Expert Systems With Applications 184 (2021) 115548

Contents lists available at ScienceDirect

Expert Systems With Applications

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

Motivation detection using EEG signal analysis by residual-in-residual convolutional neural network

Soham Chattopadhyay [a], Laila Zary [b], Chai Quek [b], Dilip K. Prasad [c,*]

[a] Department of Electrical Engineering, Jadavpur University, 12, CIT Rd, Bidhannagar Kolkata, 700054, India
[b] School of Computer Science & Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore, 639798, Singapore
[c] Department of Computer Science, UiT The Arctic University of Norway, Hansine Hansens Veg 18, Tromsø, 9019, Norway

ARTICLE INFO

Dataset link: https://codeocean.com/capsule/2144609/tree/v1, https://github.com/SohamChattopadhyayEE/RRCNN

MSC:
00-01
99-00

Keywords:
EEG
Motivation
Deep learning

ABSTRACT

While we know that motivated students learn better than non-motivated students but detecting motivation is challenging. Here we present a game-based motivation detection approach from the EEG signals. We take an original approach of using EEG-based brain computer interface to assess if motivation state is manifest in physiological EEG signals as well, and what are suitable conditions in order to achieve the goal? To the best of our knowledge, detection of motivation level from brain signals is proposed for the first time in this paper. In order to resolve the central obstacle of small EEG datasets containing deep features, we propose a novel and unique 'residual-in-residual architecture of convolutional neural network (RRCNN)' that is capable of reducing the problem of over-fitting on small datasets and vanishing gradient. Having accomplished this, several aspects of using EEG signals for motivation detection are considered, including channel selection and accuracy obtained using alpha or beta waves of EEG signals. We also include a detailed validation of the different aspects of our methodology, including detailed comparison with other works as relevant. Our approach achieves 89% accuracy in using EEG signals to detect motivation state while learning, where alpha wave signals of frontal asymmetry channels are employed. A more robust (less sensitive to learning conditions) 88% accuracy is achieved using beta waves signals of frontal asymmetry channels. The results clearly indicate the potential of detecting motivation states using EEG signals, provided suitable methodologies such as proposed in this paper, are employed.

1. Introduction

Motivation, i.e. the driving force behind an individual's actions, is an important aspect in individual's life, especially in education. Motivation is a pervasive and significant determinant that decisively affects the behaviour of students, educators and all actors in all educational levels (Pintrich & Schunk, 2002). It is therefore hardly surprising that motivation has long been an interesting topic in scientific research (Touré-Tillery & Fishbach, 2014). Various researchers have found that motivated students are likely to be more engaged, devote efforts and struggles with more challenging activities, persist longer at a certain activity and show a higher performance and better outcomes (Garris, Ahlers, & Driskell, 2002). Since motivation influences the learning process (Christophel, 1990), it is of importance to study the educational experience and understand what motivates one's behaviour in educational contexts (Elliot & Covington, 2001). According to Elliot and Covington (2001), "intervention programs and procedures that fail to take motivational considerations into account are destined for failure" (Elliot & Covington, 2001).

Various types of approaches have been made to motivate students by developing virtual class rooms (Olszewska, 2021) or by using educational robot-based learning systems (Chin, Hong, & Chen, 2014) and so on. These techniques intend to introduce a system which is more self-driven and less human dependent. Such cyber physical system based advanced techniques overcome several challenges of online learning, and bring about significant changes in educational, technological and societal means. But with the advent of these technologies like e-learning, several aspects of jettisons have come forward, which often cause significant dip in students' motivation level. The effect of e-learning on the motivation level of students has been discussed

The code (and data) in this article has been certified as Reproducible by Code Ocean: (https://codeocean.com/). More information on the Reproducibility Badge Initiative is available at https://www.elsevier.com/physical-sciences-and-engineering/computer-science/journals.

* Corresponding author.
E-mail addresses: chattopadhyaysoham99@gmail.com (S. Chattopadhyay), lailazary@gmail.com (L. Zary), ashcquek@ntu.edu.sg (C. Quek), dilip.prasad@uit.no (D.K. Prasad).

https://doi.org/10.1016/j.eswa.2021.115548
Received 22 February 2021; Received in revised form 1 July 2021; Accepted 1 July 2021
Available online 5 July 2021

in (Harandi, 2015; Nehme, 2010). Some contrasting conclusions reveal from the approaches mentioned above. Harandi (2015) emphasizes that these advanced learning techniques help a student to learn more effectively and in more motivated manner. Whereas, Nehme (2010) take more generic approach towards the affect of e-learning on students' motivation, and come up with the conclusion that in spite of many advantages, e-learning might cause serious anxiousness and significant fear of learning in students' minds, which can affect the motivation level of the students as well.

Many other educational approaches try to motivate students by implementing learning systems containing multimedia and interactivity such as games. This approach builds on the assumption that students can be motivated if the pedagogical instructions are engaging and uses recent learning content technologies. However, this is not always the case. Motivation is not always inherent to the quality of the content or the implementation of recent technologies. According to De Vicente and Pain (2002), the main focus of previous research has been on the motivation in relation to instructional design, while limited research has focused on motivation detection which means that more research is needed in the latter (De Vicente & Pain, 2002).

Measuring motivation is a prerequisite for being able to detect it. However, because motivation is a psychological concept, it cannot be directly observed nor recorded (Touré-Tillery & Fishbach, 2014). Researchers tried to measure motivation by examining different conceptualizations of motivation. However, a significant aspect before attempting to measure motivation is to determine and understand what type of motivation the researcher is willing to measure, in what context, and what are the other concepts involved besides the concept of motivation (Touré-Tillery & Fishbach, 2014). Moreover, applying a suitable motivation theory as a base for motivation detection investigations is required because there is a large amount of psychological literature on theories about emotion and motivation but only few studies have applied them (De Vicente & Pain, 2002).

The theoretical perspectives applied in this article are the reinforcement sensitivity theory (RST) and Arnold's appraisal theory of emotion. RST is based on the distinction between approach and avoidance motivation, which are the two types of motivation. In approach motivation, the behaviour is triggered by a positive and desirable event like reward, incentive and appreciation to reach desirable outcomes. On the other hand, in avoidance motivation, the behaviour is triggered by a negative undesirable stimulus like aversion, punishment and threat (Elliot, 2008; Elliot & Covington, 2001). The distinction between approach and avoidance motivation in studying and analysing motivation and behaviour is one of the oldest ideas in psychology concerning the behaviour of organisms according to Elliot (Elliot, 2008; Elliot & Covington, 2001). What is new with this distinction is the ability to explain and predict motivated behaviour (Elliot, Gable, & Mapes, 2006). Moreover, the widespread use of this distinction reveals the essential and basic roles of the approach and avoidance motivations in the human functioning (Elliot, 2008).

The variation between the approach and avoidance motivations lies in valence (pleasantness and unpleasantness of stimuli or environment). This means that motivation and emotion goes hand in hand, and therefore motivation cannot be studied separately from emotion. Emotions are considered as a motive according to Arnold's appraisal theory of emotion, which explains that emotions generate motivational states of the person. Liking generates the approach motivation and disliking generates the avoidance motivation (Arnold, 1960; Buck, 1988; Reeve, 2014).

According to De Vicente and Pain (2002), there have been numerous attempts for approaching the challenge of measuring and detecting a person's motivational states. The difference between these attempts is on the information source that has been used. The methodological traditions and formats of motivation measurements used are questionnaire, self-report methodology, behavioural approach and neurophysiological measures approach (De Vicente & Pain, 2002; Fulmer & Frijters, 2009). Self-report via questionnaires and scores is a commonly used approach when it comes to measuring motivation. However, there are many issues about these approaches. First of all, their accuracy is questionable because measuring attitudes with a tool like questionnaire or score raises many sceptical thoughts according to De Vicente and Pain (2002). They argue that this is because these approaches use a person's perceptions to measure motivation states, which results into the problem of subjectivity (De Vicente & Pain, 2002; Touré-Tillery & Fishbach, 2014). Furthermore, questionnaires are static, which means that they can only inform about the permanent characteristics of the persons but motivation states are changing (De Vicente & Pain, 2002).

Another approach is to use neurophysiological measures such as recording brain activations using electroencephalogram (EEG) (Touré-Tillery & Fishbach, 2014). However, there have been a limited number of attempts that have studied motivation using this approach (De Vicente & Pain, 2002). Further, these attempts have analysed the data from only a single feature or property derived from the EEG signals or studied other concepts that can be considered as motivational factors, such as attention and engagement. One such study used power spectral analysis of EEG data and reported that EEG waves correlated with the increase of motivation during a serious game play. Another study measured a person's attention level as a motivational factor using neurophysiological and physiological measures. Data mining, which is a sub-field of computer science that is used to discover novel and potentially valuable information from large quantities of data (Baker et al., 2010), has also been used (Derbali, Chalfoun, & Frasson, 2011) and the best result of the three data mining classifiers used in this study reached an accuracy of 73.8% in detecting person's attention level using EEG, heart rate & skin conductance. However, using an objective data source, analysing the data source in a multidimensional manner and applying suitable motivation theories as a base to capture the complexity of motivation are required. This is the gap that the current article addresses.

In this era of artificial intelligence, various deep learning techniques have evolved in order to address several real life problems. such as Esen, Inalli, Sengur, and Esen (2008a) have reported a comparative study of adaptive neuro-fuzzy inference systems (ANFIS) and artificial neural network (ANN) applied to model a ground-coupled heat pump (GCHP). In the paper the designed models are used to predict the system performance related to the air and the ground temperature. Similar to above, several such automatic approaches to learn to predict various aspects of GCHP have been developed so far (Esen, Inalli, Sengur, & Esen, 2008b, 2008c). In these aforementioned studies also various types of advanced approaches like artificial neural networks, statistical weighted pre-processing (SWP) and so on are used to achieve better performances. Similarly, a study to model a solar air heater (SAH) system (Esen, Ozgen, Esen, & Sengur, 2009) to establish a solar-assisted ground source heat pump system-based slinky type ground heat exchanger (GHE) (Esen, Esen, & Ozsolak, 2017) using various ANN, ANFIS, wavelet neural network (WNN) models have been evolved in recent times.

Similarly, Schirrmeister et al. (2017) have made thorough experiments with various convolutional neural architectures which are particularly designed to decode and execute various tasks from raw EEG signals. Dai, Zheng, Na, Wang, and Zhang (2019) have combined classical CNN with variational autoencoder (VAE) for classifying various motor-imaginary tasks via analysing raw EEG signals. The Gaussian distribution of EEG representation has been fit to the signal by the decoder part of the VAE, where the extracted CNN features are fed to the deep VAE. Özdenizci, Wang, Koike-Akino, and Erdoğmuş (2019) address the problem of specific variability during the recording session of the EEG in determining a person identity via analysing brain activity. The authors introduce an adversarial inference approach, which extends various deep learning models for session-invariant person-discriminative representations learning, from the view of an invariant representation-learning. Another approach of determining motor-imaginary tasks by analysing EEG signals have been proposed by Amin, Alsulaiman, Muhammad, Mekhtiche, and Hossain (2019). The proposed technique extracts multi-layer CNN features for having good deal of spacial and temporal features, and fused those features by using their proposed MCNN and CCNN techniques. The paper reports very decent accuracies on two popular and publicly available datasets.

Apart from that, detection of error correlates in EEG-based BCI analysis, has become an important aspect of research in recent times. Such as, Rouanne, Śliwowski, Costecalde, Benabid, and Aksenova (2021) have performed a thorough study over the detection of error correlates in the primary motor cortex at the single trial level. The paper shows that the error correlates can be detected by using popularly used classifiers such as support vector machine (SVM), multi-layer perceptrons (MLP), N-way partial least squares (NPLS) and CNN. Furthermore, the authors report very good area under curve (AUC) and ROC curve for all four aforementioned broad range of classifiers in the motor-imaginary control condition.

The areas within computer and systems science to which this article contributes to are data mining and affective computing (Calvo, 2010). The problem that this article addresses is how to accurately detect motivational states by (a) applying suitable motivation theories and related concepts as a base to motivation detection and (b) performing data mining of multidimensional EEG signal data to provide an objective alternative to the currently prevalent self-reported subjective measurements. This study aims to establish a model that detects a person's motivational states based on an optimal combination of EEG data and data mining methods. To accomplish this, we have designed a game for experimentation purpose and also proposed a novel deep learning framework for detection of motivation. It is well known that EEG dataset that can provide very large number of sample data points for training purpose is scarce, therefore it becomes difficult for deeper model to learn relevant features. Here in this study we address this problem too through our proposed Residual-in-Residual Convolution Neural Network (RRCNN) model that fits very small dataset properly in spite of having very deep architecture. The basic building block of RRCNN network is the Residual Stack (RS), which actually operates the residual-in-residual operation. Despite being a very deep model, the inclusion of small and large skip-connections for accomplishing residual-in-residual operations ensures that the proposed RRCNN does not overfit on a small dataset. In addition, we have used Bayesian optimization for tuning of various hyperparameters of RRCNN network.

The objective of the paper is to propose data mining and affective computing framework for studying motivation through the use of EEG signal and brain–computer interface. Our methodology consists of (a) designing a game-based interface which allows for stimulating motivation while recording the EEG signal of the player, (b) selecting suitable subset of channels and features from the entire EEG channel dataset, (c) learning the deep residual features that encode the state of motivation by training RRCNN on the selected and pre-processed EEG data and using the trained RRCNN to detect motivation, and (d) perform statistical analysis associated with the motivation of the participants. We undertake a variety of studies regarding the different aspects of our methodology to validate the effectiveness of our approach. Our results show improvement over the state-of-the-art and establish the utility of our approach for studying motivation through EEG signals.

This work is original in the sense that EEG signals have not been used for detecting motivation. Our unique game-based data mining approach is designed to extract multi-dimensional EEG signals in response to motivation stimuli such as reward and punishment. This has enabled us to perform data mining, hypothesis validation and statistical analysis associated with motivation of this nature for the first time. An important novelty of our work is the RRCNN, which is explicitly designed to possess both deep residual feature support and the ability to learn them using small datasets.

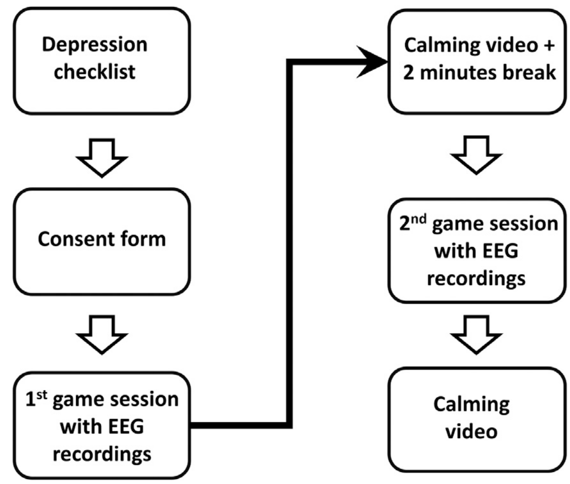

Fig. 1. The overall workflow of the game is given here. For different experiment groups, different treatments are given to the players in 1st and 2nd game sessions. The detailed information is given in 2.2.3.

2. Experiment

In this section we discuss about our data mining approach which we have implemented in order to gather EEG based data, associated to motivational states of various persons. Various subsections of this section discuss about the participants included into the experiment, the elements which were used in the experiment, the philosophy and the base of the experiment, the stimuli and the experimentation procedure (see Fig. 1).

2.1. Participants

Prior approval from institutional review board of Nanyang Technological University, Singapore has been taken before conducting this study(detail in Zary (2018)). 30 healthy participants with no history of drug use and smoking, or any neurological and psychiatric disease, were chosen for the experiment. The age of each participant was between 21 to 45 years. All of the participants were from school of computer science and engineering, school of humanities and school of social sciences. A recruitment letter was sent to them and the willing participants who fulfil the aforementioned criteria, were selected for the experiment. The recruitment was done with the help of Nanyang Technological University, school of computer science and engineering, school of humanities and school of social sciences. Thereafter, the potential students were set to answer a depression checklist, so that, their depression level and mood scale can be measured, as only the participants with low depression level were to be included in the study.

2.2. Apparatus and materials

The electroencephalography machine, which was used in data collection, is known as BrainMarker and the model of the machine is DEV12001EEG. The machine has 21 electrodes in total, 19 of which are recording electrodes, 1 ground electrode and 1 is for system reference point. The placement of the ground and the reference points are at AFz, and CPz. Where AFz is between Nasion and Fz, and CPz is between Cz and Pz. Following the universal standard of 10–20 electrodes placement system, 21 electrodes are placed on the scalp of each participant. The 10–20 electrodes placement system is used. The mentioned EEG electrodes detect the change in electrical signals with various brain activities.

Table 1
The description of the different versions of the game.

	Version A	Version B	Version C
Initial value of score field	NA	0	12
Hardcoded best score	NA	12	4
Score obtained for one match	NA	+1	0
Score obtained for one mismatch	NA	0	−1
Sound played	NA	Positive sound	Negative sound
Feedback message upon match	NA	'Pair match, well done'	None
Feedback message upon mismatch	NA	None	'Pairs do not match, it's a pity'

2.2.1. The game

A game was designed as the brain computer interface for the experimentation purpose and to evaluate the motivation state of participants. Gamification's association with motivation as investigated in Richter, Raban, and Rafaeli (2015) was used as for designing the game. The game was designed using the Unity Game Engine, which is a tool-set for creating computer games. The game has three versions, say A, B, and C. The first version is a non-treatment game sessions, which is developed for the control group (described below). This treatment contains 3 card memory game sessions. The positioning of the cards in each session were fixed and predefined such that the sequence of cards shown to each player during the game remains same. The players have to match consecutively flipped cards. For a particular level of game, if all the pairs are matched successfully, then "Proceed to Next Level" button appears on the screen. After the participant completes the 5th game session, the button "End of game" is used.

The versions B and C are created to embed audio-visual positive and negative feedback into the game. The version A is the simple non-treatment game which has been discussed earlier in this section. Version B is the same as version A but has a positive feedback after each card flip, which triggers the approach motivation into the player. Version C is similar to version B, but the only difference is that the given audio-visual feedback is negative. The negative feedback is given to trigger the avoidance motivation among the participant players. Through these versions, the reinforcement sensitivity theory and the constructs of approach and avoidance motivation are embedded in the experiment design. The specifications of different versions of the game are given by Table 1.

2.2.2. Stimulus

In this experiment, motivational factors have been used as stimuli. These motivational factors trigger the approach and avoidance motivations, which is discussed in the reinforcement sensitivity theory. Two types of cues are mentioned in the theory, namely the promotion cue and the prevention cue. Promotion cue contains success and gain, in order to trigger the approach motivation. Prevention cue consists of failure and loss, which triggers the avoidance motivation. The stimuli for these cues are described below:

- **Promotion cue condition:**
  In promotion cue condition, the game sessions are created such that, when a pair match occurs then:
  - The memory game sessions start with 0 points and the playing participants have to achieve as much points as possible. (Gain)
  - Positive sound effect is given for a match. (Success)
  - Positive text is shown for a match : "Pair match, well done" (Success)
- **Prevention cue condition:**
  In the prevention cue condition, the game sessions with mismatched pairs are created such that:
  - The game session starts with 13 points and the participants have to avoid losing points. (Loss)
  - Negative sound is played for each loss condition. (Failure)
  - For each mismatch pair, negative message is shown: "Pair do not match, it's a pity" (Failure)

2.2.3. The game procedure and the brain computer interface (BCI)

As given in Fig. 1 this experiment is performed through an open loop BCI system. 30 selected participants are divided into 3 groups. Each group consists of 10 participants. The groups are named as control group(Con), first experiment group (ER/P) and second experiment group (EP/R). Each participant has to play the game for two sessions. In between two sessions of the game, a calming music video is played to retrieve the mood of the player to neutral state, which is followed by a break of 2 min.

The participants who are assigned to the control group (Con) play the version A. The first experiment group (ER/P) play version B for the first session and version C for the next session of experiment. On the other hand, participants of second experiment group (EP/R) play the game of version C for the first session of experiment and version B for the second session. So basically for control group, no feedback is given for either matched pair or unmatched pair. For the first experiment group, positive reward is given for session 1, to trigger approach motivation and negative reward is given for the later session, which triggers avoidance motivation. Whereas for second experiment group (EP/R), the feedback of each session is reversed of that of first experiment group. Detailed information about the game and game sequence for different groups is given by Fig. 2.

3. Proposed hypothesis and validation approach

Hypothesis. If a person in motivated then he learns more and starts getting success consecutively in the game, that is the approach motivation is triggered. In contrast, consecutive failures or no fixed sequence of success is evident of triggering the avoidance motivation in the participant.

3.1. Success in the game paradigm for motivation

For the validation of the hypothesis statistically, we have considered CUMSUM or cumulative sum (Ploberger & Krämer, 1992) plots. In the experiment of card flipping, each flip which produces a matched pair, is labelled as one and, the flips which results in an unmatched pair is labelled as zero. From the CUMSUM plots, we observe the sequences of 0's and 1's with respect to the time the player is playing. There can be three sequences that can be observed from the CUMSUM plots. These are

- **Sequence of 1's :** In this case, ones start coming consecutively after some period of the game. That is, the success starts coming in sequential manner, which means the participant have learnt the game. In this case, we conclude that the approach motivation is triggered in the participant.
- **Sequence of 0's :** If consecutive zeros are present in the sequence, it implies that the participant is not learning anything. This therefore indicates that the patient is under the triggering of the avoidance motivation.
- **Sequence of random 0's and 1's :** If no fixed pattern of sequence is found, that is zeros and ones are coming at random, then also we can say that the participant is not learning, the approach motivation is not triggered prominently and the avoidance motivation dominates over the approach motivation.

We have obtained conclusive results for visualizing the dataset, from these cumulative plots, which are discussed in the **Results** section (Section 5).

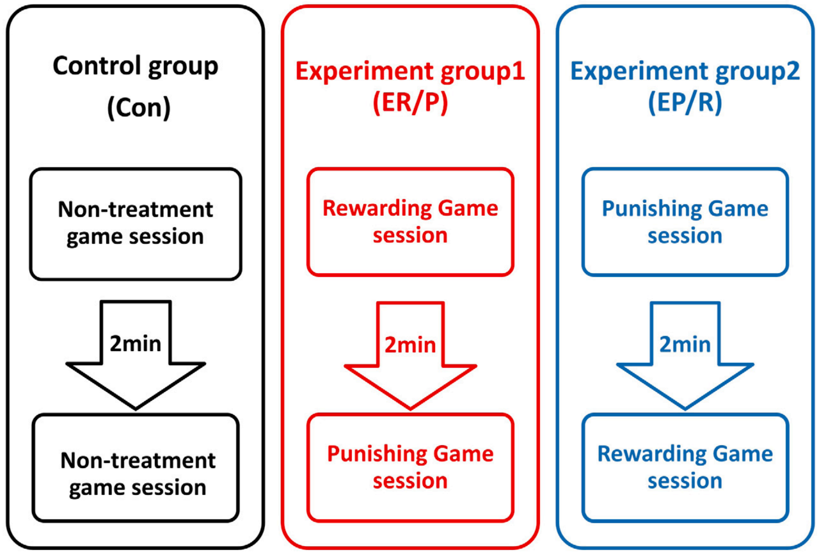

Fig. 2. The game procedure.

3.2. EEG channel sets of interest

In this current study, we have considered three different EEG channel sets (see Table 2) to validate proposed hypothesis of motivation. The channel sets selected for evaluation purpose of proposed model of motivation detection are frontal asymmetry index, emotions and 'both'. The selection of EEG channels with these aforementioned measures are discussed below. The placement of all 21 channels in 10–20 system of electrode placement is pictorially shown in Fig. 3.

3.2.1. Frontal asymmetry index

Frontal asymmetry is the measurement of electroencephalography (EEG) activity in the left frontal region versus the activity in the right frontal region. Analysis of frontal asymmetry index has always been an important domain to study the approach and avoidance behaviour of human beings (Davidson, Ekman, Saron, Senulis, & Friesen, 1990; Ng, Fishman, & Bellugi, 2015). If left sided prefrontal activation is more, then that indicates higher level in approach motivation. In contrast, higher level in avoidance motivation results in activation of right sided prefrontal lobe (Gollan et al., 2014; Horan, Wynn, Mathis, Miller, & Green, 2014; Mauss & Robinson, 2009; Ng et al., 2015; Sutton & Davidson, 1997). For analysis of frontal asymmetry index we have considered the channels F3, F7, F4 and F8.

3.2.2. Emotion

The connection between valence and frontal asymmetry of an individual's emotional state is reported in some research studies (Mauss & Robinson, 2009). For example we can consider the results obtained by Davidson et al. (1990). It is stated in Davidson et al. (1990) that "negative emotions are associated with the right-sided activation in the frontal and anterior temporal region of the brain, whereas positive emotions are associated with the left-sided activation in the frontal and anterior temporal region of the brain". Furthermore, research study by Jatupaiboon, Pan-ngum, and Israsena (2013) indicates that the channels reporting best results in classifying positive and negative emotions are T7 and T8. These channels are located on the temporal lobe and are analysed in our study for motivational state detection.

3.2.3. Both

Motivation and emotion go hand in hand and some researchers enhance the significance of analysing motivation and emotion in a combined manner (Arnold, 1960; Buck, 1988; Reeve, 2014). Therefore in this study, we have considered another measure called 'Both, which is the combination of both frontal asymmetry index and emotion measures. This is to check whether by combining both channels sets improves the accuracy of classification between motivated and non-motivated mental states. F3, F7, F4, F8, T7 and T8 channels are taken into account in order to analyse this measure.

Table 2
Measures, and EEG channels selected in the data analysis.

Measures	Brain activation area	EEG channels
Frontal asymmetry index	Greater left-sided prefrontal activation vs. Greater right-sided prefrontal activation	Approach motivation : F3, F7 Avoidance motivation: F4, F8
Emotions	The left frontal region left side of the anterior temporal region activation vs. Right frontal region, the right side of the anterior temporal region activation	Positive emotions: T7 Negative emotions: T8
Both	All brain activation areas stated above	Both approach motivation and positive emotions: F3, F7, T7 Both approach avoidance motivation and negative emotions : F4, F8, T8

3.3. Selection of EEG spectral bands

EEG records the electrical activity of our brain with the aid of different electrodes placed at appropriate positions of human scalp. The electrical brain waves recorded by the EEG machine can be categorized into four different types: beta (>13 Hz), alpha (8–13 Hz), theta (4–8 Hz), and delta (0.5–4 Hz) (Blinowska & Durka, 2006; Tatum IV, 2014). In this current study, after de-noising the EEG signal using MATLAB software, Alpha and Beta waves are extracted from the spectral contents of the raw signal by using Butterworth band-pass filter. Alpha waves are associated with the awake and relaxed state of mind, which implies that the person is active with a passive reacting state. Alpha wave also reflects particularly, cognitive and memory performance of a person. On the other hand, beta waves occur in mental effort or activation, that is doing something with extreme level of concentrations. In this research framework of detecting motivation level, we have particularly chosen alpha waves and beta waves to discriminate between different brain activations caused by passive reflective state of mind versus concentration, attention and thinking.

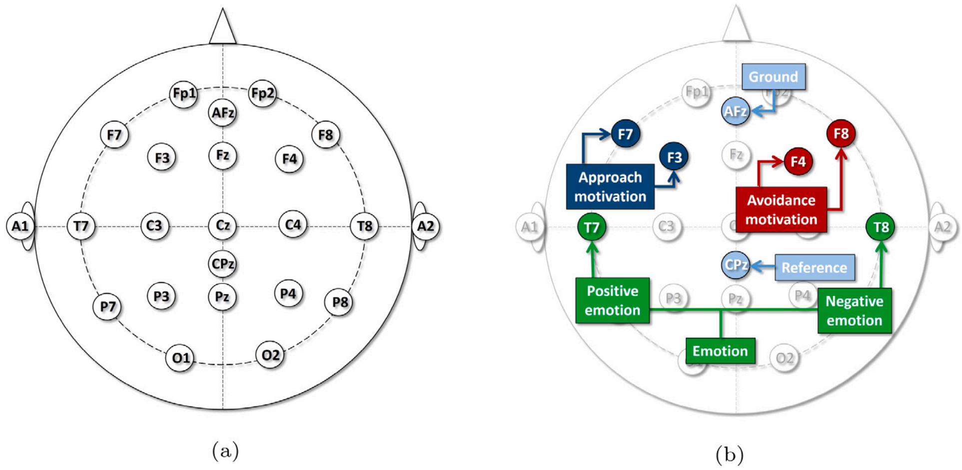

Fig. 3. In (a), 21 channels of 10–20 system of electrode placement is given. The channels which are highlighted in (b), we have evaluated in this study to detect motivation of participants. These highlighted electrodes include the ground and reference points also.

3.4. EEG and stimulus slicing

During the experiments, EEG signals are recorded continuously without introducing breaks for the transitions between sessions. Therefore, it becomes necessary to mark the exact time of usage of the stimuli in the experiments. The time stamps for each event of matched or unmatched cards are tracked and sorted in the game log files. This is a significant step to cut down 300 ms of EEG activity before the stimulus, which is the thought process and 300 ms of EEG activity after the stimulus, known as the reaction process. This is done to compare different brain activations while thinking before the stimulus and reacting after the stimulus.

4. The proposed Residual-in-Residual CNN (RRCNN)

4.1. Network architecture

Our deep neural classifier is comprised of following stages : shallow feature extraction, deep feature extraction, adaptive average pooling and flattening, and classification. A single 1-D convolution layer has been used for the shallow feature extraction from the raw EEG signal. Now suppose the input data is X_i and shallow intermediate feature is F_s. Then, these are related by Eq. (1)

$$F_s = C_{SF}(X_i) \tag{1}$$

where $C_{SF}(.)$ is the 1-D convolution operation, using which shallow features (SF) from the raw input signal is extracted.

Now, the shallow feature map is further fed to our residual portion of the network for deep feature extraction. The residual portion of the network is consisted of some stacks of residual blocks and all the residual blocks within each stack always have same feature dimensions. These stacks also have residual skip-connections between the input and output of the stack. Our residual-in-residual architecture of the residual block comprises of long and short skip-connections (see Fig. 4), which is able to fit very small dataset without problems like over-fitting, gradient vanishing and so on.

Suppose there are a such stacks, such that the residual portion of the architecture(A) that extracts deeper low dimensional features, can be expressed as $A = \{A_1, A_2, \dots ., A_a\}$. Suppose the input feature map of a stack A_1 is say I_{A1} and output is say O_{A1}, It is mentioned that the skip-connection is between I_{A1} and O_{A1}. Since a skip-connection adds the prior input feature map to the posterior output feature map, therefore the final output of the residual stack A_1 would be the addition of I_{A1} and O_{A1}. If the final output of stack A_1 is O_{f1} and if the operation of stacked residual blocks for stack A_1 is C_{s1} then

$$O_{A1} = C_{S1}(I_{A1}) \tag{2}$$

$$O_{f1} = I_{A1} + O_{A1} \tag{3}$$

In addition, let us suppose that the stack A_1 has b number of residual blocks such that,

$$A_1 = B = \{B_1, B_2, \dots .., B_b\} \tag{4}$$

where $B_1, B_2, \dots .$ are 1-D residual blocks. For a single stack say A_1, the input feature dimensions and output feature dimensions of all residual blocks (B_i) are same and can be different for different stacks but then output of each stack passes through another convolution operation before being fed to the next stack. Suppose the output feature map of A_1 stack is F_{A1}, which has suppose the feature dimension of $L_{A1} \times Ch_{A1}$. Now the convolution operations can be represented as

$$C = \{C_1, C_2, \dots \dots, C_{a-1}\} \quad | \quad C_i \in \mathcal{R}^{Ch_i \times Ch_{i+1}} \tag{5}$$

$$Ch_{i+1} \geq Ch_i \tag{6}$$

where C_1 operation is performed between residual stacks A_1 and A_2. This is how using skip-connected stacks of residual blocks, deep features are extracted from shallow feature maps.

$$F_d = C_{DF}(F_s) \tag{7}$$

where $C_{DF}(.)$ extract deep features with the aid of stacked residual layers.

Since the input time-series data of the thought process for each card flipping, has been collected in a regular fashion, therefore the time length for each data is unique and different from other. This results in a dimension mismatch error in our network. To address this problem, we have embedded an adaptive average pooling layer, which normalizes dimension deep feature (F_d) to one and outputs in features having sequence length equals to 1. Now, if the adaptive average pooling operation is $Avg(.)$ which takes the input deep feature map F_d and produces output F_{avg} having sequence length normalized to one. Due to this adaptive average pooling operation following mapping takes place.

$$F_d^{L_d \times Ch_d} \rightarrow F_{avg}^{1 \times Ch_d} \tag{8}$$

Thereafter, the feature map F_{avg} is flattened and Ch_d number of deep features are obtained from the whole network, which are further fed to the classifier for classification purpose. The residual architecture of the proposed RRCNN is pictorially described in details in Fig. 4.

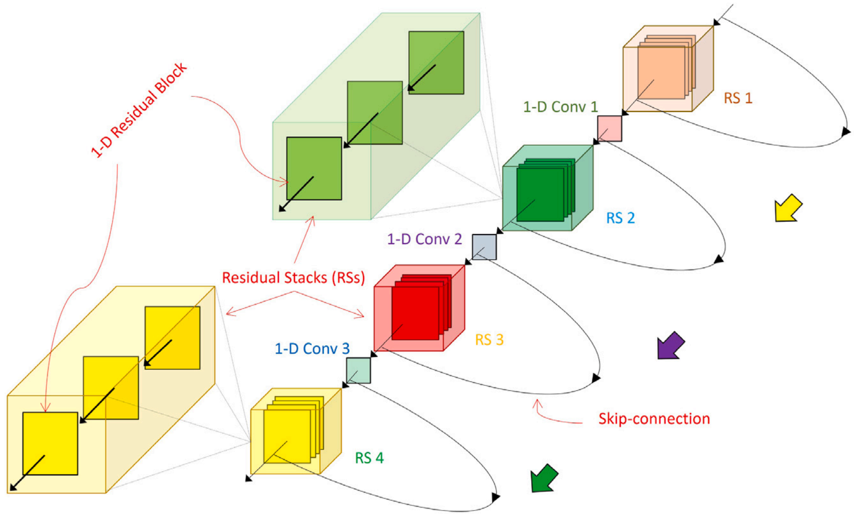

Fig. 4. The detailed architecture of RRCNN network.

4.2. 1-D residual block

The basic building block of proposed deep neural classifier is the 1-D residual block, which is shown in Fig. 5. The residual blocks are consisted of two 1-D convolution layers (C_r) and the number of input channels and output channels of each convolution layer are same. The input feature map is added to the output feature map of the second convolution layer via skip connection. Batch normalization and ReLU(.) activation function are embedded after the first convolution pooling.

Suppose the input feature map of a residual block is x_{in}. Now after two convolution operation with 1 dimensional batch-normalization and ReLU activation, suppose the output feature map is x_{out}. Say this combined operation is denoted by $C_{res}(.)$, therefore

$$x_{out} = C_{res}(x_{in}) \tag{9}$$

Now the final output x_f is the addition of x_{out} and x_{in}.

$$x_f = x_{out} + x_{in} \tag{10}$$

The pseudo-code of the residual operation is given by Algorithm 1.

Algorithm 1 Pseudocode for the 1-D residual operation. It has been mentioned earlier that the kernel size, stride, and padding of the convolution layer (C_r) are chosen to be 3, 1, and 1. Here in the algorithm '+' sign denotes the pixel-wise addition or aggregation of two feature maps having equal dimensions.

Input: $x_{in} \rightarrow$ Input feature map
$Ch_{in} \rightarrow$ Input channels
Initiate C_r = 1-D Convolution (Ch_{in}, Ch_{in}, Kernel size , Stride , Padding)
Initiate B_n = 1-D Batch normalization (Ch_{in})
Initiate Activation = ReLU(.)
for i = 1 : 2 **do**
    $x = C_r(x_{in})$
    if i ← 1 **then**
        $x = B_n(x)$
        x = Activation(x)
    end if
end for
$x_{out} = x$
$x_f = x_{in} + x_{out}$
Output: x_f

4.3. BayesIan optimization for hyperparameters tuning

In training of a deep learning model, the loss function gets optimized via backpropagation (Hecht-Nielsen, 1992) technique, which uses an optimizer like stochastic gradient descent (SGD), Adam or RMSProp and so on. The trainable parameters gets upgraded during training by the following equation.

$$\theta_{n+1} = \theta_n - \alpha \bigtriangledown L(\theta_n) \tag{11}$$

where n and α are the number of iteration and the learning rate of the optimizer and $L(\theta)$ is the loss function, which is to be optimized. The algorithm of the optimizer can oscillate when it approaches the mostly erected path of optimum. Therefore to stabilize the motion and avoid the oscillation another parameter called momentum is added to the optimizer function. So, Eq. (1) is changed to

$$\theta_{n+1} = \theta_n - \alpha \bigtriangledown L(\theta_n) + \beta(\theta_n - \theta_{n-1}) \tag{12}$$

where β is the momentum factor, which determines the contribution of the step of previous gradient to the current iteration. In addition, data augmentation, dropouts, adding regularization to the loss function term can also overcome the overfitting of DNN models. The regularization embedded loss function becomes

$$L_R(\theta) = L(\Theta) + \gamma \Phi(\omega) \tag{13}$$

where $\Phi(\omega)$ is given by

$$\Phi(\omega) = \frac{1}{2}\omega^T \omega \tag{14}$$

In above equations, γ is the regularization coefficient and ω is the weight vector. Since in our CNN model the weight layers, hidden dimensions of the residual blocks batchnorm, dropouts and such architectural hyperparameters are determined manually by performing exhaustive experimentation, therefore we are left with very few trainable hyperparameters which are to be optimized by Bayesian optimization (Snoek et al., 2015) technique. These hyperparameters are not related to the architecture of the network but to the training and evaluation process of the network. The parameters, which we have optimized with Bayesian optimization are the optimizer, learning rate, batch size and the regularization parameter. We have done 5 fold cross validation for evaluating our model, that is, we have considered 80% of the data for training, 10% for validation and 10% for testing purpose.

Now, the objective function which is to be minimized is the validation error (e_V), which is clearly not a convex function or has any closed form of expression. Therefore, gradient based approaches would not work in this case. In addition to this, since the evaluation

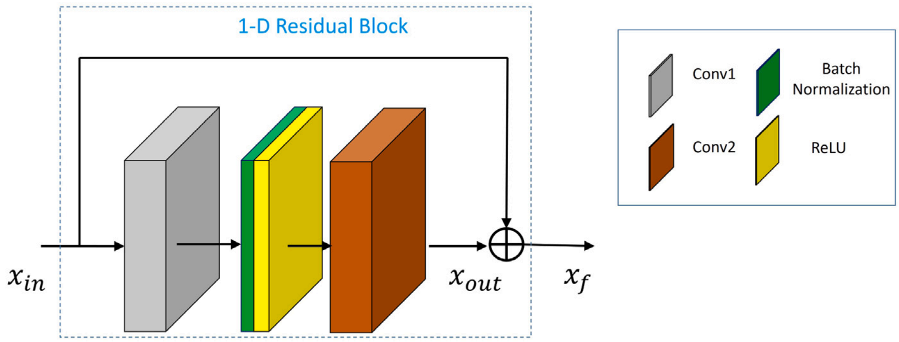

Fig. 5. The residual block of RRCNN network.

of the function is computationally much costly, therefore brute force optimization techniques like grid search or genetic algorithm will not be applicable here.

Therefore in this case, Bayesian optimization suits perfectly to find optima of the expensive objective function e_V. The algorithm works by constructing a probabilistic model, considering the objective function, which is in our case the validation error. This technique is advantageous since it takes the benefits of considering all available information from prior evaluations instead of relying on only Hessian approximations and local gradients. With the aid of Bayesian optimization technique, extrema of difficult non-convex functions can be found with relatively less iterations compare to other optimization techniques.

In Bayesian optimization, we make a fundamental assumption that the objective function e_V is defined from prior Gaussian process, that is, $e_V(\eta) \sim N(0, \chi)$, where η is a vector, which consists of the hyperparameters. Let us suppose the objective function is corrupted by Gaussian noise, which has zero mean value and standard deviation of σ_n. Therefore χ is defined as given in Eq. (15).

$$\chi = \Psi + \sigma_n^2 f \tag{15}$$

where Ψ is a matrix of dimension $f \times f$ such that the first element (element at position 1,1) is $\lambda(\eta_1, \eta_1)$ and the element at position (f, f) is $\lambda(\eta_f, \eta_f)$.

Where $\lambda(\eta, \eta')$ is known as the covariance function and I is the identity matrix. Now, let us denote the observations from previous iterations are

$$O_{1:f} = \{\eta_{1:f}, e^V_{1:f}\} \tag{16}$$

where $e^V_{1:f} = e_V(\eta_{1:f})$. Now suppose, η_{f+1} is the next point which is to be evaluate, therefore the objective function at η_{f+1} is $e^V_{f+1} = e^V(\eta_{f+1})$. Under GP prior, $e^V_{1:f}$ and e^V_{1+f} can jointly be considered as Gaussian. Therefore the we can obtain the expression for predictive distribution as

$$e^V_{f+1}|O_{1:f} \sim N(\mu(\eta_{f+1}), \sigma^2(\eta_{f+1}) + \sigma_n^2) \tag{17}$$

where

$$\mu(\eta_{f+1}) = \Lambda^T(\chi + \sigma_n^2 I)^{-1} e^V_{1:f} \tag{18}$$

$$\sigma^2(\eta_{f+1}) = \lambda(\eta_{f+1}, \eta_{f+1}) - \Lambda^T(\chi + \sigma_n^2 I)^{-1}\Lambda \tag{19}$$

$$\Lambda = [\lambda(\eta_{f+1}, \eta_1)\lambda(\eta_{f+1}, \eta_2) \ldots\ldots \lambda(\eta_{f+1}, \eta_f)]^T \tag{20}$$

Now, the predictive posterior distribution $e^V_{f+1}|O_{1:f}$ can sufficiently be characterized by the predictive variance function $(\sigma^2(\eta_{f+1}))$ and the predictive mean function $(\mu(\eta_{f+1}))$ of the distribution. These characterizing factors solely depend on the selection of the covariance function $(\lambda(\eta, \eta'))$. In our study we have used automatic relevance determination (ARD) Matern 5/2 kernel, such that the problem caused due to unrealistic smooth covariance function of ARD squared exponential kernel can be addressed.

It has been mentioned earlier that Bayesian optimization technique uses utility or acquisition function, which is derived from aforementioned predictive posterior distribution, for distribution of more computations to determine the next point for evaluation. In this current study, the utility function is the expected improvement over the best expected value, which is given by

$$\mu_{best} = \underset{\eta_l \in \eta_{1:f}}{argmin}\ \mu(\eta_l) \tag{21}$$

Under Gaussian process prior assumption, this utility function has a close form solution which is expressed as follows

$$x_{ei}(\eta_{f+1}) = \sigma(\eta_{f+1})[Y\Theta(Y) + \theta(Y)] \tag{22}$$

here $\Theta(.)$ is the cumulative distribution function, $\theta(.)$ is the PDF of the standard normal and Y is defined as

$$Y = \frac{\mu_{best} - \mu(\eta_{f+1})}{\sigma(\eta f + 1)} \tag{23}$$

The Bayesian optimization algorithm uses proxy optimization to determine which point should be evaluated next. Proxy optimization is to maximize the acquisition function, i.e., the minimization of the expected improvement (ei) value over the recent best expected value. The complete algorithm of Bayesian optimization for deep learning hyperparameter tuning is given by Algorithm 2.

Algorithm 2 Pseudocode for the Bayesian optimization for deep learning hyperparameter tuning.

Input: The range of values of the hyperparameters: lr(α), optimizer(Op), batch size(bs), regularization parameter(β) and momentum(γ).

for f = 1, … **do**
  With chosen ARD Matern 5/2 kernel, calculate $\mu(\eta_{f+1})$ and $\sigma^2(\eta_{f+1})$ (given by **Eqn.** (18) and **Eqn.** (19)).
  By optimization of the utility function over the Gaussian process prior : $\eta_{f+1} = arg \quad max_\eta x_{ei}(\eta|O_{1:f})$, find $\eta_{f+1} = (\alpha_{f+1}, Op_{f+1}, bs_{f+1}, \gamma_{f+1}, \beta_{f+1})$
  Calculate the validation error (e_V) with the aforementioned deep learning parameters.
  Augment the data $O_{1:f+1} = \{O_{1:f}, (\eta_{f+1}, e^V_{f+1})\}$
end for

Output: The best set of hyperparameters, which are given as input.

5. Results

5.1. Data analysis and visualization

It is mentioned earlier in Section 3 that for data analysis and visualization purpose, we use cumulative sum (CUMSUM) plots of different sequences of different experiment groups. Using the information obtained from the CUSMSUM plots, we have labelled the data for

Table 3
The experiment results obtained from CUMSUM plots of different groups.

Group	Number of participant	Motivated	Non-Motivated
Con	10	2	8
ER/P	10	4	6
EP/R	10	7	3

the training of subsequent deep neural classifier. The participants of different experiment groups have played the game for two sequences. The CUMSUM plot of a single participant indicates whether that person is motivated or not. Some common patterns can be found between the CUMSUM plots of participants who fall under same group and sequence. We present in Fig. 6 the representative plots of CUMSUM for both sessions of different experiment groups. Since the time of the game-play of different players are different from each other, therefore while plotting the CUMSUM, we have selected first 50 discrete time-data for analysis purpose. The common pattern in each experiment group implies that, for different persons, if similar triggering feedback is given then the motivation level gets triggered in a similar way. Explicit observations for each group are presented next.

In Fig. 6(a), the CUMSUM of 0's are rising over time at a larger rate than the CUMSUM of 1's. This means that the participant is making consecutive mistakes in finding matched pair of cards. This further implies that the person is not learning at all. Referring to our hypothesis, we may conclude that the person is not motivated. Similarly, Fig. 6(b) also indicates that the player is not much motivated to learn the game, as 0's and 1's are coming at random. On the other hand, in Fig. 6(c), we see that after some time of game-play, the participant starts getting 1's in a consecutive manner, that is, the player is motivated enough to learn the game. These three specific patterns of CUMSUM are found in all the cases of participants. It is observed that the CUMSUM plots of first sequence of experimentation for different experiment groups are quite similar and match Fig. 6(a) or (b). That is, the participants playing the first sequence of the game are not learning the game at all. On the other hand, variations are found while playing the second sequence. For control group (C) and first experiment group (ER/P), the CUMSUM plots given by Fig. 6(a) and (b) are mostly found in case of second sequence of the game play. But the number of occurrence of CUMSUM plots given by Fig. 6(c) is slightly greater for experiment group ER/P than that of C. On the other hand, the CUMSUM plot Fig. 6(c) is mostly found for sequence 2 of second experiment group EP/R, for six out of ten participants. Therefore, it is indicated from the analysis that, the participants belonging to the group ER/P learn the game better than that of the participants belonging to other groups. From the data analysis and CUMSUM plots, we are able to identify which participants are motivated and which are not. The results of the analysis, that is the number of motivated and non-motivated participants belonging to different groups are tabulated in Table 3.

5.2. Detection of motivation

5.2.1. Optimal hyperparameters determination using Bayesian optimization

We have used Bayesian optimization of optimizing the learning rate of the RRCNN (α), the momentum factor (γ), the regularization coefficient (β) and the batch size (bs) of training samples. The ranges of α, bs, γ and β are given below

$$\alpha \rightarrow [1\times10^{-6}, 1\times10^{-2}], \quad bs \rightarrow [1, 15], \quad \gamma \rightarrow [0.01, 0.99], \\ \beta \rightarrow [1\times10^{-6}, 1\times10^{-3}] \tag{24}$$

In addition the optimizers which are considered for finding the optimal set of hyperparameters are Adam, RMSProp and SGD. Different versions of RRCNN model are described in the following section. Each version is trained for 150 epochs to find out the final set of hyperparameters.

The results obtained from this optimization is given in Table 4. It is observed that for most of the versions, Adam optimizer with very slow learning rate ranging between $[10^{-5}, 10^{-4}]$ is chosen as optimum. SGD is not been selected for any of these, but RMSProp is selected for two of the versions. For all the versions, the momentum is around 0.8 and batch size remains less than 10. Since the training dataset is not too large, therefore if the batch size remains very much, thereafter some information may get lost during the 1-D batch normalization operation. So if the batch size is more, then the model may perform worse, therefore the optimization algorithm fixes the size to its appropriate.

By this final optimized hyperparameters, the results with different versions have increased from 2% to even up to 6% in classification accuracies. Therefore, Bayesian optimization has played vital role in achieving optimum performance of our model.

5.2.2. Results of RRCNN architecture variants

In this study we have evaluated different versions of the proposed RRCNN network. All the experiments are done in Python environment with the aid PyTorch front end. Among all of the versions, 6 versions are selected as they have given relatively good and comparable results when they are evaluated on our datasets, during various experimentation. The detailed architectural descriptions of the mentioned 6 versions of RRCNNs are given by Table 4.

Different versions of RRCNN are mainly different from each other by means of the number of RS, residual blocks in each RS and dimensions of feature maps of residual blocks. For some cases, where the feature maps of residual blocks of all RS are same, there we have skipped the intermediate bridging convolution layers (C_r). We can see from Table 4 that, such structures are RRCNN-B, RRCNN-C and RRCNN-D. Whereas, in spite of having same intermediate dimensions of residual blocks of different residual stacks (RS) of RRCNN-A, it has intermediate bridging convolution layers. For the versions RRCNN-E and RRCNN-F, the input and output feature maps of different residual blocks of different RS are different from each other therefore bridging convolution layers are necessary.

From Table 6, we observe that among all six combinations, RRCNN-B, RRCNN-C, RRCNN-D and RRCNN-E report promising results (>80%) in most of the cases of this binary classification task. RRCNN-A and RRCNN-F fails to achieve such good results in most cases.

5.2.3. Selection of appropriate channel sets

In this study, we have used four channel sets F3, F7 for approach motivation, and F4, F8 for avoidance motivation. Before arriving at the conclusion of using these channels sets only, we have performed experimentation on various combinations of different channel sets also, which combinations are proposed in various papers, investigated for different domains. Such as Heyat, Lai, Khan, and Zhang (2019) have analysed C4-P4 and C4-A1 channel sets for bruxism detection during deep sleep of human beings, and reports that C4-P4 channel combination outperforms C4-A1 with significant difference. In another paper (Prasad, Liu, Chen, & Quek, 2018), suicidal tendencies of students have been checked by analysing attention (Fz, F4), emotion (F3, F8) and memory (P8, P3, C3, T7) channels. The reported results show that beta waves outperform alpha waves in this case of sentiment analysis task. Similarly for the paper (Song & Sepulveda, 2018), the authors have evaluated channels P1 to P7 by selecting them by the proposed technique in self-paced brain–computer interface task onset. Similarly, (Chai et al., 2017) have selected 16 EEG channels (AF3, F3, FC1, FC5, T7, CP5, P3, O1, P4, P8, CP6, T8, FC2, F8, AF4, FP2) out of 32 channels by using independent component analysis (ICA) and scalp map projection (SMP) to classify driver fatigue by performing EEG analysis.

In this proposed work, we have to finalize the channel sets, we have performed experimentation by using aforementioned different combinations of channel sets also. Since we have used 10–20 electrode system, we have considered only those channels subset from a single channel set which are available in our 10–20 electrode system. The results (accuracy) obtained by different channel sets are reported in Table 5. It is evident from the obtained results that, the combination

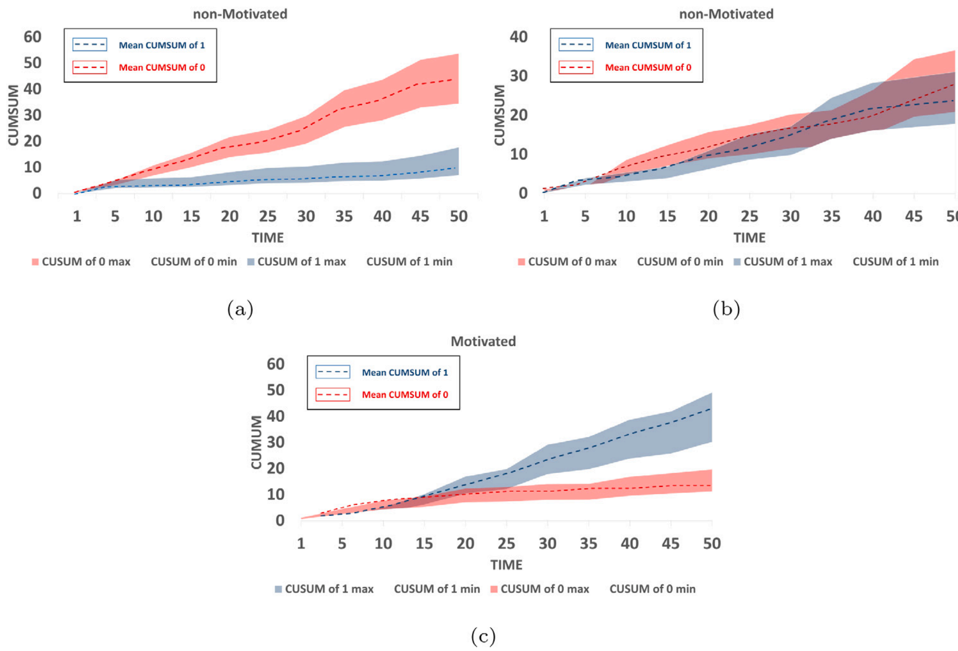

Fig. 6. Envelop plots of three different types of observations obtained from the CUMSUM strategy of data visualization is given here. It has been mentioned that EEG signals, producing plots (a) and (b) are considered as non-Motivated and plot (c) is considered as motivated.

Table 4
Different versions of RRCNN network is given with their architectural details and optimized hyperparameters which are fit for this dataset. In the table **a** and **b** are the number of residual stacks (RSs) and the number of residual blocks in each residual stack in the architecture. Ch is the number of channels in the feature maps of each residual block of RSs.

Version	a	b	Ch	Bridging conv layers (C_r)			Optimized hyperparameters				
				Conv	Input	Output	$\alpha \times 10^{-4}$	Op	bs	γ	$\beta \times 10^{-4}$
A	3	5	32	Cr1 Cr2	32 32	32 32	1.23	RMSProp	4	0.787	87.7
B	3	5	64	No bridging conv layer			0.18	Adam	6	0.812	3.257
C	3	7	64	No bridging conv layer			0.21	Adam	6	0.809	14.2
D	4	5	64	No bridging conv layer			6.23	Adam	7	0.779	0.976
E	3	5	RS-32 RS-64 RS-128	Cr1 Cr2	32 64	64 128	0.435	Adam	4	0.859	4.615
F	3	5	RS-16 RS-32 RS-64	Cr1 Cr2	16 32	32 64	2.36	RMSProp	5	0.758	0.559

of the approach and the avoidance motivation channels, that is F3, F7, F4 and F8 achieves the maximum among others. However, if the emotion channels (T7 and T8) are further added with the motivation channels, then a little dip in the performance of the RRCNN-C has been observed, since it achieves 84% accuracy which is significantly less than the maximum. Therefore we can arrive at the conclusion that, selection of motivation channels and emotion channels significantly boosts the performance of RRCNN-C in determining the motivation level of a student.

5.2.4. Channel sets

It has been mentioned earlier that we have considered two different channel sets and the combination of those channel sets for analysis purpose. The results of the analysis are given in Table 6. We derived the following observations. For all the cases, the results obtained with these brain channel sets are more than 70%. This clearly indicates that the motivation state is manifested in EEG signals. Further, the frontal asymmetry channel set provides significantly better accuracy than the other channel sets. Most of the accuracies of all cases with frontal asymmetry index are more than 80% with the highest of 89%. Whereas for emotion it is 81% and for 'Both', it is 84%. This suggests that, frontal asymmetry index carries most important information to detect whether a participant is motivated or not.

5.2.5. Role of alpha and beta waves

The accuracies achieved by alpha and beta waves are also compared in Table 6 respectively. As a whole, it is observed that, the results obtained by beta waves are better than that of the alpha waves. In 11 out of 18 cases accuracies with beta waves dominate over accuracies with alpha waves. This fact signifies that signals associated to conscious thought, carry more discriminatory and distinctive information than that of the signals which cause reflective lower activity in detection of motivation among participants.

In this study of motivation detection, we have considered the thought process data for analysis purpose. We know that beta waves are associated to works which require high level of concentration. Now if a person is motivated enough to learn the game is directly related to the fact that how much concentration he imparts in playing the game. If the player is not playing the game with enough concentration then he/she will not learn it, that is, he/she is not enough motivated.

Table 5
Comparison with various EEG channel sets, mentioned in various studies of EEG analysis.

Paper	Used channel sets	Wave	Accuracy
Heyat et al. (2019)	C4-P4	Beta	69%
Heyat et al. (2019)	C4-A1	Beta	66%
Prasad et al. (2018)	Fz, F4, F3, F8, P8, P3, C3, T7	Beta	73%
Song and Sepulveda (2018)	P3, P4, P7, P8	Alpha	70%
Chai et al. (2017)	AFz, F3, T7, T8, CPz, P3, P4, P8,	Beta	61%
RRCNN-C	F3, F7, F4, F8, T7, T8	Alpha/Beta	84%
RRCNN-C	F3, F7, F4, F8	Alpha	**89%**

Table 6
Classification results obtained using different channel sets, RRCNN versions, and alpha versus beta waves. Bold indicates row-wise best performance while underline indicates column-wise best performance.

RRCNN	Alpha waves			Beta waves		
Version	Frontal asymmetry	Emotions	Both	Frontal asymmetry	Emotions	Both
A	**78%**	72%	75%	71%	74%	71%
B	87%	78%	81%	**88%**	74%	82%
C	**89%**	78%	84%	88%	77%	80%
D	82%	79%	81%	**84%**	81%	83%
E	83%	76%	80%	**85%**	75%	84%
F	76%	73%	70%	**79%**	74%	73%

Therefore beta waves carry more discriminatory features in detection of motivation level of test participants. For frontal asymmetry index and 'Both', in four out of six cases, accuracies achieved using beta waves as the inputs, are more than that of alpha waves as the inputs. Whereas with emotion channel sets, in 50% test cases, beta waves dominate over alpha waves. Therefore, we can conclude that in determining the motivation of a participant, beta waves happen to be more effective than that of alpha waves. This study is further insightful in the sense that the classification accuracy is less dependent on the architectural optimization of the RRCNN when using beta waves, therefore more robust in conducting motivation-based studies in general.

5.3. Computational efficiency

Since from the previously mentioned experiment results we can say that RRCNN-C outperforms all other versions of RRCNN as a whole, in this section we discuss the computational efficiency of RRCNN-C only. Though the model RRCNN-C has 0.55M trainable parameters, it is tend to perform excellently good with noticeably small datasets. In the model, each RSs, and the pre-RSs layer have 0.17M and 25k trainable weights and biases in total. This deep RRCNN-C has been trained with 500 data samples for 300 epochs. In the training process, the model with minimum validation loss has been saved and further used for testing purpose. The model has been trained several times and the time taken in the training process is 143 min or 2 h and 23 min. In the testing process also, the model takes less than a minute to make prediction of 100 test samples. Therefore it is evident that the model is computationally much efficient towards analysing EEG data in an end-to-end manner.

5.4. Overall performance of the model

Previously we have mentioned that our model is less likely to overfit even in small datasets, like our EEG signal dataset of motivation detection. To justify that fact, we discuss the loss characteristics train and validation losses of RRCNN-C on alpha signals.

The variation of train and validation losses during training of the RRCNN-C is given in Fig. 7(a). For comparison purpose, similar loss plots of LSTM (Alhagry, Fahmy, & El-Khoribi, 2017) and GRU (Dey & Salemt, 2017) are also given in Fig. 7(b) and (c). The train losses of these three deep learning classifiers are found almost similar but significant variations can be observed in validation loss plots. The validation loss plot of Fig. 7(a) is much consistently decreasing in nature than the other two. It is observed from the validation loss plot of GRU, that it is not learning much as the validation loss is not converging at all, rather it oscillates. Similar characteristic is seen for the case of LSTM also. LSTM, GRU are state-of-the-art classifiers for sequential data processing, which also fail in this case, where the dataset size is very less, and proposed RRCNN-C learns fits in here too. This signifies the efficiency of proposed architecture.

5.5. Comparison with standard methods of EEG signal processing

Now to estimate the reliability and efficiency of RRCNN model, we have performed some experiments with popular signal processing techniques on the proposed dataset of motivation detection. These techniques include some deep learning models such as LSTM (Alhagry et al., 2017), GRU (Dey & Salemt, 2017) and BiLSTM (Yang, Huang, Wu, & Yang, 2020), and some traditional machine learning approaches. The deep learning models learn relevant features by themselves only, therefore we fed raw alpha and beta signals to those RNN models. On the other hand, for evaluating machine learning based models, we need to extract relevant discriminatory features from raw EEG signals. For this, we have extracted higher order crossing (HOC) (Prasad et al., 2018) features from alpha and beta frequency signals and used SVM (Prasad et al., 2018), KNN (Prasad et al., 2018) and MLP (Chatterjee & Bandyopadhyay, 2016) classifiers for classification purpose. The results obtained by different approaches mentioned above are reported in Table 7. The specifications of the models and input features as well as signals can be seen from the given Table 7. It is observed that RNN based approaches have not brought significantly good accuracies whereas machine learning classifiers such as SVM achieves impressive result. The reason can be that, to address the class imbalance problem in our data, we have considered very few data points for training the classifier. Therefore the RNN models might not have learnt enough discriminatory features to achieve impressive results. As a whole we can say that proposed RRCNN-C achieves the best with a very large margin of difference even in such a small dataset. This concludes the elevated efficiency of our model.

5.6. Evaluation of the RRCNN model across other datasets

In order to better justify the performance of the RRCNN model (RRCNN-C), the model has been evaluated on two other datasets and an ranked based statistical analysis (Friedman test) have also been performed based on the obtained results on these three datasets. The datasets on which the model has been evaluated along with the motivation dataset are the Sentiment analysis dataset proposed by Prasad et al. (2018) and BCI-IV 2a (Gaur et al., 2021) motor-imaginary competition dataset. We have extracted Higher order crossing features of different orders from the given two datasets and applied aforementioned 6 popularly adopted techniques (LSTM, GRU, BiLSTM, SVM, MLP and KNN) for the analysis purpose.

In the Sentiment analysis dataset, there are 3 classes, suicide, positive and neutral respectively. As mentioned in the paper (Prasad et al., 2018), we also have considered memory channels and HOC features of order 50 for the intention of analysing the Sentiment data. Prasad et al. (2018) have achieved maximum of 77.1% 3 class classification accuracy by considering aforementioned channels and features, and

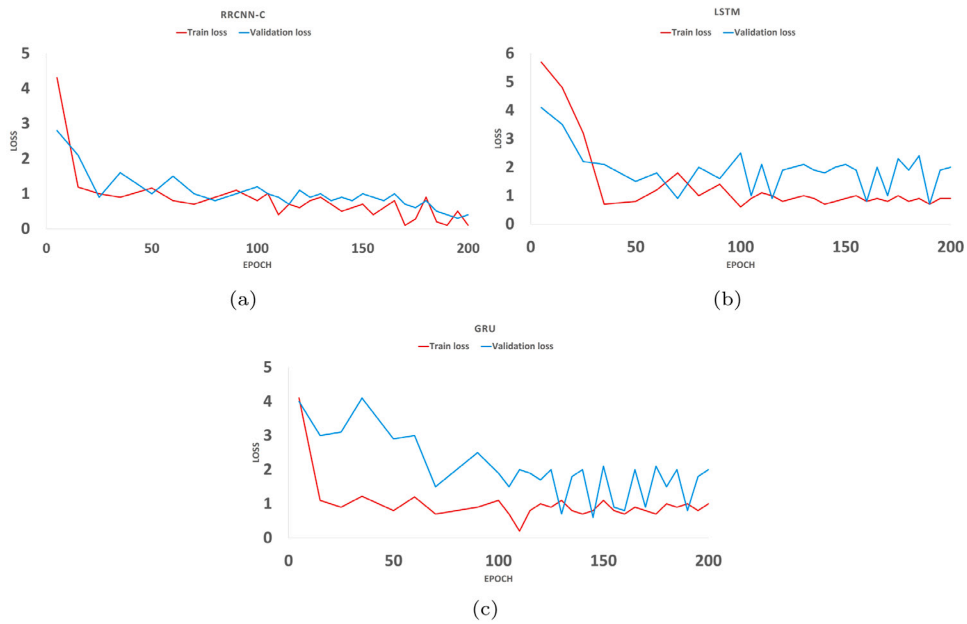

Fig. 7. The variation of train and validation losses with the epochs during the training of the model.

Table 7
A comparative study results of some popularly used signal processing techniques, achieved on the proposed motivation detection dataset.

Method	Features	Specification of the classifier	Obtained accuracy
LSTM	Alpha waves are directly fed to the RNN model	Two LSTM layers are used for classification	71%
GRU	Alpha waves achieve better results than beta waves	Two layers of GRUs have been used in the model	69%
BiLSTM	Beta waves report better results for BiLSTM	Two layers are fixed in the model	73%
SVM	Higher Order Crossing (HOC) features of beta waves with order 30 is used	The chosen kernel function and the regularization parameter are 'rbf' and 5000	81%
MLP	20 orders HOC features of beta waves	Two hidden layers each having five neurones are fixed	77%
KNN	HOC features with order 30, extracted from beta waves are used	The number of nearest neighbour is chosen 5	80%
Proposed RRCNN-C	End-to-end learning model with results better with alpha waves than beta waves	Specifications of RRCNN-C is given in Table	**89%**

with SVM classifier. In our case, the proposed RRCNN-C achieves an accuracy of 74.31% with the given consideration.

The another dataset, BCI-IV 2a dataset is a very popular publicly available motor-imaginary task detection dataset. Four different tasks have been performed while collecting the dataset, these tasks are moving the left hand, moving the right hand, moving the tongue and moving the toe. The objective is to determine the motor-imaginary movement by analysing the raw EEG signal. The data is also collected using a 10–20 electrode system and have 22 EEG channels and 3 EOG channels. The benchmark accuracies reported so far on BCI-IV 2a dataset, are 80% (Gaur et al., 2021), 79.03% (Mane et al., 2021), 81.34% (Deng, Zhang, Yu, Liu, & Sun, 2021), 79.90% (Li, Wang, Xu, & Fang, 2019) and so on. In this case our study utilizes the results reported by Ma, Qiu, and He (2020) for the channels selection purpose. As mentioned by Ma et al. (2020), we have considered only C3, C4 and Cz channels for the analysing the motor-imaginary activities of the test subjects. In this case we have fed raw signals to RRCNN-C and observed that it gives better results than HOC features of orders 30 and 50. On BCI-IV 2a dataset, RRCNN-C achieves very good results of 79.31% 4-class classification accuracy, which is very much comparable to the state-of-the-art results.

5.7. Statistical test

Now we have performed Friedman statistical test (Demšar, 2006) in order to better assess the performance of the RRCNN-C architecture. The algorithm have been compares to 6 other algorithm as mentioned above. For each of the three datasets, the algorithms are ranked on the basis of their performance, and thereafter the Friedman test has been performed and the performance evaluating criterion (χ_f^2) has been calculated by following mathematical formula

$$\chi_f^2 = \frac{12M}{N(N+1)} \sum_{n=1}^{N} [(r_n^{av})^2 - \frac{N(N+1)^2}{4}] \tag{25}$$

where M, N are number of datasets and the number of algorithms respectively, and r_n^{av} is the average rank of nth algorithm among the others, which can be calculated as follows

$$r_n^{av} = \frac{1}{M} \sum_{m=1}^{M} r_n^m \tag{26}$$

The statistical test results are reported in Table 8, where the obtained χ_f^2 is 11.59, which is more than the Friedman constant at N=7 and M=3 ($\chi^2_{N-1=7,M=3,\alpha=0.5}$ = 5.39). Therefore we can successfully reject the null-hypothesis and hence, the statistical significance of the proposed RRCNN-C has been proved.

Table 8
Friedman statistical test for justifying the results obtained RRCNN-C. Where Dataset #1, Dataset #2 and Dataset #3 are the motivation, the sentiment analysis, and the BCICIV-2a datasets respectively.

Algorithms	#1	#2	#3	Average rank	χ_f^2
RRCNN-C	1	2	1	1.33	**11.59**
LSTM	6	5	3	4.66	
GRU	7	6	6	6.33	
BiLSTM	5	7	5	5.66	
SVM	2	1	3	2	
MLP	4	4	7	5	
KNN	3	3	2	2.33	

5.8. Hypothesis validation and discussion

In this study, the proposed hypothesis has been validated with statistical approach for data visualization and data understanding, and machine learning approach for detecting motivation level. From the CUMSUM results reported by Fig. 6 and Table 3, it is evident that the participants of the group ER/P learn the most among all three groups. That is, if negative reward and positive reward both are present in a consecutive manner, then the motivation level of the participants remains high in accomplishing the task. This type of formation of different rewards makes the task more challenging, where the participants have something to loss when they make any mistake and something to gain after a success. This phenomena keeps the motivation level intact to a certain level. These conclusive results strongly support the hypothesis, that we have proposed in Section 3. Now on the basis of the results obtained by statistical analysis, we have segregated the data into two divisions-motivated and non-motivated. Our proposed RRCNN model has been evaluated on the dataset and achieved 89% classification accuracy in detecting motivation level of the participants. This is how in this paper we validate the proposed hypothesis and develop an entirely new approach of motivation detection via EEG signals processing.

6. Conclusion

Our work indicates that relatively abstract state of motivation is indeed encoded physiologically into the EEG signal. It also illustrates that motivational state can be derived through a suitable brain–computer interface and data analysis. Further, these motivational states can be used with suitably designed deep learning mechanisms to assess their physiological manifestation in the form of EEG signals.

The above conclusions were derived through a methodology with several novel and original approaches. They include a novel game-based motivation stimulating brain computer interface, a rigorous approach for identification of the most relevant EEG channels and features, a novel small-data amenable deep learning architecture RRCNN, extensive validation of the variety of aspects in our methodology, and detailed statistical analysis.

We show up to 89% accuracy in associating the EEG signals to the motivational state. The study further indicates that identification of the channel sets that indicate the motivational state is important, and in our case frontal asymmetry set comprising of F3, F4, F7, and F8 have a stronger signature of the state of motivation. We also show that even though acquiring large datasets for such studies is practically limiting and time consuming, new deep learning architectures specifically designed to deal with deep features on small datasets, such as the proposed RRCNN, can provide significant value for EEG based studies.

CRediT authorship contribution statement

Soham Chattopadhyay: Creation of models, Software, Validation, Formal analysis, Writing - original draft, Visualization, Writing of this manuscript. **Laila Zary:** Methodology, Investigation, Data curation, Writing - original draft, Writing of this manuscript. **Chai Quek:** Conceptualization, Methodology, Resources, Writing - original draft, Supervision, Project administration, Funding acquisition, Writing of this manuscript. **Dilip K. Prasad:** Conceptualization, Methodology, Software, Validation, Resources, Data curation, Writing - original draft, Supervision, Funding acquisition, Writing of this manuscript.

Declaration of competing interest

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

Code and reproducibility

We have shared the reproducible code at codeocean : https://codeocean.com/capsule/2144609/tree/v1 For more details about the source code, dataset and other details related to this work : https://github.com/SohamChattopadhyayEE/RRCNN.

Acknowledgement

SC acknowledges internship funding support from Research Council Norway's INTPART grant no. 309802.

References

Alhagry, S., Fahmy, A. A., & El-Khoribi, R. A. (2017). Emotion recognition based on EEG using LSTM recurrent neural network. *Emotion*, *8*(10), 355–358.
Amin, S. U., Alsulaiman, M., Muhammad, G., Mekhtiche, M. A., & Hossain, M. S. (2019). Deep learning for EEG motor imagery classification based on multi-layer CNNs feature fusion. *Future Generation Computer Systems*, *101*, 542–554.
Arnold, M. B. (1960). *Emotion and personality: neurological and physiological aspects, Vol. 2*. Columbia University Press.
Baker, R., et al. (2010). Data mining for education. *International Encyclopedia of Education*, *7*(3), 112–118.
Blinowska, K., & Durka, P. (2006). Electroencephalography (eeg). *Wiley Encyclopedia of Biomedical Engineering*.
Buck, R. (1988). *Human motivation and emotion*. John Wiley & Sons.
Calvo, R. A. (2010). Latent and emergent models in affective computing. *Emotion Review*, *2*(3), 288–289.
Chai, R., Naik, G. R., Ling, S. H., Tran, Y., Craig, A., & Nguyen, H. T. (2017). Channels selection using independent component analysis and scalp map projection for EEG-based driver fatigue classification. In *2017 39th annual international conference of the ieee engineering in medicine and biology society (EMBC)* (pp. 1808–1811). IEEE.
Chatterjee, R., & Bandyopadhyay, T. (2016). EEG Based motor imagery classification using SVM and MLP. In *International conference on computational intelligence and networks (CINE)* (pp. 84–89). IEEE.
Chin, K.-Y., Hong, Z.-W., & Chen, Y.-L. (2014). Impact of using an educational robot-based learning system on students' motivation in elementary education. *IEEE Transactions on Learning Technologies*, *7*(4), 333–345. http://dx.doi.org/10.1109/TLT.2014.2346756.
Christophel, D. M. (1990). The relationships among teacher immediacy behaviors, student motivation, and learning. *Communication Education*, *39*(4), 323–340.
Dai, M., Zheng, D., Na, R., Wang, S., & Zhang, S. (2019). EEG Classification of motor imagery using a novel deep learning framework. *Sensors*, *19*(3), 551.
Davidson, R. J., Ekman, P., Saron, C. D., Senulis, J. A., & Friesen, W. V. (1990). Approach-withdrawal and cerebral asymmetry: emotional expression and brain physiology: I.. *Journal of Personality and Social Psychology*, *58*(2), 330.
De Vicente, A., & Pain, H. (2002). Informing the detection of the students' motivational state: an empirical study. In *International conference on intelligent tutoring systems* (pp. 933–943). Springer.
Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, *7*, 1–30.
Deng, X., Zhang, B., Yu, N., Liu, K., & Sun, K. (2021). Advanced TSGL-eegnet for motor imagery EEG-based brain-computer interfaces. *IEEE Access*, *9*, 25118–25130.
Derbali, L., Chalfoun, P., & Frasson, C. (2011). A theoretical and empirical approach in assessing motivational factors: From serious games to an ITS. In *Twenty-fourth international flairs conference*.
Dey, R., & Salemt, F. M. (2017). Gate-variants of gated recurrent unit (GRU) neural networks. In *IEEE international midwest symposium on circuits and systems (MWSCAS)* (pp. 1597–1600). IEEE.
Elliot, A. J. (2008). *Handbook of approach and avoidance motivation*. New York: Psychology Press.

Elliot, A. J., & Covington, M. V. (2001). Approach and avoidance motivation. *Educational Psychology Review*, *13*(2), 73–92.

Elliot, A. J., Gable, S. L., & Mapes, R. R. (2006). Approach and avoidance motivation in the social domain. *Personality and Social Psychology Bulletin*, *32*(3), 378–391.

Esen, H., Esen, M., & Ozsolak, O. (2017). Modelling and experimental performance analysis of solar-assisted ground source heat pump system. *Journal of Experimental & Theoretical Artificial Intelligence*, *29*(1), 1–17.

Esen, H., Inalli, M., Sengur, A., & Esen, M. (2008a). Artificial neural networks and adaptive neuro-fuzzy assessments for ground-coupled heat pump system. *Energy and Buildings*, *40*(6), 1074–1083.

Esen, H., Inalli, M., Sengur, A., & Esen, M. (2008b). Forecasting of a ground-coupled heat pump performance using neural networks with statistical data weighting pre-processing. *International Journal of Thermal Sciences*, *47*(4), 431–441.

Esen, H., Inalli, M., Sengur, A., & Esen, M. (2008c). Performance prediction of a ground-coupled heat pump system using artificial neural networks. *Expert Systems with Applications*, *35*(4), 1940–1948.

Esen, H., Ozgen, F., Esen, M., & Sengur, A. (2009). Artificial neural network and wavelet neural network approaches for modelling of a solar air heater. *Expert Systems with Applications*, *36*(8), 11240–11248.

Fulmer, S. M., & Frijters, J. C. (2009). A review of self-report and alternative approaches in the measurement of student motivation. *Educational Psychology Review*, *21*(3), 219–246.

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming*, *33*(4), 441–467.

Gaur, P., Gupta, H., Chowdhury, A., McCreadie, K., Pachori, R. B., & Wang, H. (2021). A sliding window common spatial pattern for enhancing motor imagery classification in EEG-bci. *IEEE Transactions on Instrumentation and Measurement*, *70*, 1–9.

Gollan, J. K., Hoxha, D., Chihade, D., Pflieger, M. E., Rosebrock, L., & Cacioppo, J. (2014). Frontal alpha EEG asymmetry before and after behavioral activation treatment for depression. *Biological Psychology*, *99*, 198–208.

Harandi, S. R. (2015). Effects of e-learning on students' motivation. *Procedia-Social and Behavioral Sciences*, *181*, 423–430.

Hecht-Nielsen, R. (1992). Theory of the backpropagation neural network. In *Neural networks for perception* (pp. 65–93). Elsevier.

Heyat, M. B. B., Lai, D., Khan, F. I., & Zhang, Y. (2019). Sleep bruxism detection using decision tree method by the combination of C4-p4 and C4-a1 channels of scalp EEG. *IEEE Access*, *7*, 102542–102553.

Horan, W. P., Wynn, J. K., Mathis, I., Miller, G. A., & Green, M. F. (2014). Approach and withdrawal motivation in schizophrenia: an examination of frontal brain asymmetric activity. *PLoS One*, *9*(10), Article e110007.

Jatupaiboon, N., Pan-ngum, S., & Israsena, P. (2013). Real-time EEG-based happiness detection system. *The Scientific World Journal*, *2013*.

Li, D., Wang, J., Xu, J., & Fang, X. (2019). Densely feature fusion based on convolutional neural networks for motor imagery EEG classification. *IEEE Access*, *7*, 132720–132730.

Ma, X., Qiu, S., & He, H. (2020). Multi-channel EEG recording during motor imagery of different joints from the same limb. *Scientific Data*, *7*(1), 1–9.

Mane, R., Chew, E., Chua, K., Ang, K. K., Robinson, N., Vinod, A., et al. (2021). Fbcnet: A multi-view convolutional neural network for brain-computer interface. arxiv preprint arxiv:2104.01233.

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and Emotion*, *23*(2), 209–237.

Nehme, M. (2010). E-learning and students' motivation. *Legal Educational Review*, *20*, 223.

Ng, R., Fishman, I., & Bellugi, U. (2015). Frontal asymmetry index in williams syndrome: Evidence for altered emotional brain circuitry?. *Social Neuroscience*, *10*(4), 366–375.

Olszewska, J. I. (2021). The virtual classroom: A new cyber physical system. In *2021 IEEE 19th world symposium on applied machine intelligence and informatics (SAMI)* (pp. 000187–000192). http://dx.doi.org/10.1109/SAMI50585.2021.9378678.

Özdenizci, O., Wang, Y., Koike-Akino, T., & Erdoğmuş, D. (2019). Adversarial deep learning in EEG biometrics. *IEEE Signal Processing Letters*, *26*(5), 710–714.

Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education: theory, research, and applications*. Prentice Hall.

Ploberger, W., & Krämer, W. (1992). The CUSUM test with OLS residuals. *Econometrica*, 271–285.

Prasad, D. K., Liu, S., Chen, S.-H. A., & Quek, C. (2018). Sentiment analysis using EEG activities for suicidology. *Expert Systems with Applications*, *103*, 206–217.

Reeve, J. (2014). *Understanding motivation and emotion*. John Wiley & Sons.

Richter, G., Raban, D. R., & Rafaeli, S. (2015). Studying gamification: the effect of rewards and incentives on motivation. In *Gamification in education and business* (pp. 21–46).

Rouanne, V., Śliwowski, M., Costecalde, T., Benabid, A. L., & Aksenova, T. (2021). Detection of error correlates in the motor cortex in a long term clinical trial of ecog based brain computer interface.

Schirrmeister, R. T., Springenberg, J. T., Fiederer, L. D. J., Glasstetter, M., Eggensperger, K., Tangermann, M., et al. (2017). Deep learning with convolutional neural networks for EEG decoding and visualization. *Human Brain Mapping*, *38*(11), 5391–5420.

Snoek, J., Rippel, O., Swersky, K., Kiros, R., Satish, N., Sundaram, N., et al. (2015). Scalable bayesian optimization using deep neural networks. In *International conference on machine learning* (pp. 2171–2180).

Song, Y., & Sepulveda, F. (2018). A novel technique for selecting EMG-contaminated EEG channels in self-paced brain–computer interface task onset. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *26*(7), 1353–1362.

Sutton, S. K., & Davidson, R. J. (1997). Prefrontal brain asymmetry: A biological substrate of the behavioral approach and inhibition systems. *Psychological Science*, *8*(3), 204–210.

Tatum IV, W. O. (2014). *Handbook of eeg interpretation*. Demos Medical Publishing.

Touré-Tillery, M., & Fishbach, A. (2014). How to measure motivation: A guide for the experimental social psychologist. *Social and Personality Psychology Compass*, *8*(7), 328–341.

Yang, J., Huang, X., Wu, H., & Yang, X. (2020). EEG-Based emotion classification based on bidirectional long short-term memory network. *Procedia Computer Science*, *174*, 491–504.

Zary, L. (2018). *Modelling of multidimensional electroencephalogram signal data to detect human motivational states* (Master's thesis), Stockholm University.