





دانشکده مهندسی کامپیوتر

پایان نامه کارشناسی ارشد مهندسی هوش مصنوعی

استفاده از یادگیری عمیق برای طبقه‌بندی داده‌های EEG-fNIRS همزمان

نگارنده: حمیدرضا غنچی

اساتید راهنما

دکتر منصور فاتح

دکتر وحید ابوالقاسمی

اساتید مشاور

دکتر سعیده فردوسی

دکتر محسن رضوانی

تیرماه ۹۸

شماره: ۶۳۸، رف.ف

تاریخ: ۹۸، ۴، ۱۵

باسمه تعالی



وزارت آموزش عالی

مدیریت تحصیلات تکمیلی

فرم شماره (۳) صورتجلسه نهایی دفاع از پایان نامه دوره کارشناسی ارشد

با نام و یاد خداوند متعال، ارزیابی جلسه دفاع از پایان نامه کارشناسی ارشد خانم / آقای
با شماره دانشجویی رشته گرایش
تحت عنوان
که در تاریخ با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به
شرح ذیل اعلام می گردد:

☐ مردود ☒ قبول (با درجه: عالی) ☐ عملی ☒ نظری			
نوع تحقیق:			
امضاء	مرتبه علمی	نام و نام خانوادگی	عضو هیأت داوران
	استاد	مهندس ...	۱- استاد راهنمای اول
	استاد	فرید احمدی	۲- استاد راهنمای دوم
	رئیس دایره استاد	مهندس ...	۳- استاد مشاور
	استاد	علیه قاسمی	۴- نماینده تحصیلات تکمیلی
	استاد	مرتضی ...	۵- استاد ممتحن اول
	استاد	اسماعیل ...	۶- استاد ممتحن دوم

نام و نام خانوادگی رئیس دانشکده:

تاریخ و امضاء و مهر دانشکده:

تبصره: در صورتی که کسی مردود شود حداکثر یکبار دیگر (در مدت مجاز تحصیل) می تواند از پایان نامه خود دفاع نماید (دفاع

مجدد نباید زودتر از ۴ ماه برگزار شود).

ماحصل آموخته‌هایم را تقدیم می‌کنم به آنان که مهر آسمانشان آرام بخش آلام زمینی ام است.

به استوارترین تکیه‌گاهم، دستان پر مهر پدرم

به سبزترین محاکه زندگی ام، چشمان سبز مادرم

که هرچه آموختم در کتب عشق شما آموختم و هرچه بگویم، قطره‌ای از دریای بی‌کران مهربانیتان را
پاس نتوانم بگویم.

امروز، هستی من به امید شماست و فردا کلید باغ بهشت، رضای شما.

باشد که حاصل تلاشم، نسیم کوزه غبار خستگیان را بزداید.

بوسه بر دستان پر مهرتان.

پاس خدای عالمیان که اوست هدایت کر بی همتا

تشکر فراوان از

آقای دکتر فاتح، آقای دکتر رضوانی و بویره آقای دکتر ابوالقاسمی و خانم دکتر فردوسی

به خاطر زحمات بی دریغشان

همچنین تشکر ویژه از جناب آقای دکتر پویان که آغاز کرد و همراه همیشگی من در راستای رسیدن من به اهدافم

بودند

و

تمامی آن هایی که مرا راهنمای زندگی بوده اند.

حمیدرضا غنچه

تیرماه ۹۸

تهدنامه

اینجانب حمیدرضا غنچی دانشجوی دوره کارشناسی ارشد رشته کامپیوتر گرایش هوش مصنوعی دانشکده کامپیوتر دانشگاه صنعتی شاهرود نویسنده پایان نامه "استفاده از یادگیری عمیق برای طبقه بندی داده های EEG-fNIRS همزمان" تحت راهنمایی دکتر منصور فاتح، دکتر وحید ابوالقاسمی، دکتر محسن رضوانی و دکتر سعیده فردوسی متعهد می شوم.

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است .
- در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است .
- مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است .
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « Shahrood University of Technology » به چاپ خواهد رسید .
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می گردد.
- در کلیه مراحل انجام این پایان نامه ، در مواردی که از موجود زنده (یا بافتهای آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است .
- در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری ، ضوابط و اصول اخلاق انسانی رعایت شده است .

تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج ، کتاب ، برنامه های رایانه ای ، نرم افزار ها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می باشد . این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود . استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی باشد.

چکیده

رابط مغز کامپیوتر (BCI)، یک سیستم قدرتمند برای ارتباط بین مغز و جهان خارج است. سیستم‌های BCI سنتی تنها مبتنی بر EEG کار می‌کردند. این سیگنال‌ها به دلیل مزایایی که دارند، تحلیل و ثبت فعالیت ذهنی را آسان می‌کنند. اخیراً، محققان، سیگنال‌های EEG را به منظور بهبود عملکرد این سیستم‌ها با دیگر سیگنال‌ها ترکیب کرده‌اند. در میان این سیگنال‌ها، ترکیب سیگنال‌های EEG و fNIRS نتایج مطلوبی را بدست آورده است. در اکثر مطالعات، تنها از سیگنال‌های EEG یا fNIRS به صورت دنباله‌ای از سیگنال‌های زنجیره مانند استفاده شده است و همبستگی پیچیده‌ی بین سیگنال‌های مجاور یعنی زمان و موقعیت الکترودها در نظر گرفته نشده است. از طرف دیگر، در این راه، چالش‌هایی وجود دارند که باید رفع شوند. روش پیشنهادی این پایان‌نامه، متمرکز بر دو بخش پیش‌پردازش جهت آماده‌سازی داده‌ها و طبقه‌بندی است. پیش‌پردازش پیشنهادی، ارتباط مکانی بین سیگنال‌های EEG و fNIRS را ترکیب می‌کند. این امر به کمک موقعیت الکترودها بر روی پوست سر انجام شده است. در نتیجه، می‌توان دنباله‌هایی از این سیگنال‌های زنجیره مانند را به تانسورهای سه بعدی تبدیل کرد. در ادامه، در این پایان‌نامه، مدلی مبتنی بر شبکه عصبی عمیق، با ترکیب دو شبکه عصبی CNN و RNN، برای تشخیص دقیق اهداف ذهنی انسان و با استفاده از ویژگی‌های زمانی و مکانی ارائه شده است. نتایج حاصل از روش پیش‌پردازش پیشنهادی نشان دادند که سیگنال‌های خام، تبدیل به تصاویری با اطلاعات بیشتر می‌شوند. این تصاویر، اطلاعات مکانی سیگنال‌ها را در هر لحظه زمانی در خود دارند. در ادامه نشان داده شد که که مدل پیشنهادی دارای دقت ۹۹,۶ درصد است. این دقت نسبت به بالاترین دقت بدست آمده در مراجع دیگر، افزایش یک درصدی داشته است که نشان‌دهنده اهمیت ترکیب این دو سیگنال با یکدیگر است. از طرف دیگر، نتایج نشان دادند که پیش‌پردازش پیشنهادی و تبدیل سیگنال‌های خام به تصاویری با اطلاعات مکانی، تاثیر مطلوبی بر شبکه عصبی عمیق دارند.

کلمات کلیدی: EEG، fNIRS، یادگیری عمیق، ویژگی مکانی، ویژگی زمانی

لیست مقالات مستخرج از پایان نامه

۱- تشخیص اهداف ذهنی با استفاده از ویژگی های زمانی-مکانی از طریق یک شبکه عمیق recurrent-convolutional، بیست و پنجمین کنفرانس ملی و سومین کنفرانس بین المللی مهندسی زیست پزشکی ایران، قم، دانشگاه شهاب دانش.

فهرست مطالب

ک	فهرست شکل‌ها	۱
ل	فهرست جداول	۱
۱	فصل ۱: مقدمه	۱
۲	۱-۱ پیش‌گفتار	۲
۴	۲-۱ اندازه‌گیری فعالیت ذهنی	۴
۶	۳-۱ اندازه‌گیری فعالیت ذهنی با استفاده از EEG و fNIRS	۶
۶	۳-۱-۱ فعالیت‌های مغزی استفاده شده در سیستم‌های BCI مبتنی بر EEG	۶
۸	۳-۱-۲ فعالیت‌های مغزی استفاده شده در سیستم‌های BCI مبتنی بر fNIRS	۸
۸	۴-۱ هدف پایان‌نامه	۸
۹	۵-۱ چالش‌ها و راه‌حل‌های آن‌ها	۹
۱۲	۶-۱ نوآوری	۱۲
۱۲	۷-۱ ساختار پایان‌نامه	۱۲
۱۳	فصل ۲: کارهای پیشین	۱۳
۱۴	۱-۲ پیش‌پردازش: نیازمندی‌های بهبود سیگنال همراه با کاهش نویز	۱۴
۱۵	۱-۲-۱ روش مرجع‌دهی	۱۵
۱۵	۱-۲-۲ تجزیه و تحلیل اجزای اصلی	۱۵
۱۶	۱-۲-۳ تجزیه و تحلیل اجزای مستقل	۱۶
۱۶	۲-۲ استخراج ویژگی	۱۶
۱۶	۲-۲-۱ ویژگی‌های توان باند	۱۶
۱۷	۲-۲-۲ ویژگی‌های چگالی طیف توان	۱۷
۱۷	۲-۲-۳ روش زمان-فرکانس	۱۷
۱۷	۳-۲ طبقه‌بندی	۱۷
۱۸	۱-۳-۲ تجزیه و تحلیل تفکیک خطی	۱۸
۱۸	۲-۳-۲ ماشین بردار پشتیبان	۱۸
۱۸	۳-۳-۲ یادگیری عمیق	۱۸
۲۳	۴-۲ کارهای مرتبط	۲۳
۲۳	۱-۴-۲ ساختارهای شامل الگوریتم‌های طبقه‌بندی سنتی	۲۳

۲۶.....	۲-۴-۲ ساختارهای شامل الگوریتم‌های یادگیری عمیق
۳۲.....	۳-۴-۲ ساختارهای شامل یادگیری عمیق به همراه استفاده از ویژگی‌های زمانی و مکانی
۳۴.....	۵-۲ نتیجه‌گیری
۳۹.....	فصل ۳: ارائه روش پیشنهادی
۴۰.....	۱-۳ مقدمه
۴۰.....	۲-۳ پیش‌پردازش: آماده‌سازی دادگان برای تحلیل‌های بعدی
۴۱.....	۳-۳ آماده‌سازی داده‌ها
۴۲.....	۱-۳-۳ تبدیل سیگنال‌های یک بعدی به تانسورهای سه بعدی
۴۸.....	۲-۳-۳ پنجره‌گذاری لغزان
۴۹.....	۳-۳-۳ افزودن ویژگی زمانی به کمک روش پنجره‌گذاری لغزان
۵۱.....	۴-۳ مدل RCNN پیشنهادی
۵۵.....	فصل ۴: نتایج و آزمایش‌ها
۵۶.....	۱-۴ مقدمه
۵۶.....	۲-۴ مجموعه داده hybrid EEG-fNIRS استفاده شده در این پایان‌نامه
۵۸.....	۳-۴ خلاصه آزمایش‌ها
۵۹.....	۴-۴ نتایج مرتبط با پیش‌پردازش داده‌ها
۶۴.....	۵-۴ نتایج طبقه‌بندی مدل
۷۰.....	۶-۴ ارزیابی مدل پیشنهادی
۷۵.....	۷-۴ نتیجه‌گیری
۷۷.....	فصل ۵: نتیجه‌گیری
۷۸.....	۱-۵ جمع‌بندی و نتیجه‌گیری
۷۹.....	۲-۵ پیشنهاد برای کارهای آتی
۸۰.....	مراجع
۸۸.....	لغت‌نامه
۸۹.....	بر اساس حروف الفبای فارسی
۹۳.....	بر اساس حروف الفبای انگلیسی

فهرست شکل‌ها

شکل (۱-۱) مراحل سیستم BCI	۳
شکل (۱-۲) ساختار شبکه عصبی کانولوشن	۲۲
شکل (۱-۳) مثالی از ۲۰ کانال یک سیگنال EEG خام	۴۵
شکل (۲-۳) موقعیت الکترودهای EEG در سیستم ۵-۱۰ بین المللی	۴۷
شکل (۳-۳) فرآیند پنجره‌گذاری لغزان روی تانسورهای سه بعدی	۵۰
شکل (۴-۳) مرحله ثبت و پیش‌پردازش سیگنالهای EEG و fNIRS	۵۱
شکل (۵-۳) ساختار RCNN پیشنهادی	۵۲
شکل (۱-۴) نتیجه حاصل از روش پیش‌پردازش پیشنهادی	۶۲
شکل (۲-۴) بعد از استفاده از روش پنجره‌گذاری لغزان	۶۳
شکل (۳-۴) دقت الگوریتم پیشنهادی روی دادگان مختلف	۶۸
شکل (۴-۴) نمودار دقت بر اساس تعداد مراحل آموزش	۶۹
شکل (۵-۴) شکل کلی ماتریس درهم‌ریختگی برای یک مسئله دو کلاسه	۷۱
شکل (۶-۴) ماتریس درهم‌ریختگی برای مدل پیشنهادی	۷۲
شکل (۷-۴) نمودار ROC مرتبط با چهار فعالیت ذهنی مختلف	۷۵

فهرست جداول

جدول ۱-۱) مشخصات روش‌های ثبت سیگنال‌های مختلف [۱۹].....	۵
جدول ۱-۲) مقایسه روش‌های ذکر شده.....	۳۴
جدول ۱-۳) شرح پارامترها.....	۴۳
جدول ۲-۳) جزئیات ساختار مدل RCNN پیشنهادی استفاده شده در این پایان‌نامه.....	۵۴
جدول ۱-۴) مقایسه مدل‌ها.....	۶۵
جدول ۲-۴) دقت شبکه با توجه به اندازه مختلف پنجره لغزان.....	۷۰
جدول ۳-۴) چهار معیار صحت، حساسیت، F1 و AUC برای مدل پیشنهادی.....	۷۳

فصل ۱: مقدمه

۱-۱ پیش‌گفتار

مغز انسان، محتوای حدود ۱۰۰ میلیارد نورون را آنالیز می‌کند [۴-۱]. نخاع نیز به عنوان یک کابل واسط، اطلاعات دریافتی از اندام‌ها و اعصاب را به مغز منتقل می‌کند. این اطلاعات جهت کنترل بخش‌های مختلف بدن و حرکت آن‌ها استفاده می‌شود. افراد با صدمه‌های نخاعی، همچنان قادر به تولید سیگنال‌هایی در مغز هستند، اما این سیگنال‌ها در قسمت‌های مختلف بدن در دسترس نیستند. دلیل این رخداد، آسیب دیدگی کابل نخاعی میانی است. فناوری‌های متعددی به افراد آسیب دیده کمک می‌کنند تا یک دستگاه رباتیک یا یک ویلچر را کنترل کنند [۵-۹]. این فناوری‌ها معمولاً از حرکت سر، جهت نگاه چشم و غیره استفاده می‌کنند. به هر حال، این روش‌ها، وابسته به کنترل یکی از عضلات بدن توسط بیمار هستند. اما ارائه فناوری‌های موثر برای افراد بدون کنترل ارادی عضلات، سخت‌تر است [۱۰، ۱۱]. این‌گونه افراد در صحبت کردن و حرکت ناتوان هستند، اما می‌توانند فکر کنند و تصمیم بگیرند. از جمله بیماری‌هایی که می‌توانند منجر به فلج حرکتی شدید شوند، سکته مغزی، فلج مغزی شدید، بیماری نورون‌های حرکتی^۱، اسکروز جانبی آمیوتروفیک^۲ و التهاب مغز یا آنسفالیت^۳ هستند [۱۲]. با تکنولوژی رابط مغز-کامپیوتر^۴، می‌توان برخی افراد با مشکلات کنترل ارادی عضلات را به استقلال بیشتری در کارها رساند.

یک رابط مغز-کامپیوتر، یک سیستم ارتباطی است که به مسیرهای خروجی طبیعی مغز توسط اعصاب و عضلات بستگی ندارد [۱۳]. به بیان دیگر، BCI یک سیستم هوش مصنوعی است که می‌تواند مجموعه خاصی از الگوهای تولید شده توسط مغز را تشخیص دهد [۱۴]. سیستم BCI می‌تواند روی نگاشت، کمک، تقویت و اصلاح توابع شناختی^۵ و توابع حسی-حرکتی تمرکز کند [۱۵]. ایده سیستم BCI، استخراج الگوهای مغزی مرتبط با فعالیت‌های ذهنی (آگاهانه یا ناآگاهانه) است. این وسیله، امکان ارتباط

¹ Motor Neuron Disease (MND)

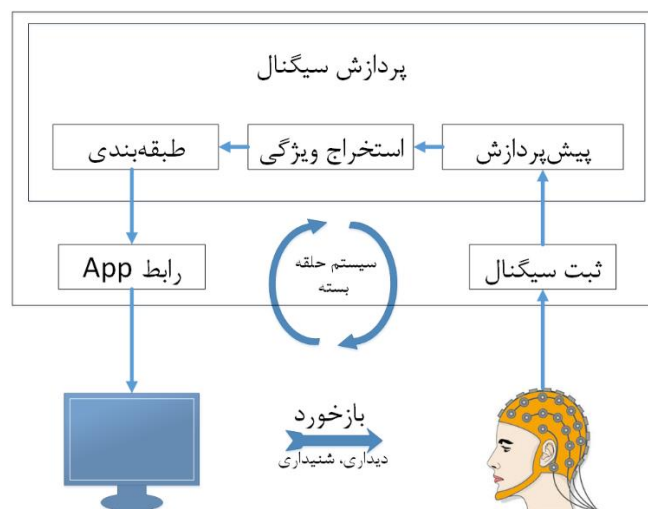
² Amyotrophic Lateral Sclerosis (ALS)

³ Encephalitis

⁴ Brain-Computer Interface (BCI)

⁵ Cognitive Function

برای بیمارانی با عدم کنترل کافی روی سیستم حرکتی (عضلات) را فراهم می‌کند. یک سیستم BCI، نیازمند یک طراحی دقیق و پیچیده است. از جمله این موارد، بهبود تکنیک‌های اکتساب سیگنال، توسعه روش‌های استخراج ویژگی، الگوریتم‌های ترجمه، توسعه سخت افزار، بهبود پروتکل‌های عملیاتی و استراتژی‌های آموزش کاربر هستند [۱۶-۱۸]. مراحل کلی یک سیستم BCI در شکل ۱-۱ نشان داده شده است. سیگنال‌های مغزی اندازه‌گیری شده و تبدیل به سیگنال‌های کنترلی برای برنامه‌های مختلف می‌شوند. سپس، دستگاه به کاربر بازخورد می‌دهد. اولین گام در این سیستم، ثبت اطلاعات یا ثبت سیگنال است. مرحله دوم، پردازش سیگنال‌های ثبت شده است. به طور کلی، مرحله پردازش سیگنال شامل پیش‌پردازش، استخراج ویژگی و طبقه‌بندی است. مرحله پیش‌پردازش مانند حذف نویز، شامل الگوریتم‌هایی به منظور بهبود کیفیت سیگنال‌های ورودی است. گام استخراج ویژگی شامل الگوریتم‌هایی برای استخراج ویژگی در سیگنال‌های ورودی است. در مرحله سوم، ویژگی‌های استخراج شده طبقه‌بندی می‌شوند. سیگنال کنترلی، به یک برنامه کاربردی منتقل شده و پس از آن، بازخورد به کاربر برمی‌گردد. این بازخورد در سیستم می‌تواند بصری یا شنوایی باشد. در سیستم BCI، این مراحل بین کاربر و برنامه ادامه می‌یابد تا زمانی که سیستم متوقف شده و خاموش شود.



شکل ۱-۱) مراحل سیستم BCI.

۱-۲ اندازه‌گیری فعالیت ذهنی

فعالیت مغزی، رفتارهای الکتروفیزیولوژیکی^۱ و همودینامیکی^۲ تولید می‌کند. سنسورهای متعددی وجود دارند که می‌توانند انواع مختلفی از فعالیت‌ها را در مغز تشخیص دهند. سیستم‌های BCI در مرحله ثبت داده، شامل دو گروه مبتنی بر تکنیک‌های تهاجمی و غیرتهاجمی هستند. تکنولوژی‌های تهاجمی به آن دسته از تکنولوژی‌ها گفته می‌شود که حسگرها مستقیماً درون مغز کار گذاشته می‌شوند. دقت و قدرت سیگنال ثبت شده توسط این تکنولوژی‌ها، بسیار بالا و نسبت سیگنال به نویز در این سیگنال‌ها، بسیار پایین است. اما از مشکلات اصلی این نوع ثبت، نیاز به عمل جراحی بر روی مغز بیمار است که ریسک مصدومیت را در بیمار بالا می‌برد. از این رو، از این روش به ندرت استفاده می‌شود. اما تکنولوژی دیگر برای ثبت سیگنال‌های مغزی، تکنولوژی‌های غیرتهاجمی است. در ثبت سیگنال‌های مغزی با کمک این تکنولوژی‌ها، سیگنال از روی پوست سر، بدون هیچگونه عمل جراحی ثبت می‌شود. این نوع سیگنال‌ها مشکلاتی نیز دارند. از جمله مشکلات این نوع سیگنال‌ها، وجود نویزهای مختلف و همچنین نسبت سیگنال به نویز بسیار پایین، در کنار نیاز به تقویت‌کننده برای مشاهده این سیگنال روی صفحه نمایش کامپیوترها است. تکنیک‌های غیر تهاجمی شامل همه‌ی روش‌های ثبت فعالیت ذهنی از خارج مرزهای بدن است [۱۹]. این روش‌ها می‌توانند دو گروه از سیگنال‌ها را اندازه‌گیری کنند. گروه اول، سیگنال‌هایی هستند که از روی فعالیت عصبی ثبت می‌شوند. به این نوع سیگنال‌ها، الکتروفیزیولوژیکی گفته می‌شود [۲۰]. در این گروه، روش‌های مبتنی بر EEG^۳ و MEG^۴ قرار دارند. در MEG، فعالیت مغز، با رزولوشن بالا اندازه‌گیری می‌شود. دستگاه‌های مورد استفاده در این روش، حجیم و گران‌قیمت هستند. به همین خاطر، گزینه نامناسبی برای استفاده در BCI هستند [۲۱]. اما EEG، فعالیت مغز را با اندازه‌گیری میدان‌های الکتریکی تولید شده از نورون‌های فعال^۵ ثبت می‌کند. این نوع سیگنال‌ها رزولوشن مکانی

^۱ Electrophysiological

^۲ Hemodynamically

^۳ Electroencephalography

^۴ Magnetoencephalography

^۵ Firing

پایین و رزولوشن زمانی بسیار بالایی (میلی ثانیه) دارند. دستگاه‌های مورد استفاده در این روش، ارزان و ساده هستند [۲۱]. از این رو، از EEG در سیستم‌های BCI به صورت گسترده استفاده شده است [۱۵].

گروه دوم، سیگنال‌هایی هستند که از فعالیت‌های همودینامیک ناشی می‌شوند. این نوع سیگنال‌ها سطح اکسیژن خون را اندازه‌گیری می‌کنند [۲۲]. این گروه سیگنال‌ها می‌توانند به وسیله روش‌های fMRI^۱ و fNIRS^۲ تشخیص داده شوند. در fMRI، سیگنال‌های وابسته به سطح اکسیژن خون^۳ اندازه‌گیری می‌شوند. سطح‌های مختلف اکسیژن خون می‌توانند توسط fNIRS نیز اندازه‌گیری شوند. مزیت این روش، قابل حمل بودن دستگاه ثبت سیگنال است. این نوع سیگنال، رزولوشن زمانی بالاتر اما رزولوشن مکانی پایین‌تری در مقایسه با fMRI دارد [۲۳]. کار با fMRI در سیستم‌های BCI از نظر اندازه‌گیری در زمان واقعی مناسب نیست و این روش برای کاربر هزینه‌بر و دشوار است. از این رو، در سال‌های اخیر از fNIRS نیز در سیستم‌های BCI استفاده می‌شود [۲۴، ۲۵]. در جدول ۱-۱ خلاصه‌ای از روش‌های مختلف ثبت سیگنال ارائه شده است.

جدول ۱-۱) مشخصات روش‌های ثبت سیگنال‌های مختلف [۱۹].

تکنیک تصویربرداری	فعالیت اندازه- گیری شده	وضوح زمانی	وضوح مکانی	خطرپذیری	قابل حمل بودن
EEG	الکتریکی	حدود ۰,۰۵ ثانیه	حدود ۱۰ میلی متر	غیر تهاجمی	قابل حمل
MEG	مغناطیسی	حدود ۰,۰۵ ثانیه	حدود ۵ میلی- متر	غیر تهاجمی	غیر قابل حمل

^۱ Functional Magnetic Resonance Imaging

^۲ Functional Near-Infrared Spectroscopy

^۳ Blood-oxygen-level Dependent

fMRI	متابولیکی	حدود ۱ ثانیه	حدود ۱ میلی-متر	غیر تهاجمی	غیر قابل حمل
fNIRS	متابولیکی	حدود ۱ ثانیه	حدود ۵ میلی-متر	غیر تهاجمی	قابل حمل

۱-۳ اندازه‌گیری فعالیت ذهنی با استفاده از EEG و fNIRS

۱-۳-۱ فعالیت‌های مغزی استفاده شده در سیستم‌های BCI مبتنی بر

EEG

مغز انسان می‌تواند سیگنال‌هایی برای فعالیت‌های مغزی مختلف ایجاد کند. عملکرد اکثر این سیگنال‌ها هنوز به صورت کامل توسط محققان و پژوهشگران آنالیز نشده است. با این حال، پدیده فیزیولوژیکی برخی از این سیگنال‌ها آنالیز شده و در برخی برنامه‌های کاربردی BCI استفاده می‌شوند. این سیگنال‌ها شامل پتانسیل‌های رخداد P300، پتانسیل‌های رخداد دیداری حالت ایستا^۱، پتانسیل‌های قشری آهسته^۲، ریتم‌های حسی-حرکتی^۳ و تصورات حرکتی^۴ هستند [۲۶]. در این پایان‌نامه، برای سیگنال EEG، از سیگنال‌های تصور حرکتی استفاده شده است. برای این منظور، در ادامه، به طور مختصر درباره این نوع سیگنال شرح داده شده است.

مطابق با حالت مغز، نوسانات مختلفی در فعالیت مغزی رخ می‌دهد. این نوسانات در چهار گروه مختلف رخ می‌دهند. این چهار گروه مبتنی بر باندهای فرکانسی در EEG، دلتا (۱ تا ۴ هرتز)، تتا (۴ تا ۸ هرتز)، مو (۸ تا ۱۳ هرتز)، بتا (۱۳ تا ۲۵ هرتز) و گاما (۲۵ تا ۴۰ هرتز) هستند. ریتم‌های حسی-حرکتی به فعالیت‌های نوسانی مشاهده شده در ناحیه‌های حرکتی و حسی-حرکتی اشاره می‌کنند. فعالیت‌های موجود در بخش‌های مختلف بدن در ناحیه‌های مختلفی در قشر حسی-حرکتی مغز نگاشت می‌شوند.

^۱ Steady State Visual Evoked Potentials (SSVEP)

^۲ Slow Cortical Potentials (SCP)

^۳ Sensory-Motor Rhythms (SMR)

^۴ Motor Imagery (MI)

فعالیت در بخش خاصی از بدن غیر از ناحیه‌های حسی-حرکتی باعث کاهش فعالیت SMR در ناحیه مغزی مربوطه می‌شود. این کاهش، ناسازگاری مرتبط با رخداد^۱ نامیده می‌شود [۲۷]. متقابلاً، سازگاری مرتبط با رخداد^۲، افزایش در فعالیت SMR در طول دوره استراحت^۳ بعد از حرکت‌های بدن است. همچنین، این فعالیت‌ها، در زمان تصور حرکت بدن رخ می‌دهند حتی اگر آن به طور واقعی حرکت داده نشود. نوسانات ERD/ERS را می‌توان در EEG در باندهای بتا و مو مشاهده کرد [۲۶].

اصطلاح تصور حرکتی، به حرکت بخشی از بدن در حالت تصور، بدون حرکت واقعی آن اشاره دارد. این تصور باعث ایجاد فعالیت‌های ERD در مغز می‌شود و می‌توان آن را در EEG مشاهده کرد. با این حال، الگوهای ERD/ERS مرتبط با تمام بخش‌های بدن را نمی‌توان در EEG مشخص کرد. الگوهای تولید شده باید به اندازه کافی بزرگ باشند تا از پس‌زمینه EEG جدا شوند. در حال حاضر، چهار نوع عمل تصور حرکتی وجود دارد که می‌توان آن‌ها را از روی EEG تشخیص داد. این عمل‌ها حرکت‌های دست چپ، دست راست، پاها و زبان هستند [۲۸].

سیگنال‌های مرتبط با MI معمولاً با استفاده از الکترودهای C3، C4 و Cz در EEG ثبت می‌شوند. فعالیت برانگیخته توسط تصور حرکت دست راست را می‌توان اکثراً در موقعیت الکتروود C3 و تصور حرکتی دست چپ را می‌توان اکثراً در موقعیت الکتروود C4 مشاهده کرد. تصورات حرکتی پای چپ و راست قابل جداسازی نیستند، چرا که ناحیه‌های این دو حرکت بسیار نزدیک هم هستند. بنابراین پتانسیل‌های اندازه‌گیری شده در پوست سر بسیار شبیه به یکدیگر هستند. هر دو مورد این فعالیت‌های برانگیخته شده، عمدتاً در ناحیه Cz دیده می‌شوند. تصورات حرکتی در محدوده گسترده‌ای از برنامه‌های کاربردی BCI استفاده می‌شوند. از این نوع سیگنال‌ها برای ارسال دستور مورد نظر به سیستم استفاده می‌شود [۲۹].

¹ Event-related desynchronization (ERD)

² Event-related synchronization (ERS)

³ Relaxation

۱-۳-۲ فعالیت‌های مغزی استفاده شده در سیستم‌های BCI مبتنی بر fNIRS

استفاده از fNIRS در مقایسه با EEG، به عنوان یک ماهیت اسکن مغزی^۱ نسبتاً جدید در نظر گرفته شده است. این تکنیک، از نور NIR^۲ در طیف قابل مشاهده الکترومغناطیسی استفاده می‌کند. دستگاه‌های ثبت این نوع سیگنال، غلظت‌های HbO₂^۳ و HHb^۴ مغزی را در یک ناحیه خاص از مغز به صورت غیر مستقیم اندازه‌گیری می‌کنند [۳۰].

مفهوم fNIRS بسیار شبیه به fMRI است. هر دو بر اساس این فرض کار می‌کنند که فعالیت مغز منجر به افزایش مصرف اکسیژن می‌شود. این عمل منجر به تغییر در HbO₂ و HHb می‌شود که می‌تواند توسط fNIRS به صورت غیر تهاجمی تشخیص داده شود. اصل اساسی برای fNIRS این است که بیشتر بافت‌های بیولوژیکی مانند پوست و استخوان برای نور NIR شفاف هستند، در حالی که HbO₂ و HHb جذب‌کننده‌های بهتری برای NIR در طیف ۷۰۰ تا ۹۰۰ نانومتر هستند [۳۰].

۱-۴ هدف پایان‌نامه

در روش‌های موجود در گذشته، برای این منظور کارهای زیادی انجام شده که شامل مشکلاتی از جمله دقت پایین، روش‌های پیش‌پردازش طولانی و استخراج ویژگی به وسیله انسان است. هدف این پایان‌نامه، رفع این مشکلات، به همراه بالا بردن دقت طبقه‌بندی برای تشخیص اهداف ذهنی است. بر این اساس، این پایان‌نامه بر سه مرحله تمرکز دارد.

در مرحله اول، پیش‌پردازشی روی سیگنال‌های EEG و fNIRS انجام می‌شود. کاهش نرخ نمونه‌برداری برای سیگنال EEG و افزایش نرخ نمونه‌برداری برای سیگنال fNIRS از مراحل پیش‌پردازش هستند. با این فرآیند، نرخ نمونه‌برداری برای هر دو سیگنال برابر شده و عملیات بعدی راحت‌تر انجام می‌شوند.

¹ Neuroimaging

² Near Infrared

³ Oxygenated-hemoglobin

⁴ Deoxygenated-hemoglobin

مرحله دوم، تبدیل سیگنال‌های سری زمانی به ماتریس‌های سه بعدی شامل ویژگی‌های مکانی و زمانی است. در این مرحله، از نقشه الکترودها، برای ثبت EEG و fNIRS برای استخراج ویژگی‌های مکانی استفاده می‌شود. سپس از یک نرمال‌سازی Z-score برای نرمال‌سازی داده‌ها استفاده شده است. در نهایت، از روش پنجره‌گذاری لغزان برای استخراج ویژگی‌های زمانی در این سیگنال‌ها استفاده شده است.

مرحله سوم، فرآیند تشخیص اهداف ذهنی با استفاده از یک مدل مبتنی بر یادگیری عمیق است. در این مرحله، با استفاده از یک مدل یادگیری عمیق ترکیبی از CNN^۱ و LSTM^۲، مدلی برای تشخیص اهداف ذهنی آموزش داده شده است که در این پایان‌نامه گزارش داده می‌شود.

۵-۱ چالش‌ها و راه‌حل‌های آن‌ها

یکی از چالش‌های مهم در طراحی سیستم‌های BCI، استفاده از سیگنال‌های الکتروفیزیولوژیکی یا همودینامیکی به صورت جداگانه در یک سیستم است. تاکنون، تحقیقات زیادی روی سیستم‌های BCI مبتنی بر سیگنال‌های الکتروفیزیولوژیکی انجام شده است. پارامترهای استفاده شده برای این سیستم‌ها شامل تصورات حرکتی، SSVEP و P300 هستند [۳۷-۳۱]. تاکنون از این سه نوع ویژگی برای کنترل ویلچر [۳۸-۴۰] و کوادکوپتر [۴۱] استفاده شده است. با این حال، روش‌های الکتروفیزیولوژیکی از رزولوشن مکانی پایین رنج می‌برند. از سوی دیگر، استفاده از روش‌های همودینامیکی به تنهایی نیز از محدودیت رزولوشن زمانی رنج می‌برد [۴۲]. در این پایان‌نامه، با ترکیب این دو سیگنال، برای حل این چالش تلاش می‌کنیم. در سال‌های اخیر، مطالعات BCI امکان ترکیب EEG را با سایر فناوری‌های اسکن مغزی نشان داده است [۴۳]. در این میان، ترکیب EEG با fNIRS نتایج خوبی را ارائه داده است [۴۲، ۴۴، ۴۵] و می‌تواند این محدودیت‌ها را برای هر سیگنال کاهش دهد [۴۶-۴۸]. یکی دیگر از دلایل استفاده از هر دو سیگنال fNIRS و EEG، دشواری ثبت همزمان fMRI با سیگنال‌های الکتروفیزیولوژیکی است. این دشواری، به دلیل ثبت این سیگنال‌ها در محیط آزمایشگاهی خاص است.

^۱ Convolutional Neural Network

^۲ Long-Short Term Memory

همچنین در این نوع ثبت، امواج الکترومغناطیسی اثر مخربی بر سیگنال‌های EEG دارند. درمقابل، سیستم ثبت fNIRS به خوبی از EEG جدا می‌شود و برای ثبت همزمان با سیگنال‌های الکتروفیزیولوژیکی (به دلیل استفاده از نور نزدیک به مادون قرمز) مناسب است. در واقع، ثبت سیگنال‌های الکتروفیزیولوژیکی با نور انجام نمی‌شود و ثبت سیگنال‌های نوری، نویز و خللی در این سیگنال‌ها ایجاد نمی‌کند. علاوه بر این، هزینه کم، نویز کم و قابل حمل بودن، استفاده همزمان سیگنال‌های EEG و fNIRS را کاربردی‌تر ساخته است. ضبط طولانی مدت EEG-fNIRS، امواج صرعی بین حمله‌ای^۱ یا صرع [۴۹-۵۲] را با استفاده از تغییرات همودینامیک قفل شده در زمان وقوع IEDها بیشتر بررسی می‌کند. ترکیب EEG و fNIRS، می‌تواند وضعیت حدود پنج دقیقه قبل از این بیماری‌ها را تشخیص دهد [۵۱]. ترکیب پتانسیل مرتبط با رخداد^۲ و پاسخ‌های همودینامیک بدست آمده توسط fNIRS نیز برای مطالعه بر روی قشر بینایی مورد استفاده قرار می‌گیرند [۵۳، ۵۴].

علاوه بر این، چندین چالش دیگر نیز وجود دارد. اولین چالش، نویز موجود در سیگنال‌ها است. بخشی از نویزهای رایج در سیگنال، تداخل فرکانس با سایر تجهیزات و اتصال نامناسب الکترودها است. در این سیگنال‌ها، فعالیت‌های فیزیولوژیکی موجب ایجاد نویز می‌شوند. این فعالیت‌ها موجب کاهش میزان نرخ سیگنال به نویز^۳ می‌شود. از جمله این فعالیت‌ها می‌توان به پلک زدن، حرکت عضلانی و ضربان قلب اشاره کرد [۵۵]. الگوریتم‌های پیش‌پردازش برای حذف نویز در سیگنال مفید هستند. چالش دوم، ابهام در ارتباط بین سیگنال‌ها و اهداف مغزی مرتبط با آن‌ها است. در حقیقت، اهداف ذهنی را نمی‌توان به طور مستقیم به وسیله سیگنال‌ها مشاهده کرد [۵۵]. برای حل این چالش، تجزیه و تحلیل و طبقه‌بندی سیگنال‌ها مهم است. برای این منظور، در این پایان‌نامه، طبقه‌بندی یادگیری عمیق مورد استفاده قرار گرفته است. چالش سوم، استخراج ویژگی برای شناسایی اهداف ذهنی توسط کاربر است. برای استخراج این ویژگی‌ها نیاز به پردازش‌های سنگین است [۵۶]. اکثر مراحل پردازش در برنامه‌های کاربردی شامل

¹ Interictal Epileptic Discharges (IED)

² Event-Related Potential (ERP)

³ Singal-to-Noise Rate (SNR)

دو مرحله، حذف نویز [۵۷] و استخراج ویژگی [۵۸] است. این مراحل برای سیستم‌های BCI به دلیل زمان‌بر بودن و وابستگی به سطح دانش حرفه‌ای یا استفاده از یک متخصص، غیر عملی است. در این پایان‌نامه استخراج ویژگی با استفاده از الگوریتم یادگیری عمیق انجام می‌شود. الگوریتم‌های یادگیری عمیق به طور خودکار، ویژگی‌های مناسب را از سیگنال استخراج می‌کنند. به لطف روش یادگیری عمیق، مرحله زمان‌بر و پر خطای استخراج ویژگی حذف شده است.

یکی از نکات مهم در این پایان‌نامه، حل مشکل دقت نامناسب در طبقه‌بندی سیگنال‌های خام است. بنابراین، انجام پیش‌پردازش بر روی داده‌ها ضروری است. ویژگی مهم سیگنال‌های بیولوژیکی که با کمک الکتروود روی پوست ثبت می‌شوند، موقعیت الکتروود روی بدن است. در این پایان‌نامه، این ویژگی به نام ویژگی مکانی نامگذاری شده است. علاوه بر این، هر سیگنال دارای یک طول زمانی است. بنابراین، زمان در سیگنال‌های بیولوژیکی نیز اهمیت دارد. این ویژگی در این پایان‌نامه نیز به عنوان ویژگی زمانی شناخته می‌شود. این ویژگی‌ها فقط در دو مرجع [۵۹، ۶۰] استفاده شده است. در این پایان‌نامه، با توجه به این دو ویژگی، با استفاده از ترکیبی از دو سیگنال EEG و fNIRS و استفاده از الگوریتم‌های یادگیری عمیق، سعی می‌شود سیگنال‌ها با دقت بالا طبقه‌بندی شوند و اهداف ذهنی کاربر با حداقل خطا مشخص شود.

کارهای اخیر تلاش کرده‌اند که سیگنال‌های EEG و fNIRS را ترکیب کنند [۱۴، ۴۲، ۴۹، ۵۰، ۵۲، ۶۱] که در میان آنها تنها یک روش از یادگیری عمیق استفاده کرده است [۶۲]. در روش پیشنهادی، سیگنال‌های بیولوژیکی خام به تانسورهای سه بعدی تبدیل شده است. این شبکه‌های سه بعدی، توپوگرافی مغز را در زمان‌های مختلف نشان می‌دهند. سپس، از پنجره لغزان برای ایجاد کلیپ‌های داده استفاده شده است. پنجره لغزان به استفاده از ویژگی‌های زمانی در سیگنال‌های بیولوژیکی کمک می‌کند. در هر کلیپ، به طور همزمان ویژگی فضایی و زمانی وجود دارد. پس از این، از یادگیری عمیق برای آموزش مدل با این تانسورهای سه بعدی استفاده شده است. در این تحقیق، با بهبود مرحله پیش‌پردازش

و تطبیق الگوریتم یادگیری عمیق به ویژه برای هدف ترکیب EEG-fNIRS، دقت تشخیص بهبود یافته است.

۱-۶ نوآوری

نوآوری این تحقیق از دو جهت قابل بررسی است. ابتدا باید بیان نمود که در این پایان نامه از دو داده EEG و fNIRS به صورت همزمان استفاده شده است. سازگارسازی این دو سیگنال، از جمله نوآوری این پایان نامه است. روش پیش پردازش و ترکیب این داده ها به صورتی که بتوان از هر دو داده به صورت یکجا برای طبقه بندی استفاده کرد، بخشی از مسئله مورد بررسی است. از طرف دیگر، در این پایان نامه، یک مدل شبکه عمیق برای طبقه بندی این داده ها با بالاترین دقت ارائه شده است.

۱-۷ ساختار پایان نامه

در ادامه این تحقیق، در بخش دوم، مروری بر روش های پیشین انجام شده است. در بخش سوم، الگوریتم های پیشنهادی برای پیش پردازش داده ها و ترکیب و تبدیل سیگنال ها به تانسورهای سه بعدی ارائه شده است. سپس مدل پیشنهادی مرتبط با یادگیری عمیق شرح داده شده است. در بخش چهارم، نتایج آزمایش ها مورد بررسی قرار گرفته است. در بخش پایانی، نتیجه گیری ارائه شده است.

فصل ۲: کاربردی ماشین

امروزه در دنیا، به دلیل کاهش مصدومیت‌ها و صدمات وارده به بیماران و همچنین کاهش ریسک عمل جراحی روی مغز انسان، به صورت گسترده از روش‌های غیرتهاجمی برای ثبت سیگنال استفاده می‌شود. روش‌های بسیار زیادی برای کاهش یا غلبه بر مشکلات این نوع سیگنال‌ها ارائه شده است که در ادامه به برخی از آن‌ها اشاره می‌کنیم. همانطور که در شکل ۱-۱ مشاهده می‌کنیم، مراحل اصلی سیستم BCI بیان شده است. این مراحل شامل الگوریتم‌ها و روش‌هایی متعددی هستند. در ادامه، برخی از مهم‌ترین و رایج‌ترین روش‌های استفاده شده توسط پژوهشگران را بررسی می‌کنیم.

۲-۱ پیش‌پردازش: نیازمندی‌های بهبود سیگنال همراه با کاهش

نویز

این بخش، نیازمندی‌های پیش‌پردازش سیگنال و روش‌های رایج ارائه شده توسط گروه‌های تحقیقاتی مختلف را شرح می‌دهد. نیاز اول مرحله پیش‌پردازش در BCI، فیلترکردن نویز یا سیگنال‌های ناخواسته است که درون سیگنال قرار گرفته است. ما در این پایان‌نامه، به دنبال تصورات حرکتی، به عنوان سیگنال اصلی هستیم. از این رو، سیگنال‌های تولید شده توسط حرکات عضلات^۱، فعالیت حاصل از حرکت چشم‌ها^۲ و فعالیت قلب در بدن را می‌توانیم به عنوان سیگنال‌های ناخواسته در نظر بگیریم. تکنیک‌های مختلفی برای حذف نویز وجود دارد. اکثر گروه‌های BCI از یک فیلتر میان‌گذر به عنوان رایج‌ترین ابزارهای پیش‌پردازش، برای حذف نویز و استخراج اطلاعات مربوط به فرکانس‌های مختلف سیگنال‌های EEG استفاده می‌کنند. ابزار پیش‌پردازش خوب می‌تواند عملکرد BCI را بهبود بخشد. این درحالی است که الگوریتم‌های مراحل بعدی، یعنی استخراج ویژگی و طبقه‌بندی روی سیگنال‌های پردازش شده انجام می‌شود. بنابراین پیش‌پردازش، اغلب مرحله‌ای بسیار مهم در نظر گرفته می‌شود. تعدادی از بهترین تکنیک‌های پیش‌پردازش استفاده شده در سیستم‌های BCI، در بخش‌های زیر شرح داده شده است [۶۴].

^۱ Electromyography (EMG)

^۲ Electrooculography (EOG)

۲-۱-۱ روش مرجع دهی^۱

روش های مرجع دهی به عنوان فیلترهای فضایی در نظر گرفته می شوند. فرض اولیه این روش ها بر این اساس است که فعالیت در سرتاسر مغز در هر لحظه با صفر جمع می شود. دو روش مرجع دهی که اغلب در BCI ها استفاده می شود، مرجع متوسط مشترک^۲ و روش لاپلاس^۳ می باشند [۶۴].

روش مرجع متوسط مشترک یک روش حذف نویز پس زمینه سیگنال است که بار محاسباتی پایینی دارد. در این روش، میانگین نمونه های همه کانال ها از تمام مقادیر الکترودها کم می شود. در این فرآیند، میانگین همه فعالیت ها، برآوردی از فعالیت در محل مرجع را نشان می دهد و در نتیجه کم کردن این میانگین در اصل یک راه حل است.

بر خلاف روش CAR، روش لاپلاس در یک موقعیت الکتروده خاص، با کم کردن ترکیبی از سیگنال ها از مجموعه ای از الکترودهای اطراف سیگنال الکتروده مرکزی، محاسبه می شود.

۲-۱-۲ تجزیه و تحلیل اجزای اصلی^۴

تجزیه و تحلیل اجزای اصلی، یک تکنیک آماری رایج برای تشخیص و ارجاع دوباره داده با نگاشت خطی است. در این روش، تعدادی از متغیرهای همبسته احتمالی، به تعداد کمتری از متغیرهای غیرهمبسته شناخته شده، به عنوان اجزای اصلی تبدیل می شوند. محور اصلی یا اولین جزء اصلی به گونه ای محاسبه می شود که شامل بیشترین میزان تغییرپذیری در داده باشد. به طور مشابه، اجزا یا محورهای بعدی با توجه به مقدار تغییرپذیری و در جهت ترتیب نزولی محاسبه می شوند. از آنجایی که داده های تغییر یافته بیشترین تغییر را در چند مولفه اول دارند، اجزای باقی مانده را می توان نادیده گرفت و ابعاد داده را کاهش داد [۶۵].

¹ Referencing

² Common Average Reference (CAR)

³ Laplace

⁴ Principal Component Analysis (PCA)

۲-۱-۳ تجزیه و تحلیل اجزای مستقل^۱

این روش یک تکنیک محاسباتی و آماری برای آشکارسازی اجزای پنهان است. این تکنیک آشکارکننده-ی مجموعه‌ای از متغیرها، اندازه‌گیری‌ها و سیگنال‌های تصادفی است. یک فرض معمول برای این روش، استقلال منابع ناشناخته از یکدیگر است. روش PCA سیگنال‌هایی را استخراج می‌کند که ناهمبسته هستند درحالی‌که ICA سیگنال‌هایی را استخراج می‌کند که به طور آماری مستقل هستند. این مناسب-ترین تمایز بین این روش‌ها است [۶۶].

۲-۲ استخراج ویژگی

استخراج ویژگی، توصیفی از سیگنال‌ها از نظر ویژگی‌های متمایز را، برای طبقه‌بندی آن‌ها، فراهم می‌کند. این بخش برخی از شناخته شده‌ترین و رایج‌ترین روش‌های استخراج ویژگی را بیان می‌کند که توسط گروه‌های تحقیقاتی بسیاری پیاده‌سازی شده است.

۲-۲-۱ ویژگی‌های توان باند^۲

توان به صورت عادی به عنوان انرژی در یک فاصله زمانی واحد تعریف می‌شود. این روش کلاسیک می‌تواند برای تعیین کمیت پدیده‌های ERD و ERS بکار برود. این روش بر اساس محاسبه مربع دامنه سیگنال در یک پنجره زمانی کوچک عمل می‌کند. از این رو، در هر پنجره، سیگنال، ساکن در نظر گرفته می‌شود. در نتیجه، در این روش ترجیحا باید از پنجره‌های کوچک‌تر استفاده کرد تا اثرات غیر ذاتی سیگنال درون آن پنجره کاهش یابد. در این روش، معمولاً از دو باند فرکانسی مو و بتا استفاده می‌شود [۶۷].

^۱ Independent Component Analysis (ICA)

^۲ Band Power (BP)

۲-۲-۲ ویژگی‌های چگالی طیف توان^۱

این روش، یک راه متفاوت برای ارزیابی توان باند است. PSD شرح می‌دهد که چگونه سیگنال در محدوده فرکانسی توزیع شده است. ویژگی PSD محتوای فرکانسی یک فرآیند تصادفی را جذب می‌کند و از نظر ریاضی، به صورت تبدیل فوریه تابع خودکار همبستگی سیگنال بیان می‌شود. تبدیل‌های فرکانسی ساده، مانند PSD، قادر به بیرون کشیدن محتوای فرکانس کلی سیگنال هستند و این به تمایز بین سیگنال‌های مختلف کمک می‌کند [۶۸].

۲-۲-۳ روش زمان-فرکانس^۲

یک سیگنال، به عنوان یک تابع در زمان، ممکن است به عنوان یک نمایش با وضوح کامل زمانی در نظر گرفته شود. به طور مشابه، تبدیل فوریه سیگنال ممکن است به عنوان نمایش با وضوح طیفی عالی، اما بدون اطلاعات زمانی در نظر گرفته شود. اما، ترکیب نمایش زمان-فرکانس، که سیگنال یک بعدی را به تابعی دو بعدی از زمان و فرکانس نگاشت می‌کند، می‌تواند برای تجزیه و تحلیل محتوای طیفی متغیر با زمان استفاده شود. تکنیک تبدیل موجک^۳ برای این روش و همچنین برای حذف نویز مناسب است. از این رو، اکثر روش‌های مبتنی بر روش زمان-فرکانس، از الگوریتم‌های استخراج ویژگی مبتنی بر موجک استفاده می‌کند. تبدیل موجک از یک پنجره متغیر در طول سیگنال استفاده می‌کند. این کار منجر به استخراج ویژگی‌های عالی از سیگنال می‌شود [۶۴].

۲-۳ طبقه‌بندی

روش‌های طبقه‌بندی برای مشخص کردن الگوهای فعالیت در سیگنال‌های بیولوژیکی بکار می‌رود. به عبارت دیگر، روش‌های طبقه‌بندی برای یادگرفتن نگاشت بین داده‌های موجود در سیگنال و کلاس‌های مرتبط با فعالیت‌های ذهنی بکار می‌روند. بنابراین، طبقه‌بند استفاده شده، کلاس یا گروه سیگنال را

^۱ Power Spectral Density (PSD)

^۲ Time-Frequency

^۳ Wavelet Transform

شناسایی می‌کند. برخی از طبقه‌بندهای اصلی در برنامه‌های کاربردی BCI، شامل تجزیه و تحلیل تفکیک خطی^۱، ماشین‌های بردار پشتیبان^۲ و یادگیری عمیق^۳ هستند. حضور نويز بیش از حد و اطلاعات ناخواسته می‌تواند منجر به شکست طبقه‌بند شود.

۲-۳-۱ تجزیه و تحلیل تفکیک خطی

تجزیه و تحلیل تفکیک خطی، یکی از موفق‌ترین طبقه‌بندهای خطی در سیستم‌های BCI می‌باشد. روش LDA یک ابر صفحه مجزا را تولید می‌کند که به صورت $f(X) = W^T X + B$ شرح داده می‌شود. در این طبقه‌بند، اگر $f(X) > 0$ باشد، X متعلق به کلاس اول است و اگر $f(X) < 0$ ، متعلق به کلاس دوم خواهد بود. این تابع، تابع تفکیک خطی نامیده می‌شود و مرز تصمیم، ابر صفحه‌ای با مجموعه نقاط $f(X) = 0$ است. به طور کلی، LDA یک روش قوی، راحت برای استفاده، دارای پیچیدگی کم و دارای نیازمندی‌های محاسباتی بسیار کم است. به طور کلی، این روش نتایج خوبی را ارائه داده است [۶۹].

۲-۳-۲ ماشین بردار پشتیبان

ایده این روش برای یافتن یک ابر صفحه بهینه است که نقاط دو کلاس مختلف را توسط بزرگ‌ترین فاصله ممکن در فضای ویژگی جدا می‌کند. این امر می‌تواند از درون همان فضای ویژگی یا با طرح ویژگی‌هایی در صفحه‌ای با ابعاد بالاتر، بدست آید [۷۰].

۲-۳-۳ یادگیری عمیق

یادگیری عمیق چالشی بوده است که در دهه گذشته، به آرامی تغییر شکل داده است. یک تعریف ساده از یادگیری عمیق، شبکه عصبی با بیش از دو لایه است. به عبارت دیگر، یادگیری عمیق نوعی از الگوریتم‌های یادگیری ماشین است که هر دو فرآیند طبقه‌بندی و استخراج ویژگی را ترکیب می‌کند

^۱ Linear Discriminant Analysis (LDA)

^۲ Support Vector Machine (SVM)

^۳ Deep Learning (DL)

[۷۵-۷۱]. برخی از حقایق درباره تکامل شبکه‌های عصبی که منجر به تولید شبکه‌های عمیق شده است در ادامه آمده است:

- نورون‌های بیشتر نسبت به شبکه‌های قبلی.
- روش‌های پیچیده‌تر ارتباط لایه‌ها و نورون‌ها در شبکه‌های عصبی.
- تحول در میزان قدرت محاسباتی در فرآیند آموزش.
- استخراج ویژگی خودکار.

استخراج ویژگی خودکار، یکی از مزایای بزرگی است که یادگیری عمیق را جایگزین الگوریتم‌های یادگیری ماشین کرده است. بوسیله استخراج ویژگی، ما تعیین می‌کنیم که فرآیند شبکه چه تصمیمی بگیرد. یعنی کدام مشخصات یک مجموعه داده می‌توانند به عنوان شاخص برای برچسب‌گذاری آن داده استفاده شوند. یادگیری عمیق برای تقریباً هر نوع داده‌ای با حداقل تنظیمات، از الگوریتم‌های متداول در مقدار دقت پیشی گرفته است. شبکه‌های عمیق را می‌توان به چهار ساختار مهم تقسیم کرد [۷۶].

- شبکه‌های از پیش آموزش دیده بدون ناظر.^۱
- شبکه‌های عصبی کانولوشن.^۲
- شبکه‌های عصبی مکرر.^۳
- شبکه‌های عصبی بازگشتی.^۴

در ادامه دو شبکه عصبی کانولوشن و مکرر را با جزئیات بیشتر بررسی می‌کنیم.

^۱ Unsupervised Pre-trained Networks

^۲ Convolutional Neural Network (CNN)

^۳ Recurrent Neural Network (RNN)

^۴ Recursive Neural Network

۲-۳-۱ شبکه عصبی کانولوشن

شبکه‌های عصبی کانولوشن [۷۷]، یکی از زیرمجموعه‌های یادگیری عمیق است. این شبکه در سال‌های اخیر مورد توجه محققان هوش مصنوعی قرار گرفته است [۷۴، ۷۸، ۷۹، ۸۰، ۸۱]. شبکه‌های عصبی کانولوشن برای نگاشت داده تصویری به یک متغیر خروجی طراحی شده است. مزیت استفاده از CNNها، توانایی آن‌ها در توسعه یک نمایش داخلی از یک تصویر دو بعدی است. این مزیت اجازه می‌دهد تا CNN موقعیت و مقیاس‌های مختلف را یاد بگیرد که هنگام کار با تصویر بسیار مهم است. CNNها برای داده‌هایی دارای ارتباطات مکانی، بهتر کار می‌کنند. اما به طور کلی، CNNها برای مواردی از جمله پیش‌بینی طبقه‌بندی و پیش‌بینی رگرسیون نیز کاربرد دارند. از این نوع شبکه در طبقه‌بندی تصاویر از جمله تصاویر پزشکی x-ray [۸۲]، تصویرسازی تشدید مغناطیسی^۱ [۷۹]، تصاویر هیستوپاتولوژی^۲ [۸۰]، تصاویر رَحِم^۳ [۷۴] و تصاویر توموگرافی تخمینی^۴ [۸۱] استفاده شده است [۷۸]. اما تحقیقات کمی در زمینه استفاده از این شبکه، برای تحلیل سیگنال‌های فیزیولوژیکی وجود دارد [۷۸]. CNN برای تشخیص خودکار آریتمی‌ها استفاده شده است. در این تشخیص، از فواصل زمانی مختلف روی سیگنال‌های ECG استفاده شده است. دقت تشخیص آریتمی، برای فواصل زمانی پنج ثانیه‌ای، ۹۲٫۵ درصد است. همچنین این دقت برای فواصل زمانی دو ثانیه‌ای، ۹۴٫۹ درصد است [۷۲]. اخیراً برای تشخیص خودکار myocardial infraction [۸۳]، بیماری‌های عروق کرونر^۵ [۷۳] و طبقه‌بندی ضربان‌های مختلف قلب [۸۴] با استفاده از سیگنال‌های ECG، از CNN استفاده شده است.

خروجی نهایی مدل CNN، مبتنی بر وزن‌ها و بایاس‌های لایه قبلی در ساختار شبکه است. بر این اساس، وزن‌ها و بایاس‌های مدل به ترتیب با استفاده از روابط (۲-۱) و (۲-۲) برای هر لایه بروز می‌شوند.

$$\Delta W_l(t+1) = -\frac{x\lambda}{r} W_l - \frac{x}{n} \frac{\partial C}{\partial W_l} + m\Delta W_l(t) \quad (2-1)$$

¹ Magnetic Resonance Imaging (MRI)

² Histopathologic Images

³ Fundus Images

⁴ Computed Tomography Images

⁵ Coronary Artery

$$\Delta B_l(t+1) = -\frac{x}{n} \frac{\partial C}{\partial B_l} + m \Delta B_l(t) \quad (2-2)$$

در روابط (۱-۲) و (۲-۲)، پارامترهای $W, B, l, \lambda, x, m, n, t$ و C به ترتیب نشان‌دهنده‌ی وزن، بایاس، شماره لایه، پارامتر تنظیم مدل^۱، نرخ یادگیری، تعداد کل نمونه‌های آموزشی، تکانه^۲، قدم برورسانی و تابع هزینه هستند. پارامترهای تنظیم مدل، نرخ یادگیری و تکانه برای آموزش مدل CNN استفاده می‌شوند. این پارامترها می‌توانند مطابق با مجموعه داده، به منظور دستیابی به عملکرد مطلوب تنظیم شوند [۸۵]. نرخ تنظیم مدل، برای اجتناب از بیش‌برازش^۳ داده استفاده می‌شود. نرخ یادگیری برای کنترل جریان سرعت یادگیری شبکه در هنگام آموزش و تکانه به همگرایی داده‌ها کمک می‌کند.

لایه‌های پایه CNN شامل کانولوشن، حداکثرسازی^۴ و لایه‌ی تماماً متصل^۵ است [۸۶، ۸۷]. به طور معمول، با افزایش عمق شبکه، تمایل به یادگیری بیشتر می‌شود [۸۷]. لایه کانولوشن، با استفاده از کرنل (پنجره) با سیگنال ورودی کانالوو می‌شود [۸۸، ۸۹]. قدم‌ها^۶، میزان حرکت کرنل (پنجره) روی سیگنال ورودی را کنترل می‌کنند. عملیات کانالوو سیگنال EEG با کرنل (پنجره)، با استفاده از رابطه (۳-۲) انجام می‌شود [۷۸].

$$y_k = \sum_{n=0}^{N-1} x_n h_{k-n} \quad (2-3)$$

در رابطه (۳-۲)، x سیگنال EEG، h کرنل (پنجره) و N تعداد عناصر x هستند. بعد از عمل کانولوشن، نقشه ویژگی برای لایه بعدی تولید می‌شود. بعد از آن، لایه نرمال‌سازی دسته‌ای^۷ برای نرمال‌سازی داده‌های آموزشی ورودی بین لایه‌های میانی اعمال می‌شود. هدف این کار، فراهم کردن یادگیری سریع‌تر است. برای کاهش اندازه نقشه ویژگی، لایه max-pooling مورد استفاده قرار می‌گیرد. در نهایت

^۱ Regularization

^۲ Momentum

^۳ Overfitting

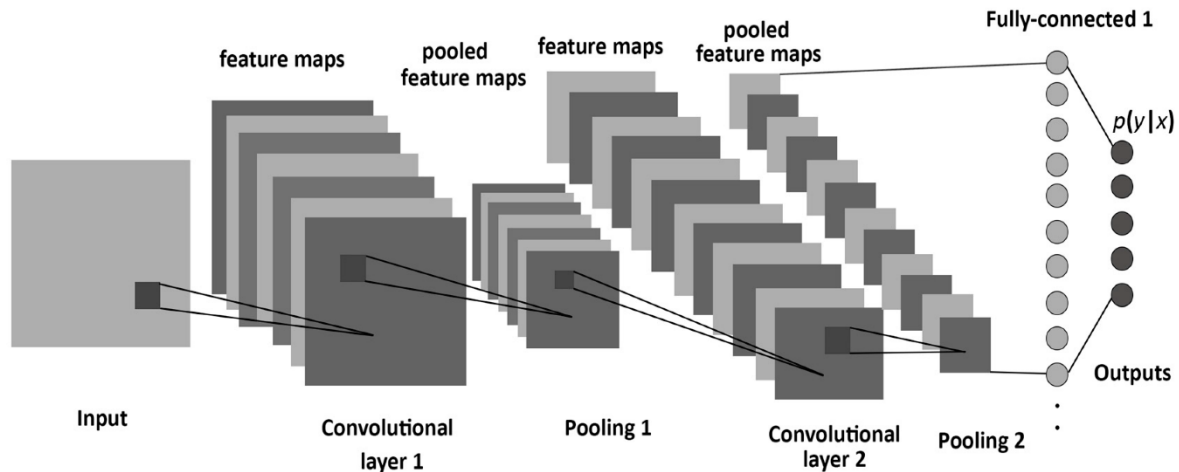
^۴ Max-Pooling

^۵ Dense

^۶ Stride

^۷ Batch Normalization

هر نورون لایه max-pooling به هر نورون لایه تماماً متصل وصل می‌شود که نتیجه خروجی سیگنال را پیش‌بینی می‌کند [۹۰، ۹۱].



شکل ۲-۱) ساختار شبکه عصبی کانولوشن

۲-۳-۳-۲ شبکه عصبی مکرر

شبکه عصبی مکرر [۹۲] از زیرمجموعه‌های شبکه‌های عمیق است. این شبکه‌ها امکان بازنمایی یک توالی با طول نامشخص را به یک بردار به طول ثابت فراهم می‌کنند. در این بازنمایی، خواص نحوی و ساختاری توالی ورودی حفظ می‌شود. به عبارت دیگر، شبکه‌های عصبی مکرر شامل یک حلقه بازگشتی هستند که اطلاعات در لحظات قبل یا اطلاعات گذشته را حفظ می‌کند. از سوی دیگر این حلقه باعث می‌شود تا اطلاعات ورودی از یک مرحله به مرحله بعد منتقل شوند. در این شبکه‌ها در هر مرحله زمانی t از $t = 1$ تا $t = n$ ، یک بردار $x^{(t)}$ از توالی ورودی $x = \langle x^{(1)}, x^{(2)}, \dots, x^{(n)} \rangle$ پردازش می‌شود. در حالت کلی، معادله‌های بروزسانی یک شبکه RNN در زمان t با استفاده از رابطه‌های (۲-۴) تا (۲-۷) انجام می‌شود.

$$a^{(t)} = Ux^{(t)} + Wh^{(t-1)} + b \quad (2-4)$$

$$h^{(t)} = \Phi(a^{(t)}) \quad (2-5)$$

$$o^{(t)} = Vh^{(t)} + c \quad (2-6)$$

$$\hat{y}^{(t)} = \text{softmax}(o^{(t)}) \quad (7-2)$$

در روابط (2-4) تا (2-7)، بردارهای b و c بایاس، و ماتریس‌های U ، V و W به ترتیب وزن یال‌های لایه ورودی به لایه پنهان، لایه پنهان به لایه خروجی و لایه پنهان به لایه پنهان هستند که پارامترهای شبکه را تشکیل می‌دهند. پارامتر Φ تابع انگیزش است. این پارامتر معمولاً به عنوان یکی از توابع ReLU یا Sigmoid انتخاب می‌شود. لایه آخر را نیز تابع Softmax تشکیل می‌دهد. این تابع، احتمال وقوع هر نشانه خروجی را مشخص می‌کند. یکی از مشکلات شبکه RNN، مشکل عدم یادگیری وابستگی بلند مدت است. شبکه LSTM نوع خاصی از شبکه‌های عصبی مکرر هستند. این نوع از شبکه‌ها، توانایی یادگیری وابستگی‌های بلند مدت را دارند [93].

۲-۴ کارهای مرتبط

در این بخش، به بررسی جدیدترین کارهای انجام‌شده در زمینه تحلیل و طبقه‌بندی سیگنال‌های EEG و fNIRS می‌پردازیم. تمرکز ما در این بخش روی دقت طبقه‌بندی و پیش‌پردازشی است که روی داده‌ها انجام شده است. تعداد زیادی از مقالات این حوزه، از الگوریتم‌های جدید یادگیری عمیق برای طبقه‌بندی داده‌های خود استفاده کرده‌اند. اما استفاده از این روش‌ها در برخی مقالات، دقت پایینی را به همراه داشته است. در ادامه به شرح تعدادی از مقالات مرتبط با حوزه طبقه‌بندی می‌پردازیم.

۲-۴-۱ ساختارهای شامل الگوریتم‌های طبقه‌بندی سنتی

در این بخش، به بررسی کارهایی می‌پردازیم که در آن‌ها از طبقه‌بندهای سنتی مانند SVM و LDA استفاده شده است. این روش‌ها شامل برخی الگوریتم‌های پیش‌پردازش برای حذف نویز و روش‌هایی برای استخراج ویژگی‌های مهم برای استفاده در این طبقه‌بندها هستند.

در مرجع [94]، یک طبقه‌بندی چند کلاسه، روی داده‌های fNIRS مبتنی بر MI بررسی شده است. ۱۰ فرد سالم برای حرکت دادن یک شیء با تصور حرکت استفاده شده‌اند. در این روش، از سیگنال‌های

قشر جلویی مغز^۱ استفاده شده است. در ابتدا از یک فیلتر باندگذر در محدوده فرکانسی ۰,۱ تا ۰,۲ هرتز استفاده شده است. سپس یک ترکیب از EEMD^۲ و ICA برای حذف نویز سیگنال استفاده شده است. ویژگی‌های میانگین سیگنال^۳ برای ورودی طبقه‌بندهای LDA و SVM استفاده شده است. دقت طبقه‌بندی با استفاده از SVM و در حالت HbO2، در پنجره ۸ تا ۲۱ ثانیه بالاتر از LDA بوده است. میانگین دقت در این روش برای چهار جهت بالا-پایین و چپ-راست به ترتیب ۴۰,۵، ۷۳ و ۷۰,۷ بوده است.

در مرجع [۶۱]، یک چهارچوب مبتنی بر رابط مغز کامپیوتر (BCI) بر اساس سیگنال‌های EEG و fNIRS، برای استفاده بیماران دارای ALS ارائه دادند. در این مقاله، اعمال مغزی، روش‌های انتخاب کانال، الگوریتم‌های استخراج ویژگی و الگوریتم‌های طبقه‌بندی مورد بررسی قرار گرفته‌اند. در ابتدا، انواع مختلف بیماری‌های شناختی و اختلالات حرکتی را برای ارزیابی مناسب سیستم BCI طبقه‌بندی کرده‌اند. قشر جلویی مغز، به عنوان ناحیه مناسب مغز برای تصور شناخته شده است. سپس، فعالیت مغزی تولیدکننده سیگنال‌های همودینامیک، بررسی شده‌اند. محاسبه ذهنی^۴ و تشکیل کلمه^۵، کارهای مناسب برای ثبت سیگنال از بیماران ALS است. سپس، از آنجایی که هدف، ناحیه مغزی خاصی است، روش‌های تعیین ناحیه برای BCI بررسی شده است. ویژگی‌های EEG و fNIRS مناسب برای سیستم BCI مورد بررسی قرار گرفته‌اند. برای یک BCI ترکیبی، ترکیب قله سیگنال، میانگین سیگنال‌های fNIRS و بالاترین توان باند سیگنال‌های EEG امیدوارکننده بوده است. برای طبقه‌بندی، LDA به صورت گسترده‌تر مورد استفاده قرار گرفته است. به طور کلی، تشخیص مناسب ناحیه و انتخاب مناسب ویژگی‌ها، دقت طبقه‌بندی را بهبود بخشیده است.

¹ Prefrontal

² Ensemble Empirical Mode Decomposition

³ Signal Average (SA)

⁴ Mental Calculation

⁵ Word Formation

در مرجع [۹۵]، با استفاده از hBCI، الگوهای سه فعالیت مغزی استخراج شده از محاسبات ذهنی، تصورات حرکتی و حالت استراحت را طبقه‌بندی کردند. این کار، با هدف بالا بردن نرخ انتقال اطلاعات با افزایش تعداد کلاس‌ها و حداقل کردن خطای طبقه‌بندی صورت گرفته است. الکترودهای EEG در قشر جلویی سر و مرکزی سر^۱ انتخاب شدند و منابع NIRS تنها در بخش پیشانی^۲ قرار داشتند. پیش-پردازش استفاده شده در این پایان‌نامه، با کاهش نرخ نمونه‌برداری داده EEG به ۲۰۰ هرتز برای کاهش بار محاسباتی و سپس اعمال فیلتر باندگذر بین ۰٫۱ تا ۵۰ هرتز بوده است. پیش‌پردازش برای داده NIRS نیز با تبدیل طول موج‌ها به تغییرات غلظت هموگلوبین و سپس اعمال یک فیلتر باندگذر باترورث^۳ با مرتبه شش بین ۰٫۰۱ تا ۰٫۰۹ بود. در این مقاله، مسئله طبقه‌بندی سه کلاس را به سه طبقه‌بندی باینری با استفاده از استراتژی طبقه‌بندی یک درمقابل یک^۴ تجزیه کردند. این کار به منظور اعمال فیلتر الگوهای مکانی رایج بانک فیلتر^۵ در داده EEG بود. از طبقه‌بند sLDA^۶ برای ارزیابی میانگین دقت‌های طبقه‌بندی برای BCI مبتنی بر EEG، BCI مبتنی بر NIRS و hBCI استفاده شده است. دقت‌های طبقه‌بندی سه کلاس برای BCI مبتنی بر EEG، BCI مبتنی بر NIRS و hBCI به ترتیب برابر ۷۶٫۱ با واریانس ۱۲٫۸، ۶۴٫۱ با واریانس ۹٫۷ و ۸۲٫۲ با واریانس ۱۰٫۲ درصد بود. دقت طبقه‌بندی روش hBCI بالاتر از دیگر روش‌ها بود. میانگین نرخ انتقال اطلاعات در روش hBCI پیشنهاد شده برابر ۴٫۷ با واریانس ۱٫۹۲ بیت بر دقیقه محاسبه شده است که ۳۴٫۳ درصد بالاتر از مطالعات گزارش شده قبلی در hBCI باینری بود.

در مرجع [۱۴]، از ترکیب EEG-fNIRS برای رمزگشایی از هشت دستور مغزی استفاده کرده‌اند. برای سیستم fNIRS از قشر جلویی مغز و برای EEG از اطراف قشرهای وابسته به پیشانی^۷، جداری^۸ و

^۱ Central

^۲ Forehead

^۳ Butterworth

^۴ One-Versus-One

^۵ Filter-bank common spatial patterns filter

^۶ shrinkage Linear Discriminant Analysis

^۷ Frontal

^۸ Parietal

دیداری^۱ استفاده شده است. در این مقاله، ریاضیات ذهنی، شمارش ذهنی، چرخش ذهنی و تشکیل کلمات کارهای مرتبط با ثبت سیگنال‌های fNIRS است. در سیگنال‌های ثبت شده fNIRS، قله سیگنال، حداقل مقدار و میانگین مقادیر HbO در پنجره‌های زمانی دو ثانیه‌ای، ویژگی‌های انتخاب شده برای طبقه‌بندی هستند. دوبار پلک زدن، سه بار پلک زدن و حرکت چشم در جهت‌های بالا-پایین و چپ-راست نیز چهار کار مرتبط با رمزگشایی سیگنال EEG است. در سیگنال‌های EEG، قله و میانگین سیگنال در پنجره‌های یک ثانیه‌ای، ویژگی‌های انتخاب شده برای طبقه‌بندی هستند. در این مقاله، از این هشت دستور تولید شده برای کنترل یک کوادکوپتر در یک فضای آزاد استفاده شده است. دقت میانگین برای چهار دستور مرتبط با سیگنال fNIRS برابر با ۷۵٫۶ درصد و برای چهار دستور دیگر مرتبط با EEG برابر ۸۶ درصد بوده است. همچنین مرجع [۹۶]، نیز روشی مبتنی بر fNIRS برای کنترل کوادکوپتر ارائه کرده است که کارایی کمتری از روش مرجع [۱۴] دارد.

در این روش‌ها، از طبقه‌بندی‌های سنتی برای طبقه‌بندی داده‌ها استفاده شده است. استفاده از این طبقه‌بندها نیازمند استخراج ویژگی‌های مفید است، به صورتی که این ویژگی‌ها بتوانند داده‌ها را به صورت کامل و قوی تمیز دهند. گاهی اوقات، ویژگی‌های استخراج شده مناسب نیستند یا داده‌ها را به درستی تمیز نمی‌دهند. در این گونه موارد، طبقه‌بند نیز قادر به جداسازی صحیح داده‌ها از یکدیگر نیست. در روش ما، ادغام مرحله استخراج ویژگی با شبکه عصبی، کمک شایانی به طبقه‌بندی داده‌ها کرده است. اما باید دقت داشت که این داده‌ها همپوشانی نداشته و قابل تمیز باشند.

۲-۴-۲ ساختارهای شامل الگوریتم‌های یادگیری عمیق

در بخش قبل، به بررسی تعدادی از روش‌ها که از طبقه‌بندهای سنتی استفاده کرده‌اند پرداختیم. در سال‌های اخیر، استفاده از یادگیری عمیق برای طراحی مدل‌هایی با دقت بالا به طور چشم‌گیری افزایش یافته است. الگوریتم‌های جدید شبکه‌های عصبی به نام یادگیری عمیق، طبقه‌بندی را بهبود بخشید. با

^۱ Visual

پیشرفت این تکنولوژی‌ها، در اکثر روش‌ها، و بدون توجه به مجموعه داده‌های مختلف، داده‌ها با دقت بیشتری طبقه‌بندی شدند. در ادامه، به برخی روش‌هایی که از یادگیری عمیق برای طبقه‌بندی استفاده کرده‌اند، اشاره می‌کنیم.

در مرجع [۹۷]، از LSTM برای بالا بردن دقت طبقه‌بندی سیگنال fNIRS استفاده کردند. سیگنال‌های fNIRS از قشر جلویی سر، از ثبت سه کار محاسبه ذهنی، شمارش ذهنی و حل پازل بدست آمده است. این روش با طبقه‌بندهای معمولی یعنی LDA و SVM و روش‌های استخراج ویژگی مختلف مقایسه شده است. در این روش، سیگنال fNIRS ابتدا توسط یک فیلتر پایین‌گذر با فرکانس قطع ۰٫۱۵ هرتز و یک فیلتر بالاگذر با فرکانس قطع ۰٫۰۱ هرتز فیلتر شده است. سپس یک ROI^۱ با استفاده از t-test با تابع پاسخ همودینامیک مورد نظر تعریف می‌شود. در ادامه از ROI ها برای هر شخص میانگیری شده است. سپس میانگین، شیب^۲، انحراف^۳، اوج سیگنال^۴ و قله سیگنال^۵ مربوط به داده ROI میانگین‌گیری شده برای هر شخص، با پنجره زمانی مشخص محاسبه می‌شود. در این مطالعه، پنجره بین دو تا هفت ثانیه در نظر گرفته شده است. میانگین دقت طبقه‌بندی برای LDA، برابر با ۳۷٫۵۰ درصد، برای SVM برابر ۳۷٫۹۶ درصد و برای LSTM برابر ۸۳٫۳ درصد می‌باشد.

در مرجع [۹۸]، تلاش کردند که دستورات تصورات حرکتی صادر شده توسط کاربر را دریافت کنند و به یک رایانه بدهند. آن‌ها می‌خواستند مطمئن شوند که رایانه می‌تواند دستور تصور حرکتی صادر شده را به درستی تشخیص دهد. این کار باید با صرف کمترین زمان ممکن در پردازش زمان واقعی انجام شود. برای این کار یک روش برای ثبت، فیلتر و پردازش سیگنال‌ها به صورت هوشمند طراحی شده است. در این روش، الکترودهای اصلی که با تصورات حرکتی سر و کار دارند، انتخاب شدند. این کار، به

^۱ Region Of Inter Interest

^۲ Slope

^۳ Skewness

^۴ Kurtosis

^۵ Peak

دلیل افزایش سرعت پردازش انجام شده است. روش طبقه‌بندی از انتخاب غیر سیستمی روی چهار آزمایش تصادفی برای هر حرکت استفاده شده است که به عنوان داده تست در نظر گرفته شده است. داده آموزشی شامل ۱۴ داده دیگر است. این کار بارها تکرار شده است. طبقه‌بندهای استفاده شده SVM و LDA هستند. در این مقاله، بعد از حذف الکترودهای اضافی، یک فیلتر باندهای حذف سیگنال-های آلوده استفاده شده است. نتایج نشان می‌دهد که روش پیشنهاد شده در این مقاله، دقت تشخیص تصورات حرکتی را افزایش می‌دهد. حذف اطلاعات اضافی باعث افزایش سرعت پردازش در این سیستم شده است.

در مرجع [۹۹]، یک سیستم تشخیص احساسات فرد محور را توسعه داده است. در این مطالعه از مجموعه داده DEAP بهره گرفته شده است. در این تحقیق، از یک شبکه عمیق با ساختار ساده برای طبقه‌بندی low-high valence و low-high arousal استفاده شده است. این شبکه، یک DNN با سه لایه مخفی است. در این تحقیق از تبدیل موجک برای استخراج ویژگی و باندهای فرکانسی مختلف استفاده شده است. ترکیب باند آلفا در الکتروود Fp1 و باند تتا در الکتروود F4 مورد توجه بوده است. دقت طبقه‌بندی valence برابر ۶۲٫۵ و برای arousal برابر ۶۴٫۲۵ است. این روش در مقابل روش‌های دیگر افزایش چشم‌گیری داشته است.

در مرجع [۱۰۰]، یک سیستم BCI با کمک OpenBCI، در یک محیط طبیعی بازخورد عصبی قدرتمند و با استفاده از یادگیری عمیق، برای بهبود عملکرد ارزیابی کردند. شرکت‌کنندگان این مطالعه، افراد سالمی بودند که از ریتم Mu آن‌ها روی قشر حسی-حرکتی استفاده شده است. در این مطالعه، تصور حرکتی دست راست و حالت استراحت انجام شده است. در این مقاله، سیگنال EEG را با کمک فیلتر باندهای باترورث بین ۰٫۱ تا ۱۰۰ هرتز، فیلتر کردند. سپس، یک فیلتر notch در ۵۰ هرتز برای حذف نویز برق شهری اعمال شده است. بعد، یک روش پنجره‌گذاری لغزان، برای تقسیم کردن داده به بخش-های کوچک‌تر با اندازه پنجره W (اندازه‌های مختلف از ۱ تا ۹) و قدم چهار میلی ثانیه اعمال شد. در

مرحله بعد، سیگنال را با استفاده از تبدیل فوریه توان طیفی در محدوده باند فرکانسی مو استخراج می‌کنند. در این مقاله، از شبکه پرسپترون تماماً متصل چند لایه برای طبقه‌بندی استفاده کرده‌اند. در این ساختار از چهار لایه تماماً متصل و یک لایه خروجی استفاده شده است. Shivanthan و همکارانش نشان دادند که روش آن‌ها بهتر از BCI‌های مبتنی بر MI در سیستم OpenBCI قبلی است. کمترین دقت برابر ۵۵,۴۶ درصد برای اندازه پنجره یک و بالاترین دقت برابر ۹۸,۹۶ درصد برای اندازه پنجره نه بدست آمده است. در مجموع، میانگین دقت تمام داوطلبان این مقاله، زمانی که اندازه پنجره برابر هفت بود، ۸۵,۲۵ درصد بدست آمده است.

در مرجع [۱۰۱]، یک روش با استفاده از شبکه‌های عصبی کانولوشن و یادگیری residual عمیق، برای پیش‌بینی حالت‌های ذهنی رانندگان از روی سیگنال‌های EEG ارائه دادند. از این رو، در این مقاله، دو مدل طبقه‌بند حالت ذهنی را با نام‌های EEG-Conv و EEG-Conv-R توسعه دادند. نویز چشمی سیگنال‌های EEG که ثبت می‌شدند، توسط ICA حذف شده‌اند. سپس یک فیلتر باند گذر بین ۱ تا ۴۰ هرتز اعمال شده است. از نرمال‌سازی z-score نیز برای نرمال‌سازی داده‌ها استفاده شده است. نتایج نشان‌دادند که هر دو مدل عملکرد بسیار خوبی را در طبقه‌بندی حالات ذهنی دارند، EEG-Conv-R برای پیش‌بینی حالت ذهنی بین افراد مناسب‌تر است و از مدل دیگر زودتر همگرا می‌شود.

در مرجع [۱۰۲]، حرکت ترکیبی بازوی راست و چپ را با استفاده از طبقه‌بند SVM طراحی کرده‌اند. در این مطالعه، از سیگنال‌های fNIRS استفاده شده است. داده‌های fNIRS ابتدا از حالت طول موج با کمک قانون لمبرت بیر اصلاح شده^۱ به تغییرات غلظت هموگلوبین تبدیل شده است. بعد از آن، این سیگنال‌ها به کمک روش فیلتر باترورث با فرکانس‌های قطع ۰,۱ تا ۰,۲ هرتز فیلتر شده‌اند. هدف آن‌ها، بررسی اثربخشی دو ویژگی مختلف، یعنی دامنه متوسط سیگنال و تفاوت بین دامنه‌های متوسط سیگنال DHbO و DHbR روی دقت بدست آمده بود. این اختلاف از فاصله زمانی ۶ تا ۱۲ ثانیه بعد از

^۱ Modified Beer Lambert Law

شروع ثبت، محاسبه شده است. ترکیب این دو ویژگی نیز مورد بررسی قرار گرفته است. ویژگی‌های استخراج شده به روش نرمال‌سازی min-max بین دامنه -۱ تا ۱ نرمال‌سازی شده‌اند. نتایج حاصل از این تحقیق نشان‌دادند که ترکیب ویژگی‌های انتخاب شده، دقت میانگین بین ۷۰ تا ۷۶٫۶۷ درصد را برای همه افراد محاسبه کرده است. تفاوت میانگین پارامترهای DHbO و DHbR به عنوان یکی از ویژگی‌های کاربردی مورد بررسی قرار گرفت که دقت متوسط ۷۰ درصد را نشان داد. این نتایج نشان می‌دهند که الگوهای متمایز در پاسخ‌های همودینامیکی ناشی از حرکات دست راست و چپ می‌توانند برای توسعه BCI مورد استفاده قرار گیرند.

در مرجع [۱۰۳]، اثر بخشی روش‌های پیشرفته مانند یادگیری عمیق برای تشخیص دوره‌های استراحت و عمل را بررسی کردند. نويز موجود در داده‌های fNIRS ثبت شده توسط یک فیلتر باندگذر بین ۰٫۳ و ۳ هرتز حذف شده است. سپس داده روی کانال مرجع میانگیری شده و نرمال‌سازی می‌شود. بعد از آن، داده‌ها را در اندازه‌های برابر (پنجره‌هایی به اندازه ۱۰۰۰ نمونه)، بدون همپوشانی بخش‌بندی کرده‌اند. در ابتدا، تغییرات ظاهری موجود در غلظت‌های HbO و HbR را در هر دو مرحله ذهنی، برای هر یک از شرکت‌کنندگان پیدا کردند. سپس یک روش برای تشخیص آن‌ها ارائه دادند. در این مقاله، از یک شبکه عصبی عمیق به نام شبکه باور عمیق استفاده شده است. شبکه آن‌ها شامل پنج لایه است. شبکه‌ی باور عمیق^۱ دارای دقت طبقه‌بندی حدود ۸۴٫۲۵ با واریانس ۹٫۱۰ درصد است.

در مرجع [۱۰۴]، تکنیکی برای ترکیب EEG و fNIRS در یک سیستم BCI طراحی کرده‌اند. این کار برای رفع برخی مشکلات سیستم‌های BCI از جمله رفع نویز و بالا بردن دقت طبقه‌بندی، برای سیستم‌های مبتنی بر EEG انجام شده است. این سیستم مبتنی بر عمل مشت کردن توسط شرکت‌کننده است. با توجه به نتایج طبقه‌بندی، ویژگی ترکیبی بر اساس ضریب موجک EEG و شیب fNIRS پیشنهاد شده است. نتایج نشان داده‌اند که نرخ تشخیص متوسط کار مشت کردن بعد از ویژگی ترکیبی

^۱ Belief Deep Network (BDN)

از سه درصد به نه درصد بالاتر از ویژگی‌های EEG و fNIRS به تنهایی بوده است. این نشان می‌دهد که fNIRS می‌تواند به طور قابل توجهی عملکرد رابط مغز کامپیوتر را افزایش دهد و کارایی آن را بهبود بخشد.

در مرجع [۶۲]، با استفاده از روش‌های یادگیری عمیق، عملکرد سیستم ترکیبی BCI مبتنی بر ترکیب EEG و fNIRS را بهبود بخشیدند. در این مقاله، طبقه‌بندی دو تصور حرکتی دست چپ و راست بررسی شده است. این آزمایش بر روی ۱۵ شخص انجام شده است. زمان هر ثبت، یک ثانیه و طول آزمایش ۱۰ دقیقه بوده است. سیگنال‌های ثبت شده از دستگاه EEG، در پنجره‌های یک ثانیه‌ای تقسیم شده و سپس بین فرکانس‌های ۸ تا ۳۰ هرتز فیلتر شده‌اند. با مربع کردن سیگنال در فرکانس‌های مرتبط با مو و بتا، توان سیگنال ارزیابی شده است. سپس ویژگی‌های ERD/ERS از آن استخراج شده است. سیگنال‌های ثبت شده از دستگاه fNIRS با استفاده از قانون لمبرت بیر اصلاح شده، به تغییرات غلظت oxy و deoxy هموگلوبین تبدیل شده‌اند. در ادامه این مقاله، با استفاده از یک شبکه DNN دقت طبقه‌بندی محاسبه شده است. شبکه عصبی آن‌ها، یک شبکه تماماً متصل با پنج لایه است و دقت محاسبه شده از این شبکه با استفاده از ویژگی‌هایی که استخراج شده برابر ۸۳,۲۸ درصد بوده است.

در مرجع [۱۰۵]، سیگنال‌های EEG-fNIRS را با استفاده از کارهای محاسبات ذهنی و زنجیره کلمات بدون آموزش اولیه کاربران، به صورت همزمان ثبت کردند و سپس آن‌ها را در حالت استراحت (baseline) قرار دادند. هر کدام از اعمال برای پنج ثانیه ثبت شده است. این زمان نسبت به مطالعات hBCI دیگر کمتر بوده است. سپس، یک طبقه‌بندی برای تمیز دادن فعالیت‌های مغزی انجام شده مبتنی بر MA و WC از BL انجام شده است. با استفاده از hBCI در تجزیه و تحلیل‌های آفلاین/pseudo آنلاین، دقت‌های میانگین ۹۰ با واریانس ۷,۱ درصد / ۸۵,۵ با واریانس ۸,۱ درصد و دقت‌های ۸۵,۸ با واریانس ۸,۶ درصد / ۷۹,۵ با واریانس ۱۳,۴ درصد به ترتیب برای MA در مقابل BL و WC در مقابل BL بدست آمده است. این موارد، در اکثر آزمایشات به طور قابل توجهی بالاتر از BCI‌های

مبتنی بر EEG یا fNIRS بوده است. با توجه به طول کوتاه آزمایش و بهبود دقت طبقه‌بندی، میانگین نرخ انتقال اطلاعات در مقابل BL، برای MA به بیش از ۹۶,۶ درصد و برای WC به بیش از ۸۷,۱ درصد در مقایسه با مطالعات گزارش شده قبلی بهبود یافته است.

در این روش‌ها از یادگیری عمیق استفاده شده است. در همه این روش‌ها، مرحله یادگیری وابسته به مرحله استخراج ویژگی است. استخراج ویژگی در الگوریتم‌های یادگیری عمیق، باعث کاهش اطلاعات مورد نیاز برای آموزش می‌شود. از این رو، ممکن است دقت طبقه‌بند را به طور قابل توجهی کاهش دهد. اما این را نیز باید در نظر داشت که مهندسی ویژگی موجود در یادگیری عمیق زمانی می‌تواند به درستی عمل کند که هر کلاس داده از دیگری متمایز باشد، یعنی داده‌های کلاس‌ها با یکدیگر همپوشانی نداشته باشد.

۲-۴-۳ ساختارهای شامل یادگیری عمیق به همراه استفاده از ویژگی-

های زمانی و مکانی

در هیچکدام از این روش‌ها، از اطلاعات مکانی و زمانی موجود در سیگنال‌ها استفاده نشده است. بیشتر ویژگی‌های استفاده شده، ویژگی‌های آماری هستند. در این مقاله، ما از این اطلاعات برای طبقه‌بندی سیگنال‌ها استفاده کرده‌ایم. در سال‌های اخیر، تنها دو مقاله از این اطلاعات برای طبقه‌بندی داده‌ها استفاده کرده‌اند که در ادامه به آن‌ها می‌پردازیم.

در مرجع [۶۰]، با استفاده از مدل‌های شبکه‌های عصبی به صورت موازی و متوالی دقت طبقه‌بندی اهداف ذهنی انسان را بررسی کرده است. در این روش، از ویژگی‌های زمانی و مکانی سیگنال EEG استفاده شده است. در مدل پیشنهادی این پایان‌نامه، همبستگی مکانی بین همسایگی‌های سیگنال‌های EEG را به سلسله مراتبی از سیگنال‌های دو بعدی شبکه مانند تبدیل کرده است. یک شبکه LSTM برای استخراج وابستگی‌های زمانی مناسب از داده‌های EEG استفاده شده است. مجموعه داده استفاده شده در این پایان‌نامه، مبتنی بر MI-EEG با ۱۰۸ فرد و حدود ۳,۱۴۵,۱۶۰ نمونه است. هر دو مدل

موازی و متوالی شبکه عمیق، دقت بالا و حدود ۹۸,۳ درصد به دست آورده است. این نتایج نشان‌دهنده افزایش ۱۸ درصدی این سناریو در مقایسه با سناریوهای مشابه می‌باشد.

در مرجع [۵۹]، تلاش کردند تا مشکل کمبود اطلاعات کافی از جمله اطلاعات مکانی و زمانی را در سیگنال‌های EEG حل کنند. برای حل این مشکل، از یک شبکه LSTM به همراه یک شبکه کانولوشن استفاده کردند. با این کار، اطلاعات زمانی و مکانی را به صورت همزمان به سیگنال‌های EEG مبتنی بر تصورات حرکتی اعمال می‌کنند. علاوه بر این، نمایشات طیفی سیگنال‌های EEG به وسیله تجزیه و تحلیل تبدیل موجک گسسته بدست آمده است. برای دستیابی به نرخ یادگیری بالاتر، قبل از آموزش از روش نرمال‌سازی دسته‌ای استفاده شده است. چهار روش برای پیش‌پردازش داده مورد استفاده قرار گرفته است. ابتدا مقدار سیگنال بدست آمده از الکتروود Cz که به عنوان الکتروود مرجع در نظر گرفته شده است از دیگر کانال‌ها کم می‌شود. سپس نه الکتروود مرتبط با EEG مبتنی بر MI انتخاب می‌شوند. با استفاده از الگوریتم‌های ICA و PCA، نویزهای موجود در سیگنال EEG حذف می‌شوند. در نهایت، با استفاده از روش فیلتر باندگذر، فرکانس‌های مو و بتا استخراج می‌شوند. برای استخراج ویژگی از تبدیل فوریه، تبدیل موجک و CSP استفاده شده است. در میان این روش‌ها، از DWT برای استخراج دنباله‌های زمان فرکانس استفاده شده است. آزمایشات مختلفی برای ارزیابی عملکرد معماری شبکه یادگیری عمیق انجام شده است. نتایج نشان می‌دهد که این روش، یک مدل مفید و قوی برای طبقه‌بندی چند کلاسه در میان بسیاری از روش‌های مختلف می‌باشد. مدل شبکه یادگیری عمیق پیشنهاد شده در این پایان‌نامه، با الگوریتم‌های طبقه‌بند معمولی یعنی RF، SVM، NB، KNN، LDA حتی LSTM و CNN به صورت مجزا مقایسه شده است. در این روش، بهترین دقت در هنگام استفاده از CNN-LSTM با استفاده از تبدیل موجک گسسته بدست آمده است. این دقت تقریباً برابر ۸۷ درصد می‌باشد.

۲-۵ نتیجه‌گیری

تنها تفاوت مرجع [۶۰] با کار ما، در استفاده از تک سیگنال برای طبقه‌بندی بود. ما سعی کردیم با افزودن یک داده جدید، یعنی fNIRS به سیگنال EEG و استفاده از هر دو اطلاعات مکانی و زمانی، دقت را بهبود بخشیم. اما در کار دوم، با اینکه از اطلاعات مکانی و زمانی استفاده شده است، اما دقت پایین است، که این ممکن است دلایل مختلفی داشته باشد. یکی از این مشکلات ممکن است به دلیل استفاده از داده‌های فرکانسی در مرحله طبقه‌بندی باشد. در این مقاله، سعی شده است تا با استخراج ویژگی‌های زمانی و مکانی در داده‌های EEG در حوزه زمان-فرکانس، آن‌ها را پیش‌پردازش و طبقه‌بندی کنند. به نظر می‌آید که مشکل در همینجا است. داده‌ها در حوزه فرکانس، ویژگی زمانی خود را از دست می‌دهند.

جدول ۲-۱، روش‌های شرح داده شده در مقالات را به طور خلاصه مقایسه می‌کند.

جدول ۲-۱) مقایسه روش‌های ذکر شده.

مرجع	روش پیشنهادی هر مرجع
مراجع با طبقه‌بندی سنتی	
[۹۴]	در این مرجع از طبقه‌بندی‌های SVM و LDA، برای طبقه‌بندی داده‌های fNIRS استفاده شده است. بالاترین دقت بدست آمده در این مقاله، ۷۳ درصد است.
[۶۱]	از طبقه‌بند LDA، برای طبقه‌بندی سیگنال‌های EEG و fNIRS استفاده کرده است. در این مقاله، سعی در تشخیص برخی بیماری‌ها با استفاده از این سیگنال‌ها داشتند.
[۹۵]	از طبقه‌بند sLDA برای طراحی یک سیستم hBCI استفاده شده است. در این مقاله، سعی داشتند تا نرخ انتقال اطلاعات را تا حد زیادی در این سیستم‌ها افزایش دهند. دقت طبقه‌بندی برابر با ۸۲٫۲ و نرخ انتقال اطلاعات نیز ۴٫۷ بیت بر دقیقه است.

[۱۴]	یک سیستم با هشت دستور مغزی را برای کنترل یک کوادکوپتر طراحی کردند. در این مقاله، از هر دو سیگنال EEG و fNIRS استفاده شده است. از طبقه‌بند LDA، برای تشخیص دستورات استفاده شده است. طراحی چنین سیستمی، برای کنترل کوادکوپتر در مرجع [۹۶] نیز آمده است که دقت پایین‌تری نسبت به این مرجع دارد. در این مقاله، دقت طبقه‌بندی بر اساس داده EEG برابر با ۸۶ درصد و برای fNIRS برابر با ۷۵٫۶ درصد است.
[۱۰۲]	از طبقه‌بند SVM، برای طبقه‌بندی سیگنال‌های fNIRS استفاده شده است. دقت میانگین بین ۷۰ تا ۷۶٫۶۷ درصد بوده است. این دقت با توجه به دو ویژگی استخراج شده پیشنهادی در این مقاله بدست آمده است. تفاوت میانگین پارامترهای DHbO و DHbR به عنوان یک ویژگی کاربردی بررسی شد که دقت متوسط ۷۰ را بدست آورد.
مراجع با طبقه‌بندهای شامل یادگیری عمیق	
[۹۷]	روشی مبتنی بر LSTM برای طبقه‌بندی سیگنال‌های fNIRS ارائه دادند. در این مقاله، روش LSTM با استخراج ویژگی‌های مختلف با روش‌های سنتی SVM و LDA مقایسه شده است. در این روش، بالاترین دقت بدست آمده برابر ۸۳٫۳ درصد است.
[۹۹]	از یک شبکه DNN برای طبقه‌بندی داده‌های احساسات ثبت شده با سیگنال EEG استفاده کرده‌اند. شبکه آن‌ها شامل سه لایه مخفی است. در این مقاله، از باند آلفا برای کانال Fp1 و باند تتا برای کانال F4 استفاده شده است. طبقه‌بندی داده‌های مرتبط با احساسات همیشه چالشی بزرگ بوده است. دقت بدست آمده در این مقاله، برابر با ۶۴٫۲۵ درصد است.
[۱۰۰]	از یک شبکه عمیق با چهار لایه تماماً متصل برای طبقه‌بندی داده‌های EEG استفاده کرده‌اند. در این مقاله، از روش پنجره‌گذاری لغزان جهت تقسیم سیگنال به بخش‌های کوچک‌تر استفاده شده است. در نتیجه، شبکه آن‌ها، داده‌ها را با دقت ۸۵٫۲۵ درصد طبقه‌بندی کرد.

[۱۰۱]	یک شبکه residual عمیق را برای طبقه‌بندی داده‌های EEG معرفی کردند. در این مقاله، نویز چشمی با استفاده از ICA حذف شده است.
[۱۰۳]	در این مقاله، یک طبقه‌بند جدید به عنوان شبکه باور عمیق، برای طبقه‌بندی داده‌های fNIRS ارائه شده است. در این روش، از روش پنجره‌گذاری لغزان، برای تبدیل داده‌ها به بخش‌های کوچک‌تر استفاده شده است. دقت بدست آمده با استفاده از این شبکه برابر با ۸۴,۲۵ درصد است.
[۶۲]	یک شبکه DNN برای طبقه‌بندی هر دو داده EEG و fNIRS ارائه شده است. شبکه آن‌ها شامل پنج لایه تماماً متصل است. دقت بدست آمده با استفاده از ویژگی‌های استخراج شده پیشنهادی در این مقاله، برابر با ۸۳,۲۸ درصد است.
مراجع با طبقه‌بندهای یادگیری عمیق با استفاده از ویژگی‌های مکانی و زمانی داده‌ها	
[۶۰]	در این مقاله، یک روش پیش‌پردازشی جدید برای داده‌های EEG ارائه شد. در این مقاله، داده‌های EEG به دنباله‌ای از کلیپ‌های داده، شامل اطلاعات مکانی و زمانی تبدیل شده‌اند. سپس، یک روش یادگیری عمیق مبتنی بر ترکیب CNN و RNN، برای طبقه‌بندی داده‌ها استفاده شد. دقت بدست آمده برای این روش، برابر با ۹۸,۳ درصد بود.
[۵۹]	در این مقاله، از ترکیب شبکه‌های CNN و LSTM برای طبقه‌بندی داده‌های EEG استفاده کردند. در این روش، از روش Referencing و سپس ICA و PCA، برای حذف نویز استفاده شده است. سپس با استفاده از یک تبدیل موجک گسسته، ویژگی‌های فرکانس-زمان را از داده‌های EEG استخراج کردند. دقت بدست آمده برای این روش، برابر با ۸۷ درصد است.

در این پایان‌نامه سعی کردیم با در نظر گرفتن ویژگی‌های زمانی و مکانی موجود در داده‌ها و همچنین ترکیب کردن دو سیگنال مختلف EEG و fNIRS و رفع چالش‌های موجود در آن، دقت تشخیص اهداف ذهنی را به طور چشم‌گیر و قابل توجهی بالا ببریم. برای این منظور، در فصل بعد، به بررسی الگوریتم

پیشنهادی و چگونگی ترکیب دو سیگنال و استخراج اطلاعات زمانی و مکانی موجود در این دو می‌پردازیم.

فصل ۳ : ارائه روش پیشنهادی

۳-۱ مقدمه

در این فصل، در ابتدا به شرح روش پیشنهادی برای آماده‌سازی سیگنال‌ها، جهت طبقه‌بندی پرداخته شده است. سپس، به شرح روشی مبتنی بر یادگیری عمیق جهت طبقه‌بندی این داده‌ها پرداخته شده است. هر سیستم BCI به طور معمول شامل پنج مرحله اصلی است. این پنج مرحله، شامل کسب داده، پیش‌پردازش، استخراج ویژگی، طبقه‌بندی و بازخورد به کاربر است. سه مرحله پیش‌پردازش، استخراج ویژگی و طبقه‌بندی در برخی از مراجع در قالب یک مرحله پردازش معرفی می‌شود. در این پایان‌نامه، تمرکز اصلی بر روی دو مرحله پیش‌پردازش داده و طبقه‌بندی آن‌ها است. در این فصل، ابتدا مراحل پیش‌پردازش بر روی دو سیگنال EEG و fNIRS بیان شده است. روش پیش‌پردازش پیشنهادی، شامل تبدیل سیگنال‌های خام به تعدادی تانسور سه بعدی است که محتوای فعالیت مغزی را در هر لحظه از زمان نمایش می‌دهد. این تانسورها شامل دو ویژگی مکانی و زمانی در مغز هستند. در ادامه این فصل و در بخش دوم، به ارائه یک الگوریتم برای طبقه‌بندی داده‌ها پرداخته شده است. این الگوریتم پیشنهادی مبتنی بر یادگیری عمیق است. با کمک این الگوریتم، می‌توان فعالیت‌های مغزی را به طور مؤثر و با دقت بالا شناسایی کرد.

۳-۲ پیش‌پردازش: آماده‌سازی دادگان برای تحلیل‌های بعدی

همانطور که به طور خلاصه در بخش ۱-۳-۱ شرح داده شد، سیگنال EEG را می‌توان به چندین محدوده فرکانسی تقسیم کرد. هر محدوده فرکانسی شامل اطلاعات خاصی است. فعالیت‌های مغزی مرتبط با تصورات حرکتی، در باندهای فرکانسی ۸ تا ۳۰ هرتز (باندهای آلفا و بتا) در سیگنال EEG ظاهر می‌شوند. از این رو، می‌توان با استفاده از یک فیلتر میان‌گذر با فرکانس قطع ۸ تا ۳۰ هرتز، دیگر فرکانس‌ها را از سیگنال حذف نمود. اما مشکل اصلی برای پیش‌پردازش داده‌های EEG و fNIRS به صورت همزمان، نرخ نمونه‌برداری ناسازگار این دو سیگنال است. در مجموعه داده جدید، نرخ نمونه‌برداری داده fNIRS برابر ۱۰ هرتز و نرخ نمونه‌برداری داده EEG برابر ۲۰۰ هرتز است. در واقع، تعداد نمونه‌های

داده EEG و fNIRS در یک ثانیه به ترتیب برابر ۲۰۰ و ۱۰ است، که این امر مانع از تطبیق مناسب و پردازش همزمان این دو داده می‌شود. از این رو، این دو نوع داده باید قبل از انجام هرگونه پردازشی سازگار شوند. برای سازگار کردن این دو داده، نرخ نمونه برداری داده EEG به ۱۲۸ هرتز کاهش و نرخ نمونه برداری داده fNIRS به ۱۲۸ هرتز افزایش می‌یابد. باید توجه داشت که کاهش نرخ نمونه برداری داده‌های EEG باعث حذف اطلاعات از این داده‌ها نمی‌شود. همچنین، با تنظیم کاهش نمونه برداری سیگنال EEG، بار پردازشی این سیگنال‌ها به طور قابل توجهی کاهش می‌یابد. از این رو، با در نظر گرفتن قانون نایکوئیست برای این سیگنال‌ها EEG، می‌توان فهمید که نرخ نمونه برداری ۱۲۸ برای این سیگنال، اطلاعات را حفظ می‌کند اما حجم داده را کاهش می‌دهد. به این دلیل که بالاترین فرکانس داده EEG به اندازه ۶۰ می‌تواند باشد که اطلاعات را برای باندهای مختلف فرکانسی حفظ می‌کند، از این رو با توجه به قانون نایکوئیست باید فرکانس ۱۲۰ تمامی اطلاعات را در خود دارد. در این پایان‌نامه، کاهش و افزایش نرخ نمونه برداری با استفاده از تبدیل فوریه انجام شده است.

حال می‌توان برای حذف برخی از نویزهای موجود در سیگنال، همانطور که در بخش ۱-۱-۲ اشاره شد، از یک فیلتر میان‌گذر باترورث استفاده کرد. سیگنال EEG بین محدوده فرکانسی ۰٫۵ تا ۳۵ هرتز و سیگنال fNIRS بین محدوده فرکانسی ۰٫۱ تا ۰٫۵ هرتز فیلتر می‌شود. با این کار، برخی از نویزها و سیگنال‌های ناخواسته متأثر از اثر برق شهری، حرکت چشم و ضربان قلب از بین می‌روند.

۳-۳ آماده‌سازی داده‌ها

انجام فرآیندهای پیش‌پردازشی مختلف روی داده‌ها و سیگنال‌های EEG و همچنین fNIRS، همیشه چالش برانگیز بوده است. همیشه نمی‌توان به طور قطع گفت که الگوریتم‌های پیش‌پردازشی برای آماده‌سازی داده، برای تجزیه و تحلیل مناسب هستند. در این بخش می‌خواهیم روشی برای پیش‌پردازش و آماده‌سازی داده‌های EEG و fNIRS ارائه دهیم که نتایج مطلوبی را در طبقه‌بندی این داده‌ها و تشخیص اهداف ذهنی به همراه داشته است. در این پایان‌نامه، برای تبدیل داده‌های یک بعدی به

تانسورهای سه بعدی و در نهایت، آماده‌سازی این داده‌ها به عنوان ورودی مدل پیشنهادی، دو مرحله پیشنهاد می‌شود. این مراحل، شامل تبدیل سیگنال‌های EEG و fNIRS به تانسورهای سه بعدی و استفاده از پنجره لغزان است.

۳-۳-۱ تبدیل سیگنال‌های یک بعدی به تانسورهای سه بعدی

استفاده از سیگنال‌های یک بعدی همیشه چالشی بوده که توجه محققان را به خود جلب کرده است. همانطور که قبلاً گفته شد، یکی از مشکلات سیگنال‌های بیولوژیکی مبهم بودن آن است. به این معنا که، از روی ظاهر سیگنال نمی‌توان تمام ویژگی‌های نهفته درون آن را تشخیص داد. به همین دلیل محققان شروع به استخراج ویژگی‌های مختلف برای متمایز کردن سیگنال‌ها و اهداف مغزی مختلف انسان کردند. در این میان اما کمتر کسی به موقعیت الکترودهای مربوط به سیگنال‌هایی مانند EEG و fNIRS توجه کرده است.

در اکثر تصاویر مغز، مخصوصاً تصاویر مرتبط با توپوگرافی مغز، فعالیت مغزی محدود به یک ناحیه کوچک نمی‌شود. به عنوان مثال، زمانی که شخصی دست راست خودش را حرکت می‌دهد، سمت چپ مغزش فعال می‌شود. ولی این فعالیت محدود به تنها یک الکتروود نیست بلکه ناحیه گسترده‌تری از سمت چپ مغز او فعال است. این فعالیت در یک ناحیه شدیدتر و در دیگر نواحی ضعیف‌تر است. اما تنها محدود به یک الکتروود نیست. با این توضیحات می‌توان نتیجه گرفت که استفاده از ویژگی مکانی و استخراج آن از اهمیت زیادی برخوردار است.

به‌طور معمول، داده‌های خام ثبت‌شده با استفاده از سیستم ثبت EEG، n کاناله در زمان t را می‌توان با یک بردار n تایی به صورت $e_t = [S_t^1, S_t^2, \dots, S_t^n]^T$ نشان داد، که در آن e_t برابر سیگنال EEG در زمان t و S_t^n برابر نمونه ثبت شده توسط دستگاه ثبت سیگنال EEG در زمان t و از کانال n م آن دستگاه است. به طور مشابه، داده‌های خام ثبت‌شده با استفاده از سیستم ثبت fNIRS در زمان t نیز

یک بردار شامل دو مقدار اکسی^۱ و دئوکسی^۲ هموگلوبین است. این بردار را می‌توان به صورت $f_t = [(E_t^1, E_t'^1), (E_t^2, E_t'^2), \dots, (E_t^m, E_t'^m)]^T$ نمایش داد که در آن f_t برابر سیگنال fNIRS در زمان t ، E_t^m برابر نمونه اکسی هموگلوبین ثبت‌شده از دستگاه fNIRS در زمان t و از کانال m آن دستگاه و $E_t'^m$ برابر نمونه اکسی هموگلوبین ثبت‌شده از دستگاه fNIRS در زمان t و از کانال m آن دستگاه است. تمام پارامترهای استفاده شده در این پایان‌نامه به صورت خلاصه در جدول ۱-۳ شرح داده شده است.

جدول ۱-۳) شرح پارامترها.

پارامتر	توضیح
e_t	بردار مرتبط با سیگنال EEG در زمان t
s_t^n	نمونه داده EEG خوانده شده از n امین کانال و در زمان t
f_t	بردار مرتبط با سیگنال fNIRS در زمان t
E_t^m	نمونه داده اکسوهموگلوبین در داده‌های fNIRS از m امین کانال و در زمان t
$E_t'^m$	نمونه داده دئوکسی هموگلوبین در داده‌های fNIRS از m امین کانال و در زمان t
n	تعداد کانال EEG
m	تعداد کانال fNIRS
m_t	شبکه‌های سه بعدی
SW_j	کلیپ‌های ایجاد شده از شبکه‌های سه بعدی در یک پنجره زمانی
SF_j	ویژگی‌های مکانی استخراج شده از شبکه CNN
TF_j	ویژگی‌های مکانی استخراج شده از شبکه RNN
h_t	حالت مخفی لایه LSTM در زمان t
x_t	ورودی در زمان t

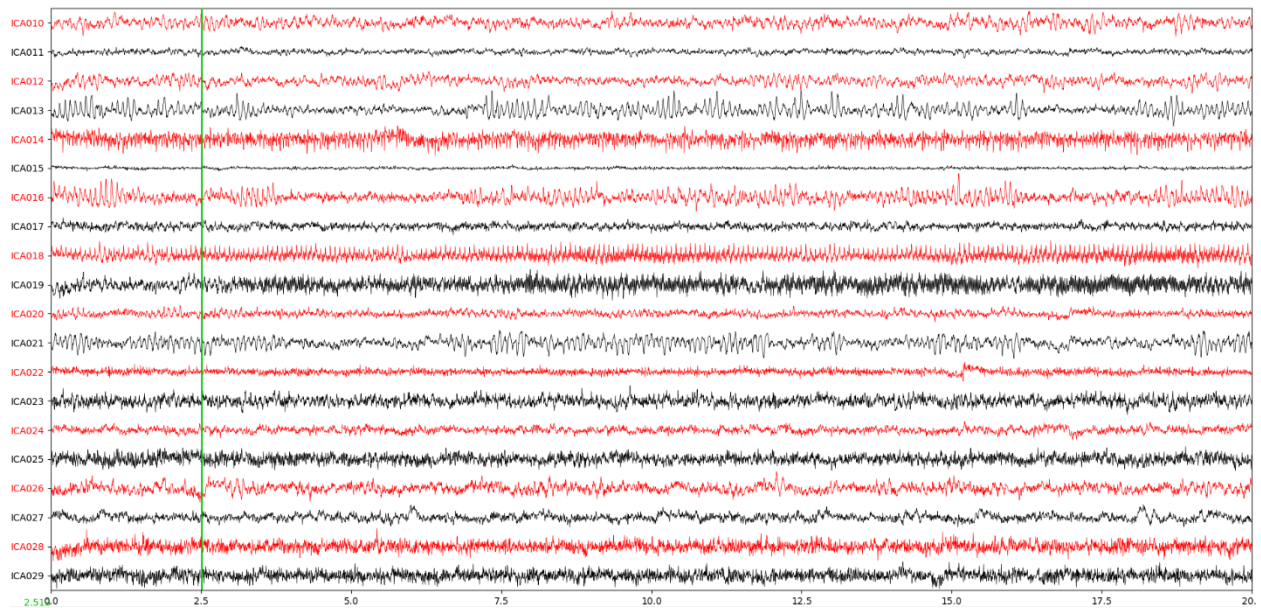
^۱ Oxygenated-hemoglobin

^۲ Deoxygenated-hemoglobin

معمولاً به طور پیش فرض، هر سیگنال از $t = 0$ شروع شده و تا انتهای ثبت سیگنال ادامه دارد. اما در این پایان نامه، برای شرح واضح تر فرآیند، زمان در سیگنال ها از t تا $t + N$ در نظر گرفته شده است. با در نظر گرفتن این محدوده، هر سیگنال دارای مقادیری به طول $(N + 1)$ است. این فرآیند برای تمام سیگنال های موجود در n کانال مرتبط با EEG و m کانال مرتبط با fNIRS انجام شده است.

شکل ۱-۳، یک مثال از سیگنال EEG خام را نشان می دهد. هر سیگنال EEG یا fNIRS به طور کلی از تعدادی نمونه در هر زمان تشکیل شده است. از این رو، می توانیم هر سیگنال را به صورت کلی به شکل $e = [e_t, e_{t+1}, \dots, e_{t+N}]$ نمایش داد که در آن e نشان دهنده

سیگنال کامل و e_t نشان دهنده مقادیر آن سیگنال در زمان t است. از آنجایی که e_t برابر یک بردار به طول تعداد کانال های سیگنال است (برای EEG برابر n و برای fNIRS برابر m)، هر سیگنال را می توان به عنوان یک ماتریس نمایش داد. هر کانال سیگنال، به عنوان یک سطر از این ماتریس در نظر گرفته شده است و ستون های این ماتریس، مقادیر نمونه های این سیگنال، به ازای هر کانال است. از این رو، دو ماتریس شکل گرفته، یکی برای EEG و دیگری برای fNIRS به ترتیب دارای ابعاد $n \times (t + N)$ و $m \times (t + N)$ هستند.



شکل ۳-۱) مثالی از ۲۰ کانال یک سیگنال EEG خام.

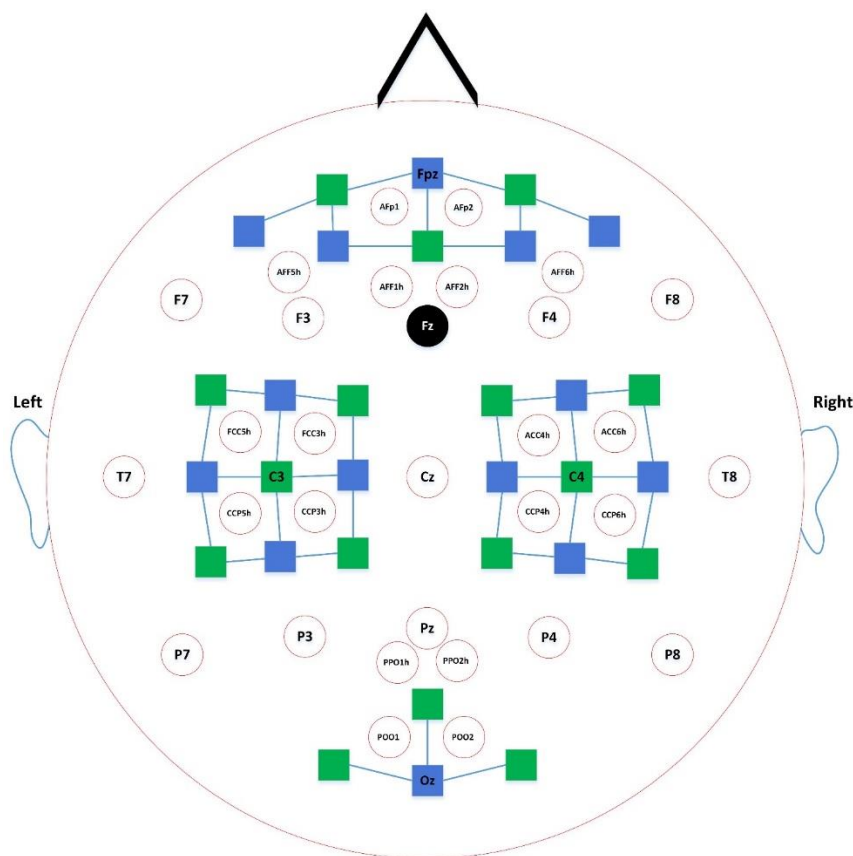
در این ماتریس، هر سطر (کانال)، تنها به دو سطر بالا و پایین خودش محدود است. به عبارت دیگر، سیگنال ثبت شده از پوست سر، همبستگی مکانی نمونه‌ها را در نظر نمی‌گیرد. به بیان ساده‌تر، در شکل ۳-۱، هیچ نشانی از مکان ثبت الکترودها دیده نمی‌شود. ثبت سیگنال و نوع ذخیره‌سازی آن، به طور کلی، وابسته به نوع سیستم ثبت یا تنظیمات آن سیستم است. از این رو، در سیگنال ثبت شده ممکن است ترتیب سطرها یا کانال‌ها بهم‌ریخته باشد. ولی در نهایت، در این سیگنال خام، همسایگی مکانی و وابستگی احتمالی بین آن‌ها بررسی نشده است.

همانطور که قبلاً گفته شد و با توجه به اهمیت موقعیت مکانی الکترودها در EEG و fNIRS و ارتباط بین ناحیه‌های مختلف مغز در فعالیت‌های مختلف، در این پایان‌نامه، از یک تابع تبدیل برای تبدیل سیگنال‌های خام به تانسورهای سه بعدی استفاده شده است. این تابع تبدیل از موقعیت الکترودها در مغز (با توجه به سیستم ثبت) کمک می‌گیرد. رابطه (۳-۱)، تابع تبدیل مربوطه را نمایش می‌دهد. ورودی این تابع، سیگنال‌های EEG و fNIRS خام است که به ترتیب با e_t و f_t نمایش داده شده است. این تابع، یک تانسور سه بعدی است که با توجه به شکل ۳-۲ طراحی شده است. در شکل ۳-۲، ناحیه‌های مربوط به ثبت سیگنال‌های EEG و fNIRS نشان داده شده است. در ادامه، این تابع مقادیر

مرتبط با هر کانال EEG و fNIRS را در تانسور مقداردهی می‌کند و در نهایت، آن تانسور را به عنوان خروجی باز می‌گرداند. اگر با دقت به رابطه (۳-۱) نگاه کنیم، متوجه می‌شویم خروجی این رابطه، یعنی تانسورها، کاملاً مطابق با مکان الکترودها در شکل ۳-۲ است. از این رو، در هر زمان t ، این تانسورها، ناحیه‌های فعال مغز را وابسته به نوع فعالیت و همچنین وابسته به مکان الکترودها نشان می‌دهند.

$$Transformation_function(e_t, f_t) = m_t = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & E_t^4 & 0 & E_t^8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & E_t^1 & 0 & E_t^2 & S_t^4 & E_t^5 & S_t^3 & E_t^6 & 0 & E_t^9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & E_t^3 & 0 & E_t^7 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & S_t^2 & 0 & 0 & S_t^5 & 0 & S_t^{10} & 0 & 0 & S_t^6 & 0 & 0 & 0 \\ 0 & S_t^1 & 0 & 0 & 0 & S_t^3 & 0 & 0 & 0 & 0 & 0 & S_t^7 & 0 & 0 & 0 & S_t^8 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & E_t^{16} & 0 & E_t^{18} & 0 & 0 & 0 & 0 & 0 & E_t^{26} & 0 & E_t^{30} & 0 & 0 \\ 0 & 0 & 0 & E_t^{14} & S_t^{13} & E_t^{17} & S_t^{14} & E_t^{23} & 0 & 0 & 0 & E_t^{25} & S_t^{25} & E_t^{29} & S_t^{26} & E_t^{28} \\ 0 & 0 & 0 & 0 & E_t^{15} & 0 & E_t^{22} & 0 & 0 & S_t^{11} & 0 & 0 & E_t^{27} & 0 & E_t^{25} & 0 \\ S_t^{17} & 0 & 0 & E_t^{13} & S_t^{15} & E_t^{20} & S_t^{16} & E_t^{24} & 0 & 0 & 0 & E_t^{28} & S_t^{27} & E_t^{33} & S_t^{28} & E_t^{24} \\ 0 & 0 & 0 & E_t^{19} & 0 & E_t^{21} & 0 & 0 & 0 & 0 & 0 & 0 & E_t^{32} & 0 & E_t^{31} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & S_t^{18} & 0 & 0 & 0 & S_t^{19} & 0 & 0 & S_t^{12} & 0 & 0 & S_t^{24} & 0 & 0 & 0 & S_t^{29} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & S_t^{20} & 0 & S_t^{23} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & S_t^{21} & E_t^{10} & S_t^{22} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & E_t^{11} & 0 & E_t^{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (3-1)$$

در این رابطه، S_t^n نشان‌دهنده نمونه داده سیگنال در زمان t و ثبت شده از کانال m سیگنال EEG است، E_t^m و $E_t'^m$ به ترتیب نشان‌دهنده اکسی و دی‌اکسی هموگلوبین برای سیگنال fNIRS در زمان t و ثبت شده از کانال m هستند.



شکل ۲-۳) موقعیت الکترودهای EEG در سیستم ۱۰-۵ بین المللی (دایره‌های قرمز و سیاه (الکتروود زمین)) و منابع fNIRS (مربع‌های آبی) و تشخیص‌دهنده‌ها (مربع‌های سبز). خط‌های آبی نشان‌دهنده کانال‌های fNIRS هستند. این تانسورهای سه بعدی با m_t نشان داده شده‌اند و شامل اطلاعات زمانی-مکانی سیگنال‌ها هستند. هر تانسور شامل دو ماتریس دوبعدی است، که به عنوان صفحه^۱ شناخته می‌شوند. در هر دو صفحه (در زمان t)، مقادیر سیگنال EEG برابر است، اما مقادیر سیگنال fNIRS متفاوت است. در یکی از این صفحات، مقدار اکسی هموگلوبین (E_t) و در دیگری مقدار دئوکسی هموگلوبین (E'_t) قرار داده شده است.

همانطور که در معادله ۱-۳ نشان داده شده است، ناحیه‌هایی که مطابق با هیچ موقعیت الکترودی نیستند، با صفر نشان داده شده‌اند. این ناحیه‌های صفر، مانند پس‌زمینه تصویر عمل می‌کنند. از این رو هیچ تاثیری بر روی آموزش شبکه عصبی ندارند. بنابراین، در ادامه، ناحیه‌های غیر صفر با استفاده از

¹ Facet

الگوریتم z-score نرمال شده‌اند. در نهایت، بعد از این مرحله، ما به ازای بردار داده‌های $[e_t, e_{t+1}, \dots, e_{t+N}]$ و $[f_t, f_{t+1}, \dots, f_{t+N}]$ ، یک تانسور سه بعدی به شکل $[m_t, m_{t+1}, \dots, m_{t+N}]$ داریم که نشان‌دهنده فعالیت مغز از زمان شروع t تا زمان پایان $t + N$ است.

در ادامه این پایان‌نامه، اطلاعات زمانی با کمک از روش پنجره‌گذاری لغزان به اطلاعات مکانی اضافه می‌شوند. برای این منظور، ابتدا، روش پنجره‌گذاری لغزان مروری می‌شود.

۳-۳-۲ پنجره‌گذاری لغزان

روش پنجره‌گذاری لغزان [۱۰۶]، روشی است که معمولاً برای استخراج ویژگی‌های مهم (به عنوان مثال میانگین) از یک مجموعه داده خام، در مقیاس بزرگ استفاده می‌شود. ایده این روش مشابه روش Moving Average است. برخی از پارامترهای استفاده شده در روش پنجره‌گذاری لغزان، به صورت جدا برای هر مجموعه داده مختلف تنظیم شده است. سه پارامتر مهم این روش در ادامه، آمده‌اند:

- اندازه پنجره: یک مقدار که اندازه پنجره را روی داده مشخص می‌کند.
- اندازه قدم: پنجره به اندازه این مقدار روی داده حرکت می‌کند.
- عملیات آماری انتخاب شده روی داده: عملیات آماری مانند میانگین، حداقل مقدار، حداکثر مقدار و غیره روی داده انجام می‌شود. این عملیات در پنجره محاسبه می‌شود.

به‌طور کلی، دو نوع رایج پنجره‌گذاری لغزان مجموع tumbling و مجموع rolling وجود دارد. مطالعات پیشین در روش پنجره‌گذاری لغزان، عدم همپوشانی بین پنجره‌ها را پیشنهاد دادند. یعنی اندازه پنجره برابر اندازه قدم باشد. اما اکنون در روش پنجره‌گذاری لغزان بین پنجره‌ها همپوشانی وجود دارد. در واقع، اندازه پنجره از اندازه قدم بزرگ‌تر است.

با توجه به دنباله $\{x_t\}$ ، دنباله جدید $\{SW_k\}$ را می‌توان با مقادیر میانگین دنباله اول استخراج کرد:

$$\{SW_k\} = \left\{ \frac{1}{w} \sum_{t=1+(k-1)m}^{kw+(k-1)m} x_t \right\} \quad (3-2)$$

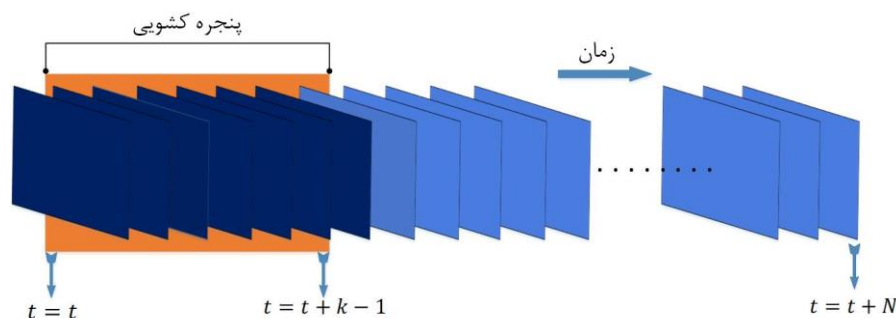
که در رابطه (۳-۲)، w نشان‌دهنده اندازه پنجره، m نشان‌دهنده اندازه قدم، k نشان‌دهنده شاخص دنباله جدید SW_k و $k = \left\lceil 1 + \frac{t-w}{m} \right\rceil$ که $\forall t \geq w$ و $w \geq m$ است. روش پنجره‌گذاری لغزان، به طور موثر می‌تواند داده‌های سیگنال در مقیاس بزرگ را به یک حجم کوچک‌تر (به نام مجموع داده جمع‌آوری شده) کاهش دهد. مجموعه داده جمع‌آوری شده، در فرآیندهای بعدی استفاده می‌شوند [۱۰۶].

روش پنجره‌گذاری لغزان در این پایان‌نامه، برای افزودن ویژگی‌های زمانی به تانسورهای سه بعدی شامل اطلاعات مکانی استفاده شده است.

۳-۳-۳ افزودن ویژگی زمانی به کمک روش پنجره‌گذاری لغزان

همانطور که تا به این جا شرح داده شد، به ازای هر نمونه در لحظه زمانی t ، یک تانسور سه بعدی m_t وجود دارد که شامل اطلاعات مکانی ثبت‌شده از آن سیگنال در آن لحظه زمانی است. از این رو، در این بخش، از روش پنجره‌گذاری لغزان استفاده شده است تا با حرکت یک پنجره با اندازه معلوم، روی این تانسورها، کلیپ‌هایی را مطابق با داده‌هایی با اطلاعات مکانی ایجاد کنیم. روش پنجره‌گذاری لغزان در اینجا، به این صورت عمل می‌کند که با حرکت دادن یک پنجره بر روی داده‌ها، داده‌های موجود در پنجره را به صورت یک کلیپ داده جدا کرده و به عنوان یک نمونه جدید برای مجموعه داده در نظر می‌گیرد.

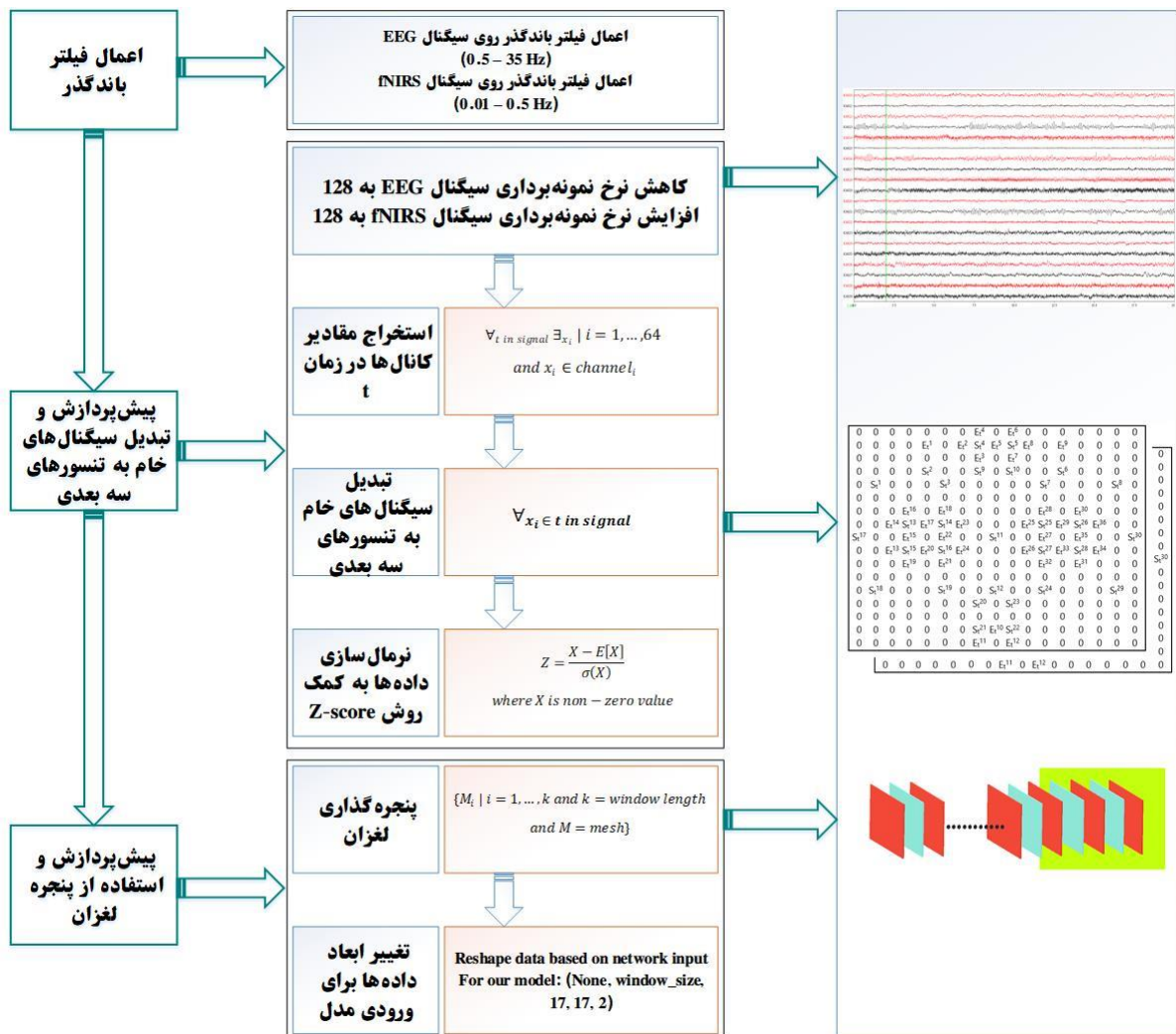
در هر پنجره، تعدادی تانسور سه بعدی از زمان t تا $t+k-1$ وجود دارد و به صورت $SW_j = [m_t, m_{t+1}, \dots, m_{t+k-1}]$ نمایش داده می‌شود که در آن SW_j برابر داده جدید، شامل تعدادی تانسور سه بعدی با اطلاعات مکانی و زمانی است. مقدار $j = 1, 2, 3, \dots, q$ است که q برابر تعداد بخش‌هایی است که در طول دوره پنجره‌گذاری مشاهده می‌شود و k برابر اندازه پنجره است. شکل ۳-۳، فرآیند اعمال پنجره‌گذاری لغزان را روی داده‌ها نشان می‌دهد.



شکل ۳-۳) فرآیند پنجره‌گذاری لغزان روی تانسورهای سه بعدی.

روش پنجره‌گذاری لغزان استفاده شده در این پایان‌نامه، تانسورهای سه بعدی m_t را به کلیپ‌های داده شامل اطلاعات زمانی و مکانی تبدیل می‌کند. منظور از کلیپ داده، هر SW_z شامل تعداد یکسانی تانسور سه بعدی است. هر کلیپ داده با کلیپ بعدی و قبلی خود ۵۰ درصد همپوشانی دارد. این همپوشانی به دلیل نقص اطلاعات درون بازه زمانی $[t, t + k - 1]$ است. به این صورت که ممکن است بخشی از حرکت شخص به صورت کامل در یک پنجره قرار نگیرد و ممکن است که بخشی از اطلاعات به پنجره بعدی منتقل شود. به همین دلیل، همپوشانی ۵۰ درصد هر پنجره با پنجره بعدی، موجب حفظ اطلاعات می‌شود و نقص اطلاعات خللی در انجام پردازش‌های بعدی ایجاد نمی‌کند.

به طور کلی، شکل ۳-۴، روند کلی روش پیش‌پردازش پیشنهادی را ارائه می‌دهد که در آن، داده خام EEG و fNIRS، در نهایت، پس از ترکیب شدن، به شکل کلیپ‌های داده با ویژگی‌های مکانی و زمانی تبدیل می‌شوند.

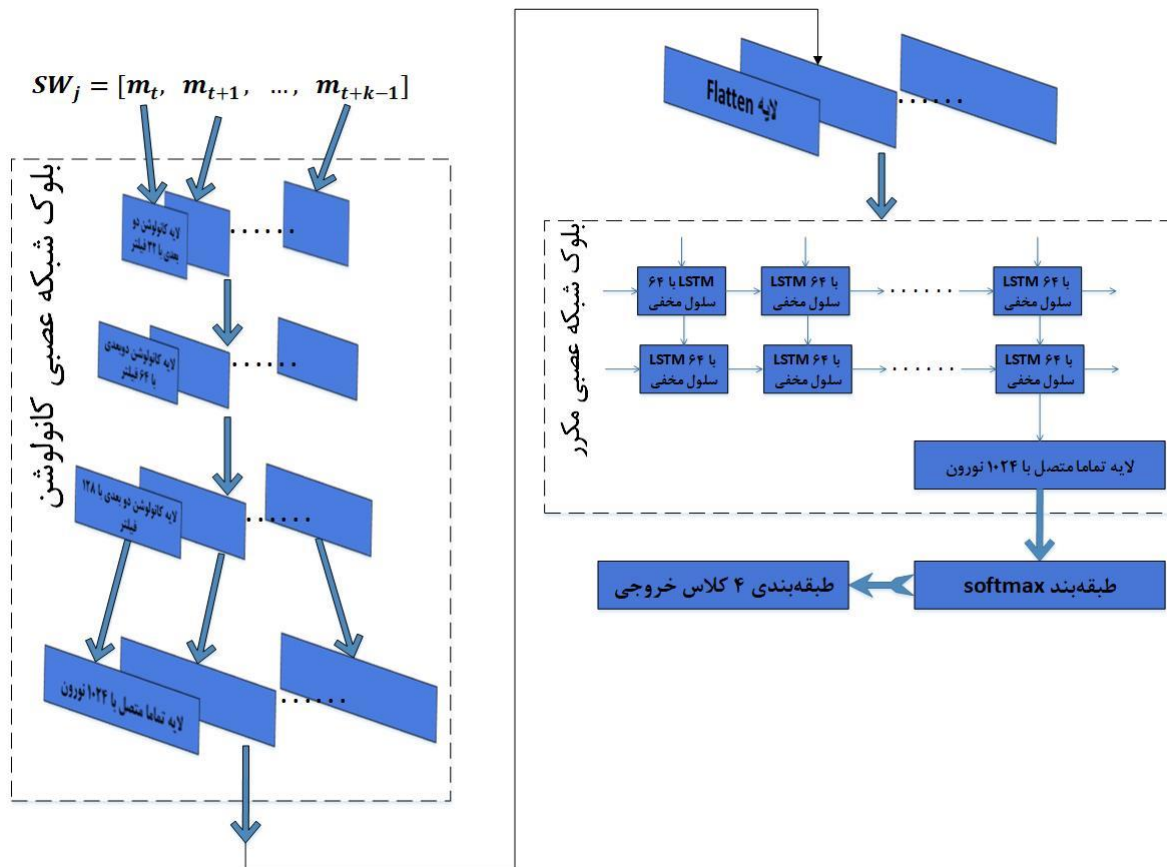


شکل ۳-۴) مرحله ثبت و پیش پردازش سیگنال های EEG و fNIRS.

۳-۴ مدل RCNN پیشنهادی

در این بخش، مدلی مبتنی بر یادگیری عمیق ارائه می دهیم. این مدل بر اساس دو روش شبکه های عصبی کانولوشن و شبکه های عصبی مکرر عمل می کند. از این مدل برای طبقه بندی اهداف ذهنی کاربر استفاده می کنیم و ساختار این مدل را با جزئیات شرح می دهیم. در این ساختار، ابتدا ویژگی های مکانی با کمک سه لایه کانولوشن از داده های ورودی استخراج می شوند. این ویژگی ها، لایه های LSTM را برای استخراج ویژگی های زمانی تغذیه می کنند. در نهایت، یک لایه Softmax، قصد کاربر را پیش بینی می کند.

ساختار مدل پیشنهاد شده در این پایان‌نامه، از کلیپ‌های داده SW_j به عنوان ورودی استفاده می‌کند. این ساختار شامل هشت لایه است. در میان این لایه‌ها، سه لایه کانولوشن دو بعدی، دو لایه تماماً متصل، دو لایه LSTM و یک لایه خروجی وجود دارد. این ساختار در شکل ۳-۵ نمایش داده شده است.



شکل ۳-۵) ساختار RCNN پیشنهادی.

ورودی لایه کانولوشن دو بعدی، داده z_{am} ورودی پنجره‌گذاری شده $SW_j = [m_t, m_{t+1}, \dots, m_{t+k-1}] \in \mathbb{R}^{k \times h \times w \times 2}$ است که در آن SW داده پنجره‌گذاری شده است که k تانسور سه بعدی در آن وجود دارد و می‌توان آن را با m_i ($i = t, t+1, \dots, t+k-1$) نمایش داد و ابعاد هر تانسور سه بعدی، برابر $h \times w \times 2$ است. که در آن h برابر طول صفحات تانسور و w عرض صفحات تانسور است. هر ورودی پنجره‌گذاری شده به لایه کانولوشن، یک بردار ویژگی

زمانی به صورت sf_i ($i = t, t + 1, \dots, t + k - 1$) استخراج می‌کند که با توجه به مقدار m_i می‌توان گفت:

$$CNN_{process}: Input(SW_j) \Rightarrow Output(SF_j) \text{ where } SF_j = [sf_t \dots sf_{t+k-1}] \quad (3-3)$$

اندازه کرنل برای هر کدام از لایه‌های کانولوشن دو بعدی، برابر 3×3 است. لایه اول دارای ۳۲ نقشه ویژگی است که برای هر لایه دو برابر می‌شود. به این ترتیب که لایه دوم دارای ۶۴ و لایه آخر دارای ۱۲۸ نقشه ویژگی است. بعد از این سه لایه، یک لایه تماماً متصل با 1024 نورون، وظیفه تبدیل ۱۲۸ نقشه ویژگی نهایی را به برداری دارای 1024 ویژگی مکانی تبدیل کند. از این رو، ویژگی مکانی نهایی، به صورت $SF_k \in \mathbb{R}^{1024}$ است. در نهایت، ویژگی مکانی استخراج شده (SF_k)، برای تغذیه شبکه RNN استفاده می‌شود.

دنباله ویژگی‌های مکانی SF_k ورودی شبکه RNN است. در این روش، از دو لایه LSTM برای ساختن یک شبکه RNN استفاده کردیم. ورودی هر لایه LSTM در این شبکه، دنباله زمانی خروجی از لایه قبلی است. در ساختار شبکه‌های RNN، حالت مخفی، واحدهای LSTM در هر زمان t را با h_t نمایش می‌دهند و h_{t-1} نیز حالت مخفی آن لایه در زمان قبلی $t - 1$ است. اطلاعات بدست آمده از زمان قبلی به حالت فعلی منتقل می‌شود و خروجی نهایی لایه را تحت تاثیر خود قرار می‌دهد. در این پایان‌نامه، ما از حالت مخفی واحد LSTM، به عنوان خروجی این لایه استفاده کردیم. در نتیجه، دنباله ورودی لایه دوم LSTM، دنباله‌ای از حالات مخفی لایه اول به صورت $[h_t, h_{t+1}, \dots, h_{t+k-1}]$ است. حالات مخفی لایه دوم را به صورت h'_t نشان می‌دهیم. تنها خروجی آخرین زمان لایه LSTM دوم (h'_{t+k-1}) برای تغذیه لایه تمام متصل استفاده می‌شود. پس داریم:

$$RNN_{process}: h'_{t+k-1} = RNN_{lstm}(SF_j) \Rightarrow TF_j, h'_{t+k-1} \in \mathbb{R}^d \quad (3-4)$$

که در رابطه بالا، d اندازه حالت مخفی یک واحد LSTM است. تعداد حالات مخفی برای هر دو لایه LSTM برابر ۶۴ است. بعد از استخراج ویژگی‌های زمانی TF_j توسط شبکه RNN، یک لایه تماماً متصل با ۱۰۲۴ نورون وجود دارد که ۶۴ حالت مخفی از ویژگی‌های زمانی ایجاد شده از این شبکه را به برداری با ۱۰۲۴ ویژگی مکانی نهایی تبدیل می‌کند. بعد از آن، لایه Softmax وظیفه پیش‌بینی احتمالی نهایی هر کلاس داده را دارد:

$$Softmax(TF_j) \in \mathbb{R}^c \quad (3-5)$$

هدف، طبقه‌بندی داده‌های ورودی در c گروه است.

علاوه بر طراحی یک ساختار مناسب، انتخاب پارامترهای بهینه، تاثیر بسیار زیادی بر تشخیص درست فعالیت‌های ذهنی دارد. توابع بهینه‌ساز زیادی برای بهینه‌سازی پارامترها وجود دارد. با این حال، تابع بهینه‌ساز Adam با نرخ یادگیری ۰,۰۰۱ بهترین عملکرد را برای این مدل نشان داد. همچنین، برای همه لایه‌ها بجر لایه آخر، از تابع فعال‌ساز ReLU استفاده شده است. جدول ۳-۲، جزئیات داخلی ساختار RCNN پیشنهاد شده در این پایان‌نامه را به صورت خلاصه بیان می‌کند.

جدول ۳-۲ جزئیات ساختار مدل RCNN پیشنهادی استفاده شده در این پایان‌نامه.

لایه	نام لایه	اندازه کرنل	تعداد خروجی	قدم
۱	Conv2D	3×3	۳۲	۱
۲	Conv2D	3×3	۶۴	۱
۳	Conv2D	3×3	۱۲۸	۱
۴	Dense	–	۱۰۲۴	–
۵	Flatten	–	–	–
۶	LSTM	–	۶۴	–
۷	LSTM	–	۶۴	–
۸	Dense	–	۱۰۲۴	–

فصل ۴ : نتیج و آزمایش؛

۴-۱ مقدمه

در این فصل، عملکرد روش پیشنهادی بررسی شده است. همچنین تحلیل آزمون‌های مختلف روی مجموعه داده EEG-fNIRS انجام شده است و از نمایش گرافیکی برای تحلیل الگوریتم پیشنهادی استفاده شده است. همچنین در این بخش، روش پیشنهادی ارائه شده با برخی از روش‌های دیگر در این زمینه مقایسه شده و نتایج گزارش شده است. در این پایان‌نامه، آزمایش‌های انجام گرفته همه بر روی مجموعه داده Technische Universität Berlin [۴۴] انجام شده است. عملکرد الگوریتم RCNN پیشنهادی با حضور fNIRS و بدون حضور آن و در ادامه با در نظر گرفتن یکی از دو موج موجود در fNIRS نیز ارزیابی شده است.

۴-۲ مجموعه داده hybrid EEG-fNIRS استفاده شده در این

پایان‌نامه

در این بخش ما قصد داریم تا مجموعه داده استفاده شده در این پایان‌نامه را شرح دهیم. این یک مجموعه داده به صورت آزاد^۱ و در دسترس است و ترکیبی از دو سیگنال EEG و fNIRS است. در این مجموعه داده دو آزمایش تصور حرکت دست چپ در مقابل دست راست و محاسبه کارهای ذهنی در مقابل حالت استراحت انجام شده است.

سیگنال‌های ثبت شده در این مجموعه داده، از ۲۸ شخص سالم راست دست و یک شخص سالم چپ دست ثبت شده است (۱۴ مرد و ۱۵ زن با میانگین سنی $28 \pm 3,7$ (سال)). در هیچ کدام از این ۲۹ فرد هیچگونه بیماری عصبی، جسمی یا صدمه‌های مرتبط با مغز ثبت نشده است. این مجموعه داده مورد تایید کمیته اخلاقی موسسه روانشناسی و ارگونومی دانشگاه صنعتی برلین با شماره ثبت SH_01_20150330 است.

¹ Open Source

داده‌های EEG و fNIRS در یک اتاق روشن و معمولی جمع‌آوری شده است. سیگنال EEG توسط تقویت‌کننده EEG چندکاناله BrainAmp با ۳۰ الکتروود فعال^۱ ثبت شده است. نرخ نمونه‌برداری برای این سیگنال برابر ۱۰۰۰ هرتز بوده است. این ۳۰ الکتروود با استفاده از سیستم بین‌المللی ۵-۱۰ ثبت شده است. سیگنال‌های fNIRS توسط سیستم NIRScout^۲ ثبت شده است. نرخ نمونه‌برداری برای این سیگنال برابر ۱۲٫۵ هرتز بود. این سیستم، ۱۴ منبع ارسال و ۱۶ تشخیص‌دهنده دارد. این سیستم ۳۶ کانال را برای ثبت سیگنال فراهم کرده است. فاصله بین هر منبع ۳۰ میلی‌متر بود. هر دو سیگنال به صورت همزمان و همه کانال‌ها بر روی یک کلاه قرار داشتند. فهرست کانال‌های EEG و fNIRS به صورت کلی و نحوه قرارگیری آن‌ها روی سر نیز در شکل ۳-۲ نشان داده شده است.

در این مجموعه داده، به طور کلی دو زیر مجموعه شامل زیر مجموعه داده مرتبط با تصورات حرکتی و زیر مجموعه داده مرتبط با محاسبات ذهنی وجود دارد. برای زیر مجموعه داده مرتبط با تصورات حرکتی، به هر فرد برای انجام تصور یک حرکت (مانند باز و بسته کردن مشت دست برای گرفتن یک توپ) دستور داده شده است. برای این کار، یک علامت جهت به سمت راست یا سمت چپ بر روی یک صفحه نمایش در مقابل داوطلب به مدت دو ثانیه ظاهر شده است. این علامت جهت با یک صدای بوق کوتاه ناپدید شده سپس کار برای داوطلب آغاز شده است. داوطلب تا زمانی که کلمه STOP روی صفحه نمایش ظاهر شده، عمل تصور را تکرار کرده است. یک دوره استراحت چند ثانیه‌ای بین هر آزمایش وجود دارد و آزمایش مجدداً تکرار شده است. این فرآیند ۲۰ مرتبه تکرار شده است.

برای زیر مجموعه داده مرتبط با محاسبات ذهنی، دو عمل محاسبه ذهنی و حالت استراحت وجود داشت. عمل محاسبه ذهنی به این صورت انجام شده که یک عملیات تفریق ساده (به عنوان مثال ۴-۳۸۴) در مرکز صفحه به مدت ۲ ثانیه ظاهر شده و سپس کاربر تا زمان دیدن علامت STOP این تفریق را در ذهنش انجام داده است.

^۱ Brain Products GmbH, Gilching (آلمان) (محصول آلمان)

^۲ NIRx GmbH (آلمان) (برلین، آلمان)

نرخ نمونه‌برداری داده‌های خام در این مجموعه داده برای EEG به ۲۰۰ هرتز و برای fNIRS به ۱۰ هرتز کاهش داده شده است. هیچ پیش‌پردازشی روی این داده‌ها انجام نشده است [۴۴].

۳-۴ خلاصه آزمایش‌ها

در بخش نتایج، به بررسی برخی اجراها و آزمایش‌ها می‌پردازیم. برای ارزیابی تاثیر ترکیب دو سیگنال EEG و fNIRS با یکدیگر، ما مدل پیشنهادی را با سیگنال‌های EEG، ترکیب سیگنال EEG با هر کدام از سیگنال‌های طول موج پایین و بالا در سیگنال fNIRS، و در نهایت ترکیب هر دو با هم آموزش داده و دقت نهایی مدل را برای هر کدام از این سیگنال‌ها شرح داده‌ایم. ماتریس درهم ریختگی برای بررسی نتایج حاصل از آزمون مدل آموزش دیده و همچنین عملکرد آن رسم شده و تحلیل شده است. نتایج مرتبط با روش پیش‌پردازش و خروجی حاصل از این روش نیز در این بخش شرح داده شده است و در نهایت دقت روش پیشنهادی با دقت روش‌های دیگر مقایسه شده است.

در این پایان‌نامه، از اندازه پنجره ۱۰ برای روش پنجره‌گذاری لغزان استفاده شده است. در این پایان‌نامه، به صورت تصادفی، ۸۰ درصد داده به عنوان داده آموزشی، ۱۰ درصد به عنوان داده ارزیابی به منظور تنظیم پارامترها و ۱۰ درصد از داده نیز به عنوان داده آزمون در نظر گرفته شده است. تاثیر اطلاعات زمانی و مکانی در بهبود عملکرد مدل پیشنهادی نیز در این بخش بررسی شده است.

همه مدل‌ها و پیاده‌سازی‌ها با استفاده از چارچوب کراس^۱ در زبان برنامه‌نویسی پایتون^۲ با پس‌زمینه تانسورفلو^۳ طراحی شده است. در این پایان‌نامه، همه پیاده‌سازی‌ها به منظور سهولت در اجرا و کاهش بار محاسباتی و زمان، در سرور HP ProLiant DL380 Gen9 واقع در دانشکده کامپیوتر دانشگاه صنعتی شاهرود انجام شده است. این سرور دارای یک پردازنده با ۵۶ هسته و حافظه داخلی ۲۵۶ گیگابایت است.

^۱ Keras

^۲ python

^۳ Tensorflow

۴-۴ نتایج مرتبط با پیش‌پردازش داده‌ها

در این بخش، نتایج اعمال روش پیش‌پردازش روی داده‌های EEG و fNIRS شرح داده شده است. در روش پیش‌پردازش شرح داده شده در بخش ۳-۳-۱، چگونگی تبدیل داده‌های EEG و fNIRS خام به تانسورهای سه بعدی شرح داده شد. این تانسورها، شامل اطلاعات مکانی مطابق با سیستم ثبت الکترودها هستند. سپس استفاده از یک روش پنجره‌گذاری لغزان، مجموعه‌ای از این تانسورهای سه بعدی را تبدیل به کلیپ‌های داده مشتمل بر اطلاعات زمانی کرده است. نکته‌ای مهم چگونگی تغییر زمانی که داده‌های EEG و fNIRS به تانسورهای سه بعدی است.

یکی از راه‌های مشاهده فعالیت‌های ذهنی، استفاده از تصاویر توپوگرافی مغز است. نقشه‌ها یا تصاویر توپوگرافی ثبت شده از مغز، نمایشی از فعالیت‌های ذهنی است که شامل اطلاعات مفیدی هستند. از روی این تصاویر می‌توان ناحیه‌های غیر فعال و برخی بیماری‌ها را تشخیص داد. این تصاویر، نمایش مغز هستند. در اکثر این تصاویر ناحیه‌های فعال به رنگ قرمز و ناحیه‌های غیر فعال به رنگ آبی نشان داده شده‌اند. از این رو، تصاویر توپوگرافی شامل اطلاعات مفیدتری نسبت به سیگنال‌های خام هستند.

فعال شدن ناحیه‌های مختلف در مغز انسان کاملاً مرتبط با فعالیتی است که شخص انجام می‌دهد. ثبت سیگنال از روی پوست سر به کمک الکترودها، امکان دریافت مکان سیگنال‌ها را می‌دهد. هر الکتروده در یک ناحیه خاص روی سر قرار می‌گیرد. اما فعالیتی که مغز انجام می‌دهد تنها مختص به یک الکتروده و تنها یک ناحیه کوچک روی سر نیست. هر فعالیت ممکن است شامل ناحیه‌هایی باشد که تاثیر آن بر روی چندین الکتروده دیده شود. اما در این حالت، یک ناحیه بیشترین فعالیت و ناحیه‌های اطراف آن به مراتب فعالیت کمتری را در خود ثبت می‌کنند. به همین خاطر استفاده از تصاویر توپوگرافی می‌تواند به تحلیل مغز و رابطه هر فعالیت با ناحیه مخصوص آن کمک کند [۹۳].

در پیش‌پردازش عنوان شده در بخش ۳-۳-۱، سیگنال‌های خام به تانسورهای سه بعدی تبدیل می‌شوند. اطلاعات این تانسورها بسیار شبیه به تصاویر توپوگرافی مغز هستند. هر تانسور شامل دو فعالیت EEG (در ناحیه‌های ذکر شده) یکسان و فعالیت fNIRS شامل دو فعالیت طول موج پایین و طول موج بالا هستند. برای درک بهتر این تانسورها، در ادامه این پایان‌نامه تانسورهای سه بعدی را به صورت دو ماتریس نمایش می‌دهیم.

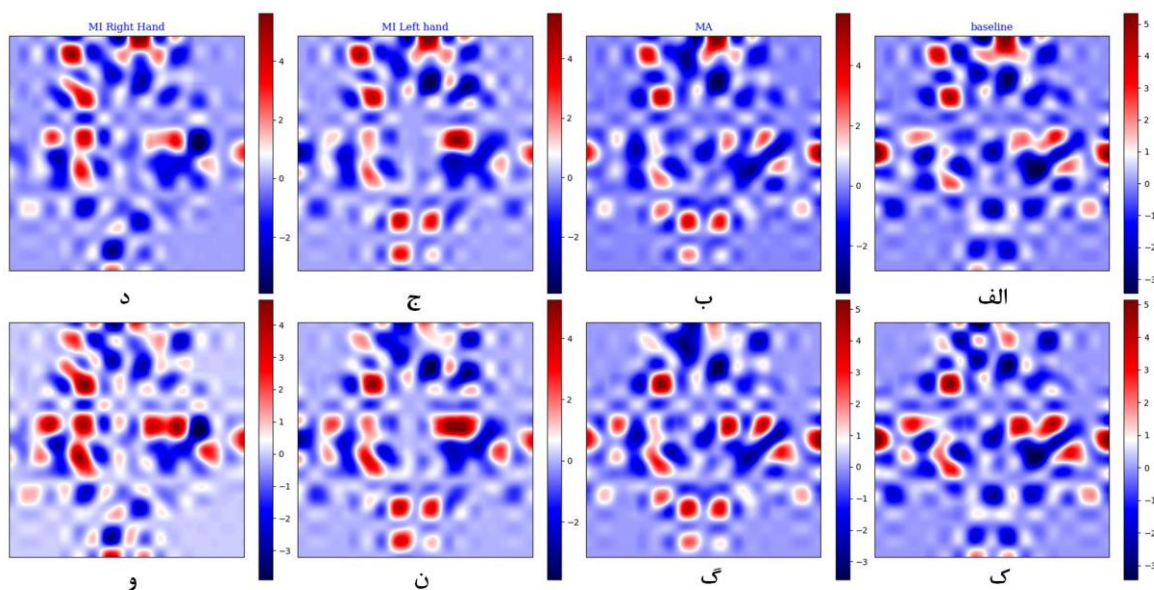
شکل ۳-۱ و شکل ۴-۱، به ترتیب سیگنال‌های سری زمانی و خروجی روش پیش‌پردازش پیشنهاد شده در این پایان‌نامه را نشان می‌دهند. در سیگنال‌های سری زمانی به صورت خام، نیاز به تخصص برای درک و استخراج ویژگی از آن وجود دارد. همانطور که در شکل ۳-۱ مشاهده می‌شود، ویژگی مکانی در این نوع سیگنال به طور کامل وجود ندارد. به این معنی که هیچ اطلاعاتی از ناحیه‌های مختلف مغز و همچنین الکترودهای آن ناحیه و فعالیت مغز در این سیگنال‌ها وجود ندارد. به صورت کلی این سیگنال‌ها، تنها شامل دو همسایه هستند که این دو همسایه نیز ممکن است در مغز، در کنار یکدیگر نباشند. اما، شکل ۴-۱ شامل اطلاعات مکانی ترکیب سیگنال‌های EEG و fNIRS است. شکل ۴-۱ نقشه‌های توپوگرافی تولید شده به کمک روش پیش‌پردازش را نمایش می‌دهد. به عبارت دیگر، این شکل نشان‌دهنده تانسورهای سه بعدی (m_t) هستند. شکل ۴-۱ فعالیت مغز را در طول کارهای مختلف نشان می‌دهد و شبیه به تصاویر توپوگرافی است. در روش پیشنهادی، داده‌های EEG و fNIRS در زمان t (یعنی e_t و f_t) به تانسورهای سه بعدی (m_t) تبدیل می‌شوند که این تانسورها مطابق فعالیت مغز در زمان t هستند. برای ایجاد این تصاویر، مغز به ناحیه‌هایی شامل الکترودهای ثبت تقسیم شده و سپس مطابق با معادله ۹، m_t ایجاد شده است. هر m_t شامل دو ماتریس با ۱۷ سطر و ۱۷ ستون است. بنابراین محور عمودی و افقی آن این مقادیر را نشان می‌دهد.

با مقایسه تصاویر ۳-۱ و ۴-۱، ما متوجه می‌شویم که تبدیل سیگنال‌های سری زمانی به تانسورهای سه بعدی اطلاعات بیشتری را ارائه می‌دهد. در پیش‌پردازش ارائه شده، ویژگی‌های مکانی استخراج شده

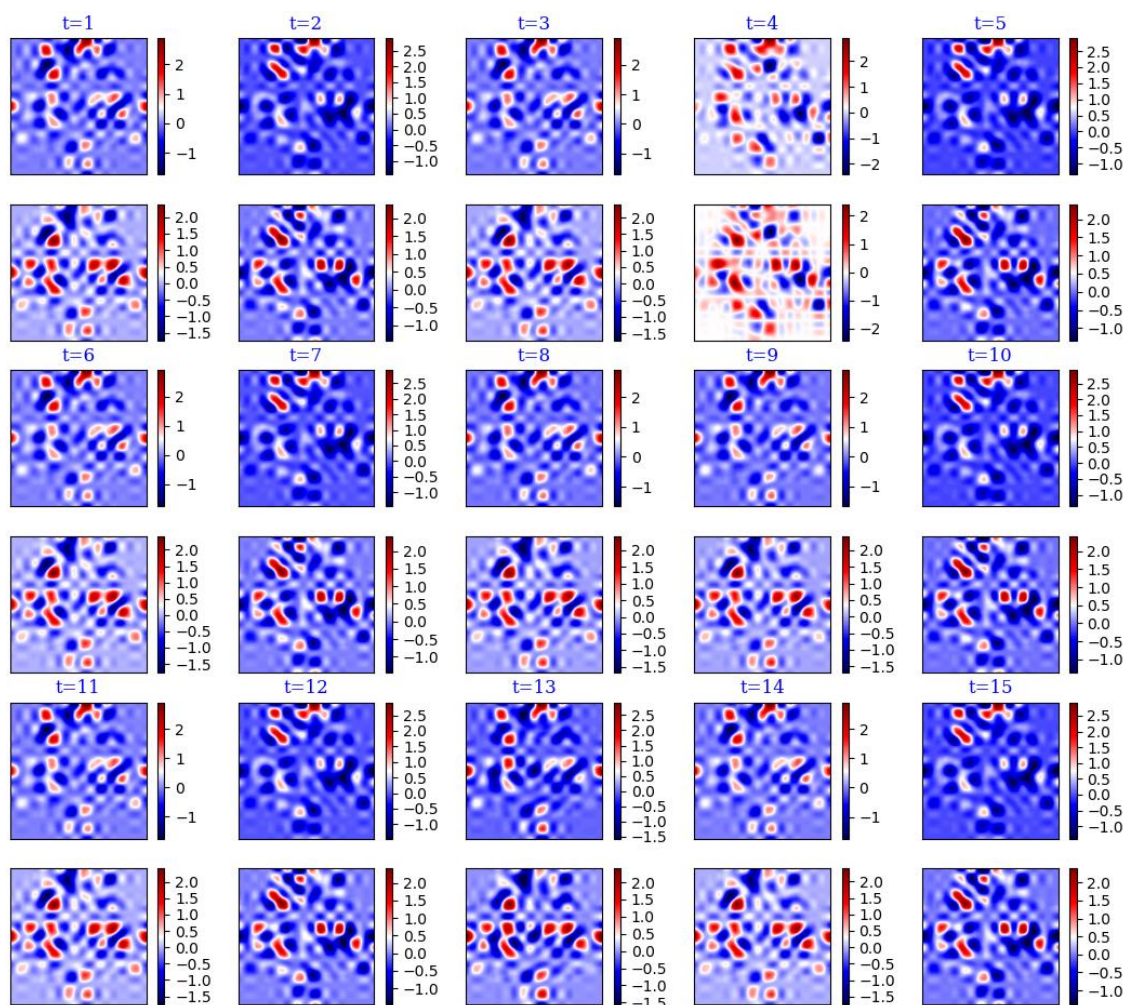
شامل اطلاعاتی درباره ناحیه‌های فعال مغز در زمان مورد نظر است. همانطور که از این تانسورهای سه بعدی مشخص است، موقعیت الکترودها مطابق با شکل ۳-۲ است. هر الکترودها و همسایگی آن به راحتی قابل تشخیص است.

از طرف دیگر، اطلاعات مکانی به تنهایی، کمی نقص دارد. این نقص به دلیل فعالیت چندگانه انسان است که داده مرتبط با فعالیت مورد نظر ما را از بین ببرد. از جمله این فعالیت‌ها حرکت سر، زبان یا پلک زدن است. با در نظر گرفتن این اتفاق، ممکن است نمایش این تانسورهای سه بعدی با خطا روبرو باشد. برای بیان بهتر، فرض کنید شخصی در لحظه زمانی t در حال انجام فعالیتی خاص باشد. انتظار می‌رود با دیدن تصاویر توپوگرافی در لحظه انجام فعالیت، ناحیه مرتبط با نیم کره چپ مغز او فعال‌تر باشد. اما وجود فعالیت‌های ناخواسته در آن زمان باعث از بین رفتن آن فعالیت می‌شود. با این اتفاق، روند مشاهده و تشخیص حرکت حتی برای سیستم‌های تشخیص نیز منجر به خطا می‌شود. به منظور رفع این مشکل می‌توانیم داده‌ها را در بازه‌های زمانی بزرگ‌تری به جای یک نمونه زمان بررسی کنیم. برای این منظور در روش پیش‌پردازش پیشنهادی از پنجره‌گذاری لغزان استفاده کرده‌ایم. با اعمال روش پنجره‌گذاری لغزان، اطلاعات زمانی نیز به این اطلاعات مکانی اضافه می‌شود که در شکل ۴-۲ نشان داده شده است.

افزودن اطلاعات زمانی موجب می‌شود که خطای ناشی از حرکت‌های ناخواسته در بدن رفع شود. به عنوان مثال اگر شخصی در بازه زمانی $t = 0$ تا $t = n$ فعالیت خاصی را انجام داده باشد و در زمان $t = k$ که $0 < k < n$ فعالیت ناخواسته را نیز انجام داده باشد، این فعالیت تنها در همان زمان از بین رفته است و در دیگر زمان‌ها به درستی قابل مشاهده است. شکل ۴-۲، تغییرات ناحیه‌های فعال مغز را در طول حرکت شخص از مجموعه داده نشان می‌دهد. این شکل، ورودی شبکه عصبی عمیق پیشنهادی است. در این شکل، فعالیت مغزی یک کلیپ داده برای فعالیت تصور حرکتی دست راست نشان داده شده است.



شکل ۴-۱) نتیجه حاصل از روش پیش‌پردازش پیشنهادی. فعالیت‌های مغزی در طول کارهای مختلف مشابه توپوگرافی مغز. الف) ترکیب EEG با طول موج کمتر سیگنال fNIRS برای تصور حرکت دست راست. ب) ترکیب EEG با طول موج کمتر سیگنال fNIRS برای تصور حرکت دست چپ. ج) ترکیب EEG با طول موج کمتر سیگنال fNIRS برای فعالیت محاسبه ذهنی. د) ترکیب EEG با طول موج کمتر سیگنال fNIRS برای حالت استراحت. ک) ترکیب EEG با طول موج بالاتر سیگنال fNIRS برای تصور حرکت دست راست. گ) ترکیب EEG با طول موج بالاتر سیگنال fNIRS برای تصور حرکت دست چپ. ن) ترکیب EEG با طول موج بالاتر سیگنال fNIRS برای فعالیت محاسبه ذهنی. و) ترکیب EEG با طول موج بالاتر سیگنال fNIRS برای حالت استراحت.



شکل ۴-۲) بعد از استفاده از روش پنجره‌گذاری لغزان. برای فعالیت تصور حرکتی دست راست در زمان‌های $t = 1$ تا $t = 15$. ویژگی‌های مکانی و زمانی برای آموزش مدل شبکه عمیق پیشنهادی استفاده شده است.

در شکل ۴-۱، تصاویر توپوگرافی مغز برای ۴ فعالیت نشان داده شده است. این فعالیت‌ها در شکل ۴-۱، از سمت چپ به ترتیب مرتبط با تصور حرکت دست راست، تصور حرکت دست چپ، محاسبه ذهنی و حالت استراحت هستند. تغییر فعالیت‌ها در ناحیه‌های مختلف مغز برای این چهار فعالیت در شکل ۴-۱ نشان داده شده است. در هر دو شکل ۴-۱ و ۴-۲، هر رنگ نشان‌دهنده طیف فعالیت در هر کانال داده است. رنگ قرمز تیره نشان‌دهنده بیشترین فعالیت و رنگ آبی تیره نشان‌دهنده کمترین فعالیت در مغز است. در تصور حرکتی دست راست، فعالیت در ناحیه چپ مغز بیشتر است. این فعالیت در ناحیه چپ مغز با توجه به شکل ۳-۱، در کانال‌های قرار گرفته در نیم کره سمت چپ مغز و در اطراف

کانال‌های C1 و C3 به مراتب بیشتر دیده می‌شود. اما در تصور حرکتی دست چپ، فعالیت در فرکانس‌های مختلف متفاوت است. این فعالیت در سرتاسر مغز دیده می‌شود. همچنین فعالیت کمی در نیم کره راست مغز نیز دیده می‌شود [۹۳]. در فعالیت ذهنی، بالاترین سطح فعالیت در ناحیه جلویی مغز دیده می‌شود و در حالت استراحت، فعالیت کمی در سرتاسر مغز مشاهده می‌شود. با توجه به این فعالیت می‌توان گفت که مغز در آرامش به سر می‌برد. توپوگرافی مغز در هنگام فعالیت حرکت دست راست، به صورت یک کلیپ در شکل ۴-۲ نمایش داده شده است. این کلیپ داده، تغییرات متوالی ناحیه‌های مختلف مغز را برای فعالیت تصور حرکتی دست راست نشان می‌دهد. این شکل شامل ۱۵ بخش برای هر زمان t می‌باشد. هر بخش شامل دو تصویر است. تصویر اول مرتبط با ترکیب سیگنال EEG با مقدار اول سیگنال fNIRS و تصویر دوم مرتبط با ترکیب سیگنال EEG با مقدار دوم سیگنال fNIRS است. تمام این اطلاعات با توجه به پیش‌پردازش انجام شده روی سیگنال‌های خام صورت گرفته است که هیچکدام در این سیگنال‌ها قابل مشاهده نیستند. از طرف دیگر، افزودن سیگنال fNIRS به EEG امکان بررسی‌های بیشتر را به ما می‌دهد. یکی از مزایای ترکیب این دو سیگنال، بیشتر شدن نمونه‌های داده است. از طرف دیگر، ناحیه‌های مختلف مغز به صورت گسترده‌تری مورد بررسی قرار می‌گیرند. با توجه به شکل ۳-۱، در هر کدام از دو ناحیه نیم کره چپ و راست مغز ۱۴ کانال در همسایگی یکدیگر قرار دارند که برای سیگنال EEG، این تعداد تنها به پنج کاهش می‌یابد. این نکته برای ناحیه جلویی سر نیز صادق است.

۴-۵ نتایج طبقه‌بندی مدل

در این بخش به بررسی نتایج حاصل از مدل پیشنهاد شده می‌پردازیم. مدل پیشنهادی یک مدل مبتنی بر یادگیری عمیق است که امروزه توجه بسیاری از محققان را به خود جلب کرده است. یادگیری عمیق تحولی عظیم را در الگوریتم‌های یادگیری ماشین بوجود آورده است. از این رو، استفاده از این الگوریتم یک مزیت بسیار مهم است. مدل پیشنهادی ترکیبی از ساختار شبکه عصبی کانولوشن و شبکه عصبی مکرر است. تبدیل ساختارهای سری زمانی به تانسورهای سه بعدی و در نتیجه تبدیل آن‌ها به کلیپ‌های

داده، استفاده از این دو نوع شبکه عصبی را مهم‌تر می‌کند. در ادامه به مزیت استفاده از این ترکیب این دو نوع شبکه عصبی می‌پردازیم.

شبکه‌های عصبی RCNN [۱۰۷]، دارای لایه‌های مختلف و گسترده‌تری هستند. به صورت تکنیکی انواع مختلفی برای ترکیب ساختارهای مختلف شبکه‌های عصبی وجود دارد. مدل‌هایی با چندین لایه CNN در ورودی و چندین لایه RNN یا LSTM در خروجی را در نظر بگیرید. این نوع شبکه می‌تواند دنباله‌ای از تصاویر ورودی مانند ویدئو را بخواند. دلیل استفاده ما از این مدل همین نکته است که خروجی روش پیشنهادی، کلیپ‌های داده هستند. این کلیپ‌ها شامل تعدادی تانسیور سه بعدی است. شبکه‌های ترکیبی برای تحلیل این کلیپ‌ها بسیار مناسب هستند.

برای ارزیابی اهمیت تجزیه و تحلیل اطلاعات مکانی و زمانی به صورت همزمان، از دو مدل شبکه عصبی کانولوشن برای در نظر گرفتن اطلاعات مکانی به تنهایی و شبکه عصبی بازگشتی برای در نظر گرفتن اطلاعات زمانی به تنهایی و ترکیب این دو استفاده شده است. در جدول ۴-۱ عملکرد این سه روش با یکدیگر مقایسه شده‌اند.

جدول ۴-۱) مقایسه مدل‌ها.

مراجع و روش - ها	EEG یا fNIRS یا ترکیب	روش طبقه‌بند	دقت (درصد)
[۶۰]	EEG	RCNN	۹۸,۳
[۱۰۱]	EEG	Deep Residual Learning	۹۲,۶۸۲
[۵۹]	EEG	RCNN with DWT	۸۶,۷
[۱۴]	EEG or fNIRS	LDA	fNIRS = ۷۵,۶ EEG = ۸۶
[۶۲]	EEG-fNIRS	DNN	۸۳,۲۸

۸۴,۲۶	Deep Belief Network	EEG-fNIRS	[۱۰۳]
۹۰	sLDA	EEG-fNIRS	[۱۰۵]
۹۸,۲۲	CNN	EEG-fNIRS	روش پیشنهادی
۹۵,۸۱	RNN	EEG-fNIRS	روش پیشنهادی
۹۹,۶۳	RCNN	EEG-fNIRS	روش پیشنهادی

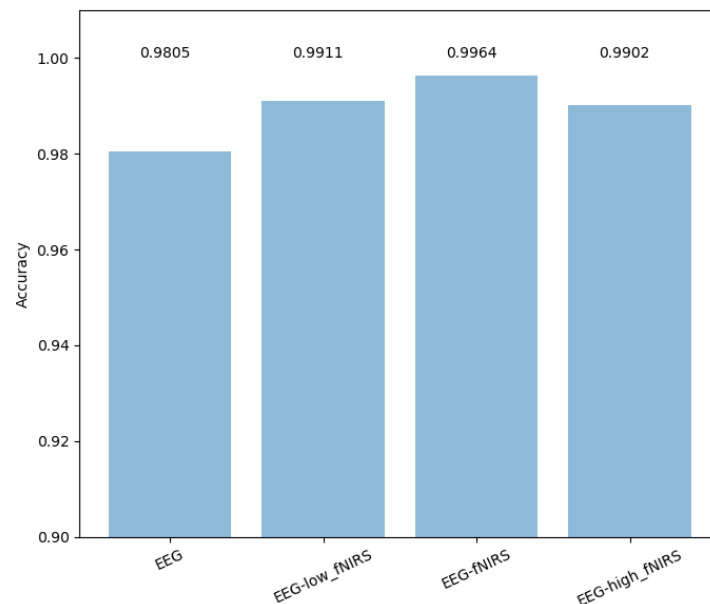
همانطور که در جدول ۴-۱ آمده است، دقت روش پیشنهادی در حالت شبکه عمیق ترکیبی به حداکثر خود رسیده است. از نتایج بدست آمده، می توان نتیجه گرفت که در نظر گرفتن اطلاعات زمانی و مکانی به صورت همزمان، به تشخیص قصد انسان کمک می کند. دقت در این حالت برابر با ۹۹,۶۳ درصد است. اما همانطور که مشهود است، دقت برای حالتی که تنها اطلاعات مکانی با کمک شبکه عصبی کانولوشن در نظر گرفته شد برابر ۹۸,۲۲ درصد است. این دقت به صورت کلی با دقت ترکیبی هر دو اطلاعات مکانی و زمانی برای سیگنال EEG در مرجع [۶۰] برابر است. دلیل بالاتر بودن دقت کلی روش پیشنهادی در مقایسه با دقت کلی مرجع [۶۰]، به دلیل وجود اطلاعات fNIRS در کنار سیگنال EEG است. انتظار می رفت که افزودن اطلاعات fNIRS به EEG دقت را در همه حالات بالا ببرد و همانطور که از نتایج پیداست، این افزایش دقت رخ داده است. از طرف دیگر، در نظر گرفتن اطلاعات زمانی به تنهایی با کمک شبکه عصبی بازگشتی دقت را به طور قابل توجهی به ۹۵,۸۱ درصد کاهش می دهد.

در [۱۴] و [۱۰۵] از یادگیری عمیق برای طبقه بندی داده ها استفاده نشده است. الگوریتم یادگیری عمیق یک تکنولوژی جدید است و در تحقیقات زیادی از این الگوریتم ها استفاده شده است. در [۱۰۱]، [۵۹]، [۶۲] و [۱۰۳] از یادگیری عمیق برای طبقه بندی داده ها استفاده شده است. این روش ها از EEG به تنهایی و هم از ترکیب دو سیگنال EEG و fNIRS استفاده کرده اند، اما دقتی به مراتب پایین تر از انتظار بدست آورده اند. دقت پایین طبقه بندی در این روش ها به دلیل نیازمندی الگوریتم های یادگیری

عمیق به داده‌های آموزشی زیاد برای آموزش است. زمانی که از الگوریتم‌های استخراج ویژگی مختلف استفاده می‌شود، حجم داده به صورت قابل توجهی کاهش می‌یابد. این کاهش داده ممکن است بر روی عملکرد شبکه عمیق تاثیر منفی بگذارد. از طرف دیگر، ویژگی‌هایی که به صورت دستی استخراج می‌شوند ممکن است شامل ویژگی‌های اشتباه یا نامناسب باشند. تمام این‌ها می‌توانند تاثیر منفی بر عملکرد یک شبکه عمیق داشته باشند.

یکی از گزینه‌هایی که از اهمیت بالایی برخوردار است، پیش‌پردازش روی داده‌ها است. معمولاً از پیش‌پردازش برای حذف نویز موجود در داده‌ها استفاده می‌شود. در [۶۰]، مرحله پیش‌پردازش شامل الگوریتم‌های حذف نویز نیست. اما از این مرحله برای تغییر و آماده‌سازی داده EEG برای مرحله طبقه‌بندی استفاده شده است. در این مقاله، سیگنال‌ها را برای استفاده از ویژگی‌های مکانی نهفته در آن‌ها تغییر داده‌اند. سپس از یک شبکه عمیق RCNN برای طبقه‌بندی داده‌های خود استفاده کرده‌اند. با توجه به جدول ۴-۱، در بیشتر روش‌ها سیگنال‌ها به صورت سری زمانی و حالت زنجیره مانند در نظر گرفته شده‌اند. در این روش‌ها سعی شده با یافتن ویژگی‌های مناسب و استخراج آن‌ها دقت افزایش یابد. اما در روش پیشنهادی، به جای استفاده از روش‌های مختلف استخراج ویژگی، تمرکز اصلی بر پیش‌پردازش و در نظر گرفتن اطلاعات زمانی و مکانی بوده است. در این مطالعه، انتخاب یک پنجره بزرگ‌تر، عملکرد مدل را به صورت قابل توجهی کاهش می‌دهد. در مقایسه با مطالعات قبلی، مدل پیشنهادی نیاز به پیش‌پردازش کمتر روی داده‌های خام دارد و آن را برای برنامه‌های کاربردی در زمان واقعی مانند BCI مناسب‌تر می‌کند. بالاترین دقت بدست آمده نیز متعلق به مرجع [۶۲] و برابر با ۹۸٫۶۳ درصد است. در روش پیشنهاد شده در این پایان‌نامه، با افزودن سیگنال fNIRS به سیگنال EEG و استفاده از اطلاعات مکانی و زمانی به صورت همزمان، دقت یک درصد افزایش داشته است.

از سوی دیگر، سیگنال fNIRS دارای دو طول موج است. ما تاثیر هر دو طول موج را به همراه سیگنال EEG بررسی کرده‌ایم. عملکرد مدل پیشنهادی برای EEG و هر کدام از طول موج‌های fNIRS در شکل ۳-۴ نشان داده شده‌اند.



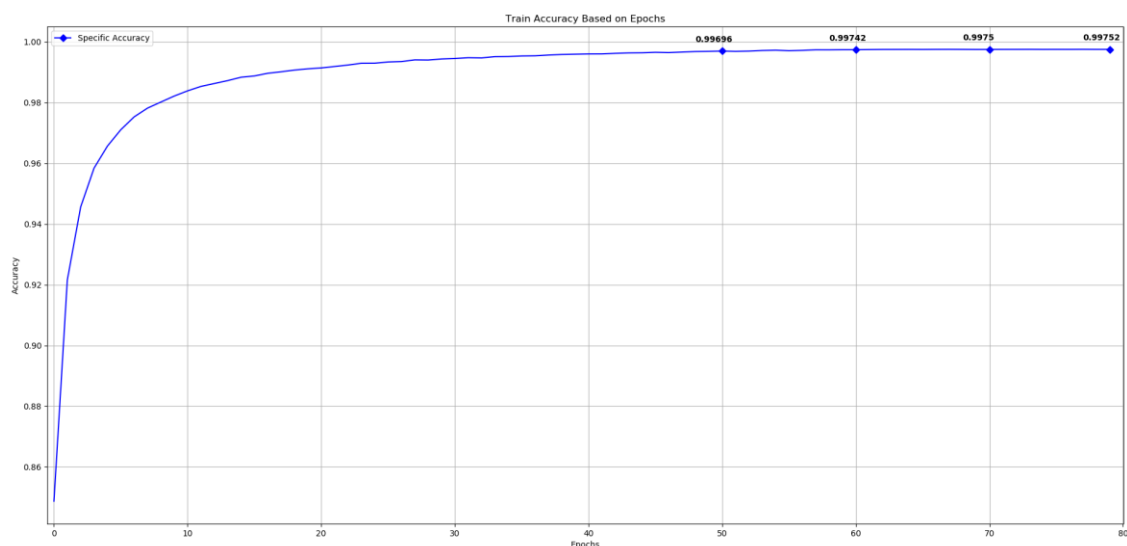
شکل ۳-۴ دقت الگوریتم پیشنهادی روی داده‌گان مختلف.

مطابق با شکل ۳-۴، دقت بدست آمده برای ترکیب سیگنال EEG با طول موج پایین‌تر fNIRS و سیگنال EEG با طول موج بالاتر fNIRS تقریباً با یکدیگر برابر هستند و می‌توان نتیجه گرفت که هر دو طول موج اطلاعات یکسانی را ارائه می‌دهد. در نتیجه ما می‌توانیم از هر کدام از این دو استفاده کنیم. مزیت استفاده از یک طول موج، کاهش بار محاسباتی سیستم است. اما باید در نظر داشت که دلیل افزایش ناچیز دقت هنگام وجود هر دو طول موج موجود در fNIRS، وجود اطلاعات بیشتر است. اما می‌توان افزایش ناچیز دقت را در نظر نگرفت و زمان اجرا را کوتاه‌تر کرد.

مطابق با شکل ۳-۴، دقت ترکیب EEG و fNIRS بالاتر از طول موج‌ها به صورت جدا است. دقت ۹۸ درصد برای سیگنال EEG و ۹۹٫۶۳ درصد برای ترکیب EEG-fNIRS نشان می‌دهد که افزودن سیگنال fNIRS به سیگنال EEG دقت تشخیص را افزایش می‌دهد. همانطور که قبلاً شرح داده شد، ثبت

سیگنال fNIRS به صورت همزمان با سیگنال EEG مشکلی ایجاد نمی‌کند. در نتیجه، استفاده همزمان از این دو سیگنال در روش پیشنهادی می‌تواند کمک بزرگی به تجزیه و تحلیل سیگنال‌ها و پیش‌بینی قصد کاربر و همچنین استفاده در سیستم‌های نیازمند دقت بالا مانند BCI کند.

یکی از مهم‌ترین پارامترها در شبکه‌های عصبی عمیق، تعداد مراحل آموزش است. یک مرحله آموزش، یک اجرای کامل از آموزش مدل است. در بیشتر موارد، انتخاب تعداد مراحل آموزش بالاتر منجر به آموزش بهتر شبکه می‌شود. اما در مراحل بالاتر، تغییرات دقت ممکن است ناچیز باشد. برای یافتن تعداد بهینه از این مراحل، باید به تغییرات دقت توجه کرد. نمودار دقت بر اساس تعداد مراحل آموزش این تغییرات را به وضوح نمایش می‌دهد.



شکل ۴-۴) نمودار دقت بر اساس تعداد مراحل آموزش.

شکل ۴-۴، نمودار دقت بر اساس تعداد مراحل آموزش را نمایش می‌دهد. همانطور که در این شکل مشاهده می‌شود، با افزایش تعداد مراحل، دقت افزایش می‌یابد. اما با افزایش تعداد مراحل از ۶۰ به ۸۰، تغییرات دقت ناچیز است و دقت کلی سیستم افزایش نمی‌یابد. از این رو، به منظور کاهش بار محاسباتی سیستم، تعداد ۶۰ مرحله برای تمامی اجراها لحاظ شده است.

یکی از پارامترهای مهم در این مسئله، پارامتر اندازه پنجره لغزان است. این اندازه تاثیر بسیار زیادی روی یادگیری و دقت شبکه عصبی عمیق تاثیر قابل توجهی دارد. جدول ۴-۲، دقت شبکه را در هنگام آموزش در اندازه‌های مختلف پنجره لغزان نشان می‌دهد. همانطور که از این جدول مشاهده می‌کنید، دقت شبکه با بزرگ‌تر شدن اندازه پنجره به صورت قابل توجهی کاهش می‌یابد. دلیل این اتفاق، وجود اطلاعات زیاد در پنجره است. اطلاعاتی که در پنجره‌های بزرگ‌تر وجود دارد، ممکن شامل چندین برچسب داده‌ای شود. از این رو کم شدن حجم اطلاعاتی و از طرف دیگر وجود اطلاعات زیادی برای آموزش یک برچسب، باعث کاهش دقت شبکه می‌شود. از طرف دیگر، کوچک‌تر کردن اندازه پنجره، آن را به سمت وجود یک داده در هر پنجره هدایت می‌کند که معنی پنجره‌گذاری را به طور کامل از بین می‌برد. از این رو، مقدار ۱۰ بهترین گزینه برای این شبکه است.

جدول ۴-۲) دقت شبکه با توجه به اندازه مختلف پنجره لغزان

اندازه پنجره لغزان (تعداد نمونه)	دقت شبکه
۱۰	۰,۹۹۶۳
۶۴	۰,۹۴۹۵
۱۰۰	۰,۹۰۹۱
۲۰۰	۰,۸۶۸۷

۴-۶ ارزیابی مدل پیشنهادی

در ادامه این بخش، به بررسی دقت روش پیشنهادی به کمک ماتریس درهم‌ریختگی پرداخته شده است. در هوش مصنوعی، ماتریس درهم‌ریختگی، نشان‌دهنده عملکرد الگوریتم‌های مربوطه است. معمولاً چنین نمایشی برای یادگیری با ناظر استفاده می‌شود. این ماتریس، پراکندگی توزیع دسته‌ها را از لحاظ درستی یا نادرستی نمایش می‌دهد. تعداد سطرها و ستون‌ها در این ماتریس به تعداد دسته‌ها وابسته

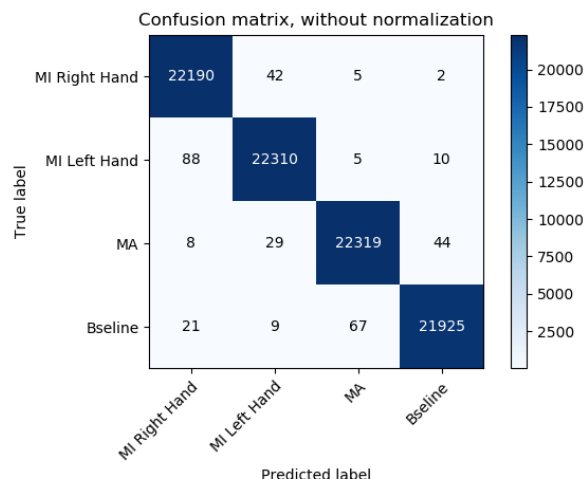
است. هر ستون از این ماتریس، نمونه پیش‌بینی شده و هر سطر، نمونه واقعی را نمایش می‌دهد. در ماتریس درهم‌ریختگی برای مدل پیشنهادی، سطر و ستون شامل برچسب‌های مرتبط با مجموعه داده هستند. این برچسب‌ها شامل تصور حرکتی دست چپ، تصور حرکتی دست راست، محاسبات ذهنی و حالت استراحت هستند. خروجی این ماتریس به صورت کلی برای یک مسئله با دو کلاس در شکل ۴-۵ نمایش داده شده است [۱۰۸].

		مقادیر واقعی	
		مثبت	منفی
درهم‌ریختگی	مثبت	درست مثبت TP	نادرست مثبت FP خطای نوع یک
	منفی	نادرست منفی FN خطای نوع دو	درست منفی TN

شکل ۴-۵) شکل کلی ماتریس درهم‌ریختگی برای یک مسئله دو کلاسه.

در حالت کلی می‌خواهیم مدل پیشنهادی، پیش‌بینی‌های نادرست را به حداقل رساند.

اطلاعات کامل طبقه‌بندی نمونه‌ها در این پایان‌نامه در ماتریس درهم‌ریختگی شکل ۴-۶ نمایش داده شده است. در شکل ۴-۶، جمعیت کل شامل ۸۹۰۷۲ داده است. تعداد داده‌های درست تشخیص داده شده، به نسبت سایر داده‌ها بسیار بزرگ‌تر است. در نتیجه انتظار می‌رود که دقت مدل نیز بالا باشد. دقت مدل با استفاده از معادله (۴-۱) محاسبه می‌شود.



شکل ۴-۶) ماتریس درهم‌ریختگی برای مدل پیشنهادی.

$$Classification\ Accuracy\ (CA) = \frac{TN + TP}{TN + FN + TP + FP} \quad (۴-۱)$$

در معادله (۴-۱)، صورت کسر نشان‌دهنده مقادیر صحیح تشخیص داده شده توسط مدل و مخرج کسر شامل نمونه‌های استفاده شده در مدل است. با جایگذاری مقادیر مربوطه از شکل ۴-۶ در این معادله، دقت مدل برابر با ۰,۹۹۶۲ است. کاملاً مشخص است که این دقت برابر با دقتی است که در بخش نتایج ذکر شده است. اما ما تنها به این معیار بسنده نمی‌کنیم و در ادامه به بررسی تعدادی دیگر از معیارهای موجود برای تشخیص عملکرد مدل پیشنهادی می‌پردازیم.

خطای طبقه‌بندی معیار دیگری است که برای عملکرد مدل استفاده شده است. این معیار بر خلاف معیار دقت، بهترین عملکرد مدل را با نزدیک شدن به صفر نشان می‌دهد. این معیار از طریق معادله (۴-۲) محاسبه می‌شود.

$$Error\ Rate = \frac{FN + FP}{TN + FN + TP + FP} = 1 - CA \quad (۴-۲)$$

در معادله (۴-۲) صورت کسر شامل تمام نمونه‌های غلط تشخیص داده شده توسط مدل و مخرج کسر، شامل تمام نمونه‌های استفاده شده در مدل است. مقدار خطا برای مدل پیشنهادی در این پایان‌نامه برابر با ۰,۰۰۳۸ است. میزان کم خطا و میزان بالای دقت نشان‌دهنده عملکرد خوب این مدل روی داده‌های پیش‌پردازش شده می‌باشد.

سه معیار مهم دیگر که برای تعیین عملکرد یک سیستم مهم است، معیارهای حساسیت یا بازخوانی، صحت و F1-score هستند. این سه معیار برای رسم نمودار ROC استفاده می‌شوند و به ترتیب با معادلات (۳-۴)، (۴-۴) و (۵-۴) نمایش داده می‌شوند. در جدول درهم ریختگی برای هر کدام از برچسب‌ها، یک مقدار از این سه معیار وجود دارد. حساسیت یا بازخوانی نشان‌دهنده تعداد نمونه‌های مثبت شناسایی شده به کل نمونه‌های مثبت است. صحت نشان‌دهنده تعداد نمونه‌های شناسایی شده مثبت به کل نمونه‌های نمونه‌های شناسایی شده است. معیار F1-score که با نام میانگین هارمونیک نیز شناخته می‌شود، یک معیار برای ترکیب دو مقدار صحت و حساسیت است تا درک بهتری نسبت به این دو مقدار رخ می‌دهد [۱۰۸].

$$Sensitivity (Recall) = \frac{TP}{TP + FN} \quad (۴-۳)$$

$$Precision = \frac{TP}{TP + FP} \quad (۴-۴)$$

$$F1 - Score: F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} = \frac{2TP}{2TP + FP + FN} \quad (۴-۵)$$

همانطور که از معادله (۵-۴) مشخص است، باید هر دو مقدار حساسیت و صحت مقداری بالا برای یک مدل باشند. در صورتی که یکی از این دو معیار کوچک باشد، صورت کسر عددی کوچک می‌شود و مقدار F1 کوچک می‌شود. از این رو هر چه دو مقدار صحت و حساسیت بالاتر و به یک نزدیک‌تر باشد، مقدار F1 بیشتر بوده و عملکرد مدل بهتر است.

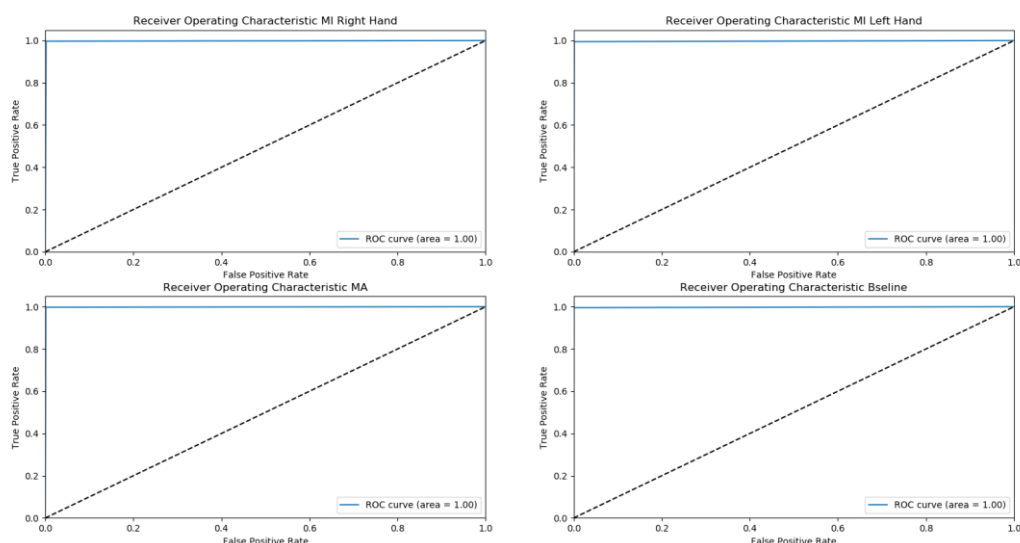
در جدول ۳-۴، مقادیر چهار معیار ذکر شده برای هر کدام از برچسب‌ها به صورت مجزا آمده است.

جدول (۳-۴) چهار معیار صحت، حساسیت، F1 و AUC برای مدل پیشنهادی.

نام کلاس فعالیت	صحت	حساسیت	F1	مساحت سطح زیر نمودار

تصور حرکتی دست راست	۰,۹۹۴۶	۰,۹۹۷۴	۰,۹۹۶۰	۰,۹۹۹۵
تصور حرکتی دست چپ	۰,۹۹۶۸	۰,۹۹۵۲	۰,۹۹۶۰	۰,۹۹۹۴
محاسبه ذهنی	۰,۹۹۵۹	۰,۹۹۷۰	۰,۹۹۶۵	۰,۹۹۹۶
استراحت	۰,۹۹۷۷	۰,۹۹۵۴	۰,۹۹۶۵	۰,۹۹۹۶
میانگین	۰,۹۹۶۲	۰,۹۹۶۲	۰,۹۹۶۲	۰,۹۹۹۵

شکل ۴-۷ نمودار ROC را برای هر کدام از کلاس‌ها به صورت جدا نمایش می‌دهد. نمودار ROC، میزان جداکنندگی یک مدل یا دقت آن را در تمیزدادن بین دو یا چند دسته نمایش می‌دهد. این نمودار از دو مقدار صحت و حساسیت تشکیل می‌شود. به طور دقیق‌تر، محور عمودی این نمودار، نرخ تولید داده‌های درست یعنی TPR و محور افقی، نرخ تولید خطا برای داده‌های مثبت یعنی FPR را نمایش می‌دهد. با این توصیف، نموداری مناسب است که محور عمودی آن به یک نزدیک‌تر و محور افقی آن به صفر نزدیک‌تر باشد. همانطور که در شکل ۴-۷ مشاهده می‌کنید، تمام تصاویر دارای نرخ تولید خطای صفر و نرخ بالای تولید داده‌های درست هستند. اما برای بررسی دقیق‌تر این نمودارها، می‌توان از مساحت سطر زیر نمودار (AUC) استفاده کرد. این مقدار عددی بین صفر و یک است که هر چه به یک نزدیک‌تر باشد، نشان از دقت بیشتر مدل در تشخیص داده‌های درست است [۱۰۸]. میزان مساحت زیر سطح نمودار برای هر کدام از کلاس‌ها در جدول ۴-۳ آمده است. این مقادیر برای هر کلاس در شکل ۴-۷ نیز ذکر شده است.



شکل ۷-۴) نمودار ROC مرتبط با چهار فعالیت ذهنی مختلف.

تمام معیارها نشان از کارا بودن عملکرد مدل پیشنهادی دارد و مدل پیشنهادی با داشتن دقت بالای ۹۹ درصد، بالاترین دقت را در بین مدل‌های دیگر در زمینه طبقه‌بندی سیگنال‌های بیولوژیکی دارد.

۷-۴ نتیجه‌گیری

در این بخش، نتایج حاصل از پیش‌پردازش و مدل یادگیری عمیق پیشنهادی بررسی شد. در روش پیش‌پردازش پیشنهادی، داده‌های EEG به تانسورهای سه بعدی تبدیل شدند. این تانسورها تصاویر توپوگرافی مغز را نمایش می‌دادند. نتایج نشان دادند که اطلاعات موجود در این تانسورهای سه بعدی، خواناتر از سیگنال‌های خام بودند. در ادامه تغییر فعالیت مغز در هنگام انجام فعالیت‌های مختلف، به کمک خروجی پیش‌پردازش پیشنهادی نشان داده شد. سپس نتایج روش یادگیری عمیق RCNN پیشنهادی بررسی شد. آزمایش‌های مختلف نشان دادند که ترکیب داده‌های EEG و fNIRS و همچنین استفاده از یادگیری عمیق، دقت را به طور چشمگیری افزایش می‌دهد. این افزایش دقت، در مقایسه با روش‌های دیگر بهترین بود. در این بخش آزمایش‌های مختلفی مورد بررسی قرار گرفت. آزمایش‌ها نشان دادند که دقت روش پیشنهادی با استفاده از EEG و fNIRS، بالاتر از مدل‌های دیگر است. در

این فصل نشان داده شد که استفاده از ویژگی‌های زمانی و مکانی در سیگنال‌ها و همچنین استفاده از یک شبکه RCNN، می‌تواند دقت طبقه‌بندی را به طور چشم‌گیری افزایش دهد.

فصل ۵ : نتیجه گیری

در این فصل، الگوریتم ارائه شده را جمع بندی و نقاط قوت و ضعف آن را بررسی می کنیم. همچنین پیشنهادهایی برای کارهای آتی ارائه می دهیم.

۵-۱ جمع بندی و نتیجه گیری

در این پایان نامه، یک سیستم طبقه بندی مبتنی بر یادگیری عمیق برای تشخیص فعالیت ذهنی با استفاده از دو سیگنال EEG و fNIRS ارائه شد. در این پایان نامه، تمرکز بر دو بخش پیش پردازش داده ها جهت آماده سازی برای مدل طبقه بند و الگوریتم طبقه بندی بود. در این پایان نامه، پیش پردازشی جهت حذف نویز احتمالی از سیگنال انجام نشد. روش پیش پردازش پیشنهادی، سیگنال را به تصاویر حاوی اطلاعات مفیدتر تبدیل کرد و آن را برای طبقه بندی و ورودی مدل آماده ساخت.

این سیستم بر اساس ویژگی های زمانی و مکانی در سیگنال ها کار می کند. روش پیش پردازش پیشنهادی، داده های EEG و fNIRS را به تانسورهای سه بعدی تبدیل می کند. این روش به کمک یک تابع تبدیل کار می کند. این تابع تبدیل، هر نمونه از دو سیگنال EEG و fNIRS را در زمان t درون یک تانسور مطابق با موقعیت الکترودها روی پوست سر مقداردهی می کند. خروجی این روش، تصاویر توپوگرافی مغز است. این تصاویر توپوگرافی نشان دهنده فعالیت های مغز در زمان های مختلف است. در این پایان نامه از روش پنجره گذاری لغزان برای تبدیل این تانسورها به کلیپ های داده استفاده شده است. در ادامه مدلی مبتنی بر یادگیری عمیق جهت طبقه بندی تانسورهای سه بعدی ارائه شد. این مدل، ترکیبی از دو بلوک CNN و RNN بود. بلوک CNN، ویژگی های مکانی را از تانسورها استخراج کرد. بلوک RNN، ویژگی های زمانی را از داده ها استخراج کرد. طبقه بند استفاده شده در این مدل Softmax بود. مدل پیشنهادی به خوبی قادر به تشخیص فعالیت های ذهنی بود.

روش پیشنهادی به خوبی قادر به رفع چالش های ذکر شده در بخش ۱-۱ بود. این روش در مقایسه با روش های دیگر، دقت بالاتری را بدست آورد. نتایج نشان دادند که استفاده از اطلاعات مکانی و زمانی، دقت تشخیص را به صورت قابل توجهی افزایش داد. از سوی دیگر، تاثیر افزودن سیگنال fNIRS به

سیگنال EEG نیز مورد بررسی قرار گرفت. نتایج نشان دادند که ترکیب این دو سیگنال بر دقت طبقه-بندی تاثیر گذاشت.

۵-۲ پیشنهاد برای کارهای آتی

در این پایان نامه، هیچ الگوریتمی جهت حذف نویز استفاده نشده است. حذف نویز کمک بزرگی به تشخیص بهتر فعالیت ذهنی می کند. در کارهای آتی می توان با استفاده از الگوریتم های مختلف حذف نویز، دقت را بهبود داد. از طرف دیگر، یافتن پارامترهای بهینه برای مدل شبکه عمیق، یکی دیگر از چالش هایی است که در کارهای آتی می تواند مورد بررسی قرار گیرد. انتخاب بهینه پارامترها یک شبکه عمیق می تواند هزینه محاسباتی سیستم، زمان اجرا و دقت را به طور چشمگیری بهبود بخشد. یکی دیگر از کارهایی که می توان در آینده به آن پرداخت، استفاده و ترکیب مجموعه داده های دیگر برای افزایش داده و همچنین افزایش کلاس ها است. افزایش کلاس ها برای یک سیستم BCI می تواند منجر به افزایش دستورات صادر شده توسط کاربر شود. با این کار یک سیستم BCI می تواند به دنیای واقعی نزدیک تر شود.



- [1] Hofman M. A., and Falk D. (Eds.) (2012) "Evolution of the primate brain: from neuron to behavior" *Elsevier*, Vol. 195.
- [2] Buzsaki G. (2006) "Rhythms of the Brain" *Oxford University Press*.
- [3] Melillo R., and Leisman G. (2010) "Neurobehavioral disorders of childhood: An evolutionary perspective" *Springer Science & Business Media*.
- [4] Illes J., and Sahakian B. J. (Eds.) (2013) "Oxford handbook of neuroethics" *Oxford University Press*.
- [5] Rebsamen B., Burdet E., Guan C., Zhang H., Teo C.L., Zeng Q., Laugier C. and Ang M.H. (2007) "Controlling a wheelchair indoors using thought" *IEEE intelligent systems* 2, pp18-24.
- [6] Guo F., Hong B., Gao X. and Gao S. (2008) "A brain-computer interface using motion-onset visual evoked potential." *Journal of neural engineering* 5.4, pp477.
- [7] Leeb R., Scherer R., Lee F., Bischof H. and Pfurtscheller G. (2004) "Navigation in virtual environments through motor imagery." *9th Computer Vision Winter Workshop, CVWW*. Vol. 4.
- [8] Ware C. and Mikaelian H.H. (1987) "An evaluation of an eye tracker as a device for computer input2." *Acm sigchi bulletin*. Vol. 17. No. SI. ACM.
- [9] Williams M.R. and Kirsch R.F. (2008) "Evaluation of head orientation and neck muscle EMG signals as command inputs to a human-computer interface for individuals with high tetraplegia." *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 16.5, pp485-496.
- [10] Bauer G., Gerstenbrand F. and Rumpl E. (1979) "Varieties of the locked-in syndrome." *Journal of neurology* 221.2, pp77-91.
- [11] Patterson J.R. and Grabois M. (1986) "Locked-in syndrome: a review of 139 cases." *Stroke* 17.4, pp758-764.
- [12] Kübler A., Kotchoubey B., Kaiser J., Wolpaw J.R. and Birbaumer N. (2001) "Brain-computer communication: Unlocking the locked in." *Psychological bulletin* 127.3, pp358.
- [13] Wolpaw J.R., Birbaumer N., Heetderks W.J., McFarland D.J., Peckham P.H., Schalk G., Donchin E., Quatrano L.A., Robinson C.J. and Vaughan T.M. (2000) "Brain-computer interface technology: a review of the first international meeting." *IEEE transactions on rehabilitation engineering* 8.2, pp164-173.
- [14] Khan M.J. and Hong K.S. (2017) "Hybrid EEG-fNIRS-based eight-command decoding for BCI: application to quadcopter control." *Frontiers in neurorobotics* 11, pp6.
- [15] Lotte F., Congedo M., Lécuyer A., Lamarche F. and Arnaldi B. (2007) "A review of classification algorithms for EEG-based brain-computer interfaces." *Journal of neural engineering* 4.2, ppR1.
- [16] Vaadia E. and Birbaumer N. (2009) "Grand challenges of brain computer interfaces in the years to come." *Frontiers in neuroscience* 3, pp15.

- [17] Blankertz B., Tangermann M., Vidaurre C., Fazli S., Sannelli C., Haufe S., Maeder C., Ramsey L.E., Sturm I., Curio G. and Mueller K.R. (2010) "The Berlin brain-computer interface: non-medical uses of BCI technology." *Frontiers in neuroscience* 4, pp198.
- [18] Van Erp J., Lotte F. and Tangermann M. (2012) "Brain-computer interfaces: beyond medical applications." *Computer* 45.4, pp26-34.
- [19] Tabar Y.R. and Halici U. (2017) "Brain Computer Interfaces for Silent Speech." *European Review* 25.2, pp208-230.
- [20] Wolpaw J.R., Loeb G.E., Allison B.Z., Donchin E., do Nascimento O.F., Heetderks W.J., Nijboer F., Shain W.G. and Turner J.N. (2006) "BCI meeting 2005-workshop on signals and recording methods." *IEEE Transactions on neural systems and rehabilitation engineering* 14.2, pp138-141.
- [21] Nicolas-Alonso L.F. and Gomez-Gil J. (2012) "Brain computer interfaces, a review." *Sensors* 12.2, pp1211-1279.
- [22] Gelenbe E., Feng Y. and Krishnan K.R.R. (1996) "Neural network methods for volumetric magnetic resonance imaging of the human brain." *Proceedings of the IEEE* 84.10, pp1488-1496.
- [23] Bauernfeind G., Leeb R., Wriessnegger S.C. and Pfurtscheller G. (2008) "Development, set-up and first results for a one-channel near-infrared spectroscopy system/Entwicklung, Aufbau und vorläufige Ergebnisse eines Einkanal-Nahinfrarot-Spektroskopie-Systems." *Biomedizinische Technik* 53.1, pp36-43.
- [24] Coyle S.M., Ward T.E. and Markham C.M. (2007) "Brain-computer interface using a simplified functional near-infrared spectroscopy system." *Journal of neural engineering* 4.3, pp219.
- [25] Power S.D., Kushki A. and Chau T. (2011) "Towards a system-paced near-infrared spectroscopy brain-computer interface: differentiating prefrontal activity due to mental arithmetic and mental singing from the no-control state." *Journal of neural engineering* 8.6, 066004.
- [26] Brumberg J.S., Nieto-Castanon A., Kennedy P.R. and Guenther F.H. (2010) "Brain-computer interfaces for speech communication." *Speech communication* 52.4, pp367-379.
- [27] Pfurtscheller G. and Da Silva F.L. (1999) "Event-related EEG/MEG synchronization and desynchronization: basic principles." *Clinical Neurophysiology*, 110, 11, pp1842-1857.
- [28] Schlögl A., Lee F., Bischof H. and Pfurtscheller G. (2005) "Characterization of four-class motor imagery EEG data for the BCI-competition 2005." *Journal of Neural Engineering*, 2, ppL14-L22.
- [29] Pfurtscheller G., Brunner C., Schlögl A. and Da Silva F.L. (2006) "Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks." *NeuroImage*, 31, pp153-159.
- [30] Teo W.P., White D. and Macpherson H. (2018) "Using noninvasive methods to drive brain-computer interface (BCI): the role of electroencephalography and functional near-infrared spectroscopy in BCI." *Smart Wheelchairs and Brain-Computer Interfaces*, pp33-63.

- [31] Trejo L.J., Rosipal R. and Matthews B. (2006) "Brain-computer interfaces for 1-D and 2-D cursor control: designs using volitional control of the EEG spectrum or steady-state visual evoked potentials." *IEEE transactions on neural systems and rehabilitation engineering* 14.2, pp225-229.
- [32] Turnip A., Hong K.S. and Jeong M.Y. (2011) "Real-time feature extraction of P300 component using adaptive nonlinear principal component analysis." *Biomedical engineering online* 10.1, pp83.
- [33] Turnip A. and Hong K.S. (2012) "Classifying mental activities from EEG-P300 signals using adaptive neural network." *Int. J. Innov. Comp. Inf. Control* 8.7, pp5839-5850.
- [34] Wang D., Miao D. and Blohm G. (2012) "Multi-class motor imagery EEG decoding for brain-computer interfaces." *Frontiers in neuroscience* 6, pp151.
- [35] Wang H., Zhang Y., Waytowich N.R., Krusienski D.J., Zhou G., Jin J., Wang X. and Cichocki A. (2016) "Discriminative feature extraction via multivariate linear regression for SSVEP-based BCI." *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 24.5, pp532-541.
- [36] Gruzelier J.H. (2014) "EEG-neurofeedback for optimising performance. I: a review of cognitive and affective outcome in healthy participants." *Neuroscience & Biobehavioral Reviews* 44, pp124-141.
- [37] Ahn M. and Jun S.C. (2015) "Performance variation in motor imagery brain-computer interface: a brief review." *Journal of neuroscience methods* 243, pp103-110.
- [38] Wang H., Li Y., Long J., Yu T. and Gu Z. (2014) "An asynchronous wheelchair control by hybrid EEG-EOG brain-computer interface." *Cognitive neurodynamics* 8.5, pp399-409.
- [39] Ramli R., Arof H., Ibrahim F., Mokhtar N. and Idris M.Y.I. (2015) "Using finite state machine and a hybrid of EEG signal and EOG artifacts for an asynchronous wheelchair navigation." *Expert Systems with Applications* 42.5, pp2451-2463.
- [40] Zhang R., Li Y., Yan Y., Zhang H., Wu S., Yu T. and Gu Z. (2016) "Control of a wheelchair in an indoor environment based on a brain-computer interface and automated navigation." *IEEE transactions on neural systems and rehabilitation engineering* 24.1, pp128-139.
- [41] Kim B.H., Kim M. and Jo S. (2014) "Quadcopter flight control using a low-cost hybrid interface with EEG-based classification and eye tracking." *Computers in biology and medicine* 51, pp82-92.
- [42] Fazli S., Mehnert J., Steinbrink J., Curio G., Villringer A., Müller K.R. and Blankertz B. (2012) "Enhanced performance by a hybrid NIRS-EEG brain computer interface." *Neuroimage* 59.1, pp519-529.
- [43] Pfurtscheller G., Allison B.Z., Bauernfeind G., Brunner C., Solis Escalante T., Scherer R., Zander T.O., Mueller-Putz G., Neuper C. and Birbaumer N. (2010) "The hybrid BCI." *Frontiers in neuroscience* 4, 3.
- [44] Shin J., von Lühmann A., Blankertz B., Kim D.W., Jeong J., Hwang H.J. and Müller K.R. (2017) "Open access dataset for EEG+ NIRS single-trial classification." *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 25.10, pp1735-1745.

- [45] Shin J., Müller K.R., Schmitz C.H., Kim D.W. and Hwang H.J. (2017) "Evaluation of a compact hybrid brain-computer interface system." *BioMed research international* 2017.
- [46] Biessmann F., Plis S., Meinecke F.C., Eichele T. and Muller K.R. (2011) "Analysis of multimodal neuroimaging data." *IEEE Reviews in Biomedical Engineering* 4, pp26-58.
- [47] Friston K.J. (2009) "Modalities, modes, and models in functional neuroimaging." *Science* 326.5951, pp399-403.
- [48] Shibasaki H. (2008) "Human brain mapping: hemodynamic response and electrophysiology." *Clinical Neurophysiology* 119.4, pp731-743.
- [49] Nguyen D.K., Tremblay J., Pouliot P., Vannasing P., Florea O., Carmant L., Lepore F., Sawan M., Lesage F. and Lassonde M. (2012) "Non-invasive continuous EEG-fNIRS recording of temporal lobe seizures." *Epilepsy research* 99.1-2, pp112-126.
- [50] Peng K., Nguyen D.K., Tayah T., Vannasing P., Tremblay J., Sawan M., Lassonde M., Lesage F. and Pouliot P. (2014) "fNIRS-EEG study of focal interictal epileptiform discharges." *Epilepsy research* 108.3, pp491-505.
- [51] Seyal M. (2014) "Frontal hemodynamic changes precede EEG onset of temporal lobe seizures." *Clinical Neurophysiology* 125.3, pp442-448.
- [52] Wallois F., Patil A., Héberlé C. and Grebe R. (2010) "EEG-NIRS in epilepsy in children and neonates." *Neurophysiologie Clinique/Clinical Neurophysiology* 40.5-6, pp281-292.
- [53] Herrmann M.J., Hutter T., Plichta M.M., Ehlis A.C., Alpers G.W., Mühlberger A. and Fallgatter A.J. (2008) "Enhancement of activity of the primary visual cortex during processing of emotional stimuli as measured with event-related functional near-infrared spectroscopy and event-related potentials." *Human Brain Mapping* 29.1, pp28-35.
- [54] Näsi T., Kotilahti K., Noponen T., Nissilä I., Lipiäinen L. and Meriläinen P. (2010) "Correlation of visual-evoked hemodynamic responses and potentials in human brain." *Experimental brain research* 202.3, pp561-570.
- [55] Elsayed N., Zaghloul Z.S. and Bayoumi M. (2017) "Brain computer interface: Eeg signal preprocessing issues and solutions" *International Journal of Computer Applications* 169 (3), pp975–8887.
- [56] Sun S. and Zhou J. (2014) "A review of adaptive feature extraction and classification methods for eeg-based brain-computer interfaces" *2014 International Joint Conference on Neural Networks (IJCNN), IEEE*, pp1746–1753.
- [57] Heydari E. and Shahbakhti M. (2015) "Adaptive wavelet technique for eeg de-noising" *2015 8th Biomedical Engineering International Conference (BMEiCON), IEEE*, pp1–4.
- [58] Yin Z., Wang Y., Liu L., Zhang W. and Zhang J. (2017) "Cross-subject eeg feature selection for emotion recognition using transfer recursive feature elimination" *Frontiers in neurorobotics* 11 pp19.
- [59] Yang J., Yao S. and Wang J. (2018). "Deep Fusion Feature Learning Network for MI-EEG Classification." *IEEE Access*, 6, pp79050-79059.

- [60] Zhang D., Yao L., Zhang X., Wang S., Chen W. and Boots R. (2017). "EEG-based Intention Recognition from Spatio-Temporal Representations via Cascade and Parallel Convolutional Recurrent Neural Networks." *arXiv preprint arXiv*, 1708.06578.
- [61] Hong K.S., Khan M.J. and Hong M.J. (2018) "Feature extraction and classification methods for hybrid fNIRS-EEG brain-computer interfaces." *Frontiers in human neuroscience* 12.
- [62] Chiarelli A.M., Croce P., Merla A. and Zappasodi F. (2018) "Deep learning for hybrid EEG-fNIRS brain-computer interface: application to motor imagery classification." *Journal of neural engineering* 15.3, 036028.
- [63] Tanaka T. and Arvaneh M. (2018) "Signal processing and machine learning for brain-machine interfaces." *Institution of Engineering & Technology*.
- [64] Gandhi V. (2014) "Brain-Computer Interfacing for Assistive Robotics: Electroencephalograms, Recurrent Quantum Neural Networks, and User-Centric Graphical Interfaces." *Academic Press*.
- [65] Pearson K. (1901) "On lines and planes of closest fit to systems of points in space" *Philos. Mag. Ser. 6* 2, pp559-572.
- [66] Jutten C. and Herault J. (1991) "Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture" *Signal Process.* 24, pp110.
- [67] Pfurtscheller G., Neuper C., Flotzinger D. and Pregenzer M. (1997) "EEG-based discrimination between imagination of right and left hand movement, Electroencephalogr." *Clin. Neurophysiol.* 103, 642651.
- [68] Stoica P. and Moses R.L. (1997) "Introduction to Spectral Analysis, Prentice Hall" *Upper Saddle River, NJ*.
- [69] Fukunaga K. (1972) "Introduction to Statistical Pattern Recognition" *Oxford, UK: Clarendon*.
- [70] Cortes C. and Vapnik V. (1995) "Support-vector networks. *Machine Learning*" 20, pp273–297.
- [71] Oh S.L., Hagiwara Y., Raghavendra U., Yuvaraj R., Arunkumar N., Murugappan M. and Acharya U.R. (2018) "A deep learning approach for Parkinson's disease diagnosis from EEG signals." *Neural Computing and Applications*, pp1-7.
- [72] Acharya U.R., Fujita H., Lih O.S., Hagiwara Y., Tan J.H. and Adam M. (2017) "Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network." *Information Sciences* 405, pp81-90.
- [73] Acharya U.R., Fujita H., Lih O.S., Adam M., Tan J.H. and Chua C.K. (2017) "Automated detection of coronary artery disease using different durations of ECG segments with convolutional neural network." *Knowledge-Based Systems* 132, pp62-71.
- [74] Tan J.H., Acharya U.R., Bhandary S.V., Chua K.C. and Sivaprasad S. (2017) "Segmentation of optic disc, fovea and retinal vasculature using a single convolutional neural network." *Journal of Computational Science* 20, pp70-79.
- [75] Acharya U.R., Fujita H., Oh S.L., Hagiwara Y., Tan J.H., Adam M. and San Tan R. (2019) "Deep convolutional neural network for the automated

diagnosis of congestive heart failure using ECG signals.” *Applied Intelligence*, Springer US, pp1-12.

- [76] Patterson J. and Gibson A. (2017) “Deep learning: A practitioner's approach.” *O'Reilly Media, Inc.*
- [77] LeCun Y., Kavukcuoglu K. and Farabet C. (2010) “Convolutional networks and applications in vision” *Proceedings of 2010 IEEE International Symposium on Circuits and Systems, IEEE*, pp253–256.
- [78] Acharya U.R., Oh S.L., Hagiwara Y., Tan J.H. and Adeli H. (2017) “Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals.” *Computers in Biology and Medicine* DOI: <https://doi.org/10.1016/j.combiomed.2017.09.017>
- [79] Pereira S., Pinto A., Alves V. and Silva C.A. (2016). “Brain tumor segmentation using convolutional neural networks in MRI images” *IEEE Transactions on Medical Imaging*, 35(5), pp1240-1251.
- [80] Hatipoglu N. and Bilgin G. (2017). “Cell segmentation in histopathological images with deep learning algorithms by utilizing spatial relationships” *Medical and Biological Engineering and Computing*, pp1-20, doi: 10.1007/s11517-017-1630-1.
- [81] Setio A.A.A., Ciompi F., Litjens G., Gerke P., Jacobs C., Van Riel S.J., Wille M.M.W., Naqibullah M., Sánchez C.I. and van Ginneken B. (2016) “Pulmonary nodule detection in CT images: False positive reduction using multi-view convolutional networks” *IEEE Transactions on Medical Imaging*, 35(5), pp1160-1169.
- [82] Kallenberg M., Petersen K., Nielsen M., Ng A.Y., Diao P., Igel C., Vachon C.M., Holland K., Winkel R.R., Karssemeijer N. and Lillholm M. (2016). “Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring” *IEEE Transactions on Medical Imaging*, 35(5), pp1322-1331.
- [83] Acharya U.R., Fujita H., Oh S.L., Hagiwara Y., Tan J.H. and Adam M. (2017) “Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals” *Information Sciences*, 415-416, pp190-198.
- [84] Acharya U.R., Oh S.L., Hagiwara Y., Tan J.H., Adam M., Gertych A. and San Tan R. (2017) “A deep convolutional neural network model to classify heartbeats” *Computers in Biology and Medicine*, DOI: 10.1016/j.combiomed.2017.08.022.
- [85] Krumholz A., Wiebe S., Gronseth G., Shinnar S., Levisohn P., Ting T., Hopp J., Shafer P., Morris H., Seiden L. and Barkley G. (2007) “Quality Standards Subcommittee of the American Academy of Neurology. Practice Parameter: evaluating an apparent unprovoked first seizure in adults (an evidence-based review): report of the Quality Standards Subcommittee of the American Academy of Neurology and the American Epilepsy Society” *American Epilepsy Society, Neurology*, 69, 21.
- [86] Acharya U.R., Fujita H., Oh S.L., Raghavendra U., Tan J.H., Adam M., Gertych A. and Hagiwara Y. (2018) “Automated identification of shockable and

- non-shockable life-threatening ventricular arrhythmias using convolutional neural network.” *Future Generation Computer Systems* 79, 3, pp952- 959.
- [87] Faust O., Hagiwara Y., Hong T.J., Lih O.S. and Acharya U.R. (2018) “Deep learning for healthcare applications based on physiological signals: A review.” *Computer Methods and Programs in Biomedicine* 161, pp1- 13.
- [88] Riesenhuber M. and Poggio T. (1999) “Hierarchical models of object recognition in cortex.” *Nature Neuroscience* 2, pp1019–1025.
- [89] Ciresan D.C., Meier U., Masci J. and Schmidhuber J. (2011) “A committee of neural networks for traffic sign classification.” 396 *Proceedings of the 2011 International Joint Conference on Neural Networks, IEEE, California* 397, pp1918-1921.
- [90] Scherer D., Müller A. and Behnke S. (2010) “Evaluation of pooling operations in convolutional architectures for object recognition.” *International conference on artificial neural networks Springer*, pp82–91.
- [91] Serre T., Wolf L. and Poggio T. (2005) “Object recognition with features inspired by visual cortex.” *Computer Vision and Pattern Recognition Conference*, pp994–1000.
- [92] Gers F.A., Schmidhuber J. and Cummins F. (1999) “Learning to forget: Continual prediction with LSTM.”, pp850-855.
- [93] Sutskever I., Vinyals O. and Le Q.V. (2014). “Sequence to sequence learning with neural networks.” *Advances in neural information processing systems*, pp3104-3112.
- [94] Peng H., Chao J., Wang S., Dang J., Jiang F., Hu B. and Majoe D. (2018). “Single-trial classification of fNIRS signals in four directions motor imagery tasks measured from prefrontal cortex.” *IEEE Transactions on NanoBioscience*.
- [95] Shin J., Kwon J. and Im C.H. (2018). A Ternary Hybrid EEG-NIRS Brain-Computer Interface for the Classification of Brain Activation Patterns during Mental Arithmetic, Motor Imagery, and Idle State. *Frontiers in neuroinformatics*, 12, 5.
- [96] Zafar A., Khan M.J., Park J. and Hong K.S. (2018). Initial-dip Based Quadcopter Control: Application to fNIRS-BCI. *IFAC-PapersOnLine*, 51(15), pp945-950.
- [97] Yoo S.H., Woo S.W. and Amad Z. (2018) “Classification of three categories from prefrontal cortex using LSTM networks: fNIRS study.” 2018 18th *International Conference on Control, Automation and Systems (ICCAS), IEEE*, pp1141-1146.
- [98] Mustafa I. and Mustafa I. (2018) “Smart thoughts: BCI based system implementation to detect motor imagery movements.” *Applied Sciences and Technology (IBCAST), 2018 15th International Bhurban Conference on, IEEE*, pp.365-371.
- [99] Pandey P. and Seeja K.R. (2019) “Subject-Independent Emotion Detection from EEG Signals Using Deep Neural Network.” *International Conference on Innovative Computing and Communications, Springer, Singapore*, pp41-46.
- [100] Yohanandan S.A., Kiral-Kornek I., Tang J., Mshford B.S., Asif U. and Harrer S. (2018) “A Robust Low-Cost EEG Motor Imagery-Based Brain-

- Computer Interface.” *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE*, pp5089-5092.
- [101] Zeng H., Yang C., Dai G., Qin F., Zhang J. and Kong W. (2018) “EEG classification of driver mental states by deep learning.” *Cognitive neurodynamics*, 12, 6, pp597-606.
 - [102] Janani A. and Sasikala M. (2018) “Classification of fNIRS Signals for Decoding Right-and Left-Arm Movement Execution Using SVM for BCI Applications.” *Computational Signal Processing and Analysis, Springer, Singapore*, pp315-323.
 - [103] Ho T.K.K., Gwak J., Park C.M., Khare A. and Song J.I. (2019) “Deep Learning-Based Approach for Mental Workload Discrimination from Multi-channel fNIRS.” *Recent Trends in Communication, Computing, and Electronics, Springer, Singapore*, pp431-440.
 - [104] Zhang M., Hua Q., Jia W., Chen R., Su H. and Wang B. (2018) “Feature Extraction and Classification Algorithm of Brain-computer Interface Based on Human Brain Central Nervous System.” *NeuroQuantology*, 16, 5.
 - [105] Shin J., Kim D.W., Müller K.R. and Hwang H.J. (2018) “Improvement of Information Transfer Rates Using a Hybrid EEG-NIRS Brain-Computer Interface with a Short Trial Length: Offline and Pseudo-Online Analyses.” *Sensors (Basel, Switzerland)*, 18, 6.
 - [106] Lee C.Y., Huang T.S., Liu M.K. and Lan C.Y. (2019) “Data Science for Vibration Heteroscedasticity and Predictive Maintenance of Rotary Bearings.” *Energies* 12.5, pp801.
 - [107] Donahue J., Anne Hendricks L., Guadarrama S., Rohrbach M., Venugopalan S., Saenko K. and Darrell T. (2015) “Long-term recurrent convolutional networks for visual recognition and description.” *Proceedings of the IEEE conference on computer vision and pattern recognition*.
 - [108] Davis J. and Goadrich M. (2006) “The relationship between Precision-Recall and ROC curves.” *Proceedings of the 23rd international conference on Machine learning. ACM*.

لغت نامه

بر اساس حروف الفبای فارسی

عنوان فارسی	عنوان انگلیسی
اسکلروز جانبی آمیوتروفیک	Amyotrophic Lateral Sclerosis (ALS)
اسکن مغزی	Neuroimaging
التهاب مغز یا آنسفالیت	Encephalitis
الکتروانسفالوگرافی	Electroencephalography
الکتروفیزیولوژیکی	Electrophysiological
امواج صرعی بین حمله‌ای یا صرع	Interictal Epileptic Discharges (IED)
باترورث	Butterworth
بیش‌برازش	Overfitting
بیماری نورون‌های حرکتی	Motor Neuron Disease (MND)
پتانسیل مرتبط با رخداد	Event-Related Potential (ERP)
پتانسیل‌های رخداد دیداری حالت ایستا	Steady State Visual Evoked Potentials (SSVEP)
پتانسیل‌های قشری آهسته	Slow Cortical Potentials (SCP)
پیشانی	Forehead
پیشانی	Frontal
تبدیل موجک	Wavelet Transform
تجزیه حالت کلی گروهی	Ensemble Empirical Mode Decomposition
تجزیه و تحلیل اجزای اصلی	Principal Component Analysis (PCA)
تجزیه و تحلیل اجزای مستقل	Independent Component Analysis (ICA)
تجزیه و تحلیل تفکیک خطی	Linear Discriminant Analysis (LDA)
تجزیه و تحلیل تفکیک خطی انقباضی	shrinkage Linear Discriminant Analysis
تشکیل کلمه	Word Formation

عنوان فارسی	عنوان انگلیسی
تصاویر توموگرافی تخمینی	Computed Tomography Images
تصاویر رَحِم	Fundus Images
تصاویر هیستوپاتولوژی	Histopathologic Images
تصاویرسازی تشدید مغناطیسی	Magnetic Resonance Imaging (MRI)
تصویرات حرکتی	Motor Imagery (MI)
تصویربرداری رزونانس مناطیسی تابعی	Functional Magnetic Resonance Imaging
تکانه	Momentum
تنظیم مدل	Regularization
توابع شناختی	Cognitive Function
توان باند	Band Power (BP)
جداری	Parietal
جلویی	Prefrontal
چگالی طیف توان	Power Spectral Density (PSD)
حافظه کوتاه و بلند مدت	Long-Short Term Memory
حرکات عضلات	Electromyography (EMG)
حرکت چشم	Electrooculography (EOG)
دوره استراحت	Relaxation
دیداری	Visual
رابط مغز-کامپیوتر	Brain-Computer Interface (BCI)
ریتم‌های حسی-حرکتی	Sensory-Motor Rhythms (SMR)
زمان-فرکانس	Time-Frequency
سازگاری مرتبط با رخداد	Event-related synchronization (ERS)
شبکه عصبی کانولوشن	Convolutional Neural Network

عنوان فارسی	عنوان انگلیسی
شبکه‌های از پیش آموزش دیده بدون ناظر	Unsupervise Pre-trained Networks
شبکه‌های عصبی بازگشتی	Recursive Neural Network
شبکه‌های عصبی کانولوشن	Convolutional Neural Network (CNN)
شبکه‌های عصبی مکرر	Recurrent Neural Network (RNN)
طیف‌سنجی نور نزدیک به مادون قرمز	Functional Near-Infrared Spectroscopy
عروق کرونر	Coronary Artery
فیلتر الگوهای مکانی رایج بانک فیلتر	Filter-bank common spatial patterns filter
قدم‌ها	Stride
لاپلاس	Laplace
لایه تماما متصل	Dense
ماشین بردار پشتیبان	Support Vector Machine (SVM)
محاسبه ذهنی	Mental Calculation
مرجع متوسط مشترک	Common Average Reference (CAR)
مرجع‌دهی	Referencing
مرکزی	Central
مگنتوانسفالوگرافی	Magnetoencephalography
منطقه مطلوب	Region Of Inter Interest
میانگین سیگنال	Signal Average (SA)
ناسازگاری مرتبط با رخداد	Event-related desynchronization (ERD)
نرخ سیگنال به نویز	Singal-to-Noise Rate (SNR)
نرمال‌سازی دسته‌ای	Batch Normalization
نزدیک مادون قرمز	Near Infrared
نورون‌های فعال	Firing

عنوان فارسی	عنوان انگلیسی
همودینامیکی	Hemodynamically
هموگلوبین با اکسیژن	Oxygenated-hemoglobin
هموگلوبین فاقد اکسیژن	Deoxygenated-hemoglobin
وابسته به سطح اکسیژن خون	blood-oxygen-level dependent
یادگیری عمیق	Deep Learning (DL)
یک درمقابل یک	One-Versus-One

بر اساس حروف الفبای انگلیسی

عنوان فارسی	عنوان انگلیسی
اسکلروز جانبی آمیوتروفیک	Amyotrophic Lateral Sclerosis (ALS)
توان باند	Band Power (BP)
نرمال سازی دسته ای	Batch Normalization
وابسته به سطح اکسیژن خون	blood-oxygen-level dependent
رابط مغز- کامپیوتر	Brain-Computer Interface (BCI)
باترورث	Butterworth
مرکزی	Central
توابع شناختی	Cognitive Function
مرجع متوسط مشترک	Common Average Reference (CAR)
تصاویر توموگرافی تخمینی	Computed Tomography Images
شبکه عصبی کانولوشن	Convolutional Neural Network
شبکه های عصبی کانولوشن	Convolutional Neural Network (CNN)
عروق کرونر	Coronary Artery
یادگیری عمیق	Deep Learning (DL)
لایه تماما متصل	Dense
هموگلوبین فاقد اکسیژن	Deoxygenated-hemoglobin
الکتروانسفالوگرافی	Electroencephalography
حرکات عضلات	Electromyography (EMG)
حرکت چشم	Electrooculography (EOG)
الکتروفیزیولوژیکی	Electrophysiological
التهاب مغز یا آنسفالیت	Encephalitis
تجزیه حالت کلی گروهی	Ensemble Empirical Mode Decomposition

عنوان فارسی	عنوان انگلیسی
ناسازگاری مرتبط با رخداد	Event-related desynchronization (ERD)
پتانسیل مرتبط با رخداد	Event-Related Potential (ERP)
سازگاری مرتبط با رخداد	Event-related synchronization (ERS)
فیلتر الگوهای مکانی رایج بانک فیلتر	Filter-bank common spatial patterns filter
نورون‌های فعال	Firing
پیشانی	Forehead
پیشانی	Frontal
تصویربرداری رزونانس مناطیسی تابعی	Functional Magnetic Resonance Imaging
طیف‌سنجی نور نزدیک به مادون قرمز	Functional Near-Infrared Spectroscopy
تصاویر رَحِم	Fundus Images
همودینامیکی	Hemodynamically
تصاویر هیستوپاتولوژی	Histopathologic Images
تجزیه و تحلیل اجزای مستقل	Independent Component Analysis (ICA)
امواج صرعی بین حمله‌ای یا صرع	Interictal Epileptic Discharges (IED)
لاپلاس	Laplace
تجزیه و تحلیل تفکیک خطی	Linear Discriminant Analysis (LDA)
حافظه کوتاه و بلند مدت	Long-Short Term Memory
تصاویرسازی تشدید مغناطیسی	Magnetic Resonance Imaging (MRI)
مگنتوانسفالوگرافی	Magnetoencephalography
محاسبه ذهنی	Mental Calculation
تکانه	Momentum
تصویرات حرکتی	Motor Imagery (MI)
بیماری نورون‌های حرکتی	Motor Neuron Disease (MND)
نزدیک مادون قرمز	Near Infrared

عنوان فارسی	عنوان انگلیسی
اسکن مغزی	Neuroimaging
یک درمقابل یک	One-Versus-One
بیش‌برازش	Overfitting
هموگلوبین با اکسیژن	Oxygenated-hemoglobin
جداری	Parietal
چگالی طیف توان	Power Spectral Desity (PSD)
جلویی	Prefrontal
تجزیه و تحلیل اجزای اصلی	Principal Component Analysis (PCA)
شبکه‌های عصبی مکرر	Recurrent Neural Network (RNN)
شبکه‌های عصبی بازگشتی	Recursive Neural Network
مرجع‌دهی	Referencing
منطقه مطلوب	Region Of Inter Interest
تنظیم مدل	Regularization
دوره استراحت	Relaxation
ریتم‌های حسی-حرکتی	Sensory-Motor Rhythms (SMR)
تجزیه و تحلیل تفکیک خطی انقباضی	shrinkage Linear Discriminant Analysis
میانگین سیگنال	Signal Average (SA)
نرخ سیگنال به نویز	Singal-to-Noise Rate (SNR)
پتانسیل‌های قشری آهسته	Slow Cortical Potentials (SCP)
پتانسیل‌های رخداد دیداری حالت ایستا	Steady State Visual Evoked Potentials (SSVEP)
قدم‌ها	Stride
ماشین بردار پشتیبان	Support Vector Machine (SVM)
زمان-فرکانس	Time-Frequency

عنوان فارسی	عنوان انگلیسی
شبکه‌های از پیش آموزش دیده بدون ناظر	Unsupervised Pre-trained Networks
دیداری	Visual
تبدیل موجک	Wavelet Transform
تشکیل کلمه	Word Formation

Abstract

The Brain Computer Interface (BCI) is a powerful system for communication between the brain and the outside world. Traditional BCI systems only worked on EEG. These signals, because of their benefits, facilitate the analysis and recording of mental activity. Recently, researchers have combined EEG signals to improve the performance of these systems with other signals. Among these signals, the combination of EEG and fNIRS signals has achieved favorable results. In most studies, only EEG or fNIRS signals are used as sequences of chain-like signals, and the complex correlation between adjacent signals, the time and position of the electrodes, is not considered. On the other hand, there are challenges that need to be addressed. The proposed method of this thesis is focused on two parts of its preprocessing for data preparation and classification. The preprocessing combines the spatial relationship between the EEG and fNIRS signals. This is done by using the position of the electrodes on the scalp. As a result, sequences of these chain-like signals can be converted into three-dimensional tensors. In this thesis, a deep neural network based on two neural networks CNN and RNN has been proposed for accurate diagnosis of human mental goals using time and spatial features. The results of the proposed preprocessing method showed that the raw signals become images with more information. These images have spatial information of the signals at any moment in time. It was shown that the proposed model has a precision of 99.6% and outperform a set of baseline methods and most recent deep learning-based EEG or EEG-fNIRS recognition models, yielding a significant accuracy increase of 1 percent in the cross-subject validation scenario. On the other hand, the results showed that the proposed preprocessing and the conversion of raw signals to images with spatial information had a favorable effect on the deep neural network.

Keywords: EEG, fNIRS, Spatial feature, Temporal feature, Deep learning



Faculty of Computer Engineering

M.Sc. Thesis in Artificial Intelligence Engineering

Deep Learning for Classification of Simultaneous EEG-fNIRS Data

By: Hamidreza Ghonchi

Supervisors

Dr. Mansoor Fateh

Dr. Vahid Abolghasemi

Advisors:

Dr. Mohsen Rezvani

Dr. Saideh Ferdowsi

June 2019