Engineering Science and Technology, an International Journal 24 (2021) 1442–1454

HOSTED BY

Contents lists available at ScienceDirect

Engineering Science and Technology, an International Journal

journal homepage: www.elsevier.com/locate/jestch

Full Length Article

Emotion recognition based on EEG feature maps through deep learning network

Ante Topic *, Mladen Russo

Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, University of Split, 21000 Split, Croatia

ARTICLE INFO

Article history:
Received 16 July 2020
Revised 28 February 2021
Accepted 24 March 2021
Available online 16 April 2021

Keywords:
Brain-computer interface
Electroencephalogram
Emotion recognition
Valence-arousal model
Deep learning
Computer-generated holography

ABSTRACT

Emotion recognition using electroencephalogram (EEG) signals is getting more and more attention in recent years. Since the EEG signals are noisy, non-linear and have non-stationary properties, it is a challenging task to develop an intelligent framework that can provide high accuracy for emotion recognition. In this paper, we propose a new model for emotion recognition that will be based on the creation of feature maps based on the topographic (TOPO-FM) and holographic (HOLO-FM) representation of EEG signal characteristics. Deep learning has been utilized as a feature extractor method on feature maps, and afterward extracted features are fused together for the classification process to recognize different kinds of emotions. The experiments are conducted on the four publicly available emotion datasets: DEAP, SEED, DREAMER, and AMIGOS. We demonstrated the effectiveness of our approaches in comparison with studies where authors used EEG signals that classify human emotions in the two-dimensional space. Experimental results show that the proposed methods can improve the emotion recognition rate on the different size datasets.

1. Introduction

Emotions are human reactions to events, and they have an influence on the whole body. They are a part of our daily life and an essential part of non-verbal communication, so it is not surprising that they have been studied by psychologists for decades. Each object we are in contact with, is the stimulus we react to with emotion. If the stimulus is positive, the emotion will consequently be positive. Regardless of the type of emotion, they can manifest through various modes of human emotional expression such as psycho-physiology, face expression, gesture, or biological reactions.

A lot of effort from researchers has been put into the invention of an intelligent emotion recognition system. Some of them are based on a group of non-physiological signals, such as speech [1], facial expression [2], and posture [3], but since these methods are quite subjective and depend on a person's age and culture, it is hard to be certain of the authenticity of emotion. Physiological signals, such as heart rate [4], skin impedance, respiration [5] or brain signals, as well as functional magnetic resonance imaging (fMRI) [6], magnetoencephalography (MEG) [7], and EEG, can give more reliable results. The key challenge in this research area is to detect human emotions based only on the physiological and biological reactions that can be captured and measured by various body sensors. An example is a Brain-Computer Interface (BCI) [8,9] which is an EEG sensor that we will focus on. It measures physiological signals that arise in the human brain, which is the central nervous system.

The area of emotion recognition based on EEG signal features is very challenging since the EEG signal is non-linear, non-stationary, and contains a significant amount of noise. Furthermore, the characteristics of the EEG signal are mainly extracted from the time, frequency, or time–frequency domain, and more recently from the spatial domain. Researchers who are constructing models for recognizing emotions based on the spatial characteristics of the signal have largely focused on the asymmetry between electrode pairs. That is, the methods in most cases refer to the differences of the signals measured by the corresponding electrode on the left and right hemispheres of the brain, respectively. Therefore, the question arises whether the spatial characteristics of the signal, which contain important information related to the recognition of emotions, can be presented in other ways.

The motivation for this work is precisely the problems that researchers face nowadays. More specifically, the goal is to create new feature maps from the EEG signal and construct an advanced model to recognize the emotional states of the user that will

* Corresponding author.
E-mail address: atopic@fesb.hr (A. Topic).
Peer review under responsibility of Karabuk University.

https://doi.org/10.1016/j.jestch.2021.03.012
2215-0986/

produce the result with a high level of accuracy. Therefore, we propose topographic (TOPO-FM) and holographic (HOLO-FM) feature maps which are created from the various characteristics of the EEG signals. Thereafter, the deep learning technique has been used as a feature extractor method on feature maps. Extracted features have been fused, and to get the classification of user emotions, a machine learning technique has been utilized. Furthermore, the idea is to verify the model on signals obtained by the experimental method and recorded by measuring devices that have different spatial resolutions with respect to the number of electrodes. For this purpose, we conducted our experiments on the four publicly available emotion datasets: DEAP, SEED, DREAMER, and AMIGOS. By using the proposed approach, we have achieved results that outperform studies that have used comparable methods.

The rest of this paper is structured as follows. In Section II the related work is described. Section III gives a description of used datasets and selected features. Further, the procedure of creating the feature maps, feature extraction, and, finally, classification methods are explained in detail in Section IV. Discussion about results is in Section V, and the conclusion is presented in Section VI.

2. Related work

With the development of BCI technology in recent years, researchers are increasingly paying attention to emotion recognition from EEG signals, and they commonly use a discrete or dimensional model for emotion classification. A discrete model with six basic emotions: anger, disgust, fear, happiness, sadness, and surprise, was proposed by Ekman [10], while Plutchik developed the wheel of eight basic emotions [11]. Each of these emotions can be expressed stronger or weaker and therefore, the two-dimensional model [12] and three-dimensional model [13], which represent emotions in Valence-Arousal, and Pleasure-Arousal-Dominance space are proposed by Russel and Mehrabian respectively.

Emotion cognition utilizes single or multimodal labeled data as input, and further deep learning to continually learn unlabeled data. Emotion detection is a part of emotion cognition and it uses single or multimodal emotional labeled data in the classification process with machine or deep learning algorithms [14]. EEG-based emotion recognition studies are typically consisting of the following steps: emotion elicitation, collecting and preprocessing signals, feature extraction, and finally classification in order to recognize specific emotion. Features can be extracted from the time domain, frequency domain, time–frequency domain, or spatial domain, and it is not unusual that authors combine features from more domains [15,16]. Similar, numerous classification algorithms have been presented so far, such as support vector machine (SVM), linear discriminant analysis (LDA), Bayesian classifier, k-nearest neighbor (k-NN) etc., but since there are more and more publicly available datasets with a large amount of brain signals experimentally acquired, researchers started to use deep learning techniques [17]. Convolutional neural networks (CNN) and deep belief networks (DBN), followed by Multi-layer perceptron neural network (MLPNN), some kind of hybrid networks, and recurrent neural networks (RNN) are the most used deep learning techniques for emotion recognition tasks [18]. CNN has demonstrated powerful capabilities for feature extraction and classification tasks. For example, in a method based on coincidence filtering approach [19], automatic extraction and classification of features [20], or emotion recognition using a multi-column CNN model [21].

In order to improve model accuracy, many studies combined deep learning techniques to extract features and supervised learning techniques for classification in various areas such as group activity recognition [22], classification of human activities based on micro-Doppler radar [23], etc., but are used in the area of emotion recognition as well. These hybrid methods most commonly use CNN for the feature extraction [24,25], but also PCANet deep learning model [26] and Bimodal Deep Auto Encoder (BDAE) [27]. Long short-term memory (LSTM) [28] and SVM [24,25,26,27] are ordinary used for classification tasks in these kind of models. The highest accuracy is reached in [24] where authors used CNN and SVM combination to recognize emotions using facial expression images for socially assistive robots.

Furthermore, it is not unusual that attentive features are used for EEG emotion recognition as the authors did in [29] where the channel-wise attention mechanism is accepted to adaptively assign the weights of different channels, and this framework is integrated into a CNN to explore spatial information of encoded EEG signals. An attention-based convolutional recurrent neural network (ACRNN) integrates extended self-attention mechanism into RNN to explore temporal information of EEG signals.

Topographic images can provide results that are comparable with other more commonly used features in the field of emotion recognition. Since DEAP is the most cited among other datasets used in this work, it is reasonable to expect that a lot of research papers are dealing with images as features exist. For instance, the method of mapping electrodes into a two-dimensional matrix [30,31] for constructing multidimensional or multi-band images proved to be good for integrating spatial characteristics, frequency domain, and temporal characteristics of the EEG signals. Feature images can be constructed using continuous wavelet transformation (CWT) for creating grid-like frames [32] or generating the EEG spectrogram with a Short-time Fourier transform (STFT) [33]. It is interesting that both methods provide an almost equal accuracy rate for valence and arousal in emotion recognition tasks. Generating power spectrum heat maps by calculating PSD on theta, alpha, and beta sub-bands is another idea for creating two-dimensional images [34]. This method tries to exploit EEG-topography to extract information from different brain regions. Finally, feature maps generated as feature distribution on the scalp under different frequency bands (theta, alpha, beta, and gamma) were constructed in the motion-related brain networks with phase-locking value (PLV) with the aim to capture the non-linear phase synchronization to statistically measure the phase synchronization between two brain areas [35].

Authors of the SEED dataset [27,36] have created feature maps from different kinds of features extracted from five frequency bands and they got the best result with differential entropy feature. Another approach is constructing spatial maps which were created from the average differential entropy calculated for each channel separately [37] or spatial maps created from Granger causality density which is the measure of a whole brain dynamical complexity [38]. It is worth noting that electrode-frequency distribution maps (EFDMs) [39] and topology-preserving differential entropy (TP-DE) [40] are two methods for creating images that will be fed into a deep learning network, can achieve the highest accuracy on the SEED dataset. Any of the aforementioned types of created feature topographic maps do not necessarily have to be constructed only from EEG signals. Fusion with other physiological signals is a common method. For instance, the spectrogram from combined EEG, Electrocardiography (ECG) and Galvanic Skin Response (GSR) signals was adopted on the AMIGOS dataset [41].

Holography is invented in 1948 by Gabor [42], which is a method of capturing and reproducing a three-dimensional image of an object using light interference and diffraction, with the help of coherent light without the use of optical lenses. We differentiate two waves in holography, an object and a reference wave. The first one is the wave that is scattered by the object and is diverging toward the recording medium, while the latter one directly illuminates the recording medium. The amplitude of the object wave,

which represents brightness or intensity, and the phase, which represents the shape of the object, contains complete spatial information about the object. It is difficult for the observer to distinguish whether it is the original object or a virtual image if the hologram is properly illuminated.

Computer-generated holography (CGH) is a technology suitable for generating holograms from a synthetic three-dimensional graphic model [43]. A digital hologram is commonly printed, but in our study, we displayed it as a two-dimensional image using [44]. During the reconstruction of the digital hologram, all object information will be provided. CGH is used in medical researchers such as [45] where authors utilized holographic microscopy and deep convolutional neural network to classify holographic images. Another example is [46], where authors used deep transfer learning (DTL) based on the VGG-19 pre-trained network to classify raw hologram constructed with a lens-free digital in-line holography (LDIH). A review of the available literature did not find any scientific research in which holography is used in the emotion recognition process. CGH records all spatial information about the object, so the idea of this paper is to map a three-dimensional object into the plane, and to show it as a two-dimensional image which will be fed into a model for recognizing the emotional states of users.

3. Experimental data

3.1. Selected datasets

This research uses the emotion EEG signals from four publicly available datasets to evaluate our method of emotion recognition. The DEAP [47], SEED [48], DREAMER [49], and AMIGOS [50] datasets are briefly introduced below, while a general comparison is given in Table 1. The raw EEG data for each

dataset was collected from every region of the brain by 14, 32, or 62 electrodes, depending on the measuring device. The EEG electrodes are placed according to the 10–20 international system [51] which shows the relationship between the underlying area of the cerebral cortex and the electrode location. The system defines the distance between adjacent electrodes as 10% and 20% of the total distance between the left and right, respectively the front and back of the head. Only EEG signals using the two-dimensional emotional space, depicted in Fig. 1, are taken as data. The two dimensions are arousal, which ranges from relaxed to aroused, and valence, which ranges from pleasant to unpleasant. Rating scale ranges were from 1 to 9 for DEAP and AMIGOS datasets, while for DREAMER dataset was from 1 to 5, so we labeled trials into two classes (low or high) using the threshold of 4.5, and 2.5 respectively. Selected datasets originally do not have the same sample rate, and therefore in order to be comparable, we used the raw signals from DREAMER and AMIGOS, and preprocessed data from DEAP dataset which have a sampling rate of 128 Hz. Moreover, we resampled the EEG signals to 128 Hz after the data has been read from the SEED dataset.

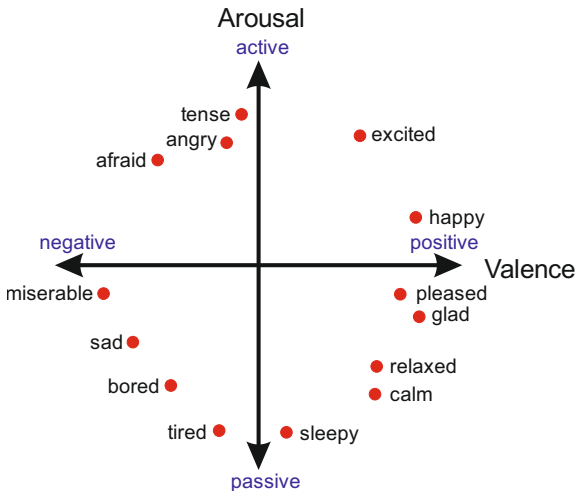

Fig. 1. Valence-arousal space.

The DEAP (Dataset for Emotion Analysis using Physiological signals) is a multimodal dataset that contains EEG, electromyography (EMG), electro-oculogram (EOG), and blood volume pulse. They were gathered based on skin temperature, plethysmograph, and galvanic skin responses, which were collected from 32 participants (16 males and 16 females) aged between 19 and 37 (mean age 26.9). Each participant was stimulated through 40 one-minute-long music video clips. After each trial/video, a self-reported emotional state was evaluated for arousal (passive/active), valence (negative/positive), liking (like/dislike) and dominance. The EEG signals were down sampled to 128 Hz and a blind source separation technique was used for removal of the EOG artefacts. Three seconds pre-trial baseline was removed, as well. Furthermore, a 4.0–45 Hz bandpass frequency filter and common average reference were also applied.

SJTU emotion EEG Dataset (SEED) contains EEG signals and eye movement signals from 15 participants (7 males and 8 females, and aged 23.27 ± 2.37 (mean ± std)). Electroencephalography signals were collected via 62 electrodes while the participants were watching fifteen Chinese film clips with three types of emotions, i.e., negative, positive, and neutral. Each class of emotion has five corresponding film clips. Three experiments were performed for each participant with an interval of about one week. The EEG data, which lasts approximately four minutes per trial, were down sampled to 200 Hz and a band-pass filter from 0 to 75 Hz was applied. In our work, we resampled signals to 128 Hz in order to be the same as in other datasets utilized in this work. Furthermore, we used only positive and negative trials from participants to be comparable with other datasets that utilize binary classification. The positive and negative affective state can be mapped as high and low valence affective state, respectively as the authors did in the studies [52,53,54,55]

Table 1
Comparison between datasets.

Dataset	Participants	Trials/ videos	Channels	Experiments	Affective states	Rating scale ranges (threshold)	Measuring device	Sampling rate [Hz]	Year of publication
DEAP	32	40	32	1	Valence/Arousal	1 – 9 (4.5)	BioSemi ActiveTwo	512	2012
SEED	15	10	62	3	Positive/Negative	N/A	ESI Neuro Scan System	1000	2015
DREAMER	23	18	14	1	Valence/Arousal	1 – 5 (2.5)	The Emotiv EPOC	128	2018
AMIGOS	40	16	14	2	Valence/Arousal	1 – 9 (4.5)	The Emotiv EPOC	128	2018

In our study, we wanted to have the input data which has recorded with the portable low-cost off-the-shelf measuring device, so we chose DREAMER and AMIGOS datasets. DREAMER is a multimodal dataset that consists of EEG and ECG signals which were acquired by 23 participants (14 males and 9 females) aged between 22 and 33 years. EEG data obtained with Emotiv EPOC wireless headset with a sampling rate of 128 Hz which has 14 electrodes while participants were watching 18 film clips that were used for eliciting nine different emotions: amusement, excitement, happiness, calmness, anger, disgust, fear, sadness and surprise. After watching emotional films, which last from 65 to 393 s, participants assessed self-emotion with the self-assessment manikins (SAM) [56] to acquire subjective assessments of valence (ranged from unhappy/sad to happy/ joyful), arousal (ranged from calm/-bored to stimulated/excited), and dominance (submissive/without control to dominance/empowered).

A dataset for Mood, personality and affect research on Individuals and GrOupS (AMIGOS) contains EEG (14 channels), ECG (2 channels), and galvanic skin response (1 channel) signals, as well as frontal video (RGB) from two experiments. In the first experiment, 40 participants watched 16 short video clips (less than 250 s), while 37 participants from the first one participated in the second experiment. They watched four long video clips (more than 14 min) while they were divided into four groups of five participants, and 17 people watched the videos individually. Participants performed self-assessment in valence, arousal, dominance, familiarity, liking, and seven basic emotions (neutral, disgust, happiness, surprise, anger, fear, and sadness), after they watched each video. For a fair comparison with other datasets from our study, we utilized only data from the first experiment. As reported in [57], seven participants (IDs: 9, 12, 21, 22, 23, 24, 29 and 33) have had missing data in their trials, therefore we did not use them in our study. Since we defined the threshold of 4.5, some participants did not have both low and high affective states, so we excluded their data as well. Participants 5 and 28 have had only high arousal while participants 11 and 30 have had only low arousal value. Besides that, the participant with ID 11 has had only high valence.

3.2. Selected features

The features describe the signal and it is therefore very important to choose the ones that will give us the best information about the signal. Commonly, features from recorded EEG signals are extracted from time domain (e.g. statistical, fractal dimension, Hjorth parameters, High Order Crossing), frequency domain (e.g. Power Spectral Density), time–frequency domain (e.g. Discrete Wavelet Transform, entropy) or spatial domain (e.g. Differential Asymmetry, Rational Asymmetry, Magnitude Squared Coherence Estimate, Differential Causality). Although researchers are using a great number of different complex features, we chose only nine features that describe the signal power, energy, complexity, irregularity, and standard deviation. Chosen features of which the fractal dimension, Hjorth activity, mobility and complexity, peak-to-peak, and the root-mean-square belong to the time domain, while the band power, the differential entropy, and the power spectral density are the features from the frequency domain. These features produce state-of-the-art accuracy in models where they have been used, and are described below.

The band power (BP) is the conventional feature that represents the sum of the squares of its time domain samples divided by the signal length [15,58,59]. Differential entropy (DE) is a measure of signal complexity that is related to minimum description length as well [60,61,62]. The differential entropy formula for Gaussian distribution can be formulated as (1)

$$h(X) = -\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} exp\frac{(x-\mu)^2}{2\sigma^2} \log\frac{1}{\sqrt{2\pi\sigma^2}} exp\frac{(x-\mu)^2}{2\sigma^2} dx = \frac{1}{2}\log 2\pi e\sigma^2 \quad (1)$$

where X submits the Gaussian distribution $N(\mu, \sigma^2)$, x is a variable, and π and e are constants.

Since the EEG signal is non-linear, there are several methods for calculating Fractal dimension (FD) which is the measure of the signal complexity and irregularity. That means the signal is more complex if the value of the fractal dimension is higher. Some of them are Minkowski bouligand or box-counting dimension [63], Fractal Brownian Motion [64] and Higuchi algorithm [65]. We utilize the latter mentioned one, since it is shown that it outperforms other methods [63]. In order to compute the fractal dimension with the Higuchi algorithm, let *X (1), X (2), …, X (N)* be a finite set of time series samples. Then, the newly constructed time series is defined as shows in (2) and (3):

$$X_k^m : X(m), X(m+k), \cdots, \left(m + \left[\frac{N-m}{k}\right] \cdot k\right) \quad (2)$$

$$(m = 1, 2, \cdots, k) \quad (3)$$

where m is the initial time and k is the interval time.
k sets of $L_m(k)$ are calculated by (4):

$$L_m(k) = \frac{\left\{\left(\sum_{i=1}^{\left[\frac{N-m}{k}\right]} |X(m+ik) - X(m+(i-1)\cdot k)|\right) \frac{N-1}{\left[\frac{N-m}{k}\right]\cdot k}\right\}}{k} \quad (4)$$

where $\langle L(k)\rangle$ denotes the average value of $L_m(k)$, and a relationship exists as follows in (5):

$$\langle L(k)\rangle \propto k^{-D} \quad (5)$$

Moreover, the fractal dimension can be obtained by logarithmic plotting between different k and its associated $\langle L(k)\rangle$. Hjorth parameters [66] for a signal x of length N are the commonly used features in signal processing [16,67]. Hjorth Activity (HA) represents the squared standard deviation of the amplitude (mean power of the signal), which is shown in (6):

$$Activity(x) = \frac{\sum_{n=1}^{N}(x(n) - \bar{x})^2}{N} \quad (6)$$

where $\bar{x}$ stands for the mean of x. Hjorth Mobility (HM) measures a standard deviation of the slope with reference to the standard deviation of the amplitude (mean frequency of the signal), and can be calculated by (7):

$$Mobility(x) = \sqrt{\frac{var(x')}{var(x)}} \quad (7)$$

where x' denotes the derivate of signal x. The third parameter, Hjorth Complexity (HC), is the number of standard slopes, i.e. measure of the deviation of the signal from the sine shape (8):

$$Complexity(x) = \frac{Mobility(x')}{Mobility(x)} \quad (8)$$

Peak-to-Peak (PP) is the difference between the maximum and minimum values in signal x. The amplitude of the signal is frequently evaluated using the peak-to-peak technique [15,68]. Moreover, the feature is frequently used in emotion recognition from audio signal [69]. Power spectral density (PSD) [36,70,71,72] is the traditional and well-known method for calculating average energy from different frequency bands. We utilized Welch's

method and took its mean value. The Root-Mean-Square (RMS) is the ordinary method that is used for measuring the amplitude of a bio-signal [73,74,75].

4. Methodology

The locations of EEG electrodes for a particular EEG band, i.e. the EEG-topography, are not taken into account in a case when the commonly used EEG features are utilized. Further, more meaningful features that perform better than traditional ones, can be extracted with deep-learning-based methods [34]. The goal of this study is to use the advantages of feature maps that can be used to represent spatial and spectral information for every of the five used EEG sub-bands (delta, theta, alpha, beta, and gamma). For this purpose, we created new feature maps from the EEG signal using nine features and constructed an advanced model with reliability to recognize the emotional states of users which are comparable with the state-of-the-art level of accuracy. Our newly constructed hybrid model uses a convolutional neural network in the process of deep learning of signal characteristics and supervised learning techniques to classify the emotional state of the user.

4.1. Feature maps creation

This research proposes two methods, called TOPO-FM and HOLO-FM, for creating feature maps that will be input into a deep learning network. Common to both methods is that for each participant's trial, every channel is decomposed into five sub-bands, delta (0–4 Hz), theta (4–8 Hz), alpha (8–16 Hz), beta (16–32 Hz), and gamma (32–64 Hz), by discrete wavelets transform (DWT) using 'db5' mother wavelet. The signal characteristics such as the band power (BP), differential entropy (DE), fractal dimension (FD), Hjorth activity (HA), complexity (HC), and mobility (HM) parameters, peak-to-peak (PP), power spectral density (PSD), and root mean square (RMS) are calculated on each sub-band. The value of each individual signal characteristic is mapped to a standard international 10–20 system describing the electrode position on the head which is shown in Fig. 3. By displaying the value of the signal characteristic at the electrode location, the position of the point in the three-dimensional space is defined.

The investigation of signal characteristics in three-dimensional space is conducted in two directions: TOPO-FM and HOLO-FM. In the first direction of research, the values of the signal characteristics calculated for a specific sub-band are displayed by a topographic map where the areas on the map that are not filled are interpolated. The second approach that we called HOLO-FM, utilizes computer-generated holography, which is used to create two-dimensional feature maps from the spatial characteristics of the signal, i.e. from points in space.

A comparison of datasets is shown in Table 1, and it is visible that all of them differ in a number of participants, trials, channels (except DREAMER and AMIGOS), and experiments. It follows that the number of feature maps per participant will vary depending on the dataset. Therefore, DEAP has 1800 feature maps per participant which is 57600 in total. For SEED is generated 1350 per participant and 20250 for all three experiments together. DREAMER and AMIGOS are datasets that have a similar total count of feature maps to the SEED. DREAMER and AMIGOS use 18630 and 20880 in total, and 810 and 720 per participant, respectively.

4.2. TOPO-FM

Values of each characteristic of the signal are positioned at the electrode locations that are defined by the 10–20 system. As shown in Fig. 2, electrode positions are mapped on the feature matrix depending on the dataset. Matrix with nine rows and nine columns has manifested as a good choice since it is possible to adequately map all electrodes [30,31]. On the other hand, it is suitable for all datasets that have less than 81 electrodes, and in our case that are all four datasets. As depicted in Fig. 3 using each characteristic of the EEG signal separately, we created a two-dimensional feature map for each of five sub-bands. Thereafter empty spaces on the matrix are interpolated, and "jet" colormap is applied to form a fully topographic map of a specific feature. These TOPO-FM images are input into a deep learning model for feature extraction.

In Fig. 4, the examples are showing the Hjorth activity feature in the alpha band for the first participant, and it is classified as a positive value on SEED as well as high valence on all other datasets. With respect to the "jet" colormap, red color represents active electrodes, while dark blue color denotes that electrode is not active at all. According to the 10–20 electrode mapping on the 9x9 matrix, that is presented in Fig. 2, coordinates for each electrode are constant, and for example, frontal F7 has (1, 7) while temporal T7 has (1, 5) coordinates. It is not fully comparable, but is visible in Fig. 4 that the area between F7 and T7 electrodes has been highly active during these trials.

4.3. HOLO-FM

Due to the fact that the hologram is able to record complete spatial information of the illuminated object in three-dimensional space, we applied it for creating holographic feature maps. For this purpose, we adopted the algorithm of an in-line hologram for computer-generated holography [44]. Firstly, the characteristics for every EEG channel sub band have been calculated and they were shown in the spatial area. The next step was to map a three-dimensional object into the plane, and to show it as a two-dimensional image.

A convenient way to describe the creation of a holographic feature map is given in Fig. 5. The source of illumination is a coherent light of wavelength λ = 532 nm which is a simulation of a green diode-pumped solid-state (DPSS) laser. Three-dimensional object scene is rotated in a way that illumination comes along the Z-axis keeping the same logic as for topographic feature maps. For each point that is normalized, we calculated a classical hologram, i.e., the interference pattern of a spherical wave (the object light), and a plane wave (the reference light). When the object is illuminated the wave is scattered by the object and its light diverges towards the hologram plane, while the reference wave directly illuminates the hologram plane. The direct light propagation method is used in the hologram plane, where these waves interfere together and create a hologram of the given object. Obviously, the object represents the feature value in the spatial area. For the purpose of creating full HOLO-FM, we repeated the aforementioned process for all channels, which is 32 for the DEAP, 62 for the SEED, and 14 for the DREAMER and AMIGOS datasets.

Each dataset in our study has a minimum of 14 channels that produce a lot of interference, resulting in a final image that contains a lot of noise. In an effort to mitigate this problem, we smoothed interference with exponential function in order to reduce the number of rips and to get clearer information about the value of the feature in the channel. Thereafter, all images created from individual object points are summed together which corresponds to a multi-exposure hologram. Finally, "jet" colormap is applied to the HOLO-FM images.

In Fig. 6, a HOLO-FM example for the first participant of all datasets is presented. High valence or positive value class is shown for Hjorth activity on the alpha band. Similar to the example for TOPO-FM, the same area in the frontal and temporal lobe is highly active when participants are stimulated with music videos.

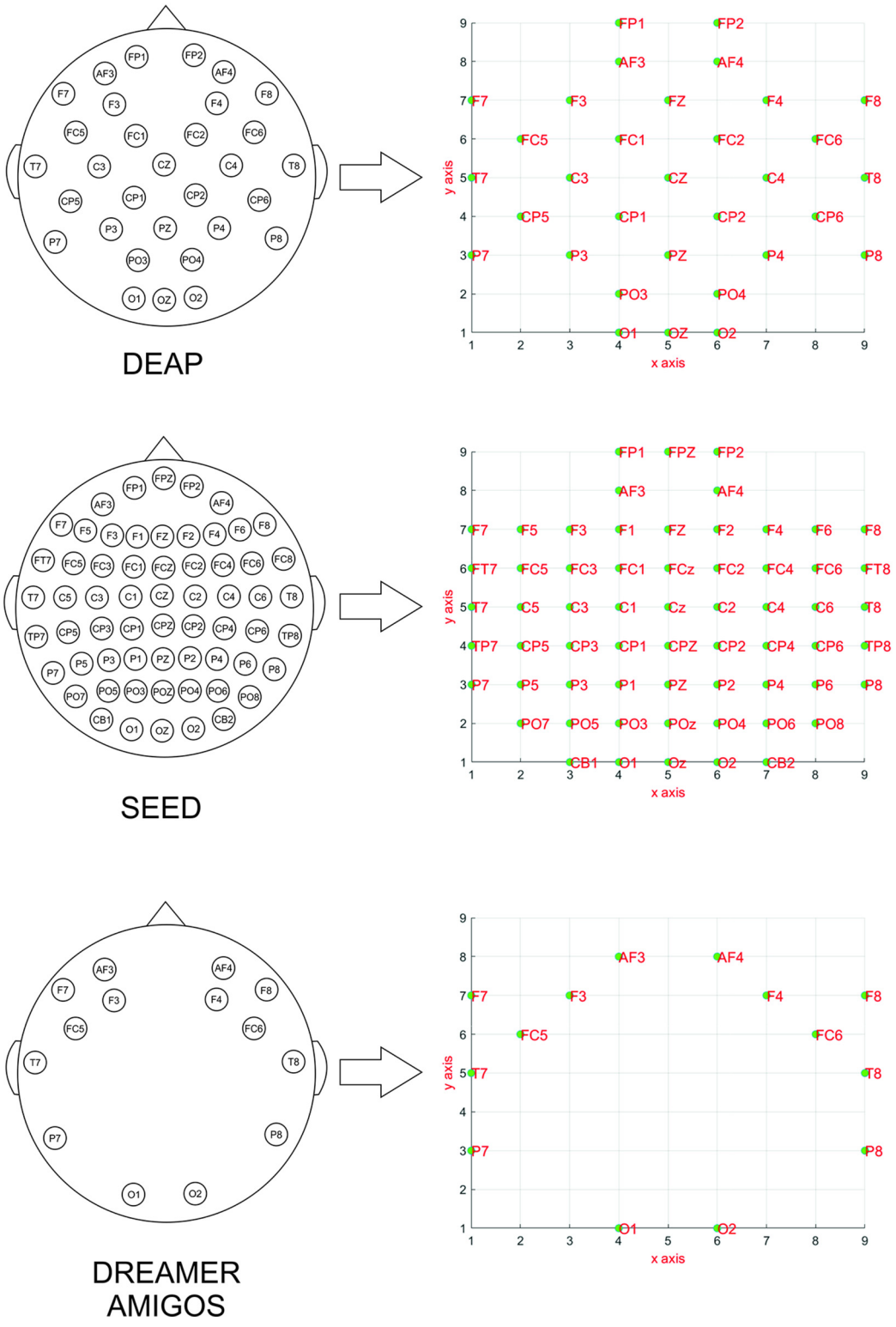

Fig. 2. 10–20 electrode mapping on the 9x9 matrix.

4.4. Model construction

Generally, the proposed model consists of three parts: extraction, fusion, and classification of the features. Features from TOPO-FM and HOLO-FM are extracted by deep learning method using separately convolutional neural network per characteristic of the EEG signal. The feature matrix is constructed by fusing extracted features that are obtained from each convolutional network. Afterward, the classification of user emotional states is performed with the machine learning method. The model is shown in the flowchart in Fig. 8, and is described with more details below.

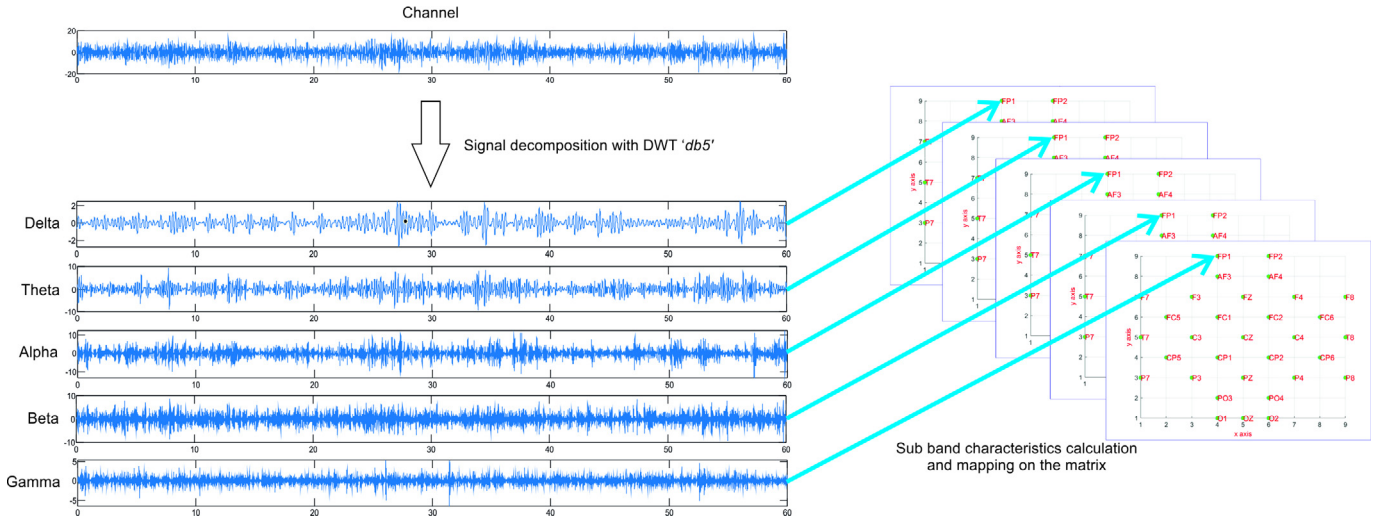

Fig. 3. Calculating characteristics from sub bands and mapping it on the feature map.

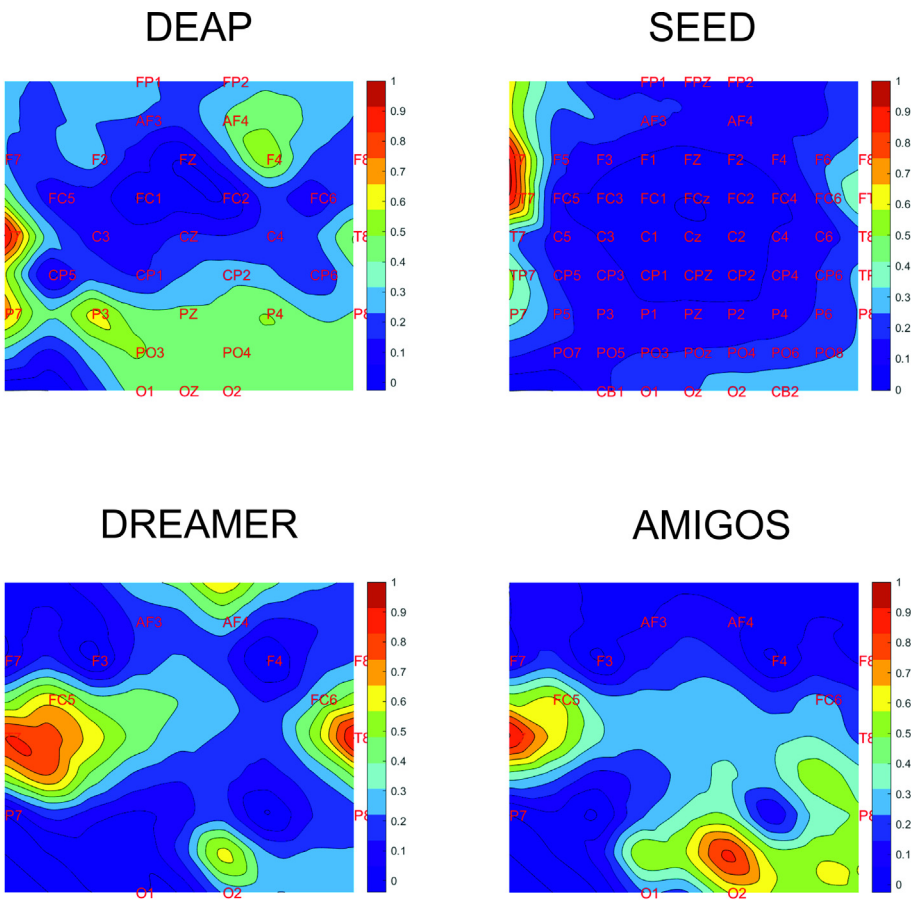

Fig. 4. TOPO-FM example.

4.5. Feature extraction

The CNN is suitable for resolving the problems of emotion recognition from EEG signals and it is widely used for learning and extracting features, as well as classification tasks in various fields. CNN is a feedforward network and it generally contains convolution layers, activation layers such as ReLU, and pooling layers. The main tasks of these layers are learning and extraction of significant features from the input object. The classification section of this kind of network is composed of a fully connected layer, and therefore we utilized only the aforementioned layers in this article.

In order to reduce the complexity of the model and to keep the execution time as short as possible, the authors used a CNN with a small number of layers [30,40]. As shown in Table 2, a deep CNN consists of seven layers to extract highly representative features. Dimensions of input color images that are fed into the network

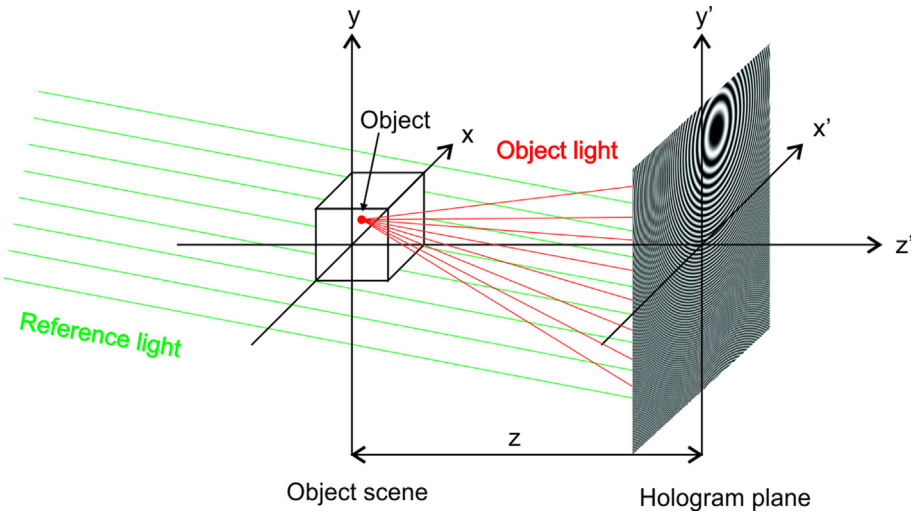

Fig. 5. Hologram scene.

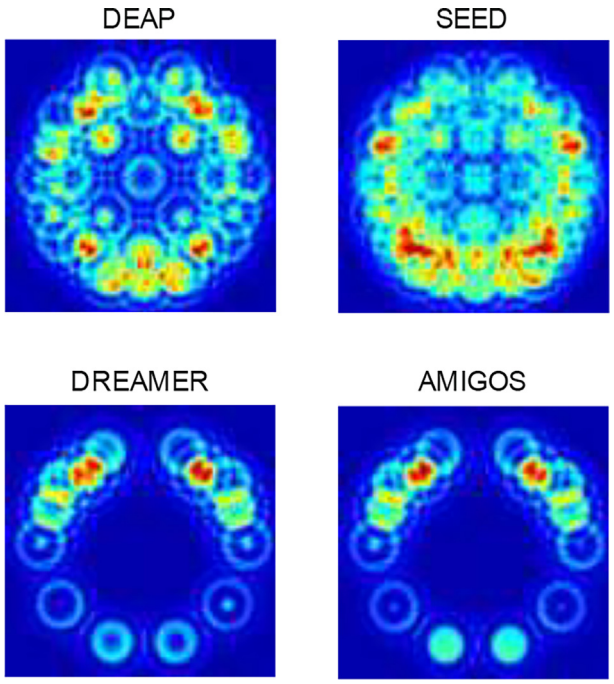

Fig. 6. HOLO-FM example.

Table 2
Details of constructed convolutional neural network.

Layer	Details	Activations	Total learnables
Image input	200x200x3, 'zerocenter' normalization	200x200x3	
Convolution	30, 2x2, stride [2 2], padding [0 0 0 0]	100x100x30	390
ReLU		100x100x30	
Max pooling	2x2, stride [2 2], padding [0 0 0 0]	50x50x30	
Convolution	10, 2x2, stride [2 2], padding [0 0 0 0]	25x25x10	1210
ReLU		25x25x10	
Fully connected		1x1x6250	39068750

are 200x200x3 with 'zerocenter' normalization. For the first two-dimensional convolution layer, we define 30 filters with the same height and width of two pixels without padding, so this layer has 390 total learnable parameters. It follows the Rectified Linear Unit (ReLU) layer and the max pooling layer with a specified 2x2 pooling region. The second convolutional layer has ten filters with the same dimensions as the first layer. It has 1210 learnable parameters, and it is followed by the second ReLU layer. Both convolutional layers, as well as, max pooling layer have a stride that equals two. At last, at the end of the CNN is a fully connected layer which consists of 39068750 most significant features, which we will later use for feature fusion, and finally for the emotion classification. Stochastic gradient descent with momentum (SGDM) is utilized for optimization, while initial learn rate, L2 regularization and mini batch size are 0.001, 0.04, and 32, respectively. Extracted features from all feature maps, are fused together, making input for the SVM machine learning classifier.

Convolutional neural networks learn high-level features in the hidden layers. The first convolutional layers learn features such as simple textures, while later ones learn high-level combinations of features learned in previous layers. Visualization of the features that have been learned by the second two-dimensional convolutional layer for the first participant in the DEAP dataset for valence affective state is shown in Fig. 7.

Since the TOPO-FM feature map consists only of the colors defined in the "jet" colormap, i.e. range from blue through green to red, the visualized features correspond accordingly.

4.6. Multi-feature fusion and classification

The learned features from multiple fully connected layers are fused together as the final representation which is used as input to supervised machine learning classification algorithm. After analyzing the literature, it can be concluded that the support vector machine is commonly used as a classificatory for extracted features with deep learning methods [24,25,26,27]. According to a survey [76] which covers 99 papers that were researching emotion recognition using EEG signals, SVM is the dominant classifier and it was used in ~59% of cases. SVM [77,78] is a powerful machine learning algorithm that is ordinarily used for classification and regression problems. It is used to distinguish different types of the label by determining an optimal separating hyper-plane. Therefore, we utilized an SVM classifier with Radial Basis Function (RBF) kernel. A 10-fold cross-validation technique was adopted to evaluate the

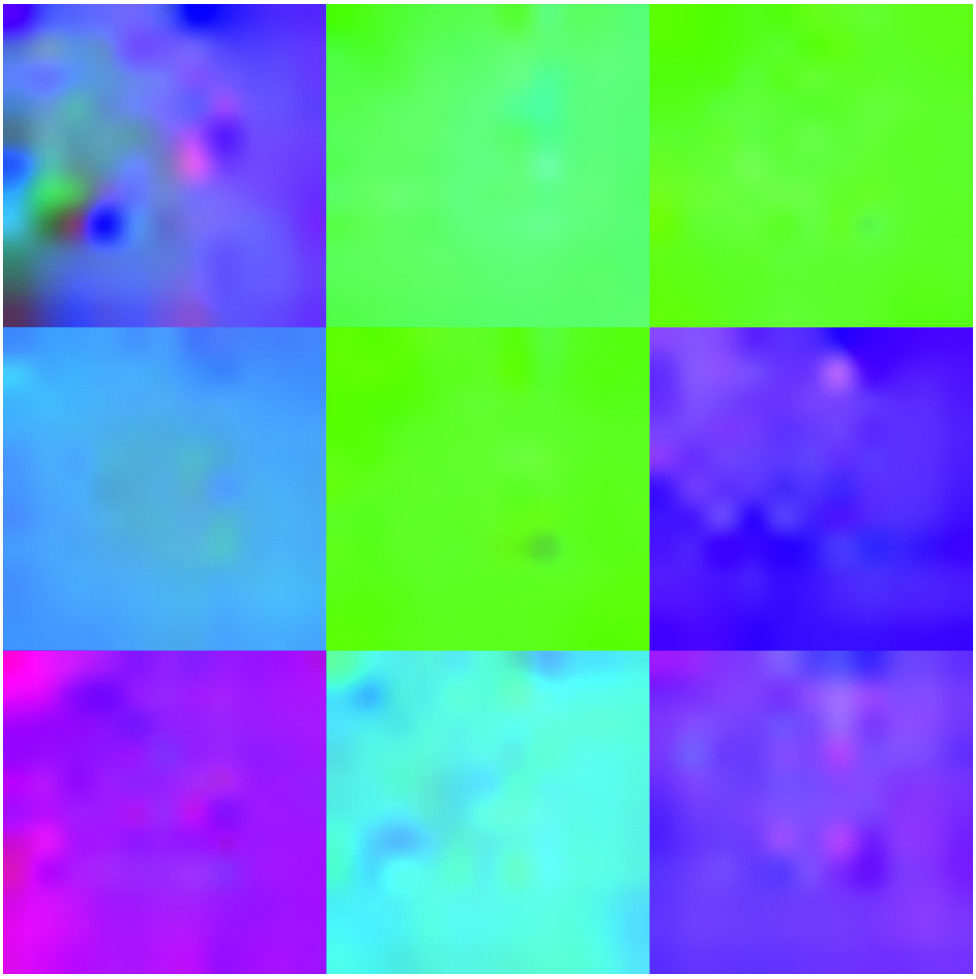

Fig. 7. Visualization of a 2D convolutional layer.

performance of our model. One of the folds was used as a test sample for each iteration, and the rest of the folds were used as training samples. This process was repeated ten times and the average classification accuracy for all participants was computed at each fold. Accuracy per participant was averaged at the end of the experiment, as well.

5. Results and discussion

In this section, we present the experimental results of our model for emotion recognition. Table 3 gives an overview of the accuracy and confidence interval with significance level $\alpha = 0.05$ for both TOPO-FM and HOLO-FM methods on all datasets in valence and arousal space. It can be clearly seen that the HOLO-FM method outperforms TOPO-FM for all datasets. While the accuracy on the DEAP dataset for these two methods is practically the same, the difference on SEED is much more significant, and amounts to 18 percentage points. DREAMER and AMIGOS differentiate their valence values for about seven, and arousal for five percentage points, respectively.

Since the feature maps are created from features calculated for each channel, these results imply that the biggest difference is on the datasets with more electrodes. Recall that SEED has 62, DEAP 32, and DREAMER and AMIGOS has 14 electrodes on their measuring devices.

Table 4 presents an additional view of using the proposed features. Features extracted from HOLO-FM and TOPO-FM, which are called TH-FM, were fused together, and then classified. The obtained results, i.e. accuracy and confidence interval with significance level $\alpha = 0.05$, do not differ much from those calculated only with TOPO-FM features.

Figs. 9–12 gives us classification accuracies per participant on each dataset. It is visible that the lines on the graphs denoting accuracy per participant are similar for TOPO-FM and HOLO-FM, but of course, for topographic maps, the values are lower. It can be said that this is expected behavior and that our methods give consistent results. The best example of this is in Fig. 10 for the SEED dataset that does not have arousal values.

In Table 5, the comparison results for each dataset are given. Beside baseline ones, which are actually studies of the dataset authors, we chose studies where authors created some kind of topographic or spectrogram images from one or more various features using only EEG signals that classify human emotions in the valence and arousal space. In studies, deep learning techniques are used for feature extraction no matter which method is used for classification. Evaluation methods are the criteria for evaluating the success of a model, and can generally be divided into the hold-out and the k-fold cross-validation methods. The most commonly used method is the k-fold cross-validation, which we utilized in this work, as well. Leave-One-Out (LOO) cross-validation is a specific case of the k-fold cross-validation where k = 1, so Leave-One-Subject-Out (LOSO) or Leave-One-Trial-Out (LOTO) can be found in the literature. Hence, Table 5 provided a comparison to other studies that used 10-fold cross validation which is common in this

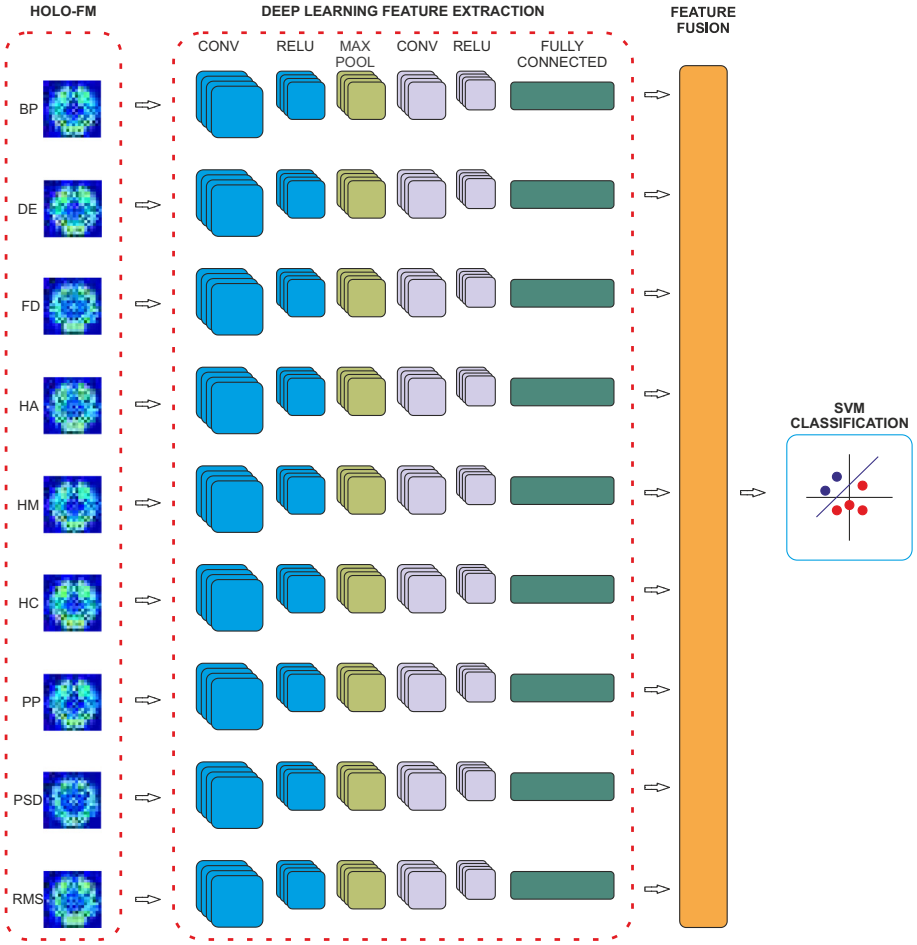

Fig. 8. Flow chart of feature extraction, fusion and classification.

Table 3
Accuracy for TOPO-FM and HOLO-FM in valence and arousal state per dataset.

	TOPO-FM		HOLO-FM	
	Valence	Arousal	Valence	Arousal
DEAP	0.7630 ± 0.0186	0.7654 ± 0.0269	0.7661 ± 0.0213	0.7772 ± 0.0287
SEED	0.7037 ± 0.0341		0.8845 ± 0.0156	
DREAMER	0.8196 ± 0.0169	0.8492 ± 0.0276	0.8820 ± 0.0084	0.9043 ± 0.0139
AMIGOS	0.8063 ± 0.0176	0.8575 ± 0.0194	0.8739 ± 0.0094	0.9054 ± 0.0102

Table 4
Accuracy for fused TOPO-FM and HOLO-FM in valence and arousal state per dataset.

	TH-FM	
	Valence	Arousal
DEAP	0.7491 ± 0.0193	0.7544 ± 0.0271
SEED	0.7311 ± 0.0302	
DREAMER	0.8125 ± 0.0173	0.8510 ± 0.0262
AMIGOS	0.7954 ± 0.0126	0.8507 ± 0.0204

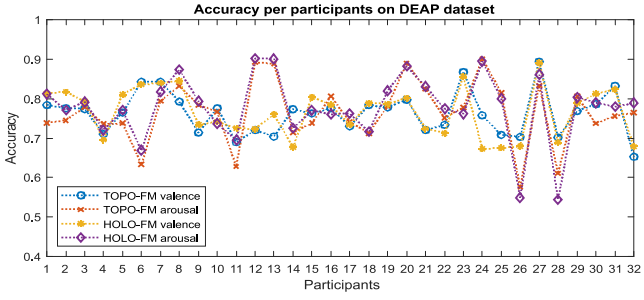

Fig. 9. Accuracy per participants on DEAP dataset.

research area [31]. AMIGOS and DREAMER datasets were published not so long ago, so there are only a few comparative works beside the baseline one.

The results in [31] on the DEAP dataset show that method for creating images by mapping electrodes on the two-dimensional matrix can give good accuracy, and authors reported 66.73% for valence, 68.28% for arousal, and 67.25% for dominance. The approach of [35] proposed a method that creates spatial maps. Graph regularized extreme learning machine (GELM) and SVM were combined together for emotion prediction, and the authors got the same accuracy rate of 62% as the authors of the DEAP dataset [47].

PSD feature was calculated in [31], as well as in [34] and [35], and in comparison, they achieved results 66.73% – 71.09% and

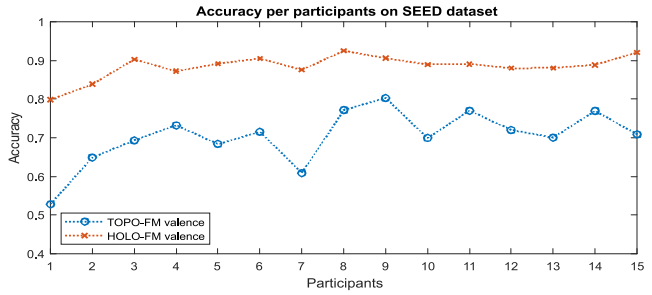

Fig. 10. Accuracy per participants on SEED dataset.

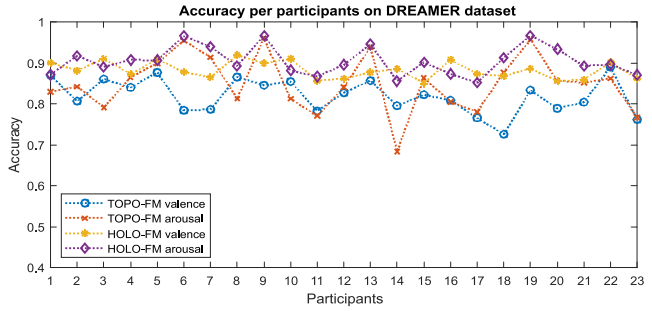

Fig. 11. Accuracy per participants on DREAMER dataset.

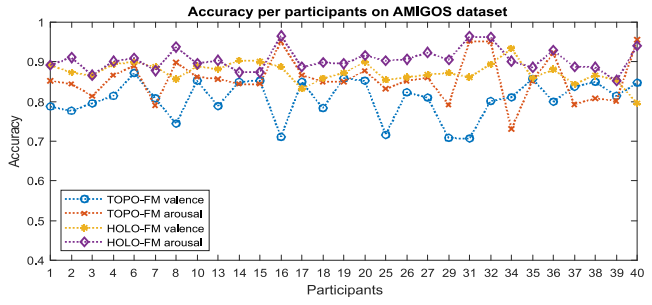

Fig. 12. Accuracy per participants on AMIGOS dataset.

68.28% – 72.58% for valence and arousal, respectively. The highest classification accuracy achieved in [35] is 88% by fusing information propagation patterns and activation differences in the brain. The SEED dataset authors [48] achieved 86.65% for emotion recognition, but [35] and HOLO-FM methods outperformed their approach with a classification accuracy of 88% and 88.45%, respectively. Valence of 62.49% and arousal of 62.17% have been reached in [49], which is fairly effective for emotion recognition, however the study [34], TOPO-FM and HOLO-FM methods achieved more significant level of accuracy. Authors of the AMIGOS dataset [50] have reached the emotion prediction of 57.60% for valence and 59.20% for arousal using spectral power features. If we consider the case in [34], we can see that they outperform accuracy results for both affective states, but still underperform both of our methods.

Computational cost on our testing machine (Windows 10, Matlab 2019, HP ProBook with integrated Intel GPU, 12 GB RAM, CPU Intel Core i5 2.40 GHz), for the Hjorth activity feature of the first participant's trial for each dataset with 10-fold evaluation method, is presented in Table 6:

The results for HOLO-FM/TOPO-FM and for Valence/Arousal are similar, so we do not show them individually. The proposed model currently cannot be used to recognize emotions in real-time, and this is one of the biggest challenges in future work.

This study has demonstrated that TOPO-FM and HOLO-FM are effective methods for emotion recognition. In DEAP, DREAMER, and AMIGOS datasets, valence and arousal emotion states are divided on low and high class, while in SEED the emotional states are divided into positive and negative values. The classification accuracies in valence and arousal space, and the overall accuracy with SEED for all three experiments, show that proposed methods perform better than comparable methods previously reported.

6. Conclusion

In this paper, we propose the usage of the EEG signal characteristics for creating topographic and holographic feature maps, called TOPO-FM and HOLO-FM, respectively. The 10–20 system is utilized to map values of each characteristic at the electrode location. Three-dimensional space is defined by displaying the value of the characteristic of the signal in the electrode position on the head. TOPO-FM method uses the values of the signal characteristics in order to display a topographic map, while HOLO-FM utilizes the computer-generated holography to create two-dimensional feature maps from points in the space. The deep learning technique was employed for feature learning from TOPO-FM and HOLO-FM, respectively. The features obtained from each individual neural network are merged into one feature matrix, whereupon machine learning-based classification was adopted to determine the affective state. DEAP, SEED, DREAMER, and AMIGOS are the datasets that were used for performance evaluation of our newly proposed model.

Based on our results, which outperform studies analyzed in this work with the same EEG data set, we concluded that proposed methods could have a significant role in future emotion recognition models. This approach can contribute to the improvement of applications developed for medical purposes, e.g. for people with disabilities ("doing by thinking"), authentication and identification of the person, detection of driver fatigue, neurogaming, etc.

In our future work, we plan to explore more different feature sets, beside the one that is given in this paper, in order to improve emotion recognition performance. Further, we will extend the proposed model by combining datasets for training and testing purposes. Finally, special efforts will be put on participant independent cross-validation with classification in three-dimensional space, i.e. valence-arousal-dominance emotional model.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Partially supported by the Croatian Science Foundation under project UIP-2014-09-3875 and by the Virtual Telemedicine Assistance - VITA, a project co-financed by the Croatian Government and the European Union through the European Regional Development Fund - the Competitiveness and Cohesion Operational Programme (KK.01.1.1.01).

Table 5
Comparison with other studies.

Dataset	Study	Used feature(s)	Classification method(s)	Evaluation method(s)	Best accuracy
DEAP	Koelstra et al. [47]	PSD	NB	LOO-CV	V: 0.6200
					A: 0.5760
	Chao et al. [31]	PSD, MFM	CapsNet	10-fold CV	V: 0.6673
					A: 0.6828
	Siddharth et al. [34]	RGB heat-map	CNN + ELM	10-fold CV	V: 0.7109
					A: 0.7258
	Li, P. et. al. [35]	Spatial map	GELM + SVM	10-fold CV	0.6200
	Our method	TOPO-FM	CNN + SVM	10-fold CV	V: 0.7630
					A: 0.7654
	Our method	HOLO-FM	CNN + SVM	10-fold CV	V: 0.7661
					A: 0.7772
SEED	Zheng et al. [48]	Feature map	kNN, LR, SVM, DBN	–	0.8665
	Li, P. et. al. [35]	Spatial map	GELM + SVM	10-fold CV	0.8800
	Our method	TOPO-FM	CNN + SVM	10-fold CV	0.7037
	Our method	HOLO-FM	CNN + SVM	10-fold CV	0.8845
DREAMER	Katsigiannis et al. [49]	PSD	SVM	10-fold CV	V: 0.6249
					A: 0.6217
	Siddharth et al. [34]	RGB heat-map	CNN + ELM	10-fold CV	V: 0.7899
					A: 0.7923
	Our method	TOPO-FM	CNN + SVM	10-fold CV	V: 0.8196
					A: 0.8492
	Our method	HOLO-FM	CNN + SVM	10-fold CV	V: 0.8820
					A: 0.9043
AMIGOS	Miranda et al. [50]	PSD, SPA	SVM	LOSO-CV	V: 0.5760
					A: 0.5920
	Siddharth et al. [34]	RGB heat-map	CNN + ELM	10-fold CV	V: 0.8302
					A: 0.7913
	Our method	TOPO-FM	CNN + SVM	10-fold CV	V: 0.8063
					A: 0.8575
	Our method	HOLO-FM	CNN + SVM	10-fold CV	V: 0.8739
					A: 0.9054

PSD: Power Spectral Density, NB: Naïve Bayes, LOO: Leave-One-Out, CV: Cross-Validation, MFM: Multiband Feature Matrix, ELM: Extreme Learning Machines, GELM: Graph regularized Extreme Learning Machine, kNN: K nearest neighbor, LR: Logistic Regression, SVM: Support Vector Machine, DBN: Deep Belief Networks, SPA: Spectral Power Asymmetry.

Table 6
Computational cost.

	Feature extraction	Features fusion	Feature classification
DEAP	8.5965 s	0.7755 s	0.3451 s
SEED	2.9945 s	0.3680 s	0.9349 s
DREAMER	4.0987 s	0.3778 s	0.6986 s
AMIGOS	3.8562 s	0.3515 s	0.7531 s

References

[1] R.A. Khalil, E. Jones, M.I. Babar, T. Jan, M.H. Zafar, T. Alhussain, Speech emotion recognition using deep learning techniques: A review, IEEE Access 7 (2019) 117327–117345.
[2] A. Alreshidi, M. Ullah, Facial Emotion Recognition Using Hybrid Features, Informatics 6 (2020) 7.
[3] F. Ahmed, A.S.M.H. Bari, M.L. Gavrilova, Emotion Recognition From Body Movement, IEEE Access 8 (2020) 11761–11781.
[4] L. Shu, Y. Yu, W. Chen, H. Hua, Q. Li, J. Jin, X. Xu, Wearable Emotion Recognition Using Heart Rate Data from a Smart Bracelet, Sensors 20 (2020) 718.
[5] L. Shu, J. Xie, M. Yang, Z. Li, Z. Li, D. Liao, X. Xu, X. Yang, A Review of Emotion Recognition Using Physiological Signals, Sensors 18 (2018) 2074.
[6] X. Chen, W. Zeng, Y. Shi, J. Deng, Y. Ma, Intrinsic Prior Knowledge Driven CICA fMRI Data Analysis for Emotion Recognition Classification, IEEE Access 7 (2019) 59944–59950.
[7] K.D. Singh, C. Fioravanti, A. Elshahabi, S. Ruiz, R. Sitaram, B. Braun, Involvement of top-down networks in the perception of facial emotions: A magnetoencephalographic investigation, NeuroImage 222 (2020) 117075.
[8] V. Chamola, A. Vineet, A. Nayyar, E. Hossain, Brain-Computer Interface-Based Humanoid Control: A Review, Sensors 20 (2020) 3620.
[9] Udovicic, G.; Topic, A.; Russo, M. Wearable technologies for smart environments: A review with emphasis on BCI. 24th International Conference on Software, Telecommunications and Computer Networks (SoftCOM), Split, 2016; pp 1-9.
[10] P. Ekman, An argument for basic emotions, Cogn. Emot. 6 (3–4) (1992) 169–200.
[11] R. Plutchik, The nature of emotions, Am. Sci. 89 (4) (2001) 344–350.
[12] J.A. Russell, A circumplex model of affect, *J. Pers. Soc. Psychol.* **Dec.** 38 (6) (1980) 1161–1178.
[13] A. Mehrabian, Comparison of the PAD and PANAS as models for describing emotions and for differentiating anxiety from depression, Journal of Psychopathology and Behavioral Assessment 19 (1997) 331–357.
[14] Chen, M.; Hao, Y. Label-less Learning for Emotion Cognition. IEEE Transactions on Neural Networks and Learning Systems July 2020, 31 (7), 2430-2440.
[15] J. Chen, B. Hu, P. Moore, X. Zhang, X. Ma, Electroencephalogram-based emotion assessment system using ontology and data mining techniques, Appl. Soft Comput. 30 (2015) 663–674.
[16] R. Jenke, A. Peer, M. Buss, Feature extraction and selection for emotion recognition from EEG, IEEE Trans. Affective Comput. 5 (3) (2014) 327–339.
[17] Z. Gao, X. Wang, Y. Yang, Y. Li, K. Ma, G. Chen, A Channel-fused Dense Convolutional Network for EEG-based Emotion Recognition, IEEE Transactions on Cognitive and Developmental Systems (2020).
[18] A. Craik, Y. He, J.L. Contreras-Vidal, Deep learning for electroencephalogram (EEG) classification tasks: a review, J. Neural Eng. 16 (3) (2019) 031001.
[19] Z. Gao, Y. Li, Y. Yang, N. Dong, X. Yang, C. Grebogi, A Coincidence Filtering-based Approach for CNNs in EEG-based Recognition, IEEE Trans. Ind. Inf. (2019).
[20] Khare, S. K.; Bajaj, V. Time-frequency representation and convolutional neural network-based emotion recognition. IEEE transactions on neural networks and learning systems 2020.
[21] H. Yang, J. Han, K. Min, A multi-column CNN model for emotion recognition from EEG signals, Sensors 19 (21) (2019) 4736.
[22] S.A. Vahora, N.C. Chauhan, Deep neural network model for group activity recognition using contextual relationship, Engineering Science and Technology, an International Journal 22 (1) (2019) 47–54.
[23] J. Zhu, H. Chen, W. Ye, A Hybrid CNN–LSTM Network for the Classification of Human Activities Based on Micro-Doppler Radar, IEEE Access 8 (2020) 24713–24720.
[24] A. Ruiz-Garcia, M. Elshaw, A. Altahhan, V. Palade, A hybrid deep learning neural approach for emotion recognition from facial expressions for socially assistive robots, Neural Comput. Appl. 29 (7) (2018) 359–373.

[25] K. Guo, H. Mei, X. Xie, X. Xu, A Convolutional Neural Network Feature Fusion Framework with Ensemble Learning for EEG-based Emotion Classification, IEEE MTT-S International Microwave Biomedical Conference (IMBioC) 1 (May 2019) 1–4.
[26] H. Chao, L. Dong, Y. Liu, B. Lu, Improved Deep Feature Learning by Synchronization Measurements for Multi-Channel EEG Emotion Recognition, Complexity (2020).
[27] W. Liu, W.L. Zheng, B.L. Lu, Emotion recognition using multimodal deep learning, International conference on neural information processing (2016) 521–529.
[28] E. Kanjo, E.M. Younis, C.S. Ang, Deep learning analysis of mobile physiological, environmental and location sensor data for emotion detection, Information Fusion 49 (2019) 46–56.
[29] W. Tao, C. Li, R. Song, J. Cheng, Y. Liu, F. Wan, X. Chen, EEG-based emotion recognition via channel-wise attention and self attention, IEEE Trans. Affective Comput. (2020).
[30] Y. Li, J. Huang, H. Zhou, N. Zhong, Human emotion recognition with electroencephalographic multidimensional features by hybrid deep neural networks, Applied Sciences 7 (10) (2017) 1060.
[31] H. Chao, L. Dong, Y. Liu, B. Lu, Emotion recognition from multiband EEG signals using CapsNet, Sensors 19 (9) (2019) 2212.
[32] X. Li, D. Song, P. Zhang, G. Yu, Y. Hou, B. Hu, Emotion recognition from multi-channel EEG data through Convolutional Recurrent Neural Network, IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (2016) 352–359.
[33] Z.M. Wang, S.Y. Hu, H. Song, Channel Selection Method for EEG Emotion Recognition Using Normalized Mutual Information, IEEE Access 7 (2019) 143303–143311.
[34] S. Siddharth, T.P. Jung, T.J. Sejnowski, Utilizing deep learning towards multi-modal bio-sensing and vision-based affective computing, IEEE Trans. Affective Comput. (2019).
[35] P. Li, H. Liu, Y. Si, C. Li, F. Li, X. Zhu, P. Xu, EEG based emotion recognition by combining functional connectivity network and local activations, IEEE Trans. Biomed. Eng. 66 (10) (2019) 2869–2881.
[36] W. Zheng, J. Zhu, B. Lu, Identifying Stable Patterns over Time for Emotion Recognition from EEG, IEEE Trans. Affective Comput. 10 (3) (2019) 417–429.
[37] S. Keshmiri, M. Shiomi, H. Ishiguro, Entropy of the Multi-Channel EEG Recordings Identifies the Distributed Signatures of Negative, Neutral and Positive Affect in Whole-Brain Variability, Entropy 21 (12) (2019) 1228.
[38] S. Keshmiri, M. Shiomi, H. Ishiguro, Emergence of the Affect from the Variation in the Whole-Brain Flow of Information, Brain sciences 10 (1) (2019) 8.
[39] F. Wang, S. Wu, W. Zhang, Z. Xu, Y. Zhang, C. Wu, S. Coleman, Emotion recognition with convolutional neural network and EEG-based EFDMs, Neuropsychologia 146 (2020) 107506.
[40] S. Hwang, K. Hong, G. Son, H. Byun, Learning CNN features from DE features for EEG-based emotion recognition, Pattern Anal. Appl. (2019).
[41] C. Li, Z. Bao, L. Li, Z. Zhao, Exploring temporal representations by leveraging attention-based bidirectional LSTM-RNNs for multi-modal emotion recognition, Inf. Process. Manage. 57 (3) (2020) 102185.
[42] D. Gabor, A New Microscopic Principle. *Nature* 161 (1948) 777–778.
[43] P.W.M. Tsang, T. Poon, Review on the State-of-the-Art Technologies for Acquisition and Display of Digital Holograms, IEEE Trans. Ind. Inf. 12 (3) (2016) 886–901.
[44] P.C.G.D.H. Lobaz, Tools, Getting started in computer generated display holography, in: Proceedings of the 11th International Symposium on Display Holography -, 2018, 2018,, pp. 159–165.
[45] Y. Jo, S. Park, J. Jung, J. Yoon, H. Joo, M.H. Kim, S.-J. Kang, M.C. Choi, S.Y. Lee, Y. Park, Holographic deep learning for rapid optical screening of anthrax spores, Sci. Adv. 3 (8) (2017) e1700606.
[46] S.J. Kim, C. Wang, B. Zhao, H. Im, J. Min, H.J. Choi, J. Tadros, N.R. Choi, C.M. Castro, R. Weissleder, H. Lee, Deep transfer learning-based hologram classification for molecular diagnostics, Sci. Rep. 8 (1) (2018) 1–12.
[47] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, I. Patras, DEAP: A database for emotion analysis; using physiological signals, IEEE Trans. Affective Comput. 3 (1) (2012) 18–31.
[48] W.-L. Zheng, B.-L. Lu, Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks, *IEEE Trans. Auton. Ment. Dev.* **Sep.** 7 (3) (2015) 162–175.
[49] S. Katsigiannis, N. Ramzan, DREAMER: A Database for Emotion Recognition Through EEG and ECG Signals from Wireless Low-cost Off-the-Shelf Devices, *IEEE J. Biomed. Health. Inf.* **Jan.** 22 (1) (2018) 98–107.
[50] J.A. Miranda-Correa, M.K. Abadi, N. Sebe, I. Patras, AMIGOS: A Dataset for Affect, Personality and Mood Research on Individuals and Groups, IEEE Transactions on Affective Computing Nov, 2018.
[51] V.L. Towle, J. Bolanos, D. Suarez, K. Tan, R. Grzeszczuk, D.N. Levin, R. Cakmur, S. A. Frank, J.-P. Spire, The spatial location of EEG electrodes: Locating the best-fitting sphere relative to cortical anatomy, *Electroencephalogr. Clin. Neurophysiol.* **Jan.** 86 (1) (1993) 1–6.
[52] Z. Lan, O. Sourina, L. Wang, R. Scherer, G.R. Müller-Putz, Domain adaptation techniques for EEG-based emotion recognition: a comparative study on two public datasets, IEEE Transactions on Cognitive and Developmental Systems 11 (1) (2018) 85–94.
[53] X. Li, D. Song, P. Zhang, Y. Zhang, Y. Hou, B. Hu, Exploring EEG features in cross-subject emotion recognition, Front. Neurosci. 12 (2018) 162.
[54] P. Pandey, K.R. Seeja, Subject independent emotion recognition system for people with facial deformity: an EEG based approach. *Journal of Ambient Intelligence and Humanized*, Computing (2020).
[55] Y. Cimtay, E. Ekmekcioglu, Investigating the Use of Pretrained Convolutional Neural Network on Cross-Subject and Cross-Dataset EEG Emotion Recognition, Sensors (Basel, Switzerland) 20 (7) (2020) 2034.
[56] J.D. Morris, Observations: Sam: The self-assessment manikin; an efficient cross-cultural measurement of emotional response, *Journal of Advertising Research* **Jan.** 35 (8) (1995) 63–68.
[57] S.-H. Wang, H.-T. Li, E.-J. Chang, (Andy) Wu, A.-Y., Entropy-Assisted Emotion Recognition of Valence and Arousal Using XGBoost Classifier, Artificial Intelligence Applications and Innovations (AIAI) (2018) 249–260.
[58] M. Murugappan, N. Ramachandran, Y. Sazali, Classification of human emotion from EEG using discrete wavelet transform, J. Biomed. Sci. Eng. 3 (4) (2010) 390.
[59] J. Atkinson, D. Campos, Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers, Expert Syst. Appl. 47 (2016) 35–41.
[60] G. Chen, X. Zhang, Y. Sun, J. Zhang, Emotion Feature Analysis and Recognition Based on Reconstructed EEG Sources, IEEE Access 8 (2020) 11907–11916.
[61] Zheng, W.; Liu, W.; Lu, Y.; Lu, B.; Cichocki, A. EmotionMeter: A Multimodal Framework for Recognizing Human Emotions. IEEE Transactions on Cybernetics March 2019, 49 (3), 1110-1122.
[62] X. Chai, Q. Wang, Y. Zhao, Y. Li, D. Liu, X. Liu, O. Bai, A Fast, Efficient Domain Adaptation Technique for Cross-Domain Electroencephalography (EEG)-Based Emotion Recognition, Sensors (Basel) 17 (5) (2017) 1014.
[63] O. Sourina, Y. Liu, A fractal-based algorithm of emotion recognition from EEG using arousal-valence model, International Conference on Bio-inspired Systems and Signal Processing (2011) 209–214.
[64] Khosrowabadi, R.; Rahman, A. W. b. A. Classification of EEG correlates on emotion using features from Gaussian mixtures of EEG spectrogram. Proceeding of the 3rd International Conference on Information and Communication Technology for the Moslem World (ICT4M) 2010, E102-E107.
[65] T. Higuchi, Approach to an irregular time series on the basis of the fractal theory, Physica D 31 (2) (1988) 277–283.
[66] B. Hjorth, EEG analysis based on time domain properties, Electroencephalogr. Clin. Neurophysiol. 29 (1970) 306–310.
[67] S.-H. Oh, Y.-R. Lee, H.-N. Kim, A novel EEG feature extraction method using Hjorth parameter, International Journal of Electronics and Electrical Engineering 2 (2) (2014) 106–110.
[68] A. Dzedzickis, A. Kaklauskas, V. Bucinskas, Human Emotion Recognition: Review of Sensors and Methods, Sensors 20 (3) (2020) 592.
[69] A.M. Bhatti, M. Majid, S.M. Anwar, B. Khan, Human emotion recognition and analysis in response to audio music using brain signals, Comput. Hum. Behav. 65 (2016) 267–275.
[70] T. Song, W. Zheng, P. Song, Z. Cui, EEG Emotion Recognition Using Dynamical Graph Convolutional Neural Networks, IEEE Trans. Affective Comput. (2018).
[71] T. Musha, Y. Terasaki, H.A. Haque, G.A. Ivamitsky, Feature extraction from EEGs associated with emotions, Artificial Life and Robotics 1 (1) (1997) 15–19.
[72] D.O. Bos, EEG-based emotion recognition, The Influence of Visual and Auditory Stimuli 56 (3) (2006) 1–17.
[73] A.A. Abdul-Latif, I. Cosic, D.K. Kumar, B. Polus, C. Da Costa, Power changes of EEG signals associated with muscle fatigue: the root mean square analysis of EEG bands, in: Proceedings of the 2004 Intelligent Sensors, Sensor Networks and Information Processing Conference, 2004, pp. 531–534.
[74] M. Murugappan, M. Rizon, R. Nagarajan, S. Yaacob, Inferring of human emotional states using multichannel EEG, European Journal of Scientific Research 48 (2) (2010) 281–299.
[75] B. Nakisa, M.N. Rastgoo, D. Tjondronegoro, V. Chandran, Evolutionary computation algorithms for feature selection of EEG-based emotion recognition using mobile sensors, Expert Syst. Appl. 93 (2018) 143–155.
[76] S.M. Alarcão, M.J. Fonseca, Emotions Recognition Using EEG Signals: A Survey, IEEE Trans. Affective Comput. 10 (3) (2019) 374–393.
[77] V.N. Vapnik, The nature of statistical learning theory, Springer, New York, 1995.
[78] C. Cortes, V. Vapnik, Support vector networks, Machine Learning 20 (1995) 273–297.