



دانشگاه صنعتی شریف
دانشکده مهندسی کامپیوتر

پایان نامه کارشناسی
مهندسی کامپیوتر

امتیازدهی خودکار ASPECTS بر روی تصاویر CT با استفاده از یادگیری ژرف

نگارش

مهتا فطرت

استاد راهنما

دکتر حمیدرضا ربیعی

تیر ۱۴۰۲

به نام خدا
دانشگاه صنعتی شریف
دانشکده مهندسی کامپیوتر

پایان نامه کارشناسی

این پایان نامه به عنوان تحقق بخشی از شرایط دریافت درجه کارشناسی است.

عنوان: امتیازدهی خودکار ASPECTS بر روی تصاویر CT با استفاده از یادگیری ژرف
نگارش: مهتا فطرت

کمیته ممتحنین

استاد راهنما: دکتر حمیدرضا ربیعی امضاء:

استاد مشاور: امضاء:

استاد مدعو: امضاء:

تاریخ:

سپاس

از استاد بزرگوارم که با کمک‌ها و راهنمایی‌های بی‌دریغشان، مرا در به سرانجام رساندن این پایان‌نامه یاری داده‌اند، تشکر و قدردانی می‌کنم. همچنین از همکاران عزیزی که با راهنمایی‌های خود در بهبود نگارش این نوشتار سهیم بوده‌اند، صمیمانه سپاسگزارم.

چکیده

سکته‌ی مغزی به عنوان دومین عامل مرگ و میر در جهان شناخته می‌شود. این عارضه می‌تواند آسیب‌های دائمی و جبران‌ناپذیری برای افراد مبتلا به همراه داشته‌باشد [۱]. بنابراین، تشخیص سریع سکته‌ی مغزی و درمان در مراحل اولیه، از اهمیت بسیار بالایی برخوردار است. امتیاز ASPECTS یک معیار برای ارزیابی وخامت سکته‌ی مغزی از روی تصاویر CT مغز می‌باشد. اما تشخیص صحیح این امتیاز از روی تصاویر اولیه، که نواحی آسیب‌دیده به خوبی در آن ظاهر نمی‌شود، حتی برای متخصصین این حوزه، امری چالش‌برانگیز است. به همین دلیل، در سال‌های اخیر، پژوهش‌های زیادی بر روی تشخیص خودکار امتیاز ASPECTS به انجام رسیده‌است. روش‌های خودکار امتیازدهی ASPECTS می‌توانند در تعیین وخامت سکته توسط متخصصین مورد استفاده قرار بگیرند و سرعت و دقت تشخیص و انتخاب روش‌های درمانی را افزایش دهند. در پژوهش حاضر، یک راهکار نظام‌مند برای تشخیص امتیاز ASPECTS مبتنی بر روش‌های یادگیری ژرف ارائه می‌شود. روش پیشنهادی این پروژه، به دقت ۰.۹۱۸ و AUC ۰.۹۶ دست یافته‌است که در مقایسه با سایر پژوهش‌های مشابه، عملکرد برجسته‌ای را نشان می‌دهد و می‌تواند در مطالعات آتی راهگشا باشد.

کلیدواژه‌ها: سکته‌ی مغزی، ASPECTS، یادگیری ژرف، CT، یادگیری انتقالی

فهرست مطالب

۱	مقدمه	۱
۱	۱-۱ تعریف مسئله	۱
۲	۲-۱ اهمیت موضوع	۲
۳	۳-۱ ادبیات موضوع	۳
۴	۴-۱ معیارهای ارزیابی	۴
۴	۵-۱ اهداف پژوهش	۴
۵	۶-۱ ساختار پایان نامه	۵
۶	۲ مفاهیم اولیه	۶
۶	۱-۲ مفاهیم پزشکی	۶
۶	۲-۱-۱ سکنه‌ی مغزی	۶
۷	۲-۱-۲ امتیاز ASPECTS	۷
۸	۲-۱-۳ نحوه‌ی امتیازدهی ASPECTS از روی تصاویر مغزی	۸
۱۰	۲-۱-۴ تصاویر پزشکی	۱۰
۱۱	۲-۲ مفاهیم فنی	۱۱
۱۱	۲-۲-۱ داده‌افزایی	۱۱
۱۲	۲-۲-۲ یادگیری انتقالی	۱۲
۱۲	۲-۲-۳ اعتبارسنجی متقابل	۱۲

۳-۲ جمع بندی ۱۴

۳ کارهای پیشین ۱۵

۳-۱ روش های مورد استفاده ۱۵

۳-۱-۱ روش ناحیه بندی و طبقه بندی ۱۵

۳-۱-۲ روش ناحیه بندی و هم پوشانی ۱۷

۳-۱-۳ روش کل نگری و طبقه بندی ۱۸

۳-۱-۴ سایر روش ها ۱۹

۳-۲ مجموعه داده و برچسب مورد استفاده ۱۹

۳-۲-۱ حجم مجموعه داده ۲۰

۳-۲-۲ نوع برچسب مجموعه داده ۲۱

۳-۳ کارهای پیشین در یک نگاه ۲۳

۳-۴ جمع بندی ۲۴

۴ روش پیشنهادی ۲۵

۴-۱ آماده سازی مجموعه داده ۲۵

۴-۱-۱ افزایش وضوح ۲۶

۴-۱-۲ کاهش نویز ۲۸

۴-۱-۳ تنظیم زاویه ۲۹

۴-۱-۴ محدود سازی تصویر به ناحیه ی مغز ۳۰

۴-۱-۵ حذف استخوان جمجمه ۳۰

۴-۲ داده افزایی ۳۲

۴-۳ ساختار ورودی و خروجی ۳۶

۴-۳-۱ ساختار ورودی ۳۶

۴-۳-۲ ساختار خروجی ۳۷

۴-۴ طراحی مدل ۳۹

۴-۵ جمع بندی ۴۰

۵ آزمایش ها ۴۱

۵-۱ آموزش مدل ۴۱

۵-۲ آزمایش مدل ۴۲

۵-۳ تشخیص های انجام شده ۴۴

۵-۴ نتایج ارزیابی ۴۶

۵-۵ مقایسه ی نتایج ۴۷

۵-۶ جمع بندی ۴۸

۶ نتیجه گیری ۴۹

۶-۱ کاربرد پذیری ۴۹

۶-۲ نقاط قوت ۵۰

۶-۳ کاستی ها ۵۱

۶-۴ پیشنهاد ادامه ی کار ۵۲

۶-۵ جمع بندی ۵۲

مراجع ۵۳

واژه نامه ۵۶

فهرست جدول‌ها

۲۳	۱-۳ فهرست کارهای پیشین
۴۴	۱-۵ اطلاعات مجموعه‌ی آزمایشی پنج لایه‌ی ارزیابی
۴۴	۲-۵ اطلاعات مجموعه‌ی آموزشی پنج لایه‌ی ارزیابی
۴۵	۳-۵ تشخیص‌های انجام‌شده توسط مدل در پنج لایه‌ی ارزیابی
۴۵	۴-۵ تشخیص‌های نهایی انجام‌شده توسط مدل
۴۶	۵-۵ معیارهای ارزیابی مدل در پنج لایه‌ی ارزیابی
۴۶	۶-۵ معیارهای ارزیابی نهایی مدل
۴۸	۷-۵ مقایسه‌ی نتایج با کارهای پیشین

فهرست شکل‌ها

۷	۱-۲ انواع سگته‌ی مغزی
۸	۲-۲ نواحی ASPECTS
۹	۳-۲ نواحی ASPECTS در ۸ برش مغز
۱۱	۴-۲ وضوح تصاویر CT
۲۰	۱-۳ توزیع امتیاز ASPECTS بیماران
۲۱	۲-۳ تعداد تصاویر به ازای امتیاز ASPECTS
۲۷	۱-۴ تصویر نمونه پیش‌پردازش
۲۸	۲-۴ تنظیم پهنا و سطح پنجره‌ی تصاویر
۲۹	۳-۴ اعمال فیلتر میانه بر تصاویر
۲۹	۴-۴ تنظیم زاویه‌ی تصاویر
۳۰	۵-۴ تنظیم ابعاد تصاویر
۳۱	۶-۴ حذف استخوان جمجمه از تصاویر
۳۲	۷-۴ مراحل کامل پیش‌پردازش
۳۳	۸-۴ داده‌افزایی به روش قرینه‌سازی تصادفی افقی
۳۳	۹-۴ داده‌افزایی به روش دوران تصادفی
۳۴	۱۰-۴ داده‌افزایی به روش جابجایی تصادفی
۳۴	۱۱-۴ داده‌افزایی به روش تغییر مقیاس تصادفی
۳۵	۱۲-۴ داده‌افزایی به روش تبدیل برشی تصادفی

۳۵	۴-۱۳ داده‌افزایی به روش زاویه‌ی دید تصادفی
۳۶	۴-۱۴ داده‌افزایی به روش مات‌سازی گاوسی تصادفی
۳۸	۴-۱۵ برش‌های انتخابی ASPECTS
۴۰	۴-۱۶ ساختار شبکه‌ی پیشنهادی
۴۳	۵-۱ شمایی از روش ارزیابی متقابل
۴۷	۵-۲ نمودار AUC-ROC مدل

فصل ۱

مقدمه

در این فصل مسئله‌ی اصلی پژوهش، یعنی امتیاز ASPECTS، به طور دقیق‌تر مورد بررسی قرار می‌گیرد و علت اهمیت بالای آن در زمینه‌ی سکتی مغزی عنوان می‌شود. سپس بررسی می‌شود که آیا این مسئله در پژوهش‌ها و محصولات مرتبط، حل شده‌است یا خیر و اینکه پژوهش حاضر چه مزیتی در این حوزه به همراه خواهد داشت. در پایان نیز ساختار کلی پایان‌نامه شرح داده می‌شود.

۱-۱ تعریف مسئله

سکتی مغزی یکی از علل مهم مرگ و میر و ناتوانی‌های اکتسابی در جهان است [۱]. امروزه روش‌های درمانی مختلفی برای بیماران مبتلا به این عارضه وجود دارد. اما تجویز روش درمانی مناسب برای هر بیمار با توجه به وضعیت وی، متفاوت است. درواقع لازم است که متخصصان، ملاک و معیاری از وضعیت پیشرفت و وخامت سکتی داشته باشند تا بتوانند یک روش درمانی را برای بیمار، مناسب یا نامناسب قلمداد کنند. یکی از مهم‌ترین این معیارها، امتیاز ASPECTS می‌باشد.

The Alberta Stroke Program Early CT Score (ASPECTS) [۲]، یک امتیاز از ۰ تا ۱۰ است که معیاری از وخامت سکتی در بیماران را به دست می‌دهد. درواقع این امتیاز از بررسی وضعیت ۱۰ ناحیه‌ی مغزی، که در دو نیم‌کره‌ی مغز به صورت متقارن وجود دارند، محاسبه می‌شود. در صورتی که هیچ عارضه‌ی ایسکمیک در مغز وجود نداشته باشد، امتیاز ASPECTS برابر ۱۰ خواهد بود و به ازای هر ناحیه‌ی آسیب‌دیده، یک امتیاز از ۱۰ کم می‌شود. به این ترتیب بیماری که هر ۱۰ ناحیه‌ی ASPECTS او، حداقل در یک نیم‌کره، آسیب دیده‌باشد، امتیاز صفر را دریافت خواهد کرد. این امتیاز سپس می‌تواند معیاری

در اختیار متخصصان قرار دهد که تشخیص بدهند آیا درمان لخته‌زدایی مکانیکی^۱ برای بیمار مناسب است یا خیر.

نکته‌ی حائز اهمیت آن است که امتیازدهی ASPECTS، حتی برای متخصصین این حوزه، یک امر چالش‌برانگیز است. به نحوی که در یک مطالعه [۳]، میزان توافق میان امتیازدهندگان، تنها ۲۸٪ محاسبه شده‌است. از طرفی، نشان داده شده‌است که ابزارهای محاسبه‌ی خودکار ASPECTS، می‌توانند میزان این توافق و سرعت امتیازدهی متخصصان را افزایش دهند [۴]. به همین جهت، این پژوهش قصد دارد با ارائه‌ی یک روش خودکار تشخیص امتیاز دوبخشی ASPECTS در راستای این بهبود دقت و سرعت، راهگشا باشد.

۱-۲ اهمیت موضوع

میان دقت، سرعت و دسترس‌پذیری در تشخیص سکته‌ی مغزی، یک بده‌بستان^۲ وجود دارد. یک‌سری تصاویر مانند MRI، علائم سکته را بهتر در خود نمایان کرده و تشخیص را برای متخصصان ساده‌تر می‌کنند. اما اخذ این تصاویر، زمان زیادی نیاز دارد و ممکن است در تمام مراکز تصویربرداری نیز در دسترس نباشند. از سوی دیگر، تصاویر CT، علائم سکته را کمتر مشخص می‌کنند و باعث می‌شوند که تشخیص، سخت‌تر و توافق میان تشخیص‌دهندگان کمتر شود. اما مزیت این مدل تصویربرداری، در سرعت اخذ تصویر و کاربرد فراگیر آن در اکثر مراکز تصویر برداری است.

اصطلاحی در این حوزه وجود دارد که عنوان می‌کند "زمان، مغز است" [۵]. این جمله به اهمیت زمان و لزوم تشخیص و درمان سریع سکته‌ی مغزی اشاره می‌کند. به طور متوسط، در بیمارانی که دچار سکته‌ی مغزی ایسکمیک شده‌اند، در هر دقیقه، ۱.۹ میلیون سلول عصبی از بین می‌رود. این عدد در مقایسه با نرخ عادی از بین رفتن سلول‌های عصبی، مانند آن است که مغز در یک ساعت، به مدت ۳.۶ سال عمر کرده‌است [۵]. به همین جهت، سرعت عمل در تشخیص سکته‌ی مغزی و آغاز هر چه زودتر درمان آن، امری حیاتی است. در نتیجه در بده‌بستان میان دقت و سرعت، این سرعت است که برتری می‌یابد و تصویربرداری CT و روش‌های تشخیصی ممکن بر روی آن را غالب می‌کند.

امتیازدهی ASPECTS نیز یک روش تشخیصی مبتنی بر تصاویر CT است. به همین دلیل است که پژوهش حول این مسئله، از اهمیت بالایی برخوردار است. اما همانطور که پیش‌تر ذکر شد، علی‌رغم سرعت بالای تشخیص در این روش، افزایش دقت حاصل از آن، یک موضوع چالش‌برانگیز است. عدم

^۱ Mechanical Thrombectomy
^۲ Tradeoff

توافق بالا میان تشخیص متخصصان نیز خبر از این مشکل دارد. مشکلی که همچنان میان متخصصان انسانی ماندگار است. هوش مصنوعی و روش‌های یادگیری ماشین می‌توانند به حل این مشکل کمک کنند. پژوهش‌هایی انجام شده‌است که نشان می‌دهد تشخیص خودکار امتیاز ASPECTS می‌تواند توافق میان متخصصان را افزایش بدهد [۴]. بنابراین ضروری است که این روش‌ها، با افزایش هرچه بیشتر دقت، در راستای بهبود سرعت و دقت خدمات درمانی سکتای مغزی، کمک‌کننده باشند.

۳-۱ ادبیات موضوع

فعالیت‌هایی که به طور مستقیم در حوزه‌ی امتیازدهی ASPECTS انجام می‌شوند را می‌توان در دو دسته‌ی کلی بررسی کرد. دسته‌ی اول، برنامه‌های کاربردی^۳ هستند که به صورت تجاری عرضه شده و در حال استفاده در مراکز درمانی می‌باشند. از جمله‌ی این برنامه‌ها می‌توان به RapidAI [۶]، Viz.ai [۷] و e-ASPECTS [۸] اشاره کرد. بعضاً این برنامه‌ها بر روی چندین میلیون تصویر از بیش از ۱۰۰ کشور دنیا آموزش دیده‌اند و به دقت بسیار مطلوبی دست یافته‌اند [۶].

دسته‌ی دوم شامل پژوهش‌هایی می‌شود که بر روی تعداد تصویرهای بسیار کوچک‌تری کار می‌کنند. مجموعه داده‌هایی^۴ که محدود به یک یا چند مرکز درمانی می‌شوند و از نظر تنوع و تعداد، با برنامه‌های فوق‌الذکر قابل مقایسه نیستند. این پژوهش‌ها سعی دارند روش‌های جدید برای تشخیص ASPECTS ارائه دهند و یا توانایی مدل‌های یادگیری پیشین را بر روی مسئله‌ی ASPECTS بررسی کنند. این مطالعات و پژوهش‌های انجام شده، هر یک با در نظر گرفتن محدودیت‌های موجود، مورد ارزیابی قرار می‌گیرند. روش‌های پیش‌رو، سپس می‌توانند در هسته‌ی محاسباتی برنامه‌های تجاری قرار بگیرند و با استفاده از ظرفیت‌های داده‌ای و محاسباتی موجود، نتایج بهتری را ارائه دهند.

بنابراین، دو دسته فعالیت‌هایی که در حوزه‌ی ASPECTS معرفی شد، یعنی برنامه‌های کاربردی توانمند و فعالیت‌های پژوهشی، هر دو نیاز هستند و به نحوی مکمل هم می‌باشند. پژوهش حاضر، در دسته‌ی دوم این فعالیت‌ها قرار می‌گیرد و در ادامه‌ی این نوشتار نیز، تنها پژوهش‌های مطالعاتی انجام شده در حوزه‌ی ASPECTS مورد بررسی، ارجاع و مقایسه قرار خواهند گرفت.

در فصل سوم، این پژوهش‌ها به تفصیل بیشتری مورد بحث قرار می‌گیرند و محدودیت‌ها و مزیت‌های هر یک بررسی می‌شود. به طور کلی، کارهای پیشین از نظر میزان داده‌ی موجود، نوع اطلاعات برچسب^۵ داده‌ها، نوع اطلاعات خروجی و ... قابل دسته‌بندی و مقایسه هستند. در این فصل ذکر خواهد شد که

Application^۳
Dataset^۴
Label^۵

پژوهش حاضر، یکی از معدود مطالعاتی است که با محدودیت‌های داده‌ای مشابه انجام شده‌است و در این زمینه به نتایج بسیار مطلوبی دست یافته‌است.

۴-۱ معیارهای ارزیابی

با توجه به نتایج گزارش شده در کارهای پیشین، مدل پیشنهادی در این پروژه بر اساس تعدادی از معیارهای پرکاربرد یادگیری ماشین در حوزه‌ی طبقه‌بندی تصاویر ارزیابی می‌شود. این معیارها عبارتند از دقت،^۶ حساسیت،^۷ تشخیص،^۸ صحت،^۹ بازیابی^{۱۰} و مساحت زیر نمودار مشخصه‌ی عملیاتی گیرنده^{۱۱}. اگرچه هرکدام از این معیارها اطلاعات به‌خصوصی را در رابطه با ویژگی‌های مدل به دست می‌دهند، دو معیار دقت و AUC می‌توانند یک درک کلی از توانایی تشخیصی مدل را فراهم کنند. در فصل آزمایش‌ها، عملکرد مدل در هر یک از این معیارها گزارش خواهد شد.

۵-۱ اهداف پژوهش

پیش‌تر ذکر شد که محدودیت‌های داده‌ای، تاثیر به‌سزایی در توانایی و عملکرد روش‌های یادگیری ماشین دارند. یکی از مهم‌ترین چالش‌های حوزه‌ی یادگیری ماشین نیز در کسب بهترین نتایج از داده‌های محدود – چه از نظر کمی و چه از نظر کیفی – می‌باشد. از طرفی فراهم کردن مجموعه‌داده‌های بزرگی که توسط متخصصان به صورت جزئی برچسب‌گذاری شده‌باشند، امری دشوار، زمان‌بر و گاه غیر عملی است. بنابراین، ارائه‌ی روش‌هایی که بتوانند از حداکثر قابلیت‌های چنین مجموعه‌داده‌هایی استفاده کنند از اهمیت بالایی برخوردار است.

این پژوهش در وهله‌ی اول می‌کوشد تا ظرفیت موجود در داده‌های یکی از مراکز درمانی را در زمینه‌ی تشخیص سکته‌ی مغزی بسنجد و سپس روش کارآمد و تکرارپذیری را در حوزه‌ی یادگیری تصاویر پزشکی، به طور خاص محاسبه‌ی ASPECTS، ارائه کند. این روش علی‌رغم محدودیت‌های موجود، به عملکرد قابل مقایسه‌ای با کارهای مشابه دست یافته‌است و به علت جامعیت بالا، با تنظیمات جزئی، قابل اعمال بر روی سایر کاربردهای پزشکی می‌باشد.

Accuracy^۶
Sensitivity^۷
Specificity^۸
Precision^۹
Recall^{۱۰}
AUC^{۱۱}

۱-۶ ساختار پایان نامه

این پایان نامه در شش فصل به شرح زیر ارائه می شود. برخی مفاهیم اولیه در رابطه با سکتی مغزی ایسکمیک و امتیاز ASPECTS در بخش دوم به اختصار اشاره شده است. این مفاهیم از آن جهت اهمیت دارند که انطباق ساختار مدل ارائه شده با روش های مورد استفاده متخصصان را بهتر مشخص می کند. همچنین در درک روش های مختلف ارائه شده در کارهای پیشین و نیامندی های داده ای هر یک راهگشا خواهد بود.

فصل سوم به مطالعه و بررسی کارهای پیشین مرتبط با امتیازدهی خودکار ASPECTS می پردازد. در فصل چهارم، روش مورد استفاده در پژوهش حاضر شرح داده می شود و در بخش پنجم، نتایج حاصله از این روش عنوان می شوند. در نهایت، فصل ششم به جمع بندی کارهای انجام شده، موفقیت ها و ناکارآمدی های متصور برای این پژوهش و ارائه پیشنهادهایی برای انجام کارهای آتی خواهد پرداخت.

فصل ۲

مفاهیم اولیه

در این فصل به ذکر برخی مفاهیم اولیه‌ی مورد ارجاع در ادامه‌ی پایان‌نامه پرداخته می‌شود. این مفاهیم در دو دسته‌ی پزشکی و فنی قابل بررسی هستند. در ابتدا، مبانی پزشکی موضوع پروژه و نکاتی حول تصاویر پزشکی مورد استفاده به اختصار شرح داده می‌شود. در ادامه نیز نکاتی در رابطه با هسته‌ی فنی پروژه و روش‌های مورد استفاده در یادگیری ماشین ذکر می‌گردد. به این ترتیب، این فصل می‌تواند در دستیابی به یک دانش مشترک در میان متخصصان هر دو حوزه اثربخش باشد.

۱-۲ مفاهیم پزشکی

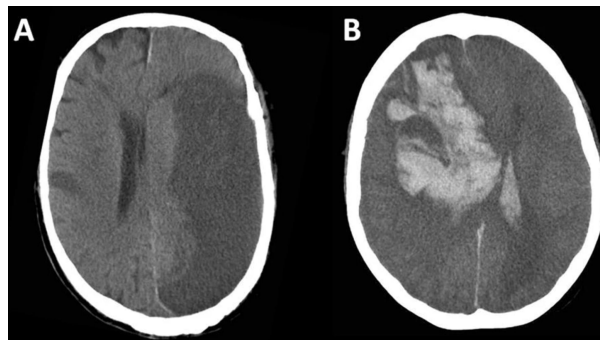
در این قسمت ابتدا نکاتی در رابطه با سکت‌های مغزی و انواع آن ذکر می‌شود. سپس امتیاز ASPECTS به عنوان یکی از مهم‌ترین روش‌های تشخیصی سکت‌ها و محور اصلی پروژه تعریف شده و به نحوه‌ی تشخیص آن اشاره می‌شود. در انتها نیز نکاتی در رابطه با تصاویر پزشکی و چالش‌های استفاده از آن‌ها در قسمت فنی مطرح می‌گردد.

۱-۱-۲ سکت‌های مغزی

انواع سکت‌های مغزی شامل دو دسته‌ی کلی سکت‌های ایسکمیک^۱ و هموراژیک^۲ هستند. سکت‌های مغزی ایسکمیک به علت قطع شدن جریان خون به بخشی از مغز رخ می‌دهد که باعث از دست رفتن ناگهانی عملکرد آن ناحیه می‌شود. در مقابل، سکت‌های مغزی هموراژیک از پاره‌شدن یک رگ خونی و یا ساختار

^۱ Ischemic
^۲ Haemorrhagic

غیر طبیعی عروقی نشأت می‌گیرد. در یک نگاه کلی، تقریباً ۸۰٪ بیماران سکته‌ی مغزی، در دسته‌ی اول، یعنی سکته‌ی ایسکمیک، قرار می‌گیرند [۱]. این دو نوع سکته‌ی مغزی، ظاهر متفاوتی در تصاویر CT به خود می‌گیرند. تصویر ۱-۲ این تفاوت را نشان می‌دهد. همانطور که در این تصویر نمایان است، عموماً تشخیص ناحیه‌ی درگیری در سکته‌ی هموراژیک ساده‌تر است و در مقابل، تشخیص این نواحی در سکته‌ی ایسکمیک، ظرافت و دقت بیشتری نیاز دارد. همانطور که در تعریف امتیاز ASPECTS خواهد آمد، پژوهش حاضر نیز، زیرمجموعه‌ی سکته‌های ایسکمیک قرار می‌گیرد و تمام مفاهیم مورد اشاره در این پایان‌نامه و تمام تصاویر مورد استفاده نیز به این نوع سکته اشاره خواهند داشت.



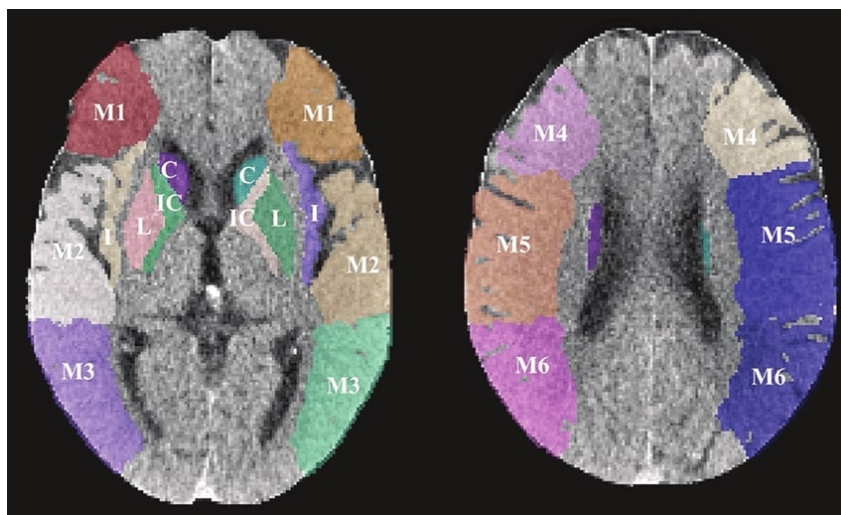
شکل ۱-۲: انواع سکته‌ی مغزی در تصاویر CT [۹]. برش مغزی A یک نمونه سکته‌ی ایسکمیک و برش B یک نمونه از سکته‌ی هموراژیک در این تصاویر را نشان می‌دهد.

۲-۱-۲ امتیاز ASPECTS

ASPECTS^۳ یک امتیاز عددی از ۰ تا ۱۰ است که میزان پیشرفت تغییرات حاصل از سکته‌ی ایسکمیک را نشان می‌دهد. امتیازدهی ASPECTS، قلمرو رگ مغزی میانی در مغز را به ۱۰ ناحیه‌ی مشخص تقسیم می‌کند (تصویر ۲-۲). امتیازدهی از ۱۰ آغاز می‌شود و به ازای هر کدام از این ۱۰ ناحیه که علائم کاهش جریان خون را نشان می‌دهند، یک امتیاز از ۱۰ کم می‌شود. این امتیاز برای تجویز لخته‌زدایی‌های درون‌رگی و برون‌رگی برای بیماران به‌کار می‌آید [۱۰].

این امتیاز در ابتدا به این منظور طراحی شد که در تشخیص بیماران که نتایج بهتری از لخته‌زدایی درون‌رگی کسب خواهند کرد، کمک کننده باشد. بعدها از این امتیاز برای تشخیص بیماران استفاده شد که برای درمان لخته‌زدایی برون‌رگی مناسب نیستند. درواقع عملیات باز کردن رگ‌ها^۴ در بیماران که علائم سکته‌ی ایسکمیک در نواحی وسیعی از مغزشان گسترده شده، می‌تواند بی‌اثر یا حتی زیان‌بار باشد. اخیراً هم این امتیاز در مجموعه‌ی دستورالعمل‌های مدیریت سکته‌ی مغزی انجمن قلب آمریکا به عنوان یک معیار کلیدی در تجویز لخته‌زدایی برون‌رگی عنوان شده‌است. به نحوی که این روش درمانی برای بیماران با

^۳The Alberta Stroke Program Early CT Score
^۴Recanalization



شکل ۲-۲: نواحی ASPECTS در دو برش از مغز [۱۱]. ۱۰ ناحیه شامل M1-M6، IC، L، C، I.

امتیاز $ASPECTS \geq 6$ توصیه می‌شود [۱۰]. با این تفسیر، مشخص می‌شود علی‌رغم ۱۰ امتیازی بودن ASPECTS، معمولاً آنچه که اهمیت دارد، تنها یک حد آستانه بر روی این امتیاز است. به این نوع از امتیازدهی که وضعیت بیماران را به دو دسته‌ی بالا و پایین یک آستانه (مثلاً $ASPECTS \geq 6$) تقسیم می‌کند، امتیاز دوبخشی شده^۵ ASPECTS می‌گویند. لازم به ذکر است که خروجی نهایی پژوهش حاصل و نتایج گزارش شده برای آن نیز، از نوع امتیازدهی دوبخشی خواهند بود.

۲-۱-۳ نحوه‌ی امتیازدهی ASPECTS از روی تصاویر مغزی

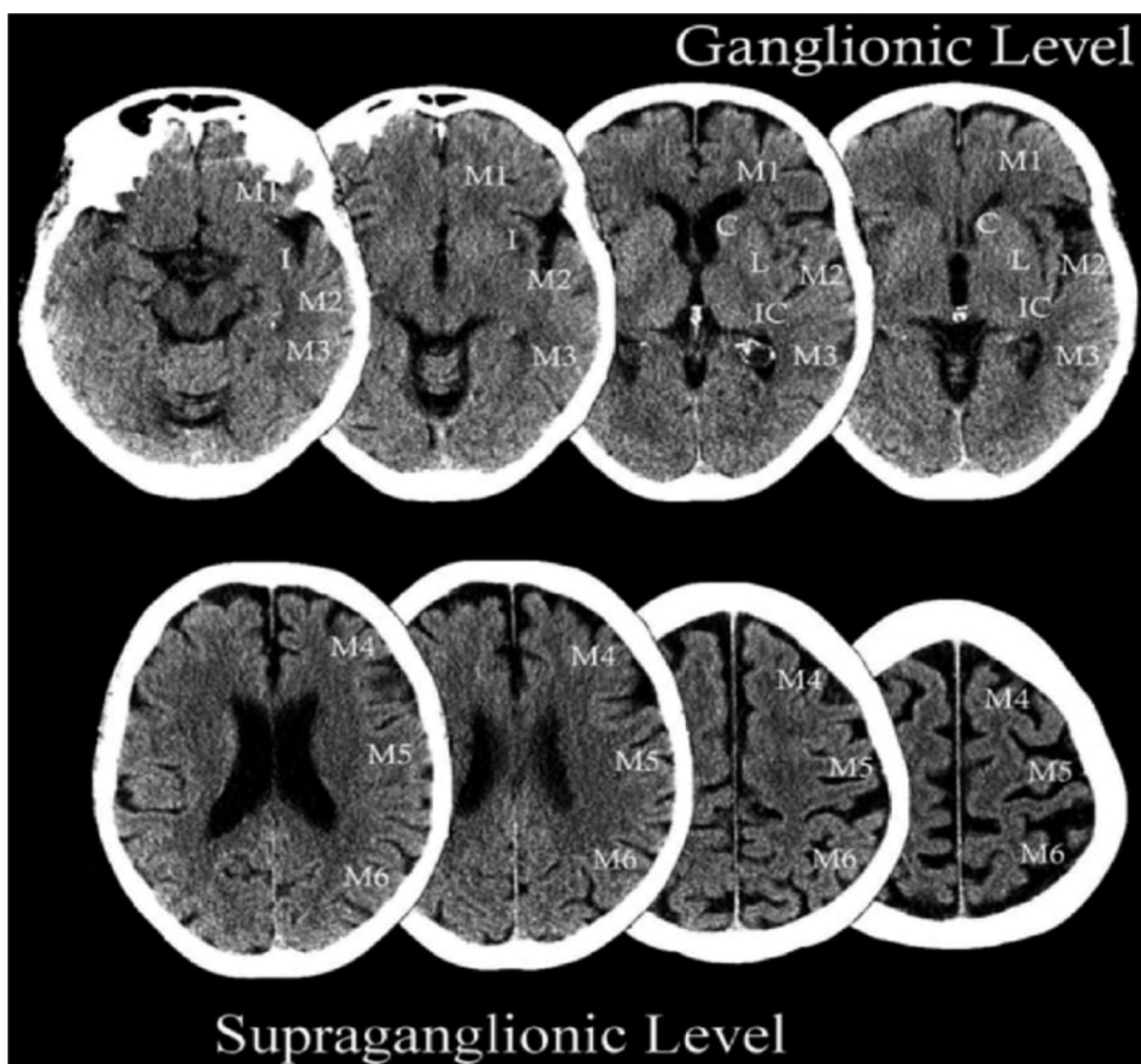
همانطور که پیش‌تر ذکر شد، امتیاز ASPECTS برای یک فرد سالم برابر با ۱۰ می‌باشد و به ازای هر یک از ۱۰ ناحیه‌ی تعیین‌شده‌ای که در اثر انسداد عروقی، آسیب دیده‌باشد، یک واحد از این امتیاز کسر می‌شود تا در حادترین وضعیت به صفر برسد. دقت داریم که هر کدام از نواحی، به صورت قرینه در دو نیم‌کره‌ی مغز وجود دارند و در هر سمتی از مغز که آسیب دیده باشند، مجموعاً تنها یک امتیاز از این ۱۰ امتیاز کم می‌شود.^۶

باید توجه داشت که هر یک از ۱۰ ناحیه‌ی ASPECTS یک حجم و ناحیه‌ی سه‌بعدی در مغز را شامل می‌شوند. تصویر ۲-۲ درواقع شمایی از این نواحی را تنها در برش‌های خاصی از مغز نمایش داده‌است. این در حالی است که هر کدام از این نواحی، در چندین برش از مغز گسترده شده‌اند. از جنبه‌ی نظری، درست آن است که در تشخیص امتیاز ASPECTS، تمام حجم مربوط به یک ناحیه در نظر گرفته‌شود. اما در عمل، معمولاً تنها چند برش از مغز به منظور تشخیص، مورد بررسی قرار می‌گیرند.

^۵ Dichotomized

^۶ البته در اکثر نمونه‌های سکته‌ی مغزی ایسکمیک، آسیب‌دیدگی تنها در یک نیم‌کره گسترش می‌یابد.

تصویر ۲-۳ گستردگی نواحی ده‌گانه‌ی ASPECTS در ۸ برش مغزی را نشان می‌دهد. اگرچه پژوهش‌هایی وجود دارند که امتیاز ASPECTS را از روی تصاویر سه‌بعدی محاسبه می‌کنند اما تشخیص انسانی ASPECTS عموماً بر روی همین تعداد محدود برش انجام می‌شود. در روش پیشنهادی این پژوهش نیز امتیازدهی ASPECTS بر روی چندین برش مهم از مغز انجام می‌شود. این در حالی است که پژوهش‌هایی که ASPECTS را به صورت سه‌بعدی محاسبه نمی‌کنند، غالباً تنها دو برش از مغز را بررسی می‌کنند که انطباق کمتری با روش انسانی دارد. جزئیات این تفاوت و روش پیشنهادی این پروژه در فصل‌های آتی خواهد آمد.



شکل ۲-۳: گستردگی نواحی ده‌گانه‌ی ASPECTS در برش‌های مغز [۱۲]

۴-۱-۲ تصاویر پزشکی

تصاویری که در این پروژه مورد استفاده قرار گرفته‌اند، به روش^۷ CT اخذ شده‌اند. اطلاعات مربوط به هر بیمار، در قالب صدها تصویر با فرمت DICOM استخراج شده و پس از یک‌سری پردازش، برای یادگیری ماشین مورد استفاده قرار گرفته‌اند. فرمت DICOM یک نوع فرمت مورد استفاده در تصاویر پزشکی است که علاوه بر پیکسل‌های تصویر، اطلاعاتی از قبیل نوع تصویربرداری، زمان اخذ تصویر، شناسه‌ی بیمار و... را در سرآیند^۸ خود نگه‌داری می‌کند. برای درک لزوم این پردازش اولیه‌ی این اطلاعات، در ادامه به برخی ویژگی‌های تصاویر پزشکی مورد استفاده به اختصار اشاره می‌شود.

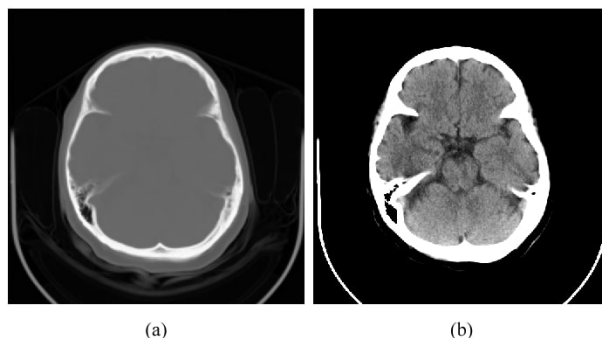
۱. **طیف رنگی:** چشم انسان تنها تعداد محدودی طیف رنگی خاکستری را می‌تواند تشخیص دهد. مثلاً طیف رنگی تصاویر سیاه‌سفید معمول، با اعدادی بین ۰ تا ۲۵۵ در هر پیکسل از تصویر مشخص می‌شود. اما تصاویر پزشکی معمولاً طیف بسیار وسیع‌تری از شدت رنگ خاکستری را شامل می‌شود. ارزش هر پیکسل در این تصاویر غالباً با واحد (HU) Hounsfield مشخص می‌شود و عموماً می‌تواند در بازه‌ی $1000 - HU$ تا $1000 HU$ باشد. اما این طیف برای انسان قابل تشخیص نیست و باید به مقدار کمتری محدود شود تا قابل مشاهده باشد. تصویر ۴-۲ لزوم این تغییر را نشان می‌دهد.

۲. **اضافات تصویر:** تصاویر مغزی CT، علاوه بر بافت اصلی مغز، شامل بخش‌های دیگری هم هستند که در یادگیری و تشخیص مورد نیاز نیستند و حتی می‌توانند برای مدل ماشین، گمراه‌کننده باشند. از جمله‌ی این موارد، مجموعه‌ی اطراف بافت اصلی مغز، قسمت‌هایی از دستگاه تصویربرداری، هوای اطراف سر بیمار و... می‌باشند. این اضافات باید از تصویر گرفته‌شوند و تنها بافت خالص مغز برای پردازش مورد استفاده قرار بگیرند.

۳. **زاویه، محل قرارگیری و فاصله‌ی سر:** در حین عکس‌برداری، زاویه‌ی سر بیمار ممکن است کاملاً مستقیم نباشد. همچنین ممکن است سر دقیقاً در مرکز تصویر قرار نداشته باشد یا نسبت به تصویر مغز سایر بیماران، دورتر و کوچک‌تر دیده‌شود. این مسئله باعث تفاوت ظاهری تصاویر مغز با هم می‌شود و لازم است یک‌دست‌سازی شود.

۴. **ناحیه‌ی تصویربرداری:** در تصاویری که از مراکز مختلف تصویربرداری (و گاه از یک مرکز) جمع‌آوری می‌شوند، تعداد برش‌های ثبت شده از مغز بیماران متفاوت است. به نحوی که در تصاویر مورد استفاده، برخی بیماران تا ۱۰ و برخی تا ۱۰۰ برش از مغز را در تصاویر خود شامل بودند. از طرفی همانطور که در بخش پیشین عنوان شد، تنها تعداد محدودی از برش‌های مغزی برای یادگیری

و تشخیص مورد توجه هستند. بنابراین لازم است این برش‌های خاص، از میان ۱۰۰ ها تصویر هر بیمار جدا شوند.



شکل ۲-۴: تفاوت توانایی تمیز جزئیات تصاویر مغز به چشم انسان، قبل و بعد از محدود کردن شدت رنگ [۱۳].
تصویر سمت چپ، تصویر اولیه را نشان می‌دهد.

در فصل روش پیشنهادی با جزئیات بیشتری خواهد آمد که هر یک از این چالش‌ها به چه صورت مدیریت شده‌اند و مراحل پردازش اولیه‌ی تصاویر نهایتاً به چه صورت تنظیم است.

۲-۲ مفاهیم فنی

در این قسمت به برخی از مهم‌ترین مفاهیم یادگیری ماشین مورد استفاده در این پروژه اشاره می‌شود. این روش‌ها به عنوان راهکاری برای مدیریت محدودیت‌های داده‌ای پروژه ارائه شده‌اند و چهارچوب کلی قسمت فنی را تشکیل می‌دهند. جزئیات مربوط به این روش‌ها در فصل روش پیشنهادی آورده شده‌است.

۲-۲-۱ داده‌افزایی

یکی از روش‌های جبران حجم کم داده‌های ورودی در یادگیری ماشین، داده‌افزایی^۹ است. در این روش، با ایجاد تغییرات جزئی بر روی تصاویر ورودی، با حفظ ویژگی‌های اصلی، تعداد تصاویر افزایش داده می‌شود. به عنوان مثال، با قرینه کردن تصویر مغزی که نیم‌کره‌ی راست آن درگیر است، می‌توان تصویر جدیدی ایجاد کرد که در آن، نیم‌کره‌ی چپ درگیر است. هر چه که مدل ماشین، تصاویر متنوع‌تری را به این ترتیب مشاهده کند، می‌تواند بهتر بیاموزد و بر روی طیف وسیع‌تری از تصاویر، تشخیص درستی بدهد. در پروژه‌ی حاضر نیز به منظور مدیریت تعداد محدود داده‌های ورودی، از روش‌های داده‌افزایی به خوبی بهره‌برده شده‌است. با این مقدمه، در فصل روش پیشنهادی، جزئیات تغییرات اعمال‌شده بر روی تصاویر خواهد آمد.

^۹ Data Augmentation

۲-۲-۲ یادگیری انتقالی

همانطور که مهارت و دقت متخصصان این حوزه، با کسب تجربه‌ی بیشتر، افزایش می‌یابد، روش‌های یادگیری ماشین بر روی تصاویر نیز مبتنی بر مشاهده‌ی تعداد زیادی نمونه‌ی ورودی هستند. البته به علت عدم هوشمندی انسانی در این مدل‌ها، نیازمندی داده‌ای به مراتب بیشتر هم هست. هرچه تعداد نمونه‌های فراگرفته‌شده توسط مدل بیشتر باشد، دقت و عملکرد تشخیصی آن نیز بیشتر خواهد بود. در مقابل، در صورتی که تعداد و تنوع داده‌ها اندک باشد، یادگیری ویژگی‌های کلیدی برای تشخیص، برای مدل دشوارتر خواهد بود و ممکن است تنها به حفظ کردن نمونه‌های مشاهده‌شده اکتفا کند.

در چنین مواردی، یکی از روش‌های مورد استفاده در حوزه‌ی یادگیری ماشین، یادگیری انتقالی^{۱۰} است. در یادگیری انتقالی، از یک مدل ماشین که توانایی‌های مشابهی با مدل مورد نیاز مسئله را دارد استفاده می‌شود. این مدل قبلاً بر روی تعداد زیادی تصویر آموزش دیده‌است و برخی مهارت‌های پایه‌ای چون تشخیص اشیاء و مرز آن‌ها در تصاویر را فراگرفته‌است. این مدل پیش‌آموزش دیده^{۱۱} سپس به عنوان هسته‌ی مدل جدید قرار می‌گیرد تا مدل جدید بتواند از توانایی‌های آن در استخراج ویژگی‌های^{۱۲} تصاویر استفاده کند و از اطلاعاتی که این مدل به دست می‌دهد، برای حل مسئله‌ی خود بهره ببرد.

تعداد قابل توجهی مدل پیش‌آموزش دیده در حوزه‌ی یادگیری تصاویر توسعه یافته‌است. این مدل‌ها تا صدها میلیون پارامتر یادگیری داشته و بر روی ده‌ها میلیون تصویر آموزش داده شده‌اند. در بخش روش پیشنهادی خواهد آمد که به کارگیری یادگیری انتقالی به کمک این مدل‌ها، چگونه به کاهش نیازمندی‌های داده‌ای، افزایش سرعت یادگیری و جامعیت روش پیشنهادی منجر شده‌است.

۳-۲-۲ اعتبارسنجی متقابل

در اعتبارسنجی مدل‌های یادگیری ماشین، توجه به این نکته ضروری است که ارزیابی باید از روی تصاویری انجام شود که تا کنون به مدل عرضه نشده‌اند. این تصاویر تحت عنوان داده‌های دیده‌نشده^{۱۳} شناخته می‌شوند. درواقع اگر مدل قبلاً برچسب یک تصویر را دیده‌باشد، می‌توانسته پارامترهای خود را به گونه‌ای تغییر دهد که این تصویر را به درستی امتیازدهی کند. اما آنچه در ارزیابی مدل، مدنظر است، توانایی مدل برای تشخیص درست بر روی تصاویر بیمارانی است که هرگز ندیده و فراگرفته‌است. تنها در این صورت است که می‌توان مدل را با اطمینان بالاتری در کاربرد واقعی به کار برد و انتظار داشت که همان عملکرد

^{۱۰} Transfer Learning

^{۱۱} Pre-trained

^{۱۲} Feature

^{۱۳} Unseen

ارزیابی شده را بر روی داده‌های جدید از خود نشان دهد.

با این مقدمه مشخص می‌شود که یکی دیگر از چالش‌ها در مواجهه با تعداد اندک مجموعه داده، در ارزیابی توانایی مدل مطرح می‌شود. چراکه بخشی از داده‌های موجود، باید به طور کامل جدا شوند و هرگز در فرایند آموزش دخالت نداشته باشند تا سپس بتوانند در ارزیابی مدل مورد استفاده قرار بگیرند. در این صورت، تعداد داده‌هایی که مدل می‌تواند بر روی آن‌ها آموزش ببیند، از پیش هم کمتر می‌شود. این مسئله توانایی مدل برای یادگیری و عملکردش بر روی تصاویر دیده‌نشده را به طرز قابل توجهی کاهش می‌دهد.

شاید یک پاسخ ساده به این مشکل، این باشد که تصاویر کمتری برای ارزیابی مدل جدا شود تا مدل بتواند بر روی تعداد بیشتری تصویر آموزش ببیند. اما این راه‌حل ممکن نیست. زیرا عملکرد مدل بر روی تنها تعداد اندکی تصویر، نمی‌تواند ملاک مناسبی برای ارزیابی آن باشد. این احتمال وجود دارد که مدل به صورت تصادفی، عملکرد بسیار خوبی از خود نشان بدهد. در این صورت نتایج حاصل از ارزیابی، گمراه‌کننده خواهد بود و ممکن است یک مدل نامناسب را به اشتباه وارد مرحله‌ی کاربردی کنند.

در این پژوهش، برای مقابله با این چالش از ایده‌ی روش اعتبارسنجی متقابل^{۱۴} استفاده شده‌است. در تناظر با نام این روش، در ادامه‌ی پایان‌نامه، روش مورد استفاده برای ارزیابی مدل، ارزیابی متقابل نامیده می‌شود. در این روش، داده‌های موجود به ۵ دسته تقسیم می‌شوند و مدل ۵ بار وارد مرحله‌ی آموزش و ارزیابی می‌شود. در هر مرحله، یک دسته به عنوان داده‌ی دیده‌نشده، کنار گذاشته می‌شود، مدل بر روی ۴ دسته آموزش می‌بیند و بر روی یک دسته ارزیابی می‌شود. به این ترتیب در طی ۵ مرحله، نهایتاً مدل بر روی تمام داده‌های موجود ارزیابی شده‌است. زمانی که عملکرد مدل بر روی تمام داده‌های ممکن ارزیابی شد و توانایی آن قابل قبول بود، مدل بر روی تمام داده‌های موجود آموزش داده می‌شود و برای کاربرد در محیط واقعی عرضه می‌شود.

به این ترتیب می‌توان گفت ارزیابی متقابل، مشکل کاهش حجم مجموعه داده به علت نیاز به ارزیابی مدل را حل می‌کند. علاوه بر این، ارزیابی متقابل، روش مطمئن‌تری را برای ارزیابی مدل ارائه می‌کند. چراکه عملکرد مدل، بر روی تمام داده‌های موجود سنجیده می‌شود و نه فقط بر روی تعدادی از تصاویر دست‌چین شده. در فصل آزمایش‌ها مشخص می‌شود که چگونه این روش در پروژه‌ی حاضر به کار گرفته شده و به چه عملکردی منجر شده‌است.

^{۱۴} Cross Validation

۲-۳ جمع‌بندی

آغاز به کار پروژه، با کسب دانش در رابطه با یک‌سری مفاهیم پزشکی همراه بوده‌است. در ادامه، این مفاهیم به حیطه‌ی اجرایی مدل شده‌اند. سپس با جستجوی راهکارهای فنی موجود، مسیری برای حل مسئله طراحی گشته‌است. آنچه در این فصل گذشت، شرح مختصری از دانش پزشکی و فنی مورد بحث در این پروژه می‌باشد. آشنایی با این مفاهیم، از دو جهت اهمیت دارد. اول آنکه باعث کسب یک دانش مشترک در میان مختصصان دو حوزه می‌شود و دوم آنکه به درک بهتر چالش‌های پروژه و روش‌های پیشنهادی آن کمک می‌کند. در فصل‌های بعد، علاوه بر نحوه‌ی مدل‌کردن مفاهیم پزشکی به حیطه‌ی اجرایی، روند دستیابی به راهکار پیشنهادی این پروژه نیز تشریح شده و مورد تحلیل قرار خواهد گرفت.

فصل ۳

کارهای پیشین

کارهای پیشین انجام شده در حوزه‌ی ASPECTS از چند نظر قابل بررسی و مقایسه هستند. اولین دیدگاه، روش مورد استفاده در این کارها است. دیدگاه دیگر نیز مجموعه داده و منابع مورد دسترسی این پژوهش‌ها می‌باشد. بررسی کارهای پیشین از این دو دیدگاه، این مزیت را دارد که جایگاه پروژه حاضر را بیشتر مشخص می‌کند و محدودیت‌ها و کارآمدی‌های آن را بهتر شرح می‌دهد. در این فصل، پژوهش‌های پیشین در این چهارچوب، مورد تحلیل قرار می‌گیرند.

۳-۱ روش‌های مورد استفاده

کارهای پیشین انجام شده از نظر روش مورد استفاده، در چند دسته‌ی کلی قابل بررسی هستند. در طی بررسی هر دسته، ابتدا روش کلی مورد استفاده در آن توضیح داده می‌شود. سپس به نمونه‌هایی از کارهای پیشین که در آن چهارچوب کار کرده‌اند اشاره می‌شود و نتایج به دست آمده توسط این کارها عنوان شده و مورد مقایسه قرار می‌گیرد.

۳-۱-۱ روش ناحیه‌بندی و طبقه‌بندی

در این روش، ده بخش مورد توجه ASPECTS در تصاویر مغزی، ناحیه‌بندی^۱ می‌شوند. به این ترتیب، مدل یادگیری ماشین به طور مستقیم از محل این نواحی در تصاویر آگاهی می‌یابد. سپس مدل آموزش داده می‌شود که هر ناحیه‌ای که می‌بیند، آیا آسیب دیده است یا خیر. یعنی یاد می‌گیرد که هر ناحیه را به

^۱ Segmentation

دو دسته‌ی آسیب‌دیده و سالم طبقه‌بندی^۲ کند. سپس با جمع امتیازات تمام ده ناحیه‌ی هر بیمار، امتیاز ASPECTS وی به دست می‌آید.

ناحیه‌بندی نواحی ASPECTS

ناحیه‌بندی ۱۰ بخش ASPECTS تصاویر به دو طریق مختلف انجام می‌شود. روش اول از یادگیری ماشین بهره می‌گیرد. در این روش، هر تصویر مغزی، برچسبی دارد که نشان می‌دهد کدام پیکسل‌های تصویر متعلق به هر ناحیه هستند. تعداد زیادی از تصاویر مغزی به همراه این برچسب‌ها به مدل ورودی داده تا ناحیه‌بندی را بیاموزد. به این ترتیب، مدل می‌تواند با دریافت یک تصویر مغزی جدید و بدون برچسب، مشخص کند که کدام پیکسل‌ها متعلق به هر ناحیه هستند.

روش دیگر ناحیه‌بندی، مبتنی بر یادگیری نیست و نیازی به تعداد زیادی تصویر به همراه برچسب ندارد. بلکه در این روش، یک یا چند تصویر مغزی استاندارد، به عنوان الگو^۳ برچسب زده می‌شوند. سپس به کمک روش‌های انطباق تصاویر^۴، تصویر الگو بر یک تصویر مغزی مورد نظر منطبق می‌شود تا نواحی مشخص شده روی آن، در تصویر جدید هم مشخص شوند.

از جمله روش‌های منطبق کردن تصویر الگو بر روی تصویر جدید، جابجایی، دوران، بزرگ‌نمایی، تغییر شکل جزئی و ... می‌باشد. تصویر الگو آن‌قدر دچار این دست تغییرات می‌شود تا معیار شباهتش با تصویر جدید به حد مطلوبی برسد. یک نمونه‌ی ساده از چنین معیاری می‌تواند مجموع اختلاف قدر مطلق دو تصویر باشد که باید کمینه شود. لازم به ذکر است که روش‌های ناحیه‌بندی به کمک انطباق تصاویر، عموماً توانایی کمتری نسبت به مدل‌های یادگیری ماشین دارند اما نسبت به آن روش‌ها نیازمندی‌های داده‌ای کمتری دارند.

استخراج ویژگی نواحی

پس از مشخص شدن محدوده‌ی هر ناحیه‌ی ASPECTS، لازم است ویژگی‌های اصلی هر ناحیه استخراج شود تا مدل بتواند از روی این ویژگی‌ها، آن ناحیه را دسته‌بندی کند. در کارهای پیشین، محاسبه‌ی چنین ویژگی‌هایی به دو طریق مختلف انجام شده است. دسته‌ی اول، استخراج ویژگی‌های هر ناحیه را به مدل یادگیری ماشین واگذار می‌کنند. یعنی تصاویر به مدل، ورودی داده می‌شوند و مدل طی چندین مرحله مشاهده‌ی نواحی به همراه برچسبشان، می‌آموزد که چه ویژگی‌هایی از تصاویر استخراج کند که بیش از همه

Classification^۲
Template^۳
Image Registration^۴

مفید باشند.

اما دسته‌ی دیگر برای استخراج ویژگی تصاویر، به جای یادگیری ماشین، روش‌های محاسباتی و پردازش تصویری را به کار می‌گیرند. درواقع یک سری ویژگی‌های آماری همچون میانگین و واریانس شدت رنگ پیکسل‌ها برای هر ناحیه محاسبه می‌شوند. پس از اینکه این ویژگی‌ها برای هر ناحیه استخراج شدند، در اختیار مدل یادگیری ماشین یا هوش مصنوعی قرار می‌گیرند تا در طبقه‌بندی نواحی، استفاده شوند.

نمونه‌ی کارهای پیشین

یکی از تازه‌ترین پژوهش‌ها در زمینه‌ی امتیازدهی خودکار ASPECTS، در همین دسته از روش‌ها قرار می‌گیرد [۱۴]. این پژوهش با ناحیه‌بندی نواحی ASPECTS و استخراج ویژگی‌های نواحی به کمک مدل یادگیری ماشین، توانسته به دقت‌های نسبتاً خوبی (تشخیص^۵ ۹۶.۶۳٪ و حساسیت^۶ ۶۲.۷۸٪ در امتیازدهی ده‌گانه و تشخیص ۷۶.۵۶٪ و حساسیت ۹۵.۴۲٪ در امتیازدهی دوبخشی) دست یابد. نمونه‌ی موفق و اخیر دیگری وجود دارد که نواحی را به کمک یادگیری ماشین ناحیه‌بندی و طبقه‌بندی می‌کند [۱۵]. این نمونه نیز نتایج بسیار خوبی (تشخیص ۹۲.۲٪ و حساسیت ۷۷.۲٪ در امتیازدهی ده‌گانه و تشخیص ۸۶.۶٪ و حساسیت ۹۵.۵٪ در امتیازدهی دوبخشی) گزارش کرده‌است. کار دیگری [۱۶] که ناحیه‌بندی را به کمک انطباق تصاویر انجام داده است، برای امتیازدهی ده‌گانه و دوبخشی به ترتیب دقت ۸۴٪ و ۹۰٪ را گزارش کرده‌است. همچنین یک نمونه از قدیمی‌ترین کارهای پیشین که دو روش برای محاسبه‌ی ASPECTS پیشنهاد داده، در روش ناحیه‌بندی و طبقه‌بندی خود، دقت ۷۰.۹٪ را اعلام کرده است [۱۷].

چند نمونه کار پیشین نیز در ادامه عنوان می‌شود که در استخراج ویژگی‌های نواحی، از روش‌های آماری استفاده کرده‌اند. یکی از موفق‌ترین نمونه‌ها در این دسته، پژوهشی نسبتاً قدیمی است که تشخیص ۹۱.۸٪، حساسیت ۶۶.۲٪ و دقت ۸۴.۹٪ در امتیازدهی ده‌گانه و تشخیص ۸۰٪، حساسیت ۹۷.۸٪ و دقت ۹۶٪ را در امتیازدهی دوبخشی به دست آورده است [۱۱]. نمونه‌های دیگری نیز از سال‌های اخیر وجود دارند [۱۸، ۱۹] که به علت مقایسه‌پذیر نبودن و یا ابهام در روش اعتبارسنجی، از ذکر نتایج آن‌ها صرف نظر می‌شود.

۳-۱-۲ روش ناحیه‌بندی و هم‌پوشانی

در این روش، دو نوع ناحیه‌بندی انجام می‌شود. نوع اول، نواحی ASPECTS و نوع دوم، بخش‌های آسیب‌دیده‌ی مغزی در اثر انسداد عروق را مشخص می‌کند. سپس هم‌پوشانی بخش‌های آسیب‌دیده با هر

Specificity^۵
Sensitivity^۶

ناحیه محاسبه می‌شود. در صورتی که نسبت مساحت آسیب‌دیده‌ی یک ناحیه، از یک حد آستانه فراتر برود، آن ناحیه به عنوان آسیب‌دیده گزارش می‌شود و در غیر این صورت، سالم شناخته می‌شود. درواقع در این روش‌ها، مدل‌های یادگیری ماشین، وظیفه‌ی اصلی ناحیه‌بندی را بر عهده دارند و نه طبقه‌بندی.

واضح است که ناحیه‌بندی نواحی آسیب‌دیده، بر خلاف ناحیه‌بندی نواحی ده‌گانه‌ی ASPECTS، به روش انطباق تصاویر ممکن نیست. زیرا الگوی ثابت و مشخصی برای نواحی آسیب‌دیده وجود ندارد. به همین دلیل این روش‌ها برای آموزش مدل ماشین، عموماً نیازمند تعداد زیادی تصویر مغزی به همراه برچسب پیکسل‌های آسیب‌دیده هستند. این نوع از برچسب‌ها، وقت و انرژی زیادی از نیروهای انسانی می‌گیرند و تهیه‌ی آن‌ها دشوارتر است.

نمونه‌ی کارهای پیشین

در میان کارهای پیشین، سه پژوهش با روش ناحیه‌بندی و هم‌پوشانی یافته‌شد. یکی از بهترین نتایج گزارش‌داده‌شده مربوط به پژوهشی در سال ۲۰۲۱ است [۲۰] که تشخیص ۹۷٪، حساسیت ۸۰٪ و دقت ۹۶٪ در امتیازدهی ده‌گانه و تشخیص ۹۲٪، حساسیت ۹۸٪ و دقت ۹۷٪ را در امتیازدهی دوبخشی گزارش کرده‌است. البته در این پژوهش به وضوح اشاره نشده‌است که این نتایج مربوط به داده‌های آموزشی هستند و یا آزمایشی. در پژوهش دیگری تشخیص ۱۷٪ تا ۸۳٪ و حساسیت ۶۹٪ تا ۱۰۰٪ در امتیازدهی نواحی ده‌گانه و تشخیص ۴۸٪ تا ۹۳٪ و حساسیت ۷۴٪ تا ۹۹٪ در یک امتیازدهی سه‌بخشی ارائه شده‌است [۱۹]. آخرین مورد مطالعه‌شده نیز نتایج را در قالب بهبودی که در توانایی تشخیصی متخصصان ایجاد می‌کند ذکر کرده‌است و از آن عبور می‌شود [۴].

۳-۱-۳ روش کل‌نگری و طبقه‌بندی

تعداد بسیار محدودی از پژوهش‌ها در این دسته قرار می‌گیرند که پژوهش حاضر نیز یکی از آن‌ها است. در این روش، تنها با در دست داشتن امتیاز ASPECTS کلی بیمار، امتیاز ASPECTS تصاویر فراگرفته می‌شود. درواقع در این روش، مدل تعداد زیادی تصویر مغزی به همراه برچسب امتیاز نهایی ASPECTS آن‌ها را مشاهده می‌کند و نمونه‌های جدید تصاویر مغزی را در یکی از دسته‌های امتیاز ASPECTS طبقه‌بندی می‌کند.

در فصل بعد خواهد آمد که این دسته از روش‌ها کمترین نیازمندی داده‌ای را دارند. به همین نسبت، دقت این روش‌ها نسبت به روش‌های قبلی، عموماً پایین‌تر است. با این حال، پژوهشی [۲۱] وجود دارد که در این دسته از روش‌ها، بالاترین دقت در میان تمام کارهای پیشین را گزارش کرده‌است. طبق بررسی‌های

انجام شده، نتایج گزارش شده معتبر نیستند چرا که در ارزیابی مدل، جدایی میان داده‌های آموزشی و آزمایشی رعایت نشده است. به عبارتی ارزیابی شامل داده‌هایی می‌شود که مدل، قبلاً پاسخ آن‌ها را مشاهده کرده است و طبعاً پیش‌بینی درست‌تری روی آن خواهد داشت. لذا از ذکر و مقایسه‌ی نتایج این پژوهش صرف نظر می‌شود.

به این ترتیب تنها یک نمونه کار دیگر با این روش در ادبیات موضوع باقی می‌ماند [۱۷]. این مدل، امتیاز ASPECTS را برای دو برش اصلی مغز می‌آموزد. این پژوهش، خطای متوسط ۰.۱۱۱۶ و خطای واریانس ۲.۵۰۸۰ را گزارش کرده است.

۳-۱-۴ سایر روش‌ها

طبیعتاً روش‌های محاسبه‌ی خودکار ASPECTS محدود به روش‌های پیشنهادی فوق نیست و هر پژوهشی را نمی‌توان لزوماً در یکی از این دسته‌ها قرار داد. در میان کارهای پیشین نیز چنین موردی وجود دارد [۲۲]. این پژوهش که جزء کارهای تازه‌تر است، روش جالبی را به کار برده است که ذکر آن خالی از لطف نیست.

در این پژوهش، نواحی ASPECTS ناحیه‌بندی می‌شوند. سپس از دو ترفند پیش‌آموزش‌دادن^۷ و تنظیم دقیق^۸ برای آموزش مدل در دو مرحله استفاده می‌شود. در گام اول، مدل با دریافت تعداد زیادی برش از مغز به همراه برچسب ASPECTS همان برش، پیش‌آموزش می‌بیند. سپس این مدل، با دریافت تعداد زیادی تصویر با برچسب‌هایی در سطح هر ناحیه، تنظیم دقیق می‌شود.

این پژوهش نهایتاً تشخیص ۸۱.۶٪، حساسیت ۶۵.۲٪ و دقت ۷۹.۷٪ در امتیازدهی ده‌گانه و تشخیص ۹۰.۷٪، حساسیت ۷۲.۲٪ و دقت ۸۸.۹٪ را در امتیازدهی دوبخشی گزارش می‌کند که می‌تواند در برخی کاربردها مناسب باشد.

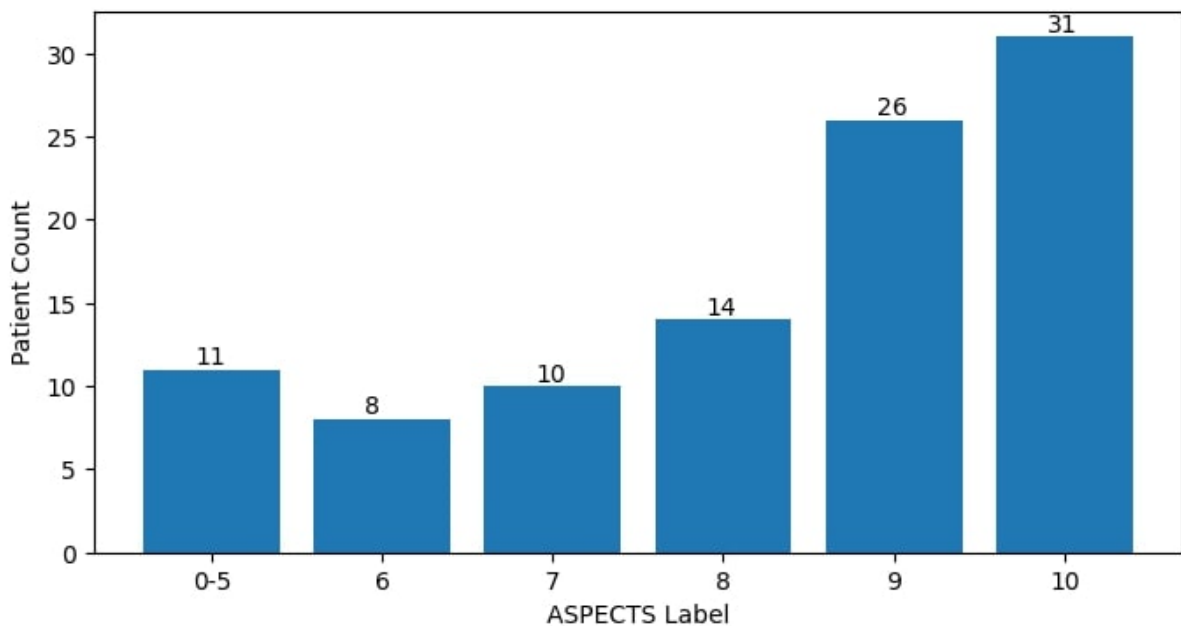
۳-۲ مجموعه داده و برچسب مورد استفاده

یک رویکرد دیگر برای مطالعه‌ی کارهای پیشین، بررسی آن‌ها از نظر مجموعه داده‌ی در دسترس و برچسب‌ها آن است. در این قسمت، کارهای پیشین از این منظر با پژوهش حاضر مقایسه می‌شوند و محدودیت‌های داده‌ای این پروژه از این دیدگاه مورد تحلیل قرار می‌گیرد.

Pre-training^۷
Tuning Fine^۸

۳-۲-۱ حجم مجموعه داده

در این پژوهش، پس از مرتب‌سازی، خالص‌سازی و حذف داده‌های نامناسب (مانند تصاویر با نویز بسیار زیاد)، مجموعاً تصاویر ۱۰۰ بیمار آماده‌ی عرضه به مدل بودند. هر یک از این بیماران، با توجه به وضعیت خود، یک امتیاز ASPECTS داشتند. شکل ۳-۱، توزیع امتیاز ASPECTS این بیماران را نشان می‌دهد. این تصویر نشان می‌دهد که در امتیازدهی دوبخشی ASPECTS در دسته‌ی امتیاز زیر ۶، در مجموع تنها ۱۱ بیمار وجود داشتند. این در حالی است که امتیاز ۱۰، به تنهایی شامل ۳۱ بیمار می‌شد. این موضوع علاوه بر حجم کوچک مجموعه داده، نشان از نامتوازن^۹ بودن آن دارد.

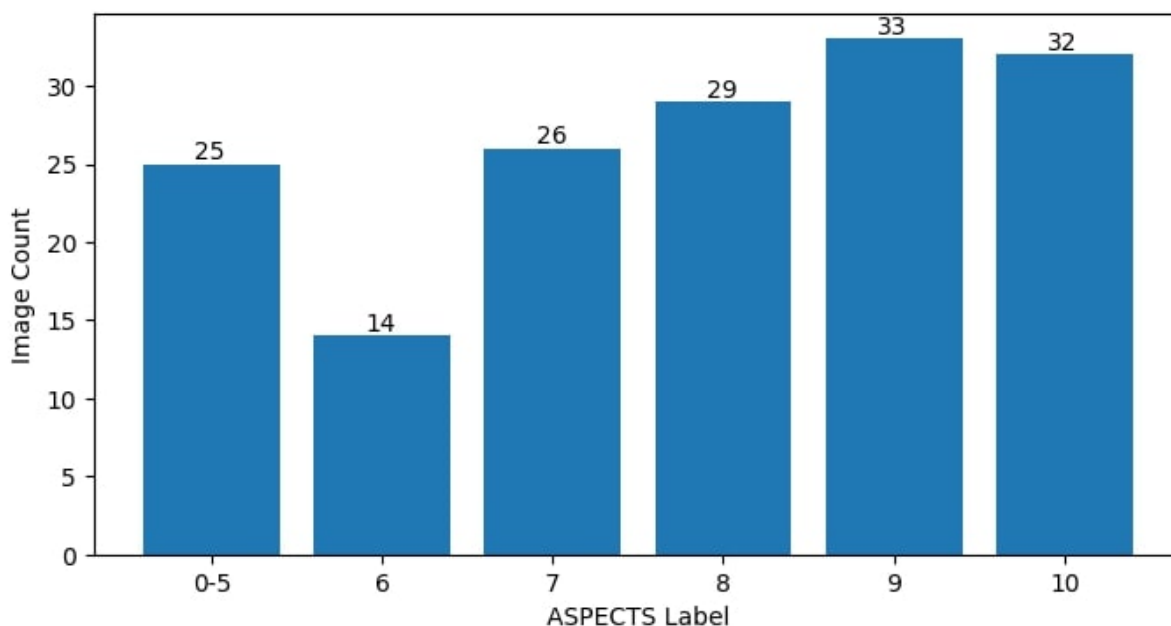


شکل ۳-۱: توزیع امتیاز ASPECTS بیماران

در مجموعه داده‌ی مورد استفاده در این پژوهش، برخی بیماران، چند نوبت تصویربرداری داشتند. بنابراین به ازای برخی بیماران، بیش از یک تصویر وجود داشت. تصاویر مختلف متعلق به یک بیمار، عموماً تفاوت‌های ظاهری اندکی داشتند اما شکل و شمایل ناحیه‌ی آسیب‌دیده در آن‌ها مشابه بود. به این ترتیب، در مجموع، ۱۵۹ تصویر از ۱۰۰ بیمار در مجموعه داده حاضر بودند. شکل ۳-۲ تعداد تصاویر موجود به ازای برچسب ASPECTS را نشان می‌دهد.

این تعداد تصویر به خصوص در برخی امتیازهای ASPECTS، تعداد کمی محسوب می‌شود. این موضوع از مقایسه‌ی حجم مجموعه داده‌ی در اختیار این پژوهش با مجموعه داده‌ی سایر پژوهش‌ها مشهود است. برخی از این پژوهش‌ها [۱۵، ۲۳، ۴] بر روی بیش از ۱۰۰۰ بیمار، برخی [۱۴، ۲۲، ۱۷، ۱۱، ۱۶] بیش از ۲۵۰ بیمار و یک مورد [۲۰] بر روی ۱۵۱ بیمار انجام شده‌اند که البته مورد اخیر، مدل

^۹ Imbalanced



شکل ۲-۳: تعداد تصویر (نه کاملاً نو) به ازای هر امتیاز ASPECTS

پیش‌آموزش دیده‌ای بر روی مجموعه داده‌ی دیگری را مورد استفاده قرار داده‌است. لازم به ذکر است، در این بین، دو مورد [۲۱، ۱۹] نیز با حجم مجموعه داده‌ی ۷۷ و ۹۰ وجود دارد که اولی به علت ارزیابی نادرست مدل و دومی به علت دقت‌های غیرقابل قبول، در ادامه مورد بحث قرار نمی‌گیرند.

۲-۲-۳ نوع برچسب مجموعه داده

نکته‌ی دیگری که لازم است اشاره شود، نوع برچسب^{۱۰} پژوهش‌های حوزه‌ی ASPECTS است. مقصود از برچسب، اطلاعات پزشکی صحیحی است که به هر تصویر نسبت داده می‌شود تا مدل از آن‌ها بیاموزد. بررسی کارهای پیشین از نظر نوع برچسبی که در اختیار داشتند، محدودیت حجم مجموعه داده و قدرت یادگیری مدل یادگیری ماشین در این پروژه را بیش از پیش مشخص می‌کند. این کارها از نظر نوع برچسب در سه دسته‌ی کلی قابل بررسی هستند.

برچسب سطح پیکسل

در این دسته، هر پیکسلی یک برچسب صفر و یکی دارد که مشخص می‌کند آن پیکسل جزء ناحیه‌ی آسیب دیده هست یا خیر. درواقع نواحی آسیب دیده، بر روی هر تصویر علامت‌گذاری می‌شوند. برخی کارهای پیشین چنین برچسب‌هایی در اختیار داشتند [۱۵، ۲۳، ۱۶، ۴]. این نوع برچسب این امکان را به دست می‌دهد

^{۱۰}Label

که مدل تنها آسیب دیدگی یا عدم آسیب دیدگی هر بخشی از بافت مغز را بیاموزد و سپس از روی میزان هم پوشانی این نواحی با نواحی ده گانه، امتیاز ASPECT تعیین شود. در واقع این روش، مستقل از نوع ناحیه‌ی مورد بررسی است و نیازی به مشاهده‌ی حالت‌های مختلف بروز سکته در نواحی مختلف ندارد.

برچسب سطح ناحیه

در این دسته، هر یک از نواحی ده گانه‌ی ASPECTS، یک برچسب صفر و یکی دارد که نشان می‌دهد آن ناحیه آسیب دیده است یا خیر. با فرض این که ناحیه‌ی سالم با برچسب یک مشخص شود، جمع امتیازات نواحی، امتیاز ASPECTS نهایی را به دست خواهد داد. برخی کارهای پیشین چنین برچسب‌هایی در اختیار داشتند [۱۴، ۱۷، ۱۱]. به این موارد، دو پژوهش دیگر را نیز می‌توان افزود. یک مورد که بنظر می‌رسد از مدل پیش آموزش دیده‌ای با برچسب سطح پیکسل استفاده کرده [۲۰] و یک مورد که برچسب سطح برش نیز دارد [۲۲]. در این نوع برچسب، این امکان وجود دارد که مدل جداگانه‌ای برای یادگیری آسیب دیدگی یا سلامت هر ناحیه آموزش داده شود. بنابراین در این حالت نیز یادگیری مدل مستقل از ارتباط نواحی با یکدیگر است و ارتباط مستقیمی با تعداد نمونه‌های موجود از هر ناحیه برای آموزش خواهد داشت.

برچسب سطح مغز

در این دسته، به ازای کل تصویر هر مغز، یک امتیاز از صفر تا ۱۰ وجود دارد که نشان دهنده‌ی ASPECTS آن بیمار است. این نوع برچسب، حداقل برچسب ممکن برای یادگیری تحت نظارت^{۱۱} ASPECTS می‌باشد و تا آن جا که در جستجو برای کارهای پیشین به دست آمد، این پژوهش، تنها مورد معتبری است که تنها برچسب‌های سطح نواحی را در دست داشته است.^{۱۲}

در این دسته برخلاف دسته‌ی قبل، برچسب مستقلی برای نواحی وجود ندارد. بنابراین یادگیری مدل، وابسته به مشاهده‌ی تعداد زیادی از حالت‌های ممکن از آسیب دیدگی نواحی می‌باشد. این در حالی است که در قسمت قبل مشاهده شد که تعداد کل بیماران با هر یک از امتیازهای صفر تا ۵ تنها ۱۱ مورد است. این تعداد تنها انواع محدودی از آسیب دیدگی‌های ممکن را می‌تواند پوشش دهد و شامل تمام حالات نمی‌شود. به عنوان مثال، امتیاز ۱ که نشان دهنده‌ی سلامت یک ناحیه است، می‌تواند در هر کدام از نواحی ده گانه رخ بدهد که به تنهایی ۱۰ حالت مختلف را در یک نیم کره ایجاد می‌کند. به این ترتیب نتیجه می‌شود که پروژه‌ی حاضر علی‌رغم نیازمندی‌های بیشتر برای یادگیری، حجم مجموعه داده‌ی کمتری نسبت به سایر

^{۱۱} Learning Supervised

^{۱۲} پژوهشی با این سطح برچسب با ۷۷ بیمار وجود دارد [۲۱] اما از آن جا که اعتبارسنجی مدل در این پژوهش به روش نادرستی انجام شده است (عدم استفاده از داده‌های کاملاً دیده نشده در فاز آزمایش)، از بررسی آن صرف نظر می‌کنیم.

کارهای پیشین دارد. با این وجود، این پروژه به نتایج مطلوب و قابل مقایسه‌ای با این پژوهش‌ها دست یافته‌است.

۳-۳ کارهای پیشین در یک نگاه

کارهای پیشین مورد ارجاع در این پایان‌نامه، مطالعات و پژوهش‌هایی هستند که در زمینه‌ی تشخیص خودکار ASPECTS به کمک روش‌های هوش مصنوعی، یادگیری ماشین و پردازش تصویر فعالیت داشته‌اند. فهرست کامل این پژوهش‌ها، از جمله مواردی که ارزیابی معتبری نداشته‌اند و تنها به روش پیشنهادی آن‌ها استناد می‌شود، در جدول ۱-۳ آمده‌است.

جدول ۱-۳: فهرست کارهای پیشین

مرجع	سال	داده‌ها	برچسب	روش	آستانه
[۱۴]	۲۰۲۳	$300 <$	ناحیه	ناحیه‌بندی و طبقه‌بندی	$5 \geq$
[۲۳]	۲۰۲۲	۱۵۰۰	پیکسل	ناحیه‌بندی و هم‌پوشانی	–
[۱۵]	۲۰۲۲	$1000 <$	پیکسل	ناحیه‌بندی و طبقه‌بندی	$7 \geq$
[۲۲]	۲۰۲۰	$250 <$	ناحیه	سایر	$6 \geq$
[۲۱]	۲۰۲۲	۷۷	مغز	طبقه‌بندی	–
[۴]	۲۰۲۲	$1300 <$	پیکسل	ناحیه‌بندی و هم‌پوشانی	–
[۲۰]	۲۰۲۱	۱۵۱	ناحیه	ناحیه‌بندی و هم‌پوشانی	$6 \geq$
[۱۸]	۲۰۲۱	–	–	ناحیه‌بندی و طبقه‌بندی	–
[۱۹]	۲۰۲۱	۹۰	–	ناحیه‌بندی و طبقه‌بندی	$8 \geq$
[۱۶]	۲۰۲۱	$250 <$	پیکسل	ناحیه‌بندی و طبقه‌بندی	$5 \geq$
[۱۱]	۲۰۱۹	$250 <$	پیکسل	ناحیه‌بندی و طبقه‌بندی	$5 \geq$
[۱۷]	۲۰۱۸	$250 <$	ناحیه	ناحیه‌بندی و طبقه‌بندی	–

۳-۴ جمع‌بندی

مطالعه‌ی کارهای پیشین در هر پژوهشی از دو جهت حائز اهمیت می‌باشد. اول آنکه دیدگاه پژوهش‌گر را نسبت به چالش‌ها و روش‌های موجود وسیع‌تر می‌کنند و امکان طراحی یک مسیر نوآورانه، بر پایه‌ی این روش‌ها، بدون تحمل چالش‌های شناخته‌شده را برای وی فراهم می‌کنند. دوم از آن جهت که معیاری برای ارزیابی روش پیشنهادی و درک جایگاه آن را به دست می‌دهند. در فصلی که گذشت، فهرست جامعی از کارهای پیشین ارائه شد و این پژوهش‌ها از نظر امکانات در دسترس و روش‌های مورد استفاده، طبقه‌بندی شده و مورد مطالعه قرار گرفتند. در فصل‌های بعدی، روش پیشنهادی این پروژه ارائه خواهد شد و نتایج آن نیز با کارهای پیشین مقایسه خواهد گشت.

فصل ۴

روش پیشنهادی

فصل حاضر به شرح مراحل عملیاتی پروژه اختصاص دارد. در این بخش ابتدا نحوه‌ی آماده‌سازی مجموعه داده شرح داده می‌شود. سپس بررسی می‌شود که این تصاویر دستخوش چه تغییراتی شده و چگونه به مدل ورودی می‌شوند. در انتها، طراحی انجام شده برای مدل یادگیری ماشین نیز توصیف خواهد شد.

۴-۱ آماده‌سازی مجموعه داده

داده‌هایی که از مراکز پزشکی دریافت می‌شوند، مسیر طولانی‌ای را طی می‌کنند تا برای آموزش مدل قابل استفاده باشند. اولین مرحله‌ی این آماده‌سازی شامل نگاشت اطلاعات بیماران به تصاویر می‌باشد. این نگاشت از طریق شماره‌ی شناسه‌ی بیمار صورت می‌گیرد و در نتیجه‌ی آن، اطلاعات پزشکی در دسترس از بیماران به همراه مسیر ذخیره‌سازی تصاویر وی به صورت ساخت‌یافته‌ای جمع‌آوری می‌شوند. در پروژه‌ی حاضر نیز تصاویر CT مغزی از یک مرکز درمانی جمع‌آوری شده و مورد پژوهش قرار گرفته‌اند.

مرحله‌ی بعد، جداسازی تصاویر مورد توجه پژوهش است. تصویربرداری‌های انجام شده از بیماران، معمولاً تنها شامل مغز نمی‌شوند و تصاویری از ریه و ... را نیز در بر دارند. به علاوه در هر مجموعه، تعداد تصاویر مغزی نیز با مجموعه‌های دیگر می‌تواند تفاوت داشته باشد. به گونه‌ای که یک بیمار ۱۵ و بیمار دیگر، ۵۰ برش از مغز را در مجموعه‌ی خود داشته باشد. این در حالی است که در امتیازدهی ASPECTS تنها چند برش خاص از مغز مورد استفاده قرار می‌گیرد. به این ترتیب، یکی از گام‌های ضروری برای آماده‌سازی داده‌ها، جداسازی این برش‌ها از میان تمام تصاویر دریافت شده از بیماران و تهیه‌ی یک نگاشت مدون از هر بیمار به برش‌های استخراج شده از تصاویر وی می‌باشد.

با انجام دو اقدام فوق که بیشتر مربوط به مرتب‌سازی و خالص‌سازی اطلاعات بودند، نوبت به مرحله‌ی پیش‌پردازش^۱ تصاویر می‌رسد. پیش‌پردازش یکی از مهم‌ترین و موثرترین گام‌های یادگیری ماشین محسوب می‌شود که ارتباط مستقیمی با عملکرد و توانایی یادگیری مدل دارد. درواقع پیش‌پردازش شامل تغییراتی در داده‌های ورودی است که باعث می‌شود تا جای ممکن، اطلاعات نامفید از داده‌ها حذف شوند و اطلاعات کلیدی نیز در قالب مناسبی به مدل عرضه شوند. به این ترتیب فرایند یادگیری برای مدل ساده‌تر خواهد بود.

نکته‌ی دیگری که باعث اهمیت بیشتر پیش‌پردازش می‌شود، قابلیت استفاده‌ی مجدد آن در پژوهش‌های دیگر است. درواقع گام‌های پیش‌پردازش تصاویر پزشکی در بسیاری از پژوهش‌ها با هم اشتراکات زیادی دارد. روش‌های ارائه شده در این پروژه نیز می‌توانند در پیش‌پردازش تصاویر CT در سایر تحقیقات راهگشا باشند. در ادامه، فرایند پیش‌پردازش مورد استفاده در این پژوهش، به صورت گام به گام بر روی یک تصویر مغزی انتخابی شرح داده می‌شوند.^۲

۴-۱-۱ افزایش وضوح

همانطور که در بخش مفاهیم اولیه ذکر شد، مقدار عددی هر پیکسل در تصاویر پزشکی، بازه‌ی بزرگی را شامل می‌شود. این در حالی است که چشم انسان تنها تعداد محدودی رنگ خاکستری را می‌تواند از هم تمیز دهد. به همین دلیل، در صورت مشاهده‌ی یک تصویر CT خام، جزئیات بافت مغز و حتی ناحیه‌ی آسیب‌دیده، قابل مشاهده نخواهد بود. با این توضیح، تصویر انتخابی برای شرح مراحل پیش‌پردازش، در ابتدا مانند شکل ۴-۱ است.^۳ بنابراین، در اولین گام، لازم است وضوح تصاویر افزایش داده‌شود.

میان مقدار هر پیکسل با بافت یا شیئی که نمایش می‌دهد ارتباط وجود دارد. به نحوی که آب، مقدار HU ، هوا مقدار $HU - ۱۰۰۰$ ، بافت استخوانی مقدار $HU ۱۰۰۰$ و بافت مغزی در حدود $HU ۵۰ - ۲۰$ می‌باشد [۲۵]. به این ترتیب، با محدود کردن مقدار پیکسل‌ها به بازه‌ی مثل $HU ۱۰۰ - ۰$ ، اطلاعات مربوط به بافت مغز از بین نمی‌رود. اما این بار به علت کاهش بازه‌ی رنگ خاکستری، اجزای تصویر از هم بهتر تمایز می‌یابند. در اصطلاح تصاویر پزشکی، به بازه‌ای که مقادیر به آن محدود می‌شوند پهنای پنجره^۴ و به مرکز این بازه سطح پنجره^۵ گفته می‌شود. با اعمال پهنای پنجره‌ی ۱۰۰ و سطح پنجره‌ی ۵۰، تصویر اولیه مانند شکل ۴-۲ می‌شود.

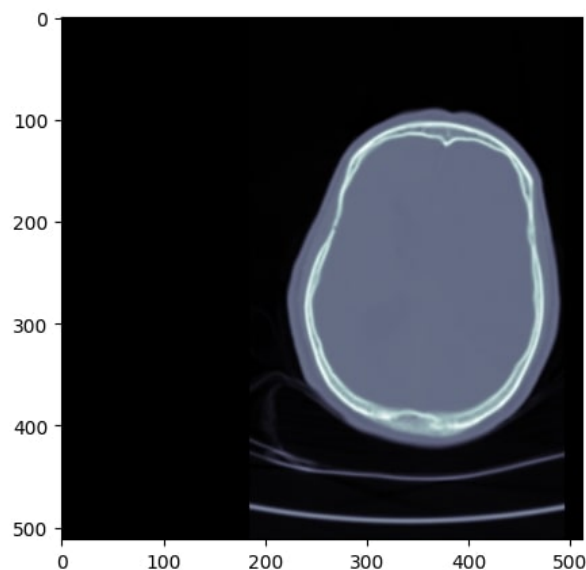
^۱Pre-processing

^۲در طراحی مراحل پیش‌پردازش این پروژه، مرجع [۲۴] به‌کار آمده‌است.

^۳ناحیه‌ی مغز برای مقاصد نمایشی، کمی جابه‌جا شده‌است.

^۴Window Width

^۵Window Level

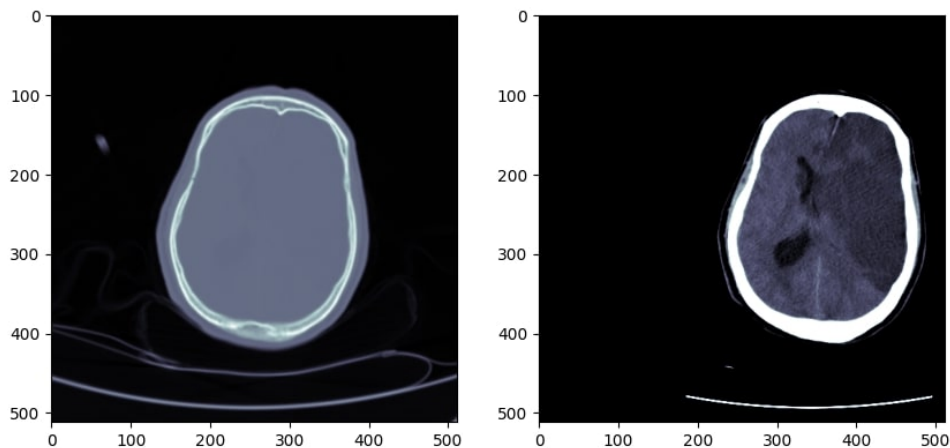


شکل ۴-۱: تصویر انتخابی برای نمایش مراحل پیش‌پردازش، در حالت اولیه.

روش معمول برای افزایش وضوح تصاویر، به همین صورت با تنظیم پهنا و سطح پنجره می‌باشد. اما باید توجه داشت که میان تصاویر بیماران مختلف، تفاوت‌های جزئی وجود دارد. برخی تصاویر ممکن است به طور کلی قدری تیره‌تر و یا روشن‌تر باشند. در نتیجه اعمال یک سیاست واحد برای پهنا و سطح پنجره ممکن است برای تمام تصاویر، بهینه نباشد و وضوح لازم را برای تصویر فراهم نکند. در این پژوهش، از یک روش پویا برای تنظیم وضوح تصاویر استفاده شده است که شرح آن در ادامه می‌آید.

در این روش، ابتدا مقدار پیکسل‌های مربوط به بافت مغزی هر بیمار به صورت یک مجموعه، استخراج می‌شوند. سپس چند درصد پایین و چند درصد بالای این مقادیر محاسبه می‌شوند. در نتیجه، یک چندک^۶ پایین و یک چندک بالا به دست می‌آید. در نهایت، مقادیر پیکسل‌های کل تصویر، محدود به بازه‌ی میان این دو چندک می‌شوند. به این ترتیب، در هر تصویری، با توجه به اطلاعات آماری همان تصویر، مقدار پیکسل‌ها محدود به بازه‌ای از رنگ خاکستری می‌شود که اطلاعات بیشتری در خود دارد. اعمال این روش پویا، نیازمند استخراج بافت مغزی است که در مراحل انتهایی پیش‌پردازش به دست می‌آید. به همین جهت، از آوردن تصویر آن در این بخش صرف نظر می‌شود اما تصویر نهایی فرایند پیش‌پردازش، نمایانگر افزایش بیشتر وضوح تصاویر نسبت به روش‌های ایستا (مانند تصویر ۴-۲) می‌باشد.

^۶Quantile



شکل ۴-۲: تصویر انتخابی برای نمایش مراحل پیش پردازش، پیش و پس از تنظیم پهنای پنجره به 100° و سطح پنجره به 50° . سمت راست، تصویر پس از تنظیم پنجره را نمایش می دهد.

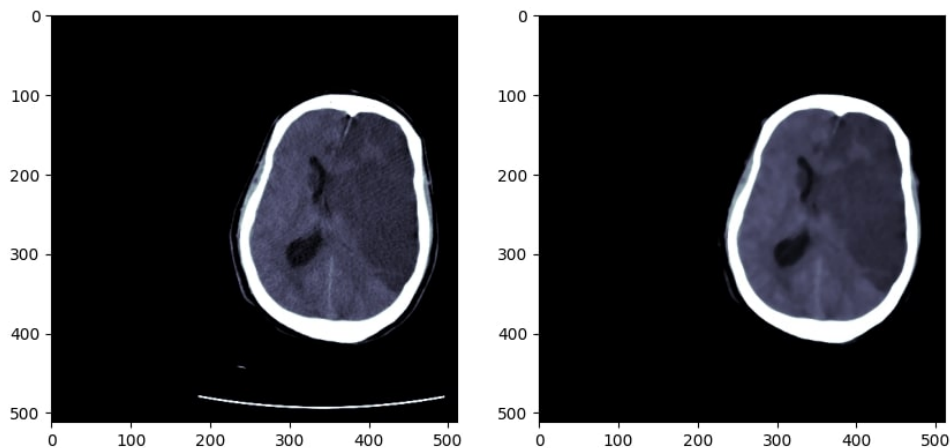
۴-۱-۲ کاهش نویز

تصاویر اخذ شده از بیماران معمولاً دارای نویز هستند. گاهی این نویز بسیار شدید است و گاهی جزئی بوده و تشخیص آن در نگاه کلی مشکل است. در این پژوهش، تصاویری که دارای نویز شدید بودند، از مجموعه داده حذف شده اند. اما در مورد سایر تصاویر، همچنان نویز اندکی باقی می ماند. در فرایند پیش پردازش پیشنهادی، این نویز با دو مرتبه اعمال فیلتر میانه^۷ در تصاویر، کاهش می یابد.

فیلتر میانه به این صورت عمل می کند که مقدار هر پیکسل را با میانه ی پیکسل های مجاورش در یک همسایگی مشخص جایگزین می کند. به این ترتیب، اگر تعداد اندکی پیکسل در آن همسایگی، مقادیر پرتی داشته باشند، با مقادیری در محدوده ی مناسب، ترمیم می شوند. شدت کاهش نویز به اندازه ی همسایگی مورد بررسی در فیلتر میانه بستگی دارد. هر چه این پنجره بزرگ تر باشد، تصویر خروجی محوتر و در هم تنیده تر می شود. در مقابل، هر چه فیلتر کوچک تر باشد، توانایی کاهش نویزش هم به جمعیت کوچک تری محدود می شود و برای رفع نویزهای سنگین مناسب نخواهد بود.

همانطور که پیش تر ذکر شد، در پیش پردازش پیشنهادی در این پژوهش، فیلتر میانه دو مرتبه اعمال می شود. یک مرتبه با پنجره ای با اندازه ی 7×7 پیکسل و یک مرتبه با اندازه ی 3×3 پیکسل. این مقادیر با آزمون و خطا بر روی تصاویر به دست آمده اند و علاوه بر کاهش نویز، باعث محو شدن و حذف جزئیات غیر مهم در عکس می شوند. این محوسازی در افزایش وضوح تصویر نیز تاثیر به سزایی دارد. چرا که در این صورت، مقادیر دورافتاده در میان پیکسل ها، حد بالا و پایین پنجره ی وضوح را محدود نمی کنند. اعمال فیلتر به این روش بر روی تصویر انتخابی در شکل ۴-۳ آمده است.

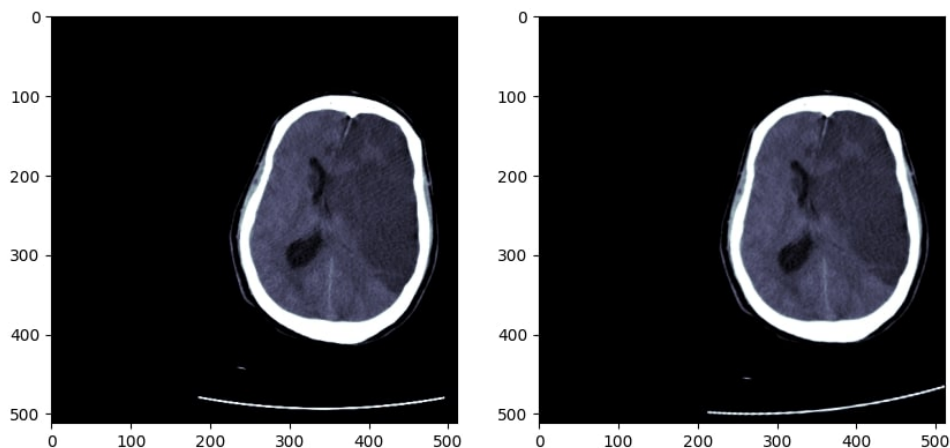
^۷Median Filtering



شکل ۴-۳: تصویر انتخابی برای نمایش مراحل پیش‌پردازش، پیش و پس از اعمال فیلتر میانه. سمت راست تصویر پس از اعمال فیلتر را نشان می‌دهد.

۴-۱-۳ تنظیم زاویه

تصاویر مغز ممکن است به دلایلی چون حرکت سر بیمار، همگی زاویه‌ی یکسانی نداشته باشند. به منظور یک‌دست سازی تصاویر از این جهت، ابتدا ناحیه‌ی سر به صورت یک بیضی تخمین زده می‌شود و سپس با دوران حول یکی از قطره‌هایش، در زاویه‌ی قائم قرار می‌گیرد. نمونه‌ی این محاسبات و تغییرات در تصویر انتخابی در شکل ۴-۴ قابل مشاهده است.

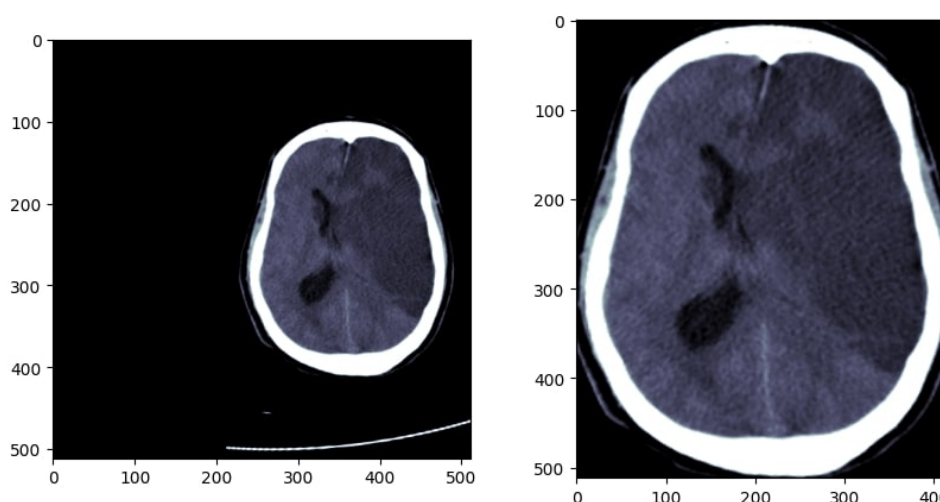


شکل ۴-۴: تصویر انتخابی برای نمایش مراحل پیش‌پردازش، پیش و پس از تنظیم زاویه. سمت راست تصویر پس از تنظیم زاویه را نشان می‌دهد.

۴-۱-۴ محدودسازی تصویر به ناحیهی مغز

به دلایل مشابه قسمت قبل، ناحیهی مغز بیمار ممکن است دقیقاً در مرکز تصویر قرار نداشته باشد. علاوه بر این، چنانکه در تصویر ۴-۴ مشخص است، در هر تصویر، هوای اطراف سر بیمار و گاهی حتی بخش‌هایی دستگاه تصویربرداری نیز حضور دارد. مقدار این اضافات نیز از تصویری به تصویر دیگر متفاوت است. یکی از مراحل پیش‌پردازش مربوط به همین مسئله می‌باشد.

در این مرحله، با برش تصویر و محدودسازی آن به ناحیهی مغز، علاوه بر اینکه مغز بیمار در مرکز تصویر قرار می‌گیرد، اضافات موجود در تصویر از آن حذف می‌شود. لازم به ذکر است از آن‌جا که ابعاد مغز بیماران متفاوت است، تصاویر حاصل از این مرحله، همگی به یک ابعاد جدید و مشخص، تغییر اندازه نیز می‌یابند. خروجی این مرحله بر روی تصویر نمونه در شکل ۴-۵ آمده‌است.



شکل ۴-۵: تصویر انتخابی برای نمایش مراحل پیش‌پردازش، پیش و پس از محدودسازی به ناحیهی مغز. سمت راست، تصویر پس از محدودسازی و استانداردسازی ابعاد را نشان می‌دهد.

۴-۱-۵ حذف استخوان جمجمه

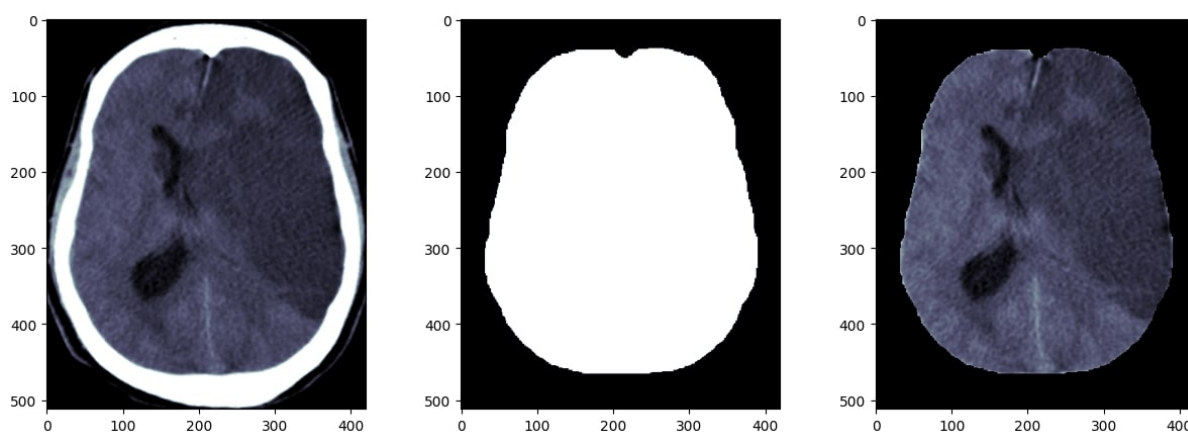
طبیعتاً تمام برش‌های مغز شامل بخشی از استخوان جمجمه در اطراف مغز هستند. اما این بافت استخوانی، به هر صورت، اطلاعات زائد و فاقد اهمیت در تشخیص سکته‌ی مغزی محسوب می‌شود. هر چه اطلاعات نامربوط کمتری به مدل عرضه شود، توانایی آن در یادگیری صحیح ویژگی‌های کلیدی نیز افزایش می‌یابد. به همین منظور، مرحله‌ی بعدی پیش‌پردازش، به حذف استخوان جمجمه^۸ و یا استخراج بافت مغز^۹ از تصاویر اختصاص دارد.

^۸Skull Stripping
^۹Brain Extraction

روش‌های مختلفی برای این کار وجود دارد. یک روش، آموزش یک مدل مجزا برای ناحیه‌بندی استخوان جمجمه می‌باشد. روش دیگر، استفاده از انطباق تصاویر است که در بخش مفاهیم اولیه به آن اشاره شد. به این معنا که هر تصویر جدید، با یک تصویر استاندارد با ناحیه‌بندی مشخص برای بافت مغز، انطباق داده شود و بافت مغزی متناظر با آن از تصویر جدید استخراج گردد. اما روش اول زمان‌بر بوده و نیازمند تعداد زیادی تصویر با برچسب ناحیه‌ی استخوان جمجمه برای هر پیکسل است. روش دوم نیز گاهی نادقیق عمل می‌کند. در این پژوهش از روش اعمال آستانه بر روی مقدار پیکسل‌ها استفاده می‌شود که شرح آن در ادامه خواهد آمد.

در این روش، از این حقیقت استفاده می‌شود که بافت استخوانی، چنانکه در بخش افزایش وضوح آمد، دارای بیشترین مقدار در واحد HU است. به این ترتیب با اعمال یک حد آستانه بر روی مقدار پیکسل‌ها، می‌توان این بافت را از تصویر حذف کرد. البته در این رابطه چالش‌هایی نیز وجود دارد. زیرا با حذف این مقادیر، یک سری پیکسل‌ها در میانه‌ی بافت مغز نیز حذف می‌شوند و حفره‌هایی در آن ایجاد می‌شود. همچنین در برخی تصاویر، ممکن است بافت‌های نازکی در اطراف استخوان جمجمه وجود داشته باشند که با اعمال حد آستانه حذف نشوند.

به منظور رفع این مشکلات، پس از اعمال آستانه، از برخی روش‌های ریخت‌شناسی^{۱۰} در پردازش تصویر استفاده می‌شود تا حفره‌های کوچک تصویر حاصل پر شوند و خطوط باریکی که به عنوان بافت مغز تشخیص داده شده‌اند، حذف گردند. نمونه‌ی استخراج بافت مغزی به این روش بر روی تصویر نمونه، در شکل ۴-۶ قابل مشاهده است.

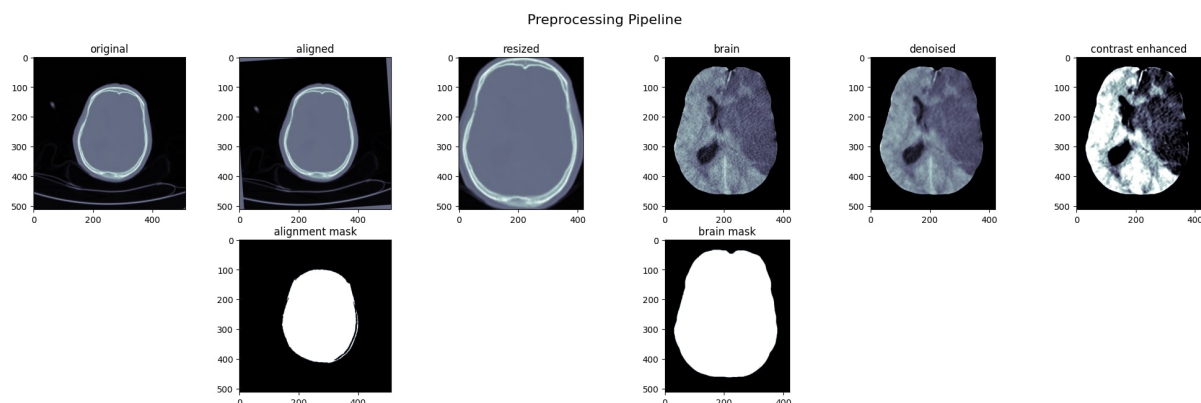


شکل ۴-۶: تصویر انتخابی برای نمایش مراحل پیش‌پردازش، پیش و پس از حذف استخوان جمجمه. تصویر وسط، ناحیه‌ی تشخیص داده شده به عنوان بافت مغز و سمت راست تصویر حاصل از حذف استخوان جمجمه را نشان می‌دهد.

همانطور که در قسمت افزایش وضوح ذکر شد پس از تکمیل این مرحله، ناحیه‌ی تشخیص داده شده

^{۱۰}Morphology

به عنوان بافت مغز، برای تحلیل‌های آماری بر روی مقدار پیکسل‌های بافت مغز مورد استفاده قرار می‌گیرد و در افزایش وضوح تصویر به‌کار می‌آید. با استفاده از این روش، وضوح تصاویر و تمایز بافت سالم از آسیب‌دیده، به طور قابل توجهی افزایش می‌یابد. شکل ۴-۷ مراحل کامل پیش‌پردازش بر روی تصویر انتخابی و نتیجه‌ی نهایی آن را نشان می‌دهد.



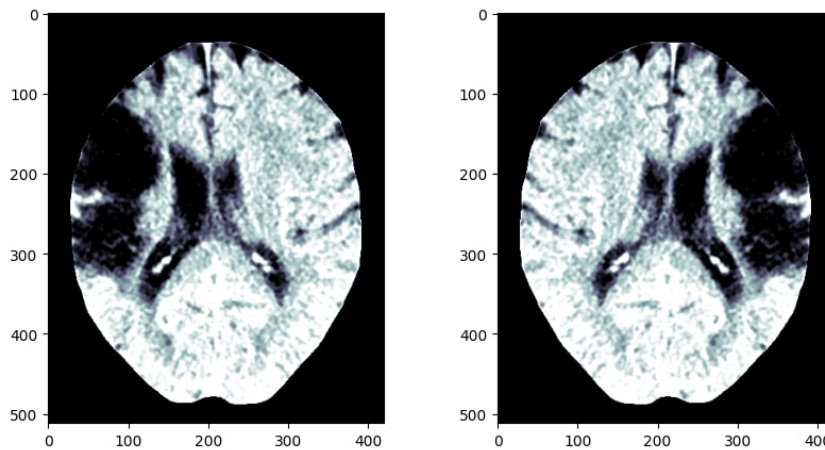
شکل ۴-۷: مراحل کامل فرایند پیش‌پردازش بر روی تصویر انتخابی. از چپ به راست، این مراحل شامل تنظیم زاویه، محدودسازی به ناحیه‌ی مغز، حذف استخوان جمجمه، کاهش نویز و افزایش وضوح می‌باشند. دو تصویر ردیف پایین نیز از چپ به راست، ناحیه‌ی تشخیص داده‌شده به عنوان سر (با تخمین سر به صورت بیضی) و ناحیه‌ی تشخیص داده‌شده به عنوان بافت مغز را نشان می‌دهند.

۴-۲ داده‌افزایی

زمانی که داده‌ها پیش‌پردازش شوند، آماده‌ی عرضه به مدل خواهند بود. اما همانطور که پیش‌تر ذکر شد، تعداد این داده‌ها محدود است و مدل نمی‌تواند یادگیری قابل اعتمادی از این تعداد تصویر داشته‌باشد. بنابراین با مقدمه‌ای که در فصل مفاهیم اولیه آمد، لازم است حجم تصاویر با روش‌های مناسبی افزایش یابد. این روش‌ها در قالب داده‌افزایی مطرح می‌شوند. در این قسمت به روش‌های داده‌افزایی مورد استفاده در این پژوهش اشاره می‌شود.

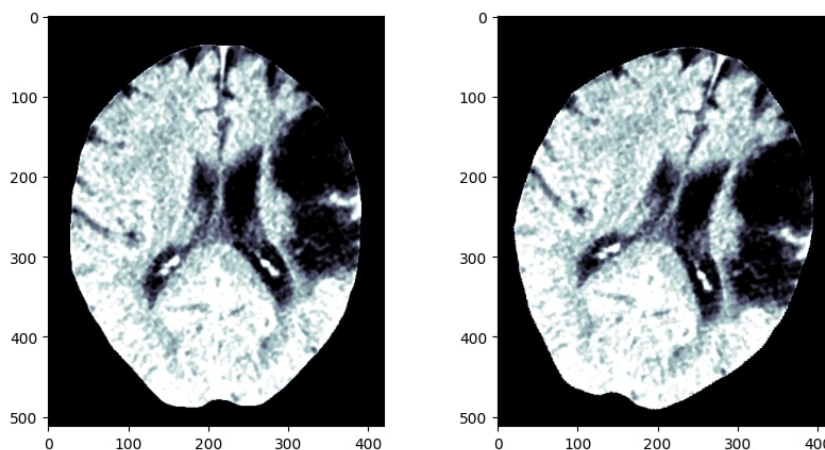
اولین تکنیک داده‌افزایی مورد استفاده، قرینه‌سازی تصادفی تصویر مغز است. به این معنا که جای نیم‌کره‌ی چپ و راست هر تصویر ورودی، در هر بار مشاهده توسط مدل، به صورت تصادفی عوض می‌شود. به این ترتیب اگر یک تصویر با آسیب دیدگی در نیم‌کره‌ی چپ در مجموعه داده وجود داشته باشد، مدل می‌تواند آسیب دیدگی در نیم‌کره‌ی راست را نیز بیاموزد. یک نمونه از این داده‌افزایی در شکل ۴-۸ قابل مشاهده است.

روش دیگری که برای افزایش داده‌ها به‌کار گرفته شد، دوران مغز است. در این تغییر، محور عمودی مغز



شکل ۴-۸: داده‌افزایی به روش قرینه‌سازی تصادفی افقی

بین 30° تا 30° درجه به صورت تصادفی دوران می‌یابد. نمونه‌ی این تغییر در تصویر ۴-۹ قابل مشاهده است.



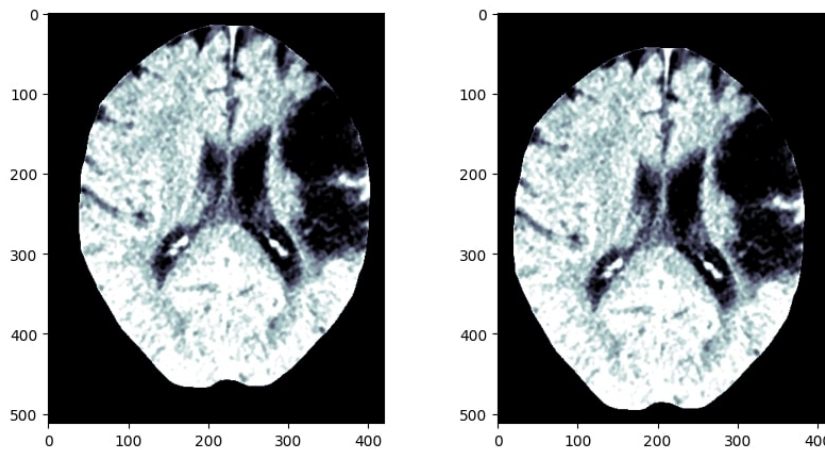
شکل ۴-۹: داده‌افزایی به روش دوران تصادفی

تکنیک دیگری که برای داده‌افزایی مورد استفاده قرار گرفته است، جابجایی جزئی مغز در تصاویر است. این جابجایی در حدود یک‌صدم اندازه‌ی تصویر صورت می‌گیرد. نمونه‌ی این تغییر در تصویر ۴-۱۰ قابل مشاهده است.

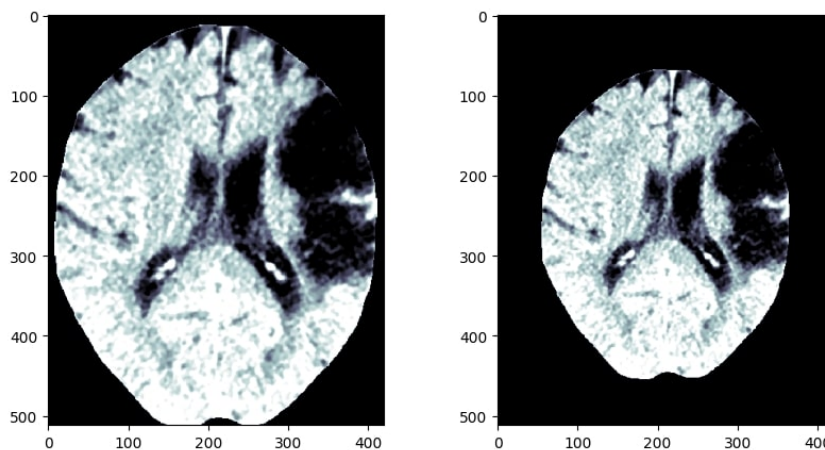
تغییر مقیاس تصادفی روش دیگری برای داده‌افزایی است که در این پروژه مورد استفاده قرار گرفته است. افزایش و کاهش مقیاس تصاویر مغزی در حدود یک‌دهم اندازه‌ی اصلی تصویر صورت می‌گیرد. یک نمونه از تبدیل مقیاس تصادفی در تصویر ۴-۱۱ قابل مشاهده است.

دو روش دیگر برای افزایش داده‌ها، تبدیل برشی^{۱۱} تصادفی و زاویه‌ی دید^{۱۲} تصادفی هستند. در یک

^{۱۱} Shearing
^{۱۲} Perspective



شکل ۴-۱۰: داده‌افزایی به روش جابجایی تصادفی

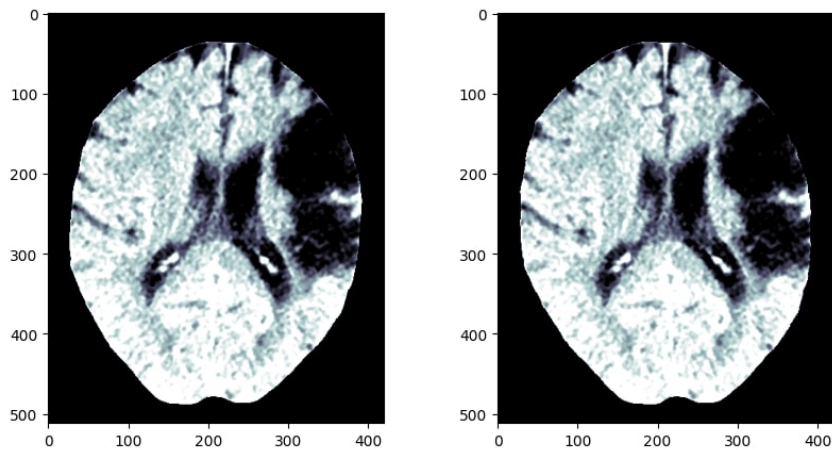


شکل ۴-۱۱: داده‌افزایی به روش تغییر مقیاس تصادفی

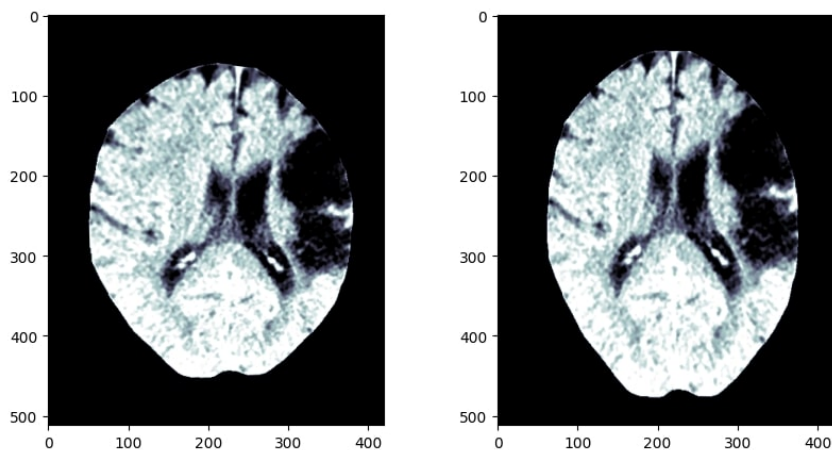
تعریف غیر رسمی، تبدیل اول، تصویر مغز را کمی در راستای افقی کج‌تر می‌کند و تبدیل دوم، زاویه‌ی نگاه متفاوت به مغز را شبیه‌سازی می‌کند. طبق آزمایش‌های انجام‌شده، تبدیل دوم برای افزایش داده‌ها تاثیر مطلوبی بر روی عملکرد مدل نداشته و در این پروژه به‌کار نیامده‌است. نمونه‌ی این دو تبدیل به ترتیب در تصاویر ۴-۱۲ و ۴-۱۳ آمده‌است.

آخرین تکنیکی که برای افزایش داده‌ها قابل استفاده می‌باشد، مات‌سازی گاوسی^{۱۳} تصادفی است. در این روش، مقدار هر پیکسل با میانگین وزن‌داری از یک همسایگی آن پیکسل جایگزین می‌شود. به این ترتیب، تصویر قدری محوتر و درهم‌تنیده‌تر دیده می‌شود. این تبدیل نیز طی آزمایش‌های صورت‌گرفته، در بهبود عملکرد مدل تاثیر مطلوبی نداشته و در این پژوهش مورد استفاده قرار نگرفته‌است. شکل ۴-۱۴ یک نمونه از این نوع داده‌افزایی را نشان می‌دهد. البته در این تصویر، به منظور نمایش بهتر، اندازه‌ی فیلتر

^{۱۳}Gaussian Blur



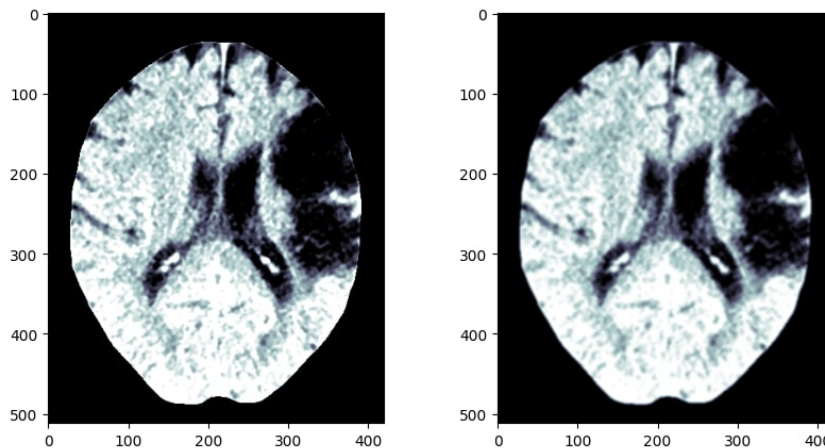
شکل ۴-۱۲: داده‌افزایی به روش تبدیل برشی تصادفی



شکل ۴-۱۳: داده‌افزایی به روش زاویه‌ی دید تصادفی

گاوسی مقدار بزرگی در نظر گرفته شده.

تمام تکنیک‌های مذکور به صورت تصادفی بر روی هر تصویر پیش‌پردازش شده اعمال شده و نتیجه‌ی آن‌ها به مدل ورودی داده می‌شوند. به این ترتیب، مجموعه داده‌ی در دسترس، با تغییراتی جزئی، افزایش حجم قابل توجهی می‌یابد. باید توجه داشت که این تغییرات، شکل آسیب‌دیدگی را تغییر نمی‌دهند و تنوع موارد سکتی مغزی‌ای که مدل می‌بیند را افزایش نمی‌دهند. بلکه تنها باعث می‌شوند نمونه‌های موجود به صورت مطمئن‌تری فراگرفته شوند. درواقع تنوع به وجود آمده، باعث می‌شود مدل، ظاهر خاص داده‌های ورودی را حفظ نکند. بلکه ویژگی‌های کلیدی آن‌ها را فراگیرد.



شکل ۴-۱۴: داده‌افزایی به روش مات‌سازی گاوسی تصادفی

۳-۴ ساختار ورودی و خروجی

همانطور که پیش‌تر ذکر شد، به ازای هر بیمار تعداد زیادی تصویر از برش‌های مختلف مغز وی وجود دارد. پژوهش‌های مختلف بر روی تعداد متفاوتی از این تصاویر کار می‌کنند. برخی از آن‌ها تمام برش‌ها را مورد بررسی قرار داده و یک مطالعه‌ی سه‌بعدی انجام می‌دهند. برخی نیز بر روی تنها دو برش خاص از مغز کار می‌کنند. در این قسمت ابتدا ساختار ورودی مدل ارائه‌شده در این پروژه ذکر می‌شود و مشخص می‌شود که این ساختار، حالت میانه‌ای از دو حالت ذکر شده می‌باشد.

در مورد فرمت خروجی مدل پژوهش‌های مشابه نیز با هم تفاوت‌هایی وجود دارد. برخی از کارهای پیشین با دریافت تصاویر یک بیمار، یک امتیاز از صفر تا ۱۰ خروجی می‌دهند و برخی، امتیاز ASPECTS را به صورت دوبخشی محاسبه می‌کنند. یعنی بالاتر و پایین‌تر بودن امتیاز ASPECTS نسبت به یک عدد آستانه را مشخص می‌کنند. در ادامه، فرمت خروجی مدل پیشنهادی نیز شرح داده می‌شود و مجدداً مشاهده می‌شود که این ساختار، حالت میانه‌ای از دو نوع ذکر شده است که برخاسته از محدودیت‌های داده‌ای می‌باشد.

۱-۳-۴ ساختار ورودی

تصویر ۲-۲ در فصل مفاهیم اولیه، دو برش به‌خصوص از مغز را نشان می‌دهد که بخش‌هایی از هر ۱۰ ناحیه‌ی ASPECTS را در برمی‌گیرند. اما حجم واقعی این نواحی، به همین دو برش محدود نمی‌شود. بلکه در چندین برش از مغز گسترده شده‌است. شکل ۲-۳ از همین بخش، ۸ برش مغز را نشان می‌دهد که تقریباً بهتری از این نواحی هستند و گستردگی آن‌ها را تا حد بسیار خوبی پوشش می‌دهند. در عمل

نیز متخصصان ناحیه‌های ASPECTS را نه فقط در دو برش خاص از مغز، بلکه در چندین برش مجاور منطبق بر کالبدشناسی تصویر ۲-۳ مشاهده می‌کنند.

بنابراین به نظر می‌رسد که بررسی این ۸ برش، به واقعیت تشخیص‌های انسانی نزدیک‌تر است. یکی از نقاط قوت این پژوهش نیز از همین موضوع نشأت می‌گیرد. در این پروژه، ۶ برش از مغز بیماران، به مدل ورودی داده می‌شود. این در حالی است که در جستجویی که در کارهای پیشین انجام شد، این کارها یا ورودی کاملاً سه‌بعدی داشتند و تمام برش‌ها را بررسی می‌کردند و یا تنها دو برش خاص از مغز را به مدل ورودی می‌دادند.

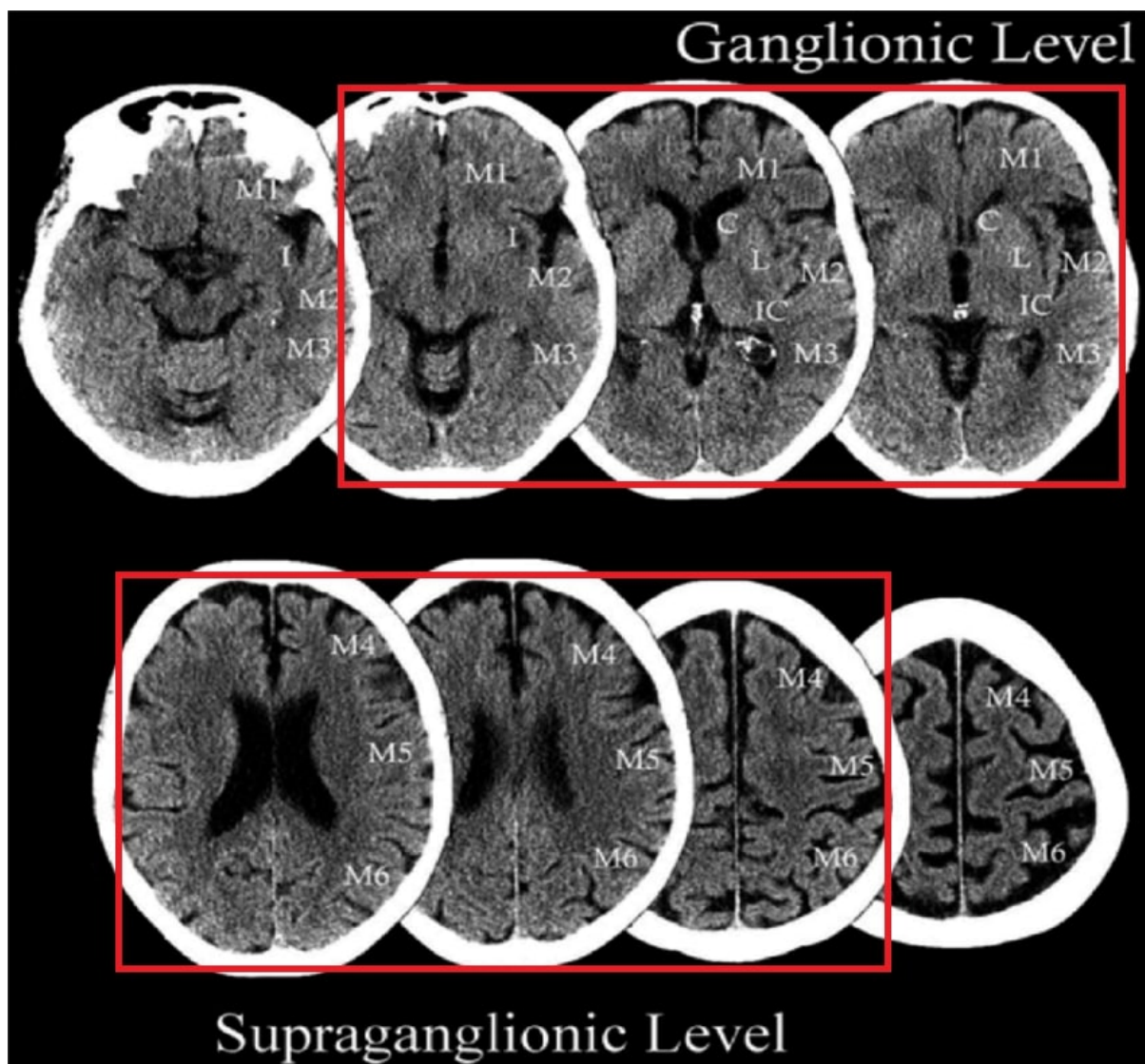
بررسی تنها دو برش از تصاویر ورودی می‌تواند قابلیت اطمینان تشخیص مدل را کاهش دهد. چرا که قسمت اعظم برخی آسیب‌دیدگی‌های مشاهده‌شده بر روی برش‌های فوقانی یا تحتانی مغز می‌باشد که از دید این مدل‌ها پنهان می‌ماند. بنابراین امید می‌رود که روش انتخابی در این پروژه که انطباق بیشتری با واقعیت دارد، عملکرد بهتری از خود نشان بدهد.

همانطور که گفته شد، مدل پیشنهادی، ۶ برش مغزی از هر بیمار را ورودی می‌گیرد. برش‌های انتخابی در تصویر ۴-۱۵ مشخص شده‌اند. عدم استفاده از برش‌های اول و هشتم به دو دلیل صورت گرفته است. اول آنکه این برش‌ها، نسبت به سایر برش‌های میانی اطلاعات کمتری دارند و حذف آن‌ها می‌تواند فرایند یادگیری را برای مدل ساده‌تر کند. به طور دقیق‌تر، کاهش تعداد ویژگی‌های ورودی، می‌تواند از بیش‌برازش^{۱۴} مدل جلوگیری کند. دوم آنکه برش اول در بسیاری از بیماران، در سطحی از مغز اتخاذ شده بود که بخشی از بافت استخوانی در بافت مغزی نفوذ کرده بود. این مسئله سبب چندبخشی شدن مغز و ایجاد تنوع زیادی در بافت استخراج‌شده مغزی می‌شد و یادگیری را برای مدل مشکل می‌کرد.

۴-۳-۲ ساختار خروجی

خروجی نهایی مدل پروژه از تصاویر هر بیمار، امتیاز دوبخشی ASPECTS است. به طور دقیق‌تر، مدل تشخیص می‌دهد که ASPECTS بیمار ≥ 6 است یا < 6 . اما در صورتی که مدل با همین خروجی آموزش داده شود، نمی‌تواند به خوبی فرا بگیرد. چرا که در این حالت، مجموعه‌ی بسیار متنوعی از تصاویر مغزی از امتیاز ۶ گرفته تا ۱۰، همگی با برچسب ≥ 6 به مدل عرضه می‌شوند. پیدا کردن یک الگوی مشترک میان این مجموعه‌ی متنوع، برای مدل دشوار است. اما در صورتی که به مدل یک راهنمایی صورت بگیرد که دقیقاً مشخص می‌کند هر تصویر متعلق به کدام امتیاز از ۶ تا ۱۰ است، مدل می‌تواند میان تصاویر مربوط به هر امتیاز، الگوی مشترکی پیدا کرده و آن امتیاز را فرا بگیرد.

^{۱۴} Overfitting



شکل ۴-۱۵: برش‌های انتخابی ASPECTS

درواقع بهتر آن است که حقایق و اطلاعات پزشکی بیشتری برای مدل فراهم شود و آموزشش در حالت ۱۰ امتیازی ASPECTS انجام شود. اما همانطور که در فصل کارهای پیشین مشاهده شد، برای هر یک از امتیازهای ۶ <، تعداد بسیار اندکی تصویر وجود دارد. به نحوی که در مجموع تنها ۱۱ بیمار با این امتیازات وجود دارد. بنابراین مدل نمی‌تواند با این تعداد نمونه‌ی اندک، دسته و مجموعه‌ای برای آن امتیاز را آموزش ببیند. به علت همین محدودیت، در این پروژه، حالت میانه‌ای از آموزش دو-کلاسه و ۱۱-کلاسه در پیش گرفته شده‌است. به نحوی که مدل بر روی ۶ دسته‌ی ۵ - ۰، ۶، ۷، ۸، ۹ و ۱۰ آموزش می‌بیند و خروجی نهایی آن، دوبخشی شده و گزارش می‌شود.

۴-۴ طراحی مدل

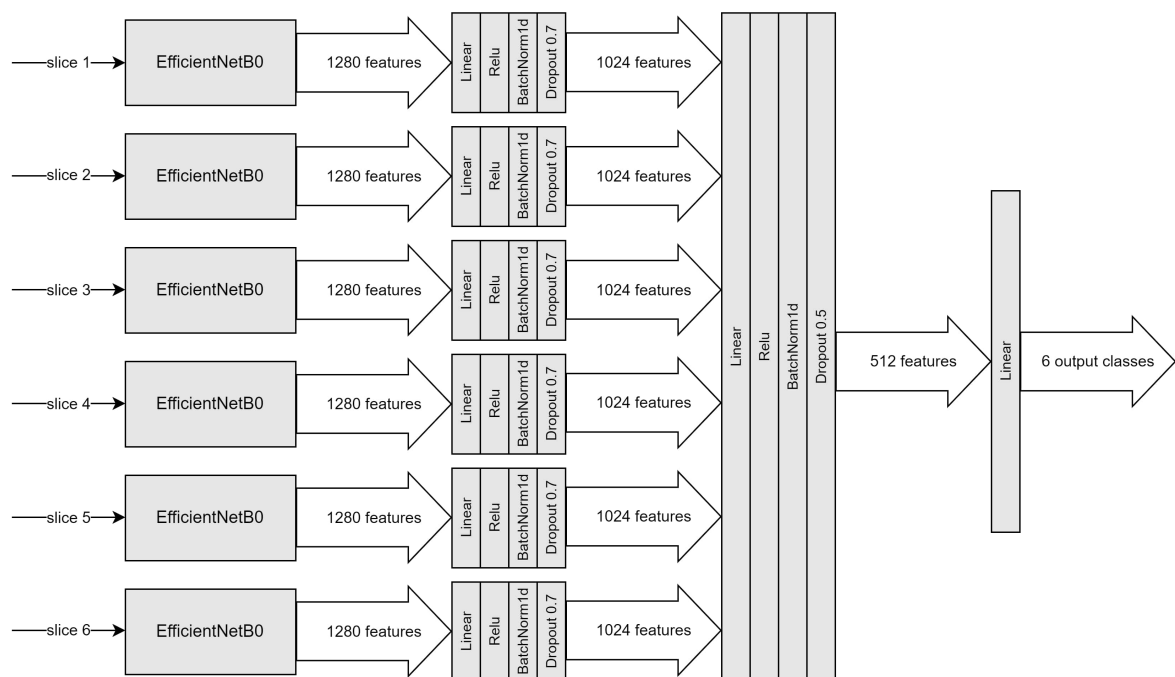
پس از پیش‌پردازش داده‌ها و بررسی ساختار ورودی و خروجی مدل، نوبت به هسته‌ی مرکزی مدل، یعنی ساختار شبکه‌ی آن می‌رسد. همانطور که در فصل مفاهیم اولیه آمد، یکی از راه‌های مدیریت حجم محدود داده‌ها، روش یادگیری انتقالی است. در این پروژه نیز، از این روش استفاده می‌شود و یک مدل پیش‌آموزش‌دیده، به عنوان پایه‌ی شبکه‌ی پیشنهادی قرار می‌گیرد. پس از بررسی مدل‌های پیش‌آموزش‌دیده‌ی مختلف، از جمله مدل‌های VGG، AlexNet، SqueezeNet، InceptionV3، ResNet، DenseNet و EfficientNetV2، مشخص شد که مدل EfficientNetB0 بهتر از همه نتیجه می‌دهد و لذا در این پروژه به‌کار گرفته شده‌است.

مدل EfficientNetB0 میلیون‌ها پارامتر دارد و بر روی میلیون‌ها تصویر آموزش دیده‌است تا آن‌ها را در ۱۰۰۰ دسته طبقه‌بندی کند. تصاویر ورودی و دسته‌های خروجی در مدل مورد نیاز ما متفاوت هستند. اما همان‌گونه که پیش‌تر گفته شد، به کمک یادگیری انتقالی می‌توان از توانایی مدل EfficientNetB0 برای تشخیص اشکال مختلف و استخراج ویژگی‌های آن‌ها بهره برد.

مدل EfficientNetB0 یک تصویر دوبعدی رنگی را ورودی می‌گیرد، یک سری ویژگی‌های کلیدی آن را استخراج می‌کند و سپس در چند لایه‌ی انتهایی مدل، که به آن قسمت طبقه‌بندی‌کننده^{۱۵} گفته می‌شود، تصویر را در یکی از دسته‌های خروجی خود قرار می‌دهد. در یادگیری انتقالی، قسمت طبقه‌بندی‌کننده از مدل پیش‌آموزش‌دیده حذف می‌شود. به این ترتیب تنها از قسمتی از مدل استفاده می‌شود که ویژگی‌ها را از تصاویر استخراج می‌کنند.

همانطور که گفته شد، مدل EfficientNetB0 تنها یک تصویر را دریافت و پردازش می‌کند. از آن‌جا که مدل پیشنهادی در این پروژه، ۶ تصویر ورودی دارد، برای استخراج ویژگی‌های این تصاویر، ۶ نمونه^{۱۶} از مدل EfficientNetB0 گرفته می‌شود. ویژگی‌های استخراج شده از هر تصویر، پشت یک لایه‌ی جدید قرار می‌گیرند تا یک مرحله پردازش شده و به منظور خاص پروژه نزدیک‌تر شوند. سپس این ویژگی‌های خاص منظوره‌تر، همگی با هم الصاق می‌شوند و طی دو لایه‌ی جدید دیگر، طبقه‌بندی نهایی تصاویر را انجام می‌دهند. با این توضیحات، ساختار نهایی شبکه در تصویر ۴-۱۶ قابل مشاهده است. جزئیات این شبکه، طی آزمایش تعداد بسیار زیادی ساختار مختلف و ارزیابی نتایج حاصل از آن‌ها به صورت تجربی به دست آمده‌است.

Classifier^{۱۵}
Instance^{۱۶}



شکل ۴-۱۶: ساختار شبکه‌ی پیشنهادی

۴-۵ جمع‌بندی

در این فصل به شرح هسته‌ی اصلی پژوهش، یعنی روش حل مسئله، پرداخته شد. ترتیب فعالیت‌های انجام‌شده برای حل مسئله در این پروژه، منطبق بر عناوین همین فصل بوده است. درواقع در ابتدا، اطلاعات دریافتی از مرکز درمانی مربوطه، گردآوری و مرتب شده‌اند. سپس چهارچوب ورودی مدل، با توجه به مفاهیم پزشکی و روش‌های مورد استفاده توسط متخصصان تعیین گشته است. همچنین با توجه به نیازمندی‌های پروژه و تا حدی محدودیت‌های حاکم بر آن، قالب خروجی مدل به صورت دویخشی انتخاب شده است. در ادامه نیز طراحی مدل انجام شده و طی آزمایش‌های بسیار، دچار تغییر و تحول گشته تا به حالت نهایی خود به صورتی که در این فصل ذکر شد، رسیده است. فصل‌های بعدی پایان‌نامه به ارزیابی و تحلیل جایگاه این روش اختصاص خواهد داشت.

فصل ۵

آزمایش‌ها

در این فصل، پس از شرح روش آموزش و آزمایش مدل، نتایج به دست آمده از ارزیابی مدل ذکر می‌شود. این نتایج به تفکیک هر لایه‌ی ارزیابی گزارش می‌شوند و جمع‌بندی عملکرد مدل بر روی کل مجموعه داده نیز عنوان می‌شود.

۵-۱ آموزش مدل

در این قسمت، جزئیات فنی آموزش مدل شرح داده می‌شود. تنظیماتی که در این قسمت ذکر می‌شوند نیز از یک جستجوی ساخت یافته بر روی حالت‌های ممکن و ارزیابی نتایج هر حالت به دست آمده است. این تنظیمات و پارامترها شامل نوع تابع زیان،^۱ نوع بهینه‌ساز،^۲ نرخ یادگیری^۳ نوع و پارامترهای زمان‌بند،^۴ تعداد دورهای آموزش،^۵ اندازه‌ی دسته^۶ و سیاست انتخاب مدل می‌باشند که مقدار آن‌ها به شرح زیر است.

- **تابع زیان:** با توجه به چند-کلاسه بودن خروجی مدل پیشنهادی، تابع زیان مورد استفاده برای آموزش مدل، تابع زیان آنتروپی متقاطع^۷ چند-کلاسه می‌باشد.

- **الگوریتم بهینه‌سازی:** الگوریتم بهینه‌سازی‌ای که در آموزش مدل به کار گرفته شده، الگوریتم Adam^۸ است. لازم به ذکر است که روش مورد استفاده در این پروژه، یادگیری انتقالی است و نه تنظیم دقیق^۹.

^۱ Loss Function

^۲ Optimizer

^۳ Learning Rate

^۴ Scheduler

^۵ Epoch

^۶ Batch Size

^۷ Crossentropy

^۸ Fine Tuning

بنابراین تمام پارامترهای مدل EfficientNetB0 ثابت می‌شوند و بهینه‌سازی، آن‌ها را دستخوش تغییر نمی‌کند.

- **نرخ یادگیری:** نرخ یادگیری در نظر گرفته شده برای بهینه‌سازی، ۰/۰۰۰۵ می‌باشد.
- **الگوریتم زمان‌بند:** الگوریتم مورد استفاده برای زمان‌بند، الگوریتم StepLR می‌باشد که در هر چند دور از آموزش، نرخ یادگیری را با یک ضریب کاهش می‌دهد. پارامترهای تنظیم شده برای این زمان‌بند به صورت زیر هستند.

– تعداد گام: ۸

– ضریب کاهش (Gamma): ۰/۱

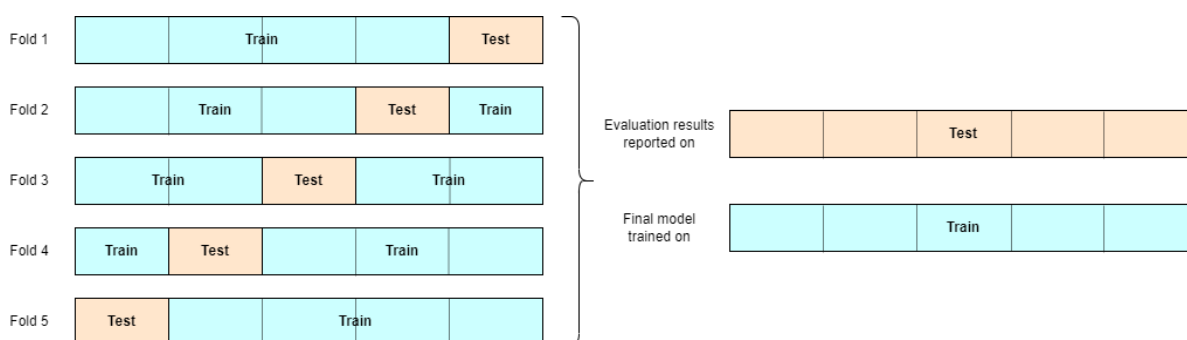
- **تعداد دوره‌های آموزش:** تعداد ۵۰ دور برای آموزش مدل در نظر گرفته شده است.
- **اندازه‌ی دسته:** تصاویر در دسته‌های ۳۶ تایی از مجموعه داده واکشی می‌شوند و در بهینه‌سازی شرکت می‌کنند.
- **سیاست انتخاب مدل:** در بسیاری از پروژه‌های یادگیری ماشین، بخشی از داده‌ها به عنوان مجموعه‌ی ارزیابی جدا می‌شوند و برای سنجش عملکرد مدل بر روی داده‌های دیده‌نشده در طی دوره‌های آموزشی مورد استفاده قرار می‌گیرند. به این ترتیب، مدل در وضعیتی که بهترین نتایج را کسب کرده است، انتخاب شده و وارد فاز آزمایش می‌شود. اما به علت محدودیت تعداد داده‌های این پژوهش، امکان جداسازی بخشی از داده‌ها به این منظور وجود ندارد. بنابراین سیاست انتخاب مدل باید تنها مبتنی بر عملکرد آن بر روی داده‌های آموزشی باشد. در این پروژه، مدلی که در یک دور آموزشی، بهترین عملکرد را در تشخیص دوبخشی امتیاز ASPECTS داده‌های آموزشی از خود نشان داده باشد، به عنوان مدل نهایی برگزیده می‌شود. نحوه‌ی ارزیابی مدل پیشنهادی که به روش ارزیابی متقابل است، این اطمینان را می‌دهد که مدل انتخاب شده به این روش، قابل اعتماد است و عملکرد خوبش بر روی داده‌های دیده شده را بر روی داده‌های دیده نشده نیز حفظ می‌کند.

۵-۲ آزمایش مدل

با مقدمه‌ای که در بخش مفاهیم اولیه در رابطه با اعتبارسنجی متقابل آمد، در روش پیشنهادی این پروژه از ایده‌ای مشابه اعتبارسنجی متقابل ۵ لایه^۹ استفاده می‌شود. به این معنا که مدل، ۵ مرتبه فرایند آموزش

^۹ 5-fold Cross Validation

و آزمایش را طی می‌کند. در هر کدام از این دفعات، تقریباً یک‌پنجم داده‌ها در مجموعه‌ی دیده‌نشده برای آزمایش قرار می‌گیرند. تا اینکه نهایتاً پس از این ۵ مرحله، عملکرد مدل بر روی تمام داده‌های موجود ارزیابی شده‌باشد. در انتها، پس از کسب اطمینان از عملکرد مدل به این روش، مدل بر روی تمام داده‌های موجود آموزش می‌بیند و به عنوان مدل نهایی ارائه می‌شود. شکل ۵-۱ شمایی از ارزیابی متقابل مورد استفاده در این پژوهش را نشان می‌دهد.



شکل ۵-۱: شمایی از روش ارزیابی متقابل. مدل در ۵ مرحله‌ی مجزا مورد آموزش و آزمایش قرار می‌گیرد تا نهایتاً بر روی تمام داده‌ها ارزیابی شده‌باشد. نتایج مدل از تجميع مجموعه‌های آزمایشی محاسبه می‌شوند و مدل نهایی نیز از آموزش بر روی تمام داده‌ها ارائه می‌شود.

در انتخاب مجموعه‌ی آزمایشی در هر لایه، از هر دسته امتیاز، به طور متوازی داده‌ی آزمایشی انتخاب شده‌است. به عنوان مثال اگر ۲۵ تصویر در یک کلاس وجود داشته، تقریباً ۵ تصویر از آن کلاس برای هر لایه انتخاب شده‌است. همچنین به این نکته توجه شده‌است که تصاویر موجود برای هر بیمار، یا تماماً در دسته‌ی آموزشی قرار بگیرند و یا تماماً در مجموعه‌ی آزمایشی باشند. به این ترتیب، جدایی و دیده‌نشده‌بودن داده‌ها در ارزیابی مدل به طور کامل رعایت شده‌است.

لازم به ذکر است روش پیشنهادی، مدل را به صورت چند-کلاس آموزش می‌دهد اما سیاست انتخاب مدل و ارزیابی مدل هر دو مبنی بر طبقه‌بندی دوبخشی هستند. به این معنا که خروجی چند-کلاسه‌ی مدل، در یکی از دو دسته‌ی ۶ \geq باشد یا ۶ $<$ قرار می‌گیرد و با امتیاز واقعی مقایسه می‌شود. نتایج گزارش شده به عنوان نتایج آزمایشی مدل نیز از همین نوع هستند.

همانطور که ذکر شد، افزاز ۵ لایه‌ای انجام‌شده بر روی مجموعه‌داده، تقریباً متوازن است. به این معنا که از هر امتیاز ASPECTS تقریباً تعداد یکسانی در هر قسمت وجود دارد. به این ترتیب مدل در فرایند آموزش هر یک از ۵ لایه، تعداد نمونه‌ی مناسبی از هر امتیاز را مشاهده کرده و فرامی‌گیرد. جدول ۵-۱ اندازه‌ی مجموعه‌ی آموزشی هر لایه و توزیع امتیازات ASPECTS در آن‌ها را نشان می‌دهد.

همانطور که مشخص است، مجموع تعداد تصویر این لایه‌ها، ۱۵۹ تصویر است که برابر با تعداد کل

جدول ۵-۱: حجم و ترکیب امتیازهای ASPECTS مجموعه‌ی آزمایشی در پنج لایه‌ی ارزیابی.

تعداد تصویر بر حسب امتیاز ASPECTS							
شماره‌ی لایه	تعداد تصویر کل	۱۰	۹	۸	۷	۶	۵-۰
۱	۳۲	۶	۷	۵	۶	۳	۵
۲	۳۱	۶	۶	۶	۵	۳	۵
۳	۳۴	۷	۷	۶	۵	۴	۵
۴	۳۱	۶	۷	۶	۵	۲	۵
۵	۳۱	۷	۶	۶	۵	۲	۵

تصاویر در دسترس در این پژوهش می‌باشد. جدول ۵-۲ نیز اطلاعات مربوط به مجموعه‌ی آموزشی هر لایه‌ی ارزیابی را نمایش می‌دهد.

جدول ۵-۲: حجم و ترکیب امتیازهای ASPECTS مجموعه‌ی آموزشی در پنج لایه‌ی ارزیابی.

تعداد تصویر بر حسب امتیاز ASPECTS							
شماره‌ی لایه	تعداد تصویر کل	۱۰	۹	۸	۷	۶	۵-۰
۱	۱۲۷	۲۶	۲۶	۲۴	۲۰	۱۱	۲۰
۲	۱۲۸	۲۶	۲۷	۲۳	۲۱	۱۱	۲۰
۳	۱۲۵	۲۵	۲۶	۲۳	۲۱	۱۰	۲۰
۴	۱۲۸	۲۶	۲۶	۲۳	۲۱	۱۲	۲۰
۵	۱۲۸	۲۵	۲۷	۲۳	۲۱	۱۲	۲۰

همانطور که مشخص است، جمع تعداد تصاویر هر مجموعه‌ی آموزشی با مجموعه‌ی متناظرش برابر با ۱۵۹ بوده و کل مجموعه‌داده را پوشش می‌دهد.

۵-۳ تشخیص‌های انجام‌شده

در هر یک از پنج لایه‌ی ارزیابی، مدل بر روی مجموعه‌ی آموزشی، یادگیری خود را تکمیل می‌کند و بر بروی مجموعه‌ی کاملاً دیده‌نشده، آزموده می‌شود. درواقع در هر دور، تمام داده‌های آزمایشی به مدل عرضه می‌شوند تا امتیاز ASPECTS دوبخشی‌شان توسط مدل تشخیص داده‌شود. تشخیص منفی به معنای

امتیاز ≥ 6 و تشخیص مثبت به معنای امتیاز $6 <$ می باشد.

در تشخیص های انجام شده توسط مدل، چهار حالت متصور است. حالت اول، مثبت صحیح^{۱۰} می باشد. این حالت تشخیص مثبتی را نشان می دهد که برچسب واقعی اش نیز مثبت است. حالت دوم، مثبت ناصحیح^{۱۱} است. این حالت، تشخیص مثبتی را نشان می دهد که برچسب واقعی اش منفی بوده است. دو معیار دیگر نیز منفی صحیح^{۱۲} و منفی ناصحیح^{۱۳} هستند که به ترتیب، تشخیص منفی با برچسب واقعی منفی و برچسب واقعی مثبت را نشان می دهند.

جدول ۳-۵، اطلاعات تشخیصی مدل در هر لایه ی ارزیابی و جدول ۴-۵، تشخیص های مدل بر روی کل داده ها را نمایش می دهد. این مقادیر برای محاسبه ی معیارهای ارزیابی مدل در قسمت بعد کاربرد دارند.

جدول ۳-۵: تشخیص های انجام شده توسط مدل در پنج لایه ی ارزیابی. F، T، N، P، FN، TN، FP، TP، و کل، به ترتیب تعداد تشخیص های مثبت صحیح، مثبت ناصحیح، منفی صحیح، منفی ناصحیح، تعداد داده های مثبت، تعداد داده های منفی، تعداد تشخیص های مثبت، تعداد تشخیص های منفی و تعداد کل داده ها را نشان می دهد.

شماره ی لایه	TP	FP	TN	FN	P	N	T	F	کل
۱	۵	۰	۲۷	۰	۵	۲۷	۳۲	۰	۳۲
۲	۵	۲	۲۴	۰	۵	۲۶	۲۹	۲	۳۱
۳	۳	۵	۲۴	۲	۵	۲۹	۲۷	۷	۳۴
۴	۵	۱	۲۵	۰	۵	۲۶	۳۰	۱	۳۱
۵	۵	۳	۲۳	۰	۵	۲۶	۲۸	۳	۳۱

جدول ۴-۵: تشخیص های نهایی انجام شده توسط مدل. F، T، N، P، FN، TN، FP، TP، و کل، به ترتیب تعداد تشخیص های مثبت صحیح، مثبت ناصحیح، منفی صحیح، منفی ناصحیح، تعداد داده های مثبت، تعداد داده های منفی، تعداد تشخیص های مثبت، تعداد تشخیص های منفی و تعداد کل داده ها را نشان می دهد.

TP	FP	TN	FN	P	N	T	F	کل
۲۳	۱۱	۱۲۳	۲	۲۵	۱۳۴	۱۴۶	۱۳	۱۵۹

TP^{۱۰}
FP^{۱۱}
TN^{۱۲}
FN^{۱۳}

۴-۵ نتایج ارزیابی

جدول ۵-۵ نتایج ارزیابی مدل در هر لایه‌ی آزمایشی و هر یک از معیارهای دقت،^{۱۴} حساسیت،^{۱۵} تشخیص،^{۱۶} صحت،^{۱۷} بازیابی^{۱۸} و مساحت زیر نمودار مشخصه‌ی عملیاتی گیرنده^{۱۹} را نشان می‌دهد. جدول ۶-۵ نیز مقدار نهایی این معیارها برای مدل را نمایش می‌دهد. این مقادیر نهایی از اجتماع تمام پیش‌بینی‌های انجام شده توسط مدل بر روی تمام داده‌ها و مقایسه‌ی آن‌ها با برچسب‌های واقعی متناظر به‌دست آمده‌اند.

جدول ۵-۵: معیارهای ارزیابی مدل در پنج لایه‌ی ارزیابی.

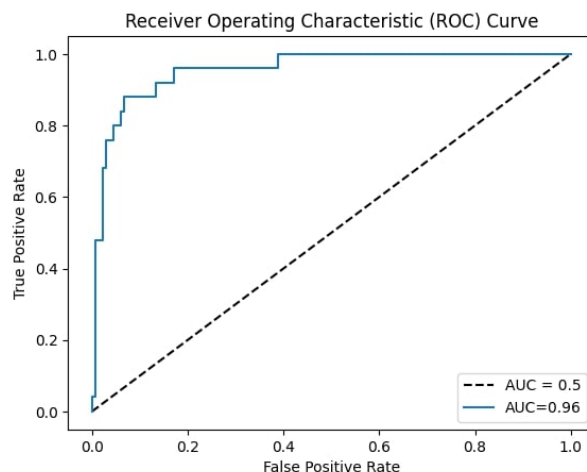
شماره‌ی لایه	دقت	حساسیت	تشخیص	صحت	بازیابی	AUC
۱	۱	۱	۱	۱	۱	۱
۲	۰.۹۳۵	۱	۰.۹۲۳	۰.۷۱۴	۱	۰.۹۹
۳	۰.۷۹۴	۰.۶	۰.۸۲۷	۰.۳۷۵	۰.۶	۰.۸۸
۴	۰.۹۶۷	۱	۰.۹۶۱	۰.۸۳۳	۱	۰.۹۹
۵	۰.۹۰۳	۱	۰.۸۸۴	۰.۶۲۵	۱	۰.۹۹

جدول ۶-۵: معیارهای ارزیابی نهایی مدل.

دقت	حساسیت	تشخیص	صحت	بازیابی	AUC
۰.۹۱۸	۰.۹۲	۰.۹۱۷	۰.۶۷۶	۰.۹۲	۰.۹۶

یکی از معیارهای مورد استفاده در ارزیابی مدل، AUC می‌باشد. AUC یکی از مهم‌ترین معیارهای ارزیابی مدل‌های یادگیری ماشین طبقه‌بندی‌کننده است و نشان می‌دهد مدل چقدر توانایی تشخیص کلاس‌ها از هم را دارد. هر چه این مدل به عدد ۱ نزدیک‌تر باشد، توانایی تشخیص هر کلاس به عنوان همان کلاس در مدل بیشتر است. AUC که کوتاه‌شده‌ی Area Under Curve است، درواقع مساحت زیر نمودار مشخصه‌ی عملیاتی گیرنده^{۲۰} می‌باشد. نمودار AUC-ROC مدل در مجموع تمام داده‌ها در تصویر ۲-۵ آمده‌است.

Accuracy^{۱۴}
Sensitivity^{۱۵}
Specificity^{۱۶}
Precision^{۱۷}
Recall^{۱۸}
AUC^{۱۹}
ROC^{۲۰}



شکل ۵-۲: نمودار AUC-ROC مدل

AUC مدل پیشنهادی در این پژوهش در مقایسه با AUC گزارش شده در تمام کارهای پیشین، بیشترین مقدار را دارد و نشان از قدرت تشخیصی بالای مدل می‌باشد.

۵-۵ مقایسه‌ی نتایج

در فصل کارهای پیشین ذکر شد که در جستجوی تحقیقات حوزه‌ی ASPECTS، پژوهش معتبری با نوع برچسب و محدودیت مشابه این پروژه به دست نیامده است. بنابراین مقایسه‌ی نتایج این پروژه با پژوهش‌های دیگر به اندازه‌ی کافی دقیق نخواهد بود. با این حال، نتایج به دست آمده نشان می‌دهد که مدل پیشنهادی، علی‌رغم تمام محدودیت‌های موجود، در کلاس پژوهش‌هایی با مجموعه داده‌های بزرگ‌تر قرار می‌گیرد و حتی در برخی معیارها از آن‌ها پیشی گرفته است.

از جمله‌ی این پژوهش‌ها می‌توان به [۱۵] اشاره کرد که پژوهش متن‌بازی بر روی بیش از هزار بیمار در سال ۲۰۲۲ بوده و دقت دوبخشی و AUC در حدود ۰.۹۱ گزارش کرده است. نمونه‌ی دیگر، پژوهش تازه‌تری در سال ۲۰۲۳ با تصاویر بیش از ۳۰۰ بیمار می‌باشد که AUC امتیازدهی دوبخشی (۴ ≤) را ۰.۸۹ اعلام کرده است [۱۴]. علاوه بر این موارد، دو پژوهش دوبخشی دیگر با اطلاعات بیش از ۲۵۰ بیمار، در سال‌های ۲۰۲۲ و ۲۰۲۱، به ترتیب دقت ۰.۸۸۹ و ۰.۹۰ و AUC ۰.۸۲ و ۰.۸۷ را اعلام کرده‌اند [۱۶، ۲۲]. یک پژوهش نیز با محدودیت داده‌ای تقریباً مشابه، دقتی در حدود ۰.۷۰ گزارش نموده است [۱۹]. در این بین، یک پژوهش قدیمی‌تر، مربوط به چند سال قبل وجود دارد که دقت ۰.۹۶ و AUC ۰.۸۹ را در بخش نتایج دوبخشی (۴ ≤) ارائه کرده است [۱۱].

پیش‌تر ذکر شد که دو مورد از کارهای پیشین که مستقیماً اشاره‌ای به روش و مجموعه داده‌ی مورد

استفاده در ارزیابی خود نکرده‌اند [۲۰] و یا جدایی داده‌های آزمایشی را رعایت نکرده‌اند [۲۱] در مقایسه‌ی نتایج ذکر نمی‌شوند. جدول ۷-۵، خلاصه‌ای از نتایج به‌دست آمده در سایر کارهای پیشین بر روی طبقه‌بندی دویبخشی ASPECTS و نتایج پروژه‌ی حاضر را نمایش می‌دهد.

جدول ۷-۵: مقایسه‌ی نتایج با کارهای پیشین. در رابطه با خانه‌هایی که با علامت - مشخص شده‌اند، اطلاعاتی گزارش نشده‌است. ستون‌های جدول از راست به چپ، مرجع مورد مقایسه، حجم مجموعه داده‌ی موجود، سطح برجسب در دسترس، آستانه‌ی دویبخشی طبقه‌بندی ASPECTS و معیارهای ارزیابی مورد مقایسه‌ی این مراجع را نشان می‌دهند.

مرجع	داده‌ها	برجسب	آستانه	دقت	حساسیت	تشخیص	AUC
[۱۴]	$300 <$	ناحیه	$5 \geq$	-	۰.۹۴۰	۰.۶۱۹	۰.۸۹
			$7 \geq$	-	۰.۹۵۴	۰.۷۶۶	۰.۸۰
[۱۵]	$1000 <$	پیکسل	$7 \geq$	۰.۹۱۳	۰.۹۵۵	۰.۸۶۶	۰.۹۱۱
[۲۲]	$250 <$	ناحیه	$6 \geq$	۰.۸۸۹	۰.۷۲۲	۰.۹۰۷	۰.۸۲
[۱۱]	$250 <$	پیکسل	$5 \geq$	۰.۹۶۰	۰.۹۷۸	۰.۸۰۰	۰.۸۹
[۱۶]	$250 <$	پیکسل	$5 \geq$	۰.۹۰۰	-	-	۰.۸۷
[۱۹]	۹۰	-	$8 \geq$	۰.۶۸-۰.۷۵	۰.۶۱-۰.۶۳	۰.۱۹-۰.۳۵	-
پژوهش حاضر	۱۰۰	مغز	$6 \geq$	۰.۹۱۸	۰.۹۲۰	۰.۹۱۷	۰.۹۶

۵-۶ جمع‌بندی

مدل پیشنهادی این پروژه، پس از طراحی و بهینه‌سازی، مورد آزمون قرار گرفته‌است. در این فصل، علاوه بر تشریح روش ارزیابی مدل، نتایج حاصل از این ارزیابی به صورت مدون ارائه شده و با کارهای پیشین مقایسه گشته‌است. ارزیابی مدل و مقایسه‌ی آن با پژوهش‌های گذشته از آن جهت اهمیت دارد که جایگاه آن را مشخص می‌کند و با در نظر گرفتن محدودیت‌های موجود می‌تواند معیاری از موفقیت آن را نیز به دست دهد. همانطور که در این فصل عنوان شد، روش پیشنهادی این مدل به نتایج رضایت‌بخشی رسیده‌است و در مقایسه با سایر مطالعات انجام‌شده، به جایگاه برجسته‌ای دست یافته‌است. با این حال، در مورد این پژوهش، کاستی‌ها و محدودیت‌هایی نیز وجود دارد که در فصل آتی تشریح خواهد شد.

فصل ۶

نتیجه‌گیری

در این فصل با جمع‌بندی نتایج روش پیشنهادی، نقاط قوت و کاستی‌های آن، جایگاه این پژوهش در حوزه‌ی تشخیص ASPECTS مورد بحث قرار می‌گیرد. در انتها نیز پیشنهاداتی برای پژوهش‌های مرتبط با این حوزه و یا ادامه‌ی پروژه‌ی حاضر ارائه می‌گردد.

۶-۱ کاربردپذیری

نتایجی که در فصل گذشته ارائه شد، نشان می‌دهد که این پژوهش در میان کارهای پیشین، جزء مطالعات پیشرو محسوب می‌شود البته همانطور که ذکر شد، نتایج برجسته‌ی روش پیشنهادی این پروژه، مربوط به حوزه‌ی پژوهشی می‌باشد و در حوزه‌ی کاربرد پزشکی به حد اطمینان قابل قبولی نرسیده‌اند. چرا که در فاز عملیاتی، کوچک‌ترین تشخیص نادرستی می‌تواند آسیب‌های جبران ناپذیری به افراد وارد کند. بعلاوه در این پژوهش، عملکرد مدل بر روی داده‌های موجود بهینه‌سازی شده‌است و دقت‌های گزارش شده، بالاترین دقت‌های ممکن بر روی این مجموعه داده هستند. این بدان معناست که مدل در مواجهه با داده‌های جدید ممکن است رفتار متفاوتی از خود نشان دهد و کوچک بودن نمونه‌ی در دسترس آن، این احتمال را بیشتر هم می‌کند.

در نتیجه ضروری است که مدل بر روی تعداد بیشتری داده آموزش و آزمایش شود. در این صورت ممکن است مدل خروجی بتواند به عنوان یک دستیار در تشخیص انسانی مورد استفاده قرار بگیرد. که البته این خود نیاز به بررسی‌های بیشتر دارد و موضوع تحقیق برخی مقالات در این زمینه بوده‌است.

لازم است توجه شود که قابل اطمینان نبودن یک مدل برای کاربرد مستقیم پزشکی، تنها به نحوه‌ی

طراحی آن مدل مربوط نمی‌شود. بلکه گاه از ذات مسئله و امکانات موجود برمی‌آید. پیش‌تر ذکر شد که دانش مدل در برجسب‌های سطح مغز، مبتنی بر مشاهده‌ی حالت‌های مختلف بروز سکت می‌باشد. اما در حجم موجود از تصاویر، مشخصاً تمام حالت‌ها جای نگرفته‌اند. بنابراین حتی اگر دقت گزارش شده از یک مدل به حد بالایی رضایت‌بخش باشد، باید به محدود بودن دانش آن و تعمیم‌ناپذیری آن در داده‌های جدید نیز توجه داشت.

۲-۶ نقاط قوت

اولین نکته‌ی مثبت در رابطه با مطالعه‌ی انجام‌شده، جامعیت و قابلیت بازاستفاده‌ی روش پیشنهادی آن در حوزه‌ی پردازش تصاویر پزشکی است. مراحل پیش‌پردازشی پیشنهادی این پروژه، به‌ویژه روش پویای ارائه‌شده برای افزایش وضوح تصاویر، می‌تواند عیناً یا با تغییر اندک در پژوهش‌های دیگر بر روی تصاویر پزشکی، به خصوص تصاویر مغزی، مورد استفاده قرار بگیرند.

علاوه بر این، ساختار مدل پیشنهادی نیز خاص امتیازدهی ASPECTS نبوده و قابل اعمال برای سایر تشخیص‌های پزشکی بر روی تصاویر می‌باشد. درواقع پژوهش حاضر توانسته نشان بدهد که ویژگی‌های استخراج‌شده توسط مدل پیش‌آموزش‌دیده‌ی EfficientNetB0 برای مقاصد طبقه‌بندی تصاویر مغزی مناسب هستند. همچنین طی یک جستجوی نظام‌مند، توانسته یک ساختار شبکه‌ی طبقه‌بندی‌کننده‌ی مناسب بر روی این ویژگی‌ها ارائه کند که می‌تواند یک الگوی اولیه برای شبکه‌ی پژوهش‌هایی باشد که در این زمینه از یادگیری انتقالی استفاده می‌کنند.

نقطه‌ی قوت دیگر، انطباق روش پیشنهادی این پروژه با روش انسانی تشخیص ASPECTS می‌باشد. طی جستجو‌هایی که در کارهای پیشین به انجام رسیده‌است، این پژوهش، اولین موردی است که تشخیص دوبعدی ASPECTS را به جای دو برش، بر روی شش برش اصلی حاوی نواحی ASPECTS انجام می‌دهد. این ساختار، پایه‌ی دانشی مدل را افزایش داده و نتایج آن را قابل اطمینان‌تر می‌کند.

همچنین این پروژه توانسته‌است روش نسبتاً نوینی را برای آموزش و آزمایش مدل‌های طبقه‌بندی دوبخشی با مجموعه داده‌های نامتوازن ارائه دهد. این پروژه با یک مسئله‌ی طبقه‌بندی دو-کلاس مواجه بوده است. در حالی که مجموعه داده‌ی موجود، اطلاعاتی بیش از ASPECTS دوبخشی را شامل می‌شد که می‌توانست مدل را در یادگیری بهتر الگوها یاری کند. اما تعداد داده‌های اندک برای برخی امتیازات در این مجموعه داده، مانع از آموزش و آزمایش کامل ۱۱- کلاس می‌شد. این پژوهش با طراحی حالت میانه‌ای^۱ از آموزش

^۱ تمام امتیازات بخش ۶ < در قالب یک کلاس گردآوری شده و آموزش بر روی این کلاس تجمیعی و ۵ کلاس امتیازات ۶ ≥ انجام شد.

دو کلاسه و چند- کلاسه نشان داد که استفاده از این اطلاعات افزوده می‌تواند یادگیری و نتایج مدل در فاز آزمایشی را بهبود ببخشد.

۳-۶ کاستی‌ها

همانطور که پیش‌تر ذکر شد، مهم‌ترین کاستی این پژوهش، حجم محدود مجموعه داده و برچسب‌های سطح بالای (برچسب تک‌عددی صفر تا ۱۰ در سطح مغز) آن بوده است. چالش دیگری که در رابطه با مجموعه داده وجود داشته، نامتوازن بودن امتیازات ASPECTS آن بوده است. به نحوی که در کل داده‌ها، تنها ۲۵ بیمار با امتیازات صفر تا ۵ موجود بوده است. این موضوع ممکن است بر نتایج به دست آمده از مدل تاثیر داشته باشد و نیاز است مورد آزمایش بیشتری قرار بگیرد.

همچنین چنانکه در قسمت قبل آمد، تعداد محدود داده‌ها موجب شده حالت‌های بسیاری از بروز سکتة در تصاویر عرضه شده به مدل وجود نداشته باشند. بنابراین لازم است با افزایش تعداد و تنوع تصاویر ورودی، علاوه بر افزایش دانش مدل، عملکرد آن نیز مورد ارزیابی دوباره قرار بگیرد.

نکته‌ی دیگری که لازم است مورد توجه قرار بگیرد، حذف داده‌های دارای نویز شدید و یا کیفیت پایین بوده است. در واقع مجموعه داده‌ی موجود، حاصل از هرس داده‌های اولیه و جدا کردن تصاویر با وضوح و مشخصات پایین می‌باشد. در نتیجه ممکن است در کاربرد عملی که تصاویر با نویز بالا و کیفیت پایین وجود دارند، عملکرد مدل بسیار متفاوت باشد.

علاوه بر این موضوع، داده‌های آموزشی و آزمایشی این پژوهش، هر دو از یک مجموعه هستند و شباهت زیادی با هم دارند. با توجه به اینکه در این پروژه از روش ارزیابی متقابل برای آموزش و سنجش مدل استفاده شده است، این امکان وجود دارد که قابلیت‌های مدل تنها بر روی همین مجموعه داده بهینه شده باشند. بنابراین لازم است که عملکرد مدل پیشنهادی این پروژه در مواجهه با داده‌هایی دیده نشده از مراکز تصویر برداری متنوع‌تر ارزیابی گردد.

نکته‌ی پایانی در این بخش، در رابطه با زمان اخذ تصاویر می‌باشد. معیاری در پژوهش‌های حوزه‌ی سکتة مغزی وجود دارد که زمان از لحظه‌ی بروز علائم تا تصویر برداری را مشخص می‌کند. مجموعه داده‌ی پژوهش حاضر از این جهت مورد تفکیک قرار نگرفته است و لازم است اطلاعات دقیق‌تری از این معیار کسب و در آموزش و آزمایش مدل، یک دست سازی شود.

۴-۶ پیشنهاد ادامه‌ی کار

با در نظر گرفتن کاستی‌های عنوان‌شده، می‌توان چند پژوهش تکمیلی در حوزه‌ی تشخیص ASPECTS را در ادامه‌ی این پژوهش متصور شد. پژوهش حاضر توانسته با امکانات داده‌ای موجود، عملکرد قابل مقایسه‌ای با پژوهش‌های موجود ارائه دهد. از طرفی آموزش مدل‌های تشخیص ASPECTS بر روی برجسب‌های سطح ناحیه و سطح پیکسل، با توجه به دانش بالاتری که به مدل می‌دهد، می‌تواند عملکرد اطمینان‌بخش‌تری را ارائه دهد. بنابراین بازسازی و آزمایش روش پیشنهادی بر روی برجسب‌های سطوح بالاتر می‌تواند در کسب نتایج بهتر و مطمئن‌تر راهگشا باشد. حتی آموزش مدل موجود بر روی تعداد و تنوع بالاتری از مجموعه‌داده ممکن است نتایج جدیدی را آشکار کند.

همچنین یک مورد از مطالعاتی که حول پروژه‌ی حاضر می‌تواند انجام شود، ارزیابی تاثیر استفاده از این مدل بر تشخیص ASPECTS متخصصان و میزان توافق میان آنان در تشخیص است. چنانکه پیش از این نیز مطالعات مشابهی بر روی روش‌های ارائه‌شده‌ی ASPECTS انجام شده‌است. انجام چنین تحقیقاتی برای هدایت روش پیشنهادی به سمت کاربردهای عملی، ضروری است.

پژوهش دیگری که می‌تواند مسیر را برای مطالعات آینده در حوزه‌ی ASPECTS هموار کند، طبقه‌بندی خودکار برش‌های مغزی بر اساس نواحی ASPECTS است. به عبارتی استفاده از یادگیری ماشین برای جداسازی تصاویر مغزی از غیر مغزی و جداسازی ۶ برش ویژه‌ی مورد بررسی از میان تصاویر مغزی. در پژوهش حاضر، به علت حجم محدود داده‌ها، این طبقه‌بندی به صورت انسانی انجام شده‌است. اما چنین ابزاری علاوه بر کمک به ایجاد مجموعه‌داده‌های بزرگ در پژوهش‌های آتی، در عملیاتی‌کردن پروژه‌ی حاضر نیز مفید خواهد بود. مجموعه‌داده‌ی مورد استفاده در این پژوهش، خود می‌تواند پایه‌ای بر یادگیری مدل مذکور باشد.

۵-۶ جمع‌بندی

در فصل انتهایی این پایان‌نامه نقاط قوت و کاستی‌های این پروژه در کنار هم ذکر و بررسی شده‌اند. درک این موفقیت‌ها و ضعف‌ها می‌تواند قابلیت‌های راهکار پیشنهادی را در مرحله‌ی عمل مشخص کند و چشم‌اندازی برای رفع کاستی‌ها و بهره‌برداری از موفقیت‌ها در کارهای آتی به دست دهد. با عنایت به آنچه که در این فصل ذکر شد، نتیجه می‌شود که این روش، علی‌رغم نوآوری‌ها، دستاوردها و نتایج برجسته، نیازمند آزمایش و اعتبارسنجی در محیط عملی و تایید نیروهای متخصص می‌باشد. امید است که پروژه‌ی انجام‌شده، با رفع کاستی‌ها و موانع موجود، در هر دو زمینه‌ی پژوهشی و کاربردی راهگشا باشد.

- [1] E. S. Donkor. Stroke in the century: a snapshot of the burden, epidemiology, and quality of life. *Stroke research and treatment*, 2018, 2018.
- [2] P. A. Barber, A. M. Demchuk, J. Zhang, and A. M. Buchan. Validity and reliability of a quantitative computed tomography score in predicting outcome of hyperacute stroke before thrombolytic therapy. *The Lancet*, 355(9216):1670–1674, 2000.
- [3] N. van Horn, H. Kniep, G. Broocks, L. Meyer, F. Flottmann, M. Bechstein, J. Götz, G. Thomalla, M. Bendszus, S. Bonekamp, et al. Aspects interobserver agreement of 100 investigators from the tension study. *Clinical Neuroradiology*, pages 1–8, 2021.
- [4] W. Chen, J. Wu, R. Wei, S. Wu, C. Xia, D. Wang, D. Liu, L. Zheng, T. Zou, R. Li, et al. Improving the diagnosis of acute ischemic stroke on non-contrast ct using deep learning: a multicenter study. *Insights into Imaging*, 13(1):1–12, 2022.
- [5] J. L. Saver. Time is brain—quantified. *Stroke*, 37(1):263–266, 2006.
- [6] Rapidai. <https://www.rapidai.com/>. Accessed: June 5, 2023.
- [7] Viz.ai - ai-powered stroke care. <https://www.viz.ai/>. Accessed: June 23, 2023.
- [8] e-ASPECTS - artificial intelligence for stroke assessment. <https://www.brainomix.com/stroke/e-aspects/>. Accessed: June 23, 2023.
- [9] L. Le Guennec, C. Cholet, F. Huang, M. Schmidt, N. Bréchet, G. Hékimian, S. Besset, G. Lebreton, A. Nieszkowska, P. Leprince, et al. Ischemic and hemorrhagic brain injury during venoarterial-extracorporeal membrane oxygenation. *Annals of intensive care*, 8:1–10, 2018.
- [10] M. Mokin, C. T. Primiani, A. H. Siddiqui, and A. S. Turk. Aspects (alberta stroke program early ct score) measurement using hounsfield unit values when selecting patients for stroke thrombectomy. *Stroke*, 48(6):1574–1579, 2017.

- [11] H. Kuang, M. Najm, D. Chakraborty, N. Maraj, S. Sohn, M. Goyal, M. Hill, A. Demchuk, B. Menon, and W. Qiu. Automated aspects on noncontrast ct scans in patients with acute ischemic stroke using machine learning. *American journal of neuroradiology*, 40(1):33–38, 2019.
- [12] A. T. Wilson, S. Dey, J. W. Evans, M. Najm, W. Qiu, and B. K. Menon. Minds treating brains: understanding the interpretation of non-contrast ct aspects in acute ischemic stroke. *Expert Review of Cardiovascular Therapy*, 16(2):143–153, 2018.
- [13] T.-L. Tan, K.-S. Sim, and A.-K. Chong. Contrast enhancement of ct brain images for detection of ischemic stroke. In *2012 International Conference on Biomedical Engineering (ICoBE)*, pages 385–388. IEEE, 2012.
- [14] S.-J. Lee, G. Park, D. Kim, S. Jung, S. Song, J. M. Hong, D. H. Shin, and J. S. Lee. Clinical evaluation of a deep-learning model for automatic scoring of the alberta stroke program early ct score on non-contrast ct. *Journal of NeuroInterventional Surgery*, 2023.
- [15] Z. Cao, J. Xu, B. Song, L. Chen, T. Sun, Y. He, Y. Wei, G. Niu, Y. Zhang, Q. Feng, et al. Deep learning derived automated aspects on non-contrast ct scans of acute ischemic stroke patients. Technical report, Wiley Online Library, 2022.
- [16] H. Kuang, B. K. Menon, S. I. Sohn, and W. Qiu. Eis-net: Segmenting early infarct and scoring aspects simultaneously on non-contrast ct of patients with acute ischemic stroke. *Medical Image Analysis*, 70:101984, 2021.
- [17] S.-m. Jung and T.-k. Whangbo. Evaluating a deep-learning system for automatically calculating the stroke aspect score. In *2018 International conference on information and communication technology convergence (ICTC)*, pages 564–567. IEEE, 2018.
- [18] N. Liu and S. Nie. Deep learning-based aspects scoring method for acute ischemic stroke. In *2021 3rd International Conference on Artificial Intelligence and Advanced Manufacture*, pages 99–103, 2021.
- [19] Z. Yu, Z. Chen, Y. Yu, H. Zhu, D. Tong, and Y. Chen. An automated aspects method with atlas-based segmentation. *Computer Methods and Programs in Biomedicine*, 210:106376, 2021.
- [20] M. Naganuma, A. Tachibana, T. Fuchigami, S. Akahori, S. Okumura, K. Yi, Y. Matsuo, K. Ikeno, and T. Yonehara. Alberta stroke program early ct score

- calculation using the deep learning-based brain hemisphere comparison algorithm. *Journal of Stroke and Cerebrovascular Diseases*, 30(7):105791, 2021.
- [21] R. M. Golkonda, V. Menon, and V. Nambiar. Automated aspects classification in acute ischemic stroke using efficientnetv2. In *2022 IEEE 7th International conference for Convergence in Technology (I2CT)*, pages 1–6. IEEE, 2022.
- [22] P.-L. Chiang, S.-Y. Lin, M.-H. Chen, Y.-S. Chen, C.-K. Wang, M.-C. Wu, Y.-T. Huang, M.-Y. Lee, Y.-S. Chen, and W.-C. Lin. Deep learning-based automatic detection of aspects in acute ischemic stroke: Improving stroke assessment on ct scans. *Journal of Clinical Medicine*, 11(17):5159, 2022.
- [23] U. Upadhyay, M. Ranjan, S. Golla, S. Tanamala, P. Sreenivas, S. Chilamkurthy, J. Pandian, and J. Tarpley. Deep-aspects: A segmentation-assisted model for stroke severity measurement. *arXiv preprint arXiv:2203.03622*, 2022.
- [24] L.-C. Ma. Medical image pre-processing with python. <https://towardsdatascience.com/medical-image-pre-processing-with-python-d07694852606>, May 2020. Accessed on June 7, 2023.
- [25] S. Kamalian, M. H. Lev, and R. Gupta. Computed tomography imaging and angiography—principles. *Handbook of clinical neurology*, 135:3–20, 2016.

واژه‌نامه

الف

آنتروپی متقاطع Crossentropy
استخراج بافت مغز Brain Extraction
اعتبارسنجی متقابل Cross Validation
اعتبارسنجی متقابل ۵ لایه 5-fold Cross Validation
الگو Template
اندازه‌ی دسته Batch Size
ایسکمیک Ischemic
انطباق تصاویر Image Registration

ب

باز کردن رگ‌ها Recanalization
بازیابی Recall
بده‌بستان Tradeoff
برچسب Label
برنامه‌های کاربردی Application
بهینه‌ساز Optimizer
بیش‌برازش Overfitting

پ

پهنای پنجره Window Width
پیش‌آموزش دادن Pre-training
پیش‌آموزش دیده Pre-trained
پیش‌پردازش Pre-processing

ت

تابع زیان Loss Function
تبدیل برشی Shearing
تشخیص Specificity
تنظیم دقیق Fine Tuning

چ

چندک Quantile

ح

حساسیت Sensitivity
حذف استخوان جمجمه Skull Stripping

خ

هموراژیک Haemorrhagic

د

داده‌افزایی Data Augmentation
دقت Accuracy
دوبخشی‌شده Dichotomized
دوره آموزش Epoch
دیده‌نشده Unseen

م

Gaussian Blur مات‌سازی گاوسی
 TP..... مثبت‌صحیح
 FP..... مثبت‌ناصحیح
 Dataset..... مجموعه‌داده
 AUC.. مساحت زیر نمودار مشخصه‌ی عملیاتی گیرنده
 ROC مشخصه‌ی عملیاتی گیرنده
 TN..... منفی‌صحیح
 FN..... منفی‌ناصحیح

ن

Segmentation ناحیه‌بندی
 Imbalanced..... نامتوازن
 Learning Rate..... نرخ یادگیری
 Instance..... نمونه

و

Feature..... ویژگی

ی

Transfer Learning یادگیری انتقالی
 Supervised Learning یادگیری تحت نظارت

ر

Modality..... روش
 Morphology ریخت‌شناسی

ز

Perspective زاویه‌ی دید
 Scheduler..... زمان‌بند

س

Header..... سرآیند
 Window Level..... سطح پنجره

ص

Precision..... صحت

ط

Classification..... طبقه‌بندی
 Classifier..... طبقه‌بندی‌کننده

ف

Median Filtering فیلتر میانه

ل

Mechanical Thrombectomy.. لخته‌زدایی مکانیکی

Abstract

Stroke is known as the second leading cause of death worldwide [1]. This dangerous disease can cause permanent and irreversible damage to those affected. Therefore, rapid diagnosis and treatment of stroke in the early stages is of utmost importance. The ASPECTS score is a criterion for evaluating the severity of stroke based on brain CT images. However, the diagnosis of this score on early images, in which the damaged areas may not be clearly visible, is a challenging task even for human experts. This has led to numerous studies on automated ASPECTS scoring in recent years. Automated ASPECTS scoring methods can aid specialists in determining the severity of stroke and therefore increase the speed and accuracy of diagnosis and treatment method selection. This study proposes a systematic solution for ASPECTS score diagnosis based on deep learning methods. The proposed method shows prominent results compared to those of similar studies and could be a valuable contribution to future research.

Keywords: Stroke, Deep Learning, ASPECTS, CT, Transfer Learning



Sharif University of Technology
Department of Computer Engineering

B.Sc. Thesis

Automated ASPECTS Scoring on CT images using deep learnig

By:

Mahta Fetrat

Supervisor:

Dr. Hamid R. Rabiee

June 2023