sensors

MDPI

Article

Investigating the Use of Pre-trained Convolutional Neural Network on Cross-Subject and Cross-Dataset EEG Emotion Recognition

Yucel Cimtay [1], *, Erhan Ekmekcioglu [1]

[1] Institute for Digital Technologies, Loughborough University London, UK; y.cimtay@lboro.ac.uk; e.ekmekcioglu@lboro.ac.uk

* Correspondence: yucel.cimtay@gmail.com; y.cimtay@lboro.ac.uk

Received: date; Accepted: date; Published: date

Abstract: EEG has great attraction in emotion recognition studies due to its resistance to deceptive actions of human. This is one of the most significant advantages of brain signals in comparison to visual or speech signals in emotion recognition context. A major challenge in EEG-based emotion recognition is that EEG recordings exhibit varying distributions for different people as well as for the same person at different time instances. This nonstationary nature of EEG limits the accuracy of it when subject independency is the priority. The aim of this study is to increase the subject-independent recognition accuracy by exploiting pre-trained state of the art Convolutional Neural Network (CNN) architectures. Unlike similar studies that extract spectral band power features from the EEG readings, raw EEG data is used in our study after applying windowing, pre-adjustments and normalization. Removing manual feature extraction from the training system overcomes the risk of eliminating hidden features in the raw data and helps leverage the deep neural networks' power in uncovering unknown features. To improve the classification accuracy further, median filter is used to eliminate the false detections along a prediction interval of emotions. This method yields mean cross-subject accuracy of 86.56% and 78.34% on SEED dataset for 2 and 3 emotion classes, respectively. It also yields mean cross-subject accuracy of 72.81% on DEAP dataset and 81.8% on LUMED dataset for 2 emotion classes. Furthermore, the recognition model that has been trained using the SEED dataset was tested with the DEAP dataset, which yields a mean prediction accuracy of 58.1% across all subjects and emotion classes. Results show that in terms of classification accuracy, the proposed approach is superior to, or on par with, the reference subject-independent EEG emotion recognition studies identified in the literature and has limited complexity due to the elimination of the need for feature extraction.

Keywords: EEG; emotion recognition; pretrained models; convolutional neural network; dense layer; subject independency; dataset independency; raw data; filtering on output

1. Introduction

EEG is the measurement of the electrical signals which is a result of brain activities. The voltage difference is measured between the actual electrode and reference electrode. There are several EEG measurement devices in the market such as Neurosky, Emotiv, Neuroelectrics and Biosemi [1] which provide different spatial and temporal resolutions. Spatial resolution is related to number of electrodes and temporal resolution is related to the number of EEG samples processed for unit time. Generally, EEG has high temporal but low spatial resolution. In terms of spatial resolution, EEG electrodes can be placed on the skull according to the 10-20 or 10-10 and 10-5 positioning standards [2].

EEG has lately been used as a powerful emotion prediction modality. It is reliable, portable and relatively inexpensive compared to other brain monitoring tools and technologies. EEG has many application areas. For the clinical applications, EEG is mostly used to investigate the patterns related to sleep [3] and epilepsy [4]. Some other applications of EEG analysis are consciousness and hyperactivity disorders [5,6], measurement of the affective components such as level of attention [7–9], mental workload [10], mood and emotions [11–14] and brain computer interfaces which is the work of transforming brain signals into direct instructions [15–17].

Human emotions have crucial effects on communication with others. Understanding the emotions of human provides controlling and regulating the behaviors. As the digital world name of emotion recognition, affective computing, is the work of emotion recognition by using various sensors and computer-based environments. This concept was originated with Rosalind Picard's paper [18] on affective computing in 1995. In EEG context, affective computing is achieved by setting up brain computer interfaces (BCI) which includes sensors, machines and coding. In BCI, the operation of affective computing starts with presenting users with stimuli which induces specific emotions. These stimuli may be video, image, music, etc. During the session, EEG data is recorded with EEG devices. The next step is typically extracting features from the recorded EEG and training a classifier to predict emotion labels. The final step is testing the trained model with new EEG data which is not used in training session. Data collection and labelling are the most important aspects which has an impact on resulting recognition accuracy. The "Brouwer recommendations" about data collection given in [19] is crucial for handling accurate data and labelling.

In the relevant emotion recognition literature, emotions have been broadly represented in two ways. The first approach classifies emotions as discrete states such as the six basic emotions proposed by Ekman and Friesen [20]. The second approach defines emotion as a continuous 4-D space of valence, arousal, dominance and liking [21, 22]. In most of the studies, this space is reduced to 2-D as valence and arousal dimensions [23,24]. The study conducted in [25] is very useful in the sense that it relates discrete and continuous approaches to each other. Discrete emotion states are mapped on to the valence-arousal circumplex model according to high number of blogposts. This study enables scientists to transform emotions from continuous space to discrete space.

In EEG data channels, typically frequency domain analysis is used. In frequency domain the most important frequency bands are delta (1-3 Hz), theta (4-7 Hz), alpha (8-13 Hz), beta (14-30 Hz) and gamma (31-50 Hz) [26]. Fast Fourier Transform (FFT), Wavelet Transform (WT), eigenvector and Autoregressive are the methods, which transform EEG signal from time domain to frequency domain [27]. The study [28] extracts several frequency-domain features like Differential entropy (DE) and energy spectrum (ES) in order to classify EEG data and [29] investigates the critical frequency bands and channels for EEG based emotion recognition. One of the major problems we observed in the context of emotion classification based on the analysis of EEG channels was that the classifier performance fluctuates remarkably across persons as well as across dataset. Most of these approaches train their classifiers using a set of features derived using the frequency domain analysis. While the classifiers' performance are sufficiently high on test data, which comprise samples belonging to the same subjects but excluded from training and validation, when the same classifier is applied on EEG data of other subjects or on data extracted from various other datasets, the performance is degraded significantly. The same problem in subject-independent analysis is not apparent in the literature in the context of emotion recognition from facial expressions or other physiological data (e.g., heart rate variability, electro-dermal activity). This observation has led us to investigate classification approaches using raw EEG data, which preserves all information and prevents the risk of removing hidden features before training the classifier.

2. Literature Review

For the classification of EEG signals, many machine learning methods such as KNN [28], SVM [28–30], DT (Decision Tree) [31], Random Forest (RF) [32] and LDA (Linear Discriminant Analysis) [33] are applied in this field. In deep learning context, DBN (Deep Belief Network) [34] and AE (Auto encoders) [35] are studied with promising results. Besides DBN and AE, CNN and LSTM structures are widely used [36–40]. Most of these models have shown good results for subject dependent analysis. In [41], KNN method is employed on DEAP dataset [66] for different numbers of channels which show accuracy between 82% and 88%. The study conducted in [42] quadratic time-frequency distribution (QTFD) is employed to handle a high-resolution time-frequency representation of the EEG and the spectral variations over time. It reports mean classification accuracies ranging between 73.8% and 86.2%. In [43] four different emotional states (happy, sad, angry and relaxed) are classified. In that study, DWT is applied on the DEAP dataset. Wavelet features are classified using an SVM classifier with Particle Swarm Optimization (PSO) [70]. The overall accuracy of 80.625% is reported with valence and arousal accuracy of 86.25% and 88.125%, respectively.

An important issue in EEG-based emotion detection is the non-linearity and non-stationarity of EEG signals. Feature sets, such as the spectral band powers of EEG channels, extracted from different peoples against the same emotional states do not exhibit strong correlation. For example, galvanic skin response is a robust indicator of arousal state, where different people's responses correlate with each other well. Training and testing data made of EEG channels' spectral band powers and their derivatives have different distributions. And it is difficult to identify sets of features from the EEG recordings of different subjects, different sessions and different datasets that exhibit more commonality. This makes the classification difficult with traditional classification methods, which assume identical distribution. In order to address this problem and provide subject independency to EEG-based emotion recognition models, deeper networks, domain adaptation and hybrid methods have been applied [44,45]. Furthermore, various feature extraction techniques have been applied, and different feature combinations have been tried [48,50].

Subject independent EEG emotion recognition, as a challenging task, has gained high interest by many researchers lately. The method called Transfer Component Analysis (TCA) conducted in [44] reproduces Kernel Hilbert Space, on the assumption that there exists a feature mapping between source and target domain. A Subspace Alignment Auto-Encoder (SAAE) which uses non-linear transformation and consistency constraint method is used in [45]. This study compares the results with TCA. It achieves a leave one out mean accuracy of 77.88% in comparison with TCA, which shows 73.82% on SEED dataset. Moreover, mean classification accuracy for session-to-session evaluation is 81.81%, an improvement of up to 1.62% compared to the best baseline TCA. In one of the studies, CNN with Deep domain confusion technique is applied on SEED dataset [65] and achieves 90.59% and 82.16 mean accuracy for conventional (subject-dependent) EEG emotion recognition and "leave one out cross validation", respectively [46]. In [47] Variational Mode Decomposition (VMD) is used as a feature extraction technique and Deep Neural Network as the classifier. It gives 61.25% and 62.50% accuracy on DEAP dataset for arousal and valence, respectively. Another study [49] is using a deep convolutional neural network with changing numbers of convolutional layers on raw EEG data which is collected during music listening. It reports maximum 10-fold-validation mean accuracy of 81.54% and 86.87% for arousal and valence, respectively. It also achieves 56.22% of arousal and 68.75% of valence accuracies for one-subject-out test. As can be seen, the reported mean accuracy levels drop considerable in one-subject-out tests due to the nature of the EEG signals.

The study [50] extracts totally 10 different linear and nonlinear features from EEG signals. The linear features are Hjorth activity, Hjorth mobility, Hjorth complexity, the standard deviation, PSD-Alpha, PSD Beta, PSD-Gamma, PSD-Theta and the nonlinear features are sample entropy and wavelet entropy. By using a method called Significance Test/sequential Backward Selection and the Support Vector Machine (ST-SBSSVM) which is a combination of the significance test, sequential backward selection, and support vector machine, it achieves 72% cross subject accuracy for DEAP dataset with High-Low valence classification. It also achieves 89% maximum cross subject accuracy for SEED dataset with positive-negative emotions. Another study [51] uses FAWT (Flexible Analytic Wavelet Transform) which decomposes EEG signals into sub bands. Random forest and SVM are used for classification. The mean classification accuracies are 90.48% for positive/neutral/negative (three classes) in the SEED dataset; 79.95% for high arousal (HA)/low arousal (LA) (two classes); 79.99% for the high valence (HV)/low valence (LV) (two classes); and 71.43% for HVHA/HVLA/LVLA/LVHA (four classes) in the DEAP dataset. In [52] transfer recursive feature elimination (T-RFE) technique is used to determine a set of the most robust EEG features for stable distribution across subjects. This method is validated on DEAP dataset, the classification accuracy and F-score for arousal is 0.7867, 0.7526 and 0.7875, 0.8077 for valence. A regularized graph neural network (RGNN) is applied in [53] for EEG-based emotion recognition, which includes inter-channel relations. The classification accuracy results on SEED dataset are 64.88%, 60.69%, 60.84%, 74.96%, 77.50%, 85.30% for delta, theta, alpha, beta, gamma and all bands. Moreover, it achieves 73.84% of accuracy on SEED IV [71] dataset.

There are several studies which apply transfer learning which aims to explore common stable features and apply to other subjects [54]. In terms of affective computing, the work is exploring some common and stable features which are invariant between subjects. This is also called domain adaptation. In [55] the scientists tried to find typical spatial pattern filters from various recording sessions and have applied these filters on the following ongoing EEG samples. Subject dependent spatial and temporal filters are derived from 45 subjects and a representative subset is chosen in [56]. The study [57] uses compound common spatial patterns which are the sum of covariance matrices. The aim of this technique is to utilize the common information which is shared between different subjects. The other important studies which apply different domain adaptation techniques on SEED dataset are [58–61]. The common properties of these domain adaptation techniques are exploring an invariant feature subspace which reduces the inconsistencies of EEG data between subjects or different sessions. In the study in [62] domain adaptation technique is applied not only in cross-subject context but also for cross datasets. The trained model in SEED dataset is tested against DEAP dataset and vice versa. It reports an accuracy improvement of 7.25%-13.40% with domain adaptation compared to the one without domain adaptation. Scientists applied an adaptive subspace feature matching (ASFM) in [63] in order to integrate both the marginal and conditional distributions within a unified framework. This method achieves 83.51%, 76.68%, 81.20% classification accuracies for the first, second and third sessions of SEED dataset, respectively. This study also conduct testing between sessions. For instance, it trains the model with the data of first session and test on the second session data. In domain adaptation method, the conversion of features into a common subspace may lead to data loss. In order to avoid this, a Deep Domain Confusion (DDC) method based on CNN architecture is used [64]. This study uses adaptive layer and domain confusion loss based on Maximum Mean Discrepancy (MMD) to automatically learn a representation jointly trained to optimize classification and domain invariance. The advantage of this is adaptive classification with retaining the original distribution information.

Having observed that the distribution of commonly derived sets of features from EEG signals show differences between subjects, sessions and datasets, we anticipate that there could be some invariant feature sets that follow common trajectories across subjects, sessions and datasets. There is a lack of studies that investigate these additional feature sets in the EEG signals that can contribute to robust emotion recognition across subjects. The aim of this study is to uncover these kinds of features in order to achieve promising cross-subject EEG-based emotion classification accuracy with manageable processing loads. For this purpose, raw EEG channel recordings after normalization and a state-of-the-art pre-trained CNN model are used. The main motivation behind choosing a pre-trained CNN architecture is due to their superiority in feature extraction and inherent exploitation of domain adaptation. To improve the emotion recognition accuracy, additionally in the test phase, a median filter is used in order to reduce the evident false alarms.

2.1 Contributions of the Work

The contributions of this work to the related literature in EEG-based emotion recognition can be summarized as follows:

- Feature extraction process is completely left to a pretrained state-of-the-art CNN model InceptionResnetV2 whose capability of feature extraction is shown as highly competent in various classification tasks. This enables the model to explore useful and hidden features for classification.

- Data normalization is applied in order to remove the effects of fluctuations in the voltage amplitude and protect the proposed network against probable ill-conditioned situations.
- Extra pooling and dense layers are added to the pretrained CNN model in order to increase its depth, so that the classification capability is enhanced.
- The output of the network is post-filtered in order to remove the false alarms, which may emerge in short intervals of time where the emotions are assumed to remain mostly unchanged.

3. Materials

The EEG datasets used in this work are SEED [65], the EEG data of DEAP [66] and our own EEG dataset [76] which is a part of multimodal emotional database LUMED (Loughborough University Multimodal Emotion Database). All the datasets are open to public access.

3.1 Overview of the SEED Dataset

SEED dataset is a collection of EEG recordings which is prepared by BCMI laboratory of Shanghai Jiao Tong University. 15 clips are chosen for eliciting (neutral, negative and positive) emotions. Each stimuli session is composed of 5 sec. of hint of movie, 4 min. of clip, 45 sec. of self-assessment and 15 sec. of rest. There are 15 Chinese subjects (7 females and 8 males) participated in this study. Each participant had 3 sessions on different days. Totally 45 session of EEG data has been recorded. The labels are given according to the clip contents (-1 for negative, 0 for neutral and 1 for positive). The data was collected via 62 channels which are placed according to 10-20 system, down sampled to 200Hz, a bandpass frequency filter from 0-75Hz was applied and presented as MATLAB "mat" files.

3.2 Overview of the DEAP Dataset

DEAP [22] is a multimodal dataset which includes the electroencephalogram (EEG) and peripheral physiological signals of 32 participants. For 22 of the 32 participants, frontal face video was also recorded. Data was recorded while one-minute long music videos were watched by participants. Totally 40 videos were shown to each participant. The videos were rated by the participants in terms of levels of arousal, valence, like/dislike, dominance and familiarity which changes between 1 and 9. EEG data is collected with 32 electrodes. The data was down sampled to 128Hz, EOG artefacts were removed, a bandpass frequency filter from 4.0-45.0Hz was applied. The data was segmented into 60 second intervals and a 3 second baseline data was removed.

3.3 Overview of the LUMED Dataset

LUMED (Loughborough University Multimodal Emotion Dataset) is a new multimodal dataset that was created in Loughborough University London (UK), by collecting simultaneous multimodal data from 11 participants (4 females and 7 males). The modalities include visual data (face RGB), peripheral physiological signals (galvanic skin response, heartbeat, temperature) and EEG. These data were collected from participants while they were presented with audio-visual stimuli loaded with different emotional content. Each data collection session lasted approximately 16 minutes long that consists of short video clips playing one after the other. The longest clip was approximately 2.5 minutes and the shortest one was 1 minute long. Between each clip, in order to provide the participant a refresh and rest, a 20 second-long gray screen was displayed. Although the emotional ground truth of each clip was estimated based on the content, in reality, a range of different emotions might be triggered for different participants. For this purpose, after each session, the participants were asked to label the clips they watched with the most dominant emotional state they felt. In this current study, we exploited the EEG modality of the LUMED dataset only. For this study, we have re-labelled the samples, such that only 2 classes were defined as negative valence and positive valence. This is done to make a fair comparison with other studies. Moreover, each channel's data was filtered in the frequency range of 0.5Hz to 75Hz to attenuate the high frequency components that are not believed to be have a meaningful correlation with the emotion classes. Normally, captured EEG signals are noisy with EMG (electromyogram) and EOG (electrooculogram) type artefacts. EMG artefacts are electrical noise resulted from facial muscle activities and EOG is electrical noise due to eye movements. For traditional classification and data analysis methods, in order to prevent heavily skewed results, these kinds of artefacts should be removed from the EEG channel data through several filtering stages. As an example, the study in [87] removes the eye movement artefacts from signal by applying ICA (Independent Component Analysis). LUMED dataset has been created initially with the purpose of training a deep-learning based emotion recognition system, described in Section 4. Depending on the type and purpose of other supervised machine learning systems, this dataset could require a more thorough pre-processing for artefact removal. In LUMED, EEG data was captured based on 10-20 system by Neuroelectrics Enobio 8 [82], an 8-channel EEG device with a temporal resolution of 500 Hz. The used channels were FP1, AF4, FZ, T7, C4, T8, P3, OZ, which are spread over frontal, temporal and center lobes of the brain.

4. Proposed Method

In this work, the emotion recognition model works on raw EEG signals without pre-feature extraction. Feature extraction is left to a state-of-the-art CNN model: InceptionResnetV2. The success of this pretrained CNN model on raw data classification was extensively outlined in [67]. Since the distribution of EEG data shows variations from person to person, session to session, and dataset to dataset, it is difficult to identify a feature set that exhibits good accuracy every time. On the other hand, pretrained CNN models are very competent in feature extraction. Therefore, this work gets use of it.

4.1 Windowing of Data

Data is split into fixed length (N) windows with an overlapping size of $N/6$ as shown in Figure 1 (displayed for 3 random channels). One window of EEG data is given in Figure 2 where M is the number of selected channels, C_{ab} is "b^{th} data point of channel a".

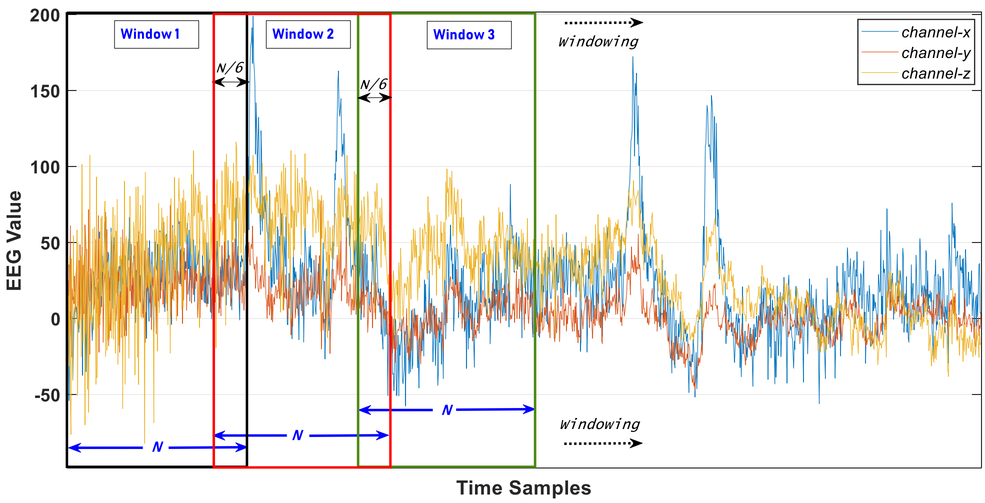

Figure 1 Windowing with overlapping on raw EEG data

4.2 Data Reshaping

EEG data is reshaped to fit the input layer properties of InceptionResnetV2 which is shown in Figure 2. KERAS, which is an open-source neural network library written in Python, is used for the training purpose. Since KERAS is used for training purpose, the minimum input size should be (N_1, N, 3) for InceptionResnetV2 where $N_1 \geq 75$, $N \geq 75$ [68]. Depending on the number of selected channels, each channel data is augmented by creating the noisy copies of it. For instance, if the number of selected channels is S, for each channel the number of noisy copies is calculated according to Equation 1 where $ceil$ operator rounds the number to the next integer (if the number is not integer) and NNC is number of noisy copies.

$$NNC = ceil\left(\frac{N1}{S}\right) - 1, \tag{1}$$

The noisy copies of each channel are created by adding random samples of a gaussian distribution of mean μ and variance σ^2 where μ and σ are chosen as 0 and 0.01, respectively. This process is given in Equation 2 where $[\tilde{C}_{a1}, \tilde{C}_{a2}, \dots, \tilde{C}_{aN}]$ is the noisy copy of the original data $[C_{a1}, C_{a2}, \dots, C_{aN}]$ and $[n_{a1}, n_{a2}, \dots, n_{aN}]$ is the noise vector.

$$[\tilde{C}_{a1}, \tilde{C}_{a2}, \dots, \tilde{C}_{aN}] = [C_{a1}, C_{a2}, \dots, C_{aN}] + [n_{a1}, n_{a2}, \dots, n_{aN}], \tag{2}$$

Since the samples are randomly chosen, each noisy copy is different from each other. N is related to the windowing size. Therefore, a window size N greater than or equal to 75 is chosen. We chose N as 300 in order to provide a standard window size for datasets. This corresponds to 1.5 secs for SEED dataset, approximately 2 secs for DEAP dataset and 0.6 sec for LUMED dataset. Moreover, we chose N_1 as 80 for all datasets. Augmentation process is repeated 3 times in order to make the data to fit KERAS input size. This work does not use interpolation between channels due to that EEG is a nonlinear signal. The reason for adding noise is mainly for data augmentation. There are several ways of data augmentation such as rotation, shifting and adding noise. In the image processing context, rotation, shifting, zooming and adding noise are used. However, we only use noise addition for EEG data augmentation in order to both keep the channel's original data and create new augmented data with limited gaussian noise. This is as if there was another electrode very close the electrode, which is augmented with additional noise. We use data augmentation instead of data duplication in order to make the network adapt to the noisy data and increase the prediction capability of it. This also prevents the network from overfitting due to data repetition. This technique was similarly applied in [83].

4.3 Normalization

Following windowing, augmentation and reshaping, each channel data is normalized by removing mean of each window from each sample. This is repeated for all channels and the noisy copies. The aim of removing the mean is to equate the mean value of each window to 0. This protects the proposed network against probable ill-conditioned situations. In MATLAB, this process is applied automatically on the input data. In KERAS, we performed this manually just before training the network. Each dimension is created separately so they are different from each other.

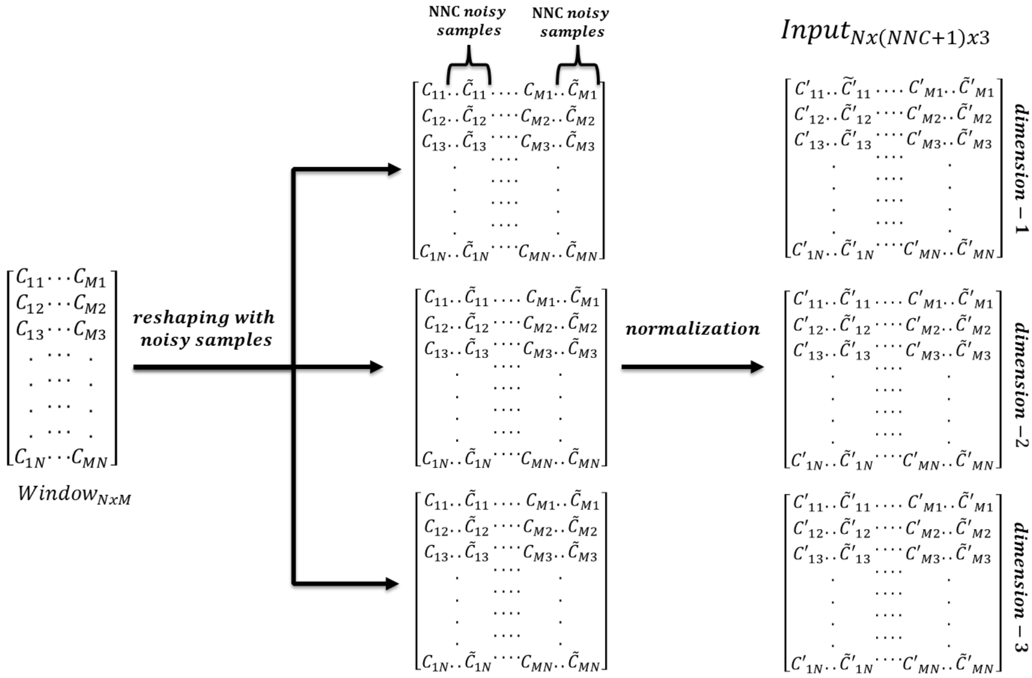

Figure 2 Windowing, Reshaping and Normalization on EEG data

4.4 Channel Selection

In this work, we concentrated on the frontal temporal lobes of the brain. As it is stated in [72,73], emotion changes mostly affect the EEG signals on the frontal and temporal lobes. Different number of channels are tried in this work and increasing the number of channels does not help improve the accuracy. Because, technically, including the channels in the model, which are not correlated with the emotion changes, does not help and on the contrary can adversely affect the accuracy. It is also known that the electrical relations between asymmetrical channels are determining the arousal and valence, hence the emotion [74,75]. Therefore, we have chosen 4 asymmetrical pairs of electrodes: AF1, F3, F4, F7, T7, AF2, F5, F8 and T8 from frontal and temporal lobes which are equally spread on the skull. The arrangement of these channels in the window is AF1, AF2, F3, F4, F5, F6, F7, T7, T8.

4.5 Network Structure

In this work a pretrained CNN network, InceptionResnetV2, is used as base model. Following InceptionResnetV2, Global Average Pooling layer is added for decreasing the data dimension and extra dense layers (fully connected layers) are added in order to increase the depth and success for classifying complex data. The overall network structure is given in Figure 3, and the properties of the layers following the CNN is described in Table 1. The training parameters are specified in Table 2. In Figure 3, Dense Layer-5 determines the number of output classes and $argMax$ selects the one with the maximum probability. We use "relu" activation function to cover the interaction effects and non-linearities. This is very important in our problem while using a deep learning model. Relu is one of the most widely used and successful activation functions in the field of artificial neural networks. Moreover, at the last dense layer we use the "softmax" activation in order to produce the class probabilities.

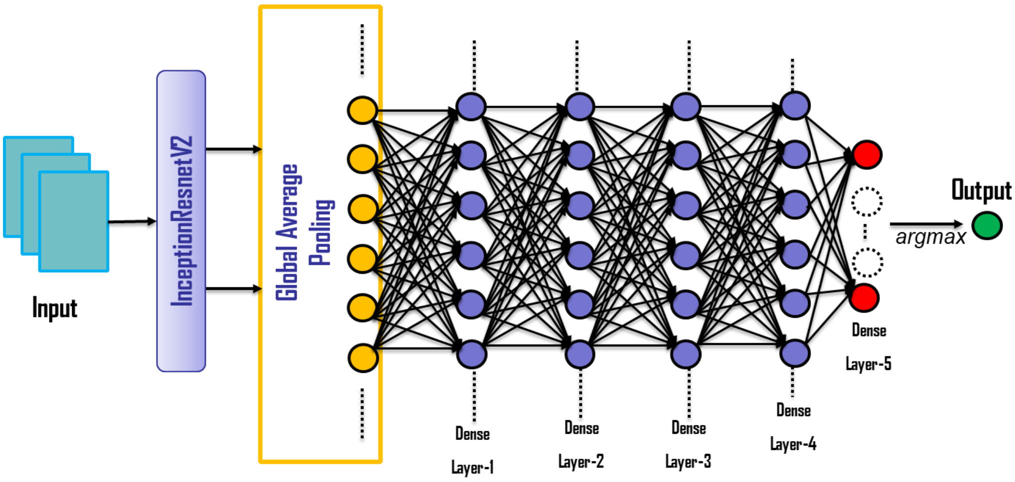

Figure 3 The structure of proposed network model.

Table 1. The properties of Layers following InceptionResnetV2 base model.

Layer (Type)	Output shape	Connected to	Activation Function
Global_Average_Pooling	(None, 1536)	convolution	-
Dense 1	(None, 1024)	Global_Average_Pooling	Relu
Dense 2	(None, 1024)	Dense 1	Relu
Dense 3	(None, 1024)	Dense 2	Relu
Dense 4	(None, 512)	Dense 3	Relu
Dense 5	(None, z [1])	Dense 4	Softmax

[1] z is set according to the number of output classes.

Table 2. The training parameters of network.

Property	Value
Base model	InceptionResnetV2
Additional layers	Global Average Pooling, 5 Dense Layers
Regularization	L2
Optimizer	Adam
Loss	Categorical cross entropy
Max. # Epochs	100
Shuffle	True
Batch size	64
Environment	Win 10, 2 Parallel GPU(s), TensorFlow
# Output classes	2 (Pos-Neg) or 3 (Pos-Neu-Neg)

4.6 Filtering on Output Classes

Since EEG is very prone to noise and different type of artifacts, filtering of EEG signals is widely studied in EEG recognition context. The study conducted in [84] compares three types of smoothing filters (smooth filter, median filter and Savitzky–Golay) on EEG data for the medical diagnostic purposes. The authors concluded that the most useful filter is the classical Savitzky–Golay since it smooths the data without distorting the shape of the waves.

Another EEG data filtering study is provided in [85]. This study employs a moving average filtering on extracted features and then classifies the signal by using SVM (Support Vector Machine). It achieves very promising accuracy results with limited processing time compared to similar studies.

Emotions change quicker than moods for healthy people [69]. However, in very short time intervals (in the range of few seconds), the emotions show lesser variance in healthy individuals with good emotion regulation. Different from the studies [84, 85], the filtering is applied on the output in our method. It is assumed that in a defined small-time interval T the emotion state does not change. Therefore, we apply a median filter on the output data inside a specific time interval with an aim of removing the false alarms and increase the overall emotion classification accuracy. This process is shown in Figure 4 where A and B stands for different classes.

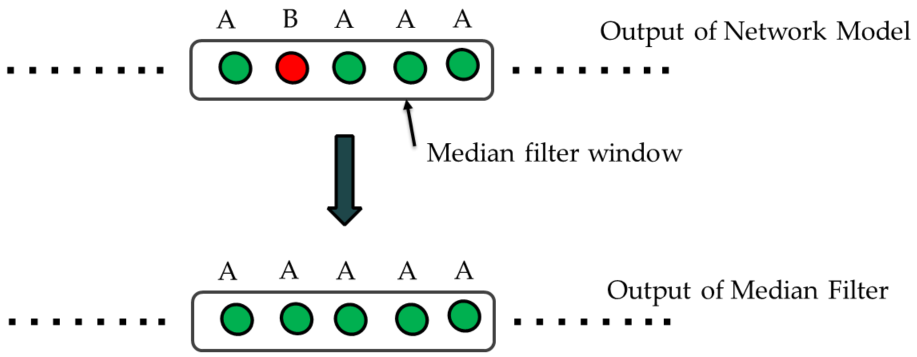

Figure 4 Filtering on Output

The overall process that describes how model training and testing is carried out is visually depicted in Figure 5.

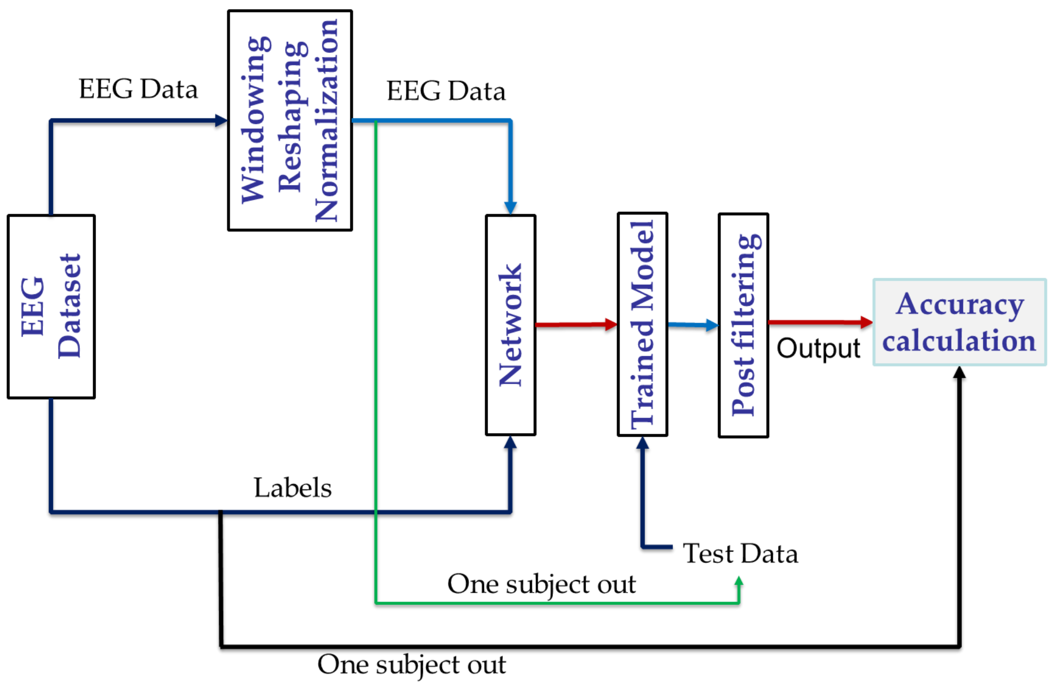

Figure 5 Overall training and testing process of the EEG-based emotion recognition model

5. Results and Discussions

In this work, for SEED dataset, classification tests are conducted for 2 categories of classification: 2-classes: Positive-Negative valence (Pos-Neg) and 3-classes: Positive-Neutral-Negative valence (Pos-

Neu-Neg). SEED dataset provides the labels as negative, neutral and positive. DEAP dataset labels the valence and arousal between 1 and 9. The valence values above 4.5 are taken as positive and the values smaller than 4.5 are taken as negative. For LUMED dataset, classification is done as either positive or negative valence. One-subject-out classification for each dataset are conducted and the results are compared to several reference studies, which provide cross-subject and cross-dataset results. In one-subject-out tests, one subject's data is excluded completely from the training set. The remaining training set is divided into training and validation sets. In this work, during training, when we do not see improvement on validation accuracy for 6 consecutive epochs, we stopped the training and applied the test data on the final model. An example is shown in Figure 6. For each user, Table 3 depicts one-subject-out tests for SEED dataset based on all sessions together, with and without normalization and with and without output filtering. We also got the accuracy results without pooling and dense layers. The mean accuracies dropped by 8.3% and 11.1% without pooling and dense layers, respectively. Applying median filter on the predicted output improves the mean accuracy by approximately 4% for SEED dataset. The filter size is empirically and set to 5. This corresponds approximately to 6 seconds of data. In this time interval it is assumed that the emotion state remains unchanged. It can be seen in Table 3 that the accuracy for some users is high and for some users it is relatively lower. This is based on the modeling of the network with the remaining training data after excluding the test data. However standard deviation is still acceptable. Another issue is that when the number of classes is increased from 2 (Pos-Neg) to 3 (Pos-Neg-Neu), the prediction accuracies drop. This is because some samples labelled as neutral might fall into the negative or the positive classes.

One of the most important characteristics of our work in this paper is that we provide the accuracy scores for each subject separately. This is not observed in most of the other reference studies that tackle EEG-based emotion recognition.

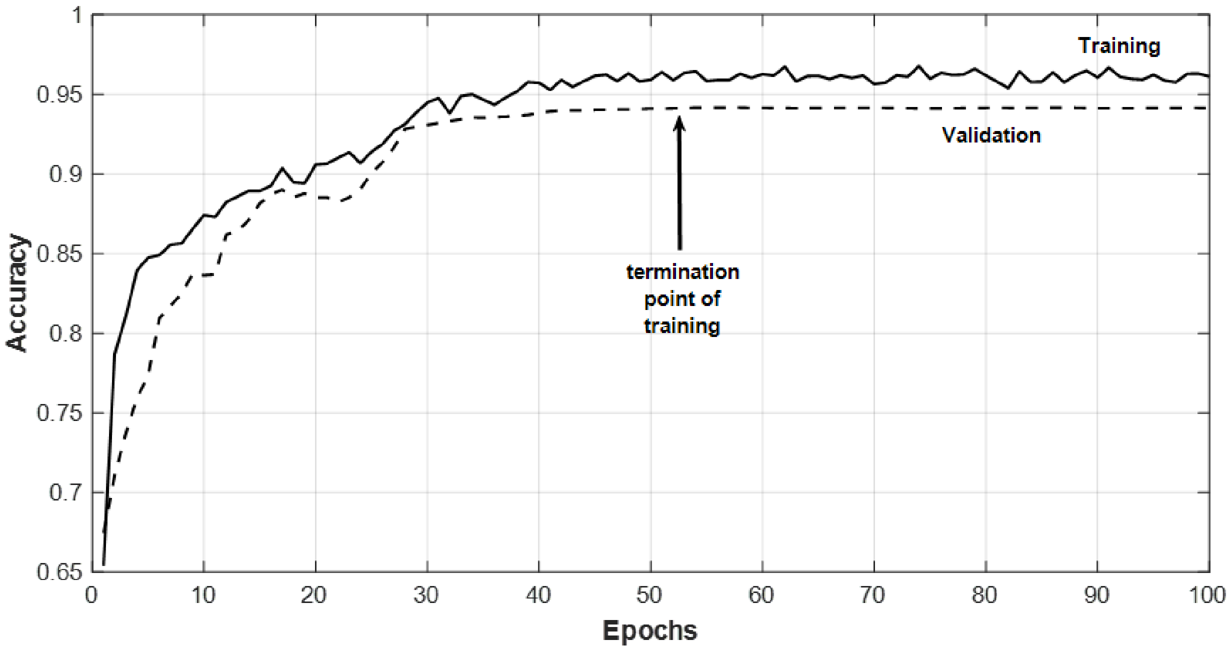

Figure 6 An example of network training

Table 3. "One subject out" classification accuracies for SEED dataset.

Users	Accuracy (Pos-Neg)[1]	Accuracy (without normalization)	Accuracy (with filtering)	Accuracy (Pos-Neu-Neg)[2]	Accuracy (without normalization)	Accuracy (with filtering)

		(Pos-Neg)	(Pos-Neg)		(Pos-Neu-Neg)	(Pos-Neu-Neg)
User 1	85.7	74.2	88.5	73.3	55.2	78.2
User 2	83.6	76.3	86.7	72.8	57.4	78.5
User 3	69.2	56.7	74.3	61.6	53.9	67.7
User 4	95.9	69.4	96.1	83.4	72.4	88.3
User 5	78.4	70.1	83.2	74.1	54.3	76.5
User 6	95.8	81.8	96.4	85.3	67.7	89.1
User 7	72.9	56.2	77.7	64.4	53.5	70.3
User 8	69.2	49.3	75.2	62.9	51.3	69.2
User 9	88.6	61.5	90.5	79.2	64.8	82.7
User 10	77.8	70.1	82.7	69.3	56.0	74.5
User 11	78.6	65.7	83.1	73.0	59.9	78.1
User 12	81.6	72.0	85.7	75.6	63.2	78.4
User 13	91.2	80.2	94.2	81.3	73.1	84.9
User 14	86.4	72.3	91.8	73.9	61.1	78.5
User 15	89.2	73.9	92.3	75.4	60.3	80.3
Average	82.94	68.64	86.56	73.7	60.27	78.34
Std.Dev.	8.32	8.88	6.94	6.80	6.59	6.11

[1] Pos-Neg: Positive-Negative, [2] Pos-Neu-Neg: Positive-Neutral-Negative.

Table 4 shows the cross-subject accuracy comparison of several top studies, which provides the results for 2-classes (Pos-Neg) or 3-classes (Pos-Neu-Neg). For Pos-Neg, the proposed method achieves 86.5% accuracy which is slightly lower than ST-SBSSVM [50]. However, our method has far less complexity, since it does not depend on pre-feature extraction and associated complex calculations. Furthermore, it is not clear in [50] if the reported maximum accuracy of ST-SBSSVM corresponds to the mean prediction accuracy of all subjects, or the maximum prediction accuracy of any subject amongst all.

Another issue is that many reference cross-subject studies use the excluded users' data for validation during the training process. In domain adaptation methods, target domain is also used with source domain to convert data into an intermediate common subspace that makes distributions of target and source domain closer. Similar approach with adding an error function is used in [46]. Using the target domain with source domain can increase the cross-subject [46,63] accuracy because of that the distributions between labelled and unlabeled data is controlled by some cost functions empirically. However, these kinds of approaches are not well-directed. Since, we should know that we only have source domain in cross-subject and/or cross-dataset classification. We aim to generate a model and feature set only from source domain which will be tested with unused target data (either labelled or unlabeled). What should be done is that validation and training data should be clearly split, where excluded subjects' data should not be used in validation. After reaching the furthest epoch where overfitting does not kick in yet, training should be stopped. Then, the final trained model should be tested with the excluded subjects' data. In our study we respected this rule.

Table 4. "One-subject-out" prediction accuracies of reference studies using the SEED dataset.

Work	Accuracy (Pos-Neg)	Accuracy (Pos-Nue-Neg)
ST-SBSSVM [50]	89.0	-
RGNN [53]	-	85.3
Proposed	**86.5**	**78.3**
CNN-DDC [46]	-	82.1
ASFM [63]	-	80.4
SAAE [45]	-	77.8
TCA [44]	-	71.6

GFK [77]	67.5
KPCA [78]	62.4
MIDA [79]	72.4

Table 5 shows the accuracy results of proposed model for DEAP database for 2-classes (Pos-Neg). Generally, the reported accuracies are lower than the ones achieved in the SEED dataset. This may be due to poorer labelling quality of the samples in the DEAP dataset. Some reference studies employ varying re-labelling strategies on the samples of the DEAP dataset to revise class labels. This automatically increases the reported prediction accuracy levels. However, we decided not to alter and respect the original labelling strategy used in that dataset. We only set the threshold in the exact midpoint of the scale of 1 to 9 to divide the samples into two classes, positive and negative. It is acceptable to achieve slightly lower accuracy values than some others as shown in Table 6. To reiterate, post median filtering improves the mean prediction accuracy by approximately 4%.

Table 5. "One-subject-out" prediction accuracies for DEAP dataset using 2-classes (Pos-Neg)

Users	**Accuracy**	**Accuracy (with median filtering)**
User 1	65.1	69.2
User 2	71.2	73.4
User 3	67.8	69.1
User 4	61.7	65.3
User 5	73.1	75.9
User 6	82.5	85.4
User 7	75.5	77.2
User 8	67.6	71.3
User 9	62.8	67.9
User 10	61.9	66.6
User 11	68.8	72.5
User 12	64.3	69.8
User 13	69.1	74.9
User 14	64.3	68.8
User 15	65.6	70.2
User 16	68.7	72.1
User 17	65.6	70.7
User 18	75.8	78.3
User 19	66.9	72.1
User 20	70.4	73.2
User 21	64.5	68.8
User 22	61.6	68.3
User 23	80.7	83.6
User 24	62.5	69.4
User 25	64.9	70.1
User 26	69.7	72.9
User 27	82.7	85.3
User 28	68.9	73.8
User 29	61.7	69.9
User 30	72.9	77.7
User 31	73.1	78.4
User 32	63.6	68.1
Average	68.60	72.81

Std. Dev.	5.85	5.07

Table 6 shows the prediction accuracies of several studies that use the DEAP dataset for 2 classes (Pos-Neg). Our proposed method yields promising accuracy results with only limited complexity (e.g., without any pre-feature extraction cycle) when compared to others. For all eight incoming EEG data channels, the windowing, reshaping, normalization and classification processes take on average 0.34 sec on the test workstation (Core i-9, 3.6 GHz, 64 Gb RAM). This is the computational time used by the python script and KERAS framework. Hence, the data classification can be achieved with roughly a delay of half a second, rendering our usable in real time systems.

Table 6. One-subject-out accuracy comparison of several studies for DEAP dataset (Pos-Neg).

Work	Accuracy
FAWT [51]	79.9
T-RFE [52]	78.7
Proposed	**72.8**
ST-SBSSVM [50]	72
VMD-DNN [47]	62.5
MIDA [79]	48.9
TCA [44]	47.2
SA [80]	38.7
ITL [81]	40.5
GFK [77]	46.5
KPCA [78]	39.8

Table 7 shows the accuracy results of our proposed model on the LUMED dataset for 2-classes (Pos-Neg). It produces a mean prediction accuracy of %81.8 with a standard deviation of 10.9. Post media filtering increases the mean accuracy by approximately 4.5%.

Table 7. One-subject-out prediction accuracies for the LUMED dataset.

	Accuracy	Accuracy (with filtering)
User 1	85.8	87.1
User 2	56.3	62.7
User 3	82.2	86.4
User 4	73.8	78.5
User 5	92.1	95.3
User 6	67.8	74.1
User 7	66.3	71.4
User 8	89.7	93.5
User 9	86.3	89.9
User 10	89.1	93.4
User 11	58.9	67.6
Average	77.11	81.80
Std. Dev.	12.40	10.92

In this work, cross-dataset tests are also conducted between the SEED-DEAP, SEED-LUMED and DEAP-LUMED datasets for positive and negative labels. Table 8 shows the cross-dataset accuracy results between SEED and DEAP. Our model is trained using the data in the SEED dataset and tested on the DEAP dataset separately. It yields 58.10% mean prediction accuracy that is promising in this context. The comparison of the cross-dataset performance of our proposed model with the other cross-dataset studies is given in Table 9. The cross-dataset accuracy of our model is consistently superior to other studies. Table 10 shows the cross-dataset results between SEED-LUMED and DEAP-LUMED. Since LUMED is a new dataset we cannot give any benchmark results with other studies. However, the mean accuracy results and standard deviations are promising.

Table 8. Cross-dataset prediction accuracy results (Trained on SEED and Tested on DEAP)

Users	Accuracy (Pos-Neg)	Accuracy (with median filtering) (Pos-Neg)
User 1	50.5	54.9
User 2	61.7	63.7
User 3	43.3	47.3
User 4	46.0	51.5
User 5	68.9	71.9
User 6	45.3	49.4
User 7	73.4	77.2
User 8	51.9	56.3
User 9	62.3	67.9
User 10	63.8	68.6
User 11	48.6	53.6
User 12	46.4	51.3
User 13	50.1	57.1
User 14	70.4	76.9
User 15	58.8	62.8
User 16	59.7	66.3
User 17	46.6	53.1
User 18	64.7	68.5
User 19	47.9	53.3
User 20	39.1	44.6
User 21	62.1	68.8
User 22	45.6	51.3
User 23	61.4	69.9
User 24	54.0	59.2
User 25	50.8	56.3
User 26	40.8	44.7
User 27	39.2	45.3
User 28	42.4	48.4
User 29	46.2	50.3
User 30	41.7	46.2
User 31	61.4	65.7
User 32	53.8	57.1
Average	**53.08**	**58.10**
Std. Dev.	9.54	9.51

Table 9. One-subject-out cross-dataset prediction accuracy and standard deviation comparison of several studies (Trained on SEED and Tested on DEAP)

Work	Accuracy (Pos-Neg)	Standard Deviation

Proposed	**58.10**	**9.51**
MIDA [79]	47.1	10.60
TCA [44]	42.6	14.69
SA [80]	37.3	7.90
ITL [81]	34.5	13.17
GFK [77]	41.9	11.33
KPCA [78]	35.6	6.97

Table 10. Cross-dataset prediction accuracy results (Trained on SEED/DEAP and Tested on LUMED)

Users (LUMED)	Trained on SEED		Trained on DEAP	
	Accuracy (Pos-Neg)	Accuracy (with median filtering) (Pos-Neg)	Accuracy (Pos-Neg)	Accuracy (with median filtering) (Pos-Neg)
User 1	68.2	72.3	42.7	48.3
User 2	54.5	61.7	50.4	56.8
User 3	54.3	59.6	49.7	54.1
User 4	59.6	64.1	51.3	57.9
User 5	44.8	53.7	87.1	89.7
User 6	67.1	73.5	52.6	58.4
User 7	53.2	60.8	53.8	59.2
User 8	64.5	71.2	46.0	49.3
User 9	48.6	50.9	84.7	85.6
User 10	64.9	76.3	51.5	58.8
User 11	57.1	64.8	63.8	67.1
Average	**57.89**	**64.44**	**57.6**	**62.29**
Std. Dev.	7.33	7.82	14.23	12.91

6. Conclusions

In many recognition and classification problems, the most time and resource consuming part is the feature extraction process. Many scientists focus on extracting meaningful features from the EEG signals either in time and/or frequency domain in order to achieve successful classification results. However, the derived feature sets, which can be useful in the classification problem for one subject, recording session or dataset can fail for different subjects, recording sessions and datasets. Furthermore, since the feature extraction process is a complex and time-consuming process, it is not particularly suitable for online and real time classification problems. In this study, we do not rely on a separate pre-feature extraction process and shift this task to the deep learning cycle that inherently employs this process. Hence, we do not manually remove any potentially useful information from the raw EEG channels. Similar approaches, where deep neural networks are utilized for recognition, were applied in different domains, such as in [86] where electromagnetic sources can be recognised. The success of CNNs has already been shown as highly competent in various classification tasks, especially in the image classification context. Therefore, we deploy of a pretrained CNN architecture called InceptionResnetV2 to classify the EEG data. We have taken the necessary steps to reshape the input data to feed into and train this network.

One of the most important issues, which influences the success of deep learning approaches is the data itself and the quality and reliability of the labels of the data. The "Brouwer recommendations" about data collection given in [19] are very useful for handling accurate data and labelling. Especially during the EEG data recording process, these recommendations should be double checked due to the EEG recording device's sensitivity to noise.

EEG signals are non-stationary and nonlinear. This makes putting forth a general classification model and a set of features based on the well-studied spectral band powers difficult. It is important to be able to identify stable feature sets between subjects, recording sessions and datasets. Since for complex classification problems, CNN is very successful in extracting not-so-obvious features from the input data, we exploit a state of the art pretrained CNN model called InceptionResnetV2 and do not filter out any information from the raw EEG signals. For robustness, we further enrich this deep network by adding fully connected dense layers. This increases the depth and prevents the network from falling into probable ill-conditions and overfitting problems.

In this work, we applied the model successfully on three different EEG datasets: SEED, DEAP and LUMED. Furthermore, we tested our model in a cross-dataset context. We have trained our model with SEED dataset, tested on DEAP and LUMED dataset. Moreover, we have trained our model with DEAP dataset and tested on LUMED dataset. We showed that the results are promising and superior to most of the reference techniques. Once we generate the fully pre-trained model, we can feed any online raw data directly as input to get the output class immediately. Since there is not a dedicated pre-feature extraction process, our model is more suitable to be deployed in real-time applications.

Author Contributions: Conceptualization, Y.C.; methodology, Y.C.; software, Y.C.; validation, Y.C.; formal analysis, Y.C.,E.E.; investigation, Y.C.; resources, Y.C.; data curation, Y.C.; writing—original draft preparation, Y.C.,E.E.; writing—review and editing, Y.C, E.E.; visualization, Y.C.,E.E.; supervision, E.E.; project administration, Y.C., E.E; funding acquisition, E.E.

Funding: This work was supported by an Institutional Links grant, ID 352175665, under the Newton - Katip Celebi partnership between the UK and Turkey. The grant is funded by the UK Department of Business, Energy and Industrial Strategy (BEIS) and The Scientific and Technological Research Council of Turkey (TUBITAK) and delivered by the British Council. For further information, please visit www.newtonfund.ac.uk.

Acknowledgments: We would like to thank the creators of DEAP and SEED datasets for openly sharing them with us and the wider research community. The authors would also like to thank all volunteering staff and students in Loughborough University London for participating in the recording sessions to generate the LUMED dataset.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Top 14 EEG Hardware Companies: Available online: https://imotions.com/blog/top-14-eeg-hardware-companies-ranked/ (accessed on 09/02/2020).
2. Valer J., Daisuke T., Ippeita D., 10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems**,** *Journal of NeuroImage,* **2007,** *Volume 34, Issue 4*, Pages 1600-1611, ISSN 1053 8119, https://doi.org/10.1016/j.neuroimage.2006.09.024.
3. Aboalayon K.A.I., Faezipour M., Almuhammadi W.S., and Moslehpour S., Sleep stage classification using EEG signal analysis: a comprehensive survey and new investigation**,** *MDPI Journal of Entropy*, **2016,** Vol. **18** 272.
4. Acharya U.R., Sree S. V., Swapna G., Martis R. J., and Suri J.S., Automated EEG analysis of epilepsy: a review, *Journal of Knowledge-Based Systems*, **2013**, Vol. 45, Pages.147–65.
5. Engemann D.A. et al, Robust EEG-based cross-site and cross-protocol classification of states of consciousness, *Journal of Brain*, **2018**, Vol. 141, Pages 3179–92.
6. Arns M., Conners C. K., and Kraemer H. C., A decade of EEG theta/beta ratio research in ADHD: a meta-analysis, *Journal of Attention Disorders*, **2013**, Vol. 17, Pages 374–83.
7. Liu N.H., Chiang C.Y., Chu H.C., Recognizing the degree of human attention using EEG signals from mobile sensors, *Journal of Sensors (Basel)*. **2013**, Vol. 13(8), Pages. 10273–10286, doi:10.3390/s130810273
8. Shestyuk A.Y., Kasinathan K., Karapoondinott V., Knight R.T., Gurumoorthy R., Individual EEG measures of attention, memory, and motivation predict population level TV viewership and Twitter engagement. *Journal of PLOS One*, **2019**, Vol. 14(3), doi:10.1371/journal.pone.0214507.
9. Mohammadpour M., and Mozaffari S., Classification of EEG-based attention for brain computer interface, *3rd Iranian Conference on Intelligent Systems and Signal Processing (ICSPIS), Shahrood*, **2017**, Pages 34-37. doi: 10.1109/ICSPIS.2017.8311585.
10. So W.K.Y., Wong S.W.H., Mak J.N., Chan R.H.M., An evaluation of mental workload with frontal EEG. *Journal of PLOS ONE*, **2017**, Vol. 12(4), https://doi.org/10.1371/journal.pone.0174949.

11. Thejaswini S., Ravikumar K.M., Jhenkar L., Aditya N., Abhay K.K., Analysis of EEG Based Emotion Detection of DEAP and SEED-IV Databases using SVM, *International Journal of Recent Technology and Engineering (IJRTE)*, **2019**, Vol. 8, Issue-1C, Pages 207-211, ISSN: 2277-3878.
12. Liu J., Meng H., Nandi A., and Li M., Emotion detection from EEG recordings, *12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, **2016**, Changsha, Pages 1722-1727.
13. Gómez A., Quintero L., López N., Castro J., An approach to emotion recognition in single-channel EEG signals: a mother child interaction, *Journal of Physics: Conference Series*, **2016**, Vol. 705.
14. Xing X., Zhenqi L., Tianyuan X., Shu L., Hu B., and Xiangmin X., SAE+LSTM: A New Framework for Emotion Recognition From Multi-Channel EEG, *Journal of Frontiers in Neurorobotics*, **2019**, Volume 13. https://doi.org/10.3389/fnbot.2019.00037.
15. Müller-Putz G., Peicha L., Ofner P., Movement Decoding from EEG: Target or Direction, *Graz Brain-Computer Interface Conference*, **2017**, doi: 10.3217/978-3-85125-533-1-63.
16. Padfield N., Zabalza J., Zhao H., Masero V., Ren J., EEG-Based Brain-Computer Interfaces Using Motor-Imagery: Techniques and Challenges. *MDPI Journal of Sensors*, **2019**, Vol. 19(6), 1423. doi:10.3390/s19061423.
17. Haufe S., Mondini A.U., Valeria M., Anna L., Cappello A., EEG-Based BCI System Using Adaptive Features Extraction and Classification Procedures, *Journal of Computational Intelligence and Neuroscience*, **2016**, https://doi.org/10.1155/2016/4562601.
18. Picard R.W., Affective Computing, MIT Media Laboratory, *Perceptual Computing;* **1995**, 20 Ames St., Cambridge, MA 02139.
19. Alarcão S.M., and Fonseca M. J., Emotions Recognition Using EEG Signals: A Survey, *IEEE Transactions on Affective Computing*, **2019**, Vol. 10, no. 3, Pages 374-393, doi: 10.1109/TAFFC.2017.2714671.
20. Ekman P., Friesen W.V., O'Sullivan M., Chan A.I., Diacoyanni T., Heider K., Krause R., LeCompte W.A., Pitcairn T., and Ricci-Bitti P. E., Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology*, **1987**, vol. 53, no. 4, Pages 712–717, doi: 10.1037/0022-3514.53.4.712.
21. Scherer K.R., What are emotions? And how can they be measured, *SAGE Journals: Social Science Information*, **2005**, Vol. 44, no. 4, Pages 695-729, https://doi.org/10.1177/0539018405058216.
22. Koelstra S., Mueh C., Soleymani M., Lee J.-S., Yazdani A., Ebrahimi T., Pun T., Nijholt A., Patras I., DEAP: A Database for Emotion Analysis using Physiological Signals, *IEEE Transactions on Affective Computing,* **2011**, Vol. 3, no. 1, Pages 18-31.
23. Russell J. A., A circumplex model of affect, *Journal of Personality and Social Psychology*, **1980**, Vol. 39, no. 6, Pages. 1161–1178**.**
24. Guzel Aydin S., Kaya T., Guler H., Wavelet-based study of valence-arousal model of emotions on EEG signals with LabVIEW, *Journal of Brain Informatics*, **2016**, Vol. 3(2), Pages 109–117, doi:10.1007/s40708-016-0031-9.
25. Paltoglou G., and Thelwall M., Seeing Stars of Valence and Arousal in Blog Posts, *IEEE Transactions on Affective Computing*, **2013,** Vol. 4, no. 1, Pages 116-123, doi: 10.1109/T-AFFC.2012.36.
26. The McGill Physiology Virtual Lab: Available online: https://www.medicine.mcgill.ca/physio/vlab/biomed_signals/eeg_n.htm (09/02/2020).
27. Al-Fahoum A.S., Al-Fraihat A., Methods of EEG Signal Features Extraction Using Linear Analysis in Frequency and Time-Frequency Domains, *ISRN neuroscience*, **2014**, doi:10.1155/2014/730218.
28. Ruo-Nan D., Jia-Yi Z., and Bao-Liang L., Differential Entropy Feature for EEG-based Emotion Classification, *6th International IEEE/EMBS Conference on Neural Engineering (NER)*, **2013**, San Diego, CA, Pages 81-84.
29. Wei-Long Z., and Bao-Liang L., Investigating Critical Frequency Bands and Channels for EEG-based Emotion Recognition with Deep Neural Networks, **2015,** *IEEE Transactions on Autonomous Mental Development (IEEE TAMD)*, Vol. 7(3), Pages 162-175.
30. George F.P., Shaikat I.M., Hossain P.S.F., Parvez M.Z., Uddin J., Recognition of emotional states using EEG signals based on time-frequency analysis and SVM classifier, *International Journal of Electrical and Computer Engineering* (IJECE), **2019,** Vol. 9, Pages 1012-1020.
31. Soundarya S., An EEG based emotion recognition and classification using machine learning techniques, *International Journal of Emerging Technology and Innovative Engineering*, **2019,** Volume 5, Issue 4, Pages 744-750.

32. Swati V., Preeti S., Chamandeep K., Classification of Human Emotions using Multiwavelet Transform based Features and Random Forest Technique, *Indian Journal of Science and Technology*, **2015**, Vol 8.
33. Bono V., Biswas D., Das S., and Maharatna K., Classifying human emotional states using wireless EEG based ERP and functional connectivity measures**,** *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, **2016,** Las Vegas, NV, 2016, pp. 200-203. doi: 10.1109/BHI.2016.7455869.
34. Nattapong T., Ken-ichi F., Masayuki N., Application of Deep Belief Networks in EEG, based Dynamic Music-emotion Recognition, *The International Joint Conference on Neural Networks* (IJCNN), **2016**, Vancouver, Canada.
35. Prieto L., and Kominkova O., Emotion Recognition using AutoEncoders and Convolutional Neural Networks. *MENDEL*. **2018**, vol. 24, no. 1, Pages 113-120, doi: 10.13164/mendel.2018.1.113
36. Chen J. X., Zhang P. W., Mao Z. J., Huang Y. F., Jiang D. M., and Zhang Y. N., Accurate EEG-Based Emotion Recognition on Combined Features Using Deep Convolutional Neural Networks, **2019**, *IEEE Access*, Vol. 7, Pages 44317-44328.
37. Tripathi S., Acharya S., Sharma R. D., Mittal S., and Bhattacharya S., Using deep and convolutional neural networks for accurate emotion classification on DEAP Dataset, *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence* , **2017**, Pages 4746-4752.
38. Salama E. S., El-Khoribi R. A., Shoman M. E., and Shalaby M. A., EEG-based emotion recognition using 3D convolutional neural networks, *International Journal of Advanced Computer Science and Applications(IJACSA)*, **2018**, vol. 9, no. 8, Pages 329-337.
39. Alhagry S., Aly A., El-Khoribi R., Emotion Recognition based on EEG using LSTM Recurrent Neural Network, *International Journal of Advanced Computer Science and Applications*, **2017**, Vol. 8, Pages 355-358.
40. Jeevan R. K., Shiva Kumar P., Srivikas M., EEG-based emotion recognition using LSTM-RNN machine learning algorithm, *1st International Conference on Innovations in Information and Communication Technology (ICIICT),* **2019,** Pages 1-4.
41. Li M., Xu H., Liu X., Lu S. Emotion recognition from multichannel EEG signals using K-nearest neighbor classification*. Journal of Technology and Health Care*, **2018**, Vol. 26, Pages 509–519, doi:10.3233/THC-174836.
42. Alazrai R., Homoud R., Alwanni H., Daoud M.I., EEG-Based Emotion Recognition Using Quadratic Time-Frequency Distribution. *MDPI Sensors***, 2018**, Vol. 18, No. 8, 2739, doi:10.3390/s18082739
43. Nivedha R., Brinda M., Vasanth D., Anvitha M., and Suma K. V., EEG based emotion recognition using SVM and PSO, *International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT),* **2017**, Kannur, Pages 1597-1600.
44. Pan S.J., et. al., Domain adaptation via transfer component analysis. *IEEE Transaction on Neural Network*, **2011**, Vol. 22, No. 2, Pages 199–210.
45. Chai X., et. al.: Unsupervised domain adaptation techniques based on auto-encoder for non-stationary EEG-based emotion recognition. *Journal of Computers in Biology and Medicine*, **2016**, Vol. 79, Pages 205–214.
46. Weiwei Z., Cross-Subject EEG-Based Emotion Recognition with Deep Domain Confusion, *Intelligent Robotics and Applications. ICIRA 2019. Lecture Notes in Computer Science*, **2019**, Vol. 11740.
47. Pallavi P., Seeja K.R., Subject independent emotion recognition from EEG using VMD and deep learning, *Journal of King Saud University - Computer and Information Sciences*, **2019**, Pages 53-58
48. Fu Y., Xingcong Z., Wenge J., Pengfei G., Guangyuan L., Multi-method Fusion of Cross-Subject Emotion Recognition Based on High-Dimensional EEG Features, *Frontiers in Computational Neuroscience*, **2019**, Volume 13.
49. Keelawat P., Thammasan N., Kijsirikul B., Numao M., Subject-Independent Emotion Recognition During Music Listening Based on EEG Using Deep Convolutional Neural Networks, *IEEE 15th International Colloquium on Signal Processing & Its Applications (CSPA)*, **2019,** Penang, Malaysia, Pages 21-26. 10.1109/CSPA.2019.8696054.
50. Yang F., Zhao X., Jiang W., Gao P., Liu G., Multi-method Fusion of Cross-Subject Emotion Recognition Based on High-Dimensional EEG Features, *Frontiers in Computational Neuroscience,* **2019**, doi:10.3389/fncom.2019.00053.
51. Gupta V., Chopda M. D., Pachori R. B., Cross-Subject Emotion Recognition Using Flexible Analytic Wavelet Transform From EEG Signals, in *IEEE Sensors Journal*, **2019**, Vol. 19, no. 6, Pages 2266-2274.
52. Zhong Y., Yongxiong W., Li L., Wei Z., Jianhua Z., Cross-Subject EEG Feature Selection for Emotion Recognition Using Transfer Recursive Feature Elimination, *Frontiers in Neurorobotics*, **2017,** Vol. 11.

53. Zhong P., Wang D., Miao C., EEG-Based Emotion Recognition Using Regularized Graph Neural Networks, arXiv.org, in Press, **2019.**
54. Pan S. J. and Yang Q., A survey on transfer learning, *IEEE Transaction on Data Eng*ineering, **2010**, Vol. 22, no. 10, Pages 1345–1359.
55. Krauledat M., Tangermann M., Blankertz B., and Müller K. R., Towards zero training for brain-computer interfacing, *Journal of PLOS ONE*, **2008,** Vol. 3, no. 8.
56. Fazli S., et. al., Subject-independent mental state classification in single trials, *Journal of Neural Networks*, **2009**, Vol. 22, no. 9, Pages. 1305–1312.
57. Kang H., Nam Y., and Choi S., Composite common spatial pattern for subject-to-subject transfer, *IEEE Signal Processing Letters*, **2009**, Vol. 16, no. 8, Pages 683–686.
58. Zheng W.L., Zhang Y.Q., Zhu J.Y., and Lu B.L., Transfer components between subjects for EEG-based emotion recognition, *2015 International Conference on Affective Computing and Intelligent Interaction (ACII),* **2015,** Xi'an, China, Pages 917–922.
59. Zheng W.-L., and Lu B.-L., Personalizing EEG-based affective models with transfer learning, *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16)*, **2016**, Pages 2732–2738.
60. Chai X., et al., Unsupervised domain adaptation techniques based on auto-encoder for non-stationary EEG-based emotion recognition, Journal of *Computers in Biology and Medicine*, **2016**, Vol. 79, Pages 205–214.
61. Chai X., et al., A fast, efficient domain adaptation technique for cross-domain electroencephalography (EEG)-based emotion recognition, *MDPI Sensors*, **2017**, Vol. 17, no. 5, Page 1014.
62. Zirui L., Sourina O., Wang L., Scherer R., Müller-Putz G., Domain Adaptation Techniques for EEG-Based Emotion Recognition: A Comparative Study on Two Public Datasets, *IEEE Transactions on Cognitive and Developmental Systems*, **2018**, Vol.11, Pages 85-94, 10.1109/TCDS.2018.2826840.
63. Chai X., et. al., Efficient Domain Adaptation Technique for Cross-Domain Electroencephalography (EEG)-Based Emotion Recognition, *MDPI Sensors***, 2017**, Vol.17, Page 1014.
64. Tzeng E., et. al., Deep domain confusion: maximizing for domain invariance, *arXiv preprint* arXiv:1412.3474, **2014.**
65. SEED Dataset. Available online: http://bcmi.sjtu.edu.cn/~seed/ (accessed on 09/02/2020).
66. DEAP Dataset. Available online: https://www.eecs.qmul.ac.uk/mmv/datasets/deap/ (accessed on 09/02/2020).
67. Pretrained Deep Neural Networks, Available online: https://uk.mathworks.com/help/deeplearning/ug/pretrained-convolutional-neural-networks.html (accessed on 09/02/2020).
68. Keras Applications. Available online: https://keras.io/applications/#inceptionresnetv2 (accessed on 09/02/2020).
69. Lane A., Beedie C., Terry P., Distinctions between Emotion and Mood. *Journal of Cognition and Emotion*. **2005**, Vol. 19, doi: 10.1080/02699930541000057.
70. Particle Swarm Optimization, Available online: https://www.sciencedirect.com/topics/engineering/particle-swarm-optimization (accessed on 09/02/2020)
71. SEED Dataset. Available online: http://bcmi.sjtu.edu.cn/~seed/seed-iv.html (accessed on 09/02/2020).
72. Salzman C. D., Fusi S., Emotion, cognition, and mental state representation in amygdala and prefrontal cortex, *Journal of Annual Review of Neuroscience*, **2010,** Vol. 33, Pages 173-202, doi: 10.1146/annurev.neuro.051508.135256.
73. Zhao G., Zhang Y., Ge Y., Frontal EEG Asymmetry and Middle Line Power Difference in Discrete Emotions, *Frontiers in Behavioral Neuroscience*, **2018,** Vol.12, Page 225.
74. Bos P. -O. D., EEG-based Emotion Recognition. The Influence of Visual and Auditory Stimuli, **2006.**
75. Al-Nafjan A., Al-Wabil A., Hosny M., Al-Ohali Y., Classification of Human Emotions from Electroencephalogram (EEG) Signal using Deep Neural Network, *(IJACSA) International Journal of Advanced Computer Science and Applications*, **2017**, Vol. 8, No. 9.
76. Loughborough University EEG based Emotion Recognition Dataset. Available online: https://www.dropbox.com/s/xlh2orv6mgweehq/LUMED_EEG.zip?dl=0
77. Gong B., Shi Y., Sha F., and Grauman K., Geodesic flow kernel for unsupervised domain adaptation *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*(CVPR), **2012,** Pages 2066-2073.

78. Schölkopf B., Smola B., and Müller K. R., Nonlinear component analysis as a kernel eigenvalue problem, *Neural computation*, **1998,** Vol. 10, no. 5, Pages 1299-1319**.**
79. Yan K., Kou L., and Zhang D., Learning Domain-Invariant Subspace Using Domain Features and Independence Maximization, *IEEE Transactions on Cybernetics*, **2018**, vol. 48, no. 1, Pages. 288-299.
80. Fernando B**.,** Habrard A., Sebban M., and Tuytelaars T., Unsupervised visual domain adaptation using subspace alignment, *Proceedings of the IEEE International Conference on Computer Vision*, **2013**, Pages 2960-2967.
81. Shi Y., and Sha F., Information-theoretical learning of discriminative clusters for unsupervised domain adaptation, *Proceedings of the International Conference on Machine Learning (ICML),* **2012**, Pages 1275–1282.
82. Enobio 8. Available online: https://www.neuroelectrics.com/solutions/enobio/8/ (accessed on 09/02/2020).
83. Wang F., Zhong S., Peng J., Jiang J., Liu Y., Data Augmentation for EEG-Based Emotion Recognition with Deep Convolutional Neural Networks, *In: Schoeffmann K. et al. (eds) MultiMedia Modeling. MMM. Lecture Notes in Computer Science,* **2018**, Vol. 10705. Springer, Cham.
84. Kawala-Sterniuk A., et. al., Comparison of Smoothing Filters in Analysis of EEG Data for the Medical Diagnostics Purposes, **2020**, *MDPI Sensors,* Vol. 20, No. 3, Page 807, doi: 10.3390/s20030807
85. Tang C., et. al., EEG-based emotion recognition via fast and robust feature smoothing, **2017**, *Proceedings of International Conference on Brain Informatics*, Pages 83-92.
86. Matuszevski J., and Pietrow D., Recognition of electromagnetic sources with the use of deep neural networks, *Proc. SPIE 11055, XII Conference on Reconnaissance and Electronic Warfare Systems*, **2019**, https://doi.org/10.1117/12.2524536.
87. Plöchl M., Ossandón J.P., König P., Combining EEG and eye tracking: identification, characterization, and correction of eye movement artifacts in electroencephalographic data. *Front Hum Neuroscience*, **2012**, Vol. 6:278. doi:10.3389/fnhum.2012.00278.