



دانشکده فنی و مهندسی - گروه کامپیوتر

پایان نامه کارشناسی

تشخیص بیماران آلزایمر از انسان های سالم از طریق پردازش تصاویر **MRI** مغزی آن ها با
استفاده از شبکه های عصبی **pre-train** شده

دانشجو: هدیه ریسی

استاد راهنما: دکتر علیرضا رضوانیان

نیم سال تحصیلی دوم ۱۴۰۱-۱۴۰۲

فهرست

1.....	فهرست	
2.....	چکیده	
2.....	مقدمه	
3.....	بیان مسئله	
3.....	پیشینه پژوهش	
7.....	روش پژوهش و پیاده‌سازی	
7.....	طراحی مدل	۱-۱
7.....	انتخاب نوع ورودی	۱-۲
8.....	مجموعه داده‌ها	۱-۳
8.....	ارزیابی	
8.....	معیارهای ارزیابی	۱-۴
9.....	تجزیه تحلیل و نتایج	۱-۵
10.....	نتایج	

چکیده

در چند سال اخیر تحقیقات انجام شده و مقالات مختلف کاربردهای وسیعی از شبکه های عصبی در تشخیص بیماری مختلف را نشان داده اند. برخی بیماری ها مانند آلزایمر که درمان خاصی برای آن ها به صورت کلی کشف نشده و تنها می توان از روند پیشرفت آن ها جلوگیری کرد، زمان تشخیص نقش مهمی را در روند بیماری ایفا می کند. در این پروژه سعی شده مدلی با استفاده از مدل های شبکه عصبی پیش تمرین داده شده (pre-train) استفاده شده تا با استفاده از روش های مختلفی که به طور مفصل در متن به آن ها اشاره شده مدلی را به وجود بیاوریم که قابلیت تشخیص بیماران آلزایمری را از انشان های سالم داشته باشد. در این پروژه از Structural MRI استفاده شده که از منبع داده ADNI به دست آمده است. تعداد افرادی که تصویر مغز آن ها در این پروژه استفاده شده شامل ۲۹۴ بیمار آلزایمری و ۲۹۶ انسان سالم می باشد. لازم به ذکر است که آخرین MRI ثبتی در ADNI برای هر فرد استفاده شد و یک مدل های ما (ViT) توانست به دقتی بیش از ۷۰ درصد در تشخیص دست بیابد که در کارهای پزشکی به خصوص با این تعداد کم از داده عدد خوبی محسوب می شود.

واژگان کلیدی: MRI, Deep SVM, Transfer learning, ResNet۱۸, Vision Transformers, شبکه عصبی

مقدمه

بیماری آلزایمر، یک اختلال عصبی پیشرو (neurodegenerative)، به طور قابل توجهی میلیون ها نفر را در سراسر جهان تحت تاثیر قرار می دهد. ایران بیش از یک میلیون مبتلا به آلزایمر داریم که البته برخی هم ممکن است شناسایی نشده باشند و در آمار ما هم نیایند. در ایران هر ۷ دقیقه یک نفر به آلزایمر مبتلا می شود و هزینه ماهانه یک فرد مبتلا به دمانس بین ۵ تا ۱۵ میلیون تومان بسته به مرحله ای که بیمار در آن قرار دارد برآورد می شود.

در نتیجه تشخیص زودهنگام و دقیق برای مدیریت موثر و برنامه ریزی درمانی بسیار مهم است. با این حال، روش های تشخیصی مرسوم با چالش هایی در حساسیت و ویژگی مواجه هستند که بر نیاز به تکنیک های پیشرفته تأکید می کند. این پروژه با هدف استفاده از قدرت هوش مصنوعی در تصویربرداری پزشکی، به ویژه تمرکز بر طبقه بندی تصاویر MRI برای متمایز کردن بیماران آلزایمر از افراد سالم است. این پروژه با استفاده از مدل های از پیش آموزش دیده، یعنی ResNet۱۸ و Vision Transformer که به دلیل کارایی خود در وظایف تشخیص تصویر مشهور هستند، به دنبال معرفی یک رویکرد جدید در تشخیص آلزایمر است.

ResNet۱۸، که از معروف ترین و کاربردی ترین معماری مدل های یادگیری عمیق می باشد، و Vision

Transformer، مدلی که معماری ترانسفورماتور را برای تجزیه و تحلیل تصویر به کار می برد، به دلیل قابلیت های

اثبات شده شان در سناریوهای طبقه بندی تصویر پیچیده انتخاب شدند. ادغام این مدل ها در فرآیند تشخیصی نوید افزایش دقت را می دهد، بنابراین تشخیص زودهنگام را تسهیل می کند و به تدوین استراتژی های درمانی هدفمند کمک می کند.

بیان مسئله

در این پروژه، سعی شده تا با شیوه های مختلف ورودی های مختلف را امتحان کرده و با آزمایش روش های مختلف بهترین روش را برای تشخیص افراد بیمار از سالم اتخاذ کنیم. در ابتدا داده ها را دانلود کردیم سپس آن ها را با استفاده از نرم افزار fast surfer که نرم افزار بسیار معروفی در زمینه پیش پردازش داده های MRI می باشد، پیش پردازش کردیم که شامل نرمال سازی تصاویر، صاف کردن تصاویر و جدا سازی مجسمه (skull scripting)، جداسازی قشر های خاکستری (Gray Matter) و قشر سفید (White Matter) و مایع مغزی نخاعی (CSF) را انجام دادیم. بعد از این ها تصمیم گرفتیم که روش های مختلف محاسباتی برای طبقه بندی این تصاویر ۳ بعدی مغزی را دریابیم. مهم ترین مسئله ای که ما با آن مواجه بودیم انتخاب نوع مدل، نحوه ی دادن ورودی و انتخاب فیچرها برای ورودی است. چون تصاویر MRI جزو تصاویر پیچیده هستند ما تصمیم گرفتیم که از شبکه های عصبی ژرف (Deep Neural Networks) استفاده کنیم. همچنین چون تعداد داده های ما برای تمرین دادن یک شبکه ی عصبی از ابتدا کم بود تصمیم گرفتیم که از شبکه های عصبی از پیش تمرین داده شده (pre-train model) که به اصطلاح به این شیوه transfer learning گفته می شود استفاده کنیم. نوع انتخاب مدل، نحوه ی دادن ورودی با توجه به این که داده های ما ۳ بعدی بودند و مانند تصاویر عادی از ۳ کانال RGB تشکیل نشده بودند و همچنین انتخاب هایپر پارامترهای دیگر مانند ضریب یادگیری و نحوه ی نرمال سازی داده و مسائلی از این دست جزو چالش هایی بود که در این پروژه با آن مواجه بودیم.

پیشینه پژوهش

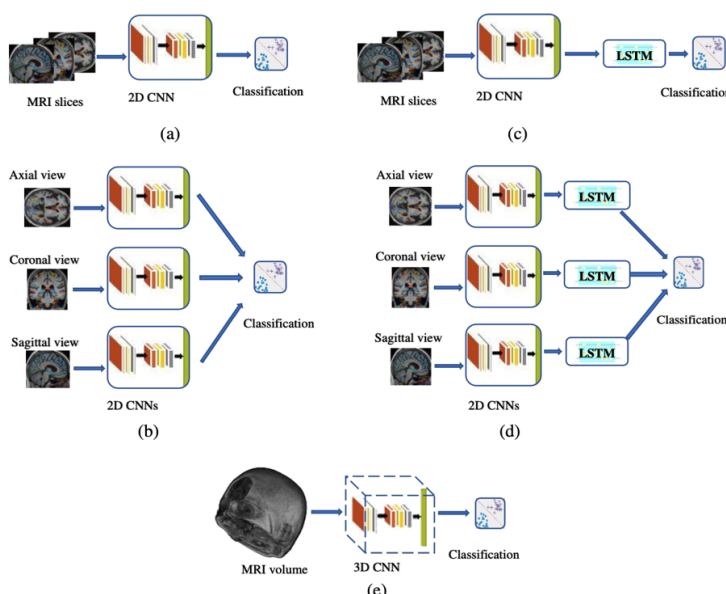
طبق مقالات مطالعه شده بیشتر کسانی که عنوان پروژه مشابهی داشتند از روش هایی که بسار معروف هستند اما با جزئیات متفاوتی استفاده کرده بودند. عمده ی مقالات در مواجهه با تصاویر ۳ بعدی MRI دو روش را اتخاذ کرده بودند:

1) روش اول که از مدل هایی که روی تصاویر ۳ بعدی تمرین داده شده بودند و شبکه های عصبی ۳ بعدی مانند resnet۱۸ که روی تصاویر ۳ بعدی pre-train شده بود

2) روش دوم که در آن پژوهشگران تصاویر سه بعدی را از ۳ جهت به تصاویر دو بعدی تبدیل کرده بودند و از مدل های pre-train معمول که برای تصاویر دو بعدی بود استفاده شده بود.

در تصویر بعدی می توانیم نمونه ای از کار های انجام شده در یک مقاله انجام شده را مشاهده می کنید.[۱]

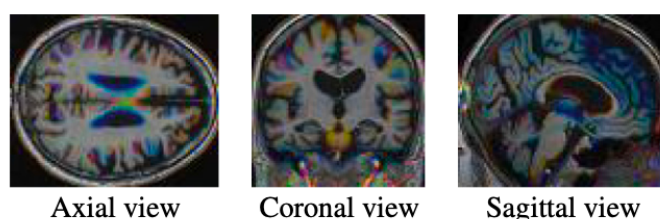
Ebrahimi and Luo: Convolutional neural networks for Alzheimer's disease detection on MRI images



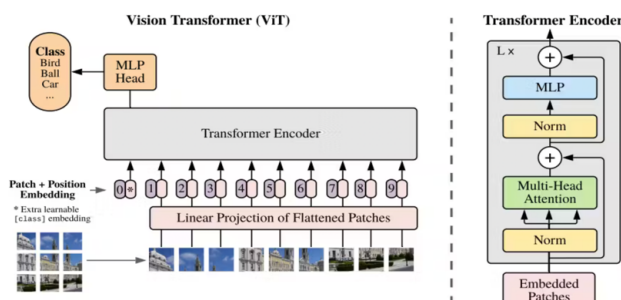
در تصویر همانطور که مشاهده می کنید روش های مختلفی اتخاذ شده که به طور مرسوم در اکثر مقالات همین گونه انجام شده است. پس از انتخاب این که تصاویر را تبدیل به ۲ بعدی کنیم مرحله ی بعدی این بود که چه چیزی را مورد مطالعه قرار دهیم و چگونه این تصاویر را به عنوان ورودی به مدل ها ارائه دهیم. همانطور که ممکن است بدانید مدل های pre-train بر روی دیتاست ImageNet تمرین داده شده اند. این تصاویر معمولی هستند که دارای سه کانال رنگی R,G,B می باشند. یکی از مشکلات بزرگ کار با این داده ها تبدیل آن ها به گونه ای که با این مدل ها همخوانی داشته باشند می باشد. همچنین انتخاب این که کدام لایه ها را به عنوان دیتا انتخاب کنیم که بیشترین اطلاعات را حاوی باشد هم از مشکلاتی بود که در مقالات مختلف جستجو کردیم. در مقاله اشاره شده کار عجیبی که صورت گرفته شده بود این بود که بعد از پیش پردازش، ابعاد هر $79 \times 95 \times 79$ MRI بود. برای CNN های سه بعدی، کل MRI از پیش پردازش شده برای آموزش استفاده شد. برای CNN های ۲ بعدی، برش های ۲ بعدی از هر نمای MRI (sagittal,coronal,axial) با دور انداختن اولین و آخرین برش های تصویر استخراج شد. با چیدن سه برش مجاور به عنوان کانال های رنگی RGB، برش های MRI باقی مانده به ۲۴ تصویر RGB از ناحیه ی Coronal، ۱۹ تصویر RGB از ناحیه ی Sagittal و ۱۶ تصویر RGB از ناحیه ی Axial تبدیل شدند. در برخی مقالات معروف دیگر باور داشتند که لایه های وسط بیشترین اطلاعات را در خود ذخیره دارند. در برخی از مقالات دیگر از میانگین چندین لایه ی وسط استفاده شده بود. همچنین در برخی مقالات دیگر چون بیماری آلزایمر تصور می رود که روی قشر خاکستری مغز بیشترین اثر را می گذارد تنها از قشر خاکستری برای تحلیل و پردازش استفاده شده است. در ادامه در مطالعات پژوهشی یافتیم مدلی که به عنوان مدل از پیش تمرین داده شده که در

بیشتر مقالات استفاده شده است ۱۸ ResNet می باشد. در مقاله ی ابراهیمی و همکاران هایپر پارامترهایی که در طراحی مدل خود نیز استفاده کرده بودند. کمی کار را برای ما راحت تر کرده بود.

Ebrahimi and Luo: Convolutional neural networks for Alzheimer's disease detection on MRI images



اما در مورد Vision Transformer ها می توان گفت که ساختار Transformer ها این روزها در بین پژوهشگران مورد توجه قرار گرفته است. این ساختار که با حضور در LLM ها و معروف ترین آن ها Chat GPT باعث شد تا پژوهشگران این مدل را برای تصاویر نیز استفاده کنند و اینگونه بود که Vision transformer برای پردازش تصاویر به وجود آمد. این معماری مانند ۱۸ ResNet دارای مدل pre-train می باشد که ما از آن استفاده کردیم. این مدل مانند ۱۸ ResNet نیز روی ImageNet تمرین داده شده است. در ادامه به شکل کلی این دو معماری می پردازیم.



An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale

این شکل معروف است که معمولاً برای توضیح Vision transformer در جاهای مختلف استفاده می شود. در ادامه به توضیحات این معماری به خصوص می پردازیم.

توضیح معماری Vision Transformer:

به طور کلی می‌توانیم بگوییم این مدل شامل چند قسمت می‌باشد:

1) پچ + جاسازی موقعیت (Patch + Position Embedding)

- تصویر ابتدا به یک سری تکه‌های با اندازه ثابت تقسیم می‌شود.
- سپس این تکه‌ها صاف می‌شوند و به صورت خطی نمایش داده می‌شوند (به یک بردار یک بعدی از ویژگی‌ها تبدیل می‌شوند).
- تعبیه‌های موقعیتی به تکه‌های مسطح اضافه می‌شوند تا اطلاعات مربوط به مکان هر وصله را در تصویر اصلی حفظ کنند، زیرا معماری ترانسفورماتور خود هیچ مفهومی از ترتیب یا توالی ندارد.

2) Transformer Encoder

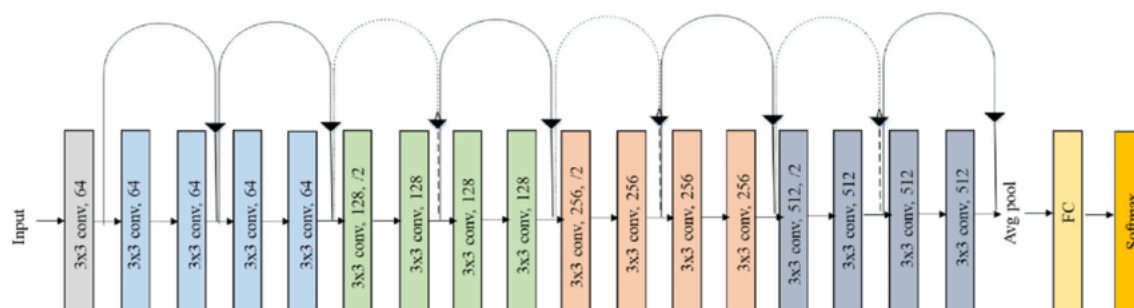
- سپس توالی تکه‌های تعبیه شده (منظور پچ‌هاست) به یک رمزگذار ترانسفورماتور (Transformer Encoder) وارد می‌شود.
- رمزگذار ترانسفورماتور از چندین لایه تشکیل شده است (که در اینجا به تفصیل نشان داده نشده است) که هر یک شامل دو لایه فرعی است:
- ۱. یک مکانیسم خودتوجهی چند سر (multi-head self-attention mechanism) که به مدل اجازه می‌دهد اهمیت بخش‌های مختلف تصویر را هنگام پیش‌بینی اندازه‌گیری کند.
- ۲. یک شبکه فید‌فوروارد (MLP: Multi-Layer Perceptron).
- یک مرحله نرمال‌سازی (Norm) بعد از هر لایه فرعی و یک اتصال باقیمانده (اضافه شدن ورودی به خروجی هر زیر لایه) قبل از نرمال‌سازی وجود دارد.
- این فرآیند به مدل اجازه می‌دهد تا وابستگی‌های دوربرد را مدیریت کند و زمینه جهانی تصویر را درک کند.

3) MLP Head

- سپس خروجی رمزگذار ترانسفورماتور از یک سر پرسپترون چند لایه (MLP) عبور داده می‌شود.
- این سر MLP معمولاً برای کارهای طبقه‌بندی استفاده می‌شود. اطلاعات رمزگذاری شده را از ترانسفورماتور می‌گیرد و یک پیش‌بینی برای کلاس تصویر (به عنوان مثال، پرنده، توپ، ماشین و غیره) خروجی می‌دهد.

Class Token (4

- علاوه بر پچ های تعبیه شده، یک «توکن کلاس» (که با «*» مشخص می‌شود) نیز در کنار پچ ها به رمزگذار ترانسفورماتور وارد می‌شود.
- نشانه کلاس برای تجمیع اطلاعات در کل دنباله پچ ها استفاده می شود و نمایش نهایی آن (پس از عبور از رمزگذار ترانسفورماتور) توسط سر MLP برای طبقه بندی نهایی استفاده می شود.



شکل کلی معماری ResNet18

معماری استفاده شده نیز در این تحقیق همانند این می باشد با این تفاوت که لایه ها منجمد شدند و لایه هایی جدید به آخر مدل اضافه شده اند.

روش پژوهش و پیاده‌سازی

۱-۱ طراحی مدل

طراحی مدل نقش مهمی را در نتیجه ایفا می کند. باتوجه به این که منابع پردازی ما کم بود و توانایی دانلود داده بیشتری نداشتیم پس محدودیت وجود داشت. در این راستا با توجه به این شرایط بهترین انتخاب استفاده از الگوریتم های یادگیری عمیقی بود که قبلا روی داده های وسیع دیگری آموزش دیده باشند. دو سری مدل معروف Vision transformer ها و ResNet ۱۸ به عنوان معماری های منتخب با توجه به پیشینه ی پژوهشی به دست آمده انتخاب شدند. در مدل ResNet۱۸ تمامی لایه ها freeze شدند تا وزن آن ها با آموزش جدید تغییر نکند. لایه هایی در آخر به آن ها اضافه شد. سه لایه خطی به ترتیب با ورودی ۵۱۲ و خروجی ۱۲۸، با ورودی ۱۲۸ و خروجی ۲ و یک لایه الگوریتم softmax در آخر برای طبقه بندی استفاده شدند. در این بین از لایه های activation با تابع Relu استفاده شد. همچنین برای جلوگیری از مشکل overfit در میان هر کدام از لایه ها به ترتیب یک dropout ۲۰ درصدی و ۱۰ درصدی صورت گرفت. در مورد معماری Vision Transformer ها نیز تمامی لایه ها به جز MLP، به صورت freeze قرار گرفته اند.

۱-۲ انتخاب نوع ورودی

یکی از مهم ترین مسائل و چالش های این پژوهش نحوه ی دادن MRI بیماران به عنوان ورودی به مدل های طراحی شده بود. از آنجایی که مدل ها بر روی دیتاست ImageNet آموزش داده شده اند، ورودی عکس دو بعدی سه

کاناله‌ی نرمال دارای کانال‌های رنگی RGB را تحویل می‌گیرند. اما MRI بیمارها را همگی به صورت ۳ بعدی و در دارای ابعاد مختلف با توجه به اسکتری که در آن برای گرفتن تصاویر بودند متفاوت بود. در اینجا ما MRI بیمارها را در ۳ بعد مختلف برش دادیم و آن‌ها را به عکس‌های دو بعدی‌ای که قبلاً به آن‌ها اشاره کرده بودیم تبدیل کردیم. سپس لایه‌ی وسط آن را به عنوان عکس‌های ورودی در نظر گرفتیم. لازم به ذکر است که در مرحله‌ی پیش پردازشی که قبل‌تر به آن اشاره کرده بودیم، وزن‌های عکس‌های MRI را نرمال کردیم و علاوه بر جداسازی استخوان جمجمه از عکس‌ها، ۳ ناحیه‌ی قشر خاکستری (Gray Matter) و قشر سفید (White Matter) و مایع مغزی نخاعی (CSF) را از داخل عکس استخراج کردیم. در این تحقیق ما دو نوع خاص از ورودی را استفاده کردیم. یک نوع آن که لایه‌های استخراج شده ذکر شده از تصاویر را به عنوان کانال‌های رنگی به ورودی امان دادیم. و در حالت دیگر تنها از قشر خاکستری مغز (GM) به عنوان ورودی استفاده کردیم. به گونه‌ای که مقادیر لایه‌ی وسط را سه بار تکرار کردیم و به عنوان R, G, B به مدل دادیم. (مثل دادن یک عکس خاکستری به مدل).

۱-۳ مجموعه داده‌ها

مجموعه داده‌های مورد استفاده در این تحقیق از پایگاه داده دسترسی آزاد طرح تصویربرداری عصبی بیماری آلزایمر که به ADNI معروف است می‌باشد. این موسسه زیر مجموعه‌ی دانشگاه کالیفرنیا جنوبی (USC) می‌باشد و به صورت تخصصی بر روی آلزایمر کار می‌کند. بعد از پر کردن درخواست دسترسی داده شد. و داده‌ها در چند مرحله به صورت زیر دانلود شد و مورد پردازش قرار گرفت. لازم به ذکر است از آنجایی که هر بیمار چند MRI ممکن بود داشته باشد آخرین MRI انجام شده از آن‌ها انتخاب شد.

جدول مجموعه داده‌ها

نوع داده‌ها	تعداد داده‌ها
بیماران آلزایمری (AD)	۲۹۴
انسان‌های سالم – گروه کنترل	۳۰۰
مجموع داده‌ها	۵۹۴

ارزیابی

۱-۴ معیارهای ارزیابی

- دقت یا accuracy: در این تحقیق از شاخص بسیار معروف دقت برای ارزیابی مدل استفاده شده است که فرمول آن به صورت زیر می‌باشد.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

- شاخص F1 یا F1 Score : شاخص f1 یکی دیگر از شاخص های معروفی است که برای ارزیابی مدل های مختلف استفاده می شود. ما از هر دو شاخص در هر فولد برای برازیابی مدل ها استفاده کردیم تا بتوانیم به طور کامل آن ها را بررسی کنیم.

$$F1 \text{ Score} = \frac{TP}{TP + \frac{1}{2}(FP + FN)}$$

۱-۵ تجزیه تحلیل و نتایج

طبقه بندی داده های ما در دو مدل مختلف ذکر شده و در دو مدل ورودی در سه جهت مختلف عکس برداری بررسی شده است. یعنی در مجموع ۱۲ نوع مختلف از ورودی ها و مدل ها بررسی شد که داده ها در هر کدام به صورت K-fold ۵ تایی بررسی شد تا از سوگیری انتخاب داده جلوگیری شود.

جدول دقت برای داده تست

نوع برش تصویر	مدل VIT با GM	مدل VIT	مدل Resnet18 با GM	مدل ResNet18
sagital	۴۶٪	۴۷٪	۵۲٪	۴۴٪
coronal	۷۱٪	۶۷٪	۵۶٪	۵۰٪
axial	۶۲٪	۵۹٪	۵۹٪	۵۰٪

همانطور که مشاهده می شود بیشترین اطلاعات در برش های coronal و axial قرار دارد که با پیشینه ی پژوهشی ما مطابقت دارد. همانطور که قابل توجه است مدلی vision transformer ای که داده های لایه ی قشر خاکستری رت دریافت می کند توانسته است به دقت ۷۰ درصدی برسد. این اعداد بیان شده درون جداول میانگین فولدهای اجرا شده می باشد.

جدول f1 score برای داده تست

نوع برش تصویر	مدل VIT با GM	مدل VIT	مدل Resnet18 با GM	مدل ResNet18
sagital	۵۳٪	۵۴٪	۴۸٪	۴۲٪
coronal	۷۷٪	۷۳٪	۵۶٪	۴۲٪
axial	۶۹٪	۶۶٪	۵۴٪	۴۲٪

نتایج

تشخیص بیمار آلزایمری از انسان سالم یکی از مهم ترین دغدغه هایی است که بشر با آن مواجه است زیرا تشخیص هر چه سریع تر آن باعث جلوگیری از پیشرفت وسیع آن در مناطق مختلف مغز می شود که همین امر زمان تشخیص را حیاتی تر می سازد و در این تحقیق ما سعی کردیم با روش های محاسباتی نوین، که شامل استفاده از الگوریتم های پیشرفته یادگیری عمیق می شد، برنامه و پروژه ای را جلو ببریم که راهگشای این مشکل باشد. ما در این پروژه توانستیم با استفاده از یادگیری عمیق و پردازش MRI بیماران به دقت مطلوبی برسیم. لازم به ذکر است که تمامی پردازش ها بر روی سرور انجام شده که قابل استناد می باشد.

لینک گیتهاب نتایج پروژه:

https://github.com/HediyeRaisy/MRI_Classification