

# PulseRide: A Robotic Wheelchair for Personalized Exertion Control with Human-in-the-Loop Reinforcement Learning

Azizul Zahid  
azahid@vols.utk.edu  
University of Tennessee Knoxville  
Knoxville, Tennessee, USA

Jason Scott  
jlsconfig@utk.edu  
University of Tennessee Knoxville  
Knoxville, Tennessee, USA

Bibek Poudel  
bpoudel3@vols.utk.edu  
University of Tennessee Knoxville  
Knoxville, Tennessee, USA

Scott Crouter  
scrouter@tennessee.edu  
University of Tennessee Knoxville  
Knoxville, Tennessee, USA

Danny Scott  
dscott57@vols.utk.edu  
University of Tennessee Knoxville  
Knoxville, Tennessee, USA

Weizi Li  
weizili@utk.edu  
University of Tennessee Knoxville  
Knoxville, Tennessee, USA

Sai Swaminathan  
sai@utk.edu  
University of Tennessee Knoxville  
Knoxville, Tennessee, USA

## ABSTRACT

Maintaining an active lifestyle is vital for quality of life, yet challenging for wheelchair users. For instance, powered wheelchairs face increasing risks of obesity and deconditioning due to inactivity. Conversely, manual wheelchair users, who propel the wheelchair by pushing the wheelchair’s handrims, often face upper extremity injuries from repetitive motions, with shoulder pain affecting 42-66%. These challenges underscore the need for a mobility system that promotes activity while minimizing injury risk. Maintaining optimal exertion during wheelchair use enhances health benefits and engagement, yet the variations in individual physiological responses complicate exertion optimization. To address this, we introduce PulseRide, a novel wheelchair system that provides personalized assistance based on each user’s physiological responses, helping them maintain their physical exertion goals. Unlike conventional assistive systems focused on obstacle avoidance and navigation, PulseRide integrates real-time physiological data—such as heart rate and ECG—with wheelchair speed to deliver adaptive assistance. Using a human-in-the-loop reinforcement learning approach with Deep Q-Network algorithm (DQN), the system adjusts push assistance to keep users within a moderate activity range without under- or over-exertion. We conducted preliminary tests with 10 users on various terrains, including carpet and slate, to assess PulseRide’s effectiveness. Our findings show that, for individual users, PulseRide maintains heart rates within the moderate activity zone as much as 71.7% longer than manual wheelchairs. Among all users, we observed an average reduction in muscle contractions of 41.86%, delaying fatigue onset and enhancing overall comfort and engagement. These results indicate that PulseRide offers a healthier, adaptive mobility solution, bridging the gap between passive and physically taxing mobility options.

## CCS CONCEPTS

• **Human-centered computing** → **Accessibility technologies.**

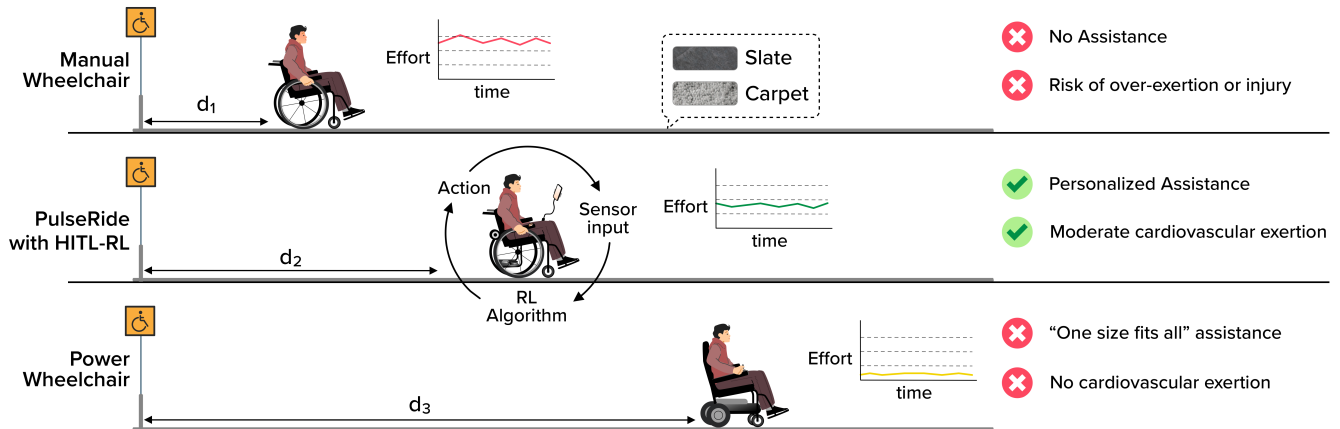
## KEYWORDS

Assistive Robotics, Reinforcement Learning

## 1 INTRODUCTION

Global demand for assistive technologies is substantial and growing. The World Health Organization (WHO) estimates that 2.5 billion people need one or more assistive products, a figure projected to reach 3.5 billion by 2050 [47]. A significant portion of this need is related to mobility impairment, requiring assistive devices like canes, walkers, and wheelchairs [61]. In particular, 80 million people rely on wheelchairs for daily mobility [71], essential for both routine tasks (e.g., moving around the home) and broader social participation. While standard manual wheelchairs provide essential mobility assistance, they rely on user effort for propulsion. Carpets, rough terrain, and even small inclines or slopes increase the effort and energy cost of mobility [30]. A growing body of evidence suggests that the repetitive stress of manual propulsion has serious health consequences. The continuous repetitive motion often leads to upper extremity injuries, with shoulder pain reported in 42% [16] to 66% of users [22] and common injuries to the rotator cuff muscles [3]. Additionally, bilateral carpal tunnel syndrome is frequently associated with handrim manipulation and engagement [6]. These injuries could be mitigated by reducing push frequency to decrease the repetitive strain associated with long-term wheelchair use [8].

Power wheelchairs offer an alternative by providing motorized assistance, alleviating the physical strain of manual propulsion. However, they often provide fixed levels of support focused solely on mobility, addressing the basic requirement of independent movement from point A to B, but generally overlook the user’s need for physical activity and its cardiovascular benefits [5, 18]. Physical inactivity disproportionately affects wheelchair users, contributing to obesity, deconditioning, and further health decline [13]. Additionally, power wheelchair users may experience limitations in community participation [24], as reduced physical activity can hinder engagement in social settings, further impacting their well-being. A third option has emerged to address the limitations of conventional manual and power mobility: power assist wheels. For



**Figure 1: Comparison of wheelchair assistance paradigms. Manual wheelchair (top) provides no assistance and requires maximum user effort, leading to early onset of fatigue and risk of over-exertion. Power wheelchair (bottom) offers fixed assistance but discourages physical activity i.e., complete reliance on motorized propulsion. PulseRide (middle) provides personalized adaptive assistance, balancing user’s physical effort (to maintain moderate cardiovascular exertion) and mobility needs. For the same effort budget, as PulseRide delays the onset of fatigue, users are able to cover more distance ( $d_2 > d_1$ ).**

instance, systems such as SmartDrive [1], Yomper [2] and e-motion [4] have shown to allow users to stroke the wheelchair pushrims to activate small, lightweight motors that drive the wheels for a brief period of time (seconds). Users must continue to stroke the pushrim as they would if they were propelling a conventional manual chair. This method requires less strength and endurance than manual wheeling and is especially helpful on inclines, uneven terrain, and carpeted surfaces. However, most power-assist systems still operate with a “one-size-fits-all” approach to assistance, lacking automatic adaptation to varying environmental conditions (e.g., slopes, ramps, gravel) and overlooking the variability in individual user capabilities and contexts. Assistance levels are typically adjusted manually using a knob with preset velocity settings, making these systems less responsive to real-time user and environment needs.

These limitations highlight a critical gap in assistive technology that could dynamically personalize assistance to match each user’s unique ability [70], physiological state and environmental context. Additionally, wheelchair users often have limited opportunities for physical activity, with manual propulsion remaining one of the few accessible forms of exercise [14]. Integrating physical activity into daily routines, beyond structured leisure time, is increasingly recommended to promote holistic health benefits [15]. Therefore, determining the appropriate level and timing of automatic assistance to meet both mobility and health goals remains an open question. Additionally, understanding how an autonomous system could achieve physical activity goals as part of daily mobility for wheelchair users is still underexplored. In response, we propose a new assistance paradigm for wheelchair users that integrates real-time physiological feedback with autonomous control to deliver personalized, exertion-aware assistance. As shown in Figure 1, this approach provides a balanced solution between manual wheelchairs that require maximum user effort and power wheelchairs that discourage physical activity. Our research focuses on the following questions:

- **RQ1:** How can we design an autonomous wheelchair system that promotes physical activity while supporting users’ daily mobility?
- **RQ2:** How can we personalize assistance levels by dynamically adapting to users’ physiological states?
- **RQ3:** How can we ensure that this adaptive assistance is effective across diverse environments?

To address these research questions, we designed a novel, autonomous wheelchair system called PulseRide. PulseRide utilizes Human-in-the-Loop Reinforcement Learning (HITL-RL), integrating real-time data on users’ physiological state including heart rate and electrocardiogram (ECG) signals, and wheelchair movement (speed) to dynamically adjust the timing and duration of automatic push assistance. This approach enables PulseRide to continuously balance physical support based on user exertion, adapting in real time to match individual needs and environmental conditions. To the best of our knowledge, PulseRide is the first wheelchair system to offer personalized, adaptive assistance that balances users’ physical effort with their mobility needs as a part of their daily mobility. Through a human subject study with 10 participants (diverse in sex, age, and body mass index) across two indoor surfaces (slate and carpet), we demonstrate the ability of PulseRide to improve: (1) time spent in the moderate heart rate zone by up to 71.7% over manual wheelchair use, (2) reduction of muscle contraction by 41.86% and 24.94% across multiple environments (slate and carpet), and (3) fatigue resistance, delaying onset of fatigue. These results demonstrate that PulseRide successfully achieves its core objectives of providing a personalized, exertion-aware daily mobility solution that promotes physical health.

## 2 BACKGROUND AND RELATED WORK

### 2.1 Physical Activity and Wheelchair Users

Current research strongly supports the beneficial effects of physical activity (PA) on human health [37] and well-being [9, 27, 34, 55].

Physical activity has been shown to help prevent chronic conditions such as cardiovascular disease, diabetes, cancer, hypertension, obesity, depression, and osteoporosis [7, 10, 33, 48, 68, 69]. For wheelchair users in particular, PA is especially valuable for maintaining or improving physical capacity [32, 42], contributing to an increased quality of life [29, 31]. Both the Centers for Disease Control and Prevention and the American College of Sports Medicine recommend a minimum of 30 minutes of moderate-intensity activity daily [49] to achieve health benefits such as decreased depression and anxiety [45, 66]. Another guideline suggests 150 minutes of moderate-intensity activity per week [43]. These recommendations are particularly important for wheelchair users, as they help reduce the risk of secondary health problems that could further limit functional independence [49]. However, despite these clear benefits, physical inactivity is especially prevalent among wheelchair users [29, 45, 64]. Many wheelchair users struggle to engage in the recommended levels of health-promoting exercise [11, 21]. For wheelchair users, opportunities to be physically active are often limited, with manual wheelchair propulsion remaining one of the few accessible forms of PA and exercise [14]. To address this gap, researchers have suggested integrating PA into daily routines [55, 66], moving beyond traditional structured leisure-time exercise.

This shift has led to a focus on "Lifestyle Physical Activity" (LPA) [15], which emphasizes incorporating moderate-to-vigorous PA into everyday life rather than relying on structured exercise alone [50]. LPA involves self-selected activities—including leisure, occupational, and household tasks—that meet at least moderate intensity levels and can be part of daily routines, whether planned or unplanned [17]. However, LPA has not yet been clearly defined for wheelchair users, and personalized PA schemes for this group remain underdeveloped. Moreover, measuring PA in wheelchair users has proven challenging due to variability among individuals and the complexity of assessing their PA levels. In addressing these measurement challenges, researchers have explored various approaches. Coulter et al. [14] used tri-axial accelerometers mounted on wheelchairs to measure physical activities through motion, a method that has become common in the literature [23, 44]. However, these devices do not directly capture the user's actual activity levels through on-body sensors, which could provide more detailed and specific information about the user's activity.

In this work, we aim to bridge the gap between existing physical activity recommendations and their practical implementation for wheelchair users by developing a system that integrates real-time physiological feedback with adaptive mobility assistance. While prior studies have explored activity measurement using wheelchair-mounted accelerometers [14, 23, 44], these approaches focus primarily on motion-based metrics and lack real-time, personalized adjustments based on user-specific physiological states. Our work addresses this limitation by incorporating data such as heart rate and ECG to provide personalized, exertion-aware assistance. By combining physiological sensing with Human-in-the-Loop Reinforcement Learning (HITL-RL), we enable dynamic adaptation of wheelchair push assistance to meet individual user needs. This personalized approach not only encourages sustained physical activity but also supports users in achieving moderate exertion levels.

## 2.2 Human-in-the-Loop Reinforcement Learning for Smart Wheelchair

Traditional assistance methods often rely on predetermined models [20] or fixed thresholds [67], which struggle to capture the dynamic nature of user fatigue and recovery patterns [38]. While current AI approaches excel at analyzing health data, they primarily focus on monitoring and prediction rather than real-time adaptation of assistance. In this context, Reinforcement Learning (RL) offers a promising solution for real-time adaptation, though its application in smart wheelchairs has primarily focused on autonomous navigation tasks [12, 39, 58, 59]. Goli et al. [25] demonstrated the importance of human feedback in optimizing navigation assistance, yet the integration of physiological indicators within RL frameworks for physical activity support remains largely unexplored.

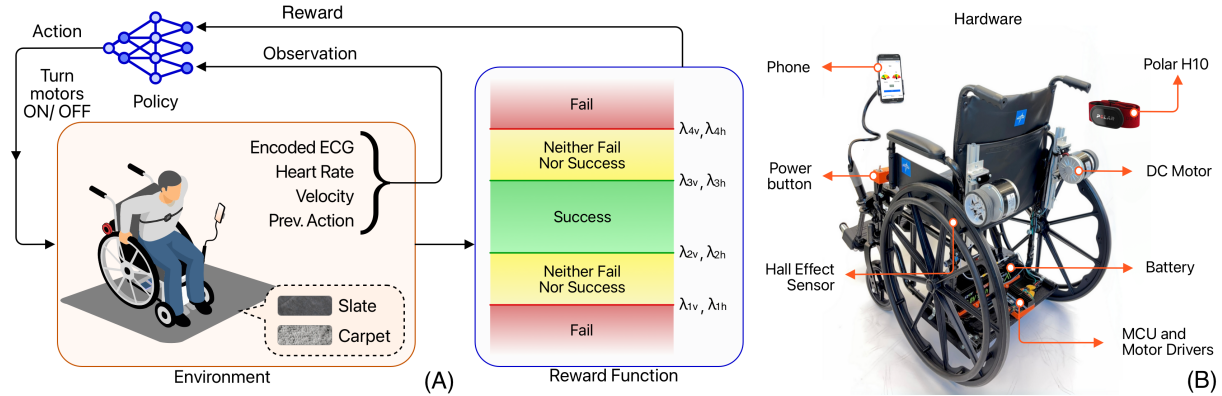
Human-in-the-Loop Reinforcement Learning (HITL-RL) extends traditional RL by incorporating explicit human input into the learning process. While conventional RL systems learn through environmental rewards alone, HITL-RL leverages direct human input to shape the policy, particularly valuable when the reward function cannot fully capture complex human preferences and comfort levels [19, 65]. This approach is especially relevant for wheelchair assistance, where user comfort, fatigue levels, and physical activity goals must be balanced dynamically. The integration of human feedback in HITL-RL offers two key advantages for wheelchair assistance systems. First, it enables the system to learn from both immediate physiological responses (like heart rate and ECG) and explicit user preferences, creating more personalized assistance policies. Second, human guidance helps address the exploration-exploitation challenges in RL [40], allowing safer and more efficient learning in real-world scenarios. This framework provides the foundation for adaptive assistance that can simultaneously support both mobility requirements and physical activity goals, while maintaining the personalization benefits of AI-driven approaches.

## 3 SYSTEM DESIGN

In this section, we provide an overview of PulseRide's design and a walkthrough of its operation. We then describe how Human-in-the-Loop Reinforcement Learning (HITL-RL) integrates real-time data on users' physiological state and wheelchair movement to dynamically adjust the level of automatic push assistance. Finally detailing other the technical aspects such neural network model architectures.

### 3.1 PulseRideHardware and Operation

PulseRide is designed using a standard off-the-shelf manual wheelchair from Medline (Model Drive Medical SSP118FA-SF). We motorized the wheelchair with two 250W 24V DC brushed motors powered by two 12V lithium batteries, mounted on aluminum extrusions behind the wheelchair (Figure 2B). The motors are connected to the wheelchair tires via grip wheels, allowing for powered assistance when needed. We chose brushed DC motors over stepper motors due to their low resistance when turned off, which enables users to manually propel the wheelchair without additional resistance. For monitoring and sensing, PulseRide employs an array of sensors. To measure human physiological responses, we use a Polar H10 sensor for ECG and heart rate. To track wheelchair velocity,



**Figure 2: (A) PulseRide’s reinforcement learning framework, using inputs such as encoded ECG, heart rate, velocity, and previous action to adapt motor activation via a reward function that encourages moderate exertion. (B) PulseRide hardware setup, including the Polar H10 heart rate sensor, phone interface, DC motors, hall effect sensor, motor drivers, MCU, power button, and battery.**

we installed a hall effect sensor along with 18 magnets on the wheel rims for RPM calculation, with each increment representing a rotation of 20 degrees ( $360^\circ/18$ ). The system supports a maximum load of 113 kg and includes a speed limit control that caps the maximum speed at 2.5 m/s to ensure safety. To further enhance safety, we incorporated a power cut-off switch near the armrest and a remote kill switch, allowing for immediate shutdown if necessary. The total cost of the attachable hardware, including the manual wheelchair, is under \$500, making it an affordable solution.

Upon starting the PulseRide system, the user connects to the wheelchair via their mobile phone, where they are presented with an interface showing important safety information. Next, before fully operating PulseRide, the user completes a three-step setup process. First is a pretraining phase, where the user propels the wheelchair for 135 seconds at three perceived exertion levels (low, medium, and high), specific to them. During the pre-training phase, PulseRide calculates user-specific heart rate and velocity thresholds, which are represented as activity zones to guide the user. These zones are visualized with green representing the moderate activity zone, yellow representing transition zones (between moderate and extreme), and red representing both low and high extreme activity zones (Figure 2A). The user then pushes the wheelchair for 20 minutes across these zones as instructed by our AI algorithm, receiving prompts on the mobile interface, such as “go slow” or “go fast,” as directed. After completing this AI-guided training, the user enters the testing mode and can operate PulseRide with adaptive assistance.

During the training phase, PulseRide uses representation learning through neural networks in two distinct stages: ECG signal encoding and the RL policy as shown in Figure 3. The ECG encoder, trained to classify activity levels based on ECG input (shown in Figure 3 (A)), learns to extract meaningful physiological patterns from the signals. HITL-RL then processes these encoded patterns along with heart rate and velocity data to determine when and how long to provide assistance to maintain moderate exertion levels. This approach allows PulseRide to automatically identify relevant patterns in user behavior and physiological responses that impact their ability to propel the wheelchair. Notably, there is no explicit

programming of which features PulseRide should analyze; potential factors include fatigue patterns, muscular strength, and heart rate response.

### 3.2 Reinforcement Learning Approach

We now go over the artificial intelligence (AI) algorithm used by PulseRide for providing dynamic assistance. As introduced in Section 2.2, we are using a variation of Reinforcement Learning called Human-in-the-Loop Reinforcement Learning (HITL-RL).

We formalize our approach using a Markov Decision Process, represented as the tuple  $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$ . Here,  $\mathcal{S}$  denotes the set of states, capturing all possible configurations of the environment and user;  $\mathcal{A}$  is the set of actions the agent can take;  $\mathcal{P}(s', r|s, a)$  defines the environment dynamics, representing the probability of transitioning to state  $s'$  and receiving reward  $r$  after taking action  $a$  in state  $s$ ;  $\mathcal{R}(s, a)$  is the reward function, which assigns a numerical value to each state-action pair based on the desirability of that action in that state; and  $\gamma$  is the discount factor that prioritizes immediate rewards over distant future rewards. The agent with a policy  $\pi_\omega$ , parameterized by  $\omega$ , aims to maximize the cumulative discounted rewards. The agent’s actions are determined based on the Q-value function, which is iteratively updated according to the equation:

$$Q_{(s,a)} = (1 - \alpha) Q_{(s,a)} + \alpha \left[ R_{(s,a)} + \gamma \max_{a'} Q_{(s',a')} \right], \quad (1)$$

where  $\alpha$  represents the learning rate and  $a'$  denotes the successor action of  $a$ . In scenarios with large state spaces, function approximation techniques, such as deep neural networks are preferred due to their ability to generalize and ensure convergence. We implement Deep Q Networks (DQN) [41] as our RL algorithm, incorporating two crucial mechanisms: experience replay and target network. Experience replay maintains a memory buffer of observed state-action-reward transitions (size=120), from which we uniformly sample batches during training. This approach reduces variance in the learning process and addresses the non-stationarity of data distribution during policy updates. The target network serves as a separate Q-value estimator whose parameters are updated every

timestep by 5% ( $\tau$ ) from the learned policy network, stabilizing the training process and preventing divergence of the policy learned by the agent.

During training, the policy learns the timing (when to assist) and duration (for how long) while relying on the user for stopping and directional control. As designed, the PulseRide policy cannot apply brakes or reverse movement. Because human responses are introduced into the control loop, an uncertain learning environment is created due to the inherent variability of such responses (e.g., when the policy reduces assistance by applying action 0 over multiple time steps, expecting increased user effort, the user might not respond as anticipated). This uncertainty introduces additional complexity to the exploration process, which is crucial for successful policy learning, as during training the policy needs to be sufficiently exposed to diverse state-action pairs to learn optimal behaviors [62]. To systematically address these challenges, we implement a two-part exploration strategy during training that combines algorithmic exploration with structured human participation. At the algorithmic level, we employ an  $\epsilon$ -greedy approach with exponential decay which produces noisy actions from the DQN algorithm. At the human level, we implement a structured training protocol where users deliberately introduce periods of varying activity levels to expose the policy to different physiological states. Importantly, users are not required to follow any such protocol after obtaining a trained policy. To further stabilize the training process, we employ smooth L1 loss and gradient clipping to prevent large policy updates.

### 3.3 Electrocardiogram Encoder

We encode the Electrocardiogram (ECG) signals using a trained Convolutional Neural Network (CNN) before passing them as input to the RL agent (Figure 3 (B)). To leverage transfer learning, we first train the CNN on an adjunct supervised learning task of classifying ECG intervals into Low, Medium, and High activity levels (Figure 3 A). This training allows the encoder to learn general ECG features that are relevant for the RL task, even though the RL agent doesn't explicitly perform activity classification. After training, we freeze the CNN's weights and use it to transform three-second ECG intervals of length 390 (sampled at 130Hz) into concise vectors of eight dimensions, as shown in Figure 3 (B). This CNN encoder serves multiple purposes in our approach. First, it significantly reduces the dimensionality of the observation space for the RL agent, which reduces the computational complexity of the control task. Second, it leverages the CNN's capacity to extract relevant features from time-series data, identifying patterns in the ECG that indicate relevant physiological features. Third, by encoding the previous three-second interval, it provides temporal context to capture meaningful patterns in the signal. By encoding the ECG signal into a vector of size eight, we strike a balance between providing the RL agent with the rich, high-fidelity ECG signal of size 390 and the simple, low-fidelity heart rate measurement of size one. This approach provides more nuanced physiological information than heart rate alone, while avoiding the computational complexity of processing entire raw ECG signals. Hence, to train the encoder, we collected data from four participants separate from our main participant pool and split the dataset into 80% training and 20% validation sets. After 500 epochs of training, the CNN encoder achieves 100% training

accuracy and 87.5% validation accuracy. We then remove the final MLP head from the model, allowing it to encode ECG signals into eight-dimensional vectors that capture meaningful physiological features. Table 1 provides a comprehensive list of hyperparameters used in the CNN encoder training.

Model	Hyper-parameter	Value
CNN Encoder	Batch size	16
	Learning rate	0.0001
	Number of epochs	500
	Optimizer	Adam
	CNN Kernel size	5
	Number of training samples	303
	Loss function	Cross-Entropy
HITL-RL	Replay buffer size	120
	Batch size	32
	Discount factor ( $\gamma$ )	0.99
	Optimizer	Adam
	Exploration rate (initial $\epsilon$ )	0.99
	Exploration rate (final $\epsilon$ )	0.05
	Activation function	ReLU
	Optimizer	Adam [35]
	Gradient clipping	$[-100, 100]$
	Target Network update ratio ( $\tau$ )	0.005
	Number of training episodes	30

**Table 1: Training parameters and hyper-parameters for CNN Encoder and HITL-RL.**

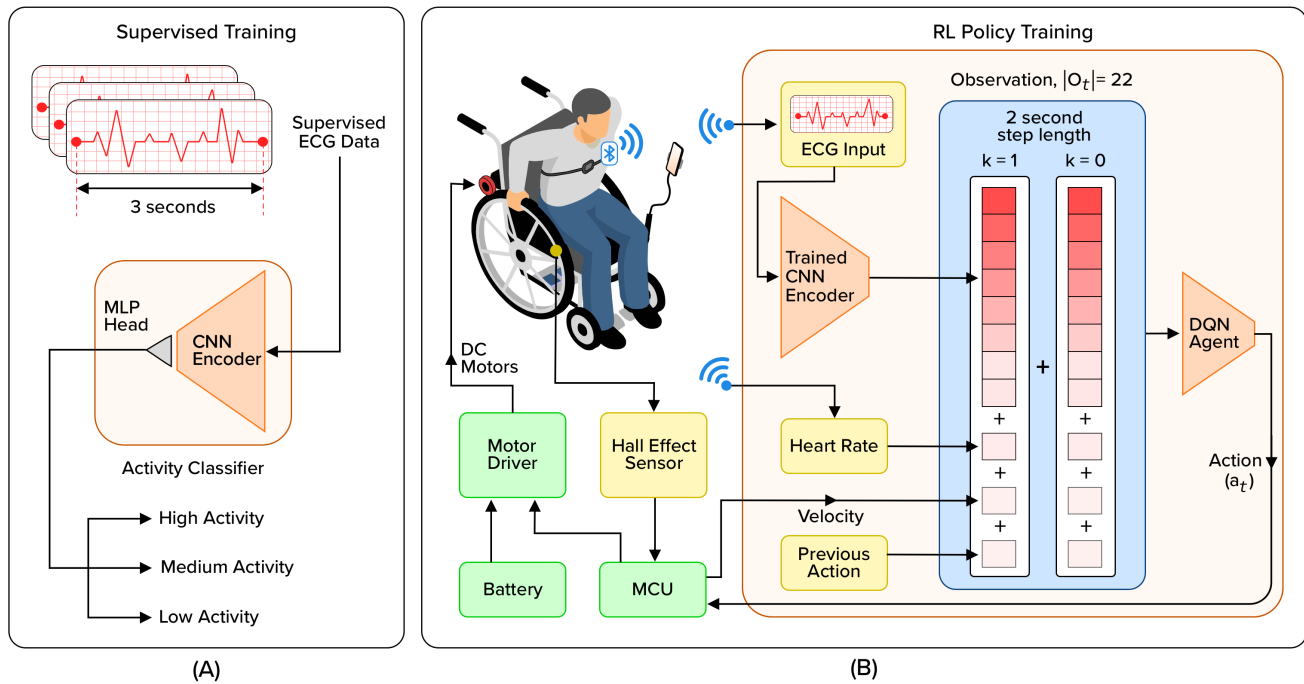
### 3.4 Markov Decision Process

The main components of our Markov decision process are outlined below:

**Observation:** To sufficiently capture the dynamic behavior of the system, we represent observations with heart rate, velocity, previous action, and the output of the trained CNN encoder. Further, to capture transitions across multiple steps, we stack information from the two previous seconds to compose an observation in each time-step. This 2-second time-step (Figure 3 (B)) enables the agent to consider recent historical data when making decisions.

**Action:** The actions are binary i.e., either turn the two motors attached to the wheels ON (pushing PulseRide forward) or OFF.

**Reward function:** To facilitate rapid learning, we employ a hybrid approach that incentivizes high-level desired outcomes (such as maintaining heart rate below the user's perceived over-exertion range) while simultaneously discouraging undesired actions (such as turning the motor ON when the system is already at high velocity). The complete reward function, shown in Algorithm 1, incorporates user-specific empirical thresholds ( $\lambda$  values) obtained during the pre-training phase, where users follow a specific protocol (defined in Section 4.2) of pushing PulseRide at their own perceived low, medium, and high intensities. This approach ensures personalized adaptation to each individual's capabilities and limitations. Specifically, the function utilizes thresholds for both velocity ( $\lambda_{1v}, \lambda_{2v}, \lambda_{3v}$ ) and heart rate ( $\lambda_{1h}, \lambda_{2h}, \lambda_{3h}$ ) to define boundaries, as shown in Figure 2, for acceptable ('success' zone shown in green and 'neither success nor failure' zone shown in yellow) and unacceptable ('failure' zone shown in red) ranges. These thresholds directly customize the system's behavior based on each user's physiological responses. The function uses the coefficient of variation of velocity, defined as  $vel_{cv} = \frac{\mu_{\text{pretraining velocity}}}{\sigma_{\text{pretraining velocity}}}$ , and calculates distances



**Figure 3: (A) Activity classifier consisting of a Convolutional Neural Network (CNN) encoder and a Multi-Layered Perceptron (MLP) head, trained on a supervised dataset to classify ECG signals into three activity classes (High, Medium, and Low). (B) The realtime, high-fidelity ECG signals are input to the trained CNN Encoder (frozen during RL policy training) which feeds the encoded output into the observations of the RL pipeline. Additionally, the heart rate from the Polar heart rate monitor, the velocity from the hall effect sensor and the previous actions are also added as observations. The DQN Agent passes its actions to the wheelchair microcontroller unit (MCU) corresponding to either turn ON or turns OFF the motor.**

from threshold values for both velocity and heart rate at every timestep during training. These calculations form the basis for determining the appropriate reward or penalty in various scenarios. Similar reward functions are commonly used in DC motor control problems [52].

### 3.5 Model Architectures

The Deep Q Network (DQN) policy network and encoder are both Convolutional Neural Networks (CNNs), which are effective in processing physiological inputs [51, 54]. The encoder architecture contains three ConvNormPool blocks that progressively process the single-channel input through 256, 128, and 64 channels. Within each ConvNormPool block, a residual connection adds the output of the first convolutional layer to the output of the third convolutional layer before normalization and Swish activation. This skip connection enhances feature preservation and facilitates better gradient flow during training [28]. After each block, a max pooling layer with kernel size 2 halves the dimensions of the feature maps. The output of three ConvNormPool blocks is passed through an adaptive average pooling layer and flattened before being passed through a fully connected layer that maps the 64 input features to eight features followed by a final linear layer (MLP head) maps these eight features to three output classes. Whereas in the HITL-RL policy network, the CNN architecture is designed to process physiological data by expressing the locality in time as locality in space (by stacking two consecutive timestep inputs), enabling the

network to detect patterns in the temporal transitions of the data. The network takes consecutive velocities, voltages, encoded ECG signals, and heart rates as input. The policy network consists of two convolutional layers that expand the single input channel first to 8 and then to 16 channels, each followed by ReLU activation. The convolutional feature maps are then flattened and passed through three fully connected layers that progressively reduce the dimensionality from  $832 \rightarrow 128 \rightarrow 32$  neurons with ReLU activation after each layer. The final output layer produces two values corresponding to the Q-values for each action. The encoder and policy network have approximately one million and 110,000 total trainable parameters respectively.

## 4 METHODOLOGY

We now go over our experimental validation of the PulseRide system, designed to evaluate its effectiveness in providing personalized, adaptive push assistance.

### 4.1 Participants

We recruited 10 participants without disabilities, aged 19-30 ( $M = 22.8$ ,  $SD = 3.65$ ), seven participants identified as male and three as female. Participants had an average height of 67.3 inches ( $SD 4.92$ ) and an average BMI of 23.63 ( $SD 4.27$ ). Recruitment was conducted through university channels and social media. The inclusion criteria required participants to be physically able of propelling a

**Algorithm 1** Reward Function

---

```

# Calculate distances
vel_dist ← max(0, velocity - λ3v)
if velocity < 0 then
    vel_dist ← max(0, |velocity| - |λ2v|)
end if
heart_dist ← max(0, heart rate - λ3h)
if heart rate < 0 then
    heart_dist ← max(0, |heart rate| - |λ2h|)
end if

# Calculate Reward
reward ← -1 + 0.5 · e-3·heart_dist + 0.5 · e-2·vel_dist
if (velocity > λ1v) or (velocity = vel_cv) or (heart rate < λ2h) or
(heart rate > λ1h) then
    reward ← -1.2 # (failed = True) Terminate the
episode
else if (λ2h < heart rate ≤ λ3h) and (velocity > λ2v) then
    if action = 1 and velocity > λ3v then
        reward ← -1 + e-(1.8·vel_dist+2.5·heart_dist)
    else
        reward ← 0.6 · e-2·heart_dist + 0.6 · e-2·vel_dist
    end if
else if (λ3h < heart rate < λ1h) and (velocity < λ3v) then
    reward ← -1.4 + 0.5 · e-3·heart_dist + 0.5 · e-2·vel_dist
end if

```

---

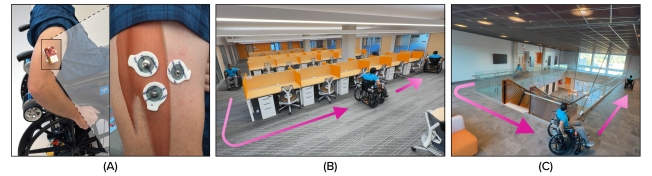
manual wheelchair and providing informed consent. During recruitment, participants completed a questionnaire about their ability to engage in moderate physical activity and any relevant medical advice they had received regarding exercise. We excluded individuals with pre-existing cardiovascular or musculoskeletal conditions, as well as those with cognitive impairments that could prevent them from understanding the study tasks. Selected participants were compensated at a rate of \$20 per hour and the experiments took approximately 1.5 hours. This research protocol was reviewed and approved by the university ethics and review board.

The decision to conduct this study with the participants without disabilities rather than wheelchair users was made after consideration of two factors. Firstly, as this is an initial investigation of PulseRide, we prioritized safety and sought to minimize potential risks to participants. Individuals without disabilities could more easily adapt to and recover from the physical demands of the study, allowing us to gather preliminary data on the system’s effectiveness without exposing vulnerable populations to untested technology. Secondly, using participants without prior wheelchair experience helped us test how well our system adapts to new users. Since all participants started with the same level of wheelchair skills (none), any differences in their performance would likely be due to our system’s effectiveness, not their prior abilities. While we acknowledge the limitations of this approach in terms of immediate real-world applicability, it serves as a crucial first step in its development. Further discussion of this consideration is presented in Section 6.3.

**4.2 Procedure**

In the experiment procedure, participants start by receiving a brief introduction to the system, followed by a hands-on demonstration and a trial run. Then, they rest for a short rest period to establish their baseline heart rate. Next, the human-in-the-loop reinforcement learning (HITL-RL) policy for each participant is trained in two consecutive stages: 1) a pre-training phase, which lasts for two minutes and 15 seconds, and 2) a training phase which takes 20 minutes. The training is followed by test period of five minutes in two different environments.

**4.2.1 Pre-Training:** During the pre-training stage, the RL policy is turned off, and participants complete three consecutive 45-second intervals, pushing the wheelchair at their perceived moderate and transition (between moderate and extreme) activity levels. These intervals serve multiple purposes that enable system personalization. First, the heart rate range (min, max) is determined. Second, heart rate and velocity statistics (mean and standard deviation) used for normalizations are determined. And third, user-specific heart rate and velocity thresholds ( $\lambda$  values used in the reward function of the training phase) that define the boundaries between moderate, transition, and extreme intensity zones are established. In Figure 2, these zones are visualized with green representing moderate activity zone shown as ‘success’, yellow representing transition zones (between moderate and extreme) shown as ‘neither success nor failure’, and red representing both low and high extreme activity zones shown as ‘failure’.



**Figure 4: (A) Placement of electromyography (EMG) sensor electrodes on the lateral head - triceps brachii muscle of the right arm. (B) Indoor testing path on a carpeted floor. Users followed a closed-loop path in both environments with different coefficients of friction. The carpet surface exhibits a higher coefficient of friction than slate, requiring greater propulsion effort from users during wheelchair operation. (C) Indoor testing path on a slate floor.**

**4.2.2 Training:** During training, participants are tasked with maintaining PulseRide velocity within their personalized moderate velocity zone, established during pre-training. The moderate velocity zone is visually represented by the green zone in the center of the user interface gauges. Two real-time analog needles continuously indicate instantaneous heart rate and velocity. As described in Section 3.2, to evaluate system response outside the moderate zones, participants receive prompts in the user interface to alter their speed and activity intensity. The prompt at training episode 15 is ‘Go slower’ and the prompt at training episode 25 is ‘Go faster’. The complete training period comprises of 30 episodes and each episode contains 20 timesteps of 2 seconds duration. Notably, all training sessions were conducted only on a slate surface, with participants following a closed-loop path of approximately 50 meters perimeter (Figure 4).

**4.2.3 Testing:** Following training, each participant underwent a test phase consisting of five episodes (30 timesteps each). Similar to training, participants followed the same closed loop paths but on two different surfaces - slate and carpet flooring (Figures 4B & 4C). The carpeted surface, with its higher coefficient of friction, required more physical effort for wheelchair propulsion compared to the smooth slate surface. Participants were tasked with maintaining their personalized moderate velocity ‘green’ zone established during pre-training. To collect data for validation, EMG sensors were placed on the triceps of participants’ both arms using MyoWare 2.0 muscle sensor development kits, which include an Electromyography (EMG) muscle sensor, a wireless communication shield, and solid gel electrodes. We attached the electrodes along the longitudinal midline of each triceps head, parallel to the muscle fibers, avoiding placement directly on motor points and at the edges of the muscle to prevent crosstalk from adjacent muscles [46] as shown in Figure 4A. The EMG sensor input measurement rate is 20Hz, further, to ensure sufficient data capture, we started the collection process earlier and stopped later than the actual test time. During analysis, the collected data was synchronized with the correct test timing. It is important to note that EMG data was only used for additional validation and analysis during test period and is not a part of the input to the RL agent, hence it is excluded from Section 3.

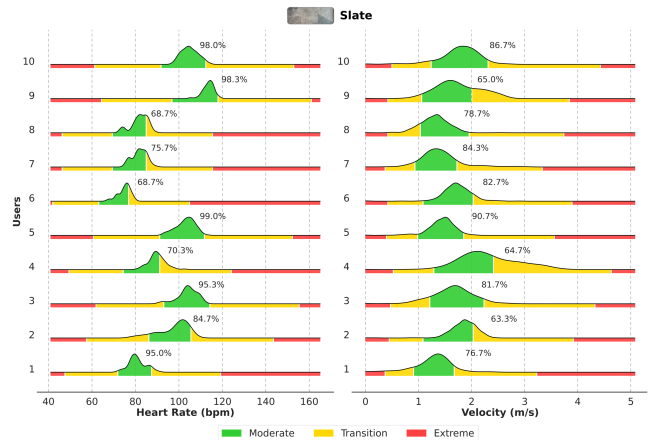
### 4.3 Measures and Analysis

Our experiments investigate PulseRide’s effectiveness across multiple dimensions with comparisons to manual wheelchair use. Specifically, we study four physiological and task outcomes:

**4.3.1 Number of muscle contractions:** We quantify the muscular effort by counting the total number of muscle contractions. A muscle contraction is detected on the raw EMG signal using root mean square (RMS) as a threshold i.e., we detect a contraction event if the raw EMG exceeds the RMS value. Since both push and recovery (pulling back and cocking the arms for another push) movements trigger muscle contractions, they directly correlate with the number of times a wheelchair is pushed.

**4.3.2 User-Specific Personalization:** We evaluate PulseRide’s adaptation to individual users by analyzing their heart rate and velocity distributions across different activity zones (moderate, transition, and extreme). Using Gaussian kernel density estimation, we construct continuous probability density functions of these physiological responses from observed data. The proportion of time spent in each zone, particularly the moderate activity zone, directly quantifies PulseRide’s effectiveness in maintaining personalized activity levels.

**4.3.3 Generalization Across Environments:** We assess PulseRide environmental adaptability by comparing physiological responses across two distinct surfaces: Slate (smooth) and Carpet (rough). The key measure is the consistency of heart rate zones across environments i.e., whether users maintain similar physiological states regardless of surface conditions. This comparison reveals the system’s ability to adapt assistance levels to varying environmental demands while maintaining user comfort.



**Figure 5: Heart rate and velocity distributions across users on Slate environment, estimated using Gaussian kernel on observed data. The x-axis presents each user’s personalized ranges corresponding to moderate, transition and extreme activity levels.**

**4.3.4 Fatigue Reduction:** We evaluate PulseRide’s ability to reduce fatigue by analyzing electromyography (EMG) signals compared to manual wheelchair use. As muscles fatigue, there is a shift in the EMG power spectrum toward lower frequencies, which manifests as a decrease in mean frequency (MNF) [57]. To quantitatively assess fatigue delay, we conducted MNF analysis on raw EMG data by dividing the total testing time of 300 seconds into 30 equal 10 seconds segments. Then for each segment, we determined the mean frequency and calculated the best-fitted approximation of the MNF data across all timeslots, allowing us to track muscle fatigue progression throughout the testing period.

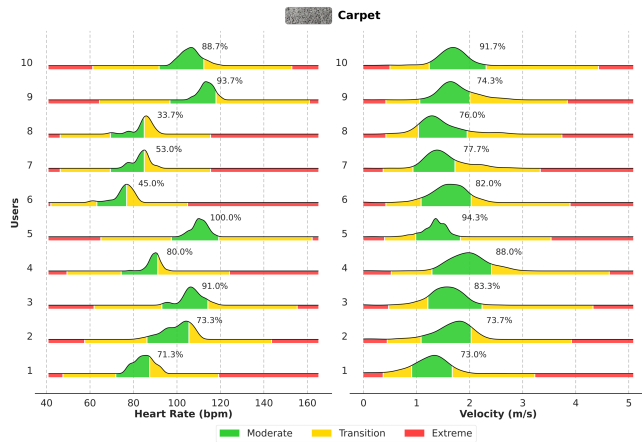
## 5 RESULTS

### 5.1 Personalization across users

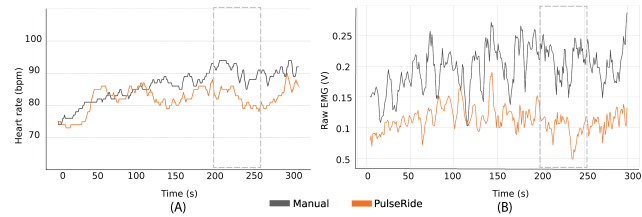
Our results indicate that PulseRide effectively personalizes heart rate activity, helping users maintain heart rates within their individual moderate activity zones across all participants. As shown in Figure 5 and Figure 6, despite each user having a unique moderate heart rate zone, PulseRide consistently enabled users to remain within these personalized zones. However, individual analysis reveals some outliers. For instance, Users 8, 7, and 6 spent only 33.7%, 53%, and 45% of the time in the moderate heart rate zone, respectively, with a higher percentage of time in the transition zone toward the upper extreme heart rate zone. This deviation can be explained by examining the velocity distribution for these users, where they spent more amount of time in the transition zone toward higher velocities. Since PulseRide does not have explicit braking capabilities, users may have been pushing themselves more in these instances.

Additionally, Figures 7A and 7B illustrate a single user’s heart rate and EMG trajectories throughout a test session. These figures show that heart rate and raw EMG values are generally higher (in boxed regions) during manual wheelchair use than with PulseRide, indicating that manual propulsion requires greater muscular exertion.





**Figure 6: Heart rate and velocity distributions across users on Carpet environment, estimated using Gaussian kernel on observed data. The x-axis presents each user’s personalized ranges corresponding to moderate, transition and extreme activity levels.**

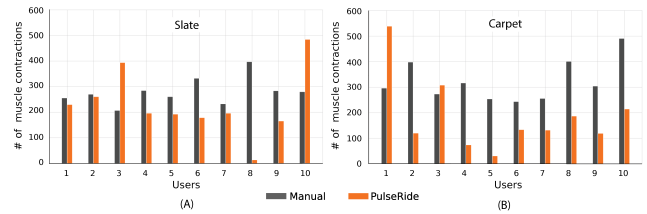


**Figure 7: A. Heart rate trajectory for one user throughout the test experiment on Slate. B. EMG trajectory for one user throughout the test experiment on Slate.**

### 5.2 Muscle Contraction Analysis: Impact of Smart Assistance on Propulsion Efficiency

Our muscle contraction analysis, shown in Figure 8 A and B, presents the number of contractions across users and reveals a consistent reduction in muscle effort required to propel PulseRide compared to the manual wheelchair condition in both environments (slate and carpet). Across users, PulseRide decreased muscle contractions by an average of 41.86% on slate and 24.94% on carpet compared to the manual wheelchair. While these reductions indicate that PulseRide generally lowers the effort needed for propulsion, some outliers emerged in our results. For example, User 3 and User 10 on slate, and User 1 and User 3 on carpet, did not experience gains in muscle contraction reduction.

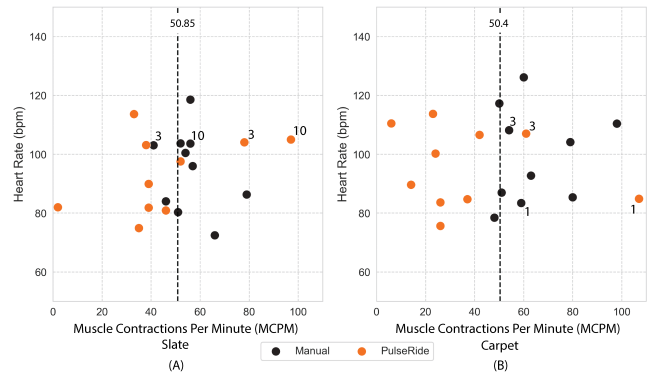
To further understand these outliers, we analyzed muscle contractions per minute relative to heart rate across users. As shown in Figure 9 A & B, User 3 in the slate environment exhibited higher muscle contractions with PulseRide than with the manual wheelchair. However, the time this user spent in the moderate heart rate activity zone with PulseRide (95.3%) was similar to the manual condition (94%). This suggests that, while muscle contractions did not decrease for certain outliers like Users 3, 10, and 1 in the PulseRide condition, the intelligent assistance provided by PulseRide was



**Figure 8: Total number of muscle contractions for each user for manual and PulseRide devices A. on Slate B. on Carpet.**

deliberate in maintaining a stable heart rate within the moderate activity zone rather than focusing solely on reducing pushes.

Finally, Figure 9 A and B highlight distinct clusters for both PulseRide and manual conditions, showing that muscle contractions per minute below 50.85% are associated with PulseRide in the slate environment and below 50.4% in the carpet environment. This clustering further supports the effectiveness of PulseRide in reducing muscle effort across the majority of users.



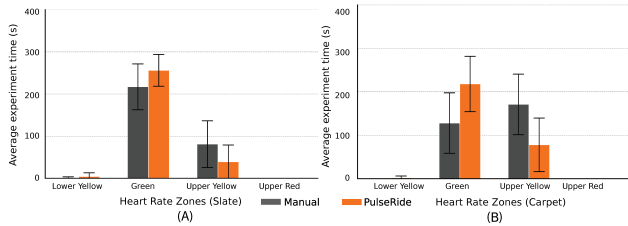
**Figure 9: A total number of user muscle contractions per minute for each user for manual and PulseRide wheelchair on the slate surface (shown in A) and carpet surface (shown in B)**

### 5.3 Generalization across environments

Our results show that across different environments—both slate and carpet—PulseRide maintained participants’ heart rates within the moderate activity zone. In the slate environment (Figure 10A), participants’ average heart rates were in the Green zone 85.6% of the time and in the Upper Yellow zone 13.1% of the time. Similarly, in the carpet environment (Figure 10B), heart rate consistency was preserved, with approximately 73.15% in the Green zone and 26.2% in the Upper Yellow zone. This highlights PulseRide’s adaptability, as it was trained only in the slate environment yet generalized effectively to the carpet environment.

In contrast, the manual wheelchair showed greater inconsistency across environments. In the slate condition (Figure 10A), participants’ heart rates were in the Green zone 72% of the time and in the Upper Yellow zone 27%, following a similar but less efficient trend compared to PulseRide. However, on carpet (Figure 10B), the manual wheelchair led to opposite results: participants’ heart rates

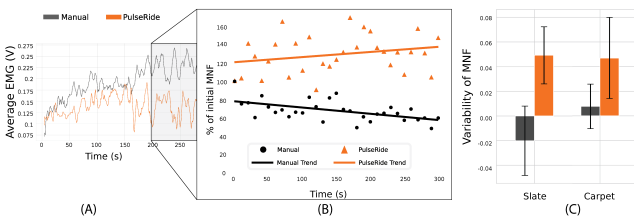
were in the Upper Yellow zone for approximately 57.2% of the time, with only 42.8% in the Green zone. This suggests that participants experienced fatigue more quickly when using a manual wheelchair on carpeted surfaces, making it difficult for them to maintain their heart rate within the moderate zone.



**Figure 10: Time spent in heart rate zones during tests (error bars: standard deviation). A. On slate, PulseRide users stayed longer in the moderate activity zone than with manual wheelchairs. B. On carpet, PulseRide also maintained moderate activity better, with manual wheelchairs leading to more intense activity.**

### 5.4 PulseRide delays fatigue

Our results demonstrate that PulseRide delays fatigue compared to manual wheelchairs. As shown in Figure 11A, the rectified average EMG values (moving average with a 5-second window) of a single participant highlight reduced muscle activation during PulseRide use. Figure 11B shows the mean frequency (MNF) trend, where the fitted line for PulseRide displays a positive slope, indicating a gradual increase in MNF over time and suggesting reduced fatigue during extended use. Conversely, the MNF trend for manual wheelchairs exhibits a negative slope, indicating increased fatigue toward the end of the session.

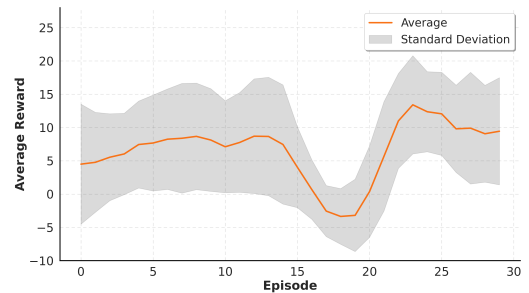


**Figure 11: (A) Raw electromyography (EMG) signal with a moving average (T=5) from the triceps brachii of a participant in the carpet environment during PulseRide use and manual wheelchair use. (B) Trend of the mean frequency (MNF) distribution for the participant’s EMG data (10-second window) in the carpet environment for PulseRide and manual wheelchair use. The Y-axis represents a percentage of initial values. A steeper negative slope in the manual wheelchair condition indicates a more rapid onset of fatigue. (C) Variability in mean frequency (MNF) across all participants in both slate and carpet environments for PulseRide and manual wheelchair use. Error bars represent the standard error of the mean.**

Figure 11C further illustrates MNF variability across all ten participants in both slate and carpet environments. PulseRide consistently demonstrated more positive MNF variability compared to manual wheelchairs, reflecting less fatigue during prolonged use across all participants. These results underscore PulseRide’s ability to provide adaptive assistance that mitigates fatigue more effectively than manual propulsion.

## 6 DISCUSSION

### 6.1 Reward plot



**Figure 12: Training reward averaged across all users. The solid orange line represents the average whereas the shaded represents standard deviation.**

Figure 12 shows the average reward per episode across participants. Initially, the average reward gradually increases up to episode 15 as the agent learns and improves with each episode. To ensure effective policy learning in our human-in-the-loop system, it is crucial that the agent experiences diverse physiological states and velocity ranges during training. This exploration is achieved by prompting the user the interface to follow a structured protocol. At episode 15, participants received a ‘Go faster’ prompt, followed by a ‘Go slower’ prompt at episode 25. These prompted phases intentionally drove participants’ velocities into transition or extreme zones, resulting in negative rewards. This explains the noticeable dip in rewards from episode 15 to 20. After this high-velocity phase, participants resumed maintaining velocity within their personalized moderate zone, leading to swift increase in reward till episode 25 followed by saturation till episode 30. While the reward trajectory suggests successful policy learning, we restricted the total number of episodes to 30 to keep the training time within 20 minutes. Further investigation is needed to understand potential additional policy improvements with extended training duration.

### 6.2 Data Scarcity

Deep Q-learning algorithms (DQN) typically require substantial environmental interactions to learn an effective policy. However, in this study, training data was limited by participant involvement, with only 600 total interactions available during the training phase. Given that wheelchair propulsion is physically demanding, it was essential to keep the training phase brief while still achieving a well-trained policy. To address this data scarcity, we incorporated domain knowledge through a pre-trained ECG encoder that extracts eight meaningful features per second based on the previous

three seconds of electrocardiogram data. This approach enriched the input features available to the primary DQN network. Additionally, we extended the agent's timesteps to two seconds, allowing it to process more data before each update. These adjustments significantly reduced the required training duration while ensuring an optimized model for each participant.

### 6.3 Participants without disabilities

Individuals with disabilities typically experience higher exertion and quicker muscle fatigue during wheelchair propulsion due to factors such as reliance on arm exercise, extent of paralysis, reduced sympathetic control, and overall activity levels [26]. While integrating physical activity into daily routines is crucial for cardiovascular fitness and muscle endurance [56], users are more likely to adhere to exercise intensities of their preference than to adjust based on strict physiological criteria [53]. With these considerations in mind, we designed our initial study as a proof of concept with participants without disabilities to evaluate PulseRide's core functionality in a controlled setting. Future studies with wheelchair users are essential to validate these findings in the target population.

## 7 CONCLUSION AND FUTURE WORK

In this study, we introduced PulseRide, an adaptive wheelchair system designed to provide personalized assistance based on users' physiological data and movement patterns. Using a Human-in-the-Loop Reinforcement Learning (HITL-RL) approach, PulseRide continuously adjusts push assistance to keep users within their moderate activity heart rate zone, offering a novel solution that bridges the gap between passive and physically taxing mobility options. Our experimental validation demonstrates that PulseRide effectively generalizes across different environments, maintaining consistent performance on both slate and carpet surfaces. Compared to manual wheelchairs, PulseRide successfully reduced muscle contractions and delayed fatigue onset, enabling users to maintain moderate activity levels over extended periods.

Moreover, our analysis shows that PulseRide personalizes assistance to match each user's unique physiological needs, providing tailored support without explicit environment-specific programming. While some outliers were observed, largely due to individual pushing preferences and the absence of braking, PulseRide's intelligent assistance remained effective in promoting sustained physical engagement without overexertion.

These findings highlight PulseRide's potential as a viable, adaptive mobility solution that can enhance health outcomes and user comfort, especially for those relying on wheelchairs for daily activities. However, there is still much room for further research, particularly in recruiting individuals with disabilities to provide real-world feedback and performance evaluations. This could broaden the scope of the study, extending its application to rehabilitation, where the system can directly track health and physical improvements over time, automatically suggesting personalized activity plans. While the DQN algorithm used in this study delivered a solid performance, future work should explore more complex algorithms, such as policy gradient [63], actor-critic [36], and proximal policy

optimization methods [60], to benchmark performance. Additionally, investigating continuous action spaces for motor thrust tuning could yield even better results.

Overall, this research opens new pathways for the development of assistive technologies driven by physiological data, ultimately promoting physical activity and health for wheelchair users. The potential for future research and system enhancement is vast, and the PulseRide system lays the foundation for next-generation personalized assistive devices that are health-conscious.

## REFERENCES

- [1] [n. d.]. SmartDrive. <https://hub.permobil.com/smartdrive#options> Accessed: 2024-08-30.
- [2] Acekare. 2023. Yomper+ and Enjo: Power add-on for manual wheelchairs. <https://www.acekare.com/power-add-on-manual-wheelchair-yomper-enjo/> Accessed: 2024-11-02.
- [3] Michael Akbar, Gabriel Balean, Manuela Brunner, Thorsten M. Seyler, Thomas Bruckner, Judith Munzinger, Thomas Grieser, Hans J. Gerner, and Markus Loew. 2010. Prevalence of rotator cuff tear in paraplegic patients compared with controls. *J. Bone Joint Surg. Am.* 92, 1 (Jan. 2010), 23–30. <https://doi.org/10.2106/JBJS.H.01373> arXiv:20048092
- [4] Alber USA. 2023. e-motion: Active drive wheels for manual wheelchairs. <https://www.alber-usa.com/us/products/active-drives/e-motion/> Accessed: 2024-11-02.
- [5] Mudasar Saleem Andrabi, Mercy Mumba, Betty Key, and Robert Motl. 2022. Physical activity programs for cardiovascular outcomes in community wheelchair users: A systematic review. *Frontiers in Rehabilitation Sciences* 3 (2022), 1007778.
- [6] Mahsa Asheghan, Mohammad Taghi Hollisaz, Taher Taheri, Hadi Kazemi, and Amidoddin Khatibi Aghda. 2016. The prevalence of carpal tunnel syndrome among long-term manual wheelchair users with spinal cord injury: A cross-sectional study. *J. Spinal Cord Med.* 39, 3 (May 2016), 265. <https://doi.org/10.1179/2045772315Y.0000000033>
- [7] Steven N Blair and Suzanne Brodney. 1999. Effects of physical inactivity and obesity on morbidity and mortality: current evidence and research issues. *Medicine and science in sports and exercise* 31, 11 Suppl (1999), S646–62.
- [8] Michael L. Boninger, Alicia M. Koontz, Sue Ann Sisto, Trevor A. Dyson-Hudson, Michael Chang, Robert Price, and Rory A. Cooper. 2005. Pushrim biomechanics and injury prevention in spinal cord injury: recommendations based on CULP-SCI investigations. *J. Rehabil. Res. Dev.* 42, 3 (May 2005), Suppl. <https://doi.org/10.1682/jrrd.2004.08.0103> arXiv:16195959
- [9] Frank W Booth, Scott E Gordon, Christian J Carlson, and Marc T Hamilton. 2000. Waging war on modern chronic diseases: primary prevention through exercise biology. *Journal of applied physiology* 88, 2 (2000), 774–787.
- [10] Claude Ed Bouchard, Roy J Shephard, and Thomas Ed Stephens. 1994. Physical activity, fitness, and health: international proceedings and consensus statement. In *International Consensus Symposium on Physical Activity, Fitness, and Health, 2nd, May, 1992, Toronto, ON, Canada*. Human Kinetics Publishers.
- [11] Johannes BJ Bussmann, Ulrich W Ebner-Priemer, and Jochen Fahrenberg. 2009. Ambulatory activity monitoring: Progress in measurement of activity, posture, and specific motion patterns in daily life. *European Psychologist* 14, 2 (2009), 142–152.
- [12] Sotirios Chatzidimitriadis and Konstantinos Sirlantzis. 2022. Deep Reinforcement Learning for Autonomous Navigation in Robotic Wheelchairs. In *International Conference on Pattern Recognition and Artificial Intelligence*. Springer, 271–282.
- [13] R. A. Cooper, L. A. Quatrano, P. W. Axelson, W. Harlan, M. Stineman, B. Franklin, J. S. Krause, J. Bach, H. Chambers, E. Y. Chao, M. Alexander, and P. Painter. 1999. Research on physical activity and health among people with disabilities: a consensus statement. *J. Rehabil. Res. Dev.* 36, 2 (April 1999), 142–154. arXiv:10661530 <https://pubmed.ncbi.nlm.nih.gov/10661530>
- [14] Elaine H Coulter, Philippa M Dall, Lynn Rochester, Jon P Hasler, and Malcolm H Granat. 2011. Development and validation of a physical activity monitor for use on a wheelchair. *Spinal cord* 49, 3 (2011), 445–450.
- [15] Rachel E Cowan, Stephanie L Silveira, Tina Helle, Uffe Læssøe, Kirstine Rosenbeck Gøeg, Jette Bangshaab, and Robert W Motl. 2022. Lifestyle physical activity in manual wheelchair users—an overlooked public health opportunity. *Spinal Cord* 60, 2 (2022), 190–192.
- [16] M. Dalyan, D. D. Cardenas, and B. Gerard. 1999. Upper extremity pain after spinal cord injury. *Spinal Cord* 37, 3 (March 1999), 191–195. <https://doi.org/10.1038/sj.sc.3100802> arXiv:10213328
- [17] Andrea L Dunn, Ross E Andersen, and John M Jakicic. 1998. Lifestyle physical activity interventions: History, short- and long-term effects, and recommendations. *American journal of preventive medicine* 15, 4 (1998), 398–412.
- [18] Kara Edwards and Annie McCluskey. 2010. A survey of adult power wheelchair and scooter users. *Disability and Rehabilitation: Assistive Technology* 5, 6 (2010),

- 411–419.
- [19] Arturo Escobedo, Anne Spalanzani, and Christian Laugier. 2013. Multimodal control of a robotic wheelchair: Using contextual information for usability improvement. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 4262–4267.
  - [20] Guoxi Feng, Lucian Buşoni, Thierry-Marie Guerra, and Sami Mohammad. 2019. Data-efficient reinforcement learning for energy optimization of power-assisted wheelchairs. *IEEE Transactions on Industrial Electronics* 66, 12 (2019), 9734–9744.
  - [21] Patrick Fougeyrollas and Luc Noreau. 2000. Long-term consequences of spinal cord injury on social participation: the occurrence of handicap situations. *Disability and Rehabilitation* 22, 4 (2000), 170–180.
  - [22] Heather D. Fullerton, Jeffrey J. Borckardt, and Alan P. Alfano. 2003. Shoulder pain: a comparison of wheelchair athletes and nonathletic wheelchair users. *Med. Sci. Sports Exerc.* 35, 12 (Dec. 2003), 1958–1961. <https://doi.org/10.1249/01.MSS.0000099082.54522.55> arXiv:14652488
  - [23] Shawn C Gendle, Mark Richardson, James Leeper, L Brent Hardin, J Matt Green, and Phillip A Bishop. 2012. Wheelchair-mounted accelerometers for measurement of physical activity. *Disability and Rehabilitation: Assistive Technology* 7, 2 (2012), 139–148.
  - [24] Edward M. Giesbrecht, Jacqueline D. Ripat, Arthur O. Quanbury, and Juliette E. Cooper. 2009. Participation in community-based activities of daily living: comparison of a pushrim-activated, power-assisted wheelchair and a power wheelchair. *Disabil. Rehabil. Assist. Technol.* 4, 3 (May 2009), 198–207. <https://doi.org/10.1080/17483100802543205> arXiv:19241234
  - [25] Aditya Goil, Matthew Derry, and Brenna D. Argall. 2013. Using machine learning to blend human and robot controls for assisted wheelchair navigation. In *2013 IEEE 13th International Conference on Rehabilitation Robotics (ICORR)*. 1–6. <https://doi.org/10.1109/ICORR.2013.6650454>
  - [26] JA Haisma, LHV Van Der Woude, HJ Stam, MP Bergen, TAR Sluis, and JBJ Bussmann. 2006. Physical capacity in wheelchair-dependent persons with a spinal cord injury: a critical review of the literature. *Spinal cord* 44, 11 (2006), 642–652.
  - [27] William L Haskell, I-Min Lee, Russell R Pate, Kenneth E Powell, Steven N Blair, Barry A Franklin, Caroline A Macera, Gregory W Heath, Paul D Thompson, and Adrian Bauman. 2007. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *Circulation* 116, 9 (2007), 1081.
  - [28] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.
  - [29] GREGORY NV HEATH and Peter H Fentem. 1997. 8 physical activity among persons with disabilities—A public health perspective. *Exercise and sport sciences reviews* 25, 1 (1997), 195–234.
  - [30] Catherine S. Holloway, Andrew Symonds, Tatsuto Suzuki, Angela Gall, Peter Smitham, and Stephen Taylor. 2015. Linking wheelchair kinetics to glenohumeral joint demand during everyday accessibility activities. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2015:2478–81., ; (2015), Annu. <https://doi.org/10.1109/EMBC.2015.7318896> arXiv:26736796
  - [31] Sungjin Hwang, Jiwoong Heo, Youngwug Cho, Jucheol Moon, Yushin Lee, Hansung Kim, Jaehyuk Cha, and Kwanguk Kim. 2024. Transportation Mode Detection Technology to Predict Wheelchair Users' Life Satisfaction in Seoul, South Korea. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 1, Article 9 (March 2024), 20 pages. <https://doi.org/10.1145/3643506>
  - [32] TW Janssen, CA Van Oers, Edith P Rozenaal, EM Willemsen, A Peter Hollander, and LH Van der Woude. 1996. Changes in physical strain and physical capacity in men with spinal cord injuries. *Medicine and science in sports and exercise* 28, 5 (1996), 551–559.
  - [33] Robert T Kell, Gordon Bell, and Art Quinney. 2001. Musculoskeletal fitness, health outcomes and quality of life. *Sports Medicine* 31 (2001), 863–873.
  - [34] Y Antero Kesaniemi, Elliot Danforth, Michael D Jensen, Peter G Kopelman, PIERRE Lefebvre, Bruce A Reeder, et al. 2001. Dose-response issues concerning physical activity and health: an evidence-based symposium. *Medicine & Science in Sports & Exercise* 33, 6 (2001), S351–S358.
  - [35] Diederik P Kingma. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).
  - [36] Vijay Konda and John Tsitsiklis. 1999. Actor-critic algorithms. *Advances in neural information processing systems* 12 (1999).
  - [37] Peng Liao, Walter Dempsey, Hillol Sarker, Syed Monowar Hossain, Mustafa Al'Absi, Predrag Klasnja, and Susan Murphy. 2018. Just-in-time but not too much: Determining treatment timing in mobile health. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 2, 4 (2018), 1–21.
  - [38] Peng Liao, Kristjan Greenewald, Predrag Klasnja, and Susan Murphy. 2020. Personalized HeartSteps: A Reinforcement Learning Algorithm for Optimizing Physical Activity. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 1, Article 18 (mar 2020), 22 pages. <https://doi.org/10.1145/3381007>
  - [39] Chen-Lung Lu, Zi-Yan Liu, Jui-Te Huang, Ching-I Huang, Bo-Hui Wang, Yi Chen, Nien-Hsin Wu, Hsueh-Cheng Wang, Laura Giarré, and Pei-Yi Kuo. 2021. Assistive navigation using deep reinforcement learning guiding robot with UWB/voice beacons and semantic feedbacks for blind and visually impaired people. *Frontiers in Robotics and AI* 8 (2021), 654132.
  - [40] Biao Luo, Zhengke Wu, Fei Zhou, and Bing-Chuan Wang. 2024. Human-in-the-Loop Reinforcement Learning in Continuous-Action Space. *IEEE Transactions on Neural Networks and Learning Systems* 35, 11 (2024), 15735–15744. <https://doi.org/10.1109/TNNLS.2023.3289315>
  - [41] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. 2015. Human-level control through deep reinforcement learning. *nature* 518, 7540 (2015), 529–533.
  - [42] Satoshi Muraki, Noriaki Tsunawake, Yasuaki Tahara, Sagaru Hiramatsu, and Masahiro Yamasaki. 2000. Multivariate analysis of factors influencing physical work capacity in wheelchair-dependent paraplegics with spinal cord injury. *European Journal of Applied Physiology* 81 (2000), 28–32.
  - [43] Mark S Nash, Rachel E Cowan, and Jochen Kressler. 2012. Evidence-based and heuristic approaches for customization of care in cardiometabolic syndrome after spinal cord injury. *The journal of spinal cord medicine* 35, 5 (2012), 278–292.
  - [44] Tom E Nightingale, Peter C Rouse, Dylan Thompson, and James LJ Bilzon. 2017. Measurement of physical activity and energy expenditure in wheelchair users: methods, considerations and future directions. *Sports medicine-open* 3 (2017), 1–16.
  - [45] Office of Disease Prevention, Health Promotion, et al. 2000. US Department of Health and Human Services: Healthy People 2010. <http://www.health.gov/healthypeople/> (2000).
  - [46] Satoshi Ohashi, Akira Shionoya, Keiu Harada, Masahito Nagamori, and Hisashi Uchiyama. 2022. Posture Estimation Using Surface Electromyography during Wheelchair Hand-Rim Operations. *Sensors* 22, 9 (2022). <https://doi.org/10.3390/s22093296>
  - [47] World Health Organization and United Nations Children's Fund. 2022. *Global report on assistive technology*. World Health Organization.
  - [48] RALPH S PAFENBARGER Jr, ROBERT T HYDE, CHUNG-CHENG HSIEH, and ALVIN L WING. 1986. Physical activity, other Life-style patterns, cardiovascular disease and longevity. *Acta Medica Scandinavica* 220, S711 (1986), 85–91.
  - [49] Russell R Pate, Michael Pratt, Steven N Blair, William L Haskell, Caroline A Macera, Claude Bouchard, David Buchner, Walter Ettinger, Gregory W Heath, Abby C King, et al. 1995. Physical activity and public health: a recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *Jama* 273, 5 (1995), 402–407.
  - [50] Katrina L Piercy, Richard P Troiano, Rachel M Ballard, Susan A Carlson, Janet E Fulton, Deborah A Galuska, Stephanie M George, and Richard D Olson. 2018. The physical activity guidelines for Americans. *Jama* 320, 19 (2018), 2020–2028.
  - [51] Polomarco. 2023. *ECG Classification: CNN-LSTM with Attention Mechanism*. <https://www.kaggle.com/code/polomarco/ecg-classification-cnn-lstm-attention-mechanism>
  - [52] Bibek Poudel, Thomas Watson, and Weizi Li. 2022. Learning to control dc motor for micromobility in real time with reinforcement learning. In *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 1248–1254.
  - [53] L Qi, M Ferguson-Pell, Z Salimi, R Haennel, and A Ramadi. 2015. Wheelchair users' perceived exertion during typical mobility activities. *Spinal cord* 53, 9 (2015), 687–691.
  - [54] Pranav Rajpurkar, Awni Y Hannun, Masoumeh Haghpanahi, Codie Bourn, and Andrew Y Ng. 2017. Cardiologist-level arrhythmia detection with convolutional neural networks. *arXiv preprint arXiv:1707.01836* (2017).
  - [55] W Jack Rejeski, Lawrence R Brawley, and Sally A Shumaker. 1996. Physical activity and health-related quality of life. *Exercise and sport sciences reviews* 24, 1 (1996), 71–108.
  - [56] D Rimaud, P Calmels, and X Devillard. 2005. Training programs in spinal cord injury. In *Annales de readaptation et de médecine physique: revue scientifique de la Société française de rééducation fonctionnelle de readaptation et de médecine physique*, Vol. 48. 259–269.
  - [57] Sandra B Rodrigues, Luis Palermo de Faria, António M Monteiro, José Luis Lima, Tiago M Barbosa, and José A Duarte. 2022. EMG signal processing for the study of localized muscle fatigue—pilot study to explore the applicability of a novel method. *International Journal of Environmental Research and Public Health* 19, 20 (2022), 13270.
  - [58] Hye-Yeon Ryu, Je-Seong Kwon, Jeong-Hak Lim, A-Hyeon Kim, Su-Jin Baek, and Jong-Wook Kim. 2021. Development of an autonomous driving smart wheelchair for the physically weak. *Applied Sciences* 12, 1 (2021), 377.
  - [59] Sushil Kumar Sahoo and Bibhuti Bhusan Choudhury. 2023. AI advances in wheelchair navigation and control: A comprehensive review. *Journal of process management and new technologies* 11, 3–4 (2023), 115–132.
  - [60] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347* (2017).
  - [61] Mandi Sehgal, Jeremy Jacobs, and Wendy S Biggs. 2021. Mobility assistive device use in older adults. *American family physician* 103, 12 (2021), 737–744.

- [62] Richard S Sutton. 2018. Reinforcement learning: An introduction. *A Bradford Book* (2018).
- [63] Richard S Sutton, David McAllester, Satinder Singh, and Yishay Mansour. 1999. Policy gradient methods for reinforcement learning with function approximation. *Advances in neural information processing systems* 12 (1999).
- [64] Michelle L Tolerico, Dan Ding, Rory A Cooper, Donald M Spaeth, Shirley G Fitzgerald, Rosemarie Cooper, Annmarie Kelleher, and Michael L Boninger. 2007. Assessing mobility characteristics and activity levels of manual wheelchair users. *Journal of Rehabilitation Research & Development* 44, 4 (2007).
- [65] Mohd Razali Md Tomari, Yoshinori Kobayashi, and Yoshinori Kuno. 2012. Development of smart wheelchair system for a user with severe motor impairment. *Procedia Engineering* 41 (2012), 538–546.
- [66] Anne Vuillemin, Stéphanie Boini, Sandrine Bertrais, Sabrina Tessier, Jean-Michel Oppert, Serge Herberg, Francis Guillemin, and Serge Briançon. 2005. Leisure time physical activity and health-related quality of life. *Preventive medicine* 41, 2 (2005), 562–569.
- [67] Nianfeng Wan, S. Alireza Fayazi, Hamed Saeidi, and Ardalan Vahidi. 2014. Optimal power management of an electric bicycle based on terrain preview and considering human fatigue dynamics. *Proceedings of the American Control Conference*, 3462–3467. <https://doi.org/10.1109/ACC.2014.6859373>
- [68] Darren ER Warburton, Norman Gledhill, and Arthur Quinney. 2001. The effects of changes in musculoskeletal fitness on health. *Canadian journal of applied physiology* 26, 2 (2001), 161–216.
- [69] Darren ER Warburton, Crystal Whitney Nicol, and Shannon SD Bredin. 2006. Health benefits of physical activity: the evidence. *Cmaj* 174, 6 (2006), 801–809.
- [70] Jacob O Wobbrock, Shaun K Kane, Krzysztof Z Gajos, Susumu Harada, and Jon Froehlich. 2011. Ability-based design: Concept, principles and examples. *ACM Transactions on Accessible Computing (TACCESS)* 3, 3 (2011), 1–27.
- [71] World Health Organization. 2023. Assistive Technology. <https://www.who.int/en/news-room/fact-sheets/detail/assistive-technology> Accessed: 2024-09-27.

Received 6 January 2025; revised 31 March 2025; accepted 30 June 2025