



Computational Intelligence in Electrical Engineering
Vol. 13, No. 1, 2022
Research Paper

Recognition of Emotion Provoked by Auditory Stimuli using EEG Signal Based on Deep Neural Networks

Sobhan Sheykhivand¹, Zohreh Mousavi², Tohid Yousefi Rezaii^{*3}, Sholeh Alaei⁴

¹ Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran

² Department of Mechanical Engineering, University of Tabriz, Tabriz, Iran

³ Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran

⁴ Department of Physics, Urmia Branch, Islamic Azad University, Urmia, Iran

Abstract:

Excitements are important for the proper interpretation of actions as well as relationships among individuals. Recognizing emotions through Electroencephalogram (EEG) allows recognition of emotional states without traditional methods including filling in the questionnaire. The automatic emotion recognition reflects the excitement of the individual without clinical examinations or need to visits, which plays a very important role in completing the Brain-Computer Interface (BCI) puzzle. One of the major challenges in this regard is first to select and extract the proper characteristics/features of the EEG signal in order to create an acceptable distinction between different emotional states. The process of finding the desirable feature is generally time consuming. This study presents a new approach for the automatic identification of 3-states of emotion (positive, negative and neutral) based on the auditory stimulation of EEG signals. In the proposed method, the raw EEG signal is directly applied to convolutional neural network-long short time memory (CNN-LSTM) network, without involving the extraction/selection feature. This has been a challenging process in previous literature. The proposed network architecture includes 10 convolutional layers with 3 LSTM layers followed by 2 fully connected layers. The simulation results of the proposed algorithm for classifying 2-stages (negative and positive) and 3-stages (negative, neutral and positive) of emotion for 12 active channels show the accuracy of 97.42% and 95.23% and Cohen's Kappa coefficient of 0.96 and 0.93 respectively.

Keywords: EEG, Emotion, Auditory Stimuli, Convulsion Neural Network (CNN).



This is an open access article under the CC BY-NC-ND/4.0/ License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).



<http://dx.doi.org/10.22108/isee.2020.122304.1357>

تشخیص هیجانات القاشده با تحریک شنوایی از سیگنال‌های EEG بر اساس شبکه‌های

یادگیری عمیق

سبحان شیخی‌وند^۱، زهره موسوی^۲، توحید یوسفی رضایی^{۳*}، شعله اعلایی^۴

۱- دانشجوی دکتری تخصصی، دانشکده مهندسی برق و کامپیوتر - دانشگاه تبریز - تبریز - ایران

s.sheikhvand@tabrizu.ac.ir

۲- دکتری تخصصی، دانشکده مهندسی مکانیک - دانشگاه تبریز - تبریز - ایران

zohreh_mousavi@tabrizu.ac.ir

۳- دانشیار، دانشکده مهندسی برق و کامپیوتر - دانشگاه تبریز - تبریز - ایران

yousefi@tabrizu.ac.ir

۴- استادیار، گروه فیزیک، واحد ارومیه - دانشگاه آزاد اسلامی - ارومیه - ایران

sh.alaei@iaurmia.ac.ir

چکیده: هیجانات، برای تفسیر درست اقدامات و نیز ارتباطات بین انسان‌ها مهم‌اند. شناخت هیجانات از طریق سیگنال الکتروانسفالوگرام (EEG)، امکان تشخیص حالات هیجانی را بدون روش‌های سنتی از جمله پرسشنامه، میسر و بدون معاینات و ویژگی‌های بالینی، هیجان مدنظر را در فرد بازگو می‌کند که نقش بسیار مهمی در تکمیل کردن پازل تعامل بین مغز و کامپیوتر (BCI) دارد. یکی از مشکلات تشخیص خودکار هیجانات، استخراج ویژگی‌های مطلوب است؛ به گونه‌ای که این ویژگی‌ها بیشترین تمایز را بین مراحل مختلف هیجانی ایجاد می‌کنند. فرآیند یافتن ویژگی‌های مناسب، عموماً امری زمان‌بر است. در این پژوهش، رویکرد جدیدی برای شناسایی خودکار ۳ حالت مثبت، منفی و خنثی از هیجانات مبتنی بر تحریک شنوایی از سیگنال‌های EEG ارائه شده است. در روش پیشنهادی، مستقیماً سیگنال EEG ثبت شده، ورودی شبکه عمیق کانولوشنال و شبکه حافظه طولانی کوتاه‌مدت (CNN-LSTM) در نظر گرفته می‌شود؛ بدون اینکه از استخراج/انتخاب ویژگی کلاسیک استفاده شود. موارد گفته شده به‌عنوان روند چالش‌برانگیز در ادبیات قبلی مطرح شده است. معماری شبکه پیشنهادی به‌صورت ۱۰ لایه کانولوشن با ۳ لایه LSTM و به دنبال آن، ۲ لایه کاملاً متصل طراحی شده است. نتایج شبیه‌سازی روش پیشنهادی برای طبقه‌بندی ۲ حالت و ۳ حالت از هیجانات به‌ترتیب صحت ۹۷/۴۲٪ و ۹۵/۲۳٪ و ضریب کاپاکوهن ۰/۹۶ و ۰/۹۳ را ارائه می‌دهند. علاوه بر این، مقایسه نتایج حاصل شده با روش‌های رایج، عملکرد مطلوب روش پیشنهادی را نشان می‌دهد.

واژه‌های کلیدی: EEG، هیجان، تحریک موسیقیایی، CNN

۱- مقدمه

تاریخچه طولانی تحقیق در زمینه ماهیت هیجانات، تا کنون توافق عامی راجع به اینکه هیجانات چیست و چگونه می‌توان آن را نشان داد، حاصل نشده است. پُل کلینجینا تعریف نسبتاً جامعی برای هیجانات ارائه کرده است. او هیجان را حاصل تعامل عوامل ذهنی، محیطی و فرآیندهای عصبی و هورمونی بدن می‌داند. هیجان، حالت فیزیولوژیکی ذهنی و مرتبط با مغز است که با طیف گسترده‌ای از احساسات، رفتار و افکار مرتبط است. پس درواقع احساس

هیجان‌ها نقش مهمی در زندگی انسان دارند. باوجود

^۱ تاریخ ارسال مقاله: ۱۳۹۹/۰۱/۰۹

تاریخ پذیرش مقاله: ۱۳۹۹/۱۰/۰۹

نام نویسنده مسئول: توحید یوسفی رضایی

نشانی نویسنده مسئول: ایران - تبریز - دانشگاه تبریز - دانشکده مهندسی برق و کامپیوتر

در مجموع، مدت ثبت سیگنال ۲۰ دقیقه بوده و سه هیجان خوش‌آیندی^۶، حالت خنثی^۷ و ناخوشایندی^۸ در آنها بررسی شده است. همچنین در این پژوهش، دو دسته ویژگی با یکدیگر مقایسه شده‌اند که با تبدیل فوریه سریع^۹ و استخراج تعدادی ویژگی‌های آماری با به‌کارگیری ماشین بردار پشتیبان^{۱۰} (SVM) به‌عنوان طبقه‌بند برای هر دو حالت، نرخ بازشناسی صحیح ۶۶٪ گزارش شده است. در [۵]، دو کلاس هیجانی آرامش^{۱۱} در حالتی آرام و با چشمانی باز آزمایش شده‌اند. در این پژوهش از بخش‌بندی داده و رگرسیون خطی^{۱۲} برای استخراج ویژگی استفاده شده است. این آزمایش روی ۴۳ نفر انجام شده است. سپس با استفاده از عملیات نرمال‌سازی سیگنال و استفاده از خوشه‌بندی‌های فازی^{۱۳} در باندهای فرکانسی مختلف، نتایج کمی به دست آمده‌اند. در [۶]، استخراج ویژگی‌ها برای شناخت هیجان‌ات از طریق سیگنال‌های EEG ۶۴ کاناله با یکدیگر، مقایسه و ویژگی‌های مهم در حوزه شناخت با استفاده از تکنیک‌های یادگیری ماشین، انتخاب شده‌اند. در این مقاله، ویژگی‌هایی در هر یک از سه حوزه زمان، فرکانس و زمان - فرکانس روی داده‌هایی متشکل از ۷ زن و ۹ مرد استخراج شده‌اند. داده‌ها شامل ۵ بار هیجانی (شادی، کنجکاوی، عصبانیت، ناراحتی و آرامش) بوده‌اند و به کمک تصاویر IAPS^{۱۴} در ۸ قسمت ۳۰ ثانیه‌ای برای هر کلاس هیجانی، استفاده و با دو بعد آزمون خودارزیابی آدمک^{۱۵} (SAM)، صحت القایشان با محک مدل سه‌بعدی هیجان‌ات اعتبار سنجی شده‌اند. در مرحله پردازش سیگنال، داده مربوط به ۵ شرکت‌کننده به علت کیفیت پایین سیگنال ضبط‌شده کنار گذاشته شده و میانگین ۱۱ فرد با ۶ استخراج ویژگی مختلف و استفاده از QDA^{۱۶} به‌عنوان طبقه‌بند صحت بازمیابی هیجان‌ات بین ۳۴٪ تا ۳۶٪ گزارش شده است. در [۷]، از شبکه یادگیری عمیق DLN^{۱۷} در بازشناسی هیجان‌ات از طریق سیگنال EEG بهره گرفته شده است. در این پژوهش، از ۳۲ نفر ثبت سیگنال با ۳۲ کانال انجام شده است. همچنین، برای استخراج ویژگی‌های مهم، الگوریتم PCA^{۱۸} اعمال شده است. نکته شایان توجه در این پژوهش، استخراج ویژگی و طبقه‌بندی به‌صورت جداگانه برای هر کدام از ابعاد انگیزتی و ظرفیت

به‌نوعی زیرمجموعه‌ای از هیجان محسوب می‌شود و در این پژوهش، بازشناسی هیجان‌ات مدنظر است [۱]. هیجان‌ات به‌نوعی در میان مسائلی مانند تجارب روزانه هر انسان، ادراک و انجام وظایف روزانه مانند آموزش، ارتباطات و حتی تصمیم‌گیری‌های روزانه شکل می‌گیرد. بازشناسی هیجان‌ات نقش مهمی در زندگی بشری داشته است و بیشتر تحقیقات سنتی و قدیمی در این زمینه از پارامترهای فیزیکی مثل حالت‌های صورت و حرکات بدن استفاده کرده‌اند. با گذشت زمان و پیشرفت علم و فناوری، شرایط برای اخذ اطلاعات به‌صورت مستقیم از مغز مهیا شده است. برای گرفتن اطلاعات از مغز گزینه‌های متعددی از جمله تصویربرداری تشدید مغناطیسی عملکردی (fMRI)، الکتروانسفالوگرافی (EEG) و طیف‌سنجی مادون‌قرمز نزدیک (NIRS) وجود دارد. در این میان، استفاده از سیگنال EEG به دلیل مزایایش نسبت به سایر روش‌های اخذ سیگنال، متداول‌تر است. EEG دارای دو مزیت عمده رزولوشن زمانی بالا، حمل‌شدنی و کم‌هزینه بودن ابعاد تجهیزات اندازه‌گیری است [۲]. شناخت هیجان‌ات از طریق سیگنال EEG، امکان تشخیص حالات هیجانی را بدون روش‌های سنتی از جمله پرسشنامه، میسر و بدون معاینات و ویژگی‌های بالینی، هیجان مدنظر را در فرد بازگو می‌کند که نقش بسیار مهمی در تکمیل کردن پازل تعامل بین مغز و کامپیوتر^{۱۹} (BCI) دارد [۲]. روش‌های متفاوتی برای القای هیجان‌ات در انسان وجود دارد. تماشای فیلم‌های احساسی، تماشای تصاویر احساسی، تصویرسازی‌های ذهنی و موسیقی‌های احساسی برخی از روش‌های القای هیجان‌ات اند [۳].

در شناخت هیجان‌ات از طریق EEG، عموماً بر مناسب بودن ویژگی‌ها موافقت زیادی وجود ندارد و در پژوهش‌های اندکی تعداد کمی ویژگی‌های مختلف باهم مقایسه شده‌اند؛ به‌طور مثال، در مطالعات انجام‌شده در این زمینه، در [۴] پنج نفر با القای تحریک‌های هیجانی از طریق تصویر آزمایش شده‌اند. تصاویر هر کدام به مدت ۶ ثانیه از فاصله ۵ متری برای افراد تحت آزمایش با یک نمایشگر نشان داده شده‌اند. با احتساب ۱۵ ثانیه زمان بین القای حالات هیجانی و ۲ ثانیه زمان بین پخش تصاویر،

استفاده کردند. این پژوهشگران از تحریک موسیقی روی ۱۵ نفر شرکت کننده در آزمایش خود استفاده کردند. آنها همچنین مدل پیشنهادی خود را با چند الگوریتم رایج مقایسه کردند. خطای مدل پیشنهادی آنها برای طبقه بندی ۲ هیجان در حدود ۰/۰۸۹ گزارش شد. پانایو و همکاران [۱۱] از شبکه های عصبی عمیق برای تشخیص دو نوع هیجان با استفاده از سیگنال های EEG استفاده کردند. این پژوهشگران ۱۲ شرکت کننده را در آزمایش خود به کار گرفتند. معماری شبکه پیشنهادی آنها از ۶ لایه کانولوشن تشکیل شده بود. آنها الگوریتم پیشنهادی خود را با SVM مقایسه کردند و به این نتیجه رسیدند که مدل پیشنهادی آنها عملکرد بهتری در تشخیص احساسات دارد. یانگ و همکاران [۱۲] از یک شبکه عصبی بازگشتی برای شناسایی خودکار هیجان از سیگنال های EEG استفاده کردند. این محققان در آزمایشات خود از مجموعه داده های مبتنی بر تحریک ویدیویی بهره گرفتند. همچنین، آنها در روش پیشنهادی خود سیگنال های تک بعدی EEG را به فریم های دوبعدی به منظور آموزش شبکه تبدیل کردند. صحت گزارش شده آنها برای هر دو کلاس ظرفیت و هیجان به ترتیب ۹۰٪ و ۹۱٪ گزارش شده است. چن و همکاران [۱۳] برای طبقه بندی خودکار هیجان دو کلاس از سیگنال های EEG استفاده کردند. این پژوهشگران، از شبکه های عصبی تکرار شونده موازی در الگوریتم پیشنهادی خود به کار گرفتند. صحت نهایی گزارش شده برای طبقه بندی کلاس ظرفیت و برانگیختگی براساس الگوریتم پیشنهادی آنها به ترتیب ۹۳/۶۴٪ و ۹۳/۲۶٪ گزارش شده است. وی و همکاران [۱۴] از تبدیل موجک دوگانه برای استخراج ویژگی های مطلوب از سیگنال های EEG به منظور بازشناسی هیجان استفاده کردند. همچنین این پژوهشگران پس از استخراج ویژگی های مطلوب، از واحدهای بازگشتی به منظور آموزش مدل خود بهره گرفتند. در نهایت آنها به ترتیب به صحت ۸۵٪، ۸۴٪ و ۸۷٪ برای کلاس های هیجان مثبت، منفی و خنثی دست یافتند.

چالش عمده در الگوریتم های بازشناسی هیجان از سیگنال های EEG، انتخاب ویژگی متمایز کننده مراحل

است. نتایج حاصل با اعمال طبقه بند SVM و طبقه بند Bayes ارزیابی شده اند. طبق گزارش ها پس از طبقه بندی مشخص شده که شبکه DLN بهتر از SVM عمل کرده است. در این پژوهش، بخش استخراج ویژگی به صورت ثابت در نظر گرفته شده و در بخش طبقه بندی چهار الگوریتم مختلف و ترکیبی استفاده شده است. برای مشخص شدن میزان کارایی روش پیشنهادی، از ۳ معیار کمی و معمول استفاده شده است. به همین منظور، برای ارزیابی عملکرد روش پیشنهادی، معیارهای شاخص حساسیت، صحت و اختصاصیت محاسبه شده اند. زمانی که بتوان داده ها را به دو گروه مثبت و منفی تقسیم کرد، حساسیت به معنی نسبتی از موارد مثبت است که آزمایش آنها را به درستی، مثبت علامت گذاری می کند. اختصاصیت به معنی نسبتی از موارد منفی است که آزمایش آنها را به درستی، منفی علامت گذاری می کند و صحت به معنی تعداد تشخیص های درست دو کلاس نسبت به تمام تشخیص های درست و نادرست دو کلاس است. نتایج طبقه بندی این پژوهش در تشخیص صحیح بعد ظرفیت ۵۳/۴۲٪ با انحراف معیار ۶۴٪ و تشخیص صحیح بعد انگیزتگی ۵۲٪ با انحراف معیار ۷۵٪ گزارش شده اند. در [۸]، سه دسته ویژگی در بازشناسی هیجان مطابق با تحریک ویدیویی براساس دیتاست ثبت شده اختصاصی با یکدیگر مقایسه شده اند. سپس ماتریس ویژگی حاصل، با SVM طبقه بندی شده است. همچنین، فرکانس نمونه برداری در این پژوهش ۱۰۰ هرتز بوده است. نتایج نهایی این مقاله با توجه به تنوع ویژگی های استفاده شده و اعمال یک مرحله هموارسازی ویژگی برای ۶ فرد با صحت ۸۷٪ گزارش شده اند. بیمین و همکاران [۹] از یک مدل مبتنی بر همبستگی برای استخراج ویژگی ها از سیگنال های EEG برای طبقه بندی هیجان مختلف (آرامش، شادی، غم و اندوه) روی ۸ شرکت کننده استفاده کردند. آنها از طبقه بندهای BP، SVM، LDA و C4.5 استفاده کردند و نتیجه گرفتند طبقه بند C4.5 برای تشخیص احساسات، بهتر از سایر طبقه بندها عمل می کند. فاطمه و همکاران [۱۰] از یک آبشار موازی فازی برای پیش بینی محتوای عاطفی هیجان از سیگنال های EEG

مختلف هیجانی از یکدیگر است. در بیشتر مطالعات پیشین برای تشخیص خودکار مراحل مختلف هیجانی از روش‌های مرسوم آماری و پردازشی، ویژگی‌های استخراج‌شده و سپس با روش‌های کاهش ویژگی، ویژگی‌های مطلوب و بهینه انتخاب می‌شدند. استخراج این ویژگی‌ها به صورت دستی در طبقه‌بندی مراحل مختلف هیجان‌ات، باعث پیچیدگی بار محاسباتی الگوریتم خواهد شد. به علاوه، ویژگی‌های مطلوب و بهینه در یک مسئله ممکن است برای مسئله‌ای دیگر، ویژگی بهینه‌ای محسوب نشود؛ بنابراین، استفاده از روشی که بتواند ویژگی‌های مناسب را بسته به نوع مسئله و داده یاد بگیرد، امری ضروری است. این مسئله، نکته کلیدی این پژوهش است. در این مقاله، یک الگوریتم تماماً خودکار دسته‌بندی (بدون نیاز به انتخاب و استخراج ویژگی به صورت دستی) برای تشخیص ۳ مرحله هیجان (مثبت، منفی و خنثی)، با صحت و سرعت پیش‌بینی بالا با استفاده از سیگنال‌های EEG بر مبنای یادگیری عمیق و شبکه CNN طراحی شده است تا سیگنال خام اولیه را به منظور یادگیری ویژگی‌ها و شناسایی خودکار مراحل مختلف هیجان‌ات پردازش کند.

ادامه مقاله به صورت زیر تدوین شده است. در بخش ۲، داده‌های آزمایشی ثبت شده با استفاده از تحریک موسیقیایی بررسی می‌شوند و شبکه عصبی کانولوشنال به همراه شبکه حافظه طولانی کوتاه مدت توضیح داده می‌شود. در بخش ۳، معماری شبکه روش پیشنهادی براساس (CNN-LSTM) ارائه می‌شود. در بخش ۴، نتایج شبیه‌سازی بررسی می‌شوند. در نهایت، بخش ۵ مربوط به نتیجه‌گیری است.

۲- مواد و روش‌ها

در این بخش، ابتدا مجموعه داده‌های ثبت شده EEG توضیح داده می‌شود، سپس شبکه‌های عصبی عمیق بررسی می‌شوند.

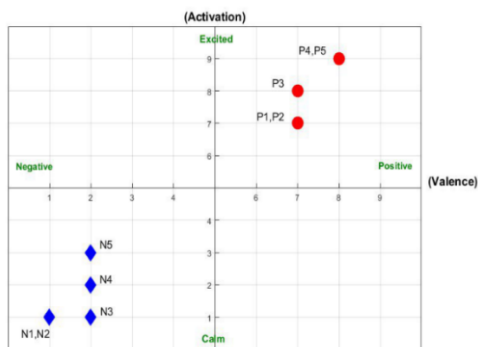
۲-۱- ثبت سیگنال EEG

برای بازشناسی هیجان‌ات از روی سیگنال EEG به

ایجاد یک پایگاه برای ۳ هیجان مثبت، منفی و خنثی اقدام شده است. برای ارزیابی ۲ هیجان مثبت و منفی از آزمون ارزیابی SAM^{۱۹*} در نسخه کاغذی و ۹ درجه‌ای در فرآیند آزمایش استفاده شد. در این آزمون، نمره پایین‌تر از ۳، پایین و نمره بالاتر از ۶ بالا محسوب می‌شود [۶]. پیش از ثبت سیگنال از تمام شرکت‌کنندگان خواسته شد تا فرم رضایت‌نامه (نداشتن سابقه بیماری روحی و روانی، نداشتن بیماری صرع، استفاده نکردن از داروهای روان‌پزشکی، داشتن خواب کافی قبل از آزمون، مصرف نکردن غذاهای چرب و مواد کافئین قبل از آزمایش و شست‌وشوی موها قبل از آزمون) را مطالعه و در صورت تمایل به شرکت در آزمون آن را امضا کنند. سپس از شرکت‌کنندگان خواسته شد تا پرسشنامه خلق‌آنی افسردگی بک^{۲۰} را تکمیل کنند. پس از تکمیل پرسشنامه، طبق استانداردهای روان‌شناسی، آن دسته از شرکت‌کنندگانی که نمره بالاتر از ۲۱ در این آزمون به دست آورده‌اند، از فرآیند پردازش و نتیجه‌گیری کنار گذاشته شده‌اند. برای ثبت سیگنال از ۱۶ نفر (۶ زن و ۱۰ مرد) در بازه سنی ۲۰ تا ۲۸ سال دعوت به همکاری شد. سیگنال EEG افراد در حال گوش دادن به موسیقی ثبت شد. کلیه ثبت‌های انجام شده در نور و دمای کنترل شده محیط (۲۴ درجه سانتی‌گراد) در ساعات ۹ صبح الی ۱۴ ظهر انجام شدند تا افراد احساس خستگی نداشته باشند. یک صندلی راحتی برای نشستن افراد به منظور القای بهتر و جلوگیری از نویزهای حرکتی تهیه شد. همچنین برای جلوگیری از نویز EOG^{۲۱} از تمامی افراد خواسته شد تا در فرآیند ثبت سیگنال، چشمانشان را بسته نگه دارند. برای ثبت سیگنال از دستگاه ۲۱ کاناله Encephalan شرکت مدیکام روسیه (Medicom) استفاده شده است. درباره آرایش الکترودها روی سر، از سیستم استاندارد بین‌المللی ۱۰-۲۰ استفاده شده است. فرکانس نمونه‌برداری در هنگام ثبت ۲۵۰ هرتز با تطبیق امپدانس ۱۰ کیلو اهم بوده است. تعداد الکترودهای دستگاه ثبت سیگنال آزمایشگاه ۲۱ الکتروود بود که در مرحله پردازش با نرم‌افزار متلب طبق استاندارد، عملاً ۱۹ کانال در اختیار قرار گرفته است. به دلیل راحتی و تنظیم صحیح الکترودها روی سر، از کلاه مخصوص آن استفاده شده



شکل (۱): ثبت سیگنال EEG از یکی از شرکت کنندگان.



شکل (۲): نتیجه اعتبارسنجی SAM درباره میزان القای مؤثر محرک هیجانی (۱۰ قطعه موسیقی) برای سوژه اول.

منظور از بهینه‌سازی مصرف انرژی انتخاب الگوها، اتخاذ و به‌کارگیری روش‌ها و سیاست‌هایی در مصرف انرژی الکتریکی است. ساختمان‌های مسکونی بخش مهمی از مصرف‌کنندگان انرژی الکتریکی به شمار می‌آیند. ورود تکنولوژی سیستم مدیریت هوشمند به ساختمان‌های مسکونی، تا حدودی مصرف انرژی الکتریکی را بهینه کرده است.

است. در شکل ۱ نمایی از ثبت سیگنال یکی از شرکت‌کنندگان نشان داده شده است. شرح نتایج آزمون افسردگی بک و خودارزیاب مانکن، در جدول ۱ نمایش داده شده است؛ برای مثال، طبق جدول ۱، فرد شماره ۳ به دلیل تطابق‌نداشتن پاسخ به سؤال کنترلی در آزمون SAM (میانگین بعد انگیزندگی القایی بالاتر از ۳ است) از روند پردازش کنار گذاشته شده است. جزئیات نتیجه اعتبارسنجی القای هیجان با آزمون خودارزیاب مانکن، برای فرد شماره ۱ در شکل ۲ نمایش داده شده‌اند. با توجه به شکل ۲ و جدول ۱، فرد شماره ۱ با میانگین بعد ظرفیت القایی هیجان مثبت ۹ (بزرگ‌تر از ۶) و میانگین بعد انگیزندگی القایی هیجان منفی (کوچک‌تر از ۳) و نمره افسردگی بک ۱۶ (۲۱ > ۱۶) وارد روند پردازش شده است.

برای تحریک هیجان مثبت و منفی در افراد از تحریک موسیقایی استفاده شد. هر قطعه موسیقی به مدت یک دقیقه پخش شده و بین هر دو قطعه پانزده ثانیه سکوت (حالت خشی) برای جلوگیری از انتقال جانب دارانه هیجان در نظر گرفته شده است. درخور ذکر است برای القای بهتر از پخش با بلندگو اجتناب شده و از هدفون با توجه به نویز کمتر برای پخش موسیقی استفاده شده است.

تم و حالت موسیقی، تأثیر عمومی و فیزیولوژیک دارد و بر هر فردی با مکانیسم ذهنی و عاطفی مختلف تأثیر می‌گذارد؛ اما اندازه و شدت این تأثیر به وضعیت سلول‌های عصبی، سابقه ذهنی و عادت شنونده بستگی دارد. برای تحریک موسیقایی در افراد آزمایش‌شده، برای دو هیجان مثبت و منفی، طبق [۱۵] تم حزین برای القای هیجان منفی و تم هیجانی (هیستریکال) برای القای هیجان مثبت استفاده شد. جدول ۲ جزئیات موسیقی‌های انتخابی را برای هر تم و شکل ۳ ترتیب و نحوه پخش محرک موسیقایی را برای افراد شرکت‌کننده نشان می‌دهد (مطابق با شکل ۳، کلمات اختصاری N و P به ترتیب به هیجان منفی و هیجان مثبت مربوط است).

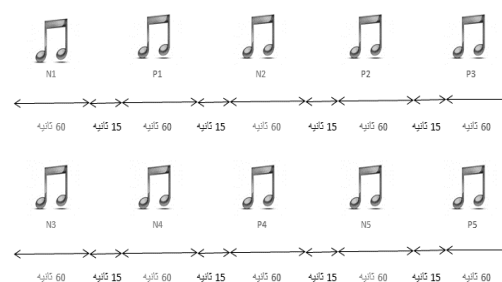
جدول (۱): اعتبارسنجی افراد شرکت‌کننده در فرآیند ثبت سیگنال EEG

شماره سوژه	جنسیت	سن	نمره افسردگی بک	میانگین بعد ظرفیت القایی هیجان مثبت	میانگین بعد انگیزندگی القایی هیجان مثبت	میانگین بعد ظرفیت القایی هیجان منفی	میانگین بعد انگیزندگی القایی هیجان منفی	نتیجه اعتبارسنجی	علت حذف سوژه از مرحله پردازش
۱	پسر	۲۵	۱۶	۹	۹	۱/۸	۱	✓	-
۲	پسر	۲۴	۲۲	۶/۸	۶/۲	۳/۶	۲	*	افسردگی بک (۲۲>۲۱)
۳	دختر	۲۷	۱۹	۶/۲	۷/۴	۴/۲	۴/۶	*	عدم تطابق پاسخ به سؤال کنترلی در آزمون SAM
۴	پسر	۲۴	۴	۷/۴	۷/۶	۲/۴	۲/۶	✓	-
۵	پسر	۲۴	۰	۵/۸	۵	۴/۴	۵/۶	*	عدم تطابق پاسخ به سؤال کنترلی در آزمون SAM
۶	پسر	۲۸	۱۰	۵/۶	۵/۴	۲	۱/۶	*	عدم القای مورد نظر در کلاس هیجانی مثبت
۷	پسر	۲۸	۱۳	۷/۲	۷/۴	۳/۸	۳/۸	*	عدم القای مدنظر در کلاس هیجانی منفی
۸	پسر	۲۰	۱۹	۷/۸	۷/۴	۲/۸	۳	✓	-
۹	پسر	۲۶	۹	۷/۴	۷	۳/۴	۵/۴	*	عدم القای مدنظر در کلاس هیجانی منفی
۱۰	دختر	۲۳	۹	۶/۸	۶/۶	۳/۸	۳/۲	*	عدم القای مدنظر در کلاس هیجانی منفی
۱۱	دختر	۲۵	۲۲	۷/۸	۸	۴/۵	۳	*	افسردگی بک (۲۲>۲۱)
۱۲	دختر	۲۷	۱	۸/۶	۸/۶	۲	۱/۲	✓	-
۱۳	دختر	۲۹	۹	۶	۶	۲	۱/۲	✓	-
۱۴	پسر	۲۶	۸	۸	۸	۱/۸	۱/۶	✓	-
۱۵	دختر	۲۵	۱۲	-	-	-	-	-	نویز حرکتی و نویز EOG
۱۶	پسر	۲۷	۰	۷/۴	۸	۱/۸	۲	✓	-

بر اساس وزن و بایاس لایه‌های قبلی در ساختار شبکه است.

جدول (۲): آهنگ‌های به کاررفته برای القای هیجان.

آهنگ استفاده شده	معنی	علامت اختصاری
پیش‌درآمد اصفهان اثر محمدرضا لطفی	القای هیجان منفی	N1
شش و هشت آذری	القای هیجان مثبت	P1
پیش‌درآمد همایون اثر فرامرز پایور	القای هیجان منفی	N2
شش و هشت آذری	القای هیجان مثبت	P2
شش و هشت بندری	القای هیجان مثبت	P3
قطعه افشاری اثر سهراب پورناظری	القای هیجان منفی	N3
پیش‌درآمد اصفهان اثر محمدرضا لطفی	القای هیجان منفی	N4
شش و هشت فارسی	القای هیجان مثبت	P4
پیش‌درآمد دشتی اثر کیهان کلهر	القای هیجان منفی	N5
شش و هشت بندری	القای هیجان مثبت	P5



شکل (۳): ترتیب پخش موسیقی برای شرکت‌کنندگان.

۲-۲- شبکه‌های عصبی کانولوشن

شبکه عصبی کانولوشنال، در واقع یک شبکه عصبی بهبودیافته است. در این شبکه، چندین لایه با روشی قدرتمند در کنار هم آموزش می‌بینند [۱۶]. این روش، بسیار کارآمد بوده و یکی از رایج‌ترین روش‌ها در کاربردهای مختلف بینایی ماشین است. همانند شبکه‌های عصبی مصنوعی (ANN)، تصمیم خروجی نهایی مدل CNN

می‌نامند. عملگر کانولوشن به شرح زیر است:

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

که x سیگنال، h فیلتر، N تعداد عناصر در x و y بردار خروجی است.

لایهٔ ادغام: این لایه به کاهش نمونه^{۳۰} هم معروف است و ابعاد نورون‌های خروجی را از لایهٔ کانولوشن کاهش می‌دهد و باعث کاهش محاسبات و نیز جلوگیری از پدیدهٔ بیش‌برازش می‌شود. در این پژوهش، از لایهٔ ادغام بیشینه^{۳۱} استفاده شده است که فقط مقادیر بیشینه در هر نگاشت ویژگی را انتخاب می‌کند و باعث کاهش تعداد نورون‌های خروجی می‌شود.

لایهٔ FC: دارای اتصال کامل به تمام فعال‌سازی‌ها در لایهٔ قبلی است.

لایهٔ حذف تصادفی: از این لایه به‌منظور جلوگیری از پدیدهٔ بیش‌برازش استفاده می‌شود [۱۷]. نحوهٔ کار آن به این صورت است که در هر مرحله از آموزش، هر نورون با احتمالی از شبکه بیرون انداخته می‌شود؛ به طوری که در نهایت یک شبکهٔ کاهش‌داده‌شده باقی می‌ماند.

لایهٔ نرمال‌سازی دسته‌ای: این لایه به‌منظور نرمال‌سازی داده‌ها در داخل شبکه انجام می‌شود [۱۸]. زمانی که محاسبات مختلف روی دادهٔ ورودی اعمال شود، توزیع داده‌ها تغییر پیدا خواهد کرد. این لایه با هدف کاهش تغییر کوواریانس داخلی، سرعت آموزش شبکه را افزایش و باعث تسریع در همگرایی می‌شود. تبدیل لایهٔ نرمال‌سازی دسته‌ای به شرح زیر است:

$$\begin{aligned} \mu_B &= \frac{1}{n} \sum_{i=1}^n y_i^{(t-1)} \\ \sigma_B^2 &= \frac{1}{n} \sum_{i=1}^n (y_i^{(t-1)} - \mu_B)^2 \\ \hat{y}^{(t-1)} &= \frac{y^{(t-1)} - \mu_B}{\sqrt{(\sigma_B^2 + \varepsilon)}} \\ z^{(t)} &= \gamma^{(t)} \hat{y}^{(t-1)} + \beta^{(t)} \end{aligned} \quad (2)$$

که μ_B و σ_B^2 به ترتیب میانگین و واریانس دسته‌اند. ε یک ثابت کوچک برای ثبات عددی، l شماره لایه، بردار ورودی به لایه نرمال‌ساز، $z^{(t)}$ بردار خروجی نرمال

در هر شبکهٔ عصبی کانولوشنال، دو مرحله برای آموزش وجود دارد؛ مرحلهٔ انتشار پیشرو^{۳۲} و مرحلهٔ پس‌انتشار^{۳۳} (BP)[۱۷]. در مرحلهٔ نخست، داده‌های ورودی به شبکه اعمال می‌شوند و این عمل چیزی جز ضرب نقطه‌ای بین ورودی و پارامترهای هر نورون و در نهایت اعمال عملیات کانولوشن در هر لایه نیست و در نهایت، خروجی شبکه محاسبه می‌شود. به‌منظور تنظیم پارامترهای شبکه یا به عبارت دیگر آموزش شبکه، از نتیجهٔ خروجی برای محاسبهٔ میزان خطای شبکه استفاده می‌شود. برای این کار، خروجی شبکه با استفاده از یک تابع خطا^{۳۴} با پاسخ صحیح، مقایسه و میزان خطا محاسبه می‌شود. BP روشی برای محاسبهٔ گرادیان تابع اتلاف نسبت به وزن‌ها است. BP سیگنال‌های خطا را در شبکه در حین آموزش پس می‌زند و باعث به‌روزرسانی وزن‌ها می‌شود. در مرحله بعد، براساس میزان خطای محاسبه‌شده، مرحلهٔ پس‌انتشار آغاز می‌شود. در این مرحله، گرادیان هر پارامتر با توجه به قاعدهٔ زنجیره‌ای محاسبه می‌شود و تمامی پارامترها با توجه به تأثیرشان بر خطای ایجادشده در شبکه، به‌روزرسانی می‌شوند. بعد از به‌روزرسانی پارامترها، مرحلهٔ بعدی انتشار پیشرو آغاز خواهد شد. بعد از تکرار تعداد مناسبی از این مراحل، آموزش شبکه به پایان می‌رسد. در این شبکه، خروجی هر لایه همان ویژگی‌ها هستند که بعد کمتری نسبت به داده اصلی دارند.

به‌طور کلی، یک شبکهٔ کانولوشنال از سه لایهٔ اصلی تشکیل می‌شود که عبارت‌اند از: لایه کانولوشن، لایه ادغام^{۳۵} و لایه تمام متصل^{۳۶} (FC) [۱۶]. برای جلوگیری از فرآیند بیش‌برازش^{۳۷} و بهبود عملکرد شبکه از لایه‌های حذف تصادفی^{۳۸} و نرمال‌سازی دسته‌ای^{۳۹} نیز استفاده می‌شود. همچنین در شبکه‌های عصبی نیاز است پس از هر لایه از تابع فعال‌سازی استفاده شود که در ادامه، این لایه‌ها و توابع به‌اجمال معرفی می‌شوند.

لایهٔ کانولوشن: شامل فیلترهایی (کرنل‌هایی) است که روی سیگنال EEG می‌لغزند. یک کرنل، یک ماتریس است که با سیگنال ورودی EEG کانوالو می‌شود. این لایه عمل کانولوشن را روی سیگنال EEG ورودی با استفاده از کرنل انجام می‌دهد. خروجی کانولوشن را نگاشت ویژگی

$$h_t = \Gamma_o \cdot \tanh(C_t) \quad (5)$$

که در آن Γ_o دروازه خروجی بوده و کنترل‌کننده میزان محتوایی است که از طریق حافظه ارائه می‌شود. دروازه خروجی از طریق رابطه (۶) محاسبه می‌شود:

$$\Gamma_o = \sigma(W_o \cdot [h_{t-1}, X_t] + b_o) \quad (6)$$

که در آن σ تابع فعال‌سازی سیگموئید است. W_o نیز یک ماتریس اوریب است. سلول حافظه C_t نیز با فراموشی نسبی حافظه فعلی و اضافه کردن محتوای حافظه جدید به صورت \hat{C}_t از رابطه (۷) به‌روزرسانی می‌شود که در آن محتوای حافظه جدید از رابطه (۸) به دست می‌آید.

$$C_t = \Gamma_f \cdot C_{t-1} + \Gamma_u \cdot \hat{C}_t \quad (7)$$

$$\hat{C}_t = \tanh(W_c \cdot [h_{t-1}, X_t] + b_c) \quad (8)$$

آن میزان حافظه فعلی که باید فراموش شود با دروازه فراموشی Γ_f کنترل می‌شود و آن مقدار از حافظه جدید که باید به سلول حافظه اضافه شود، با دروازه به‌روزرسانی (دروازه ورودی) Γ_u انجام می‌گیرد. این عمل در رابطه (۹) و (۱۰) نشان داده شده است [۲۰، ۲۱].

$$\Gamma_f = \sigma(W_f \cdot [h_{t-1}, X_t] + b_f) \quad (9)$$

$$\Gamma_u = \sigma(W_u \cdot [h_{t-1}, X_t] + b_u) \quad (10)$$

پژوهشگران در این پژوهش قصد دارند در روش پیشنهادی خود از ترکیب شبکه‌های CNN با شبکه‌های LSTM به‌منظور افزایش پایداری و کاهش نوسانات معماری پیشنهادی استفاده کنند.

۳- روش پیشنهادی

در این بخش، روش پیشنهادی مقاله ارائه می‌شود. شکل ۴ ساختار کلی روش پیشنهادی را نشان می‌دهد.

مربوط به یک نورون، $\gamma^{(l)}$ و $\beta^{(l)}$ به ترتیب پارامترهای مربوط به مقیاس و تغییر نرخ یادگیری‌اند.

تابع فعال‌سازی: پس از هر لایه کانولوشن، یک تابع فعال‌سازی اعمال می‌شود. تابع فعال‌سازی یک عملگر است که خروجی را به مجموعه‌ای از ورودی‌ها نگاشت می‌کند و برای غیرخطی کردن ساختار شبکه استفاده می‌شود [۱۹]. در این پژوهش، از تابع LeakyRelu به‌عنوان فعال‌سازی در هر لایه استفاده شده است و این ویژگی را دارد که غیرخطی بودن و تنگی را به ساختار شبکه اعمال کند؛ بنابراین، در برابر تغییرات جزئی در ورودی مقاوم است. رابطه ۳ تابع LeakyRelu را نشان می‌دهد که در حالت نرمال $\alpha = 0.01$ است.

$$f(x) = \begin{cases} x & x > 0 \\ \alpha x & x \leq 0 \end{cases} \quad (3)$$

تابع سافت مکس^{۳۲}: این تابع توزیع احتمالی کلاس‌های خروجی را محاسبه می‌کند؛ بنابراین، در لایه تمام متصل آخر از تابع سافت مکس برای پیش‌بینی اینکه کدام سیگنال ورودی به هیجان مثبت، منفی و خنثی مربوط است، استفاده می‌شود و رابطه آن به فرم زیر است:

$$p_j = \frac{e^{x_j}}{\sum_k e^{x_k}} \quad \text{for } j = 1, \dots, k \quad (4)$$

که در آن، x ورودی شبکه است و مقادیر خروجی p بین صفر و یک بوده که مجموع آنها برابر با یک است.

۲-۳- شبکه حافظه طولانی کوتاه‌مدت (LSTM)

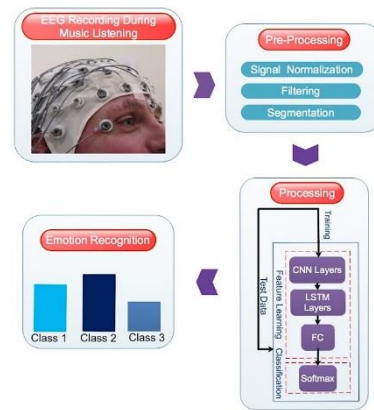
شبکه حافظه طولانی کوتاه‌مدت^{۳۳} (LSTM) از جمله شبکه‌های عصبی بازگشتی^{۳۴} (RNN) است که به‌منظور رفع ضعف‌های شبکه‌های عصبی بازگشتی کلاسیک از جمله حل مشکل پرکنندگی گرادیان به کار برده می‌شوند. برخلاف شبکه عصبی بازگشتی سنتی که صرفاً جمع متوازن سیگنال‌های ورودی را محاسبه می‌کند و سپس از یک تابع فعال‌سازی عبور می‌دهد، هر واحد LSTM از یک حافظه C_t در زمان t بهره می‌برد. فعال‌سازی (خروجی) واحد LSTM به‌صورت رابطه (۵) تعریف می‌شود.

کلاس خواهیم داشت. سپس با تکنیک هم‌پوشانی (برای جلوگیری از پدیدهٔ بیش‌برازش) داده‌های هر کانال به بازه‌های ۸ ثانیه‌ای تقسیم می‌شوند. در واقع هر الکتروود به طول ۷۵۰۰۰ (۳۰۰ ثانیه) با شیفت زمانی ۲۰۰ (۰/۸ ثانیه) با میزان هم‌پوشانی ۹۰٪ به ۳۶۰ نمونه به طول ۲۰۰۰ (۸ ثانیه) تقسیم می‌شود؛ بنابراین، برای نمونه‌های e الکتروود در نظر گرفته شده خواهیم داشت: $(e \times 360 \times 2000)$. چون ۷ فرد و ۲ کلاس داریم، پس در نهایت ابعاد نمونه‌ها و ویژگی‌های اولیه برابر با $(2 \times 7 \times 360) \times (e \times 2000)$

خواهد شد که e برابر است با تعداد الکتروودها، ۷ تعداد افراد و ۲ تعداد کلاس‌ها است. همچنین، برای سناریو دوم ۳ حالت (مثبت، خنثی و منفی) در نظر گرفته می‌شود. بدین ترتیب، ابعاد ماتریس ورودی برای سناریو دوم برابر است با $(3 \times 7 \times 360) \times (e \times 2000)$. این عملیات در شکل ۶ نشان داده شده است.

۳-۲- معماری شبکهٔ عمیق پیشنهادی

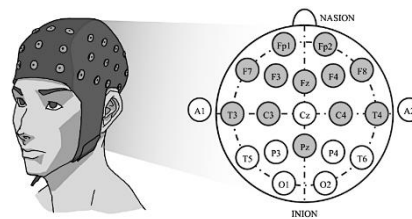
در معماری شبکهٔ پیشنهادی از ترکیب ۱۰ لایه کانولوشن یک‌بعدی و ۳ لایه LSTM استفاده شده است. برای پیاده‌سازی شبکهٔ کانولوشنال از کتابخانهٔ کراس در زبان برنامه‌نویسی پایتون استفاده شده است. معماری شبکهٔ عصبی عمیق پیشنهادی به صورت زیر انتخاب شده است: (۱) یک لایهٔ حذف تصادفی؛ (۲) یک لایهٔ کانولوشن با تابع غیرخطی LeakyRelu و یک لایهٔ ادغام بیشینه همراه با نرمال‌ساز دسته‌ای اضافه می‌شود؛ (۳) معماری مرحلهٔ قبل، ۹ بار دیگر تکرار می‌شود؛ (۴) خروجی معماری قبلی به ۳ لایهٔ شبکهٔ LSTM با توابع غیرخطی LeakyRelu متصل می‌شود که به صورت سری پشت سر یکدیگر قرار گرفته‌اند؛ (۵) خروجی معماری قبلی به یک ماتریس دوبعدی متصل می‌شود؛ (۶) دولایه تماماً متصل برای دسترسی به لایهٔ خروجی استفاده می‌شود. شکل ۷ جزئیات معماری شبکهٔ عمیق پیشنهادی را نشان می‌دهد. جدول ۳ جزئیات معماری شبکهٔ کانولوشنال پیشنهادی به کاررفته را نشان می‌دهد. طبق جدول ۳، کاهش بعد در لایه‌های پنهان از ۲۴۰۰۰ (۲۰۰۰×۱۲) (تعداد ویژگی‌های اولیه) به ۱۰۰ ادامه یافته



شکل (۴): بلوک-دیاگرام الگوریتم پیشنهادی.

۳-۱- پیش‌پردازش داده‌ها

ابتدا دادگان با استفاده از نرمال‌ساز Min-Max بین صفر و یک نرمالیزه شدند، سپس یک فیلتر شکافی^{۳۰} برای حذف فرکانس ۵۰ هرتز برق شهری و یک فیلتر میان‌گذر باترورت مرتبه ۱ با فرکانس پایین ۰/۱۵ و ۴۵ هرتز بر روی دادگان اعمال شد [۱۲-۱۴]. با توجه به این موضوع که یکی از اهداف این پژوهش ارائهٔ الگوریتمی مبتنی بر حداقل کانال‌های سیگنال‌های فیزیولوژیک است، طبق پژوهش‌های [۸-۱۰] فقط از کانال‌های Fp1, Fp2, Fp, F7, F3, Fz, F4, F8, T3, C3, C4, T4 و Pz در شبیه‌سازی استفاده شده است. شکل ۵ کانال‌های انتخاب‌شده برای شبیه‌سازی را نشان می‌دهد.

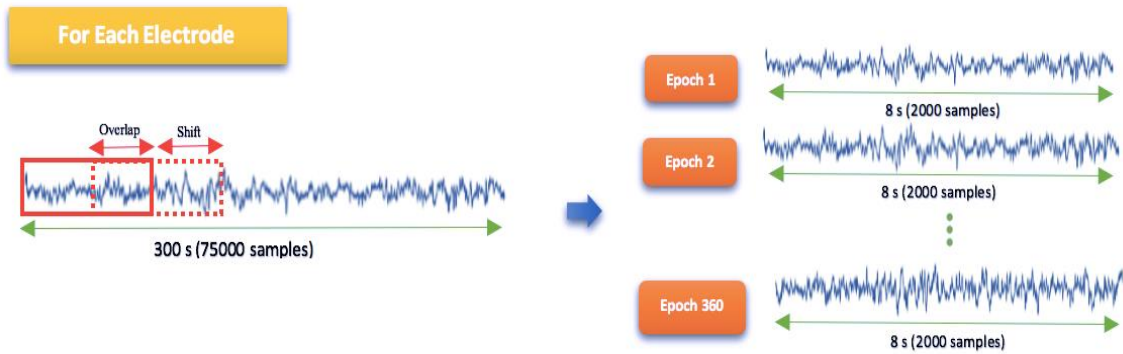


شکل (۵): کانال‌های انتخاب‌شده (خاکستری رنگ) در شبیه‌سازی.

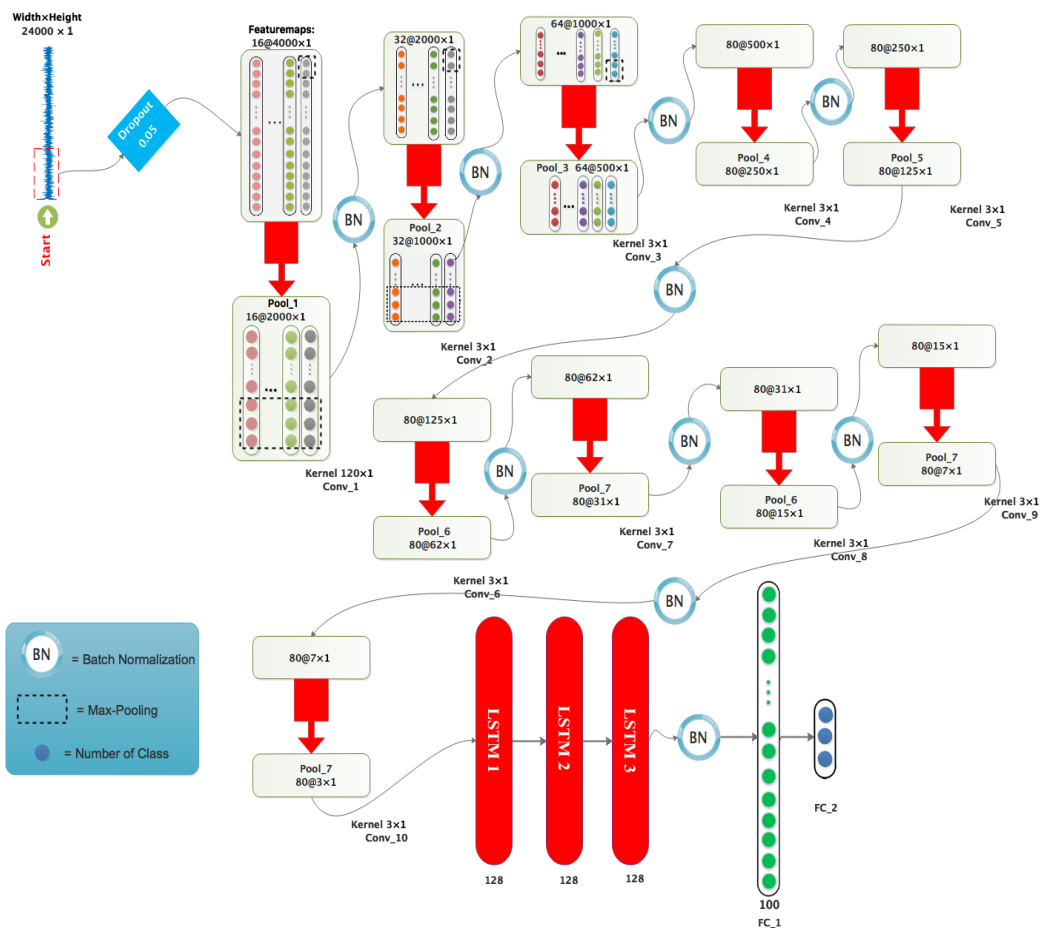
در این کار، ۲ سناریو مختلف در نظر گرفته شده است. در سناریو اول، ۲ حالت هیجانی (مثبت و منفی) در نظر گرفته می‌شود. در سناریو اول، ابتدا برای هر الکتروود، ۵ دقیقه (۳۰۰ ثانیه) از سیگنال برای حالت مثبت و حالت منفی انتخاب می‌شود؛ در این صورت ۲ نوع داده با ابعاد ۷۵۰۰۰ (فرکانس نمونه‌برداری برابر ۲۵۰ است) برای هر

تشخیص هیجانات القاشده با تحریک شنوایی از سیگنال‌های EEG بر اساس شبکه‌های یادگیری

است و در نهایت بردار ویژگی انتخاب شده به دولایه تماماً متصل با تابع غیرخطی Leaky Relu و Softmax متصل می‌شود.



شکل (۶): عملیات هم‌پوشانی (اورلپ) روی سیگنال ثبت شده.



شکل (۷): جزئیات شبکه عصبی عمیق پیشنهادی (CNN-LSTM).



شکل (۸): تخصیص داده‌های سیگنال EEG مربوط به سناریو اول و دوم.

جدول (۳): جزئیات معماری و سایز فیلترهای شبکه پیشنهادی.

L	Layer type	Activation function	Output Shape	Size of Kernel and Pooling	Strides	Number of filters and neurons	padding
0-1	Convolution1-D	Leaky ReLU	(None, 4000, 16)	120×1	6×1	16	yes
1-2	Max-Pooling1-D	-	(None, 2000, 16)	2×1	2×1	-	no
2-3	Convolution1-D	ReLU Leaky	(None, 2000, 32)	3×1	1×1	32	yes
3-4	Max-Pooling1-D	-	(None, 1000, 32)	2×1	2×1	-	no
4-5	Convolution1-D	Leaky ReLU	(None, 1000, 64)	3×1	1×1	64	yes
5-6	Max-Pooling1-D	-	(None, 500, 64)	2×1	2×1	-	no
6-7	Convolution1-D	Leaky ReLU	(None, 500, 80)	3×1	1×1	80	yes
7-8	Max-Pooling1-D	-	(None, 250, 80)	2×1	2×1	-	no
8-9	Convolution1-D	Leaky ReLU	(None, 250, 80)	3×1	1×1	80	yes
9-10	Max-Pooling1-D	-	(None, 125, 80)	2×1	2×1	-	no
10-11	Convolution1-D	Leaky ReLU	(None, 125, 80)	3×1	1×1	80	yes
11-12	Max-Pooling1-D	-	(None, 62, 80)	2×1	2×1	-	no
12-13	Convolution1-D	Leaky ReLU	(None, 62, 80)	3×1	1×1	80	yes
13-14	Max-Pooling1-D	-	(None, 31, 80)	2×1	2×1	-	no
14-15	Convolution1-D	Leaky ReLU	(None, 31, 80)	3×1	1×1	80	yes
15-16	Max-Pooling1-D	-	(None, 15, 80)	2×1	2×1	-	no
16-17	Convolution1-D	Leaky ReLU	(None, 15, 80)	3×1	1×1	80	yes
17-18	Max-Pooling1-D	-	(None, 7, 80)	2×1	2×1	-	no
18-19	Convolution1-D	Leaky ReLU	(None, 7, 80)	3×1	1×1	80	yes
19-20	Max-Pooling1-D	-	(None, 3, 80)	2×1	2×1	-	no
21-22	LSTM	Leaky ReLU	(None, 128)	-	-	128	-
22-23	LSTM	Leaky ReLU	(None, 128)	-	-	128	-
23-24	LSTM	Leaky ReLU	(None, 128)	-	-	128	-
24-25	Fully-connected	Leaky ReLU	(None, 100)	-	-	100	-
25-26	Fully-connected	Softmax	(None, 2-3)	-	-	2-3	-

آزمون انتخاب می‌شوند. همچنین، برای مجموعه آموزش ۱۰٪ از داده‌ها برای مجموعه اعتبارسنجی استفاده می‌شود. شکل ۸، تخصیص داده‌های EEG را برای مجموعه آموزش و آزمون برای سناریو اول و دوم نشان می‌دهد.

۴- نتایج و بحث

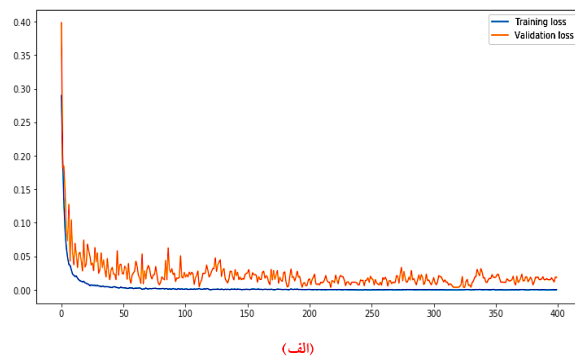
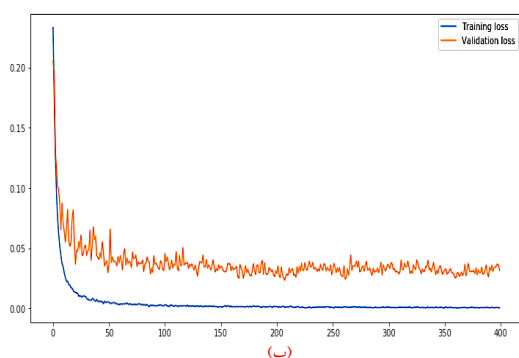
در این بخش، نتایج شبیه‌سازی شبکه عمیق پیشنهادی برای تشخیص خودکار هیجانات از سیگنال‌های EEG ارائه شده‌اند. شکل ۹ نمودار خطا را برای سناریو اول و دوم نشان می‌دهد. طبق شکل ۹ خطای شبکه برای سناریو اول با افزایش تکرار الگوریتم و در تکرار ۱۱۳۰ام تقریباً به حالت پایدار رسیده است. همچنین، با توجه به شکل ۹، نمودار

۳-۳- آموزش شبکه عمیق پیشنهادی

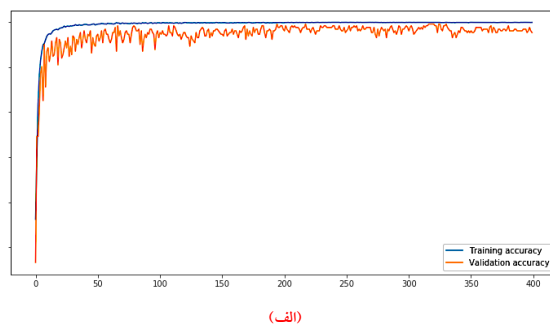
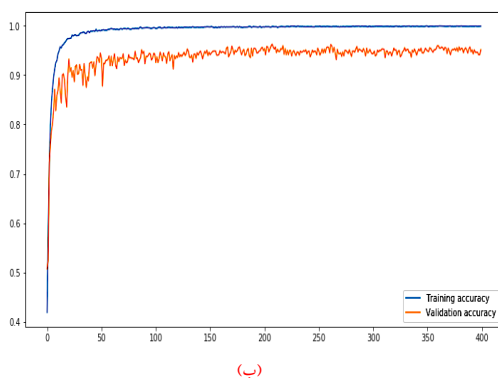
در این پژوهش، همه ابر پارامترهای شبکه پیشنهادی CNN-LSTM به‌دقت تنظیم شده‌اند تا بهترین نرخ همگرایی را به دست آورند و در نهایت، تابع هدف کراس آنروپی^{۳۶} و بهینه‌ساز Rmsprop با نرخ یادگیری ۰/۰۰۱ انتخاب شده‌اند. روش مرسوم پس انتشار خطا با اندازه دسته‌ای ۱۰ برای آموزش شبکه استفاده شده است. تعداد کل نمونه‌ها در این کار برای سناریو اول و دوم به ترتیب ۵۰۴۰ و ۷۵۶۰ است؛ از این تعداد ۶۰٪ داده‌ها به‌طور تصادفی برای آموزش شبکه (۳۰۲۴) برای سناریو اول و ۴۵۳۶ (برای سناریو دوم) و ۴۰٪ باقی‌مانده (۲۰۱۶) برای سناریو اول و ۳۰۲۴ (برای سناریو دوم) به‌عنوان مجموعه

با توجه به نمودار پراکندگی شکل ۱۳، معماری پیشنهادی برای طبقه‌بندی داده‌ها بسیار کارآمد بوده است. همچنین، برای نشان‌دادن عملکرد مطلوب الگوریتم پیشنهادی برای طبقه‌بندی، سناریو دوم (مثبت، خنثی و منفی) با سایر روش‌های رایج از جمله CNN، DBM و MLP مقایسه شد. برای شبکه CNN معماری شبکه پیشنهادی در جدول ۳ بدون در نظر گرفتن شبکه‌های LSTM به کار گرفته شد. برای شبکه DBM و MLP تعداد لایه‌های پنهان ۳ و نرخ یادگیری ۰/۰۰۱ در نظر گرفته شد. شکل ۱۴ عملکرد روش پیشنهادی (CNN-LSTM) را در مقایسه با شبکه‌های MLP و DBM برای سناریو دوم نشان می‌دهد.

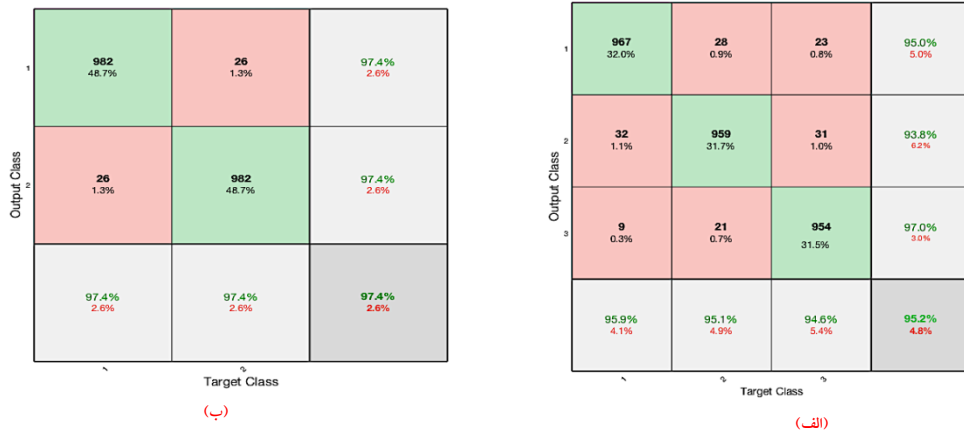
خطا برای سناریو دوم نیز در تکرار ۱۱۴۵م تقریباً ثابت شده است. شکل ۱۰ صحت روش پیشنهادی را برای سناریو اول و دوم در ۴۰۰ تکرار نشان می‌دهد. برای سناریو اول و دوم از شکل ۱۰ مشاهده می‌شود عملکرد روش پیشنهادی برای طبقه‌بندی هیجان‌ها پس از ۲۰۰ تکرار به ترتیب به صحت ۹۷/۴۲٪ و ۹۵/۲۳٪ می‌رسد. شکل ۱۱ ماتریس درهم‌ریختگی را برای سناریو اول و دوم نشان می‌دهد. مطابق شکل ۱۱، عملکرد شبکه عمیق پیشنهادی بسیار امیدوارکننده است. شکل ۱۲ نمودار بار (شامل دقت، حساسیت، صحت و اختصاصیت) را برای سناریو اول و دوم نشان می‌دهد. شکل ۱۳ نمودار پراکندگی در لایه‌های مختلف کانولوشن را برای سناریو اول و دوم نشان می‌دهد.



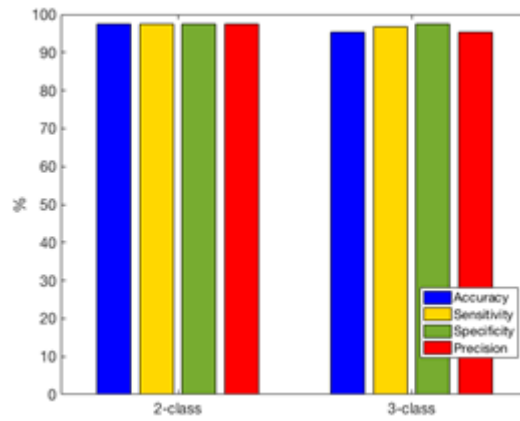
شکل (۹): نمودار خطا برای الف. سناریو اول (هیجان مثبت و منفی) و ب. سناریو دوم (هیجان مثبت، خنثی و منفی).



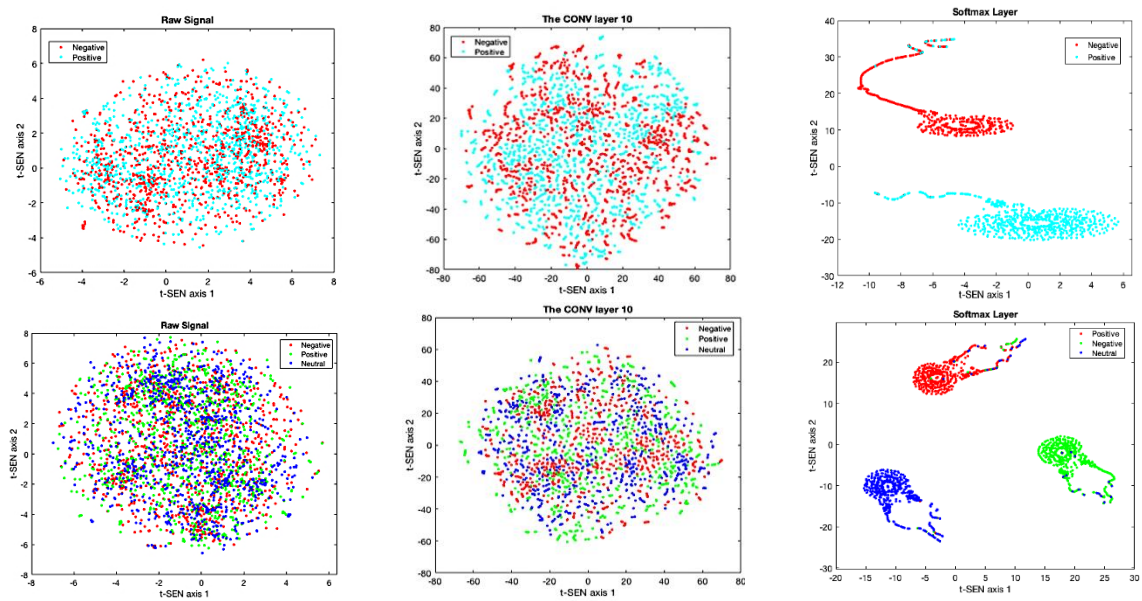
شکل (۱۰): نمودار صحت برای الف. سناریو اول (هیجان مثبت و منفی) و ب. سناریو دوم (هیجان مثبت، خنثی و منفی).



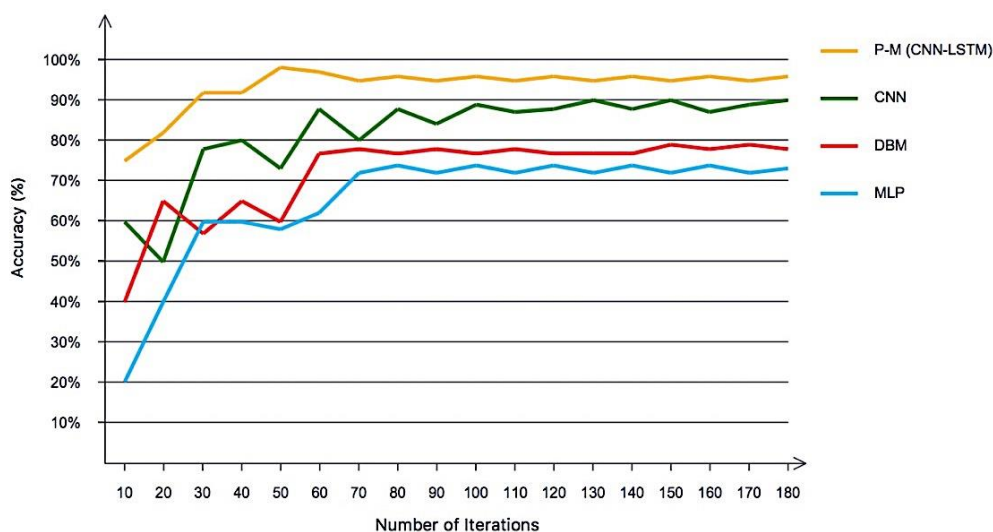
شکل (۱۱): ماتریس درهم‌ریختگی برای الف. سناریو اول (هیجان مثبت و منفی) و ب. سناریو دوم (هیجان مثبت، خنثی و منفی).



شکل (۱۲): نمودار میله‌ای (شامل صحت، حساسیت، اختصاصیت و دقت) برای سناریو اول و سناریو دوم.



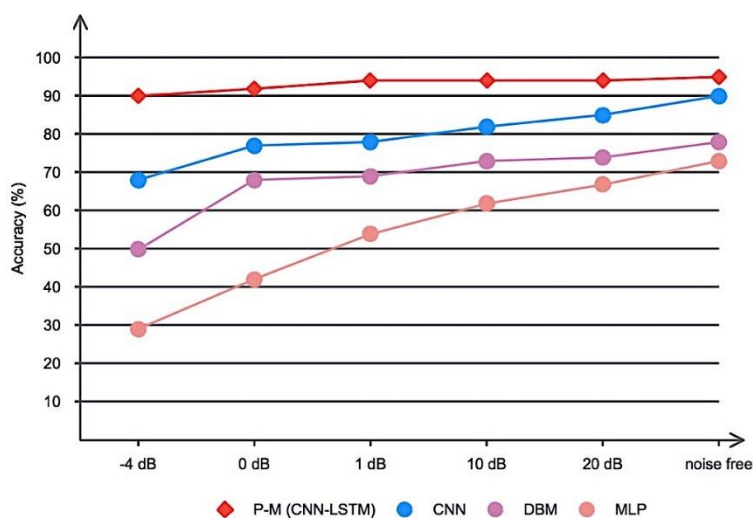
شکل (۱۳): نمودار t-SEN برای لایه‌های مختلف کانولوشن برای الف. سناریو اول و ب. سناریو دوم.



شکل (۱۴): عملکرد شبکه عمیق پیشنهادی (CNN-LSTM) در مقایسه با شبکه‌های CNN، DBM و MLP برای سناریو دوم.

جدول (۴): پیچیدگی محاسباتی برای الگوریتم پیشنهادی در مقایسه با ۳ روش دیگر.

Class	P-M (LSTM-CNN)		CNN		DBM		MLP	
	Train	Test	Train	Test	Train	Test	Train	Test
2-Stages	5400 s	5 s	5002 s	3 s	3011 s	4.5 s	909 s	2.5 s
3-Stages	12600 s	6 s	11200 s	3.5 s	6009 s	4.5 s	1201 s	2 s



شکل (۱۵): مقایسه صحت شبکه عمیق پیشنهادی در مقایسه با روش‌های دیگر در محیط نویزی.

مطابق شکل ۱۴، صحت حاصل شده برای ۳ شبکه مقایسه شده برای طبقه‌بندی به ترتیب ۹۰٪، ۷۹٪ و ۷۳٪ است. مطابق شکل ۱۴، معماری الگوریتم پیشنهادی براساس شبکه‌های CNN-LSTM در طبقه‌بندی مراحل هیجان‌ات مثبت، خنثی و منفی بسیار کارآمد بوده است. همچنین، پیچیدگی محاسباتی برای الگوریتم پیشنهادی (CNN-

مطابق شکل ۱۴، صحت حاصل شده برای ۳ شبکه مقایسه شده برای طبقه‌بندی به ترتیب ۹۰٪، ۷۹٪ و ۷۳٪ است. مطابق شکل ۱۴، معماری الگوریتم پیشنهادی براساس شبکه‌های CNN-LSTM در طبقه‌بندی مراحل هیجان‌ات مثبت، خنثی و منفی بسیار کارآمد بوده است. همچنین، پیچیدگی محاسباتی برای الگوریتم پیشنهادی (CNN-

لایه LSTM و ۲ لایه تماماً متصل استخراج شدند. نتایج نشان دادند مدل پیشنهادی، ویژگی‌هایی را انتخاب می‌کند که بیشترین تمایز را در بین کلاس‌ها ایجاد می‌کند. با توجه به بالابودن صحت الگوریتم، می‌توان از آن در کاربردهای سیستم‌های BCI استفاده نمود.

مراجع

- [1] S. Sanei and J. A. Chambers, EEG signal processing. John Wiley & Sons, 2013.
- [2] S. Sanei, "Adaptive processing of brain signals. 2013," ed: John Wiley & Sons.
- [3] K. R. Scherer, "What are emotions? And how can they be measured?," Social science information, Vol. 44, No. 4, pp. 695-729, 2005.
- [4] K. Schaaff and T. Schultz, "Towards emotion recognition from electroencephalographic signals," in Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on, 2009, pp. 1-6: IEEE.
- [5] M. Murugappan, M. Rizon, R. Nagarajan, S. Yaacob, I. Zunaidi, and D. Hazry, "EEG feature extraction for classifying emotions using FCM and FKM," International journal of Computers and Communications, Vol. 1, No. 2, pp. 21-25, 2007.
- [6] R. Jenke, A. Peer, and M. Buss, "Feature extraction and selection for emotion recognition from EEG," IEEE Transactions on Affective Computing, Vol. 5, No. 3, pp. 327-339, 2014.
- [7] S. Jirayucharoensak, S. Pan-Ngum, and P. Israsena, "EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation," The Scientific World Journal, Vol. 2014, 2014.
- [8] X.-W. Wang, D. Nie, and B.-L. Lu, "Emotional state classification from EEG data using machine learning approach," Neurocomputing, Vol. 129, pp. 94-106, 2014.
- [9] F. Hasanzadeh, M. Annabestani, and S. Moghimi, "Continuous Emotion Recognition during Music Listening Using EEG Signals: A Fuzzy Parallel Cascades Model," arXiv preprint arXiv:1910.10489, 2019.
- [10] Y. Hou and S. Chen, "Distinguishing Different Emotions Evoked by Music via Electroencephalographic Signals," Computational intelligence and neuroscience, vol. 2019, 2019.
- [11] P. Keelawat, N. Thammasan, M. Numao, and B. Kijssirikul, "Spatiotemporal Emotion Recognition using Deep CNN Based on EEG during Music Listening," arXiv preprint arXiv:1910.09719, 2019.
- [12] Y. Yang, Q. Wu, M. Qiu, Y. Wang, and X. Chen, "Emotion recognition from multi-channel EEG through parallel convolutional recurrent neural network," in Proc. Int. Joint Conf. Neural Netw. (IJCNN), pp. 1-7, Jul. 2018.
- [13] H. Yang, J. Han, and K. Min, "A multi-column CNN model for emotion recognition from EEG

می‌دهد. در مطالعات قبلی از روش‌های متداول مانند تبدیل موجک^{۳۷} (WT)، تجزیه حالت تجربی^{۳۸} (EMD) و غیره برای استخراج و انتخاب ویژگی‌های مهم سیگنال استفاده شده که شامل برخی از مشکلات رایج در رابطه با پارامترهای روش انتخاب و استخراج ویژگی مانند انتخاب نوع موجک مادر، تعداد سطوح تجزیه و غیره است. روش پیشنهادی بدون استفاده از روش‌های معمول طبقه‌بندی، به حذف بلوک انتخاب ویژگی منجر می‌شود و بدون استفاده از الگوریتم‌های استخراج ویژگی متداول، به‌طور مستقیم از روی خود سیگنال EEG به تشخیص هیجانات می‌پردازد.

جدول (۵): مقادیر کاپاکوهن برای سناریو اول و دوم.

Class	2-stages	3-stages
Kappa	0.96	0.93

به‌منظور ارزیابی عملکرد الگوریتم پیشنهادی در سناریوهای نویز آلود، نویز گوسی سفید با SNR -۴ تا ۲۰ دسی‌بل به سیگنال‌های EEG ثبت شده اضافه شد و صحت طبقه‌بندی سناریو دوم در هر طیف در مقایسه با روش‌های مقایسه شده در شکل ۱۵ بررسی شد. نتایج حاصل شده در شکل ۱۵ نمایش داده شده‌اند. با توجه به شکل ۱۵، عملکرد طبقه‌بندی الگوریتم پیشنهادی به‌طور چشمگیری نسبت به نویز در طیف گسترده‌ای از SNR نسبت به روش‌های مقایسه شده مقاوم است.

۵- نتیجه‌گیری

با توجه به مطالب بخش اول، چالش عمده در بازشناسی هیجانات، انتخاب ویژگی متمایزکننده هیجانات از یکدیگر بود که ادغام آن با روش‌های مختلف طبقه‌بندی به پدید آمدن نتایج مختلف و متفاوت منجر شده بود. روش پیشنهادی این مقاله بدون استفاده از روش‌های عمدتاً پیچیده طبقه‌بندی، به حذف بلوک انتخاب ویژگی منجر شده و بدون استخراج ویژگی‌های متداول، به‌طور مستقیم از روی خود سیگنال EEG موفق به بازشناسی ۲ سناریو مختلف با صحت بالای ۹۰٪ شده است. در این پژوهش، ویژگی‌ها به‌صورت سلسله‌مراتبی با استفاده از ۷ لایه کانولوشن، ۳

- neural networks by preventing co-adaptation of feature detectors," arXiv preprint arXiv:1207.0580, 2012.
- [18] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," arXiv preprint arXiv:1502.03167, 2015.
- [19] N. Siddique and H. Adeli, Computational intelligence: synergies of fuzzy logic, neural networks and evolutionary computing. John Wiley & Sons, 2013.
- [20] Z. Mousavi., et al. "Deep convolutional neural network for classification of sleep stages from single-channel EEG signals." Journal of neuroscience methods 108-312. 2019 .
- [21] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural computation, Vol. 9, No. 8, pp. 1735-1780, 1997.
- signals," Sensors, Vol. 19, No. 21, p. 4736, Oct, 2019
- [14] J. Chen, D. Jiang, Y. Zhang, and P. Zhang, "Emotion recognition from spatiotemporal EEG representations with hybrid convolutional recurrent neural networks via wearable multi-channel headset," Comput. Commun, Vol. 154, pp. 58-65, Mar. 2020.
- [15] Mohammadzadeh. A, Applications of music therapy in the fields of psychiatry, medicine and psychology, including music and mysticism, classification of music themes, Classification of music themes, music therapy and global unity. Tehran: Secrets of Knowledge, Vol. 104, pp. 10-20, Mar. 2005.
- [16] Goodfellow. I., Bengio. Y., Courville. A., 2016. Deep Learning. MIT Press, <http://www.deeplearningbook.org>.
- [17] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving

1 Emotions

2 Functional Magnetic Resonance Imaging (fMRI)

3 Electroencephalography (EEG)

4 Near-Infrared Spectroscopy (NIRS)

5 Brain-Computer Interface (BCI)

6 Pleasant

7 Neutral

8 Unpleasantness

9 Fast Fourier Transform (FFT)

10 Support Vector Machine (SVM)

11 Relaxation

12 Linear Regression

13 Fuzzy Clustering

14 International Affective Picture system (IAPS)

15 Self-Assessment Manikin (SAM)

16 Quadratic Discriminant Analysis (QDA)

17 Deep Learning Network (DLN)

18 Principal Component Analysis (PCA)

^{۱۹}آزمون خودارزیابی تصاویر آدمک (SAM) نمایش تصویری از ابعاد سه‌گانه است که لانگ به‌عنوان جایگزینی برای مقیاس‌های خودگزارش‌دهی ابداع کرد. نمایش‌های آن در هر بعد با یک ویژگی تصویری در یک مقیاس ۹ درجه‌ای است که پاسخ‌دهنده می‌تواند آنچه را احساس می‌کند، انتخاب کند. SAM از شکل‌های آدمک در یک مقیاس برای هر یک از ابعاد عاطفی استفاده می‌کند. از SAM به‌عنوان یک مقیاس کاربردی تصویری برای سنجش ابعاد هیجان مثبت، منفی و خنثی استفاده می‌شود.

20 Beck Depression Inventory

21 Electrooculography (EOG)

22 Feedforward

23 Back Propagation

24 Loss Function

25 Pooling Layer

26 Fully Connected

27 Over Fitting

28 Dropout

29 Batch Normalization

30 Down Sampling

- 31 Max-Pooling
- 32 Softmax
- 33 Long Short Term Memory (LSTM)
- 34 Recurrent Neural Network (RNN)
- 35 Notch Filter
- 36 Cross Entropy
- 37 Wavelet Transform
- 38 Empirical Mode Decomposition

