



طبقه بندی بهینه تومورهای مغزی در تصاویر MRI به کمک یادگیری

عمیق

زهرة عربی^۱، امید مهدی یار^۲، مهدی تقی زاده^{*۳}

- ۱- دانشکده مهندسی برق - واحد کازرون، دانشگاه آزاد اسلامی، کازرون، ایران
Arabi.zohre@gmail.com
- ۲- دانشکده مهندسی برق - واحد کازرون، دانشگاه آزاد اسلامی، کازرون، ایران
M.taghizadeh@kau.ac.ir
- ۳- دانشکده مهندسی برق - واحد کازرون، دانشگاه آزاد اسلامی، کازرون، ایران
Amir_mahdiyar@yahoo.com

چکیده: فناوری های تصویربرداری پزشکی و بیولوژیک، اطلاعات تصویری ارزشمندی از ساختار و عملکرد یک ارگان را از سطح مولکول ها تا کل جسم فراهم می کنند. مغز پیچیده ترین عضو در بدن است و با توسعه سریع فناوری های تصویربرداری پزشکی و بیولوژیکی، توجهات تحقیقاتی فزاینده ای را به خود جلب می کند. از شایع ترین بیماری های مغز می توان به ایجاد بافت ناهنجار در سلول های مغزی اشاره کرد که منجر به تشکیل تومورهای مغزی می شود. از آنجایی که تومورهای مغزی با خطر مرگ و میر قابل توجهی مرتبط هستند و پیش بینی دقیق و سریع این بیماری در روند درمان تاثیر مستقیم دارد، لذا در این تحقیق از تعداد زیادی داده های تصویربرداری MRI تومور مغزی برای شناسایی سرطان های مغز و یافتن روشی با استفاده از تکنیک های یادگیری عمیق استفاده شد. برای تشخیص خودکار از چند مدل یادگیری عمیق استفاده شد و طبقه بندی سه نوع تومور مغزی، متشکل از گلیوم، مننژیوم و هیپوفیز نیز با این الگوریتم ها انجام شد. بر اساس نتایج آزمون های انجام شده، بهترین دقت نتایج به دست آمده در این تحقیق ۹۶ درصد بود که با در نظر گرفتن نسبت ۶۰ درصد برای داده های آموزشی و ۴۰ درصد برای داده های آزمون حاصل شد.[□]

واژه های کلیدی: تصویربرداری پزشکی، تومورهای مغزی، یادگیری عمیق، تصاویر MRI، گلیوم، مننژیوم، هیپوفیز

Optimal Classification of Brain Tumors in MRI Images Using Deep Learning Techniques

Zohreh Arabi¹, Omid Mahdiyar², Mehdi Taghizadeh^{*3}

- ¹ Department of Electrical Engineering, Kazerun Branch, Islamic Azad University, Kazerun, Iran
Arabi.zohre@gmail.com
- ² Department of Electrical Engineering, Kazerun Branch, Islamic Azad University, Kazerun, Iran
M.taghizadeh@kau.ac.ir
- ³ Department of Electrical Engineering, Kazerun Branch, Islamic Azad University, Kazerun, Iran
Amir_mahdiyar@yahoo.com

Abstract:

Medical and biological imaging technologies provide valuable image information of the structure and function of an organ from the level of molecules to the whole body. The brain is the most complex organ in the body and is attracting increasing research attention with the rapid development of medical and biological imaging technologies. One of the most common brain diseases is the creation of abnormal tissue in brain cells, which leads to the formation of brain tumors.

*نویسنده مسئول؛ نشانی: ایران - کازرون - دانشگاه آزاد اسلامی کازرون - دانشکده مهندسی برق



Since brain tumors are associated with a significant risk of death and the accurate and rapid prediction of this disease has a direct impact on the treatment process, therefore, in this research, a large number of brain tumor MRI imaging data was used to identify brain cancers and find a method. Deep learning techniques were used. Several deep learning models were used for automatic diagnosis, and the classification of three types of brain tumors, consisting of glioma, meningioma, and pituitary, was also done with these algorithms. Based on the results of the conducted tests, the best accuracy of the results obtained in this research was 96%, which was obtained by considering the ratio of 60% for training data and 40% for test data.

Keywords: Medical Imaging, Brain tumors, Deep learning, glioma, meningioma, pituitary.

DOI: 00.00000/0000

تاریخ چاپ مقاله: ۱۴۰۲/۰۳/۲۱

تاریخ پذیرش مقاله: ۱۴۰۲/۰۲/۲۸

نوع مقاله: پژوهشی

تاریخ ارسال مقاله: ۱۴۰۲/۰۱/۱۵

می‌شود. تصویربرداری تشدید مغناطیسی متداول‌ترین فناوری مورد استفاده برای تشخیص سرطان مغز است و یکی از قدرتمندترین و پرکاربردترین روش‌های تصویربرداری در پزشکی بالینی است. عملکرد این فناوری به این صورت است که با استفاده از نیروی مغناطیسی قوی در اطراف اندام بیمار، از اعضای بدن بیمار عکس می‌گیرد.

در طول سه دهه گذشته، MRI پیشرفت‌های متعددی را به همراه داشته است که می‌تواند اطلاعات کمی و کاربردی مانند ریزمعماری شبکه و جریان پرفیوژن را ارائه دهد. MRI به دلیل توانایی این فناوری در عدم تهدید افراد آسیب پذیر، در مقایسه با سی تی اسکن، برای تشخیص تومورهای مغزی استفاده می‌شود. علاوه بر این، کیفیت تصاویر به دست آمده در حال پیشرفت بوده و تشخیص زود هنگام بیماری را تسهیل می‌کند. تصاویر MRI با سهولت بیشتری در اینترنت در دسترس است و در نتیجه محققانی که بر روی این موضوع تحقیق می‌کنند را تشویق به ایجاد محصولات برای کمک به کارکنان پزشکی می‌کند. فرآیند تشخیص توسط کادر پزشکی، تصمیم‌گیری و نتیجه‌گیری در مورد بیماران تومور مغزی به زمان زیادی نیاز دارد. در مقابل، درمان بیماران تومور مغزی نیاز به اقدام فوری دارد [۲].

با استفاده از روش‌های یادگیری ماشین و یادگیری عمیق برای ایجاد سیستم‌های طبقه‌بندی، در فرآیند تشخیص با استفاده از تصاویر MRI، مدت زمان تشخیص کوتاه‌تر شده و نتیجه دقیق‌تری ارائه می‌شود. در سال‌های اخیر، مطالعات متعددی به طور خاص با هدف طبقه‌بندی تومورهای مغزی با استفاده از تصاویر MRI از بیماران انجام شده است. [۷،۵]. در پنج سال گذشته، مطالعات متعدد تأیید کردند که MRI می‌تواند بیماری تومور مغزی را با نتایج سریع‌تری طبقه‌بندی کند. [۸-۱۲]. مطالعه دیگری نشان داده است که دقت خوب در سناریوی اعمال شده به دو بخش شامل مدل‌های مختلف معماری برای مجموعه داده‌های مختلف تقسیم می‌شود. [۱۳-۱۸]. مطالعه‌ای توسط سلطان و همکاران [۶] جهت استفاده از الگوریتم CNN انجام شد. مجموعه داده مورد استفاده در سناریوی اول از مجموعه داده‌ای استفاده می‌کند که از بیمارستان نانفنگ و بیمارستان عمومی، دانشگاه پزشکی تیانجین، چین، از سال ۲۰۰۵ تا ۲۰۱۰ جمع‌آوری شده بود. تحقیق مرتبط بعدی توسط گوامی و همکاران [۷] انجام شد که تشخیص تومور مغزی را از مجموعه داده‌های تصویر MRI

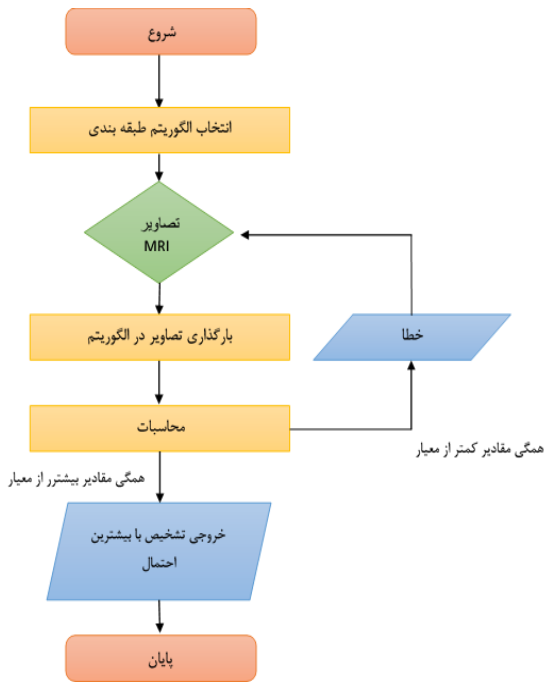
۱- مقدمه

مغز به عنوان یک عضو مهم در بدن انسان مانند سایر اعضای بدن عمل می‌کند، اما می‌توان بیان نمود مغز نقش غلب در زندگی روزمره را ایفا می‌کند. مغز با اندازه ۱۳۵۰ سی سی بیش از ۱۰۰ میلیون سلول عصبی دارد که این سلول‌ها تمام فعالیت‌های انسان را کنترل می‌کنند. قشر مغز به بخش بیرونی مغز انسان اشاره دارد که عملکرد حسی را پردازش می‌کند، این قسمت از مغز فعالیت‌های حرکتی و شناختی سطح بالایی را تولید می‌کند [۱]. با توجه به مطالب بیان شده، مغز به یکی از مهم‌ترین اندام‌ها در حمایت از زندگی انسان تبدیل می‌شود. اگر مغز بمیرد، سلول‌های عصبی به درستی کار نمی‌کنند و منجر به مرگ بیمار می‌شود.

شایع‌ترین بیماری در مغز، تومور مغزی است که به دلیل رشد بافت غیرطبیعی در مغز ایجاد می‌شود. یکی از شایع‌ترین تومورهای مغزی در کودکان و بزرگسالان گلیوم یا آستروسیتوم است [۲]. بر اساس علت وقوع، بافت غیرطبیعی در مغز به دو نوع اولیه و ثانویه تقسیم می‌شود [۳]. تومور نوع اولیه مغزی به دلیل ایجاد سلول‌های توموری است که از خود بافت مغز منشأ می‌گیرند. به همین صورت، تومورهای مغزی‌ای که به عنوان تومور مغزی ثانویه طبقه‌بندی می‌شوند در نتیجه بافت سرطانی در سایر قسمت‌های بدن هستند که به مغز گسترش می‌یابد. تومورهای مغزی به دلیل جهش‌های ژنتیکی در سلول‌های مغز رشد می‌کنند، علی‌رغم اینکه در مورد جهش‌های ژنتیکی علت نامشخص است. با این حال، عوامل متعددی مانند سن، وراثت و رادیوتراپی وجود دارد که می‌تواند خطر ابتلا به تومور مغزی را افزایش دهد. در مجموع، کشنده‌ترین نوع تومور اولیه مغزی که معمولاً در بزرگسالان رخ می‌دهد، گلیوم است [۴-۵].

بر اساس داده‌های آژانس بین‌المللی تحقیقات سرطان در سال ۲۰۱۴، بیش از ۱۲۴۰۰۰ نفر از تومورهای مغزی رنج می‌برند که منجر به مرگ بیش از ۹۷۰۰۰ بیمار شده است که نشان می‌دهد این بیماری برای بقای فرد مبتلا خطرناک تلقی می‌شود. در این زمان، به دلیل دقت بالای فناوری فعلی و پیشرفت سریع آن، در اکثر تشخیص‌های بیماری، فناوری به‌عنوان یک پشتیبان برای معاینه توسط پزشک در نظر گرفته





شکل (۱): فلوچارت مدل تشخیص تومور

همانگونه که مشخص است در ابتدا الگوریتم جهت تشخیص تومور انتخاب می‌شود. پس از این مرحله تصاویر MRI موجود در الگوریتم جهت آموزش بارگزاری می‌شوند. پس از آموزش کامل، خروجی‌ها دسته‌بندی شده و عملکرد الگوریتم با توجه به نتایج آموزش و همچنین نتایج اعتبارسنجی مورد ارزیابی و مقایسه قرار می‌گیرد.

۲-۱- مجموعه داده تشخیص تومور

این مجموعه داده که در سایت Kaggle ارائه شده است به هدف کمک به افراد برای ساخت مدل‌های یادگیری ماشین جهت تشخیص تومورهای مغزی است. این مجموعه داده شامل اسکن MRI از مغز است. مجموعه داده به سه پوشه تقسیم می‌شود. (۱) yes- این پوشه حاوی ۱۵۰۰ عدد اسکن های MRI است که تومور دارند. (۲) no- این پوشه حاوی ۱۵۰۰ عدد اسکن های MRI است که تومور ندارند. (۳) pred- این پوشه حاوی ۶۰ عدد اسکن های MRI بدون برچسب برای اهداف آزمایشی است. در شکل (۲) تعدادی از تصاویر پوشه‌های yes و no که به صورت تصادفی انتخاب شده است ارائه می‌گردد.

ایجاد شده توسط چنگ انجام می‌داد، این مجموعه داده شامل ۳۰۶۴ تصویر MRI مغز بودند که به ۱۴۲۶ تصویر مننژیو، ۷۰۸ گلیوما و ۹۳۰ تصویر هیپوفیز تقسیم می‌شدند. روش ماشین یادگیری به شدت منظم شده^۱ مورد استفاده در این مطالعه به دقت ۹۲/۴۳ درصد منجر شد. سومین مطالعه با موضوع تومورهای مغزی در سال ۲۰۱۹ انجام شد که دقت بسیار بالایی نیز داشت، تحقیق انجام شده توسط دیپاک و همکاران [۸] ارائه گردید که با مدل CNN از پیش آموزش دیده دقت ۹۳ درصد را به دست آورد. در این تحقیق الگوریتم گوگل نت^۲ به عنوان روشی برای استخراج ویژگی از تصاویر MRI و ماشین بردار پشتیبان^۳ به عنوان روشی برای طبقه‌بندی تصاویر تومورهای مغزی استفاده شد. مجموعه داده مورد استفاده در این مطالعه از یک مجموعه داده ایجاد شده توسط چنگ که از فیگشیر^۴ به دست آمده استفاده کرد [۹]. تحقیقات انجام شده توسط سواتی و همکاران [۱۰] نیز از روش یادگیری انتقالی برای طبقه بندی تصاویر MRI از تومورهای مغزی استفاده کرد، روش VGG19، بهترین دقت ۹۲/۸۲٪ را به دست آورد. مجموعه داده مورد استفاده در این مطالعه از مجموعه داده‌ای با مجموع ۳۰۶۴ تصویر که از ۲۳۳ بیمار تومور مغزی به دست آمده بهره می‌برد [۹]. مطالعه مرتبط دیگری توسط نورین و همکاران [۱۱] انجام شد که از مجموعه داده‌های مرجع [۹] استفاده می‌نمود و شامل الگوریتم‌های InceptionV3 و DenseNet201 بود، که هر یک از الگوریتم‌ها به ترتیب دقت ۹۳/۹۴٪ و ۹۵/۱۱٪ را به دست آوردند. نتایج به دست آمده از چندین مطالعه قبلی نشان دهنده دقت روش‌ها بود. جهت ایجاد دقت بهتری نسبت به تحقیقات قبلی روشی که در آخرین تحقیقات بیان گردید، دقت بالای ۹۵٪ را ارائه نمود.

بر اساس مشکلات بیان شده، هدف نویسندگان این مقاله ایجاد مدلی است که بتواند تومورهای مغزی را با استفاده از تصاویر تشدید مغناطیسی (MRI) و از طریق استخراج ویژگی هر تصویر با استفاده از روش‌های یادگیری عمیق، تشخیص داده و طبقه‌بندی کند. لذا، در این مقاله یک کتابخانه منبع باز تنسورفلو ارائه می‌شود که مجموعه ای از الگوریتم‌های پایه و پیشرفته شامل الگوریتم‌های AlexNet، GoogLeNet، VGG15، VGG19 و ResNet50 را جهت تشخیص و طبقه بندی تصاویر MRI مغز اجرا نموده و در نهایت دقت هر یک از الگوریتم‌ها بررسی خواهد شد.

۲- داده‌ها و روش‌های به کار رفته

همانگونه که در بخش قبل بیان گردید هدف نهایی در این مقاله تشخیص و طبقه‌بندی تومورهای مغزی با استفاده از الگوریتم‌های شبکه‌های عصبی کانولوشنی است که در بخش‌های بعد به صورت جامع تشریح خواهند شد. جهت تشخیص وجود یا عدم وجود تومورهای مغزی مدلی طراحی شد که در شکل (۱) فلوچارت آن ارائه می‌شود.

^۲ SVM
^۴ Figshare

^۱ RELM
^۲ GoogleNet



۲-۲- مجموعه داده طبقه‌بندی تومور

این مجموعه داده ترکیبی از سه مجموعه داده زیر است:

- Figshare
- SARTAJ
- Br35H

این مجموعه داده شامل ۷۰۲۲ تصویر از تصاویر MRI مغز انسان است که به ۴ کلاس گلیوما - مننژیوم - بدون تومور و هیپوفیز طبقه‌بندی می‌شوند.

هیچ تصویر کلاس تومور از مجموعه داده Br35H گرفته نشد.

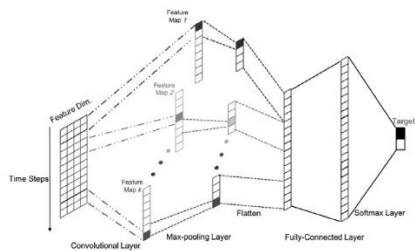
هیچ تصویر کلاس گلیوما از مجموعه داده SARTAJ گرفته نشد.

توزیع تصاویر در این مجموعه داده به صورت شکل (۳) می‌باشد و نمونه تصاویر در شکل (۴) ارائه شده است.

۲-۳- شبکه‌های عصبی کانولوشن

در این بخش، به طور خلاصه شبکه عصبی کانولوشن پایه (CNN) و انواع آن که اغلب در تجزیه و تحلیل تصاویر پزشکی استفاده می‌شود، ارائه می‌گردد. وظایفی که معمولاً توسط مدل‌های CNN در تصویربرداری پزشکی انجام می‌شود را می‌توان تقریباً به انجام سه وظیفه اصلی تقسیم کرد: طبقه‌بندی، تشخیص و تقسیم‌بندی. به عنوان مثال، مدل‌های معمولی مانند AlexNet, GoogleNet, VGG و ResNet با الحاق لایه‌های کاملاً متصل به طبقه‌بندی‌کننده‌های مختلف، هدف طبقه‌بندی را دارند. مدل‌هایی از جمله R-CNN, Faster RCNN و YOLO معمولاً برای تشخیص بیماری‌ها استفاده می‌شوند و مدل‌های دیگر مانند FCNN و U-Net به طور گسترده برای کار تقسیم‌بندی استفاده می‌شوند. همه مدل‌ها از بخش کانولوشن معماری‌های بنیادی CNN به عنوان استخراج‌کننده ویژگی استفاده می‌کنند (به عنوان مثال AlexNet, GoogleNet, VGG, ResNet) و لایه‌های کاملاً متصل را با تکنیک‌های دیگر برای کارهای خاص جایگزین می‌کنند.

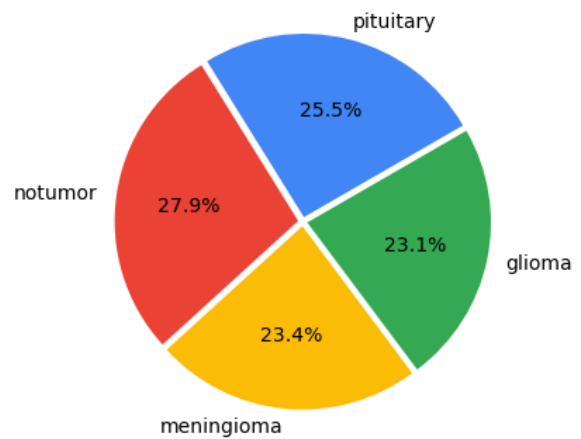
یک شبکه عصبی کانولوشن از سه لایه تشکیل شده است - لایه ورودی، لایه پنهان و لایه خروجی، که شبیه به شبکه‌های عصبی سنتی مانند شبکه‌های عصبی مصنوعی (ANN) است. تفاوت بین CNN و ANN در لایه پنهان ارائه شده است. لایه پنهان یک مدل CNN معمولاً از سه لایه تشکیل شده است، یعنی لایه کانولوشن، لایه زیر نمونه (لایه ادغام) و لایه کاملاً متصل، همانطور که در شکل (۵) ارائه شده است.



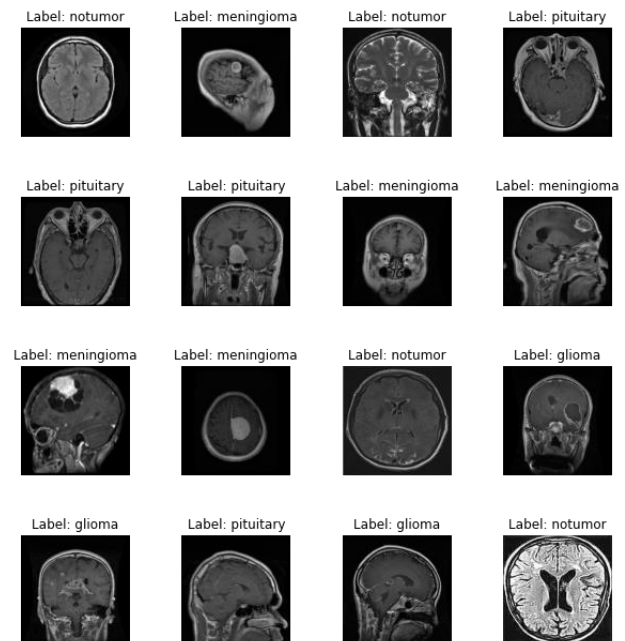
شکل (۵): شبکه عصبی کانولوشن ساده



شکل (۲): تعدادی از تصاویر پوشه‌های yes و no که به صورت تصادفی انتخاب شده است.



شکل (۳): توزیع تصاویر در مجموعه داده طبقه‌بندی



شکل (۴): نمونه‌ای از تصاویر مجموعه داده و برچسب‌های آن



این شبکه در کل از ۲۲ لایه تشکیل شده است. معماری شبکه ارائه شده در این مقاله کاملاً با VGGNet، ZFNet و AlexNet متفاوت است. این الگوریتم شامل یک لایه کانولوشن 1×1 در وسط شبکه است، و به جای استفاده از لایه‌های کاملاً متصل، از ادغام میانگین کل شبکه (میانگین جهانی) در انتهای شبکه استفاده می‌شود.

۲-۳-۳- بررسی الگوریتم VGG

شبکه VGG در دو معماری مختلف با عنوان‌های VGG 16 و VGG 19 ارائه شده است. ابتدا شبکه VGG 16 پیشنهاد شد و سپس با تغییراتی جزئی در شبکه VGG 16، شبکه VGG 19 مطرح گردید. شبکه VGG 16 همان‌طور که در شکل (۸) نشان داده شده، شامل ۱۶ لایه کانولوشنی یا ۱۶ لایه پارامتری است. شبکه VGG 16 شامل دو لایه کانولوشنی با $64 \times 3 \times 3$ فیلتر که پشت سر هم قرار گرفته‌اند. سپس، یک لایه ماکس پولینگ 2×2 با پرش به اندازه ۲ قرار گرفته است. این لایه ماکس پولینگ علاوه بر نمونه برداری، وظیفه کاهش بعد ویژگی‌ها به نصف را هم دارد. در ادامه، دو لایه کانولوشنی دیگر با $128 \times 3 \times 3$ و یک لایه ماکس پولینگ 2×2 با پرش ۲ قرار گرفته‌اند. به‌طور مشابه، سه لایه کانولوشنی با $256 \times 3 \times 3$ فیلتر و یک لایه ماکس پولینگ 2×2 با پرش ۲ قرار گرفته‌اند. ۳ لایه کانولوشنی با $512 \times 3 \times 3$ فیلتر و یک لایه ماکس پولینگ ادامه این شبکه هست که البته دو بار تکرار می‌شود. در نهایت، ویژگی‌ها تبدیل به یک بردار ویژگی می‌شوند تا در اختیار لایه‌های نورونی یا تمام‌اتصال قرار گیرند. دو لایه نورونی به ابعاد 4096 پشت سر هم قرار گرفته‌اند. در نهایت، یک لایه نورونی به ابعاد 1000 که متناظر با تعداد کلاس‌های کاربردی ما هست، در نظر گرفته شده است.

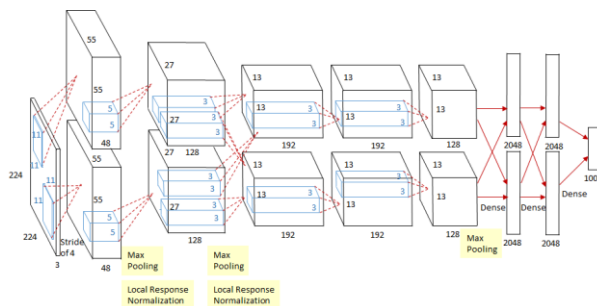
۲-۳-۴- بررسی الگوریتم ResNet

این الگوریتم تلاش می‌کند تا فرآیند یادگیری را برای الگوریتم‌های بهینه‌سازی ساده‌تر کند و عمق شبکه عصبی را به هر میزان که خواهد زیاد کند. پیشنهاد شد که وقتی شبکه نمی‌تواند به صورت مستقیم لایه هویت را یاد بگیرد به کمک یک مسیر جایگزین به او در این زمینه کمک شود. شکل (۹) این موضوع را نشان می‌دهد. همان‌طور که در این شکل نشان داده شده است علاوه بر مسیر اصلی که شامل لایه‌های کانولوشنی، عادی‌سازی دسته‌ای و تابع فعال‌سازی است، داده ورودی به صورت مستقیم و بدون هیچ تغییری با خروجی لایه‌های غیرخطی استفاده شده جمع می‌گردد.

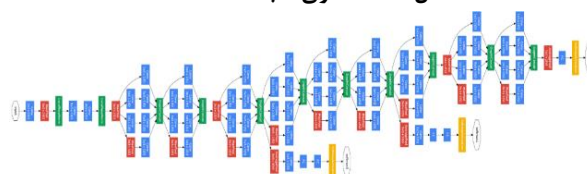
۳- نتایج

بخش نتایج به دو قسمت تقسیم می‌شود:

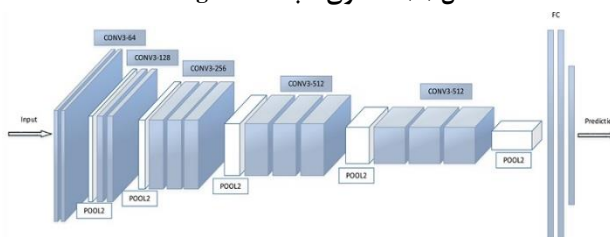
- ✓ تشخیص تومور توسط الگوریتم‌ها
- ✓ طبقه‌بندی داده‌های MRI توسط الگوریتم‌ها



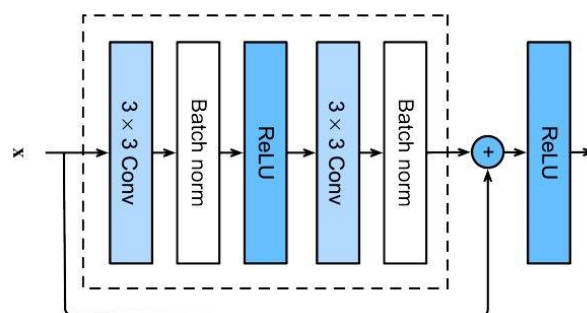
شکل (۶): معماری شبکه AlexNet



شکل (۷): معماری شبکه GoogleNet



شکل (۸): معماری شبکه VGG16



شکل (۹): استفاده از مسیر جایگزین

۲-۳-۱- بررسی الگوریتم AlexNet

الگوریتم AlexNet یک شبکه عصبی کانولوشنی عمیق است که به‌منظور شناسایی و طبقه‌بندی تصاویر رنگی با سایز $224 \times 224 \times 3$ با ساختار شکل (۶) ارائه شده است. به‌طور کلی این شبکه‌ی عصبی 62 میلیون پارامتر یادگیری و 11 لایه دارد.

۲-۳-۲- بررسی الگوریتم GoogleNet

از نام "GoogLeNet" به وضوح مشخص است که ارائه دهنده این الگوریتم Google است. و همچنین "GoogLeNet" حاوی کلمه "LeNet" برای ادای احترام به الگوریتم LeNet پروفیسور یان لیکن است. به این الگوریتم Inception v1 نیز می‌گویند زیرا نسخه‌های ۲، ۳ و ۴ نیز وجود دارند (شکل (۷)).



۳-۱-۳- فراخوانی (حساسیت)

یکی از مهمترین پارامترهای ارزیابی، پارامتر فراخوانی است که به آن «نرخ پاسخ‌های مثبت درست» نیز می‌گویند. تمرکز اصلی معیار فراخوانی بر خلاف معیار صحت (که نتایج مثبت کاذب را نیز بررسی می‌نمود) بر روی داده‌هایی است که واقعاً «مثبت» بوده‌اند. فراخوانی به معنی نسبتی از موارد مثبت صحیح به کل نتایج مثبت (مجموع مثبت صحیح و منفی کاذب) موجود است. این پارامتر به صورت زیر محاسبه می‌شود:

$$Recall = Sensivity = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (۳)$$

۳-۱-۴- معیار تشخیص‌پذیری

از متداول‌ترین پارامترها که معمولاً در کنار فراخوانی بررسی می‌شود، پارامتر خاصیت تشخیص‌پذیری است که به آن «نرخ پاسخ‌های منفی درست» نیز می‌گویند. این خاصیت به معنی نسبتی از موارد منفی است که آزمایش آن‌ها را به درستی به عنوان نمونه منفی تشخیص داده است. این پارامتر به صورت زیر محاسبه می‌شود:

$$Specifity = \frac{N_{TN}}{N_{TN} + N_{FP}} \quad (۴)$$

۳-۱-۵- معیارهای ارزیابی امتیاز F1 یا F-measure

یکی از بهترین معیارها برای بررسی دقت یک مدل معیار امتیاز F1 است. این معیار صحت و فراخوانی را با هم در نظر می‌گیرد. معیار F1 در بهترین حالت، یک و در بدترین حالت صفر است.

$$F - measure = \frac{2 \times (recall \times Perciston)}{recall + Percision} \quad (۵)$$

۳-۱-۶- مساحت زیر نمودار منحنی مشخصه عملکرد

سیستم^۱

یکی از روش‌های بررسی و ارزیابی عملکرد دسته‌بندی دو دویی، «نمودار مشخصه عملکرد» است. کارایی الگوریتم‌های «دسته‌بندی دو دویی» معمولاً به وسیله شاخص‌هایی مثل «فراخوانی» و «تشخیص‌پذیری» سنجیده می‌شوند. اما در نمودار مشخصه عملکرد هر دوی این شاخص‌ها ترکیب شده و به صورت یک منحنی نمایش داده می‌شوند. اغلب برای بررسی کارایی الگوریتم‌های دسته‌بندی یا ایجاد داده‌های رسته‌ای از منحنی مشخصه عملکرد استفاده می‌کنند. معیار ارزیابی مساحت زیر نمودار منحنی مشخصه عملکرد سیستم اطلاعاتی را در مورد اینکه یک مدل کار خود در جداسازی موارد را به خوبی انجام می‌دهد ارائه می‌دهد.

در هر دو قسمت شبیه‌سازی‌ها برای الگوریتم‌های تشریح شده ارائه شد. اما برای بررسی عملکرد الگوریتم‌ها نیاز به تعریف شاخص‌های عملکردی است که در ادامه ارائه خواهد شد.

۳-۱-۱- معیارهای ارزیابی الگوریتم‌های هوش مصنوعی

جهت بررسی کیفیت عملکرد الگوریتم‌های مختلف، معیارهای ارزیابی تعریف می‌شوند. این معیارها کارایی الگوریتم را برای موارد استفاده مختلف تفسیر و توصیف می‌نمایند. در ارزیابی نتایج معمولاً معیارهای زیر متصور است و بیشترین استفاده را نسبت به معیارهای ارزیابی دیگر دارد.

- ✓ معیار ارزیابی دقت
- ✓ معیار ارزیابی صحت
- ✓ معیار ارزیابی فراخوانی
- ✓ معیار ارزیابی تشخیص‌پذیری
- ✓ معیار ارزیابی امتیاز F1
- ✓ معیار ارزیابی مساحت زیر نمودار منحنی عملکرد سیستم
- ✓ معیار ارزیابی MCC

۳-۱-۱-۱- معیار ارزیابی دقت

برای بررسی میزان درست بودن نتایج خروجی پارامتر دقت تعریف می‌شود. با بررسی این پارامتر می‌توان به صورت کلی کارایی مدل را بررسی نمود. به علاوه این پارامتر، بیانگر چگونگی آموزش مدل است. نکته مهمی که درباره این معیار می‌توان بیان نمود عدم ارائه اطلاعات جزئی در مورد کارایی مدل بوده و فقط حد آموزش یافتن مدل را نشان می‌دهد. این پارامتر از تقسیم مجموع نتایج مثبت و منفی پیش‌بینی شده صحیح تقسیم بر تعداد کل نتایج که شامل نتایج مثبت و منفی صحیح و غلط است به دست می‌آید.

$$Accuracy = \frac{N_{TN} + N_{TP}}{N_{TN} + N_{FN} + N_{TP} + N_{FP}} \quad (۱)$$

۳-۱-۲- صحت

جهت ارزیابی نتایج مثبت مدل، معیار صحت مورد بررسی قرار می‌گیرد. در واقع این معیار بیان می‌نماید چند درصد جواب‌های مثبت الگوریتم درست است. برای محاسبه این معیار، نسبت نتایج موارد صحیح طبقه‌بندی‌شده توسط الگوریتم به کل تعداد مواردی که الگوریتم چه به‌صورت صحیح و چه به‌صورت غلط، در آن کلاس مثبت طبقه‌بندی کرده است تقسیم می‌شود:

$$Percision = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (۲)$$



۳-۱-۷- معیار ارزیابی MCC^۱

را در مجموعه آموزشی ارزیابی می کند. قابل ذکر است که مجموعه آموزشی بخشی از مجموعه داده است که برای آموزش اولیه مدل استفاده می شود. ضریب زیان اعتبارسنجی میزان خطای مدل بر روی داده های دیده نشده را بیان می نماید. از نظر محاسباتی، ضریب زیان آموزش با در نظر گرفتن مجموع خطاها برای هر مثال در مجموعه آموزشی محاسبه می شود. هر بار که یک مجموعه داده از یک الگوریتم عبور می کند، گفته می شود که یک آپک^۲ را به پایان رسانده است. بنابراین، آپک در یادگیری ماشین، به انتقال کل داده های آموزشی از طریق الگوریتم اشاره دارد. آپک یک فراپارامتر است که روند آموزش مدل یادگیری ماشین را تعیین می کند.

همانگونه که در شکل (۱۰) مشخص است، ضریب زیان الگوریتم AlexNet در مرحله آموزش کمتر از ۰/۱۱ درصد و در مرحله اعتبارسنجی تقریباً برابر ۰/۰۹ درصد است. ضریب دقت در الگوریتم AlexNet در مرحله آموزش تقریباً ۹۸ درصد و در مرحله اعتبارسنجی تقریباً ۹۹ درصد است. این نتایج بیانگر آن است که در مرحله آموزش، خطای مشاهده شده پس از ۱۲۰ تکرار آموزش ۰/۱۱ درصد است و پس از آموزش الگوریتم، خطای مشاهده شده برای مجموعه داده های اعتبارسنجی ۰/۰۹ کاهش می یابد. به علاوه ضریب دقت که بیانگر تشخیص صحیح الگوریتم پس از ۱۲۰ بار تکرار آموزش است برابر ۹۸ درصد است و برای اعتبارسنجی الگوریتم پس از آموزش، ضریب دقت یا به عبارتی دیگر تشخیص صحیح توسط الگوریتم ۹۹ درصد است.

در ادامه ماتریس درهم ریختگی برای سه مرحله آموزش، آزمایش و اعتبارسنجی ارائه می شود (شکل ۱۱). همانگونه که در قسمت قبل بیان شد ماتریس درهم ریختگی بیانگر تعداد تشخیص صحیح مثبت یا منفی و همچنین تشخیص کاذب مثبت و یا منفی است. در قطر اصلی نتایج صحیح مثبت و منفی مشخص می شوند. در این مقاله، سطر اول تعداد منفی ها صحیح و کاذب ارائه می شود و در سطر دوم منفی کاذب و صحیح بیان خواهد شد. مجموعه داده ۳۰۰۰ عکس MRI تومورهای مغزی با توجه به قرار گرفتن ۶۰ درصد این عکس ها در مجموعه داده آموزش و ۴۰ درصد در مجموعه داده اعتبارسنجی به صورت ذیل بررسی خواهد شد. تعداد ۱۸۰۰ عکس به صورت تصادفی از مجموعه برای آموزش انتخاب می شوند. ۹۰۰ عکس دارای تومور و ۹۰۰ عکس فاقد تومور است. جهت بررسی عملکرد الگوریتم آموزش یافته، دو مجموعه اعتبارسنجی و آزمایش هر دو با تعداد ۶۰۰ عکس ایجاد می شود. ۳۰۰ عکس دارای تومور و ۳۰۰ عکس فاقد تومور است. در مرحله آموزش پس از ۱۲۰ تکرار، الگوریتم ۸۹۳ عکس از تصاویر بدون تومور را صحیح تشخیص می دهد، اما ۷ عدد از این تصاویر را مثبت کاذب تشخیص می دهد. در مورد تصاویر با تومور، الگوریتم ۳ عدد از تصاویر را منفی کاذب تشخیص داده و ۸۹۷ عدد را به درستی با تومور تشخیص می دهد. همچنین، در مرحله اعتبارسنجی تعداد صحیح منفی ۲۹۷ و تعداد صحیح مثبت ۲۹۹ است و تعداد منفی کاذب ۳ و تعداد مثبت کاذب ۱ تصویر است.

epoch^۲

پارامتر دیگری که برای ارزیابی کارایی الگوریتم های یادگیری ماشین از آن استفاده می شود معیار ارزیابی MCC است. این پارامتر بیانگر کیفیت کلاس بندی برای یک مجموعه باینری است. و سنجای است که بیانگر بستگی مابین مقادیر مشاهده شده از کلاس باینری و مقادیر پیش بینی شده از آن می باشد. مقادیر مورد انتظار برای این کمیت در بازه ۱- و ۱ متغیر می باشد. مقدار +۱، نشان دهنده پیش بینی دقیق و بدون خطای الگوریتم یادگیر از کلاس باینری می باشد. مقدار ۰، نشان دهنده پیش بینی تصادفی الگوریتم یادگیر از کلاس باینری می باشد. مقدار -۱، نشان دهنده عدم تطابق کامل مابین موارد پیش بینی شده از کلاس باینری و موارد مشاهده شده از آن می باشد. مقدار این پارامتر را به طور صریح، با توجه به مقادیر ماتریس آشفستگی به شرح زیر، می توان محاسبه نمود:

$$MCC = \frac{N_{TP} \times N_{TN} - N_{FP} \times N_{FN}}{\sqrt{(N_{TP} + N_{FP})(N_{TP} + N_{FN})(N_{TN} + N_{FP})(N_{TN} + N_{FN})}} \quad (۶)$$

۳-۱-۸- ماتریس درهم ریختگی

در حوزه هوش مصنوعی، ماتریس درهم ریختگی به ماتریسی گفته می شود که در آن عملکرد الگوریتم های مربوطه را نشان می دهند. معمولاً چنین نمایشی برای الگوریتم های یادگیری با ناظر استفاده می شود، اگرچه در یادگیری بدون ناظر نیز کاربرد دارد. معمولاً به کاربرد این ماتریس در الگوریتم های بدون ناظر ماتریس تطابق می گویند. هر ستون از ماتریس، نمونه ای از مقدار پیش بینی شده را نشان می دهد. در صورتی که هر سطر نمونه ای واقعی (درست) را در بر دارد. اسم این ماتریس نیز از آنجا به دست می آید که امکان این را فراهم می نماید که آسان تر اشتباه و تداخل بین نتایج را مشاهده کرد. خارج از حوزه هوش مصنوعی، این ماتریس معمولاً ماتریس پیشابندی یا ماتریس خطا نامیده می شود.

۳-۲- تشخیص تومور توسط الگوریتم ها

جهت تشخیص تومور از الگوریتم های AlexNet، GoogleNet، VGG19، VGG16 و ResNet50 استفاده شد.

۳-۲-۱- نتایج الگوریتم AlexNet

جهت بررسی عملکرد الگوریتم ضرایب دقت و زیان تعریف می شوند. پس از توضیحاتی پیرامون این ضرایب، نتایج الگوریتم AlexNet مورد بررسی قرار خواهد گرفت. ضریب دقت و زیان معیارهایی برای ارزیابی چگونگی مطابقت یک مدل یادگیری عمیق با داده های مورد بررسی است. ضریب دقت آموزش، بررسی عملکرد مدل بر روی داده هایی است که با آن ها آموزش دیده است، در حالی که ضریب دقت اعتبارسنجی میزان عملکرد مدل بر روی داده های دیده نشده است. ضریب زیان آموزش، خطای مدل

Matthew's correlation coefficient^۱



در مرحله آزمایش نیز تعداد صحیح منفی ۲۹۷ و تعداد صحیح مثبت ۲۸۹ تصویر است. در این حالت تعداد منفی کاذب ۳ و تعداد مثبت کاذب ۱۱ عدد توسط الگوریتم تشخیص داده شده است.

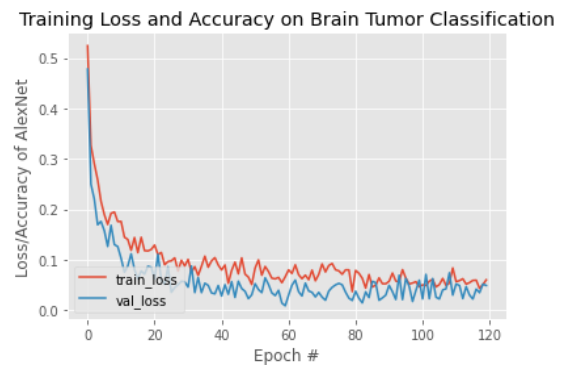
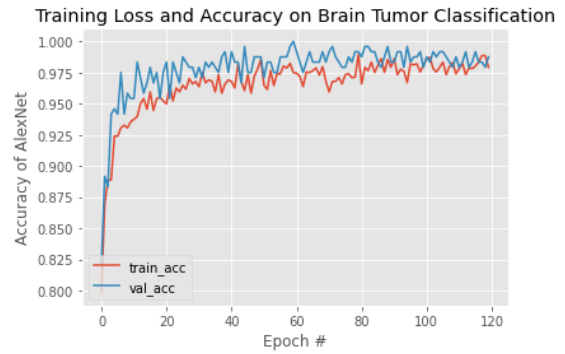
۲-۲-۳- نتایج الگوریتم GoogleNet

ضریب زیان الگوریتم GoogleNet در مرحله آموزش کمتر از ۰/۵ درصد و در مرحله اعتبارسنجی تقریباً برابر ۶ درصد است. ضریب دقت در الگوریتم GoogleNet در مرحله آموزش تقریباً ۹۳ درصد و در مرحله اعتبارسنجی تقریباً ۶۵ درصد است. این نتایج بیانگر آن است که در مرحله آموزش خطای مشاهده شده پس از ۱۲۰ تکرار روند آموزش مدل، ۰.۵ درصد است و پس از آموزش الگوریتم، خطای مشاهده شده برای مجموعه داده‌های اعتبارسنجی مقدار ۶ درصد خواهد بود. به علاوه ضریب دقت که بیانگر تشخیص صحیح الگوریتم پس از ۱۲۰ بار تکرار آموزش است برابر ۹۸ درصد است و برای اعتبارسنجی الگوریتم پس از آموزش، ضریب دقت یا به عبارتی دیگر تشخیص صحیح توسط الگوریتم ۹۹ درصد است (شکل ۱۲).

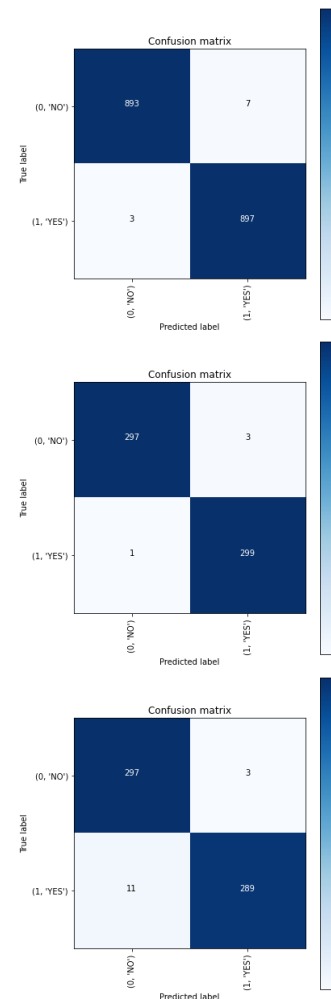
در ادامه ماتریس درهم ریختگی برای سه مرحله آموزش، آزمایش و اعتبارسنجی ارائه می‌شود. در مرحله آموزش پس از ۱۲۰ تکرار، الگوریتم ۷۵۱ عکس از تصاویر بدون تومور را صحیح تشخیص می‌دهد، اما ۶۵ عدد از این تصاویر را مثبت کاذب تشخیص می‌دهد. در مورد تصاویر با تومور، الگوریتم ۱۳۴ عدد از تصاویر را منفی کاذب تشخیص داده و ۷۶۶ عدد را به درستی با تومور تشخیص می‌دهد. در مرحله اعتبارسنجی تعداد صحیح منفی ۲۴۸ و تعداد صحیح مثبت ۲۵۰ است و تعداد منفی کاذب ۵۲ و تعداد مثبت کاذب ۵۰ تصویر است. در مرحله آزمایش نیز تعداد صحیح منفی ۲۳۵ و تعداد صحیح مثبت ۲۴۷ تصویر است. در این حالت تعداد منفی کاذب ۶۵ و تعداد مثبت کاذب ۵۳ عدد توسط الگوریتم تشخیص داده شده است (شکل ۱۳).

۳-۲-۳- نتایج الگوریتم VGG16

ضریب زیان الگوریتم VGG16 در مرحله آموزش کمتر از ۰/۴۵ درصد و در مرحله اعتبارسنجی تقریباً برابر ۵/۸ درصد است. ضریب دقت در الگوریتم VGG16 در مرحله آموزش تقریباً ۹۸ درصد و در مرحله اعتبارسنجی تقریباً ۹۳ درصد است. این نتایج بیانگر آن است که در مرحله آموزش خطای مشاهده شده پس از ۱۲۰ تکرار روند آموزش مدل، ۰/۴۵ درصد است و پس از آموزش الگوریتم، خطای مشاهده شده برای مجموعه داده‌های اعتبارسنجی مقدار ۵/۸ درصد خواهد بود. به علاوه ضریب دقت که بیانگر تشخیص صحیح الگوریتم پس از ۱۲۰ بار تکرار آموزش است برابر ۹۸ درصد است و برای اعتبارسنجی الگوریتم پس از آموزش، ضریب دقت یا به عبارتی دیگر تشخیص صحیح توسط الگوریتم ۹۳ درصد است (شکل ۱۴).



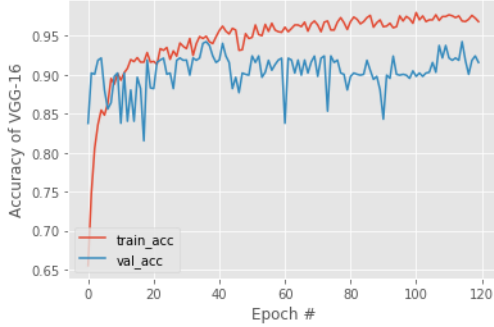
شکل (۱۰): ضریب زیان و دقت الگوریتم AlexNet



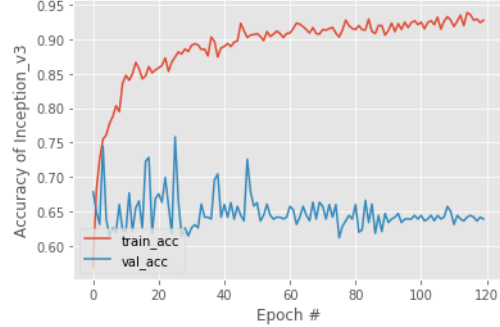
شکل (۱۱): ماتریس درهم ریختگی الگوریتم AlexNet برای سه مرحله آموزش، آزمایش، اعتبارسنجی



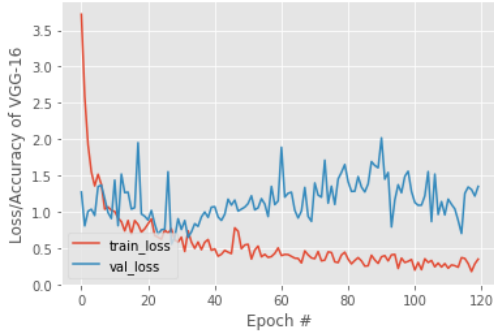
Training Loss and Accuracy on Brain Tumor Classification



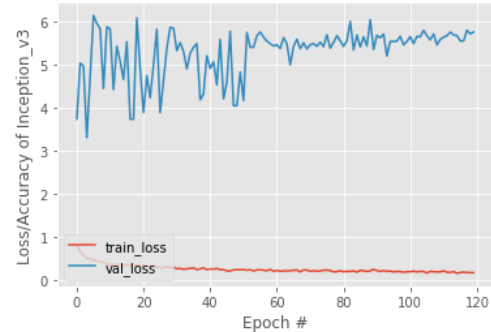
Training Loss and Accuracy on Brain Tumor Classification



Training Loss and Accuracy on Brain Tumor Classification

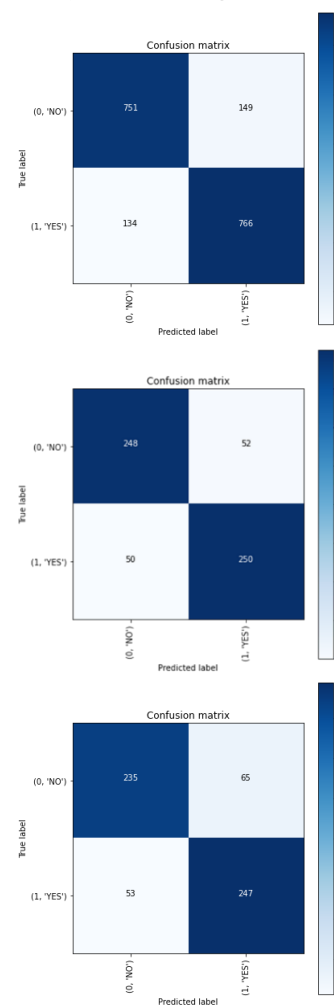
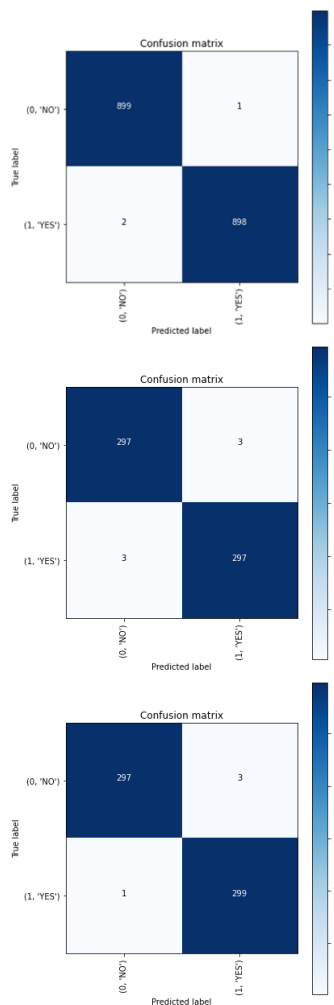


Training Loss and Accuracy on Brain Tumor Classification



شکل (۱۴): ضریب زبان و دقت الگوریتم VGG16

شکل (۱۲): ضریب زبان و دقت الگوریتم GoogleNet

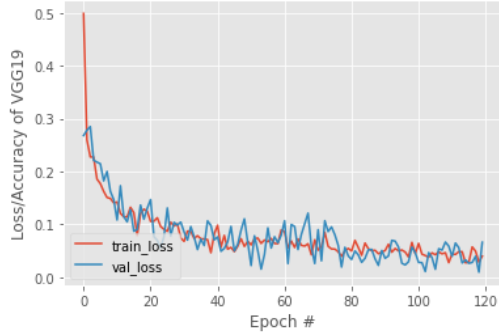


شکل (۱۵): ماتریس در هم ریختگی الگوریتم VGG16 برای سه مرحله آموزش، آزمایش، اعتبارسنجی

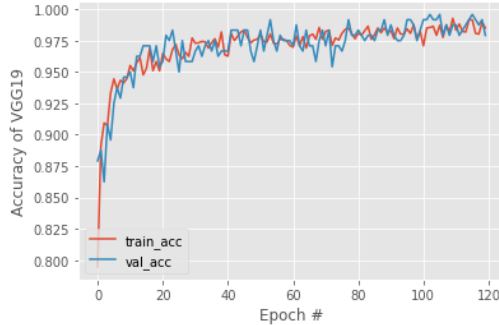
شکل (۱۳): ماتریس در هم ریختگی الگوریتم GoogleNet برای سه مرحله آموزش، آزمایش، اعتبارسنجی



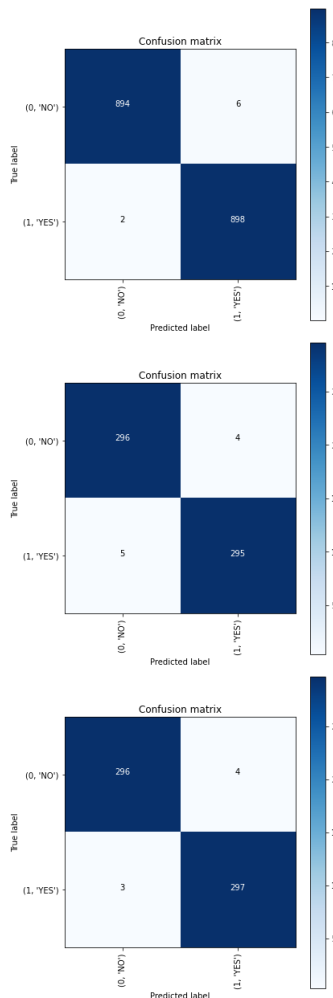
Training Loss and Accuracy on Brain Tumor Classification



Training Loss and Accuracy on Brain Tumor Classification



شکل (۱۶): ضریب زیان و دقت الگوریتم VGG19



شکل (۱۷): ماتریس درهم ریختگی الگوریتم VGG19 برای سه مرحله آموزش، آزمایش، اعتبارسنجی

در ادامه ماتریس درهم ریختگی برای سه مرحله آموزش، آزمایش و اعتبارسنجی ارائه می‌شود. در مرحله آموزش پس از ۱۲۰ تکرار، الگوریتم ۸۹۹ عکس از تصاویر بدون تومور را صحیح تشخیص می‌دهد، اما ۱ عدد از این تصاویر را مثبت کاذب تشخیص می‌دهد. در مورد تصاویر با تومور، الگوریتم ۲ عدد از تصاویر را منفی کاذب تشخیص داده و ۸۹۸ عدد را به درستی با تومور تشخیص می‌دهد. در مرحله اعتبارسنجی تعداد صحیح منفی ۲۹۷ و تعداد صحیح مثبت ۲۹۷ است و تعداد منفی کاذب ۳ و تعداد مثبت کاذب ۳ تصویر است. در مرحله آزمایش نیز تعداد صحیح منفی ۲۹۷ و تعداد صحیح مثبت ۲۹۹ تصویر است. در این حالت تعداد منفی کاذب ۳ و تعداد مثبت کاذب ۱ عدد توسط الگوریتم تشخیص داده شده است (شکل ۱۵).

۴-۲-۳- نتایج الگوریتم VGG19

ضریب زیان الگوریتم VGG19 در مرحله آموزش کمتر از ۰/۰۵ درصد و در مرحله اعتبارسنجی تقریباً برابر ۰/۰۸ درصد است. ضریب دقت الگوریتم VGG19 تقریباً برابر ۹۸/۳ درصد است. این نتایج بیانگر آن است که در مرحله آموزش، خطا پس از ۱۲۰ تکرار روند آموزش مدل، ۰/۰۵ درصد است و پس از آموزش الگوریتم، خطا برای مجموعه داده‌های اعتبارسنجی مقدار ۰/۰۸ درصد خواهد بود. به علاوه، ضریب دقت برابر ۹۸/۳ درصد است و برای اعتبارسنجی الگوریتم پس از آموزش، ضریب تشخیص صحیح توسط الگوریتم ۹۸ درصد است (شکل ۱۶).

در ادامه ماتریس درهم ریختگی ارائه می‌شود. در مرحله آموزش پس از ۱۲۰ تکرار، الگوریتم ۸۹۴ عکس از تصاویر بدون تومور را صحیح تشخیص می‌دهد، اما ۶ عدد را مثبت کاذب تشخیص می‌دهد. در مورد تصاویر با تومور، الگوریتم ۲ تصویر را منفی کاذب تشخیص داده و ۸۹۸ عدد را به درستی با تومور تشخیص می‌دهد. در مرحله اعتبارسنجی، تعداد صحیح منفی ۲۹۶، تعداد صحیح مثبت ۲۹۵، تعداد منفی کاذب ۴ و تعداد مثبت کاذب ۵ تصویر است. در مرحله آزمایش نیز تعداد صحیح منفی ۲۹۶، تعداد صحیح مثبت ۲۹۷، منفی کاذب ۴ و مثبت کاذب ۳ عدد توسط الگوریتم تشخیص داده شده است (شکل ۱۷).

۵-۲-۳- نتایج الگوریتم ResNet50

ضریب زیان الگوریتم ResNet50 در مرحله آموزش کمتر از ۰/۱ درصد و در مرحله اعتبارسنجی تقریباً برابر ۲ درصد است. ضریب دقت الگوریتم ResNet50 در مرحله آموزش تقریباً ۹۹ درصد و در مرحله اعتبارسنجی تقریباً ۸۳ درصد است. این نتایج بیانگر آن است که در مرحله آموزش، خطای آموزش مدل، ۰/۰۱ درصد است و پس از آموزش الگوریتم، خطای مشاهده شده برای مجموعه داده‌های اعتبارسنجی مقدار ۲ درصد خواهد بود. به علاوه، ضریب دقت که بیانگر تشخیص صحیح الگوریتم پس از ۱۲۰ بار تکرار آموزش است برابر ۹۹ درصد است و برای اعتبارسنجی الگوریتم پس از آموزش، ضریب دقت یا به عبارتی دیگر تشخیص صحیح توسط الگوریتم ۸۳ درصد است.



در ادامه ماتریس درهم ریختگی برای سه مرحله آموزش، آزمایش و اعتبارسنجی ارائه می‌شود. در مرحله آموزش پس از ۱۲۰ تکرار، الگوریتم ۸۹۹ عکس از تصاویر بدون تومور را صحیح تشخیص می‌دهد، اما ۱ عدد از این تصاویر را مثبت کاذب تشخیص می‌دهد. در مورد تصاویر با تومور، الگوریتم ۱ عدد از تصاویر را منفی کاذب تشخیص داده و ۸۹۹ عدد را به درستی با تومور تشخیص می‌دهد. در مرحله اعتبارسنجی تعداد صحیح منفی ۲۹۹ و تعداد صحیح مثبت ۲۹۷ است و تعداد منفی کاذب ۱ و تعداد مثبت کاذب ۳ تصویر است. در مرحله آزمایش نیز، تعداد صحیح منفی ۲۹۷ و تعداد صحیح مثبت ۲۹۹ تصویر است. در این حالت تعداد منفی کاذب ۳ و تعداد مثبت کاذب ۱ عدد توسط الگوریتم تشخیص داده شده است.

۳-۲-۶- عملکرد الگوریتم‌های تشخیص تومورهای مغزی

نتایج حاصل از به کارگیری الگوریتم‌های AlexNet، GoogLeNet، VGG16، VGG19 و ResNet50 برای مجموعه داده‌های در نظر گرفته شده در این مقاله، در جدول (۱) ارائه شده است. همانگونه که مشخص است در بین الگوریتم‌ها، نتایج بیانگر دقت بالای الگوریتم‌های VGG16 و ResNet50 در تشخیص تومور در تصاویر MRI می‌باشد. هرچند می‌توان با در نظر گرفتن معیار مقدار مساحت زیر نمودار منحنی مشخصه عملکرد سیستم، بیان نمود الگوریتم VGG16 دقت بالاتری نسبت به سایر الگوریتم‌ها دارد. علاوه بر این، کمترین دقت در بین الگوریتم‌های ارائه شده مربوط به الگوریتم GoogleNet است. می‌توان علت نتایج را به معماری الگوریتم‌ها نسبت داد که با توجه به وجود ۱۶ لایه کانولوشن در الگوریتم VGG16 یکی از دقیق‌ترین الگوریتم‌هاست. الگوریتم ResNet50 نیز به واسطه استفاده از ۱۵۲ لایه و تکنیک "اتصالات پرش" یکی از سریع‌ترین و دقیق‌ترین الگوریتم‌هاست.

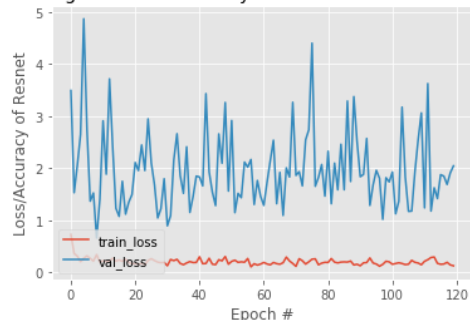
۳-۳- طبقه‌بندی داده‌های MRI

در طبقه‌بندی، کد نوشته شده علاوه بر این که باید تشخیص دهد این تصویر دارای تومور است نوع تومور را نیز باید مشخص نماید. جهت انجام طبقه‌بندی نیاز به مجموعه داده‌ای است که دارای انواع مختلف تومور بوده و برچسب هر تصویر، نوع تومور را مشخص نماید. با توجه به توضیحات بیان شده مجموعه داده پیشین ناقص است و باید با مجموعه داده دوم جایگزین شود. در ادامه عملکرد الگوریتم‌های مختلف در طبقه‌بندی تصاویر مورد بررسی قرار خواهد گرفت.

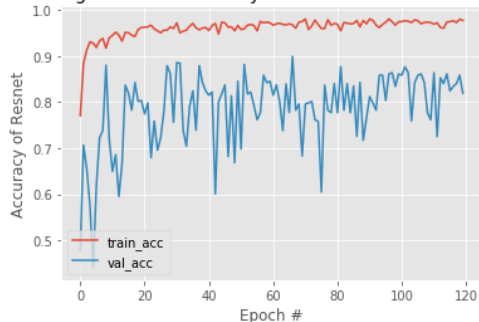
جدول (۱): عملکرد الگوریتم‌ها

ROC AUC	F1 Score	Specificity %	Precision %	Accuracy %	الگوریتم
99.67	97.67	97.67	97.77	97.67	AlexNet
88.29	80.32	80.33	80.38	80.33	GoogLeNet
99.98	99.33	99.33	99.34	99.33	VGG16
99.2	98.83	98.83	9.83	98.83	VGG19
99.86	99.33	99.33	99.34	99.33	ResNet50

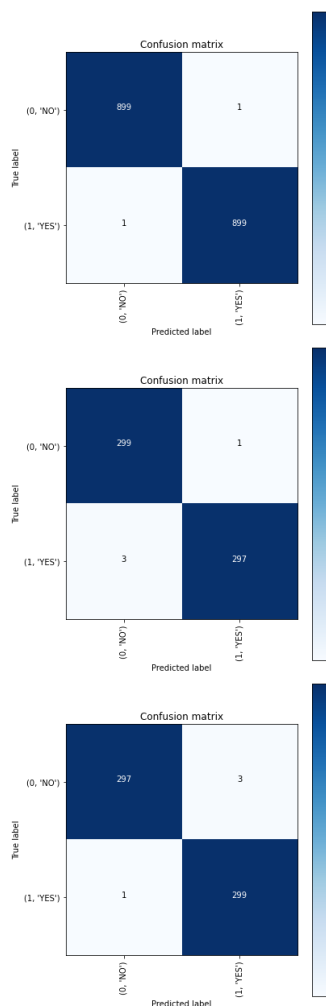
Training Loss and Accuracy on Brain Tumor Classification



Training Loss and Accuracy on Brain Tumor Classification



شکل (۱۸): ضریب زیان و دقت الگوریتم ResNet50



شکل (۱۹): ماتریس درهم ریختگی الگوریتم ResNet50 برای سه

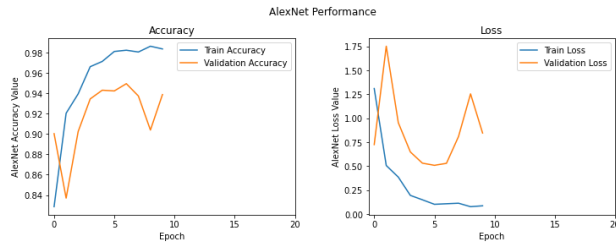
مرحله آموزش، آزمایش، اعتبارسنجی



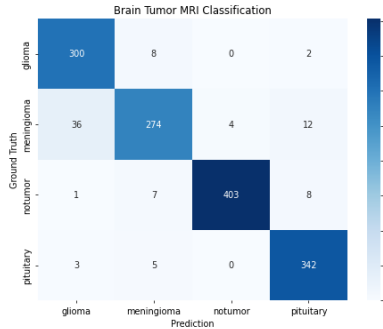
۱-۳-۳- طبقه‌بندی مجموعه داده به کمک الگوریتم

AlexNet

همانگونه که در نمودارهای عملکردی الگوریتم AlexNet مشاهده می‌شود، دقت در فاز آموزش تقریباً ۹۹ درصد و در فاز اعتبارسنجی برابر ۹۴ درصد است، خطا نیز در فاز آموزش تقریباً ۰/۱ درصد و در فاز اعتبارسنجی ۰/۸ درصد است. همانگونه که مشخص است این الگوریتم نرخ خطای کمتر از یک درصد را ثبت نموده است. به علاوه در ماتریس درهم‌ریختگی نیز مشخص است این الگوریتم با تشخیص صحیح تعداد ۴۰۳ تصویر بدون تومور از تعداد ۴۱۹ تصویر بدون تومور بهترین عملکرد را در طبقه‌بندی تصاویر بدون تومور داشته است. به همین ترتیب دقت طبقه‌بندی‌ها در هیپوفیز بالاتر از گلیوما و در گلیوما بالاتر از مننژیوم است. دقت کلی الگوریتم در طبقه‌بندی تصاویر تومورهای مغزی نیز ۹۴ درصد است.



شکل (۲۰): نمودارهای عملکردی الگوریتم AlexNet

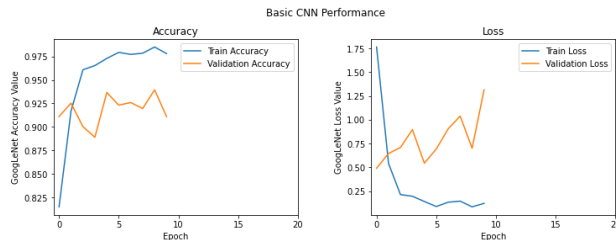


شکل (۲۱): ماتریس در هم‌ریختگی الگوریتم AlexNet

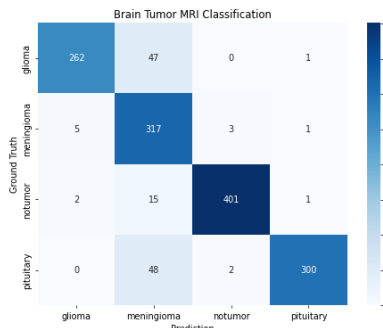
۲-۳-۳- طبقه‌بندی مجموعه داده به کمک الگوریتم

GoogLeNet

نمودارهای عملکردی الگوریتم GoogLeNet، دقت در فاز آموزش را تقریباً ۹۷ درصد و در دقت در فاز اعتبارسنجی را ۹۰ درصد بیان می‌نماید، خطا نیز در فاز آموزش تقریباً ۰/۱ درصد و در فاز اعتبارسنجی ۱/۳۸ درصد است. در ماتریس درهم‌ریختگی مشخص است این الگوریتم با تشخیص صحیح تعداد ۴۰۱ تصویر بدون تومور از تعداد ۴۱۹ تصویر بدون تومور بهترین عملکرد را در طبقه‌بندی تصاویر بدون تومور داشته است. به همین ترتیب دقت طبقه‌بندی‌ها در هیپوفیز بالاتر از مننژیوم و در مننژیوم بالاتر از گلیوما است. دقت کلی الگوریتم در طبقه‌بندی تصاویر تومورهای مغزی نیز ۹۱ درصد است.



شکل (۲۲): نمودارهای عملکردی الگوریتم GoogLeNet



شکل (۲۳): ماتریس در هم‌ریختگی الگوریتم GoogLeNet

۳-۳-۳- طبقه‌بندی مجموعه داده به کمک الگوریتم

VGG16

نمودارهای عملکردی الگوریتم VGG16، دقت در فاز آموزش را تقریباً ۹۹ درصد و در دقت در فاز اعتبارسنجی را ۹۵ درصد بیان می‌نماید، خطا نیز در فاز آموزش تقریباً ۰/۰۱ درصد و در فاز اعتبارسنجی ۰/۱ درصد است. جدول شاخص‌های عملکردی الگوریتم VGG16 برای هر چهار کلاس طبقه‌بندی که به ترتیب گلیوما، مننژیوم، تصاویر بدون تومور و هیپوفیز است، بیانگر دقت بالای الگوریتم در تشخیص کلاس تصاویر بدون تومور و تومورهای هیپوفیز و دقت پایین‌تر الگوریتم در تشخیص تصاویر تومورهای مننژیوم است. دقت کلی الگوریتم در طبقه‌بندی تصاویر تومورهای مغزی نیز ۹۶ درصد است.

جدول (۲): عملکرد الگوریتم AlexNet

F1 Score	Specificity %	Precision %	کلاس
92	97	88	گلیوما
88	84	93	مننژیوم
98	96	99	بدون تومور
96	98	94	هیپوفیز

جدول (۳): عملکرد الگوریتم GoogLeNet

F1 Score	Specificity %	Precision %	کلاس
91	85	97	گلیوما
84	97	74	مننژیوم
97	96	99	بدون تومور
92	86	99	هیپوفیز

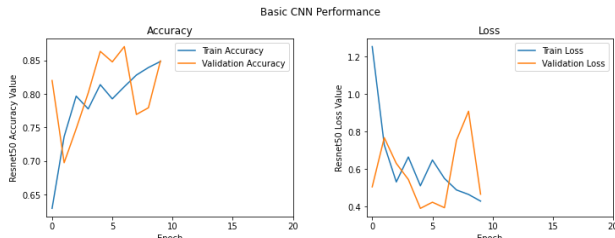


۳-۳-۴- طبقه‌بندی مجموعه داده به کمک VGG19

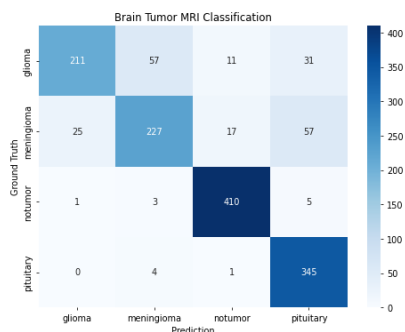
نمودارهای عملکردی الگوریتم VGG19، دقت در فاز آموزش را تقریباً ۹۹ درصد و در دقت در فاز اعتبارسنجی را ۹۳ درصد بیان می‌نماید، خطا نیز در فاز آموزش تقریباً ۰/۰۱ درصد و در فاز اعتبارسنجی ۰/۱۸ درصد است. جدول شاخص‌های عملکردی الگوریتم VGG19 برای هر چهار کلاس طبقه‌بندی که به ترتیب گلیوما، مننژیوم، تصاویر بدون تومور و هیپوفیز است، بیانگر دقت بالای الگوریتم در تشخیص کلاس تصاویر بدون تومور و تومورهای هیپوفیز و دقت پایین‌تر الگوریتم در تشخیص تصاویر تومورهای مننژیوم است. دقت کلی الگوریتم در طبقه‌بندی تصاویر تومورهای مغزی نیز ۹۴ درصد است.

۳-۳-۵- طبقه‌بندی مجموعه داده به کمک ResNet50

نمودارهای عملکردی الگوریتم VGG19، دقت در فاز آموزش را تقریباً ۹۹ درصد و در دقت در فاز اعتبارسنجی را ۹۳ درصد بیان می‌نماید، خطا نیز در فاز آموزش تقریباً ۰/۰۱ درصد و در فاز اعتبارسنجی ۰/۱۸ درصد است. جدول شاخص‌های عملکردی الگوریتم ResNet50 برای هر چهار کلاس طبقه‌بندی که به ترتیب گلیوما، مننژیوم، تصاویر بدون تومور و هیپوفیز است، بیانگر دقت بالای الگوریتم در تشخیص تصاویر بدون تومور و دقت پایین‌تر الگوریتم در تشخیص تصاویر تومورهای مننژیوم است. دقت کلی الگوریتم در طبقه‌بندی تصاویر تومورهای مغزی نیز ۸۵ درصد است.



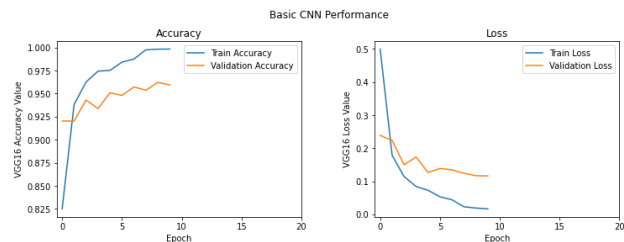
شکل (۲۸): نمودارهای عملکردی الگوریتم ResNet50



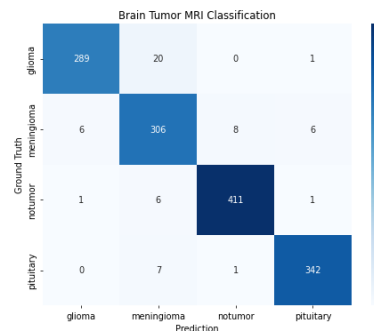
شکل (۲۹): ماتریس در هم‌ریختگی الگوریتم ResNet50

جدول (۶): عملکرد الگوریتم ResNet50

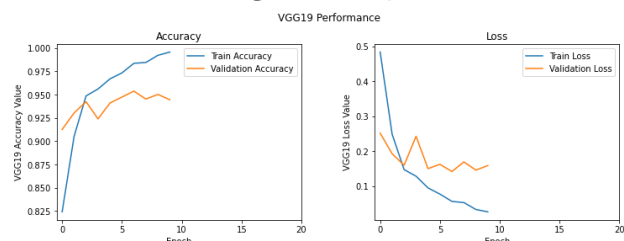
F1 Score	Specificity %	Precision %	کلاس
77	68	89	گلیوما
74	70	78	مننژیوم
96	98	93	بدون تومور
88	99	79	هیپوفیز



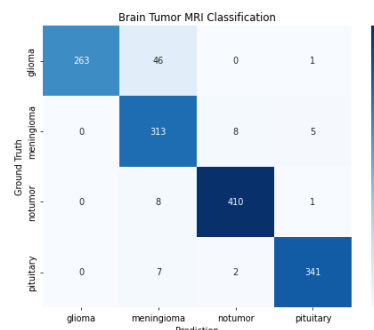
شکل (۲۴): نمودارهای عملکردی الگوریتم VGG16



شکل (۲۵): ماتریس در هم‌ریختگی الگوریتم VGG16



شکل (۲۶): نمودارهای عملکردی الگوریتم VGG19



شکل (۲۷): ماتریس در هم‌ریختگی الگوریتم VGG19

جدول (۴): عملکرد الگوریتم VGG16

F1 Score	Specificity %	Precision %	کلاس
95	93	98	گلیوما
92	94	90	مننژیوم
98	98	98	بدون تومور
98	98	98	هیپوفیز

جدول (۵): عملکرد الگوریتم VGG19

F1 Score	Specificity %	Precision %	کلاس
92	85	1	گلیوما
89	96	84	مننژیوم
98	98	98	بدون تومور
98	97	98	هیپوفیز



۴- نتیجه‌گیری

بخش‌بندی تصاویر پزشکی و به‌ویژه تقسیم‌بندی تصاویر مغز باید بسیار دقیق و دقیق‌تر از سایر الگوریتم‌های تقسیم‌بندی تصاویر غیرپزشکی باشد. با تجزیه و تحلیل داده‌های پزشکی به کمک الگوریتم‌های یادگیری عمیق می‌توان نتیجه‌گیری نمود، اگرچه این الگوریتم‌ها چالش‌های زیادی در ارتباط با معرفی یادگیری عمیق در محیط‌های بالینی دارند، اما جزء اولین گزینه‌های انتخابی بوده و روش‌های مختلف یادگیری عمیق به دلیل قدرت بیان فوق‌العاده آن‌ها، نتایج ارزشمندی را تولید می‌کنند. یادگیری عمیق در بررسی تصاویر پزشکی تحت تأثیر پیشرفت‌های بنیادی رایانه‌ای قرار گرفته است. مسابقه‌های بررسی تصاویر پزشکی از جمله مسابقه MICCA و دیگر مسابقات، محققان بیشتری را از جامعه بنیادی رایانه به حوزه بررسی تصاویر پزشکی ترغیب نموده است. بهبودهای بیشتر در شبکه‌های عمیق، زمان آموزش این الگوریتم‌های سه بعدی را کاهش می‌دهد.

در مقاله حاضر تشخیص و طبقه‌بندی تصاویر MRI مورد بررسی قرار گرفت و مشخص شد در مرحله تشخیص الگوریتم‌های VGG16 و ResNet50 با دقت ۹۹/۳۳ بیشترین دقت را داشتند و کمترین دقت در این مرحله به الگوریتم GoogLeNet با دقت ۸۸/۳۳ مربوط می‌شد. در مرحله طبقه‌بندی الگوریتم VGG16 بالاترین دقت را در بین الگوریتم‌ها با مقدار ۹۶ درصد داشت و کمترین دقت مربوط به الگوریتم ResNet50 با دقت ۸۵ درصد بود. لازم به ذکر است این نتایج مربوط به مجموعه داده‌های مشخص شده بوده و ممکن است با تغییر مجموعه داده تغییر نماید. اما مسلم است که الگوریتم VGG16 نسبت به سایر الگوریتم‌ها در هر دو مجموعه داده عملکرد بهتری را ارائه نموده است. همانگونه که در بخش ۳-۲-۶ نیز بیان گردید، تفاوت عملکردی و به ویژه عملکرد بهینه الگوریتم VGG16 را باید به واسطه معماری منحصر به فرد آن و وجود ۱۶ لایه کانولوشنی این الگوریتم دانست. هر چند وجود ۱۲۸ میلیون پارامتر شبکه باعث روند کند در آموزش و تشخیص ویژگی‌ها توسط این الگوریتم می‌شود. باید در نظر داشت دقت بالای این الگوریتم دیگر کمبودها را پوشش می‌دهد. در ضمن می‌توان بیان نمود بهینه‌ترین الگوریتم که با دقت بالا سرعت بیشتری نسبت به VGG16 ارائه می‌دهد، الگوریتم ResNet50 است. این الگوریتم به واسطه وجود ۱۵۲ لایه و استفاده از تکنیک اتصالات پرش دقت بالا و سرعت بهینه را به صورت همزمان ارائه می‌دهد.

مراجع

- [1] A. E. Minarno, M. H. C. Mandiri, Y. Munarko, and H. Hariyady, "Convolutional Neural Network with Hyperparameter Tuning for Brain Tumor Classification," *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, vol. 6, no. 2, pp. 127–132, May 2021, doi: 10.22219/KINETIK.V6I2.1219.
- [2] D. H. Gutmann and H. Kettenmann, "Microglia/Brain Macrophages as Central Drivers of Brain Tumor Pathobiology," *Neuron*, vol. 104, no. 3, pp. 442–449, Nov. 2019, doi: 10.1016/J.NEURON.2019.08.028.
- [3] T. Nejo, A. Mende, and H. Okada, "The current state of immunotherapy for primary and secondary brain tumors: similarities and differences," *Japanese Journal of Clinical Oncology*, vol. 50, no. 11, pp. 1231–1245, Oct. 2020, doi: 10.1093/JJCO/HYAA164.
- [4] W. Tomaszewski, L. Sanchez-Perez, T. F. Gajewski, and J. H. Sampson, "Brain tumor microenvironment and host state: Implications for immunotherapy," *Clinical Cancer Research*, vol. 25, no. 14, pp. 4202–4210, Jul. 2019, doi: 10.1158/1078-0432.CCR-18-1627.
- [5] C. K. Kuhl, "Abbreviated Magnetic Resonance Imaging (MRI) for Breast Cancer Screening: Rationale, Concept, and Transfer to Clinical Practice," *Annu Rev Med*, vol. 70, pp. 501–519, Jan. 2019, doi: 10.1146/ANNUREV-MED-121417-100403.
- [6] H. H. Sultan, N. M. Salem, and W. Al-Atabany, "Multi-Classification of Brain Tumor Images Using Deep Neural Network," *IEEE Access*, vol. 7, pp. 69215–69225, 2019, doi: 10.1109/ACCESS.2019.2919122.
- [7] A. Gumaei, M. M. Hassan, M. R. Hassan, A. Alelaiwi, and G. Fortino, "A Hybrid Feature Extraction Method with Regularized Extreme Learning Machine for Brain Tumor Classification," *IEEE Access*, vol. 7, pp. 36266–36273, 2019, doi: 10.1109/ACCESS.2019.2904145.
- [8] S. Deepak and P. M. Ameer, "Brain tumor classification using deep CNN features via transfer learning," *Computers in Biology and Medicine*, vol. 111, p. 103345, Aug. 2019, doi: 10.1016/J.COMPBIOMED.2019.103345.
- [9] "Brain_tumor_Dataset" https://figshare.com/articles/dataset/brain_tumor_dataset/1512427 (accessed Jun. 09, 2022).
- [10] Z. N. K. Swati et al., "Brain tumor classification for MR images using transfer learning and fine-tuning," *Comput Med Imaging Graph*, vol. 75, pp. 34–46, Jul. 2019, doi: 10.11016/J.COMPMEDIMAG.2019.05.001.
- [11] N. Noreen, S. Palaniappan, A. Qayyum, I. Ahmad, M. Imran, and M. Shoaib, "A Deep Learning Model Based on Concatenation Approach for the Diagnosis of Brain Tumor," *IEEE Access*, vol. 8, pp. 55135–55144, 2020, doi: 10.1109/ACCESS.2020.2978629.
- [12] H. Hairani, K. E. Saputro, and S. Fadli, "K-means-SMOTE untuk menangani ketidakseimbangan kelas dalam klasifikasi penyakit diabetes dengan C4.5, SVM, dan naive Bayes," *Jurnal Teknologi dan Sistem Komputer*, vol. 8, no. 2, pp. 89–93, Apr. 2020, doi: 10.14710/JTSSISKOM.8.2.2020.89-93.
- [13] S. H. Wang and Y. D. Zhang, "DenseNet-201-Based Deep Neural Network with Composite Learning Factor and Precomputation for Multiple Sclerosis Classification," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 16, no. 2s, Jun. 2020, doi: 10.1145/3341095.
- [14] S. Deepak and P. M. Ameer, "Brain tumor classification using deep CNN features via transfer learning," *Computers in Biology and Medicine*, vol. 111, p. 103345, Aug. 2019, doi: 10.1016/J.COMPBIOMED.2019.103345.
- [15] N. Noreen, S. Palaniappan, A. Qayyum, I. Ahmad, M. Imran, and M. Shoaib, "A Deep Learning Model Based on Concatenation Approach for the Diagnosis of Brain Tumor," *IEEE Access*, vol. 8, pp. 55135–55144, 2020, doi: 10.1109/ACCESS.2020.2978629.



- [16] M. C. Xena and Z. Wang, "Brain Tumour Detection and Classification using Multi-Level Ensemble Transfer Learning in MRI Dataset," *icABCD 2021 - 4th International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems, Proceedings*, Aug. 2021, doi: 10.1109/ICABCD51485.2021.9519361.
- [17] N. M. Dipu, S. A. Shohan, and K. M. A. Salam, "Deep Learning Based Brain Tumor Detection and Classification," *2021 International Conference on Intelligent Technologies, CONIT 2021*, Jun. 2021, doi:10.1109/CONIT51480.2021.9498384.
- [18] R. M. Prakash and R. S. S. Kumari, "Classification of MR brain images for detection of tumor with transfer learning from pre-trained CNN models," *2019 International Conference on Wireless Communications, Signal Processing and Networking, WiSPNET 2019*, pp.508–511, Mar. 2019, doi: 10.1109/WISPNET45539.2019.9032811.
- [19] Y. Guo, S. Liu, Z. Li, X. Shang, "BCDForest: a boosting cascade deep forest model towards the classification of cancer subtypes based on gene expression data." *BMC Bioinform* 19(5),118 (2018)
- [20] S. Athisayamani, R.S. Antonyswamy, V. Sarveshwaran, M. Almeshari, Y. Alzamil, V. Ravi, "Feature Extraction Using a Residual Deep Convolutional Neural Network (ResNet-152) and Optimized Feature Dimension Reduction for MRI Brain Tumor Classification." *Diagnostics* 2023, 13, 668. <https://doi.org/10.3390/diagnostics13040668>
- [21] S.Z. Kurdi, M.H. Ali, M.M. Jaber, T. Saba, A. Rehman, R. Damaševičius. " Brain Tumor Classification Using Meta-Heuristic Optimized Convolutional Neural Networks." *J. Pers. Med.* 2023, 13, 181. <https://doi.org/10.3390/jpm13020181>
- [22] M.A. Haq, I. Khan, A. Ahmed, S.M. Eldin, A. Alshehri, N.A. Ghamry." DCNNBT: A novel deep convolution neural network based on brain tumor classification model." *Fractals* 2023.
- [24] B.S., Abd El-WahabNasr, S. Khamis, et al. " BTC-fCNN: Fast Convolution Neural Network for Multi-class Brain Tumor Classification." *Health Inf Sci Syst* 11, 3 (2023). <https://doi.org/10.1007/s13755-022-00203-w>

