AN ONLINE BALANCE TRAINING APPLICATION USING POSE ESTIMATION AND AUGMENTED REALITY

Amirhossein Etaat

A thesis
in
The Department
of
Engineering and Computer Science

Presented in Partial Fulfillment of the Requirements
For the Degree of Master of Applied Computer Science
Concordia University
Montréal, Québec, Canada

January 2022

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: **Amirhossein Etaat**

Entitled: **An Online Balance Training Application using Pose Estimation and Augmented Reality**

and submitted in partial fulfillment of the requirements for the degree of

Master of Applied Computer Science

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining commitee:

______________________________ Chair
Dr. Lata Narayanan

______________________________ Examiner
Dr. Tiberiu Popa

______________________________ Examiner
Dr. Tristan Glatard

______________________________ Supervisor
Dr. Marta Kersten-Oertel

Approved ______________________________
Dr. Dhrubajyoti Goswami, Graduate Program Director

January 7th, 2022 ______________________________
Dr. Mourad Debbabi, Dean
Gina Cody School of Engineering and Computer Science

Abstract

An Online Balance Training Application using Pose Estimation and Augmented Reality

Amirhossein Etaat

The evolution of digitally connected devices and artificial intelligence has opened the door for novel health and fitness applications that can be used by individuals at a time and in an environment convenient to them. The purpose of our research was to develop a platform that requires no additional hardware to provide an online balance training program. Balance exercises are often prescribed for healthy aging to keep the body active, improve balance and coordination, and prevent falls and injuries, as well as, for those doing rehabilitation after injuries or diseases such as stroke. We developed a simple web application (BaART: Balance Augmented Reality Trainer) that uses PoseNet to determine a user's location and pose to count the number of repetitions that were done successfully. Furthermore, we looked at how augmented reality, and specifically adding a virtual chair, might impact a user's sense of balance. In a study of 20 participants with and without balance disorders, we found that the developed system was easy to use and many would consider using such a system, particularly our older participants who spend more time at home. However, we also found that the virtual object (i.e. chair) was not used by most people. Furthermore, those with balance issues felt they required a real chair for balance and some even felt that the virtual object was distracting from the exercise. In the future, we plan to explore other uses of augmented reality, such as feedback on exercise quality, gaming features, and a virtual avatar trainer.

Acknowledgments

I am very pleased to have the opportunity to be acquainted with Dr. Marta Kersten-Oertel and do my research under her supervision. She is one of the most kind-hearted and generous professors I have ever met. I am deeply grateful to Dr. Marta Kersten-Oertel, my supervisor. Her moral support, patience, and assistance through my education are unforgettable. I could not have had a better mentor and role model. Besides my supervisor, I would like to thank my fellow labmates in the Applied Perception Lab for sharing their expertise and knowledge and providing a friendly environment to do research in. Last but not least, I would like to thank my supportive family and friends for supporting me spiritually throughout my life. I always believe nothing is impossible. I always believe nothing is impossible everything can be achieve your approach and effort will decide your result.

Contribution of Authors

I am the first author of the manuscript presented in Chapter 3 of this dissertation and as such have performed all of the methodological developments, experimental design, data collection and analysis of results. The contributions of co-authors include supervision, technical discussions and review of the manuscript.

- Chapter 3: An Online Application using Pose Estimation and Augmented Reality for Balance Training. Authors: Amirhossein Etaat, Negar Haghbin, Marta Kersten-Oertel; Contributions: Guarantors of integrity of the study: all authors; study and design concepts: A.E. N.H. and M.K.-O.; software development: A.E and; data collection: A.E.; data preparation and analysis: A.E. and M.K-O.; supervision: M.K.-O.; manuscript preparation: A.E.; manuscript revision: all authors; editing and final version: A.E., N.H and M.K.-O.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In 2018, the World Health Organization (WHO) reported that Canadians can expect to live 82.8 years, an increase of over 10 years from the 1960s. The *healthspan* or the period of life prior to chronic sickness or a degenerating condition, however, ends over a decade earlier than lifespan [1]. The advancement of connected and pervasive technologies combined with artificial intelligence is beginning to show the potential to increase healthspan and improve quality of life. For example, smartphones can allow the elderly to age in place, wearables can monitor health and inform caretakers if and when problems arise, and gaming and mixed reality technologies can motivate those aging to remain active.

One frequent event that leads to severe and even fatal injuries, and a rapid decrease of healthspan, is falling. In fact, according to the WHO, every year one out of three adults aged 65+ experiences a fall [2]. Among 80 year olds, this number increases to two out of three individuals. Reports indicate that falls make up for over half of all hospitalizations among the elderly [3]. Depending on the type and severity of injuries, patients' stays in the hospital can be anywhere from moderately long, e.g. twenty days for those going through hip fractures, to troubling, life-long stays. Moreover, even patients who are not hospitalized may suffer from subsequent daily life limitations due to their falls. Such limiting complications are famously categorized under post-fall syndrome [4] and include but are not limited to losing the ability to move easily, seclusion, isolation, and subsequently depression, and becoming highly dependent on receiving care from others. In general, falls can lead to unpleasant, and lasting physical and mental complications, putting considerable pressure on the healthcare system. It is, however, possible to decrease the potential of falls and the associated by using preventive measures, one of the most important of which is balance training and exercise.

1.1 Balance Training

Weakened postural and ambulation control and a higher injury and fall risk are expected outcomes due to declines in sensory functions in older individuals [5]. Gait and balance disorders can significantly influence daily life on a physical and psychological level and have been shown to be one of the main causes of mortality among older adults. For this reason, balance training has been suggested not only for fall prevention but also in rehabilitation from diseases such as stroke or Parkinson's that may cause balance disorders. Various factors induce balance disorders, including medical conditions, drugs, inner ear issues, or neurological disorders.

Health organizations, such as the American and British Geriatric Society (ABGS) [6] and the National Institute of Clinical Excellence (NICE) have published clinical instructions to prevent and evaluate older adults' falls [7]. Customized exercise programs for strength, balance, gait and coordination training have been shown to significantly decrease the likelihood of falls in older adults [8]. Furthermore, home-based exercise programs have been shown to decrease falls and the associated injuries by 35-45% [9]. These types of programs involve walking, muscle strengthening, and balance training, and can be specifically tailored for an individual by physiotherapists. However, a lack of experts for monitoring such sessions has limited these types of preventive measures. Furthermore, studies have found that many adults do not do follow exercises or training programs that can help or prevent ailments due to a lack of motivation, time and/or the high cost of physiotherapy, gyms, and aids [10].

1.2 Health applications for the elderly

Despite the popular belief that older adults are less proficient in using technology and technology-related products, they are much more familiar and able to work with technology compared to previous generations. In fact, the current older generation has a high tendency towards smartphones with high speed connections and smart functions [11] and a strong desire to learn how to use modern smartphones, particularly to learn health-related information [12]. This among other factors has led to many mobile and online health platforms directed towards the elderly.

Many factors are contributing to the disappearance of conventional physiotherapy and fitness training including the changing needs of the population, difficultly in accessing specialized facilities, not to mention the current pandemic. These and other factors have led to poor adherence of fitness programs. An online virtual training and fitness coach has the potential to resolve some of these challenges.

1.3 Mixed Reality

Mixed reality is defined as the blending of physical and digital worlds or elements. On one end of the spectrum is a purely digital world, i.e. virtual reality, and on the other the real physical world. Augmented reality, which lies within the spectrum, has been defined as adding virtual or digital elements into the real world. Mixed reality technologies are increasingly being used in fitness and health applications, for example with virtual trainers or augmented reality guidance applications. In fitness and exercising scenarios specifically, mixed reality is being used to help users better understand exercises and the proper way of doing them and/or for adding gaming elements to make exercising more enjoyable. In this thesis, we explore adding augmented reality elements in the form of a virtual chair into the real world (the live web camera feed images) to allow the user to "use" the virtual chair to improve their balance.

1.4 Motivation

Many people may not be able to do personalized exercise routines or receive physiotherapy due to both time and monetary costs. Yet, the number of older adults needing directed physical exercise is rising and thus creating a need to use new technologies to accelerate the adoption of fitness routines to help ease the pressure off the health cares system [13, 14]. Research studies have looked at rehabilitation and fitness applications for those diagnosed with balance disorders as a way to motivate individuals to adhere to their programs. However, many require specialized equipment (wearables, motion trackers, etc.) which are not always available to the end-user.

The purpose of the research presented in this thesis was to explore a simple low-cost online application for balance training that requires no specialized equipment. The developed prototype, "**Ba**lance **A**ugmented **R**eality **T**rainer" (BaART) can be used on any connected device with a web camera, i.e. a computer, mobile device, or TV, and it can be used at home, work, or wherever is convenient to the user. To determine a user's pose and to count the number of repetitions, PoseNet (a real-time pose detection library that uses deep learning) was used. Furthermore, augmented reality was incorporated into the application to see if this can aid users to perform balance exercises. The overall idea of the BaART system is shown in Figure 1.

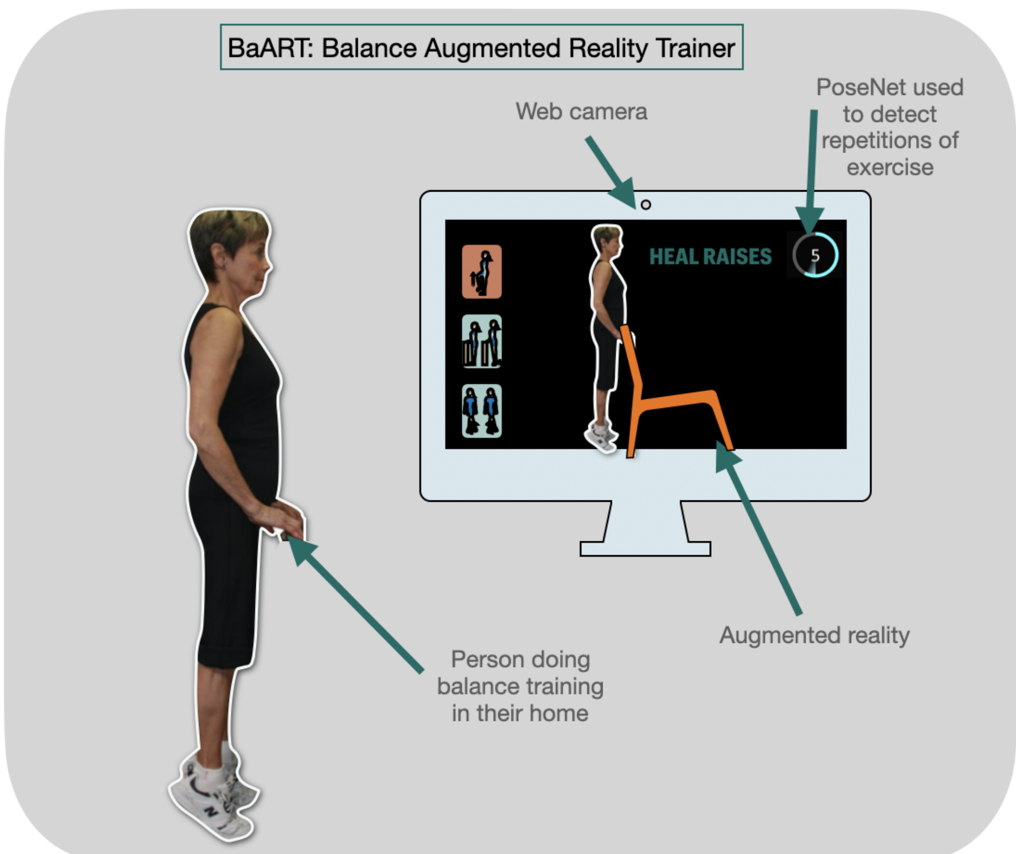


Figure 1: "**Ba**lance **A**ugmented **R**eality **T**rainer" (BaART) detects human body parts using a simple camera. The user gets instructions on the balance exercises, for example, to do heel raises. The keypoint data (from PoseNet) is used to determine if the exercises are executed correctly and count the exercises. The virtual chair is placed in front of the determined location of the user and the user can pretend to hold it for better balance.

1.5 Contributions

In the course of conducting the research described in this thesis, we developed an online trainer "**Ba**lance **A**ugmented **R**eality **T**rainer" (BaART) prototype for doing balancing exercises at home-based.[1] In doing so, we focused on (1) ensuring wide accessibility, (2) ease of use, and (3) augmented reality visualization for improving a user's sense of balance. For the first goal, we used PoseNet which is able to run in a browser and thus requires no additional hardware beyond that of a connected digital device with a web camera. To ensure easy usability, we provide a simple interface with instructional videos and allow for voice commands. To study the impact of augmented reality on balance, we added virtual objects that users could focus on to maintain balance.

A user study was conducted to determine the ease of use of the system and the impact of augmented reality on balance training. The results of our study showed:

1. PoseNet is effective for recognizing exercises to count repetitions and time of doing an exercise.

2. The BaART web application is easy to use (System Usability Score of 80 or very good) and intuitive.

3. Virtual objects such as chairs that are introduced into an augmented reality view of the person doing an exercise do not improve the user's sense of balance and can even distract them from performing the exercises.

1.6 Organization

The remainder of the thesis is organized as follows: in the next chapter, we provide a background to human pose estimation and different methods for doing human pose estimation. In Chapter 3, we describe our developed online training platform and the results of the study that was used to evaluate the prototype. Chapter 4 concludes the thesis and provides avenues for future work.

[1] Github: "https://github.com/Amirhossein-Etaat/BalanceTrainingApplication" Application Link:"https://bit.ly/3znbwM1"

Chapter 2

Background

Computer vision is an interdisciplinary field in which computers are used to get an understanding of the content of digital images or videos. Human action recognition (HAR) is one area of computer vision where actions or an individual's activities (e.g. walking or throwing a ball) are determined using sensor data and/or video or image data [15]. HAR has promising potential and extensive applications in various fields including human-computer interaction, entertainment, surveillance and security, clinical environments, and fitness and rehabilitation [16]. One component of HAR, and the focus of this chapter, is Human Pose Estimation (HPE). Human Pose Estimation (HPE) is the process of detecting the location of an individual's body keypoints (e.g. joints) in space by analyzing images or videos [17, 18]. The localized joints can then be used to recognize the activity a person or multiple persons are performing. In the following chapter, we describe varying methods for motion capture and different approaches to HPE.

2.1 Human Pose Estimation

Human Pose Estimation (HPE) is an effective low-cost and accurate way to capture the movement of people and detect the points of human actions. Thus it has become an important research field of research and is used in many applications including tracking people, action recognition, virtual reality, movies and animation, human-computer interaction, medical assistance, sports, and video games [19]. For example, HPE has been used in the digital entertainment industry to animate digital character models in computer animation. In physical education, it can provide a quantitative analysis of action details so that trainers can make more objective evaluations of exercises. HPE is also being used in medical applications such as education with virtual reality surgical simulators and robotic surgery. More recently, HPE is being used in health, fitness and, clinical applications. For example, researchers have proposed the utilization of motion capture systems and HPE for studying the kinematics of the human body articulations for anatomy in the context of training, physiotherapy and rehabilitation.

Human pose estimation works by making observations to find an articulated human body's pose, consisting of joints and rigid parts [20] as depicted in Figure 2. Pose estimation can be conducted in 2D or 3D and there are many different methods that have been proposed to determine pose, including marker-based methods and image-based methods. The later of which use machine learning methods such as support vector machines (SVMs) and relevance vector machines (RVMs) [21] and deep learning methods. We describe some of these in the following sections.

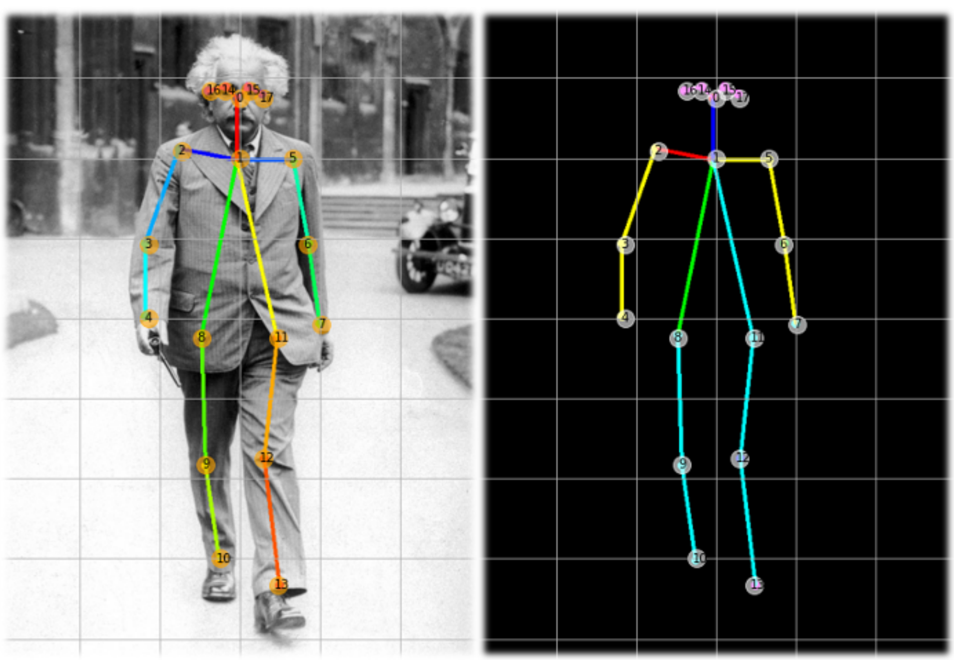


Figure 2: An example of pose estimation of a single-person. The position and location of an individual's keypoints (e.g. anatomical landmarks/joints) have been detected [22].

2.2 Marker-based Human Pose Estimation

Marker-based HPE makes use of motion capture systems that record objects or human movement using cameras. Marker-based motion capture systems comprise a number of hardware components including camera(s) and sensors that collect data for calculations, processing, interpolation, filtering, and storing and transmitting them in real-time to a client application. There are three main types of motion capture technologies: mechanical, magnetic, and optical (see Figure 3), which are described next.

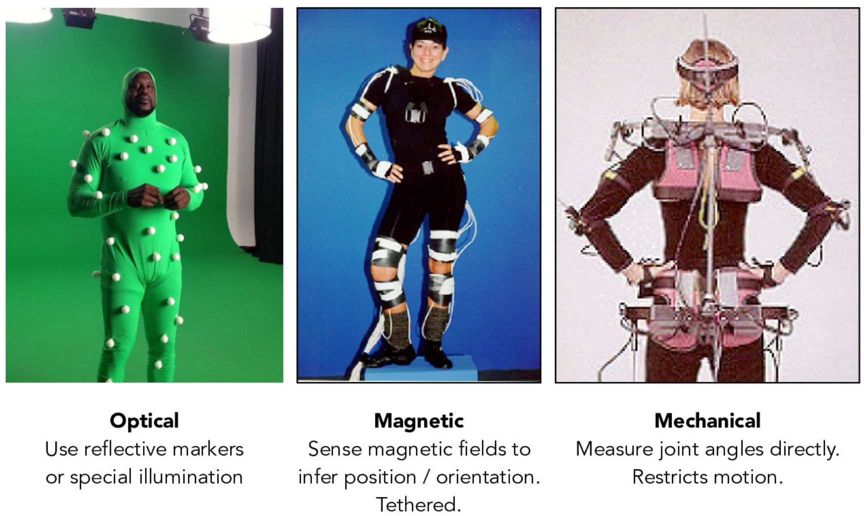


Figure 3: Three main types for motion capture technologies: (a) Optical, (b)Magnetic, (c) Mechanical [23].

2.2.1 Mechanical Motion Capture

In mechanical motion capture, an actor, puts on a bodysuit that is structured like a human skeleton, i.e. an exoskeleton (see Figure 4). The actor's joints are connected to angular encoders that send movement data (the rotation and position of the joints) to a workstation for processing. The movement simulation procedure includes a mechanism by which the computer, in charge of processing, compares the relative position of the encoders with the new data coming from them as the actor is moving. The computer records the value of movement of each encoder by knowing the relative position of the encoders and/or joints. Using this information pose and movements can be reconstructed in the software environment. It should be noted that an offset is applied to each encoder because it is very difficult to precisely match their position with that of the real relationship [24, 25].

Figure 4: An exoskeleton as depicted above is used in mechanical motion capture systems [26].

Mechanical motion capture has two main advantages: (1) the data capturing process is highly accurate and (2) the data capture is not susceptible to external sources, e.g. low quality of the cameras and the number of cameras used for capturing. However, as the encoders are mechanically constrained not all details of every movement can be encoded. Furthermore, the exoskeleton can be cumbersome, particularly the conventional exoskeletons which use wires to connect to the encoders, which in turn use wires to send the data to the computer [27].

2.2.2 Magnetic Motion Capture

Magnetic motion capture technology uses a network of electric sensors and transmitters that have three orthogonal coils. By measuring relative disturbances in an induced magnetic field the orientation and rotation of the joints can be calculated [28, 24]. The sensors are attached to a bodysuit as can be seen in Figure 5.

Magnetic motion capture provides excellent data accuracy and is less computationally expensive than mechanical systems. The biggest disadvantage of these types

Figure 5: Magnetic motion capture was used in the Fox filming of ”The Lord of The Rings” [29].

of systems is the susceptibility to magnetic field disturbances caused by the presence of metal objects, which can heavily distort the data [27].

2.2.3 Optical Motion Capture

Optical motion capture systems use a set of cameras to capture a person wearing a bodysuit, i.e., a “MoCap Suit”, onto which optical markers are attached. Thus, one can imagine that optical motion capture works similar to radar. The cameras emit waves, typically infrared electromagnetic waves, toward the actor who is wearing a bodysuit with attached markers. These markers (i.e. infrared reflective spheres) reflect the infrared to the cameras that in this stage, function as receivers. The cameras capture the markers from different angles to determine the coordinates of the markers. The data collected from all cameras (at least two) are sent to a processing unit to determine the position and movements of the actor [27]. Specifically, by using two or more cameras it is possible to triangulate the 3D position of each marker. It should be noted that optical motion capture systems should be calibrated in order to deal with optical parallax and distortion and ensure high data precision.

The collection of data in this technique is done conventionally through the markers attached to the outfit of actors, however, modern capturing systems can also collect more precise data from each particular subject via dynamic surface tracking methods.

Optical motion capture systems can be extremely accurate and have the benefit

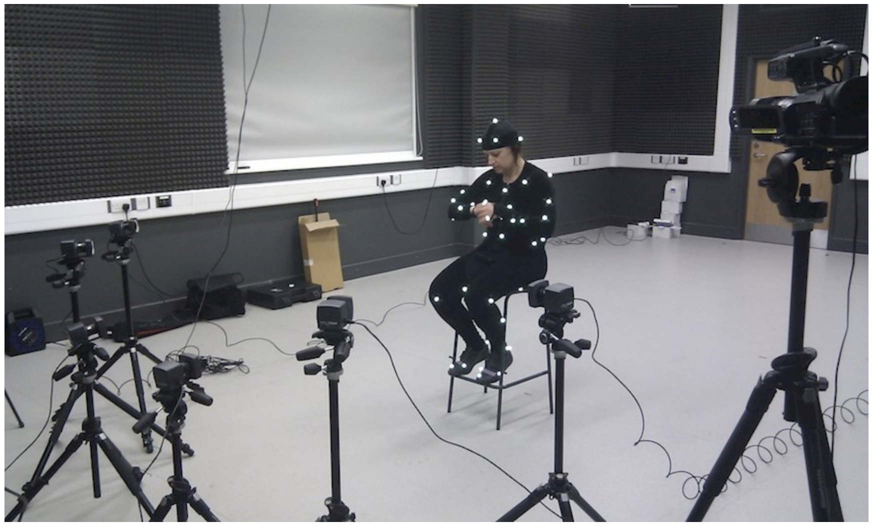


Figure 6: A dancer wearing a suit used in an optical motion capture system [30].

of being very configurable. A large number of markers and a large active area is possible, depending on budget and space limitations. However, optical systems need extensive post-processing, so operating costs are high, the hardware is expensive and the systems are sensitive to reflective noise and yellow light[25].

Kinect

One popular, consumer-grade, optical motion capture system is the Microsoft Kinect. The Kinect is a motion-sensing input device that was first released in 2010 by Microsoft [31] (see Figure 7). The Kinect is an RGB-D sensor providing synchronized colour and depth images that includes RGB cameras, depth sensors, microphone arrays, and infrared projectors that can determine depth using structured light which is able to provide depth signals, audio signals, and RGB images simultaneously [32]. This allows the Kinect to do real-time motion detection for such things as gesture-based interaction and human pose estimation for activity recognition.

There are a few articles that assess the Kinect's performance from both a hardware and software perspective. According to the experimental results, the Kinect is more accurate than a time-of-flight camera (ToF camera) and comparable to a medium-resolution stereo camera. The Kinect works very well in low-light rooms, however does not work well outdoors. One can obtain the precise depth image without difficulty and use it to precisely track human poses [34]. The Kinect camera is also really

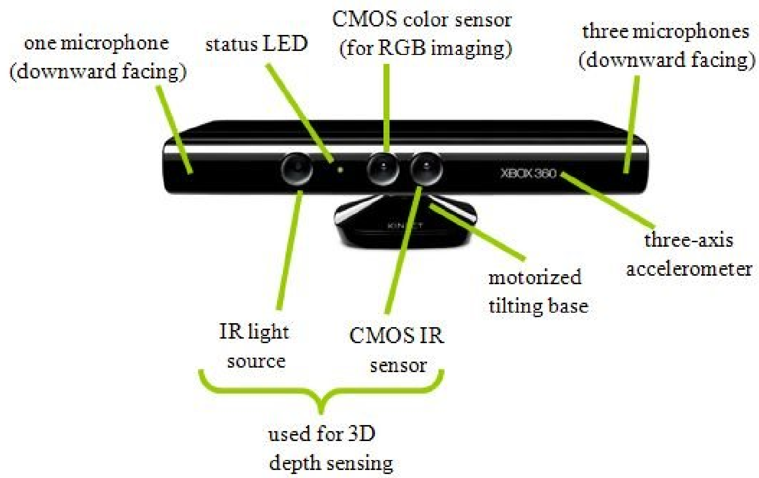


Figure 7: A Kinect sensor as an input device which captures the color image and the depth map [33]

beneficial for various HCI applications, but it should be noted that the distance for detecting objects is restricted to under 3.5 meters [34].

2.2.4 Summary of Motion Capture Systems

In general, marker-based motion capture systems require specific hardware and specialized software programs to receive and process motion data. The costs of the software, equipment, needed space and the personnel required can be significant, and thus such systems are not always ideal. For this reason, several innovative research studies in computer vision have introduced creative techniques that eliminate the dependency of motion capture on specialized markers, suits, and cameras and rather use simple images or videos for HPE. These are described in the following section.

2.3 Computer Vision Based Pose Estimation

Computer vision based techniques for HPE employ special computer algorithms to analyze streams of video or image inputs. These algorithms determine the position of specific joints (e.g. knee) or anatomy (e.g. nose) to identify the position of the body and make it possible to determine actions or track movements. Thus, computer vision-based methods predict the pose without the presence of an actor wearing any type of special hardware, equipment, or suit but rather by determining the location of specific keypoints of a single (or multiple) persons in a video and/or image. One of the main challenges of computer vision based HPE methods includes estimating the keypoints of a pose when there is: no visible joints (e.g. due to clothing), background noise (e.g. other people or objects), self-occlusions, and/or there is no consistent lighting. HPE can refer to single or multi-person pose estimation and can be 2D or 3D but in general has a similar pipeline regardless of the method as described below.

2.3.1 Single and Multi-person Pose Estimation

Pose estimation may be classified based on the number of people tracked, i.e., one person or *single-person pose estimation* (SPPE) or more than one, *multi-person pose estimation* (MPPE). Single-person pose estimation is an easier problem as only one person is present in the image or video frame. Multi-person pose estimation is more complex for a number of reasons, one of which is the issue of inter-person occlusion, where figures in the scene may overlap.

2.3.2 2D versus 3D Human Pose Estimation

The pose estimation problem may also be classified as either 2D or 3D depending on the required output dimension. In 2D HPE, the task is to predict the location of body joints in an image in terms of the pixel values of those locations. In 3D, the task is to estimate the articulated spacial 3D joint locations or in other words the spatial arrangement of the joints in three dimensions. Thus 3D HPE is considered to be relatively more difficult in comparison with 2D HPE. Although many 3D pose estimation models first predict 2D Pose and then extend this information to 3D pose some methods for directly predicting 3D pose do exist [19].

Traditional 2D HPE algorithms used hand-craft feature extraction and body models, e.g. rigid kinematic skeleton models with a specific number of joints to determine poses [19]. For images or frames with multiple persons, pre-processing is required to crop the source image so that there is only one single person and processing is done on each cropped image [19]. Popular 2D human pose estimation libraries include OpenPose, CPN, AlphaPose, and HRNe.

With 3D HPE, not only can one predict the 3D joints in 3D space, but also there are some methods that are able to determine the 3D human mesh from images or video frames. Owing to this, there has been much interest in 3D HPE as it could be used in animation, virtual and augmented reality and 3D action prediction. 3D multi-person HPE is more complex and few methods have been proposed to solve the problem with those have using deep learning approaches. One toolkit that uses neural networks for real-time 3D human pose estimation (SPPE and MPPE) is OpenPose, described later in this Chapter.

2.3.3 Human Body Modelling

Three-dimensional HPE can use model-free or model-based methods. The model-free methods do not use human models, whereas model-based methods use a parametric body model or pattern to estimate pose and human joints from the images [19]. In other words, the locations of human body parts or keypoints in an image or video are used to build a human body representation. There are three types of models for human body modeling: kinetic or skeleton-based, planar or volumetric. Kinetic models can be represented as a graph with vertices representing joints and edges encoding constraints and priors about the body model structure. These can be used in both 2D and 3D pose estimation. Planar or contour-based models, which can be used for 3D pose estimation, represent the shape of a human body such that body parts are represented as shapes such as circles, rectangles, cylinders, etc., that approximate the human body shape. Lastly, volumetric models have been used for 3D pose estimation particularly in contexts where 3D meshes are needed, for example, character animation. Numerous 3D human body models exist and are used for deep learning based 3D human pose estimation to recover a 3D human mesh. GHUM and GHUML(ite) for example are deep learning pipelines trained on a dataset of full-body scans of over 60k human configurations.

2.3.4 Human Pose Estimation Pipeline

In the following section we give the main steps involved in HPE which involves pre-processing, extracting features, and inferring the features.

Pre-processing: A number of pre-processing steps may be needed prior to detecting the pose. These can include removing the background image from the figure,

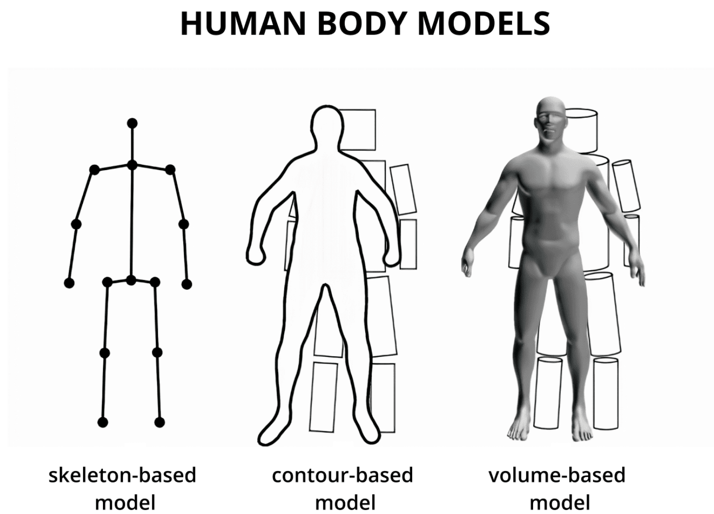


Figure 8: Human body models that are commonly applied. a skeleton-based model, b contour-based models, and c volume-based models [35].

denoising, and for MPPE determining each human in the image or the frame. In some algorithms a bounding box may be used around each human in order to process them individually. Camera calibration and image registration may also be needed if multiple cameras were used, e.g. in the case of 3D Human Pose Estimation.

Figure 9: One pre-processing step may include creating a bounding box for every human in a given image or frame. Each bounding box may then be individually evaluated for Human Pose [36].

Feature Extraction In the context of HPE and machine learning, feature extraction involves creating derived values from raw data, in this case, images or videos that can then be used in a learning algorithm that detects poses. Features can be explicit, e.g. Scale Invariant Feature Transform (SIFT), or implicit, e.g. features generated automatically by deep learning feature maps.

Pose Inference There are two main approaches to inferring pose: top-down and bottom-up. Bottom-up methods first estimate each body joint and then connect the joints to construct a human body in a particular pose. When the algorithm detects the different parts and joints of the human body a formulation of the graph is used to determine the human body and pose. In the top-down method, a person detector is first used and then the body joints within each detected bounding box are determined. Two popular methods for top-down pose estimation are generative body model based methods, which fit a body-model on the image, and deep learning based methods.

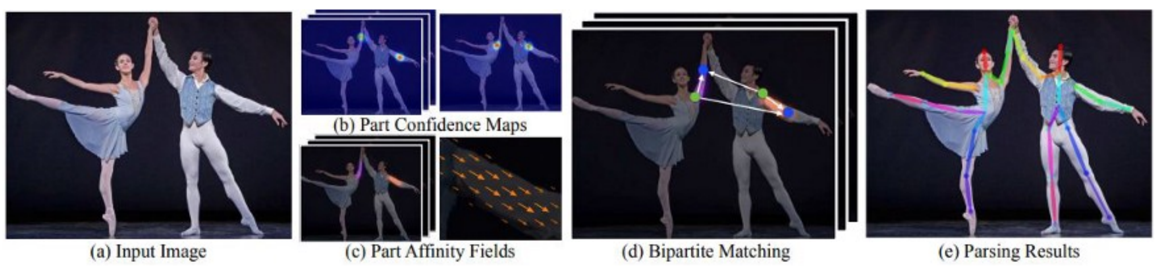


Figure 10: An example of the bottom up approach to HPE [37]

Post-processing Lastly, post-processing may be used to determine any abnormal human poses and correct or reject them. A number of post-processing algorithms exist which evaluate the result of the pose estimation pipeline and score the results based on the likelihood that the result is a valid pose.

2.4 Pose Classification using Deep Learning

In the last couple of years, artificial intelligence has given encouraging results in the field of HAR systems. Neural networks and numerous deep learning algorithms have been proposed to develop models in order to automatically track human body skeletons. Human pose estimation has different deep learning categories: (1) human body model-based and model-free which can use generative and discriminate characteristics; (2) from high-level abstraction to low-level pixel evidence which can be top-down and bottom-up, and (3) regression-based (i.e. directly mapping from input images to body joint positions) or detection-based (generating intermediate image patches or heatmaps of joint locations); and (4) end-to-end training and stage-by-stage training [19]. These are briefly described below.

Regression-based and detection-based HPE methods can be either top-down or bottom-up. Top-down approaches begin by detecting humans and generating their locations in bounding boxes at a high-level of abstraction after which each pose is estimated. Bottom-up approaches, on the other hand, estimate all body components of each person in the input image before grouping them using human body model fitting [19].

Regression-based methods seek to learn a mapping from input image to pose using an end-to-end framework that directly generates joint coordinates. Detection-based methods predict positions of body parts and joints, using a sequence of rectangular windows, each for a specific body part. Heatmaps can also be used each showing one joint position by a 2D Gaussian distribution at the joint location [19].

Regional multi-person pose estimation (RMPE) is one of the most favored available top-down methods of pose estimation. Top-down methods notoriously rely on the accuracy of the person detector used, which can be a complex problem when the image area is full of people. In an area with many people, wrong localization readings and redundant bounding box predictions are inevitable and negatively impact the detection performance of algorithms [38].

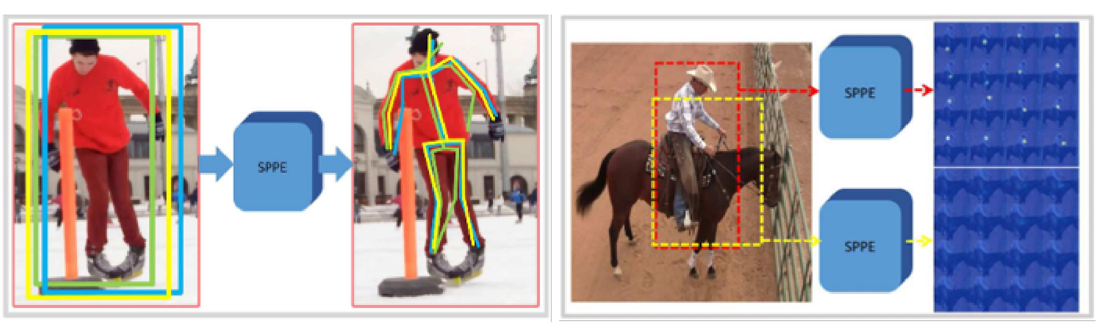


Figure 11: Impact of duplicate predictions and low confidence bounding rectangle windows [38].

To overcome the challenge of obtaining a high-quality single-person region from an incorrectly given bounding box, Fang *et al.* proposed the symmetric spatial transformer network (SSTN) [38]. In this method, a single person pose estimator is put in charge of designating the region in question to locate the pose for a given person. For converting the estimated pose to the original coordinate system of the image, a spatial de-transformer network (SDTN) is employed, and in the end, a parametric pose non-maximum suppression (NMS) technique facilitates the elimination of pose deductions. A 'Pose Guided Proposals Generator' can also be used to improve training samples to enhance the training of the SPPE and SSTN networks.

OpenPose

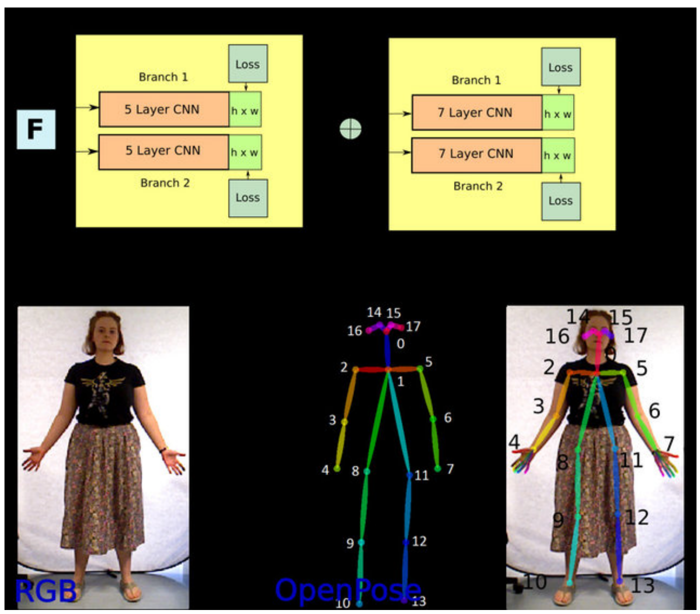


Figure 12: The OpenPose architecture has a multistage CNN with two branches. Confidence maps are predicted by the first brannch and Part Affinity Fields (PAFs) by the second branch. The predictions from the two branches, as well as the picture features, are concatenated [39].

OpenPose has become one of the most widely used bottom-up approaches for MPPE. In the first step, OpenPose detects parts or keypoints, then associates them with individual subjects in the image. The OpenPose network analyzes the first few layers of an image to determine the properties that are then supplied into two parallel branches of convolutional layers (see Figure 12). The first branch detects confidence maps, and the second branch detects Part Affinity Fields (PAFs), after the first stage the predictions from two branches exchange to the next stage which repeats these stages, and finally, produces a collection of 2D vector fields that encode the location and orientation of keypoints [40] as output. A total of 18 maps are predicted for each keypoint (e.g. eyes, ears, nose) in the human pose skeleton to find their location. Note that a confidence map is a two-dimensional depiction of the confidence points that a specific body component is at a given pixel. The confidence maps can be used to create bipartite graphs between pairs (the connection between parts) of parts (keypoints).

PoseNet

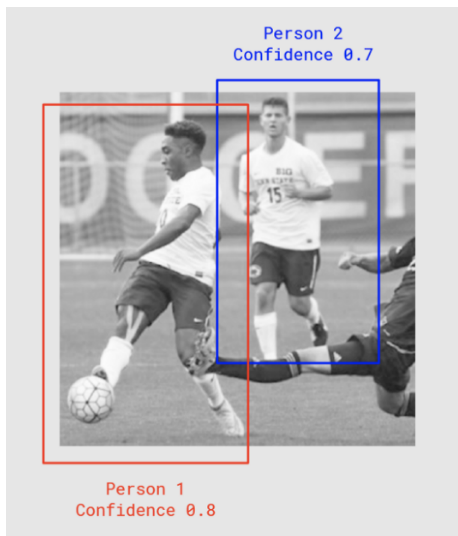


Figure 13: PoseNet gives confidence score for each pose and keypoint detected [41].

PoseNet is a deep learning TensorFlow model that can estimate and track human poses by detecting 17 keypoints (e.g. knees, ankles, wrists, elbows, hips, nose, eye, etc). Note that TensorFlow is an open-source machine learning platform that provides an ecosystem of tools for developing ML applications. PoseNet can be used for single or multi-person detection and determines poses, pose confidence scores, keypoint positions, and keypoint confidence scores from the model outputs [42]. More specifically, PoseNet provides a posture object containing a list of keypoints and an instance-level confidence score for each detected individual [43]. The confidence score is based on how well it can recognize the image or a certain key point within that image [42]. For example, in Figure 13 one person in the image has a confidence score of 0.8, while the second person has a confidence score of 0.7 due to occlusions and the distance of the subject. The keypoint position and keypoint confidence scores are also given. All of the keypoint positions in the input picture include the x and y coordinates and can be mapped directly onto the image.

PoseNet was the first learning-based architecture that used the concept of regressing the absolute pose with a deep architecture. In doing so, it has several advantages when compared to previous methods. These include the fact that the architecture has short inference times (milliseconds instead of minutes), a low memory footprint (megabytes compared to gigabytes), and it doesn't need conventional sensor tools or physical devices. The TensorFlow.js version of PoseNet also allows for real-time human pose estimation in the browser and thus was chosen for the developed prototype.

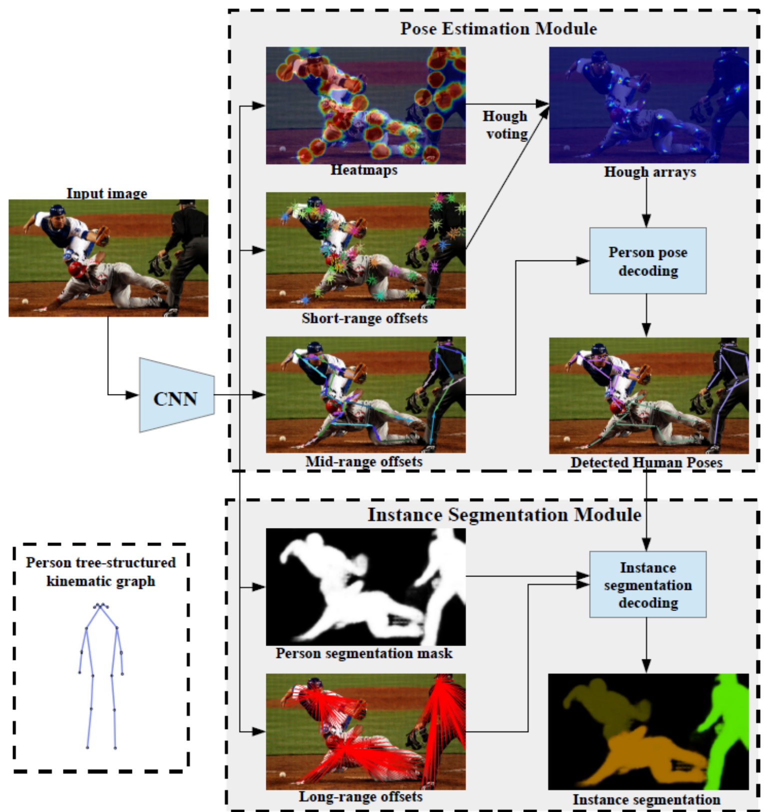


Figure 14: PoseNet: It is based on a fully convolutional way, and the model employs a convolutional network to detect individual keypoints, and predict displacements. The Pose Estimation Module employs the keypoint heatmaps and short- and mid-range offsets to detect human poses. The Instance Segmentation Module uses the person segmentation maps and long-range offsets with the human pose detection to determine the person instance segmentation masks [43].

In the next chapter, we describe how we used this version to develop a web application to explore the use of augmented reality in the context of remote balance training exercises.

Chapter 3

An Online Application using Pose Estimation and Augmented Reality for Balance Training

A version of this chapter was submitted to the International Conference for Information and Communication Technologies for Aging Well and E-health (ICT4AWE 2022).

3.1 Introduction

As smart mobile phones and internet connections become more accessible than ever, mobile health (mHealth) applications are becoming ever more popular with many seeking to take their health into their own hands within their own environments. This has led to growing availability of web and mobile applications for rehabilitation, virtual coaches for training and exercise, online doctor visits, and mental health, meditation, and fitness applications.

In this paper, we focus specifically on rehabilitation and balance exercises. Balance exercises which work the core muscles, lower back, and legs are often prescribed for healthy aging to keep the body active, improve balance and coordination, and prevent falls and injuries. They are also used in rehabilitation after injuries or diseases such as stroke. New sensing (e.g. wearables) and communication technologies (e.g. internet of things) are positively influencing the expansion of training and rehabilitation programs outside of standard healthcare facilities. It is now feasible to do guided online exercises ensuring that they are done correctly and even track a person's progress over time in the comfort of one's own environment without direct contact with a physiotherapist or trainer. This can be done using different motion-tracking devices such as wearable sensors, depth cameras and most recently using *pose estimation* from simple web camera feeds. Pose estimation is the task of using machine learning models to estimate the pose of a person from an image or a video by determining the spatial location of keypoints (i.e. knee, elbow, nose, etc). This set of keypoints can then be connected to describe the pose of the person. Pose estimation is beginning to be widely used in training robots, motion tracking for game consoles, augmented reality, animation, and health and fitness applications.

Gamification, i.e. the introduction of game elements to traditionally non-gaming situations, has also proved to be an important aspect of mobile health applications, which can encourage and motivate users to not only comply with their exercise regimes but also to help them to enjoy doing so [44]. Mixed reality (combining virtual and real elements) has also been used both for aiding users to better understand exercises and the proper way of doing them and/or for adding gaming elements to make exercising more enjoyable.

In this paper, we focus specifically on delivering balance training exercises into the comfort of a user's own home using simple and low-cost technology (only a webcam and connected digital device are needed). Furthermore, we explore the use of mixed reality elements, i.e. adding a virtual chair for balance, to see if this can aid users to perform balance exercises. In order to study these questions, we developed a balance training exercise prototype system, "**Ba**lance **A**ugmented **R**eality **T**rainer" (BaART) and tested it with 20 subjects to determine the impact of the AR elements and in general the ease of use of the system.

3.2 Related Work

In the following section, we describe related works in terms of rehabilitation, balance training mHealth applications, and pose estimation.

3.2.1 Pose Estimation Health Applications

Encouraging people to exercise by exploiting technology at home can be a strategy that may be cost-effective, particularly if consumer-level equipment is used. A number of research groups have looked at using simple applications with webcams and using pose estimation algorithms to determine the location of the user's different joints in order to help guide and monitor exercises. Pose estimation specifically refers to computer vision techniques that can be used to detect human figures in images and videos. By taking a processed camera image as the input, pose estimation models outputs information about the keypoints (e.g. elbow, nose, knee, etc.) which can then further be used to determine the pose of a person.

Moreira *et al.* [45] systematically reviewed mobile applications proposed for analyzing human posture. Their results showed that the use of human pose estimation on mobile applications is reliable, and can assist medical clinics, and physiotherapists, especially in the case of evaluating physical treatments. Moreira *et al.* [46] also proposed a prototype application using PoseNet (a real-time pose detection library) and the TensorFlow library for automatic identification of Anatomical and Segment Points (ASPs). Their results show that PoseNet can be used to develop applications concerning the physical assessment process and diagnosis of disorders related to postural and movement changes. Herrera *et al.* [47] developed a web application, using Posenet and TensorFlow, for sedentary workers to help them take *active pauses* (i.e. pausing work to do exercises) in order to reduce the risk of job-related muscular and skeletal injuries and diseases. The application accurately determined the position of the user and whether they were doing the exercise correctly.

Yan *et al.* [48] presented a method for continuous human rehabilitation action recognition using OpenPose and a Fully Convolutional Neural Network (FCN). Their method used a Kalman filter to more accurately detect poses from RGB video streams. Their experimental results have shown that the method is able to deal with background noise and various body types with an accuracy in the recognition rate of the pose of 85.6%. Deb *et al.* [49] proposed a platform-man interaction paradigm for the learning of dance routines. OpenPose (a pose detection library) was used to capture the human body and the key points of the skeleton in real-time.

3.2.2 Balance Training

A number of previous research studies have employed various technologies to help aging adults to train balance, with an aim to maintain health and reduce the likely hood of falls [50]. Mostajeran *et al.* [51] presented an application for balance training at home using the Microsoft Kinect Sensor V2. The application has a virtual coach who gives balance training instructions and demonstrated the exercises. The authors found that adults have a more positive reaction towards using a virtual coach for balance training compared to traditional health care approaches. In a similar work, Kouris *et al.* [52] proposed a system that uses a virtual reality avatar and wearable sensors for physiotherapy balance exercises. The wearable sensors can monitor user activity and determine the correctness of the balance exercises in real-time. Vonstad *et al.* [53] developed a 2D custom balance training exergame (i.e. a type of exercise that mixes exercise and video games) using a deep learning based pose estimation system to detect human body parts and estimate three-dimensional (3D) body positions. In their work, they compared three systems: a Microsoft Kinect Sensor V2, a marker-based three-dimensional Motion Capturing system and a deep learning system using a digital camera. In a study where participants played a balance training game, the deep learning method had similar performance to both the Kinect and the marker-based system. This work demonstrates the feasibility of using less complex hardware and sensors for these types of applications.

Ogonowski *et al.* [54] developed a system that used individualized physical fitness training, gamification and wearable sensors (senior mobility monitor device) with a Microsoft Kinect and simple TV. Their work illustrated the possibility of incorporating such systems into the daily life of older adults. Hardy *et al.* [55] also developed an exergame to motivate older adults to do balance exercises and gain training based on the adaptation and exergame analysis. Their findings suggest that elderly individuals accept multimedia training and the adaptability notion improves the system's accessibility.

Smartphones are also beginning to be tested for use for balance, rehabilitation and physiotherapy exercises. For example, Androutsou *et al.* [56] proposed a smartphone application for patients with balance disorders that enables users to self-evaluate their activity and progress, communicate with others using the system, and get real-time feedback about their training, activities and progress over time.

3.2.3 Mixed Reality Health Applications

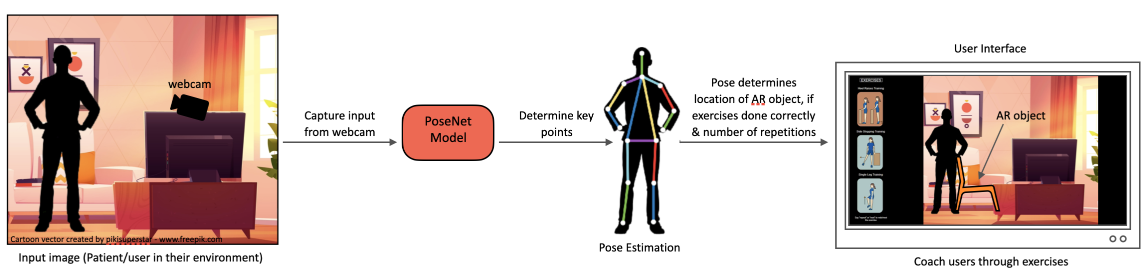
A number of researchers have explored the use of mixed reality in the context of fitness and rehabilitation applications and specifically the use of virtual coaches or trainers [57]. Virtual coaches have been explored to help clinicians monitor the use of medications and seniors' adherence to specific guidelines for medication use [58] or to keep the elderly physically active [59, 60]. For example, Felberbaum *et al.* [61] explored aspects of an AR-based virtual coach that would improve older adults' mobility. Features of such a virtual coach would include virtual friends to walk with, having interactive guidance to help define and reach goals and online monitoring to determine risks such as falls. Bickmore *et al.* [59] demonstrated that the combination of a virtual coach and pedometer can increase the amount of walking among older adults.

Virtual training assistants or coaches do not necessarily have to take on the form of a human avatar. Albaina *et al.* [60] proposed an animated flower that is used as a virtual trainer to boost the motivation of individuals to walk. Albaina *et al.* used two focus groups and conducted one field study and found that elderly users appreciated the flower virtual trainer and had a positive tendency to use it over a longer period of time. The results of their studies also demonstrated that the virtual flower trainer was effective in boosting the acceptance of such a system among older adults, despite its uncommon form.

Other studies have incorporated AR technology into the daily lives of older adults [62]. For instance, Bianco *et al.* [63] developed a tablet-based AR application for fall preventive home modifications, e.g. the installation of more handrails at desired locations at home. Ku *et al.* [64] used the 3D-ARS (3D interactive augmented reality) system to determine its impact on people with balance disorders. A randomized controlled study with 36 people who could walk on their own and could stand on one leg was done to assess the effect of the AR technology. In comparing the control group, who did a conventional physical fitness program three times a week for 1 month, to the experimental group, who used 3D-ARS training three times a week for 4 weeks, they found that although both groups experienced improvements those who used of 3D-ARS had better results in terms of stability index, weight distribution index, fall risk index, and Fourier transformations index of posturography.

3.3 System Description

Using intervention and assessment tools *at home* can considerably decrease the risk of falls and enhance one's quality of life through increased mobility and in turn healthy aging [65]. Balance training exercises and therapies are commonly used to improve both balance and gait conditions. According to the American Geriatrics Society [66] a proper and standard balance exercise plan can play a vital role in health and rehabilitation. However, studies have found that many patients and users do not do physiotherapy exercises or training that can help or prevent ailments due to a lack of motivation, time and/or the high cost of physiotherapy, gyms, and hiring personal trainers [10]. Building on the previous works described above, our goal was to develop a prototype system that requires no specialized equipment but rather can be used in the comfort of one's desired environment by simply clicking on a url on a computer, mobile phone, tablet, Smart TV or other connected device. We also aimed to make the system easy to use by providing a simple interface that includes voice commands and instructional videos. Our developed home-based rehabilitation online app uses (1) pose estimation to determine the users pose and count the number or repetitions of a given exercise and (2) explores the ability of augmented reality to help with balance exercises.



S

Figure 15: The user opens the application on a browser on any connected digital device. The web browser fetches data from the server (written in HTML, CSS, JavaScript codes including Posenet Library and P5js library). The keypoints are detected using the PoseNet model. The web application processes the captured inputs to determine exercises correctness and counts repetitions, as well as, displays the AR object in the correct position relative to the user.

3.3.1 PoseNet

Our prototype uses PoseNet [43], a library built on top of TensorFlow, which is an open-source machine learning platform that provides an ecosystem of tools for developing machine learning applications. PoseNet, which is trained in the MobileNet Architecture (a convolutional neural network (CNN) developed by Google and trained on ImageNet) has pre-trained models that detect user gestures and poses from given images. An advantage of PoseNet in comparison to other API-dependent libraries for pose estimation is that the pre-trained models run in a browser thus private data (i.e. images and video stream from webcam) are protected and not sent to a server for processing. Furthermore, anyone with a webcam equipped desktop or TV, tablet or phone can use PoseNet within their web browser. Taking advantage of PoseNet allowed us to keep the system simple and highly accessible, without the need of devices such as Kinect, wearables or motion capture systems that can add complexity and may have high installation and preparation costs.

3.3.2 BaART: Balance AR Trainer Online Application

To develop the BaART application, we used Javascript, HTML and CSS to implement the user interface and PoseNet and TensorFlow's models to recognize and estimate the main human poses using the web camera stream. Pose estimation was specifically used to count the number of repetitions that were done by the end user and check if the exercise has been done correctly. We also used p5.Speech (an extension of p5.js which is an open-source JavaScript library) for implementing speech-to-text and voice commands.

The application pipeline is shown in Figure 15. As mentioned, the application uses a simple webcam without any further specialized equipment. After opening the application via the provided URL on a computer, laptop, or mobile device, RGBA input is captured from the web camera and the user's pose is estimated via PoseNet. After successful keypoints detection, the exercise routine begins. The keypoint data (e.g. knee, ankles, hips, shoulders, eyes, etc.) is used to determine the location of the user to both place the virtual object and determine if an exercise is being done. While exercises are being performed, the keypoints and movements are tracked so that the system can evaluate if an exercise is being executed correctly and the number of repetitions the user has done. Results are displayed to the user in real-time through the interface, which shows the count and or repetitions (either number of seconds or repetitions successfully completed). The user can interact with the application using either voice commands: “next exercises”, “re-do”, “refresh” or “previous” or with a mouse on a computer or gestures on a mobile device.

Exercises

For the first prototype of our system we selected three exercises that improve balance and prevent falls based on the clinical guidelines and recommendations of the National Institute of Clinical Excellence (NICE) and the joint American and British Geriatric Society (ABGS) for the prevention and assessment of falls in older people [7, 6]. Specifically, we chose: *heel raises*, *side stepping* and *single leg standing* exercises.

Tyagi *et al.* [67] studied the adoption of technology-based telerehabilitation (TR) and found that although younger users preferred TR, older ones preferred daily visits to the rehabilitation center. As in person visits are not always feasible particularly for daily visits (or during a global pandemic), Tyagi *et al.* recommended using videos in TR programs for older patients. Thus for each of the exercises in the BaART application, an instructional video (taken from www.Physitrack.com) is first shown to explain how to do the exercise. After the video, the user then does the given exercise following along with both verbal and onscreen written instructions. As well as giving instructions, the voice of the “trainer” tries to motivate the user with encouraging statements and counts the repetitions the user has done. We briefly describe how each of the exercises was implemented in the application next.

In the first step, based on the user’s image resolution, the placement of virtual system objects and the user interface is determined using a pivot based on a variable coefficient, which changes as the page shrinks and enlarges. In the next step, we receive all the key points through Posent real-time, and we check if they have a confidence score above $C > 0.5$ ($C=$ Confidence score is based on 1-0) so we collect them. Then, we want to make sure the user is standing in the right position and their knees are below a quarter of the image resolution, so we display a line and a box at the beginning. After the users stand upright, the users will be shown how to do this exercise.

Heel raises: For heel raises, the keypoint information is used to track the position of the hips (see the green line in Figure 16 bottom) of the user while standing and uses the raising and lowering of the hips to determine the number of repetitions. Specifically, we consider the hip keypoint such that if the user's hips are at position $hips_1$, and the user raises their heels, leading to the hips being at position $hips_2$ (the second position of hip in y coordinates) then $hips_2$ should be larger than $hips_1 + t$, where t=threshold(5% of the amount of hip in the Y coordinates) such that $hips_2 > hips_1 + t$. We set a timer to make sure that the user stays on their toes for at least 2 seconds. Using this we can determine whether the user does the exercise properly and can count the number of repetitions.

Often times a chair is used for extra balance in heel raise exercises, thus we use a virtual chair to study if "resting" one's hands on a virtual chair can invoke a sense of balance (see Figure 16). The virtual chair is placed in front of the determined location of the user. During the exercise, the user raises both of their heels off the ground to stand on their tiptoes. The user raises their heels 10 times, each time the application says "raise your heels", and then slowly lowers their heels back to the floor when the application says "release".

Side Step: For side stepping, the keypoint information of the hips and shoulders are used to ensure the user is stepping to the side correctly and to count the repetitions. Specifically, from the initial position is A and $hips_1$, and when the user steps to the right right, we have position B as $hips_2$ (in x coordinates) then m = amount moved in x. Similarly, we do the same for a step to the left. Further mores we can use the two shoulder positions to detect if the user rotates, or does not stand properly, and thus determine that the user does the exercise correctly and count the number of correct repetitions.

Similar to the first exercise, a virtual chair is shown in front of the user so that the user can "rest" their hands for better balance. This exercise starts with the feet together and knees slightly bent, the user then moves one foot to the side and then back to join the other foot. The instructions given are "step right" to take a full right step and then "release" to step back to the centre and bring the feet back together. The exercise is repeated 10 times.

Single leg standing: For this exercise, the most challenging of the three, knee keypoint information is used to determine how high the leg is raised relative to the

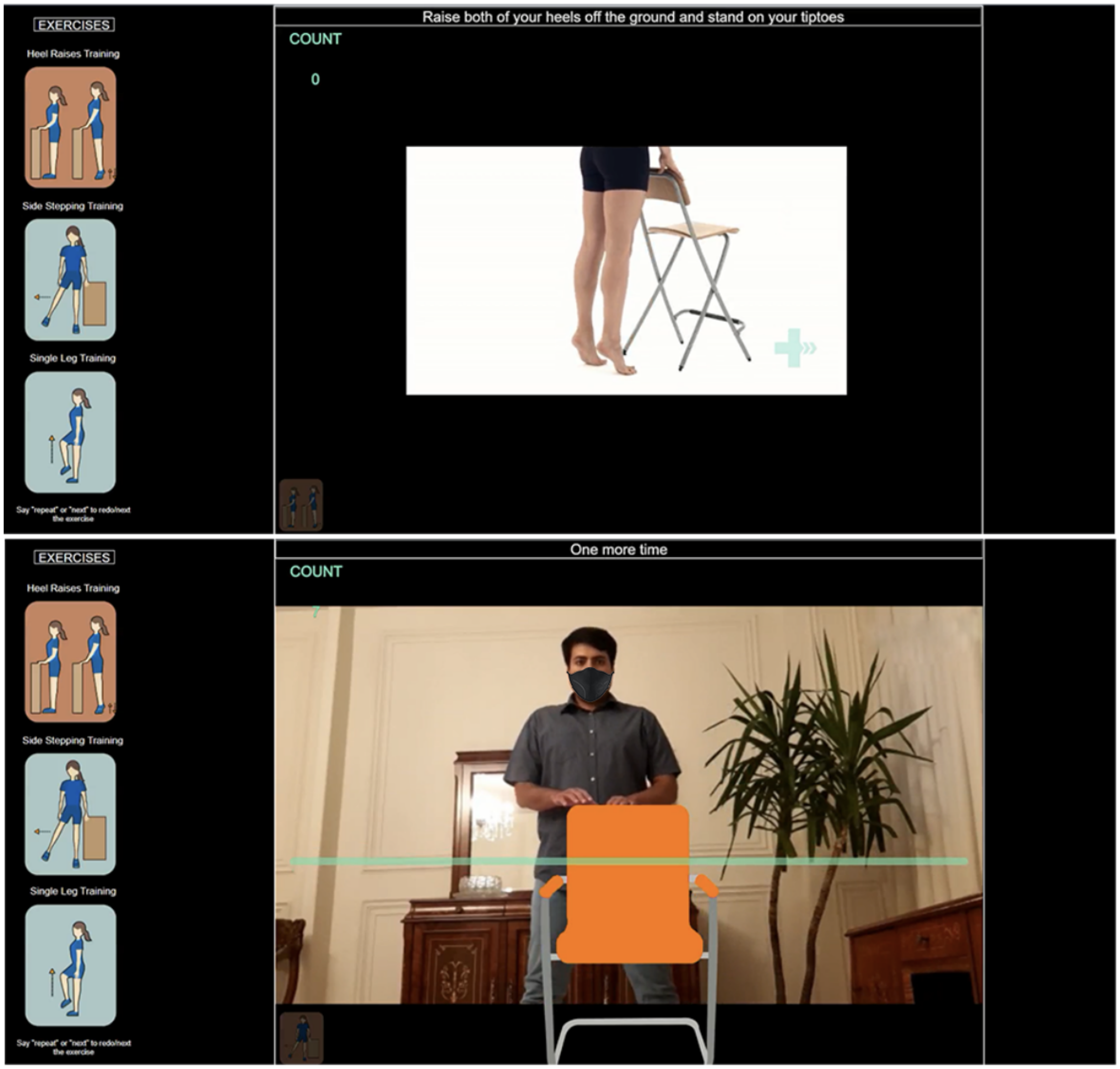


Figure 16: **Top:** The instructional video describing how to do heel raises. **Bottom:** A user doing the heel raises holding on to a virtual chair. The green line represents the position of the hips. Icons on the left are adapted from: www.saebo.com

.

distance to the pelvis. Specifically, we display a red line between the user's knees and hip, and if the user lifts his foot off the ground, the foot should be above the red line, i.e. k (knee) should be above the threshold line t in y. When this is the case the system will start counting for 15 seconds, and if the user loses the proper position the system pauses the counting.

For this exercise, the user is asked to rotate the body by 90 degrees to the right or left in order to be able to best detect the knees. The user is then instructed to stand

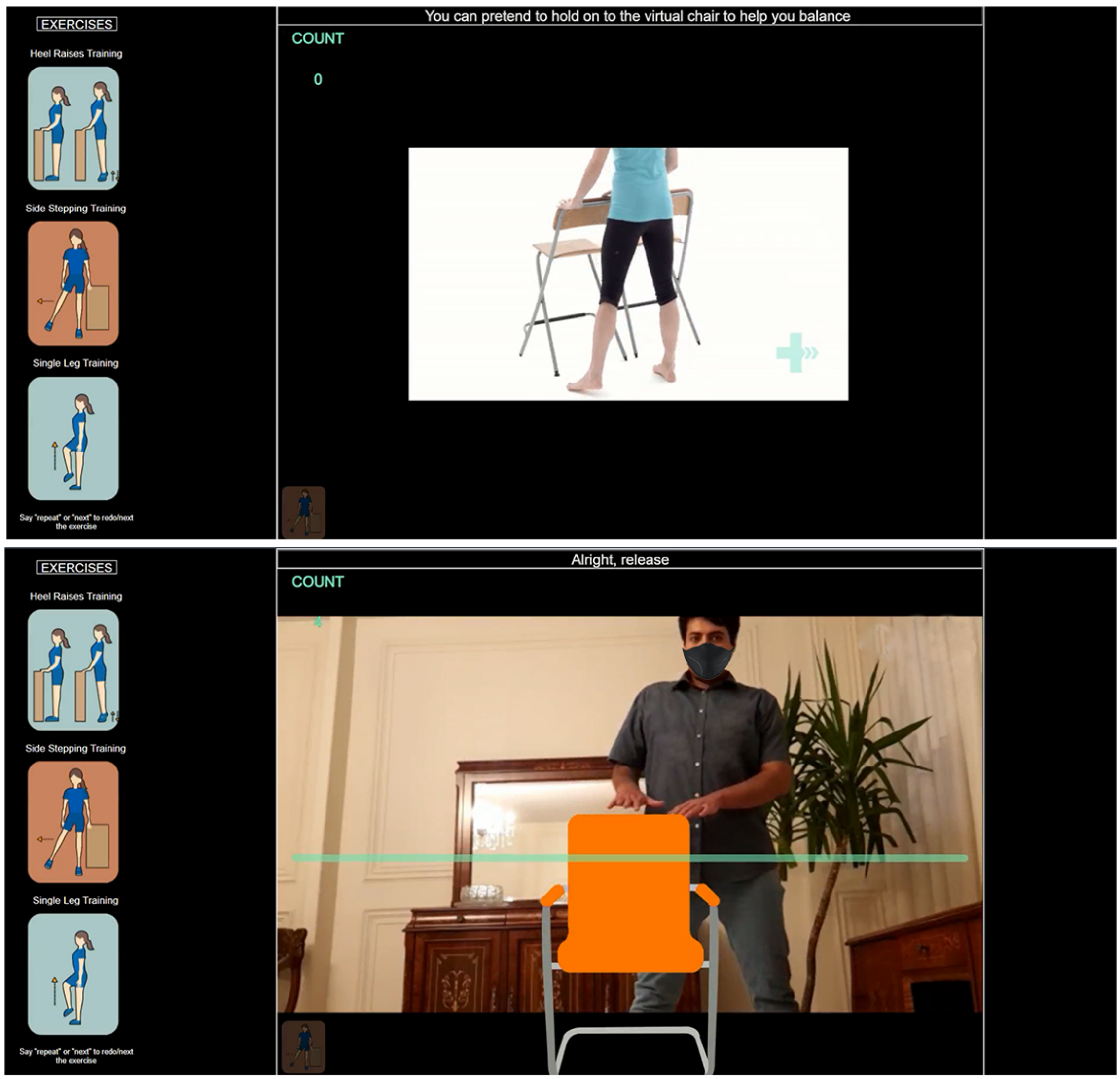


Figure 17: **Top:** The instructional video demonstrating the side stepping exercise. **Bottom:** A user doing the side stepping holding on to a virtual chair. The green line represents the position of the hips.

upright with their feet together, and try to slowly lift their foot off the ground so their thigh is perpendicular to the floor. In order to help users get a sense of balance we provide a virtual chair for them to hold on to (see Figure 18). The user is instructed to do single leg standing for 15 seconds for the left and then right leg.

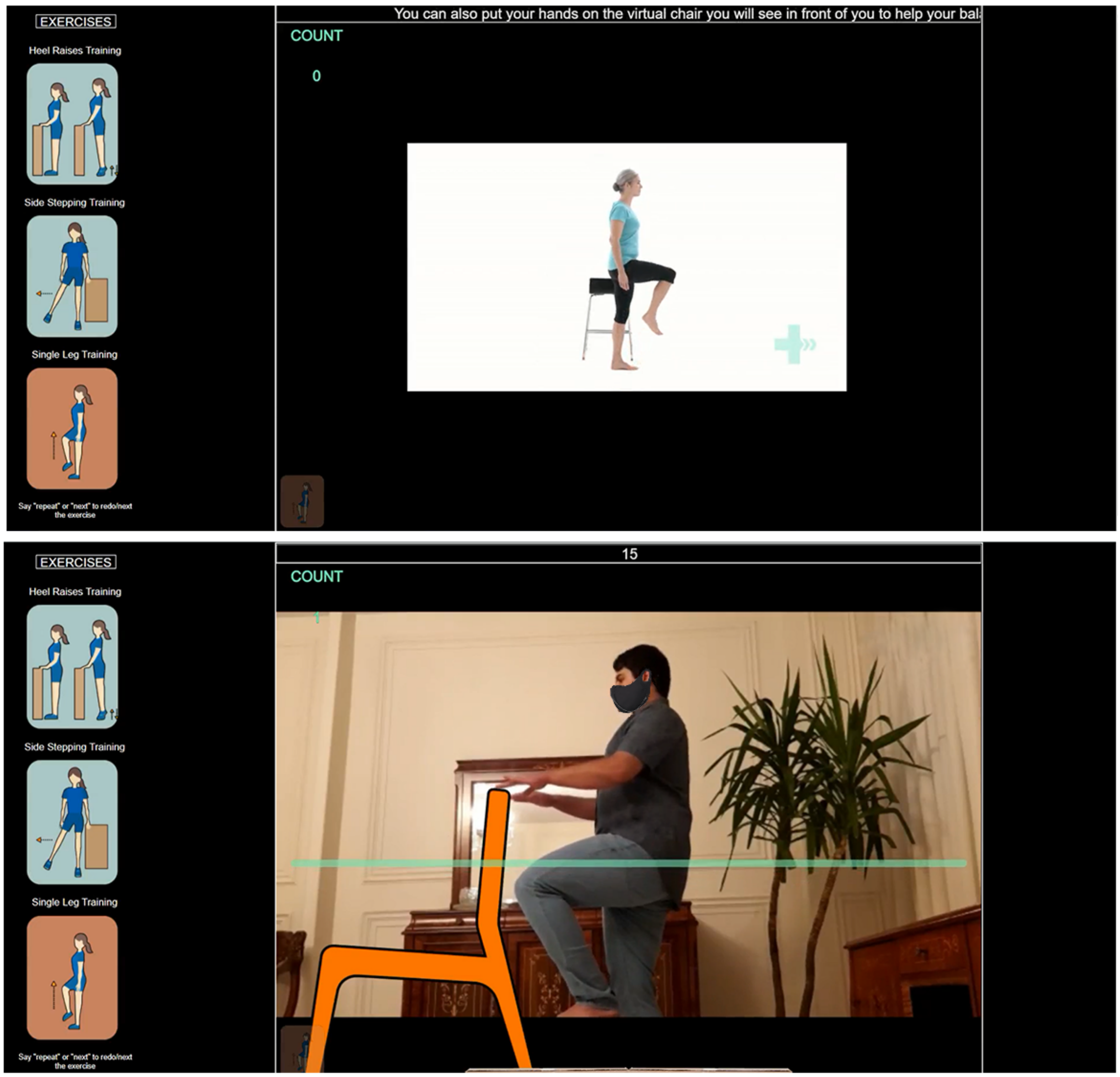


Figure 18: **Top:** The instructional video demonstrating the single leg exercise. **Bottom:** A user try to lift his foot off the ground and hold for 15 seconds, The green line represents the position of the hips

User Interface

The user interface (UI) is one of the most important factors for effective communication with users and encouraging them to redo the exercises. We developed a simple UI so that users of any age can easily interact with the application. For instance, verbal communication and images were used to convey information to the user. For those with hearing problems, the words and sentences are transcribed and displayed at the top of the interface. Also, we tried to encourage and motivate the user after each correct set using motivational and prompting words and phrases, e.g. "good job!", similar to a real trainer.

3.4 BaART Evaluation

The BaART application was evaluated with 20 participants aged 18 to 77 with 14 males and 6 females. A distribution of the ages is shown in Figure 19. Participants used the application online at their convenience. Those who agreed to participate were first sent a consent form, a description of the application, instructions on how to use the BaART application and a link to the application that they could run on the device of their choice at their convenience.

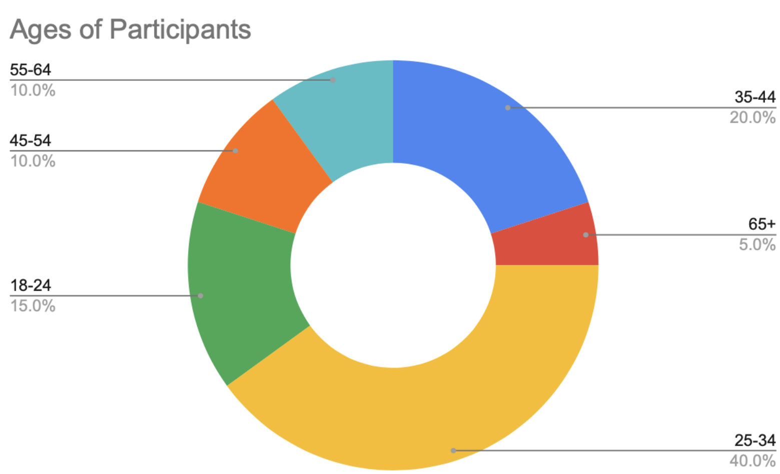


Figure 19: The BaART application was evaluated with 20 participants aged 18 to 77 with 14 males and 6 females.

3.4.1 Questionnaire: Participant Information

After the study, participants filled out a questionnaire with both background questions and questions about their experience with the application. Note that the pre-test and post-test questionnaires were combined to facilitate navigation between instructions, application and questionnaires. The summary of the analysis of the questionnaire are presented in the following sections.

Technology Use

To determine comfort with new technologies, participants were questioned on how they felt about their practical knowledge of technology, and how often they used a smartphone, tablet, or computer. The majority of participants (75%) use some form of digital device daily, whereas 25% use a device only weekly. In terms of their practical knowledge of technology, which was answered on a scale from 1 (I don't know anything about technology) to 5 (I'm very technologically savvy), the majority of participants rated their knowledge as 4 or 5, thus they felt they were technologically savvy (see Figure 20).

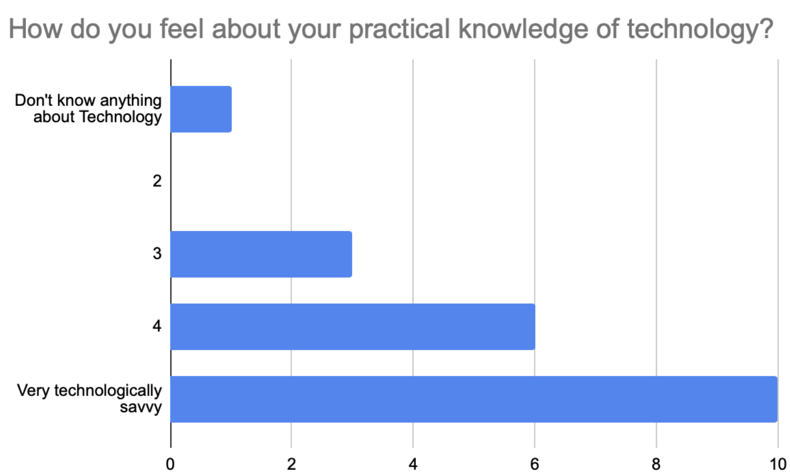


Figure 20: The majority of our participants (50%) felt they were very technologically savvy or quite savvy (30%).

We also asked how keen the participants are on trying out new technologies from 1 (Not at all) and 5 (Yes I'm usually very keen) and found that 35% were very keen, 40% were quite keen, 10% were neither keen or not keen and 15% were not keen. Lastly, we asked participants about the degree to which it was easy and simple for them to use new technologies such as mobile apps and found that the majority find it either very easy or somewhat easy (see Figure 21).

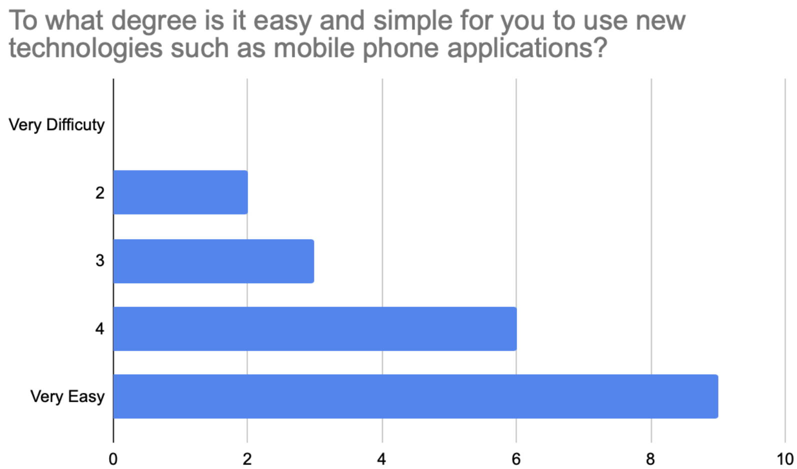


Figure 21: The majority of participants have an easy time learning new technologies.

Sport and Health Activities

The second part of the questionnaire queried users about health and fitness activities. Specifically, participants were asked if they considered themselves physically active as well as how often they participated in fitness activities such as sports, aerobics, training, etc. In a range of 1 (Not at all physically active) to 5 (Yes I'm very physically active), only 1 person (5%) considered themselves very active, 30% considered themselves somewhat active and 40% considered themselves neither very active or inactive. Lastly, 25% of participants considered themselves somewhat inactive. In terms of how often participants exercised 40% said 1-2 times a week, 40% 3-4 times a week, and 20% answered that they hardly even exercise. None of our participants exercised 5-6 times per week or everyday (see Figure 22). In terms of users ideal location for exercising, 30% prefer exercising at home, 35% at the gym, 30% prefer exercising outside and 5% said they didn't have a preference.

We also questioned participants about the main factors that lead them to exercise less frequently than they would like to. Here participants could select multiple answers. Of our population, 95% (19 people) answered "I don't have time", 6 answered "I forget to exercise", one person answered "I don't like exercising", one other said "I don't feel I need to exercise", and one person answered that they did not exercise

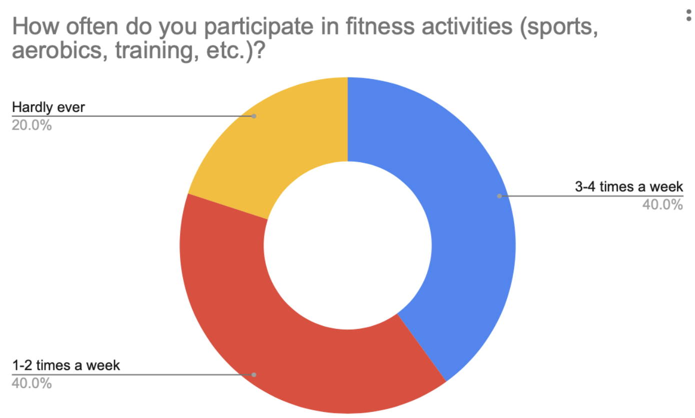


Figure 22: Of our participants, 40% exercise 1-2 times a week, 40% 3-4 times a week, and 20% hardly ever exercise. None of our participants exercised 5-6 times per week or everyday.

because they did not have someone else to exercise with.

We also asked participants about their use of online exercise training (e.g. Apple Fit+, Wii, YouTube online class) and found that most participants (70%) had not previously done online fitness classes or exercises whereas as 30% had. Lastly, we asked participants if they have ever done any rehabilitation, physiotherapy or balance training. For rehabilitation and/or physiotherapy 20% had done this, whereas 80% had not. For balance training which included balance exercises within yoga and rehabilitation/physiotherapy routines, we had 20% who had done balance training in yoga or uni-cycling routines and 75% who previously had not done any balance exercises and one person (5%) who was not sure.

Balance Disorders

As our application concerns balance exercises, we queried the participants about their experience of dizziness, falls and loss of balance. Specifically, we asked if participants had experienced a loss of balance or had a balance problem leading to falls or near falls in the last six months. The majority of participants (65%) did not have any balance problems. More than one-third of participants reported that they suffered from dizziness and loss of balance. Specifically, 10% had experienced a fall and 20% had nearly fallen. One person had experienced both a near fall and had fallen in the last six months.

3.4.2 Questionnaire: Application Evaluation

In the following section, we describe the results of the questionnaire relating to the application. First, we asked participants the type of device they use to test the BaART application. The majority (80%) tested the application on laptops, 15% tested on workstation computers and 5% (1 person) used a mobile device.

Ease of Use

Participants were asked to rate the application on a scale of 1 (Hard to understand) to 5 (Easy to understand), for the (1) video instructions and (2) auditory coaching instructions. We found that 60% found the video instructions easy to understand (rated with 5), 35% found the video instructions somewhat easy to understand and 5% found them neither hard or easy. For the spoken instructions, 55% found them easy to understand, 40% found them somewhat easy to understand and 5% found them difficult to understand. We also asked if the subjects used the voice commands (next, repeat, previous) while doing the study. Most of the participants, 80% used voice commands, 15% did not and one person (5%) answered maybe. Of those that used the voice commands, 11.5% found them very easy to use, 47% found them somewhat easy to use, 11.5% found them difficult to use, and 30% found them very difficult to use.

In terms of the overall system ease of use 50% of the participants found it very easy, 35% found it somewhat easy to use, 10% found it neither hard or easy and one participant found it difficult to use (see Figure 23).

Augmented Reality

As one of our goals was to determine if augmented reality (provided by virtual objects introduced into the video images) could aid in improving a sense of balance, we specifically asked users about their impression of this. First users were asked if they tried using the virtual object for the different exercises, we found that 60% of participants did use the virtual object, 20% did not, and 20% answered maybe. Of those that used the virtual object, we asked if it gave them an impression of helping them with their balance as they performed their exercises. On a scale of 1 (It didn't help at all) to 5 (It helped me to balance), we found that 21% found it did not help at all, 10.5%

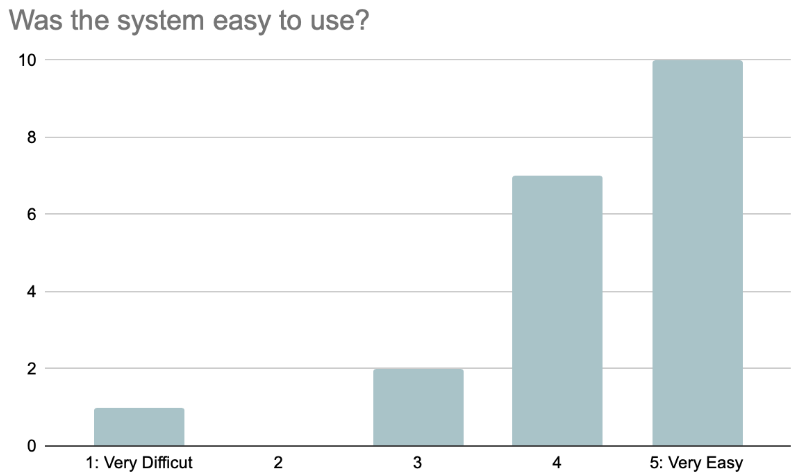


Figure 23: Fifty percent of the participants found the BaART application very easy to use, 35% found it somewhat easy to use, 10% found it neither hard or easy and one participant (5%) found it difficult to use

found it did not help, 32% were not sure, 26% thought it helped a little to balance and 10.5% found it helped them to balance. Furthermore, we asked participants if they had any comments about the use of the virtual objects for exercises, most of the participants didn't have any issue with using virtual objects, but we received some comments including *"I do not know how to use a virtual object to maintain my balance"*, *"The virtual objects didn't improve my performance"*, as well as, *"It did not help much and sometimes made me lose my focus on the exercise itself just because I wanted to keep my hands on it"*.

System Usability Scale

Finally, we provided participants with a version of the System Usability Scale (SUS) [68]. The questionnaire has ten questions positive and negative questions relating to a system's usability. The questions are answered on a scale of 1 to 5, with 1 representing Strongly Disagree and 5 representing Strongly Agree with the question's statement. The BaART application received a SUS score of 80. The average SUS score is 68, and given that a SUS value of 75 is considered good, and over 85 is excellent, our system was rated very highly with participants finding who found it easy to use, easy to learn, and felt confident using it. The results of the individual SUS questions are represented in Table 1.

System Usability Scale Questions	Average
I think that I would like to use this system frequently	3.65
I found the system unnecessarily complex	1.85
I thought the system was easy to use	4.3
I think that I would need the support of a technical person to be able to use this system	1.55
I found the various functions in this system were well integrated	4.55
I thought there was too much inconsistency in this system	1.6
I would imagine that most people would learn to use this system very quickly	4.25
I found the system very cumbersome to use	1.85
I felt very confident using the system	4.2
I needed to learn a lot of things before I could get going with this system	2.05

Table 1: The 10 questions of the System Usability Scale (SUS) with the average score from 1 (strongly disagree) to 5 (strongly agree).

Using BaART in the longer term

We asked participants to answer on a scale from 1 (very unlikely) to 5 (very unlikely) if they would use BaART if the prototype was extended with more exercises and features. Thirty percent of participants were not sure if they would use such a system, 20% percent and 50% said they would very likely use it.

We also queried specifically if the participants felt that such an application can motivate them to exercise whenever they did not feel like it or were busy, by sending notifications to their phone, on a scale of 1 (No would not help at all) to 5 (Yes it would very much help). We found that 5% said it would not help at all and another 5% said it would not help and 15% said they were not sure. However, 35% thought it would help and 40% thought it would help very much.

Lastly, we had an open question to gather the participant's overall opinions about the prototype system. Some commented on the exercises, e.g. "Increase exercises", "I think the second exercise which was step siding could be replaced to be a more effective exercise". A few users had difficulties with some aspects of the app such as their pose not being recognized although they mentioned they were able to resolve the issues by moving the camera or their position relative to the camera but this made them rate the usability lower. We also had the following comments in terms of the platform in general: *"It was good for me because I'm 77 years old and it's hard* *for me to go to the gym. But I needed a real chair or someone to help me with doing exercises."* and *"Since I am a housewife and I am usually at home, I think it is a good idea if I can do exercises at home with this app".* In general, we found that it was are older participants, who tend to spend more time at home, who were more excited about using such a system.

3.5 Discussion and Future Work

The results of our work suggest that online home exercise routines for balance training may be desirable for some users, though it may not be the ideal platform for all. Specifically, we were happy to see that our older population was excited about such a platform. Although our participant pool was very tech savvy, surprisingly 70% had not used online training applications before, but did rate the BaART application as easy and convenient to use.

In terms of the augmented reality visualization, although our sample population did not have a significant portion of users with balance disorders, those that had them, believed that they would require a real object to help them for balance rather than a virtual object. Furthermore, only 60% used the virtual object and one person even mentioned it was distracting from the exercises. We believe this still may be an interesting avenue of future research, however, it might require a better synchronization between verbal instructions and the visualization. For example, asking the users to imagine having their hands on a real chair while they step to the side. In future work, we plan to add more AR elements not necessarily for balancing but to guide the exercises and add gaming elements. One such example is an apple to pick from a tree when raising on ones tiptoes. Adding gamification elements such as this may make the application more engaging. Furthermore, a virtual trainer or avatar that does the exercises with you and adjusts their pace to the end user may also be interesting.

Given this was a first prototype, we only had three exercises and found users would be interested in having more. In the future, we plan to extend the platform to have more exercises and include other features as suggested by the participants. Additional features including: being able to either watch the tutorials again or skip them, improving the voice commands which some users had issues with (perhaps due to microphone issues or accents) and into different languages to accommodate non-native English speakers, enlarge the text instructions, and send notifications to the user to ensure they perform the exercises regularly.

Lastly, although we capture the pose and determine that the user is doing the exercise correctly in order to count the repetition, the user does not receive any visual or auditory feedback on the quality of the exercise and how to improve it if they are doing it wrong. In future work, we plan to relay this information to the user. In general, collecting, analyzing, and delivering performance-related data of users in charts could motivate users further by showing them the progress they have made.

As well as, the future directions described above it will be important to do a longer term study with more participants to look at not only how we can encourage these healthy habits in the short term but change user habits for sustained use.

3.6 Conclusion

In this paper, we describe a developed prototype for doing balancing exercises at home based on the National Institute of Clinical Excellence (NICE) guidelines for fall prevention and British Geriatric Society (ABGS) set of clinical instructions to prevent and evaluate older adults falls. We conducted a study with 20 users, both younger and older, and with and without balance disorders. The results of our study indicate that the participants liked the concept in general and found the application easy to use. Furthermore, they are interested in using this application, especially if it were extended to have more exercises and features such as notifications.

In terms of the augmented reality aspect, most users found the virtual object did not help them to balance. Although this is the case, we posit that there are additional possibilities to be explored for using AR in such a system, including adding visual feedback on the quality and correctness of the exercise, adding a virtual coach, and adding gaming elements.

The participants of the study also felt that they do not exercise as much as they like because it is time-consuming and costly. Online fitness and health applications are helping to mitigate these issues by allowing users to have personal trainers to work with them whenever and wherever they are.

Chapter 4

Conclusion

Research has shown that physical activities can improve the symptoms of balance control in order to prevent falls, particularly in adults aged 65+ [69]. However, many do not adhere to exercise and fitness routines for various reasons, including costs, and a lack of time, limited access to therapists and clinicians and a lack of motivation. The emergence of wearables, connected devices and more powerful computer vision and machine learning methods has provided new means to collect movement data and provide fitness on demand opportunities which may mitigate some of these issues.

In this thesis, we developed a prototype for doing balancing exercises at home using the National Institute of Clinical Excellence (NICE) guidelines for fall prevention and the British Geriatric Society (ABGS) set of clinical instructions to prevent and evaluate falls in older adults. The focus of our work was to ensure wide accessibility, an easy to use platform and to explore if augmented reality can be used to enhance a user's sense of balance. To make the platform easily accessible, PoseNet, which is able to run in a browser and thus requires no additional hardware beyond that of a connected digital device, was used. For usability, we developed a simple interface with voice commands and instructional videos of the exercises. Lastly, to study the impact of augmented reality on balance, we added virtual objects (i.e. chairs) that users could focus on to maintain balance.

To test the developed BaART prototype, we conducted a study with 20 users some of whom had balance issues and had experienced falls in the past year. In our sample population, the majority of participants are keen to stay active, however, fail to exercise as much as they would like due to time constraints, costs, and lack of motivation. The results of our study showed, that participants liked the concept of online balance training and found the program to be simple to use. Furthermore, they were enthusiastic about using the BaART application particularly if it was expanded to include more exercises and features like sending notifications or analyzing their progress.

In terms of augmented reality visualization, one elderly user with a significant balance disorder believed that they would require a real object to help them for balance rather than a virtual object. Furthermore, only a small majority (60%) used the virtual object and one person mentioned it was distracting from the exercises. It would be interesting to see in the future if virtual objects might be more suitable for exercises such as yoga where participants are already working with different visualization techniques and focus methods to balance or whether mixed reality elements, in general, are better suited for feedback and gamification.

4.1 Future Work

Given this was the first prototype of our system, we only had three exercises and found users would be interested in having a more complete online training program. In the future, we plan to extend the platform to have more exercises and include other features as suggested by the participants. This would include having more types of training programs, such as exercises for improving and preventing backache, neck ache, or other types of fitness routines aimed at specific ailments or disorders.

We are also interested in adding blockchain ideas to this application, for example creating a smart contract on Ethereum for getting reward tokens by doing the exercises frequently and correctly. Adding gaming elements would then allow users to have financial benefits from exercising. Added gaming elements could both guide the user in their exercises and make doing exercises more fun. For example, one can imagine having to pick an apple from a tree when doing heal raises. Such elements can make the application more engaging and may motivate users to do their exercises more frequently. Furthermore, a virtual trainer or avatar that does the exercises with you and adjusts their pace to the end user may also be interesting.

Furthermore, the application should be extended to collect, analyze, and deliver performance-related data of users, for example in charts and badges, to further motivate users. Adding notifications about progress could also help in motivation and even to just remind users to do their exercises.

In future work, we can also allow users to customize the exercises, for example, in the single-leg exercise, users can specify the difficulty of the exercise and select the amount of angle and degree that they can raise their legs and for how long. Also a module could be developed so that trainers or users could add custom exercises into the system by giving a video of a new exercise and a human activity tracking algorithm could be used to determine the key poses at different key frames of the exercises and then add these to a training set.

There are also improvements that can be made to the pose detection algorithms. In a preliminary study, one user mentioned that their knee was not being properly recognized in the single leg standing exercise. PoseNet does have some accuracy issues and sometimes fails in recognizing certain keypoints like knees and feet. Furthermore, there are sometimes position jumps in keypoint locations. In future work, we plan to develop a trained model using machine learning that especially focuses on fixing these issues.

Online fitness applications are becoming ever more important and available for people who wish to take their health into their own hands. Applications such as BaART can offer personalized fitness activities whenever and wherever a user is, thus allowing them to save both costs and time. The results of our preliminary study suggest that a future expanded version of this platform would be simple enough to use and could encourage people to do be active in order to prevent them from illness and disease.

Bibliography

[1] Tim Peterson. Healthspan is more important than lifespan, so why don't more people know about it? In *Institute for Public Health*, May 2017.

[2] World Health Organization, World Health Organization. Ageing, and Life Course Unit. *WHO global report on falls prevention in older age*. World Health Organization, 2008.

[3] Vicky Scott, Marty Pearce, and Cate Pengelly. Technical report: Hospitalizations due to falls among canadians age 65 and over an analysis of data from the discharge abstract database as presented in: Report on seniors' falls in canada (section 2.2). *British Columbia*, 2005.

[4] John Murphy and Bernard Isaacs. The post-fall syndrome. *Gerontology*, 28(4):265–270, 1982.

[5] Stephen R Lord and Russell D Clark. Simple physiological and clinical tests for the accurate prediction of falling in older people. *Gerontology*, 42(4):199–203, 1996.

[6] Summary of the updated american geriatrics society/british geriatrics society clinical practice guideline for prevention of falls in older persons. *American Geriatrics Society*, pages 148–57, 2011.

[7] Falls: Assessment and prevention of falls in older people. 2013.

[8] Tiffany E Shubert. Evidence-based exercise prescription for balance and falls prevention: a current review of the literature. *geriatric physical therapy (2001)*, 2011.

[9] M Clare Robertson, A John Campbell, Melinda M Gardner, and Nancy Devlin. Preventing injuries in older people by preventing falls: a meta-analysis of individual-level data. *American Geriatrics Society*, pages 905–11, 2002.

[10] Brian Caulfield Rob Argent, Ailish Daly. Patient involvement with home-based exercise programs: Can connected health interventions influence adherence? *JMIR Mhealth Uhealth. 2018 Mar 1*, 2018.

[11] Blanka Klimova and Petra Maresova. Elderly people and their attitude towards mobile phones and their applications—a review study. In *Advanced Multimedia and Ubiquitous Engineering*, pages 31–36. Springer, 2016.

[12] Xia Zhao, Lina Wang, Chenxi Ge, Xueting Zhen, Zheng Chen, Jie Wang, and Yuanyuan Zhou. Smartphone application training program improves smartphone usage competency and quality of life among the elderly in an elder university in china: A randomized controlled trial. *International journal of medical informatics*, 133:104010, 2020.

[13] Hon Prof John R Beard and David E Bloom. Towards a comprehensive public health response to population ageing. *Lancet (London, England)*, 385(9968):658, 2015.

[14] Rachel Proffitt and Belinda Lange. Considerations in the efficacy and effectiveness of virtual reality interventions for stroke rehabilitation: moving the field forward. *Physical therapy*, 95(3):441–448, 2015.

[15] Michalis Vrigkas, Christophoros Nikou, and Ioannis A. Kakadiaris. A review of human activity recognition methods. *Frontiers in Robotics and AI*, 2:28, 2015.

[16] Josep Maria Carmona and Joan Climent. Human action recognition by means of subtensor projections and dense trajectories. *Pattern Recognition*, 81:443–455, 2018.

[17] Sven Kreiss, Lorenzo Bertoni, and Alexandre Alahi. Pifpaf: Composite fields for human pose estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 11977–11986, 2019.

[18] Sudharshan Chandra Babu. A 2019 guide to human pose estimation with deep learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 11977–11986, 2019.

[19] Yucheng Chen, Yingli Tian, and Mingyi He. Monocular human pose estimation: A survey of deep learning-based methods. *Computer Vision and Image Understanding*, 192:102897, 2020.

[20] Xianjie Chen and Alan L Yuille. Articulated pose estimation by a graphical model with image dependent pairwise relations. In *Advances in neural information processing systems*, pages 1736–1744, 2014.

[21] D. Nguyen and H. Le. Kinect gesture recognition: Svm vs. rvm. In *2015 Seventh International Conference on Knowledge and Systems Engineering (KSE)*, pages 395–400, 2015.

[22] Marcelo Rovai. Realtime multiple person 2d pose estimation using tensorflow2.x. pages 11977–11986, 2020.

[23] Jonathan Ragan-Kelley and Ren Ng. kinematics and motion capture. 2017.

[24] Maureen Furniss. Motion capture. *posted at http://web. mit. edu/mit/articles/index—furniss. html on Dec*, 19, 1999.

[25] Alberto Menache. *Understanding motion capture for computer animation and video games*. Morgan kaufmann, 2000.

[26] Jianting Ma. Design of omnidirectional mobile platform for analysis of postural balance in transport. 09 2014.

[27] M Rahul. Review on motion capture technology. *Global Journal of Computer Science and Technology*, 2018.

[28] Margaret S Geroch. Motion capture for the rest of us. *Journal of Computing Sciences in Colleges*, 19(3):157–164, 2004.

[29] Rebecca Perry. Out-of-body workspaces: Andy serkis and motion capture technologies, Mar 2019.

[30] Motion capture studio. posted at Middlesex University Londonhttps://www.mdx.ac.uk/__data/assets/image/0016/120229/Motion-Capture-Studio.jpg.

[31] Kinect camera. Microsofthhttp://www.xbox.com/en-US/kinect/default.htm.

[32] Jungong Han, Ling Shao, Dong Xu, and Jamie Shotton. Enhanced computer vision with microsoft kinect sensor: A review. *IEEE Transactions on Cybernetics*, 43(5):1318–1334, 2013.

[33] Ahmed Ben Jmaa, Walid Mahdi, Yousra Ben Jemaa, and Abdelmajid Ben Hamadou. A new approach for hand gestures recognition based on depth map captured by rgb-d camera. *Computacion y Sistemas*, 20:709–721, 12 2016.

[34] Jungong Han, Ling Shao, Dong Xu, and Jamie Shotton. Enhanced computer vision with microsoft kinect sensor: A review. *IEEE transactions on cybernetics*, 43(5):1318–1334, 2013.

[35] Yucheng Chen, Yingli Tian, and Mingyi He. Monocular human pose estimation: A survey of deep learning-based methods. *Computer Vision and Image Understanding*, 192:102897, 03 2020.

[36] Saghir Alfasly, Beibei Liu, Yongjian Hu, Yufei Wang, and Chang-Tsun Li. Auto-zooming cnn-based framework for real-time pedestrian detection in outdoor surveillance videos. *IEEE Access*, 7:105816 – 105826, 08 2019.

[37] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. Realtime multi-person 2d pose estimation using part affinity fields. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[38] Hao-Shu Fang, Shuqin Xie, Yu-Wing Tai, and Cewu Lu. Rmpe: Regional multi-person pose estimation. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Oct 2017.

[39] Emily-Jane Rolley-Parnell, Dimitrios Kanoulas, Arturo Laurenzi, Brian Delhaisse, Leonel Rozo, Darwin Caldwell, and Nikos Tsagarakis. Bi-manual articulated robot teleoperation using an external rgb-d range sensor. 09 2018.

[40] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh. Openpose: realtime multi-person 2d pose estimation using part affinity fields. *arXiv preprint arXiv:1812.08008*, 2018.

[41] Microsoft coco.

[42] Getting started with posenet. Tensorflow sciencehttps://blog.tensorflow.org/2018/05/real-time-human-pose-estimation-in.html, 2018.

[43] Liang-Chieh Chen Spyros Gidaris Jonathan Tompson Kevin Murphy George Papandreou, Tyler Zhu. Personlab: Person pose estimation and instance segmentation with a bottom-up, part-based, geometric embedding model. 2018.

[44] Negar Haghbin. and Marta Kersten-Oertel. Multimodal cueing in gamified physiotherapy: A preliminary study. In *Proceedings of the 7th International Conference on Information and Communication Technologies for Ageing Well and e-Health - ICT4AWE,*, pages 137–145. INSTICC, SciTePress, 2021.

[45] Rayele Moreira, Ariel Teles, Renan Fialho, Rodrigo Baluz, Thalyta Cibele Santos, Rômulo Goulart-Filho, Laiane Rocha, Francisco José Silva, Nishu Gupta, Victor Hugo Bastos, et al. Mobile applications for assessing human posture: A systematic literature review. *Electronics*, 9(8):1196, 2020.

[46] Rayele Moreira, Ariel Teles, Renan Fialho, Thalyta Cibele Passos Dos Santos, Samila Sousa Vasconcelos, Itamara Carvalho de Sá, Victor Hugo Bastos, Francisco Silva, and Silmar Teixeira. Can human posture and range of motion be measured automatically by smart mobile applications? *Medical hypotheses*, 142:109741, 2020.

[47] Fabian Herrera, Rodrigo Niño, Carlos Enrique Montenegro-Marín, Paulo Alonso Gaona-García, Iñigo Sarría Martínez de Mendívil, and Rubén González Crespo. Computational method for monitoring pauses exercises in office workers through a vision model. *Journal of Ambient Intelligence and Humanized Computing*, pages 1–9, 2020.

[48] Hang Yan, Beichen Hu, Gang Chen, and E Zhengyuan. Real-time continuous human rehabilitation action recognition using openpose and fcn. In *2020*

3rd International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE), pages 239–242. IEEE, 2020.

[49] Suman Deb, Alpana Sharan, Shivangi Chaturvedi, Ankit Arun, and Aayush Gupta. Interactive dance lessons through human body pose estimation and skeletal topographies matching. *International Journal of Computational Intelligence & IoT*, 2(4), 2018.

[50] Martin G Jorgensen, Uffe Laessoe, Carsten Hendriksen, Ole Bruno Faurholt Nielsen, and Per Aagaard. Efficacy of nintendo wii training on mechanical leg muscle function and postural balance in community-dwelling older adults: a randomized controlled trial. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, 68(7):845–852, 2013.

[51] Fariba Mostajeran, Nikolaos Katzakis, Oscar Ariza, Jann Philipp Freiwald, and Frank Steinicke. Welcoming a holographic virtual coach for balance training at home: two focus groups with older adults. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 1465–1470. IEEE, 2019.

[52] I Kouris, Michail Sarafidis, Thelma Androutsou, and Dimitris Koutsouris. Holobalance: an augmented reality virtual trainer solution forbalance training and fall prevention. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 4233–4236. IEEE, 2018.

[53] Elise Klæbo Vonstad, Xiaomeng Su, Beatrix Vereijken, Kerstin Bach, and Jan Harald Nilsen. Comparison of a deep learning-based pose estimation system to marker-based and kinect systems in exergaming for balance training. *Sensors*, 20(23):6940, 2020.

[54] Corinna Ogonowski, Konstantin Aal, Daryoush Vaziri, Thomas Von Rekowski, Dave Randall, Dirk Schreiber, Rainer Wieching, and Volker Wulf. Ict-based fall prevention system for older adults: qualitative results from a long-term field study. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 23(5):1–33, 2016.

[55] Sandro Hardy, Stefan Göbel, and Ralf Steinmetz. Adaptable and personalized game-based training system for fall prevention. In *Proceedings of the 21st ACM international conference on Multimedia*, pages 431–432, 2013.

[56] Thelma Androutsou, IN Kouris, Athanasios Anastasiou, Sotiris Pavlopoulos, Fariba Mostajeran, Doris-Eva Bamiou, Gregory J Genna, Sergi G Costafreda, and Dimitrios Koutsouris. A smartphone application designed to engage the elderly in home-based rehabilitation. *Frontiers in Digital Health*, 2, 2020.

[57] Christiana Tsiourti, Emilie Joly, Cindy Wings, Maher Ben Moussa, and Katarzyna Wac. Virtual assistive companions for older adults: qualitative field study and design implications. In *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*, pages 57–64, 2014.

[58] Isa Brito Félix, Mara Pereira Guerreiro, Afonso Cavaco, Ana Paula Cláudio, Anabela Mendes, João Balsa, Maria Beatriz Carmo, Nuno Pimenta, and Adriana Henriques. Development of a complex intervention to improve adherence to antidiabetic medication in older people using an anthropomorphic virtual assistant software. *Frontiers in Pharmacology*, 10, 2019.

[59] Bickmore Timothy W, Silliman Rebecca A, Nelson Kerrie, Cheng Debbie M, Winter Michael, Henault Lori, and Paasche-Orlow Michael K. A randomized controlled trial of an automated exercise coach for older adults. *Journal of the American Geriatrics Society*, 2009.

[60] Inaki Merino Albaina, Thomas Visser, Charles A.P.G. van der Mast, and Martijn H. Vastenburg. Flowie: A persuasive virtual coach to motivate elderly individuals to walk. 8 2009.

[61] Yasmin Felberbaum, Joel Lanir, and Patrice L Weiss. Challenges and requirements for technology to support mobility of older adults. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–6, 2018.

[62] Jung Younbo. Promoting active ageing through virtual augmented reality. 2019.

[63] Michael Lo Bianco, Sonja Pedell, and Gianni Renda. Augmented reality and home modifications: A tool to empower older adults in fall prevention. In

Proceedings of the 28th Australian Conference on Computer-Human Interaction, page 499–507, New York, NY, USA, 2016. Association for Computing Machinery.

[64] Jeonghun Ku, Yeun Joon Kim, Sangwoo Cho, Teo Lim, Hye Sun Lee, and Youn Joo Kang. Three-dimensional augmented reality system for balance and mobility rehabilitation in the elderly: A randomized controlled trial. *Cyberpsychology, Behavior, and Social Networking*, 22(2):132–141, 2019. PMID: 30596530.

[65] American Geriatrics Society, Geriatrics Society, American Academy Of, and Orthopaedic Surgeons Panel On Falls Prevention. Guideline for the prevention of falls in older persons. *Journal of the American Geriatrics Society*, 49(5):664–672, 2001.

[66] J Am Geriatr Soc. Panel on prevention of falls in older persons, american geriatrics society and british geriatrics society. summary of the updated american geriatrics society/british geriatrics society clinical practice guideline for prevention of falls in older persons. pages 59(1):148–157. doi:10.1111/j.1532–5415.2010.03234.x, 2011.

[67] Shilpa Tyagi, Daniel SY Lim, Wilbert HH Ho, Yun Qing Koh, Vincent Cai, Gerald CH Koh, and Helena Legido-Quigley. Acceptance of tele-rehabilitation by stroke patients: perceived barriers and facilitators. *Archives of physical medicine and rehabilitation*, 99(12):2472–2477, 2018.

[68] John Brooke. Sus-a quick and dirty usability scale (in "usability evaluation in industry", pw jordan, b thomas, i mclelland, ba weerdmeester (eds)). pages 189–194, 1996.

[69] Physical activity. World Health Organizationhttps://www.who.int/en/news-room/fact-sheets/detail/physical-activity.