Decoding Visual Stimuli and Visual Imagery Information from EEG Signals Utilizing Multi-Perspective 3D-CNN Based Hierarchical Deep-Fusion Learning Network

Fatma Yusranur Emanet and Kazim Sekeroglu
Department of Computer Science, Southeastern Louisiana University, Hammond, LA 70402, U.S.A.

Keywords: Hierarchical Deep Learning, Brain Computer Interface, Fusion Learning, Spatiotemporal Pattern Recognition, Multi-Perspective Learning.

Abstract: Brain-Computer Interface Systems (BCIs) facilitate communication between the brain and machines, enabling applications such as diagnosis, understanding brain function, and cognitive augmentation. This study explores the classification of visual stimuli and visual imagery using electroencephalographic (EEG) data. The proposed method utilizes 3D EEG data generated by transforming 1D EEG data into 2D Spatiotemporal EEG image mappings for feature extraction and classification. Additionally, a multi-perspective 3D CNN-based hierarchical deep fusion learning network is employed to classify multi-dimensional spatiotemporal EEG data, decoding brain activity for visual and visual imagery stimulation. The findings show that the suggested multi-perspective fusion method performs better than a standalone model, indicating promising progress in using BCIs to understand and utilize brain signals for visual and imagined stimulation.

1 INTRODUCTION

The brain-computer interface systems (BCIs) are one of the crucial technologies in recent years that aim to establish communication between the brain and machines. Besides the use of BCI systems in many scientific research areas, the main purpose of BCI systems is to enable people to develop applications where they can control various devices including computers, prosthetic limbs, robots, and even video games by using only power of human thought (Lebedev, 2017).

BCI is a system that deals with the brain activities of a living thing (human or animal) and turns these activities into meaningful information about the cognitive, perceptual, or motor processes associated with neural activity patterns. This process is also known as brain decoding. Meaningful knowledge obtained thanks to brain decoding can be used for studies such as developing brain-computer interfaces, diagnosing disorders, understanding human brain function, and even augmenting cognition (Tan, 2010).

In recent years, BCI technologies have started to show their presence in fields such as medicine, neuroscience, and gaming and are used for revolutionary innovations in these fields. Especially thanks to the BCI innovations made in the medical world, many people with disabilities and limited mobility have started to meet their needs without the need for any physical activity (Miralles, 2015).

With many applications developed so far, this field frequently updates itself and is very open to developments. Therefore, it has the potential to understand the brain and its working principles, which is increasing day by day. This potential has attracted the attention of scientists and it has recently become a hot topic in the world of science and technology.

We can give examples of these applications; communication devices for people with disabilities (Millán, 2010), controlling devices in hazardous environments (Douibi, 2021), enhancing cognitive performance (Papanastasiou, 2020), prosthetic limbs that can be controlled by the user's thoughts (Vilela, 2020), and even brain-controlled video games (Nijholt, 2009). In addition to many detailed and successful studies conducted in this area, developing reliable and robust decoding algorithms, and obtaining consistent neural activity patterns by brain decoding is still a challenge today due to some reasons related to the brain such as the complexity of the brain signals due to its nature, its dynamic

Emanet, F. and Sekeroglu, K.
Decoding Visual Stimuli and Visual Imagery Information from EEG Signals Utilizing Multi-Perspective 3D-CNN Based Hierarchical Deep-Fusion Learning Network.
DOI: 10.5220/0012568500003660
Paper published under CC license (CC BY-NC-ND 4.0)
In *Proceedings of the 19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2024) - Volume 4: VISAPP*, pages 381-388
ISBN: 978-989-758-679-8; ISSN: 2184-4321
Proceedings Copyright © 2024 by SCITEPRESS – Science and Technology Publications, Lda.

structure, and being affected by environmental factors.

Brain decoding can be used for visual stimuli classification. Visual stimuli classification refers to the process of identifying the category or features of a visual stimulus such as an image or video clip. It uses the response of the brain which is the patterns of neural activity that stimuli evoke in the brain for identifying the category of a visual stimulus (Bigdely-Shamlo, 2008).

Using various machine learning techniques, models with high performance can be created and successful results can be obtained to classify visual stimuli. These models can predict the category of a new, unseen visual stimulus based on the feature map which is most relevant for the cognitive task at hand by training the algorithm with known categories of visual stimuli, such as images of faces, letters, or simple shapes (Aggarwal, 2022).

This study aims to develop a model that utilizes brain decoding for the purpose of classifying not only visual stimuli but also visual imagery. Specifically, EEG data collected from participants in an experimental setup designed for this study will be utilized for classification purposes.

Brain-computer interfaces are known for their potential to provide solutions to a wide range of issues in both scientific and everyday contexts. The primary objective of this project was to address a research-based problem related to the classification of EEG signals within the context of brain-computer interfaces. While this project was focused on addressing this specific issue, the insights and findings obtained through this work could be applied to a broader range of problems in various domains. This study contributes by analyzing EEG signals generated by visual stimuli and visual imagery. This involves using a 2D spatiotemporal EEG image representation, investigating 3D EEG data for feature extraction, and classifying based on 2D spatiotemporal EEG (ST-EEG) maps. The approach also incorporates a 3D convolutional neural network (CNN) based multi-perspective hierarchical deep fusion model for the classification of 3D EEG representations of visually evoked and visual imagery signals.

1.1 Literature Review

Brain-computer interface has been a hot topic in the world of science and technology in recent years. Although it is a topic that is widely talked about today, it is known that studies on BCI were first studied with animals in the 1970s (Kawala-Sterniuk, 2021). The latest studies in BCI have concentrated on how to improve the accuracy and speed of brain signal decoding. With many machine learning methods, particularly deep learning, brain signals are analyzed and examined. As seen in a study by Zhang et al., the use of deep learning has noticeably increased the accuracy in classifying different types of brain signals (Sun, 2020).

Another option provided by BCI systems is the classification of brain signals obtained using visual stimuli. Allison, B. Z., et al. (Jin, 2012) focused on the use of a BCI system to change frequency bands in EEG signals via visual stimuli. The scientists discovered that, with their claimed approach, participants were able to effectively change their brain signals. In another study that used visual stimuli, Kavasidis et al. (Kavasidis, 2017) studied the translation of visually evoked EEG signals into meaningful images. The approach they proposed is based on generating images using visually evoked brain signals recorded through an electroencephalograph (EEG). They implemented a deep learning framework consisting of an LSTM stacked with a generative method. They pointed out that GAN, in general, outperforms VAE and recommended that the study should continue by combining these two. They also recommended acquiring fMRI data to complement EEG data.

Similarly, Hayashi and Kawata (Hayashi, 2018) proposed a methodology to reconstruct the images which had been recorded from the monkey brain. They implemented a linear decoder that predicts visual features of viewed images at a higher-order layer of a deep convolutional neural network, so called CaffeNet (Jia, 2014). They refined the images to photorealistic images through a deep generator network (Dosovitskiy, 2016). Their approach lacks efficient choosing critical visual features for the subject for image reconstruction within a reasonable time frame.

Liu, Shuang, et al. (Liu, 2014) explored the individual identification through the extraction of features from both resting EEG and visual evoked potential signals. The features in this identification consisted of fourth-order AR parameters, power spectrum in the time and frequency domain, and phase locking value. For the classification, the extracted features were fed into an SVM. Thanks to this study, as a result of the identifications made with features, they come with the result that visually evoked tasks show better results in identifying individuals compared to relax tasks. Tirupattur and Rawat (Tirupattur, 2018) introduced a GAN architecture to generate class-specific images from brain activities, achieving good results with small datasets and they emphasize the potential for visualizing brain signals as a video stream. Additionally, Zhang et al. (Zhang, 2019) also used GANs to reconstruct shapes evoked by EEG signals, focusing on simple geometrical shapes and suggesting future exploration of more complex shapes. They also performed a feature extraction from EEG data using CNN.

In another GAN-based study, Fares and Zhong (Fares, 2020) proposed a novel DCLS-GAN framework to integrate brain and visual features, transforming EEG descriptions into class-relevant images. Similarly, Wang et al. (Wang, 2021) introduce the α-GAN approach that combines standard GAN structure with variational auto-encoder to reconstruct images from the EEG framework and the fMRI framework. Rashkov et al. (Rashkov, 2019) offered closed-loop BCI system that reconstructs the observed or imagined stimuli images from the co-occurring brain wave parameters. This paradigm contains the visual-based cognitive test for individual stimuli set selection as well as state-of- art deep learning-based image reconstruction model for native feedback presentation. Additionally, Qu et al. (Qu, 2021) suggested an algorithmic idea of extracting, selecting, and decoding the EEG features related with the stimuli based on the supervision of the decoding feature of the original stimulus image. In this study, they pointed out the lack of clear evidence to prove that humans are in visual processing tasks.

Similarly, Palazzo et al. (Palazzo, 2020) proposes a model, EEG-ChannelNet, to learn a brain manifold for EEG classification. And that, they introduce a multimodal approach that uses deep mage and EEG encoders, trained in a Siamese configuration, to learn a joint feature space for images and EEG signals recorded while users look at pictures on a screen. They trained two encoders in a siamese configuration and maximize the compatibility score between the corresponding images and EEGs. They also pointed out that identifying different responses in brain activity corresponding to different objects, patterns, or categories is a field for future study in their study.

For EEG-based brain imaging classification, Jiang et al. (Jiang, 2019) proposed a novel deep framework. The proposed framework provides multimodal brain imaging classification by using not only the strength of integrated multiple modalities but also the advantages of the added consistency test. Additionally, Spampinato et al. (Spampinato, 2017) introduce a deep learning approach to classify EEG data as well as propose the first automated classification approach employing visual descriptors extracted directly from human neural processes involved in visual scene analysis. They emphasize that there will be a greater need for very complex deep learning networks in the future and that the studies conducted in this direction will be used to distinguish brain signals produced from many image classes. Fares et al. (Fares, 2019) create a novel region-level stacked bi-directional deep learning framework for visual object classification. In this framework, there are 3 stages including the region-level information extraction stage, the feature encoding stage, and the classification stage. In their study, they predict that multimedia content information can be reconstructed through the proposed EEG representations for future studies.

2 METHODOLOGY

2.1 Dataset

In this study, EEG was used to capture brain activity patterns. A new dataset was created by collecting EEG data from seven volunteer participants using "Enobio 32" device by "NEUROELECTRICS," a medical company specializing in non-invasive brain stimulation ((n.d.), 2023). This device offers 32 electrodes and the option of dry electrodes for quick setup. The sampling rate is fixed at 500. A 10-10 international electrode placement system is used in the device (Krol, 2020).

For data collection, we gathered data from 7 participants, following a specific data collection procedure. Participants were seated in a well-lit, quiet room, ensuring minimal distractions. Positioned approximately 50 cm away from an LCD computer screen, the participants were instructed to remain seated throughout the duration of the experiment, with an EEG cap placed on their heads to record brain waves. The experiment encompassed two sections conducted on the same day. During each section, the EEG recorder captured the participants' brain waves while a slideshow randomly displayed stimulus. Throughout the experiment, the EEG cap remained on the participants' heads for consistency between the two sections. The stimuli presented in both sections consisted of the letters A, B, and C displayed in random order. In the first section, participants were directed to focus on the computer screen with their eyes open for 30 seconds, followed by 30 seconds of eyes-closed rest. Subsequently, the letters A, B, and C appeared on the screen for 10 seconds each, preceded by a 1.5-second interval of a black screen to minimize the influence of the previous letter. This sequence was repeated 20 times for each letter in random order. In the second section, participants were presented with a blank white screen throughout. An auditory cue prompted participants to mentally visualize a specific letter, which they maintained for 10 seconds. They were then instructed to imagine another randomly selected letter for another 10 seconds. This process was repeated 20 times for each of the letters, A, B, and C, resulting in multiple instances of imagined letters for each participant. The process of each experiment section is depicted in Fig. 1 and Fig. 2.

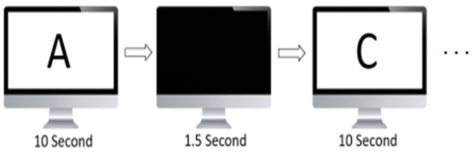


Figure 1: The First Experiment Section-Visual Stimuli Phase.

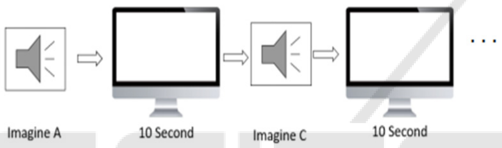


Figure 2: The Second Experiment Section-Visual Imagery Phase.

2.2 Data Transformation

The raw EEG data comprises one-dimensional time series data for each channel, which reflects the electrical activity in specific locations (referred to as channels) of the brain over time. The recording device utilized for gathering brain signals features 32 distinct electrodes, making the complete raw data manifest as a 2D matrix. This matrix encompasses 1D time series data encapsulating the electrical activities for 32 different locations. Refer to Fig. 5 for visual representation. In this context, the matrix 'S' constitutes a two-dimensional array that encompasses all the EEG data collected from 'n' channels over a duration of 't+N' time. Each row within the matrix corresponds to a specific channel for the entire duration, while the columns signify the EEG data recorded from all channels at a specific time, denoted as 't'.

In this proposed method, we took into consideration the potential impact of channel proximity and neighbourhood relations on the spatiotemporal plane. To achieve this, we transformed the data shape into 2D spatiotemporal

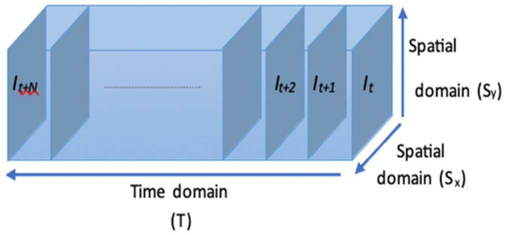


Figure 3: Creation of 2D Spatiotemporal EEG image sequence.

EEG mappings. This transformation essentially depicts each signal as if a top-down image of the brain was captured. Consequently, 2D spatiotemporal EEG maps were generated for each signal, effectively producing a three-dimensional data set. This was achieved by arranging these generated maps consecutively along the temporal plane. In essence, the altered EEG data is now represented as 2D spatiotemporal EEG maps, culminating in a 3D dataset with two dimensions in the spatial domain and one dimension in the time domain. Refer to Fig. 3 for a visual depiction of this representation. EEG data from all channels was mapped to a 9x9 matrix based on the precise locations of the electrodes on the scalp where the data were recorded. Channel locations in the 9X9 matrix are illustrated in Fig. 4 below.

nan	nan	nan	FP1	nan	FP2	nan	nan	nan
nan	nan	AF3	nan	nan	nan	AF4	nan	nan
F7	nan	F3	nan	FZ	nan	F4	nan	F8
nan	FC5	nan	FC1	nan	FC2	nan	FC6	nan
T7	nan	C3	nan	CZ	nan	C4	nan	T8
nan	CP5	nan	CP1	nan	CP2	nan	CP6	nan
P7	nan	P3	nan	PZ	nan	P4	nan	P8
nan	nan	PO3	nan	nan	nan	PO4	nan	nan
nan	nan	nan	O1	OZ	O2	nan	nan	nan

Figure 4: Channel Locations in the 9X9 Matrix.

The parts shown as nan in the 9x9 matrix represent the places where the electrodes don't exist. In order to capture the neighborhood relation, cubic interpolation was made for the empty ones between the neighboring electrodes, except for the corners. This transformation process and the transformed 2D ST-EEG map I at time stamp t, I_t, are illustrated in Fig. 5 below.

2.3 Proposed Model

All 3 different perspectives of 3D EEG data were taken into consideration and a 3D CNN model was

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

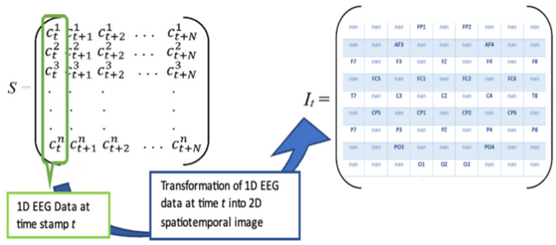


Figure 5: Transformation Process of 1D Temporal EEG Data into 2D Spatiotemporal Image.

created for each perspective. As given in Fig. 3, two dimensions of 3D EEG data correspond to the spatial domain, and the third dimension pertains to the temporal domain. The SxSy plane view offers insights into the collected data from all channels at time t, while the TSx and TSy planes provide information regarding the collected data from specific channels over a time period. Hence, this study explores three distinct views: the main view based on the SxSy plane, the second view based on the TSx plane, and the third view based on the TSy plane.

In the initial phase of the proposed fusion architecture, 3D CNN networks are employed to identify patterns from the specific viewpoint of the multidimensional spatiotemporal EEG data. Subsequently, at the second stage of the architecture, the identified patterns from the 3D CNN networks for the second and third viewpoints are merged and forwarded to the subsequent layer for further fusion. The central concept is to integrate the patterns of EEG data that share temporal information derived from the second and third viewpoints. Finally, at the concluding layer of the fusion architecture, the extracted pattern from the primary perspective and the temporal fusion layer are consolidated and merged via the output layer to achieve the final spatiotemporal fusion. Fig. 6 below provides a block diagram of the proposed multi-view hierarchical deep learning model.

As seen in Fig. 6 above, there are three 3D CNN models in the first layer of the proposed hierarchical model: main view, side view-1 and side view-2. Detailed architecture of the 3D CNN model used in main view and side views are shown in Fig. 7 below. Whereas Fig. 8 shows the architecture of the temporal fusion model and spatiotemporal fusion model. In this study, we used multi-layer perceptron for the temporal and spatiotemporal fusion, However, different supervised machine learning models such as Support Vector Machine, Bayesian Network, Decision Tree, or a multidimensional classification model can be used.

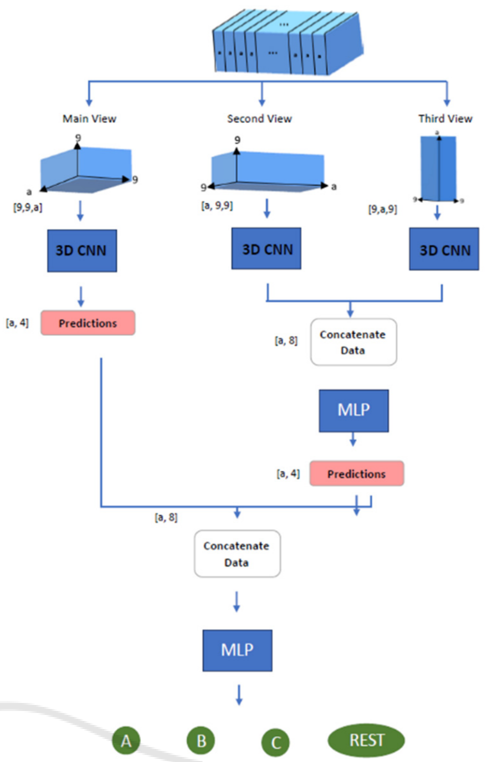


Figure 6: Block diagram of the proposed multi-perspective hierarchical deep learning model.

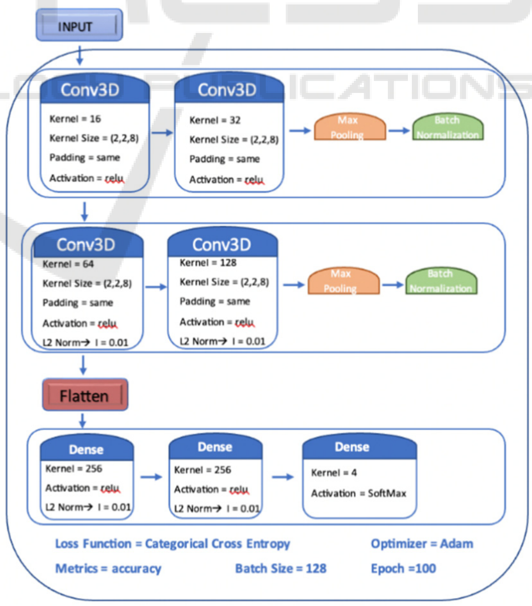


Figure 7: Main view and Side view model blog diagram.

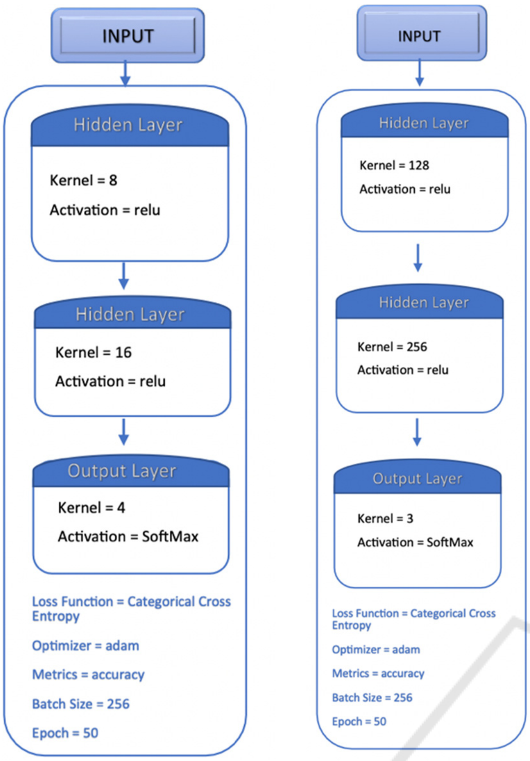
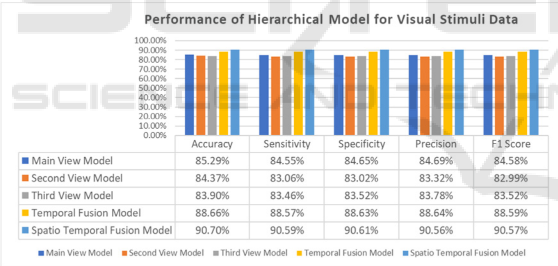
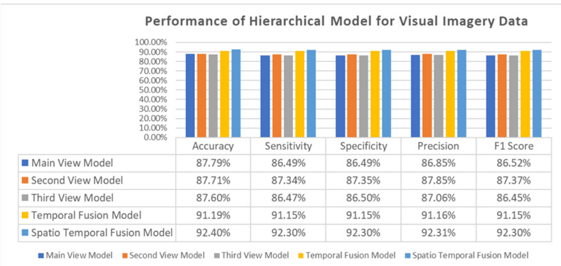


Figure 8: Temporal Fusion Model (left), Spatiotemporal Fusion Model (right).



	Accuracy	Sensitivity	Specificity	Precision	F1 Score
Main View Model	85.29%	84.55%	84.65%	84.69%	84.58%
Second View Model	84.37%	83.06%	83.02%	83.32%	82.99%
Third View Model	83.90%	83.46%	83.52%	83.78%	83.52%
Temporal Fusion Model	88.66%	88.57%	88.63%	88.64%	88.59%
Spatio Temporal Fusion Model	90.70%	90.59%	90.61%	90.56%	90.57%

Figure 9: Performance measures of Hierarchical Model for Visual Stimuli Data.



	Accuracy	Sensitivity	Specificity	Precision	F1 Score
Main View Model	87.79%	86.49%	86.49%	86.85%	86.52%
Second View Model	87.71%	87.34%	87.35%	87.85%	87.37%
Third View Model	87.60%	86.47%	86.50%	87.06%	86.45%
Temporal Fusion Model	91.19%	91.15%	91.15%	91.16%	91.15%
Spatio Temporal Fusion Model	92.40%	92.30%	92.30%	92.31%	92.30%

Figure 10: Performance measures of Hierarchical Model for Visual Imagery Data.

3 RESULTS AND DISCUSSION

In our initial experimentation, we evaluated the proposed modular fusion learning architecture, employing a 3D CNN-based network to extract patterns from 3D spatiotemporal EEG data. The outcomes of our assessment are depicted in Figure 9 which presents the classification results for visual stimuli, and Figure 10, illustrating the classification outcomes for the imagined visual stimuli. For the visual stimuli classification, we utilized simple geometric representations of the letters A, B, and C, while for the imagined visual stimuli, we instructed the subjects to mentally visualize the corresponding letters. As evident from the results, the Multi-perspective and hierarchical fusion learning approach substantially enhanced the classification accuracy.

In our preliminary findings, the utilization of the main view alone yielded an accuracy of 85.29% for visual stimuli and 87.79% for the imagined visual stimuli. However, the integration of the multi-perspective and hierarchical fusion model led to a notable increase in accuracy, reaching 90.7% for visual stimuli and 92.4% for the imagined visual stimuli. These results suggest that the proposed model is proficient in extracting patterns from multidimensional spatiotemporal EEG data, making it suitable for classification purposes and the generation of both provided and imagined visual information from EEG data.

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

4 CONCLUSION AND FUTURE WORK

The current research introduces hierarchical deep learning models that were trained to recognize patterns in 2D spatiotemporal EEG images, for the purpose of classifying visual stimuli and visual imagery data. The results of the study show that the proposed model achieved strong performance in multi-class classification of 3D EEG data.

The investigation placed a significant emphasis on analyzing visually evoked EEG signals and visual imagery EEG signals, utilizing a 2D spatiotemporal EEG image representation, and aimed to extract features and perform classification based on these representations.

The study also aimed to explore the benefits of a fusion architecture and a multi-view approach in learning the 2D ST-EEG maps. The experimental results indicate that, in general, a fusion architecture outperforms a stand-alone model.

While BCI technology has gained a lot of attention in recent years, there are still many unanswered questions and areas that require further research. This study can be seen as a preliminary study for many potential studies to be conducted in the future. By addressing the gaps in the current literature, researchers can build upon the findings of this study and expand our knowledge of BCI technology, potentially leading to new and innovative applications in the future.

Initially, the focus of this study was on identifying four specific classes, namely A, B, C, and Rest. However, in future studies, the potential exists to expand the number of classes to encompass the entirety of the alphabet.

Moreover, we employed a common dataset that was divided into training and testing sets in our study. However, there is potential for further exploration on how to improve the detection performance of models trained on data collected at different times when tested on data gathered at a later point. By doing so, real-time applications can be developed, particularly for individuals with communication difficulties who may lack the ability to speak and could benefit from a system that allows them to communicate their words through BCI technology.

ACKNOWLEDGEMENTS

Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant number P20GM103424-21.

REFERENCES

Lebedev, M. A., & Nicolelis, M. A. (2017). Brain-machine interfaces: From basic science to neuroprostheses and neurorehabilitation. Physiological reviews, 97(2), 767-837.

Tan, D., & Nijholt, A. (2010). Brain-computer interfaces and human-computer interaction (pp. 3-19). Springer London.

Miralles, F., Vargiu, E., Dauwalder, S., Solà, M., Müller-Putz, G., Wriessnegger, S. C., ... & Lowish, H. (2015). Brain computer interface on track to home. The Scientific World Journal, 2015.

Millán, J. D. R., Rupp, R., Mueller-Putz, G., Murray-Smith, R., Giugliemma, C., Tangermann, M., ... & Mattia, D. (2010). Combining brain–computer interfaces and assistive technologies: state-of-the-art and challenges. Frontiers in neuroscience, 161.

Douibi, K., Le Bars, S., Lemontey, A., Nag, L., Balp, R., & Breda, G. (2021). Toward EEG-based BCI applications for industry 4.0: challenges and possible applications. Frontiers in Human Neuroscience, 15, 705064.

Papanastasiou, G., Drigas, A., Skianis, C., & Lytras, M. (2020). Brain computer interface based applications for training and rehabilitation of students with neurodevelopmental disorders. A literature review. Heliyon, 6(9), e04250.

Vilela, M., & Hochberg, L. R. (2020). Applications of brain-computer interfaces to the control of robotic and prosthetic arms. Handbook of clinical neurology, 168, 87-99

Nijholt, A. (2009). BCI for games: A 'state of the art'survey. In Entertainment Computing-ICEC 2008: 7th International Conference, Pittsburgh, PA, USA, September 25-27, 2008. Proceedings 7 (pp. 225-228). Springer Berlin Heidelberg.

Bigdely-Shamlo, N., Vankov, A., Ramirez, R. R., & Makeig, S. (2008). Brain activity- based image classification from rapid serial visual presentation. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 16(5), 432-441.

Aggarwal, S., & Chugh, N. (2022). Review of machine learning techniques for EEG based brain computer interface. Archives of Computational Methods in Engineering, 1-20.

(n.d.). Enobio 32. Neuroelectrics. Retrieved April 16, 2023, from https://www.neuroelectrics.com/solutions/enobio/32.

Krol, L. R. (2020, November 25). EEG 10-10 system with additional information. Wikimedia. https://commons.wikimedia.org/wiki/File:EEG_10- 10_system_with_additional_information.svg

Kawala-Sterniuk, A., Browarska, N., Al-Bakri, A., Pelc, M., Zygarlicki, J., Sidikova, M., ... & Gorzelanczyk, E. J. (2021). Summary of over fifty years with brain-computer interfaces—a review. Brain Sciences, 11(1), 43.

Sun, B., Zhao, X., Zhang, H., Bai, R., & Li, T. (2020). EEG motor imagery classification with sparse spectrotemporal decomposition and deep learning. IEEE Transactions on Automation Science and Engineering, 18(2), 541-551.

Jin, J., Allison, B. Z., Wang, X., & Neuper, C. (2012). A combined brain–computer interface based on P300 potentials and motion-onset visual evoked potentials. Journal of neuroscience methods, 205(2), 265-276.

Kavasidis, I., Palazzo, S., Spampinato, C., Giordano, D., & Shah, M. (2017, October). Brain2image: Converting brain signals into images. In Proceedings of the 25th ACM international conference on Multimedia (pp. 1809-1817).

Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., ... & Darrell, T. (2014, November). Caffe: Convolutional architecture for fast feature embedding. In Proceedings of the 22nd ACM international conference on Multimedia (pp. 675- 678).

Dosovitskiy, A., & Brox, T. (2016). Generating images with perceptual similarity metrics based on deep networks. Advances in neural information processing systems, 29, 658-666.

Hayashi, R., & Kawata, H. (2018, October). Image reconstruction from neural activity recorded from monkey inferior temporal cortex using generative adversarial networks. In 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 105-109). IEEE

Liu, S., Bai, Y., Liu, J., Qi, H., Li, P., Zhao, X., ... & Ming, D. (2014). Individual feature extraction and identification on EEG signals in relax and visual evoked tasks. In Biomedical Informatics and Technology: First International Conference, ACBIT 2013, Aizu-Wakamatsu, Japan, September 16-17, 2013. Revised Selected Papers (pp. 305-318). Springer Berlin Heidelberg.

Tirupattur, P., Rawat, Y. S., Spampinato, C., & Shah, M. (2018, October). Thoughtviz: Visualizing human thoughts using generative adversarial network. In Proceedings of the 26th ACM international conference on Multimedia (pp. 950-958).

Zhang, X., Chen, X., Dong, M., Liu, H., Ge, C., & Yao, L. (2019). Multi-task Generative Adversarial Learning on Geometrical Shape Reconstruction from EEG Brain Signals. arXiv preprint arXiv:1907.13351.

Fares, A., Zhong, S. H., & Jiang, J. (2020, October). Brain-media: A Dual Conditioned and Lateralization Supported GAN (DCLS-GAN) towards Visualization of Image- evoked Brain Activities. In Proceedings of the 28th ACM International Conference on Multimedia (pp. 1764-1772).

Wang, P., Zhou, R., Wang, S., Li, L., Bai, W., Fan, J., ... & Guo, Y. (2021). A General Framework for Revealing Human Mind with auto-encoding GANs. arXiv preprint arXiv:2102.05236.

Rashkov, G., Bobe, A., Fastovets, D., & Komarova, M. (2019). Natural image reconstruction from brain waves: a novel visual BCI system with native feedback. bioRxiv, 787101.

Qu, L., Chen, D., & Yin, K. (2021, June). Research on EEG Feature Decoding Based on Stimulus Image. In 2021 IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC) (Vol. 4, pp. 467-470). IEEE.

Palazzo, S., Spampinato, C., Kavasidis, I., Giordano, D., Schmidt, J., & Shah, M. (2020). Decoding brain representations by multimodal learning of neural activity and visual features. IEEE Transactions on Pattern Analysis and Machine Intelligence.

Jiang, J., Fares, A., & Zhong, S. H. (2019). A context-supported deep learning framework for multimodal brain imaging classification. IEEE Transactions on Human- Machine Systems, 49(6), 611-622.

Spampinato, C., Palazzo, S., Kavasidis, I., Giordano, D., Souly, N., & Shah, M. (2017). Deep learning human mind for automated visual classification. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 6809-6817).

Fares, A., Zhong, S. H., & Jiang, J. (2019). EEG-based image classification via a region- level stacked bi-directional deep learning framework. BMC medical informatics and decision making, 19(6), 1-11.

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS