

# طراحی یک مدل مبتنی بر یادگیری عمیق برای طبقه‌بندی بیماری آلزایمر

زهرا آقازارتی

دانشجوی کارشناسی ارشد، گروه مهندسی فناوری اطلاعات، دانشکده مهندسی صنایع،  
دانشگاه خواجه نصیرالدین طوسی، تهران، ایران

منیره حسینی\*

دانشیار، گروه مهندسی فناوری اطلاعات، دانشکده مهندسی صنایع،  
دانشگاه خواجه نصیرالدین طوسی، تهران، ایران

مدیریت

اطلاعات

دوره ۱۱، شماره ۱

بهار و تابستان ۱۴۰۴

**چکیده:** بیماری آلزایمر به عنوان شایع‌ترین علت زوال عقل، یک بحران رو به رشد در بهداشت جهانی است. تشخیص دقیق و به‌موقع این بیماری، به‌ویژه در مراحل اولیه، نقشی حیاتی در مدیریت علائم و بهبود کیفیت زندگی بیماران دارد. در سال‌های اخیر، الگوریتم‌های یادگیری ماشین، یادگیری عمیق و به‌طور خاص شبکه‌های عصبی پیچشی، پتانسیل بالایی برای تحلیل خودکار تصاویر عصبی با دقت بالا از خود نشان داده‌اند. با این حال، اعتبار بسیاری از مطالعات پیشین به دلیل خطای نشت داده مورد تردید قرار می‌گیرد. این پژوهش با هدف غلبه بر این چالش، یک چارچوب یادگیری عمیق معتبر و کارآمد برای طبقه‌بندی گروه بیماری آلزایمر از افراد سالم با استفاده از تصویربرداری تشدید مغناطیسی ساختاری از مجموعه داده طرح تصویربرداری عصبی بیماری آلزایمر ارائه می‌کند. نوآوری اصلی این تحقیق در طراحی یک معماری اختصاصی است. این معماری به جای استفاده از مدل‌های از پیش آموزش‌دیده، از پایه طراحی شده است که با استفاده از کانولوشن‌های جداپذیر عمقی، توازن بین کارایی محاسباتی و قدرت استخراج ویژگی برقرار می‌کند. برای اعتبار نتایج، این پژوهش از بروز هرگونه نشت اطلاعات جلوگیری کرده است. ارزیابی مدل نهایی بر روی مجموعه داده آزمون، عملکرد بسیار خوبی را از خود نشان داد و به دقت ۹۵.۱۲٪ و سطح زیر منحنی مشخصه عملکرد گیرنده معادل ۰.۹۴ دست یافت. این نتایج نشان می‌دهد که یک معماری هدفمند و روش‌شناسی معتبر می‌تواند به ابزاری قابل اطمینان برای کمک به تشخیص زودهنگام بیماری آلزایمر تبدیل شود.

**کلیدواژه‌ها:** بیماری آلزایمر، طبقه‌بندی، یادگیری عمیق، شبکه‌های عصبی پیچشی، نشت داده.

## مقدمه

زوال عقل یکی از معضلات رو به رشد بهداشت عمومی در سطح جهان است و هفتمین علت اصلی مرگ و میر در جهان محسوب می‌شود. این بیماری اهمیت ویژه‌ای دارد چراکه به تدریج حافظه، تفکر، رفتار و توانایی انجام فعالیت‌های روزمره را مختل می‌کند. بار اقتصادی آن نیز بسیار سنگین است؛ به گونه‌ای که در سال ۲۰۱۹ هزینه جهانی زوال عقل حدود ۱.۳ تریلیون دلار برآورد شد (سازمان جهانی بهداشت<sup>۱</sup>، ۲۰۲۵). بیماری آلزایمر نیز مهم‌ترین علت زوال عقل است که به سرعت به یکی از بیماری‌های پرهزینه و کشنده قرن بیستم تبدیل می‌شود (شلتنز<sup>۲</sup> و همکاران، ۲۰۲۱). آلزایمر یک فرآیند غیرقابل برگشت از تحلیل رفتن مغز است که می‌تواند منجر به اختلال در حافظه، شناخت، شخصیت و سایر عملکردهای حیاتی و در نهایت مرگ شود. این بیماری برخلاف آن چه عموم تصور می‌کنند، بخشی طبیعی از روند پیری نیست (بنیاد برایت‌فوکوس<sup>۳</sup>، ۲۰۲۵). مراحل بالینی آلزایمر به ترتیب شامل شناخت طبیعی، اختلال شناختی خفیف و زوال عقل است، که ممکن است حدود ۱۵ الی ۲۵ سال طول بکشد (شلتنز و همکاران، ۲۰۲۱). تشخیص زودهنگام آلزایمر نقشی حیاتی در کند کردن روند بیماری، افزایش کیفیت زندگی بیماران و کاهش بار اقتصادی و روانی آن دارد (بانرجی<sup>۴</sup> و همکاران، ۲۰۲۰).

یکی از راه‌هایی که پزشکان برای تشخیص آلزایمر استفاده می‌کنند، بررسی تصاویر مغزی مثل تصویربرداری تشدید مغناطیسی<sup>۵</sup> یا توموگرافی گسیل پوزیترون<sup>۶</sup> است. با این حال، به علت حجم عظیم و پیچیدگی داده‌های حاصل از تصویربرداری عصبی، تحلیل دستی آن‌ها با چالش‌های فراوانی روبرو است (آریا<sup>۷</sup> و همکاران، ۲۰۲۳). هوش مصنوعی، پتانسیل فوق‌العاده‌ای برای خودکارسازی تشخیص این بیماری از خود نشان داده است. این تکنولوژی با توانایی در پردازش حجم بزرگی از داده‌ها، به پزشکان کمک می‌کند تا تشخیص‌های دقیق‌تر و سریع‌تری انجام دهند. در اینجاست که یادگیری ماشین<sup>۸</sup> و یادگیری عمیق<sup>۹</sup> به عنوان ابزارهای قدرتمند وارد میدان می‌شوند. این تکنیک‌ها قادرند الگوهای پیچیده و پنهان در داده‌ها را استخراج کرده و به تشخیص، پیش‌بینی و طبقه‌بندی بیماران با دقت بالا کمک کنند (آریا و همکاران، ۲۰۲۳).

<sup>1</sup> World Health Organization

<sup>2</sup> Scheltens

<sup>3</sup> BrightFocus Foundation

<sup>4</sup> Banerjee

<sup>5</sup> Magnetic Resonance Imaging (MRI)

<sup>6</sup> Positron Emission Tomography (PET)

<sup>7</sup> Arya

<sup>8</sup> Machine Learning

<sup>9</sup> Deep Learning

## پیشینه پژوهش

مطالعاتی که با هدف تشخیص آلزایمر انجام شده‌اند و از تصاویر MRI بهره می‌برند، معمولاً از برش‌های دوبعدی، اسکن‌های سه‌بعدی به‌عنوان داده‌ی ورودی استفاده کرده‌اند (والیانی و سونی<sup>۱</sup>، ۲۰۱۷؛ کورولف<sup>۲</sup> و همکاران، ۲۰۱۷). رویکرد دوبعدی به دلیل سادگی و حجم زیاد داده‌های تولیدی از اسکن سه‌بعدی، محبوب‌تر است. برش‌های دوبعدی MRI می‌توانند از نماهای مختلفی شامل افقی<sup>۳</sup>، کرونال<sup>۴</sup> و ساژیتال<sup>۵</sup> استخراج شوند و هر یک از این نماها اطلاعات مکملی از ساختار مغز ارائه می‌دهند. والیانی و سونی (۲۰۱۷) یکی از کارها در این زمینه بودند که با استفاده از برش‌های دوبعدی MRI و یک معماری شبکه‌های عصبی پیچشی<sup>۶</sup> از پیش آموزش دیده‌شده<sup>۷</sup>، به دقت ۸۱.۳٪ در طبقه‌بندی بیماران AD<sup>۸</sup> در مقابل افراد سالم دست یافتند. رشید<sup>۹</sup> و همکاران (۲۰۲۳) از یک معماری ترکیبی مبتنی بر VGG-16 استفاده کردند و با به‌کارگیری یک ماژول Biceph، به دقت ۱۰۰٪ در صفحات افقی و ساژیتال و دقت ۹۹.۴۵٪ در صفحه کرونال در تقسیم‌بندی CN<sup>۱۰</sup> و AD دست یافتند. ژانگ<sup>۱۱</sup> و همکاران (۲۰۲۲) با استفاده از تکنیک افزایش داده<sup>۱۲</sup> رندوم دو مرحله‌ای، توانستند عملکرد مدل‌های CNN مانند EfficientNet-B1 و DenseNet-169 را بهبود بخشیده و به دقت ۹۳.۲٪ در طبقه‌بندی بیماران AD در برابر افراد سالم دست یابند. این مطالعه بر اهمیت افزایش داده برای مقابله با کمبود داده‌های پزشکی تأکید دارد. آتنافو و دیچیوتی<sup>۱۳</sup> (۲۰۲۴) از معماری VGG-16 از پیش آموزش‌دیده استفاده کردند و به دقت ۷۱.۶۲٪ در مجموعه اعتبارسنجی رسیدند. گوئل<sup>۱۴</sup> و همکاران (۲۰۲۴) رویکردی نوآورانه را با ترکیب CNN و حافظه کوتاه‌مدت بلند<sup>۱۵</sup> پیشنهاد دادند. این مدل ترکیبی با دستیابی به دقت متعادل ۹۸.۳۳٪، عملکرد فوق‌العاده‌ای را نشان داد. برخی مطالعات با استفاده از مدل‌ها و رویکردهای پیچیده‌تر، سعی در بهبود عملکرد طبقه‌بندی دارند. مالو<sup>۱۶</sup> و همکاران (۲۰۲۴) با پیشنهاد شبکه CirMNet، یک رویکرد ترکیبی برای استخراج ویژگی‌ها ارائه دادند. معماری CirMNet دارای دو مسیر پردازش موازی است. یک مسیر از CNN برای استخراج ویژگی‌های

<sup>1</sup> Valliani & Soni<sup>2</sup> Korolev<sup>3</sup> Axial<sup>4</sup> Coronal<sup>5</sup> Sagittal<sup>6</sup> Convolutional Neural Networks (CNNs)<sup>7</sup> Pre-Trained<sup>8</sup> Alzheimer's Disease<sup>9</sup> Rashid<sup>10</sup> Cognitively Normal<sup>11</sup> Zhang<sup>12</sup> Data Augmentation<sup>13</sup> Atnafu & Diciotti<sup>14</sup> Goel<sup>15</sup> Long Short-Term Memory (LSTM)<sup>16</sup> Malu

بافتی و یک مسیر از CMSMD برای استخراج ویژگی‌های ساختاری استفاده می‌کند. این مدل توانست به دقت ۹۷.۳۴٪ در طبقه‌بندی چهار مرحله‌ی مختلف AD دست یابد که نشان‌دهنده مزیت استخراج همزمان ویژگی‌های چندگانه است. اسلام و ژانگ<sup>۱</sup> (۲۰۱۸) و هو<sup>۲</sup> و همکاران (۲۰۲۱) از روش رای‌گیری استفاده کردند. این رویکرد با ترکیب نتایج طبقه‌بندی چندین برش یا مدل، تصمیم نهایی را اتخاذ کرده و به ترتیب دقت ۹۳.۱۸٪ و ۹۳.۱۰٪ را گزارش دادند. این روش به طور مؤثری خطای ناشی از طبقه‌بندی یک برش را کاهش می‌دهد. کارکاگنی<sup>۳</sup> و همکاران (۲۰۲۳) روی داده‌های ADNI-2 و OASIS-1 کار کردند و با استفاده از CNN دوبعدی، دقت ۷۷٪ را گزارش کردند. برخی دیگر از مقالات نیز از تصاویر سه‌بعدی به عنوان ورودی مدل استفاده کردند. استفاده از اسکن‌های سه‌بعدی کامل به مدل اجازه می‌دهد تا اطلاعات ساختاری و حجمی مغز را به طور کامل درک کند. نتامپاکیس<sup>۴</sup> و همکاران (۲۰۲۴) از یک رویکرد سه‌بعدی به دقت ۹۴.۱۲٪ رسیدند. همچنین، این مطالعه با استفاده از مجموعه داده طرح تصویربرداری عصبی بیماری آلزایمر<sup>۵</sup> برای اعتبارسنجی، قابلیت تعمیم‌پذیری مدل را تأیید کرد، که یک نکته مهم در پژوهش‌های پزشکی است. ساراتچاگا<sup>۶</sup> و همکاران (۲۰۲۱) هر دو رویکرد دوبعدی و سه‌بعدی را بررسی کردند. آن‌ها با تمرکز بر ۱۰ برش مرکزی، به دقت متعادل ۹۳٪ دست یافتند. کورولف و همکاران (۲۰۱۷) نیز با معرفی مدل VoxCNN، یک شبکه سه‌بعدی را برای پردازش MRI سه‌بعدی توسعه دادند و دقت حدود ۷۹٪ و AUC<sup>۷</sup> حدود ۰.۸۸ گزارش کردند. دیناگار<sup>۸</sup> و همکاران (۲۰۲۳) از یک رویکرد کامل سه‌بعدی برای پردازش MRI استفاده کردند و به AUC حدود ۰.۸۹ دست یافتند. این مطالعه نشان داد که استفاده از کل حجم MRI می‌تواند اطلاعات ساختاری و حجمی مغز را با عملکرد خوبی استخراج کند، هرچند هنوز عملکرد آن از بعضی مدل‌های دوبعدی با داده‌افزایی پیشرفته پایین‌تر است.

همان‌طور که مرور شد، مطالعات متعددی در حوزه‌ی تشخیص این بیماری به نتایج امیدوارکننده‌ای دست یافته‌اند. با این حال، طبق پژوهش ون<sup>۹</sup> و همکاران (۲۰۲۰)، بسیاری از مطالعات پیشین با خطای نشت داده<sup>۱۰</sup> و یا شفاف نبودن عدم رخداد این خطا، روبرو هستند که اعتبار نتایج آن‌ها را زیر سؤال می‌برد. این خطا می‌تواند دقت ظاهری مدل را به طور کاذب بین ۲۹ تا ۵۵ درصد افزایش دهد (یاگیش<sup>۱۱</sup> و همکاران،

<sup>1</sup> Islam & Zhang

<sup>2</sup> Hu

<sup>3</sup> Carcagnì

<sup>4</sup> Ntampakis

<sup>5</sup> Alzheimer's Disease Neuroimaging Initiative (ADNI)

<sup>6</sup> Saratzaga

<sup>7</sup> Area Under the Curve

<sup>8</sup> Dhinagar

<sup>9</sup> Wen

<sup>10</sup> Data Leakage

<sup>11</sup> Yagis

(۲۰۲۱) و منجر به نتایجی بیش از حد خوش‌بینانه شود. این یافته‌ها در مطالعه مروری یانگ<sup>۱</sup> و همکاران (۲۰۲۵) نیز تأیید شد. این پژوهش با بررسی مطالعات یادگیری عمیق در حوزه تشخیص آلزایمر، رابطه‌ای معکوس میان دقت گزارش‌شده و صحت روش‌شناختی آشکار کرد؛ به‌طوری که مطالعاتی که تقسیم‌بندی صحیح در سطح بیمار را رعایت کرده‌اند، دقت‌هایی در بازه ۶۶ تا ۹۰٪ گزارش داده‌اند، در حالی که مطالعات با ریسک بالای نشت داده ادعای دقت ۹۵ تا ۹۹٪ داشتند.

از دلایل بروز این خطا می‌توان به تقسیم اشتباه مجموعه داده، تقسیم دیر هنگام مجموعه داده، یادگیری انتقالی دارای سوگیری<sup>۲</sup> و نبود مجموعه آزمون مستقل اشاره کرد (ون و همکاران، ۲۰۲۰). این مسئله یک شکاف تحقیقاتی مهم در این حوزه به شمار می‌رود. آیا می‌توان بدون استفاده از مدل‌های از پیش آموزش‌دیده و با پایبندی به پروتکل‌های جلوگیری از نشت داده، به دقتی قابل رقابت با سایر مدل‌ها دست یافت؟ این پژوهش با درک این چالش، به دنبال ارائه یک چارچوب معتبر و قابل اعتماد برای تشخیص بیماری آلزایمر است. هدف اصلی این مطالعه، توسعه و ارزیابی یک معماری CNN دوبعدی نوین است که از پایه برای طبقه‌بندی بیماران مبتلا به آلزایمر از افراد سالم با استفاده از تصویربرداری تشدید مغناطیسی ساختاری<sup>۳</sup> از مجموعه داده ADNI طراحی شده است. این تحقیق بر پیاده‌سازی یک پروتکل ارزیابی بی‌نقص برای تضمین اعتبار نتایج تأکید دارد و تلاش می‌کند توازن بین دقت بالا و کارایی محاسباتی برقرار کند.

## روش پژوهش

این پژوهش یک مطالعه کمی و تجربی است که در آن یک مدل یادگیری عمیق برای طبقه‌بندی تصاویر پزشکی توسعه و ارزیابی می‌شود.

## نشت داده

یکی از حساس‌ترین مراحل در تحقیقات یادگیری ماشین و یادگیری عمیق، تقسیم‌بندی صحیح داده‌ها برای جلوگیری از نشت اطلاعات است. برای جلوگیری از این مشکل، در این پژوهش یک استراتژی تقسیم‌بندی سخت‌گیرانه در سطح بیمار اتخاذ شد (ون و همکاران، ۲۰۲۰). تمامی برش‌های مربوط به یک بیمار باید تنها در یکی از مجموعه‌های آموزش، اعتبارسنجی یا آزمون قرار گیرند و به هیچ‌وجه نباید برش‌های متعلق به یک فرد در میان این مجموعه‌ها پراکنده شوند. تقسیم‌بندی داده‌ها باید پیش از ایجاد برش‌ها و در مرحله‌ای که داده‌ها به‌صورت تصاویر سه‌بعدی هستند انجام شود. چراکه تقسیم داده باید در سطح بیمار صورت گیرد، نه در سطح برش، تا از نشت اطلاعات بین مجموعه‌ها جلوگیری شود. چنین

<sup>1</sup> Young

<sup>2</sup> Biased Transfer Learning

<sup>3</sup> Structural Magnetic Resonance Imaging (sMRI)

رویکردی اطمینان حاصل می‌کند که مدل در مرحله‌ی آزمون تنها با داده‌هایی از بیمارانی مواجه شود که در طول مراحل آموزش و تنظیم مدل هرگز مشاهده نشده‌اند، و در نتیجه، ارزیابی عملکرد مدل به‌صورت واقعی‌تر و قابل اعتمادتر انجام شود.

همچنین در مواردی که برای یک بیمار چندین تصویر MRI سه‌بعدی در دسترس است، ترجیح بر آن است که تمام تصاویر مربوط به آن بیمار تنها در یکی از این مجموعه‌ها قرار گیرند. با این حال، رویکرد این پژوهش حذف تصاویر تکراری و نگهداشتن تنها یک اسکن از هر بیمار بوده است تا از وجود چند تصویر متعلق به یک فرد جلوگیری شود.

### مجموعه داده

داده‌های مورد استفاده از مجموعه داده عمومی ADNI استخراج شد. ADNI یک مطالعه‌ی چندمرکزی، طولی و بزرگ‌مقیاس است که از سال ۲۰۰۴ توسط مؤسسه ملی سالمندی آمریکا آغاز شد. هدف اصلی این پروژه، توسعه و ارزیابی نشانگرهای زیستی برای تشخیص زودهنگام و پایش پیشرفت بیماری آلزایمر است. این داده‌ها شامل تصویربرداری‌های مغزی، آزمایش‌های ژنتیکی، داده‌های بالینی و تست‌های شناختی هستند و به صورت عمومی در اختیار پژوهشگران سراسر جهان قرار می‌گیرند. ADNI-1 اولین فاز رسمی این پروژه است که در سال‌های ۲۰۰۴ تا ۲۰۱۰ اجرا شد. در این فاز، حدود ۸۰۰ شرکت‌کننده شامل سه گروه اصلی CN، MCI<sup>۱</sup> و AD وارد مطالعه شدند. تمرکز این فاز بیشتر بر توسعه و اعتبارسنجی استانداردهای تصویربرداری و استخراج نشانگرهای قابل استفاده در کارآزمایی‌های بالینی بود.

پس از اعمال فیلترهای مشخص شده در جدول ۱ بر روی مجموعه داده و حذف اسکن‌های متعدد برای هر بیمار، مجموعه داده نهایی شامل ۳۴۳ تصویر از افراد منحصربه‌فرد شد. از این تعداد، ۱۷۴ نفر زن و ۱۶۹ نفر مرد در بازه سنی ۵۵.۲ الی ۹۱ سال با مقدار آزمون کوتاه وضعیت شناختی<sup>۲</sup> برابر ۱۸ الی ۳۰ بودند. مشخصات توزیع نمونه در جدول ۲ خلاصه شده است.

جدول ۱. فیلترهای انتخاب داده از ADNI این جدول معیارهای انتخاب تصاویر از پایگاه داده ADNI را نشان می‌دهد.

| فیلتر            | مقدار انتخاب شده |
|------------------|------------------|
| پروژه            | ADNI 1           |
| گروه مورد مطالعه | CN و AD          |
| ویزیت            | ADNI Screening   |
| نوع تصویر        | MRI              |
| توضیح            | MPR; GradWarp    |
| شدت میدان        | ۱.۵ تسلا         |
| ضخامت برش        | ۱.۲ میلی متر     |

<sup>۱</sup> Mild Cognitive Impairment

<sup>۲</sup> Mini-Mental State Examination (MMSE)

| مقدار انتخاب شده | فیلتر   |
|------------------|---------|
| T1               | وزن دهی |

جدول ۲. مشخصات توزیع نمونه. توزیع مجموعه داده بین زیرمجموعه‌های آموزش، اعتبارسنجی و آزمون. تقسیم‌بندی در

سطح بیمار انجام شده تا از نشت اطلاعات جلوگیری شود.

| مجموعه داده | درصد از کل | نقش                             | سطح تقسیم‌بندی | تعداد نمونه | تعداد آزمودنی‌های AD | تعداد آزمودنی‌های CN |
|-------------|------------|---------------------------------|----------------|-------------|----------------------|----------------------|
| آموزش       | ۸۰٪        | آموزش وزن‌های مدل               | بیمار          | ۲۷۴         | ۱۲۶                  | ۱۴۸                  |
|             |            |                                 | برش            | ۳۵۶۲        |                      |                      |
| اعتبارسنجی  | ۸٪         | تنظیم فرآپارامترها و انتخاب مدل | بیمار          | ۲۸          | ۱۳                   | ۱۵                   |
|             |            |                                 | برش            | ۳۶۴         |                      |                      |
| آزمون       | ۱۲٪        | ارزیابی نهایی عملکرد            | بیمار          | ۴۱          | ۱۹                   | ۲۲                   |
|             |            |                                 | برش            | ۵۳۳         |                      |                      |

پس از تقسیم‌بندی صحیح داده‌ها، ادامه‌ی فرآیند شامل مراحل پیش‌پردازش، طراحی معماری، آموزش و ارزیابی مدل است.

### پیش‌پردازش داده‌ها

این مرحله شامل سه گام اصلی است که با استفاده از ابزارهای شناخته‌شده مانند FreeSurfer (فیشال)<sup>۱</sup>، FSL (جنکینسون)<sup>۲</sup> و همکاران، ۲۰۱۲) انجام شد:

۱. حذف جمجمه<sup>۳</sup>: جداسازی بافت مغز از بافت‌های غیرمغزی.
۲. ثبت تصویر<sup>۴</sup>: هم‌تراز کردن تمامی تصاویر در یک فضای مختصات استاندارد.
۳. استخراج برش‌های دوبعدی و نرمال‌سازی: تصاویر سه‌بعدی به مجموعه‌ای از ۱۳ برش دوبعدی از نمای افقی تبدیل و پس از نرمال‌سازی به روش Min-Max، ذخیره شدند.

پس از انجام مراحل پیش‌پردازش، فرایند افزایش داده بر روی مجموعه‌ی آموزش اعمال شد. در این پژوهش، افزایش داده شامل تبدیلات تصادفی الهام‌گرفته از مطالعه‌ی ژانگ و همکاران (۲۰۲۲) — مانند چرخش، تغییر کنتراست و افزودن نویز — بود. هدف از افزایش داده، افزایش تنوع داده‌ها و کاهش خطر بیش‌برازش است. پس از افزایش داده، به عنوان گام انتهایی، نرمال‌سازی نهایی به روش Z-score بر روی تصاویر اعمال شد و مجموعه داده برای آموزش و ارزیابی مدل، آماده گردید. یک نکته بسیار مهم در این فرآیند، محاسبه میانگین و انحراف معیار تنها بر روی داده‌های مجموعه آموزش است. این مقادیر سپس

<sup>1</sup> Fischl

<sup>2</sup> Jenkinson

<sup>3</sup> Skull Stripping

<sup>4</sup> Image Registration

برای نرمال سازی هر سه مجموعه (آموزش، اعتبارسنجی و آزمون) به کار گرفته شدند. این رویکرد از نشت اطلاعات از مجموعه های اعتبارسنجی و آزمون به فرآیند پیش پردازش جلوگیری می کند.

### معماری مدل پیشنهادی

معماری پیشنهادی، یک شبکه عصبی پیچشی عمیق است که از پایه و بدون استفاده از یادگیری انتقالی - به منظور جلوگیری از هرگونه نشت داده - برای طبقه بندی تصاویر SMRI مغز طراحی شده است. این مدل یک معماری هیبریدی است که با هدف بهینه سازی، کارایی محاسباتی برگرفته از شبکه های سبک مانند MobileNet (هاوارد<sup>۱</sup> و همکاران، ۲۰۱۷)، را با قدرت استخراج ویژگی و پایداری آموزش معماری های عمیق تر مانند ResNet (هی<sup>۲</sup> و همکاران، ۲۰۱۶) ترکیب می کند. شبکه از کنار هم قرار گرفتن اجزای زیر تشکیل شده است:

۱. بخش اولیه: شبکه با یک بخش اولیه شامل دو لایه کانولوشن متوالی با ۲ گام<sup>۳</sup> آغاز می شود. این بخش به سرعت ابعاد فضایی تصویر ورودی را از  $256 \times 256$  به  $64 \times 64$  کاهش می دهد. این کاهش ابعاد سریع، میدان دریافتی نورون ها را به سرعت افزایش داده و بار محاسباتی را در لایه های بعدی کاهش می دهد. جزئیات کامل لایه های بخش اولیه در جدول ۳ آمده است.
۲. بدنه اصلی: این بخش از یک پشته چهارتایی از CustomBlock ها تشکیل شده است. بلوک اول ابعاد فضایی را حفظ می کند (تعداد گام برابر یک است)، در حالی که سه بلوک بعدی هر کدام ابعاد فضایی را نصف می کنند (تعداد گام برابر دو است). همزمان با کاهش ابعاد فضایی، عمق (تعداد کانال ها) به تدریج افزایش می یابد.

هر بلوک CustomBlock از سه جزء اصلی شامل استخراج گر ویژگی دومسیره، مکانیزم توجه کانالی و اتصال باقی مانده تشکیل شده است. در مرحله نخست، نقشه ویژگی ورودی به دو مسیر موازی ارسال می شود. مسیر اصلی از کانولوشن های استاندارد برای استخراج ویژگی های محلی، جزئیات ساختاری و الگوهای با وضوح بالا استفاده می کند. در مقابل، مسیر زمینه ای با بهره گیری از بلوک های CustomConv، ویژگی های زمینه ای و اطلاعات کلی تر تصویر را با هزینه محاسباتی کمتر استخراج می کند. در شرایطی که مقدار گام برابر با ۱ باشد، مسیر زمینه ای علاوه بر استخراج ویژگی، شامل یک مرحله نمونه کاهی و سپس نمونه افزایی است. این فرآیند باعث می شود شبکه اطلاعات را در مقیاس فشرده تری پردازش کرده و نمایش های مقاوم تر و کلی تری از ساختار تصویر بیاموزد. خروجی دو مسیر پس از یکسان سازی ابعاد فضایی در امتداد بعد کانال با یکدیگر الحاق می شوند و یک نمایش ترکیبی از ویژگی ها را تشکیل می دهند. سپس این نمایش وارد مکانیزم توجه کانالی می شود تا اهمیت نسبی هر کانال به صورت تطبیقی تعیین گردد. برای افزایش توان تمایز ویژگی ها، از یک مکانیزم توجه کانالی الهام گرفته از معماری Squeeze-and-Excitation

<sup>1</sup> Howard

<sup>2</sup> He

<sup>3</sup> Stride



(هو و همکاران، ۲۰۱۸) استفاده شده است. در این مکانیزم، ابتدا با استفاده از لایه Adaptive Average Pooling اطلاعات فضایی هر کانال در یک مقدار منفرد فشرده شده و یک بردار توصیف‌گر کانالی تولید می‌شود. سپس این بردار از دو لایه کانولوشن  $1 \times 1$  عبور می‌کند. لایه اول ابعاد بردار را کاهش داده و لایه دوم آن را مجدداً به ابعاد اولیه بازمی‌گرداند. خروجی نهایی پس از عبور از تابع فعال‌سازی سیگموید، وزن اهمیت هر کانال را در بازه صفر تا یک تعیین می‌کند. در نهایت، این ضرایب بر روی نقشه‌های ویژگی اعمال می‌شوند تا کانال‌های حاوی اطلاعات مهم تقویت و کانال‌های کم‌اهمیت تضعیف شوند. به منظور تسهیل جریان گرادینان و جلوگیری از مشکلات محوشدگی گرادینان در شبکه‌های عمیق، در هر CustomBlock از اتصال باقی‌مانده استفاده شده است. در صورتی که ابعاد ورودی و خروجی بلوک یکسان باشند، از اتصال همانی استفاده می‌شود و ورودی بدون هیچ تغییری به خروجی بلوک افزوده می‌شود. در حالتی که ابعاد فضایی یا تعداد کانال‌ها تغییر کند، یک مسیر پروجکشن شامل AvgPool2D و کانولوشن  $1 \times 1$  برای تطبیق ابعاد ورودی و خروجی به کار گرفته می‌شود. خروجی مسیر میان‌بر پس از تطبیق ابعاد، به‌صورت عنصری با خروجی مسیر اصلی جمع شده و سپس از تابع فعال‌سازی GELU عبور می‌کند. مشخصات این بلوک در جدول ۴ قابل مشاهده است.

۳. سر طبقه‌بندی‌کننده: سر طبقه‌بندی‌کننده شامل یک لایه Linear با BatchNorm1d، GELU و Dropout با نرخ ۰.۳ و یک لایه Linear نهایی است. مجموع پارامترهای قابل آموزش مدل ۳,۵۴۲,۸۴۹ است. معماری کلی مدل در شکل ۱ و معماری CustomBlock در شکل ۲ قابل مشاهده است.

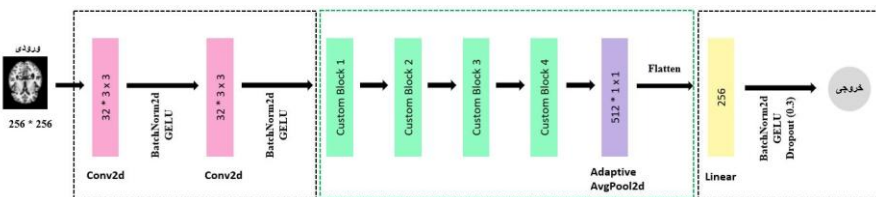
**جدول ۳. مشخصات بخش اولیه.** این بخش شامل دو لایه کانولوشن متوالی است که ابعاد تصویر ورودی  $256 \times 256$  را به  $64 \times 64$  کاهش می‌دهد. کرنل: اندازه فیلتر؛ گام: میزان جابجایی فیلتر در هر مرحله؛ لایه‌گذاری: BatchNorm2d، نرمال‌سازی: دست‌های: GELU. تابع فعال‌سازی.

| لایه          | کانال ورودی | کانال خروجی | کرنل         | گام | لایه‌گذاری | نرمال‌سازی /<br>فعال‌سازی |
|---------------|-------------|-------------|--------------|-----|------------|---------------------------|
| Conv2d<br>اول | ۱           | ۳۲          | $3 \times 3$ | ۲   | ۱          | BatchNorm2d +<br>GELU     |
| Conv2d<br>دوم | ۳۲          | ۶۴          | $3 \times 3$ | ۲   | ۱          | BatchNorm2d +<br>GELU     |

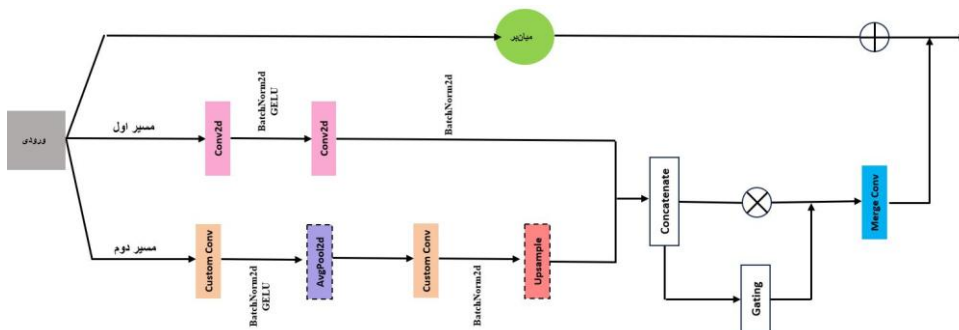
**جدول ۴. مشخصات چهار بلوک CustomBlock در بدنه اصلی شبکه.** هر بلوک ابعاد فضایی تصویر را کاهش و تعداد کانال‌ها را افزایش می‌دهد. گام: میزان جابجایی؛ AvgPool: میانگین‌گیری؛ Upsample: بزرگ‌نمایی؛ GELU: تابع فعال‌سازی.

| بلوک | کانال ورودی | کانال<br>خروجی | گام | لایه‌گذاری/نمونه‌افزایی                | ابعاد فضایی<br>خروجی | تابع فعال‌سازی |
|------|-------------|----------------|-----|----------------------------------------|----------------------|----------------|
| ۱    | ۶۴          | ۱۲۸            | ۱   | AvgPool(2) +<br>Upsample( $\times 2$ ) | $64 \times 64$       | GELU           |
| ۲    | ۱۲۸         | ۲۵۶            | ۲   | —                                      | $32 \times 32$       | GELU           |
| ۳    | ۲۵۶         | ۳۸۴            | ۲   | —                                      | $16 \times 16$       | GELU           |

|   |     |     |   |   |     |      |
|---|-----|-----|---|---|-----|------|
| ۴ | ۳۸۴ | ۵۱۲ | ۲ | — | ۸×۸ | GELU |
|---|-----|-----|---|---|-----|------|



**شکل ۱. معماری کلی مدل پیشنهادی.** این شکل ساختار کلی شبکه عصبی پیچشی را نشان می‌دهد. ورودی تصویر SMRI با ابعاد  $256 \times 256$  پیکسل (یک کانال خاکستری) از سمت چپ وارد شبکه می‌شود. بخش اولیه شامل دو لایه کانولوشن متوالی که ابعاد تصویر را از  $256 \times 256$  به  $64 \times 64$  کاهش می‌دهند. بلوک‌های CustomBlock (رنگ سبز)، چهار بلوک پردازشی پیشرفته که ویژگی‌های تصویر را به تدریج استخراج می‌کنند و ابعاد فضایی را کاهش می‌دهند. سر طبقه‌بندی‌کننده دو لایه کاملاً متصل که خروجی نهایی را به یک مقدار احتمال برای تشخیص AD یا CN تبدیل می‌کند.



**شکل ۲. معماری بلوک CustomBlock.** این شکل ساختار داخلی هر بلوک CustomBlock را نشان می‌دهد. هر بلوک از سه جزء اصلی تشکیل شده است: (۱) استخراج گر ویژگی دومسیره: ورودی به دو مسیر موازی تقسیم می‌شود؛ مسیر اصلی از کانولوشن‌های استاندارد برای ویژگی‌های محلی و مسیر دوم از بلوک‌های CustomConv برای اطلاعات کلی‌تر استفاده می‌کند. (۲) مکانیزم توجه کانالی (رنگ بنفش): الهام‌گرفته از معماری Squeeze-and-Excitation اهمیت نسبی هر کانال را به صورت تطبیقی تعیین کرده و کانال‌های مهم را تقویت می‌کند. (۳) اتصال باقی‌مانده (رنگ سبز): ورودی بلوک مستقیماً به خروجی آن اضافه می‌شود تا از محوشدگی گرادینان جلوگیری شود.

اطلاعات مرتبط با فرآیندهای آموزش در جدول ۵ قابل مشاهده است. از بهینه‌ساز AdamW (لوشچیلوف و هاتر<sup>۱</sup>، ۲۰۱۷) و زمان‌بند نرخ یادگیری CosineAnnealingLR (لوشچیلوف و هاتر، ۲۰۱۶) و تابع زیان BCEWithLogitsLoss استفاده شده است.

**جدول ۵. خلاصه فرآیندهای آموزش**

| مقدار | فرآیندها      |
|-------|---------------|
| ۷۵    | تعداد دوره ها |
| ۱۶    | اندازه دسته   |

<sup>1</sup> Loshchilov & Hutter

|                   |         |
|-------------------|---------|
| نرخ یادگیری اولیه | ۰.۰۰۰۱  |
| تنظیم وزن         | ۰.۰۰۰۱  |
| طول چرخه زمان‌بند | ۱۰ دوره |
| حداقل نرخ یادگیری | 1e-6    |

### اطلاعات تکمیلی پیاده‌سازی

در طول آموزش، عملکرد مدل پس از هر دوره بر روی یک مجموعه داده اعتبارسنجی ارزیابی شد. این ارزیابی به‌صورت برش-محور انجام گرفت. هدف از این مرحله، نظارت بر یادگیری مدل، تشخیص بیش‌برازش و انتخاب بهترین نقطه توقف بود. در این سطح، هر برش تصویر به عنوان یک نمونه مستقل در نظر گرفته شده و معیارهای عملکرد بر اساس پیش‌بینی‌های تک‌تک برش‌ها محاسبه شد.

همچنین آزمون نهایی و اصلی مدل بر روی مجموعه آزمون انجام شد که مدل هیچ‌گاه در طول آموزش آن را مشاهده نکرده است. این ارزیابی به‌صورت بیمار-محور صورت گرفت که یک رویکرد بالینی معتبرتر است:

۱. تمام اسلایس‌های متعلق به یک بیمار به مدل داده شد.
  ۲. احتمال خروجی مدل (پس از اعمال تابع سیگموید) برای تمام این اسلایس‌ها محاسبه شد.
  ۳. میانگین این احتمالات به عنوان احتمال نهایی برای آن بیمار در نظر گرفته شد.
  ۴. برچسب نهایی بیمار بر اساس این احتمال میانگین (با آستانه ۰.۵) تعیین شد.
- این روش بسیار قوی‌تر از تصمیم‌گیری بر اساس یک برش تنهاست، زیرا نویز یا ویژگی‌های غیرمعمول در یک برش را تعدیل کرده و بر اساس اطلاعات کلی از یک ناحیه مهم مغز تصمیم‌گیری می‌کند. برای اطمینان از اینکه نتایج این پژوهش کاملاً قابل تکرار هستند، یک مقدار اولیه ثابت برای تمام کتابخانه‌هایی که از مولدهای اعداد تصادفی استفاده می‌کنند تنظیم شد.
- بخش پیش‌پردازش در سیستم شخصی انجام شد و بخش اجرای کد اصلی در محیط رایگان Google Colab با مشخصات سخت‌افزاری و نرم‌افزاری زیر پیاده‌سازی و اجرا شد.

سخت‌افزار:

- پردازنده مرکزی (CPU): Intel(R) Xeon(R) CPU @ 2.30GHz
- پردازنده گرافیکی (GPU): Tesla T4
- حافظه اصلی (RAM): 13289424 kB

نرم‌افزار و کتابخانه‌ها:

- سیستم عامل: Linux posix
- زبان برنامه‌نویسی: Python 3

- NumPy: برای محاسبات عددی.
- Scikit-learn: برای تقسیم‌بندی داده‌ها و محاسبه معیارهای ارزیابی.
- Pillow: برای پردازش اولیه تصاویر.
- Torchvision: برای تبدیلات استاندارد تصویر.
- Matplotlib: برای مصورسازی نتایج.
- tqdm: برای نمایش نوار پیشرفت.

### تجزیه و تحلیل یافته‌ها

مدل نهایی، پس از اتمام فرآیند آموزش، بر روی مجموعه آزمون که شامل ۴۱ بیمار کاملاً جدید است، ارزیابی شد. عملکرد کلی مدل پیشنهادی بسیار بالا است و مدل توانست با موفقیت بیماران مبتلا به آلزایمر را از افراد سالم تشخیص دهد. نتایج کمی در جدول ۶ خلاصه شده است. مدل به دقت کلی ۹۵.۱۲٪ دست یافت که نشان می‌دهد از هر ۱۰۰ فرد، به طور متوسط ۹۵ نفر را به درستی طبقه‌بندی کرده است. مقدار AUC نیز برابر با ۰.۹۴ است که قدرت تمایز بالای مدل را در جدا کردن دو گروه AD و CN تأیید می‌کند.

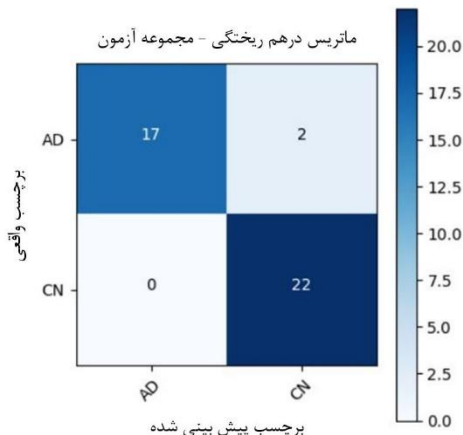
جدول ۶. نتایج عملکرد نهایی مدل بر روی مجموعه آزمون

| مقدار  | تعریف                                                                                    | نماد | معیار ارزیابی                                  |
|--------|------------------------------------------------------------------------------------------|------|------------------------------------------------|
| ۹۵.۱۲٪ | نسبت کل پیش‌بینی‌های صحیح                                                                | ACC  | دقت                                            |
| ۸۹.۴۷٪ | چه درصدی از بیماران واقعی به درستی تشخیص داده شده‌اند.                                   | SEN  | حساسیت                                         |
| ۱۰۰٪   | از میان تمام مواردی که مدل به عنوان بیمار پیش‌بینی کرده، چه درصدی واقعاً بیمار بوده‌اند. | PRE  | دقت مثبت                                       |
| ۹۴.۴۴٪ | میانگین هماهنگ حساسیت و دقت                                                              | F1   | امتیاز F1                                      |
| ۰.۹۴   | توانایی تمایز کلی مدل                                                                    | AUC  | سطح زیر منحنی مشخصه عملکرد گیرنده <sup>۱</sup> |
| ۰.۲۹   | اختلاف بین مقدار واقعی و پیش‌بینی مدل                                                    | Loss | زیان                                           |

ماتریس درهم‌ریختگی در شکل ۳ نیز تحلیل دقیق‌تری از خطاها ارائه می‌دهد:

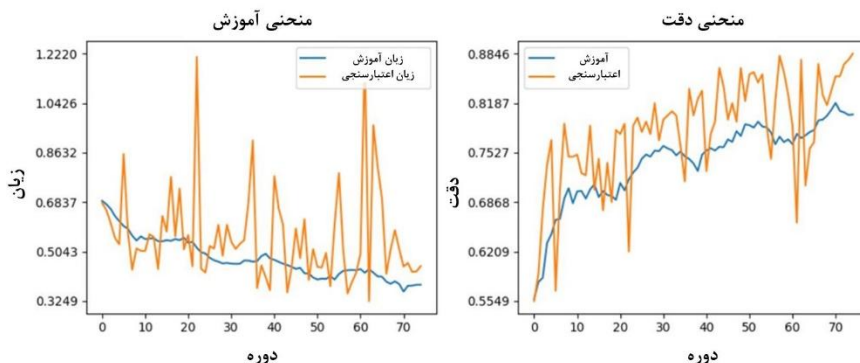
- مثبت واقعی: مدل ۱۷ بیمار مبتلا به AD را به درستی به عنوان AD طبقه‌بندی کرده است.
- منفی کاذب: مدل ۲ بیمار مبتلا به AD را به اشتباه به عنوان فرد سالم طبقه‌بندی کرده است.
- منفی واقعی: مدل ۲۲ فرد سالم را به درستی به عنوان CN طبقه‌بندی کرده است.
- مثبت کاذب: مدل هیچ فرد سالمی را به اشتباه به عنوان بیمار طبقه‌بندی نکرده است.

<sup>1</sup> Receiver Operating Characteristic Curve (ROC)



شکل ۳. ماتریس درهم ریختگی برای مجموعه آزمون. هر خانه تعداد بیمارانی را نشان می‌دهد که در آن دسته قرار گرفته‌اند. ردیف‌ها نشان‌دهنده تشخیص واقعی و ستون‌ها نشان‌دهنده پیش‌بینی مدل هستند. رنگ تیره‌تر بیانگر تعداد بیشتر است.

نمودارهای دقت و زیان حین آموزش و اعتبارسنجی نیز در شکل ۴ قابل مشاهده است. مطابق نمودار، دقت آموزش به طور پیوسته افزایش یافته که نشان‌دهنده ظرفیت کافی مدل برای یادگیری داده‌ها است. دقت اعتبارسنجی نیز با وجود نوسانات، یک روند کلی صعودی را طی کرده است. همگرایی هر دو منحنی زیان و دقت، نشان می‌دهد که فرآیند آموزش با موفقیت به پایان رسیده و مدل نهایی انتخاب‌شده نماینده یک نقطه بهینه از عملکرد مدل است.



شکل ۴. منحنی‌های زیان (چپ) و دقت (راست) آموزش و اعتبارسنجی. هر دوره یک بار کامل عبور مدل از تمام داده‌های آموزشی است. نمودار سمت چپ (منحنی زیان): کاهش تدریجی خطای مدل در طول آموزش را نشان می‌دهد؛ مقادیر پایین‌تر نشان‌دهنده عملکرد بهتر است. نمودار سمت راست (منحنی دقت): درصد نمونه‌هایی که مدل به درستی طبقه‌بندی کرده نشان می‌دهد. خط آبی: عملکرد بر روی مجموعه آموزش. خط نارنجی: عملکرد بر روی مجموعه اعتبارسنجی. همگرایی دو منحنی نشان‌دهنده آموزش موفق بدون بیش‌برازش است.

## نتیجه‌گیری

این پژوهش با هدف توسعه‌ی یک مدل یادگیری عمیق معتبر و کارآمد برای تشخیص بیماری آلزایمر انجام شد. مدل پیشنهادی با دستیابی به دقت ۹۵.۱۲٪ و AUC برابر با ۰.۹۴، نشان‌دهنده عملکرد قابل توجهی در مقایسه با بسیاری از رویکردهای موجود در ادبیات است. با این حال، هدف اصلی این مطالعه نه صرفاً دستیابی به بیشترین دقت، بلکه ارائه یک چارچوب روش‌مند و قابل اعتماد مبتنی بر جلوگیری از نشت داده و عدم استفاده از پیش‌آموزش خارجی بوده است.

در مقایسه با مطالعات پیشین، بسیاری از رویکردهای مبتنی بر CNN دوبعدی مانند اسلام و ژانگ (۲۰۱۸)، هو و همکاران (۲۰۲۱)، ساراتچاگا و همکاران (۲۰۲۱) و ژانگ و همکاران (۲۰۲۲) معمولاً دقت‌هایی در حدود ۹۳٪ گزارش کرده‌اند. برخی مطالعات پیشرفته‌تر مانند گوئل و همکاران (۲۰۲۴) و مالو و همکاران (۲۰۲۴) با بهره‌گیری از معماری‌های ترکیبی یا حافظه‌دار، به دقت‌های بالاتری در حدود ۹۷٪ تا ۹۸٪ دست یافته‌اند. در این میان، عملکرد مدل پیشنهادی در محدوده رقابتی این روش‌ها قرار می‌گیرد، با این تفاوت که ارزیابی آن تحت یک پروتکل سخت‌گیرانه جلوگیری از نشت داده انجام شده است.

در مقایسه با رویکردهای سه‌بعدی نیز، اگرچه برخی مطالعات مانند نتامپاکیس و همکاران (۲۰۲۴)، کورولف و همکاران (۲۰۱۷) و دیناگار و همکاران (۲۰۲۳) از اطلاعات حجمی کامل MRI استفاده کرده‌اند، مدل پیشنهادی با استفاده از داده‌های دوبعدی توانسته است عملکردی قابل مقایسه ارائه دهد، در حالی که از نظر محاسباتی بسیار سبک‌تر و مناسب‌تر برای کاربردهای عملی است. این موضوع نشان می‌دهد که طراحی معماری مناسب می‌تواند تا حدی وابستگی به داده‌های سه‌بعدی و منابع محاسباتی سنگین را کاهش دهد.

هرچند برخی مطالعات مبتنی بر یادگیری انتقالی یا پیش‌آموزش‌شده، گزارش عملکردهای بالاتری داشته‌اند، در این پژوهش به‌طور آگاهانه از استفاده از مدل‌های از پیش آموزش‌دیده خودداری شده است تا از هرگونه سوگیری ناشی از انتقال دانش بین دامنه‌ای جلوگیری شود. بنابراین، نتایج به‌دست‌آمده را می‌توان در چارچوب یک ارزیابی محافظه‌کارانه و از نظر روش‌شناختی قابل اعتماد تفسیر کرد.

یکی از نتایج مهم این مطالعه، دستیابی مدل به دقت مثبت بالا است که از دیدگاه کاربردهای بالینی اهمیت ویژه‌ای دارد. این ویژگی نشان می‌دهد که مدل در تشخیص موارد مثبت عملکرد پایداری داشته است، اگرچه برای تأیید نهایی قابلیت تعمیم آن، انجام ارزیابی بر روی داده‌های مستقل و چندمرکزی ضروری است.

در مجموع، این پژوهش نشان می‌دهد که یک معماری CNN دوبعدی طراحی‌شده به‌صورت اختصاصی، همراه با رعایت دقیق اصول جلوگیری از نشت داده، می‌تواند عملکردی رقابتی با بسیاری از روش‌های موجود در ادبیات ارائه دهد. با این حال، برای دستیابی به یک نتیجه‌گیری قطعی درباره برتری نسبت به سایر روش‌ها، انجام مطالعات گسترده‌تر و ارزیابی روی داده‌های متنوع‌تر پیشنهاد می‌شود.

با این حال، این پژوهش با محدودیت‌هایی نیز مواجه بود. نخست، مدل تنها بر روی مجموعه داده ADNI ارزیابی شده و برای تأیید قابلیت تعمیم‌پذیری، نیازمند سنجش عملکرد آن بر روی داده‌های خارجی و

مستقل است؛ البته به شرط آنکه اطمینان حاصل شود نشت داده‌ای رخ نخواهد داد. دوم، این مطالعه بر طبقه‌بندی دوتایی تمرکز داشته، در حالی که یک چالش بالینی مهم دیگر، تمایز میان اختلال شناختی خفیف و پیش‌بینی تبدیل آن به آلزایمر است. از دیگر محدودیت‌های این مطالعه، عدم تحلیل جداگانه عملکرد مدل بر اساس جنسیت است. کلینگنبرگ<sup>۱</sup> و همکاران (۲۰۲۳) نشان دادند که مدل‌های CNN مبتنی بر SMRI از مجموعه داده ADNI برای زنان عملکرد بهتری نسبت به مردان دارند، حتی در شرایطی که داده‌های آموزش از نظر جنسیت متعادل باشند. با توجه به اینکه مجموعه داده مورد استفاده در پژوهش حاضر نیز شامل ۱۷۴ زن و ۱۶۹ مرد است، بررسی احتمال وجود چنین سوگیری‌ای در مدل پیشنهادی به عنوان کار آینده توصیه می‌شود.

برای پژوهش‌های آتی، پیشنهاد می‌شود مدل بر روی طبقه‌بندی چندکلاسه گسترش یابد. همچنین، ادغام داده‌های چندوجهی مانند تصاویر PET، اطلاعات ژنتیکی و بالینی می‌تواند به یک سیستم تشخیصی جامع‌تر و دقیق‌تر منجر شود. استفاده از تکنیک‌های تفسیرپذیری مانند Grad-CAM (سلواراجو<sup>۲</sup> و همکاران، ۲۰۱۷) و هوش مصنوعی توضیحی<sup>۳</sup> (آریتا<sup>۴</sup> و همکاران، ۲۰۲۰) نیز می‌تواند به افزایش اعتماد بالینی به مدل کمک کند.

در نهایت، این پژوهش با ارائه یک معماری نوین و یک چارچوب روش‌شناختی معتبر که از نشت داده جلوگیری می‌کند، نشان داد که می‌توان به ابزارهای تشخیصی هوشمند، کارآمد و قابل اعتماد برای بیماری آلزایمر دست یافت. این تحقیق گامی مهم در جهت پر کردن شکاف موجود در ادبیات و پیشبرد دانش در زمینه طراحی شبکه‌های عصبی برای کاربردهای خاص تصویربرداری پزشکی است.

---

<sup>1</sup> Klingenberg

<sup>2</sup> Selvaraju

<sup>3</sup> Explainable AI

<sup>4</sup> Arrieta

## منابع

- Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>
- Arya, A. D., Verma, S. S., Chakarabarti, P., Chakrabarti, T., Elngar, A. A., Kamali, A.-M., & Nami, M. (2023). A systematic review on machine learning and deep learning techniques in the effective diagnosis of Alzheimer's disease. *Brain Informatics*, 10(1), 17. <https://doi.org/10.1186/s40708-023-00195-7>
- Atnafu, S. W., & Diciotti, S. (2024). Development of an interpretable deep learning system for the identification of patients with Alzheimer's disease. In *Green Energy and Technology* (pp. 27–37). Springer Nature Switzerland.
- Banerjee, D., Muralidharan, A., Hakim Mohammed, A. R., & Malik, B. H. (2020). Neuroimaging in dementia: A brief review. *Cureus*, 12(6), e8682. <https://doi.org/10.7759/cureus.8682>
- BrightFocus Foundation. (2025). *Alzheimer's disease: Facts & figures*. BrightFocus Foundation. <https://www.brightfocus.org/alzheimers/facts-figures/>
- Carcagnì, P., Leo, M., Del Coco, M., Distant, C., & De Salve, A. (2023). Convolution neural networks and self-attention learners for Alzheimer dementia diagnosis from brain MRI. *Sensors (Basel, Switzerland)*, 23(3). <https://doi.org/10.3390/s23031694>
- Dhinagar, N. J., Thomopoulos, S. I., Laltoo, E., & Thompson, P. M. (2023). Efficiently training vision transformers on structural MRI scans for Alzheimer's disease detection. In *arXiv [eess.IV]*. <https://doi.org/10.48550/ARXIV.2303.08216>
- Fischl, B. (2012). FreeSurfer. *NeuroImage*, 62(2), 774–781. <https://doi.org/10.1016/j.neuroimage.2012.01.021>
- Goel, S., Sharma, N., & Zheng, Z. (2024). *Diagnosis of Alzheimer's disease using 3D and 2D convolutional neural networks*. Retrieved October 2, 2025, from <https://cs231n.stanford.edu/2024/papers/diagnosis-of-alzheimers-disease-using-3d-and-2d-convolutional-ne.pdf>
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770–778. <https://doi.org/10.1109/CVPR.2016.90>



- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv*. <https://arxiv.org/abs/1704.04861>
- Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 7132–7141. <https://doi.org/10.1109/CVPR.2018.00745>
- Hu, Z., Wu, Q., Chen, C., Xiao, L., & Jin, S. (2021). Alzheimer's disease diagnosis method based on convolutional neural network using key slices voting. *2021 11th International Conference on Information Science and Technology (ICIST)*.
- Islam, J., & Zhang, Y. (2018). Brain MRI analysis for Alzheimer's disease diagnosis using an ensemble system of deep convolutional neural networks. *Brain Informatics*, 5(2), 2. <https://doi.org/10.1186/s40708-018-0080-3>
- Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, 62(2), 782–790. <https://doi.org/10.1016/j.neuroimage.2011.09.015>
- Klingenberg, M., Stark, D., Eitel, F., Budding, C., Habes, M., Ritter, K., & Alzheimer's Disease Neuroimaging Initiative. (2023). Higher performance for women than men in MRI-based Alzheimer's disease detection. *Alzheimer's Research & Therapy*, 15, 84. <https://doi.org/10.1186/s13195-023-01225-6>
- Korolev, S., Safiullin, A., Belyaev, M., & Dodonova, Y. (2017). Residual and plain convolutional neural networks for 3D brain MRI classification. In *arXiv [cs.CV]*. <http://arxiv.org/abs/1701.06643>
- Loshchilov, I., & Hutter, F. (2016). SGDR: Stochastic gradient descent with warm restarts. *arXiv*. <http://arxiv.org/abs/1608.03983>
- Loshchilov, I., & Hutter, F. (2017). Decoupled weight decay regularization. *arXiv*. <http://arxiv.org/abs/1711.05101>
- Malu, G., Uday, N., Sherly, E., Abraham, A., & Kuber Bodhey, N. (2024). CirMNet: A shape-based hybrid feature extraction technique using CNN and CMSMD for Alzheimer's MRI classification. *IEEE Access: Practical Innovations, Open Solutions*, 12, 80491–80504. <https://doi.org/10.1109/access.2024.3408311>
- Ntampakis, N., Diamantaras, K., Chouvarda, I., Argyriou, V., & Sarigiannidis, P. (2024). Enhanced deep learning methodologies and MRI selection techniques for dementia diagnosis in the elderly population. In *arXiv [eess.IV]*. <http://arxiv.org/abs/2407.17324>

- Rashid, A. H., Gupta, A., Gupta, J., & Tanveer, M. (2023). Biceph-net: A robust and lightweight framework for the diagnosis of Alzheimer's disease using 2D-MRI scans and deep similarity learning. *IEEE Journal of Biomedical and Health Informatics*, 27(3), 1205–1213. <https://doi.org/10.1109/JBHI.2022.3174033>
- Saratxaga, C. L., Moya, I., Picón, A., Acosta, M., Moreno-Fernandez-de-Leceta, A., Garrote, E., & Bereciartua-Perez, A. (2021). MRI deep learning-based solution for Alzheimer's disease prediction. *Journal of Personalized Medicine*, 11(9), 902. <https://doi.org/10.3390/jpm11090902>
- Scheltens, P., De Strooper, B., Kivipelto, M., Holstege, H., Chételat, G., Teunissen, C. E., Cummings, J., & van der Flier, W. M. (2021). Alzheimer's disease. *Lancet*, 397(10284), 1577–1590. [https://doi.org/10.1016/s0140-6736\(20\)32205-4](https://doi.org/10.1016/s0140-6736(20)32205-4)
- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 618–626. <https://doi.org/10.1109/ICCV.2017.74>
- Valliani, A., & Soni, A. (2017). Deep residual nets for improved Alzheimer's diagnosis. *Proceedings of the 8<sup>th</sup> ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*.
- Wen, J., Thibeu-Sutre, E., Diaz-Melo, M., Samper-González, J., Routier, A., Bottani, S., Dormont, D., Durrleman, S., Burgos, N., Colliot, O., Alzheimer's Disease Neuroimaging Initiative, & Australian Imaging Biomarkers and Lifestyle flagship study of ageing. (2020). Convolutional neural networks for classification of Alzheimer's disease: Overview and reproducible evaluation. *Medical Image Analysis*, 63(101694), 101694. <https://doi.org/10.1016/j.media.2020.101694>
- World Health Organization. (2025, March 31). *Dementia*. <https://www.who.int/news-room/fact-sheets/detail/dementia/>
- Yagis, E., Atnafu, S. W., García Seco de Herrera, A., Marzi, C., Scheda, R., Giannelli, M., Tessa, C., Citi, L., & Diciotti, S. (2021). Effect of data leakage in brain MRI classification using 2D convolutional neural networks. *Scientific Reports*, 11(1), 22544. <https://doi.org/10.1038/s41598-021-01681-w>
- Young, V. M., Gates, S., Garcia, L. Y., & Salardini, A. (2025). Data leakage in deep learning for Alzheimer's disease diagnosis: A scoping review of methodological rigor and performance inflation. *Diagnostics*, 15(18), 2348. <https://doi.org/10.3390/diagnostics15182348>
- Zhang, F., Pan, B., Shao, P., Liu, P., Alzheimer's Disease Neuroimaging Initiative, Australian Imaging Biomarkers Lifestyle flagship study of ageing, Shen, S., Yao,

P., & Xu, R. X. (2022). A single model deep learning approach for Alzheimer's disease diagnosis. *Neuroscience*, 491, 200–214.  
<https://doi.org/10.1016/j.neuroscience.2022.03.026>

## Introducing a Deep Learning-Based Model for Classification of Alzheimer's Disease

**Zahra Aghaziarati**

*M.Sc. Student, Department of Information Technology Engineering, Faculty of Industrial Engineering, K. N. Toosi University of Technology, Tehran, Iran*

**Monireh Hosseini<sup>1\*</sup>**

*Associate Prof., Department of Information Technology Engineering, Faculty of Industrial Engineering, K.N. Toosi University of Technology, Tehran, Iran*

### Abstract

Alzheimer's disease, the most common cause of dementia, is a growing global health concern. Accurate and timely diagnosis of this disease, especially in the early stages, plays an important role in managing symptoms and improving the quality of life of patients. In recent years, machine learning algorithms, deep learning, and in particular convolutional neural networks, have shown great potential for automated analysis of neuroimaging with high accuracy. However, the validity of many previous studies has been questioned due to data leakage errors. Aiming to overcome the challenge of data leakage, this study presents a valid and efficient deep learning framework for distinguishing patients with Alzheimer's disease from control subjects using structural magnetic resonance images from ADNI dataset. The main innovation of this study is the design of a dedicated architecture. Instead of using pre-trained models, this architecture is designed from the ground up, using depthwise separable convolutions to strike an optimal balance between computational efficiency and feature extraction power. To validate the results, this research has avoided any data leakage. Evaluation of the final model on the test dataset showed excellent performance, achieving 95.12% accuracy and an area under the ROC curve of 0.94. These results indicate that a carefully designed architecture and a valid methodology can become a robust and reliable tool to aid in the early diagnosis of Alzheimer's disease.

**Keywords:** Alzheimer's Disease, Classification, Deep Learning, Convolutional Neural Networks, Magnetic Resonance Imaging.

---

<sup>1</sup> Corresponding Author: hosseini@kntu.ac.ir