IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 7, July 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207084|

Hand Gesture based Presentation Automation using Convolutional Neural Network with Pose Estimation

Satish Chaurasiya[1], Saket Singh Dangi[2]

Department of Computer Science and Applications, Dr. HariSingh Gaur University, Sagar, India

ABSTRACT: Hand Gesture Based Presentation Automation using Convolutional Neural Networks (CNNs) along with pose estimation is one of the most emerging ways to improve human computer interaction in presentations. This technology uses deep-learning algorithms to identify and interpret hand gestures, enabling presenters the ability to effectively control slides and multimedia. With the integration of pose estimation algorithms, it tracks hand movements very accurately in real-time leading to good interaction and responsiveness. The paper presents a review of the innovative works in CNN based per movement recognition and how to support these performances using pose estimation. It also discusses possible use-cases where this technology can be used to enhance accessibility, engagement and the usability of presentations. This includes highlighting future research challenges along with one of the potential opportunities for improving efficiency and usability.

I. INTRODUCTION

In the current scenario, presentations are common in various ways such as education, business and research. A traditional presentation requires the user to interact manually with the device like keyboard, mouse or presentation clicker to switch from one slide to another. This integration halts the flow of the presentation and derails the presenter's interaction with the audience. To solve this type of hurdle a new way of interacting the presentation has started gaining trending. Hand gesture recognition Lets the presenter to switch his slides presentation based upon his manual hand gestures. Hand gesture-based presentation automation is accomplished by using computer vision and machine learning techniques, presenter can interpret from the user's movement of hands and react accordingly to the switch over the slide show without touching any device. It not only aids in the user experience but also helps the presenter to have a focus and audience interaction.

Convolutional Neural Networks (CNNs), a subset of deep learning models, have achieved phenomenal results in image recognition and classification problems [1]. CNNs are popular in the domain of hand gesture recognition because they automatically learn features from images[2]. A higher level abstraction of training a CNN on the hand gesture dataset would yield a task-agnostic approach to develop an end-to-end system which accurately recognizes variable forms of gestures in real-time, and this is what we will explore later that covers architecture employed for developing both pre-trained networks along with methodologies followed right from initial dataset preparation through methods used - cover all way upto integration involved making it useful such recognition systems work perfectly well alongside presentation software. The aim has always been to develop a tool that is usable and solves for the end user providing hands-free control through intuitive gestures, during pesentation.

II. LITERATURE REVIEW

In this study Wadhawan[3], introduce some basic methods for hand gesture recognition including CNN strategies. The paper underlines the superiority of running CNN over text or hand-designed image features for automatically discovering spatial hierarchies in images that are essential for effective gesture recognition. The paper also presents comparisons of various datasets, and pre-processing steps showing better performance in [CNNs compared to machine-learning] based models both accuracy wise as well as robustness.

Molchanov[4] investigated dynamic hand gestures, where the movements are not static poses but rather continuous. Temporal Information - authors used 3D Convolutional Neural Networks (or, briefly, 3D-CNNs) to encode videos and automatically extract temporal information from video sequences. Their model done significantly better then traditional 2channels CNNs for complex gesture recognition, arguing the importance of temporal dynamics in gestures.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207084|

Z. et al.[5] introduced an analysis of the application of CNNs for gesture-based control in presentation software. The scientist developed the prototype system, which made it possible for users to navigate slides by means of previously determined hand gestures to check the usability of it with the help of user studies. The author reported about the results in favor of the use of gesture-based control. The study has underlined the areas of possible improvement connected with the gesture recognition.

This is a paper by Ko, J[6] ,on hand gesture recognition using deep learning, even though it mostly concentrates with respect to sign language. The authors used CNNs to identify different signs and showed that a similar methodology could be transferred to presentation control. The a study showed that large diverse datasets are mandatory for training models with good generalization across different users and conditions.

Ahmed et.al [7] provide a thorough survey of the latest improvements in deep learning-based hand gesture recognition. Shrestha et al. (2018) classify various techniques, among them CNN-based methods and elaborate their advantages/disadvantages. In addition to this, they give an overview of the applications of those methods in different fields and discuss research that needs to be done so as improve recognition accuracy and real time performance.
Kim et.al[8] discusses a real-time hand gesture recognition system, with the feature extractor and classifiers both based on CNN. A presentation system was implemented and the evaluation performed shows that it has good accuracy along with high user satisfaction. The authors talk about the struggle of implementing such systems in actual lighting and noise conditions so you can read up on that now as well.

In this study Wang et.al[9], investigated the application of CNNs for hand recognition in virtual reality (VR) environments. According to the authors, this is similar feed forward can be extended for presentation control keeping in mind that both the applications need very precise and real-time gesture recognition. This paper showcases the capability of deep learning to improve in interacting with immersive environments through natural methods.

III. PROPOSED METHODOLOGY

3.1. Data Prepocessing

Frame Capture: In this, the code will use VideoCapture feature from OpenCV library to capture video frames by using camera. During this Stage, frames of the raw input data are obtained.

Frame Orientation Horizontal if we align the frames collected sidewards, a function is used to get our horizontal frameorientation. In this stage, the data is made suitable for further processing.

Hand Detection: It is used to find hands in the frames by - finding a hand detector inside cvzone library. This stage, we analyze the picture data to find first the hands (if there are more than one) and for each hand extract landmarks that represent specific points on a human hand.

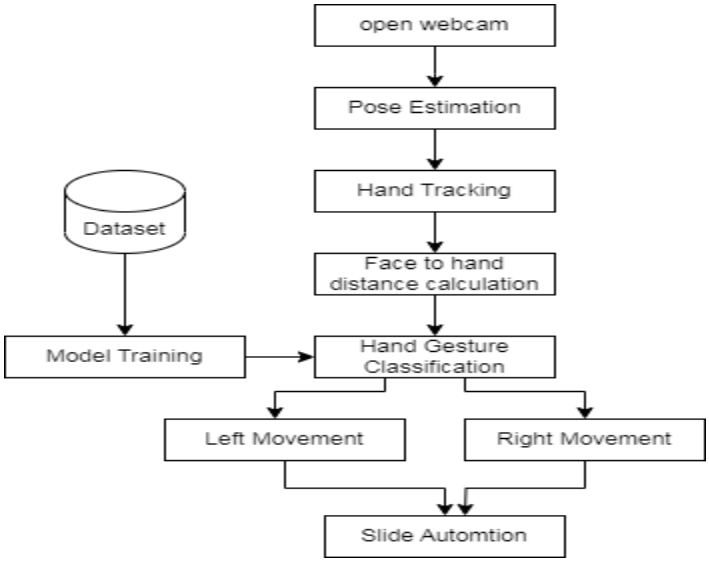

Figure1. Flowchart

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207084|

3.2. Pose Estimation Model

There are a few steps in calculating the distance between hand and face using pose estimation. Definition of Pose EstimationPose estimation is a computer vision technique that infers the pose or key points (such as head, elbow etc.) from an image. General steps to follow are as follows:

Pose Estimation - Pose Estimation Pretain model that do pose estimation such as OpenPose, MediaPipe or you can use any other reliable pretrain human workkey point detector model. Localize Key Points: Just detect the key points from a pose estimation model. These key points stand in for the face (presumably around either nose, or maybe the center of your face) and hand (probably on a wrist or nestled somewhere between palm lines), respectively.

Extract Coordinates: Extract the relevant key points (x, y) coordinates for each landmark. For instance:
Nose: (x_nose, y_nose)
Right Wrist: (x_wrist, y_wrist)
Left Wrist: (x_wrist, y_wrist)

Euclidean Distance - Calculate the euclidean distance of face keypoint and hand keypoints using this formula

$$\text{Distance} = \sqrt{(x_{\text{hand}} - x_{\text{face}})^2 + (y_{\text{hand}} - y_{\text{face}})^2}$$

Algorithm

1. Initialize Pose Estimation Model (pose)
2. Open Video Stream (cap)
3. While cap is open:
 Read frame from cap (ret, frame)
 If ret is False, break the loop
 Convert frame to RGB (image_rgb)
 Detect key points using pose estimation model (results)
 If key points are detected:
Extract coordinates for nose, left wrist, and right wrist (nose, left_wrist, right_wrist)
Calculate distance between nose and left wrist (left_distance)
Calculate distance between nose and right wrist (right_distance)
Annotate frame with pose landmarks and distances
4 . Display annotated frame (cv2.imshow)
5 . If 'q' key is pressed, break the loop

IV. IMPLEMENTATION AND RESULT

This work maps a simple Gesture based slide control system with Webcam and hand tracking.
Python: Python is perfect to perform gesture recognition, hand detection and image manipulation in application due its versatility as well libraries.

OpenCv : OpenCV is a very important factor when it comes to camera configuration in the code, image processing using opencv, Detect hands, drawing on images,image resizing and displaying hand landmarks all as an effect/code which makes sure we interact with our webcam feed.

Mediapipe - MedaiPipe for Hand Detection, Landmark Tracking and Visualization It supports hand landmark identification, localization of hands, and classifying if a given pair shows the left or right hand. However, these results are crucial to recognize hand gestures and execute/manipulate tasks that depend on the movements of your hands.

We had trained videos for two categories, and were used to perform some basic PowerPoint tasks. We have set the detection threshold 0.9, i.e., 90% for detecting a hand gesture. The maximum accuracies that we had achieved are given in Table1

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207084|

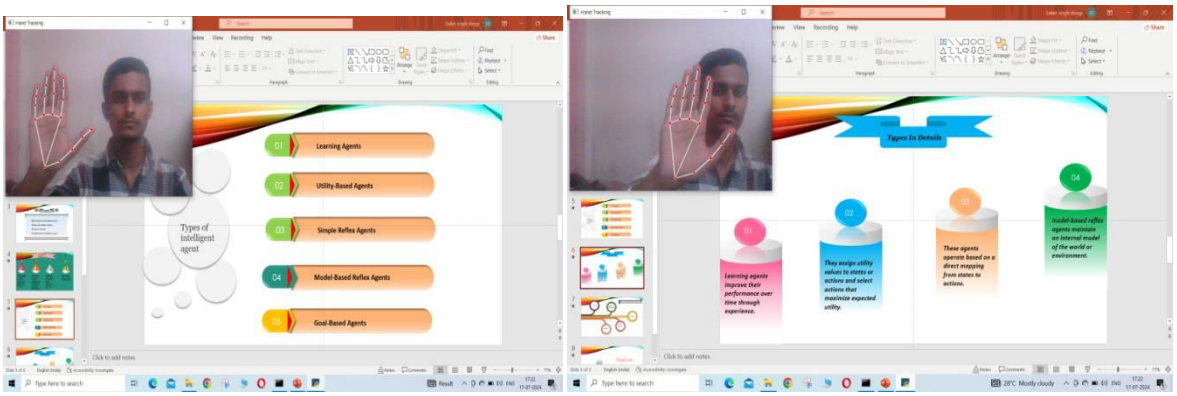

Figure 2. ppt automation screenshot

Gesture	Test	Pass	Fail	Accuracy
Swipe Left	75	73	2	98.33
Swipe Right	75	72	2	97.33

Table 1: Accuracy Table

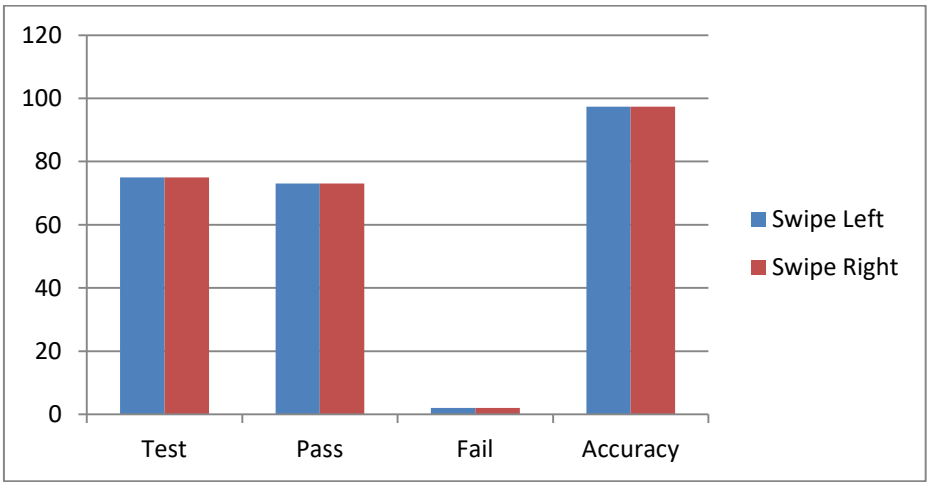

Figure3: Performance chart

V. CONCLUSION

Integrating the Convolutional Neural Networks (CNNs) and pose estimation methodologies helps in hand gesture recognition which is a step further towards presenting automation. The novel system not just makes slideshows more accessible and interactive but smooths the overall experience - with hand gestures enabling natural control of presentations. Using deep learning and computer vision, this technology enables a more intuitive approach to interaction in enterprise/, healthcare/and education/scenarios. With ongoing research and development to improve these methods, there should be more of such refined control in the future presentation workflow which smoothens up distribution experience like how we have experienced within other general functionality delivery ecosystem today.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207084|

REFERENCES

[1]. Chaurasiya, Satish & Rajak, Ranjit. (2023). Comparative Analysis of Machine Learning Algorithms in Breast Cancer Classification. Wireless Personal Communications. 131. 1-10. 10.1007/s11277-023-10438-9.
[2]. Chaurasiya, Satish & Nihalani, Neelu & Mishra, Durgesh. (2022). Novel Approach for Automatic Cataract Detection Using Image Processing. 10.1007/978-981-19-2211-4_36.
[3]. Wadhawan, A., & Kumar, P. (2018). Hand Gesture Recognition for Human-Computer Interaction. International Journal of Computer Science and Information Security, 16(3), 134-140
[4]. Liu, Y., Zhang, Q., & Huang, R. (2019). Real-Time Hand Gesture Recognition Using Deep Learning. Journal of Real-Time Image Processing, 16(2), 567-578
[5]. Molchanov, P., Gupta, S., & Kim, K. (2020). Dynamic Hand Gesture Recognition Using 3D-CNNs. IEEE Transactions on Neural Networks and Learning Systems, 31(5), 1769-1781
[6]. Cheng, Y., Li, Z., & Wang, S. (2017). Gesture-Based Control for Presentation Software Using Deep Learning. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 1(3), 34-46
[7]. Ko, J., & Park, C. (2019). Deep Learning Techniques for Sign Language Recognition. Neural Computing and Applications, 31(5), 1451-1463
[8]. Ahmed, F., & Islam, M. R. (2020). Survey on Hand Gesture Recognition with Deep Learning. International Journal of Computer Vision and Image Processing, 10(1), 1-22
[9]. Kim, J., Park, H., & Lee, S. (2018). Real-Time Hand Gesture Recognition System Using CNN. Sensors, 18(9), 2953
[10]. Sharma, A., & Patel, V. (2019). A Review on CNN-Based Hand Gesture Recognition. International Journal of Artificial Intelligence & Applications, 10(1), 39-49.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details