FEDERAL UNIVERSITY OF UBERLÂNDIA
FACULTY OF ELECTRICAL ENGINEERING
PROGRAM OF POST-GRADUATION IN ELECTRICAL ENGINEERING

APPLICATION OF NEURAL NETWORK TO ASSESS WHEELCHAIR DRIVING ABILITIES BY POWER MOBILITY ROAD TEST

Felipe Roque Martins

Uberlândia – Minas Gerais
2022

Felipe Roque Martins

APPLICATION OF NEURAL NETWORK TO ASSESS WHEELCHAIR DRIVING ABILITIES BY POWER MOBILITY ROAD TEST

Tese de doutorado apresentada ao Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal de Uberlândia, como parte dos requisitos para obtenção do título de **DOUTOR EM ENGENHARIA ELÉTRICA**.

Orientador: Prof. Dr. Eduardo Lázaro Martins Naves

Uberlândia – Minas Gerais
2022

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

M386
2022

Martins, Felipe Roque, 1992-
Application of neural network to assess wheelchair driving abilities by power mobility road test [recurso eletrônico] / Felipe Roque Martins. - 2022.

Orientador: Eduardo Lázaro Martins Naves.
Tese (Doutorado) - Universidade Federal de Uberlândia, Pós-graduação em Engenharia Elétrica.
Modo de acesso: Internet.
Disponível em: http://doi.org/10.14393/ufu.te.2022.679
Inclui bibliografia.
Inclui ilustrações.

1. Engenharia elétrica. I. Naves, Eduardo Lázaro Martins,1970-, (Orient.). II. Universidade Federal de Uberlândia. Pós-graduação em Engenharia Elétrica. III. Título.

CDU: 621.3

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:
Gizele Cristine Nunes do Couto - CRB6/2091
Nelson Marcos Ferreira - CRB6/3074

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
Coordenação do Programa de Pós-Graduação em Engenharia Elétrica
Av. João Naves de Ávila, 2121, Bloco 3N - Bairro Santa Mônica, Uberlândia-MG, CEP 38400-902
Telefone: (34) 3239-4707 - www.posgrad.feelt.ufu.br - copel@ufu.br

ATA DE DEFESA - PÓS-GRADUAÇÃO

<table>
<tr><td>Programa de Pós-Graduação em:</td><td colspan="5">Engenharia Elétrica</td></tr>
<tr><td>Defesa de:</td><td colspan="5">Tese de Doutorado, 313, PPGEELT</td></tr>
<tr><td>Data:</td><td>Vinte e dois de dezembro de dois mil e vinte e dois</td><td>Hora de início:</td><td>14:00</td><td>Hora de encerramento:</td><td>17:00</td></tr>
<tr><td>Matrícula do Discente:</td><td colspan="5">11723EEL010</td></tr>
<tr><td>Nome do Discente:</td><td colspan="5">Felipe Roque Martins</td></tr>
<tr><td>Título do Trabalho:</td><td colspan="5">Application of neural network to assess wheelchair driving abilities by power mobility road test</td></tr>
<tr><td>Área de concentração:</td><td colspan="5">Processamento da Informação</td></tr>
<tr><td>Linha de pesquisa:</td><td colspan="5">Engenharia Biomédica</td></tr>
<tr><td>Projeto de Pesquisa de vinculação:</td><td colspan="5">Coordenador do projeto: Eduardo Lázaro Martins Naves Título do projeto: Rehabnet: Rede de Telereabilitação por meio de Realidade Virtual e Aumentada Agência financiadora: CNPq Número do processo na agência financiadora: 307754/2020-0 Vigência do projeto: 01/03/21 a 28/02/24</td></tr>
</table>

Reuniu-se por meio de videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Engenharia Elétrica, assim composta: Professores Doutores: Angela Abreu Rosa de Sá - PNPD/FEELT/UFU; Maria Aparecida Ferreira de Mello - UNIFESP; Guilherme Fernandes Souza Miguel - UFES; Daniel Stefany Duarte Caetano - UNIESSA; Eduardo Lázaro Martins Naves - FEELT/UFU, orientador(a) do(a) candidato(a).

Iniciando os trabalhos o(a) presidente da mesa, Dr(a). Eduardo Lázaro Martins Naves, apresentou a Comissão Examinadora e o candidato(a), agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor(a) presidente concedeu a palavra, pela ordem sucessivamente, aos(às) examinadores(as), que passaram a arguir o(a) candidato(a). Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o(a) candidato(a):

Aprovado.

Esta defesa faz parte dos requisitos necessários à obtenção do título de Doutor.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.

Documento assinado eletronicamente por **Maria Aparecida Ferreira de Mello**, **Usuário Externo**, em 22/12/2022, às 16:31, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Guilherme Fernandes de Souza Miguel**, **Usuário Externo**, em 22/12/2022, às 16:45, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **DANIEL STEFANY DUARTE CAETANO**, **Usuário Externo**, em 22/12/2022, às 16:49, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Angela Abreu Rosa de Sá**, **Usuário Externo**, em 22/12/2022, às 16:49, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Eduardo Lazaro Martins Naves**, **Professor(a) do Magistério Superior**, em 22/12/2022, às 16:49, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **4131787** e o código CRC **B7BBBD20**.

ACKNOWLEDGMENTS

The research presented in this document would not be concluded without several contributions, major or otherwise, from the most diverse sources. I would like to thank my parents, Carolina and Elzionor, and my sister Thalita, for being always the first to celebrate my victories and help me recover from my failures. Gratitude to my wife Kathryn, for providing unlimited support and being at my side during every step of the way. My heartfelt thank you. You all are my most precious treasures.

I would also like to extend the gratitude to my friends and family, all of those who were at the right place and time. Life is often composed of so many small elements that, if changed ever so slightly, may alter the whole big picture in so many different ways. It is a wonder how each of you aligned so perfectly in such a chaotic and beautiful world.

My profound gratitude to my advisor, Eduardo Martins, for providing guidance when the directions were not so clear, and to all members of the Assistive Technology Lab, in special Caroline Valentini and Giovana Reis, both of which were invaluable to the execution of the research. Thank you to each of the participants of the research, for providing assistance willingly while expecting nothing in return. Wherever the results of the research will lead, it will be thanks to your participation.

To the members of the examination board, thank you for so many important suggestions, all of which will further contribute to the development of the study. It is an honor having you be part of this fundamental step of my life.

My gratitude to the Federal University of Uberlândia and to the Coordination for the Improvement of Higher Education Personnel (CAPES), for providing the resources, be it technical, theoretical or financial, necessary to fulfill the objectives of this research.

RESUMO

O procedimento padrão de prescrição de cadeiras de rodas motorizadas para pessoas com deficiência físicas envolve diversas etapas, incluindo avaliação das habilidades de condução da cadeira de rodas por parte de um profissional da área da saúde. Tal profissional utiliza ferramentas clínicas desenvolvidas por pesquisadores para auxiliar no processo de avaliação. Porém, um dos principais problemas é que a maior parte de tais ferramentas dependem do julgamento e da experiência do profissional, podendo portanto ser subjetivas. Existem estudos que avaliam a possibilidade de utilizar métricas objetivas para determinar se o indivíduo possui as habilidades necessárias para usar a cadeira de rodas. O uso de tecnologias de realidade virtual permite a obtenção de tais métricas, que de outra forma poderiam ser complexas de serem obtidas em uma situação com uma cadeira de rodas real, ou até mesmo oferecer riscos à segurança do usuário. Os dados obtidos necessitam então de processo de tomada de decisão, e embora tais análises possam ser feitas pelo profissional da saúde, elas requerem conhecimento e entendimento do significado de tais métricas. Dessa forma, algoritmos de *machine learning* (aprendizado de máquina) vem com o intuito de automatizar o processo de avaliação pela utilização de redes neurais com treinamento supervisionado. O objetivo da presente tese é desenvolver e avaliar um sistema estruturado a partir de redes neurais, utilizando quatro métricas obtidas através do simulador de cadeira de rodas EWATS, modelado de acordo com as tarefas fornecidas pelo Power Mobility Road Test para avaliar as habilidades de condução do usuário. As métricas selecionadas foram: tempo de execução da tarefa, número de comandos enviados ao joystick que controla a cadeira de rodas motorizada, número de colisões apresentado durante o percurso, e o valor da raiz do erro médio quadrático (RMSE, que avalia a distância da trajetória de um dado objeto em relação à menor trajetória ou trajeto otimizado). Foram feitos experimentos usando dois grupos distintos, um para obtenção de dados para treinar a rede neural no processo de avaliação da tarefa e outro para testar a rede neural após treinada, ambos supervisionados por um profissional da saúde. Três modelos de classificadores foram comparados através de um teste dos postos sinalizados de Wilcoxon: um Multi-layer Perceptron (MLP), um SVM (Support Vector Machine) e um KNN (k-Nearest Neighbors). Verificou-se com significância estatística que o SVM obteve melhor precisão que os outros dois modelos (80%), mas vários atributos podem ser explorados no design da MLP para melhorar sua precisão. Além disso, mais testes com uma amostra maior e com maior representatividade dos dados também são necessários para se obter um melhor resultado de classificação.

Palavras-chave: Redes neurais; cadeira de rodas motorizada; realidade virtual; ferramentas de avaliação

ABSTRACT

The standard procedure for prescribing electric-powered wheelchairs for people with physical disabilities involves several steps, including the assessment of wheelchair driving skills by a healthcare professional who uses clinical tools developed by a researcher to assist in the evaluation process. However, one of the main problems found in the clinical setting is that these tools are generally dependent on the professional's judgment and experience, and can therefore be subjective. There are studies that evaluate the possibility of using objective metrics to determine whether an individual has the necessary skills to use a wheelchair. The use of virtual reality technologies allows obtaining such parameters; collecting this data could otherwise be too complex while using a real wheelchair, or may even present risks to the user's safety. The data obtained can then be used to aid the process of decision-making, and although such analyses may be performed by the health professional, it demands knowledge and an understanding of their meaning. Machine learning algorithms are presented as an alternative to automate the evaluation process by using neural networks with supervised training. The objective of the present thesis is the development and assessment of a system created using neural networks, using four objective metrics obtained through the EWATS wheelchair simulator, which was modeled according to tasks provided by the Power Mobility Road Test to evaluate the driving skills of the user. The selected metrics were: time elapsed during the task execution, number of commands sent to the joystick that controls the motorized wheelchair, number of collisions that occurred during the task, and the value of the root-mean-square error (RMSE, which evaluates the distance of the trajectory of a given object relative to the shortest or optimized trajectory). Experiments were carried out using two different groups, the first to provide data to train the neural network in the task evaluation process and the second to test the neural network post-training. Both groups were supervised by a health professional. Three classifier models were compared using a Wilcoxon Signed-Rank test: Multi-layer Perceptron (MLP), SVM (Support Vector Machine) and KNN (*k*-Nearest Neighbors). It was verified with statistical significance that the under the testing conditions the SVM obtained better prediction accuracy than both other models (80%), but several attributes could be explored in the design of the MLP to improve its accuracy. Further tests with a larger sample size and with greater representativity of the data are also necessary to obtain better classification results.

Keywords: Neural networks; electric-powered wheelchair; virtual reality; assessment tools

DERIVATIVE WORKS

Martins, F.R., Naves, E.L.M., Morère, Y. et al. (2022). Preliminary assessment of a multimodal electric-powered wheelchair simulator for training of activities of daily living. J Multimodal User Interfaces 16, 193–205. https://doi.org/10.1007/s12193-021-00385-9

ILLUSTRATION INDEX

INDEX OF TABLES

ABBREVIATIONS AND SYMBOLS

ADL: Activities of Daily Living

AI: Artificial Intelligence

AR: Augmented Reality

AT: Assistive Technology

CNN: Convolutional Neural Networks

CP: Cerebral Palsy

DoF: Degrees of Freedom

EPW: Electric-Powered Wheelchair

FERS: Functional Evaluation Rating Scale

HMD: Head Mounted Display

KNN: *k*-Nearest Neighbors

miWE: McGill Immersive Wheelchair

MLP: Multi-layer Perceptron

OPCM: Optimal Preview Control Model

PCDA: Power-Mobility Community Driving Assessment

PIDA: Power-Mobility Indoor Driving Assessment

PMFET: Power Mobility Functional Evaluation Tasks

PMRT: Power-Mobility Road Test

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

PWS: Power Wheelchair Simulator

QDM: Quantitative Driving Metrics

RNN: Recurrent Neural Networks

RMSE: Root-Mean-Square Error

SVM: Support Vector Machine

TBI: Traumatic Brain Injury

VAHM: Véhicule Autonome pour Handicapés Moteurs

VAHM-2: Véhicule Autonome pour Handicapés Moteurs-2

ViEW: Virtual Electrical Wheelchair

VR: Virtual Reality

VRSIM: Virtual Reality-based SIMulator

VRSIM-2: Virtual Reality-based SIMulator 2

WST-Q: Wheelchair Skills Test Questionnaire

WSTP: Wheelchair Skills Training Program

WTS: Wheelchair Training System

TABLE OF CONTENTS

1. INTRODUCTION

There are many assistive technology (AT) devices whose goal is to assist individuals with disabilities in their activities of daily living. The electric-powered wheelchair (EPW) is an AT device that provides mobility for individuals with severe motor disabilities caused by diverse pathology and injuries. However, a very specific set of motor, visual and cognitive skills is required in order to properly drive an EPW (Lange & Grieb, 2015). Some of these skills are optimization of the completion of a given task, the ability to perform certain maneuvers (such as handling curves and driving in reverse) and being able to avoid collisions with obstacles (Niniss & Inoue, 2006; Harrison et al, 2010). Further training may also be required to improve the aforementioned skills (Lange & Grieb, 2015). In fact, the lack of access to a proper assessment and training may increase the likelihood of accidents involving the wheelchair user and their surrounding environment (Fehr et al, 2000).

In order to minimize the risks found in driving an EPW without the proper training, solutions such as Virtual Reality (VR) technology were implemented to create a controlled and safe environment where the wheelchair user can learn and practice the relevant skills. In the health field, VR is commonly applied to reproduce real-life situations without any risks to patients, due to the possibility of simulating and visualizing certain actions that may not be perceived in the real world (Nunes et al, 2011). VR is also used as a tool for training, rehabilitation and education, among others (Sánchez et al, 2011). Several EPW simulators were created to address the lack of assessment and training of wheelchair users (Abellard et al, 2010; Nunnerley et al, 2016; John et al, 2018; Morère et al, 2015; Archambault et al, 2011b).

The task of assessing the capabilities of the individual driving a wheelchair are generally performed by health professionals in the field of AT. These professionals have a number of tools at their disposal that can assist in the evaluation process. The Power-Mobility Indoor Driving Assessment (PIDA) was created to evaluate performance of the individual in indoor tasks, such as using facilities like bathrooms and elevators, or driving through narrow doorways (Dawson et al, 1994). Similarly, the Power-Mobility Community Driving Assessment (PCDA) is used to assess the individual's driving skills in an outdoor environment (Letts et al, 1998). The Wheelchair Skills Test (WST) was originally conceived to assist in the assessment of manual wheelchair skills. However, later studies began to focus on the tool's ability to assess EPW skills as well (Kirby et al, 2015). The Power-Mobility Road Test (PMRT) is based on driving tasks from the WST. It is composed of 16 tasks, 12 predictable tasks and four unpredictable ones that inform the user's ability to interact with the environment (Massengale et al, 2005).

A common factor found in these assessment methods is their dependence on the skills of the professional performing the evaluation. While they function as guidelines, providing instructions on how to score the execution of a certain activity, the actual task of determining if a user possesses the necessary skills to suitably drive an electric-powered wheelchair is left entirely in the hands of the evaluator. Consequently, the same task could be scored differently by two different evaluators, as everything depends on their respective judgments and experience with the assessment tools. Furthermore, the process of prescribing the wheelchair using these methods remains a manual task, being time-consuming and inefficient, and delaying access to the technology to those who need it. As such, its structure could greatly benefit from a certain degree of automation, while maintaining the decision-making process under supervision of health professionals. Currently, one of the most well-accepted methods of process automation currently is the incorporation of artificial intelligence into a system.

Deep learning is an ever-growing field with practical applications and many active research topics. Software with artificial intelligence (AI) are used to automate routine labor, understand speech or images, make diagnoses in medicine and support basic scientific research. The use of AI to tackle problems involving knowledge of the world and make decisions that appear subjective is called machine learning (Goodfellow et al, 2016). The goal of machine learning is to design and create algorithms that allow the use of empirical data, experience and training to evolve and adapt to changes that occur in their environment (Fu & Hao, 2012).

In situations in which data is introduced to the system perfectly labeled and categorized (the corresponding correct outputs), and the machine then trains on its own, the machine learning is called supervised (Jhaveri et al, 2022). One of the particular uses for supervised machine learning is classification problems, in which the output of features only admits discrete, unordered values. Classification problems appear in many diverse real world applications, such as spam detection, churn prediction, sentiment analysis and dog breed detection. It becomes possible to obtain a discrete output given a series of features independent from human judgment, and therefore, less prone to subjectivity.

1.1 Justification and Contributions

The correct assessment of the motorized wheelchair driving skills is fundamental to the process of prescribing a wheelchair that suits the users' needs, while also pinpointing weaknesses that may require further training. While a number of tools were created over time to facilitate this role, the fact remains that they are task-dependent, and that the result of any consequent evaluation is subjective; the outcome depends on the observations and expertise of the health professional, and may vary from one professional to another, as stated previously.

Over the last few decades, objective methods have been created in an attempt to quantify the EPW driving abilities and remove subjectivity from the process of evaluation. While there are still discussions over which parameters best constitute a model that represents the users' driving skills, advancements in computational prowess have since enabled access to several metrics that can objectively measure speed and accuracy, which could be difficult in a clinical setting (Archambault et al, 2011b). Furthermore, the application of objective parameters in a virtual environment was tested with a positive outcome, providing high inter-rater reliability (Kamaraj et al, 2016).

The works of Mahajan (2012) in the University of Pittsburgh, for instance, calculate multiple metrics at once while trying to determine the performance of users with at least one year post-traumatic brain injury at completing a task of driving the wheelchair in a two-dimensional environment. The RMSE (root-mean-square error, showing the deviation of the midline of the driving path), time to execute the trial, movement offset, movement error, and significant changes in direction are all entry parameters used to obtain an estimate of the ability of the users in performing the task set before them. In his 2020 dissertation, Kamaraj tried to establish his Quantitative Driving Metrics (QDM) as a set of parameters to evaluate wheelchair driving skill in the virtual environment. They were composed of a set of kinematic variables, namely the time elapsed, number of collisions, linear and angular velocity and root mean squared deviation from the midline of a task. The results showed that the QDM selected are stable and able to represent driving skills in an EPW simulator (Kamaraj, 2020). Hernandez-Ossa et al consider the time spent while executing a task, path following error given by the RMSE value and number of movement commands made with the selected input interface to assess the user's performance while using their wheelchair simulator, the SimCadRom (Hernandez-Ossa et al, 2020).

However, it is important to highlight that these studies all have yet to provide an automated option to the outcome for those clinical assessment tools. Despite being able to acquire several quantitative parameters, the results are still under the supervision of a health professional and/or a specialist with the expertise to handle the decisions from the data obtained.

Techniques of machine learning, more specifically neural networks with supervised training, are algorithms able to receive (as input) the quantitative data obtained from the clinical assessment tools and provide an immediate answer to the evaluation. The reliability of a neural network increases over time as more data is fed into the system, and the parameters can be reconfigured provided that there is previous and adequate training. Using neural networks to classify the driving performance may in turn make it possible to automate the assessment and therefore speed up the prescription of new wheelchairs, further facilitating the entire process.

This thesis proposes the preliminary study for the inclusion of neural networks into the process of assessment to handle the quantitative data obtained while using a wheelchair simulator. After the supervised training of the neural network using data collected from the experiments, the algorithm is expected to be able to classify the performance of a given task without the need of a health professional's input. To the best of the author's knowledge, it is the first study to incorporate neural networks in assisting the healthcare professional with the decision-making process of evaluation of wheelchair driving skills.

1.2 Objectives

The objectives of the current research are divided into two parts: one main, overarching objective that defines the research question; and several specific ones, smaller in scope, whose goal is to provide the steps necessary to reach the main objective.

1.2.1 General Objective

The general objective of the research is to develop an automated method of assessment of human electric-powered wheelchair driving capabilities for new wheelchair users, using a combination of neural network techniques and quantitative parameters obtained from a task being executed in the wheelchair.

1.2.2 Specific Objectives

- Develop a method to evaluate the ability to properly drive an electric-powered wheelchair using a specific set of parameters.
- Create a virtual platform that provides a safe environment for testing with individuals.
- Collect and analyze the data of the parameters from tests in a controlled environment.
- Implement and train a neural network to use the parameters to automatically classify the results of tested individuals.
- Integrate the neural network into the virtual platform to provide real-time classifications for the individuals tested.
- Compare the accuracy of the neural network with the score obtained in the Power Mobility Road Test (PMRT) from health professionals.

1.3 Hypotheses

The following hypotheses will be evaluated in this research:

- The implemented neural network managed to classify the PMRT score correctly, using the set of metrics chosen for the experiment.
- It was not possible to confirm the hypothesis.

2. LITERATURE REVIEW

The present research is supported by scientific foundations for the development of the technology, as well as verification and assessment of both parameters selected to represent the wheelchair driving capabilities of an individual and the neural network trained using such data as input information, with the expected output of providing a score independently of the healthcare professional's supervision.

For this reason, the following sections will define the terms, concepts and models that compose the theoretical framework to support research whose goal is the technological maturity and fulfillment of the objectives proposed in this thesis.

2.1 Wheelchairs

Assistive technology (AT) devices have been increasingly available for individuals with several disabilities. Typical wheelchair users are individuals with spinal cord injuries, balance disorders, as well as older and frail people. Having access to the device has given independent mobility, frequently being the primary means of transportation, and may increase independence in activities of daily living (ADL). Currently, the development of wheelchairs is seen as a responsibility of both medical professional and biomedical engineers (Mikołajewska et al, 2013).

The earliest physical representation of wheeled chairs is established to be an image on a 6th century Chinese sarcophagus, but it is estimated that the history of the wheelchair dates back to 3500 BCE. It was not until the 16th century that more information about wheelchairs appeared, when an artist made a drawing of the King of Spain sitting in a very elaborate wheelchair pushed by others. The inventor of such a device remains unknown (Rodrigo & Herrera, 2008). The first reference of a self-propelled wheelchair is that of Steven Farfler, a 22-years-old watchmaker with physical disability who built himself a stable chair mounted on a three-wheeled chassis (Cooper, 1995).

From the 18th century onward, wheelchairs started to be present in a clinical setting, appearing at surgical and medical instrument catalogs as vehicles to transport patients. With a design similar to those of armchairs, they were made of wood, wicker and/or iron, often with the drawback of being ornate, heavy and cumbersome. It was around the 1950s when lightweight tubular-steel, folding wheelchairs started appearing. Its lighter designs facilitated travel and afforded access for many wheelchair users for the first time (Woods & Watson, 2004).

2.1.1 Electric-powered wheelchairs

Electric-powered wheelchairs were first recorded under development in the United States of America in 1903, powered by a 10-cell Edison battery and a Westinghouse 12-volt motor; however, the initial objective of the EPW was not to transport people with disabilities, however. Great Britain was the first country to start developing wheelchairs specifically designed for people with disabilities with models that could go up to 16 kph over distances between 48 and 64 km on a single charge. At this point in time, there was less of a demand for the electric versions of wheelchairs, since they could be expensive, unreliable and could only cover small distances before requiring a charge (Woods & Watson, 2003).

The discovery of penicillin in 1929 was a key factor in advances of the technology for electric-powered wheelchairs, since it drastically increased the survivability of many people with severe impairments. At the same time, Europe and North America both struggled to find technical solutions to the problems of disabilities as a result of World War II. Furthermore, the polio epidemic, thalidomide and the Vietnam War were additional factors that propelled engineering research and development into assistive devices, specially powered wheelchairs. As a result, the technology evolved quickly. In 1953, George Klein, a mechanical engineer from the National Research Council in Canada was responsible for bringing four big innovations to the EPW: (a) an electrical system with low current, solving problems such as rapid burning of switch contacts; (b) two independent drive motors instead of a single one for both models, a feature that made it possible to engage and disengage the drives separately; (c) a new type of controller that made use of the different driving states of each of the motors; and (d) a new gearbox incorporating spur gears and a friction pulley driving each main wheel of the chair, better utilizing the reversible motors and ensuring consistency of speed (Woods & Watson, 2003).

The typical operation for an interface device to operate the wheelchair uses four quadrants to indicate the direction given by the user. The first quadrant indicates the forward-right region, the second forward-left, third quadrant for reverse-left and last quadrant for reverse-right, as shown in Illustration 1. Controllers use the vector velocity of the two (or more) motors as the primary control variable. Forward and reverse speeds are programmable within the controller and are set by limiting the maximum speed of each motor (Cooper et al, 2006).

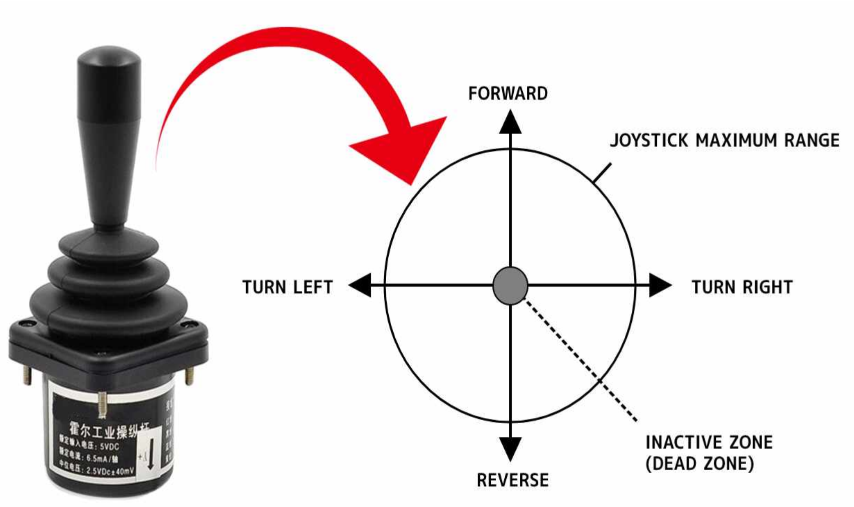

Illustration 1: Representation of a typical four-quadrant control for an EPW.

Advances in computational prowess have made it possible to combine technologies originally developed for mobile robots to create "smart wheelchairs". Smart wheelchairs are defined by the combination of a standard power wheelchair and a computer with a set of sensors. Smart wheelchairs have been designed to provide navigation assistance in different ways. For example, a smart wheelchair could ensure safe navigation, aid the performance of ADL or autonomously transport the user to a set location, reducing the physical, cognitive and perceptive load on individuals with severe disabilities (Zhang et al, 2021).

2.2. Wheelchair safety

Even though the development of wheelchairs has exponentially grown over the last 150 years, the fact remains that wheelchairs are assistive technology devices that require proper training and understanding to be utilized correctly. While roughly a tenth of wheelchair users need and/or use an electric-powered wheelchair, those models are linked to at least 25% of the accidents that occur with the users. Falling and tipping over are the most common cause of accidents involving a manual wheelchair, whereas collisions generally present the biggest challenge for EPW (Leblong et al, 2021). The frequency and conditions of the resulting injuries suggest the importance of stability in wheelchair design, prescription and training.

In regards to safety-related design, for example, it is possible to use sensor systems to assist in autonomous navigation of mobile robots. These approaches have been incorporated in electric-powered wheelchairs to assist disabled people in their mobility. The general architecture of a wheelchair navigation control system includes proper sensors to provide measurement of obstacles, a localization module that detects the wheelchair's position, a navigation module that generates the control variables, and a control module that turns those variables into commands for the motors (Desai et al, 2017).

There are several studies that assess the effectiveness of training for wheelchair users. Those training programs are efficacious, safe and practical in improving wheelchair-related skills (Tu et al, 2017). For example, the Wheelchair Skills Training Program (WSTP) has been tested with EPW users to test the wheelchair skills in comparison with a control group that receives standard care, along with other parameters, such as goal achievement, satisfaction with training, retention of skills, injury rate, confidence with wheelchair use and participation (Kirby et al, 2015). The assessment was done at different time periods (baseline, post-training and 3 months post-training), and along with the Wheelchair Skills Test Questionnaire (WST-Q) showed modest transient post-training improvements and a positive view on the whole training process.

Another essential step in ensuring safety and proper use of the technology comes from the correct assessment and prescription of the wheelchair. It requires matching an individual's needs with their environment. There are five important guidelines when evaluating an individual for a wheelchair: (a) patient history; (b) clinical impairments; (c) functional abilities; (d) unique body shape of the patient; and (e) any existing wheelchair. This test can be a trial (where the wheelchair is borrowed) or a simulation (when the patient is placed in a similar physical situation) (Batavia, 2010). The prescription is often overseen by a clinician (such as an occupational therapist) that is responsible for ensuring that the potential user possesses the combination of motor, visual and cognitive skills that are necessary to drive an EPW (Lange & Grieb, 2015).

2.2.1 Clinical assessment tools

Health professionals can rely on several tools to assist with the prescription of wheelchairs for new users. One of the attempts at standardizing the process of assessment was the Wheelchair Skills Test (WST), created by Kirby et al (2002). The proposal of the WST is to measure if wheelchairs are safe and effective for the users in their own environment, and involved users performing tasks in a standardized and obstacle-laden environment, presenting results in a reasonable period of time, in an environment where documentation of those results are possible, where safety is guaranteed, and with the presence of wheelchairs experts that could explore alternative components and make adjustments. The WST uses a scoring system to evaluate users based on whether the driver is able or not to perform each of the tasks.

The PIDA was developed in 1994 to help determine with accuracy the user's competency and safety while using an EPW with accuracy. It was created with the help of wheelchair experts including consumers and occupational therapists. There are a total of 30 tasks included in the final version, and tests indicated moderately strong intra-rater reliability and very good inter-rater reliability (Dawson et al, 1994). However, the conclusions suggest that raters (such as the health professional), subjects and qualities of the instruments factor in the reliability of the score.

While the goal of the PIDA is to assess the capabilities of the user in indoor situations, the PCDA was created by Letts et al to verify the user's abilities displayed in outdoor environments. Similar to the PIDA, the PCDA was developed using a modified nominal group consensus method. It evaluates the person-environment interaction as the individuals drive electric-powered wheelchairs in a variety of community settings. It was created in response to the increasing number of power-mobility devices, which presented itself as both an increased range of choices in equipment and potential risks for accidents (Letts et al, 1998). The scoring process is once again subjective since it depends on how the evaluator rates the capabilities of given tasks, such as approaching ramps and crossing entrances.

Hasdai et al developed the Functional Evaluation Rating Scale (FERS), which uses a scoring system similar to the PIDA and is used by other researchers to evaluate performance in simulators (Hasdai et al, 1998). The initial purpose of the study was to evaluate wheelchair driving skills in a simulator for children with disabilities. The resulting score in the simulation was significantly increased after a training period. Similarly to the FERS, the purpose of the Power Mobility Functional Evaluation Tasks (PMFET) purpose was to compare the performance models of wheelchairs for children with disabilities in a variety of indoor activities. Functional positioning at work surfaces, access to environments, and movement were addressed in simulated home and school settings (Deitz et al, 1991).

The Power Mobility Road Test was developed in 2001. Its tasks were adapted from a combination of three different clinical assessment tools: the PIDA (Dawson et al, 1994), the Functional Evaluation Rating Scale (FERS) (Hasdai et al, 1998) and the Power Mobility Functional Evaluation Tasks (Deitz et al, 1991). A total of 16 tasks were included, divided between two sections: structured and unstructured tasks. Structured tasks (tasks 1 to 12) are predictable and are as follows:

- Approaching people/furniture without bumping into them;
- Starting and stopping wheelchair at will;
- Passing through doorways without hitting walls (0.9m doorways);
- Turning around a 90 right hand corner (90 right turn);
- Turning around a 90 left hand corner (90 left turn);
- Driving straight forward (4.5m) in an open area;
- Driving straight backwards (3m) in an open area;
- Turning 180;
- Starting and stopping wheelchair upon request;
- Turning right and left upon command;
- Driving straight forward (4.5m) in a narrow corridor without hitting walls;
- Maneuver between objects.

On the other hand, the unstructured section consists of four tasks (tasks 13 to 16) that are unpredictable and unknown to the participant. The following tasks are included:

- Avoid unexpected obstacles (ball);
- Avoid unexpected obstacles (person entering hallway);
- One person coming towards participant in hallway;
- "Wet Floor" sign, choosing to wait or speed up.

The PMRT scoring is based on the PIDA scoring system, and uses a 4-point scale. A score of four indicates complete independence, while a score of three indicates that the user was hesitant about completing the tasks, requiring multiple trials and/or having minor accidents. A score of two shows that the participant committed serious accidents that could harm themselves or other people. Finally, a score of one shows an inability to complete a task. The total score (called total composite score) is expressed as a percentage and a score of over 95% indicates safe driving (Massengale et al, 2005).

2.3. Virtual reality and wheelchair simulators

The first computational interfaces started appearing around the 1940s and 1950s, and were based on the use of switches and lights. Initially, communication with the computer was done using machine language. It was not until the early 1960s that the first rudimentary graphical interfaces started appearing. With the advent of microprocessors, microcomputers became more accessible in the 1970s, and started using a command-based interface. The evolution of the command-based interface resulted in Windows, an operational system that continues to be used today. However, the Windows interface is restricted to the monitor and therefore needs specific representations, such as those for menus and icons (Kirner & Siscouto, 2007).

Virtual reality (VR) then emerged in the 1960s from a new generation of computer interfaces, using three-dimensional representations similar to those from the user's perspective, and allows crossing over the monitor barrier, while enabling more natural interactions with the system. It allows visualization, movement and user interaction, in real-time, with computer-generated three-dimensional environments. The sense of sight is usually predominant during its applications, but other senses such as touch and hearing may be used to enrich the user experience (Kirner & Siscouto, 2007).

Virtual reality applications are developed by using an appropriate set of tools for each case, such as languages, libraries and visual development environments. The correct selection of which tool to use depends on knowledge of the application domain and the availability of hardware and software resources, as well as financial resources when dealing with commercial tools (Guimarães et al, 2007). Some of the most used graphical engines for development of VR applications are Unity 3D, Unreal Engine, Frostbite Engine and CryEngine, although there are a variety of others that serve more specific applications.

One of the many characteristics of virtual reality is its ability to simulate real-life situations with a high degree of fidelity, without any of the risks generally associated with their real-life counterparts. VR has already demonstrated its potential for training vehicular manipulation, such as cars, bicycles or wheelchairs, and has proven to be safe, cost-effective and highly motivating (Pithon et al, 2009).

Several simulators have since been developed to cover different aspects of wheelchair use. The VRSIM, for example, was created to evaluate driving performance of individuals with cerebral palsy, identifying differences between two joystick designs, a standard movement sensing joystick and a novel isometric joystick (Dicianno et al, 2012). It was also developed with the intent to evaluate driving performance of wheelchair users in a VR environment. The works of Abellard et al (2010) and Arlati et al (2019) are two scoping reviews that describe the characteristics of the state of art for wheelchair simulators being developed in research settings. One of the few common traits for the majority of those settings is that they offer a first-person camera perspective, which contributes to the sense of immersion for the user.

Even if the development of wheelchair simulators has rapidly provided many interesting applications that are otherwise difficult or even impossible to achieve in the real world, a scientific gap was found in the capability to automatically evaluate an individual in driving the electric-powered wheelchair using such tools. Several quantitative parameters can easily be obtained in the virtual environment, but they act as indicators; once again, the onus is on the healthcare professional to determine whether the user possesses the necessary skills to drive an EPW to an adequate degree. This process of classification and scoring could be automated, for example, through use of machine learning techniques.

2.4. Machine learning and neural networks

The concept of artificial intelligence can be traced back to the 1940s, when English mathematician Alan Turing developed a code-breaking machine called The Bombe for the British government, considered the first working electro-mechanical computer, with the sole purpose of breaking a German code during World War II. The procedure in which The Bombe worked made Turing wonder about the intelligence of such machines, and in 1950, he published his work "Computing Machinery and Intelligence", laying the foundation for what is still considered the standard for identifying intelligence of an artificial system, the Turing Test, as indicated in Illustration 2. The term artificial intelligence appeared only six years later, in work by Stanford scientists Minsky and McCarthy, in an event whose goal was to create a new research area focused on machines capable of simulating human intelligence (Haelein & Kaplan, 2019).

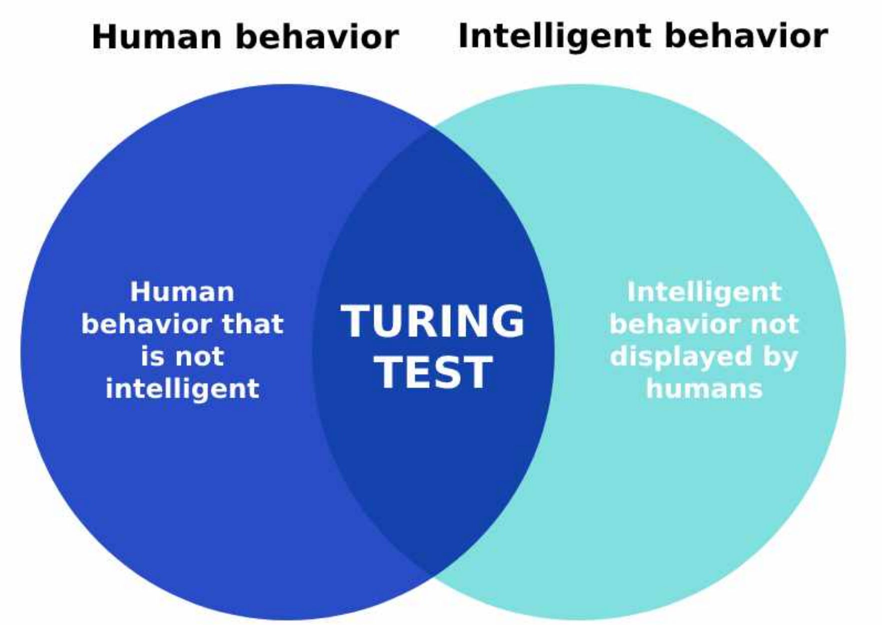

Illustration 2: Illustrated concept of the Turing Test to identify intelligence in an artificial system.

While artificial intelligence was quickly able to solve problems that may be difficult for humans, such as a chess match or breaking codes, tasks that humans solve intuitively (such as face recognition or recognition of spoken words) remained challenging for computers, since those are not bound by a list of formal, mathematical rules. It was not until the concept of deep learning was created that those intuitive tasks could finally be solved by computers (Goodfellow et al, 2016).

Deep learning is a subfield of artificial intelligence that approaches intuitive problems as a hierarchy of concepts, where each concept is defined through its relation to simpler concepts. By being able to learn from experience, it was possible to solve problems while avoiding the need for a knowledge-based approach, when human operators are required to specify all the knowledge needed beforehand. Instead, the AI system is able to acquire its own knowledge by extracting patterns from raw data. This capability is known as machine learning (Goodfellow et al, 2016).

The term machine learning refers to the automated detection of meaningful patterns of data. They can be classified according to the learning paradigm that they present: supervised or unsupervised, depending if the labels of the data are known (for example, to organize a set of data as one type or the other versus receiving all the raw data and trying to identify a pattern); active or passive (depending whether the system interacts with the environment during training, such as by posing questions, or does not interact at all); learning through a helpful or adversarial "teacher" (if the human operator determines that the training data is generated by some random process versus training under the assumption that the data is organized in an attempt to mislead the system); and finally online versus batch learning protocols (such as when the system has to provide an output throughout the process of learning, or if all the data can be processed before an output is required) (Shalev-Shwartz & Ben-David, 2009).

Neural network can be defined as an interconnected assembly of simple processing elements, units or nodes, whose functionality is loosely based on the animal neuron (Zador, 2019). Artificial neural networks are formal computation constructs that are modeled after a simplified model of the brain. They can be described as a directed graph whose nodes correspond to the neurons and edges are representations of the links between those neurons. Each neuron receives as input a weighted sum of the outputs of the neurons connected to its incoming edges, and their output is similarly used for neurons whose inputs are connected to those outgoing edges (Shalev-Shwartz & Ben-David, 2009).

There are several applications for neural networks. For example, recognition has always been a demanding task for humans, and since it lacks formal structure and mathematical methods, has been almost impossible for computers. Character recognition is seen as the ability to detect segmenting and identifying characters from images. Neural networks allow the implementation of multi-layered networks in order to be able to learn complex mappings in high dimensional spaces, therefore making pattern recognition a possibility for computers. They also can be used in conjunction with chemical sensing (sensor arrays or spectrometers) to mimic the mammalian olfactory system in humans (Kadurumba et al, 2020).

In the medical field, neural networks can be applied for interpretation (where the data is presented to the network which produces an output that can be directly interpreted), enhancement (in which the output is used to enhance the image or signal) and compression (where the data can be compressed by having fewer hidden layer neurons than inputs, facilitating the process of transmission and storage of information). Neural networks also may be used for artifact removal, image segmentation, classification of data and normalization of medical images, among other applications (Singh et al, 2020). Deep machine learning has found a steady source of incentive for advancements in the medical field, since there is a constant demand for better and more accurate results, with special attention given to healthcare in general.

3. RELATED WORKS

Given the increasing demand on assistive technology devices, including those aimed at mobility such as electric-powered wheelchairs, the research field is always expanding, and new applications for wheelchair simulators are constantly being proposed and developed. There are several works with emphasis on EPW training using virtual environments and which performance indicators are able to determine whether the necessary motor, visual and cognitive requirements are present in an individual.

The next sections will attempt to summarize the most relevant findings that identify, justify and validate the research direction taken in this document, including the scientific gap that the proposed system intends to fill.

3.1. Quantitative assessment of performance

Attempts to evaluate a wheelchair user's performance are not new in the literature. In fact, several clinical tools were created with such objectives in mind and were previously discussed, such as the WST (Kirby et al, 2002), the PIDA (Dawson et al, 1994), the FERS (Hasdai et al, 1998) and the PMRT (Massengale et al, 2005). Whilst effective and widely used in healthcare, the reality is that they remain qualitative methods of assessment; that fact makes it very difficult to determine a standardized procedure to evaluate a performance, since each of those tests may use a different set of criteria to classify the individual's skills.

There has been a motion in the research community ever since to try and shift the focus of assessment methods in favor of quantitative parameters rather than qualitative ones. Quantitative parameters are discrete, precise, and do not depend on interpretation of a third party to translate their meaning. Abellard et al (2010) provide one of the earliest studies to identify the drawbacks of the subjectivity of qualitative parameters, and simultaneously discuss the advantages of utilizing wheelchair simulators to solve problems such as material financial cost and safety concerns. As the authors explored different wheelchair simulators being developed up to that date, they briefly made mention of several objective parameters recorded by each of those simulators.

For instance, the Virtual Electric Power Wheelchair Driving initially developed in 2003 was capable of recording the number of collisions, time spent executing the proposed scenario and trajectory of the wheelchair (Ding et al, 2003). Another of the works discussed, the VAHM2 (Véhicule Autonome pour Handicapés Moteurs 2) simulator, was first presented as a virtual environment modeled after the VAHM (Véhicule Autonome pour Handicapés Moteurs) smart wheelchair proposed in 1998. The research contributes to the main topic discussed through two major observations: first, further studies with the simulator indicated the necessity of recording and using quantitative parameters to evaluate users' capabilities, namely duration (time elapsed), number of collision, number of times the wheelchair was stopped and trajectory of the wheelchair (Niniss & Nadif, 2000); second, some of the authors involved in the development and research of the VAHM wheelchair would later be responsible for the creation of the ViEW (Virtual Electric

Wheelchair) simulator, one of the forerunners in the current scenario of wheelchair simulator development (Morère et al, 2015). While much of the research presented would ultimately be considered outdated by modern standards, they allowed a glimpse at the correlation beginning to appear between research and development of EPW simulators and experimentation on quantitative variables to better represent wheelchair driving capabilities.

Following the discussion of performance assessment raised by Abellard, Morère et al (2018) questioned the possibility of associating performance indicators with qualitative methods, such as questionnaires and functional evaluation rating scales, to evaluate an individual's wheelchair driving skills. In reference to an earlier work (Morère et al, 2015), the metrics first incorporated to the evaluation process were: time taken to complete a task; mean and standard deviation of the trajectory; average absolute amplitude; and joystick angle. However, the newer iteration of the study made changes to those parameters, using instead overall distance covered by the participant, rather than the RMSE as described in Kamaraj et al (2014) since, according to the authors, the latter is ideal when the trajectory is particularly relevant to the analysis, whereas in an outdoor context the RMSE would have provided biased results. The time elapsed was omitted as well, being cited as potentially biased as the users could freely control the driving speed of the wheelchair. The jerk factor was another performance indicator established in the 2018 study, which describes the jerk factor value to be inversely correlated to the comfort level displayed by users. The mean amplitude of the joystick previously used was retained as means to show the confidence of the participant with the joystick (Morère et al, 2018).

Kamaraj presented a set of kinematic parameters he defined as QDM (Quantitative Driving Metrics) to act as indicators of driving capacity within the virtual environment. The variables that constitute the QDM are: time to execute a given task, number of collisions, average linear and angular velocities and RMSE. The results displayed high stability and construct validity, although the authors declare the further inquiry is necessary to prove the content validity of the set of variables. One important aspect found in the research was that the capability of differentiating between new and experienced wheelchair users are similar to those of the PMRT. In another section of the study, the author was able to create an experimental setup to replicate the acquisition of the QDM data in a real-world environment. For safety reasons, the number of collisions was removed from the computation of the variables; instead, the author used the jerk factor (calculated from the third derivative of the position) to quantify the smoothness of the EPW's trajectory. So far, the author has obtained good results in quantifying the driving experience, although he displayed reservations regarding some characteristics of driving that may not be easily computed in a given system (Kamaraj, 2020).

The work of Sá et al serves as a reference on what has been explored as quantitative parameters in the literature so far. While limited to wheelchair simulators as opposed to real-life electric-powered wheelchairs, they have found through the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) methodology 29 different quantitative indicators that have been used in several works. The most used indicators were: time elapsed (found in 23 out of the 42 studies analyzed), number of collisions (appearing in 20 studies), trajectory and speed (10 out of the 42 studies for each of the parameters) and number of joystick movements (seen in 8 of those studies). Even though these five parameters appeared most often in the literature, the fact that there is such a high number of performance indicators being used shows that no gold standard for quantification of the driving capabilities of wheelchair users exists as of yet, requiring a better effort to standardize the variables to compose the mathematical model that would represent such task (Sá et al, 2022).

3.2. Main developments in wheelchair simulators

As stated previously, the advancements on research in wheelchair simulation in virtual reality are intrinsically connected to the increased interest in the application of quantitative parameters in the assessment of driving capabilities. Therefore, many of the aforementioned works should be present or are somehow related to the works that appear in this section as well.

One of the early works of most significance for this thesis first appeared in the paper published by Niniss & Nadif in 2000, the VAHM2, itself an extension of a previous work by Bourhis & Agostini (1998). The VAHM2 was created to perform tests with the proposed smart wheelchair device, while taking into consideration problems of availability of people with disabilities in clinical settings, time-efficiency and safety of its users. The main goal of the study was to verify the efficiency of the integrated navigation system (with three modes of driving: manual, half-automatic, and full automatic mode). It utilized a HMD (head mounted display) interface in an 3D environment having the user seated in the VAHM wheelchair, and modeled 2D sensors to mimic the functioning of ultrasonic sensors and detect the presence of obstacles in the virtual world to adjust its trajectory. The kinematics of the wheelchair were controlled via calculation of the angular speed in each of its motorized wheels. By applying different combinations of angular speed in both left and right wheels (+ω, 0 and -ω), eight directions could be obtained (Niniss & Nadif, 2000).

Although the VAHM2 simulator is no longer in development, it sets a precedent for characteristics found on later wheelchair simulators. One of those simulators is the Virtual Electrical Wheelchair (ViEW) simulator. Initially designed as a 2D interface, later studies indicated the limitations of the two-dimensional approach in wheelchair simulation. Therefore, the version proposed in the 2015 study used a 3D modeling and 3D engine software (3DVia Virtools™ and 3D Studio Max™, respectively) for the creation of the virtual environment. The data recorded during sessions included: time, position of the wheelchair, position of the joystick, number of collisions, orientation of the head using a head-tracking device and gaze position using an eye-tracking device.

The study performed experiments to test the inability to drive an electric-powered wheelchair in the presence of cognitive or driving disorders, and it discusses the advantages and limitations of wheelchair simulators, such as providing a level of assistance not possible in real-life situations at the cost of lack of realistic behavior in certain elements of the scenario (Morère et al, 2015). A different research question was prompted in a 2018 study, where the ViEW simulator was used as an intervention performed in a six-month long course with individuals with cerebral palsy (CP) under supervision of an occupational therapist. The authors highlighted the benefits of using an EPW simulator for training, citing its stimulating benefits to both incentive participation and adherence (Morère et al, 2018). The ViEW simulator was also used in other works, such as validating the Optimal Preview Control Model (OPCM) strategy to deduce significant quantitative parameters determined by driving behavior (Zatla et al, 2018).

The VRSIM (Virtual Reality-based SIMulator) is another simulator that displayed extensive changes and applications over the years. Initially disclosed in the study of Spaeth et al, the first iteration of the VRSIM presented a bird's eye-view perspective in a 2D environment, with special focus given to obtain realistic steering and inertial behaviors. The main goal of the study was to analyze the feasibility of the simulator in experiments with participants with traumatic brain injury (TBI) (Spaeth et al, 2008). Subsequent studies aimed to discern the correlation between virtual training and real-world wheelchair training performance.

A newer, improved version of the VRSIM was proposed by Mahajan in his 2012 dissertation called VRSIM-2 (Virtual Reality-based SIMulator-2). It was written using C++ interfaced with the engine Multigen Paradigm Vega Prime. His experiment tasked 34 subjects (17 with CP and 17 in a control group) to perform a series of tasks involving turning left and right in corridors. Once again, a bird's eye-view camera was used, albeit in a 3D environment, limiting the sense of presence of the users. The sense of presence can be defined as the psychological sensation of being in the virtual environment (Slater & Wilbur, 1997).

The author further explores the development of the VRSIM-2 (which was dubbed "VRSIM-3.0", even though it remains identified as VRSIM-2 in all subsequent works), this time abandoning the bird's eye-view camera in favor of an immersive, first-person perspective. This latest iteration was modeled after the set of tasks presented in the PMRT, whose reliability in the virtual environment is established through experimentation as well. The performance indicators recorded during the tasks were RMSE, Cartesian coordinates of the wheelchair, speed, and number of collisions. Through ridge regression, the study attempted to predict real-life scores of the PMRT based on the user score in the virtual environment (Mahajan, 2012).

Further research was made using the VRSIM-2 with Kamaraj's attempt to assess the inter-rater reliability of the PMRT inside the simulator in his work (Kamaraj et al, 2016). He also applied some earlier findings regarding quantitative indicators with good representativity of a wheelchair user's driving capabilities (Kamaraj et al, 2014) into the best practices for the simulator, culminating in the definition of the QDM mentioned in the previous section (time, collisions, average linear and angular velocities and RMSE) (Kamaraj, 2020).

SimCadRom (Simulador de Cadeira de Rodas Motorizada) is an EPW simulator developed by Hernandez-Ossa et al in 2017 which aims to provide a safe environment for training driving capabilities and testing new control interfaces, taking in consideration users with severe disabilities unable to use the traditional EPW joystick. The SimCarRom was originally developed to provide a comparison between real-world driving and virtual environment, using qualitative methods of assessment (5-point Likert scale) as well as recording some quantitative data, such as time elapsed between tasks and times the virtual wheelchair left the boundaries established in the virtual environment (Hernandez-Ossa et al, 2017). Subsequent studies incorporated eye-tracking control to the system (Montenegro-Couto et al, 2018), and the quantitative performance indicators were changed to time, RMSE and number of commands (Hernandez-Ossa et al, 2020).

The work of Archambault is often used as a reference in wheelchair simulation. The McGill Immersive Wheelchair (or miWe) simulator was one of the early adopters to incorporate 3D first-person perspective in its design, referring to the sense of presence as an important concept to take into consideration in virtual environments. The first version of the miWe simulator was conceptualized in (Archambault et al, 2008), in which it is proposed as an application of the 6-DoF (degrees of freedom) movable platform developed by the research team, using linear and centrifugal accelerations to provide motion feedback to the user. Later studies attempted to validate the tool (developed in Unreal Engine™ and modeled after a subset of the WSTP) by comparing values of smoothness of joystick control and time between the virtual scenario and real-world wheelchairs. Interestingly enough, the results showed participants had more difficulty completing the VR tasks than those given in real-life (Archambault et al, 2011a). Further improvements closed the gap in reliability of the simulator and achieved better results (Archambault et al, 2011b). Another study used qualitative data, as well as time and number of collisions, to provide an assessment of modifications on the technology (Archambault et al, 2016). Presence of augmented feedback and skills learning retention were explored in a later study, and it showed moderate results in experiments performed with 40 participants (Bigras et al, 2019). A modified version of the simulator (miWe-C) was used in an experiment with children with CP, neuromuscular disease and spinal cord injury to contrast the virtual environment with the real-world tasks, and its results corroborated previous findings to validate the tool (Gefen et al, 2019).

The AccesSim project could be considered a deviation of the norm in terms of goals for wheelchair VR simulations. Rather than focusing on the development of the driving skills of users, it focused instead on detecting accessibility issues for wheelchair users; it is also aimed at architecture engineers and designers, as opposed to healthcare professionals. The AccesSim was based on the WSTP, and was designed to replicate the most common obstacles found by wheelchair users in terms of transportation (slopes and moving up and down a curb), and it registers linear acceleration and angular velocity as parameters to indicate the performance of a specific task (Gonçalves et al, 2014).

Headleand et al created the Wheelchair-Rift with the goal of incorporating serious games elements in its structure to raise interest and involvement of the user while performing the activities presented. The simulator was created using the Unity3D game engine and focused on the use of HMD technologies, more specifically the Oculus Rift™, to increase immersion and sense of presence. The setup also included a hand gesture-tracker device, the Leap Motion™ (attached to the HMD) to detect and reproduce the hand movements in the virtual environment and allow a more natural interaction with objects in such an environment. The virtual environment was modeled to reproduce a multi-story building, with each floor representing an increase in the level of difficulty of the tasks. While initial tests were performed to test the system, the simulator is yet to be validated in a clinical setting (Headleand et al, 2016).

The WTS (Wheelchair Training System) was proposed by Nunnerley et al as a means to address the lack of user-centered design solutions in VR simulation. The actual tool was developed in collaboration with several organizations in New Zealand, and the experiments were conducted with 12 participants (five experienced wheelchair users and seven healthcare) in the form of questionnaires with semi-structured questions regarding the user experience (Nunnerley et al, 2016). Vailland et al developed another wheelchair simulator called PWS (Power Wheelchair Simulator) in 2019, also focusing on user-centered design. The PWS was an extension of an early study (Devigne et al, 2017), which describes the development of a generic VR wheelchair simulator using the Unity3D game engine to test its semi-autonomous driving algorithm. The cybersickness felt by participants in Devigne's study was addressed in the research proposed by Vailland et al through use of different rendering interfaces. The latter, similarly to Nunnerley's study, emphasized the importance of participation of healthcare professionals and wheelchair users in the design and testing of the tools (Vailland et al, 2019).

3.2. Table comparison and research gap

Given the wide variety of findings that corroborate with the objectives of this thesis, the results of the literature review were compiled and shown in Table 1, to facilitate understanding and comparison between different works. The main products involving simulation of electric-powered wheelchairs are organized in accordance with the order they were presented in the previous sections. They are also categorized by what interfaces are used to display the simulator, and HMD are often selected in an attempt to increase the sense of presence inside the virtual environment. There is a direct correlation between the use of HMD devices and presence of cybersickness, but its study is beyond the scope of this document.

As many of the wheelchairs presented were involved in more than one research (often at the same time), and therefore its application could have shifted over time as well, only the objective of the first iteration of the simulator was included. In a similar fashion, only the first work to mention the design and/or development of the virtual environment is listed, although several derivative works were described in the previous sections. Another important information included was regarding what quantitative indicators were measured in the simulators. It serves to indicate how the parameters may change from simulator to simulator, and that various quantitative variables have been explored in the literature in an attempt to quantify the driving performance of individuals, as observed by Sá et al (2022).

Table 1: Summary of findings on electric-powered wheelchair simulators

Simulator name	Interfaces used to display VE	Initial applications	First appeared in	Quantitative variable measured	Neural network
VAHM2	HMD	Assisted driving	Niniss & Nadif, 2000	Angular speed	No
ViEW	Desktop screen, HMD	Study of feasibility	Morère et al, 2011	Time, Cartesian coordinates of joystick/wheelchair, collisions, head orientation and gaze position	No
VRSIM	Projected flat screen	Study of feasibility	Spaeth et al, 2008	RMSE and speed	No
VRSIM-2	Three plasma screens	Comparison between CP and control	Mahajan, 2012	QDM (RMSE, time, linear and angular speed, acceleration, collisions)	No
miWe	Desktop screen	Test 6-DoF platform	Archambault et al, 2008	Acceleration	No
AccesSim	Projected flat screen	Detect accessibility issues	Gonçalves et al, 2014	Acceleration, angular speed	No
Wheelchair-Rift	HMD	Study of feasibility	Headleand et al, 2016	None	No
WTS	HMD	Study of feasibility	Nunnerley et al, 2016	None	No
PWS	HMD, immersive rooms, desktop screen	Assisted driving	Devigne et al, 2017	Time, commands, Cartesian coordinates of wheelchair, speed, collisions	No
SimCadRom	HMD	Study of feasibility	Hernandez-Ossa et al, 2017	Time, RMSE, commands	No
EWATS	Desktop screen, HMD	Study of feasibility	Martins, 2017	Time, collisions, commands	No

A common denominator in all of those studies is that, while there is an attempt of quantifying the driving performance using several virtual and real-life quantitative parameters, there is a lack of a method to automate the task. Henceforth, the main contribution of this thesis is found in the effort to utilize classification techniques (based in neural networks) to learn, based on the performance indicators here selected, and provide an acceptable prediction of the score of users in a given task. The closest attempt at a prediction model was found in the work of Mahajan (2012). However, the focus of the later part of the study was to compare the score obtained in real-world PMRT tasks with those obtained from the ridge regression done from the data acquired.

4. MATERIAL AND METHODS

The present study comprises the development of a neural network capable of calculating the PMRT score through use of objective parameters. It is part of the research project titled RehabNet for the National Council for Scientific and Technological Development (CNPq) under the registration number 307754/2020-0.

The characteristics and structures that compose the system developed in this document will be explained in detail in the following sections, including the reasoning behind the design choices made during each step of the development.

4.1. System architecture and components

The figure shown in Illustration 3 provides a general overview of the system proposed. Its main structure can be divided into three different components: the user interface, the virtual reality and the machine learning components. Each component is intertwined with the other two and can be defined as the following:

- **User interface:** the user interface component is defined by the user themselves, along with the input method used to interface with the virtual reality system. It is dependent on the users' capabilities and limitations, for example using a traditional wheelchair joystick, electromyographic signals or even through gaze analysis.
- **Virtual reality:** this component refers to the wheelchair simulator used in the research, which in turn follows a specific set of instructions as defined in the clinical tool it was modeled after. The virtual reality component is responsible for translating the user's intent while using the input method into a graphical representation of the real-world.
- **Machine learning:** the last component of the system refers to the combination of the performance indicators acquired during the execution of a given task and the algorithm for the neural network responsible for processing the indicators and providing the expected output at the end.

For this study, the input method for the user interface was restricted to an adapted EPW joystick (so it can connect to a functioning computer) in order to minimize the bias created from different skill levels on each input method for an individual. As the main goal is to verify the accuracy of the algorithm of a neural network using a specific set of parameters, having several input methods at once would detract focus from the efficacy of the training and the performance indicators chosen to be used in the system.

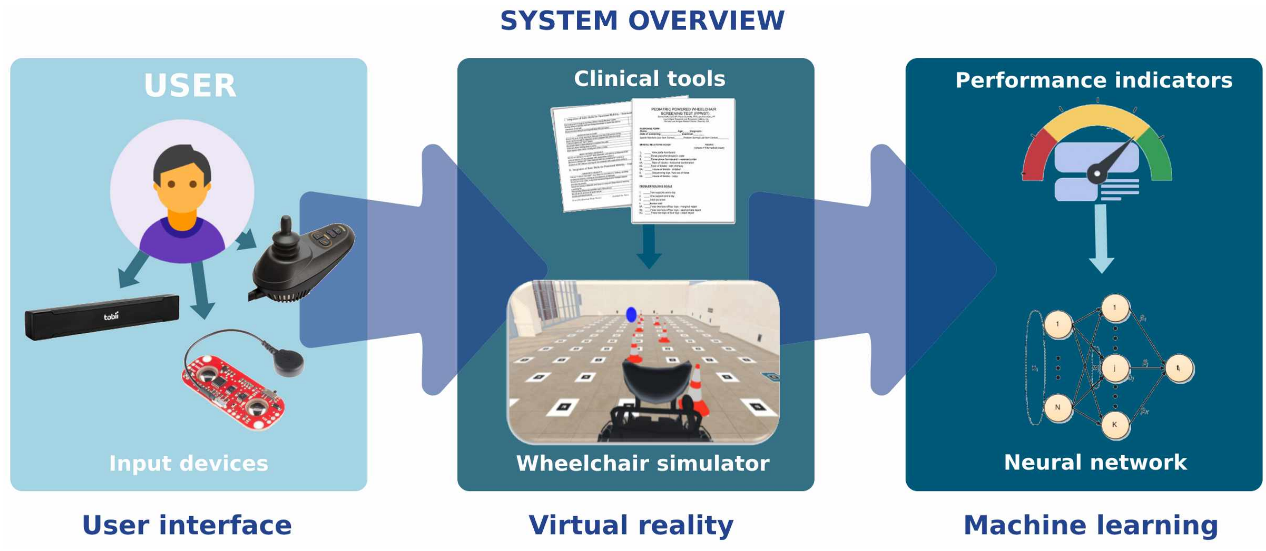

Illustration 3: Different components that define the architecture of the system proposed.

4.2. Performance indicators

As established in the previous sections, one of the main advantages that a simulated environment provides is the access to a series of measurement data that can serve as indication of the driving skills. There are different options when it comes to which performance indicator to use, for example: the time spent to complete a task and the number of joystick movements (Archambault et al, 2011b); the length of a determined trajectory or the number of collisions (Webster, 2001); comparison to an optimal trajectory (Niniss & Inoue, 2006); even multiple parameters at once, such as time, velocity, total of collisions and RMSE calculated from a specified path (Kamaraj et al, 2014).

There is no gold standard when it comes to the correct quantitative performance indicators. The parameters established to indicate the users' driving abilities in the experimental setup are derivative of the QDM proposed by Kamaraj. In that work, the QDM selected are: total number of collisions, time to complete a task, speed and RMSE obtained from deviation of the best trajectory (Kamaraj, 2020). Given both the discrete nature of neural network training and the limited space given to most tasks performed in the virtual environment (as those are indoor driving conditions), it did not feel appropriate to use speed as an indicator of performance. Instead, the number of movements made with the joystick was chosen to represent the general understanding of how to properly operate the device.

Although there are several quantitative parameters available through the literature (Sá et al, 2022), the setup described here was designed to be used with a smaller set of metrics. Having fewer parameters allows faster configuration of the predictive algorithm, which is expected not only to provide an answer in real-time to the execution of tasks, but will also generate a training database that will grow with each use, therefore reducing processing speed as the volume of data increases. For the reasons described, the performance indicators were limited to four in this first iteration of the prediction system: number of collisions, time, number of commands and RMSE, and each of those indicators will be described in greater detail in the following subsections.

4.2.1. Number of collisions

The number of collisions is a descriptor calculated by the number of times after the start of a task that the wheelchair made contact with an obstacle such as traffic cones, chairs and/or walls. Its value is directly correlated to the safety of the user in an environment, whether real or virtual, since it may represent the risk of accidents caused by lack of control and spatial notion of the EPW. One of the big advantages of using a VR simulator over a clinic experiment is that this parameter can be safely measured without actually creating any risks to the well-being of the individual. Values with a tendency to zero would be considered ideal, as it would indicate that the user is capable of interacting with the environment and reacting to unexpected situations without putting themselves in danger. It is also a parameter used in several of the works involving wheelchair simulators discussed previously (Morère et al, 2018; Zatla et al, 2018; Niniss & Inoue, 2006; Mahajan, 2012; John et al, 2018).

4.2.2. Time to complete a task

Monitoring the time an individual takes to complete a certain task has a direct correlation to the ability to make quick judgment calls and proceed towards a goal, showing fast reactions and good spatial notion. Generally speaking, the less time spent to conclude a given task, the better understanding of how to drive the wheelchair is displayed by the user. When combined with a low or non-existent number of collisions, those parameters are able to indicate that the individual is capable of driving the wheelchair safely and with high accuracy.

The measurement of the time elapsed is taken by recording the data on the precise frame as the system first detects an input given by the user (via joystick), and also in the frame as the wheelchair concludes the task being executed. The time is calculated separately for each of the tasks, and its use as a performance indicator can be seen in many different works (Morère et al, 2018; Zatla et al, 2018; Silva et al, 2018; Niniss & Inoue, 2006; Mahajan, 2012; Hernandez-Ossa et al, 2017). In fact, according to the review conducted by Sá et al, it is the most used quantitative metric found in the literature (Sá et al, 2022).

4.2.3. Number of commands

Similarly to the time required to execute a task, counting the number of commands is indicative of how much the user understands the operation of a wheelchair. It means understanding when to stop, the best time to start turning the wheelchair and how precise the commands should be when navigating an environment with obstacles. The number of commands account for how many directions in a specific interval of time were given to the EPW, and are measured from the moment when the first command is given, as well as every different command after until the end of the task. As such, as long as the user maintains the same direction, only one command will be counted.

For the study, a total of nine commands are included, as shown in Illustration 4.

- Move forward;
- Engage in reverse;
- Turn to the right;
- Turn to the left;
- Move forward while turning to the right;
- Move forward while turning to the left;
- Move in reverse while turning to the right;
- Move in reverse while turning to the left;
- Stop the wheelchair.

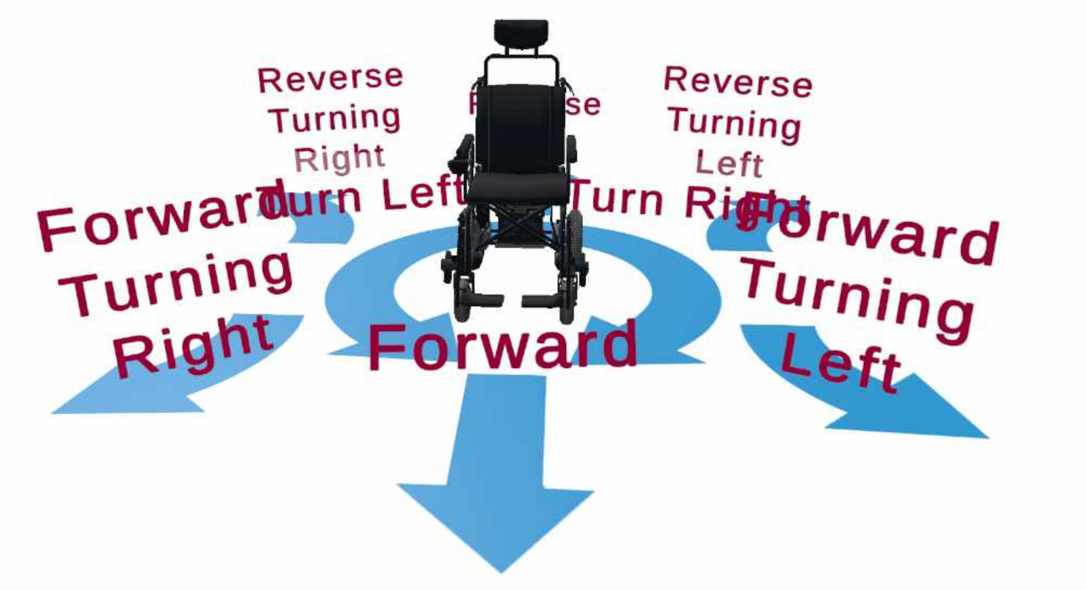

Illustration 4: Three-dimensional representation of the possible direction inputs given to the wheelchair.

The joystick used in the experiments was adapted from a real wheelchair joystick, where the original circuitry was replaced by a circuit composed of two analog potentiometers (one for each of the X and Y axes) attached to an Arduino™ microcontroller. The microcontroller was connected to the computer running the simulator via USB serial port. The reason for the adaptation was to increase the perception that the user was driving an EPW. Following the same line of thought, each of the participants on the experiments were asked to sit in a Freedom wheelchair as the tasks were being completed.

As with the other parameters included in the algorithm, measurement of the number of commands was done in other works as well (Morère et al, 2018; Archambault et al, 2011a; Silva et al, 2018).

4.2.4. Root-mean-square error (RMSE)

In statistics, the root-mean-squared error is commonly used to measure the differences between values predicted by some model against the values observed in a given experiment. The higher the value of the RMSE, the further away from the predicted model the set of data is. The formula that represents a generic model for the RMSE can be seen in Equation (1).

$$RMSE = \sqrt{\frac{\sum (Predicted_i - Actual_i)^2}{N}} \quad (1)$$

For the calculation of the RMSE in this study, the position of the wheelchair is taken using a set of Cartesian coordinates on a plane (x,y) at a frequency of 60Hz. The value is then compared to the ideal position for that frame according to Equation (1), resulting in the RMSE for that specific frame. The sequence of coordinates combined are defined as nominal trajectory, and the best possible sequence of coordinates is referred to as ideal trajectory. The RMSE is calculated from comparing the nominal trajectory to the ideal trajectory, as shown in Illustration 5, where the blue line defines the optimal trajectory for an electric-powered wheelchair model with only four directions (forward, reverse, turn left or right), hence the sharp lines. The other trajectories indicated deviation from the optimal and that can be seen in the increased time taken to complete the task, which indicates a positive correlation between RMSE values and time to finish the task.

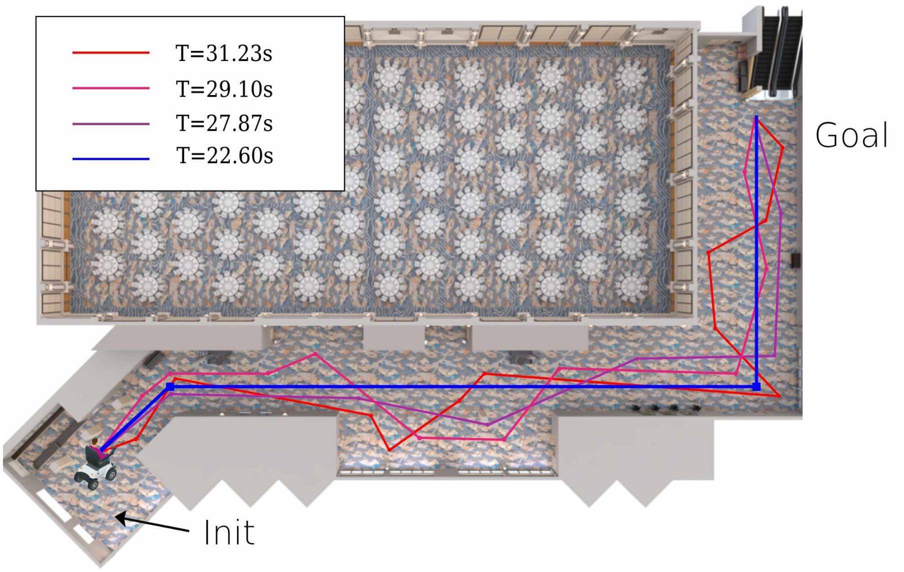

Illustration 5: Difference between ideal trajectory and nominal trajectory in the calculation of the RSME.

The intention of the RMSE value is to measure the ability of the wheelchair user in keeping a steady progression in a determined path, along with any deviations from that trajectory. It is an assessment of the accuracy of the driving, and combined with the other parameters gives a concrete understanding of the driving capabilities of the individual. It is a very informative parameter, used in diverse works, mostly using wheelchair simulation, since it is not as trivial a task to calculate RMSE in the virtual environment as opposed as with a real-world EPW (Morère et al, 2018; Niniss & Inoue, 2006; Mahajan, 2012; Kamaraj, 2020).

4.3. Neural network

After obtaining the parameters determined in the previous subsection, the resulting dataset (set of data including all parameters selected in this study) must be used to train a neural network to be able to recognize and classify the driving skills of an individual. Some of the most used classes of artificial neural networks are Multi-layer Perceptron (MLP), Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN). While CNN are very suitable for prediction of problems involving image data as the input, and RNN were created to work on problems such as sequence prediction, MLP are very adept at classification prediction problems with known inputs or labels.

As the objective of the neural network is to receive the data obtained from the execution of the task and to determine the score according to the PMRT, the decision was made to utilize MLP with back-propagation for the system. There are four inputs (time, collisions, commands and RMSE), one hidden layer and three possible outputs (scores from 2 to 4, since a score of 1 is given when the individual is unable to finish the task, and therefore the data should not be computed in the final calculation of the score as to not contaminate the rest of the data). While there is not any rule set on how many neurons should compose the hidden layer, an informal guideline suggests a starting point at 2/3 the size of the input layer, plus the size of the output layer. Rounding that value up resulted in a total of six neurons in the hidden layer. Illustration 6 demonstrates a model for the proposed neural network.

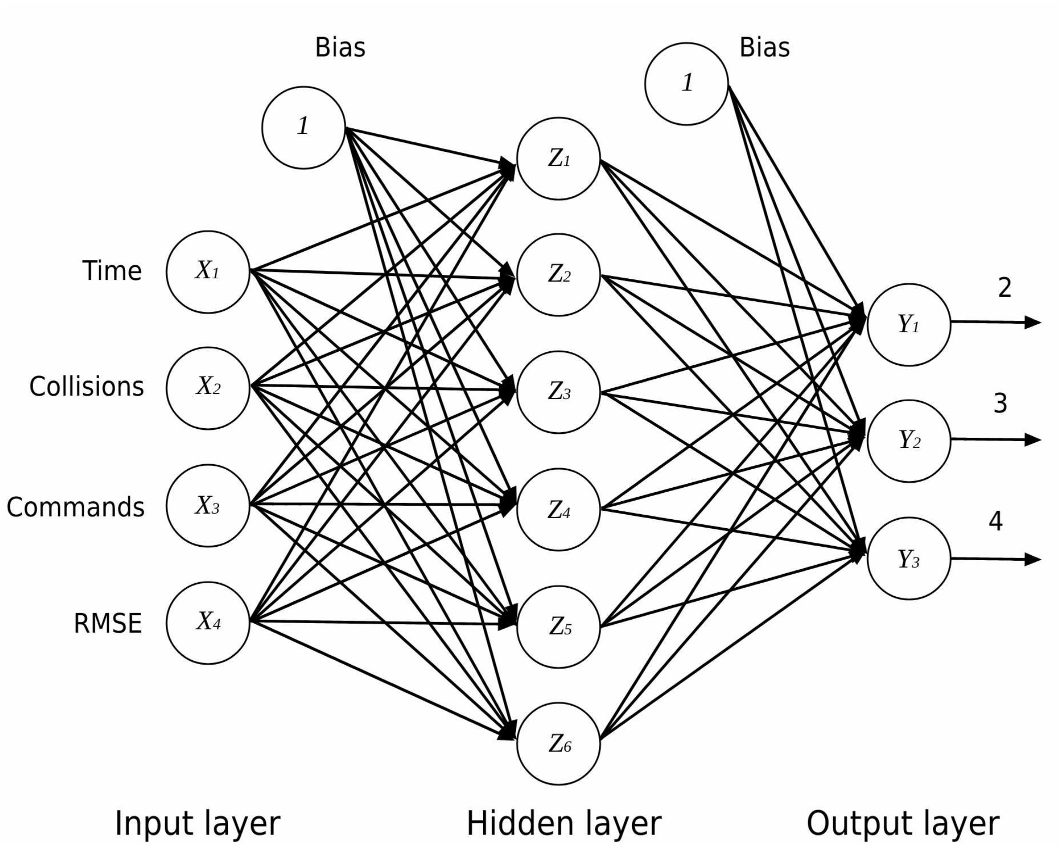

Illustration 6: Preliminary model of the proposed MLP neural network.

Each score in the output layer is given as a probability vector ranging from 0 to 1. In an ideal classification scenario, the values of the three output neurons would be 1, 0 and 0 (respectively) for the score of 2; 0, 1 and 0 (respectively) for scores of 3; and 0, 0, 1 (respectively) for the 4 score. The final score is decided based on the highest value between the neurons (an array of 0.2, 0.5 and 0.3 in the output layer would result in the classification of the score being 3).

For the preliminary MLP model, the error threshold (the difference between the desired output and the output obtained) was established at 0.01, and the number of iterations or epochs as 1000. Both parameters are utilized as reference as when to stop training the artificial neural network. The learning rate used is α=0.5, and the activation function selected for the hidden layer is hyperbolic tangent (tanH) function. The parameters of the MLP were defined experimentally using a controlled dataset with n inputs (n=10, to mimic the expected training conditions for the experiments), where all the labels were to ensure representativity of the outputs. Using the controlled dataset, increasing the number of hidden layers and/or neurons did not seem to present any significant effects in either classification accuracy or processing time. As such, the parameters established initially were kept for the experimental phase.

After the experiments were executed, and therefore real data was made available, the MLP model underwent a second step of redesign. First, to improve data analysis, an external algorithm was implemented with the same parameters instantiated, using the mathematical tool Mathworks Matlab™, as described in Gafford's implementation (Gafford, 2022). The Matlab model was used to experiment different setups with the real-life dataset.

A couple of observations were taken into account regarding other attributes of the MLP: while some of the iterations converged very quickly (less than 200 epochs), others could not achieve convergence even with a larger threshold for number of epochs (10000+). As such, the cutoff value was kept at 1000 epochs. Another empirical observation was that the learning rate, initially set for α=0.5, often caused the weights not to converge at the end of the execution. As such, the learning rate was lowered to α=0.00033, as it had displayed better convergence under most conditions.

Several trials involved increasing the number of neurons in the hidden layer and the number of hidden layers themselves, but no combination showed better or worse accuracy than the one initially set. Considering foreseeable performance slow-down as the dataset grows in size, coupled with no obvious trade-off with regards to accuracy suggested keeping the values already in place. Given the common occurrence of local minimums being used as stopping criteria, the error threshold was removed from the design of the neural network parameters.

The dataset obtained from task 8 (further details about each characteristic of each task are presented in the section below) was used to test different configurations of the neural network. Illustration 7 is a small example of tests performed with the dataset for validation and training.

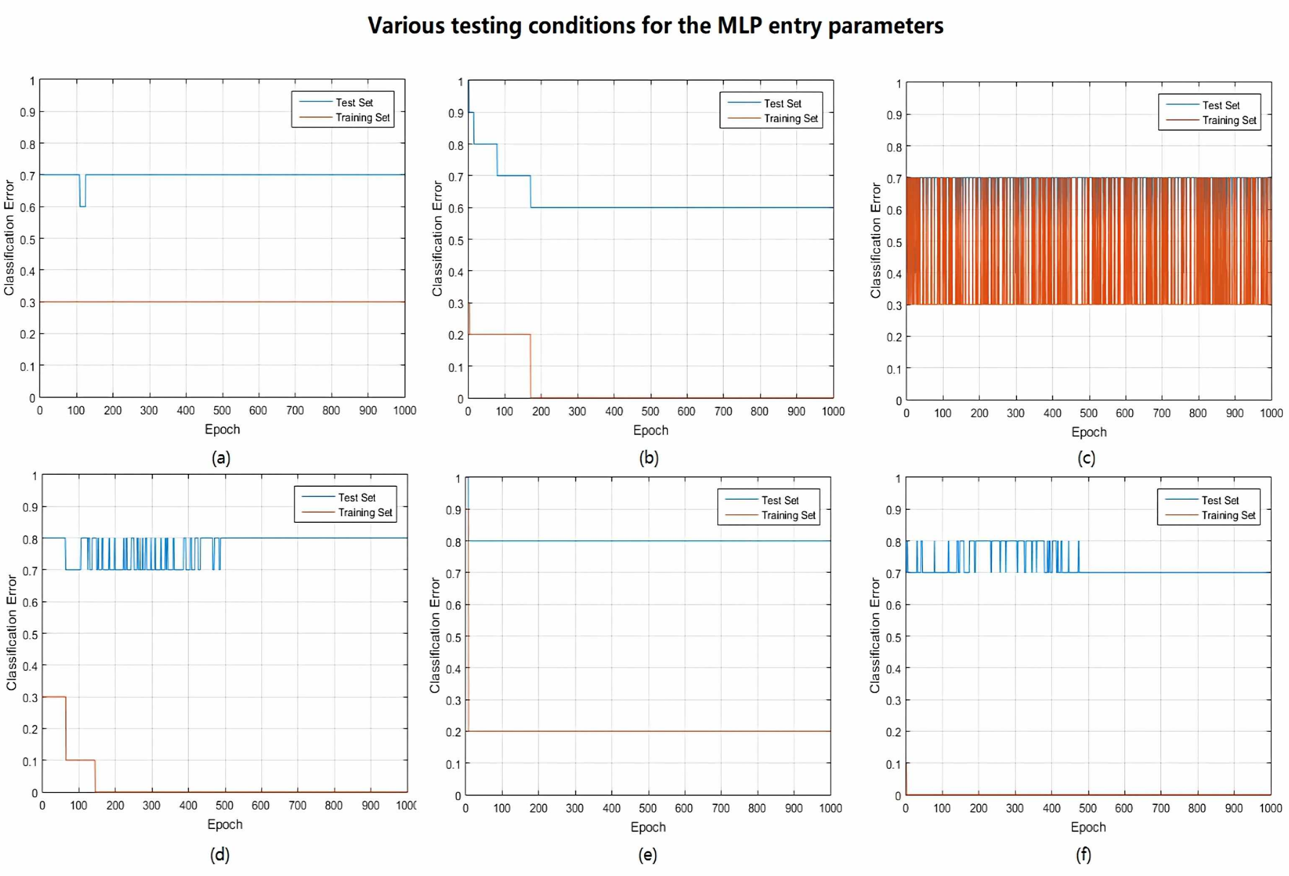

Illustration 7: Several attempts to experimentally increase the fitness of the MLP. Various learning rates were tested: (a) shows very low learning rate at α=0.00001; (b) for medium at α=0.00033; high at (c) at α=0.5; (d) increases the total of hidden layers to three; (e) changes the number of neurons in the hidden layer to 2; and (f) tests an Leaky ReLU activation function.

For the loss function for the back-propagation, categorical Cross-Entropy with Softmax() activation function in the output layer seems to be widely used in classification problems with simpler classification tasks (as opposed to dealing with multiple, more complex classifications at once). Equation (2) shows the calculation of the loss function in the output layer, where a_n^H is the nth neuron in the last layer (H).

$$J_{(x,y)} = -\sum \left(y_n . \ln \left(a_n^H \right) \right) \quad (2)$$

One last consideration with regards to the redesign of the neural network is that, since it is dynamically generated at each initialization of the prediction process, as it is meant to adapt to an ever-increasing dataset, its reproducibility is therefore compromised; the weights used the by neural network are constantly changing values. Moreover, the MLP is calculated differently for each task, since the predictors used to train the algorithm vary from task to task.

Illustration 8 shows the classification results obtained by the MLP post-redesign, using the training and test datasets obtained from the experiments.

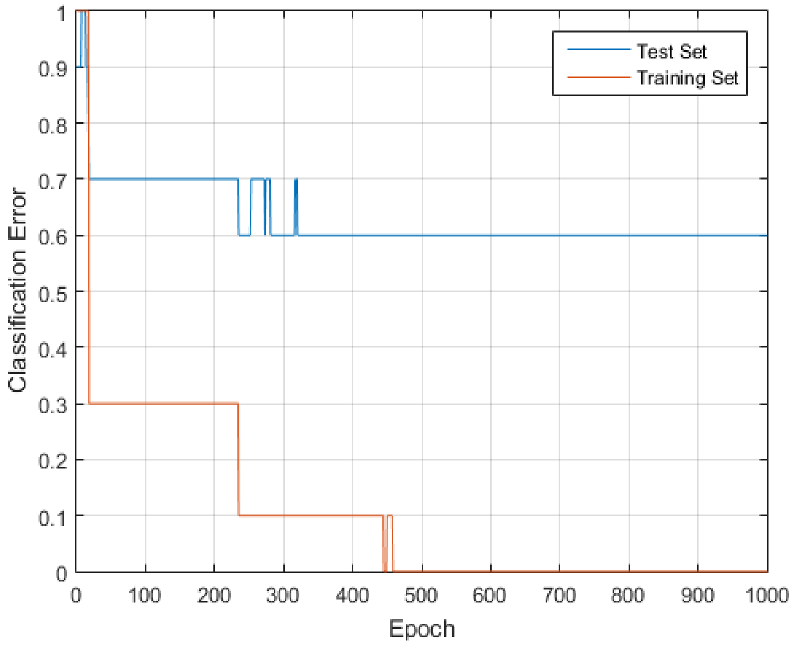

Illustration 8: Testing results of final design for the MLP using the training and test datasets.

4.4. Protocol for wheelchair driving assessment

For the clinical assessment tool selected for the current study, the Power-Mobility Road Test (PMRT) seems to be widely used in several clinical trials, both in the real world and in a simulated environment. It also has displayed favorable results when incorporated in applications with wheelchair simulators (Mahajan, 2012; Kamaraj et al, 2016; Kamaraj, 2020).

Taking into consideration the lack of structured characteristics of the last four tasks, this study was limited to the first 12 structured tasks provided by the PMRT (Massengale et al, 2015). As described by Valentini, reports collected from wheelchair users indicated that the four last tasks of that assessment clinical tool (referred to in the document as unstructured) are, by definition, harder to establish a proper standard, and that could generate bias in the quantitative measurements (Valentini, 2019). As such, the tasks included in this work were adapted from that study, and are discussed in Table 2.

Table 2: Description of the adaptations of the PMRT in the virtual environment

Task number	Descriptor	Design choices in the virtual environment
1	Approaching people/ furniture without bumping into them	Wheelchair starts 2m away from a 1.5mx1.5m square wooden table, positioned in a straight line from the starting point, and the user must approach and park the device properly under the table, while avoiding collision
2	Starting and stopping wheelchair at will	Three coordinates marked by X (first one 1m ahead of the wheelchair; second 1m further and 1m to the right of the first one; and last 1m more meter further and 1m to the left of the second). The user must stand on each marker and stop the wheelchair completely for 5s. After auditory cue, proceed to next marker until the end
3	Crossing doorways without hitting walls (0.9m doorways)	A door frame is placed 1m in a straight line from the starting point. The user must cross the portal and drive straight ahead for 1m more to finish the task
4	Turning around a 90 right hand corner (90 right turn)	User must drive the wheelchair straight for 2m, and after crossing an invisible checkpoint (to ensure the path set will be obeyed), make a 90 degree turn to the right and drive straight for 1m
5	Turning around a 90 left hand corner (90 left turn)	Similar to the above, except the user is required to make a 90 degree turn to the left instead
6	Driving straight forward (4.5m) in an open area	User is required to drive the wheelchair in a straight line until reaching the end of the lane marked with blue tape in the ground
7	Driving straight backwards (3m) in an open area	Similar to the above, except the task is performed in reverse, and therefore visual cues are limited. The user must utilize the surroundings and/or speed to estimate the travel distance necessary to reach the end of the task
8	Turning 180 degrees	Wheelchair must drive straight for 1m, turn 90 degrees to the right and drive straight for 1m more. Then, perform an U-turn in place, completely inverting the forward direction of the wheelchair, and drive forward for 1m. Lastly, turn the wheelchair to the turn for 90 degrees and drive forward for 1m to finish the task
9	Starting and stopping wheelchair upon request	User freely drive the wheelchair for 40s. An auditory cue (chime) will inform the user to stop the wheelchair for 5s. Another chime indicates the start of the next phase, lasting 30s. Again, a chime informs that the wheelchair is required to stop for another 5s. Lastly, another chime restarts the free driving for 50s
10	Turning right and left upon command	Wheelchair follows a straight path for 1m, then turns 90 degrees to the right and drives straight for 1m more, then performs another 90-degree turn to the left. The task is concluded after driving straight for 1m
11	Driving straight forward (4.5m) in a narrow corridor not hitting walls	Similar to Task 6, except 3m-tall wood panels are placed in sequence on both sides of the course, providing a corridor of 1m between, which must be followed while avoiding collision with the panels
12	Maneuver between objects	Six wooden chairs placed at specific distances in two sets ahead of the wheelchair. First set is straight ahead 1.5m, 4.5m and 7.5m of the starting point. Second set starts 50cm to the right of the wheelchair, and is distributed at a distance forward of 0m, 3m and 6m. The wheelchair needs to perform a zigzag pattern to avoid collision with the wooden chairs, while user attempts to reach the end of the task close to the furthest obstacle (7.5m moving straight)

Since Task 9 involves free-form driving, in which users are allowed to move wherever they want in-between stops, it is impossible to determine an optimal trajectory and therefore calculate its deviation; therefore, the RMSE value of the Task 9 was set to null for all participants, as it is non-existent. As such, the RMSE is not taken into consideration during training of the neural network for the task in question. Every other task allowed enough structure to properly measure the time, number of commands, number of collisions and RMSE to train the neural network adequately.

4.5. Wheelchair simulator

In order to simulate the real-world conditions of driving an EPW, a simulator was developed using the game engine Unity 3D, offering a first-person view from the perspective of the user. The simulator was originally designed to test different forms of wheelchair control, similarly to (Hernandez-Ossa et al, 2017), but has since been adapted to include the necessary algorithms to perform the experiments. The simulator was developed within the project: "Multimodal system for distance training in a virtual or augmented reality environment for users of electric-powered wheelchairs" with the unique identifier CAPES/88887.091034/2014-01, and was named Electric-Powered Wheelchair Assessment and Training Simulator (EWATS) (Martins, 2017). The EWATS was further validated in subsequent studies (Martins et al, 2021).

The objective of the simulator is to perform a series of maneuvers over a course with obstacles. The 3D environment used was a generic gym scene, since it was open enough to perform all the tasks without being constricted by lack of space (unless such conditions were part of the tasks themselves), while also being familiar enough so that users could have a better sense of presence. As described in the previous section, each of the 12 tasks presented in the PMRT were modeled inside the environment according to the specifications found in Table 2 (Valentini, 2019). To provide assistance during execution of the task, users were given an option to toggle an UI guidance system, composed of a marker in the display indicating direction and distance from the next checkpoint. Recording of the performance started as soon as the user sent the first command to the wheelchair, signaling the beginning of the task. The virtual environment with Task 12 and the guidance system active can be seen in Illustration 9.

Although the scenario provided ample space for maneuvering, users were encouraged to stay inside the boundaries set by the task (generally represented with blue tapes stretched in the ground, delimiting the course). This was done in order to ensure certain invisible checkpoints would be crossed as the task progressed. The checkpoints were used as reference for measurement of some of the quantitative parameters. Because of this design, those checkpoints could not be skipped, as it would fundamentally change the overall structure of the task. In fact, the guidance system used the coordinates of the checkpoints to provide directions to the user.

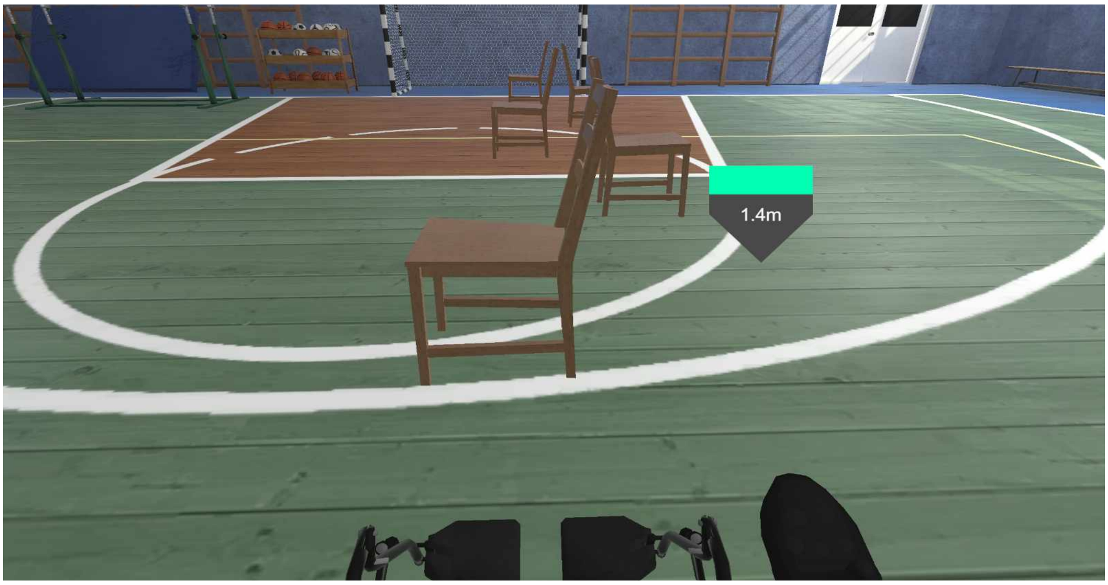

Illustration 9: Example of a PMRT task in the simulator, with the guidance system indicating direction and distance from the next checkpoint.

During the execution of each task, the four performance indicators were collected through internal calculations in the system. Those parameters were then provided as input to the neural network immediately after the completion of the task (unless the user was unable to complete the task, in which case the score in the PMRT would be automatically set to 1). Since by definition the neural network required supervised training before it was able to calculate the final score, a minimum total of 10 tasks were required be executed until completion under supervision of a healthcare professional, and for those 10 initial (or as it was later referred to as the training dataset) the score had to be set manually. However, after achieving the desired size for the training dataset the neural network was able to judge the parameters and determine any subsequent input of information. Pending confirmation (decision-making following the healthcare professional's judgment), newer entries were also designed to be included in the calculation of the final score, with the intent that, as the amount of data increases with usage, so does the accuracy of the classification.

There were several reasons for selecting the EWATS simulator as tool to perform the experimental protocol: first, many of the EPW simulators described are not readily (if at all) available to third-party use, limiting choice of the simulation tool; in second place, even if the usage rights were obtained, those applications are given as black box, as-is models, meaning they are preemptively set up to perform a certain task, and in-depth access to their code and functionalities may be denied. As such, it would be impossible to perform the modifications necessary to incorporate the neural network into the system; and lastly, using a tool previously developed can facilitate acquisition of the performance indicators proposed for the study, as all the details of the implementation are known and easily reachable.

4.6. Subject recruitment

The research involved a total of 20 healthy volunteers of both sexes separated in two groups (training and trial). Individuals with mobility disabilities that require an EPW may be recruited in future iterations of the study. The participants were screened by a healthcare professional to ensure the compliance with the criteria of eligibility before inclusion in the study. The presence of any exclusion criteria resulted in the immediate dismissal of the participant. The sample size was chosen by convenience, following the safety guidelines placed in effect in the indoor environment as a consequence of the COVID-19 pandemic from 2020.

The experimental protocol also requires the presence of one healthcare professional in AT with experience both in using clinical tools to assess the driving capabilities of new wheelchair users and with the scoring in the PMRT, thereby referred to as professional. The role included overseeing both trial and test groups and providing the required score to train and subsequently test the neural network. The professional recruited to participate in the experiment is an occupational therapist with several years of experience in wheelchair prescription, and has undergone extensive training using the simulator and its functionalities as well.

4.6.1. Inclusion criteria

- Subject must be between 18 to 40 years old;
- Subject must not have a history of neurological or musculoskeletal diseases;
- Subject must not have displayed other visual, acuity and perception problems;
- Subject must be able to provide informed consent;

4.6.2. Exclusion criteria

- Subject presents diseases that may compromise their ability to exert approximately 2N of force required to utilize the joystick;
- Subject with history of seizures;
- Subject already participating in another study (clinical trial) involving rehabilitation or use of an experimental drug.

4.7. Research protocol

The experimental protocol proposed in this thesis divides the experiments into two major phases: one for the training group and one of the trial group. A diagram detailing the steps included in the process is shown in Illustration 10.

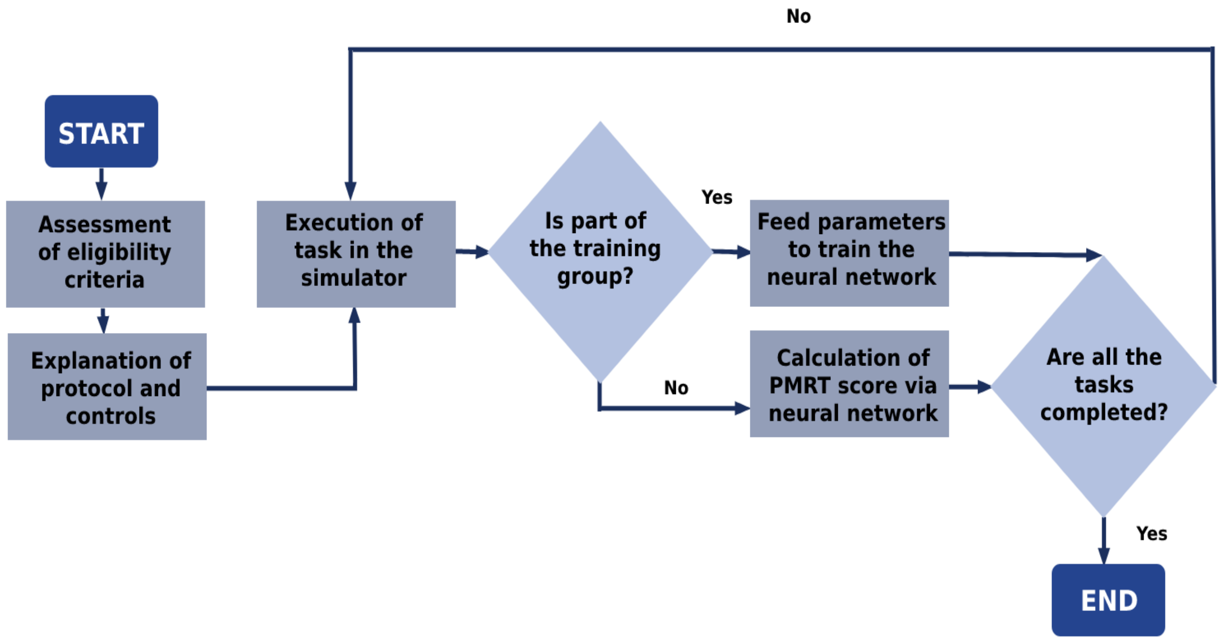

Illustration 10: Diagram of the research protocol executed during experiments.

The first phase of the experiment involved all 10 participants of the training group and the presence of the health professional. Subjects were first given an overview of the system along with instructions as to how to operate the joystick. They were then asked to sit in the Freedom electric-powered wheelchair, which was placed in front of a computer screen. The participants performed 12 tasks in the virtual environment (EWATS) presented in the same order as described in the PMRT, resulting in a total of 120 tasks. During the entire phase, the healthcare professional was seated beside the participant, overseeing the execution of each of the tasks. After each task was successfully completed, the results screen with the variables measured was shown to both participant and professional. At this point, the professional was asked to evaluate the score as per guidelines provided by the PMRT (Massengale et al, 2005). The resulting score was disclosed and subsequently recorded in the system along with the four indicative parameters (time, collisions, commands and RMSE) required to later train the neural network. An unique identifier was also generated to facilitate access to those parameters.

The second phase of the experimental protocol was virtually identical to the first. The 10 participants from the trial group were screened, following an explanation of the objectives of the research and what was required of them. The sequence of which the PMRT tasks were presented in the simulator remained the same, resulting in another 120 trials. Similarly to the first phase, the healthcare professional was required to be present and observe the experiments. However, the rating was performed blind, meaning the score was kept with the professional rather than inserted into the system. The reason behind that decision was to prevent bias created by having prior knowledge of the target score before the fitting and classifying the data using the neural network (that has been trained using the training dataset obtained in the first phase). Only after using the algorithm to predict the scores the access to the real scores established by the health professional was granted for comparison.

4.8. Data analysis

As stated in the previous section, there will be two datasets resulting from the execution of the experimental protocol, one dataset for training and one dataset for testing. The classification accuracy will be calculated using Matlab, as well as statistical variables such as average and standard deviations.

Aside from the MLP implementation in the Matlab, two other classification algorithms will be designed using the mathematical tool as well, in order to compare classification accuracy of the MLP model created. The two classification algorithms are: Support Vector Machine (SVM), supervised learning models characterized for being non-probabilistic binary linear classifiers; and *k*-Nearest Neighbors (KNN), an algorithm that relies on distance for classification of its data. Both are common methods of classification using machine learning techniques, serving as reference for comparison with the MLP developed in the study. A statistical test (10x2-fold cross-validation with the Wilcoxon Signed-Rank test) will be performed to validate the statistical significance of the results. The implementation of the test will be described in the following sections.

5. RESULTS

The demographics of the subjects of the study are summarized in Table 3. All the participants were recruited from the Federal University of Uberlândia. The experiments were conducted between November and December 2022. The occupational therapist responsible for overseeing the trials has at least 12 years of experience in EPW provision to new users.

Table 3: Demographic information about population

Demographics		**Training**	**Test**
Participants	Male	3	7
	Female	7	3
Total participants		10	10
Average age (years±SD)		24.5±5.3	23.2±2.6

All the parameters obtained from the participants during both phases of the experiments are listed in Table 4. To preserve their identity, the subjects are assigned an unique identifier (TR01-TR10 for the training phase group and TE01-TE10 for the test or trial phase group). The parameters displayed in the table are the sum of variables obtained in each of the 12 tasks. For example, subject TR08 displayed 5 collisions across all tasks. The composite score is the normalized value of the total score obtained, adjusted for 12 tasks rather than16 tasks originally present in the PMRT.

Table 4: Performance parameters for participants of the research

Subject ID	**Time elapsed (s)**	**Collisions**	**Commands**	**RMSE**	**Adjusted composite score**
TR01	402	2	343	51.86	0.96
TR02	466	3	512	150.70	0.94
TR03	415.5	0	339	416.12	0.90
TR04	413.4	3	600	16.60	0.92
TR05	328.7	1	236	69.80	0.96
TR06	449.6	0	399	134.17	0.96
TR07	483.8	0	535	619.37	0.92
TR08	450	5	416	13.64	0.92
TR09	437.3	1	247	459.36	0.94
TR10	411.3	0	333	465.53	0.94
TE01	426.2	3	619	27.98	0.92
TE02	370.4	0	285	19.24	0.98
TE03	474.7	5	557	39.66	0.94
TE04	465.2	0	416	210.64	0.96
TE05	443.8	5	595	1125.40	0.88
TE06	387.4	3	353	27.36	0.94
TE07	453.6	8	565	2688.85	0.81
TE08	362.4	5	290	348.80	0.92
TE09	583.8	2	709	2708.72	0.73
TE10	395.5	0	338	79.23	0.98

In order to compare the classification capabilities of the model proposed using the data collected during the experiments, two other classification models were trained using the same dataset as the simulator, aside from the previously established MLP created with the same parameters to facilitate fine-tuning of characteristics. The first one is a SVM (in which the kernel function is set to linear), and the second function is a KNN (set for euclidean distance with number of neighbors=1).

Illustrations 11-22 show the results of the classification with all four algorithms for each of the tasks in the form of confusion matrices, where the predicted values are plotted against the real value, which, as stated, was given access to only after the experiments were concluded and the neural network was trained so the comparison could be made without generating bias. Since the structure of each task may differ from one another, it made sense to analyze each task separately, since one different model was created using each dataset from every one of the 12 tasks. Illustration 11, for instance, shows that all models made the same prediction, where two instances of which the true score was 3 were predicted as 4.

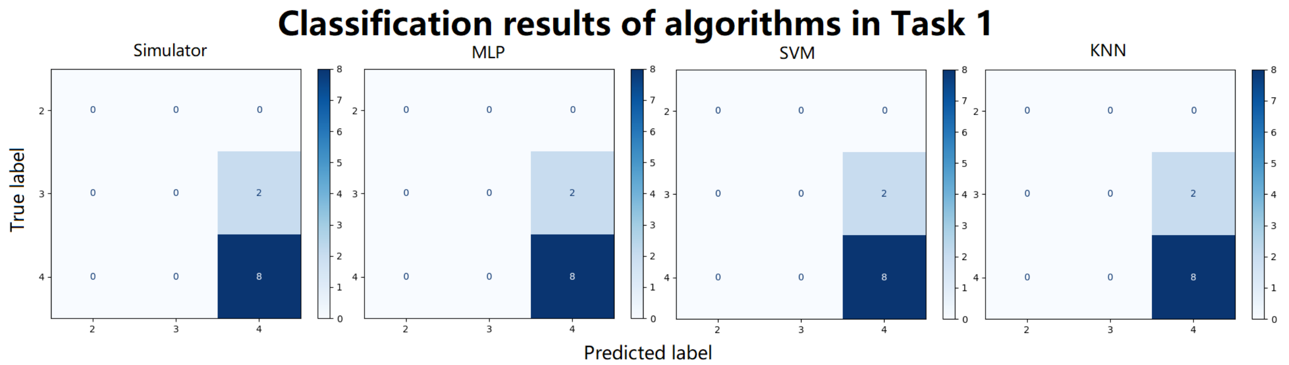

Illustration 11: Confusion matrices for all four algorithms tested for Task 1.

Illustration 12 shows variation between the predicted scores for Task 2. Both the MLP implemented in the simulator and the KNN achieved similar predictions. The SVM, on the other hand, had a much lower prediction rate.

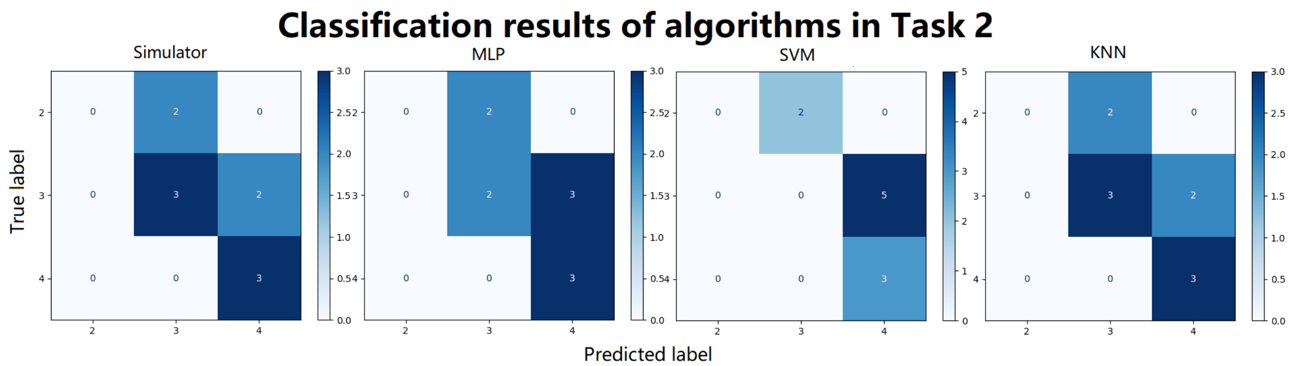

Illustration 12: Confusion matrices for all four algorithms tested for Task 2.

Illustration 13 once again showed similar prediction results between all four algorithms in Task 3, where users were asked to drive across a door frame without colliding.

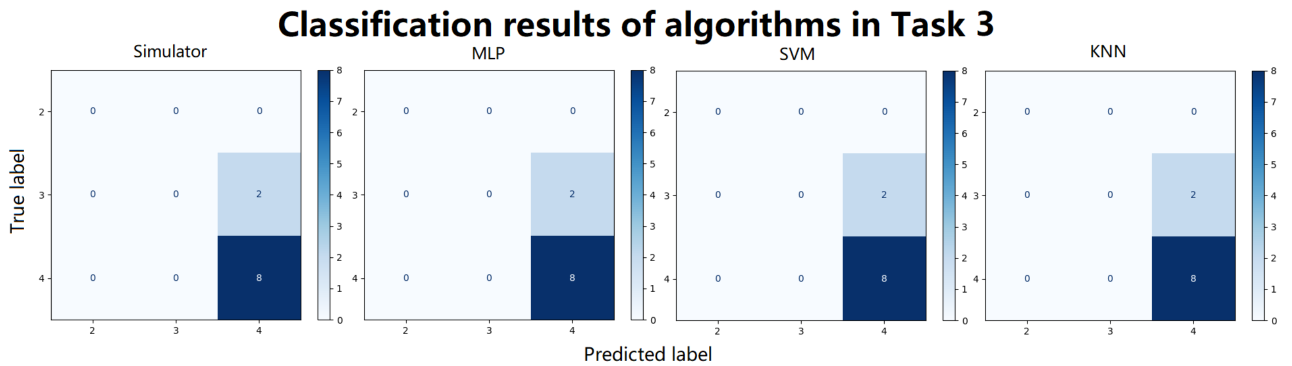

Illustration 13: Confusion matrices for all four algorithms tested for Task 3.

The behavior displayed was repeated in Task 4, as seen in Illustration 14. Yet, no algorithm was capable of achieving perfect accuracy in predicting the tasks so far.

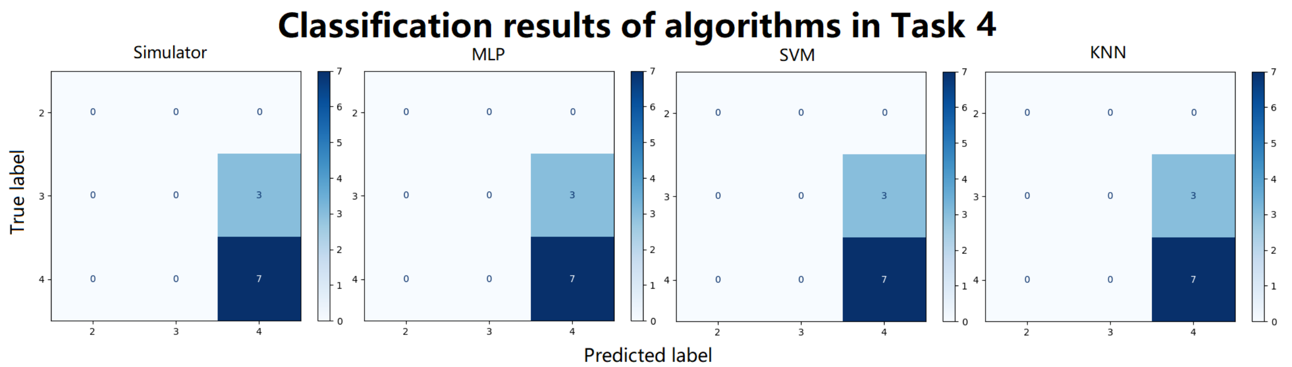

Illustration 14: Confusion matrices for all four algorithms tested for Task 4.

Illustration 15 shows an oddity with regards to the classification skills, and, possibly, to the dataset itself. Despite being virtually the same task as the previous one (except users were asked to turn left rather than turn right), the prediction models showed discrepancy in the results, as opposed to what was seen in Illustration 14.

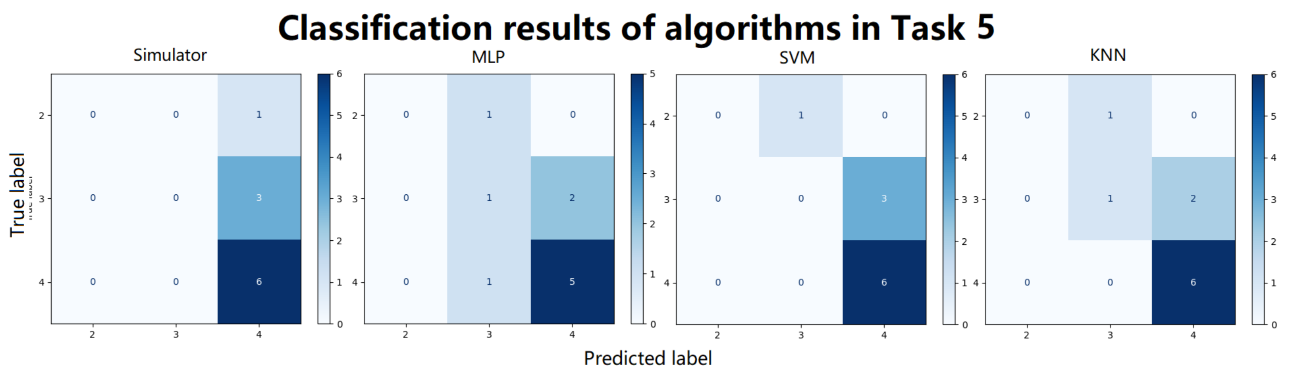

Illustration 15: Confusion matrices for all four algorithms tested for Task 5.

Illustration 16 indicates similar classification results between three of the four models: the MLP in the simulator, the MLP designed for testing, and the SVM. KNN, however, showed poor accuracy when compared with the previous models, only providing the correct score 4 out of the 10 times for participants of the trial group.

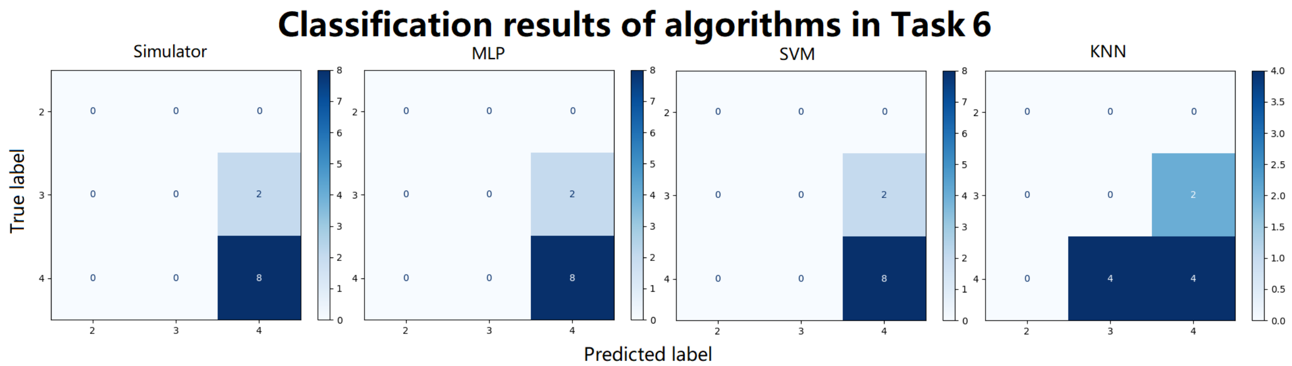

Illustration 16: Confusion matrices for all four algorithms tested for Task 6.

The KNN and SVM obtained similar results in classifying the dataset from Task 7. While the resulting accuracy was the same between both MLP models, the distribution of predictions is different between them, as seen in Illustration 17.

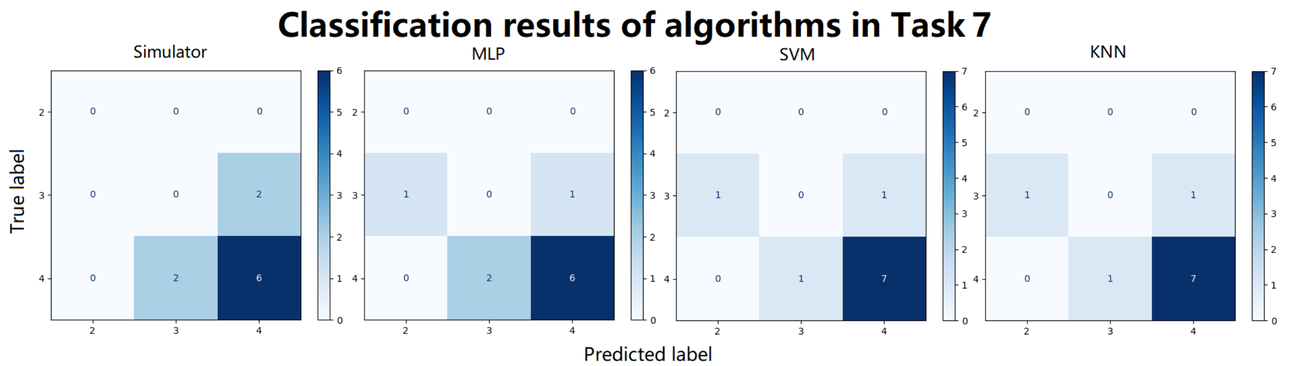

Illustration 17: Confusion matrices for all four algorithms tested for Task 7.

Illustration 18 indicates the performance of the models in predicting the score for Task 8 was poor, with a maximum of 5 out of 10 obtained by the KNN. The instructions for completing Task 8 were rather simple (turning 180 degrees), but the structure of the task itself was complex and could have contributed to the generally lower scores. However, it does not change the fact that the algorithms were not able to account for the low scores in their predictions.

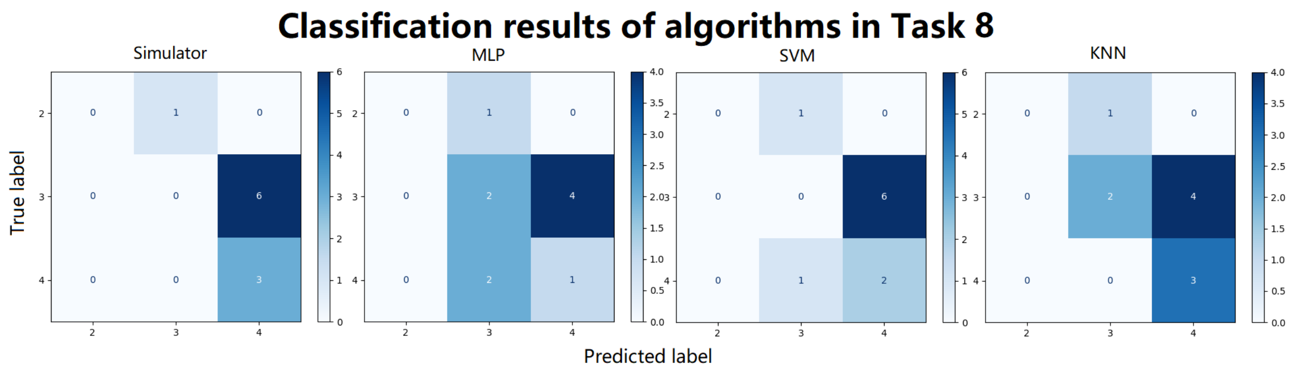

Illustration 18: Confusion matrices for all four algorithms tested for Task 8.

Task 9 prediction results were calculated differently for each of the machine learning algorithms as well, including the two similar MLP models. In fact, the MLP algorithm implemented in Matlab achieved the best classification accuracy out of all of them, despite not being able to correctly predict the instances in which the true score was 3. This behavior can be seen in Illustration 19.

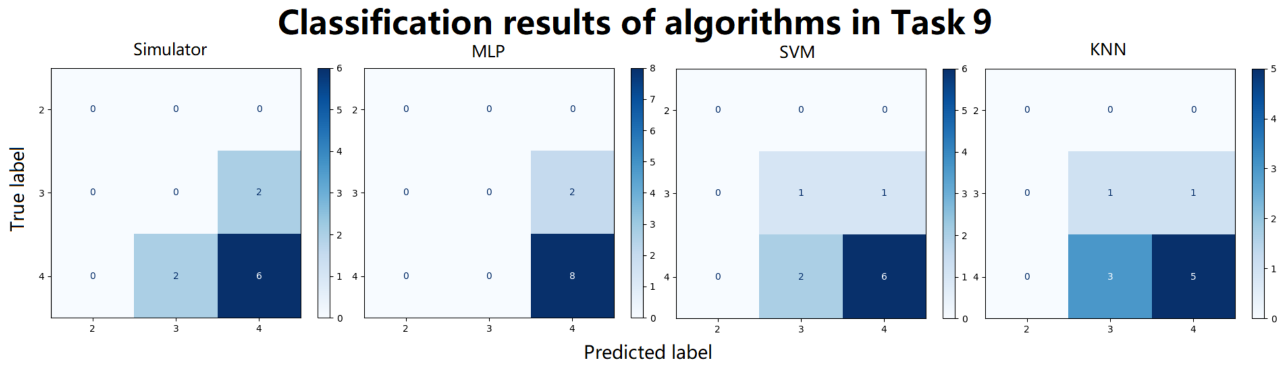

Illustration 19: Confusion matrices for all four algorithms tested for Task 9.

Illustration 20 shows that all models achieved the same general accuracy, and that both pairs between simulator and KNN, and MLP and SVM displayed similar classification tendencies for Task 10.

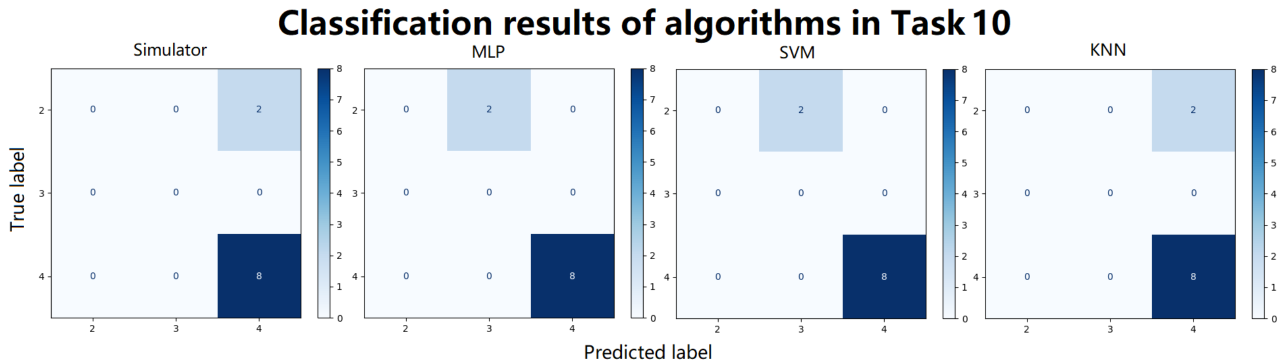

Illustration 20: Confusion matrices for all four algorithms tested for Task 10.

Illustration 21 showcases a situation in which both MLP resulted in the same predictive behavior, whereas the SVM and KNN had a lower accuracy for the data in which they were tested. 9 out of 10 was also the highest accuracy value displayed so far including all of the models.

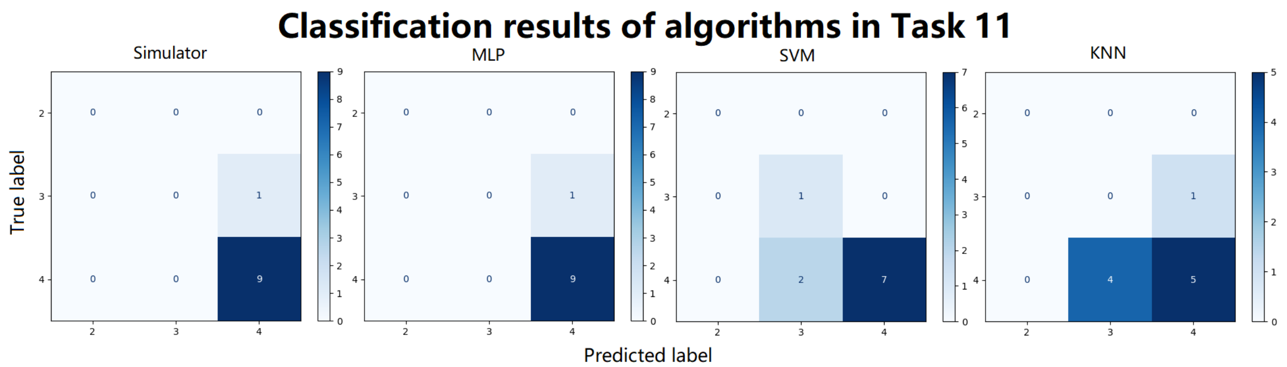

Illustration 21: Confusion matrices for all four algorithms tested for Task 11.

Task 12 was another case in which the prediction accuracy was low for all the algorithms, particularly for the MLP and the KNN. Given the complexity of the task, which included maneuvering between obstacles, the results assigned by the professional leaned on the average side (6 people obtained the score of 3), as shown in Illustration 22.

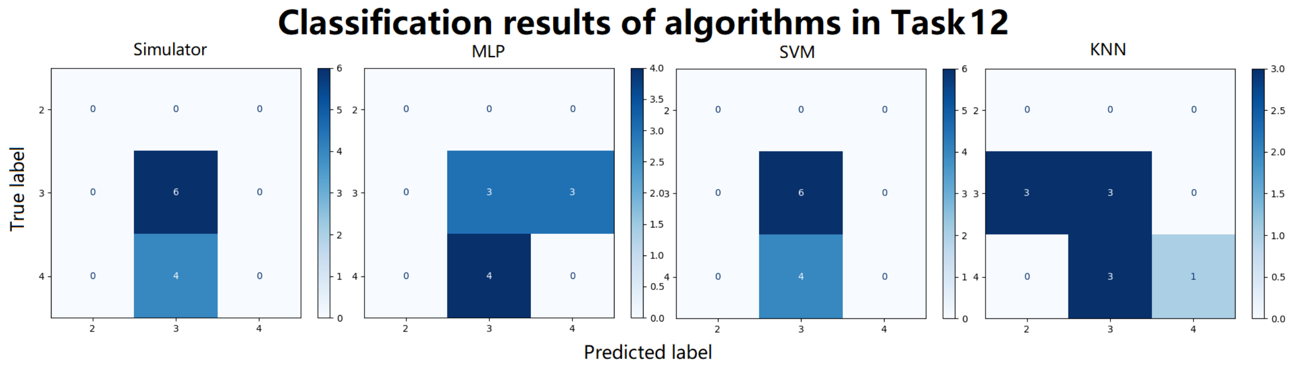

Illustration 22: Confusion matrices for all four algorithms tested for Task 12.

While the confusion matrices displayed in the previous figures allow more detailed analysis of the individual behavior for each of the classifiers included in this research after being trained with the dataset obtained in the training phase (along with the target score provided by the healthcare professional), the information presented in Table 5 provide an general overview of the results obtained for all 12 tasks with each of the models. The prediction accuracy is indicated by how many times the predicted score matched the target score in the trial (or test) dataset.

The prediction accuracy between all models are the same in Tasks 1, 3, 4 and 10, although, as it can be seen in Illustration 11, Illustration 13, Illustration 14 and Illustration 20, only in the first three tasks the prediction was the same. They are also cases in which the score 4 was predicted for all the trials. Table 5 also lists the average accuracy for each model, calculated from the accuracy obtained from all tasks. While the results are comparable, a statistical test would be required to determine with significance if either of the predictive models is superior to the others, which will be described below.

Table 5: Prediction accuracy table for all algorithms

Task number	Simulator (%)	MLP (%)	SVM (%)	KNN (%)
Task 1	80	80	80	80
Task 2	30	50	30	60
Task 3	80	80	80	80
Task 4	70	70	70	70
Task 5	60	60	60	70
Task 6	80	80	80	40
Task 7	60	60	70	70
Task 8	30	30	20	50
Task 9	60	80	70	60
Task 10	80	80	80	80
Task 11	90	90	80	50
Task 12	60	30	60	40
Overall accuracy	65	65.8	65	62.5

Another relevant information obtained from the experiments is regarding the final weights calculated by the MLP. As stated during the section explaining the decisions behind the parameter values selected for the design of the algorithm, it is not possible to provide a single value for each of the weights on each neuron because the MLP itself is calculated dynamically. To verify that statement, several trials were performed using the same training dataset and tested against the same trial dataset as before, obtained from Task 10. Out of those, 25 trials were selected in which the prediction results matched for all participants (even if they were incorrect). The final weight for all neurons in those 25 trials were then extracted and organized in two tables. Table 6 lists the average and standard deviations from the neurons between the input layer and the hidden layer (with each neuron identified as H1-H6). Considering the standard deviation value is higher than the average in all of the neurons, it indicates it is nearly impossible to predict the final weights since they are constantly shifting between iterations of the neural network.

Table 6: Estimate mean and standard deviation of weights between input layer and hidden layer 25 samples taken from Task 10

Hidden layer	Time (weight±SD)	Collision (weight±SD)	Command (weight±SD)	RMSE (weight±SD)
H1	-0.09±0.36	-0.05±0.25	-0.12±0.40	-0.04±0.45
H2	0.00±0.38	-0.14±0.30	0.06±0.36	0.03±0.50
H3	-0.16±0.33	0.04±0.32	-0.11±0.43	-0.12±0.49
H4	0.01±0.49	0.07±0.31	0.08±0.31	-0.06±0.44
H5	0.04±0.36	0.00±0.32	0.14±0.45	-0.11±0.46
H6	-0.02±0.31	-0.04±0.24	-0.04±0.41	-0.10±0.46

Table 7 shows exactly the same behavior with the weights of the neurons between the hidden layer, once again identified as H1-H6, and the output layer with the prediction results (2, 3 and 4). The 1 value does not figure in the output layer since, as explained previously, it refers to instances in which the user is unable to complete the task. As such, including the quantitative parameters of incomplete tasks in the dataset calculations would contaminate the results, since they could display abnormally low values of commands, time, number of collisions or RMSE (depending on the reason why the participant was unable to perform the task.

Table 7: Estimate mean and standard deviation of weights between hidden layer and output of the classifier 25 samples taken from Task 10

Output	H1 (weight±SD)	H2 (weight±SD)	H3 (weight±SD)	H4 (weight±SD)	H5 (weight±SD)	H6 (weight±SD)
2	0.53±2.00	-0.06±2.01	0.79±1.85	-0.06±1.81	-0.04±1.99	0.23±1.84
3	0.32±0.81	-0.03±0.86	-0.10±0.97	0.01±0.87	-0.35±0.88	-0.20±1.10
4	-0.22±0.84	-0.00±1.06	-0,13±1.03	0.11±1.04	0.50±0.88	0.13±1.27

In order to validate the differences between the classifiers, a statistical test is required. One of the most commonly and recommended methods of statistical analysis for comparison between classification algorithms is the 5x2-fold cross-validation with paired Student's *t*-test, which consists in randomly dividing the whole dataset (all 20 entries from all participants) in half (so that the number of folds is *k*=2) and using one half to train the algorithm and the other to calculate the accuracy, and doing the opposite (using the second half to train and the first to test the classification). The process is then repeated five times (hence the 5x2-fold) and the differences between accuracy scores are used in a paired t-test.

However, one of the assumptions of the test is that the distribution of the data is normal, which is not the case for the dataset established in the experiments. As such, it was necessary to use the non-parametric equivalent of the paired *t*-test, the Wilcoxon Signed-Rank test. To have enough samples to utilize the Wilcoxon statistical test, the 2-fold processing of the data was repeated 5 more times, resulting in a 10x2-fold cross-validation with the Wilcoxon Signed-Rank test. The results of the accuracy obtained by each of the algorithms, along with the results of the Wilcoxon test are shown in Table 8.

Table 8: Statistical analysis using a 10x2-fold cross-validation with Wilcoxon Signed-Rank test

	Average accuracy (%)			**MLP X SVM**		**MLP X KNN**		**SVM X KNN**	
	MLP	**SVM**	**KNN**	**h_0**	**p-value**	**h_0**	**p-value**	**h_0**	**p-value**
Task 1	84	88.5	78.5	*failed*	0.09	*reject*	0.03*	*reject*	<0.01*
Task 2	45.5	57.5	52.5	*reject*	0.03*	*reject*	0.04*	*failed*	0.30
Task 3	85.5	89	82.5	*failed*	0.29	*failed*	0.23	*reject*	0.02*
Task 4	76	88	71	*reject*	<0.01*	*reject*	0.04*	*reject*	<0.01*
Task 5	58.5	77	71.5	*reject*	<0.01*	*reject*	<0.01*	*failed*	0.38
Task 6	71	86	73	*reject*	<0.01*	*failed*	0.66	*reject*	<0.01*
Task 7	53	72.5	63.5	*reject*	<0.01*	*reject*	<0.01*	*reject*	0.04*
Task 8	43	58.5	50	*reject*	<0.01*	*reject*	0.03*	*failed*	0.08
Task 9	76	78.5	62	*failed*	0.49	*reject*	<0.01*	*reject*	<0.01*
Task 10	85.5	94	83.5	*reject*	0.02*	*failed*	0.13	*reject*	<0.01*
Task 11	82.5	92	85	*reject*	<0.01*	*failed*	0.27	*reject*	<0.01*
Task 12	53.5	69.5	44.5	*reject*	<0.01*	*reject*	<0.01*	*reject*	<0.01*
Overall accuracy	67.83	79.25	68.16						

**Statistically significant, p<.05*

All the information (datasets, algorithms, statistical tests, among others) are stored for the sake of the reproducibility of the study, as well as for any future applications involving the research tools and procedures.

The interface displaying the results of performance in each task can be seen in Illustration 23. After the data is loaded, the interface displays when the task started, when it was concluded, how long it took to complete the task, how many commands were given to the joystick, how many collisions happened along the course, the value of the RMSE, and if the session was concluded. It also displays how to the performance was scored in the PMRT as a suggestion, and the rehabilitation professional can update the score in case it was calculated incorrectly; this new updated score will then be used to train the neural network in future iterations of the simulator.

Illustration 23: Interface displaying the results of performance post-task.

6. DISCUSSIONS

The Wilcoxon Signed-Rank test provided statistically significant information regarding the classification capabilities of the three methods discussed (since both MLP implementations were similar). The paired *t*-test initially generated statistically impossible results; that was what actually suggested the population did not have a normal distribution, which in turn would make impossible to obtain a statistical analysis as the data was non-parametric. The Wilcoxon test was sought out as a non-parametric alternative to the Student's *t*-test, and it only failed to give results of significance because of limitations of the sample size. The test was originally developed by Wilcoxon (1945) and further improved by Mann & Whitney (1947), and it has since been used when the population does not present a normal distribution.

In fact, it was observed that the sample size was the main limitation of the study, which will be discussed in full shortly. Having generated enough datasets using the *k*-fold cross-validation technique, it was then possible to determine the relevant findings of the experiment. The 2-fold cross-validation was repeated 10 times, generating a total of 20 different datasets (in which one half mirrors the other half). The Wilcoxon test itself is based on the null hypothesis that there are no significant differences between the models. As such, the comparisons presented in Table 8 determined if the conditions for rejecting the null hypothesis were met.

The results listed in Table 8 showed that in all cases, the accuracy results of the SVM (Linear Support Vector Machine) were superior to both the Multi-layer Perceptron and the *k*-Nearest Neighbors. The statistical significance of those results were proven for almost all the tasks (for Tasks 2, 4, 5, 6, 7, 8, 10, 11 and 12 between the MLP and SVM and Tasks 1, 3, 4, 6, 7, 9, 10, 11 and 12 between the KNN and the SVM) at a significance level of 5%. The comparison results between the MLP and KNN were mixed, with some of the task scores predicted better using the MLP and some with the KNN. While it is not possible to affirm with confidence that the MLP was better at predicting the scores than the KNN, it is certainly true that the SVM outperformed both of the others for all tasks, reaching a total accuracy score of almost 80%. However, there are attributes chosen during both phases of design for the MLP, such as randomly generated weights, activation function or even absence of a momentum function (for example, the ADAM algorithm) (Kingma & Ba, 2014), that could improve the performance of the MLP to a level comparable to the accuracy obtained by the SVM model.

Even the best average accuracy achieved (80%) would not be considered ideal to configure the method itself as entirely reliable, where accuracy values of 95% would be considered a good cutoff for prediction systems. This led to the analysis of the dataset itself, and some particular details were of note. The first and most important was regarding the representativity of the classification data. For example, in Task 1 (which consisted of driving in a straight line and parking the wheelchair under a wooden table placed 2m away from the starting point) all the participants of the training phase received a score of 4. As such, when during the trial phase two of the 10 participants were scored with 3, none of the classifiers were able to recognize the value of the indicative parameters that would lead to such classification. In fact, as they never trained with any target score different from 4, every single result they would ever produce (and this argument remains true for any classification algorithm) would be a score of 4.

As it can be observed from the statement above, the lack of enough evidence for each class or score was a problem observed in all of the tasks in the training dataset, which originated from the sample size itself. The only viable solution for such a problem would be to increase the labeled data, which is the dataset used for training, making sure there are enough cases for each class so that the classification model is able to identify what significant characteristics in the input parameters would lead to each of the scores, and adjust itself accordingly.

With regards to the scores themselves, the ceiling effect of the PMRT mentioned by both studies of Mahajan (2012) and Kamaraj (2020) was observed. A total of 80 tasks in the labeled data (testing dataset) were scored as 4 according to the PMRT (Massengale et al, 2005) (representing 2/3 or 66.67% of the sample). However, the quantitative parameters on those tasks with a score of 4 varied considerably (even among the data obtained from the same task). That could present a challenge to any of the classification methods available, as it becomes difficult to determine a properly mathematical model that can represent the significance of each variable in the classification process.

The Tables 6 and 7 exemplify an important characteristic of the design of the MLP. The weight that correlates the relevance of the relationship among neurons (between the input and the hidden layer in Table 6 and between the hidden layer and the output in Table 7) is generated dynamically for each iteration of the neural network. This variability can be attributed to the random values given to the initial weights (between -0.5 and 0.5). Since the training is performed whenever the MLP is executed, a new set of initial weights is generated each time. As such, even if the classification results are similar between two executions of the MLP, the weights that define the relationship between each neuron and the output vary, significantly altering the convergence properties of the neural network. Given the high variability of the values shown in both tables, it is hard to establish, based on those results alone, the significance of the relationship of each of the quantitative parameters. The present study considered the time, number of collisions, number of commands, and RMSE, which consists of a variation of the QDM proposed by Kamaraj (2020), as the input parameters. The possible outputs for the classifier were the scores of 2, 3 or 4, which depend on the results obtained from the executions. Other studies could analyze this relationship between input and output, and its results could serve as reference to identify the best parameters to be measured that can determine the score in the PMRT.

As the focus of the experiments was given exclusively to the quantitative parameters and accuracy of classification methods, observations regarding usability factors were not taken into consideration for this step of the research. All of the notes provided by the subjects during the execution of the experimental protocol were archived to be discussed in further studies. They could provide critical information on how to improve the tests and the wheelchair simulator's interface, which could contribute to an increase in the sense of presence felt by participants during use.

7. FINAL CONSIDERATIONS

This thesis examined the possibilities of integrating neural network techniques into clinical assessment tools for automation of the individual's scoring based exclusively on quantitative variables. The use of neural network (and, as a matter of fact, any machine learning method) has grown exponentially in the last few decades, because of their capacity of providing fast and reliable solutions to problems of classification, while not being affected by human factors, such as emotional or psychological states, fatigue, lack of experience or judgment, among others. Wheelchair prescription is a laborious and time-consuming process, and it depends on availability of healthcare professionals with credentials to perform the assessment. Incorporating a certain degree of automation into the process can expedite the assessment, and make it easier to grant wheelchairs to individuals that require the AT device for their mobility.

Various works that are in some manner correlated with the research conducted in this document were described, including their goals, strengths and limitations, and were all relevant to guide the methodology proposed. While some of those systems are no longer in development, at least according to the literature review, several others are constantly under improvement, with new applications and findings being presented given the growing interest in the field. As demonstrated in Sá et al (2022), the amount of papers published following the theme more than tripled in the past decade or so, in comparison with the previous decade. This growing interest can be justified by diverse reasons, such as improvements in wheelchair and VR technology and increased demand for assistive technologies due to the aging of the population.

The hypothesis that performance indicators such as time and number of collisions could be used to classify the data with enough accuracy was met, albeit partially, given limitations on the sample size, thus invalidating the alternative hypothesis. The SVM algorithm seemed particularly relevant according to the statistical analysis, but the MLP also displayed its capabilities of classifying the dataset obtained in this research. Further studies are required with a bigger sample and better representativity of the data. In fact, in accordance with the characteristics of the MLP developed, the dataset that was generated as a result of the test phase can now be used as part of the training dataset as well. The expected behavior is that, with the increased sample size and more representativity of the classification targets, subsequent experiments could achieve even higher accuracy results, which is a common characteristic among classifiers.

The neural network chosen for the development of the prediction system was the MLP, which is one of the most commonly used algorithms to solve classification problems. As there is no gold standard on which parameters to use in designing a neural network, the MLP parameters were initially selected based on experimental trials performed with a fictional dataset. After the experiment protocol was performed, the MLP underwent a process of redesign, using the real training dataset obtained from the first phase of the experiment. While some adjustments were made (in particular to the learning rate and error threshold), no statistically significant improvements were observed in the classification results between before and after the redesign. However, some design characteristics were either not implemented (such as momentum function, as the MLP uses a backpropagation algorithm), not considered fully (finding a fixed set of initial weights to obtain consistent training even with changing dataset) or could be changed or improved (such as the activation function selected in the design). The implementation of the changes described could further improve the accuracy of the MLP, indicating the potential of the algorithm in the classification of the results of the PMRT.

Under the current training circumstances, the Support Vector Machine classifier achieved the highest classification accuracy with using the 10x2-fold cross-validation method between all three algorithms (MLP, SVM and KNN), obtaining a total of 79.25% accuracy (against 67.83% from the MLP and 68.16% for the KNN). Considering that the attributes used in the SVM algorithm were standard, since the intended goal was to use to compare the accuracy with the MLP implemented in the simulator, an alternative and concurrent analysis may be conducted aimed at fine-tuning the design parameters of the SVM, followed by more studies to test for improvements in its classification capabilities. The reason given to why such analysis was not included in this research was a result of the study constraints, time being one of the most relevant ones.

Future works may include, but are not limited to, tests with an increased sample size that includes better representation of the scores and improvements on the framework (simulator) which was used during the experimental protocol. This divides the focus between two major contents: relevance of the scientific contribution and improvement of the experience for the users (and rehabilitation professionals). The previous is defined by statistical analysis with the larger sample to further validate the performance indicators selected, both in comparison with varied classification methods available in the literature, and to validate the choice of the performance indicators themselves. While some of the continuous variables described in the study of Sá et al (2022) could pose a challenge to classification methods that use discrete values as input in the dataset, several others can be included to generate different models based on various combinations of quantitative parameters. The latter, however, requires a study of usability using the proper methodologies (such as Nielsen's heuristics) to assess the user interface design and how it could be improved in terms of presentation and sense of presence.

The final aim of the research is to create a reliable method of assessment that can then be provided to healthcare professionals and healthcare services (such as rehabilitation centers) to facilitate the access to electric-powered wheelchair devices to all users in need of the technology, and increase their ability to perform the activities of daily living and, as a consequence, their quality of life in general. The experiments performed with the MLP indicated that, under the correct circumstances, it is possible to achieve the expected results.

REFERENCES

Abellard, P., Randria, I., Abellard, A., Ben Khelifa, M.M., & Ramanantsizehe, P. (2010). Electric wheelchair navigation simulators: Why, when, how? *Mechatronic Systems Applications*. https://doi.org/10.5772/8927

Archambault, P.S., Routhier, F., Hamel, M., & P. Boissy. (2008). Analysis of movement to develop a virtual reality powered-wheelchair simulator. 2008 Virtual Rehabilitation, 133–138. https://doi.org/10.1109/ICVR.2008.4625149

Archambault, P.S., Chong, J.N.F., Sorrento, G., Routhier, F., & Boissy, P. (2011a). Comparison of powered wheelchair driving performance in a real and in a simulated environment. 2011 International Conference on Virtual Rehabilitation, 1–7. https://doi.org/10.1109/ICVR.2011.5971807

Archambault, P.S., Tremblay, S., Cachecho, S., Routhier, F., & Boissy, P. (2011b). Driving performance in a power wheelchair simulator. *Disability and Rehabilitation: Assistive Technology*, *7*(3), 226–233. https://doi.org/10.3109/17483107.2011.625072

Archambault, P. S., Blackburn, É., Reid, D., Routhier, F., & Miller, W.C. (2016). Development and user validation of driving tasks for a power wheelchair simulator, Disability and Rehabilitation, 39:15, 1549–1556. https://doi.org/10.1080/09638288.2016.1226423

Arlati, S., Colombo, V., Ferrigno, G., Sacchetti, R., & Sacco, M. (2019). Virtual reality-based wheelchair simulators: A scoping review. *Assistive Technology*, *32*(6), 294–305. https://doi.org/10.1080/10400435.2018.1553079

Batavia, M. (2010). The Wheelchair Evaluation: A Clinician's Guide. (2nd ed.) Jones and Bartlett

Bigras, C., Kairy, D., & Archambault, P.S. (2019). Augmented feedback for powered wheelchair training in a virtual environment. J NeuroEngineering Rehabilitation 16, 12. https://doi.org/10.1186/s12984-019-0482-3

Bourhis, G. & Agostini, Y. (1998). The VAHM Robotized Wheelchair: System Architecture and Human-Machine Interaction. Journal of Intelligent and Robotic Systems, 22(1), 39–50. https://doi.org/10.1023/A:1007934111358

Cooper, R.A. (1995). Rehabilitation engineering applied to mobility and manipulation. Bristol: Institute of Physics Publishing

Cooper, R. [Rory], Cooper, R. [Rosemarie], Tolerico, M., Guo, S., Ding, D., & Pearlman, J. (2006). Advances in electric-powered wheelchairs. *Topics in Spinal Cord Injury Rehabilitation*, *11*(4), 15–29. https://doi.org/10.1310/ACUK-KFYP-ABEQ-A30C

Dawson, D., Chan, R., & Kaiserman, E. (1994). Development of the power-mobility indoor driving assessment for residents of long-term care facilities: A preliminary report. *Canadian Journal of Occupational Therapy*, *61*(5), 269–276. https://doi.org/10.1177/000841749406100507

Deitz, J., Jaffe, K.M., Wolf, L.S., Massagali, T.L., & Anson, D. (1991). Pediatric power wheelchairs: Evaluation of function in the home and school environments. Assistive Technology, 3, 24–31. https://doi.org/10.1080/10400435.1991.10132177

Desai, S., Mantha, S.S., & Phalle, V.M. (2017). Advances in smart wheelchair technology. 2017 International Conference on Nascent Technologies in Engineering (ICNTE), 1–7. https://doi.org/10.1109/ICNTE.2017.7947914

Devigne, L., Babel, M., Nouviale, F., Narayanan, V.K., Pasteau, F., & Gallien, P. (2017). Design of an immersive simulator for assisted power wheelchair driving. 2017 International Conference on Rehabilitation Robotics (ICORR), 995–1000. https://doi.org/10.1109/ICORR.2017.8009379

Dicianno, B.E., Mahajan, H., Guirand, A.S., & Cooper, R.A. (2012). Virtual Electric Power Wheelchair driving performance of individuals with spastic cerebral palsy. *American Journal of Physical Medicine & Rehabilitation*, *91*(10), 823–830. https://doi.org/10.1097/PHM.0b013e31825a1497

Ding, D., Cooper, R.A., Guo, S., & Corfman, T.A. (2003). Robust velocity control simulation of a power wheelchair, Proceedings of the RESNA 26th International Annual Conference, Atlanta, GA, USA

Fehr, L., Langbein, W.E., & Skaar, S.B. (2000). Adequacy of power wheelchair control interfaces for persons with severe disabilities: a clinical survey. Journal of rehabilitation research and development, 37(3), 353–360

Fu, F., & Hao, Q. (Eds.). (2012). *Intelligent Sensor Networks: The integration of sensor networks, Signal Processing and machine lear*. CRC Press

Gafford, J. (2022). Multi-Layer Perceptron (MLP) Class (https://www.mathworks.com/matlabcentral/fileexchange/74695-multi-layer-perceptron-mlp-class), MATLAB Central File

Exchange. Retrieved: November 22, 2022

Gefen, N., Rigbi, A., Archambault, P.S., & Weiss, P.L. (2019). Comparing children's driving abilities in physical and virtual environments. Disability and Rehabilitation: Assistive Technology. https://doi.org/10.1080/17483107.2019.1693644

Gonçalves, F. & Trenoras, L.A., Monacelli, E., & Schmid, A. (2014). Motion adaptation on a wheelchair driving simulator. https://doi.org/10.1109/VAAT.2014.6799463

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press, Cambridge, MA, USA

Guimarães, M.P., Gnecco, B.B., & Damazio, R. (2007). Ferramentas para desenvolvimento de aplicações de Realidade Virtual e aumentada. *Realidade Virtual e Aumentada - conceitos, projeto e aplicações*

Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of Artificial Intelligence. *California Management Review*, *61*(4), 5–14. https://doi.org/10.1177/0008125619864925

Harrison, C.S., Grant, P.M., & Conway, B.A. (2010). Enhancement of a virtual reality wheelchair simulator to include qualitative and quantitative performance metrics. *Assistive Technology*, *22*(1), 20–31. https://doi.org/10.1080/10400430903520223

Hasdai, A., Jessel, A.S., & Weiss, P.L. (1998). Use of computer simulator for training children with disabilities in the operation of a powered wheelchair. American Journal of Occupational Therapy, 52, 215–220. https://doi.org/10.5014/ajot.52.3.215

Headleand, C.J., Day, T., Pop, S.R., Ritsos, P.D., & John, N.W. (2016). A Cost-Effective Virtual Environment for Simulating and Training Powered Wheelchairs Manoeuvres. Studies in Health Technology and Informatics, 220, 134–41

Hernandez-Ossa, K.A., Longo, B., Montenegro-Couto, E.H., Romero-Laiseca, M.A., Frizera-Neto, A., & Bastos-Filho, T. (2017). Development and pilot test of a virtual reality system for electric powered wheelchair simulation. 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC). 2355–2360. https://doi.org/10.1109/SMC.2017.8122974

Hernandez-Ossa K.A, Montenegro-Couto, E.H., Longo, B., Bissoli, A., Sime, M.M., Lessa, H.M., Enriquez, I.R., Frizera-Neto, A., & Bastos-Filho, T. (2020). Simulation System of Electric-Powered

Wheelchairs for Training Purposes. Sensors (Basel). 2020 Jun 24;20(12):3565. https://doi.org/10.3390/s20123565

Jhaveri, R.H., Revathi, A., Ramana, K., Raut, R., & Dhanaraj, R.K. (2022). A review on machine learning strategies for real-world engineering applications", Mobile Information Systems, vol. 2022, Article ID 1833507. https://doi.org/10.1155/2022/1833507

John, N.W., Pop, S.R., Day, T.W., Ritsos, P.D., & Headleand, C.J. (2018). The implementation and validation of a virtual environment for training powered wheelchair manoeuvres. *IEEE Transactions on Visualization and Computer Graphics*, *24*(5), 1867–1878. https://doi.org/10.1109/TVCG.2017.2700273

Kadurumba, C., Nwaiwu, U. & Nwasuka, N.C. (2020). *Neural network applications*

Kamaraj, D.C., Dicianno, B.E., Schmid, M., Boyanoski, T., & Cooper, R.A. (2014). Quantifying power wheelchair driving ability. In Conference Proceedings, RESNA (p. 1). Retrieved from https://www.researchgate.net/profile/Deepan_C_Kamaraj/publication/272683427_Quantifying_Power_Wheelchair_Driving_Ability/links/54ec00ac0cf2082851bf310b.pdf

Kamaraj, D.C., Dicianno, B.E., Mahajan, H.P., Buhari, A.M., & Cooper, R.A. (2016). Interrater reliability of the Power Mobility Road Test in the virtual reality–based simulator-2. *Archives of Physical Medicine and Rehabilitation*, *97*(7), 1078–1084. https://doi.org/10.1016/j.apmr.2016.02.005

Kamaraj, D.C. (2020). Quantifying Electric Powered Wheelchair Driving Ability. Doctoral Dissertation, University of Pittsburgh (Unpublished)

Kingma, D.P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization. CoRR, abs/1412.6980.

Kirby, R.L., Swuste, J., Dupuis, D.J., MacLeod, D.A., & Monroe, R. (2002). The Wheelchair Skills Test: a pilot study of a new outcome measure. Archives of physical medicine and rehabilitation, 83(1), 10–18. https://doi.org/10.1053/apmr.2002.26823

Kirby, R.L., Miller, W.C., Routhier, F., Demers, L., Mihailidis, A., Polgar, J.M., Rushton, P.W., Titus, L., Smith, C., McAllister, M., Theriault, C., Thompson, K., & Sawatzky, B. (2015). Effectiveness of a wheelchair skills training program for powered wheelchair users: A randomized controlled trial. *Archives of Physical Medicine and Rehabilitation*, *96*(11). https://doi.org/10.1016/j.apmr.2015.07.009

Kirner, C., & Siscouto, R. (Eds.) (2007). Realidade Virtual e aumentada: conceitos, projeto e aplicações. *Sociedade Brasileira de Computação*

Lange, M.L., & Grieb, E. (2015). *Optimizing power wheelchair use through mobility training*. Rehab Management. Retrieved from http://www.rehabpub.com/2015/10/optimizing-power-wheelchair-use-mobility-training/

Leblong, E., Fraudet, B., Devigne, L., Babel, M., Pasteau, F., Nicolas, B., & Gallien P. (2021). SWADAPT1: assessment of an electric wheelchair-driving robotic module in standardized circuits: a prospective, controlled repeated measure design pilot study. J Neuroeng Rehabil. 16;18(1):140. https://doi.org/10.1186/s12984-021-00923-2

Letts L., Dawson, D., & Kaiserman-Goldenstein, E. (1998). Development of the power-mobility community driving assessment. *Canadian Journal of Rehabilitation,11(3),* 123–129

Mahajan, H.P. (2012). Development and validation of simulators for power wheelchair driving evaluations (PhD thesis). University of Pittsburgh, Pittsburgh, PA (Unpublished)

Mann, H., & Whitney, D. (1947). On a test of whether one of two random variables is stochastically larger than the other, Ann. Math. Statist. 18(1), 50–60. https://doi.org/10.1214/aoms/1177730491

Martins, F.R. (2017). Simulador para treinamento de cadeirantes em ambiente virtual acionado por comandos musculares e/ou visuais. (Unpublished) https://doi.org/10.14393/ufu.di.2017.504

Martins, F.R., Naves, E.L.M., Morère, Y., & Sá, A.A.R. (2021). Preliminary assessment of a multimodal electric-powered wheelchair simulator for training of activities of daily living. J Multimodal User Interfaces 21:1–29. https://doi.org/10.1007/s12193-021-00385-9

Massengale, S., Folden, D., McConnell, P., Stratton, L., & Whitehead, V. (2005). Effect of visual perception, visual function, cognition, and personality on power wheelchair use in adults. *Assistive Technology*, *17*(2), 108–121. https://doi.org/10.1080/10400435.2005.10132101

Mikołajewska, E., Mikołajewski, D., & Rozwój. (2013). Wheelchair development from the perspective of physical therapists and biomedical engineers. *Advances in Clinical and Experimental Medicine*

Montenegro-Couto, E.H., Hernandez-Ossa, K., Bissoli, A., Sime, M., & Bastos, T. (2018). Towards an assistive interface to command robotic wheelchairs and interact with environment through eye gaze. https://doi.org/10.29327/cobecseb.78867

Morère, Y., Abdelkader, M.A.H., Meliani, S.M., & Bourhis, G. (2011). Powered wheelchair driving analysis on a simulator. pages 679–685. AAATE2011

Morère, Y., Bourhis, G., Cosnuau, K., Guilmois, G., Blangy, E., & Rumilly, E. (2015). View, a wheelchair simulator for driving analysis. *2015 International Conference on Virtual Rehabilitation (ICVR)*. https://doi.org/10.1109/ICVR.2015.7358574

Morère, Y., Bourhis, G., Cosnuau, K., Guilmois, G., Rumilly, E., & Blangy, E. (2018). ViEW: A wheelchair simulator for driving analysis. Assistive Technology. 3;32(3):125–135. https://doi.org/10.1080/10400435.2018.1503204

Niniss, H., & Nadif, A. (2000). Simulation of the behaviour of a powered wheelchair using virtual reality, Proceedings of the 3rd International Conference on Disabilities, Virtual Reality and Associated Technologies, 9–14, Alghero, Italy

Niniss, H., & Inoue, T. (2006). Assessment of driving skills using Virtual Reality: Comparative Survey on experts and unskilled users of electric wheelchairs. *Technology and Disability*, *18*(4), 217–226. https://doi.org/10.3233/TAD-2006-18409

Nunes, F., Costa, R.M., Machado, L., & Moraes, R.M. (2011). Realidade virtual para saúde no Brasil: Conceitos, desafios e oportunidades. *Revista Brasileira De Engenharia Biomédica*, *27*(4), 243–258. https://doi.org/10.4322/rbeb.2011.020

Nunnerley, J., Gupta, S., Snell, D., & King, M. (2016). Training wheelchair navigation in immersive virtual environments for patients with spinal cord injury – end-user input to design an effective system. *Disability and Rehabilitation: Assistive Technology*, *12*(4), 417–423. https://doi.org/10.1080/17483107.2016.1176259

Pithon, T., Weiss, T., Richir, S., & Klinger, E. (2009). Wheelchair simulators: A review. *Technology and Disability*, *21*(1-2), 1–10. https://doi.org/10.3233/TAD-2009-0268

Rodrigo, S.E., & Herrera, C.V. (2008). Wheelchairs: history, characteristics, and technical specifications. Smart Wheelchairs and Brain-Computer Interfaces, 257–290. https://doi.org/10.1016/B978-0-12-812892-3.00011-X

Sá, A.A.R., Morère, Y., & Naves, E.L.M. (2022). Skills assessment metrics of electric powered wheelchair driving in a virtual environment: a survey. Med Biol Eng Comput;60(2):323–335. https://doi.org/10.1007/s11517-022-02500-8

Sánchez, J., Cobb, S., Sharkey, P., & Merrick, J. (2011). Virtual reality and assistive technologies for people with disabilities. *International Journal on Disability and Human Development*, *10*(4). https://doi.org/10.1515/IJDHD.2011.065

Singh, S.P., Wang, L., Gupta, S., Goli, H., Padmanabhan, P., & Gulyás, B. (2020). 3D deep learning on medical images: A review. Sensors (Basel, Switzerland), 20(18), 5097. https://doi.org/10.3390/s20185097

Silva, Y., Souza, V., Naves, E., & Bastos. T. (2018). Teleoperation training environment for new users of electric powered wheelchairs. 8th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcares (ICTH 2018), 343–350. https://doi.org/10.1016/j.procs.2018.10.191

Shalev-Shwartz, S., & Ben-David, S. (2009). Understanding machine learning. https://doi.org/10.1017/CBO9781107298019

Slater, M., Wilbur, S. (1997). A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. Presence: Teleoperators and Virtual Environments. 6 (6): 603–616. https://doi.org/10.1162/pres.1997.6.6.603

Spaeth, D. M., Mahajan, H., Karmarkar, A., Collins, D., Cooper, R. A., & Boninger, M. L. (2008). Development of a wheelchair virtual driving environment: trials with subjects with traumatic brain injury. Archives of physical medicine and rehabilitation, 89(5), 996–1003. https://doi.org/10.1016/j.apmr.2007.11.030

Tu, C.J., Liu, L., Wang, W., Du, H.P., Wang, Y.M., Xu, Y.B., & Li, P. (2017). Effectiveness and safety of wheelchair skills training program in improving the wheelchair skills capacity: a systematic review. Clinical Rehabilitation, 31(12), 1573–1582. https://doi.org/10.1177/0269215517712043

Vailland, G., Grzeskowiak, F., Devigne, L., Gaffary, Y., Fraudet, B., Leblong, É, Nouviale, F., Pasteau, F., Breton, R., Guégan, S., Gouranton, V., Arnaldi, B., & Babel, M. (2019). User-centered design of a multisensory power wheelchair simulator: towards training and rehabilitation applications," 2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR), 77–82. https://doi.org/10.1109/ICORR.2019.8779496

Valentini, C.A.M. (2019). Protocolo para condução de cadeira de rodas motorizada usando realidade virtual. 118 f. Dissertation. Federal University of Uberlândia. Uberlândia. (Unpublished) https://doi.org/10.14393/ufu.di

Webster, J.S., McFarland, P.T., Rapport, L.J., Morrill, B., Roades, L.A., & Abadee, P.S. (2001). Computer-assisted training for improving wheelchair mobility in unilateral neglect patients. Archives of Physical Medicine and Rehabilitation, 82(6), 769–775. https://doi.org/10.1053/apmr.2001.23201

Wilcoxon, F. (1945). Individual comparisons by ranking methods, Biometrics Bull. 1(1), 80–83. https://doi.org/10.2307/3001968

Winn, A.K., & Julius, A.A. (2013). Optimization of human generated trajectories for safety controller synthesis. *2013 American Control Conference*. https://doi.org/10.1109/ACC.2013.6580513

Woods, B., & Watson, N. (2003). A short history of powered wheelchairs. *Assistive Technology*, *15*(2), 164–180. https://doi.org/10.1080/10400435.2003.10131900

Woods, B., & Watson, N. (2004). The social and technological history of wheelchairs. *International Journal of Therapy and Rehabilitation*, *11*(9), 407–410. https://doi.org/10.12968/ijtr.2004.11.9.19587

Zador, A.M. (2019). A critique of pure learning and what artificial neural networks can learn from animal brains. Nat Commun 10, 3770. https://doi.org/10.1038/s41467-019-11786-6

Zatla, H., Morère, Y., Hadj-Abdelkader, A., Bourhis, G., Demet, K., Guilmois, G., Bigaut, N., K., & Cosnuau, K. (2018), Preview distance index for the analysis of powered wheelchair driving, IRBM, 39 (3), 194–205. https://doi.org/10.1016/j.irbm.2018.03.001

Zhang, X., Hui, L., Wei, L., Song, F., & Hu, F. (2021). A bibliometric analysis of human-machine interaction methodology for electric-powered wheelchairs driving from 1998 to 2020. Int J Environ Res Public Health. 18(14):7567. https://doi.org/10.3390/ijerph18147567