تشخیص MS از روی تصاویر MRI

سید امین ناجی اصفهانی*[1]، محمدامین اسدی‌نیا[2]، سید مجتبی میرقادری[3]، محمد کاظمی[4]

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

[2]دانشگاه اصفهان، m.a.asadinia@eng.ui.ac.ir

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

[4]دانشگاه اصفهان، m.kazemi@eng.ui.ac.ir

*سرپرست تیم

چکیده - مولتیپل اسکلروزیس یک بیماری خودایمنی است که در آن سیستم ایمنی بدن غلاف میلین نورون‌های مغز و نخاع را تخریب می‌کند. این نواحی به صورت پلاک‌هایی در تصویربرداری تشدید مغناطیسی ظاهر می‌شوند. در روش سنتی تشخیص این پلاک‌ها توسط متخصص مغز و اعصاب انجام می‌شود اما امروزه می‌توان با استفاده از هوش مصنوعی این‌کار را با سرعت و دقت بیشتری انجام داد. به این منظور شبکه‌ای برای حل این چالش طراحی شده که از سه زیرشبکه متفاوت استفاده می‌کند. در لایه اول، از یک شبکه‌ی تقسیم‌بندی استفاده شده تا پلاک‌ها را تشخیص دهد. در لایه دوم برای بهبود تشخیص پلاک‌ها از شبکه کدگذار خودکار و سازوکار تمرکز بر نواحی مهم بهره گرفته شده است. در لایه‌ی آخر نیز شبکه‌ی حافظه کوتاه‌مدت طولانی گنجانده شده است که برای تشخیص محل رخداد پلاک‌ها استفاده می‌شود. شبکه‌ی اصلی با استفاده از تصاویر استخراج و کدگذاری شده از هر زاویه به صورت جداگانه آموزش داده می‌شود. برای ایجاد خروجی نهایی، ماسک‌های تولیدشده‌ی هر زاویه توسط شبکه عصبی دیگری ادغام می‌شوند تا مقدار خطای کلاس‌بندی به حداقل مقدار خود برسد. دقت این مدل روی داده‌های آموزش بالای ۰.۹۹ درصد و خطای کلاس‌بندی کمتر از ۰.۰۰۲ است.

کلید واژه- بیماری MS، تصویربرداری MRI، شبکه تقسیم‌بندی، شبکه حافظه‌ی کوتاه‌مدت طولانی، شبکه عصبی عمیق

۱- مقدمه

امروزه هوش مصنوعی[1] در اکثر حوزه‌های مختلف مورد استفاده قرار می‌گیرد. نوآوری و دقت، دو معیاری هستند که روش‌های مبتنی بر هوش مصنوعی را نسبت به روش‌های سنتی برتر می‌سازد. علم پزشکی نیز حوزه‌ای است که با هوش مصنوعی درآمیخته است و به پزشکان کمک می‌کند تا معاینه‌ی بهتری را در زمان کمتری انجام دهند.

۱-۱- پیش‌گفتار

مولتیپل اسکلروزیس[2] یک بیماری خودایمنی است که سیستم عصبی مرکزی را درگیر می‌کند. تشخیص زودرس این بیماری کمک می‌کند تا جلوی پیشروی آن گرفته شود. یکی از روش‌های مرسوم برای تشخیص قطعی بیماری MS استفاده از تصویربرداری تشدید مغناطیسی[3] است. در MRI با تزریق مواد رادیواکتیو (گادولینیم)، پلاک‌هایی به صورت حلقه‌های ناقص روشن در فاز T1 دیده می‌شوند. این پلاک‌ها می‌توانند در نواحی پری‌ونتریکولار[4]، اینفراتنتوریال[5]، ماده سفید[6] و جاکستاکورتیکال[7] مغز باشند. تعیین تعداد و محل پلاک‌ها بر اساس شاخص مک دونالد در قطعیت تشخیص MS بسیار مهم است.

تشخیص پلاک‌ها و محل قرارگیری آن‌ها در چهار ناحیه یادشده توسط پزشک متخصص انجام می‌شود که با صرف زمان و گاهاً خطا همراه است. در نتیجه مسئله‌ی اصلی ارائه‌ی روشی با بهره‌گیری از هوش مصنوعی است تا بتواند این فرایند را به‌صورت خودکار در مدت زمان کمتری انجام دهد و همچنین دقت بالاتری ارائه کند و یا در حالت‌های خاص‌تر به پزشک معالج راهنمایی دهد. روش موردنظر باید بتواند علاوه بر تشخیص پلاک‌ها در تصویر MRI، محل رخداد آن‌ها را نیز کلاس‌بندی[8] کند.

۱-۲- مرور کارهای پیشین

پژوهش‌های زیادی در زمینه‌ی تشخیص بیماری MS از روی داده‌های

[1] Artificial Intelligence (AI)
[2] Multiple Sclerosis (MS)
[3] Magnetic Resonance Imaging (MRI)
[4] Periventricular
[5] Infratentorial
[6] White matter
[7] Juxtacortical
[8] Classification

تصویری و سیگنالی انجام گرفته است. مارزولو و همکارانش با طراحی شبکه‌ی عصبی کانولوشنی[9]، وضعیت بیماران MS را بر اساس مقیاس وضعیت ناتوانی گسترده[10] تعیین کردند. خطای جذر میانگین مربعات[11] این شبکه برای تعیین وضعیت ۸۹ بیمار MS، ۰.۰۹ درصد گزارش شده است. در پژوهشی دیگر شبکه‌ی عصبی طراحی‌شده با ۲۵ لایه‌ی مختلف توانسته است با دقت ۹۶.۸۸ درصد بیماران MS را از افراد سالم تشخیص دهد. وانگ و همکارانش نیز در پژوهشی مشابه و با طراحی شبکه‌ی عصبی توانستند بیماران MS را با دقت ۹۸.۷۷ درصد شناسایی کنند. شبکه‌ی طراحی‌شده از ۱۴ لایه‌ی شبکه‌ی عصبی کانولوشنی استفاده می‌کند که هر لایه، ویژگی‌های خاصی را برای تصمیم‌گیری نهایی استخراج می‌کند. در سال ۲۰۲۱ میلادی، شروان و گوپتا دو محقق هندی شبکه‌ای را طراحی و معرفی کردند که با وجود لایه‌های بسیار کم قادر بود بیماران MS و افراد سالم را با دقت بالایی کلاس‌بندی کند. این شبکه تنها از ۳ لایه کانولوشنی برای استخراج ویژگی‌های تصاویر MRI استفاده می‌کند و در نهایت توانسته است با استفاده از بهینه‌ساز[12] شیب کاهشی تصادفی[13] به دقت ۹۹.۵۵ درصد برسد. این درحالی است که شبکه‌های دیگر از لایه‌های بیشتری استفاده می‌کنند و به همین علت نیز پیچیدگی بیشتری دارند.

شبکه‌های یادشده همگی از شبکه‌های عصبی کانولوشنی دوبعدی استفاده می‌کنند. این درحالی است که با توجه به پیوستگی تصاویر MRI می‌توان از شبکه‌های سه‌بعدی نیز بهره گرفت. در همین رابطه یو و همکارانش با استفاده از سه لایه‌ی شبکه‌ی عصبی کانولوشنی سه‌بعدی و بهینه‌ساز آدادلتا[14] شبکه‌ای طراحی کردند که می‌تواند میزان پیشرفت MS را هنگام تشخیص زودرس این بیماری پیش‌بینی کند. این شبکه با استفاده از ۱۴۰ داده‌ی کلینیکی متفاوت توانسته است به دقت ۷۲.۹ درصد برسد که نسبت به شبکه‌های قبلی دقت پایینی محسوب می‌شود. در پژوهشی مشابه ژانگ توانسته است شبکه‌ای طراحی کند که بیماران MS را با دقت ۹۸.۲۳ درصد تشخیص دهد. این شبکه از ۷ لایه شبکه‌ی عصبی کانولوشنی سه‌بعدی برای استخراج ویژگی‌های تصویر سه‌بعدی MRI مغز در قسمت اول استفاده می‌کند.

اکثر شبکه‌های طراحی شده برای تشخیص بیماری MS بر روی مجموعه‌ی کوچکی از تصاویر آموزش می‌بینند و به همین علت ممکن است تعمیم‌پذیری مناسب را نداشته باشند. به صورت ساده‌تر، شبکه ممکن است هنگام بررسی داده‌های متنوع‌تر دچار خطا شود و لذا استفاده‌ی بهتر و بهینه‌تر از پیوستگی تصاویر MRI مغز برای تشخیص این بیماری با توجه به کمبود داده تاثیرگذار خواهد بود.

[9] Convolutional Neural Network (CNN)
[10] Expanded Disability Status Scale (EDSS)
[11] Root Mean Square Error (RMSE)
[12] Optimizer
[13] Stochastic Gradient Descent (SGD)
[14] AdaDelta

۲- روش پیشنهادی برای حل چالش

۲-۱- چالش‌های مسئله

همان‌طور که پیش‌تر هم توضیح داده شد برای تشخیص قطعی بیماری MS استفاده از تصویر MRI مغز بسیار موثر است. در حقیقت تعداد و محل پلاک‌های ایجادشده در مغز پس از تزریق مواد رادیواکتیو می‌تواند به عنوان معیاری برای تشخیص این بیماری قرار گیرد. از آن‌جا که بررسی هرکدام از تصاویر MRI برای شناسایی پلاک‌ها و تعیین محل آن توسط پزشک بسیار زمان‌بر است، استفاده از روش‌های هوش‌مصنوعی می‌تواند مثل آن‌چه در بخش ۱-۲ شرح داده شد مؤثر باشد و در کنار پزشک معالج راهنمایی‌های مفیدی ارائه کند.

برای طراحی شبکه‌ای با قابلیت تشخیص پلاک‌ها و کلاس‌بندی آن‌ها به چهار ناحیه‌ی پری‌ونتریکولار، اینفراتنتوریال، ماده سفید و جاکستاکورتیکال باید به دو نکته‌ی زیر توجه داشت؛

- پیوستگی تصاویر MRI می‌تواند چه برای شناسایی پلاک‌ها و چه برای مشخص کردن نواحی رخداد آن‌ها موثر باشد. برای همین لازم است تا شبکه‌ی مورد نظر برای طراحی دارای حافظه باشد که بتواند رابطه تصاویر متوالی را استخراج کند.
- طبق توضیحات داده شده، تصاویر MRI برای این چالش دارای ابعاد ۲۳۶×۲۵۶×۱۶۰ هستند. برای استخراج ماسک خروجی تصاویر داده شده، استفاده از یک بعد به تنهایی کفایت می‌کند. به عنوان مثال می‌توان با بررسی ۲۳۶ تصویر با ابعاد ۲۵۶×۱۶۰ و ایجاد ماسک‌های مرتبط مسئله را حل کرد. ماسک سه‌بعدی نهایی هم از به‌هم چسباندن ۲۳۶ ماسک تولیدشده به‌دست می‌آید. با این حال استفاده از تصاویر هر سه بعد می‌تواند به دقت بیشتری منجر شود چرا که ممکن است هر زاویه اطلاعات بهتری از یک ناحیه ارائه کند.

در ادامه به بررسی ساختار شبکه‌ی طراحی‌شده برای حل این چالش می‌پردازیم و دقت را روی ۱۰ درصد از داده‌های آموزش محاسبه می‌کنیم.

۲-۲- ساختار شبکه طراحی شده

با توجه به داده‌های آموزش بهترین راهکار برای تشخیص پلاک‌های ایجادشده در تصویر MRI، استفاده از مدل‌های تقسیم‌بندی[15] به عنوان هسته‌ی اصلی شبکه خواهد بود. مدل‌های تقسیم‌بندی، تصویر را به بخش‌های متفاوت تشکیل‌دهنده اجسام یا صحنه تقسیم می‌کنند که هرکدام از نواحی به رنگ خاصی نمایش داده می‌شود. با این کار نواحی مختلف از یکدیگر متمایز و قابل تشخیص خواهند بود. اکثر مدل‌های تقسیم‌بندی از چندین لایه‌ی کانولوشنی عمیق برای استخراج ویژگی‌ها و تولید ماسک استفاده می‌کنند.

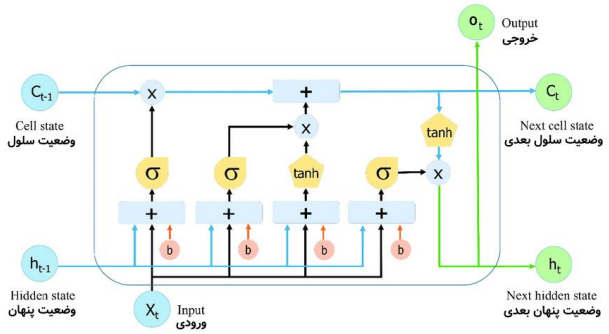
همان‌طور که قبل‌تر هم توضیح داده شد یکی از نکات اصلی برای حل این چالش، توجه به پیوستگی تصاویر MRI است. در نتیجه شبکه‌ی مورد نظر

[15] Segmentation

بهتر است دارای حافظه باشد و اطلاعات داده‌های متوالی را در خود ذخیره کند. بهترین راه‌حل، استفاده از شبکه عصبی بازگشتی[16] است. این شبکه با توجه به مسئله، مقداری از اطلاعات داده‌های پیشین را در واحدهای حافظه ذخیره می‌کند و هنگام استخراج ویژگی‌های فعلی از آن‌ها استفاده می‌کند. این شبکه‌ها، بسته به مسئله انتخابی می‌تواند اطلاعات یک یا چند داده پیشین را در خود ذخیره کند. از این شبکه برای پیش‌بینی سری‌های زمانی، داده‌های متنی، بورس و غیره استفاده می‌شود.

با توجه با توضیحات داده شده شبکه طراحی شده برای حل این چالش از دو هسته‌ی اصلی برای تشخیص پلاک‌های مغزی و تعیین ناحیه رخداد آن‌ها استفاده می‌کند. در هسته ابتدایی از یک شبکه‌ی تقسیم‌بندی استفاده شده است که ماسک اولیه را ایجاد می‌کند. این شبکه از ۵ لایه‌ی کانولوشنی متوالی استفاده می‌کند که هر لایه نسبت به لایه‌ی قبلی خود از ۲ برابر فیلتر بیشتر برای استخراج ویژگی‌های تصویر MRI استفاده می‌کند. در نتیجه بردار تعداد فیلترهای هر لایه برای شبکه طراحی‌شده به این صورت تعریف می‌شود؛ (۳۲, ۶۴, ۱۲۸, ۲۵۶, ۵۱۲) یعنی عمیق‌ترین لایه‌ی شبکه از ۵۱۲ فیلتر استفاده می‌کند و می‌تواند ویژگی‌های بیشتر و پیچیده‌تری استخراج کند. ورودی این شبکه در هر مرحله یک تصویر تک کاناله با ابعاد ۱۲۸×۱۲۸ پیکسل است. برای یافتن ویژگی‌های بهتر از دو لایه‌ی تماماً متصل استفاده شده است که در نهایت بردار خروجی توسط تابع فعال‌ساز سافت‌مَکس[17] تولید می‌شود و درایه‌ی i ام این بردار نشان‌دهنده احتمال رخداد پلاک تشخیص داده شده در ناحیه i ام است. در نتیجه ماسک تولیدشده در این مرحله علاوه بر نمایش پلاک‌ها، آن‌ها را به چهار کلاس مختلف تقسیم‌بندی می‌کند. با این‌حال خروجی کلاس‌بندی شبکه‌ی تقسیم‌بندی دقت بالایی ندارد و لازم است تا از شبکه‌ی دیگری برای بهود آن استفاده کرد.

برای رفع مشکل کلاس‌بندی، در مرحله‌ی بعدی از یک شبکه شبکه‌ی عصبی بازگشتی استفاده می‌کنیم تا با درنظر گرفتن ارتباط بین نواحی مختلف در تصاویر متوالی MRI، دقت کلاس‌بندی پلاک‌ها را افزایش دهد. به همین منظور از یک شبکه حافظه کوتاه‌مدت طولانی[18] که جزو شبکه‌های بازگشتی محسوب می‌شود استفاده می‌کنیم. هر واحد شبکه‌ی حافظه‌ی کوتاه‌مدت طولانی دارای سه دروازه اصلی است که می‌تواند براساس داده‌های ورودی و اطلاعات واحد قبلی، ویژگی‌های جدیدی استخراج کند. ساختار یک واحد شبکه حافظه‌ی کوتاه‌مدت طولانی در شکل ۱ نمایش داده شده است. همان‌طور که قابل مشاهده است داده‌های جدید وارد دروازه ورودی می‌شوند. اطلاعات واحد قبلی نیز از طریق دروازه وضعیت سلول و وضعیت پنهان وارد واحد فعلی می‌شوند که با استفاده از عملیات‌های مختلف و توابع فعال‌ساز، وضعیت سلول بعدی و وضعیت پنهان بعدی محاسبه می‌شود. به همین ترتیب می‌توان با افزایش تعداد واحدهای شبکه حافظه‌ی کوتاه‌مدت طولانی، حافظه‌ی مدل را افزایش داد تا هر بلوک از اطلاعات پیشین بیشتری برای تصمیم‌گیری و تولید خروجی استفاده کند.



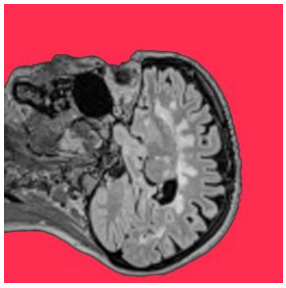
شکل ۱: یک واحد شبکه حافظه‌ی کوتاه‌مدت طولانی

شبکه‌ی طراحی شده برای حل این چالش نیز از یک شبکه‌ی حافظه‌ی کوتاه‌مدت طولانی استفاده می‌کند که در هر عمقی از شبکه‌ی تقسیم‌بندی وجود دارد و به تعداد نصف فیلترهای هر عمق از واحدهای حافظه استفاده می‌کند. به این صورت شبکه‌ی اصلی از ۵ هسته‌ی حافظه‌ی کوتاه‌مدت طولانی استفاده می‌کند که در هر عمق، ویژگی‌های مربوط به همان سطح را ذخیره می‌کند و برای تصاویر بعدی از آن‌ها بهره می‌گیرد. این‌کار باعث می‌شود شبکه حرکت تصاویر MRI مغز را با شروع از یک سمت به سمت دیگر درک کند و ۴ ناحیه مورد نظر را برای کلاس‌بندی پلاک‌ها بهتر بشناسد. تا اینجا در مورد هسته‌های اصلی شبکه بحث کردیم و مشخص شد که درون این شبکه از دو هسته‌ی تقسیم‌بندی و حافظه‌ی کوتاه‌مدت طولانی برای تشخیص پلاک‌ها و شناسایی محل رخداد آن‌ها استفاده کردیم. در ادامه اما در مورد تکنیک‌ها و سازوکارهای اضافی استفاده شده برای بهبود دقت خروجی توضیح می‌دهیم.

یکی از روش‌های به‌کار برده شده برای بهبود کارایی شبکه، استفاده از الگوریتم مقیاس‌دهی اِفیشِنت‌نِت‌بی۷[19] است. این الگوریتم از معماری شبکه عصبی کانولوشنی استفاده می‌کند و مدل را در تمام ابعاد مختلف آن یعنی عمق، عرض و وضوح گسترش می‌دهد. این‌کار باعث می‌شود شبکه بتواند جزئیات بیشتری را برای حل مسئله درنظر بگیرد و دقت خروجی بهبود پیدا کند با این حال مشخص است که با افزایش ابعاد مدل، پیچیدگی محاسباتی شبکه نیز بسیار بیشتر می‌شود و زمان بیشتری برای آموزش یا آزمایش مدل نیاز است. از طرف دیگر برای کاهش حجم محاسباتی از سازوکار تمرکز[20] بر نواحی مهم بهره گرفته شده است. از آنجایی که هسته‌های اصلی شبکه، محاسبات را در سطح پیکسل انجام می‌دهند و با توجه به اینکه در تصاویر MRI نواحی خارج از جمجه و کم‌اهمیت بسیار است، مشخص است پردازش‌هایی اضافی وجود دارد که پیچیدگی محاسباتی شبکه را بی‌جهت افزایش می‌دهد. برای مثال نواحی کم‌اهمیت یک تصویر MRI مغزی در شکل ۱ به رنگ قرمز نمایش داده شده است. به همین علت این سازوکار می‌تواند بدون کاهش دقت شبکه و یا حذف جزئیات مهم، محاسبات کلی شبکه را کاهش دهد. نحوه‌ی عملکرد به این صورت است که عملیات‌های کانولوشنی و غیره را صرفا در نواحی مهم انجام می‌دهد و خروجی بقیه نواحی

[16] Recurrent Neural Network (RNN)
[17] Softmax
[18] Long Short-Term Memory
[19] EfficientNetB7
[20] Attention Mechanism

را صفر درنظر می‌گیرد. شناسایی نواحی مهم نیز به تدریج انجام می‌شود؛ یعنی در ابتدا کل تصویر به عنوان نواحی مهم درنظر گرفته می‌شود و کم‌کم این ناحیه تا جایی کوچکتر می‌شود که تاثیری بر دقت خروجی نداشته باشد. این روش برای شبکه‌ی عصبی عمیق علاوه بر کاهش محاسبات، باعث جلوگیری از بیش‌برازش[21] هم می‌شود. در حقیقت زمانی که بیش‌برازش رخ می‌دهد مدل به جای یادگیری شروع به حفظ کردن می‌کند و در نتیجه هنگام مواجهه با داده‌ی متنوع دقت بسیار پایینی خواهد داشت.



شکل ۲: نواحی کم اهمیت تصویر MRI مغز که به رنگ قرمز است

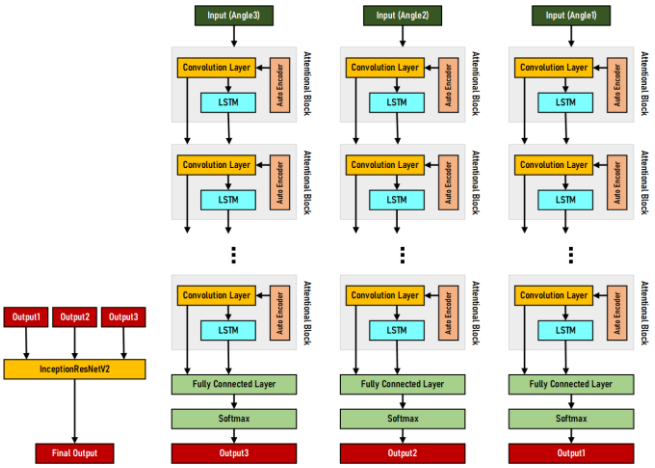
تکنیک دیگری که برای بهبود شبکه استفاده کردیم استفاده از شبکه کدگذار خودکار[22] با قابلیت استخراج ۴۰۹۶ ویژگی متفاوت است. این شبکه از کدگذار و کدبردار تشکیل می‌شود که تصویر ورودی توسط کدگذار به یک فضای پنهان[23] نگاشت و یا به تعبیر دیگری کد می‌شود. این کدگذاری به صورتی است که ویژگی‌های مهم تصویر ورودی حفظ می‌شود تا در هنگام کدبرداری، تصویر اصلی با همان جزئیات اولیه قابل بازسازی باشد. لذا یکی دیگر از روش‌های استخراج ویژگی‌های تصویر MRI مغز، آموزش شبکه‌ی کدگذار خودکاری است که بتواند تصاویر را با کیفیت و جزئیات زیادی بازسازی کند. در نتیجه ماتریس ایجاد شده توسط کدگذار می‌تواند در صورتی که خروجی کدبردار بسیار شبیه به ورودی باشد به عنوان معیاری برای شبکه‌ی تقسیم‌بندی و تولید ماسک درنظر گرفته شود.

در صورت کمبود داده برای آموزش شبکه‌ی اصلی، می‌توان از کدگذار خودکار متغیر و یا شبکه‌ی مولد متخاصم[24] برای تولید داده‌های بیشتر استفاده کرد. به دلیل افزایش بسیار کم دقت خروجی هنگام استفاده از شبکه GAN برای تولید داده‌های بیشتر، از تکنیک افزایش داده صرف نظر کردیم.

روش دیگری که برای کلاس‌بندی بهتر پلاک‌ها به‌کار برده‌ایم استفاده از تکنیک کدگذاری وان هات[25] است. همان‌طور که در توضیحات این چالش آمده است، مقدار پیکسل‌های تصاویر ماسک، می‌تواند در بازه‌ی ۰ تا ۴ قرار بگیرد که اعداد ۱ تا ۴ نشان‌دهنده‌ی چهار ناحیه‌ی مغز و عدد ۰ نشان‌دهنده نواحی بی‌اهمیت است. همان‌طور که قبلاً توضیح داده شد، ورودی شبکه‌ی اصلی، تصویر با ابعاد ۱۲۸×۱۲۸ پیکسل است و این یعنی تصاویر ماسک داده شده نیاز به تغییر اندازه دارند چون تصاویر سه‌بعدی ماسک داده شده ابعادی برابر با ابعاد تصاویر سه‌بعدی MRI مغز دارند. تغییر ابعاد تصویر با استفاده از روش‌های معمولی باعث ایجاد مقادیر جدید و یا مقداردهی اشتباه می‌شود. استفاده از روش OHE می‌تواند در کاهش این خطا مؤثر باشد. در این روش ابتدا یک بردار به طول مقادیر مختلف در تصاویر ماسک درنظر می‌گیریم. در این مسئله لازم است تا با توجه به مقادیر ۰ تا ۴ در تصاویر ماسک، برداری به طول ۵ در نظر بگیریم. در مرحله‌ی بعدی به هرکدام از مقادیر مختلف یک بردار یکتا نسبت می‌دهیم به نحوی که تنها از اعداد ۰ و ۱ تشکیل شده باشد و همچنین تنها یک درایه‌ی ۱ در هر بردار وجود داشته باشد. برای مثال می‌توان به صورت زیر عمل کرد.

0 = [1, 0, 0, 0, 0]
1 = [0, 1, 0, 0, 0]
2 = [0, 0, 1, 0, 0]
3 = [0, 0, 0, 1, 0]
4 = [0, 0, 0, 0, 1]

برای تولید ماسک نهایی، استفاده از تصاویر MRI در یک زاویه کافی است با این حال با توجه به اینکه هر زاویه اطلاعات بهتری از یک ناحیه ارائه می‌دهد، تعداد پلاک‌های شناسایی شده در هر زاویه ممکن است متفاوت باشد. به همین علت تصاویر ماسک تولیدشده برای هر زاویه وارد شبکه‌ی عصبی دیگری می‌شوند تا براساس خطای کلاس‌بندی، ماسک نهایی با ادغام بهینه خروجی هر زاویه به‌دست آید. در نهایت، ساختار نهایی شبکه طراحی شده به‌صورت شکل ۳ است.



شکل ۳ : معماری کلی شبکه‌ی طراحی شده

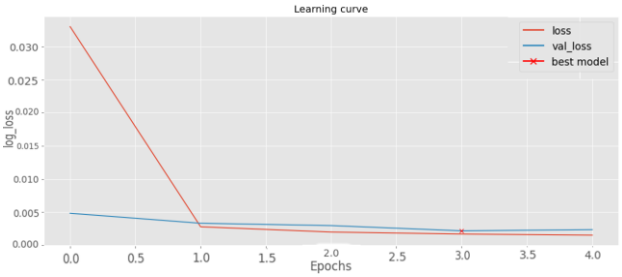
۲-۳- آموزش شبکه

در قسمت قبلی ساختار کلی شبکه‌ی طراحی‌شده برای حل چالش مسابقه توضیح داده شد. برای آموزش شبکه‌ی طراحی‌شده، ۵۵ تصویر سه‌بعدی MRI مغز در اختیار داریم. در مرحله‌ی اول صرفاً از ۵۰ داده برای آموزش استفاده کردیم و ۵ داده‌ی دیگر را برای آزمودن مدل درنظر گرفتیم. تصاویر مغز برای هر داده در هر زاویه به‌صورت جداگانه استخراج می‌شود و تصاویر با زوایای یکسان به ترتیب توالی صحیح در یک پوشه قرار می‌گیرند. در نتیجه شش پوشه برای داده‌های آموزش داریم که سه پوشه برای تصاویر مغز و سه پوشه برای تصاویر ماسک درنظر گرفته شده است. در نتیجه در هر پوشه،

[21] Overfitting
[22] AutoEncoder
[23] Latent space
[24] Generative Adversarial Network (GAN)
[25] One Hot Encoding (OHE)

تصاویر مربوط به یک زاویه از مغز یا ماسک افراد مختلف قرار دارد. در مرحله‌ی بعدی با استفاده از روش‌های گفته شده تصاویر موجود در هر پوشه را به ابعاد ۱۲۸×۱۲۸ پیکسل تغییر می‌دهیم و وارد شبکه می‌کنیم. داده‌های ورودی به دو دسته‌ی آموزش و صحت‌سنجی[۲۶] تقسیم می‌شوند که به ترتیب شامل ۹۰ و ۱۰ درصد از کل ۵۰ داده می‌شوند. به این ترتیب ۴۵ داده برای آموزش و ۵ داده برای صحت‌سنجی استفاده می‌شود. شبکه‌ی طراحی‌شده در ۲۰ ایپاک[۲۷]، اندازه دسته[۲۸] برابر ۴ و برای هر زاویه به صورت جداگانه آموزش داده می‌شود. الگوریتم آدام[۲۹] با نرخ یادگیری[۳۰] ۰.۰۰۰۵ به عنوان بهینه‌ساز درنظر گرفته شده است که سعی می‌کند مقدار تابع ضرر[۳۱] را کمینه کند. مقدار آنتروپی متقابل کلاس‌بندی شده[۳۲] به عنوان معیاری برای محاسبه مقدار ضرر پس از پردازش هر دسته استفاده می‌شود. دقت خروجی و میزان همگرایی کولبک-لایبلر[۳۳] نیز به عنوان معیارهای شبکه درنظر گرفته شده‌اند تا مدل به سمتی حرکت کند که این دو نیز بهینه شوند.

نمودار یادگیری برای داده‌های آموزش و صحت‌سنجی و برای ۵ ایپاک در شکل ۴ رسم شده است. همان‌طور که واضح است شبکه در همان چهار ایپاک ابتدایی به دقت بالا و ضرر پایینی رسیده است.



شکل ۴: نمودار یادگیری پس از آموزش شبکه برای ۵ ایپاک

در ادامه این‌کار را برای داده‌های دو زاویه‌ی دیگر نیز انجام می‌دهیم. در نتیجه پس از اتمام مرحله‌ی آموزش سه ماتریس وزن داریم که از هرکدام برای پیش‌بینی و تولید ماسک‌های همان زاویه استفاده می‌کنیم. صرف نظر از اینکه این‌کار ممکن است کمی زمان‌بر باشد، نتیجه‌ی نهایی از ادغام سه ماسک تولید شده در زوایای مختلف به‌دست می‌آید. واضح است استفاده از ماسک‌های به‌دست‌آمده در یک زاویه برای تولید تصویر سه‌بعدی نهایی کفایت می‌کند با این حال از آن‌جایی که هر زاویه دید و اطلاعات بهتر و بیشتری از یک ناحیه ارائه می‌کند، استفاده از ماسک‌های به‌دست‌آمده در هر سه زاویه می‌تواند دقت خروجی را افزایش دهد.

همان‌طور که پیش‌تر هم توضیح داده شد مرحله‌ی آخر برای حل این چالش یکی‌کردن تصاویر به‌دست آمده در هر سه زاویه است. با این حال این فرایند چالش‌هایی دارد؛ در وهله‌ی اول باید توجه داشت که احتمالاً در سه تصویر، ناحیه‌های یکسانی وجود دارد که کلاس متفاوتی دارند. برای مثال ممکن است پیکسل ردیف ۱۰۰ام و ستون ۱۰۰ام در یکی از تصاویر به کلاس ۲ و در دو تصویر دیگر به کلاس ۱ تعلق داشته باشد. هرچند تعداد چنین پیکسل‌هایی کم است و بیشتر پیکسل‌ها با یکدیگر همخوانی دارند، با این حال موجب می‌شود یکی‌کردن سه تصویر با مشکلاتی همراه باشد. برای تصمیم‌گیری در مورد اینکه مقدار پیکسل متناظر کدام‌یک از سه تصویر ارجحیت بیشتری دارد از شبکه اینسپشن‌رزنت‌وی۲[۳۴] استفاده کردیم. این شبکه جزو شبکه‌های عصبی عمیق محسوب می‌شود که با ۱۶۴ لایه می‌تواند کلاس‌بندی ۱۰۰۰ کلاس را انجام دهد. در این چالش، شبکه InceptionResNetV2 به گونه‌ای آموزش می‌بینند که هنگام یکی‌کردن سه تصویر علاوه بر بهبود دقت تشخیص پلاک‌ها، مقدار کلاس‌بندی اشتباه[۳۵] به حداقل مقدار خود برسد. در نهایت شبکه با مقادیر نوشته شده در جدول ۱ آموزش را به پایان می‌رساند که بهترین مدل در ایپاک سوم به‌دست آمده است.

جدول ۱: مقادیر معیارهای مدل پس از آموزش به ترتیب افزایش val_loss

	loss	mean_squared_error	accuracy	precision	recall	kullback_leibler_divergence	val_loss
3	0.001663	0.000143	0.999542	0.999574	0.999517	0.001661	0.002150
4	0.001494	0.000136	0.999563	0.999589	0.999542	0.001492	0.002325
2	0.001965	0.000153	0.999514	0.999548	0.999489	0.001963	0.002941
1	0.002760	0.000174	0.999456	0.999487	0.999432	0.002757	0.003293
0	0.033056	0.001845	0.997457	0.999072	0.992627	0.033052	0.004808

۳- نتایج شبکه روی داده‌های موجود و برخی اصلاح‌ها

پس از آموزش شبکه و تولید وزن‌ها طبق آن‌چه در بخش ۲-۳ گفته شد در مرحله‌ی بعدی مدل آموزش دیده را با استفاده از ۵ داده‌ی دیگر تست می‌کنیم. البته ذکر این نکته ضروری است که نمودارهای رسم شده در قسمت قبلی به ازای ۵ ایپاک به‌دست آمده است و این درحالی است که شبکه‌ی نهایی برای حل چالش مسابقه طی ۲۰ ایپاک آموزش می‌بینند. نتایج پیش‌بینی مدل در تشخیص و کلاس‌بندی پلاک‌ها برحسب معیارهای اصلی در جدول ۲ آورده شده است. با توجه به این نتایج، شبکه توانسته است این چالش را با دقت خوبی حل کند.

جدول ۲: نتایج نهایی برحسب معیارهای مدل روی داده‌های تست

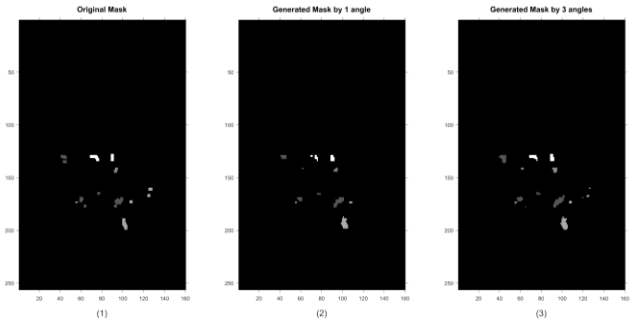
	Evaluate_Result
loss	0.000335
mean_squared_error	0.000034
accuracy	0.999887
precision	0.999887
recall	0.999887
kullback_leibler_divergence	0.000331

در ادامه نیز ماسک‌های تولیدشده در هر سه زاویه با استفاده از شبکه‌ی عصبی InceptionResNetV2 ادغام می‌شوند. در شکل ۵ یک نمونه از این فرایند نشان داده شده است. مشخص است تصویر ادغام شده دقت بالاتری چه در

[26] Validation
[27] Epoch
[28] Batch size
[29] ADAM
[30] Learning Rate (LR)
[31] Loss function
[32] Categorical-CrossEntropy
[33] Kullback–Leibler divergence (KL divergence)

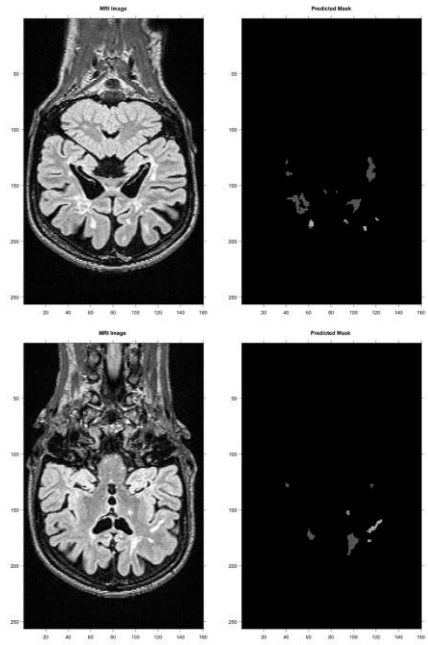
[34] InceptionResNetV2
[35] Misclassification

تشخیص پلاک‌ها و چه در کلاس‌بندی آن‌ها دارد. هرچند استفاده از این روش کمی زمان‌بر است با این‌حال با توجه به نتیجه‌ی خروجی استفاده از آن باعث بهبود شبکه شده است.



شکل ۵: (۱) تصویر ماسک اصلی، (۲) تصویر ماسک تولیدشده با یک زاویه، (۳) تصویر ماسک تولیدشده با ادغام نتایج هر سه زاویه

تا به اینجا دقت شبکه‌ی طراحی‌شده بررسی شد. نتایج نشان می‌دهد که مدل طراحی شده می‌تواند با دقت خوبی پلاک‌های مغزی را تشخیص دهد و آن‌ها را به ۴ ناحیه کلاس‌بندی کند. در نتیجه در مرحله‌ی نهایی برای تولید ماسک تصاویر تست داده شده که شامل ۲۰ تصویر سه‌بعدی MRI مغزی است، شبکه را با تمام داده‌های آموزشی داده شده که شامل ۵۵ داده است آموزش می‌دهیم. در مرحله‌ی قبل ۵ داده از کل ۵۵ داده را برای تست شبکه درنظر گرفته بودیم که با توجه نتایج مطلوب شبکه، نیازی به داده‌های تست نداریم و بهتر است که در مرحله نهایی برای بهترشدن شبکه از تمام ۵۵ داده برای آموزش استفاده کنیم. چند نمونه از تصویر MRI مغز از یک زاویه به همراه ماسک تولیدشده‌ی آن با استفاده از شبکه‌ی طراحی شده در شکل ۶ نمایش داده شده است.



شکل ۶: نمونه ماسک‌های تولیدشده برای داده‌های تست

۴- نتیجه‌گیری

شبکه‌ی عصبی طراحی شده برای تشخیص بیماری MS از روی تصاویر MRI با استفاده از دو هسته‌ی مرکزی برای پردازش تصاویر دقت خوبی دارد. این شبکه با استفاده از دو شبکه‌ی تقسیم‌بندی و حافظه‌ی کوتاه‌مدت طولانی می‌تواند پلاک‌های موجود در فاز T1 تصاویر MRI را شناسایی و با توجه به پیوستگی عکس‌برداری MRI، پلاک‌های تشخیص داده شده را در چهار ناحیه‌ی پری‌ونتریکولار، اینفراتنتوریال، ماده سفید و جاکستاکورتیکال کلاس‌بندی کند. دقت این شبکه روی داده‌های آموزش بالای ۹۹ درصد و خطای شبکه زیر ۰.۰۰۲ است که نشان می‌دهد شبکه توانسته است در حل این چالش نتیجه خوبی داشته باشد.

منابع

[1] Marzullo, A., Kocevar, G., Stamile, C., Calimeri, F., Terracina, G., Durand-Dubief, F., & Sappey-Marinier, D. (2019, July). Prediction of Multiple Sclerosis Patient Disability from Structural Connectivity using Convolutional Neural Networks. In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 2087-2090). IEEE.A. Author 1 and B. Author 2, "Title of the conference paper," *Proc. Int. Conf. on Power System Reliability*. Singapore, pp. 100-105, 1999.

[2] Siar, H., & Teshnehlab, M. (2019, January). Diagnosing and classification tumors and MS simultaneous of magnetic resonance images using convolution neural network. In 2019 7th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS) (pp. 1-4). IEEE.

[3] Shrwan, R., & Gupta, A. (2021). Classification of Pituitary Tumor and Multiple Sclerosis Brain Lesions through Convolutional Neural Networks. In IOP Conference Series: Materials Science and Engineering (Vol. 1049, No. 1, p. 012014). IOP Publishing.

[4] Yoo, Y., Tang, L. W., Brosch, T., Li, D. K., Metz, L., Traboulsee, A., & Tam, R. (2016). Deep learning of brain lesion patterns for predicting future disease activity in patients with early symptoms of multiple sclerosis. In Deep Learning and Data Labeling for Medical Applications (pp. 86-94). Springer, Cham.

[5] Roy, S., Butman, J. A., Reich, D. S., Calabresi, P. A., & Pham, D. L. (2018). Multiple sclerosis lesion segmentation from brain MRI via fully convolutional neural networks. arXiv preprint arXiv:1803.09172.

[6] Valverde, S., Cabezas, M., Roura, E., González-Villà, S., Pareto, D., Vilanova, J. C., ... & Lladó, X. (2017). Improving automated multiple sclerosis lesion segmentation with a cascaded 3D convolutional neural network approach. NeuroImage, 155, 159-168.