



تشخیص بیماری‌های ماکولای چشم از روی تصاویر OCT

محمدامین اسدی‌نیا^۱، سید مجتبی میرقادری^۲، سید امین ناجی اصفهانی^{۳*}، محمد کاظمی^۴

دانشگاه اصفهان، m.a.asadnia@eng.ui.ac.ir

دانشگاه صنعتی اصفهان، m.mirghaderi@ec.iut.ac.ir

دانشگاه صنعتی خواجه نصیرالدین طوسی، s.najiesfahani@email.kntu.ac.ir

دانشگاه اصفهان، m.kazemi@eng.ui.ac.ir

*سرپرست تیم

چکیده - تصویر شبکیه چشم با تکنیک مقطع نگاری همدوسی اپتیکی (OCT) می‌تواند برای تشخیص و پیش‌بینی بیماری‌های متعددی مورد استفاده قرار گیرد که تمرکز در این پژوهش، تشخیص بیماری ورم ماکولای دیابتی (DME) از افراد سالم و یا دارای سایر بیماری‌های مرتبط است. پردازش ماشینی این تصاویر از دیرباز مورد توجه بوده است، اما با توسعه شبکه‌های عصبی عمیق و ارائه نتایج بسیار خوب آن‌ها، این زمینه بیش از پیش اهمیت پیدا می‌کند. در این پژوهش، شبکه ارائه شده در [۱] که یک ساختار کانولوشنی اصلاح شده است برای داده‌های ارائه شده در چالش و تولید خروجی‌های مدنظر مورد آموزش قرار گرفت. دقت شبکه با استفاده از داده‌های در دسترس چالش به بالای ۹۵ درصد می‌رسد.

کلیدواژه - تصاویر OCT شبکیه، شبکه‌های عصبی کانولوشنی، بیماری‌های ماکولای چشم

به‌صورت کوتاه در زیر بخش بعدی مرور شده‌اند. طبیعتاً عمده‌ی تلاش‌ها برای بهبود دقت تشخیص این الگوریتم‌ها بوده است، چراکه حتی یک مورد تشخیص درست بیشتر در این سیستم‌ها ارزشمند است. یکی از بهترین روش‌ها در مقاله [۱] ارائه شده است که برای چالش مد نظر پاسخ بسیار خوبی دارد و می‌تواند بیماری‌های ماکولای چشم و افراد سالم را براساس توضیحات چالش کلاس‌بندی کند. این شبکه روی داده‌های در اختیار قرار گرفته برای چالش آموزش داده می‌شود که جزئیات آن در قسمت‌های بعدی آورده شده است. قسمت‌های بعدی مقاله به این‌صورت است که در زیر بخش بعدی مرور کوتاهی از کارهای انجام شده در این حوزه ارائه می‌شود. در بخش دوم، روش استفاده شده برای حل چالش بیان و جزئیات آن توضیح داده می‌شود. در قسمت سوم، نتایج عملکرد شبکه آموزش داده شده ارائه می‌گردند و در نهایت مقاله در بخش چهارم نتیجه‌گیری می‌شود.

۱-۲- مرور کارهای پیشین

اولین رویکرد برای تشخیص و طبقه‌بندی بیماری‌های شبکیه از روی تصاویر شامل تکنیک‌های پردازش تصویر متعدد و به دنبال آن استخراج و طبقه‌بندی ویژگی بود [۷]. یکی از این روش‌های خودکار شامل یافتن ناهنجاری‌هایی مانند آنوریسم‌های میکرو، خونریزی‌ها، آگزودا و لکه‌های پشمی از تصاویر فوندوس شبکیه بود [۸]. این رویکرد از یک الگوریتم کاهش نویز و محوکردن استفاده می‌کند تا مسئله با چهار کلاس متفاوت را به دو مورد از یک مشکل دو کلاسه تقسیم کند. از آنجا به بعد، تفاضل پس‌زمینه و به دنبال آن تخمین شکل برای استخراج

۱- مقدمه

تشخیص بیماری‌ها با استفاده از الگوریتم‌های پردازش تصویر و یادگیری ماشین همواره مورد توجه پژوهشگران بوده است که می‌تواند جایگزین پزشک و یا به عنوان راهنما برای پزشک بسیار سودمند باشد. در تصاویری مانند OCT که ساختار پیچیده‌ای دارند و بررسی آن‌ها با چشم پزشک در برخی موارد بسیار سخت است، الگوریتم‌های ماشینی بیشتر مورد استفاده قرار می‌گیرند. قسمت ماکولای شبکیه چشم می‌تواند خود دچار مشکل گردد و یا بر اثر عوارض بیماری‌های دیگر آسیب ببیند. از بیماری‌های ماکولای چشم می‌توان به ورم دیابتی ماکولا^۱، تخریب ماکولای وابسته به سن^۲، نئوواسکولاریزیشن کروئیدال^۳ و سوراخ ماکولا^۴ اشاره کرد. در این پژوهش تمرکز بر تشخیص بیماری DME و AMD از بقیه بیماری‌های مرتبط و همچنین افراد سالم، به کمک پردازش تصاویر OCT است.

۱-۱- پیش‌گفتار

می‌توان گفت که بیش از ۲۰ سال است که کاربرد روش‌های مبتنی بر هوش مصنوعی برای تشخیص بیماری‌های ماکولای چشم مورد توجه قرار گرفته است و پیشرفت‌های قابل توجهی را نیز تجربه کرده است که

¹ DME

² AMD

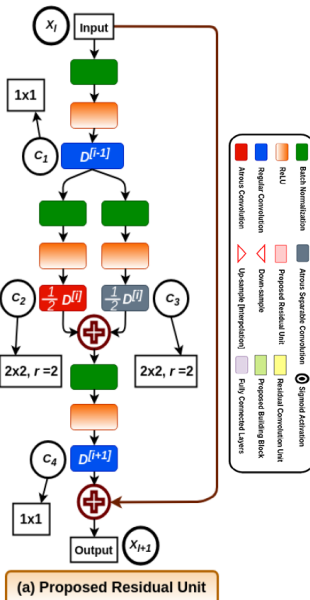
³ CNV

⁴ MH

استخراج ویژگی‌های تصویر OCT و طبقه‌بندی نهایی استفاده می‌کند. در معماری این شبکه از دو سازوکار ویژه استفاده شده است که در حقیقت نقش شبکه‌ی عصبی بازگشتی را ایفا می‌کنند. در ادامه به معرفی آن‌ها می‌پردازیم.

۱-۲-۲- سازوکار یادگیری ثانویه

این سازوکار ورودی‌های جدید را با استفاده از سه لایه‌ی کانولوشنی پردازش می‌کند؛ درحالی است که ورودی جدید با خروجی پردازش شده جمع می‌شود. این سه لایه کانولوشنی به صورت عملیات‌های کانولوشنی (1×1) ، (3×3) و (1×1) هستند که بر روی تصویر ورودی اعمال می‌شوند. جایگزینی کانولوشن (3×3) در این ساختار با عملیات‌های کانولوشنی دیگر می‌تواند به صورت بالقوه تاثیر زیادی روی نحوه یادگیری، پیچیدگی محاسباتی و عملکرد مدل داشته باشد. برای حل این چالش، از دو لایه‌ی کانولوشنی موازی به جای کانولوشن (3×3) استفاده می‌شود که کانولوشن آتروس^{۱۰} و کانولوشن آتروس جداشونده نام دارد. در نهایت نیز خروجی این دو لایه با یکدیگر جمع شده و وارد لایه‌ی بعدی می‌شوند. اضافه کردن این دو لایه باعث می‌شود شبکه‌ی طراحی شده بتواند با استفاده از پارامترهای کمتر نسبت به کانولوشن عادی به دقت بالاتری برسد. در حقیقت تفاوت کانولوشن آتروس با کانولوشن عادی در مقدار اتساع^{۱۱} شبکه خلاصه می‌شود که در بعضی از مسائل می‌تواند از مشکلات مهمی مثل بیش‌برازش و یا محوشدگی شیب^{۱۲} جلوگیری کند. در شکل ۱ ساختار یک واحد از این سازوکار نمایش داده شده است.



شکل ۱: ساختار یک واحد از سازوکار یادگیری ثانویه^[۱]

ویژگی‌های مهم استفاده می‌شود. در نهایت، از آن ویژگی‌ها برای طبقه‌بندی هر یک از چهار ناهنجاری استفاده شد. به طور مشابه، سایر تکنیک‌های مبتنی بر ویژگی برای تشخیص ماکولای دیابتی و نئوواسکولاریزاسیون مشیمیه استفاده شدند. برای مثال در [۹] تصاویر بر روی پنج پارامتر متمایز متمرکز شدند: ضخامت شبکیه، میزان ضخیم شدن شبکیه، حجم ماکولا، مورفولوژی شبکیه و رابطه شبکیه شیشه‌ای. روش‌های دیگر، کلاس‌بندی آماری را با الگوریتم‌های تشخیص لبه برای تشخیص لبه‌های تیز ترکیب کردند [۱۰]. الگوریتم سانچز و همکاران [۱۰] به دقت ۷۹.۶ درصد برای طبقه‌بندی DME دست یافت. رویکرد ایگ و همکارانش [۸] که شامل کلاس‌بندی‌کننده‌ی ماهالانوبیس^۵ بود، میکروآنوریزم، خونریزی، آگزودا و لکه‌های پشم پنبه را با حساسیت ۶۹، ۸۳، ۹۹ و ۸۰ درصد تشخیص داد. کاملاً مشهود است که هر یک از این تکنیک‌ها پیشرفت‌های جزئی نشان می‌دهند، اما از نظر دقت به نتایج مطلوبی نرسیدند. با این حال تمام کارهای انجام شده برای افزایش چند درصدی دقت تشخیص و کلاس‌بندی بیماری‌های ماکولای چشم سودمند است و می‌تواند به عنوان راهبردی برای طراحی شبکه‌های بهتر و دقیق‌تر باشد.

۲- روش استفاده شده برای حل چالش

۱-۲-۱- چالش‌های مسئله

تصویربرداری OCT برای تشخیص بیماری‌های چشمی انجام می‌شود که جزئیات لایه‌ی شبکیه‌ی چشم را در قالب تصویرهای متعددی استخراج می‌کند. در نتیجه لزوماً تعداد تصویرهای گرفته شده برای بیماران مختلف یکسان نیست. علاوه بر این، تصویربرداری OCT در طول زمان انجام می‌شود و برای تشخیص وجود بیماری چشمی لازم است تمام تصاویر گرفته شده بررسی شوند. به همین علت برای کلاس‌بندی داده‌ها لازم است از شبکه‌های حافظه‌دار استفاده شود و هر بیمار داخل کلاسی قرار گیرد که اکثر تصاویر OCT بیمار به آن کلاس تعلق داشته باشد.

با توجه به شباهت زیاد تصاویر OCT برای هر بیمار، استفاده تنها از شبکه‌های کلاس‌بندی معروف مثل رزنت^۶، وی‌جی‌جی^۷، افیشنت^۸ و غیره باعث بیش‌برازش^۹ و گیج‌شدن شبکه می‌شود. در نتیجه نیاز به شبکه‌ای داریم که بتواند ویژگی‌های تصویر را بهتر و دقیق‌تر استخراج کند.

۲-۲- ساختار شبکه استفاده شده^[۱]

شبکه‌ی استفاده شده برای حل این چالش از ۱۲ لایه‌ی مختلف برای

^{۱۰} Atrous

^{۱۱} Dilated

^{۱۲} Vanishing gradient

^۵ Mahalanobis

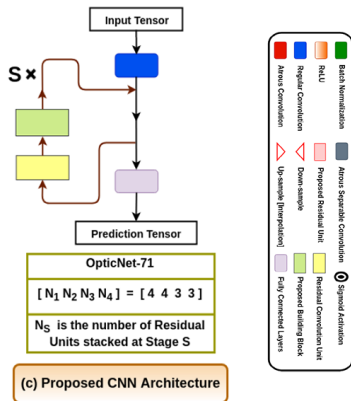
^۶ ResNet

^۷ VGG

^۸ EfficientNet

^۹ Overfitting

۲-۲-۲- واحسازی و انتشار سیگنال



شکل ۳: معماری نهایی شبکه‌ی استفاده شده^[۱]

تعداد لایه‌های کانولوشنی در ۴ واحد ثانویه به صورت (۳, ۳, ۴, ۴) است یعنی در اولین واحد ثانویه ۴ لایه کانولوشنی وجود دارد که با استفاده از روش واحسازی انباشته می‌شوند. هر لایه کانولوشنی در هر واحد ثانویه از تعداد فیلتر مشخصی برای استخراج ویژگی‌های تصویر OCT استفاده می‌کند. سپس خروجی وارد لایه‌ی ادغام‌کننده و در نهایت توسط دو لایه‌ی تماماً متصل به ۳ گروه مختلف کلاس‌بندی می‌شود که در ... اطلاعات کامل تمام لایه‌های شبکه‌ی استفاده شده برای حل این چالش آمده است.

ساختار استفاده شده و نتایج ارائه‌شده نشان می‌دهد این مدل دقت خوبی برای تشخیص انواع بیماری‌های ماکولای چشم مثل AMD، DME و غیره دارد. به همین علت برای حل این چالش همین مدل را روی داده‌های ارسال شده آموزش می‌دهیم و از مدل آموزش داده شده برای پیش‌بینی و کلاس‌بندی داده‌های تست استفاده می‌کنیم. در ادامه به توضیح فرایند آموزش و پارامترهای تنظیم‌شده می‌پردازیم.

جدول ۱: ستون آخر لایه‌های شبکه استفاده شده را معرفی می‌کند^[۱]

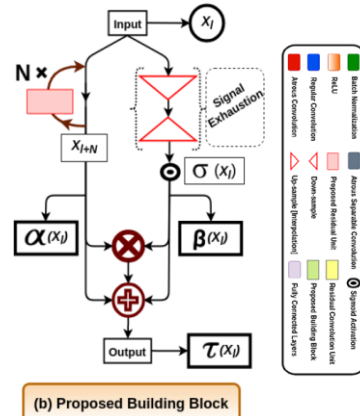
Layer Name	ResNet50 V1 [28]	OpticNet71 [Ours]
Conv 7 × 7	$[64] \times 1$	$[64] \times 1$
Stage1: Res Conv	$[64, 64, 256] \times 1$	$[64, 64, 256] \times 1$
Stage1: Res Unit	$[64, 64, 256] \times 2$	$[32, 32, 32, 256] \times 4$
Stage2: Res Conv	$[128, 128, 512] \times 1$	$[128, 128, 512] \times 1$
Stage2: Res Unit	$[128, 128, 512] \times 3$	$[64, 64, 64, 512] \times 4$
Stage3: Res Conv	$[256, 256, 1024] \times 1$	$[256, 256, 1024] \times 1$
Stage3: Res Unit	$[256, 256, 1024] \times 5$	$[128, 128, 128, 1024] \times 3$
Stage4: Res Conv	$[512, 512, 2048] \times 1$	$[512, 512, 2048] \times 1$
Stage4: Res Unit	$[512, 512, 2048] \times 2$	$[256, 256, 256, 2048] \times 3$
Global Avg Pool	2048	2048
Dense Layer 1	K (Classes)	256
Dense Layer 2	-	K (Classes)
Parameters	25.64 Million	12.50 Million
Required FLOPs	3.8×10^9	2.5×10^7
CNN Memory	98.20 MB	48.80 MB

۲-۳- آموزش شبکه

تصاویر OCT داده شده به شرکت‌کنندگان مسابقه دارای ابعاد 300×300 پیکسل هستند که برای آموزش شبکه نیاز است ابعاد تصاویر به 224×224 پیکسل کاهش پیدا کند. برای آموزش شبکه تصاویر به دو

این روش از ساختار نسبتاً پیچیده‌ای برای تولید خروجی استفاده می‌کند. در مرحله اول، ورودی وارد دو مسیر موازی می‌شود که در یکی از دو مسیر از N واحد ثانویه به صورت بازگشتی و انباشته استفاده می‌شود و در مسیر دیگر ابتدا نمونه‌ها کاهش می‌یابند و در ادامه با استفاده از درون‌یابی خطی مجدداً نمونه‌ها به مقدار اولیه خود افزایش پیدا می‌کنند. این کار در ظاهر بیهوده به نظر می‌رسد اما در عمل از محوشدگی شیب جلوگیری می‌کند. در نهایت نیز سیگنال‌های دو مسیر با بهره‌گیری از عملیات جمع و ضرب با یکدیگر ادغام می‌شوند و خروجی نهایی ایجاد می‌شود.

مزیت مهم این سازوکار زمانی مشخص می‌شود که در یکی از دو مسیر، محوشدگی یا تخریب شیب برای یک دسته کوچک^{۱۳} رخ دهد. در این صورت مسیر دیگر می‌تواند سیگنالی منتشر کند که محوشدگی یا تخریب شیب ندارد و در نهایت هنگام ادغام با مسیر دیگر، شیب نهایی را اصلاح کند. از آنجایی که تصاویر OCT برای هر بیمار شباهت زیادی با یکدیگر دارد احتمال رخداد تخریب شیب زیاد است. در نتیجه استفاده از این ساختار می‌تواند تا حد زیادی این مشکل را مرتفع سازد. در شکل ۲ نحوه‌ی واحسازی به این صورت را مشاهده می‌کنید.

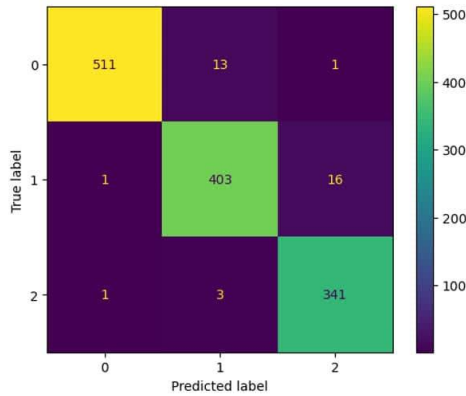


شکل ۲: ساختار سازوکار واحسازی و انتشار سیگنال^[۱]

معماری نهایی از ۴ لایه واحد ثانویه به همراه روش واحسازی استفاده می‌کند و ویژگی‌های تصویر را در سطح پیکسل استخراج می‌کند. در معماری کلی شبکه نمایش داده شده است.

¹³ Mini-batch

در دو کلاس دیگر قرار می‌دهد ماتریس درهم‌ریختگی^{۲۳} را برای داده‌های صحت‌سنجی رسم کردیم که در شکل ۶ نمایش داده شده است.



شکل ۶: ماتریس درهم‌ریختگی برای داده‌های صحت‌سنجی

همان‌طور که از ماتریس درهم‌ریختگی مشخص است کمترین خطای مدل هنگام کلاس‌بندی تصاویر در کلاس ۲ ام است. خطای این دسته تنها ۴ مورد در ۳۴۵ تصویر است.

نکته‌ی مهم این است که برای کلاس‌بندی هر بیمار لازم است تمام تصاویر مربوط به آن بیمار با استفاده از شبکه پیش‌بینی شود. در نهایت بیمار به کلاسی تعلق می‌گیرد که تعداد بیشتری از تصاویر در یک کلاس خاص باشد. برای مثال اگر بیماری ۲۰۰ تصویر OCT داشته باشد و ۱۵۰ تصویر آن در کلاس ۱ قرار گیرد؛ در نتیجه بیمار در کلاس ۱ قرار می‌گیرد.

۳- نتایج شبکه روی داده‌های تست

همان‌طور که گفته شد برای پیش‌بینی کلاس هر بیمار ابتدا تمام تصاویر مربوط به بیمار توسط شبکه پیش‌بینی می‌شود و در یکی از سه کلاس قرار می‌گیرد. در نهایت تعداد تصاویر متعلق به هر کلاس برای هر بیمار شمرده می‌شود و به تعداد کل تصاویر بیمار تقسیم می‌شود. بیمار به کلاسی تعلق خواهد داشت که مقدار به‌دست آمده برای آن کلاس بزرگ‌تر باشد. برای مثال پیش‌بینی شبکه برای ۵ سوژه اول در جدول ۲ آمده است. برای مثال شبکه برای داده‌ی اول تست پیش‌بینی می‌کند که با احتمال ۵۰ درصد به کلاس دوم تعلق دارد. به همین ترتیب هرچه این مقدار بزرگ‌تر باشد نشان می‌دهد شبکه به صورت قاطع‌تری کلاس آن داده را پیش‌بینی کرده است.

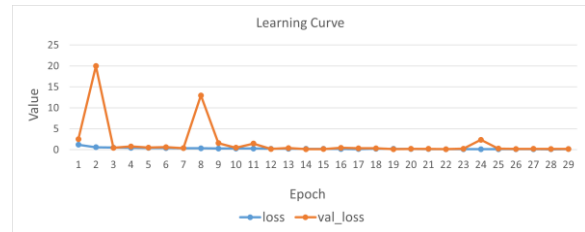
جدول ۲: خروجی برای ۵ مورد اول از داده‌های تست داده شده

case 1	case 2	case 3	case 4	case 5
50%	71%	93%	85%	94%
2	1	0	0	1

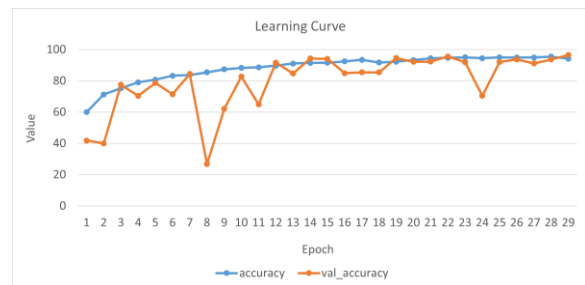
²³ Confusion matrix

دسته آموزش و صحت‌سنجی^{۱۴} تقسیم می‌شوند که داده‌های آموزش شامل ۱۵۰۰۰ تصویر و داده‌های صحت‌سنجی شامل ۱۳۰۰ تصویر است. در نهایت شبکه با اندازه دسته^{۱۵} برابر ۱۶ و برای ۳۰ دوره^{۱۶} آموزش می‌بیند. مقدار آنتروپی متقابل کلاس‌بندی شده^{۱۷} به عنوان معیاری برای محاسبه تابع ضرر^{۱۸} استفاده می‌شود که با استفاده از الگوریتم آدام^{۱۹} و نرخ یادگیری^{۲۰} برابر ۰.۰۰۰۱ بهینه‌سازی می‌شود. در نهایت هم تابع فعال‌ساز^{۲۱} سافت‌مکس^{۲۲} احتمال قرار گرفتن هر تصویر در هر کلاس را محاسبه و به عنوان خروجی نهایی تحویل می‌دهد. پس از پایان فرایند آموزش، دقت روی داده‌های آموزش بالای ۹۵ درصد و دقت روی داده‌های صحت‌سنجی بالای ۹۳ درصد است. همچنین مقدار ضرر در بهترین حالت ۰.۱۲ برای داده‌های آموزش و کمتر از ۰.۱۹ برای داده‌های صحت‌سنجی محاسبه شده است. نمودار یادگیری پس از پایان فرایند آموزش را در شکل ۴ و شکل ۵ به ترتیب برای میزان ضرر و دقت شبکه مشاهده می‌کنید.

بهترین وزن برای پیش‌بینی نهایی به نحوی از بین وزن‌های محاسبه شده انتخاب می‌شود که مدل در کمترین میزان ضرر و بالاترین دقت باشد. در این مسئله وزن‌های دوره‌ی ۳۰ ام به عنوان وزن‌های نهایی برای پیش‌بینی داده‌های تست انتخاب شده است.



شکل ۴: نمودار یادگیری برای میزان ضرر



شکل ۵: نمودار یادگیری برای میزان دقت

برای بررسی بهتر اینکه شبکه پیش‌بینی داده‌ها را به چه نحوی انجام می‌دهد؛ یعنی چه تعداد از تصاویر را در کلاس درست و چه تعدادی را

¹⁴ Validation
¹⁵ Batch size
¹⁶ Epoch
¹⁷ Categorical-CrossEntropy
¹⁸ Loss function
¹⁹ ADAM
²⁰ Learning Rate (LR)
²¹ Activation Function
²² Softmax



۴- نتیجه‌گیری

شبکه‌ی استفاده شده برای حل این چالش با استفاده از تکنیک‌های بازگشتی و واحدهای مختلف، می‌تواند بیماری‌های ماکولای چشمی را برای داده‌های در دسترس مسابقه تشخیص دهد و یا افراد سالم را از آنان تفکیک کند. این شبکه به دقت بالای ۹۵ درصد روی داده‌های آموزش و به دقت بالای ۹۳ درصد روی داده‌های صحت سنجی رسیده است و می‌توان گفت که شبکه‌ی استفاده شده می‌تواند به عنوان سیستمی خودکار برای شناسایی و تفکیک بیماری‌های ماکولای چشمی با استفاده از تصاویر OCT شناخته شود.

منابع

- [1] S. A. Kamran, S. Saha, A. S. Sabbir and A. Tavakkoli, "Optic-Net: A Novel Convolutional Neural Network for Diagnosis of Retinal Diseases from Optical Tomography Images," 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA), Boca Raton, FL, USA, 2019, pp. 964-971, doi: 10.1109/ICMLA.2019.00165.
- [2] X. C. MeindertNiemeijer, L. Z. K. Lee, M. D. Abramoff, and M. Sonka, "3d segmentation of fluid-associated abnormalities in retinal oct: Probability constrained graph-search-graph-cut," IEEE Transactions on Medical Imaging, vol. 31, no. 8, pp. 1521–1531, 2012.
- [3] M. Awais, H. Müller, T. B. Tang, and F. Meriaudeau, "Classification of sd-oct images using a deep learning approach," in 2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA). IEEE, 2017, pp. 489–492.
- [4] 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.
- [5] R. Kafieh, H. Rabbani, and S. Kermani, "A review of algorithms for segmentation of optical coherence tomography from retina," Journal of medical signals and sensors, vol. 3, no. 1, p. 45, 2013.
- [6] c. Lemaître, M. Rastgoo, J. Massich, C. Y. Cheung, T. Y. Wong, E. Lamoureux, D. Milea, F. Meriaudeau, and D. Sidib'e, "Classification of sd-oct volumes using local binary patterns: experimental validation for dme detection," Journal of ophthalmology, vol. 2016, 2016.
- [7] H. Nguyen, A. Roychoudhry, and A. Shannon, "Classification of diabetic retinopathy lesions from stereoscopic fundus images," in Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 'Magnificent Milestones and Emerging Opportunities in Medical Engineering' (Cat. No. 97CH36136), vol. 1. IEEE, 1997, pp. 426–428.
- [8] B. M. Ege, O. K. Hejlesen, O. V. Larsen, K. Møller, B. Jennings, D. Kerr, and D. A. Cavan, "Screening for diabetic retinopathy using computer based image analysis and statistical classification," Computer methods and programs in biomedicine, vol. 62, no. 3, pp. 165–175, 2000.
- [9] G. Panozzo, B. Parolini, E. Gusson, A. Mercanti, S. Pinackatt, G. Bertoldo, and S. Pignatto, "Diabetic macular edema: an oct-based classification," in Seminars in ophthalmology, vol. 19, no. 1-2. Taylor & Francis, 2004, pp. 13–20.
- [10] C. I. Sánchez, R. Hornero, M. I. Lopez, and J. Poza, "Retinal image analysis to detect and quantify lesions associated with diabetic retinopathy," in The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 1. IEEE, 2004, pp. 1624–1627.