

## Emotion Recognition Based on Improved VGG-16 Network

Sara Motamed <sup>\*1</sup>

Department of Computer Engineering, FSh.C., Islamic Azad University, Fouman, Iran

### ABSTRACT

#### RESEARCH PAPER

**Received:**

23 March 2025

**Accepted:**

27 June 2025

**Keywords:**

Emotion Recognition,

Electroencephalogram Signals,

Transfer Learning,

VGG-16 Network,

Fuzzy Layer.

**Background and Objectives:** In recent years, the use of deep learning methods to analyze electroencephalogram and physiological signals for emotion recognition has attracted the attention of many researchers and scholars. However, the lack of expertise in extracting features from electroencephalogram signals and classifying emotions based on these signals is considered one of the main challenges of this discussion.

**Methods:** In this study, signal-to-image conversion will be used to prepare input data for a pre-trained network. The innovation of the research is in how to prepare the input data to the pre-trained network. The innovation of the proposed method is also the conversion of signal to image as input features to the learning model in the recorded data for emotion recognition. Also, a hybrid architecture based on VGG-16 network and fuzzy layer is used to optimally solve the problem and increase the emotion recognition rate.

**Findings:** Given the problems and limitations of data collection and the sensitivity of electroencephalogram data to environmental noise and similar problems, recording electroencephalogram signals in an isolated environment is required. Considering the above, it is better to use a dataset whose accuracy has been previously confirmed by researchers and scholars. In this article, the public DEAP dataset was used for research and hypothesis testing.

**Conclusion:** By examining the results of implementing the proposed method, the emotion recognition rate shows an accuracy of 90.89 percent.

<sup>1</sup> **Corresponding author:**



[motamed.sarah@gmail.com](mailto:motamed.sarah@gmail.com)

Copyright © Author(s).



## نشریه تخصصی آرمان پردازش، دوره ۶، شماره ۲، سال ۱۴۰۴

فصلنامه تخصصی آرمان پردازش  
(APJ)Homepage: [www.armanprocessjournal.ir](http://www.armanprocessjournal.ir)

شاپای الکترونیکی: ۴۵۴۹-۲۷۸۳



## تشخیص احساسات مبتنی بر شبکه VGG-16 بهبود یافته

سارا معتمد<sup>۱،\*</sup><sup>۱</sup> گروه مهندسی کامپیوتر، واحد فومن و شفت، دانشگاه آزاد اسلامی، فومن، ایران

## چکیده

## مقاله پژوهشی

**پیشینه و اهداف:** در سال‌های اخیر، استفاده از روش‌های یادگیری عمیق برای تجزیه و تحلیل سیگنال‌های الکتروانسفالوگرام و فیزیولوژیکی برای شناخت احساسات توجه بسیاری از پژوهشگران و محققان را به خود جلب کرده است. با این حال، کمبود تخصص در استخراج ویژگی از سیگنال‌های الکتروانسفالوگرام و دسته‌بندی احساسات بر اساس این سیگنال‌ها یکی از چالش‌های اساسی این بحث محسوب می‌شود.

تاریخ دریافت مقاله:

۱۴۰۳/۱/۳

**روش‌ها:** در این پژوهش جهت آماده‌سازی داده‌های ورودی به شبکه‌ی از پیش آموزش دیده، از تبدیل سیگنال به تصویر استفاده خواهد شد. نوآوری پژوهش در چگونگی آماده‌سازی داده‌های ورودی به شبکه‌ی از پیش آموزش دیده است. نوآوری روش پیشنهادی نیز تبدیل سیگنال به تصویر به عنوان ویژگی‌های ورودی به مدل یادگیری در داده‌های ثبت شده است که جهت تشخیص احساسات است. همچنین از یک معماری ترکیبی مبتنی بر شبکه VGG-16 و لایه فازی، برای حل بهینه مسئله و بالابردن نرخ تشخیص احساسات استفاده می‌شود.

تاریخ پذیرش:

۱۴۰۴/۴/۷

**یافته‌ها:** با توجه به مشکلات و محدودیت‌های جمع‌آوری داده و حساسیت داده‌های الکتروانسفالوگرام به نویز محیط و مشکلاتی از این قبیل باعث می‌شود ثبت سیگنال‌های الکتروانسفالوگرام در محیطی ایزوله انجام شود. با توجه به موارد ذکر شده، بهتر است از مجموعه داده‌ای استفاده شود که صحت داده‌های ثبت شده‌ی آن پیش‌تر توسط پژوهشگران و محققان تایید شده است. در این مقاله از مجموعه داده‌ی عمومی DEAP برای پژوهش و بررسی فرضیات استفاده شده است.

واژگان کلیدی:

شناسایی احساسات،  
سیگنال‌های الکتروانسفالوگرام،  
یادگیری انتقالی،

**نتیجه‌گیری:** با بررسی نتایج حاصل از پیاده‌سازی روش پیشنهادی، نرخ تشخیص احساسات دقت ۸۹/۹۰ درصد را نشان می‌دهد.

شبکه VGG-16، لایه فازی.

Copyright © Author(s).



## ۱- مقدمه

شناسایی احساسات از روی فعالیت و عملکرد مغز توسط الگوریتم‌های یادگیری ماشین و یادگیری عمیق از مسائل مهم چند سال اخیر است. یکی از مهمترین مزایای سیگنال‌های مغزی در مقایسه با سیگنال‌های دیداری یا گفتاری در زمینه تشخیص احساسات، توانایی آن در عدم مخفی کردن احساسات توسط افراد با تغییر حالت چهره یا تغییر لحن صدا هنگام صحبت کردن است. الکتروانسفالوگرام اندازه‌گیری سیگنال‌های الکتریکی است که نتیجه‌ی فعالیت‌های مغز را نشان می‌دهد. این اندازه‌گیری توسط اختلاف ولتاژ بین الکتروود واقعی و الکتروود مرجع صورت می‌گیرد. روش‌های مختلفی برای تشخیص احساسات مانند تجزیه و تحلیل حالات چهره، زبان بدن، گفتار و حرکات چشم وجود دارد که شاخص‌های فیزیکی هستند. برای علائم فیزیولوژیکی می‌توانیم از سیگنال‌های EEG، سیگنال‌های EMG، GSR و سایر داده‌های بیومتریک استفاده کنیم که نیاز به ابزارهای تخصصی برای ضبط دارند. برای ثبت سیگنال‌های EEG باید از تعدادی الکتروود که روی پوست سر قرار می‌گیرند استفاده کنیم. این سیگنال‌ها پاسخ‌های مغز در برابر محرک‌ها، افکار و احساسات را به عنوان یک سیگنال منعکس می‌کنند. این سیگنال‌ها برای تشخیص و بررسی اختلالات مغزی و بیماری‌های عصبی استفاده می‌شود. یکی از بزرگترین مزایای سیگنال‌های EEG توانایی آنها در ارائه نتایج دقیق است. یکی دیگر از مزیت‌های EEG در مقایسه با روش‌های دیگر، غیر تهاجمی بودن آن است و امکان جمع‌آوری اطلاعات در زمان واقعی را فراهم می‌کند. برخلاف تکنیک‌های تصویربرداری مانند fMRI و سی‌تی‌اسکن که به تجهیزات سنگین و گران قیمت نیاز دارند، EEG سبک وزن است و می‌توان از آن در تنظیمات مختلف استفاده کرد. علاوه بر این، ضبط سیگنال‌های EEG بسیار سریع و دقیق است. این سیگنال‌ها حتی قبل از اینکه انسان از تغییرات احساسات خود آگاه شود نشان داده و ثبت می‌شود.

تشخیص احساسات با استفاده از یادگیری ماشین، یادگیری عمیق و سیگنال‌های EEG پاسخ‌های سریع و بی‌درنگ به داده‌ها و واکنش‌های احساسی را ممکن می‌سازد. این فناوری را می‌توان در خدمات پزشکی، خدمات مشتری، امنیت، روانشناسی و علوم اعصاب به کار برد. در [۱]، به طراحی یک حالت شبکه عصبی پیچشی عمیق برای تشخیص پنج نوع مختلف از احساسات پرداخته است. دقت این روش ۷۸/۰۴ درصد گزارش شده است. در [۲] یک روش یادگیری عمیق برای تشخیص هفت احساس مختلف خشم، نفرت، خنثی، ترس، شادی، غم و تعجب پیشنهاد شد. داده‌های مورد استفاده در این مطالعه تصاویر چهره بودند. مدل پیشنهادی در این مقاله شبکه‌های عصبی پیچشی بود که روی دو مجموعه داده FERC-2013 و JAFFE بررسی شد. نتایج ارائه شده به ترتیب دقت ۷۰/۱۴٪ را در FERC-2013 و ۹۸/۶۵٪ را در مجموعه

داده JAFFE نشان داد. در [۳]، مدلی ساده و مؤثر برای تشخیص احساسات گفتار گزارش شد. این مدل از ویژگی فرکانس Mel به عنوان ورودی شبکه پرسپترون استفاده کرد. مدل پیشنهادی عملکرد بالایی را در دستگاه‌هایی با حافظه محدود نشان دادند. در [۴]، تشخیص احساسات گفتار، به ویژه حالت هیجانی ترس، مورد بررسی قرار گرفت. روش کار بدینصورت بود که ابتدا ویژگی‌های فرازبانی از داده‌های صوتی خام استخراج شده، سپس طبقه‌بندی احساسات و تولید سیگنال‌های هشدار دهنده مطرح شد. در [۵]، روشی را برای تشخیص احساسات با استفاده از داده‌های متنی، صوتی و تصویری ارائه کردند. مدل پیشنهادی ویژگی‌های چند لایه را با شبکه‌های پیچشی سبک وزن در مجموعه داده‌های FER-2013 و AffectNet ترکیب کرده و دقت ۷۴/۳۵٪ و ۵۶/۶۷٪ را گزارش کردند که این نتایج تعادل خوبی بین دقت و سرعت مدل را نشان می‌دادند.

در [۶]، دقت تشخیص احساسات با استفاده از سیگنال‌های EEG توسط یادگیری عمیق پیچشی مورد بررسی قرار گرفت. در [۷] مدل پیچشی برای تشخیص احساسات بر اساس سیگنال‌های EEG پیشنهاد شد و عملکرد آن را با استفاده از روش‌های ارزیابی مختلف محاسبه کردند. در [۸] نظمی و همکاران، با استفاده از SVM و KNN به تشخیص احساسات با استفاده از سیگنال‌های EEG پرداختند. نتیجه این روش‌ها دقت ۹۷/۳۳ درصد برای SVM و ۹۸/۳۷ درصد برای KNN بود. در [۹] بمنظور استخراج ویژگی از ویژگی‌های انرژی، آنالیز و تبدیل موجک استفاده شده است. همچنین برای دسته‌بندی داده‌ها KNN، SVM و ANN بکار برده شد. در [۱۰] یک مدل پیچشی با هسته‌های چند مقیاسی برای تشخیص احساسات پیشنهاد شد. مدل معرفی شده به دقت طبقه‌بندی باینری ۹۸/۲۷ درصد برای برانگیختگی و ۹۸/۳۶ درصد برای ظرفیت دست یافت، که مشخص می‌کند استفاده از همبستگی‌های کانال و باندهای فرکانسی به طور مؤثر تشخیص احساسات را بهبود می‌بخشد. در [۱۱]، مدل‌های پیچشی دوبعدی و داده‌های چندوجهی بمنظور افزایش دقت طبقه‌بندی احساسات ارائه شد. در [۱۲] یک رویکرد جدید برای تشخیص بهتر احساسات با استفاده از سیگنال‌های EEG پیشنهاد شد. این شبکه ترکیبی متشکل از یک شبکه عصبی کانولوشن همراه با مدل‌های پنهان مارکوف بود که در آن طبقه‌بندی کننده‌های مدل، جایگزین طبقه‌بندی کننده‌های عصبی برای تشخیص احساسات می‌شدند. نتایج تجربی اثربخشی رویکردهای پیشنهادی ۷۹/۷۷ درصد دقت را برای برانگیختگی، ۸۳/۰۹ درصد برای ظرفیت و ۸۱/۸۳ درصد برای تسلط گزارش نمود. در [۱۳] مدلی بر اساس یک شبکه پیچشی سه بعدی معرفی شد. این مدل مبتنی بر نمایش مکانی و زمانی سیگنال‌های EEG برای تسهیل استخراج ویژگی‌های دقیق‌تر و مؤثرتر در تشخیص احساسات بود. نتایج

پژوهش [۱۴]، از شبکه‌ی مولد هموردی شرطی<sup>۱</sup> (cWGAN) برای تولید داده‌های الکتروانسفالوگرام به منظور افزایش مجموعه داده استفاده شده است. در این پژوهش کیفیت داده‌های تولید شده با استفاده از سه معیار مورد بررسی قرار گرفته و داده‌هایی با کیفیت بالا به داده‌ی اصلی اضافه شده است. در پژوهش [۱۵]، سه روش برای افزایش داده‌های الکتروانسفالوگرام جهت افزایش عملکرد مدل‌های تشخیص احساسات پیشنهاد شده است. روش‌های پیشنهادی بر اساس دو مدل مولد ژرف؛ رمزگذار خودکار متنوعی<sup>۲</sup> و شبکه‌ی مولد هموردی و دو روش افزایش داده است. این سه روش، شبکه‌ی مولد هموردی شرطی، رمزگذار خودکار متنوعی انتخابی و شبکه‌ی مولد هموردی انتخابی نامیده می‌شوند. در پژوهش [۱۶]، بمنظور تشخیص احساسات از سه مجموع داده‌ی DEAP، DREAMER و یک مجموع داده‌ی اختصاصی استفاده شده است. همچنین آنها این سه مجموعه داده را روی مدل پیشنهادیشان ترکیب کرده و به نتایج قابل توجهی در تشخیص احساسات دست یافتند. در پژوهش [۱۷] از مجموعه داده‌ی DEAP و مدل TRFE-LSSVM-SG برای دسته‌بندی استفاده شده است. ویژگی‌های استخراج شده در بُعد فرکانس در مجموع ۲۱۶ ویژگی است که توان سیگنال را در فرکانس‌های تتا، آلفای کوچک، آلفا، بتا و گاما برای همه‌ی ۳۲ کانال محاسبه کرده‌اند. همچنین، تفاضل توان سیگنال برای جفت کانال‌های متقارن نیز محاسبه شده است. ویژگی‌های استخراج شده در بُعد زمان در مجموع ۲۲۴ است که در کل ۴۴۰ ویژگی در بُعد زمان و فرکانس از ۳۲ کانال استخراج شده است. در [۱۸]، از یک روش یادگیری با نظارت برای تشخیص احساسات از روی مجموعه داده‌ی DEAP استفاده شده است. آن‌ها از روش M-TRFE برای انتخاب ویژگی به منظور کاهش زمان اجرای TRFE و از LSSVM برای دسته‌بندی و از مدل دو بعدی نمودار برانگیختگی و ظرفیت برای تشخیص احساسات استفاده کرده‌اند. در پژوهش [۱۹]، از شبکه‌ی از پیش آموزش دیده‌ی<sup>۳</sup> عصبی پیچشی برای دسته‌بندی داده‌های الکتروانسفالوگرام خام و بدون استخراج ویژگی از مجموعه داده‌ی DEAP، SEED و LUMBED استفاده کرده‌اند. نتایج حاصل از این پژوهش نشان داد که دقت دسته‌بندی برای افراد مختلف متفاوت بوده و همچنین برای مجموعه داده‌های متفاوت، مقادیر متفاوتی به دست آمده است. متوسط دقت دسته‌بندی برای مجموعه داده‌ی SEED برای دو کلاس و سه کلاس به ترتیب ۸۶/۵۶ و ۷۸/۳۴ درصد و دقت متوسط برای مجموعه داده‌ی DEAP دو کلاس ۷۲/۸۱ درصد و برای مجموعه داده‌ی LUMBED دو کلاس ۸۱/۱ درصد بوده است. در پژوهش [۲۰]، مدل رده‌بند گروه چند لایه همجوشی مبتنی بر رمزگذار خودکار انباشته برای شناخت

گزارش شده نشان داد که این رویکرد به دقت بالایی برای طبقه‌بندی‌های مختلف دست یافته است. توجه به این نکته مهم است که ترکیب کانولوشن سه بعدی پیچیدگی مدل را افزایش می‌دهد، به همین دلیل در این مقاله از ترکیبی VGG-16 با لایه فازی برای تشخیص احساسات صوتی استفاده شده است.

در این مقاله، از مجموعه داده DEAP استفاده شده و نوآوری مدل پیشنهادی در این مقاله نیز در چگونگی آماده‌سازی داده‌های ورودی به شبکه‌ی از پیش آموزش دیده است. از مزایا و اهداف مدل پیشنهادی می‌توان به موارد زیر اشاره کرد:

- استفاده از VGG-16 برای استخراج ویژگی‌های قوی از سیگنال‌های

#### EEG

- تبدیل صدا به تصویر باعث می‌شود بتوانیم از قدرت مدل‌های بینایی استفاده کنیم.
- VGG-16 که روی ImageNet آموزش دیده، به‌خوبی ویژگی‌های الگوهای پیچیده (مثل طیف فرکانسی صدا) را استخراج می‌کند.
- در این مدل افزایش دقت و پایداری نسبت به شبکه‌های ساده‌تر مثل CNN معمولی وجود دارد.
- افزودن لایه فازی جهت تفسیرپذیری و انعطاف‌پذیری بالا
- لایه فازی مدل را به سمت استدلال انسانی‌تر و قابل تفسیر هدایت می‌کند.
- در شرایطی که داده‌ها نویزدار، غیرقطعی یا مبهم باشند، لایه فازی کمک می‌کند تصمیم‌گیری نرم‌تر و هوشمندانه‌تر باشد.
- قابلیت فازی در یادگیری قوانین ترکیبی بین ویژگی‌ها باعث افزایش دقت در داده‌های خاکستری می‌شود.
- توانایی تطبیق با داده‌های صوتی واقعی و پیچیده
- ترکیب VGG (برای یادگیری خودکار ویژگی‌های پیچیده) با منطق فازی (برای تنظیم پاسخ‌های خروجی) مدل شناسایی هیجان را برای داده‌های طبیعی، گفتار واقعی، لهجه‌دار، هیجانی و غیره قوی و مناسب‌تری نماید.

#### ۲- کارهای مرتبط

مشکل کمبود داده الکتروانسفالوگرافی به خصوص در تشخیص احساسات منجر به ایجاد مشکل در ساخت یک مدل احساسی با دقت بالا با استفاده از الگوریتم‌های یادگیری ماشین و شبکه‌های عصبی ژرف می‌شود. در پژوهش‌های [۱۴] و [۱۵] که هر دو حاصل تحقیقات مربوط به یک گروه پژوهشی در سال‌های ۲۰۱۸ و ۲۰۲۰ است، برای حل مشکل کمبود داده آموزشی از شبکه‌های عصبی مولد به منظور تولید داده‌های الکتروانسفالوگرام برای بهبود عملکرد مدل استفاده شده است. در

<sup>۱</sup> Variational Autoencoder (VAE)  
Pretrained network<sup>۲</sup>

<sup>۱</sup> Conditional Wasserstein Generative Adversarial Network

داده‌های DEAP مورد ارزیابی قرار گرفت. با توجه به بررسی‌های صورت گرفته، تمامی مدل‌های ارائه شده دارای یکسری معایب و مزایا هستند. در این مقاله هدف حل چالش‌های موجود در مدل‌های بررسی، روی مجموعه داده‌ی عمومی DEAP می‌باشد.

### ۳- مجموعه داده

در این پژوهش برای قابلیت اطمینان از صحت و کیفیت داده‌ها، از مجموعه داده‌ی عمومی DEAP استفاده شده است. از بین مجموعه داده‌های موجود، مجموعه داده‌ی عمومی DEAP بیشترین تعداد شرکت‌کننده‌ها را دارد. همچنین علاوه بر داده‌های ثبت شده الکتروانسفالوگرام، ۸ کانال داده فیزیولوژیکی ثبت شده نیز در این مجموعه داده وجود دارد. تعداد افرادی که در این دیتاست مشارکت دارند ۳۲ نفر (۱۶ مرد و ۱۶ زن در بازه سنی ۱۷ تا ۳۲) و از هر فرد ۴۰ فیلم گرفته شده است. این مجموعه داده توسط گروهی از متخصصان گردآوری شده و محققان و پژوهشگران بسیاری از آن در تحقیقات خود استفاده کرده‌اند [۲۵].

#### ۳-۱- آموزش مدل پیشنهادی

از میان شبکه‌های عمیق موجود، شبکه‌ی از پیش آموزش دیده VGG16 به دلیل کم بودن داده‌های آموزش انتخاب شده است. برای کاهش ابعاد داده به کمک اسپکتروگرام، داده‌های الکتروانسفالوگرام با ابعاد بالا به تصاویر با ابعاد کمتر تبدیل شده‌اند [۲۶]. روند تشخیص احساسات با تبدیل سیگنال به تصویر و استفاده از روش یادگیری انتقالی در شکل ۱ نمایش داده شده است.

احساسات ارائه شد. آن‌ها ویژگی‌های استخراج شده از سیگنال‌های الکتروانسفالوگرام ویژگی‌های توان سیگنال از کانال‌های مختلف، ویژگی‌های اختلاف توان بین کانال‌های متقارن، ویژگی‌های بعد زمان و فرکانس سیگنال‌های فیزیولوژیکی از جمله: الکترواوکولوگرافی<sup>۱</sup>، الکترومیوگرافی<sup>۲</sup>، دمای پوست<sup>۳</sup>، پاسخ پوست گالوانیک<sup>۴</sup>، فشار خون<sup>۵</sup> و تنفس<sup>۶</sup> به عنوان ورودی مدل برای پیش‌بینی پارامترهای برانگیختگی و ظرفیت، برای تشخیص احساسات از روی مدل دو بعدی برانگیختگی-ظرفیت داده استفاده کرده و به نتایج قابل توجهی دست یافتند. در پژوهش [۲۱] از مجموعه داده‌ی DEAP و شبکه‌ی عمیق حافظه طولانی-کوتاه مدت برای استخراج ویژگی استفاده کردند. مدل پیشنهادی آنها از یک لایه‌ی Dense، برای دسته‌بندی پارامترهای برانگیختگی و ظرفیت استفاده شده بود. دقت دسته‌بندی داده‌های آزمایش برای این پارامترها به ترتیب ۸۵/۶۵، ۸۵/۴۵ و ۸۷/۹۹ درصد تخمین زده شد. در پژوهش [۲۲] از شبکه‌ی عصبی پیچشی سه بعدی و مجموعه داده‌ی DEAP برای تشخیص احساسات از روی سیگنال‌های الکتروانسفالوگرام چند کاناله استفاده کرده و برای بهبود کارایی مدل پیشنهادی، یک فاز افزایش داده را در نظر گرفت. در پژوهش [۲۳]، شبکه‌ی عصبی عمیق برای تشخیص احساسات از سیگنال‌های الکتروانسفالوگرام با استفاده از مجموعه داده‌ی DEAP پیشنهاد شد. دو نوع معماری شبکه‌ی عصبی در این پژوهش مورد بررسی قرار گرفت: شبکه‌های عصبی پیچشی و شبکه‌ی عصبی عمیق بازگشتی. هر دو مدل هنگام آموزش روی داده‌های از پیش پردازش شده، در دسته‌بندی احساسات کاربر بسیار مؤثر عمل کرد. در پژوهش [۲۴] از مدل GELMY برای شناسایی الگوهای پایدار در طول زمان و ارزیابی پایداری مدل شناسایی احساسات استفاده شد. در این پژوهش انتخاب ویژگی‌ها و دسته‌بندی الگوی احساسات در مجموعه داده‌های SEED و مجموعه

<sup>۱</sup> Blood Volume pressure<sup>۱</sup>

<sup>۲</sup> Respiration<sup>۲</sup>

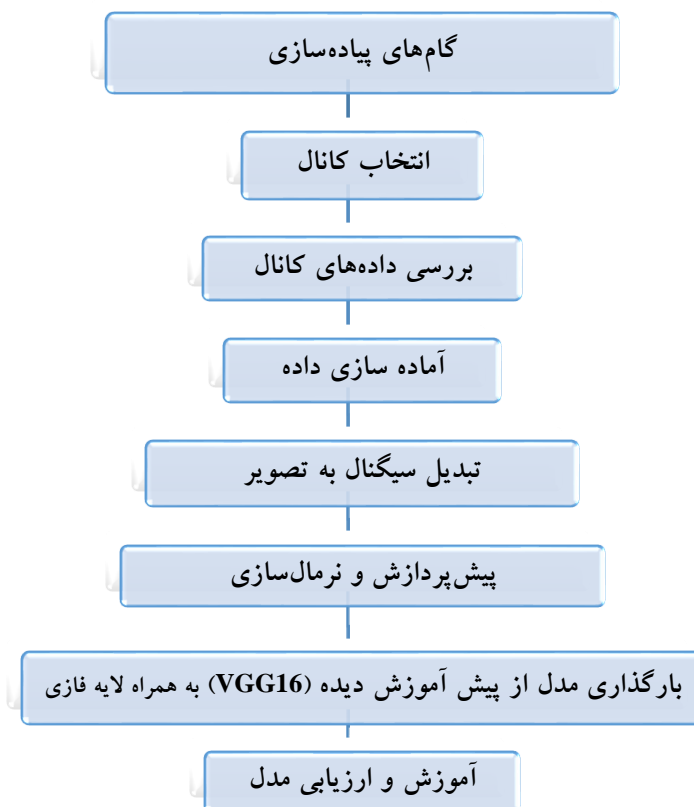
<sup>۳</sup> Graph Extreme Learning Machine

<sup>۱</sup> Electrooculography (EOG)

<sup>۲</sup> Electromyography(EMG)

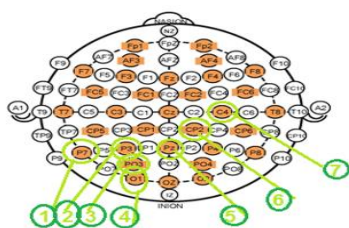
<sup>۳</sup> Skin Temperature

<sup>۴</sup> Galvanic skin response



شکل ۱: تشخیص احساسات مبتنی بر تبدیل سیگنال به تصویر و یادگیری انتقالی

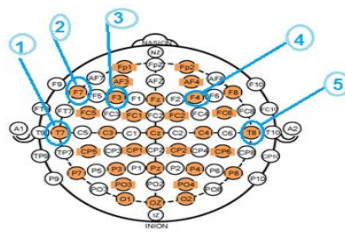
به بهبود یادگیری مدل نشده است. از این رو کانال‌هایی که برای تشخیص احساسات اهمیت دارد را پایش و شناسایی دقیق و مدون کرده و از آن‌ها برای تحلیل و شناسایی احساسات استفاده شده است. مطابق [۲۷، ۲۸] کانال‌های F3، F4، F7، T7، T8 و کانال‌های P3، P7، P03، O1، C4، PZ و CP2 بیشترین ارتباطات را با حالات هیجانی دارند (شکل ۲ الف و ب) [۲۷، ۲۸].



در ادامه سلسله روال مدل پیشنهادی مورد بحث و بررسی قرار می‌گیرد.

### ۱-۳-۱- انتخاب کانال

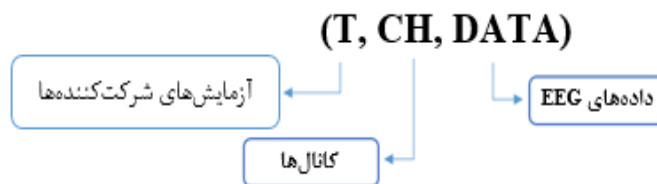
انتخاب کانال به دو منظور انجام می‌گیرد یکی برای کاهش ابعاد داده و دیگری حذف داده‌های اضافی که نه تنها به یادگیری مدل کمک نمی‌کند بلکه نتیجه را بدتر هم می‌کند. در [۲۷، ۲۸]، کانال‌های بسیاری امتحان شده ولی نتایج نشان داده که افزایش تعداد کانال‌ها الزاماً منجر



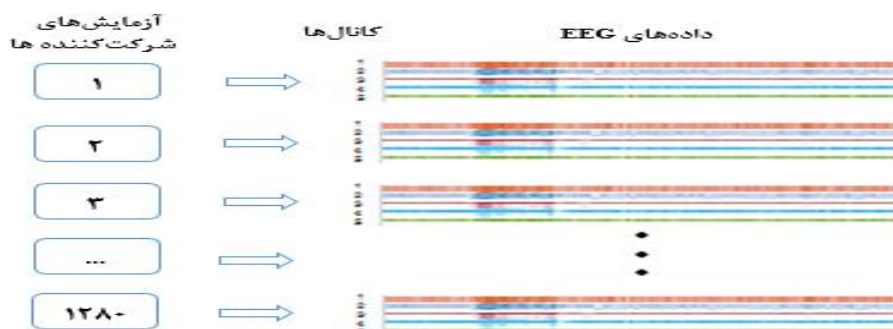
شکل ۲: مجموعه کانال‌های انتخاب شده [۲۸، ۲۷]

های یادگیری عمیق، بتوان احساسات مختلف را دسته‌بندی کرد. داده‌های مربوط به ۵ کانال، در یک آرایه‌ی سه بعدی که به ترتیب هر بعد بیانگر آزمایش‌های شرکت کننده‌ها، تعداد کانال‌ها و داده‌های الکتروانسفالوگرام ثبت شده از هر کانال است، در شکل‌های ۳ و ۴ نشان داده شده است.

در ادامه آماده‌سازی داده‌ها صورت می‌گیرد. تمامی آزمایش‌های مربوط به ۳۲ شرکت‌کننده در مجموع ۱۲۸۰ نمونه‌ی آزمایش است. صرف نظر از جزئیات الگوی احساسی سیگنال‌های الکتروانسفالوگرام، یک اشتراک در الگوی سیگنال تولید شده برای احساسات مشابه وجود دارد که تا حد زیادی انتظار می‌رود با شناسایی این الگوهای احساسی توسط مدل



شکل ۳: آرایه‌ی سه بُعدی داده‌های مرتب شده



شکل ۴: آرایه‌ی سه بُعدی داده‌های ثبت شده

معمول به حذف نویز از داده‌ها برای نزدیک شدن به سیگنال‌های عصبی واقعی اشاره دارد. ابعاد داده‌های الکتروانسفالوگرام بسیار بالا است و تفسیر و توصیف آن دشوار است. تصمیم‌گیری برای این که این داده‌ها به شکلی مرتب شوند که بتوان از آن برای مدل‌های یادگیری استفاده کرد کار دشواری است. آرایش مجموعه داده‌ی DEAP در شکل ۵ نمایش داده شده است.

## ۲-۱-۳- آموزش داده‌ها

به طور کلی، پیش‌پردازش، فرایند تبدیل داده‌های خام به قالبی است که برای تجزیه و تحلیل بهتر و همچنین تبدیل به قالبی قابل تفسیر برای کاربر است. در مورد داده‌های الکتروانسفالوگرام، پیش‌پردازش به طور



شکل ۵: آرایش مجموعه داده DEAP

نه تنها موجب بهبود دقت نمی‌شوند بلکه گاهی باعث کاهش دقت می‌گردد. از این رو سعی می‌شود با حذف یک سری کانال‌ها که اطلاعات زیادی در اختیار قرار نمی‌دهند ابعاد داده را کاهش داده و کانال‌هایی انتخاب شوند که متناسب با هدف تشخیص احساسات هستند.

برای استخراج ویژگی، اول باید داده‌ها به آرایشی مناسب مرتب شوند تا قابل تحلیل باشند. در هر آزمایش ۸۰۶۴ نمونه در مدت زمان ۶۳ ثانیه با فرکانس ۱۲۸ هرتز ثبت شده است. ۳ ثانیه‌ی اول مربوط به زمان آماده‌سازی و ۶۰ ثانیه طول دوره‌ی آزمایش است. داده‌های ۳ ثانیه‌ی اول حذف شده‌اند. با این کار ابعاد داده از ۸۰۶۴ به ۷۶۸۰ کاهش یافته است. با توجه به این که چه تعداد کانال برای بررسی داده‌ها انتخاب شوند ابعاد داده می‌تواند بیشتر و یا کمتر شوند. نتایج حاصل از تحقیقات پیشین و تجربیات این پژوهش نشان داده است که افزایش تعداد کانال‌ها

## ۳-۳-۲- تبدیل سیگنال به تصویر

شبکه یک آرایه‌ی بسیار بزرگ است چه بسا که با در نظر گرفتن داده‌های چند کاناله به مراتب آرایه‌ی ورودی شبکه بزرگ و بزرگتر می‌شود. از این رو با تبدیل سیگنال‌های الکتروانسفالوگرام به تصویر این مشکل حل می‌شود و حجم داده نیز کاهش می‌یابد.

– **محاسبه‌ی اسپکتروگرام:** برای تولید اسپکتروگرام، یک سیگنال در حوزه‌ی زمان به بخش‌های کوتاه‌تر با طول برابر تقسیم می‌شود. سپس، تبدیل سریع فوریه روی هر بخش اعمال می‌شود. اسپکتروگرام در مقایسه با تبدیل سریع فوریه، الگوهای تغییر انرژی را نشان می‌دهد که ممکن است در تبدیل سریع فوریه و چگالی طیف توان، قابل مشاهده نباشد. اسپکتروگرام تبدیل فوریه‌ی کوتاه مدت یک سیگنال را محاسبه می‌کند [۲۹].

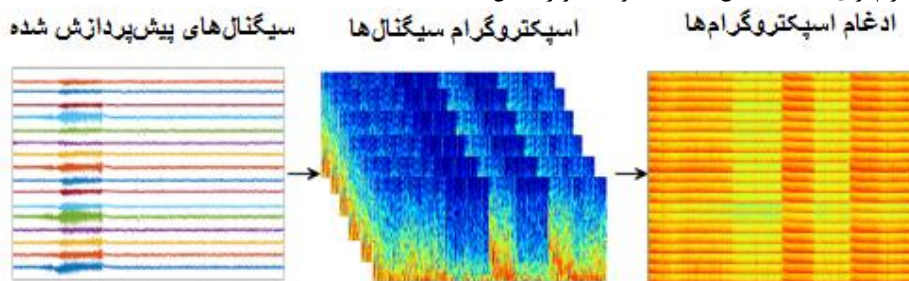
داده‌های ثبت شده از هر کانال در آزمایش یک آرایه‌ی بزرگ است و حافظه‌ی بسیاری را اشغال می‌کنند. جدا از بحث حافظه، بزرگ شدن آرایه‌ی داده، در سرعت پردازش اثر گذاشته و سرعت پردازش را کاهش می‌دهد. انتخاب روشی که در عین حال با حفظ محتوای اطلاعات، حجم داده را برای پردازش و بررسی کاهش دهد یکی از موضوعات اساسی این پژوهش است. مدل‌های یادگیری عمیق به دلیل لایه‌ها و تعداد نورون‌هایی که دارند به تنهایی برای یادگیری، مدت زمان زیادی را صرف می‌کنند. از آنجاییکه مجموع داده‌های ثبت شده از هر کانال برای ورودی

$$[S, F, T] = \text{spectrogram}(\text{signal}) \quad (1)$$

شده است. سپس تصویر را ذخیره کرده و تصویر حاصل از هر کانال را زیر هم چسبانده و تصاویر مربوط به ۵ کانال را در یک تصویر به ابعاد (۲۲۴،۲۲۵،۳) با فرمت jpeg ذخیره شده است. روند تبدیل سیگنال به تصویر در شکل ۶ نشان داده شده است.

در رابطه‌ی ۱، بردار فرکانس‌های  $F$  و بردار زمان‌های  $T$  را که در آن اسپکتروگرام محاسبه می‌شود، بر می‌گرداند. طول  $F$  برابر تعداد ردیف‌های  $S$  است. طول  $T$  برابر تعداد سگمنت‌ها است.

– **تبدیل سیگنال‌های هر کانال به تصویر:** پس از انتخاب کانال و آماده‌سازی داده، اسپکتروگرام برای تک تک کانال‌ها محاسبه و به تصویر تبدیل



شکل ۶: تبدیل سیگنال به تصویر

## ۳-۳-۳- پیش‌پردازش و نرمال‌سازی

داده‌های الکتروانسفالوگرام پس از تبدیل به تصویر مقادیری بین ۰ تا ۲۵۵ دارند و با استفاده از رابطه‌ی ۲، به مقادیر بین ۱- و ۱ تبدیل شده‌اند.

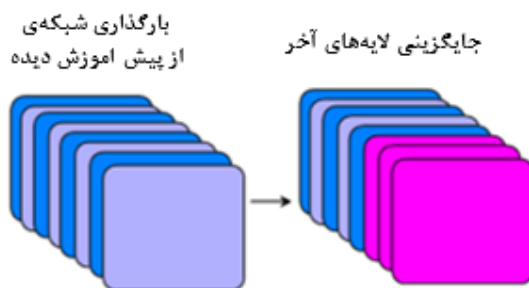
$$X = ((X.\text{astype}('float32')/255.) \times 2) - 1 \quad (2)$$

تا ۱ تبدیل خواهند شد. سپس مقادیر پس از نرمال‌سازی و پیش‌پردازش آماده‌ی ورود به شبکه خواهند شد.

$X$  آرایه‌ای از پیکسل‌های تصویر با مقادیری بین ۰ تا ۲۵۵ است. ابتدا  $X$  را به مقادیر اعشاری تبدیل کرده سپس با تقسیم بر ۲۵۵ مقادیر بین ۰ تا ۱ خواهند بود و بعد با ضربدر ۲ و تفریق حاصل از ۱، مقادیر بین ۱-

به داده‌ی زیادی نیاز ندارند و زمانی که سایز مجموعه داده کوچک باشد، انتخاب مناسبی هستند [۳۲].

جزئیات پیاده‌سازی مدل از پیش آموزش دیده: در این مقاله از کتابخانه‌های keras و Tensorflow در پایتون، استفاده شده است. مطابق شکل ۷ می‌توان با بارگذاری این مدل‌ها و جایگزینی لایه‌های آخر متناسب با مسأله‌ی خود، از شبکه برای حل مسئله استفاده کرد.

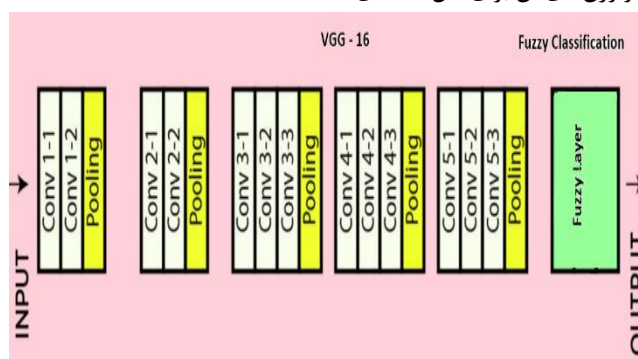


شکل ۷: آموزش با انتقال به کمک شبکه‌های از پیش آموزش دیده

استفاده شده است. در شکل ۸ لایه‌های مدل پیشنهادی نمایش داده شده است.

شبکه از پیش آموزش دیده روی یک مجموعه داده آموزش داده می‌شود. سپس از روی وزن‌های شبکه دانش کسب می‌کند. می‌توان این وزن‌ها را استخراج کرده و به هر شبکه‌ی عصبی دیگری منتقل کرد و به جای آموزش شبکه‌ی عصبی دیگر از ابتدا، ویژگی‌های آموخته شده را انتقال داد. یک مدل از پیش آموزش دیده مدلی است که توسط دیگران برای حل یک مسأله ایجاد شده است. به جای ساختن یک مدل از ابتدا برای حل یک مسأله مشابه، می‌توان از مدل آموزش دیده روی مسأله دیگر به عنوان نقطه شروع استفاده کرد. مدل‌های از پیش آموزش دیده اغلب

سیگنال‌های الکتروانسفالوگرام که از پیش به تصویر تبدیل شده‌اند پس از نرمال‌سازی و پیش‌پردازش به ورودی شبکه‌ی از پیش آموزش دیده به نام VGG16 داده شده و از معماری و وزن‌های آن برای حل مسأله‌ی



شکل ۸: معماری مدل پیشنهادی

بسیار بالا و Zero-Crossing Rate بالا و تغییر سریع فرکانس باشد، قانون ۴: تشخیص ترس: اگر Pitch نوسانی و انرژی متوسط به پایین و سرعت متوسط باشد. همچنین تابع عضویت در این مقاله تابع گاوسی می‌باشد.

قابل توجه است که ادغام شبکه VGG با لایه‌های فازی مزایای قابل توجهی دارد. یکی از مهم‌ترین این مزایا، توانایی مدیریت عدم قطعیت در داده‌ها است. لایه‌های فازی به خوبی با داده‌های نامشخص یا ناهمگن کنار می‌آیند که این امر به‌ویژه در کاربردهای واقعی، جایی که داده‌ها همیشه دقیق نیستند، بسیار مفید است. علاوه بر این، استفاده از منطق فازی باعث می‌شود که مدل‌ها قابل تفسیرتر شوند، زیرا می‌توانند نتایج را به صورت عبارات فازی ارائه دهند و این مساله در تحلیل خروجی‌ها و

همانطور که در شکل ۸ توضیح داده شد، در مدل پیشنهادی خروجی‌های لایه‌های عمیق VGG، که ویژگی‌های تصویری را استخراج می‌کنند، به لایه‌های فازی متصل می‌شوند. این لایه‌ها به صورت تطبیقی با استفاده از فرآیندهای آموزشی نظارت‌شده به یادگیری، قوانین فازی‌ای ایجاد می‌کنند تا نرمی و عدم قطعیت داده‌ها را در نظر بگیرند. برای آموزش مدل، از یک تابع هزینه استفاده می‌شود که هم خروجی‌های قطعی و هم خروجی‌های فازی را در نظر می‌گیرد. از قوانین فازی‌ای که در این بخش استفاده شده می‌توان به موارد زیر اشاره کرد:

قانون ۱: تشخیص شادی: اگر انرژی بالا و Pitch بالا و سرعت گفتار سریع باشد، قانون ۲: تشخیص ناراحتی: اگر انرژی پایین و Pitch پایین و سرعت گفتار آهسته باشد، قانون ۳: تشخیص عصبانیت: اگر انرژی

است: ۱- استفاده از مدل‌های پیش‌آموزش دیده، ۲- تغییر وزن‌ها و پارامترهای مدل بر اساس داده‌های جدید و ۳- آموزش محدود روی مجموعه داده‌های کوچک‌تر و خاص‌تر بوده است. همچنین در آزمایشات از مجموعه اعتبارسنجی متقاطع K-Fold برابر ۱۰ برای تنظیم دقیق مدل استفاده می‌شود. دلیل استفاده از مجموعه اعتبارسنجی این است تا نشان دهیم با هر گونه افزایش دقت در مجموعه داده‌های آموزش، دقت در مجموعه داده‌ای که قبلاً به شبکه نشان داده نشده است افزایش می‌یابد یا خیر. زیرا اگر دقت در مجموعه داده‌های آموزش افزایش یابد، اما دقت در مجموعه داده‌های اعتبارسنجی ثابت بماند یا کاهش یابد؛ یعنی مدل در حال بیش‌برازش است و باید آموزش را متوقف کنیم. همچنین جزئیات مربوط به تنظیم ابرپارامترهای مدل پیشنهادی در جدول ۱ نشان داده شده است.

تصمیم‌گیری به ما کمک می‌کند. همچنین، ترکیب قابلیت یادگیری عمیق VGG با پردازش فازی می‌تواند دقت نتایج را در تشخیص‌های پیچیده افزایش دهد. این ترکیب اجازه می‌دهد تا مدل‌ها به‌طور دینامیک با شرایط مختلف انطباق پیدا کنند و نیاز به تنظیمات دقیق پارامترها را کاهش دهد. ضمن اینکه، لایه‌های فازی می‌توانند به کاهش مشکل یادگیری بیش‌ازحد در مدل کمک کنند، زیرا این قابلیت را دارند که بر الگوهای کلی‌تر تمرکز کنند و بدین ترتیب عملکرد کلی مدل‌های مبتنی بر شبکه‌های عصبی را در زمینه‌های مختلف بهبود بخشند. برای جلوگیری از بیش‌برازش و افزایش تعمیم‌پذیری مدل از تکنیک‌های Global Average Pooling، حذف تصادفی، منظم‌سازی، نرمال‌سازی دسته‌ای، اعتبارسنجی متقابل و توقف زودهنگام استفاده شده است.

#### ۴- نتایج آزمایشات

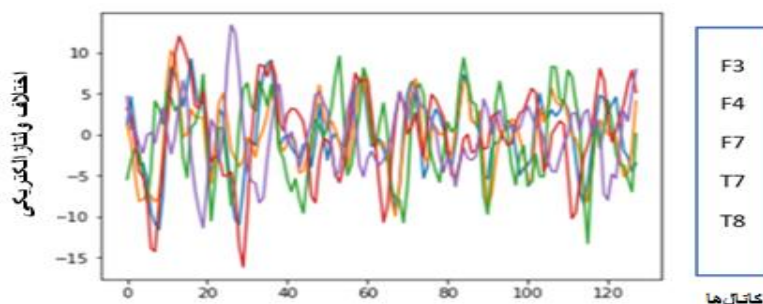
در پیاده‌سازی مدل پیشنهادی از نرم‌افزار پایتون و فرآیند fine tuning استفاده شده است. مراحل فرآیند fine tuning بدین‌صورت

جدول ۱: تنظیم فرآیندهای مدل

نرخ یادگیری	بهینه‌ساز	تابع زیان	اندازه دسته	تعداد گام آموزش
۰/۰۰۰۱	Adam	Binary_CrossEntropy	۱۶	۱۰

زمانی برای کانال‌ها در شکل ۹ نمایش داده شده است. مقایسه‌ی آن‌ها مشخص می‌کند که مجموعه‌ای که انتخاب شده اطلاعات گوناگونی در اختیار قرار می‌دهد یا خیر.

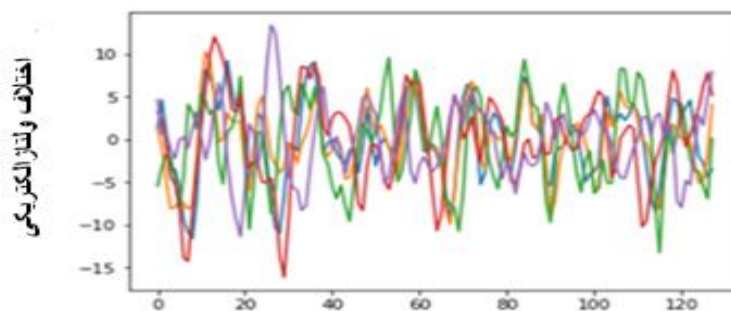
سیگنال‌های ثبت شده از مجموعه کانال‌های انتخاب شده در شکل‌های ۹ تا ۱۲ نمایش داده شده است. برای مقایسه‌ی الگوی سیگنال‌ها از کانال‌های گوناگون یک شرکت‌کننده، سیگنال ثبت شده در یک بازه‌ی



شکل ۹: نمونه داده‌های ۵ کانال از یک شرکت‌کننده در یک آزمایش

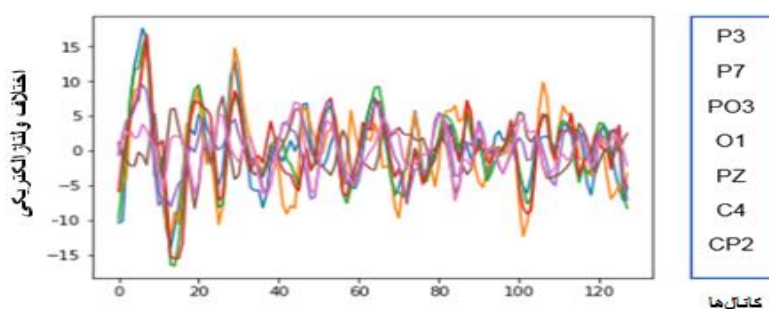
که هر یک از این کانال‌ها اطلاعات متفاوتی را در اختیار قرار می‌دهد. از این رو انتظار می‌رود که بتوان از مجموع این کانال‌ها اطلاعات مفیدی استخراج کرد.

در شکل ۹ داده‌های ثبت شده از کانال‌های F3 و F4, F7, T7, T8 در ۱ ثانیه برای یک شرکت‌کننده در یک آزمایش نشان داده شده است. الگوی سیگنال‌های این کانال‌ها با یکدیگر فرق دارد به این معنی است



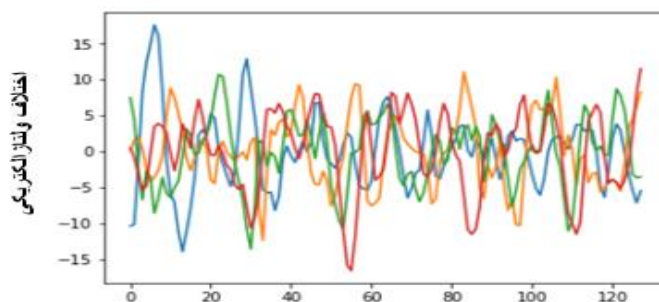
شکل ۱۰: نمونه داده‌های کانال F3 برای یک شرکت کننده در ۴ آزمایش

در شکل ۱۰ داده‌های کانال F3 برای یک شرکت کننده در آزمایش‌های گوناگون در مدت زمان ۱ ثانیه نمایش داده شده است. همانطور که در شکل ۱۰ مشخص است سیگنال‌های ثبت شده از یک کانال در آزمایش‌های گوناگون متفاوت است و نشان دهنده‌ی آن است که در اثر تحریک‌های مختلف الگویی که در مغز ایجاد می‌شود متفاوت است.



شکل ۱۱: نمونه داده‌های ۷ کانال از یک شرکت کننده در یک آزمایش

در شکل ۱۱ داده‌های ثبت شده از کانال‌های P3، P7، PO3، O1، PZ، C4 و CP2 در ۱ ثانیه برای یک شرکت کننده در یک آزمایش نشان داده شده است. همچنین در شکل ۱۲ نمونه داده‌های کانال P3 برای یک شرکت کننده در آزمایش‌های گوناگون در مدت زمان ۱ ثانیه نمایش داده شده است.



شکل ۱۲: نمونه داده‌های کانال P3 برای یک شرکت کننده در ۴ آزمایش

پیاده‌سازی روش پیشنهادی به کمک یادگیری انتقالی و لایه فازی و بدون لایه فازی در جداول ۲ و ۳ آمده است. همان‌طور که مشخص است، استفاده از لایه فازی روی شبکه VGG16 منجر به افزایش دقت دسته‌بندی شده است.

با توجه به اشکال ۹ تا ۱۲ و صرف نظر از جزئیات الگوی احساسی سیگنال‌های الکتروانسفالوگرام، یک اشتراک در الگوهای سیگنال‌های تولید شده برای احساسات مشابه وجود دارد که تا حد زیادی انتظار می‌رود با شناسایی این الگوهای احساسی توسط مدل‌های یادگیری عمیق، بتوان احساسات مختلف را دسته‌بندی کرد. نتایج حاصل از

جدول ۲: نتایج مدل از پیش آموزش دیده VGG16 بدون لایه فازی

کانال‌ها	ظرفیت		برانگیختگی	
	دقت متوسط (%)	انحراف معیار (%)	دقت متوسط (%)	انحراف معیار (%)
Images_5_CH	۹۰/۵۶	±۴/۳۸	۰/۰۷	±۲/۳۷
Images_7_CH	۹۱/۵۷	±۳/۳۱	۰/۰۷	±۴/۴۴
Images_32_CH	۸۲/۶۵	±۳/۳۲	۰/۰۸	±۴/۶۱
Physiological_images	۸۹/۰۱	±۳/۹۳	۰/۰۷	±۵/۳۴

جدول ۳: نتایج مدل از پیش آموزش دیده VGG16 با لایه فازی

کانال‌ها	ظرفیت		برانگیختگی	
	دقت متوسط (%)	انحراف معیار (%)	دقت متوسط (%)	انحراف معیار (%)
Images_5_CH	۹۱/۷۷	±۲/۱۲	۰/۰۱	±۱/۳۴
Images_7_CH	۹۳/۵۰	±۱/۸۱	۰/۱۱	±۲/۷۸
Images_32_CH	۸۹/۶۳	±۲/۴۲	۰/۱۱	±۳/۴۱
Physiological_images	۹۰/۸۹	±۲/۶۱	۰/۱۱	±۲/۱۸

همچنین جدول ۴ مقایسه مدل‌های مختلف را با مدل پیشنهادی یکبار با در نظر گرفتن دسته‌بندی فازی و یکبار بدون در نظر گرفتن دسته‌بندی فازی نشان می‌دهد.

جدول ۴: مقایسه مدل پیشنهادی با سایر کارهای انجام شده

مراجع	برانگیختگی (%)	ظرفیت (%)
Choi et al. [31]	۷۴/۶۵	۷۸
Alhagry et al.[32]	۸۵/۶۵	۸۵/۴۵
Acharya-CNN[30]	۸۴/۷	۸۵/۰۱
Acharya-LSTM[30]	۸۱/۹۱	۸۴/۳۹
بندی فازی مدل پیشنهادی بدون در نظر دسته	۸۶/۹۸	۸۵/۲۳

اصلی در این مقاله تحلیل سیگنال‌های مغزی، روش‌های مختلف پردازش و تحلیل سیگنال، آشنایی بیشتر با روش‌های یادگیری ماشین برای تحلیل سیگنال‌های مغزی، احساسات و اهمیت شناخت آن‌ها و شناسایی احساسات از روی تغییر الگوی سیگنال‌های الکتریکی مغز به کمک یادگیری عمیق است. در روش پیشنهادی از این مقاله، برای مقابله با کمبود داده‌ها، از یادگیری انتقالی استفاده شده و برای غلبه به مشکل ابعاد بالای داده‌های الکتروانسفالوگرام تکنیک تبدیل سیگنال به تصویر مورد استفاده قرار گرفته است. با توجه به نتایج حاصل از روش پیشنهادی در این مقاله، مشخص شده که استفاده از مدل از پیش آموزش دیده شده VGG16 در کنار لایه فازی می‌تواند منجر به افزایش قابل توجه دقت دسته‌بندی سیگنال‌های مغزی شود. با توجه به نتایج حاصل از تحقیقات و مطالعات انجام شده در زمینه‌ی تشخیص احساسات از روی سیگنال‌های الکتروانسفالوگرام، روش‌های استخراج و انتخاب ویژگی مناسب برای یافتن یک آرایه از ویژگی‌های مشترک میان افراد مختلف در برانگیختگی احساسی مشابه در برابر یک محرک یکسان، می‌تواند قدرت پیش‌بینی مدل‌هایی که برای شناسایی احساسات ارائه می‌شوند را تا حد زیادی بهبود بخشد. پس از فرآیند شناسایی ویژگی‌های مناسب، روش‌های مختلف افزایش داده می‌تواند به بهبود تشخیص و شناسایی احساسات توسط مدل‌های یادگیری کمک بیشتری کند. علاوه بر بحث تعداد نمونه‌ها، برچسب‌گذاری داده‌های ثبت شده نیز بسیار اهمیت دارد. این مسأله کیفیت تشخیص مدل را تحت تاثیر قرار می‌دهد و منجر می‌شود تعداد پیش‌بینی‌های اشتباه افزایش یابد. تعداد این موارد هر چقدر بیشتر باشد، دقت پیش‌بینی کمتر می‌شود. موارد ذکر شده جزو مسائلی است که برای بالابردن نرخ تشخیص احساسات در تحقیقات آتی مورد بحث و بررسی قرار خواهند گرفت.

### تعارض منافع

نویسندگان این مقاله اعلام می‌دارند که هیچ‌گونه تعارض منافع توسط نویسندگان بیان نشده است.

### مراجع

- [1] Pranav E, Kamal S, Chandran C S, Supriya M H, facial emotion recognition using deep convolutional neural network, 6th International Conference on Advanced Computing & Communication System (ICACCS), Kochi, India, 2020.
- [2] Jaiswal A, Facial Emotion Detection Using Deep Learning, International Conference for Emerging Technology (INCET), Belgaum, India, 2020.
- [3] Akinpelu S, Lightweight Deep Learning Framework for Speech Emotion Recognition, IEEE Access, 2023.

با مشاهده جدول ۴ می‌توان نتیجه گرفت که روش پیشنهادی بهتر از سایر روش‌های رقیب در تشخیص احساسات عمل می‌کند. انگیزه اصلی در انتخاب شبکه‌ی از پیش آموزش دیده به دلیل برتری آن‌ها در استخراج ویژگی است. از آنجایی که توزیع داده‌های الکتروانسفالوگرام تغییرات از فردی به فرد دیگر، جلسه به جلسه و مجموعه داده به مجموعه داده را نشان می‌دهد، شناسایی مجموعه ویژگی‌هایی که هر بار دقت خوبی را نشان می‌دهند دشوار است. از طرف دیگر، مدل‌های از پیش آموزش دیده در استخراج ویژگی بسیار شایسته هستند. برای بهبود دقت تشخیص احساسات، در مرحله‌ی آزمون، از فیلتر میان‌گذر برای کاهش پیش‌بینی‌های غلط مشهود استفاده می‌شود. فرایند استخراج ویژگی به طور کامل به یک شبکه‌ی عصبی پیچشی از پیش آموزش دیده واگذار شده تا قابلیت استخراج ویژگی در کارهای مختلف دسته‌بندی را نشان داده شده است. این قابلیت، مدل را قادر می‌سازد تا ویژگی‌های مفید و پنهان را برای دسته‌بندی کشف کند. نرمال‌سازی داده‌ها به منظور حذف اثرات نوسانات دامنه‌ی ولتاژ و محافظت از شبکه‌ی پیشنهادی در برابر شرایط احتمالی اعمال می‌شود. خروجی شبکه به منظور حذف پیش‌بینی‌های غلطی که ممکن است در بازه‌های زمانی کوچک اتفاق بیفتد فیلتر می‌شود. همان‌طور که در این مقاله بیان شده است، تغییرات عاطفی بیشتر روی سیگنال‌های الکتروانسفالوگرام در لوب‌های پیشانی و گیجگاهی تأثیر می‌گذارد.

### ۵- نتیجه‌گیری

در این مقاله مدلی برای دسته‌بندی پارامترهایی جهت تشخیص احساسات از روی سیگنال‌های الکتروانسفالوگرام، ارائه شده است. سیگنال‌های الکتروانسفالوگرام، داده‌هایی با ابعاد بالا و تعداد کم هستند. ماهیت ناپایدار سیگنال‌های الکتروانسفالوگرام در طول زمان و متغیر با زمان بودن آن‌ها، تحلیل و دسته‌بندی را دشوار می‌کنند. هدف اصلی در شناسایی احساسات از روی سیگنال‌های الکتروانسفالوگرام یافتن یک الگوی مشترک میان سیگنال‌های ثبت شده در افراد مختلف برای تحریک‌های یکسان و برانگیختگی احساس مشابه است. نتایج حاصل از مطالعات و تحقیق‌های پیشین و این پژوهش نشان می‌دهد که الگوهای احساسی در بین افراد مختلف متفاوت است. الگوی احساسی برای یک فرد در زمان‌های مختلف نیز متفاوت است. از این رو پیدا کردن یک الگوی مشترک با دقت بالا کار دشواری است. چرا که هر اتفاقی در پیرامون یا حتی در درون یک فرد، روی تغییرات الگوی سیگنال تاثیرگذار است و همچنین برای تحریک افراد با یک محرک یکسان ممکن است میزان تحریک‌پذیری در افراد مختلف یکسان نباشد. هدف

- [19] Cimtay Y, Ekmekcioglu E. Investigating the use of pretrained convolutional neural network on cross-subject and cross-dataset EEG emotion recognition, *Sensors*, 2020; 20(7): 20-34.
- [20] Yin Z, et al. Recognition of emotions using multimodal physiological signals and an ensemble deep learning model, *Computer methods and programs in biomedicine*, 2017; 140: 93-110.
- [21] Alhagry S, Fahmy A A, El-Khoribi R A. Emotion recognition based on EEG using LSTM recurrent neural network, *International Journal of Advanced Computer Science and Applications*, 2017; 8(10).
- [22] Salama E S. EEG-based emotion recognition using 3D convolutional neural networks, *International Journal of Advanced Computer Science and Applications*, 2018; 9(8).
- [23] Zheng W L, Zhu J-Y, Lu B-L, Identifying stable patterns over time for emotion recognition from EEG, *IEEE transactions on affective computing*, 2017; 10(3): 417-429.
- [24] Li, X. Emotion recognition from multi-channel EEG data through convolutional recurrent neural network, *IEEE international conference on bioinformatics and biomedicine (BIBM)*, 2016.
- [25] Tripathi S. Using deep and convolutional neural networks for accurate emotion classification on DEAP data, in *Proceedings of the AAAI Conference on Artificial Intelligence*. 2017.
- [26] Aslan Z, Akin M. Automatic Detection of Schizophrenia by Applying Deep Learning over Spectrogram Images of EEG Signals, *Traitement du Signal*, 2020; 37(2).
- [27] Raghu S. EEG based multi-class seizure type classification using convolutional neural network and transfer learning, *Neural Networks*, 2020; 124: 202-212.
- [28] Bizopoulos P, Lambrou G I, Koutsouris D. Signal2image modules in deep neural networks for eeg classification, in *2019 41st Annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. 2019.
- [29] Prasomphan S. Improvement of speech emotion recognition with neural network classifier by using speech spectrogram, in *2015 International Conference on Systems, Signals and Image Processing (IWSSIP)*. 2015.
- [30] Acharya R. Multi-Class Emotion Classification Using EEG Signals, in *Advances in Computing*, Singapore, 2021; 1367: 474-485.
- [31] Choi E. K. D. Arousal and valence classification model based on long short-term memory and deap data for mental healthcare management, *Healthc*, 2018; 4(24): 309-316.
- [32] Alhagry S. Emotion recognition based on EEG using LSTM recurrent neural network, *Emotion*, 2017; 8(10): 355-358.
- [4] Tarunika P A K, "Applying Machine Learning Techniques for Speech Emotion Recognition, in *2018 9th International Conference on Contemporary Computing and Communication Networks (ICCCNT)*, Bengaluru, India, 2018.
- [5] Torres M, Evaluation of Interpretability for Deep Learning Algorithms in EEG Emotion Recognition: A Case Study in Autism, in *Artificial Intelligence in Medicine*, 2023.
- [6] Bazgir O, Emotion recognition with machine learning using EEG signals, *Journal of Electrical and Computer Engineering*, 2023.
- [7] Raveendrababu Vempati L D S, A systematic review on automated human emotion recognition using electroencephalogram signals and artificial intelligence, *Resulted in Engineering*, 2023; 18: 101027.
- [8] Suhaimi N J T, EEG-based emotion recognition: A state of the art review of current trends and opportunities, *Computational Intelligence and Neuroscience*, 2022: 8875426.
- [9] Yuwei Chen J H, Deep Learning-Based Emotion Detection", *Journal of Computer and Communications*, 2022; 10: 57-71.
- [10] Phan T, EEG-based emotion recognition by convolutional neural network with multi-scale kernels, *Sensors*, 2021; 21(15).
- [11] Kwon H, Electroencephalography Based Fusion Two-Dimensional (2D)-Convolution Neural Networks (CNN) Model for Emotion Recognition System, *Sensors*, 2018; 18(5).
- [12] Zhong Q, Electroencephalogram access for emotion recognition based on a deep hybrid network, *Frontiers in Human Neuroscience*, 2020.
- [13] Cho H H J. Spatio-temporal representation of an electroencephalogram for emotion recognition using a three-dimensional convolutional neural network, *Sensors*, 2020; 20(12).
- [14] Zhang A, et al. EEG data augmentation for emotion recognition with a multiple generator conditional Wasserstein GAN, *Complex & Intelligent Systems*, 2021: 1-13.
- [15] Luo Y, et al. Data augmentation for enhancing EEG-based emotion recognition with deep generative models, *Journal of Neural Engineering*, 2020; 17(5): 056021.
- [16] Gonzalez H A, Yoo J, Elfadel I M. EEG-based emotion detection using unsupervised transfer learning, in *2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, 2019, IEEE.
- [17] Yin Z, et al. Cross-subject EEG feature selection for emotion recognition using transfer recursive feature elimination, *Frontiers in neurorobotics*, 2017; 11: 19.
- [18] Cai J, Chen W, Yin Z. Multiple transferable recursive feature elimination technique for emotion recognition based on EEG signals, *Symmetry*, 2019; 11(5): 683.