Information Processing and Management 60 (2023) 103262

Contents lists available at ScienceDirect

Information Processing and Management

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

SmartRolling: A human–machine interface for wheelchair control using EEG and smart sensing techniques

Landu Jiang [a], Cheng Luo [a], Zexiong Liao [a], Xuan Li [b], Qiuxia Chen [c], Yuan Jin [d], Kezhong Lu [a], Dian Zhang [a,*]

[a] Shenzhen University, Shenzhen, China
[b] McGill University, Montreal, Canada
[c] Shenzhen Polytechnic, Shenzhen, China
[d] Popular Health, Shenzhen, China

ARTICLE INFO

Keywords:
EEG signal
Deep CNN
Wheelchair control
Smart healthcare

ABSTRACT

Smart wheelchairs based on brain–computer interface (BCI) have been widely utilized recently to address certain mobility problems for people with disability. In this paper, we present SmartRolling, an intuitive human–machine interaction approach for the direct control of robotic wheelchair that jointly leverages EEG signals and motion sensing techniques. Specifically, SmartRolling offers two wheelchair-actuation modes for users with different physical conditions: (1) head motion only — people who are severely disabled but able to do basic tasks using eyes and head, and (2) head and hands motion — in addition to type 1, people who can use functioning hands/arms for extra tasks. The system issues operation commands by recognizing different EEG patterns elicited by motor execution (ME) tasks including eye blink, jaw clench, and fist open/close, while at the same time estimates users' steering intentions based on their facing direction by leveraging inertial measurements and computer vision techniques. The experiment results demonstrate that the proposed system is robust and effective to meets the individual's needs and has great potential to promote better health.

1. Introduction

Over the past decades, there has been a rapid growth in the usage of assistive devices/robots such as electric wheelchairs that promote mobility to support people who have difficulties in walking and thus enhance the quality of human life. According to the study by the World Health Organization (WHO) (World Health Organization, 0000), there are over 1 billion people worldwide that have the problem in body function or structure, while at least 100 million people require a wheelchair. Though robotic devices could benefit older adults and people with disabilities, most of conventional control/input solutions like keyboard, mouse, joystick or touchscreen are designed for the healthy-user scenarios, they are not always easy to use when an individual is severely disabled (e.g., a person with spinal cord injuries resulting in quadriplegia or paraplegia, muscular dystrophy, multiple sclerosis, or strokes.). Therefore, an effective interface that enables users with physical limitations in executing a task or action to convey their intentions or operations to assistive devices is appropriate.

On the contrary, low-level motion control systems translate EEG signals into motion commands that enable users to control the wheelchair directly (Al-Qaysi, Zaidan, Zaidan, & Suzani, 2018). For example, Tanaka, Matsunaga, and Wang (2005) presented a brain controlled electric wheelchair which is the first work (as claimed) that uses only EEG signals. The user could decide the direction of

* Corresponding author.
E-mail address: zhangd@szu.edu.cn (D. Zhang).

https://doi.org/10.1016/j.ipm.2022.103262
Received 15 June 2022; Received in revised form 19 December 2022; Accepted 26 December 2022
Available online 17 January 2023
0306-4573/

the next move (i.e., to the left or right region) based on MI tasks — imagining left or right limb movements. Moreover, Choi (2012) also proposed a similar EEG approach but can perform 3-way movements including turning left and right, as well as going forward (by imagining foot motion). The main disadvantage of using MI signals is that the brain rhythms for each motion (imagination) may be variable between different subjects even after extensive training, thus it is not practical in real world scenarios. In order to achieve a higher accuracy, SSVEP and P300 BCIs are utilized to issue low-level motion commands such as turning left, turning right, going forward, and going backward (Mandel, Lüth, Laue, Röfer, Gräser, & Krieg-Brückner, 2009; Pires, Castelo-Branco, & Nunes, 2008). The SSVEP response to flashing light (associate with commands) can be detected quite robustly, and the visual stimulus for SSVEP can be implemented as an LED panel of a number of diodes oscillating at different frequencies (Prueckl & Guger, 2010). While P300 BCIs display direction options on the screen (i.e., visual stimulus), users can fixate attention to the target that they want to follow. It usually takes longer time than MI and SSVEP to issue commands due to processing latency, but have a higher accuracy and a more natural graphical user interface (e.g., arrow icons on the screen). However, both two techniques require a high visual demand that the user has to be continuously focusing on the mission or repeating the movement. The people may be tired or suffer from sore eyes after a trip, and long term visual stimulus could also cause safety issues such as distraction and fatigue (Rebsamen et al., 2010). Though hybrid BCIs have been provided to increase the operation efficiency that users can switch the wheelchair mode between high-level automatic navigation and low-level motion control (Mandel et al., 2009; Puanhvuan & Wongsawat, 2012), they have higher computational complexity compare with the single scheme method. In addition, the hybrid scheme does not add any new command for steering motion, it still takes a large number of steps to reach a destination with basic directions (e.g., front, back, right, and left) that need to be further investigated (Rebsamen et al., 2010).

Moving along this direction, in this paper we present SmartRolling, an intuitive human–machine interaction approach for the direct control of robotic wheelchair that jointly leverages EEG signals and human motion sensing techniques. The key insight of our work is that estimating users' steering intentions based on their facing directions allowing for a more natural interaction way that could help them to continuously control the wheelchair. Specifically, SmartRolling offers two wheelchair-actuation modes for users with different physical conditions : (1) head motion only — people who are severely disabled but able to do basic tasks using eyes and head, and (2) head and hands motion — in addition to condition 1, people who can use functioning hands/arms for extra tasks. To achieve the goal, we use the off-the-shelf BCI headset to collect EEG data and adopts machine learning algorithms (i.e., Inception CNN and SVM) to recognize signal patterns elicited by different motor execution (ME) tasks[1] including eye blink, jaw clench, head turn, and fist open/close, while at the same time a wearable device with the front-facing camera and IMU sensors (e.g., smart glasses) is employed to track the user's head movements by using computer vision techniques. SmartRolling aims to provide an effective and practical paradigm that adds a new degree of freedom for the user to control the wheelchair without the exhausting mental processes, and thus can assist the people with disabilities in terms of both convenience and safety.

While the idea of combining EEG signals and multi-sensory data for robotic wheelchair control seems simple, the practical implementation of SmartRolling entails substantial challenges. Firstly, it is important to select correct brain regions corresponding to different human motion intents (wheelchair moving commands). Reducing the number of electrodes could also help improve the classification efficiency and offer less expensive solution for system design. Secondly, users may have various motion styles (e.g., strength and speed), which could imprint different hallmarks in the brain activities and head movements. Finally, EEG signals can be heavily contaminated noise caused by user arbitrary movements and environment conditions (Barua, Ahmed, Ahlstrom, Begum, & Funk, 2017). Therefore, we first adapt wavelet transform to explore time–frequency features of EEG signals and then propose a modified version of squeeze-and-excitation (SE) block (Hu, Shen, & Sun, 2018) to select most representative eeg electrodes/channels. The channels with lower weights will be pruned to make the classification more efficient. Finally, an optical flow-based method is used to adjust user moving directions that filters out unstable data on real road.

The main contributions of this paper are summarized as follows:

1. We propose a hybrid sensing boosted human–machine interaction approach for assistive robot/wheelchair control. The system issues operation commands by recognizing different EEG patterns elicited by motor execution (ME) tasks, while at the same time estimates users' steering intentions based on their facing direction by leveraging inertial measurements and computer vision techniques.
2. We explore time–frequency features of EEG signals by adapting wavelet transform, and further employ SVM and deep CNN models (i.e., Inception V3) for operation event identification. A modified squeeze-and-excitation (SE) block is also deployed to perform channel (electrode) selection, which extracting the most effective electrodes as input to improve system computational efficiency.
3. We use optical flow based methods to adjust human head states and align two coordinate systems — headset and wheelchair, which could help users to initialize the moving direction intuitively after turns or system reset.
4. We conduct well-designed real-world moving experiments, involving 9 participants (3 females and 6 males, age from 16–53) in indoor and outdoor environments. SmartRolling achieved an average accuracy over 99% for ME tasks recognition and 98% for vision based direction estimation. The experiment results demonstrate that the proposed system is robust and effective, which has great potential to promote better healthcare services and quality of life.

[1] Motor execution (ME) and motor imagery (MI) correspond to event-related desynchronization/synchronization (ERD/ERS) EEG respectively, which could cause oscillatory activities observed in different regions in the sensorimotor cortex of the brain (Pfurtscheller & Da Silva, 1999), more details will be presented in Section 2.1.

To the best of our knowledge, instead of using EEG signals or pure vision based solutions to issue steering commands, SmartRolling is the first approach that combines multi-sensory data (i.e., EEG, Camera, IMU) following the human natural to control the wheelchair movements intuitively. Thus, our work could provide a road map to other studies on robotic wheelchair design to assist people with body disabilities.

2. Design consideration

2.1. Background and motivation

Recently, research and development of brain-controlled assistive robots such as intelligent wheelchairs have received a great attention in both academia and industry since they are able to bring mobility to people with disabilities, and thus improve the quality of life for these users. In order to deliver the subject to the destination successfully, a brain-controlled wheelchair should fulfill two basic conditions (Rebsamen et al., 2010): (1) The control must be ergonomic and safe: the wheelchair should provide an efficient and intuitive BCI that allows certain freedom to users. If a BCI requires long time concentration or considerable mental efforts, it may cause fatigue and distraction during the trip, which will exhaust the user and result in safety issues on road (e.g., collisions). (2) The deployment must be easy. The system equipped with real-time laser scanner or requires predefined locations may limit the usage of wheelchairs. Developers should minimized the human efforts and computation cost of the system to perform necessary functions that makes the wheelchair adequate for individuals across the globe (Gomez Torres, Parmar, Aggarwal, Mansur, & Guthrie, 2019).

Among various techniques, EEG BCI has become one of the most popular non-invasive solutions for robotic wheelchair control. There are mainly three types of EEG signals that have been widely used in brain-controlled robot including P300, steady state visually evoked potential (SSVEP) and event related desynchronization/event related synchronization (ERD/ERS). P300 is a positive potential deflection that responses to desired external stimulus such as visual, auditory, or tactile stimulus, and it is primarily measured in the parietal lobe at a latency of approximately 300 ms (Iturrate, Antelis, Kubler, & Minguez, 2009). The activity of brain signals based on the P300 can be translated into low-level steering or high-level navigation commands with a high accuracy. On the other hand, SSVEP is one of the critical brain signals that responses to a visual stimulus modulated at a frequency range from 6 Hz to 75 Hz, which can be measured in the occipital area (Choi, 2012). It has been developed as the wheelchair control and BCI-based speller scheme by taking advantage of fast adaptation, high accuracy and minimal training effort from the subject. However, both of these two methods require concentration on the target or produce a visual stimulus for a long period of time, which will be tiring for their users. In the case of ERD and ERS, ERD refers to the oscillations of the mu and beta rhythms of the sensorimotor cortex that changes when a movement is being prepared or during a movement — motor imagery (MI), while ERS refers to the increase of oscillations after a movement occurs — motor execution (ME) (Aljalal, Ibrahim, Djemal, & Ko, 2020). Since it may be difficult for users to imagine different movements which always need a large amount of time for training, the performance of MI based systems are quite unstable. The ME paradigm is able to train subjects with minimum time and effort — it is easier to perform an actual motion than imagination. The signals of ME tasks could be detected by the system accurately, however, the main drawback of ME is that not all subjects could do actual motion due to different disability conditions.

2.2. Steering scenarios and execution protocols

In this paper we leverage ME EEG to built an intuitive human–machine interaction that enable inexperienced users to control the wheelchair effectively. In order to support the people who have difficulties to perform certain ME tasks (e.g. moving hands or arms), we estimate user steering intentions by using IMU sensors and the front-facing camera on wearable devices (i.e., smart glasses) instead of EEG signals. The basic strategy of our method is turning the wheelchair into the direction where the user is focusing on. Since a person usually looks into the direction to go when walking or driving, this is a natural way to indicate the desired direction (Matsumotot, Ino, & Ogsawara, 2001).

More specifically, we provide two types of execution protocols for users with different physical conditions: (1) head motion only and (2) head and hands motion.

2.2.1. Head motion only

The mode is designed for people who have lost most voluntary muscle control, but still be able to use their eyes and head. There are two basic ME tasks — eye blinks and jaw clench. Initially, the user sits on the still wheelchair wearing the EEG headset and smart glasses accordingly, and the system state is waiting command (i.e., not performing any action). In order to start the control interface, the user has to face forward (the same direction as the camera facing) and perform 3 times of eye-blinking within 3 s and no less than 0.5 s between two blinks. If the user wants the wheelchair to move forward, he/she has to do the same task again. This mechanism will help to avoid the case that the wheelchair accidentally starts to hit any obstacle. If the user wants to make a turn, the wheelchair has to be stopped first. A jaw clench motion will stop the moving wheelchair. Then the user can turn his/her head to the desired direction (i.e., left or right) smoothly and continuously, the system will calculate the angle of the rotation based on IMU sensing data. The user has to keep facing the desired direction and perform 3 times of eye-blinking again to confirm the turning command. After the turn is finished, the user could choose to move forward or turn again using the same methods mentioned above. In addition, the camera is used for coordinates alignment which will be discussed in Section 2.3.

2.2.2. Head and hands motion

In addition to the Head Motion Only execution protocol, we also provide another way for the users who are capable of moving their arms/hands. We examine 3 ME tasks including fist open/close, arbitrary arm/hand motion and turn back movement. The user will have a similar operation procedure as execution protocol 1 but can control the wheelchair more conveniently. More specifically, user can clench the fist (either right/left or both) to issue stop command when the wheelchair is moving forward, and then use eye-blinking command to restart the wheelchair. As such behaviors reveal the communication between hand muscles and the nervous system of human (Wohlfert, Blader, & Wilson, 2017), as clenching the fist will be a natural response when we need a stop action on road. The system should also be able to detect meaningless gestures like arbitrary arm/hand motion or turn back movements, filtering out such gestures could improve the user safety. Two protocols share the same control scheme/command for start and stop to avoid user confusion. Additionally, the system is able to turn the exact angle by using IMU sensor data (details in Section 4.4).

In order to implement above execution protocols, we address the issue by detecting and recognizing ME tasks including eye blinks, jaw clench, fist open/close, arbitrary arm/hand motion and extreme motion (turn back). A set of non-related motion like human arbitrary movements should also be considered as interference. If the proposed models could successfully identify the target EEG signals, it would demonstrate that our design is efficient for brain-controlled wheelchairs.

2.3. Subject motion monitoring

In this paper, We use a commodity wireless Electroencephalogram (EEG) headset to monitor human brain activities with 32 EEG channels (electrodes) which adheres to the 10/20 system standard. As each electrode is designed to access brainwave data at a specific brain location that serves different brain functions. Thus incorporating all available channels could be redundant that adds extra interference and significant computational overhead to the system especially for time–frequency analysis. In addition, a vision based real time video analysis is desired to adjust user head pose and facing direction for coordinates alignment. At the initial step, the user is facing forward and the headset/glasses (headset and glasses share the same coordinate system since they are both worn by the user) coordinate system is along with the wheelchair, the moving direction is the defined as the positive direction. When there is a turn, the headset rotates with the user head for direction estimation. But the angle will not be calculated by the wheelchair, thus after the wheelchair rotation the positive direction for headset/glasses is still the old moving direction (before the turn). If we do not reset the coordinates, the system cannot estimate the accurate head rotation direction next time. Thus a vision based solution is deployed using optical flow based method to adjust if the user is facing forward during moving. Based on the output, we can reset the positive direction and align two coordinate systems (headset/glasses and wheelchair). For the IMU based angle estimation, we are able to get the value using the API of smart glasses, thus is not the major contribution we considered in this paper.

3. System design

In this section, we present a detailed discussion of the SmartRolling architecture and algorithms.

3.1. System overview

SmartRolling monitors the human brain activities by processing EEG signals captured from the wireless electrode headset, while at the same time detecting user turning intention via IMU sensors and front-facing cameras. The key components of SmartRolling are illustrated in Fig. 1 including (1) Data collection and preprocessing; (2) Time–frequency analysis, (3) Direction Steering and (4) Command Recognition.

The Data Collection and Preprocessing module is responsible for sampling the raw EEG signals in a fixed time interval (e.g., 3 s) when a user issues a command, and use denoising method to the time series data for time–frequency analysis. Two steps are included in this module: (1) obtaining a fixed rate of EEG signals format ready for the Wavelet transform and (2) filtering out burst noise and impulse. While at the same, the IMU and front-facing camera record the data and frames for rotation estimation and facing direction adjustment. In the Feature Extraction module, we first apply Wavelet transform for time–frequency analysis and obtain a informative spectrogram. We separate the eye blinks and other movements based on the observation in time–frequency domain. Finally, in the classification module, for eye blink detection, a CNN-SVM model is applied since we do not need a complex deep model for a binary classification. For other motion, a modified Squeeze-and-Excitation Block is added in the input layer of CNN model for EEG channel (electrode) selection, our purpose is to reduce the computational cost and achieve same level accuracy with less input data channels. We use a lightweight version of Inception V3 model for EEG signal pattern classification, the SmartRolling system is able to issue proper wheelchair-control commands based on the classification results and the rotation estimation outputs.

3.2. Data collection and preprocessing

3.2.1. Data formatting

In order to improve the computational efficiency of command recognition, we need to extract the most representative channels from 32 EEG electrodes. In this module, 16 electrodes (AF3, AF4, F7, F3, FZ, F4, F8, FC1, FC2, C3, CZ, C4, P7, PZ, P8 and O2) are selected based on current EEG studies. For the data recording, a sliding window of 3.0 s with 1 s stride at sampling rate 500 Hz is

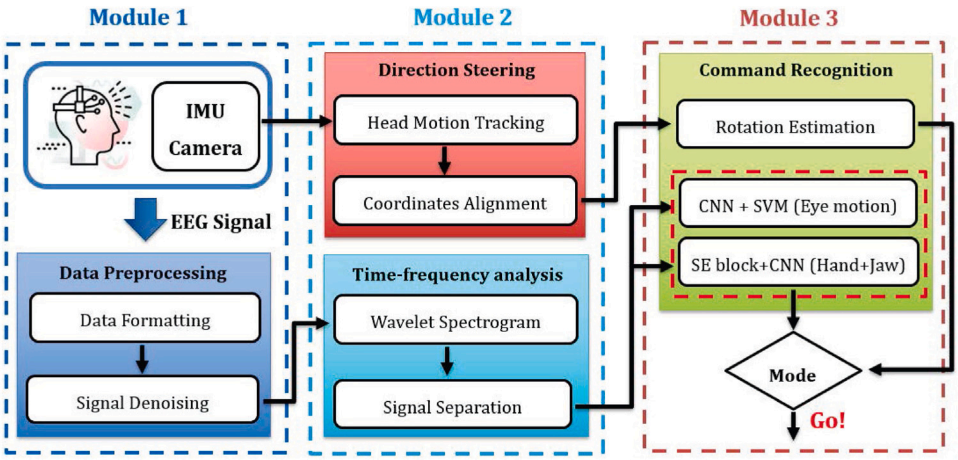

Fig. 1. System overview.

employed on input signals. We first performed 0.5–100 Hz band-pass filtering to remove basic system noise, and then down-sampled the data from 500 Hz to 250 Hz and re-referenced them to the averaged earlobes.

For IMU and Video data collection, a commodity mixed reality device (e.g., Microsoft HoloLens device) will be worn by participants sitting on the wheelchair. The system will keep reading the sensor data and video frames with a sampling rate at 24 HZ.

3.2.2. Signal denoising

To handle the burst noise and impulse during the recording, we leverage Hampel identifier to detect and eliminate the outliers of the input signals. For a recorded sample S, Hampel identifier algorithm check whether the EEG value of S is outside the region of Y - the value distance from the estimated median distribution of neighboring samples. For example, S has l neighbors on each side, Hampel identifier computes the median of a window with a length of $2l+1$. Then it measures the deviation of the samples in the window and compare with median absolute deviation (MAD) of Y, as illustrated in Eq. (1):

$$\left|Y_i - Median(Y)\right| > \lambda \cdot MAD, \tag{1}$$

If the distance between Y_i and the median value exceeds the threshold ($\lambda\cdot$ MAD, where λ is 3 by default), it will be replaced by the median .

In addition, there are also other activities happening in different wheelchair control conditions. We are able to filter out the input if it does not belong to the task domain by monitoring wheelchair dynamics via IMU sensors — accelerometer and gyroscope. However, it is not the main contribution considered in this paper, and there have been a number of studies done well on it (Jiang, Chen, & He, 2016; Jiang, Xie, Zhang, & Gu, 2022).

3.3. Time–frequency analysis

3.3.1. Wavelet transform

As EEG signal is a collection of non-stationary time series corresponding to measurements of human brain activities across different frequency bands, thus feature extraction and analysis in the time–frequency domain is necessary. For non-stationary time series data representation and processing, two window-based techniques (1) Short-Time Fourier Transform (STFT) and (2) Wavelet transform (WT) have outperformed others for human motion study in time–frequency domains. STFT methods first divide signals into fixed length subwindows/segments and then apply a sequence of Fourier transforms (e.g., FFT) separately on each window, where the data can be considered as segment-wise stationary. However, it is crucial but hard to find an optimal window size to obtain satisfactory time and frequency resolutions. On the other hand, WT methods (known as discrete or continuous WT) allow multi-resolution signal analysis at different scale components. It shows excellent results on the analysis of multi-band time series data like EEG, Wi-Fi CSI, etc. In this paper, we use cgau8 as the kernel and set scale range 1–64, thus output a 64 × 750 (250 × 3) × 16 (channels) data matrix (spectrogram) of each window for CNN model.

3.3.2. Eye movements spotting

As discussed in Section 2.3, EEG signals are distinguishable from other human movements. As shown in Fig. 2, we can see a clear pattern of eye movements (blinks) in both time domain and time–frequency spectrogram, while the components in spectrogram are much easier to be extracted from the original signal — eye blink signals are below 30 HZ and most of other activities are above 35 HZ. In addition, since the eye states information is also very important for user behavior analysis, we stored it as back up data/hints

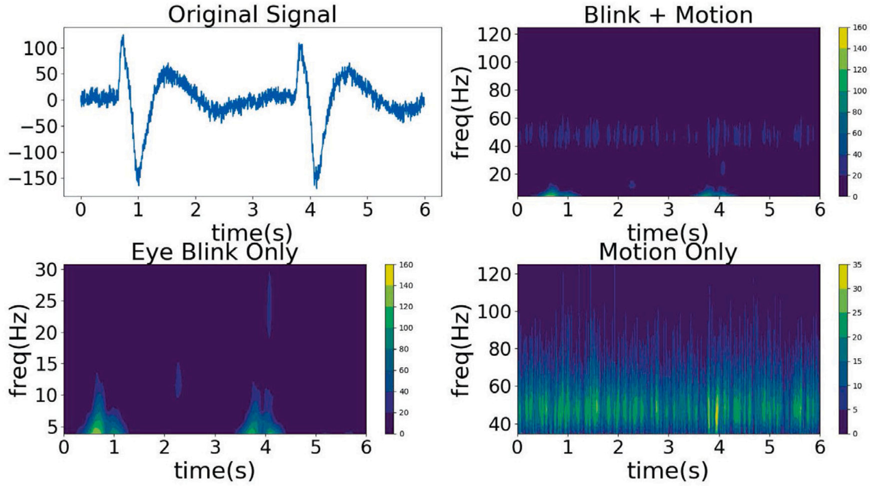

Fig. 2. Eye movements in EEG signals: (1) time domain, (2) human motion in time–frequency domain, (3) eye blinks only in time–frequency domain and (4) other motion.

for the future study. That is why we did not filter it out in the preprocessing module. After eye movements extraction, we further select 3 channels for blink detection — AF4 F7 F3, thus the data size for normal motion is 32 (frequency) × 750 (time/samples) × 16 (channels), and the data size for eye blink detection is 32 (frequency) × 750 (time/samples) × 3 (channels).

3.4. Direction steering

3.4.1. Facing direction

In our application, there are generally 4 face direction states (Guinness, 2020):

- Dolly: when the user moves forward with the camera facing forward as well.
- Truck: when the user moves forward with the camera facing left or right.
- Pan: when the user is still but rotating his/her head left or right.
- Steady: when there is no movement.

We propose a two-step strategy to address the challenge. In the first step, a dense optical flow-based approach is used to compute the translation magnitude and angle on specified grid points. Then a K-means clustering algorithm is adopted to fit the translation magnitude into 3 clusters. In the end, we extract the angle mode, translation mode, cluster ratios (cluster 1/cluster 3, cluster 1/cluster 2, and cluster 2/cluster 3), as well as the mean translation magnitude as the features to perform our classification.

3.4.2. Optical flow based method

In the first approach, we calculate the dense optical flow using (Farnebäck, 2003). To calculate the optical flow, we use two adjacent frames, and estimate the displacement field. Since each pixel's neighborhood can be approximated by a polynomial function: $f(x) = x^T Ax + b^T x + c$ where A is a symmetric matrix, b is a vector, c is a scalar and x is the displacement coordinate. We use the least square estimation to calculate the coefficients. The result is calculated from a weighted average of all coordinates from the neighborhood. We refer to the original paper for detailed implementation. Based on the optical flow, we calculate the translation magnitude and angle of the translation for each grid point.

To classify the category to which a given short period of frames belongs, we analyze:

1. In the case of Dolly, the top-left quarter of the image should have angles between 0 to 90 degrees. The top-right quarter of the image should have angles between 90 to 180 degrees. The bottom-left quarter of the image should have angles between 0 to −90 degrees. The bottom-right quarter of the image should have angles between −90 to −180 degrees. When looking at the translation, due to the points close to the vanishing point should have the smallest translation, whereas the far away points should have the largest translation. When calculating the variance of the translation, the variance should be the largest in this scenario.
2. In the case of Truck movement, all the points should have similar angles and translation magnitudes as the camera moves in parallel with the scene. The region-based variances should have the smallest values among the 3 moving scenarios.
3. In the case of Panning movement, center region points move faster than those are further. So, we should have larger translations in the center region, and smaller else where.
4. In the case of Steady (no movement), translation magnitudes are all close to 0 for all points.

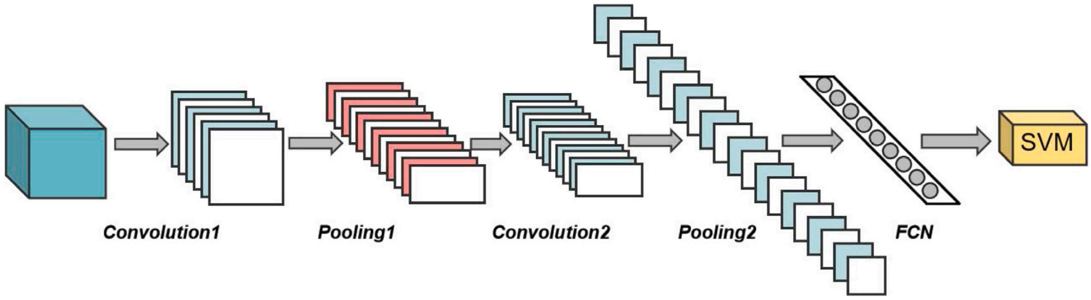

Fig. 3. CNN-SVM model.

Table 1
The outline of the proposed network architecture — CNN-SVM.

Type	Patch size/Stride	Input size
Conv	5 × 5	13 × 32 × 250
Pool	5 × 5	23 × 28 × 246
Conv	3 × 3	23 × 10 × 119
Linear	Logits	3 × 4 × 59
Linear-SVM	Features	1 × 1 × 708

3.4.3. K-means clustering for feature extraction

Before a K-means clustering algorithm, we filter out all grid points whose translation magnitude is 0. For the remaining points, we calculate the mean translation and store it as a feature $Still$. Next, we discard outlier points whose translation is in the 10% on either tail. Then, we record the modes of the angle and translation for the rest points, store them as two features $Angle$, $Translation$.

Secondly, we use K-means clustering to fit the translation magnitudes with 3 clusters. As we consider small, medium, and large translation clusters should give us good insights. We did tests using different numbers of clusters from 2 to 5, and number 3 achieved best performance. We then compute the ratio between any two cluster centers and use them for classification (Random forest).

3.4.4. Random forest for direction estimation

We leverage the learning model using Random Forest (Breiman, 2001) for face direction estimation. There are several reasons for choosing Random Forest. For one, Random Forest is a nonparametric ensemble leaning method which will be less affected by the presence of outliers. During the training phase, it builds a multitude of individual decision trees considering random subsets of the samples and then aggregates their decisions to form a final facing direction prediction. In real-world scenarios, we have a variety of data recorded with different view angle and body height. As each decision tree only has access to a certain set of training samples, the ensemble nature of Random Forest algorithm increases the diversity in the forest and leads to more robust overall predictions.

In addition, another advantage of Random Forest algorithm is that it is less prone to overfitting. Bagging in Random Forest operates in two levels—data selection and variable selection. Individually, outputs of a decision tree have low bias and high variance. Given enough trees in the forest, the classifier would form a consensus between trees that are trained using a different set of data. Again, by taking the average over all trees, each fitted to a subset of the original data set, the algorithm yields at one bagged classifier that is more stable and less prone to overfitting.

3.5. CNN-SVM model

As shown in Fig. 3, a set of regular CNN layers is added to SVM classifier for identifying eye-blinking event. The detailed structure information is shown in Table 1. Specifically, we combines a convolutional neural network (CNN) and a linear SVM. The support vector machine (SVM) was developed by Vapnik (Cortes & Vapnik, 1995) for binary classification. Its objective is to find the optimal hyperplane to separate two classes in a given database. According to the study (Tang, 2013), the test accuracies of CNN-Softmax and CNN-SVM are almost the same. But based on our observation for EEG data testing, SVM performs slightly better than CNN-Softmax. In addition, there will be a significant extra workload for data collection and training a deep model that will not fit our need - a lightweight binary classifier. Therefore, a binary SVM classifier is considered as a better solution for our goal — effectively distinguish eye blinks from other activities. But we do not use CNN-SVM model for other ME task recognition due to lower classification accuracy as illustrated in Table 5 in the Evaluation 4.

3.6. CNN model with SE block

Generally speaking, people tend to extract features from data by deepening the depth of the network (e.g., CNN model). However, with the increase of the depth/size, the neighborhood grows exponentially, and there is a "information bottleneck" phenomenon:

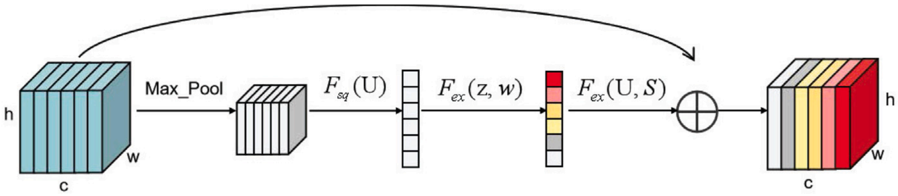

Fig. 4. SE-Inception module architecture.

a large amount of information from many fields has to be squeezed into a single node feature vector, resulting degraded model performance. It could be even worse if the input data is a set of 32-channel EEG spectrogram, thus a CNN model with SE module is proposed in this section to handle such issues.

3.6.1. Squeeze-and-Excitation block

In order to extract information from sparse spectrum, we leverage Squeeze-and-Excitation (SE) block (Hu et al., 2018) to handle the channel reduction issue. The goal of SE block is to improve the representational power of a network that adaptively recalibrates channel-wise feature. The structure of the SE block is shown in Fig. 4. When given a feature map U of input X via any given transformation F_{tr}, a corresponding SE block can be constructed. The feature map U is first passed through a so called $Squeeze$ operation, which produces a channel descriptor that aggregates global spatial information (h×w) by using global average pooling to generate channel-wise statistics. Formally, a statistic z can be generated by shrinking U through its spatial dimensions h×w, for example, the i th element of z is calculated by:

$$z_i = F_{sq}(U_i) = \frac{1}{h \times w} \sum_{n=1}^{h} \sum_{m=1}^{w} u_{n,m}, \tag{2}$$

At the second step, the aggregation is followed by an $Excitation$ operation to fully capture channel-wise dependencies. It takes the embedding as input and produces a collection of per-channel modulation weights, the detailed process is illustrated in Fig. 4:

$$S = F_{ex}(z, W) \tag{3}$$

Then, the weights are applied to the feature maps U to generate the output of the SE block which can be directly fed into the next layer of CNN.

$$\tilde{X} = F_{scale}(U, S) \tag{4}$$

SE block is computationally lightweight and imposes only a slight increase in model complexity and computational burden, which has been widely tested in different deep model approaches, as well as for EEG based studies (Li et al., 2020). However, it cannot be used directly for input EEG channel reduction since the existing SE block structure aims to adjusting the weights of feature maps e.g., the output of a convolution. While after pass a convolution layer, the tag of the EEG channel (electrode) will be erased. Tough we can remove the F_{tr} transformation and connect the input matrix to SE block, the spatial information (h × w) may be lost due to the global pooling — brain signals somehow follow the gaussian distribution and the global average value is close to Zero.

To address this issue, we apply a Max Pooling layer to extract local maximums as features and then feed them to the SE block. The weights of each EEG channel will be updated during training (the backpropagation). The detailed results are presented in Section 4.

3.6.2. Inception networks

Since many of the related conclusions are from empirical observation, we evaluate on several networks with different numbers of layers and reach to a conclusion that Inspired by Inception Networks (Szegedy, Vanhoucke, Ioffe, Shlens, & Wojna, 2016). We arrive at a modified (lightweight) version of Inception model, which is shown in Fig. 5. The most important reason that we choose Inception V3 structure as the CNN model is that it mainly focuses on saving computational power by modifying the previous Inception Networks. In addition, it has been tested that the SE block achieved state-of-the-art performance with Inception models, which is considered as our future work for EEG signal analysis.

In addition, we perform two rounds training to examine channel selection performance. The proposed SE block with Inception Model will produce a set of per-channel modulation weights generated from pre-collected 16 electrodes. We will select 8 (half of the channel number) channels with highest weights as the input for the new model sharing the same structure with the old one, and then select 4 and 2 channels accordingly. The network architectures is illustrated in Table 2. In the evaluation, we will provide a comprehensive study on channel selection using different combination of the datasets.

4. Evaluation

In this section, we present the evaluation of SmartRolling in both laboratory and outdoor with mixed reality environments. Specifically, we first describe the experiment setup and datasets, then present the results of face direction estimation and ME task recognition with different channel settings — 16, 8, 4 and 2. Finally, we also discuss how to improve the system performance and possible extensions.

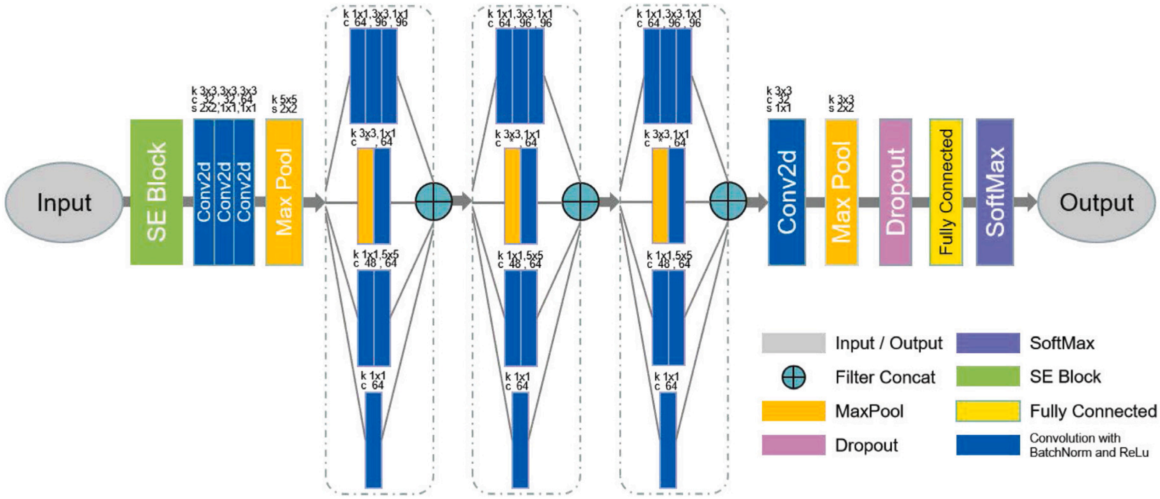

Fig. 5. Inception network structure.

Table 2
The outline of the proposed network architecture — 16 channels.

Type	Patch size/Stride	Input size
SE Block	Section 3.6	16 × 32 × 750
Inception	Inception block	112 × 32 × 750
Inception	Inception block	364 × 5 × 184
Conv	3 × 3/1	288 × 5 × 184
Linear	Logits	1 × 1 × 2880
Linear	Classifier	1 × 1 × 6

Fig. 6. Wheelchair setup.

4.1. Experimental setup

We conduct the experiments using a mixed reality environment in laboratory conditions and out-office environments. Specifically, SmartRolling utilizes a wireless EEG headset from Guger Technologies (model name: g.Nautilus RESEARCH Headset) to collect brain activity data, at the same time, the participant will wear a Microsoft HoloLens mixed reality device with front-facing camera and IMU sensors. (as shown in Fig. 6).

We manually labeled ground-truth events and the system detection results. In our experiments, we recruited 9 participants - 3 females and 6 males from age 16 to 53. Each participant was asked to perform 7 ME activities in laboratory conditions (also worn the HoloLens device) and out-office environments including eye blink, jaw clench, fist open/close and non-related movements, while at the same time they were also asked to perform head moving tasks including dolly, truck left/right, pan left/right and steady, they sat on the wheelchair and one of our staff will stand by the wheelchair for safety concerns. There are 6750 EEG samples and

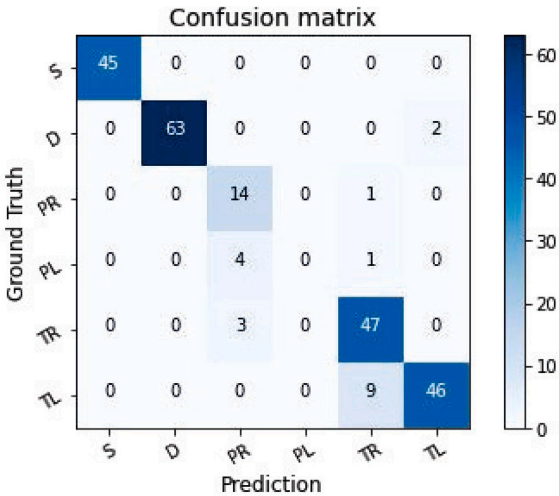

Fig. 7. Face direction estimation.

Table 3
Eye blink detection results.

Events	Blink	Non-related
Blink	85	5
Non-related	9	81

200 video clips collected in total. We implemented deep model on a desktop with a Intel i7-6850K cpu and a NVIDIA GeForce GTX 1080 Ti GPU. The raw EEG data were stored offline for further analysis.

4.2. Face direction estimation

In this section, we provide a comprehensive analysis of our vision based model by using 10-fold cross validation. As shown in Fig. 7, where S: steady, D: Dolly, PR (PL): Pan Right (Pan Left), TR (TL): Truck Right (Truck Left). Though the number of events is limited, thus the data is not well balanced. But we can see that the model is able to distinguish moving forward frames from other samples, thus can efficiently align the device coordinates and reset positive direction for the system and user. Moreover, our method can be implemented on any wearable device that equipped with a front-facing camera and IMU sensors which can be much cheaper than the mixed reality device used in this project. By using a mixed reality device, we can further add new visual and acoustic features to improve the user experience.

4.3. ME task recognition

4.3.1. Eyeblink recognition

The result for eye blinks detection is shown in Table 3, a 10-fold cross validation is applied, and we selected the worst case to show the performance of our model. Compare to other ME tasks, the duration of eye blinks is much shorter but easy to be detected. In addition, users may become tired if they repeat the motion frequently, thus we did not collect eye blink samples as many as other tasks (around 200 samples in total). The results shows that our proposed method is able to detect most of eye blinking events. The system wrongly detect 6 non-related events as eye blinks, it is because that we ask participants to perform arbitrary movements that some of them did eye blinks naturally without notice. Moreover, in our design we have very strict scheme to issue a command that the user must do 3 times of eye blinks in a limited time and facing forward. Thus the model is able filter non-related motion out and will not affect the system control performance.

4.3.2. Other execution commands recognition

As shown in Figs. 8 and 9, the confusion matrices of wheelchair control event detection with different channel selection are presented. We can see that the 4-electrode model achieved the highest average accuracy of 95.4%. For 8 channels, we select ('FP1', 'F7', 'FZ', 'FC2', 'CP1', 'CP2', 'F5' and 'O1'), for 4 channels we select ('FC2', 'CP1', 'CP2' and 'F5') and for 2 channels we select ('FC2' and 'CP2'). And Event 1 is fist close, Event 2 is turning back, Event 3 is jaw clenching, Event 4 is arbitrary hand motion and Event 5 is fist open (relax).

Though 16 channels provide richful information, much more noise and diverse patterns among different participants may also exist. By checking the average classification accuracy in Table 4, we can see that 16 or 8 channels can perform well on certain users,

Table 4
Performance comparison of different channel selection.

People	Model							
	2ch		4ch		8ch		16ch	
	Accuracy	kappa	Accuracy	kappa	Accuracy	kappa	Accuracy	kappa
User 1	98.70	98.38	98.15	97.69	100.00	100.00	99.60	99.50
User 2	95.25	94.06	98.40	98.00	99.10	98.88	98.55	98.19
User 3	96.70	95.88	97.30	96.63	90.95	88.69	97.45	96.81
User 4	99.60	99.50	99.95	99.94	80.75	75.94	86.35	82.84
User 5	76.45	70.56	82.15	77.69	67.90	59.88	74.95	68.69
User 6	91.45	89.31	95.60	94.50	67.30	59.13	71.65	64.56
User 7	92.45	90.56	95.40	94.25	63.70	54.63	59.40	49.25
User 8	90.95	88.69	93.40	91.75	80.20	75.25	88.40	85.50
User 9	93.65	92.06	98.00	97.50	80.00	75.00	77.20	71.50
Average	92.80	91.00	95.37	94.22	81.10	76.38	83.73	79.65

Table 5
Comparison results of ablation study for ME task Recognition.

People	Model		
	SVM	Conv+SVM	Ours-4
User 1	80.60	95.40	**98.15**
User 2	90.40	94.80	**98.40**
User 3	72.60	86.20	**97.30**
User 4	77.20	98.40	**99.95**
User 5	50.60	78.80	**82.15**
User 6	69.00	86.20	**95.60**
User 7	61.80	96.40	95.40
User 8	50.60	87.20	**93.40**
User 9	58.60	83.80	**98.00**
Average	67.93	89.69	95.37

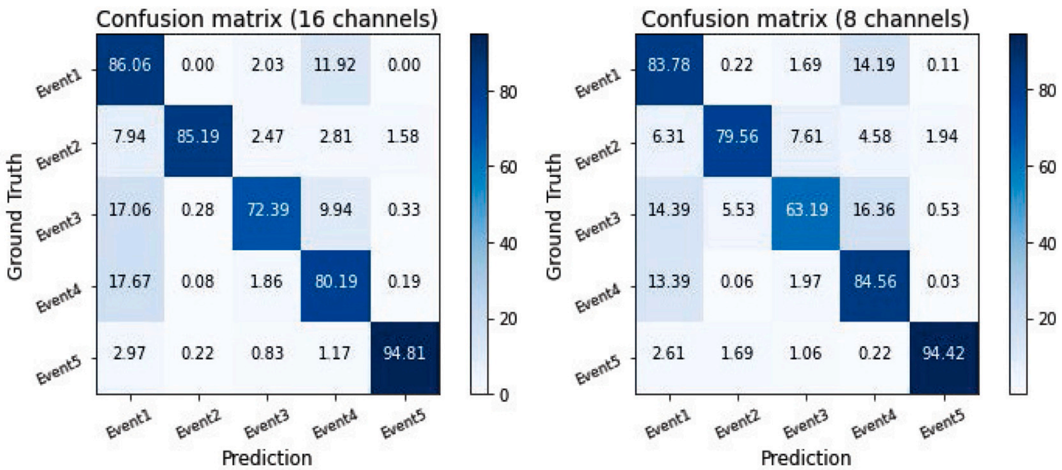

Fig. 8. Comparison of different channel selection (16 channel versus 8 channel).

however, the performance may change sharply due to domain variety. On the other hand, 4 or 2 channels performs much more stable among different user domains. 2-channel model is less accurate than 4-channel, which is mainly because that 4 electrodes contains more information that can provide higher dimension data to be learned for distinguishing different task patterns. We can see from Fig. 9, Event 1 and Event 4 are always mixed together that confused the classifier. It does make sense that fist close may have certain motion patterns being similar with the arbitrary hand motion, these two types of hand/arm gestures may also share same muscle functionalities. A mix-up method may help to address such issues. Based on the results, the model uses less electrodes achieved better accuracy, while the accuracy of 2 channels become lower but acceptable, while the computational time and training load is much smaller than 16 or 8 channels. Moreover, a larger dataset may help to further distinguish these signals.

Additionally, we also present the comparison results of ablation study as illustrated in Table 5. It shows that our proposed model (Inception-CNN) could outperform CNN-SVM model in most cases except for user 7. Actually, the classification accuracy of

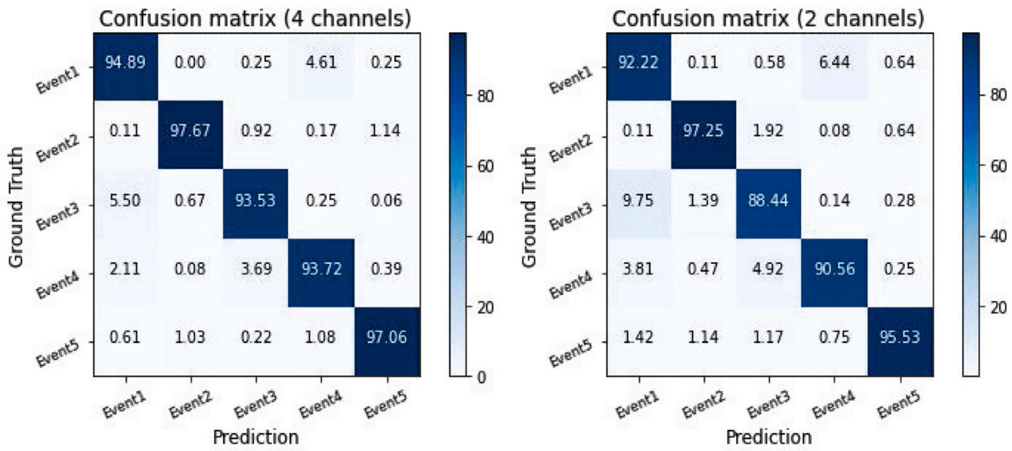

Fig. 9. Comparison of different channel selection (4 channel versus 2 channel).

each model for user 7 is very close — 96.4% and 95.4%. The comparison results also demonstrate that CNN-SVM design could achieve stable and better recognition results than other basic machine learning models (e.g., basic SVM), which is capable of binary classification for eye-blinking identification as mentioned in Section 3.5. However, CNN-SVM is not as good as deep Inception-CNN on more complex recognition scenarios — other ME task classification.

4.4. Discussion

Based on the evaluation results, our proposed system is able to provide low-level motion control service for robotic wheelchair. Actually, we did not fully used the capability of the device as well as the model. For example, as mentioned in Section 2.2, the vision based solution can be used for an accurate wheelchair rotation — making the wheelchair just turn into the direction (i.e., angle) where the user wants to go. To achieve this, the work (Matsumotot et al., 2001) using the system with stereo cameras which are less practical in real world scenarios. In our design, we can install one extra camera (a normal one) on the wheelchair facing forward (the wheelchair moving direction). When the system is on, we always compare the optical flow features of two video frames — one is from the wheelchair and the other one is from the smart glasses. If they are all facing the same direction, the feature distance should be small, otherwise, it will exceed a threshold when there is a turn or significant vibration. The implementation is considered in our future work.

5. Related work

With the development of communication technologies and rapid growth of Internet of Things (IoT), many EEG based assistive robot control systems have been proposed recently (Aljalal et al., 2020; Bi, Fan, & Liu, 2013). By taking the advantage of P300 BCI, a high-level navigation work is proposed by Iturrate et al. (2009). They use menu-based system that list a number of predefined location for user to choose, a user can relax during the transition without heavy mental process. Moreover, SSVEP have been used to explore the motor movement control (Turnip, Soetraprawata, & Tamba, 2015). However, since the system mainly relies on small square blinking stimuli which can influence exhaustion if BCI is used for a long period. Moreover, a limited number of commands exists due to harmonic frequencies (Al-Qaysi et al., 2018). Mandel et al. (2009) utilized SSVEP BCI are utilized to issue low-level motion commands including turning left, turning right, going forward, and going backward, where the stimulus is implemented as an LED-panel of four different diodes oscillating at 13, 14, 15, and 16 Hz, respectively. In addition, a proof of concept system called the Affordable Smart Wheelchair (ASW) for indoor use is designed by Gomez Torres et al. (2019). The fully autonomous functions of the wheelchair includes room-to-room and predetermined docking location navigation. The system uses a voice control interface to achieve autonomous navigation through affordable components. Above approaches well examined the capability of smart sensing platforms, however, none of them took users' thoughts into account since people with disabilities may want to be in the charge of the wheelchair.

Motor imagery enables users to control an object by imagining the body movement, Choi (2012) propose an MI BCI for smart wheelchair control which can perform three motion commands including turning left and right and going forward, and validated this robot in a real world. Badajena, Sethi, Dash, Rout, and Sahoo (2022) developed a system that receive, process, and classify EEG signals using naive Bayes, support vector machine (SVM) and decision tree. What is more, Dong, Zhang, Zhu, Du, and Tong (2022) designed a multi-modal platform for wheelchair control by distinguishing between SSVEP and MI potentials. The main disadvantage of such MI BCIs is that they require extensive training and users may not be able to perform imagination accurately. Ferreira, Silva, Celeste, Bastos Filho, and Sarcinelli Filho (2007) presented a human–machine interface (HMI) based on the signals generated by eye blinks (EMG) or brain activity corresponding to visual demands (EEG). However, driving a wheelchair with frequent usage of eyes can be difficult for users (may get exhausted) and potentially dangerous, thus it is not practical in real life scenarios. On the other hand, Alhakeem, Ali, and Abd-Alhameed (2020) non-brain source signals, such as eye-blinking, teeth clenching, jaw squeezing to increase the number of commands. They used Discrete Wavelet Transform with Auto regressive to extract the signal's features. These features are classified by using a linear SVM classifier.

The multi-modal BCI system is also becoming popular recently, Dong et al. (2022) proposed a threshold discrimination based solution that users can continuously control the wheelchair to turn left or right via MI signals, while control the wheelchair to start/stop and forward/backward by using the SSVEP stimulation panel. However, calibration data typically contain insufficient discriminability, resulting in unreliable feedback in the earliest stage of the training, which may decrease subjects' motivation. To improve the performance in the early stages of MI training, a novel hybrid BCI paradigm based on MI and P300 is provided by Zuo et al. (2020). Authors use reliable MI classifications to correct the unreliable P300 classifications. Toward the most related work in detecting user facing directions, Matsumotot et al. (2001) developed a robust vision-based system to detect face and gaze information using a stereo camera system. Based on the face and gaze information they are able to estimate the direction intentions of the user and thus steering the wheel chair. However, the system requires specialized infrastructures and the user still has to focus on the object during moving, which is less practical for real world scenarios.

To the best of our knowledge, SmartRolling is the first human–machine interaction approach that jointly leverages EEG signals targeting ME tasks while at the same time using IMU sensors and cameras to estimate user head motion and facing direction. Existing multi-sensor based solutions either mainly use different types of EEG signals or only focus eye gaze tracking method for wheelchair control, which are less intuitive. More importantly, we also propose a modified structure of deep model for data channel reduction, which could improve the system performance significantly.

6. Conclusion

In this paper, we aim to build a EEG based smart wheelchair control system, SmartRolling, that jointly leverages EEG BCI and motion sensing techniques. Specifically, SmartRolling offers two wheelchair-actuation modes for users with different physical conditions : (1) head motion only — people who are severely disabled but able to do basic tasks using eyes and head, and (2) head and hands motion — in addition to type 1, people who also have functioning hands. The system issues operation commands by recognizing different EEG patterns elicited by motor execution (ME) tasks including eye blink, jaw clench, head turn, and fist open/close, while at the same time estimates users' steering intentions based on their facing direction by leveraging inertial measurements and computer vision techniques. The experiment results demonstrate that the proposed system is robust and effective to meets the individual's needs and has great potential to promote better health.

CRediT authorship contribution statement

Landu Jiang: Conception and design of study, Acquisition of data, Analysis and/or interpretation of data, Writing – original draft, Writing – review & editing. **Cheng Luo:** Conception and design of study, Acquisition of data, Analysis and/or interpretation of data, Writing – original draft. **Zexiong Liao:** Acquisition of data, Analysis and/or interpretation of data, Writing – original draft. **Xuan Li:** Conception and design of study, Writing – review & editing. **Qiuxia Chen:** Analysis and/or interpretation of data, Writing – review & editing. **Yuan Jin:** Acquisition of data, Analysis and/or interpretation of data, Writing – review & editing. **Kezhong Lu:** Conception and design of study, Writing – original draft. **Dian Zhang:** Conception and design of study, Analysis and/or interpretation of data, Writing – original draft, Writing – review & editing.

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.

Data availability

The data that has been used is confidential.

Acknowledgments

This research was supported in part by NSFC, China 61872247, Shenzhen Peacock Talent Grant 827-000175, Guangdong Natural Science Fund 2019A1515011064, Shenzhen Polytechnic Fund 6020320004K, Key Technology Research and Development Program of China 2019YFE0196800, Overseas High-level Talents of Shenzhen KQJSCX20180329191021388. All authors approved the version of the manuscript to be published.

References

Al-Qaysi, Z., Zaidan, B., Zaidan, A., & Suzani, M. (2018). A review of disability EEG based wheelchair control system: Coherent taxonomy, open challenges and recommendations. *Computer Methods and Programs in Biomedicine*, *164*, 221–237.

Alhakeem, Z. M., Ali, R. S., & Abd-Alhameed, R. A. (2020). Wheelchair free hands navigation using robust DWT_AR features extraction method with muscle brain signals. *IEEE Access*, *8*, 64266–64277.

Aljalal, M., Ibrahim, S., Djemal, R., & Ko, W. (2020). *Comprehensive review on brain-controlled mobile robots and robotic arms based on electroencephalography signals*. Springer.

Badajena, J. C., Sethi, S., Dash, A., Rout, P., & Sahoo, R. K. (2022). Smart wheelchair using brain waves through machine learning. In *Intelligent systems* (pp. 445–457). Springer.

Barua, S., Ahmed, M. U., Ahlstrom, C., Begum, S., & Funk, P. (2017). Automated EEG artifact handling with application in driver monitoring. *IEEE Journal of Biomedical and Health Informatics*, *22*(5), 1350–1361.

Bi, L., Fan, X.-A., & Liu, Y. (2013). EEG-based brain-controlled mobile robots: a survey. *IEEE Transactions on Human-Machine Systems*, *43*(2), 161–176.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Choi, K. (2012). Control of a vehicle with EEG signals in real-time and system evaluation. *European Journal of Applied Physiology*, *112*(2), 755–766.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297.

Dong, E., Zhang, H., Zhu, L., Du, S., & Tong, J. (2022). A multi-modal brain–computer interface based on threshold discrimination and its application in wheelchair control. *Cognitive Neurodynamics*, 1–11.

Farnebäck, G. (2003). Two-frame motion estimation based on polynomial expansion. In *Scandinavian conference on image analysis* (pp. 363–370). Springer.

Ferreira, A., Silva, R., Celeste, W., Bastos Filho, T., & Sarcinelli Filho, M. (2007). Human–machine interface based on muscular and brain signals applied to a robotic wheelchair. *90*, In *Journal of Physics: Conference Series*. (1), IOP Publishing, Article 012094.

Gomez Torres, I., Parmar, G., Aggarwal, S., Mansur, N., & Guthrie, A. (2019). Affordable smart wheelchair. In *Extended abstracts of the 2019 Chi conference on human factors in computing systems* (pp. 1–6).

Guinness, K. (2020). The 16 types of camera shots and angles (2020 video guide). Boords.Comhttps://boords.com/blog/16-types-of-camera-shots-and-angles-with-gifs.

Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. In *Proceedings of the IEEE Conference on computer vision and pattern recognition* (pp. 7132–7141).

Iturrate, I., Antelis, J. M., Kubler, A., & Minguez, J. (2009). A noninvasive brain-actuated wheelchair based on a P300 neurophysiological protocol and automated navigation. *IEEE Transactions on Robotics*, *25*(3), 614–627.

Jiang, L., Chen, X., & He, W. (2016). SafeCam: Analyzing intersection-related driver behaviors usingmulti-sensor smartphones. In *2016 IEEE International conference on pervasive computing and communications (PerCom)* (pp. 1–9).

Jiang, L., Xie, W., Zhang, D., & Gu, T. (2022). Smart diagnosis: Deep learning boosted driver inattention detection and abnormal driving prediction. *IEEE Internet of Things Journal*, *9*(6), 4076–4089.

Li, Y., Liu, Y., Cui, W.-G., Guo, Y.-Z., Huang, H., & Hu, Z.-Y. (2020). Epileptic seizure detection in EEG signals using a unified temporal-spectral squeeze-and-excitation network. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *28*(4), 782–794.

Mandel, C., Lüth, T., Laue, T., Röfer, T., Gräser, A., & Krieg-Brückner, B. (2009). Navigating a smart wheelchair with a brain-computer interface interpreting steady-state visual evoked potentials. In *2009 IEEE/RSJ International conference on intelligent robots and systems* (pp. 1118–1125). IEEE.

Matsumotot, Y., Ino, T., & Ogsawara, T. (2001). Development of intelligent wheelchair system with face and gaze based interface. In *Proceedings 10th IEEE International workshop on robot and human interactive communication. ROMAN 2001 (Cat. No. 01TH8591)* (pp. 262–267). IEEE.

Pfurtscheller, G., & Da Silva, F. L. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology*, *110*(11), 1842–1857.

Pires, G., Castelo-Branco, M., & Nunes, U. (2008). Visual P300-based BCI to steer a wheelchair: a Bayesian approach. In *30th Annual international conference of the IEEE engineering in medicine and biology society* (pp. 658–661). IEEE.

Prueckl, R., & Guger, C. (2010). Controlling a robot with a brain-computer interface based on steady state visual evoked potentials. In *The 2010 International joint conference on neural networks* (pp. 1–5). IEEE.

Puanhvuan, D., & Wongsawat, Y. (2012). Semi-automatic P300-based brain-controlled wheelchair. In *2012 ICME International conference on complex medical engineering* (pp. 455–460). IEEE.

Rebsamen, B., Guan, C., Zhang, H., Wang, C., Teo, C., Ang, M. H., et al. (2010). A brain controlled wheelchair to navigate in familiar environments. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *18*(6), 590–598.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on computer vision and pattern recognition* (pp. 2818–2826).

Tanaka, K., Matsunaga, K., & Wang, H. O. (2005). Electroencephalogram-based control of an electric wheelchair. *IEEE Transactions on Robotics*, *21*(4), 762–766.

Tang, Y. (2013). Deep learning using linear support vector machines. arXiv preprint arXiv:1306.0239.

Turnip, A., Soetraprawata, D., & Tamba, T. A. (2015). Eeg-ssvep signals extraction with nonlinear adaptive filter for brain-controlled wheelchair. In *2015 15th International conference on control, automation and systems* (pp. 1870–1873). IEEE.

Wohlfert, E. A., Blader, I. J., & Wilson, E. H. (2017). Brains and brawn: toxoplasma infections of the central nervous system and skeletal muscle. *Trends in Parasitology*, *33*(7), 519–531.

World Health Organization 0000. Wheelchair Service Training Package, https://www.who.int/disabilities/technology/wheelchairpackage/en/.

Zuo, C., Jin, J., Yin, E., Saab, R., Miao, Y., Wang, X., et al. (2020). Novel hybrid brain–computer interface system based on motor imagery and P300. *Cognitive Neurodynamics*, *14*(2), 253–265.