



Combining Convolutional Neural Network (CNN) and Grad-CAM for Parkinson's Disease Prediction and Visual Explanation

Reyhaneh Dehghan¹ , Marjan Naderan Tahan²  and Seyyed Enayatallah Alavi³ 

1. Msc of Artificial Intelligence, Department of Computer Engineering, Faculty of Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran, Email: r-dehghan@stu.scu.ac.ir
2. Corresponding Author, Associate Professor, Department of Computer Engineering, Faculty of Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran, Email: m.naderan@scu.ac.ir
3. Assistant Professor, Department of Computer Engineering, Faculty of Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran, Email: se.alavi@scu.ac.ir

Article Info

Article type:
Research Article

Article history:
Received 8 June 2024
Received in revised form 4
Sept 2024
Accepted 5 Sept 2024
Published online 21 Sept
2024

Keywords:
Convolutional neural
network (CNN), Data
augmentation, Grad-CAM,
Parkinson's disease, Spect
images.

ABSTRACT

Parkinson's disease is one of the types of neurological diseases that is caused by the destruction of brain cells that produce dopamine. Early detection of Parkinson's disease is an important factor in slowing the progression of the disease. In this study, a Convolutional Neural Network (CNN) namely ConvNet, is used to discriminate Parkinson's patients based on Single Photon Emission Computed Tomography (SPECT) images acquired from the PPMI database. Since the dataset is limited, after a pre-processing stage, two data augmentation techniques are used. Finally, the Grad-CAM technique is used to obtain visual interpretation from the predictions of the proposed CNN. To evaluate the proposed method, different measures such as accuracy, sensitivity (recall) and f1-score are used. Simulation results according to the measures shows that when the classic data augmentation method is used accuracy is increased to 98.50% and more efficient classification is performed.

Cite this article: Dehghan, R. & Others, (2024)., Combining Convolutional Neural Network (CNN) and Grad-CAM for Parkinson's Disease Prediction and Visual Explanation. *Engineering Management and Soft Computing*, 10 (1). 1-13. DOI: <https://doi.org/10.22091/jemsc.2024.10828.1180>



© The Author(s)
DOI: <https://doi.org/10.22091/jemsc.2024.10828.1180>

Publisher: University of Qom

ترکیب شبکه عصبی کانولوشن (CNN) و Grad-CAM برای پیش بینی و تفسیر پذیری بصری بیماری پارکینسون

ریحانه دهقان^۱، مرجان نادران طحان^۲ و سیدعنایت اله علوی^۳

۱. کارشناس ارشد هوش مصنوعی، گروه مهندسی کامپیوتر، دانشکده مهندسی، دانشگاه شهید چمران اهواز، اهواز، ایران. رایانامه: r-dehghan@stu.scu.ac.ir

۲. نویسنده مسئول، دانشیار گروه مهندسی کامپیوتر، دانشکده مهندسی، دانشگاه شهید چمران اهواز، اهواز، ایران. رایانامه: m.naderan@scu.ac.ir

۳. استادیار گروه مهندسی کامپیوتر، دانشکده مهندسی، دانشگاه شهید چمران اهواز، اهواز، ایران. رایانامه: se.alavi@scu.ac.ir

اطلاعات مقاله	چکیده
نوع مقاله: مقاله پژوهشی	بیماری پارکینسون یکی از انواع بیماری‌های عصبی است که در اثر تخریب سلول‌های مغزی تولیدکننده دوپامین ایجاد می‌شود. تشخیص زودهنگام بیماری پارکینسون عامل مهمی در کاهش سرعت پیشرفت بیماری است. در این مطالعه، از یک شبکه عصبی کانولوشن (CNN) به نام ConvNet برای طبقه‌بندی افراد سالم و افراد مبتلا به بیماری پارکینسون بر اساس تصاویر توموگرافی رایانه‌ای تک‌فوتونی (SPECT) از پایگاه داده PPMI استفاده شده است. از آنجاییکه این مجموعه داده محدود است، پس از یک مرحله پیش‌پردازش داده‌ها، از دو تکنیک افزایش داده کلاسیک برای جلوگیری از بیش‌برازش و عملکرد بهتر مدل استفاده شده است. در نهایت از تکنیک Grad-CAM جهت تفسیر پیش‌بینی‌های انجام‌شده توسط شبکه عصبی کانولوشن پیشنهادی استفاده شده است. برای ارزیابی روش پیشنهادی از معیارهای متنوعی همچون دقت، حساسیت و f1-score استفاده شده است. نتایج شبیه‌سازی برحسب معیارهای ارزیابی نشان می‌دهد که با استفاده از تکنیک افزایش داده کلاسیک می‌توان طبقه‌بندی موثرتری انجام داد، به گونه‌ای که روش پیشنهادی از نظر دقت طبقه‌بندی به دقت ۹۸.۵ درصد دست یافت.
تاریخ دریافت: ۱۴۰۳/۰۳/۱۹	
تاریخ بازنگری: ۱۴۰۳/۰۶/۱۴	
تاریخ پذیرش: ۱۴۰۳/۰۶/۱۵	
تاریخ انتشار: ۱۴۰۳/۰۶/۳۱	
کلیدواژه‌ها: Grad-CAM، افزایش داده، بیماری پارکینسون، تصاویر SPECT، شبکه عصبی کانولوشن	

استناد: دهقان، ریحانه؛ نادران طحان، مرجان و علوی، سیدعنایت‌اله. (۱۴۰۳). «ترکیب شبکه عصبی کانولوشن (CNN) و Grad-CAM برای پیش بینی و

تفسیر پذیری بصری بیماری پارکینسون». مدیریت مهندسی و رایانش نرم، دوره ۱۰ (۱). صص: ۱۳-۱.

<https://doi.org/10.22091/jemsc.2024.10828.1180>



© نویسنده‌گان.

ناشر: دانشگاه قم

(۱) مقدمه

بیماری پارکینسون، اختلالی پیشرونده است که بر سلول‌های عصبی قسمت‌های عمیق مغز به نام جسم سیاه تأثیر می‌گذارد. سلول‌های عصبی در این قسمت از مغز، مسئول تولید ماده شیمیایی به نام دوپامین هستند (اپینینی و گوپتا (۲۰۲۳)). دوپامین به عنوان یک پیام‌رسان بین بخش‌های مغز و سیستم عصبی عمل می‌کند و به کنترل و هماهنگی حرکات بدن کمک می‌کند. به دلایلی نامعلوم، در برخی از افراد سلول‌های عصبی تولیدکننده دوپامین در جسم سیاه از بین می‌روند. هنگامیکه ۸۰ درصد دوپامین از بین برود، علائم بیماری پارکینسون مشاهده می‌شوند (فائول و لاورتی (۱۹۶۹)). شیوع بیماری پارکینسون با افزایش سن بیشتر می‌شود بطوریکه این بیماری اغلب در افراد بالای ۶۰ سال تشخیص داده می‌شود. اگر این بیماری در سنین پایین‌تر تشخیص داده شود به آن پارکینسون زودرس می‌گویند که در افراد بین ۲۰ تا ۴۰ سال بروز می‌کند (پئوه و همکاران (۲۰۱۷)). علائم و نشانه‌های بیماری پارکینسون می‌تواند از فردی به فرد دیگر متفاوت باشد.

بسیاری از مردم فکر می‌کنند که علائم اولیه پارکینسون نشانه‌های طبیعی پیری است بنابراین کمتر به علائم آن توجه می‌کنند. علاوه بر علائم حرکتی که شامل لرزش در هنگام استراحت، مشکل در راه رفتن، کندی حرکت، سفتی عضلات، صدای نازک یا آرام، تغییرات نوشتاری و سایر علائم مشابه است، ممکن است علائم غیر حرکتی نیز وجود داشته باشد که شامل تغییرات خلقی، اختلال در بویایی و چشایی، اختلالات خواب، سرگیجه و غش، مشکلات دستگاه ادراری و یبوست است (ونگ و همکاران (۲۰۲۰)). بیماری‌های تخریب‌کننده سلول‌های عصبی درمان ندارند اما اگر این گونه بیماری‌ها به موقع تشخیص داده شوند، می‌توان این بیماری‌ها را کنترل کرد.

روش‌های مختلف تصویربرداری عصبی مانند تصویربرداری تشدید مغناطیسی (MRI)، تصاویر توموگرافی رایانه‌ای تک‌فوتونی (SPECT) و تصویربرداری توموگرافی انتشار پوزیترون (PET) برای تشخیص زودهنگام بیماری پارکینسون استفاده شده‌اند (وایله و همکاران (۲۰۱۷)). اسکن اسپکت یک نوع تصویربرداری پزشکی است که برای تشخیص بیماری‌های مختلف استفاده می‌شود. تحقیقات نشان می‌دهد در افرادی که به بیماری پارکینسون مبتلا شده‌اند، کاهش جذب نسبی ماده حاجب در پوتامن بیشتر از هسته دمی است و این موضوع ممکن است در شناسایی افراد مبتلا به پارکینسون با استفاده از تصاویر SPECT قبل از ظهور علائم بیماری مفید باشد. به همین دلیل در این مطالعه از تصاویر SPECT برای تشخیص بیماری پارکینسون استفاده شده است (مارک و همکاران (۱۹۹۶)).

از طرف دیگر روش‌های یادگیری ماشین و شبکه‌های عصبی عمیق نیز در سال‌های اخیر در تشخیص بیماری پارکینسون مورد استفاده قرار گرفته‌اند و به دقت‌های قابل‌قبولی در تشخیص این بیماری دست یافته‌اند (اپینینی و گوپتا (۲۰۲۳))، محمد و همکاران (۲۰۲۱) - هاتالیا و همکاران (۲۰۲۲)). از میان آنها روش‌های تشخیصی مختلفی براساس دستخط، سیگنال‌های صوتی، تصاویر MRI مغزی و تصاویر اسپکت مغزی ارائه شده‌است که روش‌های مبتنی بر تصاویر اسپکت کمتر کار شده‌است. در تحقیق حاضر، در ابتدا به نرمال‌سازی داده‌های موجود پرداخته شده و سپس با استفاده از تکنیک‌های مرسوم داده‌افزایی تصاویر، تصاویر جدیدی تولید شده‌اند. در ادامه طبقه‌بندی و شناسایی افراد سالم و افراد

مبتلا به بیماری پارکینسون با استفاده از شبکه عصبی کانولوشن انجام شده و در انتها، از روش Grad-CAM برای تفسیر پیش بینی های انجام شده توسط مدل پیشنهادی استفاده می شود. مهمترین نوآوری های این مقاله عبارتند از:

- استفاده از یک شبکه عصبی کانولوشن جدید برای طبقه بندی افراد مبتلا به بیماری پارکینسون با دقت بالا
- تولید تصاویر مصنوعی با استفاده از روش های افزایش داده بر اساس تصاویر موجود در مجموعه داده
- استفاده از روش Grad-CAM برای تفسیر پیش بینی های انجام شده توسط شبکه عصبی کانولوشن

پیشنهادی

در بخش دوم، مطالعات پیشین این حوزه به اجمال بررسی می شوند. در بخش سوم، روش پیشنهادی ارائه می شود. در بخش چهارم، معیارهای ارزیابی و نتایج شبیه سازی و مقایسه با روش های پیشین ارائه شده است. در نهایت در بخش پنجم، نتیجه گیری بیان خواهد شد.

۲) پیشینه تحقیق

همانطور که پیش تر ذکر شد، در سال های اخیر روش های مبتنی بر یادگیری ماشین و شبکه های عصبی عمیق در زمینه های مختلف تصویربرداری پزشکی از جمله تشخیص بیماری پارکینسون مورد استفاده قرار گرفته است. در ادامه تعدادی از کارهای پیشین که از این روش ها برای تشخیص افراد مبتلا به بیماری پارکینسون استفاده کرده اند، بطور خلاصه بررسی شده اند.

اپینینی و گوپتا (۲۰۲۳) از شبکه عصبی کانولوشن InceptionV3 برای طبقه بندی افراد سالم و افراد مبتلا به بیماری پارکینسون استفاده کردند. مجموعه داده استفاده شده شامل ۷۰۸ تصویر SPECT استخراج شده از پایگاه داده PPMI است و دقت مدل پیشنهادی ۷۴.۶۵ درصد گزارش شده است.

محمد و همکاران (۲۰۲۱) از شبکه عصبی کانولوشن با معماری AlexNet برای طبقه بندی افراد سالم و افراد مبتلا به بیماری پارکینسون استفاده کردند. مجموعه داده استفاده شده شامل تصاویر SPECT استخراج شده از پایگاه داده PPMI است. مدل پیشنهادی به دقت ۹۹.۳۴ درصد و حساسیت ۹۹.۰۴ درصد دست یافت.

گویندو و پالوه (۲۰۲۳) از ماشین بردار پشتیبان (SVM)، جنگل تصادفی (RF)، k نزدیکترین همسایه (KNN) و رگرسیون لجستیک (LR) برای تمایز بین افراد مبتلا به بیماری پارکینسون و افراد سالم استفاده کردند. مجموعه داده استفاده شده در این مطالعه، شامل سیگنال صوتی ۳۱ نفر است که از پایگاه داده PPMI جمع آوری شده است. نتایج نشان می دهد که طبقه بند جنگل تصادفی عملکرد بالاتری نسبت به سایر طبقه بندها با دقت تشخیص ۹۱.۸۳ درصد و حساسیت ۰.۹۵ درصد ارائه کرده است.

تونسر و همکاران (۲۰۲۰) از روش تجزیه مقادیر منفرد (SVD) برای استخراج ویژگی های برجسته از سیگنال های صوتی برای تشخیص بیماری پارکینسون استفاده کردند. آنها همچنین از الگوریتم KNN برای طبقه بندی افراد سالم و افراد مبتلا به بیماری پارکینسون استفاده کردند و مدل پیشنهادی به دقت ۹۲.۴۶ درصد منجر شد. با انجام پس پردازش برای به دست آوردن نتایج منحصر به فرد برای هر فرد، دقت را به ۹۶.۸۳ درصد افزایش دادند.

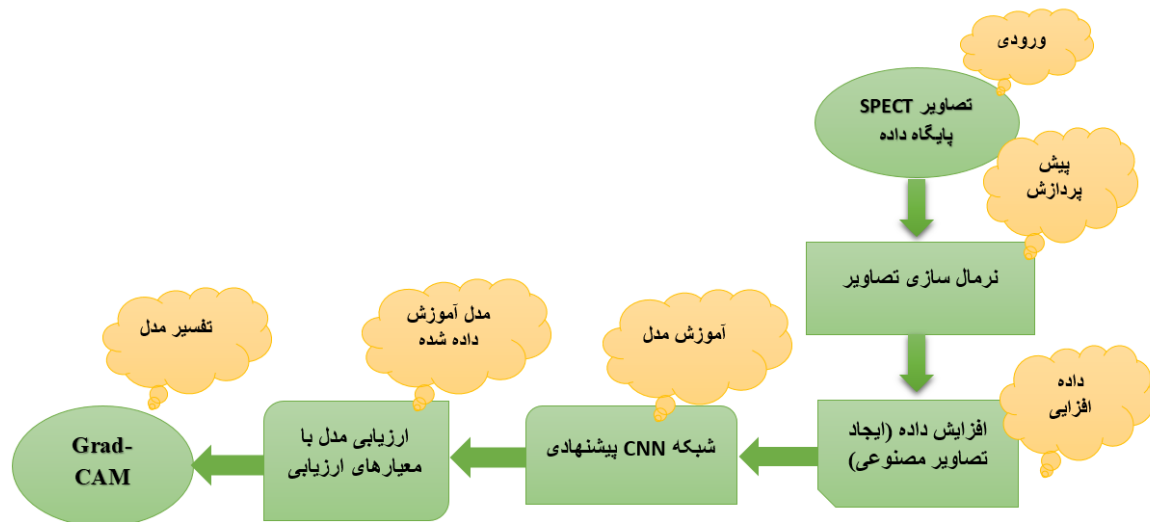
مگش و همکاران (۲۰۲۰) از شبکه عصبی کانولوشن با معماری VGG16 برای طبقه‌بندی افراد سالم و افراد مبتلا به بیماری پارکینسون استفاده کردند. مجموعه داده استفاده شده شامل ۶۴۲ تصویر SPECT استخراج شده از پایگاه داده PPMI است. دقت مدل پیشنهادی ۹۵.۲ درصد گزارش شده است.

هاتالیا و همکاران (۲۰۲۲) یک مدل مبتنی بر شبکه عصبی کانولوشن ارائه کردند که می‌تواند افراد مبتلا به بیماری پارکینسون را براساس تصاویر PSECT طبقه‌بندی کند. مدل پیشنهادی با یک مجموعه داده بزرگ با ۵۸۶۹۲ تصویر برای آموزش و ۱۱۳۳۸ تصویر برای اعتبارسنجی و ۷۸۲۶ تصویر برای تست، آموزش داده شده است. مدل پیشنهادی به دقت ۸۸.۹ درصد و حساسیت ۹۰.۱۲ درصد دست یافت.

همانطور که از کارهای پیشین مشاهده می‌شود، روش‌های تشخیصی مختلفی روی سیگنال‌های صوتی و تصاویر اسپکت وجود دارد. از آنجا که تشخیص از سیگنال‌های صوتی وقتی قابل بررسی است که فرد به پارکینسون مبتلا شده است، تشخیص زودهنگام بیماری برای آن امکان‌پذیر نیست. در مقابل تصاویر اسپکت مغزی در فازهای اولیه بیماری کاهش دوپامین را نشان می‌دهند و می‌توان از آن برای تشخیص زودهنگام بیماری استفاده کرد. بنابراین در سال‌های اخیر، محققان بیشتر به سمت استفاده از این تصاویر رفته‌اند و در این تحقیق نیز از همین تصاویر استفاده می‌شود.

(۳) روش تحقیق

نمودار ۱، نمودار روش پیشنهادی برای تشخیص افراد مبتلا به بیماری پارکینسون و طبقه‌بندی به دو کلاس افراد سالم و افراد بیمار را نشان می‌دهد.



نمودار ۱. فلوچارت روش پیشنهادی

(۳-۱) مجموعه داده

مجموعه داده مورد استفاده در این مطالعه از پایگاه داده PPMI جمع آوری شده است. پایگاه داده PPMI از سال ۲۰۱۰ برای ایجاد یک مجموعه داده با دسترسی باز و کتابخانه‌ای از نمونه‌های زیستی برای سرعت بخشیدن به پژوهش در حوزه

پارکینسون فعالیت می کند (مارک و همکاران (۲۰۱۱)). دیتاست استفاده شده در این مطالعه، شامل ۶۵۰ تصویر SPECT است که به دو دسته سالم (۲۱۳) و بیمار (۴۳۷) تقسیم شده اند. مجموعه داده بطور تصادفی به سه گروه آموزش و اعتبارسنجی و تست تقسیم شده است که ۸۰ درصد داده ها به عنوان داده های آموزش و ۱۰ درصد داده ها به عنوان داده های اعتبارسنجی و ۱۰ درصد به عنوان مجموعه داده تست در نظر گرفته شده است.

۳-۲) پیش پردازش داده ها

نرمال سازی داده یکی از گام های حیاتی در آماده سازی داده ها برای یادگیری ماشین و یکی از مهمترین مراحل پیش پردازش است. در پردازش تصویر، نرمال سازی فرآیندی است که محدوده مقادیر شدت پیکسل را تغییر می دهد. هدف این فرآیند در کاربردهای مختلف معمولاً تبدیل محدوده تصویر به محدوده ای است که با کل مجموعه تصاویر در مجموعه داده سازگارتر باشد. در واقع تغییرات با شدت بالا ممکن است به دلیل پارامترها و اسکنرهای مختلف استفاده شده در هنگام گرفتن تصویر رخ دهد و نرمال سازی تصاویر منجر به بهبود عملکرد مدل و افزایش سرعت یادگیری می شود (مارتینز و همکاران (۲۰۱۸)). در این مطالعه از روش نرمال سازی min-max استفاده شده که مقادیر شدت پیکسل ها حداقل صفر و حداکثر یک هستند.

۳-۳) افزایش داده

افزایش داده، فرآیند تولید مصنوعی داده های جدید از داده های موجود است تا تنوع و حجم داده های آموزشی افزایش یابد. در حوزه پردازش تصویر تکنیک هایی نظیر بزرگنمایی تصاویر، تغییر روشنایی تصاویر، اعمال نویز، چرخش تصویر برای تولید داده ها در شبکه های عصبی استفاده می شود. دلیل اصلی افزایش داده این است که داشتن یک پایگاه داده بزرگ برای حصول یک عملکرد خوب مدل حیاتی است. افزایش داده در Keras از طریق کلاس ImageDataGenerator انجام می پذیرد. این کلاس این امکان را فراهم می کند تا با تعریف پارامترهای مختلف مانند چرخش، انتقال، زوم، تصاویری متنوع از یک تصوی اصلی تولید شود (ماش و همکاران (۲۰۱۶)). در این مطالعه، به دلیل محدودیت تعداد داده ها، از تکنیک های معکوس کردن تصویر در جهت محور افقی و چرخش تصویر برای تولید تصاویر جدید استفاده شده است. با تعیین و مقداردهی `horizontal_flip=True` در تابع سازنده بطور تصادفی جهت برخی تصاویر از چپ به راست معکوس می شوند. با تعیین و مقداردهی `rotation_range = 20` در تابع سازنده بطور تصادفی برخی تصاویر ۲۰ درجه در جهت عقربه ساعت می چرخند.

۳-۴) شبکه کانولوشن پیشنهادی (ConvNet)

معماری شبکه کانولوشن پیشنهادی که پیش تر (دهقان و همکاران (۲۰۲۲)) ارائه کرده بودیم، تحت عنوان ConvNet پیشنهادی به صورت مختصر توضیح داده می شود: در لایه اول ۶۴ فیلتر کانولوشن با ابعاد ۷x۷، در لایه دوم ۶۴ فیلتر کانولوشن با ابعاد ۳x۳، در لایه سوم، پنجم و هفتم ۱۲۸ فیلتر کانولوشن با ابعاد ۳x۳، در لایه چهارم و هشتم Max- Pooling با ابعاد ۲x۲، در لایه نهم Flatten و در لایه آخر یک لایه کاملاً متصل با یک نورون و تابع فعال سازی سیگموئید

برای طبقه‌بندی استفاده شده است. در این مقاله از بهینه‌ساز Adagrad با نرخ یادگیری ۰/۰۱ و تابع هزینه Binary cross-entropy برای آموزش مدل پیشنهادی استفاده شده است. جزئیات مدل پیشنهادی در جدول ۱ ارائه شده است.

جدول ۱. جزئیات مدل ConvNet پیشنهادی

لایه ها	نوع	سایز پولینگ	سایز کرنل	گام	عمق	ابعاد خروجی
(۱)	Convolutional Layer	-	۷×۷	۲×۲	۶۴	۶۴×۵۵×۵۵
(۲)	Convolutional Layer	-	۳×۳	۱×۱	۶۴	۵۵×۵۵×۶۴
(۳)	Convolutional Layer	-	۳×۳	۱×۱	۱۲۸	۵۵×۵۵×۱۲۸
(۴)	Max-Pooling Layer	۲×۲	-	۲×۲	--	۲۷×۲۷×۱۲۸
(۵)	Convolutional Layer	-	۳×۳	۱×۱	۱۲۸	۲۷×۲۷×۱۲۸
(۶)	Max-Pooling Layer	۲×۲	-	۲×۲	--	۱۳×۱۳×۱۲۸
(۷)	Convolutional Layer	-	۳×۳	۱×۱	۱۲۸	۱۳×۱۳×۱۲۸
(۸)	Max-Pooling Layer	۲×۲	--	۲×۲	--	۶×۶×۱۲۸
(۹)	Flatten Layer	-	--	--	--	۴۶۰۸
(۱۰)	Fully connected	-	--	--	۲۵۶	۱×۲۵۶
(۱۱)	Fully connected	-	--	--	۱۲۸	۱×۱۲۸

۳-۵) تفسیر شبکه عصبی پیچشی با استفاده از Grad-CAM

شبکه‌های عصبی پیچشی، از جمله روش‌هایی هستند که به مدل‌های جعبه سیاه معروف هستند. به این معنی که در این دسته از روش‌ها، کاربر فقط از تعداد ورودی‌ها، معماری شبکه و تعداد خروجی‌های قابل تولید خبر دارد اما از مکانیزم‌های استفاده‌شده جهت تبدیل ورودی‌ها به خروجی‌ها و چگونگی تولید پیش‌بینی‌های نهایی مرتبط با داده‌های ورودی اطلاعی ندارد. برای تفسیر تصمیمات اتخاذشده توسط مدل‌های یادگیری عمیق در بینایی ماشین و کاربردهای مرتبط، مجموعه‌ای از تکنیک‌ها و ابزارها طراحی شده‌اند. در کاربردهای مختلف یادگیری عمیق مرتبط با تصویربرداری پزشکی، بسیار مهم است که مدل یادگیری عمیق را تفسیر کرد (ژو و همکاران (۲۰۱۶)). در این تحقیق، برای تفسیر خروجی‌ها یا پیش‌بینی‌های انجام‌شده توسط مدل پیشنهادی از تکنیک Grad-CAM استفاده شده است. این تکنیک در واقع مناطقی از تصویر ورودی را برجسته می‌کند که مدل در طی فرآیند طبقه‌بندی به آن مناطق توجه زیادی کرده است. در واقع برای تشخیص افراد سالم و افراد مبتلا به بیماری پارکینسون باید به جسم مخطط در تصاویر SPECT توجه کرد. تصاویر SPECT افراد مبتلا به بیماری پارکینسون یک ناحیه مخطط کوچکتر و گرد را نشان می‌دهند و در مقابل، تصاویر SPECT افراد سالم، شکل C بسیار روشن خود را حفظ می‌کنند.

مراحل تکنیک Grad-CAM به صورت زیر است:

۱- برای مشخص کردن پیکسل‌های مرتبط با هر کلاس خاص در تصاویر ورودی، کافی است تا گرادیان‌های تابع هدف نسبت به خروجی‌های لایه‌های کانولوشن با استفاده از (۱) محاسبه شوند (سلواراجو و همکاران (۲۰۱۷)):

$$p = \frac{\partial \gamma^c}{\partial A_{ij}^k} \quad (1)$$

که در آن Y^C تابع هدف، A^K خروجی‌های آخرین لایه کانولوشن (نقشه ویژگی) را نشان می‌دهد. بنابراین k تا نقشه ویژگی وجود دارد که هر نقشه ویژگی i سطر و j ستون دارد.

۲- برای به دست آوردن وزن نقشه‌های ویژگی، باید میانگین گرادیان‌ها محاسبه شود. بنابراین اگر k نقشه ویژگی وجود دارد، باید میانگین را برای هر k نقشه ویژگی مطابق (۲) محاسبه کرد (سلواراجو و همکاران (۲۰۱۷)):

$$= \frac{1}{Z} \sum_i \sum_j \frac{\partial Y^C}{\partial A_{ij}^K} \alpha_K^C \quad (2)$$

که در آن Z یک عدد ثابت است (تعداد پیکسل‌های موجود در نقشه فعال سازی است)

۳- یک نقشه حرارتی باید برای نقشه‌های ویژگی ترسیم شود و هر نقشه حرارتی باید در وزن‌های متصل به نرون «برنده» ضرب شود و سپس نتایج با هم جمع شوند. طبق (۳)، تابع فعال‌سازی ReLU باید استفاده شود زیرا فقط ویژگی‌هایی را در نظر می‌گیرد که تاثیر مثبتی بر کلاس مورد علاقه دارند (سلواراجو و همکاران (۲۰۱۷)).

$$= \text{ReLU} \left(\sum_K \alpha_K^C A^K \right) L_{\text{Grad-CAM}}^C \quad (3)$$

۴- در انتها باید از تکنیک افزایش نرخ نمونه برداری (Up sampling) به منظور افزایش اندازه Grad-CAM برای مطابقت با اندازه تصویر ورودی استفاده کرد. سپس Grad-CAM به دست آمده بر روی تصویر اصلی قرار داده می‌شود (سلواراجو و همکاران (۲۰۱۷)).

۴ یافته‌ها

۴-۱ معیارهای ارزیابی

یکی از مهمترین مراحل بعد از طراحی و ساخت یک مدل، ارزیابی عملکرد آن مدل است. ارزیابی مدل در یادگیری ماشین فرآیندی است که در آن عملکرد مدل با استفاده از معیارهای از پیش تعیین شده، ارزیابی می‌شود. در این مقاله از معیارهای زیر استفاده شده است:

۴-۱-۱ ماتریس درهم ریختگی (Confusion Matrix)

ماتریس درهم ریختگی خروجی و عملکرد کامل مدل را توصیف می‌کند. به این صورت که هر ردیف نشان‌دهنده نمونه‌ای از یک کلاس واقعی و هر ستون نشان‌دهنده نمونه‌ای از یک کلاس پیش‌بینی شده است. معمولاً چنین ماتریسی برای الگوریتم‌های یادگیری با ناظر استفاده می‌شود اگرچه در یادگیری بدون ناظر نیز کاربرد دارد. سطرها و ستون‌های این ماتریس عبارتند از (کریستینک و همکاران (۲۰۲۰)):

- a) مثبت واقعی (True Positive): نمونه عضو دسته مثبت باشد و عضو همین کلاس تشخیص داده شود.
- b) منفی واقعی (True Negative): نمونه عضو کلاس منفی باشد و عضو همین کلاس تشخیص داده شود.
- c) مثبت کاذب (False Positive): نمونه عضو کلاس منفی باشد و عضو کلاس مثبت تشخیص داده شود.
- d) منفی کاذب (False Negative): نمونه عضو کلاس مثبت باشد و عضو کلاس منفی تشخیص داده شود.

۴-۱-۲) دقت (Accuracy)

این معیار رایج‌ترین معیار کارآیی برای الگوریتم‌های طبقه‌بندی است و عبارتست از تعداد پیش‌بینی‌های صحیح صورت گرفته به نسبت همه پیش‌بینی‌های صورت گرفته است (کریستینک و همکاران (۲۰۲۰):

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FN+FP} \quad (۴)$$

۴-۱-۳) حساسیت (Recall یا sensitivity)

عبارتست از نسبت نمونه‌های مثبت واقعی (TP) به تمام نمونه‌های مثبت واقعی و مثبت‌هایی که به اشتباه منفی تشخیص داده شده‌اند (کریستینک و همکاران (۲۰۲۰):

$$\text{Recall} = \frac{TP}{TP+FN} \quad (۵)$$

۴-۱-۴) ویژگی (Specificity)

عبارتست از نسبت تعداد نمونه‌های منفی به تعداد کل نمونه‌های منفی پیش‌بینی شده (کریستینک و همکاران (۲۰۲۰):

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (۶)$$

۴-۱-۵) صحت (Precision)

عبارتست از نسبت نمونه‌های مثبت واقعی (TP) به تمام نمونه‌های مثبت واقعی و مثبت کاذب (کریستینک و همکاران (۲۰۲۰):

$$\text{Precision} = \frac{TP}{TP+FP} \quad (۷)$$

۴-۱-۶) F1-score

عبارتست از ترکیبی از معیارهای صحت و حساسیت که معمولاً به‌عنوان معیار مناسبی برای ارزیابی مجموعه داده‌های نامتعادل استفاده می‌شود و در بهترین حالت یک و در بدترین حالت صفر است (کریستینک و همکاران (۲۰۲۰):

$$\text{F1-Score} = \frac{2 \times \text{precision} \times \text{Recall}}{\text{precision} + \text{Recall}} \quad (۸)$$

۴-۲) ارزیابی عملکرد روش پیشنهادی

در جدول شماره ۲، نتایج مدل پیشنهادی بدون استفاده از تکنیک افزایش داده و با استفاده از تکنیک افزایش داده کلاسیک مقایسه شده است. نتایج این جدول روی مجموعه داده‌های تست نشان می‌دهد که استفاده از تکنیک کلاسیک افزایش داده برای تولید تصاویر جدید برای مدل پیشنهادی منجر به افزایش مقادیر معیارهای ارزیابی شده است. جدول شماره ۳، نتایج روش‌های پیشنهادی را در مقایسه با کارهای قبلی از نظر دقت نشان می‌دهد. نتایج نشان می‌دهد که روش‌های ارائه شده در این مقاله از دقت بسیار بهتری نسبت به روش‌های ارائه شده برخوردار هستند.

جدول ۲. ارزیابی عملکرد روش پیشنهادی از نظر معیارهای ارزیابی

روش	دقت	ویژگی	حساسیت	f1-score	TP	FP	TN	FN
مدل پیشنهادی بدون افزایش داده	۹۷.۰۱٪	۹۶.۶۱٪	۹۶.۶۱٪	۹۶.۶۱٪	۲۱	۱	۴۴	۱
مدل پیشنهادی با افزایش داده	۹۸.۵٪	۹۷.۸۲٪	۹۸.۸۸٪	۹۸.۳۲٪	۲۱	۱	۴۴	۰

جدول ۳. مقایسه روش های پیشنهادی با سایر روش های ارائه شده

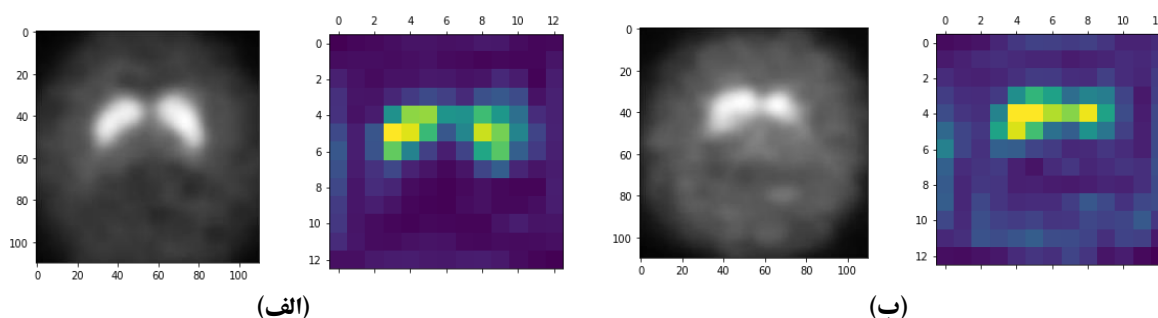
روش	مدل	مجموعه داده	دقت
مدل پیشنهادی بدون افزایش داده	ConvNet	SPECT	۹۷.۰۱٪
مدل پیشنهادی با افزایش داده	CovNet_Rotation	SPECT	۹۸.۵۰٪
روش [۱۱]	CNN	SPECT	۸۸.۹٪
روش [۱]	InceptionV3	SPECT	۷۴.۶۵٪
روش [۱۰]	VGG16	SPECT	۹۵.۲٪

از آنجاکه در کارهای پیشین هیچ کدام اشاره ای به سرعت یا زمان آموزش شبکه یا حافظه نداشته اند، در این تحقیق نیز امکان مقایسه سرعت و حافظه وجود نداشت. لازم به ذکر است که بیشتر کارهای پیشین از معیار دقت استفاده کردند که در جدول ۳ گزارش شده است. علاوه بر این، معیارهای حساسیت، ویژگی و امتیاز f1 نیز برای روش پیشنهادی این مقاله در جدول ۲ ارائه شده است. از بین تمامی معیارهای ارزیابی، امتیاز f1 جامع تر است زیرا ترکیبی از حساسیت و صحت است و در کنار معیار دقت می تواند تا حد زیادی نشان دهنده کارایی مدل باشد. مطابق جدول ۲ نتایج، مدل پیشنهادی که با داده افزایشی همراه باشد، با امتیاز f1 برابر ۹۸ درصد بهترین نتیجه را داده است.

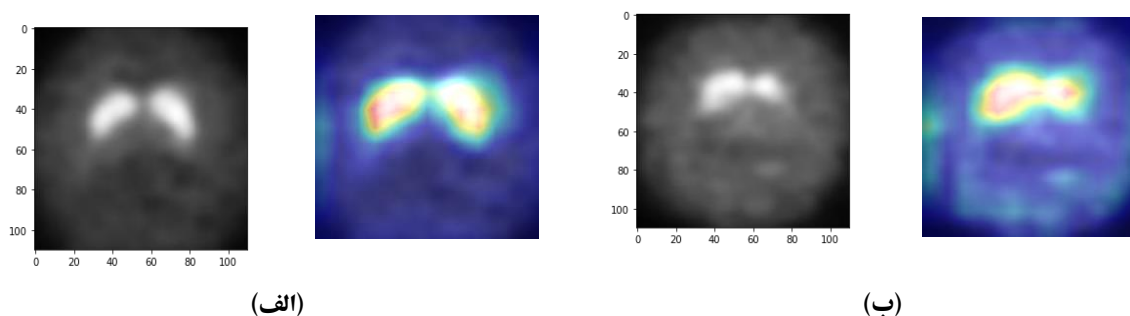
۳-۴) نتایج تفسیر مدل با استفاده از Grad-CAM

۳-۴-۱) نتایج تفسیر مدل بدون استفاده از تکنیک افزایش داده

در این بخش، نتایج حاصل از نقشه حرارتی و روش Grad-CAM برای تفسیر پیش بینی ها ارائه شده است. نمودار ۲ نقشه حرارتی افراد سالم و افراد مبتلا به بیماری پارکینسون را برای مدل پیشنهادی ConvNet نشان می دهد. در نقشه های حرارتی سردترین رنگ (آبی) قسمت هایی از تصویر را نشان می دهد که هیچ تاثیری بر تولید پیش بینی های CNN ندارد. در مقابل قسمت هایی از تصویر که بیشترین تاثیر را در تولید پیش بینی های CNN دارند، در نقشه حرارتی، زرد یا قرمز رنگ هستند. نمودار ۳ نتایج روش Grad-CAM را برای تفسیر شبکه پیشنهادی ConvNet نشان می دهد. نتایج ارائه شده نشان می دهد که شبکه کانولوشن پیشنهادی طبقه بندی را به درستی انجام داده است، زیرا افراد سالم را براساس جسم مخطط از افراد مبتلا به بیماری پارکینسون جدا کرده است.



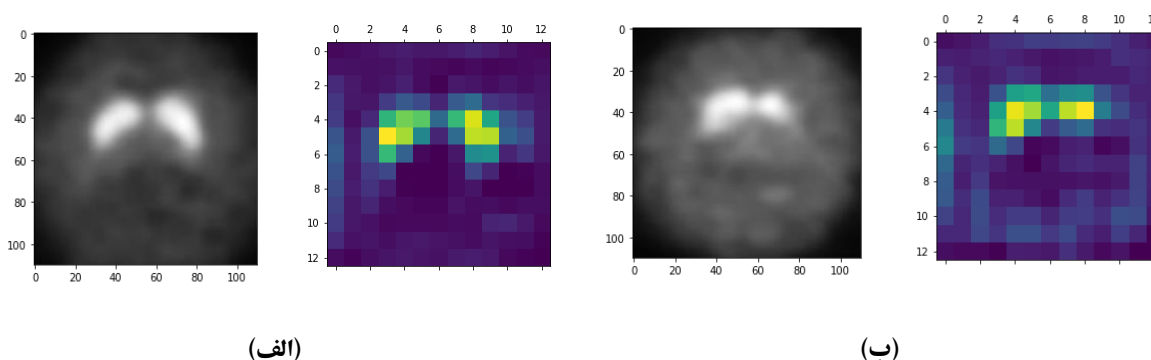
نمودار ۲. نقشه های حرارتی (الف) ConvNet برای فرد سالم، (ب) ConvNet برای فرد PD



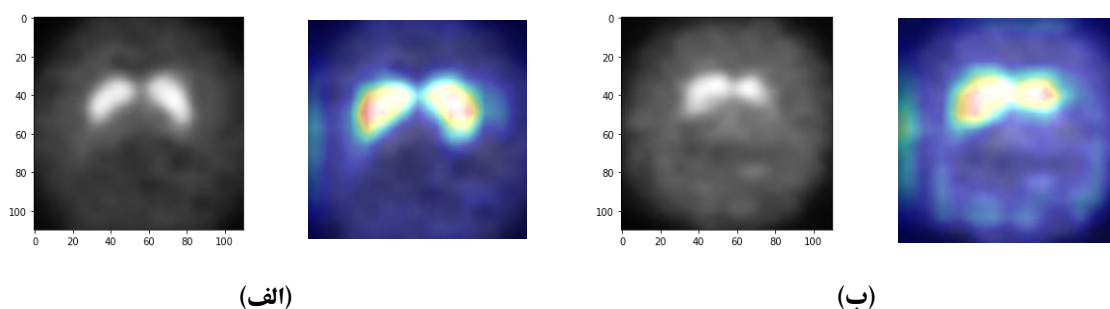
نمودار ۳. نتایج Grad-CAM به ازای (الف) ConvNet برای فرد سالم، (ب) ConvNet برای فرد PD

۴-۳-۲) نتایج تفسیر مدل با افزایش داده

نمودار ۴ نقشه حرارتی افراد سالم و افراد مبتلا به بیماری پارکینسون را برای مدل ConvNet با استفاده از تکنیک افزایش داده کلاسیک و نمودار ۵ نتایج روش Grad-CAM را برای تفسیر شبکه ConvNet با تکنیک افزایش داده کلاسیک نشان می دهد. نتایج نشان می دهد که مدل پیشنهادی توانسته است به خوبی افراد سالم را از افراد مبتلا به بیماری پارکینسون براساس بخش جسم مخطط جدا کند.



نمودار ۴. نقشه های حرارتی با تکنیک افزایش داده (الف) ConvNet برای فرد سالم، (ب) ConvNet برای فرد PD



نمودار ۵. Grad-CAM (الف) ConvNet برای فرد سالم، (ب) ConvNet برای فرد PD

۵) نتیجه گیری و پیشنهادها

در این مطالعه، در مرحله اول نرمال سازی تصاویر برای حصول نتایج بهتر انجام شده سپس به دلیل محدودیت تعداد داده ها، از تکنیک های معکوس کردن تصویر در جهت محور افقی و چرخش تصویر برای تولید تصاویر جدید برای عملکرد بهتر مدل استفاده شده است. سپس از یک شبکه عصبی کانولوشن به نام ConvNet برای طبقه بندی افراد سالم و افراد مبتلا به بیماری پارکینسون بر اساس تصاویر SPECT جمع آوری شده از پایگاه داده PPMI استفاده شده است. در نهایت از روش Grad-CAM برای تفسیر پیش بینی های انجام شده توسط مدل پیشنهادی استفاده شده است. نتایج شبیه سازی مدل پیشنهادی بدون استفاده از تکنیک افزایش داده در مقایسه با سایر روش های طبقه بندی، دقت ۹۷.۰۱ درصد، حساسیت ۹۶.۶۱ درصد و امتیاز $f1$ ۹۶.۶۱ درصد را ارائه کرده است. علاوه بر این، نتایج شبیه سازی مدل پیشنهادی همراه با تکنیک افزایش داده نیز دقت ۹۸.۵۰ درصد، حساسیت ۹۸.۸۸ درصد و امتیاز $f1$ ۹۸.۳۲ درصد را نشان می دهد که نسبت به کارهای پیشین بهبود یافته است.

منابع

- Appineni, A. and Gupta, A. (2023). Preemptive Diagnosis of Parkinson's Disease through DaT Scans using InceptionV3-based Convolutional Neural Networks. <https://doi.org/10.20944/preprints202312.2268.v1>
- Faull, R. and Laverly, R. (1969). Changes in dopamine levels in the corpus striatum following lesions in the substantia nigra. *Experimental neurology*, 23(3): p.332-340. [https://doi.org/10.1016/0014-4886\(69\)90081-8](https://doi.org/10.1016/0014-4886(69)90081-8)
- Poewe, W., et al., (2017). Parkinson disease. *Nature reviews Disease primers*, 3(1): p. 1-21. <https://doi.org/10.1038/nrdp.2017.13>
- Wang, W., et al., (2020). *Early detection of Parkinson's disease using deep learning and machine learning*. IEEE Access, 8: p. 147635-147646. <https://doi.org/10.1038/nrdp.2017.13>
- Wile, D.J., et al., (2017). Serotonin transporter binding and motor onset of Parkinson's disease in asymptomatic LRRK2 mutation carriers: a cross-sectional study. *The Lancet. Neurology*, 16(5): p. 351. [https://doi.org/10.1016/S1474-4422\(17\)30056-X](https://doi.org/10.1016/S1474-4422(17)30056-X)
- Marek, K., et al., (1996). [sup 123 I] beta-CIT/SPECT imaging demonstrates bilateral loss of dopamine transporters in hemi-Parkinson's disease. *Neurology*, 46(1): p. 231-237. <https://doi.org/10.1212/WNL.46.1.231>
- Mohammed, F., He, X. and Lin, Y. (2021). Retracted: An easy-to-use deep-learning model for highly accurate diagnosis of Parkinson's disease using SPECT images. *Computerized Medical Imaging and Graphics*. <https://doi.org/10.1016/j.compmedimag.2020.101810>
- Govindu, A. and Palwe, S. (2023). Early detection of Parkinson's disease using machine learning. *Procedia Computer Science*, 218: p. 249-261. <https://doi.org/10.1016/j.procs.2023.01.007>
- Tuncer, T., Dogan, S. and Acharya, U.R. (2020). Automated detection of Parkinson's disease using minimum average maximum tree and singular value decomposition method with vowels. *Biocybernetics and Biomedical Engineering*, 40(1): p. 211-220. <https://doi.org/10.1016/j.bbe.2019.05.006>

- Magesh, P.R., Myloth, R.D. and Tom, R.J. (2020). An explainable machine learning model for early detection of Parkinson's disease using LIME on DaTSCAN imagery. *Computers in Biology and Medicine*, 126: p. 104041. <https://doi.org/10.1016/j.compbiomed.2020.104041>
- Hathaliya, J., et al., (2022). Convolutional neural network-based Parkinson disease classification using SPECT imaging data. *Mathematics*, 10(15): p. 2566. <https://doi.org/10.3390/math1015256>
- Marek, K., et al., (2011). The Parkinson progression marker initiative (PPMI). *Progress in neurobiology*, 95(4): p. 629-635. <https://10.1016/j.pneurobio.2011.09.005>
- Martinez-Murcia, F.J., et al., (2018). Convolutional neural networks for neuroimaging in Parkinson's disease: Is preprocessing needed?, *International journal of neural systems*, 28(10): p. 1850035. <https://doi.org/10.1142/S0129065718500351>
- Mash, R., Borghetti, B. and Pecarina, J. (2016). Improved aircraft recognition for aerial refueling through data augmentation in convolutional neural networks. in *Advances in Visual Computing: 12th International Symposium, ISVC 2016, Las Vegas, NV, USA, December 12-14, Proceedings, Part I 12*. 2016. Springer. https://doi.org/10.1007/978-3-319-50835-1_11
- Dehghan, R., Naderan, M. and Alavi, S.E. (2022). Detection of Parkinson's disease using Convolutional Neural Networks and Data Augmentation with SPECT images. in *2022 12th International Conference on Computer and Knowledge Engineering (ICCKE)*. IEEE. 17-18 Nov., Mashhad, Iran. <https://doi.org/10.1109/ICCKE57176.2022.9960085>
- Zhou, B., et al. (2016). Learning deep features for discriminative localization. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. <https://doi.org/10.1109/CVPR.2016.319>
- Selvaraju, R.R., et al. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. in *Proceedings of the IEEE international conference on computer vision*. 22-29 October, Venice, Italy. <https://doi.org/10.1109/ICCV.2017.74>
- Krstinić, D., et al., (2020). Multi-label classifier performance evaluation with confusion matrix. *Computer Science & Information Technology*, 1: p. 1-14. <https://10.5121/csit.2020.100801>