HIGH SCHOOL EDITION
Journal of Student Research
Volume 13 Issue 3 (2024)

Generalizing EEG-Based Classification for User-Independent Brain-Computer Interface

Ian Baek[1] and Sojung Min[1#]

[1]Korean Minjok Leadership Academy
[#]Advisor

ABSTRACT

The emergence of brain-computer interface technology has changed the way of interacting between people and the computer. By applying this technology to individuals with impairments, it is possible to help them regain their mobility. Consequently, exoskeleton robots, guided by electroencephalograms (EEG), have been studied to provide assistance to these individuals. However, previous methods have struggled to achieve accurate classification of user intentions, often displaying an excessive sensitivity to input noise. Thus, there is a need to develop methods that are robust to noise and yield highly accurate results. In this research, I proposed a noise robust system for classifying user intentions based on EEG signals. The proposed system takes EEG signals as input and outputs commands that guide exoskeleton robots in assisting individuals with impairments. These commands encompass a range of fundamental movements, including running, forward and backward walking, maintaining a stationary position, and more. Through comprehensive experiments, the results obtained by the proposed method substantiate its superiority over prior approaches. I expect that this method holds the potential to significantly aid individuals in need, particularly those with impairments or undergoing rehabilitation.

1. Introduction

1.1 Electroencephalography

EEG stands for Electroencephalography which is a non-invasive neuroimaging technique used to measure and record the electrical activity generated by the brain. This method involves placing electrodes on the scalp to detect and record the electrical signals produced by neurons (nerve cells) in the brain. These signals are often referred to as brainwaves. EEG is commonly used in clinical settings to diagnose and monitor various neurological conditions, such as epilepsy, sleep disorders, and brain injuries.

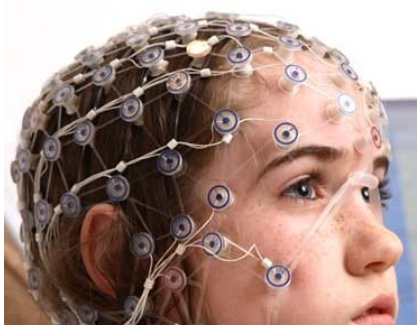

Figure 1. A device for capturing electroencephalography (EEG) signal

Figure 1 shows the process of capturing EEG signals through the use of electrodes. During an EEG recording, electrodes are attached to specific locations on the scalp using a cap or adhesive. The electrical activity is measured in the form of voltage fluctuations, which are then amplified and recorded. The resulting EEG data is analyzed to identify patterns and abnormalities.

As depicted in Figure 1, the EEG signal is acquired from multiple channels. An EEG channel refers to a specific recording from a single electrode or a group of electrodes placed on the scalp to capture electrical activity generated by the brain. In EEG recordings, multiple channels are used simultaneously to provide a comprehensive view of the brain's electrical patterns. Each EEG channel corresponds to the electrical activity detected by an individual electrode or a set of electrodes placed at specific locations on the scalp. Different channels may capture distinct aspects of brain activity. For example, frontal channels may be more relevant for cognitive processes, while occipital channels may be more sensitive to visual processing.

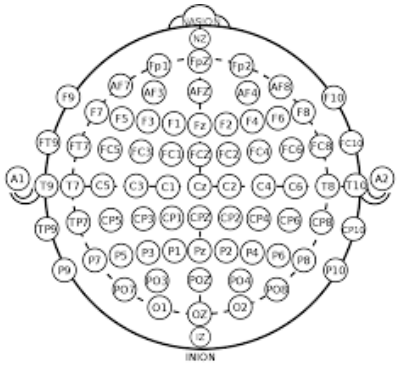

Figure 2. An illustration of various EEG channels

These EEG signals hold considerable utility and value as they allow us to discern the intentions behind the actions of individuals with disabilities. This information becomes instrumental in controlling exoskeletons, facilitating assistance in their movements. The goal is to translate these neural signals into actionable commands for the exoskeleton which enables individuals with disabilities to control the device through their brain activity.

In this study, to create such a system, I introduce an EEG-based machine learning classification system for action recognition. I present a practical and effective training approach. Leveraging representation learning, the proposed machine learning model is trained to capture and isolate action-related features, enhancing the precision and robustness of the classification process.

2. Related Work

2.1 EEG Signal Classification

Electroencephalography (EEG) signal classification is a significant aspect of neuroinformatics and brain-computer interface (BCI) research. EEG records the electrical activity of the brain through electrodes placed on the scalp, capturing neural signals associated with various cognitive and motor functions. The goal of EEG signal classification is to develop algorithms or models that can interpret these signals and categorize them into meaningful classes or actions.

Figure 3. An individual controls the external device using brain activity (Hinss et al. 2023)

As shown in Figure 3, EEG signal classification allows individuals to control external devices or software using their brain activity. This has applications in assistive technology for individuals with motor disabilities. The development of accurate and efficient classification models is crucial for enhancing the precision and reliability of brain-computer interfaces (BCIs) in real-world scenarios.

In this research, I have developed an action classification system based on EEG signals tailored for individuals with impairments. Through the implementation of this system, I expect that individuals with impairments will gain the capability to command external machines or devices, such as exoskeletons designed to assist them in their movements.

2.2 Unsupervised Learning

Unsupervised learning is a category of machine learning where the algorithm is tasked with identifying patterns, relationships, or structures within data without explicit guidance or labeled outcomes. In other words, unlike supervised learning where the algorithm is provided with labeled training data, unsupervised learning involves working with unlabeled data to uncover inherent patterns or intrinsic structures. This approach mainly focuses on capturing and extracting meaningful representations or features from raw, unlabeled data. The goal is to enable the model to learn a compact and informative internal representation of the data's underlying structure without the need for explicit supervision.

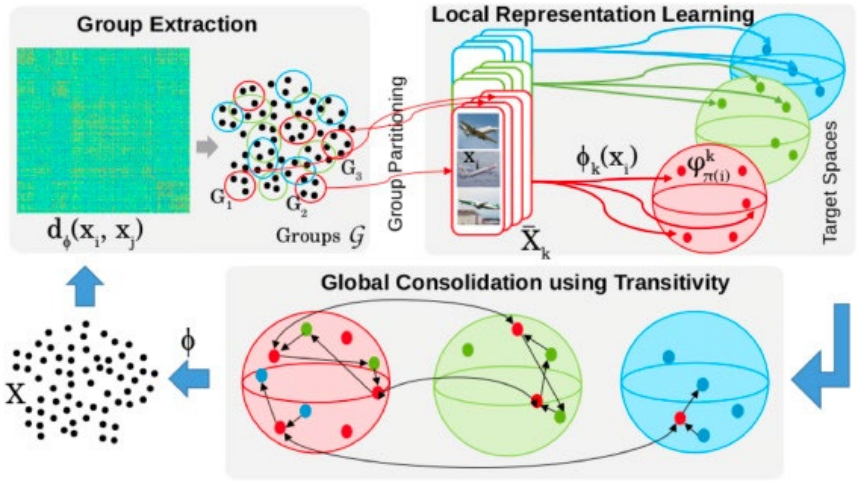

Figure 4. Explanation of representation learning (Milbich et al. 2020)

Representation learning, a subset of unsupervised learning, involves the automatic discovery of relevant features or patterns within the data. These meaningful features, often latent representations, are disentangled from the initially entangled feature space. This disentanglement enables the trained model to yield more accurate results in downstream tasks.

This research aims to disentangle action-related features from other elements extracted from the EEG signal. The trained model, incorporating the proposed representation learning, will be employed for action classification as a downstream task. Comparative assessments with previous methods will be conducted to evaluate the effectiveness of the proposed approach.

3. Proposed Method

In this chapter, I present detailed information about the proposed method which focuses on disentangling action-related features to enhance the accuracy and robustness of the action classification system. The proposed method comprises two sequential training processes: representation learning and transfer learning.

3.1 Disentanglement of Action-Related Features

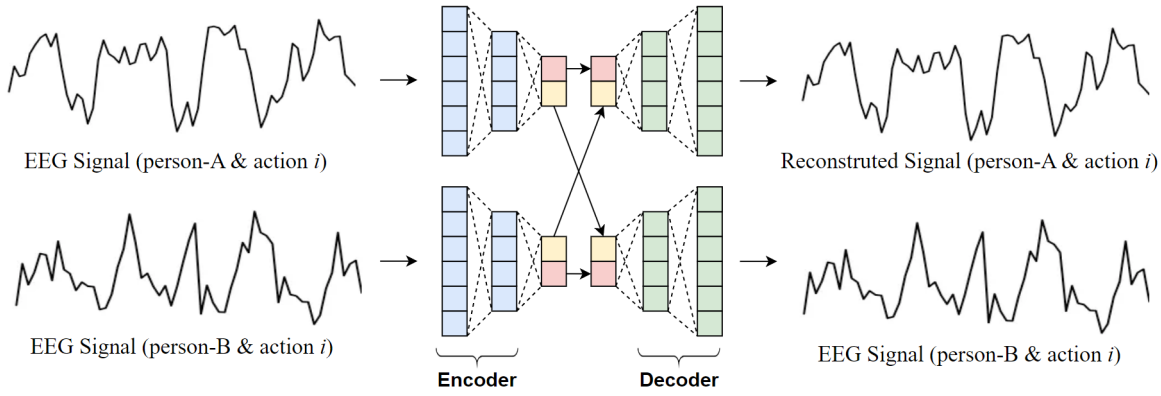

Figure 5. Overview of the proposed representation learning for disentanglement of action-related features

Figure 5 illustrates the proposed representation learning framework designed to disentangle action-related features. The architecture employs a convolutional neural network-based autoencoder which is composed of an encoder and a decoder. The network takes two EEG signals as input, both corresponding to the same action but emanating from different individuals. Despite exhibiting distinct patterns in the time series domain, these EEG signals may harbor latent features associated with the action category due to their shared nature.

The objective of the proposed representation learning is to isolate the action-related features from the feature maps generated by the encoder. To achieve this, I introduce a novel feature-switching approach that involves swapping action-related features from different sources. As illustrated in Figure 5, the decoder utilizes the feature maps with the exchanged action-related features as input to reconstruct the original EEG signal.

Despite the interchange of action-related features, the decoder is designed to reconstruct the original EEG signal consistently, given that these features originate from the same action in the EEG signal. For the training, I employ the L1 loss function which is a commonly utilized metric in training autoencoder (Baldi et al. 2012) architectures.

Equation 1: L1 loss function

$$L_1 = \left|I_i^A - \widehat{I_i^A}\right| + \left|I_i^B - \widehat{I_i^B}\right|$$

In Equation 1, I_i^A and I_i^B denote the EEG signal from person-A and person-B, respectively. Note that these two EEG signals belong to the same action category. The variable $\widehat{I_i^A}$ and $\widehat{I_i^B}$ represent the reconstructed EEG signals that should be identical correspondence with I_i^A and I_i^B. The L1 loss function quantifies the element-wise disparity between two signals. The optimal value is zero which indicates that the reconstructed signal is identical to the original signal. This result signifies that the encoder has successfully disentangled the action-related features.

3.2 Action Classification

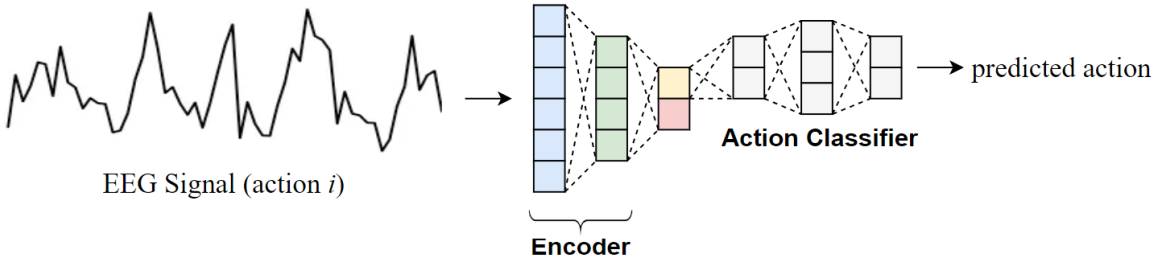

Figure 6. Overview of the proposed transfer learning (action classification)

Figure 6 shows the proposed transfer learning that leverages the pre-trained encoder obtained from the representation learning phase. Instead of training the action classifier from scratch, the proposed method employs transfer learning. Given the pre-trained encoder's ability to consistently disentangle action-related features from EEG signals, this transfer learning process enables rapid and accurate training. The effectiveness of this approach is detailed in the experimental results chapter through a series of comprehensive experiments.

For training the action classifier, the cross-entropy loss function, a standard choice for classification networks, is employed. Equation 2 introduces the cross-entropy function.

Equation 2: Cross-entropy loss function

$$L = -log_e y$$

In Equation 2, y denotes the predicted action probability. It quantifies the difference between the predicted probability distribution and the true probability distribution of the target classes. The cross-entropy loss increases as the predicted probabilities diverge from the actual distribution. In other words, it penalizes the model more when it is confident about incorrect predictions. The goal during training is to minimize this loss, and as the model improves, the cross-entropy loss should approach zero.

3.3 Parameter Explanation for Reproduce

In this chapter, I present the parameter settings required to reproduce the proposed approach. Regarding the architecture, I implement a 1-D convolutional neural network based on ResNet-50 (He et al. 2016) to create the encoder. For the decoder, an inverted ResNet-50 is developed, wherein the max-pooling layers are substituted with upsampling

layers. I conducted representation learning training for 140 epochs, employing a learning rate of 0.0001. Subsequently, for transfer learning, the model underwent training for 15 epochs with the same learning rate setting. A consistent batch size of 512 was utilized for both representation and transfer learning phases.

4. Experimental Results

4.1 EEG Dataset

In this research, an EEG dataset was utilized to investigate the generalization of EEG-based classification for a user-independent Brain-Computer Interface. The dataset encompasses signals recorded during six distinct categories of activities, each captured from an EEG device. The targeted activities are step up, stand up, sit down, move forward, turn left, and turn right. I collected 100 samples per individual for each action category, resulting in a total of 1200 samples for each category. The entire dataset consists of 7200 samples acquired from all participants across the six activity categories.

4.2 Evaluation Metrics and Experiment Design

I employed K-fold cross validation which is a widely adopted technique in machine learning for evaluating the performance and generalizability of predictive models. As illustrated in Figure 7, this cross validation method entails dividing a dataset into ***K*** equally sized folds or subsets. The model is then iteratively trained ***K*** times with each iteration utilizing different folds as the training set and the remaining data as the validation set.

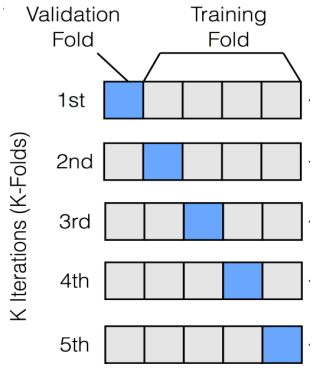

Figure 7. K-fold cross validation

For the evaluation metrics, I used accuracy, precision, recall, and f1-score which are the key metrics of evaluation object classification problems. Accuracy measures the overall correctness of the classification model and is calculated as the ratio of correctly predicted instances to the total instances. Recall quantifies the ability of a model to capture all relevant instances of a particular class. It is the ratio of correctly predicted positive instances to the total actual positive instances. Precision gauges the accuracy of positive predictions made by the model and is the ratio of correctly predicted positive instances to the total predicted positive instances. Finally, f1-score is the harmonic mean of precision and recall. It provides a balanced measure between precision and recall which offers a single metric that considers both false positives and false negatives.

	equation	definition
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$	the proportion of correct predictions among the total number of cases examined.
Precision	$\frac{TP}{TP+FP}$	the proportion of true positive predictions in the total predicted positives.
Recall	$\frac{TP}{TP+FN}$	the proportion of actual positives that were identified correctly.
F1-score	$2\times\frac{P\times R}{P+R}$	the harmonic mean of precision and recall

Figure 8. Evaluation metrics used in this research

4.3 Quantitative Evaluation

Table 1. Performance comparison with the state-of-the-art methods

	Accuracy	**Recall**	**Precision**	**F1-Score**
VGG19 based (Simonyan et al. 2014)	0.7489 (±0.0007)	0.7491 (±0.0006)	0.7486 (±0.0005)	0.7463 (±0.0007)
MobileNetV2 based (Sandler et al. 2018)	0.7520 (±0.0012)	0.751 (±0.0015)	0.7513 (±0.0012)	0.7467 (±0.0011)
DenseNet121 based (Huang et al. 2017)	0.7701 (±0.0011)	0.7692 (±0.0009)	0.7684 (±0.0016)	0.7658 (±0.0014)
Resnet50 based (He et al. 2016)	0.7707 (±0.0006)	0.7713 (±0.0012)	0.7708 (±0.0008)	0.7687 (±0.0012)
Ours	**0.8016 (±0.0009)**	**0.8016 (±0.0010)**	**0.8009 (±0.0008)**	**0.7983 (±0.0007)**

Table 1 presents a summary of the performance comparison between the proposed method and state-of-the-art convolutional neural network architectures. To process the 1-D EEG signal, I replaced all 2-D convolutional layers with their 1-D counterparts. As shown in Table 1, architectures based on VGG19 (Simonyan et al. 2014) and MobileNetV2 (Sandler et al. 2018) yielded poor results due to their relatively shallow network layers. In contrast, models based on DenseNet121 (Huang et al. 2017) and ResNet50 (He et al. 2016) demonstrated slightly higher accuracy benefitting from their deeper layer structures. Notably, the proposed method outperformed all the considered supervised learning methods and achieved the highest accuracy among the evaluated approaches.

I believe the proposed method outperforms over other approaches because during the proposed representation learning process, the trained model can extract meaningful features which makes the prediction results more accurate by disentangling the action-related features. A more in-depth explanation of the effectiveness of the proposed method is provided in Chapter 4.4.

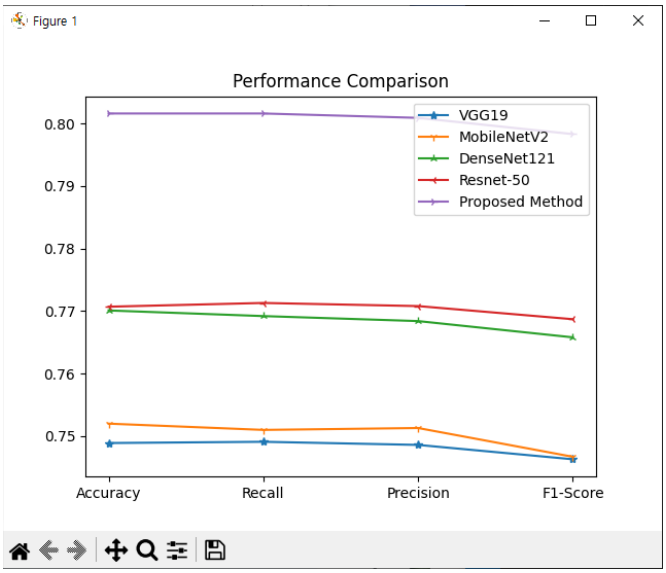

Figure 9. Performance comparison with the state-of-the-art methods

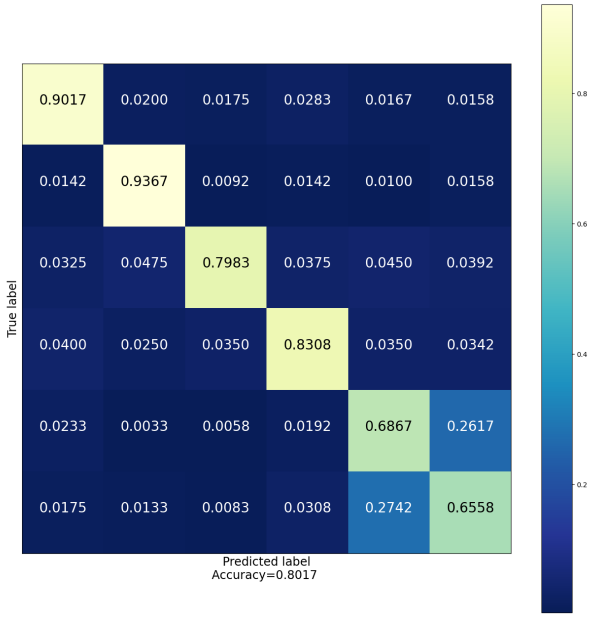

Figure 10. Confusion matrix of the proposed method

Figure 10 depicts the confusion matrix evaluation of the proposed method. Robust results are observed for the first four actions (step up, stand up, sit down, move forward). However, for the turn left and turn right actions, the trained model exhibits relatively lower accuracy due to the similarity in signal patterns within EEG signals for these turning actions. This outcome underscores a limitation in the proposed method's ability to effectively disentangle and distinguish between these two actions. Future studies will focus on addressing this challenge by further investigating methods to disentangle these similar EEG signal patterns.

4.4 Ablation Study for Different Convolutional Neural Network Architectures

Table 2. Ablation study result (Convolutional neural network architecture replacement)

	Accuracy (baseline)	**Accuracy (proposed method)**
VGG19 based	0.7489 (±0.0007)	0.7720 (±0.0009)
MobileNetV2 based	0.7520 (±0.0012)	0.7784 (±0.00010)
DenseNet121 based	0.7701 (±0.0011)	0.8015 (±0.0012)
Resnet50 based	0.7707 (±0.0006)	0.8016 (±0.0009)

In this chapter, I performed an ablation study to assess the effectiveness of the proposed approach. Initially, I trained four distinct convolutional neural network architectures without the proposed approach. Note that these models were trained in a supervised manner. After that, I trained the same four architectures with the proposed approach (representation learning). The performance disparity between these two sets of models is detailed and compared in Table 2.

As shown in Table 2, the proposed method enhanced accuracy across all four architectures. This outcome provides evidence that the proposed representation learning approach effectively contributed to disentangling action-related features. This eventually led to an improvement in the model's classification performance.

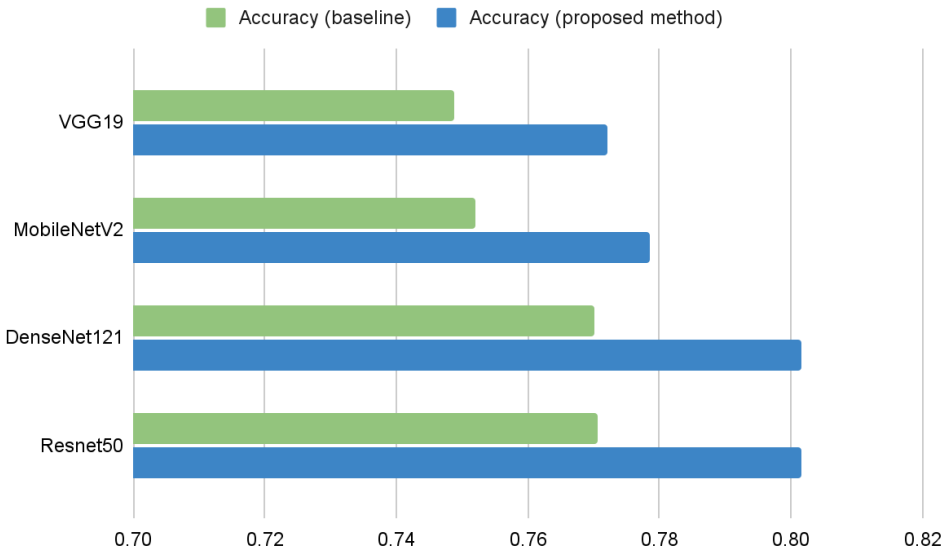

Figure 11. Ablation study result (Convolutional neural network architecture replacement). Figure visualizes the results of the ablation study using a bar graph to offer a clearer interpretation of the experimental outcomes.

5. Conclusion

In this research, I proposed a EEG-based classification for a user-independent Brain-Computer Interface. The goal of this research is to generalize EEG-based classification, emphasizing user independence, through the proposed representation learning. The target action categories are step up, stand up, sit down, move forward, turn left, and turn right

which are captured from 12 individuals. The proposed method demonstrated promising outcomes, surpassing state-of-the-art convolutional neural network architectures in terms of accuracy. Leveraging representation learning, the proposed approach showcased its efficacy in disentangling action-related features, particularly enhancing the model's performance for the first four actions—step up, stand up, sit down, and move forward. Additionally, an ablation study further validated the significance of the proposed representation learning approach. By training four different convolutional neural network architectures, both with and without the proposed method, it is observed that a consistent improvement in accuracy across all architectures when the representation learning was employed. Despite these advancements, challenges persist, notably in accurately classifying turn left and turn right actions due to their similar signal patterns. This limitation underscores the need for ongoing research to refine the model's ability to disentangle such intricate distinctions in EEG signals.

References

Baldi, P. (2012, June). Autoencoders, unsupervised learning, and deep architectures. In Proceedings of ICML workshop on unsupervised and transfer learning (pp. 37-49). JMLR Workshop and Conference Proceedings.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778). https://doi.org/10.48550/arXiv.1512.03385

Hinss, M. F., Jahanpour, E. S., Somon, B., Pluchon, L., Dehais, F., & Roy, R. N. (2023). Open multi-session and multi-task EEG cognitive Dataset for passive brain-computer Interface Applications. Scientific Data, 10(1), 85.

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708). https://doi.org/10.48550/arXiv.1608.06993

Milbich, T., Ghori, O., Diego, F., & Ommer, B. (2020). Unsupervised representation learning by discovering reliable image relations. Pattern Recognition, 102, 107107.

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4510-4520). https://doi.org/10.48550/arXiv.1801.04381

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. https://doi.org/10.48550/arXiv.1409.1556