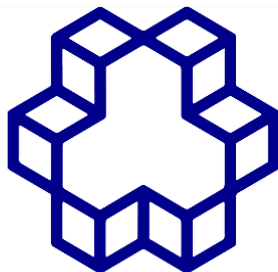


به نام خدا



دانشگاه صنعتی خواجه نصیرالدین طوسی

دانشکده مهندسی مکانیک

پایان نامه کارشناسی

گروه بایومکانیک

عنوان:

تشخیص بیماری چشم به کمک هوش مصنوعی

نگارش:

پارسا آقائی 9819603

مصطفی امیری 9819783

استاد راهنما:

دکتر مهکامه شربت دار

بهمن 1402

چکیده:

در دهه های گذشته شاهد توسعه هوش مصنوعی و کاربردهای آن بوده ایم. در سالیان اخیر محققان زیادی تلاش کرده اند تا این امر را به فرآیند های پزشکی شامل تشخیص و درمان بیماری ها نیز توسعه دهند. یکی از این زمینه های تشخیصی پزشکی، چشم پزشکی است. تشخیص بیماری چشم یکی از مهم ترین وظایف چشم پزشکان است. تشخیص بیماری چشم با استفاده از هوش مصنوعی یکی از هزاران زمینه ای است که روز به روز محققان در حال تحقیق و توسعه آن هستند.

در این پروژه بعد از معرفی بیماری های Dry AMD و Wet AMD و همچنین توضیح درباره هوش مصنوعی سعی می شود تا به کمک هوش مصنوعی، مدلی بسازیم تا عملکرد مناسبی در تشخیص این بیماری ها در مقابل چشم سالم داشته باشد. برای مدل سازی در این پروژه از شبکه های عصبی پیچشی از پیش آموزش دیده استفاده شده است. این مدل سازی بر اساس داده های جمع آوری شده از بیمارستان فوق تخصصی چشم نگاه بوده است. نتایج استخراج شده در این پروژه، نتایج قابل قبولی هستند. برای مثال دقت مدل های ساخته شده در این پروژه در بین 85 تا 95 درصد می باشد.

فهرست:

1	فصل اول: مقدمه
1-1	معرفی مسئله تحقیق:
2-1	اهمیت و ضرورت تحقیق:
3-1	هدف از تحقیق:
4-1	پیشینه پژوهش:
2	فصل دوم: بیماری های چشم
1-2	توضیح کلی بیماری:
1-1-2	دژنراسیون ماکولای خشک (Dry AMD):
2-1-2	دژنراسیون ماکولای مرطوب (Wet AMD):
3-1-2	علائم و نشانه ها:
4-1-2	علل و عوامل خطر:
5-1-2	تشخیص بیماری:
6-1-2	درمان:
7-1-2	پیشگیری:
8-1-2	نحوه نظارت بر بیماری:
2-2	تصویر برداری پزشکی از چشم:
3	فصل سوم: هوش مصنوعی
1-3	تاریخچه هوش مصنوعی:
2-3	طبقه بندی هوش مصنوعی:
3-3	شبکه های عصبی مصنوعی:
1-3-3	شبکه عصبی پیچشی (CNN):
2-3-3	شبکه عصبی پیچشی از پیش آموزش دیده شده:

29	4-3 معیار های سنجش شبکه های عصبی:
36	4.فصل چهارم: نتایج و بحث
36	1-4 دیتاست:
36	1-1-4 معرفی دیتاست:
36	2-1-4 پیش پردازش داده ها:
37	3-1-4 طبقه بندی دیتاست:
38	4-1-4 تعداد داده ها و نحوه تقسیم آن ها در آموزش و تست:
39	2-4 جزئیات و نتایج مدل سازی:
40	1-2-4 مدل سازی با استفاده از شبکه VGG19:
44	2-2-4 مدل سازی با استفاده از شبکه ResNet:
46	3-2-4 مدل سازی با استفاده از شبکه InceptionV3:
50	4-2-4 مدل سازی با استفاده از شبکه Xception:
54	3-4 مقایسه مدل های ساخته شده:
55	5.فصل پنجم: جمع بندی و نتیجه گیری
55	1-5 آموخته های مدل سازی:
55	2-5 چشم انداز آینده:

فهرست اشکال:

- شکل ۱-۱: عملکرد سیستم مقاله‌ی تشخیص آلزایمر..... ۴
- شکل ۲-۱: مشخص کردن نقاط غیرعادی مغز با کمک تشخیص اشیا..... ۵
- شکل ۳-۱: استخراج لایه‌های شبکه با کمک CNN..... ۶
- شکل ۴-۱: نمونه‌ای از سیستم‌های یاری‌رسان نابینایان به کمک عینک هوشمند..... ۶
- شکل ۵-۱: تعدادی از مقالات در زمینه AMD همراه اطلاعات نوع داده‌ها، دقت و اهدافشان..... ۷
- شکل ۱-۲: مقایسه‌ی بینایی چشم عادی و AMD خشک و مرطوب..... ۸
- شکل ۲-۲: تست شبکه Amsler برای تشخیص AMD..... ۱۱
- شکل ۳-۲: مقایسه عکس‌های OCT چشم عادی و AMD خشک و مرطوب..... ۱۴
- شکل ۴-۲: دستگاه‌های مختلف ممکن است نتیجه‌ی نسبتاً متفاوتی از یک نوع عکس (Fundus) ارائه بکنند..... ۱۶
- شکل ۵-۲: مقایسه SLO و OCT..... ۱۶
- شکل ۶-۲: مقایسه Fundus (چپ رنگی و وسط سیاه و سفید) با OCT (راست)..... ۱۶
- شکل ۱-۳: توسعه و گسترش هوش مصنوعی..... ۱۹
- شکل ۲-۳: روند یک شبکه عصبی..... ۲۱
- شکل ۳-۳: تعدادی از توابع فعال‌ساز..... ۲۱
- شکل ۴-۳: انواع لایه‌های شبکه‌های عصبی..... ۲۲
- شکل ۵-۳: نمونه‌ای از Max Pooling..... ۲۴
- شکل ۶-۳: فرایند یک شبکه CNN..... ۲۵
- شکل ۷-۳: معماری لایه‌های شبکه ResNet..... ۲۶
- شکل ۸-۳: معماری لایه‌های شبکه InceptionV3..... ۲۷
- شکل ۹-۳: معماری لایه‌های شبکه Xception..... ۲۷
- شکل ۱۰-۳: معماری لایه‌های شبکه VGG16 و VGG19..... ۲۸
- شکل ۱۱-۳: نمایی تعریفی از FP, FN, TP, TN..... ۲۹
- شکل ۱۲-۳: نمایشی از یک Confusion Matrix دارای 3 کلاس..... ۳۰
- شکل ۱۳-۳: ماتریس یک دیتاست Imbalanced و Accuracy آن..... ۳۱
- شکل ۱۴-۳: نمایی تعریفی از Overfitting و Underfitting..... ۳۳
- شکل ۱۵-۳: نمایی از Accuracy و Loss در طول Train و Validation..... ۳۳
- شکل ۱۶-۳: نحوه محاسبه‌ی Precision..... ۳۴
- شکل ۱۷-۳: نحوه محاسبه‌ی Recall..... ۳۴
- شکل ۱-۴: نمونه‌هایی از عکس‌های اولیه..... ۳۷

شکل ۴-۲: نمونه هایی از عکس ها بعد از فرآیند پیش پردازش	۳۷
شکل ۴-۳: نمونه ای از تصاویر هر کلاس از دیتاست	۳۸
شکل ۴-۴: جزئیات مدل سازی با استفاده از شبکه VGG19	۴۰
شکل ۴-۵: نمودار دقت و تابع هزینه بر حسب دوره برای داده های آموزش و تست در مدل VGG19	۴۱
شکل ۴-۶: ماتریس در هم ریختگی در مدل VGG19	۴۲
شکل ۴-۷: گزارش طبقه بندی در مدل VGG19	۴۳
شکل ۴-۸: جزئیات مدل سازی با استفاده از شبکه ResNet با فریز کردن تمامی لایه ها	۴۴
شکل ۴-۹: نمودار دقت و تابع هزینه بر حسب دوره برای داده های آموزش و تست در مدل ResNet با فریز کردن تمامی لایه ها	۴۴
شکل ۴-۱۰: ماتریس در هم ریختگی در مدل ResNet با فریز کردن تمامی لایه ها	۴۵
شکل ۴-۱۱: گزارش طبقه بندی در مدل ResNet با فریز کردن تمامی لایه ها	۴۵
شکل ۴-۱۲: جزئیات مدل سازی با استفاده از شبکه InceptionV3	۴۶
شکل ۴-۱۳: نمودار دقت و تابع هزینه بر حسب دوره برای داده های آموزش و تست در مدل InceptionV3	۴۷
شکل ۴-۱۴: ماتریس در هم ریختگی در مدل InceptionV3	۴۸
شکل ۴-۱۵: گزارش طبقه بندی در مدل InceptionV3	۴۹
شکل ۴-۱۶: جزئیات مدل سازی با استفاده از شبکه Xception	۵۰
شکل ۴-۱۷: نمودار دقت و تابع هزینه بر حسب دوره برای داده های آموزش و تست در مدل Xception	۵۱
شکل ۴-۱۸: ماتریس در هم ریختگی در مدل Xception	۵۲
شکل ۴-۱۹: گزارش طبقه بندی در مدل Xception	۵۳

فهرست جداول:

- جدول ۱-۲: مقایسه‌ی دستگاه‌های تصویربرداری چشم ۱۷
- جدول ۱-۴: نام گذاری کلاس ها ۳۸
- جدول ۲-۴: پراکندگی داده های دیتاست در هر کلاس ۳۸
- جدول ۳-۴: مشخصات لایه های اضافه شده به مدل های از پیش آموزش دیده ۳۹

1. فصل اول: مقدمه

1-1 معرفی مسئله تحقیق:

تشخیص بیماری چشم به کمک هوش مصنوعی یکی از هزاران زمینه ای است که در سالیان اخیر توسعه زیادی داشته است.

در این پروژه سعی می شود تا به کمک هوش مصنوعی، مدلی بسازیم تا عملکرد مناسبی در تشخیص بیماری های Wet AMD و Dry AMD در مقابل چشم سالم داشته باشد.

1-2 اهمیت و ضرورت تحقیق:

تشخیص بیماری چشم یکی از مهم ترین وظایف چشم پزشکان است. تسریع این فرآیند می تواند یکی از موضوعات قابل اهمیت برای پزشکان باشد. تشخیص بیماری چشم به کمک هوش مصنوعی از مواردی است که می تواند تشخیص اولیه ای قبل از مراجعه به پزشک در اختیار بیمار قرار دهد. همچنین می تواند به تشخیص صحیح پزشکان، با هدف کاهش خطای پزشکی، کمک شایانی کند.

1-3 هدف از تحقیق:

هدف از انجام این پروژه معرفی ابزاری بسیار کاربردی به پزشکان و محققان و استفاده از قابلیت های آن جهت تسریع فرآیند تشخیص بیماری های چشم است. موضوع بیماری هایی که در این پروژه مورد بررسی قرار می گیرند، بیماری های Wet AMD و Dry AMD در مقابل چشم سالم است.

1-4 پیشینه پژوهش:

استفاده از هوش مصنوعی و یادگیری ماشین برای مقاصد پزشکی بیشتر از 10 سال قدمت دارد. این زمینه اما هنوز در روند تحقیق^۱ قرار دارد و تا توسعه^۲ فاصله زیادی دارد. در واقع استفاده‌ی عمومی، گسترده و تاثیرگذار یادگیری ماشین در درمان و تشخیص بیماری‌ها هنوز به صورت کاربردی صورت نگرفته است. توسعه این موضوع اما اصلاً دور از ذهن نیست، در نتیجه تحقیقات در مورد این زمینه همچنان در حال گسترش است.

تحقیقات این زمینه برای تشخیص سرطان‌های مختلف ([1] [2] [3])، تومورهای مغزی ([4] [5])، تحلیل سیگنال‌های EEG برای تشخیص وضعیت احساسی و بینایی [6]، شکستگی‌های ستون فقرات [7]، وضعیت مینای دندان [8]، شمارش تعداد سلول‌های نمونه‌های آزمایشگاهی [9] و در نهایت انواع بیماری‌های چشم ([10] [11] [12]) انجام شده است.

قابل توجه است که مبنای تقریباً تمامی این تحقیقات و حتی روش انجام اکثر آنها به یک صورت بوده است و همه آنها از شبکه‌های عصبی پیچشی^۳ استفاده کرده‌اند. روند همه تحقیقات مورد بررسی به این صورت است که تعداد مناسبی عکس از کلاس‌های مدنظر داریم که بعد از پیش پردازش^۴ به مدل می‌دهیم. بعد یادگیری مدل اگر نیاز باشد از پسا پردازش^۵ استفاده می‌شود. برای تعیین محل تومور، شکستگی یا بیماری از تشخیص اشیاء^۶ استفاده می‌کنیم. در نهایت هم با استفاده از معیارهایی مانند Accuracy و F1 Score عملکرد شبکه را ارزیابی می‌کنیم. در ادامه تعدادی از مقاله‌های منتشرشده در این زمینه را مرور می‌کنیم تا نگاهی کلی به تحقیقات در این زمینه داشته باشیم:

در مقاله‌ای که در سال ۲۰۱۹ منتشر شده [1] دوره‌ای نسبتاً کامل از تحقیقات انجام‌شده در زمینه‌ی سرطان و هوش مصنوعی داریم. این مقاله نگاهی به نحوه‌ی پیش‌پردازش، بخش بندی^۷ و پسا پردازش، انواع روش‌های تشخیص و کلاس‌بندی^۸، کدهای موجود برای یادگیری استفاده از شبکه‌های عصبی و دیتاست‌های مرجع برای انواع سرطان‌ها می‌کند. همچنین توضیح مختصر روش‌های غیر از CNN (FCN, U-Net, GANs, RNN, AEs, RBM, LTSM) می‌دهد و در نهایت به بررسی مقالات منتشرشده در مباحث سرطان‌های سینه ([13] [14])، ریه [15]، پروستات [16]، مغز ([17] [18]) و پوست ([19] [20]) می‌پردازد.

¹ research

² Development

³ CNN network

⁴ preprocessing

⁵ Post Processing

⁶ Object detection

⁷ Segmentation

⁸ Classification

بر اساس مطالعات بررسی شده، CNN بهترین عملکرد را در بین تمام معماری ها دارد.

هنگام پیاده سازی یادگیری عمیق برای تشخیص سرطان، یکی از چالش های اصلی کمبود مجموعه داده ها است. هر الگوریتم یادگیری به مقدار زیادی داده برای آموزش جهت سنجش عملکرد نیاز دارد. بسیاری از محققان از مجموعه داده های کوچکی برای الگوریتم های خود استفاده می کنند. برای مقابله با مسئله مجموعه داده محدود بسیاری از محققان از افزایش داده^۹ استفاده می کنند که شامل تکنیک هایی مانند چرخش، برش و فیلتر برای افزایش تعداد داده های موجود است.

کنتراست پایین و SNR^{10} تصاویر پزشکی مسئول عملکرد ضعیف الگوریتم های یادگیری عمیق هستند. بنابراین، مسئله دیگر این است که اگر داده ها کنتراست پایین و SNR ضعیفی داشته باشند، چگونه می توان عملکرد مدل پیشنهادی را بهبود بخشید.

موضوع دیگری که مشاهده شده، نابرابری توزیع داده های آموزشی است. برابری داده های آموزشی بسیار مهم است. بسیاری از این مشکلات در روند تحقیق ما هم به چشم خورد. سختی به دست آوردن داده و دیتاست مناسب از بیمارستان های کشور که طبقه بندی درستی برای عکس های بیماران چشم نداشتند، شدیداً سرعت روند هرگونه تحقیق در این زمینه را کاهش می دهد. تفاوت عکس های خروجی OCT^{11} از دستگاه های مختلف که با مشکل کنتراست عکس های سرطان مقاله ی مذکور شباهت دارد. نابرابری داده ها از این لحاظ که بعضی از بیماری های بینایی بسیار کمیاب است و اگر نیاز به تحقیق در آن زمینه باشد هم زمان و امکان انجام هر پروژه ای مختل می شود. مقالات بعدی که به آن می پردازیم نگاه بهتری از تحقیقات نزدیک به کاربرد واقعی به ما می دهد. درواقع دیوار و مرز بزرگی در دنیای پزشکی جلوی این گونه کاربردها را نگرفته و توسعه و استفاده ی بهینه از آنها آنچنان دور به نظر نمی رسد.

در مقاله ی که در سال ۲۰۲۰ منتشر شده [21]، از یادگیری عمیق و CNN برای تشخیص بیماری آلزایمر استفاده شده است. در پروژه ی مذکور از عکس های ۳ بعدی MRI^{12} مغز استفاده شده است. این عکس ها از ۴۶۹ فرد گرفته شده اند و ۸۵ درصد آنها (۲۵۴۴ عکس) به داده های آموزش و ۱۵ درصد (۴۶۹ عکس) به داده های تست اختصاص داده شده است. در نهایت از 2D-CNN و 3D-CNN برای مدل استفاده شده است. در شکل 1-1 نمایی از نتایج و شکل مدل ارائه شده است:

⁹ Data Augmentation

¹⁰ Signal to Noise Ratio

¹¹ Optical Coherence Tomography

¹² Magnetic Resonance Imaging

Table 3. The comparison for accuracy, sensitivity, and specificity of three kinds of binary classification for different CNN models.

Classification	Model	Accuracy(%)	Sensitivity(%)	Specificity(%)
AD versus NC	2D-CNN	88.90 ± 4.50	84.70 ± 7.30	87.10 ± 4.50
	3D-CNN	89.40 ± 2.20	82.10 ± 7.30	84.70 ± 4.10
	3D-CNN-SVM	99.10 ± 1.13	99.80 ± 0.37	98.40 ± 1.17
AD versus MCI	2D-CNN	65.20 ± 3.40	62.40 ± 3.50	67.90 ± 9.10
	3D-CNN	86.50 ± 2.70	74.00 ± 4.20	78.90 ± 9.10
	3D-CNN-SVM	89.40 ± 6.90	86.70 ± 9.10	84.00 ± 4.80
MCI versus NC	2D-CNN	61.80 ± 6.90	55.20 ± 8.40	70.30 ± 9.83
	3D-CNN	81.20 ± 5.30	74.70 ± 8.30	80.30 ± 9.83
	3D-CNN-SVM	98.90 ± 2.78	98.90 ± 3.69	98.80 ± 0.63
Ternary Classification	2D-CNN	82.57 ± 7.35	N/A	N/A
	3D-CNN	89.76 ± 8.67	N/A	N/A
	3D-CNN-SVM	92.11 ± 2.31	N/A	N/A

شکل ۱-۱: عملکرد سیستم مقاله‌ی تشخیص آلزایمر

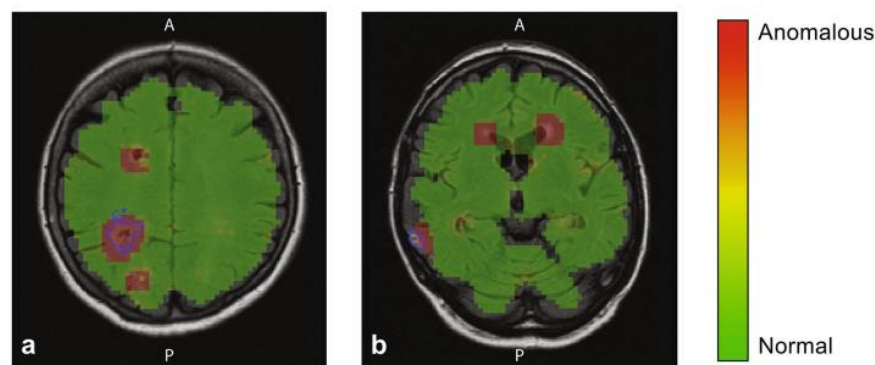
در شکل 1-1 ارزیابی عملکرد مدل رو در ۳ حالت باینری و ۱ حالت ۳ کلاسه مشاهده میکنیم.

تشخیص سنتی آلزایمر معمولاً بر اساس استخراج دستی ویژگی‌های بافت است که به زمان و منابع انسانی غیرعملی برای پیش پردازش تصویر نیاز دارد. الگوریتم استفاده شده در این مقاله می تواند به رادیولوژیست ها یا تکنسین های تصویربرداری در غربالگری و تشخیص، بدون خطر قرار گرفتن در معرض تشعشعات کمک کند. این مطالعه غربالگری آلزایمر و ¹³MCI را در یک جمعیت عمومی در ابتدا با هدف تشخیص زود هنگام و مداخله بالقوه آسان تر و کاربردی تر می کند. به طور خلاصه این مقاله و مقالات مشابه آن مشخص می کند که اتوماسیون این گونه شبکه ها برای تشخیص زودهنگام و بالقوه ی بیماری در حال حاضر کاملاً عملی و کاربردیست.

همچنین مشخص است که عملکرد 3D-CNN ها بسیار قابل توجه است. مشکل این شبکه ها در حجم بالای آنها و نیاز به سیستم های کارآمد برای عملی کردن یا به اصطلاح Run کردن آنهاست. عکس های OCT که ما از آنها در این پروژه استفاده کردیم دو بعدی و با استفاده از 2D-CNN ها مدل شده اند. مقاله ی مذکور نشان می دهد که با استفاده از عکس های ۳ بعدی امکان بهتر کردن عملکرد شبکه ی عصبی خود در این پروژه را داریم. البته جمع آوری عکس های ۳ بعدی چشم در بیمارستان ها بسیار محدودتر از عکس های ۲ بعدی است.

مقاله ی بعدی ای که به آن می پردازیم [22] با استفاده از روش تشخیص اشیا (Object Detection) آنفارکتوس ها یا سگته ها و لخته های مختلف سلولی را در مغز مشخص می کند.

¹³ Mild Cognitive Impairment



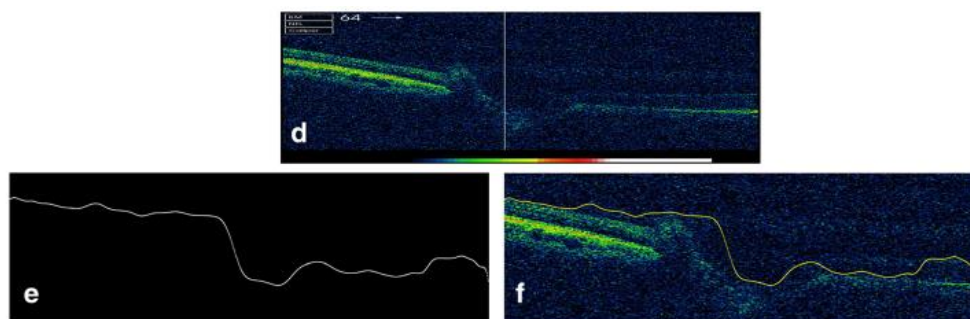
شکل ۱-۲: مشخص کردن نقاط غیرعادی مغز با کمک تشخیص اشیا

همانطور که در شکل ۱-۲ مشخص است بعد از روند همیشگی ساخت شبکه‌های عصبی می‌توان در تشخیص بیماری‌ها به صورت بصری از تشخیص اشیا هم استفاده کرد.

همچنین شبکه‌ی عصبی این مقاله تعدادی لخته‌ی مغزی پیدا کرد (۵ درصد موارد غیرعادی) که قبلاً توسط رادیولوژیست‌ها یافته نشده بودند. “که نشان می‌دهد خطاهای خواندن در ۳ تا ۵ درصد موارد در عمل رادیولوژی روزانه رخ می‌دهد. این خطاها می‌توانند به دلیل سوگیری بی‌توجهی رخ دهند، جایی که رادیولوژیست‌ها بر مرکز تصویر متمرکز می‌شوند، در حالی که یافته‌هایی که در لبه‌های تصویر به دست آمده اند را نادیده می‌گیرند [23]. از آنجایی که حجم کار در رادیولوژی به دلیل بار بیشتری از تصاویر رادیولوژی که باید ارزیابی شوند، روز به روز در حال افزایش است، نادیده گرفتن آسیب شناسی‌های مغزی بسیار شایع تر است. [24]

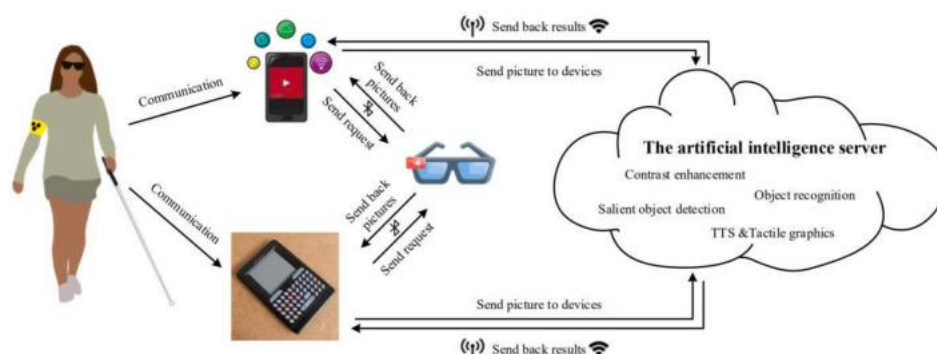
CNN همچنین به ما اجازه می‌دهد راه‌های بیشتری برای اندازه‌گیری و تشخیص رو به روی خود ببینیم. برای مثال طبقه‌بندی و تحلیل سیگنال‌های EEG [6] برای تشخیص وضعیت تمرکز و احساسی افراد یا شناخت احساسات آنها هنگام گوش دادن به موسیقی با استفاده از شبکه‌های CNN انجام می‌شود. [25]

در سال‌های اخیر تحقیقات بسیار زیادی در زمینه‌ی هوش مصنوعی و بیماری‌های چشمی و اختلالات بینایی صورت گرفته است. در مقاله‌ای [12] به صورت کلی و کامل به تعدادی از این تحقیقات و انواع آنها اشاره شده است. استفاده از CNN برای بیماری‌های دیابت ([26] [27] [28] [29])، گلوکوم یا آب سیاه ([30] [31] [32]) و CNV [10] نمونه‌هایی از آنها هستند.



شکل ۱-۳: استخراج لایه‌های شبکه‌ای با کمک CNN

همچنین ساخت دستگاه‌هایی برای کمک به افراد کم‌بینا و نابینا با کمک هوش مصنوعی به سرعت در حال افزایش است؛ سیستم‌هایی که با استفاده از سنسورهایی در عصا [33] [34]، یا شناخت اشیاء به کمک یک دوربین [35] می‌توانند رفاه زندگی این افراد را افزایش دهند.



شکل ۱-۴: نمونه‌ای از سیستم‌های یاری‌رسان نابینایان به کمک عینک هوشمند

در نهایت هدف پروژه‌ی ما تشخیص بیماری AMD و شناسایی بین ۳ کلاس چشم سالم، بیماری Wet AMD و بیماری Dry AMD است. استفاده از عکس‌های OCT برای تشخیص بیماری چشم به کمک یادگیری ماشین مسئله‌ی بسیار مورد توجهی است. برای اینکه نگاهی کلی به تحقیقات انجام‌شده در زمینه‌ی AMD داشته باشیم می‌توانیم به مقاله‌ی [11] مراجعه کنیم که تعداد زیادی از این مقالات را در جدولی گردآوری کرده.

اکثر این مقالات برای پیش‌بینی تغییرات وضعیت چشم از وضعیت Dry به Wet مربوطند [36] [37] و از دیتاست‌های از پیش‌آماده استفاده می‌کنند. پروژه‌ی ما علاوه بر اینکه موضوع متفاوتی دارد و هدفش تشخیص صرف بین این دو بیماری و چشمان عادیست، از عکس‌های OCT بیمارستان داخلی (بیمارستان نگاه) مدل شده است.

Condition	Imaging analysed	Database (n)	AI tool	Task	Performance (metrics provided by each paper)
AMD	Colour	407 eyes with nonadvanced AMD	DL	Distinguishes between low and high-risk AMD by quantifying drusen location, area, and size	For drusen area: ICC > 0.85; for diameter: ICC = 0.69; for AMD risk assessment: ROC = 0.948 and 0.954
	Colour, OCT, and IR	278 eyes with/without reticular pseudodrusen (RPD)	DL	Automatic quantification of RPD	ROC = 0.94 and 0.958; κ agreement = 0.911; ICC = 0.704
	Colour	2951 subjects from AREDS (834 progressors)	DL	Association between genetic variants and transition to advanced AMD	AUC: 5 years = 0.885; 10 years = 0.915
	Colour and OCT	280 eyes from 140 participants	DL	Prediction of progression to late AMD	AUC = 0.85
	Colour and microperimetry	280 eyes from 140 participants	DL	Predictive value of pointwise sensitivity and low luminance deficits for AMD progression	AUC = 0.8
	Colour	> 4600 participants from AREDS	DL	Predict progression to advanced dry or wet AMD	Accuracy = 0.86 (1 year) and 0.86 (2 years); specificity = 0.85 (1 year) and 0.84 (2 years); sensitivity = 0.91 (1 year) and 0.92 (2 years)
	Colour	1351 subjects from AREDS (> 31,000 images)	DL	Predict progression to advanced dry or wet AMD	AUC = 0.85
	Colour	67,401 colour fundus images from 4613 study participants	DL	Estimate 5-year risk of progression to wet AMD and geographic atrophy based on 9-step AREDS severity scale	Weighted κ scores = 0.77 for the 4-step and 0.74 for the 9-step AMD severity scales
	Colour	4507 AREDS participants and 2169 BMES participants	DL	Validation of a risk scoring system for prediction of progression	Sensitivity = 0.87; specificity = 0.73
	OCT	2795 patients	DL	Prediction to nAMD within a 6-month window	AUC = 0.74 (conversion scan ground truth) and 0.886 (1st injection ground truth)
	OCT	671 AMD fellow eyes with 13,954 observations	DL	Predict progression to wet AMD	AUC = 0.96 ± 0.02 (3 months); 0.97 ± 0.02 (21 months)
	OCT	686 fellow eyes with non-neovascular AMD at baseline	DL	Predict conversion from non-neovascular to neovascular AMD	Drusen are within 3 mm of fovea (HR = 1.45); mean drusen reflectivity (HR = 3.97)

شکل ۵-۱: تعدادی از مقالات در زمینه AMD همراه اطلاعات نوع داده‌ها، دقت و اهدافشان

در ادامه این گزارش در فصل دوم به توضیحات پزشکی بیماری‌های چشم Wet AMD و Dry AMD و روش‌های تصویر برداری پزشکی پرداخته می‌شود. در فصل سوم در رابطه با انواع هوش مصنوعی و کاربرد های آن و همچنین درباره شبکه های عصبی مصنوعی توضیحاتی ارائه می‌گردد. در فصل چهارم نتایج مدل سازی ارائه و درباره آن‌ها بحث می‌گردد. در فصل پنجم نیز جمع بندی نهایی و نتیجه گیری ارائه می‌شود.

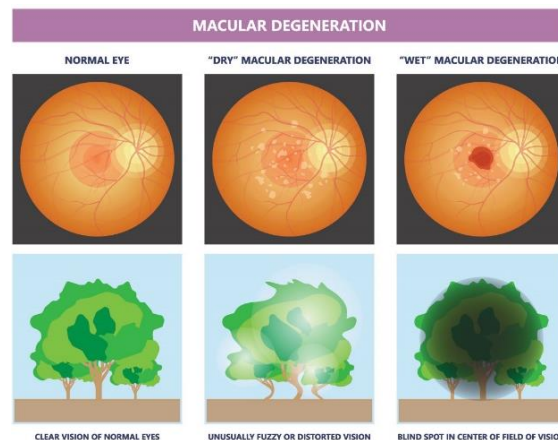
2. فصل دوم: بیماری های چشم

دژنراسیون ماکولا وابسته به سن (AMD^{14}) علت اصلی نابینایی در میان افراد مسن بالای 55 سال است. [38] انتظار می رود تعداد افراد مبتلا به AMD در طی 10 سال آینده به دلیل پیری جمعیت، فشار خون بالا و دلایل دیگر 1.5 برابر شود. [39] [40] این یک بیماری نورودژنراتیو پیشرونده است که تحت تأثیر عوامل خطر محیطی و ژنتیکی قرار دارد. [41] شدت AMD عمدتاً با تصاویر رنگی فوندوس در یک محیط بالینی توسط چشم پزشکان تشخیص داده می شود. AMD دیررس به دو شکل ظاهر می شود:

1: آتروفی جغرافیایی (GA) که به عنوان AMD خشک¹⁵ نیز شناخته می شود، با انحطاط تدریجی و ناپدید شدن اپیتلیوم رنگدانه شبکیه، سلول های گیرنده نوری و لایه کوریوکاپیلاریس در شبکیه مرکزی مشخص می شود.

2: نئوواسکولاریزاسیون مشیمیه (CNV) که به نام AMD مرطوب¹⁶ نیز شناخته می شود، با رشد عروق خونی جدید و نشستی در شبکیه مشخص می شود که باعث از دست دادن گسترده گیرنده نوری و در نهایت کاهش سریع در حدت بینایی می شود.

برخی از بیماران با مرحله اولیه یا میانی AMD بینایی خود را برای مدت طولانی با پیشرفت آهسته بیماری در طول زمان حفظ می کنند، اما برخی دیگر به سرعت به یک یا هر دو نوع AMD دیررس پیشرفت می کنند. [37] [42]



شکل 2-1: مقایسه ی بینایی چشم عادی و AMD خشک و مرطوب

¹⁴ Age related Macular Degeneration

¹⁵ Dry AMD

¹⁶ Wet AMD

2-1 توضیح کلی بیماری:

شبکیه لایه ای از سلول هاست که سطح پشتی چشم انسان را می پوشاند. شبکیه اطلاعات را به مغز منتقل می کند که بینایی انسان را ممکن می کند.

ماکولا ناحیه ای در مرکز شبکیه است. این ناحیه مسئول دید مرکزی است که برای دیدن جزئیات ضروری است. در طول AMD، سلول های زیر ماکولا تحلیل می روند و می میرند، که توانایی چشم برای دیدن مرکزی را مختل می کند. افراد مبتلا به AMD هنوز دید محیطی خود را دارند، اما در دیدن اشیاء روبروی خود مشکل دارند.

2-1-1 دژنراسیون ماکولای خشک (Dry AMD):

بین 85 تا 90 درصد موارد AMD نسخه خشک این بیماری است. [43] بیماری AMD خشک معمولاً از نظر بینایی کمتر از AMD مرطوب تاثیرگذار است و شامل رشد عروق خونی غیرطبیعی نمی شود.

ظهور بسیاری از رسوبات زرد رنگ که دروسن^{۱۷} نامیده می شوند، معمولاً نشانه ای از خشکی AMD است. دروسن حاوی مواد زائد و چربی های رسوب شده از سلول های شبکیه چشم است. با رشد و تکثیر دروسن در شبکیه، شبکیه ممکن است کمتر قادر به تولید سیگنال بصری باشد.

افرادی که AMD خشک دارند ممکن است مشکلات بینایی مانند یک نقطه تیره یا تار در بینایی مرکزی خود را تجربه کنند. با این حال، علائم افراد مبتلا به خشکی صورت AMD معمولاً کمتر از افراد مبتلا به AMD مرطوب است. تقریباً در 10٪ از افراد مبتلا به AMD خشک، وضعیت آنها ممکن است به AMD مرطوب تبدیل شود. افرادی که هر یک از علائم AMD را تجربه می کنند.

2-1-2 دژنراسیون ماکولای مرطوب (Wet AMD):

AMD مرطوب همچنین به عنوان AMD "اگزوداتیو" شناخته می شود. پزشکان این نسخه از AMD را کمتر از AMD خشک تشخیص می دهند. حدود 10 تا 15 درصد از کل موارد AMD را شامل می شود.

در AMD مرطوب، عروق خونی غیر معمول در اطراف ماکولا و شبکیه ظاهر می شوند. این رگ ها می توانند پروتئین ها یا لیپیدهای منبع قابل اعتماد را نشت کنند و همچنین ممکن است باعث ایجاد جای زخم شوند. [44]

¹⁷ Drusen

بدون درمان، اکثر افراد مبتلا به AMD مرطوب از دست دادن بینایی قابل توجه و دائمی را تجربه خواهند کرد. درمان می تواند به کاهش درجه از دست دادن بینایی کمک کند، اما همیشه از کاهش بینایی به طور کامل جلوگیری نمی کند.

2-1-3 علائم و نشانه ها:

در مراحل اولیه AMD خشک، افراد ممکن است هیچ علامتی نداشته باشند. هنگامی که AMD خشک به مرحله میانی پیشرفت می کند، برخی از علائم ممکن است شامل تاری جزئی در بینایی مرکزی یا مشکلات دیدن در فضاهای تاریک باشد.

علائم AMD خشک پیشرفته مشابه علائم AMD مرطوب است. افرادی که هر یک از این شرایط را دارند ممکن است به بیماری AMD مبتلا باشند:

- شروع به درک خطوط مستقیم به صورت موج دار کنید
- نقاط خالی را در دید خود ببینید
- با دید در نور کم مشکل بیشتری دارند
- تاری دید مرکزی را تجربه کنید

2-1-4 علل و عوامل خطر:

افرادی که سابقه خانوادگی AMD دارند بیشتر احتمال دارد که خودشان به این بیماری مبتلا شوند. تحقیقات نشان می دهد افرادی که سیگار می کشند یا بیش از حد در معرض اشعه ماوراء بنفش قرار می گیرند نیز در معرض خطر بیشتری هستند. برخی دیگر از علل شایع AMD و عوامل خطر عبارتند از:

تغذیه ضعیف

اضافه وزن

فشار خون بالا

عدم فعالیت بدنی کافی

سن مهمترین عامل خطر AMD است. افراد بالای 55 سال در معرض خطر بیشتری برای AMD هستند و این خطر با گذشت زمان افزایش می یابد. در طول فرآیند پیری، تغییراتی در شبکه رخ می دهد که احتمال ایجاد AMD را افزایش می دهد.

در نهایت، برخی از مطالعات نشان داده اند که AMD در بین زنان شایع تر است.

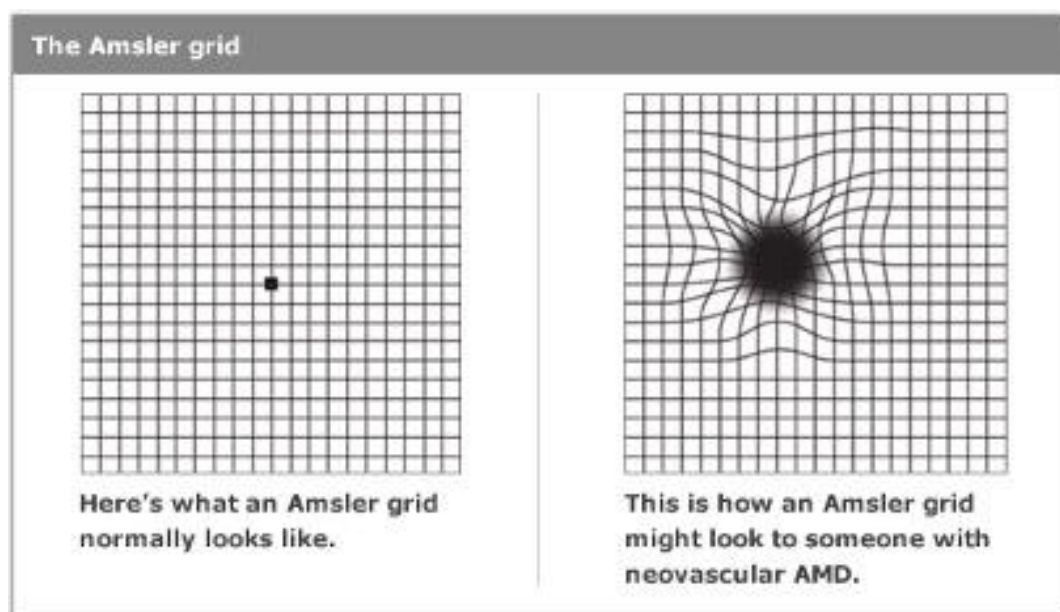
2-1-5 تشخیص بیماری:

افراد مبتلا به AMD خشک تغییراتی را در لایه سلولی زیر شبکه نشان خواهند داد. آنها همچنین احتمالاً دارای رسوبات دروسن در چشم خواهند بود و ممکن است به خود شبکه آسیب برسد.

افراد مبتلا به AMD مرطوب ممکن است تجمع مایع در اطراف شبکه و همچنین رسوبات در اطراف ماکولا داشته باشند. برخی از این افراد ممکن است به دلیل خونریزی در زیر ماکولا، تغییر رنگ خاکستری یا سبز نشان دهند. در نهایت، ممکن است شواهدی از خونریزی در داخل یا نزدیک ماکولا وجود داشته باشد.

برای تشخیص AMD، یک متخصص پزشکی احتمالاً از تست شبکه Amsler استفاده می کند. در طول این آزمایش، فرد شبکه ای از خطوط مستقیم را بررسی می کند. اگر هر یک از این خطوط منحنی یا مخدوش به نظر برسد، ممکن است فرد علائم اولیه AMD را نشان دهد.

شکل 2-2 نمایی از تست شبکه Amsler برای تشخیص AMD را نشان می دهد.



شکل ۲-۲: تست شبکه Amsler برای تشخیص AMD

شبکه را در فاصله 12 اینچی از صورت خود نگه دارید و هر بار یک چشم را بپوشانید، به نقطه مرکز شبکه خیره شوید. اگر به نظر می رسد هر نقطه تاری وجود داشته باشد، یا اگر خطوط ناهنجار به نظر می رسند، ممکن است AMD وجود داشته باشد.

یکی دیگر از تست های تشخیصی رایج آنژیوگرافی فلورسین فوندوس (FFA) است. FFA شامل تزریق رنگی است که به داخل رگ های خونی چشم می رود. این آزمایش از تشعشعات یونیزان استفاده نمی کند. این رنگ می تواند به پزشکان کمک کند تا تشخیص دهند که آیا نشتی در رگ های خونی چشم وجود دارد یا خیر. اگر چنین است، این نشانه واضحی از AMD مرطوب است.

2-1-6 درمان:

تغییرات سبک زندگی مانند ترک سیگار و رژیم های پیشگیرانه مانند مکمل های ویتامین برای بیماران در معرض خطر به منظور کاهش سرعت پیشرفت بیماری توصیه می شود. [45]

در حال حاضر هیچ درمانی برای درمان AMD خشک وجود ندارد. [46] [47] با توجه به تأثیر آن بر عوارض و عدم وجود درمان های اثبات شده، آتروفی جغرافیایی یک چالش بالینی مداوم است و درمان های اصلاح کننده بیماری برای AMD خشک یک زمینه تحقیقاتی مهم است. [48] [49]

با این حال، مطالعات نشان می دهد که مصرف برخی ویتامین ها می تواند به جلوگیری از بدتر شدن AMD خشک کمک کند. این ویتامین ها عبارتند از:

ویتامین C و E، فلز روی، بتاکاروتن، فلز مس

افرادی که مبتلا به AMD مرطوب هستند، گزینه های درمانی بیشتری دارند. از آنجایی که AMD مرطوب نتیجه نشت غیرطبیعی رگ های خونی است، می توان آن را با داروهای ضد فاکتور رشد اندوتلیال عروقی (VEGF) درمان کرد. این داروها را می توان برای جلوگیری از خونریزی یا نشت رگ های خونی غیرطبیعی به چشم تزریق کرد. درمان با تزریق داخل چشمی فاکتور رشد اندوتلیال ضد عروقی (anti-VEGF) موثر است و خطر از دست دادن بینایی ناشی از شکل نئوواسکولار AMD دیررس را کاهش می دهد، اما هزینه ها و تعداد تزریق های مورد نیاز، بار قابل توجهی بر سیستم های مراقبت های بهداشتی و درمانی ایجاد می کند. [50] [51] [52]

2-1-7 پیشگیری:

مشکلات AMD از آسیبی که به سلول های زیر شبکیه روی می دهد به وجود می آیند. در نتیجه، هر استراتژی که بتواند از چشم ها در برابر این نوع آسیب محافظت کند، می تواند به کاهش خطر ابتلا به AMD در فرد کمک کند.

از آنجایی که سیگار یکی از بزرگترین عوامل خطر AMD است، ترک سیگار یا شروع نکردن آن می تواند به پیشگیری از AMD کمک کند.

ورزش منظم همچنین می تواند به کاهش خطر AMD کمک کند. حفظ یک سبک زندگی سالم می تواند به افراد کمک کند فشار خون و کلسترول خود را کاهش دهند و خطر ابتلا به AMD را بیشتر کاهش دهند.

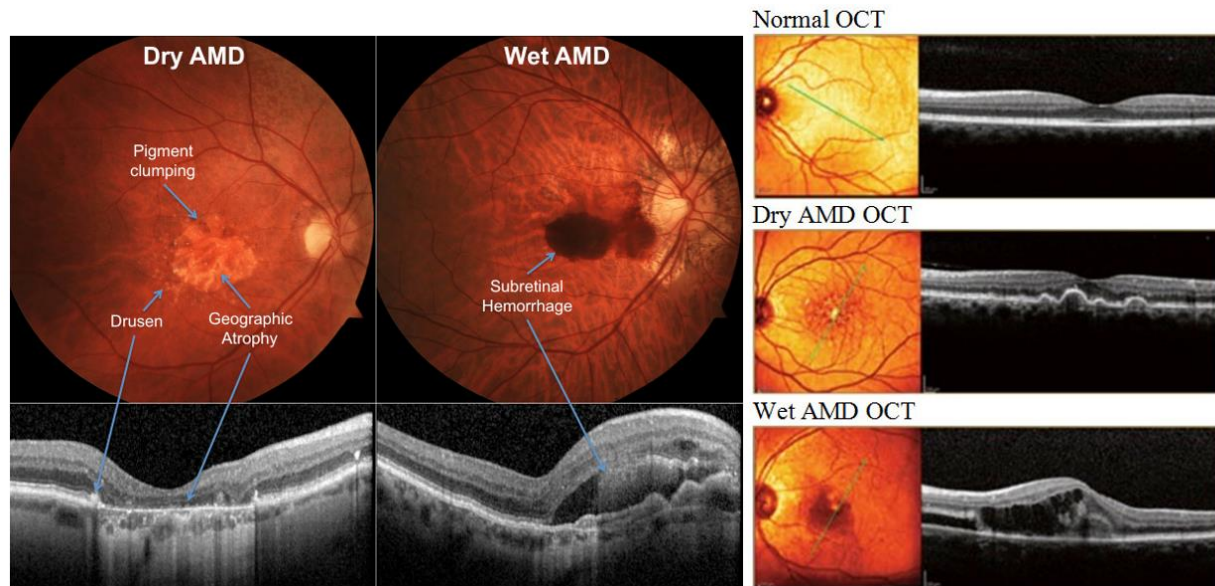
2-1-8 نحوه نظارت بر بیماری:

شناسایی و طبقه بندی چشم ها با AMD عمدتاً با استفاده از تصاویر رنگی فوندوس با تعیین دستی اندازه و گسترش دروسن انجام می شود ([53] [54])، با این حال، روش های تصویربرداری دیگر، مانند توموگرافی انسجام نوری (OCT)، نیز در حال افزایش است. [55] طبقه بندی توسط ناظر انسانی (رادیولوژیست یا دکتر و متخصص) زمان بر است و مستعد تغییرات بین ناظران است. [56]

توموگرافی انسجام نوری تصاویر مقطعی از شبکیه را ارائه می کند که امکان شناسایی مایع و سایر ویژگی های پاتومورفولوژیک مانند دروسن، مواد هیپرانعکاسی زیر شبکیه (SHRM) یا آتروفی را فراهم می کند. [57] علاوه بر این، وسعت (اندازه و توزیع) این ویژگی ها را می توان از OCT اندازه گیری کرد و اطلاعات تشخیصی و پیش آگهی بهتری ارائه می دهد. [58] بنابراین، تصمیمات درمانی در روال بالینی فعلی تا حد زیادی بر پارامترهای تخمینی متکی است، که به ویژه در مقایسه بازدیدهای بعدی، مستعد خطا و سوگیری هستند. ([59] [60])

دروسن ماکولا در زمینه درجه بندی AMD مهم است و برخی ویژگی های دروسن با پیشرفت به سمت مرحله نهایی AMD مرتبط هستند. [61] در عکاسی فوندوس آنها به صورت لکه های زرد مایل به سفید ظاهر می شوند و فنوتیپ های مختلف دروسن را می توان تشخیص داد. [62]

در شکل 2-3 جمع شدن لکه های زرد و سفید دروسن و خونریزی را در AMD مرطوب میتوان مشاهده کرد.



شکل ۲-۳: مقایسه عکسهای OCT چشم عادی و AMD خشک و مرطوب

2-2 تصویر برداری پزشکی از چشم:

انواع مختلفی از دستگاه های تصویربرداری چشم در چشم پزشکی و بینایی سنجی برای تشخیص و نظارت بر بیماری های چشم استفاده می شود. این دستگاه ها به متخصصان مراقبت از چشم کمک می کنند تا قسمت های مختلف چشم را تجسم کرده و هرگونه ناهنجاری را شناسایی کنند. برخی از انواع رایج دستگاه های تصویربرداری چشم عبارتند از:

1. دوربین های فوندوس: این دوربین ها تصاویر دقیقی از پشت چشم، از جمله شبکیه، دیسک بینایی و رگ های خونی می گیرند. آنها اغلب برای تشخیص و نظارت بر شرایطی مانند رتینوپاتی دیابتی، دژنراسیون ماکولا و گلوکوم استفاده می شوند.

2. OCT¹⁸ (توموگرافی منسجم نوری): OCT یک تکنیک تصویربرداری غیر تهاجمی است که تصاویر مقطعی از شبکیه را ارائه می دهد و امکان تجسم دقیق لایه های آن را فراهم می کند. معمولاً برای تشخیص و مدیریت شرایطی مانند ادم ماکولا، گلوکوم و سایر بیماری های شبکیه استفاده می شود.

¹⁸ Optical Coherence Tomography

3. اسکن لیزری افتالموسکوپ SLO: این تکنیک یک تصویربرداری مبتنی بر لیزر است که تصاویری با وضوح بالا از شبکه ایجاد می کند. اغلب برای تشخیص و نظارت بر بیماری های شبکه ای، مانند دژنراسیون ماکولا مرتبط با سن و اختلالات عروقی شبکه ای استفاده می شود

4. آنژیوگرافی فلورسین^{۱۹}: این تکنیک تصویربرداری شامل تزریق یک رنگ فلورسنت به جریان خون است که رگ های خونی در شبکه ای را برجسته می کند. سپس دوربین های تخصصی تصاویر را می گیرند و به تشخیص بیماری هایی مانند دژنراسیون ماکولا، رتینوپاتی دیابتی و ناهنجاری های عروقی کمک می کنند.

5. توپوگرافی قرنیه^{۲۰}: که برای نقشه برداری از سطح قرنیه استفاده می شود، توپوگرافی قرنیه به ارزیابی انحنا و شکل قرنیه کمک می کند. برای برنامه ریزی جراحی های انکساری مانند لیزیک و تشخیص بیماری های مانند قوز قرنیه بسیار مهم است.

6. تصویربرداری اولتراسوند^{۲۱}: از سونوگرافی A-scan و B-scan می توان برای تجسم ساختارهای داخلی چشم استفاده کرد، به خصوص زمانی که روش های نوری محدود هستند. آنها اغلب در موارد محیط های غیر شفاف مانند آب مروارید یا خونریزی زجاجیه استفاده می شوند.

7. میکروسکوپ کانفوکال^{۲۲}: این تکنیک تصویربرداری از لیزر اسکن برای گرفتن تصاویر دقیق از قرنیه و سایر ساختارهای شفاف چشم در سطح سلولی استفاده می کند. در مطالعه بیماری های قرنیه و ارزیابی سلامت اندوتلیوم قرنیه ارزشمند است.

8. پاشی متری^{۲۳}: این دستگاه ضخامت قرنیه را اندازه گیری می کند که برای ارزیابی شرایطی مانند گلوکوم و برنامه ریزی برای جراحی های انکساری مهم است.

این دستگاه های تصویربرداری نقش مهمی در تشخیص زودهنگام، برنامه ریزی درمان و نظارت بر شرایط مختلف چشم ایفا می کنند و به مراقبت بهتر از بیمار در زمینه چشم پزشکی و بینایی سنجی کمک می کنند.[63]

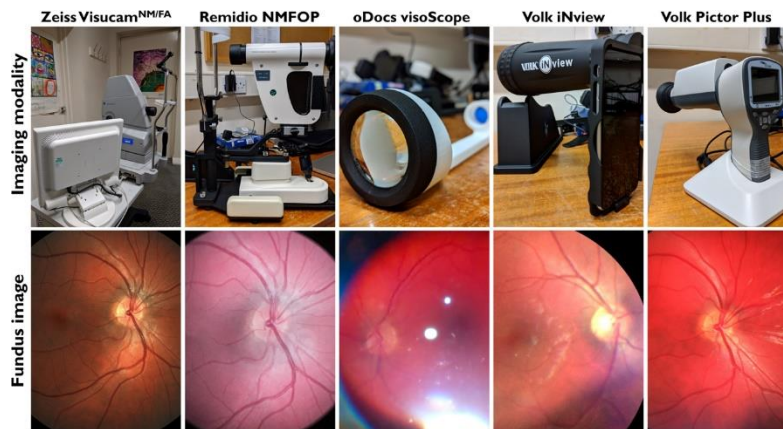
¹⁹ Fluorescein Angiography

²⁰ Corneal Topography

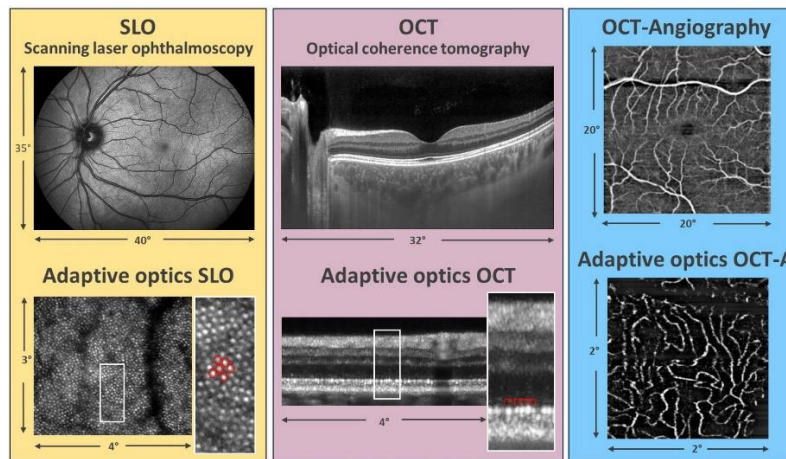
²¹ Ultrasound Imaging (A&B-scan)

²² Confocal Microscopy

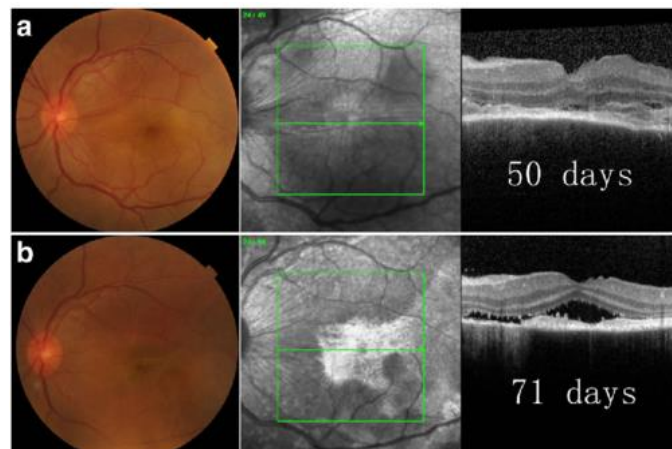
²³ Pachymetry



شکل ۲-۴: دستگاه های مختلف ممکن است نتیجه ی نسبتاً متفاوتی از یک نوع عکس (Fundus) ارائه بکنند [64]



شکل ۲-۵: مقایسه SLO و OCT [65]



شکل ۲-۶: مقایسه Fundus (چپ رنگی و وسط سیاه و سفید) با OCT (راست) [66]

جدول 1-2 مقایسه ای از دستگاه های تصویربرداری چشم ارائه می دهد.

جدول ۲-۱: مقایسه‌ی دستگاه‌های تصویربرداری چشم

روش ثبت تصویر	معایب	مزایا	کاربرد	دستگاه تصویربرداری
عکاسی دیجیتال غیر تهاجمی از پشت چشم	نیاز به گشودن دهانه‌های زیاد، دید محدود	تصاویر با دقت بالا از شبکیه	بررسی شبکیه، تصویربرداری عروق	دوربین Fundus
اسکن بر اساس تداخل نوری از نور بازتابی انجام می‌شود	تجهیزات گرانقیمت، عمق نفوذ محدود	دیدهای مقطع، رزولوشن بالا	تصویربرداری شبکیه و ماکولا	توموگرافی نوری
اسکن لیزر برای مشاهده دقیق شبکیه	دید محدود، تکنولوژی پیچیده	تصاویر با کیفیت بالا، کنتراست افزوده شده	تصویربرداری دقیق شبکیه	اس ال او
تصاویر متوالی برداری از فلورسانس در طول زمان	تهاجمی به دلیل تزریق رنگ، احتمال عوارض جانبی	نمایش ناهنجاری‌های عروق	تصویربرداری عروق خونی در شبکیه	آنژیوگرافی فلورسین
پروژکشن حلقه‌ها بر روی کره چشم و تجزیه و تحلیل الگوهای بازتاب	محدود به سطح کره چشم، برای ساختارهای داخلی مناسب نیست	ارزیابی دقیق انحنای کره چشم	نقشه‌برداری سطح کره چشم برای جراحی لیزیک	توپوگرافی کرنی
استفاده از امواج فولتاژ، با A-scan برای طول محور و B-scan برای تصاویر مقطع	تهاجمی، رزولوشن محدود، وابسته به عامل انسانی	نفوذ به واسطه‌های مات، تصویربرداری به صورت زمان واقعی	مشاهده ساختار داخلی چشم	تصویربرداری اولتراسوند (A&B-scan)
اسکن لیزر برای تصویربرداری دقیق از ساختارهای سلولی	محدود به ساختارهای سطحی، تجهیزات گران‌قیمت	تصاویر با رزولوشن بالا از بافت‌های شفاف	تصویربرداری سلولی از کره چشم	میکروسکوپی هموار
استفاده از امواج فولتاژ یا روش‌های نوری برای اندازه‌گیری ضخامت کره چشم	محدود به ضخامت کره چشم، برای ساختارهای دقیق مناسب نیست	غیر تهاجمی، اندازه‌گیری سریع	اندازه‌گیری ضخامت کره چشم	پاکیمتری

3. فصل سوم: هوش مصنوعی

هوش مصنوعی (AI) هوشی است که توسط ماشین‌ها ظهور پیدا می‌کند، در مقابل هوش طبیعی که توسط جانوران شامل انسان‌ها نمایش می‌یابد. اما پیش از هر چیز باید این موضوع را دانست که کلمه هوش، نشان دهنده امکان استدلال است و این که آیا هوش مصنوعی می‌تواند به توانایی استدلال دست یابد یا خیر، خود موضوع اختلاف محققان است. هوش مصنوعی به توانایی ماشین‌ها برای انجام کارهایی که نیازمند هوش انسانی است، اطلاق می‌شود. هوش مصنوعی یکی از جالب‌ترین حوزه‌های تحقیقاتی در چند دهه اخیر بوده است. تاریخچه هوش مصنوعی قدمت زیادی دارد و از دهه ۱۹۵۰ میلادی آغاز شد. مفهوم هوش مصنوعی به عنوان ایده انتقال هوش انسانی به ماشین‌ها گسترش پیدا کرد. این ایده مبهم، از آن زمان تاکنون توسعه زیادی را تجربه کرده است، و بنابراین، به ابزاری درخشان در طراحی‌های دنیای مدرن تشخیص بیماری‌ها، ربات‌های هوشمند، وسایل نقلیه، اندام مصنوعی و ... تبدیل شده است.

در این فصل تاریخچه مختصری از هوش مصنوعی بیان خواهد شد. در ادامه توضیحات مفصلی در مورد هوش مصنوعی و شاخه‌های مختلف آن ارائه شده است.

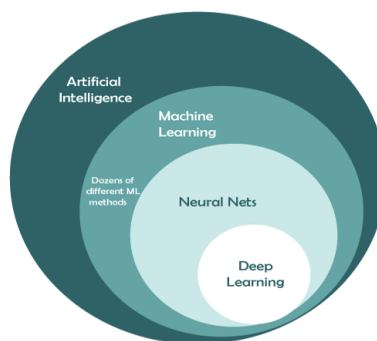
3-1 تاریخچه هوش مصنوعی:

علیرغم این واقعیت که مفهوم هوش مصنوعی برای اولین بار در دهه 1950 بیان شد، اولین افکار در مورد وجود موجودی باهوش (به طور دقیق، باهوش مانند نوع بشر) غیر از نژاد بشر، به دوران باستان بازمی‌گردد که یک موضوع بحث برانگیز موضوع مربوط به ماهیت شناخت و عملیات شناختی در میان فیلسوفان به ویژه در یونان باستان بسیار مورد توجه بود. فیلسوفان مشهوری مانند ارسطو، توماس آکویناس، رنه دکارت، توماس هابز با طرح سؤالاتی در مورد تعاریف استدلال و هوش تأثیرات قابل توجهی در تکمیل مفهوم هوش مصنوعی در طول زمان گذاشتند. اگرچه این سؤالات منجر به تولد هوش مصنوعی به عنوان یک نظریه می‌شد، اما آنها باید منتظر می‌ماندند تا بستر فنی مورد نیاز توسط پیشرفت علم فراهم شود، که تا دهه 1940 با ایجاد اولین رایانه‌ها محقق نشده بود. در دهه 1950 آلن تورینگ با بیان نظریه ای به نام «بازی تقلید» که از طریق آن امکان تفکر ماشین‌ها در آینده توضیح داده شد، زمینه ساز تولد هوش مصنوعی به عنوان یک پدیده علمی شد. [67]

3-2 طبقه بندی هوش مصنوعی:

در هوش مصنوعی مسایل مختلف با روش‌های مختلف حل می‌شود. بنابراین مهم است که یک رویکرد هوش مصنوعی مناسب برای مسئله انتخاب شود. به این معنی که روش انتخابی باید با توجه به مقیاس دشواری مشکل، بهترین تناسب را داشته باشد. طبقه بندی مناسبی برای این هوش مصنوعی وجود دارد که به صورت زیر است. همچنین در شکل 3-1 این طبقه بندی به خوبی نمایش داده شده است.

- **هوش مصنوعی (AI²⁴):** هوش مصنوعی هر تلاشی برای تقلید از رفتار یک انسان است. هوش مصنوعی را می‌توان به عنوان کل دنیای فناوری محاسباتی در نظر گرفت که هر چیزی را که از راه دور شبیه هوش انسانی است، نشان می‌دهد.
- **یادگیری ماشین (ML²⁵):** الگوریتم یادگیری ماشینی الگوریتمی است که قادر به یادگیری از طریق داده‌ها است.
- **یادگیری عمیق (DL²⁶):** یادگیری عمیق نیز نوعی تقلید از مغز انسان خواهد بود. در یادگیری عمیق، نیازی نیست که همه چیز را به صراحت برنامه ریزی کنیم.
- **شبکه عصبی مصنوعی (ANN²⁷):** شبکه عصبی واحد عملیاتی در یادگیری عمیق محسوب می‌شود که برای حل مسائل پیچیده داده محور، رفتاری شبیه به رفتار مغز انسان دارد. از شبکه عصبی در حوزه‌های مختلف مبتنی بر هوش مصنوعی، از تشخیص گفتار و تشخیص چهره گرفته تا مسائل مربوط به مراقبت‌های بهداشتی و بازاریابی استفاده می‌شود. در شبکه عصبی داده‌های ورودی با عبور از لایه‌های مختلف و گره‌های مصنوعی متصل به هم، پردازش می‌شوند تا در نهایت خروجی مورد انتظار تولید شود.



شکل 3-1: توسعه و گسترش هوش مصنوعی

²⁴ Artificial Intelligence

²⁵ Machine Learning

²⁶ Deep Learning

²⁷ Artificial Neural Network

3-3 شبکه‌های عصبی مصنوعی:

تا کنون کاربردهای مختلفی برای شبکه‌های عصبی از جمله در بینایی کامپیوتر، تجزیه و تحلیل تصاویر پزشکی، و پردازش زبان طبیعی پیشنهاد شده است. برای درک مفهوم شبکه‌های عصبی ابتدا لازم است توضیحی از ساختار معماری شبکه عصبی و اجزای آن ارائه شود و سپس به عملکرد تابع فعال سازی در شبکه‌های عصبی و انواع آن‌ها پرداخته شود. اجزای اصلی شبکه عصبی عبارت‌اند از گره (نورون)^{۲۸} و لایه^{۲۹}.

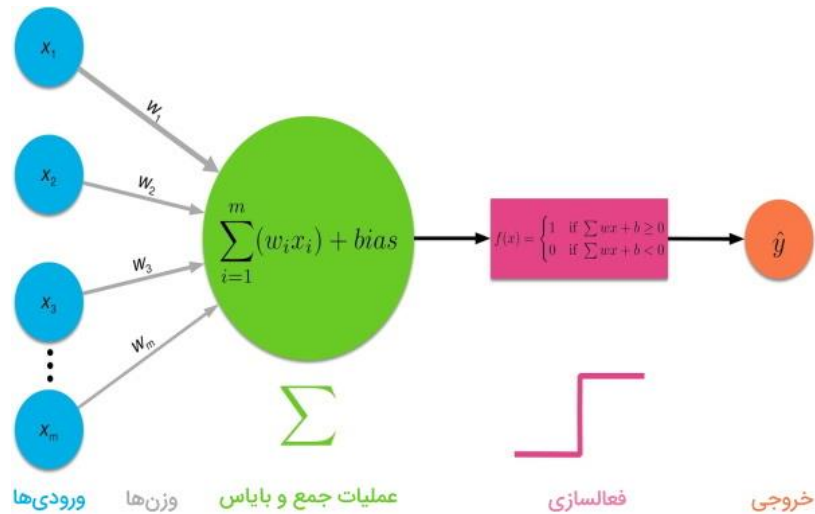
گره در شبکه عصبی مصنوعی: شبکه‌های عصبی از واحدهای مستقلی به نام گره یا نود ساخته شده‌اند که این گره‌ها رفتاری شبیه به رفتار مغز انسان دارند. به عبارتی، گره‌های شبکه‌های عصبی، عملکردی مشابه با عملکرد نورون‌های مغز انسان دارند که مجموعه‌ای از سیگنال‌های ورودی (محرک‌های خارجی) را دریافت می‌کنند. گره، از اجزای مختلفی ساخته شده است؛ این اجزا در ادامه فهرست شده‌اند و در شکل 3-2 روند آن‌ها مشخص شده است:

- **ورودی گره در شبکه عصبی:** ورودی گره شامل مجموعه‌ای از ویژگی‌ها است که برای روند یادگیری شبکه عصبی، به مدل وارد می‌شوند.
- **وزن گره در شبکه عصبی:** از وزن‌ها به منظور تعیین میزان اهمیت ویژگی‌های ورودی گره استفاده می‌شود. به عبارتی، با اعمال ضرب داخلی یا اسکالر مقدار ورودی گره و ماتریس وزن، میزان اهمیت هر ویژگی ورودی مشخص می‌شود.
- **تابع فعال‌سازی^{۳۰} در شبکه عصبی:** در یک شبکه عصبی، یک تابع فعال‌سازی ورودی را نرمالایز می‌کند و یک خروجی تولید می‌کند که سپس به لایه بعدی منتقل می‌شود. توابع فعال‌سازی، غیر خطی بودن را به خروجی اضافه می‌کنند که شبکه‌های عصبی را قادر می‌سازد مسائل غیر خطی را حل کنند (تعداد کلاس بیشتر از ۲). به عبارت دیگر، یک شبکه عصبی بدون تابع فعال‌سازی اساساً فقط یک مدل رگرسیون خطی (Linear Regression) است. توابع فعال‌سازی رایج عبارتند از Sigmoid، Linear، Tanh، Softmax و ReLU اما بسیاری دیگر وجود دارند. شکل 3-3 نمودار تعدادی از توابع فعال‌ساز را نشان می‌دهد.
- **بایاس گره در شبکه عصبی:** نقش بایاس در شبکه عصبی این است که خروجی نهایی تابع فعال‌سازی را تغییر دهد. نقش بایاس مشابه نقش مقدار ثابت در تابع خطی است.

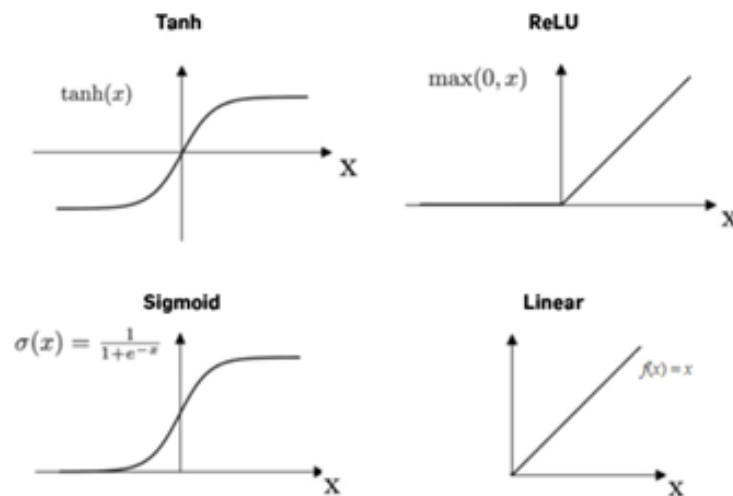
²⁸ Node

²⁹ Layer

³⁰ Activation Function



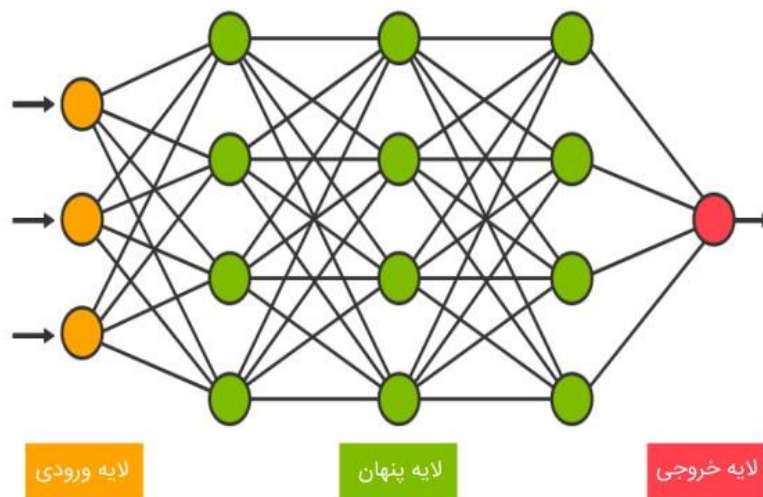
شکل 3-2: روند یک شبکه عصبی



شکل 3-3: تعدادی از توابع فعال‌ساز

لایه در شبکه های عصبی مصنوعی: زمانی که در شبکه عصبی، چندین گره در یک ردیف کنار هم قرار بگیرند، یک لایه را تشکیل می‌دهند. زمانی که چندین لایه از گره‌ها در کنار یکدیگر قرار بگیرند، شبکه عصبی ساخته می‌شود. انواع لایه های شبکه های عصبی در ادامه فهرست شده اند. همچنین شکل 3-4 انواع لایه های شبکه های عصبی را نشان می دهد.

- **لایه ورودی:** در شبکه های عصبی، همیشه یک لایه ورودی برای دریافت ویژگی های ورودی وجود دارد. به عبارتی، اطلاعات دنیای خارج، از طریق این لایه وارد شبکه عصبی می شوند. در این لایه، هیچ عملیات محاسباتی انجام نمی شود و صرفاً اطلاعات (ویژگی ها) از طریق این لایه، به لایه های پنهان منتقل می شوند. تعداد گره های لایه ورودی به ابعاد داده های ورودی بستگی دارند.
- **لایه پنهان:** این لایه، بین لایه ورودی و لایه خروجی قرار دارد و به همین دلیل به این لایه، لایه میانی یا پنهان گفته می شود. به طور معمول، از تابع فعال سازی مشابه برای تمامی لایه های پنهان شبکه عصبی استفاده می شود. با این حال، براساس هدف و کاربرد شبکه عصبی، تابع فعال سازی لایه خروجی شبکه معمولاً با سایر توابع فعال سازی لایه های پنهان مغایرت دارد.
- **لایه خروجی:** این لایه، خروجی نهایی شبکه عصبی را تولید می کند. در شبکه های عصبی، وجود یک لایه خروجی الزامی است. لایه خروجی، ورودی خود را از لایه قبل خود دریافت می کند و با استفاده از گره های خود، محاسباتی روی ورودی ها انجام می دهد و مقدار خروجی شبکه عصبی را محاسبه می کند. تعداد گره ها در لایه خروجی برابر با تعداد کلاس ها است.



شکل ۳-۴: انواع لایه های شبکه های عصبی

3-3-1 شبکه عصبی پیچشی (CNN³¹):

شبکه‌های عصبی پیچشی شاخه‌ای از شبکه‌های عصبی عمیق هستند که معمولاً برای طبقه‌بندی تصاویر در یادگیری ماشین استفاده می‌شوند. شبکه‌های عصبی پیچشی نسبت به بقیه رویکردهای دسته‌بندی تصاویر به میزان کمتری از پیش پردازش استفاده می‌کنند. این امر به معنی آن است که شبکه یاد می‌گیرد که گره‌ها را از طریق یادگیری خودکار بهینه کند و همچنین شبکه معیارهایی را فرا می‌گیرد که در رویکردهای قبلی به صورت دستی فراگرفته می‌شدند. این استقلال از دانش پیشین و دستکاری‌های انسانی در شبکه‌های عصبی کانولوشن یک مزیت اساسی است.

در میان شبکه‌های عصبی، شبکه‌ی عصبی کانولوشنی یکی از بهترین‌ها برای حل مسائل حوزه‌ی بینایی ماشین^{۳۲}، مانند شناسایی تصاویر^{۳۳}، طبقه‌بندی تصاویر^{۳۴}، تشخیص چهره^{۳۵} و ... است.

یک شبکه عصبی پیچشی از دو بخش کلی تشکیل شده است:

- استخراج ویژگی^{۳۶}
- طبقه‌بندی

استخراج ویژگی‌های یک عکس توسط لایه‌های زیر انجام می‌شود.

لایه‌ی کانولوشن^{۳۷}: لایه‌ی کانولوشن اولین گام در استخراج ویژگی عکس ورودی است. در لایه‌ی کانولوشن، شبکه با استفاده از فیلترها^{۳۸} که ماتریس‌هایی کوچک (به طور مثال 3×3) هستند، ویژگی‌های عکس را یاد می‌گیرد. در این حال فیلتر مثل یک پنجره‌ی کوچک روی عکس ورودی قرار می‌گیرد. این فیلتر روی هر بخش از عکس ورودی با یک گام^{۳۹} مشخص که قرار می‌گیرد مقدار هر پیکسل آن با مقدار پیکسل متناظرش در ماتریس عکس ورودی ضرب می‌شود؛ حاصل جمع ضرب تمامی پیکسل‌ها با پیکسل‌های متناظرشان ماتریس جدیدی را ایجاد می‌کند که به آن ویژگی کانولوشن گفته می‌شود.

³¹ Convolutional Neural Network

³² Computer Vision

³³ Image Detection

³⁴ Image Classification

³⁵ Face Recognition

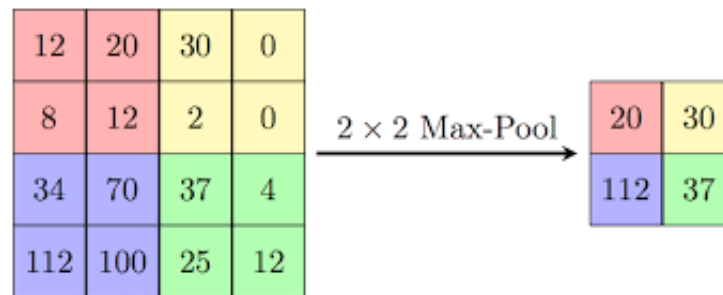
³⁶ Feature Extraction

³⁷ Convolutional Layer

³⁸ kernel

³⁹ Stride

لایه‌ی ادغام^{۴۰}: گام بعدی در مرحله استخراج ویژگی لایه‌ی ادغام است. در این مرحله هدف این است که ابعاد عکس ورودی را کاهش دهیم و فقط اطلاعات یا پیکسل‌های مهم را نگه داریم و باقی را حذف کنیم. از انواع آن می‌توان به روش حداکثر ادغام^{۴۱} و روش میانگین ادغام^{۴۲} و روش جمع ادغام^{۴۳} اشاره کرد. در روش حداکثر ادغام پیکسلی که مقدار بزرگتری در مقایسه با باقی دارد، در روش میانگین ادغام مقدار میانگین و در روش جمع ادغام جمع تمامی مقادیر در هر ویژگی کانولوشن انتخاب می‌شود. شکل 3-5 عملیات حداکثر ادغام را نشان می‌دهد.



شکل 3-5: نمونه‌ای از Max Pooling

در شبکه‌های عصبی کانولوشن از یک سری وزن‌های اولیه استفاده می‌شود. شبکه، وزن‌ها را در لایه‌های کانولوشن به اشتراک می‌گذارد که باعث می‌شود حداقل حافظه و بیشترین کارایی به دست آید.

به طور کلی زمانی که یک عکس به یک شبکه‌ی CNN وارد می‌شود، ابتدا به مرحله‌ی استخراج ویژگی وارد می‌شود. در این مرحله هر عکس ورودی از چندین سری لایه‌ی کانولوشن و تابع فعال‌ساز ReLU و لایه‌ی ادغام عبور می‌کند. سپس عکس‌های ورودی به طبقه‌بندی وارد می‌شوند. در این مرحله ابتدا مسطح‌سازی^{۴۴} صورت می‌گیرد و سپس به لایه‌های کاملاً متصل^{۴۵} وارد می‌شوند. رویکرد لایه کاملاً متصل در اصل مشابه کاری است که در شبکه عصبی MLP^{۴۶} انجام می‌شود. در نهایت یک فعال‌ساز Softmax برای مسائل طبقه‌بندی چندکلاسه و یا فعال‌ساز Sigmoid برای مسائل طبقه‌بندی باینری روی آن اعمال می‌شود تا عکس‌ها براساس مقادیر احتمالی طبقه‌بندی شوند. در شکل 3-6 نمونه‌ای از این فرایند را می‌توان مشاهده کرد.

⁴⁰ Pooling Layer

⁴¹ Max Pooling

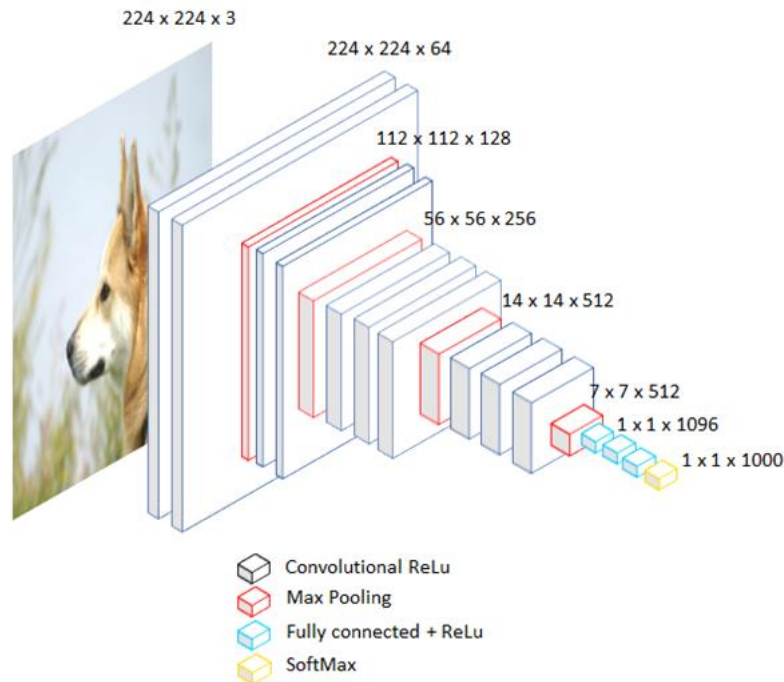
⁴² Average Pooling

⁴³ Sum Pooling

⁴⁴ Flatting

⁴⁵ Fully connected Layer

⁴⁶ Multi-Layer Perceptron



شکل ۳-۶: فرایند یک شبکه CNN

3-2-3 شبکه عصبی پیچشی از پیش آموزش دیده شده^{۴۷}:

شبکه های عصبی کانولوشنال از پیش آموزش دیده مدلی هستند که بر روی مجموعه داده های بزرگ آموزش دیده اند و می توانند برای کاربردهای مختلف مانند طبقه بندی تصویر، تشخیص اشیا یا تشخیص چهره مجدداً استفاده شوند. با استفاده از یک مدل از پیش آموزش دیده، می توانید با داده های کمتر برچسب گذاری شده به عملکرد خوبی دست پیدا کنید. زیرا این مدل ها قبلاً از یک مجموعه داده بزرگ یاد گرفته اند.

برخی از این دیتاست های بزرگ عبارت اند از:

- ImageNet: پایگاه داده ImageNet شامل حدود 16 میلیون تصویر در 1000 کلاس مختلف است و مسابقات آن هر ساله برگزار می شود و تبدیل به جام جهانی بینایی کامپیوتر شده است.
- CIFAR
- MNIST

⁴⁷ Pre-trained CNN model

از معروف ترین شبکه های عصبی کانولوشنال از پیش آموزش دیده می توان به شبکه های زیر اشاره کرد:

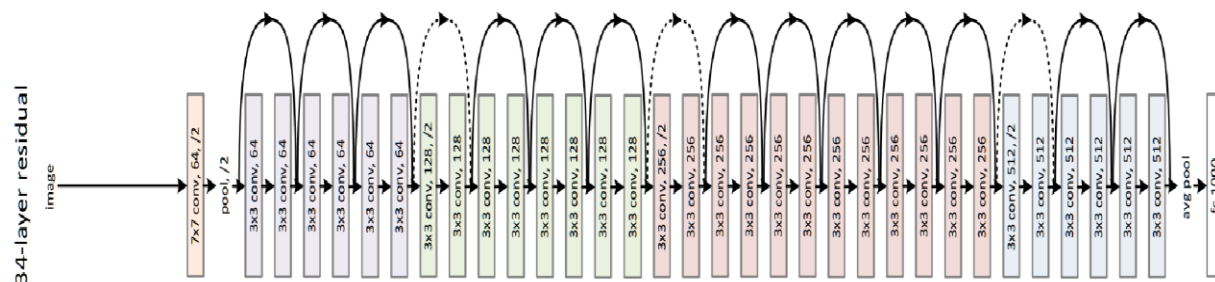
- ResNet
- InceptionV3
- Xception
- VGG16
- VGG19
- MobileNet
- EfficientNet

در ادامه توضیحاتی درباره مدل های از پیش آموزش دیده ای که در این پروژه استفاده شده اند، ارائه شده است.

مدل ResNet50^{۴۸}:

شبکه عصبی عمیق ResNet50 را محققانی به نام شیوکی‌نگ رن^{۴۹}، کیمینگ هی^{۵۰}، ژان سان^{۵۱} و زیانگیا ژوئنگ^{۵۲} در سال ۲۰۱۵ معرفی کردند. مدل ResNet تاکنون یکی از محبوب ترین و موفق ترین مدل های یادگیری عمیق بوده است. این مدل برنده ی چالش ILSVRC در سال ۲۰۱۵ بود. دلیل موفقیت شبکه ی ResNet این است که به ما امکان آموزش شبکه های عصبی بسیار عمیق با بیش از ۱۵۰ لایه را می دهد.

شکل ۳-۷ معماری لایه های این شبکه ها را نشان می دهد.



شکل ۳-۷: معماری لایه های شبکه ResNet

⁴⁸ Residual Network

⁴⁹ Shaoqing Ren

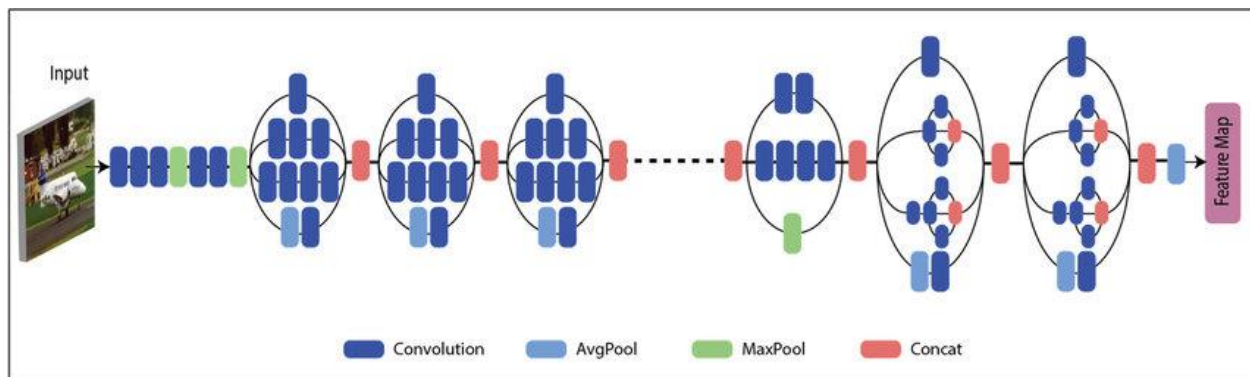
⁵⁰ Kaiming He

⁵¹ Jian Sun

⁵² Xiangyu Zhang

مدل InceptionV3 :

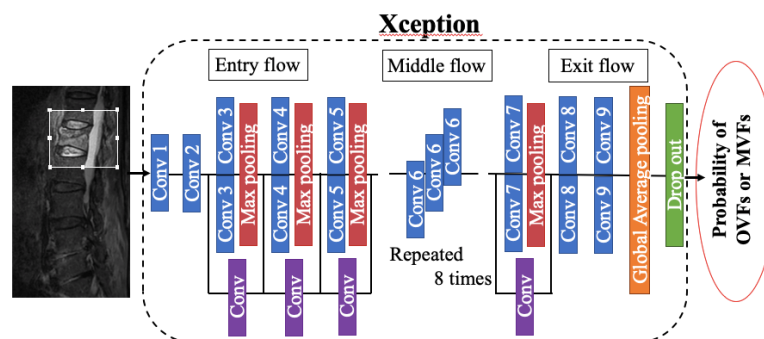
شبکه عصبی عمیق InceptionV3 برای کمک به تجزیه و تحلیل تصویر و تشخیص اشیا است و به عنوان یک ماژول برای GoogLeNet شروع به کار کرد. این شبکه سومین نسخه از شبکه عصبی کانولوشنال آغازین گوگل است که در ابتدا در مسابقه معروف ImageNet معرفی شد. این شبکه در مجموعه داده ImageNet دقتی بیش از 78 درصد دارد و نقطه اوج بسیاری از ایده های توسعه یافته توسط محققان متعدد در طول سال ها است. شکل 3-8 معماری لایه های این شبکه ها را نشان می دهد.



شکل 3-8: معماری لایه های شبکه InceptionV3

مدل Xception :

شبکه عصبی عمیق Xception یک شبکه عصبی کانولوشن است که 71 لایه عمق دارد. شکل 3-9 معماری لایه های این شبکه ها را نشان می دهد.

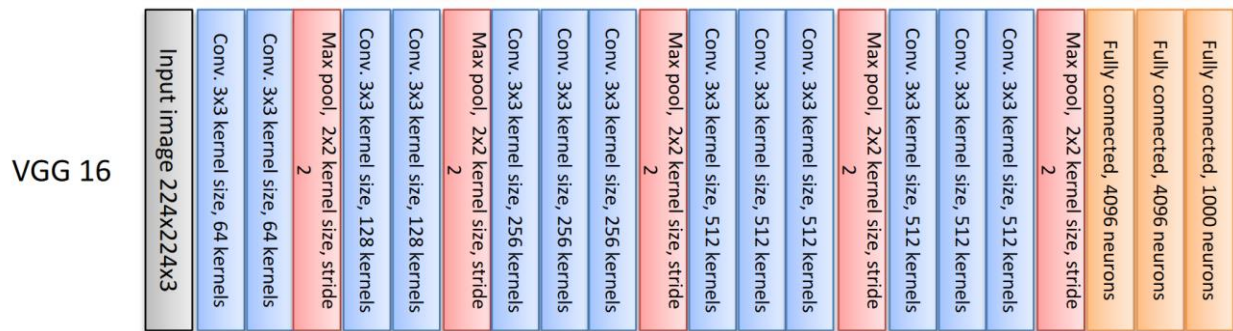


شکل 3-9: معماری لایه های شبکه Xception

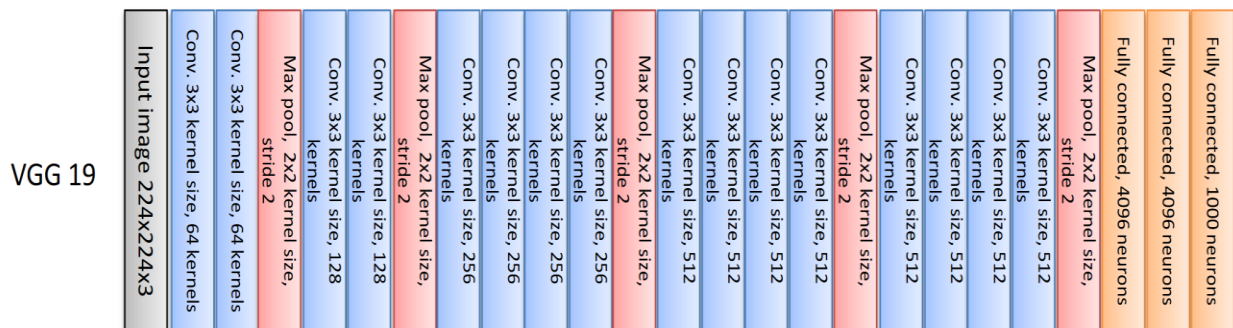
مدل VGG^{۵۳}:

شبکه عصبی عمیق VGG یک شبکه عصبی کانولوشنی است که توسط محققانی به نام سیمونز^{۵۴} و زیسرمن^{۵۵} از دانشگاه آکسفورد پیشنهاد شد. این شبکه جزو اولین شبکه‌های عصبی کانولوشنی است که توانست دقت بالایی در مسابقه معروف ImageNet بدست آورد. باتوجه به پیشرفت‌های اخیر شبکه‌های عصبی عمیق، شاید دیگر نتوان شبکه VGG را یک شبکه مدرن دانست، اما به دلیل ساختار خوب، سادگی، تعداد لایه‌های نه‌چندان زیاد هنوز در بسیاری از زمینه‌های بینایی کامپیوتر از این شبکه استفاده می‌شود.

شبکه عصبی عمیق VGG دو نوع معماری 16 لایه ای و 19 لایه ای دارد که به ترتیب به نام های VGG16 و VGG19 شناخته می شوند. شکل های 3-10 معماری لایه های این شبکه ها را نشان می‌دهند.



الف: معماری لایه های شبکه VGG16



ب: معماری لایه های شبکه VGG19

شکل 3-10: معماری لایه های شبکه های VGG16 و VGG19

⁵³ Visual Geometry Group⁵⁴ K. Simonyan⁵⁵ A. Zisserman

3-4 معیارهای سنجش شبکه های عصبی:

برای بررسی سطح و صحت عملکرد یک شبکه ی عصبی معیارهای متفاوتی تعریف شده است. هرکدام از این معیارها به صورت کمی یا نمایشی عملکرد مدل را ارزیابی می کنند. تعدادی از مهم ترین و متداول ترین این معیارها را در ادامه توضیح خواهیم داد:

ماتریس در هم ریختگی^{۵۶}:

برای اینکه ایده ی مشخصی از تمام معیارها درک کنیم، ابتدا به ماتریس در هم ریختگی می پردازیم. این ماتریس نمایشی کمی و بصری از عملکرد مدل روی داده های آزمایشی یا تست است. به این صورت که برای ساخت و روند یادگیری مدل از درصد زیادی از داده ها (عموما بیشتر از ۸۰ درصد) استفاده می شود. به این داده ها و روند یادگیری مدل آموزش^{۵۷} گفته می شود. بخش باقی مانده از داده ها برای ارزیابی استفاده می شود که به آن بخش تست گفته می شود.

به طور ساده داده های تست برای تشخیص درستی کلاس ها به مدل داده می شود. شکل 3-11 تعریف و بررسی معیارهای TP^{58} , FP^{59} , FN^{60} , TN^{61} را نمایش می دهد که برای ارزیابی عملکرد مدل استفاده می شود.

True label	Class 1	<div>TP for class 1</div>	<div>FP for class 2, FN for class 1</div>
	Class 2	<div>FP for class1, FN for class 2</div>	<div>TP for class 2</div>
		Class 1	Class 2
		Predicted	

شکل 3-11: نمایی تعریفی از FP, FN, TP, TN

⁵⁶ Confusion Matrix

⁵⁷ Train

⁵⁸ True Positive

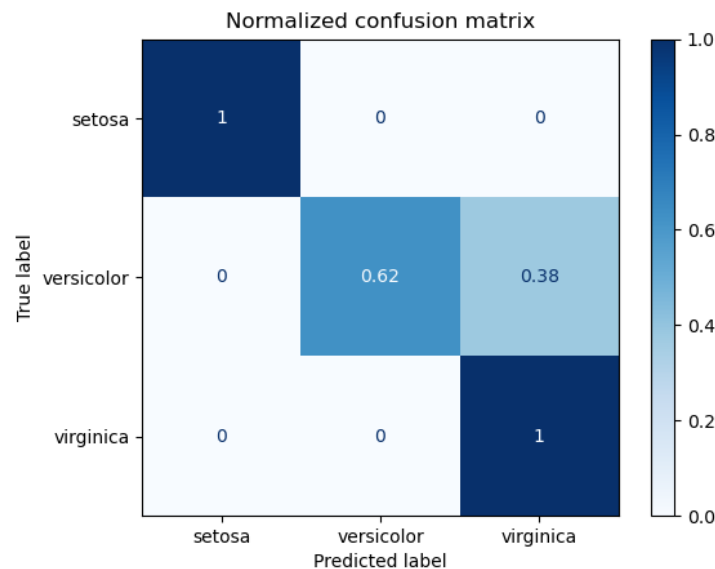
⁵⁹ False Positive

⁶⁰ False Negative

⁶¹ True Negative

در شکل 3-11 شاهد یک مدل باینری یا ۲ کلاسه هستیم. در مدل‌های ۲ کلاسه ما به طور ساده و زبان کامپیوتر با صفر و یک درگیریم. TP و TN یعنی مدل ما داده‌های هر کلاس را به درستی تشخیص داده است. FP و FN یعنی جابه‌جا تشخیص داده است.

به کمک این پارامترها می‌توانیم برای هر تعداد کلاس دلخواه، یک ماتریس با نام ماتریس در هم ریختگی بسازیم. در شکل 3-12 نمایشی از یک ماتریس در هم ریختگی دارای 3 کلاس به صورت نرمالایز شده نمایش داده شده است.



شکل ۳-۱۲: نمایشی از یک Confusion Matrix دارای 3 کلاس

دقت^{۶۲}:

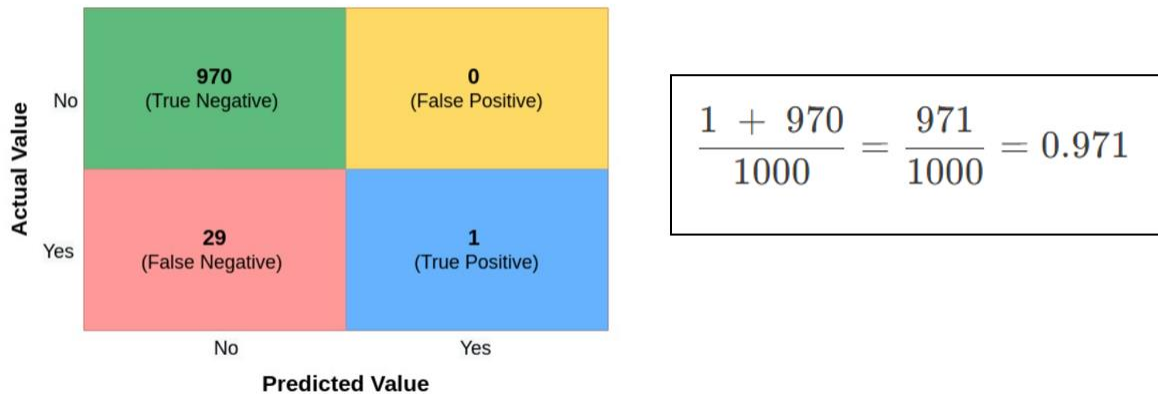
این معیار به صورت زیر به دست می‌آید:

$$Accuracy = \frac{TP + TN}{Total Predictions}$$

همانطور که مشخص است داده‌های قطری را بر همه‌ی داده‌ها تقسیم کرده‌ایم. مشکل بزرگ این معیار در دیتاست‌های غیر متعادل^{۶۳} است که در تعداد دیتاها تفاوت زیادی وجود دارد برای مثال ماتریس شکل 3-13 و دقت آن را در نظر بگیرید.

⁶² Accuracy

⁶³ Data imbalance



شکل ۳-۱۳: ماتریس یک دیتاست Imbalanced و Accuracy آن

همانطور که از شکل 3-13 مشخص است دقت ما ۹۷ درصد است؛ اما عملکرد واقعی مدل بسیار غیردقیق است، چون به ازای همه ورودی‌ها، تشخیص مدل، کلاس ۱ یا همان کلاسی است که داده‌های آن بیشتر است.

یک مشکل بزرگ دیگر دقت هنگام آموزش این است که اگر دیتاست بسیار متفاوتی داشته باشیم که تشخیص تفاوت‌ها بین آن‌ها برای مدل عملاً غیر ممکن باشد، به طور مثال داده‌ها به هم مربوط نباشند، ممکن است به اشتباه دقت بالا به دست آید؛ این معیار در اینجا در واقع نشان از این دارد که مدل ما ویژگی‌هایی مطلوب را یاد نگرفته و به جای آن همه‌ی صفات منفی و بی‌فایده‌ی دیتاست را فراگرفته است. این مشکل را در دیتاست‌هایی که دقت بسیار بالا و F1 score بسیار پایین دارند، مشاهده می‌کنیم. F1 score یکی از معیارهای سنجش شبکه‌های عصبی است که در ادامه معرفی می‌گردد.

تابع هزینه^{۶۴}:

تابع هزینه یا زیان^{۶۵} احتمالات یا عدم قطعیت یک پیش‌بینی را بر اساس این که چقدر پیش‌بینی از مقدار واقعی متفاوت است، در نظر می‌گیرد. این معیار به ما دید دقیق‌تری از عملکرد خوب مدل می‌دهد.

برخلاف دقت، تابع هزینه یک درصد نیست؛ بلکه خلاصه‌ای از خطاهای ایجاد شده برای هر نمونه در مجموعه‌های آموزشی یا اعتبار سنجی است. تابع هزینه اغلب در فرآیند آموزش برای یافتن بهترین مقادیر پارامتر برای مدل استفاده می‌شود (به عنوان مثال وزن‌ها در شبکه عصبی). در طول فرآیند آموزش، هدف به حداقل رساندن تابع هزینه است.

⁶⁴ Loss function

⁶⁵ Cost function

مهم ترین توابع هزینه عبارت اند از: Binary Cross-entropy ، Categorical Cross-entropy ، Hinge ، Poisson و ...

بهینه ساز^{۶۶}:

بسیار مهم است که وزن های مدل را در طول فرآیند آموزش تغییر دهیم تا پیش بینی های خود را تا حد امکان صحیح و بهینه کنیم. اما دقیقاً چگونه پارامترهای مدل را به چه میزان و چه زمانی تغییر دهیم؟

بهترین پاسخ یک بهینه ساز است. بهینه ساز، تابع هزینه و پارامترهای مدل را با به روزرسانی مدل در پاسخ به خروجی تابع ضرر به هم گره می زنند. به عبارت ساده تر، بهینه سازها با استفاده از وزن ها، مدل را به دقیق ترین شکل ممکن می سازند. تابع هزینه به عنوان راهنما عمل می کند و برای بهینه ساز مشخص می کند چه زمانی در جهت درست یا غلط حرکت می کند.

بیش برآزشی^{۶۷} و زیربرآزشی^{۶۸}:

این اصطلاحات دو افراط متضاد را توصیف می کنند که هر دو منجر به عملکرد ضعیف مدل می شوند. بیش برآزش به مدلی اشاره دارد که بیش از حد بر روی جزئیات داده های آموزش دقت کرده است و نویز موجود در مجموعه داده را یاد می گیرد. مدلی که بیش برآزش است، روی داده های جدید و دیده نشده عملکرد خوبی نخواهد داشت. بیش برآزشی مسلماً رایج ترین مشکل در یادگیری ماشینی کاربردی است و به ویژه مشکل ساز است، زیرا مدلی که به نظر می رسد بسیار دقیق است، عملکرد ضعیفی در واقعیت و آزمودن دارد. در واقع مدل ما داده ها را یاد نگرفته و حفظ کرده است.

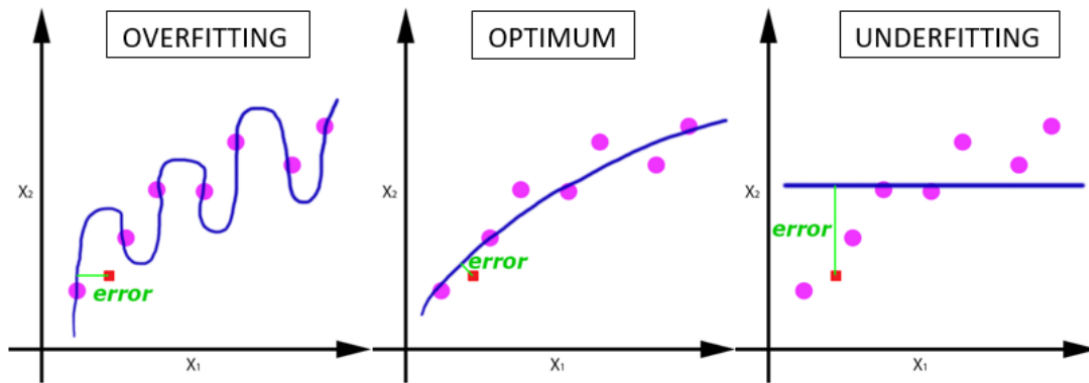
زیربرآزش معمولاً به مدلی اشاره دارد که به اندازه کافی آموزش ندیده است. این می تواند به دلیل زمان ناکافی آموزش یا مدلی باشد که به سادگی آموزش داده نشده است. مدل زیربرآزش بر روی داده های آموزشی و همچنین داده های جدید و دیده نشده به طور یکسان عملکرد ضعیفی خواهد داشت.

هردوی این اتفاقات، عملکرد ضعیفی را به همراه خواهد داشت و نقطه ایده آل ما در بین این دو حالت افراطی است. با افزایش تعداد تکرارهای آموزشی، پارامترهای مدل به روز می شوند و منحنی از حالت زیربرآزش به بهینه و سپس به بیش برآزش می رود. در شکل 3-14 نمایی از بیش برآزشی و زیربرآزشی نمایش داده شده است.

⁶⁶ Optimizer

⁶⁷ Overfitting

⁶⁸ Underfitting

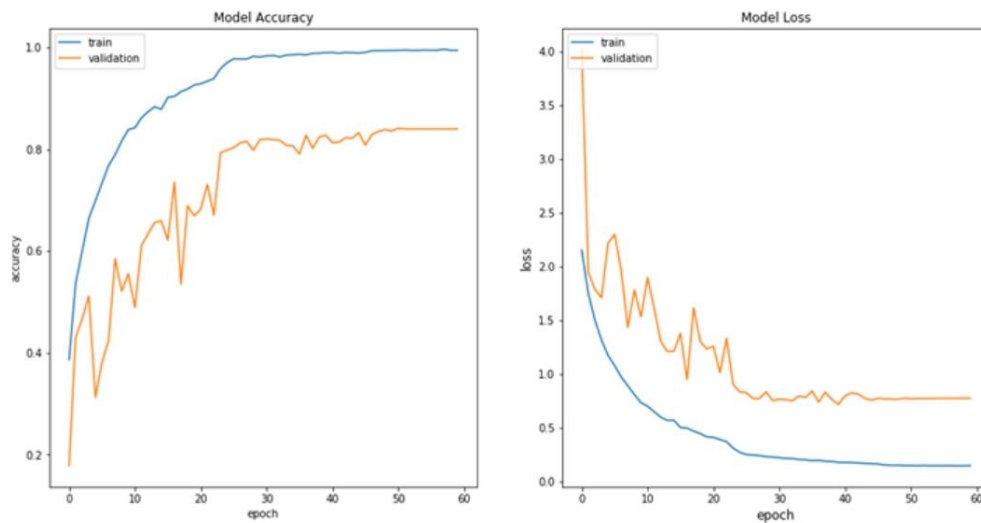


Source: Sagar Sharma / Towards Data Science

شکل ۳-۱۴: نمایی تعریفی از Overfitting و Underfitting

رابطه‌ی دقت و تابع هزینه با بیش‌برازشی و زیربرازشی:

در حالت کلی باید دقت در مرحله‌ی ارزیابی پایین‌تر از حالت آموزش باشد چون ما داده‌های تست کمتری داریم. به همین ترتیب، تابع هزینه مرحله ارزیابی باید بیشتر از آموزش باشد. در غیر این صورت‌ها در واقع بیش‌برازشی داریم و داده‌های تست توسط مدل حفظ شده‌اند. اگر تابع هزینه خیلی بالا یا دقت خیلی پایین هم داشته باشیم، احتمال زیربرازشی یکی از چندین احتمالی است که باید بررسی کنیم. همچنین در نمودارهای بیش‌برازش فراز و نشیب‌های زیادی مشاهده می‌شود. شکل ۳-۱۵ نمایی از دقت و تابع هزینه در طول آموزش و ارزیابی را نمایش می‌دهد.



Source: Microsoft

شکل ۳-۱۵: نمایی از Loss و Accuracy در طول Train و Validation

صحت^{۶۹}:

با توجه به شکل 3-16 این معیار برای هر کلاس به صورت زیر محاسبه می‌شود:

		Predicted	
		Negative	Positive
Actual	Negative	True Negative	False Positive
	Positive	False Negative	True Positive

شکل 3-۱۶: نحوه محاسبه‌ی Precision

$$Precision = \frac{TP}{TP + FP} = \frac{TP}{Total\ Predictions}$$

زمانی که هدف مینیمم کردن موارد مثبت کاذب FP است، این معیار اهمیت بالایی دارد. برای مثال تشخیص ایمیل‌های مخرب^{۷۰}.

پوشش^{۷۱}:

با توجه به شکل 3-17 این معیار برای هر کلاس به صورت زیر محاسبه می‌شود:

		Predicted	
		Negative	Positive
Actual	Negative	True Negative	False Positive
	Positive	False Negative	True Positive

شکل 3-۱۷: نحوه محاسبه‌ی Recall

$$Recall = \frac{TP}{TP + FN} = \frac{TP}{Total\ Actual\ Positive}$$

زمانی که هدف مینیمم کردن موارد منفی کاذب FN است، از این معیار استفاده می‌کنیم. درواقع پوشش ساده‌ترین و بنیادی‌ترین شکل بررسی خطا است و در بررسی کلاس‌های مختلف می‌توان به راحتی معنی این معیار را فهمید.

⁶⁹ Precision

⁷⁰ Spam

⁷¹ Recall

F1 score

این معیار درواقع میانگین هارمونیک صحت و پوشش است و نگاه کلی از عملکرد مدل به ما می‌دهد.

$$F1\ score = \frac{2 * Precision * Recall}{Precision + Recall}$$

نحوه‌ی بیان رایج عملکرد یک مدل شبکه‌ی عصبی:

عموماً برای نمایش درستی کارکرد یک مدل دقت، F1 Score و برای نمایش بصری ماتریس در هم ریختگی آن را به نمایش می‌گذارند. هر سه‌ی این معیارها عملکرد سیستم را در هر کلاس و به صورت میانگین^{۷۲} به نمایش می‌گذارند. به طور خلاصه، دقت نشان می‌دهد مدل آموزش دیده است یا خیر. همچنین F1 Score نشان می‌دهد مدل آموزش داده شده‌ی ما کلاس داده‌های جدید را تشخیص می‌دهد یا خیر.

⁷² Weighted average

4. فصل چهارم: نتایج و بحث

در این فصل ابتدا به معرفی دیتاستی که در فرایند مدل سازی استفاده می شود پرداخته می شود. سپس درباره جزئیات مدل های ساخته شده توضیحات مختصری داده می شود و نتایج آن ها ارائه و درباره آن ها بحث می شود. نتایج ارائه شده در این فصل شامل نمودار دقت بر حسب دوره^{۷۳} برای داده های آموزش و تست، نمودار تابع هزینه بر حسب دوره برای داده های آموزش و تست، گزارش طبقه بندی^{۷۴} و ماتریس در هم ریختگی است.

4-1 دیتاست:

4-1-1 معرفی دیتاست:

دیتاستی که بر روی آن مدل سازی ها انجام می شود، شامل عکس های OCT⁷⁵ است که برای تشخیص بیماری های چشم از آن ها استفاده می شود. این دیتاست از بیمارستان فوق تخصصی چشم نگاه جمع آوری شده است. این دیتاها طبقه بندی نشده بودند. بنابراین لازم بود تا کلاس هر عکس مشخص شود. فرایند طبقه بندی و لیبل زدن دیتاها نیز در همین بیمارستان و به کمک دکتر رحیمی انجام شده است. لازم به ذکر است که توضیحات مربوط به نحوه تصویربرداری پزشکی و دستگاه های آن در فصل دوم آورده شده است.

4-1-2 پیش پردازش داده ها^{۷۶}:

قبل از مدل سازی لازم است بر روی داده ها تغییراتی را اعمال کنیم تا این عکس ها را به صورت مطلوب آماده ورود به شبکه کنیم. این پیش پردازش داده ها ممکن است به روش های مختلف انجام شود. برای مثال crop کردن، زوم کردن، چرخاندن تصاویر و... حتی ممکن است در این فرایند لازم باشد تا تصاویری که در یک کلاس با بقیه داده های آن کلاس هم خوانی ندارند را از دیتاست حذف کنیم.

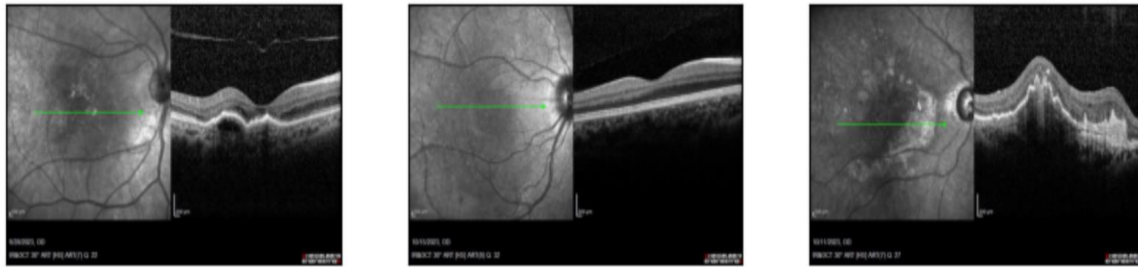
نمونه هایی از عکس های اولیه قبل از پیش پردازش در شکل 4-1 نمایش داده شده است.

⁷³ epochs

⁷⁴ Classification report

⁷⁵ Optical Coherence Tomography

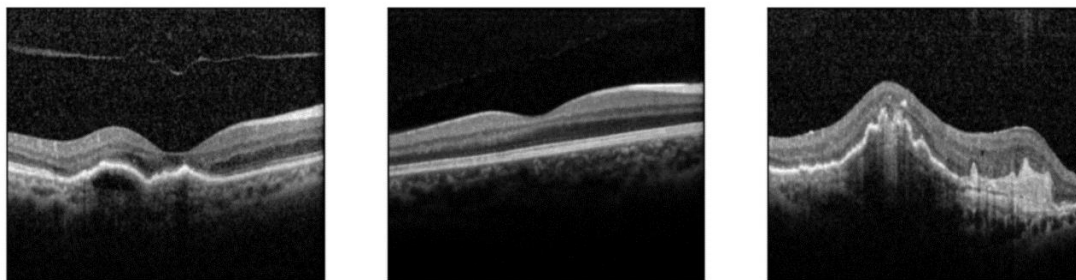
⁷⁶ Preprocessing Date



شکل ۴-۱: نمونه هایی از عکس های اولیه

سایز همه این عکس (596 , 1008) است. قسمت سمت راست این تصاویر، عکس OCT چشم است و برای مدل سازی فقط قسمت سمت راست تصاویر مطلوب است. همچنین به حاشیه های تصاویر نیاز نداریم، بنابراین در فرآیند پیش پردازش داده ها لازم است عکس ها را crop کنیم.

سایز همه عکس ها بعد از فرآیند پیش پردازش به (435 , 508) تغییر می کند. لازم به ذکر است که در هر مدل باید سایز عکس ها را متناسب با سایز ورودی آن مدل تغییر دهیم. نمونه هایی از عکس های crop شده بعد از پیش پردازش در شکل ۴-۲ نمایش داده شده است.



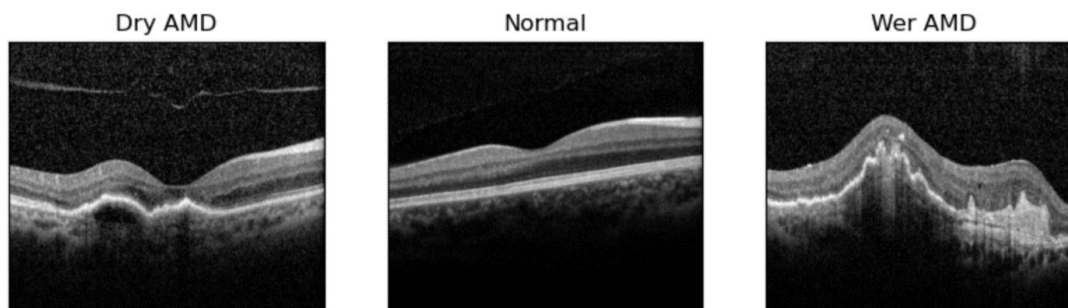
شکل ۴-۲: نمونه هایی از عکس ها بعد از فرآیند پیش پردازش

4-1-3 طبقه بندی دیتاست:

این دیتاست شامل 3 دسته است که وضعیت بیماری چشم را مشخص می کند. این سه دسته عبارت اند از:

- چشم سالم (Normal)
- بیماری Dry AMD
- بیماری Wet AMD

توضیحات مربوط به این بیماری ها در فصل دوم آورده شده است. نمونه ای از تصاویر هر کلاس از دیتاست در شکل ۴-۳ آورده شده است.



شکل ۴-۳: نمونه ای از تصاویر هر کلاس از دیتاست

در مدل سازی نام هر کلاس به صورت یک عدد شناخته می شود. نام گذاری عدد مربوط به هر کلاس مطابق جدول 4-1 می باشد.

جدول ۴-۱: نام گذاری کلاس ها

نام کلاس	عدد مربوط به کلاس در مدل سازی
Dry AMD	0
Normal	1
Wet AMD	2

4-1-4 تعداد داده ها و نحوه تقسیم آن ها در آموزش^{۷۷} و تست^{۷۸}:

این دیتاست شامل 285 عکس OCT است که در 3 کلاس تقسیم شده اند. برای جلوگیری از عدم تعادل دیتاها تعداد 24 عکس از هر کلاس را به داده های تست اختصاص می دهیم و سایر عکس های هر کلاس مربوط به داده های آموزش هستند. پراکندگی داده ها در هر کلاس در جدول 4-2 نمایش داده شده است.

جدول ۴-۲: پراکندگی داده های دیتاست در هر کلاس

نام کلاس	تعداد کل داده ها	تعداد داده های آموزش	تعداد داده های تست
Dry AMD	89	65	24
Normal	80	56	24
Wet AMD	116	92	24
مجموع	285	213	72

⁷⁷ Train

⁷⁸ Test

4-2 جزئیات و نتایج مدل سازی:

در این قسمت جزئیات و نتایج مدل سازی های مختلف ارائه و درباره آن ها بحث می شود. همچنین مدل سازی های انجام شده با هم مقایسه می شوند.

قبل از شروع فرایند مدل سازی، با استفاده از ابزار ImageDataGenerator، داده های آموزش و تست را که از قبل در پوشه های جداگانه قرار دادیم، به مدل معرفی و تکنیک های افزایش داده ها^{۷۹} را روی آن ها اعمال می شود. با توجه به کمبود دیتا، برای مدل سازی از شبکه های عصبی از پیش آموزش دیده استفاده می کنیم. شبکه های عصبی از پیش آموزش دیده مورد استفاده در این پروژه عبارت اند از:

- VGG19
- ResNet
- InceptionV3
- Xception

در هر کدام از موارد بالا، مدل سازی یک بار با فریز کردن همه لایه های شبکه از پیش آموزش دیده و یک بار با از فریز خارج کردن 5 لایه آخر آن ها انجام می شود. لازم به ذکر است که لایه هایی که فریز می شوند، پارامتر های آن ها قابل آموزش دیدن نیستند. در تمامی مدل سازی هایی که در ادامه ارائه شده است، ابتدا با استفاده از شبکه های از پیش آموزش دیده ویژگی های دیتا ها استخراج می شود. سپس در ابتدای لایه های این شبکه ها یک لایه ورودی^{۸۰} اضافه می شود. همچنین به انتهای لایه های این شبکه ها یک لایه مسطح ساز^{۸۱} و دو لایه کاملاً متصل^{۸۲} و یک لایه خروجی^{۸۳} به مشخصات موجود در جدول 3-4 اضافه شده و فرایند مدل سازی آغاز می شود. اضافه کردن لایه های ورودی و خروجی که در هر مدل سازی لازم است، اما لایه های مسطح ساز و لایه های کاملاً متصل به دلیل افزایش پارامتر های قابل آموزش و در نتیجه بهبود عملکرد مدل سازی است.

جدول ۳-۴: مشخصات لایه های اضافه شده به مدل های از پیش آموزش دیده

تعداد گره ها	تابع فعال ساز	نوع لایه	نام لایه
300	Relu	Dense	Fully Connected 1
200	Relu	Dense	Fully Connected 2
3 (برابر با تعداد کلاس ها)	Softmax	Dense	Output

⁷⁹ Data Augmentation

⁸⁰ Input Layer

⁸¹ Flatting

⁸² Fully connected Layer

⁸³ Output Layer

4-2-1 مدل سازی با استفاده از شبکه VGG19:

در این مدل سازی با استفاده از شبکه VGG19 ویژگی های دیتا ها استخراج می شود. سپس در ابتدای لایه های این شبکه یک لایه ورودی و به انتهای لایه های این شبکه یک لایه مسطح ساز و دو لایه کاملاً متصل و یک لایه خروجی به مشخصات موجود در جدول 3-4 اضافه شده و فرایند مدل سازی آغاز می شود. لازم به ذکر است که سائز عکس های ورودی به شبکه VGG19 باید (224, 224) باشد که در لایه ورودی تنظیم می شود.

جزئیات این مدل در شکل 4-4 نمایش داده شده است.

Model: "model"

Layer (type)	Output Shape	Param #
input_2 (InputLayer)	[(None, 224, 224, 3)]	0
vgg19 (Functional)	(None, 7, 7, 512)	20024384
flatten (Flatten)	(None, 25088)	0
dense (Dense)	(None, 300)	7526700
dense_1 (Dense)	(None, 200)	60200
dense_2 (Dense)	(None, 3)	603

=====
 Total params: 27611887 (105.33 MB)
 Trainable params: 7587503 (28.94 MB)
 Non-trainable params: 20024384 (76.39 MB)

الف: مدل سازی با کردن تمامی لایه ها

Model: "model"

Layer (type)	Output Shape	Param #
input_2 (InputLayer)	[(None, 224, 224, 3)]	0
vgg19 (Functional)	(None, 7, 7, 512)	20024384
flatten (Flatten)	(None, 25088)	0
dense (Dense)	(None, 300)	7526700
dense_1 (Dense)	(None, 200)	60200
dense_2 (Dense)	(None, 3)	603

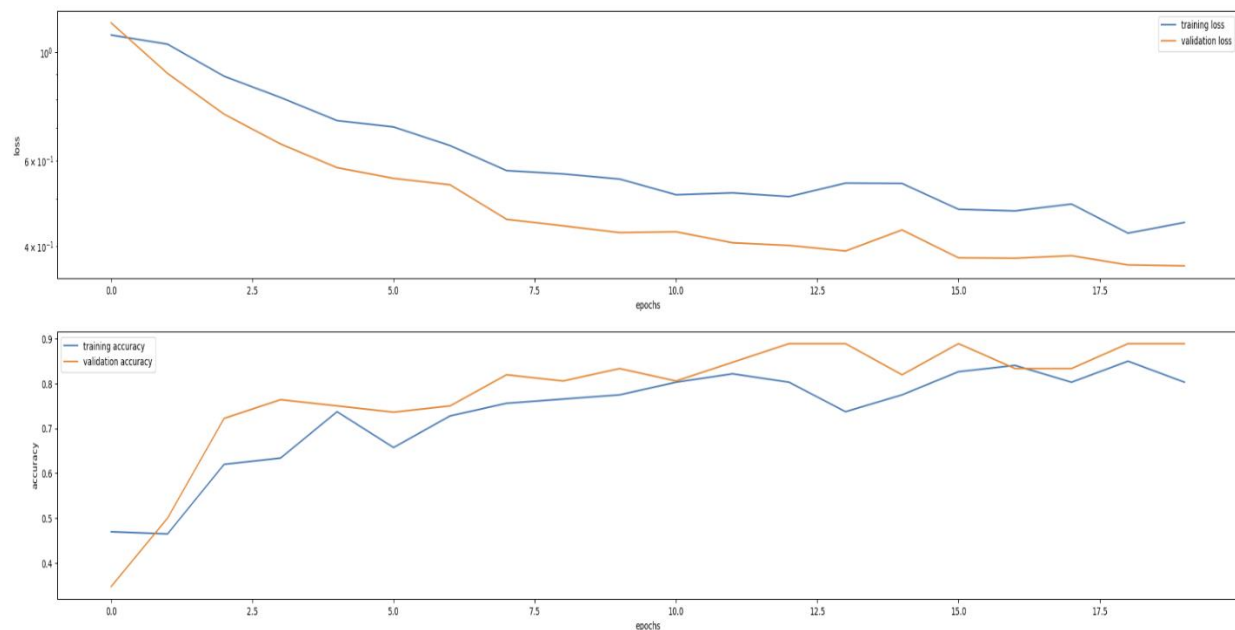
=====
 Total params: 27611887 (105.33 MB)
 Trainable params: 17026735 (64.95 MB)
 Non-trainable params: 10585152 (40.38 MB)

ب: مدل سازی با از فریز خارج کردن 5 لایه آخر

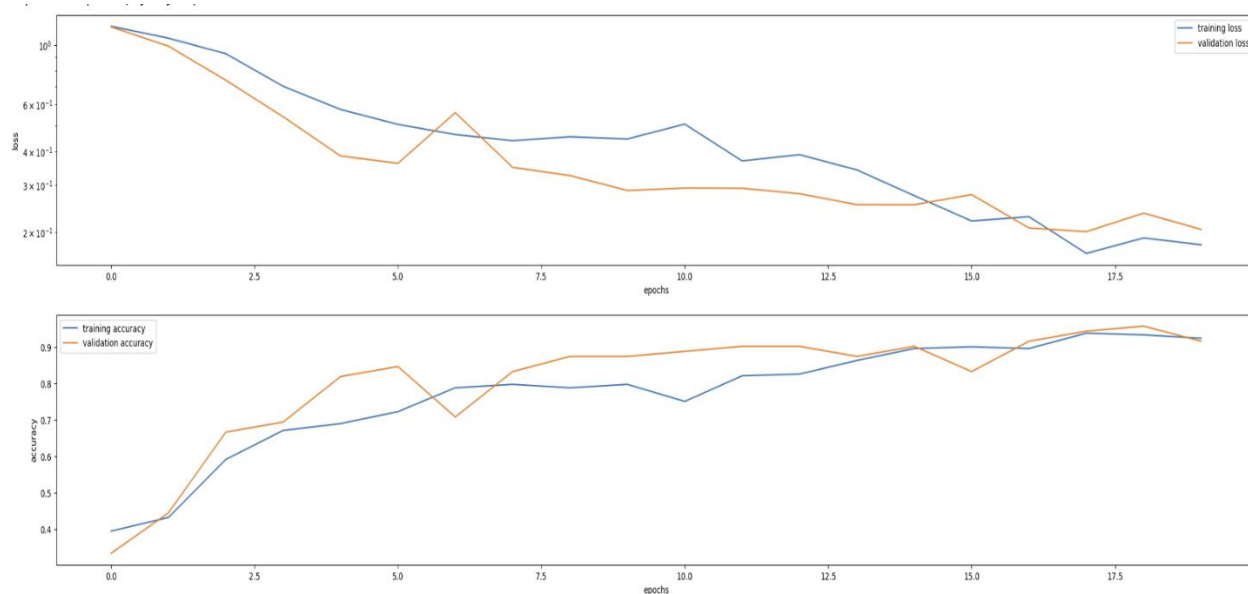
شکل 4-4: جزئیات مدل سازی با استفاده از شبکه VGG19

نتایج مدل سازی با استفاده از شبکه VGG19 در شکل های 4-5، 4-6 و 4-7 ارائه شده است.

نمودار دقت و تابع هزینه بر حسب دوره برای داده های آموزش و تست:



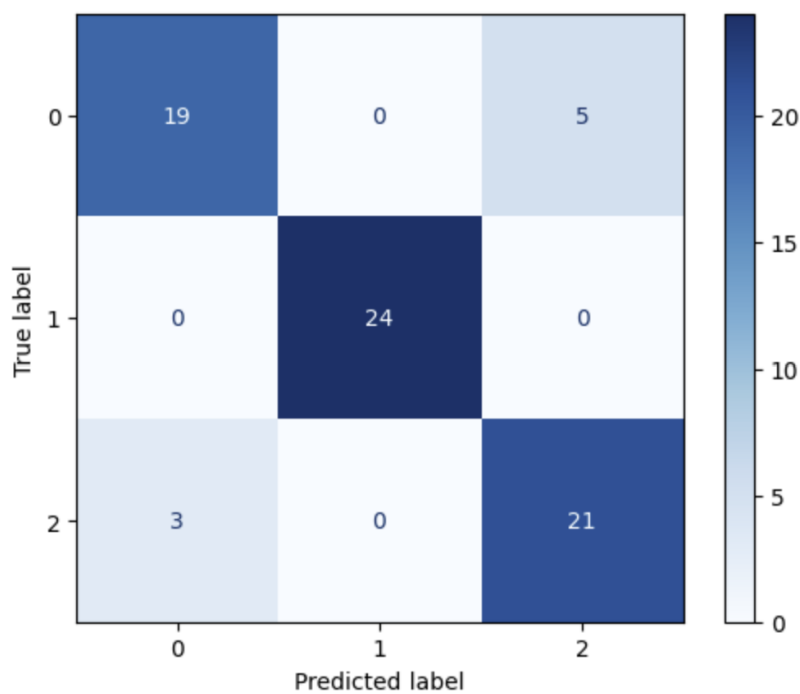
الف: مدل سازی با کردن تمامی لایه ها



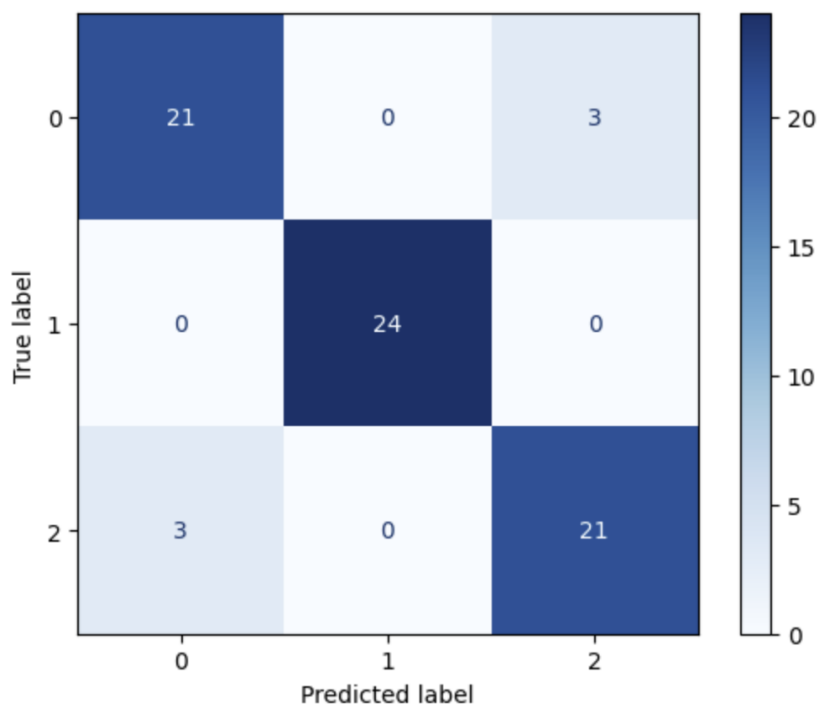
ب: مدل سازی با از فریز خارج کردن 5 لایه آخر

شکل 4-۵: نمودار دقت و تابع هزینه بر حسب دوره برای داده های آموزش و تست در مدل VGG19

ماتریس در هم ریختگی:



الف: مدل سازی با کردن تمامی لایه ها



ب: مدل سازی با از فریز خارج کردن 5 لایه آخر

شکل ۴-۶: ماتریس در هم ریختگی در مدل VGG19

گزارش طبقه بندی:

	precision	recall	f1-score	support
0	0.86	0.79	0.83	24
1	1.00	1.00	1.00	24
2	0.81	0.88	0.84	24
accuracy			0.89	72
macro avg	0.89	0.89	0.89	72
weighted avg	0.89	0.89	0.89	72

الف: مدل سازی با کردن تمامی لایه ها

	precision	recall	f1-score	support
0	0.88	0.88	0.88	24
1	1.00	1.00	1.00	24
2	0.88	0.88	0.88	24
accuracy			0.92	72
macro avg	0.92	0.92	0.92	72
weighted avg	0.92	0.92	0.92	72

ب: مدل سازی با از فریز خارج کردن 5 لایه آخر

شکل ۴-۷: گزارش طبقه بندی در مدل VGG19

بحث درباره نتایج مدل سازی با استفاده از شبکه VGG19:

همانطور که در شکل های 4-5 الف و ب مشخص است، دقت داده های آموزش از دقت داده های تست کمتر است و همچنین تابع هزینه داده های آموزش از تابع هزینه داده های تست بیشتر است. بنابراین می توان نتیجه گرفت که مدل های ساخته شده با استفاده از شبکه VGG19 در هر دو حالت (هم با فریز کردن همه لایه ها و هم با از فریز کردن 5 لایه آخر)، بیش برآزش شده اند.

همانطور که شکل های 4-6 الف و ب مشخص است مدل های ساخته شده با استفاده از شبکه VGG19 در هر دو حالت (هم با فریز کردن همه لایه ها و هم با از فریز کردن 5 لایه آخر)، پیش بینی درستی از داده ها مخصوصاً در داده های کلاس 1 (چشم سالم) داشته اند. اما با توجه به اینکه مدل ها بیش برآزش شده اند ممکن است در پیش بینی داده های جدید به درستی عمل نکنند.

با توجه به شکل های 4-7 الف و ب نیز می توان نتیجه گرفت که مدل های ساخته شده با استفاده از شبکه VGG19 در هر دو حالت (هم با فریز کردن همه لایه ها و هم با از فریز کردن 5 لایه آخر)، دقت مناسبی در پیش بینی داده ها دارند. اما به دلیل بیش برآزش شدن مدل این دقت ها قابل قبول نیستند.

2-2-4 مدل سازی با استفاده از شبکه ResNet:

در این مدل سازی با استفاده از شبکه ResNet ویژگی های دیتا ها استخراج می شود. سپس در ابتدای لایه های این شبکه یک لایه ورودی و به انتهای لایه های این شبکه یک لایه مسطح ساز و دو لایه کاملاً متصل و یک لایه خروجی به مشخصات موجود در جدول 3-4 اضافه شده و فرایند مدل سازی آغاز می شود. لازم به ذکر است که سایز عکس های ورودی به شبکه ResNet باید (224 , 224) باشد که در لایه ورودی تنظیم می شود.

این مدل کارایی لازم را بر روی دیتاست نداشت. بنابراین از ساخت شبکه با از فریز کردن 5 لایه آخر صرف نظر کرده و به ساخت شبکه با فریز کردن تمامی لایه ها بسنده می شود.

جزئیات این مدل در شکل 4-8 نمایش داده شده است.

Model: "model"

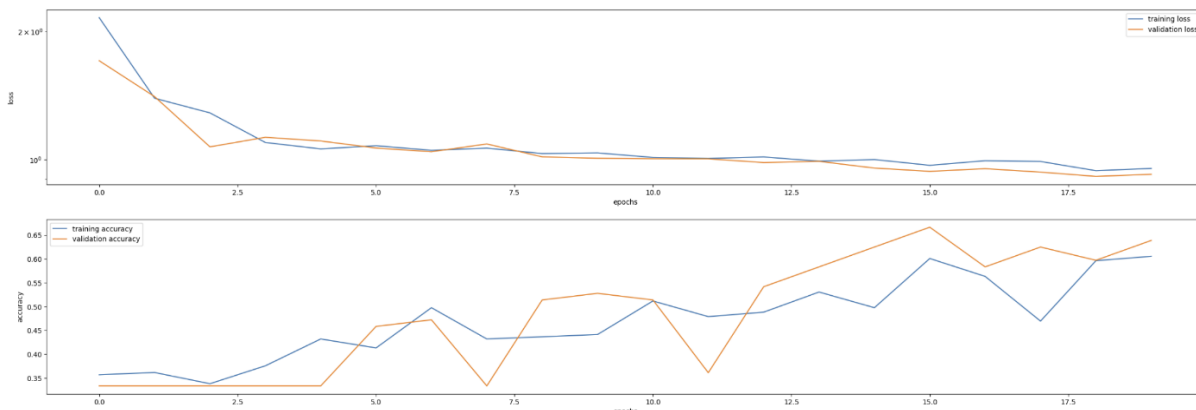
Layer (type)	Output Shape	Param #
input_2 (InputLayer)	[(None, 224, 224, 3)]	0
resnet50 (Functional)	(None, 7, 7, 2048)	23587712
flatten (Flatten)	(None, 100352)	0
dense (Dense)	(None, 300)	30105900
dense_1 (Dense)	(None, 200)	60200
dense_2 (Dense)	(None, 3)	603

=====
 Total params: 53754415 (205.06 MB)
 Trainable params: 30166703 (115.08 MB)
 Non-trainable params: 23587712 (89.98 MB)

شکل 4-۸: جزئیات مدل سازی با استفاده از شبکه ResNet با فریز کردن تمامی لایه ها

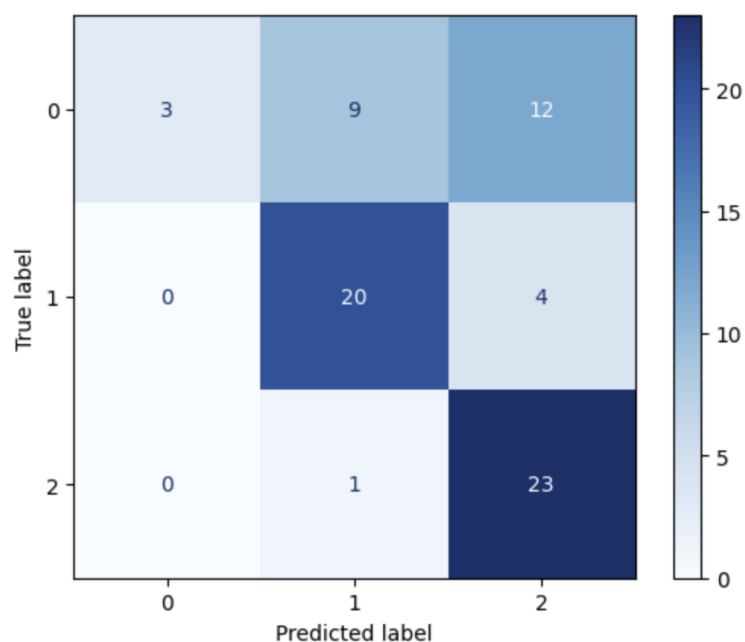
نتایج مدل سازی با استفاده از شبکه ResNet در شکل های 4-9، 4-10 و 4-11 ارائه شده است.

نمودار دقت و تابع هزینه بر حسب دوره برای داده های آموزش و تست:



شکل 4-۹: نمودار دقت و تابع هزینه بر حسب دوره برای داده های آموزش و تست در مدل ResNet با فریز کردن تمامی لایه ها

ماتریس در هم ریختگی:



شکل ۴-۱۰: ماتریس در هم ریختگی در مدل ResNet با فریز کردن تمامی لایه ها

گزارش طبقه بندی:

	precision	recall	f1-score	support
0	1.00	0.12	0.22	24
1	0.67	0.83	0.74	24
2	0.59	0.96	0.73	24
accuracy			0.64	72
macro avg	0.75	0.64	0.56	72
weighted avg	0.75	0.64	0.56	72

شکل ۴-۱۱: گزارش طبقه بندی در مدل ResNet با فریز کردن تمامی لایه ها

بحث درباره نتایج مدل سازی با استفاده از شبکه ResNet:

با توجه به شکل ۴-۹ مدل ساخته شده با استفاده از شبکه ResNet به طور مناسب آموزش دیده نشده است. همچنین همان طور که در شکل ۴-۱۰ مشخص است، مدل ساخته شده با استفاده از شبکه ResNet بیشتر داده ها را در کلاسی پیش‌بینی می‌کند که بیشترین داده های آموزش را دارد و نشان دهنده این است که مدل ساخته شده مدلی زیر برآزش است. از شکل ۴-۱۱ هم پایین بودن دقت مدل ساخته شده کاملاً مشهود است. بنابراین مدل ساخته شده با استفاده از شبکه ResNet به طور کلی مدل قابل قبولی نیست.

3-2-4 مدل سازی با استفاده از شبکه InceptionV3:

در این مدل سازی با استفاده از شبکه InceptionV3 ویژگی های دیتا ها استخراج می شود. سپس در ابتدای لایه های این شبکه یک لایه ورودی و به انتهای لایه های این شبکه یک لایه مسطح ساز و دو لایه کاملاً متصل و یک لایه خروجی به مشخصات موجود در جدول 3-4 اضافه شده و فرایند مدل سازی آغاز می شود. لازم به ذکر است که سائز عکس های ورودی به شبکه InceptionV3 باید (299, 299) باشد که در لایه ورودی تنظیم می شود.

جزئیات این مدل در شکل 4-12 نمایش داده شده است.

Model: "model"

Layer (type)	Output Shape	Param #
input_2 (InputLayer)	[(None, 299, 299, 3)]	0
inception_v3 (Functional)	(None, 8, 8, 2048)	21802784
flatten (Flatten)	(None, 131072)	0
dense (Dense)	(None, 300)	39321900
dense_1 (Dense)	(None, 200)	60200
dense_2 (Dense)	(None, 3)	603

=====
 Total params: 61185487 (233.40 MB)
 Trainable params: 39382703 (150.23 MB)
 Non-trainable params: 21802784 (83.17 MB)

الف: مدل سازی با کردن تمامی لایه ها

Model: "model"

Layer (type)	Output Shape	Param #
input_2 (InputLayer)	[(None, 299, 299, 3)]	0
inception_v3 (Functional)	(None, 8, 8, 2048)	21802784
flatten (Flatten)	(None, 131072)	0
dense (Dense)	(None, 300)	39321900
dense_1 (Dense)	(None, 200)	60200
dense_2 (Dense)	(None, 3)	603

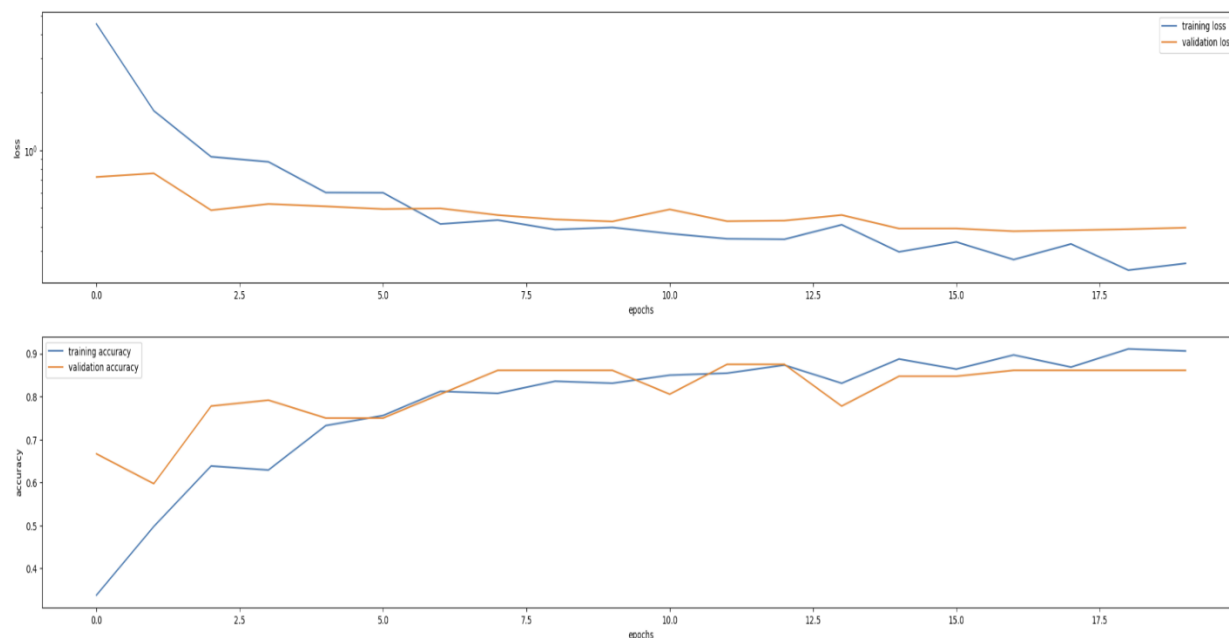
=====
 Total params: 61185487 (233.40 MB)
 Trainable params: 39382895 (150.23 MB)
 Non-trainable params: 21802592 (83.17 MB)

ب: مدل سازی با از فریز خارج کردن 5 لایه آخر

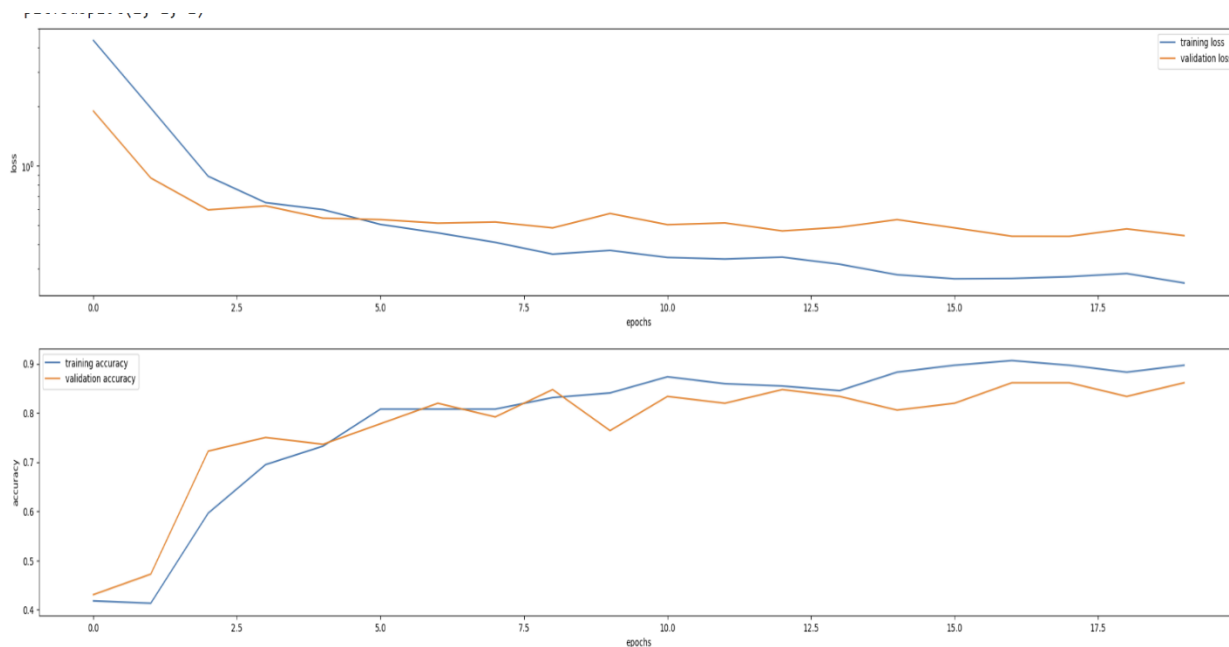
شکل 4-۱۲: جزئیات مدل سازی با استفاده از شبکه InceptionV3

نتایج مدل سازی با استفاده از شبکه InceptionV3 در شکل های 4-13 ، 4-14 و 4-15 ارائه شده است.

نمودار دقت و تابع هزینه بر حسب دوره برای داده های آموزش و تست:



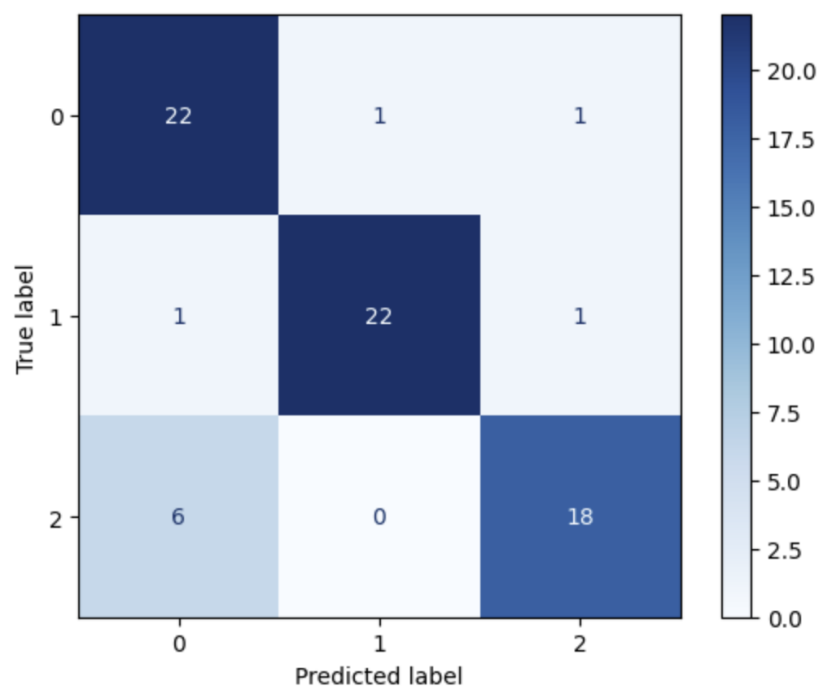
الف: مدل سازی با کردن تمامی لایه ها



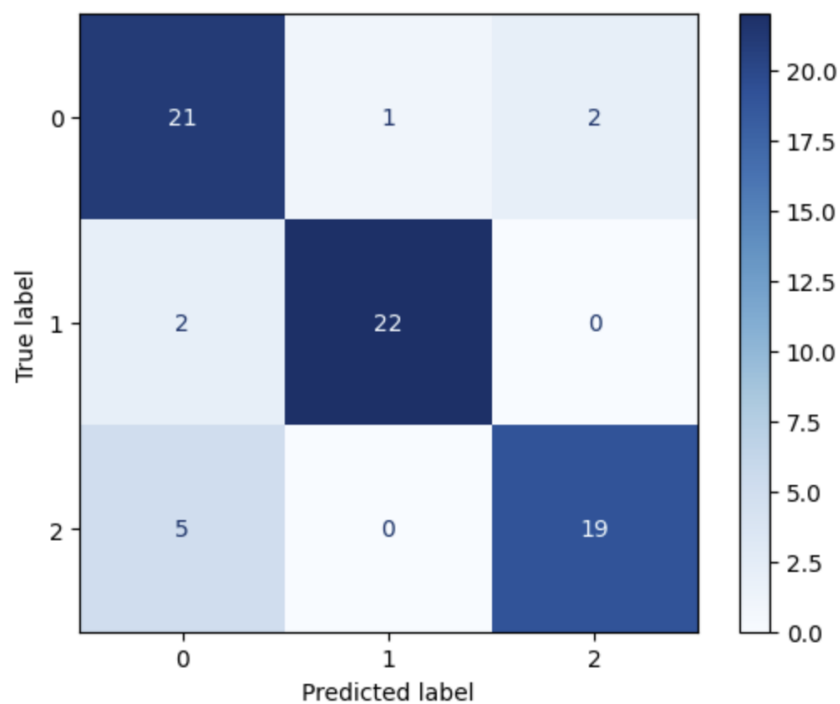
ب: مدل سازی با از فریز خارج کردن 5 لایه آخر

شکل 4-۱۳: نمودار دقت و تابع هزینه بر حسب دوره برای داده های آموزش و تست در مدل InceptionV3

ماتریس در هم ریختگی:



الف: مدل سازی با کردن تمامی لایه ها



ب: مدل سازی با از فریز خارج کردن 5 لایه آخر

شکل ۴-۱۴: ماتریس در هم ریختگی در مدل InceptionV3

گزارش طبقه بندی:

	precision	recall	f1-score	support
0	0.76	0.92	0.83	24
1	0.96	0.92	0.94	24
2	0.90	0.75	0.82	24
accuracy			0.86	72
macro avg	0.87	0.86	0.86	72
weighted avg	0.87	0.86	0.86	72

الف: مدل سازی با کردن تمامی لایه ها

	precision	recall	f1-score	support
0	0.75	0.88	0.81	24
1	0.96	0.92	0.94	24
2	0.90	0.79	0.84	24
accuracy			0.86	72
macro avg	0.87	0.86	0.86	72
weighted avg	0.87	0.86	0.86	72

ب: مدل سازی با از فریز خارج کردن 5 لایه آخر

شکل ۴-۱۵: گزارش طبقه بندی در مدل InceptionV3

بحث درباره نتایج مدل سازی با استفاده از شبکه InceptionV3:

با توجه به شکل های 4-13 الف و ب می توان نتیجه گرفت که مدل های ساخته شده با استفاده از شبکه InceptionV3 در هر دو حالت (هم با فریز کردن همه لایه ها و هم با از فریز کردن 5 لایه آخر)، به طور قابل قبولی آموزش دیده اند.

همانطور که شکل های 4-14 الف و ب مشخص است مدل های ساخته شده با استفاده از شبکه InceptionV3 در هر دو حالت (هم با فریز کردن همه لایه ها و هم با از فریز کردن 5 لایه آخر)، پیش بینی قابل قبولی از داده ها داشته اند.

با توجه به شکل های 4-15 الف و ب نیز می توان نتیجه گرفت که مدل های ساخته شده با استفاده از شبکه InceptionV3 در هر دو حالت (هم با فریز کردن همه لایه ها و هم با از فریز کردن 5 لایه آخر)، دقت قابل قبولی در پیش بینی داده ها دارند. همچنین مشاهده می شود که هر دو مدل دقت برابری دارند.

4-2-4 مدل سازی با استفاده از شبکه Xception:

در این مدل سازی با استفاده از شبکه Xception ویژگی های دیتا ها استخراج می شود. سپس در ابتدای لایه های این شبکه یک لایه ورودی و به انتهای لایه های این شبکه یک لایه مسطح ساز و دو لایه کاملاً متصل و یک لایه خروجی به مشخصات موجود در جدول 3-4 اضافه شده و فرایند مدل سازی آغاز می شود. لازم به ذکر است که سائز عکس های ورودی به شبکه Xception باید (299, 299) باشد که در لایه ورودی تنظیم می شود. جزئیات این مدل در شکل 4-16 نمایش داده شده است.

Model: "model"

Layer (type)	Output Shape	Param #
input_2 (InputLayer)	[(None, 299, 299, 3)]	0
xception (Functional)	(None, 10, 10, 2048)	20861480
flatten (Flatten)	(None, 204800)	0
dense (Dense)	(None, 300)	61440300
dense_1 (Dense)	(None, 200)	60200
dense_2 (Dense)	(None, 3)	603

=====
 Total params: 82362583 (314.19 MB)
 Trainable params: 61501103 (234.61 MB)
 Non-trainable params: 20861480 (79.58 MB)

الف: مدل سازی با کردن تمامی لایه ها

Model: "model"

Layer (type)	Output Shape	Param #
input_2 (InputLayer)	[(None, 299, 299, 3)]	0
xception (Functional)	(None, 10, 10, 2048)	20861480
flatten (Flatten)	(None, 204800)	0
dense (Dense)	(None, 300)	61440300
dense_1 (Dense)	(None, 200)	60200
dense_2 (Dense)	(None, 3)	603

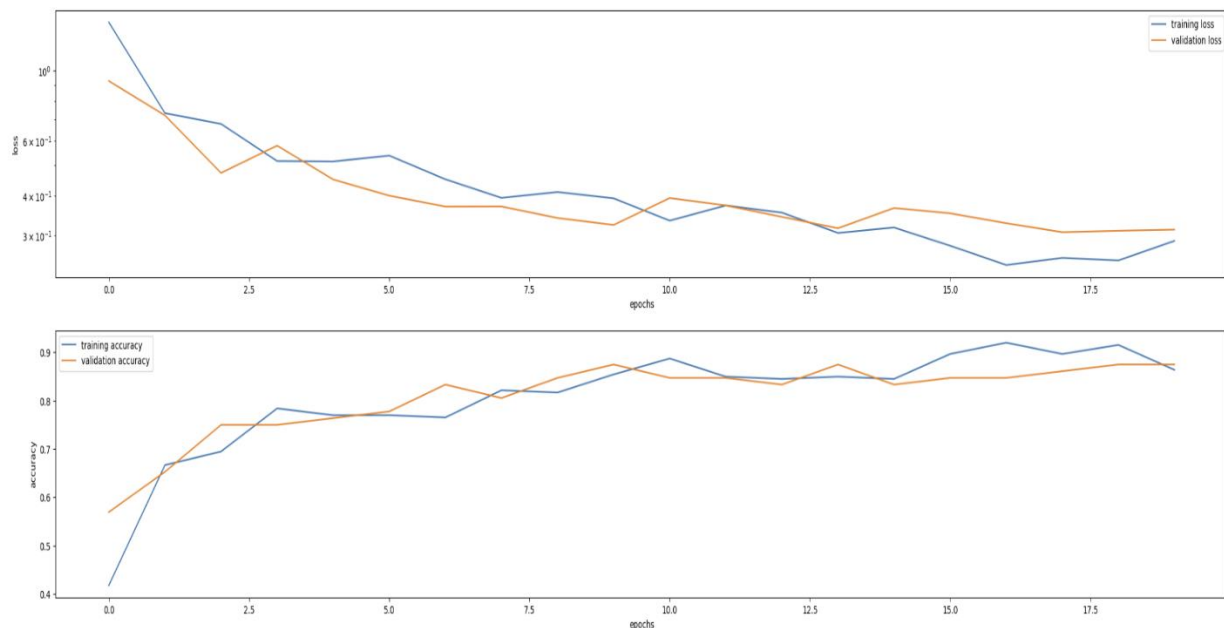
=====
 Total params: 82362583 (314.19 MB)
 Trainable params: 66249903 (252.72 MB)
 Non-trainable params: 16112680 (61.46 MB)

ب: مدل سازی با از فریز خارج کردن 5 لایه آخر

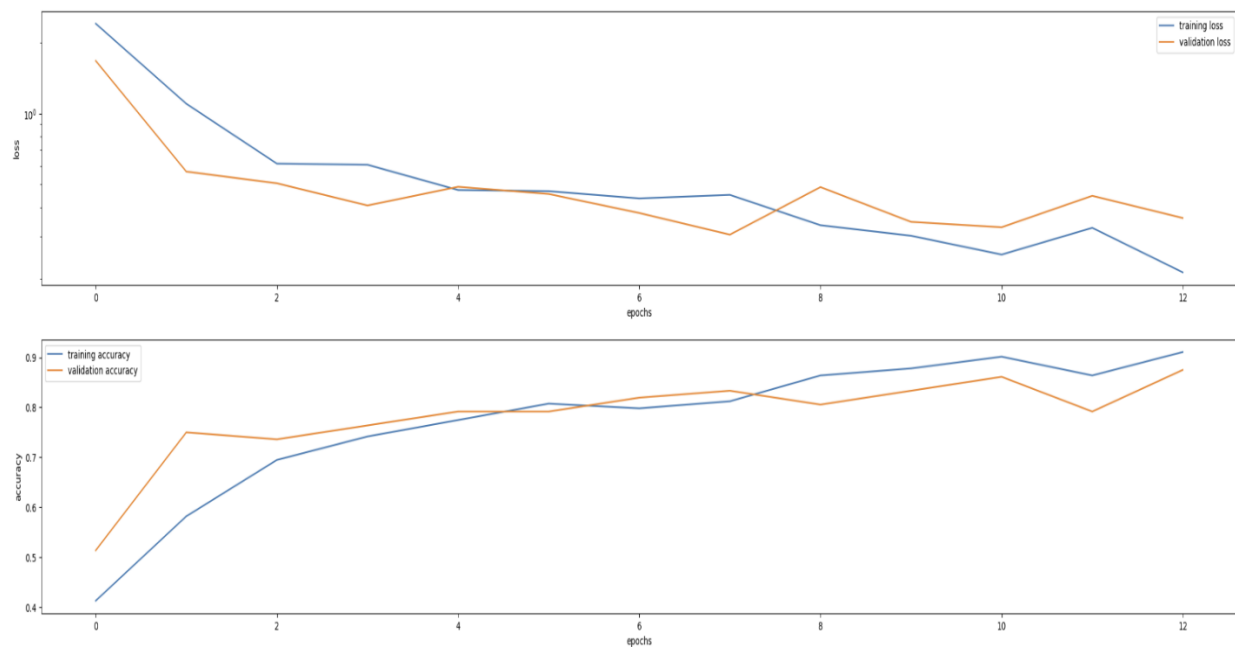
شکل ۴-۱۶: جزئیات مدل سازی با استفاده از شبکه Xception

نتایج مدل سازی با استفاده از شبکه Xception در شکل های 4-17، 4-18 و 4-19 ارائه شده است.

نمودار دقت و تابع هزینه بر حسب دوره برای داده های آموزش و تست:



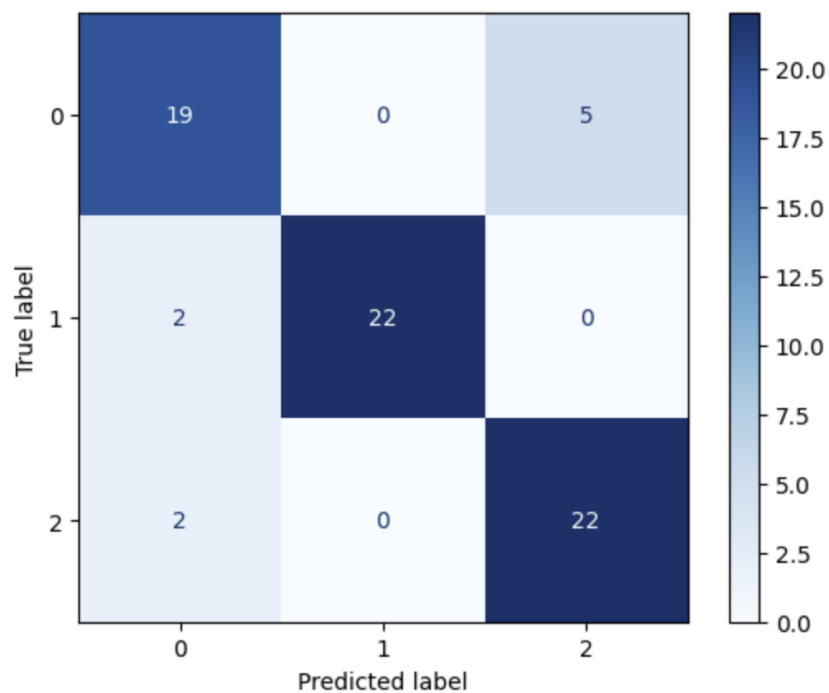
الف: مدل سازی با کردن تمامی لایه ها



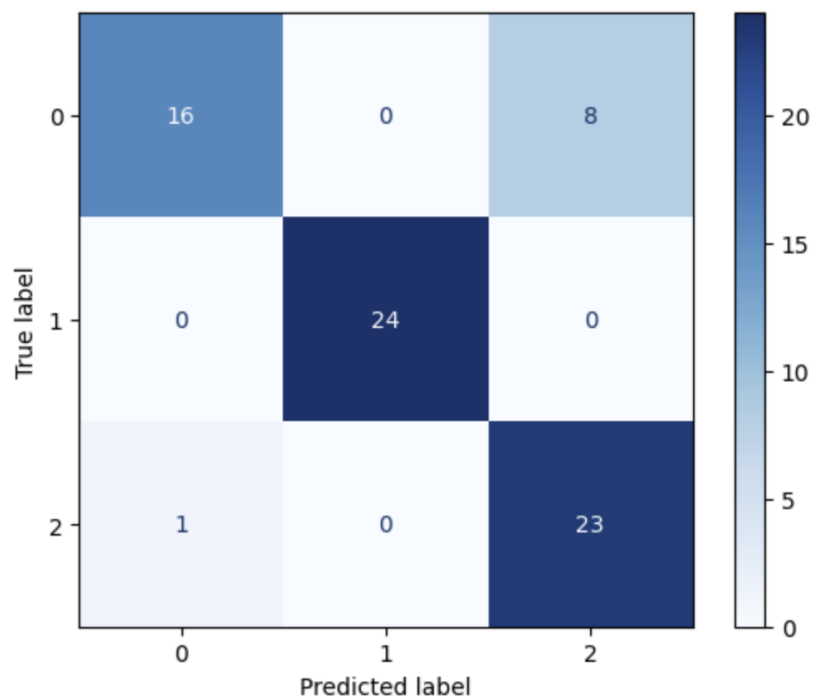
ب: مدل سازی با از فریز خارج کردن 5 لایه آخر

شکل 4-۱۷: نمودار دقت و تابع هزینه بر حسب دوره برای داده های آموزش و تست در مدل Xception

ماتریس در هم ریختگی:



الف: مدل سازی با کردن تمامی لایه ها



ب: مدل سازی با از فریز خارج کردن 5 لایه آخر

شکل ۴-۱۸: ماتریس در هم ریختگی در مدل Xception

گزارش طبقه بندی:

	precision	recall	f1-score	support
0	0.83	0.79	0.81	24
1	1.00	0.92	0.96	24
2	0.81	0.92	0.86	24
accuracy			0.88	72
macro avg	0.88	0.88	0.88	72
weighted avg	0.88	0.88	0.88	72

الف: مدل سازی با کردن تمامی لایه ها

	precision	recall	f1-score	support
0	0.94	0.67	0.78	24
1	1.00	1.00	1.00	24
2	0.74	0.96	0.84	24
accuracy			0.88	72
macro avg	0.89	0.88	0.87	72
weighted avg	0.89	0.88	0.87	72

ب: مدل سازی با از فریز خارج کردن 5 لایه آخر

شکل ۴-۱۹: گزارش طبقه بندی در مدل Xception

بحث درباره نتایج مدل سازی با استفاده از شبکه Xception:

با توجه به شکل های ۴-۱۷ الف و ب می توان نتیجه گرفت که مدل های ساخته شده با استفاده از شبکه Xception در هر دو حالت (هم با فریز کردن همه لایه ها و هم با از فریز کردن ۵ لایه آخر)، به طور قابل قبولی آموزش دیده اند.

همانطور که شکل های ۴-۱۸ الف و ب مشخص است مدل های ساخته شده با استفاده از شبکه Xception در هر دو حالت (هم با فریز کردن همه لایه ها و هم با از فریز کردن ۵ لایه آخر)، پیش بینی قابل قبولی از داده ها مخصوصاً داده های کلاس ۱ و ۲ (چشم سالم و بیماری Wet AMD) داشته اند.

با توجه به شکل های ۴-۱۹ الف و ب نیز می توان نتیجه گرفت که مدل های ساخته شده با استفاده از شبکه Xception در هر دو حالت (هم با فریز کردن همه لایه ها و هم با از فریز کردن ۵ لایه آخر)، دقت قابل قبولی در پیش بینی داده ها دارند. همچنین مشاهده می شود که هر دو مدل دقت برابری دارند.

3-4 مقایسه مدل های ساخته شده:

با مقایسه خروجی های مدل های ساخته شده می توان به راحتی نتیجه گرفت که مدل سازی با استفاده از شبکه InceptionV3 و Xception بهترین عملکرد را در بین مدل های ساخته شده دارد. همچنین مدل سازی با استفاده از شبکه ResNet عملکرد ضعیفی در پیش بینی داده ها دارد و مدل قابل قبولی نیست.

در حالت کلی مقایسه مدل های ساخته شده به صورت زیر است:

$Xception > InceptionV3 > VGG19 > ResNet$

5. فصل پنجم: جمع بندی و نتیجه گیری

5-1 آموخته های مدل سازی:

در این پروژه با استفاده از انواع الگوریتم های شبکه عصبی مصنوعی سعی بر این شد تا با داده های نسبتاً کمی که در اختیار داشتیم مدل های مختلفی بسازیم و مدلی با بهترین عملکرد را پیدا و آن را به عنوان خروجی این پروژه در نظر بگیریم. مقایسه مدل ها در انتهای فصل پنجم انجام شده است و مدلی که بهترین عملکرد را در میان مدل های ساخته شده داشت مدل سازی با استفاده از شبکه VGG19 با از فریز خارج کردن 5 لایه آخر است.

طبیعی است که با افزایش دیتا ها و جمع آوری تعداد بیشتری عکس OCT از بیمارستان می توانیم مدل هایی با دقت بالاتر و عملکرد بهتری بسازیم.

همچنین می توان از الگوریتم های سطح پیشرفته هوش مصنوعی یا به صورت پیشرفته تر از الگوریتم های یادگیری عمیق استفاده کرد استفاده کرد و مدلی با عملکرد بهتر ساخت. ذکر این نکته قابل توجه است که استفاده از الگوریتم های یادگیری عمیق به مجموعه داده بزرگتری نیاز دارد.

5-2 چشم انداز آینده:

مسلماً نخستین چشم انداز ما برای آینده بهبود عملکرد مدل سازی با گسترش مجموعه داده است.

گام بعدی، تعیین شدت بیماری های Wet AMD و Dry AMD در کنار تشخیص آن ها با عملکردی مناسب است. با استفاده از تعیین شدت بیماری ها می توان روند بهبود و وضعیت بیماران را مشخص کرد.

1. Munir, K., et al., *Cancer diagnosis using deep learning: a bibliographic review*. Cancers, 2019. **11**(9): p. 1235.
2. Haq, A.U., et al., *Detection of breast cancer through clinical data using supervised and unsupervised feature selection techniques*. IEEE Access, 2021. **9**: p. 22090-22105.
3. Kumar, G. and H. Alqahtani, *Deep Learning-Based Cancer Detection-Recent Developments, Trend and Challenges*. CMES-Computer Modeling in Engineering & Sciences, 2022. **130**(3).
4. Menaouer, B., et al., *Detection and classification of brain tumors from MRI images using a deep convolutional neural network approach*. International Journal of Software Innovation (IJSI), 2022. **10**(1): p. 1-25.
5. Jalali, V. and D. Kaur, *A study of classification and feature extraction techniques for brain tumor detection*. International Journal of Multimedia Information Retrieval, 2020. **9**(4): p. 271-290.
6. Bird, J.J., et al., *A study on CNN image classification of EEG signals represented in 2D and 3D*. Journal of Neural Engineering, 2021. **18**(2): p. 026005.
7. Zhou, J., et al., *Automatic vertebral body segmentation based on deep learning of Dixon images for bone marrow fat fraction quantification*. Frontiers in endocrinology, 2020. **11**: p. 555036.
8. Jaiswal, P. and S. Bhirud. *Study and analysis of an approach towards the classification of tooth wear in dentistry using machine learning technique*. in *2021 IEEE International Conference on Technology, Research, and Innovation for Betterment of Society (TRIBES)*. 2021. IEEE.
9. AKKOYUN, F. and A. ÖZÇELİK, *Rapid characterization of cell and bacteria counts using computer vision*. Türk Doğa ve Fen Dergisi, 2021. **10**(1): p. 269-274.
10. Nejad, R.B., et al. *A Multiclass Retinal Diseases Classification Algorithm using Deep Learning Methods*. in *2022 10th RSI International Conference on Robotics and Mechatronics (ICRoM)*. 2022. IEEE.
11. Daich Varela, M., et al., *Artificial intelligence in retinal disease: clinical application, challenges, and future directions*. Graefes Archive for Clinical and Experimental Ophthalmology, 2023. **261**(11): p. 3283-3297.
12. Wang, J., S. Wang, and Y. Zhang, *Artificial intelligence for visually impaired*. Displays, 2023. **77**: p. 102391.
13. Wichakam, I. and P. Vateekul. *Combining deep convolutional networks and SVMs for mass detection on digital mammograms*. in *2016 8th international conference on knowledge and smart technology (KST)*. 2016. IEEE.
14. Spanhol, F.A., et al. *Breast cancer histopathological image classification using convolutional neural networks*. in *2016 international joint conference on neural networks (IJCNN)*. 2016. IEEE.
15. Balagurunathan, Y., et al., *Reproducibility and prognosis of quantitative features extracted from CT images*. Translational oncology, 2014. **7**(1): p. 72-87.
16. Yan, K., et al. *Automatic prostate segmentation on MR images with deep network and graph model*. in *2016 38th Annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. 2016. IEEE.

17. Zhao, L. and K. Jia. *Deep feature learning with discrimination mechanism for brain tumor segmentation and diagnosis*. in *2015 international conference on intelligent information hiding and multimedia signal processing (IIH-MSP)*. 2015. IEEE.
18. Pereira, S., et al., *Brain tumor segmentation using convolutional neural networks in MRI images*. IEEE transactions on medical imaging, 2016. **35**(5): p. 1240-1251.
19. Pomponiu, V., H. Nejati, and N.-M. Cheung. *Deepmole: Deep neural networks for skin mole lesion classification*. in *2016 IEEE international conference on image processing (ICIP)*. 2016. IEEE.
20. Majtner, T., S. Yildirim-Yayilgan, and J.Y. Hardeberg. *Combining deep learning and hand-crafted features for skin lesion classification*. in *2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA)*. 2016. IEEE.
21. Feng, W., et al., *Automated MRI-based deep learning model for detection of Alzheimer's disease process*. International Journal of Neural Systems, 2020. **30**(06): p. 2050032.
22. Van Hespen, K.M., et al., *An anomaly detection approach to identify chronic brain infarcts on MRI*. Scientific Reports, 2021. **11**(1): p. 7714.
23. Busby, L.P., J.L. Courtier, and C.M. Glastonbury, *Bias in radiology: the how and why of misses and misinterpretations*. Radiographics, 2018. **38**(1): p. 236-247.
24. McDonald, R.J., et al., *The effects of changes in utilization and technological advancements of cross-sectional imaging on radiologist workload*. Academic radiology, 2015. **22**(9): p. 1191-1198.
25. Keelawat, P., et al., *Spatiotemporal emotion recognition using deep CNN based on EEG during music listening*. arXiv preprint arXiv:1910.09719, 2019.
26. Macsik, P., et al., *Local binary CNN for diabetic retinopathy classification on fundus images*. Acta Polytech. Hung, 2022. **19**(7): p. 27-45.
27. Tufail, A.B., et al., *Diagnosis of diabetic retinopathy through retinal fundus images and 3D convolutional neural networks with limited number of samples*. Wireless Communications and Mobile Computing, 2021. **2021**: p. 1-15.
28. Skouta, A., et al., *Hemorrhage semantic segmentation in fundus images for the diagnosis of diabetic retinopathy by using a convolutional neural network*. Journal of Big Data, 2022. **9**(1): p. 78.
29. Vinayaki, V.D. and R. Kalaiselvi, *Multithreshold image segmentation technique using remora optimization algorithm for diabetic retinopathy detection from fundus images*. Neural Processing Letters, 2022. **54**(3): p. 2363-2384.
30. Latif, J., et al., *ODGNet: a deep learning model for automated optic disc localization and glaucoma classification using fundus images*. SN Applied Sciences, 2022. **4**(4): p. 98.
31. Mahum, R., et al., *A novel hybrid approach based on deep CNN to detect glaucoma using fundus imaging*. Electronics, 2021. **11**(1): p. 26.
32. Elmoufidi, A., et al., *CNN with multiple inputs for automatic glaucoma assessment using fundus images*. International Journal of Image and Graphics, 2023. **23**(01): p. 2350012.
33. Elsonbaty, A.A., *Smart blind stick design and implementation*. International Journal of Engineering and Advanced Technology (IJEAT), 2021. **10**(5): p. 17-20.
34. Masud, U., et al., *Smart assistive system for visually impaired people obstruction avoidance through object detection and classification*. IEEE access, 2022. **10**: p. 13428-13441.

35. Kadhim, M.R. and B.K. Oleiwi, *Blind assistive system based on real time object recognition using machine learning*. Engineering and Technology Journal, 2022. **40**(1): p. 159-165.
36. Bhuiyan, A., et al., *Artificial intelligence to stratify severity of age-related macular degeneration (AMD) and predict risk of progression to late AMD*. Translational vision science & technology, 2020. **9**(2): p. 25-25.
37. Yan, Q., et al., *Deep-learning-based prediction of late age-related macular degeneration progression*. Nature machine intelligence, 2020. **2**(2): p. 141-150.
38. Klein, R., B.E. Klein, and K.L. Linton, *Prevalence of age-related maculopathy: the Beaver Dam Eye Study*. Ophthalmology, 1992. **99**(6): p. 933-943.
39. Wong, T.Y., G. Liew, and P. Mitchell, *Clinical update: new treatments for age-related macular degeneration*. The Lancet, 2007. **370**(9583): p. 204-206.
40. Wong, T.Y. and S.L. Rogers, *Statins and age-related macular degeneration: time for a randomized controlled trial?* American journal of ophthalmology, 2007. **144**(1): p. 117-119.
41. Grassmann, F., S. Fauser, and B.H. Weber, *The genetics of age-related macular degeneration (AMD)–novel targets for designing treatment options?* European Journal of Pharmaceutics and Biopharmaceutics, 2015. **95**: p. 194-202.
42. Holz, F.G., et al., *Progression of geographic atrophy and impact of fundus autofluorescence patterns in age-related macular degeneration*. American journal of ophthalmology, 2007. **143**(3): p. 463-472. e2.
43. Gass, J.D.M., *Drusen and disciform macular detachment and degeneration*. Archives of Ophthalmology, 1973. **90**(3): p. 206-217.
44. Hobbs, S.D. and K. Pierce, *Wet age-related macular degeneration (Wet AMD)*, in *StatPearls [Internet]*. 2022, StatPearls Publishing.
45. Mares, J.A., et al., *Healthy lifestyles related to subsequent prevalence of age-related macular degeneration*. Archives of ophthalmology, 2011. **129**(4): p. 470-480.
46. Lim, L.S., et al., *Age-related macular degeneration*. The Lancet, 2012. **379**(9827): p. 1728-1738.
47. Ahlers, C., et al., *Imaging of the retinal pigment epithelium in age-related macular degeneration using polarization-sensitive optical coherence tomography*. Investigative ophthalmology & visual science, 2010. **51**(4): p. 2149-2157.
48. Holz, F.G., et al., *Imaging protocols in clinical studies in advanced age-related macular degeneration: recommendations from classification of atrophy consensus meetings*. Ophthalmology, 2017. **124**(4): p. 464-478.
49. Chiang, A. and D. Lally, *Assessment of progression of geographic atrophy in the FILLY study*. Investigative Ophthalmology & Visual Science, 2020. **61**(7): p. 4307-4307.
50. Rosenfeld, P.J., et al., *Ranibizumab for neovascular age-related macular degeneration*. New England Journal of Medicine, 2006. **355**(14): p. 1419-1431.
51. Colijn, J.M., et al., *Prevalence of age-related macular degeneration in Europe: the past and the future*. Ophthalmology, 2017. **124**(12): p. 1753-1763.
52. Wong, W.L., et al., *Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis*. The Lancet Global Health, 2014. **2**(2): p. e106-e116.
53. Smith, W., et al., *Risk factors for age-related macular degeneration: pooled findings from three continents*. Ophthalmology, 2001. **108**(4): p. 697-704.

54. Klein, R., et al., *The Wisconsin age-related maculopathy grading system*. Ophthalmology, 1991. **98**(7): p. 1128-1134.
55. Chiu, S.J., et al., *Validated automatic segmentation of AMD pathology including drusen and geographic atrophy in SD-OCT images*. Investigative ophthalmology & visual science, 2012. **53**(1): p. 53-61.
56. Scholl, H.P., et al., *Inter-and intra-observer variability in grading lesions of age-related maculopathy and macular degeneration*. Graefes archive for clinical and experimental ophthalmology, 2003. **241**: p. 39-47.
57. Schmidt-Erfurth, U. and S.M. Waldstein, *A paradigm shift in imaging biomarkers in neovascular age-related macular degeneration*. Progress in Retinal and eye Research, 2016. **50**: p. 1-24.
58. Sadda, S.R., et al., *Consensus definition for atrophy associated with age-related macular degeneration on OCT: classification of atrophy report 3*. Ophthalmology, 2018. **125**(4): p. 537-548.
59. Toth, C.A., et al., *Identification of fluid on optical coherence tomography by treating ophthalmologists versus a reading center in the comparison of age-related macular degeneration treatments trials*. Retina, 2015. **35**(7): p. 1303-1314.
60. Bogunović, H., et al., *RETOUCH: The retinal OCT fluid detection and segmentation benchmark and challenge*. IEEE transactions on medical imaging, 2019. **38**(8): p. 1858-1874.
61. Bird, A.C., et al., *An international classification and grading system for age-related maculopathy and age-related macular degeneration*. Survey of ophthalmology, 1995. **39**(5): p. 367-374.
62. Bressler, N.M., et al., *Clinicopathologic correlation of drusen and retinal pigment epithelial abnormalities in age-related macular degeneration*. Retina, 1994. **14**(2): p. 130-142.
63. Liu, W. and H.F. Zhang, *Photoacoustic imaging of the eye: a mini review*. Photoacoustics, 2016. **4**(3): p. 112-123.
64. Das, S., et al., *Feasibility and clinical utility of handheld fundus cameras for retinal imaging*. Eye, 2023. **37**(2): p. 274-279.
65. Shirazi, M.F., et al., *Multi-modal and multi-scale clinical retinal imaging system with pupil and retinal tracking*. Scientific Reports, 2022. **12**(1): p. 9577.
66. Zhao, C., et al., *Longitudinal observation of subretinal fibrosis in Vogt-Koyanagi-Harada disease*. BMC ophthalmology, 2018. **18**: p. 1-8.
67. Shieber, S.M., *The Turing test: verbal behavior as the hallmark of intelligence*. 2004: Mit Press.