Imaginary Control of a Mobile Vehicle Using Deep Learning Algorithm: A Brain Computer Interface Study

AMIN HEKMATMANESH1, HAMED MOHAMMADI AZNI2, HUAPENG WU1, MOHSEN AFSHARCHI2, MING LI1, AND HEIKKI HANDOOS1, (Member, IEEE)

1Laboratory of Intelligent Machines, LUT University, 53850 Lappeenranta, Finland
2Department of Computer Engineering, University of Zanjan, Zanjan 45371-38891, Iran

Corresponding author: Amin Hekmatmanesh (amin.hekmatmanesh@lut.fi)

ABSTRACT Controlling a remote mobile vehicle using electroencephalograph (EEG) signals is still a challenge specially achieving a high degree of accuracy and precision. In the present study, the focus is on implementing an efficient feature space in a deep-based learning (DL) algorithm for a single trial application. More specifically, a boosting feature algorithm by means of long short-term memory (LSTM) networks are implemented in a deep auto-encoder (DAE) algorithm for producing an effective feature space tp identify event related desynchronization/event related synchronization (ERD/ERS) patterns in EEG signals. For this purpose, three different DL-based algorithms are implemented that the models are based on a convolutional neural network (CNN), DAE, and LSTM networks to extract and boost the main features. In addition, our previous improved support vector machine (SVM)-based algorithm is employed to consider the potential of SVM and implemented DL-based algorithms for a two classes identification. To consider the efficiency of our implemented methods, algorithms are employed for control of a remote mobile vehicle in an imaginary right-hand opening and making a right-hand fist task. In our experiment, eleven subjects participated in an imaginary movement task. In the experiment, the displayed movement pictures were colored in yellow and red colors for stimulating brain to generate stronger ERD/ERS patterns. Results showed that the proposed algorithm by using the boosting technique significantly increased the accuracy with a higher precision of 73.31% ± 0.03. The proposed method enabled the DL algorithm to be used in single trial experiments.

INDEX TERMS Brain–computer interface, EEG signal processing, event related desynchronization, deep learning.

I. INTRODUCTION

The brain is the main organ that controls the human body and several techniques have been developed to decode the brain neurons activities to explore how the human body controlled. Several mathematical-based studies have been investigated on different aspects of human brain activity such as stress during driving [1], voluntary actions [2], involuntary actions such as breathing [3], and the sleep process [4]. Brain stroke disability is a terrible situation for people who have had an active role in society. Hence, these patients need full-time assistance for their normal daily activities. Brain computer interface (BCI) science has been developed to be a lifesaving solution.

Different neurons in different locations in the brain produce different patterns, which are identifiable using electroencephalograph (EEG) signals [5]. Here, we focus on the central area which is related to the (imaginary) movement patterns, namely event related desynchronization/event related synchronization (ERD/ERS). One important aspect of research focuses on the intention to move (ERD pattern) and the onset of real body movements (ERS pattern). Some of the related known patterns to the movements are error related potential [6], readiness potential [3], ERD/ERS [7], [8], and event related potentials (ERP) patterns such as P300 and steady state visually evoked potential patterns [9].

Neurons have the potential to produce imaginary movement patterns similar to real movement patterns with smaller amplitudes [10]. The investigated topics related to imaginary...
movement patterns include: controlling a mobile vehicle [11], controlling a prosthetic hand (wearable robots, exoskeleton robots) [12], [13], controlling appliances such as lamps [14], controlling aerial vehicles such as quadcopters [15], [16], control of a vehicle in the main directions [17], break assistant commands also named emergency breaks [18], [19], change lane commands [20] and steering control [11], [21]. In the above-mentioned topics, two considerable issues have been investigated, namely accuracy and time delay. Several methods have been developed for the automatic identification of early imaginary patterns from 5 ms to 2 s for the applications [16], [22].

Some relative BCI studies to our method and application are explained in below. In a primitive study, Haufe et al. [23] designed an algorithm to identify the ERD/ERS patterns of EEG related to the emergency brake intention. Simultaneously, electrocardiogram (EMG) signals measured the muscular reactions to emergency braking tasks. In the algorithm, the selected features were the area under the identified ERP patterns from the EEG and EMG signals. Features are then classified using a regularized linear discriminant analysis (RLDA) classifier. One key point of the study was combining the EMG and EEG features to increase the accuracy and precision of identifying emergency break patterns. The reported results showed significant achievements in accuracy and response time to emergency braking. Additionally, the location of the neurons related to the emergency braking were determined. The research limitations was recognized as follows: 1) employing a low number of subjects; 2) extracting a small number of features for training a classifier; and 3) utilizing a binary classifier [23].

Afterwards, Kim et al. [24] developed the idea by Haufe et al. [23] by increasing the number of identification classes and number of features using different patterns. In the algorithm, ERD/ERS and ERP patterns were extracted by means of applying frequency filtering between 5-35 Hz. Next, the Hilbert transformation was employed for feature extraction. The features were then identified using the RLDA classifier. The achieved accuracy increased due to using different types of features and extracting a larger number of features, but low precision (high accuracy variation) was reported. In conclusion the study, covered some limitations of the previous studies, but usage of the RLDA classifier was a limitation, because it is more effective in binary identification conditions than multi-class identification [20], [25].

In our previous series of studies, we developed different algorithms for control of a mobile vehicle for moving forward and breaking (stopping) [12], [26]-[28]. Additionally, we implemented methods to solve the above-mentioned limitations such as the number of subjects, limited number of features, feature selection and optimization of the classifiers. For example, we [8] extracted a discriminative filterbank common spatial pattern (DFBCSP) features, which were optimized by a discriminative sensitive learning vector quantization (DSLVQ) training algorithm. In the DSLVQ optimization algorithm, features were updated based on the number of their repetitions, which means the values with high numbers of repetitions were recognized as the EEG background and diminished by multiplication of a small coefficient and the values with low numbers of repetitions were identified as the aim voluntary patterns and ascended by multiplication of a large coefficient. Then, the linear discriminant analysis (LDA) and principal component analysis feature selection algorithms were applied to select effective features. Finally, 14 supervised classifiers were employed to identify the ERD/ERS patterns. The best selected classifier was the optimized soft margin support vector machine (SSVM) classifier, which is used by the generalized radial bases functions (GRBF) kernel (SSVM-GRBF) function [8], [28]. The SSVM algorithm was an optimized regularization support vector machine (SVM) method for finding the best soft margin area for decision-making in the feature space. In addition, three free parameters were added to the Gaussian function that enabled promising coverage of the scattered features in each class. This was named generalizing the radial bases functions.

Then, we employed the DSLVQ feature optimizer and the SSVM-GRBF classifier in our studies. The significant cons of our studies were the limitation of the CSP for two classes, besides, the SSVM is a binary classifier which is a constraint. Several methods have been developed to use the CSP and SVM as multi-class solutions in which the results showed a drop in the results [1], [29], [30].

Later on, we designed an algorithm based on a customized mother wavelet for a wavelet packet to identify ERD/ERS patterns of individual subjects. The aim of the study was to solve the limitations in [23] and [24], which involved using a constant mother wavelet to find the ERD/ERS patterns which vary continuously and cause low precision [23]. For this purpose, we designed a customized mother wavelet based on an individual subject’s ERD/ERS patterns for a wavelet-packet with 20 different frequencies (variations). The components related to the ERD/ERS were then computed and a new signal was calculated using a detrended fluctuation analysis that enabled us to extract long-term correlation features [12], [21]. In the wavelet-based study, the SSVM-GRBF classifier [8], [28] was then used to classify the features. The results improved significantly, but a considerable limitation exists, which was at least two second delay in the real-time experiment.

Afterwards, we implemented a chaotic-based feature extraction algorithm to identify the moving forward and stopping (brake) classes. The largest Lyapunov exponent (LLE) feature is a known feature in nonlinear identification and prediction systems [5], [31]. The LLE concept in our study was produced a trajectory of a nonlinear system during imagination of opening a hand and make a fist by reconstructing a phase space based on a delayed EEG signal. The obtained trajectories were not distinctive for all of the imaginations. Therefore, the principals of the LLE (mutual information and false nearest neighbors) were optimized using chaotic tug of war and water drop optimizer algorithms [26], [27].
The results showed a distinctive vision for the reconstructed trajectories in the phase space in comparison with the traditional LLE. The main limitation of the algorithm was the time consuming optimization algorithm which is required for all new incoming data, that causes large delay for real-time systems. The recent investigation employed Deep Learning (DL) as a successful method for classification in EEG [32], images [33], and speech processing [34]. Recently, several studies have been published which identify the ERD/ERS using DL for different purposes such as control of a prosthetic hand [32] and braking assistant applications [19]. For example, Zhang et al. [19] employed wavelet components to compute the power spectrum density and canonical correlation analysis features. Then, DL, SVM and ensemble classifiers were used that the ensemble classifier achieved the best result. The drawback of the study was in utilizing the traditional SVM, not enough input data for DL classifiers without optimization. Additionally, the authors used a time-consuming algorithms such as wavelets, which are not suitable for the real-time systems. By comparing studies, the DL algorithm has potential of achieving higher accuracies if enough data feed to the DL algorithm. The limitation of DL is a large number of measured input values is required for training a DL algorithm.

The next generation of studies, employed DL algorithms with different sensors to measure the environment- and vehicle situation by means of external sensors to identify emergency brake and navigating vehicles. Therefore, a complementary information is added to the feature space (based on the biosignals) and enough inputs is provided for the DL algorithm. In short, external sensors have the potential to provide informative information such as auditory information [35], the air condition (foggy, sunny and rainy) vehicle condition (acceleration, velocity, wheel angle and gas pedal angle), and camera data [36] to increase the accuracy with significant precision [9], [37]–[40]. Comprehensive details of the algorithms for brain-controlled mobile vehicles and aerial vehicles are available in our review paper [41].

In the present study, our contribution involves stimulating imaginary movement in the brain using imagination and colors. Then, a DL algorithm is implemented to extract imaginary movement patterns for control of a mobile vehicle, as a single trial application. The implemented DL model includes deep auto-encoder (DAE). The DAE features are then boosted using the long short-term memory (LSTM) algorithm. Next, a feature space using the boosted and DAE features is formed and then features classified using DL-based classifiers. Regarding the algorithms three different classifiers were designed and results were compared. Additionally, our best achieved method in the previous study [28] was used here to identify the imaginary patterns and then compared with the DL-based methods. The rest of paper is continues as follows: II- data acquisition and experimental setup; III- methods; IV- results; V- discussion; and VI- conclusion.

II. DATA ACQUISITION AND EXPERIMENTAL SETUP
The employed technique to record data in the present study is the same technique as in our previous investigations in [8], [17], [21]. In short, the executed task took place in four steps as follows: 1) showing a fixation cross at the center of a black screen to attract the subject’s attention to the center of the screen, 2) showing a sketch of a closed fist and open-hand, which are colored in red and yellow, respectively, 3) imagination of the seen colored sketch after disappearing the sketch pictures, and 4) resting for a random period of time. The reason for using colors in the sketches is simultaneous stimulation of emotion (colored-base) and imaginary areas of the brain [42], [43] to produce more distinctive patterns. The experiment was repeated for 150 cycles for eleven subjects, 75 tasks were given each for making a fist and opening the right hand. The employed amplifier for recording the EEG was the Enobio32 portable gel-electrode with a sampling rate of 500 Hz. The 32 electrodes where installed on a stretchable cap based on the international 10/20 system. The algorithm implemented in Google Colab using the Python language with the Tensorflow and Keras libraries.

III. METHODS
In the present study, a new configuration of the DL based on the combination of the DAE with the CNN and LSTM was implemented to generate informative features to improve the accuracy and precision. The concept of the algorithm is illustrated in Figure 1 and the details of the method are illustrated in Figure 2. To this end, the first step is the preprocessing which is explained in the following section:

A. PREPROCESSING
Due to the large size of the data, we employed batch technique to read the EEG data in multiple batch packages. The 32 EEG channel signals were first divided into segments of 2500 msec (500 msec before hiding the pictures and 2000 msec after the moment of hiding the pictures). Each batch was passed through a fifth order band pass IIR Butterworth filter with the edges set to 7 Hz to 30 Hz and then normalized. The filter edges were selected experimentally based on the obtained accuracies and previous studies [44]. Therefore, the obtained filtered matrix size for individual subjects was $2500 \times 32 \times 150$. The algorithm configured 150 batches of $2500 \times 32$ matrices for each individual subject. In the following section, our DL configuration is explained.

B. CONVOLUTIONAL NEURAL NETWORK CONSTRUCTION
The DAE includes three main parts, encoder, middle layer and decoder. The input layer feeds the encoder to map the input into a feature space which has the same dimensions as the input EEG signal. Then, multiple hidden layers consisting of pairs of CNNs and pooling layers are set to reach the optimum feature space (Figure 2). 2) The middle layer includes the optimum feature values, which are used as the main features for the next computations. In our algorithm, the DAE has four
pairs of layers of the CNN and max-pooling for the encoder and also has similar layers for the inverted structure as the decoder layers. Shortly, four encoders, one middle layer and four decoders. In each hidden layer, max-pooling halved the length of the data ($A_{2500 \times 32}$ to $A_{1250 \times 32}$) to reach an optimum size, as illustrated in Figure 2. The optimum size of features means features have the potential to regenerate the input EEG data using a decoder that is a sign of storing the main EEG information. In our computations, the determined optimum size of the feature space for each trial of opening the hand or making a fist was $A_{512 \times 78}$.

In detail for each layer, the rectified linear units (ReLU) function was set in the layers and weights updated using a forward- and backward propagation technique. On the CNN layers, a filter size of $1 \times 5$ is convolved to the input data that generate the same size data, and then applied on the signal. High-level features are then computed by selecting half of data using a max-pooling approach. This procedure is performed for each pair of the CNN and max-pooling layers. In the training phase, a dropout technique is used to prevent overfitting, therefore 30% of the connected ReLU functions are removed randomly from updating the computations.
Afterwards, a fully connected Softmax layer neural network is then followed for the classification step.

To consider the computations, assume the size of the neurons in the CNN layer is \( N \times N \) and the filter \( (\omega) \) size for the CNN weights is \( l \times l \), then the formulation for the forward propagation procedure is formulated as follows [45]:

\[
x_{ij}^l = \sigma \left( \sum_{a=0}^{N-l} \sum_{b=0}^{N-l} \omega_{ab} y_{i+a}^{l-1}(j+b) \right),
\]

where \( y \) is the output of the neuron’s activation function, \( m \) is the number of hidden layers, and \( \sigma \) is the weighting matrix. In the max-pooling part, the segment’s size \( (l \times l) \) is selected and the maximum value is then replaced. The final size of the input would be \( \frac{N-a+1}{l} \times \frac{N-b+1}{l} \). In our case, the down sampling involved max-pooling with \( l = 2 \) (matrix size \( 1 \times 2 \)). The error \( (Er) \) value in the forward propagation is computed by:

\[
\frac{\partial Er}{\partial \omega_{ab}} = \sum_{i=0}^{N-n} \sum_{j=0}^{N-n} \frac{\partial Er}{\partial y_{ij}^l} \frac{\partial y_{ij}^l}{\partial \omega_{ab}} = \sum_{i=0}^{N-n} \sum_{j=0}^{N-n} \frac{\partial Er}{\partial y_{ij}^l} \omega_{ab} y_{i+a}^{l-1}(j+b),
\]

where the gradient formula is computed as follows:

\[
\frac{\partial Er}{\partial y_{ij}^l} = \frac{\partial Er}{\partial y_{ij}^l} \frac{\partial y_{ij}^l}{\partial x_{ij}^l} = \frac{\partial Er}{\partial y_{ij}^l} \sigma \left( x_{ij}^l \right),
\]

Finally, the back propagation error \( (\frac{\partial Er}{\partial y_{ij}^l}) \) is computed using Eq. (4):

\[
\sum_{a=0}^{n-l} \sum_{b=0}^{n-l} \frac{\partial Er}{\partial x_{i+a}^{l-1}(j+b)} \frac{\partial x_{i+a}^{l-1}(j+b)}{\partial y_{ij}^l} = \sum_{a=0}^{n-l} \sum_{b=0}^{n-l} \frac{\partial Er}{\partial y_{ij}^l} \omega_{ab}.
\]

After updating the weights, the optimized weights in the middle layer are computed. The optimized weights are the compressed features named DAE features. In the next part, we explain how the AE algorithm works.

### C. AUTO ENCODER

To obtain the DAE features, the prepared EEG matrix is fed into a deep neural network (DNN) which has nine layers. The middle layer contains the DAE features, as shown in Figure 2. In the DNN procedure, each hidden layer’s weight is computed using the CNN and then down sampled using max-pooling filters. The weights are then updated and optimized utilizing a forward-backward propagation technique. The encoder part of the algorithm reduces the input feature-space size until it reaches the optimum size with the capability of regenerating the input EEG signal by means of a set of decoder layers. The DAE features \([A_{l}^{12 \times 78}]\) are fed into the LSTM algorithm and, simultaneously a copy of the DAE features is kept to be combined with the LSTM features. In the next step, we present how the DAE features in the middle layer are boosted by using the LSTM and recurrent neural network (RNN) algorithms.

![Figure 3. The RNN configuration [47].](image)

### D. RNN AND LSTM FEATURE EXTRACTION

LSTM is a successful method for speech feature extraction, which was introduced by [46] in a speech signal processing study. The combination of LSTM features with a deep RNN classifier achieves promising results, which are shown in Figure 3 and explained as follows:

1) **RECURRENT NEURAL NETWORK**

The RNN network is a series of single RNN networks as shown in Figure 3. A single RNN connection is computed as follows [47]: in the algorithm, \( X(i) \) represents the DAE features, \( H(i) \) is the hidden vector, \( Y(i) \) is the RNN output and the sampling instant is \( i = 0, \ldots, 77 \).

\[
H(i) = \gamma \left( UX(i) + SH(i-1) + \text{bias}_1 \right)
\]

\[
Y(i) = \gamma \left( VH(i) + \text{bias}_2 \right)
\]

where \( U, S \) and \( V \) are the connection weights which are updating in each trial (shown in Figure 3), \( \gamma \) is an activation function, and \( \text{bias}_1 \) and \( \text{bias}_2 \) are constant bias values. In order to implement the LSTM feature extraction based on the RNN, hidden layer neurons in the RNN are replaced by the LSTM blocks. The advantage of the LSTM modification is the good handling of the gradient vanishing problem [48]. The following section shows how the RNN and LSTM are integrated.

2) **LONG SHORT-TERM MEMORY**

To implement our LSTM model regarding the RNN configuration, a series of 77 LSTM blocks were set in the algorithm. A single block diagram of the LSTM is illustrated in Figure 2, so that each LSTM includes three inputs as follows: 1- sequence feature \( X(i), \ i = 0, \ldots, 77 \), 2- memory \( C_{i-1} \), and 3- computed feature \( H_{i-1} \). The LSTM has two outputs of memory \( C_i \) and a computed feature \( H_i \), which are passed to the next blocks. As illustrated in Figure 2, the LSTM is a dependent algorithm based on the previous- states of memory \( C_{i-1} \) and features \( H_{i-1} \). The input memory \( C_i \) is computed as follows:

\[
C_i = \tanh \left( S_{CX} X(i) + S_{CH} H_{i-1} + b_c \right).
\]
Then, the memory \((C_l)\) and hidden features \((H_l)\) are updated by means of the following algorithms:
\[
C_l = F_l \odot C_{l-1} + I_l \odot C_l,
\]
\[
H_l = O_l \odot \tanh \left(C_l\right),
\]
where \(\odot\) is the point-wise multiplication operator; \(F_l\), \(I_l\) and \(O_l\) are the forget, input and output gates, respectively. As shown in Figure 2, inside the LSTM \(F_l\), \(I_l\) and \(O_l\) are computed as follows:
\[
F_l = \gamma \left( S_{FX}X_l + S_{F}H_{l-1} + \text{bias}_F \right),
\]
\[
I_l = \gamma \left( S_{IX}X_l + S_{I}H_{l-1} + \text{bias}_I \right),
\]
\[
O_l = \gamma \left( S_{OX}X_l + S_{O}H_{l-1} + \text{bias}_O \right).
\]
Finally, the computed features for the individual blocks are sorted in a matrix and combined with the obtained DAE features for the classification step. The output layer of the deep algorithm for the classification is a fully connected layer of Softmax functions that are presented as follows:
\[
P(z_t | x) = \frac{P(z_t) P(x | z_t)}{\sum_{l=1}^{L} P(z_t) P(x | z_l)} = \frac{\exp(a_t)}{\sum_{l=1}^{L} \exp(a_l)},
\]
where \(P(c_l)\) and \(P(x | z_l)\) are the prior probability of a class and the conditional probability of a class in \(a_l = \ln(P(z_l) P(x | z_l))\), respectively.

IV. RESULTS
In order to find the best possible algorithm combination, four different classifiers were implemented. Three of the methods were different configurations of the DL, named 1- DAE-LSTM, 2- DAE, and 3- CNN-LSTM (traditional DL). The fourth method is the SSVM-GRBF from our previous studies [17], [28]. The ERD/ERS pattern identification results are presented in Table 1, which are based on the accuracy and statistical analysis of the paired t-test and ANOVA values. The statistical analysis were used to determine the significance of the computed features and achieved accuracies.

V. DISCUSSION
The efficiency of DL identification technique depends on the quality and quantity of data. It should be considered that the DL may not solve all constraints of a study concurrently, but it could be a well-promising solution for identification problems if enough data is fed into the algorithm [17], [28]. In the present study, we proposed a new feature space configuration to solve the DL-based algorithm constraint, which is a high input data is required for training the DL algorithm while a limited input data is available, specifically for single trial applications such as control of a mobile vehicle, bionic hand and ankle foot orthosis.

In the previous studies [1], [13], the feature extraction step counted as a critical step in the identification methods. The next challenge involves the algorithms for choosing the informative features from a feature pool which contain the main information related to the aim patterns, such as an ERD/ERS wave. In the traditional methods, it was necessary to compute many features in one step, and then apply feature selection algorithms in the next step. In practice, sometimes informative features are removed in the feature selection part and then insignificant results were obtained. Additionally, accuracy reduction occurs in comparison with the same algorithm without feature selection step. The other challenging part in the EEG signal processing was to limit the number of sensors in the processing and training of the classifiers. Sensor limitation is applied regarding the area of the brain neural activation depending on the applications. For example, in our experiment we worked on the sensory motor cortex area of the brain [8], [26], [27]. Therefore, in the channel selection part the most effective sensors are selected and the others ignored in the processing, due to increasing the process and time consuming procedures in the real-time systems.

In the next development step, the constraint of the abovementioned manual feature extraction is facilitated by developing a DL algorithm. The DL method computes features automatically in a recurrent training approach to find the optimum features, which means the obtained features have the potential of reflecting the behavior of the patterns. DL is an automatic effective solution for the feature extraction and selection steps. Therefore, the DL usage growing widely for the identification of EEG patterns in different applications such as epilepsy [49], drowsiness [50] and imaginary movements [51]. The DL constraints includes a large number of input values are required to achieve significant results and the training phase is a time-consuming procedure. The advantages of the DL approach enable us to utilize all

| EEG Signal | Accuracy % of the methods | DAE-LSTM | DAE | CNN-LSTM | SSVM-GRBF [28] |
|------------|--------------------------|---------|-----|----------|----------------|
| Case 1     | 0.7476                   | 0.6904  | 0.7375 | 0.7966   |
| Case 2     | 0.7333                   | 0.6791  | 0.6944 | 0.7433   |
| Case 3     | 0.7125                   | 0.6803  | 0.7014 | 0.6833   |
| Case 4     | 0.7216                   | 0.6122  | 0.6597 | 0.7033   |
| Case 5     | 0.7235                   | 0.6775  | 0.6736 | 0.7300   |
| Case 6     | 0.7464                   | 0.6424  | 0.7361 | 0.7200   |
| Case 7     | 0.7866                   | 0.6806  | 0.7529 | 0.8833   |
| Case 8     | 0.7933                   | 0.5993  | 0.6667 | 0.8800   |
| Case 9     | 0.7211                   | 0.6748  | 0.6981 | 0.7833   |
| Case 10    | 0.7186                   | 0.7101  | 0.7014 | 0.6658   |
| Case 11    | 0.7142                   | 0.6619  | 0.6736 | 0.6533   |
| Average    | 0.7380 ±0.03             | 0.6644 ±0.55 | 0.6996 ±0.04 | 0.7554 ±0.16 |
the recorded signals for the identification processing, which means the computed features includes all the details of the evoked potential patterns. In our current experiment, we take advantages of the ERD/ERS patterns in 32 EEG channels. Therefore, the number of features increased highly, and in case of extra number of features, the DL algorithm enable us to control the number of features by using defining different sizes of hidden layers in the DAE structure.

In the present study, we implement three DL-based algorithms to identify the imaginary movement patterns. Our proposed method is explained in section III that includes two steps for computing features, namely DAE feature extraction and feature boosting. The main feature extraction part is the DAE algorithm as shown in Figure 1. In the DAE feature extraction procedure, each hand-opening or fist-making signal is down sampled into a matrix size of 512 × 78 as shown in Figure 2. Therefore, each subject has 150 matrices with an optimum size of 512 × 78, which means the matrix has the potential to reconstruct the original ERD/ERS. The optimum size obtained experimentally is shown as the middle layer of the DAE gold in Figure 2. The second step for the feature computation procedure, named feature boosting, is based on the DAE extracted features. In the feature boosting algorithm, one LSTM block is designed for each vector of 512 × 78 to provide new features (boosted features), therefore, a series of 78 LSTM blocks is set for each DAE feature matrix. The reason for selecting the LSTM as a feature boosting was that the LSTM algorithm keep the time chronological order in the EEG signal for extracting temporal properties. In addition, the LSTM enable handling the gradient vanishing problem due to using the RNN structure [48].

Here, we explain why we focus on feature boosting algorithms. Studies [21], [27], [28] and reported challenges [52] showed that the boosting feature algorithms produce more accurate results. Also, regarding our previous studies [27], [28], we computed common spatial patterns (CSPs) for the EEG signals to detect the ERD/ERS patterns. In the algorithm, a DSVQ training algorithm was employed to boost features by weighting them in a training procedure [53] and the results improved significantly. From our point of view, feature boosting is a critical point in a single trial and the results improved significantly. From our point of view, feature boosting is a critical point in a single trial classification. To investigate the efficiency, the DAE features were boosted using the LSTM algorithm to maintain the EEG time-sequence in the features and then fed into the classifier. Insignificant improvements in the accuracy show that the boosted features do not include enough information for classification, in fact again some significant information is missed in the LSTM procedure.

Regarding the achievements, we employed the LSTM algorithm for boosting the EEG features. To find the effects of a boosting algorithm in our study, the DAE with the LSTM (DAE-LSTM) boosting algorithm were implemented and results were compared with the DAE and CNN-LSTM algorithms (Table 1). The other key point of our study was the combination of the DAE features and time-based boosted features in one feature space. The reason for combining the features is the boosting algorithm misses some informative information during the process by considering the accuracy of the DAE results in Table 1. After computational considerations, we obtained significant results by combining the features before and after boosting in a feature space in comparison with the other implemented methods. The loss of information in the classification could be compensated by combining the DAE (original features) and the boosted features with respect to the time sequence of EEG signal in the classification. It is noticeable that including the time-dependent features have informative effects on constructing a feature space.

The last employed classification was the SSVM-GRBF combination, which is also the best-optimized method in our previous studies. The SSVM-GRBF achieved significant results in comparison with the other modifications of the SVM-based classifiers [17], [28]. By comparing the SSVM-GRBF result (average = 0.7554% ± 0.16) in our study with the implemented DL-based results in Table 1, it is evident that the average accuracy of the SSVM-GRBF is higher than the DAE-LSTM (0.7380% ± 0.03), but with small difference. On the other side, the precision of the SSVM-GRBF method is significantly lower (higher variation in accuracy) than the DAE-LSTM method. Regarding the achievements, regardless of the small difference in the SSVM-GRBF average accuracy,
because of the high level of precision the DAE-LSTM algorithm could be counted as the best average result among the other approaches, but it still has limitations because it is time-consuming in training, even though it was a single trial experiment.

In order to consider the significance of the achieved features and accuracy results, a statistical paired t-test and a repeated measures ANOVA test with a Post-hoc test using a Tukey correction analysis were employed [54]. The t-test analysis was applied on the normalized computed features to consider if the feature alterations between the imaginary hand opening and making a fist were meaningful. Then, the repeated measures ANOVA test with the post-hoc test using Tukey correction was then employed to consider the significance of the accuracy and precision results. The paired t-test results revealed that the \( p \) values for the extracted features of the nine subjects for the CNN-LSTM and DAE-LSTM algorithms were significant (\( p < 0.05 \)). In details, The DAE approach produced features with insignificant alterations for one subject (paired t-test \( p > 0.05 \)) and the alterations of features for two subjects (CSP-based features) in the SSVM-GRBF approach found insignificant, paired t-test \( p > 0.05 \). In the next step, the ANOVA test was computed for the proven significant features. This revealed that the significant features (ANOVA, \( p < 0.05 \)) led to significant accuracies and precisions (ANOVA, \( p < 0.05 \)).

Noticeable advantages of the DL in comparison with the SVM-based classifiers are presented as follows: 1) DL has potential for multi-class identification and may achieve more accurate results in comparison with SVM when it modified as a multi-classifier; 2) the DL has higher potential for handling large-scale data classification than the SVM-based algorithms. The weakness of the DL are 1) time-consuming computations in the training phase, and 2) the DL has potential of trapping in local minima that causes low performance if enough data is not fed in. On the other hand, the SVM-based algorithms are not time-consuming algorithms in the training phase and achieve more accurate results with a low number of features in comparison with the DL-based algorithms. We proposed our method to solve this limitation in the DL algorithm with a different configuration in the feature space. The results show that the obtained DAE-LSTM accuracy is close to the SSVM-GRBF accuracy with better precision, which means the DAE-LSTM has the potential of being used in single trial applications if the feature space is well-organized. The current study had limitations that included: 1) a low number of subjects for better training of the DL algorithm; 2) more classes should be added for a complete navigation of the mobile vehicle; and 3) designing a distinctive task to generate solid and stronger ERD/ERS patterns.

In order to process the data we employed a system with the following specifications: tesla K80, ram 12 GB, disk 320 G. From the time consuming computational point of view, the approximate time for the computations was as follows: DAE-LSTM took 22.5 min; the DAE took 21 min; CNN-LSTM took 20.5 min and SSVM-GRBF took 4.5 min. From speed point of view, the SVM was impressively faster training phase speed than the DL method. The DL had higher level of computations but the next generation of hardware technology might help for the use of DL in training for new incoming data during real-time processing.

**VI. CONCLUSION**

The presented study considered the identification of imaginary ERD/ERS patterns using DL-based methods for control of a mobile vehicle in a single trial experiment. Three different methods based on the DL concept and one method based on the optimized SVM classifier were implemented and the achieved accuracies and precisions were considered. It was found that the best method, named DL-LSTM included a feature space with the combination of the main DAE features and boosted DAE features using the LSTM algorithm. The algorithm achieved higher accuracy results with higher precision of 73.31\% ± 0.3 with \( p < 0.05 \). It is concluded that the boosting feature algorithms has a significant impact on the identifying results, but it may remove some important weights from the feature space. Combination of the features before boosting with the features after boosting generate a complementary feature space that has higher accuracy with significant precision. The organized feature space in the DAE-LSTM method enabled the use of DL for the single trial applications.

**REFERENCES**

[1] A. Hekmatmanesh, V. Zhidchenko, K. Kauranen, K. Siitonen, H. Handroos, S. Soutukorva, and A. Kilpeläinen, “Biosignals in human factors research for heavy equipment operators: A review of available methods and their feasibility in laboratory and ambulatory studies,” IEEE Access, vol. 9, pp. 97466–97482, 2021.

[2] R. C. Moioli, P. H. J. Nardelli, M. Taynnan Barros, W. Saad, A. Hekmatmanesh, P. Gória, A. S. de Sena, M. Dzaferagic, H. Siljak, W. van Leekwijk, D. Carrillo, and S. Latré, “Neurosciences and 6G: Lessons from and needs of communicative brains,” 2020, arXiv:2004.01834.

[3] H.-D. Park, C. Barnoud, H. Trang, O. A. Kannape, K. Schaller, and O. Blanke, “Breathing is coupled with voluntary action and the cortical readiness potential,” Nature Commun., vol. 11, no. 1, pp. 1–8, Dec. 2020.

[4] A. Hekmatmanesh, S. M. R. Noori, and M. Mikalii, “Sleep spindle detection using modified extreme learning machine generalized radial basis function method,” in Proc. 22nd Iranian Conf. Electr. Eng. (ICEE), May 2014, pp. 1898–1902.

[5] S. M. R. Noori, A. Hekmatmanesh, M. Mikaeili, and K. Sadeghniat-Haghighi, “K-complex identification in sleep EEG using MELM-GRBF classifier,” in Proc. 21st Iranian Conf. Biomed. Eng. (ICBME), Nov. 2014, pp. 119–123.

[6] H. Zhang, R. Chavarriaga, Z. Khaliliardali, L. Gheorghe, I. Iturrate, and J. D. R. Millán, “EEG-based decoding of error-related brain activity in a real-world driving task.” J. Neural Eng., vol. 12, no. 6, Dec. 2015, Art. no. 066028.

[7] E. M. Sokhadze, M. F. Casanova, E. L. Casanova, E. Lamina, D. P. Kelly, and I. Khachidze, “Event-related potentials (ERP) in cognitive neuroscience research and applications.” NeuroRegulation, vol. 4, no. 1, p. 14, 2017.

[8] A. Hekmatmanesh, F. Jamaloo, H. Wu, H. Handroos, and A. Kilpeläinen, “Common spatial pattern combined with kernel linear discriminate and generalized radial basis function for motor imagery-based brain computer interface applications,” in AIP Conf. Proc., vol. 1956, Jan. 2018, Art. no. 020003.

[9] X.-A. Fan, L. Bi, T. Teng, H. Ding, and Y. Liu, “A brain–computer interface-based vehicle destination selection system using P300 and SSVEP signals,” IEEE Trans. Intell. Transp. Syst., vol. 16, no. 1, pp. 274–283, Feb. 2015.
A. Myronenko, “3D MRI brain tumor segmentation using autoencoder regularization,” in Proc. Int. MICCAI Brainlesion Workshop. Springer, 2018, pp. 311–320.

M. Pregenzer, “Distinction sensitive learning vector quantization (DSLVQ)—application as a classifier based feature selection method for a brain computer interface,” in Proc. 4th Int. Conf. Artif. Neural Netw., 1995, pp. 433–436.

L. Surhone, M. Timpledon, and S. Marseken, “Wilcoxon signed-rank test,” Tech. Rep., 2010.

AMIN HEKMATMANESH received the bachelor’s degree in electrical engineering from the Science and Research of Fars University, Shiraz, Iran, 2010, the master’s degree in biomedical engineering from Shahed University, Tehran, Iran, 2013, and the Ph.D. degree in brain-controlled ankle foot and hand orthosis and mobile vehicle robots using the EEG with the Laboratory of Intelligent Machines, Lappeenranta University of Technology, from 2016 to 2019. His master’s thesis was about analyzing sleep EEG signal processing, learning and negative emotional memory. Since 2020, he has been holding a postdoctoral position on heavy machine operator’s health monitoring and signal processing for horse simulators with the Laboratory of Intelligent Machines, Lappeenranta University of Technology.

HAMED MOHAMMADI AZNI received the bachelor’s and master’s degrees in 2009 and 2013, respectively. He is currently pursuing the Ph.D. degree with the Department of Computer Engineering, University of Zanjan. His research interests include pattern recognition, times series data, and machine learning, and he is particularly interested in deep learning for medical’s image and signal analysis.

HUAPENG WU was born in China, in 1964. He received the D.Sc. (Tech.) degree from the Lappeenranta University of Technology, Finland, in 2001. He has been an Associate Professor (since October 2004) and a Professor (January 2008–June 2011) with the Lappeenranta University of Technology, Finland. His research interests include robotics, AI control, mechatronics, mechanical manufacturing, and automation, and publications include four books and more than 100 publications in his research areas.

MOHSEN AFSHARCHI received the M.Sc. degree in computer science from the Iran University of Science and Technology, in 1996, and the Ph.D. degree in artificial intelligence from the University of Calgary, Canada, in 2006. Since 2006, he has been with the Department of Computer Engineering, University of Zanjan, Iran, where he leads the Multi-Agent Systems Lab (MASLab). He is currently an Associate Professor with the Department of Computer Engineering, University of Zanjan. His research interests include multi-agent learning, probabilistic reasoning, distributed constraint optimization, and software testing.

MING LI received the bachelor’s and master’s degrees from the Hubei University of Technology, in 2006 and 2008, respectively, and the D.Sc. degree from the Lappeenranta University of Technology (LUT), Finland, in 2014. He has been holding a postdoctoral position at LUT, since 2015, in the field of robotics and artificial intelligence. Since 2015, he has also been a Research Engineer through the EUROfusion Engineering Grants under implementation of the Fusion Roadmap in Horizon 2020. Till date, he has published over 20 peer-reviewed journal and conference articles and four book sections and holds two patents. His research interests include robotics, artificial intelligence, control system design, embedded systems, and software development. He keeps an excellent academic record as an early career researcher, and started to publish research articles since his early bachelor study period.

HEIKKI HANDROOS (Member, IEEE) received the M.Sc. (Eng.) and D.Sc. (Tech.) degrees from the Tampere University of Technology, in 1985 and 1991, respectively. He has been a Professor of machine automation with the Lappeenranta University of Technology, since 1992. He has also been a Visiting Professor with the University of Minnesota, Peter the Great St. Petersburg Polytechnic University, and the National Defense Academy, Japan. His research interests include modeling, design, and control of mechatronic transmissions to robotics and virtual engineering. He has published over 250 international scientific papers and supervised around 30 D.Sc. (Tech.) theses. He has been an Associate Editor of Journal of Dynamic Systems, Measurement, and Control (ASME), since 2014. He has led several important domestic and international research projects.