Preference Classification Using Electroencephalography (EEG) and Deep Learning

Jason Teo, Chew Lin Hou and James Mountstephens
Faculty of Computing & Informatics, Universiti Malaysia Sabah,
Jalan UMS, 88400 Kota Kinabalu, Sabah, Malaysia.
jtwteo@ums.edu.my

***Abstract*—Electroencephalogram (EEG)-based emotion classification is rapidly becoming one of the most intensely studied areas of brain-computer interfacing (BCI). The ability to passively identify yet accurately correlate brainwaves with our immediate emotions opens up truly meaningful and previously unattainable human-computer interactions such as in forensic neuroscience, rehabilitative medicine, affective entertainment and neuro-marketing. One particularly useful yet rarely explored areas of EEG-based emotion classification is preference recognition [1], which is simply the detection of like versus dislike. Within the limited investigations into preference classification, all reported studies were based on musically-induced stimuli except for a single study which used 2D images. We present two EEG-based preference classification studies: using (1) kNN for a 10-subject EEG classification problem; (2) deep learning for an expanded 16-subject EEG classification problem. We show that inter-subject variability introduces significant classification problems when larger cohorts of test subjects are used and that deep learning shows promising results in terms of addressing this inter-subject variability problem in EEG-based preference classification.**

***Index Terms*—Deep Learning; Electroencephalography; Emotion Classification; Preference Recognition; Visual Aesthetics.**

I. INTRODUCTION

Being emotional is a crucial part of what makes us human. Indeed, it has been argued that human consciousness evolved from the primordial emotions of early man [2]. Having the ability to sense our immediate situation and respond with the corresponding emotion is what human brains have been evolved to perform on a predominantly subconscious level [3]. The human brain constantly processes the sensory inputs coming in from around us and provides us with the subsequent emotional output arising from that particular experience. One of the most effective approaches to human emotion classification is based on the analysis and interpretation of electroencephalograms [4]. Feelings of anger, sadness, happiness, surprise and fear are some of the more commonly studied emotions in EEG-based classifications [5, 6, 7, 8].

As part of the diverse emotions that humans feel, one of the most basic emotions that we experience is the feeling of like versus dislike. Everyday experiences such as tasting a new dish, watching a new movie or visiting a new shop will all certainly evoke some preference response. Although being one of the most frequently encountered emotions, surprisingly few attempts have been made to investigate the use of EEG as a means to recognize human preferences. Moreover, in the limited studies that have been reported on EEG-based preference classification, EEG-based preference classification studies have predominantly relied on the use of music as the primary stimuli [9, 10, 11, 12, 13, 14] with only a solitary study which uses 2D images [15].

In our ongoing investigations on EEG-based preference classification, we used novel stimuli in the form of 3D rotating objects [1]. Our ultimate goal is to create a thought-based 3D computer-aided design (CAD) system to assist designers in rapid prototyping of 3D objects for 3D printing, hence the novel use of 3D rotating objects as the visual stimuli in our experiments. Our early work reported promising results of up to 80%. However, these preliminary experiments involved only a small test group of 5 subjects. In order to test its generalizability, scaling the study up to a larger cohort of test subjects would be required. It is well-known that EEG-based studies involving larger cohorts generally tend to have significantly reduced accuracies due to inter-subject variability [16] in addition to intra-subject variability [17]. Therefore, the main objective of the current study is to expand the investigation to larger cohorts of 10 and 16 subjects respectively and attempt to maintain the level of classification accuracy.

This paper is organized as follows: Section 2 presents some background material on machine learning and deep learning in EEG-based emotion classification; Section 3 explains the setup of our EEG system for classification of aesthetic preference; Section 4 discusses the classification results obtained from the various machine learning techniques investigated; and finally Section 5 summarizes the main findings of our work and suggests some useful avenues to pursue as future work.

II. RELATED MATERIAL

EEG-based classification of human emotions using machine learning techniques can be broadly classified into three categories: those that employ features from the (1) time, (2) frequency, and (3) a combination of time and frequency domains. Emotion classification based on the time domain relies on the detection of event-related potentials (ERPs), which can be further classified into signals that manifest with a short, medium and long latency post exposure to the stimuli. ERPs were used to automatically identify the test subjects' preferences based on valence and arousal, obtaining average classification rates of 55.7% for arousal and 58.8% for valence [13].

Emotion classification based on the frequency domain relies on the analysis of spectral power from the delta, theta, alpha, beta and gamma bands. Music preference was classified using features extracted with the Common Spatial

Patterns (CSPs) method, obtaining an accuracy of 74.8% with linear support vector machines (SVMs) [10]. Similarly, music preference was also successfully classified with an accuracy rate of 85.7% using SVMs with features extracted using a standard Fast Fourier Transform (FFT) [12]. In a more recent and only study which uses non-musically-based stimuli, radial SVMs were able to achieve 88.5% classification accuracy for 2D image preference using spectral analysis [15].

Finally, emotion classification using both the time and frequency (TF) domains relies on the analysis of spectral power at specific time windows that span the entire duration of the measurement period. Three different TF analysis techniques were tested in conjunction with a number of machine learning algorithms to classify music preference where it was found that the k-nearest neighbors (kNN) classifier performed the best with 86.5% accuracy [9]. A follow-up study by the same researchers using a finer grained model which grouped the stimuli into familiar versus unfamiliar music obtained a higher accuracy of 91.0%, also using a kNN classifier [11]. A TF study utilizing Short-Time Fourier Transform (STFT) reported classification rates of 98.0% for music preference using a kNN classifier [14].

Deep learning in the form of Deep Belief Networks (DBNs) was used to classify the preferences of 32 subjects when viewing short music videos [18]. Deep learning is achieved in DBNs by stacking a number of Restricted Boltzmann Machines (RBMs) on each other, then using the output from a lower-level RBM to serve as the input to a higher-level RBM and so on within the multi-layer stack of RBMs. It was found to outperform a range of different SVMs and standard RBMs achieving an average of 77.8% classification accuracy. DBNs were also used to classify the emotions of 6 subjects when viewing short video clips that were either eliciting positive or negative emotions [19]. A novel critical selection method was employed to identify and use only the top 5 EEG channels for the classification tasks. It was reported that DBN's with the critical channel selection approach achieved accuracies of 87.6%, which were slightly better than Extreme Learning Machines (ELMs) and SVMs while significantly better than kNNs. In these two studies, it is worth noting that the training and prediction tasks were done on a per-subject basis and not over the entire cohort. This means that the classification process requires retraining for each new subject before prediction can take place. This approach effectively removes the challenge of dealing with inter-subject variability and only caters for intra-subject variability.

In the only deep learning preference classification study which attempts to perform classification over the entire cohort, a combination of unsupervised learning in the form of stacked autoencoders (AEs) with supervised learning of softmax classifier was used to predict the emotional states of 32 subjects based on valence and arousal. Despite using a very large number of neurons for the deep learner, which were reported by the authors to result in an extended amount of computational time required for training, and further augmenting the system with additional Principal Component Analysis (PCA) and covariate shift adaptation (CSA) to preprocess the features, the best prediction rates using leave-one-out cross-validation (LOOCV) were very low at 53.4% for valence and 52.0% for arousal [20]. This shows that inter-subject variability significantly adds to the difficulty of EEG-based preference classification in addition to intra-subject variability.

III. METHODS

10 subjects (5 female and 5 male, mean age 22.40) and 16 subjects (8 female and 8 male, mean age = 22.44) respectively participated in this study. The subjects do not have any known history of psychiatric illnesses and had normal or corrected-to-normal vision. The subjects are briefed on the data acquisition process before the experiments commenced. The ABM B-Alert X10 with 9 electrode channels (POz, Fz, Cz, C3, C4, F3, F4, P3 and P4) was used as the EEG acquisition device. A subject wearing the headset is as shown in Figure 1(a). The programming languages used to develop this system were MATLAB, Java and R: Java was used for displaying the visual stimuli; MATLAB was used for the integration with the B-Alert X10's SDK; and signal preprocessing, feature extraction, and classification were carried out using R.

(a)

(b)

Figure 1: (a) A user fitted with the EEG acquisition device; (b) Process flow of data acquisition.

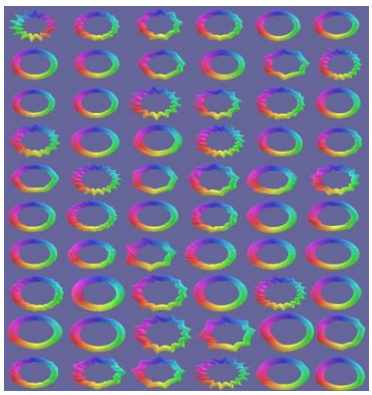
Figure 1(a) shows the process ow for data acquisition. At the beginning of the data acquisition process, a blank screen of 3s is displayed as a resting state in order to avoid any brain activities related to the previous trial. This is followed by 5 to 15s of viewing the 3D stimuli with the minimum and maximum time of 5s and 15s respectively. After the minimum time, the subject is able to proceed to the rating state based on their free will while at the maximum time, the system will automatically proceed to the rating state. The purpose of enabling the subject to decide on their viewing time for the stimuli is to avoid the subject from becoming bored and fatigued during the data acquisition process. Asking the user to view repetitively at fixed intervals and rate the stimuli could lead boredom [21] which may then further cause fatigue [22]. Hence, allowing the user to have the option to move to the next stimuli could save time and avoid fatigue since the subject no longer needs to wait until the maximum time in order to conduct the rating and moving on to the next stimuli. At the end of the shape viewing process, a rating with a scale from 1-5 (1: like very much, 2: like, 3: undecided, 4: do not like, 5: do not like at all) is displayed to the subject, which is adopted from related studies [9, 11].

A 3D shape in the form of a bracelet-like object generated

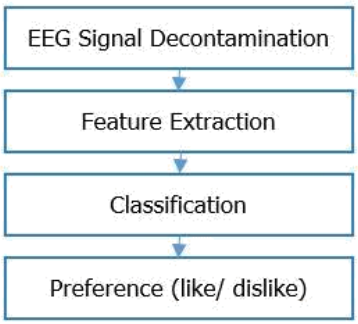
using the Gielis Superformula [23] as described in Equation 1 was used as the stimuli in this study. Our motivation for using this shape as the 3D visual stimuli is to assess the aesthetics of jewellery-type objects since visual aesthetics is the key motivating factor for the decision to purchase such items.

Modification of the parameters of the superformula allows the generation of various natural and elegant-looking 3D shapes. 60 bracelet models as shown in Figure 2(a) were generated by assigning the various superformula parameters with randomly generated values. Various ranges were selected for the different parameters in order to be able to generate shapes that appear closest to a bracelet-like object. These bracelet-like 3D shapes were then displayed virtually on a computer and with rotations on different axes of the presented stimuli so that it could be viewed at different angles in order for the subject to fully visualize the generated 3D bracelet-like object.

In the data processing stage, the collected signals are processed according to the steps as shown in Figure 2(b) to conduct the preference classification. Firstly, the EEG signals are decontaminated from environmental and physiological artifacts. These artifacts are removed automatically using the SDK provided by ABM for the B-Alert X10 headset in MATLAB. Environmental artifacts are removed by applying a 50 Hz notch filter. 5 different physiological artifacts (electromyography (EMG), eye blinks, excursions, saturations, and spikes) are also automatically removed in real-time, where the excursions, saturations, and spikes are replaced by zero values. Spline interpolation was later used to fill in the zero values generated by the automatic artifact removal system.



(a)



(b)

Figure 2: (a) 60 bracelet-like shapes generated using the Gielis superformula; (b) Process flow of signal processing.

The cleaned EEG signals are then transformed into the TF domain using an STFT as per the method suggested by [9, 11]. The STFT process decomposes each of the 9 acquisition channels into five bands: delta (1-3Hz), theta (4-6 Hz), alpha (7-12 Hz), beta (13-30 Hz) and gamma (31-64Hz); thereby giving a total of 45 input features. From the 960 observations obtained during the data acquisition of 16 subjects each viewing 60 of the 3D visual stimuli, only 208 observations which recorded the strongest ratings on each end of the preference spectrum scale were used for the classification process, which were the ratings of 1 (like very much) and 5 (do not like at all). Therefore a matrix of 47 columns consisting of the observation ID, rating, and each of the 45 input TF features, and 208 rows of observations served as the input to the respective classifiers. Additionally, all the subjects' baseline readings obtained during the resting state were first subtracted from the viewing state values prior to the classification process.

IV. Experimental Results & Discussion

A. 10-subject EEG Preference Classification

In this first study, a cohort size of 10 was used. Here only the k-nearest neighbor (kNN) classifier was used to classify 2 preference classes and the nearest neighbor parameter (k) from 1 to 9 was tested. In conjunction with the kNN approach, two different power spectral density (PSD) signal transforms were used here in this first study, which were the Burg and Welch methods.

A brute-force search is applied to the classifier to test all the combinations of the obtained features. However, to search through all the combinations would require a long period of time to complete, so all the combinations of the features are tested for 1 to 5 features and Table I shows the accuracies at 68.67% and above on either method of PSD. The features in Table II with occurrence higher than 1 are used to train and test the classifier.

The accuracy of different combinations on features is as shown below, where the accuracy of 71% and above using Burg method is as shown in Table 2 while Welch method is shown in Table 3. The highest accuracy for Burg methods is 73% with nearest neighbor set to 1 along with features Czgamma, C3beta, C3gamma, C4gamma, C4 theta, F3beta, F3gamma, F3deltam F4beta and P3alpha whereas the highest accuracy for the Welch methods is 74% with nearest neighbor set to 1 along with features Fzdelta, C3beta, C3gamma, C4gamma, C4theta, F3beta, F3delta, F4alpha, F4beta, F4gamma, and P3alpha, also with Fzdelta, Czgamma, C3beta, C3gamma, C4gamma, C4theta, F3beta, F3delta, F4alpha, F4beta, F4gamma and P3alpha.

Table 4 shows the confusion matrix for the best accuracy (74%) obtained. Through the use of features Fzdelta, C3beta, C3gamma, C4gamma, C4theta, F3beta, F3delta, F4alpha, F4beta, F4gamma, and P3alpha, the true positive (TP), true negative (TN), false positive (FP) and false negative (FN) were 52, 59, 24 and 15 respectively. Meanwhile for the features with Fzdelta, Czgamma, C3beta, C3gamme, C4gamma, C4theta, F3beta, F3delta, F4alpha, F4beta, F4gamma and P3alpha, the TP, TN, FP and FN were 53, 58, 24 and 14 respectively. The total testing data were 150. The confusion matrix for both combinations of features which obtained 74% accuracy are similar whereby there is only one feature difference between both combination of features.

Table 1
Classification accuracy obtained using brute force search based on 1-5 features combination

Combination of features	Accuracy
Fzbeta,C3beta,C4gamma,P3alpha Czgamma,C3beta,F3delta,P3delta POztheta,Fzbeta,C4gamma,F3delta,P3delta POzdelta,C3beta,F4alpha,F4beta,F4gamma Fzalpha,Czalpha,F3beta,F3delta,F4beta Fzbeta,C4gamma,C4theta,F3delta,F4alpha Fzgamma,Fzdelta,C3alpha,C3beta,C4theta Fztheta,Fzdelta,F3beta,F4alpha,F4beta C4gamma,F3gamma,F3delta,F4gamma	68.67%
C4gamma,F3delta,F4gamma,P3theta POztheta,Czbeta,C4gamma,F3delta,P3delta POztheta,C3gamma,C3theta,F3delta,F4alpha Fzbeta,Czgamma,C3beta,C4gamma,P3alpha Fzbeta,C3beta,C4gamma,P3alpha,P4gamma Fzdelta,C4gamma,F4alpha,F4gamma,F4theta Fzdelta,F3beta,F4alpha,F4beta,P3alpha Czalpha,C3gamma,C4theta,F3delta,F4gamma Cztheta,C4gamma,F4gamma,P3alpha,P3beta	69.33%
Fztheta,Fzdelta,F3beta,F4alpha,F4beta	70.00%
POztheta,Fzbeta,C3gamma,F3delta,P3delta	70.67%
C3gamma,C4gamma,F3gamma,F3delta,F4gamma	72.67%

Table 2
Accuracy of different combination of features using the burg method

Features	k	Accuracy
Czγ,C3β,C3γ,C4γ,C4θ,F3β,F3γ,F3δ,F4β,P3α	2	73**%**

α indicates alpha, β indicates beta, θ indicates theta, γ indicates gamma, and δ indicates delta rhythm.

Table 3
Accuracy of different combination of features using Welch method

Features	k	Acc.
Fzδ,C3β,C3γ,C4γ,C4θ,F3β,F3δ,F4α,F4β,F4γ,P3α	1	74%
Fzδ,Czγ,C3β,C3γ,C4γ,C4θ,F3β,F3δ,F4α,F4β,F4γ,P3α	1	74%
Fzδ,Czγ,C3β,C3γ,C4γ,C4θ,F3β,F3δ,F4α,F4β,P3α	1	73%
Fzδ,Czγ,C3β,C4γ,C4θ,F3β,F3δ,F4α,F4β,P3α	1	73%
Fzδ,C3β,C3γ,C4γ,C4θ,F3β,F3δ,F4α,F4β,P3α	1	73%
Fzδ,C3β,C4γ,C4θ,F3β,F3δ,F4α,F4β,F4γ,P3α	1	73%
Fzβ,Fzδ,C3β,C4γ,C4θ,F3β,F3δ,F4α,F4β,F4γ,P3α	1	73%
Fzδ,C3β,C3γ,C4γ,C4θ,F3β,F3γ,F3δ,F4α,F4β,P3α	1	73%
Fzβ,Fzδ,Czγ,C3β,C3γ,C4γ,C4θ,F3β,F3δ,F4α,F4β,F4γ,P3α	1	73%
Fzβ,Fzδ,C3β,C4γ,C4θ,F3β,F3δ,F4α,F4β,P3α	1	72%
Fzβ,Fzδ,C3β,C4θ,F3β,F3γ,F3δ,F4α,F4β,P3α	2	72%
Fzβ,Fzδ,C3β,C3γ,C4γ,C4θ,F3β,F3δ,F4α,F4β,P3α	1	72%
Fzβ,Fzδ,C3γ,C4γ,C4θ,F3β,F3γ,F3δ,F4α,F4β,P3α	1,2	71%
Fzδ,Czγ,C3β,C3γ,C4γ,F3β,F3γ,F3δ,F4α,F4β,P3α	2	72%
Fzδ,Czγ,C3β,C4γ,C4θ,F3β,F3γ,F3δ,F4α,F4β,P3α	2	72%
Fzδ,C3γ,C4γ,C4θ,F3β,F3γ,F3δ,F4α,F4β,F4γ,P3α	1	72%
Fzβ,Fzδ,Czγ,C3β,C3γ,C4γ,C4θ,F3β,F3δ,F4α,F4β,P3α	1	72%
Fzβ,Fzδ,Czγ,C3β,C4γ,C4θ,F3β,F3γ,F3δ,F4α,F4β,P3α	2	72%
Fzδ,Czγ,C3β,C3γ,C4γ,C4θ,F3β,F3γ,F3δ,F4α,F4β,P3α	1	72%
Fzβ,Fzδ,C3β,C3γ,C4γ,C4θ,F3β,F3γ,F3δ,F4α,F4β,F4γ,P3α	1	72%
Fzθ,Fzδ,F3β,F4α,F4β,F4γ	6	71%
POzθ,Fzβ,C3β,C3γ,F3δ,F4γ,P3δ	1	71%
Fzβ,Fzδ,C3β,C4γ,C4θ,F3β,F4α,F4β,P3α	2	71%
Fzβ,Fzδ,C4γ,C4θ,F3β,F3δ,F4α,F4β,P3α	1	71%
Fzδ,C3β,C3γ,C4θ,F3β,F3δ,F4α,F4β,P3α	1,2	71%
Fzδ,C3β,C4γ,C4θ,F3β,F3δ,F4α,F4β,P3α	1,2	71%
Fzβ,Fzδ,C3γ,C4θ,F3β,F3γ,F3δ,F4α,F4β,P3α	2	71%
Fzδ,C3β,C3γ,C4γ,F3β,F3γ,F3δ,F4α,F4β,P3α	1	71%
Fzδ,C3β,C4γ,C4θ,F3β,F3γ,F3δ,F4α,F4β,P3α	1	71%
Fzδ,C3β,C4θ,F3β,F3γ,F3δ,F4α,F4β,F4γ,P3α	2	71%
Fzβ,Fzδ,Czγ,C3β,C4θ,F3β,F3γ,F3δ,F4α,F4β,P3α	2	71%
Fzβ,Fzδ,C3β,C4γ,C4θ,F3β,F3γ,F3δ,F4α,F4β,P3α	2	71%
Fzδ,Czγ,C3β,C4γ,C4θ,F3β,F3δ,F4α,F4β,F4γ,P3α	1	71%
Fzδ,C3β,C4γ,C4θ,F3β,F3γ,F3δ,F4α,F4β,F4γ,P3α	1	71%
Fzβ,Fzδ,Czγ,C3β,C4γ,C4θ,F3β,F3δ,F4α,F4β,F4γ,P3α	1	71%
Fzβ,Fzδ,C3β,C3γ,C4γ,C4θ,F3β,F3γ,F3δ,F4α,F4β,P3α	1,2	71%
Fzβ,Fzδ,C3β,C3γ,C4γ,C4θ,F3β,F3δ,F4α,F4β,F4γ,P3α	1	71%
Fzδ,C3β,C3γ,C4γ,C4θ,F3β,F3γ,F3δ,F4α,F4β,F4γ,P3α	1	71%
Fzβ,Fzδ,Czγ,C3β,C3γ,C4γ,C4θ,F3β,F3γ,F3δ,F4α,F4β,F4γ,P3α	1,2	71%
Fzδ,Czγ,C3β,C3γ,C4γ,C4θ,F3β,F3γ,F3δ,F4α,F4β,F4γ,P3α	1	71%
Fzβ,Fzδ,Czγ,C3β,C3γ,C4γ,C4θ,F3β,F3γ,F3δ,F4α,F4β,F4γ,P3α	1	71%

α indicates alpha, β indicates beta, θ indicates theta, γ indicates gamma, and δ indicates delta rhythm.

Table 4
Confusion matrix for best accuracy

Fzδ, C3β, C3γ, C4γ, C4θ, F3β, F3δ, F4α, F4β, F4γ, P3α	Actual/Predicted	Like	Dislike
	Like	52	15
	Dislike	24	59
Fzδ, Czγ, C3β, C3γ, C4γ, C4θ, F3β, F3δ, F4α, F4β, F4γ, P3α	Actual/Predicted	Like	Dislike
	Like	53	14
	Dislike	25	58

α indicates alpha, β indicates beta, θ indicates theta, γ indicates gamma, and δ indicates delta rhythm.

Moreover, the Welch method performed better than the Burg method especially using a nearest neighbor of one and two. Meanwhile, the features in the combination obtained accuracy of 71% and above are mostly from channels Fz, C3, C4, F3, F4 and P3.

Left parietal (P3) is believed to be involved in mental rotation, which corresponds to the subjects mentally observing the rotating objects being shown on the screen. Meanwhile, the left and right central strip (C3 and C4) are believed to be active during visualization of movements. Additionally, the frontal lobe (Fz, F3 and F4) is believed to be involved in decision making, memory processing and attention. Dorsolateral prefrontal cortex (F3) in the frontal lobe is also related to aesthetic perception.

B. 16-subject EEG Preference Classification

In this second study, a number of different classifiers were used and as the results show, most of the classifiers performed poorly due to inter-subject variability for the EEG-based preference classification. As such, we proceeded to include deep learning as a potential approach to overcome these poor learning outcomes from the commonly found classifiers.

Ten-fold cross-validation was used to test the different algorithms for the classification task. Deep neural networks were initialized with 200 hidden neurons with 2 hidden layers using the uniform adaptive method [24] for the initialization of the weight matrix. The rectified linear activation function [25] was used with an adaptive learning rate method [26]. The deep neural networks were run for 10 epochs using cross-entropy as the error function. The deep neural networks were tested against 10 state-of-the-art classifiers provided in the R package known as "caret" comprising support vector machines, random forests, Naïve Bayes, various decision tree as well as k-nearest neighbor classification models. The default settings for the parameters that were used are as listed in the parentheses of the corresponding classifiers.

Table 5 shows the results of the experimental runs. Based on 10-fold cross-validation, only two classifiers were able to obtain classification rates of above 60%, with the deep neural networks performing the best at almost 64% while the linear SVM was only able to achieve 60.19% accuracy. The other 9 classifiers were only able to achieve classification rates of between 56-59% with the radial SVM achieving 59.67% and k-nearest neighbor performing the worst at 56.29%.

As can be seen clearly from this experiment, with a larger cohort size of 16 subjects, the inter- and intra-subject variability in terms of EEG-based preference classification is indeed very significant and resulting in low classification accuracy rates. Also shown here is the promise of deep neural networks as machine learning algorithms that are able to perform better than any of the other 10 commonly used classifiers in this highly challenging task.

In conclusion, we have shown through this systematic and

comprehensive empirical comparison, deep neural networks are a highly promising approach in terms of dealing with large-scale EEG datasets that comprise of significant noise arising from variations between and within subjects. Further investigations into the different deep learning neural network activation functions and their corresponding architectures would be beneficial in further improving the performance of deep learning in EEG-based applications such as preference classification.

Table 5
Classification results for preference classification using ten-fold cross-validation

Classifier	Accuracy
Deep Net	63.99%
SVM Linear (C = 1)	60.19%
SVM Radial (sigma = 0.04, C = 1)	59.67%
OneR	59.00%
Adaboost	58.65%
Random Forest (mtry = 14)	57.74%
NNet (decay = 0, size = 3)	57.71%
JRip (NumOpt = 2)	57.21%
Naive Bayes (usekernel = true, fl = 0)	56.79%
C5.0 (type = tree, winnow = true, trials = 1)	56.74%
kNN (k = 5)	56.29%

V. CONCLUSION

In conclusion, we have shown through this systematic and comprehensive empirical comparison; deep neural networks are a highly promising approach in terms of dealing with large-scale EEG datasets that comprise of significant noise arising from variations between and within subjects. An initial study using kNN provided sufficiently good results in a 10-subject study. However, when expanded to a larger cohort size of 16 subjects, the results were not encouraging. However, the use of deep learning was able to observably overcome some of the difficulties presented by inter-subject variability posed by larger cohort sizes in EEG-based preference classification. Further investigations into the different deep learning neural network activation functions and their corresponding architectures would be beneficial in further improving the performance of deep learning in EEG-based applications such as preference classification.

ACKNOWLEDGEMENT

This project is supported by the FRGS research grant scheme refs: FRG0349 & FRG0435 from the Ministry of Higher Education, Malaysia.

REFERENCES

[1] L. H. Chew, J. Teo, J. Mountstephens, Aesthetic preference recognition of 3D shapes using EEG, Cognitive neurodynamics 10 (2) (2016) 165-173.
[2] D. Denton, The primordial emotions: The dawning of consciousness.
[3] N. Watanabe, M. Haruno, Effects of subconscious and conscious emotions on human cue-reward association learning, Scientific reports 5.
[4] X.-W. Wang, D. Nie, B.-L. Lu, Emotional state classification from EEG data using machine learning approach, Neurocomputing 129 (2014) 94-106.
[5] M. Murugappan, N. Ramachandran, Y. Sazali, et al., Classification of human emotion from EEG using discrete wavelet transform, Journal of Biomedical Science and Engineering 3 (04) (2010) 390.
[6] M.-K. Kim, M. Kim, E. Oh, S.-P. Kim, A review on the computational methods for emotional state estimation from the human EEG, Computational and mathematical methods in medicine 2013.
[7] Y. Liu, O. Sourina, M. K. Nguyen, Real-time EEG-based human emotion recognition and visualization, in: Cyberworlds (CW), 2010 International Conference on, IEEE, 2010, pp. 262-269.
[8] M. Li, B.-L. Lu, Emotion classification based on gamma-band EEG, in: 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2009, pp. 1223-1226.
[9] S. K. Hadjidimitriou, L. J. Hadjileontiadis, Toward an EEG-based recognition of music liking using time-frequency analysis, IEEE Transactions on Biomedical Engineering 59 (12) (2012) 3498-3510.
[10] Y. Pan, C. Guan, J. Yu, K. K. Ang, T. E. Chan, Common frequency pattern for music preference identification using frontal EEG, in: Neural Engineering (NER), 2013 6th International IEEE/EMBS Conference on, IEEE, 2013, pp. 505-508.
[11] S. K. Hadjidimitriou, L. J. Hadjileontiadis, EEG-based classification of music appraisal responses using time-frequency analysis and familiarity ratings, IEEE Transactions on Affective Computing 4 (2) (2013) 161-172.
[12] K. C. Tseng, B.-S. Lin, C.-M. Han, P.-S. Wang, Emotion recognition of EEG underlying favourite music by support vector machine, in: Orange Technologies (ICOT), 2013 International Conference on, IEEE, 2013, pp. 155-158.
[13] A. Yazdani, J. S. Lee, J.-M. Vesin, T. Ebrahimi, A ECT recognition based on physiological changes during the watching of music video, ACM Transactions on Interactive Intelligent Systems 2 (EPFL-ARTICLE-177741) (2012) 1-26.
[14] J. Moon, Y. Kim, H. Lee, C. Bae, W. C. Yoon, Extraction of user preference for video stimuli using EEG-based user responses, ETRI Journal 35 (6) (2013) 1105-1114.
[15] Y. Kim, K. Kang, H. Lee, C. Bae, Preference measurement using user response electroencephalogram, in: Computer Science and its Applications, Springer, 2015, pp. 1315-1324.
[16] S. Goncalves, J. De Munck, P. Pouwels, R. Schoonhoven, J. Kuijer, N. Maurits, J. Hoogduin, E. Van Someren, R. Heethaar, F. L. Da Silva, Correlating the alpha rhythm to bold using simultaneous EEG/FMRI: inter-subject variability, Neuroimage 30 (1) (2006) 203-213.
[17] G. Pfurtscheller, C. Brunner, A. Schlogl, F. L. Da Silva, Mu rhythm (de) synchronization and EEG single-trial classification of different motor imagery tasks, Neuroimage 31 (1) (2006) 153-159.
[18] K. Li, X. Li, Y. Zhang, A. Zhang, Affective state recognition from EEG with deep belief networks, in: Bioinformatics and Biomedicine (BIBM), 2013 IEEE International Conference on, IEEE, 2013, pp. 305-310.
[19] W.-L. Zheng, J.-Y. Zhu, Y. Peng, B.-L. Lu, EEG-based emotion classification using deep belief networks, in: 2014 IEEE International Conference on Multimedia and Expo (ICME), IEEE, 2014, pp. 1-6.
[20] S. Jirayucharoensak, S. Pan-Ngum, P. Israsena, EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation, The Scientific World Journal 2014.
[21] V. Shackleton, Boredom and repetitive work: a review, Personnel Review 10 (4) (1981) 30-36.
[22] A. Craig, Y. Tran, N. Wijesuriya, H. Nguyen, Regional brain wave activity changes associated with fatigue, Psychophysiology 49 (4) (2012) 574-582.
[23] J. Gielis, A generic geometric transformation that unifies a wide range of natural and abstract shapes, American journal of botany 90 (3) (2003) 333-338.
[24] D. Nguyen, B. Widrow, Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights, in: Neural Networks, 1990. 1990 IJCNN International Joint Conference on, IEEE, 1990, pp. 21-26.
[25] V. Nair, G. E. Hinton, Rectified linear units improve restricted Boltzmann machines, in: Proceedings of the 27th International Conference on Machine Learning (ICML-10), 2010, pp. 807-814.
[26] M. D. Zeiler, Adadelta: an adaptive learning rate method, arXiv preprint arXiv:1212.5701.