

Proposal of an Augmented Reality Telerehabilitation System for Powered Wheelchair User's Training

Daniel Caetano  , Caroline Valentini, Fernando Mattioli  , Paulo Camargos  , Thiago Sá  , Alexandre Cardoso  , Edgard Lamounier  and Eduardo Naves 

Abstract—Many people worldwide have been experimenting a decrease in their mobility as a result of aging, accidents and degenerative diseases. In many cases, a Powered Wheelchair (PW) is an alternative help. Currently in Brazil, patients can receive a PW from the Unified Health System, following prescription criteria. However, they do not have an appropriate previous training for driving the PW. Consequently, users might suffer accidents since a customized training protocol is not available. Nevertheless, due to financial and/or health limitations, many users are unable to attend a rehabilitation center. To overcome these limitations, we developed an Augmented Reality Telerehabilitation System Architecture based on the Power Mobility Road Test (PMRT) for supporting PW user's training. In this system, the therapists can remotely customize and evaluate training tasks and the user can perform the training in safer conditions. The video stream and data transfer between each environment were made possible through UDP (User Datagram Protocol). To evaluate and present the system architecture potential, a preliminary test was conducted with 3 spinal cord injury participants. They performed 3 basic training protocols defined by a therapist. The following metrics were adopted for evaluation: number of control commands; elapsed time; number of collisions; biosignals. Also, a questionnaire was used to evaluate system features. Preliminary results demonstrated the specific needs of individuals using a PW thanks to adopted metrics (qualitative and emotional). Results have shown the need for a training system with customizable protocols to fulfill these needs. User's evaluation demonstrates that the combination of AR techniques with PMRT adaptations, increases user's well-being after training sessions. Furthermore, a training experience helps users to overcome their displacement problems, as well as for appointing challenges before large scale use. This system architecture allows further studies on telerehabilitation of PW users training.

Index Terms—Mobility, Powered Wheelchair Training, Augmented Reality, Telerehabilitation, Power Mobility Road Test, Biosignals

I. INTRODUCTION

IN Brazil, more than 45 million people have some motor limitation among which 2.33 percent (1 million people) have a severe motor disability, according to the 2010 census

Daniel Caetano, Fernando Mattioli, Paulo Camargos, Thiago Sá, Alexandre Cardoso, Edgard Lamounier and Eduardo Naves is with the Department of Electrical and Computer Engineering, Faculty of Electrical Engineering from Federal University of Uberlândia, Brazil e-mail: {sdc.daniel, mattioli.fernando, paulocamargos12, thiagosspaiva,}@gmail.com and {alexandre, lamounier, eduardonaves}@ufu.br). Caroline Valentini is with Disabled Child Care Association of Uberlândia email: ckti.carolinevalentini@gmail.com

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 (Project CAPES PGPTA 3461/2014).

Digital Object Identifier (DOI):10.14209/jcis.2020.5

[1]. Among them, mostly are elderly people whose autonomy is seriously affected by a decline in their motor function and cognitive performance. Also, it includes individuals who have suffered a stroke or injury [2]–[4]. According to the census of England and Wales carried out in 2011, 1.9 percent of the population use a wheelchair, an estimated 1.2 million people [5]. Other countries will have proportionally similar numbers within their population.

In this context, a Powered Wheelchair (PW) is one of the most important Assistive Technology Devices (ATDs) to aid in recovering their autonomy and mobility in daily activities [4]. However, not everybody has conditions to acquire such device [2], [3], [6].

The Unified Health System in Brazil is responsible for providing a PW for users, attending criteria established by Ordinance n. 1272 through rehabilitation centers Brazil [7], [8]. Although users get a simple training on the PW, it is not enough to prepare them, leading to misuse and eventual accidents [7], [8]. Nevertheless, many users face different problems to overcome their displacement of rehabilitation center like financial resources, time available and others [9].

Authors have been exploring Virtual Reality (VR) in many protocols and several situations safely [5], [6], [10], [11]. A Virtual Reality Environment (VRE) for PW driving training, which is focused only on indoor challenges are shown by John et al.[5], Kamaraj et al. [10] and Mahajan et al. [11]. A VRE with indoor and outdoor scenarios, but not yet focused in training is shown by Vailland et al.[6]. An important feature for these VRE's, that increase user training adherence, is the immersion level. Usually, it is related to the quality of 3D models, the number of senses transferred to the VRE, e.g., a head-mounted display (HMD) or desktop monitor, used to preview the VRE, although it can lead the user to have cybersickness [5], [6], [10]. Notwithstanding, these VRE's have some limitations such as were tested only with healthy users, poor quality of VR immersion, some real-world situations are not easy to reproduce, and the existence of hidden objects in some cases. When eligible users are involved, there are many causes of limitations and not every user is able to wear an HMD. Thus, an adaptive and flexible environment, that allows the therapist to set up individualized protocols and to assess users distinctly, caring about their impairments in real-time is needed [12].

On the other hand, Augmented Reality (AR) can provide the user with better control, where user interaction can be achieved in a more realistic and intuitive way and safely [13]–[17].

In doing so, the system augments the real world with digital information (e.g., pictures, videos, instructions, clinical data) enhancing the user's experience [14]. AR systems can allow a health professional to monitor users' performance visually. For example, haptic machines and wearable sensors can record quantitative information while users exercise [14].

Driving performance was assessed in many different ways in presented simulators. A non-standard statistic method to asses comparative metrics among the users from each group was used by John et al. [5]. The PMRT was chosen by Kamaraj et al.[10], Mahajan et al. [11] and Massengale et al. [18] as a standard and reliable methodology (training protocol and assessment) applied for each user distinctly. The PMRT consists of 2 groups of tasks: structured and unstructured. The 12 structured tasks are predictable. The 4 unstructured/dynamic tasks are unpredictable and require the user to make decisions about interacting with the environment, such as avoiding a person walking down the hallway or avoiding a therapy ball in the way[18]. All tasks evaluated on PMRT are based on visual perception. Moreover, user's emotional state associated with a training session was not evaluated. Biosignals data processing allows looking for clues related to poor driving performance connected to that emotional state [19]–[22].

The user displacement challenge to address to the rehabilitation center has been overcome due to the advances in computer network technology, a new technique has brought the possibility of delivering rehabilitation services to users far from the rehabilitation centers. This technique — known as telerehabilitation[23] has among others, the goal of extending and improving user care [23], [24].

In light of the aforementioned, in this paper, the authors presents a telerehabilitation architecture developed to be used as a complementary tool for PW users training, based on the PMRT [10], [18]. This architecture allows a remote support by the therapist to define customized protocols using virtual (static and animated) objects to be rendered in different markers. During the entire training session, user's biosignals (Electrodermal Activity - EDA and Blood Volume Pulse — BVP) were collected to assess the impacts of the different protocols on each user and provide clues related to performance improvements [19]–[22].

This paper is organized as follows. We introduce the telerehabilitation system architecture designed including the environments description in Section II-A, while the system modeling detail to support the architecture is shown in Section II-B. Biosignal data acquisition and processing is presented in Section II-C. Preliminary test scope is presented in Section II-D to evaluate user experience and their training with the system. The results are presented and discussed in Section III and IV. Finally, in Section V we draw our conclusions and highlight some potential future studies.

II. MATERIAL AND METHODS

A. Architecture Designed

Based on three main modules of computational system used by Burdea et al. [25], he authors present the solution outline on Figure 1, composed by [26]:

- A training site, where the exercises protocols will be executed;
- A patient site, where the user remotely controls the wheelchair;
- A therapist site in which a health professional can customize, follow and evaluate trials executed by the user and access performance data;

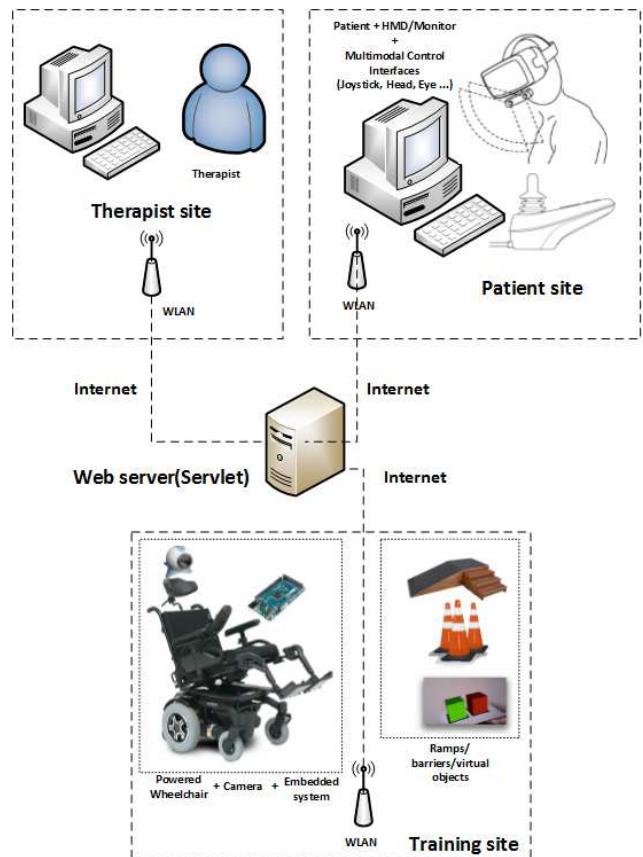


Fig. 1: AR Telerehabilitation System Architecture

1) *Training site*: The controlled environment where each structural part was built based on 60 surveys filled out by PW users with different disabilities. The area presented in Figure 2 is 14.76x7.16 m and is currently located in a classroom at Federal University of Uberlândia and can be previewed in 360° image¹.

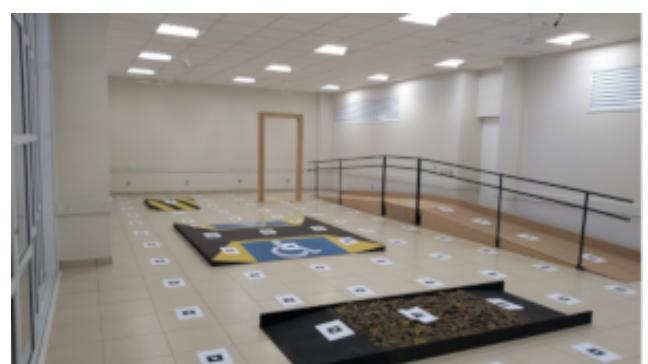


Fig. 2: Remote training site

¹<https://www.swheelchairth.com.br:8443/servletserver/room.jsp>

The room is composed by physical objects such as curb and a high access ramp, uneven surface area, spine, portal and a PW (Figure 3a) without diagonal movements.

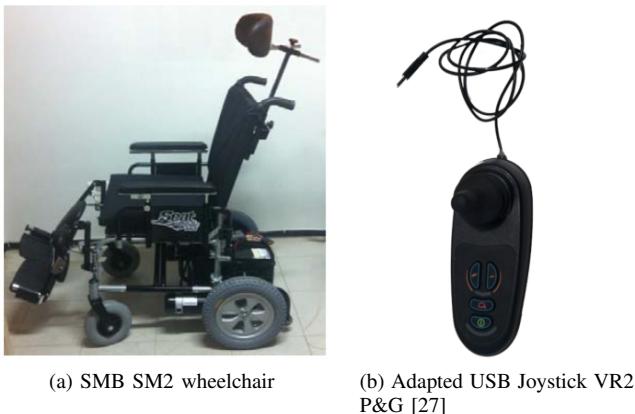


Fig. 3: Training and Patient environment components

2) *Patient site:* From a remote place, the user controls the PW using different interfaces such as a keyboard, joystick (Figure 3b) and others. From traditional monitors the patient has an augmented preview in a semi-immersive level.

3) *Therapist site:* As in the patient environment from a remote place, the health professional is able to aid the user on performing the tasks described in Section II-A and also interacts with the PW using the arrow keys to move or space key to stop.

B. Implementation details

The web server application² was implemented using Java™ Servlets technology that provides a consistent mechanism and makes possible the development of an architecture with video streams (Multi-WebRTC) and different channels to receive/redirect data. Thus, an independent system platform was designed to keep different sites connected over Internet. The development is based on the MVC (Model - View - Control) which is very useful for building interactive software [28]. The MVC is divided into Model (Java class), Controller (Java Servlet) and View (JSP pages). This makes it easy to incorporate different libraries (JSArtoolkit, Bootstrap and WebGL) and features with a reliable web application. Servlets are used in MVC model to control events that come from JSP page through POST, GET requests and also external data coming from sockets (UDP). In this system, there are two basic actors' roles: therapist and patient.

1) *Training environment:* The training environment has a PW connected to a microcontroller responsible for receiving user's commands, activate the PW and collect its speed. The data transfer between the web server and this environment was made possible through UDP socket connection using a 4G router modem. Also, a smartphone was coupled to the PW, responsible to provide a pure video streaming data to the therapist and patient environment. An emergency stop was implemented in order to avoid possible crashes in case

of data interruptions in data transmission greater than 3 seconds [9]. The patient permission to interact and visualize the environment is given by the therapist after starting the video streaming.

2) *Therapist role:* One of the most important actions in the Therapist role is “**create a new protocol**”. Other available actions are register a new user (therapist or patient), setup, begin, tracking, evaluate, take notes of a training session and follow the training session. In Figure 4, the therapist is able to insert or clear activities, make a protocol description and define an empty fiducial AR markers map.

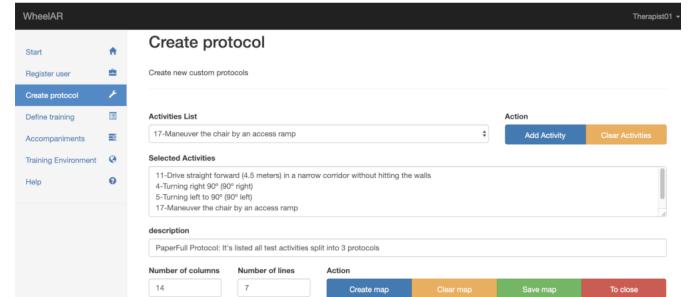


Fig. 4: Defining the activities for each protocol

All red buttons shown in Figure 5 represent physical objects inside the room that can be used by the therapist on training protocol.

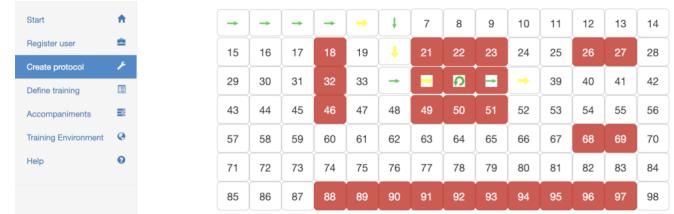


Fig. 5: Customized protocol through AR techniques

Virtual objects can be used as a guidance (green arrows) or avoidance. From the list presented in Figure 6 it is possible to add a static or animated object to compose each activity of the training protocol as presented in Figure 5.



Fig. 6: Virtual objects (animated/statics) list

Yellow arrows are used to signalize the end of an activity proposed by the therapist. This adaptation was implemented based on PMRT in order to allow the therapist to customize the protocols activities (structural or non-structural) providing the user with their needs.

²<https://www.swheelchairth.com.br:8443/servletserver/>

The training workflow consists of receiving the patient request and defining the protocol, initiating stream, tracking session, evaluate training and saving notes (therapist role). After the training request, the patient is redirected to a “**Waiting page**” as presented in Figure 7. In this page the patient is instructed to read all information of existing virtual objects and which action to perform.

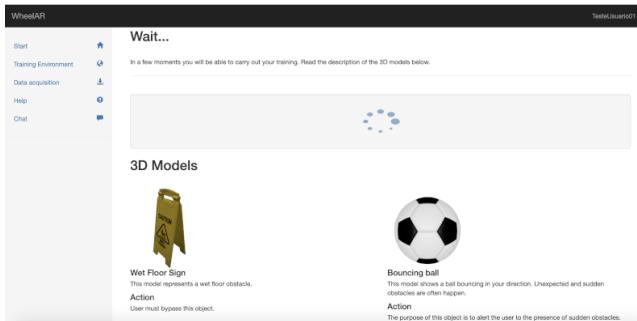


Fig. 7: Waiting page: Virtual objects information

Figure 8 present the training sessions requested by the patients and possible actions and also realize the training environment status.

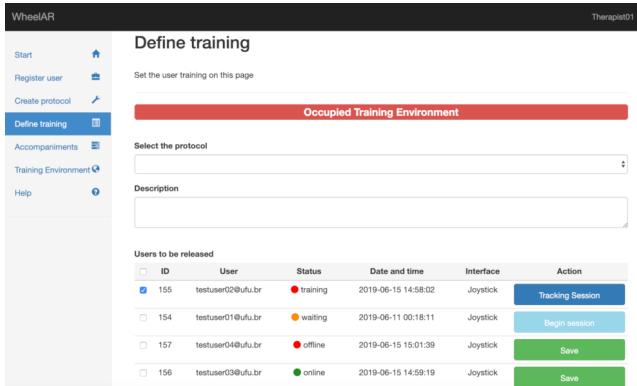


Fig. 8: Training protocols request status and actions

After the therapist has started a new training session his is able to track it and mark each activity finished by the patient from the viewport presented in Figure 9.

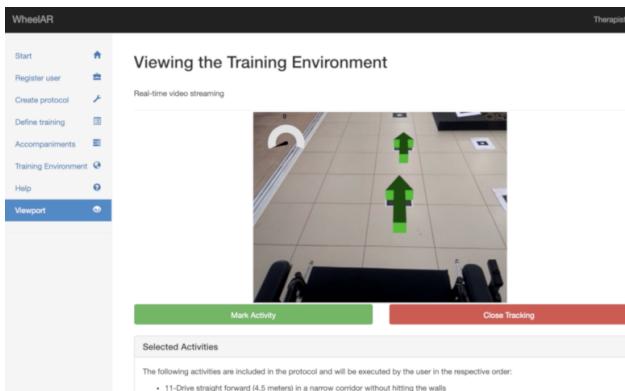


Fig. 9: Tracking session and Therapist ViewPort page

Once the training protocol is finished, the therapist is able to evaluate the training session by the adaptation realized in a

PMRT Assessment form as presented by Figure 10. After the workflow fulfillment he can also use the bar chart to follow the patient evolution after each protocol.

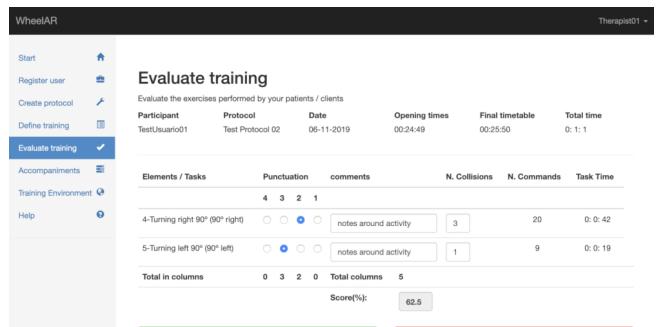


Fig. 10: Adapted PMRT Evaluation Form

3) *Patient role:* A cross-platform Java™ application shown in Figure 11a and Figure 11b must be downloaded from the patients home page presented in Figure 12.



(a) Starting dataflow (b) Close dataflow

Fig. 11: Data acquisition GUI

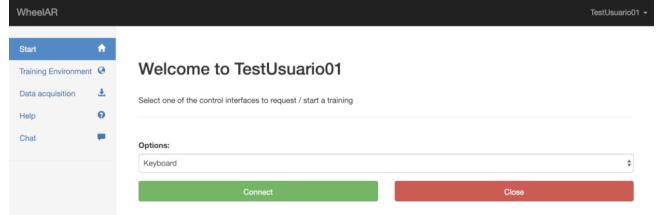


Fig. 12: Patient's page

This application is responsible for establishing a connection with the system server and relay commands produced by the patient in a joystick [29]. Requesting a new training session, the patient must select the control interface. After the therapist released the training the patient is able to interact and preview the augmented training environment shown in Figure 13.

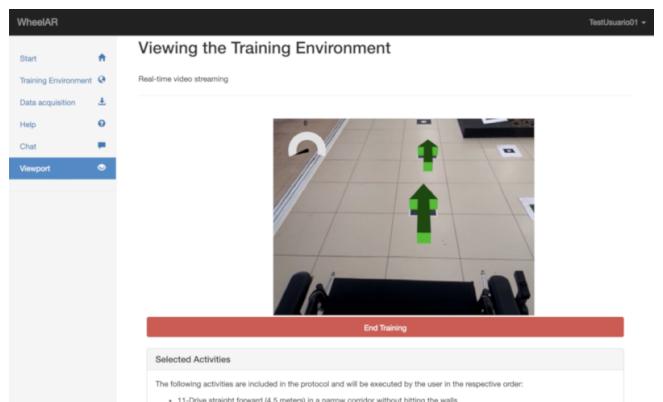


Fig. 13: Real-time streaming preview page

C. Biosignal data acquisition and processing

Biosignal data acquisition was performed using the E4 wristband (Figure 14), manufactured by Empatica®. The instrument is an easy to wear wristband that can measure various biosignals, among them EDA and BVP.



Fig. 14: The E4 wristband for biosignal data acquisition

EDA data processing was conducted using [30], as shown in Figure 15, to decompose the signal in two components: phasic and tonic. For evaluation of the event related skin conductance responses (each command given by the participant is considered an event in this preliminary study), the phasic component was used, considering an event response window starting 1s after each event and finishing 4s after this same event. The Continuous Decomposition Analysis (CDA) method was used, given its robustness on decomposing EDA biosignals in continuous tonic and phasic data [31].

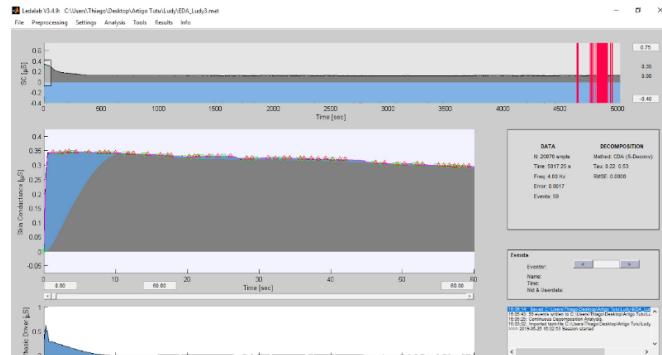


Fig. 15: Ledalab software used for EDA decomposition into phasic and tonic components

Amongst output variables extracted from Ledalab to evaluate individuals responses, the Integrated Skin Conductance Response (ISCR) was used. This metric consists of the time integral of the phasic driver extracted by Ledalab within the response window (1s to 4s after event). This variable was chosen since it considers both magnitude and duration of responses (it's a time integral), while other available variables (count of SCRs, average SCR, sum of amplitudes of SCRs, SCR Phasic Max Response) take into account more unidimensional aspects of the responses. All data was z scored to reduce subject variability.

Data processing of BVP was performed using Kubios [32], a software used to analyze Heart Rate Variability. It utilizes Inter Beat Interval (IBI), extracted from BVP measured by the E4 wristband [33] for each participant on each protocol. It returns several output variables, among them the stress index

(SI) calculated as the square root of Baevsky's stress index [34].

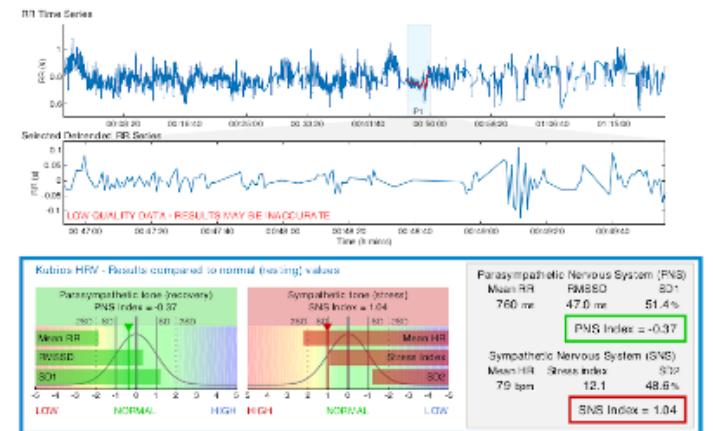


Fig. 16: A section of a report generated by the Kubios software for HRV

D. Preliminary test

1) *Participants:* General and Clinical information about the participants in this research is detailed in Table II and I.

TABLE I: Participants general information

No.	Description	Detail
1	Participants	Consist of 3 adults aged 22 to 57 years;
2	Gender	1 woman and 2 men with spinal cord injury [35], [36]; Collection was made in a rehabilitation center 10 kilometers far from the training site;
3	Location	
4	Consent form	All participants were asked to fill a consent form;
5	Ethical Committee	The approval to conduct this research was obtained from the Federal University (377566140.0000.5152).

TABLE II: Participants clinical information

Participant	Detail
01	He has 2 years of injury, has 18 months of experience in driving the Ottobock PW, drove a car previously, the user performance is satisfactory, the performance of ADL (feeding) / AIVD (change of posture, transport), have transportation (not adapted), house adapted to move around and needs joystick adaptation;
02	He has 5 year of injury, has 48 months of experience in driving the Ortobrás PW, drove a car previously, the user performance is satisfactory, the performance of ADL (Dress, undress, put on.) / AIVD (SI), have adapted transportation, house fully adapted and requested a joystick position adaptation;
03	She has 3 years of injury, has 6 months of experience in driving the Freedom PW, did not drive cars previously, the user performance is satisfactory but uses adaptation in the trachea, the Activities of Daily Living (ADL) / Instrumental Activities of Daily Life (AIVD) performance is semi independent (SI) and SI, it does not have transport, have significant spasms that affect postural control and PW control performance. She is not well positioned in the PW. Required posture revision and joystick adjustment;

2) *Procedure and process:* The initial process steps followed for each participant every first time, instructed and supported by a therapist, is described in Table III.

The survey questionnaire is shown in Tables IV and V.

TABLE III: Procedure and Processes

Step	Procedure/Process
1	Ensure that the environment temperature is conditioned at 25°C for good quality of biosignals [20];
2	Participants will be registered into the system by the therapist as a patient. Participants had the project explained again for purposes of clarification and for a chance to withdraw the study;
3	Participants will be instructed about how to log into the system;
4	Participants will be instructed about system functionalities and how to make the download of the application, responsible for transmitting the joystick commands to control the PW remotely;
5	The remote environment will be presented for the participants in 360° and explained how he/she must proceed to interact with;
6	Explain to the patient how to request a training session;
7	Participants will be instructed about how to request a training session;
8	Let the patient rest by 7 minutes before starting each training session to ensure a good emotional state and relax [20];
9	Wearing the "E4 wristband" wearable device as shown in Figure 14;
10	Request the training sessions;
11	While the therapist is selecting one of the preliminaries training protocols(Figure 5) splitted in: drive straight forward (4.5 meters) in a narrow corridor without hitting the walls; turning right and left 90° and maneuver the chair by an access ramp, the patient read all information about how to proceed in from of each virtual object;
12	At the end of each protocol, participants are invited to fill out a survey questionnaire with advantages, disadvantages and suggestions or observations about the protocol performed while the therapist is evaluating the trial performed;
13	Participants have to rest for 5 minutes to ensure comfort and absence of side effects before request a new training session [20];
14	After the last protocol, participants are invited to fill out other questions related to his/her own individual profile and system requisites;
15	The study ended.

TABLE IV: Survey Questionnaire

N.o	Questions Asked
6	How was it to lean to use the system? 5 4 3 2 1 Very Easy Easy Relatively Easy Difficult Very Difficult
7	How do you evaluate the graphical interface of the system? 5 4 3 2 1 Great Very good Good Median Bad
8	How was it to use the system? 5 4 3 2 1 Very easy Easy Relatively easy Difficult Very difficult
9	Did the system meet the navigation needs? 5 4 3 2 1 Excellent Great Good Median Bad
10	How do you consider the quality of the image presented? 5 4 3 2 1 Excellent Great Good Median Bad
11	Do you consider that the AR (virtual objects) help to carrying out the training? 5 4 3 2 1 Very Moderate Medium Little None
12	How was the processing time (response time)? 5 4 3 2 1 Very fast Fast Moderate Slow Very slow
13	How much do you consider this tool assists in the development of driving skills? 5 4 3 2 1 Intensely Very Moderately Little Nothing
14	How do you evaluate your well-being after training? 5 4 3 2 1 Relaxed Little tired Tired Very tired Exhausted
15	Are you satisfied with the system features? 5 4 3 2 1 Very Satisfied Satisfied Indifferent Dissatisfied Very dissatisfied

TABLE V: Survey Questionnaire

N.o	Questions Asked
16	Does the system do the right thing? 5 4 3 2 1 Extremely well Very well Well Relatively well Nothing

3) *Data collection and analysis:* The data information extracted during the trials is shown in Table VI and the analysis performed in this data is described in Table VII.

TABLE VI: Data collection information

N.o	Description
1	Parameters like the number of input controls, elapsed time, collisions number were measured in order to evaluate the participants trials performance;
2	A survey questionnaire was also used to provide qualitative information about the system from the participants point of view as well the therapist's comments about each activity performed;
3	Participants biosignals were also collected to analyze protocols impacts during the trials

TABLE VII: Data analysis information

N.o	Description
1	The PMRT Methodology is used for trials performance evaluation;
2	Bar charts shown the participants questions evaluation and Word clouds [37] represent the most relevant reasons for the patient's evaluations based on participants and therapist's comments/notes;
3	Protocols data analysis were divided in: 1) EDA Analysis, using ISCR as variable for statistical analysis between protocols. All statistical evaluations were made using the R-software™ [38]. The Shapiro-Wilk test was used, and the data distribution was found to be non-normal [39]; 2) Stress Index Analysis, using the IBI time series(Inter-Beat Interval - Time between individual heart beats) was used to obtain the Stress Index for each protocol individually using the software Kubios [40];

III. RESULTS

A. Trials performance evaluation information

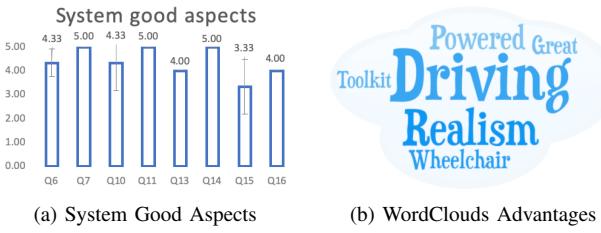
Table VIII presents the performance evaluation realized by therapist.

TABLE VIII: PMRT Evaluation Summary

Participant	Prt	Command	Collision	Time(s)	Score(%)
01	1	38	2	199	50
	2.1	37	1	328	50
	2.2	7	0	57	50
	3	54	0	354	75
02	1	27	0	282	75
	2.1	17	0	127	75
	2.2	21	0	67	75
	3	66	0	178	75
03	1	14	0	227	75
	2.1	9	0	98	50
	2.2	13	0	51	50
	3	34	0	236	75

B. System requirements information

Figure 17a and Figure 18a represent the participants' survey questionnaire evaluation. Figure 17b and Figure 18b represents visual qualitative information about advantages, disadvantages and comments.

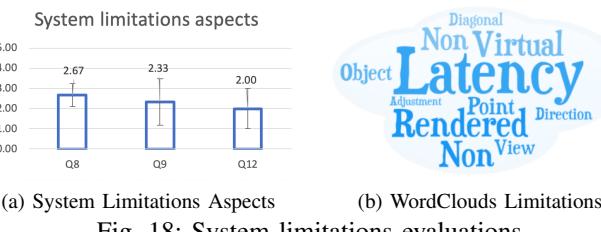


(a) System Good Aspects



(b) WordClouds Advantages

Fig. 17: Good evaluations



(a) System Limitations Aspects



(b) WordClouds Limitations

Fig. 18: System limitations evaluations

C. Protocols evaluation

1) *EDA Analysis:* The ISCR (obtained after processing EDA) Descriptive Statistics is presented in Table IX. It is possible to observe that Skin Conductance Responses are skewed and can produce non-normal distributions [41], as observed in the difference between ISCR average and ISCR median.

TABLE IX: ISCR Descriptive Statistics

Group	Count	Mean	SD	Median
Protocol 01	83	-4.65e-18	0.988	-0.110
Protocol 02	111	-9.01e-12	0.991	-0.263
Protocol 03	123	-2.44e-11	0.992	-0.274

Columns meaning:

- Count: number of observations
- Mean: ISCR average on each protocol
- SD: Standard Deviation of ISCR on each protocol
- Median: ISCR Median of each protocol

Results of statistical analysis of ISCR as a variable in each protocol are: **Kruskal-Wallis chi-squared = 0.012995, df = 2, p-value = 0.9935**, showing no statistical significant differences between each protocol applied related to ISCR [42]. These results refer only to the sample utilized, to demonstrate system's capabilities. A higher number of volunteers in future studies will be necessary to draw broader and more general conclusions. Therefore, future studies should consider a higher number of participants.

2) *Stress Index Analysis:* In order to obtain the stress index for each protocol and each participant, we used the IBI time series prepared for processing using Kubios.

TABLE X: Stress zone boundaries

Stress Zones	SI
Very High	>30
High	22.4 - 30
Elevated	12.2 - 22.4
Normal	7.1 - 12.2
Low	<7.1

TABLE XI: Stress zone participants comparison

User	SI Protocol 1	SI Protocol 2	SI Protocol 3
Participant 01	12.0769	11.4139	09.9194
Participant 02	17.5959	14.8135	16.9240
Participant 03	06.0430	05.2265	08.0705

Table X present the Stress Zone and corresponding SI (Stress Index) band values [40].

Participants comparison results are presented in Table XI:

Finally, data presented on Table XI is translated to stress zones in Table XII:

TABLE XII: Stress zone participants translation

User	SI Protocol 1	SI Protocol 2	SI Protocol 3
Participant 01	Normal	Normal	Normal
Participant 02	Elevated	Elevated	Elevated
Participant 03	Low	Low	Normal

IV. DISCUSSION

To better understand the participant's acceptance of the developed tool/architecture and its requirements, the following considerations is presented.

Questions 6, 7 (Figure 17a) and 8 (Figure 18a) seek to evaluate usability. It is possible to note that the participants considered the GUI handy and that the system is easy to learn. According to participants, the two main faults in the system were the latency (response time - Q12[Figure 18a]) and failure to render virtual objects (Figure 18b) shown in Figure 19.



(a) Rendering failure-01



(b) Rendering failure-02



(c) Rendering failure-03

Fig. 19: Virtual objects rendering failure

In Figure 19a the guidance arrow is rendered at a wrong place, in Figure 19b and 19c the guidance and final activity arrows was not rendered properly, due to image quality, marker size or illumination. Other facts can be considered, with less relevance such as the PW does not possess diagonal movements Q9 (Figure 18a) and, adjustments from the participant (Figure 18b).

On the other hand, questions 10, 11, 13, 15 and 16 (Figure 17a) confirm that the application of telerehabilitation techniques fused with the AR techniques presents itself as a great tool to elevate the driving skills of the patients without leaving their home(Figure 17b). This is due to the realism, by controlling a real PW, through a real scenario, in addition to listening to the sounds of triggering and pausing the PW during the realization of the protocol, as shown in Figure 20 .



Fig. 20: Virtual objects static and animated

During the execution of each training protocol, the therapist is able to monitor the participants' performance in the proposed activities in real time. Upon completion of the training process, the therapist must complete the adapted PMRT assessment form. Information such as number of commands and up time are retrieved from the database. The therapist must still fill in the number of collisions and comment on the participant's performance before assigning a note to the activity. After the activity notes are applied, the final score of the protocol is calculated. The participant is considered approved in a protocol when he/she reaches a score higher than or equal to 95% [18], as the summary presented by Table VIII. Through the comparative bar charts shown in Figure 21, the therapist can follow the evolution in the development of the participant's abilities. Therefore, it can define the design of future protocols that will prepare the user to drive a PW in their daily activities confidently.

Biosignals can provide important information related to participant's performance connected to the emotional state that can prove useful in future training applications and on strategies for improving safety during PW use.

The participant's emotional state can constitute a barrier to good driving performance, and it is important to understand, whenever these issues arise, if the underlying cause is related to the participant's motor skills or associated to their emotional state, looking to correctly address these issues. Additionally, in training protocols, it is desired that the user is not overwhelmed by the training difficulty level, which could be observed through his biosignals with appropriate methodology and detection algorithms.

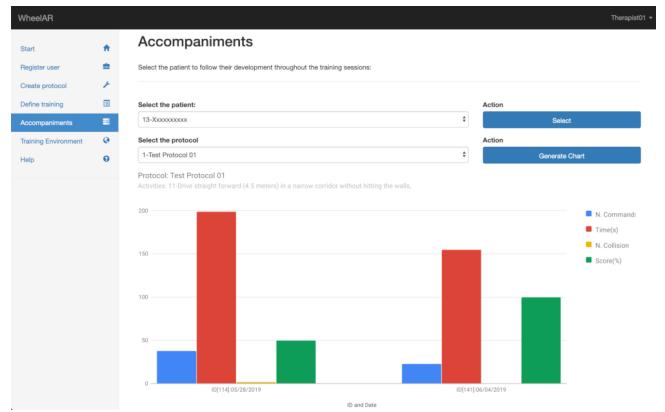


Fig. 21: Patient training protocol comparison

This means that, in this specific case, there was no relevant difference on participants performance considering the three different protocols. However, this does not represent a final conclusion given that the sample was small and protocol duration is considered short - longer protocols might present different results. The intention of this preliminary test is to showcase the system architecture capabilities.Having a different impact across individuals may emphasize the importance of having individualized or personalized protocols that could be based on therapist observations complemented by biosignal data.

V. CONCLUSION

Based on the participants' survey answers, we believe that the purpose of the system architecture was achieved. In this preliminary study, all tests were conducted in a rehabilitation center. However, in the future, after the issues found in this research have been fixed, only the first session will be performed in a rehabilitation center and futures sessions might be accomplished at home, in agreement with the patient conditions. Since the beginning of the system architecture development, we opted to use only open-source tools and libraries because we hope in the future to incorporate this system on the rehabilitation centers connected with Brazilian Unified Health System and share³ this project with the community as open-source for non-commercial license use.

From these results we conclude that the protocols did not show difference on the participants analyzed. Participant 3 was the only one changing stress zones. However, we believe it represents a normal variation towards a resting state on the normal stress zone, not caused by the protocols.

Different subject segmentation based on injury type, severity, age and experience using assistive technologies can be implemented to evaluate the learning curve connected to participant biosignals evolution during time and to identify roadblocks on each protocol. Patient's stress monitoring is also important not only to address learning but also to assure safety during normal use. Also, further studies could be done in this direction to classify patient's stress better and provide a risk

³<https://github.com/dantutu/servletserver>

assessment in order to develop improved strategies to prevent accidents.

Future work can be divided into two strands: increase users' experience and scientific contributions. In order to correct the rendering failure of the augmented objects, the fiducial markers will be removed and replaced by rendering based on the Cartesian position of the PW [43]. By implementing a proximity sensor, all virtual objects within range are rendered, as in the Poke-mon GO game [44]. Otherwise, the use of deep learning techniques can be investigated in order to reduce rendering failure [45]. Studies will be conducted, to improve the control signals and video stream latency. With the release of 5G Networks, many problems related to latency faced in this work due to the 4G Internet and Internet speed provide by concessionary, in the rehabilitation center can be addressed [46]. Finally, replace the model of PW used by one with less time of breaking of inertia and also with diagonal movements.

Although the results were satisfactory, there are limitations in this research-study, such as the small number of volunteers and the need to have more sessions and training during the experiment. However, due to latency during the trials in many different moments was not possible to complete the sessions. As mentioned before, the participants have many restrictions, as described in Table II and also their displacement to the rehabilitation center. The majority of works cited did not make evaluations using eligible participants, only healthy users.

REFERENCES

- [1] d. c. Cartilha, "Disabled people," Disabled people. Luiza Maria Borges Oliveira / Secretary of Human Rights of the Presidency of the Republic (SDH / PR) / National Secretary for the Promotion of the Rights of Persons with Disabilities (SNPD)[Portuguese version], 2012, p. 17.
- [2] S. Arlati, V. Colombo, G. Ferrigno, R. Sacchetti, and M. Sacco, "Virtual reality-based wheelchair simulators: A scoping review," *Assistive Technology*, 2019, pp. 1–12, doi: 10.1080/10400435.2018.1553079.
- [3] C. C. Caro, J. D. Costa, and D. M. C. Cruz, "The use of assistive devices for mobility and functional independence in subjects with stroke / the use of mobility assistive devices and the functional independence in stroke patients," vol. 26, no. 3. [Portuguese version], 2018, doi: 10.4322/2526-8910.ctoAO1117.
- [4] M. J. Scherer and S. Federici, "Why people use and don't use technologies: Introduction to the special issue on assistive technologies for cognition/cognitive support technologies," *NeuroRehabilitation*, vol. 37, no. 3, 2015, pp. 315–319, doi: 10.3233/NRE-151264.
- [5] N. W. John, S. R. Pop, T. W. Day, P. D. Ritsos, and C. J. Headleand, "The implementation and validation of a virtual environment for training powered wheelchair manoeuvres," *IEEE transactions on visualization and computer graphics*, vol. 24, no. 5, 2018, pp. 1867–1878, doi: 10.1109/TVCG.2017.2700273.
- [6] G. Vailland, F. Grzeskowiak, L. Devigne, Y. Gaffary, B. Fraudet, E. Leblong, F. Nouviale, F. Pasteau, R. Le Breton, S. Guegan et al., "User-centered design of a multisensory power wheelchair simulator: towards training and rehabilitation applications," 2019, doi: 10.1109/ICORR.2019.8779496.
- [7] Brazil, "Includes wheelchair procedures and wheelchair postural adaptation in the unified health system table of procedures, drugs, orthotics, prosthetics and special materials (opm)," in *Ordinance 1272*. Accessed: 04/04/19, [Portuguese version], 2013. [Online]. Available: [http://bvsms.saude.gov.br/bvs/saudelegis/gm/2013/prt1272_25_06_2013.html\\$](http://bvsms.saude.gov.br/bvs/saudelegis/gm/2013/prt1272_25_06_2013.html$)
- [8] ———, "Supplementary info," in *Ordinance No. 1272 - Annex I*. Accessed: 04/04/19, [Portuguese version], 2013. [Online]. Available: [http://bvsms.saude.gov.br/bvs/saudelegis/gm/2013/anexo/anexo_prt1272_26_06_2013.pdf\\$](http://bvsms.saude.gov.br/bvs/saudelegis/gm/2013/anexo/anexo_prt1272_26_06_2013.pdf$)
- [9] S. Mitsumura, Y. Akimoto, E. Sato-Shimokawara, and T. Yamaguchi, "Development of remote controllable visiting robot considering local situation," in *Micro-NanoMechatronics and Human Science (MHS)*, 2014 International Symposium on. IEEE, 2014, pp. 1–4, doi: 10.1109/MHS.2014.7006142.
- [10] D. C. Kamaraj, B. E. Dicianno, H. P. Mahajan, A. M. Buhari, and R. A. Cooper, "Interrater reliability of the power mobility road test in the virtual reality-based simulator-2," *Archives of physical medicine and rehabilitation*, vol. 97, no. 7, 2016, pp. 1078–1084, doi: 10.1016/j.apmr.2016.02.005.
- [11] H. P. Mahajan, B. E. Dicianno, R. A. Cooper, and D. Ding, "Assessment of wheelchair driving performance in a virtual reality-based simulator." *The journal of spinal cord medicine*, vol. 36, no. 4, 2013, pp. 322–332, doi: 10.1179/2045772313Y.0000000130.
- [12] M. K. MacGillivray, B. J. Sawatzky, W. C. Miller, F. Routhier, and R. L. Kirby, "Goal satisfaction improves with individualized powered wheelchair skills training," *Disability and Rehabilitation: Assistive Technology*, vol. 13, no. 6, 2018, pp. 558–561, doi: 10.1080/17483107.2017.1353651.
- [13] L. Manring, J. Pederson, D. Potts, B. Boardman, D. Mascarenas, T. Harden, and A. Cattaneo, "Augmented reality for interactive robot control," in *Special Topics in Structural Dynamics & Experimental Techniques, Volume 5*. Springer, 2020, pp. 11–18, doi: 10.1007/978-3-030-12243-0_2.
- [14] A. Borresen, C. Wolfe, C.-K. Lin, Y. Tian, S. Raghuraman, K. Nahrstedt, B. Prabhakaran, and T. Annaswamy, "Usability of an immersive augmented reality based telerehabilitation system with haptics (artesh) for synchronous remote musculoskeletal examination," *International Journal of Telerehabilitation*, vol. 11, no. 1, 2019, pp. 23–32, doi: 10.5195/ijt.2019.6275.
- [15] S. B. Abdallah, F. Ajmi, S. B. Othman, S. Vermandel, and S. Hammadi, "Augmented reality for real-time navigation assistance to wheelchair users with obstacles' management," in *International Conference on Computational Science*. Springer, 2019, pp. 528–534, doi: 10.1007/978-3-030-22750-0_47.
- [16] M. D. Wiederhold, "Augmented reality: Poised for impact," 2019, doi: 10.1089/cyber.2019.29140.mdw.
- [17] M. Eckert, J. Volmerg, and C. Friedrich, "Augmented reality in medicine: Systematic and bibliographic review. *jmir mhealth and uhealth*, 7 (4), e10967," 2019, doi: 10.2196/10967.
- [18] S. Massengale, D. Folden, P. McConnell, L. Stratton, and V. Whitehead, "Effect of visual perception, visual function, cognition, and personality on power wheelchair use in adults," *Assistive Technology*, vol. 17, no. 2, 2005, pp. 108–121, doi: 10.1080/10400435.2005.10132101.
- [19] A. Affanni, R. Bernardini, A. Piras, R. Rinaldo, and P. Zontone, "Driver's stress detection using skin potential response signals," *Measurement*, vol. 122, 2018, pp. 264–274, doi: 10.1016/j.measurement.2018.03.040.
- [20] A. Lanatà, G. Valenza, A. Greco, C. Gentili, R. Bartolozzi, F. Bucchi, F. Frendo, and E. P. Scilingo, "How the autonomic nervous system and driving style change with incremental stressing conditions during simulated driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 3, 2014, pp. 1505–1517, doi: 10.1109/TITS.2014.2365681.
- [21] J. Hernandez, D. McDuff, X. Benavides, J. Amores, P. Maes, and R. Picard, "Autoemotive: bringing empathy to the driving experience to manage stress," in *Proceedings of the 2014 companion publication on Designing interactive systems*. ACM, 2014, pp. 53–56, doi: 10.1145/2598784.2602780.
- [22] J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Transactions on intelligent transportation systems*, vol. 6, no. 2, 2005, pp. 156–166, doi: 10.1109/TITS.2005.848368.
- [23] A. Barriga, J. M. Conejero, J. Hernández, E. Jurado, E. Moguel, and F. Sánchez-Figuerola, "A vision-based approach for building telecare and telerehabilitation services," *Sensors*, vol. 16, no. 10, 2016, p. 1724, doi: 10.3390/s16101724.
- [24] Y. Lin, S. Song, and M. Q.-H. Meng, "The implementation of augmented reality in a robotic teleoperation system," in *2016 IEEE International Conference on Real-time Computing and Robotics (RCAR)*. IEEE, 2016, pp. 134–139, doi: 10.1109/RCAR.2016.7784014.
- [25] G. C. Burdea, "Invited paper: Low-cost telerehabilitation," in *Proceedings of Third International Workshop on Virtual Rehabilitation*, 2004, pp. 3297–3300.
- [26] D. Caetano, F. Mattioli, E. Lamounier, and A. Cardoso, "[demo] on the use of augmented reality techniques in a telerehabilitation environment for wheelchair users' training," in *Mixed and Augmented Reality*

- (ISMAR), 2014 IEEE International Symposium on. IEEE, 2014, pp. 329–330, doi: 10.1109/ISMAR.2014.6948473.
- [27] J. PGVR2, “PG Drivers Technology - Wheelchair Control System.” Accessed: 07/04/16, 2016. [Online]. Available: http://www.cw-industrialgroup.com/getattachment/439a3b95-d06f-4c98-bbd8-7f28cb9b1f71/vr2_brochure
- [28] S. Hansen and T. V. Fossum, “Refactoring model-view-controller,” *Journal of Computing Sciences in Colleges*, vol. 21, no. 1, 2005, pp. 120–129.
- [29] D. Caetano, F. Mattioli, E. Lamounier, A. Cardoso, and E. Naves, “Adaptation of a usb interface for vr2 joystick applied to the training of wheelchair users,” *Assistive Technology: Development and Application*, vol. 1, no. 1, 2018, pp. 261–267, Portuguese version.
- [30] M. Benedek and C. Kaernbach, “Ledalab.” Accessed: 16/06/19, 2016. [Online]. Available: <http://www.ledalab.de/>
- [31] ——, “A continuous measure of phasic electrodermal activity,” *Journal of neuroscience methods*, vol. 190, no. 1, 2010, pp. 80–91, doi: 10.1016/j.jneumeth.2010.04.028.
- [32] “Heart rate variability - Kubios HRV,” 2017. [Online]. Available: www.kubios.com
- [33] “E4 data - IBI expected signal – Empatica Support.” Accessed: 13/02/20, 2020. [Online]. Available: <https://support.empatica.com/hc/en-us/articles/360030058011-E4-data-IBI-expected-signal>
- [34] R. Baevsky and A. Berseneva, “Use of kardivar system for determination of the stress level and estimation of the body adaptability,” *Methodical recommendations. Standards of measurements and physiological interpretation*. Moscow-Prague, 2008.
- [35] A. Rizzo, S. Thomas Koenig, and T. B. Talbot, “Clinical results using virtual reality,” *Journal of Technology in Human Services*, 2019, pp. 1–24, doi: 10.1080/15228835.2019.1604292.
- [36] J. Nunnerley, S. Gupta, D. Snell, and M. King, “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*, vol. 12, no. 4, 2017, pp. 417–423, doi: 10.1080/17483107.2016.1176259.
- [37] K. V. Bletzer, “Visualizing the qualitative: making sense of written comments from an evaluative satisfaction survey,” *Journal of educational evaluation for health professions*, vol. 12, 2015, doi: 10.3352/jeehp.2015.12.12.
- [38] J. Jurečková, J. Picek, and M. Schindler, *Robust statistical methods with R*. Chapman and Hall/CRC, 2019.
- [39] T. U. Islam, “Ranking of normality tests: An appraisal through skewed alternative space,” *Symmetry*, vol. 11, no. 7, 2019, p. 872, doi: 10.3390/sym11070872.
- [40] hrvKubios, “Kubios hrv (ver. 3.1) user’s guide,” in *Kubios Oy*. Accessed: 11/07/19, 2018. [Online]. Available: https://www.kubios.com/downloads/Kubios_HRV_Users_Guide.pdf
- [41] J. J. Braithwaite and D. G. Watson, “Issues Surrounding the Normalization and Standardisation of Skin Conductance Responses (SCRs). Technical Research Note. Selective Attention & Awareness Laboratory (SAAL), Behavioural Brain Sciences Centre, School of Psychology, University of Birmingham,” 2015.
- [42] S. Bodkin, J. Hart, and B. C. Werner, “Common statistical tests,” in *Basic Methods Handbook for Clinical Orthopaedic Research*. Springer, 2019, pp. 153–161, doi: 10.1007/978-3-662-
- [43] R. Alves, C. R. Lopes, and M. Veloso, “Checkpoint coverage with wifi localization for indoor service robots,” in *JCAI 2016 WS Autonomous Mobile Service Robots*, 2016, pp. 10–17. [Online]. Available: <https://www.dropbox.com/sh/wp6abcswxstce3f/AADtfgehd7j0I9cNgZP2Zoua?dl=0>
- [44] Y. El Miedany, “Virtual reality and augmented reality,” in *Rheumatology Teaching*. Springer, 2019, pp. 403–427.
- [45] F. Mattioli, D. Caetano, A. Cardoso, E. Naves, and E. Lamounier, “An experiment on the use of genetic algorithms for topology selection in deep learning,” *Journal of Electrical and Computer Engineering*, vol. 2019, 2019, doi: 10.1155/2019/3217542.
- [46] S. Rashid and S. A. Razak, “Big data challenges in 5g networks,” in *2019 Eleventh International Conference on Ubiquitous and Future Networks (ICUFN)*. IEEE, 2019, pp. 152–157, doi: 10.1109/ICUFN.2019.8806076.



Daniel Caetano received the B.S. degree (2009) and the M.S. degree (2012) in Computer Engineering, from the Federal University of Uberlândia, Uberlândia, Brazil. He is currently working toward the Ph.D. degree in Computer Graphics and Biomedical Engineering in the Federal University of Uberlândia, Uberlândia, Brazil. His research interests include virtual/augmented reality, rehabilitation engineering, intelligent systems and brain-computer interfaces.



Caroline Valentini received the B.S. degree (2008), M.S. degree (2019) Biomedical Engineering, from the Federal University of Uberlândia, Uberlândia, Brazil. Has been working as an Occupational Therapist (OT) in the private sector and the Brazilian Unified Health System for 10 years. Her research interests include rehabilitation services and practices, Assistive Technology (AT).



Fernando Mattioli received the B.S. degree (2009), M.S. degree (2012) and Ph.D degree (2019) in Computer Engineering, from the Federal University of Uberlândia, Uberlândia, Brazil. His research interests include artificial intelligence, software engineering and virtual/augmented reality.



Paulo Camargos holds B.S. degree (2019) in Biomedical Engineering from Federal University of Uberlândia. He is currently working toward the M.S. degree in Computational Systems and Devices applied to Health. His research interest include Virtual and Augmented Reality applied in to the Biomedical Engineering.



Thiago Sá received the B.S. degree (2006) in Mechanical Engineering from University of São Paulo, São Carlos. He is currently working toward the M.S. in Biomedical Engineering at Federal University of Uberlândia. His research interest include rehabilitation robotics, biosignal data processing, control engineering and assistive technology.



Alexandre Cardoso received the B.S. degree (1987) in Electrical Engineering, M.S. degree (1991) and Ph.D degree (2002) in Electrical Engineering. He is an Associate Professor at Faculty of Electrical Engineering, Federal University of Uberlândia. His main research areas are virtual reality, augmented reality, education, virtual environments, human-computer.



Edgard Lamounier holds a B.S. degree (1984) in Electrical Engineering, M.S. degree (1989) and a Ph.D degree (1996) from Leeds University, England. He is an Associate Professor at Faculty of Electrical Engineering, Federal University of Uberlândia. His main research areas are Virtual and Augmented Reality applications in Mobile Learning, Engineering Simulation and Human Rehabilitation.



Eduardo Naves holds a B.S. degree (1994) in Electrical Engineering, M.S. degree (2001) and a Ph.D degree (2006) in Biomedical Engineering From Federal University of Uberlândia. He is an Associate Professor at Faculty of Electrical Engineering, Federal University of Uberlândia. His main research areas are Assistive Technology, Rehabilitation Engineering and Biomechanical Analysis of Human Movement.

Research Article

An Experiment on the Use of Genetic Algorithms for Topology Selection in Deep Learning

Fernando Mattioli , **Daniel Caetano**, **Alexandre Cardoso**, **Eduardo Naves**, and **Edgard Lamounier**

Faculty of Electrical Engineering, Federal University of Uberlândia, Uberlândia, MG, Brazil

Correspondence should be addressed to Fernando Mattioli; mattioli.fernando@gmail.com

Received 1 June 2018; Revised 26 October 2018; Accepted 3 December 2018; Published 15 January 2019

Guest Editor: Francesco Camastra

Copyright © 2019 Fernando Mattioli et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The choice of a good topology for a deep neural network is a complex task, essential for any deep learning project. This task normally demands knowledge from previous experience, as the higher amount of required computational resources makes trial and error approaches prohibitive. Evolutionary computation algorithms have shown success in many domains, by guiding the exploration of complex solution spaces in the direction of the best solutions, with minimal human intervention. In this sense, this work presents the use of genetic algorithms in deep neural networks topology selection. The evaluated algorithms were able to find competitive topologies while spending less computational resources when compared to state-of-the-art methods.

1. Introduction

Many digital signals, including images, present an hierarchical nature, in which higher level features are obtained through composition of lower level ones [1]. Deep neural networks (DNN) are a particular class of artificial neural networks (ANN), composed by stacked layers which explore this hierarchical behavior through automatic feature extraction [2]. With the addition of more layers and more processing units in each layer, DNN are able to achieve significant performance in problems of increasing complexity [3]. DNN have presented impressive results, specially in classification and regression applications, including image recognition and computer vision [4].

The design of neural network topologies depends on previous domain knowledge and expertise. Recently, the development of methods which minimize human interference has been discussed by researchers and practitioners [5]. Currently used approaches include random search, grid search, and transfer learning. In random search, the solution space is randomly sampled, whereas in grid search, all possible combinations of a reduced solution

subspace are evaluated [6]. In transfer learning, networks exhaustively trained with standard datasets are adapted to another application, by fine-tuning its weights with the target dataset [7].

These approaches present some limitations which narrow down their application. In pure random search, the effort of finding a good solution is not rewarded since other solutions will be equally sampled from the search space. Grid search can be impractical if the search space is too big and also subjected to local optimum convergence if a non-representative subspace is selected. In transfer learning, the complexity of existing and available topologies can make them unsuitable for the required application, for example, if severe real-time constraints are to be satisfied. In this sense, the development of methods to assist the design of new ANN topologies, with minimal intervention and limited computational resources, remains an important research topic.

The definition of a DNN topology for a given problem can be considered an optimization problem, given the high number of parameters to be chosen [8]. Some of these parameters, including number of layers, neurons per layer, activation functions, and learning algorithms, form a vast

search space making exhaustive search practically impossible [9]. Therefore, the need for robust optimization algorithms arises.

Genetic algorithms (GA) provide direct search, inspired by the theory of evolution, being widely applied in complex optimization problems with lots of parameters [10, 11]. GA are a subclass of evolutionary computation, which has been successfully used to assist ANN design [12, 13]. Through genetic operators, GA are capable of exploring promising search subspaces (by combining good solutions) and also escape local optima, reducing the amount of computational resources needed when compared to exhaustive search [9, 13].

The context presented in the above paragraphs motivated the investigation presented in this work. Two main objectives were defined: (i) To assess the efficiency of a search method using GA, with minimal human intervention, in DNN topology selection. (ii) To evaluate the contribution of a fitness prediction method in the algorithm performance. We believe that the investigation of these research objectives should help filling two gaps, namely, the appreciation of good solutions (found in the search space) and the economy of computing resources by discarding (i.e., not evaluating) solutions predicted to be potentially bad.

2. Background and Related Work

Deep learning allows computational models with multiple computing layers to process data in different levels of abstraction [1]. In addition, deep learning models provide automatic feature extraction capabilities, by learning the representation (features) together with the input/output mapping [2, 3].

Training a deep learning model from scratch requires two main steps: selection of the model topology and presentation of input/output data. The fundamental building blocks of a network topology are the processing layers, composed by neurons, convolutional filters, activation functions, dropout rate, and spatial reductions such as pooling and stride [14]. These building blocks form the model hyperparameters, and their choice is known to directly affect the network learning speed and performance [15].

Genetic algorithms are a set of evolutionary computing algorithms in which the solution space is explored through recombination of previous explored solutions [16]. In an analogy to natural selection, better solutions have more probability to be chosen for recombination, while bad solutions tend to be discarded [10]. The result of this recombination process is an offspring of new solutions, in which individuals inherit different characteristics from their parents [9].

In standard GA, at first, a population of candidate solutions is randomly generated. Each solution is represented by a data structure (chromosome), containing the parameters to be tuned. These solutions are evaluated by a fitness function, representing the performance of the solution on the target problem with higher values being returned by the best solutions. In the next step, a selection algorithm is used

to choose existing solutions for recombination. The chosen individuals are then subjected to the genetic operators of crossover and mutation to create an offspring of new solutions (generation). In crossover, elements of selected chromosomes are exchanged, creating new solutions inheriting characteristics of their parents. In mutation, an element of the chromosome is randomly changed, producing a new solution. The solutions in the new generation will be evaluated and will replace the previous population, optionally keeping some of its best individuals through elitism [11, 17].

Recently, the use of GA in deep learning hyperparameter optimization has been addressed by many researchers. In [4], a GA is used to evolve topologies of large-scale DNN. An encoding schema, based on the configuration of each DNN layer, is proposed. The adaptation of the genetic operators of crossover and mutation for the topology selection problem is also presented. Reference [9] presents the use of evolutionary optimization and convolutional neural networks in a cancer diagnosis application. A fixed topology of 3 layers is used, and the hyperparameters of these 3 layers (filters, kernel size, and pooling window) are encoded in the chromosome. A genetic programming method to evolve convolutional neural network architectures is also presented in [13]. In the proposed method, solutions are represented by a set of predetermined building blocks, together with their connections. Solutions are evaluated based on their error rates and number of parameters.

The results reported in these works confirm the benefits of the application of GA to deep learning hyperparameter optimization. However, the computational cost of these approaches is still high, given the number of solutions that must be evaluated by the GA. Therefore, techniques to optimize the use of computation resources in this class of applications remain an important area of inquiry. An example of such a technique is presented in [18]. In their work, a method based on design space exploration to optimize DNN hyperparameters is presented. The proposed method uses an ANN to predict the performance of a given topology. Together with Pareto efficiency, this prediction prevents the waste of computational resources by not evaluating potentially bad solutions. In [19], another approach to minimize the number of hyperparameters of a deep learning model is presented. In their work, the authors demonstrate how pyramidal-structured convolutional neural networks present competitive results when compared to state-of-the-art models, while scaling down the number of parameters and reducing the use of computational resources.

This work aims at contributing to the field of deep learning hyperparameter optimization by providing an evolutionary method for network topology selection with minimal human intervention. The proposed method might be used by researchers and practitioners to define an initial working topology, without requiring previous expertise on the field. The initial topology can be further tuned and optimized according to specific needs such as dataset size and available computational resources. In the next sections of this paper, details of the proposed selection method will be presented.

3. Materials and Methods

In the experiments presented in this paper, four topology search methods (described later in this section) were evaluated. The previously reported MNIST and CIFAR-10 datasets were used to support this study and are available at DOI 10.1109/MSP.2012.2211477 [20] and in [21]. For the MNIST dataset, 1024 samples were randomly extracted for training, while 256 samples (distinct from the training set) were chosen to test the neural networks. With the CIFAR-10 dataset, 2000 samples for training and 500 samples for test were used. To provide fair comparisons, all search methods used the same data. Model performance was measured as the test accuracy, i.e., the number of test patterns correctly classified divided by the total number of test patterns, as follows:

$$\text{acc}(\%) = \frac{n_C}{n_C + n_W} \times 100, \quad (1)$$

where n_C is the number of test patterns correctly classified and n_W is the number of wrong classifications.

In order to represent model complexity, the number of connections in the model was used. These two metrics—test accuracy and number of weights—will be used to evaluate a solution with respect to the Pareto front, in a multiobjective optimization problem: maximize the test accuracy and minimize the number of weights.

The Caffe framework was used to implement the deep learning models evaluated in this work [22]. A customized genetic algorithm was developed using the C++ programming language and can be used as a reference implementation of the presented method. A personal computer without GPU acceleration was used to train the models.

In all evaluated search methods, each solution is represented by two stacks of layers, the first being convolutional layers and the second fully connected layers. Each layer is presented as a string descriptor, as shown in the examples of Figure 1. To facilitate parsing these descriptors, elements are separated by semicolon characters.

3.1. Random Search. To evaluate random search performance, a set of 4000 distinct solutions was randomly generated, following the search space restrictions presented in Table 1. These same restrictions were used by the GA-based methods.

In order to preserve the random nature of the search method, no guidance was set for model generation. When invalid models are created (for example, because of spacial reduction), they are replaced with valid ones. Each model is then trained for 100 epochs, using the standard Adam optimization method, with parameters $\alpha = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = 10^{-8}$ [23]. Batch sizes of 64 for the MNIST dataset and 50 for the CIFAR-10 dataset were used. To prevent resource waste, an early stopping of 10 epochs was set up, so that training is interrupted when no improvement in the test accuracy observed for 10 consecutive epochs.

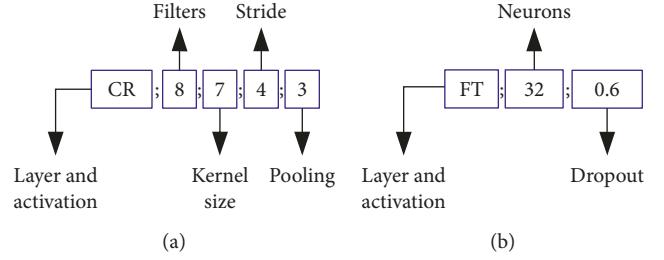


FIGURE 1: Layer descriptor examples. In (a), a convolutional layer, activated by ReLU, with 8 filters, a 7×7 kernel, stride of 4, and pooling of 3. In (b), a fully connected layer, activated by TanH, with 32 neurons and a dropout rate of 0.6.

TABLE 1: Search space restrictions.

Description	Value(s)
# convolutional layers	0–3
Convolutional filters	{2, 4, 8, 16, 32}
Kernel size	{3, 5, 7, 9, 11}
Stride	{1, 2, 3, 4}
Pooling	{2, 3, 4}
Activation functions	{Linear, Sigmoid, TanH, ReLU}
# fully connected layers	0–2
Neurons	{8, 16, 32, 64}
Dropout	{0.0, 0.2, 0.4, 0.6}
Activation functions	{Linear, Sigmoid, TanH, ReLU}

3.2. Grid Search. The objective of grid search is the exhaustive exploration of a solution space. However, this is impractical in deep learning topology optimization, given the high number of hyperparameters to be configured. Therefore, to provide a fair comparison setup, a more restrictive subset of the search space presented in Table 1 was chosen. This subset, presented in Table 2, contains a total of 3654 models, resulting from all combinations of its hyperparameter values. In grid search evaluation, discarded invalid solutions were not replaced, to preserve strict exploration of the reduced search space.

3.3. Genetic Algorithm. The GA-based search was conducted using the parameters presented in Table 3. By defining a population size of 200 and 20 generations, a maximum of 4000 solutions is to be evaluated. Crossover and mutation probabilities were set to 70% and 20%, respectively. An elitism parameter was configured so that the 10% fittest solutions are automatically propagated to the next generation. The remaining solutions will be obtained through recombination of the current population. The standard roulette wheel was used as the selection method so that each solution has a selection probability proportional to its fitness.

3.3.1. Chromosome. Each solution, s_i , $i = 1, \dots, 200$, is represented by two stacks of layer descriptors, one for convolutional layers and the other for fully connected layers. Therefore, each chromosome is composed of two vectors of strings, each string being the descriptor of a layer as

TABLE 2: Search space restrictions for grid search.

Description	Value(s)
# convolutional layers	0–3
Convolutional filters	{16, 32}
Kernel size	{3, 5}
Stride	1
Pooling	2
Activation function	ReLU
# fully connected layers	0–2
Neurons	{32, 64}
Dropout	0.4
Activation functions	{Sigmoid, TanH, ReLU}

TABLE 3: GA parameters.

Description	Value(s)
Population size	200
Generations	20
Crossover probability	0.7
Mutation probability	0.2
Elitism	0.1
Selection method	Roulette

presented in Figure 1. Convolutional layers descriptors contain the activation function, number of filters, kernel size, stride, and pooling size. Fully connected layer descriptors contain the number of neurons and the dropout probability. These solutions compose the initial population P , presented as follows:

$$P = \{s_1, s_2, \dots, s_{200}\}. \quad (2)$$

3.3.2. Evaluation. The fitness function, used to evaluate a given solution, is represented by the test accuracy of the solution in the corresponding classification problem. The solution's layer descriptors are used to build a deep learning model, which is trained using the same procedure described in Section 3.1. The best accuracy of the model when classifying test samples (samples which are not present in training), calculated using equation (1), is then used as the solution fitness. Once a model is trained, it is stored in a local database, to prevent it from being re-evaluated if it is present in further generations.

3.3.3. Selection. Solutions are selected for recombination based on their fitness by using the roulette wheel selection method. To provide better solutions with higher chances of being selected, a probability q_i is assigned to each solution s_i according to the following equation [11]:

$$q_i = \frac{f(s_i)}{\sum_{k=1}^{200} f(s_k)}, \quad i = 1, 2, \dots, 200, \quad (3)$$

where $f(s_i)$ is the fitness (i.e., test accuracy) of the i_{th} solution. The cumulative probability \hat{q}_i for solution s_i is defined as [11]

$$\hat{q}_i = \sum_{k=1}^i q_k, \quad i = 1, 2, \dots, 200. \quad (4)$$

After randomly generating a floating point $p \in [0, 1]$, solution s_i is chosen if $\hat{q}_{i-1} < p \leq \hat{q}_i$, given $\hat{q}_0 = 0$.

3.3.4. Crossover. Crossover is performed by swapping layers between 2 parent solutions, according to the randomly chosen cut points. Since each solution is composed by two stacks of layers (convolutional and fully connected), the recombination is conducted independently for each stack, using the standard single-point crossover [17]. For each layer stack, a random floating point $p \in [0, 1]$ is generated, and crossover is conducted if $p \leq 0.7$.

In Figure 2(a), two parent solutions (S1 and S2) are presented. Solution S1 has 3 convolutional and 1 fully connected layers, while S2 has 2 convolutional and 3 fully connected layers. The cut-points of both stacks of each parent are indicated using arrows in Figure 2(b). Finally, in Figure 2(c), the two children solutions resulting from the crossover operation are presented: S3 has one convolutional layer and 3 fully connected layers, while S4 presents 4 convolutional and 1 fully connected layers.

3.3.5. Mutation. The mutation operation consists on adding or removing layers at random positions in the solution's chromosome. As with crossover, mutation is conducted independently, for convolutional and fully connected layers. For each layer stack, a random floating point $p \in [0, 1]$ is generated. If $p \leq 0.2$, a second random floating point $p_{\text{add}} \in [0, 1]$ will be generated, indicating the type of mutation. If $p_{\text{add}} \leq 0.5$, a random layer will be added at a random position, in the corresponding layer stack. Otherwise, one of the existing layers will be removed from the corresponding layer stack.

In Figure 3(a), the two solutions previously obtained after crossover are presented. The arrows in Figure 3(b) indicate mutation points: for solution S3, a new random layer will be added after the arrow; for S4, the layer pointed by the arrow will be removed. The two solutions presented in Figure 3(c) are the resulting solutions of these operations.

After crossover and mutation, valid produced solutions are evaluated and added to the population. Invalid solutions are discarded, and new recombinations take place until the population is complete with 200 individuals.

3.4. Genetic Algorithm + Fitness Predictor. Many of the solutions evaluated in the previous methods present a low test accuracy after deep learning model training. This is an undesirable effect since computational resources were allocated to train a model which will be further discarded. For this reason, a modified version of the GA-based search was experimented. In this version, new generated solutions are subjected to a performance prediction step. The predicted performance will then be used to determine if this solution is worth evaluation. To achieve this, an ANN-based predictor

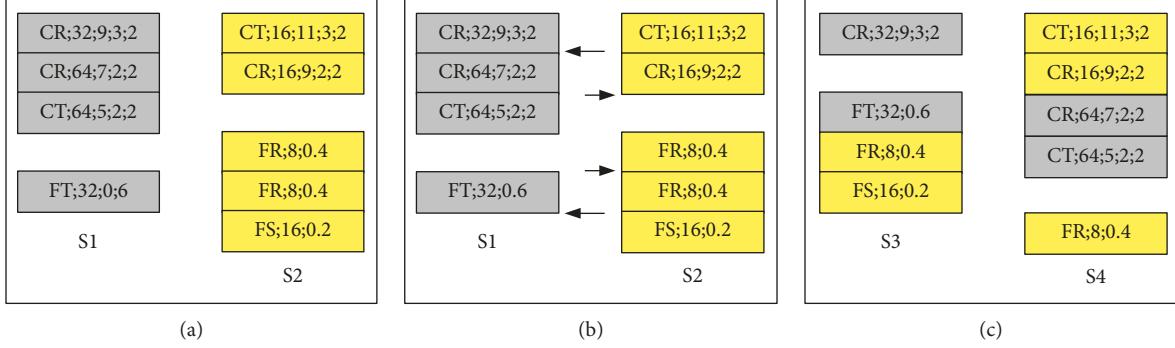


FIGURE 2: Crossover example: solutions S3 and S4 are produced from crossover of parents S1 and S2.

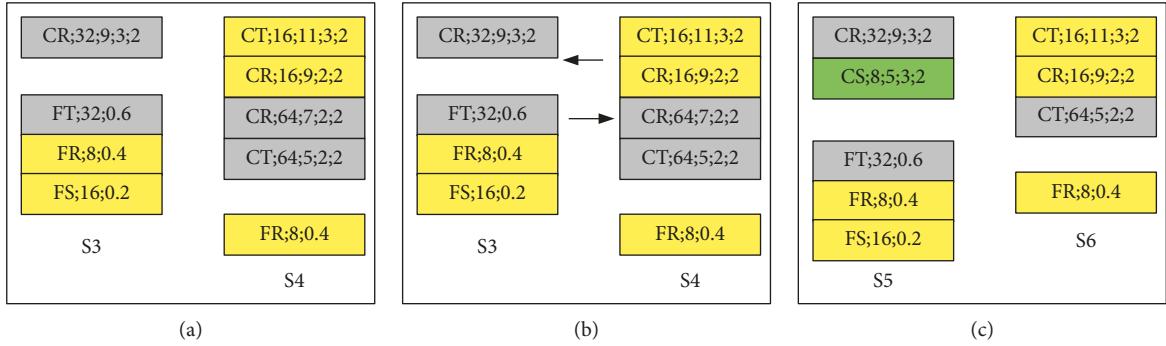


FIGURE 3: Mutation example: solutions S5 and S6 are produced after mutation of S3 and S4.

was set up, using knowledge gained from previous evaluations to estimate performance of new models.

The fitness predictor was setup with a fixed topology of 3 fully connected layers, with 64 neurons each, and a dropout rate of 0.4. The inputs of the predictor are formatted as a vector of integers containing the descriptors of the model layers. Each convolutional layer is represented by a 9-element vector, containing the number of filters, kernel size, stride, pooling, normalization, and activation function. Each fully connected layer is represented by a 6-element vector, containing the number of neurons, activation function, and the dropout rate. The final input vector contains 39 integers, composed of 3 convolutional descriptors and 2 fully connected descriptors. Empty layers are represented by a series of zeros in the corresponding positions. Figure 4 presents examples of the predictor inputs.

The fitness predictor was trained for 100 epochs, with an early stopping of 10 epochs, using the same optimizer as the MNIST and CIFAR-10 classifiers. The available training data (set of models previously trained and the corresponding fitness) is split into training (80% of the available data) and test (20% of the available data) patterns. When a solution is evaluated, it feeds the ANN predictor, which is retrained with the new pattern.

The decision on whether to evaluate a model or not is made upon the comparison of the predicted fitness with the already evaluated models. To this end, the Pareto efficiency was used, considering two dimensions: the solution fitness (performance of the classifier) and the number of weights

(complexity of the model). Dominant solutions (i.e., solutions that are superior to others in all attributes) have a 90% probability of being evaluated, while dominated solutions have a 10% probability of being evaluated. The complete algorithm used in this search method is described in the next paragraphs.

In Figure 5, the main tasks for generating the initial population are presented.

- (1) Initially, 10 solutions are randomly generated and evaluated.
- (2) The Pareto front is created, containing the dominant solutions among these 10 solutions.
- (3) The fitness predictor is trained, using the 10 solutions as input patterns. 8 solutions are randomly chosen to be used as training patterns, while the 2 remaining solutions are used as test patterns.
- (4) While the number of solutions in the population is less than the desired population size (200), new solutions are randomly generated.
- (5) The predictor described in task 3 is used to predict the fitness of the new generated solution. This predicted fitness, together with the number of weights of the solution, will be used to determine if its network model will be trained or not. The procedure for solution evaluation is presented in Figure 6.
- (6) If the solution (with predicted fitness) belongs to the Pareto front (dominant solution), it has a 90%

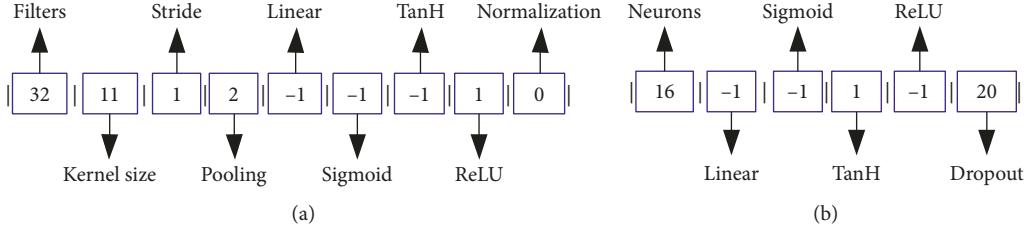


FIGURE 4: Predictor input examples. In (a), the descriptor CR;32;11;1;2 is represented. In (b), the descriptor FT;16;0.2 is represented.

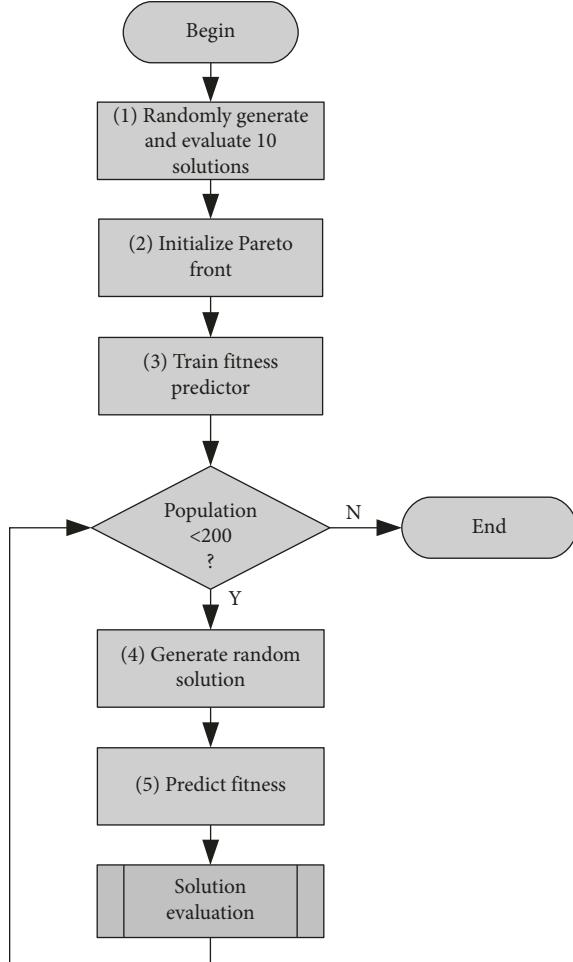


FIGURE 5: Initial population generation.

probability of being trained. Otherwise (dominated solution), this probability decreases to 10%. This step has the objective of providing more resources to potentially good solutions, while preventing resource wasting by bad solutions. If a random generated number (p) is below the respective threshold, the solution's model is trained.

- (7) After model training, the fitness predictor is updated, adding the new solution to the input patterns and being retrained from scratch. 80% of the available patterns are used for training, while the remaining patterns are used for the test.

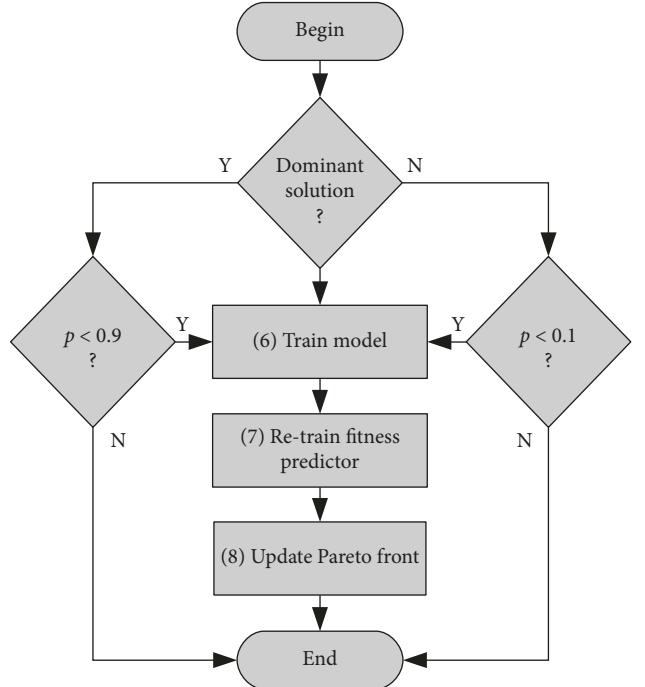


FIGURE 6: Solution evaluation.

- (8) Finally, the Pareto front is updated with the new solution.

After the initial population is generated, it is evolved through the desired number of generations. In the results reported in this work, the population is evolved for 20 generations. The steps for population evolution are presented in Figure 7.

- (9) If the number of generations is not achieved, a new generation is initialized with an elite of the 20 best solutions (10%) of the current population.
- (10) Elite solutions which were not previously evaluated (predicted solutions) have their models trained.
- (11) While the number of solutions in the current generation is less than the population size (200), two solutions are selected from the current population, using the roulette selection algorithm.
- (12) The crossover operator (as described in Section 3.3.4) is applied, with a probability of 70%.
- (13) The mutation operator (as described in Section 3.3.5) is applied, with a probability of 20%.

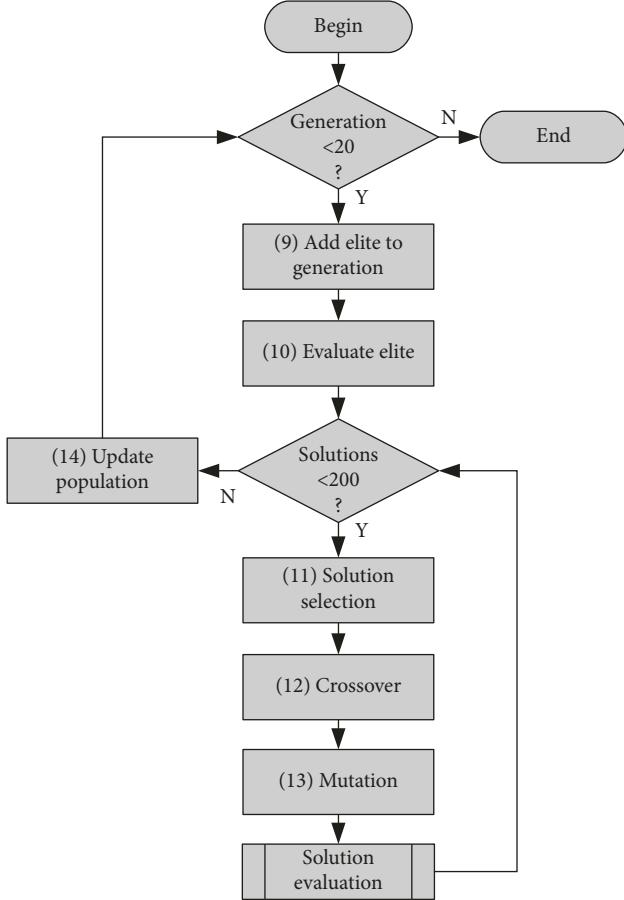


FIGURE 7: Population evolution.

- (14) When the number of desired solutions is achieved, the population is updated with the solutions in the current generation.

4. Results and Discussion

After testing the four search methods described in the previous section, it was possible to identify some contributions of the proposed method, which will be discussed in this section. In Table 4, the correlation between the test accuracy and some hyperparameters of the neural networks is presented.

The results presented in Table 4 show that no strong correlation between the isolated hyperparameters and the test accuracy can be observed. This behavior indicates that the performance of the network in the target problem must be affected by the ensemble of hyperparameters, i.e., network layers and layer blocks, as the hyperparameters of adjacent layers can directly impact the contribution of a given layer in the network topology.

Figure 8 presents the evaluated solutions for the 4 search methods with the MNIST and CIFAR-10 datasets, highlighting the Pareto front obtained with each one. It is possible to notice the concentration of evaluated solutions near the Pareto front, for all experiments. In grid search, the exploration of a restricted solution subspace causes the

TABLE 4: Correlation between model hyperparameters and test accuracy.

Hyperparameters	Correlation coefficient (ρ)	
	MNIST	CIFAR-10
# convolutional layers	-0.08	-0.25
# fully connected layers	-0.23	-0.21
# convolutional filters	0.12	0.13
Kernel size	0.13	0.11
# fully connected neurons	0.24	0.20
Dropout rate	0.02	0.00

prevalence of solutions with more weights, diverging from the other approaches which randomly explore the entire solution space. It is important to highlight that the Pareto front was determined using equal weights for both attributes (fitness and weights). Therefore, in these results, a small model with low accuracy is equally important as a big model with higher accuracy. To change this behavior, weights can be applied to each attribute, making, for example, accuracy more important than model complexity.

For the CIFAR-10 dataset, although some models achieved 100% accuracy on training samples, all evaluated models presented test accuracy below 60%. As the training and test samples were randomly chosen from the original dataset (preserving equal distribution of samples per class), we believe this results from the inherent complexity of the CIFAR-10 dataset, which would require more training samples to improve testing accuracy. To investigate this hypothesis, the best models were chosen to a second training stage, with more training samples from the original dataset. The results of this evaluation will be presented further in this section.

Figure 9 presents the evolution of both the population of solutions and its best individual across all generations for GA and GA + Predictor search methods. No significant difference was observed when comparing GA and GA + Predictor results, except for the fact that, with the MNIST dataset in the GA + Predictor experiment, the best solution was found earlier during evolution, which did not occur in the CIFAR-10 dataset. In all cases, the best solutions were kept in the next generations, given the use of elitism. Tables 5 and 6 summarize the results of the four search methods applied to both datasets.

The GA-based search methods were able to find competitive (and even superior) models compared to the random and grid search methods. However, the number of evaluated models was drastically reduced, as a result of the genetic operators which replicate good solutions in the further generations of the population, causing its convergence to the best evaluated solutions. In grid search, the reduction of the number of evaluated solutions is due to the selected solution space, which contained invalid solutions. This result would be different depending on the search space selection, but this would require prior domain knowledge, differently from the other methods which minimize human intervention.

Execution times were measured considering the time spent with optimization (population initialization, selection,

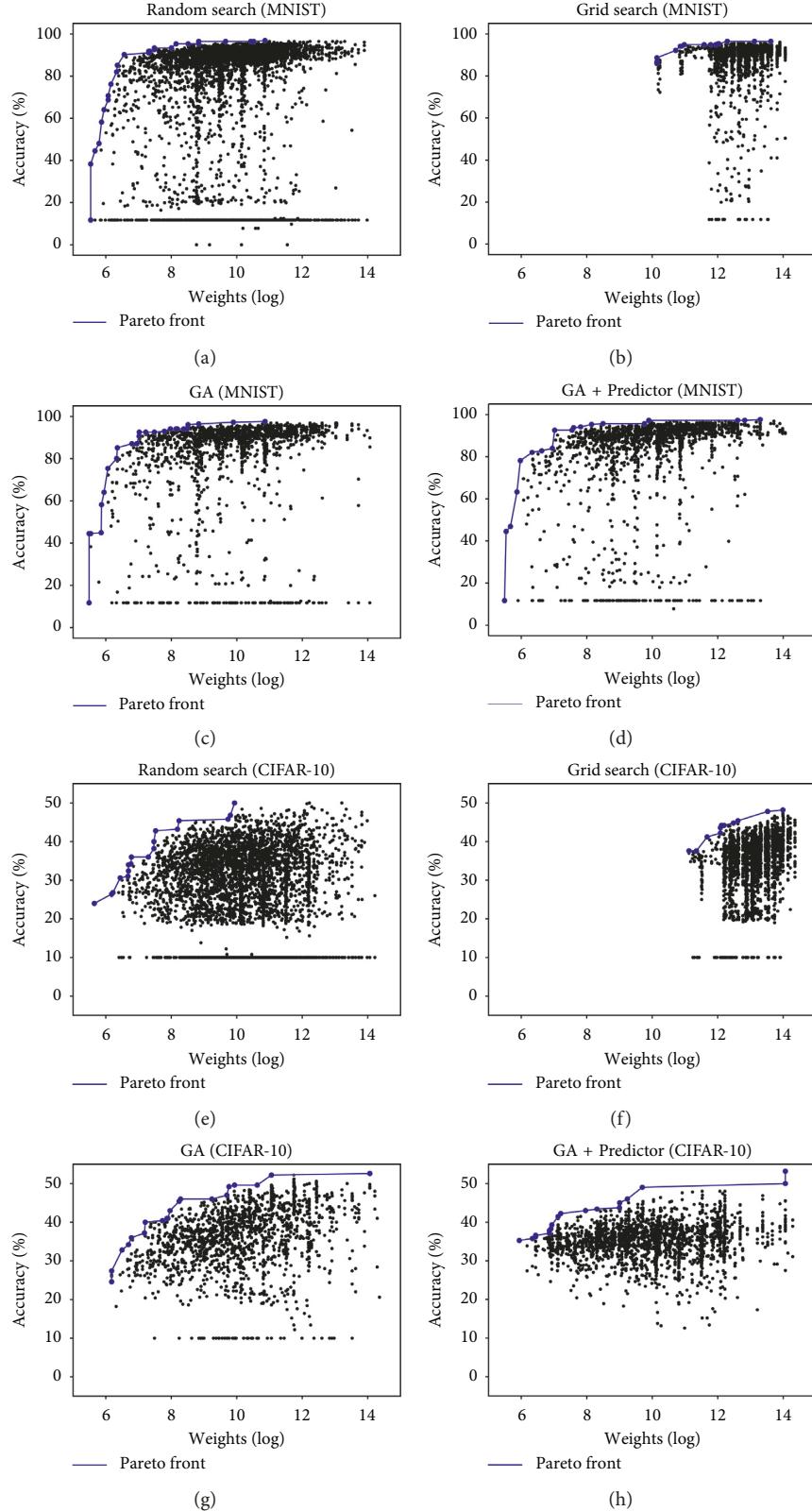


FIGURE 8: Pareto front obtained with each search method for the MNIST and CIFAR-10 datasets.

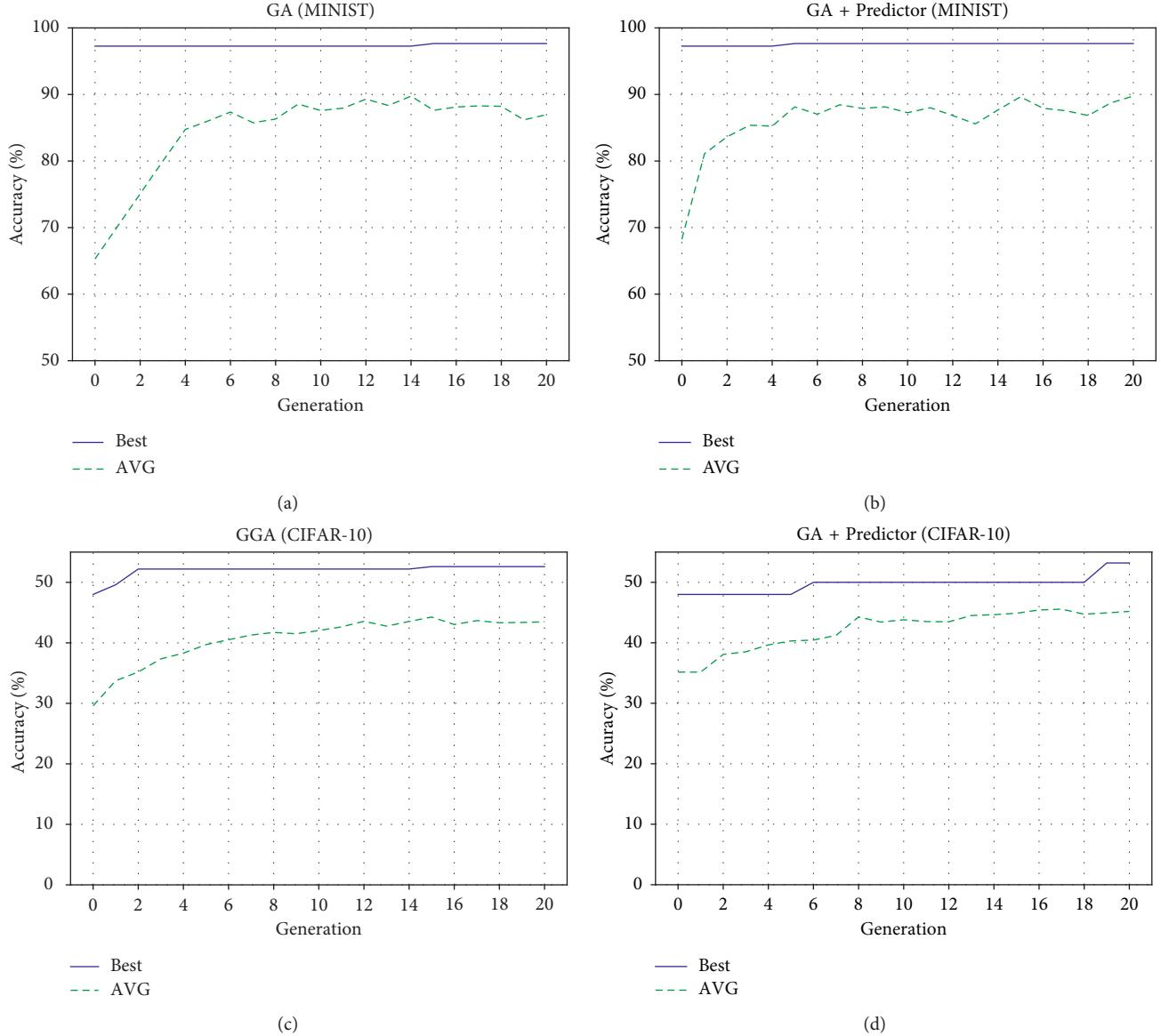


FIGURE 9: Population evolution for the MNIST and CIFAR-10 datasets.

TABLE 5: MNIST summary.

Method	Evaluated solutions	Model	Best solution	
			Accuracy (%)	Weights
Random search	4000	CTN;32;5;2;3 FL;16;0.6	97	52218
Grid search	2278	CR;32;3;1;2 CR;16;5;1;2 FS;32;0.4	96	218298
GA	1533	CT;32;5;2;3 FL;16;0.2	98	52218
GA + Predictor	1528	CRN;32;9;1;4 FL;64;0.4	98	595201

TABLE 6: CIFAR-10 summary.

Method	Evaluated solutions	Model(s)	Best solution	
			Accuracy (%)	Weights
Random search	4000	CT;16;5;2;4	50	20586
		CT;32;5;1;4	50	202442
		CT;16;7;1;4	50	544778
		FT;64;0.6		
Grid search	2754	CR;16;3;1;2 CR;32;5;1;2 FT;64;0.4	48	119364
GA	1733	CT;32;5;1;4 FT;64;0.6	53	1283146
GA + Predictor	975	CT;32;5;1;4 FS;64;0.4	53	1283146

genetic operators, etc.), fitness predictor training, and solution evaluation (model training). With the MNIST dataset, the GA search method was executed for 15 hours, with 12 minutes spent by the GA. The GA + Predictor method was executed for 14 hours, with 19 minutes spent by the GA and 9 minutes spent with predictor training. With the CIFAR-10 dataset, GA search was executed for 45 hours (8 minutes spent by the GA) and GA + Predictor search was executed for 24 hours (13 minutes spent by the GA and 8 minutes spent with predictor training). In all these cases, remaining time (total execution time minus optimization and predictor training) was spent with model training. These values show the minimal cost associated with optimization (GA and GA + Predictor search) and fitness prediction (GA + Predictor only), when compared to deep learning model training.

The use of the predictor did not result in significant reduction in the number of evaluated models and execution time for the MNIST dataset, when compared to the standard GA. Also, the best solution found by the GA + Predictor method is far more complex than the one found by the standard GA, which indicates that the prediction efficiency must be improved for this dataset. However, with the CIFAR-10 dataset, a reduction of 43% on the number of evaluated models was observed, in addition to a significant decrease in execution time. The best solution found by the GA + Predictor method is very similar to the solution found by the standard GA method, except for the activation function and dropout rate in the fully connected layer. These results demonstrate the economy of computational resources observed when prediction accuracy is good enough to discard potentially bad models and indicate that further investigation on fitness prediction must be conducted to extend this benefit to other datasets.

In order to compare the performance of the solutions optimized using the GA with existing models, a larger subset of the MNIST and CIFAR-10 datasets was used. For the MNIST dataset, 10048 training samples and 2048 test samples were randomly selected, while preserving the other configuration parameters. For the CIFAR-10 dataset, 10000 training samples and 2000 test samples were used. As in [19], the optimized models were compared to some of the state-of-the-art models, namely, LENET [20] and the Caffe Cifar-10 model (C10). In addition, two pyramidal structured models presented in [19]—SPyr_Rev_LENET and

SPyr_Rev_C10—were also evaluated. To provide fair comparisons, all models were trained from scratch using the same training parameters, for a maximum of 100 epochs early stopping when no improvement on the test accuracy was observed for 10 consecutive epochs. Training and test accuracy, measured using equation (1), as well as the time required to train each model from scratch (Δt) were calculated. The results of this experiment are presented in Tables 7 and 8.

With the MNIST dataset samples, both standard GA (GANet) and GA + Predictor (GA + PNet) optimized models presented competitive performance when compared to state-of-the-art models, considering both accuracy and training time. In addition, GANet was the fastest model to train, presenting a decrease of 0.4% and 0.5% in training and test accuracy, when compared to the best model (LENET). These optimized solutions were obtained without human intervention, through the evolution of an initial population of randomized solutions. For CIFAR-10, the GANet model presented inferior performance, when compared to the benchmark results. The GA + PNet model showed slightly better results, with a decrease of 2-3% when compared to the benchmark average in the evaluated samples, despite requiring less training time than 2 of the benchmark models. However, as for the MNIST experiment, the proposed method's objective is to provide researchers and practitioners with competitive starting points, without requiring prior expertise or intervention. These starting models can and should be subjected to a fine-tuning process, in order to improve their performance for the target problem. Also, the classification improvement observed in this last experiment, when compared to the optimization stage, indicates that more challenging datasets might require larger training sets or bigger populations of solutions during optimization, at the cost of higher usage of computational resources. Further work will be conducted to investigate these hypotheses, as well as the improvement of the fitness predictor demonstrated in this work.

5. Conclusion

In this paper, a method for DNN topology selection using genetic algorithms was presented. The evolutionary-based

TABLE 7: Optimized and reference models (MNIST).

Model	Descriptor	Accuracy(%)		Weights	Δt (s)
		Train	Test		
LENET	CL;20;5;1;2	100	98.8	8131080	373
	CL;50;5;1;2				
	FR;500;0.0				
C10	CR;32;5;1;3	99.1	97.6	488042	4121
	CR;32;5;1;3				
	CR;64;5;1;3				
	FL;64;0.0				
SPyr_Rev (LENET)	CL;50;5;1;2	100	98.6	3271830	652
	CL;20;5;1;2				
	FR;500;0.0				
SPyr_Rev (C10)	CR;64;5;1;3	99.4	98.3	284042	3136
	CR;32;5;1;3				
	CR;32;5;1;3				
	FL;64;0.0				
GANet (MNIST)	CT;32;5;2;3	99.6	98.3	52218	242
GA + PNet (MNIST)	CRN;32;9;1;4 FL;64;0.4	99.7	98.8	595201	429

TABLE 8: Optimized and reference models (Cifar-10).

Model	Descriptor	Accuracy(%)		Weights	Δt (s)
		Train	Test		
LENET	CL;20;5;1;2	100	100	12132080	1514
	CL;50;5;1;2				
	FR;500;0.0				
C10	CR;32;5;1;3	97.8	98.6	882858	6953
	CR;32;5;1;3				
	CR;64;5;1;3				
	FL;64;0.0				
SPyr_Rev (LENET)	CL;50;5;1;2	99.5	99.2	3271830	2005
	CL;20;5;1;2				
	FR;500;0.0				
SPyr_Rev (C10)	CR;64;5;1;3	89.8	88.5	483850	13204
	CR;32;5;1;3				
	CR;32;5;1;3				
	FL;64;0.0				
GANet (CIFAR-10)	CTN;32;5;1;4 FT;64;0.6	75.6	75.1	1283146	3794
GA + PNet (CIFAR-10)	CT;32;5;1;4 FS;64;0.4	94.1	93.9	1283146	4870

techniques were able to achieve competitive results with minimal human intervention and using less computational resources, when compared to random and grid search. The optimized solutions were compared to state-of-the-art models, offering promising starting points given the vast search space of DNN hyperparameters.

The use of a fitness predictor in the model evaluation stage resulted in a significant reduction on the number of evaluated models and execution time for one of the explored datasets. This fact demonstrates the benefits of fitness prediction but suggests that the use of other prediction methods, as well as the optimization of the predictor itself may lead to different results. Such methods should include

accuracy prediction from early epochs and accuracy estimation from smaller subsets.

Future work will be conducted in order to evaluate the application of the proposed search methods in other datasets and applications, including image regression problems. Improvement of the fitness predictor and scalability of the search method to bigger datasets will also be investigated.

Data Availability

The previously reported MNIST and CIFAR-10 datasets were used to support this study and are available at <https://doi.org/10.1109/MSP.2012.2211477> [20] and in [21].

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

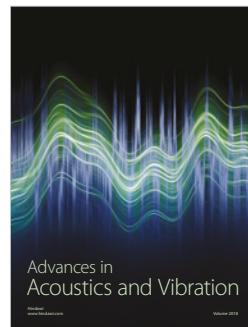
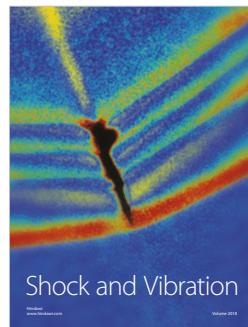
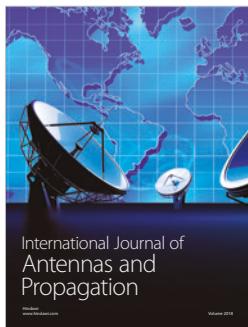
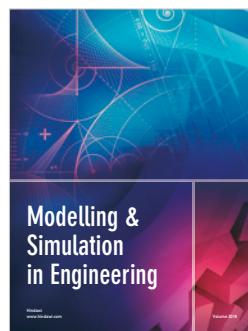
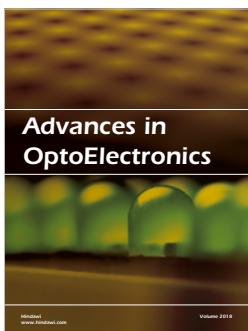
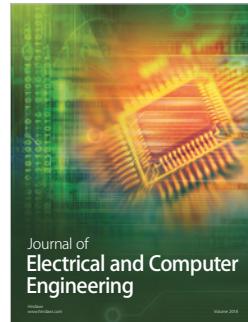
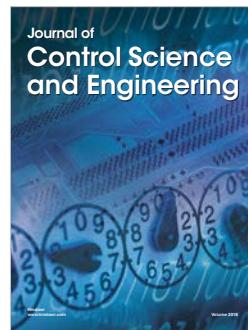
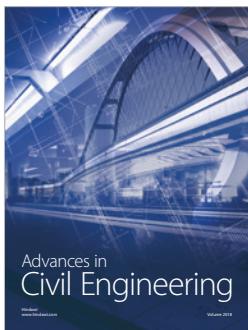
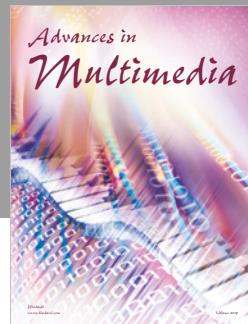
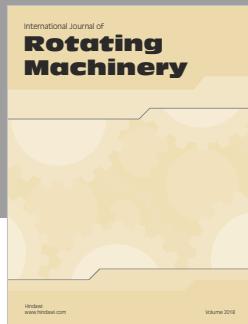
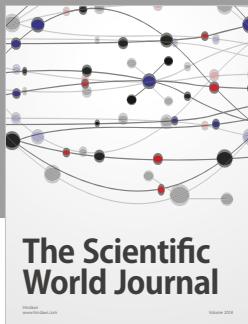
This research was supported by CAPES/Brazilian Ministry of Education & Culture, to whom the authors are deeply grateful.

Supplementary Materials

The source code used to support the findings of this study is included within the supplementary material of this paper. This material includes two C++ projects (ga-dnn-mnist and ga-dnn-cifar10), each one containing the source files used in dataset processing, deep neural network training, GA operations, and fitness prediction. A README file within each project provides details about individual source files and compilation instructions. Any additional request concerning the use of this material can be addressed to the corresponding author. (*Supplementary Materials*)

References

- [1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [2] J. Patterson and A. Gibson, *Deep Learning: A Practitioner's Approach*, O'Reilly Media, vol. 1, 1st edition, 2017.
- [3] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, Cambridge, MA, USA, 2016, <http://www.deeplearningbook.org>.
- [4] F. Assunção, N. Lourenço, P. Machado, and B. Ribeiro, "Evolving the topology of large scale deep neural networks," in *Proceedings of European Conference on Genetic Programming*, pp. 19–34, Springer, Parma, Italy, April 2018.
- [5] S. Huang, X. Li, Z. Cheng, Z. Zhang, and A. Hauptmann, "Gnas: a greedy neural architecture search method for multi-attribute learning," 2018, <http://arxiv.org/abs/1804.06964>.
- [6] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *Journal of Machine Learning Research*, vol. 13, pp. 281–305, 2012.
- [7] J. Yang, Y.-Q. Zhao, and J. C.-W. Chan, "Learning and transferring deep joint spectral-spatial features for hyperspectral classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 8, pp. 4729–4742, 2017.
- [8] B. Baker, O. Gupta, N. Naik, and R. Raskar, "Designing neural network architectures using reinforcement learning," 2016, <http://arxiv.org/abs/1611.02167>.
- [9] A. L. Rincon, A. Tonda, M. Elati, O. Schwander, B. Piwowarski, and P. Gallinari, "Evolutionary optimization of convolutional neural networks for cancer mirna biomarkers classification," *Applied Soft Computing*, vol. 65, pp. 91–100, 2018.
- [10] J. H. Holland, *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, MIT Press, Cambridge, MA, USA, 1992.
- [11] F. H. F. Leung, H. K. Lam, S. H. Ling, and P. K. S. Tam, "Tuning of the structure and parameters of a neural network using an improved genetic algorithm," *IEEE Transactions on Neural Networks*, vol. 14, no. 1, pp. 79–88, 2003.
- [12] K. O. Stanley and R. Miikkulainen, "Evolving neural networks through augmenting topologies," *Evolutionary Computation*, vol. 10, no. 2, pp. 99–127, 2002.
- [13] M. Suganuma, S. Shirakawa, and T. Nagao, "A genetic programming approach to designing convolutional neural network architectures," in *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 497–504, ACM, Berlin, Germany, July 2017.
- [14] S. Dutta and E. Gros, "Evaluation of the impact of deep learning architectural components selection and dataset size on a medical imaging task," in *Proceedings of Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications*, vol. 10579, March 2018.
- [15] J. Safarik, J. Jalowiczor, E. Gresak, and J. Rozhon, "Genetic algorithm for automatic tuning of neural network hyper-parameters," in *Proceedings of Autonomous Systems: Sensors, Vehicles, Security, and the Internet of Everything*, vol. 10643, article 106430Q, International Society for Optics and Photonics, Orlando, FL, USA, May 2018.
- [16] S. Russel and P. Norvig, *Artificial Intelligence: A Modern Approach*, Pearson, London, UK, 3rd edition, 2009.
- [17] M. Mitchell, *An Introduction to Genetic Algorithms*, MIT Press, Cambridge, MA, USA, 1998.
- [18] S. C. Smithson, G. Yang, W. J. Gross, and B. H. Meyer, "Neural networks designing neural networks: multi-objective hyper-parameter optimization," in *Proceeding of 2016 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pp. 1–8, IEEE, Austin, TX, USA, November 2016.
- [19] I. Ullah and A. Petrosino, "About pyramid structure in convolutional neural networks," in *Proceeding of 2016 International Joint Conference on Neural Networks (IJCNN)*, pp. 1318–1324, IEEE, Vancouver, BC, USA, July 2016.
- [20] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [21] A. Krizhevsky and G. Hinton, "Learning multiple layers of features from tiny images," Tech. Rep., University of Toronto, Toronto, ON, Canada, 2009.
- [22] Y. Jia, E. Shelhamer, J. Donahue et al., "Caffe: convolutional architecture for fast feature embedding," 2014, <http://arxiv.org/abs/1408.5093>.
- [23] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," 2014, <http://arxiv.org/abs/1412.6980>.





Av. Beira Gonçalves, 9500 - Setor 4 - Prédio 43.612 - Sala 219 - Bairro Agronomia
CEP 91.500-500 - Porto Alegre - RS - Tel. (51) 3308-4835 - Fax (51) 3308-7142
Home: www.sbc.org.br - E-mail: sbc@sbc.org.br

Paulo Roberto Ferreira Carvalho (UFPEL)

Presidente

Lisandro Zamboninetti Graxille (UFRGS)

Vice-Presidente

Renato de Mattos Galvão (UFRGS)

Diretor Administrativo

Carlos André Guimarães Ferraz (UFV)

Diretor de Finanças

Altigran Soares da Silveira (UFAM)

Diretor de Eventos e Criações Especiais

Mirila Moura Mero (UFMG)

Diretora de Educação

José Viterbo Filho (UFF)

Diretor de Publicações

Claudia Lage Rebello da Motta (UFFL)

Diretora de Plano de Estudos e Programas Especiais

Marcelo Dadachian Feitosa (CETEP)

Diretora de Secretaria Regional

Edson Norberto Cáceres (UFMS)

Diretor de Divulgação e Marketing

Roberto da Silva Siqueira (UFMG)

Diretor de Relações Profissionais

Ricardo de Oliveira Andrade (UNICAMP)

Diretor de Competições Científicas

Eduardo José de Araújo Macêdo (UFBA)

Diretor de Cooperação com Sociedades Científicas

Anílio Francisco Zocca (PUC-RS)

Diretor de Articulação de Empresas

Tendências e Técnicas em Realidade Virtual e Aumentada



EDITORES:

Marcelo de Paiva Guimarães
Enrico Pereira dos Santos Nunes
Rafael Rieder
Marcelo da Silva Housell

PRODUÇÃO GRÁFICA

Canal 6 Projetos Editoriais - www.canal6.com.br

Dados Internacionais de Catalogação na Publicação (CIP)

Tendências e Técnicas em Realidade Virtual e Aumentada/
Sociedade Brasileira de Computação - SBC - v. 1 (CD00)
- Porto Alegre: a Instituição, 2013.

Anual
ISSN 2177-6776

I. Realidade Virtual, Realidade Aumentada.
I. Gaiáneiros, Marcela de Paiva. II. Nunes, Enrico Pereira dos
Santos. III. Rieder, Rafael. IV. Housell, Marcelo da Silva.

CD0 006

Índice para catálogo sistemático:

I. Realidade Virtual e Aumentada: Ciência da Computação 006

TODOS OS DIREITOS RESERVADOS. Permitida a reprodução total ou parcial desde que citada a fonte (Tendências e Técnicas em Realidade Virtual e Aumentada, Porto Alegre-BS, Brasil). Na ausência da indicação da fonte ou expressa autorização da Instituição, é proibida a sua reprodução total ou parcial, por qualquer meio ou processo, especialmente por sistemas gráficos, microfotografia, fotográfico, fotostatico, fotocópia ou videofotográfico. Válida a manutenção dessa regras de proteção total ou parcial, bem como a inclusão de qualquer parte dessa obra em qualquer sistema de processamento de dados. Essas produções aplicam-se também às características da obra e à sua editoração. A violação dos direitos autorais é punível como crime (art. 184 e ss. do Código Penal, cf. Lei nº. 9.610, de 17-12-1998) com pena de prisão e multa, conjuntamente com busca e apreensão e indenizações diversas (arts. 22, 123, 124 e 126, da Lei nº. 5.988 de 34-12-1973, Lei dos Direitos Autorais).

Este periódico foi especialmente editado,
em tiragem limitada, a partir de comentários
desenvolvidos para os módulos e tópicos
apresentados durante o XVII Simpósio de
Realidade Virtual e Aumentada, realizado em
São Paulo, de 25 a 28 de Maio de 2015, promovido pela Sociedade Brasileira de Compu-
tação e organizado pelo Impet.

São Paulo, SP
2015

Prefácio.....	7
A Realidade Virtual e o Motor do Jogo Unity	9
Mario Popolin Neto, Isabele Andrade Agostinho, Diego Roberto Colombo Dias, Hélio Aparecido Redelli, José Renzo Ferreira Brega	
Self Quantificado e a Integração de Dados Cotidianos de Usuários em Aplicações para RA	24
Lúcio Silveira	
Sistemas de Realidade Virtual: Que Hardware Utilizar?	43
Jales Neto Bogoni, Otávio Basso Gomes, Márcio Saragiola Pinho	
Uso de Intereração Natural e Serious Games em Processos de Reabilitação Motora	65
Renan Vinícius Aranha, Clarissa Avelino Xavier de Camargo, Marcos Wagner de Souza Ribeiro	
Uma Experiência Espaço-temporal num Edifício Arquitetônico através da Realidade Misturada	87
Maria Amélia Eliseu, Luiz Henrique Falavigna, Darlan Silva Alves Delmondes, Jorge Henrique Garcia	
O Uso de Realidade Aumentada em Treinamento de Cadeirantes por Telereabilitação	102
Daniel Caetano, Fernando Mattioli, Alexandre Cardoso, Alcimar Soares, Edgard Lemosnir Jr.	

Chapter

1

O Uso de Realidade Aumentada em Treinamento de Cadeirantes por Telereabilitação

Daniel Caetano, Fernando Mattioli, Alexandre Cardoso, Alcimar Soares e Edgard Lamounier Jr.

Abstract

This work aims to investigate the use of Augmented Reality in wheelchairs for telerehabilitation purposes. At first, a prototype has been developed to enable the conduction of communication between three distinct environments – control, training and supervision. In the control environment, the patient visualizes the training environment and remotely controls a wheelchair (currently simulated by a robotic vehicle). In the training environment, the robotic vehicle is positioned against some real and virtual components, representing challenges faced by wheelchair users on their daily activities. Finally, on the supervision environment, the therapist can follow the execution of exercises by the patient and customize the training protocol. In addition, this chapter presents brief discussion on the use of Virtual / Augmented Reality on rehabilitation systems.

Resumo

Este trabalho propõe investigar o uso de Realidade Aumentada na telereabilitação de cadeirantes. Primeiramente, um protótipo foi desenvolvido para permitir a realização da comunicação entre três ambientes distintos - controle, treinamento e supervisão. No ambiente de controle, o paciente visualiza o ambiente de treinamento e controla remotamente uma cadeira de rodas (atualmente, simulada por um veículo robótico). No ambiente de treinamento, o veículo robótico está posicionado, bem como alguns componentes reais e virtuais, representando desafios enfrentados pelos usuários de cadeiras de rodas em suas atividades diárias. Finalmente, no ambiente de supervisão, o terapeuta pode seguir a execução de exercícios pelo paciente e personalizar o protocolo de treino. Além disso, este capítulo apresenta uma breve discussão sobre o uso de Realidade Virtual e Aumentada em sistemas de reabilitação.

1.1. Introdução

Um dos conceitos mais relevantes nos dias atuais, seja na promoção de grandes eventos, no projeto de novos equipamentos ou na construção de novas edificações é a acessibilidade [Lima, 2006]. Graças aos avanços relacionados às tecnologias empregadas para promover a acessibilidade, esta já não é vista como um diferencial, mas sim um pré-requisito nos mais diversos domínios incluindo aqueles relacionados ao trabalho, educação e lazer. Neste sentido, em muitos casos, limitações nas habilidades de locomoção já não impedem que cadeirantes realizem suas atividades diárias com eficiência.

Nos últimos anos, não é difícil perceber mudanças no cenário urbano, com o objetivo de prover acessibilidade a cadeirantes. Dentre estas mudanças estão calçadas mais acessíveis, rampas utilizadas como alternativas a escadas, elevadores elétricos utilizados no transporte público, além de banheiros públicos adaptados.

A cadeira de rodas, por sua vez, teve um papel fundamental na evolução da acessibilidade. Em todo o mundo, pode-se perceber os avanços que as cadeiras de rodas trouxeram para a qualidade de vida de seus usuários. No Brasil, em particular, mais de 24 milhões de pessoas possuem algum tipo de deficiência, muitos destes fazendo uso diário de cadeiras de rodas para sua locomoção [Schwarz, Haber, 2006].

Apesar de seu alto custo, cadeiras de rodas eletrônicas saíram, definitivamente, dos laboratórios de pesquisa para conquistar um crescente espaço nos ambientes urbanos. A cada nova geração de componentes eletrônicos, novas abordagens de controle são desenvolvidas e testadas. Além disso, a miniaturização torna possível acoplar a cadeira eletrônica a um computador embarcado, pesando menos que 500 gramas.

A rápida evolução da tecnologia traz consigo novos desafios: com estratégias de controle robustas e uma vasta gama de funcionalidades disponíveis, a adaptação dos usuários ao controle das cadeiras eletrônicas não é trivial. Desta forma, uma fase de treinamento passa a ser parte essencial do processo global de reabilitação. Para que este treinamento seja efetivo, torna-se necessário expor o usuário em treinamento às mesmas dificuldades encontradas em seu cotidiano.

No entanto, uma abordagem formal de treinamento, para cadeirantes, ainda é não é algo usual [Jenkins, 2002], [Karmarkar et. al., 2009], [Salatin et. al., 2010]. Dentre as principais dificuldades encontradas no treinamento de cadeirantes estão **restrições no tempo disponível para treinamento**, limitações quanto aos **recursos financeiros disponíveis**, limitações quanto ao **espaço físico disponível** e falta de um **processo padronizado de treinamento** [Routhier et. al., 2012].

Sendo assim, um primeiro desafio é apresentado: a necessidade de se construir um ambiente de treinamento que **reproduza fielmente as dificuldades encontradas pelos cadeirantes em seu dia-a-dia**, sem expor o paciente em treinamento aos riscos do mundo real. É desejável, ainda, que este ambiente de treinamento **permita um certo grau de customização e adaptação às necessidades** de cada paciente. Outra característica desejável é que este ambiente de treinamento possa **ser acessado remotamente**, uma vez que o acesso ao ambiente físico de treinamento é, por sua vez, uma dificuldade a ser superada. Uma vez que o ambiente de treinamento esteja disponível remotamente, vários usuários poderão acessar o mesmo ambiente de

treinamento, em momentos distintos, maximizando a taxa de utilização deste ambiente.

1.2. Fundamentações

Esta seção visa fundamentar conceitos básicos relacionados a Realidade Virtual e Realidade Aumentada. Em seguida, são apresentados alguns dos trabalhos relacionados que complementam o desenvolvimento da atual proposta.

1.2.1. Realidade Virtual (RV)

Dentre as várias definições dadas à Realidade Virtual (RV), pode-se destacar a seguinte: “Interface avançada para aplicações computacionais, onde o usuário pode **navegar e interagir**, em tempo real, em um ambiente tridimensional gerado por computador, usando dispositivos multissensoriais” [Cardoso, Kirner, Kelner, 2007].

Essencialmente, esta tecnologia permite a criação de um ambiente artificial, no qual o usuário tem a **impressão de não somente estar dentro deste ambiente**, mas também munido da capacidade de navegar no mesmo e interagir com os seus objetos, podendo até **alterar a realidade que o envolve**.

A fim de ilustrar este conceito, nota-se na Figura 1.2a, um usuário em um laboratório munido de um capacete (HMD). Por meio de técnicas de RV, este usuário é transportado para um ambiente virtual similar ao de uma cozinha - Figura 1.2.b

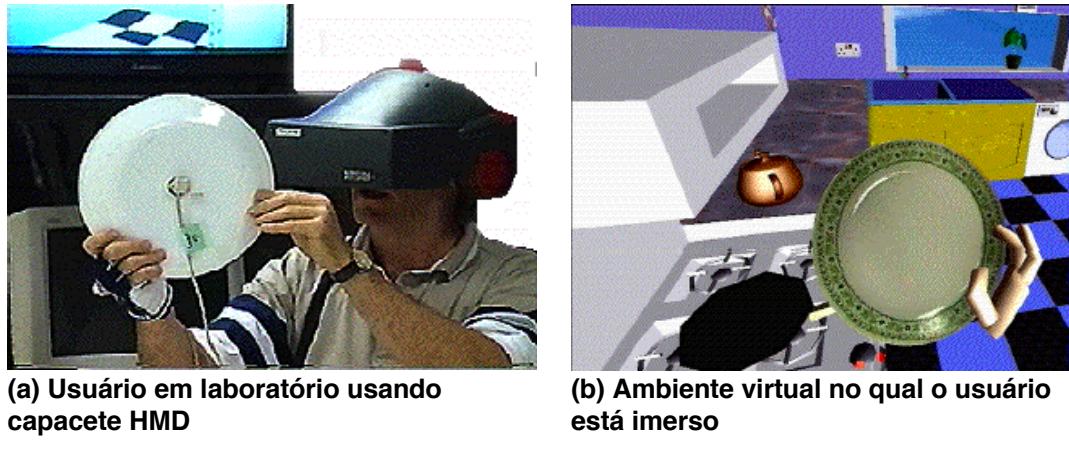
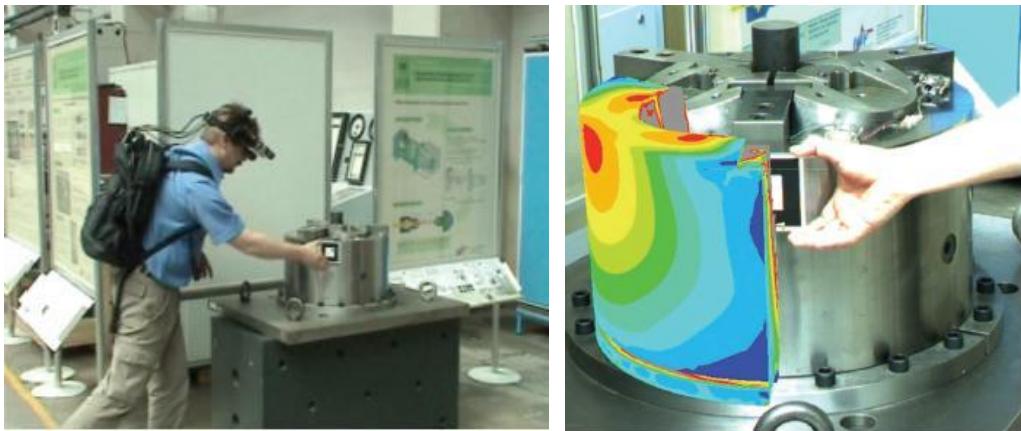


Figura 1.2. – Características de um ambiente em RV [Cardoso, Kirner, Kelner, 2007].

1.2.2. Realidade Aumentada (RA)

Realidade Aumentada (RA) consiste na inserção de objetos virtuais no ambiente físico, apresentada ao usuário **em tempo real**, com o apoio de **algum dispositivo tecnológico**, usando a **interface do ambiente real adaptada para visualizar e manipular os objetos reais e virtuais** [Kirner, Kirner, 2007]. Esta definição pode ser ilustrada por meio da Figura 1.3, que apresenta uma aplicação na qual um engenheiro pode visualizar as camadas de temperatura em determinada peça através de técnicas de RA [Weidlich, Scherer, Wabner 2008].



(a) Colocando um marcador em um sistema (b) Visualizando a malha de Elementos Finitos para estudo e análise de temperatura

Figura 1.3. – Características de um ambiente em RA [Weidlich, Scherer, Wabner 2008].

1.3. Trabalhos Relacionados

1.3.1. Telereabilitação para pacientes com AVC

O trabalho proposto por [Holden, Dyar, Cimadoro, 2007] descreve a arquitetura de um ambiente virtual remoto de treinamento (telereabilitação) para atender as necessidades de pacientes que tenham sofrido acidente vascular cerebral (AVC) e que não possuam, por razões diversas, **condições de se locomover até um centro de reabilitação**.

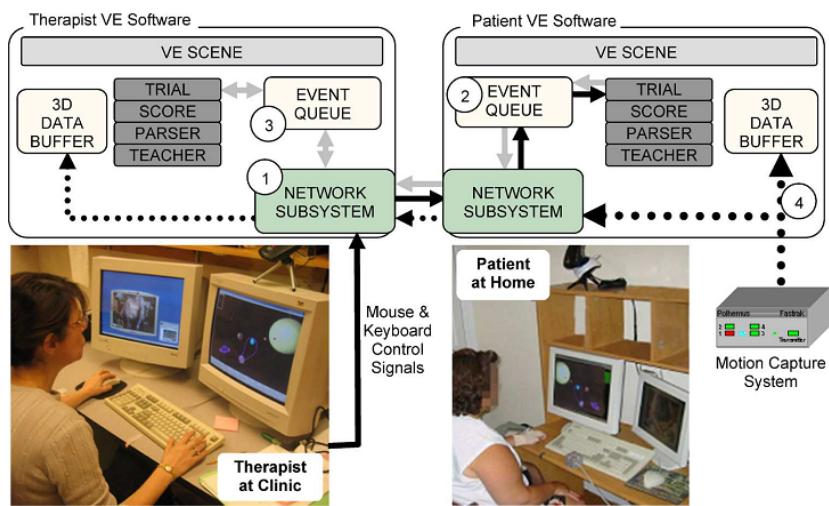


Figura 1.4. - Sistema de telereabilitação virtual. Paciente e terapeuta podem se comunicar através de um link de teleconferência e um monitor [Holden, Dyar, Cimadoro, 2007].

O sistema desenvolvido visa auxiliar na recuperação das funções da extremidade superior do paciente, para isto, **o paciente fica em frente de um computador com dois monitores**. Um monitor exibe a cena do **ambiente virtual** e o outro é utilizado para

videoconferência do terapeuta, conforme mostrado na Figura 1.4.

Em cada sessão de treinamento, o terapeuta determina quais movimentos serão realizados pelo paciente no ambiente virtual. Uma vez executado o movimento, módulos de avaliação atribuem uma nota ao movimento realizado.

Com base na arquitetura da Figura 1.4 são realizadas as seguintes ações:

- O equipamento de captura de movimento transmite as informações sobre os movimentos do braço do paciente para a exibição no ambiente virtual;
- O terapeuta controla, da clínica, o software e visualiza a cena idêntica àquela apresentada ao paciente na sua casa (segundo monitor);
- A câmera de vídeo permite ao terapeuta visualizar remotamente o espaço de trabalho do paciente [Holden, Dyar, Schwamm, Bizzi, 2003];

Os módulos principais e fluxo de dados durante as sessões de treinamento são mostradas no topo. São eles: TRIAL, SCORE PARSER, TEACHER, EVENTQUEUE, 3-D DATABUFFER. Estes módulos interagem para produzir as cenas no ambiente virtual.

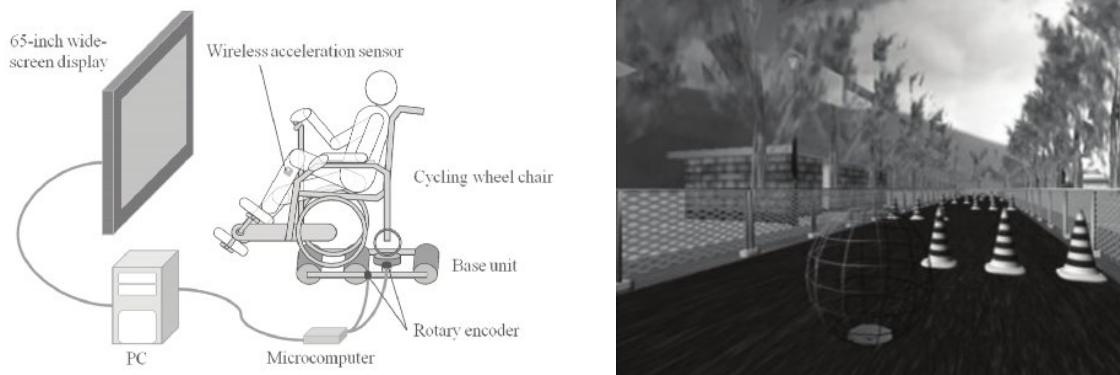
- O subsistema de rede passa dados entre o computador do paciente e terapeuta;
- As setas sólidas pretas mostram o fluxo de sinal de controle do teclado antes que os efeitos da mudança sejam exibidos no computador paciente;
- As setas cinza mostram o fluxo posterior do referido sinal de controle para sincronizar as cenas do ambiente virtual do terapeuta;
- As setas tracejadas mostram o fluxo de dados em 3-D, que são transmitidos em um fluxo contínuo (pontos menores indicam reduzida resolução temporal para os dados transmitidos);

Os resultados obtidos após os testes realizados com os pacientes apontam diversas vantagens do sistema de telereabilitação tais como eliminação do **tempo de deslocamento** e esforços secundários, permitindo que os pacientes fossem submetidos a sessões mais longas de treinamento com maior repetição dos exercícios durante cada sessão.

1.3.2. Realidade Virtual para avaliar habilidades na condução de cadeiras de rodas

No trabalho desenvolvido em [Sugita et. al., 2012], é apresentado um sistema de treinamento em RV para usuários de cadeiras de rodas manuais. A Figura 1.5.a apresenta a estrutura utilizada neste trabalho.

Foram desenvolvidos 4 cenários de teste, abordando as principais ações relacionadas à utilização de uma cadeira manual. Os exercícios realizados nestes cenários foram utilizados para avaliar 4 aspectos principais: equilíbrio, controle da velocidade, esterçao e desvio de obstáculos. A Figura 1.5.b apresenta um dos cenários de teste, utilizado na avaliação da capacidade de desvio de obstáculos.



(a) Sistema de Realidade Virtual para o treinamento de usuários de cadeiras manuais

(b) Teste de desvio de obstáculos

Figura 1.5. – Sistema de treinamento [Sugita et. al., 2012].

Os resultados obtidos pelos pesquisadores neste trabalho indicaram que a habilidade dos voluntários no controle das cadeiras foi aprimorada com a prática dos exercícios propostos. No entanto, foram observados diferentes padrões de melhoria quando comparados os resultados dos voluntários mais velhos com os resultados dos voluntários mais novos.

1.3.3. Realidade Virtual em cadeiras de rodas motorizadas

Neste estudo, desenvolvido por [Harrison et al., 2000], foi apresentada uma avaliação a respeito da utilização de dois ambientes virtuais não-imersivos no treinamento de usuários de cadeiras eletrônicas. Para tal, voluntários (aptos e cadeirantes) desempenharam uma série de atividades relacionadas à exploração de ambientes, manobra da cadeira e orientação, tanto no ambiente real como no ambiente virtual.

O primeiro ambiente virtual desenvolvido para este estudo representa uma sala fechada de um hospital, com algumas mesas e cadeiras posicionadas estrategicamente para que o usuário desvie destes obstáculos. O segundo ambiente virtual representa um andar inteiro de um hospital, composto por uma enfermaria, quatro quartos, uma sala de espera, uma cozinha, vários corredores, dentre outros. O ambiente virtual inclui ainda vários funcionários do hospital, além de outros pacientes em cadeiras de roda, todos se movimentando aleatoriamente pelo ambiente.

No primeiro ambiente, são realizados exercícios relacionados à realização de manobras na cadeira. São eles:

- Guiar a cadeira em linha reta, para frente, por 10 metros;
- Guiar a cadeira em linha reta, para trás, por 2 metros;
- Entrar com a cadeira de frente em um espaço estreito;
- Sair com a cadeira deste espaço estreito, em um movimento para trás;
- Realizar um giro de 180° em volta de um objeto estacionário;
- Realizar um movimento de “zigue-zague” em torno de uma série de objetos estacionários;

- Parar a cadeira repentinamente, sob comando.

No segundo ambiente, os usuários realizam tarefas de orientação. Para tal, é pedido que eles se dirijam a determinado ambiente do hospital, seguindo as placas de orientação nas paredes e desviando das pessoas e obstáculos no caminho.

Os dados obtidos nestes experimentos indicaram que, apesar do realismo apresentado pelos ambientes virtuais e da correta escolha das tarefas, a execução destas nos ambientes virtuais apresentou-se mais desafiadora em comparação ao ambiente real. Dentre os principais problemas apontados pelos voluntários está a ausência de visão periférica, principalmente nas tarefas relacionadas à reversão do sentido da cadeira e ao movimento de “zigue-zague”. Além disso, entre os dados coletados, o número de colisões e o número de manobras executadas durante a realização de uma dada tarefa apresentaram os melhores resultados para medida da eficiência do usuário no controle da cadeira.

1.3.4. Uso de ondas cerebrais para controle de cadeiras de rodas

O estudo desenvolvido por [Huang et al., 2012], propõe um novo paradigma eficaz e prático de uma interface cérebro-computador capaz de controlar uma cadeira de rodas virtual em 2D. O sistema utiliza como interface de controle os sinais biológicos EEG e EMG. Estes sinais definem as “tarefas mentais” ou comandos a serem executados.

O experimento consiste em guiar a cadeira de rodas virtual dentro de um cenário simulado com dimensões de 20x20 m, até que o paciente atinja o alvo ou colida com as laterais do ambiente. A qualquer momento, o terapeuta pode interferir interrompendo o experimento, caso necessário. Cada sessão consistia em dois conjuntos de atividades, realizados em média em 30 minutos.

A Figura 1.6 exibe todas as etapas de um experimento realizado por um usuário. conforme descrito abaixo:

- a. A cadeira de rodas começou no centro com e se encontra “Parada”;
- b. A cadeira de rodas começou a se mover para cima ao longo da barra de direção (cor azul) depois de receber o comando de Iniciar/Parar;
- c. A cadeira de rodas começou a fazer curva à direita após receber o comando vire à direita;
- d. A cadeira parou de girar e começou a se mover ao longo da barra de direção (cor azul) depois de receber o comando de Iniciar/Parar;
- e. A cadeira de rodas parou de se mexer e começou a fazer curva à direita após receber o comando vire à direita;
- f. A cadeira parou de girar e começou a se mover ao longo da barra de direção (cor azul) depois de receber o comando Iniciar/Parar;
- g. A cadeira atingiu o alvo;
- h. A simulação da cadeira de rodas de controle reiniciado;

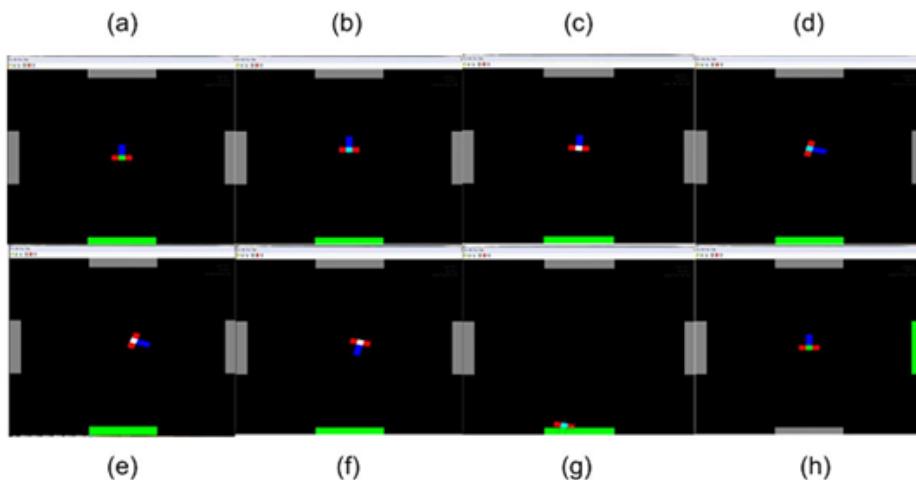


Figura 1.6. – Imagens de uma execução do controle da cadeira de rodas [Huang et al., 2012].

Os resultados alcançados pelos autores comprovam que a média da taxa de acerto para os pacientes submetidos ao primeiro conjunto de atividades é de 98,4%. Para os pacientes submetidos ao segundo conjunto de jogos, esta média foi de 100%, o que comprova que todos os pacientes obtiveram um aumento significativo em suas habilidades de controlar cadeiras de rodas.

1.4. Comparação entre RV ou RA para Treinamento de Cadeirantes

Vários pesquisadores tais como: [Holden et al., 2007], [Sugita et al., 2012], [Huang et al., 2012] e [Harrison et al., 2000] têm investigado a aplicação de técnicas de Realidade Virtual em reabilitação humana.

Estes ambientes virtuais de treinamento podem ser utilizados, por exemplo, em reabilitação cognitiva e treinamento de usuários de próteses. Dentre as principais características destes ambientes, pode-se destacar a preservação da integridade física do paciente. No entanto, muitos pacientes reportam dificuldades ao utilizar este tipo de sistema, dadas as limitações no grau de imersão que alguns ambientes virtuais proveem [Harrison et al., 2000]. A Tabela 1.1 resume algumas características dos estudos apresentados nas seções anteriores, relevantes no escopo do presente projeto de pesquisa.

A partir destas observações, pode-se notar a importância de se fornecer uma resposta em tempo real pelo sistema de treinamento. Por meio do uso de técnicas de Realidade Aumentada, torna-se possível otimizar a experiência do usuário, em comparação à imersão proporcionada pelos ambientes puramente virtuais. Além disso, considerando os requisitos da aplicação e as características de disponibilidade inerentes à Realidade Aumentada, pode-se notar que o mesmo ambiente remoto de treinamento pode ser compartilhado por vários usuários, em diferentes localizações [Silva, 2011].

Tabela 1.1. Resumo dos trabalhos correlatos

Trabalhos Relacionados	Tele reabilitação	Realidade Virtual	Realidade Aumentada	Processamento em Tempo Real
[Holden et. al, 2007]	Sim	Não	Não	Sim
[Sugita et. al, 2012]	Não	Sim	Não	Sim
[Harrison et. al, 2000]	Não	Sim	Não	Sim
[Huang et. al, 2012]	Não	Não	Não	Sim

1.5. Telereabilitação imersiva aplicada ao treinamento de cadeirantes

Por meio das fundamentações e comparações realizadas entre cada técnica a ser aplicada e as limitações vivenciadas pelos usuários e pesquisadores, busca-se desenvolver um protótipo de *Framework* que possua as seguintes vantagens:

- **Reduzido espaço físico necessário para treinamento:** uma vez que o ambiente remoto esteja construído e configurado, o paciente pode estar localizado em uma pequena sala, equipada com um computador, um HMD e dispositivos de interação (teclado, mouse, joystick etc.);
- **Acessibilidade do espaço de treinamento:** pelas características apresentadas no item anterior, percebe-se que o espaço de treinamento pode estar localizado em lugares mais acessíveis aos pacientes, podendo estar localizado até em sua própria residência;
- **Disponibilidade do ambiente remoto:** um mesmo ambiente de treinamento pode estar acessível 24 horas por dia, sendo disponibilizado para diferentes pacientes em diferentes fuso-horários;
- **Customização do ambiente de treinamento:** na presente proposta, a Realidade Aumentada possibilita que profissionais de saúde que acompanham os pacientes insiram objetos virtuais no ambiente de treinamento. Desta forma, de acordo com a necessidade de cada paciente, o profissional de saúde poderá programar exercícios com diferentes níveis de dificuldade e intensidade, para que o paciente evolua, gradativamente, na realização do seu treinamento;
- **Relatórios de treinamento:** o sistema utilizado para o treinamento de cadeirantes estará conectado a um banco de dados, o que permitirá o armazenamento de informações referentes à sessão de treinamento. Dentre as informações de maior importância, pode-se citar, por exemplo, o tempo gasto para realizar determinado exercício ou os índices de sucesso/falha na realização de determinada tarefa. Estes dados serão usados para produzir relatórios de treinamento que serão disponibilizados aos profissionais de saúde que acompanham o paciente;

Segurança: apesar de controlar uma cadeira real, o paciente estará realizando os exercícios em um ambiente seguro, sem estar exposto aos riscos que o ambiente de treinamento oferece.

A Figura 1.7 esboça os componentes necessários para o desenvolvimento desta proposta.

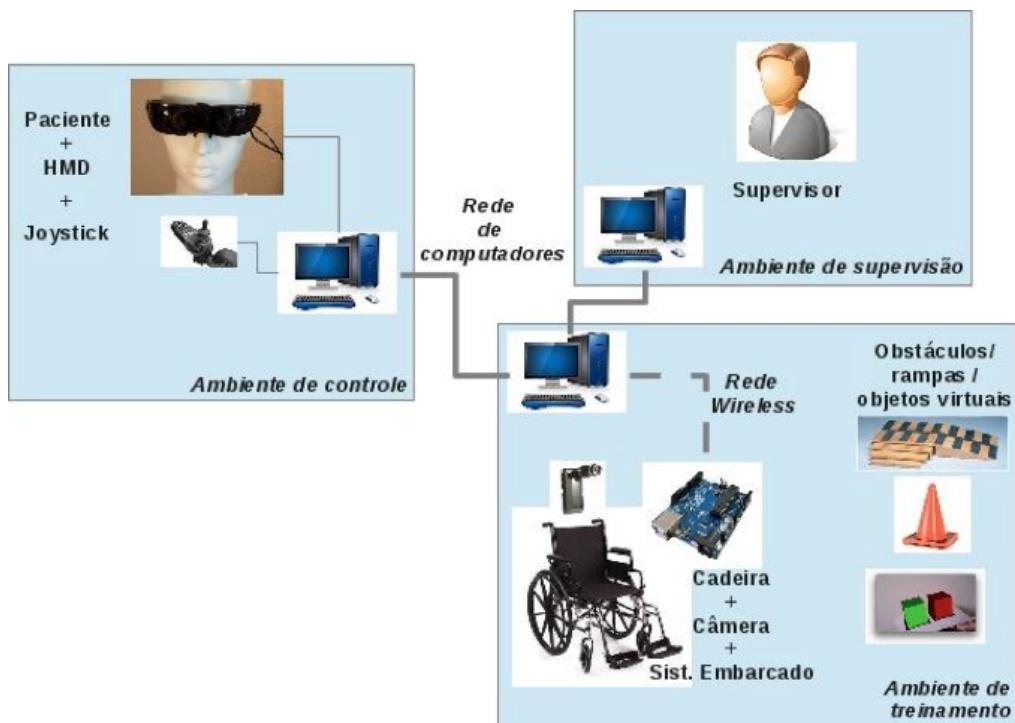


Figura 1.7. – Esboço dos componentes para o *Framework* proposto

Na Figura 1.7 são apresentados os 3 ambientes que compõe a arquitetura proposta:

- Ambiente de treinamento, no qual estão localizados a cadeira (simulada por um veículo robótico), objetos virtuais e os obstáculos físicos como por exemplo uma rampa;
- Ambiente controle, no qual o usuário visualiza, através de um HMD, a sala de treinamento, enriquecida com os obstáculos virtuais;
- O ambiente de supervisão, de onde o terapeuta poderá acompanhar todas as atividades realizadas pelo usuário, além de personalizar as tarefas realizadas pelo mesmo.

A seguir, serão detalhadas as características de cada ambiente.

1.5.1. Ambiente de controle

Neste ambiente, o paciente possuirá um computador ligado à Internet, no qual estarão conectados um HMD e um joystick. O HMD (responsável por aumentar o nível de imersão) proverá ao paciente um feedback visual do ambiente de treinamento onde estarão posicionados os objetos virtuais e reais contidos no mesmo. Por meio do joystick (similar aos utilizado em cadeiras eletrônicas), o paciente enviará ao ambiente de treinamento o sinais responsáveis por controlar a cadeira de roda/veículo.

A aplicação em execução neste ambiente é responsável por estabelecer um canal de conexão com os demais ambientes, permitindo o envio e recebimento de informações.

1.5.2. Ambiente de treinamento

Neste ambiente, estarão posicionados uma cadeira/veículo de comando eletrônico (sem nenhum usuário) à qual estará acoplada uma câmera (responsável por coletar as informações visuais do ambiente), alguns objetos virtuais (representados pelos marcadores) e alguns objetos físicos (obstáculos e rampa), utilizados para reproduzir as dificuldades encontradas pelos cadeirantes em seu cotidiano. Além destes componentes o ambiente possuirá também um desktop conectado a Internet.

O veículo/cadeira de rodas é controlado por um sistema embarcado. Este sistema recebe os sinais de controle recebido no desktop via wireless. O desktop, por sua vez, recebe estes sinais do ambiente de controle.

1.5.3. Ambiente de supervisão

Neste ambiente, o terapeuta ou profissional da área poderá acompanhar todas as atividades realizadas pelo paciente no ambiente de treinamento. O terapeuta poderá também personalizar os exercícios a serem realizados pelo paciente, alterando os objetos virtuais associados aos marcadores. Por exemplo: um cenário, com alguns objetos virtuais, será apresentado ao paciente através do HMD. Em seguida, será solicitado ao usuário que utilize o joystick para capturar certos objetos e contornar outros.

1.5.4. Processo de treinamento

O processo de treinamento em si consiste na execução de um protocolo de treinamento. Neste processo, o terapeuta (supervisor) pode adicionar ou remover os objetos virtuais dentro do ambiente de treinamento, a fim de, avaliar o processo de desenvolvimento do paciente. Este processo é dividido em três segmentos:

- Avaliar a evolução do paciente em um dado intervalo de tempo;
- Avaliar a evolução individual do paciente em um dado intervalo de tempo;
- Avaliar a evolução do paciente em relação a de um grupo;

Por meio do ambiente de treinamento o paciente, irá executar as tarefas de controle do veículo/cadeira de rodas, de modo a evitar os objetos posicionados. Para utilizar estas aproximações em um sistema de telereabilitação, algumas alterações devem ser consideradas. O paciente deverá executar os comandos para controlar veículo/cadeira de rodas remoto, da mesma maneira que seria realizado no ambiente físico. Por esta razão, um canal de comunicação de duplo sentido é necessário, para que o mesmo possa ao mesmo tempo enviar os sinais de controle e receber de volta o feedback visual dos mesmos. No ambiente de treinamento apresentado pela Figura 1.7, o veículo possui um sistema embarcado responsável por estabelecer a comunicação wireless entre o computador veículo.

A utilização de técnicas de RA, permitem que o supervisor possa adicionar/remover objetos no ambiente, conforme as necessidades ou evolução do paciente.

1.5.5 Processo de comunicação

O processo de comunicação é um objeto crítico na arquitetura proposta. Por meio dos canais de comunicação o paciente pode ao mesmo tempo controlar o veículo/cadeira de rodas e ainda visualizar o ambiente de treinamento assim como o supervisor. O caso de uso exibido na Figura 1.8, ilustra estas interações.

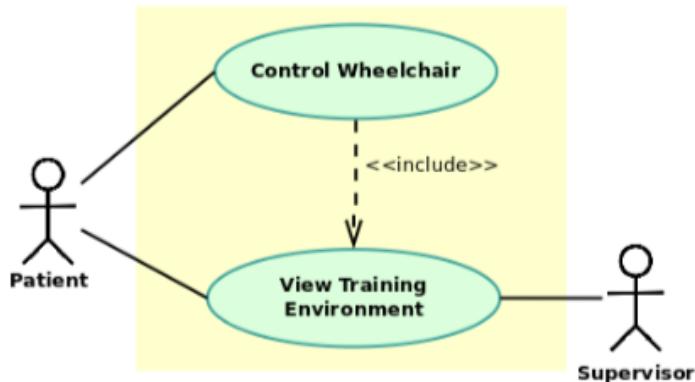


Figura 1.8. – Caso de uso para comunicação.

Sinais de controle de alto nível são enviados para o ambiente de treinamento, a fim de movimentar o veículo, e podem ser divididos em quatro classes de sinal: para frente, para traz, para esquerda e para a direita. O ideal é que todos estes sinais sejam amostrados de forma analógica não conflitantes.

Apesar da simplicidades destes sinais, eles são críticos à aplicação por afetarem diretamente a experiência do usuário no processo. Por esta razão requisitos de tempo real e um protocolo de comunicação confiáveis necessários.

Por outro lado, a visualização do ambiente de treinamento provê ao paciente e ao supervisor um feedback visual de todas atividades executadas neste ambiente. Embora seja desejável, uma alta qualidade de imagem não é obrigatória. Imagens de alta qualidade necessitam de grande capacidade de transferência de dados. Eventuais falhas na transferência de dados podem levar a perdas de pacotes, o que afeta diretamente a experiência do usuário.

1.6. Experimentos realizados e resultados

Alguns experimentos foram realizados a fim de testar as abordagens de comunicação.

O seguinte cenário de teste foi montado, com a seguinte configuração:

- Um robô com quatro rodas, controlado por Arduino¹ foi usado para simular uma cadeira de rodas electrónica;
- Um smartphone, rodando o sistema operacional Android², foi anexado ao robô,

¹ <http://arduino.cc/>

² <http://www.android.com/>

para fornecer o recurso de câmera;

- Um notebook, simulando o ambiente de controle, foi utilizado para emitir comandos de movimento e visualizar vídeos capturados no ambiente de treinamento (pela câmera do smartphone);

O protótipo desenvolvido assim como as configurações do ambiente de treinamento são exibidos na Figura 1.9.

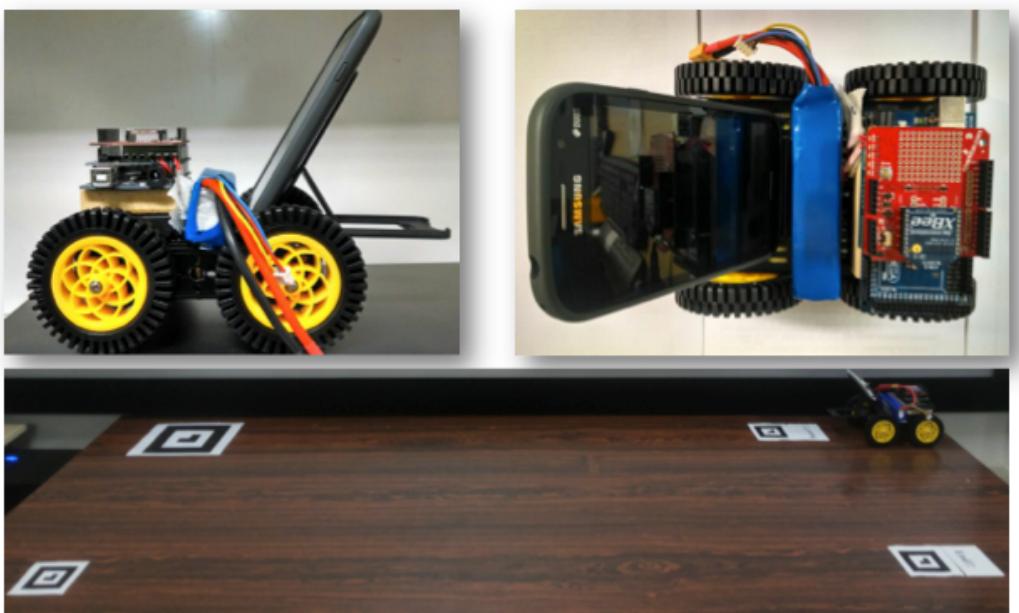


Figura 1.9. – Robô com quatro rodas com smartphone anexado.

1.6.1. Comandos de controle

O primeiro experimento realizado consistiu no envio de comandos de controle (comandos de movimento), através de uma conexão de rádio entre o ambiente de controle (representado pelo notebook) e do ambiente de treinamento (em que o robô foi colocado). Esta primeira experiência foi conduzida usando um Arduino Xbee Shield³, que implementa a especificação ZigBee, com base no padrão IEEE 802.15 [Heile, 2014]. A escolha para este módulo e especificação foi com base na disponibilidade e experiência do grupo de pesquisa com este componente.

1.6.2. Transmissão de vídeo utilizando RTSP

Para experimentar a transmissão de vídeo entre o dispositivo móvel e o notebook foi desenvolvido um aplicativo de teste implementado o protocolo *Real-Time Streaming Protocol* (RTSP). RTSP é um protocolo em nível de aplicação, desenvolvido para fornecer a transmissão em tempo real de dados de mídia, como áudio e vídeo [Schulzrinne, Rao e Lanphier, 1998]. Em comparação com a abordagem socket, a RTSP fornece uma interface de nível superior, encapsulando as responsabilidades de nível inferior tal como o empacotamento de informação e controle de fluxo.

³ <http://arduino.cc/en/Main/ArduinoXbeeShield>

Em primeiro lugar, uma biblioteca de código aberto (LibStreaming⁴) foi utilizada para realizar esses testes. No entanto, alguns problemas de incompatibilidade foram enfrentados durante a integração com a Realidade Aumentada, tais como o acesso exclusivo necessário para a câmera. Com essa restrição, não foi possível utilizar a câmera, simultaneamente, para Realidade Aumentada e transmissão de vídeo. Para superar este problema, uma biblioteca RTSP foi desenvolvida, baseado na implementação apresentada por Kurose e Ross [Kurose and Ross, 2013]. A Figura 1.10 apresenta os principais componentes da biblioteca desenvolvidos.

A classe **RtspServer** é implantada para o dispositivo móvel, fornecendo as funcionalidades de transmissão do servidor. As informações são empacotadas nas instâncias do **RTPpacket**, antes de serem enviadas através da rede. Um componente de interface do usuário - VideoStream - foi desenvolvido para implementar captura de vídeo, usando a câmera do dispositivo móvel. O **RtspClient** é implantado em estações de ambos os ambientes de controle e supervisão. Através destes clientes, terapeutas e pacientes podem visualizar o ambiente de formação, a partir do ponto de vista da cadeira de rodas.

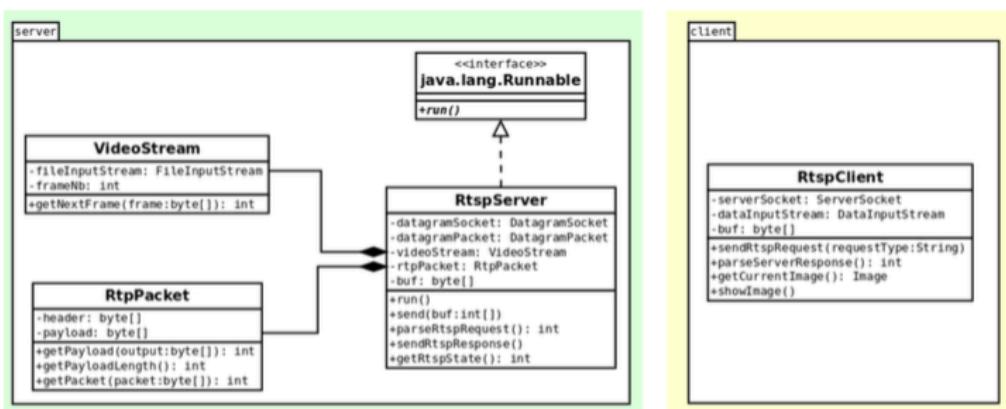


Figura 1.10. – Classe principal RTSP – Cliente / Servidor

1.6.3. Tele reabilitação aumentada

As funcionalidades do uso de técnica de Realidade Aumentada foram investigadas através de duas abordagens diferentes: no primeiro, os objetos virtuais são adicionadas ao vídeo pelo dispositivo móvel. Esta abordagem será referenciada neste trabalho como "Realidade Aumentada Remota". Na segunda abordagem, o vídeo bruto (sem objetos virtuais) é transmitido ao cliente. O cliente, por sua vez, aumenta os objetos virtuais para o vídeo transmitido. Esta abordagem será referenciado neste trabalho como "Realidade Aumentada Local".

1.6.4. Resultados alcançados

Várias classes de testes foram realizados tais como taxa de frames por segundo e utilização de recursos (processador, memória, e rede). Estes foram empregados com

⁴ <https://github.com/fyhertz/libstreaming>

objetivo de avaliar qual das duas abordagens citadas na seção anterior é melhor.

As configurações do dispositivo móvel utilizado nestes testes foram:

- Clock de 830 MHz CPU;
- 290 MB RAM;

Para mensurar a variação de frames por segundo, foi utilizada a abordagem a seguir:

1. Quando a aplicação é inicializada, o tempo de sistema (em nano segundos) são coletados e armazenados em uma variável temporária (*lastSystemTime*);
2. Quando um novo *frame* chega no cliente, a hora atual do sistema é coletada e armazenada em outra variável temporária (*currentSystemTime*).
3. A taxa de *frames* por segundo é calculada usando a Equação 1.1:

$$FPS = \frac{1.10^9}{currentSystemTime - lastSystemTime}$$

4. Este valor é armazenado na memória;
5. A variável *lastSystemTime* é atualizada para *currentSystemTime*;
6. Os passos de 1 a 5 são repetidos até atingir um conjunto de 200 amostras;

A Figura 1.11 permite observar que a taxa de frames por segundo, no ambiente de Realidade Aumentada Remota foi muito inferior a alcançada no Ambiente Local, que foi aproximadamente 30 frames por segundo. Pode-se facilmente compreender o que ocorreu por meio de uma observação na Figura 1.12. Esta Figura demonstra qual foi o nível de uso do processamento em cada dispositivo. Percebe-se que no ambiente de Realidade Aumentada Remoto o nível de uso chegou a média de 80% da sua capacidade máxima enquanto no ambiente de Realidade Aumentada Local, não ultrapassou 2%.

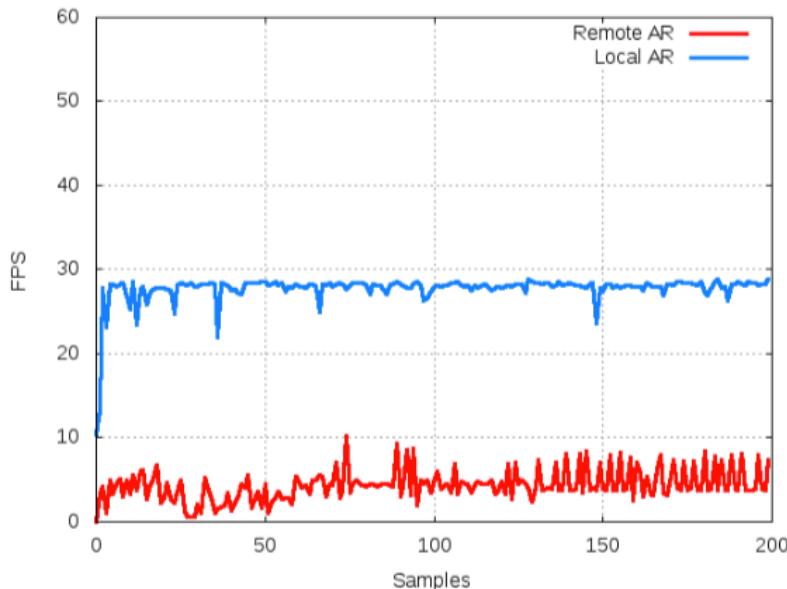


Figura 1.11. – Frames por segundo nos ambientes RA Remota e Local.

Sendo assim, pode-se perceber que a segunda abordagem de Tele reabilitação Aumentada foi a mais indicada neste escopo de teste.

Novos teste poderão ser realizados no futuro, com outros dispositivos móveis com maior capacidade de processamento, além de, incorporar ao framework os demais componentes de interface nos três ambientes melhorando a experiência no uso do framework tanto para o supervisor quanto para o paciente.

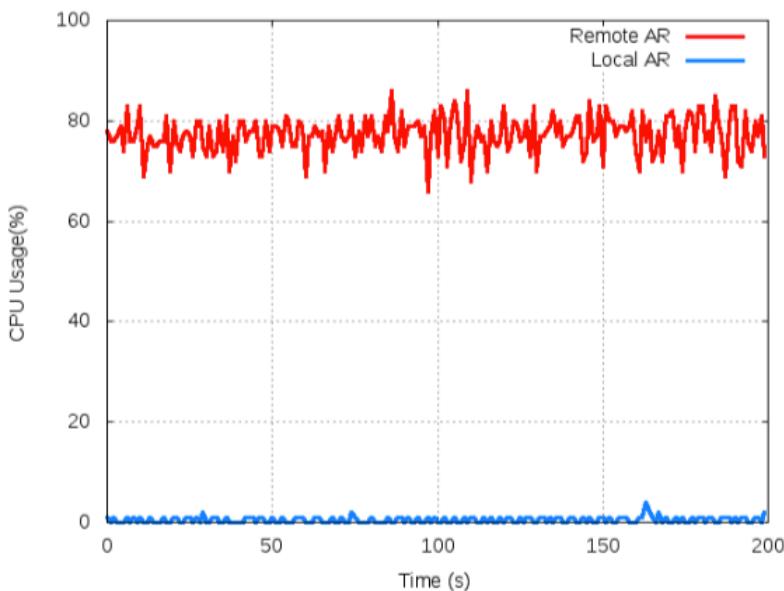


Figura 1.12. – Uso do processador nos ambientes de RA Remota e Local.

1.7. Conclusões e trabalhos futuros

Neste trabalho, é proposto o desenvolvimento de um *framework* para telereabilitação utilizando Realidade Aumentada. O *framework* é baseado em uma arquitetura distribuída, em telereabilitação e é fornecida em três ambientes distintos: um ambiente de treinamento, um ambiente de controle e um ambiente de supervisão.

Dois modelos arquitetônicos foram testados. Em ambos os modelos, o feedback é fornecido através do uso de clientes e servidores RTSP. No primeiro modelo, a Realidade Aumentada foi processada em um site remoto (ambiente de treinamento), através de um dispositivo móvel. No segundo modelo, vídeos puros são transmitidos a partir do local remoto, e Realidade Aumentada foi processada localmente.

Os experimentos realizados indicam que é preferível obter um fluxo de vídeos puros e que a execução do processamento da Realidade Aumentada seja realizada localmente. Os experimentos também tornaram possível a construção de uma plataforma de avaliação, que pode ser avaliada por profissionais de saúde sobre o seu uso como, uma plataforma para telereabilitação. Do ponto de vista técnico, os componentes desenvolvidos são adequados para uso em tais aplicações.

Os trabalhos futuros serão conduzidos de modo a melhorar os componentes do *framework* de acordo com os resultados obtidos. Além disso, a avaliação de outros dispositivos móveis, com mais recursos de processamento, será realizada. Finalmente,

os componentes de interface (para uso em ambientes de controle e supervisão) serão integrados ao trabalho atual.

1.8. Agradecimentos

Esta pesquisa é apoiada pela FAPEMIG (Fundação de Amparo à Pesquisa do Minas Gerais) através de projeto PPM-00511-13 para a qual os autores são profundamente gratos, bem como a CAPES/Brasil, Ministério da Educação e Cultura.

1.9. Referências

- A. Cardoso, C. Kirner, E. L. Júnior, and J. Kelner. Tecnologias e ferramentas para o desenvolvimento de sistemas de realidade virtual e aumentada. Editora Universitária UFPE, pages 1–19, 2007.
- A. Harrison1, G. Derwent, A. Enticknap1, F. Rose, and E. Attree. Application of virtual reality technology to the assessment and training of powered wheelchair users. ICDVRAT 2000, 2000.
- B. Heile, IEEE 802.15 Working Group for WPAN. 2014. Disponível em:
<http://www.ieee802.org/15/>.
- M. K. Holden, T. Dyar, L. Schwamm, and E. Bizzi. Home-based telerehabilitation using a virtual environment system. In Proceedings of the 2nd International workshop on virtual rehabilitation, pages 4–12, 2003.
- M. K. Holden, T. A. Dyar, and L. Dayan-Cimadoro. Telerehabilitation using a virtual environment improves upper extremity function in patients with stroke. Neural Systems and Rehabilitation Engineering, IEEE Transactions on, 15(1):36–42, 2007.
- D. Huang, K. Qian, D.-Y. Fei, W. Jia, X. Chen, and O. Bai. Electroencephalography (eeg)- based brain–computer interface (bci): A 2-d virtual wheelchair control based on event- related desynchronization/synchronization and state control. Neural Systems and Rehabilitation Engineering, IEEE Transactions on, 20(3):379–388, 2012.
- S. Jenkins. Wheelchair training provision by nhs wheelchair services. London: Whizz-Kidz, 2002.
- A.M.Karmarkar,D.M.Collins,T.Wichman,A.Franklin,S.G.Fitzgerald,B.E.Dicianno, P. F. Pasquina, and R. A. Cooper. Prostheses and wheelchair use in veterans with lower-limb amputation. J Rehabil Res Dev, 46(5):567–76, 2009.
- C. Kirner and T. G. Kirner. Virtual reality and augmented reality applied to simulation visualization. Simulation and Modeling: Current Technologies and Applications, 1:391– 419, 2007.
- J. F., Kurose,; K. W., Ross, Computer Networking: A top-down approach. 6. ed. [S.l.]: Pearson, 2013.

- L. H. Lima. Acessibilidade para pessoas portadoras de deficiências: requisito da legalidade, legitimidade e economicidade das edificações públicas. Jus Navigandi, Teresina, ano, 11, 2006.
- Marlus Dias Silva. “Realidade aumentada distribuída na Web com independência de servidor para manipulação de objetos virtuais na camada cliente”. Diss. de mestrado. Uberlândia: Universidade Federal de Uberlândia, abr. de 2011.
- F. Routhier, R. L. Kirby, L. Demers, M. Depa, and K. Thompson. Efficacy and retention of the french-canadian version of the wheelchair skills training program for manual wheelchair users: a randomized controlled trial. Archives of physical medicine and rehabilitation, 93(6):940–948, 2012.
- B. Salatin, I. Rice, E. Teodorski, D. Ding, and R. A. Cooper. A survey of outdoor electric powered wheelchair driving. In Proc 33th RESNA Int Conf. Las Vegas, USA, 2010.
- A. Schwarz and J. HABER. População com deficiência no brasil–fatos e percepções. São Paulo: Coleção Febran de Inclusão Social, 2006.
- N.Sugita,Y.Kojima,M.Yoshizawa,A.Tanaka,M.Abe,N.Homma,K.Seki, and N.Handa. Development of a virtual reality system to evaluate skills needed to drive a cycling wheel- chair. In Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE, pages 6019–6022. IEEE, 2012.
- D. Weidlich, S. Scherer, and M. Wabner. Analyses using vr/ar visualization. Computer Graphics and Applications, IEEE, 28(5):84–86, 2008.