

یک چارچوب یادگیری عمیق بهبودیافته مبتنی بر VGG برای طبقه‌بندی گلیوما در تصاویر تشدید مغناطیسی مغز

آسیه خسروانیان^۱

مقاله پژوهشی

چکیده

مقدمه: تشخیص زودهنگام تومورهای مغزی، می‌تواند نقش مهمی در برنامه‌ریزی‌های درمان داشته باشد. در تصاویر تشدید مغناطیسی مغز، شباهت‌های زیادی بین بافت سالم و سرطانی وجود دارد، لذا تشخیص دقیق تومورها به یک چالش مهم تبدیل شده است. همچنین طبقه‌بندی دستی این تصاویر، زمان‌بر و مستعد خطای انسانی است؛ لذا توسعه روش‌های طبقه‌بندی خودکار می‌تواند کمک مؤثری به پزشکان ارایه دهد.

روش بررسی: در این پژوهش، معماری جدیدی از شبکه عصبی عمیق برای طبقه‌بندی خودکار تومور مغزی در تصاویر تشدید مغناطیسی پیشنهاد شده است. در روش پیشنهادی، معماری شبکه عصبی عمیق VGG-16 بازنطراحی شده است. به منظور حفظ رزولوشن مکانی ویژگی‌ها در عمق شبکه و تشخیص نواحی کوچک تومورها و شناسایی مرزهای مبهم، لایه Max-pooling در معماری VGG-16 حذف گردید.

یافته‌ها: روش پیشنهادی با استفاده از مجموعه داده BraTS2020 با روش ۱۰-fold cross validation ارزیابی بر اساس معیارهای Precision، Accuracy، Recall، F1 Score و AUC نشان داد که معماری پیشنهادی نسبت به نسخه اصلی شبکه عصبی VGG-16، عملکرد بهتری دارد و توانسته در معیارهای ذکر شده به ترتیب به مقادیر ۰/۰۳۵±۰/۰۱۲، ۰/۰۲۶±۰/۰۰۹۷۳۹، ۰/۰۴۴±۰/۰۰۹۵۲۴، ۰/۰۵۱±۰/۰۰۹۴۹۰، ۰/۰۲۷±۰/۰۰۹۶۳۰، ۰/۰۴۴±۰/۰۰۹۵۱۴ دست یابد.

نتیجه‌گیری: نتایج حاصل از روش پیشنهادی در این پژوهش، تایید می‌کند که حفظ اطلاعات مکانی و شناسایی مرزهای مبهم می‌تواند نقش مهمی در افزایش دقت طبقه‌بندی داشته باشد. به علاوه تغییرات اعمال شده منجر به کاهش نمونه‌برداری بیش از حد (Over-Downsampling) شده است.

واژه‌های کلیدی: طبقه‌بندی؛ یادگیری عمیق؛ تصویربرداری تشدید مغناطیسی؛ شبکه عصبی عمیق.

پیام کلیدی: استفاده از روش پیشنهادی در این پژوهش، به پزشکان کمک می‌کند تا تومورهای مغزی را سریع‌تر و دقیق‌تر در تصاویر MRI شناسایی کنند. این مدل هوشمند با حفظ جزئیات مکانی و تشخیص مرزهای مبهم بین بافت سالم و سرطانی، خطای انسانی را کاهش داده و می‌تواند به تصمیم‌گیری بهتر درمانی و برنامه‌ریزی دقیق‌تر جراحی‌ها کمک کند. به کارگیری این روش در مراکز تصویربرداری پزشکی، موجب صرفه‌جویی در زمان و افزایش دقت تشخیص در بیماران مبتلا به تومور مغزی می‌شود.

دریافت مقاله: ۱۴۰۴/۷/۲۱

پذیرش مقاله: ۱۴۰۴/۹/۳۰

تاریخ انتشار: ۱۴۰۴/۱۰/۱۰

ارجاع: خسروانیان آسیه. یک چارچوب یادگیری عمیق بهبودیافته مبتنی بر VGG برای طبقه‌بندی گلیوما در تصاویر تشدید مغناطیسی مغز. مدیریت اطلاعات سلامت ۲۲:۱۴۰۴ (۴): ۱۹۹-۲۱۴.

مقدمه

تومورهای مغزی انواع مختلفی دارند که یکی از شایع‌ترین آن‌ها تومورهای گلیوما است. تومورهای گلیوما به دو دسته گلیوما با درجه پایین (LGG: Low-Grade Glioma) و گلیوما با درجه بالا (HGG: High-Grade Glioma) تقسیم می‌شوند (۱). برای تشخیص آسیب‌های ایجاد شده در بافت مغز از تصویربرداری تشدید مغناطیسی استفاده می‌شود. این تصاویر به سبب فیزیک تصویربرداری، نویز یا آرتیفکت حرکتی بیمار، مستعد ناهمگنی شدت روشنایی هستند. این پدیده با کاهش وضوح تصویر کار تشخیص را برای پزشکان مشکل می‌کند. لذا در سال‌های اخیر محققان هوش مصنوعی، به گسترش روش‌های خودکار و نیمه‌خودکار بر پایه یادگیری عمیق برای تشخیص انواع آسیب‌های مغزی در این تصاویر پرداختند.

در مسایل طبقه‌بندی تومورهای مغزی، رویکردهای یادگیری ماشین سنتی و یادگیری عمیق وجود دارد. انتخاب ویژگی مرحله حساسی در رویکردهای یادگیری ماشین سنتی است (۲) در مقابل روش‌های یادگیری عمیق وجود دارد که استخراج ویژگی را به صورت خودکار انجام می‌دهند و امروزه مورد توجه محققان قرار گرفته است.

در سال‌های اخیر شبکه‌های مبتنی بر CNN متعددی برای طبقه‌بندی تصاویر مغزی به کار رفته‌اند (۳-۵). در مطالعه Ganesh و همکاران، رویکردی ترکیبی برای بخش‌بندی تومورهای مغزی پیشنهاد شد. آن‌ها از شبکه عصبی VGG-16 برای استخراج ویژگی استفاده کردند (۶). معماری yolo-v5 در پژوهش Saranya و همکاران عملکرد قابل توجهی در دسته‌بندی چند کلاسه تومور مغزی نشان داد (۷). Graph Convolutional Gated Recurrent Network (GCGRN) توسط Chanemougavel و همکاران معرفی شد که دقت مناسبی در طبقه‌بندی تومور ارایه داد (۸). Nancy و همکاران، مدلی مبتنی بر شبکه عصبی VGG همراه با مکانیزم‌های توجه و ماژول‌های Model Agnostic Concept Extractor (MACE) پیشنهاد دادند. تنظیم دقیق وزن‌های انتقالی و پیچیدگی محاسباتی ماژول توجه، از محدودیت‌های این روش بود (۹).

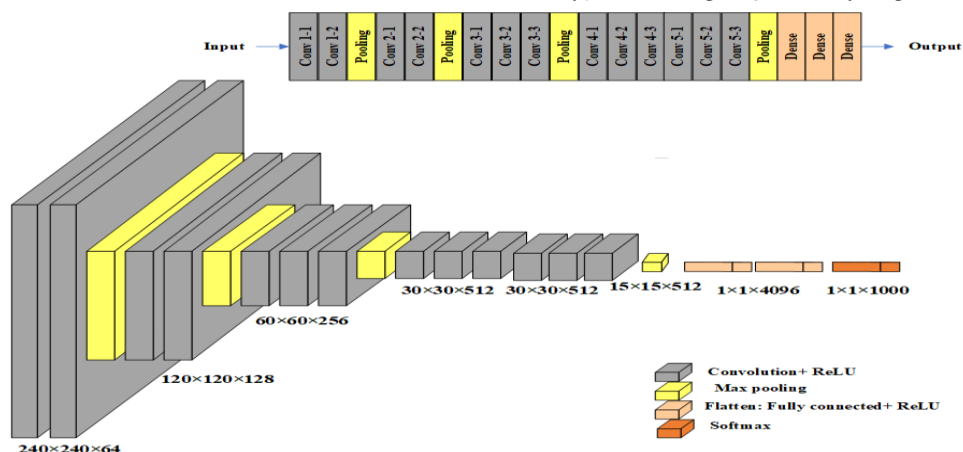
۱. استادیار، گروه مهندسی کامپیوتر، دانشکده فنی و مهندسی، مجتمع آموزش عالی لارستان، لارستان، ایران.

نویسنده طرف مکاتبه: آسیه خسروانیان، استادیار، گروه مهندسی کامپیوتر، دانشکده فنی و مهندسی، مجتمع آموزش عالی لارستان، لارستان، ایران
Email: khosrovanian@lar.ac.ir

به ۳۶۹ بیمار است که از این تعداد ۷۶ بیمار دارای گلیوما درجه پایین و ۲۹۳ بیمار دارای گلیوما درجه بالاست. تصاویر مرجع در این مجموعه داده، توسط ۴ رادیولوژیست بررسی شده و ماسک نهایی به عنوان تصویر مرجع برای هر بیمار در مجموعه داده قرار گرفته است.

شبکه عصبی عمیق VGG: شبکه عصبی عمیق VGG-16 یکی از معماری‌هایی است که در حوزه دسته‌بندی تصاویر، موفق عمل کرده است. این معماری در سال ۲۰۱۴ معرفی شد (۱۹) و ایده آن، استفاده از لایه‌های کانولوشنی با اندازه کرنل ۳ و افزایش تدریجی تعداد کانال‌ها در عمق شبکه است. این شبکه دارای ۱۳ لایه کانولوشنی و ۳ لایه تمام متصل است که به دلیل وجود ۱۶ لایه در آن، VGG-16 نامیده شده است. ساختار این شبکه دارای ۵ بلوک است که هر بلوک شامل لایه‌های کانولوشنی با تابع فعال‌ساز ReLU است که پس از آن، لایه Max-pooling برای کاهش ابعاد مکانی قرار می‌گیرد. ویژگی‌های استخراج شده در طول لایه‌های شبکه، به سه لایه تمام متصل ارسال می‌گردد. در لایه خروجی Softmax قرار داده شده است که می‌تواند توزیع احتمال کلاس‌ها را تخمین بزند. در این پژوهش معماری VGG 16 به‌عنوان مدل پایه انتخاب شد. دلیل این انتخاب، ساختار ساده، ماژولار و در عین حال قدرتمند این معماری است که آن را به یکی از شبکه‌های مرجع در بسیاری از مطالعات بینایی ماشین و تحلیل تصاویر پزشکی تبدیل کرده است.

در این پژوهش، معماری اصلی شبکه عصبی VGG-16، با هدف بهبود عملکرد طبقه‌بندی در تصاویر پیچیده تشدید مغناطیسی تغییر داده شده است. به‌منظور پیشگیری از کاهش ابعاد و حفظ جزئیات و ساختارهای ظریف در تصویر، لایه Max-pooling موجود در بلوک چهارم حذف گردیده است. از آنجا که این لایه فاصله کمی تا لایه‌های تمام متصل برای تصمیم‌گیری شبکه دارد، حذف این لایه با هدف کاهش خطر بیش نمونه‌برداری و حفظ وضوح تصویر در مرزهای مبهم انجام شده است. سایر بخش‌های معماری شبکه بدون تغییر باقی‌مانده است تا اثر این حذف انتخابی، ارزیابی گردد. شکل ۱ معماری تغییر یافته این شبکه را نمایش می‌دهد.



شکل ۱: معماری پیشنهادی شبکه عصبی عمیق VGG-16

یک رویکرد دو مرحله‌ای توسط Shedbalkar و همکاران پیشنهاد شد. در این پژوهش، ابتدا با استفاده از شبکه عصبی U-Net تصاویر مغزی ناحیه‌بندی شدند و خروجی این مرحله برای دسته‌بندی به شبکه VGG داده شد (۱۰). در پژوهش Gayathri و همکاران، نیز یک رویکرد ترکیبی معرفی شد که برای بخش‌بندی از شبکه عصبی CNN و برای طبقه‌بندی از شبکه VGG استفاده شد (۱۱). در پژوهش Pokhriyal و همکاران، یک رویکرد مبتنی بر VGG پیشنهاد شد که از یادگیری انتقالی برای آموزش شبکه استفاده می‌کرد (۱۲). مطالعات متعددی نشان دادند که لایه Max-pooling در ساختار شبکه‌های عصبی کانولوشنی، اطلاعات با وضوح بالا را از بین می‌برد و می‌تواند منجر به شبه نمونه‌برداری شود که دقت شناسایی ساختارهای ظریف، مانند مرزهای مبهم را کاهش می‌دهد (۱۳-۱۵). مرور پیشینه پژوهش نشان می‌دهد که محققان برای غلبه بر این چالش، سعی کردند این لایه را با pooling تطبیقی یا راهکارهای دیگری جایگزین کنند. درحالی که تاثیر مستقیم حذف این لایه بدون جایگزینی در معماری‌های عمیق مانند VGG-16 برای پردازش تصاویر پیچیده، کمتر مورد بررسی قرار گرفته است.

هدف اصلی این پژوهش، طراحی معماری تغییر یافته VGG-16 و ارزیابی تاثیر این تغییر ساختاری، بر عملکرد شبکه عصبی عمیق در طبقه‌بندی تومورهای نوع گلیوما است تا بتواند زمینه‌ساز ایجاد معماری‌هایی با حفظ جزئیات مکانی برای پردازش تصاویر پزشکی باشد.

روش بررسی

مجموعه داده: در این پژوهش، از مجموعه داده (۱۶-۱۸) BraTS 2020 استفاده می‌شود. این مجموعه داده، شامل اسکن‌های سه بعدی از تصاویر تشدید مغناطیسی است که شامل تومورهای نوع گلیوما درجه بالا و درجه پایین است. این تصاویر در چهار مدالیته مختلف T1، T2، FLAIR و T1c از ۱۹ مرکز درمانی با استفاده از اسکنرهای مختلف به دست آمده است. اندازه هر تصویر ۱۵۵×۲۴۰×۲۴۰ و کسل با وضوح ۱×۱ میلی متر است. این مجموعه داده شامل اسکن‌های سه‌بعدی مربوط

همان‌طور که در شکل ۱ نشان داده شده است این شبکه مانند معماری پایه، شامل ۵ بلوک کانولوشنی، لایه‌های تماماً متصل و لایه‌هایی Softmax است با این تفاوت که در بلوک چهارم لایه Max-pooling حذف شده است. این حذف باعث می‌شود که نقشه ویژگی بعد از بلوک چهارم تغییر اندازه نداشته باشد و مستقیماً به لایه‌های بعدی متصل شود. استفاده از Max-pooling در بلوک‌های اول تا سوم، در حذف نویز و کاهش هزینه محاسباتی نقش مهمی دارند؛ اما در بلوک چهارم، این لایه سبب کاهش کیفیت ویژگی‌ها به ویژه در تصاویر حساس تشدید مغناطیسی می‌شود. در واقع، عملیات pooling در این بلوک، بخش زیادی از فرکانس‌های بالای تصویر را حذف می‌کند و این درست همان جایی است که برای تشخیص در تصاویر پیچیده با مرزهای مبهم دارای اهمیت است. آنجا که Max-Pooling فاقد پارامتر است، حذف آن هیچ اثری بر تعداد پارامترهای کل شبکه ندارد. در معماری پایه، عملیات pooling اندازه نقشه ویژگی را از $512 \times 30 \times 30$ به $128 \times 15 \times 15$ کاهش می‌دهد؛ در حالی که در مدل پیشنهادی این مقدار بدون کاهش به بلوک پنجم منتقل می‌شود. با توجه به اینکه هزینه محاسباتی کانولوشن متناسب با مساحت نقشه ویژگی است، حذف pooling منجر به چهار برابر شدن FLOPs بلوک پنجم و در نتیجه افزایش تقریباً ۷۰ تا ۷۵ درصدی هزینه محاسباتی کل مدل و زمان آموزش می‌شود. این افزایش هزینه به‌صورت هدفمند پذیرفته شده است، چرا که حذف pooling در این مرحله موجب حفظ ویژگی‌های با فرکانس بالاتر و جلوگیری از، از دست رفتن اطلاعات ظریف ساختارهای توموری در تصاویر MRI می‌شود.

در معماری پیشنهادی، در بلوک‌های اول تا سوم به ترتیب ۶۴، ۱۲۸ و ۲۵۶ فیلتر استفاده شده است و برای بلوک چهارم از ۵۱۲ فیلتر استفاده شده است. از آنجا که برای آموزش شبکه عصبی عمیق داده‌های زیادی مورد نیاز است، از برش‌های دوبعدی استفاده شده است به این منظور تعداد ۱۵۳۰۰ تصویر دوبعدی از مدالیته FLAIR از مجموعه داده انتخاب شده است. لازم به ذکر است که این تصاویر دوبعدی از اسکن‌های سه‌بعدی بیماران موجود در مجموعه داده BraTS استخراج شده‌اند و هر اسکن سه‌بعدی شامل ۱۵۵ برش دوبعدی می‌باشد. دلیل اصلی انتخاب رویکرد مبتنی بر برش‌های دوبعدی، کاهش پیچیدگی محاسباتی مدل، کاهش تعداد پارامترهای قابل آموزش و امکان آموزش شبکه در زمان مناسب با منابع سخت‌افزاری در دسترس بوده است.

مجموعه داده BraTS پیش از انتشار، توسط برگزارکنندگان چالش، تحت مراحل استاندارد پیش‌پردازش شامل هم‌ثبت‌سازی (Co-registration) مدالیته‌ها، حذف جمجمه و همسان‌سازی فضایی قرار گرفته است. در این پژوهش نیز پیش از ورود تصاویر به شبکه، مقادیر شدت پیکسل‌ها نرمال‌سازی شد تا پایداری فرآیند آموزش و همگنی داده‌های ورودی افزایش یابد.

هدف از این آموزش، دسته‌بندی تومورهای مغزی به دو گروه گلیوما درجه پایین و گلیوما درجه بالا بود. برای ارزیابی عملکرد مدل، کل داده‌ها در سطح بیمار سازمان‌دهی شدند تا از بروز هرگونه نشت اطلاعات جلوگیری شود. در این چارچوب، از اعتبارسنجی متقاطع ده‌بخشی (10-fold cross validation) استفاده شد. بدین ترتیب، بیماران به ده گروه تقسیم شدند و در هر تکرار، داده‌های مربوط به بیماران نه گروه، برای آموزش و

داده‌های بیماران گروه باقی‌مانده برای آزمون به کار گرفته شد. این فرآیند موجب کاهش حساسیت مدل نسبت به یک تقسیم خاص از داده‌ها و افزایش اطمینان‌پذیری نتایج شد. در هر تکرار از اعتبارسنجی متقاطع، تمامی برش‌های مربوط به بیماران موجود در مجموعه آموزش برای آموزش مدل و تمامی برش‌های بیماران مجموعه آزمون برای ارزیابی استفاده شدند. این رویکرد، تضمین می‌کند که اسلایس‌های یک بیمار، هرگز به‌طور همزمان در هر دو مجموعه آموزش و آزمون قرار نمی‌گیرند. برای جلوگیری از ورود اسلایس‌های فاقد اطلاعات مفید، ابتدا برش‌هایی که فاقد ساختار مغزی یا فاقد ناحیه تومور بودند حذف شدند. سپس از میان برش‌های معتبر، نمونه‌ها از سطوح مختلف اسکن در امتداد محور Z انتخاب شدند تا پوشش مناسبی از نواحی مختلف مغز و ساختار تومور فراهم شود. به‌منظور افزایش تنوع ساختاری و بافتی، انتخاب اسلایس‌ها به‌صورت تصادفی از میان اسلایس‌های معتبر انجام شد. لازم به ذکر است که این مرحله بخشی از فرآیند آماده‌سازی داده‌ها بوده و جایگزینی برای روش‌های مقابله با عدم‌توازن کلاس‌ها در مرحله آموزش محسوب نمی‌شود. به دلیل عدم توازن ذاتی میان کلاس‌های گلیوما درجه بالا و گلیوما درجه پایین در مجموعه داده BraTS، مجموعه‌ای از راهکارها برای کاهش سوگیری مدل به سمت کلاس غالب در نظر گرفته شد. نخست، در مرحله انتخاب برش‌های آموزشی، از نمونه‌برداری متوازن استفاده شد تا توزیع نمونه‌های هر دو کلاس، در دسته‌های کوچک آموزشی تا حد امکان متعادل باشد. دوم، در فرآیند آموزش از وزن‌دهی به کلاس‌ها در تابع هزینه استفاده شد تا وزن بیشتری به کلاس اقلیت (گلیوما درجه پایین) اختصاص یابد و خطای مربوط به این کلاس اهمیت بیشتری پیدا کند. علاوه بر این، برای افزایش تنوع داده‌های کلاس اقلیت، از افزایش داده هدفمند، شامل تبدیلاتی مانند چرخش، وارونگی و تغییرات کنترل‌شده شدت تصویر استفاده شد. این اقدامات به مدل کمک می‌کند تا در حضور عدم توازن داده‌ها، عملکرد متوازن‌تری در تشخیص هر دو کلاس ارائه دهد.

در فرایند آموزش شبکه، از بهینه‌ساز Adam با نرخ یادگیری 0.001 و کاهش تدریجی استفاده شده است. اندازه دسته برابر با ۳۲ در نظر گرفته شده است. معیار توقف تعداد ۵۰ دوره بود که بر اساس عدم بهبود در معیار validation loss طی ۱۰ دوره متوالی، در نظر گرفته شد و از تابع ضرر cross-entropy استفاده شده است.

پیاده‌سازی در بستر تانسورفلو به همراه کراس انجام شده است و از یک سرور با پردازنده گرافیکی NVIDIA GeForce RTX 2080 Ti و ۱۱ گیگابایت حافظه VRAM استفاده شده است. به منظور ارزیابی عملکرد مدل پیشنهادی، از معیارهای ارزیابی متداول در دسته‌بندی تصاویر استفاده شده است. معیارهای صحت، دقت، بازخوانی، اختصاصیت، شاخص F1، به ترتیب در روابط ۱ تا ۵ نشان داده شده است. منحنی ROC و مساحت زیر منحنی (AUC) نیز برای ارزیابی به کار رفته‌اند.

- (1) $Accracy = (TP+TN)/(TP+TN+FP+FN)$
- (2) $Precision = TP/(TP+FP)$
- (3) $Recall = TP/(TP+FN)$
- (4) $Specificity = TN/(TN+FP)$
- (5) $F1-Score = 2 \times (Precision \times Recall) / (Precision + Recall)$

معیار حساسیت، عملکرد شبکه در شناسایی صحیح نمونه‌های منفی را نشان می‌دهد و شاخص F1 میانگین وزنی دو معیار دقت و بازخوانی است که عملکرد شبکه بر اساس تعادل بین این دو معیار را نشان می‌دهد. AUC معیاری است که با محاسبه مساحت زیر منحنی توانایی تفکیک مدل بین کلاس‌های مثبت و منفی را نشان می‌دهد.

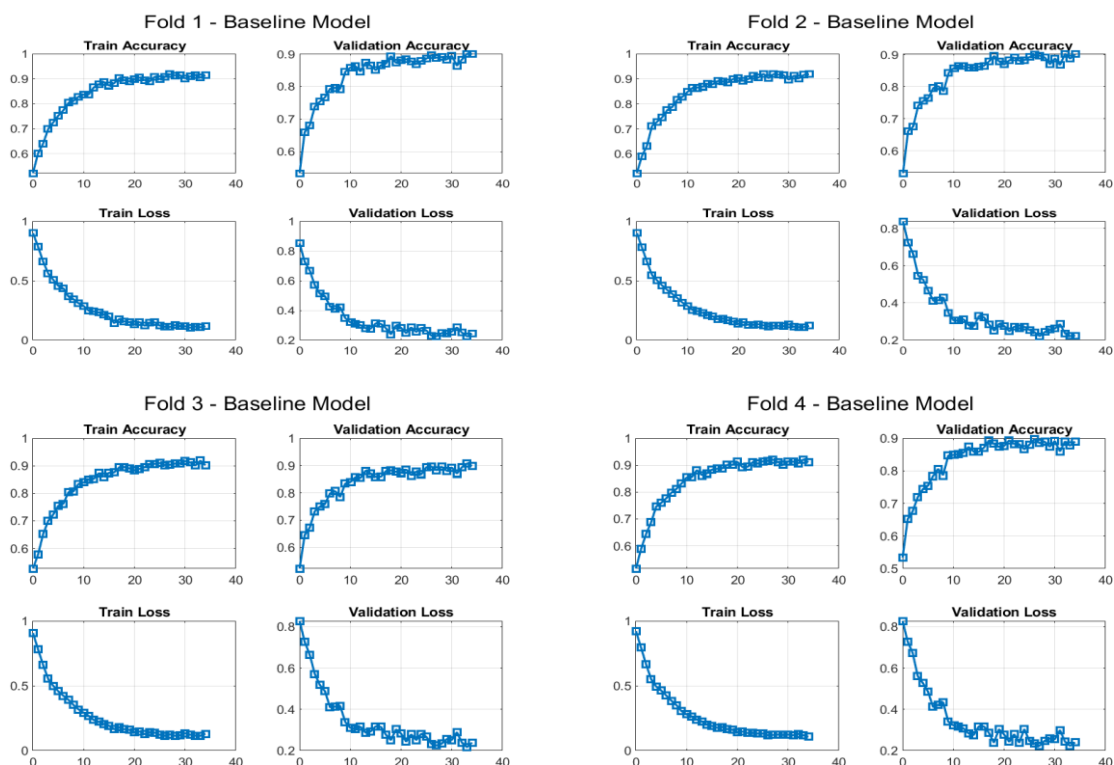
یافته‌ها

نتایج معماری اصلی شبکه عصبی عمیق VGG

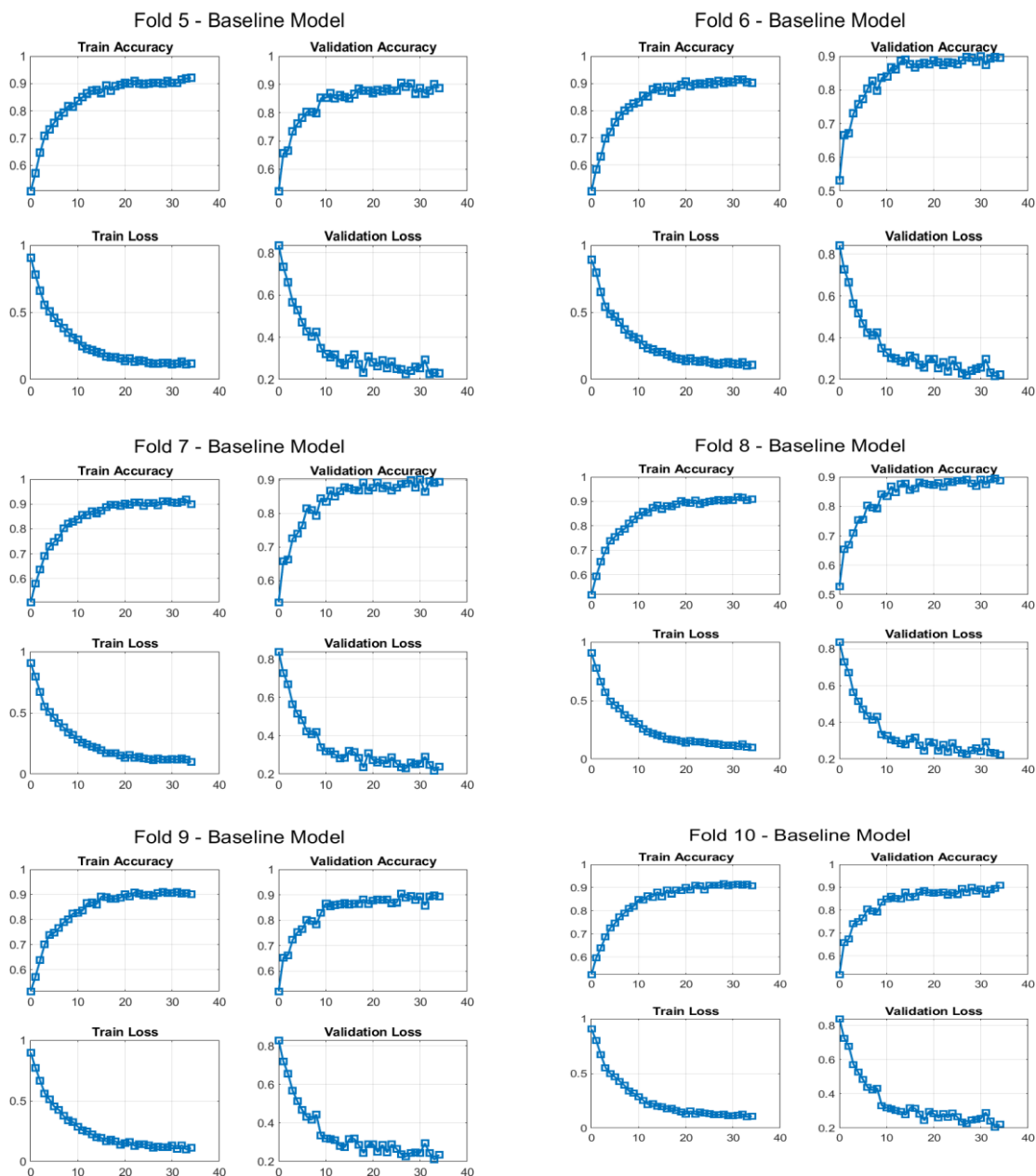
به‌منظور بررسی عملکرد شبکه عصبی پایه VGG-16، در طبقه‌بندی تصاویر تومور مغزی گلیوما، ساختار اصلی این شبکه با تنظیمات آموزشی که شرح داده شد مورد آموزش و آزمون قرار گرفت. در این ارزیابی از معیارهای بیان شده در هر دسته استفاده شده است به‌علاوه میانگین و انحراف معیار کل دسته‌ها نیز در نظر گرفته شده است تا ارزیابی جامع‌تری صورت گیرد. شکل ۲، منحنی‌های یادگیری مدل پایه VGG-16 در هر دسته را نمایش می‌دهد...

در روابط ذکر شده، TP به معنی مثبت درست است و نشان‌دهنده تعداد تصاویری است که واقعا متعلق به کلاس HGG بودند و شبکه عصبی نیز آن‌ها را به درستی به کلاس HGG متعلق می‌داند. TN به معنی منفی درست است و تعداد تصاویری را نشان می‌دهد که در واقعیت متعلق به کلاس LGG بودند و شبکه نیز آن‌ها را به درستی در کلاس LGG دسته‌بندی کرده است. FP به معنی مثبت کاذب است و نشان‌دهنده تعداد تصاویری است که متعلق به کلاس LGG بوده است و شبکه به اشتباه آن را در کلاس HGG دسته‌بندی کرده است. FN به معنی منفی کاذب است و نشان‌دهنده تعداد تصاویری است که در واقعیت HGG بودند اما شبکه به اشتباه آن‌ها را در دسته LGG قرار داده است.

معیار صحت، نشان‌دهنده تعداد پیش‌بینی‌های صحیح به کل نمونه‌ها است که عملکرد کلی شبکه را نشان می‌دهد. معیار دقت، بیانگر این است که نمونه‌هایی که شبکه مثبت پیش‌بینی کرده چند درصد واقعا مثبت بوده است. معیار بازخوانی، درصد نمونه‌های واقعا مثبت را که شبکه توانسته به درستی شناسایی کند را نشان می‌دهد.



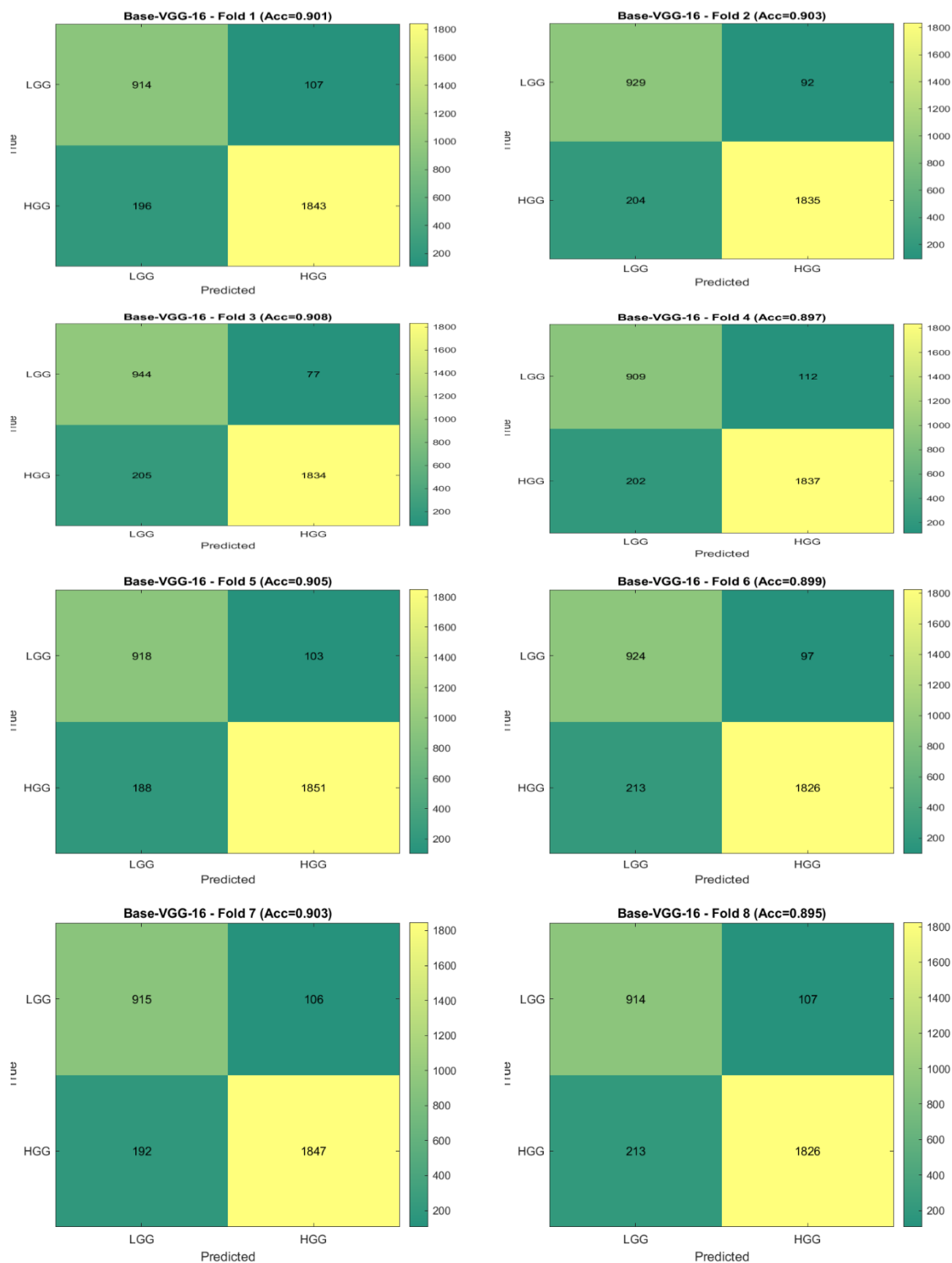
شکل ۲- منحنی‌های یادگیری مدل پایه VGG-16 در هر دسته



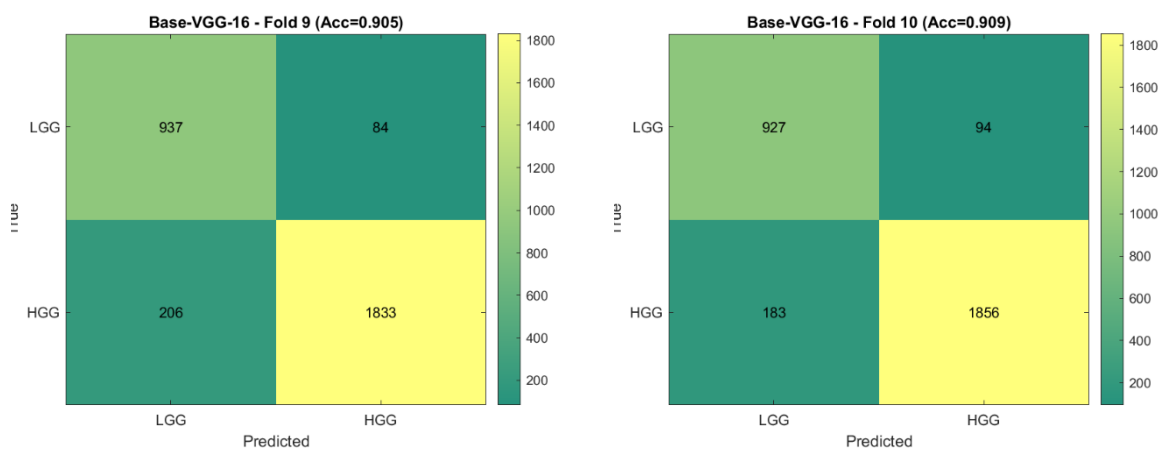
ادامه شکل ۲- منحنی‌های یادگیری مدل پایه VGG-16 در هر دسته

این امر، ناشی از توزیع داده‌ها در هر دسته است. در مجموع شکل ۳، ماتریس کانفیوژن در هر دسته را نمایش می‌دهد. هر ماتریس نشان دهنده دسته‌بندی تومورها به دو دسته LGG و HGG است که مواردی که به درستی دسته‌بندی شده‌اند در قطر اصلی ماتریس و مواردی که به درستی طبقه‌بندی نشدند در قطر فرعی ماتریس نشان داده شده است...

همان‌طور که در شکل ۲ مشاهده می‌شود با افزایش تعداد دوره‌های آموزش، دقت آموزش و همچنین دقت اعتبارسنجی افزایش یافته است که نشان دهنده عدم بیش‌برازش در آموزش مدل است. مشاهده نمودارهای خطا نشان دهنده این است که خطا نیز به صورت یکنواخت کاهش یافته است. اگر چه نوسانات جزئی در برخی دسته‌ها نشان داده شده است،



شکل ۳: ماتریس کانفیوژن مدل پایه VGG-16 در هر دسته

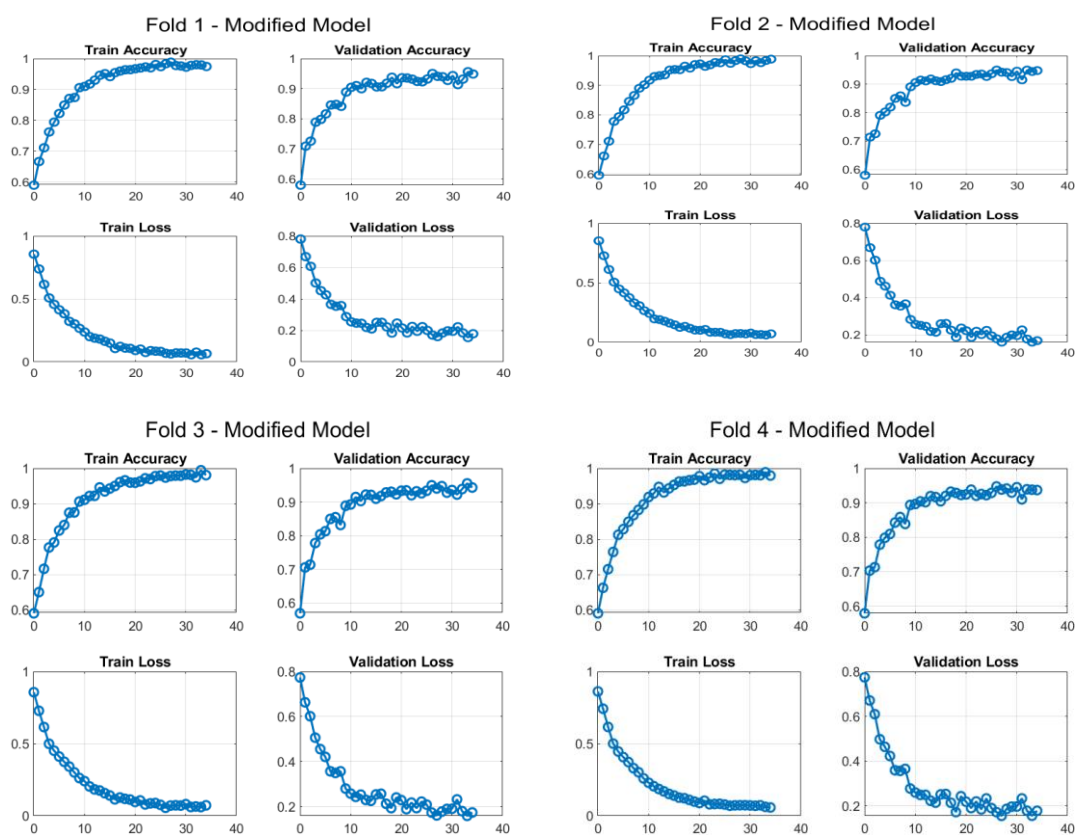


ادامه شکل ۳: ماتریس کانفیوژن مدل پایه VGG-16 در هر دسته

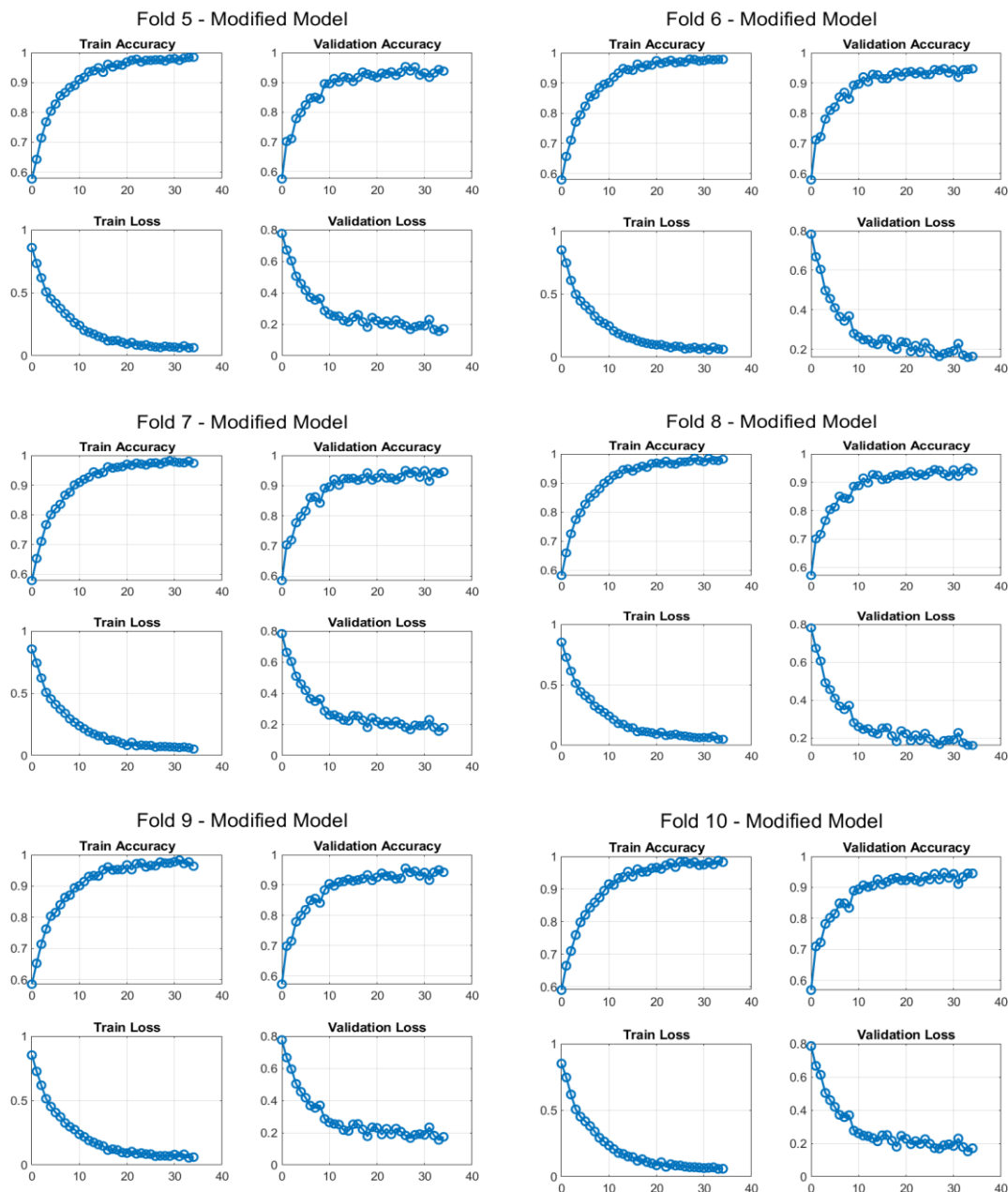
تحت همان شرایط آموزشی، مجموعه داده و روش اعتبارسنجی متقاطع ده بخشی که برای مدل پایه استفاده شده بود ارزیابی گردید. شکل ۴، منحنی‌های یادگیری مدل پیشنهادی VGG-16 در هر دسته را نمایش می‌دهد.

نتایج معماری پیشنهادی شبکه عصبی عمیق VGG

نتایج حاصل از نسخه اصلاح شده معماری شبکه عصبی عمیق VGG-16 با حذف هدفمند لایه Max-pooling در این بخش ارایه می‌گردد. این معماری پیشنهادی،



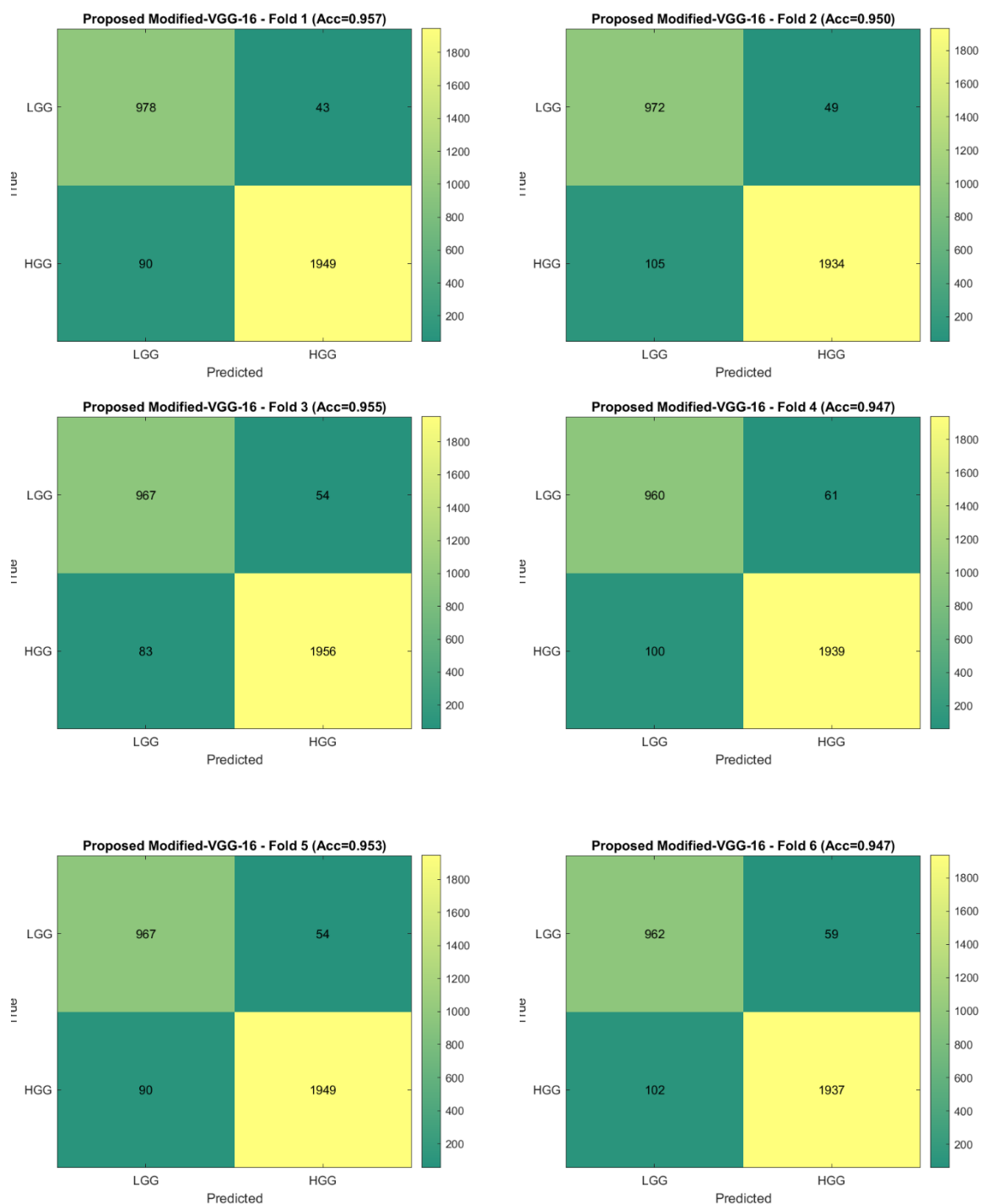
شکل ۴- منحنی‌های یادگیری مدل پیشنهادی VGG-16



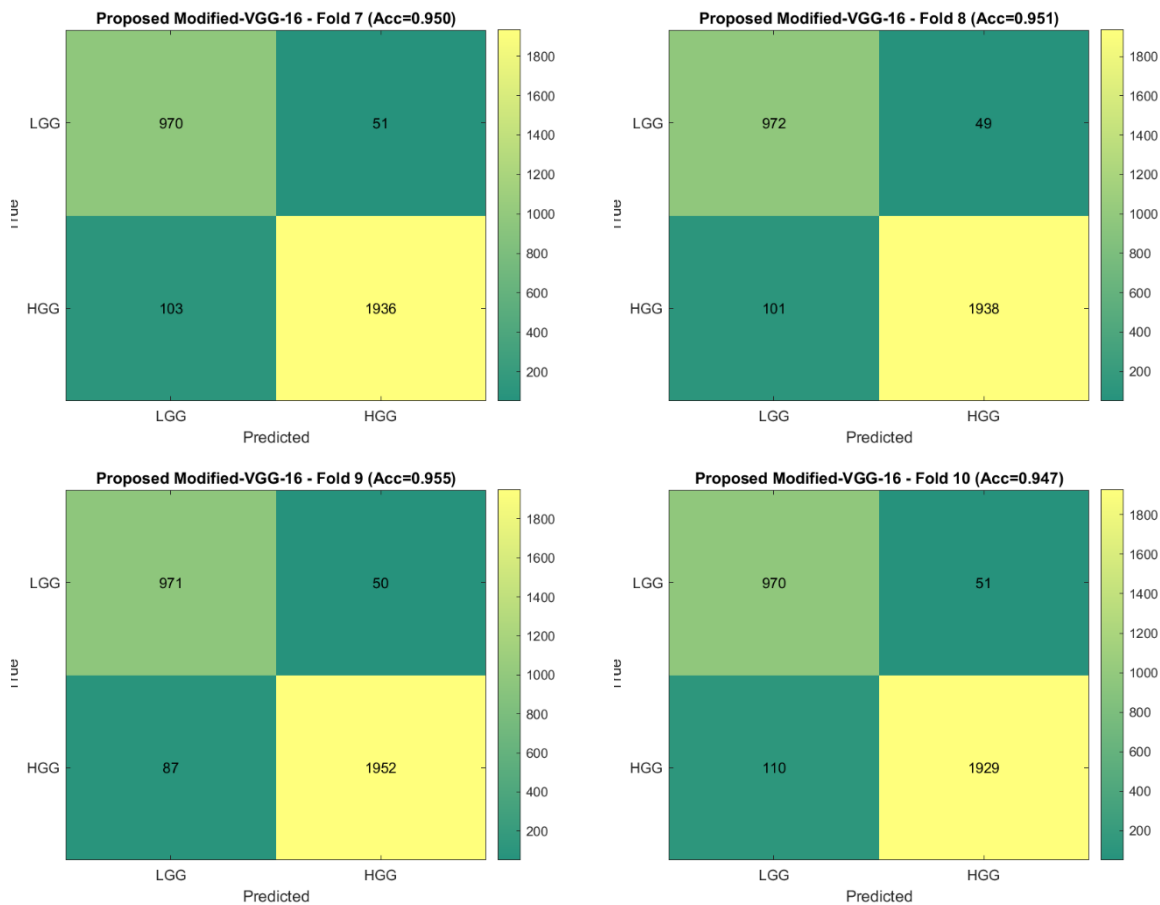
ادامه شکل ۴- منحنی های یادگیری مدل پیشنهادی VGG-16

نیز اتفاق نیفتاده است. شکل ۵، ماتریس کانفیوژن مدل پیشنهادی را نشان می‌دهد. مقایسه شکل ۵ با ماتریس کانفیوژن مدل پایه، برای هر دو کلاس LGG و HGG نرخ‌های بالای تشخیص درست را نشان می‌دهد. این امر نشان می‌دهد که مدل پیشنهادی به جزئیات ساختاری ظریف حساس‌تر بوده است و پراکندگی کمتری نسبت به مدل پایه دارد.

همان‌طور که در شکل ۴، مشاهده می‌شود دقت آموزش و اعتبار سنجی در مدل پیشنهادی، بیشتر از مدل پایه است. همچنین، خطای آموزش به‌طور پیوسته کاهش یافته است. مقایسه نمودار با نمودارهای مربوط به آموزش شبکه عصبی پایه نشان می‌دهد که خطای اعتبارسنجی در سطوح پایین‌تر از مدل پایه به حالت پایدار رسیده است. این امر نشان می‌دهد که خطای بیش برآزش در آموزش مدل پیشنهادی



شکل ۵: ماتریس کانفیوژن مدل پیشنهادی VGG-16 در هر دسته



ادامه شکل ۵: ماتریس کانفیوژن مدل پیشنهادی VGG-16 در هر دسته

است اما در چند فولد (به‌ویژه فولدهای ۶، ۷ و ۱۰) مدل پیشنهادی زودتر از مدل پایه به بیشترین دقت اعتبارسنجی رسیده است. این نتایج، نشان می‌دهد که مدل اصلاح‌شده در مجموع دارای روند همگرایی روان‌تر و سریع‌تری نسبت به مدل پایه است. جدول ۱ مقایسه سرعت همگرایی بین مدل پایه و مدل پیشنهادی را نشان می‌دهد.

جدول ۱: مقایسه سرعت همگرایی بین مدل پایه و مدل پیشنهادی (۱۰ فولد اعتبارسنجی متقابل)

معیار (Metric)	میانگین و انحراف معیار مدل پایه	میانگین و انحراف معیار مدل پیشنهادی
دوره متناظر با بیشترین دقت اعتبارسنجی	$3/14 \pm 3/3$	$3/19 \pm 29/1$

تحلیل سرعت همگرایی: در این مطالعه، به‌منظور ارزیابی کمی سرعت

همگرایی، برای هر فولد، دوره (epoch) متناظر با بیشترین دقت اعتبارسنجی استخراج شد. میانگین این مقدار در ده فولد برای مدل پیشنهادی برابر با ۲۹/۱ دوره و برای مدل پایه برابر با ۳۰/۳ دوره بود. اگرچه اختلاف بین دو مدل از نظر عددی نسبتاً کوچک

مقاطع ده‌بخشی و بر روی یکسان‌ترین تقسیم‌بندی داده‌ها، آموزش و ارزیابی شدند. در هر فولد، مقادیر معیارهای Accuracy, Precision, Recall, Specificity, F1-Score و AUC به‌صورت مستقل برای هر دو مدل محاسبه گردید. به این ترتیب، برای هر معیار دو مجموعه زوجی شامل ۱۰ مقدار (یک مقدار برای مدل پایه و مقدار متناظر برای مدل پیشنهادی بر روی همان فولد) فراهم شد. آزمون Wilcoxon signed-rank بر روی این داده‌های زوجی اعمال گردید و نتایج آن در جدول ۲ گزارش شده است.

همان‌گونه که در جدول ۱ مشاهده می‌شود، مدل پیشنهادی به‌طور متوسط حدود ۱/۲ دوره زودتر به بیشترین دقت اعتبارسنجی می‌رسد و نرخ همگرایی آن تقریباً ۴ درصد سریع‌تر از مدل پایه است. مقایسه و تحلیل آماری دو معماری برای مقایسه آماری عملکرد معماری پایه VGG-16 و معماری پیشنهادی، از آزمون Wilcoxon signed-rank استفاده شد. هر دو مدل با استفاده از اعتبارسنجی

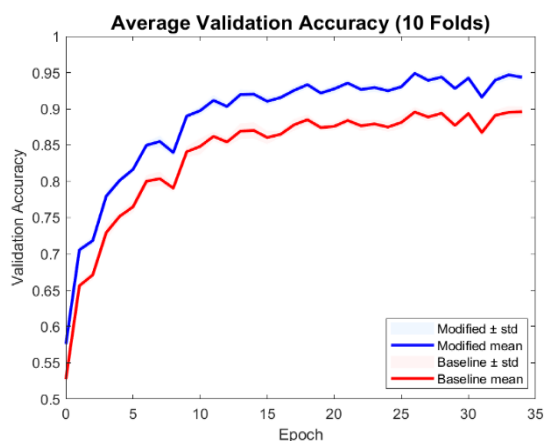
جدول ۲: آزمون ویلکاکسون برای مقایسه معماری پایه و معماری پیشنهادی

شاخص	سطح معنی داری (p)	اندازه اثر	میانگین مدل پایه	میانگین مدل پیشنهادی
Accuracy	۰/۰۰۰۹۹۷۷	۰/۹۵۱۸	۰/۹۰۲۶	۰/۹۵۱۲
Precision	۰/۰۰۰۹۹۷۷	۰/۹۵۱۸	۰/۹۴۹۵	۰/۹۷۳۹
Recall	۰/۰۰۰۹۹۷۷	۰/۹۵۱۸	۰/۹۰۱۸	۰/۹۵۲۴
Specificity	۰/۰۰۰۹۹۷۷	۰/۹۵۱۸	۰/۹۰۴۱	۰/۹۴۹۰
F1	۰/۰۰۰۹۹۷۷	۰/۹۵۱۸	۰/۹۲۵۰	۰/۹۶۳۰
AUC	۰/۰۰۰۹۹۷۷	۰/۹۵۱۸	۰/۹۰۱۶	۰/۹۵۱۴

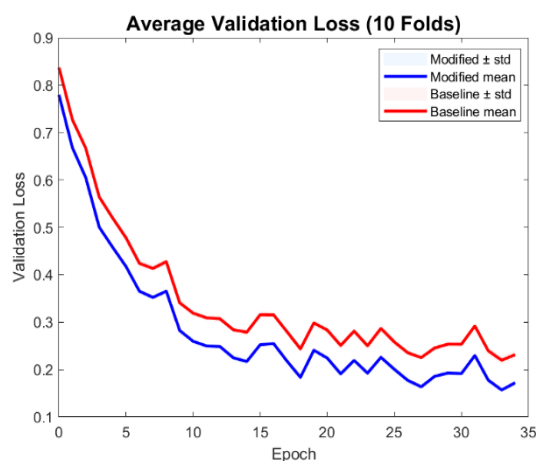
است. شایان ذکر است که در تمامی فولدهای اعتبارسنجی متقاطع، مقادیر معیارهای عملکرد مدل پیشنهادی نسبت به مدل پایه بیشتر بوده است. از این رو الگوی اختلاف‌های زوجی در آزمون Wilcoxon برای همه شاخص‌ها مشابه بوده و در نتیجه مقدار آماره آزمون، سطح معنی‌داری (p-value) و اندازه اثر برای معیارهای مختلف مقادیر یکسانی به دست آمده‌اند.

مقایسه میانگین دقت اعتبارسنجی و خطای اعتبارسنجی معماری پایه VGG-16 و معماری پیشنهادی در شکل ۶ نشان داده شده است.

همان‌گونه که در جدول ۲ نشان داده شده، تمامی معیارها دارای مقدار p کمتر از ۰/۰۰۱ بوده و اندازه اثر نیز بر اساس طبقه‌بندی کوهن در محدوده اثر بزرگ قرار دارد؛ امری که بیانگر برتری معنادار آماری معماری پیشنهادی نسبت به معماری پایه است. به علاوه، روش پیشنهادی در تمامی شاخص‌ها عملکرد بهتری داشته است. میزان بهبود بین ۴/۴۸ درصد (در معیار AUC) تا ۵/۱۷ درصد (در معیارهای Recall، Specificity و AUC) متغیر بوده است. مقدار سطح معنی داری، در همه دسته‌ها یکسان و برابر کمترین مقدار ممکن در آزمون ویلکاکسون برای $N=10$ است که نشان دهنده سازگاری مطلق بهبودها و سطح معنی داری $p < 0.001$ از نظر آماری



ب



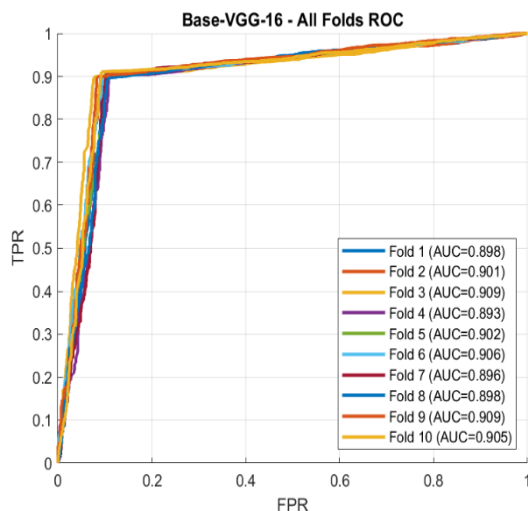
الف

شکل ۶- مقایسه میانگین دقت اعتبارسنجی و خطای اعتبارسنجی معماری پایه VGG-16 و معماری پیشنهادی در ده دسته الف: خطای اعتبارسنجی، ب: دقت اعتبارسنجی

تومورهای LGG و HGG توسط مدل پیشنهادی، با حفظ جزئیات مکانی و ساختارهای ظریف در تصاویر پیچیده مغز است.

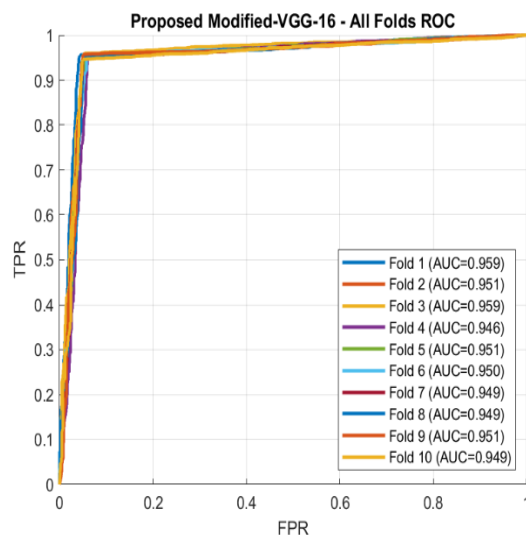
همانطور که در شکل ۶، مشاهده می‌شود مدل پیشنهادی به دقت بالاتر و خطای کمتر در اعتبارسنجی دست یافته است که بیانگر قدرت تفکیک بهتر

پایه به 0.9514 در مدل پیشنهادی تغییر یافته است که نشان‌دهنده برتری مدل پیشنهادی است. از طرفی در نمودار مدل پیشنهادی مشاهده می‌شود که پراکندگی در دسته‌ها به طور قابل توجهی کمتر از مدل پایه است که نشان‌دهنده پایداری مدل در تمام فولدها است.



ب

شکل ۷ منحنی ROC در هر دسته برای دو مدل پایه و پیشنهادی را نشان می‌دهد. همانطور که مشاهده می‌شود مدل پیشنهادی (نمودار سمت راست)، حساسیت و اختصاصیت بالاتری را نشان می‌دهد چرا که در همه دسته‌ها منحنی به گوشه بالا سمت چپ نزدیک‌تر بوده است. مقدار میانگین AUC از 0.9016 در مدل



الف

شکل ۷- منحنی ROC در هر دسته برای دو مدل پایه و پیشنهادی، الف: معماری پیشنهادی، ب: معماری پایه

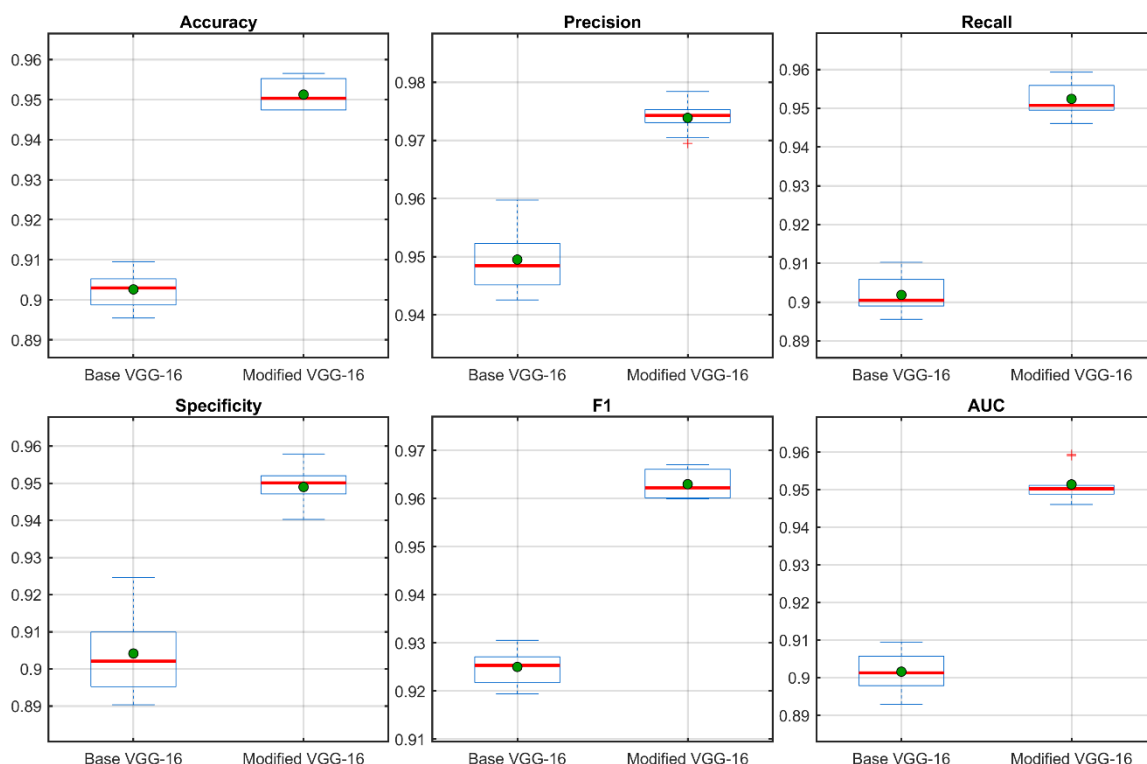
جدول ۳: میانگین و انحراف معیار شاخص‌های عملکردی در مدل پایه و مدل پیشنهادی

شاخص	مدل پایه VGG-16	مدل پیشنهادی VGG-16
Accuracy	0.9026 ± 0.0045	0.9512 ± 0.0035
Precision	0.9495 ± 0.0055	0.9739 ± 0.0026
Recall	0.9018 ± 0.0050	0.9524 ± 0.0044
Specificity	0.9041 ± 0.0110	0.9490 ± 0.0051
F1-Score	0.9250 ± 0.0035	0.9630 ± 0.0027
AUC	0.9016 ± 0.0055	0.9514 ± 0.0044

مدل پایه در معیارهای ذکر شده در این پژوهش، در شکل ۸ نشان داده شده است. در این نمودار، جعبه آبی رنگ محدوده بین چارک اول و چارک سوم را نمایش می‌دهد. خط افقی قرمز درون هر جعبه، نشان‌دهنده میانه است. دایره سبز رنگ معرف میانگین است. و علامت‌های به علاوه قرمز، مقادیر پرت (Outlier)ها را نشان می‌دهند.

جدول ۳، میانگین شاخص‌های عملکردی در مدل پایه و مدل پیشنهادی را نشان می‌دهد. روش پیشنهادی، در تمامی معیارهای ارزیابی، عملکرد بهتری داشته است. بیشترین بهبود در شاخص صحت با $4/86$ درصد است. داده‌های جدول نشان می‌دهد که در سایر معیارها نیز بهبود وجود دارد. در ادامه برای نمایش بهتر نتایج، نمودارهای جعبه‌ای آمده است. نمودار جعبه‌ای عملکرد مدل پیشنهادی در مقایسه با

Base VGG-16 vs Modified VGG-16 Performance Metrics



شکل ۸- نمودار جعبه‌ای عملکرد مدل پیشنهادی در مقایسه با مدل پایه در شش معیار ارزیابی

همگرایی سریع‌تر مدل پیشنهادی را تأیید می‌کند. این موضوع، همراه با بهبود مشاهده شده در ماتریس‌های کانفیوژن، بیانگر آن است که کاهش کنترل شده در عملیات pooling باعث ساده‌تر شدن فضای جستجوی شبکه و تسهیل بهینه‌سازی شده است.

برای ارزیابی علمی‌تر تفاوت عملکرد دو مدل، آزمون آماری غیرپارامتریک ویلکاکسون بر اساس نتایج ده فولد اعتبارسنجی انجام شد. مقدار p-Value کمتر از ۰/۰۰۱ و اندازه اثر بزرگ، تفاوت معنی‌دار میان دو معماری را در شاخص‌هایی نظیر AUC تأیید می‌کند. نمودارهای جعبه‌ای نیز کاهش پراکندگی عملکرد و افزایش ثبات میان فولدها را نشان می‌دهند که بیانگر ارتقای تعمیم‌پذیری مدل پیشنهادی است.

از آنجا که مدل حاضر مبتنی بر داده‌های دوبعدی و تک‌مدالیت است، مقایسه مستقیم آن با معماری‌های سه‌بعدی یا چندمدالیت از نظر فنی معتبر نیست. به همین دلیل، مقایسه با روش‌های سه‌بعدی صرفاً با هدف ارایه چشم‌انداز کلی از وضعیت پژوهش‌های اخیر آورده شده است و مقایسه کمی مستقیم تنها با روش‌های دوبعدی انجام شده است. مقایسه عملکرد مدل پیشنهادی با مطالعات (۲۰-۲۳) نشان می‌دهد که روش پیشنهادی به دقت ۹۵ درصد رسیده که بالاتر از سایر روش‌های طبقه‌بندی تومور دو بعدی بوده است. به این ترتیب، مدل پیشنهادی، در برابر روش‌های دوبعدی

همان‌طور که مشاهده می‌شود در تمام شاخص‌ها، میانه و دامنه میان چارکی در مدل پیشنهادی، بالاتر بوده که نشان‌دهنده عملکرد بهتر مدل است. جعبه و خطای بیرونی (Whisker) در مدل پیشنهادی فشرده‌تر هستند که نشان‌دهنده پراکندگی کمتر نتایج میان دسته‌هاست. این پایداری، نتایج آماری جدول ۲ را تأیید می‌کند که برای همه شاخص‌ها اندازه اثر بزرگ (۰/۹۵۱۸) به‌دست آمده است.

بحث

در این پژوهش، یک نسخه اصلاح‌شده از معماری VGG-16 برای طبقه‌بندی خودکار گلیوماهای مغزی ارایه شد. نوآوری اصلی مدل پیشنهادی، حذف هدفمند لایه Max-Pooling در بلوک چهارم شبکه است؛ تغییری که با هدف کنترل کاهش رزولوشن فضایی و حفظ ویژگی‌های بافتی در لایه‌های عمیق‌تر اعمال شد. نتایج تجربی نشان می‌دهد که این تغییر ساختاری، موجب تقویت دقت استخراج ویژگی‌ها و بهبود قابلیت تمایز میان گلیوماهای درجه پایین و بالا شده است.

حفظ رزولوشن فضایی در لایه‌های عمیق‌تر، امکان یادگیری الگوهای حساس‌تری از مرزهای تومور و ساختارهای بافتی را فراهم کرده است. تحلیل منحنی‌های یادگیری، پایداری بیشتر و سرعت همگرایی مطلوب‌تر معماری اصلاح‌شده را نسبت به نسخه استاندارد VGG نشان می‌دهد. علاوه بر این، مقایسه تعداد دوره‌های لازم برای دستیابی به بیشترین دقت اعتبارسنجی در ده فولد،

استفاده از مدل‌های سه‌بعدی و ترکیب چند مدالیته می‌تواند اطلاعات مکانی و بافتی بیشتری فراهم کند.

در مجموع، یافته‌های این پژوهش نشان می‌دهد که اصلاحات هدفمند بر روی لایه‌های میانی شبکه‌های عصبی عمیق، می‌تواند راهکاری کم‌هزینه، موثر و عملی برای ارتقای عملکرد مدل‌های تشخیصی پزشکی باشد؛ راهکاری که می‌تواند زمینه‌ساز توسعه سامانه‌های هوش مصنوعی قابل‌اعتمادتر برای استفاده در محیط‌های بالینی شود.

نتیجه‌گیری

در این پژوهش، نسخه‌ای اصلاح‌شده از معماری VGG-16 با هدف بهبود عملکرد در طبقه‌بندی خودکار گلیوماهای مغزی ارائه شد. حذف هدفمند لایه Max-Pooling در بلوک چهارم شبکه، موجب حفظ بهتر ویژگی‌های مکانی در لایه‌های عمیق‌تر و در نهایت افزایش دقت، پایداری و سرعت همگرایی مدل نسبت به معماری پایه گردید. یافته‌ها نشان می‌دهد که اعمال اصلاحات ساختاری سنجیده، در شبکه‌های عصبی عمیق می‌تواند راهبردی موثر و کم‌هزینه برای ارتقای عملکرد مدل‌ها در کاربردهای تصویربرداری پزشکی باشد. به‌طور کلی، نتایج این مطالعه نشان می‌دهد که تنظیم دقیق اجزای میانی شبکه‌های عمیق، می‌تواند جایگزینی موثر برای افزایش غیرضروری پیچیدگی مدل‌ها باشد و در نهایت به طراحی سامانه‌های تشخیصی قابل‌اعتمادتر، برای به‌کارگیری در محیط‌های بالینی کمک کند.

پیشنهاده‌ها

پیشنهاد می‌شود از مدل اصلاح‌شده VGG-16 در سامانه‌های کمک‌تشخیصی MRI برای شناسایی دقیق‌تر گلیوماهای مغزی استفاده شود. همچنین، به‌جای افزایش پیچیدگی شبکه، می‌توان با اصلاح هدفمند لایه‌های میانی دقت و کارایی مدل را بهبود داد. به‌کارگیری این رویکرد در محیط‌های بالینی می‌تواند به تصمیم‌گیری سریع‌تر و مطمئن‌تر پزشکان کمک کند.

تشکر و قدردانی

مقاله حاصل تحقیق مستقل و بدون حمایت مالی و سازمانی می‌باشد.

تضاد منافع

در انجام پژوهش حاضر، نویسندگان هیچ‌گونه تضاد منافی نداشتند.

موجود، عملکرد رقابتی و در مواردی برتر ارائه کرده است. این موضوع، در کنار همگرایی سریع‌تر و پایداری بالاتر، مزیت روش پیشنهادی را تقویت می‌کند.

برای ارائه یک نمای کلی از پیشرفت‌های اخیر در حوزه تشخیص تومور مغزی، مجموعه‌ای از مطالعات سال‌های اخیر (۲۳-۳۲) که عمدتاً مبتنی بر داده‌های سه بعدی، چندمدالیته یا معماری‌های پیچیده‌تر مانند ResNet، DenseNet و D^۳ CNN هستند، مورد بررسی قرار گرفتند. از آنجا که این روش‌ها از داده‌های حجیم‌تر، اطلاعات چندرشته‌ای و معماری‌های بسیار عمیق‌تر استفاده می‌کنند، مقایسه کمی مستقیم میان آن‌ها و روش پیشنهادی انجام نشده است.

انتخاب معماری VGG-16، چندین ملاحظه فنی وجود داشت؛ نخست، این شبکه ساختاری سبک و منظم دارد و برای مجموعه داده‌های پزشکی با اندازه محدود مناسب‌تر از مدل‌های عمیق‌تر و پیچیده‌تری مانند ResNet، DenseNet یا Vision Transformers است. دوم، سادگی لایه‌ها امکان تحلیل تأثیر دقیق تغییرات پیشنهادی را فراهم می‌کند؛ در حالی که ارزیابی اثر یک تغییر واحد در معماری‌های مدرن با مسیرهای موازی متعدد، به‌مراتب دشوارتر است. افزون‌براین، VGG در بسیاری از مطالعات کلاسیک و حتی پژوهش‌های جدید حوزه تصویربرداری پزشکی، به‌عنوان پایه مقایسه استفاده می‌شود و این امر امکان انجام مقایسه‌های منصفانه را فراهم می‌کند. با وجود این، توسعه نسخه‌های سبک‌وزن از معماری‌های پیشرفته‌تر، از جمله مدل‌های سه‌بعدی و چندمدالیته، مسیر مناسبی برای پژوهش‌های آینده به شمار می‌رود.

این مطالعه دارای محدودیت‌هایی است. به دلیل استفاده از افزایش داده‌سازی تصادفی و در لحظه، امکان گزارش مستقیم نتایج در سطح بیمار وجود نداشت؛ با این حال، استفاده از اعتبارسنجی متقاطع بیمار-محور تضمین می‌کند که هیچ‌یک از اسلایس‌های مربوط به یک بیمار در هر دو مجموعه آموزش و آزمون قرار نگیرد. بازطراحی زیرساخت لاگ‌گذاری برای ثبت پیش‌بینی‌ها در سطح بیمار، در برنامه پژوهش‌های آینده قرار دارد. همچنین، اگرچه یک مطالعه تحلیل حذف (Ablation Study) کامل بر روی تمامی بلوک‌ها انجام نشد، تمرکز بر بلوک چهارم بر اساس تحلیل نقش عملیات pooling در حفظ ویژگی‌های بافتی صورت گرفت. بررسی جایگزین‌هایی نظیر pooling تطبیقی یا لایه‌های مبتنی بر stride می‌تواند در تحقیقات آینده بیش عمیق‌تری درباره عملکرد معماری فراهم سازد. استفاده از برش‌های دوبعدی و تنها یک مدالیته می‌تواند یک محدودیت باشد در حالی که

References

1. Louis DN, et al. The 2021 WHO classification of tumors of the central nervous system: a summary. *Neuro-oncol.* 2021;23(8):1231–1251.
2. Zahoor MM, Khan SH, Alahmadi TJ, et al. Brain tumor MRI classification using a novel deep residual and regional CNN. *Biomedicine.* 2024;12(7):1395.
3. Sánchez-Moreno L, Perez-Peña A, Duran-Lopez L, et al. Ensemble-based convolutional neural networks for brain tumor classification in MRI: enhancing accuracy and interpretability using explainable AI. *Comput Biol Med.* 2025;195:110555.
4. Li C, Zhang F, Du Y, Li H. Classification of brain tumor types through MRIs using parallel CNNs and firefly optimization. *Sci Rep.* 2024;14(1):15057.
5. Reyes D, Sánchez J. Performance of convolutional neural networks for the classification of brain tumors using magnetic resonance imaging. *Heliyon.* 2024;10(3):e110555.
6. Ganesh S, Gomathi R, Kannadhasan S. Brain tumor segmentation and detection in MRI using convolutional neural networks and VGG16. *Cancer Biomark.* 2025;42(3):18758592241311184.

7. Saranya M, Praveena R. Accurate and real-time brain tumour detection and classification using optimized YOLOv5 architecture. *Sci Rep.* 2025;15:25286.
8. Chanemougavel V, Jayanthi K. HAIBTC-TLMO: hybrid artificial intelligence-based brain tumor classification using transfer learning and marine predators optimization. *Int J Elect Comput Eng.* 2025;12(7):13–23.
9. Nancy AM, Maheswari R. Brain tumor segmentation and classification using transfer learning based CNN model with model agnostic concept interpretation. *Multimedia Tools Appl.* 2025;84(5):2509–2538.
10. Shedbalkar J, Prabhushetty K. Deep transfer learning model for brain tumor segmentation and classification using UNet and chopped VGGNet. *Indones J Electr Eng Comput Sci.* 2024;33:1405–1415.
11. Gayathri T, Kumar S. Brain tumor segmentation and classification using CNN pre-trained VGG-16 model in MRI images. *IIUM Eng J.* 2024;25(2):196–211.
12. Pokhriyal N, Jain A. Brain tumor classification through MRIs using transfer learning with VGG16 model. *Res Appl Sci Eng Technol.* 2025;13(1):1821–1827.
13. Zhao L, Zhang Z. An improved pooling method for convolutional neural networks. *Sci Rep.* 2024;14(1):1589.
14. Hyun J, Seong H, Kim E. Universal pooling – a new pooling method for convolutional neural networks. *Exp Syst Appl.* 2021;180:115084.
15. Williams T, Li R. Wavelet pooling for convolutional neural networks. In: *ICLR*; 2018 Feb 15.
16. Awsaf A. BraTS 2020 Dataset (Training + Validation). Kaggle [Internet]. Available from: <https://www.kaggle.com/datasets/awsaf49/brats20-dataset-training-validation>.
17. Menze BH, Jakab A, Bauer S, et al. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans Med Imaging.* 2015;34(10):1993–2024. <https://pubmed.ncbi.nlm.nih.gov/25494501/>
18. Bakas S, Reyes M, Jakab A, et al. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. *arXiv preprint arXiv:1811.02629.* 2018. Available from: <https://www.med.upenn.edu/sbia/brats2020.html>
19. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556.* 2014.
20. Zahid U, Ashraf I, Khan MA, et al. BrainNet: optimal deep learning feature fusion for brain tumor classification. *Comput Intel Neurosci.* 2022;2022:1–13.
21. Kalaiselvi T, Padmapriya T, Sriramakrishnan P, Priyadharshini V. Development of automatic glioma brain tumor detection system using deep CNNs. *Int J Imaging Syst Technol.* 2020;30(4):926–938.
22. Shahzadi I, Tang TB, Meriadeau F, Quyyum A. CNN-LSTM: cascaded framework for brain tumour classification. In: *IEEE-EMBS Conf Biomed Eng Sci*; 2018; Sarawak, Malaysia. IEEE; 2018. p. 633–637.
23. Wang X, Yang S, Wu X. Automatic glioma grading based on two-stage networks by integrating pathology and MRI images. In: *MICCAI Brainlesion Workshop*; 2020 Oct 4; Cham: Springer. p. 455–464.
24. Yin B, Cheng H, Wang F, Wang Z. Brain tumor classification based on MRI images and noise reduced pathology images. In: *MICCAI Brainlesion Workshop*; 2020 Oct 4; Cham: Springer. p. 465–474.
25. Lerousseau M, Deutsch E, Paragios N. Multimodal brain tumor classification. In: *MICCAI Brainlesion Workshop*; 2020 Oct 4; Cham: Springer. p. 475–486.
26. Pei L, Hsu WW, Chiang LA, et al. A hybrid convolutional neural network method for brain tumor classification using MRI and WSI. In: *MICCAI Brainlesion Workshop*; 2020 Oct 4; Cham: Springer. p. 487–496.
27. Zhao B, Huang J, Liang C, et al. CNN-based fully automatic glioma classification with multi-modal medical images. In: *MICCAI Brainlesion Workshop*; 2020 Oct 4; Cham: Springer. p. 497–507.
28. Hamidinekoo A, Pieciak T, Afzali M, et al. Glioma classification using multimodal radiology and histology data. In: *MICCAI Brainlesion Workshop*; 2020 Oct 4; Cham: Springer. p. 508–518.
29. Liu X, Yao C, Chen H, et al. BTSC-TNAS: transformer for brain tumor segmentation and classification. *Comput Med Imaging Graph.* 2023;110:102307.
30. Aloraini M, Khan A, Aladhadh S, et al. Combining transformer and convolution for brain tumor classification. *Appl Sci.* 2023;13(6):3680.
31. Rahman MA, Masum MI, Hasib KM, et al. GliomaCNN: lightweight CNN for brain tumor classification using explainable AI. *Comput Model Eng Sci.* 2024.
32. Asad R, Rehman S, Imran A, et al. Early melanoma brain tumor detection using AI. *Biomedicines.* 2023;11(1):184...

An Improved VGG Based Deep Learning Framework for Glioma Classification in Brain MRI Images

Asieh Khosravanian ¹

Original Article

Abstract

Introduction: Early detection of brain tumors plays a significant role in treatment planning. In brain magnetic resonance imaging (MRI), the considerable similarities between healthy and cancerous tissues make accurate tumor detection a major challenge. Besides, manual classification of these images is time consuming and prone to human error. Accordingly, developing automated classification methods could provide substantial support to physicians.

Methods: In this study, a novel deep neural network architecture is proposed for automatic classification of brain tumors in MRI images. In the proposed method, the VGG 16 architecture is redesigned. To preserve the spatial resolution of features in deeper layers and to enable detection of small tumor regions as well as identification of ambiguous boundaries, the max pooling layer in the original VGG 16 architecture is removed.

Results: The proposed method was evaluated using the BraTS2020 dataset with a 10-fold cross-validation approach. The evaluation results, based on the metrics Accuracy, Precision, Recall, Specificity, F1 Score, and AUC, suggested that the architecture achieved better performance compared to the original VGG-16 neural network. More specifically, it attained values 0.9512 ± 0.0035 , 0.9739 ± 0.0026 , 0.9524 ± 0.0044 , 0.9490 ± 0.0051 , 0.9630 ± 0.0027 , and 0.9514 ± 0.0044 , respectively, for the mentioned metrics.

Conclusion: The obtained results in this study confirm that preserving spatial information and identifying ambiguous boundaries can play a significant role in improving classification accuracy. In addition, the applied modifications led to reduced over-downsampling.

Keywords: Classification; Deep Learning ; Magnetic Resonance Imaging; Neural Networks

Received: 13 Oct; 2025

Accepted: 21 Dec; 2025

Published: 31 Dec; 2025

Citation: Khosravanian A. **An Improved VGG Based Deep Learning Framework for Glioma Classification in Brain MRI Images.** Health Inf Manage 2025; 22(4): 199-214.

This article resulted from This article is the result of independent research.

1. Assistant Professor, Department of Computer Engineering, University of Larestan, Lar, Iran

Corresponding Author: Asieh Khosravanian; Assistant Professor, Department of Computer Engineering, University of Larestan, Lar, Iran. Email: khosravanian@lar.ac.ir