Journal of Electrical and Computer Engineering Innovations
(JECEI)
Journal homepage: http://www.jecei.sru.ac.ir

Research paper

Utilizing Normalized Mutual Information as a Similarity Measure for EEG and fMRI Fusion

Z. Rabiei, H. Montazery Kordy *

Faculty of Electrical and Computer Engineering, Babol Noshirvani University of Technology, Babol, Iran.

Article Info

Article History:
Received 23 June 2024
Reviewed 11 August 2024
Revised 28 September 2024
Accepted 15 October 2024

Keywords:
Data fusion
Coupled matrix tensor factorization
Electroencephalogram (EEG)
functional Magnetic Resonance Imaging (fMRI)
Normalized mutual information (NMI)

*Corresponding Author's Email Address:
hmontazery@nit.ac.ir

Abstract

Background and Objectives: Neuroscience research can benefit greatly from the fusion of simultaneous recordings of electroencephalogram (EEG) and functional magnetic resonance imaging (fMRI) data due to their complementary properties. We can extract shared information by coupling two modalities in a symmetric data fusion.

Methods: This paper proposed an approach based on the advanced coupled matrix tensor factorization (ACMTF) method for analyzing simultaneous EEG-fMRI data. To alleviate the strict equality assumption of shared factors in the common dimension of the ACMTF, the proposed method used a similarity criterion based on normalized mutual information (NMI). This similarity criterion effectively revealed the underlying relationships between the modalities, resulting in more accurate factorization results.

Results: The suggested method was utilized on simulated data with correlation levels of 50% and 90% between the components of the two modalities. Despite different noise levels, the average match score improved by 20% compared to the ACMTF model, as demonstrated by the results.

Conclusion: By relaxing the strict equality assumption, we can identify shared components in a common mode and extract shared components with higher performance than the traditional methods. The suggested method offers a more robust and effective way to analyze multimodal data sets. The findings highlight the potential of the ACMTF method with NMI-based similarity criterion for uncovering hidden patterns in EEG and fMRI data.

Introduction

Joint analysis of neuroimaging data such as EEG and fMRI has the potential to gain a better understanding of brain functioning. The primary objective of analyzing multiple modalities is to utilize common and distinct information from complementary modalities to understand neural activities better. EEG and fMRI data fusion can provide researchers with a more comprehensive understanding of the brain's spatial and temporal functions [1], [2].

The synchronous electrical activity of brain neurons over time can be measured using EEG. While EEG has the perfect temporal resolution, this technique has poor spatial information due to the number of electrodes employed. On the other hand, blood oxygenation level-dependent (BOLD) imaging is a technique that is commonly used to measure brain activity indirectly using fMRI. Although it measures BOLD signals at a millimeter range, this technique is sluggish compared to brain activity [3], [4]. Therefore, EEG and fMRI can be fused to improve the localization of brain activity in time and space due to their complementary spatiotemporal resolutions. In recent years different types of fusion methods have been developed; the majority of them focus on the matrix factorization of EEG and fMRI into different components.

Independent component analysis (ICA) [5], principal component analysis (PCA), canonical component analysis (CCA) [6], [7], and independent vector analysis (IVA) [8] are methods for decomposing matrixes. Extensive studies have been conducted on decomposing EEG and fMRI datasets into different components using ICA and PCA methods. These methods define components using only two dimensions: time and space [5]. Whereas, EEG and fMRI datasets are typically multidimensional including time, voxels or channels, frequency, trial, and participant. One potential solution to this issue is to explore alternative methods for analyzing these datasets that preserve the interactions between different modalities. One approach uses tensor decomposition techniques designed to apply to multi-way data structures [9].

Non-physiological assumptions like orthogonality and statistical independence are the basis of matrix factorization models. However, the uniqueness of higher-order tensor decompositions is obtained by relaxed conditions (without any nonphysiological assumptions), making interpreting the extracted components easier [10]. By applying tensor decomposition algorithms such as Tucker decomposition or Parallel factorial analysis CANDECOMP/PARAFAC (CP) decomposition, researchers can extract more meaningful and interpretable components from EEG and fMRI datasets without losing important interactions between different dimensions [11], [12]. Coupled matrix tensor factorization (CMTF) is the most common method for fusing EEG and fMRI datasets using tensor decomposition methods [13]. In the CMTF model, data definition involves a third-order tensor of EEG coupled with a matrix fMRI. Gradient-based optimization algorithms are used to factorize the coupled data, and CP is utilized to model higher-order tensors. In the fusion of EEG and fMRI, the assumption is that there is one or more common modes of variation between the two modes, such as time or subject. The main disadvantage of the CMTF method is its reliance on equal shared components in the common dimension. Several methods have been introduced to alleviate this restrictive assumption. The advanced coupled matrix-tensor factorization (ACMTF) was developed in [14], identifying shared and unshared components. While the ACMTF can estimate the weights of the components and identify the factor matrices, it assumes that the shared components between the two modalities are identical. The notion of equality concerning brain signals could be confining. In [15] the CMTF model has been used to analyze the joint decomposition of EEG data at source level with fMRI along with a common spatial profile. This method can identify both common and discriminative subspaces compared to the CMTF method. A relaxed form of ACMTF was presented by the authors in [16]. This method overcomes the equality assumption of shared factors in a common dimension. This method uses the l_1-norm and l_2-norm to express similarity and then apply it to the components and their first and second derivatives. In [17] a tensor decomposition model was proposed in which a soft coupling method (Euclidean distance) was implemented for fusion EEG and fMRI. [18] has used the maximum correlation between the shared components of EEG and fMRI. Although Pearson's correlation coefficient (ρ) used in [18] is one of the most popular dependence measures with many desirable features, it only evaluates linear relationships. To assess relationships and dependencies between variables in a general sense, we need a metric, not only for linear or monotonic relationships. In contrast to the Pearson's correlation coefficient, mutual information (MI) takes into account both linear and non-linear relationships between variables, making it a more comprehensive measure of dependence. Additionally, MI can capture complex dependencies that may not be captured by the correlation coefficient alone. This makes MI a valuable tool for analyzing relationships in a wide range of fields. Overall, while the correlation coefficient is useful for measuring linear relationships, MI provides a more nuanced and flexible approach to understanding the dependencies between variables [19].

However, the estimation of MI and entropy values can be challenging. MI-based measures need appropriate estimation methods as the underlying probability distributions are unknown. The most commonly used technique for estimating MI is histogram-based density estimation [17]. Despite not always being the most accurate method, histogram-based density estimation has acceptable accuracy. Normalization of mutual information is essential because MI values can vary widely depending on the scale of the variables involved. By transforming MI into a standardized range, we can compare and interpret the information content more accurately. As a result, we used the normalized mutual information (NMI) as a similarity metric in our study [20], [21]. The contributions of the proposed method are summarized as follows:

- Using normalized mutual information as a similarity measure, our proposed method can relax the restrictive equality assumption of shared components in the ACMTF method.
- As a comprehensive approach, our method can estimate the weight of each component and identify identical and similar components with various correlation levels.
- Our proposed method, compared to other methods based on similarity criteria, can identify components that are linearly or nonlinearly related to each other.

The following is the structure of this paper. We first explain tensor decomposition, the ACMTF method, and

HRF modeling. The proposed method and the calculation of normalized mutual information are presented in the second part. Then, a simulation study is used to validate the performance of the presented method. Finally, the paper is completed with a discussion and conclusion.

Material and Methods

A. Notation

Vectors, matrices, and higher-order tensors in this study are identified using italic lower-case, italic upper-case, and italic calligraphic upper-case letters respectively. For a matrix *A*, $\bar{A}$ denotes its transpose. The symbol $\odot$ signifies the Khatri-Rao product of two matrices, A$\in \mathbb{R}^{I\times R}$ and B$\in \mathbb{R}^{J\times R}$, namely, $A\odot B = [a_1 \otimes b_1, a_2 \otimes b_2, \dots a_R \otimes b_R]$, with a_i and b_i being the ith columns of A and B respectively, and $\otimes$ denoting Kronecker prodeuct.

B. Tensor Decomposition

In mathematics, a tensor is described as a numerical array with multiple indexes, and the order of a tensor is the number of its modes or dimensions. The Canonical Polyiadic Decomposition (CP or CPD) model is briefly discussed in this section. A third-order tensor $\boldsymbol{\chi} \in \mathbb{R}^{I\times J\times K}$ with the modes of trial, frequency, and channel represents EEG data and a matrix $Y \in \mathbb{R}^{I\times L}$ (trial (scan) by voxels) indicates fMRI signal. Fig. 1 shows the EEG coupled with fMRI in the trial mode.

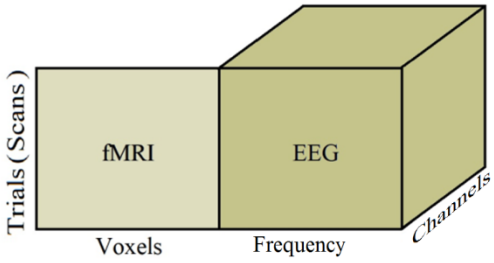


Fig. 1: A 3rd-order tensor EEG signal coupled with fMRI matrix in the trial mode.

CP is thought of as an extension of singular value decomposition (SVD) to higher-order tensors. It represents a 3rd-order tensor $\boldsymbol{\chi} \in \mathbb{R}^{I\times J\times K}$ as a linear combination of rank-one tensors:

$$\boldsymbol{\chi} = [\![\lambda; A, B, C]\!] = \sum_{r=1}^{R} \lambda_r a_r \circ b_r \circ c_r \qquad (1)$$

where $[\![.]\!]$ and $\circ$ indicate the full multilinear and vector outer product respectively. The vectors a_r, b_r and c_r are as rank-one components form the factor matrices A$\in \mathbb{R}^{R\times 1} = [a_1 \dots a_R]$, B$\in \mathbb{R}^{R\times 1} = [b_1 \dots b_R]$, and C$\in \mathbb{R}^{R\times 1} = [c_1 \dots c_R]$ respectively. The terms factor and component mention to the rank-one matrices or higher-order rank-one tensors. R signifies the number of factors and $\lambda \in \mathbb{R}^{R\times 1}$ is weights of rank-one components. The CP or PARAFAC model is one of the most popular tensor decomposition models. This model is used alongside models like Block Term Decomposition and the Tucker decomposition model [22].

C. Advanced Coupled Matrix and Tensor Factorization

We assume that the EEG data is structured as a third-order tensor to present variations across the trial, spectral, and spatial dimensions. At the same time, the fMRI matrix characterizes variations across the trial and spatial dimensions. Using the Advanced Coupled Matrix Tensor Factorization (ACMTF) model we can jointly factorize the 3rd-order tensor $\boldsymbol{\chi}$ coupled with a matrix Y in trial mode. The common mode between the EEG and fMRI is trial-to-trial (scan-to-scan) covariations in brain activity [23]. According to the definition of EEG and fMRI, a temporal relationship between EEG and fMRI data is considered in the form of hemodynamic response function (HRF) [4]-[24]. Thus, the ACMTF model [14] can be utilized to formulate an optimization problem:

$$\begin{aligned} & f(\lambda, \sigma, T_{ee}, F_{ee}, M_{ee}, M_{fm}) = \|\boldsymbol{\chi} - \\ & [\![\lambda; T_{ee}, F_{ee}, M_{ee}]\!]\| + \left\|Y - HT_{ee}\Sigma {M_{fm}}^T\right\|^2 + \\ & \beta\|\lambda\|_1 + \beta\|\sigma\|_1 \qquad (2) \\ & s.t. \|t_{eer}\| = \|f_{eer}\| = \|m_{eer}\| = \|m_{fmr}\| = \\ & 1 \quad for\ r = 1, \dots, R \end{aligned}$$

where the tensor $\boldsymbol{\chi}$ and matrix Y are decomposed based on the CANDECOMP/PARAFAC (CP) and singular value decomposition (SVD) models, respectively. The factor matrix $T_{ee} \in \mathbb{R}^{I\times R}$ (trial-to-trial variation) is common between EEG and fMRI. Moreover, hemodynamic trials of fMRI could be predicted using the convolution of T_{ee} in EEG with known HRF $h(t)$. The Toeplitz matrix H contains samples of $h(t)$ on its diagonals [4].

Also, $F_{ee} \in \mathbb{R}^{J\times R}$ and $M_{ee} \in \mathbb{R}^{K\times R}$ are factor matrices corresponding to the frequency and channel topography of the EEG signal, respectively; $M_{fm} \in \mathbb{R}^{L\times R}$ is the factor matrix corresponding to the spatial maps of voxels; and $\lambda \in \mathbb{R}^{R\times 1}$ and $\sigma \in \mathbb{R}^{R\times 1}$ are weights of rank-one components in the third-order tensor and the matrix, respectively. The $\Sigma \in \mathbb{R}^{R\times R}$ is a diagonal matrix, with σ forming its diagonal. Also, $\|.\|$ and $\|.\|_1$ represent the Frobenius norm and l_1-norm, respectively; $\beta \geq 0$ is a penalty parameter; and $t_{eer}, f_{eer}, m_{eer}$, and m_{fmr} are the r^{th} columns of T_{ee}, F_{ee} , M_{ee}, and M_{fm}, respectively. The weights λ and σ are sparsified using the l_1-norm terms. Thus, the components with significant weights in both modalities are considered shared, while unshared components have weights of almost zero in one of the datasets.

D. HRF Modeling

One method of analyzing the task-related fMRI data is to estimate the shape of the time courses corresponding to the considered stimulus. Among the various methods used to estimate these time courses, the linear

convolutional model is the most common technique. In this approach, the task-related time course can be modeled as a convolution between the stimulus function and a particular impulse response function known as the hemodynamic response function (HRF) [25]. One of the most popular techniques used to analyze fMRI signals is the general linear model (GLM). It models the BOLD signal as a linear combination of several various predictors. The GLM method requires an accurate estimate of the HRF. Several different models of the HRF are used in the analysis of fMRI signals [25], [26].

The most widely used model for the functional shape of the HRF is the double gamma distribution model, usually referred to as canonical which is used in SPM software. There are several models of canonical HRF in literature. HRF waveform shapes can be controlled by one to nine free parameters, based on the model [27]. The canonical HRF demonstrated in Fig. 2 has been used in this study. Some HRF studies use a basis function approach under the GLM framework; for example two sine basis functions or a product of a sine function and exponential function. Woolrich et al. presented constrained linear basis sets for HRF modeling using Variational Bayes. The proposed HRF was parameterized using six parameters. Several techniques exist to choose a basis set with the constraints so that the subspace spanned by the basis function creates a plausible HRF waveform [28].

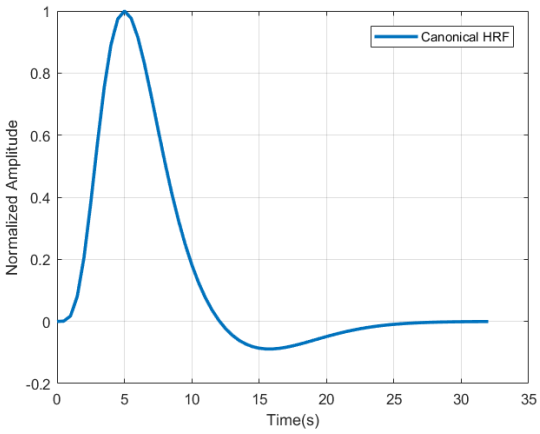


Fig. 2: Representation of the canonical HRF in SPM.

Bayesian methods are another approach for incorporating prior knowledge or uncertainty into the modeling of HRF [29]. The deconvolution method and a semiparametric approach based on finite impulse response (FIR) are other methods with more flexibility in modeling HRF [30].

Although HRF waveform varies across different subjects and different brain regions, in most studies, the HRF is considered invariant and assumed to be known. H is a convolution matrix defined as follows, where *l* is the length of the signal of T_{ee}, m is the length of HRF and the dimension of H is $(m + l - 1) \times l$.

$$H = \begin{bmatrix} h_1 & 0 & \cdots & 0 & 0 \\ h_2 & h_1 & \cdots & \vdots & \vdots \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ h_m & h_{m-1} & \cdots & 0 & h_2 \\ 0 & h_m & \ddots & h_{m-2} & \vdots \\ 0 & 0 & \cdots & h_{m-1} & h_{m-2} \\ \vdots & \vdots & \cdots & h_m & h_{m-1} \\ 0 & 0 & 0 & \cdots & h_m \end{bmatrix} \quad (3)$$

E. Normalized Mutual Information (NMI)

Measurement of the relationship between two variables can be done using information theory (IT). A generalized criterion of dependency is mutual information (MI). MI quantifies the amount of information shared between two variables. In addition, It excels at capturing complex interdependencies that may not be adequately represented by traditional measures like Pearson's correlation coefficient. Overall, while Pearson's correlation coefficient is useful for measuring linear relationships, mutual information provides a more comprehensive understanding of the dependencies between variables. As a result, MI can be used to determine how similar or dissimilar the shared components in two datasets are.

If X and Y are two random variables, $p(x, y)$ is the joint distribution, and $p(x)$ and $p(y)$ are the marginal distributions, the mutual information is calculated as follows:

$$I(X,Y) = \sum_{x,y} p(x,y) log(\frac{p(x,y)}{p(x)p(y)}) = H(X) + H(Y) - H(X;Y) \quad (4)$$

where $I(X,Y)$ is the mutual information, $H(X)$ and $H(Y)$ are Shannon's entropies of the discrete random variables X and Y, respectively, and $H(X;Y)$ is the joint entropy [21].

Estimating the entropy and MI values requires careful consideration of the data distribution and sample size. MI-based measures require to be estimated due to the unknown underlying probability distributions. Additionally, choosing the appropriate method for estimating entropy and MI can impact the reliability and interpretability of the results. Compared to various MI estimation methods such as kernel density estimation or Gaussian mixture models, histogram-based density estimation is a simple and effective technique [20]. However, care must be taken to choose the appropriate number of bins to avoid underestimating or overestimating the true MI values. Before adding MI to the objective function, it needs to be normalized. The normalization ensures that all variables are on a consistent scale and make information more meaningful. Specifically, it is transformed into a number within the range of [0, 1]. The normalized mutual information (NMI) can be defined as

$$NMI(X,Y)=\frac{I(X;Y)}{\frac{1}{2}(H(X)+H(Y))}=\frac{2I(X;Y)}{(H(X)+H(Y))} \quad (5)$$

F. Generalized Coupled Matrix Tensor Factorization

The ACMTF has been enhanced by incorporating the NMI criteria to address the equality constraint of the shared components. Now, according to generalized coupled matrix tensor factorization (GCMTF), the cost function is as follows:

$$\begin{aligned} &f(\lambda,\sigma,T_{ee},F_{ee},M_{ee},T_{fm},M_{fm})=\|\boldsymbol{\chi}-\\ &[\![\lambda;T_{ee},F_{ee},M_{ee}]\!]\|+\left\|Y-HT_{fm}\Sigma {M_{fm}}^{T}\right\|^{2}+\\ &\gamma\sum_{r=1}^{R}\left(1-e^{-(\lambda_r\sigma_r)^2/\varepsilon}\right)\left(1-NMI(t_{eer},t_{fmr})\right)+\\ &\beta\sum_{r=1}^{R}\sqrt{{\lambda_r}^2+\epsilon}+\beta\sum_{r=1}^{R}\sqrt{{\sigma_r}^2+\epsilon} \end{aligned} \quad (6)$$

$$s.t.\ \|t_{eer}\|=\left\|t_{fmr}\right\|=\|f_{eer}\|=\|m_{eer}\|=\left\|m_{fmr}\right\|=1 \quad for\ r=1,\dots,R$$

where γ is the penalty parameter; $t_{eer}, f_{eer}, m_{eer}$, t_{fmr}, and m_{fmr} are the r^{th} columns of T_{ee}, F_{ee}, M_{ee}, T_{fm}, and M_{fm}, respectively; and NMI is the normalized mutual information between t_{eer} and t_{fmr}. By selecting a sufficiently small enough $\epsilon>0$, the l_1-norm of λ and σ has been replaced with their differentiable equivalents. The expression $(1-e^{-(\lambda_r\sigma_r)^2/\varepsilon})$ is the smoothed l_0-norm, where ε is a tunable and small parameter to approximate l_0-norm [19]-[31]. This term is used to identify the shared components and avoid the maximization of the NMI between the unshared components.

Alternating Least Square (ALS) is the traditional approach for optimizing the objective function. In [11], non-conjugate gradient methods achieve faster convergence than ALS. In this approach, it is necessary to compute the gradients of the objective function with respect to their parameters. Hence, NMI gradients with respect to factor matrices in the objective function need to be computed. The Score Functions (SFs) defined in [20]-[32] were used to compute the NMI gradient.

If we have a bounded random vector X and a small enough Δ vector of the same dimension, it is demonstrated:

$$I(X+\Delta)-I(X)=E\{\Delta^T\beta_X(X)\}+o(\Delta) \quad (7)$$

where $\beta_X(X)=\psi_x(X)-\varphi_x(X)$ is the Score Function Difference (SFD) of X and $o(\Delta)$ represents the higher-order expressions in Δ. The terms $\psi_x(X)$ and $\varphi_x(X)$ are the Marginal Score Functions (MSFs) and Joint Score Functions (JSFs) of vector X, which are defined as follows:

$$\psi_x(X)=-\frac{d}{dx_i}lnp_{x_i}(x_i)=-\frac{\acute{p}_{x_i}(x_i)}{p_{x_i}(x_i)} \quad (8)$$

$$\varphi_x(X)=-\frac{\partial}{\partial x_i}ln\,p_X(X)=-\frac{\frac{\partial}{\partial x_i}p_X(X)}{p_X(X)} \quad (9)$$

where $p_{x_i}(x_i)$ is the marginal probability density function (PDF) of x_i and $p_X(X)$ is the joint PDF of random vector X. SFD estimation is our main concern since it is the gradient of mutual information. Histogram estimation is the preferred method among the various techniques used to estimate SFD due to its acceptable accuracy despite its simplicity [20]. The GCMTF method is graphically depicted in Fig. 3.

The cost function gradient is computed using these equations:

$$\begin{aligned} &\partial l_0/\partial\lambda_r=(2/\varepsilon){\sigma_r}^2\lambda_r e^{-(\lambda_r\sigma_r)^2/\varepsilon}\left(1-NMI(t_{eer},t_{fmr})\right)\\ &\partial l_0/\partial\sigma_r=(2/\varepsilon)\lambda_r{}^2\sigma_r e^{-(\lambda_r\sigma_r)^2/\varepsilon}\left(1-NMI(t_{eer},t_{fmr})\right)\\ &\partial f/\partial T_{ee}=(T_{(1)}-X_{(1)})(\lambda^T\odot M_{ee}\odot F_{ee})+\alpha(T_{ee}-\bar{T}_{ee})\\ &\quad-\gamma\sum_{r=1}^{R}\left(1-e^{-(\lambda_r\sigma_r)^2/\varepsilon}\right)\partial NMI/\partial t_{eer}\\ &\partial f/\partial F_{ee}=(T_{(2)}-X_{(2)})(\lambda^T\odot M_{ee}\odot T_{ee})+\alpha(F_{ee}-\bar{F}_{ee})\\ &\partial f/\partial M_{ee}=(T_{(3)}-X_{(3)})(\lambda^T\odot f_{ee}\odot T_{ee})\\ &\quad+\alpha(M_{ee}-\bar{M}_{ee})\\ &\partial f/\partial T_{fm}=H^THT_{fm}\Sigma M_{fm}^TM_{fm}\Sigma^T-H^TY\Sigma M_{fm}\\ &\quad+\alpha(T_{fm}-\bar{T}_{fm})\\ &\quad-\gamma\sum_{r=1}^{R}\left(1-e^{-(\lambda_r\sigma_r)^2/\varepsilon}\right)\partial NMI/\partial t_{fmr}\\ &\partial f/\partial M_{fm}=M_{fm}\Sigma T_{fm}^TH^THT_{fm}\Sigma^T-Y^THT_{fm}\Sigma\\ &\quad+\alpha(M_{fm}-\bar{M}_{fm})\\ &\partial f/\partial\lambda_r=(\tau-\chi)\times_1 t_{eer}\times_2 f_{eer}\times_3 m_{eer}\\ &\quad+\left(\beta/2\right)\lambda_r/\sqrt{{\lambda_r}^2+\epsilon}+\partial l_0/\partial\lambda_r\\ &\partial f/\partial\sigma_r=HT_{fm}^T(HT_{fm}\Sigma M_{fm}^T-Y)M_{fm}\\ &\quad+\left(\beta/2\right)\sigma_r/\sqrt{{\sigma_r}^2+\epsilon}+\partial l_0/\partial\sigma_r \end{aligned} \quad (10)$$

where $\tau=[\![\lambda;T_{ee},F_{ee},M_{ee}]\!]$, $X_{(n)}$ is the tensor $\boldsymbol{\chi}$ unfolded in the nth mode, $\times_n$ defines the tensor-vector product in the nth mode, $\odot$ signifies the Khatri-Rao product and $\bar{T}_{ee}$ corresponds to T_{ee} with columns divided by their l_2-norms. Also term l_0 refers to the smoothed l_0-norm.

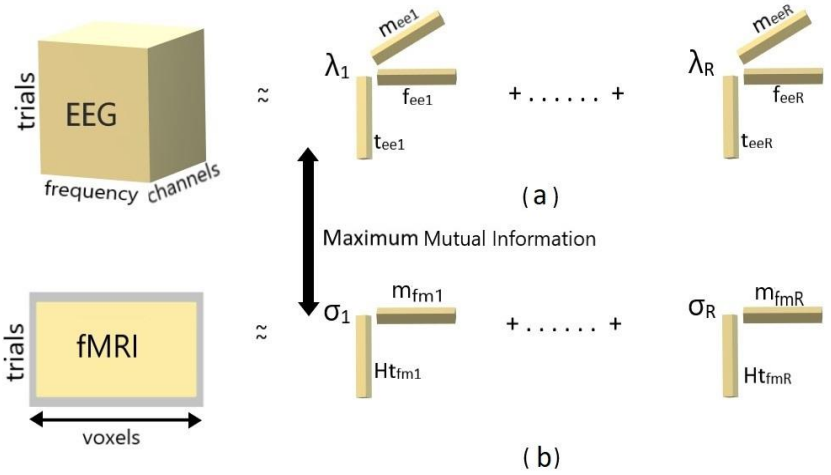


Fig. 3: Graphical representation of the GCMTF method with coupling in trial mode. (a) 3rd-order tensor EEG is represented by trials (t_{ee}), frequency (f_{ee}), and channels (m_{ee}). (b) Trials (t_{fm}) and voxels (m_{fm}) are used to index the matrix fMRI.

Results

The dataset used in this study is an oddball auditory paradigm derived from 17 healthy subjects [33]. The EEG data were denoted as a third-order tensor trial×frequency×channel and fMRI data as a matrix trial (scan) ×voxels. The EEG tensor for simulated data was generated using the same values as the real data [33]. The defined time windows (before and after the stimulus) were applied to the signals of each channel. These windows were transformed into a spectrogram using the Fourier transform. Then, based on the Canonical Polyadic Decomposition (CPD), these factors are multiplied to form the EEG tensor. To generate matrix fMRI, the trials were first convolved with canonical HRF and then multiplied with spatial factor (voxels). It is necessary to select the rank of the dataset before applying the method. Hence, using the Corcondia test, the number of components is selected to be 3. Now, it is assumed that all three components have significant values in both modalities. Thus, [1 1 1] was chosen as the values of λ and σ.

To evaluate the presented method's ability to identify shared components with different linear correlation levels, the components were selected as follows. The first two shared components have a linear correlation of 90%, the second ones have a linear correlation of 50%, and the third ones are considered to be the same. The temporal components of the two modalities with different linear correlation levels are shown in Fig. 4. The results of the GCMTF method were compared to the ACMTF method. To evaluate the robustness of our proposed method against noise, white Gaussian noise of different SNRs has been added to both the EEG and fMRI datasets. The SNR levels added to the data were selected from -15 dB to +15 dB. Although both methods have acceptable performance against different noise levels, the GCMTF method is more effective than the ACMTF method.

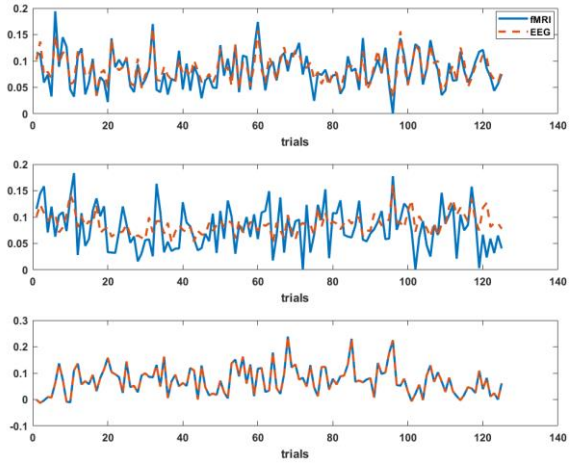


Fig. 4: From up to down, the temporal components of both modalities exhibit a linear correlation of 90%, 50%, and 100% (the same components).

Estimating the true components has decreased due to the assumption of equal components in the ACMTF method. Fig. 5 illustrates the weights of λ and σ estimated using the proposed GCMTF and ACMTF methods. Also, the shared components estimated by each method in the presence of high-level noise are depicted in Fig. 6 and Fig. 7.

In Fig. 8 the performance of the GCMTF method has been compared with the ACMTF method. The Match Score $MS = \frac{1}{R}\sum_{r=1}^{R} \hat{a}_r{}^T a_r \Big/ \|\hat{a}_r{}^T\| \|a_r\|$ is used to evaluate the results. In the MS relationship, $\hat{a}_r$ and a_r are the estimated and true values, respectively. The average match score for each simulated factor is illustrated in Fig. 8. The results indicate that the GCMTF method outperforms the ACMTF method. The average Match Score was raised by approximately 20% in the GCMTF model compared to the ACMTF model.

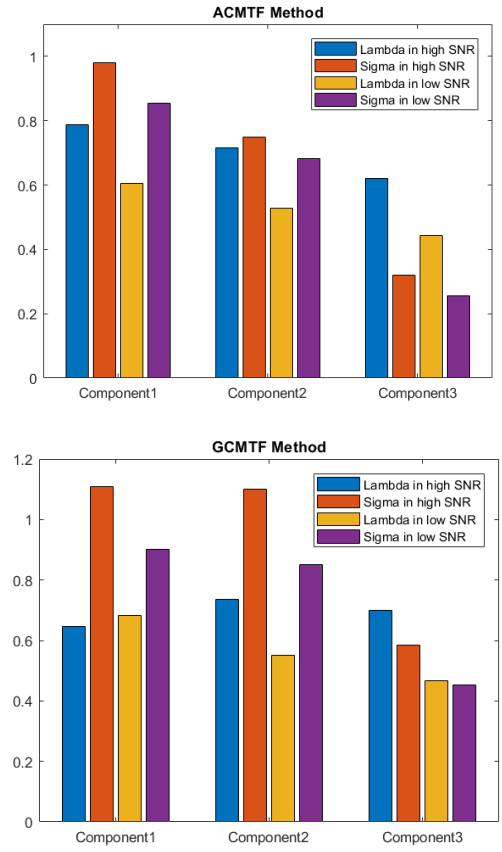


Fig. 5: The weights of λ and σ estimated using the proposed GCMTF and ACMTF methods.

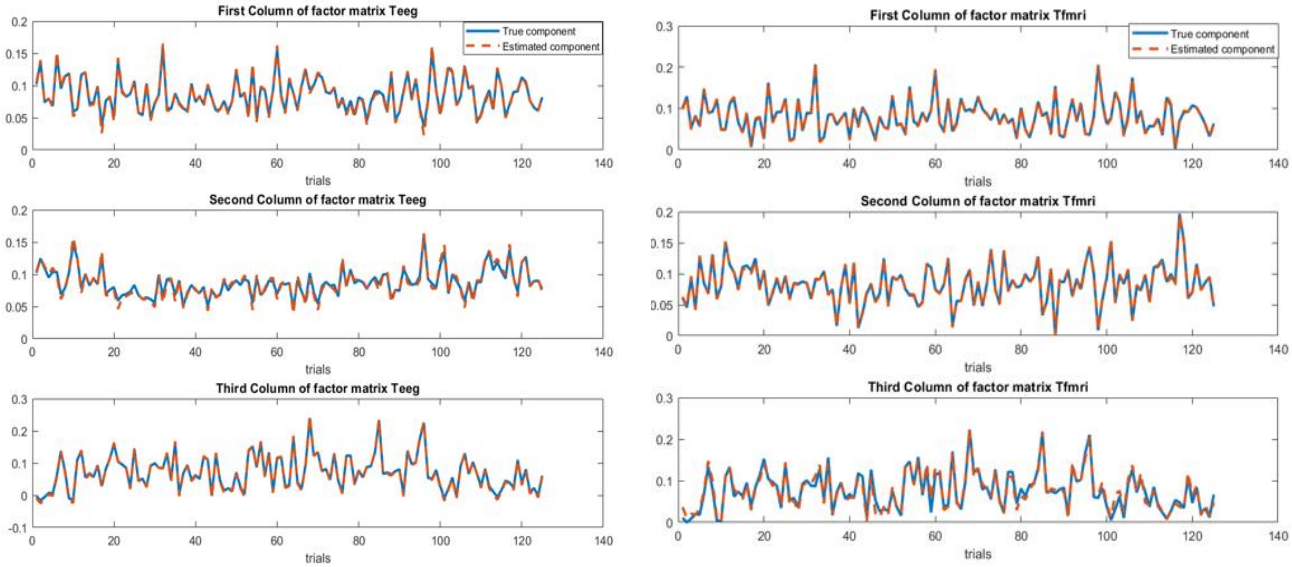
Discussion

The shared components in the common mode are assumed to be identical when fusing EEG and fMRI. However, this assumption is restrictive. Our proposed method replaced the equality assumption with a similarity measure.

Using normalized mutual information as a similarity measure, we can capture the differences between EEG and fMRI data that may not be fully accounted for by assuming identical components. This allows us to more accurately fuse information from both modalities and potentially uncover new insights that may have been overlooked with the traditional approach. Additionally, by quantifying the level of similarity between components, we can better understand the relationship between EEG and fMRI.

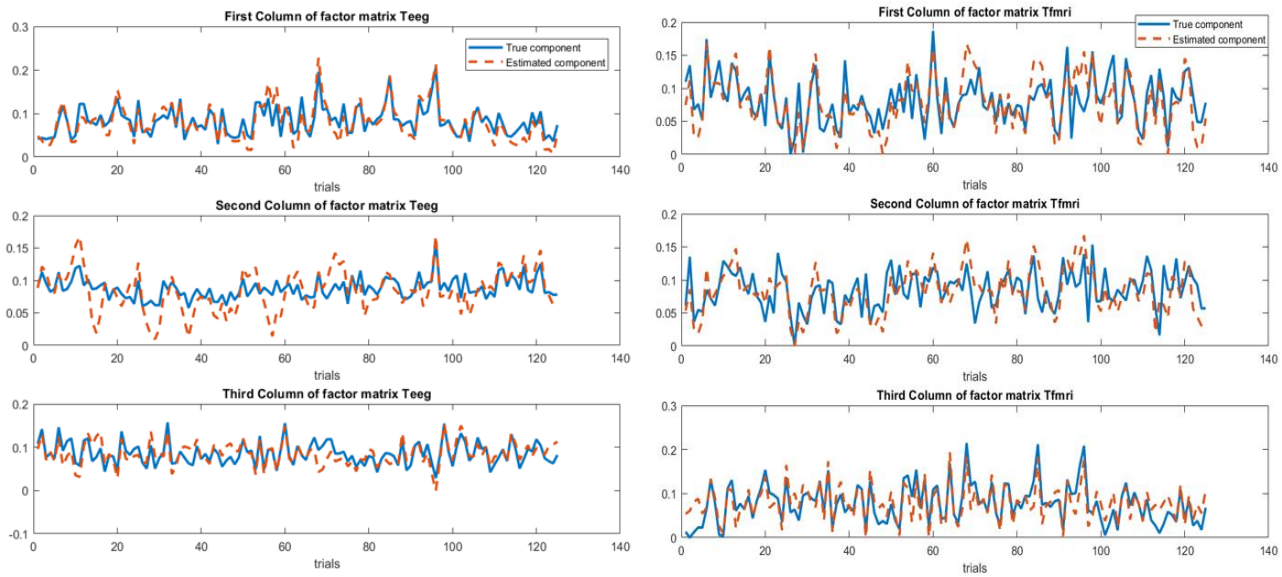
An NMI value approaching 1 indicates significant similarity, but when it nears zero, it means the opposite. Our simulations take into account three different levels of correlation.

The results indicate that our proposed GCMTF method significantly improves accurately estimating shared components with correlated temporal modes compared to existing methods like ACMTF. Moreover, the GCMTF method accurately estimates the weight values for each component corresponding to their existence in the dataset.



(a) (b)

Fig. 6: The components estimated by GCMTF method: (a) The estimated EEG components (T_{ee}) (b) The estimated fMRI components (T_{fm}).



(a) (b)

Fig. 7: The components estimated by ACMTF method: (a) The estimated EEG components (T_{ee}) (b) The estimated fMRI components (T_{fm}).

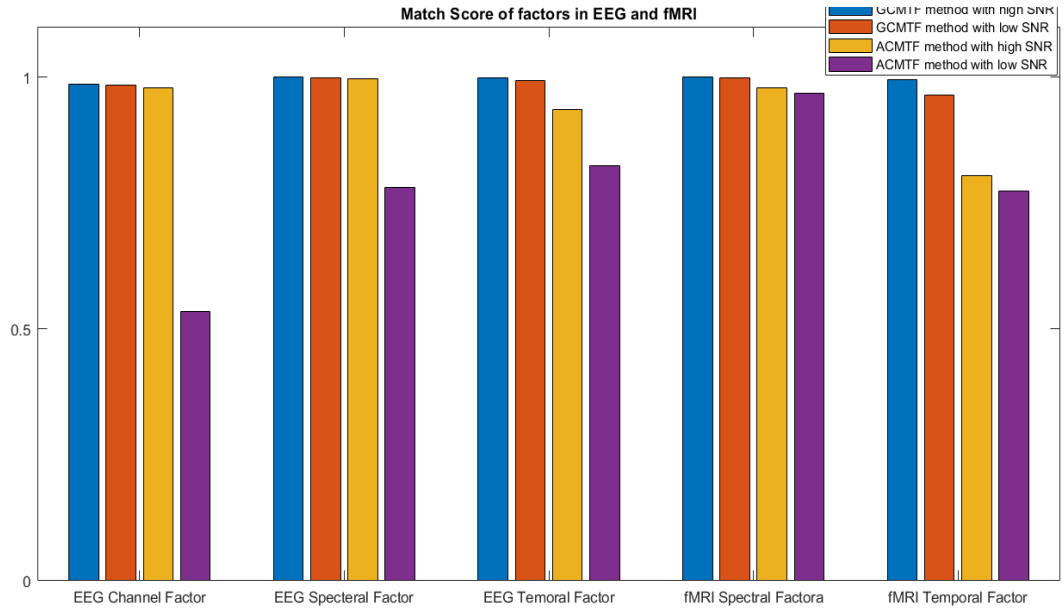


Fig. 8: Average match score (MS) between extracted factors by ACMTF and GCMTF and their ground truth at low and high SNR.

Conclusion

Overall, our proposed method offers a flexible and versatile approach to the fusion of multimodal data, providing a better understanding of the relationship between the two modalities. The application of this method can improve our capacity to study brain function in real EEG and fMRI data and opens up new possibilities for studying complex cognitive processes and neurological disorders. Although the GCMTF method is superior to the ACMTF, some modifications need to be made to improve its performance. Our method assumes that the HRF waveform is invariant for all brain voxels. However, the model can be more flexible by considering the variability in HRF across different subjects and brain regions in real data.

Furthermore, adopting other techniques to estimate MI rather than the histogram-based method may enhance the accuracy of estimating MI and entropy values. One alternative technique that could be explored

is kernel density estimation, which can provide a smoother and more continuous estimate of the underlying distribution. Additionally, advanced techniques such as neural networks or support vector machines could result in more precise and reliable estimations of mutual information and entropy in data analysis.

Author Contributions

Z. Rabiei: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data Curation, Visualization, Writing, Original Draft. H. Montazery Kordy: Conceptualization, Validation, Visualization, Editing the final version of the paper, and Supervision.

Acknowledgment

The authors would like to express their sincere gratitude to Dr. Raziyeh Mosayebi for her invaluable support, knowledge, and constructive feedback throughout this research project.

Conflict of Interest

The authors declare no potential conflict of interest regarding the publication of this work. The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Abbreviations

EEG	Electroencephalogram
fMRI	functional magnetic resonance imaging
BOLD	Blood oxygenation level-dependent
HRF	Hemodynamic Response Function
GLM	General linear model
ACMTF	Advanced coupled matrix tensor factorization
NMI	Normalized mutual information
GCMTF	Generalized coupled matrix tensor factorization
MS	Match score

References

[1] Z. Jiang, Y. Liu, W. Li, Y. Dai, L. Zou, "Integration of simultaneous fMRI and EEG source localization in emotional decision problems," Behav. Brain Res., 448: 114445, 2023.

[2] D. Lahat, T. l. Adali, C. Jutten, "Multimodal data fusion: an overview of methods, challenges and prospects," Proc. IEEE, 103: 1449-1477, 2015.

[3] T. Warbrick, "Simultaneous EEG-fMRI: what have we learned and what does the future hold?," Sensors, 22: 2262, 2022.

[4] S. Van Eyndhoven, B. l. Hunyadi, L. De Lathauwer, et al., "Flexible fusion of electroencephalography and functional magnetic resonance imaging: Revealing neural-hemodynamic coupling through structured matrix-tensor factorization," in Proc. 25th European Signal Processing Conference (EUSIPCO): 26-30, 2017.

[5] H. K. Aljobouri, "Independent component analysis with functional neuroscience data analysis," J. Biomed. Phys. Eng., 13: 169, 2023.

[6] J. Sui, D. Zhi, V. D. Calhoun, "Data-driven multimodal fusion: approaches and applications in psychiatric research," Psychoradiology, 3: kkad026, 2023.

[7] L. Du, H. Wang, J. Zhang, S. Zhang, L. Guo, J. Han, A. S. D. N. Initiative, "Adaptive structured sparse multiview canonical correlation analysis for multimodal brain imaging association identification," Sci. China Inf. Sci., 66: 142106, 2023.

[8] R. F. Silva, S. M. Plis, T. l. Adali M. S. Pattichis, V. D. Calhoun, "Multidataset independent subspace analysis with application to multimodal fusion," IEEE Trans. Image Process., 30: 588-602, 2020.

[9] Y. Jonmohamadi, S. Muthukumaraswamy, J. Chen et al., "Extraction of common task features in EEG-fMRI data using coupled tensor-tensor decomposition," Brain Topogr., 33(5): 636-650, 2020.

[10] G. R. Poudel, R. D. Jones, "Multimodal neuroimaging with simultaneous fMRI and EEG," in Handbook of Neuroengineering: Springer, pp. 1-23, 2021.

[11] E. Acar, T. G. Kolda, D. M. Dunlavy, "All-at-once optimization for coupled matrix and tensor factorizations," arXiv preprint arXiv:1105.3422, 2011.

[12] E. Acar, M. A. Rasmussen, F. Savorani, et al., "Understanding data fusion within the framework of coupled matrix and tensor factorizations," Chemom. Intell. Lab. Syst., 129: 53-63, 2013.

[13] E. Acar, G. z. Gurdeniz, M. A. Rasmussen et al., "Coupled matrix factorization with sparse factors to identify potential biomarkers in metabolomics," Int. J. Knowl. Discovery Bioinf. (IJKDB), 3(3): 1-22, 2012.

[14] E. Acar, E. E. Papalexakis, G. z. Gurdeniz et al., "Structure-revealing data fusion," BMC Bioinf., 15(1): 1-17, 2014.

[15] E. Karahan, P. A. Rojas-Lopez, M. L. Bringas-Vega et al., "Tensor analysis and fusion of multimodal brain images," Proc. IEEE, 103(9): 1531-1559, 2015.

[16] B. Rivet, M. Duda, A. Guerin-Dugue, et al., "Multimodal approach to estimate the ocular movements during EEG recordings: a coupled tensor factorization method," in Proc. 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society: 6983-6986, 2015.

[17] C. Chatzichristos, E. Kofidis, L. De Lathauwer et al., "Early soft and flexible fusion of EEG and fMRI via tensor decompositions," arXiv preprint arXiv:2005.07134, 2020.

[18] R. Mosayebi, G. A. Hossein-Zadeh, "Correlated coupled matrix tensor factorization method for simultaneous EEG-fMRI data fusion," Biomed. Signal Process. Control, 62: 102071, 2020.

[19] Y. Maeda, H. Kawaguchi, H. Tezuka, "Estimation of mutual information via quantum kernel method," arXiv preprint arXiv:2310.12396, 2023.

[20] M. Babaie-Zadeh, C. Jutten, "A general approach for mutual information minimization and its application to blind source separation," Signal Process., 85: 975-995, 2005.

[21] T. O. Kvalseth, "On normalized mutual information: measure derivations and properties," Entropy, 19: 631, 2017.

[22] A. Cichocki, D. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, H. A. Phan, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," IEEE Signal Process. Mag., 32: 145-163, 2015.

[23] N. M. Correa, T. Eichele, T. l. Adali, et al., "Multi-set canonical correlation analysis for the fusion of concurrent single trial ERP and functional MRI," Neuroimage, 50(4): 1438-1445, 2010.

[24] S. Van Eyndhoven, P. DuPont, S. Tousseyn, et al., "Augmenting interictal mapping with neurovascular coupling biomarkers by

structured factorization of epileptic EEG and fMRI data," Neuroimage, 228: 117652, 2021.

[25] M. Morante, "A lite parametric model for the hemodynamic response function," arXiv preprint arXiv:2004.13361, 2020.

[26] D. A. Handwerker, J. M. Ollinger, M. D'Esposito, "Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses," Neuroimage, 21: 1639-1651, 2004.

[27] Z. Y. Shan, M. J. Wright, P. M. Thompson, K. L. McMahon, G. G. Blokland, G. I. De Zubicaray, N. G. Martin, A. A. Vinkhuyzen, D. C. Reutens, "Modeling of the hemodynamic responses in block design fMRI studies," J. Cereb. Blood Flow Metab., 34: 316-324, 2014.

[28] M. W. Woolrich, T. E. Behrens, S. M. Smith, "Constrained linear basis sets for HRF modeling using Variational Bayes," Neuroimage, 21: 1748-1761, 2004.

[29] C. Gössl, L. Fahrmeir, D. P. Auer, "Bayesian modeling of the hemodynamic response function in BOLD fMRI," Neuroimage, 14: 140-148, 2001.

[30] C. Goutte, F. A. Nielsen, K. Hansen, "Modeling the hemodynamic response in fMRI using smooth FIR filters," IEEE Trans. Med. Imaging, 19: 1188-1201, 2000.

[31] H. Mohimani, M. Babaie-Zadeh, C. Jutten, "A fast approach for overcomplete sparse decomposition based on smoothed l_0-norm," IEEE Trans. Signal Process., 57: 289-301, 2008.

[32] M. Babaie-Zadeh, C. Jutten, K. Nayebi, "Differential of the mutual information," IEEE Signal Process. Lett., 11: 48-51, 2004.

[33] J. M. Walz, R. I. Goldman, M. Carapezza, J. Muraskin, T. R. Brown, P. Sajda, "Simultaneous EEG-fMRI reveals a temporal cascade of task-related and default-mode activations during a simple target detection task," Neuroimage, 102: 229-239, 2014.

Biographies

Zahra Rabiei received her B.Sc. degree in Electronic Engineering from KN Toosi University of Technology,Tehran, Iran in 2001 and M.Sc. degree in Control Engineering from Ferdowsi university of Mashhad, Iran in 2005. She is currently a Ph.D. student in Biomedical Engineering in Babol Noshirvani University of Technology, Babol, Iran. The focus of her research is on biomedical signal processing, EEG, fMRI, and data fusion.

- Email: z.rabiei@stu.nit.ac.ir
- ORCID: 0000-0002-2048-8942
- Web of Science Researcher ID: NA
- Scopus Author ID:NA
- Homepage: NA

Hussain Montazery Kordy received his B.S. degree in Electronic Engineering from Mazandaran University, Babol, in 2000, the M.S. degree in Biomedical Engineering from Sharif University of Technology, in 2003 and the Ph.D. degree in Biomedical Engineering from Tarbiat Modarres University, Tehran, Iran, in 2009. Since 2010, he has been a member of the Electrical and Computer Engineering Faculty, Babol Noshirvani University of Technology, Babol, Iran. His research focuses on computer aided diagnosis, feature selection and extraction, and biomedical signal and image processing.

- Email: hmontazery@nit.ac.ir
- ORCID: 0000-0002-2010-4945
- Web of Science Researcher ID: AAD-3933-2022
- Scopus Author ID: 54386905700
- Homepage: https://ostad.nit.ac.ir/home.php?sp=389010

How to cite this paper:
Z. Rabiei, H. Montazery Kordy, "Utilizing normalized mutual information as a similarity measure for EEG and fMRI fusion," J. Electr. Comput. Eng. Innovations, 13(1): 141-150, 2025.

DOI: 10.22061/jecei.2024.10984.754

URL: https://jecei.sru.ac.ir/article_2207.html