A MOBILE SOLUTION TO ANKLE RANGE OF MOTION (ROM) MEASUREMENT USING IMAGE PROCESSING TECHNIQUES

by
Gülşah Gökhan Gökçek

Submitted to Graduate School of Natural and Applied Sciences
in Partial Fulfillment of the Requirements
for the Degree of Master of Science in
Computer Engineering

Yeditepe University
2022

A MOBILE SOLUTION TO ANKLE RANGE OF MOTION (ROM) MEASUREMENT USING IMAGE PROCESSING TECHNIQUES

APPROVED BY:

Assist. Prof. Dr. Tacha Serif ....
(Thesis Supervisor)
(Yeditepe University)

Prof. Dr. H.Fatih Uğurdağ ....
(Özyeğin University)

Assist. Prof. Dr. Onur Demir ....
(Yeditepe University)

DATE OF APPROVAL: ..../..../2022

I hereby declare that this thesis is my own work and that all information in this thesis has been obtained and presented in accordance with academic rules and ethical conduct. I have fully cited and referenced all material and results as required by these rules and conduct, and this thesis study does not contain any plagiarism. If any material used in the thesis requires copyright, the necessary permissions have been obtained. No material from this thesis has been used for the award of another degree.

I accept all kinds of legal liability that may arise in case contrary to these situations.

Name, Last name Gülşah Gökhan Gökçek

Signature …..……………………….………………………..

ACKNOWLEDGEMENTS

I acknowledge Assist. Prof. Dr. Tacha Serif's support and assistance with great gratitude. Pursuing my thesis under his supervision taught a lot to someone like me who entered the field of computer engineering later onn.

I also thank all Research Assistants of the Yeditepe University Computer Engineering department for their advice and friendship.

Finally, I would like to give special thanks to my family especially my husband Umutcan and my little sunshine Aras for their continuous support and understanding.

ABSTRACT

A MOBILE SOLUTION TO ANKLE RANGE OF MOTION (ROM) MEASUREMENT USING IMAGE PROCESSING TECHNIQUES

According to the literature, the estimated number of people with and without foot pain varies between 17 percent and 30 percent. Looking at this data, it can be easily said that there is a very large population in society suffering from foot pain. Therefore, due to its high proliferation, foot pain is considered a major factor reducing the quality of life for many. Considering the above, this thesis proposes a prototype system to help people with reduced mobility, due to foot pain, or because they are living in a rural area, to submit their current feet condition and seek a doctor's and/or physiotherapist's consultation. Accordingly, the developed system is designed as a mobile e-health application so that users can register/login, add their health history, and regularly upload the latest status of their feet In addition, the developed system allows users to take videos of their ankles while bending their feet forward and backward (dorsiflexion and plantarflexion) and upload them to the system so that they can automatically calculate the range of motion. Consequently, the prototype is evaluated by 8 healthy and young participants. The measurements of the same participants is also performed by a physiotherapist with a goniometer. These two 16-measure datasets are first evaluated by looking at their intraclass correlation. Although there is no statistically significant correlation between them, these two data sets are analyzed with the Bland & Altman plot and found to be in agreement with fixed bias. This result shows that the developed system can be used instead of a goniometer with the improvements to be made. Usability tests are also conducted among the participants using the system. The effect of usability scores on the difference between measurements are examined and reported by linear regression analysis.

ÖZET

GÖRÜNTÜ İŞLEME TEKNİKLERİ İLE AYAK BİLEĞİ AÇIKLIĞI ÖLÇÜMÜNE MOBİL BİR ÇÖZÜM

Literatüre göre ayak ağrısı olan ve olmayan kişilerin tahmini sayısı yüzde 17 ile yüzde 30 arasında değişmektedir. Bu verilere bakıldığında toplumda ayak ağrısı çeken çok büyük bir nüfusun olduğu rahatlıkla söylenebilir. Bu nedenle, yüksek proliferasyonu nedeniyle, ayak ağrısı birçokları için yaşam kalitesini azaltan önemli bir faktör olarak kabul edilir. Yukarıdakiler göz önünde bulundurularak, bu tez, ayak ağrısı nedeniyle veya kırsal bir bölgede yaşadıkları için hareket kabiliyeti kısıtlı kişilerin mevcut ayak durumlarını sunmalarına ve bir doktor ve/veya fizyoterapist danışmanlığına başvurmalarına yardımcı olacak bir prototip sistem önermektedir. Buna göre geliştirilen sistem, kullanıcıların kayıt/giriş yapabilmeleri, sağlık geçmişlerini ekleyebilmeleri ve ayaklarının son durumunu düzenli olarak yükleyebilmeleri için mobil bir e-sağlık uygulaması olarak tasarlanmıştır. Ayrıca geliştirilen sistem, kullanıcıların ayaklarını öne ve arkaya doğru bükerken (dorsifleksiyon ve plantarfleksiyon) ayak bileklerinin videolarını çekmelerine ve hareket açıklığını otomatik olarak hesaplayabilmeleri için sisteme yüklemelerine olanak tanıyor. Sonuç olarak, prototip 8 sağlıklı ve genç katılımcı tarafından değerlendirilir. Aynı katılımcıların ölçümleri de fizyoterapist tarafından gonyometre ile yapılmaktadır. Bu iki 16-verilik veri seti, önce sınıf içi korelasyonlarına bakılarak değerlendirilir. Aralarında istatistiksel olarak anlamlı bir ilişki olmamasına rağmen, bu iki veri seti Bland & Altman grafiği ile analiz edildi ve sabit yanlılıkla uyum içinde olduğu bulundu. Bu sonuç, yapılacak iyileştirmelerle geliştirilen sistemin gonyometre yerine kullanılabileceğini göstermektedir. Sistemi kullanan katılımcılar arasında kullanılabilirlik testleri de yapılmaktadır. Kullanılabilirlik puanlarının ölçümler arasındaki farka etkisi doğrusal regresyon analizi ile incelenir ve raporlanır.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS/ABBREVIATIONS

AI	Artificial Intelligence
AWS	Amazon Web Services
B&A	Bland & Altman
CNN	Convolutional Neural Networks
DF	Dorsiflexion
ICC	Intra class correlation coefficent
IMU	Inertial Measurement Unit
JSON	JavaScript Object Notation
LSTM	Long short-term memory
MEMS	Micro-electromechanical
PF	Plantar Flexion
RCNN	Region Based Convolutional Neural Networks
ROM	Rotation of Motion
SPSS	Statistical Package for Social Sciences
UG	Universal Goniometer

1. INTRODUCTION

There is a fairly large population in the community suffering from foot pain. According to the literature, people suffering from foot pain have a distribution between 17 percent and 30 percent in the society [1], [2]. Thus, it is considered an important factor that reduces the quality of life for many. Since this problem reduces mobility, it is difficult for people with foot pain to receive treatment. Mobile technologies, which are improving and diversifying day by day, are considered as a solution to this problem.

In order to create a solution to the identified problem, a system to help people with foot pain or limited mobility due to rural areas to post updates of their current status and get consultation from a doctor and/or physiotherapist. The system is designed as an e-health application where users can register/login, add their health history and current status. Furthermore, this system allows users to take videos of their feet to evaluate ankle rotation of motion.

In this thesis, the validity of the ankle ROM measurement of the developed system is identified comparing with UG measurements taken by a physiotherapist. This comparison is performed by calculating intra class correlation coefficient of two methods. Also, the agreement between the two measurements is also investigated with the Bland & Altman plot with fixed bias. From another point of view, it is very much important to evaluate its usability and user experience issues and its pitfalls since all the data collection is done through a mobile phone interface. The usability is measured with system usability scale and user scores were collected. Further analysis with additional statements such as efficiency, learnability and user aesthetics are also accomplished

Accordingly, this paper is structured as follows; section 2 provides a detailed background and literature review on the topic of ankle ROM, ankle ROM measurement developments. Section 3 lists available methodologies that have been used to detect ankle ROM and existing tools that have been developed and used to identify this problem. Section 4 describes the analysis of the proposed solution and depicts the design of the system. Consequently, section 5 details the implementation steps of the prototype mobile application and the web services. Section 6 details the testing and evaluation methods and highlights the findings of these tests.

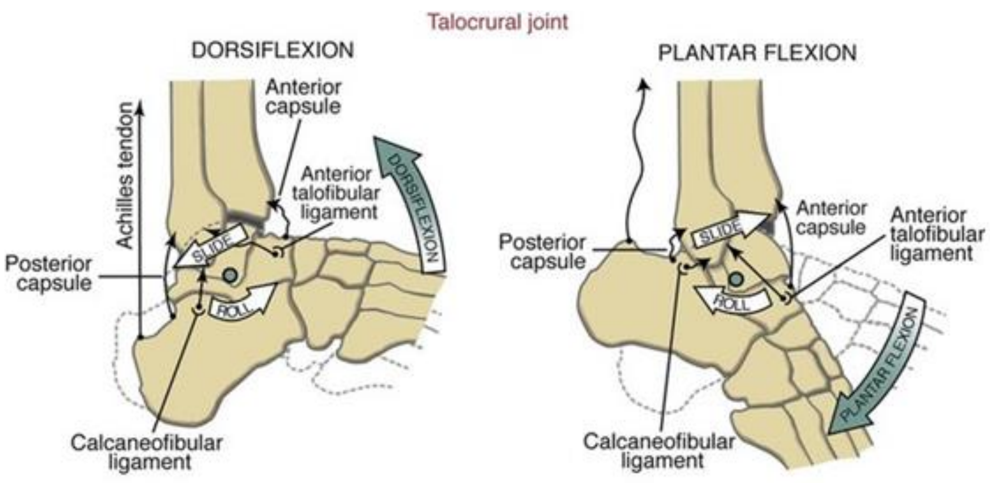
Last but not least, section 7 draws the conclusions 2 and proposes recommendations as future development items on the prototype by underlining its current shortcomings.

2. BACKGROUND

This section elaborates on the ankle anatomy and presents the ankle ROM measurements in the literature. Additionally, studies that involve technological measurement techniques in this field and their findings are discussed.

2.1. ROTATION OF MOTION

The ankle is popularly known as the junction of the foot and leg. To give a more scientific definition, an ankle is defined as a complex structure consists of two joints: true ankle joint and subtalar joint [3]. The true ankle joint has a hinge structure in which dorsiflexion and plantarflexion movement of the foot is produced (Figure 2.1). The first movement, dorsiflexion, occurs when the foot and leg approach each other (Figure 2.1a). The other flexion movement, plantar flexion, of the foot can be performed when the top of the foot moves away from the leg (Figure 2.1b). Plantar and dorsiflexion are the most common movements of the ankle in the sagittal plane. Therefore, the ankle range of motion (ROM) is considered as the maximum range where the tibiofibular joint can extend making these flexion moves.



(a) (b)

Figure 2.1. Talocrural (ankle) joint movements in sagittal plane [4]

Limited ankle range of motion may be a risk factor associated with loss of balance. For functional tasks like walking, which needs an angle of 8 to 10 degrees of dorsiflexion, a specific amount of ankle ROM is required [5]. In addition, any limitation in the ankle dorsiflexion is perceived as a risk factor or a variety of lower extremity illnesses, such as plantar fasciitis and Achilles tendinopathy. Plantar flexion limitations are generally associated with ankle injuries, including sprains and fractures [6]. Therefore, in the treatment process of these lower limb pathologies, clinicians measure range of motion to monitor the recovery process [7]. Measurement of range of motion can be an important step of the rehabilitation process because it can affect the selection of procedural interventions, the progression in rehabilitation process, and the end point of the treatment [8]. Due to its importance, various measurement techniques and tools have been developed.

2.2. ROM MEASUREMENT STUDIES IN REHABILITATION

In the clinical environment, the universal goniometer, inclinometer, and tape measure are the most commonly used tools. There are some parameters that affect the ankle ROM measurement when using these three tools. The measurement position and experience level of the rater are two examples of these parameters. The effects of these parameters on the measurements are revealed by inter and intra rater reliability and intra class correlation (ICC) analyzes. Inter rater reliability shows the measurement reliability of more than one rater while intra rater reliability is the consistency of the measurements taken by one rater repeatedly. To obtain the validity of the measurements ICCs are calculated statistically. In this section, clinical ROM measurement studies are presented by revealing the effects of these parameters on measurements.

The universal goniometer is an inexpensive and widely accessible tool used in many practice settings to measure ROM angle [4]. To evaluate the validity of the goniometer [9] conducted a comparison between visual estimation and goniometric measurement of ankle dorsiflexion. As part of the evaluation, they asked 12 orthopedic physiotherapists to measure ankle dorsiflexion first by visual estimation, then with a universal goniometer. The result of their study showed that inter-tester variation is twice that of goniometric measurement. Numerically, measurement error with goniometer was 5 degrees while variation in measurement was detected 11 degrees with visual estimation.

Another tool for measuring ankle dorsiflexion ROM is the inclinometer. One of the reliability studies of inclinometer is conducted by Venturni et al. [10] comparing with goniometer. Active ankle range of motion measurements are taken from healthy individuals. Two examiners performs the dorsiflexion measurements with these two techniques consecutively. Each examiner performed both measurement methods consecutively. As a result of this research, it is demonstrated that active dorsiflexion ROM measurement of the ankle using the digital inclinometer is more reliable than the one obtained with the universal goniometer. Also, they reported that inter-tester reliability is higher with an inclinometer as well.

Tape measure is the other simple tool used in the clinic environment to evaluate ankle ROM. Inter-rater and intra-rater reliability of tape measure conducted by Benell et al. [11]. The ankles of 13 healthy subjects are measured in a weight-bearing dorsiflexion (DF) lunge position. They performed their test using the inclinometer and tape measure and the testers had the changing experience levels. Their findings indicate that the inclinometer and tape measure can be affected by heel contact circumstances, lower leg bone form, and lower leg length, and that the inclinometer and tape measure are dependent on the subject's alignment. In addition, they have reported that the repeatability of DF measurement in lunge position did not appear to be affected by the raters' skill level or experience.

Finally, as an inclusive study of aforementioned three tools, Konor et al. [12] aimed to determine the reliability of ankle ROM measurements using three different techniques. Ankle dorsiflexion of twenty healthy subjects were measured in weight bearing position by a novice rater. They found that the reliability coefficients for the digital inclinometer and tape measure techniques were higher compared to the goniometer. Therefore, it can be interpreted that goniometer measurement needs a more experienced rater.

2.3. ROM EVALUATIONS WITH ELECTRONIC DEVICES

The need for more accurate and reliable ankle ROM measurement and advances in technology have led to the diversification of measurement techniques. Not only digital versions of simple tools mentioned in the previous section are developed, but also it has been seen that devices in clinical measurements or in the home environment have been diversified.

The following section; highlights and details the aforementioned devices and related studies conducted on them.

One device developed by Worsley et al. [13] and they presented a study designed to evaluate the reliability and validity of a new standardized method with a D-Flex device (Figure 2.2) to measure ankle range of motion. A group of 20 healthy subjects were occupied to quatify ankle range of motion in the weight-bearing position. They used an inclinometer, a goniometer, and the D-Flex measurement devices for measurements. The result of the study shows that the D-flex had ICCs 0.76–0.95, the highest inter-rater and intra-rater reliability, compared with the values of 0.55–0.85 and 0.32–0.71 for the goniometer and inclinometer, respectively. Consequently, it is stated that when testing ankle dorsiflexion ROM in healthy subjects, the D-Flex system offers a valid, portable, and easy to use alternative to the weight-bearing lunge test.

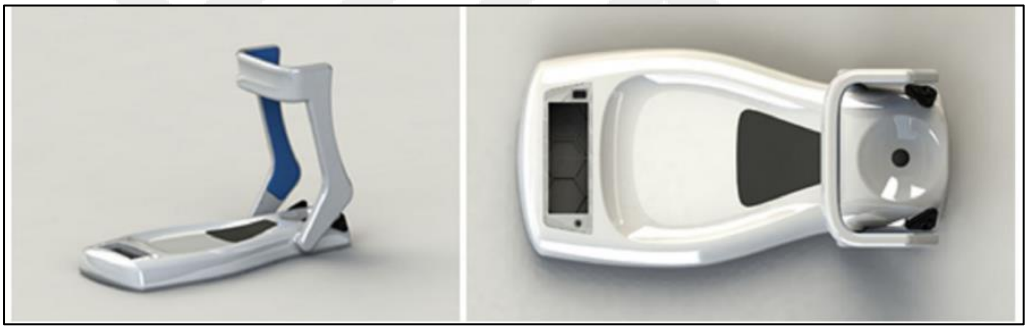


Figure 2.2. D-Flex device [13]

Another device introduced by Chen et al. [14] which is a more complex rehabilitation system which entailed a robotic ankle ROM measurement. They developed an electronic system to make sure the subject’s position and by considering the position of the subject to measure ankle ROM more accurately- as can be seen in Figure 2.3. As part of the study, the authors performed an experiment where a zero-torque control is applied to acquire the subject’s extreme joint angle and largest resistance torque when measuring the range of motion. Moreover, a controller switch is applied to ensure that the extreme joint angle is appropriate. The dorsiflexion of 10 healthy subjects measured with the proposed system and a physiotherapist. The maximum dorsiflexion angles measured with protractor were 32.1 ± 6.8 degrees while proposed system detect the agnles as 37.3 ± 6.5 degrees. The average discrepancy between the protractor and suggested system measurements was 17.35 with

11.05 percent deviation. The maximum plantar flexion angle of protractor and proposed system were reported as 55.8±12.2 degrees and 57.6±14.1 degrees, respectively. These results showed that the proposed system provides more accurate ROM assesment than a protractor. Also, they emphasized that ROM is an important parameter in ankle rehabilitation by indicating that accurate measurement of ROM helps effective stretching.

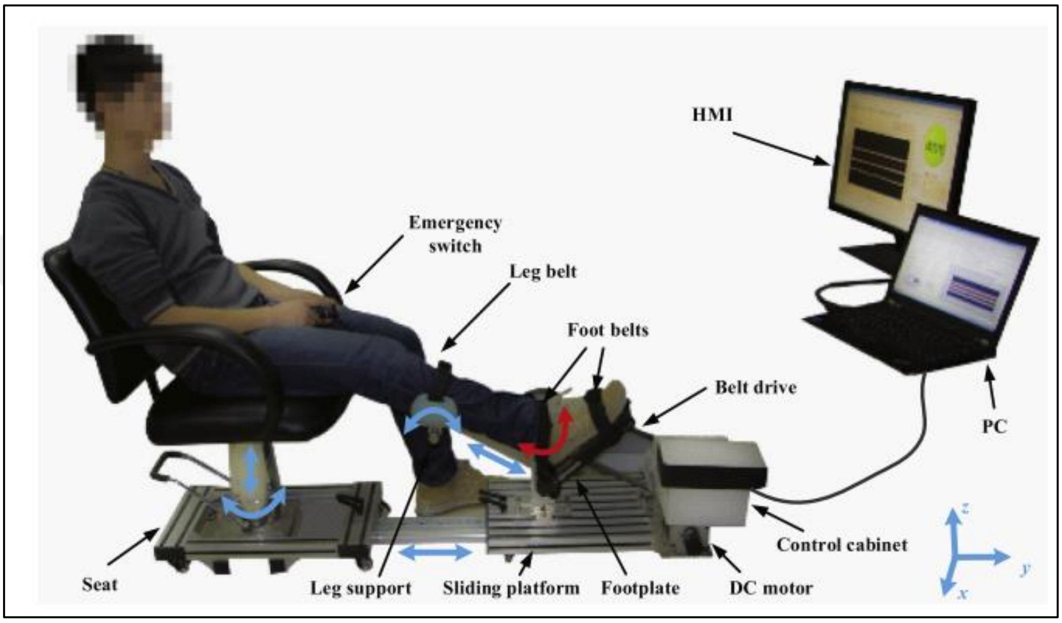


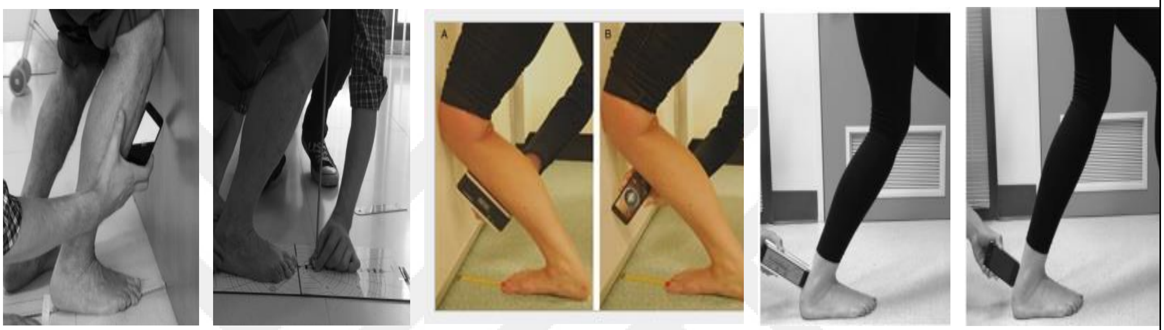
Figure 2.3. Rehabilitation system proposed by Chen et al. [14]

Furthermore, a study conducted by Ahmadian et al. [15] introduces a wearable system equipped with sensors in order to quantify ROM during triple single-leg hop trials. Three measurement units are placed to ten able-bodied participants. The strap-down integration approach was used to calculate 3D knee and ankle angles, and the results were compared to camera-based joint angles. From received signals ankle ROM measurement can be calculated. They found out that ROM error medians of the measurement are reported that when compared to the camera-based system correlation coefficients for both joints are less than 2.3 and 3.2 degrees, while correlation coefficients for all hop phases are greater than 0.92. Therefore, they interpreted that the new approach was accurate enough to detect post-injury changes in hopping kinematics and significant differences in ROM amongst clinical samples.

2.4. ANKLE ROM MEASUREMENTS WITH MOBILE APPLICATIONS

Although high measurement accuracy and reliability is reported for the electronic devices mentioned in the previous section, these devices are not portable. With the widespread use of smartphones, the mobile applications (apps) specifically designed to measure ROM offer clinicians a quick and easy method to examine flexibility [16]. These measurement apps mostly get inspired by simple tools. This section details ROM measurements using mobile and non-invasive solutions.

Inclinometer-based measurements are popular among mobile apps because mobile devices have gyro sensors which enable detecting the slope of desired surface. In literature, there are studies investigating some of these mobile apps in terms of their validity and reliability. First one is a study conducted by Morales et al. [17] which assesses the validity and reliability of the Leg Motion device for measuring ankle dorsiflexion ROM in older adults. In their work, authors take measurements of ankle dorsiflexion from the 33 healthy subjects with Leg Motion device (Figure 2.4.b), tape measure, UG and the iPhone's built-in inclinometer application (Figure 2.4.a). The mean values and standard deviations are calculated for measurements of Leg Motion device, tape measure, goniometer and inclinometer app. Findings of study shows that Leg Motion device is a valid and reliable tool as an alternative to the classic weight-bearing lunge test for measuring ankle dorsiflexion ROM in older adults. Another study is conducted to analyse the measurement reliability of iHandy mobile app by Vohralik et al. [18]. To achieve this measurements of iHandy and a digital inclinometer are compared (Figure 2.4.c-d). The ankle dorsiflexion range of motion was measured in 20 healthy persons using a convenience sample. Both the inclinometer's and the smartphone app's has ICCs and 95 percent CIs in excellent range, with narrow CIs. There was only a 0.07-degree and 0.2-degree difference between sessions for the inclinometer and the app, respectively. In the light of their findings, they concluded that a smartphone with the iHandy app can be used in clinical practice and research as an easy and convenient alternative to an inclinometer. Third and last inclinometer-based smartphone application is for the iPhone named Tiltmeter. Williams et al. [19] undertook a study where they compared the digital inclinometer and TiltMeter (Figure 2.4.e-f). Two raters (novice and experienced) conducted the measurements among 20 subjects in both a bent knee and straight leg position to determine the intra-rater and inter-rater reliability. Except for one intra-rater measure, the intra-rater and inter-rater reliability for the devices and in both positions were all above ICC 0.8. (Digital inclinometer, novice, ICC 0.65). The digital inclinometer and the tiltmeter have near-perfect inter-rater reliability, with an ICC of 0.96. The ICC between the two devices for concurrent validity was 0.83. As a result, the Tiltmeter app was proven to be a reliable and cost-effective method of determining the possible ankle range of motion. Besides, authors made a warning for health practitioners about using caution in applying these findings to other smart phone equipment if surface areas are not comparable.



(a) (b) (c) (d) (e) (f)

Figure 2.4. Inclinometer based mobile ankle ROM measurements [17]–[19]

From a different perspective, there are studies that tried to evaluate ROM utilizing mobile phone cameras. In this type of a solution, they aim to perform the measurement by putting virtual markers to the image taken by the phone. In addition, there are studies where real markers are stuck on the subject's body and then the ROM angle is calculated in reference to these markers using implemented software.

Some of these smartphone applications are evaluated by Wang et al. [20]. They measured ankle plantar flexion and dorsiflexion of 24 patients with ankle disease using the DrGoniometer app, Clinometer app, Goniometer app, and a universal goniometer. They utilized a 2-way mixed model to evaluate the interrater reliability and it showed good to excellent interrater reliability for each app for plantar flexion and dorsiflexion. Moreover, they apply a usability test to participants, and it is reported that the participants feel comfortable with measurement performed by mobile device. Hence, from findings of the experiments, they reported that the use of smartphone apps to assess ankle ROM may be more convenient than using UG. Physicians can use these apps to measure the patient’s dorsiflexion and plantar flexion from a picture sent by them remotely, or a patient could measure his or her own ROM at home using a personal smartphone.

Another study including smartphone measurement is conducted by Awatani et al. [21]. They examine the inter-rater reliability, validity, and error of angle measurements for ankle dorsiflexion range of motion using an image-based smartphone application. Eighteen volunteers participated in the study. Three examiners (two examiners are qualified certified athletic trainers and third one is not a certified trainer) performed smartphone application measurements. Two evaluation processes are followed in the experiment. Two image taken with a smartphone app with and without placing tape markers to the participants body. In first process, app's virtual markers are placed referring to the tape markers while the body parts are taken as reference points in the second process. They made statistical analysis from measurements taken from apps and radiographic images. Three examiners found a strong correlation coefficient between radiographic measurement and smartphone app measurement, as well as significant ICCs. Three examiners discovered that absolute difference upper 95 percent CI = 4.5 to 4.9 degrees, rate of absolute difference upper 95 percent CI = 15.8 percent –17.2 percent for measurement errors between radiography measures and smartphone app measurements. As a result, their study indicates that the inter-rater reliability of the ankle dorsiflexion range of motion measurements utilizing an application was high, with minor inter-rater measurement errors.

Another ankle ROM measurement research using markers is Romero Franco et al. [22]'s study. They analyzed the concurrent validity and reliability of an image-based iPhone application for assessing range of motion and joint position sense in ankle and knee joints. For knee and ankle, the intra-rater reliability of the iPhone app for measuring ROM and joint position sense was excellent and positive. Similarly, for ankle and knee ROM and knee JPS, inter-rater reliability was above ICC>0.80, but significant for ankle joint position sense. For validity, it was determined that the maximum value of bias observed in ankle joint position sense was -0.31 degrees, and the maximum range of bias in ankle ROM was -4.3 to +4.3 degrees. Hence, they state that this new iPhone application is valid and reliable for evaluating ankle and knee ROM and joint position sensation, while extra care is required during ankle evaluation to avoid inaccuracies.

Although Cunha et al. [23] did not evaluate the ankle ROM, they assess the ROM measurement from a video taken from a mobile app called Angles. Authors assessed the 11 convergent validity and reliability of joint angle measurements (elbow, shoulder hip, knee) from a video goniometer iPhone/iPad application in three groups - namely in adults, older and young children. Angles app can measure the ankle ROM by allowing the user to select a frame from a video and to place virtual markers to the selected frame. As part of their study, they evaluated the range of motion with this app for multiple body parts. They found out that there are high correlations for repeated measures between the app and the gold standard angle measurement instruments. Thus, they suggest that the app is a valid and reliable tool for assessing joint angles during functional activity

2.5. IMPACTS OF ARTIFICIAL INTELLIGENCE, MACHINE LEARNING AND DEEP LEARNING ON ROM MEASUREMENT

Human motion tracking for rehabilitation has been an active research topic since the 1980s[24]. The state of art method in this field is human pose estimation. Although it is mostly used to track human activity when performing exercises, it is possible to detect ROM angle using the same approach. In this section, the literature review is presented on this topic.

Kinect is a device popularly used in games that use body and body movement measurement. Due to this popularity, it has attracted the attention of many researchers.. For this reason, Ma et al. [25] investigate whether Kinect is a valid and reliable clinical gait analysis tool for children with cerebral palsy. Also, it is questioned that linear regression and LSTM recurrent neural network approaches may improve its performance. On 10 children with cerebral palsy, a gait study is performed. The correctness of kinematics derived using the Kinect, according to their findings, is inferior to that achieved using a typical marker-based 3DGA system. While the measurement errors of every ROM decreased after using the linear regression and LSTM calibration methods, the hip and knee sagittal angle trajectories matched the reference marker-based motion capture system the most. Although the Kinect's overall validity has been explored, they believe the Kinect-based gait analysis system has the potential to increase its capacity to measure lower limb kinematics.

Measurement with Kinect is costly and need preparation. Moreira et al. [26] presented that automatic identification of Anatomical and Segment Points can be detected using machine learning algorithms and computer vision models. As Moreira said, there are various studies [21], [27]–[35] with different methods in human pose estimation, using machine learning or deep learning algorithm. However, some of these studies does not contain footankle region, so they are not suitable for measuring ankle ROM.

One of the studies including ankle ROM measurement is a research by Ota et al. [36] who is questioning whether OpenPose-based gait analysis is valid or not. In their work, they investigate VICON and Openpose based detection algorithms. VICON is a sensor-based system used to detect sensor movement on the other hand Openpose is the first real-time multi-person system to jointly detect human body, hand, facial, and foot key points on single images. This study included twenty-four healthy young people. During walking and running, the subjects were evaluated. During treadmill walking and running, VICON was used to measure pelvic segment angles as well as hip, knee, and ankle joint angles. Digital cameras were used to collect photographs from both the right and back sides at the same time. The corresponding angles were measured from the estimated joint points after processing with OpenPose. Linear regression analysis was used to validate these estimates, and intraclass correlation coefficients between OpenPose and VICON data were calculated. It is found that for most of the ankle ROM, the coefficients of determinations were high. Although the ICCs showed moderate correlations, the 95 percent confidence intervals ranged widely as a result of the Bland &Altman test with fixed bias. In other words, they stated that the ankle joint angles measured using OpenPose and VICON are in agreement significantly.

Another study evaluating the ROM of the lower body including ankle ROM is Li et al.'s [37]. They develop an in-home lower body rehabilitation system based on a novel lightweight human pose estimation model. They implement this light-weighted model to a mobile device. To evaluate this system, they compared the key points found by the application with those marked in the COCO keypoint dataset [37]. The reliability of the proposed system is not validated with a clinician or a gold standard for ROM measurement. Nevertheless, their study shows that ankle ROM can be calculated using mobile devices with a deep learning approach.

3. METHODOLOGY

In this section, the existing ROM measurement techniques and methodologies are presented. Moreover, the usability testing methodologies and statistical analysis methodologies are also investigated.

3.1. QUANTIFYING ANKLE ROM

This section highlights the existing ROM measurement techniques and methodologies. Beginning with the introduction of traditional ROM measurement methods, different measurement techniques are presented. The last subsection introduces the artificial intelligence approaches that can be utilized to achieve a non-invasive ROM measurement.

3.1.1. Widespread Methods in Clinics

The goniometer measurement necessitates the greatest level of technical expertise due to the requirement to align the axis with the joint support and set the two arms at designated reference locations [38]. The UG's fixed arm is aligned with the lateral midline of the fibula, which is parallel to the fibula's head. At the intersection of the lines between the lateral midline of the fibula and the lateral midline of the fifth metatarsal, the axis should be distal to but aligned with the lateral malleolus. As shown in Figure 3.1, the moveable arm is positioned in the lateral midline of the fifth metatarsal.

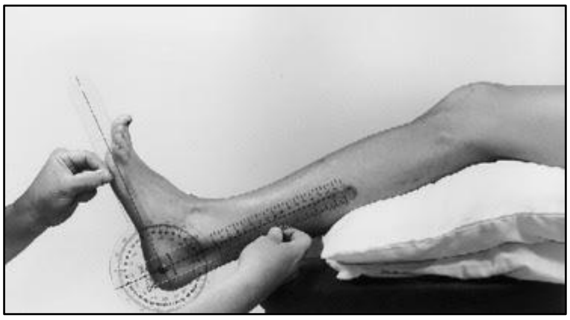


Figure 3.1. Goniometric measurement of ankle ROM [38]

The inclinometer measurement is done in the weight-bearing lunge position to acquire the most precise result of ankle dorsiflexion ROM [11]. The maximum distance of the toe from the wall while keeping contact with the wall and knee without lifting the heel was defined as maximal dorsiflexion in this position. The rater monitored heel contact with the ground by lightly resting their fingers on the heel and feeling for movement, as well as visually inspecting the heel for movement. A digital inclinometer (Figure 3.2a) was mounted at the tibial tuberosity and used to measure the angle of the tibia relative to the ground once the subject had achieved the final lunge posture. Measuring plantar flexion with the inclinometer placed at the previously marked tibial midpoint, the barefoot subjects were instructed to maximally PF their foot. The rater then elevated the test lower limb until the inclinometer read 0 degrees (Figure 3.2b). With the foot still pointed, a second reading was taken at the talonavicular joint on the dorsum of the foot (Fig. 3.2c).

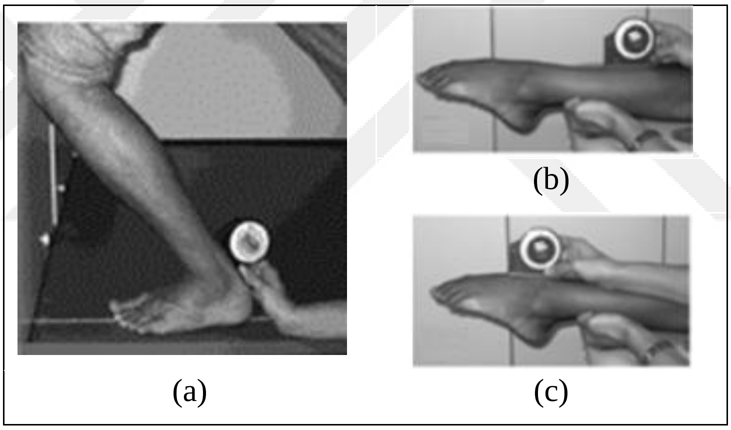


Figure 3.2. Inclinometer measurement [39]

The tape measurement can only be used to assess dorsiflexion. The patient lunges forward until the knee touches the wall, then places the test foot perpendicular to the wall on a tape measure. The foot is pulled away from the wall until just the knee makes little contact with the wall and the foot remains flat on the ground. The distance between the great toe and the wall is measured in centimeters, with each centimeter equivalent to about 3.6 degrees of ankle dorsiflexion.

3.1.2. Methods Using Digital Systems

Inclinometer-like measurements are performed using gyroscope and accelerometer sensors of mobile devices. Also, image marking methods also utilized to eval ankle ROM.

3.1.2.1. Sensor Measurements

A gyroscope is a device used for quantifing orientation and angular velocity [37]. There are different types of gyroscopes. A conventional gyroscope is a device that consists of a wheel mounted on two or three gimbals with pivoting supports that allow the wheel to revolve around a single axis. (Figure 3.3). A system of three gimbals, one positioned on top of the other with orthogonal pivot axes, can be utilized to allow a wheel mounted on the innermost gimbal to maintain its orientation regardless of its support's orientation in space.

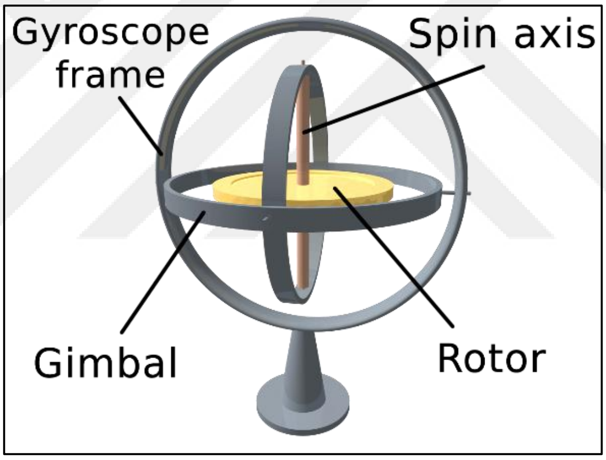


Figure 3.3. Gyroscope components [40]

On the other hand, Traditional mechanical gyroscopes employ wheels and gimbals, but the gyroscopes within cellphones do not. Instead, smartphones have MEMS (Micro-Electro-Mechanical Systems) gyroscopes. A MEMS gyroscope can measure the angle of orientation or the angular rate of rotation. The idea is based on a vibrating structure hanging in such a way that Coriolis forces may be detected while the mass rotates in relation to inertial space [41].

The integration of the gyroscope has enabled for more precise movement detection in 3D space than the prior lone accelerometer seen in a number of smartphones. For more robust direction and motion detection, gyroscopes are typically paired with accelerometers (acceleration sensors) in consumer electronics.

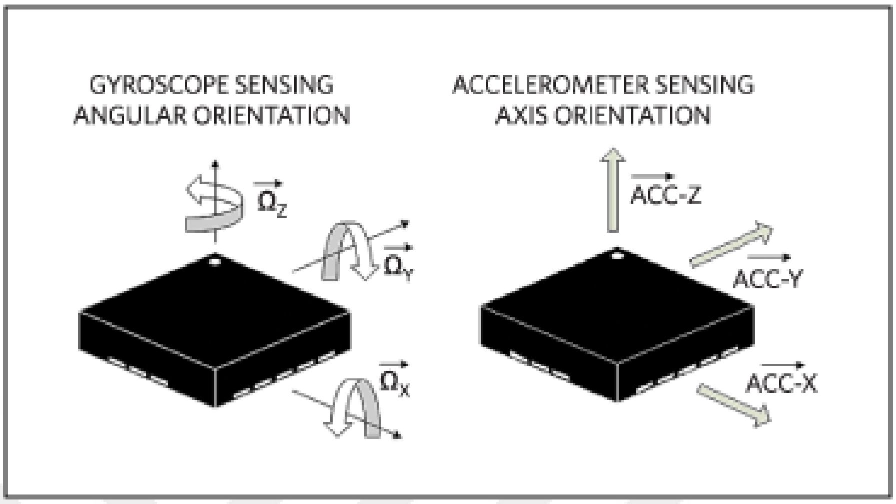


Figure 3.4. Gyroscope and Accelerometer MEMS [42]

An accelerometer is a sensor that determine the proper acceleration [43]. Proper acceleration refers to a body's acceleration in its own immediate rest frame [44]; as opposed to coordinate acceleration, which refers to acceleration in a fixed coordinate system. For example, accelerometers in free fall will measure zero. Accelerometers are included in several smartphones, digital music players, and personal digital assistants for user interface control; the accelerometer is frequently utilized to provide orientation of the device's screen depending on how the device is handled. Every generation of Apple devices has featured an accelerometer. Devices with Android OS also have accelerometer sensor. Accelerometers in mobile devices can be utilized as pedometers in conjunction with specialist applications, in addition to adjusting the orientation view. In addition to all these, ankle ROM can be measured with accelerometer sensor as well. In an unweight bearing position, device with an accelerometer is placed on the dorsum of the foot. When the foot flexing measurement is taken. With the gyro sensor, it is possible to measure ankle ROM just like the inclinometer described in the previous section. Dorsiflexion and plantar flexion can be detected using a device with a gyroscope instead of an inclinometer as shown in Figure 3.3. Hybrid methodologies also exists. IMUs with accelerometers and gyroscope sensor is put on the dorsum of the foot. With this method, ankle ROM can be calculated during hop test [15] .

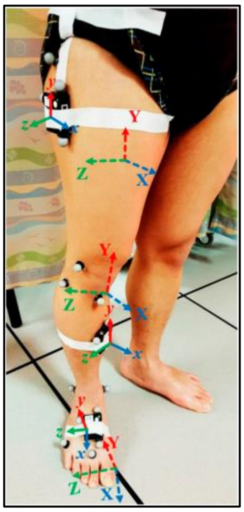


Figure 3.5. An example of a sensor measurement [15]

3.1.2.2. Marking techniques

The ankle ROM measurement can be performed on images or videos using marking techniques. It has been seen that the ankle ROM angle can be determined using virtual or real markers.

One method is that the angle is determined using photo editors such as Adobe Photoshop after taking a photo of the relevant joint identified using real markers. Before taking the photo, markers are placed on the participant's body (Figure 3.6). With the help of special systems, markers are detected to measure ROM. When an image is processed, markers lead to more accurate measurement. Therefore, it becomes possible to calculate ankle ROM by combining these joints.

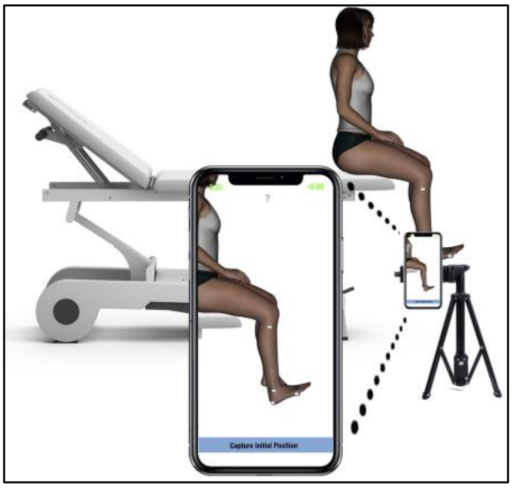


Figure 3.6. Real markers placed on human body [22]

Another method is to mark the required joints with virtual markers after the photo is taken. This process is generally done with mobile applications. Pieces of marker images are displayed to place on the photograph taken using the mobile application (Figure 3.7). The user places these markers at the appropriate positions on the joints they want to identify. Angles can be calculated using the positions of the markers on the screen. It has been observed that the same process is used for videos as well. It has been reported that it is possible to detect angles with markers on the selected frame from the video.

Figure 3.7. Virtual marker placed on an image [45]

3.1.2.3. Human Pose Estimation

The challenge of estimating the articulated joint locations of a human body from one image or a sequence of photos of that person is known as human pose estimation. This challenge has been tried to be solved with the methods developed over time. Thus, it has been observed that advances in computer learning provide more accurate human body part detection.

Edges, color histograms, contours, and histogram of oriented gradients were used in the early works of HPE for determining accurate locations of body parts by extracting the features. Color spaces were utilized to extract the color or luminosity features. For HPE, skin color is the most relevant feature to be extracted from a given image. Converting the color space of the image from RGB to HSV [46]–[50] and/or YCrCb [46], [51], [52] are widely facilitated to identify skin color in images. The conversions can be performed as formulated in Basha Shaik et al.'s study [53]:

$$H = \arccos \frac{\frac{1}{2}(2R - G - B))}{\sqrt{(R-G)^2 - (R-B)(G-B)}} \tag{3.1}$$

$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} \tag{3.2}$$

$$V = \max(R, G, B) \tag{3.3}$$

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0{,}419 & -0{,}081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{3.4}$$

The range specified for skin color is altering according to data in the literature but the range defined in Kolkur et al. 's study [54] combines two color spaces. For HSV and YCbCr color spaces, the range can be determined according to constraints in Table 3.1.

Table 3.1. Skin Color Ranges of RGB, HSV, YCbCr

Color Space	**Range**
RGB	R > 95 and G > 40 and B > 20 and R > G and R > B and | R - G | > 15 and A > 15
HSV	0.0 <= H <= 50.0 and 0.23 <= S <= 0.68
YCbCr	Cb > 85 and Y > 80 and Cr <= (1.5862*Cb)+20 and Cr >=(0.3448*Cb)+76.2069 and Cr >= (-4.5652*Cb)+234.5652 and Cr <= (-1.15*Cb)+301.75 and Cr <= (-2.2857*Cb)+432.85

Histograms can be used to extract features where histogram equalizations can be used to preprocess the images. Contrast Limited Adaptive Histogram Equalization (CLAHE) allows for histogram equalization without adding to the noise introduced by contrast amplification. The contrast amplification in the vicinity of a certain pixel value in CLAHE is determined by the slope of the transformation function. The slope of the neighborhood cumulative distribution function determines the value of the histogram at that pixel value (CDF). CLAHE limits the amplification before computing the CDF by clipping the histogram at a predefined value. The CDF's slope, and hence the transformation function, is constrained. The normalization of the histogram, and hence the size of the neighborhood region, define the clip limit, or value at which the histogram is clipped [55].

After feature extraction and enhancements, using regression, generation, part relations and other reasoning models, the human pose is estimated. These methods have quite a few drawbacks such as lack of representation of all constraints, low performance in extracting features and high reasoning complexity.

Pictorial Structures Model [28], [29], [56] and Flexible Mixture-of Parts [57] are two methods to estimate human Pose with machine learning algorithms. Pictorial Structures Model expresses the spatial correlation of rigid body parts as a tree-structured model to estimate the location of the joints. On the other hand, the Flexible Mixture-of Parts method uses the deformable part models. A mixture of small and non-oriented parts is formed considering the diversity of location and orientation of the body parts.

(a) (b)

Figure 3.8. Pictorial Structures Model [56] & Flexible Mixture-of Parts methods [57]

In recent years, deep learning has been widely applied in the fields of image classification, object detection, semantic analysis, human action recognition, and pose estimation.

Considering all deep learning methods to estimate human pose [21], [27]–[35], Convolutional Neural Networks (CNN) led to improvements in the HPE area. CNN is a deep learning model for processing data with a grid pattern, such as images, that is meant to learn spatial hierarchies of characteristics from low-level to high-level patterns automatically and adaptively. Convolutional neural networks are made up of three types of layers: convolution, pooling, and fully connected layers. Using these layers single-stage and two-stage detectors are created to detect objects [58]. While single-stage detectors make predictions by learning the class probabilities and bounding box coordinates, two-stage detectors first find the region of interest then perform the classification on that region. Two-stage detectors like Faster RCNN [59], have higher accuracy than single-stage detectors. Predictably, two-stage detectors are slower.

A two-stage CNN algorithm can be adapted to human pose estimation problems by adding the Mask branch and the key point detection branch [58]. For this purpose Mask RCNN was created.

Mask RCNN[59] is a deep neural network aimed to solve instance segmentation problems primarily. Figure 3.9 shows the Mask R-CNN framework for instance segmentation. As seen in this figure, feature maps are extracted from Faster R-CNN which uses a CNN. Feeding with the extracted feature maps, Region Proposal Network (RPN) returns the candidate bounding boxes. Regions are decided according to the Intersection of Union. To bring all candidates to the same size, a region of interest pooling layer is applied. These regions are given as an input to a fully connected network so that the class label and bounding boxes are predicted. Segmentation mask for each region that contains an object is also returned.

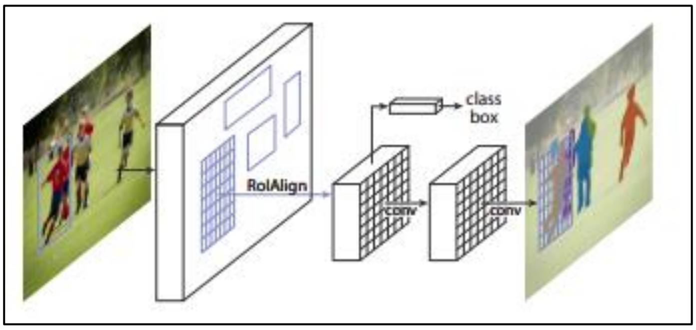


Figure 3.9. The Mask R-CNN framework for instance segmentation [59]

In order to adapt to human pose estimation, the position of a key point is modeled as a single-hot mask. Masks of K masks, one for each type of K keypoint (eg, left shoulder, right elbow) can be estimated using RCNN.

Thinning algorithms were also used to determine the foot posture within the scope of this thesis. Although there are various algorithms for thinning, most common ones are iterative algorithms using morphological operations [60] such as hit-or-miss or opening and erosion, Zhang Suen fast parallel thinning algorithm [61], Guo & Hall's two sub-iteration parallel Thinning algorithm [62].

Morphological methods such as erosion and opening were employed to remove selected foreground pixels from binary images. It operates on a hit-or-miss basis. The hit-and-miss transform is a binary morphological operation that can be used to search an image for certain patterns of foreground and background pixels. It is the most fundamental binary morphological operation, as practically all other binary morphological operators can be deduced from it.

The Zhang-Suen algorithm operates on black pixels with eight neighbors. This indicates that pixels near the image's border and corners aren't evaluated. Starting with a black pixel, it determines which of its neighbors will be white under certain conditions.

In the Guo & Hall thinning algorithm, contour pixels are examined for deletion in an iterative process. The decision is based on a 3 × 3 neighbourhood. In each iteration one cycle deletes the north and east pixels while the other turns south and west pixels into black.

Lastly, reminding that datasets are also an important factor in estimating the human pose in images and in frames. Although body datasets diversified in the HPE area, the most well-known three datasets are listed in Table 3.1. The first dataset in the table, the COCO, is widely used with CNN models in the literature. The dataset includes images annotated with 17 key points for each person. The second dataset MPII is another body dataset that consists of 16 key points for a person. The most important difference between COCO and MPII is while the former has support for multi-person HPE, the latter does not. The last dataset in the table is the AI Challenger dataset It is the largest one and it consists of high-resolution images. However, among these three, it has the least key points.

Table 3.2. Human Pose Estimation Datasets

	Total number of images	Human images	Key Points
COCO [63]	200k	250k	Up to 17
MPII [64]	25k	40k	Up to 16
AI Challenger Human Keypoint Detection [65]	600k	300k	14

3.2. USABILITY TESTING METHODS

Human-centred design, according to ISO 9241-210, is an approach to systems design and development that focuses on the use of the system and applies human factors/ergonomics and usability knowledge and techniques to make interactive systems more usable [66]. In this section test and evaluation methods are presented.

3.2.1. Test Methods

Usability test environment and the methods of performing the test is important. Usability.gov defines the behaviour models of subjects performing the usability test under 4 headings; (i)concurrent think aloud, (ii)in retrospective think aloud, (iii)concurrent probing, (iv)retrospective probing [67]. First model is applied by letting participants talk about the product when the test is going on. This approach aims to keep the participant's consciousness alive. For the second one, retrospective think aloud, participants watch the video replay of their actions after the test session is complete. In the third model, moderators ask follow-up questions to participants when they act differently from expected. The last one requires waiting and asking follow-up questions after the test session is complete. Generally, the last one is applied with a combination of other models.

Moreover, there are several metrics which is useful to collect during the course of testing.

- Successful Task Completion: Each scenario involves the collection of unique information that would normally be used in a routine activity. As the participant signals that they have found the solution or fulfilled the task goal, the scenario is considered complete. You might wish to ask participants multiple-choice questions in some circumstances. Include the questions and answers in the exam plan and make them available to note-takers and observers.
- Time On Task: The time it takes for the participant to finish the task.
- Critical Errors: Critical mistakes are departures from the scenario's completion objectives. For example, owing to the participant's workflow, reporting the incorrect data value. In essence, the participant will be unable to complete the assignment. The task aim may or may not be inaccurate or partial, and the participant may or may not be aware of it.
- Non-Critical Errors: Non-critical errors are ones that the participant recovers and that do not preclude the participant from successfully completing the task. As a result of these inaccuracies, the work is accomplished in a less efficient manner. Non-critical mistakes include exploratory activities such as opening the wrong navigation menu item or wrongly utilizing a control.
- Error-Free Rate: The error-free rate is the percentage of test takers who complete the job without making any errors.
- Subjective Measures: These are self-reported participant ratings for satisfaction, ease of use, discovering information, and other factors, with participants rating the item on a 5- to 7-point Likert scale.
- Likes, Dislikes and Recommendations: Participants share what they loved most about the site, what they disliked the most, and suggestions for how to improve it.

3.2.2. Evaluation Methods

Usability is a combination of factors including; intuitive design, ease of learning, efficiency, memorability, error frequency and satisfaction [68]. Intuitive design entails a near-intuitive grasp of the site's architecture and navigation. The speed with which a user who has never seen the user interface before can do basic tasks is referred to as ease of learning. The efficiency of use is how quickly a skilled user can complete activities. Memorability refers to a user's ability to recall information from a website long enough to use it successfully on subsequent visits. The frequency and severity of errors users make when using the system, as well as the severity of the faults and how users recover from them, are all factors to consider. The user enjoys using the system, as evidenced by subjective satisfaction. The System Usability Scale (SUS) survey [69] consists of statements that evaluates above factors. It has 10 statements and it uses a Likert Scale between 1 to 5. The statements that are odd 1 are subtracted from the score the user provides for the statement. For the even statements, the given score is subtracted from 5. At the end, sum all these numbers are multiplied by 2.5.

Moreover, an ISO standard introduces additional factors such as user interface aesthetics, accessibility and operability. Adding some statements to the survey gives more information about the usability of the system.

3.3. STATISTICAL ANALYSIS METHODS

When performing statistical analysis, first of all, the reliability of the data to be analyzed should be measured. The most common method for this measurement is Cronbach's Alpha [70]. Let X_i denote the observed score of item i and X the sum of all k items in a test. Let σij denote the covariance between X_i and X_j, σ_i^2 (=σ_{ii}) indicate the variance of X_i, and σ_X^2 indicate the variance of X. σ_X^2 including the item variances and inter-item covariances:

$$\rho\tau = \frac{\mathrm{k}}{\mathrm{k}-1}\left(1 - \frac{\sum_{i=1}^{k} \sigma_i^2}{\sigma_X^2}\right) \tag{3.5}$$

The higher ρτ value shows that the statements are reliable for defining the concept. Also, it is possible to capture the reverse direction by analysing correlations between statements.

Another analysis method for small sample size is the t-test. Under the null hypothesis, it is any statistical hypothesis test in which the test statistic follows a Student's t-distribution. The t-test is often preferred when the test statistic has a normal distribution and the value of a scaling component in the test statistic is known. When the scaling term is unknown and is replaced with an estimate based on the data, the test results follow a Student's t distribution.

For example, the t-test may be used to evaluate if the means of two sets of data are significantly different.

$$t = \frac{Z}{s} = \frac{X - \mu}{\sigma/\sqrt{n}} \tag{3.6}$$

distribution. For example, the t-test may be used to evaluate if the means of two sets of data are significantly different.

By creating bounds of agreement, the Bland-Altman (B&A) plot describes agreement between two quantitative measurements. The mean and standard deviation of the discrepancies between two measurements are used to calculate these statistical boundaries. They employed a graphical approach to examine the assumptions of normality of differences and other properties. Each of the n samples is represented on the graph by the mean of two measurements as the x-value and the difference between two values as the y-value value for a sample of n observations as in Figure 3.10.

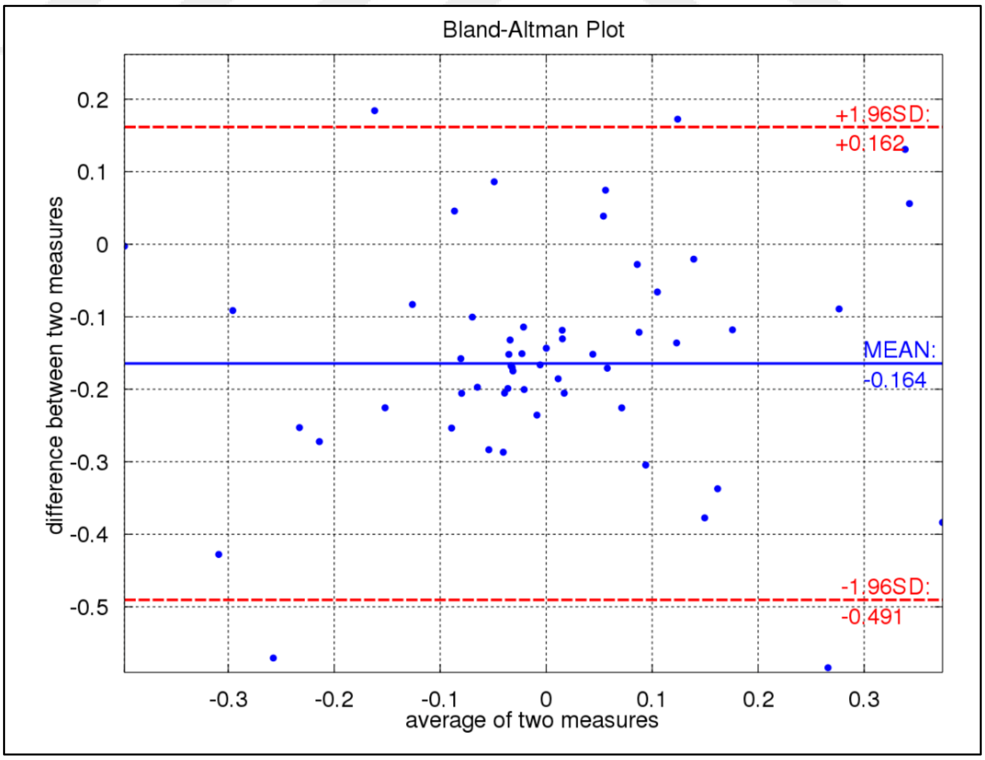


Figure 3.10. Bland Altman plot [71]

The Pearson Correlation Test is a method for detecting if two sets of data have a linear relationship. The Pearson correlation coefficient is the ratio of two variable' covariance to the product of their standard deviations and this coefficient is always between -1 and 1. The measure, like covariance, can only show a linear correlation of variables and eliminates many other forms of interaction or association. The age and height of a group of high school students, for example, should have a Pearson correlation value that is considerably more than 0, but less than 1.

$$\rho_{X,Y} = \frac{\mathrm{cov}(\mathrm{X},\mathrm{Y})}{\sigma_X \sigma_Y} \tag{3.7}$$

Where cov is the covariance, σ_X is the standard deviation of X and σ_y is the standard deviation of Y.

4. ANALYSIS AND DESIGN

In this section, the problem that is the subject of the thesis research is described. From these stages of the research, some functional and non-functional requirements are determined. The identified requirements and the solution domain are presented.

4.1. PROBLEM DEFINITION

The problem definition process is carried out with a team of an orthopedic doctor and two experienced physiotherapists. It has been stated that patients with reduced mobility due to foot disorders such as plantar fasciitis and Achille tendinopathy have difficulty in coming to the clinics. For this reason, patients may not receive any treatment, or their healing process is prolonged. Clinicians have stated that there is no remote system that can be easily used for ankle ROM measurement, which is known to be a risk factor for the previously mentioned diseases.

As discussed in the previous section, techniques exist to remotely measure ankle ROM non-invasively. These techniques are performed through photography or video. Some require finding joint points and marking them but this needs experience in finding joints. Others calculate the Rom angle by detecting the human pose. However, these techniques require additional equipment such as a high-resolution camera, Kinect or sensor system. It is possible to estimate human pose by using a smartphone camera but there is no application or system which is validated for measuring ankle ROM.

These identified problems led to the thesis research. The proposed system can be used measuring ankle ROM remotely. In addition, whether usability of the system affects the measurement quality.

4.2. REQUIREMENTS

The functional requirements are detected with a use case scenario defined by clinicians. The scenario is below:

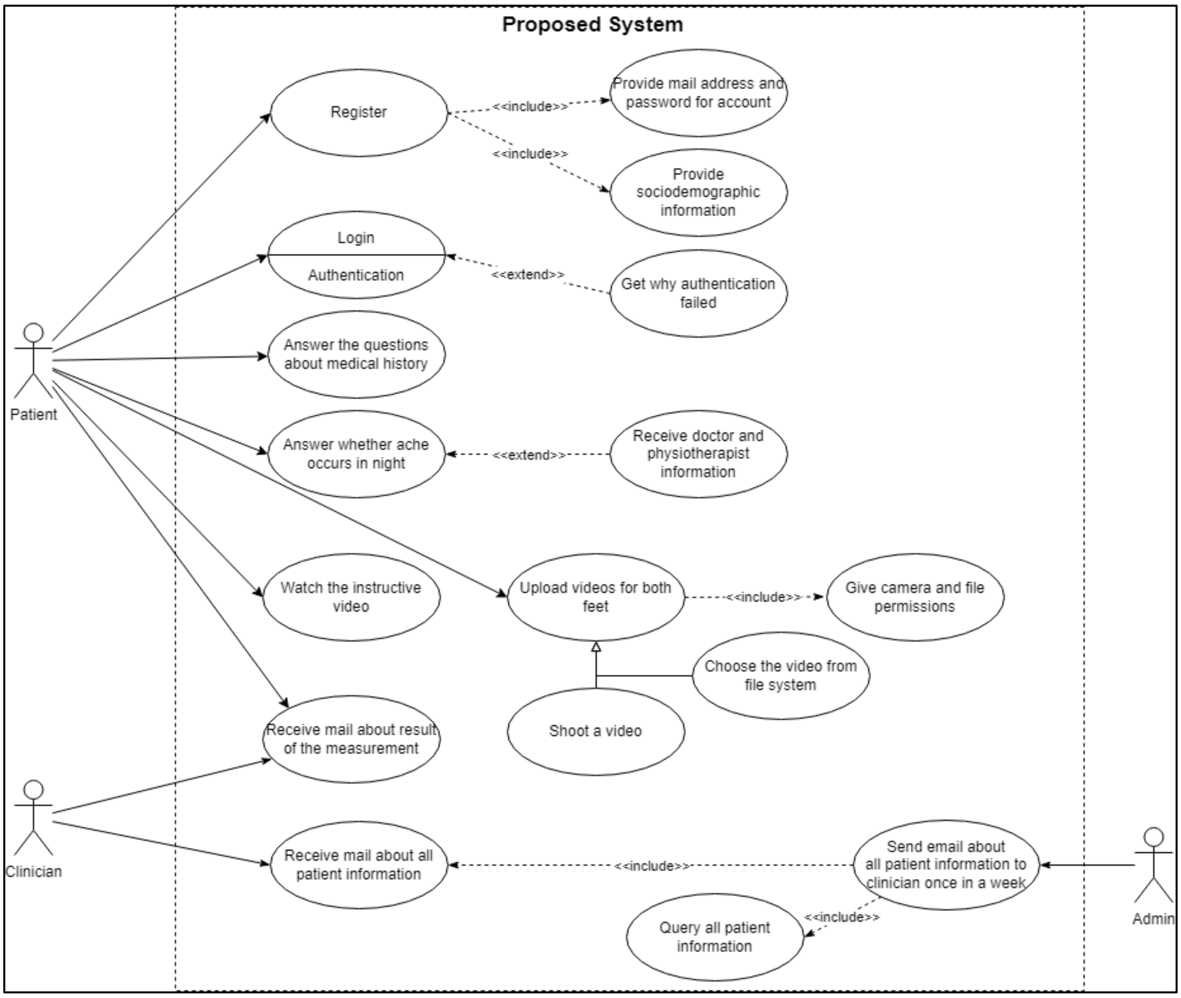


Figure 4.1. Use Case Diagram of the Proposed System

The patient who can speak Turkish should use a system with authentication. Both user and sociodemographic information should be gathered. S/he should fill questions about the medical history. Then, information about the level of her/his ailment should be gathered for both sides of the feet In the meantime, a question should be raised to ask whether the patient feels a hindfoot pain at night or not. Patient should continue to use the system in both cases but if her/his answer is yes, the system should display the clinician's information. After dealing with all questions, ankle ROM should be detected non-invasively. The system should show the patient how to perform measurement with an explanatory text and a tutorial video. At the end, both clinician and patient should be informed about measured ankle ROM.

Within the scope of the use case shown above (see Figure 4.1), it is aimed to develop a system to gather information about hindfoot pain related diseases, and to calculate the risk factor (limitation in ankle ROM) . In order to propose a solution to the problem, it is required to implement a client-server system because the data should be collected remotely from patients and can be accessible to physiotherapists.

In this scenario, the client should be a mobile application that can gather the required data. Each patient should be associated with a user account to be able to authenticate her/himself. If internet connection exists, the user should register and login to the system. If there is no connection, the user cannot perform the authentication. However, the application should run offline as a quality requirement. In this case, a previously logged in user should be able to use the application. Consequently, to store the data when connection is lost, there should be a local database. Moreover, methods with camera components should be utilized to quantify ankle ROM in a non-invasive way. Therefore, the device that is hosting the application should have a camera component. The videos taken by the patient using the camera component should be stored in the file system of the device because of the offline feature of the application. When the connection is established, the patient should be able to upload the previously recorded video. All collected data including the video should be able to be transmitted over mobile data or WiFi channels.

On the other end, server-side, multiple services should be provided to meet the scenario. In addition to services such as user management, and data storage, the server should also provide ROM angle measurement service. For the user management and the data storage services, a relational database should be facilitated. Furthermore, data storage service involves archiving the video files. Last but not least, the ankle ROM measurement service should be able to detect human foot from a video file. For this purpose, the server should be capable of running the machine learning frameworks such as Tensorflow, and Keras. Also, the server should be compatible with CUDA which is a parallel computing platform and application programming interface that allows software to use certain types of graphics processing unit (GPU) for general purpose processing.

4.3. DESIGN

As described in the requirements, a client-server model is designed. Client side includes mobile devices that can run on Android operating system, and run the mobile application, which is designed to collect socio-demographic information and foot-ankle-specific information of individuals in Turkish language. The application is intended to provide interaction between the system to be developed and users. It is planned to transmit user data to the platform over the same system infrastructure.

The server side infrastructure to be developed consists of a web server, database server and an application server which provide services running on a Linux-based operating system. In addition to services such as user management, data storage and web interface, the server is also intended to provide ROM range measurement. It is aimed to calculate (detect) risk factors for hindfoot disorders with ROM range measurement. Figure 4.2 shows how ROM measurement service interacts with clients and other system components.

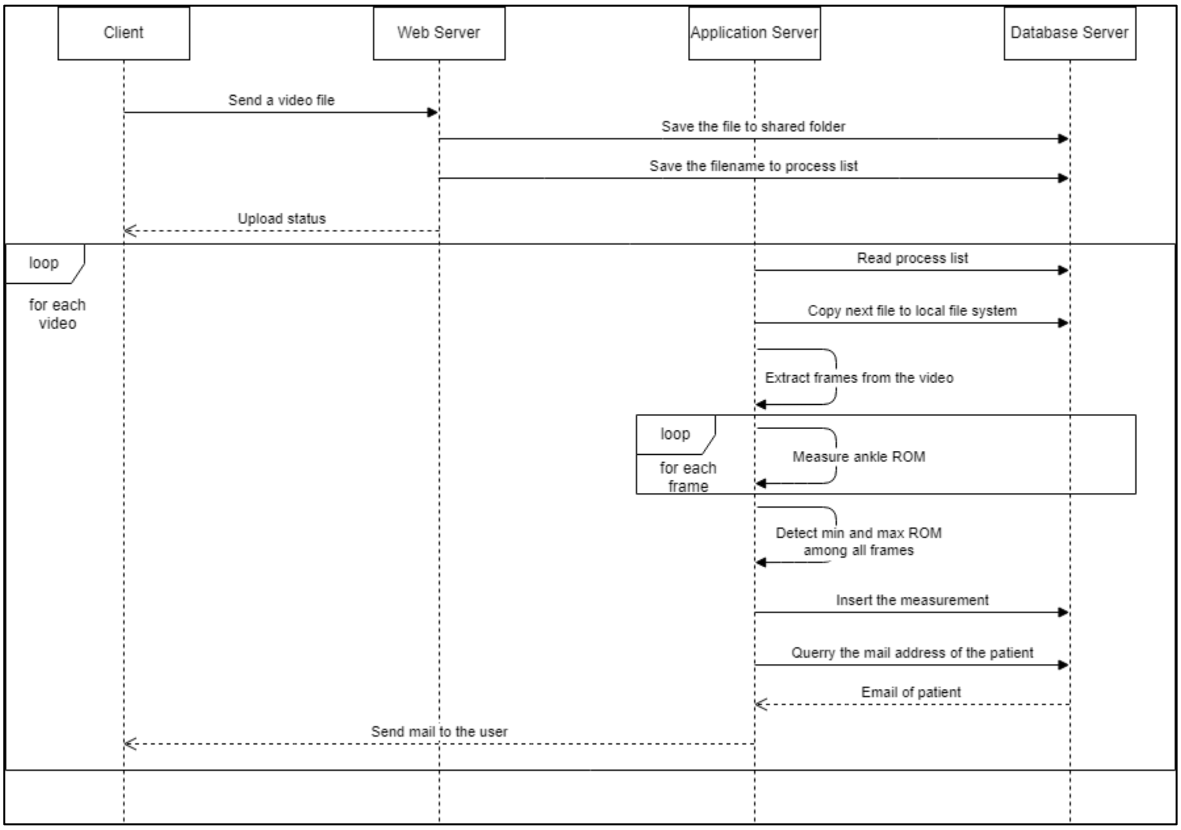


Figure 4.2. Sequence diagram of ROM Measurement service

In the designed system (Figure 4.2), the service is started when the customer sends a video to the server. The web server receives the sent packet and writes the video to the public folder. It also adds the filename of the video to the list of processes in the public folder. Since the measurement time of the measurement algorithm may be long, the web server returns a response to the customer at this point. In the next part of the service, the measurement algorithm is triggered by time and reads the file names written to the process list sequentially.

It copies the file corresponding to each name it reads to the local file system in order to preserve the original file. The ROM angle is calculated for each frame by dividing the video file in the local file into frames. Minimum and maximum measurements are made for each video file and the difference between them is written to the database. The measurements made by querying the user email address database are notified to the user by email.

5. IMPLEMENTATION

Considering the previous section, a client-server system is implemented as shown in Figure 5.1. Bearing in mind the requirements highlighted in the previous section, the system consists of two main components: Mobile and Cloud.

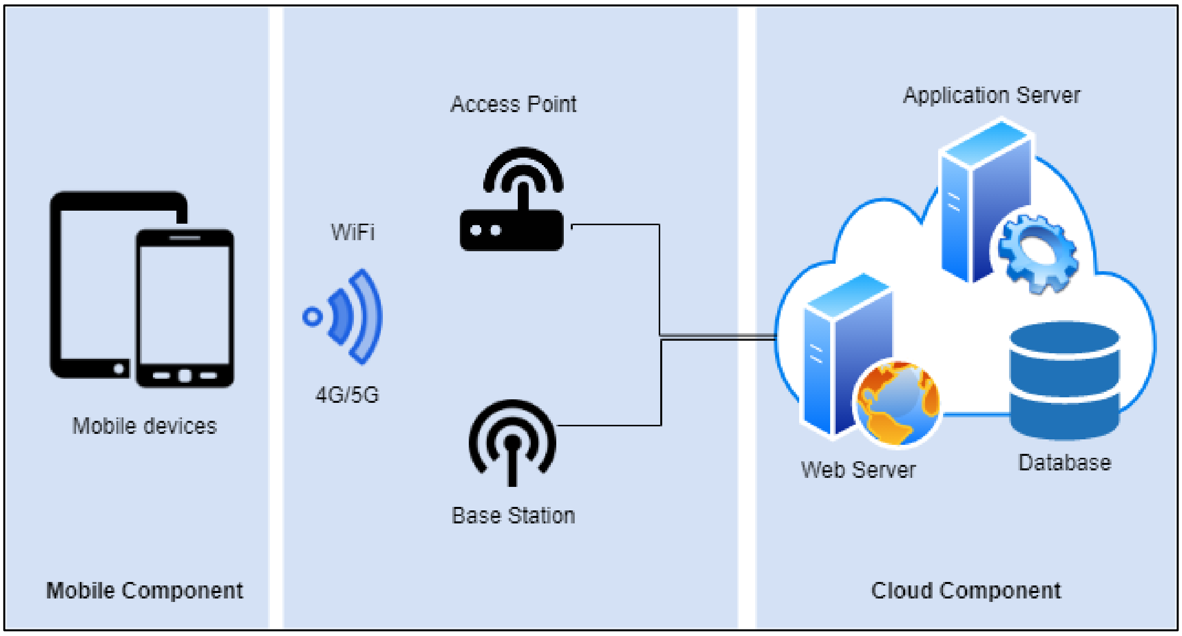


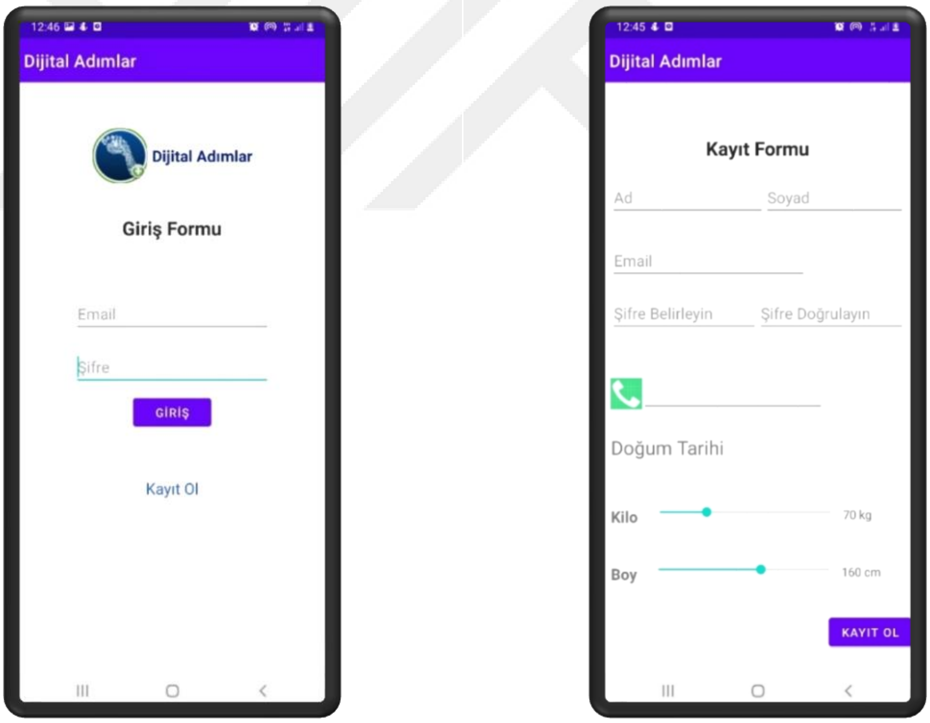
Figure 5.1. Overall architecture of proposed system

5.1. MOBILE COMPONENT

The mobile component will consist of the mobile application developed in accordance with the mobile device and device operating system on which the software will run. Accordingly, the mobile component will consist of software that creates a user interface suitable for the screen sizes of the device used, instantly stores the collected user data, communicates with the cloud component, and safely directs the stored data to the right service.

5.1.1. User Interface

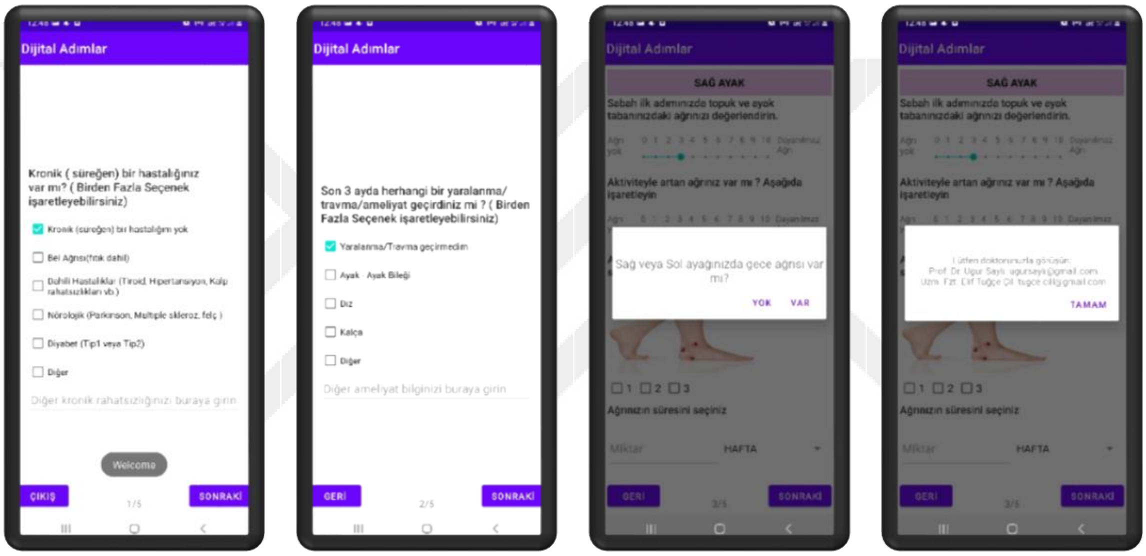
It is envisaged that the mobile application will consist of 7 screens. User login form is designated to be the first screen among the created interfaces. Two options are offered on this form; one is that new users can access the registration screen, the other is that registered users can log into the system. People who want to create a new user registration are able to reach the registration screen by clicking the “Kayıt Ol” text in blue color (Figure 5.2a). The registration page is created as a form screen where users can enter their sociodemographic (name, surname, contact, date of birth, height and weight) information and create a user account by setting a password (Figure 5.2b). In this screen, UI components are utilized to simplify the data input. Text data is collected with EditText objects. The Date of birth is selected using the DatePickerDialog object. Also, users identify their height and weight with the SeekBar object. The selected value is displayed at TextView located at the end of the SeekBar. After all sections are filled using these UI components, the information will be saved with the “Kayıt Ol” button. The registered user is able to log in to the system pressing the “Giriş" button with the account information s/he has created after this stage.



(a) (b)

Figure 5.2. Mobile app login/register screens

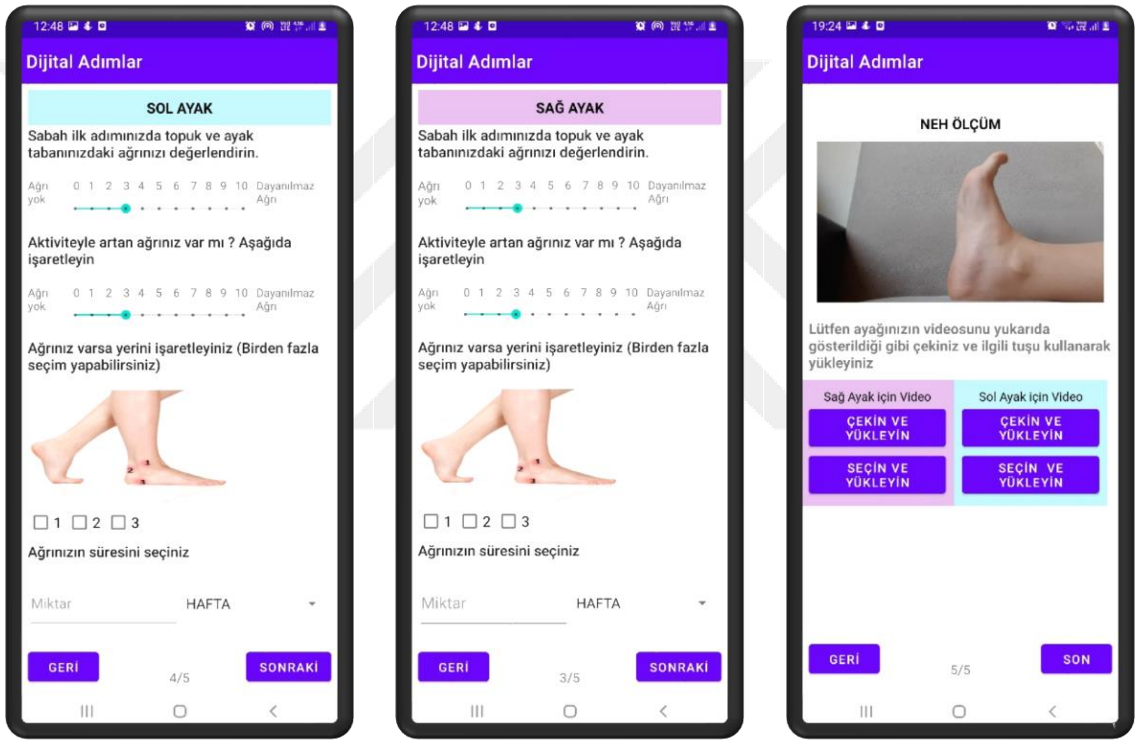
When the user logs in to the system with the account information previously created, s/he is faced with screens showing questions about chronic disease (diabetes mellitus, complications of neurological and internal disorders) information and the history of surgery/injury (Figure 5.3a-b). In order to facilitate data entry in the information collection process and increase the usability of the application, it is planned to use checkboxes for these questions. Upon answering these, the information entered is saved into the local database by clicking the "Sonraki" button and it is transferred to the hind foot pain screen. As soon as this screen changes, , the night pain dialog box will appear on the screen (Figure 5.3c). Users who state that they do not have night pain will be directed to the “Sağ Ayak” screen, which is the next screen in the application. On the other hand, users who state that they have night pain will be able to contact their doctor using the communication options provided (Figure 5.3-d).



(a) (b) (c) (d)

Figure 5.3. Medical background and status screens

The screen for collecting data about foot pain has been created in a structure where the user can easily enter the pain degrees, regions, and durations for the right and left feet in various situations (for example, when waking up in the morning, inactivity) (Figure 5.4a-b). The layout of these screens, which were created to collect right and left foot data, was kept the same in terms of ease of use and learnability. For the sake of distinguishing one from the other, colors are assigned to the sides of the foot. In other words, screens or tasks related to the right foot have pink color while those related to the left foot have a blue color. If the required data is entered and the “Sonraki” button is pressed, the user is directed to the risk factors data collection screen, which is the next screen in the application. In the risk factors screens, the user will be greeted with the screen designed for the ROM measurement service (Figure 5.4c). In order to explain how the video should be taken for the measurement, a tutorial video demonstration is placed on the screen using a VideoView. This instructive video is displayed when an internet connection exists. After watching the tutorial, users can select to upload the video via shooting the video or choosing the video from the local file system of the device. For each foot, there are two buttons corresponding to these two options. Users press the "Çekin ve Yükleyin" button to open the camera of the device. If the "Seçin ve Yükleyin " button is used, then the gallery of the phone is opened for choosing the desired video. With the "Son" button the data except the video file is sent to the server and a Toast message is displayed indicating that the measurements will be emailed when it is calculated.



(a) (b) (c)

Figure 5.4. ROM measurement screen mock-ups

5.1.2. Local Database

In order to have an offline feature, the SQLite database engine is facilitated. Sociodemgrphical and user data does not stored in the local database because it is not possible to register without a connection. When the user register, two entries are inserted into the user table and patient table of the main database. After that when the user login to the system, patient_id is sent from cloud component to the mobile application. This data is stored in the patient table of the local database. In addition, it has the tables to store the data collected from the screens about chronic pain, surgery, detailed pain information for both sides of the feet (Figure 5.5). As seen in the figure, night pain is also stored in the patient table because these two has 1 to 1 relationship between each other. Except patient table, private key of all tables are autoincremented.

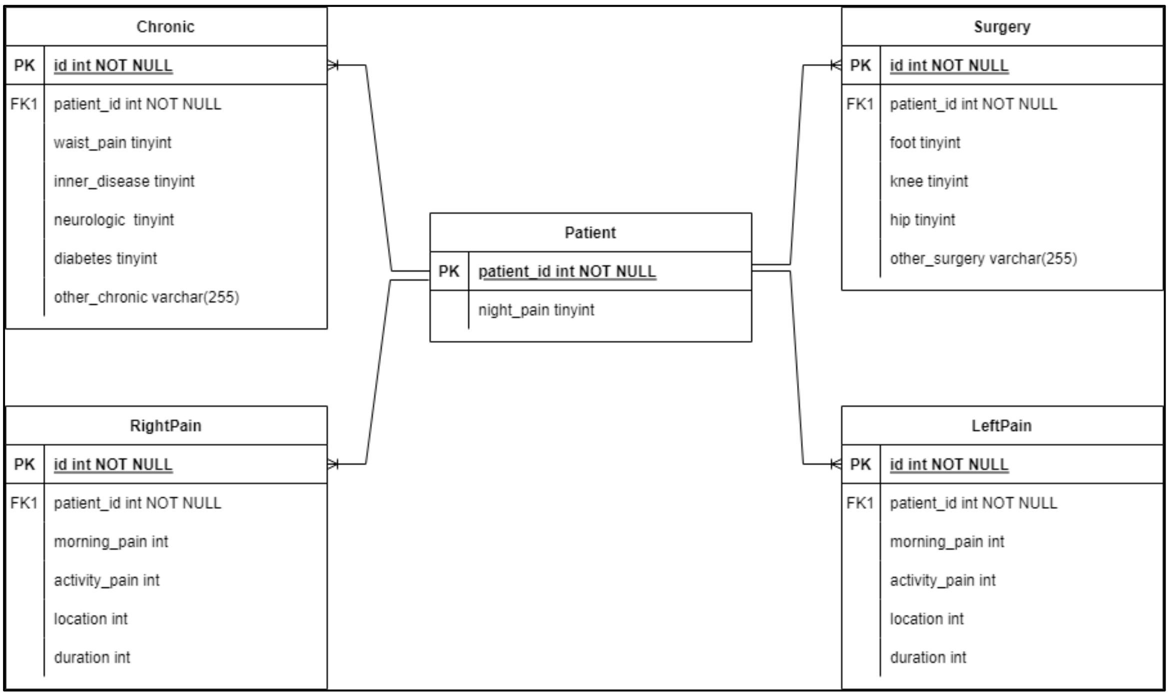


Figure 5.5. Entity Relationship Diagram of local database

Furthermore, videos taken by the user will be saved in the file system of the mobile device. In case of connection problems, users do not lose the data, and this leads to the application being easier to use.

5.1.3. Backend

The developed mobile application and related mobile device components are shown on Figure 5.6. The mobile application is a software which performs the identified tasks in previous sections by utilizing the components of the mobile device. In order to manage the data the user provides, the mobile application creates queries to insert data into the local database and retrieve data from it. For the ROM measurement case, it guides the users to the camera component in the necessary sections and enables them to take videos. These are saved in the file system of the device in mp4 (Motion Pictures Experts Group Layer 4 format) format. This format is preferred because it is the default option for devices with Android OS and also provides compression to the files. Therefore, there is no need to convert the video file which requires more resources. All the data that is collected with the mobile component and sent to the cloud component is converted into JSON objects by the mobile application using JSON notation. These created JSON objects are calculated and transmitted via the hardware control unit in the mobile device. This unit provides routing of the data via mobile data or WiFi options regarding the channel suitable for the user's situation. On the other hand, the data coming from the cloud component to the mobile component is also directed to the mobile application through this unit.

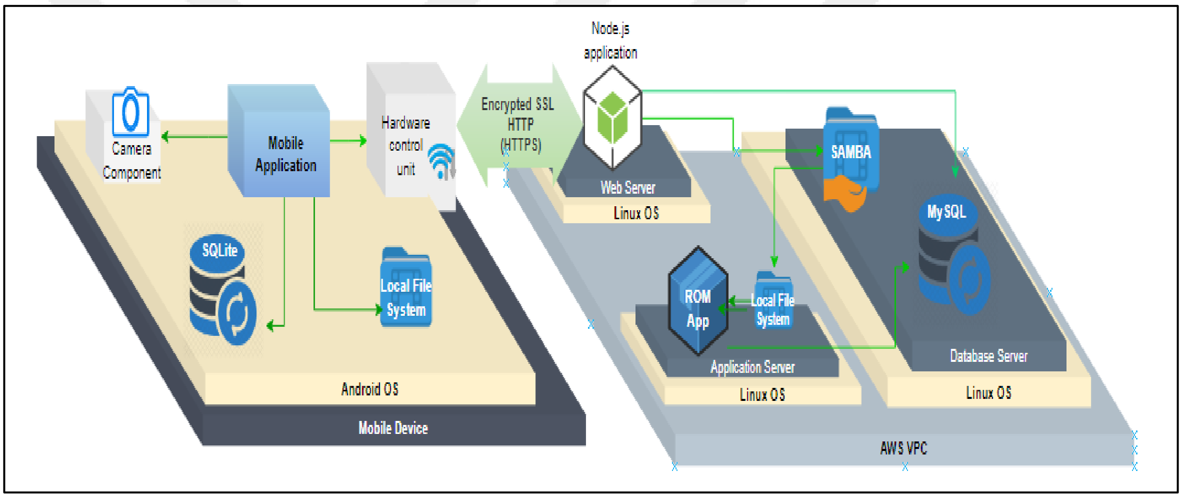


Figure 5.6. Detailed architechture diagram

5.2. CLOUD COMPONENT

Cloud component runs on Amazon Web Services (AWS). It consists of three EC2 Ubuntu servers which are functionally a web server, an application server, and a database server in a virtual private cloud (VPC). Amazon Web Server’s t2.micro product is used for web server and database server while application server is running on t2.medium. The details of server virtual machine type will be discussed in the following section. Furthermore, in this section, implementation of these three servers are described.

5.2.1. Web Server

As mentioned before this server is running on AWS t2.micro which has a capacity of 1 virtual CPU (a hyperthread of an Intel Xeon core) and the memory volume is 30 GiB. It runs Ubuntu 20.04 operating system. An Nginx web server is running on this instance and the server application is developed using Node.js framework.

Considering functionality of this server as seen in Figure 5.6, it welcomes the data packets coming from the mobile component to the cloud component. The server is accessed from "romm.persystlab.org" hostname. Because it is the interface for cloud components it has a static IP associated with the hostname.

For authentication functionality, the server application listens to "POST" requests for the register process and "GET" requests for the login process from "/auth" path. When the registration packet is received from the mobile component, it includes user information and patient information in one JSON object. The application parses the JSON and creates two queries to insert the data in. Then, a response is created in a JSON format consisting of "status" and "message" keys and corresponding values. After a successful registration, the user sends a JSON object to login to the system. In this case, the application queries the database and obtains the registered values to compare. Response format created for this transaction is the same as the format of the registration. Moreover, in case of successful login, patient_id is written to the response as the value of the message.

For data storage, the mobile component sends all chronic, surgery and pain information as one JSON object to "/info" path with a post request. Corresponding patient_id is also added to this JSON object to identify the owner of the data. The server application parses the received JSON object and the JSON arrays included. Then, it creates an SQL query for inserting the data into the relevant table. In all cases, a response is sent in the same format, the content of which changes depending on the status.

For video demonstration, the server application streams the video from the "/video" path. On the other hand, the user uploads the video files to the same path with a "POST" request whose content-type is multipart/form-data. This multipart also includes patient_id and a side information which are added to the packet as a JSON object. Therefore, when the server application receives this packet, it stores the video into the shared folder with a name in format of patientID_side_timestamp.mp4. Furthermore, the name of the file is written to the process list which is located in the shared folder. Storing video and writing the name of the video to the list are performed as child-processes because the shared folder is accessed with sudo permission. After all processes are completed, the server application sends a response.

5.2.2. Database Server

The database server has the same hardware features as the web server. The operating system that this server runs upon is also Ubuntu 20.04. Functionality of this server is different from Web Server. The infrastructure in which user, patient, medical history information and risk factor measurement data stored in the cloud component is a relational database management system namely MySQL DBMS. Considering all the data, an entity relationship design has been carried out and is presented in Figure 5.7.

The information collected during user registration (Figure 5.2a) is stored in the user table and patient table in the database. There is a one-to-one relationship between these two tables. Data about users' chronic diseases is kept in the chronic table and surgery/injury information (Figure 5.3a) is kept in the surgery table. Due to this information can be changed through time so the relationship between the two tables is one-to-many (1-N). All pain information of the patient for both feet is stored in the pain table, different from the local database. Additionally the night pain is also stored in the pain table in the main database. The relationship between the pain table and the patient table is more than zero because the pain can be changed through time. It is needed to store the changing values because of monitoring the patient status. On the other hand, in the rom_measurement table, it is stored information namely side of the foot, path of the media files (Figure 5.4) and the results of the measurements performed on the videos. The media is uploaded at different times for each patient and it is required to keep both data to record measurements through time. Therefore, the relationship of the measurement results of the media in question with the patients is also arranged in a one-to-many way.

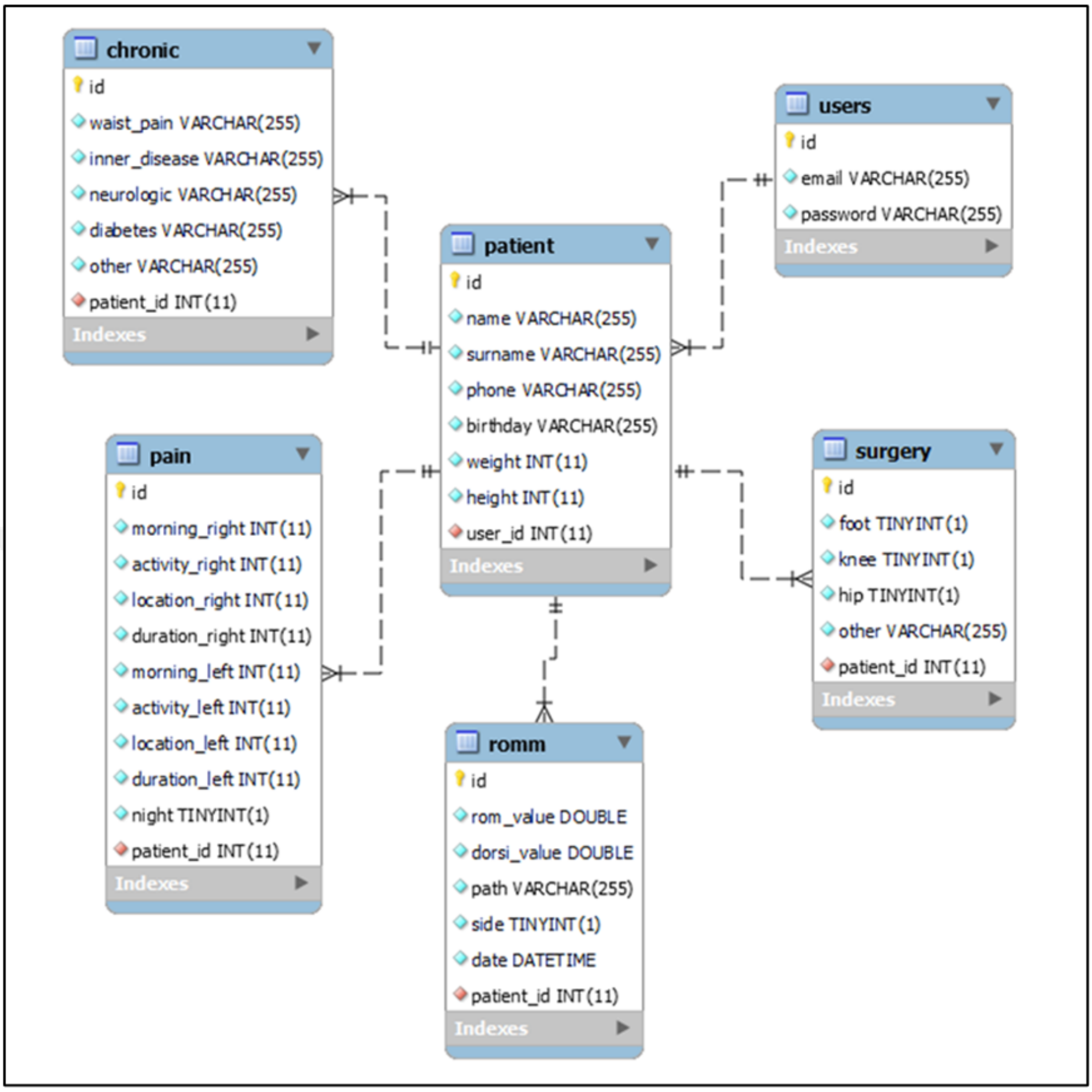


Figure 5.7. ER Diagram for the main database of the proposed system

Furthermore, the database server has a responsibility of archiving the video files sended by the user. However, the video files are received by the web server first and it is needed to transfer the files to the database server. This transfer was achieved with SAMBA protocol. In order to perform the data share with this protocol, samba server is configured on the database server. The shared folder is mounted to the /mnt directory of the other two servers. Therefore, whenever a video file is received by a web server, it is stored to the database server. The application server has access to the share folder through its /mnt folder.

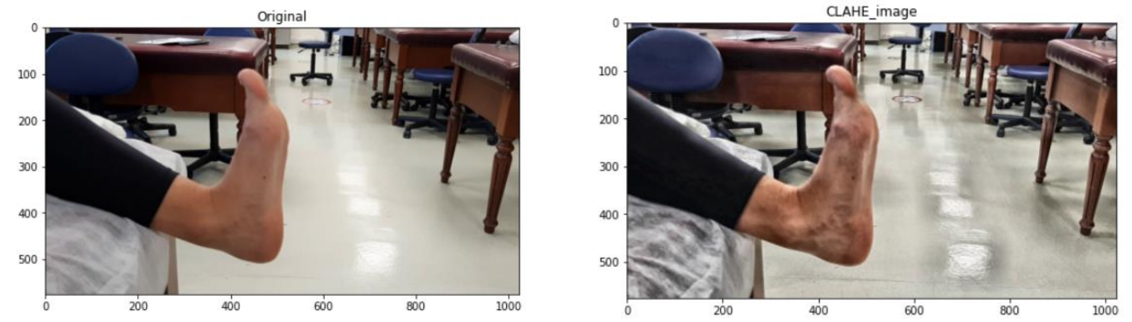
5.2.3. Application Server

Application Server has the capability of t3.micro product of AWS. Consequently, there are two vCPUs on this server. Memory volume is higher than the other two numerically 100 GiB. Deep Learning Amazon Machine Image running on this instance so it has the Ubuntu 18.04 operating system with necessary frameworks for deep learning applications. Also, it is compatible with NVIDIA CUDA which provides parallel processing and improves the performance of deep learning applications.

This server provides the service of measuring ankle ROM. To be able to do this, the shared folder is mounted to the filesystem of this server as a samba client. Because this configuration is reset when the instance is rebooted, a chronejob is added to the list to set the samba connection at reboot. When the connection is established with the shared folder the server reads the process list every 5 minutes. The reason behind this, output of the ROM measurement application is ready in approximately 2 minutes. Therefore, the ROM application is triggered by a chronejob, not by a request. When the filename is read from the process list, it is copied to the local file system to protect the original file. In the meantime, the corresponding line is marked to show the measurement for this file in progress. After these preparations the application developed for measuring ankle ROM is called. The output of the application is written to the database and an e-mail is sent to the user.

5.3. ANKLE ROM MEASUREMENT IMPLEMENTATION

It is aimed that the ankle ROM measurement application takes the videos from the file system as input and gives the ankle ROM range result as output. In order to make this measurement, a list of images will be extracted from each video by dividing the videos into frames with ffmpeg. Specifically, two frames are derived every second. The model developed for ROM measurement is designed to work on images. In this model, the region expressing the foot shape in each image is determined by using multiple segmentation algorithms . Before segmentation, the luminosity enhancement is made for the frame due to the high segmentation performance. Using the OpenCV library, the color space of the frame was shifted from BGR(blue,green,red) to LAB(lightness, (red,green), (blue,yellow)). In this color space, CLAHE is applied to the L chanel of the frame to equalize the histogram of the lightness but limiting contrast (Figure 5.8b).



(a) (b)

Figure 5.8. CLAHE enhancement

From the resulting frame, foot shape is detected with the Mask RCNN [59]algorithm trained with the COCO dataset [63] is applied. is applied. Using this algorithm, a mask of the region, a bounding box, and a prediction rate are created. In our case, the mask may not be precise enough because the input video is created via a mobile device camera. To improve the measurement, skin color filters are applied to the masked images via color segmentation. Skin color mask is detected on both HSV and YCrYcb color spaces. Channel ranges used for both spaces are in Table 5.1.

Table 5.1. Color ranges used for skin detection

Color Space	**Minimum**	**Maximum**
HSV	(1, 30, 30)	(33, 235, 235)
YCbCr	(5,133,77)	(235,200,150)

These two skin color masks are combined with bitwise and operation to capture similarities of both masks. Gaussian blur and dilation are applied to the resulting mask to close the gaps in the masked area. After all segmentations and operations foot form is attained as the image as in Figure 5.9c.

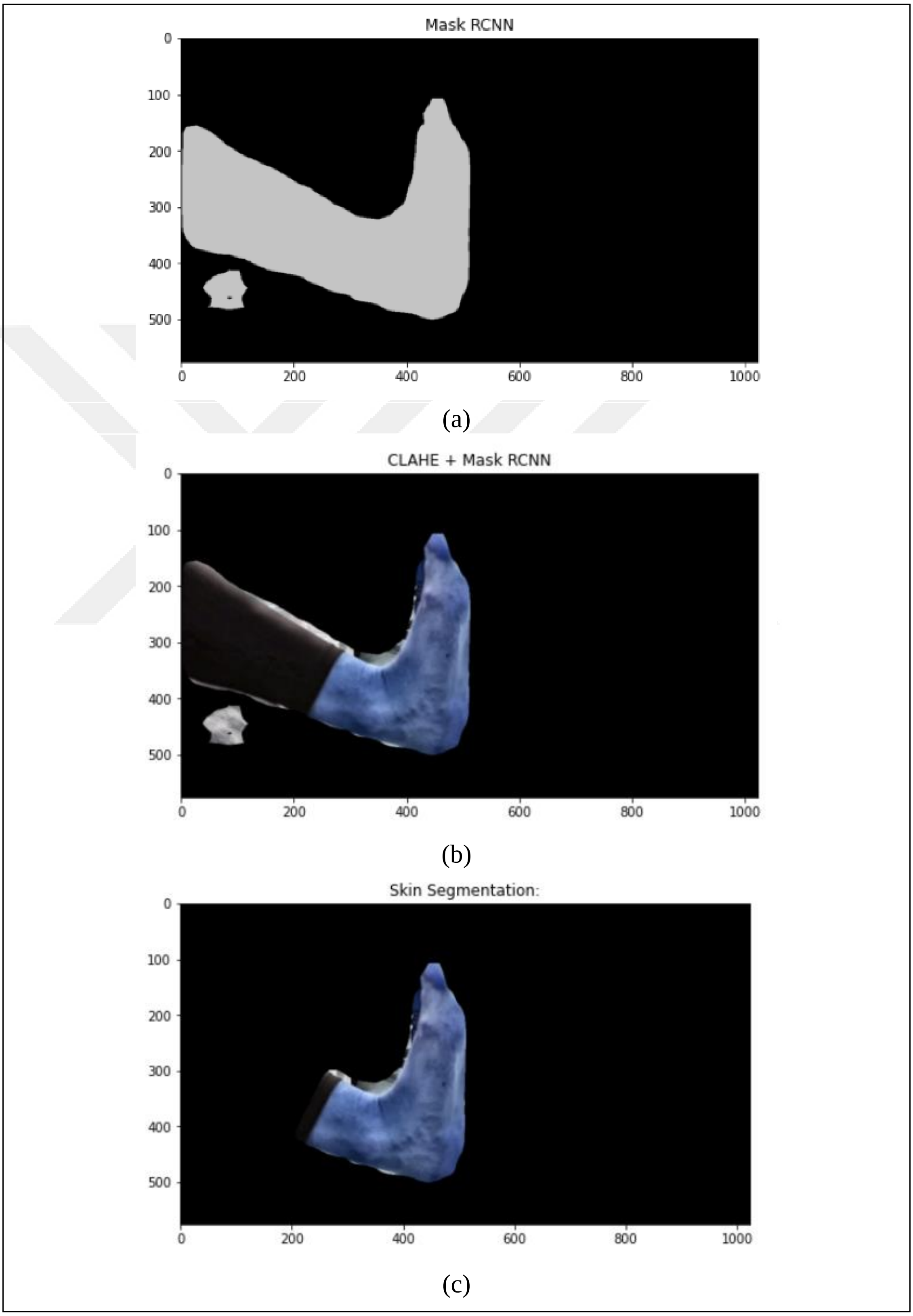


Figure 5.9. Mask RNN and Skin segmentation in BGR color space

To calculate ankle ROM, extreme points are determined on the foot image by grabbing contour on the frame. Rightmost, Leftmost, Bottom, and Top points are identified applying contour technique to the image in Figure 5.9b. To find the ankle point, Guo-Hall thinning algorithm is applied to the image in Figure 5.9c. This algorithm creates fork points in the corners (Figure 5.10a). In our case, which is the shape of the foot, the fork occurs approximately on the ankle. To detect this intersection point, a list of 9x9 filters is created via combination. In each filter, the center pixel and 3 or 4 pixels of the border are occupied. The extreme points and ankle point is seen in Figure 5.10b. Color codes for points are in Table 5.2.

Table 5.2. Color Codes for extreme points

Extreme Point	**Color**
Rightmost	Green
Leftmost	Blue
Bottom	Yellow
Top	Red
Middle	Orange

The location of these extreme points can change in different positions of the foot and leg. For example, when plantar flexion is performed, the location of the top point may not be the toe, unlike Figure 5.10b. In order to regulate this situation, a decision mechanism was created. This mechanism decides which points will be included, based on their distances to each other. Again, considering the plantar flexion position, the top point maybe somewhere on the leg in Figure 5.10b. In this case, the distance between the top and the leftmost point is smaller than the distance between the top and rightmost point. It is seen that the top point may not be appropriate for calculating plantarflexion angle. The angle between the leftmost, bottom and rightmost points is closer to the actual value than the middle and top points.

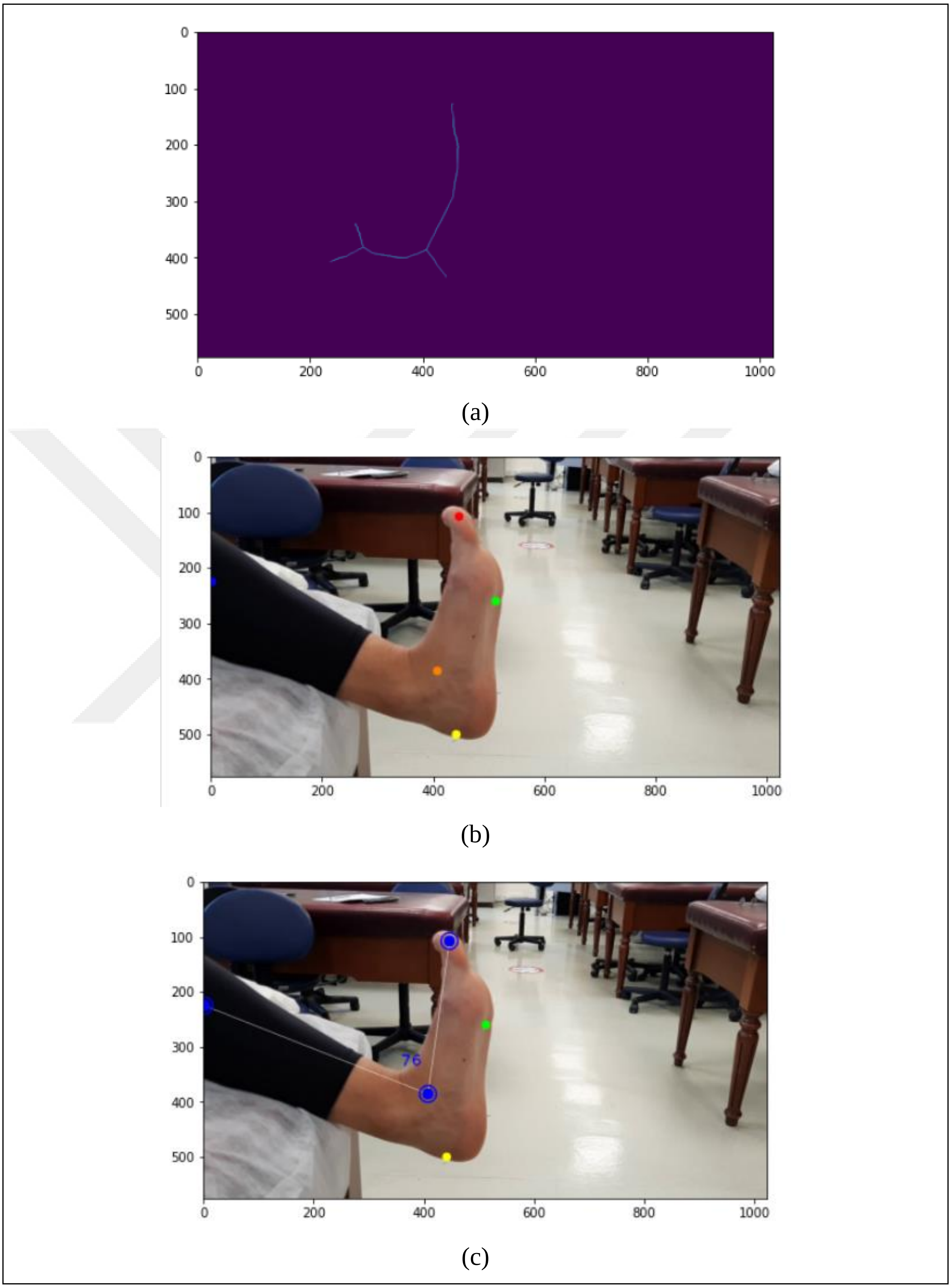


Figure 5.10. Finding skeleton, detecting extreme points, calculating angle

6. TEST AND RESULTS

In this section, description of the test environment, statistical analysis of test and usability evaluation of the system are presented. All data analysis given in the section are calculated with the SPSS tool.

6.1. USABILITY

Usability test is conducted among 15 people. The descriptive data of the participants including age, height and weight are shown in Table 6.1.

Table 6.1. Descriptive data of usability test participants

N=15	Minimum	Maximum	Mean ±SD
Age (years)	18	24	20.33±1.34
Height (cm)	157	183	171.18±8.91
Weight (kg)	47	91	64.93±13.82

Usability testing was carried out in accordance with retrospective probing conditions. In order to collect successful task completion and time on task data, all subjects are assigned to perform three tasks; (i)register/login, (ii)answer medical information, and (iii)ROM measurement. An introductory document explaining how to complete each task is given to the subject before the test starts (APPENDIX A). The completion time is measured for each task. After the test, a survey is conducted in order to evaluate the subjective data like the user aesthetics, learnability,satisfaction and efficiency of the system. Also, SUS questions are added to the survey to detect the usability score of the system (APPENDIX B).

First finding is the usability score which is calculated according to its definition. Average score is found as 76.5 /100 which means that the application usability is high and participants acclaim the system. Furthermore, it is questioned whether SUS score affects the difference between measurements taken by the system and the physiotherapist. After regression analysis, it is found that the SUS score does not affect the measurement validity.

In order to facilitate the data of the subjective statements, a reliability and internal consistency analysis is performed. Chronbach's Alpha analysis is chosen to depict the reliability and internal consistency of the statements. Results are found greater than 0.7 so that all analyzed subjective expressions were internally consistent for measuring the defined concepts.

Case Processing Summary

		N	%
Cases	Valid	14	93,3
	Excluded[a]	1	6,7
	Total	15	100,0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,882	,887	5

Item Statistics

	Mean	Std. Deviation	N
aestetics1	4,50	,650	14
aestetics2	3,93	,917	14
aestetics3	4,36	,633	14
aestetics4	4,00	,679	14
aestetics5	4,43	,646	14

(a)

Case Processing Summary

		N	%
Cases	Valid	15	100,0
	Excluded[a]	0	,0
	Total	15	100,0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,780	,814	6

Item Statistics

	Mean	Std. Deviation	N
learnability1	1,4000	,50709	15
learnability2	2,1333	1,18723	15
learnability3	1,6000	,91026	15
learnability4	1,7333	1,03280	15
SUS10	2,2667	1,22280	15
learnability5	1,7333	,79881	15

(b)

Case Processing Summary

		N	%
Cases	Valid	15	100,0
	Excluded[a]	0	,0
	Total	15	100,0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,741	,755	3

Item Statistics

	Mean	Std. Deviation	N
efficiency1	2,47	1,125	15
efficiency2	2,00	,756	15
effieciency3	1,73	1,033	15

(c)

Figure 6.1. Cronbach's Alpha Reliability test results for user aestetics, learnability and efficiency

After finding consistency of the statements, it is deduced that the system has approximately moderate efficiency and outstanding aesthetics and learnability features by analysing the mean values of the corresponding data. These mean values are reported in Table 6.2.

Table 6.2. Mean values of subjective statement data

	User aestetics	Learnability	Efficiency
Scores	4.26/5	4.28/5	2.93

To investigate the efficiency data further, the correlation between efficiency and time to perform the task is examined. The correlation is inspected to apply Pearson’s correlation test because it is found that two data is normal. It appears that there is a correlation between one of tasks and the efficiency. As seen in Table 6.3, the correlation is negative. It implies that more than half of the users think that application is not efficient while they do not spare time to complete the answering medical information task.

Table 6.3. Pearson Correlation value between efficiency and Tasks

	Register/Login task	Answering Medical Information task	ROM measurement task
Efficiency mean	Cannot found any significant correlation	-0.523	Cannot found any significant correlation

6.2. ROM MEASUREMENT

In this section, the analysis about ROM measurement are presented. The ROM measurement is calculated for 8 people of 16 feet Measurements are taken in a physiotherapy laboratory environment with experienced physiotherapists. Different from usability test, we have lesser number of participants. Gender of the participants also varies. Participant’s more detailed descriptive data are presented in Table 6.4. It shows mean age, mean height and mean weight with the standard deviations.

Table 6.4. Descriptive data of participants

N=8	**Minimum**	**Maximum**	**Mean ±SD**
Age (years)	18	24	20.68±1.72
Height (cm)	157	183	170.27±9.21
Weight (kg)	47	91	65.09±14.52

The measurement performed by the developed system is compared with physiotherapist goniometric measurement. The normality of two data is analysed and they found normal. Table 6.5 shows the normality test results for two data groups.

Table 6.5. Tests of Normality ROM Measurements

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
romApp	.160	16	.200*	.904	16	.092
romUG	.159	16	.200*	.954	16	.563
*. This is a lower bound of the true significance.						
a. Lilliefors Significance Correction						

Because Shapiro-Wilk is not significant, it can be accepted that these two data have normal distribution. Also, Skewness and Kurtosis values are between -1,5 and +1,5, the normality of two data groups is verified (Table 6.6). Although having a normal data distribution lets us perform Pearson's Correlation test results, any correlation is not found between ROM range measurements.

Table 6.6. Descriptives of the ROM measurement data

		Statistic	Std. Error
romApp	Skewness	.844	.564
	Kurtosis	-.062	1.091
romUG	Skewness	.574	.564
	Kurtosis	.050	1.091

Furthermore, dorsiflexion and plantar flexion data are analysed separately. However, DF measured by the physiotherapist (dfUG) are significant according to Shapiro-Wilk and Kolmogorov-Smirnov. Skewness and Kurtosis values also show that we cannot perform Pearson correlation test on these data (Table 6.7). Therefore, Spearman correlation is performed for this data. However, statistically significant correlation is not found between two measurement techniques.

Table 6.7. Tests of Normality for plantar flexion and dorsiflexion data

	Kolmogorov-Smirnov[a]			Shapiro-Wilk				
	Statistic	df	Sig.	Statistic	df	Sig.	Skewness	Kurthosis
dfApp	.151	16	.200*	.929	16	.235	-.984	65.9475
pfApp	.141	16	.200*	.961	16	.682	-.233	-.975
dfUG	.297	16	<.001	.815	16	.004	.431	33.75
pfUG	.187	16	.140	.939	16	.333	.421	-.675
a. Lilliefors Significance Correction								

As a last analysis, agreement on the data group is examined with the Bland & Altman plot. First, ROM values measured by app and physiotherapist are investigated. Mean and differences of data groups are found. Then, we should test the difference value with one sample t-test to state that there is no statistically difference between measuring with two methods. Table 6.8 shows the results of one sample t-test for ROM difference(difference), plantar flexion difference (pdiff) and dorsiflexion difference (ddiff).

Table 6.8. One Sample t-test for difference values of ROM, PF and DF

	Test Value = 0						
	t	df	Significance		Mean Difference	95 percent Confidence Interval of the Difference	
			One-Sided p	Two-Sided p		Lower	Upper
difference	1.215	15	.122	.243	15.10750	-11.3967	41.6117
pdiff	1.671	15	.058	.115	32.19750	-8.8622	73.2572
ddiff	-1.203	15	.124	.247	-16.84000	-46.6680	12.9880

From the definition of Bland & Altman plot, upper and lower bounds of confidence intervals are added to the plot. Figure 6.3, 6.4 and 6.5 show the B&A plots of ROM measurement, dorsiflexion measurement and plantar flexion measurement. As a result, it is seen that all data are within the confidence interval. Therefore, we can conclude that measurements taken by the system and the physiotherapist are in agreement when measuring dorsiflexion, plantarflexion and ROM range.

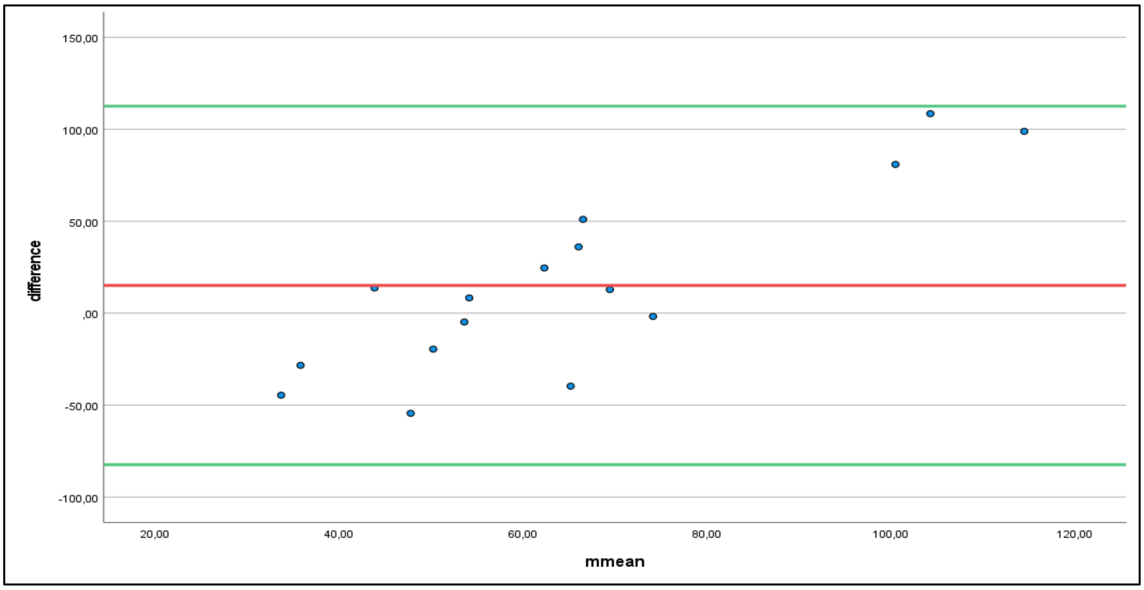


Figure 6.2. Bland & Altman Plot for ROM range measurement

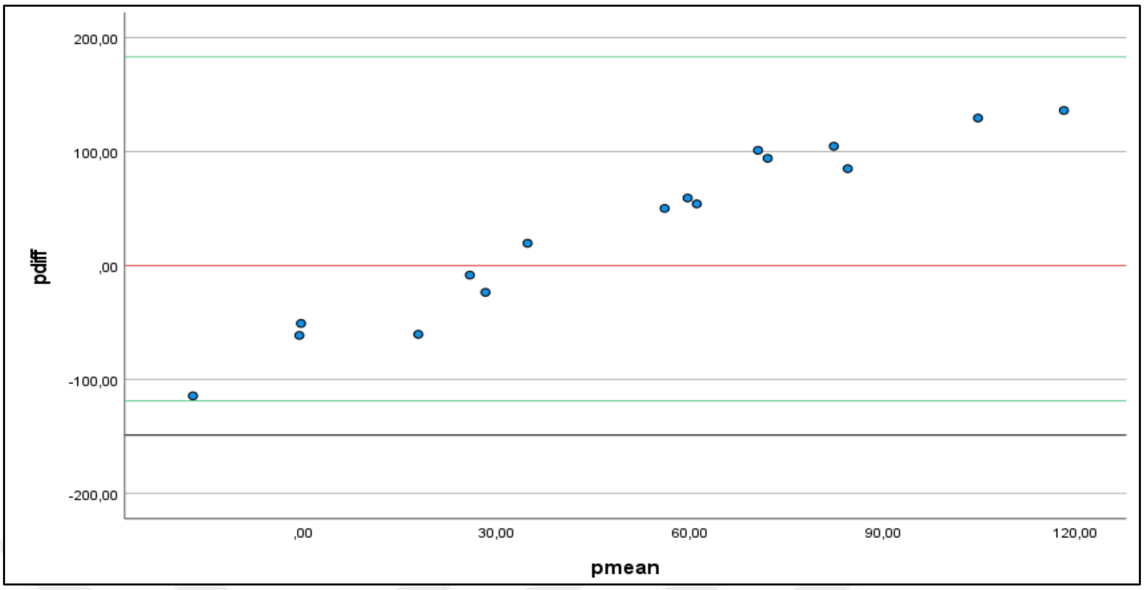


Figure 6.3. Bland & Altman Plot for plantar flexion range measurement

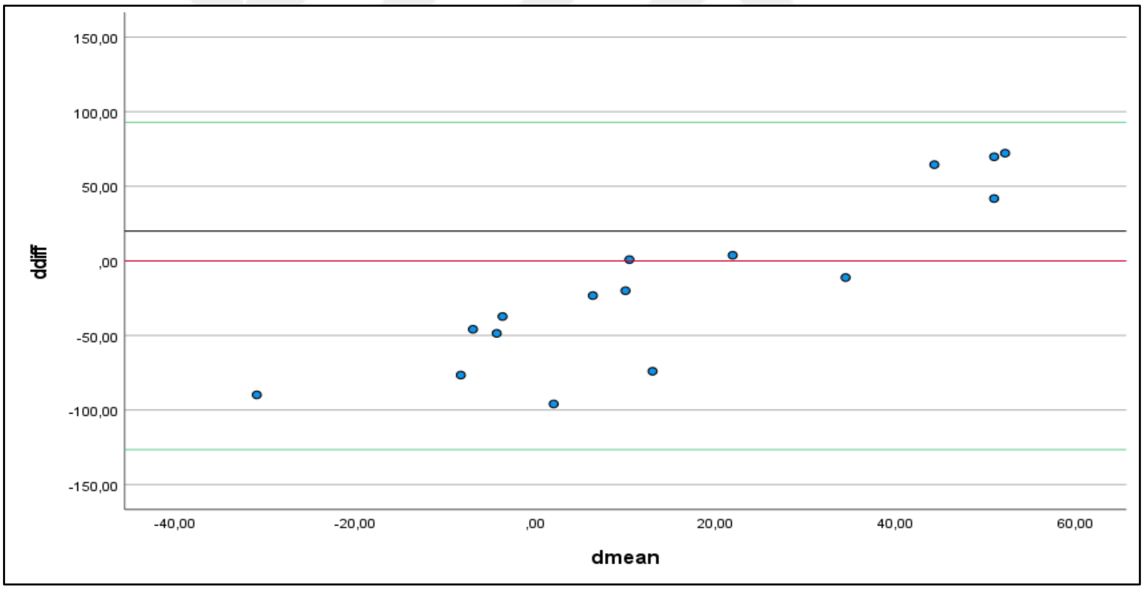


Figure 6.4. Bland & Altman Plot for dorsiflexion measurement

7. DISCUSSION AND CONCLUSION

In this thesis, a system is developed to detect the ankle ROM from a video containing only the foot region. Because the foot has a very low number of features on the sides, it is hard to detect only the foot in a video or in an image. Even if we consider the most well-known human pose estimation algorithms, it can be clearly seen that they fail to detect the foot region in a given picture frame when the image does not contain the rest of the body. For example, BlazePose [72], which is a lightweight human pose estimation algorithm, can detect feet's key points in an image. However, if the source image does not contain the full feet - i.e. also contain the knee itself, even the BlazePose algorithm cannot detect the key of a foot properly. Figure 7.1a lists the predefined key points of the BlazePose; and also Figure 7.1b/c provides screen captures of BlazePose results when executed on a pre-recorded video. The big red points on images shown in Figures 7.1b and 7.1c correspond to the ankle point -27th point from the predefined list.

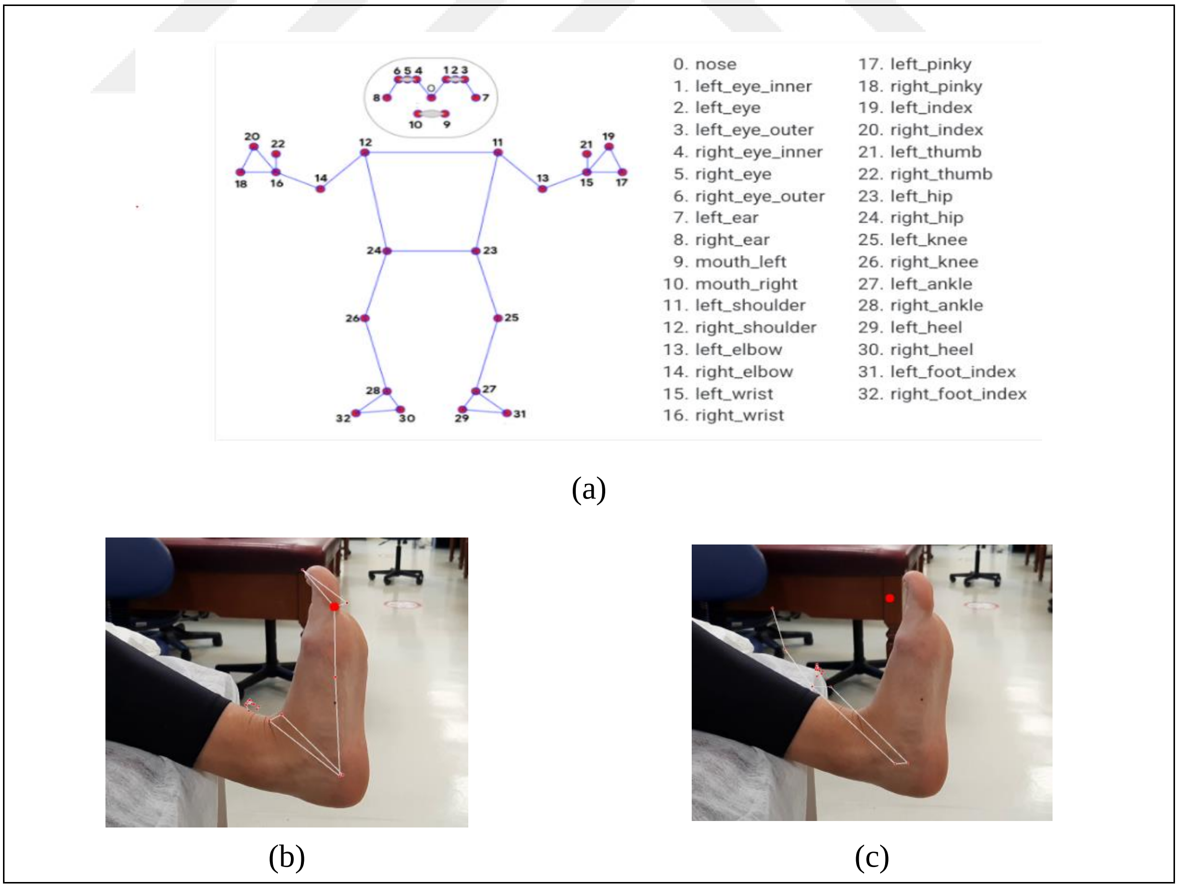


Figure 7.1. BlazePose keypoints and implementation on the input images

In addition, the OpenPose algorithm is also used to determine whether a foot area or person can be found in a source video. As can be seen in the following example, even the OpenPose can have issues in detecting a person in an image or video shot too close-by to person or body part (Figure 7.2).

Figure 7.2. Unsuccessful detection of human using OpenPose

On the other hand, the algorithm developed in the scope of this thesis enables detecting the foot region independent of background and other body parts and calculating the ankle ROM angle. In order to measure the validity of the algorithm and usability of the system, a real-life study was conducted. Eight people participated in this study, and their ankle ROM ranges were calculated using both the developed system and the physiotherapists' usual measurement method. In addition to the above objective evaluations, the system is also evaluated by the end users from a subjective perspective. Hence, the participants are asked to fill a usability questionnaire after using the prototype system.

Initially, the first question attacked by this study was whether the prototype system's and the physiotherapist's calculations had a normal distribution. Pearson and Spearman correlation tests are applied to the data and it was found to be normal, with no statistically significant correlation. Since the data set was small, the analysis was also refined with the Bland & Altman plot. To apply B&A, first of all, the data is analyzed with a one-sample t-test and the difference between the data sets are proven to be different from 0. Thus, it has emerged that a B&A plot can be performed. In this plot, it is determined that all data is within the confidence interval. This situation can be interpreted as there is an agreement between the two data sets.

Since the data is provided by the users, it is thought that the usability may also be effective in the measurements. For this reason, various analyses related to usability are made. The first of these analyses is the usability score measurement. This measurement result yielded to be quite high as 76.5/100. Additionally, efficiency, learnability and user aesthetics factors were also questioned, and it was determined that the measurements made with these factors were higher than Cronbach's Alpha. The efficiency score was low compared to other statements, so it is compared to users' task completion times. The findings of this comparison shows that completion times do not support user responses. In this section, it was determined that 52 percent of the users gave low scores even though they performed the task very quickly. Finally, it is found that usability score does not affect the difference between system's and physiotherapist's measurements.

As a result, our objective and subjective test show that the system can be used to obtain the ankle ROM angle by inexperienced users outside the clinical setting. Especially in the home environment, videos provided by inexperienced users may have problems such as human skin-like background or close-by shoots. As explained above, while OpenPose cannot find a person in test videos containing only foot region, the developed system can find a person object from same test videos with 100 percent rate. When angle measurement was made with BlazePose, it was seen that it could not place the key points correctly and could not calculate the angle in between. On the other hand, the developed system is able to measure angles for each test video by placing key points

Although the system has limited accuracy in measuring ankle ROM against UG, it has potential to be more accurate. At this point, it is remembered that goniometric measurements are performed while one arm of the UG is pointing at the head of the tibia while the other arm should point at the fifth metatarsal. However, the tibia head is not seen in the videos sent to the developed system. For this reason, the most left pixel entering the frame is used instead of the tibia head. This ensures that the margin of error changes depending on the incoming input. It is thought that the margin of error can be reduced by conducting a study that can predict the tibial head in future studies.

It has also been observed that some videos are not angularly aligned correctly. Such situations also make it difficult to determine the angle correctly. In order to solve this problem, it is possible to shoot video while the phone angle is correct by using the sensors of the phones such as the gyroscope and accelerometer specified in the methods section. Thus, it is expected that the input quality will increase.

To conclude, the system developed in the scope of this thesis can be used as a remote tool for ankle ROM range. For future work, improving the accuracy can be achieved by enhancing the identified problems. With a larger dataset and enhancements, usability study can be reconducted to evaluate the effects of usability on the validity of ROM.

REFERENCES

1. Gill TK, Menz HB, Landorf KB, Arnold JB, Taylor AW, Hill CL. Predictors of foot pain in the community: the North West Adelaide health study. *J Foot Ankle Res*. 2016 Jul 13;9(1):23.

2. Garrow AP, Silman AJ, Macfarlane GJ. The Cheshire Foot Pain and Disability Survey: a population survey assessing prevalence and associations. *Pain*. 2004 Jul 1;110(1):378–84.

3. WebMD. *Webster's New World Medical Dictionary*, 3rd Edition. Houghton Mifflin Harcourt; 2009. 482 p.

4. Neumann DA. *Kinesiology of the Musculoskeletal System* - E-Book: Foundations for Rehabilitation. Elsevier Health Sciences; 2013. 750 p.

5. Murray MP, Kory RC, Clarkson BH, Sepic SB. Comparison of free and fast speed walking patterns of normal men. *Am J Phys Med*. 1966 Feb;45(1):8–23.

6. Lynch SA. Assessment of the Injured Ankle in the Athlete. *J Athl Train*. 2002;37(4):406–12.

7. Shultz S, Olszewski A, Ramsey O, Schmitz M, Wyatt V, Cook C. A Systematic Review Of Outcome Tools Used To Measure Lower Leg Conditions. *Int J Sports Phys Ther*. 2013 Dec;8(6):838–48.

8. Kim BH, Hong SH, Oh IW, Lee YW, Kee IH, Lee SY. Measurement of Ankle Joint Movements Using IMUs during Running. *Sensors*. 2021 Jan;21(12):4240.

9. Croxford P, Jones K, Barker K. Inter-tester comparison between visual estimation and goniometric measurement of ankle dorsiflexion. *Physiotherapy Theory and Practice*. 1998 Jan;14(2):107–13.

10. Venturni C, André A, Aguilar B, Giacomelli B. Reliability of two evaluation methods of active range of mo- tion in the ankle of healthy individuals. *Acta Fisiatr*. 2006 Jan 1;13.

11. Bennell KL, Talbot RC, Wajswelner H, Techovanich W, Kelly DH, Hall AJ. Intra-rater and inter-rater reliability of a weight-bearing lunge measure of ankle dorsiflexion. *Aust J Physiother*. 1998;44(3):175–80.

12. Konor M, Morton S, Eckerson J, Grindstaff T. Reliability of three measures of ankle dorsiflexion range of motion. *International journal of sports physical therapy*. 2012 Jun 1;7:279–87.

13. Worsley PR, Conington C, Stuart H, Patterson A, Bader DL. A randomised cross over study to evaluate the performance of a novel ankle dorsiflexion measurement device for novice users. *Journal of Foot and Ankle Research*. 2018 Jul 31;11(1):45.

14. Chen G, Zhou Z, Wang N, Wang Q. Range-of-motion measurement with therapist-joined method for robot-assisted ankle stretching. *Robotics and Autonomous Systems*. 2017 Aug 1;94:34–42.

15. Ahmadian N, Nazarahari M, Whittaker JL, Rouhani H. Instrumented triple single-leg hop test: A validated method for ambulatory measurement of ankle and knee angles using inertial sensors. *Clinical Biomechanics*. 2020 Dec 1;80:105134.

16. Wellmon RH, Gulick DT, Paterson ML, Gulick CN. Validity and Reliability of 2 Goniometric Mobile Apps: Device, Application, and Examiner Factors. *J Sport Rehabil*. 2016 Dec;25(4):371–9.

17. Morales CR, Lobo CC, Sanz DR, Corbalán IS, Ruiz BR, López DL. The concurrent validity and reliability of the Leg Motion system for measuring ankle dorsiflexion range of motion in older adults. *PeerJ*. 2017 Jan 3;5:e2820.

18. Vohralik SL, Bowen AR, Burns J, Hiller CE, Nightingale EJ. Reliability and validity of a smartphone app to measure joint range. *Am J Phys Med Rehabil*. 2015 Apr;94(4):325–30.

19. Williams CM, Caserta AJ, Haines TP. The TiltMeter app is a novel and accurate measurement tool for the weight bearing lunge test. *Journal of Science and Medicine in Sport*. 2013 Sep 1;16(5):392–5.

20. Wang KY, Hussaini SH, Teasdall RD, Gwam CU, Scott AT. Smartphone Applications for Assessing Ankle Range of Motion in Clinical Practice. *Foot & Ankle Orthopaedics*. 2019 Jul 1;4(3):247301141987477.

21. Awatani T, Enoki T, Morikita I. Inter-rater reliability and validity of angle measurements using smartphone applications for weight-bearing ankle dorsiflexion range of motion measurements. *Physical Therapy in Sport*. 2018 Sep 1;34.

22. Romero-Franco N, Jiménez-Reyes P, González-Hernández JM, Fernández-Domínguez JC. Assessing the concurrent validity and reliability of an iPhone application for the measurement of range of motion and joint position sense in knee and ankle joints of young adults. *Physical Therapy in Sport*. 2020 Jul 1;44:136–42.

23. Cunha AB, Babik I, Harbourne R, Cochran NJ, Stankus J, Szucs K, et al. Assessing the Validity and Reliability of a New Video Goniometer App for Measuring Joint Angles in Adults and Children. *Archives of Physical Medicine and Rehabilitation*. 2020 Feb 1;101(2):275–82.

24. Zhou H, Hu H. Human motion tracking for rehabilitation—A survey. *Biomedical Signal Processing and Control*. 2008 Jan 1;3(1):1–18.

25. Ma Y, Mithraratne K, Wilson NC, Wang X, Ma Y, Zhang Y. The Validity and Reliability of a Kinect v2-Based Gait Analysis System for Children with Cerebral Palsy. *Sensors (Basel)*. 2019 Apr 7;19(7):1660.

26. Moreira R, Teles A, Fialho R, dos Santos TCP, Vasconcelos SS, de Sá IC, et al. Can human posture and range of motion be measured automatically by smart mobile applications? *Medical Hypotheses*. 2020 Sep 1;142:109741.

27. Ferrari V, Marin-Jimenez M, Zisserman A. Progressive search space reduction for human pose estimation. *IEEE Conference on Computer Vision and Pattern Recognition*. 2008. p. 1–8.

28. Andriluka M, Roth S, Schiele B. Pictorial structures revisited: People detection and articulated pose estimation. *IEEE Conference on Computer Vision and Pattern Recognition*. 2009. p. 1014–21.

29. Fihl P, Moeslund TB. Pose Estimation of Interacting People using Pictorial Structures. *7th IEEE International Conference on Advanced Video and Signal Based Surveillance*. 2010. p. 462–8.

30. Newell A, Yang K, Deng J. Stacked Hourglass Networks for Human Pose Estimation. In: Leibe B, Matas J, Sebe N, Welling M, editors. *Computer Vision – ECCV 2016. Cham: Springer International Publishing*; 2016. p. 483–99. (Lecture Notes in Computer Science).

31. Cao Z, Simon T, Wei S-E, Sheikh Y. Realtime Multi-person 2D Pose Estimation Using Part Affinity Fields. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017. p. 1302–10.

32. Yang W, Li S, Ouyang W, Li H, Wang X. Learning Feature Pyramids for Human Pose Estimation. *IEEE International Conference on Computer Vision (ICCV)*. 2017. p. 1290–9.

33. Chen Y, Wang Z, Peng Y, Zhang Z, Yu G, Sun J. Cascaded Pyramid Network for Multi-person Pose Estimation. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2018. p. 7103–12.

34. Xiao B, Wu H, Wei Y. Simple Baselines for Human Pose Estimation and Tracking. *Computer Vision – ECCV 2018*. *Springer International Publishing*; 2018. p. 472–87. (Lecture Notes in Computer Science; vol. 11210).

35. Wei S-E, Ramakrishna V, Kanade T, Sheikh Y. Convolutional Pose Machines. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016. p. 4724–32.

36. Ota M, Tateuchi H, Hashiguchi T, Ichihashi N. Verification of validity of gait analysis systems during treadmill walking and running using human pose tracking algorithm. *Gait & Posture*. 2021 Mar 1;85:290–7.

37. Li Y, Wang C, Cao Y, Liu B, Tan J, Luo Y. Human pose estimation based in-home lower body rehabilitation system. *International Joint Conference on Neural Networks (IJCNN). Glasgow, United Kingdom: IEEE*; 2020. p. 1–8.

38. Themes UFO. Measurement Of Range Of Motion Of The Ankle And Foot [Internet]. Musculoskeletal Key. 2016 [cited 2022 Jan 11]. Available from: https://musculoskeletalkey.com/measurement-of-range-of-motion-of-the-ankle-and-foot/

39. Dickson D, Hollman-Gage K, Ojofeitimi S, Bronner S. Comparison of Functional Ankle Motion Measures in Modern Dancers. *Journal of Dance Medicine & Science*. 2012 Sep 1;16.

40. Gyroscope. In: Wikipedia [Internet]. 2022 [cited 2022 Jan 11]. Available from: https://en.wikipedia.org/w/index.php?title=Gyroscope&oldid=1063557582

41. Control of MEMS Coriolis Vibratory Gyroscopes. : 9.

42. Accelerometer and Gyroscopes Sensors: Operation, Sensing, and Applications [Internet]. [cited 2021 Dec 31]. Available from: https://www.maximintegrated.com/en/design/technical-documents/app-notes/5/5830.html

43. Tinder RF. *Relativistic Flight Mechanics and Space Travel: A Primer for Students, Engineers and Scientists*. Morgan & Claypool Publishers; 2007. 141 p.

44. Rindler W. *Essential Relativity: Special, General, and Cosmological*. Springer; 2013. 332 p.

45. Wang C, Li Y, Xiong Z, Luo Y, Cao Y. Lower Body Rehabilitation Dataset and Model Optimization. In: 2021 *IEEE International Conference on Multimedia and Expo (ICME)*. 2021. p. 1–6.

46. Zarit BD, Super BJ, Quek FKH. Comparison of five color models in skin pixel classification. *Proceedings International Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems In Conjunction with ICCV'99*. 1999. p. 58–63.

47. Birchfield S. Elliptical head tracking using intensity gradients and color histograms. *Proceedings 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 1998. p. 232–7.

48. McKenna SJ, Gong S, Raja Y. Modelling Facial Colour And Identity With Gaussian Mixtures. *Pattern Recognition.* 1998;31(12):1883–1892.

49. Sigal L, Sclaroff S, Athitsos V. *Estimation and prediction of evolving color distributions for skinsegmentation under varying illumination*. *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2000.)* 2000. p. 152–9 vol.2.

50. Albiol A, Torres L, Delp E. Optimum color spaces for skin detection. *Proceedings 2001 International Conference on Image Processing.* 2001. p. 122–4 vol.1.

51. Son Lam Phung, Bouzerdoum A, Chai D. A novel skin color model in YCbCr color space and its application to human face detection. *Proceedings International Conference on Image Processing. Rochester, NY, USA: IEEE*. 2002. p. I-289-I–292.

52. Menser B, Wien M. *Segmentation and tracking of facial regions in color image sequences*. 2000. p. 731–40.

53. Shaik KB, Ganesan P, Kalist V, Sathish BS, Jenitha JMM. Comparative Study of Skin Color Detection and Segmentation in HSV and YCbCr Color Space. *Procedia Computer Science*. 2015 Jan 1;57:41–8.

54. Kolkur S, Kalbande D, Shimpi P, Bapat C, Jatakia J. Human Skin Detection Using RGB, HSV and YCbCr Color Models. *Proceedings of the International Conference on Communication and Signal Processing 2016 (ICCASP 2016). Lonere, India: Atlantis Press;* 2017.

55. Min, Byong-seok et al. A Novel Method of Determining Parameters of CLAHE Based on Image Entropy. *International Journal of Software Engineering and its Applications 7* 2013. p. 113-120.

56. Felzenszwalb PF, Huttenlocher DP. Pictorial Structures for Object Recognition. *International Journal of Computer Vision*. 2005 Jan;61(1):55–79.

57. Yang Y, Ramanan D. Articulated pose estimation with flexible mixtures-of-parts. *CVPR 2011*. 2011. p. 1385–92.

58. Shi M, Cai H, Gao Y. Optimization of Human Pose Detection Based on Mask RCNN. *2021 2nd International Symposium on Computer Engineering and Intelligent Communications (ISCEIC)*. 2021. p. 273–7.

59. He K, Gkioxari G, Dollár P, Girshick R. Mask r-cnn. *Proceedings of the IEEE international conference on computer vision 2017* .pp. 2961-2969.

60. Fisher R, Perkins S, Walker A, Wolfart E. Hypermedia Image Processing Reference (HIPR). Artificial Intelligence - AI. 1996 Jan 1;

61. 1. Zhang TY, Suen CY. A fast parallel algorithm for thinning digital patterns. *Commun ACM*. 1984 Mar;27(3):236–9.

62. Guo Z, Hall RW. Fast fully parallel thinning algorithms. *CVGIP: Image Understanding*. 1992 May 1;55(3):317–28.

63. Lin, Tsung-Yi, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. Microsoft coco: Common objects in context. *European conference on computer vision*, pp. 740-755. Springer, Cham, 2014.

64. Andriluka M, Pishchulin L, Gehler P, Schiele B. 2D Human Pose Estimation: New Benchmark and State of the Art Analysis. *2014 IEEE Conference on Computer Vision and Pattern Recognitio. Columbus, OH, USA: IEEE;* 2014. p. 3686–93.

65. Wu J, Zheng H, Zhao B, Li Y, Yan B, Liang R, et al. AI Challenger : A Large-scale Dataset for Going Deeper in Image Understanding. *2019 IEEE International Conference on Multimedia and Expo (ICME)*. 2019 Jul;1480–5.

66. Bevan N, Carter J, Earthy J, Geis T, Harker S. New ISO Standards for Usability, Usability Reports and Usability Measures. *Proceedings, Part I, of the 18th International Conference on Human-Computer Interaction*. 2016. p. 268–78.

67. Running a Usability Test | Usability.gov [Internet]. [cited 2021 Dec 31]. Available from: https://www.usability.gov/how-to-and-tools/methods/running-usability-tests.html

68. Planning a Usability Test | Usability.gov [Internet]. [cited 2021 Dec 31]. Available from: https://www.usability.gov/how-to-and-tools/methods/planning-usability-testing.html

69. System Usability Scale (SUS) | Usability.gov [Internet]. [cited 2021 Dec 31]. Available from: https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html

70. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951 Sep 1;16(3):297–334.

71. Bland–Altman plot - Wikipedia [Internet]. [cited 2021 Dec 31]. Available from: https://en.wikipedia.org/wiki/Bland%E2%80%93Altman_plot 72.

72. Bazarevsky V, Grishchenko I, Raveendran K, Zhu T, Zhang F, Grundmann M. Blazepose: On-device real-time body pose tracking. *arXiv preprint arXiv:2006*.10204. 2020 Jun 17.

APPENDIX A: Usability Tasks

Task 1: Register and Login

1. Wait for the start command of the interviewer.
2. You will be greeted by the opening screen of the DigitalSteps application.
3. Register using the directions in the app.
4. Log into the system with the account you created.
5. When you see a welcome message on the screen, inform the interviewer.

Task 2: Answer medical information questions

1. Wait for the start command of the interviewer.
2. Answer the questions you come across in the application by following the directions until you reach the 5th screen of the application.
3. When you see the “NEH measurement” title on the screen, inform the interviewer.

Task 3: ROM measurement

1. Wait for the start command of the interviewer.
2. Upload a video to the system for both feet by watching the informative video.
3. Finish using the application using the “Son” key.
4. Inform the interviewer.

APPENDIX B: Usability Questionnaire

DigitalSteps Mobile Application Evaluation Form

Dear participant,

This survey study has been prepared to be used in the outputs of the “Digital Steps: Remote Evaluation and Rehabilitation of Back Foot Pains” project and the aforementioned project jointly carried out by Yeditepe University Computer Engineering and Yeditepe University Physiotherapy and Rehabilitation Department. The instrumentalization is purely academic and the data obtained from the study will be used for scientific purposes. The data you create by answering the questionnaire will be stored in accordance with the Law on Protection of Personal Data No. 6698. Thank you for your contribution to the work.

Section 1:

1. Name Surname

____

2. Age

- 18-25
- 26-40
- 41-64
- 65+

3. Education

- None
- Elementary
- High School
- University
- Msc/Phd

Section 2:

		Strongly Disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
1	The font of the text in this app was legible.	☐	☐	☐	☐	☐
2	The application contains instructions to facilitate use.	☐	☐	☐	☐	☐
3	It takes a long time to load visual elements using the app.	☐	☐	☐	☐	☐
4	The text size used in the app was very small.	☐	☐	☐	☐	☐
5	To perform a task, I have to remember a lot of information over several actions.	☐	☐	☐	☐	☐
6	The informative videos in the app take a long time to load.	☐	☐	☐	☐	☐
7	Learning to use the app was difficult.	☐	☐	☐	☐	☐
8	The images used in this application were beautiful.	☐	☐	☐	☐	☐
9	While using the application, I did not know what to write in the written fields.	☐	☐	☐	☐	☐
10	I find the app design beautiful.	☐	☐	☐	☐	☐
11	I had a hard time navigating to previous/next screens in the application.	☐	☐	☐	☐	☐
12	The icons and buttons used are beautiful and recognizable.	☐	☐	☐	☐	☐

Section 3:

		Strongly Disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
1	I think that I would like to use this system frequently.	☐	☐	☐	☐	☐
2	I found the system unnecessarily complex.	☐	☐	☐	☐	☐
3	I thought the system was easy to use.	☐	☐	☐	☐	☐
4	I think that I would need the support of a technical person to be able to use this system.	☐	☐	☐	☐	☐
5	I found the various functions in this system were well integrated.	☐	☐	☐	☐	☐
6	I thought there was too much inconsistency in this system.	☐	☐	☐	☐	☐
7	I would imagine that most people would learn to use this system very quickly.	☐	☐	☐	☐	☐
8	I found the system very cumbersome to use.	☐	☐	☐	☐	☐
9	I felt very confident using the system.	☐	☐	☐	☐	☐
10	I needed to learn a lot of things before I could get going with this system.	☐	☐	☐	☐	☐