Jaw-Operated Human Computer Interface Based on EEG Signals via Artificial Neural Networks

Muhammet Serdar Bascil
Department of Electrical and Electronics Engineering, Yozgat Bozok University, Yozgat 66200, Turkey

Corresponding Author Email: serdar.bascil@bozok.edu.tr

https://doi.org/10.18280/ria.340103

Received: 18 September 2019
Accepted: 5 December 2019

Keywords:
EEG, Jaw-Machine interface (JMI), PCA, MLNN+LM, PNN

ABSTRACT
Man and machine interfaces help paralyzed peoples to communicate and control their environment. This work purpose to introduce and shows a novel kind of machine interface using the horizontal signs of conscious jaw motions on brain signals stored in electroencephalogram (EEG). Electrical functions of the brain are extricated and transformed to control commands. Jaw-Machine Interface (JMI) serve a new functionality for tetraparesis to run peripheral devices with the help of a computer using only horizontal jaw motions. In this study, mean absolute deviation (MAD) and entropy (S) values are derived of EEG and hemispherical designs are valued and examined as offline analysis approach. Principle component analysis (PCA) is used to reduce redundant information from data and two types of artificial neural networks which are Multilayer Neural Network with Levenberg Marquardt training algorithm (MLNN+LM) and Probabilistic Neural Network (PNN) via k-fold method are run to find out horizontal jaw patterns on brain waves.

1. INTRODUCTION

Man and machine interactions enable handicapped persons to control and communicate environmental devices such as a computer, a cursor, a wheelchair, a tv etc. [1]. Patients suffer from the tetraparesis do not use both of the arms and legs [2]. Jaw-machine interface (JMI) is a jaw-operated system to manage assistive devices by voluntary jaw motions to help tetraparesis and makes their life easier.

In the literature, some research utilizes EEG based on jaw motions. But all of them focus to extract dominantly appearing and recognizing changings on both EEG and electromyogram (EMG) signals. A wearable head band was designed by Wei et al., to detect facial movements on EMG and electrooculogram (EOG) signals. They tried to control a wheelchair with 5 different stimulates named movement patterns [3]. In their next work they introduced a new control way to detect facial movements by using facial image processing together with EMG signals. By using a route following an experimental setup, they evaluate the performance of an electrical wheelchair [4]. Wei and Hu reported that they developed their previous studies on EMG and image processing by creating software simulation environment to extract brain intentions and declared high performance [5]. Emotive EPOC headset that measures EEG activity and recorded head movements and facial expressions was used Rechy-Ramirez and Hu. A hands-free electrical wheelchair benefited was executed via a graphical user interface [6]. Jeong et al., utilized surface electromyography (sEMG) to obtain high quality EMG signals to control a quadrotor properly. They recorded these signals any region of the body part motions such as forearm, forehead, back of the neck, index finger, jaw clenching and mouth motions [7]. A wearable headband which has twenty passive electrode networks was produced by Paul et al. They recorded horizontal eye movements, gestures and jaw clenching signals to extract user intentions on EMG and EOG. They presented that a cursor or keyboard functions can be control [8]. Costa et al., showed to control a robotic arm in two dimensions by extracting EMG signals stored in EEG recordings produced with 5 different jaw clenching [9]. Zeilfelder et al., placed a BMP280 digital pressure sensor in the outer ear canal to recognize tongue and jaw motions. They tried the system on six participant reported that tongue and jaw movements are suitable in put mechanisms for man and machine interfaces [10]. According to literature, the systems operated by the jaw have some equipment around the jaw. These devices may disturb disable, uncomfortable and interfere with the speech [11, 12].

This paper aims to enlarge of jaw machine interface (JMI) was the first introduced to literature in the previous study [13]. The signals recorded as differential multi-channel from scalp and features of them extracted by MAD and S values in time domain analysis. The hemispherical patterns were recognized by using MLNN+LM and PNN

2. METHODOLOGY

2.1 Study protocol

The EEG recordings of 18 differential channels were measured over the scalp according to the 10-20 electrode placement system [14]. A1-A2 ear electrodes described as reference. Each channel was sampled at 1024 Hz, filtered 0.3 Hz high pass, 70 Hz low pass and digitized 16 bits. Also, 50 Hz notch filter was applied for elimination of power line noise. EEG signal features on the raw dataset of voluntary jaw movements were achieved by MAD and S values. PCA method is used to reduce feature dimension by selecting most significant ones. The best chosen 8 PCA vectors were then
conveyed to MLNN+LM and PNN algorithms to recognize hemispherical changes. After all, the results are discussed and concluded. The workflow of JMI was presented in Figure 1.

2.1 Data collection

In this study, an experimental paradigm was made to 7 male and 3 female volunteers of healthy subjects with ages in the range of 25-35 years without any nervous system impairment. All the subjects were seated in front of the LCD which have just direction stimulus and instructed not to move any part of the body except jaw motions as given in Figure 2. JMI system has two output that voluntary control (VC) status and no control (NC) status. JMI system output is activated on VC status by the user and has no output on NC status. As seen on Figure 1, subjects move their jaws distinctly and serially (closed lips and no teeth grinding) from the case of 2 to 1 to perform the left control task and from the case of 2 to 3 to perform the right control task.

The experimental paradigm shown in Figure 2, has two trial. All the subjects implement initially 1st trial moves from left to right and then 2nd trial moves inverse. Each recording sequences of two trial last 120s starts with subject relaxation and each consists of 8 VC parts (4 right and 4 left) of 10s separated with 5s NC status. Subjects were performed VC status nearly 20 times during 10s. All of them performed 160 times left and 160 times right (20x8=160) jaw movements through experimental paradigm. So, every control task consists of 512 (1s/2*1024) samples. At last, EEG recordings was created as 320 different dataset (160 right VC and 16 left VC) each has 18 channel and each channel has 512 samples for each subject.

3. Feature Extraction

The informative and non-redundant feature vector was formed form the raw data signal on this topic. In this study, the mean-absolute deviation (MAD) and entropy (S) values were performed. The MAD is used for calculation of demand variability [15] as the mathematical presentation defined in Eq. (1).

\[
MAD = \frac{1}{N} \sum_{i=1}^{N} |x_i|
\]

(1)
where, \( X_i = 1, 2, 3 \ldots N \) is a time sequence, and \( N \) stands for the length of samples in the data array.

The term of entropy was introduced by Claude Shannon in his 1948 paper [16]. It holds the information of how to distinguish each signal in an orbit according to another signal behaviours [17] as given the formula below.

\[
S(X) = - \sum_{i=1}^{N} p(X_i) \log p(X_i) \tag{2}
\]

where, \( X_i \) is time series of the signal, \( p(X_i) \) possible probabilities and \( N \) is length of the signal.

High dimensional features of EEG dataset \((320 \times 512)\) are effectively extracted \((320 \times 18)\) by the aid of MAD and S concepts.

### 4. PRINCIPAL COMPONENT ANALYSIS

Principal component analysis (PCA) is a statistical method removing the redundant information from data and reduces dimensionality [18]. Converting the higher dimensional data \((X)\) into a lower dimensional one \((S)\) is carried out by PCA determining the eigenvalues and eigenvectors of the covariance matrix \((C)\). The equations of PCA is shown in (3-5).

\[
C(X) = \sum_{1}^{N} \frac{(X_i X^T)}{N} \tag{3}
\]

\[
u_i = Cu_i, \quad i = 1, 2, 3 \ldots m \tag{4}
\]

where, \( \lambda_i \) stands for the eigenvalue of the covariance matrix \((C)\) and \( u_i \) is the corresponding eigenvector.

\[
S_i(i) = u_i^T X_i, \quad i = 1, 2, 3 \ldots m \tag{5}
\]

where, \( S_i(i) \) describes the principal components of the data set \((X)\) and more information about PCA can be found in the study [19]. In this work, the highest variance values were selected to create a new data set using only 8 dimensions \((320 \times 8)\) by PCA indicating 98.54% of the information shown in Figure 3. Thus, it provides an easier calculation for machine learning algorithms.

![Figure 3. Feature selection by PCA](image)

### 5. MACHINE LEARNING METHODS

Multilayer Neural Network with Levenberg Marquardt training algorithm (MLNN+LM) and Probabilistic Neural Network (PNN) are the kinds of neural network and used in classification and pattern recognition processes. In the study, both of the structures incorporate in k-cross validation technique. All dataset is divided into k pieces \((k=10)\) randomly with this technique. Neural networks are repeated k-times in the independence of selection for samples [20]. Mathematical equations of classification accuracy given in (6), (7) and (8):

\[
\text{Accuracy}(TS) = \frac{\sum_{i=1}^{TS} \text{estimate}(n)}{|TS|}, \quad n_i \in TS \tag{6}
\]

\[
\text{estimate}(n) = \begin{cases} 1, & \text{if estimate}(n) = cn \\ 0, & \text{otherwise} \end{cases} \tag{7}
\]

\[
\text{Classification accuracy (ML)} = \frac{\sum_{i=1}^{k} \text{accuracy}(TS_i)}{|k|} \tag{8}
\]

where, \( TS \) refers the test data set to be classified, while \( n \in TS \), \( c_n \) is the class of \( n \) and \( \text{estimate}(n) \) stands for the classification result of \( n \) estimated by neural networks.

#### 5.1 Multilayer neural network

Multilayer neural network is formed an input layer, hidden layers and an output layer. The logarithmic sigmoid activation function is applied to weighted input vector and the output is generated [21, 22].

MLNN with Levenberg Marquardt (LM) training algorithm was used to recognize left and right jaw patterns. Levenberg Marquardt is a very fast and efficient training methods [23]. It helps to provide lower hidden layers and maximize the generalization ability. In this work, two hidden layers approach \((20-50 \text{ neurons})\) was chosen because it has the better convergence than the first order. Also, hidden neuron numbers were selected randomly while obtaining the best results [23, 24].

#### 5.2 Probabilistic neural network

PNN was developed by Donald Specht [25] and known as a distance-based neural network and offers faster classification results. The first layer is input layer of PNN, the second layer is radial bases layer realizes the distances between the input vector and rows in the weight matrix, and the las is competitive layer determines the classification with maximum probability of correctness. The smoothing parameter is crucial in the PNN classifier hence an appropriate smoothing parameter can often be dependent on data [26]. In this work, a random search method between 0.1 and 1 was made in 0.01 steps to ensure the best performance of the PNN and to find the optimal value of the spread of activation functions.

### 6. RESULTS AND DISCUSSION

In this study, one dimensional (horizontal) jaw movements has been examined to improve and enlarge the previous study [13] on JMI. Subjects-8 (Sub-8) has the best and Subject-3 (Sub-3) has the worst classification results. The explanation of these participant results compared and clarified to avoid confusion.

The raw EEG dataset \((320 \times 512)\) of each participant were converted \(320 \times 18\) dimension with the help of MAD and S characteristics. This new feature vector fed into PCA to eliminate redundant data and select most significant information. PCA removed the redundant information from data and reduced dimensionality as \(320 \times 8\) given in Figure 3 indicating 98.54% of the information. At last, MLNN+LM and
PNN structures performed on the size of 320x8 MAD and S features to obtain classification results. Also, together with neural networks results true positive rates (TPRs) and false positive rates (FPRs) were calculated to check classification results given in (9), (10) respectively.

$$\text{TPR} = \frac{TP}{TP + FN}$$  \quad (9)
$$\text{FPR} = \frac{FP}{FP + TN}$$  \quad (10)

TPRs mean the rate of the correctly classified VC status and FPRs mean the rate of the misclassified NC status. All the performance results of classifiers are given in Table 1.

As seen on Table 1, Sub-8 has the best classification performance, TPRs and the minimum FPRs. For that reason, Sub-8 is named as the best participant but, it can’t be said for Sub-3. He has the lowest performances and he is marked as the least successful participant of all seen on Table 1.

Sub-8 reached the classification accuracy (Acc.) with 90.03% on horizontal jaw movement task, the highest TPR with 94.12% and the lowest FPR with 5.74% via PNN on MAD features and Sub-3 has 81.69% Acc., 85.05% TPR and 13.77% FPR. The lowest results of Sub-8 were appeared with 84.33% and Sub-3 performed to the worst performance of all with 76.58% in MLNN+LM on S features.

It is obvious that, MAD enables better performance than S on neural networks for all subjects and PNN arrives higher accuracies than MLNN+LM. When the Average results compared, the advantage of PNN on MLNN+LM is observed (1.4%) on MAD once again. Also, TPRs and FPRs displayed a similar tendency. The last column of Table 1 indicates the globally trained classifier (Sub-1+2+…+10) performance and it suggests that PNN is more suitable than MLNN+LM on MAD once and again and the global data performance 80.28% is the lowest of all. It can be deduced that personal usage of the system is more suitable to recognize voluntary jaw patterns. Figure 4 shows the best classification results of Sub-8, Sub-3, Average, and Global Data.

The neural network results of JMI are supported with the brain mapping results given below for Sub-8 and Sub-3. It is known that the right hemisphere of the brain controls the left motor activities and the other way around. The existence of delta frequencies seen on frontal and occipital lobes in changing powers anti-symmetrically on the brain hemispheres. In the literature, jaw clenching, teeth grinding and jaw movements dominantly seen on EEG named as artifacts [27-29]. But recent studies showed that they are not only qualified as artifact physical force but also have significant information on delta frequency band [13, 30-32].

![The Best Performances (PNN on MAD %)](image)

**Figure 4.** The best results of Sub-8, Sub-3, average and global data

Figure 5 and 6 show brain signal intensities with EEG bands on conscious jaw motions on the right and left of Sub-8 respectively. The left hemisphere of the brain is active on right motions and the other way around in parallel with the literature and the delta frequencies are dominantly seen on the frontal and occipital lobes of the brain. The intensities of delta on occipital lobes can be understood that the Sub-8 followed successfully experimental paradigm on LCD. The different intensities of delta on the scalp can be explained as Sub-8 jaw fatigue during trials. Also, it can be said that Sub-8 may be right-handed according to the more delta intensities on left hemisphere than the right.

Figure 7 and 8 show brain signal intensities with EEG bands on conscious jaw motions on the right and left of Sub-3 respectively. The left hemisphere of the occipital is active as it is expected and Sub-3 followed experimental paradigm sufficiently on LCD but, it is not the same on the left hemisphere of the frontal lobe. It can be explained that Sub-3 was collapsed on the right motion of jaw during trials. Also, the delta waves seen on temporal lobes similarly on frontal lobes. These results can be interpreted as Sub-3 is more susceptible to external stimulus or have distractibility.

It is obvious that Sub-8 brain signal intensities of voluntary jaw motion on the right and left hemisphere are clearer and order than Sub-3 has. By this means and thanks to the neural network results given in Table 1, Sub-8 is named as the best successful participant and Sub-3 is the worst.

### Table 1. Performance results for all subject

<table>
<thead>
<tr>
<th>Methods</th>
<th>Sub-1</th>
<th>Sub-2</th>
<th>Sub-3</th>
<th>Sub-4</th>
<th>Sub-5</th>
<th>Sub-6</th>
<th>Sub-7</th>
<th>Sub-8</th>
<th>Sub-9</th>
<th>Sub-10</th>
<th>Average</th>
<th>Global Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAD (MAD)</td>
<td>Acc.</td>
<td>83.65</td>
<td>88.27</td>
<td>81.05</td>
<td>85.66</td>
<td>87.33</td>
<td>82.56</td>
<td>84.12</td>
<td>89.05</td>
<td>81.55</td>
<td>86.94</td>
<td>85.01</td>
</tr>
<tr>
<td>TPRs</td>
<td>88.28</td>
<td>92.86</td>
<td>84.94</td>
<td>90.81</td>
<td>92.56</td>
<td>87.38</td>
<td>89.12</td>
<td>92.94</td>
<td>86.11</td>
<td>91.33</td>
<td>89.63</td>
<td>83.55</td>
</tr>
<tr>
<td>FPRs</td>
<td>8.11</td>
<td>5.68</td>
<td>13.48</td>
<td>6.94</td>
<td>6.13</td>
<td>8.76</td>
<td>7.55</td>
<td>6.11</td>
<td>6.12</td>
<td>7.1</td>
<td>15.64</td>
<td></td>
</tr>
<tr>
<td>S MAD (MAD)</td>
<td>Acc.</td>
<td>81.55</td>
<td>84.15</td>
<td>76.58</td>
<td>83.03</td>
<td>83.22</td>
<td>80.88</td>
<td>82.44</td>
<td>84.33</td>
<td>80.05</td>
<td>83.83</td>
<td>81.99</td>
</tr>
<tr>
<td>TPRs</td>
<td>86.48</td>
<td>89.76</td>
<td>81.77</td>
<td>87.95</td>
<td>88.38</td>
<td>85.05</td>
<td>87.05</td>
<td>89.04</td>
<td>84.28</td>
<td>88.55</td>
<td>86.83</td>
<td>80.33</td>
</tr>
<tr>
<td>FPRs</td>
<td>9.87</td>
<td>7.05</td>
<td>14.44</td>
<td>8.12</td>
<td>7.28</td>
<td>10.16</td>
<td>9.16</td>
<td>6.45</td>
<td>10.94</td>
<td>7.86</td>
<td>9.13</td>
<td>17.88</td>
</tr>
<tr>
<td>S MAD (MAD)</td>
<td>Acc.</td>
<td>84.92</td>
<td>89.1</td>
<td>81.69</td>
<td>87.76</td>
<td>88.88</td>
<td>84.33</td>
<td>86.05</td>
<td>90.03</td>
<td>83.08</td>
<td>88.33</td>
<td>86.41</td>
</tr>
<tr>
<td>TPRs</td>
<td>89.03</td>
<td>93.52</td>
<td>85.05</td>
<td>91.64</td>
<td>93.01</td>
<td>88.16</td>
<td>90.79</td>
<td>94.12</td>
<td>86.94</td>
<td>92.68</td>
<td>89.49</td>
<td>84.94</td>
</tr>
<tr>
<td>FPRs</td>
<td>8.59</td>
<td>6.02</td>
<td>13.77</td>
<td>7.77</td>
<td>6.94</td>
<td>9.22</td>
<td>8.15</td>
<td>5.74</td>
<td>10.2</td>
<td>7.21</td>
<td>8.36</td>
<td>13.28</td>
</tr>
<tr>
<td>PNN (%)</td>
<td>Acc.</td>
<td>81.55</td>
<td>84.68</td>
<td>77.05</td>
<td>83.18</td>
<td>84.02</td>
<td>81.66</td>
<td>82.12</td>
<td>85.05</td>
<td>80.96</td>
<td>83.56</td>
<td>82.38</td>
</tr>
<tr>
<td>TPRs</td>
<td>86.72</td>
<td>89.05</td>
<td>82.14</td>
<td>87.65</td>
<td>89.38</td>
<td>85.55</td>
<td>86.55</td>
<td>90.67</td>
<td>84.33</td>
<td>88.44</td>
<td>87.04</td>
<td>81.96</td>
</tr>
</tbody>
</table>
Figure 5. Brain signal intensities of Sub-8 on conscious jaw motion to the right

Figure 6. Brain signal intensities of Sub-8 on conscious jaw motion to the left

Figure 7. Brain signal intensities of Sub-3 on conscious jaw motion to the right
ACKNOWLEDGMENT

The author would like to thanks the subjects of the University of Bozok for providing the participation for this research.

The study was approved by the Ethical Committee of Bozok University. All procedures performed in study involving human participants were in accordance with the ethical standards of the institutional and/or national research committee. Informed consent was obtained from all individual participants included in the study.

REFERENCES


7. CONCLUSION

This study means to enlarge and improve the previous work on one dimensional JMI. The main goal of the work is to upgrade quality of life and to present a new communication pathway for tetraparesis by not having some obtrusive equipments around the jaw such as electrodes.

The brain waves stored on EEG were acquired from ten participants by generating conscious jaw motions distinctly and serially with the help of experimental paradigm. MAD and S characteristics of these signals were extracted and the most important features were selected by PCA. Then, MLNN+LM and PNN classifiers were run on these features to find out horizontal jaw patterns on EEG and to get 1D control. At last, it is realized that PNN on MAD structure is the best choice for this aim according to results given in Fig. 6. Also, the personal usage of the system is more convenient to implement 1-D control by using voluntary jaw motions according to the lowest classification results of global data given in Table 1.

After all, PNN on MAD structure has fairly parallel results to the SVM on RSM structure studied in previous study. It offers a different view point, makes a contribute to the literature and encourage the JMI studies for feature developments.


