An Assistive Robotic System for Cognitive State Assessment in Individuals with Spinal Cord Injury

Ashish Jaiswal
The University of Texas at Arlington
USA

Gaurav Nale
The University of Texas at Arlington
USA

Qiyuan An
The University of Texas at Arlington
USA

Hamza Reza Pavel
The University of Texas at Arlington
USA

Enamul Karim
The University of Texas at Arlington
USA

Sneh Acharya
The University of Texas at Arlington
USA

Fillia Makedon
The University of Texas at Arlington
USA

ABSTRACT

We introduce an innovative assistive robotic setup tailored for assessing the cognitive state of individuals with spinal cord injuries (SCI) during their daily activities. Utilizing physiological sensors such as ECG, EEG, and EDA, along with cameras for facial expression, our system is designed to detect and evaluate cognitive fatigue in participants as they engage with a collaborative robot. Specifically, two tasks, **Cooking Pasta Sauce** and **Getting Ready for Work**–have been crafted to gather data on cognitive states. Participants interact with the robot using natural language (*English*) to perform tasks, while their physiological responses, facial expressions, and activities are recorded. The study comprises three phases of cognitive fatigue: baseline, moderate, and severe. Cognitive fatigue is induced through the N-back task paradigm, and its severity is assessed using the Visual Analogue Scale for Fatigue (VAS-F) questionnaire. Our system is designed to intervenes during the tasks based on the detected cognitive fatigue levels. In this paper, we concentrate on validating the cognitive fatigue detection system using only physiological sensors during task performance, achieving an accuracy of **85.7%** and a recall of **0.87**. We provide detailed insights into the system design and present a preliminary analysis of the gathered data.

CCS CONCEPTS

• **Human-centered computing** → User interface toolkits.

KEYWORDS

physiological sensors, human computer interface, multimodal system, cognitive fatigue, assistive robots

ACM Reference Format:
Ashish Jaiswal, Gaurav Nale, Qiyuan An, Hamza Reza Pavel, Enamul Karim, Sneh Acharya, and Fillia Makedon. 2024. An Assistive Robotic System for Cognitive State Assessment in Individuals with Spinal Cord Injury. In *The PErvasive Technologies Related to Assistive Environments (PETRA) conference (PETRA '24), June 26–28, 2024, Crete, Greece.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3652037.3652068

PETRA '24, June 26–28, 2024, Crete, Greece

ACM ISBN 979-8-4007-1760-4/24/06
https://doi.org/10.1145/3652037.3652068

1 INTRODUCTION

This paper introduces the design and development of an end-to-end personalized assistive robotic system, named **iRCSA** (*Intelligent Robotic Cooperation for Safe Assistance*), with the primary goal of recognizing, assessing, and responding to the Cognitive Fatigue (CF) levels in individuals with Spinal Cord Injury (SCI) during Human-Robot Cooperation (HRC) tasks. In light of the growing prevalence of robotics and Artificial Intelligence (AI), assistive robots hold great promise for improving the independence and quality of life for individuals with disabilities. While existing research predominantly focuses on ensuring safe HRC in industrial settings, there remains a notable gap in understanding the cognitive states of individuals engaging with robots in their daily lives.

To bridge this gap, our project endeavors to design and develop the iRCSA system, integrating a multi-sensory system to detect participants' CF levels and an assistive robot capable of providing corresponding support. Physiological data (ECG, EDA, EEG) along with audio and video is collected from individuals with SCI during HRC tasks. Utilizing advanced machine learning algorithms, pertinent features are extracted from the collected data, automatically assessing the individual's CF level. Based on this assessment, the iRCSA system is planned to dynamically adjust the robot's behavior to offer personalized support. However, in this paper, we mainly focus on understanding the cognitive state of the participants while performing the task.

The development and evaluation of iRCSA adheres to the Participatory Action Research (PAR) approach, involving SCI subjects at every stage of the project. Their invaluable insights and feedback is taken to ensure the acceptability and usability of the proposed system. HRC scenarios, encompassing daily tasks such as cooking and preparing for work, are orchestrated to facilitate cooperative interactions between individuals with SCI and the assistive robot. The potential outcomes of this research are significant, promising to elevate the quality of life for individuals with SCI by enabling assistive robots to comprehend and respond to their cognitive state. By addressing the cognitive aspect of HRC, the iRCSA system stands to enhance the safety, efficiency, and effectiveness of assistive robotic systems in delivering support and care to individuals with SCI.

This paper begins by reviewing existing research in assistive robotics, physiological sensors, and cognitive fatigue assessments (section: 2). Subsequently, a comprehensive exploration of the system's design is provided, explaining each module's functionality: the mobile robot assistant, physiological sensor setup, facial expressions recording, and speech recognition (section: 3). The methodology employed in the study is then detailed, encompassing data collection and labeling processes (sections: 4 & 5). A preliminary analysis of the acquired sample data is conducted, centering on the cognitive fatigue detection system (section: 6). The findings are presented alongside a comparative assessment against prior methods (section: 7). The paper concludes by outlining potential applications and suggesting future directions for research (section: 8).

2 RELATED WORK

The convergence of assistive robots, physiological sensors, facial emotion analysis, and human activity recognition technologies is reshaping our comprehension of cognitive states, particularly in the context of individuals with disabilities. Feelings of frustration, isolation, anxiety, and even depression can emerge, often stemming from the loss of mobility and autonomy. This emotional landscape can further interact with cognitive states, highlighting the intricate relationship between mental and physical well-being [21]. Assistive robots, defined as *"machines designed to restore or enhance physical abilities"*, offer a ray of hope in healthcare. Research, exemplified by the work of Krebs et al. [16], underscores their potential to not only improve physical well-being but also restore a sense of autonomy, alleviating feelings of dependency.

On the other hand, physiological sensors in the domain of human-computer interaction such as EEG hold promise for gauging cognitive states in people [25]. For instance, Schirrmeister et al. [26] utilize deep learning with convolutional neural networks to assess cognitive state of human subjects from EEG signals, enabling the rise of user-centric systems. Similarly, Bashivan et al. [2] further extended on the work the verify the potential of neural networks in decoding physiological signals. However, the intrusive nature of some sensors pose challenges to continuous monitoring feasibility. To mitigate the intrusive issue with physiological sensors, Baltrusaitis et al. [1] developed a deep learning algorithm to decode facial micro-expressions to provide insights into an individual's emotional state. The non-intrusive nature of cameras make them a valuable and viable option to monitor a person's cognitive state.

Similarly, to understand human emotion, natural language has been extensively research to understand the underlying emotion behind text and speech from a person. For instance, [10] utilizes speech between pilots and air traffic controllers to detect fatigue early and prevent accidents, while [11] explores the relationship between speech and illness severity. Thus, speech has demonstrated a close connection with fatigue levels, making it crucial to detect fatigue early in certain occupations to avert potential dangers or accidents.

Hence, while numerous studies have explored the different domains above, a palpable research gap exists, particularly in their combined application for people with disabilities and spinal cord injuries. Our system aims to bridge that gap by combining multiple modalities to understand human emotion and cognitive state. By interweaving assistive robots, physiological sensors, and facial analysis, our aim is to gain a holistic understanding of the daily cognitive experiences of individuals with SCIs, moving beyong mere assistance to craft a deeply empathetic and responsive system.

3 SYSTEM DESIGN

This section outlines the specifics of the experimental setup. We devise a system wherein participants engage in two Human-Robot Collaboration (HRC) scenarios featuring robotic assistance, simulating real-world situations. The designated tasks include ***Cooking a Pasta Sauce*** and ***Getting Ready for Work***. The objective of the robotic assistant is to aid wheelchair-bound participants in carrying out everyday activities. To ensure the successful completion of these tasks, a robotic manipulator, as illustrated in Fig. 2, is employed to fetch objects for participants as needed. The robotic system seamlessly integrates with a multi-sensory system to collect physiological data, facilitating the evaluation of participants' cognitive fatigue states. Additionally, a speech assistant module is implemented to assist participants in interacting with the robot using natural language. The entire participant activity, including facial expressions, is recorded using RGB cameras.

The experimental setup is illustrated in Fig. 2, featuring two tables: one designated for task performance by the subject and the other for the placement of necessary items. The figure highlights the ingredients essential for the ***Cooking Pasta Sauce*** task. In this task, participants are presented with a list of cooking ingredients (such as tomato sauce, mushrooms, salt, etc.) that they must memorize before commencing the task. Subsequently, they instruct the robot to retrieve each ingredient individually, initiating the pasta sauce cooking process. The sequence and quantity of ingredients procured by the robot are entirely dependent on the participant's instructions. Similarly, in the ***Getting Ready for Work*** task, the robot aids the participant in obtaining six common items essential for preparing to leave for work. These items encompass a cellphone, laptop, headphone, keys, etc.

3.1 Mobile Robot Assistant

The mobile robotic assistant module of our system comprises two primary components: the Summit XL omnidirectional mobile robot and the 7-DOF Franka Emika Panda robotic manipulator, commonly referred to as the Panda arm. The Summit XL robot facilitates free movement in confined spaces, offering agility and versatility. On top of its base, the Panda arm is equipped with a 2-finger gripper capable of picking and fetching objects with dimensions of up to 80mm in width and weighing up to 3 kg.

The Panda arm is equipped with torque sensors on each joint, ensuring precise interactions with objects and providing a heightened level of safety when operating in close proximity to humans. This safety feature is crucial for collaborative tasks. Additionally, the robotic arm is fitted with an RGBD camera, enhancing its capability for grasping and picking objects by providing visual information.

To support navigation and interaction, the robot base is equipped with LiDAR sensors and cameras. This sensor suite aids the robot

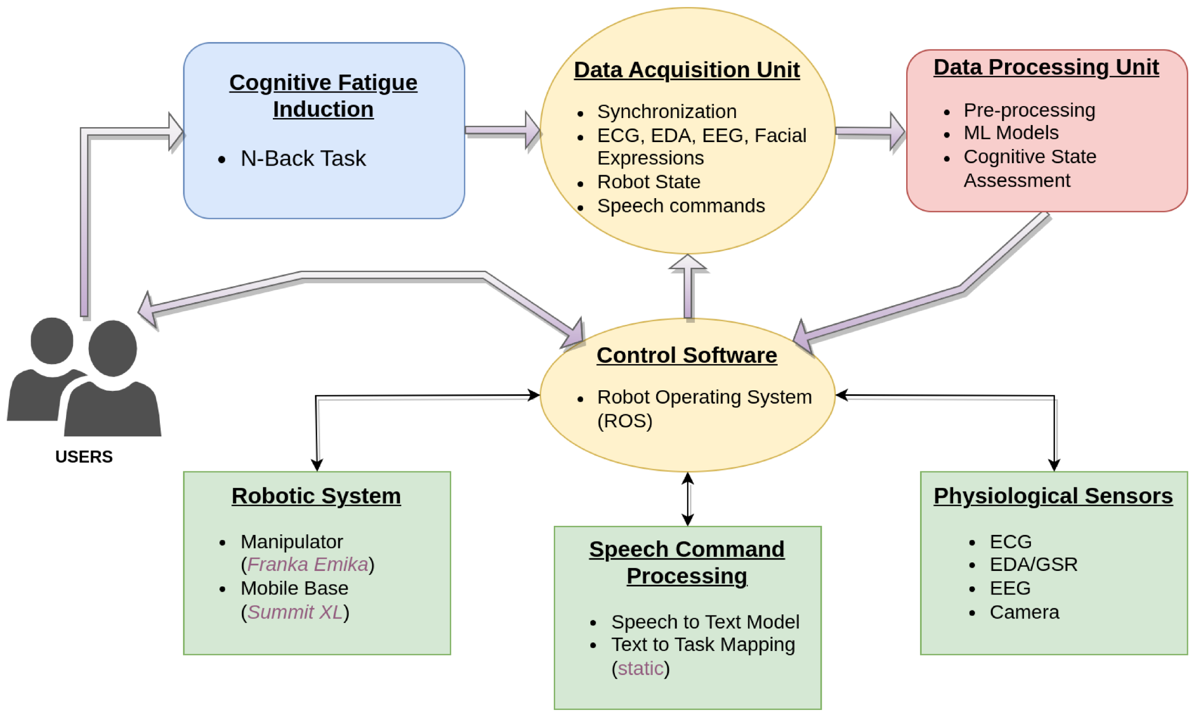

Figure 1: System Design: A robotic system that can monitor the user's cognitive state using physiological, vision, and audio sensors and adapt its behavior accordingly.

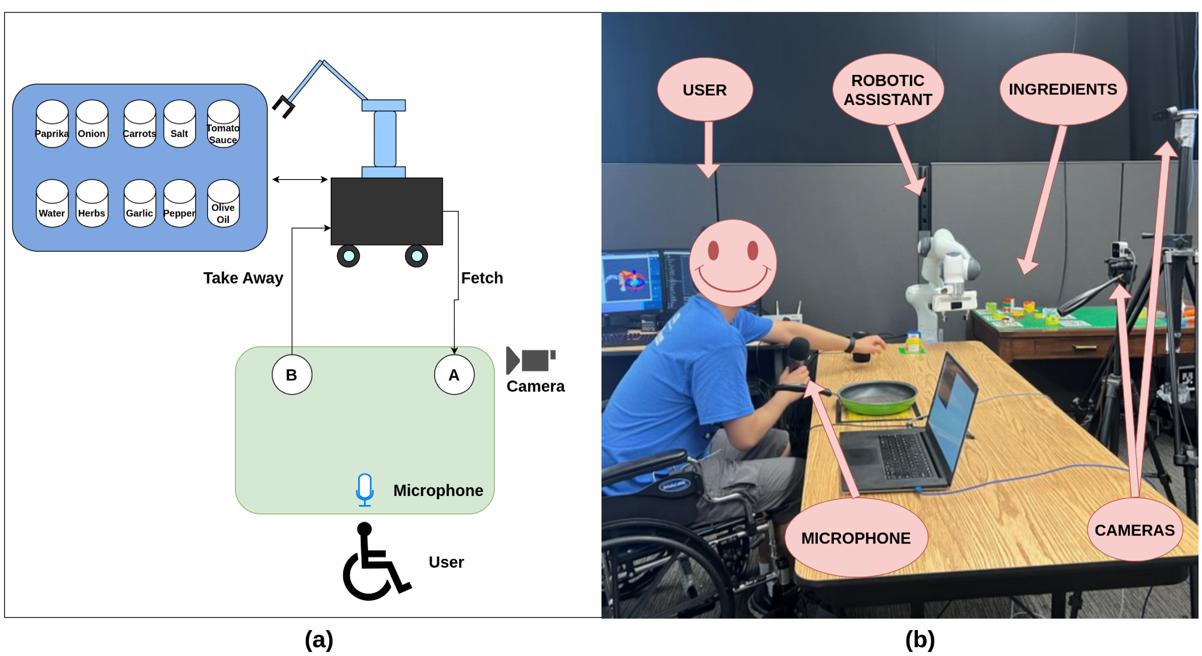

Figure 2: a. Overview of the experimental setup. b. A subject with SCI performing the simulated Pasta Sauce cooking task.

in perceiving its surroundings and navigating through the environment effectively. Subjects have the flexibility to guide the robots through predefined sequences of actions or make real-time adjustments using spoken commands, allowing for intuitive and dynamic interaction with the mobile robotic assistant module.

3.2 Physiological Sensor Setup

Our human-centric framework incorporates a diverse array of physiological sensors seamlessly integrated into a multi-sensory system. This system is adept at collecting electroencephalogram (EEG), electrocardiogram (ECG), and electrodermal activity (EDA). These physiological sensors play a crucial role in evaluating cognitive fatigue (CF) levels during Human-Robot Collaboration (HRC) scenarios. The PLUX Biosignals sensor module [4] is utilized for gathering ECG and EDA data, while the Muse S headset [19] is employed for collecting EEG signals. These sensors actively monitor electrical activity in the heart, skin, and the brain and are significant in detecting the cognitive state of a person [14, 15, 17].

During various tasks performed by subjects, EEG signals are captured using the Muse S headset, which features four electrodes (AF7, AF8, TP9, TP10) strategically positioned on different areas of the head. The EEG signals provide quantifiable data on brain electrical activity, categorized into five frequency bands: alpha, beta, delta, gamma, and theta, each corresponding to a distinct brain state. As illustrated in Fig. 3, electrodes for ECG and EDA from the PLUX Bluetooth module are attached to different points on the body, with red and white dots denoting the front and the black dot denoting the back. Both the Muse headset and the PLUX module are connected to the ROS system via Bluetooth, ensuring a continuous stream of data to their respective topics.

3.3 Facial Expressions

In addition to physiological sensors, our configuration includes two types of camera sensors to intricately capture the physical activities and facial responses of individuals with spinal cord injury (SCI) during their daily tasks. These RGB-D cameras, equipped with both color and depth capabilities, are designed to offer a comprehensive view of the environment. One camera captures the overall activity area (the table) where the participant performs tasks, while the other camera provides a close-up, dynamic view of the participant during the ***Cooking Pasta Sauce*** and ***Getting Ready for Work*** tasks.

An ancillary objective of our system is to extract crucial features such as human body position and facial landmarks by harnessing cutting-edge libraries like OpenPose [6] and OpenFace [1]. Recognizing human activity is pivotal, and our system strives to predict and interpret captured behaviors using powerful computer vision algorithms. These inferred actions can yield essential insights into cognitive fatigue levels, especially during interactions with the robot. For instance, by monitoring eye motion, blink rate, and utilizing facial landmarks, we can discern different eye movement patterns, which serve as vital indicators of cognitive fatigue. However, it's important to note that the exploration of vision data is beyond the scope of this paper.

3.4 Speech Recognition

To bridge the communication gap between humans and the robot, we implement a speech recognition module for receiving commands from participants and controlling the mobile robot. The speech recognition pipeline, implemented through the Google Cloud Speech library for Python [1], involves the following steps: i) adapting to ambient noise, ii) recognizing a trigger word and identifying the task keyword, and iii) dispatching the pick-and-place command to the robotic system. In the initial step, the speech recognizer acclimates to the ambient environmental noise for 0.5 seconds to enhance the recognition of speech commands in the subsequent step. During the command recognition stage, the participant is required to:

- Begin by saying, "Hi/Hey Robot" to prompt the robot to listen to the command,
- Subsequently, state, "I would like to get an *item_name*" as the command to instruct the robot to fetch the specified item.

The items are fetched one at a time, with the robot initiating the task by responding with "Sure, fetching *item_name* for you." After delivering the item to the designated location, the robot notifies the user of task completion. The predetermined items for the "Cooking Pasta Sauce" and "Getting Ready for Work" tasks are fixed and include:

- **Cooking Pasta Sauce**: *pasta, cheese, carrots, tomato sauce, green beans, mushrooms, garlic, chili, butter, salt, bell pepper, and corn*
- **Getting Ready for Work**: *cellphone, coffee, keys, calculator, laptop, and headphones*

Upon issuing the command, we capture both the speech audio and transcribed text commands that can be used to facilitate the downstream cognitive fatigue detection task in the future. The speech audio is stored in WAV format, while the transcribed text is obtained from the Google Cloud Speech library and saved in CSV format. The rationale behind collecting speech data from participants stems from previous studies indicating that fatigue can lead to potential dangers, accidents, or a decline in life quality [8, 10].

4 METHODOLOGY

4.1 N-Back Task Paradigm

The N-Back tasks employed in this study involves participants to a series of stimuli, such as letters, by indicating whether the current stimulus matched the one present N trials earlier as shown in Fig. 4. The 2-Back task is chosen for its efficacy in inducing cognitive fatigue without overwhelming the participants [12]. The continuous performance nature of the task taxes working memory and executive functions, leading to increased cognitive load and fatigueness. Fig. 4 shows the GUI designed for the 2-Back task along with the play screen (right) showing letters one at a time.

4.2 Visual Analogue Score of Fatigue (VAS-F) Questionnaire

The VAS-F questionnaire is a simple and widely used questionnaire to assess fatigue [7]. It has been used in a variety of settings, including clinical research and occupational health [3, 13]. It is administered at key points during the experiment to subjectively assess cognitive fatigue levels. The scale consists of 18 items relating to the subjective experience of fatigue. Each item asks respondents to place an "X," representing how they currently feel, along a visual analogue line that extends between two extremes (e.g., from "not at

[1]https://cloud.google.com/speech-to-text/docs/speech-to-text-client-libraries

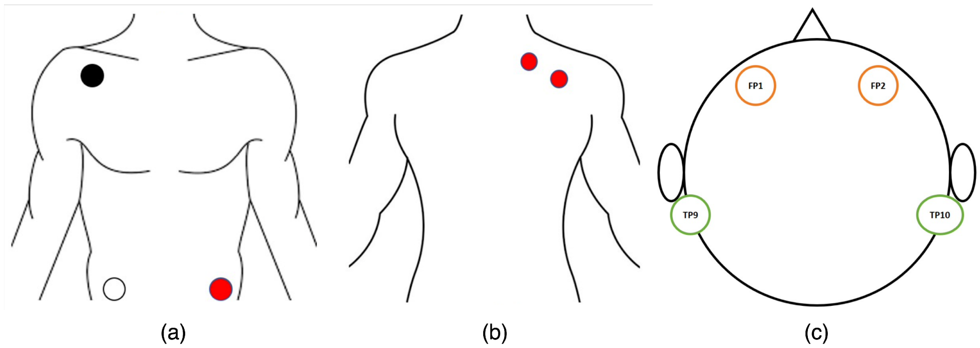

Figure 3: Illustration of the placement of sensors on the human body: (a) ECG sensors located on the right shoulder, left hip, and right hip, forming Einthoven's triangle [27]; (b) EDA/GSR electrodes positioned on the left shoulder to capture skin conductivity; and (c) EEG sensor positions based on the 10–10 electrode system employed by MUSE. This system records data from the TP9, AF7, AF8, and TP10 positions in the electrode configuration

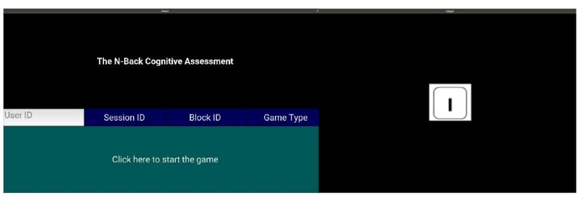

Figure 4: Graphical User Interface built for the N-Back tasks with an example image of a letter during a game round on the right.

all tired" to "extremely tired"). The participant is asked to select an option to indicate their current level of fatigue. In this study, VAS-F score between 40-70 has been considered Moderate CF and scores over 70 as Severe CF.

4.3 Experimental Phases

The experimental study is designed to be conducted in three distinct phases to systematically evaluate the impact of cognitive fatigue on the performance. Each phase aims to capture the participant's cognitive state at different levels of cognitive fatigue. Participants start the experiment in a rested state, free from any cognitive load. During this phase, baseline measurements are obtained from the physiological sensors (ECG, EMG, and EEG) and the facial expressions through the cameras. The baseline data enables us to normalize the data signals for each participant. Following the baseline phase, participants are asked to engage in N-Back tasks (2-Back). N-back tasks have been widely recognized for inducing cognitive fatigue and mental workload [22]. After every few rounds of the N-back task, participants are required to fill a VAS-F questionnaire that determines their level of cognitive fatigue. When moderate cognitive fatigue is induced, the participants are asked to perform the daily activity tasks (cooking pasta and getting ready for work). Finally, when the VAS-F score is administered to cross the 70 mark, the participant is considered in a severe CF phase where they perform the daily activity tasks one last time.

While performing the Cooking Pasta and Getting Ready for Work tasks, physiological signals are constantly recorded along with the facial expressions of the participants. The robotic system assists the participants in completing the tasks. Speech commands are issued by the participant in order to interact with the robot. All the data modalities are recorded during the tasks, providing a comprehensive dataset for analysis.

5 DATA COLLECTION AND LABELING

Participant recruitment was a collaborative initiative between the Student Access & Resource (SAR) and the Office of Accessible Education (OAE) at the university. We specifically reached out to members of the UTA basketball team who use wheelchairs. For our initial study, we invited eight participants, ensuring a balanced representation by sex. The primary goal was to gather feedback on the design of two Human-Robot Collaboration (HRC) tasks: **Cooking Pasta Sauce** and **Getting Ready for Work**. Participants, while assisted by the mobile robot assistant, engaged in these daily activities and interacted with the robot using natural language (English) through speech. The study was designed to evaluate task performance at three distinct phases of cognitive fatigue: ***baseline***, ***moderate***, and ***severe***.

Cognitive fatigue was induced through multiple rounds of N-back tasks, with participants completing a minimum of six rounds. Additional rounds were administered if the desired level of ***severe*** fatigue was not reached. The VAS-F questionnaire gauged the level of cognitive fatigue experienced by participants after each N-back task. Physiological sensor data, collected using Muse EEG headbands, ECG sensors, and Bioplux EMG sensors, provided insights into cognitive workload, heart rate variability, and skin sensitivity, respectively. Facial expressions captured by a camera will be used to analyze for signs of stress, fatigue, and changes in emotional state. Speech commands were recorded and cane be used to assess variations in participants' vocal characteristics, potentially correlating with cognitive fatigue levels.

Sensor data were synchronized using timestamps from the robot operating system (ROS). Preprocessing steps were applied to EEG, ECG, and EDA data to eliminate noise, artifacts, and baseline shifts.

Spectral analysis was conducted on EEG data to extract cognitive load-related frequency bands. ECG data were processed to compute heart rate variability parameters, and EDA data were filtered and normalized for skin conductivity quantification.

The VAS-F questionnaire, yielding subjective cognitive fatigue scores between 0-100, provided a general rating of overall fatigue intensity perceived by participants. Scores falling below **40** were categorized as *No Fatigue (< 40)*, those between **40 and 70** as *Moderate Fatigue (> 40 and < 70)*, and anything surpassing **70** as *Severe/Extreme Fatigue (> 70)*. The VAS-F scores served as a benchmark for evaluating the induced cognitive fatigue's effectiveness and validating the multi-modal data analysis.

6 COGNITIVE FATIGUE DETECTION SYSTEM: PRELIMINARY ANALYSIS

This paper primarily centers on detecting the cognitive fatigue state of participants during Human-Robot Collaboration (HRC) tasks. While vision data, speech transcriptions, and robot state are recorded during the experiments, their primary purpose is to contribute to the development of an intervention system, which will be further explored in future research. In this section, we explain the pre-processing pipeline for physiological signal data and elaborate on their significance in detecting the three pre-defined levels of cognitive fatigue.

6.1 EEG

The acquisition of EEG signals during experimental tasks is conducted utilizing the MUSE S headset. This device is equipped with four electrodes positioned at AF7, AF8, TP9, and TP10, establishing contact with specific regions of the head, as illustrated in Fig. 3(c). The EEG signals measure electrical activity in the brain, and in this experiment, we analyze them by decomposing them into five distinct frequency bands: alpha, beta, gamma, and theta, as presented in Fig. 5. Each frequency band corresponds to a distinct cerebral state. For instance, delta waves occur in the frequency range of 0.5 Hz to 4 Hz and are present during sleep, while beta waves occur between 13 Hz to 30 Hz and are associated with active thinking. Similarly, other waves and their associated states include alpha waves (8-12 Hz) denoting normal wakefulness, gamma waves (30-80 Hz) signifying sensory perception integration, and theta waves (4-7 Hz) indicating drowsiness and the early stages of sleep. In addition, 50-60 Hz frequencies are pre-processed beforehand to mitigate potential power line interference on the EEG signals.

6.2 ECG and EDA/GSR

6.2.1 ECG. Electrocardiogram (ECG) signals provide insights into changes in the cardiovascular system by reflecting the heart's electrical activity. These signals contain crucial information regarding cardiac pathologies that impact the heart, characterized by five peaks known as fiducial points labeled P, Q, R, S, and T [24]. Numerous studies have established a correlation between fatigue and alterations in the body's cardiovascular response [20]. The ECG signals are recorded using Einthoven's triangle approach [9], an imaginary formation of three limb leads in electrocardiography, creating a triangle encompassing the two shoulders and the pubis as shown in Fig. 3(a).

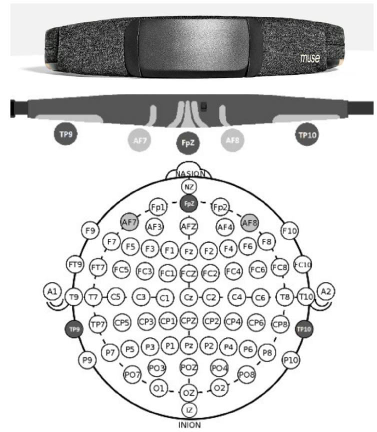

Figure 5: Electrode placement comparison between MUSE and the international 10-20 system. Top: Commercial MUSE S Headband. Bottom: 10-20 Electrode Placement System

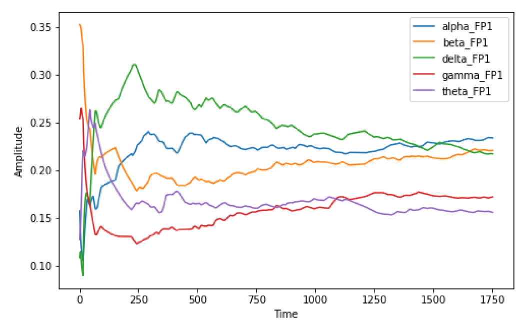

Figure 6: Raw Amplitude plot of Frequency bands extracted from the electrode at AF7 position (on MUSE) from a sample raw EEG signal from one of the subjects. The readings were collected during one of the 2-Back tasks undergoing for a little under 3 minutes.

To eliminate undesired noise from the ECG signals, the Pan and Tompkins QRS detection algorithm is employed [23]. Initially, the signals are cleaned through a high-pass Butterworth filter with a fixed cut-off frequency of 0.5 Hz. Subsequently, a notch filter is applied to the cleaned signal to eliminate components with a frequency of 50 Hz, thereby mitigating power line interference. Following this, RR intervals are derived from the signal's R_Peaks and further refined by removing outliers as shown in Fig. 7. The missing values are then imputed using the linear interpolation method. Ultimately, a comprehensive set of 113 time-domain and frequency-domain features, encompassing metrics such as heart rate variability (HRV), are extracted for the purpose of training machine learning models.

6.2.2 EDA/GSR. Conversely, Electrodermal Activity (EDA), commonly known as galvanic skin response (GSR), mirrors the sympathetic nervous system's activity. This system is influenced by physiological and emotional stimuli, gauging the skin conductivity

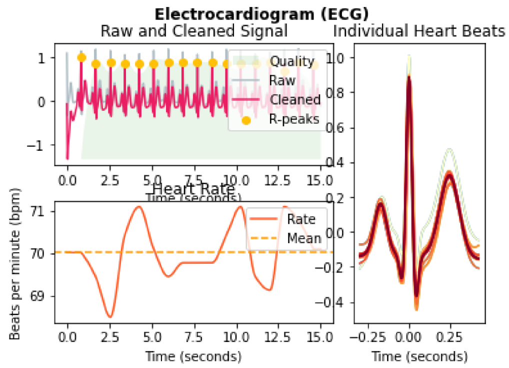

Figure 7: A 15-second ECG signal sample after noise removal from one of the subjects recorded at 250 samples/second during a recorded heart rate of 70 BPM. The R-peaks are used as one of the ECG features for training machine learning models along with other Heart-Rate Variability (HRV) features.

of the body. EDA signals can additionally indicate the intensity of one's emotional state, enabling the identification of psychological or emotional arousal episodes.

The EDA signals undergo initial processing by applying a low-pass Butterworth filter with a cut-off frequency of 3 Hz. EDA signals can be deconstructed into two distinct components: phasic and tonic [5]. Given that the phasic component represents the more rapidly changing aspect of the signal influenced by physiological responses to stimuli, it was selectively extracted for further analysis. Features, specifically Skin Conductance Response (SCR) peaks, were then derived from the refined EDA signals. Similarly, both time-domain and frequency-domain features were extracted from the EMG signals. The majority of the feature extraction procedures were conducted using the Neurokit2 package [18] for all three signals.

7 EXPERIMENTATION AND RESULTS

For the detection of different levels of cognitive fatigue, 100 statistical features are extracted from EEG signals and 129 combined features from ECG and EDA signals. These encompass diverse features at different frequency levels such as peaks, rates, onsets, offsets, and more. To train the machine learning models, rather than processing the entire signal for a task as a single input, we partition the temporal signals into multiple slices based on various window sizes (5 seconds, 10 seconds, and 20 seconds). Each signal slice inherits the same label as its parent signal, and features are then extracted. This approach augments the volume of input data points for ML model training. However, we also evaluate the models using complete signal blocks as inputs. Likewise, during inference, the input signal is subdivided into smaller slices based on the window size established during training. Each slice is classified individually by the model, and ultimately, the entire signal block is classified based on the predominant class among the classified slices. This technique enhances the model's resilience to noise or outliers in the signals, as noise within certain slices may have minimal impact on the final classification result.

Despite the limited sample size of eight participants in our current preliminary study, we employ transfer learning to leverage data acquired from previous studies for cognitive fatigue detection [14]. The dataset originates from a similar experimental setup involving N-back tasks designed to induce cognitive fatigue. We possess physiological sensor data (ECG, EEG, EDA, and EMG) from 32 healthy participants, who provided self-reported subjective VAS-F scores after each round of N-back. The features extracted from these samples are incorporated into the dataset from our preliminary study to facilitate the classification of the three intended levels of cognitive fatigue.

The entire aggregated dataset, comprising data from a total of 40 subjects, is randomly partitioned into training (70%, 28 subjects), validation (15%, 6 subjects), and test (15%, 6 subjects) sets. Stratified sampling is employed during the partitioning process to address potential imbalances in the dataset. Additionally, 5-fold cross-validation is performed for each of the models. Four distinct machine learning models—Logistic Regression (Log Reg.), Support Vector Machines (SVM), Random Forest (RF), and Long Short-Term Memory (LSTM) recurrent neural network—are employed in the analysis. Various combinations of features extracted from the signals are utilized to predict cognitive fatigue.

Logistic Regression, SVM, and RandomForest classifiers are trained on the features extracted from the physiological signals. However, the LSTM models (with 256 hidden layers) are trained on the raw signal due to its ability to process time-series data. We use the similar window-based method to train the LSTM models, where the input size of the EEG signals provided is *t x 20 x 1* (five frequency bands from each of the electrodes). On the other hand, ECG and EDA signals are combined to form inputs of size *t x 2 x 1*. Finally, the LSTM is trained on *t x 23 x 1* inputs for all signals combined. Here, "t" represents the number of timesteps in the signal, which varies based on the window size.

The Avg. Recall presented in the tables is the average recall for the **Moderate Fatigue** and **Severe Fatigue** conditions obtained across 5-fold cross-validation for each ML model. The best-performing value for each model among different window sizes is considered. Notably, the LSTM model outperforms others with an **85.7%** prediction accuracy in detecting cognitive fatigue states. The recall value of **0.87** indicates that actual fatigue cases are correctly identified 87% of the time, with only a 13% false positive rate.

8 CONCLUSION

This paper makes a significant contribution to the field of Human-Computer Interaction (HCI) by presenting a comprehensive framework that integrates multi-modal sensors for assessing cognitive fatigue in individuals with Spinal Cord Injury (SCI). The proposed system shows promise in enhancing our understanding of how cognitive fatigue impacts task performance and overall well-being. The study's results illuminate the challenges and opportunities in designing assistive systems that facilitate efficient task completion while prioritizing the cognitive well-being of users. In future research, we plan to explore the significance of acitivity videos, facial expressions, robotic state, and speech data in detecting cognitive fatigue, in addition to analyzing physiological signals. Furthermore, our future goal is to develop a real-time cognitive state analysis

Table 1: Detection of Cognitive Fatigue (CF) with EDA/GSR + EMG Features

Model	Accuracy (Window Size)				Avg. Recall
	5s	10s	20s	Full Block	
Log Reg.	68.1%	68.7%	68.9%	71.8%	0.59
SVM	**77.3%**	79.7%	80.1%	82.1%	0.68
RF	71.1%	**79.9%**	76.4%	80.9%	0.73
LSTM	68.6%	79.2%	**84.2%**	**84.5%**	**0.77**

Table 2: Detection of Cognitive Fatigue (CF) with EEG + EDA/GSR + EMG Features

Model	Accuracy (Window Size)				Avg. Recall
	5s	10s	20s	Full Block	
Log Reg.	64.2%	64.9%	66.7%	66.7%	0.69
SVM	**77.1%**	**80.3%**	80.3%	80.9%	0.77
RF	73.7%	77.8%	78.9%	78.8%	0.70
LSTM	68.8%	77.1%	**84.4%**	**85.7%**	**0.87**

Table 3: Comparison of different models with the state-of-the-art algorithms

Model	**Accuracy**	**Avg. Recall**	**Ref.**
RF	64.69%	0.65	[17]
RF	66.20%	0.66	[17]
LSTM	84.1%	0.90	[14]
LSTM (Ours)	**85.7%**	0.87	Table 2

system that enables the robot to intervene during Human-Robot Collaboration (HRC) tasks. This work lays the groundwork for future endeavors in designing personalized and responsive robotic assistance tailored to individuals with diverse cognitive states and abilities.

REFERENCES

[1] T. Baltrusaitis, A. Zadeh, Y. C. Lim, and L. P. Morency. 2018. OpenFace 2.0: Facial behavior analysis toolkit. In *IEEE International Conference on Automatic Face & Gesture Recognition*. IEEE.

[2] P. Bashivan, I. Rish, M. Yeasin, and N. Codella. 2015. Learning Representations from EEG with Deep Recurrent-Convolutional Neural Networks. *arXiv preprint arXiv:1511.06448* (2015).

[3] Angelique Belmont, Nathalie Agar, and Philippe Azouvi. 2009. Subjective fatigue, mental effort, and attention deficits after severe traumatic brain injury. *Neurorehabilitation and neural repair* 23, 9 (2009), 939–944.

[4] Biosignalsplux. [n. d.]. *Researcher Kit.* https://biosignalsplux.com/products/kits/researcher.html.

[5] Jason J Braithwaite, Derrick G Watson, Robert Jones, and Mickey Rowe. 2013. A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments. *Psychophysiology* 49, 1 (2013), 1017–1034.

[6] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. 2017. Realtime multi-person 2d pose estimation using part affinity fields. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 7291–7299.

[7] Trudie Chalder, G Berelowitz, Teresa Pawlikowska, Louise Watts, S Wessely, D Wright, and EP Wallace. 1993. Development of a fatigue scale. *Journal of psychosomatic research* 37, 2 (1993), 147–153.

[8] Carla Aparecida de Vasconcelos, Maurílio Nunes Vieira, Göran Kecklund, and Hani Camille Yehia. 2019. Speech analysis for fatigue and sleepiness detection of a pilot. *Aerospace medicine and human performance* 90, 4 (2019), 415–418.

[9] Willem Einthoven, G Fahr, and A De Waart. 1950. On the direction and manifest size of the variations of potential in the human heart and on the influence of the position of the heart on the form of the electrocardiogram. *American heart journal* 40, 2 (1950), 163–211.

[10] Harold P Greeley, Eric Friets, John P Wilson, Sridhar Raghavan, Joseph Picone, and Joel Berg. 2006. Detecting fatigue from voice using speech recognition. In *2006 IEEE International Symposium on signal processing and information technology*. IEEE, 567–571.

[11] Jing Han, Kun Qian, Meishu Song, Zijiang Yang, Zhao Ren, Shuo Liu, Juan Liu, Huaiyuan Zheng, Wei Ji, Tomoya Koike, et al. 2020. An early study on intelligent analysis of speech under COVID-19: Severity, sleep quality, fatigue, and anxiety. *arXiv preprint arXiv:2005.00096* (2020).

[12] Susanne M Jaeggi, Ria Seewer, Arto C Nirkko, Doris Eckstein, Gerhard Schroth, Rudolf Groner, and Klemens Gutbrod. 2003. Does excessive memory load attenuate activation in the prefrontal cortex? Load-dependent processing in single and dual tasks: functional magnetic resonance imaging study. *NeuroImage* 19, 2 (2003), 210–225.

[13] Ashish Jaiswal, Ashwin Ramesh Babu, Mohammad Zaki Zadeh, Glenn Wylie, and Fillia Makedon. 2023. Detecting Cognitive Fatigue in Subjects with Traumatic Brain Injury from FMRI Scans Using Self-Supervised Learning. In *Proceedings of the 16th International Conference on PErvasive Technologies Related to Assistive Environments*. 83–90.

[14] Ashish Jaiswal, Mohammad Zaki Zadeh, Aref Hebri, Ashwin Ramesh Babu, and Fillia Makedon. 2023. A Smart Sensor Suit (SSS) to Assess Cognitive and Physical Fatigue with Machine Learning. In *International Conference on Human-Computer Interaction*. Springer, 120–134.

[15] Enamul Karim, Hamza Reza Pavel, Ashish Jaiswal, Mohammad Zaki Zadeh, Michail Theofanidis, Glenn Wylie, and Fillia Makedon. 2023. An EEG-based Cognitive Fatigue Detection System. In *Proceedings of the 16th International Conference on PErvasive Technologies Related to Assistive Environments*. 131–136.

[16] H. I. Krebs, N. Hogan, M. L. Aisen, and B. T. Volpe. 2008. Robot-aided neurorehabilitation. *IEEE transactions on rehabilitation engineering* 6, 1 (2008), 75–87.

[17] Hongyu Luo, Pierre-Alexandre Lee, Ieuan Clay, Martin Jaggi, and Valeria De Luca. 2020. Assessment of fatigue using wearable sensors: a pilot study. *Digital biomarkers* 4, Suppl 1 (2020), 59–72.

[18] Dominique Makowski, Tam Pham, Zen Juen Lau, Jan Brammer, François Lespinasse, Hung Pham, Christopher Schölzel, and SH Chen. 2021. NeuroKit2: A Python toolbox for neurophysiological signal processing. *Behavior Research Methods* 53 (02 2021). https://doi.org/10.3758/s13428-020-01516-y

[19] MUSE. [n. d.]. MUSE S - The Next Generation of MUSE S. "https://choosemuse.com/muse-s/"

[20] Richard Nelesen, Yasmin Dar, KaMala Thomas, and Joel E Dimsdale. 2008. The relationship between fatigue and cardiac functioning. *Archives of internal medicine* 168, 9 (2008), 943–949.

[21] Ellinor Nordin, Erik Rosendahl, and Lillemor Lundin-Olsson. 2006. Timed "Up & Go" test: reliability in older people dependent in activities of daily living—focus

on cognitive state. *Physical therapy* 86, 5 (2006), 646–655.

[22] Adrian M Owen, Kathryn M McMillan, Angela R Laird, and Ed Bullmore. 2005. N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human brain mapping* 25, 1 (2005), 46–59.

[23] Jiapu Pan and Willis J Tompkins. 1985. A real-time QRS detection algorithm. *IEEE transactions on biomedical engineering* (1985), 230–236.

[24] David Richley. 2013. New training and qualifications in electrocardiography. *British Journal of Cardiac Nursing* 8, 1 (2013), 38–42.

[25] Sakar Rijal, Ludovica Corona, M Scott Perry, Eleonora Tamilia, Joseph R Madsen, Scellig SD Stone, Jeffrey Bolton, Phillip L Pearl, and Christos Papadelis. 2023. Functional connectivity discriminates epileptogenic states and predicts surgical outcome in children with drug resistant epilepsy. *Scientific Reports* 13, 1 (2023), 9622.

[26] Robin Tibor Schirrmeister, Jost Tobias Springenberg, Lukas Dominique Josef Fiederer, Martin Glasstetter, Katharina Eggensperger, Michael Tangermann, Frank Hutter, Wolfram Burgard, and Tonio Ball. 2017. Deep learning with convolutional neural networks for EEG decoding and visualization. *Human brain mapping* 38, 11 (2017), 5391–5420.

[27] Frank N Wilson, Franklin D Johnston, Francis F Rosenbaum, and Paul S Barker. 1946. On Einthoven's triangle, the theory of unipolar electrocardiographic leads, and the interpretation of the precordial electrocardiogram. (1946).