

Posture Detection and Comparison of Different Physical Exercises Based on Deep Learning Using Media Pipe, Opencv

Prof. Dr. Sunil Kale*, Nipun Kulkarni**, Sumit Kumbhkarn**, Atharva Khuspe**, Shreyash Kharde**
Dept. of Computer Engineering, Pimpri Chinchwad College of Engineering, Pune.

Abstract:

The occasion of this paper is to improve the body posture during exercise. The AI based smart system to suggest better body posture by live image and video sensing is implemented. Mental health is directly dependent on our daily routine and physical workout. The physical exercise practices are very important to keep the hormone level at normal and stay mindful. Thus the physical practices can be done in a proper way without causing any harm to your body. So the way of exercise should be monitored all the time and corrected if some changes are observed. Study AI-based exercise monitoring systems by using python moduli like mediapipe, tensorflow, matplotlib, OpenCV, etc. Which can read the input and display the image with precise output.

Keywords: Physical fitness and its importance, Python, Image and Video Processing, Mental health, image detection.

I. Introduction:

Artificial Intelligence is the future of technology. As per the technological leap, we have seen things get processed quickly; this is possible because of AI and powerful libraries in Python. The motive behind this paper is to understand how AI-enabled Healthcare systems work combined with Artificial intelligence and machine learning algorithms and how these systems have a significant impact on human life in terms of comfort, cost-effectiveness, and environmental friendliness. AI based intelligent systems have a vast application in various fields, like agriculture, waste management, home security, and healthcare, etc. Image and video processing in exercise prediction is a broader view of a better lifestyle. Machine learning and artificial intelligence have seen the upliftment of neural networks in the last few years. Neural networks are made up of several layers, which are called deep neural networks. There are many deep learning models specialized in solving many tasks. Mental health is always directly dependent on physical fitness. To help in recovering from depression, daily exercises are suggested by practitioners. The challenge with physical exercise is that it is essential to do it correctly because any wrong position maintained during an exercise session might render the activity ineffective and possibly cause inconvenience. The attendance of a coach is required so that the meeting can be monitored and the individual's posture can be adjusted. Because not every client approaches or has access to a trainer, a computerized reasoning-based application might be utilized to detect exercise positions and provide accuracy to assist people in improving their structure. The system would be necessary because only some clients approach or have access to a trainer.

II.Literature survey:

Physical fitness and its importance:

[1] The most publicized comparative study of American and European youth conducted utilizing standardized fitness schools. 42% of American youth met the minimal fitness of the European youth at 92%. Almost 50% of American children cannot be physically fit, and many are getting soft. The softness of every individual will lead to stripping and destroying the nation's liveliness, energy, and spirit. The hostile media criticized this practice and focused on keeping schoolchildren's fitness scores low. The motto of critics to prove that fitness testing felt embarrassing, uncomfortable, and not so meaningful. To improve the psychological importance of physical fitness, researchers suggested giving rewards and taking it as academic achievement.

Then the performance indicator and a reward tag can be the motivational aspect for students of all kinds of potential. The cognitive evaluation theory keeps the youth motivated and enthused about regular fitness practices.

As we know that the previous generation had more fitness than us. However, the current generation is still trying to improve their mental and physical fitness. Better nutritious food and daily fitness practices will lead to the gain of good fitness remarks. The recent revival of critics for and against the value of physical fitness and potential testing created vast public awareness of the emergency in the surrounding area.

[2] The critical concepts for the complete coverage of physical fitness the author had stated in this. The concept focuses on assessment and its relation with health-related fitness and varying health per age. Early debate in physical assessment questioned the utility of fitness tests. However, a clear view by public health experts and medical experts concerning the assessment's importance, as well as previously taken fitness tests in school physical education programs as part of academics. The author also focuses on the person's aerobics capacity, which may be independent of age. One of the results also stated that aerobic capacity is directly proportional to the amount of oxygen used and taken in by the body.

The strength of any individual can be governed by these ten key concepts of the fitness assessment. The importance of fitness assessment and its involvement in daily routine is considered for fitness programming and effective coordination.

Python and its use:

[3] The author describes the course of image processing, which should be subdivided into modules such as the basics of Python, image processing and its fundamentals, digital and current techniques of image processing, Image and video acquisition, detection of the main object of the model in motion, Image and pixel formation, borders, image segmentation, and the optic instruments. The representation of GUI(Graphic User Interface) and the PBL(Project Based Learning) improve students' involvement in the learning process. As it is directly the viewpoints of the problems, one could traverse through all the perceptions the questions like What? How? and For what?

The problem-solving techniques and strategies and specific skills development are possible through teaching in specified disciplines. The workshop planning in class and creating a healthy environment for the new technology is part of the teaching process. This will make one adapt quickly to the upcoming techniques in the world.

Image and Video Processing:

[4] The author used pre-trained models for 19 different types of objects on the "travel" data for the PASCAL Visual Object Classes Challenge 2009. The robust approach applied provides easy object detection for static images, further increasing the efficiency of video detection. The model trained by selecting one frame per second and running an object detector for each of them to collect the scores. The probability distribution is calculated using Platt scaling and map, then calibrated probabilities. The author applied four different approaches: windowing, Pos, which is a parser of tagging, Parsing, and Parsing II. The model works on rules and neural networks.

The data set is to be divided into many disjoint and training sets to get the valuable test for object recognition. The activity detection has limitations outside of the set recognizer. It trained the model according to the data sets and added the text according to only what we had in datasets. The model could not detect objects in the pictures that were not mentioned in the data sets.

[5] The author describes the ergonomic evaluation of the posture, which ultimately helps the workers to know their postures are not good enough. This will decrease the effect of back pains and ache on the workers and design suitable workstations for them. The Ergonomic methods are mainly used to test or assess risk at workstations. Many experts have different opinions on this posture evaluation method. The MOCAP(Motion Capture) equipment analyzes and differentiates between variable postures. The AI model can be trained by human data feed, and humans can assist in its predictions. The valuable resource to business is humans, and it can be kept with excellent efficiency if it is physically fit for the work.

The research aims towards the combined use of MOCAP and AI systems to investigate the advantage of better posture as well as the disadvantage of bending back and many other structures. The OWAS(Ovako Working Posture Assessment System) method for evaluation is subjective. So it builds an AI-based solution to help the observer and predict more consistently for classification.

[6] The paper focused on developing methods of measuring bodies placed in front of the optic detector. If the actual size and the captured size in the Image are given, it gets easier to calculate the distance between them. Nevertheless, in many cases, the actual size is not given. Thus, the paper gives a methodology to calculate distance from a given object whenever the real size is not mentioned. The camera used for demonstration should be fast to capture more accurate real-time data. The single and multi-shot options depended on the preciseness and accuracy of the solution one wanted.

The measurement between the camera and the object is demonstrated through the basic principles of the working of a camera. The author also used the basics of optics and the laws of reflection to obtain a calculated solution.

Mental health:

[7] Methamphetamine(MA) ties the dopamine, serotonin, and norepinephrine receptors as well as transporters to the neurons leading to better and faster accumulation of monoamine neurotransmitters in the gap between two cells called synapses. The MA induces euphoria by abusing the mind/brain perk pathway by releasing neurotransmitters. The prolonged consumption of such drugs severely damages the functioning of monoamine transmitters. Physical work and exercises reduce the high demand for hormones and drugs. The demands remain in considerable amounts still the aerobic exercises have the most considerable and beneficial effect on improvement. The sleepy and unenthusiastic lifestyle is the result of the overconsumption of MA. Muscle atrophy, BMI, and posture imbalance are the significant signs of MA consumption.

[8] The paper reviews mental health and the interrelated benefits of exercise. Many mental health practitioners recommend exercise as an effective way to address mental health problems. It is always recommended to exercise daily for individuals with mental disorders. It is hard to implement heavy exercise and practices daily, but it leads to a tremendous positive impact on mental health. Mental health always stays on a continuum of having no symptoms of severe issues or mental disorders. The author classified the disorder symptoms as emotional state, behavior, problem in thinking, and approach towards the situation in routine tasks. The lack of confidence, anxiety, and panic creates more diseases than actual adverse conditions for a body.

Antipsychotic medication and exercise are the critical components for psychotic disorders. For psychotic disorders, antipsychotic medication is a crucial component of treatment, but it always goes along with psychosocial intervention as well as practical support.

Motion Detection:

[9] One of the great options is to trap the video frames at the camera or optic device itself, especially if all necessary energy and storage are available for computation and recording in the endpoint environment. Hence, no usual visits to record the disc, which clips should be kept at the end. The author proposed a "camera trap," which is a self-triggering device due to some specific activity due to some specialized sensors. The unique step towards this technology needs a program controller at the capture end, but when the camera trap can make the suggestion and detect when to record to disk and when to not. The controllers range from microprocessors to high capacity and deep techno computers—the range of programming language and hardware headers for interface with an external camera lens and the sensors. The camera trap technology has also profited from environmental detector data that will be detected along with the compliances.

The author proposed an optical device that can make conversation with a voluntary detector that finds temperature, moisture, brilliance, height from the sea shore, GPS positioning, timezone, and, deduced with the help of this, exposure and height of the sun from the horizon level. The detection gets more manageable with the RPi boards. The central reliability of this system is the PiCamera library, which provides an interactive Python interface through the RPi modules.

[10] One should increase the thickness of the therapeutic videotape controller development. Consequently, styles and algos from different image-processing orders will be suitable to be developed on the basis of this approach. Also, underestimating primary focuses, but designing the therapeutic videotape controller planned for perpetration makes sense in the being realities of supposing only about CDSS controller that aids the croaker in decision- timber. The most important thing is that the methodology should aid in increasing the perceptivity and particularity of the croaker – the stoner of the system.

One of the obstacles in the ultramodern enhancement of the latest-image processing styles for therapeutic systems is the need for expert perspectives and limited participation of croakers in the process.

One of the obstacles in the recent development of the latest image processing styles for therapeutic systems is less expert opinion and not expected participation of croakers in the detection and processing. The issue is the development algorithms, along with that for doing practical work.

[11] The system places a detector on the ceiling and, when the person is at rest, focuses more on the head plate of a person rather than the main body. In the controller, if the mean height is lower than 0.5 m, the optic device will suppose the person is in rest to sleep. Therefore, the system can find the rest or moving state and start the heartbeat discovery before the stage. However, we further use the DB- checkup algorithm for clustering the point shadows, If the person is not found as moving.

In this system, pinpoints with at least five contiguous points of a distance of 28 cm will be considered a group. When the person is moving or resting, the waves reflected from the head are more impactful than those reflected by the remaining part.

Therefore, based on the clustering algorithms, when the person is standing or sitting, we can praise the head and calculate the mean height of the head to represent the height of the mortal.

[12] As a consistent result, covariance analyses mentioned that Intervention activities soften the stress and negative thoughts, ultimately affecting general and internal health. Similarly, the retrogression analysis handed base for supposition piece of exercising practices, the use is some functions or moving continuously. The relationship between stress, mental health, and physical fitness are directly interconnected.

Overall, the experimental research base for the supposition that physic-exercise keeps the goods of stress on general and internal health. Remarkably, the goods were set up for the actors of the stress relief training program. Therefore, unborn outcomes need to further investigate the stress-softening goods of different exercise types as well as the goods of volition stress-daily interventions to inspect which interventions are the most capable of relieving.

As a harmonious result, covariance analyses mentioned that The relation between stress internal health and physical fitness is directly connected. Overall, the experimental exploration base for the supposition that physical- exercise keeps the goods of stress on general and internal health.

[13] Initially, the idea of the gesture recognition system is to identify applied multivariate gaussian distribution. Its applications were less because it collects the 3D features of human body parts. Recognition is mainly divided into image acquisition, preprocessing, extraction and classification.

Hand Gestures should be located and tracked. It has two methods. Taking video input and analyzing each frame or segmented only image input. Video measurements are then combined to form a single linear recognized image. Kalman filtering uses a series of measurements observed over time. Initially, gesture recognition feeded complex because of the need for modern technology.

[14] Gesture recognition systems for disabled people can be built by using MATLAB. Sample images are being processed. The feature extraction method changes the input data into the set of features. It extracts necessary information and removes unwanted features. Finally, the system is left with redundant data. It aims for the natural interaction between humans and computers.

The proposed system for signal processing. It converts hand gestures to commands. The system will read it directly and not through human-to-human conversation.

[15] Visually impaired individuals benefit from the touch screen and hand gesture interactivity technology. The authors have created a braille drawing using motion sensors based on hand gestures coupled with digital devices to assist blind persons with navigation. The suggested device accepts input variables such as finger motion and calculates values of x and y coordinates, swiping speeds, and pixel rates related to user inputs. The given inputs were evaluated using an ANN, an artificial neural network (ANN). Also, the crow search algorithm (CSA) provides the predicted outcome for blind persons.

[16] Employing deep learning techniques to recognize hand motion movements In this study, the authors used the Haar feature and the AdaBoost algorithm to segment the hand gesture data. Here, background noises are removed from photos of human hand gestures using the CamShift algorithm. In order to identify realistic human needs, We need to apply CNN in real-time hand/body motion data for experimental purposes. Results showed that 98.3% accuracy is achieved by the suggested CNN method for hand/body gesture recognition.

[17] A more effective Grad-CAM (GCAM) model proposed to recognize hand movements for 3D micro-Doppler feature creation. The Grad-CNN and CAM model evaluates relevant features and regions of 3D gestures by removing unimportant characteristics from noise regions.

The authors utilized two layers (convolution) that integrate critical azimuth and elevation angle information from micro-Doppler datasets of multi-channel. The demonstration results indicate that the proposed categorization methodology has 96.61% precision. For improved performance, this approach could be expanded to take temporally significant hand motion aspects into account.

[18] The Raspberry Pi 3 Model B and Python have been selected due to the programming language's inexpensive hardware and free availability. In this laboratory course, students learn how to manage hardware and software, design, implement, and debug an embedded image processing system, use Python as an alternative to MATLAB, and analyze image signals. Python and Raspberry Pi are introduced at the beginning of the lab course. By comparing the results of pre-and post-lab examinations, instructors can determine whether or not students have a comprehensive understanding of embedded image and video processing algorithms.

[19] This study describes a method for counting pushups in real time using 2D video footage. Then, it examines important motion characteristics related to counting the pushups. 147,840 samples were gathered from 220 pushup videos, each shot from two distinct angles. Half of the videos were used to model the suggested method, and the other half to evaluate its efficacy. The research provided is wholly dependent on recognizing the precision of pushups. It examines various pushup regulations from different nations. Further study will be conducted to increase the accuracy of the deep learning approach that uses 3D human body analysis in military medical examinations.

[20] In this paper, It is suggested that a system for tracking one's sitting posture relies on machine learning algorithms to anticipate one's posture. The test results show that when other machine learning methods are compared, the precision of the Random Forest Approach (RFA) has a precision of (98.70%) with 30 trees and an accuracy of (99.19%). However, because of its complicated computation, the forecast time is higher at 67.7 milliseconds. The SVM has the quickest 0.64ms prediction time.

[21] During robotic grasping activities, different objects frequently appear in the Image in varied positions and orientations, making it difficult to provide a functional graphical solution for robotic learning. In this research, we present a robotic grasping technique based on a 3D detection network that minimizes camera orientation's impact on picture identification. Finally, robotic control is used to grab real-world things.

[22] Computer vision is one of the most promising technologies for acquiring information. The cornerstone for resolving identification based on computers creates issues in developing artificial intelligence (AI) systems that process images of multiple images to highlight crucial details. The computer-made system can efficiently process massive data without sacrificing quality. The computer vision system's RAM and microprocessor dynamically highlight structures and save intermediate results.

Interaction at a low level with this block is facilitated via an open distributed real-time operating system. The analog signal data from an advanced metering device is the source data for the computer vision system, and it accesses the analytical system via the video sensor. This work presents the algorithm for a computer vision system's operation. When constructing the application, we utilized the vision library (OpenCV) open-source computer of algorithms image processing.

[23] Determining a material's qualities is necessary to confirm its acceptability, but doing so can occasionally be time-consuming, expensive, and complicated. To solve this issue, free Python libraries such as pymatgen, matminer, and others are used with the Materials API to collect and operate datasets. A machine learning model may be constructed when combined with machine learning libraries such as Sklearn.

[24] This research presents a novel Python-based data processing framework for Human Activity Pose Tracking. It gives the capability to rapidly process raw video data for human pose tracking acquired in unrestricted contexts. In addition, PyHAPT enables interpolation to restore missing joint data and data visualization that provides insights into spatial-temporal skeletal information.

[25] The building of a Python-based, general-purpose data study tool for OpenFOAM is described. Our approach is centered on creating OpenFOAM utilizing Python data analysis library bindings. Using the NumPy C-API, OpenFOAM multiple precision data is converted to a NumPy array, allowing Python modules to do unrestricted data study analysis and changes on flow-field data. We show how the recommended wrapper can be applied to an in-situ online singular value decomposition built in Python and get through the PimpleFOAM (OpenFOAM solver). Lastly, we demonstrate the application of cutting-edge machine learning techniques within the Python ecosystem by deploying a deep neural network for compacting all the flow-field data using an auto-encoder.

[26] The author has analyzed the pushup forms using a video-based system; previously, sensors were used to count the pushups, and there were many limitations to the system, which were having a high cost, as well as the accuracy of the system needed to be improved. To overcome all these challenges, the author suggests using a visual-based real-time image-capturing system using the OpenPose software. The author uses a 2D human pose where important motion features are analyzed for correct and incorrect pushup posture. Front view and flank view, two input views are taken. The posture is analyzed by considering body parts like the head, shoulders, and legs.

After analyzing the system, the author found that the vision-based system had great accuracy in analyzing pushups, and its potential use could be in military tests. A deep learning approach used to improve the reliability of the system.

[27] This paper analyzes the user movements using the movement identification system where the physical movements are tracked using the help tracking the joints and angle between them. An algorithm is used to identify and compare the repetitions with the original one.

The system is able to identify the physical movements.

[28] The author addresses the pose estimation using various approaches using Open Pose using 2D visuals as inputs to analyze the skeletal pose of the person; the identification system with the help of a neural network model, where important points are captured using heat maps. Moreover, their interaction as well as isolating movements according to the surroundings. Unique pointing featuring techniques were used to identify the joints and train them in a neural network.

The model's accuracy is up to 35%, but with each refinement stage, it increases, and at stage five, the model's accuracy is 48%. The author provided a heavily optimized neural network-based solution that can be used in real life to identify human posture.

[29] The author proposes a 3D pose estimation system where markerless movement capturing is used and analyzes various movements in men and women, like walking, jumping, and ball throwing. The system used Open Pose. The poses are identified using multiple cameras connected in sync to track all the movements. The captured motion is compared, the differences are noted in the corresponding joint position and the errors are recorded.

The author has worked on human posture analysis where the neural network is trained with a complexity increase of 6.5 multiple.

[30] In this paper, the author studies the range of movement of the shoulders. Here the author uses five tests to study the range of motion. The first test estimated the range for passive movements like flexion, abduction, external rotation, and placing the hand behind the back. The second test measured the active movements for abduction, external rotation, and flexion. For the third test, a polaroid camera was used with still photography to label the important points on the subjects. The fourth test, about maximum overhead reach in a standing position which is measured against a metric. After analyzing all the movements, abduction, flexion, and external rotation, the error of the measurement was found between 11-21 degrees for visual prediction, 14 to 23 degrees with goniometry, and 13 to 22 degrees with still photography. The primary purpose of this study was to compare and analyze the shoulders' active and passive range of motion.

[31] The author of this paper analyzes the various grips in lat pull-down movements, and the effectiveness of the grips is analyzed. The lat pull-down exercise focuses on the back, biceps and forearms muscles, Strengthening these muscles will develop a solid pulling movement in the person. The research was carried out with various grips and wrist positions like wide grip and close grip, and the wrists are pronated and supinated.

The results were that the closed grip has more effect on the mid back muscles and the back muscles, and the wide grip has a more significant effect on the lat muscles. Furthermore, the pronated grip focuses more on the forearms than the supinated grips.

[32] This research paper's author talks about 3D modeling using the RGB D sensor's image processing in deep learning. The RGB d sensors recognize the human posture through visual input. The human posture in this paper is analyzed using a CNN-based method, and two approaches are demonstrated. A decision-in model is a convolutional neural network model with the help of RGB. The depth of images and the body posture are analyzed. In the third stage, the camera will collect the input in the RGB format and also analyze the depth of the images, and these images will be freely processed in a data set. After analyzing these images, a training method parameter will be set; by considering this parameter, a neural network format will be created to analyze the human body posture, and it will be recognized. This system marks the main human skeleton points like head, shoulder, abdomen, and hip joints needs and angles. The system analyzes moments of the human body on a real-time basis and classifies whether the man is standing in a building or sitting, or walking. Analyzing various posters with quite a great accuracy, it analyzes the standing position with 92.3%, bending position with 91.8%, sitting position with 97.6%, walking position with 89.9%, and crouching position with 93.7% accuracy.

[33] In this paper, the author uses sensor-based equipment to identify the physical movements using the PCBA analysis. This is a sensor-based computer that identifies the human movement, which assists them in analyzing whether the performed exercise is right or wrong. The system collects data from the sensors and processes it with the sample analysis with the real-time feedback six core with the skill level preferred

in the system and detailed report of the scale generated. This system generates output according to pressure applied on the sensor and renders and filters the input according to the parameter set. The output is rendered according to the specific parameter. These parameters help identify whether the person is putting more force on the left or right leg. This way, the authors state that we can track the exercises.

[34] In this paper, the author talks about the significance and benefits of health and how fitness is getting expertly expensive nowadays. Hence, the author suggests a system based on sensors to identify whether the forming exercises are a writer. It uses the Telescopic sensors and the EMC sensor module to perform some exercises to check whether the form of the exercise is right or wrong. Suggested three main architectures of the system: first, data collection, where all the data is collected and noted. The second is feature extraction with all the features and the home added and data being transformed, and the third is the learning model where Recognition of the neural networks and fuzzy logic is used to get the output. The author uses the T-bar exercise and the bicep curl exercise. The system detects the user's future and guides them with their scope model. It helps the user to avoid muscle injury as well as joint injury caused due to the wrong form of the exercise, and its performance accuracy is about 89%.

[35] This Literature Review is on deep structure learning. The researchers have been studying for ten years, and they cover a total of 93 research papers and study them. So we can say that this concept is unique, and it's not easy to study. So the neural network is based on deep learning, which is RNN, DNN, and CNN, so these methods divide in video processing. Human action identification, anomaly detection, and behavior analysis are the main areas of video processing study. Users of various networks, such as YouTube, Twitter, Facebook, etc., frequently choose video data as their preferred format. Currently, it is also the data type with the quickest growth rate. Every day, YouTube receives millions of new video uploads. The complexity and amount of video data make it difficult to interpret and evaluate. Deep learning algorithms are appropriate for managing massive amounts of video data because they can process and interpret millions of information gathered from dispersed sensors. The author reviewed all research papers and found a solution in this section that examines the traits, approaches, dangers, countermeasures, and deep learning algorithms stated in the chosen publications. Although there has been much advancement, it has yet to be done accurately. Future video datasets should likewise be widely accessible for free. This research paper has some drawbacks. We can say that if the user inputs a low-quality video, sometimes the machine focuses on the background more than the actual object in front of them, and some technical issues are also found in this research paper.

[36] Researchers included some object detection techniques called "You Only Look Once(YOLO)." This algorithm is based on their network. It relies on video based on their network. Also, researchers include some concepts here. They are saying that the Image is processed before they go to the final results, and due to this Image, we will get very sharp, and the background around that Image is not focussed as the main objects in that Image. The scientific community has extensively studied motion estimation from image sequences. Optical flow estimation uses time-varying picture intensity to approximate the motion field. It is preferable to get accurate findings in real-time while using methods that allow problem-specific customization.

So due to help of this research paper, the user will get to know about some concepts which are discussed above and which are based on image processing and video processing in the practical world where it is used; we see some applications of Images, as well as video processing like in speed gun which is fitted on the highway to detect the speed of cars, is the main application behind this paper, and we can track the live objects. In the above research paper, some technical errors may happen due to the motion concept that comes here, and due to these errors may happen in the final results.

[37] This paper modifies human activity as sentences created by a language composed of atomically small body positions. Only a series of silhouettes taken from various angles serve as a storage mechanism for the knowledge of body stance. Individual body components are not recognized, and there are no specific 3D postures or body models. In this research paper, language helps us think, create, and make decisions, which is part of deep learning. Hence, we have to train our model with the help of a mathematical model or any logic or a particular algorithm to do all these things. We have to consider the worst case of that particular scenario to perfectly train the model. After that, our model is ready to perform the operations, which is the target, like we have to detect the actions which humans perform and analyze them. Now our model is fully ready to detect any actions.

From the researcher's opinion, the best place to start a conversation about actions and how to recognize them is by defining what we mean by action. This model should identify all human activities, as well as some scenarios like when there is something happens in the room while detecting their actions; they need to know so they have good knowledge of their surrounding which thing is doing which task, and also to identify some interruptions like anybody can enter and exit in the room, so they need to detect that particular thing. These are some human interruptions or unpredictable things that can happen during the detection of actions. So the author says that they have a solution for these things, and it is called some "verbs." Except

for a particular object or whenever there is the movement of the body, the product or system only focuses on that part. It is called a Visual verb, which is important according to the author's point of view. The biggest drawback of this paper is that it does not recognize the individual body parts. Human actions are done using various body parts and their motions Eg-yoga includes Surya namaskar, which involves various body part motions like legs, hands, back, eyes, neck, etc.

[38] Python language is preferred in this research paper because it is open source, accessible, and easy to understand and learn. Additionally, it has a standard library with modules specifically for threading, networking, databases, etc., geared at programming. Other programs are also used in the lab, such as plotting and handling photographs. With many companies using Python as their primary programming language, Python provides students enrolled in technical education programs with a strong foundation for future employment. Users use Raspberry Pi to create hardware projects, automate their houses, manage Kubernetes clusters, take advantage of Edge computing, and even commercial applications. This article presents a newly developed experimental program, "Image and Video Signal Processing for Embedded Systems." Students must use the Python programming language to complete various tasks on the Raspberry Pi. It includes two more experiments in the future: background replacement and object detection. In this research paper, only one issue is that background replacement, and object detection needs to be optimized.

[39] One of the primary publications to mix image processing and acquisition, it gives readers a solid foundation in both areas. The book will increase readers' understanding of picture capture methods and related image processing, helping them conduct experiments more successfully and affordably and evaluate and quantify data more accurately. Python is employed in many real-world scenarios and has long been regarded as one of the easiest programming languages for non-programmers to learn.

1. Explains the way to collect pictures physically and analyze them analytically to comprehend the technology involved in the Image.
2. Provides illustrations, thorough derivations, and functional Python samples of the concepts.
3. Provides helpful advice on image capture and processing.
4. Provides several tasks to check the reader's knowledge of Python programming and image processing.

[40] In this research, the paper author is saying that there is some inner disturbance while we take the record of particular shots or videos. This is natural, and any disturbance can harm the video or Image. So according to the author they say that, first of all, identify the type of disturbance in that video, and we can search in

that video which type of disturbance is there, and this is called a high-level video event, which of two types like one is used while there is a lot of peoples and other is used for some knowledge-based activities. High-level event identification is the process of automatically detecting certain high-level events within a video stream. This can often be daunting, especially when shooting video in free environments. While the solutions in use today are different, we have identified the key elements they have in common and provided insight into all of them. High-level video event identification searches video clips to identify events of interest automatically. As described by the authors, high-level or complex events are long-term, physically, and temporally dynamic object interactions in a particular scene setting. Two general categories of complex events are social gatherings and educational activities. These events are generally complex, so it is a little difficult task.

In this research paper, the author has deep and clear knowledge of their topic, reviewing many research papers and coming to a solution for this concept of high-level video detection. There are many techniques involved in this, like various approaches are there which include algorithms, kernel, and visual concepts. There also they have to identify the audio and visual tasks here. Some fusion strategies are there to improve the algorithms. Also, many features are available in the market. They have mentioned that these techniques are involved in their previous evaluations and strategies called TRECVID, which are detected here.

[41] In this paper, the author has fundamental knowledge of deep learning concepts used in neural technology. From this, they are trying to find the solution for the computer vision concept for the detection of vehicles on the roads and the faces of people. Additionally, many possible study areas demand constant investigation and excavation. We continuously refine the model to extend the algorithm's speed and accuracy. Vehicle identification has recently become necessary as a tool for efficient traffic management. We suggested a deep neural network of convolution with a minimum of nine layers in this study. The foundation for deep learning A vehicle data set collected from various perspectives is used to evaluate the suggested approach using Caffe. The suggested model employs deep convolutional neural networks, which have a better performance when compared to classic machine learning-based vehicle recognition. It requires vehicle location and has subpar fault precision.

[42] Action recognition is one of the technologies that enable interactions between humans and computers, video surveillance, and video scene interpretation. An appropriate feature extraction approach is necessary for solving action recognition difficulties. To deal with these problems in action recognition tasks, Several local space-time visual representations have been suggested. Long Recurrent Convolutional Networks (LRCNs) and Deep Convolutional Neural Networks (DCNNs) have demonstrated considerable promise in various fields. This paper uses motion maps and combines a C3D network with an LRCN network to review the problem of human action recognition.

There are several applications for the present area of computer vision research known as human activity recognition. Recurrent neural networks (RNN) and deep convolutional networks (DCN) have recently attracted more interest in multimedia research and produced cutting-edge findings. During this study, We provide a fresh framework that expertly combines LSTM and 3D-CNN networks. First, a "motion map" is created by integrating the discrimination data from the video into a deep 3-D network of convolution

(C3D). A motion map and the frame after it in the training video may be merged to produce a new motion map by gradually prolonging the movie.

[43] In this research, the paper author is trying to identify the faces of people in low-condition videos, and these videos have low pixels or resolutions, which is unclear to us. Now they included various techniques, which are hard parts of deep learning approaches called neural technology used here. This is called a MERN network, and it is a multimodal technique that first takes data from the user and that selects the dataset for their model; after processing that data, they transfer to further operations and then lastly will get outcomes. Basically, they capture images for low light at different instances, and they work on that images while they collect different images; they work on them using deep learning techniques and slowly make the Image of that person's face and finally create the Image of a particular face how interesting is that! Various types of datasets are available, like PaSC, IJB-S, and YTF, which have different features. Most videos that have low quality in their capture are in the IJB-S method.

Face photos of inferior quality hurt face recognition ability. However, aggregating the data in video frames can provide more discriminative features for poor-quality video sequences. We propose the application of a MARN for face recognition in low-resolution movies captured in the real world (MARN). In contrast to other recurrent networks, MARN learns to aggregate pre-trained embeddings, making it resistant to overfitting (RNNs). In contrast to quality-aware aggregation techniques. Utilizing the video context, MARN learns many attention vectors in an adaptable manner. Low-satisfactory video dataset is superbly stepped forward via MARN; in step, there are findings on video which includes faces as discussed above and have three different types. At the same time, it produces results that are equivalent to high-quality video datasets.

[44] The author's deep knowledge of their area of expertise means in the "Yoga" field. There are many types of yoga included in this paper, and the author comes to the conclusion that the ideal posture for that type of yoga is carried out by the people, and it detects that the person is doing the right things or not finding the best postures for people. The author also says that yoga improves our physical health and mental imagination power, which broadly impacts our life. Yoga is a necessity of today's world. Everybody, children, adults, and aged people, do yoga regularly to maintain their lifestyle. Yoga includes various body parts of the body like hands, legs, toes, head, shoulders, etc. it also includes dance and side plank. Yoga posture recognition software may recognize the practitioner's present stance and then get training resources from the Internet to remind the user of the current pose. The system also uses a Sensor computer to gather the user's body map and recover the body form.

[45] The dataset for this study comprises exercises, dips, and pull-ups and is based on data from three UWB sensors and supplementary inertial data. A thorough performance examination of the CNN conducted, and the NB and DT's recognition accuracy up to 89.4 and 92.9, respectively. The accuracy was found to be more than 95% for ENN and at a level of 94.81 for CNN. This study proposed an innovative way to dimension and includes various gym exercises that involve legs, pushups, etc. This approach was followed by a novel approach to conditioning identification using conventional classifiers, CNNs, and ultimately ENN. Based on the outcomes of the trials, the pre-trained

original CNN and forthcoming ENN might be validated. The primary benefit of the suggested strategy is the efficiency attained by our solution, which is the ENN conforming of 9 networks in a computationally challenging setting.

[46] The web and mobile platform Feast In is to satisfy consumer demand for home cooking. They might restrict their search for recipes with a better search algorithm. Furthermore, they can post a photo of food users have discovered, and they return the result as a list. Researchers worked on Vs. Code, MongoDB atlas, GitHub, etc. tools. The study aims to create an image recognition model that will work with the platform. Machine learning will be used to develop an image recognition model using training photos. Since it is a new platform currently being implemented, it might receive less traffic. The goal of Feast In is to serve as a global platform for users. Users will be able to recognize a recipe from an image with the help of an image recognition model. The custom feature developed as a result of this study will enable users to modify recipes as they see fit.

[47] Enhancing a picture or video's aesthetic value is known as an aesthetic enhancement. Digital photo and video editing have both benefited from computational aesthetics optimization. Using the same features, designing image processing methods can enhance the Quality of Experience (QoE). We provide a framework for interactive enhancement that will increase the visual appeal of photos and movies. The suggested method is used in a domain that has undergone wavelet transformation.

Different enhancement techniques can be developed to get the desired outcome by altering how wavelet coefficients are decreased or amplified. RGB plus Depth (RGB-D) films can raise their perceived quality by using an interactive video editing application and an algorithmic improvement technique. The results are that the framework for improving specific video materials has been implemented. It is based on a Laguerre-Gauss wavelet domain multi-resolution representation of edges. An automated RGB-D video enhancement technique, as well as an interactive object-based video editing system, have been tested.

III. Study of Algorithms:

When the user launches the system, it takes input from the camera device. It analyzes the content in the frame and detects distinct body parts, and marks them using the Mediapipe library. Pose estimation is a computer vision technique to track the movements of a person or an object. Pose estimation is actively used in augmented reality, animation, gaming, and robotics. There are several models present today to perform pose estimation.

Blaze pose offers several applications, including fitness and yoga trackers. These applications can be implemented by using an additional classifier like the one we will build in this article.

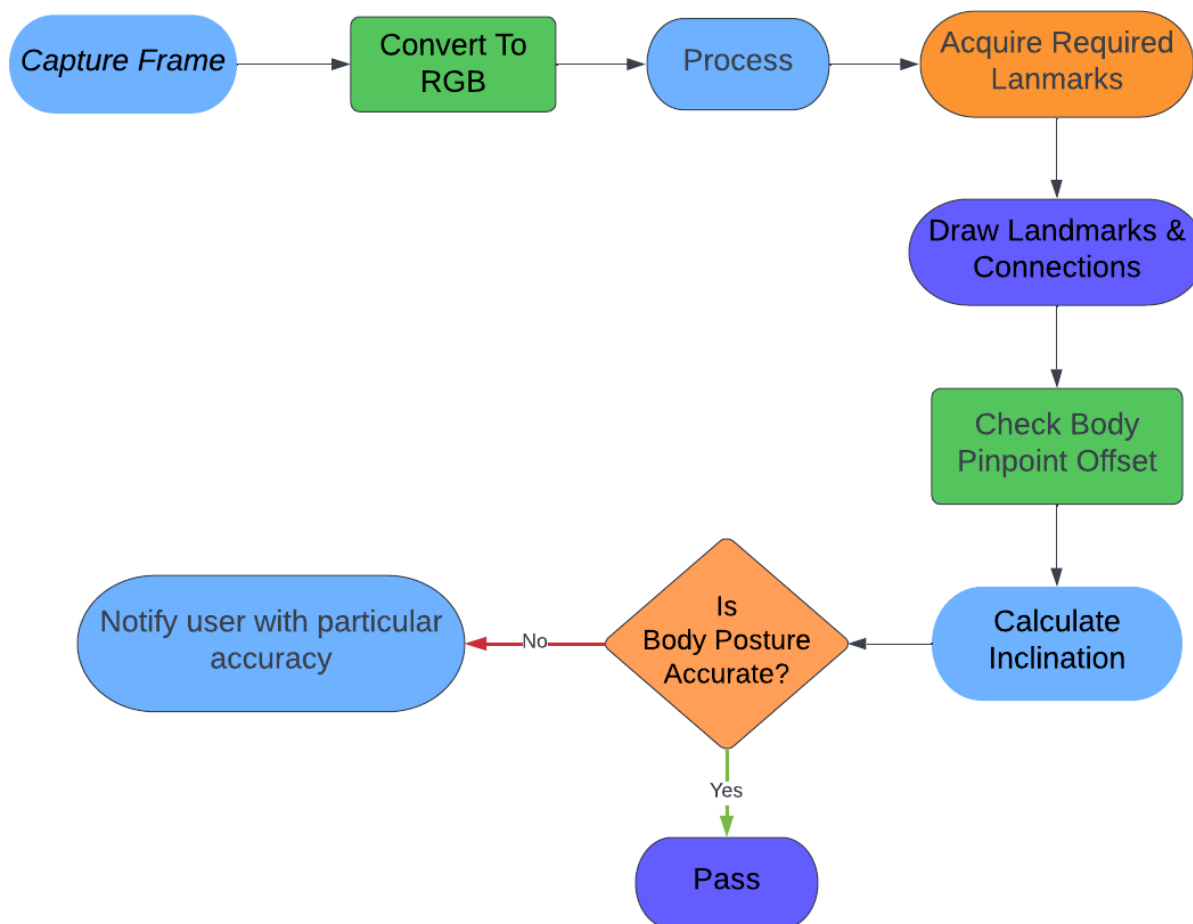


Fig 1. THE ABOVE LOGIC IS ABOUT TRACKING POSTURES

Results and conclusion of Existing system:

The important thing which the review came to know is this; the approach should help to increase particularity in exercises by the system. The problem is the algorithm improvisation in addition to doing the needful work. The author[11] placed a detector on the top ceiling; thus, when a person is not doing a task or the person's body is in a rest position, that captures the head part data rather than the complete body so that the system can find the outcome by the AI and it is also. The MOCAP(Motion Capture) [5] equipment analyzes and differentiates between variable postures. The AI model can be trained by human data feed, and humans can assist in its predictions. When the person stands or sits, it reflects a more robust signal for detecting one's body posture.[11] When the person's body is at rest, the signals reflected by the person are more impactful than those of a moving one.

Stress, anxiety, and depression[7] tend to negatively impact the body's fitness and hormone levels [12]. The improvement of object detection for better suggestions also captures the actual distance detection. Thus the appropriate prediction of body posture through clustering algorithms during exercise. The essential

data will be divided into numerous disjoint and uncommon data sets to get a useful test for object or body posture recognition. Many of the data sets were provided in a limited manner, so the model needed to be trained.

IV. Methodology:

To estimate the human posture in 2-d images by means of OpenCV[11] and MediaPipe.

The System architecture consists of 5 stages: Executing entered commands(By Jupyter Notebook), Making the optic devices accessible, Grabbing the input from that webcam, and analysis of posture to get exact pinpoints to correlate with the already existing data sets. After this, the live video images are converted into image frames[4]. Then the data set will compare to inbuilt poses. However, the results will get displayed in percentage. The difference will be compared for computing accuracy in a specific exercise. The algorithm used for the analysis and result calculation is done in the neural network functions and methods for each pixel[9].

A win discovery model for operating on the entire images, which returns hand-oriented box bondings. The hand corner model works mainly on the win sensor's cropped images portion, which gives the 3d high dedication coordinates[11]. Mediapipe gives 3D milestones from just a single frame. OpenCV consists of a comprehensive set of classics as well as state-of-the-art computer vision and machine literacy algos.[3] . These algorithms can be used to describe the wrong body posture, identify body bendings, and classify mortals in live video. It sews the high-resolution images of the entire scene, finding similar images from previously entered in the existing database. Making detection from feeds like hand and body position. The system uses the CV3 to get the device optic media access and the mediapipe for sketching the pinpoints on the body, hand, and legs. Apply the styling notations like thickness, circle radius, and color. After successfully recognizing all body posture poses, we will destroy the camera window of the device by command `cv2.destroyAllWindows()`.

V. Algorithmic Survey:

The steps of this algorithm are explained below :

Step 1: The algorithm inputs the following optic device image [3].

Step 2: The Image is passed through layers of different processing libraries like mediapipe or openpose.

Step 3: Then, a min pooling layer with stride two is applied to point out the real-time endpoints of the moving object[6].

Step 4: Make detection from the feed on the criteria Min_detection and Min_tracking.

Step 5: 2 fully connected layers are added.

Step 6: The final output layer is a softmax layer that puts up the image pattern. Apply basic styling and coloring to the detected points so that they can be better visible to human eyes categories.

VI. Comparative study:

- Theory :

Table 1. Method and application comparison

Previous Methodology	Proposed Methodology
Implemented Using Raspberry Pi and Python language.	Implemented Using Python, python modules and Jupyter notebook.
Hardware requirements are more.	Suitable for low end devices.
Requires less time to train.	Requires more time to train.
Implemented by using deep learning.	Implemented by using machine learning.
Large data sets are required for training.	It is possible to deal with less data.
Object detection is not there.	Object detection is done by machine learning.

Openpose used by researchers for old research. Our system uses Mediapipe. Here is the difference between Mediapipe & Oppose.

Table 2. Library Comparison

MediaPipe	OpenPose
Holistic mode included	Realtime 2D multi-person keypoint detections mode
33 body/foot landmarks	15,18,27 body/foot landmarks
468 facial landmarks	70 facial landmarks
Customizable ML solution	Non-customizable

VII. Proposed model:

1. Introduction:

In a developing country like India, where people have busy schedules, mental and physical health is ignored. Mental[7] and physical health are highly interrelated.

AI-based Smart systems play a vital role in providing efficient solutions at cheap costs. There is less human intervention[5] regarding smart systems, which leads to hassle-free life.

Good health Practices can be done by improving the level of recommendation and suggesting a better way of living with peaceful mental health[8].

Now, most things are controlled by AI, So smart implementation in healthcare is much needed to serve the people who can not afford the high training costs[1].

2. *Gratified Content:*

In this paper, we have analyzed pull-ups form using a video-based system. In the past sensor-based systems were used to analyze the proper form of the exercises[3]. However, there are too many limitations to the sensor-based systems, such as high cost, not providing accurate data as well as delays in significant changes. To overcome these challenges, a vision-based system will be used for counting the pull-ups as well as analyzing the form of the pull-ups. A real-time imagery system will use body point capture and analysis of key points marked in the joints to capture an essential range of motion.[4] Here human pose will be analyzed using real-time imagery for the correct pull-up form. A rear view of the user will be taken for the analysis. The posture and counting of the pull-ups will be carried out by analyzing the critical movements of the vital body parts like the head, shoulders, back, arms, and legs.

Human posture estimate from a video is beneficial for various applications, including measuring physical activities, identifying sign languages, and controlling full-body gestures. As an illustration, it may serve as the foundation for applications such as yoga, dancing, and fitness. In augmented reality, it can also enable the superimposition of digital material and information on top of the natural environment.

Using BlazePose research, which also drives the ML Kit Posture Detection API, MediaPipe Pose is a machine-learning solution for high-fidelity body pose tracking. It does this by inferring 33 3D landmarks and a background segmentation mask on the entire body from RGB video frames. Our technique delivers real-time performance on most recent mobile phones, desktops/laptops, in Python, and even on the web. In contrast, state-of-the-art systems rely primarily on intense desktop environments for inference.

3. *Algorithm:*

[48] Make the system ready to take input from media devices.

1. Check the availability of optic[3] sensing devices and take access from the system for capturing.
2. VideoCapture(X)
 $X \rightarrow$ Vice code for your input-taking device
 If (WebCam is open)
 Read the input
 Return the Image to the screen (frame)
 Release the WebCam after taking input
3. Make detection from the feed on the criteria Min_detection and Min_tracking

4. Apply basic styling and coloring to the detected points to be better visible to human eyes.
5. Destroy the window once the user finishes it.

Let's understand it with an example:

If one is performing the pull-ups, initially, the user's dead hang position will be captured [29], and it will be analyzed where all the points of the back will be analysed whether they are in a straight line. Then the user's hand position will be observed, and the grip will be identified[31]; after that, when the user performs the exercise, the system will monitor it. It will analyze it with some parameters[32], like the angle between the forearms and biceps; let it be "a" will be observed, and it must be less than 45 degrees and greater than 10 degrees. The shoulder and wrist points should be in a straight line[30], which can be analyzed using the equation of a straight line where the slope is "s."

If $10 < a < 40$ & $s=0$

"It is a perfect pull-up."

Else

"Need improvements"

Hardware needed:

- The high resolution and high-speed camera are essential to capture live imaging and faster processing for optic flow[6].
- The processing system should be capable enough to work with Python [3] modules like OpenCV, Tensorflow, MATplotlib, Mediapipe, and other libraries.
- Better wifi or internet connectivity if cloud storage is connected.

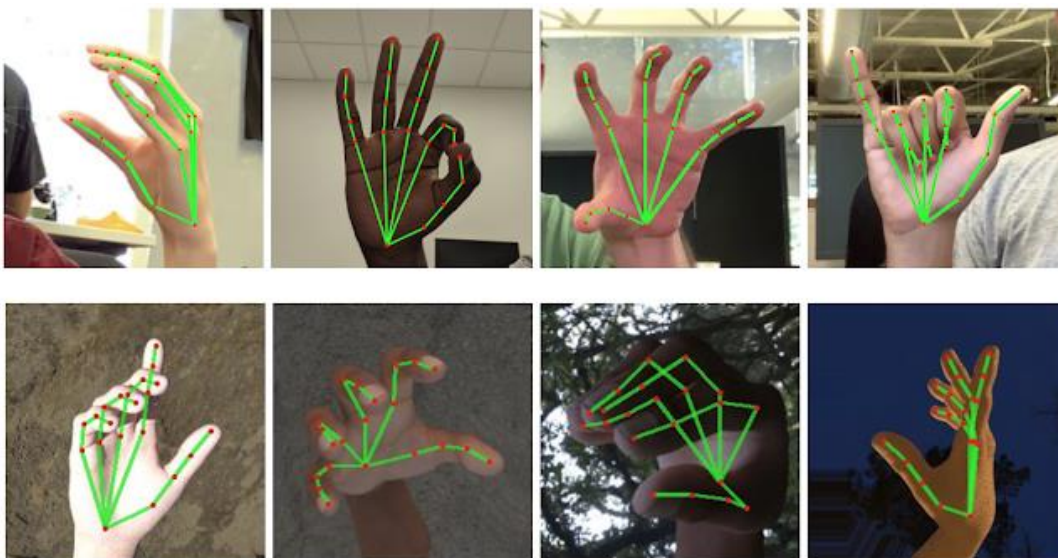


Fig 2. DIFFERENT HAND POSITIONS DETECTED

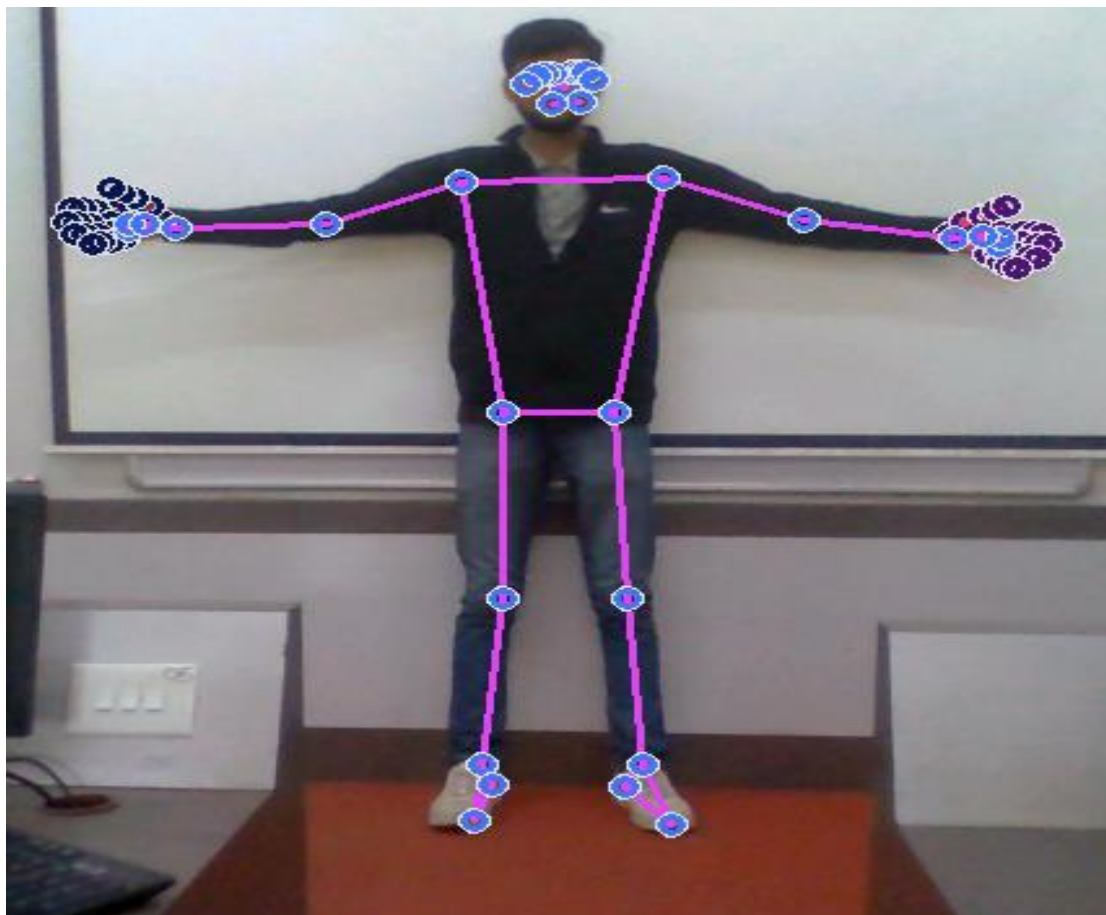
4. *Proposed architecture/ system diagram:*

MediaPipe supports OpenGL ES up to version 3.2 on Android/Linux and up to ES 3.0 on iOS. In addition, MediaPipe also supports Metal on iOS. OpenGL ES 3.1 or greater is required (on Android/Linux systems) for running machine learning inference calculators and graphs.

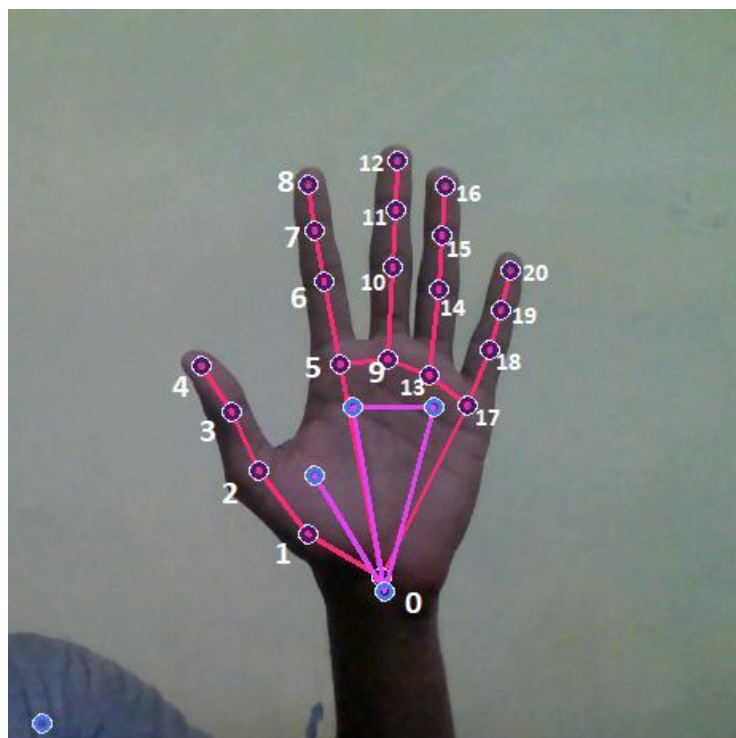
- Install MSYS2 and edit the Path variables of the environment
- Install necessary packages
- Install Python & allow the executable to edit the PATH environment variable
- Install Visual Studio Code or Jupyter Notebook
- Visit GitHub and refer Mediapipe repository
- Configure OpenCV for getting video input

1. Left_shoulder
2. Right_shoulder
3. left elbow
4. right elbow
5. Left hip
6. Right hip
7. Left knee
8. Right knee
9. Left heel
10. Right heel
11. left ankle
12. right ankle
13. Left_foot_index
14. right foot_index
15. Left palm
16. Right palm
17. Face
18. Others

Fig 2. DIFFERENT BODY PIN-POINTS DETECTED



- In the beginning, the stoner, who is now perched on the hanging bar, will dissect the straightness of the reversal. For the stoner to gain credit for a pull-up, he must bring his head over the wrist line on the bar. The next big thing in technology is going to be artificial intelligence. The purpose of this paper is to gain an understanding of how AI-enabled healthcare systems operate in conjunction with artificial intelligence and machine literacy algorithms and how these systems have a significant impact on mortal life in terms of the level of comfort they provide, the amount of money they save, and the amount of environmental goodwill they do.
- AI-based smart systems have a wide range of applications across various industries. The use of image and videotape processing in exercise research represents a more holistic approach to improving human existence. Neural networks have been brought to greater prominence in recent years thanks to advances in machine literacy and artificial intelligence. Deep neural networks are neural networks that have several layers and are hence so named because of their construction. There are several different models of deep literacy; each specializes in working on a different job.
- Recognizing people's faces is rapidly becoming an aspect of in-depth knowledge.



- 0.Wrist
- 1.THUMB-CMC
- 2. THUMB_MCP
- 3.THUMB_IP
- 4.THUMB_TIP
- 5.INDEX_FINGER_MCP
- 6.INDEX_FINGER_PIP
- 7.INDEX_FINGER_DIP
- 8. INDEX_FINGER_TIP
- 9. MIDDLE_FINGER_MCP
- 10. MIDDLE_FINGER_PIP
- 11. MIDDLE_FINGER_DIP
- 12. MIDDLE_FINGER_TIP
- 13. RING_FINGER_MCP
- 14. RING_FINGER_PIP
- 15. RING_FINGER_DIP
- 16. RING_FINGER_TIP
- 17. PINKY_MCP
- 18. PINKY_PIP
- 19. PINKY_DIP
- 20. PINKY_TIP

Fig 4. Identified Palm Points.

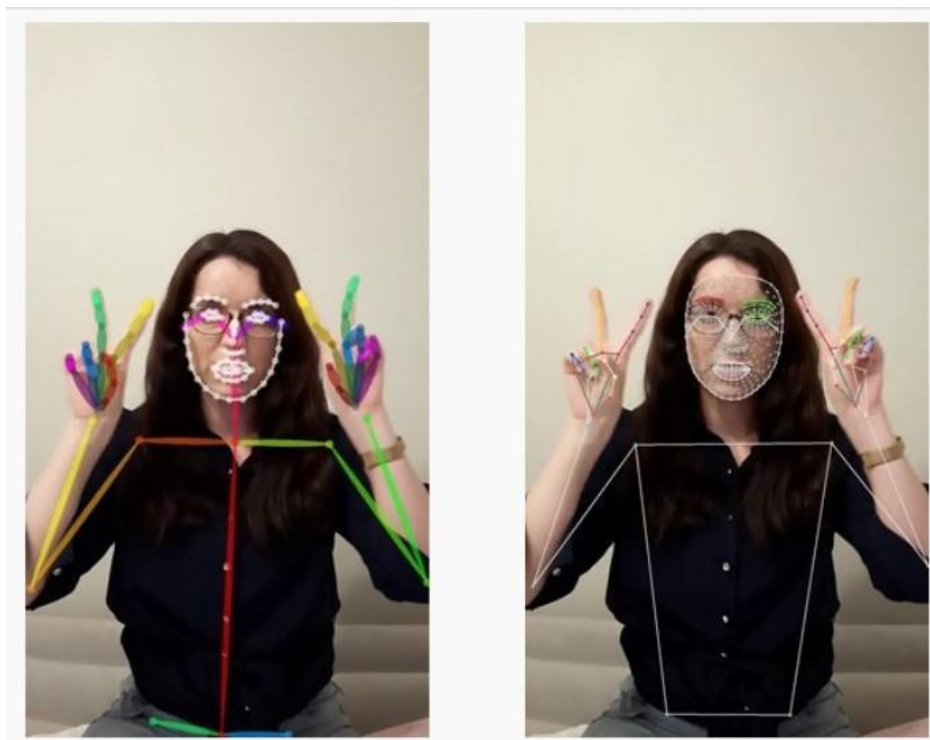


Fig 5. Differ between Open pose and Mediapipe

4. Conclusion:

In this paper, we have studied how AI-based Smart systems work. We also learned about many exciting fields, like psychology and its direct relation to daily physical exercises.

The model will work on suggesting a better way of exercise, and proper posture avoids harm to the body. The use of python libraries like OpenCV, Tensorflow, Matplotlib, and Mediapipe. We also come to the outcome of many people wanting a personal trainer who cannot afford it due to the high fees, so this will be a revolutionary system in the healthcare field.

References:

- [1] Lenny D. Wiersma & Clay P. Sherman (2008) The Responsible Use of Youth Fitness Testing to Enhance Student Motivation, Enjoyment, and Performance, Measurement in Physical Education and Exercise Science, 12:3, 167-183, DOI: 10.1080/10913670802216148
- [2] Charles B. Corbin, Gregory J. Welk, Cheryl Richardson, Catherine Vowell, Dolly Lambdin & Scott Wikgren (2014) Youth Physical Fitness: Ten Key Concepts, Journal of Physical Education, Recreation & Dance, 85:2, 24-31, DOI: 10.1080/07303084.2014.866827

- [3] A. F. Jiménez López, M. C. Prieto Pelayo and Á. Ramírez Forero, "Teaching Image Processing in Engineering Using Python," in IEEE Revista Iberoamericana de Tecnologías del Aprendizaje, vol. 11, no. 3, pp. 129-136, Aug. 2016, doi: 10.1109/RITA.2016.2589479.
- [4] Motwani, T.S., & Mooney, R.J. (2012). Improving Video Activity Recognition using Object Recognition and Text Mining. *ECAI*.
- [5] Igelmo, Victor & Syberfeldt, Anna & Högberg, Dan & Rivera, Francisco & Perez Luque, Estela. (2020). Aiding Observational Ergonomic Evaluation Methods Using MOCAP Systems Supported by AI-Based Posture Recognition. 10.3233/ATDE200050.
- [6] Real Distance Measurement Using Object Detection of Artificial Intelligence Jae Moon Lee*, Kitae Hwang, In Hwan Jung, Received: 11 November 2020; Accepted: 27 December 2020; Published online: 05 April 2021.
- [7] Huang J, Zheng Y, Gao D, Hu M, Yuan T. Effects of Exercise on Depression, Anxiety, Cognitive Control, Craving, Physical Fitness and Quality of Life in Methamphetamine-Dependent Patients. *Front Psychiatry*. 2020 Jan 28;10:999. doi: 10.3389/fpsyt.2019.00999. PMID: 32047445; PMCID: PMC6997340.
- [8] Morgan, Amy J., et al. "Exercise and mental health: an exercise and sports science Australia commissioned review." *Journal of Exercise Physiology Online*, vol. 16, no. 4, Aug. 2013, pp. 64+. *Gale OneFile: Health and Medicine*, link.gale.com/apps/doc/A361184771/HRCA?u=anon~33e6715&sid=googleScholar&xid=1bb87a40. Accessed 21 Nov. 2022.
- [9] Miklas Riechmann, Ross Gardiner, Kai Waddington, Ryan Rueger, Frederic Fol Leymarie, Stefan Rueger, Motion vectors and deep neural networks for video camera traps, *Ecological Informatics*, Volume 69, 2022, 101657, ISSN 1574-9541,
- [10] Nataliia Obukhova, Alexandr Motyko, Alexandr Pozdeev, Personalized approach to developing image processing and analysis methods for medical video systems., *Procedia Computer Science*, Volume 176, 2020, Pages 2030-2039, ISSN 1877-0509,
- [11] Jiacheng Wu, Naim Dahnoun, A health monitoring system with posture estimation and heart rate detection based on millimeter-wave radar, *Microprocessors and Microsystems*, Volume 94, 2022, 104670, ISSN 0141-9331,
- [12] Sandra Klaperski, Reinhard Fuchs, Investigation of the stress-buffering effect of physical exercise and fitness on mental and physical health outcomes in insufficiently active men: A randomized controlled trial, *Mental Health and Physical Activity*, Volume 21, 2021, 100408, ISSN 1755-2966,

- [13] Dhanashri Patil, Prof. S. M. Kulkarni (2016). Hand Gesture Recognition Using Neural Network. Vol no.4 Issue no.06, http://ijates.com/images/short_pdf/1465967240_297ijates.pdf
- [14] G. R. S. Murthy and R. S. Jadon, "Hand gesture recognition using neural networks," 2010 IEEE 2nd International Advance Computing Conference (IACC), 2010, pp. 134-138, doi: 10.1109/IADCC.2010.5423024.
- [15] S. M. Aslam and S. Samreen, "Gesture Recognition Algorithm for Visually Blind Touch Interaction Optimization Using Crow Search Method," in IEEE Access, vol. 8, pp. 127560-127568, 2020, doi: 10.1109/ACCESS.2020.3006443.
- [16] Sun, Jing-Hao & Ji, Ting-Ting & Zhang, Shu-Bin & Yang, Jia-Kui & Ji, Guang-Rong. (2018). Research on the Hand Gesture Recognition Based on Deep Learning. 1-4. 10.1109/ISAPE.2018.8634348.
- [17] Du C, Zhang L, Sun X, Wang J, Sheng J. Enhanced Multi-Channel Feature Synthesis for Hand Gesture Recognition Based on CNN With a Channel and Spatial Attention Mechanism. IEEE Access 2020;8:144610-20. [DOI: 10.1109/access.2020.3010063] [Cited by in Crossref: 7] [Cited by in F6Publishing: 7] [Article Influence: 3.5] [Reference Citation Analysis]
- [18] Karina Jaskolka, Jürgen Seiler, Frank Beyer, André Kaup, A Python-based laboratory course for image and video signal processing on embedded systems, Heliyon, Volume 5, Issue 10, 2019, E02560, ISSN 2405-8440, <https://doi.org/10.1016/j.heliyon.2019.e02560>. (<https://www.sciencedirect.com/science/article/pii/S2405844019362206>)
- [19] H. -J. Park, J. -W. Baek and J. -H. Kim, "Imagery based Parametric Classification of Correct and Incorrect Motion for Push-up Counter Using OpenPose," 2020 IEEE 16th International Conference on Automation Science and Engineering (CASE), 2020, pp. 1389-1394, doi: 10.1109/CASE48305.2020.9216833.
- [20] Ferdews Tlili, Rim Haddad, Ridha Bouallegue, Raed Shubair, Machine Learning Algorithms Application For The Proposed Sitting Posture Monitoring System, Procedia Computer Science, Volume 203, 2022, Pages 239-246, ISSN 1877 0509, <https://doi.org/10.1016/j.procs.2022.07.031>. (<https://www.sciencedirect.com/science/article/pii/S1877050922006366>)
- [21] Junyan Ge, Jinlong Shi, Zhiqiang Zhou, Zhi Wang, Qiang Qian, A grasping posture estimation method based on 3D detection network, Computers and Electrical Engineering, Volume 100, 2022, 107896, ISSN 0045-7906, <https://doi.org/10.1016/j.compeleceng.2022.107896>. (<https://www.sciencedirect.com/science/article/pii/S0045790622001823>)

[22] Stepan Sivkov, Leonid Novikov, Galina Romanova, Anastasia Romanova, Denis Vaganov, Marat Valitov, Sergey Vasiliev, The algorithm development for operation of a computer vision system via the OpenCV library, *Procedia Computer Science*, Volume 169, 2020, Pages 662-667, ISSN-1877-0509-
<https://doi.org/10.1016/j.procs.2020.02.193>
(<https://www.sciencedirect.com/science/article/pii/S1877050920303161>)

[23] Ward, Logan & Dunn, Alexander & Faghaninia, Alireza & Zimmermann, Nils & Bajaj, Saurabh & Wang, Qi & Montoya, Joseph & Chen, Jiming & Bystrom, Kyle & Dylla, Maxwell & Chard, Kyle & Asta, Mark & Persson, Kristin & Snyder, G. & Foster, Ian & Jain, Anubhav. (2018). Matminer: An open source toolkit for materials data mining. *Computational Materials Science*. 152. 60-69. 10.1016/j.commatsci.2018.05.018.

[24] Hao Quan, Andrea Bonarini, PyHAPT: A Python-based Human Activity Pose Tracking data processing framework, *Software Impacts*, Volume 13, 2022, 100305, ISSN 2665-9638, <https://doi.org/10.1016/j.simpa.2022.100305>(<https://www.sciencedirect.com/science/article/pii/S2665963822000446>)

[25] Romit Maulik, Dimitrios K. Fytanidis, Bethany Lusch, Venkatram Vishwanath, Saumil Patel, PythonFOAM: In-situ data analyses with OpenFOAM and Python, *Journal of Computational Science*, Volume 62, 2022, 101750, ISSN 1877-7503, <https://doi.org/10.1016/j.jocs.2022.101750>(<https://www.sciencedirect.com/science/article/pii/S1877750322001387>)

[26] Park, Ho-Jun & Baek, Jang-Woon & Kim, Jong-Hwan. (2020). Imagery based Parametric Classification of Correct and Incorrect Motion for Push-up Counter Using OpenPose. 1389-1394. 10.1109/CASE48305.2020.9216833.

[27] Madanayake, P. & Wickramasinghe, W. & Liyanarachchi, H. & Herath, H. & Karunasena, Anuradha & Perera, Tharindu. (2016). Fitness Mate: Intelligent workout assistant using motion detection. 1-5. 10.1109/ICIAFS.2016.7946559.

[28] Osokin, Daniil. (2018). Real-time 2D Multi-Person Pose Estimation on CPU: Lightweight OpenPose.

[29] Nakano, Nobuyasu & Sakura, Tetsuro & Ueda, Kazuhiro & Omura, Leon & Arata, Kimura & Iino, Yoichi & Fukushima, Senshi & Yoshioka, Shinsuke. (2020). Evaluation of 3D Markerless Motion Capture Accuracy Using OpenPose With Multiple Video Cameras. *Frontiers in Sports and Active Living*. 2. 10.3389/fspor.2020.00050.

- [30] Hayes, Kimberley & Walton, Judie & Szomor, Zoltan & Murrell, George. (2001). Reliability of five method for assessing shoulder range of motion. The Australian journal of physiotherapy. 47. 289-94. 10.1016/S0004-9514(14)60274-9.
- [31] Lusk SJ, Hale BD, Russell DM. Grip width and forearm orientation effects on muscle activity during the lat pull-down. J Strength Cond Res. 2010 Jul;24(7):1895-900. doi: 10.1519/JSC.0b013e3181ddb0ab. PMID: 20543740.
- [32] Elforaici, Mohamed El Amine & Chaaraoui, Ismail & Bouachir, Wassim & Ouakrim, Youssef & Mezghani, Neila. (2018). Posture Recognition Using an RGB-D Camera: Exploring 3D Body Modeling and Deep Learning Approaches. 69-72. 10.1109/LSC.2018.8572079.
- [33] Moller, Andreas et al. "GymSkill: A Personal Trainer for Physical Exercises." 2012 IEEE International Conference on Pervasive Computing and Communications (2012): n. pag. Web.
- [34] Hannan, Abdul & Shafiq, Muhammad & Hussain, Faisal & Pires, Ivan. (2021). A Portable Smart Fitness Suite for Real-Time Exercise Monitoring and Posture Correction. Sensors. 21. 10.3390/s21196692.
- [35] V. Sharma, M. Gupta, A. Kumar and D. Mishra, "Video Processing Using Deep Learning Techniques: A Systematic Literature Review," in IEEE Access, vol. 9, pp. 139489-139507, 2021, doi: 10.1109/ACCESS.2021.3118541.
- [36] Botella, G., García, C. Real-time motion estimation for image and video processing applications. *J Real-Time Image Proc* 11, 625–631 (2016). <https://doi.org/10.1007/s11554-014-0478-y>
- [37] Abhijit S. Ogale, Alap Karapurkar, and Yiannis Aloimonos Computer Vision Laboratory, Dept. of Computer Science University of Maryland, College Park, MD 20742 USA {ogale,karapurk,yiannis}@cs.umd.edu
- [38] Jaskolka, Karina & Seiler, Jürgen & Beyer, Frank & Kaup, André. (2019). A Python-based laboratory course for image and video signal processing on embedded systems. Heliyon. 5. e02560. 10.1016/j.heliyon.2019.e02560.
- [39] Chityala, R., & Pudipeddi, S. (2020). Image Processing and Acquisition using Python (2nd ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9780429243370>
- [40] Jiang, YG., Bhattacharya, S., Chang, SF. *et al.* High-level event recognition in unconstrained videos. *Int J Multimed Info Retr* 2, 73–101 (2013). <https://doi.org/10.1007/s13735-012-0024-2>
- [41] Luo, X., Shen, R., Hu, J., Deng, J., Hu, L., & Guan, Q. (2017). A Deep Convolution Neural Network

Model for Vehicle Recognition and Face Recognition. In *Procedia Computer Science* (Vol. 107, pp. 715–720). Elsevier B.V. <https://doi.org/10.1016/j.procs.2017.03.153>

[42] S. Arif, J. Wang, T. Ul Hassan, and Z. Fei, "3D-CNN-based fused feature maps with LSTM applied to action recognition", *Future Internet*, vol. 11, no. 2, pp. 42, Feb. 2019. 3D-CNN-based Fused function Maps with LSTM applied to action Recognition

[43] S. Gong, Y. Shi, and A. Jain, "Low-quality video face recognition: Multi-mode aggregation recurrent network (MARN)", *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop (ICCVW)*, pp. 1027-1035, Oct. 2019. Face Recognition in Low-Quality Video: Recurrent Multimode Aggregation Network (MARN)

[44] An approach to sport activities recognition based on an inertial sensor and deep learning Grzegorz Pajak a,* , Pascal Krutz b,* , Justyna Patalas-Maliszewska a , Matthias Rehm b , Iwona Pajak a , Martin Dix b. An approach to sport activities recognition based on an inertial sensor and deep learning

[45] An approach to sport activities recognition based on an inertial sensor and deep learning Grzegorz Pajak a,* , Pascal Krutz b,* , Justyna Patalas-Maliszewska a , Matthias Rehm b , Iwona Pajak a , Martin Dix b. An approach to sport activities recognition based on an inertial sensor and deep learning

[46] Lee Ann, Evelyn Toh, et al. "Feast in: A Machine Learning Image Recognition Model of Recipe and Lifestyle Applications | MATEC Web of Conferences." *Feast in: A Machine Learning Image Recognition Model of Recipe and Lifestyle Applications | MATEC Web of Conferences*, 25 Jan. 2021, www.matec-conferences.org/articles/mateconf/abs/2021/04/mateconf_eureca2020_04006/mateconf_eureca2020_04006.html.

[47] "Selective Video Enhancement in the Laguerre–Gauss Domain." *Selective Video Enhancement in the Laguerre–Gauss Domain* - ScienceDirect, 13 Oct. 2022, www.sciencedirect.com/science/article/pii/S0923596522001552.

[48] Guide on Openpose Library-<https://github.com/CMU-Perceptual-Computing-Lab/openpose>