

رده‌بندی تصور گفتار مبتنی بر سیگنال‌های مغزی با افزودن داده‌ها و شبکه تطبیق دامنه هم‌اوردانه با استفاده

از یادگیری عمیق در کاربرد رابط مغز-رایانه

مرضیه انجرائی^۱، میر محسن پدرام^{۲*}، میترا میرزا رضایی^۳

^۱دانشجوی دکتری، گروه مهندسی کامپیوتر، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

^۲دانشیار، گروه مهندسی کامپیوتر و برق، دانشگاه خوارزمی، تهران، ایران

^۳دانشیار، گروه علمی مهندسی کامپیوتر، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

مقاله پژوهشی

چکیده

رابط مغز-رایانه مبتنی بر گفتار راهبردهای ارتباط صوتی مؤثری را برای کنترل دستگاه‌ها از طریق دستورات گفتاری که از سیگنال‌های مغزی تفسیر می‌شوند، ارائه می‌کنند. یکی از چالش‌های مهم در مسئله رابط مغز-رایانه رده‌بندی سیگنال‌های مغزی مبتنی بر الکتروانسفالوگرافی است. الکتروانسفالوگرافی یک سیگنال مغزی غیرتاهجمی است که از سطح پوست سر از طریق الکترودها ضبط می‌شود. سیگنال‌های به‌دست آمده با استفاده از تجهیزات آسان و ارزان دارای وضوح مکانی نسبتاً پایین و وضوح زمانی بالا هستند که برای دستیابی به نتایج بهینه مناسب‌ترین روش استخراج و رده‌بندی ویژگی‌ها باید استفاده شود. همچنین جمع‌آوری داده‌های کافی برای آزمودنی جدید زمان و تلاش زیادی می‌طلبد که در این مقاله افزودن داده‌ها با استفاده از مدل مولد هم‌اوردانه به‌منظور بهبود عملکرد رده‌بندی سیگنال‌های مغزی پیشنهاد شده است. همچنین مدلی برای رده‌بندی تصور گفتار مبتنی بر سیگنال‌های مغزی آزمودنی جدید با یادگیری انتقالی با استفاده از روش مولد هم‌اوردانه مبتنی بر تطبیق دامنه متمایزکننده هم‌اوردانه ارائه شده است. به‌منظور شناسایی تصور گفتار از پایگاه داده KaraOne استفاده شده است. روش پیشنهادی با سایر روش‌های جدید بر اساس معیارهای صحت و کاپا مورد ارزیابی قرار گرفتند. طبق نتایج به‌دست آمده روش پیشنهادی با صحت ۸۶٪ و ۶۰/۲۱٪ به ترتیب تصور کلمات و واج‌ها را رده‌بندی می‌کند. مدل پیشنهادی، مستقل از سیگنال‌های مغزی هر فرد است که با آموزش مدل بر روی سیگنال‌های مغزی افزوده شده آزمودنی‌ها در دامنه منبع می‌توان به‌طور مؤثر سیگنال‌ها را در دامنه هدف بدون نیاز به داده‌های برچسب‌گذاری شده از آزمودنی جدید رده‌بندی کرد.

تاریخ دریافت:

۱۴۰۳/۱۲/۱۰

تاریخ پذیرش:

۱۴۰۴/۱/۳۱

کلیدواژه‌ها:

تصور گفتار، افزودن داده، تطبیق دامنه متمایزکننده هم‌اوردانه، یادگیری عمیق، یادگیری انتقالی، شبکه عصبی پیچشی، رابط مغز-رایانه.

نویسنده مسئول:

pedram@khu.ac.ir

doi: 10.22034/ABMIR.2025.22852.1105

E-ISSN: [2821-2037](https://doi.org/10.22034/ABMIR.2025.22852.1105) /© 2024. Published by Yazd University This is an open access article

under the CC BY 4.0 License (<https://creativecommons.org/licenses/by/4.0/>).



۱- مقدمه

سامانه‌های رابط مغز-رایانه^۱ یک سامانه ارتباطی است که به هیچ فعالیت عضلانی نیاز ندارد. درواقع سامانه رابط مغز-رایانه، فرد را برای ارسال دستورات به یک دستگاه الکترونیکی تنها با استفاده از فعالیت مغزی قادر می‌سازد. سامانه‌های رابط مغز-رایانه در ابتدا برای کمک به افراد ناتوان در برقراری ارتباط با محیط طراحی شده بودند. چنین دستگاه‌هایی برای افراد سالم نیز کاربردهایی دارند؛ مانند سرگرمی، درمان‌های پیشگیرانه، ارتباطات شخصی، بازی‌ها و غیره. گفتار پرکاربردترین شکل ارتباطی در انسان است. بااین‌حال، افرادی که از اختلال حرکتی عضلانی رنج می‌برند (بیماران قفل‌شده) نمی‌توانند از این ارتباط طبیعی استفاده کنند. سامانه‌های رابط مغز-رایانه مبتنی بر تصور گفتار^۲، رویکردهای ارتباطی کارآمدی را برای نظارت بر محیط اطراف از طریق دستورالعمل‌های تفسیر شده از داده‌های مغزی مربوط به تصور گفتار ارائه می‌کنند که می‌تواند زندگی افراد را بهبود ببخشد.

اولین قدم در سامانه‌های رابط مغز-رایانه ثبت سیگنال‌های مغزی است. روش‌های مختلفی برای ثبت و ذخیره سیگنال‌های مغزی وجود دارند از جمله الکتروانسفالوگرافی^۳، مگنتوآنسفالوگرافی، الکتروکورتیکوگرافی و تصویربرداری تشدید مغناطیسی عملکردی. الکتروانسفالوگرافی یکی از روش‌های پرکاربرد است که به دلیل غیرتهاجمی، قابل حمل و نسبتاً ارزان بودن آن برای کاربردهای عملی رابط مغز-رایانه مناسب است. به‌منظور بهبود کارایی سامانه رابط مغز-رایانه لازم است رمزگشایی مناسبی از سیگنال الکتروانسفالوگرافی انجام شود؛ بنابراین استخراج ویژگی و رده‌بندی^۴ دو چالش اصلی در سامانه‌های رابط مغز-رایانه مبتنی بر سیگنال‌های الکتروانسفالوگرافی است. امروزه الگوریتم‌های یادگیری عمیق با موفقیت در زمینه رمزگشایی تصور گفتار پیاده‌سازی شده‌اند که در مقایسه با روش‌های یادگیری ماشین سنتی، کارایی و سادگی بالاتری را ایجاد می‌کنند [۱].

رمزگشایی تصورات انسان مبتنی بر سیگنال‌های الکتروانسفالوگرافی در چند سال اخیر مورد توجه زیادی قرار گرفته

که منجر به گسترش طراحی سامانه‌های رابط مغز-رایانه مبتنی بر تصور گفتار شده است. تمرکز اصلی تحقیقات در زمینه رده‌بندی تصور گفتار روی کوچک‌ترین واحدهای زبان مانند هجاها، واج‌ها، صامت، مصوت و کلمات محدود بوده است. در سال ۲۰۰۹ یک سامانه رابط مغز-رایانه مبتنی بر تصور مصوت‌ها (/a/ و /u/) ارائه شد که از الگوی مشترک فضایی و ماشین بردار پشتیبان غیرخطی به ترتیب به‌عنوان بردار ویژگی و رده‌بند استفاده شد که صحت^۵ رده‌بندی در محدوده ۶۸٪ تا ۷۸٪ برای سه آزمودنی^۶ بود [۲]. در سال ۲۰۱۰ الگوریتمی مبتنی بر ضرایب خودرگرسیون به‌عنوان ویژگی‌ها و الگوریتم k-نزدیک‌ترین همسایه برای رده‌بندی تصور دو هجای /ba/ و /ku/ ارائه شد که صحتی در حدود ۶۱٪ گزارش کردند [۳]. همچنین در [۴] برای دو هجای /ba/ و /ku/ از روش شناسایی کور مرتبه دوم برای حذف مصنوعات و کاهش ابعاد داده‌های خام استفاده شد. سپس، تبدیل هیلبرت - هوانگ اعمال و طیف هیلبرت برای به‌دست‌آوردن ویژگی‌ها نرمال شد. از یک رده‌بند بیزین مبتنی بر تحلیل تفکیک خطی چند رده‌ای استفاده شد که صحت در محدوده بین ۱۶/۸۰٪ و ۳۰/۸۲٪ گزارش کردند.

در [۵] تصور پنج نوع واج مورد تجزیه و تحلیل قرار گرفت که در گفتار، بیان صوتی متفاوتی دارند. این واج‌ها شامل فک پایین (/aa/ و /ae/)، زبانی (/l/ و /r/)، لبی (/uu/ و /ow/)، بینی (/m/ و /n/) و اصطکاکی (/s/ و /z/) بودند. روش‌های بیز ساده و تحلیل تفکیک خطی برای رده‌بندی دو رده‌ای بین تصور انواع واج‌ها اعمال شد که صحت بین ۶۶/۴٪ و ۷۶٪ بود. در سال ۲۰۱۳ الگوریتمی برای تمایز بین دو نویسه چینی تک هجا به معنای "چپ" و "یک" ارائه شد [۶]. بردارهای ویژگی سیگنال‌های الکتروانسفالوگرافی توسط الگوهای مشترک فضایی استخراج و سپس این بردارها توسط ماشین بردار پشتیبان رده‌بندی شدند که صحت بین ۶۶/۸۷٪ و ۸۳/۹۷٪ گزارش شد. تشخیص تصور پنج مصوت ژاپنی (/i/، /a/، /u/، /e/، /o/ و /v/) با استفاده از ماشین بردار مربوط با هسته گوسین به‌جای ماشین بردار پشتیبان برای کاهش هزینه محاسبات اعمال که منجر به صحت متوسط ۷۹٪ شد. درحالی‌که در [۸]

⁴ Classification

⁵ Accuracy

⁶ Subject

¹ Brain-computer interface (BCI)

² speech imagery

³ Electroencephalography (EEG)

عصبی با سه‌لایه پنهان بود. بالاترین صحت حدود ۸۶٪ گزارش شد.

پیشرفت در زمینه تصور گفتار با مجموعه داده‌های دسترسی باز به دست می‌آید که یکی از پایگاه داده در دسترس، 'Kara One' [۱۵] است. مجموعه داده‌های قابل دسترس به تسهیل تحقیقات تصور گفتار کمک می‌کنند. در سال ۲۰۱۵، پایگاه داده 'Kara One' در مؤسسه توان‌بخشی تورنتو ایجاد و منتشر شد که شامل سیگنال‌های جمع‌آوری شده از ۱۴ فرد سالم در طول تفکر و صحبت برای هفت واج /i/، /u/، /pi/، /di/، /ti/، /m/ و /n/ و چهار کلمه 'pat'، 'knew' و 'gnaw' است. این محرک‌ها برای داشتن تعداد نسبتاً برابری از مصوت‌ها، حروف منفجره^۲، بینی^۳ و همچنین واج‌های صدادار و بی‌صدا انتخاب شدند. از دیدگاه واج‌شناختی پنج رده‌بندی دودویی، مصوت در مقابل صامت، وجود واج بینی، وجود واج دوزبانی، وجود مصوت جلوی بالا (/i/) و وجود مصوت پشت بالا (/u/) انجام دادند. آن‌ها ویژگی‌های آماری مختلف را روی ۱۰٪ پنجره‌ها با ۵۰ درصد همپوشانی محاسبه کردند. حداکثر صحت ۷۹/۱۶٪ در رده‌بندی (/u/) با رده‌بند بردار ماشین پشتیبان گزارش شد [۱۵].

در سال ۲۰۱۶، یک شبکه عصبی مبتنی بر مدل‌سازی گفتار برای پایگاه داده 'Kara One' ارائه شد که در آن از سه ساختار شبکه برای نگاشت سیگنال‌های مغزی به واج‌ها استفاده شد. مدل پیشنهادی شامل یک‌لایه استخراج ویژگی مبتنی بر مدل زبان، یک‌لایه نگاشت ویژگی صوتی و یک ماشین بولتزمن محدود شده مبتنی بر لایه یادگیری ویژگی است. بهترین نتیجه در رده‌بندی ۷ واج و ۴ کلمه با استفاده از شبکه عصبی با صحت ۴۱/۵٪ به دست آمد که نشان می‌دهد مدل پیشنهادی از روش ماشین بردار پشتیبان برای رده‌بندی واج‌ها بهتر عمل می‌کند [۱۶]. در سال ۲۰۱۸ روشی برای تشخیص تصور گفتار از سیگنال‌های الکتروانسفالوگرافی با استفاده از ضرایب کپسترا ل بَسامد مل به عنوان استخراج ویژگی و رده‌بند ماشین بردار پشتیبان ارائه شد که به میانگین صحت ۲۰/۸۰٪ برای رده‌بندی واحدهای زبانی در پایگاه داده 'Kara One' رسید [۱۷].

³ Nasal

رده‌بند ماشین بردار پشتیبان برای تصور دو مصوت (/a/ و /u/) می‌تواند عملکردی بین ۷۷/۵٪ و ۱۰۰٪ با استخراج ویژگی‌های واریانس، آنالیز و انرژی سیگنال به دست آورد.

در سال ۲۰۱۷ رویکردی مبتنی بر ویژگی‌های چندگانه ریمانی برای رده‌بندی چهار مجموعه مختلف ارائه کردند که شامل تصور مصوت‌ها (/a/، /i/ و /u/)، کلمات کوتاه (out.in)، کلمات طولانی (independent.cooperate) و کلمات کوتاه و بلند (cooperate.in) است. صحت گزارش شده برای چهار مجموعه به ترتیب ۴۹/۲٪، ۵۰/۱٪، ۶۶/۲٪ و ۸۰/۱٪ است [۹]. در [۱۰] یک سامانه رابط مغز-رایانه مبتنی بر تصور گفتار با استفاده از کلمات بله و خیر برای آزمودنی‌های دوزبانه انگلیسی و هندی بررسی شد. به آزمودنی‌ها آموزش داده شد که پاسخ یک سؤال دودویی را به زبان انگلیسی یا هندی تصور کنند. تحلیل مؤلفه اصلی برای کاهش ابعاد داده‌ها و شبکه عصبی مصنوعی به عنوان رده‌بند استفاده که صحت کلی رده‌بندی گفتار دوزبانه ۷۵/۳۸٪ گزارش شد. مشابه کار قبل، در [۱۱] الگوریتمی بر اساس ویژگی‌های استخراج شده با استفاده از تبدیل موجک گسسته و شبکه‌های عصبی برای رده‌بندی تصورات بله و نه ارائه شد که آن‌ها صحتی در حدود ۶۷٪ گزارش کردند. اولین سامانه آنلاین شناسایی گفتار تصور شده در [۱۲] معرفی شد که کلمات انتخاب شده بله و نه بودند. روش آن‌ها با استفاده از ماشین بردار پشتیبان و ترکیبی از ویژگی‌های طیفی و زمان - بَسامد برای رده‌بندی کلمات بله و نه به میانگین صحت ۶۹/۳٪ رسید.

در [۱۳] از مجموعه داده‌ای استفاده شده که شامل تصور مصوت‌ها (/a/، /i/ و /u/)، کلمات کوتاه (out.in)، کلمات طولانی (independent.cooperate) است. آن‌ها ترکیب موازی شبکه عصبی پیچشی^۱ و شبکه‌های عصبی برگشتی را اعمال و سپس با استفاده از رمزگذار خودکار عمیق به شناسایی کلمات و مصوت‌ها پرداختند علی‌رغم پیچیدگی مدل، آن‌ها بهبود ۲۳ درصدی در صحت را گزارش کرده‌اند. همچنین در [۱۴] از شبکه‌های عصبی عمیق برای رده‌بندی تصور گفتار استفاده شد. هدف رده‌بندی تصور کلمات (in.cooperate) با تبدیل موجک گسسته و یک شبکه

¹ Convolutional neural networks (CNN)

² Plosive

بسامد به‌عنوان استخراج ویژگی و از رده‌بند ترکیبی شبکه عصبی پیچشی با یک شبکه حافظه بلند-کوتاهمدت استفاده کردند. نتایج نشان می‌دهد که ترکیب شبکه عصبی پیچشی با یک شبکه حافظه بلند-کوتاهمدت با افزایش صحت از ۳۷٪ به ۴۳٪ بهتر از شبکه عصبی پیچشی عمل می‌کند. در [۲۵] از بسامد لحظه‌ای و آنتروپی طیفی توان به‌عنوان ویژگی‌های سیگنال‌های الکتروانسفالوگرافی و از شبکه عصبی دوقلو برای مقایسه شباهت دسته‌های مختلف و از الگوریتم k-نزدیک‌ترین همسایه برای رده‌بندی تصور دو رده‌ای کلمات در پایگاه داده 'Kara One' استفاده که متوسط صحت حدود ۸۳٪ گزارش شده است. در [۲۶] بر پیش‌بینی متون از سیگنال‌های مغزی و پیشنهاد حروف بعدی از طریق رده‌بند شبکه حافظه دوطرفه بلند-کوتاهمدت و شبکه عصبی پیچشی مبتنی بر بهینه‌سازی ودینگ کروس انجام می‌شود که صحت برای 'Kara One' ۹۴/۲۹۷٪ گزارش شده است.

در [۲۷] رویکردی برای انتخاب ویژگی چند رده‌ای با استفاده از الگوریتم دودویی الگوی مشترک فضایی پیشنهاد شد. زمینه تصور گفتار به الگوریتم‌های سریعی برای محاسبه سیگنال‌های مغزی و تعامل با محیط نیاز دارد که این روش خیلی سریع نیست. در [۱] به بررسی یادگیری ویژگی‌های عمیق از سیگنال‌های الکتروانسفالوگرافی مربوط به تصور کلمات برای طراحی یک سامانه رابط مغز-رایانه بر اساس افکار انسان است. نمودارهای توپولوژیکی سیگنال‌های الکتروانسفالوگرافی با میانگین زمانی در تمام آزمایش‌های مربوط به تصور یک کلمه خاص به‌عنوان ورودی به شبکه عصبی عمیق تغذیه می‌شوند. معماری یادگیری عمیق طراحی شده دارای ادغام شبکه عصبی پیچشی دوبعدی و شبکه حافظه بلند-کوتاهمدت است که از قابلیت‌های هر دو معماری شبکه عصبی کمک می‌گیرد. صحت مدل طراحی شده با نمودارهای توپولوژیکی به‌عنوان ویژگی‌های ورودی ۲۰ تا ۲۵ درصد بیشتر از صحت حالت تصادفی برای مجموعه داده 'Kara One' گزارش شده است.

به‌طور خلاصه، بیشتر مقاله‌های پیشین، الگوریتمی برای رده‌بندی دودویی تصور گفتار با استفاده از ویژگی‌های دامنه زمان و بسامد

و شبکه عصبی پیچشی به ترتیب به‌عنوان استخراج ویژگی و رده‌بند استفاده شود بهبود می‌یابد و برای رده‌بندی واج‌ها به صحت ۲۴/۱۹٪ می‌رسد [۱۸]. همچنین آن‌ها در [۱۹] سه نوع واج با استفاده از ضرایب کپسترال بسامد مل و کدگذاری پیش‌بینی خطی همراه با شبکه عصبی پیچشی با میانگین مقدار $f1 = ۳۸/۳۳$ ٪ رمزگشایی کردند. آن‌ها نشان دادند که ضرایب کپسترال بسامد مل صحت بهتری نسبت به کدگذاری پیش‌بینی خطی ارائه می‌دهند و با الحاق آن‌ها به یک ماتریس دوکاناله می‌توان عملکرد بهتری نسبت به ترکیب آن‌ها در ماتریس یک کاناله به دست آورد.

در [۲۰] برای رده‌بندی دودویی در پایگاه داده 'Kara One' از راه‌برد شبکه عصبی عمیق ترکیبی متشکل از یک شبکه عصبی پیچشی، یک شبکه حافظه بلند-کوتاهمدت و یک رمزگذار خودکار عمیق استفاده کردند. شبکه عصبی عمیق سلسله‌مراتبی از ماتریس کوواریانس متقاطع به‌عنوان ماتریس ویژگی ورودی استفاده کرد. نتایج به دست آمده صحت کلی رده‌بندی وظایف دودویی را ۲۲/۵٪ افزایش داد و به میانگین صحت ۷۷/۹٪ در رده‌بندی دودویی دست یافت. در کار اخیرشان [۲۱] آن‌ها روش قبلی را کمی اصلاح کردند و شبکه بلندمدت را با ترکیب مکانی و زمانی از شبکه عصبی پیچشی برای ارزیابی عملکردشان جایگزین کردند. این روش به میانگین صحت ۸۳/۴۲٪ برای رده‌بندی ۲ رده‌ای دست یافت که ۵/۵۲٪ بیشتر از روش قبلی بود. رویکرد پیشنهادی در [۲۲] از کانال‌های الکتروانسفالوگرافی در مناطق خاصی از مغز برای طبقه‌بندی استفاده و بردارهای ویژگی متمایز با ترکیبی از ویژگی‌های دامنه زمان و موجک برای هر یک از کانال‌ها استخراج شد. برچسب رده نهایی هر آزمایش^۱ با اعمال اکثریت رأی در نتایج رده‌بندی کانال‌های جداگانه به دست آمد که برای رده‌بندی واج‌ها و کلمات در 'Kara One' به میانگین صحت ۵۷/۱۵٪ رسیدند.

در [۲۳] بهترین عملکرد گزارش شده تا ۳۷ درصد صحت برای تمام ۱۱ واج و کلمات مختلف هنگام استفاده از کوواریانس متقاطع برای یک پنجره ۰٫۲۵ ثانیه و یک شبکه عصبی پیچشی حاوی دولایه کانولوشن با فیلترهای ۶۴ و ۱۲۸ متصل به یک‌لایه متراکم با ۶۴ نورون است. آن‌ها در [۲۴] از کوواریانس متقاطع در حوزه

¹ Trial

واج‌ها را رده‌بندی می‌کند. مدل پیشنهادی، مستقل از سیگنال‌های مغزی هر فرد است که با آموزش مدل بر روی سیگنال‌های مغزی افزوده‌شده آزمودنی‌ها در دامنه منبع می‌توان به‌طور مؤثر سیگنال‌ها را در دامنه هدف بدون نیاز به داده‌های برچسب‌گذاری شده از آزمودنی جدید رده‌بندی کرد. از این‌رو، مدل پیشنهادی صحت رده‌بندی را به میزان قابل‌توجهی بهبود می‌بخشد و زمان به‌دست‌آوردن سیگنال‌های مغزی برای آزمودنی جدید را کاهش می‌دهد که برای کاربردهای رابط مغز-رایانه در دنیای واقعی می‌تواند موردتوجه باشد.

نوآوری‌ها و مشارکت‌های اصلی این مقاله به‌صورت زیر است:
(۱) افزایش سیگنال‌های الکتروانسفالوگرافی با استفاده از مدل هم‌وردانه مولدی که باعث بهبود عملکرد رده‌بندی سیگنال‌های مغزی مبتنی بر تصور گفتار می‌شود.

(۲) مدلی برای رده‌بندی تصور گفتار مبتنی بر سیگنال‌های مغزی آزمودنی جدید با یادگیری انتقالی با استفاده از روش مدل مولد هم‌وردانه مبتنی بر تطبیق دامنه متمایزکننده هم‌وردانه ارائه‌شده است.

(۳) مدل پیشنهادی، مستقل از سیگنال‌های مغزی هر فرد است که با آموزش مدل بر روی سیگنال‌های مغزی افزوده‌شده آزمودنی‌ها در دامنه منبع می‌تواند به‌طور مؤثر سیگنال‌ها را در دامنه هدف بدون نیاز به داده‌های برچسب‌گذاری شده از آزمودنی جدید رده‌بندی کند.

(۴) مدل پیشنهادی در مقایسه با سایر مدل‌های جدید بالاترین میانگین صحت و کاپا برای رده‌بندی ۷ واج و ۴ کلمه برای تصور گفتار در پایگاه داده KaraOne ارائه کرد.

(۵) مدل پیشنهادی صحت رده‌بندی را به میزان قابل‌توجهی بهبود می‌بخشد و زمان به‌دست‌آوردن سیگنال‌های مغزی برای آزمودنی جدید را کاهش می‌دهد که برای کاربردهای رابط مغز-رایانه در دنیای واقعی کاربردی است.

و همچنین از ترکیب شبکه‌های عصبی استفاده کرده‌اند که رده‌بند با صحت بالایی برای رده‌بندی تمام واج‌ها و کلمات مختلف برای مجموعه داده 'Kara One' گزارش نشده است. از آنجایی‌که جمع‌آوری داده‌های کافی برای آزمودنی جدید زمان و تلاش زیادی می‌طلبد مجموعه داده‌های سیگنال‌های مغزی معمولاً شامل تعداد محدودی از نمونه‌های آموزشی است. با افزایش تعداد رده‌ها و تعداد محدود نمونه‌ها برای آزمودنی‌های مختلف و باتوجه به خصوصیات سیگنال‌های الکتروانسفالوگراف در سامانه‌های رابط مغز-رایانه عملکرد رده‌بندی سیگنال‌های مغزی مبتنی بر تصور گفتار به‌طور قابل‌توجهی مسئله‌ای چالش‌برانگیز است؛ بنابراین در این مقاله افزودن داده‌ها^۱ با استفاده از شبکه مولد هم‌وردانه^۲ و همچنین از یادگیری انتقالی با استفاده از روش تطبیق دامنه متمایزکننده هم‌وردانه^۳ پیشنهاد شده است. باتوجه به مطالعات انجام شده تاکنون روش افزایش سیگنال‌های الکتروانسفالوگرافی با استفاده از مدل مولد هم‌وردانه و همچنین استفاده از یادگیری انتقالی با روش مدل مولد هم‌وردانه مبتنی بر تطبیق دامنه متمایزکننده هم‌وردانه برای رده‌بندی سیگنال‌های مغزی مبتنی بر تصور گفتار ارائه نشده است.

بنابراین در این مقاله افزودن داده‌ها با استفاده از شبکه مولد هم‌وردانه به‌منظور بهبود عملکرد رده‌بندی سیگنال‌های مغزی پیشنهاد شده است. برای هر رده و هر آزمودنی دو برابر داده اولیه با صحت بالا ایجاد شده است. همچنین استفاده از سیگنال‌های الکتروانسفالوگراف سایر آزمودنی‌ها به‌عنوان دامنه منبع^۴ برای ایجاد مدلی برای رده‌بندی سیگنال‌های مغزی آزمودنی جدید به‌عنوان دامنه هدف^۵ با یادگیری انتقالی با استفاده از روش تطبیق دامنه هم‌وردانه ارائه‌شده است. به‌منظور شناسایی تصور گفتار از پایگاه داده KaraOne با ۷ واج و ۴ کلمه استفاده شده است. روش پیشنهادی با سایر روش‌های جدید بر اساس معیارهای صحت و کاپا^۶ مورد ارزیابی قرار گرفتند. طبق نتایج به‌دست‌آمده روش پیشنهادی با صحت ۸۶٪ و ۶۰/۲۱٪ به ترتیب تصور کلمات و

⁴ Source domain

⁵ Target domain

⁶ Kappa

¹ Data augmentation

² Generative adversarial network (GAN)

³ Adversarial discriminative domain adaptation (ADDA)

۵- از تمام داده‌های آموزشی و سیگنال‌های مشابه افزوده شده با استفاده از شبکه عصبی پیچشی ویژگی‌ها استخراج می‌شوند.

۶- رده‌بندی تصور گفتار برای واج‌ها و کلمات مختلف برای آزمودنی جدید با استفاده از یادگیری انتقالی انجام می‌شود.

شبکه‌های مولد هم‌وردانه رویکردی برای تولید نمونه‌ها بر اساس فن‌های یادگیری عمیق مانند شبکه‌های عصبی عمیق پیچشی هستند. شبکه‌های مولد هم‌وردانه از فرایندهای هم‌وردانه بین شبکه‌های مولد^۱ و متمایزکننده^۲ برای آموزش هم‌زمان آن‌ها استفاده می‌کنند. شبکه مولد در شبکه‌های مولد هم‌وردانه یک بردار تصادفی می‌گیرد و نمونه‌هایی را تولید می‌کند. شبکه متمایزکننده نمونه‌ای از مجموعه آموزشی یا نمونه‌های تولیدشده از شبکه مولد را می‌گیرد و سپس نمونه‌ها را به دو رده با برچسب‌های واقعی یا جعلی (تولیدشده) رده‌بندی می‌کند. پس از طی مراحل آموزشی، از شبکه مولد برای تولید نمونه‌های جدید استفاده می‌شود. همان‌طور که در شکل (۲) نمایش داده شده است، در این پژوهش از شبکه مولد هم‌وردانه عمیق پیچشی^۳ با معماری و پارامترهای خاص برای تولید سیگنال‌های الکتروانسفالوگرافی تصور گفتار استفاده شده است. به منظور شناسایی تصور گفتار از پایگاه داده 'KaraOne' با هفت واج /a/، /i/، /u/، /pi/، /di/، /ti/، /m/ و /n/ و چهار کلمه 'pat'، 'knew' و 'gnaw' استفاده شده است.

استفاده از سیگنال‌های الکتروانسفالوگرافی سایر آزمودنی‌ها به عنوان دامنه منبع برای ساخت مدلی برای رده‌بندی سیگنال‌های الکتروانسفالوگرافی یک آزمودنی جدید به عنوان دامنه هدف، یک فن قابل‌دستیابی در توسعه یک رابط مغز-رایانه است. عملکرد رده‌بندی آزمودنی موردنظر با استفاده از آموزش بر روی داده‌های منبع بدون تطبیق دامنه به دلیل ویژگی‌های سیگنال‌های الکتروانسفالوگرافی کاهش می‌یابد. برای کاهش اختلاف توزیع بین دامنه‌های منبع و هدف برای سیگنال‌های الکتروانسفالوگرافی تصور گفتار در کاربرد مغز-رایانه، در این پژوهش از یادگیری انتقالی با استفاده از روش تطبیق دامنه هم‌وردانه ارائه شده است. یادگیری یک رده‌بند با عملکرد تعمیم قوی در دامنه هدف، مقصود روش‌های تطبیق دامنه است. برای رده‌بندی داده‌های سیگنال‌های

ادامه مقاله به شرح زیر سازمان‌دهی شده است: در بخش دوم جزئیات مدل پیشنهادی برای افزودن سیگنال‌های مغزی با استفاده از شبکه مولد هم‌وردانه ارائه و همچنین سه جزء اصلی شبکه تطبیق دامنه هم‌وردانه به منظور یادگیری انتقالی بیان می‌شود. در بخش سوم مجموعه داده و معیارهای ارزیابی توضیح داده شده است. در بخش چهارم نتایج ارزیابی و در نهایت در بخش پنجم نتیجه‌گیری کلی این پژوهش و کارهای آینده ارائه شده است.

۲- روش پیشنهادی

در این بخش ابتدا مراحل اولیه مدل پیشنهادی معرفی می‌شود. سپس الگوریتم افزودن داده با استفاده از شبکه مولد هم‌وردانه بیان می‌شود. در پایان، الگوریتم تطبیق دامنه متمایزکننده هم‌وردانه به منظور یادگیری انتقالی مبتنی بر شبکه مولد هم‌وردانه ارائه می‌شود.

۲-۱ طرح کلی مدل

نمودار بلوکی مدل پیشنهادی برای رده‌بندی تصور گفتار مبتنی بر سیگنال‌های مغزی با افزودن داده‌ها و شبکه تطبیق دامنه هم‌وردانه در کاربرد رابط مغز-رایانه در شکل (۱) نشان داده شده است. بلوک‌های آبی بیانگر نوآوری مقاله است. این مدل شامل شش مرحله اصلی به شرح زیر است:

- ۱- سیگنال‌های مغزی مبتنی بر تصور گفتار با روش غیرتهاجمی از سطح پوست سر از طریق الکترودها ثبت می‌شود.
- ۲- سیگنال‌های اولیه برای افزایش نسبت سیگنال به نویز و حذف مصنوعات پیش‌پردازش می‌شوند.
- ۳- به دلیل نیاز به داده‌های الکتروانسفالوگراف کافی برای هر آزمودنی، با استفاده از مدل مولد هم‌وردانه، سیگنال‌های مشابه داده آموزشی ایجاد می‌شود.
- ۴- با روش تطبیق دامنه هم‌وردانه از سیگنال‌های الکتروانسفالوگراف سایر آزمودنی‌ها به عنوان دامنه منبع برای رده‌بندی سیگنال‌های مغزی آزمودنی جدید به عنوان دامنه هدف از روش یادگیری انتقالی استفاده می‌شود.

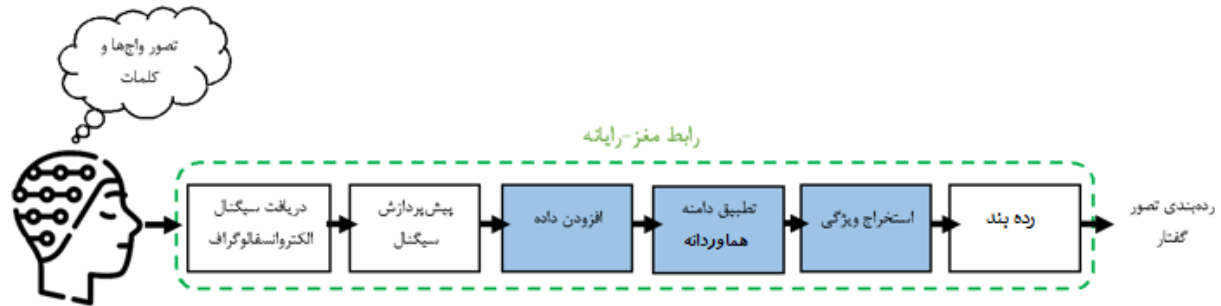
³ Deep convolutional GAN (DCGAN)

¹ Generator

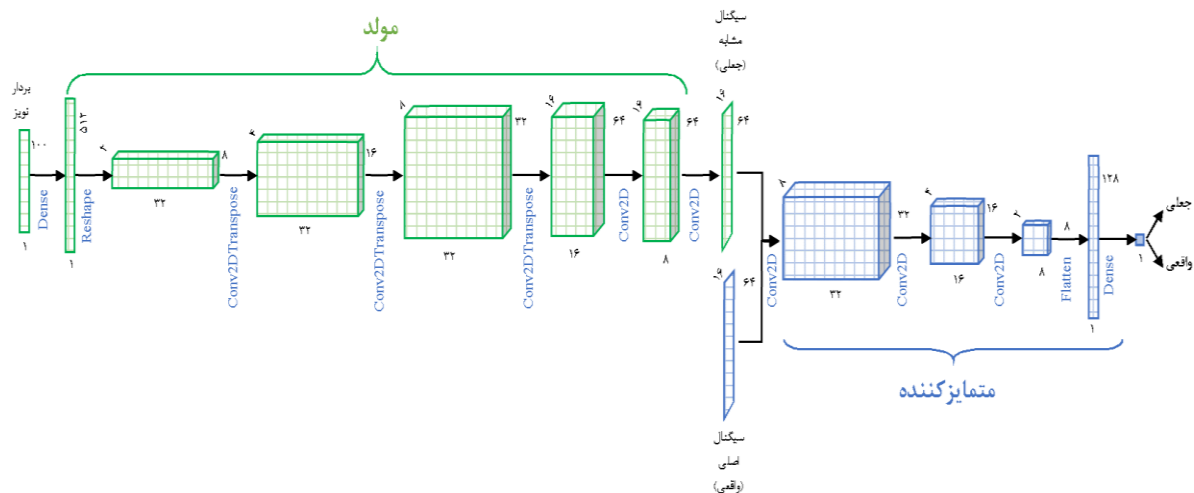
² Discriminator

شکل نشان داده شده است که از سه جزء اصلی تشکیل شده است: دو رمزگذار به عنوان استخراج کننده ویژگی ها، یک متمایز کننده دامنه و یک رده بند. در ادامه این سه جزء معرفی می شود.

الکتروانسفالوگرافی تصور گفتار از دامنه هدف با برچسب داده های سیگنال های تصور گفتار از دامنه منبع، در این پژوهش یک شبکه تطبیق دامنه هماوردانه پیشنهاد می شود. نمودار بلوکی مدل تطبیق دامنه متمایز کننده هماوردانه مبتنی بر شبکه مولد هماوردانه در



شکل (۱): نمایش ساده از مدل پیشنهادی برای رده بندی تصور گفتار مبتنی بر سیگنال الکتروانسفالوگراف



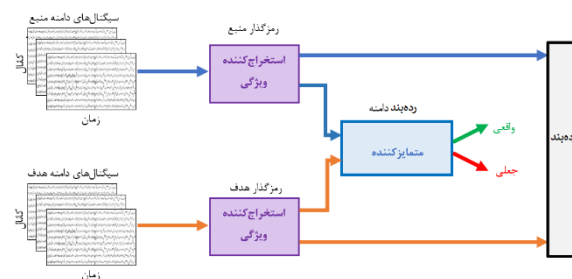
شکل (۲): مدل پیشنهادی شبکه مولد هماوردانه عمیق پیچشی برای تصور گفتار

جلوگیری از بیش برازش، یک لایه ادغام^۱ بعد از هر لایه پیچشی در شبکه قرار می گیرد. جدول ۱ مشخصات طراحی استخراج کننده ویژگی را برای مدل تطبیق دامنه متمایز کننده هماوردانه مبتنی بر شبکه مولد هماوردانه نشان می دهد

استخراج کننده ویژگی برای سیگنال های الکتروانسفالوگرافی تصور گفتار با استفاده از شبکه عصبی پیچشی پیاده سازی شده است. هر نمونه از سیگنال های الکتروانسفالوگرافی $x \in R^{E \times T}$ حاوی اطلاعات زمانی و مکانی به عنوان ورودی به شبکه عصبی پیچشی داده می شود. شبکه عصبی پیچشی از عملیات کانولوشن دوبعدی استفاده می کند تا فیلترهای زمانی و مکانی را ترکیب کند، بنابراین ویژگی ها با ترکیب کانال ها در طول زمان استخراج می شوند. به منظور

^۱ Pooling

فرض می‌شود. در نهایت، استخراج‌کننده ویژگی (رمزگذار هدف) آموزش داده می‌شود تا شبکه متمایزکننده را فریب دهد تا شبکه متمایزکننده نتواند تمایز بین ویژگی‌های استخراج‌شده دامنه‌های منبع و هدف قائل شود؛ بنابراین، استخراج‌کننده ویژگی (رمزگذار هدف) توانایی نگاشت داده‌های سیگنال‌های الکتروانسفالوگراف را در فضای ویژگی مشترک توسط آزمون‌های منبع و هدف به دست می‌آورد. در مدل تطبیق دامنه متمایزکننده هم‌آوردانه مبتنی بر شبکه مولد هم‌آوردانه، شبکه متمایزکننده به‌عنوان رده‌بند دامنه، تعیین می‌کند که آیا ویژگی‌های استخراج‌شده از دامنه منبع یا هدف، متناظر با خروجی‌های واقعی یا جعلی هستند. شبکه متمایز دامنه شامل چهارلایه کاملاً متصل با تعداد نوروهای ۱۲۸، ۶۴، ۳۲ و ۱ است. در لایه آخر، تابع فعال‌سازی سیگموئید برای پیش‌بینی این که ورودی متعلق به کدام دامنه است استفاده می‌شود.



شکل (۳): مدل تطبیق دامنه متمایزکننده هم‌آوردانه مبتنی بر

شبکه مولد هم‌آوردانه برای رده‌بندی تصور گفتار

شبکه متمایزکننده دامنه با استخراج‌کننده ویژگی به‌عنوان رمزگذار هدف از طریق یادگیری هم‌آوردانه آموزش می‌بیند. پس از آموزش استخراج‌کننده ویژگی به‌عنوان رمزگذار منبع و رده‌بند با استفاده از داده‌های برچسب‌دار دامنه منبع، خروجی استخراج‌کننده ویژگی به‌عنوان داده واقعی در نظر گرفته می‌شود. همچنین، خروجی استخراج‌کننده ویژگی (رمزگذار هدف) به‌عنوان داده‌های جعلی

جدول (۱): معماری استخراج‌کننده ویژگی با ساختار شبکه عصبی پیچشی برای مدل تطبیق دامنه متمایزکننده هم‌آوردانه مبتنی بر شبکه مولد

هم‌آوردانه

خصوصیات	بعد خروجی	نگاشت ویژگی	کرنل	لایه	شبکه
	(16,64,1)			Input	استخراج ویژگی با ساختار شبکه عصبی پیچشی دو لایه‌ای
padding='same', activation='tanh'	(16,64,4)	4	(2,2)	Convolution	
	(8,32,4)		(2,2)	Max Pooling	
padding='same', activation='tanh'	(8,32,4)	4	(2,2)	Convolution	
	(4,16,4)		(2,2)	Max Pooling	
padding='same', activation='tanh'	(4,16,2)	2	(2,2)	Convolution	
	(2,8,2)		(2,2)	Max Pooling	
خصوصیات	(32)			Flatten	

کاهش می‌یابد. برای پرداختن به این چالش، در این پژوهش از مدل تطبیق دامنه متمایزکننده هم‌آوردانه مبتنی بر شبکه مولد هم‌آوردانه برای کاهش اختلاف توزیع بین دامنه‌های منبع و هدف برای سیگنال‌های الکتروانسفالوگرافی تصور گفتار در کاربرد رابط مغز-رایانه استفاده کردیم. در ابتدا استخراج‌کننده ویژگی (رمزگذار منبع) و رده‌بند با استفاده از داده‌های برچسب‌دار دامنه منبع آموزش

هدف رده‌بند رمزگشایی از بازنمایی ویژگی‌های به‌دست‌آمده با استفاده از استخراج‌کننده ویژگی و پیش‌بینی برچسب‌های تصور گفتار برای هر نمونه از سیگنال‌های الکتروانسفالوگراف است. باتوجه به عدم تطابق دامنه‌های منبع و هدف، کارایی رده‌بند که صرفاً بر روی داده‌های منبع بدون تطبیق دامنه، آموزش داده شود، هنگام اعمال بر روی سیگنال‌های الکتروانسفالوگرافی داده‌های هدف

رده‌بندی تصور گفتار مبتنی بر سیگنال‌های مغزی با افزودن داده‌ها و شبکه تطبیق دامنه هم‌آوردانه

با استفاده از یادگیری عمیق در کاربرد رابط مغز-رایانه

مدل پیشنهادی شبکه مولد هم‌وردانه عمیق پیچشی برای سیگنال‌های الکتروانسفالوگرافی تصور گفتار شامل یک شبکه مولد و یک شبکه متمایزکننده مبتنی بر شبکه عصبی پیچشی است. شبکه مولد یک بردار تصادفی با ۱۰۰ عنصر از توزیع گاوسی می‌گیرد. لایه متراکم به‌عنوان اولین لایه پنهان برای نگاشت یک بردار تصادفی به چندین سیگنال با وضوح پایین استفاده می‌شود. سپس این گره‌ها تغییر شکل به ۳۲ نگاشت ویژگی متفاوت با ابعاد 2×8 می‌دهند. در مرحله بعد، فرایند نمونه‌برداری بالا توسط لایه‌های Conv2DTranspose با فعال‌سازی LeakyReLU اعمال می‌شود. در نهایت، لایه Conv2D به‌عنوان لایه خروجی برای ایجاد سیگنال‌های الکتروانسفالوگرافی مشابه با شانزده کانال و ۶۴ نقطه نمونه‌برداری است. معماری دقیق شبکه مولد در جدول ۲ آمده است.

شبکه متمایزکننده نمونه سیگنال‌های الکتروانسفالوگرافی تصور گفتار با ابعاد (۱، ۶۴، ۱۶) را از مجموعه آموزشی یا خروجی شبکه مولد می‌گیرد. این مدل شامل سه لایه Conv2D با فعال‌سازی LeakyReLU است. لایه‌های Dropout به‌عنوان یک ترفند ضروری برای جلوگیری از بن‌بست در مدل شبکه مولد هم‌وردانه استفاده می‌شود. در نهایت، لایه متراکم با یک گره به‌عنوان لایه خروجی با فعال‌سازی سیگموئید قرار دارد تا پیش‌بینی کند که سیگنال ورودی واقعی یا جعلی است. معماری دقیق شبکه متمایزکننده در جدول ۳ فهرست شده است.

داده می‌شوند. سپس استخراج‌کننده ویژگی (رمزگذار هدف) با استفاده از بازخورد شبکه متمایزکننده، به‌صورت هم‌وردانه آموزش داده می‌شود. در نهایت، رده‌بند آموزش دیده شده با داده‌های دامنه منبع، به‌صحت برچسب داده‌های دامنه هدف را بدون استفاده از داده‌های برچسب‌گذاری شده هدف پیش‌بینی می‌کند. در این مطالعه، داده‌های دامنه منبع و هدف با استفاده از شبکه کاملاً متصل با دولایه رده‌بندی می‌شود.

۲-۲ مدل پیشنهادی افزودن داده با استفاده از شبکه

مولد هم‌وردانه

یکی از چالش‌های رده‌بندی مجموعه داده‌های تصور گفتار، تعداد نمونه‌های کم برای هر آزمودنی است. مقدار ناکافی داده منجر به بیش برآزش مدل‌های مختلف می‌شود. برای حل این مشکل، محققان از افزودن داده‌ها به‌عنوان یک رویکرد مناسب برای جلوگیری از بیش برآزش و بهبود عملکرد تصور گفتار چند رده‌ای در رده‌بندی رابط مغز-رایانه مبتنی بر سیگنال الکتروانسفالوگرافی استفاده می‌کنند [۲۸]. در این پژوهش از فن یادگیری عمیق برای افزودن مجموعه داده سیگنال‌های الکتروانسفالوگرافی برای بهبود عملکرد رابط مغز-رایانه مبتنی بر تصور گفتار استفاده شد. در مدل پیشنهادی با استفاده از یادگیری بدون نظارت آموزش داده و تعداد نمونه‌ها را بدون نیاز به برچسب رده‌ها افزایش داده است. در ادامه عملکرد مدل افزودن داده‌ها برای رده‌بندی سیگنال‌های الکتروانسفالوگرافی تصور گفتار مورد بررسی قرار گرفته است.

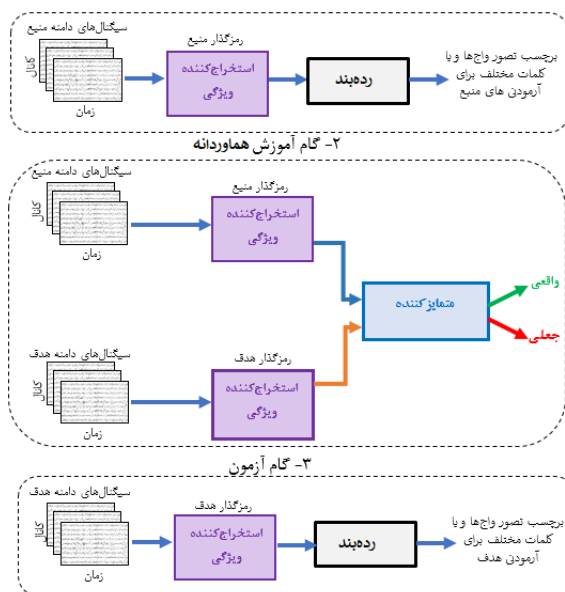
جدول (۲): معماری مولد در شبکه مولد هم‌وردانه عمیق پیچشی

لایه‌ها	کرنل	نگاشت ویژگی	بعد خروجی	فعال‌سازی	خصوصیات
Input			100		
Dense			$32 \times 2 \times 8 = 512$	LeakyReLU(alpha=0.2)	
Reshape			(2, 8, 32)		
Conv2DTranspose	(4,4)	32	(4,16,32)	LeakyReLU(alpha=0.2)	strides=(2,2),padding='same'
Conv2DTranspose	(4,4)	32	(8,32,32)	LeakyReLU(alpha=0.2)	strides=(2,2),padding='same'
Conv2DTranspose	(4,4)	16	(16,64,16)	LeakyReLU(alpha=0.2)	strides=(2,2),padding='same'
Conv2D	(4,4)	8	(16,64,8)	Tanh	padding='same'
Conv2D	(2,2)	1	(16,64,1)	Tanh	padding='same'

جدول (۳): معماری متمایزکننده در شبکه مولد هم‌آوردانه عمیق پیچشی

لایه‌ها	کرل	نگاشت ویژگی	بعد خروجی	فعال‌سازی	خصوصیات
Input			(16,64,1)		
Conv2D	(4,4)	32	(8,32,32)	LeakyReLU(alpha=0.2)	strides=(2,2),padding='same'
Dropout			(8,32,32)		Drop rate=0.4
Conv2D	(4,4)	16	(4,16,16)	LeakyReLU(alpha=0.2)	strides=(2,2),padding='same'
Dropout			(4,16,16)		Drop rate=0.4
Conv2D	(4,4)	8	(2,8,8)	LeakyReLU(alpha=0.2)	strides=(2,2),padding='same'
Dropout			(2,8,8)		Drop rate=0.4
Flatten			128		
Dense			1	Sigmoid	خصوصیات

هنگامی که خطای رده‌بند به حداقل برسد، پارامترهای رمزگذار منبع برای مقداردهی اولیه پارامترهای رمزگذار هدف استفاده می‌شود.
۱- گام پیش از آموزش



شکل (۴): مراحل مدل تطبیق دامنه متمایزکننده هم‌آوردانه مبتنی بر شبکه مولد هم‌آوردانه برای رده‌بندی تصویر گفتار

گام آموزش هم‌آوردانه: استخراج‌کننده ویژگی (رمزگذار هدف) و شبکه متمایزکننده هر دو به روش هم‌آوردانه آموزش داده می‌شوند. با استفاده از یادگیری بدون نظارت، شبکه متمایزکننده برای تمایز بین سیگنال‌های دامنه‌های منبع و هدف آموزش داده می‌شود.

۲-۳ مدل پیشنهادی تطبیق دامنه متمایزکننده هم‌آوردانه برای تصور گفتار

استفاده از سیگنال‌های سایر آزمودنی‌ها به عنوان دامنه منبع برای ساخت مدلی برای رده‌بندی سیگنال‌های آزمودنی جدید به عنوان دامنه هدف، یک فن قابل‌دستیابی است. در مدل تطبیق دامنه متمایزکننده هم‌آوردانه مبتنی بر شبکه مولد هم‌آوردانه، سیگنال‌های الکتروانسفالوگرافی تصور گفتار دامنه منبع و برچسب‌های آن از توزیع منبع فراهم هستند. همچنین سیگنال‌های الکتروانسفالوگرافی تصور گفتار دامنه هدف بدون برچسب‌گذاری از توزیع هدف وجود دارد. با وجود فقدان برچسب سیگنال‌ها در دامنه هدف، مقصود آموزش رده‌بند برای پیش‌بینی دقیق برچسب‌های داده‌های هدف است. هدف روش‌های تطبیق دامنه هم‌آوردانه کاهش اختلاف بین توزیع دامنه‌های منبع و هدف است. در نتیجه، همان‌طور که در شکل ۴ نشان داده شده است، در این پژوهش مدل تطبیق دامنه متمایزکننده هم‌آوردانه مبتنی بر شبکه مولد هم‌آوردانه برای رده‌بندی سیگنال‌های الکتروانسفالوگرافی تصور گفتار در کاربرد رابط مغز-رایانه پیشنهاد شده است که شامل ۳ مرحله پیش از آموزش^۱، آموزش هم‌آوردانه و آزمون است.

گام پیش از آموزش: در یادگیری نظارت‌شده، از مجموعه داده منبع برای آموزش استخراج‌کننده ویژگی و رده‌بند استفاده می‌شود.

¹ Pre-training

به‌منظور توسعه کاربرد رابط مغز-رایانه برای رده‌بندی تصور گفتار از قسمت تصور گفتار برای هر آزمودنی استفاده شده است.

کانال‌های سیگنال‌های الکتروانسفالوگرافی بر اساس نواحی آناتومیک مغز درگیر در مفهوم‌سازی گفتار، طرح‌های بیانی، شروع و هماهنگی محرک حرکتی برای ارسال به عوامل مؤثر انتخاب شدند. باتوجه‌به نواحی آناتومیک مغز درگیر در گفتار به‌خوبی شناخته شده است که ناحیه بروکا نقش مهمی در تولید گفتار دارد، قشر حرکتی اولیه سیگنال‌هایی را تولید می‌کند که اجرای حرکات را کنترل می‌کند که با تخیل حرکتی نیز مرتبط است و ناحیه حرکتی ثانویه مسئول برنامه‌ریزی حرکتی است [۲۴]؛ بنابراین در این پژوهش بر اساس نواحی آناتومیک مغز برای تولید گفتار کانال‌های T7, FT7, FC5, FC3, FC1, FZ, F1, F3, F5, F7, AF3, C1, C3, C5 و CZ انتخاب شدند.

در مدل پیشنهادی افزودن داده با استفاده از شبکه مولد هم‌اوردانه از بهینه‌ساز RMSProp با نرخ یادگیری 0.0004 و از تابع فقدان binary_crossentropy برای شبکه‌های مولد و متمایزکننده استفاده شد. هنگامی که صحت تشخیص سیگنال‌های واقعی و جعلی در شبکه متمایزکننده به حدود ۱۰۰٪ رسید به‌عبارت دیگر زمانی که شبکه مولد توانست شبکه متمایزکننده را با صحت بسیار بالایی فریب دهد آن‌گاه از شبکه مولد به‌عنوان تولید نمونه‌های سیگنال الکتروانسفالوگرافی تصور گفتار با صحت بالا استفاده می‌شود. برای هر رده از تصور گفتار و هر آزمودنی به‌طور جداگانه سیگنال‌های مشابه مجموعه داده آموزشی تولید می‌کنیم که در این پژوهش ۲ برابر سیگنال‌ها الکتروانسفالوگرافی تصور گفتار با استفاده از شبکه مولد هم‌اوردانه تولید می‌شود.

در مدل تطبیق دامنه متمایزکننده هم‌اوردانه مبتنی بر شبکه مولد هم‌اوردانه دو مجموعه داده از سیگنال‌های الکتروانسفالوگرافی تصور گفتار به‌عنوان دامنه‌های منبع و هدف برای آزمودنی‌ها در نظر گرفته می‌شود. هر یک از چهارده آزمودنی در مجموعه داده 'KaraOne' می‌توانند در مجموعه داده دامنه‌های منبع و هدف باشند. هنگامی که یک آزمودنی به‌عنوان دامنه هدف انتخاب شد، سیزده آزمودنی باقیمانده به‌عنوان دامنه منبع در نظر گرفته می‌شوند.

درحالی که رمزگذار هدف برای فریب دادن شبکه متمایزکننده آموزش داده می‌شود؛ بنابراین پس از آموزش، رمزگذار هدف، سیگنال‌های الکتروانسفالوگرافی تصور گفتار دامنه هدف را با توزیع‌های سیگنال‌های الکتروانسفالوگرافی تصور گفتار دامنه منبع هماهنگ می‌کند.

گام آزمون: پس از آموزش هم‌اوردانه، رده‌بند که در گام پیش از آموزش، با مجموعه سیگنال‌های دامنه منبع آموزش دیده می‌تواند برچسب سیگنال‌های الکتروانسفالوگرافی تصور گفتار دامنه هدف را پیش‌بینی کند. در مرحله آزمون، مجموعه داده‌های دامنه هدف به رده‌بند از پیش آموزش دیده از طریق رمزگذار هدف داده می‌شود تا برچسب سیگنال‌های الکتروانسفالوگرافی تصور گفتار دامنه هدف را شناسایی کند.

۳- آزمایش‌ها

مدل پیشنهادی افزودن داده و همچنین مدل تطبیق دامنه متمایزکننده هم‌اوردانه مبتنی بر شبکه مولد هم‌اوردانه در پایتون با استفاده از Keras API با چارچوب TensorFlow پیاده‌سازی شدند. همچنین پیش‌پردازش‌های لازم برای سیگنال‌های الکتروانسفالوگرافی تصور گفتار در نرم‌افزار متلب انجام شد. آزمایش‌ها روی مک‌بوک پرو رتینا با پردازنده Core i5 ایتل با بسامد $2/7$ گیگاهرتز انجام شدند. در این بخش ابتدا پیش‌پردازش‌های انجام شده روی سیگنال الکتروانسفالوگرافی تصور گفتار بیان سپس مجموعه داده و معیارهای ارزیابی شرح داده می‌شود.

ابتدا سیگنال‌های الکتروانسفالوگراف با استفاده از EEGLAB به‌منظور حذف مصنوعات چشمی پردازش می‌شوند. سپس سیگنال‌ها بین ۱ تا ۵۰ هرتز فیلتر شدند و میانگین مقادیر از هر کانال کم شد. همچنین با استفاده از همسایگی کانال‌های مجاور، فیلتر لاپلاسین روی داده‌ها اعمال کردیم. سپس الگوریتم تحلیل مؤلفه‌های مستقل اعمال شد. داده‌های سیگنال‌های الکتروانسفالوگرافی به قسمت‌های مختلف تقسیم شدند و هر آزمایش به ۴ حالت (استراحت، تحریک، تصور گفتار، صحبت) که در بخش مجموعه داده (۳-۱) توضیح داده می‌شود تقسیم شد.

۱-۳ مجموعه داده

از شرکت‌کنندگان در دانشگاه تورنتو ثبت داده‌های سیگنال‌های مغزی انجام شده است. تمام شرکت‌کنندگان دست راست بودند. هیچ‌کدام از شرکت‌کنندگان دارای اختلالات بینایی، شنوایی یا حرکتی نبودند و هیچ مشکل نورولوژیکی و سوءمصرف مواد نداشتند. هر جلسه ثبت داده در محیط اداری در مؤسسه توان‌بخشی تورنتو انجام شد. هر شرکت‌کننده روی یک صندلی پشت مانیتور رایانه نشسته است. دوربین فیلم‌برداری کینکت مایکروسافت (v.1.8) در کنار صفحه نمایش قرار داده شد تا اطلاعات صورت و سخنرانی شرکت‌کننده را ضبط کند. از دستگاه neuroscan ۶۴ کاناله استفاده شده است که در آن، قرارگیری الکتروده از سامانه ۱۰-۲۰ پیروی می‌کند. تمام داده‌های سیگنال الکتروانسفالوگرافی با استفاده از تقویت‌کننده SynAmps RT ثبت شده است و با ۱ کیلوهرتز نمونه‌گیری انجام شده است. مقاومت هر الکتروده معمولاً کمتر از ۱۰ کیلو اهم نگه‌داشته می‌شود.

پس از راه‌اندازی و تنظیمات دستگاه سیگنال الکتروانسفالوگرافی به شرکت‌کننده دستور داده می‌شود تا به مانیتور رایانه نگاه کند و تا حد ممکن کم‌ترین حرکت را در هنگام ثبت داده سیگنال‌های الکتروانسفالوگرافی داشته باشد. در طول ۳۰ تا ۴۰ دقیقه، دستورها روی صفحه‌نمایش یکی‌یکی نمایش داده می‌شود. از ۷ واج (/iy/, /i/, /uw/, /piy/, /tiy/, /diy/, /m/, /n/ pat, pot, knew, و ۴ کلمه gnaw) استفاده شده است.

هر آزمایش شامل ۴ حالت متوالی است:

- ۱- یک حالت استراحت ۵ ثانیه‌ای است که در آن به شرکت‌کننده دستور داده شده که استراحت کنند و ذهن خود را از هر اندیشه پاک کنند.
- ۲- حالت بعدی تحریک است در اینجا متن روی صفحه‌نمایش داده می‌شود و بر روی بلندگو رایانه دستور موردنظر پخش می‌شود. این یک دوره ۲ ثانیه بوده که در آن شرکت‌کننده خود را آماده برای بیان آن دستور می‌کند.
- ۳- یک حالت تصور گفتار ۵ ثانیه‌ای است که در آن شرکت‌کننده دستور نمایش داده شده را تصور می‌کند.

۴- حالت صحبت است که در آن شرکت‌کننده با صدای بلند

دستور نمایش داده شده را می‌خواند. سنسور دوربین هر دو ویژگی صوتی و صورت را در این مرحله ثبت می‌کند. به‌طور طبیعی، باتوجه به تأثیر حرکت در سیگنال الکتروانسفالوگرافی، انتظار می‌رود نویز در سیگنال حالت گفتاری ایجاد شود. هنگامی که شرکت‌کننده به صحبت خود پایان داد آزمایش بعدی برای دستور بعدی شروع می‌شود. هر یک از دستورها ۱۲ بار و در مجموع ۱۳۲ بار تکرار می‌شود. ابتدا واج‌ها تکرار می‌شوند و سپس کلمات ارائه می‌شود. آزمایش‌ها به صورت تصادفی در هر یک از این دو بخش تکرار می‌شود. پس از هر ۴۰ آزمایش به شرکت‌کننده فرصتی برای استراحت داده می‌شود [۱۵].

۲-۳ معیارهای ارزیابی

برای مقایسه عملکرد مدل پیشنهادی تطبیق دامنه متمایزکننده هماوردها برای رده‌بندی تصور گفتار با سایر روش‌های جدید از معیارهای صحت و کاپا به عنوان معیارهای ارزیابی استفاده شده است. متداول‌ترین معیاری که برای ارزیابی عملکرد مسئله رده‌بندی استفاده می‌شود، میزان صحت رده‌بندی است که با استفاده از رابطه (۱) به دست می‌آید:

$$Classification Accuracy = \left(\frac{N_C}{N_T} \right) \times 100\% \quad (1)$$

که در آن N_C تعداد واج‌ها و کلمات تصور گفتاری که به درستی رده‌بندی شده و N_T تعداد کل آزمایش‌ها است.

معیار کاپا یک معیار ارزشمند در رده‌بندی چند رده‌ای است، زیرا نه تنها رده‌بندی درست بلکه رده‌بندی اشتباه را نیز در نظر می‌گیرد و همچنین اثرات خطاهای تصادفی را حذف می‌کند. علاوه بر این، مقدار کاپا نشان‌دهنده تعمیم مدل و سازگاری در رده‌بندی مسائل چند رده‌ای است [۲۹].

مقدار کاپا با استفاده از رابطه (۲) محاسبه می‌شود:

$$kappa = \frac{p_0 - p_e}{1 - p_e} \quad (2)$$

که p_0 و p_e به ترتیب صحت رده‌بندی و احتمال تصادفی هستند.

۴- نتایج آزمایش‌ها

عملکرد روش پیشنهادی تطبیق دامنه متمایزکننده هماوردها برای رده‌بندی تصور گفتار با افزودن ۲ برابری داده با استفاده از شبکه

کلمه با روش افزودن ۲ برابری داده حدود ۱۰ درصد بیشتر از روش بدون افزودن داده است. نتایج برای آزمودنی‌های مختلف نشان‌دهنده آن است که افزودن داده برای تصور کلمات منجر به بهبود عملکرد رده‌بندی با روش یادگیری انتقالی مبتنی بر مدل تطبیق دامنه متمایزکننده همواردانه است.

مقادیر کاپا به‌طور جداگانه برای هر آزمودنی دامنه هدف در مجموعه داده 'KaraOne' برای تصور کلمات در شکل (۵) نمایش داده شده است. همان‌طور که از نمودار قابل مشاهده است مقدار کاپا برای روش پیشنهادی مدل تطبیق دامنه متمایزکننده همواردانه با افزودن ۲ برابری داده مبتنی بر شبکه مولد همواردانه به‌طور چشمگیری بیشتر از روش بدون افزودن داده است. میانگین مقدار کاپا باتوجه به ۱۴ آزمودنی برای روش افزودن ۲ برابری داده برابر با ۰.۸۰ درحالی‌که در روش بدون افزودن داده برابر با ۰.۶۷ با مدل تطبیق دامنه متمایزکننده همواردانه است.

به‌منظور بررسی و ارزیابی مدل پیشنهادی با تعداد بیشتری از تصور گفتار از مجموعه داده 'KaraOne' برای رده‌بندی ۷ واج /iy/, /uw/, /piy/, /tiy/, /diy/, /m/, /n/ استفاده شده است. صحت رده‌بندی و مقدار کاپا تصور واج‌ها با مدل پیشنهادی تطبیق دامنه متمایزکننده همواردانه در حالت بدون افزودن داده و افزودن ۲ برابری داده مبتنی بر شبکه مولد همواردانه در جدول ۵ نشان داده شده است. میانگین صحت رده‌بندی تصور واج‌ها با روش پیشنهادی با افزودن ۲ برابری داده حدود ۶۰٪ است که حدود ۱۰ درصد بیشتر از روش بدون افزودن داده است. نتایج برای آزمودنی‌های مختلف نشان‌دهنده آن است که افزودن داده برای تصور واج‌ها منجر به بهبود عملکرد رده‌بندی با روش تطبیق دامنه متمایزکننده همواردانه است و بیش برآزش مدل را به علت ترکیب سیگنال‌های الکتروانسفالوگرافی آزمودنی‌های دامنه منبع کاهش می‌دهد. این نتایج نشان می‌دهد که روش تطبیق دامنه متمایزکننده همواردانه می‌تواند به‌طور مؤثر سیگنال‌های الکتروانسفالوگرافی تصور گفتار آزمودنی دامنه هدف را بدون نیاز به داده‌های برچسب‌گذاری شده با استفاده از سیگنال آزمودنی‌های دامنه منبع رده‌بندی کند.

مولد همواردانه و بدون افزودن داده با معیارهای صحت رده‌بندی و کاپا ارزیابی و با سایر روش‌های جدید مقایسه شدند. به‌منظور شناسایی تصور گفتار از پایگاه داده KaraOne با ۷ واج و ۴ کلمه و ۱۴ آزمودنی استفاده شده است. به‌منظور بررسی مدل پیشنهادی عملکرد رده‌بندی چند رده‌ای تصور کلمات، تصور واج‌ها و تصور کلمات و واج‌ها گزارش شده است. همچنین از دیدگاه واج‌شناختی پنج رده‌بندی دودویی، مصوت در مقابل صامت (C/V)، وجود واج بینی (±Nasal)، وجود واج دوزبانی (±Bilab)، وجود مصوت جلوی بالا (±iy/) و وجود مصوت پشت بالا (±uw/) با استفاده از مدل پیشنهادی مورد ارزیابی قرار گرفتند.

جدول (۴): صحت رده‌بندی تصور کلمات در مجموعه داده

'KaraOne' با روش پیشنهادی بدون افزودن داده و افزودن ۲ برابری داده در مدل تطبیق دامنه متمایزکننده همواردانه مبتنی بر

شبکه مولد همواردانه

آزمودنی	صحت رده‌بندی تصور کلمات	
	روش پیشنهادی افزودن ۲	روش پیشنهادی بدون افزودن
دامنه	برابری داده با تطبیق دامنه	داده با تطبیق دامنه
هدف		
S0	85.41	79.16
S1	89.58	83.33
S2	93.75	81.25
S3	91.66	81.25
S4	83.33	77.08
S5	83.33	79.16
S6	83.33	81.25
S7	79.16	64.58
S8	89.58	70.83
S9	87.5	85.41
S10	89.58	75.00
S11	89.58	79.16
S12	83.33	72.91
S13	75.00	56.25
میانگین	86.00	76.18

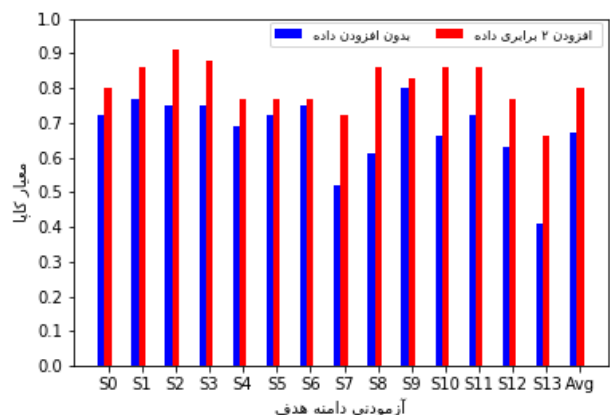
نتایج صحت رده‌بندی تصور کلمات (gnaw, knew, pot, pat) در مجموعه داده 'KaraOne' با روش پیشنهادی مدل تطبیق دامنه متمایزکننده همواردانه با افزودن ۲ برابری داده مبتنی بر شبکه مولد همواردانه و همچنین بدون افزودن داده در جدول ۴ نشان داده شده است. در مدل تطبیق دامنه متمایزکننده همواردانه مبتنی بر شبکه مولد همواردانه هنگامی که یک آزمودنی به‌عنوان دامنه هدف انتخاب شد، بقیه آزمودنی‌ها به‌عنوان دامنه منبع در نظر گرفته می‌شوند. میانگین صحت رده‌بندی برای ۱۴ آزمودنی در تصور ۴

افزودن داده و افزودن ۲ برابری داده مبتنی بر شبکه مولد هماوردانه در جدول ۶ آمده است. میانگین صحت رده‌بندی برای ۱۴ آزمودنی در تصور واج‌ها و کلمات با افزودن ۲ برابری داده و بدون افزودن داده به ترتیب برابر با $51/72\%$ و $44/04\%$ است. همچنین میانگین مقدار کاپا برای ۱۴ آزمودنی با روش پیشنهادی تطبیق دامنه متمایزکننده هماوردانه در حالت بدون افزودن داده و افزودن ۲ برابری داده به ترتیب برابر $0/38$ و $0/46$ است. نتایج نشان می‌دهد که ویژگی‌های متمایز با به‌حداقل‌رساندن تغییر دامنه بین دامنه‌های منبع و هدف استخراج‌شده‌اند که عملکرد رده‌بندی را برای تصور واج‌ها و کلمات در سیگنال‌های الکتروانسفالوگرافی تصور گفتار در کاربرد رابط مغز-رایانه بهبود می‌بخشد.

جدول (۶): صحت رده‌بندی و مقدار کاپا تصور واج‌ها و کلمات در مجموعه داده 'KaraOne' با روش پیشنهادی افزودن ۲ برابری داده در مدل تطبیق دامنه متمایزکننده هماوردانه مبتنی بر شبکه مولد هماوردانه

آزمودنی دامنه هدف	صحت رده‌بندی و مقدار کاپا در تصور ۷ واج و ۴ کلمه			
	مقدار کاپا روش	مقدار کاپا روش	صحت روش	صحت روش
	پیشنهادی افزودن ۲ برابری داده با تطبیق دامنه	پیشنهادی بدون افزودن داده با تطبیق دامنه	پیشنهادی افزودن ۲ برابری داده	پیشنهادی بدون افزودن داده با تطبیق دامنه
S0	0.52	0.44	56.81	50.00
S1	0.42	0.34	47.72	40.15
S2	0.43	0.34	48.48	40.15
S3	0.50	0.43	55.30	48.48
S4	0.48	0.37	53.03	43.18
S5	0.42	0.35	47.72	40.90
S6	0.41	0.34	46.96	40.15
S7	0.52	0.38	56.81	43.93
S8	0.57	0.51	61.36	56.06
S9	0.40	0.36	46.21	42.42
S10	0.46	0.41	51.51	46.96
S11	0.40	0.35	46.21	40.90
S12	0.45	0.34	50.75	40.15
S13	0.50	0.37	55.30	43.18
میانگین	0.46	0.38	51.72	44.04

برای نگاهی دقیق‌تر به نتایج، ماتریس درهم‌ریختگی برای آزمودنی دامنه هدف S2 که بالاترین صحت رده‌بندی در تصور کلمات را داشته در جدول شماره ۷ نشان داده شده است. همان‌طور که قابل مشاهده است کلمات gnaw و pot به‌درستی رده‌بندی شده‌اند. چنین مشاهده‌ای در واقع نشان می‌دهد که افزودن ۲ برابری داده در



شکل (۵): مقایسه معیار کاپا برای آزمودنی‌های دامنه هدف با روش پیشنهادی بدون افزودن داده و افزودن ۲ برابری داده در مدل تطبیق دامنه متمایزکننده هماوردانه مبتنی بر شبکه مولد هماوردانه در مجموعه داده 'KaraOne'

جدول (۵): صحت رده‌بندی و مقدار کاپا تصور واج‌ها در مجموعه داده 'KaraOne' با روش پیشنهادی افزودن ۲ برابری داده در مدل تطبیق دامنه متمایزکننده هماوردانه مبتنی بر شبکه مولد هماوردانه

آزمودنی دامنه هدف	صحت رده‌بندی و مقدار کاپا تصور واج‌ها			
	مقدار کاپا روش	مقدار کاپا روش	صحت روش	صحت روش
	پیشنهادی افزودن ۲ برابری داده با تطبیق دامنه	پیشنهادی بدون افزودن داده با تطبیق دامنه	پیشنهادی افزودن ۲ برابری داده	پیشنهادی بدون افزودن داده با تطبیق دامنه
S0	0.45	0.37	53.57	46.42
S1	0.48	0.36	55.95	45.23
S2	0.68	0.49	72.61	57.14
S3	0.45	0.30	53.71	40.47
S4	0.47	0.36	54.76	45.23
S5	0.61	0.41	66.66	49.99
S6	0.69	0.54	73.80	60.71
S7	0.47	0.43	54.76	51.19
S8	0.48	0.44	55.95	52.38
S9	0.48	0.44	55.95	52.38
S10	0.62	0.34	67.85	44.04
S11	0.55	0.48	61.90	55.95
S12	0.51	0.49	58.33	57.14
S13	0.50	0.44	57.14	52.38
میانگین	0.53	0.42	60.21	50.76

مقدار کاپا و صحت رده‌بندی ۱۱ رده‌ای تصور واج‌ها و کلمات با روش پیشنهادی تطبیق دامنه متمایزکننده هماوردانه در حالت بدون

جدول (۸): ماتریس درهم ریختگی برای بالاترین صحت رده‌بندی واج‌ها با آزمودنی دامنه هدف S6 با مدل پیشنهادی تطبیق دامنه متمایزکننده هماوردانه مبتنی بر شبکه مولد هماوردانه

دقت	/di:/	/pi:/	/n/	/m/	/iy/	/tiy/	/uw/	روش افزودن ۲ برابر ری داده با تطبیق دامنه
83.33 %	0	0	0	1	0	1	1	/u w/
66.66 %	1	0	1	1	0	8	1	/tiy /
66.66 %	1	0	1	2	8	0	0	/iy/
58.33 %	0	1	0	7	2	1	1	/m/
83.33 %	0	0	1	0	1	0	1	/n/
75%	0	9	1	0	0	2	0	/pi y/
83.33 %	1	0	0	1	1	0	0	/di y/
برچسب‌های پیش‌بینی شده								

صحت رده‌بندی تصور واج‌ها از دیدگاه واج‌شناختی با پنج رده‌بندی دودویی، مصوت در مقابل صامت (C/V)، وجود واج بینی (±Nasal) وجود واج دوزبانی (±Bilab)، وجود مصوت جلوی بالا (±iy/) و وجود مصوت پشت بالا (±uw/) با استفاده از مدل پیشنهادی تطبیق دامنه متمایزکننده هماوردانه با افزودن ۲ برابری داده و بدون افزودن داده در جدول ۱۰ مورد ارزیابی قرار گرفتند. طبق نتایج بالاترین میانگین صحت رده‌بندی برای واج /uw/± با افزودن ۲ برابری داده برای ۱۴ آزمودنی دامنه هدف برابر با ۹۰/۸۱٪ است. نتایج برای آزمودنی‌های مختلف نشان‌دهنده آن است که افزودن داده برای تصور واج‌ها منجر به بهبود عملکرد رده‌بندی با روش یادگیری انتقالی مبتنی بر مدل تطبیق دامنه متمایزکننده هماوردانه است. همچنین نتایج نشان می‌دهد که یادگیری متمایزکننده مناسب قابلیت تعمیم رمزگذار هدف را بهبود می‌بخشد و در نتیجه عملکرد رده‌بند را افزایش می‌دهد.

مدل تطبیق دامنه متمایزکننده هماوردانه نقش مهمی برای دستیابی به رده‌بندی دقیق تصور گفتار ایفا می‌کند.

جدول (۷): ماتریس درهم ریختگی برای بالاترین صحت رده‌بندی کلمات با آزمودنی دامنه هدف S2 با مدل پیشنهادی تطبیق دامنه متمایزکننده هماوردانه مبتنی بر شبکه مولد هماوردانه

دقت ^۱	pot	knew	pat	gnaw	روش افزودن ۲ برابری داده با تطبیق دامنه
100%	0	0	0	12	gnaw
83.33%	1	0	10	1	pat
91.66%	1	11	0	0	knew
100%	12	0	0	0	pot
برچسب‌های پیش‌بینی شده					

برای نگاهی دقیق‌تر به نتایج چند رده‌ای برای مدل تطبیق دامنه متمایزکننده هماوردانه، ماتریس درهم‌ریختگی برای تصور ۷ واج برای آزمودنی دامنه هدف S6 در جدول ۸ گزارش شده است. مقادیری که در قطر این ماتریس قرار گرفته‌اند نشان‌دهنده تعداد تصور واجی است که به درستی رده‌بندی شده‌اند. تصور واج‌های /diy/ /n/ /uw/ بیشترین تعداد تصور واج درست را نشان می‌دهد.

به‌منظور بررسی دقیق‌تر، نتایج ۱۱ رده‌ای برای مدل تطبیق دامنه متمایزکننده هماوردانه با افزودن ۲ برابری داده برای آزمودنی‌های دامنه منبع، ماتریس درهم‌ریختگی برای تصور ۷ واج و ۴ کلمه برای آزمودنی دامنه هدف S8 که بالاترین صحت رده‌بندی در تصور واج‌ها و کلمات را داشته در جدول ۹ ارائه شده است. تصور واج /tiy/ بیشترین تعداد تصور واج درست و همچنین کلمات pat و knew بیشترین تعداد تصور کلمه درست را نشان می‌دهند. چنین مشاهده‌ای در واقع نشان می‌دهد که افزودن ۲ برابری داده در مدل تطبیق دامنه متمایزکننده هماوردانه افزودن ۲ برابری داده مبتنی بر شبکه مولد هماوردانه نقش مهمی برای دستیابی به رده‌بندی دقیق تصور گفتار ایفا می‌کند.

¹ Precision

جدول (۹): ماتریس درهم ریختگی برای بالاترین صحت رده‌بندی واج‌ها و کلمات با آزمودنی دامنه هدف S8 با مدل پیشنهادی تطبیق دامنه

متمایزکننده هماوردانه مبتنی بر شبکه مولد هماوردانه

دقت	pot	knew	Pat	gnaw	/diy/	/piy/	/n/	/m/	/iy/	/tiy/	/uw/	روش افزودن ۲ برابری داده با تطبیق دامنه
58.33%	0	0	0	1	0	0	0	2	2	0	7	/uw/
83.33%	1	0	0	0	0	1	0	0	0	10	0	/tiy/
50%	1	0	1	0	0	0	1	1	6	0	2	/iy/
58.33%	0	0	0	0	0	1	0	7	1	1	2	/m/
75%	0	0	0	1	0	1	9	0	0	1	0	/n/
33.33%	1	0	0	0	0	4	0	1	0	3	3	/piy/
66.66%	1	0	1	0	8	0	1	1	0	0	0	/diy/
58.33%	1	0	1	7	0	1	1	0	0	1	0	Gnaw
66.66%	1	1	8	0	0	0	0	0	0	0	2	Pat
66.66%	0	8	0	1	2	0	0	0	0	0	1	Knew
58.33%	7	4	0	0	0	0	0	0	0	0	1	Pot

برچسب‌های پیش‌بینی شده

جدول (۱۰): صحت رده‌بندی تصور واج‌ها با پنج رده‌بندی دودویی در مجموعه داده 'KaraOne' با روش پیشنهادی بدون افزودن داده و افزودن

۲ برابری داده در مدل تطبیق دامنه متمایزکننده هماوردانه مبتنی بر شبکه مولد هماوردانه

آزمودنی دامنه هدف	صحت رده‌بندی تصور واج‌ها از دیدگاه واج‌شناسی									
	C/V		±Nasal		±Bilab		±iy/		±uw/	
	بدون افزودن داده با تطبیق دامنه	افزودن ۲ برابری داده با تطبیق دامنه	بدون افزودن داده با تطبیق دامنه	افزودن ۲ برابری داده با تطبیق دامنه	بدون افزودن داده با تطبیق دامنه	افزودن ۲ برابری داده با تطبیق دامنه	بدون افزودن داده با تطبیق دامنه	افزودن ۲ برابری داده با تطبیق دامنه	بدون افزودن داده با تطبیق دامنه	افزودن ۲ برابری داده با تطبیق دامنه
S0	82.14	88.09	76.19	78.57	74.99	80.95	71.42	88.09	86.90	90.47
S1	80.95	86.90	83.33	86.90	76.19	89.28	69.04	85.71	88.09	89.28
S2	76.19	85.71	72.61	78.57	79.76	89.28	78.57	91.66	91.66	95.23
S3	73.80	91.66	78.57	84.52	84.52	90.47	75.00	88.09	86.90	90.47
S4	79.76	85.71	78.57	79.76	82.14	88.09	79.76	92.85	90.47	92.85
S5	79.76	89.28	76.19	82.14	72.61	79.76	72.61	86.90	85.71	89.28
S6	72.61	88.09	73.80	83.33	79.76	83.33	76.19	91.66	84.52	89.28
S7	80.95	85.71	77.38	83.33	79.76	83.33	72.61	88.09	89.28	92.85
S8	78.57	91.66	80.95	83.33	76.19	82.14	76.19	86.90	83.33	88.09
S9	80.95	89.28	75	85.71	79.76	84.52	80.95	85.71	90.47	91.66
S10	76.19	85.71	79.76	82.14	73.80	80.95	73.80	84.52	85.71	89.28
S11	78.57	89.28	79.76	86.90	78.57	85.71	83.33	88.09	85.71	88.09
S12	77.38	86.90	80.14	82.14	73.80	78.57	75.00	83.33	88.09	91.66
S13	78.57	86.90	82.14	83.33	83.33	88.09	80.95	88.09	88.09	92.85
میانگین	78.31	87.92	78.17	82.90	78.22	84.60	76.10	87.83	87.49	90.81

مقدار کاپا باتوجه به ۱۴ آزمودنی با مدل تطبیق دامنه متمایزکننده هماوردانه و افزودن ۲ برابری داده برای واج /uw/ برابر با ۰/۸۱ است. نتایج برای آزمودنی‌های مختلف نشان‌دهنده آن است که افزودن داده برای تصور واج‌ها منجر به بهبود عملکرد رده‌بندی می‌شود و همچنین نشان می‌دهد که روش تطبیق دامنه متمایزکننده هماوردانه می‌تواند به‌طور مؤثر سیگنال‌های الکتروانسفالوگرافی

مقادیر کاپا به‌طور جداگانه برای هر آزمودنی دامنه هدف در مجموعه داده 'KaraOne' برای پنج رده‌بندی دودویی تصور واج‌ها در جدول ۱۱ آمده است. همان‌طور که قابل‌مشاهده است مقادیر کاپا برای روش پیشنهادی مدل تطبیق دامنه متمایزکننده هماوردانه با افزودن ۲ برابری داده مبتنی بر شبکه مولد هماوردانه به‌طور چشمگیری بیشتر از روش بدون افزودن داده است. میانگین

رده‌بندی تصور گفتار مبتنی بر سیگنال‌های مغزی با افزودن داده‌ها و شبکه تطبیق دامنه هماوردانه

با استفاده از یادگیری عمیق در کاربرد رابط مغز-رایانه

تمتایزکننده با به حداقل رساندن تغییر دامنه بین دامنه منبع و هدف استخراج شده‌اند که عملکرد رده‌بندی برای آزمودنی جدید را بدون نیاز به داده‌های برچسب‌گذاری شده بهبود می‌بخشد.

مقایسه میانگین صحت رده‌بندی تصور ۷ واج و ۴ کلمه برای روش پیشنهادی تطبیق دامنه متمایزکننده هم‌اورده گفتار با افزودن ۲ برابری داده و بدون افزودن داده و روش‌های [۲۴، ۲۳، ۲۱، ۱۷، ۱۶] در شکل ۶ نمایش داده شده است. همان‌طور که از نتایج رده‌بندی ۱۱ رده‌ای قابل مشاهده است روش پیشنهادی تطبیق دامنه متمایزکننده هم‌اورده گفتار با افزودن ۲ برابری داده مبتنی بر شبکه مولد هم‌اورده حدود ۱۰٪ بیشتر از بهترین روش [۱۶] که شامل یک لایه استخراج ویژگی مبتنی بر مدل زبان، یک لایه نگاشت ویژگی صوتی و یک ماشین بولتزمن محدود شده مبتنی بر لایه یادگیری ویژگی است. مدل پیشنهادی تطبیق دامنه متمایزکننده هم‌اورده گفتار طی سه مرحله با استفاده از یادگیری هم‌اورده اختلاف بین توزیع دامنه‌های منبع و هدف را به طور قابل توجهی کاهش می‌دهد.

تصور گفتار آزمودنی دامنه هدف را بدون نیاز به داده‌های برچسب‌گذاری شده با استفاده از سیگنال آزمودنی‌های دامنه منبع رده‌بندی کند.

میانگین صحت رده‌بندی تصور واج‌ها از دیدگاه واج‌شناختی با پنج رده‌بندی دودویی با روش پیشنهادی مدل تطبیق دامنه متمایزکننده هم‌اورده با افزودن ۲ برابری داده مبتنی بر شبکه مولد هم‌اورده و بدون افزودن داده با روش‌های [۲۱، ۲۰، ۱۶، ۱۵] در جدول ۱۲ مقایسه شده است. از پنج رده‌بندی دودویی روش پیشنهادی مدل تطبیق دامنه متمایزکننده هم‌اورده با افزودن ۲ برابری داده میانگین صحت رده‌بندی بالاتری نسبت به روش‌های دیگر دارد در حالی که در رده‌بندی مصوت در مقابل صامت (C/V) روش [۲۱] میانگین صحت رده‌بندی بالاتری دارد. همان‌طور که از نتایج مشخص است، تفاوت ویژگی‌هایی که توسط مدل انطباق دامنه هم‌اورده استخراج شده نسبت به ویژگی‌های روش‌های ذکر شده این که از تطبیق دامنه استفاده شده است. این نشان می‌دهد که ویژگی‌های

جدول (۱۱): مقدار کاپا تصور واج‌ها با پنج رده‌بندی دودویی در مجموعه داده 'KaraOne' با روش پیشنهادی بدون افزودن داده و

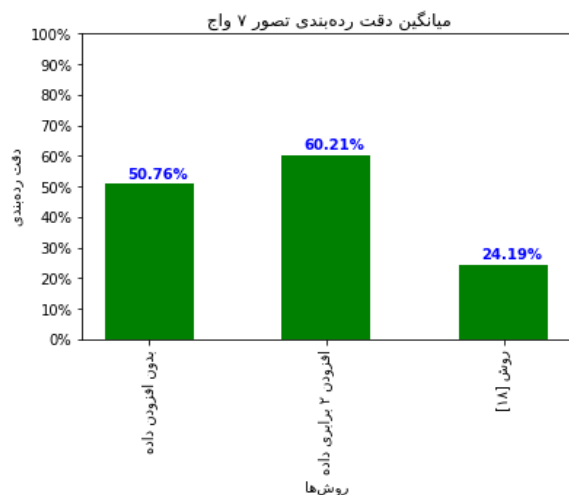
افزودن ۲ برابری داده در مدل تطبیق دامنه متمایزکننده هم‌اورده مبتنی بر شبکه مولد هم‌اورده

آزمودنی دامنه هدف	مقدار کاپا تصور واج‌ها از دیدگاه واج‌شناسی									
	C/V		±Nasal		±Bilab		±/iy/		±/uw/	
	بدون افزودن داده با تطبیق دامنه	افزودن ۲ برابری داده با تطبیق دامنه	بدون افزودن داده با تطبیق دامنه	افزودن ۲ برابری داده با تطبیق دامنه	بدون افزودن داده با تطبیق دامنه	افزودن ۲ برابری داده با تطبیق دامنه	بدون افزودن داده با تطبیق دامنه	افزودن ۲ برابری داده با تطبیق دامنه	بدون افزودن داده با تطبیق دامنه	افزودن ۲ برابری داده با تطبیق دامنه
S0	0.64	0.76	0.52	0.57	0.49	0.61	0.42	0.76	0.73	0.80
S1	0.61	0.73	0.66	0.73	0.52	0.78	0.38	0.71	0.76	0.78
S2	0.52	0.71	0.45	0.57	0.59	0.78	0.57	0.83	0.83	0.90
S3	0.47	0.83	0.57	0.69	0.69	0.80	0.5	0.76	0.73	0.80
S4	0.59	0.71	0.57	0.59	0.64	0.76	0.59	0.85	0.80	0.85
S5	0.59	0.78	0.52	0.64	0.45	0.59	0.45	0.73	0.71	0.78
S6	0.45	0.76	0.47	0.66	0.59	0.66	0.52	0.83	0.69	0.78
S7	0.61	0.71	0.54	0.66	0.59	0.66	0.45	0.76	0.78	0.85
S8	0.57	0.83	0.61	0.66	0.52	0.64	0.52	0.73	0.66	0.76
S9	0.61	0.78	0.5	0.71	0.59	0.69	0.61	0.71	0.80	0.83
S10	0.52	0.71	0.59	0.64	0.47	0.61	0.47	0.69	0.71	0.78
S11	0.57	0.78	0.52	0.73	0.57	0.71	0.66	0.76	0.71	0.76
S12	0.54	0.73	0.61	0.64	0.47	0.57	0.5	0.66	0.76	0.83
S13	0.57	0.73	0.54	0.66	0.66	0.76	0.61	0.76	0.76	0.85
میانگین	0.56	0.75	0.54	0.65	0.56	0.68	0.51	0.75	0.74	0.81

رده‌بندی تصور گفتار مبتنی بر سیگنال‌های مغزی با افزودن داده‌ها و شبکه تطبیق دامنه هم‌اورده

با استفاده از یادگیری عمیق در کاربرد رابط مغز-رایانه

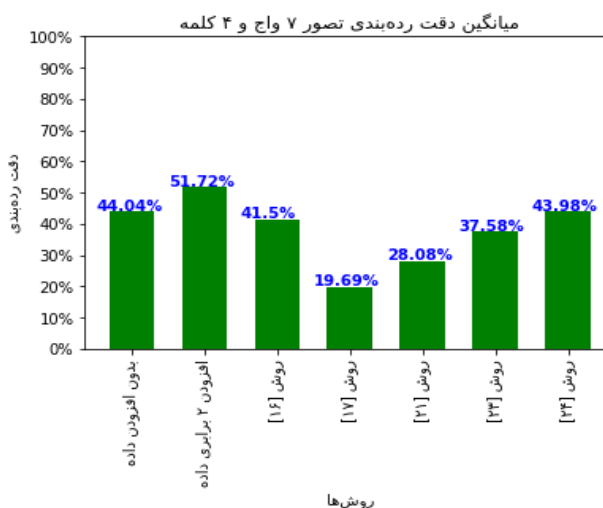
کپستروال بسامد مل و شبکه عصبی پیچشی به ترتیب به عنوان استخراج ویژگی و رده‌بند استفاده شده است. همچنین روش افزودن داده حدود ۱۰٪ بیشتر از روش بدون افزودن داده است. این نشان می‌دهد که یادگیری متمایزکننده مناسب قابلیت تعمیم رمزگذار هدف را بهبود می‌بخشد و در نتیجه عملکرد رده‌بند را افزایش می‌دهد.



شکل (۷): مقایسه میانگین صحت رده‌بندی تصور واج‌ها با روش پیشنهادی بدون افزودن داده و افزودن ۲ برابری داده در مدل تطبیق دامنه متمایزکننده هم‌آوردانه مبتنی بر شبکه مولد هم‌آوردانه با روش‌های دیگران در مجموعه داده 'KaraOne' مقایسه میانگین صحت رده‌بندی تصور ۴ کلمه برای روش پیشنهادی و روش [۱] در شکل ۸ آمده است. در [۱] نمودارهای توپولوژیکی سیگنال‌های الکتروانسفالوگرافی با میانگین زمانی در تمام آزمایش‌های مربوط به تصور یک کلمه خاص به عنوان ورودی به شبکه عصبی عمیق تغذیه می‌شوند. معماری یادگیری عمیق طراحی شده دارای ادغام شبکه عصبی پیچشی دوبعدی و شبکه حافظه بلند-کوتاه مدت است که از قابلیت‌های هر دو معماری شبکه عصبی کمک می‌گیرد. میانگین صحت روش پیشنهادی تطبیق دامنه متمایزکننده هم‌آوردانه گفتار با افزودن ۲ برابری داده حدود ۴۳٪ بیشتر است؛ بنابراین افزودن داده مبتنی بر روش‌های مولدی همراه با روش تطبیق دامنه‌های منبع و هدف عملکرد رده‌بندی تصور گفتار را به طور قابل توجهی بهبود می‌بخشد.

جدول (۱۲): مقایسه میانگین صحت رده‌بندی تصور واج‌ها با پنج رده‌بندی دودویی با روش پیشنهادی بدون افزودن داده و افزودن ۲ برابری داده در مدل تطبیق دامنه متمایزکننده هم‌آوردانه مبتنی بر شبکه مولد هم‌آوردانه با روش‌های دیگران در مجموعه داده

روش‌های پیشنهادی	مقایسه میانگین صحت رده‌بندی تصور واج‌ها از دیدگاه واج‌شناسی				
	C/V	$\pm Nasa_1$	$\pm Bila_b$	$\pm iy/$	$\pm uw/$
بدون افزودن داده با تطبیق دامنه	78.31	78.17	78.22	76.10	87.49
افزودن ۲ برابری داده با تطبیق دامنه	87.92	82.90	84.60	87.83	90.81
[15]	18.08	63.50	56.64	59.60	79.16
[16]	41	74	71	76	87
[20]	85.23	73.45	75.55	73.30	81.99
[21]	89.16	78.33	81.67	87.20	85

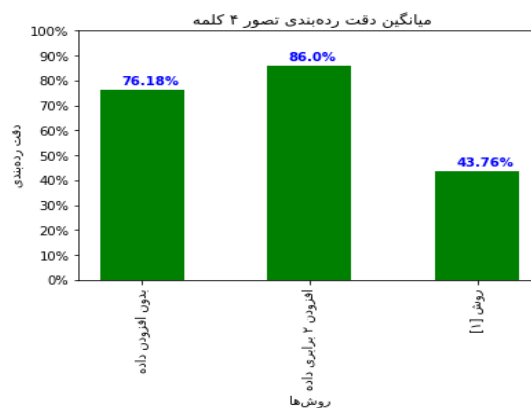


شکل (۶): مقایسه میانگین صحت رده‌بندی تصور کلمات و واج‌ها با روش پیشنهادی بدون افزودن داده و افزودن ۲ برابری داده در مدل تطبیق دامنه متمایزکننده هم‌آوردانه مبتنی بر شبکه مولد هم‌آوردانه با روش‌های دیگران در مجموعه داده 'KaraOne' مقایسه میانگین صحت رده‌بندی تصور ۷ واج برای روش پیشنهادی و روش [۱۸] در شکل ۷ آمده است. میانگین صحت روش پیشنهادی تطبیق دامنه متمایزکننده هم‌آوردانه گفتار با افزودن ۲ برابری داده حدود ۳۶٪ بیشتر است از روش [۱۸] که از ضرایب

۵- نتیجه‌گیری

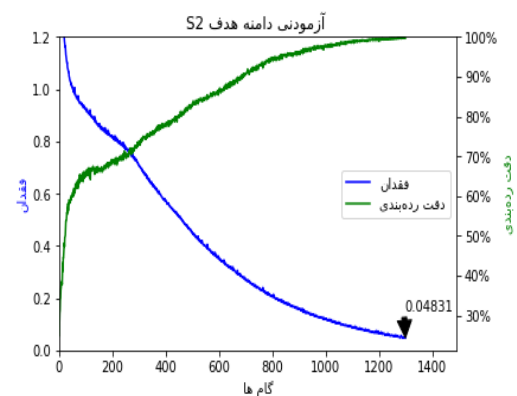
در این مقاله مدل تطبیق دامنه متمایزکننده هماوردانه برای رده‌بندی سیگنال‌های مغزی تصور گفتار با افزودن داده‌ها مبتنی بر مدل مولد هماوردانه برای کاربرد رابط مغز-رایانه ارائه شد. این مطالعه از یادگیری انتقالی با استفاده از روش تطبیق دامنه هماوردانه برای ارائه مدلی به‌منظور رده‌بندی سیگنال‌های الکتروانسفالوگراف آزمودنی جدید به‌عنوان دامنه هدف با برچسب داده‌های سایر آزمودنی‌ها به‌عنوان دامنه منبع برای تصور گفتار با به‌حداقل رساندن اختلاف توزیع بین دامنه‌های منبع و هدف از طریق یادگیری هماوردانه استفاده کرد. مدل پیشنهادی شامل سه جزء است: استخراج‌کننده ویژگی، متمایزکننده دامنه و رده‌بند. استخراج‌کننده ویژگی برای سیگنال‌های الکتروانسفالوگرافی تصور گفتار با استفاده از شبکه عصبی پیچشی پیاده‌سازی شده است. شبکه متمایزکننده دامنه با استخراج‌کننده ویژگی به‌عنوان رمزگذار هدف از طریق یادگیری هماوردانه آموزش می‌بیند. درنهایت، رده‌بند آموزش‌دیده شده با داده‌های دامنه منبع، به‌صحت برچسب داده‌های دامنه هدف را بدون استفاده از داده‌های برچسب‌گذاری شده پیش‌بینی می‌کند. به‌منظور شناسایی تصور گفتار از پایگاه داده KaraOne با ۷ واج و ۴ کلمه استفاده شده است. روش پیشنهادی با سایر روش‌های جدید بر اساس معیارهای صحت و کاپا مورد ارزیابی قرار گرفتند. طبق نتایج به‌دست‌آمده روش پیشنهادی تطبیق دامنه متمایزکننده هماوردانه گفتار با افزودن ۲ برابری داده با صحت ۸۶٪ و ۶۰/۲۱٪ به ترتیب تصور کلمات و واج‌ها را رده‌بندی می‌کند. مدل پیشنهادی، مستقل از سیگنال‌های مغزی هر فرد است که با آموزش مدل بر روی سیگنال‌های مغزی افزوده‌شده آزمودنی‌ها در دامنه منبع می‌توان به‌طور مؤثر سیگنال‌ها را در دامنه هدف بدون نیاز به داده‌های برچسب‌گذاری شده از آزمودنی جدید رده‌بندی کرد. ازاین‌رو، مدل پیشنهادی صحت رده‌بندی را به میزان قابل‌توجهی بهبود می‌بخشد و زمان به‌دست‌آوردن سیگنال‌های مغزی برای آزمودنی جدید را کاهش می‌دهد که برای کاربردهای رابط مغز-رایانه در دنیای واقعی می‌تواند مورد توجه باشد.

نمودار فقدان و صحت رده‌بندی آزمودنی دامنه هدف S2 که بالاترین صحت رده‌بندی در تصور کلمات را داشته در شکل ۹ نمایش داده‌شده است. همان‌طور که قابل‌مشاهده است بعد از حدود ۱۳۰۰ گام آموزش فقدان رده‌بند با داده‌های دامنه‌های منبع برابر با ۰/۰۴ است درحالی‌که صحت رده‌بندی حدود ۱۰۰٪ است. این نشان می‌دهد که ویژگی‌های متمایزکننده و کارایی با استفاده از روش تطبیق دامنه متمایزکننده هماوردانه استخراج‌شده که با استفاده از یادگیری هماوردانه اختلاف بین توزیع دامنه‌های منبع و هدف را کاهش می‌دهد.



شکل (۸): مقایسه میانگین صحت رده‌بندی تصور کلمات با روش پیشنهادی بدون افزودن داده و افزودن ۲ برابری داده در مدل تطبیق دامنه متمایزکننده هماوردانه مبتنی بر شبکه مولد هماوردانه با

روش‌های دیگران در مجموعه داده 'KaraOne'



شکل (۹): فقدان و صحت رده‌بندی آزمودنی دامنه هدف S2 برای تصور کلمات با مدل پیشنهادی تطبیق دامنه متمایزکننده هماوردانه با افزودن ۲ برابری داده مبتنی بر شبکه مولد هماوردانه در مجموعه داده

'KaraOne'

در آینده، چالش رده‌بندی تصور گفتار برای کاربردهای رابط مغز-رایانه با استفاده از مدل تعمیم دامنه هم‌اوردانه بدون نیاز به سیگنال‌های الکتروانسفالوگراف آزمودنی جدید ارائه خواهد شد.

References

- [1] M. Bisla and R. Anand, "EEG Based Brain Computer Interface System for Decoding Covert Speech using Deep Neural Networks," in *2023 IEEE 12th International Conference on Communication Systems and Network Technologies (CSNT)*, Bhopal, India, 2023.
- [2] C. S. DaSalla, H. Kambara, M. Sato and Y. Koike, "Single-trial classification of vowel speech imagery using common spatial patterns," *Neural Networks*, vol. 22, no. 9, pp. 1334-1339, 2009.
- [3] K. Brigham and B. V. Kumar, "Imagined Speech Classification with EEG Signals for Silent Communication: A Preliminary Investigation into Synthetic Telepathy," in *2010 4th International Conference on Bioinformatics and Biomedical Engineering*, Chengdu, China, 2010.
- [4] S. Deng, R. Srinivasan, T. Lappas and M. D'Zmura, "EEG classification of imagined syllable rhythm using Hilbert spectrum methods," *Journal of neural engineering*, vol. 7, no. 4, 2010.
- [5] X. Chi, J. B. Hagedorn, D. Schoonover and M. D'Zmura, "EEG-Based Discrimination of Imagined Speech Phonemes," *International Journal of Bioelectromagnetism*, vol. 13, no. 4, pp. 201-206, 2011.
- [6] L. Wang, X. Zhang, X. Zhong and Y. Zhang, "Analysis and classification of speech imagery EEG for BCI," *Biomedical Signal Processing and Control*, vol. 8, no. 6, pp. 901-908, 2013.
- [7] M. Matsumoto and J. Hori, "Classification of silent speech using support vector machine and relevance vector machine," *Applied Soft Computing*, vol. 20, pp. 95-102, 2014.
- [8] S. Iqbal, Y. U. Khan and O. Farooq, "EEG based classification of imagined vowel sounds," in *2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom)*, New Delhi, India, 2015.
- [9] C. H. Nguyen, G. K. Karavas and P. Artemiadis, "Inferring imagined speech using EEG signals: a new approach using Riemannian manifold features," *Journal of Neural Engineering*, vol. 15, 2017.
- [10] Balaji, A. Haldar, K. Patil, T. S. Ruthvik, V. CA, M. Jartarkar and V. Baths, "EEG-based classification of bilingual unspoken speech using ANN," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Jeju, Korea (South), 2017.
- [11] R. Sereshkeh, R. Trott, A. Bricout and T. Chau, "EEG Classification of Covert Speech Using Regularized Neural Networks," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 12, pp. 2292 - 2300, 2017.
- [12] R. Sereshkeh, R. Trott, A. Bricout and T. Chau, "Online EEG classification of covert speech for brain-computer interfacing," *International journal of neural systems*, vol. 27, no. 8, p. 1750033, 2017.
- [13] P. Saha and S. Fels, "Hierarchical Deep Feature Learning For Decoding Imagined Speech From EEG," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019.
- [14] J. T. Panachakel, A. Ramakrishnan and T. Ananthapadmanabha, "A novel deep learning architecture for decoding imagined speech from eeg," *arXiv preprint arXiv:2003.09374*, 2020.
- [15] S. Zhao and F. Rudzicz, "Classifying phonological categories in imagined and articulated speech," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, South Brisbane, QLD, Australia, 2015.
- [16] P. Sun and J. Qin, "Neural Networks based EEG-Speech Models," *arXiv preprint arXiv:1612.05369*, 2016.
- [17] Cooney, R. Folli and D. Coyle, "Mel Frequency Cepstral Coefficients Enhance Imagined Speech Decoding Accuracy from EEG," in *2018 29th Irish Signals and Systems Conference (ISSC)*, Belfast, UK, 2018.
- [18] A.-L. Rusnac and O. Grigore, "Generalized Brain Computer Interface System for EEG Imaginary Speech Recognition," in *2020 24th International Conference on Circuits, Systems,*

- Communications and Computers (CSCC)*, Chania, Greece, 2020.
- [19] A.-L. Rusnac and O. Grigore, "Convolutional Neural Network applied in EEG imagined phoneme recognition system," in *2021 12th International Symposium on Advanced Topics in Electrical Engineering (ATEE)*, Bucharest, Romania, 2021.
- [20] P. Saha, S. Fels and M. Abdul-Mageed, "Deep learning the EEG manifold for phonological categorization from active thoughts," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brighton, UK, 2019.
- [21] P. Saha, M. Abdul-Mageed and S. Fels, "SPEAK YOUR MIND! Towards Imagined Speech Recognition With Hierarchical Deep Learning," *arXiv:1904.05746*, 2019.
- [22] J. T. Panachakel, A. G. Ramakrishnan and T. V. Ananthapadmanabha, "Decoding imagined speech using wavelet features and deep neural networks," in *2019 IEEE 16th India Council International Conference (INDICON)*, Rajkot, India, 2019.
- [23] A.-L. Rusnac and O. Grigore, "CNN Architectures and Feature Extraction Methods for EEG Imaginary Speech Recognition," *Sensors*, vol. 22, no. 13, p. 4679, 2022.
- [24] A.-L. Rusnac and O. Grigore, "Imaginary Speech Recognition Using a Convolutional Network with Long-Short Memory," *Applied Sciences*, vol. 12, no. 22, p. 11873, 2022.
- [25] Q. Heting and G. Nuo, "Research on the Classification Algorithm of Imaginary Speech EEG Signals Based on Twin Neural Network," in *2022 7th International Conference on Signal and Image Processing (ICSIP)*, Suzhou, China, 2022.
- [26] V. S. Rathod, A. Tiwari and O. G. Kakde, "Wading corvus optimization based text generation using deep CNN and BiLSTM classifiers," *Biomedical Signal Processing and Control*, vol. 78, p. 103969, 2022.
- [27] Alizadeh and H. Omranpour, "EM-CSP: An efficient multiclass common spatial pattern feature method for speech imagery EEG signals recognition," *Biomedical Signal Processing and Control*, vol. 84, p. 104933, 2023.
- [28] K. Zhang, G. Xu, Z. Han, K. Ma, X. Zheng, L. Chen, N. Duan and S. Zhang, "Data Augmentation for Motor Imagery Signal Classification Based on a Hybrid Neural Network," *Sensors*, vol. 20, no. 16, p. 4485, 2020.
- [29] J. Xu, H. Zheng, J. Wang, D. Li and X. Fang, "Recognition of EEG Signal Motor Imagery Intention Based on Deep Multi-View Feature Learning," *Sensors*, vol. 20, no. 12, p. 3496, 2020.

Speech Imagery Classification based on Brain Signals with Data Augmentation and Adversarial Domain Adaptation Network using Deep Learning in Brain-Computer Interface Application

Marzieh Anjerani¹, Mir Mohsen Pedram^{2*}, Mitra Mirzarezaee³

¹Ph.D. Student, Department of Computer Engineering, SR.C., Islamic Azad University, Tehran, Iran

²Associate Professor, Department of Electrical and Computer Engineering, Faculty of Engineering, Kharazmi University, Tehran, Iran

³Associate Professor, Department of Computer Engineering, SR.C., Islamic Azad University, Tehran, Iran

Article Information

Original Research Paper

Received:

2025 February 28

Accepted:

2025 April 20

Keywords:

Speech Imagery, Data Augmentation, Adversarial Discriminative Domain Adaptation, Deep Learning, Transfer Learning, Convolutional Neural Network, Brain-Computer Interface

Corresponding Author*:

pedram@khu.ac.ir

Abstract

Speech-based brain-computer interfaces provide effective voice communication strategies for controlling devices through spoken commands interpreted from brain signals. One of the major challenges in the brain-computer interface problem is the classification of brain signals based on electroencephalography. Electroencephalography is a non-invasive brain signal that is recorded from the scalp surface through electrodes. The signals obtained using easy and cheap equipment have relatively low spatial resolution and high temporal resolution, which requires the most appropriate feature extraction and classification method to achieve optimal results. Also, Collecting sufficient data for a new subject requires a lot of time and effort, so in this paper, data augmentation using a generative adversarial model is proposed to improve the performance of brain signal classification. Also, a model for classifying speech imagery based on brain signals of a new subject with transfer learning using a generative adversarial method based on adversarial domain adaptation method is presented. In order to identify speech imagery, the KaraOne database was used. The proposed method was evaluated with other new methods based on accuracy and kappa criteria. According to the results, the proposed method classifies word imagery and phonemes with 86% and 60.21% accuracy, respectively. The proposed model is independent of each individual's brain signals, which can be effectively classified in the target domain by training the model on the augmented brain signals of the subjects in the source domain without the need for labeled data from the new subject.

 : 10.22034/ABMIR.2025.22852.1105

E-ISSN: [2821-2037](#) /© 2024. Published by Yazd University This is an open access article under the CC BY 4.0 License (<https://creativecommons.org/licenses/by/4.0/>).

