Cognitive Analysis of Mental States of People According to Ethical Decisions Using Deep Learning Approach

[1]Dipayan Dewan, [1]Lidia Ghosh, [2]Biswadeep Chakraborty, [1]Abir Chowdhury, [1]Amit Konar, and [3]Atulya K. Nagar
[1]Artificial Intelligence Laboratory, ETCE Department, Jadavpur University, Kolkata, India
[2]Dept. of Electrical and Computer Engineering, Georgia Institute of Technology, USA
[3]Dept. of Math and Computer Science, Liverpool Hope University, UK
diiipayan93@gmail.com, lidiaghosh.bits@gmail.com, biswadeep.ju.etce@gmail.com, abir.chwdhry@gmail.com, konaramit@yahoo.co.in and nagar@hope.ac.uk

***Abstract*—Human behavior is a complex action which has provoked the thoughts of many people for a long time. However, very little about the cognitive aspect of an individual person's personality is known till the date. In order to differentiate people with significant differences in personality from their brain responses, we must at first be able to classify them based on their thought processes. From the viewpoint of classical ethics, people can broadly be classified into two main classes, namely, Categorical and Consequentialist. In this paper, we conduct several experiments where the subjects experience various ethical dilemmas and their ethical values in response to the presented stimuli are investigated through a question-answer session. The brain responses of the subjects are acquired using electroencephalography, which is then fed to an attention based parallel Convolutional Bi-directional Long Short-Term Memory (AConvBi-LSTM-NN) network with an ultimate aim to classify people into two above mentioned categories.**

Apart from the application point of view, the novelty of the paper lies in representing the EEG time-series into a sequence of multispectral 2D images which contain the spatial information of the acquired signal. The spectral information, along with the EEG time series (temporal information) are then used to train the proposed convolutional Bi-LSTM Network. The experimental results demonstrate promising results in classifying people based on their ethical values from their brain responses, with high classification accuracy. This provides scope for a new direction of research which can be further explored.

Keywords— BCI, electroencephalography, Deep Learning, Long Short-Term Memory Network, Ethical Decision

I. INTRODUCTION

Human behavior has been an intriguing aspect of human cognition and much has gone into the research of finding out why people behave in a particular manner. Though some great strides have been made in this field, however, the biological underpinnings of ethical decision making using brain signal processing are still unexplored. Many of the daily decisions taken by us fall within the category of consequentialist ethics. Consequentialism is a theory in the domain of moral philosophy which is based on whether something is right by looking into what the consequences for the action are. For example, though lying is a wrong deed, if done so to save someone's life is considered a right thing according to the consequentialist theory. Consequentialism focuses decision making upon the potential outcomes of an action where the outcome, coupled with the intent of the action, becomes the standard for morality.

The paper proposes a novel idea of classifying people according to their ethical decision making using their brain responses. By consequentialism theory, people can be broadly classified into 2 types; consequential and categorical. Although there exist several non-invasive brain signal acquisition modalities, we select electroencephalography (EEG) to acquire electrical signals from the human scalp, due to its high temporal resolution, cost-effectiveness and portability measures.

The most challenging research component of the work lies in the feature extraction and classifier design. There exist several classification techniques involving EEG signals as input to the classifier [1]-[3]. In this paper, we propose an attention based Convolutional Bi-directional Long Short-Term Memory (Bi-LSTM) network-based architecture for the present classification task. In the light of Artificial Intelligence era in this decade, Deep Learning models have immensely found its way into the Brain-Computer Interfacing (BCI) domain for its greater ability to learn complex correlation among the acquired brain responses of different lobes, thus offering provisions for automated generation of independent (correlation-free) features [4]. Among the well-known deep learning algorithms, Convolutional Neural Network (CNN) [5], Recurrent Neural Network (RNN) [6] and Long Short-term Memory (LSTM) [7] network have shown promising results in time-series prediction, image processing as well as classification. Although there are traces of Deep Learning techniques in electroencephalographic (EEG) time-series data analysis [8], there is hardly any work on classification problem using EEG time-series data and EEG topographic maps as image data jointly as input to the deep neural network classifiers. This paper fills the void. The novelty of the paper lies in designing an attention-based convolutional Bi-LSTM network which is capable of learning both the spatial and temporal information of the EEG signal and producing

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

features that are less sensitive to the subjective variations and the EEG artifacts.

The classical CNN gives the best performance dealing with image data as input to the network and used in almost every type of image dataset classification task. The challenge here is to represent EEG time-series data into 2D multispectral images, which can provide the similar efficacy. The novelty of the paper lies in converting the EEG time-series into 2-dimensional images by suitable projection and interpolation techniques and the series of 2D images, thus obtained, are then fed to CNN for obtaining the best spatial feature. Besides that, the EEG time-series data are fed to the attention-based Bi-LSTM network which is responsible for extracting the temporal feature. Features obtained from the respective networks then fuse together. Experiments undertaken show that the proposed architecture, named as AConvBi-LSTM-NN provides the highest accuracy for the present classification problem as compared with the state-of-the-art classifier algorithms.

The paper is divided into 5 sections. The proposed principles & methodology of this paper are discussed in section II. Section III deals with the Experimental Framework, EEG data acquisition during the ethical decision-making task, and the results obtained from the experiment are discussed in section IV. Conclusions are listed in section V.

II. Principles and Methodologies

The present paper proposes a novel attention based parallel Convolutional Bi-LSTM Neural Network (AConvBi-LSTM-NN), which integrates the classical image-dataset classification approach with the flavor of EEG time-series prediction. Convolutional Neural Network (CNN) [9] is a powerful tool for image data classification. The existing literature on image data classification (MNIST, Imagenet, CIFAR10, CIFAR100) shows that CNN outperformed the other state-of-the-art classifiers by a significant margin. Additionally, Bi-LSTM has the efficacy to extract temporal features of EEG time-series data [10]. Therefore, the authors are taking a keen interest to integrate the joint benefit of CNN and Bi-LSTM to capture both the spatial and temporal characteristics of EEG time-series data. The EEG time-series are first converted into a sequence of 2-dimensional multi-spectral images by a suitable approach described below, which captures spatial information and then fed to the CNN to extract the spatial characteristics. The attention-based Bi-LSTM is parallelly trained with the sequences of EEG signals to extract the temporal characteristics. A block-diagram representation of the proposed approach is shown in Fig. 1.

A. *Generating 2D multi-spectral images from EEG time-series*

EEG captures the electrical signals generated from the different brain regions, with the help of several electrodes placed around the cortex. Although EEG signals have a spatial dimension, most of the significant EEG features are extracted in the time-frequency domain as EEG provides high temporal resolution and poor spatial resolution [11]. For the present classification problem, EEG signals are acquired from the subjects' brain when they engage themselves in making their ethical decisions. As there is evidence of theta (4-8Hz), alpha (8-14Hz) and beta(14-30Hz) activation for decision making [12], the Short-time Fourier Transform (STFT) [13] of the EEG signals in the three different frequency bands are calculated for each trial to obtain the respective frequency

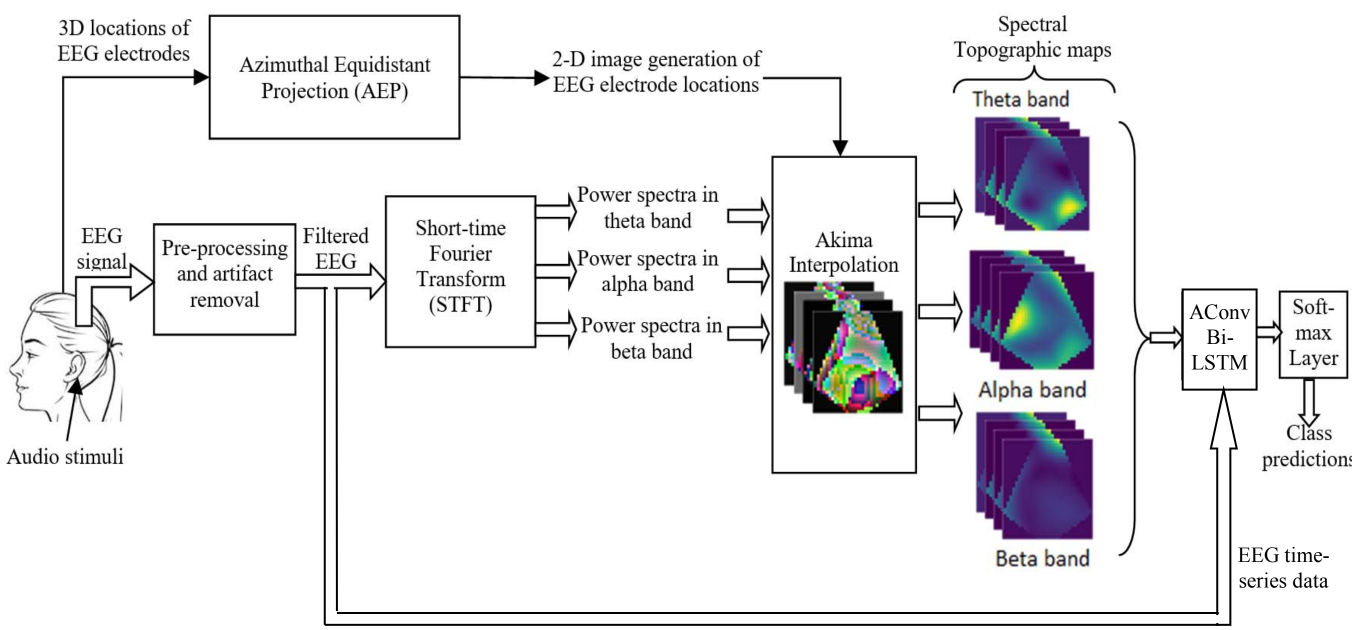


Fig. 1. Block-diagram Representation of the proposed framework

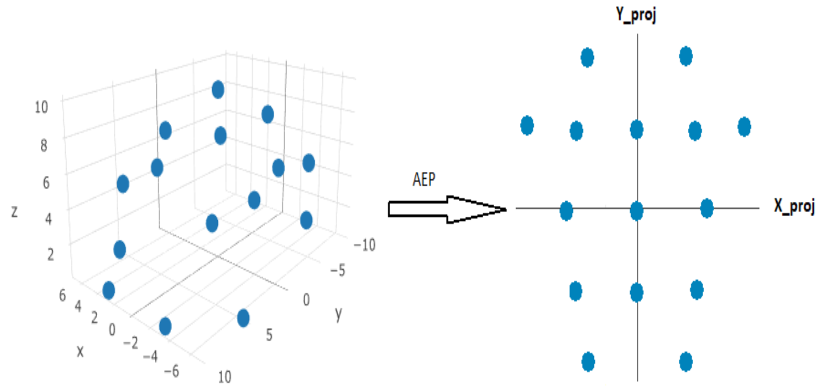


Fig. 2. 2D projection of 3D electrode location using AEP

spectrum.

In addition, as the EEG electrodes are scattered over the 3-dimensional space of the human scalp, conversion from 3D space to 2D surface electrode representation requires a suitable projection method. While projecting the 3D electrode positions into a 2D one, significant emphasis should be given to preserving the relative distance between the EEG electrodes. This is realized by a popular polar projection technique, named as *Azimuthal Equidistance Projection* (AEP) [14]. The most important property of AEP is that it preserves the distance and direction of each electrode intact from the center as it was in the 3D surface. Fig. 2 shows the electrode location in 3D space and its 2D projection respectively.

The frequency spectrum of each band is then interpolated with a 2D EEG electrode space to form 2D topographical images of the EEG time-series [15], [16]. Thus, the temporal information of the EEG time-series is represented in the form of sequential images.

Akima interpolation [17] method is applied here to estimate the value in between the two electrodes to get a smooth image of the topographical map. Repeating this procedure for three different frequency bands we are working with, we get three different topographical images (3 channels) for a single trial. These three images can be interpreted as the three channels for color images and are fed to the CNN.

B. Classification using Attention-based Convolutional Bi-directional Long Short-Term memory Neural Network

Here we propose a novel attention based parallel Convolutional Bi-LSTM deep neural network architecture as shown in Fig. 3. The motive behind developing this architecture is to extract the temporal and the spatial feature of the EEG-signal separately and then fuse those features to classify the data. The images generated from the EEG signals are fed to a 2D-CNN model to extract the spatial feature. Simultaneously the Bi-LSTM network is employed to extract the temporal feature from the EEG-signal itself.

1) **Attention-based CNN:** Here we use the attention-based CNN (ABCNN) as described in [18]. ABCNN is employed to capture the spatial feature of the images generated from the EEG signals. These images are generated by processing the 21 channel EEG data into a frame of a one-second window and then a 2D projection of scalp is made from this frame. The model here in this experiment is implemented using tensorflow and Keras [19] library in Python. In classical CNN, we use 4 convolution layers and 3 pooling layers. Rectifier linear unit (ReLU) is used as an activation function in the convolution layer. Attention is added in the output of each convolution layer to reweight the convolution output by influencing average pooling through direct attention weighting, as shown in Fig. 3.

We take all the images generated from the EEG signals as input images. The spatial information of the input EEG images is then extracted in the first convolution layer of CNN using 32 filters of 3 × 3 dimensions. This filter size is fixed for all the convolution layers in this model with a stride of 1 pixel. We are using zero-paddings to keep the dimension of the input image intact after the convolution operation. Each of the 4 convolution layers is followed by a MaxPool layer of 2 × 2 window size with a stride of 1 pixel. The length of the convolution filter is increased every time it goes into a deeper layer. At the end of the convolution layer, a fully connected or dense layer is applied to create the final spatial feature vector.

2) **Proposed Attention-based Bi-LSTM Network:** Here, a Bi-LSTM network is employed to extract the temporal feature from the EEG signals. Bi-LSTM network contains two baseline LSTM units [20] to process the temporal sequence of information in two opposing directions simultaneously. One LSTM unit processes the input time-sequence in the forward direction i.e., from first time-point to the end time-point and the other LSTM unit processes the same input sequence in the backward direction, as shown in Fig. 3. The final output of the Bi-LSTM network is determined by combining the output of the two baseline-LSTM units together. Each baseline LSTM unit consists of four major components: memory cell, input gate, forget gate and output gate [27]. Suppose, the input to the Bi-LSTM be $S = [\mathbf{s}_1, \mathbf{s}_2, ...\mathbf{s}_{t-1}, \mathbf{s}_t, \mathbf{s}_{t+1}, ..., \mathbf{s}_T]$, output set be $Y = [\mathbf{y}_1, \mathbf{y}_2, ...\mathbf{y}_{t-1}, \mathbf{y}_t, \mathbf{y}_{t+1}, ..., \mathbf{y}_T]$ for $t = [1, T]$ time-steps.

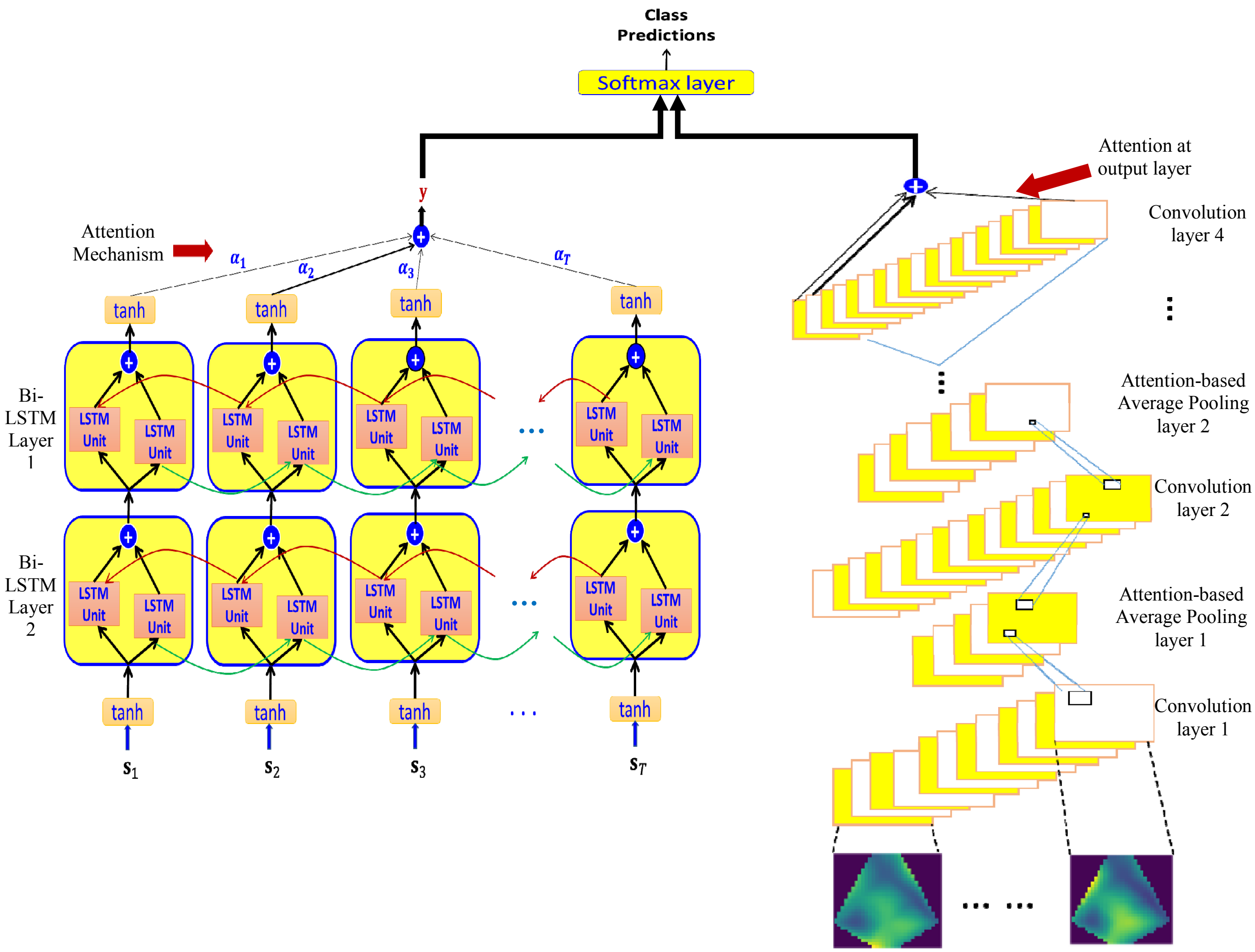


Fig. 3. Architecture of the proposed AConvBi-LSTM-NN

Each baseline LSTM unit processes the same input S in the following manner.

Using the current input $\mathbf{s}_t$ and the output of the previous hidden layer, $\mathbf{h}_{t-1}$, the memory cell of each baseline LSTM decides which information of the previous state needs to be kept or which to forget using the forget gate $\mathbf{f}_t$, given by (1), at time t.

$$\mathbf{f}_t = sig(\mathbf{W}_{fs}.\mathbf{s}_t + \mathbf{W}_{fh}.\mathbf{h}_{t-1} + \mathbf{b}_f), \quad (1)$$

where, $\mathbf{W}_{ab}$ denotes the weight matrix from layer a to layer b and $\mathbf{b}_f$ is the forget gate bias. Depending on the input gate information $\mathbf{i}_t$, following (2) and previous cell state $\mathbf{c}_{t-1}$, the memory cell computes how much information should be stored in the current state $\mathbf{c}_t$ into two steps by (3) and (4).

$$\mathbf{i}_t = sig(\mathbf{W}_{is}.\mathbf{s}_t + \mathbf{W}_{ih}.\mathbf{h}_{t-1} + \mathbf{b}_i) \quad (2)$$

$$\hat{\mathbf{c}}_t = \tanh(\mathbf{W}_{cs}.\mathbf{s}_t + \mathbf{W}_{ch}.\mathbf{h}_{t-1} + \mathbf{b}_c) \quad (3)$$

$$\mathbf{c}_t = \mathbf{i}_t.\hat{\mathbf{c}}_t + \mathbf{f}_t.\mathbf{c}_{t-1} \quad (4)$$

Finally, the output gate $\mathbf{o}_t$ decides the information flow from the current cell to the next cell and is represented by the hidden state information $\mathbf{h}_t$. The definition of each parameter of baseline LSTM cell is given below:

$$\mathbf{o}_t = sig(\mathbf{W}_{os}.\mathbf{s}_t + \mathbf{W}_{oh}.\mathbf{h}_{t-1} + \mathbf{b}_o), \quad (5)$$

$$\mathbf{h}_t = \mathbf{o}_t.\tanh(\mathbf{c}_t). \quad (6)$$

In the above equations (5)-(6), sigmoid (*sig*) and tanh are the two non-linear functions, defined as $sig(f) = \frac{1}{1+e^{-f}}$ and $\tanh(f) = \frac{2}{2+e^{-2f}} - 1$. The weight matrices $\mathbf{W}_f$, $\mathbf{W}_i$, $\mathbf{W}_c$, $\mathbf{W}_o$ and the biases $\mathbf{b}_f, \mathbf{b}_i, \mathbf{b}_c, \mathbf{b}_o$ are trained by back-propagation through time algorithm.

Unlike baseline LSTM, in Bi-LSTM, there are forward sequences $\vec{\mathbf{h}}_t$, defined by (7) and backward sequences $\overleftarrow{\mathbf{h}}_t$, defined by (8) in the hidden layers at time *t*.

$$\vec{\mathbf{h}}_t = \tanh(\mathbf{G}_{\vec{h}}.\mathbf{s}_t + \mathbf{W}_{\vec{h}}.\vec{\mathbf{h}}_{t-1} + \mathbf{b}_{\vec{h}}) \tag{7}$$

$$\overleftarrow{\mathbf{h}}_t = \tanh(\mathbf{G}_{\overleftarrow{h}}.\mathbf{s}_t + \mathbf{W}_{\overleftarrow{h}}.\overleftarrow{\mathbf{h}}_{t-1} + \mathbf{b}_{\overleftarrow{h}}) \tag{8}$$

Here, **G** is the weight matrix connecting the input layer and the hidden layer. Simultaneously the output layer of Bi-LSTM, $\mathbf{h}_t$, at time *t*, can be defined as the composite effect of both the forward and backward sequences and given by the equation below.

$$\mathbf{h}_t = \tanh(\mathbf{U}_{\vec{h}}.\vec{\mathbf{h}}_t + \mathbf{U}_{\overleftarrow{h}}.\overleftarrow{\mathbf{h}}_t + \mathbf{b}_y), \tag{9}$$

where, **U** is the weight matrix connecting the hidden layer and the output layer. In the proposed Bi-LSTM network, we use 2 Bi-LSTM layers with 128 hidden units and 1 drop-out layer.

Attention Mechanism in the Bi-LSTM Network: Integrating the Attention mechanism with the LSTM enables the network to pay attention on certain time-points of the input information with maximum discrimination. In traditional Bi-LSTM network, the last hidden layer is considered as the final output. On the other hand, in the attention mechanism, the last hidden layer outputs are multiplied by the trainable weights, thus seizing features with maximum discrimination at a given time-point. In the proposed scheme, an attention layer is added at the output of the Bi-LSTM network to capture the output of the each hidden state of the Bi-LSTM network, where the attention weights are defined as

$$\alpha_t = \frac{\exp(\mathbf{h}_t)}{\sum_{t=1}^{T} \exp(\mathbf{h}_t)}. \tag{10}$$

The output of the attention layer thus can be defined as

$$\mathbf{y} = \sum_{t=1}^{T} \alpha_t.\mathbf{h}_t \tag{11}$$

and a dense layer(of the same dimension as applied in the CNN network) is applied to obtain the final temporal feature vector.

Finally the required Spatio-temporal feature is achieved by fusing the feature vector acquired from the respective network. This fused Spatio-temporal feature vector then connected with a dense layer of output dimension. In this layer, a softmax classifier [21] is used to classify the desired class labels (here, two classes: Categorical and Consequentialist) based on probability measurements. Softmax layer gives the predicted class for a given input by checking the maximum probability distribution from the output layer. We use categorical cross-entropy as a loss function and Adam optimizer [22] with learning rate 0.0001 to train the model.

C. Decision-Based Classification System

According to the types of moral decision a person makes, people can be broadly classified into two categories viz., Categorical and Consequential. Categorical decision making indicates morality in certain duties and rights. On the other hand, according to the consequentialist theory, people who fall in that category indicate morality in the consequences of the act. During classifier training, the proposed AConvBi-LSTM-NN is trained with the EEG responses of the two different categories with an aim to produce the above mentioned two classes.

III. Experimental Framework

The section briefly narrates about the experimental set-up, participants and the experiments undertaken.

A. Participants

Eighteen healthy subjects (8 female and 10 male), aged 20 to 26 years and all attending college or graduated, participated in the experiment. The subjects had no medical history of neurological, psychiatric and/or motor diseases and had a normal or corrected-to-normal vision. Participants provided written consent prior to participation and the experiments were conducted according to the Helsinki protocol [22].

B. EEG Data Acquisition

The experiments are performed using a 21-channel Nihon Kohden system, with a sampling frequency of 500 Hz, with 21 electrodes positioned according to the 10-20 system arranged in Fp1, Fp2, F7, F3, Fz, F4, F6, T1, T2, T3, T4, T5, T6, C3, Cz, C4, P3, Pz, P4, O1, O2 electrode locations. The reference electrode for left hemisphere is left earlobe electrode A1and right earlobe electrode A2 is used as reference for right hemisphere. The ground electrode is placed in Fpz.

C. Ethical Decision Task

The EEG signals are acquired from the subjects' brain while they are presented with an audio stimulus of 10 minutes duration, which narrates a story of ethical dilemmas. Three different scenarios are used as audio stimuli, among which only 2 are presented here due to space-economy. After each audio stimulus presentation, a question-answer session is carried out to judge the subjects' value of ethics. A rest period of 10 minutes duration between consecutive audio stimuli is given to the subject to eliminate the effect of previous stimulations in the next one. The stimulus diagram with timing details is given in Fig. 4. The ethical questions are inspired from the first lecture of the Harvard course on Justice by Prof. Michael Sandel titled "THE MORAL SIDE OF MURDER".

The two scenarios used to make ethical decision are given below.

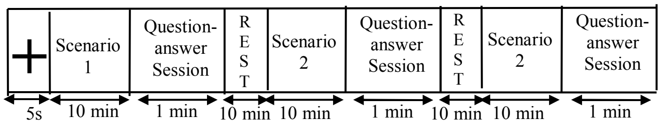


Fig. 4. Stimulus diagram

Scenario 1: In the first scenario, we gave the subjects the classic runaway railroad problem. In this scenario, there is a railway engine whose control is lost and is heading towards a junction. The subject is put into a hypothetical situation where he is an observer and has a lever in front of him which can change the course of the runaway railway engine. Moreover, the situation also entails that if the user does not switch the handle and the engine is let to run its free course, it will kill 5 people who are there on the tracks. On the other hand, if the lever is pulled, the tracks will change and as a result, only 1 person will die. The subject was then asked the first ethical question whether he would like to pull the lever or let it be if such a situation ever arose.

Scenario 2: In the next scenario, the point of view of the subject is changed. In this case, there is no lever which the subject has to pull and instead, he is watching the runaway train from a bridge overhead and is accompanied by a huge man. The situation is such that the subject knows that if he pushes that big man to fall on the tracks, he would die but stop the train and save the lives of the 5 people working on the tracks. The subject is then asked the question whether he or she would be likely to push the man. This, though logically has the same logic of killing one person to save five, has different results as the person is the directly involved in killing somebody. After this, question, the subject was farther asked what he would do if instead of pushing, he indeed had a lever which he could pull to kill the person like the former case.

D. Artifact Removal of Raw EEG data

To make the EEG response artifact-free, the raw EEG signal is passed through an Elliptical band pass filter (BPF) [23] of order 4 with the pass-band frequencies of 4 to 50 Hz. Next, the Independent Component Analysis (ICA) [24] has been performed to remove the eye-blinking artifacts.

IV. Results

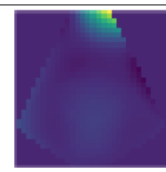
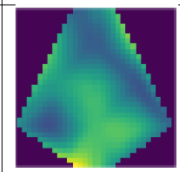
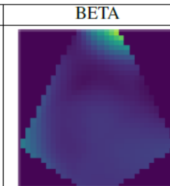
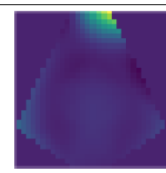
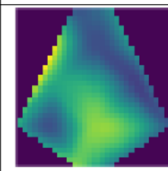
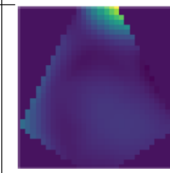
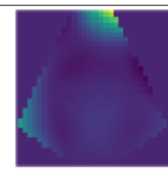
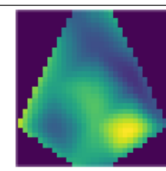
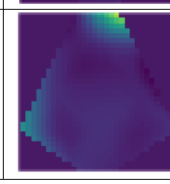
In this section, we discuss the performance evaluation of the proposed technique with the existing ones by using an experimental basis of analysis.

A. Frequency band selection for the Ethical Decision-making Task

This experiment deals with the frequency band selection for the present ethical decision making task. To perform the experiment, the topographical maps obtained for three different frequency bands: theta, alpha, and beta are compared. The topographic maps of theta, alpha, and beta frequency bands are presented in Table-I for 3 different experimental trials due to the space economy. In the colored topographic map, yellow color indicates the highest activation; green indicates moderate activation and blue indicates the lowest activation. It is evident from the table that, the alpha band shows the highest activation with more yellow colored regions as compared to the other two for all the 3 trials, which reveals the alpha band activation is strongly associated with ethical decision-making task.

Table-I

Comparison of Activations in the Theta, Alpha and Beta Bands

	THETA	ALPHA	BETA
TRIAL 1			
TRIAL 2			
TRIAL 3			

B. Performance Analysis of the proposed AConvBi-LSTM-NN Classifier

Computer simulations are performed to investigate the relative performance of the proposed AConvBi-LSTM-NN algorithm against the other existing standard classifiers with respect to the three performance metrics: percentage Classifier Accuracy (CA), the True Positive Ratio (TPR) and the False positive Ratio (FPR) [25]-[26]. It is evident from the table that the proposed AConvBi-LSTM-NN algorithm yields highest values of CA, TPR and FPR. Thus the proposed framework outperforms all the existing classifier algorithms for the present classification problem.

Table-II

Comparative Evaluation of the Proposed AConvBi-LSTM-NN Classifier

Classifiers	CA (%)	TPR	FPR
Proposed AConvBi-LSTM-NN	**87.79**	**0.84**	**0.89**
Bi-LSTM [20]	80.56	0.81	0.85
LSTM [27]	82.29	0.80	0.78
RNN [28]	78.44	0.82	0.75
CNN [9]	80.02	0.81	0.77
SVM [29]	76.59	0.76	0.72
kNN [30]	72.22	0.78	0.70
LDA [31]	74.02	0.79	0.75
QDA [32]	76.75	0.78	0.77

C. *Parameters of the Proposed Model and a Comparative study with other deep learning networks*

Since the proposed architecture integrates the characteristics of both CNN and Bi-LSTM, the parameters and their dimensions of each of the network module is tabulated in Table III. The overall learning rate and dropout rate of the AConvBi-LSTM-NN are 0.0001 and 0.2 respectively. The number of learning epochs of the network is 400 and the optimizer here chosen is ADAM.

TABLE-III
PARAMETERS OF ATTENTION-BASED CNN

Modules of AConvBi-LSTM-NN	Parameters	Values
CNN	Number of Convolution Layers	4
	Number of Attention-based average pooling layers	3
	Number of filters	32,80,200
	Filter size	4 × 4
	Pooling size	2 × 2
Bi-LSTM	Number of LSTM layers	300
	Size of Hidden Layer	128

TABLE-IV
COMPARATIVE STUDY WITH EXISTING DEEP LEARNING ALGORITHMS

Deep Learning Algorithms	Parameters	Input	Accuracy (%)
AConvBi-LSTM-NN	**2 layers of CNN and Bi-LSTM**	**Image + Time-series (Spatiotemporal)**	**87.79**
Parallel CNN-BLSTM [33]	2 layers of CNN and Bi-LSTM	Spatio-temporal	84.17
Cascade CNN-BLSTM [33]	2 layers of CNN and Bi-LSTM	Spatio-temporal	80.53
DMVST-NET [34]	2 layers of CNN and LSTM	Spatio-temporal	81.54
Bi-LSTM [20]	2000 hidden layers	Time-series (Temporal)	76.71
DCRNN [35]	2000 hidden layers	Spatio-temporal	80.03
ABCNN [18]	2000 hidden layers	Spatio-temporal	79.16
LSTM	2000 hidden layers	Temporal	80.01
RNN	2000 hidden layers	Temporal	73.84
CNN	2000 hidden layers	Spatio-temporal	72.17

TABLE V
STATISTICAL VALIDATION

Reference Algorithm: Proposed AConvBi-LSTM Classifier		
Algorithm for comparison	**Z**	**Acceptance/ Rejection of the null hypothesis**
Parallel CNN-BLSTM [33]	3.684	**Accept**
Cascade CNN-BLSTM [33]	4.29	Reject
DMVST-NET [34]	4.06	Reject
Bi-LSTM [20]	6.83	Reject
DCRNN [35]	7.19	Reject
ABCNN [18]	4.02	Reject
LSTM	5.38	Reject
RNN	6.99	Reject
CNN	8.17	Reject
SVM	12.67	Reject
kNN	17.92	Reject
LDA	16.09	Reject
QDA	14.88	Reject

A comparative study is undertaken here to evaluate the performance of the proposed network with the existing predominant deep learning algorithms, tested on the same dataset. Table IV includes the percentage accuracy of each algorithm along with the parametrs used. The table clearly explains the suprimacy of the proposed AConvBi-LSTM-NN over the others.

D. *Statistical Test*

The performance of the proposed classifier is statistically validated employing the well-known McNemar's test. To do so, the conventional z-score is computed for the two algorithms: the proposed AConvBi-LSTM-NN as the reference algorithm and any one of the state-of-the-art algorithms listed in Table V. If the estimated z-value is greater than the z-value (at 95% confidence level), then it can be stated that the difference in accuracy between the two classifier algorithms is statistically significant and is comparable with Parallel CNN-BLSTM only.

V. CONCLUSION

This paper examines a deep neural network based technique to differentiate between people with stark differences in personalities from their brain responses. Existing classification techniques mostly rely on the temporal characteristics of EEG signals but AConvBi-LSTM-NN takes into account both the temporal and spatial characteristics of brain EEG data, making it one of the more accurate classifiers for EEG analysis. Experimental results show promising results with regards to the classification accuracy of 87.79\% when compared to other state-of-the-art. The novel technique proposed by the authors may provide a novel approach for EEG-based analysis.

ACKNOWLEDGMENT

We thankfully acknowledged the fund provided by Ministry of Human Resource Development for RUSA-II project granted to Jadavpur University, India.

References

[1] A. Craik, Y. He, and J. L. Contreras-Vidal. "Deep learning for electroencephalogram (EEG) classification tasks: a review." *Journal of neural engineering* 16, no. 3 (2019): 031001.

[2] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger. "A review of classification algorithms for

EEG-based brain–computer interfaces: a 10 year update." *Journal of neural engineering* 15, no. 3 (2018): 031005.

[3] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi. "A review of classification algorithms for EEG-based brain–computer interfaces." *Journal of neural engineering* 4, no. 2 (2007): R1.

[4] X. Zhang, L. Yao, X. Wang, J. Monaghan, and D. Mcalpine. "A survey on deep learning based brain computer interface: Recent advances and new frontiers."*arXiv preprint arXiv:1905.04149* (2019).

[5] A. Supratak, H. Dong, C. Wu, and Y. Guo, "Deepsleepnet: A model for automatic sleep stage scoring based on raw single-channel eeg," IEEE.

[6] Kuanar, Shiba, Vassilis Athitsos, Nityananda Pradhan, Arabinda Mishra, and Kamisetty R. Rao. "Cognitive analysis of working memory load from EEG, by a deep recurrent neural network." In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2576-2580. IEEE, 2018.

[7] Alhagry, Salma, Aly Aly Fahmy, and Reda A. El-Khoribi. "Emotion recognition based on EEG using LSTM recurrent neural network." *Emotion* 8, no. 10 (2017): 355-358.

[8] Ghosh, Lidia, Sayantani Ghosh, Amit Konar, Pratyusha Rakshit, and Atulya K. Nagar. "Decoding of EEG Signals Using Deep Long Short-Term Memory Network in Face Recognition Task." In *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 477-483. IEEE, 2018.

[9] LeCun, Yann, and Yoshua Bengio. "Convolutional networks for images, speech, and time series." *The handbook of brain theory and neural networks* 3361, no. 10 (1995): 1995.

[10] Hu, Xinmei, and Qi Yuan. "Epileptic EEG Identification Based on Deep Bi-LSTM Network." In *2019 IEEE 11th International Conference on Advanced Infocomm Technology (ICAIT)*, pp. 63-66. IEEE, 2019.

[11] Gevins, Alan, Harrison Leong, Michael E. Smith, Jian Le, and Robert Du. "Mapping cognitive brain function with modern high-resolution electroencephalography." *Trends in neurosciences* 18, no. 10 (1995): 429-436.

[12] G. Deuschl, A. Eisen *et al.,* "Recommendations for the practice of clinical neurophysiology: guidelines of the international federation of clinical neurophysiology," 1999.

[13] Griffin, Daniel, and Jae Lim. "Signal estimation from modified short-time Fourier transform." *IEEE Transactions on Acoustics, Speech, and Signal Processing* 32, no. 2 (1984): 236-243.

[14] Anderson, Paul S. "An oblique, poly-cylindrical, orthographic azimuthal equidistant cartographic projection: its purpose, construction and theory." *Cartography* 8, no. 4 (1974): 182-186.

[15] Yue-Hei Ng, Joe, Matthew Hausknecht, Sudheendra Vijayanarasimhan, Oriol Vinyals, Rajat Monga, and George Toderici. "Beyond short snippets: Deep networks for video classification." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4694-4702. 2015.

[16] Bashivan, Pouya, Irina Rish, Mohammed Yeasin, and Noel Codella. "Learning representations from EEG with deep recurrent-convolutional neural networks." *arXiv preprint arXiv:1511.06448* (2015).

[17] Akima, Hiroshi. "A new method of interpolation and smooth curve fitting based on local procedures." *Journal of the ACM (JACM)* 17, no. 4 (1970): 589-602.

[18] Yin, Wenpeng, Hinrich Schütze, Bing Xiang, and Bowen Zhou. "Abcnn: Attention-based convolutional neural network for modeling sentence pairs." *Transactions of the Association for Computational Linguistics* 4 (2016): 259-272.

[19] Gulli, Antonio, and Sujit Pal. *Deep learning with Keras*. Packt Publishing Ltd, 2017.

[20] S. Hochreiter, and J. Schmidhuber, "Long Short-Term Memory", Journal Neural Computation, MIT Press, vol 9 , pp 1735-1780, Nov 1997.

[21] A. Graves, and J. Schmidhuber. "Framewise phoneme classification with bidirectional LSTM and other neural network architectures." *Neural Networks* 18, no. 5-6, pp: 602-610, 2005.

[22] World Medical Association. "World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects." *Bulletin of the World Health Organization* 79, no. 4 (2001): 373.

[23] Christiano, Lawrence J., and Terry J. Fitzgerald. "The band pass filter." *international economic review* 44, no. 2 (2003): 435-465.

[24] Hyvärinen, Aapo, and Erkki Oja. "Independent component analysis: algorithms and applications." *Neural networks* 13, no. 4-5 (2000): 411-430.

[25] M. Sokolova and G. Lapalme. "A systematic analysis of performance measures for classification tasks."*Information processing & management* 45, no. 4 (2009): 427-437.

[26] C.Ferri, J.Hernández-Orallo, and R. Modroiu. "An experimental comparison of performance measures for classification." *Pattern Recognition Letters* 30, no. 1 (2009): 27-38.

[27] Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." *Neural computation* 9, no. 8 (1997): 1735-1780.

[28] Mikolov, Tomáš, Martin Karafiát, Lukáš Burget, Jan Černocký, and Sanjeev Khudanpur. "Recurrent neural network based language model." In *Eleventh annual conference of the international speech communication association*. 2010.

[29] Suykens, Johan AK, and Joos Vandewalle. "Least squares support vector machine classifiers." *Neural processing letters*9, no. 3 (1999): 293-300.

[30] Dudani, Sahibsingh A. "The distance-weighted k-nearest-neighbor rule." *IEEE Transactions on Systems, Man, and Cybernetics* 4 (1976): 325-327.

[31] Ye, Jieping, Ravi Janardan, and Qi Li. "Two-dimensional linear discriminant analysis." In *Advances in neural information processing systems*, pp. 1569-1576. 2005.

[32] Zhang, Michael Q. "Identification of protein coding regions in the human genome by quadratic discriminant analysis."*Proceedings of the National Academy of Sciences* 94, no. 2 (1997): 565-568.

[33] Ma, Xiaolei, Jiyu Zhang, Bowen Du, Chuan Ding, and Leilei Sun. "Parallel architecture of convolutional bi-directional lstm neural networks for network-wide metro ridership prediction."*IEEE Transactions on Intelligent Transportation Systems* 20, no. 6 (2018): 2278-2288.

[34] Yao, Huaxiu, Fei Wu, Jintao Ke, Xianfeng Tang, Yitian Jia, Siyu Lu, Pinghua Gong, Jieping Ye, and Zhenhui Li. "Deep multi-view spatial-temporal network for taxi demand prediction." In *Thirty-Second AAAI Conference on Artificial Intelligence*. 2018.

[35] Li, Yaguang, Rose Yu, Cyrus Shahabi, and Yan Liu. "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting." *arXiv preprint arXiv:1707.01926* (2017).