

Decoding Brain Patterns for Colored and Grayscale Images using Multivariate Pattern Analysis

Raheel Zafar^{1,*}, Muhammad Noman Malik², Huma Hayat¹ and Aamir Saeed Malik³

¹ Department of Engineering

*Corresponding author raheelsatti@gmail.com

² Department of Computer Science

^{1,2} National University of Modern Languages, Islamabad, Pakistan
mnauman@numl.edu.pk

³ Centre for Intelligent Signal & Imaging Research (CISIR),
Universiti Teknologi PETRONAS, Perak, Malaysia

*Corresponding author: Raheel Zafar

*Received May 5, 2019; revised August 10, 2019; accepted January 21, 2020;
published April 30, 2020*

Abstract

Taxonomy of human brain activity is a complicated rather challenging procedure. Due to its multifaceted aspects, including experiment design, stimuli selection and presentation of images other than feature extraction and selection techniques, foster its challenging nature. Although, researchers have focused various methods to create taxonomy of human brain activity, however use of multivariate pattern analysis (MVPA) for image recognition to catalog the human brain activities is scarce. Moreover, experiment design is a complex procedure and selection of image type, color and order is challenging too. Thus, this research bridge the gap by using MVPA to create taxonomy of human brain activity for different categories of images, both colored and gray scale. In this regard, experiment is conducted through EEG testing technique, with feature extraction, selection and classification approaches to collect data from prequalified criteria of 25 graduates of University Technology PETRONAS (UTP). These participants are shown both colored and gray scale images to record accuracy and reaction time. The results showed that colored images produces better end result in terms of accuracy and response time using wavelet transform, t-test and support vector machine. This research resulted that MVPA is a better approach for the analysis of EEG data as more useful information can be extracted from the brain using colored images. This research discusses a detail behavior of human brain based on the color and gray scale images for the specific and unique task. This research contributes to further improve the decoding of human brain with increased accuracy. Besides, such experiment settings can be implemented and contribute to other areas of medical, military, business, lie detection and many others.

Keywords: Multivariate pattern analysis (MVPA), Human Brain, Taxonomy, EEG.

1. Introduction

Decoding of human brain activity patterns is challenging, though emerging area of research due to its enormous range of applications. This is one of the reason many current and previous studies introduced several novel methods and techniques [1-5], to decode/classify the human brain activities. Decoding of human brain activity is the pressing concern due to its various applications from clinical till defense. For example, it is useful in medical science to identifying brain tumor, stroke and dementia [6], to identify the terrorists in the field of defense and security [7], to identify the better product for business commercial purposes [8, 9] and can be used to identify the suspect in courts [10].

Human brain is a complex organ of the body and it is difficult to study the pattern of human brain activities. That's why every step of analysis has its own importance since the beginning e.g. experiment design and preprocessing till the end i.e. statistical analysis. Each part of the analysis can contribute in the extraction of brain information; that's why several methods are used at each level in current studies. In this regard, several studies intended electroencephalogram (EEG) data analysis along with various techniques for feature extraction, including Wavelet Transform (WT), Principal Component Analysis (PCA), Independent Component Analysis (ICA) and many others [11-14].

Plethora of research is conducted to collect and analyzed the EEG data. Since different steps are involved in the classification of brain data e.g. experiment design, preprocessing, feature extraction, feature selection and classification; so in each study specific problem is considered. The studies have been focusing to improve the results at that specific step. Different techniques, have been used in feature extraction and selection procedure. Some studies have used experimental design in which several colored and grayscale images are discussed along with the response time. The grayscale images usually have less computational requirements and are simple in vision, on the other hand color images have additional color impact during the image presentation [15]. Moreover, color and contrast has an influenced during image recognition [16] that can affect largely on the brain activities. Response time or appearance on the screen with colored or grayscale images are also highlighted as an important factor in recent studies [2, 17]. Besides, it is also notable from literature that studies have used different classification methods like Nearest Neighbor, Logistic Regression, Naïve Bayes Classifier, Support vector machine and many others [18].

Recently, color vision is considered another important aspect and connected with our emotions [19]. Author of that study provides diagnostic cues using natural scene to inform a dangerous situation. Moreover, response of brain is also found for colored and grayscale images using EEG in the above discussed study. They found heart rate, neural activity and retinal focus for danger vs non-threatening images, however their hypothesis was restricted to emotions and not to find the difference in brain behavior for color vs grayscale images.

Based on above, it can be argued that it is difficult to extract maximum information from the brain, and machines are limited with their functionalities to measure all the information of the brain. The measured brain information is only the part of total activations occurred in the brain and are helpful in different field especially in medical and research. As there is a

hardware restriction in measuring the brain activities, so the particular focus is on analysis and experiment designs. Since the primary objective of neuroscientist is to extract maximum and affective information from the brain in minimum time; therefore several experiment designs are already used along with pre-processing and statistical techniques. These experiment designs have different parameters and image types like response time, response type and image type e.g. colored or grayscale. A good experiment design is always considered better along with appropriate hardware, efficient pre-processing techniques and in-depth statistical analysis. In this regard different preprocessing approaches [20]. Moreover, analysis techniques [21, 22] are useful and have been used in existing studies along with color/grayscale images which has higher or lower impact in object recognition [23, 24]. All the steps and the corresponding techniques are important during the analysis phase and have a strong combined effect in improving the decoding accuracy.

In this paper, the primary focus is to investigate that how can the decoding accuracy be increased in terms of experiment design and analysis. The behavior of brain against different colored and grayscale images are measured and analyzed comprehensively. The motivation of this paper is to use Multivariate Pattern Analysis (MVPA) approach for both data sets and compared the results of different techniques and experiments i.e. colored and grayscale images. In addition to the vast applications of brain decoding, this research could eventually help industry software designers, graphical engineers and others to foresee that how much mix of colors they can used to get a positive response from human brain in their respective context. Presently, MVPA is a popular and effective technique, thus in this study a comprehensive comparison of colored and grayscale images is conducted based on MVPA to get better output in both cases. This research contributes to explore the patterns to classify the human brain, that how brain reacts against various objects in terms of accuracy, relevancy and with color behavior on their identification.

2. Materials and Methods

In this study, experiment methodology is conducted because of its comprehensive characteristics and its usage earlier studies [17, 25, 26]. Experiment methodology includes experiment design, data collection, analysis and the results. The detail of each experimental module is explained in the following sections.

2.1 Participants

Initially, participant selection was done carefully by considering the factors of age, color blindness. A total of thirty participants were selected initially for this experiment which were graduate student of University Technology PETRONAS (UTP). It was assured that age of participant must fall 25 to 35 years. Thus, average age of the subjects was 31 years. The written consent form was collected from each subject before the start of the experiment.

Since, the task includes both colored and grayscale sessions so it was carefully seen that none of the subject should be color blind. Among 30 participants, 3 participants did not reply for right/wrong category on time for most of the images and two participants were found irrelevant in their results. Hence by discarding the five participants, the final analysis was

done with the data of 25 subjects. The experiment and participant selection criterion was approved by the ethics committee of UTP and the data was also collected at UTP.

2.2 Materials

Providing images as subject material in this experiment is crucial. In this regard, a total of 210 images were presented in which half were target and the rest were non-target images. The target images were from 7 categories means each had 30 images with 15 target and 15 non-target images. Besides, two different sessions were conducted, one for colored images and other for grayscale images. All the images were taken from internet in colored form and grayscale version was changed in MATLAB because according to one of the study, conversion of color images to grayscale has little impact on recognition performance [27]. These images were shown to participants on computer one by one. It was assured that picture quality must be standardized and should not have any negative impact in identification and understanding.

2.3 Experimental Conduction

2.3.1 Briefing: Before the experiment conduction, participants were briefed about the procedure such as participants were informed that both colored and grayscale tasks were response based in which they have to respond for right and wrong category by pressing the button from the keyboard. They should reply as soon as possible with maximum time of 800 milliseconds. After that there was gap of 1 sec which had been used as baseline during analysis.

2.3.2 Conduction: The experiment paradigm is shown in Fig. 1 for colored images; whereas grayscale images session also have the same experiment with grayscale images. Since it takes 200ms to reach information from eye to visual cortex [15], so 800ms time gives the earliest response of the participant and it will give the better picture of the brain. The parameters used in this study like response time, number of images and the rest time are mostly taken from existing studies [2, 14]. During the experiment, the name of the category was appeared on the screen before the start of the category like word “animal” appeared on the screen. After that, the subject must reply for right or wrong category by pressing 1 or 2 from the keyboard which mentioned that whether the animal picture is on the screen or not. For half of the participants, initially colored task was performed followed by the gray scale and for rest of the participants the order of task was reversed.

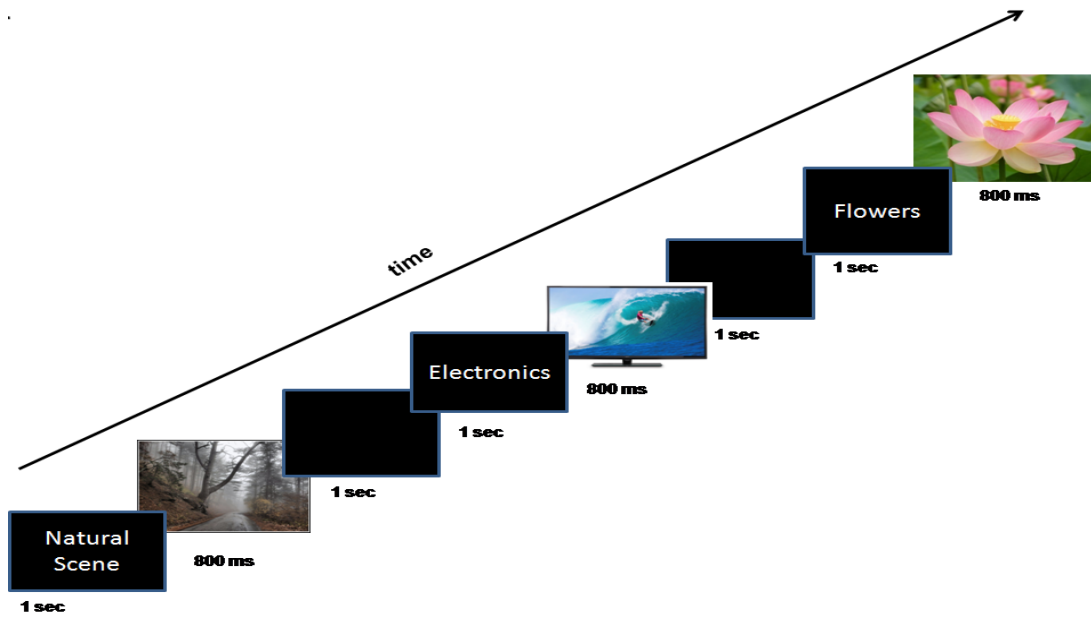


Fig. 1. EEG experiment design

2.4 Data Analysis

After collecting the data the first step is the pre-processing of data. The pre-processing is followed by the feature extraction, feature selection and classification. In current study, features of EEG data are extracted and selected using wavelet transform (WT) and t-test simultaneously against several color and grayscale images. Classification of independent subjects is done for both colored and grayscale images against the baseline. It is also observed that either user response is better with color images or gray scale. Lastly, response time is also measured for both type of images. The stimulus set has seven categories including human, natural scenes, jewelry, fruit, transport, electronic devices and flower which are resembling with daily routine images. The analysis is done using various wavelet transform techniques like Haar, Symlet2 and Daubechies4 (Db4), but Db4 is finally taken for analysis due to better accuracy. The experiment has two different sessions, one with colored images and the other with gray-scale images; while the data is taken from 30 same subjects for both sessions.

2.4.1 Pre-processing of data

The continuous EEG data was recorded with Electrical Geodesics Incorporated (EGI) system having 128 channels with a sampling frequency of 250 Hz with impedance of 50 K Ω . The data was recorded using Net Station software [28].

The preprocessing of EEG data was done using Brain Electrical Source Analysis (BESA) in which the signal was filtered between 0.3 to 30 Hz using a bandpass filter as all the relevant frequencies like alpha, beta, delta and theta lies in this range. Finally the pre-processed data was exported to MATLAB for further analysis. In MATLAB, data against

individual image was analyzed separately and no response or data was excluded for fair analysis.

Analyzing data for better classification results is considered on feature extraction, selection of key features and finally classification of significant data according to the particular class. During the final analysis, data of 25 subjects were used to make the analysis reliable.

2.4.2 Wavelet Transform

In Continuous wavelet transform (CWT), the continuous signals can be transformed into another resulting signal with two variables; translation and scaling. This transformed signal is easy to interpret and significant for time frequency analysis. The CWT of continuous function $x(t)$ is described as

$$W_{\psi}(s, \tau) = \int_{-\alpha}^{+\alpha} x(t)\psi_{s,\tau}^*(t)dt \quad (1)$$

$$\text{Where } \psi_{s,\tau}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-\tau}{s}\right) \quad (2)$$

s and τ are known as scaling and translation parameters respectively. $W_{\psi}(s, \tau)$ represents the wavelet transform coefficients and Ψ is the basic mother wavelet.

In Discrete wavelet transform (DWT), the same phenomenon is followed as in CWT but the only difference is that in DWT the scaling and position values based on the power of 2. In this way, the values of s become $s = 2^j$ and the value of $\tau = k * 2^j$ where $(j, k) \in \mathbb{Z}^2$, so the above equation (2) can be written as

$$\psi_{s,t}(t) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{t-k*2^j}{2^j}\right) \quad (3) \quad [29]$$

2.4.3 Feature Extraction

After preprocessing of EEG data, the next step is to extract the features. Three type of features were taken for analysis, raw EEG data, sample entropy and wavelet transform. Sample entropy is a modification of approximate entropy and can calculate the time series signals [30, 31]. It is already utilized in some studies to investigate the signals especially the dynamics of heart rate [32]. In existing studies, many EEG features are described including the above one; and few of them comprehensively discussed and used wavelet transform (WT) [2, 33]. Raw data is also taken as feature in couple of studies and compared with other features like Short time Fourier transform (STFT), WT and many others [14], however in recent studies discrete wavelet transform (DWT) is considered as best among other techniques [2, 34]. In the above section, wavelet transform is already explained mathematically and the difference of DWT and CWT is discussed too. The DWT has become a powerful technique in biomedical signal processing and used often as already discussed. In WT, the signal is decomposed into different frequency components by scaling and shifting. In DWT, the wavelets are sampled in discrete form and the signal is passed through a series of high and low pass filters to analyze the frequencies for the purpose of decomposition and approximation details of the signal [35]. The number of decomposed features from a single image using DWT is mentioned in [Table 1](#).

Table 1. Number of decomposed features from a single image using DWT.

Wavelet	Db4
D1	128 × 128
D2	128 × 67
D3	128 × 37
D4	128 × 22
D5	128 × 14
A5	128 × 14
Total	128 × 282

In this study, MVPA is used to extract the information from distributed brain regions using EEG, as all classification techniques, particularly decoding are based on multivariate analysis. In MVPA, patterns of neural activity are associated with various cognitive states and particular information in a region can be investigated [1, 26, 36, 37]. The analysis was done based on individual subjects and the data of every channel was arranged in the row vector against each image. The experiment was response based, and the EEG data against every response (both right and wrong) was taken for further analysis along with the corresponding baseline (resting time between two stimuli when no stimulus was presented).

2.4.4 Feature Selection

For feature selection, two methods were compared, i.e. t-test and Bhattacharyya distance. T-test is normally used to determine if two population means are equal. A common application is to test if a new process or treatment is superior to a current process or treatment [38]. A t-test is a statistical hypothesis test in which the test statistic follows a student t-distribution. The Bhattacharyya distance measures the similarity of two probability distributions and is related to Bhattacharyya coefficient occurred due to the overlap between two statistical populations or samples [39]. In this analysis, the t-test was selected for the selection of features and significant features with $p < 0.05$ were taken for further analysis. Different number of features was taken for analysis i.e. from 50 to 1000, and among them the best result were considered only.

2.4.5 Classification and Cross Validation

After selection of significant features, Support vector machine (SVM) and linear discriminant analysis (LDA) were used for the purpose of classification. Before classification, the whole data was normalized with z-score and finally the statistical test was done using cross validation. In this study, Mote-Carlo cross validation was applied in which 90% data was used for training purpose and rest of the data for the purpose of testing. The data was shuffled 100 times during cross validation and the average was taken for final results. In addition to the accuracy for examining for colored and gray scale images, response time for colored and gray scale images were also observed along with the correct responses for both type of images. The detail of experimental results are reported in the next section.

2.4.6 Confusion Matrix

A confusion matrix is used to describe the performance of the prediction model on a set of test vector for which the true values are known. The confusion matrix is explained in **Table 2** which includes true positive, true negative, false positive (type I error) and false negative (type II error) while the equations of sensitivity, specificity and accuracy are discussed in equations 4, 5 and 6 respectively. Sensitivity, specificity and accuracy are described in terms of True positive (TP), True negative (TN), False positive (FP) and False negative (FN).

Table 2. Table of confusion matrix

	True condition	
Predicted condition	True positive (TP)	False positive (FP) Type I error
	False negative (FN) Type II error	True negative (TN)

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (4)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (5)$$

$$\text{Accuracy} = \frac{(TN+TP)}{(TN+TP+FN+FP)} \quad (6)$$

3. Results

The results are discussed comprehensively in this section.

3.1 Results with specific brain region (Univariate analysis)

Univariate analysis is a simplest and traditional way to analyze the brain data. It is still used in recent studies especially during region of interest analysis (ROI) and for the purpose of comparison [5]. In this study univariate analysis is done only for comparing results with Multivariate pattern analysis (MVPA) and data from specific brain regions were taken for the analysis. Two brain regions were used for ROI analysis; occipital and temporal as the task was a visual task with recognition so these two brain regions have significant information for this particular task. The separate analysis was done for both brain regions

and significant results were determined from both regions i.e. average accuracy in occipital 82.3% for colored and 80.5% for grayscale and in temporal 80.1% in colored and 78.2% in grayscale as mention in **Table 3**. Since both the results are meaningful but lesser than the MVPA in which EEG data was taken from several brain regions.

Table 3. Decoding Result of Univariate analysis

Brain Region	Colored Accuracy	Grayscale Accuracy
Occipital	82.3%	80.5%
Temporal	80.1%	78.2%

3.2 Results with Multivariate pattern analysis (MVPA)

In this subsection, Results are reported using MVPA in which experimental results are described for both data sets i.e. colored and gray scale. Multivariate model is applied during the analysis with the data of 25 subjects. The detailed analysis was done on both the sessions separately i.e. colored and gray scale images. Initially the features were extracted using sample entropy, raw data, Global field Power [40] and discrete wavelet transform [2]. After that features were selected using t-test and Bhattacharya distance, finally the classification was done using SVM and linear discriminant analysis (LDA). Since data analysis consists of various steps so in each step different existing techniques are applied to compare the results with the proposed method which makes the study reliable.

The accuracy, specificity and sensitivity of both data sets are found using each method and the detail comparison of all existing techniques is described in the discussion section. Currently, the best techniques are reported here which includes DWT (for feature extraction), t-test (for feature selection) and SVM (for classification). In **Fig. 2**, the decoding results of both colored and grayscale data sets are compared for each participant along with the overall average of all participants. The specificity and sensitivity of individual participant is shown in **Fig. 3**.

Fig. 2 clearly specified that average accuracy is better in case of colored images; moreover in case of individual subjects accuracy is higher compared to colored images mostly. In **Fig. 2**, Avg is representing the average accuracy of all subjects. Specificity and Sensitivity are statistical measures and are widely used; also known as classification function in statistics. In **Fig. 3**, Sensitivity and Specificity is shown that illustrate higher in most of the subjects for colored images compared to the gray images.

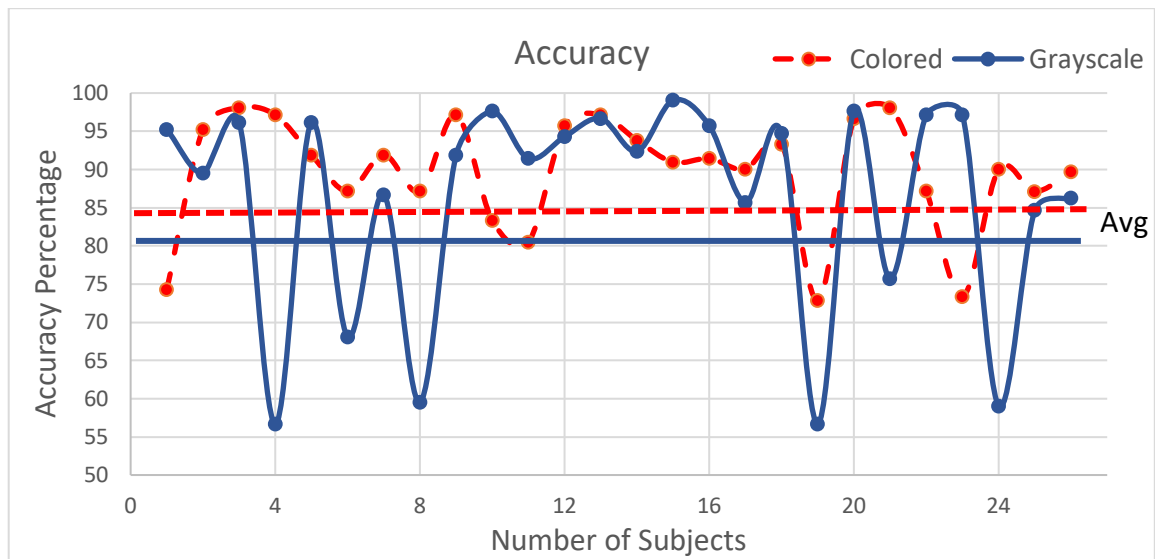


Fig. 2. Classification accuracy for colored and grayscale images

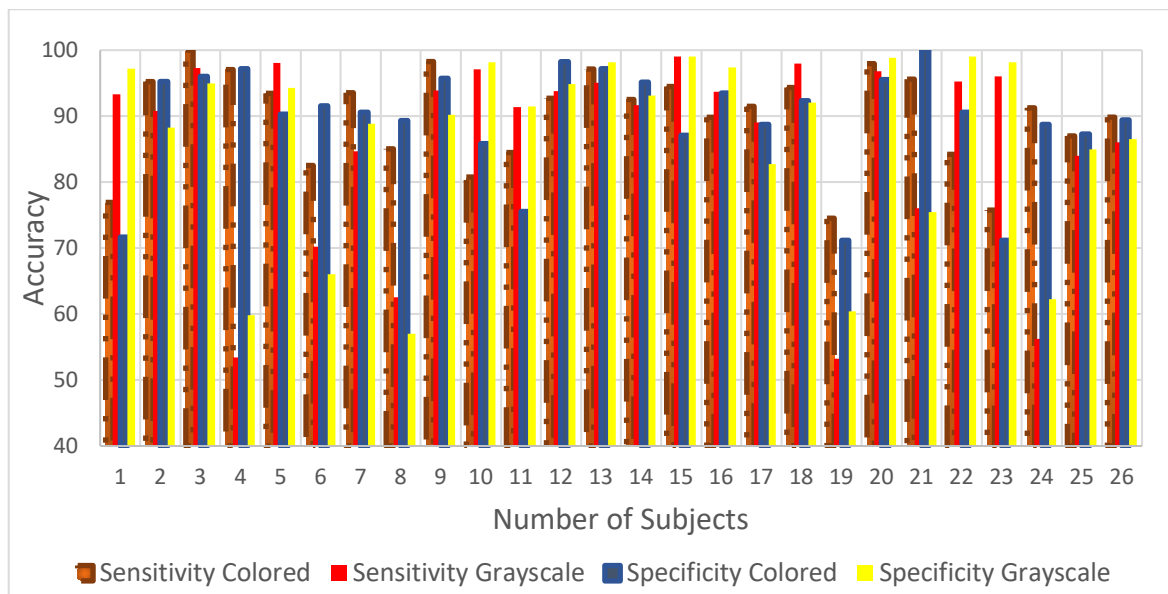


Fig. 3. Sensitivity and Specificity for Colored and Grayscale Images

The number of features plays a vital role in the analysis as it improves the efficiency of the algorithm as well as helps in interpretation of the results. In this regard, the most significant features were taken. Besides, MVPA classification accuracy was computed using different voxels which vary between 50 to 5,000 voxels. In **Fig. 4**, the best accuracy result against the number of features is shown for each individual participant. The most significant features are those taken based on the highest t-values from the GLM of the training data set.

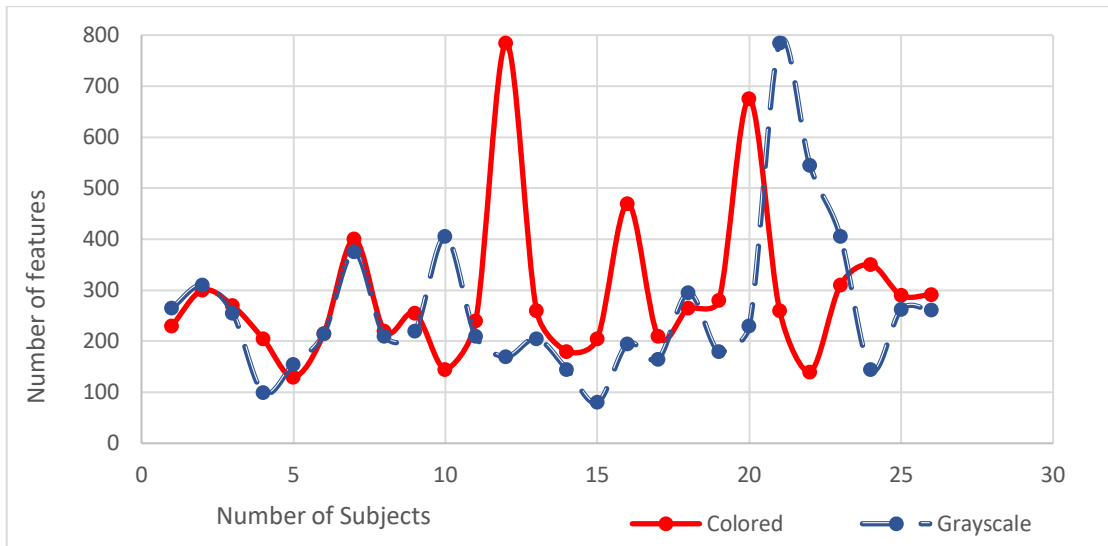


Fig. 4. Number of features for colored and grayscale images

3.3 Result of Correct responses

In addition to the above results, correct responses are also measured for both data sets. Correct responses means that how many times the subject has selected the right option. Since the aim of this study is to do a comprehensive analysis, thus focus was to find the every aspect of the participant’s response. In correct response, the number of right responses for each experiment was calculated. The right response of each subject is mentioned in Fig. 5, which clearly indicates that in case of color images number of right responses were more than the grayscale images.

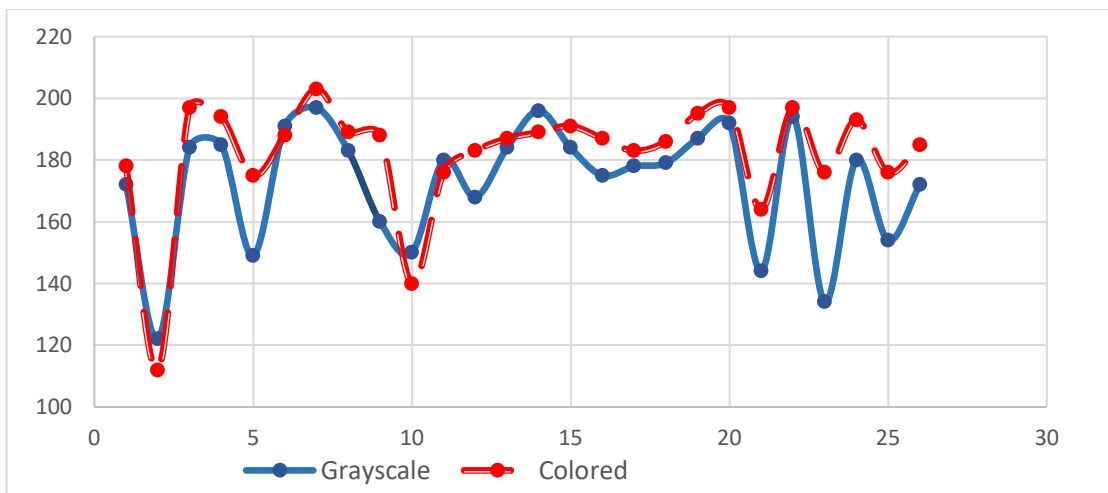


Fig. 5. Number of Correct Responses for Colored and Grayscale images

Since the purpose of this study is to do a comprehensive comparison between the colored and gray-scale stimuli so every aspect of the analysis is considered. Although the accuracy and correct responses are already found; finally, the response time for both data sets is calculated and shown in Fig. 6. It can be clearly visible in the Fig. 6 that grayscale images have more response time compared to the colored which means that participant's response is fast in colored case.

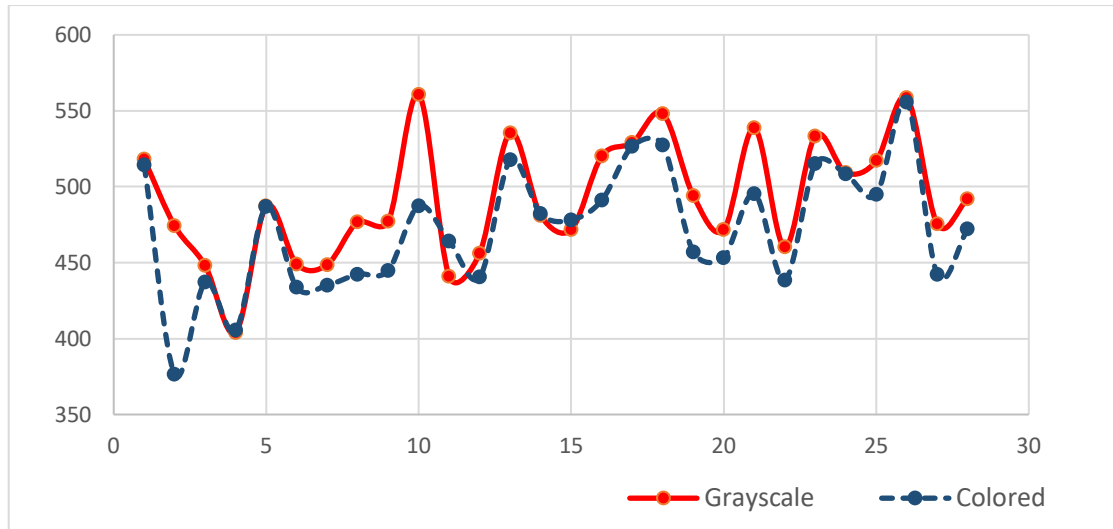


Fig. 6. Response time (millisecond) for Colored and Grayscale images

3.4 Behavior of features for both data sets

Furthermore, it is also observed that in most of the cases the number of features are less in case of grayscale compared to colored, but in few cases of grayscale images, more number of significant features are used which showed better results. During analysis, it is observed that it is true in most of the cases, however it is noticed that those subjects which used more features for grayscale have better accuracy compared to color. This result is observed for two subjects i.e. 1 and 10; in which more significant features are used and have better accuracy. The accuracy against various number of significant features were found and only the best results are mentioned in Fig. 2, however the number of features for best results are shown in Fig. 4 for all subjects.

3.5 Results within Categories

In addition to the comparison of colored and grayscale images, we have also compared the accuracy, response time and correct responses within the categories. A pattern is observed between accuracy, response time and correct responses like the category which showed good classification results have more correct responses with less response time. Similarly, the category which showed good results in colors like human and flower also have good accuracy in grayscale. Similarly, the category which have least accuracy in colored like electronic devices also have low accuracy in grayscale as shown in Fig. 7.

In the above discussion, the comprehensive results are defined in every aspect. Firstly, the univariate analysis was done, followed by multivariate pattern analysis. The number of significant features are also found and in most of the participants, these features are less than 400 for both colored and grayscale. Moreover, colored images used slightly more significant features than grayscale. The number of correct responses are also calculated and are better in case of colored images than grayscale in most of the cases. The average response time of all participants are determined and fast response is found in colored images compared to grayscale in most of the participants.

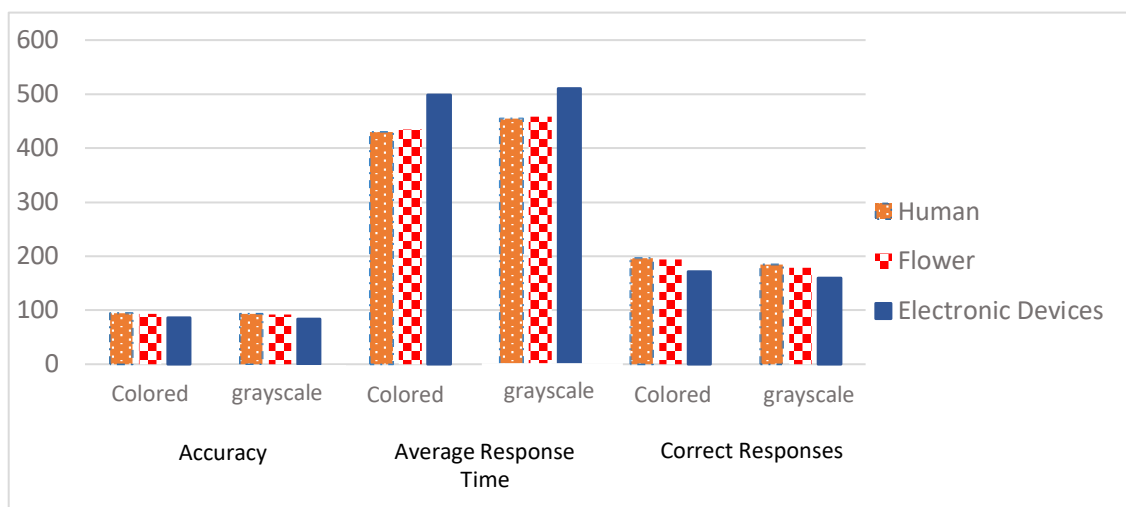


Fig. 7. Accuracy, Average response time and correct responses of colored and grayscale images within categories

4. Discussion

In this study, it is examined whether color images can better facilitate the subject in terms of recognition. Accuracy, right response and early detection of colored and grayscale images are investigated and compared. This study explores the classification performance using various feature extraction and classification techniques for two data sets i.e. colored and gray scale. In existing classification and decoding studies [2, 17, 26], both colored and gray scale images were used as stimuli but less research has compared the effect of brain, response time and right/wrong response of subjects for both types of images. In this study, we have found and matched the effect of brain for both type of images comprehensively. Although in both data sets same images were presented but in different colors and order. For half of the subjects, initially colored task was performed followed by the gray scale and for rest of the subjects the order of task was reversed.

The accuracy, specificity and sensitivity for both data sets are compared. Results showed that the accuracy is better for colored stimuli compared to gray scale. Although, couple of subjects showed better accuracy with grayscale images, however more number of significant features is engaged in these cases which mean significant features also have role in

classification of human brain activity. In addition to above analysis, the right responses and response time was also measured which were better for colored images. This means the subjects distinguished the colored stimuli more quickly, accurately and rightly. In response based task, subjects should reply for every image but in some cases subjects did not reply due to late recognition of the category. During the analysis, it is also observed that in most of the cases the subjects did not reply more for grayscale images compared to colored images. The total non-reply images were 547 in case of grayscale compared to 352 in case of colored for all 25 subjects. Finally, in this study it is found that EEG can better be classified with DWT compared to other feature extraction techniques along with SVM and MVPA.

In few studies, color faces were detected earlier than grayscale faces [41]. Importantly, however, color faces were also detected faster than grayscale faces in the same color context, but in the current study we have not only prove the existing concept but also examined the brain behavior for both colored and grayscale images along with the classification accuracy. The subjects were chosen randomly and did not have any previous experience of data collection, this may be the reason that subjects were more comfortable with colored images. In neuroscience studies, it is common to first train the subject before the collection of data. It is done in several fMRI studies [17, 25, 42] and achieved good results with gray scale images as subject become familiar with the environment and stimuli.

The detailed comparison of different feature extraction and selection techniques for proposed and existing methods are described in Table 4. It consists of 4 columns, in which feature extraction, feature selection methods are mentioned along with the average accuracy of colored and grayscale images using SVM and LDA. First column is representing the feature extraction methods in which three different methods are used i.e. raw data, sample entropy and discrete wavelet transform. For feature selection three methods i.e. Bhattacharya distance, t-test and global field power methods are used and mentioned in second column. Average accuracy of colored images using SVM and LDA is shown in third column while for grayscale images, average accuracy is mentioned in the fourth column using SVM and LDA.

Table 4. Comparison of different feature extraction and selection methods using SVM

<i>Feature Extraction method</i>	<i>Feature Selection method</i>	<i>Average Accuracy of Colored images SVM / LDA</i>	<i>Average Accuracy of Grayscale images SVM / LDA</i>
Raw data	Bhattacharya distance	83.8% / 81.8%	81.4% / 79.7%
Raw Data	Global field Power	82.9% / 80.8%	81.7% / 79.5%
Raw data	t-test	85.8% / 84.1%	83.2% / 81.1%
Sample Entropy	Bhattacharya distance	85.4% / 83.7%	82.9% / 80.2%
Sample Entropy	Global field Power	83.9% / 81.7%	80.6% / 78.4%
Sample Entropy	t-test	86.4% / 82.1%	83.8% / 81.6%
Discrete wavelet	Bhattacharya distance	87.3% / 85.8%	83.9% / 81.9%

transform			
Discrete wavelet transform	Global field Power	86.5% / 84.7%	82.3% / 80.4%
Discrete wavelet transform	t-test	89.6% / 87.9%	86.2% / 81.8%

Several current methods are implemented on colored and grayscale data sets. Three different feature extraction methods are used with three different feature selection methods using two different classifiers on two different data sets. Raw data is used with Bhattacharya distance, t-test and global field power using SVM and LDA for colored and grayscale images. Similarly sample entropy and discrete wavelet transform are used with Bhattacharya distance, t-test and global field power using SVM and LDA for colored and grayscale images. The result for all combinations are summarized in [Table 3](#). Although good results are achieved in each case, however the combination of discrete wavelet transform and t-test along with SVM gives the best result as compared to all other combinations. Raw data and Bhattacharya distance is giving 83.8% using SVM and 81.8% using LDA for colored images while 81.4% using SVM and 79.7% using LDA for grayscale images. Similarly the best results, discrete wavelet transform and t-test give 89.6% with SVM and 87.9% with LDA for colored images and 86.2% with SVM and 81.8% with LDA for grayscale images.

The study in hand has not only focused on experiment design but also discussed several methods of feature extraction and selection. The analysis of EEG data is quite complex starting from experiment design, pre-processing, feature extraction, feature selection and classification. The finest analysis is required at each step to get the overall best results. This study not only explain the basics of EEG data analysis but also different methods and cases of experimentation design. Moreover, human behavior is discussed with different cases including response time, right response and several other conditions. As human brain is a complex play, it is rather impossible to completely define the behavior of the brain due to many restrictions like hardware, response of the subject and many others. However, the data acquired through various steps can be transformed into useful information through several statistical and mathematical techniques which are discussed in this paper.

5. Conclusion

It is a challenging task to extract the significant feature from the human brain as it involves complex data and artifacts. Moreover, MVPA is a better way for the analysis of EEG data as this method helps in extracting more information from the brain. Experiment design also has a key role which decides either the task is simple or response based, may have colored or gray scale images and many other limitations as each step has the ability to extract information from the brain. The present study much focused on colored vs gray scale images along with the different analysis techniques. Moreover, experiment design is helpful in improving results by extracting large amount of information from the brain including the type and color of image i.e. better results can be achieved by using colored images instead of gray scale in terms of accuracy, specificity and sensitivity. In addition, it is determined that subject is more comfortable with colored images as more right responses and less response

time is observed with colored images. Lastly, extraction, selection and classification of data using advanced techniques is essential for worthy results; since in neuroscience every single process has the key role in final analysis. This is the reason for better overall results every single analysis step need proper processing and performance.

Acknowledgment

This research is supported by the Ministry of Education (MOE), Malaysia under the Grant of Higher Institution Centre of Excellence (HICoE) for CISIR (Ref: 0153CA-002).

References

- [1] A. Gardumi, D. Ivanov, L. Hausfeld, G. Valente, E. Formisano, and K. Uludağ, "The effect of spatial resolution on decoding accuracy in fMRI multivariate pattern analysis," *NeuroImage*, vol. 132, pp. 32-42, 2016. [Article \(CrossRef Link\)](#)
- [2] M. Taghizadeh-Sarabi, M. R. Daliri, and K. S. Niksirat, "Decoding objects of basic categories from electroencephalographic signals using wavelet transform and support vector machines," *Brain topography*, vol. 28, pp. 33-46, 2015. [Article \(CrossRef Link\)](#)
- [3] P. Agrawal, D. Stansbury, J. Malik, and J. L. Gallant, "Pixels to voxels: modeling visual representation in the human brain," *arXiv preprint arXiv:1407.5104*, 2014. [Article \(CrossRef Link\)](#)
- [4] R. Zafar, S. C. Dass, A. S. Malik, N. Kamel, M. J. U. Rehman, R. F. Ahmad, J. M. Abdullah, and F. Reza, "Prediction of human brain activity using likelihood ratio based score fusion," *IEEE access*, vol. 5, pp. 13010-13019, 2017. [Article \(CrossRef Link\)](#)
- [5] R. Zafar, N. Kamel, M. Naufal, A. S. Malik, S. C. Dass, R. F. Ahmad, J. M. Abdullah, and F. Reza, "A study of decoding human brain activities from simultaneous data of EEG and fMRI using MVPA," *Australasian physical & engineering sciences in medicine*, vol. 41, pp. 633-645, 2018. [Article \(CrossRef Link\)](#)
- [6] R. Zafar, A. S. Malik, N. Kamel, S. C. Dass, J. M. Abdullah, F. Reza, and A. H. Abdul Karim, "Decoding of visual information from human brain activity: A review of fMRI and EEG studies," *Journal of integrative neuroscience*, vol. 14, no. 2, pp. 155-168, 2015. [Article \(CrossRef Link\)](#)
- [7] J.-D. Haynes and G. Rees, "Decoding mental states from brain activity in humans," *Nature Reviews Neuroscience*, vol. 7, pp. 523-534, 2006. [Article \(CrossRef Link\)](#)
- [8] S. M. McClure, J. Li, D. Tomlin, K. S. Cypert, L. M. Montague, and P. R. Montague, "Neural correlates of behavioral preference for culturally familiar drinks," *Neuron*, vol. 44, pp. 379-387, 2004. [Article \(CrossRef Link\)](#)
- [9] G. Van Doorn and B. Miloyan, "The Pepsi Paradox: A review," *Food quality and preference*, vol. 65, pp. 194-197, 2018. [Article \(CrossRef Link\)](#)
- [10] "Deceiving the law," *Nat Neuroscience*, vol. 11, pp. 1231-1231, 2008. [Article \(CrossRef Link\)](#)
- [11] M. Akin, "Comparison of wavelet transform and FFT methods in the analysis of EEG signals," *Journal of Medical Systems*, vol. 26, pp. 241-247, 2002. [Article \(CrossRef Link\)](#)
- [12] M. Murugappan, N. Ramachandran, and Y. Sazali, "Classification of human emotion from EEG using discrete wavelet transform," *Journal of Biomedical Science and Engineering*, vol. 3, p. 390, 2010. [Article \(CrossRef Link\)](#)
- [13] A. Subasi and M. I. Gursoy, "EEG signal classification using PCA, ICA, LDA and support vector machines," *Expert Systems with Applications*, vol. 37, no. 12, pp. 8659-8666, 2010. [Article \(CrossRef Link\)](#)

- [14] R. Zafar, S. C. Dass, and A. S. Malik, "Electroencephalogram-based decoding cognitive states using convolutional neural network and likelihood ratio based score fusion," *PloS one*, vol. 12, p. e0178410, 2017. [Article \(CrossRef Link\)](#)
- [15] M. Bindemann and A. M. Burton, "The role of color in human face detection," *Cognitive Science*, vol. 33, no. 6, pp. 1144-1156, 2009. [Article \(CrossRef Link\)](#)
- [16] Z. Lu, B. Guo, A. Boguslavsky, M. Cappello, W. Zhang, and M. Meng, "Distinct effects of contrast and color on subjective rating of fearfulness," *Frontiers in psychology*, vol. 6, p. 1521, 2015. [Article \(CrossRef Link\)](#)
- [17] K. N. Kay, T. Naselaris, R. J. Prenger, and J. L. Gallant, "Identifying natural images from human brain activity," *Nature*, vol. 452, pp. 352-355, 2008. [Article \(CrossRef Link\)](#)
- [18] M. A. Rahman, W. Ma, D. Tran, and J. Campbell, "A comprehensive survey of the feature extraction methods in the EEG research," in *Proc. of International Conference on Algorithms and Architectures for Parallel Processing*, pp. 274-283, 2012. [Article \(CrossRef Link\)](#)
- [19] S. Freiders, S. Lee, D. Statz, and T. Kim, "The Influence of Color on Physiological Response," [Article \(CrossRef Link\)](#)
- [20] S. Sanei and J. A. Chambers, *EEG signal processing*, John Wiley & Sons, 2013. [Article \(CrossRef Link\)](#)
- [21] A. S. Al-Fahoum and A. A. Al-Fraihat, "Methods of EEG Signal Features Extraction Using Linear Analysis in Frequency and Time-Frequency Domains," *International Scholarly Research Notices*, vol. 2014, 2014. [Article \(CrossRef Link\)](#)
- [22] M. Hajinorozi, T.-P. Jung, C.-T. Lin, and Y. Huang, "Feature extraction with deep belief networks for driver's cognitive states prediction from EEG data," in *Proc. of Signal and Information Processing (ChinaSIP), 2015 IEEE China Summit and International Conference on*, pp. 812-815, 2015. [Article \(CrossRef Link\)](#)
- [23] W. Zhu, J. Drewes, and K. R. Gegenfurtner, "Animal detection in natural images: effects of color and image database," *PLoS One*, vol. 8, p. e75816, 2013. [Article \(CrossRef Link\)](#)
- [24] M. W. Self and S. Zeki, "The integration of colour and motion by the human visual brain," *Cerebral Cortex*, vol. 15, pp. 1270-1279, 2004. [Article \(CrossRef Link\)](#)
- [25] T. Naselaris, R. J. Prenger, K. N. Kay, M. Oliver, and J. L. Gallant, "Bayesian reconstruction of natural images from human brain activity," *Neuron*, vol. 63, no. 6, pp. 902-915, 2009. [Article \(CrossRef Link\)](#)
- [26] J. A. Mumford, B. O. Turner, F. G. Ashby, and R. A. Poldrack, "Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses," *NeuroImage*, vol. 59, no. 3, pp. 2636-2643, 2012. [Article \(CrossRef Link\)](#)
- [27] C. Kanan and G. W. Cottrell, "Color-to-grayscale: does the method matter in image recognition?," *PloS one*, vol. 7, p. e29740, 2012. [Article \(CrossRef Link\)](#)
- [28] EGI, "Net Station 5 Software," 2016. Available at: <https://www.egi.com/research-division/geodesic-eeeg-system-components/eeeg-software>.
- [29] R. Haddadi, E. Abdelmounim, M. E. Hanine, and A. Belaguid, "Discrete Wavelet Transform Based Algorithm for Recognition of QRS Complexes," in *Proc. of 2014 International Conference on Multimedia Computing and Systems (ICMCS)*, 2014. [Article \(CrossRef Link\)](#)
- [30] G. J. Jiang, S.-Z. Fan, M. F. Abbod, H.-H. Huang, J.-Y. Lan, F.-F. Tsai, H.-C. Chang, Y.-W. Yang, F.-L. Chuang, and Y.-F. Chiu, "Sample entropy analysis of EEG signals via artificial neural networks to model patients' consciousness level based on anesthesiologists experience," *BioMed research international*, vol. 2015, 2015. [Article \(CrossRef Link\)](#)
- [31] D. Bai, T. Qiu, and X. Li, "sample entropy and its application in EEG based epilepsy detection," *Sheng wu yi xue gong cheng xue za zhi= Journal of biomedical engineering= Shengwu yixue gongchengxue zazhi*, vol. 24, pp. 200-205, 2007.
- [32] M. Weippert, M. Behrens, A. Rieger, and K. Behrens, "Sample entropy and traditional measures of heart rate dynamics reveal different modes of cardiovascular control during low intensity exercise," *Entropy*, vol. 16, pp. 5698-5711, 2014. [Article \(CrossRef Link\)](#)

- [33] R. E. Yohanes, W. Ser, and G.-B. Huang, "Discrete Wavelet Transform coefficients for emotion recognition from EEG signals," in *Proc. of Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pp. 2251-2254, 2012. [Article \(CrossRef Link\)](#)
- [34] I. Simanova, M. Van Gerven, R. Oostenveld, and P. Hagoort, "Identifying object categories from event-related EEG: toward decoding of conceptual representations," *PloS one*, vol. 5, p. e14465, 2010. [Article \(CrossRef Link\)](#)
- [35] R. Merry, "Wavelet theory and applications: a literature study," 2005.
- [36] M. Cauchoix, G. Barragan-Jason, T. Serre, and E. J. Barbeau, "The neural dynamics of face detection in the wild revealed by MVPA," *The Journal of Neuroscience*, vol. 34, pp. 846-854, 2014. [Article \(CrossRef Link\)](#)
- [37] R. Zafar, N. Kamel, M. Naufal, A. S. Malik, S. C. Dass, R. F. Ahmad, J. M. Abdullah, and F. Reza, "Decoding of visual activity patterns from fMRI responses using multivariate pattern analyses and convolutional neural network," *Journal of integrative neuroscience*, vol. 16, pp. 275-289, 2017. [Article \(CrossRef Link\)](#)
- [38] P. Baldi and A. D. Long, "A Bayesian framework for the analysis of microarray expression data: regularized t-test and statistical inferences of gene changes," *Bioinformatics*, vol. 17, pp. 509-519, 2001. [Article \(CrossRef Link\)](#)
- [39] T. N. Tran, R. Wehrens, and L. M. Buydens, "Clustering multispectral images: a tutorial," *Chemometrics and Intelligent Laboratory Systems*, vol. 77, pp. 3-17, 2005. [Article \(CrossRef Link\)](#)
- [40] S. M. Crouzet, N. A. Busch, and K. Ohla, "Taste quality decoding parallels taste sensations," *Current Biology*, vol. 25, pp. 890-896, 2015. [Article \(CrossRef Link\)](#)
- [41] M. Lewis and A. Edmonds, "Searching for faces in scrambled scenes," *Visual Cognition*, vol. 12, pp. 1309-1336, 2005. [Article \(CrossRef Link\)](#)
- [42] K. N. Kay, J. Winawer, A. Mezer, and B. A. Wandell, "Compressive spatial summation in human visual cortex," *Journal of Neurophysiology*, vol. 110, no. 2, pp. 481-494, 2013. [Article \(CrossRef Link\)](#)



RAHEEL ZAFAR is born in Rawalpindi, Pakistan in 1980. He did his PhD from Universiti Teknologi PETRONAS, Malaysia in Electrical & Electronic Engineering with majors in Bio-Medical Engineering. He did Master in 2006 from University of Engineering & Technology Taxila Pakistan. Currently, he is serving as Assistant Professor in department of engineering at National University of Modern Languages Islamabad Pakistan. He is working there since February 2007.



MUHAMMAD NOMAN MALIK. He is born in Bahawalpur, Pakistan, in 1982. He received the B.Sc. in computer science from Islamia University, Bahawalpur, in 2001, the MIT degree from the National University of Modern Languages, Islamabad, Pakistan, in 2004, the M.S. degree in engineering management (software project management) from the Center of Advanced Studies in Engineering, University of Engineering and Technology, Taxila, in 2009, and the Ph.D. degree from the Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia, in 2015. He has been with the Department of Computer Sciences since 2005. He teaches several courses, including software project management, research methodology, human-computer interaction, and web programming. He also directs master's and PhD student research projects. He serves as Assistant Professor with the Department of Computer Sciences, Faculty of Engineering and Computer Sciences, National University of Modern Languages. His current research is on software crowdsourcing, sustainable software development, human aspects in software development, human-computer interaction, and big social data. He is a member of IEEE and the International Association of Computer Science and Information Technology.



HUMA HAYA KHAN. She is born in Attok, Pakistan, in 1985. She received the B.S. degree in computer science from the National University of Modern Languages, Islamabad, Pakistan, in 2007, the M.S. degree in software engineering from International Islamic University, Islamabad, in 2011, and the Ph.D. degree in software engineering from Universiti Teknologi Malaysia in 2015. She has been with the Department of Software Engineering, National University of Modern Languages, since 2015. She teaches several courses, including software process and quality, requirements, design, software project management, and software construction. She also directs master's and bachelor's student projects. She serves as an Assistant Professor with the National University of Modern Languages. Her research interests include software engineering process and quality, software measurement, usability evaluation, software requirement engineering, and global software development. Her current research is on emotion based requirement engineering in global software development.



AAMIR SAEED MALIK completed his PhD from Gwangju Institute of Science & Technology, South Korea. Currently, he is Associate Professor and the ex-director of Biomedical Technology Group, one of the Mission Oriented Research at Universiti Teknologi PETRONAS (UTP).