



بررسی روش‌های یادگیری عمیق جهت تشخیص صرع در تصاویر MRI

نگین شفیعی زاده^۱، هاجر دانش^{۲*}، فرهاد خسروی^۳، شکوفه یراقی^۴

^۱ گروه برق و مهندسی پزشکی، دانشکده فنی مهندسی، دانشگاه شهید اشرفی اصفهانی، اصفهان، ایران negin@ashrafi.ac.ir

^۲ گروه برق و مهندسی پزشکی، دانشکده فنی مهندسی، دانشگاه شهید اشرفی اصفهانی، اصفهان، ایران، h.danesh@ashrafi.ac.ir

^۳ گروه برق و مهندسی پزشکی، دانشکده فنی مهندسی، دانشگاه شهید اشرفی اصفهانی، اصفهان، ایران، fkha@ashrafi.ac.ir

^۴ گروه مهندسی کامپیوتر، دانشکده فنی مهندسی، دانشگاه شهید اشرفی اصفهانی، اصفهان، ایران، sh.yaraghi@ashrafi.ac.ir

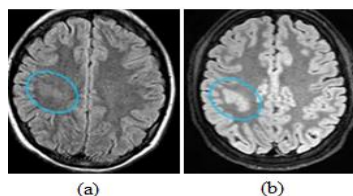
چکیده

صرع از شایع‌ترین اختلالات نورولوژیک است که تشخیص زودهنگام آن برای جلوگیری از آسیب‌های عصبی ضروری است. تصویربرداری MRI به دلیل وضوح بالا و غیرتهاجمی بودن، ابزار استاندارد در شناسایی ضایعات مغزی مرتبط با صرع محسوب می‌شود، اما تفسیر دستی آن زمان‌بر و مستعد خطاست. در سال‌های اخیر، شبکه‌های عصبی کانولوشنی (CNN) امکان تحلیل خودکار تصاویر MRI و شناسایی الگوهای پیچیده را فراهم کرده‌اند. این مطالعه مروری به بررسی عملکرد مدل‌های Fc-Net، 3D-CNN، DAG-CNN، U-Net، DNN و m-CNN در تشخیص صرع پرداخته است. نتایج نشان می‌دهند که مدل‌های CNN بالاترین دقت را داشته‌اند در برخی موارد تا بیش از 96٪ دقت تشخیصی حاصل شده است. با وجود این پیشرفت‌ها، چالش‌هایی مانند ناهمگونی داده‌ها، عدم شفافیت مدل‌ها و نیاز به سخت‌افزار قدرتمند همچنان مطرح است و نیازمند تحقیقات بیشتر می‌باشد.

واژه‌های کلیدی: صرع، یادگیری عمیق، شبکه عصبی کانولوشنی، MRI، تشخیص خودکار.

1. مقدمه

صرع یکی از شایع‌ترین اختلالات عصبی مزمن سیستم عصبی مرکزی است که ناشی از تخلیه‌های الکتریکی پاتولوژیک، ناگهانی و همگانی یا موضعی در نورون‌های مغزی می‌باشد. این بیماری با حملات تشنجی غیرقابل پیش‌بینی و عودکننده همراه است که ممکن است شامل اختلال در هوشیاری، حرکات غیرارادی غیرعادی یا تغییرات رفتاری باشد. بر اساس برآوردهای جهانی، بیش از 70 میلیون نفر در سراسر جهان به این بیماری مبتلا هستند [1]. طبق گزارشی از اتحادیه بین‌المللی مبارزه با صرع (ILAE)¹، صرع حداقل با دو تشنج، آن هم در فاصله بیش از 24 ساعت تشخیص داده می‌شود [2]. صرع بر اساس محل شروع فعالیت تشنجی در مغز به چهار دسته تقسیم می‌شود که شامل صرع کانونی، صرع منتشر، صرع کانونی-منتشر و صرع با منشأ نامشخص است [3، 4]. در صرع کانونی، فعالیت تشنجی از یک ناحیه محدود در یکی از نیمکره‌های مغز آغاز می‌شود؛ در حالی که در صرع منتشر، فعالیت پاتولوژیک از ابتدا هر دو نیمکره را درگیر می‌کند. نوع کانونی-منتشر شامل حملاتی است که ابتدا به صورت موضعی شروع شده و سپس به سایر نواحی مغز گسترش می‌یابند. در مواردی که منشأ تشنج قابل شناسایی نباشد، بیماری به عنوان صرع با منشأ نامشخص طبقه‌بندی می‌شود [5]. در میان انواع صرع کانونی، صرع لوب تمپورال (TLE)² شایع‌ترین نوع است و معمولاً از ساختارهای داخلی لوب تمپورال، به ویژه هیپوکامپ، منشأ می‌گیرد [6]. برای تشخیص دقیق نوع و منشأ تشنج، از روش‌های تصویربرداری عصبی متعددی از جمله³ MRI،⁴ PET،⁵ SPECT و⁶ CT استفاده می‌شود. هر یک از این روش‌ها مزایا و محدودیت‌های خاص خود را دارند. در این میان MRI به عنوان روش استاندارد طلایی در ارزیابی ساختاری مغز شناخته می‌شود؛ زیرا بدون استفاده از پرتو یونیزان، امکان نمایش با وضوح بالا از ساختارهای آناتومیک و شناسایی ضایعات ریز یا ناهنجاری‌های زیربالیانی را فراهم می‌کند [7]. با این حال، سایر روش‌ها دارای محدودیت‌های قابل توجهی هستند. PET اگرچه قادر به نمایش تغییرات متابولیک مناطق مغزی است، اما در مواجهه با تغییرات غیراختصاصی مستعد خطا بوده و ممکن است منجر به سوءتشخیص یا تأخیر در برنامه‌ریزی درمانی شود [8]. SPECT فرآیندی زمان‌بر است و نیازمند تزریق عامل رادیواکتیو در حین حمله تشنجی و آماده‌سازی دقیق بیمار می‌باشد که کاربرد بالینی آن را محدود می‌کند [9]. EEG با وجود داشتن وضوح زمانی بسیار بالا و ضرورت در ثبت فعالیت‌های الکتریکی لحظه‌ای مغز، در ارائه اطلاعات مکانی دقیق از کانون تشنجی دچار محدودیت است و اغلب نمی‌تواند محل دقیق ضایعه را مشخص کند [10]. CT نیز علیرغم کاربرد در شرایط اورژانسی، به دلیل استفاده از پرتو یونیزان و حساسیت پایین‌تر نسبت به MRI در شناسایی ضایعات اپی‌لپتوژنیک، نقش محدودی در ارزیابی ساختاری بیماران مبتلا به صرع ایفا می‌کند [7]. با وجود پیشرفت‌های چشمگیر در فناوری‌های تصویربرداری، تفسیر دستی تصاویر MRI همچنان با چالش‌هایی مواجه است. حجم بالای داده‌های سه‌بعدی، پیچیدگی آناتومی مغز و وجود تغییرات ساختاری ظریف، این فرآیند را به کاری زمان‌بر، وابسته به تجربه بالینی متخصص و مستعد خطای انسانی تبدیل کرده است [11]. همان‌طور که در شکل (1) نشان داده شده است، تفاوت‌های ساختاری بین تصاویر MRI یک فرد سالم (A) و یک فرد مبتلا به صرع کانونی (B) اغلب بسیار ظریف بوده و نیازمند تحلیل دقیق و عمیق‌تری است.



شکل 1- (A) تصویر MRI فرد سالم و (B) تصویر MRI فرد مبتلا به صرع کانونی [12]

¹International League Against Epilepsy

² Temporal Lobe Epilepsy

³ Magnetic Resonance Imaging

⁴ Positron Emission Tomography

⁵ Single Photon Emission Computed Tomography

⁶ Computed Tomography

⁷ Electroencephalography

در این راستا، روش‌های محاسباتی پیشرفته، به‌ویژه یادگیری عمیق (DL)¹ به‌عنوان زیرشاخه‌ای از یادگیری ماشین، جایگاه فزاینده‌ای در تحلیل تصاویر پزشکی یافته‌اند [13]. در این میان، شبکه‌های عصبی کانولوشنی (CNN)² به‌دلیل توانایی در استخراج خودکار ویژگی‌های سلسله‌مراتبی و شناسایی الگوهای پیچیده در داده‌های تصویری، ابزاری قدرتمند در تحلیل تصاویر MRI و تشخیص ناهنجاری‌های مرتبط با صرع محسوب می‌شوند [14].

2. مواد و روش‌ها

این پژوهش به عنوان یک مطالعه مروری، با هدف ارزیابی و مقایسه عملکرد معماری‌های مختلف یادگیری عمیق در تشخیص صرع از تصاویر MRI انجام شده است. جهت شناسایی مطالعات مرتبط، جستجوی جامعی در پایگاه‌های علمی معتبر از جمله Google Scholar، IEEE Xplore، ScienceDirect و MDPI انجام گرفت. کلیدواژه‌های اصلی مورد استفاده در این جستجو شامل «یادگیری عمیق»، «تشخیص صرع»، «شبکه عصبی کانولوشنی» و «تحلیل تصاویر MRI» بودند که با به‌کارگیری عملگرهای بولی به‌صورت ترکیبی مورد استفاده قرار گرفتند. بازه زمانی مورد بررسی مقالات منتشر شده، بین سال‌های 2015 تا 2024 در نظر گرفته شد تا اطمینان حاصل شود که پیشرفت‌های اخیر در این حوزه به‌طور کامل پوشش داده شده‌اند. پس از غربالگری دقیق مقالات بر اساس عنوان و چکیده، و سپس مطالعه متن کامل، 47 مطالعه مطابق معیارهای ورود و خروج انتخاب شدند. معیارهای ورود شامل مطالعاتی بود که به‌طور مستقیم از روش‌های یادگیری عمیق به‌ویژه شبکه‌های عصبی کانولوشنی برای تحلیل تصاویر MRI در تشخیص صرع استفاده کرده بودند و دارای گزارش معیارهای ارزیابی کمی مانند دقت یا ضریب دایس بودند که بر روی داده‌های انسانی انجام شده بودند. از سوی دیگر، مقالاتی که تنها از روش‌های یادگیری ماشین سنتی بهره می‌بردند، داده‌های غیرتصویری داشتند و یا از نظر روش‌شناختی فاقد شفافیت و کیفیت کافی بودند، از تحلیل نهایی حذف شدند. اهداف اصلی این مطالعه شامل ارزیابی دقت تشخیص صرع، شناسایی محل کانون تشنج و تمایز آن از سایر اختلالات عصبی از طریق تحلیل تصاویر MRI است. به این منظور، معماری‌های مختلف که در مطالعات مورد بررسی به کار گرفته شده‌اند از جمله شبکه کانولوشنی چندلایه (m-CNN)، شبکه کانولوشنی سه‌بعدی (3D-CNN)، شبکه (U-Net)، شبکه گراف جهت‌دار (DAG-CNN)، شبکه کانولوشنی با کرنل مختلط (Fc-Net) و شبکه عصبی عمیق (DNN) مورد تحلیل و مقایسه قرار گرفتند. انتخاب این مدل‌ها بر اساس پتانسیل بالای آن‌ها در بهبود دقت طبقه‌بندی، استخراج خودکار ویژگی‌های معنادار و کاهش خطای تفسیر انسانی صورت گرفت. فرآیند تحلیل در این مطالعه در قالب یک چرخه ساختاریافته چهارمرحله‌ای طراحی شد. در گام نخست، داده‌های تصویری مرتبط با صرع از منابع عمومی و تخصصی بررسی گردید. سپس، روش‌های پیش‌پردازش شامل حذف نویز، نرمال‌سازی شدت پیکسل‌ها، جداسازی بافت‌های مغزی و حذف ساختارهای غیرمرتبط معرفی گردید و در انتها انواع روش‌های طبقه‌بندی براساس شبکه‌های عصبی مورد بررسی قرار گرفت. این چارچوب تحلیلی، امکان ارزیابی و مقایسه دقیق عملکرد معماری‌های مختلف را فراهم کرده و مسیر انتخاب بهینه‌ترین روش‌های تشخیصی در شرایط بالینی را هموار می‌سازد.

2-1-2 مراحل تشخیص تومور مغزی با استفاده از تکنیک‌های یادگیری عمیق

1-1-2 پایگاه داده: پایگاه‌های داده معتبر مورد استفاده در مطالعات مرتبط با تشخیص صرع، به طور عمده توسط دانشگاه‌ها، مراکز تحقیقاتی-تخصصی و کلینیک‌های عصبی-جراحی جمع‌آوری و مدیریت می‌شوند. این منابع شامل داده‌های بالینی، تصاویر MRI و گاهی اطلاعات ژنتیکی بیماران هستند و نقش محوری در توسعه روش‌های نوین تشخیصی و درمانی

¹ Deep Learning

² Convolutional Neural Network

ایفا می‌کنند. برخی از مهم‌ترین این منابع شامل داده‌های جمع‌آوری شده توسط دانشگاه‌های کالیفرنیا (سانفرانسیسکو و سن دیگو)، اموری، پنسیلوانیا، کارولینای جنوبی و مراکز تحقیقاتی وابسته به پروژه‌های انسانی می‌باشند. علاوه بر این، از پایگاه‌های داده عمومی مانند پایگاه طرح تصویربرداری عصبی بیماری آلزایمر (ADNI)^۱ (برای مطالعات مقایسه‌ای با بیماری آلزایمر) و پایگاه اطلس ژنوم سرطان (TCGA)^۲ (در مطالعات مرتبط با سرع ناشی از تومورهای مغزی) نیز در برخی تحقیقات استفاده شده است [12]. لیست کامل پایگاه‌های داده، تعداد نمونه‌ها و نویسندگان مطالعات مورد بررسی در جدول (1) ارائه شده است.

جدول 1- پایگاه داده‌های مورد استفاده در مقالات بررسی شده

تعداد تصاویر	پایگاه داده	نام نویسندگان
589 نفر (TLE) و 589 نفر سالم	دانشگاه‌های: کالیفرنیا سانفرانسیسکو، راش، کالیفرنیا سن دیگو، پیتسبورگ، پنسیلوانیا، نورث ول، بن، کارولینای جنوبی، اموری، نیویورک، لیورپول و پروژه‌های انسانی	Erike kaestner و همکاران [15]
95 بیمار مبتلا به MTLE ³ و 202 نفر سالم	دانشگاه‌های: کارولینای جنوبی، اموری و بن	Ezequiel Gleichgerrch و همکاران [16]
157 نفر (TLE) مقاوم به دارو، 73 نفر مبتلا به الزایمر و 251 نفر سالم در سنین بالای 18 سال	دانشگاه‌های: کارولینای جنوبی، اموری و بن، ADNI	Allen J Chang و همکاران [17]
359 نفر (TLE)	7 مرکز جراحی	Erike kaestner و همکاران [18]
105 نفر مبتلا به سرع تومور مغزی و 155 نفر سالم	TCGA	Jayathi vajiram و همکاران [19]
32 نفر مبتلا به سرع و 2132 نفر سالم	مرکز سرع بزرگسالان دانشگاه واشنگتن	Patrick H Lockett و همکاران [20]
45 MRI (TLE) و 40 MRI سالم	دانشگاه پنسیلوانیا، مرکز پزشکی دانشگاه وندربیل	T. Campbell Arnold و همکاران [21]
85 نفر نوع 1 ILAE ⁴ و MTLE ³ و 56 نفر سالم	بیمارستان NHO Nishiniigata Chuo	Yosuke Ito و همکاران [22]

2-1-2 پیش پردازش :

تکنیک‌های پیش پردازش، از جمله حذف نویز و بهبود کیفیت تصویر، نقش حیاتی در افزایش تفسیرپذیری تصاویر MRI دارند و برای دستیابی به تشخیص دقیق و قابل اعتماد ضروری هستند [23]. فرآیند حذف نویز معمولاً شامل مراحل تنظیم کنتراست، هموارسازی و نرمال‌سازی است. اگر این مراحل به درستی انجام نشوند، می‌توانند منجر به خطاهای جدی در طبقه‌بندی نهایی شوند [24]. برای پیش پردازش تصاویر MRI و افزایش دقت مدل‌های یادگیری عمیق، از ابزارهای استاندارد و تخصصی متعددی استفاده می‌شود. از جمله این ابزارها می‌توان به SPM، CAT12 و FSL/FAST اشاره کرد.

¹ Alzheimer's Disease Neuroimaging Initiative

² The Cancer Genome Atlas

³ Mesial Temporal Lobe Epilepsy

⁴ Hippocampal Sclerosis

SPM یک جعبه‌ابزار قدرتمند برای تحلیل آماری داده‌های تصویربرداری عصبی است که به محققان امکان می‌دهد فعالیت مغز را در پاسخ به محرک‌های مختلف ارزیابی کنند [25]. CAT12^۱ نیز شامل مجموعه‌ای از ابزارهای پیشرفته برای کنترل کیفیت، تقسیم‌بندی بافت‌ها بر اساس وکسل و سطح و انجام تحلیل‌های آناتومیکی با دقت بالا می‌باشد همچنین، FAST^۲ که بخشی از مجموعه نرم‌افزاری FSL است، به‌طور اختصاصی برای تقسیم‌بندی بافت‌های مغزی در تصاویر ساختاری MRI طراحی شده است [26]. در مراحل پیش‌پردازش، به‌ویژه در هموارسازی فضایی، از پارامتر FWHM^۳ به‌عنوان معیاری برای تعیین اندازه کرنل گاوسی استفاده می‌شود [27]. FAST^۴ بخشی از مجموعه نرم‌افزار FSL^۵ می‌باشد. این ابزار برای تقسیم‌بندی بافت‌های مغزی در تصاویر ساختاری MRI استفاده می‌شود [28]. همچنین، الگوریتم Canny Edge Detection به‌عنوان یکی از دقیق‌ترین روش‌های تشخیص لبه در پردازش تصویر، در حوزه MRI برای شناسایی مرزهای بین بافت‌های مغزی به کار می‌رود [29]. علاوه بر این، نرم‌افزار ITK-SNAP^۵ به‌عنوان یک ابزار تعاملی برای تقسیم‌بندی نیمه‌خودکار تصاویر پزشکی سه‌بعدی، به‌طور گسترده در MRI برای شناسایی و جداسازی ساختارهای آناتومیکی یا پاتولوژیک مورد استفاده قرار می‌گیرد [30]. در مقاله Kaestner و همکاران [15] تصاویر MRI با وزن T1 با استفاده از ابزارهای Nii_preprocess، SPM12 و CAT12 پیش‌پردازش شدند. مراحل شامل نرمال‌سازی به فضای استاندارد MNI152، تقسیم‌بندی بافت‌ها، هموارسازی سه‌بعدی با FWHM، استخراج اسلایس‌ها و برچسب‌گذاری بودند. در مطالعه ای از Gleichgerricht و همکاران [16] برای نرمال‌سازی تصاویر به فضای استاندارد MNI و تقسیم‌بندی بافت‌های مغزی، از ابزارهای SPM و FAST استفاده شد. نقشه‌های ماده خاکستری سپس با استفاده از تابع smooth در SPM و با FWHM هموارسازی شدند تا تفاوت‌های آناتومیکی بین افراد کاهش یابد. همچنین در مقاله Chang و همکاران [17] مراحل پیش‌پردازش شامل نرمال‌سازی به فضای استاندارد MNI152، تقسیم‌بندی بافت‌های مغزی با استفاده از SPM و CAT12، هموارسازی سه‌بعدی با FWHM، آستانه‌گذاری، استخراج اسلایس محوری مرکزی و برچسب‌گذاری بر اساس تشخیص بیماری بود. Kaestner و همکارانش [18] در مطالعه ای دیگر، تصاویر MRI با استفاده از ابزار CAT12 در محیط SPM12 پیش‌پردازش شدند. مراحل شامل تقسیم‌بندی بافت‌ها، نرمال‌سازی به فضای MNI و کنترل کیفیت خودکار بود. همه تصاویر با امتیاز کیفیت ≤ 75 پذیرفته شدند. در مقاله Vajiram و همکارش [19] حذف بافت‌های غیرمغزی با استفاده از تکنیک‌های مورفولوژیکی انجام شد. از مدل رگرسیون لجستیک جهت انتخاب ویژگی‌های تصویری استفاده شد و همچنین الگوریتم Canny برای تشخیص لبه‌ها همراه با محاسبه گرادیان تصویر به کار گرفته شدند. در مقاله Luckett و همکاران [20] مراحل پیش‌پردازش شامل تصحیح حرکت، حذف واریانس‌های مزاحم (شامل روند خطی، نویز حرکتی)، فیلترگذاری پایین‌گذر زمانی، هموارسازی و نرمال‌سازی تصاویر به فضای استاندارد MNI بود. Arnold و همکاران [21] در این مطالعه، ناحیه رزکشن در تصاویر پس از جراحی توسط متخصص نورورادیولوژی و با نرم‌افزار ITK-SNAP به‌صورت دستی تقسیم‌بندی شد. برای افزایش تعمیم‌پذیری مدل، از تکنیک‌های افزایش داده شامل چرخش، وارونه‌سازی و جابه‌جایی تا ۱۰٪ استفاده شد. داده‌ها نیز با اعتبارسنجی متقابل پنج‌تایی به مجموعه‌های آموزش، اعتبارسنجی و آزمون تقسیم شدند. در مقاله Ito و همکاران [22] از هر بیمار، ۱۶ اسلایس FLAIR در محور اکسیال موازی با هیپوکامپ تهیه شد. برای آموزش، سه اسلایس و برای آزمون، اسلایس مرکزی انتخاب گردید. پس از نرمال‌سازی مقادیر پیکسلی در بازه (۰-۱)، تکنیک‌های افزایش داده جهت کاهش overfitting به کار رفت.

^۱ Computational Anatomy Toolbox

^۲ FMRIB's Automated Segmentation Tool

^۳ Full Width at Half Maximum

^۴ Fully Automated Statistical Thresholding

^۵ FMRIB Software Library

2-1-3 استخراج ویژگی

فرآیند استخراج ویژگی از تصاویر MRI با استفاده از معماری‌های مختلف شبکه‌های عصبی عمیق انجام می‌شود و نقش محوری در تمایز دقیق بین بیماران مبتلا به صرع، افراد سالم و سایر اختلالات عصبی مانند بیماری آلزایمر ایفا می‌کند. این مرحله به صورت خودکار و بدون نیاز به مهندسی دستی ویژگی‌ها، قادر به شناسایی الگوهای پیچیده و ظریف در داده‌های تصویری است. در مقاله Kaestner و همکاران [15] برای استخراج ویژگی‌ها، از مدل FCNet الهام گرفته از پژوهش Peng و همکاران [31] استفاده شد. این مدل شامل چهار لایه کانولوشنی سه‌بعدی متوالی بود که هر کدام همراه با نرمال‌سازی دسته‌ای، لایه Pooling و تابع فعال‌سازی ReLU عمل می‌کردند. پس از استخراج ویژگی‌ها، از تکنیک‌های تفسیرپذیری مانند Grad-CAM برای شناسایی نواحی مغزی مؤثر در تصمیم‌گیری مدل، استفاده شد. در مقاله Gleichgerricht و همکاران [16] نقشه‌های ماده خاکستری مغز به عنوان ورودی به دو مدل CNN و DAG-CNN داده شدند. در مقاله Chang و همکاران [17] از یک شبکه عصبی کانولوشنی دوبعدی برای استخراج ویژگی‌ها استفاده شد. معماری این شبکه شامل سه ماژول کانولوشنی بود که هر کدام از یک لایه کانولوشن 3×3 ، نرمال‌سازی دسته‌ای، تابع فعال‌سازی ReLU و لایه Pooling تشکیل شده بودند. پس از این بخش، دو لایه کاملاً متصل و یک لایه Softmax برای انجام طبقه‌بندی نهایی به کار گرفته شدند. در مقاله دیگر از Kaestner و همکاران [18] از یک مدل CNN با معماری 2.5 بعدی استفاده شد. برای آموزش مدل از تابع خطا Binary Cross-Entropy، بهینه‌ساز Adam و تنظیم‌کننده L2 استفاده شد تا از بیش‌برازش جلوگیری شود. همچنین، برای شناسایی نواحی مغزی مؤثر در طبقه‌بندی، از تکنیک Grad-CAM بهره گرفته شد. در مقاله Vajiram و همکاران [19] استخراج ویژگی‌ها با ترکیب روش‌های رادیومیکس و یادگیری عمیق انجام شد سپس، با استفاده از مدل رگرسیون لجستیک، ویژگی‌های مهم برای شناسایی خوشه‌های اپی‌لپتوژنیک انتخاب شدند. برای بازسازی سطح سه‌بعدی ضایعات از داده‌های حجمی MRI، از الگوریتم Marching Cubes استفاده شد. در مقاله Luckett و همکاران [20] برای استخراج ویژگی‌های عملکردی مغز از یک شبکه عصبی کانولوشنی سه‌بعدی با معماری متراکم (DenseNet) استفاده شد. ورودی مدل شامل داده‌های سری زمانی BOLD (نمایانگر فعالیت مغز در طول زمان) بود. پس از آموزش مدل، از تکنیک Grad-CAM استفاده شد. در مقاله Arnold و همکاران [21] از یک الگوریتم ترکیبی مبتنی بر سه شبکه عصبی کانولوشنی U-Net استفاده شد. در این معماری، بخش رمزگذار (Encoder) با استفاده از شبکه پیش‌آمخته EfficientNet-B1 جایگزین شد تا استخراج ویژگی‌ها بهینه‌تر شود. برای ترکیب خروجی سه شبکه و تولید نتیجه نهایی، از روش رأی اکثریت (Majority Vote) استفاده شد. در مقاله Ito و همکاران [22] به منظور افزایش دقت و جلوگیری از بیش‌برازش، از تکنیک‌های نرمال‌سازی دسته‌ای در لایه‌های اولیه استفاده شد. همچنین، الگوریتم RMSProp به عنوان تابع بهینه‌سازی برای آموزش مدل به کار گرفته شد.

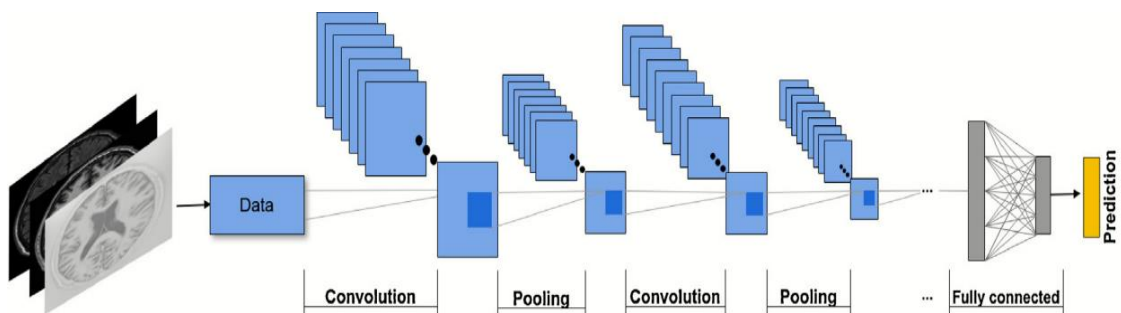
2-1-4 طبقه بندی

پس از انجام مراحل پیش‌پردازش و استخراج ویژگی‌های معنادار از تصاویر MRI، فرآیند طبقه‌بندی با بهره‌گیری از معماری‌های مختلف یادگیری عمیق آغاز می‌شود. این بخش به عنوان هسته اصلی سیستم‌های کمک تشخیص مبتنی بر هوش مصنوعی عمل می‌کند و نقش تعیین‌کننده‌ای در تصمیم‌گیری بالینی دارد. در مطالعات بررسی‌شده، از معماری‌های متنوعی از جمله CNN، DAG-CNN، 3D-CNN، DNN، m-CNN و U-Net برای طبقه‌بندی استفاده شده است. هر یک از این مدل‌ها با توجه به ساختار خاص خود، الگوهای تصویری مرتبط با بیماری را شناسایی می‌کنند. به عنوان مثال، مدل‌های مبتنی بر 3D-CNN با بهره‌گیری از معماری چندلایه و فیلترهای سه‌بعدی، قادر به تحلیل داده‌های حجمی MRI و شناسایی محل آغاز تشنج با دقت بالا هستند. از سوی دیگر، معماری DAG-CNN با استفاده از ساختار سلسله‌مراتبی و چندمقیاسی، توانایی

شناسایی تغییرات بافتی در ماده خاکستری مغز به‌ویژه در نواحی مرتبط با صرع لوب تمپورال را دارد. برای بهینه‌سازی فرآیند طبقه‌بندی، از روش‌های استاندارد مانند انتخاب ویژگی‌های رادیومیک و تکنیک‌های حذف نویز استفاده شده است. همچنین، به‌منظور افزایش تفسیرپذیری و شفافیت در تصمیم‌گیری مدل‌ها، از تکنیک‌هایی مانند Grad-CAM بهره گرفته شده است. معماری‌هایی نظیر U-Net در کاربردهای بالینی نیز مورد استفاده قرار می‌گیرند، اگرچه که تمرکز اصلی آن‌ها بر تقسیم‌بندی دقیق نواحی آسیب‌دیده در تصاویر پزشکی می‌باشد. این قابلیت به پزشکان امکان می‌دهد تا با دقت بالا مرزهای آناتومیکی ناهنجاری‌ها را شناسایی کرده و برنامه‌ریزی درمانی مؤثرتری انجام دهند. در مجموع، عملکرد هر یک از این معماری‌ها در این مطالعه مروری از دو منظر دقت تشخیص و پایداری نتایج مورد ارزیابی قرار گرفته است. یافته‌ها نشان می‌دهند که انتخاب معماری مناسب، به‌طور مستقیم بر کیفیت خروجی نهایی تأثیر می‌گذارد و نیازمند تطبیق با نوع داده، هدف تشخیصی و شرایط بالینی است.

معماری CNN

CNN از معماری‌های مؤثر در تحلیل تصاویر پزشکی، به‌ویژه MRI مغز، محسوب می‌شود [32]. ساختار سلسله‌مراتبی آن امکان یادگیری خودکار ویژگی‌ها از سطوح ساده تا پیچیده را بدون نیاز به مهندسی دستی فراهم می‌سازد. معماری استاندارد CNN شامل لایه‌های کانولوشنی (Convolutional Layers)، پولینگ (Pooling Layers) و کاملاً متصل (Fully Connected Layers) است. در لایه‌های کانولوشنی، فیلترهای کوچک ویژگی‌های اولیه مانند لبه‌ها و بافت‌ها را استخراج کرده و در لایه‌های عمیق‌تر به ساختارهای مغزی مانند هیپوکامپ و قشر مخ تبدیل می‌شوند. لایه‌های پولینگ، به‌ویژه نوع بیشینه‌گیری (Max-Pooling)، با کاهش ابعاد داده و افزایش مقاومت مدل در برابر تغییرات جزئی، از بیش‌برازش (Overfitting) جلوگیری می‌کنند. داده‌ها سپس به‌صورت بردار یک‌بعدی (Flattened Vector) به لایه‌های کاملاً متصل منتقل شده و فرآیند طبقه‌بندی نهایی انجام می‌شود. شکل (2) نمایی کلی از معماری مذکور را نشان می‌دهد [33].

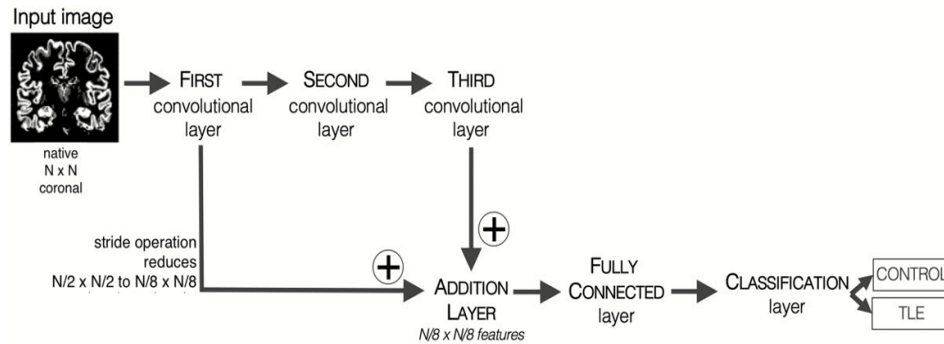


شکل 2- نمایی کلی از معماری CNN [33]

معماری DAG-CNN

DAG-CNN معماری پیشرفته‌ای در شبکه‌های عصبی کانولوشنی است که به‌جای ساختار خطی، از گراف جهت‌دار بدون دور (Directed Acyclic Graph) بهره می‌برد. این ساختار با ایجاد مسیرهای موازی بین لایه‌های غیرمتوالی، جریان گرادینان را بهبود داده و همگرایی مدل را در شبکه‌های عمیق تسهیل می‌کند. توانایی DAG-CNN در استخراج ویژگی‌های چندمقیاسی، آن را برای تحلیل تصاویر MRI مغز به‌ویژه ساختارهایی با تنوع فضایی مانند هیپوکامپ و قشر مخ مناسب ساخته است. هر بلوک شامل کانولوشن، Batch Normalization و ReLU بوده و موجب افزایش پایداری و سرعت آموزش می‌شود. در معماری شکل (۳)، سه لایه کانولوشنی متوالی طراحی شده‌اند. لایه اول با فیلترهای بزرگ، ویژگی‌های کلی نواحی قشری و

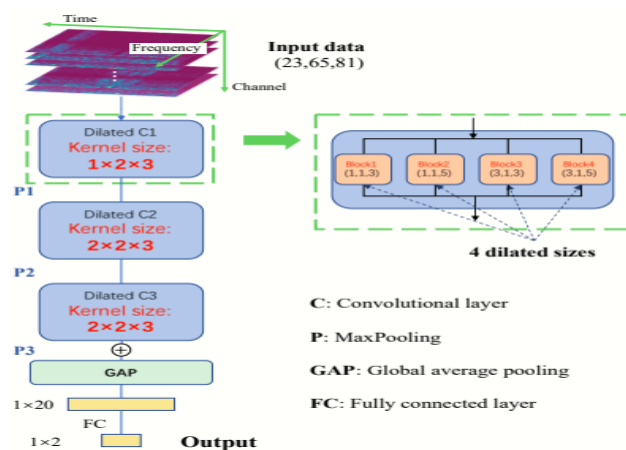
زیرقشری را شناسایی می‌کند؛ لایه دوم و سوم با فیلترهای کوچک‌تر، به استخراج دقیق‌تر از نواحی خاص می‌پردازند. خروجی لایه اول و سوم در لایه ادغام (Fusion Layer) ترکیب شده و اطلاعات چندمقیاسی برای تصمیم‌گیری نهایی فراهم می‌شود. داده‌های ادغام‌شده به لایه کاملاً متصل (Fully Connected) و طبقه‌بندی (Classification Layer) ارسال شده و برچسب «مبتلا به صرع» یا «سالم» تخصیص می‌یابد. این معماری با بهره‌گیری از ساختار گرافی و ویژگی‌های چندسطحی، دقت بالایی در شناسایی تغییرات ساختاری مرتبط با صرع نشان داده است [34].



شکل 3- نمایی کلی از معماری DAG-CNN [16]

معماری 3D-CNN

شبکه عصبی کانولوشنی سه‌بعدی (3D-CNN) از معماری‌های مؤثر در پردازش داده‌های حجمی مانند تصاویر MRI مغز است. برخلاف مدل‌های دوبعدی، این شبکه با فیلترهای سه‌بعدی، روابط فضایی بین وکسل‌ها را در تمامی ابعاد مدل‌سازی کرده و پیوستگی آناتومیکی ساختارها را حفظ می‌کند. این ویژگی در تشخیص صرع کانونی که نیازمند تحلیل یکپارچه داده‌های سه‌بعدی است، اهمیت بالایی دارد [35]. معماری شبکه 3D-CNN در شکل (۴) نشان داده شده است. در برخی از این معماری‌ها، خروجی لایه‌های مختلف پس از لایه نهایی کانولوشنی با یکدیگر ترکیب می‌شوند تا ویژگی‌های چندسطحی از تمامی عمق شبکه بهره‌مند شوند. سپس، از تکنیک (GAP)^۱ برای کاهش ابعاد بردار ویژگی‌ها و کاهش تعداد پارامترهای قابل آموزش استفاده می‌شود. در نهایت، داده‌های تولیدشده به یک لایه کاملاً متصل و سپس به یک لایه Softmax ارسال می‌شوند [36].

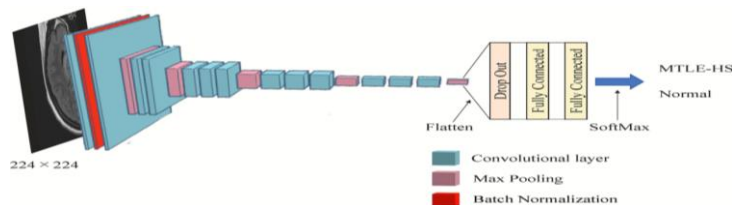


شکل 4- معماری 3D-CNN [36]

^۱ Global Average Pooling

مدل m-CNN

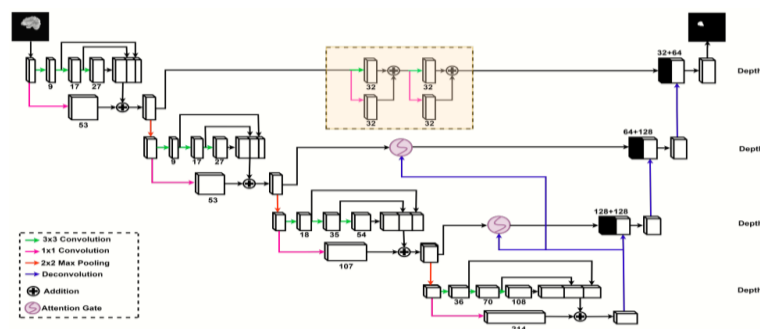
مدل m-CNN یک نسخه توسعه یافته و تخصصی شده از معماری VGG-16 است که به طور خاص برای بهبود عملکرد در تشخیص بیماری های عصبی به ویژه صرع لوب تمپورال متوسط (MTLE) و اسکروز هیپوکامپی (HS) طراحی شده است. یکی از ویژگی های کلیدی این مدل، استفاده از کانولوشن های مختلط (Complex-Valued Convolution) است. این رویکرد امکان پردازش همزمان داده ها را در دو حوزه مختلف فراهم می کند: حوزه فرکانسی (k-space) و حوزه مکانی [37]. برای افزایش تعمیم پذیری مدل و کاهش بیش برازش، از تکنیک های نرمال سازی دسته ای (Batch Normalization) در لایه های حساس استفاده شد. همچنین، با بهره گیری از ImageDataGenerator در Keras، داده از طریق تغییرات تصادفی مانند تغییر ابعاد، برش، بزرگ نمایی و چرخش افقی انجام گرفت. در فرآیند آموزش، وزن های اولیه مدل از شبکه پیش آموخته VGG-16 که بر روی مجموعه داده ImageNet آموزش دیده است انتقال یافته اند. سپس، کل شبکه به صورت پایان به پایان (End-to-End) با پارامترهای بهینه سازی دقیق مجدداً آموزش دیده اند. این پارامترها با استفاده از الگوریتم بهینه ساز RMSProp بهینه می شوند. این استراتژی آموزشی منجر به همگرایی پایدار و دقیق مدل شده و توانایی آن را در شناسایی تغییرات ساختاری ظریف مانند کاهش حجم هیپوکامپ در بیماران مبتلا به MTLE-HS به طور مؤثری افزایش داده است. معماری نهایی m-CNN در شکل (5) به صورت شماتیک نمایش داده شده است [22].



شکل 5- نمایش کلی از معماری m-CNN [22]

معماری U-Net

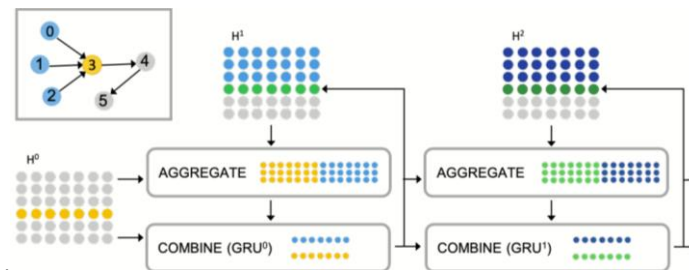
U-Net یکی از معماری های مؤثر در تقسیم بندی تصاویر پزشکی، به ویژه MRI مغز، است که با دقت بالا در تعیین مرزهای ساختارهای آناتومیکی شناخته می شود. این مدل شامل دو مسیر انقباضی برای استخراج ویژگی های سطح بالا و مسیر انبساطی برای بازسازی فضای مکانی است. اتصالات پرشی بین این دو مسیر، اطلاعات مکانی دقیق را حفظ کرده و با ویژگی های عمیق ترکیب می کنند؛ قابلیتی که در شناسایی ضایعات کوچک و کم کنتراست مرتبط با صرع اهمیت ویژه ای دارد [38]. نسخه های پیشرفته تر این معماری، مانند Attention U-Net، Residual U-Net و U-Net++، با افزودن مکانیسم ها و توپولوژی های غنی تر، به بهبود دقت و حساسیت در شناسایی نواحی کوچک و کم کنتراست پرداخته اند [39]. در معماری شکل (۶)، از بلوک های MultiRes برای استخراج ویژگی های چندمقیاسی استفاده شده است. این بلوک ها با کانولوشن های 3×3 و ترکیب نهایی با کانولوشن 1×1 ، ویژگی ها را به صورت یکپارچه ادغام می کنند. مسیر باقیمانده شامل کانولوشن های 3×3 و 1×1 ، شکاف معنایی بین رمزگذار و رمزگشا را کاهش می دهد [40].



شکل 6- نمایش کلی از معماری U-NET [40]

معماری DNN

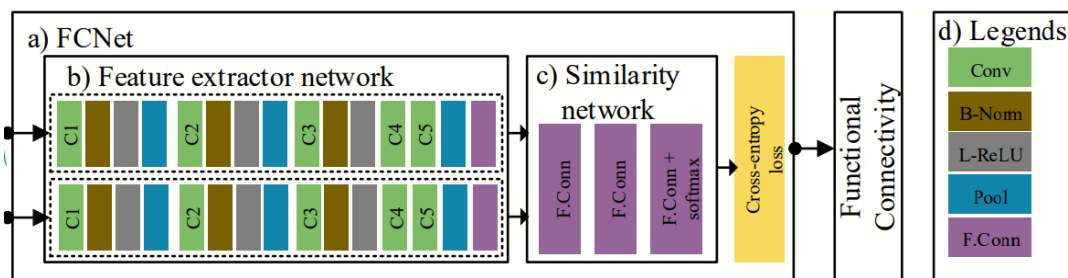
DNN به عنوان یکی از پایه‌های اصلی یادگیری عمیق، توانایی استخراج خودکار ویژگی‌های سلسله‌مراتبی از داده‌های پیچیده را فراهم کرده و عملکرد طبقه‌بندی تصاویر MRI مغز را در مقایسه با روش‌های سنتی یادگیری ماشین به طور چشمگیری بهبود می‌بخشد. DNN‌ها با یادگیری مستقیم از داده خام، به طور خودکار ویژگی‌های معنادار را در سطوح مختلف انتزاع جذب کرده و نیاز به مهندسی ویژگی را به طور کامل حذف می‌کنند. معماری DNN مورد بررسی در این مطالعه شامل چندین لایه کانولوشنی با فیلترهای گسترده، توابع فعال‌سازی ReLU و لایه‌های کاملاً متصل است. یکی از ویژگی‌های کلیدی این مدل، استفاده از معماری رمزگذار-رمزگشا است. این معماری به طور همزمان امکان حفظ ویژگی‌های ساختاری ضروری تصویر و حذف نویز و اطلاعات غیرمرتبط را فراهم می‌کند. معماری نمونه‌ای از این شبکه و فرآیند پردازش یک گره در شکل (7) نشان داده شده است [41].



شکل 7- نمونه‌ای از فرآیند پردازش با معماری DNN [41]

شبکه Fc-Net

Fc-Net یک شبکه عصبی کانولوشنی است که به طور خاص برای تحلیل داده‌های fMRI^۱ طراحی شده است. برخلاف روش‌های سنتی تحلیل اتصال عملکردی که عمدتاً بر اساس همبستگی زمانی بین نواحی مغزی و فاصله فیزیکی آن‌ها عمل می‌کنند، Fc-Net به صورت مستقیم و بدون نیاز به مراحل پیش‌پردازش مرسوم، اتصال عملکردی را از سیگنال‌های خام سری زمانی fMRI استخراج می‌کند. معماری Fc-Net از دو جزء اصلی تشکیل شده است: یک بخش استخراج ویژگی که الگوهای موضعی و سراسری در سیگنال‌های fMRI را شناسایی می‌کند و یک شبکه کاملاً متصل که در یک پیکربندی سیامی (Siamese) قرار گرفته است. پیکربندی سیامی به شبکه اجازه می‌دهد تا شباهت بین جفت نمونه‌ها مثلاً بین یک فرد سالم و یک بیمار مبتلا به صرع را به صورت موثر و مستقیم محاسبه کند. این ساختار با مقایسه مستقیم نمونه‌ها، دقت طبقه‌بندی را به طور چشمگیری افزایش می‌دهد [46]. نمایی شماتیک از ساختار الگوریتم Fc-Net در شکل (۸) ارائه شده است [42].



شکل 8- نمایی از الگوریتم FC-NET [42]

^۱ Functional Magnetic Resonance Imaging

3-نتایج

به منظور ارزیابی جامع عملکرد معماری‌های مختلف یادگیری عمیق در تشخیص صرع، مقالات پژوهشی معتبری که به‌طور مستقیم به این موضوع پرداخته‌اند، مورد بررسی و تحلیل قرار گرفتند. یافته‌های این مطالعات در جدول (۲) خلاصه شده‌اند و شامل اطلاعاتی درباره نوع مدل، پایگاه داده، مراحل پیش‌پردازش، روش استخراج ویژگی، الگوریتم طبقه‌بندی، دقت نهایی تشخیص و مقایسه‌ای کلی از ویژگی‌های کلیدی، مزایا و معایب هر معماری می‌باشد. تحلیل نتایج نشان می‌دهد که مدل‌های مبتنی بر شبکه‌های عصبی کانولوشنی به‌ویژه نسخه‌های پیشرفته آن عملکرد قابل توجهی در تشخیص صرع دارند. در میان این مدل‌ها، CNN در مطالعه Vijram و همکاران [19] با دقت 96/09٪ بالاترین دقت را به ثبت رسانده است و همچنین 3D-CNN در مطالعه Luckett و همکاران [22] به دقت چشمگیر 96٪ دست یافته است. مدل m-CNN نیز با استفاده از کانولوشن‌های مختلط و اقتباس از معماری VGG-16، دقت 87/7٪ را در مطالعه Ito و همکاران [24] کسب کرده و در استخراج الگوهای پیچیده در حوزه فضایی-فرکانسی (k-space) و حوزه تصویر، عملکرد مؤثری از خود نشان داده است. این مدل به‌ویژه در تشخیص صرع لوب تمپورال متوسط (MTLE) و اسکروز هیپوکامپی (HS) کارایی بالایی دارد. معماری DAG-CNN با ساختار گراف جهت‌دار و اتصالات پرشی (Skip Connections)، دقت 85٪ را در مطالعه Gleichgerricht و همکاران [17] گزارش کرده است. این مدل با امکان استخراج ویژگی‌های چندمقیاسی و بهبود جریان گرادینت در شبکه‌های عمیق، از نقاط قوت خود بهره می‌برد؛ هرچند ساختار پیچیده آن نیازمند تنظیم دقیق پارامترها و محاسبات سنگین‌تری است. مدل Fc-Net که اصلاً برای تحلیل داده‌های fMRI طراحی شده است، با بهره‌گیری از یک شبکه سیامی (Siamese) و استخراج مستقیم اتصال عملکردی از سیگنال‌های خام، دقت 86/72٪ را در تشخیص بیماران مبتلا به صرع به دست آورده است [20]. مدل DNN با معماری رمزگذار-رمزگشا، دقت 79/26٪ را در مطالعه Vajiram و همکاران [20] ارائه کرده است. این مدل با توانایی حذف نویز و بازسازی ویژگی‌های معنادار، مناسب شرایطی است که کیفیت تصاویر پایین‌تر یا نویز بالاتری دارند؛ اما عملکرد آن در مقایسه با مدل‌های پیشرفته‌تر، محدودتر است. در مطالعاتی که از U-Net استفاده شده است، این مدل بیشتر برای تقسیم‌بندی ناحیه ضایعات و نه طبقه‌بندی مستقیم به کار رفته است. با استفاده از ترکیب U-Net با مدل‌های پیش‌آموزش‌دیده مانند EfficientNet-B1، معیار ضریب دایس (DSC) برابر با 0/84 در مطالعه Arnold و همکاران [23] گزارش شده است. این عدد نشان‌دهنده دقت بالای این مدل در تعیین مرزهای دقیق ضایعات مغزی و کاربرد بالینی آن در برنامه‌ریزی جراحی است. تفاوت معماری و روند اجرای مدل‌های CNN در تشخیص صرع و اختلالات مغزی به‌وضوح قابل مشاهده می‌باشد: در مطالعه Gleichgerricht و همکاران [16]، از یک مدل CNN دوبعدی با معماری ساده استفاده شده که تصاویر MRI را به‌صورت اسلایس‌های دوبعدی پردازش می‌کند؛ تمرکز اصلی بر ارزیابی امکان‌سنجی تشخیص صرع با داده‌های محدود بوده است. در مطالعه Chang و همکاران [17] از یک معماری عمیق‌تر بهره گرفته که شامل چندین لایه کانولوشنی و Fully Connected است و توانسته بین صرع لوب تمپورال، آلزایمر و افراد سالم تمایز ایجاد کند؛ این مدل از داده‌های چندکلاسه و تنظیمات دقیق‌تری در آموزش استفاده کرده است. در مقاله Kaestner و همکاران [18]، یک مدل 3D-CNN طراحی شده که با فیلترهای سه‌بعدی، روابط فضایی بین وکسل‌ها را در سه محور تصویر MRI مدل‌سازی می‌کند؛ این ساختار برای تعیین سمت شروع تشنج بسیار مؤثر بوده است. در نهایت، Vajiram and Sivakumar [19] از CNN در کنار الگوریتم‌های حجمی مانند Marching Cube و Skull Stripping، Radiomic و Feature Selection استفاده کرده‌اند تا حجم ماده خاکستری، سفید و مایع مغزی نخاعی را در بیماران مبتلا به تومور صرعی تحلیل کنند؛ مدل آن‌ها با دقت 96/09٪ عملکرد بالایی داشته است. در مجموع، تفاوت اصلی این سه مدل CNN در نوع داده ورودی، عمق شبکه، وجود لایه‌های پیشرفته و نحوه استخراج ویژگی‌هاست که مستقیماً بر دقت و کاربرد بالینی آن‌ها تأثیر گذاشته‌اند. همچنین، یافته‌های این مطالعه نشان می‌دهد که تحلیل تصاویر MRI با استفاده از الگوریتم‌های یادگیری عمیق، علاوه بر تشخیص کلی صرع، در تمایز این بیماری از سایر اختلالات عصبی نیز مؤثر است. به‌عنوان مثال، در مطالعه Chang و همکاران [18]، یک مدل CNN توانسته است صرع لوب تمپورال را از بیماری آلزایمر با دقت 86٪ و از افراد سالم با دقت 81٪ تشخیص دهد.

جدول ۲- مروری بر مقالات بررسی شده جهت تشخیص صرع

روش طبقه بندی	نویسندگان	پیش پردازش	استخراج ویژگی	مزایا	معایب	ویژگی های کلیدی	دقت (%)
3D-CNN	Lockett و همکاران[20]	نرمال سازی ، حذف نویز، انتخاب اسلایس دو بعدی	Grad CAM[43]	تحلیل حجمی ، دقت بسیار بالا	محاسبات سنگین ، نیاز به سخت افزار قوی	فیلتر های سه بعدی برای مدل سازی فضایی کامل	96
	Gleichgerrcht و همکاران[44]	نرمال سازی تصاویر در فضای استاندارد، هم تراز و حذف نویز	یادگیری خودکار، نقشه های saliency برای شناسایی نواحی مغزی مهم				85/9
	Lasagni و همکاران[45]	هم تراز و تصاویر، نرمال سازی، حذف نویز، آماده سازی	یادگیری خودکار، Grad-CAM				80/3
Fc-Net	Kaestner و همکاران[15]	نرمال سازی ، حذف نویز ، تقسیم بندی دستی	استفاده از تفاوت ساختاری هیپوکامپ	استخراج اتصال عملکردی مستقیم	محدود در تحلیل ساختار	شبکه Siamese برای مقایسه عملکرد مغز	86/72
DAG-CNN	Gleichgerrch و همکاران[16]	نرمال سازی ، هم تراز و ، حذف ساختار غیر مغزی	ترکیب ویژگی های بافتی توسط CNN	استخراج چند مقیاسی ، انتقال بهتر ویژگی ها	ساختار پیچیده ، نیاز به تنظیم دقیق	گراف جهت دار با مسیر های میان لایه ای	85
CNN	Chang و همکاران[17]	نرمال سازی ، تقسیم بندی ، هموار سازی	استفاده از 2D-CNN	دقت بالا ، پردازش داده های حجیم	نیاز به داده های زیاد ، خطر بیش برآزش ، تفسیر پذیری نسبتا پایین و وابسته به کیفیت تصویر	یادگیری سلسله مراتبی ، انعطاف پذیری معماری ، قابلیت تعمیم پذیری	86
	Kaestner و همکاران[18]	نرمال سازی ، حذف نویز ، انتخاب بخش های کلیدی	استفاده از CNN و Grad-CAM	کاهش وابستگی به متخصص انسانی و کاهش خطای انسانی			78
	Vajiram و همکاران[19]	نرمال سازی ، تقسیم بندی دستی	استفاده از Radiomix و DL	96/09			
DNN	vajiram و همکاران[19]	نرمال سازی ، تقسیم بندی دستی	استفاده از Radiomix و DL	حذف نویز، مناسب برای تصاویر با کیفیت پایین	دقت کمتر ، وابسته به تنظیمات	رمزگذار و رمزگشا با لایه های Fully Connected	79/26
	Buttar و همکاران[46]	کاهش نویز ، نرمال سازی شدت تصویر، انتخاب اسلایس های با بیشترین شدت	استخراج ویژگی های بافتی و ساختاری				98/44
	Maher و همکاران[47]	ساخت کانکتوم از داده ها ، نرمال سازی ماتریس اتصال، حذف نویز	استفاده از وزن های اتصال بین نواحی مغز، تبدیل کانکتوم به بردار ویژگی،				67/86
U-Net	Thomas و همکاران[40]	نرمال سازی شدت تصویر ، برش ناحیه مغز ، تبدیل به 3D	استفاده از attention blocks	تقسیم بندی دقیق ، حفظ اطلاعات مکانی	کاهش حساسیت به نواحی کوچک و داده های نامتعادل	مسیر دوگانه با skip connection و attention	89/02
m-CNN	Ito و همکاران[22]	نرمال سازی ، تبدیل به JPEG	انتخاب برش	استخراج الگو های پیچیده فضایی-فرکانسی	حساس به نوع داده ، نیاز به تنظیمات خاص	کانولوشن مختلط ، اقتباس از VGG-16	87/7

3. نتیجه گیری

مدل‌های یادگیری عمیق، به‌ویژه شبکه‌های عصبی کانولوشنی و معماری‌های پیشرفته آن‌ها، تحولی چشمگیر در تشخیص و تحلیل بیماری‌های نورولوژیک از جمله صرع ایجاد کرده‌اند. این مدل‌ها با قابلیت استخراج خودکار ویژگی‌ها از داده‌های تصویری و بدون نیاز به مهندسی دستی، امکان تحلیل دقیق ساختارهای پیچیده مغز و شناسایی ضایعات ریز و غیرقابل مشاهده با چشم غیرمسلح را فراهم می‌کنند. یافته‌های این مطالعه مروری نشان می‌دهد که این فناوری‌ها نه تنها دقت تشخیص صرع را به‌طور چشمگیری افزایش داده‌اند، بلکه امکان تمایز این بیماری از سایر اختلالات عصبی مانند بیماری آلزایمر را نیز فراهم کرده‌اند.

از منظر کاربردی، این مدل‌ها در تشخیص زودهنگام صرع، تعیین دقیق محل کانون تشنج، برنامه‌ریزی جراحی‌های عصبی از طریق تقسیم‌بندی دقیق ضایعات و توسعه سیستم‌های پایش و هشدار بالینی نقش برجسته‌ای ایفا می‌کنند. مهم‌ترین مزایای این رویکرد شامل دقت بالا در شناسایی تغییرات ساختاری مرتبط با صرع، توانایی پردازش داده‌های سه‌بعدی و چندمنبعی، کاهش زمان تحلیل و بهبود تصمیم‌گیری بالینی است. این فناوری‌ها خطاهای ناشی از تفسیر دستی تصاویر را کاهش داده و زمینه را برای توسعه سیستم‌های کمک‌تشخیص هوشمند و برنامه‌ریزی درمانی دقیق‌تر فراهم می‌سازند.

با این حال، چالش‌های مهمی نیز در مسیر کاربرد گسترده این مدل‌ها وجود دارد. از جمله این چالش‌ها می‌توان به نیاز به داده‌های برچسب‌خورده در حجم زیاد، وابستگی به کیفیت تصاویر و مراحل پیش‌پردازش، هزینه محاسباتی بالا، نیاز به سخت‌افزار قدرتمند و ماهیت جعبه سیاه این مدل‌ها اشاره کرد. این موارد اعتماد بالینی پزشکان به خروجی سیستم‌ها را محدود می‌کنند. بنابراین، توسعه الگوریتم‌های قابل تفسیر که بتوانند منطق تصمیم‌گیری خود را به متخصصان نشان دهند، ضرورتی اجتناب‌ناپذیر است. چشم‌انداز آینده این حوزه به سمت ادغام داده‌های چندمنبعی از جمله EEG، MRI، داده‌های ژنتیکی و بالینی حرکت می‌کند. این رویکرد می‌تواند دقت و قابلیت تعمیم مدل‌ها را به‌طور چشمگیری افزایش دهد. همچنین، استفاده از روش‌هایی مانند یادگیری فدرال می‌تواند به حفظ حریم خصوصی بیماران کمک کند، در حالی که مدل‌های قابل تفسیر شفافیت و اعتمادپذیری سیستم‌ها را افزایش می‌دهند.

در نهایت، یکپارچه‌سازی این فناوری‌های هوشمند با سیستم‌های تصمیم‌یار بالینی، مسیری نوین را در جهت پزشکی دقیق و شخصی‌سازی شده باز خواهد کرد. به‌طور کلی، یافته‌های این پژوهش گواهی بر پتانسیل بالای مدل‌های یادگیری عمیق در ارتقای دقت تشخیص صرع، کاهش زمان پردازش، تمایز این بیماری از سایر اختلالات عصبی و بهبود برنامه‌ریزی درمانی بیماران مبتلاست. ادغام پردازش تصاویر MRI با الگوریتم‌های پیشرفته یادگیری عمیق، نه تنها چشم‌انداز نوینی از پزشکی مبتنی بر داده را ترسیم می‌کند، بلکه به متخصصان ابزارهایی دقیق، سریع و قابل اتکا برای تحلیل بیماری‌های مغزی ارائه می‌دهد.

4. منابع و مراجع

1. Ilakiyaselvan, N., A.N. Khan, and A. Shahina, *Deep learning approach to detect seizure using reconstructed phase space images*. Journal of biomedical research, 2020. 34(3): p. 240.
2. Fisher, R.S., et al., *ILAE official report: a practical clinical definition of epilepsy*. Epilepsia, 2014. 55(4): p. 475-482.
3. Havasimehr, M., et al., *A review on the animal models of seizure*. 2018.
4. Scheffer, I.E., et al., *ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology*. Epilepsia, 2017. 58(4): p. 512-521.
5. Fisher, R.S., et al., *Operational classification of seizure types by the International League Against Epilepsy: Position Paper of the ILAE Commission for Classification and Terminology*. Epilepsia, 2017. 58(4): p. 522-530.
6. DeFelipe, J., et al., *Neuroanatomical and psychological considerations in temporal lobe epilepsy*. Frontiers in neuroanatomy, 2022. 16: p. 995286.

7. Yoganathan, K., et al., *Neurological update: structural and functional imaging in epilepsy surgery*. Journal of Neurology, 2023. 270(5): p. 2798-2808.
8. Qian, Z., et al., *Evaluation of MRI post-processing methods combined with PET in detecting focal cortical dysplasia lesions for patients with MRI-negative epilepsy*. Seizure-European Journal of Epilepsy, 2024. 117: p. 275-283.
9. Rogeau, A., et al., *Identifying and troubleshooting the pitfalls of ictal/interictal brain perfusion SPECT studies*. Nuclear Medicine Communications, 2023. 44(11): p. 1053-1058.
10. Shahraki, G. and E. Irankhah. *Diagnosis of epilepsy disease with MRI images analysis and EEG signal processing*. in *Proceedings of the 2nd International Conference on Electronics, Biomedical Engineering, and Health Informatics: ICEBEHI 2021, 3–4 November, Surabaya, Indonesia*. 2022. Springer.
11. Litjens, G., et al., *A survey on deep learning in medical image analysis*. Medical image analysis, 2017. 42: p. 60-88.
12. <A_Finite_Element_Method_for_Deformable_Models.pdf>.
13. Yazdani, A., S. Costa, and B. Kroon, *Artificial intelligence: Friend or foe?* 2023, Wiley Online Library. p. 127-130.
14. Shen, D., G. Wu, and H.-I. Suk, *Deep learning in medical image analysis*. Annual review of biomedical engineering, 2017. 19(1): p. 221-248.
15. Kaestner, E., et al., *Adding the third dimension: 3D convolutional neural network diagnosis of temporal lobe epilepsy*. Brain Communications, 2024. 6(5): p. fcae346.
16. Gleichgerrcht, E., et al., *Radiological identification of temporal lobe epilepsy using artificial intelligence: a feasibility study*. Brain Communications, 2022. 4(2): p. fcab284.
17. Chang, A.J., et al., *MRI-based deep learning can discriminate between temporal lobe epilepsy, Alzheimer's disease, and healthy controls*. Communications Medicine, 2023. 3(1): p. 33.
18. Kaestner, E., et al., *Convolutional neural network algorithm to determine lateralization of seizure onset in patients with epilepsy: A proof-of-principle study*. Neurology, 2023. 101(3): p. e324-e335.
19. Vajiram, J. and S. Sivakumar, *Quantitative Assessment of Brain Tumor Epilepsy by AI Models*.
20. Luckett, P.H., et al., *Deep learning resting state functional magnetic resonance imaging lateralization of temporal lobe epilepsy*. Epilepsia, 2022. 63(6): p. 1542-1552.
21. Arnold, T.C., et al., *Deep learning-based automated segmentation of resection cavities on postsurgical epilepsy MRI*. NeuroImage: Clinical, 2022. 36: p. 103154.
22. Ito, Y., et al., *Deep learning-based diagnosis of temporal lobe epilepsy associated with hippocampal sclerosis: An MRI study*. Epilepsy Research, 2021. 178: p. 106815.
23. Singh, N.T., et al. *Preprocessing of medical images using deep learning: A comprehensive review*. in *2023 Second International Conference on Augmented Intelligence and Sustainable Systems (ICAISS)*. 2023. IEEE.
24. Messadi, M., A. Bessaid, and A. Taleb-Ahmed, *Extraction of specific parameters for skin tumour classification*. Journal of medical engineering & technology, 2009. 33(4): p. 288-295.
25. Tierney, T.M., et al., *SPM 25: open source neuroimaging analysis software*. arXiv preprint arXiv:2501.12081, 2025.
26. Kauppi, J.-P., J. Pajula, and J. Tohka, *A versatile software package for inter-subject correlation based analyses of fMRI*. Frontiers in neuroinformatics, 2014. 8: p. 2.
27. Candemir, C., *Spatial smoothing effect on group-level functional connectivity during resting and task-based fMRI*. Sensors, 2023. 23(13): p. 5866.
28. Zhang, Y., M. Brady, and S. Smith, *Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm*. IEEE transactions on medical imaging, 2002. 20(1): p. 45-57.
29. Thiele, F., et al., *P15. 09. A Canny edge detection algorithm for quantitative differentiation between diffuse and circumscript glioma growth patterns on MRI*. Neuro-Oncology, 2022. 24(Suppl 2): p. ii85.

30. Yushkevich, P.A., Y. Gao, and G. Gerig. *ITK-SNAP: An interactive tool for semi-automatic segmentation of multi-modality biomedical images*. in 2016 38th annual international conference of the IEEE engineering in medicine and biology society (EMBC). 2016. IEEE.
31. Peng, H., et al., *Accurate brain age prediction with lightweight deep neural networks*. Medical image analysis, 2021. 68: p. 101871.
32. Raisi, A., M. Nasiri, and H.J.J.o.I.U.o.M.S. Danesh, *Diagnostic Techniques and Artificial Intelligence for Multiple Sclerosis Identification*. 2025. 33(3): p. 113-148.
33. Bernal, J., et al., *Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: a review*. Artificial intelligence in medicine, 2019. 95: p. 64-81.
34. Yang, S. and D. Ramanan. *Multi-scale recognition with DAG-CNNs*. in *Proceedings of the IEEE international conference on computer vision*. 2015.
35. Çiçek, Ö., et al. *3D U-Net: learning dense volumetric segmentation from sparse annotation*. in *International conference on medical image computing and computer-assisted intervention*. 2016. Springer.
36. Wang, Z., J. Yang, and M. Sawan. *A novel multi-scale dilated 3D CNN for epileptic seizure prediction*. in *2021 IEEE 3rd International Conference on Artificial Intelligence Circuits and Systems (AICAS)*. 2021. IEEE.
37. Ren, Y., W. Jiang, and Y. Liu. *A complex-valued dual-domain dilated convolution neural network for brain mri reconstruction*. in *2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. 2022. IEEE.
38. Nguyen, M.K., P.H. Tran, and T.T. Phan, *Using U-Net models in deep learning for brain tumor detection from MRI scans*. CTU Journal of Innovation and Sustainable Development, 2024. 16(Special issue: ISDS): p. 91-97.
39. Azad, R., et al., *Medical image segmentation review: The success of u-net*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024.
40. Thomas, E., et al., *Multi-res-attention UNet: a CNN model for the segmentation of focal cortical dysplasia lesions from magnetic resonance images*. IEEE journal of biomedical and health informatics, 2020. 25(5): p. 1724-1734.
41. Thost, V. and J. Chen, *Directed acyclic graph neural networks*. arXiv preprint arXiv:2101.07965, 2021.
42. Riaz, A., et al. *Fcnet: a convolutional neural network for calculating functional connectivity from functional mri*. in *International Workshop on Connectomics in Neuroimaging*. 2017. Springer.
43. Selvaraju, R.R., et al. *Grad-cam: Visual explanations from deep networks via gradient-based localization*. in *Proceedings of the IEEE international conference on computer vision*. 2017.
44. Gleichgerrcht, E., et al., *Redefining diagnostic lesional status in temporal lobe epilepsy with artificial intelligence*. Brain, 2025. 148(6): p. 2189-2200.
45. Lasagni, L., et al., *Focal cortical dysplasia type II detection using cross modality transfer learning and grad-CAM in 3D-CNNs for MRI analysis*. arXiv preprint arXiv:2504.07775, 2025.
46. Buttar, A.M., et al., *Enhanced neurological anomaly detection in MRI images using deep convolutional neural networks*. Frontiers in Medicine, 2024. 11: p. 1504545.
47. Maher, C., et al., *Deep learning distinguishes focal epilepsy groups using connectomes: Feasibility and clinical implications*. medRxiv, 2023: p. 2023.02. 09.23285681.

Deep Learning Approaches for Epilepsy Detection in MRI Images

Negin shafizade¹, Hajar Danesh², Farhad Khosravi³, Shekofeh Yaraghi⁴

¹Department of Electrical and Biomedical Engineering, Faculty of Engineering and Technology, Shahid Ashrafi Esfahani University, Isfahan, Iran, negin@ashrafi.ac.ir

² Department of Electrical and Biomedical Engineering, Faculty of Engineering and Technology, Shahid Ashrafi Esfahani University, Isfahan, Iran, h.danesh@ashrafi.ac.ir

³ Department of Electrical and Biomedical Engineering, Faculty of Engineering and Technology, Shahid Ashrafi Esfahani University, Isfahan, Iran, fkha@ashrafi.ac.ir

⁴ Department of Computer, Faculty of Engineering, Shahid Ashrafi Esfahani University, Isfahan, Iran, sh.yaraghi@ashrafi.ac.ir

Abstract— Epilepsy is one of the most prevalent neurological disorders, and its early diagnosis is critical to preventing neurological damage. Magnetic resonance imaging (MRI) is regarded as the standard diagnostic tool for identifying epilepsy-related brain lesions due to its high resolution and non-invasive nature. However, manual interpretation of MRI scans is time-consuming and prone to human error. In recent years, convolutional neural networks (CNNs) have enabled automated analysis of MRI data and the detection of complex pathological patterns. This review evaluates the performance of several deep learning architectures—including Fc-Net, 3D-CNN, DAG-CNN, U-Net, DNN, and m-CNN—in the context of epilepsy detection. Findings indicate that CNN-based models achieve the highest diagnostic accuracy, with some studies reporting performance exceeding 96%. Despite these advances, significant challenges remain, including data heterogeneity, model interpretability, and the requirement for high-performance computing resources, all of which necessitate further investigation.

Keywords: Epilepsy, Deep Learning, Convolutional Neural Network, MRI, Automated Diagnosis