

# ***A System for True and False Memory Prediction based on 2D and 3D Educational Contents and EEG Brain Signals***

Saeed Bamatraf<sup>1</sup>, Muhammad Hussain<sup>1</sup>, Hatim Aboalsamh<sup>1</sup>, Emad-Ul-Haq Qazi<sup>1</sup>, Aamir Malik<sup>3</sup>,  
HafeezUllah Amin<sup>3</sup>, Hassan Mathkour<sup>1</sup>, Ghulam Muhammad<sup>2</sup>, Hafiz Muhammad Imran<sup>1</sup>

<sup>1</sup>Department of Computer Science, <sup>2</sup>Department of Computer Engineering, College of Computer  
& Information Sciences, King Saud University, Riyadh, Saudi Arabia

<sup>3</sup>Department of Electrical & Electronics Engineering, Universiti Teknologi PETRONAS, Tronoh,  
Perak, Malaysia  
(mhussain@ksu.edu.sa)

## ***Abstract:***

The advances in 3D consumer technology has raised the question whether 3D educational contents are more effective than their 2D counterparts for learning and memory retention/recall. Electroencephalography (EEG) brain signals have been extensively studied to decode human brain activations performing different tasks. We studied the impact of 2D and 3D educational contents on learning and memory recall using EEG brain signals. For this purpose, a classification approach has been adopted that predicts true and false memories in case of both short term memory (STM) and long term memory (LTM) and helps to decide whether there is a difference between the impact of 2D and 3D educational contents on learning and memory recall. In the proposed approach, EEG brain signals are converted into topomaps (scalp-maps) and then discriminative features are extracted/selected from these topomaps and finally support vector machine (SVM) is employed to predict brain states as true and false memories. For data collection, sixty eight healthy individuals volunteered, half of which watched the learning material in 2D format whereas the second half watched the same material in 3D format. After learning task, memory recall tasks were performed after 30 minutes (for STM) and two months (for LTM), and EEG signals were recorded. In case of STM, 97.5% prediction accuracy was achieved for 3D and 96.6% for 2D and in case of LTM, it was 100% both for 2D and 3D. The statistical analysis of the results suggested that for learning and memory recall both 2D and 3D education materials do not have much difference in case of STM and LTM.

*Keywords:* EEG brain signal, Memory retention and recall, Feature extraction, Support vector machines, 2D and 3D contents

## ***1. Introduction***

The multimedia technology has widely taken over the education system because it makes the concepts easy to understand. The multimedia principle by Mayer states that people learn more deeply from words and pictures than from words alone [1]. Currently, the dominant multimedia resources are mainly in 2D form. With recent advances in science and technology, 3D technology is becoming common and soon it will be available for educational purposes. As such, the educational contents can be presented to the students either in 2D or in 3D format. In 2D format, 3D objects are visualized by projected them on 2D space. With the advent of 3D technology, nowadays, 3D objects can be visualized as 3D with the help of 3D devices. These advances in technology have now made us wonder whether the 3D material is more effective than the 2D content in terms of learning and memory retention and recall.

Electroencephalography (EEG) is an imaging technique that captures brain activation patterns detected as electrical activity on the scalp [2, 3]. EEG signals reflect the brain behavior during mental activities. Since the procedure of EEG is pain free and noninvasive, it has been widely used in normal adults and children to study brain activation patterns associated with tasks like memory retention/recall, perception, attention and emotions.

Some researchers studied the effects of 2D and stereoscopic 3D on tasks like viewer's experience of the movie content [4], spatial cognition [5, 6, 7], deepening the understanding of PC hardware [8], understanding and knowledge acquisition [9, 10], spatial visualization skills [11], and education learning processes [12]. These studies used subjective approach without using EEG brain signals. Also, no one explicitly studied the effects of 2D and 3D educational contents on learning and memory recall. Though the subjective approach based on the statistical analysis of the answers of questions can be used to assess the effects of 2D and 3D educational contents, EEG signals give direct insight into brain states and can lead to more accurate conclusions. Intuitively, the brain states are different while giving answers based on learned information and guess.

EEG technology has been employed by a number of researchers to study brain activations during different tasks such as cognitive tasks [13], playing video games on large screens [14], and playing video games on small and large displays [15]. Some researchers studied the EEG brain signals to locate brain regions and the components responsible for memory functions [16-19]. Also some studies have been done recently on attention, learning and memory using EEG brain activations [20-23]. To our best knowledge there does not exist any study so far, which focused on studying the impact of 2D and 3D educational contents on learning and memory recall using direct brain behavior through EEG technology.

We studied the effects of 2D and 3D educational contents on memory retention and recall with the aid of brain images (topomaps) captured using EEG technology. For this purpose, we developed pattern recognition systems to predict true memory (remembered/recalled) and false memory (forgotten) using direct brain responses via topomaps for 2D and 3D educational contents and then used them for assessing their effects on learning and memory retention and recall. For developing the systems, the data was collected from sixty eight healthy individuals. Half of them watched the learning material in 2D format whereas the second half watched the same material in 3D format. After learning task, memory retention and recall tasks were performed in the form of multiple choice questions (MCQs) related to the watched contents after 30 minutes (for STM) and two months (for LTM), and EEG signals were recorded.

The proposed pattern recognition systems comprise feature extraction and classification stages. Features are extracted from EEG signals corresponding to correct answer (true memory) and wrong answer (false memory) and SVM is used for prediction. For feature extraction, a simple and robust technique has been introduced that first of all creates topomaps from EEG signals and removes redundancy using city-block distance. One system was developed for 2D and 3D each to predict true and false memories, which can be used to predict the memory retention and recall abilities of individuals. Each system encodes the brain states and their performance has been analyzed to decide whether there is a difference between the impacts of 2D and 3D educational contents on memory retention and recall. In case of STM, the mean prediction accuracy is 96.5% for 2D and 97.5% for 3D, whereas it is 100% for both 3D and 2D for LTM. Further statistical analysis of the results revealed that there is no significant difference in using 3D and 2D materials as far as memory retention and recall is concerned. The initial results of this study were presented in [24]; this paper extends the study to LTM and gives thorough analysis of the proposed methods.

**The main contributions of the paper are:**

- a pattern recognition system for predicting the state of memory, whether the memory is true or false and
- assessment of the effects of 2D and 3D educational contents on memory retention and recall.

The remaining paper is organized as follows. The proposed methodology is elaborated in Section 2. The results are presented and discussed in Section 3 and Section 4 gives the statistical analysis. Section 5 concludes the paper.

## 2. Methods and Materials

We developed one pattern recognition system for 2D and 3D each for assessing the brain states while true and false answers of MCQs i.e. true and false memories. Each system was trained with the labeled EEG brain signals that encode the brain behavior during true answer (true memory) and false answer (false memory). As both the systems have the same architecture and configuration, and have been trained in the same way, so the trained systems encode the brain behaviors corresponding to 2D and 3D contents and their prediction accuracies can be used to assess the effectiveness of 2D and 3D for learning and memory recall. To assess the effects of 2D and 3D, we developed the hypothesis that performance of the system associated with 3D is better than that of 2D. Precisely, the hypothesis is modeled as follows:

**Null hypothesis H0:** performance of the system associated with 3D is better than that of 2D

**Alternate hypothesis H1:** There is no difference in the performance of the systems associated with 2D and 3D.

Using hypothesis testing, we conclude whether the null hypothesis holds or not.

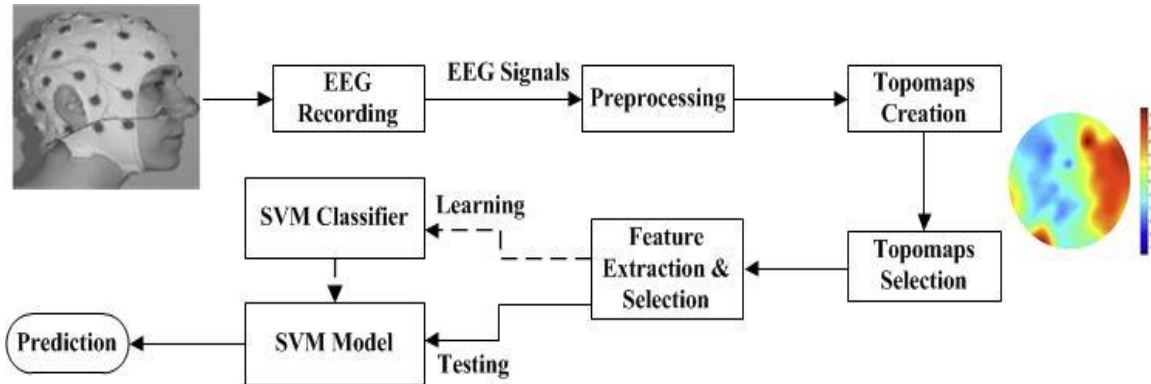
**Our hypothesis about the difference between 2D and 3D groups is based on the following observations:**

- For data collection, the subjects were selected in such a way that the age, IQ levels and the background knowledge of the subjects in each group were same.
- Each group watched the same educational contents and responded to the questionnaires containing the same questions. The only difference was that of 2D and 3D formats.
- Modeled the brain states of each group as a classification system, which is based on the same feature extraction technique and the same classification method.
- Each system was implemented on the same platform using the same environment and was learned in the same way and was evaluated using the same procedure.
- There is complete similarity from data collection to learning and evaluating the systems. If there is any difference between 2D and 3D educational contents in respect of memory retention and recall, it should appear in the performance of the systems.

In the rest of this section, first, we give a general overview of the pattern recognition system. Then, we zoom into each component of the system. Feature extraction is an important component of the system, and discriminative features are needed to represent the brain behavior reliably corresponding to the answer of a question. We introduce a simple and robust method for feature extraction.

### 2.1 Proposed System

An overview of the proposed system is shown in Figure 1. First, subjects are selected to participate in the learning and memory recall tasks. EEG signals are recorded from the subjects during the tasks. These signals are preprocessed to remove the artifacts and noisy signals. Topomaps are then created from the clean signals, and after that, the most representative topomaps are selected from



**Figure 1. The proposed system for predicting true and false memories.**

the huge number of topomaps per sample (question). Different feature extraction approaches are used to extract the features from topomaps. Feature selection is needed to reduce the higher dimensionality of feature spaces and selects the most representative features, which subsequently are passed to the classifier to produce the classification model. Using the train classification model, we assess true and false memories and afterwards analyze the effects of 2D and 3D educational contents on learning and memory retention and recall.

In the following subsections, we give detail of each step in the proposed system.

### ***2.1.1 Collection of Data***

The details related to the data collection process is presented in this section.

#### ***Selection of Participants***

Total number of volunteers who took part in the experiments were 68 and their ages were in the range of 18 to 30 years. The volunteers were having normal or close to normal eye sight and were not suffering with any form of neurological disorders, which could affect the results. Two groups were formed for 2D and 3D based on metrics of age and their knowledge.

#### ***Learning Contents and Tasks Related to Experiments***

The same learning contents for both 2D and 3D related to the experiments in this study, were acquired from Eureka 3D system [25]. The contents were of biology in nature and the volunteers had no existing knowledge in biological area. Learning and memory recall were two building blocks of the experiments performed. In learning phase, the participants watched 2D or 3D learning contents depending upon their group for the time spans of 8 to 10 minutes. In the phase of memory recall, two retention periods were employed, i.e., one was for STM in which the retention period was of 30 minutes and the other was for LTM in which the retention period was of two months. In this process, twenty MCQs were inquired from the participants and each MCQ had four choices for answers. The same MCQs were asked from participants of both the groups, i.e., 2D and 3D. While performing the recall experiment, EEG signals were recorded for study. The 41-inch TV screens were used for the display of learning contents and MCQs of the recall phase and those were placed at the distance of 1.5m from the eyes of the participants. The participants had to answer the questions in the max time of 30 seconds. In the case of right answers, ‘1’ was assigned and in the opposite case ‘0’ was assigned for the wrong ones.

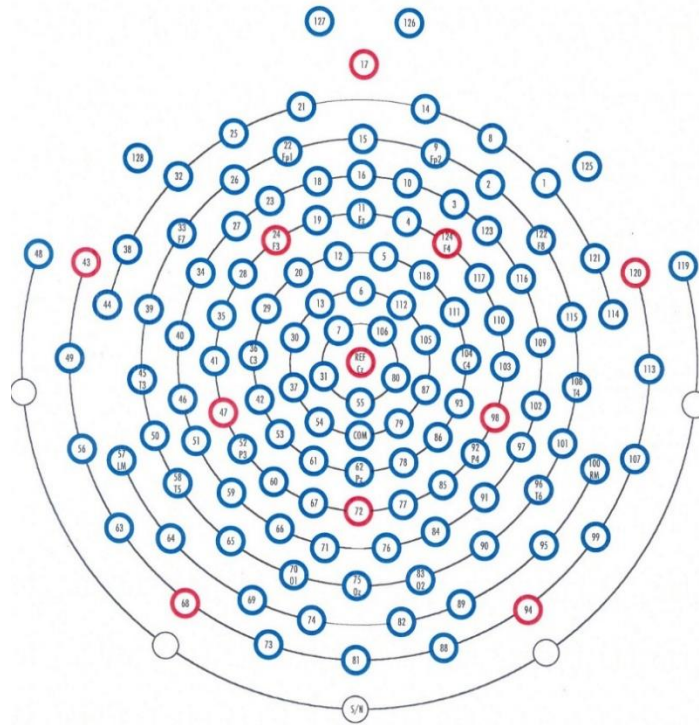
#### ***Recording of EEG Signals***

The recording of EEG signals was done in the period of 30 seconds, which started when the participants were shown the questions and ended when they chose answers from the four choices for the answer. The starting and ending time points of the time period specified for answering the question were recorded in a file, which were used eventually for the extraction of the part of EEG signal related to the question. Sampling rate of 250 samples/sec and 128 channels Hydro Cel Geodesic Net (see Figure 2) were used for the recording of EEG data. The EEG signals are spatio-temporal data which are characterized by the temporal evolution and the adjacent spatial potential distributions. These signals are represented with  $X(k) \in \mathbb{R}^{M \times T}$  where  $k$  is the question identifier,  $M$  is the number of channels and  $T$  is the number of sampled data points along temporal evolution. The selection of electrodes was performed based on their location on the scalp, i.e., the outermost electrodes were omitted and in total 93 were chosen out of 128 electrodes for further analysis.

### 2.1.2 Preprocessing

The noise in the form of artifacts, i.e., eye blinking, etc., is present in the recorded EEG data, which adversely affects the performance of feature extraction and eventually the prediction accuracy. It is utmost important to remove such noise for the overall performance of the system. In preprocessing phase, such noise is removed. The noise is present in various forms in the recorded EEG data, which includes the EEG activity that is not the result of response of stimuli; noise due to the variability in ERP components which is a result of neural and cognitive activity variations; another common source of noise is the presence of bioelectric activities like movement of eyes, blinking, movement of muscles, etc.; and the final source of noise is due to the electric equipment like display devices, etc.

The artifacts in the raw EEG data were detected using a band pass filter (1-48Hz), and then exported into Matlab files (.mat) format using Netstation software of EGI. Ocular artifacts were removed by using Gratton and Coles method [26] and visual inspection.

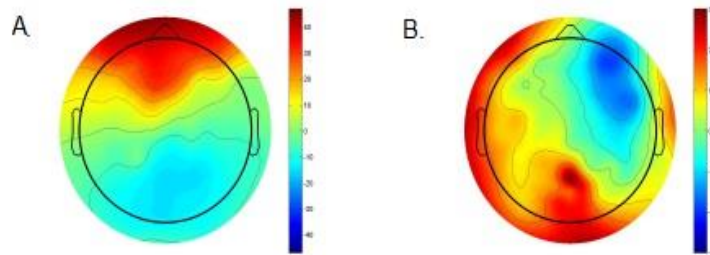


**Figure 2. 128-channels Hydro Cel Geodesic Net**

### 2.1.3 Creation of Topomaps

One possibility for feature extraction is EEG power spectral density, but it is computed in Fourier domain where time information is lost. We know that brain is a nonlinear dynamic system [30] and to keep track of the evolution of brain states over time is important. As such, we used EEG potential because it represents the changes of brain states over time. Recent studies have shown that time domain analysis play an important role in understanding the EEG signals [31-34]. These studies have applied many time analysis methods on EEG recorded during performing various tasks (such as problem solving [32]) and brain states (e.g. emotional states [33], and epileptic seizure [35]). Hence, we have focused on the EEG potential rather than spectral analysis.

EEG signal at a particular time point  $t$  is a collection of the potentials at all electrode positions i.e.  $x^t = (x_1^t, x_2^t, x_3^t, \dots, x_M^t)$  and is represented as a topomap (a digital image describing brain activation at  $t$ ), two such topomaps are shown in Figure 3. EEG signals corresponding to the answer of each question are converted into topomaps. Features can be extracted from the topomaps using image processing and analysis techniques. The EEG signals are recorded from each subject while answering the questions and saved in a Matlab raw file (.mat) and the time points at which the subject starts answering each question until finishes answering are saved in an event file. Using .mat file together with the event file, we create the topomaps corresponding to the EEG recordings of each question. We used EEGLAB toolbox [27] for this step.



**Figure 3. Two sample topomaps, A: Correct Answer, B: Incorrect Answer.**

The quantity of topomaps depends upon the time which is taken by the participants for answering the questions. The number of topomaps corresponding to a question can be up to 7500 in total as the maximum time allocated for answering a question is 30 sec and sampling rate is 250.

### 2.1.4 Selection of Topomaps

Topomaps which occur consecutively in time contain redundant information, which do not aid in achieving good prediction accuracy. The more the discriminative information, the more the prediction accuracy may be achieved. In order to have maximum discriminative information, the consecutive topomaps are analyzed using a distance metric and as a result the most discriminant ones are selected. A number of distance metrics have been proposed in the research community but cityblock distance metric is well known for its simplicity and effectiveness and therefore we chose it for the selection of most discriminative topomaps. Using cityblock distance metric, the distance between the pairs of topomaps is calculated and put in descending order, and finally the topomaps which are highly dissimilar are selected as shown in Figure 4.

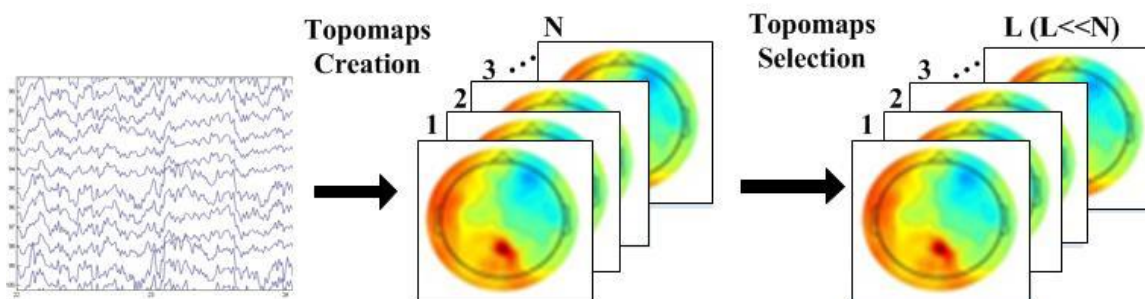
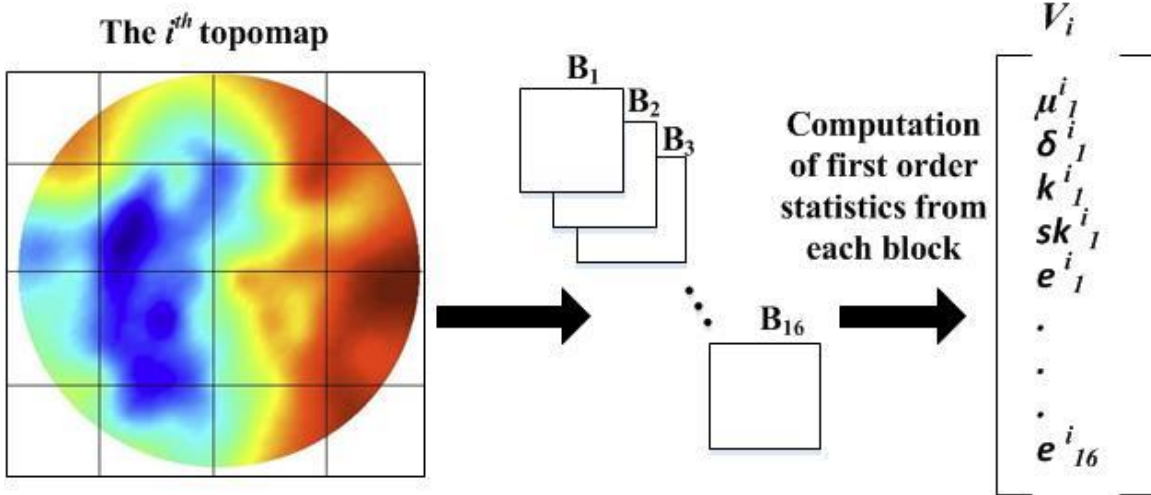


Figure 4. Selection of L topomaps with the highest dissimilarity

### 2.1.5 Feature Extraction

The next step after removing the redundant information is to extract discriminative features from the selected topomaps. First order statistics are used for the extraction of features from topomaps. In order to discriminate the structures in images, texture plays a major role. Texture information from the selected topomaps is extracted using first order statistics: mean, standard deviation, entropy, skewness and kurtosis. First order statistics calculated from a topomap are global features and less discriminative because the localization information of the texture micropatterns is lost. For capturing localized texture information, each topomap is divided into a number of blocks, and then the statistical features are extracted from each block and merged into a vector that forms the representation of the topomap being processed. This process is shown in Figure 5. For our experiments, we tested different numbers of blocks but we found the acceptable number to be  $8 \times 8$  and  $16 \times 16$  blocks. We examined different combinations of five statistical features.

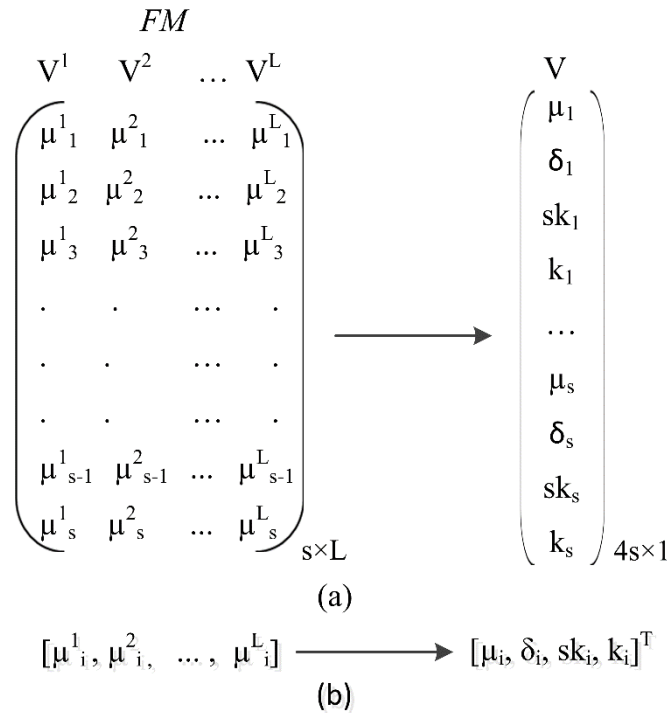
Similarly, we compute the feature vector for each selected topomap and combine them into one feature vector that represents the brain state while answering a question. We can concatenate the vectors corresponding to all selected topomaps to form the feature vector, but in this case the dimension of the feature space will be excessively large causing curse of dimensionality problem. For instance, if 100 topomaps are selected, each topomap is divided into  $16 \times 16$  blocks and 5 features are calculated from each block then the dimension of the feature space will be  $100 \times 16 \times 16 \times 5 = 128000$ . We employ a different approach that reduces the dimension of the feature space significantly while keeping the most discriminative information and does not depend on the



**Figure 5. Features Extraction from 4×4 Blocks of topomaps.**

number of selected topomaps. Using this approach, the dimension of feature space reduces to  $4 \times 16 \times 16 \times 5 = 5120$  for the above case, detail is given below.

After forming the vectors  $V^1, V^2, \dots, V^L$  (each of dimension  $s$ ) which contain features from the  $L$  most discriminant selected topomaps related to one question, the feature matrix  $FM$  of order  $(s \times L)$  is built, where the vectors  $V^1, V^2, \dots, V^L$  forms its columns. Figure 6(a) shows  $FM$  for the case when only one feature (mean  $\mu$ ) is computed from each block. Now from each row of this matrix, the discriminative information is captured by computing four first order statistics (mean, standard



**Figure 6. (a) FM of order  $s \times L$  is the feature matrix for all topomaps corresponding to a question,  $V$  is the feature vector computed from matrix FM, (b) from the  $i$ th row of the matrix FM, four statistical features are extracted.**



deviation, skewness and kurtosis), see Figure 6(b) and these are concatenated to form the feature vector as shown in Figure 6(a).

### ***2.1.6 Feature Selection***

If a topomap is divided into  $16 \times 16$  blocks and five statistical features are computed from each block, then each topomap will be represented by  $(5 \times 16 \times 16 = 1280)$  features and the total number of features will be  $1280 \times 4 = 5120$  per question, which is high. High dimensional feature space has a number of disadvantages like high computational cost and affects the classification accuracy. It is utmost important to employ some feature selection methodology in order to reduce the high dimension of the feature space by reducing the redundancy and selecting the most discriminant features. In the proposed approach, ROC curve is used to select the most important features. **The receiver operating characteristic (ROC) curve measures the class separability of a certain feature. It quantifies the overlap between the distributions of the feature in two classes and is computed as an area between two curves. This area is zero for complete overlap and is 0.5 for complete separation. A feature is discriminatory if the area for this feature is close to 0.5, for detail see [40].**

### ***2.1.7 Classification***

Learning and memory recall process incorporates two classes of data, i.e., correct and incorrect answers. For this problem, many learning models can be used, which includes but not limited to, artificial neural networks, decision trees, support vector machines (SVM) etc. Out of these, support vector machines are regarded as the state of the art classification models which can achieve outstanding accuracies. This is a linear classifier which achieves the maximum margin between the classes of data and can work for the very few training samples and at the same time works for classifying non-linear data as well by taking the data to higher dimensional space using kernel trick. Various kinds of kernels can be used with SVM like radial basis function (RBF) and polynomial. As RBF kernel is known to give good results, we choose it for the classification purposes. The parameters of the kernel are learned using grid-search method and the well-known libsvm [28] library is used for the implementation of SVM. **Note that SVM is useful for applications where the number of features is much larger than the number of samples [36-39].**

## **3. Results and Discussion**

In this section, the results of the proposed approach are presented and subsequently discussed. First, we describe the evaluation policy that we used to evaluate the systems. Then, we present an overview of the subjects and their answers to the questions in both STM and LTM sessions. After that, we present the result of the methods we used on STM and LTM. In the next section, we give the statistical analysis of the result of STM and LTM for 2D and 3D to see if there is a difference between them or not.

### ***3.1 Evaluation Policy***

In this section, the details about the dataset used for experiments are presented along with the evaluation methodology and the metrics used to measure the prediction accuracies are discussed. For the experiments performed, the dataset is selected to make sure that the samples used for training and testing phases fairly represent the two classes. The selected set comprises 200

questions, out of which 100 are from correct class. The same procedure is used for data selection in both 2D as well as 3D learning contents for both the cases of STM as well as LTM.

Classification is performed between subjects so that the system is general, not depends on a particular subject. Each system is learned and tested using the data across different subjects of the same group.

In order to estimate the prediction accuracy of the classifiers, we used 10-fold cross validation [29]. Cross validation is a renowned validation technique that is employed for such purpose. In 10-fold cross validation, data is randomly partitioned into 10 parts in which training is performed on nine parts and the left-over fold is used for testing purposes. This procedure is applied 10 times and in each turn different fold is used for testing purposes while the rest are used for training. At the end, to estimate the prediction accuracy of the system, mean and standard deviation of 10 iterations are calculated. This is the state of the art methodology for estimating the prediction accuracy of a classifier and ensures there is no overfitting because every time the system is trained with different data set and tested with different data set, not used in training.

In order to measure the performance of the systems, the well-known metrics of accuracy and the AUC were used. As we collected the data corresponding to 2D and 3D educational contents in the same way, and also used the same classification models, which were trained and tested in the same way, the prediction performance of the classifiers is going to help in investigating the effectiveness of 2D or 3D learning contents for learning and memory recall.

### 3.2 Results for Long-Term Memory (LTM)

We modeled two systems, one for 2D educational content and one for 3D case. We trained and tested each system using the method described in Section 3.1.1. In this case, each system involves three parameters, which include number of blocks, the extracted features from each block, and the number of selected topomaps.

Initially, we selected 20 topomaps and partitioned each topomap into  $8 \times 8$  and  $16 \times 16$  blocks. The results are shown in Tables 1 and 2. The best accuracy obtained in case of  $8 \times 8$  is 90.5% for 3D with two features and 91.5% for 2D with five features, and it is 91.5% for 3D with two features and 95.5% for 2D with five features in case of  $16 \times 16$  blocks. In case of  $16 \times 16$  blocks, there is little increase in the accuracy but the overall trend is the same, i.e. the best accuracy for 3D case is obtained with 2 features and that for 2D case is with features per block.

**Table 1: Results with  $8 \times 8$  blocks and 20 selected topomaps**

| # Features | Accuracy                         |                                  | AUC             |                 |
|------------|----------------------------------|----------------------------------|-----------------|-----------------|
|            | 2D                               | 3D                               | 2D              | 3D              |
| 1          | $74 \pm 10.8$                    | $81 \pm 9.4$                     | $0.73 \pm 0.13$ | $0.84 \pm 0.11$ |
| 2          | $91.5 \pm 5.8$                   | <b><math>90.5 \pm 6.4</math></b> | $0.90 \pm 0.09$ | $0.89 \pm 0.11$ |
| 3          | $90 \pm 4.7$                     | $89 \pm 5.7$                     | $0.93 \pm 0.07$ | $0.90 \pm 0.05$ |
| 4          | $89.5 \pm 3.7$                   | $89.5 \pm 6$                     | $0.90 \pm 0.08$ | $0.91 \pm 0.07$ |
| 5          | <b><math>91.5 \pm 4.7</math></b> | $88.5 \pm 4.7$                   | $0.93 \pm 0.06$ | $0.93 \pm 0.05$ |

**Table 2: Results with  $16 \times 16$  blocks and 20 selected topomaps**

| # Features | Accuracy     |                | AUC             |                 |
|------------|--------------|----------------|-----------------|-----------------|
|            | 2D           | 3D             | 2D              | 3D              |
| 1          | $84 \pm 6.2$ | $85.5 \pm 8.6$ | $0.82 \pm 0.12$ | $0.84 \pm 0.14$ |

|   |                   |                   |             |             |
|---|-------------------|-------------------|-------------|-------------|
| 2 | 92 ± 7.2          | <b>91.5 ± 5.3</b> | 0.93 ± 0.07 | 0.94 ± 0.07 |
| 3 | 95.5 ± 6          | 88 ± 7.2          | 0.96 ± 0.08 | 0.89 ± 0.08 |
| 4 | 95 ± 4.1          | 88.8 ± 8.2        | 0.96 ± 0.04 | 0.90 ± 0.06 |
| 5 | <b>95.5 ± 4.4</b> | 88.5 ± 7.1        | 0.96 ± 0.04 | 0.89 ± 0.08 |

Next the number of selected topomaps was increased to 30 and features are extracted from 8x8 and 16x16 blocks. The results are shown in Tables 3 and 4. The best accuracy obtained in case of 8x8 is 96.5% for 3D with three features and 92.5% for 2D with three features, and it is 96% for 3D with two features and 97.5% for 2D with four features in case of 16x16 blocks. It indicates that by increasing the number of selected topomaps increases the accuracy.

**Table 3: Results with 8x8 blocks and 30 selected topomaps**

| # Features | Accuracy          |                   | AUC         |             |
|------------|-------------------|-------------------|-------------|-------------|
|            | 2D                | 3D                | 2D          | 3D          |
| 1          | 73.5 ± 10         | 89 ± 7.8          | 0.73 ± 0.15 | 0.87 ± 0.11 |
| 2          | 92 ± 4.2          | 96 ± 4.6          | 0.89 ± 0.07 | 0.97 ± 0.06 |
| 3          | <b>92.5 ± 6.4</b> | <b>96.5 ± 5.8</b> | 0.94 ± 0.06 | 0.96 ± 0.06 |
| 4          | 92 ± 5.4          | 96 ± 3.9          | 0.92 ± 0.08 | 0.97 ± 0.03 |
| 5          | 91 ± 5.2          | 95.5 ± 4.4        | 0.92 ± 0.06 | 0.96 ± 0.04 |

**Table 4: Results with 16x16 blocks and 30 selected topomaps**

| # Features | Accuracy          |                 | AUC         |             |
|------------|-------------------|-----------------|-------------|-------------|
|            | 2D                | 3D              | 2D          | 3D          |
| 1          | 82.5 ± 9.2        | 93.5 ± 5.3      | 0.80 ± 0.12 | 0.94 ± 0.05 |
| 2          | 95.5 ± 3.7        | <b>96 ± 3.2</b> | 0.96 ± 0.06 | 0.97 ± 0.04 |
| 3          | 97 ± 3.5          | 94.5 ± 5.5      | 0.96 ± 0.05 | 0.95 ± 0.07 |
| 4          | <b>97.5 ± 4.3</b> | 93.5 ± 5.8      | 0.97 ± 0.06 | 0.96 ± 0.05 |
| 5          | 97 ± 2.6          | 94 ± 2.1        | 0.97 ± 0.03 | 0.95 ± 0.03 |

In view of this trend, we increased the number of selected topomaps to 50, 100, 120 and 150. With 120 and 150 selected topomaps, we got the maximum accuracy. The results with 150 topomaps are shown in Tables 5 and 6. The best mean accuracy of 100% was achieved with 16x16 blocks and one feature (i.e. mean) per block for both 2D and 3D cases.

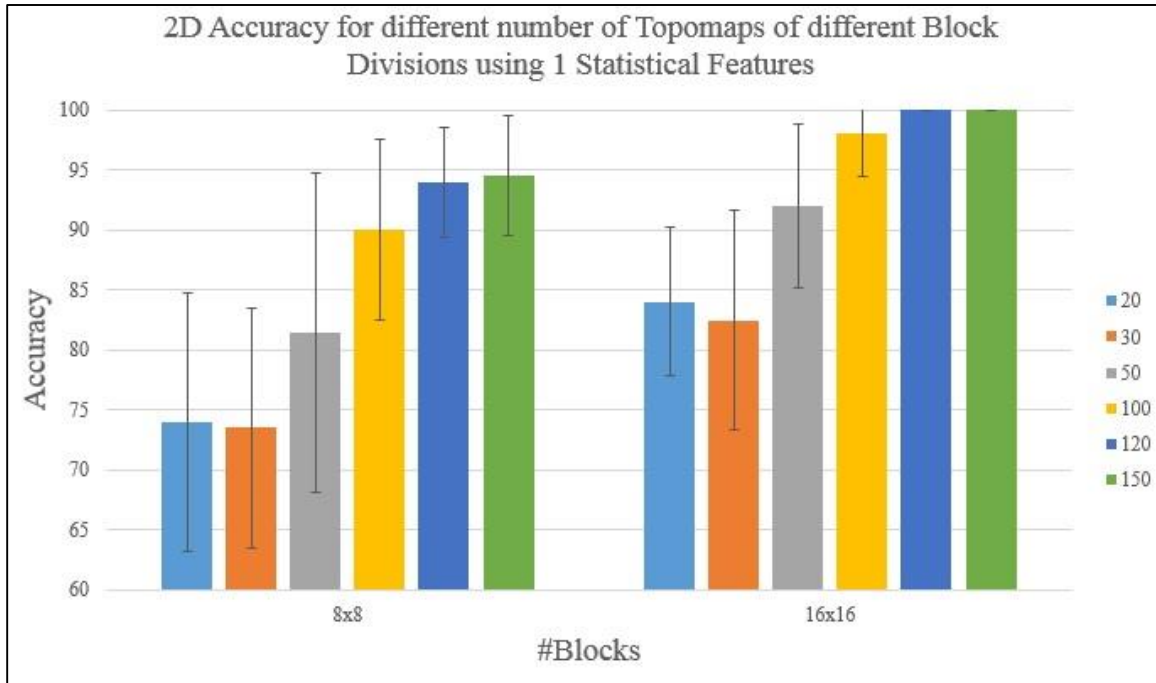
**Table 5: Results with 8x8 blocks and 150 selected topomaps**

| # Features | Accuracy        |            | AUC         |             |
|------------|-----------------|------------|-------------|-------------|
|            | 2D              | 3D         | 2D          | 3D          |
| 1          | 94.5 ± 5        | 97 ± 4.2   | 0.94 ± 0.07 | 0.97 ± 0.04 |
| 2          | 98.5 ± 2.4      | <b>100</b> | 0.99 ± 0.02 | 1           |
| 3          | 98.5 ± 2.4      | 99.5 ± 1.6 | 0.98 ± 0.04 | 0.99 ± 0.01 |
| 4          | <b>99 ± 2.1</b> | 99 ± 2.1   | 0.99 ± 0.04 | 0.99 ± 0.03 |
| 5          | 99 ± 2.1        | 99 ± 2.1   | 0.99 ± 0.02 | 1           |

**Table 6: Results with 16x16 blocks and 150 selected topomaps**

| # Features | Accuracy |    | AUC |    |
|------------|----------|----|-----|----|
|            | 2D       | 3D | 2D  | 3D |

|   |            |            |             |   |
|---|------------|------------|-------------|---|
| 1 | <b>100</b> | <b>100</b> | 1           | 1 |
| 2 | 99.5 ± 1.6 | 100        | 0.99 ± 0.01 | 1 |
| 3 | 99.5 ± 1.6 | 100        | 1           | 1 |
| 4 | 99.5 ± 1.6 | 99.5 ± 1.6 | 1           | 1 |
| 5 | 99 ± 2.1   | 99.5 ± 1.6 | 0.99 ± 0.02 | 1 |



**Figure 7: The best case accuracies in case of LTM for different selected topomaps for 2D educational material.**

Generally, the more the number of topomaps and the number of blocks, the better the accuracy for both 2D and 3D cases as shown in Figures 7 and 8. The best average accuracy (100%) was achieved by 120 and 150 topomaps. The number of selected topomaps, which gives the best accuracy is much less than the total number of topomaps per question (less or equal 7500). It indicates that the discriminative information about the brain behavior for true or false memory is embedded in few topomaps. Further the best performing systems select 150 discriminative topomaps, divide each

topomap into 16×16 blocks and compute one feature (i.e. mean ) from each block. The other performance metric i.e. AUC also shows the similar performance for both 2D and 3D cases.

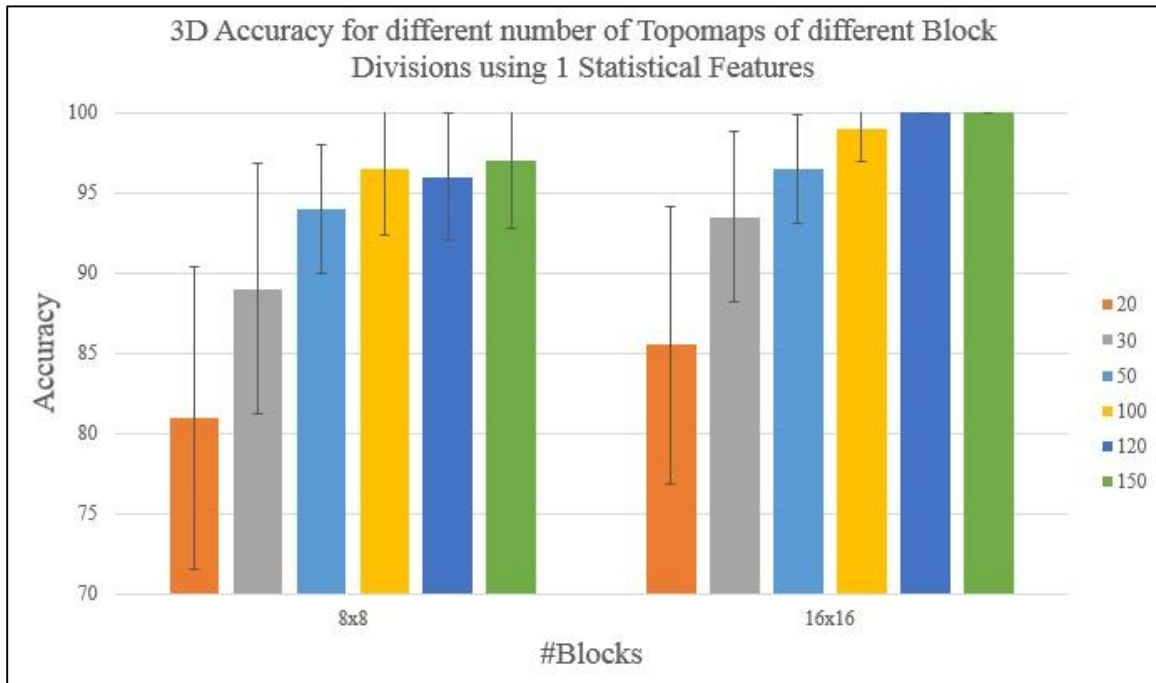
The reason that the average accuracy of 100 was reached in case of LTM is that after two months of retention, subjects either still remember the answers or forgot them at all i.e. clear separation of correct and incorrect answers.

### 3.3 Results for Short-Term Memory (STM)

The results for STM were presented in [24], the detail can be found there. We tried 20, 30, 50, 100, 120 and 150 topomaps with 8x8 and 16x16 blocks. We present here only the results of the best case, which were obtained using 150 topomaps. The results are shown in Tables 7 and 8. The best accuracy obtained in case of 8×8 is 97.5% for 3D with three features and 95.5% for 2D with two features, and it is 97.5% for 3D with two features and 96.5% for 2D with also two features in case of 16×16 blocks. In this case the configuration of the best performing systems for the prediction of true and false memories is 150 most discriminative topomaps, 16×16 block division and two features (mean and standard deviation) from each block.

**Table 7: Results of 8×8 blocks using 150 selected topomaps**

| # Features | Accuracy          |                   | AUC         |             |
|------------|-------------------|-------------------|-------------|-------------|
|            | 2D                | 3D                | 2D          | 3D          |
| 1          | 91 ± 7            | 87.5 ± 6.8        | 0.93 ± 0.06 | 0.87 ± 0.10 |
| 2          | <b>95.5 ± 2.8</b> | 96.5 ± 3.4        | 0.95 ± 0.05 | 0.96 ± 0.07 |
| 3          | 93 ± 5.4          | <b>97.5 ± 3.6</b> | 0.92 ± 0.07 | 0.99 ± 0.01 |
| 4          | 93.5 ± 5.8        | 96.5 ± 4.7        | 0.92 ± 0.08 | 0.98 ± 0.04 |
| 5          | 93.5 ± 3.4        | 97 ± 2.6          | 0.94 ± 0.07 | 0.97 ± 0.04 |



**Figure 8. The best case accuracies in case of LTM for different selected topomaps for 3D educational material.**

**Table 8: Results of 16×16 blocks using 150 selected topomaps**

| # Features | Accuracy          |                   | AUC         |             |
|------------|-------------------|-------------------|-------------|-------------|
|            | 2D                | 3D                | 2D          | 3D          |
| 1          | 95.5 ± 4.4        | 96 ± 3.9          | 0.97 ± 0.05 | 0.97 ± 0.04 |
| 2          | <b>96.5 ± 3.4</b> | <b>97.5 ± 3.5</b> | 0.95 ± 0.06 | 0.98 ± 0.03 |
| 3          | 86 ± 7.4          | 92.5 ± 5.9        | 0.87 ± 0.09 | 0.93 ± 0.08 |
| 4          | 85.5 ± 6.4        | 92.5 ± 7.6        | 0.86 ± 0.06 | 0.93 ± 0.08 |
| 5          | 86.5 ± 6.7        | 92 ± 6.3          | 0.88 ± 0.07 | 0.93 ± 0.07 |

It was observed that prediction accuracy is directly proportional to the number of selected topomaps and reached to its peak for 150 topomaps. The primary factor for this behavior is that when more number of topomaps are selected, we get more discriminative information. The results of the best case for all the selected topomaps with two different block divisions have been shown in a comprehensive way in Figures 9 and 10. Generally, the more the number of topomaps and the number of blocks, the more the accuracy of 2D and 3D.

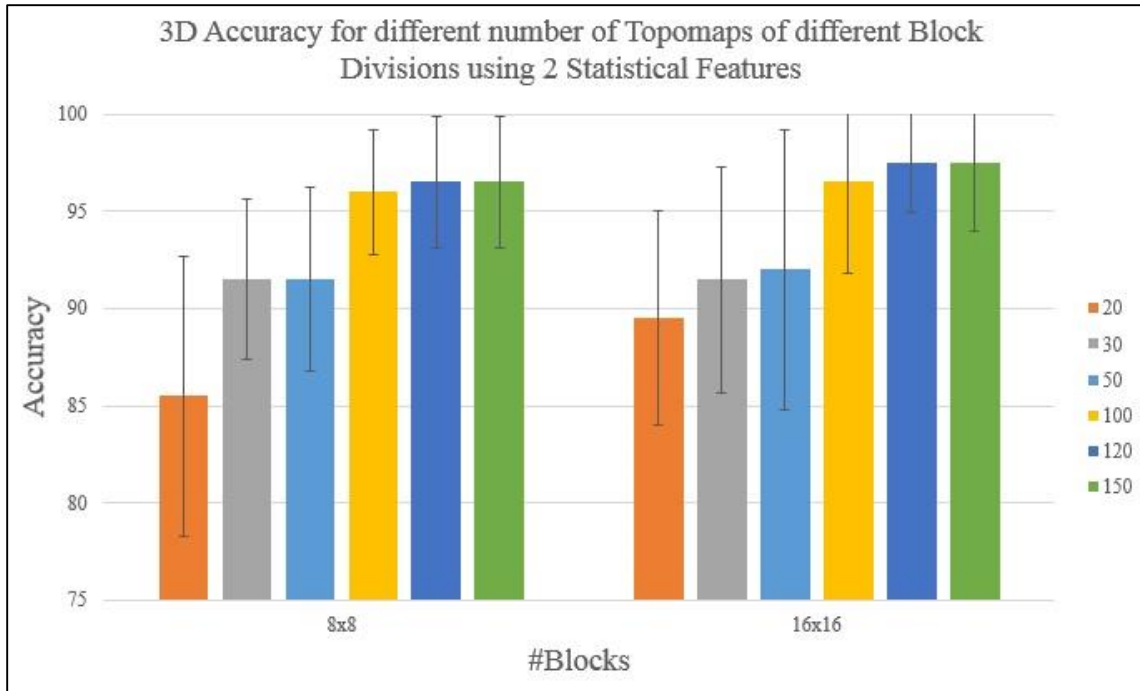


Figure 9. The best case accuracies in case of STM for different selected topomaps for 2D educational material.

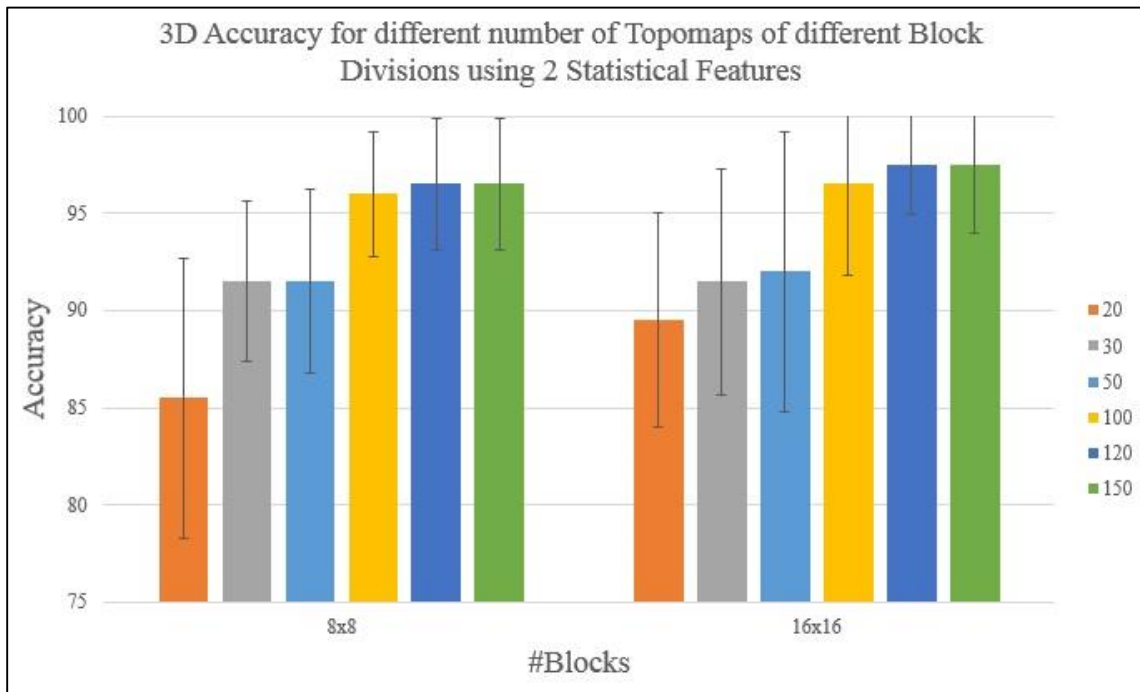


Figure 10. The best case accuracies in case of LTM for different selected topomaps for 3D educational material.

### **3.4 STM+LTM**

Now, we merged the 200 samples for STM with the 200 samples for LTM together and extract the statistical features. The average prediction accuracy is still excellent; we got an average accuracy of 97% for both 2D and 3D with the use of five features. This indicates that the proposed system is robust and is not affected by the size of the data.

## **4. Which is better 2D or 3D?**

In order to come to any conclusion about the effectiveness of 2D or 3D educational contents for learning and memory retention/recall, we tested the hypothesis described in Section 3. For this purpose, we need enough samples of prediction accuracy values. To achieve this, we executed the systems five times with 10-fold cross validation using the best parameters and every time randomizing the datasets so that we can get average accuracy values for 50 samples having different combinations of true and false memories. After getting 50 accuracy values for the systems developed for 2D and 3D, we used SPSS for hypothesis testing. .

Using SPSS software, an independent t-test was applied depending on the assumption that normality is acceptable. The results for 2D and 3D groups are approximately same, i.e., in the case of 2D, M is 96.6 and SD is 3.7, while in 3D case M is 97.2 and SD is 3.3. The value of t is -0.847 and p is 0.399, which is greater than 0.05. Therefore, the null hypothesis is rejected and it can be said that statistically 2D and 3D contents are same for learning and memory recall in the case of STM. Similar tests were performed in the case of LTM and same can also be said that 2D and 3D contents are statistically the same for learning and memory recall. In the case of 2D, M is 99.4 and SD is 1.6 and for 3D M is 99.3 and SD is 2.0, while the value of p is 0.787 and t is 0.271

The observed results did not contradict with the results based on the answers to the questions of the 2D and 3D content for both STM and LTM. If we analyze the results, we find that in case of STM, the difference between the correct answers is 23 for 2D and 3D group, which is equal to 3% of all answers, whereas this difference is only 3 (i.e. 0.468%) in case of LTM. This percentage is not significantly different if we take into account that some correct answers may be due to guess simply and not true memory. This supports our finding that 2D and 3D educational contents are approximately same for learning, memory retention and recall. These findings are similar to the previous subjective findings on the effectiveness of 2D and 3D contents on learning and knowledge acquisition [11. 14] and education learning processes [21] without using EEG technology.

## **5. Conclusion**

We studied the effectiveness of 2D and 3D educational contents on learning and memory retention/recall using EEG brain signals. For this purpose, we proposed pattern recognition systems for predicting true and false memories and then used them for assessing the effects of 2D and 3D educational contents on learning and memory retention/recall by analyzing the brain behavior during learning and memory recall tasks using EEG signals.

For modeling pattern recognition systems, first, we collected the data. Sixty-eight healthy volunteers participated in data collection. Two groups corresponding to 2D and 3D were formed, which were based on the participants' ages and their knowledge. The data collection was done in two phases: learning and memory recall. During the learning phase, depending upon the group, the participants watched 2D or 3D contents for durations of 8 to 10 minutes. In order to check memory recall, two retention periods were used, i.e., 30 minute period in the case of STM and 2 months period in the case of LTM. Each participant answered 20 MCQs which were same for both 2D and



3D groups and were about the learned contents. During the recall process, EEG signals were recorded.

For the pattern recognition systems, we introduced a simple and robust feature extraction technique that first converts the EEG signals (corresponding to the response of a question) into topomaps, selects the most discriminative topomaps and then captures the localized texture information from the selected topomaps. For classification, we employed SVM with RBF kernel. The proposed systems can reliably predict true and false memories employing the direct brain activations in case of STM and LTM. We developed separate systems with the same architecture and configuration for 2D and 3D educational materials. As each system encodes the brain activations during memory retention/recall tasks, the performance of the system can be analyzed to assess the impact of 2D and 3D educational materials. Statistical analysis reveals that there is no significant difference between the effects of 2D and 3D educational materials on learning and memory retention/recall. Our findings are in accordance with the previous studies (without using direct brain activations) on the effects of 2D and 3D materials. However, our approach is based on the pattern recognition systems. As such our research has the added advantage of proposing pattern recognition systems, which can be used to predict true and false memories and can be adopted to assess the learning levels of students. Though there is no difference in the effectiveness of 2D and 3D educational contents on learning and memory retention/recall from the performance perspective, different brain regions can be activated for 2D and 3D educational material. To investigate which brain regions are activated for 2D and 3D educational materials is the subject for future work.

### ***Acknowledgement***

This project was supported by NSTIP strategic technologies programs, grant number 12-INF2582-02 in the Kingdom of Saudi Arabia.

### ***References***

- [1] Mayer, R.E. and R. Moreno, *A cognitive theory of multimedia learning: Implications for design principles*. Journal of Educational Psychology, 1998. **91**(2): p. 358-368.
- [2] Teplan, M., *Fundamentals of EEG measurement*. Measurement science review, 2002. **2**(2): p. 1-11.
- [3] Niedermeyer, E. and F.L. da Silva, *Electroencephalography: basic principles, clinical applications, and related fields* 2005: Lippincott Williams & Wilkins.
- [4] Carrier, L.M., et al., *Pathways for Learning from 3D Technology*. International Journal of Environmental & Science Education, 2012. **7**(1).
- [5] Cockburn, A. and B. McKenzie, *Evaluating spatial memory in two and three dimensions*. International Journal of Human-Computer Studies, 2004. **61**(3): p. 359-373.
- [6] Price, A. and H.-S. Lee, *The effect of two-dimensional and stereoscopic presentation on middle school students' performance of spatial cognition tasks*. Journal of Science Education and Technology, 2010. **19**(1): p. 90-103.
- [7] Tavanti, M. and M. Lind. *2D vs 3D, implications on spatial memory*. in *Information Visualization, 2001. INFOVIS 2001. IEEE Symposium on*. 2001. IEEE.
- [8] Mukai, A., et al., *Effects of Stereoscopic 3D Contents on the Process of Learning to Build a Handmade PC*. Knowledge Management & E-Learning: An International Journal (KM&EL), 2011. **3**(3): p. 491-506.

- [9] Huk, T. and C. Floto. *Computer-animations in education: the impact of graphical quality (3D/2D) and signals*. in *World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*. 2003.
- [10] Rias, R.M. and W. Khadijah Yusof. *Animation and prior knowledge in a multimedia application: A case study on undergraduate computer science students in learning*. in *Digital Information and Communication Technology and it's Applications (DICTAP), 2012 Second International Conference on*. 2012. IEEE.
- [11] Wang, H.-C., C.-Y. Chang, and T.-Y. Li, *The comparative efficacy of 2D-versus 3D-based media design for influencing spatial visualization skills*. *Computers in Human Behavior*, 2007. **23**(4): p. 1943-1957.
- [12] Perez, F., S. Vanegas, and P. Cortes. *Evaluation of learning processes applying 3D models for constructive processes from solutions to problems*. in *Education Technology and Computer (ICETC), 2010 2nd International Conference on*. 2010. IEEE.
- [13] Amin, H.U., et al. *Brain activation during cognitive tasks: An overview of EEG and fMRI studies*. in *Biomedical Engineering and Sciences (IECBES), 2012 IEEE EMBS Conference on*. 2012. IEEE.
- [14] Malik, A.S., et al. *Investigating brain activation with respect to playing video games on large screens*. in *Intelligent and Advanced Systems (ICIAS), 2012 4th International Conference on*. 2012. IEEE.
- [15] Malik, A.S., et al. *Disparity in brain dynamics for video games played on small and large displays*. in *Humanities, Science and Engineering Research (SHUSER), 2012 IEEE Symposium on*. 2012. IEEE.
- [16] Babiloni C, Babiloni F, Carducci F, Cappa SF, Cincotti F, Del Percio C, et al. *Human cortical responses during one-bit short-term memory. A high-resolution EEG study on delayed choice reaction time tasks*. *Clin Neurophysiol* 2004; 115: 161-170.
- [17] Klimesch, W., *EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis*. *Brain research reviews*, 1999. **29**(2): p. 169-195.
- [18] CHEN, X.-l., B. YAO, and D.-y. WU, *Application of EEG Non-linear Analysis in Vision Memory Study*. *Chinese Journal of Rehabilitation Theory and Practice*, 2006. **6**: p. 012.
- [19] Amin, H.U., et al. *Brain Behavior in Learning and Memory Recall Process: A High-Resolution EEG Analysis*. in *The 15th International Conference on Biomedical Engineering*. 2014. Springer.
- [20] Brian A. Coffman, Vincent P. Clark, Raja Parasuraman, "Battery powered thought: Enhancement of attention, learning, and memory in healthy adults using transcranial direct current stimulation", *NeuroImage* 85 (2014) 895–908
- [21] Bon-Mi Gu, Hedderik van Rijn, Warren H. Meck, "Oscillatory multiplexing of neural population codes for interval timing and working memory", *Neuroscience and Biobehavioral Reviews* 48 (2015) 160–185
- [22] Elizabeth L Johnson and Robert T Knight, "Intracranial recordings and human memory", *Current Opinion in Neurobiology* 2015, 31:18–25
- [23] Liang-Tien Hsieh, Charan Ranganath, "Frontal midline theta oscillations during working memory maintenance and episodic encoding and retrieval", *NeuroImage* 85 (2014) 721–729
- [24] Bamatraf, S., et al. *A System based on 3D and 2D Educational Contents for True and False Memory Prediction using EEG Signals*. in *7th Annual International IEEE EMBS Conference on Neural Engineering*. 2015. Montpellier, France: IEEE.
- [25] *Eureka.in*. Available from: <http://www.designmate.com/eurekain.html>.
- [26] Gratton, G., M.G. Coles, and E. Donchin, *A new method for off-line removal of ocular artifact*. *Electroencephalography and clinical neurophysiology*, 1983. **55**(4): p. 468-484.

- [27] Delorme, A. and S. Makeig, *EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis*. Journal of neuroscience methods, 2004. **134**(1): p. 9-21.
- [28] Chang, C.-C. and C.-J. Lin, *LIBSVM: a library for support vector machines*. ACM Transactions on Intelligent Systems and Technology (TIST), 2011. **2**(3): p. 27.
- [29] Kohavi, R. *A study of cross-validation and bootstrap for accuracy estimation and model selection*. in *IJCAI*. 1995.32.
- [30] D. P. Subha, P. K. Joseph, R. Acharya, and C. M. Lim, "EEG signal analysis: a survey," Journal of Medical Systems, vol. 34, pp. 195-212, 2010.
- [31] W. Lian, R. Talmon, H. Zaveri, L. Carin, and R. Coifman, "Multivariate time-series analysis and diffusion maps," Signal Processing, vol. 116, pp. 13-28, 2015.
- [32] M. H. Díaz, F. M. Córdova, L. Cañete, F. Palominos, F. Cifuentes, C. Sánchez, and M. Herrera, "Order and Chaos in the Brain: Fractal Time Series Analysis of the EEG Activity During a Cognitive Problem Solving Task," Procedia Computer Science, vol. 55, pp. 1410-1419, 2015.
- [33] S. Akdemir Akar, S. Kara, S. Agambayev, and V. Bilgiç, "Nonlinear analysis of EEGs of patients with major depression during different emotional states," Computers in Biology and Medicine, vol. 67, pp. 49-60, 2015.
- [34] Z. Li, X. Jin, and X. Zhao, "Drunk driving detection based on classification of multivariate time series," Journal of Safety Research, vol. 54, pp. 61.e29-64, 2015.
- [35] U. R. Acharya, F. Molinari, S. V. Sree, S. Chattopadhyay, K.-H. Ng, and J. S. Suri, "Automated diagnosis of epileptic EEG using entropies," Biomedical Signal Processing and Control, vol. 7, pp. 401-408, 2012.
- [36] Hsu, Chih-Wei, Chih-Chung Chang, and Chih-Jen Lin. "A practical guide to support vector classification." (2003).
- [37] Joachims, Thorsten. Text categorization with support vector machines: Learning with many relevant features. Springer Berlin Heidelberg, 1998.
- [38] Chapelle, Olivier, Patrick Haffner, and Vladimir N. Vapnik. "Support vector machines for histogram-based image classification." Neural Networks, IEEE Transactions on 10.5 (1999): 1055-1064.
- [39] Cortes, Corinna, and Vladimir Vapnik. "Support-vector networks." Machine learning 20.3 (1995): 273-297.
- [40] Theodoridis, S. and K. Koutroumbas, Chapter 5 - Feature Selection, in Pattern Recognition (Fourth Edition), S. Theodoridis and K. Koutroumbas, Editors. 2009, Academic Press: Boston. p. 261-322.