# Improved EEG Classification by factoring in sensor topography

Lubna Shibly Mokatren, *Student Member, IEEE,* Rashid Ansari, *Fellow, IEEE,* Ahmet Enis Cetin, *Fellow, IEEE,*
Alex D. Leow, *Member, IEEE,* Olusola Ajilore, Heide Klumpp, *Member, IEEE,*
and Fatos T.Yarman Vural, *Senior Member, IEEE*

*Abstract*—Electroencephalography (EEG) serves as an effective diagnostic tool for mental disorders and neurological abnormalities. Enhanced analysis and classification of EEG signals can help improve detection performance. This work presents a new approach that seeks to exploit the knowledge of EEG sensor spatial configuration to achieve higher detection accuracy. Two classification models, one which ignores the configuration (model 1) and one that exploits it with different interpolation methods (model 2), are studied. The analysis is based on the information content of these signals represented in two different ways: concatenation of the channels of the frequency bands and an image-like 2D representation of the EEG channel locations. Performance of these models is examined on two tasks, social anxiety disorder (SAD) detection, and emotion recognition using DEAP dataset. Validity of our hypothesis that model 2 will significantly outperform model 1 is borne out in the results, with accuracy 5–8% higher for model 2 for each machine learning algorithm we investigated. Convolutional Neural Networks (CNN) were found to provide much better performance than SVM and kNNs.

*Index Terms*—EEG, SAD, emotion recognition, deep learning, classification.

## I. INTRODUCTION[1]

**T**HE use of electroencephalography (EEG) is a popular mechanism for diagnosing mental states and brain disorders. EEG provides measurements of brain activities acquired using electrodes placed over the scalp. While other brain imaging techniques, such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), are used in diagnosis, EEG has important attributes in that it captures the temporal activity of the brain and is more affordable compared with other methods [2]. The EEG waveform is divided into five main frequency bands [3]: Delta ($\delta$: up to 4 Hz), Theta ($\theta$: 4-8 Hz) waves, Alpha ($\alpha$: 8 - 15 Hz) waves, Beta ($\beta$: 15-32 Hz) waves, and Gamma ($\gamma$: $\geq$ 32 Hz) waves. EEG is a very popular, noninvasive monitoring method which plays an important role as a diagnostic tool in braincomputer interface (BCI) applications. It is used to evaluate various mental disorders, such as Alzheimers disease, strokes, migraine, sleep disorders, and Parkinsons disease [4]. However, the analysis

L. Shibly Mokatren, R. Ansari and A. Cetin are with the Department of Electrical and Computer Engineering, University of Illinois at Chicago, Chicago

A. D. Leow,O. Ajilore and H. Klumpp are Department of Psychiatry, University of Illinois at Chicago, Chicago

F. T.Yarman Vural was with Department of Computer Engineering, Middle East Technical University, Ankara, Turkey

[1]A preliminary version of this manuscript is presented in IEEE EMBS Conference on Neural Engineering [1]

process of EEG is not always accurate as the data are complex and degraded by noise and artifacts. Therefore, providing a new model to help improve the accuracy of EEG analysis is very crucial. In the past, many classification algorithms were devised for using EEG data [5], such as, linear discriminant analysis, neural networks, SVM, nonlinear bayesian classifiers, kNN, hidden markov model, combination of classifiers, and other algorithms for EEG-based BCI which are mostly based on machine learning [6]. In their study, Samiee et al. [7] presented a low cost approach to classify long-term epileptic EEG records using SVM. Flumeri et al. [8] analyzed the cognitive and mental workload in real driving settings using automatic-stop-StepWise Linear Discriminant classifier. In their work, Oh et al. [9] presented an automated system for Parkinson's disease detection using deep neural networks. While there are many studies involving EEG classification, none considered the spatial locations and configuration of the EEG channel sensors as a means to possibly achieve better accuracy in analysis or classification tasks. Factoring in the sensor topography was a key driving factor in our research. Two classification models, one which ignores the configuration (model 1) and one that exploits it with different interpolation methods (model 2), were considered. We hypothesize that model 2 would significantly outperform model 1. Validity of our hypothesis is borne out in the results, with average accuracy 5–8% higher for model 2 for each machine learning algorithm we investigated. Convolutional Neural Networks (CNN) were found to provide much better performance than SVM and kNNs. To demonstrate the significance of EEG analysis, two major classification tasks are examined: Social Anxiety Disorder (SAD) detection, and emotion recognition using DEAP dataset, which are detailed next.

### A. SAD task

SAD, the world's third largest mental health problem, affects 7% of the population [10]. It is characterized by extreme avoidance of social situations and the fear of negative evaluations from others [11]. The diagnosis process of SAD was first characterized in 1980 by Diagnosis and Statistical Manual for Mental Disorders (DSM-III). However, the criteria evolved and the most recent description appears in the fifth edition of the manual (DSM-5) [12]. In the field of psychiatry, the reliability and quality of the diagnostic process of SAD DSM-5 are critical for getting an accurate assessment of the disorder [13]. The use of EEG in SAD diagnosis has

seen only limited study. Identifying SAD patients by visual detection of differences in the EEG signals is impractical. Therefore, automated detection using techniques such as machine learning are usually employed, which could lead to more accurate diagnosis, better connectivity analysis, and improved understanding of treatment responses in SAD [14].

### B. Emotion recognition task

The study of EEG-based emotion recognition is very popular in many fields such as psychology, neuroscience, and computer science. Emotions are a very important factor in correct interpretation of actions and play a crucial role in all-day communication. There exist several recent research studies about EEG-based emotion recognition systems. Piho and Tjahjadi [15] investigated reduced EEG data of emotions using a mutual information-based adaptive windowing and achieved average accuracy of $89.61\%$ and $89.84\%$ for valence and arousal, respectively. Chao et al. [16] integrated deep belief networks with glia chains learning framework using multichannel EEG data and achieved average accuracy of $76.83\%$ and $75.92\%$ for valence and arousal states classification, respectively. In their study, Xu J. et al. [17] proposed a baseline strategy of using power spectral density feature extraction methods and CNNs and obtained $81.14\%$ and $77.69\%$ for valence and arousal, respectively. Another recent study on DEAP dataset was conducted by Ganapathy and Swaminathan [18] using electrodermal activity signals and multiscale deep CNNs to achieve a classification accuracy of $81.25\%$ and $83.75\%$ for valence and arousal, respectively. Although there are numerous studies related to this field, the efficiency of some of the algorithms is limited.

EEG-based emotion recognition task can be subject-dependent or subject-independent [19]. In this paper, both subject-dependent and subject-independent approaches are investigated.

## II. METHOD

### A. EEG recording

*1) SAD dataset:* The EEG dataset used in this paper was acquired from the Department of Psychiatry at the University of Illinois at Chicago (UIC). The acquisition of multi-channel EEG is done using an Electro-Cap with electrodes (sensors) positioned at 34 different locations i.e. channels. The data are gathered from a total of 64 subjects divided into control and SAD patient groups. For this study, the brain activity being analyzed is at resting state, without the introduction of stimuli or task instructions. The duration of the EEG recording varied from 2-7 minutes. The digitized signals are acquired at a sampling frequency of 1024 Hz.

*2) DEAP dataset:* The DEAP dataset is used for emotion analysis. Specifically, in this research, valence and arousal were assessed. DEAP is a publicly available EEG dataset [20] which contains signals from 32 participants. Each participant watched 40 one-minute long videos and evaluated themselves on the basis of four emotion states: arousal, valence, liking, and dominance on a scale of 1-9. The data is recorded over 32 channels on the scalp. For our analysis, a classification task
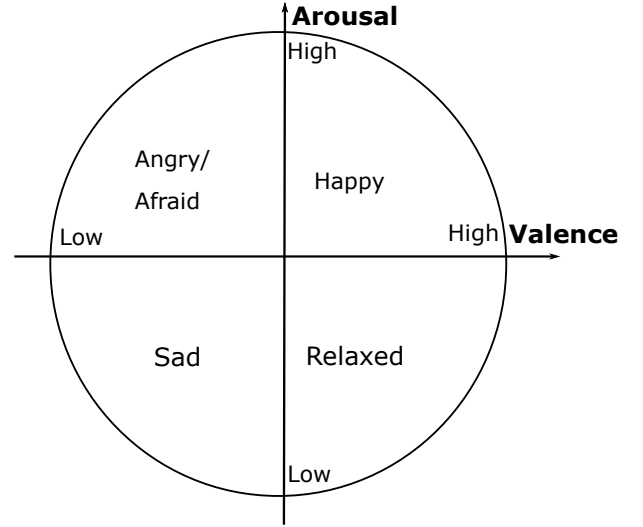


Fig. 1. Valence-Arousal space

of two output labels is considered. A rating value greater than or equal to 5 is set to 1 (aroused; pleasant), otherwise, it is set to 0 (relaxed; unpleasant). The 2D valence-arousal emotion space is shown in Fig. 1.

### B. Data Preprocessing

*1) SAD data-set:* EEG data are complex, as multiple processes take place simultaneously causing various artifacts and noise to appear, such as residual eye movements and EMG artifacts. Hence, it needs to be preprocessed before analysis and cleaned for better interpolation during analysis. EEG data preprocessing stages are implemented using Fieldtrip and Brain Vision Analyzer (Brain Products, Gilching Germany) software. Manual inspection of EEG data for the presence of EMG artifacts or eye movements is performed, and epochs where artifacts appear are removed from the analysis. The frequencies of interest are in the range of 1-64 Hz, covering the five different frequency bands ($\delta,\theta,\alpha,\beta,\gamma$).

*2) DEAP dataset:* The EEG data are downsampled to 128 Hz. EOG and EMG artifacts are removed. As delta waves usually corresponds to deeper sleep [21], useful and informative data for emotion analysis are known to lie at 4-45 Hz frequency. Hence, a bandpass filter was applied leaving the first band, delta [0-4] Hz out of the analysis process. Eye artifacts were removed using a blind source separation technique, and the data were averaged to the common average reference (CAR) where the common average of the entire electrode is subtracted from a specific channel of interest, resulting in a zero-mean voltage distribution [22]. The data are segmented into 60-seconds trials for each video, and a 3 seconds pre-trial baseline is removed.

### C. Data Analysis and Feature extraction

Nonstationary phenomena are present in EEG data due to the constant switching of the meta-stable states of neurons assembling during brain functioning, causing signal changes in the form of spikes and momentary events. In our data

analysis, each channel signal is normalized and divided into N windows in which the data are assumed to be stationary. For each window, a dyadic discrete wavelet decomposition (DWT) is applied to extract the five frequency bands, where the energy and entropy of these bands will be extracted as features. After testing multiple sizes of window segments based on the sampling frequency and the statistical properties, a window size of N=5 seconds was found to yield the best results for detection for the SAD dataset and a window size of N=4 seconds for the DEAP dataset. The analysis of the signal content of each of the EEG frequency bands can be utilized to estimate the subjects' cognition and emotional states.

The energy and entropy of the content of each windowed segment is computed for the k frequency bands (k=5 for SAD, k=4 for DEAP) separately as described later in equations 1 to 4. The analysis is based on the energy and entropy content of these signals represented in two different ways: concatenation of the channels of the k frequency bands and an image-like 2D representation of the EEG channel locations. The latter method is discussed in Section II.E.

### D. Wavelet decomposition

After segmenting the data into multiple windows, the discrete wavelet transform is applied to extract the EEG features, as seen in the DWT tree in Fig. 2. Since EEG signals are non-stationary, Fourier methods are not adequate enough for the time-frequency analysis of such signals. However, wavelet transforms can capture the local behavior of the signal, and can obtain both frequency and time information of transient non-stationary signals. Hence, they are more appropriate and preferable to use for EEG analysis and to decompose the signal into different bands [23].
For the SAD dataset, the data are downsampled to 128 Hz and a dyadic wavelet transformation with 4 levels [24] is used for the signal decomposition into subbands corresponding to the five frequency bands using DWT. [0-4] Hz for $\delta$ band, [4-8] Hz for $\theta$, [8-16] Hz for $\alpha$, [16-32] Hz for $\beta$, and [32-64] Hz for $\gamma$. All five frequency bands are used for the analysis of SAD.
In the case of the DEAP dataset, for each data segment, a 4-level DWT was applied to the input with a filter bank using Daubechies 4 (db4) wavelet. This is illustrated in Table I that contains the decomposition of EEG signals with a sampling frequency of 128 Hz. At the first level, the gamma frequency band was found from the detail coefficients (from the high-pass filter), and the rest are found by repeating the decomposition, by cascading the low and high pass filter over the approximation coefficients (from each lowpass output of the previous level). Only 4 frequency bands, Gamma, Beta, Alpha and Theta are used for the emotion recognition task.

It should be noted that db4 wavelet is chosen due to its orthogonality and smoothing features, which are used for optimal detection of changes in EEG signal [25]. Moreover, according to [26], extraction of EEG signal features using these wavelets could be more efficient as it has near-optimal time-frequency localization properties.

TABLE I
DISCRETE WAVELET DECOMPOSITION

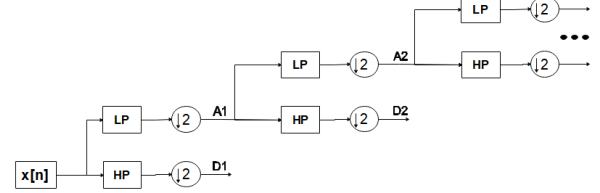| Wavelet Decomposition | | |
| --- | --- | --- |
| Frequency band | Frequency range(Hz) | Decomposition level |
| Gamma | $32 - 64$ Hz | D1 |
| Beta | $16 - 32$ Hz | D2 |
| Alpha | $8 - 16$ Hz | D3 |
| Theta | $4 - 8$ Hz | D4 |
| Delta | $0 - 4$ Hz | A4 |



Fig. 2. Discrete wavelet transform

In both datasets, the energy and entropy content are both extracted as features. The mean wavelet energy $E_j$ of detail coefficients at decomposition level $j$ is defined as

$$\bar{E}_j = \frac{\sum_k |D_{j,k}|^2}{N_j} \quad (1)$$

where $N_j$ is the number of wavelet detail coefficients at level $j$. The total energy is defined as

$$E_{tot} = \sum_{j=1}^{N} \bar{E}_j \quad (2)$$

and the relative wavelet energy is calculated as follows

$$q_j = \frac{\bar{E}_j}{E_{tot}} \quad (3)$$

The wavelet entropy is defined as

$$w_j = -q_j log q_j \quad (4)$$

### E. Image representation of the EEG data

The data are acquired using $M$ electrodes placed over different areas on the scalp, Frontal (F), Central (C) Temporal (T), Parietal (P) and, Occipital (O), where $M$ is 34 for SAD data and 32 for DEAP data. It is believed that knowledge of the location of the channels can provide improved detection accuracy in the analysis of the data. To examine this assumption, two main data models are examined using $M$ channels and $B$ extracted features. These features can be the energy of the frequency bands, or a combination of energy and entropy. $B \in \{5; 10\}$ for SAD, where $B$=5 features when the energy of the 5 frequency bands is used for the analysis, and $B$=10 when both energy and entropy are used. $B \in \{4; 8\}$ for DEAP as only 4 frequency bands were used in the analysis. First, the $M$ channels of the $B$ features are concatenated by creating a $M{\times}B$ feature matrix over each window, without accounting for the location of the channel electrodes ($34{\times}B$
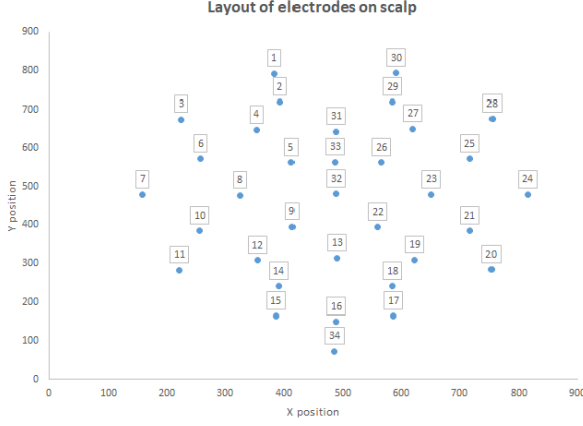
Fig. 3. Layout of 34 electrodes on scalp

for SAD, and $32{\times}B$ for DEAP). Second, a 3-D array of size $K{\times}K{\times}B$ is created, where the first two dimensions represent an image of $K{\times}K$ pixels corresponding with the channels positioning over the scalp while the third dimension represents the number of features B. We chose $K$=15, so that the location is adequately captured without making the image size too large for computational load. In the latter method, the locations that not exactly correspond to any of the $M$ channels are filled using different interpolation techniques.

For both datasets, an image of size $15 \times 15$ was derived to construct an image-like representation of the channels layout. The $M$ channels are mapped to specific pixels in the image based on their locations. A layout of the channels' location for SAD data is shown in Fig. 3. To fill in the missing pixel feature values, the Inverse Distance Weighting (IDW) interpolation method was used [27]: for an interpolated value $e$ at point $x$, the samples $u_i = u(x_i)$ for $i = 1, ...n_i$, which lay within a distance less than $d_{max}$ from point $x$, are only used to calculate interpolated value using the following weighted average.

$$u(x) = \frac{\sum_{i=1}^{n_i} w_i(x)u_i}{\sum_{i=1}^{n_i} w_i(x)} \qquad (5)$$

where $w_i(x) = \frac{1}{d(x,x_i)}$ and $d(x,x_i)$ represents the distance between points $x$ and $x_i$. In areas where no energy values exist within a distance of $d_{max}$ i.e. "Border Points" (BP), the nearest value is simply repeated. $d_{max}$ is chosen empirically.

Other interpolation methods were also considered, such as IDW with zero values at border points (IDW with 0 BP), nearest neighbor interpolation, bilinear interpolation, and cubic b-spline interpolation [28]. However, they were all found to be inferior to the method mentioned above. Table IX in Section IV summarizes the average performance of the different interpolation methods, tested on the main CNN with energy and entropy features on SAD dataset.

## III. EXPERIMENTS

After the data are collected and preprocessed, two models for each dataset are used as previously discussed.

SAD data: In the first model, the energy and entropy for the five frequency bands are calculated separately in each window

of 5 seconds. Hence, for each window a feature matrix of dimensions $34{\times}B$ is constructed. There are 34 channels or electrodes and 2 values of energy and entropy are extracted from each of the 5 frequency bands, yielding $B \in \{5; 10\}$. Each row in the matrix represents one channel. The second model is based on the image representation of the EEG data, as discussed in Section II.E. For each window a 3D energy array of $15{\times}15{\times}B$ is built, where the third dimension represents the energy and entropy values extracted from the five frequency bands.

DEAP data: The models are built in a manner similar to SAD models. However, the dimensions are different. For the first model, the energy and entropy for 4 frequency bands only are calculated separately in each window of 4 seconds. Hence, for each window an energy matrix of dimensions $32{\times}B$ is constructed, as 32 channels were considered, $B \in \{4; 8\}$. In the second model, a 3D energy array of $15{\times}15{\times}B$ is built for each window. For both models in each dataset, each matrix is considered as a single sample for the training or testing data-set.

### A. Acquisition of training and testing EEG data

*1) SAD dataset:* To train the network, a stratified 8-fold cross-validation is applied. A total of 7 folds (56 subjects) are used for training and validation (about $10\%$ validation, $90\%$ training) and the remaining fold (8 subjects) is used for testing. The classifier is trained 8 separate times and in each trial a different fold is used for testing. In both models, for every window of size $N$, a feature matrix is constructed and considered as a single sample, where $N = 5$ seconds. For each subject, multiple samples are gathered by sliding a moving window of size $N$ with no overlap of windows over all channels. The samples collected for the training, validation and testing sets do not overlap, i.e. different samples collected from specific subject cannot be used for training and testing at a certain trial. Samples are labeled 1 if they belong to SAD patients, and 0 otherwise. For every trial, each testing subject is evaluated as follows:

$$prediction = \begin{cases} patient & \text{if } x_i = \frac{p_i}{t_i} \geq Th \\ control & \text{else} \end{cases} \qquad (6)$$

where $Th = 0.45$, $p_i$ is number of samples classified as 1 for subject $i$, $t_i$ is total number of samples for subject $i$. The threshold was lowered to 0.45 due to the fact that EEG data contain unwanted signals.

*2) DEAP dataset:* A very important aspect in emotion recognition task is whether it is subject-dependent or independent. A subject-dependent task means partitioning the training and testing data from same participants. However, different samples taken from the same video cannot be used in both the training and testing data. Subject-independent task means that testing is performed on a group of participants which is different from the training group. According to [29] there exists a physiological linkage with emotion recognition, which makes the recognition depend on age, culture, and gender. Both the structure of the training and testing data, and the physiological linkage make the performance accuracy
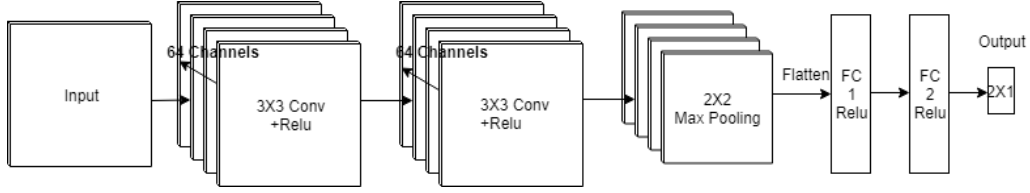
Fig. 4. Convolutional neural network structure

in subject-independent task lower than subject-dependent. A summary of the comparison is provided in Section IV.

A stratified 8-fold cross-validation is applied on the DEAP dataset. Again, 7 folds are used for training and validation, and the remaining fold is used for testing. The average accuracy is found after the classifier is trained 8 times. In the subject-independent task, the 7 folds correspond to 28 participants and the last fold corresponds to the 4 remaining participants. The window size taken to create a sample is $N = 4$ seconds. For every 60 seconds of video, the samples are gathered from all 32 channels by sliding a window of size $N$ with $N/2$ overlap.

The choice of the shift size for each dataset was made empirically as it yielded better results when compared with other window slides. The window shifts that are tested are $N/4$, $N/2$, $N$, and $3N/2$. Early stopping was applied by monitoring the validation loss to avoid overfitting.

### B. Data Augmentation

One of the key challenges in machine learning algorithms in general and deep neural networks specifically, is not having sufficient training data to properly perform a classification task [30]. Training with small datasets might cause the model to be highly biased to the data in the training set, making the model perform poorly on the validation or testing set, as it cannot generalize what it learned to unseen samples. These models suffer from overfitting. Regularization, dropout, batch normalization, and data augmentations are some of the methods used to tackle the problem of overfitting [31]. Image augmentation technique is introduced to help improve the classification performance by creating more robust models with the ability to generalize. Data augmentation refers to artificially generating data by creating new samples to expand the training dataset. It is done by performing transformations on the original images while preserving the label which is invariant to certain variations. In this work, horizontal and vertical shift augmentation is used to expand the training set. This translation is done by moving the image along the X or Y directions, specifically a shift of 1 or both 1 and 2 pixels is done on the image while preserving the image size. For Model 1, images are shifted by a certain number of pixels to the right, left, up and down. The deleted rows or columns are simply replaced by the previous row or column respectively. For Model 2, the shift is done over the first 2 dimensions only. The third dimension which corresponds to the energy or entropy of the frequency bands is filled according to the feature value it has at that specific location. In Section IV, Table X and Table XI summarize the effectiveness of data augmentation for Model 1 and Model 2, respectively.

### C. CNN Network Structure

In recent years, the use of deep learning solutions has become very popular in many applications. Deep network-based algorithms have significantly improved the performance for many problems compared with other machine learning algorithms [32]. CNNs have turned out to be the most powerful deep learning architectures for image-related problems [33], and they have rapidly become a methodology of choice for medical images processing. CNNs are a specific kind of feedforward deep neural networks. Their architecture is characterized by arranging convolutional layers, pooling layers, and fully-connected layers. The input to the network is arrays of data containing energy and entropy values of subband signals which can be viewed as images.

A sequential model is built for SAD classification task as shown in Fig. 4. The first layer in this model is a 2D convolution layer with kernel size 3×3, 64 output filters, and ReLu activation over the outputs. It is then followed by a similar 2D convolution layer, and a max-pooling layer with pool size of 2. A dropout with rate=0.25 is performed. The input is then flattened and fed to a fully connected layer with 128 output dimension and ReLu activation. Another dropout is performed with rate=0.2 followed by a final fully connected layer with Softmax activation and output dimension that equals to 2 corresponding to the two labels in the recognition task. Dropout is a regularization method that is used to reduce overfitting. A similar model configuration is used for emotion recognition task, with a different number of filters, fully connected layers, and parameters such as the dropout rate. Only one max-pooling layer was used in both cases rather than two, as it yields better accuracy. The parameters of the network were determined by comparing accuracies of other models with different structures. Among all trials, the parameters that yielded the best performance were selected. It should be noted that batch normalization was applied to the input from both datasets and the convolutional layers to reduce the internal covariate shift [34].

## IV. RESULTS

In the study, inputs are constructed based on two models. These models are investigated for their classification performance using different classifiers, different types of feature values, and two different datasets. Model 1 is built by concatenating the channels of the frequency bands with no consideration for the spatial configuration of EEG electrodes. Model 2 takes the electrode configuration over the scalp into account.

TABLE II
CONFUSION MATRIX FOR SAD - MODEL 1

| Model 1 | | Actual positive | Actual negative | |
|---|---|---|---|---|
| **Energy features** | Predicted positive | 26 | 6 | Accuracy= 81.25% |
| | Predicted negative | 6 | 26 | Sensitivity=81.25% Specifity=81.25% |
| **Energy and entropy Features** | Predicted positive | 28 | 6 | Accuracy= 84.38% |
| | Predicted negative | 4 | 26 | Sensitivity=87.5% Specifity=81.25% |

TABLE III
CONFUSION MATRIX FOR SAD - MODEL 2

| Model 2 | | Actual positive | Actual negative | |
|---|---|---|---|---|
| **Energy features** | Predicted positive | 29 | 4 | Accuracy= 89.06% |
| | Predicted negative | 3 | 28 | Sensitivity=90.63% Specifity=87.5% |
| **Energy and entropy Features** | Predicted positive | 30 | 3 | Accuracy= 92.19% |
| | Predicted negative | 2 | 29 | Sensitivity=93.75% Specifity=90.63% |

TABLE IV
AVERAGE PERFORMANCE ON SAD DATASET

| Classifier | energy features | | energy & entropy | |
|---|---|---|---|---|
| | Model 1 | Model 2 | Model 1 | Model 2 |
| CNN | 81.25% | 89.06% | 84.38% | 92.19% |
| SVM | 73.43% | 80.5% | 76.56% | 81.25% |
| kNN(k=3) | 70.31% | 76.56% | 71.87% | 78.13% |

TABLE V
LVHV SUBJECT-INDEPENDENT TASK AVERAGE PERFORMANCE

| Classifier | Energy | | Energy & entropy | |
|---|---|---|---|---|
| | Model 1 | Model 2 | Model 1 | Model 2 |
| kNN (k=3) | 73.16% | 78.05% | 75.53% | 81.48% |
| kNN (k=5) | 72.34% | 77.82% | 75.04% | 80.83% |
| CNN | 82.54% | 87.29% | 85.16% | **91.33%** |

The inputs to the classifiers of SAD and DEAP datasets are built using two groups of features: energy of the frequency bands, or a combination of energy and entropy.

### A. SAD dataset

The main results for SAD dataset classification are presented using confusion matrices. In this case, accuracy is defined as the ability to correctly classify a subject, sensitivity is the ability to correctly identify patients when SAD is present, and specificity is the ability to correctly identify control within a healthy group. The analysis is subject-independent. The confusion matrices for Model 1 and Model 2 can be seen in Table II and Table III, respectively. The top part of each table represents inputs built using energy of the frequency bands as features, and the bottom part represents inputs built using both energy and entropy. For each table, the first entry of the third column has the number of actual SAD patients predicted as positive (i.e., patients), the second entry of the third column has the number of actual patients predicted as negative (i.e., healthy subjects), etc.

The accuracy, sensitivity, and specificity are all higher for the proposed approach (Model 2), regardless of the number of features used. Thus, Model 2 that takes advantage of the location of the channels is found to provide superior performance, with a maximum classification accuracy of 92.19% for Model 2 and 84.38% for Model 1. Another important observation is that the number of features influences the accuracy, and it is higher when using both entropy and energy features.

In this classification task the EEG data are also fed to SVM with radial basis kernel function (RBF) with kernel parameter $\sigma = 0.4$ using LIBSVM tool. The parameters were properly chosen by estimating the performance of SVM with different kernels and hyper-parameters tuning using the validation set. Moreover, the data are applied to k-NN with k=3 and k=4, but the proposed network structure stood out as superior in terms of overall classification accuracy, as seen in Table IV for data using energy and entropy features.

It should be noted that Model 2 gave significantly higher accuracy than Model 1, regardless of the classifier type or feature values. Establishing the importance of 2D representation of the spatial configuration of EEG sensors.

### B. DEAP dataset

In this subsection, the binary valence and arousal states, low/high valence (LVHV) and low/high arousal (LAHA), are estimated. The classification is evaluated under two study cases, subject-dependent and subject-independent. The classifiers used for emotion recognition task are the CNN network described in Section III.C, kNN classifier with k=3 neighbors, and kNN with k=5. The average performance of Model 1 and Model 2 is evaluated using energy of decomposed frequency bands as features, or both energy and entropy. Table V and Table VI represent performance of subject-independent task for valence and arousal, respectively. Table VII and Table VIII summarize the accuracies for subject-dependent task on valence and arousal, respectively. It is detected from the mentioned tables that Model 2 outperforms Model 1 in all cases, regardless of feature selection, classifier, or nature of the study case (subject dependent/independent). For instance, in subject-independent valence recognition, the highest average accuracy is $91.33\%$ for Model 2 and only $85.16\%$ for Model 1.

By comparing the average accuracies presented in the tables, it is included that including the entropy in the feature selection scheme has improved the performance of the classifiers.

Another important observation is the noticeable differences in performance between the subject-dependent and subject-independent cases. The manner of construction of the training and testing data in these two cases makes the subject-dependent classifier achieve higher accuracies. However, the lack of generalization of such classifiers is the price one has to pay.

TABLE VI
LAHA SUBJECT-INDEPENDENT AVERAGE PERFORMANCE

| Classifier | Energy | | Energy & entropy | |
|---|---|---|---|---|
| | Model 1 | Model 2 | Model 1 | Model 2 |
| kNN (k=3) | 72.93% | 77.87% | 75.04% | 80.47% |
| kNN (k=5) | 72.63% | 76.71% | 74.08% | 79.64% |
| CNN | 81.76% | 87.33% | 84.18% | **90.87%** |

TABLE VII
LVHV SUBJECT-DEPENDENT AVERAGE PERFORMANCE

| Classifier | Energy | | Energy & entropy | |
|---|---|---|---|---|
| | Model 1 | Model 2 | Model 1 | Model 2 |
| kNN (k=3) | 78.77% | 82.47% | 81.42% | 86.06% |
| kNN (k=5) | 77.65% | 82.51% | 78.85% | 84.17% |
| CNN | 84.24% | 89.6%2 | 87.69% | **94.38%** |

TABLE VIII
LAHA SUBJECT-DEPENDENT AVERAGE PERFORMANCE

| Classifier | Energy | | Energy & entropy | |
|---|---|---|---|---|
| | Model 1 | Model 2 | Model 1 | Model 2 |
| kNN (k=3) | 78.92% | 82% | 80.31% | 84.46% |
| kNN (k=5) | 76.46% | 81.13% | 78.88% | 84.14% |
| CNN | 85.03% | 90.15% | 86.84% | **93.43%** |

TABLE IX
AVERAGE PERFORMANCE OF DIFFERENT INTERPOLATION METHODS

| Interpolation Comparison | |
|---|---|
| Interpolation method | Average recognition perf. |
| Nearest Neighbor | 75% |
| Bilinear | 76.56% |
| IDW with 0 BP | 84.37% |
| Cubic B-spline | 89.06% |
| IDW with NN BP | 92.19% |

The EEG channels' locations are represented as an image in Model 2. There are 34 channels in SAD dataset, and 32 in DEAP. To fill out the empty pixels, a few interpolation schemes are tested, as mentioned in Section II.E. For both models the Inverse Average Weighted interpolation technique is used as it yielded the best results in the classifiers performance. The average accuracies of different interpolation methods tested are presented in Table IX for SAD dataset with energy and entropy features.

It should be noted that for both datasets, a smaller number of channels was also used in the analysis. However using all 34 channels for SAD and 32 for DEAP gave superior results in terms of classification accuracies. Moreover, various image sizes needed to represent the first two dimensions in Model 2 were also tested but 15×15 was found to give sufficient accuracies.

### C. Effect of data augmentation

The performance results listed in Table I to Table IX belong to classifiers with larger training sets, which were expanded using data augmentation. The significance of creating more robust model by increasing the training dataset is reflected in the higher classification accuracies of training with augmented samples achieved in both models. The results are summarized

TABLE X
DATA AUGMENTATION AFFECT MODEL 1

| Classification Accuracy | | |
|---|---|---|
| Model 1 | no augmentation | with augmentation |
| SAD | 81.25% | 84.38% |
| Arousal | 81.55% | 84.18% |
| Valence | 82.71% | 85.16% |

TABLE XI
DATA AUGMENTATION AFFECT MODEL 2

| Classification Accuracy | | |
|---|---|---|
| Model 2 | no augmentation | with augmentation |
| SAD | 89.06% | 92.19% |
| Arousal | 86.47% | 90.87% |
| Valence | 88.34% | 91.33% |

in Table X for Model 1 and Table XI for Model 2. In these tables, the second column represents performance of CNN using original training set, and the third column using the augmented training set.

## V. CONCLUSION

In this paper, a new EEG-based classification approach was proposed. Unlike other approaches, this approach factored in the spatial configuration of the EEG sensors in signal analysis for the first time to our knowledge. A model that takes advantage of the knowledge of the locations of the electrodes was created to construct the EEG dataset. In order to assess the effectiveness of the model, two EEG datsets were used for analysis: SAD for patient/control classification and DEAP for emotion recognition. Overall results showed that the performance of various classifiers based on this model was 5%-8% higher in accuracy, compared with the same classifiers which ignored the configuration. In addition, use of the entropy along with the energy as relevant features in the EEG based classification task, reduced the error rate for the different classifiers (CNN, SVM, kNN) in both datasets. It is also found that data augmentation for the training set is very important to further enhance the performance. This improvement is especially noticeable for Model 2. The effectiveness of the proposed approach is reflected in the results which prove its superiority over other approaches, regardless of the classifier type, features, and most importantly, for EEG data.

## REFERENCES

[1] L. S. Mokatren, R. Ansari, A. E. Cetin, A. D. Leow, O. Ajilore, H. Klumpp, and F. T. Yarman-Vural, "Eeg classification based on image configuration in social anxiety disorder," *The 9th International IEEE EMBS Conference on Neural Engineering*.

[2] S. Sanei and J. A. Chambers, *EEG signal processing*. John Wiley & Sons, 2013.

[3] M. Abo-Zahhad, S. Ahmed, and S. N. Seha, "A new eeg acquisition protocol for biometric identification using eye blinking signals," vol. 07, pp. 48–54, 05 2015.

[4] S. Siuly, Y. Li, and Y. Zhang, "Significance of eeg signals in medical and health research," in *EEG Signal Analysis and Classification*. Springer, 2016, pp. 23–41.

[5] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for eeg-based brain–computer interfaces," *Journal of neural engineering*, vol. 4, no. 2, p. R1, 2007.

[6] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rako-tomamonjy, and F. Yger, "A review of classification algorithms for eeg-based brain–computer interfaces: a 10 year update," *Journal of neural engineering*, vol. 15, no. 3, p. 031005, 2018.

[7] K. Samiee, S. Kiranyaz, M. Gabbouj, and T. Saramäki, "Long-term epileptic eeg classification via 2d mapping and textural features," *Expert Systems with Applications*, vol. 42, no. 20, pp. 7175–7185, 2015.

[8] G. Di Flumeri, G. Borghini, P. Aricò, N. Sciaraffa, P. Lanzi, S. Pozzi, V. Vignali, C. Lantieri, A. Bichicchi, A. Simone *et al.*, "Eeg-based mental workload neurometric to evaluate the impact of different traffic and road conditions in real driving settings," *Frontiers in human neuroscience*, vol. 12, p. 509, 2018.

[9] S. L. Oh, Y. Hagiwara, U. Raghavendra, R. Yuvaraj, N. Arunkumar, M. Murugappan, and U. R. Acharya, "A deep learning approach for parkinson?s disease diagnosis from eeg signals," *Neural Computing and Applications*, pp. 1–7, 2018.

[10] T. A. Richards. (2017) The social anxiety association. [Online]. Available: http://socialphobia.org/

[11] F. Leichsenring and F. Leweke, "Social anxiety disorder," *New England Journal of Medicine*, vol. 376, no. 23, pp. 2255–2264, 2017.

[12] S. G. Hofmann and M. W. Otto, *Cognitive Behavioral Therapy for Social Anxiety Disorder: Evidence-Based and Disorder Specific Treatment Techniques*. Routledge, 2017.

[13] H. C. Kraemer, D. J. Kupfer, D. E. Clarke, W. E. Narrow, and D. A. Regier, "Dsm-5: how reliable is reliable enough?" *American Journal of Psychiatry*, vol. 169, no. 1, pp. 13–15, 2012.

[14] D. A. Moscovitch, D. L. Santesso, V. Miskovic, R. E. McCabe, M. M. Antony, and L. A. Schmidt, "Frontal eeg asymmetry and symptom response to cognitive behavioral therapy in patients with social anxiety disorder," *Biological Psychology*, vol. 87, no. 3, pp. 379–385, 2011.

[15] L. Piho and T. Tjahjadi, "A mutual information based adaptive windowing of informative eeg for emotion recognition," *IEEE Transactions on Affective Computing*, 2018.

[16] H. Chao, H. Zhi, L. Dong, and Y. Liu, "Recognition of emotions using multichannel eeg data and dbn-gc-based ensemble deep learning framework," *Computational intelligence and neuroscience*, vol. 2018, 2018.

[17] J. Xu, F. Ren, and Y. Bao, "Eeg emotion classification based on baseline strategy," in *2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS)*. IEEE, 2019, pp. 43–46.

[18] N. Ganapathy and R. Swaminathan, "Emotion recognition using electrodermal activity signals and multiscale deep convolution neural network." *Studies in health technology and informatics*, vol. 258, pp. 140–140, 2019.

[19] Y. Liu and O. Sourina, "Real-time subject-dependent eeg-based emotion recognition algorithm," in *Transactions on Computational Science XXIII*. Springer, 2014, pp. 199–223.

[20] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "Deap: A database for emotion analysis; using physiological signals," *IEEE transactions on affective computing*, vol. 3, no. 1, pp. 18–31, 2012.

[21] M. Teplan *et al.*, "Fundamentals of eeg measurement," *Measurement science review*, vol. 2, no. 2, pp. 1–11, 2002.

[22] D. J. McFarland, L. M. McCane, S. V. David, and J. R. Wolpaw, "Spatial filter selection for eeg-based communication," *Electroencephalography and clinical Neurophysiology*, vol. 103, no. 3, pp. 386–394, 1997.

[23] H. Adeli, Z. Zhou, and N. Dadmehr, "Analysis of eeg records in an epileptic patient using wavelet transform," *Journal of neuroscience methods*, vol. 123, no. 1, pp. 69–87, 2003.

[24] N. J. Fliege, *Multirate digital signal processing*. John Wiley New York, 1994, vol. 994.

[25] P. Jahankhani, V. Kodogiannis, and K. Revett, "Eeg signal classification using wavelet feature extraction and neural networks," in *IEEE John Vincent Atanasoff 2006 International Symposium on Modern Computing (JVA'06)*. IEEE, 2006, pp. 120–124.

[26] M. Murugappan, N. Ramachandran, and Y. Sazali, "Classification of human emotion from eeg using discrete wavelet transform," *Journal of biomedical science and engineering*, vol. 3, no. 04, p. 390, 2010.

[27] B. A. Eckstein, "Evaluation of spline and weighted average interpolation algorithms," *Computers & Geosciences*, vol. 15, no. 1, pp. 79–94, 1989.

[28] T. M. Lehmann, C. Gonner, and K. Spitzer, "Survey: Interpolation methods in medical image processing," *IEEE transactions on medical imaging*, vol. 18, no. 11, pp. 1049–1075, 1999.

[29] J. A. Soto and R. W. Levenson, "Emotion recognition across cultures: The influence of ethnicity on empathic accuracy and physiological linkage." *Emotion*, vol. 9, no. 6, p. 874, 2009.

[30] J. Lemley, S. Bazrafkan, and P. Corcoran, "Smart augmentation learning an optimal data augmentation strategy," *IEEE Access*, vol. 5, pp. 5858–5869, 2017.

[31] J. Wang and L. Perez, "The effectiveness of data augmentation in image classification using deep learning," *Convolutional Neural Networks Vis. Recognit*, 2017.

[32] L. Deng, D. Yu *et al.*, "Deep learning: methods and applications," *Foundations and Trends® in Signal Processing*, vol. 7, no. 3–4, pp. 197–387, 2014.

[33] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[34] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.