A Remote Pulmonary Rehabilitation System Using the Wearable RESpeck Monitor

Teodora Georgescu

MInf Project (Part 2) Report
Master of Informatics
School of Informatics
University of Edinburgh

2020

Abstract

Chronic Obstructive Pulmonary Disorder (COPD) is a prevalent disease of the lungs and is particularly common amongst people of age 50 and above. Patients are often hospitalised following an exacerbation, and are sometimes advised to enrol in a Pulmonary Rehabilitation (PR) group. PR is a physical exercise program which has been shown to improve components of BODE index in COPD patients [4]. Many patients are not aware of the benefits of PR, and of those who are referred to a PR program, fewer than a half complete it, due to location, prices or even worsening of symptoms. In this project, we developed an end-to-end system that helps patients conduct PR at home using a simple Android application, and allows physicians and nurses to track their patients' progress remotely through a dashboard. We developed an activity recognition module which detects whether the PR exercises are performed correctly with an accuracy of 92.83%, using Convolutional Neural Networks. Moreover, we developed a dashboard functionality that displays information about the breathing rate during and after the exercises, as well as patients' resting time and correctness scores for a PR session.

Acknowledgements

I would like to thank my professor, D.K. Arvind, for his continuous guidance and discussions over new ideas. A big thank you to Andrew Bates, who was the technical guide throughout this project.

Lastly, I would have not completed this journey without my partner Sebastian (literally) by my side, my best friend Adela and my family. Thank you all for the endless talks and your promptness to help whenever I needed you.

Table of Contents

Acronyms

6MWT 6-minute Walk Test. 17, 53

CCQ Clinical COPD Questionnaire. 6, 7, 29

CNN Convolutional Neural Networks. 13, 21, 22

COPD Chronic Obstructive Pulmonary Disorder. 5, 6, 9, 10, 15–17, 33, 53, 54

CPM Continuous Patient Monitoring. 9

DNN Deep Neural Network. 13

ECG Electrocardiogram. 9

HAR Human Activity Recognition. 12, 13, 17–22, 25, 54

ICU Intensive Care Unit. 9

LOSOXV Leave-One-Subject-Out Cross-Validation. 21, 22, 36, 38, 44, 45

LR Logistic Regression. 12

PR Pulmonary Rehabilitation. 5–8, 10–12, 15, 16, 20, 25–31, 33, 43, 53, 54

RF Random Forest. 12

RNN Recurrent Neural Networks. 13

RPM Remote Patient Monitoring. 9

RR Respiratory Rate. 9, 11

Chapter 1

Introduction

Chronic Obstructive Pulmonary Disorder (COPD) is an umbrella term for chronic lung conditions that cause breathing difficulties [19]. The most common are emphysema and chronic bronchitis, which produce irreversible, long-term damage of the lungs and inflammation of the airways. This leads to symptoms such as breathlessness, wheezing, coughing and frequent chest infections [41]. Most diagnosed patients are over the age of 40 and prevalence increases with age [22]. COPD cases make up for 12.5% of the NHS emergency admissions to hospital, being the second largest cause of admission in the UK [43] and a leading cause of death worldwide [7]. Moreover, about a third of COPD patients will be readmitted to the hospital within three months [44].

A sudden worsening of the symptoms is called an *acute exacerbation* and is usually the reason COPD patients need to repeatedly seek medical help at a hospital [23]. Exacerbations are dangerous because they cause further damage to the lungs. Pulmonary Rehabilitation (PR) therapy has been shown to ease the general symptoms of COPD, help prevent exacerbations and improve patients' exercise capacity and quality of life [4]. Even with its proven benefits, PR is underutilised worldwide and frequently unaccessible to patients, with an approximate completion rate of under 10% [27]. Reasons include "insufficient funding, limited resources for PR programs, lack of healthcare professional, patient and caregiver awareness and knowledge regarding the process and benefits of PR", travel time and variable health of patients which causes them to miss PR sessions on the days they are unwell [54].

Although the MInf1 project was initially intended to only last one year, it can be easily integrated within the contents of the MInf2 project, since the two cover the same area of research. The two parts of the project are explained in the following sections.

1.1 MInf1

The primary task in the MInf1 project was to develop a model that could accurately identify periods of coughing from a person wearing the RESpeck monitor. This was achieved using a Random Forest Classifier trained on windows of accelerometer and gyroscope data from subjects performing several actions: breathing, coughing, talking, eating, laughing, singing and different types of movement. Episodes of stationary coughing were differentiated from non-coughing episodes and movement with an accuracy of 87.2%, and coughing during movement was correctly identified with an accuracy of 81.49%.

Already existing methods of cough monitoring use microphones or invasive tools such as chest straps or nasal cannulas. The RESpeck is, in contrast, a lightweight sensor the size of a 50p coin, worn on the lower left ribcage. Given that the frequency of coughing episodes is a good indicator of a COPD patient's health state, accurate cough detection is a powerful tool for continuous patient monitoring.

In the context of PR, the frequency and length of coughing episodes have been shown to decrease after the completion of PR therapy [57]. The ability to track and quantify the strength of cough episodes can be a strong indicator of the effectiveness of PR uptake in COPD patients.

1.2 MInf2

The aim of the MInf2 project was divided into four parts, and each part is explained in the paragraphs below:

1. Develop a mobile application for enabling patients to perform a set of PR exercises at home and analyse the adequacy of the performed exercises.
2. Develop a mobile application where patients can complete a daily CCQ test, a questionnaire assessing the perceived quality of life.
3. Develop a live dashboard accessible by medical staff, through which patient data would be made available.
4. Investigate the effectiveness and accuracy of the system using healthy volunteers.

The following systems works under the assumption that the patient is wearing the RESpeck monitor V5, which contains a 3-axis accelerometer, on their lower-left rib cage. The RESpeck sensor is shown in Figure 1.1 compared to a European coin.

Figure 1.1: Respeck dimensions compared to a €1 coin.

The PR Android application is named *Rehab3* and it allows users to choose from a list of PR exercises and a duration for each exercise. The app then displays a helpful

GIF of the current exercise to be performed and a countdown indicating the remaining time. The application is augmented with a Deep Learning model which classifies, in real-time, the correctness with which the exercises are completed. The model is trained on data collected from 15 volunteers, who were asked to perform each exercise in the PR program under the author's supervision at two different speeds: normal and slow. This method allowed us to analyse the exercises, taking into account that recovering patients might complete them at slower speeds.

The second Android application, named *Rehab Diary*, is a simple one-page application displaying the 10 questions of the Clinical COPD Questionnaire (CCQ) [31] and taking input from the user in the form of a score from 0 to 5 for each question. The patients complete the questionnaire each day and the medical staff can view their cumulative score and individual answers through the *Dashboard*.

The *Dahsboard* was developed using Google Cloud and is displaying charts of time series data for each selected patient. The information is pulled from Google Datastore and is displayed in a hierarchical view. The topmost view presents averages of the breathing rate, activity level, PR resting time between exercises and PR session correctness. When clicking on one particular datapoint in the graphs, the user can view more detailed data for each session, such as per-exercise correctness and momentary breathing rates between each PR exercise.

Finally, the validity of the system was tested using 3 volunteers who where asked to wear the RESpeck monitor and perform the PR routine at the two different speeds with the author's supervision. The automatically calculated results were compared to the author's notes of their PR performance. Furthermore, we conducted trials to test the accuracy of a step counting algorithm developed by a former student, particularly on slow walking and shuffling.

The achievements of this project were the following:

- We developed a fully functional, end-to-end PR system which helps patients adhere to their PR routine at home.
- We automated the process of filling in the daily CCQ, by including it in an automated *Rehab Diary* application.
- We created a *Dashboard* through which medical staff can access their patients' data and see their progress in real-time.
- We pushed the boundaries of currently available automatic PR system by developing a model which correctly identifies performed exercises with an average accuracy of 92.83% using Convolutional Neural Networks.
- We perform real-time classification of movements during the PR exercises and use these to predict a correctness score for each exercise, with an average accuracy of 94.19%.
- To the author's knowledge, there are currently no other PR systems that classify the correctness of each exercise using a single accelerometer while simultaneously recording the breathing rate of patients.

1.3 Outline

The report is structured as follows:

- **Chapter 2** presents an account of related work in the areas of remote patient monitoring, human activity recognition and pulmonary rehabilitation methods.
- In **Chapter 3**, more background about pulmonary rehabilitation is provided, as well as short discussions about deep learning for human activity recognition.
- **Chapter 4** presents the implementation details of the PR system, including the communication between the apps and automatic uploading of data.
- In **Chapter 5** we discuss the implementation of the classification algorithm, including the data collection for training the activity recognition models and the hyperparameter optimisation of the models.
- The analysis of the system validity is explained in **Chapter 6**.
- Finally, we discuss current limitations and further improvements for the system in **Chapter 7**, and provide a conclusion to this project.

Chapter 2

Related Work

This chapter provides a literature review of the methods applied for remote patient monitoring systems, universal and specific to COPD patients, human activity recognition and accurate step counting.

2.1 Remote Patient Monitoring

Continuous Patient Monitoring (CPM) had its inception in the Intensive Care Unit (ICU) in hospitals, where it is still being used to detect critical changes in patients' health signals [29], adjust therapy and improve patient safety [28]. These methods, however, limit the patient to the hospital rooms and do not provide any use when the patient has been discharged.

Remote Patient Monitoring (RPM) is a fast-growing area of technology and research, which allows patients to have normal daily activities outside of the hospital while still being closely monitored with the use of wearable sensors and/or ambient sensors. The greatest advantages of RPM are the real-time detection of illness, prevention of worsening illness (and deaths) and reduced number of hospitalisations [35]. Furthermore, many of these sensors disturb a patient's normal activity to the minimum because they are non-invasive, for example, wrist-worn sensors. RPM has been shown to improve outcomes for patients with various conditions, one of which is obstructive pulmonary disease [46].

In the context of specialised RPM for COPD patients, a number of technologies are currently available:

- **O_2 monitoring** - devices that monitor the level of oxygenation with minimal blood draws, providing information about a patient's lung condition [14].
- **Electrocardiogram (ECG) monitoring** - for example, the patch ECG [61]. Cardiovascular diseases are the most frequent comorbidities with COPD and monitoring a patients' cardiac condition is key in detecting early symptoms [15].
- **Respiratory Rate (RR) monitoring** - some devices estimate the RR from the ECG [8], or pulse oximetry and heart rate [39]. The RESpeck has been shown to

obtain very accurate respiratory rates which match measurements taken through obstructive methods such as nasal cannulas [3].

- **Spirometry** - by measuring the volume of air inspired and expired by a patient, the lung function can be assessed and airflow limitation can be detected. Severe airflow limitation confirms the presence of COPD [49]. Portable spirometers are available on the market but measurements taken by the patients on their own need to be carefully investigated by trained medical staff [12].
- **Environmental sensors** - present in the patients' environments and often embedded into everyday objects, they measure environmental temperature and air quality. There is a positive correlation between quality of life and air quality [11], and polluted air has been shown to be an important COPD risk factor [55].
- **Activity monitoring** - patients with COPD display significantly lower levels of physical activity than healthy subjects, particularly before the occurrence of an exacerbation [65]. Most patients do not follow physical activity recommendations, but remote monitoring has had a proven motivating effect on COPD patients, which lead to the increase in their physical activity [1]. Activity monitoring is usually achieved through the use of pedometers and accelerometers. The RESpeck sensor has also been shown to accurately identify the activity level of a patient [36].

However, none of the traditional methods provide a means for encouraging the patients to take up and adhere to PR at home.

2.2 Remote Pulmonary Rehabilitation

Several remote pulmonary rehabilitation systems have been studied in the past years. Each one of these systems relies on continuously monitoring a bodily metric, be it SpO_2 or the heart rate, to get a sense of the patient's state during the execution of the exercises.

2.2.1 Heart Rate and SpO_2 Monitoring for PR

Tang et. al [60] enrolled 37 healthy participants in a remote PR exercise session each, conducted using the eHAB telerehabilitation system [50]. They tested the feasibility of remotely monitoring the subjects' oxygen saturation and heart rate and found an 80% agreement between remotely transmitted data and the participant's self-reported measurements. In addition, the participants found the system usable and were confident they could use it at home for an extended period of time. This study proved that taking remote measurements of a patient's data is feasible, especially when trying to deliver PR to less accessible areas.

Marshall et. al [37] developed a pilot application inspired by a standard PR program used within the UK National Health Service. Patients were instructed to perform 12 exercises for 1-5 minutes each, while heart rate and SpO_2 data was sampled every 30 seconds from a Nonin 4100 Bluetooth Pulse Oximeter. The application was not deployed on a full patient trial, however, real COPD patients were interviewed by their physiotherapists with regards to the usability of the application, and their opinions were taken into account during the design phase. The stand-out characteristic of the application was showing the patient images of the movements and the duration for which they were supposed to perform the exercise.

One study investigated the effects of an app-based PR program for patients with advanced lung cancer, where a significant improvement in some symptoms was observed (activity level, exercise capacity, emotional and social functioning, fatigue, depression and anxiety) [48]. The patients were instructed to wear an accelerometer-based device on their wrist, which tracked their activity level and heart rate during the exercises. Although this study proved the efficiency of remote PR programs, the patients' performance for each exercise was not analysed, nor were the patients' respiratory rates recorded throughout the trial.

2.2.2 Exercise Classification for PR

Zhou [68] developed a remote PR application as part of their MSc thesis, which served as an inspiration for this project. The application was designed for ease of use at home by COPD patients, showing a range of instructions for exercises, timers and breathing rate information. The breathing rate and conformance to the exercise regime data were provided to the physiotherapists. The interface is shown in Figure 2.1.

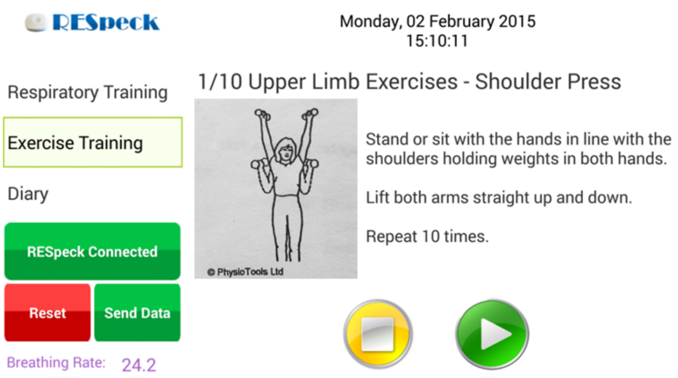


Figure 2.1: Interface of the initial Pulmonary Rehabilitation app using the RESpeck [68].

The patients' RR and activity levels were collected by a RESpeck sensor during the exercise sessions. The raw data was collected on the Android device and later transferred to the server manually. For detecting exercises, the author used a simple neural network trained on a supervised exercise session. The performance of the classifier is evaluated in two separate experiments. First, a separate classifier was trained for three different speeds, using *a single subject's data*. The performance of the classifiers was between 83.3%-90.6%, averaged over the three speeds. The accuracy of the classifiers when trained at one speed and training another speed were between 34%-58%. Secondly, the data from one group of patients was used to train a separate classifier, which was then used to classify the exercises performed by another group. The average classification accuracy was 66.8%, suggesting that exercise signals vary considerably among subjects.

To account for the shortcomings in the previous project, we collected data from a wide range of subjects, asking them to perform the exercises at different speeds, and trained a universal classifier for the system.

2.2.3 Breathing Rate Monitoring

In his thesis, Fischer [16] presented a novel way of assessing the well-being of patients with COPD by predicting the CAT score using the breathing rate signals during the PR session resting time. The method was validated on 31 COPD patients who were wearing the RESpeck monitor and completed PR exercises at home. The patients also filled in COPD Assessment Tests daily, which estimated their disease state. The respiratory rate and the tidal volume were shown to highly correlate to the CAT scores. This proved that the resting respiratory rate of a patient during a PR session, as measured by the RESpeck, can be used as a predictor for the well-being of a patient. The breathing rate is estimated using the RESpeck as shown in [3], and a higher breathing rate during rest periods is highly correlated with a worse patient condition.

This project builds upon the previous studies and fills the gaps in current research by continuously monitoring the patients' activity level and respiratory rate throughout the day, and by analysing the patients' performance during the exercise sessions, as well as their breathing rates during resting times.

2.3 Human Activity Recognition

An important part of the system we developed for this project was the ability to accurately detect completed exercises and in what proportion they were performed correctly.

Human Activity Recognition (HAR) is an active field of research, with new advancements with the advent of deep learning methods and it involves classifying or predicting the movement of a person based on sensor data. The most popular sensors for HAR are accelerometers, and many current methods rely on manual feature extraction from the accelerometer data, such as filtering the signal, extracting mean value and variance of the signal and applying Fourier and wavelet transforms [34]. A popular set of activities that is often used as a benchmark are daily activities such as walking, climbing stairs, doing housework or riding the bike, with classifiers such as Random Forest (RF) and Logistic Regression (LR) obtaining accuracies of up to 94% on certain datasets [62]. The disadvantage of these models is the dependence on domain-specific knowledge in signal processing to correctly engineer features fit for a machine learning model, and the generalisation power of the models is restricted [66].

Velloso et al [64] developed a model for analysing the correctness of weightlifting exercising by placing three Inertial Magnetic Units (IMUs) on subjects' arm, belt and glove, and an additional IMU on the dumbbell. They recorded 5 different alterations of one exercise, only one of which was the correct one. The other 4 variations were common mistakes beginners make when weightlifting. Using a Random Forest Classifier, this method resulted in a 98.2% accuracy but, as the authors suggested, it is not a scalable approach to record all possible mistakes for a single exercise.

Deep Neural Network (DNN) methods enable automatic feature learning and are currently achieving state-of-the-art results for HAR. The greatest advantage of DNNs is that they can learn features directly from the raw data, eliminating the issue of feature engineering, which speeds up both the process of training and analysing live data.

There are two main types of DNNs that are most suitable to HAR problems: Recurrent Neural Networks (RNN) and Convolutional Neural Networks (CNN). RNNs are efficient at recognising short activities that contain an inherent order, for example gait cycles (stance and swing phases) [53], while CNNs are better at learning features contained in recursive patterns, such as the repetitions of a physical exercise.

One study that had particularly good results when using CNNs for activity classification was conducted by Um et. al [63]. The authors used a forearm-worn IMU to classify 50 gym exercises recorded by professional athletes. The raw acceleration and orientation data was aggregated into a 2D image, and the x, y and z axes of the data were treated as the channels of the 2D image. The images were then passed through a 3-layer CNN with up to 128 feature maps per layer and an accuracy of 92.1% on the test set was obtained.

2.3.1 Justification for using CNNs

One of the main features of the *Rehab3* app was to perform real-time classification of the exercises, and send these summaries to the server so that physiotherapists can have up-to-date information about the activity of their patients. At the time of writing, CNNs had a simple and efficient implementation for Android apps in Tensorflow, as opposed to RNNs. Furthermore, given our goal of correctly recognising patterned activities instead of long-term activities, we decided to concentrate on the usage of CNNs.

Chapter 3

Background

This chapter introduces the reader to ideas and methods used throughout the report. First, a short overview of the medical concepts is provided: Pulmonary Rehabilitation, the BODE index and the 6-minute Walk Test. Next, a summary of Human Activity Recognition methods is presented, followed by a brief overview of Deep Neural Networks, in particular, Convolutional Neural Networks.

3.1 Pulmonary Rehabilitation

Pulmonary Rehabilitation (PR) is a physical exercise and educational program designed for patients with severe respiratory symptoms, such as breathlessness. It helps patients better understand and manage their disease through physical and breathing exercises. Most patients that enrol in PR programs are COPD patients, but people with other long-term lung dysfunctions, such as emphysema or bronchitis, also benefit from it [42]. PR has been shown to significantly improve components of the quality of life (physical function, health perception, depression), dyspnea and functional capacity, regardless of the severity of the illness at the start of the program uptake [56], [21].

A PR course is usually six to eight weeks long and is taught in groups of up to 16 people by trained healthcare professionals, such as physiotherapists and nurses [18]. The sessions are typically held in local hospitals, community halls and other health centres, which means that the patient needs to travel to these meeting points. Examples of exercises include bicep curls, leg extensions, heel raises and, most importantly, walking [13]. All prescribed exercises are designed to increase the lung capacity by strengthening the breathing muscles and the overall physical condition of the patient.

As mentioned in Chapter 1, the overall uptake of PR, even with its proven benefits, is disappointingly low. A national survey in the US revealed that 57% of COPD patients with heavy symptoms were not even aware of PR [2]. It is estimated that only 6% of COPD patients enrol in PR programs and, from those who start a program, only 50% complete it [40]. Reasons include location, time off work and degrading health condition. By developing an application that patients can use in the comfort of their own home could make the program more accessible and increase the adherence and completion rates.

3.1.1 BODE Index

The BODE index is a scoring system for COPD that has been shown to be a good predictor of mortality [51]. Four different factors are evaluated in the BODE index:

- **B**ody Mass Index (BMI) - a ratio of the height and weight of a patient, estimating how underweight or overweight they are.
- Airway **O**bstruction - measured by testing the forced expiratory volume (FEV_1), i.e. the volume of air forcefully exhaled in one second, after a deep breath.
- **D**yspnea - the degree of breathlessness.
- **E**xercise tolerance - measured in the amount of steps taken during a 6-minute walk test.

PR has been shown to improve some components of the BODE index by decreasing dyspnea and increasing exercise capacity [10].

3.1.2 PR exercises

For this project, we used a set of 10 PR exercises, recommended by the British Lung Foundation in their PR take-home leaflets [17]. The exercises are shown in Appendix A, Figures A.1 - A.10 and explained below:

1. Sit to stand - standing up from a sitting position and sitting back down on a chair. This strengthens a patient's quads, glutes and improves their cardio.
2. Leg extensions - while sitting down on a chair, extending one leg at a time parallel to the ground. This exercises the upper legs. The patient should switch legs halfway through the exercise.
3. Wall squats - sliding up and down while leaning on a wall, keeping the back straight. This strengthens the legs and upper body.
4. Heel raises - keeping hands on a still surface for balance, rising up on the tips of the toes and coming back down. This exercises the calf muscles.
5. Bicep curls - filling up a bottle or a carton of milk with water or sand and raising it up to the shoulders. This works out the upper arms of the patient.
6. Shoulder presses - raising the arms up above the head, and bringing them back down. This movement works the arms and upper back.
7. Wall push-offs - standing at arm's reach from the wall, then slowly leaning into the wall, like a vertical push-up. This movement exercises the arms and back of a patient.
8. Leg slides - keeping hands on a still surface for balance, sliding each leg at a time to the side and bringing it back in the middle. This works the core, upper legs and balance. The patient should switch legs halfway through the exercise.

9. Step ups - stepping up and down a stair step or exercise block. This movement develops the leg muscles and improves cardio condition. The patient should alternate the stepping foot throughout the exercise.

10. Walking - typically, the 6-minute Walk Test (6MWT). This exercise is optimal for the maintenance of leg and upper body muscles, as well as general exercise tolerance.

All the exercise images are shown as GIFs in the *Rehab3* app, to encourage the user to perform them as correctly as possible.

3.1.3 The 6-minute Walk Test

Walking is one of the most commonly performed day-to-day activities and the most severely impacted in patients with COPD. The 6-minute Walk Test (6MWT) has been shown to reliably assess the functional exercise capacity in patients with COPD [24]. This type of assessment has traditionally been done through questionnaires, or simply asking the patients how far they can walk. Unfortunately, human recollection is not reliable and patients can often provide under- or over-estimations of their true exercise tolerance. The 6MWT is a simple, reliable objective measure of the distance that a patient can walk on a flat, hard surface in a period of 6 minutes. The patients are allowed to choose their own intensity of exercise and may even stop and rest during the test. The results of such a self-paced test can better reflect the daily physical activity tolerance, since most daily activities are not performed at the maximum level of extortion [47].

Burioka et. al [6] showed in their study that the number of steps walked per second (NSPS) is an equally good indicator of the efficacy of PR for COPD patients. Therefore, we implemented a Peak Detection algorithm for counting the number of steps taken during a 6MWT, inspired by Jordi Sorribas' master thesis [58]. Sorribas demonstrated the efficacy of this algorithm in his thesis, citing an accuracy of 86.8%, and stating its efficiency especially on slow walking patterns.

We report the total number of steps taken during the 6MWT, as well as an estimate of the average number of steps per second during the session. This is obtained by simply dividing the number of steps by the duration of the exercise.

3.2 Human Activity Recognition

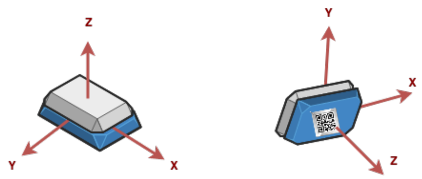
Human Activity Recognition is an emerging research area in the field of machine learning and pattern recognition. Its main goal is to allow researchers to automatically infer the subjects' activities from body-worn inertial sensors. HAR shares many methods with related fields, such as computer vision or natural language processing, but also presents a unique set of challenges that require other computational approaches. For example, computer vision and language processing tasks can have very clear problem definitions, such as the detection of a certain object in an image or a spoken word in a sentence. In contrast, in HAR, especially long-term monitoring, the set of relevant activities is often impossible to define at the beginning of the study. Moreover, the variability with which people can perform the same action is one of the main reasons for which a HAR system requires a different set of performance measures rather than just accuracy.

The following sub-sections present the details of the most commonly used sensor for HAR, the accelerometer and the RESpeck in particular, with a discussion of expected challenges in a HAR task.

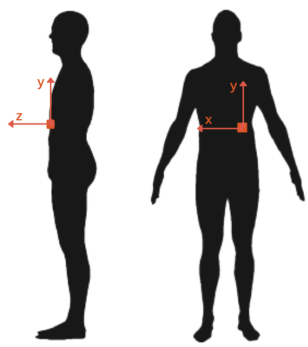
3.2.1 Accelerometers

An accelerometer measures the linear acceleration, i.e. the rate of change of the velocity of an object in the three orthogonal directions. Measurements are made in m/s^2 and we refer to the three directions as the axes: $accel_x$, $accel_y$, $accel_z$. There are two types of acceleration forces: static, such as the force of gravity which always points downwards, and dynamic, caused by the action of an external agent. An accelerometer can identify its orientation in space by finding out which direction is ”downwards” - measuring the amount of gravitational acceleration. Similarly, by measuring the amount of dynamic acceleration in each direction, we can analyse the direction in which a device is moving [59].

The RESpeck (V5) is a small chest-worn device containing a 3-axis accelerometer, that can be paired with an Android App and send packets of accelerometer data via Bluetooth Low Energy at a rate of 12.5Hz. The sensor is mounted on the front of a subject’s chest, right below the left ribs. This choice of placement is optimal for both respiratory and activity monitoring purposes - the chest is easily the best location of the body for the extraction of respiratory signals, while also being very close to the human body centre of mass, which translates into well-defined movement for most activities. Figure 3.1a illustrates the three axes of measurement of the device and Figure 3.1b shows the sensor mounting on a subject.



(a) Axes of measurement.



(b) Sensor placement.

Figure 3.1: RESpeck axes of measurement and placement on a subject’s body.

3.2.2 Challenges

Bulling et. al [5] provide a comprehensive discussion on the methods of HAR. The paper identifies three main types of challenges: common research challenges - shared with other fields of pattern recognition, challenges specific to HAR and application challenges.

Common research challenges include:

- *Intra-class variability*. The same activity might be performed very differently among subjects, and it can even occur when the activity is performed by the same person. Factors that influence this diversity could be emotional (stress, fatigue) or environmental (walking on a flat surface vs. walking uphill). This is partly overcome by training the HAR model on a wide range on subjects.
- *Inter-class similarity*. Two very different activities can show very similar characteristics in accelerometer data. For example, standing and sitting might give the exact same accelerometer signals, even if the activities are clearly different. This can be fixed either by placing additional sensors in the subject's environment (e.g. another accelerometer on the leg to indicate the leg position) or by using another key action to identify the activity (for example, the previous action of sitting down will indicate that the subject is not standing).
- *The NULL class*. Usually, HAR systems are only concerned with a small part of the continuous input data stream they receive from the sensor, for example, in the case of continuous monitoring. This results in a huge class imbalance between relevant and irrelevant (NULL class) actions, and may lead to an increased number of false positives. One way to reduce the impact of the class imbalance is to include some noisy activities in the classifier training.

Some unique challenges when working on a HAR system are:

- *Definition of physical activities*. Clearly defining which activities and characteristics are most important for the task at hand is often one of the first obstacles in developing a HAR system.
- *Class imbalance*. Many activities occur infrequently and for a short period of time, such as brushing one's teeth or drinking a glass of water, while very few activities occur for a longer period of time: sleeping, sitting down and walking are among the most usual ones. This produces a considerable class imbalance if the researches rely only on continuous data gathering. This can be addressed by recording additional training data for the smaller classes, to reduce the imbalance and improve the performance of the model.
- *Ground truth labels*. Manually annotating HAR data is an expensive and tedious task, especially if the annotator is only shown raw accelerometer data. To improve this process, data should be annotated automatically during collection, by asking the participants to perform a certain activity and label the whole recording time as that particular activity. This might, however, introduce bias in the way a subject is performing the activity and is also subject to outliers - for example if a subject suddenly stops during a walk.

Finally, challenges that arise when designing the application system itself are:

- *Variability in sensor data*. A problem in reusing previously collected datasets is that each type of sensor yields different signals. This depends on the make and type of sensor and its placement on the subject's body (some sensors are wrist worn, some sensors are placed in pockets).
- *Tradeoffs in system design*. When designing a HAR system, there are always tradeoffs between accuracy, latency, processing power and energy usage. Many real-world applications of HAR (such as gesture recognition) require very low latency, while for other types of HAR (such as continuous monitoring) offline classification with a one-day latency might be enough, so more resources could be spent on computational complexity.

All these challenges were carefully considered during the development of the system. The *intra-class variability* was addressed by recruiting a wide range of participants and asking them to perform each exercise in two different ways - slower, as a patient might, and faster, as a healthy person would. The *inter-class similarity* was resolved by collecting enough data for potentially misleading actions - for example, we suspected that bicep curls might be confused with simply standing, so we collected standing data from participants. This also helped with the *NULL class* problem, even though our system did not need significant amounts of noisy data to achieve a good performance.

Regarding the *definition of physical activities*, each of the ten PR exercises is well defined and unique as a physical activity (as described in Section 3.1.2) for the purposes of identification. The drawback was that there were no accessible *ground truth labels* for the exercises, that is, we had no guidance from a medical professional with regards to how correct exercises were to be performed. We merely followed the instructions given out in the British Lung Foundation leaflets, as a normal patient would.

Finally, due to *variability in sensor data*, we could not reuse any previously collected data sets, for example, for walking. One other reason was the special placement of our sensor, on the lower left coastal margin, which has no other equivalents in open-source data. For the purpose of this project, we devised a separate data collection protocol and personally collected data from volunteers to match our needs. A particular problem faced in this project was the *system design trade-off* between classification accuracy, processing power and energy consumption. While the accuracy is completely determined by the choice of model, the processing power of the Android phone had an impact on its energy consumption. We needed to find an algorithm that would be accurate enough for clinical usage, but preserve the Android battery as much as possible, so that the patients would not need to charge it all the time (or possibly run out of battery during a PR session).

3.2.3 Data processing techniques

In a typical HAR task, the raw sensor data is first preprocessed using signal smoothing methods to filter out noisy signals and outliers. In this project we used a simple moving average filter, which takes the average of every N consecutive samples of the wavelet passed as input [38]. This algorithm has a very low computational complexity and was a perfect trade-off between performance and latency for our task. Figure 3.2 shows the effect of different sizes of N on the shape of the signal. We can see that a small N is not effective enough at removing signal noise, while a too large N alters the shape of the signal altogether. For smoothing the walking signals, we used an experimentally determined $N = 10$.

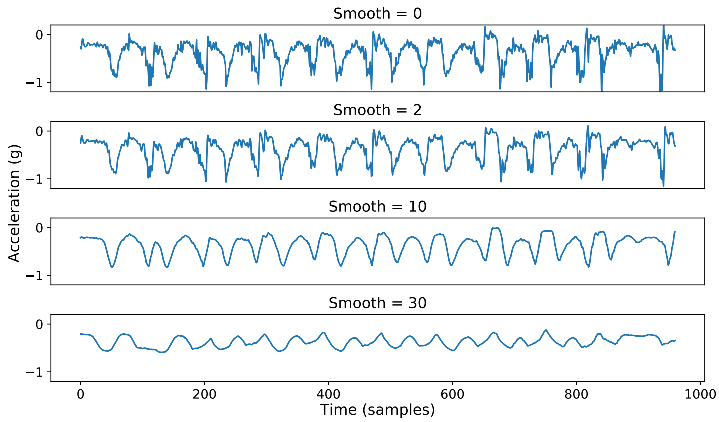


Figure 3.2: The effect of different sizes of smoothing windows.

The preprocessed data is then segmented into windows of data. The size of the window depends on the activity that needs to be identified. Small, quick movements such as steps may require only a small window size, since the action of interest is not long in duration. When trying to identify sequences of moves, however, the optimal window size might be larger, for example, if we wish to include both a peak and a trough of the signal. For this project, the window size was determined experimentally, taking into consideration the fact that some activities might be performed slower by actual patients. The window size is fixed only for walking, where we consider 1-minute windows of data. At a sampling frequency of 12.5Hz, the window size equals 750 data points.

Next, relevant features are extracted from the windowed data. This can either be done manually, by choosing features such as the range of a signal or its standard deviation from the mean, or automatically, by feeding the windowed data into a neural network and let the algorithm identify the relevant features. Last year's project, summarised in Section 1.1, focused on manually extracting these features from the raw data. This year's project aimed at using Convolutional Neural Networks (CNN) to automatically identify the best features, as described in the next sections.

Finally, the relevant features from the data and the activity true labels are passed through a classification algorithm. Due to the *intra-class variability*, the classification is run using the Leave-One-Subject-Out Cross-Validation (LOSOXV) method to correctly assess how the model will perform on future, unseen data. LOSOXV is a statistical technique which uses one data point as the validation set, and the rest of the data as training input, repeating this process for all data points in the set. The performance is averaged over all trials to provide an estimate of the generalisation performance of the model [30]. In a HAR task, we treat recordings from a single subject

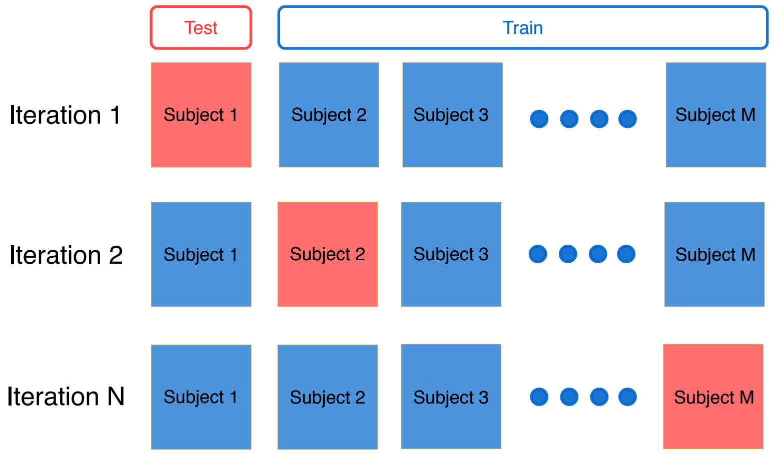


Figure 3.3: Leave-One-Subject-Out Cross-Validation (LOSOXV) for a HAR task.

as a unique data point. The data set is therefore partitioned into subject-specific data, and the model is iteratively trained on the partition not containing one subject's data [67]. This process is illustrated in Figure 3.3.

3.3 Convolutional Neural Networks

CNNs, introduced as "neurocognitrons" at their beginnings and later popularised by LeCun et. al [32] were inspired by the visual nervous system in vertebrates [20]. Their main advantage is the ability to recognise geometrical patterns irrespective of their position in an image and shape distortions.

Similar to neural networks, the building blocks of CNNs are neurons with learnable weights and biases. CNNs are trained using three additional concepts: local receptive fields, shared weights and pooling [45].

In a classical neural network, each neuron is fully connected to the previous layer. In a CNN, each hidden neuron is connected to a small region of the input neurons, called the *local receptive field*. These fields share the same weights and biases in each layer, meaning they are designed to detect the same geometric feature regardless of its location on the inputs image. Local receptive fields are also known as filters or kernels, and are slid (*convolved*) over the entire input image, as shown in Figure 3.4. The dot product between the weights of the filter and the input part of the image at the kernel gives the output of the layer.

A typical CNN also contains pooling layers after successive convolutional layers, which serve as a dimensionality reduction unit for the number of parameters and computation in the network. The filter of the pooling layer slides over the input with a certain stride and takes the maximum (or average) of all values in that region. This way, the dimensionality of a layer can be reduced by a factor equal to the stride.

CNNs are particularly useful for image classification, and in this project, we leveraged their ability to detect local features by treating a window of accelerometer signals as a three-channel one-dimensional image. Figure 3.4 illustrates the process of one convolutional layer. Each input channel corresponds to one of the $accel_x$, $accel_y$ and $accel_z$ axes. We then slide a kernel over each channel of our 1D image to obtain a convolved signal in a number N of output channels, where N is equal to the number of kernels slid over the input. The convolved signal then becomes the input to the second convolutional layer.

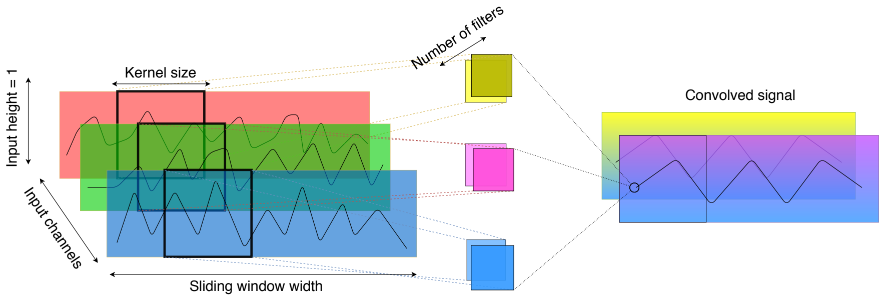


Figure 3.4: Convolution layer.

One particular preprocessing step that CNNs benefit from is input normalisation [33]. In a neural network trained using backpropagation and gradient descent, the cost function E_{train} can be computed by measuring the average of each E_p:

$$E_p = C(D_p, M(Z_p, W))$$
$$E_{train} = \frac{1}{P} \sum_{p=1} E_p$$

where Z_p is the p^{th} input pattern, W is a set of learnable parameters for all inputs to the system, and E_p measures the difference between the desired output D_p for the pattern Z_p and the output M produced by the system. Simply put, the learning problem consists of finding an appropriate set W which minimises $E_{train}(W)$. The gradient descent algorithm adjusts the parameters iteratively as follows:

$$W(t) = W(t-1) - \eta \frac{\delta E}{\delta W}$$

where η is usually a scalar constant, knows as the *learning rate*.

Convergence to a good minimum of the E_{train} function is faster if the average of the training set input variables is close to zero, and also if they are scaled so that they have about the same covariance. For example, take two example features, X_1 ranging from $[1,10]$ and X_2 ranging from $[0,1]$. Their corresponding weights W_1 and W_2 will also have different scales because the network is learning how to find a linear combination of the inputs. The resulting loss function topology will be an ellipse skewed in the direction of the larger weight W_1 and the gradient descent algorithm will place more emphasis on the larger parameter gradient. This leads to a zig-zag descent which is slow, inefficient and might miss the local minimum, as shown in Figure 3.5.

Moreover, having the inputs in the same range of values helps the input nodes learn good weights at a fast enough rate. Therefore, we needed to standardise and normalise

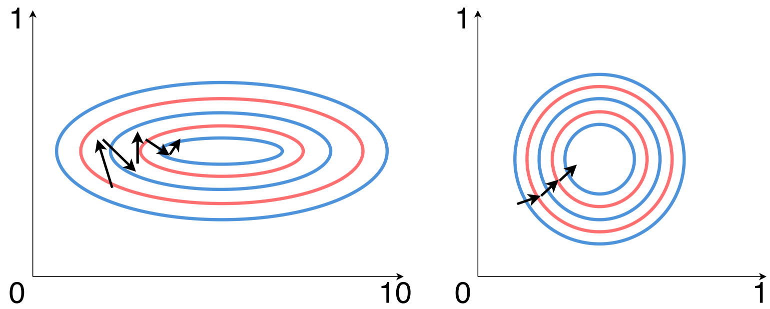


Figure 3.5: Gradient descent for unscaled values (left) and scaled values (right).

the inputs to our CNN, and also use Batch Normalisation layers [26] after each Convolutional layer in order to keep the scaling consistent throughout the network.

Chapter 4

Pulmonary Rehabilitation System

In this chapter, we discuss the structure of the Rehab system. First, the *Rehab3* app is presented, along with the development of the HAR model which allows the app to identify correctly performed PR exercises. We also discuss how the raw data is stored and subsequently uploaded to the Cloud. Next, we present the *Rehab Diary* application and the types of questions being presented to the patients. Finally, the Dashboard is presented, with a discussion of the various types of information summarised for the physiotherapists.

4.1 Application Ecosystem

The *Rehab3* app, *Rehab Diary* app and the *Dashboard* were all implemented within the existing Speckled ecosystem. Figure 4.1 illustrates the communication between system components: the RESpeck transmits the accelerometer data of the patient to the Android phone, which subsequently sends summaries of the data to Google Cloud, assuming a stable internet connection. The Dashboard then pulls data from the Google Cloud database and displays it to the physiotherapists in charge of the patients.

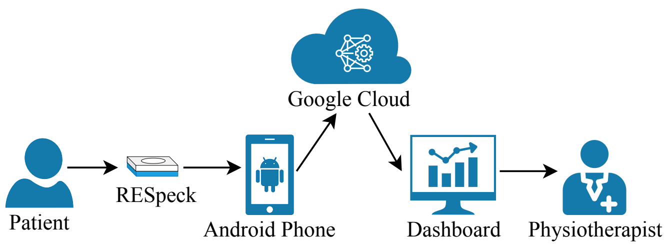


Figure 4.1: Communication between system components.

Figure 4.2 describes the lower-level interactions between the existing application, Air-Respeck, and the *Rehab3* and *Rehab Diary* applications. First, the Android phone is paired with one RESpeck via the Pairing App - this establishes a continuous BLE connection between the two devices. Then, the existing AirRespeck application receives packets of raw accelerometer data from the RESpeck and uses them to perform daily activity monitoring. The raw data is safely stored on the phone's SD card.

Whenever a user starts a PR session and opens the *Rehab3* app, AirRespeck forwards the accelerometer data to *Rehab3* through broadcasting. *Rehab3* records raw data separately and saves it on the phone memory. After the PR session is over, *Rehab3* sends a summary of the session back to AirRespeck in the form of a JSON object. The summary includes the duration and correctness of each exercise, the resting time and the breathing rate during the resting times. AirRespeck uploads this summary and its own daily activity summaries to the Cloud Datastore. Uploading happens once a minute and, in the event of network disconnection, AirRespeck queues up the packets and sends them once the connection has been reestablished. When a user completes a *Rehab Diary* entry, the scores for each individual question are sent to the AirRespeck, which also uploads these together with the rest of the summarised data.

The raw data stored on the phone's SD card is regularly uploaded manually, by downloading the files to a local computer and running an upload script. On the Cloud side, the raw data is automatically formatted and appended to the Datastore.

Finally, the *Dashboard* draws data from the Datastore and displays it in different levels of abstractions. The *Dashboard* users, typically the physiotherapists, can select a range of dates for which to show the data. In a time-hierarchical view, we display the breathing rate and activity levels throughout the selected time period, the average resting time between PR exercises, the average correctness of the performed exercises and the daily average CCQ score.

The system is designed for international use, therefore all developed applications are translated based on the selected locale. The current possible languages are English, Dutch and Italian.

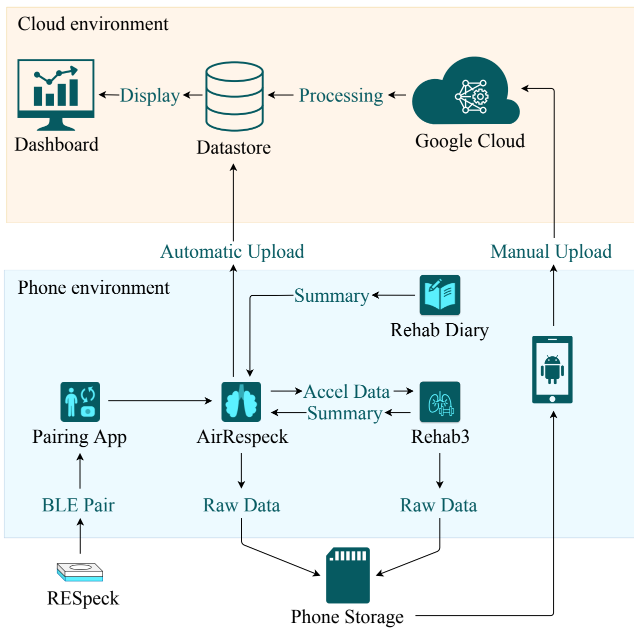


Figure 4.2: Application ecosystem.

4.2 Rehab3 App

The *Rehab3* app was first developed as part of Zhou's master project [68]. As discussed in Section 2.2.2, the old version of the application had a differently structured interface. The newer version developed in this project was designed to run on Android phones with the screen of at least 5", to accommodate the large buttons and writing. We followed most of the principles laid out in [68] for designing for older users: keeping messages as short as possible, reducing choice wherever possible, use easy to understand images and keep the text in a large font.

The application can be launched either from the home screen or from the AirRespeck app. Figure 4.3 shows the screen in AirRespeck from which the patients can start both the *Rehab3* app and the *Rehab Diary* app. The screen also shows the connection status of the RESpeck and allows patients to troubleshoot in case the RESpeck is not connected.

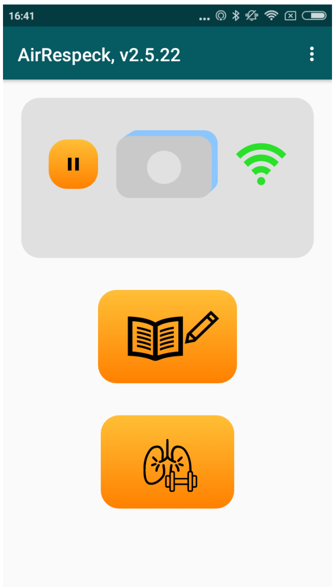


Figure 4.3: AirRespeck "Subject Mode" screen, where users can see the RESpeck connection status (top), start the *Rehab Diary* app (middle) and the *Rehab3* app (bottom).

Once the *Rehab3* app is launched, the user is asked to select at least one exercise from a list of the 10 options discussed in Section 3.1.2. Then, the user is asked to select a duration for the selected exercises, and number of repetitions. If the user selects *Walking* as one of the exercises, they will be prompted with an additional screen where they can select the duration of the walk. This sequence of choices is illustrated in Figure 4.4.

The PR session is then started and the users are shown a screen with the name of the exercise they selected, the number of reps they want to perform the exercise for and the expected time it should take them. In the middle of the screen, they are shown a GIF with the appropriate PR exercise. The complete list of GIFs can be found in Appendix A. After the time for the exercise runs out, the users are told to rest until their breathing comes back to normal. We disabled the "Next" button during the exercise time because

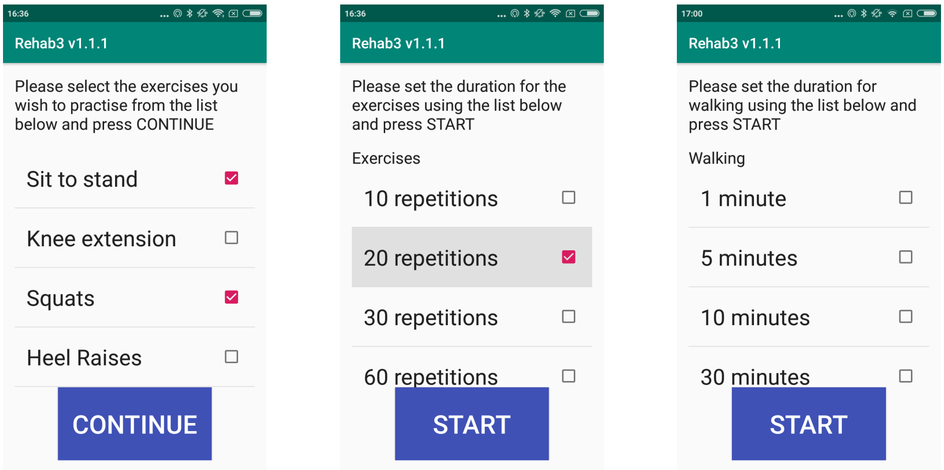


Figure 4.4: *Rehab3* interface for selecting exercises from a list (left), selecting the number of repetitions for each exercise (middle) and the duration for walking (right).

we wanted to encourage the users to finish the selected exercise. After the time is up and they are resting, they can press the "Next" button whenever they are ready to start the next exercise. When the user finishes all exercises, they can press on the "Finish" button and be taken to a summary screen which shows their average breathing rate during resting time for the past 3 PR sessions. Figure 4.5 shows the sequence of these screens.

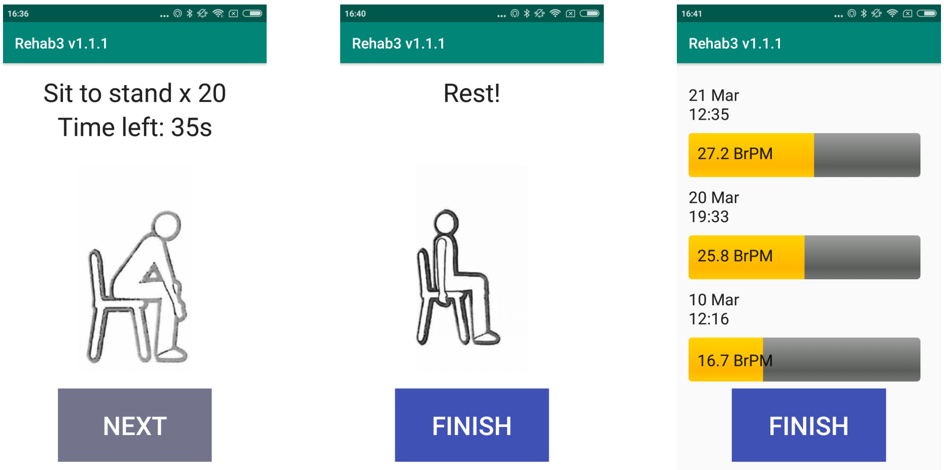


Figure 4.5: Rehab3 interface for performing a selected exercise (left), taking a rest until breathing comes back to normal (middle) and the summary screen (right).

During the execution of the exercises, a Machine Learning model is run on the Android phone which analyses the raw accelerometer data and classifies the recorded movements. If the movement classification matches the exercise that is currently selected on the interface, the application assigns a correctness score to the exercise. Further information about the development of this model is presented in Section 5.4.

When the exercise session is finished, a JSON object containing summaries of the session is sent back to the AirRespeck app. The summaries contain:

- A timestamp of the PR session.
- A list of the performed exercises containing:
  - Duration for each exercise
  - Resting time for each exercise
  - Breathing rate during the resting time for each exercise
- Average breathing rate for the session
- Standard deviation of the breathing rate for the session

The AirRespeck app subsequently uploads this to the Datastore on Google Cloud where the data is processed to be suitable for showing on the dashboard.

4.3 Rehab Diary Application

The purpose of the *Rehab Diary* app is to help the patient take a daily subjective assessment of their quality of life by completing a *Clinical COPD Questionnaire (CCQ)*. The CCQ consists of 10 questions in total and is divided into three parts: questions related to symptoms, functional activity and mental health. Each question is scored between 0 and 5 and the total score is calculated by adding up the scores. Higher scores indicate a worse perceived quality of life [31]. The example CCQ test used for this project was taken from the official Dutch CCQ website [52], which provides translations of the CCQ in over 60 languages.

The Rehab Diary application was built upon an existing version which was presenting the COPD Assessment Test (CAT) to subjects in other Speckled projects. The application skeleton consisted of a single activity, a TextView for displaying the current question and a pop-up alert with choices for the answers. The TextView and Alert are updated with each CCQ question from the set of 10 pre-determined questions. The answers are recorded in a simple array, which is sent to the AirRespeck app once all the questions have been completed. Figure 4.6 shows a possible user flow through the screens.

4.4 Dashboard View

All the recorded data is ultimately presented to the physiotherapists on the *Dashboard*. The data is organised by subject ID, and the users can select a date range for the data, as well as whether they wish to see the automatically uploaded data or the manually uploaded data. If the subject ID is present in the database, the RESpeck data will be loaded up and the user can see it displayed in graphs by clicking on one of the “Show dashboard” buttons.

The dashboard is split into 6 graphs, each of them containing hierarchical views of the data. At the top-level, the one automatically loaded when opening the dashboard, the user is presented with averaged data from the patient’s daily activity and PR sessions,

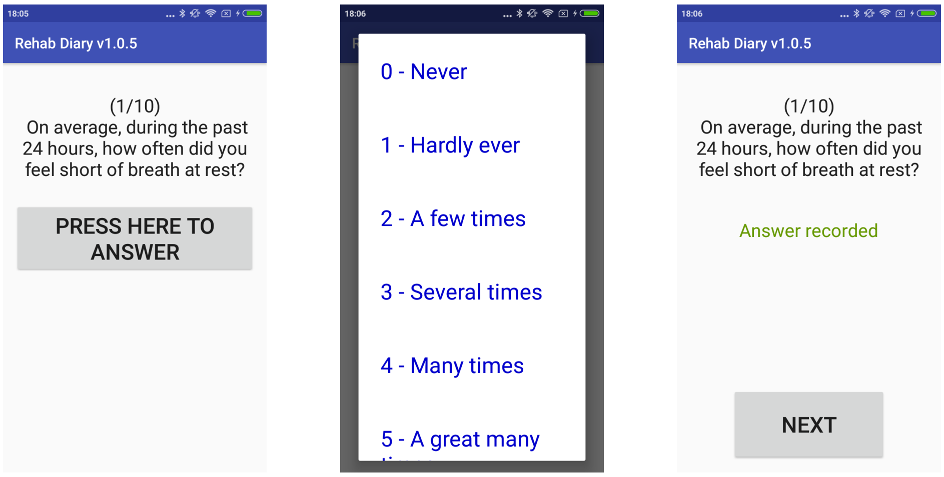


Figure 4.6: *Rehab Diary* interface displaying a CCQ question (left), taking input for the answer (middle) and confirming the answer was recorded (right).

as shown in Figure 4.7. The first column presents the average breathing rate throughout the day, a colour-coded activity level graph and the CCQ diary scores. The second column presents the Rehab sessions average breathing rate, the average resting time and the average correctness for each Rehab session. The graphs are synchronised, so hovering over one point on one of the graphs will show the corresponding data points on the other 5 graphs.

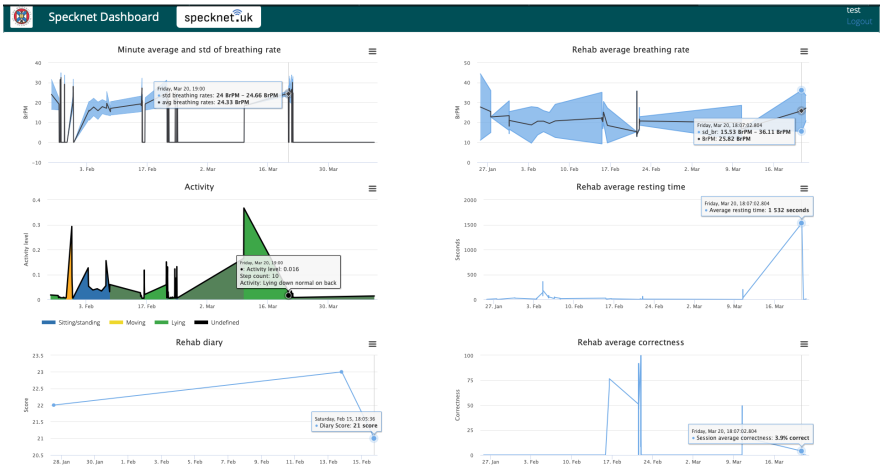


Figure 4.7: Top-level hierarchy of the dashboard. For a specific subject and time frame, averaged data is presented

The user may click on any point on the graphs (representing a Rehab session) and this will open up the second-level hierarchy of the dashboard, shown in Figure 4.8. This level of graphs shows the information about one specific PR session. The left column shows the minute average and standard deviation of the breathing rate, the activity level throughout the session and the CCQ diary score for that day. The right column presents the breathing rate datapoints during the rest time for each exercise in the session, the resting time for each exercise in the session and the correctness for each exercise in the session.

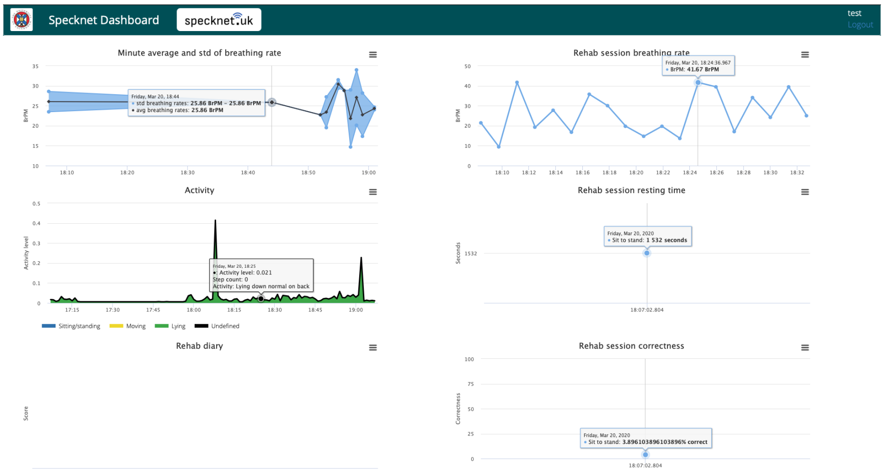


Figure 4.8: Bottom-level hierarchy of the dashboard. For a specific subject and time frame, session data is presented.

The *Rehab Diary* data is also available for download in `csv` format so that physiotherapists can track the score changes of each individual question.

The dashboard system uses Python Pyramid for its scripts and displays the graphs using JavaScript Highcharts. The system is designed to receive updates within minutes of the patient completing a PR session, thus providing timely information to the physiotherapists overseeing the process.

Chapter 5

Exercise Classification

Exercise detection is one of the key features of the Pulmonary Rehabilitation system. Not only can the system tell whether a subject is performing the correct activity at the right time, but the correctness of the performance of each exercise can also be estimated from the live data. The next sections present the training and hyperparameter tuning process for the model.

5.1 Data Collection

In order to train a model that could recognise correctly performed exercises, we needed to conduct our own data collection. We recruited 15 participants from the University of Edinburgh. No participants reported pre-existing lung conditions and all were in good physical shape. The data collection study was approved according to the Informatics Research Ethics Process with RT number #4118. The participant information sheet is attached in Appendix B, and the participant consent form is shown in Appendix C.

Participants were asked to perform the 10 PR exercises twice, at different speeds. The first speed was normal, as a healthy person would perform the exercises. The second speed was slower, aiming to reproduce movements of a person that has a low physical exercise tolerance. This was done to introduce some variability in the dataset, in the absence of actual COPD patients that could help with the data collection. Each exercise was performed for 10 repetitions. For the exercises which included alternate leg movement, the subjects were asked to perform half of the repetitions with one leg, and half of the repetitions with the other. Table 5.1 summarises the details about the collected data.

The distribution of timings per exercise, for each speed, averaged for all subjects, is illustrated in Figure 5.1. Each slow speed exercise took, on average, 20 seconds longer than its normal speed counterpart. The figure also shows that the dataset is not balanced. This was a conscious decision, further challenging our model to be able to draw good decision boundaries even in the presence of unbalanced class counts. Furthermore, setting the number of repetitions rather than the number of seconds to perform an exercise gave us an estimate of how long each exercise should last for a healthy

Participant ID	M/F	Age	Total time	Normal speed	Slow speed
P001	M	21	13m, 8s	4m, 16s	8m, 52s
P002	M	25	12m, 41s	4m, 45s	7m, 56s
P003	F	27	9m, 12s	4m, 24s	4m, 47s
P004	M	25	10m, 21s	5m, 1s	5m, 19s
P005	M	22	10m, 10s	3m, 56s	6m, 13s
P006	M	24	7m, 25s	3m, 18s	4m, 6s
P007	F	24	11m, 57s	4m, 18s	7m, 38s
P008	M	23	8m, 10s	3m, 33s	4m, 37s
P009	F	23	11m, 8s	4m, 32s	6m, 35s
P010	F	25	11m, 14s	4m, 40s	6m, 34s
P011	M	23	14m, 32s	5m, 33s	8m, 59s
P012	M	23	14m, 22s	4m, 2s	10m, 20s
P013	F	22	11m, 15s	3m, 22s	7m, 53s
P014	M	22	9m, 43s	4m, 21s	5m, 22s
P015	M	21	13m, 14s	5m, 32s	7m, 41s
Total	Ratio M/F	Mean age	Average times		
15	2:1	23.33 ± *1.61*	11m, 14s ± *2m, 3s*	4m, 22s ± *39.6s*	6m, 52s ± *1m, 45s*

Table 5.1: Data collection recording statistics.

and for a recovering patient.

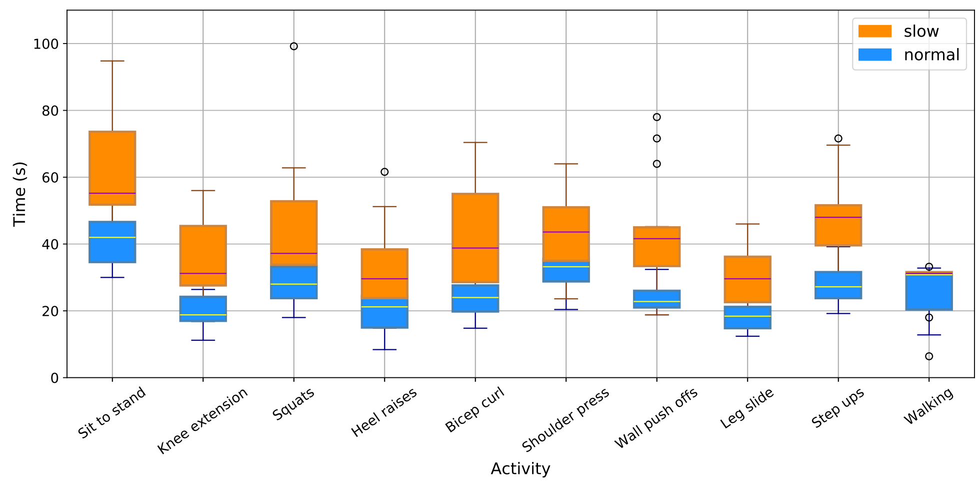


Figure 5.1: Average duration of each exercise, for the normal and slow speed.

5.2 Data exploration and preprocessing

The first step in developing a classification model was to visualise the collected data. For each of the 10 exercises, the accelerometer data was plotted in order to identify similarities and differences that might indicate where the feature engineering should begin. To combine the signals from all axes in a way which preserves the signal shapes, we applied PCA on the recordings of each subject and plotted the first principal component of the signal. Figure 5.2 shows a comparison between the resulting signals for the first 5 PR exercises, for both speeds: sit to stand, knee extension, squats, heel raises and bicep curls.

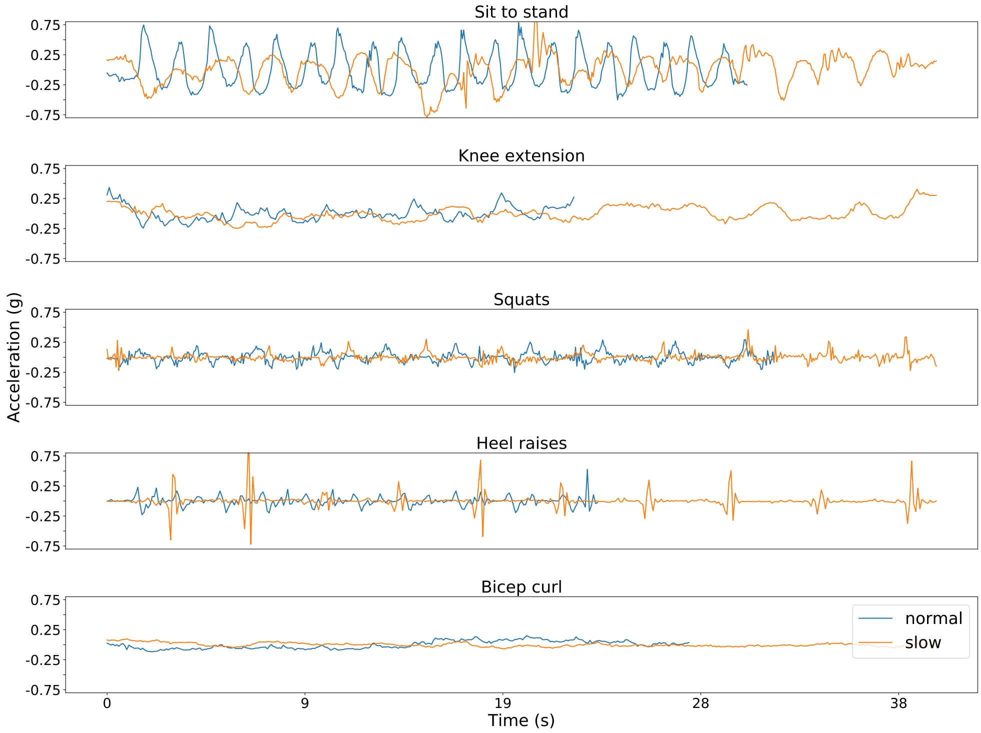


Figure 5.2: Comparative PCA curves of the first 5 PR exercises.

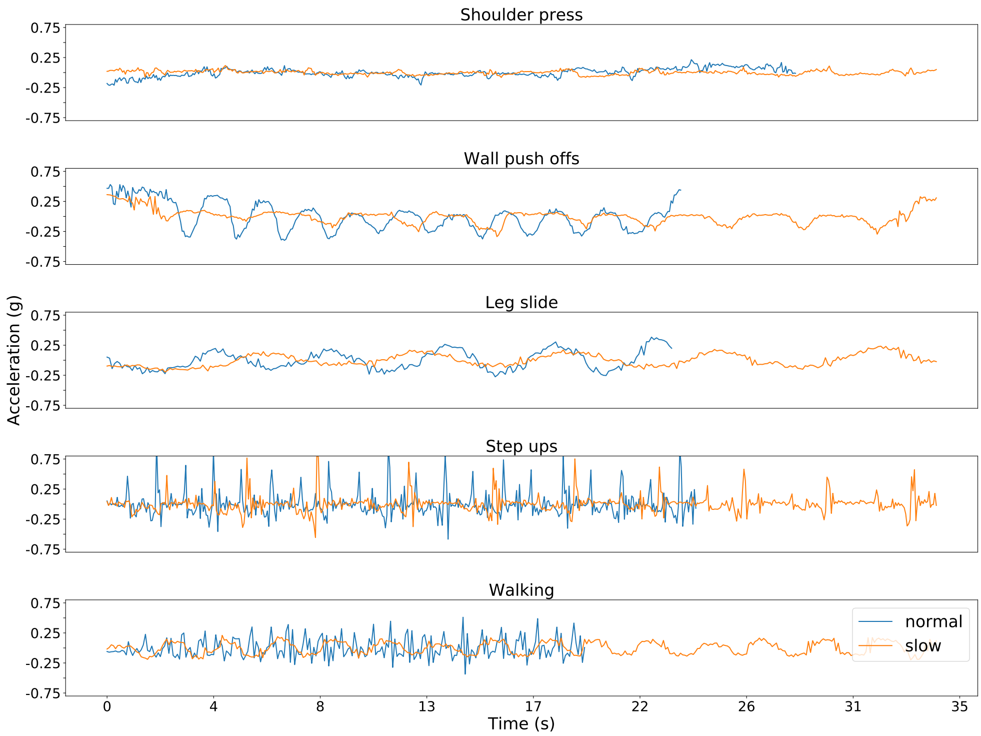


Figure 5.3: Comparative PCA curves of the last 5 PR exercises.

While *Sit to stand* and *Heel raises* present very distinctive patterns, the curves for *Bicep curls* and *Knee extensions* are rather flat. This can be explained by the fact that the chest of a subject is not moving much during these exercises. We can also observe that the slower speed exercises preserve the curve shapes of their normal speed counterparts, but the shapes are stretched along the x-axis. This was one of the first indications that a CNN might be appropriate for our task, since they are known to be able to identify shapes regardless of their location or skewness. Similarly, *Squats* present slight peaks and troughs in the signal, corresponding to the subject going up and down the wall.

The last 5 exercises are displayed in Figure 5.3. These are more dynamic exercises and, with the exception of the *Shoulder press*, which corresponds to a rather flat line, the other exercises have very distinctive shapes.

This suggests that a classification model might perform quite well on the last 5 exercises. An interesting aspect of the *Step ups* and *Walking* exercises was that the slower speed flattens the acceleration curves considerably, especially in the case of walking, there the peaks and the troughs are not present at all in the slow version. This imposes a challenge on accurately counting steps when a patient might be dragging their feet or walking at a very slow pace.

Next, we slice the datasets into sliding windows of variable sizes. The size of a sliding window is given by the number of samples in that window. For example, a window of size 25 contains 25 datapoints, which are sampled at 12.5Hz, therefore, the window will be 2 seconds long. Each window is treated as a 1D image, and the three accelerometer axes are treated as the image channels (analagous to RGB channels). Each window then becomes a datapoint for the classification model. Figure 5.4 shows this process.

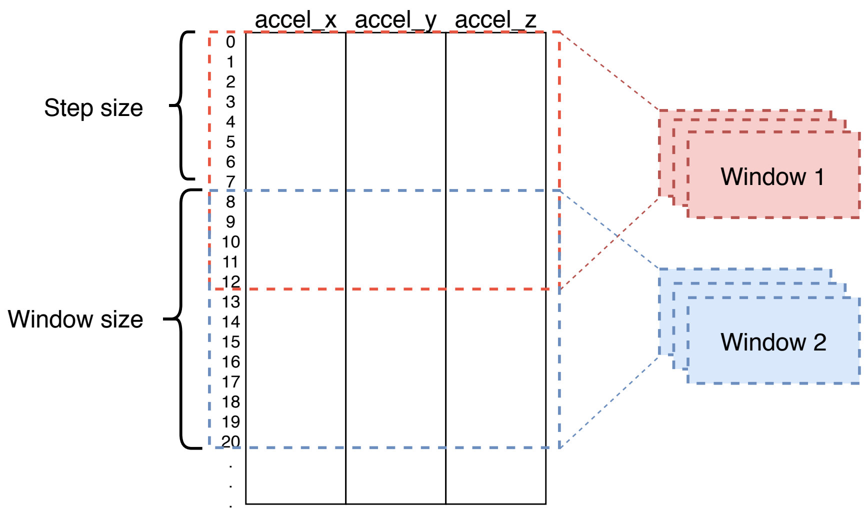


Figure 5.4: The process of turning raw data into sliding windows.

To address the intra-class variability, the dataset is partitioned by subject. Each model is then trained iteratively on 12 subjects, validated on 2 subjects and finally tested on 1 left-out subject, in a LOSOXV manner, as explained in Section 3.2.3. The training is complete once all the subjects have been left out once. We report the final classification accuracy as the averaged accuracy across all Left-Out-Subject folds.

5.3 Alternative models

One of the first techniques we considered at the beginning of the project was to develop a One-vs-Rest classifier for each one of the PR exercises. We would therefore train 10 separate classifiers, and each classifier would output one of two classes: positive, when the performed exercise is the expected one, and negative, whenever the data does not match the expected exercise.

We trained a 3-layer CNN, with the architecture shown in Figure 5.6, and with the following hyperparameters chosen experimentally: a window size of 38 with a step size of 19, normalised features, 128 filters with a kernel size of 3, a learning rate of 0.0001, a batch size of 32 and ReLU as the activation function. The best achieved accuracy was 92.8% however, when inspecting the average confusion matrix across all exercises, we noticed a high proportion of false positives, as shown in Figure 5.5a.

We pursued this option for a short time, in parallel with developing the overall classifier. The biggest issue when training these models was the huge class imbalance, which we attempted to correct by downsampling the data in the Negative class, as well as adding class weights for both the Positive and Negative classes, to encourage the learning algorithm to optimise for better f-scores. The best classification accuracy obtained with this model was 93.79% with an f-score of 80.09%. The final confusion matrix is shown in Figure 5.5b. The rate of false positives was still too high at 43%, so we decided to use an overall classifier instead.

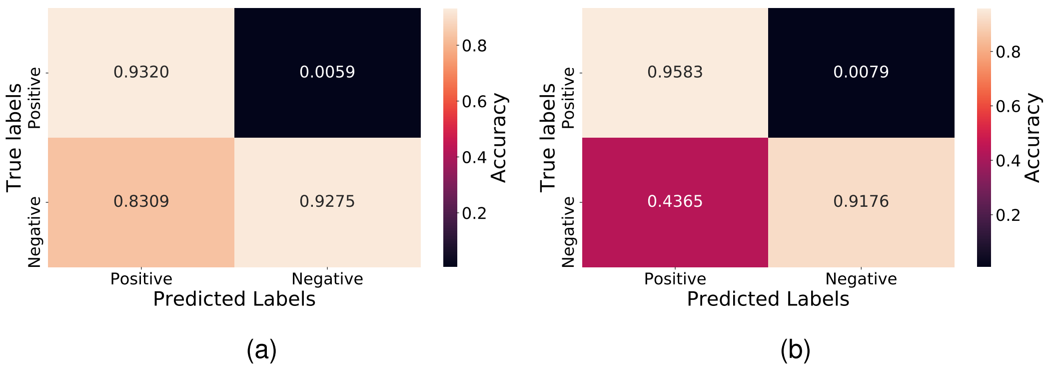


Figure 5.5: One-vs-Rest classifier averaged performance across all exercises, in its baseline form (a) and optimised form (b).

5.4 Baseline model

As explained in Section 1.1, last year's project focused on cough detection using Random Forest Classifiers, which achieved over 85% accuracy, but included handcrafting appropriate features to feed into the classifier. Moreover, implementing these features for live classification on an Android phone proves to be difficult, especially in the case of the PCA transformation of the signal.

Therefore, we began our experiments using a simple three-layer CNN, with the architecture presented in Figure 5.6.

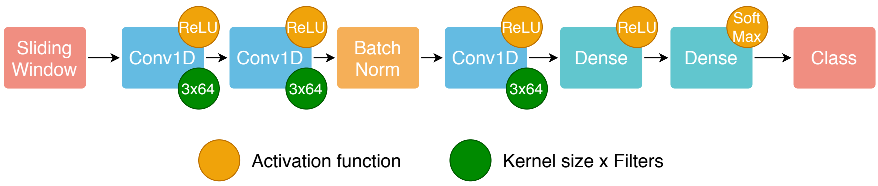


Figure 5.6: Baseline 3-layer CNN architecture.

We first generated the training dataset using the raw accelerometer features and observed that, although the training accuracy goes up to 54% in 250 epochs, the validation accuracy does not exceed 40%. Figure 5.7 shows the averaged confusion matrix for all LOSOXV folds. While the *Sit to stand* and *Wall push off* exercises seem to be mostly correctly classified, the other exercises are heavily misclassified. In particular, the *Shoulder press* is often confused with *Knee extensions* and *Bicep curls*, which probably stems from the fact that the accelerometer signals look very similar for these three exercises (see Figures 5.2 and 5.3). The final test accuracy was 39.01%.

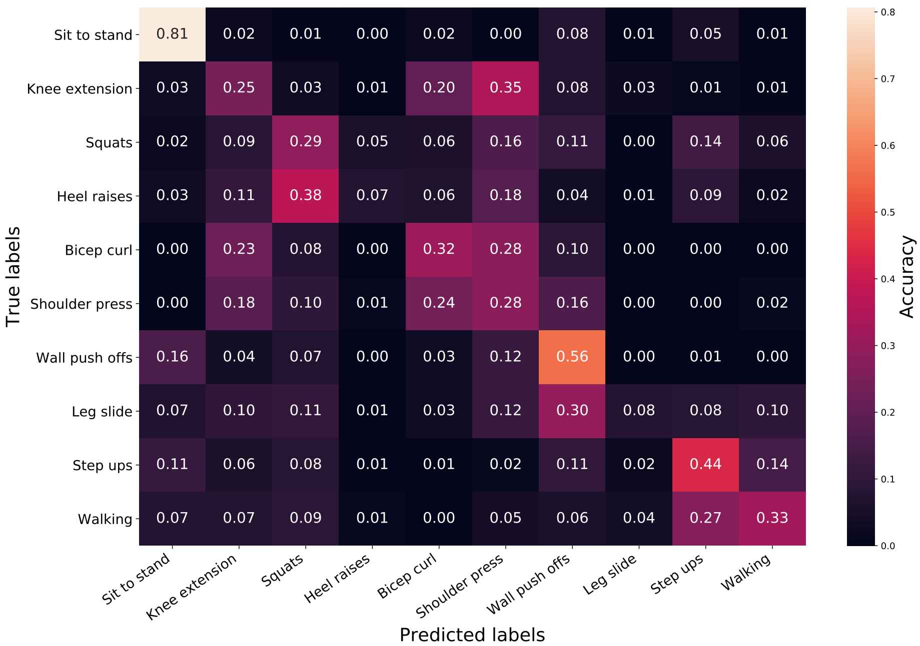


Figure 5.7: Confusion matrix for baseline experiment.

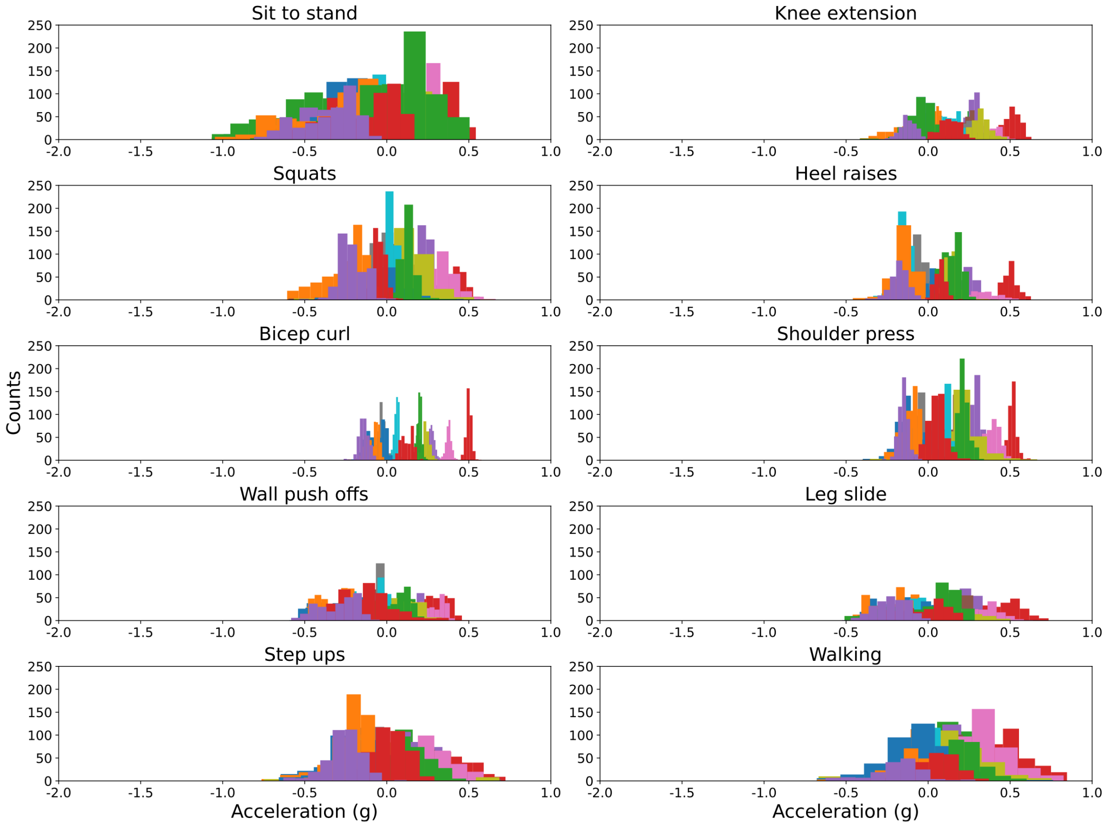
5.4.1 Standardising Features

When inspecting the distributions of the accelerometer values for each subject, we observed a lot of variation within a single class. As discussed in Section 3.2.2, this is what causes the inter-class variability, where the same activity might be performed very differently by different subjects. This is illustrated in Figure 5.8a, where we plotted the variability just for the $accel_x$ axis for all activities and subjects.

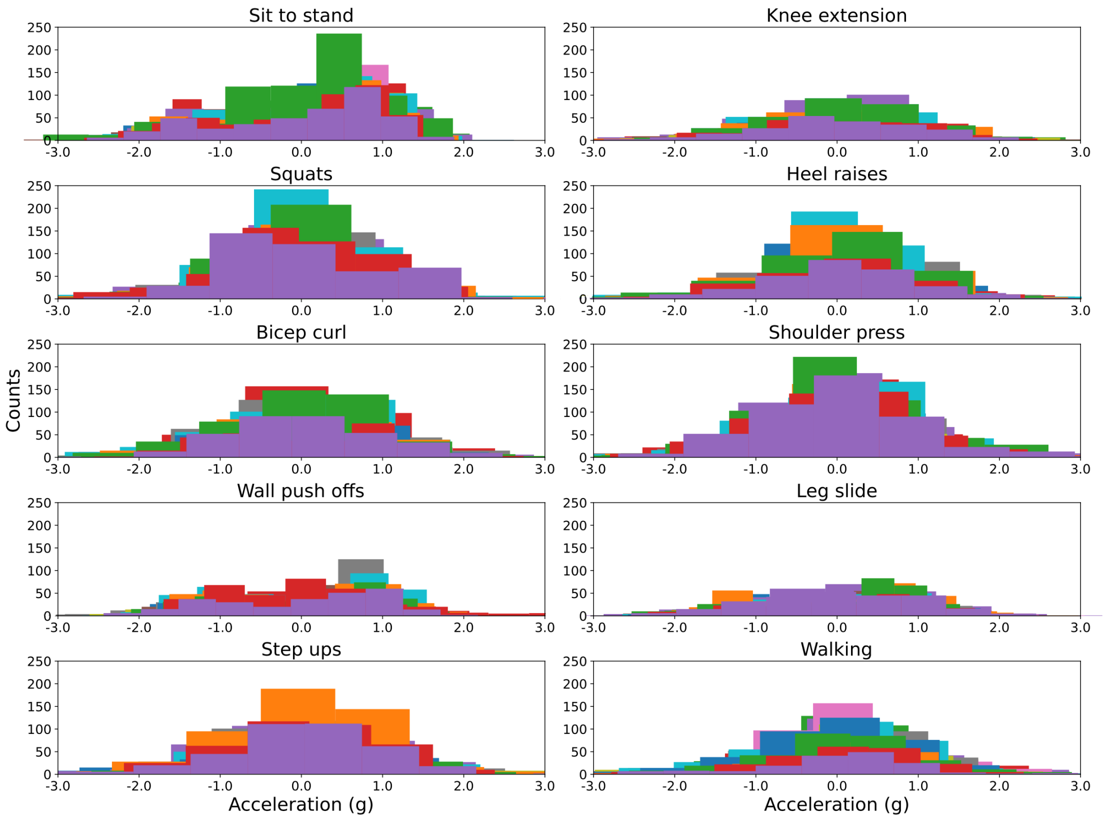
To tackle the intra-class variation problem, we standardised the data for each subject and for each activity, as shown in Figure 5.8b. Standardisation is achieved by bringing the mean to 0 and the standard deviation to 1:

$$X_{standardised} = \frac{X - \mu}{\sigma}$$

where μ is the mean of all samples in the window and σ is the standard deviation. The same model was trained on standardised features and resulted in an accuracy of 59.80% on the training set and 40.50% on the test set. This was a slightly higher generalisation performance than the former 39.01%, but it was a poor score nonetheless.



(a) Raw accelerometer axes distributions for the 10 exercises.



(b) Standardised accelerometer axes distributions for the 10 exercises.

Figure 5.8: Distributions of the $accel_x$ axis for all activities and subjects. Each differently coloured histogram represents the exercise recordings from a single subject.

5.4.2 Normalising Features

Given that neural networks typically work well with normalised data, we investigated the distribution of values for the $accel_x$, $accel_y$ and $accel_z$ axes for each of the exercises, for all subjects and speeds. This is shown in Figure 5.10a. We observed that each of the axes is contained in a roughly limited interval: $accel_x \in [-1,1]$, $accel_y \in [-2,-0.5]$ and $accel_z \in [-0.5,0.5]$. As explained in Section 3.3, differently scaled features might lead to a slow, zig-zagging gradient descent, which in turn slows the learning process of the model and hinders its ability to arrive at an optimal minimum in the loss function.

To improve the classification performance even further, we normalise the data separately for each activity, so that the potential model receives features from the same scale. To the best of the author's knowledge, no HAR paper has mentioned activity-specific normalisation. The application of this project is, however, under a very controlled circumstance - the user will select an exercise they are about to perform, and using this knowledge we can select the appropriate normaliser when it comes to classifiying live signals. Figure 5.10b shows the signal distributions after normalising to $[0,1]$.

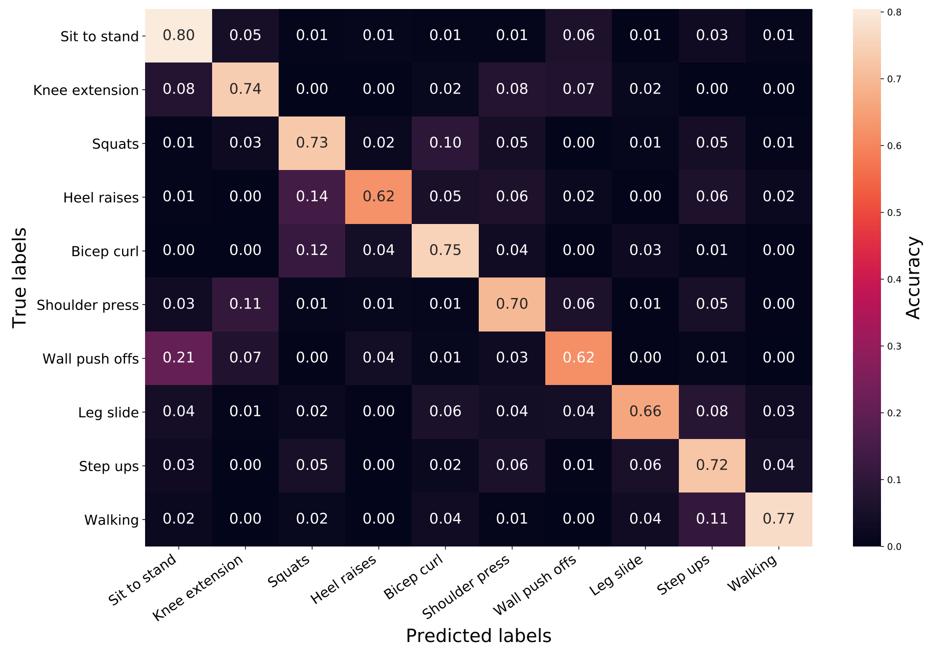
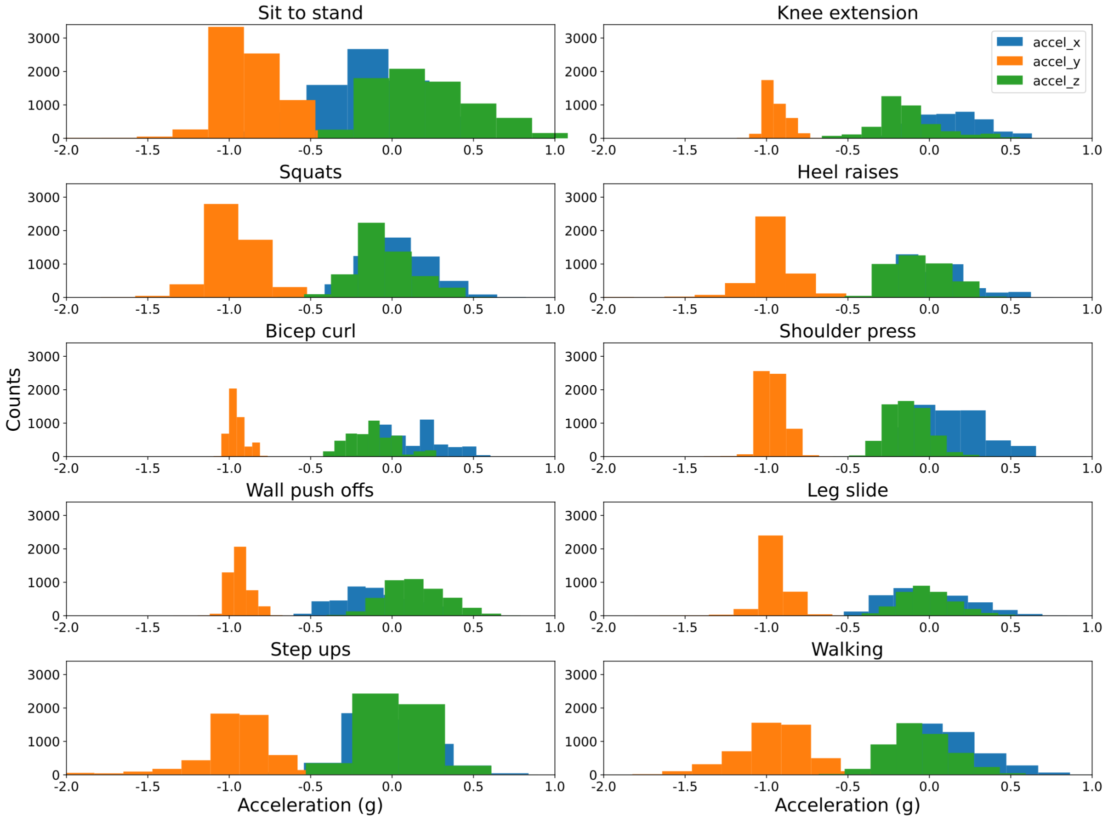
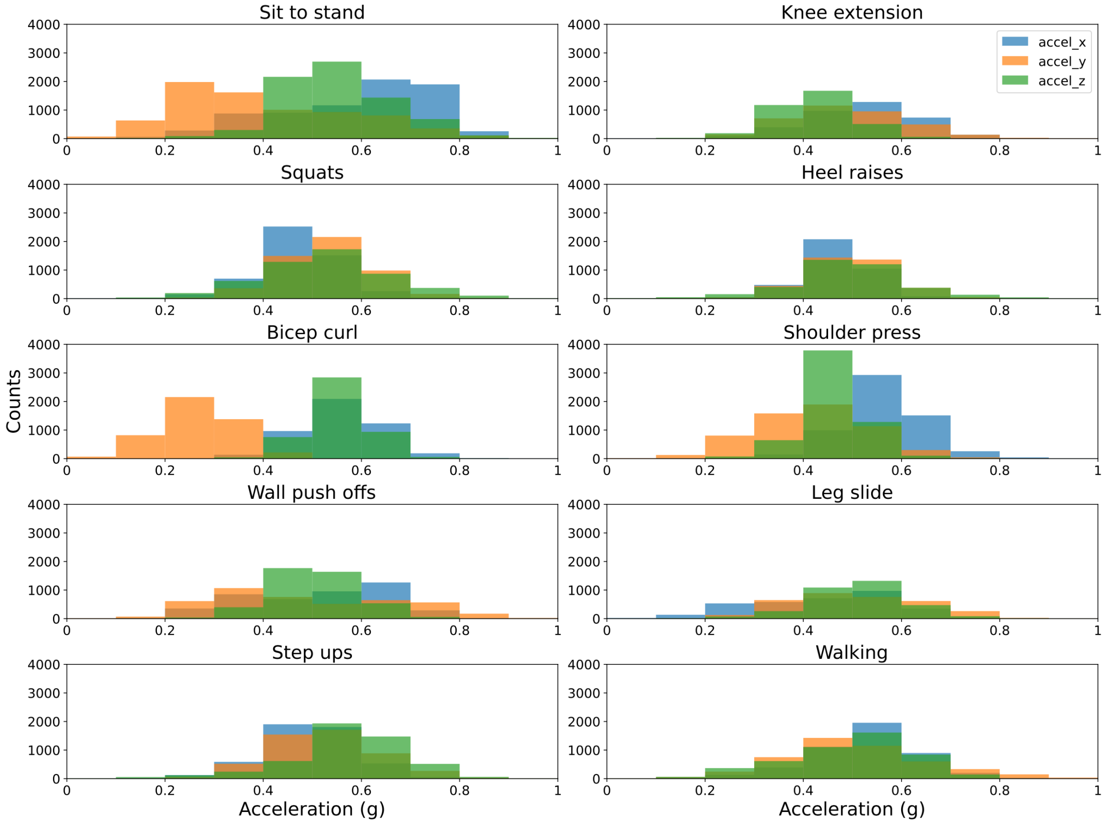


Figure 5.9: Confusion matrix for the baseline model with normalised features.

Normalising the $accel_x$, $accel_y$ and $accel_z$ axes resulted in a big improvement: 76.85% training accuracy and 71.65% testing accuracy, an average improvement of 30%. The confusion matrix is shown in Figure 5.9 and begins to be more concentrated over the main diagonal. Not only is the classification accuracy increasing, but so are the precision and recall.



(a) Raw accelerometer axes distributions for the 10 exercises.



(b) Normalised signal distributions for the 10 exercises.

Figure 5.10: Distributions for accelerometer signals, averaged for all subjects and shown for each exercise.

5.5 Hyperparameter Tuning

The next step in our experiments was to find out how the hyperparameters affect the classification accuracy.

5.5.1 Window and Step Size

First, we looked at the effect of the window and step size on the accuracy of the classifier. Since the live classification must work during the PR session itself, the window size cannot be larger than 20s (the minimum amount of time that a patient can spend on an exercise). Furthermore, in the future case where we might want to provide live feedback to the user, a smaller window size is more desirable. The step size dictates how much of the training data is replicated throughout the training examples. A small step size results in a large overlap, which means that much of the data is replicated and that might falsely result in a higher accuracy. On the other hand, a step size that is too small might result in not enough training data, so a balance between the two must be reached.

With the baseline architecture, we ran experiments using window sizes of 38, 63 and 50 datapoints (corresponding from 2.5s to 5s approximately). We kept the step size at a minimum of 19, and varied it for the values 19, 25 and 30. Experimentally, a window size of 63 with a step size of 50 yielded the best results for the baseline architecture. The results are presented in Figure 5.11 and contrasted to the first three baseline experiments.

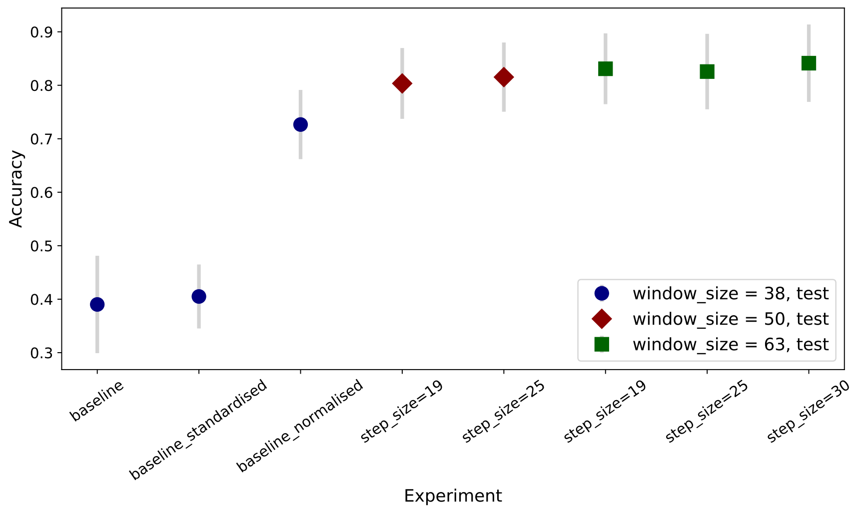


Figure 5.11: Experiment accuracies for varying window and step sizes.

5.5.2 Number of Filters and Batch Size

Next, we increased the number of filters and the batch size iteratively. The number of filters in a CNN correspond to the number of feature maps slid across each input to the network. In theory, the higher the number of maps, the larger the interpretational power of the network. A larger batch size allows the network to better estimate the gradient of the error function, and theoretically helps the network converge faster.

The numbers of filters tried were 64, 128 and 256, and the batch sizes were 32 and 64. A greater number of filters clearly improved the accuracy of the classifier, however, a larger batch size always resulted in a drop of at least 10% in accuracy. This could be because the training data was not sufficient, and one batch size of 64 was too large a chunk from the training set. Figure 5.12 compares the resulting accuracies and standard deviations after applying LOSOXV.

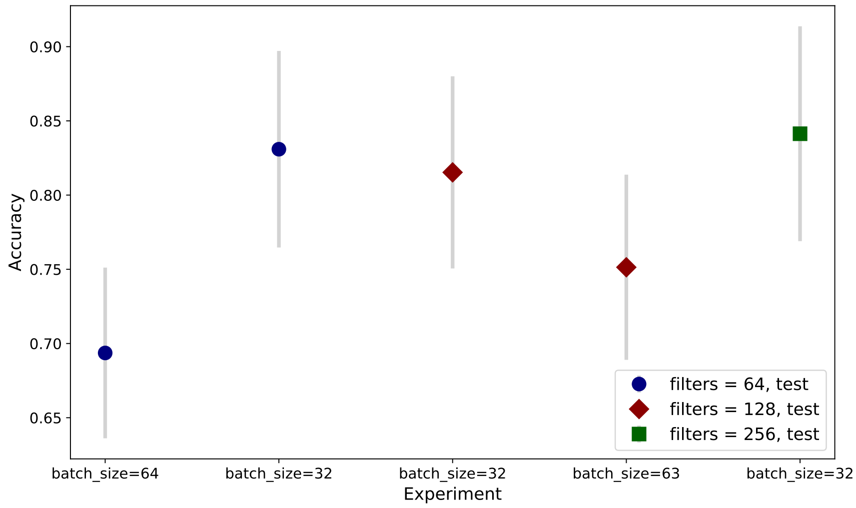


Figure 5.12: Experiment accuracies for varying the number of filters and the batch size.

To conclude, the best hyperparameters found were a batch size of 32, 256 filters, a window size of 63 and a step size of 30. All further experiments used these settings as their baseline.

5.6 Architecture Experiments

Next, we investigated the effect of the depth and complexity of the architecture on the classifier accuracy and speed of training. We built three variations of the baseline architecture:

- A deeper CNN, with 5 Convolutional layers
- A 5-layer CNN with 2 Dropout layers
- A 5-layer CNN with 2 Dropout layers and 3 MaxPooling layers

Each variation was first run with the best setting of the hyperparameters, as found in the previous section.

5.6.1 Deeper CNN

The deeper CNN architecture is shown in Figure 5.13. We added two additional convolutional layers, with a Batch Normalisation layer after each one of them, in order to prevent the input to each layer changing its distribution away from the normalised one.

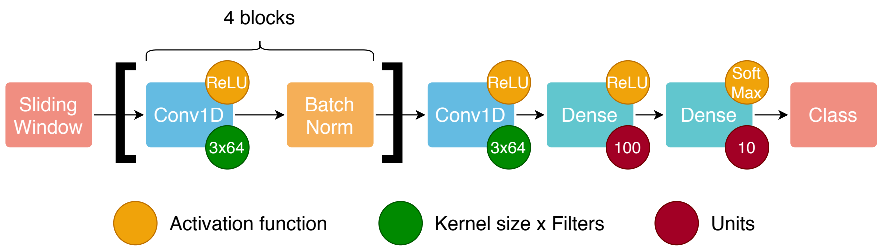


Figure 5.13: Deeper CNN architecture with 5 Convolutional layers.

We again varied the window size and the number of filters to see the rate at which they influence the accuracy improvement. As expected, the largest window size and the largest amount of filters result in an average accuracy of 86.62% across all activities.

5.6.2 Dropout

The architecture for the following set of experiments was constructed by adding two Dropout layers between the Convolutional layers, with a dropout rate of 0.2. Dropout layers help improve the generalisation performance of the model by randomly cutting off connections with neurons from the previous layer with a certain probability (dropout rate). Figure 5.14 shows the structure of this architecture.

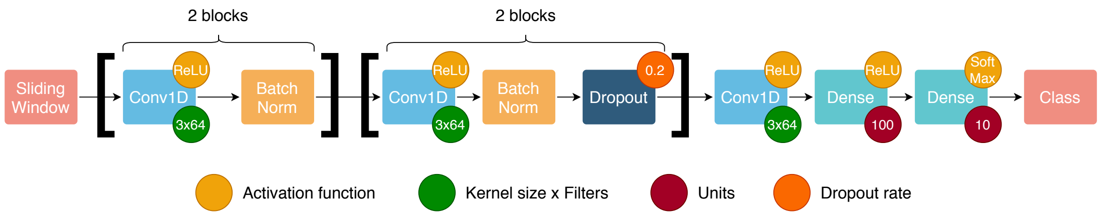


Figure 5.14: Deeper CNN architecture with two Dropout layers.

Adding two dropout layers to our architecture and using the highest performing hyperparameters (window size of 75, step size of 25 and 256 filters), achieved an increase in accuracy by 3.1% to 89.65%. We tried increasing the window size again, to 125, and the step size to 30. Additionally, we changed the optimizer from SGD to Adam, and set the learning rate to 1e-5. This resulted in the best achieving accuracy across all our experiments: 92.83% after LOSOXV.

5.6.3 Pooling

Finally, we added pooling layers after each convolutional layer, to reduce the complexity of the network and attempt to improve its generalisation performance even further. The architecture is shown in Figure 5.15.

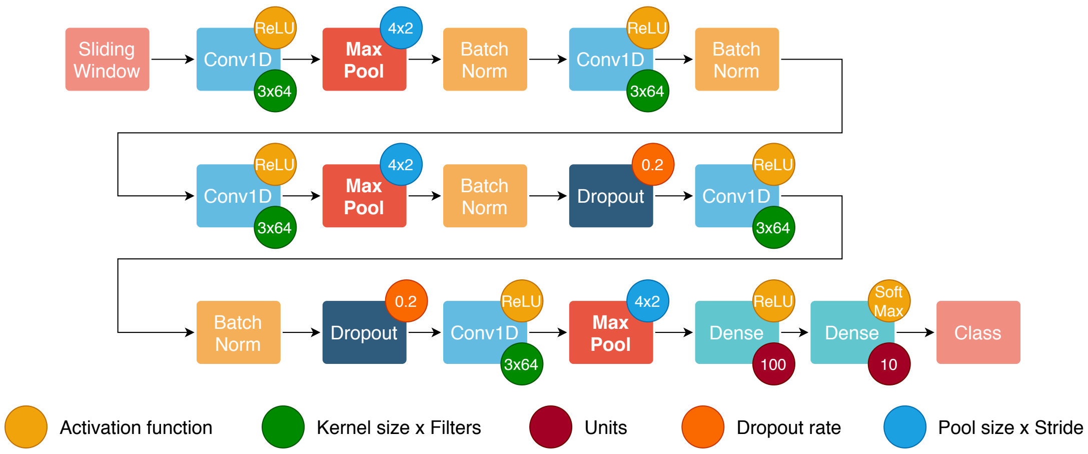


Figure 5.15: Deeper CNN architecture with MaxPool layers.

Unfortunately, pooling layers did not seem to help with our task, reducing the test accuracy to a surprising 87.78%. We suggest this happened because, given the window size was only 125 points long, reducing its dimensions by a factor of 2 thrice along the CNN structure leaves very little room for prediction accuracy, with around 16 datapoint remaining for the final two layers of the CNN. Pooling layers might be a good addition to data with large dimensions, but does not help in the task at hand.

5.7 Best Performing Model

Our experiments concluded that the best performing model was a CNN with 5 convolutional layers with a batch normalisation layer after each one of them and two dropout layers with a dropout rate of 0.2. The best hyperparameter settings were a window size of 125 and a step size of 30, 256 filters and a batch size of 32. The best performing optimiser was Adam, bringing the accuracy up to 92.83% after 250 epochs. The model confusion matrix is shown in Figure 5.16.

5.8 Step counting

As explained in Section 3.1.3, this part of the algorithm was inspired by Sorribas' master thesis [58]. We use a two-level hierarchical classification model which allows us to only count steps if the detected activity is walking, and apply the correct algorithm parameters depending on the type of walk - fast, slow or shuffling.

The hierarchical model is illustrated in Figure 5.17. First, we extract the predicted activity from a window of data. If the predicted activity is walking, we deploy a second-level pre-trained classification model which specifies whether the performed walk is fast, slow or shuffling. Finally, depending on the type of walk detected, we apply a peak detection algorithm with varying hyperparameters such as a threshold and a minimum peak distance.

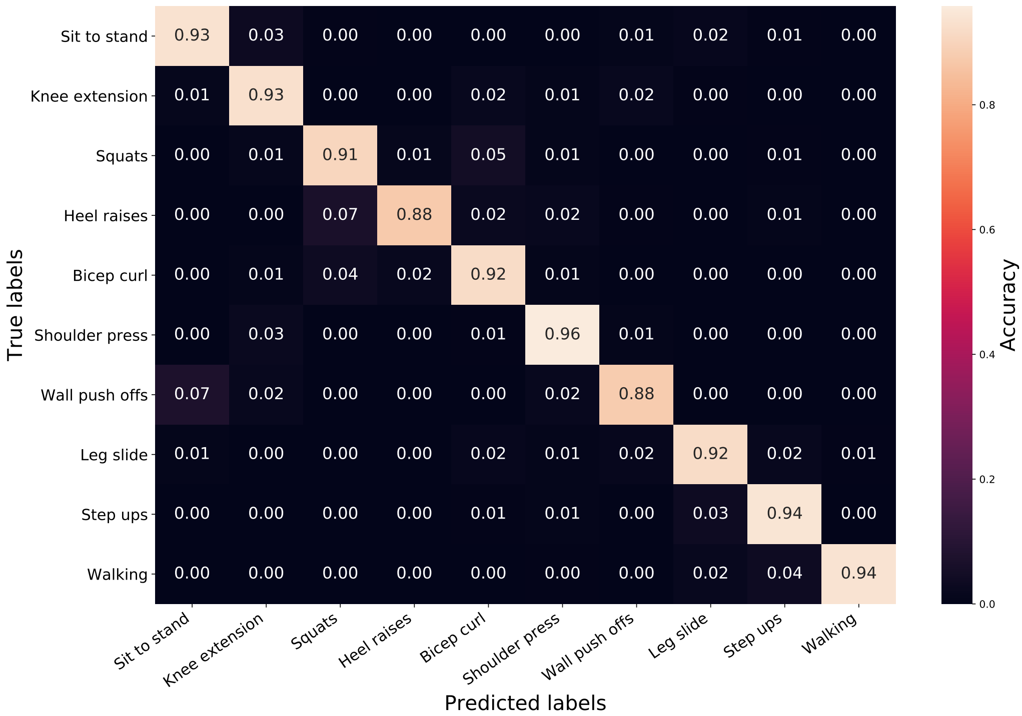


Figure 5.16: Best performing model confusion matrix.

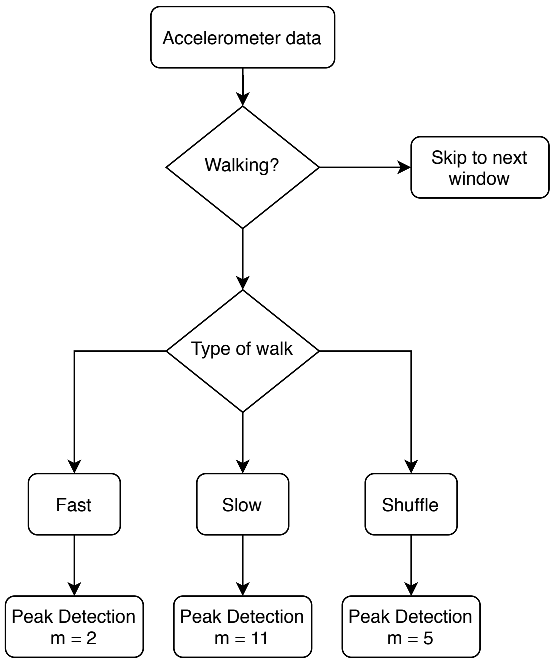


Figure 5.17: Hierarchical model for step counting.

We collect 1-minute windows of data for this particular exercise, then classify the window using the model described above. The peak detection algorithm first smooths the vertical component of the acceleration (the y-axis in the case of RESpeck), using a simple moving average smoothing, with a window size of 10. We then count the number of peaks within the 1-minute time window that are at least m peaks away from the previously detected peak, where:

- $m = 2$ for fast walking
- $m = 11$ for slow walking
- $m = 5$ for shuffling

The total number of peaks gives the number of steps taken during the timeframe. Figure 5.18 shows the three stages of processing the data.

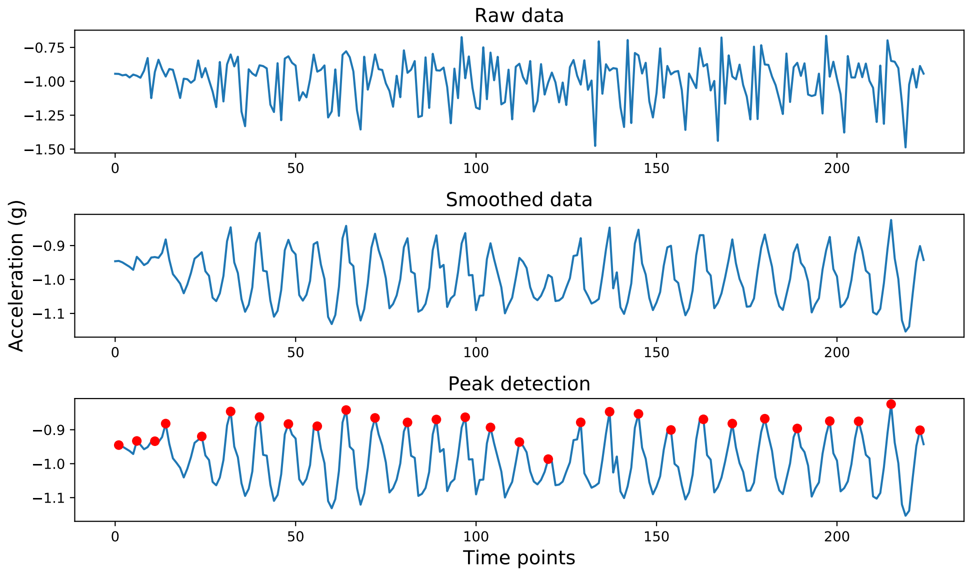


Figure 5.18: The three stages of the step counting algorithm.

Sorribas' reported accuracy was 86.8%. Due to the COVID-19 crisis, we could only implement his model offline (as opposed to implementing it in the Android app), and ran it on data that had already been collected. Furthermore, when validating the system performance we compared the performance of this model with other step counters on the market and observed a very similar accuracy. Further details are given in the following chapter.

Chapter 6

System Validation

This chapter describes the implementation of live classification directly on the Android phone, as well as the experiments run to validate the real system performance, including the usability of the dashboard functionalities.

6.1 Live classification

After finding the best configurations for the classifier, we trained it on all available data (without leaving any subjects out) and exported the final model in `.h5` format, which is the Tensorflow version of mobile-run classification algorithms. We then imported this into our Android application. To achieve the same normalisation effect, we exported the minimum and maximum values for each activity across all subjects. During live classification, we can normalise each window by applying the following linear scaling formula:

$$X_{normalised} = \frac{X - X^i_{min}}{X^i_{max} - X^i_{min}}$$

where i is the index for each type of PR exercise.

To run the live classification, once the user starts the PR session, we start aggregating accelerometer data into windows of the specified size (125) and step (30). We then standardise each window, bringing it to 0 mean and unit variance, and normalise it using the formula above. Note that each window of data is normalised twice, using parameters for both normal speeds and slow speeds. Next, we pass the two normalised windows of data (one for normal speed and one for slow speed) through the classifier and, if at least one of the classifier outputs matches the currently selected PR exercise, we record the window as being correctly executed. If the output does not match, we assign a correctness of 0 to the current window. This way of double-classifying a window ensures that our algorithm can identify movements and every type of speed. At the end of the PR exercise, the total correctness score is calculated as the proportion of sliding windows classified as the selected exercise. Figure 6.1 illustrates this process.

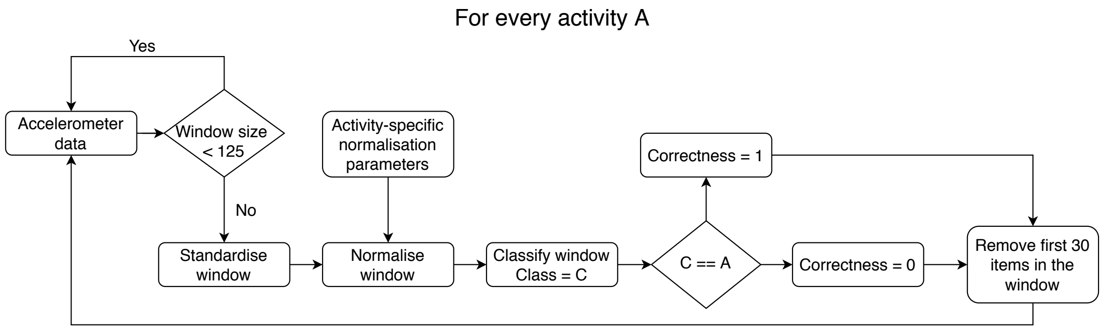


Figure 6.1: Live classification algorithm for PR exercises.

6.2 Rehab3 Classification Performance

Due to the unfortunate circumstances of COVID-19, our initial plan of testing the system extensively on at least 5 new subjects were delayed. The strict rules of social distancing meant that we could not have contact with other people than the author's flatmates. The live classification system was therefore tested on the author's 3 flatmates. Given that they were also part of the initial data collection, a different model was trained for each one of them. For each subject, the model was trained on all available data except the data collected from that subject themselves. That model was then uploaded onto the Android application and used for testing.

Each participant was asked to undertake two separate PR sessions, and perform the exercises at the two different speeds: normal and slow.

For each type of session, we looked at the rate of agreement between the live classification, offline classification and human annotator. The results are compared window by window, as shown in Figure 6.2. For example, we can see that the live classifier using the normalisation for slow exercises is correctly picking up all movements, while the real-time classifier using the fast normalisation data confuses *Heel raises* with *Sit to stand* in some instances.

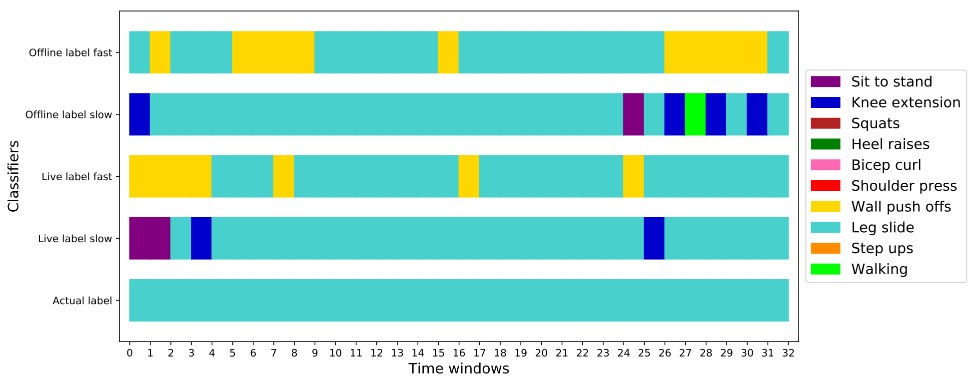


Figure 6.2: Example pixelgram from a performed PR exercise, comparing the predictions of the two classifiers (real-time and offline), using the two recorded speeds as normalisation factors.

The accuracy of each type of classifier is averaged for the entire session. The complete results are displayed in Table 6.1, broken down by subject, speed and classifier.

Normal speed				
	Live classification		Offline classification	
Participant	Normal	Slow	Normal	Slow
1	87.89%±*0.17*	**98.75%±*0.03***	82.25%±*0.18*	95.25%±*0.06*
2	84.02%±*0.25*	**90.25%±*0.27***	76.54%±*0.31*	89.58%±*0.25*
3	86.42%±*2.43*	**93.57%±*2.03***	74.28%±*3.33*	90.71%±*2.28*
Average accuracy	86.11%	**94.19%**	77.69%	91.84%
Slow speed				
	Live classification		Offline classification	
Participant	Normal	Slow	Normal	Slow
1	78.68%±*2.66*	**97.76%±*0.52***	62.04%±*3.44*	92.09%±*1.56*
2	82.74%±*0.19*	**95.93%±*0.11***	64.95%±*0.36*	85.62%±*0.20*
3	78.51%±*2.47*	**92.69%±*1.75***	69.16%±*28.68*	88.71%±*1.48*
Average accuracy	79.97%	**95.46%**	65.38%	88.80%

Table 6.1: Live classifier performance vs. offline classifier performance, for three different types of PR sessions.

In addition, we compare the performance for each exercise in Figure 6.3. We notice that *Heel raises* and *Leg slides* cause the most confusion for the classifiers, as they are the most often misclassified. After further investigation, we observed that *Leg slides* are most often misclassified as *Bicep curls* which can be explained by the lack of movement of the chest between the two exercises. The big difference between the performance of the live classifier and the offline classifier observed in the training phase of this project might stem from the different approach we take to the standardisation on inputs. In the training phase, we standardised the data for each subject and activity, whereas in the live classifier, we standardise each window of data.

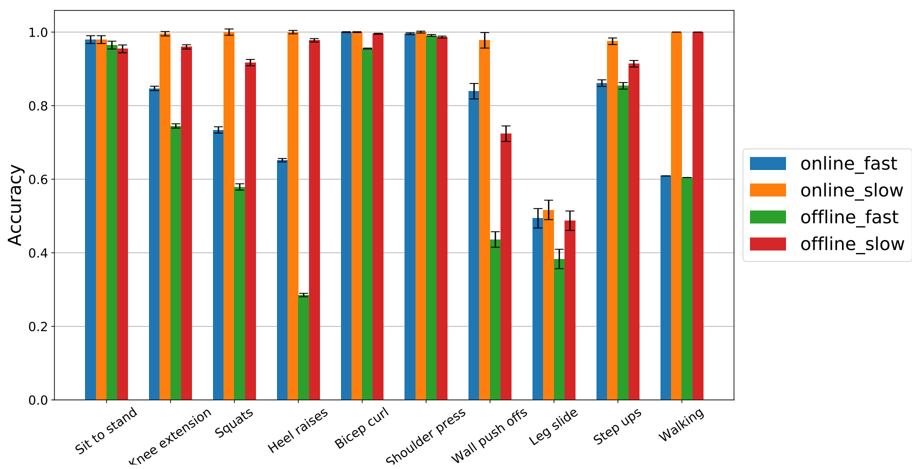


Figure 6.3: Comparative accuracies for the four classifiers, broken down by exercise.

6.3 6-minute walk test

The step counting algorithm was only developed for offline use, due to the difficulties introduced by COVID-19, but was nonetheless tested with data gathered during the validation phase. The subjects were asked to perform a 1-minute walk 3 times:

- One walk at a normal pace,
- One walk at a slow pace,
- One walk by dragging their feet, i.e. shuffling.

During the data recording, the subjects were also wearing a Fitbit Charge 3 [25] on their left wrist and an iPhone in their right pocket, with the Pedometer [9] application turned on. Additionally, both the author and the participant manually counted the steps taken during the 1-minute walk. Table 6.2 summarises all 9 trials, 3 for each of the 3 subjects, and compares the numbers obtained by applying the appropriate step counting algorithm with the other three manual measures.

		Normal Walk			
Participant	Manual count	Fitbit	Pedometer	Peak Detection	Error
1	89	90	89	90	+1 step
2	111	115	121	112	+1 step
3	89	90	92	91	+2 steps
		Slow Walk			
Participant	Manual count	Fitbit	Pedometer	Peak Detection	Error
1	32	8	14	39	+7 steps
2	76	91	110	76	0 steps
3	56	22	29	55	-1 step
		Shuffling			
Participant	Manual count	Fitbit	Pedometer	Peak Detection	Error
1	44	0	0	49	+5 steps
2	73	0	68	68	-5 steps
3	55	0	0	51	-4 steps

Table 6.2: Normal and slow walk and shuffling data, as recorded by a human annotator, a Fitbit, a Pedometer application and the Peak Detection algorithm. The error between the Peak Detection algorithm and the manually annotated steps is shown.

While the Peak Detection algorithm has an average error of 2.88 steps across all types of walking, it almost always outperforms other commercial step counters. This is especially in the case of *Slow walking* and *Shuffling*, where other step counters underestimate the number of steps or do not even detect them at all. The latter can be observed in the recordings of Participants 1 and 3, when performing *Shuffling*.

Chapter 7

Conclusion and Future Work

In this report, we presented a working system which helps COPD patients conduct PR at home. The system uses the wearable RESpeck monitor to collect accelerometer data, identifies correctly performed exercises and logs the patient's breathing rate in the resting periods between exercises, through the *Rehab3* app. In addition, patients are provided with an additional application called *Rehab Diary*, where they can fill in a daily questionnaire about their perceived well-being. The exercise session summaries and diary entries are sent to a secure Cloud hosted database and shown to trained medical staff through an interactive *Dashboard*.

7.1 Results

We collected PR data from 15 volunteers and trained a Convolutional Neural Network to recognise each of the ten exercises, with a final average accuracy of 92.83%. We then exported this model onto the *Rehab3* app, to perform real-time classification of the exercises. Due to unfortunate international circumstances of COVID-19, the system was not widely used on real patients. Instead, it was validated using 3 volunteers, who performed a series of PR sessions in the author's presence. The manually annotated activities were compared to the automatically labeled ones, resulting in an overall accuracy of 94.19% for the live classifier.

The 6MWT was validated against popular step tracking devices and applications and the results showed that the Peak Detection algorithm clearly outperforms other commercial devices, especially in the case of slow walking and shuffling, which is what recovering patients are likely to perform rather than walking.

7.2 Strengths and weaknesses

This project presented a novel approach to the prevailing problem of low PR uptake and completion rates. It provides an entirely automated physical exercise guidance for the patient, and collects meaningful data summarised for medical staff, who can monitor the patients remotely and track their exercise performance and breathing rate responses. To the best of the author's knowledge, there is no other existing system that performs automated activity recognition of the PR exercises.

Not only does the system provide PR guidance, but also it is a continuous patient monitoring system, keeping track of the patient's average breathing rate and activity level throughout the day. As detailed in Chapter 1, a sudden decrease in the level of activity is a very good indicator of an upcoming COPD exacerbation, so having such a system in place could aid in the early detection of the onset of symptoms. All of this information is gathered from the patient automatically and summarised for the medical staff within minutes: the patient is not required to be in the presence of their physiotherapist to relate their symptoms and the bias introduced by giving them questionnaires is also removed.

The biggest weakness of this project was that, due to limited time and resources, the training data for the HAR model was collected from volunteers with entirely different characteristics to actual potential users of the *Rehab3* app. They differed in age and physical conditioning, which could be one of the reasons for which the system might underperform in the future. A robust data collection process would have included subjects from a more diverse demographic. However, his would only have been feasible given the appropriate time and resources, since older participants or real COPD patients are difficult to enrol in such studies. To compensate for this bias in the data, the volunteers for this project were asked to perform each exercise twice, at different speeds, to account for the fact that recovering patients might complete the exercises at a slower pace.

Another weakness of the system and a point for future improvement is the lack of user studies with regards to the user interface of the *Rehab3* and *Rehab Diary* applications. Though the user interfaces were inspired by older, similar projects, there was no inquiry as to whether older or more unexperienced users find the interaction with the apps useful or logical. During the design phase, we did take into account well established rules of interface design for our target users, such as large fonts and simple, uncluttered screens. However, this was just an estimation and future problems regarding the usability of the application are to be expected.

7.3 Future work

Future work on this project should aim to firstly extend the data collection to a wider range of participants, aiming for volunteers that are particularly similar to the target audience of this system. This could give true insights of which exercises are performed correctly and whether there are any exercises in particular that all participants struggle with. This could also lead to an improvement of the instructions shown to the user upon app startup. Moreover, having a calibration period for each individual patient, especially in the presence of a medical advisor, would highly improve the performance of the classification model. To further strengthen the classification model, the *NULL class* should be included in the training process. Example elements of the *NULL class* could be random noise or altered signals.

Secondly, a good future addition to the project would be conducting user studies on the usability of the current interface and trying to either improve it, or query participants of all ages and technical knowledge about features they find useful/not useful. The timing functionality on the *Rehab3* app could also be adapted to the speed of each patient, by including another speed detection mechanism alongside the exercise classifier.

Last but not least, further developments can be brought to the *Dashboard*. Without cluttering the workspace unnecessarily, the dashboard menu can present the medical staff with a wider range of viewing options - for example, viewing the summaries data from all patients at once, so that they can easily pick out the problematic cases. This would require a refactoring of the dashboard implementation, but offering more flexibility would increase the user productivity on the website and could result in very early detection of the onset of an exacerbation.

Bibliography

[1] Wytske A Altenburg, Nick HT ten Hacken, Linda Bossenbroek, Huib AM Kerstjens, Mathieu HG de Greef, and Johan B Wempe. Short-and long-term effects of a physical activity counselling programme in copd: a randomized controlled trial. *Respiratory medicine*, 109(1):112–121, 2015.

[2] R Graham Barr, Bartolome R Celli, Fernando J Martinez, Andrew L Ries, Stephen I Rennard, John J Reilly Jr, Frank C Sciurba, Byron M Thomashow, and Robert A Wise. Physician and patient perceptions in copd: the copd resource network needs assessment survey. *The American journal of medicine*, 118(12):1415–e9, 2005.

[3] Andrew Bates, Martin J Ling, Janek Mann, and Damal K Arvind. Respiratory rate and flow waveform estimation from tri-axial accelerometer data. In *2010 International Conference on Body Sensor Networks*, pages 144–150. IEEE, 2010.

[4] Charlotte E Bolton, Elaine F Bevan-Smith, John D Blakey, Patrick Crowe, Sarah L Elkin, Rachel Garrod, Neil J Greening, Karen Heslop, James H Hull, William DC Man, et al. British thoracic society guideline on pulmonary rehabilitation in adults: accredited by nice. *Thorax*, 68(Suppl 2):ii1–ii30, 2013.

[5] Andreas Bulling, Ulf Blanke, and Bernt Schiele. A tutorial on human activity recognition using body-worn inertial sensors. *ACM Computing Surveys (CSUR)*, 46(3):1–33, 2014.

[6] Naoto Burioka, Akari Imada, Akiko Kiyohiro, Fumika Sugitani, Takenori Fujii, Akari Hosaka, Sachiko Nakamoto, Takashi Amisaki, and Eiji Shimizu. Modified six-minute walk test: number of steps per second. *Yonago acta medica*, 57(1):61, 2014.

[7] Bartolome R Celli, WATS MacNee, AATS Agusti, Antonio Anzueto, B Berg, A Sonia Buist, Peter MA Calverley, N Chavannes, T Dillard, B Fahy, et al. Standards for the diagnosis and treatment of patients with copd: a summary of the ats/ers position paper. *European Respiratory Journal*, 23(6):932–946, 2004.

[8] Peter H Charlton, Timothy Bonnici, Lionel Tarassenko, David A Clifton, Richard Beale, and Peter J Watkinson. An assessment of algorithms to estimate respiratory rate from the electrocardiogram and photoplethysmogram. *Physiological measurement*, 37(4):610, 2016.

[9] Cross Forward Consulting. Pedometer++. `https://pedometer.app`, 2019. Accessed: 2020-03-31.

[10] CG Cote and BR Celli. Pulmonary rehabilitation and the bode index in copd. *European Respiratory Journal*, 26(4):630–636, 2005.

[11] Murat Darçın. Association between air quality and quality of life. *Environmental Science and Pollution Research*, 21(3):1954–1959, 2014.

[12] Eric Derom, Chris Van Weel, Giuseppe Liistro, Johan Buffels, Tjard Schermer, Ernst Lammers, Emiel Wouters, and Marc Decramer. Primary care spirometry. *European Respiratory Journal*, 31(1):197–203, 2008.

[13] Carol DerSarkissian. Smart exercises for people with copd. `https://www.webmd.com/lung/copd/ss/slideshow-copd-exercises`, 2018. Accessed: 2020-03-07.

[14] Ch Domingo, E Canturri, M Luján, A Moreno, H Espuelas, and A Marín. Transcutaneous measurement of partial pressure of carbon dioxide and oxygen saturation: validation of the sentec monitor. *Archivos de Bronconeumología ((English Edition))*, 42(5):246–251, 2006.

[15] Jeremy A Falk, Steven Kadiev, Gerard J Criner, Steven M Scharf, Omar A Minai, and Philip Diaz. Cardiac disease in chronic obstructive pulmonary disease. *Proceedings of the American thoracic society*, 5(4):543–548, 2008.

[16] Darius Fischer. Predicting the well-being of copd patients with respiratory data from pulmonary rehabilitation. unpublished, 2016.

[17] British Lung Foundation. Your exercise handbook - for poeple living with a lung condition. `https://cdn.shopify.com/s/files/1/0221/4446/files/EH1_Exercise_Handbook_2018_Final_Artwork_PRINT_-_sample.pdf?11900238372730823857`, 2018. Accessed: 2020-03-28.

[18] British Lung Foundation. Pulmonary rehabilitation or pr. `https://www.blf.org.uk/support-for-you/keep-active/pulmonary-rehabilitation`, 2019. Accessed: 2020-03-07.

[19] COPD Foundation. What is copd? `https://www.copdfoundation.org/What-is-COPD/Understanding-COPD/What-is-COPD.aspx`. Accessed: 2020-01-25.

[20] Kunihiko Fukushima. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological cybernetics*, 36(4):193–202, 1980.

[21] R Garrod, J Marshall, E Barley, and PW Jones. Predictors of success and failure in pulmonary rehabilitation. *European Respiratory Journal*, 27(4):788–794, 2006.

[22] RJ Halbert, JL Natoli, Anacleto Gano, E Badamgarav, A Sonia Buist, and DM Mannino. Global burden of copd: systematic review and meta-analysis. *European Respiratory Journal*, 28(3):523–532, 2006.

[23] Healthline. Copd exacerbation. https://www.healthline.com/health/copd/exacerbation-symptoms-and-warning-signs. Accessed: 2020-01-26.

[24] Liang-Hui Huang and Yu-Ju Chen. The 6-minute walk test to assess exercise capacity of patients with chronic obstructive pulmonary disease, 2016.

[25] Fitbit Inc. Fitbit charge 3 user manual. https://staticcs.fitbit.com/content/assets/help/manuals/manual_charge_3_en_US.pdf, 2019. Accessed: 2020-03-31.

[26] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift, 2015.

[27] Sarah E Jones, Stuart A Green, Amy L Clark, Mandy J Dickson, Ann-Marie Nolan, Clare Moloney, Samantha SC Kon, Faisal Kamal, Joy Godden, Cathy Howe, et al. Pulmonary rehabilitation following hospitalisation for acute exacerbation of copd: referrals, uptake and adherence. *Thorax*, 69(2):181–182, 2014.

[28] Ashish K Khanna, Phillip Hoppe, and Bernd Saugel. Automated continuous noninvasive ward monitoring: future directions and challenges. *Critical Care*, 23(1):194, 2019.

[29] Eric Kipnis, Davinder Ramsingh, Maneesh Bhargava, Erhan Dincer, Maxime Cannesson, Alain Broccard, Benoit Vallet, Karim Bendjelid, and Ronan Thibault. Monitoring in the intensive care. *Critical care research and practice*, 2012, 2012.

[30] Ron Kohavi et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*, volume 14, pages 1137–1145. Montreal, Canada, 1995.

[31] Samantha SC Kon, Deniz Dilaver, Manvi Mittal, Claire M Nolan, Amy L Clark, Jane L Canavan, Sarah E Jones, Michael I Polkey, and William DC Man. The clinical copd questionnaire: response to pulmonary rehabilitation and minimal clinically important difference. *Thorax*, 69(9):793–798, 2014.

[32] Yann LeCun, Léon Bottou, Yoshua Bengio, Patrick Haffner, et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[33] Yann A LeCun, Léon Bottou, Genevieve B Orr, and Klaus-Robert Müller. Efficient backprop. In *Neural networks: Tricks of the trade*, pages 9–48. Springer, 2012.

[34] Li Liu, Yuxin Peng, Shu Wang, Ming Liu, and Zigang Huang. Complex activity recognition using time series pattern dictionary learned from ubiquitous sensors. *Information Sciences*, 340:41–57, 2016.

[35] Lakmini P Malasinghe, Naeem Ramzan, and Keshav Dahal. Remote patient monitoring: a comprehensive study. *Journal of Ambient Intelligence and Humanized Computing*, 10(1):57–76, 2019.

[36] Janek Mann, R Rabinovich, Andrew Bates, S Giavedoni, W MacNee, and DK Arvind. Simultaneous activity and respiratory monitoring using an accelerometer. In *2011 International Conference on Body Sensor Networks*, pages 139–143. IEEE, 2011.

[37] Alison Marshall, Oleg Medvedev, and Alexey Antonov. Use of a smartphone for improved self-management of pulmonary rehabilitation. *International journal of telemedicine and applications*, 2008, 2008.

[38] MathWorks. Signal smoothing. `https://uk.mathworks.com/help/signal/examples/signal-smoothing.html`. Accessed: 2019-03-24.

[39] Medtronic. Nellcor™ bedside respiratory patient monitoring system, pm1000n. `https://www.medtronic.com/covidien/en-us/products/pulse-oximetry/nellcor-bedside-respiratory-patient-monitoring-system-pm1000n.html`. Accessed: 2020-02-01.

[40] Andrea Méndez, Patricio Labra, Rafael Pizarro, and Nelly Baeza. Low rates of participation and completion of pulmonary rehabilitation in patients with chronic obstructive pulmonary disease in primary health care. *Rev Med Chile*, 146:1304–1308, 2018.

[41] NHS. Chronic obstructive pulmonary disease (copd). `https://www.nhs.uk/conditions/chronic-obstructive-pulmonary-disease-copd/`. Accessed: 2020-01-26.

[42] NHS. Pulmonary rehabilitation. `https://www.england.nhs.uk/ourwork/clinical-policy/respiratory-disease/pulmonary-rehabilitation/`, 2020. Accessed: 2020-03-06.

[43] Information Services Division NHS. Acute hospital activity and nhs beds information in scotland, 2018. `https://www.isdscotland.org/Health-Topics/Hospital-Care/Publications/2017-11-07/2017-11-07-Annual-QuarterlyAcuteActivity-Report.pdf`. Accessed: 2019-01-25.

[44] National Institute for Health NICE and Care Exellence. Chronic obstructive pulmonary disease (copd) update, 20 may 2015. `https://www.nice.org.uk/guidance/qs10/documents/briefing-paper`. Accessed: 2020-01-26.

[45] M.A. Nielsen. *Neural Networks and Deep Learning*. Determination Press, 2015.

[46] Benjamin Noah, Michelle S Keller, Sasan Mosadeghi, Libby Stein, Sunny Johl, Sean Delshad, Vartan C Tashjian, Daniel Lew, James T Kwan, Alma Jusufagic, et al. Impact of remote patient monitoring on clinical outcomes: an updated meta-analysis of randomized controlled trials. *NPJ digital medicine*, 1(1):1–12, 2018.

[47] ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories et al. Ats statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med*, 166:111–117, 2002.

[48] Sojung Park, Ji Youn Kim, Jae Cheol Lee, Hyeong Ryul Kim, Seungjae Song, Hee Kwon, Wonjun Ji, and Chang Min Choi. Mobile phone app–based pulmonary rehabilitation for chemotherapy-treated patients with advanced lung cancer: Pilot study. *JMIR mHealth and uHealth*, 7(2):e11094, 2019.

[49] Romain A Pauwels, A Sonia Buist, Peter MA Calverley, Christine R Jenkins, and Suzanne S Hurd. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: Nhlbi/who global initiative for chronic obstructive lung disease (gold) workshop summary. *American journal of respiratory and critical care medicine*, 163(5):1256–1276, 2001.

[50] Nancye M Peel, Trevor G Russell, and Leonard C Gray. Feasibility of using an in-home video conferencing system in geriatric rehabilitation. *Journal of Rehabilitation Medicine*, 43(4):364–366, 2011.

[51] DJ Powrie. The bode index: a new grading system in copd. *Thorax*, 59(5):427–427, 2004.

[52] PhD Professor Thys van der Molen, MD. Clinical copd questionaire. `https://ccq.nl`, 2018. Accessed: 2020-04-01.

[53] Protokinetics. Understanding phases of the gait cycle. `https://www.protokinetics.com/2018/11/28/understanding-phases-of-the-gait-cycle/`, 2018. Accessed: 2020-03-01.

[54] Carolyn L Rochester, Ioannis Vogiatzis, Anne E Holland, Suzanne C Lareau, Darcy D Marciniuk, Milo A Puhan, Martijn A Spruit, Sarah Masefield, Richard Casaburi, Enrico M Clini, et al. An official american thoracic society/european respiratory society policy statement: enhancing implementation, use, and delivery of pulmonary rehabilitation. *American journal of respiratory and critical care medicine*, 192(11):1373–1386, 2015.

[55] Sundeep S Salvi and Peter J Barnes. Chronic obstructive pulmonary disease in non-smokers. *The lancet*, 374(9691):733–743, 2009.

[56] Praful Schroff, Jason Hitchcock, Christopher Schumann, J Michael Wells, Mark T Dransfield, and Surya P Bhatt. Pulmonary rehabilitation improves outcomes in chronic obstructive pulmonary disease independent of disease burden. *Annals of the American Thoracic Society*, 14(1):26–32, 2017.

[57] Graziele Coissi Sobrinho, Teresa James, Zoe Pond, Lindsay Welch, Susan Batchelor, Rhona Copp, Helen Potterton, Tracy Nash, Thomas Havelock, and Tom M Wilkinson. Impact of pulmonary rehabilitation on cough and sputum symptom perception in patients with copd, 2015.

[58] Jordi Esteve Sorribas. Using the respeck for gait analysis. unpublished, 2019.

[59] SparkFun. Accelerometer basics. `https://learn.sparkfun.com/tutorials/accelerometer-basics/all`. Accessed: 2019-03-23.

[60] Jonathan Tang, Allison Mandrusiak, and Trevor Russell. The feasibility and validity of a remote pulse oximetry system for pulmonary rehabilitation: a pilot study. *International journal of telemedicine and applications*, 2012, 2012.

[61] Ivan Tomasic, Viktor Avbelj, and Roman Trobec. Smart wireless sensor for physiological monitoring. In *pHealth*, pages 295–301, 2015.

[62] Niall Twomey, Tom Diethe, Xenofon Fafoutis, Atis Elsts, Ryan McConville, Peter Flach, and Ian Craddock. A comprehensive study of activity recognition using accelerometers. In *Informatics*, volume 5, page 27. Multidisciplinary Digital Publishing Institute, 2018.

[63] Terry Um, Vahid Babakeshizadeh, and Dana Kulic. Exercise motion classification from large-scale wearable sensor data using convolutional neural networks. pages 2385–2390, 09 2017.

[64] Eduardo Velloso, Andreas Bulling, Hans-Werner Gellersen, Wallace Ugulino, and Hugo Fuks. Qualitative activity recognition of weight lifting exercises. In *AH*, 2013.

[65] Sigrid NW Vorrink, Helianthe SM Kort, Thierry Troosters, and Jan-Willem J Lammers. Level of daily physical activity in individuals with copd compared with healthy controls. *Respiratory research*, 12(1):33, 2011.

[66] Jindong Wang, Yiqiang Chen, Shuji Hao, Xiaohui Peng, and Lisha Hu. Deep learning for sensor-based activity recognition: A survey. *Pattern Recognition Letters*, 119:3–11, 2019.

[67] Ganggang Xu, Jianhua Z Huang, et al. Asymptotic optimality and efficient computation of the leave-subject-out cross-validation. *The Annals of Statistics*, 40(6):3003–3030, 2012.

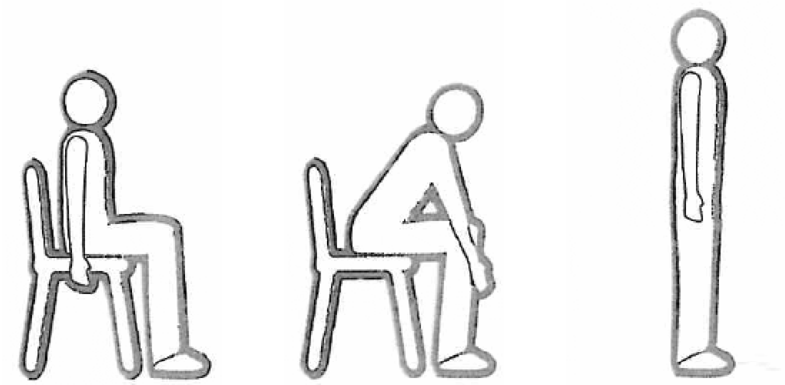
[68] Qi Zhou. A pulmonary rehabilitation system for copd patients on mobile devices using the wearable respeck breathing and activity monitor. unpublished, 2015.

Appendices

Appendix A

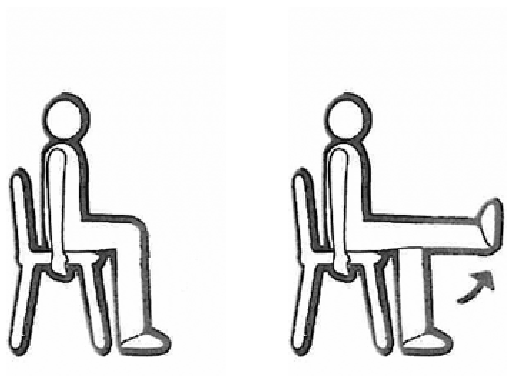
PR exercises

The figures below show each type of exercise in the *Rehab3* app and an explanation in the figure description. The images are shown as moving GIFs in the app.



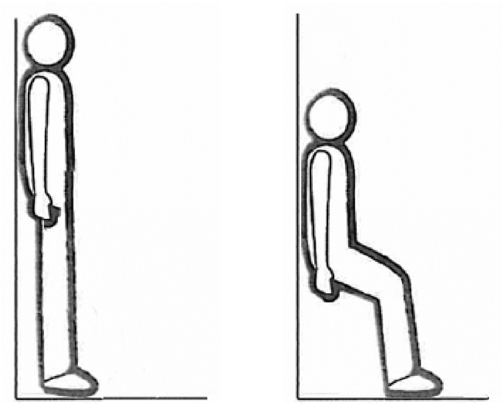
(a) Initial position. (b) Leaning forward. (c) Standing up.

Figure A.1: Sit to stand exercise.



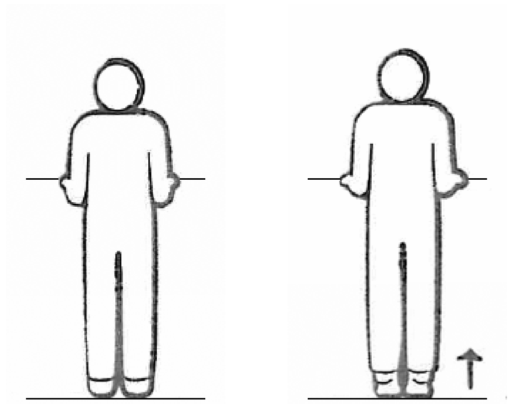
(a) Leg down. (b) Leg extended.

Figure A.2: Leg extensions exercise.



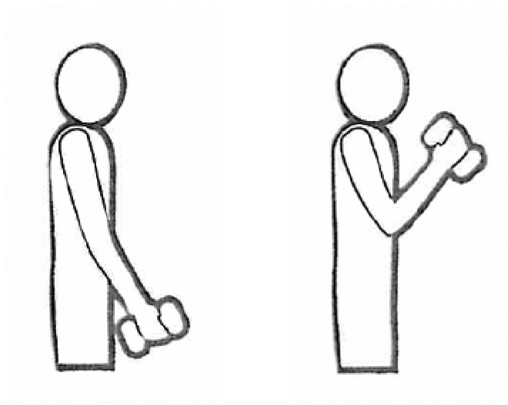
(a) Starting position. (b) Squatting down.

Figure A.3: Wall squats exercise.



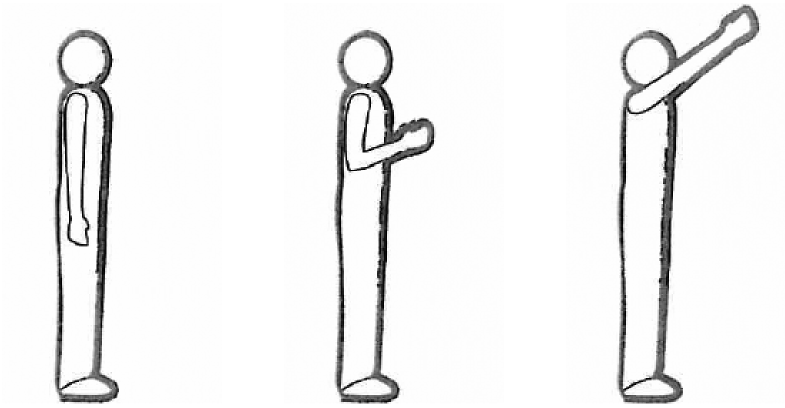
(a) Heels down. (b) Heels up.

Figure A.4: Heel raises exercise.



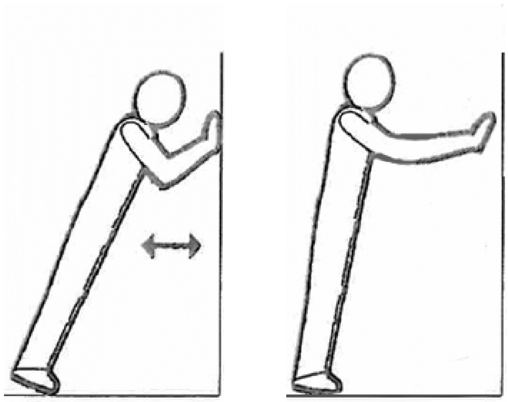
(a) Arms down. (b) Arms up.

Figure A.5: Bicep curls exercise.



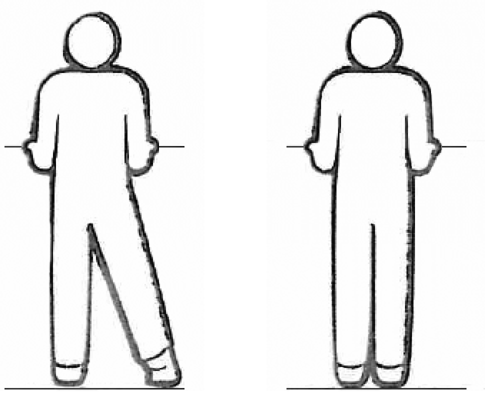
(a) Arms down. (b) Arms to chest. (c) Extending arms.

Figure A.6: Shoulder press exercise.



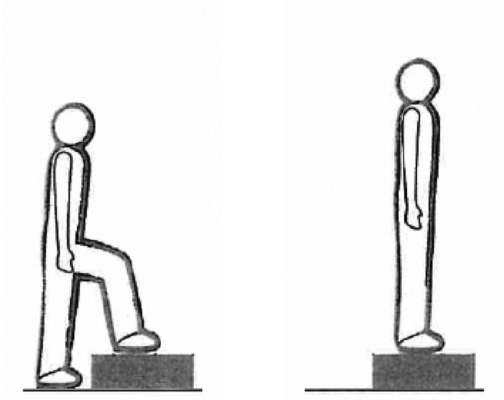
(a) Lean into wall. (b) Push away.

Figure A.7: Wall push-offs exercise.



(a) Leg in. (b) Leg out.

Figure A.8: Leg slides exercise.



(a) Foot on step. (b) Lift up.

Figure A.9: Step up exercise.

Figure A.10: Walking exercise.

Appendix B

Data collection participant information sheet

This is the participant information sheet presented to volunteers who helped with the data collection for the Pulmonary Rehabilitation system.

Participant Information Sheet

Project title:	A Remote Pulmonary Rehabilitation System Using the Wearable Respeck Monitor
Principal investigator:	Prof. DK Arvind
Researcher collecting data:	Teodora Georgescu

This study was certified according to the Informatics Research Ethics Process, RT number #4118. Please take time to read the following information carefully. You should keep this page for your records.

Who are the researchers?

The main researcher is Teodora Georgescu, who collects the data within the contents of her MInf2 project. She is a final year Master of Informatics student at the University of Edinburgh's School of Informatics. In conjunction with the Centre for Speckled Computing in the School of Informatics, she is performing a study on pulmonary rehabilitation exercises and their long-term effect on COPD patients.

Professor D K Arvind is the project supervisor and principal investigator. Andrew Bates a Research Fellow working in the research team and both are providing technical support in the MInf2 project.

What is the purpose of the study?

The research team aims to develop an app that can help COPD patients conduct Pulmonary Rehabilitation (PR) exercises at home, by wearing the Respeck sensor, which can track breathing and activity rates. We aim to train a model for each of the PR exercises and automate the recognition of correctly performed exercises.

Why have I been asked to take part?

You have been asked to take part in this study because you are a student at the University of Edinburgh.

Do I have to take part?

No – participation in this study is entirely up to you. You can withdraw from the study at any time, without giving a reason. Your rights will not be affected. If you wish to withdraw, contact the PI. We will stop using your data in any publications or presentations submitted after you have withdrawn consent. However, we will keep copies of your original consent, and of your withdrawal request.

What will happen if I decide to take part?

You will be asked to wear one sensor, the Respeck, which will collect your activity and breathing signals and send the data to an Android Application developed for data collection.

The Respeck device should be placed below the lower left ribcage, encased in a small disposable plastic bag, with the flat blue surface against your skin, and secured using the medical tape provided. Please ensure the device is the right way up, i.e., you can read the text on the flat side of the device.

You will be provided an Android phone for collection of the Respeck data. You will then be asked to perform a number of PR exercises while recording the data:

- Sitting down/standing up from a chair
- Leg raising while sitting down on a chair
- Wall squats
- Calf raises
- Bicep curls with small weights (1kg)
- Arm extensions
- Wall push-ups
- Side steps
- Step-ups on a low-height box
- Walking

Before and after each exercise, you will be asked to sit down and breahe normally for 30 seconds. Each exercise consists of a number of repetitions. You will be asked to perform the exercise routine twice. I will guide you through all the tasks for this experiment and verify that the data is being collected appropriately.

At any point in time, if you feel that you do not wish to continue with the study, then please feel free to let me know and the study will be stopped immediately.

Are there any risks associated with taking part?

You'll be asked to wear the Respeck device. This is a CE-marked device and as part of this process it has undergone the necessary safety tests. Participants with known plaster/plastic allergy will be excluded. The device is enclosed in a disposable plastic bag and is not in direct contact with the skin. The Respeck device is cleaned and sterilised once returned. There are no significant risks associated with participation.

Are there any benefits associated with taking part?

No.

What will happen to the results of this study?

The results of this study may be summarised in published articles, reports and presentations. Quotes or key findings will always be anonymous. With your consent, information can also be used for future research. Your data may be archived for a minimum of 5 years.

With your consent, the research team might share the fully anonymised data of this study with other researchers outside of the University of Edinburgh as part of publications.

Data protection and confidentiality.

Your sensor data will be processed in accordance with Data Protection Law. All information collected about you will be kept strictly confidential. Your data will be referred to by a unique participant number rather than by name.

Your sensor data will only be viewed by the research team: Teodora Georgescu, Andrew Bates and Professor D K Arvind for this MSc project. Your anonymised data may be used in other ethically approved research projects supervised by Professor D K Arvind or be made available to other researchers outside of the University of Edinburgh as part of publications. By signing the consent form, you agree to such usage.

Summaries of anonymised sensor data will be stored on Google Cloud Services (Google storage and Google datastore), where only the research team has access, and which is secured by two-factor authentication from Google. A copy of the anonymised sensor data is stored on the University's secure encrypted cloud

storage services *datasync* (https://www.ed.ac.uk/information-services/computing/desktop-personal/datasync), for which the research team has writing access and MInf and Year 4 project students supervised by Professor Arvind will have reading access. We only store summaries of accelerometer data, not personal information such as name, age or location.
Your consent information will be kept separately from your responses in order to minimise risk.

What are my data protection rights?
The University of Edinburgh is a Data Controller for the information you provide. You have the right to access information held about you. Your right of access can be exercised in accordance Data Protection Law. You also have other rights including rights of correction, erasure and objection. For more details, including the right to lodge a complaint with the Information Commissioner's Office, please visit www.ico.org.uk. Questions, comments and requests about your personal data can also be sent to the University Data Protection Officer at dpo@ed.ac.uk.

Who can I contact?
If you have any further questions about the study, please contact Teodora Georgescu (s1530344@sms.ed.ac.uk).

If you wish to make a complaint about the study, please contact:
Professor D K Arvind (dka@inf.ed.ac.uk) or the Informatics Ethics Panel (inf-ethics@inf.ed.ac.uk).

When you contact us, please provide the study title and detail the nature of your complaint.

Updated information.
If the research project changes in any way, an updated Participant Information Sheets will be made available on request from Teodora Georgescu (s1530344@sms.ed.ac.uk).

Alternative formats.

To request this document in an alternative format, such as large print or on coloured paper, please contact Teodora Georgescu (s1530344@sms.ed.ac.uk).

General information.

For general information about how we use your data, go to: edin.ac/privacy-research

Appendix C

Data collection participant consent form

The following document is the participant consent form presented to volunteers who helped with the data collection study for the Pulmonary Rehabilitation system.

Participant number:________________________

Participant Consent Form

Project title:	A Remote Pulmonary Rehabilitation System Using the Wearable Respeck Monitor
Principal investigator (PI):	Prof. DK Arvind
Researcher:	Teodora Georgescu, s1530344@sms.ed.ac.uk
PI contact details:	dka@inf.ed.ac.uk

Please tick yes or no for each of these statements.

		Yes	No
1.	I confirm that I have read and understood the Participant Information Sheet for the above study, that I have had the opportunity to ask questions, and that any questions I had were answered to my satisfaction.		
2.	I understand that my participation is voluntary, and that I can withdraw at any time without giving a reason. Withdrawing will not affect any of my rights.		
3.	I consent to my anonymised data being used in academic publications and presentations.		
4.	I understand that my anonymised data can be stored for a minimum of two years		
5.	I allow my data to be used in future ethically approved research.		
6.	I agree to take part in this study.		

Name of person giving consent	Date dd/mm/yy	Signature

Name of person taking consent	Date dd/mm/yy	Signature

THE UNIVERSITY *of* EDINBURGH informatics

Appendix D

System validation experiment participant information sheet

The following document is the participant information sheet for the system validation experiment.

Participant Information Sheet

Project title:	A Remote Pulmonary Rehabilitation System Using the Wearable Respeck Monitor
Principal investigator:	Prof. DK Arvind
Researcher collecting data:	Teodora Georgescu

This study was certified according to the Informatics Research Ethics Process, RT number #4118. Please take time to read the following information carefully. You should keep this page for your records.

Who are the researchers?

The main researcher is Teodora Georgescu, who collects the data within the contents of her MInf2 project. She is a final year Master of Informatics student at the University of Edinburgh's School of Informatics. In conjunction with the Centre for Speckled Computing in the School of Informatics, she is performing a study on pulmonary rehabilitation exercises and their long-term effect on COPD patients.

Professor D K Arvind is the project supervisor and principal investigator. Andrew Bates a Research Fellow working in the research team and both are providing technical support in the MInf2 project.

What is the purpose of the study?

The research team aims to develop an app that can help COPD patients conduct Pulmonary Rehabilitation (PR) exercises at home, by wearing the Respeck sensor, which can track breathing and activity rates. We aim to train a model for each of the PR exercises and automate the recognition of correctly performed exercises.

Why have I been asked to take part?

You have been asked to take part in this study because you are a student at the University of Edinburgh.

Do I have to take part?

No – participation in this study is entirely up to you. You can withdraw from the study at any time, without giving a reason. Your rights will not be affected. If you wish to withdraw, contact the PI. We will stop using your data in any publications or presentations submitted after you have withdrawn consent. However, we will keep copies of your original consent, and of your withdrawal request.

What will happen if I decide to take part?

You will be asked to wear one sensor, the Respeck, which will collect your activity and breathing signals and send the data to an Android Application developed for data collection.

The Respeck device should be placed below the lower left ribcage, encased in a small disposable plastic bag, with the flat blue surface against your skin, and secured using the medical tape provided. Please ensure the device is the right way up, i.e., you can read the text on the flat side of the device.

You will be provided an Android phone for collection of the Respeck data. You will then be asked to perform a number of PR exercises while recording the data:

- Sitting down/standing up from a chair
- Leg raising while sitting down on a chair
- Wall squats
- Calf raises
- Bicep curls with small weights (1kg)
- Arm extensions
- Wall push-ups
- Side steps
- Step-ups on a low-height box
- Walking

Before and after each exercise, you will be asked to sit down and breahe normally for 30 seconds. Each exercise consists of a number of repetitions. You will be asked to perform the exercise routine twice. I will guide you through all the tasks for this experiment and verify that the data is being collected appropriately.

At any point in time, if you feel that you do not wish to continue with the study, then please feel free to let me know and the study will be stopped immediately.

Are there any risks associated with taking part?

You'll be asked to wear the Respeck device. This is a CE-marked device and as part of this process it has undergone the necessary safety tests. Participants with known plaster/plastic allergy will be excluded. The device is enclosed in a disposable plastic bag and is not in direct contact with the skin. The Respeck device is cleaned and sterilised once returned. There are no significant risks associated with participation.

Are there any benefits associated with taking part?

No.

What will happen to the results of this study?

The results of this study may be summarised in published articles, reports and presentations. Quotes or key findings will always be anonymous. With your consent, information can also be used for future research. Your data may be archived for a minimum of 5 years.
With your consent, the research team might share the fully anonymised data of this study with other researchers outside of the University of Edinburgh as part of publications.

Data protection and confidentiality.

Your sensor data will be processed in accordance with Data Protection Law. All information collected about you will be kept strictly confidential. Your data will be referred to by a unique participant number rather than by name.

Your sensor data will only be viewed by the research team: Teodora Georgescu, Andrew Bates and Professor D K Arvind for this MSc project. Your anonymised data may be used in other ethically approved research projects supervised by Professor D K Arvind or be made available to other researchers outside of the University of Edinburgh as part of publications. By signing the consent form, you agree to such usage.

Summaries of anonymised sensor data will be stored on Google Cloud Services (Google storage and Google datastore), where only the research team has access, and which is secured by two-factor authentication from Google. A copy of the anonymised sensor data is stored on the University's secure encrypted cloud

storage services *datasync* (https://www.ed.ac.uk/information-services/computing/desktop-personal/datasync), for which the research team has writing access and MInf and Year 4 project students supervised by Professor Arvind will have reading access. We only store summaries of accelerometer data, not personal information such as name, age or location.
Your consent information will be kept separately from your responses in order to minimise risk.

What are my data protection rights?
The University of Edinburgh is a Data Controller for the information you provide. You have the right to access information held about you. Your right of access can be exercised in accordance Data Protection Law. You also have other rights including rights of correction, erasure and objection. For more details, including the right to lodge a complaint with the Information Commissioner's Office, please visit www.ico.org.uk. Questions, comments and requests about your personal data can also be sent to the University Data Protection Officer at dpo@ed.ac.uk.

Who can I contact?
If you have any further questions about the study, please contact Teodora Georgescu (s1530344@sms.ed.ac.uk).

If you wish to make a complaint about the study, please contact:
Professor D K Arvind (dka@inf.ed.ac.uk) or the Informatics Ethics Panel (inf-ethics@inf.ed.ac.uk).

When you contact us, please provide the study title and detail the nature of your complaint.

Updated information.
If the research project changes in any way, an updated Participant Information Sheets will be made available on request from Teodora Georgescu (s1530344@sms.ed.ac.uk).

Alternative formats.

To request this document in an alternative format, such as large print or on coloured paper, please contact Teodora Georgescu (s1530344@sms.ed.ac.uk).

General information.

For general information about how we use your data, go to: edin.ac/privacy-research

Appendix E

System validation experiment participant consent form

The following document is the participant consent form presented to the participants in the second experiment for the validation of the system.

Participant number:________________________

Participant Consent Form

Project title:	A Remote Pulmonary Rehabilitation System Using the Wearable Respeck Monitor
Principal investigator (PI):	Prof. DK Arvind
Researcher:	Teodora Georgescu, s1530344@sms.ed.ac.uk
PI contact details:	dka@inf.ed.ac.uk

Please tick yes or no for each of these statements.

		Yes	No
1.	I confirm that I have read and understood the Participant Information Sheet for the above study, that I have had the opportunity to ask questions, and that any questions I had were answered to my satisfaction.		
2.	I understand that my participation is voluntary, and that I can withdraw at any time without giving a reason. Withdrawing will not affect any of my rights.		
3.	I consent to my anonymised data being used in academic publications and presentations.		
4.	I understand that my anonymised data can be stored for a minimum of five years.		
5.	I allow my data to be used in future ethically approved research.		
6.	I agree to take part in this study.		

Name of person giving consent	Date dd/mm/yy	Signature

Name of person taking consent	Date dd/mm/yy	Signature

THE UNIVERSITY *of* EDINBURGH informatics