

## استفاده از معماری شبکه عصبی پیچشی برای طبقه‌بندی احساسات با مطالعه سیگنال‌های الکتروانسفالوگرافی

محمد رضا فرجی

کارشناسی مهندسی کامپیوتر، دانشکده فنی فومن، دانشکده‌گان فنی، دانشگاه تهران، فومن، ایران  
پست الکترونیکی: m.rezafaraji@ut.ac.ir

عاطفه حسن‌زاده\*

استادیار، دانشکده فنی فومن، دانشکده‌گان فنی، دانشگاه تهران، فومن، ایران  
پست الکترونیکی: hasanzadeh.a@ut.ac.ir

### چکیده

طبقه‌بندی احساسات استفاده شده است که پس از انجام یک سری پیش‌پردازش‌ها مانند تبدیل موجک پیوسته<sup>۱</sup> بر روی سیگنال‌ها، آن‌ها را به‌عنوان ورودی به شبکه عصبی پیچشی سه‌بعدی داده که میزان دقت به‌دست آمده از شبکه طراحی شده در داده‌های آموزشی برابر ۰/۹۸۰۷۲۹ و بر داده‌های آزمایشی برابر ۰/۹۸۸۹ است. این امر نشان می‌دهد استفاده از این معماری خطای کمی دارد و وظیفه<sup>۲</sup> محول شده به آن را با دقت بالا و خطای کمی انجام می‌دهد. **واژه‌های کلیدی:** تشخیص احساسات، شبکه عصبی پیچشی، سیگنال الکتروانسفالوگرافی، تبدیل موجی پیوسته، محیط کولب در پایتون.

### مقدمه

احساس به‌عنوان واکنش یا آگاهی نسبت به محرک‌های خارجی تعریف می‌شود و نقش مهمی در زندگی روزمره ایفا می‌کند زیرا بر روند زندگی افراد تأثیر می‌گذارد. عواطف اساسی مانند شادی، خشم و غم و اندوه، به‌طور داوطلبانه

مسئله تشخیص احساسات حائز اهمیت فراوانی است و در حوزه‌های گوناگون دارای کاربرد است که از جمله آن‌ها می‌توان به موارد پزشکی اشاره کرد که می‌تواند به درمان بیماران سرعت ببخشد. این مطالعات نشان داده است که از سیگنال‌های مغزی می‌توان برای طبقه‌بندی بسیاری از حالات عاطفی استفاده کرد. این روند دشوار به نظر می‌رسد، به‌خصوص که سیگنال‌های مغز پایدار نیستند چراکه واکنش به حالت‌های مختلف احساسی که روی سیگنال‌های مغز تأثیر می‌گذارند متنوع‌اند. بنابراین، عملکرد سیستم‌های تشخیص احساسات توسط سیگنال‌های مغزی به کارایی الگوریتم‌های مورد استفاده بستگی دارد. اخیراً، مطالعه روی سیگنال بندی الکتروانسفالوگرافی<sup>۱</sup> به دلیل در دسترس بودن مورد توجه بسیاری قرار گرفته است. در این پژوهش از معماری شبکه‌های عصبی پیچشی<sup>۲</sup> و به‌طور دقیق‌تر شبکه‌های عصبی پیچشی سه‌بعدی<sup>۳</sup> برای

\* نویسنده مسئول

- 1- Electroencephalographic signaling
- 2- Convolutional Neural Network (CNN)
- 3- 3D Convolutional Neural Network (3D\_CNN)

4- continuous wavelet transform (CWT)

5- task

یا غیرارادی توسط افراد منعکس می‌شود و این امر به‌طور قابل توجهی بر موقعیت افراد در جامعه تأثیر می‌گذارد. افراد دارای عواطف منفی می‌توانند از جامعه طرد شوند که به‌نوبه خود از منظر فیزیولوژیکی و روانی تأثیرگذار خواهد بود [۱]. علاوه بر این، افرادی که دارای احساسات مثبت هستند، استانداردهای زندگی بهتری دارند و عمر طولانی‌تری دارند [۲]. مطالعات زیادی در مورد تحلیل احساسات در ادبیات وجود دارد. اخیراً محققان به‌طور فعال در این زمینه برای درک و تجزیه و تحلیل رفتار احساسات فعالیت می‌کنند. اما انتزاعی بودن و متفاوت بودن عواطف از فردی به فرد دیگر باعث می‌شود مطالعات ناکارآمد باشد و پژوهشها در این زمینه به حد مطلوب نرسد [۳]. علاوه بر این، جمع‌آوری انبوهی از داده‌ها (مربوط به حالت‌های چهره، سیگنال‌های صوتی، سیگنال‌های مغز، زبان بدن و ...) باعث پیچیده شدن داده‌ها و طولانی شدن فرآیند تجزیه و تحلیل می‌شود. به همین دلایل، نیاز به استفاده از روش‌های تحلیل مبتنی بر هوش مصنوعی و کامپیوتری به وجود آمده است [۴]. احساسات را می‌توان با روش‌های مختلفی از جمله سیگنال‌های صوتی، حالات چهره، زبان بدن و سیگنال‌های فیزیولوژیکی به دست آورد. افراد احساسات خود را از طریق صداها، کلمات و رفتار منتقل می‌کنند. پیش‌بینی احساسات با تجزیه و تحلیل شدت صدا، قدرت صدا، سطح گفتار و سرعت گفتار انجام می‌شود [۵]. روش دیگری که برای به دست آوردن احساسات استفاده می‌شود، تجزیه و تحلیل حالات چهره است. در این روش موارد تحت آزمون در مقابل دوربین می‌ایستند و واکنش آن‌ها به محرک‌ها مشاهده می‌شود [۶].

به‌طور مشابه، احساسات با زبان بدن و حرکات فیزیکی بررسی می‌شوند. به‌عنوان مثال، یک فرد عصبانی لب‌های خود را محکم به هم می‌فشارد، یک فرد بی‌حوصله از پنجره به بیرون نگاه می‌کند یا پاهای خود را می‌کوبد [۷]. با این حال، دست‌کاری مداوم احساسات به دست آمده با این روش‌ها ممکن است موجب شود داده‌ها نادرست

و غیرقابل اعتماد باشند. افراد به راحتی می‌توانند حالت‌ها، لحن‌ها و حرکات خود را دست‌کاری کنند [۸]. به این دلایل، نیاز به یک سیستم قابل اعتمادتر ایجاد شده و اهمیت سیگنال‌های فیزیولوژیکی افزایش یافته است. رایج‌ترین سیگنال‌های مورد استفاده در مطالعات تشخیص احساسات عبارت‌اند از الکتروانسفالوگرافی<sup>۶</sup>، الکتروکاردیوگرافی<sup>۷</sup>، فتوپلتیسموگرافی<sup>۸</sup> و دمای پوست<sup>۹</sup> [۹]. با این حال، در بین این روش‌ها سیگنال‌های الکتروانسفالوگرافی ترجیح داده شده است. دلایل اصلی این وضعیت سهولت استفاده و مقرون به صرفه بودن در مقایسه با سایر روش‌ها است [۱۰].

الکتروانسفالوگرافی تکنیکی برای جمع‌آوری اطلاعات از یک فرد با قرار دادن یک دستگاه پوشیدنی در اطراف سر به روشی غیرتهاجمی است. این دستگاه حاوی الکترودهایی است که مطابق با سیستم ۱۰-۲۰ عدد در اطراف سر قرار گرفته‌اند، این دستگاه بر فعالیت مغز نظارت می‌کند و سیگنال‌های مغز را ثبت می‌کند. این دستگاه‌ها سیگنال‌های ضبط‌شده را از طریق ارتباطات سیمی یا بی‌سیم به منظور تجزیه و تحلیل به رایانه‌ها منتقل می‌کنند. الکتروانسفالوگرافی روشی آسان و ارزان است که برای ثبت فعالیت مغز استفاده شده است و همچنین ثابت کرده است که از طریق فعالیت مغز است که می‌توان احساسات را تشخیص داد. اندازه‌گیری فعالیت مغز از طریق دستگاه‌های الکتروانسفالوگرافی به دانشمندان کمک کرده است تا احساسات افراد دارای معلولیت را درک کنند که تشخیص احساسات از طریق حالات چهره آن‌ها دشوار است [۱۳].

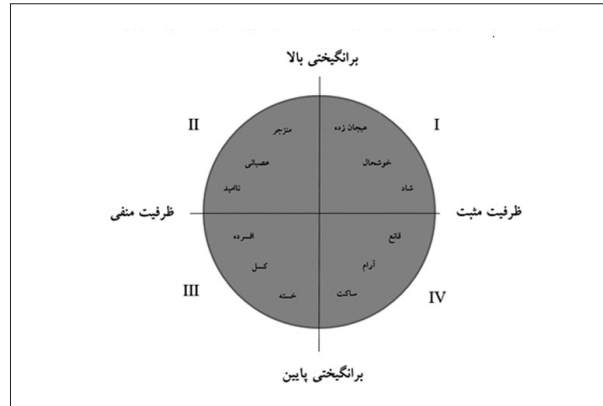
دو نوع مدل احساسی در نوشتار وجود دارد، گسسته و بعدی. مدل احساسات گسسته شامل هشت هیجان اساسی (خشم، ترس، غم، انزجار، تعجب، انتظار، اعتماد و شادی) دربردارنده هیجان‌ات مثبت و منفی است [۱۴].

6-Electroencephalographic (EEG)

7- Electrocardiography (ECG)

8- Photoplethysmogram (PPG)

9- skin temperature



شکل ۱. احساسات انسان بر پایه مدل‌های ظرفیت و برانگیختگی [۱۶]

برعکس، در مدل ابعادی، احساسات نه با نام خود، بلکه با موقعیت آن‌ها در سطح برانگیختگی-ظرفیت بیان می‌شوند [۱۵]. احساسات در چهار ناحیه مختلف در سطح برانگیختگی-ظرفیت قرار می‌گیرند. صفحه ظرفیت نشان‌دهنده احساسات منفی و مثبت است. سمت راست صفحه ظرفیت مثبت را مشخص می‌کند، درحالی‌که سمت چپ نشان‌دهنده ظرفیت منفی است. در سطح برانگیختگی، احساسات از احساسات غیرفعال به احساسات فعال از پایین به بالا طبقه‌بندی می‌شوند. در شکل ۱، یک ساختار کلی از مدل هیجانی بعدی ارائه شده است.

همان‌طور که از شکل ۱ می‌توان دریافت، صفحه برانگیختگی-ظرفیت به چهار ناحیه مجزا تقسیم می‌شود. درحالی‌که احساسات مثبت با برانگیختگی بالا در منطقه اول قرار می‌گیرند، احساسات منفی با برانگیختگی بالا در منطقه دوم وجود دارد و بالعکس احساسات منفی با برانگیختگی کم در نواحی سوم و چهارم وجود دارد که ناحیه سوم شامل احساسات منفی با برانگیختگی کم است و منطقه چهارم از احساسات مثبت با برانگیختگی کم تشکیل شده است. با کمک این ساختار، احساسات نه با نام آن‌ها بلکه با موقعیت آن‌ها بیان می‌شود. بنابراین، این مدل به‌عنوان یک مدل جهانی بیان می‌شود [۱۶]. علاوه بر این، در ادبیات، سه نوع محرک مختلف وجود دارد (محرک‌های شنیداری، دیداری، شنیداری/دیداری) که در آن‌ها از محرک‌ها برای تحریک فرد برای رسیدن به احساساتی که باید استخراج

شوند، استفاده می‌شود [۱۷]. محرک‌های شنیداری از صداها ساخته شده‌اند و معمولاً شامل موسیقی و کلیپ‌های صوتی است. محرک‌های بینایی مانند تصاویری هستند که بانام‌های مختلف برچسب‌گذاری شده‌اند. در مورد محرک‌های شنیداری/دیداری (که از بهترین محرک‌ها برای استخراج احساسات محسوب می‌شوند)، در اغلب مطالعات، هم صداها و هم تصاویر بصری را در برمی‌گیرد [۱۸-۱۹]. در این پژوهش، فرآیند تشخیص هیجان با استفاده از مدل احساسات بعدی و مدل هیجان گسسته انجام شده است. مطالعه شامل چهار مرحله است:

در مرحله اول، داده‌های احساسات الکتروانسفالوگرافی از مجموعه داده عمیق<sup>۱</sup> به‌دست آمده است. در مرحله دوم، سیگنال‌های الکتروانسفالوگرافی از طریق تبدیل موجک پیوسته پیش‌پردازش می‌شود. در مرحله سوم این موجک‌ها که از طریق سیگنال به دست آمده به قاب‌هایی به‌اندازه ۱ ثانیه تقسیم می‌شود و سپس به‌صورت سه‌بعدی در کنار هم قرار داده می‌شود و نتیجه به‌عنوان ورودی به شبکه عصبی پیچشی سه‌بعدی داده می‌شود که در داده‌های آموزشی دقت متوسط ۹۸ درصد و بر داده‌های آزمایشی دقت تقریبی ۹۲ درصد حاصل می‌شود.

به‌این‌ترتیب بخش‌های اصلی این مقاله را می‌توان به شرح زیر فهرست کرد:

- ۱) بررسی مطالعات مرتبط و تقسیم‌بندی آن‌ها به دو روش یادگیری ماشین و یادگیری عمیق
- ۲) بررسی روش که شامل مطالعه مجموعه داده مورد استفاده، معماری شبکه‌های سه‌بعدی و نحوه انجام کار در آن‌ها و پیش‌پردازش‌های انجام‌شده بر روی سیگنال‌ها و بررسی معماری سه‌بعدی است که از آن برای دسته‌بندی احساسات استفاده شده است.
- ۳) نتایج حاصل از شبکه پیچشی سه‌بعدی و مقایسه نتایج به‌دست‌آمده با مطالعات پیشین
- ۴) جمع‌بندی و نتیجه‌گیری کلی در مورد موضوع و بررسی چالش‌های موجود در این حوزه

## ۲. پژوهش‌های مرتبط در مطالعات پیشین

معمولاً دو راهبرد یادگیری ماشین برای تجزیه و تحلیل سیگنال‌های الکتروانسفالوگرافی وجود دارد: یادگیری ماشینی گام‌به‌گام و یادگیری عمیق [۲۰]. یادگیری ماشینی گام‌به‌گام عمدتاً شامل سه مرحله است:

- مرحله اول به دست آوردن داده‌های دیجیتال با نمونه‌برداری از سیگنال‌های آنالوگ است که به‌عنوان پیش‌پردازش سیگنال شناخته می‌شود.

- مرحله دوم استخراج ویژگی مصنوعی است که با استفاده از فرمول‌های استخراج ویژگی، ویژگی‌ها را محاسبه می‌کند.

- در مرحله سوم، ویژگی‌ها با استفاده از روش‌های یادگیری ماشینی برای دستیابی به نتیجه طبقه‌بندی احساسات طبقه‌بندی می‌شوند.

تبدیل موجک و معیارهای آنتروپی به‌طور گسترده در استخراج ویژگی استفاده می‌شود [۲۱]. موروپاپان<sup>۱۱</sup> و همکاران از تابع موجک برای استخراج مجموعه‌ای از ویژگی‌های متعارف و مبتنی بر انرژی اصلاح‌شده از سیگنال‌های الکتروانسفالوگرافی برای طبقه‌بندی احساسات استفاده کردند [۲۲]. پل<sup>۱۲</sup> و همکاران از روش تجزیه و تحلیل نوسانات بدون روند برای استخراج ویژگی‌ها استفاده کردند و یک ماشین بردار پشتیبان برای طبقه‌بندی فضای ویژگی الکتروانسفالوگرافی مربوط به حالات احساسی مختلف در رده‌های مربوطه را مورد استفاده قرار دادند [۲۳]. جیانگ<sup>۱۳</sup> و همکاران از یادگیری انتقال برای کاهش تفاوت در توزیع داده‌ها بین داده‌های آموزشی و آزمایشی استفاده کردند [۲۴]. همچنین در مطالعه‌ای دیگر یک مدل خوشه‌بندی فازی مقاوم با فضای پنهان انتقال متقابل دامنه مشترک پیشنهاد کردند [۲۵].

با این حال، پوشش دادن همه ویژگی‌های ضمنی با استخراج دستی دشوار است و فرمول‌های مورد استفاده

11- Murugappan  
12- Paul  
13- Jiang

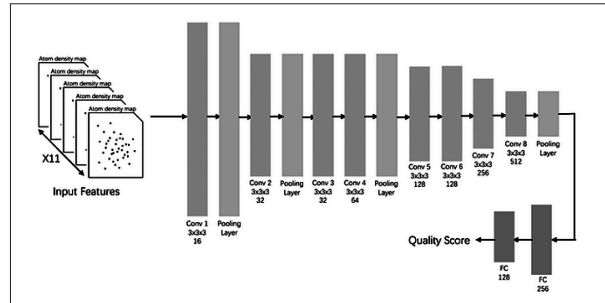
برای استخراج ویژگی‌های حوزه زمان و دامنه فرکانس اغلب بسیار پیچیده هستند. علاوه بر این، سیگنال‌های الکتروانسفالوگرافی مستعد نوفه‌هایی مانند مصنوعات الکترومیوگرافی هستند که تداخل جدی در رویکرد یادگیری ماشین ایجاد می‌کنند. با توجه به شرایط فوق، برخی از روش‌های یادگیری عمیق برای حل این مشکلات استفاده می‌شود. الهاگری<sup>۱۴</sup> و همکاران مدل حافظه طولانی کوتاه‌مدت<sup>۱۵</sup> را برای یادگیری ویژگی‌های سیگنال‌های الکتروانسفالوگرافی پیشنهاد کردند. دقت طبقه‌بندی برای برچسب‌های مختلف به ۸۵/۶۵، ۸۵/۴۵ و ۸۷/۹۹ درصد رسید [۲۶].

شرمایستر<sup>۱۶</sup> و همکاران از شبکه‌های عصبی پیچشی برای رمزگشایی و تجسم سیگنال‌های الکتروانسفالوگرافی استفاده کردند که پتانسیل زیادی را در تشخیص احساسات سرتاسری بر اساس سیگنال‌های الکتروانسفالوگرافی نشان داده‌اند [۲۷]. ژانگ<sup>۱۷</sup> و همکاران استخراج ویژگی کاملاً خودکار طبقه‌بندی حجم کاری ذهنی<sup>۱۸</sup> را بهبود بخشیدند که در مقایسه با روش‌های یادگیری ماشین سنتی عملکرد بالایی داشت [۲۸]. از آنجایی که مدل‌های یادگیری عمیق برای تشخیص احساسات مبتنی بر الکتروانسفالوگرافی هنوز در مراحل ابتدایی خود هستند، همچنان ظرفیت بسیاری برای تنظیم ساختار مدل و تنظیمات پارامتر وجود دارد.

## ۳. روش انجام کار

در این مقاله از یکی از معماری‌های یادگیری عمیق برای تشخیص احساسات با استفاده از سیگنال الکتروانسفالوگرافی استفاده می‌شود. روشی که برای پیاده‌سازی در نظر گرفته شده استفاده از شبکه عصبی پیچشی سه‌بعدی است. قبل از این که به بیان جزئیات پردازیم، ابتدا اشاره‌ای به شبکه مذکور خواهیم داشت تا

14- Alhagry  
15- long Short-Term Memory (LSTM)  
16- Schirrmeyer  
17- Zhang  
18- Mental Workload (MWL)



شکل ۲: طرحی از پیش سه بعدی

متوجه شویم که اساس کار آن به چه شکل است تا بهتر بتوان استفاده از آن در طول پیاده‌سازی را درک نمود.

### ۱.۳ شبکه عصبی پیش سه بعدی

یک شبکه عصبی پیش سه بعدی از یک فیلتر سه بعدی برای انجام پیش یا هم آمیختی استفاده می‌کند. هسته قادر به لغزش در سه جهت است، درحالی‌که در شبکه عصبی پیش دو بعدی می‌تواند در دو بعد حرکت کند. ورودی‌ای که سه کانال یا صفحه دارد موجب می‌شود، فیلتر مورد نظر هم سه کانال داشته باشد. یعنی به‌عنوان مثال به‌جای یک فیلتر  $3 \times 3$  برای پیش یا هم آمیختی در ورودی دو بعدی، اکنون یک فیلتر  $3 \times 3 \times 3$  وجود دارد. دلیل این امر این است که هر صفحه از ورودی به‌صورت مجزا برای خودش یک فیلتر دارد. سه صفحه از فیلتر به‌صورت موازی باهم روی سه صفحه از ورودی حرکت می‌کنند. با فرض این‌که فیلتر  $5 \times 5$  باشد، در هر لحظه ۲۵ ضرب در هر صفحه انجام می‌شود و در مجموع برای سه صفحه ۷۵ ضرب صورت می‌گیرد. در نهایت خروجی برابر است با جمع این ۷۵ حاصل ضرب. در نهایت نتیجه حاصل ضرب سه صفحه باهم جمع می‌شود و خروجی نهایی یک تصویر  $100 \times 100 \times 3$  نیست، بلکه یک تصویر  $100 \times 100$  است (شکل ۲).

به‌عنوان مثال، اگر ورودی به جای ۳ صفحه، ۱۰ صفحه داشته باشد، فیلتر باید ۱۰ صفحه داشته باشد و اگر فیلتر  $5 \times 5$  در نظر گرفته شود، اندازه فیلتر به شکل  $10 \times 5 \times 5$  می‌باشد. هر یک از این ۱۰ صفحه، در صفحه متناظر با صفحه تصویر تلفیق می‌شود.

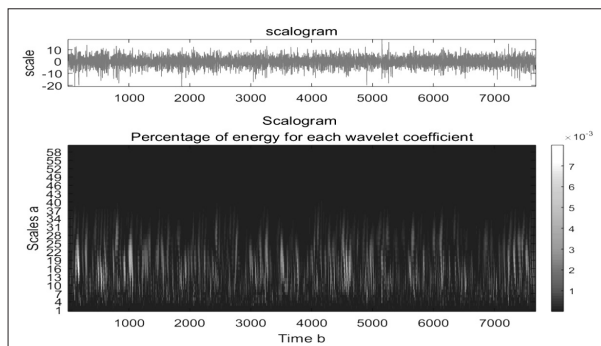
پس از آشنایی با شبکه‌های عصبی پیش سه بعدی و پیش‌های سه بعدی، آن شبکه با مجموعه داده مورد استفاده برای طراحی شبکه آموزش<sup>۱۹</sup> داده می‌شود. مجموعه داده مورد استفاده در این مقاله، مجموعه داده عمیق است که ابتدا به معرفی این مجموعه داده پرداخته می‌شود.

### ۲.۳ مجموعه داده عمیق

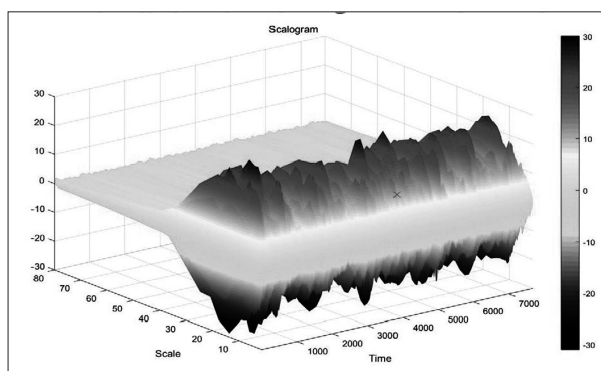
مجموعه داده عمیق یک مجموعه داده دارای چند حالت برای تجزیه و تحلیل حالات عاطفی انسان ارائه می‌دهد [۲۹]. الکتروانسفالوگرام و سیگنال‌های فیزیولوژیکی محیطی مربوط به ۳۲ شرکت‌کننده درحالی‌که آن‌ها هر یک از ۴۰ ویدیوی یک دقیقه‌ای از ویدیوهای نشان داده شده را تماشا می‌کردند، ضبط شده است. شرکت‌کنندگان هر ویدئو را از نظر سطح تحریک، قدرت، مانند دوست‌داشتن، سلطه و آشنایی رتبه‌بندی می‌کنند. برای ۲۲ نفر از ۳۲ شرکت‌کننده فیلم چهره جلویی نیز ثبت شده است [۳۰]. در شکل ۳ نمایی از امتیازدهی کاربران به ویدیوها نشان داده شده است.

پس از آشنایی با مجموعه داده و نحوه به وجود آمدن آن‌ها، در مرحله بعدی بر روی این مجموعه داده‌ها یک سری پیش پردازش انجام می‌شود تا آن‌ها به‌صورت استاندارد وارد شبکه شده و بتوان داده‌های حاصل را آموزش داد. برای این منظور، داده‌های سیگنال اصلی از طریق تبدیل موجک پیوسته پیش‌پردازش می‌شود و از یک معماری شبکه عصبی پیش سه بعدی به‌عنوان طبقه‌بندی کننده استفاده می‌شود. لازم به ذکر است که رابطه  $Wt = CWt(x)$  تبدیل موجک پیوسته  $x$  را برمی‌گرداند. ورودی یعنی  $x$  یک بردار یا یک عدد مختلط یا یک جدول زمانی متغیر منفرد است که به‌طور منظم نمونه‌برداری می‌شود و حداقل باید چهار نمونه داشته باشد.

تبدیل موجک پیوسته با استفاده از موجک تحلیلی مورس با پارامتر تقارن (گاما) برابر با ۳ و پهنای باند زمان برابر با ۶۰ به دست می‌آید [۳۱]. تبدیل موجک پیوسته



شکل ۵: تبدیل داده‌های اصلی به ضرایب موجک و انرژی موجک



شکل ۶: نمودار سطحی اسکالوگرام بر اساس داده‌های ورودی پس از پیش‌پردازش داده‌ها

این مرحله، شکل داده از ۳۲ کانال و ۸۰۶۴ نقاط نمونه به ۳۲\*۶۴ (مقیاس) و ۷۶۸۰ نقطه نمونه تبدیل می‌شود (شکل‌های ۵ و ۶).

در این مرحله از نرم‌افزار متلب برای تبدیل سیگنال‌های اصلی برای ضرایب موجک استفاده می‌شود. سپس به مراحل بعدی پیش‌پردازش داده‌ها پرداخته می‌شود.

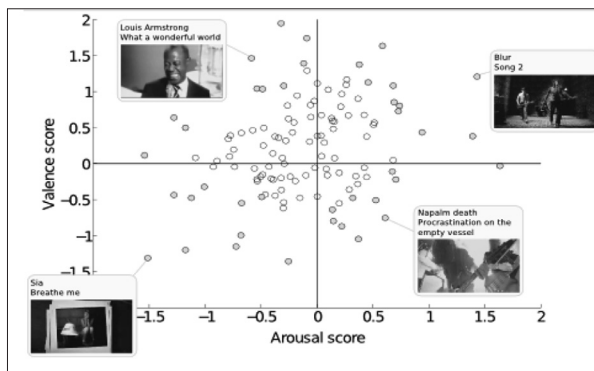
(۲) برش دادن قاب‌ها<sup>۲۰</sup>

سپس 1s به عنوان طول قاب تنظیم می‌شود، بنابراین می‌توان ۶۰ قاب را در یک فیلم 60s دریافت کرد. شکل هر قاب ۳۲(کانال)\*۶۴(مقیاس) است [۳۴].

(۳) انتخاب مقیاس<sup>۲۱</sup>

سپس میانگین نسبت بهره‌وری انرژی<sup>۲۲</sup> برای همه ۶۴ مقیاس در ۳۲ کانال محاسبه می‌شود و مقیاس 8th ~ 39th برای کاهش محاسبات انتخاب می‌شود. با توجه به برنامه نوشته‌شده برای انتخاب مقیاس‌ها در محیط نرم‌افزار متلب، مطابق شکل ۷، می‌توان میانگین نسبت بهره‌وری انرژی بر

20- Cut frames  
21- Select scale  
22- Energy Efficiency Ratio (EER)



شکل ۴: نمونه‌ای از تبدیل موجک پیوسته

از ۱۰ صدا در هر اکتاو استفاده می‌کند. حداقل و حداکثر مقیاس‌ها به‌طور خودکار بر اساس گسترش انرژی موجک در فرکانس و زمان تعیین می‌شوند. اگر  $X$  دارای مقدار حقیقی باشد،  $Wt$  یک ماتریس دوبعدی است که هر ردیف با یک مقیاس مطابقت دارد. اندازه ستون  $Wt$  برابر با طول  $X$  است. اگر  $X$  مختلط باشد،  $Wt$  یک ماتریس سه‌بعدی است که در آن صفحه اول تبدیل موجک پیوسته برای مقیاس‌های مثبت (قسمت تحلیلی یا خلاف جهت عقربه‌های ساعت) و صفحه دوم تبدیل موجک پیوسته برای مقیاس‌های منفی (قسمت ضد تحلیلی یا در جهت عقربه‌های ساعت) است. تابع  $CWT$  از نرمال‌سازی  $L1$  استفاده می‌کند. با نرمال‌سازی  $L1$ ، اگر مؤلفه‌های نوسانی با دامنه برابر در مقیاس‌های مختلف در داده‌ها وجود داشته باشد، اندازه آن‌ها در تبدیل موجک پیوسته برابر است [۳۲]. استفاده از نرمال‌سازی  $L1$  نمایش دقیق‌تری از سیگنال را نشان می‌دهد. نمونه‌ای از این تبدیل موجک‌ها را می‌توان در شکل ۴ مشاهده نمود.

### ۳.۳. پیش‌پردازش داده‌ها

بعد از بررسی مفهوم تبدیل موجک پیوسته به مسیر اصلی خود یعنی پیش‌پردازش داده‌های موجود در مجموعه داده می‌پردازیم. پیش‌پردازش داده‌ها شامل چهار مرحله است که عبارت‌اند از:

(۱) تحلیل تبدیل موجک پیوسته

خط پایه حذف‌شده و داده‌های اصلی سیگنال از طریق تبدیل موجک پیوسته به ضرایب موجک تبدیل‌شده و سپس به انرژی موجک (اسکالوگرام) تبدیل می‌شوند [۳۳]. در

جدول ۱. مشخصات گردآوری

epoch	12
optimizer	Adam
Learning rate	0.0001
Batch Size	12
Loss function	Categorical cross entropy

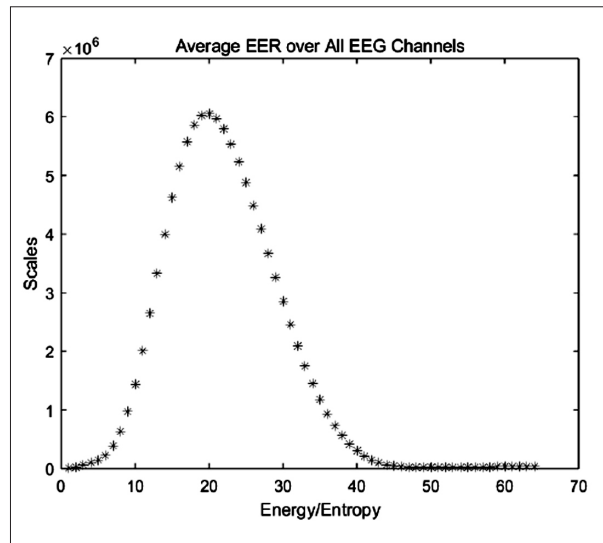
در آن برچسب یک به معنای احساسات خوب و برچسب صفر به معنای احساسات ناخوشایند است. در ادامه کار شبکه طراحی شده مطابق مشخصات موجود در جدول ۱ گردآوری<sup>۲۶</sup> می‌شود.

به‌طورکلی مراحل انجام این پروژه در نمودار شکل ۹ قابل مشاهده است که بر اساس آن ابتدا کتابخانه‌های موردنیاز فراخوانی می‌شود، سپس داده‌های موردنیاز در محیط کولب<sup>۲۷</sup> بارگذاری می‌شود. در گام بعد اقدام به نرمال‌سازی داده‌ها می‌شود، سپس یک بعد سوم برای داده‌ها در محیط کولب فراهم می‌شود. داده‌ها برچسب‌گذاری می‌شود و به دو بخش داده‌های آموزشی و آزمایشی تقسیم می‌شود که بر اساس آن ۹۰ درصد داده‌ها به داده‌های آموزشی و ۱۰ درصد به داده‌های آزمایشی اختصاص می‌یابد و در پایان طراحی معماری شبکه سه‌بعدی در خروجی داده‌های طبقه‌بندی شده را خواهیم داشت.

#### ۴. نتایج

حال با توجه به مشخصات فوق شبکه طراحی شده در ۱۲ اپیاک<sup>۲۸</sup> آموزش داده می‌شود و می‌توان نتایج را به‌وضوح در شکل ۱۰ مشاهده نمود. بر اساس مشاهدات صورت گرفته هنگام اجرای کد، در بخش آموزش داده‌ها دقت ۰/۹۸۷۲۹ و در قسمت آزمودن دقت ۰/۹۸۸۹ حاصل شد که دقت خوبی محسوب می‌شود.

همچنین در شکل ۱۱ می‌توان نمودار خطا در دو بخش آموزشی و آزمایشی را مشاهده نمود که بر آن اساس



شکل ۷: میانگین نسبت بهره‌وری انرژی بر روی همه کانال‌های الکتروانسفالوگرافی

روی همه کانال‌های الکتروانسفالوگرافی را مشاهده نمود [۳۴].

#### ۴) تکه سه‌بعدی<sup>۲۳</sup>

چندین قاب انتخاب می‌شود و به‌صورت تکه سه‌بعدی در کنار هم قرار داده می‌شود [۳۵]. این آخرین مرحله پیش‌پردازش داده‌ها است تا بتوان این داده‌ها را به‌عنوان ورودی به شبکه عصبی پیچشی سه‌بعدی طراحی شده اختصاص داد. پیش‌پردازش به‌طور کامل در محیط متلب انجام می‌شود.

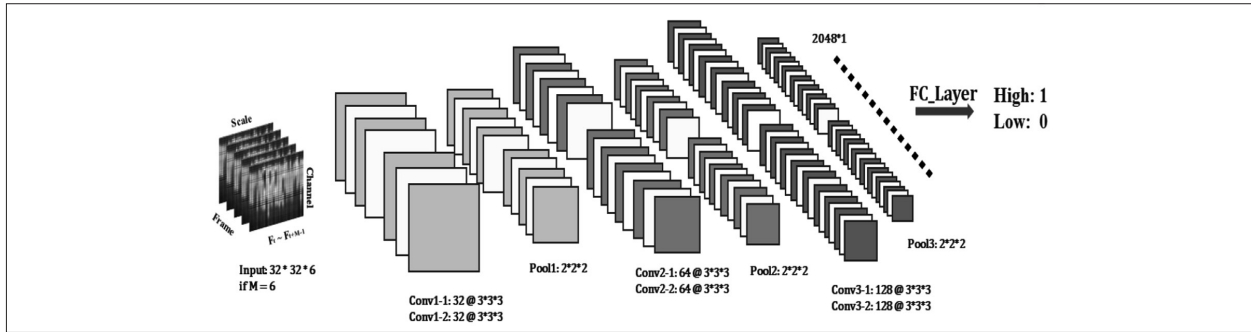
#### ۳.۴. شبکه عصبی پیچشی سه‌بعدی طراحی شده

در ادامه کار بایستی معماری شبکه تعریف شود. معماری تعریف شده مطابق شکل ۸ است که در آن از شش لایه پیچش؛ سه لایه حداکثر ادغام<sup>۲۴</sup> و یک تابع بیشینه هموار<sup>۲۵</sup> استفاده شده است.

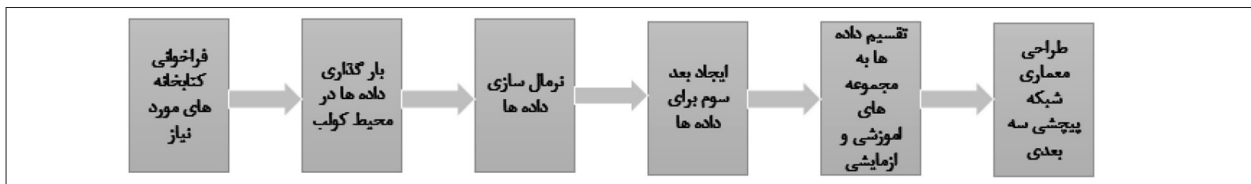
پس از استخراج مجموعه ویژگی‌ها توسط لایه‌های پیچش و انتخاب ویژگی‌های مهم توسط لایه حداکثر ادغام، ویژگی‌های نهایی جهت تعیین نوع احساسات وارد تابع بیشینه هموار می‌شوند که این تابع ویژگی‌های خام شبکه عصبی را به بردار احتمالات تبدیل می‌کند، که اساساً یک توزیع احتمال روی کلاس‌های ورودی دودویی است که

26- compile  
27- <https://github.com/5560405930/Detecting-emotions-using-the-EEG-signal>  
28- Epoch

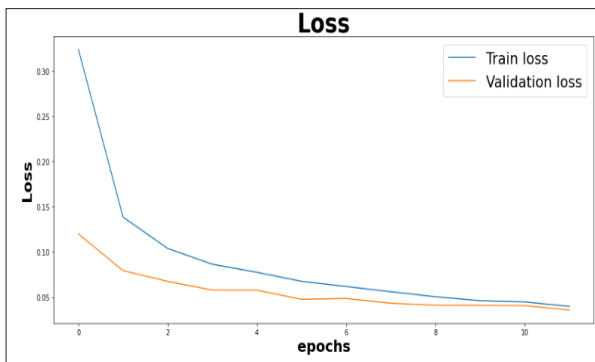
23- 3D chunk  
24- Max Pooling  
25- Soft Max



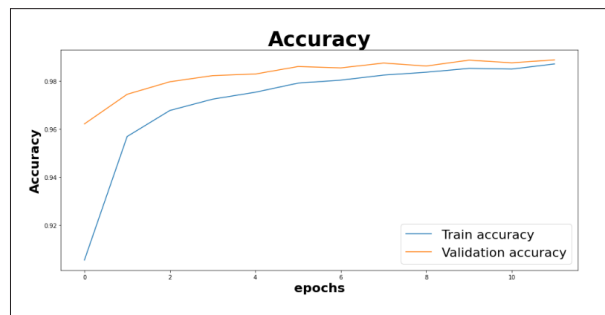
شکل ۸. معماری شبکه عصبی پیچشی سه بعدی



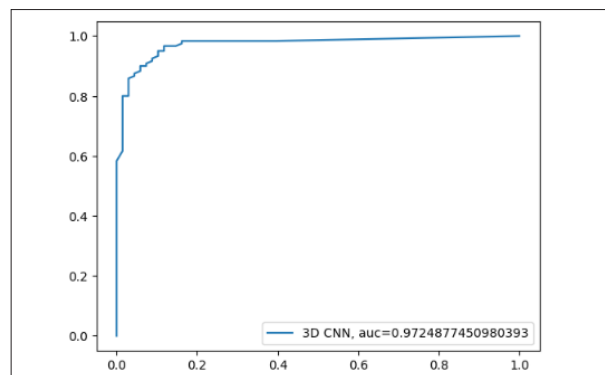
شکل ۹: مراحل انجام پروژه در محیط کولب



شکل ۱۱: خطای شبکه در آموزش و آزمایش



شکل ۱۰: دقت شبکه در آموزش و آزمایش



شکل ۱۲: نمودار راک شبکه عصبی پیچشی سه بعدی

آستانه‌های مختلف طبقه‌بندی ترسیم می‌کند و مقدار سطح زیر منحنی را نشان می‌دهد. از آنجایی که سطح زیر منحنی معیاری است که از احتمالات پیش‌بینی‌های رده استفاده می‌کند، می‌توان به مدل دارای سطح زیر منحنی بالاتر نسبت به مدلی با سطح زیر منحنی پایین‌تر، حتی اگر دقت مشابهی داشته باشد، اطمینان بیشتری داشت [۳۶]. بر اساس شکل ۱۲، مقدار سطح زیر منحنی الگوریتم طراحی شده برابر  $0.9724$  است. به‌طور کلی در مبحث سطح زیر منحنی،  $0.5$  تفاوتی را نشان نمی‌دهد، مقادیر بین  $0.7$  تا  $0.8$  قابل قبول، مقادیر بین  $0.8$  تا  $0.9$  خوب و مقدار بیش از  $0.9$  عالی در نظر گرفته می‌شود [۳۷]. به این ترتیب مقدار سطح زیر منحنی به دست آمده برابر با  $0.9724$  در این الگوریتم، عملکرد عالی را به ثبت رسانده است. در جدول ۲ برخی از بررسی‌های مشابه صورت

میزان خطا در داده‌های آموزشی به  $0.399$  و در داده‌های آزمایشی به  $0.395$  می‌رسد.

همچنین در شکل ۱۲ می‌توان منحنی راک<sup>۲۹</sup> مربوط به شبکه سه‌بعدی را مشاهده نمود که بر اساس آن مقدار سطح زیر منحنی<sup>۳۰</sup> شبکه طراحی شده برابر  $0.9724$  است. این نمودار نرخ مثبت واقعی به مثبت کاذب را در

29- ROC-Curve  
30- Area under curve (AUC)



جدول ۲: خلاصه‌ای از روش‌های استفاده‌شده برای تشخیص و طبقه‌بندی احساسات

مرجع	طبقه‌بندی	مجموعه داده	نتایج
Alhagry et al.[26]	یادگیری عمیق	مجموعه داده DEAP	۸۵/۶۵٪، ۸۵/۴۵٪ و ۸۷/۹۹٪ با برانگیختگی، ظرفیت و رده‌های موردعلاقه
Mehmood et al.[38]	روش یادگیری عمیق ترکیبی	IAPS	بالاترین متوسط دقت ۷۶/۶۲٪
Li et al.[39]	شبکه عصبی پیچشی سلسله مراتبی	SEED	بلندی موج بتا ۸۶/۲٪ و بلندی موج گاما ۸۸/۲٪
Bazgir et al.[40]	شبکه عصبی مصنوعی، ماشین بردار پشتیبان و k نزدیک‌ترین همسایه	مجموعه داده DEAP	دقت ۹۱/۳٪ برای برانگیختگی و ۹۱/۱٪ برای صحت
Putra et al.[41]	تجزیه موجک و k نزدیک‌ترین همسایه	مجموعه داده DEAP	دقت ۵۷/۵٪
Zangeneh Soroush et al.[42]	یادگیری ماشین	فیلم‌های ایرانی	تشخیص ۹۰٪
-Santamaria Granados et al.[43]	شبکه عصبی پیچشی عمیق	مجموعه داده AMIGOS	دقت ۹۹٪
George et al.[44]	ویژگی‌های آماری دامنه زمان - فرکانس	مجموعه داده DEAP	دقت ۹۲/۳۶٪
Chen et al.[45]	شبکه عصبی پیچشی	مجموعه داده DEAP	۵۸٪ در ظرفیت و ۳/۲۹٪ در برانگیختگی
روش پیشنهادی در این مقاله	شبکه عصبی پیچشی سه‌بعدی	مجموعه داده DEAP	دقت ۹۸/۸۹٪

از داده‌ها استخراج می‌کند که به‌نوبه خود موجب کاهش خطای شبکه می‌شود.

#### مراجع

- [1] Naji M, Firoozabadi M, Azadfallah P, Emotion classification during music listening from forehead biosignals, SIViP, 9, 2013, pp. 1365–1375.
- [2] Diener E, Chan MY, Happy people live longer: Subjective well-being contributes to health and longevity, Appl. Psychol. Health Well Being, 3, 2011, pp. 1–43.
- [3] Alakuş TB, Türkoğlu İ, Determination of effective EEG channels for discrimination of positive and negative emotions with wavelet decomposition and support vector machines, Int. J. Inform. Technol, 12, 2019, pp. 229–237.
- [4] Turnip A, Simbolon AI, Amri MF, Sihombing P, Setiadi RH, Mulyana E, Backpropagation neural networks training for EEG-SSVEP classification of emotion recognition, Internetw. Indones, 9, 2017, pp. 53–57.
- [5] Hu Y, Dong J, Zhou S, Audio-textual emotion recognition based on improved neural networks, Math. Probl. Eng. 2019, 2019, pp. 1–9.
- [6] Lasri I, Solh AR, Belkacemi ME, Facial emotion recognition of students using convolutional neural network, In Proceedings of the 3rd International Conference on Intelligent Computing in Data Sciences, Marrakech, Morocco, 29–30 October 2019.
- [7] Garber-Barron M, Si M, Using body movement and posture for emotion detection in non-acted scenarios, In Proceedings of the IEEE International Conference on Fuzzy Systems, Brisbane, QLD, Australia, 10–15 June 2012.
- [8] Sassenrath C, Sassenberg K, Ray DG, Scheiter K, Jarodzka H, A mo-

گرفته، جمع‌آوری شده است که در آن‌ها از روش‌های مختلفی استفاده شده و نتایج متفاوتی حاصل شده است. در این جدول مجموعه داده مورداستفاده نویسندگان، روش پیشنهادی به‌کاررفته در این مقاله و نتایج حاصل نیز قید شده است. بر اساس جدول ۲ روش پیشنهادی در این مقاله، به دقت بهتری نسبت به مطالعات قبلی دست پیدا کرده است.

#### ۵. جمع‌بندی

در این مقاله سیگنال‌های الکتروانسفالوگرافی برای تشخیص حالات انسانی معرفی شد. همچنین، از شبکه‌های عصبی پیچشی سه‌بعدی برای طبقه‌بندی احساسات صورت که بر اساس سیگنال‌های الکتروانسفالوگرافی به‌دست آمده بودند استفاده شد. مطابق آنچه در جدول ۲ نشان داده شد، برخی از روش‌های مطالعه شده به دلیل محدودیت در استخراج ویژگی عملکرد ضعیفی دارند و قادر به تشخیص احساسات با دقت بالا نیستند. اما شبکه طراحی‌شده در این مقاله از دقت و صحت بالایی برخوردار بود و همان‌طور که بررسی شد ویژگی‌های زیادی را

learning techniques, *Sensors*, 19(7), 2019, 1738.

[31] Lishani AO, Person recognition using gait energy imaging, University of Northumbria at Newcastle, United Kingdom, 2018.

[32] Gurley K, Kareem A, Applications of wavelet transforms in earthquake, wind and ocean engineering, *Engineering structures*, 21(2), 2019, pp. 149-167.

[33] Yanik H, Değirmenci E, Büyükkakılı B, Karpuz D, Kılınc OH, Gürgül S, Electrocardiography (ECG) analysis and a new feature extraction method using wavelet transform with scalogram analysis. *Biomedical Engineering/Biomedizinische Technik*, 65(5), 2020, pp. 543-556.

[34] Yao G, Lei T, Zhong J, Jiang P, Learning multi-temporal-scale deep information for action recognition, *Applied Intelligence*, 49(6), 2019, pp. 2017-2029.

[35] Bolles RC, Horaud P, 3DPO: A three-dimensional part orientation system, In *Three-dimensional machine*, 1987.

[36] Bülow E, Hahn U, Andersen IT, Rolfson O, Pedersen AB, Hailer NP, Prediction of early periprosthetic joint infection after total hip arthroplasty. *Clinical Epidemiology*, 2022, pp. 239-253.

[37] Mamun M, Farjana A, Al Mamun M, Ahammed MS, Lung cancer prediction model using ensemble learning techniques and a systematic review analysis, In *2022 IEEE World AI IoT Congress (AIoT)*. IEEE, 2022, pp. 187-193.

[38] Mehmood RM, Du R, Lee HJ, Optimal Feature Selection and Deep Learning Ensembles Method for Emotion Recognition From Human Brain EEG Sensors, *IEEE Access*, 5, 2017, pp. 14797-14806.

[39] Li J, Zhang Z, He H, Hierarchical Convolutional Neural Networks for EEG-Based Emotion Recognition, *Cogn. Comput.*, 2017, 10, pp. 368-380.

[40] Bazgir O, Mohammadi Z, Habibi SAH, Emotion Recognition with Machine Learning Using EEG Signals. In *Proceedings of the 2018 25th National and 3rd International Iranian Conference on Biomedical Engineering (ICBME)*, Qom, Iran, 2019, pp. 29-30.

[41] Putra AE, Atmaji C, Ghaleb F, EEG-Based Emotion Classification Using Wavelet Decomposition and K-Nearest Neighbor. In *Proceedings of the 2018 4th International Conference on Science and Technology (ICST)*, Yogyakarta, Indonesia, 2018, pp. 1-4.

[42] Zangeneh SM, Maghooli K, Setarehdan SK, Nasrabadi AM, Emotion Classification through Nonlinear EEG Analysis Using Machine Learning Methods. *Int. Clin. Neurosci. J.* 5, 2018, pp. 135-149.

[43] Santamaria-Granados L, Muñoz-Organero M, Ramírez-González G, Abdulhay E, Arunkumar N, Using Deep Convolutional Neural Network for Emotion Detection on a Physiological Signals Dataset (AMI-GOS). *IEEE Access*, 7, 2018, pp. 57-67.

[44] George FP, Mannafee I, Hossain P, Parvez MZ, Uddin J, Recognition of emotional states using EEG signals based on time-frequency analysis and SVM classifier. *Int. J. Electr. Comput. Eng. (IJECE)*, 9, 2019, 1012.

[45] Chen JX, Zhang PW, Mao ZJ, Huang YF, Jiang, DM, Zhang YN, Accurate EEG-Based Emotion Recognition on Combined Features Using Deep Convolutional Neural Networks, *IEEE Access*, 7, 2019, pp. 44317-44328.

titational determinant of facial emotion recognition: Regulatory focus affects recognition of emotions in faces. *PLoS ONE*, 9, 2014, e112383.

[9] Wioleta S, Using physiological signals for emotion recognition. In *Proceedings of the 6th Conference on Human System Interactions*, Sopot, Poland, 6–8 June 2013

[10] Yan J, Chen S, A EEG-based emotion recognition model with rhythm and time characteristics. *Brain Inf.* 6, 2019, 7.

[11] Casson AJ, Yates DC, Smith SJM, Duncan JS, Rodriguez-Villegas, E. Wearable electroencephalography, *IEEE Eng. Med. Biol. Mag.* 2010, 29, pp. 44–56.

[12] Bashivan P, Rish I, Heisig S, Mental state recognition via wearable EEG. *arXiv* 2016, arXiv:1602.00985, 2016.

[13] Hurlings R, Datcu D, Rothkrantz LJM, Emotion recognition using brain activity. In *Proceedings of the 9th International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing*, Gabrovo, Bulgaria, 12–13 June 2008.

[14] Gu S, Wang F, Patel NP, Bourgeois, J.A.; Huang, J.H. A model for basic emotions using observations of behavior in *drosophila*. *Front. Psychol.* 10, 2019, pp. 1–13.

[15] Hania WMB, Lachiri Z, Emotion classification in arousal-valence dimension using discrete affective keywords tagging, In *Proceedings of the International Conference on Engineering & MIS*, Monastir, Tunisia, 8–10 May 2017.

[16] Stickel C, Ebner M, Steinbach-Nordmann S, Searle G, Holzinger A, Emotion detection: Application of the valence arousal space for rapid biological usability testing to enhance universal access, In *Proceedings of the International Conference on Universal Access in Human-Computer Interaction*, San Diego, CA, USA, 19–24 July 2009.

[17] Bradley MM, Lang PJ, International affective digitized sounds (IADS): Stimuli, instruction manual and affective ratings. In *Technical Report, No: B-2; The Center for Research in Psychophysiology*, University of Florida: Gainesville, FL, USA, 1999.

[18] Raval D, Sakle M, A literature review on emotion recognition system using various facial expression, *IJARIE*, 1, 2015, pp. 326–329.

[19] Alakuş TB, Türkoğlu İ, Database for an emotion recognition system based on EEG signals and various computer games—GAMEEMO, *Biomed. Signal Process. Control.* 60, 2020, 101951.

[20] Yang P, Wang D, KAGN ZJ, Li T, FU LH, YU, YR, Prediction model of paroxysmal atrial fibrillation based on pattern recognition and ensemble CNN-LSTM, *Journal of Zhejiang University (Engineering Science)*, 54(5), 2020, pp. 1039-1048.

[21] Zhang Y, Ishibuchi H, Wang S, Deep Takagi–Sugeno–Kang fuzzy classifier with shared linguistic fuzzy rules, *IEEE Transactions on Fuzzy Systems*, 26(3), 2017, pp. 1535-1549.

[22] Murugappan M, Ramachandran N, Sazali Y, Classification of human emotion from EEG using discrete wavelet transform, *Journal of biomedical science and engineering*, 3(04), 2010, p. 390.

[23] Paul S, Mazumder A, Ghosh P, Tibarewala DN, Vimalarani G, EEG based emotion recognition system using MFDDFA as feature extractor, In *2015 International Conference on Robotics, Automation, Control and Embedded Systems (RACE)*. IEEE, 2015, pp. 1-5.

[24] Jiang Y, Gu X, Wu D, Hang W, Xue J, Qiu S, Lin CT, A novel negative-transfer-resistant fuzzy clustering model with a shared cross-domain transfer latent space and its application to brain CT image segmentation, *IEEE/ACM transactions on computational biology and bioinformatics*, 18(1), 2020, pp. 40-52.

[25] Jiang Y, Chung FL, Wang S, Deng Z, Wang J, Qian P, Collaborative fuzzy clustering from multiple weighted views, *IEEE transactions on cybernetics*, 45(4), 2014, pp. 688-701.

[26] Alhagry S, Fahmy AA, El-Khoribi RA, Emotion recognition based on EEG using LSTM recurrent neural network, *International Journal of Advanced Computer Science and Applications*, 8(10), 2017, pp. 355-358.

[27] Schirrmeyer RT, Springenberg JT, Fiederer LDJ, Glasstetter M, Eggenberger K, Tangermann M, Ball T, Deep learning with convolutional neural networks for EEG decoding and visualization, *Human brain mapping*, 38(11), 2017, pp. 5391-5420.

[28] Zhang J, Li S, Yin Z, Pattern classification of instantaneous mental workload using ensemble of convolutional neural networks, *IFAC-PapersOnLine*, 50(1), 2017, pp. 14896-14901.

[29] Chen J, Siu SW, Machine learning approaches for quality assessment of protein structures, *Biomolecules*, 10(4), 2020, 626.

[30] Bălan O, Moise G, Moldoveanu A, Leordeanu M, Moldoveanu F, Fear level classification based on emotional dimensions and machine