویر منفرد قابل تشخیص نباشند. مدل‌هایی مانند DeepDR Plus^۱، با تحلیل بیش از ۷۰۰۰۰۰ تصویر فوندوس از بیماران دیابتی، توانسته‌اند زمان پیشرفت بیماری را تا پنج سال آینده با دقت بالا پیش‌بینی کنند [۸]. همچنین، الگوریتم‌هایی نظیر ADRPPA^۲ با استفاده از تصاویر طولی، توانسته‌اند تغییرات ظریف بین دو تصویر متوالی را شناسایی کرده و خطر پیشرفت بیماری را به‌صورت دقیق‌تری ارزیابی کنند[۹]. این رویکردها نه تنها به بهبود دقت تشخیص کمک می‌کنند، بلکه امکان شخصی‌سازی فواصل غربالگری و مداخلات درمانی را نیز فراهم می‌سازند. در نتیجه، بهره‌گیری از تحلیل تصاویر طولی در کنار مدل‌های یادگیری عمیق، افق‌های جدیدی را در پیش‌بینی و مدیریت مؤثرتر رتینوپاتی دیابتی گشوده است[۷].

۲-۴- نابرابری در توزیع داده‌های تصویری و اهمیت افزایش داده

یکی از چالش‌های کلیدی در توسعه مدل‌های دقیق تشخیص رتینوپاتی دیابتی، نابرابری در توزیع داده‌های تصویری میان طبقات مختلف بیماری است. در اغلب پایگاه‌های داده معتبر، تصاویر متعلق به طبقات متوسط^۷ یا شدید^۸ از نظر تعداد غالب هستند، در حالی که تصاویر مربوط به مراحل خفیف^۹ و تکثیری^{۱۰} به طور محسوسی کمتر دیده می‌شوند. این عدم توازن طبقه‌ای باعث می‌شود تا مدل‌های یادگیری ماشین به طبقات پر نمونه تمایل بیشتری نشان دهند و در نتیجه دقت و حساسیت در شناسایی طبقات کم نمونه کاهش یابد. از دیدگاه پزشکی، چنین خطایی به معنای نادیده گرفتن بیماران در مراحل ابتدایی یا پیشرفته بیماری است که نیاز به مداخله فوری دارند. برای مقابله با این معضل، رویکردهای متنوعی پیشنهاد شده است که مهم‌ترین آن‌ها

¹ Macula

² Deep Neural Networks

³ Vitreous Hemorrhage

⁴ Retinal Detachment

⁵ Advanced Feature Extraction

⁶ Rare Class Emphasis

⁷ Moderate

⁸ Severe

⁹ Mild

¹⁰ Proliferative

استفاده از تکنیک های افزایش داده^۱ است. در این روش، با ایجاد تصاویر جدید از طریق چرخش جابه جایی، بزرگ نمایی، فیلترگذاری و تغییر شدت نور و کنتراست، مجموعه داده های مصنوعی ولی واقع گرایانه ایجاد می شود که می تواند تنوع آماری داده ها را بهبود بخشد [۱۰]. مطالعات اخیر در حوزه بینایی ماشین پزشکی نشان می دهند که استفاده از الگوریتم های پیشرفته افزایش داده مانند GAN^۲ و VAE^۳ ها به شکل چشمگیری دقت مدل ها را در طبقات نادر افزایش داده اند. در عین حال، پژوهشگران پزشکی هشدار داده اند که فرآیند افزایش داده باید به گونه ای انجام شود که ساختار فیزیولوژیک شبکه حفظ شود و منجر به تولید نمونه های غیرواقعی نشود [۱۱]، زیرا این مسئله می تواند منجر به خطای تشخیصی در محیط بالینی گردد. بنابراین، طراحی صحیح فرآیند افزایش داده با مشارکت متخصصان چشم پزشکی و مهندسان داده یکی از محورهای مهم در پژوهش های بین رشته ای کنونی محسوب می شود. هدف این مقاله نیز بهره گیری از افزایش داده هوشمند برای حل همین چالش بنیادی است [۱۲].

۲-۵- پیشینه مدل های پردازش تصویر به کار گرفته شده در درجه بندی رتینوپاتی دیابتی

مطالعات نشان داده اند که استفاده از تصاویر طولی بیماران که در بازه های زمانی مختلف ثبت شده اند، می تواند به طور قابل توجهی تشخیص تغییرات تدریجی در شبکه را بهبود بخشد و پیش بینی دقیق تری از پیشرفت رتینوپاتی دیابتی ارائه دهد. داده های طولی این امکان را به مدل ها می دهند که نه تنها ناهنجاری های ایستا، بلکه روندهای پویا مانند ظهور یا رشد رگ برآمدگی های ریز^۴، رگ زایی جدید^۵ یا خونریزی ها را نیز ارزیابی کنند که برای پیش آگاهی بالینی حیاتی هستند. بهره گیری از داده های زمانی می تواند راهبردهای مداخله زودهنگام را بهبود بخشد و برنامه های درمانی را بر اساس دوره بیماری هر فرد، شخصی سازی کند [۷].

در جدول ۱، چندین مدل پیشروی یادگیری عمیق که برای تشخیص و درجه بندی رتینوپاتی دیابتی استفاده شده اند، به همراه ویژگی های معماری و عملکرد گزارش شده ی آنها خلاصه شده است. این مدل ها شامل شبکه های عصبی کانولوشنی سنتی و همچنین معماری های پیشرفته تری مانند EfficientNet، ResNet و ترنسفورمرهای بینایی هستند که هر یک مزایای منحصر به فردی در استخراج ویژگی ها، حساسیت طبقه بندی و کارایی محاسباتی ارائه می دهند. [۱۳] اثربخشی این مدل ها اغلب به عواملی مانند کیفیت و تنوع داده های آموزشی، استفاده از تکنیک های افزایش داده، وجود مکانیسم های توجه و همچنین یکپارچه سازی داده های تصویربرداری طولی بستگی دارد [۱].

جدول ۱: مدل های یادگیری و ویژگی های آنها [۱۴]

Table 1: Learning models and their features

عملکرد	ویژگی ها	مدل یادگیری عمیق
دقت بالا در تشخیص مراحل مختلف رتینوپاتی دیابتی	مدل ترکیبی برای بهبود دقت طبقه بندی	VGG16 + XGBoost
عملکرد مناسب در تشخیص رتینوپاتی دیابتی با داده های متنوع	استفاده از بلوک های باقی مانده برای آموزش عمیق تر	EfficientNet[۱۵]
در مدل سازی ساختارهای پیچیده خوب عمل می کند؛ ولی نسبت به تنظیمات حساس است	حساس به عدم توازن، نیازمند به منابع محاسباتی بالا، حساسیت بالا به نویز و کنتراست در تصاویر	ResNet50

^۱ Data Augmentation

^۲ Generative Adversarial Networks

^۳ Variational Autoencoders

^۴ Microaneurysms

^۵ Neovascularization

۷ افزایش دقت تشخیص رتینوپاتی دیابتی خفیف و تکثیری با استفاده از تصاویر شبکیه افزوده شده / بابک اسماعیل پور قوچانی

این مدل‌ها با به‌کارگیری تکنیک‌های مختلف یادگیری عمیق، دقت و کارایی بالایی در تشخیص و درجه‌بندی رتینوپاتی دیابتی از تصاویر فوندوس چشم نشان داده‌اند. معماری‌های پیچیده کانولوشنی و شبکه‌های مبتنی بر مکانیزم توجه، توانایی قابل‌توجهی در استخراج ویژگی‌های مرتبط شبکه‌ای و تمایز بین مراحل مختلف بیماری از خود نشان داده‌اند. نقطه قوت این مدل‌ها در توانایی پردازش الگوهای بصری پیچیده و تغییرات ظریف در آناتومی شبکه‌ای نهفته است که برای تشخیص زودهنگام و تصمیم‌گیری بالینی حیاتی است [۱].

با این حال، چالش‌های متعددی همچنان وجود دارد که مانع از بهره‌برداری کامل از پتانسیل این سیستم‌ها می‌شود. عدم تعادل کلاس‌ها در مجموعه داده‌ها، به‌ویژه کم‌نمایی مراحل اولیه و تکثیری رتینوپاتی دیابتی، می‌تواند به پیش‌بینی‌های سوگیرانه و کاهش حساسیت در تشخیص شرایط نادر اما از نظر بالینی مهم منجر شود [۱۶]. همچنین، تنوع در کیفیت تصاویر ناشی از تفاوت در دستگاه‌های تصویربرداری و شرایط نوردهی، نویزی ایجاد می‌کند که ممکن است مدل را دچار اشتباه کند. وابستگی به مجموعه داده‌های بزرگ با حاشیه‌نویسی دستی نیز یک محدودیت اساسی محسوب می‌شود، زیرا برچسب‌زنی توسط متخصصان زمان‌بر و مستعد تفاوت بین ارزیاب‌های مختلف است. این محدودیت‌ها نیاز مداوم به تحقیقات بیشتر را برجسته می‌سازد، از جمله توسعه خطوط پردازش اولیه قوی‌تر، تولید داده‌های مصنوعی و روش‌های یادگیری نیمه‌نظارت‌شده. مدل‌های آینده همچنین باید هدف قرار دهند که داده‌های طولی و متادیتای بالینی را برای افزایش دقت تشخیصی و قابلیت اطمینان در کاربردهای پزشکی واقعی ادغام کنند. این پیشرفت‌ها می‌توانند به سیستم‌های هوش مصنوعی کمک کنند تا به طور مؤثرتری در محیط‌های بالینی پیاده‌سازی شوند و در نهایت منجر به مراقبت‌های بهتر از بیماران دیابتی شوند [۱۶].

۲-۶- پیشینه روش‌های پیش پردازش تصاویر شبکیه :

در تحلیل تصاویر فوندوس شبکیه برای تشخیص رتینوپاتی دیابتی، مرحله پیش‌پردازش نقش حیاتی در بهبود کیفیت تصاویر و افزایش دقت مدل‌های یادگیری عمیق ایفا می‌کند. با توجه به تنوع در شرایط تصویربرداری، کیفیت تصاویر و وجود نویزهای مختلف، استفاده از تکنیک‌های پیش‌پردازش مناسب می‌تواند تأثیر بسزایی در عملکرد سیستم‌های تشخیص خودکار داشته باشد [۱۷].

جدول ۲: تکنیک‌های پیش‌پردازش تصاویر به کار گرفته شدن در تحقیقات مرتبط با رتینوپاتی دیابتی

Table 2: Image preprocessing techniques used in research related to diabetic retinopathy

منبع	توضیح	تکنیک پیش‌پردازش
[۱۷]	بهبود کنتراست محلی تصاویر برای برجسته‌کردن ویژگی‌های ظریف شبکه‌ای	افزایش کنتراست (CLAHE)
[۱۸]	یکسان‌سازی روشنایی تصاویر دارای نورپردازی	نرمال‌سازی شدت روشنایی
[۱۷]	استفاده از کانال سبز تصویر RGB برای بهبود تمایز ساختارهای عروقی	استخراج کانال سبز
[۱۹]	حذف نویز و تقویت لبه‌ها برای بهبود وضوح ساختارهای شبکه‌ای	فیلتر میانگین و لاپلاسیان
[۲۰]	استانداردسازی ابعاد تصاویر برای ورود به شبکه‌های عصبی	تغییر اندازه و برش تصاویر
[۲۱]	ایجاد تنوع در داده‌ها از طریق چرخش، آینه‌سازی و برش برای جلوگیری از بیش‌برازش	افزایش داده (Data Augmentation)

مطالعات اخیر نشان داده‌اند که روش‌هایی مانند افزایش کنتراست، حذف نویز با فیلترهای مختلف، نرمال‌سازی شدت روشنایی، و استخراج کانال سبز تصویر می‌توانند به طور مؤثری ویژگی‌های مهم شبکه‌ای را برجسته کرده و زمینه را برای تحلیل دقیق‌تر فراهم

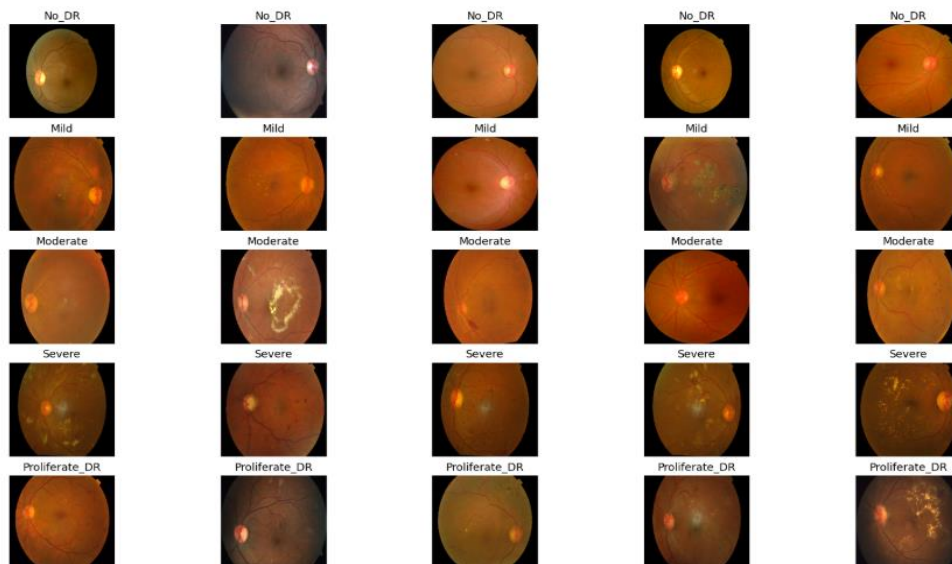
کنند [۲۲]. در جدول زیر، برخی از تکنیک های پیش پردازش رایج در تحلیل تصاویر رتینوپاتی دیابتی به همراه توضیحات و منابع مرتبط ارائه شده است [۲۱]. استفاده از این تکنیک های پیش پردازش، به ویژه در ترکیب با مدل های یادگیری عمیق، می تواند دقت تشخیص مراحل مختلف رتینوپاتی دیابتی را به طور قابل توجهی افزایش دهد. با این حال، انتخاب و تنظیم مناسب این تکنیک ها بر اساس ویژگی های خاص هر مجموعه داده، برای دستیابی به بهترین عملکرد ضروری است.

۳- روش پیشنهادی:

در روش پیشنهادی، برای بررسی و تحلیل دقیق تصاویر رتینوپاتی دیابتی پنج کلاس مختلف از بیماری شامل Mild, No_DR, Moderate, Severe و Proliferate_DR در نظر گرفته شده است که به ترتیب عدم وجود رتینوپاتی دیابتی، ضعیف، متوسط، شدید و رتینوپاتی دیابتی پیشرفته نامیده می شود. [۲۳].

۳-۱- دیتاست:

مجموعه داده **APTOS 2019 Blindness Detection** یکی از مهم ترین و جامع ترین منابع برای تشخیص رتینوپاتی دیابتی است که در رقابت های Kaggle مورد استفاده قرار گرفته است. این مجموعه شامل ۳۶۶۲ تصویر با وضوح بالا از تصاویر تصاویر شبکه چشم است که به پنج کلاس مختلف بر اساس شدت بیماری تقسیم بندی شده اند. این پنج کلاس شامل عدم وجود، خفیف، متوسط، شدید و پیشرفته هستند. توزیع تصاویر در این مجموعه به صورت نامتوازن است به طوری که بیشتر تصاویر در کلاس های عدم وجود و خفیف قرار دارند، در حالی که کلاس های شدید و پیشرفته دارای تعداد کمتری از نمونه ها هستند. به طور دقیق، کلاس عدم وجود دارای حدود ۱۸۰۵ تصویر، کلاس خفیف حدود ۳۷۰ تصویر، کلاس متوسط حدود ۹۹۹ تصویر، کلاس شدید حدود ۱۹۳ تصویر و کلاس پیشرفته در حدود ۲۹۵ تصویر است. این توزیع نامتعادل، یکی از چالش های اساسی در توسعه مدل های یادگیری عمیق برای تشخیص DR محسوب می شود، زیرا ممکن است مدل ها به دلیل تعداد کمتر نمونه ها در کلاس های شدیدتر، در تشخیص صحیح این موارد دچار مشکل شوند. برای مقابله با این مشکل، روش های افزایش داده استفاده می شوند تا تعادل بهتری بین کلاس ها ایجاد شود. همچنین، این تصاویر در اندازه ها و کیفیت های مختلف هستند و معمولاً نیاز به پردازش های اولیه مانند کاهش نویز، افزایش کنتراست مانند CLAHE و استخراج کانال سبز برای بهبود کیفیت تصاویر دارند. علاوه بر این، متادیتای مرتبط با هر تصویر نیز شامل شناسه تصویر و سطح بیماری است که به شکل فایل های CSV در این مجموعه موجود است. به دلیل این تنوع و چالش های مرتبط با داده ها، مجموعه APTOS 2019 به عنوان یک منبع ارزشمند برای ارزیابی روش های پیشرفته در تشخیص خودکار رتینوپاتی دیابتی در تحقیقات بینایی ماشین و یادگیری عمیق شناخته می شود [۲۳]. همچنین برای یکسان سازی تصاویر همه تصاویر به ابعاد پیش فرض ۲۲۴*۲۲۴ در سه کانال رنگی (3*224*224) تغییر داده شده اند.



شکل ۱: نمونه ای از تصاویر هر کلاس در دیتاست نمونه [۲۳]
Figure 1: An example of images from each class in the sample dataset

۲-۳- مدل پیشنهادی:

مدل پردازش تصویر ارائه شده در این پژوهش باهدف بهبود دقت تشخیص و کاهش مشکل عدم توازن داده‌ها در کلاس‌های مختلف طراحی شده است. در گام نخست، از تکنیک‌های داده‌افزایی برای افزایش تنوع داده‌ها استفاده می‌شود. این تکنیک‌ها شامل چرخش تصاویر، انتقال سطری و ستونی، چرخش ۱۸۰ درجه و ترکیب با تصاویر همسان از همان کلاس‌های داده‌های کم‌تعداد هستند. این روش‌ها با ایجاد تغییرات تصادفی در موقعیت و چرخش تصاویر، به شبکه عصبی کمک می‌کنند تا ویژگی‌های متنوع‌تری از داده‌ها یاد بگیرد و از برازش بیش از حد^۱ جلوگیری شود. پس از داده‌افزایی، مرحله پیش‌پردازش آغاز می‌شود که در آن شدت نور تصاویر تنظیم می‌شود تا تفاوت‌های غیرضروری در روشنایی حذف گردد. همچنین، کانال سبز تصاویر به دلیل اطلاعات بیشتر در زمینه تغییرات بافت و رنگ به‌عنوان ویژگی اصلی استخراج می‌شود. این روش باعث کاهش ابعاد ویژگی‌ها و بهبود کارایی مدل در استخراج ویژگی‌های مهم می‌گردد.

فرض می‌شود مجموعه داده شامل N تصویر شبکه‌ای چشم به‌صورت زیر باشد:

$$D = \{(X_i, y_i)\}_{i=1}^N \quad (1)$$

که $X_i \in \mathbb{R}^{224 \times 224 \times 3}$ که نشان دهنده تصویر رنگی RGB با ابعاد ثابت $224 \times 224 \times 3$ است و $y_i \in \{0, 1, 2, 3, 4\}$ است که بیانگر برچسب شدت بیماری شامل پنج کلاس «عدم رتینوپاتی، خفیف، متوسط، شدید و پرولیفراتیو» است. لازم به تأکید است که در نسخه نهایی مقاله، ابعاد تصویر به‌صورت یکتا و ثابت تعریف شده و هرگونه ابهام قبلی در این خصوص برطرف شده است. در مرحله بعدی پیش‌پردازش جهت افزایش کنتراست تصاویر صورت گرفته است در گام نخست پیش‌پردازش کلیه تصاویر ورودی به ابعاد استاندارد 224×224 تغییر اندازه داده شده و به منظور کاهش ناپایداری عددی، تسریع همگرایی فرآیند یادگیری و جلوگیری

¹ overfitting

از غالب شدن ویژگی هایی با مقیاس بزرگ تر، تحت عملیات نرمال سازی شدت پیکسل قرار گرفته اند. فرض شود تصویر ورودی پس از تغییر اندازه به صورت یک تانسور سه بعدی به شکل زیر تعریف شود:

$$I \in \mathbb{R}^{224 \times 224 \times 3}$$

که در آن هر عنصر $I(x, y, c)$ مقدار شدت پیکسل در مختصات مکانی (x, y) و کانال رنگی

$$c \in \{R, G, B\}$$

را نشان می دهد. با توجه به این که تصاویر ورودی از نوع ۸ بیتی هستند، دامنه مقادیر شدت پیکسل ها به صورت زیر خواهد بود:

$$I(x, y, c) \in [0, 255]$$

برای نگاشت یکنواخت مقادیر شدت پیکسل ها به بازه پیوسته $[0, 1]$ ، از نرمال سازی خطی مبتنی بر بیشینه دامنه استفاده شده است. به این ترتیب، تصویر نرمال شده I^{norm} به صورت زیر محاسبه می شود:

$$I^{norm}(x, y, c) = \frac{I(x, y, c)}{255} \quad (2)$$

پس از این تبدیل، دامنه مقادیر شدت پیکسل ها به شکل زیر تغییر می یابد:

$$X^{norm}(x, y, c) \in [0, 1]$$

این نوع نرمال سازی باعث می شود که تمامی کانال های رنگی در مقیاسی یکسان به شبکه عصبی کانولوشنی ارائه شوند و در نتیجه، گرادین های محاسبه شده در لایه های ابتدایی شبکه از نوسانات شدید جلوگیری کنند. از منظر تئوریک، این عملیات منجر به بهبود پایداری فرآیند بهینه سازی و افزایش سرعت همگرایی الگوریتم یادگیری می شود، به ویژه زمانی که از توابع فعال سازی غیرخطی نظیر $ReLU$ استفاده می شود. لازم به ذکر است که نرمال سازی به صورت یکنواخت بر روی تمامی تصاویر مجموعه داده و بدون استفاده از اطلاعات برچسب ها انجام شده است، از این رو هیچ گونه نشت اطلاعاتی در فرآیند آموزش و ارزیابی مدل رخ نداده است.

باتوجه به شواهد بالینی و مطالعات پیشین، کانال سبز تصویر حاوی بیشترین اطلاعات مرتبط با ضایعات میکروواسکولار نظیر میکروآنورسم ها و خونریزی ها است. از این رو، به منظور تقویت ویژگی های تشخیصی بدون تخریب اطلاعات رنگی کلی، تبدیل CLAHE به صورت انتخابی تنها بر روی کانال سبز اعمال به صورت زیر اعمال می شود.

$$X_i^{(g)} = CLAHE(X^{norm}_i(:, :, 2)) \quad (3)$$

و تصویر نهایی بازسازی شده به صورت زیر تعریف می شود:

$$X_i^{\sim} = [X^{norm}_i(:, :, 1), X_i^{(g)}, X^{norm}_i(:, :, 3)] \quad (4)$$

این راهبرد موجب افزایش وضوح نواحی آسیب دیده شبکه بدون ایجاد نویز یا بیش تقویتی ناخواسته در سایر کانال ها می شود.

در گام بعدی وظیفه متعادل سازی داده ها و افزایش داده که یکی از چالش های اصلی عدم توازن ذاتی کلاس ها در مجموعه داده تشخیص رتینوپاتی دیابتی است. به منظور کاهش بایاس مدل نسبت به کلاس های غالب، از افزایش داده تطبیقی مبتنی بر تبدیلات هندسی استفاده شده است. این تبدیلات شامل چرخش تصادفی، بزرگ نمایی، جابه جایی افقی و عمودی و وارونگی افقی تصاویر

۱۱ افزایش دقت تشخیص رتینوپاتی دیابتی خفیف و تکثیری با استفاده از تصاویر شبکه‌ای افزوده شده / بابک اسماعیل پور قوچانی

هستند. افزایش داده صرفاً بر روی داده‌های آموزش اعمال شده است تا از نشت اطلاعات¹ به مجموعه آزمون جلوگیری شود. بدین ترتیب، تعداد نمونه‌ها در هر کلاس به صورت کنترل شده متعادل شده و پایداری فرایند یادگیری افزایش یافته است.

مدل پیشنهادی شبکه عصبی کانولوشن سبک‌وزن بوده و ۹ لایه قابل یادگیری است که به صورت سلسله‌مراتبی ویژگی‌های مکانی تصاویر را استخراج می‌کند. معماری مدل شامل چهار بلوک کانولوشن - Pooling متوالی و دو لایه تمام‌متصل در انتها است. لایه‌های کانولوشنی با فیلترهای ۳×۳ و توابع فعال‌سازی ReLU طراحی شده‌اند تا توانایی مدل در استخراج ویژگی‌های غیرخطی افزایش یابد. در لایه نهایی، از تابع Softmax برای تولید توزیع احتمال کلاس‌ها استفاده شده است.

تابع نگاشت مدل به صورت زیر تعریف می‌شود:

$$f_{\theta}: \mathbb{R}^{224 \times 224 \times 3} \rightarrow [0,1]^5$$

که در آن θ مجموعه پارامترهای قابل‌یادگیری شبکه است. برای هزینه و فرایند آموزش از تابع هزینه Cross-Entropy چند کلاسه به صورت زیر استفاده شده است.

$$L = -\frac{1}{N} \sum_{i=1}^N \sum_{k=1}^5 y_i^k \log(y_i^k) \quad (5)$$

در این رابطه، y_i^k احتمال پیش‌بینی شده برای کلاس k و y_i^k برچسب واقعی به صورت one-hot است. به منظور بهینه‌سازی تابع هزینه، از الگوریتم Adam با نرخ یادگیری 10^{-3} استفاده شده است. پارامترهای آموزش شامل اندازه دسته برابر با ۳۲ و تعداد تکرارهای آموزش برابر با ۲۵ در نظر گرفته شده‌اند. برای جلوگیری از بیش‌برازش، از لایه Dropout با نرخ ۰.۵ در بخش تمام‌متصل استفاده شده است. در مسائل طبقه‌بندی چندکلاسه با عدم توازن داده‌ها، استفاده از تابع هزینه آنتروپی متقاطع استاندارد می‌تواند منجر به غالب شدن نمونه‌های آسان و کاهش حساسیت مدل نسبت به طبقات کم‌نمونه شود. به منظور رفع این مشکل، در این پژوهش از تابع هزینه فوکال استفاده شده است که با اعمال یک ضریب وزنی تطبیقی، سهم نمونه‌های سخت را در فرآیند یادگیری افزایش می‌دهد. فرض شود برای هر نمونه ورودی X_i بردار برچسب واقعی به صورت one-hot به شکل:

$$y_i = (y_i^1, y_i^2, \dots, y_i^C)$$

است و خروجی مدل پس از اعمال تابع Softmax برابر با

$$p_i = (p_i^1, p_i^2, \dots, p_i^C)$$

باشد که C در رابطه بالا همان تعداد کلاس‌ها می‌باشد. ابتدا آنتروپی متقاطع چند کلاسه به صورت زیر تعریف می‌شود:

$$\mathcal{L}_{FL}(i) = -\sum_{c=1}^C y_i^c \log(p_i^c) \quad (6)$$

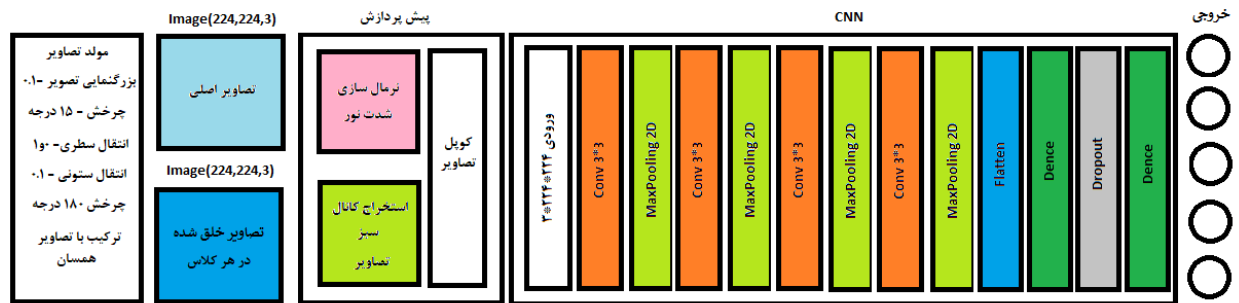
تابع هزینه فوکال با افزودن یک ضریب کاهنده وابسته به میزان اطمینان پیش‌بینی، به صورت زیر تعریف می‌شود:

$$\mathcal{L}_{FL}(i) = -\sum_{c=1}^C \alpha y_i^c - (1 - p_i^c)^{\gamma} \log(p_i^c) \quad (7)$$

که در آن p_i^c احتمال پیش بینی شده برای کلاس c می باشد و مقدار y_i^c مقدار برچسب واقعی عکس و α ضریب وزن دهی کلاس ها به منظور کنترل اثر توازن داده ها است و γ پارامتر تمرکز برای کاهش سهم نمونه های آسان می باشد. برای جلوگیری از ناپایداری عددی ناشی از لگاریتم مقادیر نزدیک به صفر، مقادیر p_i^c در بازه $[1 - \epsilon, \epsilon]$ محدود می شوند که در آن ϵ یک مقدار بسیار کوچک مثبت است. در نهایت، مقدار تابع هزینه برای هر نمونه با جمع بر روی تمامی کلاس ها محاسبه می شود:

$$\mathcal{L}_{FL}(i) = -\sum_{c=1}^C \alpha y_i^c (1 - p_i^c)^\gamma (-\log(p_i^c)) \quad (8)$$

این تابع هزینه موجب می شود نمونه هایی که به درستی و با اطمینان بالا طبقه بندی شده اند، سهم کمتری در فرایند یادگیری داشته باشند، در حالی که نمونه های سخت و کم نمونه تأثیر بیشتری بر به روزرسانی پارامترهای مدل خواهند داشت.



شکل ۲: مدل پیشنهادی شبکه عصبی کانولوشن ۱۲ لایه با کنتراست رنگ CLAHE و استخراج کانال سبز و بهره گیری از تابع زیان فوکال
Figure 2: Proposed 12-layer convolutional neural network model with CLAHE color contrast, green channel extraction, and use of focal loss function

راهبرد تقسیم داده به صورت زیر است که مجموعه داده به صورت متوازن و طبقه بندی شده به سه بخش تقسیم شده است: ۸۰ درصد برای آموزش، ۱۰ درصد برای اعتبارسنجی و ۱۰ درصد برای آزمون نهایی. داده های اعتبارسنجی از مجموعه آموزش استخراج شده اند و در طول فرایند یادگیری برای پیش عملکرد مدل مورد استفاده قرار گرفته اند. این راهبرد امکان ارزیابی منصفانه و قابل تکرار عملکرد مدل را فراهم می کند. وآوری اصلی این پژوهش در ترکیب هدفمند سه مؤلفه کلیدی نهفته است: (۱) اعمال انتخابی CLAHE بر کانال سبز تصویر به عنوان مؤثرترین منبع اطلاعات تشخیصی، (۲) متعادل سازی تطبیقی کلاس ها از طریق افزایش داده کنترل شده، و (۳) طراحی یک معماری CNN سبک وزن با پیچیدگی محاسباتی پایین و مناسب برای کاربردهای بالینی. این هم افزایی به صورت یکپارچه در مطالعات پیشین گزارش نشده و منجر به بهبود پایداری و دقت تشخیص شدت رتینوپاتی دیابتی شده است.

جدول ۳: معرفی ویژگی های مدل های مورد ارزیابی قرار گرفته شده

Table 3: Introduction of the features of the evaluated models

تعداد تصاویر آموزشی	بهینه ساز	متوازن سازی داده	پیش پردازش	تعداد لایه	معماری	مدل
۲۰۰۰	VGG16				XGBoost	
۲۰۰۰	Adam	ندارد	ندارد	۱۲	Conv2D, MaxPooling, Sequential Dense	CNN ساده
۲۰۰۰	Adam	دارد	CLAHE	۱۲	Conv2D, MaxPooling, Sequential Dense	CNN ساده
۲۰۰۰	Adam	دارد	فیلتر لاپلاسی	۱۲	Conv2D, MaxPooling, Sequential Dense	CNN ساده
۲۰۰۰	Adam	دارد	CLAHE و کانال سبز	۱۲	Conv2D, MaxPooling, Sequential Dense	CNN ساده

Focal CNN Loss	Conv2D, MaxPooling, Sequential Dropout	۱۲	CLAHE و کانال سبز	دارد	Adam + Focal Loss	۲۰۰۰
EfficientNet	EfficientNetB0	۸۰	CLAHE و کانال سبز	دارد	Adam	۲۰۰۰
ResNet50	ResNet50	۸۰	CLAHE و کانال سبز	دارد	Adam + Focal Loss	۲۰۰۰

نتایج حاصل از ارزیابی مدل‌ها نشان‌دهنده بهبود دقت تشخیص و کاهش نرخ خطا در طبقه‌بندی سطوح مختلف بیماری رتینوپاتی دیابتی بوده است. جدول --- جزئیات هر مدل از جمله معماری، پیش‌پردازش، اندازه داده‌ها، و دقت نهایی را ارائه می‌دهد:

۴- ارزیابی نتایج:

نتایج به‌دست‌آمده از مدل پیشنهادی نشان می‌دهد که تمرکز اصلی این رویکرد نه صرفاً تشخیص وجود یا عدم وجود رتینوپاتی دیابتی، بلکه تفکیک دقیق شدت بیماری در مراحل بالینی حساس است. این تمایز از آن جهت اهمیت دارد که در بسیاری از مطالعات پیشین، گزارش عملکرد عمدتاً به شاخص‌های کلی محدود شده و توان مدل‌ها در تشخیص مراحل شدید کمتر مورد توجه قرار گرفته است. در این پژوهش، استفاده هم‌زمان از استخراج کانال سبز تصاویر شبکه‌ای، بهبود کنتراست موضعی با تکنیک CLAHE و به‌کارگیری تابع زیان فوکال، یک چارچوب هدفمند برای یادگیری ویژگی‌های بحرانی فراهم کرده است. ارزیابی مبتنی بر ماتریس درهم‌ریختگی و منحنی‌های ROC نشان می‌دهد که اگرچه عملکرد مدل در کلاس سالم بسیار بالا و مشابه سایر مدل‌هاست، اما تفاوت واقعی در کلاس‌های با پیچیدگی تشخیصی بیشتر نمایان می‌شود. این الگو بیانگر آن است که معماری و راهبرد یادگیری مدل به‌گونه‌ای طراحی شده که به‌جای بهینه‌سازی ساده دقت کلی، بر کاهش خطاهای پرهزینه از نظر بالینی تمرکز دارد؛ بنابراین، نتایج باید در چارچوب هدف کاربردی مدل تفسیر شوند و نه صرفاً بر اساس معیارهای کلی عملکرد.

مقایسه مستقیم مدل مبتنی بر تابع زیان فوکال با مدل مشابه بدون این تابع، نشان‌دهنده نقش تعیین‌کننده آن در بهبود عملکرد در کلاس‌های نامتوازن است. نتایج ماتریس به‌هم‌ریختگی در شکل زیر نشان می‌دهد که در کلاس‌های سالم و متوسط، تفاوت محسوسی میان دو مدل مشاهده نمی‌شود؛ امری که با ماهیت این کلاس‌ها، به‌عنوان گروه‌های پرتعداد و آسان‌تر برای یادگیری، هم‌خوانی دارد. با این حال، در کلاس‌های شدید و پیشرفته، مدل دارای تابع زیان فوکال عملکرد بهتری از خود نشان داده است. این موضوع بیانگر آن است که فوکال لاس با کاهش وزن نمونه‌های آسان و افزایش تمرکز بر نمونه‌های دشوار، توانسته است فرایند یادگیری را به سمت تمایز بهتر الگوهای پیچیده عروقی هدایت کند. از منظر تحلیلی، این یافته با ادبیات موجود در حوزه یادگیری عمیق برای داده‌های پزشکی نامتوازن هم‌راستاست، جایی که تمرکز بر کاهش خطای نوع دوم در کلاس‌های بحرانی اهمیت ویژه‌ای دارد؛ بنابراین، استفاده از تابع زیان فوکال در این پژوهش نه یک انتخاب تزینی، بلکه یک تصمیم روش‌شناختی مبتنی بر ماهیت داده و هدف بالینی مدل بوده است.

در مقایسه با مدل‌های عمیق مبتنی بر یادگیری انتقالی، نتایج ماتریس به‌هم‌ریختگی نشان می‌دهد که EfficientNet80 علی‌رغم کارایی مناسب در تشخیص نمونه‌های سالم، تمایل قابل‌توجهی به سوگیری به سمت کلاس غالب دارد. این رفتار موجب کاهش توان آن در تفکیک مراحل شدید و پیشرفته بیماری شده است [۲۴]. ResNet50 نیز، با وجود عمق زیاد و ظرفیت بالای یادگیری، در تمایز دقیق میان کلاس‌های خفیف و پیشرفته عملکرد پایداری از خود نشان نداده است. این مسئله می‌تواند ناشی از

حساسیت بالای این معماری به توزیع داده و نبود راهبردهای کنترلی مؤثر در فرآیند آموزش باشد. در مقابل، مدل پیشنهادی با معماری هدفمندتر و ترکیب پیش پردازش های تخصصی شبکه، توانسته است توازن بهتری میان حساسیت و اختصاصیت در کلاس های بحرانی ایجاد کند. این نتایج نشان می دهد که پیچیدگی معماری به تنهایی تضمین کننده عملکرد بهتر نیست و تطبیق راهبرد یادگیری با مسئله بالینی نقش تعیین کننده ای دارد. از این منظر، مدل پیشنهادی نسبت به معماری های انتقالی عمومی، سازگاری بیشتری با مسئله درجه بندی شدت رتینوپاتی دیابتی دارد.

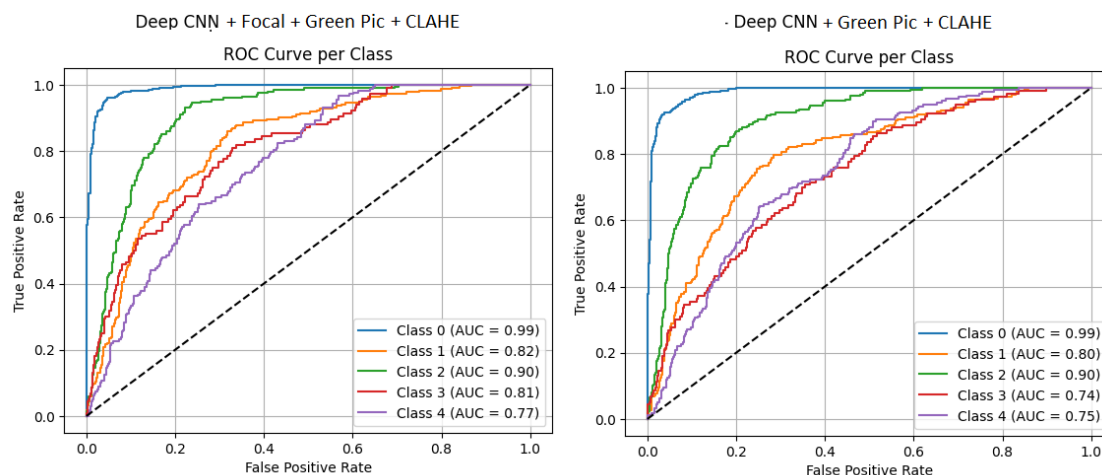


شکل ۳: ماتریس کانفیوژن مدل های یادگیری مورد ارزیابی قرار گرفته

Figure 3: Confusion matrix of the evaluated learning models

مدل **XGBoost** به‌عنوان نماینده‌ای از روش‌های یادگیری ماشین کلاسیک [۱۹]، همانطور که در ماتریس به هم ریختگی دیده می‌شود در شناسایی کلاس سالم عملکرد قابل قبولی از خود نشان داده است، اما در تفکیک مراحل مختلف بیماری، به‌ویژه در کلاس‌های شدید و پیشرفته، دچار هم‌پوشانی قابل توجه شده است. تحلیل نتایج نشان می‌دهد که این مدل تمایل دارد بسیاری از نمونه‌ها را در کلاس متوسط طبقه‌بندی کند، که می‌تواند ناشی از ناتوانی آن در استخراج الگوهای مکانی پیچیده از تصاویر شبکه‌ای باشد. این رفتار مشابه عملکرد شبکه عصبی کانولوشنی با پیش‌پردازش ساده لاپلاسی است که آن نیز فاقد توان کافی برای تمایز دقیق الگوهای عمیق عروقی بوده است. این یافته‌ها نشان می‌دهد که مدل‌های غیرکانولوشنی یا مبتنی بر ویژگی‌های دستی، اگرچه در مسائل ساده‌تر کارآمد هستند، اما در کاربردهای تصویربرداری پزشکی با پیچیدگی بالا محدودیت ذاتی دارند. در نتیجه، برتری مدل پیشنهادی بیش از آنکه ناشی از افزایش صرف دقت باشد، به دلیل تناسب ساختاری آن با ماهیت داده و هدف تشخیصی مسئله است.

بر اساس تحلیل منحنی‌های ROC زیر ارائه‌شده، می‌توان نتیجه گرفت که مدل پیشنهادی مبتنی بر Deep CNN همراه با کانال سبز، CLAHE و تابع زیان فوکال توانسته است الگوی یادگیری هدمندتری نسبت به نسخه بدون فوکال لاس ایجاد کند. همان‌گونه که منحنی‌ها نشان می‌دهند، عملکرد مدل در کلاس سالم در هر دو حالت تقریباً یکسان و نزدیک به حد ایده‌آل است که بیانگر توان بالای مدل در تشخیص نمونه‌های غیر بیمار و جلوگیری از مثبت کاذب است. با این حال، تفاوت اصلی و معنادار در کلاس‌های با پیچیدگی تشخیصی بیشتر، به‌ویژه کلاس‌های شدید و پیشرفته، مشاهده می‌شود؛ به‌طوری که استفاده از تابع زیان فوکال منجر به جابه‌جایی محسوس منحنی ROC به سمت گوشه بالا - چپ شده و بیانگر افزایش حساسیت مدل در نرخ خطای ثابت است. این رفتار نشان می‌دهد که مدل پیشنهادی توانسته است نمونه‌های دشوار و کم‌تعداد را بهتر در فضای ویژگی تفکیک کند و از غلبه کلاس‌های پر تعداد جلوگیری نماید. از منظر بالینی، این ویژگی اهمیت بالایی دارد، زیرا خطا در تشخیص مراحل شدید و پیشرفته می‌تواند پیامدهای جبران‌ناپذیری برای بیمار به همراه داشته باشد؛ بنابراین، تحلیل ROC تأیید می‌کند که نوآوری اصلی مدل نه در بهبود تشخیص موارد ساده، بلکه در افزایش قابلیت اعتماد مدل در تصمیم‌گیری‌های حساس بالینی نهفته است که آن را به گزینه‌ای مناسب برای سیستم‌های پشتیبان تشخیص در غربالگری و مدیریت رتینوپاتی دیابتی تبدیل می‌کند.



شکل ۴: منحنی ROC برای مدل‌های پیشنهادی

Figure 4: ROC curve for the proposed models

در مجموع، همان طور که در جدول مقایسه ای زیر قابل رویت است نتایج این پژوهش نشان می دهد که اگر هدف صرفاً تشخیص وجود یا عدم وجود رتینوپاتی دیابتی باشد، اکثر مدل های بررسی شده عملکرد نسبتاً مشابهی دارند. باین حال، در سناریوهایی که تفکیک دقیق شدت بیماری، به ویژه در مراحل شدید و پیشرفته، اهمیت بالینی بالایی دارد، مدل پیشنهادی برتری نسبی و معناداری از خود نشان می دهد. این ویژگی می تواند نقش مهمی در پشتیبانی از تصمیم گیری های بالینی، مانند اولویت بندی بیماران برای مداخلات درمانی فوری، ایفا کند. در عین حال، محدودیت هایی نظیر چالش در تمایز میان کلاس های سالم و خفیف همچنان وجود دارد که می تواند ناشی از شباهت ظاهری این مراحل و تمرکز کمتر مدل بر نمونه های آسان باشد. این محدودیت مستقیماً به طراحی تابع زیان فوکال مرتبط است و نشان می دهد که بهبود عملکرد در یک بخش، ممکن است به بهای کاهش حساسیت در بخش دیگر تمام شود؛ بنابراین، نتایج باید به صورت متوازن تفسیر شوند و مدل به عنوان یک ابزار کمکی تخصصی برای تشخیص مراحل بحرانی بیماری در نظر گرفته شود، نه یک سامانه جایگزین کامل برای ارزیابی بالینی.

جدول ۲: مقایسه مدل ها بر اساس نتایج شبیه سازی

Table 4: Comparison of models based on simulation results

مدل	نقاط قوت اصلی	نقاط ضعف اصلی	عملکرد در کلاس های شدید	پیامد بالینی
مدل پیشنهادی (CNN + CLAHE + Green + Focal Loss)	تمرکز بر نمونه های دشوار، بهبود تشخیص مراحل بحرانی	ضعف نسبی در تمایز سالم و خفیف	بالا و پایدار	مناسب برای پشتیبانی تصمیم درمانی
CNN بدون Focal Loss	سادگی آموزش، عملکرد مناسب در کلاس های آسان	افت دقت در کلاس های نادر	متوسط	محدود در کاربرد بالینی پیشرفته
EfficientNetB0	دقت بالا در تشخیص سالم	سوگیری به کلاس غالب	پایین	مناسب برای غربالگری اولیه
ResNet50	ظرفیت یادگیری بالا	ناپایداری در تفکیک کلاس ها	متوسط تا پایین	نیازمند تنظیمات پیچیده
XGBoost	عملکرد خوب در کلاس سالم	ناتوان در استخراج الگوهای مکانی	پایین	نامناسب برای درجه بندی شدت

۵- نتیجه گیری:

نتایج ارزیابی نشان می دهد که هر دو معماری مبتنی بر شبکه عصبی کانولوشنی که از پیش پردازش کانال سبز تصاویر شبکه و تکنیک بهبود کنتراست موضعی CLAHE بهره می برند، در تشخیص کلاس سالم عملکردی نزدیک به حد ایده آل از خود نشان داده اند. مقدار بالای AUC در این کلاس بیانگر آن است که مدل ها قادرند با دقت بالا موارد فاقد نشانه های بیماری را شناسایی کرده و از ورود بیماران سالم به چرخه درمانی غیر ضروری جلوگیری کنند. این ویژگی از منظر بالینی اهمیت ویژه ای دارد، زیرا کاهش خطای مثبت کاذب نقش مستقیمی در کاهش بار کاری سیستم های غربالگری و افزایش اعتماد پزشکان به سامانه های تصمیم یار دارد. باین حال، تشخیص کلاس سالم به دلیل وضوح ساختاری بیشتر تصاویر، چالش اصلی مدل محسوب نمی شود و ارزیابی توان واقعی مدل باید بر اساس عملکرد آن در مراحل بالینی حساس تر صورت گیرد.

تحلیل دقیق تر نتایج نشان می دهد که تفاوت اصلی میان مدل دارای تابع زیان فوکال و مدل فاقد آن، در طبقه بندی مراحل شدید و پیشرفته بیماری آشکار می شود. در این کلاس ها، مدل پیشنهادی با استفاده از فوکال لاس توانسته است تمرکز یادگیری را از

نمونه‌های آسان به سمت نمونه‌های دشوار و کم نمونه هدایت کند و در نتیجه تفکیک بهتری میان الگوهای پاتولوژیک پیچیده ایجاد نماید. این رفتار مستقیماً با ماهیت تابع زیان فوکال همخوانی دارد، زیرا کاهش سهم نمونه‌های پرتعداد باعث می‌شود مدل در برابر عدم توازن داده‌ها مقاوم‌تر عمل کند. از منظر بالینی، این بهبود در تشخیص مراحل شدید و پیشرفته حائز اهمیت است، چرا که این مراحل بیشترین ریسک کاهش غیرقابل‌بازگشت بینایی را به همراه دارند و حتی بهبود نسبی در دقت تشخیص آن‌ها می‌تواند تأثیر قابل‌توجهی بر زمان‌بندی مداخلات درمانی داشته باشد. در مقابل، هر دو نسخه مدل در تفکیک مراحل خفیف و متوسط با محدودیت‌هایی مواجه بوده‌اند. این مسئله را می‌توان به شباهت بالای الگوهای ظاهری این مراحل نسبت داد، به گونه‌ای که تغییرات عروقی در این سطوح اغلب ظریف و پراکنده بوده و حتی برای متخصصان انسانی نیز تشخیص آن‌ها چالش‌برانگیز است. علاوه بر این، استفاده از تابع زیان فوکال اگرچه در افزایش حساسیت به نمونه‌های دشوار مؤثر بوده، اما به طور ذاتی تمرکز مدل را از نمونه‌های ساده‌تر کاهش می‌دهد که این امر می‌تواند یکی از دلایل عدم بهبود محسوس در کلاس‌های اولیه باشد. این یافته نشان می‌دهد که مزیت فوکال لاس وابسته به سطح پیچیدگی کلاس هدف بوده و اثر آن در مراحل میانی بیماری محدودتر است.

مقایسه با مدل‌های مرجع مورد استفاده در این پژوهش، تصویر شفاف‌تری از جایگاه مدل پیشنهادی ارائه می‌دهد. نتایج نشان داد که EfficientNet با وجود برخورداری از معماری عمیق و یادگیری انتقالی، تمایل شدیدی به طبقه‌بندی نمونه‌ها در کلاس سالم داشته و در تشخیص مراحل غیر سالم دچار سوگیری شده است؛ پدیده‌ای که احتمالاً ناشی از عدم سازگاری وزن‌های از پیش‌آموزش‌دیده با الگوهای خاص تصاویر شبکه‌ای است. ResNet50 نیز اگرچه در برخی نمونه‌ها عملکرد قابل‌قبولی داشته، اما در تفکیک مراحل خفیف، شدید و پیشرفته ناپایداری قابل‌توجهی از خود نشان داده است. در سوی دیگر، مدل XGBoost تنها در تشخیص کلاس سالم موفق عمل کرده و بخش قابل‌توجهی از نمونه‌های غیر سالم را به کلاس متوسط نگاشت کرده است که نشان‌دهنده ضعف این مدل در استخراج ویژگی‌های مکانی - ساختاری پیچیده تصاویر شبکه‌ای است. در مقایسه، مدل پیشنهادی توانسته است توازن بهتری میان حساسیت و تفکیک‌پذیری در کلاس‌های بالینی حساس برقرار کند. با وجود عملکرد مطلوب مدل پیشنهادی در مراحل پیشرفته بیماری، برخی محدودیت‌ها همچنان پابرجا هستند که مستقیماً بر نتایج مشاهده‌شده تأثیر گذاشته‌اند. عدم توازن ذاتی داده‌ها در کلاس‌های شدید و پیشرفته، حتی با وجود استفاده از افزایش داده و تابع زیان فوکال، می‌تواند منجر به نوسان در دقت طبقه‌بندی این مراحل شود. همچنین، اتکای مدل به یک کانال تصویری واحد، اگرچه موجب تمرکز بهتر بر عروق شبکه‌ای شده است، اما ممکن است باعث از دست رفتن اطلاعات مکمل موجود در سایر کانال‌ها شود. این محدودیت‌ها نشان می‌دهد که بهبود عملکرد مدل در آینده نیازمند رویکردهای ترکیبی‌تر در سطح داده و معماری است. در مجموع، نتایج این مطالعه حاکی از آن است که مدل پیشنهادی به‌ویژه در تشخیص مراحل بالینی بحرانی، نسبت به مدل‌های مقایسه‌شدن برتری نسبی دارد، اما برای دستیابی به یک سامانه بالینی پایدار و قابل‌تعمیم، توسعه‌های بیشتری ضروری است.

فهرست مراجع:

- [1] V. Gulshan *et al.*, "Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs," *JAMA*, vol. 316, no. 22, p. 2402, Dec. 2016, doi: 10.1001/jama.2016.17216.
- [2] D. S. W. Ting *et al.*, "Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes," *JAMA*, vol. 318, no. 22, p. 2211, Dec. 2017, doi: 10.1001/jama.2017.18152.
- [3] M. D. Abràmoff, P. T. Lavin, M. Birch, N. Shah, and J. C. Folk, "Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices," *NPJ Digit Med*, vol. 1, no. 1, p. 39, Aug. 2018, doi: 10.1038/s41746-018-0040-6.
- [4] NHS England, "Diabetic eye screening programme: standards," NHS England.
- [5] D. S. W. Ting, Y. Liu, P. Burlina, X. Xu, N. M. Bressler, and T. Y. Wong, "AI for medical imaging goes deep," *Nat Med*, vol. 24, no. 5, pp. 539–540, May 2018, doi: 10.1038/s41591-018-0029-3.
- [6] Kaggle, "Blindness-detection."
- [7] R. Gargeya and T. Leng, "Automated Identification of Diabetic Retinopathy Using Deep Learning," *Ophthalmology*, vol. 124, no. 7, pp. 962–969, Jul. 2017, doi: 10.1016/j.ophtha.2017.02.008.
- [8] L. Dai *et al.*, "A deep learning system for predicting time to progression of diabetic retinopathy," *Nat Med*, vol. 30, no. 2, pp. 584–594, Feb. 2024, doi: 10.1038/s41591-023-02702-z.
- [9] V. Y. Wang, M.-T. Lo, T.-C. Chen, C.-H. Huang, A. Huang, and P.-C. Wang, "A deep learning-based ADRPPA algorithm for the prediction of diabetic retinopathy progression," *Sci Rep*, vol. 14, no. 1, p. 31772, Dec. 2024, doi: 10.1038/s41598-024-82884-9.
- [10] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *J Big Data*, vol. 6, no. 1, p. 60, Dec. 2019, doi: 10.1186/s40537-019-0197-0.
- [11] M. Buda, A. Maki, and M. A. Mazurowski, "A systematic study of the class imbalance problem in convolutional neural networks," *Neural Networks*, vol. 106, pp. 249–259, Oct. 2018, doi: 10.1016/j.neunet.2018.07.011.
- [12] P. Costa *et al.*, "End-to-End Adversarial Retinal Image Synthesis," *IEEE Trans Med Imaging*, vol. 37, no. 3, pp. 781–791, Mar. 2018, doi: 10.1109/TMI.2017.2759102.
- [13] A. Jungo, O. Scheidegger, M. Reyes, and F. Balsiger, "pymia: A Python package for data handling and evaluation in deep learning-based medical image analysis," *Comput Methods Programs Biomed*, vol. 198, p. 105796, Jan. 2021, doi: 10.1016/j.cmpb.2020.105796.
- [14] V. K. R. Poranki and B. S. Rao, "DRG-Net: Diabetic Retinopathy Grading Network using Graph Learning with Extreme Gradient Boosting Classifier," *Informatica*, vol. 48, no. 2, May 2024, doi: 10.31449/inf.v48i2.5078.

- [15] N. Shelke, R. Somkunwar, A. Pimpalkar, S. Maurya, and S. Chhabria, "Ensemble EfficientNet: a novel technique for identification, classification and prediction of diabetic retinopathy," *Automatika*, vol. 66, no. 3, pp. 543–558, Jul. 2025, doi: 10.1080/00051144.2025.2514884.
- [16] T. Li, Y. Gao, K. Wang, S. Guo, H. Liu, and H. Kang, "Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening," *Inf Sci (N Y)*, vol. 501, pp. 511–522, Oct. 2019, doi: 10.1016/j.ins.2019.06.011.
- [17] B. Ramasubramanian and S. Selvaperumal, "A comprehensive review on various preprocessing methods in detecting diabetic retinopathy," in *2016 International Conference on Communication and Signal Processing (ICCSP)*, IEEE, Apr. 2016, pp. 0642–0646. doi: 10.1109/ICCSP.2016.7754220.
- [18] M. Alshahrani, M. Al-Jabbar, E. M. Senan, I. A. Ahmed, and J. A. M. Saif, "Hybrid Methods for Fundus Image Analysis for Diagnosis of Diabetic Retinopathy Development Stages Based on Fusion Features," *Diagnostics*, vol. 13, no. 17, p. 2783, Aug. 2023, doi: 10.3390/diagnostics13172783.
- [19] C. Mohanty *et al.*, "Using Deep Learning Architectures for Detection and Classification of Diabetic Retinopathy," *Sensors*, vol. 23, no. 12, p. 5726, Jun. 2023, doi: 10.3390/s23125726.
- [20] T. R. Gadekallu, N. Khare, S. Bhattacharya, S. Singh, P. K. R. Maddikunta, and G. Srivastava, "Deep neural networks to predict diabetic retinopathy," *J Ambient Intell Humaniz Comput*, vol. 14, no. 5, pp. 5407–5420, May 2023, doi: 10.1007/s12652-020-01963-7.
- [21] G. Alwakid, W. Gouda, and M. Humayun, "Deep Learning-Based Prediction of Diabetic Retinopathy Using CLAHE and ESRGAN for Enhancement," *Healthcare*, vol. 11, no. 6, p. 863, Mar. 2023, doi: 10.3390/healthcare11060863.
- [22] S. Venkatesan, G. Sundaramali, R. Velu, and N. Dilip Raja, "Evaluation of lead time for contour and roundness measuring equipment," *Mater Today Proc*, vol. 46, pp. 7115–7118, 2021, doi: 10.1016/j.matpr.2020.10.372.
- [23] Retina images to detect diabetic retinopathy, "Retina images to detect diabetic retinopathy," Kaggle. Accessed: Dec. 12, 2025. [Online]. Available: <https://www.kaggle.com/datasets/sovit Rath/diabetic-retinopathy-224x224-2019-data>
- [24] A. Bilal, L. Zhu, A. Deng, H. Lu, and N. Wu, "AI-Based Automatic Detection and Classification of Diabetic Retinopathy Using U-Net and Deep Learning," *Symmetry (Basel)*, vol. 14, no. 7, p. 1427, Jul. 2022, doi: 10.3390/sym14071427.