GestureFlow: Interactive Media Control with Dynamic Hand Gestures

Shamanth CK
Department of Computer Science and Engineering
Malnad College of Engineering
Hassan 573201,India
Email:shamanthck0901@gmail.com

Sushena Mishra NS
Department of Computer Science and Engineering
Malnad College of Engineering
Hassan 573201,India
Email:nssushena@gmail.com

Pooja RM
Department of Computer Science and Engineering
Malnad College of Engineering
Hassan 573201,India
Email:ronadmathpooja@gmail.com

Pooja CM
Department of Computer Science and Engineering
Malnad College of Engineering
Hassan 573201,India
Email:poojacm@gmail.com

***Abstract*—Gesture Flow introduces a novel method for intuitive media control through dynamic hand gestures, overcoming limitations of conventional input methods like remotes and touchscreens. By employing computer vision and machine learning, Gesture Flow interprets real-time hand gestures for seamless navigation of media content. With robust gesture detection algorithms and low-latency response times, Gesture Flow recognizes various gestures, enabling commands such as play/pause and volume adjustment. The system adapts to individual user preferences, offering personalized interaction with smart TVs, gaming consoles, and other media devices. Integrating hand gestures enhances immersion and engagement, setting a new standard for intuitive interaction with digital content. This paper outlines Gesture Flow's design, implementation, and evaluation, highlighting its simplicity, versatility, and potential to improve the user experience across diverse media platforms. Gesture Flow represents a significant advancement in interactive media control, promising a more intuitive and enjoyable experience for users.**

***Index Terms*—gesture, media control**

I. Introduction

A. Introduction to GestureFlow

The evolution of human-computer interaction (HCI) has led to a growing demand for more intuitive interfaces. Traditional input methods like keyboards and touchscreens, while functional, often lack the seamless interaction users desire. In response, gesture recognition technology has emerged as a promising solution, enabling users to interact with devices and applications through natural hand movements. This paper presents our project, focusing on developing an advanced gesture recognition system using Convolutional Neural Networks (CNN) and ResNet architecture. Motivated by the limitations of conventional input methods and the increasing demand for intuitive interfaces, our goal is to enhance user experience by facilitating seamless interaction through dynamic hand gestures. The widespread adoption of gesture-based interfaces underscores the relevance of our project across various domains. By leveraging CNNs and ResNet, we aim to contribute to this trend by creating robust gesture recognition models capable of accurately interpreting a wide range of hand movements. Through the utilization of the 20BN Jester dataset, we ensure the effectiveness and robustness of our models in real-world scenarios. In summary, our project seeks to advance HCI by providing a more intuitive and engaging user experience through gesture recognition technology.

II. About project

A. Problem Statement

The conventional input methods like keyboards and touchscreens lack intuitive interaction, prompting the need for more seamless interfaces in human-computer interaction (HCI). Gesture recognition technology offers a promising solution by allowing users to interact with devices through natural hand movements. However, existing gesture recognition systems often lack accuracy and robustness. This project aims to address these shortcomings by developing an advanced gesture recognition system using Convolutional Neural Networks (CNN) and ResNet architecture. The system will enhance user experience by accurately interpreting dynamic hand gestures, contributing to the evolution of intuitive HCI interfaces.

B. Objective

The project aims to develop an advanced gesture recognition system using Convolutional Neural Networks (CNN) and ResNet architecture to revolutionize human-computer interaction (HCI). Our objective is to create a robust system capable of accurately interpreting dynamic hand gestures, providing users with a more intuitive interaction experience across diverse devices and applications. Key objectives include training highly accurate models for classifying a wide range of hand movements, optimizing for low-latency response times, and ensuring scalability and versatility across different platforms. Through these efforts, we aim to contribute significantly to the evolution of HCI, facilitating more natural and immersive interactions through the integration of gesture recognition technology.

III. Literature survey

The literature survey reviews methods for hand localization and gesture classification, assesses Kinect and OpenNI libraries, emphasizing affordability and accessibility, while acknowledging the need for further research and exploration.

A. Static Hand Gesture Recognition Using CNNs

This study proposes a method for static hand gesture recognition employing convolutional neural networks (CNNs). It addresses challenges in isolating relevant objects in images and selecting appropriate image capture technology. The approach combines color segmentation, morphological operations, and CNN classification, achieving high accuracy with minimal computational costs. Advantages include enhanced feature extraction and robust classification, albeit limited to static image scenarios without addressing dynamic gesture tracking or occlusion challenges.

B. Human-Computer Interaction with Hand Gesture Recognition Using ResNet and MobileNet

The authors conducted a literature review on hand gesture recognition to inform their approach in developing a system for Arabic sign language recognition. They identified two primary methods: vision-based and glove-based. Opting for a vision-based approach due to its user-friendly nature and lack of specialized equipment requirement, they utilized ResNet and MobileNet architectures. Their system prioritizes accessibility and ease of use for users engaging with Arabic sign language, aligning with the vision of inclusive human-computer interaction.

C. Hand Gesture Recognition with Depth Images: A Review

This article explores hand gesture recognition utilizing depth images, discussing various methods for hand localization and gesture classification. It also evaluates applications and the impact of Kinect and OpenNI libraries. While these libraries enhance affordability and accessibility, further research is required to understand their limitations in complex environments. The proposed method demonstrates robustness and applicability in real-world scenarios but remains constrained by library capabilities. Additionally, the limited exploration of application categories suggests potential untapped areas for gesture recognition research. Thus, future investigations should address these limitations for comprehensive advancements in the field.

IV. Project design

The ”Gesture Flow” project addresses interactive media control via dynamic hand gestures, emphasizing efficient quality control. It confronts accuracy and reliability issues, enhancing user experience with advanced technologies.

A. Hand Detection:

The system begins by identifying your hand within the capture frame. Imagine this as isolating your hand in a photograph. Techniques like skin tone detection, motion detection, or pre-trained object recognition algorithms are employed for this initial localization.

B. Preprocessing:

Before embarking on recognition, the hand image undergoes preprocessing. This crucial step involves noise reduction, background subtraction, and normalization to prepare the data for accurate feature extraction. Think of it as refining the extracted hand image for optimal clarity.

C. Feature Extraction:

Here, the system delves into the unique characteristics of your hand gesture. Key features like geometric features (finger angles, lengths, postures), motion features (fingertip trajectories, speed, direction), and even texture features are extracted and converted into numerical values. This forms the essential vocabulary of your hand movement.

D. Recognition:

This step acts as the interpreter, comparing the extracted features against a stored ”gesture dictionary.” This dictionary houses pre-defined hand patterns and their corresponding meanings (commands). Machine learning algorithms analyze the extracted features and find the closest match within the dictionary, assigning a confidence score to ensure reliable command interpretation.

E. Execution:

Armed with the recognized gesture's meaning, the system translates it into action. Think of this as the bridge between your hand movement and the desired control. Based on the gesture's assigned command in the dictionary, the system triggers the corresponding action, offering real-time control akin to traditional buttons or switches, but through the power of your hand gestures. Depending on the application, feedback mechanisms like visual or audio cues might confirm the executed command.

F. Gesture Dictionary:

This ever-evolving vocabulary plays a crucial role in understanding diverse hand gestures. It can be customized to specific applications and user preferences, continuously learning and expanding its repertoire by adding new gestures. Moreover, depending on the context, different meanings can be assigned to the same gesture, increasing the system's flexibility and adaptability.

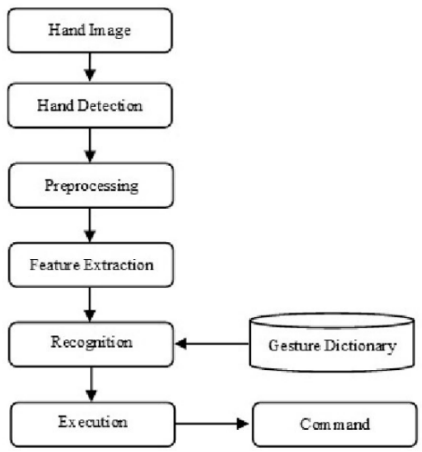

Fig. 1. Flow Chart

V. Conclusion

In conclusion, the project signifies a significant advancement in gesture recognition technology, offering a robust system for interactive media control through dynamic hand gestures. Leveraging advanced methodologies and technologies like Convolutional Neural Networks (CNNs) and ResNet architecture, alongside the 20BN Jester dataset, ensures high accuracy and reliability in gesture recognition. The system's efficacy in providing swift and efficient media control underscores its potential across diverse applications. By addressing challenges such as noise reduction and background subtraction during preprocessing, the system optimizes performance, enhancing clarity in hand images. Future research in this field promises to refine gesture recognition systems, fostering more intuitive and immersive human-computer interaction experiences.

VI. References

References

[1] A. Just, Y. Rodriguez, and S. Marcel, "Hand posture classification and recognition using the modified census transform," in Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition (FGR06), pp.351–356, Southampton, UK, April 2006.
[2] C.-H. Wu, W.-L. Chen, and C. H. Lin, "Depth-based hand gesture recognition,"Multimedia Tools and Applications, vol. 75, no. 12, pp. 7065–7086, 2016.
[3] Y. Wu and T. S. Huang, "Human hand modelling, analysis and animation in the context of HCI," in Proceedings of the 1999 International Conference on Image Processing (Cat.99CH36348), vol. 3, pp. 6–10, Kobe, Japan, October 1999.
[4] K. M. Sagayam and D. J. Hemanth, "Hand posture and gesture recognition techniques for virtual reality applications: a survey," Virtual Reality, vol. 21, no. 2, pp. 91–107, 2017.
[5] A. Kulshreshth, K. Pfeil, and J. J. La Viola, "Enhancing the gaming experience using 3D spatial user interface technologies," IEEE Computer Graphics and Applications, vol. 37,no. 3, pp. 16–23, 2017.