A Novel Approach for Transfer Learning Motor Imagery Classification Based on IVA

Caroline P. A. Moraes[1], Denis G. Fantinato[2], Aline Neves[1]
[1]*Center for Engineering, Modeling and Applied Social Sciences*
Federal University of ABC (UFABC)
Santo André, Brazil
[2]*Department of Computer Engineering and Industrial Automation*
State University of Campinas (UNICAMP)
Campinas, Brazil
caroline.moraes@ufabc.edu.br

***Abstract*—Motor imagery (MI) classification based on electroencephalogram (EEG) signals performs an important role in neurological rehabilitation for therapeutic proposes. Independent Component Analysis (ICA) is a set of techniques with a solid framework and is widely used in the signal processing area. Inspired by ICA, Independent Vector Analysis (IVA) is an extension of the problem for multiple datasets and explores the correlation between different datasets through the use of Mutual Information (Mutinf). The statistical dependency between datasets through Mutinf could help in MI classification since it allows a generic and homogeneous treatment of the whole data and a possible knowledge transfer between patients. This paper proposes an innovative approach for the Transfer Learning MI task by exploring the minimization of mutual information through IVA applied to motor imagery. The results show a high correlation and small standard deviation cross-subjects.**

***Index Terms*—Independent Vector Analysis, Transfer Learning, Motor Imagery, Brain-Computer Interface, Electroencephalogram**

I. Introduction

Joint Blind Source Separation (JBSS) is a multi-model approach that has attracted the attention of the scientific community due to its wide range of applications [1]–[3]. Inspired by the Blind Source Separation (BSS) problem, the goal of JBSS problem is to provide source separation of unknown sources through a set of observed mixtures, i.e., to recover the independent latent signals of each subject exploiting the possible correlation between signals recorded from different subjects [4]. Independent Vector Analysis (IVA) is a technique, from a set of JBSS methods, that generalizes Independent Component Analysis (ICA) to multiple datasets. Such an approach has been applied in the biomedical context using signals from fMRI (functional Magnetic Resonance Imaging) [5], [6] and electroencephalogram (EEG) [7], [8], showing a substantial potential in the field.

In Brain-Computer Interface (BCI), the brain wave signals from the subject are recorded using sensors and electrodes, which allows a connection between the brain and the external world. Such a system could be categorized into three types: Invasive BCIs, Partially-invasive BCIs, and Non-invasive BCIs. EEG is a non-invasive BCI method that brings forward the ability to capture brain activity in real-time, at the level of milliseconds, and an affordable method for measuring brain waves on the surface of the scalp. Since the EEG presents such benefits, a wide application of this method has been disseminated in the biomedical area [9]–[11]. One of the EEG applications is the Motor Imagery (MI) classification task that aims to identify the imagined movements from the brain's electrical signals collected by BCI equipment without any real motor execution [12], [13]. Moreover, recent works also explore MI classification in the Transfer Learning (TL) background where a model designed for one subject is reused on a second subject.

In [14] the authors combined Euclidean alignment (EA) with deep learning techniques such as EEGNet [15] and ShallowNet [16] to classify the MI movements from BCI Competiton IV Dataset 1 (DS1) using TL. The results demonstrated the influence of EA, reaching an accuracy of 69.2% for EEGNet and 68.9% for ShallowNet. In the work of Siwei Liu et al. [17], the authors compare several TL methods, between them, the CNNnet [18], EEGNet, and ShallowNet, achieving an accuracy of 63%, 69% and 78%, respectively.

Despite the growing use of IVA in biomedical signals processing that naturally addresses cross-subject issues, based on our current understanding and expertise, the method was never applied as a TL technique. The appealing approach of exploring the independent components across subjects seems a suitable method for TL applications. In that sense, we propose the application of IVA as a Transfer Learning technique for MI classification. The method is applied to the DS1 dataset and uses three traditional classifier algorithms, Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), and Multilayer Perceptron (MLP) to evaluate the performance of the method.

In Section II, we describe the JBSS concept and IVA method. Section III shows a brief explanation of the autoregressive (AR) model used for dimensionality reduction and the

This study was financed by grants #2020/10014-2 and #2020/09838-0, São Paulo Research Foundation (FAPESP), grant #88887.595656/2020-00 Coordination for the Improvement of Higher Education Personnel (CAPES)

classifiers adopted. Section IV presents the proposed method employing IVA as a TL method. In Section V the simulation results are presented and evaluated. Finally, we conclude this paper in Section VI.

II. Independent Vector Analysis

In the following, we describe the general concept of the JBSS problem with K datasets, each containing N samples, formed from linear mixtures of M independent sources. The mixing process can be modeled by: $\mathbf{x}^{[k]}(n) = \mathbf{A}^{[k]}\mathbf{s}^{[k]}(n),\ 1 \leq n \leq N,\ 1 \leq k \leq K$, $\mathbf{s}^{[k]}(n) = [s_1^{[k]}(n), \ldots, s_M^{[k]}(n)]^T \in \mathbb{R}^M$ is the concatenated source vectors in each dataset and $\mathbf{A}^{[k]} \in \mathbb{R}^{M\times M}$ is the k-th invertible mixing matrix, where superscript T denotes transpose, and both are unknown.

For retrieving the sources, the prewhitening procedure is recommended [19], where the whitening matrix $\mathbf{V}^{[k]}$ was obtained by computing $\mathbf{V}^{[k]} = \mathbf{E}^{[k]}\mathbf{D}^{[k]^{-1/2}}\mathbf{E}^{[k]^T}$, $\mathbf{D}^{[k]}$ is a diagonal matrix with the eigenvalues, and $\mathbf{E}^{[k]}$ is a matrix with the eigenvectors of the correlation matrix from the mixing vector $\mathbf{x}^{[k]}$ for each dataset. With $\mathbf{V}^{[k]}$, the whitening process results in $\mathbf{z}^{[k]}(n) = \mathbf{V}^{[k]}\mathbf{x}^{[k]}(n)$. The demixing system is given by: $\mathbf{y}^{[k]}(n) = \mathbf{W}^{[k]}\mathbf{z}^{[k]}(n),\ 1 \leq n \leq N,\ 1 \leq k \leq K$. The goal is to obtain K matrices $\mathbf{W}^{[k]}$ and the corresponding source vectors estimated $\mathbf{y}^{[k]}(n)$ for each dataset. Fig. 1 shows an illustration of the concept.

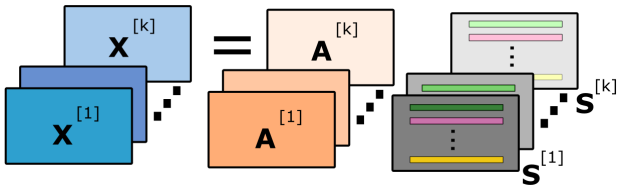

Fig. 1. Joint Blind Source Separation concept.

IVA was developed based on ICA concepts and could be considered as an extension of the latter for multiple datasets through the mutual information minimization among the source component vectors (SCV) [19]. The m-th SCV for k-th dataset can be defined as: $\mathbf{y}_m = [y_m^{[1]}\ y_m^{[2]}\ \ldots\ y_m^{[k]}] \in \mathbb{R}^K$. In IVA, SCVs are made maximally dependent within an SCV cross K datasets and maximally independent with respect to other SCVs. When we concatenate the mth SCV from different K dataset, it is possible to define the Source Component Matrix (SCM), through the concatenation of each row of $\mathbf{y}^{[k]}$, as $\mathbf{SCM}_m = [\mathbf{y}_m^{[1]}, \mathbf{y}_m^{[2]}, \ldots, \mathbf{y}_m^{[k]}]^T \in \mathbb{R}^M$. This process is exemplified by Fig. 2 based on the concept shown in Fig. 1.

The IVA cost function is given by:

$$\begin{aligned}\mathcal{I}_{IVA} &\triangleq \mathcal{I}[\mathbf{y}_1; \ldots; \mathbf{y}_M] \\ &= \sum_{m=1}^{M} \mathcal{H}[\mathbf{y}_m] - \sum_{k=1}^{K} \log|\det(\mathbf{W}^{[k]})| - C_1,\end{aligned} \tag{1}$$

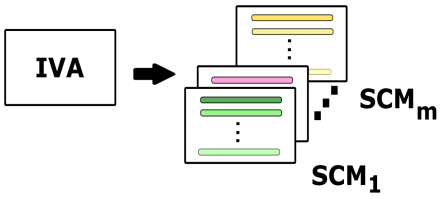

Fig. 2. Source Component Matrix.

where $\mathcal{I}[\mathbf{y}_1; \ldots; \mathbf{y}_M]$ is the mutual information within the SCVs, $\mathcal{H}(\cdot)$ is entropy and C_1 is a constant term that depends only on $\mathbf{z}^{[K]}$.

The minimization of the cost function (1) simultaneously minimizes the entropy of all components and maximizes the mutual information within each estimated SCV [19]. In this paper, we work with IVA-G [5] that assumes a multivariate Gaussian distribution for the SCVs, and thus only takes second-order statistical information into account.

III. Dimension Reduction and Classifiers

The IVA is able to obtain a new set of features. However, due to the dimension of EEG signals, it is necessary to apply some intermediate processes to enable proper data classification. In the sequel, we briefly describe the dimensional reduction and classifiers methods used to implement the application of IVA as a TL method.

A. Autoregressive model

AR model is frequently used to represent a random process in a different dimension. This is possible due to the model structure, where the output variable linearly depends on its own previous values [20]. In this paper, the benefits of using the AR model is twofold: it obtains important attributes from the time series given by the IVA outputs, and it also reduces data dimension due to the AR model size used. AR parameters were extracted from each IVA output signal.

B. LDA, SVM and MLP

Linear Discriminant Analysis (LDA) is a classical machine learning method that classifies the data in terms of statistical measures based on the mean value and variance of the training dataset [21]. On the other hand, Support Vector Machine (SVM) is an efficient supervised algorithm that determines a hyperplane in which the classes of the original problem are "separable" [22]. Inspired by the structure of the human brain, more specifically, in the biological neurons, the Multilayer perceptron (MLP) is characterized by several layers of input nodes connected as a directed graph between the input and output layers. In this work, when IVA is combined as a TL criterion with SVM, LDA, or MLP, the algorithm is called IVATL-S, IVATL-L, and IVATL-M, respectively. The classifiers inputs are given by the AR parameters obtained as explained in Section III.A, for each subject separately.

IV. IVA as a Transfer Learning approach

In the preprocessing stage, the EEG signals were submitted to a holdout evaluation technique using 80% of the data for training and 20% for testing, and following [19], the data also was pre-whitened for each subject independently.

Considering the training data, the IVA was applied for each class C (left hand (L) and right hand/foot (R/F)) to obtain the $\mathbf{W}_c^{[k]}$ matrices, where $k = \{1, 2, 3, 4\}$ and $c = \{1, 2\}$, that correspond to the weights values of the features extracted for the c-th class and k-th subject. Next, considering the k-th subject, the training and test data were multiplied by $\mathbf{W}_1^{[k]}$ and $\mathbf{W}_2^{[k]}$, followed by each corresponding whitening matrix $\mathbf{V}_1^{[k]}$ and $\mathbf{V}_2^{[k]}$, which result in the IVA components $\mathbf{y}_c^{[k]}$. Such an approach is necessary to maintain the blind assumption about the test data (class is unknown) [7].

As a contribution of this paper and extending the work in [7], the Transfer Learning approach will be grounded in an IVA ranking obtained from the SCMs correlation cross-subject matrices, using training data. IVA ranking was built based on the correlation between the independent components $\mathbf{y}_c^{[k]}$ cross subject for each class, in this case, $\mathbf{y}_1^{[k]}$ and $\mathbf{y}_2^{[k]}$. In that sense, the classifier model selection for the test data is based on the IVA ranking cross-subject obtained from SCMs using the training data. The cross-subject will be selected to transfer their classifier training weights model to the test one. Having chosen the optimal model for each subject, the IVA features previously stacked, are used as inputs for the MI classifiers. The whole procedure is exemplified by Fig. 3.

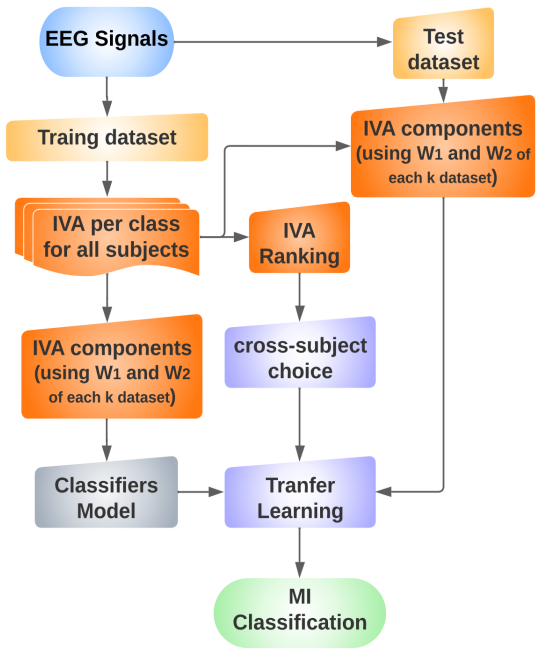

Fig. 3. Algorithm block diagram for Transfer Learning using IVA ranking.

V. Experiments and Results

In order to evaluate the performance of the algorithms, in this section, the method was evaluated with respect to the IVA adaptation step size (μ), the number of AR coefficients (p), and the number of hidden layers (h) and neurons (nr) from MLP. In addition, all the results were obtained based on an average of 20 simulations.

TABLE I
An Example of the IVA ranking and the criterion selection TL for class Left (L) and class Right/Foot (R/F).

	Cross-Sub.	IVA ranking L	IVA ranking R/F	Selection
"a"	"b"	0.4369	0.4988	"f"
	"f"	0.5149	0.5180	
	"g"	0.3974	0.3730	
"b"	"a"	0.4369	0.4988	"f"
	"f"	0.4847	0.6144	
	"g"	0.4415	0.3995	
"f"	"a"	0.4847	0.5180	"b"
	"b"	0.5149	0.6144	
	"g"	0.5109	0.4613	
"g"	"a"	0.3974	0.4613	"f"
	"b"	0.4415	0.3995	
	"f"	0.5109	0.4613	

A. BCI Competition IV Dataset1

DS1 was provided by B. Blankertz *et al.* [23] and was used as a competition dataset to recognize human subjects from the artificial data for MI movements. Since IVA explores the correlations between the components and we aim to analyze the real features, in this work, we used only the 4 human subjects. Each subject chose two classes of MI, the first class was left hand and the second class could be chosen among right hand, and foot or optionally, both feet. The procedure protocol was established by a visual arrow cue indicating the MI task for 4s, and a brief resting before the next cue. In total, for each subject, 100 trials per class were gathered. In this work, we use the calibration datasets from the 4 human subjects (labeled "a", "b", "f", and "g"), where each EEG dataset contains 59 channels and the signal was downsampled (100Hz). Signals were filtered in the 4–45 Hz bandwidth using a finite impulse response (FIR) filter and were then averaged and referenced.

B. IVA Ranking analysis

Based on the SCMs correlations matrices extracted from IVA, it is possible to build a correlation ranking between subjects. As an example, Tab. I shows the correlations extracted from IVA with respect to class Left and Right/Foot from one of the simulation scenarios. After building the IVA ranking, for each subject, we selected the cross-subject with the higher correlation score, e.g., the transfer option for subject "a" was subject "f" due to the highest correlation of 0.5180; For subject "f", the subject chosen was "b" with a correlation of 0.6144.

Fig. 4 presents the cross-subject selection frequency for each subject. For subjects "a", "b" and "g", the subject "f" was more often selected as a similar cross-subject, and for subject "f" the selected one was subject "b". These results show that the subject "f" is a universal transfer feature subject for DS1, which could lead to a future global TL model.

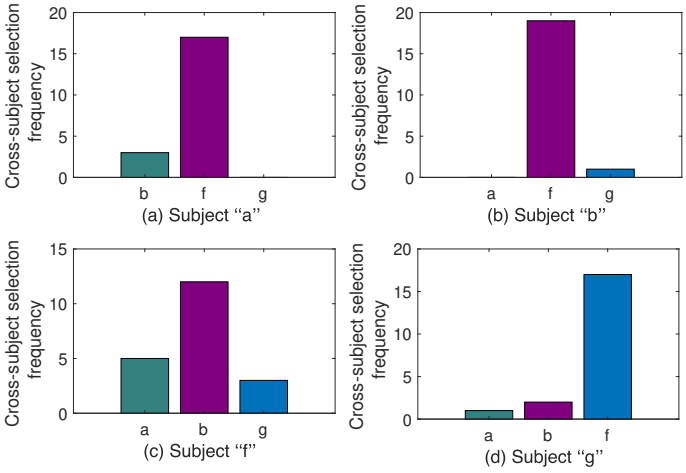

Fig. 4. Cross-subject selection frequency for each subject: (a) subject "a"; (b) subject "b";(c) subject "f";(d) subject "g".

C. Coefficients and Adaptation Step Size Analysis

Since the method is related to the AR model, it is possible to analyze the algorithm performance based on the number of AR coefficients. Initially, we fixed the IVA adaptation step size in $\mu = 1$, while the number of coefficients p sweeps from 1 to 10. Fig. 5(a) shows the results obtained. For IVATL-L the algorithm achieves the maximum accuracy of $64,9\%$, while the IVATL-S presents a smooth increase result of 68.5%. In both cases, as the number of coefficients increases, the accuracy increases as well, until the limit of $p = 10$. The best parameters choices were $p = 8$ for IVA-L and $p = 9$ for IVA-S.

In the sequel, the IVA adaptation step size was investigated. Based on the previous coefficients analyses and choosing the best parameters for each classifier (IVATL-S $p = 9$ and IVATL-L $p = 8$), μ was varied from 0.01 to 1.5. In all cases, as shown in Fig. 5(b), as the adaptation step size increases the classification accuracy also increases until the maximum value with $\mu = 1$, thereafter the performance decreases. Thus, in both cases, the best results were achieved with $\mu = 1$.

D. MLP Parameters Analysis

Having selected the optimal parameters, $p = 9$ and $\mu = 1$, we now investigate the IVA TL approach as input features to the MLP classifier. MLP architecture was implemented using the Scikit-learn (Sklearn) library for machine learning, with a few modifications. It sweeps the number of hidden layers and neurons to obtain the highest performance possible using the activation function "ReLU" and the stochastic gradient-based optimizer "Adam". To evaluate the performance of the

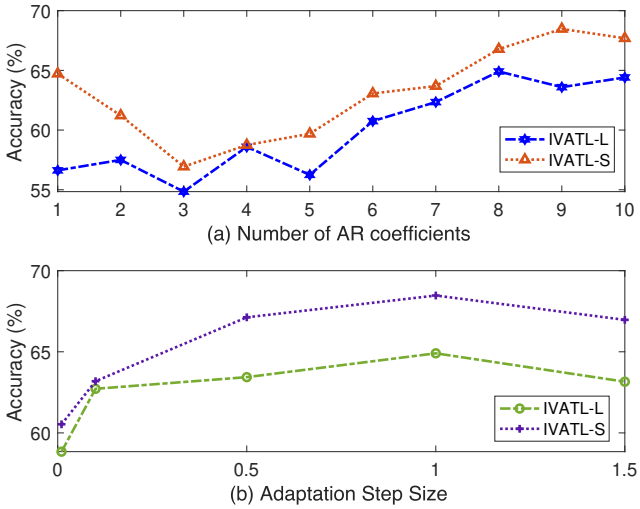

Fig. 5. (a) IVATL-S and IVATL-L performance based on the number of AR coefficients; (b) IVATL-S and IVATL-L performance based on the adaptation step size.

algorithm, the number of hidden layers was fixed in $h = 200$, while the number of neurons nr swept from 50 to 1000. Fig. 6(a) presents the performance of the IVATL-M varying the number of neurons. In this case, when the number of neurons was above 500, the performance decreased rapidly, obtaining a maximum mean accuracy of 70.2% with 300 neurons. Sweeping the number of hidden layers from 25 to 400, at first, the algorithm is able to improve its performance, but after 1000 layers, IVATL-M suffers a smooth decrease, maintaining the accuracy around 70.1% as shown in Fig. 6(b). The best result of $70,5\%$ was achieved with $h = 100$.

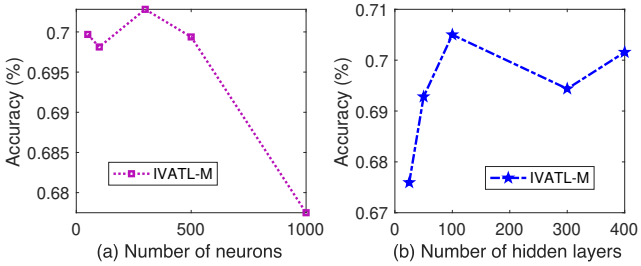

Fig. 6. (a) IVATL-M performance based on the number of hidden layers; (b) IVATL-M performance based on the number of neurons.

E. Discussion

Tab. II shows the results obtained using the IVATL-L classifier with the optimal parameters of $p = 8$ and $\mu = 1$, IVATL-S classifier with the optimal parameters of $p = 9$ and $\mu = 1$, IVATL-M classifier with the optimal parameters of $p = 9$, $\mu = 1$, $nr = 300$ and $h = 100$. The obtained performances were also compared with existing literature results, mainly TL approaches based on deep learning using CNNnet, EEGnet, and ShallowNet [14], [17].

Considering the three tested classifiers, IVATL-M achieved the best overall performance, with an accuracy of 70.5%.

TABLE II
A COMPARISON BETWEEN THE PROPOSED METHOD AND THE EEGNET, AND SHALLOWNET AVERAGE AND STANDARD DEVIATION (SD) CLASSIFICATION ACCURACY (%).

Method	"a"	"b"	"f"	"g"	Average ± SD
CNNnet	71	60	63	58	63 ± 4.9
EEGnet	82	62	73	59	69 ± 9.1
ShallowNet	84	67	89	72	78 ± 8.8
IVATL-L	67.8	65.6	72	61.6	66.8 ± 4.3
IVATL-S	70	66.8	71.8	65.1	68.5 ± 3.1
IVATL-M	71.6	**69.4**	74	67	70.5 ± **3**

Even though it is still lower than the performance obtained by ShallowNet, the obtained standard deviation is much better: 3 against 8.8, suggesting that the obtained results across subjects are statistically similar. It is also interesting to notice that IVA-based methods were able to increase the accuracy on the classification of subjects known to be difficult such as subjects "b" and "g", achieving the highest result in the first case and only losing to ShallowNet in the second one. This suggests that IVA TL manages to transfer relevant features from the easiest subjects to classify to those that are harder. On the other hand, we can also notice that the method degraded the accuracy obtained for subjects that are known to be easy to classify such as subject "a". A probable explanation is that the hardest subjects interfere with the feature extraction obtained by IVA since they are all processed together. Nevertheless, the results show that IVA has the potential to be used in TL frameworks.

VI. CONCLUSION

In this work, we presented a pioneer Transfer Learning approach for EEG motor imagery classification based on IVA. Since IVA explores the minimization of the mutual information to achieve independent vector analysis cross-datasets, it is almost intuitive that IVA has potential as a TL technique. The experimental results showed that IVA-M presented the lowest standard deviation, showing a homogeneous distribution classification cross-subjects. Moreover, we showed how the choice of the parameters of the algorithm such as step sizes, number of AR coefficients, neurons, and hidden layers could affect the accuracy. A fine adjustment of such parameters is an essential step. For future perspectives, we consider extending the work using more complex datasets and classifiers.

REFERENCES

[1] M. Ge, G. Cui, and L. Kong, "Mainlobe jamming suppression for distributed radar via joint blind source separation," *IET Radar, Sonar & Navigation*, vol. 13, no. 7, pp. 1189–1199, 2019.

[2] X. Chen, Z. J. Wang, and M. McKeown, "Joint blind source separation for neurophysiological data analysis: Multiset and multimodal methods," *IEEE Signal Processing Magazine*, vol. 33, no. 3, pp. 86–107, 2016.

[3] D. Sugumar, P. Vanathi, and S. Mohan, "Joint blind source separation algorithms in the separation of non-invasive maternal and fetal ecg," in *2014 International Conference on Electronics and Communication Systems (ICECS)*. IEEE, 2014, pp. 1–6.

[4] A. Matthew, A. Tülay, and X.-L. Lin, "Joint blind source separation with multivariate gaussian model: Algorithms and performance analysis," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1672–1683, 2012.

[5] I. Lehmann, E. Acar, T. Hasija, M. A. B. S. Akhonda, V. D. Calhoun, P. J. Schreier, and T. Adali, "Multi-task fmri data fusion using iva and parafac2," in *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2022, pp. 1466–1470.

[6] S. Ma, V. D. Calhoun, R. Phlypo, and T. Adalı, "Dynamic changes of spatial functional network connectivity in healthy individuals and schizophrenia patients using independent vector analysis," *NeuroImage*, vol. 90, pp. 196–206, 2014.

[7] C. P. A. Moraes, B. Aristimunha, L. H. d. Santos, W. H. L. Pinaya, R. Y. d. Camargo, D. G. Fantinato, and A. Neves, "Applying independent vector analysis on eeg-based motor imagery classification," in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023, pp. –.

[8] X. Chen, H. Peng, F. Yu, and K. Wang, "Independent vector analysis applied to remove muscle artifacts in eeg data," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 7, pp. 1770–1779, 2017.

[9] M. Soufineyestani, D. Dowling, and A. Khan, "Electroencephalography (eeg) technology applications and available devices," *Applied Sciences*, vol. 10, no. 21, p. 7453, 2020.

[10] A. Khosla, P. Khandnor, and T. Chand, "A comparative analysis of signal processing and classification methods for different applications based on eeg signals," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 2, pp. 649–690, 2020.

[11] S. Asadzadeh, T. Y. Rezaii, S. Beheshti, A. Delpak, and S. Meshgini, "A systematic review of eeg source localization techniques and their applications on diagnosis of brain abnormalities," *Journal of Neuroscience Methods*, vol. 339, p. 108740, 2020.

[12] N. Padfield, J. Zabalza, H. Zhao, V. Masero, and J. Ren, "Eeg-based brain-computer interfaces using motor-imagery: Techniques and challenges," *Sensors*, vol. 19, no. 6, p. 1423, 2019.

[13] G. Pfurtscheller, C. Neuper, D. Flotzinger, and M. Pregenzer, "Eeg-based discrimination between imagination of right and left hand movement," *Electroencephalography and clinical Neurophysiology*, vol. 103, no. 6, pp. 642–651, 1997.

[14] D. Wu, X. Jiang, and R. Peng, "Transfer learning for motor imagery based brain-computer interfaces: A tutorial," *Neural Networks*, 2022.

[15] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces," *Journal of neural engineering*, vol. 15, no. 5, p. 056013, 2018.

[16] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for eeg decoding and visualization," *Human brain mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.

[17] S. Liu, J. Zhang, A. Wang, H. Wu, Q. Zhao, and J. Long, "Subject adaptation convolutional neural network for eeg-based motor imagery classification," *Journal of Neural Engineering*, vol. 19, no. 6, p. 066003, 2022.

[18] H. Cecotti and A. Graser, "Convolutional neural networks for p300 detection with application to brain-computer interfaces," *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 3, pp. 433–445, 2010.

[19] M. Anderson, T. Adali, and X.-L. Li, "Joint blind source separation with multivariate gaussian model: Algorithms and performance analysis," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1672–1683, 2011.

[20] R. Brüggemann, *Model reduction methods for vector autoregressive processes*. Springer Science & Business Media, 2012, vol. 536.

[21] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of eugenics*, vol. 7, no. 2, pp. 179–188, 1936.

[22] V. Vapnik, *Statistical learning theory*. Wiley New York, 1998.

[23] B. Blankertz, G. Dornhege, M. Krauledat, K.-R. Müller, and G. Curio, "The non-invasive berlin brain–computer interface: fast acquisition of effective performance in untrained subjects," *NeuroImage*, vol. 37, no. 2, pp. 539–550, 2007.