AN INTELLIGENT FITNESS POSTURE CORRECTION AND SUGGESTION SYSTEM FOR VISUALLY IMPAIRED USING COMPUTER SCIENCE AND IOT SYSTEM (INTERNET OF THINGS)

Sean Li[1], Andrew Park[2]

[1]Sage Hill School, 20402 Newport Coast Dr, Irvine, CA 92603
[2]Computer Science Department, California State Polytechnic University, Pomona, CA 91768

ABSTRACT

Visual impairments are an unfortunate trait that can be obtained during birth, unexpected incidents, or natural causes like age. These reasons shouldn't be the determining factor of whether the individual can remain physically healthy or not. Our goal is to provide an assistant to improve fitness journeys of those who need guidance on form and varieties of exercises. We intend to solve these problems with a dedicated fitness device that has voice guidance and a built-in camera to help lead users into building good exercise habits to prevent the risk of injury and help them remain in their best possible shape [2]. Some key technologies would be the Raspberry Pi, USB speaker, a mini microphone, and the use of AI [3]. Some challenges that we faced would be how to make the AI comprehend the words clearly and ignore outside noises. We fixed this issue with the use of OpenCV and MediaPipe. This idea is ultimately something that people should use because it provides the tools to a successful physical transformation or just daily routines to keep someone in shape, giving the visually impaired a key to a more convenient life [4].

KEYWORDS

Computer Vision, Accessibility, Fitness, Voice Command

1. INTRODUCTION

Having proper technique and form is crucial when performing a physical exercise to avoid incurring injuries as well as getting the most out of the action; however, this is significantly more difficult to verify for someone with a visual impairment. Visual impairment is a disability that an individual is born with or developed later in life. The World Health Organization estimates a 7%-10% increase in blindness every five years [1]. In 2020, an overall estimate of 43.3 million suffer from visual impairment of all ages and 295 million with MSVI [19]. These statistics are higher than what the WHO estimated originally at 277 million people with MSVI and 38.5 million for blindness. Having more resources to help people with blindness on their fitness journey will serve as a major boost to their motivation and lead to better fitness outcomes. Physical fitness is something that everyone of all ages should enjoy and experience safely and at their own pace regardless of handicap.

David C. Wyld et al. (Eds): NATL, SEAPP, SIGEM, MLTEC, CSEA – 2024
pp. 29-36, 2024. - CS & IT - CSCP 2024 DOI: 10.5121/csit.2024.14204

In all 3 methodologies, each individual technology has the goal to help someone who is impaired in some way. Methodology A, is inspired by the passion for yoga and helps teach yoga to people with visual impairments, giving all people a chance to feel inner peace and have a good stretch. Methodology B, is similarly influenced by yoga for the impaired by creating a smart mat that uses vibrations to lead the individual to have correct form. Methodology C, is a device in the base of your shoe which helps the impaired know when and where to turn by the use of vibrations, though it's purpose is not for yoga, it is for everyday use. All of these devices help the impaired with fitness and navigation. Some of the shortcomings in common with all these devices would be the limitations such as Methodology A and B is that it is only exclusive for yoga and doesn't include any other exercises while Methodology C, relies on vibrations and sometimes that could be misleading due to oncoming traffic or obstacles [18]. Our project uses skeletal and voice guidance for a variety of exercises ranging from weights to dynamic stretching, our project aims to improve on overall fitness such as muscle growth and healthy habits.

To address this problem, we propose a system that combines both pose estimation along with a voice-command based interface to allow people with visual impairments to receive guidance on proper physical fitness technique. Such a system essentially works through a device with a microphone, speaker, and camera which walks the user through what exercise they would like the machine to look at. The camera then records their motion with that specific exercise in mind, pipes that feed into a pose estimation AI, and provides corrective advice in real time as they are trying to perform the action [5]. This addresses the need for a visual analysis of what the participant is doing in order to understand what needs to be corrected which said participant would not otherwise be able to do given their lack of eyesight. We believe this to be a better solution than some of the systems currently out there as it provides an all-in-one package for the end user that avoids requiring vision to implement; it should also be noted that there are not very many solutions tackling the issue of assisting blind people in the context of physical exercise to begin with.

2. CHALLENGES

In order to build the project, a few challenges have been identified as follows.

2.1. Visual Analysis

One of the major components that needs to be addressed is that there must be a way for the device to perform visual analysis on behalf of a user who may not be able to do so on their own. Our analysis would need to receive the input data in a format that is compatible with the AI, which a video feed is not immediately able to do. To that end, we can use solutions like the opencv library to interact with the camera hardware to extract a live video feed and format it into a form we can use for analysis [6].

2.2. Interpret Data

The next stage after receiving video input would be to figure out what pose the user is currently in. This requires that our system is able to take an individual frame and interpret an otherwise nondescript array of pixel data to figure out both if a person is present in the frame and what body parts there are. We can accomplish such a task through the use of libraries like mediapipe, which provides a set of machine learning models that were trained to understand such information and mark accordingly. Through the use of AI, we would then be able to get the pose landmark data to perform proper analysis of the user's performance [17].

2.3. User Experience

To achieve the accessibility goal of the project, the user experience flow must be easy to understand and interact with. This would require a robust voice recognition system that is able to both give out instructions and queries as well as understanding what the user is trying to say while also filtering out false positives that could lead to frustration or otherwise unintended actions by the machine [16]. The statements made by the device should be easy to understand and also able to account for various edge cases and tasks a user would want to get done quickly.

3. SOLUTION

To achieve the accessibility goal of the project, the user experience flow must be easy to understand and interact with. This would require a robust voice recognition system that is able to both give out instructions and queries as well as understanding what the user is trying to say while also filtering out false positives that could lead to frustration or otherwise unintended actions by the machine [7]. The statements made by the device should be easy to understand and also able to account for various edge cases and tasks a user would want to get done quickly.

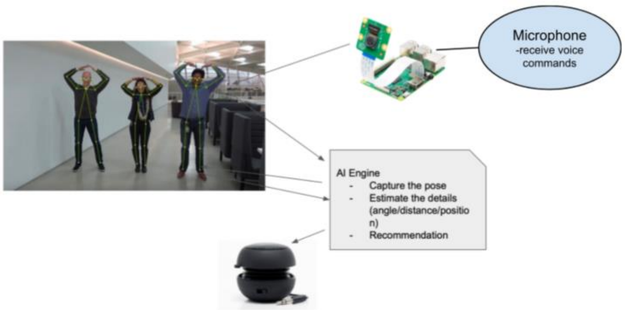

Figure 1. Overview of the solution

The first major component of BlindVision is the hardware itself. Such a system needs a camera and microphone to handle inputs and a speaker to handle outputs alongside a central computing device to power it all [8].

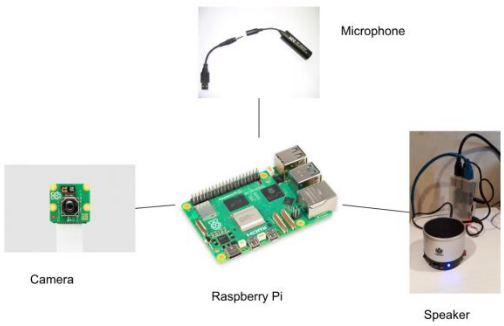

Figure 2. The component

```python
def validate_response(self, message, mode=0):
    message = message['text'].strip()
    if mode == 0: # checking yes
        affirmatives = ["yes", "yea", "yeah", "sure"]
        for word in affirmatives:
            if word in message:
                return True
    elif mode == 1: # checking no
        affirmatives = ["no", "nope"]
        for word in affirmatives:
            if word in message:
                return True
    elif mode == 2: # checking name
        banned_phrases = {"hello there", "yes", "no", "there"}
        if 1 <= len(message) <= 20 and message not in banned_phrases and len(message.split()) <= 10:
            return True
    elif mode == 3: # checking age
        affirmatives = {"thirteen", "fourteen", "fifteen", "sixteen", "seventeen", "eighteen", "nineteen", "twenty", "twenty one'
        if message in affirmatives:
            return True
    return False
```

Figure 3. Screenshot of code 1

This function is intended to validate the response said by the user. The code runs in the program when the user responds to a question by the device. The message is what the user puts out and what the AI is comprehending [9]. The mode is the expectation of the system, so that it can be primed to listen for a certain type of response. The intent of this function is to allow the user to say something in a different way without the device failing.

The pose estimation system is the second required component of BlindVision. In order to actually analyze how well a user is exercising, it needs data like what angles and degrees the user's position shows to evaluate correctness of form.

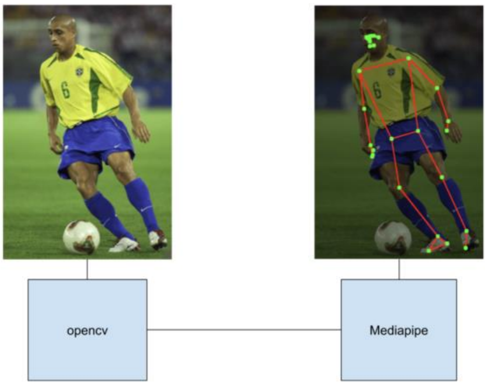

Figure 4. The pose estimation system

```python
import cv2
import numpy as np
import posefile as pose

def analyze_bicep_curl(voice, video_path=0, left=False, speak_count=False, speak_warning=False):
    cap = cv2.VideoCapture(video_path)
    detector = pose.PoseDetection()
    direction = 0 # 0 is up, 1 is down
    count = 0
    last_count = 0
    last_percentage = 0
    delay = 0

    print("Checking video input...")

    while cap.isOpened():
        success, frame = cap.read()

        if not success:
            print("Could not get frame.")
            break

        frame = cv2.resize(frame, (720, 480))
        frame = detector.find_pose(frame, False)
```

Figure 5. Screenshot of code 2

This code is to analyze the bicep curl exercise. Before we do the analysis, the function sets up the camera device and the detector, as well as relevant variables such as which direction the limb is moving or should be moving and how many repetitions it has done. While the camera is able to see something, we try to find the pose of the user using mediapipe.

The final required component of BlindVision is a proper analysis engine that can take the landmark data from the pose estimation system to then figure out if the user's exercise form is correct [10]. This is paired with a speaker to provide verbal feedback to help you recognize mistakes and improve over time with guidance.

Bench Press: "Got it, please position yourself where I can see you to begin [bench presses]. To stop, just tell me to stop." 5 seconds later: "I still cannot see you, please move to the left/right/up/down." "I still cannot see you, please move into the view of the camera." "Thank you, we will begin in 3…2…1."
Stop/Done/Finished/Complete: "Ending the current session for [insert exercise]. [Exercise specific summary]"

Figure 6. Screenshot of feedback

```python
if len(landmark_list) != 0:
    if left:
        # find left angle
        angle = detector.find_angle(frame, 11, 13, 15)
    else:
        # find right angle
        angle = detector.find_angle(frame, 12, 14, 16)

    percentage = np.interp(angle, (25, 155), (0, 100))
    # print(percentage, angle)

    if direction == 0:
        # raising the counter for delay
        if last_percentage > percentage:
            delay += 1

        # notifying upon reaching limit
        if delay >= 5: # change as needed for timing
            print("Lift your arm higher")
            if speak_warning:
                voice.speak("Lift your arm higher")
            delay = 0
    elif direction == 1:
        # raising the counter for delay
        if last_percentage < percentage:
            delay += 1

        # notifying upon reaching limit
        if delay >= 5: # change as needed for timing
            print("Lower you arm")
            if speak_warning:
                voice.speak("Lower your arm")
            delay = 0

    if percentage == 100 and direction == 0:
        delay = 0
        direction = 1
        count += 0.5
    elif percentage == 0 and direction == 1:
        delay = 0
        direction = 0
        count += 0.5

    last_percentage = percentage

    if speak_count and last_count != int(count):
        last_count = count
        voice.speak(last_count)
```

Figure 7. Screenshot of code 3

If the detector finds your body then depending on if you are analyzing your left or right arm, it will get the angle of that arm and calculate the bent degree. We then check if the arm is moving in the direction it should be going and warn them if the user is going in the wrong direction. Once the user completes one range of motion for an exercise, the direction is then switched back down. As we do this, the system will count and say to the total rep count.

4. EXPERIMENT

4.1. Experiment 1

One potential blind spot we will want to address is how well the system works for a blind person given the project's focus on tailoring to them. It is meaningfully different from the experience of a person who can see?

To conduct this experiment, participants will be separated into three distinct groups. The first group consists of people who are not blind, the second group consists of people who were born blind, and the final group consists of people who either became blind later in life or were given a blindfold to simulate someone who became blind later on. Each participant is going to go through the exact same process and will be trying the same exercises. Their session will be recorded so that we can then analyze how their behavior compares to other groups. To supplement this study, we will also conduct a survey with the participants to get information regarding demographics.

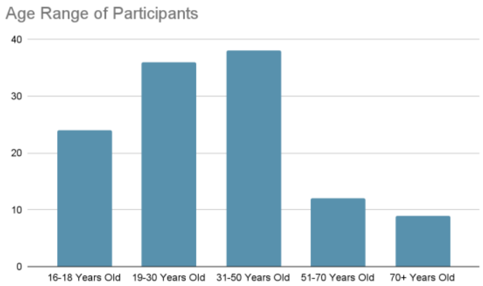

Figure 8. Figure of experiment 1

4.2. Experiment 2

Another aspect we will need to explore is the general impact on exercise frequency the BlindVision device has on the general user compared to how they would otherwise.

To properly evaluate the general impact of the device, participants will be gathered that have similar levels of fitness to one another and split into two different groups. One group will be given nothing while the other will be given the device. The participants will then be instructed to report how often they exercised each day over the course of twelve weeks. It is important to let participants know that there will be no reward or penalty based on how often they exercise, but to just exercise however much they want.

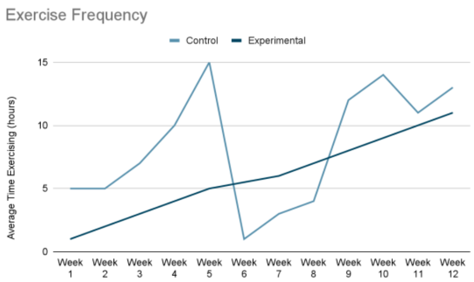

Figure 9. Figure of experiment 2

5. RELATED WORK

Rector, Kyle & Bennett, Cynthia & Kientz, Julie created a system that uses a microsoft kinect along with a skeletal tracking system and auditory feedback to help teach yoga poses even for people with visual impairments [11]. Their solution was generally effective and seen positively by the participants. It is important to note that the number of participants was 16, which while enough to get meaningful information from, a larger group would help better evaluate effectiveness. There is also the issue of their exergame using the Microsoft Kinect, which can be expensive and unyielding in terms of usage. Another limitation of Eyes-Free Yoga is that they only focus on 6 different yoga poses. We improve on what is tried by this team by expanding the range of exercises the system can analyze beyond just yoga poses. This appeals to a larger audience and also for those who are looking for similar tools outside of yoga.

Han, D. (2022) created a smart yoga mat system for visually impaired people. The way it essentially works is through a layer of PU rubber with a layer of force-sensitive resistors below it [12]. Internal pumps then expand or contract to change the friction characteristic of the mat’s surface to help guide the user through whatever yoga move they are trying to do. This is combined with a wearable gyroscope, voice guidance, vibration, and music to continue reminding the user of what they need to do. It is uncertain to what degree this solution is effective as the creator does not mention tangible testing results with a physical prototype. In terms of limitation, their solution is specifically targeting yoga, which BlindVision is looking to go beyond into general fitness. Another limitation is that this requires several specialized pieces of equipment that also include things you need to wear. We seek to improve upon this by delivering a device that is much more portable and does not require the user to strap anything to themselves.

Tachiquin, R. et al (2021) created an assistive device for the blind that uses a combination of a smartphone and an electronic module controlling a tactile display embedded into shoes [13]. They aim to essentially utilize the GPS and computing capabilities of a smartphone combined with a means to deliver tactile feedback on the user’s foot to help them navigate around. Test results from 20 subjects indicated that this solution is very effective at helping the blind navigate around with 100% of users correctly interpreting the forward and backward signals while the left had a mean success rate of 95% and the right had a success rate of 93%. Of course, the scope of this solution is not necessarily directed at fitness, but is nonetheless relevant as it allows individuals to actively navigate around outside. It should be noted that this navigation system is designed to be complementary to other guides like sticks or dogs and not an outright replacement. Our system serves to target a more focussed application in the context of exercise independent of any other aids while also being relatively inexpensive to do so.

6. CONCLUSIONS

A notable design restriction for interacting with our system is that it requires our custom device for a seamless experience. There is also the risk of certain edge cases in terms of voice commands that can lead to some unintended behaviors. While some people are physically healthier or flexible, our system does not yet account for such physical differences when guiding people through the various exercises. Future iterations of BlindVision will seek to address many of the associated issues with the aforementioned limitations and will also incorporate more exercises and variations, as well as an improved voice system in terms of quality and usability [14]. We can add exercises like tricep pulldowns alongside its variants such as rope pulldowns and overhead skull-crushers. The voice system itself can be improved to sound more human through the use of more advanced voice synthesizers alongside more active response checks to be able to react faster to user commands [15].

Continuous improvements and iterations for BlindVision will be targeted toward increasing usability while also optimizing options for distribution so that those suffering from vision impairment will have another option to explore going forward.

REFERENCES

[1] Alcaraz-Rodríguez, Virginia, et al. "Influence of physical activity and sport on the inclusion of people with visual impairment: A systematic review." International journal of environmental research and public health 19.1 (2021): 443.

[2] Rector, Kyle, Cynthia L. Bennett, and Julie A. Kientz. "Eyes-free yoga: an exergame using depth cameras for blind & low vision exercise." Proceedings of the 15th international acm sigaccess conference on computers and accessibility. 2013.

[3] Zhao, Ziyi, et al. "3d pose based feedback for physical exercises." Proceedings of the Asian Conference on Computer Vision. 2022.

[4] Aguilar-Ortega, Rafael, et al. "Uco physical rehabilitation: New dataset and study of human pose estimation methods on physical rehabilitation exercises." Sensors 23.21 (2023): 8862.

[5] Ginis, Kathleen A. Martin, et al. "Participation of people living with disabilities in physical activity: a global perspective." The Lancet 398.10298 (2021): 443-455.

[6] Wei, Guixiang, et al. "Spatial – Temporal Self-Attention Enhanced Graph Convolutional Networks for Fitness Yoga Action Recognition." Sensors 23.10 (2023): 4741.

[7] Richardson, Mike, Karin Petrini, and Michael J. Proulx. "Access to exercise for people with visual impairments during the Coronavirus-19 pandemic." British Journal of Visual Impairment 41.2 (2023): 448-463.

[8] Izquierdo, Mikel, Gustavo Duque, and John E. Morley. "Physical activity guidelines for older people: knowledge gaps and future directions." The Lancet Healthy Longevity 2.6 (2021): e380-e383.

[9] Sweeting, Joanna, et al. "Physical activity interventions for adults who are visually impaired: A systematic review and meta-analysis." BMJ open 10.2 (2020): e034036.

[10] Surakka, Airi, and Tero Kivela. "Motivating visually impaired and deaf-blind people to perform regular physical exercises." British journal of visual impairment 26.3 (2008): 255-268.

[11] Brunes, Audun, Eli Krokstad, and Liv Berit Augestad. "How to succeed? Physical activity for individuals who are blind." British Journal of Visual Impairment 35.3 (2017): 264-274.

[12] Rogge, Ann-Kathrin, et al. "Balance, gait, and navigation performance are related to physical exercise in blind and visually impaired children and adolescents." Experimental brain research 239 (2021): 1111-1123.

[13] Augestad, Liv Berit, and Lin Jiang. "Physical activity, physical fitness, and body composition among children and young adults with visual impairments: A systematic review." British Journal of Visual Impairment 33.3 (2015): 167-182.

[14] Kamelska, Anna Malwina, and Krzysztof Mazurek. "The assessment of the quality of life in visually impaired people with different level of physical activity." Physical Culture and Sport. Studies and Research 67.1 (2015): 31-41.

[15] Gaetano, Raiola, Altavilla Gaetano, and Gomez Paloma Filippo. "Effects of physical activity and sports in the reduction of stereotypy in blind subjects." Sport Sci 8 (2015): 52-55.

[16] Han, Dan. A smart yoga mat system designed for visually impaired people. Rochester Institute of Technology, 2022.

[17] Tachiquin, Ricardo, et al. "Wearable urban mobility assistive device for visually impaired pedestrians using a smartphone and a tactile-foot interface." Sensors 21.16 (2021): 5274.

[18] Darwin, Darwin. "The Effectiveness Of Implementation Of Development Plan Consultations In The Development Of Work Plan For The Regional Government Of Rokan Hilir Regency." (2022).

[19] Vision Loss Expert Group of the Global Burden of Disease Study. "Global estimates on the number of people blind or visually impaired by glaucoma: A meta-analysis from 2000 to 2020." Eye 38.11 (2024): 2036.