SPWID 2019

The Fifth International Conference on Smart Portable, Wearable, Implantable and Disability-oriented Devices and Systems

ISBN: 978-1-61208-734-4

July 28 – August 2, 2019

Nice, France

SPWID 2019 Editors

Jaime Lloret, Universitat Politecnica de Valencia, Spain

SPWID 2019

Forward

The Fifth International Conference on Smart Portable, Wearable, Implantable and Disability-oriented Devices and Systems (SPWID 2019), held between July 28, 2019 and August 02, 2019 in Nice, France, continued a series of events focused on concepts and communities dealing with specialized implantable, wearable, near-body or mobile devices, including artificial organs, body-driven technologies, and assistive services. Mobile communications played by the proliferation of smartphones and practical aspects of designing such systems and developing specific applications raise particular challenges for a successful acceptance and deployment.

We welcomed academic, research, and industrial contributions, technical papers presenting research and practical results, position papers addressing the pros and cons of specific proposals.

We take here the opportunity to warmly thank all the members of the SPWID 2019 technical program committee, as well as all the reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and effort to contribute to SPWID 2019. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

We also thank the members of the SPWID 2019 organizing committee for their help in handling the logistics and for their work that made this professional meeting a success.

We hope that SPWID 2019 was a successful international forum for the exchange of ideas and results between academia and industry and to promote further progress in the area of smart portable, wearable, implantable and disability-oriented devices and systems. We also hope that Nice, France provided a pleasant environment during the conference and everyone saved some time to enjoy the charm of the city.

SPWID 2019 Chairs

SPWID Steering Committee

Marius Silaghi, Florida Institute of Technology, USA
Jun-Dong Cho, SungKyunKwan University, Korea
Lenka Lhotska, Czech Institute of Informatics, Robotics and Cybernetics | Czech Technical University in Prague, Czech Republic

SPWID Industry/Research Advisory Committee

Warner ten Kate, Philips Research, the Netherlands

SPWID 2019

Committee

SPWID Steering Committee

Marius Silaghi, Florida Institute of Technology, USA
Jun-Dong Cho, SungKyunKwan University, Korea
Lenka Lhotska, Czech Institute of Informatics, Robotics and Cybernetics | Czech Technical University in Prague, Czech Republic

SPWID Industry/Research Advisory Committee

Warner ten Kate, Philips Research, the Netherlands

SPWID 2019 Technical Program Committee

Giovanni Albani, Istituto Auxologico Italiano - IRCCS, Verbania, Italy
Jesús B. Alonso Hernández, Institute for Technological Development and Innovation in Communications (IDeTIC) | University of Las Palmas de Gran Canaria (ULPGC), Spain
Viacheslav Antsiperov, Kotel'nikov Institute of Radio-engineering and Electronics (IRE) of Russian Academy of Sciences (RAS), Russia
Rohan Banerjee, Tata Consultancy Services, India
Roberto Beghi, University of Milan, Italy
Katharina Bredies, University of the Arts Berlin, Germany
Filip Bergquist, Sahlgrenska Academy | University of Gothenburg, Sweden
Juan V. Capella Hernández, Universitat Politècnica de València, Spain
Filippo Cavallo, The BioRobotics Institute | Scuola Superiore Sant'Anna, Pisa, Italy
Amitava Chatterjee, Jadavpur University, Kolkata, India
Claude Chaudet, Webster University Geneva, Switzerland
Biswajit Das, KIIT University, India
Jun-Dong Cho, SungKyunKwan University, Korea
Cesario Di Sarno, COSIRE Group, Aversa, Italy
Jian Du, Carnegie Mellon University, Pittsburgh, USA
Mauro Falcone, Fondazione Ugo Bordoni, Italy
Ramin Fallahzadeh, Stanford University, USA
Biyi Fang, Michigan State University, USA
Gianluigi Ferrari, University of Parma, Italy
Alessia Garofalo, COSIRE Group, Aversa, Italy
Franck Gechter, Univ. Bourgogne Franche-Comté, France
Vivian Genaro Motti, George Mason University, USA
Arfan Ghani, Coventry University, UK
Athanasios Gkelias, Imperial College London, UK
Chris Gniady, University of Arizona, USA
Mario Goldenbaum, Princeton University, USA

Raffaele Gravina, University of Calabria, Italy
Nil Gurel, Georgia Institute of Technology, USA
Jan Havlík, Czech Technical University in Prague, Czech Republic
Joshua Ho, University of Ottawa, Canada
Carmen Horrillo Güemes, Instituto de Tecnologías Físicas y de la Información (ITEFI) | Consejo Superior de Investigaciones Científicas (CSIC), Spain
Gema Ibáñez Sánchez, ITACA - Universitat Politècnica de València, Spain
Ki-Il Kim, Chungnam National University, Republic of Korea
Andrew Kusiak, University of Iowa, USA
Lenka Lhotska, Czech Institute of Informatics, Robotics and Cybernetics | Czech Technical University in Prague, Czech Republic
Tianliang Li, National University of Singapore, Singapore
Jiang Lu, University of Houston - Clear Lake, USA
Jianwei Ma, Harbin Institute of Technology, China
Parbati Kumar Manna, Intel Corporation, USA
Farokh Marvasti, Sharif University of Technology, Iran
Abhinav Mehrotra, University College London, University of Birmingham, UK
Kunal Mitra, Florida Institute of Technology, USA
Hossein Mohamadipanah, University of Wisconsin, USA
Carlos A. Mugruza Vassallo, Universidad Nacional Tecnológica de Lima Sur / Universidad de Lima, Peru / SINAPSE, UK
Sugata Munshi, Jadavpur University, India
Tadashi Nakano, Osaka University, Japan
Brendan O'Flynn, Tyndall National Institute, Ireland
Gregory O'Hare, University College Dublin (UCD), Ireland
Douglas O'Shaughnessy, INRS-EMT, Canada
Lorena Parra, Universitat Politècnica de València, Spain
Alejandro Quintero, Polytechnique de Montreal, Canada
Chandan Karadagur Ananda Reddy, Microsoft Corp., Redmond, USA
Daniel Roggen, University of Sussex, UK
Seyed-Ali Rokni, Washington State University, USA
Ramyar Saeedi, Washington State University, USA
Jacob Scharcanski, UFRGS - Universidade Federal do Rio Grande do Sul, Brazil
Rasoul Shafipour, University of Rochester, USA
Pietro Siciliano, Institute for Microelectronics and Microsystems IMM-CNR, Lecce, Italy
Marius Silaghi, Florida Institute of Technology, USA
Mu-Chun Su, National Central University, Taiwan
Kenji Suzuki, Tokyo Institute of Technology, Japan
Ryszard Tadeusiewicz, AGH University of Science and Technology, Krakow, Poland
Adrian Tarniceriu, PulseOn SA, Switzerland
Warner ten Kate, Philips Research, Netherlands
Vicente Traver, ITACA - Universitat Politècnica de València, Spain
Carlos M. Travieso-González, University of Las Palmas de Gran Canaria, Spain
Alessio Tugnolo, University of Milan, Italy
Timothy Tyndall, RAIN Live Oak Health Information Exchange and Telemedicine Network, USA
Wenwu Wang, University of Surrey, UK
Wang Wen, Institute of Acoustics - Chinese Academy of Sciences, China
Hui Wu, University of New South Wales, Australia

Table of Contents

MyAccessible+ Math: Shining Light on Math Concepts for Visually Impaired Students

Abhishek Jariwala

Department of Computer Science and
Software Engineering
Auburn University
Auburn, USA
Email: avj0003@auburn.edu

Daniela Marghitu

Department of Computer Science and
Software Engineering
Auburn University
Auburn, USA
Email: marghda@auburn.edu

***Abstract*—For some of us, a full understanding of complex mathematical concepts is only achieved through a lifetime of practice. For students with visual impairment, this process is hindered by the inability to process complex mathematical formulae. While computer technologies have successfully transformed and enhanced the learning process, the potential promised by these technologies has not become the reality for visually impaired students. For years, the most effective way to communicate ideas to a blind person has been through either audible conversation or braille writing, both methods having their shortcomings for more complex mathematical analysis. This is the purpose of the MyAccessible+ Math Project; bridging the gap between math professors and students who simply need instructions from a different perspective. The motivation of this study is to introduce a prototype of our web application that will help visually impaired high school students to evaluate their mathematical skills. We hope that the use and further development of our prototype shall open doors to these students in areas of academia and beyond that have until now seemed eternally sealed.**

Keywords*–*Vision-impaired students; Accessibility; Speech recognition.

I. INTRODUCTION

According to World Health Organization [1], approximately 250 million people in the world have moderate to severe visual impairment that cannot be cured. The National Institutes of Health (NIH) study [2] has found that 14 million Americans aged 12 and older, are visually impaired. The inability to process visual elements is an obstacle for many vision-impaired students. This disadvantage leads to a big knowledge gap between students with vision-impairment and students without disabilities.

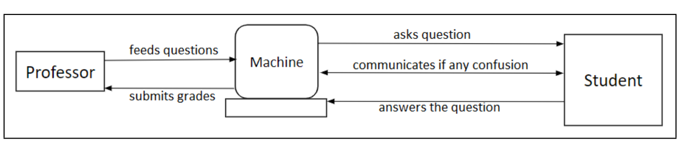

Figure 1. MyAccessible+ Math Design Structure.

The aim of this study is to introduce the prototype of our Web application, MyAccessible+ Math, that helps vision-impaired high school students evaluate their mathematical skills. The basic design structure of the project is shown in Figure 1. Entities associated with the module are explained below:

- Professor: The role of the professor is to add questions and set points for questions. An open-source math visual editor, Wiris MathType [3] is embedded on the website, which makes easier for the professor to add questions for different math topics.
- Machine: The machine represents the ”evaluating module” of the project. This module provides questions to the student. The module also provides help on each question if the student seeks for it. All the communication between the student and this module is done through Annyang.js [4], an open-source speech recognition library. Additionally, the module calculates the overall grade for each student and reports it to the professor.
- Student: The student is the most important entity of this project. The student can use basic navigation commands to iterate through the website. The most useful commands are ’Go to login page’, ’help me’, and ’logout’. Annyang.js [4] library is used for speech recognition. In the future, more commands can be added to extend the scope of the project.

The rest of the paper is organized as follows: *Section II* consists of literature review, which discusses existing techniques that are currently available that facilitate teachers to understand teaching and/or testing challenges and experiences involving vision-impaired students. *Section III* demonstrates the technologies used in the prototype. *Section IV* provides a detailed implementation of the proposed prototype. *Section V* concludes the project along with suggestions for future enhancements.

II. BACKGROUND AND RELATED WORK

For years, the most effective way to communicate ideas to a blind person has been through either tactile methods or audio methods. For example, Braille is a tactile writing system in which characters are physically raised. Audio methods include speech-to-text and text-to-speech representation using speech synthesizer tools.

MathSpeak [5] is an audio method for representing mathematical formulae to blind students. The first step is to scan the mathematical formula and render it in the notation of MathML or LaTeX. The input to the system can be either LaTeX, MathML or AMS-TEX and the system can read equations like text. The system can recognize all alphanumeric characters, all

Greek letters, and other frequently used symbols in the math formulae. One major advantage of the system is to be able to navigate through different sections of the math equation.

MathPlayer [6] is a plug-in for Microsoft Internet Explorer which was designed for rendering visualization of MathML. It can easily be integrated with screen readers.

The LAMBDA project [7] is funded by the European Union to overcome the problem of accessibility to mathematical formulae. The system consists of a MathML markup language that can be directly translated to the 8-point braille system, which is an extension of the 6-point braille system.

Braille code is used to encode mathematical and scientific notation linearly, by using 6-dot Braille cells. While Braille is suitable for the text representation, mathematical equations are multidimensional, and they may contain fraction, algebra, series, log, and exponentiation [5]. Also, 6-dot braille can represent only alphanumeric characters and a small set of special characters by the 64 combinations of possible dot placements. Thus, by extending the 6-dot Braille system to 8-dot Braille system, 64 possible combinations can be extended to 256 combinations.

This is an excellent choice for a certain context, but by today's digital standards, the use of Braille is expensive [8].

III. Technologies Used

For this project, our main goal is to develop a robust and flexible prototype of an open-source Web application in a structured manner and constantly refine it.

The current version of the application supports a limited number of math topics. New math topics, math questions, and math tests can be added in the later versions. The technologies used for the prototype are explained below:

A. *Annyang.js*

Annyang [4] is an open-source JavaScript Speech Recognition library that makes adding voice commands to any website super-easy. The student is the most important entity of the project. Navigation through Web pages is done through voice commands.

Table I shows the list of commands current version of the prototype supports. With Annyang, more custom commands can be added in the future to extend the scope of the project.

TABLE I. VOICE COMMANDS LIST

Voice Commands	Action
'new user'	Redirect to the registration page
'let me in'	Redirect to login page
'practice question'	Redirect to the practice questions list
'attempt random'	Randomly selects the question and open it in a new tab
'skip question'	Skips the question and fetches the next one
'help'	Provides more information on a question while practicing a question
'next hint'	Speaks the next hint
'Log out'	Logs out of the system
'Hint one'	Provides the first hint while practicing a question

B. *Wiris MathType Editor*

Wiris MathType Editor [3] is embedded on Web pages for the professor to add math questions. Math formulae can be exported to multiple formats and are compatible with LaTex and MathML.

C. *ResponsiveVoice.js*

ResponsiveVoice [9] is an open-source text-to-speech library written in JavaScript, offering an easy way of adding voice to any website or application.

D. *Linear Equation Parser*

The current version of the prototype supports two types of math problems: Linear Equation Solver and Linear Equation Simplifier.

Jep Java expression parser [10] is used to evaluate the mathematical expressions in this prototype. Jep Java parses and evaluates the mathematical expressions with only a few lines of code. This package allows users to enter a formula as a string, and instantly evaluates it. Jep supports user-defined variables, constants, and functions.

IV. Implementation

MyAccessible+ Math is a Java-based Web application developed using HTML5 and CSS3 as front-end and JSP and Servlets as back-end. In what follows, we discuss the development process and functional requirements of the prototype.

Initially, all the requirements were gathered and analyzed based on Evolutionary Prototyping (EP). EP allows a continuous refinement of the system and is based on the understanding of the requirements by the developers.

A. *Functional Requirements*

- Home Page (index.jsp): This is a Web page where professors and/or students can login. The machine would recognize if the user is a student or professor based on the username stored in the MySQL database. The home page provides welcome text about the application for vision-impaired students. User can say "new user" to navigate to the registration page. Students will be redirected to dashboard_student.jsp. Professors will be redirected to dashboard_professor.jsp.
- Student Dashboard (dashboard_student.jsp): Only students can access this page. This is a dashboard for students after logging in. Students can use voice navigation commands to go to any page.
- Attempt Questions Page (attempt_questions.jsp): This page provides a list of all practice questions. Student can start practicing random questions by saying "attempt random". Student can say "help" for further info. Figure 5 shows the attempt questions page for students.
- Each Question Page (each_question.jsp): Student lands on this page after choosing a question to attempt. Web page speaks the question on page load. User can say "Repeat please" to listen to the question again or can say "skip question" to skip the question and move to the next question. If needed, the user can ask for the first hint by saying "hint one". Student can ask for more hints by saying "next hint". Figure 6 shows each question page for students where the student can attempt questions one by one.
- Professor Dashboard (dashboard_professor.jsp): Only professors can access this page. Professor can quickly add a question, see the students list, and add a math

topic on this page. Professor can also redirect to those pages for a detailed view. Figure 2 shows the dashboard for the professor. In the future release, the professor would be able to see the overall performance of each student.

- Math Questions Page (math_questions.jsp): Figure 3 shows the math questions page. Only professors can access this page. The professor can add or delete questions for selected math topics.
- Math Topics Page (math_topics.jsp): Only professors can access this page. The professor can add or delete math topics. Once the topic is added, the professor can add questions for that topic.
- Math Quizzes Page (math_quizzes.jsp): Figure 4 shows the math quizzes page. Professor can create a quiz by selecting questions from the list. Since the project is still in a development state, students cannot attempt quizzes yet.

B. Conceptual Design

Following are screenshots of some of the pages that are developed based on functional requirements. In the future, these pages will be modified based on information and feedback from students.

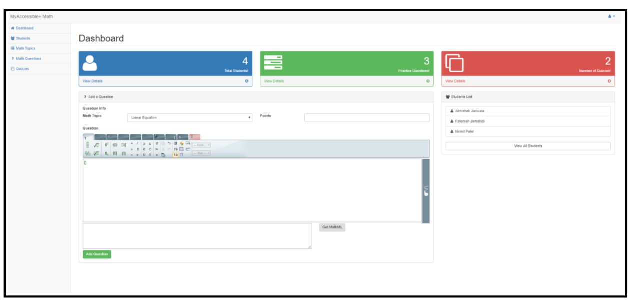

Figure 2. Dashboard professor web page design

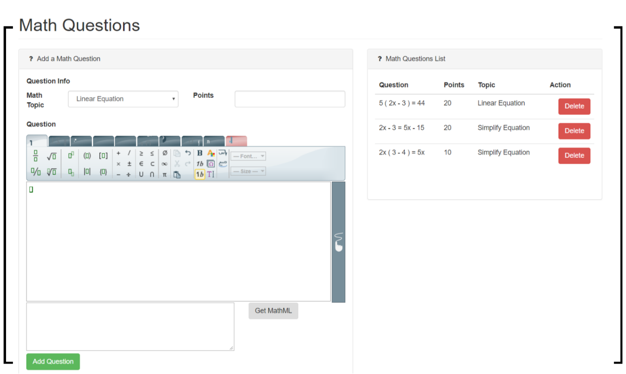

Figure 3. Math questions web page design

1) Professor View: In MyAccessible+ Math, several web pages were designed for the professor to add questions, add math topics, check the students' list, and to create quizzes. The web pages were developed in a way that had good usability, likeliness, and ease of use.

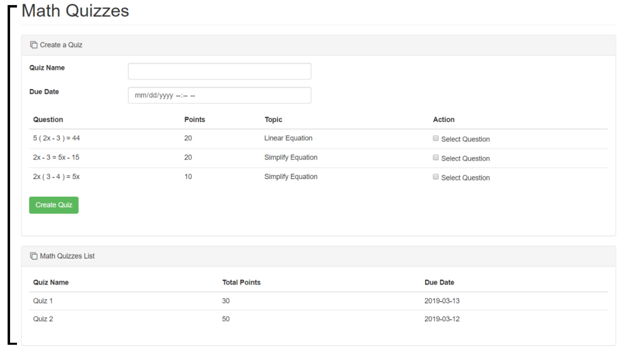

Figure 4. Math quizzes web page design

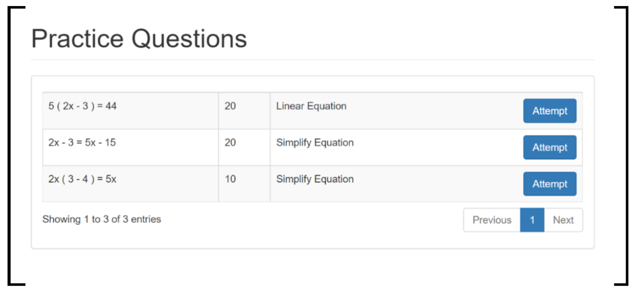

Figure 5. Attempt questions web page design

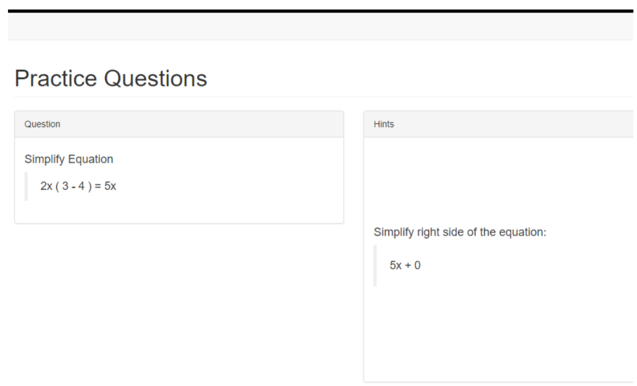

Figure 6. Each question web page design

2) Student View: Students are an important entity in this project. This application was designed and developed for visually impaired students to test their mathematical knowledge. Therefore, all the web pages that student can access have voice navigation enabled.

V. Conclusion and Future Work

Education is one space that still has the potential to be transformed by technology. Though this project is still under development, it has great potential not only for improving the education of students with visual impairments but also for inspiring the next generation of engineers and scientists.

This prototype is developed to improve math education for vision-impaired high school students. The following improvements could be added in the future release of the work:

 ISBN: 978-1-61208-734-4

- Evaluation is important while working on an application for vision-impaired students. We are planning to conduct a study at Auburn University with 15 visually impaired students enrolled in a program to evaluate their knowledge in math. We aim to assess their mathematical skills by evaluating their performance using this application.
- The prototype in this research has shown that it is possible to present mathematical expressions to students with little or no vision and test their knowledge in mathematics. The upcoming version of this prototype will include teaching module of the application where students can learn mathematics.
- The current version of the application includes solving and simplifying linear equations. However, different math topics will be added in the future release of the application.
- Verifying and updating information, resetting password for students will be added in the future release of the application.

REFERENCES

[1] "Blindness and vision impairment," 2018, URL: https://www.who.int/en/news-room/fact-sheets/detail/blindness-and-visual-impairment [accessed: 2019-03-02].

[2] "A study by National Institutes of Health," 2006, URL: https://www.nih.gov/news-events/news-releases/study-finds-most-americans-have-good-vision-14-million-are-visually-impaired [accessed: 2019-03-02].

[3] "MathType demo - For Developers," URL: http://www.wiris.com/editor/demo/en/developers [accessed: 2019-07-15].

[4] "Annyangjs - easily add speech recognition to your site," URL: https://www.talater.com/annyang/ [accessed: 2019-07-15].

[5] A. Nazemi, I. Murray, and N. Mohammadi, "MathSpeak: An Audio Method for Presenting Mathematical Formulae to Blind Students," Jun. 2012, ISBN: 978-0-7695-4894-4, ISSN: 2158-2254.

[6] "MathPlayer," URL: https://www.dessci.com/en/products/mathplayer/ [accessed: 2019-07-15].

[7] "Overview — Lambda Project," URL: http://www.lambdaproject.org/ [accessed: 2019-07-15].

[8] L. Ferres and J. Fuentes-Sepulveda, "Improving accessibility to mathematical formulas: The wikipedia math accessor," vol. 18, Jan. 2011, p. 25.

[9] "ResponsiveVoice.JS Text to Speech," URL: https://responsivevoice.org/ [accessed: 2019-07-15].

[10] "Jep Java - Math Expression Parser," URL: http://www.singularsys.com/jep/ [accessed: 2019-07-15].

 ISBN: 978-1-61208-734-4

Monitoring of Carried Weight During Walk Using a Wearable Pedobarography System

Per Anders Rickard Hellstrom
Mälardalen University
Embedded Sensor Systems for Health (ESS-H)
Västerås, Sweden
e-mail: per.hellstrom@mdh.se

Mia Folke
Mälardalen University
Embedded Sensor Systems for Health (ESS-H)
Västerås, Sweden
e-mail: mia.folke@mdh.se

***Abstract*— Personal health monitoring is advantageous in heavy work environments to reduce the risk of wear and tear and acute injuries. The study of forces between the plantar surface of the foot and a supporting structure, pedobarography, is a promising candidate for monitoring carried weight during walk. The aim of this study was to evaluate the cost effective pedobarography measurement system, IngVaL. Two aspects are evaluated, namely, how well IngVaL can monitor carried weight during walk and if the novel implementation increased the durability. Fifteen test persons made five treadmill walks with a carried weight of 10, 20, 0, 15, and 5 kg. The equipoise analysis method was used. The Root Mean Square Error (RMSE) for estimation of the carried weight was 13.8 kg. A study with the earlier version of the measurement system had a RMSE of 23.3 kg. The earlier system, as well as commercial systems using this kind of sensors, have problems with sensor durability. The new sensor implementation, where the active sensor area boundary was no longer affected by mechanical stress, resulted in no broken sensors. This study shows an increased performance of carried weight estimation compared with earlier work, together with an improved sensor durability.**

Keywords- pedobarography; carried weight; portable; wearable; insole; in-shoe.

I. INTRODUCTION

Pain in the lower back is one of the most common health problems today and is expected to become even more frequent in the future [1]. About a third of all employees in Sweden, during the year 2015, had pain in their lower back every week [2]. Heavy work load and the total amount of lifted weights and lift frequency are moderate to strong risk factors for lower back pain [3]. The year 2015, 16% of the employed men and 10% of the employed women in Sweden lifted more than 15 kg several times a day [2]. The carried weight will vary during the work time.

Monitoring of the conditions in heavy work environments is important to reduce the risk of wear and tear and acute injuries. A wearable system would make it possible to monitor workers that are not stationary.

Pedobarography, the study of forces acting between the plantar surface of the foot and a supporting surface, has been used for weight estimation while standing still [4] and is a promising candidate for estimation of carried weight while walking [5]. IngVaL (Identifying Velocity and Load) [6] is a pedobarography measurement system designed to be a robust and low cost system for monitoring of health related walk parameters. IngVaL is an improved design of an earlier research prototype. The earlier system has been validated for monitoring of walking speed [7] and carried weight [5].

The aim of this study was to evaluate the cost effective pedobarography measurement system, IngVaL. Two aspects are evaluated, namely, (1) how well IngVaL can monitor carried weight during walk and (2) if the novel sensor implementation can make the sensors more durable.

In Section II, the methods for this study are explained. Results are presented in Section III. Section IV is the discussion and the conclusion is in Section V.

II. METHODS

The design of this study was cross-sectional. This section is split into the three sub-sections of hardware, experiment setup and data analysis.

A. Hardware

The IngVaL system consists of sensors (force sensing resistors), electronics for signal conditioning and a microcontroller based data acquisition unit. The data was sent via Bluetooth 4.0 to a Windows tablet.

Four force sensing resistors of model A401 (Tekscan Inc., Boston, MA, USA) are sandwiched in each shoe insole between a base foundation made of Ethylene-Vinyl Acetate (EVA), cork, and leather as upper layers for protecting the sensors and also providing a comfortable, less perspiration inducing, interface with the foot. EVA is a firm but flexible material that is often used in sports equipment and insoles.

The heel, the lateral and medial sides of the metatarsal pad and the big toe pad were chosen as locations for the sensors, see Figure 1. The sensor locations are chosen like this due to the bone structure of the foot. Each sensor has a boundary for the active sensor area and this boundary is sensitive for mechanical stress that can short-circuit and damage the sensor. An earlier prototype of the insoles had problems with sensor durability due to mechanical stress on the boundary of the active sensor area. Besides the higher risk of sensor damage, there were also some peaks in the data from the prototype system due to the short-circuiting of the sensors.

 ISBN: 978-1-61208-734-4

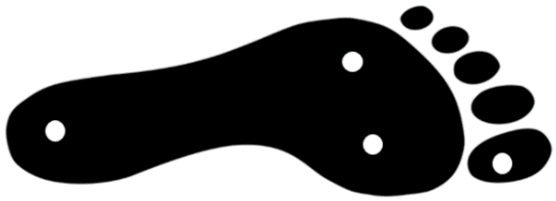

Figure 1. Showing the location of the four insole force sensing resistors located at the big toe, the inner and outer metatarsal pad and the heel.

In version two of the system, called IngVaL, EVA material was removed directly under the sensor's boundaries to prevent the mechanical stress and resulted in a disc structure under each sensor's active sensor area. The block diagram, from force sensing resistors to the data analysis, for the IngVaL system, is shown in Figure 2.

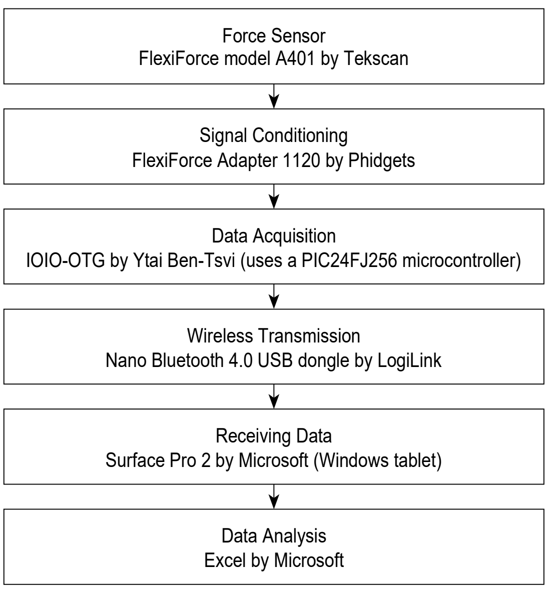

Figure 2. Block diagram for the IngVaL measurement system.

A dynamic calibration was used since the application is to measure during walking conditions [8] and the calibration functions were chosen as fourth order polynomials [9]. A Tedea-Huntleigh 1006 single point load cell (Vishay Precision Group, Malvern, USA) was used for the calibration of the force sensing resistors and the cell was in turn calibrated by using calibrated weights. The calibrating force was applied perpendicular to the active sensor area.

The data sampling (at 200 Hz) and wireless data transmission were made using an IOIO-OTG (SparkFun Electronics Inc., Niwot, CO, USA) which is based on a PIC24FJ256 microcontroller. The name IOIO-OTG comes from naming the first device IOIO, since it has many inputs and outputs, while OTG refers to the Universal Serial Bus (USB) standard On-The-Go (OTG). The IOIO-OTG was connected to the two insoles using elastic cables, secured with Velcro straps, around the ankle and around the lower thigh. A modified version of the java program ioiometer-pc [10] received and saved the data on a Windows tablet.

B. *Experiment Setup*

Fifteen test persons were recruited from the university staff. The test persons had an average weight of 83.9 kg. Inclusion criteria were that they had European Union (EU) shoe size 43-44, were healthy and able to walk naturally when carrying the extra weight. All test persons used the same shoes with the insoles inside, including force sensing resistors. All the test persons performed five walks at a speed of 1.0 m/s on a treadmill (Comfort Track Prime 97690, LifeGear Ltd., Taiwan) after an initial test walk to see that all sensors were activated and that the test person felt comfortable walking on a treadmill.

During each walk, the test persons carried a backpack loaded with a pseudo-randomly chosen extra weight of 10, 20, 0, 15, and 5 kg, see Figure 3. The first author is shown walking on the treadmill with the shoes, including the insoles with the sensors, and with the backpack loaded with extra weight. Extra padding was used inside the backpack along the spine to reduce the risk of injuries.

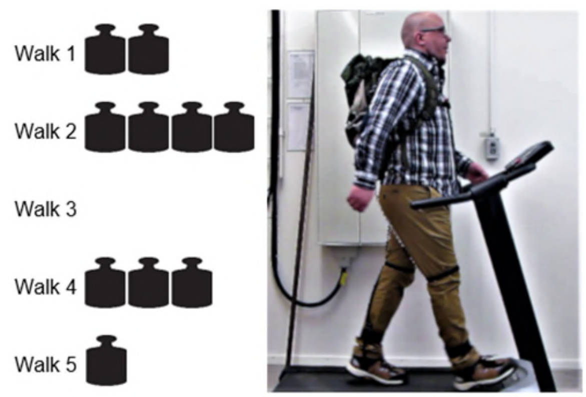

Figure 3. Each test person made five walks with different amount of carried weights (each black weight symbol represents 5 kg).

Data was recorded from the sensors in the pedobarography system during one minute per walk. Acceleration and deceleration phases were not part of the recorded data. The reference weight was measured using a GS 42 BMI electronic floor scale (Beurer GmbH, Ulm, Germany).

C. *Data Analysis*

The estimation of carried weight was made using the equipoise method [5]. Equipoise happens once during each stance phase (between heel strike and toe-off) and is defined as 0.5 when half of the weight is distributed on the heel sensor and the other half on the forward sensors. Examined data was chosen in the equipoise range of 0.5±0.1. Data samples are chosen if no weight was on the other foot at the same time to make sure all weight was on one foot. Microsoft Excel was used to calculate an average of the forces for the equipoise samples.

Figure 4 shows an overview of the three steps of the data analysis: (1) calculate the equipoise ratio, (2) select data in the 0.5±0.1 equipoise range when only one foot is in contact with the ground, and (3) calculate the average force and then calibrate for each individual.

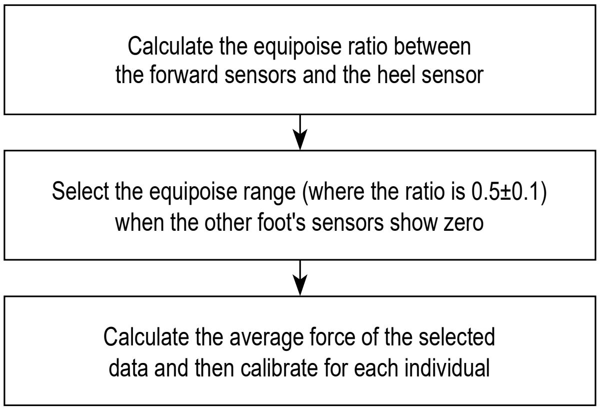

Figure 4. Overview of the three steps of the data analysis.

Two different methods were used for doing the individual calibration. Method 1 used the walk without extra added weight while method 2 also added the maximum carried weight (20 kg). The estimated carried weight was then subtracted with the known weight as measured with a reference floor scale to calculate the error. Finally, the root mean square errors were calculated.

III. RESULTS

The root mean square error was 17.2 kg when method 1 was used and 13.8 kg when method 2 was used, see Figure 5. Dots show errors using the walk without extra carried weight for the individual calibration. Circles show errors when also using the maximum carried weight (20 kg) for individual calibration. There are 15 fewer dots for method 2 since double the amount of data is used for the individual calibration.

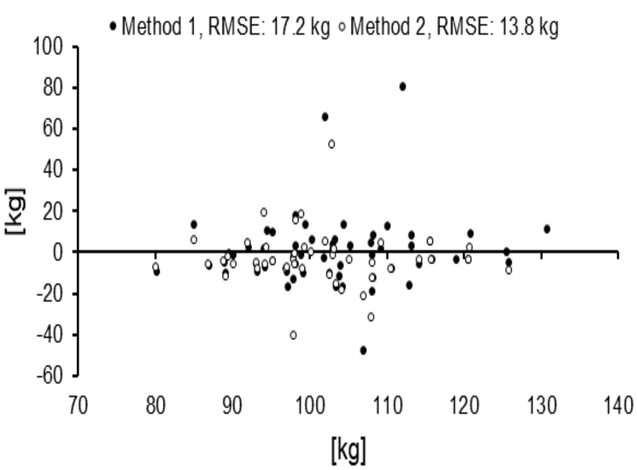

Figure 5. Reference weight (measured with a floor scale) versus weight estimation errors.

None of the sensors broke during more than 350 minutes of walking, as well as taking off and putting on the shoes 15 times.

IV. DISCUSSION

Personal health monitoring using wearable measurement systems is a promising way to be able to monitor health outside of the hospital setting [11][12].

The pedobarography system, IngVaL, is designed to be a robust and low cost measuring system for monitoring of health related walk parameters. The aim of this study was to evaluate the cost effective pedobarography measurement system IngVaL. Two aspects are evaluated, namely, (1) how well IngVaL can monitor carried weight during walk and (2) if the novel sensor implementation can make the sensors durable. The improvements from the first version of the system are mainly in the sensor implementation and sensor calibration. The improvement in the analysing method is how the analysis also uses the maximum carried weight for calibration.

This study used 15 test subjects and five different carried weights for a more thorough experimental examination than in an earlier study where ten test subjects and three different carried weights were used [5]. The same equipoise method was used in this study and resulted in a RMSE of 17.2 kg. A further method improvement resulted in a RMSE of 13.8 kg. This was a good improvement from 23.3 kg in the earlier study (recalculated because Mean Average Error (MAE) was used in that publication). This shows that the new system version (IngVaL) performed better than the previous prototype system. There is, however, room for improvement, and one possibly way to reduce the RMSE is to improve the data analysis. One challenge is to keep the thickness of the insole from becoming larger than a normal insole. The current insole is similar to a normal insole and has a thickness around 5-6 mm.

To the best understanding of the authors, there are no other wearable systems for monitoring carried weight while walking. There is, however, related work where the estimation was made after coming to a standstill after walking and that study presented a RMSE of 10.5 kg using nine test subjects [4]. The need of standing still during measurement makes it unsuitable for monitoring during a workday to see the load of the work over time.

Forces are distributed proportionally over all regions of the foot regardless of foot arch type [13]. This enables the use of fewer sensors instead of a more expensive sensor matrix. It is important to design a durable system for monitoring of heavy work environments. IngVaL used a new way of implementing the sensors into the insoles and this made them more durable. This resulted in no broken sensors during more than 350 minutes of use. Four sensors broke during 80 minutes, when using the earlier version of the system. Sensor replacement would also mean that a new calibration of the sensor is needed and this is a concern if the system is to be commercialized in the future. The durability issue made the earlier prototype system unsuitable.

A potential limitation in this study could have been the use of a treadmill, which might result in a less natural walking style compared to on a flat floor. On the other hand, the equipoise is measured when one foot has equal pressure on the forefoot sensors and the heel sensor while the other

 ISBN: 978-1-61208-734-4

foot is in the air. This part of the stance phase (when the foot is in contact with the ground) is expected to be minimally, if at all, affected by walking style because the foot is not in direct motion during this particular moment. The treadmill is instead an advantage since it allows a constant walking speed of 1.0 m/s. In order to avoid the influence of different types of shoes, all of the 15 test subjects in the study used the same shoes during the measurements. They also used the same insole in the shoes.

The IngVaL system has earlier shown to be able to measure walking speed. Together, monitoring of the carried weight and the walking speed, enable estimation of energy expenditure [14][15].

V. CONCLUSION

In this study, the cost effective pedobarography measurement system called IngVaL has been evaluated considering two aspects, namely, (1) how well IngVaL can monitor carried weight during walk, and (2) if the novel sensor implementation can make the sensors durable. This study shows that the root mean square error has been decreased from 23.3 kg to 13.8 kg and validates that the new measurement system version (IngVaL) performs better than the previous system regarding monitoring of carried weight during walk. The new implementation of the sensors has made them more durable and resulted in no broken force sensing resistors during the experiment.

ACKNOWLEDGMENT

This study was supported by the Swedish Knowledge Foundation, KKS.

REFERENCES

[1] D. Hoy et al., "A systematic review of the global prevalence of low back pain," Arthritis Rheum., vol. 64, pp. 2028-2037, 2012.

[2] The Swedish Work Environment Authority (Arbetsmiljöverket), Arbetsmiljöstatistik 2016:2, 2016.

[3] H. Heneweer, F. Staes, G. Aufdemkampe, M. van Rijn, and L. Vanhees, "Physical activity and low back pain: a systematic review of recent literature," Eur. Spine J., vol. 20, pp. 826-845, 2011.

[4] N. A. Sazonova, R. Browning, and E. S. Sazonov, "Prediction of bodyweight and energy expenditure using point pressure and foot acceleration measurements," Open Biomed. Eng. J., vol. 5, pp. 110-115, 2011.

[5] P. Hellstrom, M. Folke, and M. Ekström, "Wearable Weight Estimation System," Procedia Com. Sci., vol. 64, pp. 146-152, 2015.

[6] P. A. R. Hellstrom, A. Åkerberg, M. Ekström, and M. Folke, "Evaluation of the IngVaL Pedobarography System for Monitoring of Walking Speed," J. Healthcare Informatics Research, vol. 24, pp. 118-124, 2018.

[7] P. A. Hellstrom, A. Akerberg, M. Ekstrom, and M. Folke, "Walking Intensity Estimation with a Portable Pedobarography System," Stud. Health Technol. Inform., vol. 224, pp. 27-32, 2016.

[8] J. Florez and A. Velasquez, "Calibration of force sensing resistors (fsr) for static and dynamic applications," 2010 IEEE ANDESCON, Year, pp. 1-6.

[9] J. M. Brimacombe, D. R. Wilson, A. J. Hodgson, K. C. Ho, and C. Anglin, "Effect of calibration method on Tekscan sensor accuracy," J. Biomech Eng., vol. 131, pp. 034503, 2009.

[10] Johannes Rieke, ioiometer-pc (for Windows), in https://github.com/jrieke/ioiometer-pc, 2016.

[11] B. H. Dobkin and A. Dorsch, "The promise of mHealth: daily activity monitoring and outcome assessments by wearable sensors," Neurorehabil. Neural Repair, 2011, pp. 788-798.

[12] K. Hurt, R. J. Walker, J. A. Campbell, and L. E. Egede, "mHealth interventions in low and middle-income countries: a systematic review," Glob. J. Health Sci., vol. 8, pp. 183, 2016.

[13] S. L. Goffar et al., "Changes in dynamic plantar pressure during loaded gait," Phys. Ther., vol. 93, pp. 1175-1184, 2013.

[14] G. J. Bastien, P. A. Willems, B. Schepens, and N. C. Heglund, "Effect of load and speed on the energetic cost of human walking," Eur. J. Appl. Physiol., vol. 94, pp. 76-83, 2005.

[15] J. M. Falola, N. Delpech, and J. Brisswalter, "Optimization Characteristics of Walking with and without a Load on the Trunk of the Body," Percept. Mot. Skills, vol. 91, pp. 261-272, 2000.

Edge Machine Learning for Energy Efficiency of Resource Constrained IoT Devices

Dawit Mengistu
Department of Computer Science
Kristianstad University
Kristianstad, Sweden
Email: dawit.mengistu@hkr.se

Fredrik Frisk
Department of Computer Science
Kristianstad University
Kristianstad, Sweden
Email: fredrik.frisk@hkr.se

***Abstract*— The recent shift in machine learning towards the edge offers a new opportunity to realize intelligent applications on resource constrained Internet of Things (IoT) hardware. This paper presents a pre-trained Recurrent Neural Network (RNN) model optimized for an IoT device running on 8-bit microcontrollers. The device is used for data acquisition in a research on the impact of prolonged sedentary work on health. Our prediction model facilitates smart data transfer operations to reduce the energy consumption of the device. Application specific optimizations were applied to deploy and execute the pre-trained model on a device which has only 8 KB RAM size. Experiments show that the resulting edge intelligence can reduce the communication cost significantly, achieving substantial savings in the energy used by the IoT device.**

Keywords- Edge intelligence; IoT; Smart Sensors; RNN.

I. INTRODUCTION

Several IoT applications have emerged in healthcare with advances in wearable electronics [6], [11]. Miniaturized devices having sensing, computing and communication capabilities transformed the healthcare sector, enabling the realization of new services. Wearable devices can collect data for monitoring patients remotely to get insights on symptoms or trends, and provide better treatment.

Typical healthcare IoT services use wearable devices in combination with smartphones. Physical and physiological data collected by the wearable sensors is sent to the smartphones where it is aggregated and transferred to backend applications for further processing. The backend often consists of a number of Cloud services for data storage, analytics and machine learning needed to provide actionable information to physicians and patients [10].

The motivation for this work comes from an ongoing research aimed at mitigating Musculo-Skeletal Disorders (MSD) problems in sedentary work environment. In an effort to establish a large dataset for this research, participating subjects were identified for collecting posture data. The data collection is performed continuously for several hours a day, over a long period. In order to build a comprehensive dataset, out of work activities requiring sedentary postures (such as driving) will also be included in the data collection [4].

Energy efficiency is a major challenge in the adoption of wearable IoT for such studies because most devices used in these applications are energy constrained, often running on low capacity power sources. In our case, multiple, coin-cell battery operated wireless devices equipped with inertial sensors are worn by the subjects. Communication between the devices and the smartphone takes place via a Bluetooth Low Energy (BLE) interface. However, the batteries of the devices last few hours only because of the volume of data they transfer. For example, a wearable motion sensor with 9 channels reading 50 samples per second generates over 100MB of data per day.

In this paper, we shall present an approach to improve energy efficiency of wearable sensors through Edge Machine Learning techniques. Our goal is to reduce the volume of communication between the sensor devices and the smartphones to the minimum needed. The machine learning implementation shall recognize eventful data and transfer it to the smartphone only when it is necessary. Implementing machine learning algorithms on resource constrained devices is often not practical because the algorithms require adequate computing power and large storage memory, both of which are not available on most wearable devices. However, with the emergence of edge computing, it has been possible to handle most of the computational and storage burden of machine learning far away from the source of the data.

We investigated different machine learning algorithms to identify the ones that suit our task. Our findings show that RNNs can be implemented on a resource-constrained edge device and give the desired accuracy in real-time. As a proof-of-concept, we evaluated the execution performance and accuracy of a pre-trained RNN model on an Atmega640 microcontroller. The Atmega640 is an 8-bit microcontroller with 16MHZ clock, 64KB boot (code) memory and 8KB data memory (static RAM). This microcontroller has lower specifications (processor speed and memory) than typical devices used for such applications. It can therefore be said that the results of our experiment can be applied to devices already adopted by the wearables industry.

A Python based Machine Learning library was used to build and pre-train our RNN model. Experiments were run to determine the optimal set of model parameters that fit in the device without sacrificing accuracy significantly.

Posture data collected for the research is used to train the model. We then developed a program in C to implement the pre-trained RNN and deployed it on the sensor device for evaluation. The model's real-time performance on the edge device is found to be satisfactory for posture monitoring in sedentary work environment.

The rest of this paper is organized as follows. Section II gives a brief background on MSD research and the state-of-art in the area. Section III discusses Edge Machine Learning, its challenges and contemporary research in the field and the

 ISBN: 978-1-61208-734-4

approach proposed for the task at hand. Section IV describes the experimental setup for this study. Section V discusses the results of the experiments followed by analysis of the results. Section VI discusses related work in Edge Machine Learning research. We conclude this paper by highlighting important results and citing directions of future work.

II. MSD RESEARCH AND APPLICATION OF IOT

A. A Brief Background to the Research Problem

Sedentary work environment was recognized as one of the major causes of MSD. Studies on the issue found that prolonged seating and poor body postures can reduce blood flow in the cervical region and cause inflammation of muscles and tissues [1]. Poor postures and work positions result in back, neck and shoulder pain. MSD problems may also lead to chronic ailment and even complete immobility. Studies explain the need to maintain the right body posture in work and different daily-life situations, owing to the fact that bad posture and prolonged sitting in one position cause back, neck and shoulder pains that can get worse and develop to chronic diseases with age [3]. Studies have shown the link between sedentary life and the risk of obesity, diabetes, cardiovascular disease, and all-cause mortality.

Increasing healthcare costs, absence from work and the associated negative psychosocial complications are some of the problems caused by MSD with severe impact at societal level. According to the UK National Health Services, the country lost 31 million days in 2014 alone, due to sickness related to back, neck and muscle pain [4]. Recognizing the severity of MSD problems, the National Institute for Occupational Safety and Health in the United States (NIOSH) identified the problem as an important research agenda [8]. NIOSH identified several research directions, among which mechanisms for reducing the impact of MSD is a priority area.

B. IoT in MSD Research

The dataset created in earlier MSD researches were either incomplete or inaccurate due to the method of data collection they employed. In many cases, the data gathering process was based on physical observation or self-reported information [2]. Later researches made use of video recording and tagging [7]. With advances in sensor and communication technologies, it was possible to set up body sensor networks (BSN) that connect multiple devices worn on the subject's body to detect and label movements and postures [5]. Different types of sensors are used today to collect physical and physiological data to capture information on movements, postures, spinal loads, sit-stand frequency, metabolic processes, etc. [1].

The emergence of IoT transformed the field of healthcare by facilitating real-time data collection, monitoring and analysis with greater convenience and ease of use. Cables that were once used to connect wearables to a central data acquisition unit are now replaced with a wireless interface, such as BLE integrated into the devices.

A schematic depicting the setup of an IoT environment for acquisition and storage of data is shown in Figure 1.

The devices needed for this specific study are placed on the center of the upper back area and on the upper part of the left leg. This placement is sufficient to identify sedentary postures and detect whether the subject is sitting or standing. The devices have inertial sensors, (accelerometer, magneto-meter and gyro) and a BLE unit for communication with the smartphone.

The data gathered is stored in a Microsoft Azure Cloud storage as a time series consisting of 9 features (along x, y, z axes for each inertial sensor unit) per sensor node or device. Every row of data is timestamped and contains posture labeling as well.

III. MACHINE LEARNING AT THE EDGE

Edge devices used for sensors are often resource constrained and therefore not capable of running classical machine learning applications on their data. For such devices, both training the model and inference are often carried out far from the origin of the data, in the Cloud. The main innovation in Edge Machine Learning (Edge ML) is that inference can take place on the edge device itself, at the source of the data. Edge ML has several benefits, such as reduction of communication latency, improving energy efficiency, security, personalization and customization of services [9]. Achieving energy requires reducing the energy cost of communication between the sensor edge and the smartphone, which in turn requires reducing the flow of redundant data from the edge device (sensor).

There are important steps that should be investigated to gain from the mentioned benefits of Edge ML. First, identifying the right machine learning algorithm for the dataset at hand requires expertise and skills. Selecting the right algorithms often comes with the dilemma to choose between performance and accuracy.

Second, the capability of the target device to support the algorithm is not a trivial problem. In order to address this issue, several models have to be tested on the target device until a satisfactory one is found. One is often forced to sacrifice the accuracy if the target hardware does not have sufficient processing power or memory.

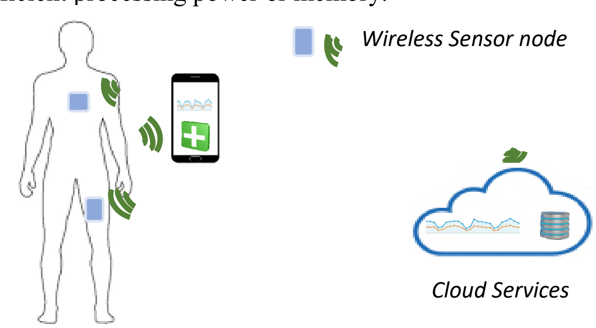

Figure 1. Placement of inertial measurement sensors

Finally, the availability of development tools for machine learning is also a major challenge. Embedded software for many low-end edge devices are written in C. Development environments of some microcontrollers are proprietary with limited flexibility, making importing available libraries very difficult.

Machine learning often requires complex software packages and libraries that can only be executed on powerful processors. Until recently, deep learning was outside the realm of low-end processors on which many IoT devices are based. Because edge devices are resource constrained, they require customized implementations for most machine learning algorithms. The following techniques can be applied in special cases to realize edge ML:

- Offloading the computational work to more powerful devices, for example by performing the training and validation phases on the Cloud;
- Reducing the precision of model parameters and approximating computations with more efficient arithmetic operations wherever possible [20];
- Using lookup tables for activation functions instead of run-time computations.

A. Deep Learning for Temporal Data

One limitation of classical sensor data analysis is the need for manual feature engineering work. Most supervised-learning algorithms are not computationally efficient for deployment on resource constrained devices. Algorithms, such as K-Nearest-Neighbour (KNN) have large storage requirements that can only be met by desktop computers or servers. Furthermore, these algorithms are not suited to detect patterns or contexts hidden in the temporal data collected by the sensors.

RNNs are effective for data with temporal or contextual sequence, such as natural language processing and time series prediction [16]. Their ability to read variable-length sequences of input samples and merge the prediction for each sample into a single prediction for the entire window makes RNN suitable for the posture monitoring application under consideration. A generalized schematic of the RNN architecture is shown in Figure 2. Current hidden states are generated using the input and the previous hidden state. This cyclic behavior in the hidden layer gives the network the ability to learn temporal sequences.

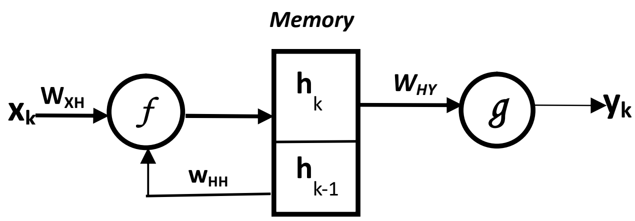

Figure 2. Representation of a recurrent neural network

The mathematical model of the network is represented with the following equations:

$$\mathbf{h}_k = f(\mathbf{W}_{XH}\mathbf{x}_k + \mathbf{W}_{HH}\mathbf{h}_{k-1} + \mathbf{b}_H) \quad (1)$$

$$\mathbf{y}_k = g(\mathbf{W}_{HY}\mathbf{h}_k + \mathbf{b}_y) \quad (2)$$

where

k represents time sequence;

$\mathbf{x}$, $\mathbf{y}$, and $\mathbf{h}$ are the input output and hidden state vectors respectively;

$\mathbf{W}_{XH}$, $\mathbf{W}_{HH}$ and $\mathbf{W}_{HY}$ are the input-to-hidden, hidden-to-hidden and hidden-to-output weight matrices respectively;

$\mathbf{b_H}$ and $\mathbf{b_y}$ are the bias vectors for the hidden and output states respectively (not shown in the figure);

f and g are non-linear activation functions.

RNN models can often be inaccurate and unstable for long input sequences and time series, due to the exploding and vanishing gradient problems [17]. The Long Short-Term Memory (LSTM) variant of RNN was proposed as a solution to overcome the problem. LSTM has achieved impressive results with sequential and time series data in applications, such as text generation, sequence prediction and anomaly detection [16][21].

B. Realization of RNN on Constrained Edge Devices

Deploying a deep learning model on resource constrained edge device, such as an 8-bit microcontroller requires significant optimizations. We shall explore where these optimizations can be applied for our specific use case. Since the number of input features and outputs is already decided by the application, one has to identify other areas to look into. Several models have to be built and tested to arrive at an acceptable one.

One optimization measure is to determine the number of neurons in the hidden layer because the computational complexity and memory requirements for neural network grow exponentially with it. Models of different sizes should be evaluated experimentally for acceptable accuracy and matching the edge device's resource capabilities. It is also possible to achieve a lower count on model parameters by pruning edges with negligibly small weights [12] [19].

Further optimization is achieved through input data reduction. The sampling rates of the sensors are often too high for the microcontroller to make inference from the acquired data within a sampling interval. Applying computationally inexpensive low pass filters helps to reduce the volume of data, to improve the quality of the data and get sufficient time interval for making inference.

Another optimization opportunity is simplifying the computation of activation functions. If the microcontroller does not have built-in floating point capabilities, evaluation of functions, such as ***sigmoid*** and ***tanh*** is expensive. Sacrificing the accuracy of these functions to a reasonable level reduces the time needed for inference significantly.

IV. EXPERIMENTS

We planned two specific tasks in this experiment. The first is to realize an optimized implementation of the algorithm and the model parameters that can fit into the available memory of the target microcontroller. Several models are built to evaluate the tradeoff between accuracy and model size. The second task is to evaluate the performance tradeoff between inference accuracy and the achievable saving in energy on the edge device.

A. Data Collection

A smartphone app is also developed for this experiment, to take care of the data received from sensors, as shown in Figure 1. The main functions of the app are labelling, time stamping, aggregating and uploading data to Cloud storage.

A sampling frequency of 50HZ is used as the base rate. We decided not to increase the sampling rate beyond this as the sensor node's microcontroller would not be able to make predictions in real time if the rate is increased. The sensors measurements are different, owing to the local movement of the body area they are placed on.

The data used in this experiment was collected over a period of 30 minutes with the subject assuming different predetermined postures alternately. This data is used to train a machine learning model which should detect the temporal instances of posture transitions. When training and testing is completed, the model can be deployed on the target.

One of the investigators supervised the subjects to wear the sensors at the correct position and guided them to change postures every 2 to 3 minutes. During a transition, i.e., when the subject changes her posture, a label for the new posture is entered via the app's user interface so that all data received from this time on will automatically have the new label until the next posture change occurs.

The postures assumed are labelled as follows:

- *Sit -upright*
- *Sit -lean left*
- *Sit -lean right*
- *Sit -lean forward*
- *Sit -lean backward*
- *Stand*

B. Training the Model

We implemented our machine learning model using the Tensorflow deep learning framework in Python. The model has 9 feature inputs, one hidden layer (LSTM) and one scalar output, for each sensor node. Since we are interested in posture transition only, a binary classifier is sufficient to detect local posture changes. There are 90000 records in the dataset, split into train (90%) and test (10%). A 70:30 train-to-test ratio was also used later for comparison. The python code was executed for different number of neurons in the hidden layer. The model size depends on the number of input features and the number of neurons in the hidden layer. The experiment shows that the number of parameters can be obtained from the equation:

$$P = (4n+1)(n+1) + 4nk \qquad (3)$$

where

P = number of model parameters (weights)
k = number of input features
n = number of neurons in the hidden layer

The storage requirement for the model is 4P bytes with each parameter represented as a 4-byte floating point value.

Another parameter of interest in the experiment is the duration of the time lag window used by the LSTM layer for prediction. Larger window width is not practical for resource constrained devices as the model prediction time becomes unacceptably long. Smaller width on the other side compromises the accuracy of the prediction. The analytical computation of optimum window size is complex because it depends on several factors that cannot be easily quantified. We therefore determined this value empirically.

C. Deploying the Model on Target Device

The microcontroller version of the RNN code was written in C. The compiled version of the code takes 34KB of flash memory. The model parameters were combined with the source code and compiled. However, they were stored in the SRAM.

Modifications were made in the data acquisition part to include a low pass moving average filter and use a sampling rate of 50HZ. The filter serves the purpose of reducing the volume of data processed by the model in addition to stabilizing the data (against noise). The time taken to execute the model's inference task for different number of neurons in the hidden layer is evaluated to determine whether inference can be achieved in real time. Similar experiments are also run for different window sizes.

The optimum number of neurons n depends on the amount of SRAM available and the number of input features. Having fewer neurons is not desired as it would compromise the accuracy of the model. The results are summarized in the next section.

V. RESULTS AND DISCUSSION

A. Evaluation of Model Training

Different hidden layer sizes were tested in the experiment. However, owing to the limitations in the target device, it is not practical to deploy large models. For example, the number of parameters for a model containing 50 neurons in the hidden layer is about 11,851. This requires about 44KB of memory on the target. The Atmel640 microcontroller has only 8KB static RAM that can be used for all temporary data. After accounting for the memory required to store a few seconds of sensor data, working memory and stack for intermediate computations, the available memory for the model parameters is just under 3.5KB. Applying equation (3), we can train a maximum of 11 neurons in the hidden layer.

Limiting the model size has also the additional benefit of eliminating the risk of overfitting. Our evaluation also shows that larger models are not suitable for the data. We got satisfactory performance with models having as few as 8 neurons. Figures 3 and 4 demonstrate this by comparing the Mean Square Error (MSE) losses for 50 and 8 neurons respectively in the hidden layer. These results show that the smaller model (8 neurons) has in fact a better accuracy because it is a close match to the number of input features.

B. Evaluation of Model Execution on the Target

The performance of the inference model was evaluated on the microcontroller for different sizes of time lag window. It is clear that the inference accuracy improves if a larger time window is used. However, this incurs a large computation cost. For a rate of 50 samples per second, the 20 milliseconds interval is very short to perform data acquisition and inference (prediction). As can be seen in Figure 6, it takes about 35 milliseconds to execute the inference step alone for a window size of 3 seconds.

It is necessary to find an acceptable tradeoff between inference accuracy and real-time response. We applied a low pass filter to stabilize the data by averaging 4 samples at a time, instead of feeding the entire sensor data stream to the ML model. The overhead of this filter is low compared to the inference operations. To our surprise, we got impressive accuracy even when a smaller window size is used. As can be seen in Figure 5, the difference in accuracy between a window size of 2 seconds and 3 seconds is not significant (both have over 95% accuracy). However, the 2 seconds window takes much lower time for inference.

It is possible to deploy a larger model in the flash memory since the SRAM would not be enough. Although the flash memory has a slower access time than that of SRAM, its performance is still acceptable.

C. Analysis of Energy Efficiency

Larger window sizes do not improve the model accuracy significantly. In fact, the improvement, if any, is outweighed by the overhead of the inference step. Because the inference overhead grows exponentially with the size of the window, the solution is not feasible for higher sampling rates. This can be seen in Figure 6.

The energy saving is calculated as the difference between the reduction in data transfer costs and the extra computation incurred by the inference step to achieve this reduction.

It follows that this saving is significant if the frequency of posture changes is low as is the case with sedentary work. The excess computation depends on the sensor data rate and the window size. In our experiment, a sampling rate of 50HZ and averaging every 4 samples is used. This gives the model an interval of 80 milliseconds per inference.

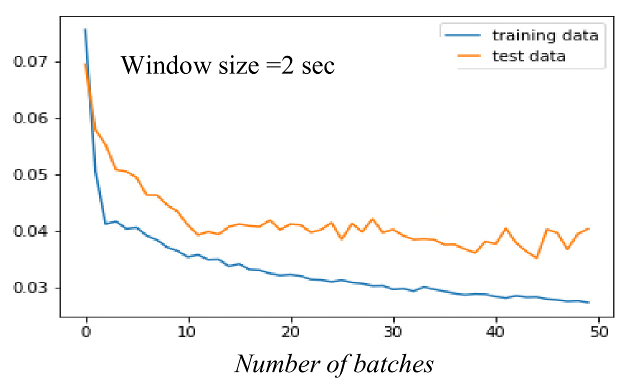

Figure 3. MSE loss (%) for number of neurons=50

As can be seen in Figure 6, if a 2 seconds window is applied, it takes about 9 milliseconds to execute the model (inference code). This achieves the desired result with about 11% increase in computation.

The BLE interface draws an average current of 8.53mA over its connection interval of 2.675 milliseconds with an empty payload [15]. According to the datasheet, the device draws a current of 17.5mA for a full payload data transfer. Android phones support a maximum of 4 packets with a-payload of 20 bytes for a minimum connection interval of 7.5 milliseconds [13].

To get an estimate of the energy saving, we can consider a case where posture changes occur every 10 seconds on average. If edge intelligence were not applied, 4.5KB of data would have to be transferred in the 10 seconds interval (from 9 channels at 50 samples per second and 1 byte per sample). With the above BLE throughput information, a data transfer period of 420 milliseconds over the 10 second duration at an average current of 17.5mA.

With the proposed approach, however, it would be enough to transfer only 225 bytes, the data for one time window only. This achieves a 95% reduction in data transfer costs. The energy consumed for the additional computational overhead is quite modest with the microcontroller drawing less than 2mA, about 10% increase.

VI. RELATED WORK

The emergence of low cost, yet powerful devices has brought machine learning to the edge. Encouraged by this development, researchers in the field have managed to achieve interesting outcomes in edge intelligence in the last few years. In most of the studies we found, the investigators tested their learning algorithms on powerful devices that are not suitable for wearable sensors.

Yazici et. al. investigated porting three different machine learning algorithms to Raspberry Pi running an embedded version of the Android OS [9]. They evaluated the performance of the algorithms for speed, accuracy and power consumption. However, their solution is realized only on powerful edge devices, not on resource-constrained 8-bit microcontrollers.

Gupta et. al. developed a KNN implementation that can run on 8-bit devices such as Arduino [18].

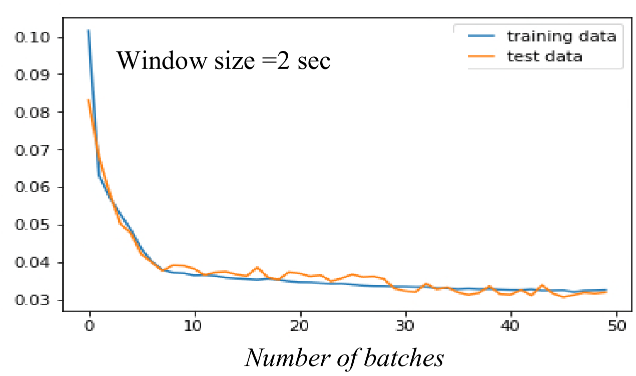

Figure 4. MSE loss (%) for number of neurons=8

 ISBN: 978-1-61208-734-4

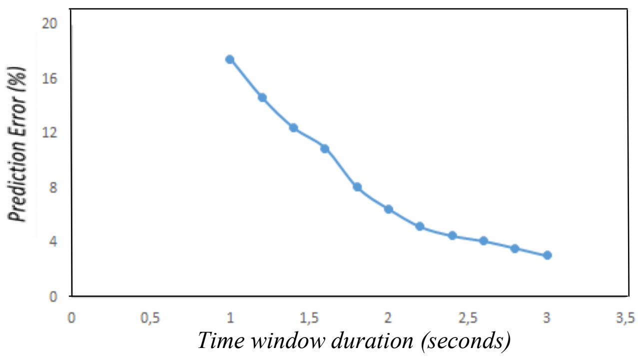

Figure 5. Model accuracy versus time window size

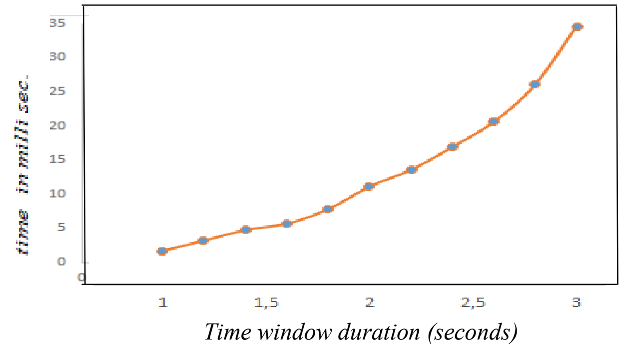

Figure 6. Execution time for inference model vs time window size

KNN is not, however, suitable for cases like ours where large datasets are required for training because the device does not have sufficient storage for the data.

Malhotra et. al. presented stacked LSTM networks for anomaly detection in time series. They evaluated their algorithms on different sensor data sets [16]. However, their algorithm is not ported to edge devices.

VII. Conclusions and Future Work

This study has shown that an LSTM-based Edge Machine Learning can bring about substantial improvement in energy efficiency of resource constrained IoT devices. Advanced machine learning platforms have significantly simplified the practical application of deep learning models by facilitating rapid prototyping and testing.

Due to physical and physiological differences in human beings, the models should be trained on an individual's own data. In our next study, we shall investigate personalized models rather than one-size-fits-all generic ones.

Though not empirically validated yet, we see that further optimizations can be made on the model. Pruning edge weights, applied by Han et al. [19], and exploiting inherent data types of sensor values can result in reduction in the computational overhead of the model. Utilizing binary neural networks [20] could also give interesting results. With this, it can be possible to deploy inference models on devices with even lower capabilities than the one used in this experiment.

References

[1] A. A. Thorp, "Prolonged sedentary time and physical activity in workplace and non-work contexts: a cross-sectional study of office, customer service and call center employees", International Journal of Behavioral Nutrition and Physical Activity Vol. 9, 2012.

[2] M. Graf, U. Guggenbuhl, and H. Krueger. "An assessment of seated activity and postures at five workplaces", International Journal of Industrial Ergonomics Vol. 15, pp. 81-90, 1995.

[3] P. Côté et. al., "The Burden and Determinants Of Neck Pain In Workers", Journal of Manipulative and Physiological Therapeutics Vol. 32, Iss. 25, 2009.

[4] J. Lennon, "Flexible study to fit around work", Occupational Health, Sutton, Vol. 67, Iss. 6, pp. 27-29, 2015.

[5] S. Kozey-Keadle et al., "The Feasibility of Reducing and Measuring Sedentary Time among Overweight, Non-Exercising OfficeWorkers", Hindawi Publishers, Journal of Obesity, 2012.

[6] S. M. R. Islam et al., "Internet of Things for Health Care: A Comprehensive Survey", IEEE Access, vol. 3, pp. 678-708, 2015.

[7] C. Marinac et al., "The Feasibility of Using SenseCams to Measure the Type and Context of Daily Sedentary Behaviors" International SenseCam Conference, Sandiego USA, 2013.

[8] W. S. Marras et al., "National occupational research agenda future directions in occupational musculoskeletal disorder health research", Elsevier Applied Ergonomics, 2009.

[9] M. T. Yazici, S. Basurra, and M. M. Gaber "Edge Machine Learning: Enabling Smart Internet of Things Applications", Big Data Cogn. Comput. 2018, 2, 26

[10] S. B. Baker, W. Xiang and I. Atkinson, "Internet of Things for Smart Healthcare: Technologies, Challenges and Opportunities", IEEE Access, vol. 5, pp. 26521-44, 2017.

[11] A. Godfrey et al., "From A to Z: Wearable technology explained", Elsevier, Maturitas, Vol. 113, pp. 40-47, 2018.

[12] I. Goodfellow et al., "Deep Learning", MIT Press, USA, 2016.

[13] C. Gomez, J. Oller, and J. Paradells. "Overview and Evaluation of Bluetooth Low Energy: An Emerging Low-Power Wireless Technology", Sensors Vol. 12, pp.11734-53, 2012.

[14] A. Kumar, S. Goyal and M.Varma. "Resource-efficient Machine Learning in 2 KB RAM for the Internet of Things", Proceedings of the 34 th International Conference on Machine Learning, Australia, 2017.

[15] "Measuring Bluetooth Low Energy Power Consumption", Texas Instruments http://www.ti.com/lit/an/swra347a/swra347a.pdf visited June 2019.

[16] Malhotra et al., "Long Short Term Memory Networks for Anomaly Detection in Time Series", Proc.s of 25th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Belgium, 2017.

[17] F. Chollet,"Deep Learning with Python", Manning Publ, 2018

[18] C. Gupta et al., "ProtoNN: Compressed and Accurate kNN for Resource-scarce Devices", Proceedings of the 34th International Conference on Machine Learning, Sydney, Australia, 2017.

[19] S. Han, H. Mao and W. J. Dally, "Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding", International Conference on Learning Representations, San Juan, Puerto Rico 2016.

[20] M. Courbariaux et al., "Binarized neural networks: Training deep neural networks with weights and activations constrained to +1 or -1", arXiv preprint arXiv:1602.02830, 2016.

[21] W. Yin, K. Kann, M. Yu, and H. Schütze. "Comparative Study of CNN and RNN for Natural Language Processing", arXiv:1702.01923, 2017.