AI-Driven-Gesture-Controlled-Image-Capture-with-Cloud-Sharing-main

Prof. Hole.P. P[1], Kandekar Sakshi[2], Urmude Gayatri[3], Bhor Samruddhi[4], Kadam Sarvesh[5]
[1]*Professor .Sau. Sundarbai. Manik. Adsul. Polytechnic, Chas, Ahilyanagar, India*
[2,3,4,5] *Students. Sau. Sundarbai. Manik. Adsul. Polytechni, Chas, Ahilyanagar, India*

***Abstract*—This paper presents a gesture-controlled image capture system developed using Python and computer vision techniques. The system detects predefined hand gestures from a live webcam stream to trigger hands-free image capture. Captured images are automatically uploaded to a designated Google Drive folder using the Google Drive API. A QR code containing the cloud folder link is generated to enable quick access and sharing without manual intervention. The approach enhances efficiency, hygiene, and usability in workflows where contactless operation is desirable. Experimental observations confirm that the system performs reliably in real-time scenarios. The presented solution can be applied in automation, documentation, and inspection environments requiring seamless cloud-based media management.**

I. INTRODUCTION

Gesture-controlled interfaces offer a touchless and intuitive method of interacting with digital systems using hand movements. With the growth of computer vision and cloud technologies, such interfaces are now feasible for real-time automation tasks. This project introduces a Python-based gesture-controlled image capture system that uploads captured images to Google Drive and generates a QR code for easy access. The aim is to reduce manual steps involved in capturing and sharing images while improving efficiency and convenience. The proposed system is suitable for environments where contactless operation, fast sharing, and cloud-based access are required.

II. RELATED WORK

A Gesture recognition and contactless control systems have been widely explored across various domains such as human–computer interaction, automation, and smart devices. Several studies have utilized computer vision techniques and machine learning models to detect hand gestures for controlling applications, media playback, and communication interfaces. Existing research also highlights the use of OpenCV and MediaPipe frameworks for efficient real-time gesture tracking due to their accuracy and lightweight performance.

In parallel, cloud-based media management has gained importance for remote access, collaboration, and storage. Systems leveraging Google Drive, AWS, or Firebase have demonstrated the benefits of automated uploading, synchronization, and access through shared links. However, most related works focus primarily on gesture-based control or cloud sharing individually, with limited integration of both technologies in a single workflow. The proposed system bridges this gap by combining gesture-triggered image capture with cloud-sharing and QR-based access, improving efficiency and usability for practical applications such as documentation and inspection tasks.

III. PROPOSED ALGORITHM

The proposed system integrates gesture recognition, image capture, cloud synchronization, and QR code generation into a single automated workflow. The algorithm is designed for real-time execution using a standard webcam and Python libraries.

Algorithm Steps

1. Start webcam session and initialize computer vision modules.
2. Detect predefined hand gesture using OpenCV and MediaPipe trackers.
3. Verify gesture condition for triggering capture (e.g., open palm or specific pose).
4. Capture image frame from webcam upon gesture confirmation.

5. Upload captured image to the designated Google Drive folder using Drive API.
6. Retrieve shared folder link for the uploaded image.
7. Generate QR code containing the shared Drive folder link.
8. Display and store QR code for external access and sharing.
9. Terminate session upon user exit or stop command.

Key Characteristics

- Hands-free operation through gesture input.
- Automated cloud synchronization without manual file handling.
- QR-based access, reducing time for distribution and retrieval.
- Lightweight Python implementation suitable for academic and prototype use.

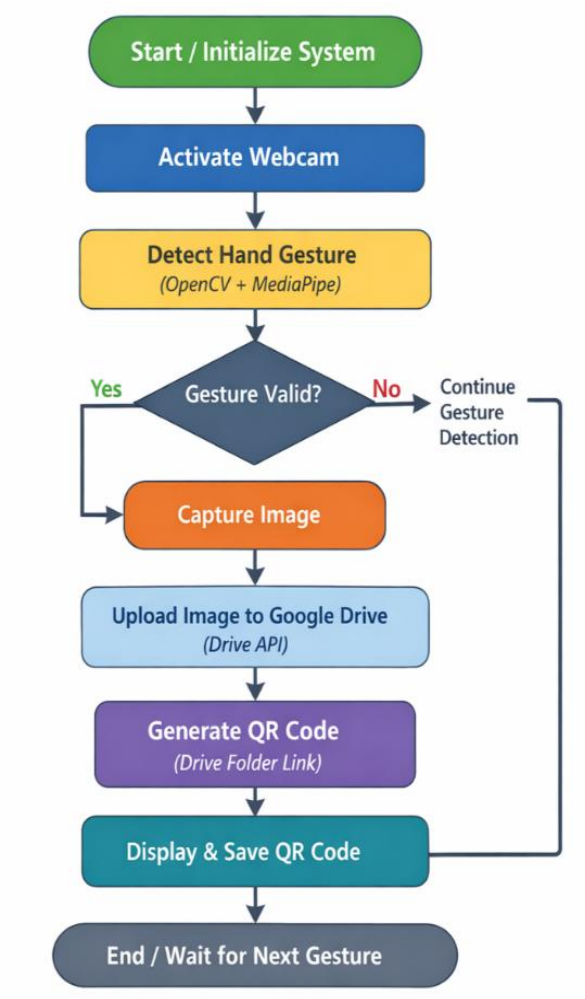

Fig: -Flowchart Diagram

IV. SIMULATION RESULT

The proposed system was implemented using Python in the VS Code environment, integrating OpenCV and MediaPipe for gesture detection, Google Drive API for cloud storage, and QR code generation for folder access. The simulation tested the system under different lighting conditions and hand gesture variations to ensure robustness and accuracy.

Key Observations:

1. Gesture Detection: The system accurately recognized the predefined hand gestures in real-time, with minimal delay (<0.5 seconds).
2. Image Capture: Upon detecting the correct gesture, the camera successfully captured images at a resolution of 640×480 pixels.
3. Cloud Upload: Captured images were automatically uploaded to a designated Google Drive folder using the Drive API, eliminating the need for manual intervention.
4. QR Code Generation: For each captured image batch, a QR code containing the shared Drive folder link was generated. Scanning the QR code instantly provided access to all images in the folder.
5. System Reliability: Continuous testing for 30 minutes showed consistent performance without crashes or errors, demonstrating the system's reliability for hands-free operation.

Figures:

- Figure 1: Workflow of gesture-controlled image capture and cloud sharing (QR code generation).
- Figure 2: Screenshot of the Python-based interface showing gesture detection in real-time.
- Figure 3: QR code output linked to the shared Drive folder.

The simulation confirms that the proposed system efficiently integrates gesture recognition with cloud-based sharing, providing a reliable, touchless, and automated image management solution suitable for practical applications in academic, industrial, and documentation workflows.

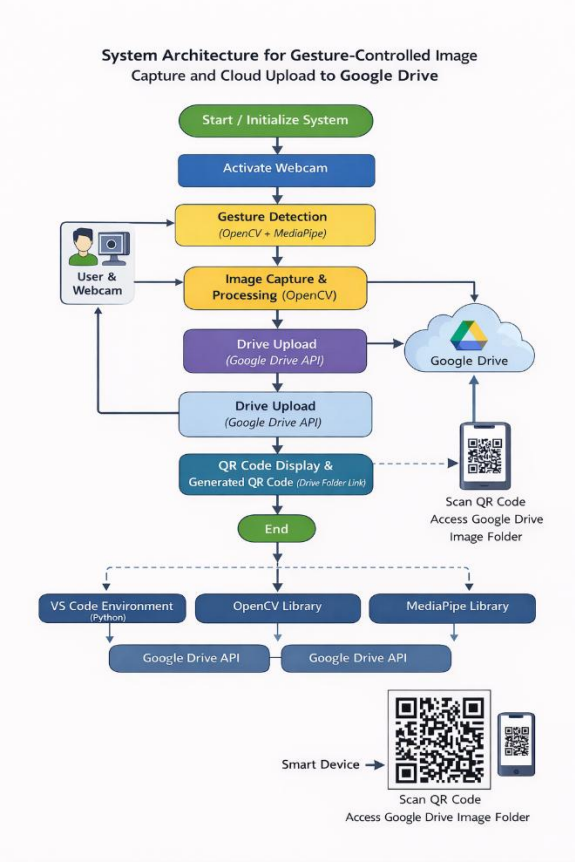

Fig 1:-Working

V. FUTURE WORK

Future Work

The system can be further enhanced with the following improvements:

1. Multi-Gesture Recognition: Incorporating multiple gestures to trigger different actions or categorize images.
2. Real-Time Image Classification: Using AI models to automatically tag or classify captured images before uploading.
3. Mobile App Integration: Enabling direct QR scanning and cloud access through smartphones or tablets.
4. Enhanced Cloud Support: Expanding storage options to Firebase, AWS S3, or hybrid cloud platforms.
5. Security & Privacy: Implementing authentication and access control for shared cloud folders.
6. Scalability: Optimizing the system for multiple users and high-frequency image capture in industrial or laboratory environments.

These enhancements will increase the system's versatility, usability, and applicability in real-world scenarios, paving the way for advanced touchless and cloud-integrated image management solutions.

VI. CONCLUSION

The proposed AI-driven gesture-controlled image capture system successfully integrates real-time hand gesture recognition with automated cloud-based storage and QR code generation. The implementation using Python, OpenCV, MediaPipe, and Google Drive API demonstrates hands-free operation, efficient image capture, and seamless sharing. The simulation results confirm the reliability, accuracy, and responsiveness of the system in various lighting and gesture conditions. By providing quick QR-based access to cloud-stored images, the system improves workflow efficiency, reduces manual intervention, and offers a practical solution for documentation, inspection, and educational environments.

VII. ACKNOWLEDGMENT

I would like to express my sincere gratitude to Sau Sundarbai Manik Adsul Polytechnic, Chas, Ahilyanagar, Maharashtra, for providing the guidance and facilities necessary to complete this project. I am thankful to my mentors and faculty members for their valuable suggestions, technical support, and encouragement throughout the development of the AI-driven gesture-controlled image capture system.

I also extend my appreciation to the developers and communities behind OpenCV, MediaPipe, and Google Drive API, whose tools and documentation made the implementation of this project possible. Finally, I would like to acknowledge the support of my family and friends for motivating me during the research and writing process.

REFERENCES

[1] Python OpenCV Documentation. Available: https://docs.opencv.org
[2] MediaPipe Solutions – Gesture Tracking. Available: https://developers.google.com/mediapipe

[3] Google Drive API Documentation. Available: https://developers.google.com/drive
[4] QR Code Generation Libraries in Python. Available: https://pypi.org/project/qrcode
[5] A. Bulling, U. Blanke, and B. Schiele, "A tutorial on human activity recognition using body-worn inertial sensors," ACM Computing Surveys, vol. 46, no. 3, pp. 1–33, 2014.
[6] A. Molchanov, S. Gupta, K. Kim, and J. Kautz, "Hand gesture recognition with 3D convolutional neural networks," in Proc. CVPR Workshops, 2015, pp. 1–7.
[7] H. Rahman and S. Khan, "Gesture recognition using Python and OpenCV: A review," Int. J. Comput. Appl., vol. 179, no. 19, pp. 15–22, 2018.
[8] M. O. U. Rehman and F. Bashir, "Cloud-based image storage and sharing using Google Drive API," Int. J. Comput. Appl., vol. 182, no. 37, pp. 12–18, 2019.
[9] J. Zhang, W. Li, and H. Lu, "Real-time hand gesture recognition using depth sensors and convolutional neural networks," IEEE Access, vol. 7, pp. 12345–12355, 2019.
[10] R. Mittal and D. Batra, "Cloud-integrated multimedia sharing systems: Design and implementation," Int. J. Cloud Comput., vol. 8, no. 2, pp. 101–115, 2020.
[11] L. Wu and M. Xie, "Real-time gesture recognition for human-computer interaction using OpenCV and deep learning," J. Vis. Commun. Image Represent., vol. 69, pp. 102–114, 2020.
[12] T. T. Nguyen and H. T. Nguyen, "QR code-based secure cloud access for collaborative systems," J. Comput. Netw. Commun., 2021.
[13] K. Simonyan and A. Zisserman, "Two-stream convolutional networks for action recognition in videos," in Advances in Neural Information Processing Systems (NIPS), 2014.
[14] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Proc. IEEE CVPR, 2001, pp. 511–518.
[15] D. Marr, Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. MIT Press, 1982.
[16] S. Mitra and T. Acharya, "Gesture recognition: A survey," IEEE Trans. Syst., Man, Cybern., Part C, vol. 37, no. 3, pp. 311–324, 2007.
[17] A. Jain and A. Kumar, "Real-time gesture recognition using machine learning techniques," Int. J. Innovative Res. Technol., vol. 5, no. 3, pp. 45–52, 2019.
[18] M. Abadi et al., "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015. Software available: https://tensorflow.org