Hand Recognition Solution as Conference Room Intelligent Lighting Control Technique

Abdulvahit KARAİL
Institute of Graduate Studies in Pure and Applied Sciences, Electrical and Electronics Engineering, Marmara University
Istanbul, Turkey
abdulvahitkarail@marun.edu.tr

Veysel Gökhan BÖCEKÇİ
Department of Electrical and Electronics Engineering, Faculty of Technology, Marmara University
Istanbul, Turkey
ORCID: 0000-0003-4559-7173

İsmail KIYAK
Department of Electrical and Electronics Engineering, Faculty of Technology, Marmara University
Istanbul, Turkey
ORCID: 0000-0002-5061-6378

ABSTRACT—In communication technologies, it can be solved in systems that are extremely simple, detailed, among the tools for use in the electronic environment, internet-connected, development and integrated systems. With the rapid development of technology, smart led lamps, smart office, smart conference room lighting, road lighting fixtures and therefore smart lighting systems have been the reason why smart lighting systems can enter daily life rapidly. Sensors which are the components of smart lighting systems, feedback from users, cloud services, smart algorithms become more widespread with artificial intelligence in an easier, faster, and more economical way, causing a change in traditional lighting control and improvements in energy saving, functionality, and user centered lighting. In this study, the mediapipe framework is intended to be used for office conference room. MediaPipe is indeed an open-source framework that enables the processing of computer vision data, including video and audio. The lighting system was built using the artificial intelligence supported mediapipe hand recognition solution, and artificial intelligence and image processing algorithms were used. Each hand gesture corresponds to a different command to control the light. There are 3 main menus are determined for the hand and defined to the system. Main menus can be reached with every hand movement. After entering the main menu, intermediate menus are processed. Lighting application can be realized through control commands obtained by hand movements, digital/analog information conversion and Dali control system. It is concluded that the ease of control brought by this method is more advantageous in terms of its purpose than other methods because it is contactless and does not require any hardware device in the hands of the presenter.

Keywords—artificial intelligence, smart lighting, mediapipe, image processing, lighting control

I. INTRODUCTION

Hand pose prediction and gesture recognition have started to attract great attention in recent years due to their human-computer natural interaction HCI (Human-computer interaction) features. Humans and machines communicate through the use of gestures [1].

Various methods have been suggested for gesture recognition systems to obtain the required information [2][3]. Examples of these methods include Hidden Markov Model (HMM), Support Vector Machine (SVM), and Principal Component Analysis (PCA).

Recent reviews have exclusively covered gesture recognition system applications [4] within the domains of human-computer interaction and robot control, utilizing a variety of algorithms [3][5]. Examples of these algorithms are Artificial Neural Network, Canny Edge Detection Algorithm, edge and skin detection algorithms.

Hand gesture recognition systems have found application in various fields, as referenced in [3][6]. These applications encompass virtual environments, robot control, medical systems, and more, showcasing the versatility of gesture recognition technology.

Hand gesture tracking is a critical component in AR/VR to provide a natural way to interact and communicate. Hand detection has found applications across various domains, including the interpretation of sign language, offering valuable support for individuals with disabilities [7]. Within the medical field, it has been applied to monitor user's real-time movements and enhance the comprehensibility of medical data visualization [8]. With a research [9], in this regard, its application in 3D games has been researched. In another study [10], real-time hand movement was used to control robot, for makeeasy robot-human interaction. Another research project explored ways to interact with a mobile device without requiring sensors attached to the human body [11]. In a separate investigation [12], the usability of gesture-based interaction technology in virtual reality (VR) was assessed, alongside a discussion of the existing challenges associated with gesture interaction.

In gesture recognition, hand gesture recognition is the most popular, due to the abundance of application areasand its easy application. It is one of the popular areas.

The main methods used in this area are listed as Kinect,Leap motion, and time of flight (ToF) [13].

Fang and friends, in their study, they implemented a video-based hand gesture detection and tracking system. They drew attention to the background problem in video-based studies. They used optical flow and color marking methods for hand recognition and tracking. They successfully described 6 hand movements with the multiple algorithms that they used [14].

Intelligent lighting control systems are systems that are equipped with sensors for human-centered lighting as well as presence and daylight control, and need smart algorithms and applications such as artificial intelligence for control.

Improving the quality of the light produced, regulating circadian rhythm [15], improving efficiency factor [16], since they can provide many benefits such as accelerating growth for plants [17], it is stated that they have a positive effect on research on industrial implementations, architecture, and human physiology, and they can find wide application areas in this context. To create healthy working environments, it is essential to provide a sufficient and stable illumination level.

In this study, the lighting of an environment is intended to be done controlled using MediaPipe, and an artificial intelligence-based system have used to determine thequality of lighting in a closed area.

2023 7th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT) | 979-8-3503-4215-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISMSIT58785.2023.10304992

Long-term measurements with measuring devices, which cannot be calculated with classical calculation methods, are expected to be made in a short time with minimum error thanks to artificial intelligence.

Thus, the lighting control was made more convenient and more efficient with a camera and an integrated system, due to the absence of light level problems in traditional image processing, the absence of a control button in the hand and the elimination of the need to be connected to the lectern.

As a result, it is aimed to control the lighting control of a room or office with a solution that is applicable and impressive for everyone. Computer programs and models for hand pose recognition, lighting control will be investigated, and this will be done using a deep neural network solution called MediaPipe.

II. METADOLOGY

The hand tracing solution leverages a machine-learning pipeline, which includes a model that operates on the cropped region supplied by the palm detector, delivering precise landmarks.

Supplying accurately cropped palm images to the hand landmark model significantly minimizes the necessity for rotations, translations, and scaling, enabling the network to allocate the majority of its capabilities to enhance landmark localization accuracy.

A. Hand Landmark Model

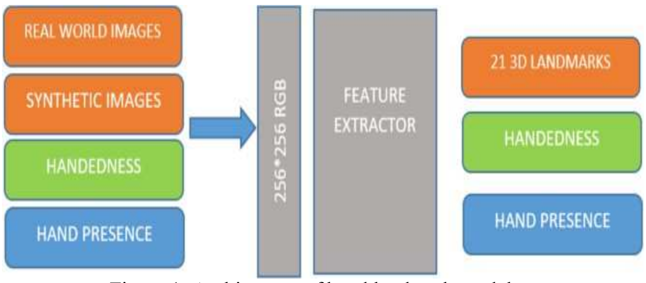

Figure 1: Architecture of hand landmark model.

The hand signal yields three distinct outputs. Each is trained with the corresponding datasets marked with the same color. Subsequent hand landmark model is determined through regression, this model excels at accurately locating 21 coordinates representing landmarks within the identified hand regions. It has the capacity to acquire a dependable internal representation of hand pose. Additionally, this model demonstrates resilience in cases of partially visible hands and self-occlusions. The model's three outputs are illustrated in Figure 1.

1. The 21 hand landmarks encompass x, y, and relative depth information.
2. A hand flag is employed to signify the likelihood of a hand being present in the input image.
3. It includes a binary classification indicating handedness, distinguishing between left or right hand.

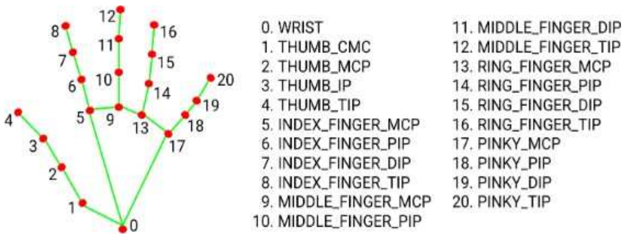

Figure 2: Hand Landmarks [25]

Commands were written according to the hand landmarks shown in figure 2. Figure 2 is taken from Mediapipe's official website.

For the 21 landmarks, the same architecture as described in reference [18] is employed. These coordinates are acquired through learning from real-world images and datasets.

To address tracking failures, an additional model output, similar to the approach described in reference [19], has been developed. This output is designed to estimate the probability that a properly aligned hand is present within the given crop, helping to recover from tracking issues.

B. MediaPipe

MediaPipe is indeed an open-source framework specifically designed for constructing pipelines that allow for computer vision inference on a diverse range of sensory data, including video and audio inputs. MediaPipe is indeed a framework designed for constructing machine-learning pipelines to process time-series data, which includes video, audio and more. Google open-sourced MediaPipe in July 2019.

MediaPipe is primarily utilized for rapidly prototyping perception pipelines that involve AI models for inference and other reusable components. Additionally, it aids in deploying computer vision applications in demos and applications across various hardware platforms.

MediaPipe was developed for machine learning (ML) teams and software developers working on production-ready ML applications, as well as for students and researchers who share their code and prototypes as part of their research efforts. MediaPipe's cross-platform capabilities extend to Desktop/Server environments, android, iOS, and embedded devices like Raspberry Pi, making it versatile for a wide range of applications.

With the assistance of MediaPipe [20], the hand-tracking pipeline can be structured as a directed graph consisting of modular components referred to as Calculators.

MediaPipe, being an open-source framework, is developed for building complex sensing pipelines that utilize accelerated inference on platforms like GPU or CPU. It currently offers fast, accurate, and standalone solutions for various tasks [21].

An object detection algorithm is employed to locate the palm of the hand, and the data from this algorithm is extended to identify the fingers by means of the hand sign model. The aim of this study is to apply neural network solutions with computer programs to be used for hand gesture detection. MediaPipe hands provide the user with real-time hand tracing using a camera. Mediapipe uses machine learning. Machine learning is utilized to extract 21 pieces of 3D landmarks from a single image. This involves a two-stage process. The first stage is the palm-sensing model, and the second stage is the hand sign model [22].

III. SYSTEM DESCRIPTION

A. Modelling

Three different datasets were selected to train the MediaPipe models. The datasets include the Wild dataset, Google's internally collected dataset, and a Synthetic dataset.

Using the Wild dataset for the palm detector is a logical choice, as it offers sufficient variability and accuracy for hand localization due to its diverse and real-world data. The important thing is that all of the datasets are used to train the hand sign model. Consequently, as a result, training the palm detector is a logical choice, as it is more straightforward to estimate bounding boxes around solid objects like palms and fists, making the process more efficient and accurate. Encoder, decoder feature is used for greater location situational awareness, such as FPN (Feature Pyramid Network), eventually, reducing the loss of focus is done to promote a greater degree of variability or large-scale variance in the context of the task at hand [23].

Steps of the object detection module:

These steps involve making assumptions about bounding boxes, resampling properties for each box, and incorporating a high-quality classifier, which collectively contribute to accurate object detection.

Another model is the hand-pointing model. After palm detection across the entire image, the hand gesture model utilizes regression to precisely localize 21 pieces of 3D coordinates within specific hand regions.

21 pieces of 3D cues are output by hand cue models, the hand tag is used to know the probability of finding it, and the hand is used to know which hand is on the right or left. The topology similar to the multi-image preload used is used to determine 21 landmarks, especially to improve the performance of a particularly important key point detector [24].

Light control will be done with my 3 main menus in hand gesture. With each hand movement, the main menus will be accessed. After entering the main menu, the intermediate menus will be processed.

3 main menus:

1. Illuminance Level
2. Color Temperature
3. Microphone control

First, we can realize the control of the lamp. The lamp can be turned on and off. Lighting level increase and decrease will be done from the 1st menu. The color temperature will be adjusted from the 2nd menu in the office. Microphone on/off control will be done from menu 3. There are 3 buttons as numbers on screen. The menu we want can be selected with these buttons as numbers.

B. Smart Lighting System Control Strategy

The hand recognition solution was applied to control the smart lighting system lighting system as part of the study's scope. This method can basically be expressed as the production of a control strategy that will turn on and off the lighting system, increase the illuminance level by dimming it between 0% and 100%, and reduce the illuminance level by dimming it between 100% and 0%.

Color change between our hand position and the colors on the color scale. I mean we change implement it to 270 degree hand rotation.

Figure 3: The smart lighting system control strategy is also seen.

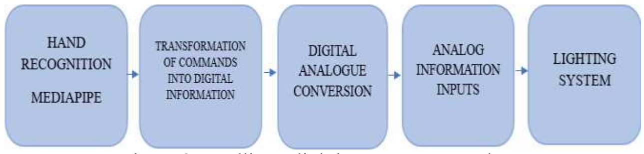

Figure 3: Intelligent lighting system control strategy.

In this context, the commands obtained in this section have been transformed into digital information.

C. Results

The project consisted of four phases. Creating a data set for lighting, using the created data in the mediapipe hand recognition solution, defining the codes corresponding to the meaning for each hand movement, making the software and controlling the lighting. After creating datasets for lighting control for an office or hall, hand gesture recognition using MediaPipe, defining what the gesture means, and using them for lighting control by the program.

There are two models and datasets used to train the data when using MediaPipe Hands, described below. Comparisons were made between points 5 and 8 on the hand. Lighting on or off command is given according tothe position of 5 to 8 points.

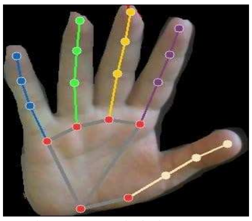

Figure 4: Lighting on

Figure 4: If the point 8 on the hand is higher than the point number 5, that is, if it is bigger, this means that thelighting is on.

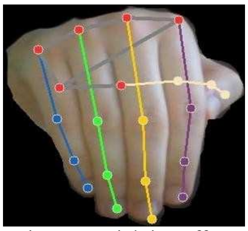

Figure 5: Lighting off

Figure 5: If the point 8 on the hand is smaller than the point number 5, that is, if it is below, this means that the lighting is off.

Comparisons were made between points 4 and 0 on the hand. Lighting on or off command is given according tothe position of 4 to 0 points.

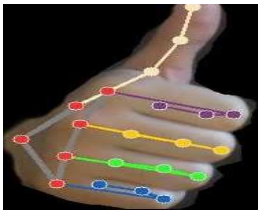

Figure 6: Increase illumination level

Figure 6: If the 4 point on the hand is above the 0 point, that is, if it is larger, it means an increase in the illumination level.

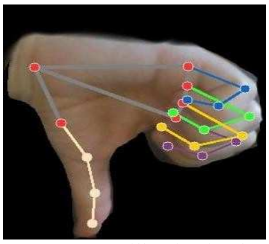

Figure 7: Decrease illumination level

Figure 7: If the 4 point on the hand is below the 0 point,that is, if it is smaller, it means a decrease in the illumination level.

IV. CONCLUSION

In this study, we controlled the light of the office environment from the hand movements we defined.Each movement was made sense and the lighting of the environment was controlled.

As a result, the lighting control of a room or office has been controlled with a solution that is feasible and impressive for everyone.

Hand pose recognition has been successfully performed using computer programs and models for lighting control and a deep neural network solution called MediaPipe. Accuracy evaluations of the brightness commands with 100 repetitions are shown in figure 8.

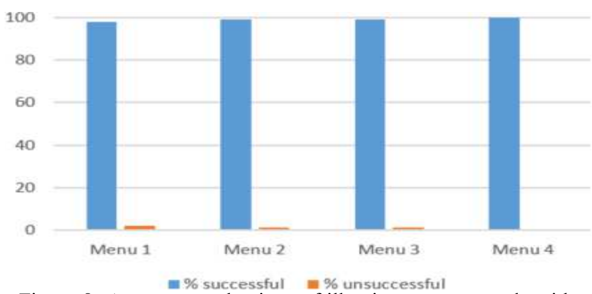

Figure 8: Accuracy evaluations of illuminance commands with 100repetitions

For menu 1 (the level of illumination controlled), the move was tried 100 times and the success rate was 98 percent. For menu 2(color temperature controlled), the move was tried 100 times and the success rate was 99 percent. For the 3rd menu (microphone controlled), the move has been tried 100 times and thesuccess rate is 100 percent.

REFERENCES

[1] Mokhtar M. Hasan, Pramoud K. Misra, (2011). "Brightness Factor Matching For Gesture Recognition System Using Scaled Normalization", International Journal of Computer Science & Information Technology (IJCSIT), Vol. 3(2)

[2] Simei G. Wysoski, Marcus V. Lamar, Susumu Kuroyanagi, Akira Iwata, (2002). "A Rotation InvariantApproach On Static-Gesture Recognition UsingBoundary Histograms And Neural International Journalof Artificial Intelligence & Applications (IJAIA), Vol.3, No.4, July 2012 173 Networks," IEEE Proceedings of the 9th International Conference on Neural InformationProcessing, Singapura.

[3] Joseph J. LaViola Jr., (1999). "A Survey of Hand Posture and Gesture Recognition Techniques andTechnology", Master Thesis, Science and Technology Center for Computer Graphics and Scientific Visualization, USA.

[4] Rafiqul Z. Khan, Noor A. Ibraheem, (2012). "Survey on Gesture Recognition for Hand Image Postures", International Journal of Computer And Information Science, Vol. 5(3), Doi: 10.5539/cis.v5n3p110

[5] Thomas B. Moeslund and Erik Granum, (2001). "A Survey of Computer Vision-Based Human Motion Capture," Elsevier, Computer Vision and Image Understanding, Vol. 81, pp. 231–268.

[6] S. Mitra, and T. Acharya. (2007). "Gesture Recognition: A Survey" IEEE Transactions on systems,Man and Cybernetics, Part C: Applications and reviews,vol. 37 (3), pp. 311- 324, doi: 10.1109/TSMCC.2007.893280.

[7] Ade Silvia and Nyayu Latifah Husni. Hand contour recognition in language signs codes using shape-based hand gestures methods. In Hand Contour Recognition inLanguage Signs Codes Using Shape Based Hand Gestures Methods. Sriwijaya University, 1st International Conference on Computer Science andEngineering, 9 2014.

[8] Helman Stern, Yael Edan, Michael Gillam, Craig Feied, Mark Smith, and Jon Handler. A Real-Time Hand Gesture Interface for Medical Visualization Applications, volume 36, pages 153–162. Springer, 04 2010.

[9] Harpreet Kaur and Jyoti Rani. A review: Study of various techniques of hand gesture recognition. In 2016IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), pages 1–5, 2016.

[10] Qing Gao, Yongquan Chen, Zhaojie Ju, and Yi Liang. Dynamic hand gesture recognition based on 3d hand pose estimation for human-robot interaction. IEEESensors Journal, PP:1–1, 05 2021.

[11] Atsushi Shimada, Takayoshi Yamashita, and Rin- ichiro Taniguchi. Hand gesture based tv control system towards both user- amp; machine-friendly gesture applications.In The 19th Korea-Japan JointWorkshop on Frontiers of Computer Vision, pages 121–126, 2013.

[12] Yang LI, Jin HUANG, Feng TIAN, Hong-An WANG, and Guo-Zhong DAI. Gesture interaction in virtual reality. Virtual Reality & Intelligent Hardware, 1(1):84–112, 2019.

[13] Cheng, H., Yang, L., & Liu, Z. (2015). Survey on 3D hand gesture recognition. IEEE transactions on circuits and systems for video technology, 26(9), 1659- 1673.

[14] Fang, Y., Wang, K., Cheng, J., & Lu, H. (2007, July). A real-time hand gesture recognition method. In 2007 IEEE International Conference on Multimedia andExpo (pp. 995-998). IEEE.

[15] J.H. Oh, S.J. Yang, Y.R. Do, "Healthy, natural, efficient and tunable lighting: four-package white leds for optimizing the circadian effect, color quality and vision performance", Light Sci. Appl. 3 (2) (2014) e141.

[16] R.F. Karlicek, "Smart lighting-beyond simple illumination," 2012 IEEE Photonics Society Summer Topical Meeting Series, IEEE (2012) 147–148.

[17] G.D. Massa, H.-H. Kim, R.M. Wheeler, C.A. Mitchell, "Plant productivity in response to led lighting", HortScience 43 (7) (2008) 1951–1956.

[18] Tomas Simon, Hanbyul Joo, Iain A. Matthews, andYaser Sheikh. Hand keypoint detection in single imagesusing multiview bootstrapping. CoRR, abs/1704.07809,2017.

[19] Yury Kartynnik, Artsiom Ablavatski, Ivan Grishchenko, and Matthias Grundmann. Real-timefacial surface geometry from monocular video on mobile gpus. CoRR, abs/1907.06724, 2019.

[20] Camillo Lugaresi, Jiuqiang Tang, Hadon Nash,Chris Mc-Clanahan, Esha Uboweja, Michael Hays, FanZhang, Chuo-Ling Chang, Ming Guang Yong, Juhyun Lee, Wan-TehChang,Wei Hua, Manfred Georg, and Matthias Grundmann.Mediapipe: A framework for building perception pipelines. volume abs/1906.08172, 2019.

[21] Ivan Grishchenko and Valentin Bazarevsky. Mediapipe holistic - simultaneous face,hand and pose prediction, on device, Dec 2020.

[22] https://google.github.io/mediapipe/solutions/hands.html
[23] Fan Zhang, Valentin Bazarevsky, Andrey Vakunov, Andrei Tkachenka, George Sung, Chuo-Ling Chang, and Matthias Grundmann. Mediapipe hands:On-device real-time hand tracking, 06 2020.
[24] Tomas Simon, Hanbyul Joo, Iain Matthews, and Yaser Sheikh. Hand keypoint detection in single images using multiview bootstrapping. In Hand Keypoint Detection in Single Images Using Multiview Bootstrapping, pages 4645–4653, 07 2017.
[25] https://google.github.io/mediapipe/solutions/hands.html