International Journal of Applied Engineering Research ISSN 0973-4562 Volume 19, Number 4 (2024) pp. 280-288
https://dx.doi.org/10.37622/IJAER/19.4.2024.280-288

Classification of EEG Signal by Using Transformer Network.

[1]Krashana D. Nirde*, [2]Dr. R. R. Manthalkar, [3]Manasa Gunda, [4]Dr. S. S. Gajre
[(1, 2, 3, 4)]SGGSIE&T, Electronics & Telecommunication Engg. Nanded, (Maharashtra), India, 431606

Abstract
Transformer networks have lately been successfully applied to computer vision, generative adversarial networks (GAN), and reinforcement learning due to their extraordinary effectiveness in natural language processing (NLP). Electroencephalogram (EEG) data classification has proven difficult, and researchers have relied too much on manual feature extraction and pre-processing. Deep learning hasn't been completed for EEG yet, despite automated feature extraction being performed in several other domains. This study investigates the transformer network's effectiveness in classifying unprocessed EEG data. (Filtered and pre-processed). A publicly accessible MAT dataset was used to assess the transformer networks' performance. The workload categorization of No Task and Arithmetic Task (mental arithmetic task) is optimized for the classifier using free access raw workload EEG data (MAT). The proposed transformer network achieves an accuracy comparable to state-of-the-art accuracy on MAT dataset is 98.92% for two workload levels. Without feature extraction, the accuracy value was obtained from raw EEG data. According to the results, transformer-based deep learning models can successfully reduce the requirement for extensive feature extraction from EEG data to achieve successful classification.

Keywords: Classification, EEG, Mental Arithmetic Task, Deep learning, Transformer,

1. Introduction

The electroencephalogram (EEG), heart rate, breathing, eye movement, electromyogram, and skin are some of the categories into which the frequently utilized physiological signals can be generally categorized. Because of its great temporal resolution, affordability, ease of use, and security, EEG is one of the most used signals among them. Thus, we concentrate on EEG-based mental arithmetic task workload recognition in this research paper.

One of the most significant mental or functional states that affect people is the cognitive workload, which is described as "the relation between the function relating the mental resources demanded by a task and those resources available to be supplied by the human." It can be affected by a variety of factors, including social and environmental influences, task complexity, individual variability, and variations in function state [1]. Cognitive workload recognition is used in many different types of work environments, including education (e.g., online learning and web browsing by students), public transportation (e.g., driving a car, an airplane, managing air traffic), and special working situations (e.g., engineers of nuclear power plants that require a lot of attention). The cognitive workload has recently been used to computer-aided diagnostics of conditions like autism spectrum disorder, depression, cancer, and schizophrenia.

Different cognitive workload states are often experienced by subjects by requiring them to engage in tasks of varying degrees of difficulty. A cognitive task in a controlled laboratory setting or running a machine in an actual or simulated world are two common paradigms used to feel cognitive workload. Subsequently, we divide the paradigms into two categories: cognitive and operant.

Task paradigms that are focused on cognitive processes. In most cases, subjects had to remain still while performing simple cognitive activities. A few examples of these paradigms are the visual deterioration task, IQ test, mental arithmetic problem, and n-back working memory challenge. The participants in the mental computation task must keep track of the outcomes of the given formulas (such as 4+5) and determine whether the supplied number (such as 09) corresponds with the results they previously calculated. Intermediate outcomes must be temporally stored, and information stored in the mental workspace must be retrieved. Literature makes use of both mathematical addition and subtraction. Additional problems can be made more difficult by carrying numbers or by using numerous digits.

This work investigated the classification of raw electroencephalography (EEG) data using transformer networks [2]. EEG has long been used for different purposes like disease diagnosis [3], BCI [4], neurofeedback [5], workload estimation [6], motor imagery classification [7], and even brain-to-brain communication [8]. EEG records the electrical activity at the human scalp produced by the ionic currents resulting from simultaneous activation of multiple neurons in the brain [9]. There are several advantages of using EEG to measure brain activity, including but not limited to its low cost, portability, and being non-invasive with a very high temporal resolution [10]. With the recent advances in deep learning [11], one expects that the classifier should learn the required features automatically and not rely on hand-engineered features. However, because of the limitations of EEG, most studies have used hand-crafted features from EEG data, even when classifying with complex neural network architectures like deep recurrent convolutions, deep-stacked autoencoders, frequency-dependent CNN, spatial-frequency CNN, etc. [12]. To build an end-to-end classification system for EEG, we chose to explore the highly effective transformer networks. To better evaluate the efficiency of transformer networks in classifying the raw EEG data, we used one public dataset (MAT; workload classification) [13].

The Contributions Of This Paper Are:

Transformer network is utilized for the classification of raw EEG data (cleaned and preprocessed; no feature extraction). To the best of our knowledge, transformer networks have never been applied to the classification of raw EEG data.

The proposed framework achieved a state-of-the-art accuracy of 98.92%. It also achieved an accuracy of 98.92% (two workload levels on a public dataset of workload classification). The classification results (accuracy, precision, recall, and F1-score) are also compared with the state-of-the-art attention networks. It has been shown that the transformer networks are a better choice, especially without needing any hand-crafted feature extraction.

The rest of the research paper is organized as follows. Section II presents a Methodology used in this research paper (MAT Dataset). Section III describes Model used for the whole experiment and the transformer network's details. Section IV reports the results, and discussion. Section V report conclusion.

EEG can assist us provide an objective, absolute, and temporally sensitive workload measure. Therefore, it has been used time and again for workload classification. Qu et a [14] used support vector machine (SVM) for classifying EEG data collected from 13 participants while doing the NASA MATB-II task [15]. They used energy from the independent components extracted from independent component analysis (ICA) as features and achieved an accuracy of 79.8% for the three workload levels (low, medium, and high). Different machine learning algorithms such as linear discriminant analysis [16], random forest [17], KNN, and SVM [18] have been utilized for workload classification from EEG data.

Different neural network architectures have already been shown to classify mental workload successfully. Hefron et al. [19] used stacked LSTMs for classifying data of six participants using the MATB-II. They achieved an accuracy of 93%. Sun et al. [20] proposed WLNet specifically for classifying mental workload and compared it to temporally constrained sparse group spatial patterns (TCSGSP, a deep version of the CSP method proposed by Zhang et al. [21]) and EEGNet [22] (proposed for BCI applications, consisting of depth-wise and separable convolution layers). WLNet was a shallow neural network consisting of one-dimensional convolution layers, thereby requiring fewer parameters. They did binary workload classification on the dataset provided by Fabien Lotte research group at INRIA [17] and found that WLNet out-performed TCSGSP and EEGNet. Saha et al. [23] used a stacked denoising autoencoder to classify three workload levels and achieved 89.51% accuracy.

Specifically for the STEW dataset, different neural networks have been applied and compared. Chakladar et al. [24] compared the perormances of stacked LSTM, BLSTM, CNN-LSTM, stacked autoencoders with LSTM/BLSTM, and their proposed model of hybrid BLSTM-LSTM network. Their proposed model achieved state-of-the-art accuracy of 86.33% and 82.57% for binary and tertiary classification on the STEW dataset. Lim et al. [13] made use of artifact subspace reconstruction

(ASR) and features were extracted using power spectral density (PSD) and neighborhood component analysis (NCA). Using the support vector regression (SVR) model, the accuracy of 69.20% was achieved for the "SIMKAP multi-task" classification. Zhu et al. [25] constructed oblique visibility graphs (OVG) and frequency domain features (mean degree, clustering coefficient, and degree distribution) were extracted. Decision tree (DT) and SVM were used for classification, where SVM performed better and yielded 89.60% and 79.50% for binary and tertiary classification, respectively. Attention-based networks have also been used for EEG classification. Zhu et al. [26] proposed a neural network based on CNN and attention mechanism to perform automatic sleep staging. Zheng and Chen [27] proposed an end-to-end attention-based Bi-LSTM method, named Bi-LSTM-AttGW for visual cognition task. The attention weighting method is applied to Bi-LSTM output.

None of the above-cited literature for the given dataset worked with raw EEG data. The input fed to the deep neural network has varied from simple band segregated data (alpha, beta, gamma, etc.) to simple features like mean, variance, etc. of the band power to much more complex features like PLV and PLI. Thus, we can conclude that EEG poses a unique challenge of complex pre-processing and feature selection mechanisms.

In this paper, we explored the possibility of using raw EEG data for classification. Given the available literature, it was evident that we needed to use a proficient neural network architecture that can automatically extract the relevant features. Recently, transformer networks have seen a phenomenal rise and are used in state-of-the-art accomplishments in almost every domain [28,29]. In the context of EEG, transformer networks have been rarely used but have been shown to generate state-of-the-art results for sleep stage classification [30] and speech recognition [31]. Therefore, we decided to use a transformer network for the classification.

2. Methodology

MAT Dataset

2.1 EEG Recording:

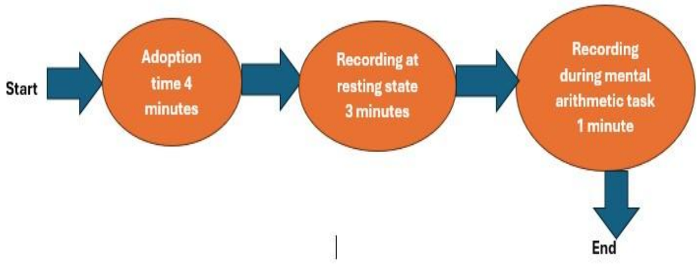
A 23-channel monopolar EEG system from Ukraine called the XAI-MEDICA was used to record the EEG data. The International 10/20 system was followed when placing silver/silver chloride electrodes on the scalp. Fp1, Fp2 electrodes correspond to prefrontal, F3, F4, Fz, F7, F8 electrodes correspond to frontal, C3, C4, Cz electrodes correspond to central, P3, P4, Pz electrodes correspond to parietal, O1, O2 electrodes correspond to occipital, and T3, T4, T5, T6 correspond to temporal. These electrodes were attached to ear reference electrodes for their calibration. The inter-electrode impedance was regularly kept below 5 k to guarantee data quality. 500 Hz was used for each channel's sampling rate. The following filters were used: a power line notch filter set at 50 Hz, a high-pass filter and a low-pass filter, both with cutoff frequencies of 0.5 Hz and 45 Hz, respectively. The amplification tract also had a time constant of 0.3 s [32].

2.2 Description of Experimental Task:

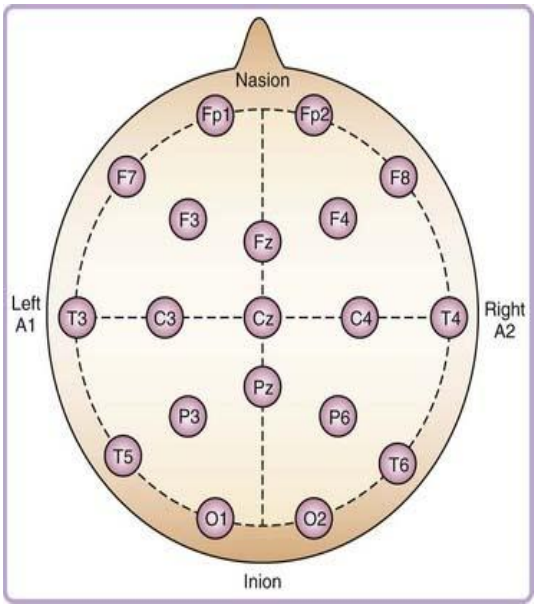
The experimental task begins with adaptation of the experiment for 0–3 min. Next 3–6 min includes subjects performing resting task where they are instructed to sit, relax and informed about how to perform the task. The last 6–10 min includes performing mental arithmetic tasks (serial subtraction). The mental arithmetic task includes performing oral communication of subtraction between 4-digit minuend and 2-digit subtrahend and the process is repeated.

2.3 EEG Collection:

Each recording consists of two distinct EEG segments, with the resting state section lasting 180 s and the mental counting segment lasting 60 s. Thirty of the original 66 subjects had to be removed from the database following a careful visual review of the EEG data by a trained electro neurophysiologist. Their EEG data was of poor quality and contained several oculographic and myographic artifacts, which led to their elimination. During the data preparation step, the artifacts (muscle, eyes and cardiac overlapping of the heart pulsation) were removed using independent component analysis (ICA). The final dataset therefore includes 36 subjects with artifacts free EEG segments. Detailed data collection protocol can be found in Fig 1 (a).



a) Data Collection Protocol



b) 10-20 EEG electrode Placement

Figure. 1

3. Model

This section discusses the proposed transformer network for the classification of mental arithmetic task workload estimation. After acquisition and pre-processing is done on EEG data, epochs are created with a window size of 0.30 s. The transformer network is then trained for classification on this raw EEG data.

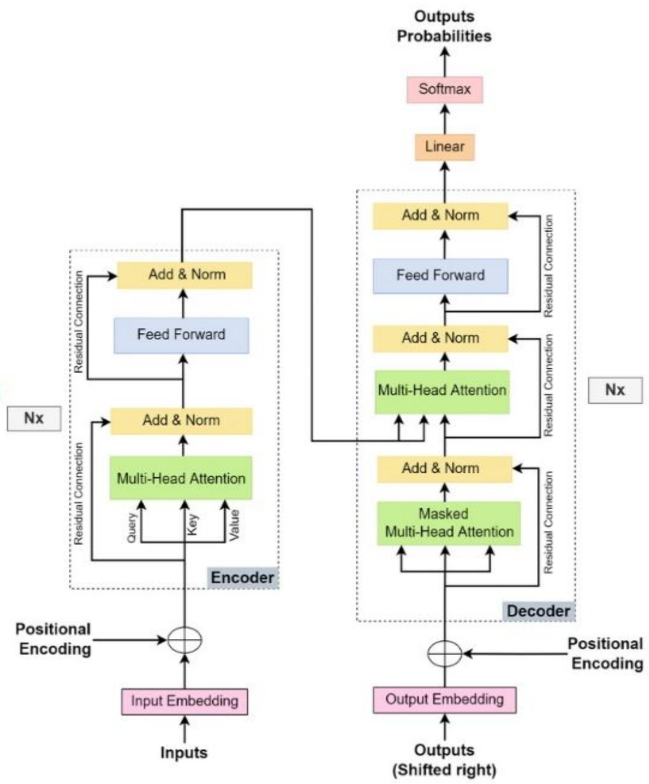
3.1 Transformer

The transformer network [33] is a neural network based on a self- attention mechanism. It has proved higher quality and lower computation requirements for language translation tasks than recurrent and convolution models. It applies a self-attention mechanism that directly models relationships between all input parts, regardless of their position. The result of these comparisons is an attention score for every other part in the input, which determines their contributions. The transformer network architecture (shown in Fig.2(a)) consists of encoders and decoders. The encoders are stacked in the transformer model as shown in Fig.2(b). Each encoder has two sub-layers, a multi-head self-attention mechanism, and a position-wise fully connected feed-forward network. There exists a residual connection around both sub-layers, followed by a normalization layer. The decoder is not used here.

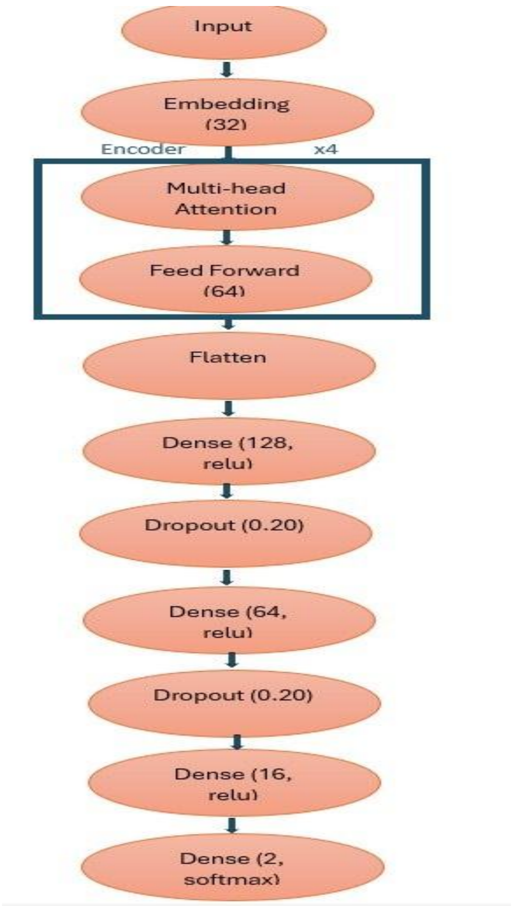
Embedding and positional encoding are done before sending the input to the encoders. Encoders operate with vectors and therefore, the input is converted into embedding vectors using a fixed representation. Unlike recurrent neural networks (RNN) and LSTM, transformer net- works do not have a way to capture relative positions of the input. To provide this contextual information, positional encoding is used with each input vector. Positional encoding is not part of the architecture of the model but the pre-processing. The positional encoding vector is generated to be the same size as the embedding for each input. The positional encoding vector is added to the embedding vector after calculation. The 'injected' pattern into the embedding vector allows the algorithm to learn this spatial information.

3.2 Classification

For classification of the MAT dataset, the network architecture for"No Task vs. Arithmetic Task" is shown in Fig.2(b). The last layer is modified as Dense (2) for the "MAT softmax" network. For embedding, 21 dimensions are used with 4 attention heads. The hidden layer size of the feed-forward network in the transformer is 64. Adam is used as an optimizer with default parameters and binary cross-entropy is used for loss. MAT dataset is used as a 65-20-15 percent split for train, validation, and test sets during the experimentation. MAT dataset is evaluated for two classification scenarios. First, for "No Task vs. Arithmetic task', the dataset has been categorized into two workload levels, at No Task as 0 and performing Arithmetic Task as 1. This network (2 class) is trained for 100 epochs with early stopping also a batch size of 10. At middle hidden layer Relu activation function is used and at an output layer softmax activation function is used.



a) Transformer network architecture



b) Proposed architecture for MAT dataset

Figure. 2

Table 1. Classification of EEG data for Local (age and gender) and public Stew dataset.

References	Model Age and Gender dataset	Features	Classifier	Accuracy Gender (02 Class)
[34]	Kaur et al.	Statistical	Random Forest	96.66
[35]	Kaushik et al.	Frequency	BLSTM-LSTM	97.50
[2]	Gourav et al.	Not used	Transformer	94.53
Proposed work	Proposed work	Not used	Transformer	98.92

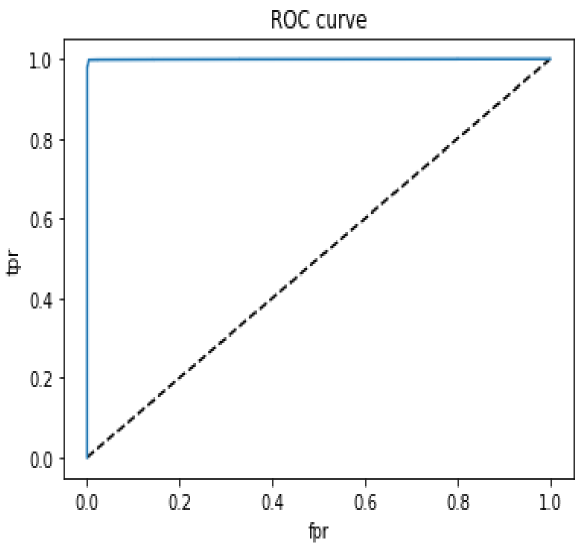
References	STEW dataset	Features	Classifier	Accuracy No Vs SIMKAP Task (02 Class)
[13]	Lim et al.	PSD, NCA	SVR	-
[25]	Zhu et al.	Frequency	DT, SVM	89.60
[24]	Chakladar et al. [	Frequency, Time	BLSTM-LSTM	86.33
[2]	Gourav et al.	Not Used	Transformer	95.28
Proposed work	Proposed work	Not used	Transformer	98.92

Table 2. Comparison of proposed method with existing state of art methods for the evaluation of different matrices

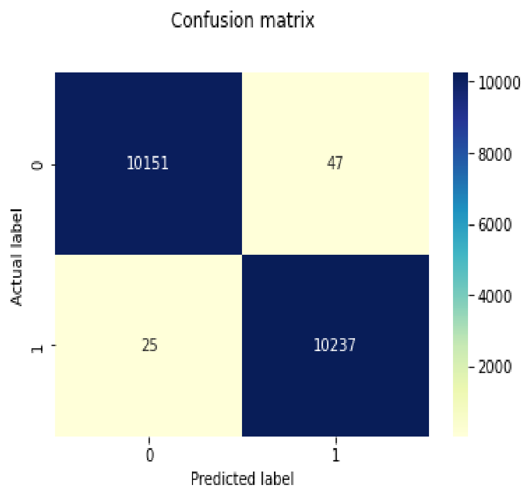
References	Model Age and Gender dataset	Method	Accuracy (02 Class)	Precision (02 Class)	Recall (02 Class)	F1-Score (02 Class)
[25]	Zhu et al.	CNN	86.38	87.72	80.21	82.64
[27]	Zheng and Chen et.al.	Bi-LSTM-AttGW	86.57	85.74	82.25	83.66
[2]	Gourav et al.	Transformer	94.53	94.39	92.82	93.55
Proposed work	Proposed work	Transformer	98.92	98.93	98.92	98.92

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 19, Number 4 (2024) pp. 280-288
https://dx.doi.org/10.37622/IJAER/19.4.2024.280-288

References	**STEW dataset**	**Method**	**Accuracy (02 Class)**	**Precision (02 Class)**	**Recall (02 Class)**	**F1-Score (02 Class)**
[25]	Zhu et al.	CNN	93.73	93.74	93.73	93.73
[27]	Zheng and Chen et al.	Bi-LSTM-AttGW	94.02	94.02	94.01	94.02
[2]	Gourav et al.	Transformer	95.28	95.29	95.28	95.28
Proposed work	Proposed work	Transformer	98.92	98.93	98.92	98.92



a) ROC for No Task Vs Mental Task



b) Confusion Matrix for No Task Vs Mental Task

Figure. 3

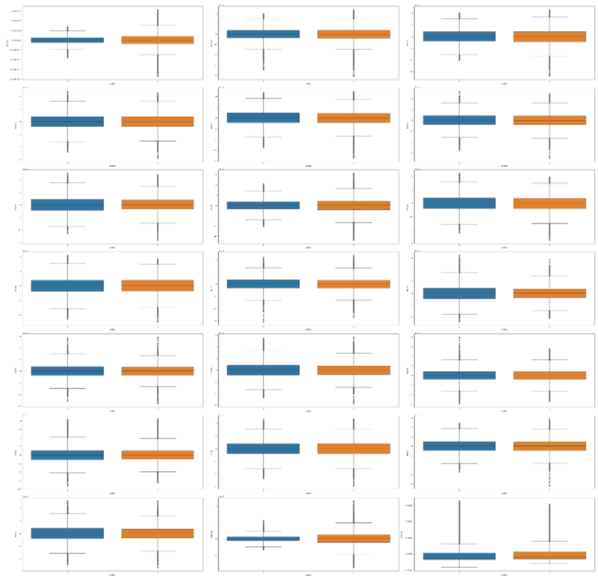
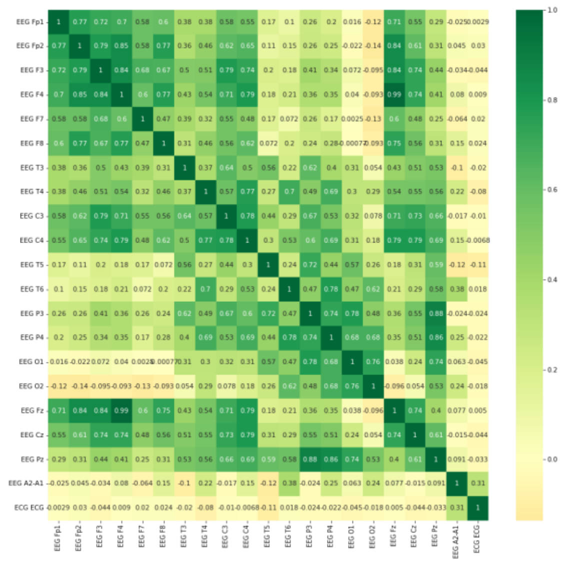


Figure. 4 Electrode variation for different task (No Task Vs Mental Task)



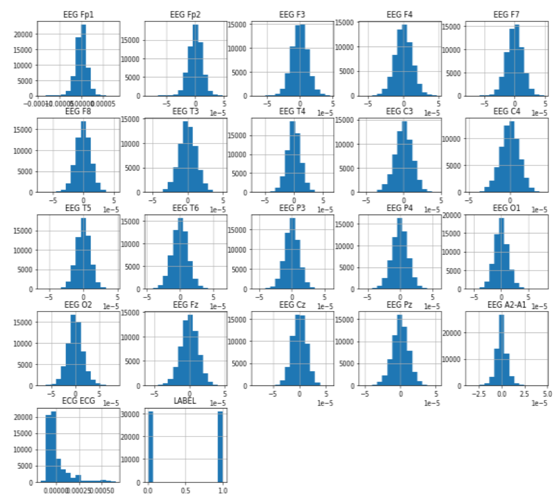


Figure. 5 Effect on different brain lobe while doing No task & Arithmetic task

For the “Arithmetic task” condition, from the Fig.5, we observe that mental activity is concentrated around the different regions of the brain such as anterior frontal, frontal, temporal, and central Fp1, Fp2, F3, F4, F7, and F8, T3 and T4, C3 and C4, with the values of 0.6- 1.0

For the “No task” condition, from Fig.5, we observe that rest activity is concentrated around the different regions of the brain such as temporal, parietal, occipital, T5 and T6, P3 and P4, O1 and O2 with the values of 0.0-0.4.

4. Results and Discussion

This section discusses the comparative analysis between the proposed model and other approaches, including machine learning and popular deep neural networks. The results are summarized in Table 1, confusion matrix is shown in Fig. 3(b) and receiver operating characteristics (ROC) is shown in Fig. 3(a), and comparison with state of art methods with different evaluation matrices is summarized in Table 2.

The proposed work make a use of transformer network (Fig. 2(b)) for classification and achieved test accuracy of 98.92% for ‘‘No Task vs. Arithmetic Task’’ classification on raw EEG data. The corresponding confusion matrix is shown in Fig. 3(b) and ROC is shown in Fig. 3(a). Transformer network trained with pre-processed EEG data clearly out-performs the existing methods with feature extraction for the classification scenarios on the MAT dataset (Fig. 1(a) and (b)). One of the reasons could be the effective learning of embedding and positional encoding used which provides contextual information.

For the classification purpose highly, successful networks proposed by Zhu et al. [25] and Zheng and Chen [27] are used. These are trained on MAT dataset used in this work. The accuracy achieved by the proposed model achieved better results than both the methods for MAT datasets. The improved accuracy might result from multi-head attention, which learns better and faster than a single attention layer used in these networks, resulting in high performance. The comparison of results with these attention-based networks is summarized in Table 2. It is observed that the proposed transformer network for EEG outperforms existing deep learning networks on raw EEG data.

Conclusion

This study explored the use of transformer networks for the classification of raw (filtered data) EEG data. To evaluate such models’for the accuracy purpose, proposed network’s performance was tested on a public dataset (MAT). The robustness of the classifier was tested on MAT dataset and had an Accuracy of 98.92% (two workload levels).

The proposed method gave state-of-the-art accuracy for MAT dataset for the classification of EEG signal (two workload levels). This proves the effectiveness of the transformers for processing raw EEG data (Filtered data) and can reduce the need for feature extraction in EEG as deep learning has done in other domains such as image processing and signal processing. Transformer networks seem to be effective in classifying EEG data as well. They might also reduce the over-reliance on the feature extraction still used in the EEG domain.

Acknowledgements

We would like to give thanks and acknowledge to all research scholars of Electronics and Telecommunication Engineering Department SGGSIE&T Nanded.

Conflict of interest

None

Data Availability

The dataset is publicly available (online Kaggle).

References List

[1] Zhou, Yueying, et al. "Cognitive workload recognition using EEG signals and machine learning: A review." *IEEE Transactions on Cognitive and Developmental Systems* 14.3 (2021): 799-818.

[2] Siddhad, Gourav, et al. "Efficacy of transformer networks for classification of EEG data." *Biomedical Signal Processing and Control* 87 (2024): 105488.

[3] Seal, Ayan, et al. "DeprNet: A deep convolution neural network framework for detecting depression using EEG." *IEEE Transactions on Instrumentation and Measurement* 70 (2021): 1-13.

[4] Chen, Xiaogang, et al. "Combination of high-frequency SSVEP-based BCI and computer vision for controlling a robotic arm." *Journal*

of neural engineering 16.2 (2019): 026012.

[5] Dousset, Clémence, et al. "Preventing relapse in alcohol disorder with EEG-neurofeedback as a neuromodulation technique: a review and new insights regarding its application." *Addictive behaviors* 106 (2020): 106391.

[6] Zheng, Wei-Long, and Bao-Liang Lu. "A multimodal approach to estimating vigilance using EEG and forehead EOG." *Journal of neural engineering* 14.2 (2017): 026017.

[7] Sadiq, Muhammad Tariq, et al. "Toward the development of versatile brain–computer interfaces." *IEEE Transactions on Artificial Intelligence* 2.4 (2021): 314-328.

[8] Pérez, A., M. Carreiras, and J. A. Dunabeitia. "Brain-to-brain entrainment: EEG interbrain synchronization while speaking and listening. Sci. Rep. 7, 4190." (2017).

[9] Henry, J. Craig. "Electroencephalography: basic principles, clinical applications, and related fields." *Neurology* 67.11 (2006): 2092-2092.

[10] Hämäläinen, Matti, et al. "Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain." *Reviews of modern Physics* 65.2 (1993): 413.

[11] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." nature 521.7553 (2015): 436-444.

[12] Gao, Zhongke, et al. "Complex networks and deep learning for EEG signal analysis." Cognitive Neurodynamics 15.3 (2021): 369-388.

[13] Lim, Wei Lun, Olga Sourina, and Lipo P. Wang. "STEW: Simultaneous task EEG workload data set." IEEE Transactions on Neural Systems and Rehabilitation Engineering 26.11 (2018): 2106-2114.

[14] Qu, Hongquan, et al. "Mental workload classification method based on EEG independent component features." Applied Sciences 10.9 (2020): 3036.

[15] Santiago-Espada, Yamira, et al. The multi-attribute task battery ii (matb-ii) software for human performance and workload research: A user's guide. No. L-20031. 2011.

[16] Bagheri, Mahsa, and Sarah D. Power. "EEG-based detection of mental workload level and stress: the effect of variation in each state on classification of the other." Journal of Neural Engineering 17.5 (2020): 056015.

[17] Pei, Zian, et al. "EEG-based multiclass workload identification using feature fusion and selection." IEEE Transactions on Instrumentation and Measurement 70 (2020): 1-8.

[18] Gupta, Shankar S., and Ramchandra R. Manthalkar. "Classification of visual cognitive workload using analytic wavelet transform." Biomedical Signal Processing and Control 61 (2020): 101961.

[19] Hefron, Ryan G., et al. "Deep long short-term memory structures model temporal dependencies improving cognitive workload estimation." *Pattern Recognition Letters* 94 (2017): 96-104.

[20] Sun, Zhe, et al. "WLNet: Towards an approach for robust workload estimation based on shallow neural networks." *Ieee Access* 9 (2020): 3165-3173.

[21] Zhang, Yu, et al. "Temporally constrained sparse group spatial patterns for motor imagery BCI." *IEEE transactions on cybernetics* 49.9 (2018): 3322-3332.

[22] Lawhern, Vernon J., et al. "EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces." *Journal of neural engineering* 15.5 (2018): 056013.

[23] Saha, Anushri, et al. "Classification of EEG signals for cognitive load estimation using deep learning architectures." *Intelligent Human Computer Interaction: 10th International Conference, IHCI 2018, Allahabad, India, December 7–9, 2018, Proceedings 10*. Springer International Publishing, 2018.

[24] Chakladar, Debashis Das, et al. "EEG-based mental workload estimation using deep BLSTM-LSTM network and evolutionary algorithm." *Biomedical Signal Processing and Control* 60 (2020): 101989.

[25] Zhu, Guohun, et al. "Cognitive load during multitasking can be accurately assessed based on single channel electroencephalography using graph methods." *IEEE Access* 9 (2021): 33102-33109.

[26] Zhu, Tianqi, Wei Luo, and Feng Yu. "Convolution-and attention-based neural network for automated sleep stage classification." *International Journal of Environmental Research and Public Health* 17.11 (2020): 4152.

[27] Zheng, Xiao, and Wanzhong Chen. "An attention-based bi-LSTM method for visual object classification via EEG." *Biomedical Signal Processing and Control* 63 (2021): 102174.

[28] Dosovitskiy, Alexey, et al. "An image is worth 16x16 words: Transformers for image recognition at scale." *arXiv preprint arXiv:2010.11929* (2020).

[29] Jiang, Yifan, Shiyu Chang, and Zhangyang Wang. "Transgan: Two transformers can make one strong gan." *arXiv preprint arXiv:2102.07074* 1.3 (2021).

[30] Bagheri, Mahsa, and Sarah D. Power. "EEG-based detection of mental workload level and stress: the effect of variation in each state on classification of the other." *Journal of Neural Engineering* 17.5 (2020): 056015.

[31] Krishna, Gautam, et al. "EEG based continuous speech recognition using transformers." *arXiv preprint arXiv:2001.00501* (2019).

[32] Zyma, Igor, et al. "Electroencephalograms during mental arithmetic task performance." *Data* 4.1 (2019): 14.

[33] Vaswani, Ashish. "Attention is all you need." *arXiv preprint arXiv:1706.03762* (2017).

[34] Kaur, Barjinder, Dinesh Singh, and Partha Pratim Roy. "Age and gender classification using brain–computer interface." *Neural Computing and Applications* 31.10 (2019): 5887-5900.

[35] Kaushik et al ‘Model, Using Deep BLSTM-LSTM Network. "EEG-based Age and Gender Prediction Using Deep BLSTM-LSTM Network Model."