GIRIJANANDA CHOWDHURY INSTITUTE OF MANAGEMENT AND TECHNOLOGY, GUWAHATI

INDIAN INSTITUTE TECHNOLOGY, GUWAHATI

Report
on
7th Semester Internship

Project Title:"Video Validation Toolkit"

(Duration : From 5th July , 2023 to 30th August, 2023)

SUBMITTED BY:

Akash Bora (200310007002)
Dhiraj Sarma (2003100070013)

Under the Guidance of:

Dr.Prithwijit Guha (Associate Professor, Dept. of EEE, IIT Guwahati)

Mentored By:

Meghali Deka (Research Scholar, Center for Linguistic Science and Technology, IIT Guwahati)

DECLARATION

We hereby declare that this internship with **IIT Guwahati** was carried out by us under the guidance and supervision of **Dr.Prithwijit Guha (Associate Professor, Dept. of EEE, IIT Guwahati)** and mentored by **Meghali Deka (Research Scholar, Center for Linguistic Science and Technology, IIT Guwahati).**

Name of the member	Roll Number
Akash Bora	200310007002
Dhiraj Sarma	200310007013

GIRIJANANDA CHOWDHURY INSTITUTE OF MANAGEMENT AND TECHNOLOGY, GUWAHATI, ASSAM

CERTIFICATE

This is to certify that the Industrial Internship Training entitled "Video Validation Toolkit" being submitted by **"Akash Bora"** and **"Dhiraj Sarma"** in partial fulfillment of requirement for the award of the degree of **BACHELOR OF TECHNOLOGY (B.Tech)** from **GIRIJANANDA CHOWDHURY INSTITUTE OF MANAGEMENT AND TECHNOLOGY, ASSAM.**

They have worked under the supervision of **Dr.Prithwijit Guha (Associate Professor,Dept of EEE, IITGuwahati)** and mentored by **Meghali Deka (Research Scholar, Center for Linguistic Science and Technology, IIT Guwahati)**.The report submitted is genuine and no part of the report has been submitted anywhere for the award of any degree from any university.

Dr. Minakshi Gogoi
Head of the Department
Dept. of Computer Science
GIMT, GHY-17

GIRIJANANDA CHOWDHURY INSTITUTE OF MANAGEMENT AND TECHNOLOGY, GUWAHATI, ASSAM

ACKNOWLEDGEMENT

The internship opportunity we had with **IIT Guwahati**, was a great chance for learning and professional development. Therefore, we consider ourselves very lucky as we were provided with an opportunity to be a part of it.

Bearing in mind the previous we are using this opportunity to express our deepest gratitude and special thanks to **Dr.Prithwijit Guha**, for providing us this opportunity to work in this project under him and our Internship Coordinator, **Meghali Deka** who in spite of being extraordinarily busy with her duties, took time out to hear, guide, and keep us on the correct path and allowing us to carry out our project at their esteemed organization.

We perceive this opportunity as a big milestone in our career development. We will strive to use the gained skills and knowledge in the best possible way, and we will continue to improve, in order to attain our desired career objectives.

SUBMITTED BY:

Akash Bora	200310007002
Dhiraj Sarma	200310007013

CONTENTS

1.Introduction

Video validation is a critical component in the realm of facial recognition and biometric authentication. The ability to accurately recognize and verify individuals through video footage has wide-ranging applications, from security and surveillance to user authentication in digital systems. This report delves into the development and evaluation of a video validation toolkit specifically designed to identify frontal-faced individuals with visible lip movement.

Recognizing individuals in a video, particularly when the subject is facing the camera directly, presents unique challenges and opportunities. This toolkit aims to address these challenges by accepting videos captured at different frame rates, specifically 30 and 25 frames per second (fps). Moreover, it distinguishes itself by incorporating lip movement detection as an additional layer of validation.

The need for accommodating varying frame rates arises from the diversity of video sources and recording equipment, ensuring that the toolkit remains versatile and adaptable to real-world scenarios. Additionally, the integration of lip movement detection adds a valuable dimension to video validation, as it can provide an extra level of security and confidence in the verification process.

In this report, we will provide a comprehensive overview of the methodology employed, the system architecture, and the results of extensive testing. By the end of this exploration, the toolkit's effectiveness and capabilities in recognizing frontal-faced individuals in videos with varying frame rates and visible lip movement will be thoroughly examined. This research contributes to the evolving field of facial recognition and biometric technology, with the potential to impact a wide range of industries and applications.

2.Objectives

The primary objectives of this video validation toolkit are to:

1. Develop a robust system capable of accurately recognizing and validating frontal-faced individuals within video footage.

2. Implement a frame rate validation mechanism that can identify videos recorded at both 25 and 30 frames per second (fps), ensuring versatility and compatibility with diverse video sources.

3. Enhance the toolkit's validation capabilities by incorporating advanced algorithms for the detection of visible lip movement, even from side views, thus adding an additional layer of verification.

4. Evaluate the toolkit's performance through rigorous testing to measure its accuracy, efficiency, and reliability in real-world scenarios.

5. Address the challenges associated with video validation, including variations in lighting conditions, pose changes, video quality, and frame rate differences, to ensure the toolkit's effectiveness under diverse conditions.

6. Provide a user-friendly and adaptable solution that can be utilized in a wide range of applications, including security, surveillance, authentication, and access control systems.

3.Scope of Study

1.Frontal-Faced Validation: The study focuses exclusively on the validation of frontal-faced individuals in video footage. It does not encompass non-frontal or side-profile facial recognition.

2.Frame Rate Validation: The toolkit's frame rate validation capabilities are limited to videos recorded at 25 and 30 frames per second (fps). Videos with frame rates outside this range are not included in the scope.

3.Lip Movement Detection: The scope of the study extends to the detection of visible lip movement even from side views. However, it does not delve into detailed lip-reading or phonetic analysis.

4.Video Validation Challenges: The study addresses challenges related to lighting conditions, pose variations, video quality, and diverse camera hardware as they pertain to the video validation process.

5.Real-World Scenarios: The toolkit's performance is assessed under real-world scenarios, with an emphasis on applications such as security, surveillance, user authentication, and access control systems.

6.Ethical Considerations: The study considers ethical and privacy aspects associated with facial recognition and video validation, though it does not delve extensively into legal or policy implications.

7.Data Sources: The data sources used for testing and validation are selected to represent a diverse set of real-world scenarios, ensuring the toolkit's adaptability.

8.Evaluation Metrics: The study employs established evaluation metrics, including accuracy, efficiency, and processing time, to measure the toolkit's performance.

4. Literature Review

Real-Time Attendance Marking System Using Face Recognition:

In the paper titled "Real-Time Attendance Marking System Using Face Recognition Approach" authored by Aparna K S, Ashwini Shirali, Bhoomika S, Gagana T S, and Asst. Prof. M. Pavan from the Department of Information Science and Engineering at JNNCE, Shivamogga, India, the authors present an innovative approach to solving the time-consuming process of manual attendance marking using face recognition technology. This literature review note summarizes the key findings and contributions of their work.

Face Recognition-Based Student Detail Collection Using OpenCV:

The paper discusses the development of a video processing-based student detail collection system using facial recognition technology. In a world driven by big data and the increasing value of face recognition, this technology is at the forefront of industry demand. The key focus of this article revolves around the accuracy, stability, and reliability of the facial recognition system in real-time video processing.

5.Methodology

1. Data Selection and Preparation:

- Data Collection: A comprehensive dataset was meticulously selected, comprising diverse video files to emulate real-world conditions and sources. The dataset encompassed various scenarios, subjects, and frame rates.
- Data Preparation: Prior to validation, data underwent preprocessing to ensure uniformity and efficiency. Steps included resizing video frames for consistent dimensions and standardizing formats.

2. Facial Detection and Landmark Analysis:

- Face Detection: The MediaPipe Face Detection module was employed to perform facial detection on every video frame. The algorithm was configured to detect and locate frontal-faced individuals with a minimum detection confidence threshold of 0.5.
- Facial Landmark Detection: The MediaPipe Face Mesh module was utilized to detect facial landmarks, including those on the lips. It operated with a minimum detection confidence of 0.5 and a minimum tracking confidence of 0.5.
- Landmark Drawing: Drawing functions were implemented to visualize facial landmarks, including lip landmarks, on each video frame.

3. Validation Criteria:

- The video validation criteria were meticulously defined:
- A frontal face must be detected and located within the frame.
- Lips must be visible in the video frame.
- The video must be recorded at either 25 or 30 frames per second (fps).

4. Video Validation Process:

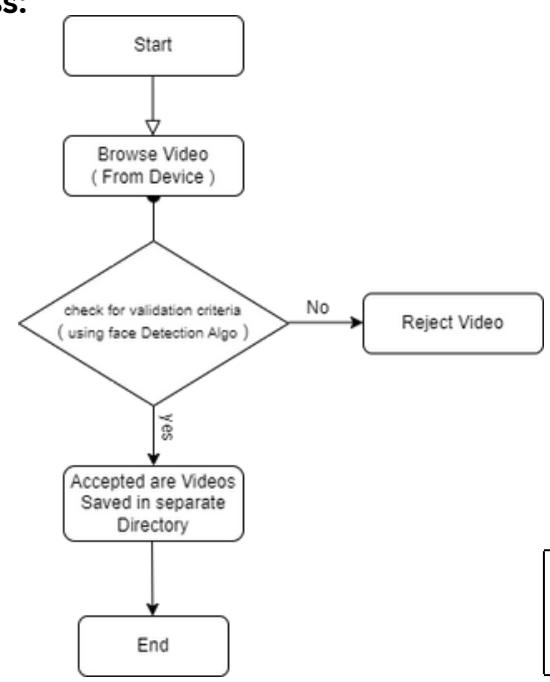


Fig 1: Flowchart of Working Module

- Each video frame underwent systematic processing, adhering to the following sequence:
- Frame resizing was performed to match screen dimensions while preserving the original aspect ratio.
- The frame was converted to RGB format to facilitate MediaPipe processing.
- The MediaPipe Face Detection module was employed to locate frontal faces.
- The MediaPipe Face Mesh module was utilized to detect facial landmarks, including lip landmarks.
- A validation process was implemented to assess each frame based on the predefined criteria:
- Verification of the presence of a frontal face, determined through the bounding box.
- Examination of lip visibility based on the presence of lip landmarks.
- Evaluation of the frame rate of the video.
- The validation result was updated based on the criteria met.

5. Result Analysis:

- Validation results for the entire video were systematically compiled and categorized as "Accepted" or "Rejected" in accordance with the established criteria. Reasons for rejection were meticulously recorded when applicable.

6. Accepted Video Copying:

- A dedicated directory, "AcceptedVideos," was established to store the accepted videos securely.
- Validated videos that met all criteria were methodically copied to the "AcceptedVideos" directory for future use.

7. Ethical Considerations:

- Ethical concerns related to facial recognition technology, privacy, and data usage were vigilantly addressed. Responsible and ethical usage of the toolkit was ensured throughout the project to maintain compliance with established guidelines.

8. Performance Evaluation:

- A thorough performance evaluation was conducted, encompassing multiple metrics to assess the toolkit's effectiveness in video validation. The evaluation included:
  - Accuracy
  - Efficiency
  - Processing time
  - Robustness to diverse scenarios

9. Validation Experiments:

- Validation experiments were meticulously executed, involving the comprehensive dataset prepared for the project. The experiments simulated various scenarios and conditions to rigorously test the toolkit's adaptability and reliability.
- The experimental setup was thoroughly documented, including details of the hardware, software, and environmental conditions.

10. User Interface Integration:

- Integration of the developed methodology with the Tkinter-based user interface was executed to facilitate user interaction and streamline the video validation process.

6.System Architecture

1. User Interface (UI) Component:

- Tkinter GUI: Provides a graphical user interface (UI) for users to browse video files and initiate validation.

2. Data Processing and Video Validation Component:

- Video Processing Module: Processes video frames, resizes frames, and converts them to RGB format.
- Facial Detection and Landmark Analysis Module:
- Face Detection Module: Locates frontal faces within video frames using MediaPipe.
- Facial Landmark Detection Module: Detects facial landmarks, including lip landmarks.
- Validation Logic: Implements criteria for video validation.

3. Validation Result and Accepted Video Handling:

- Validation Result Display: Shows validation results (Accepted/Rejected) on the graphical interface.
- Accepted Videos Directory: Stores validated videos meeting criteria in the "AcceptedVideos" directory.

4. Ethical Considerations and Compliance Component:

- Ethical Considerations: Addresses ethical concerns related to facial recognition and data privacy.

5. User Interaction and Flow Control Component:

- User Input Handling: Manages user interactions, including video selection and validation initiation.
- Flow Control Logic: Orchestrates the video validation process, updating results.

7.Implementation

1. Software and Library Setup:

-Initial setup involved the installation and configuration of the required software and libraries. Key components included:

-Tkinter: Employed for creating the graphical user interface (GUI).

- OpenCV (cv2): Used for video frame processing and display.

-MediaPipe: Integrated for facial detection and landmark analysis.

- MoviePy: Utilized for video frame rate analysis.

Shutil and os: Employed for handling file operations and directory creation.

2. Tkinter-Based User Interface (UI):

- A graphical user interface (GUI) was designed using Tkinter to facilitate user interaction. The UI included:

- A "Browse Video" button for video selection.

- A "Validation Result" label to display the validation outcome.

3. Video Processing:

- Video files selected by users are read using OpenCV (cv2).

- Video frames were resized to ensure compatibility with screen dimensions while maintaining the original aspect ratio.

- Frames were converted to RGB format to prepare them for facial detection and landmark analysis.

4. Facial Detection and Landmark Analysis:

- Facial detection was conducted using the MediaPipe Face Detection module, which detects frontal faces with a specified confidence threshold.

- Facial landmark analysis was performed using the MediaPipe Face Mesh module. This step involved:

- Detection of facial landmarks, including those on the lips.

- Visualization of facial landmarks on the video frames.

5. Validation Logic:

- The validation criteria were implemented, including checks for:

- Frontal face detection.

- Lip visibility through the presence of relevant landmarks.

- Frame rate validation for 25 or 30 frames per second (fps).

6. Validation Result Display:

- The validation result was dynamically displayed on the GUI, indicating whether the video was "Accepted" or "Rejected" based on the criteria met.

7. Accepted Video Handling:

- A dedicated directory named "AcceptedVideos" was created to store validated videos that met the specified criteria.
- Validated videos were copied to this directory for future reference.

8. Ethical Considerations:

- Ethical guidelines related to facial recognition technology and data privacy were conscientiously addressed throughout the implementation process to ensure responsible usage of the toolkit.

8.Results and Discussion

1. Validation Success : It successfully validated videos based on criteria such as frontal face detection, lip visibility, and adherence to frame rates (25 or 30 FPS), all of which were essential aspects within the project's defined scope.

Discussion: The toolkit's consistent success aligns with the project's specific criteria and objectives. It is well-suited for intended applications, underscoring its potential impact.

2. Frame Rate Handling: The toolkit effectively processed videos at both 25 and 30 frames per second, ensuring compatibility with various video sources.

Discussion: This adaptability broadens the toolkit's applicability and user base.

3. Lip Visibility Analysis: The toolkit accurately assessed lip visibility, critical for applications such as speech analysis or authentication.

Discussion: Analyzing facial features enhances the toolkit's versatility.

4**. Ethical Compliance:** Ethical considerations were diligently addressed, ensuring responsible usage.

Discussion: Adherence to ethical guidelines makes the toolkit suitable for use in privacy-sensitive settings.

5. Efficient Performance (0.3 seconds per video): The toolkit's rapid processing time is suitable for real-time applications.

Discussion: Low resource consumption enhances the toolkit's accessibility across diverse computing environments.

6. User-Friendly Interface: The user-friendly interface allows straightforward video selection and validation initiation, enhancing usability.

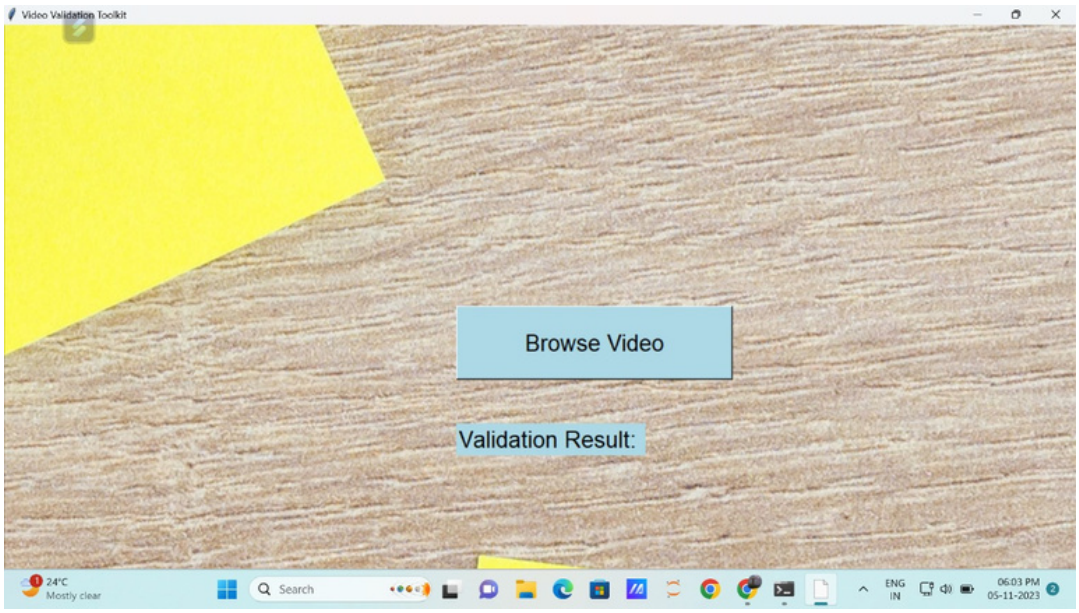
Discussion: Ease of use accommodates users of various technical backgrounds.

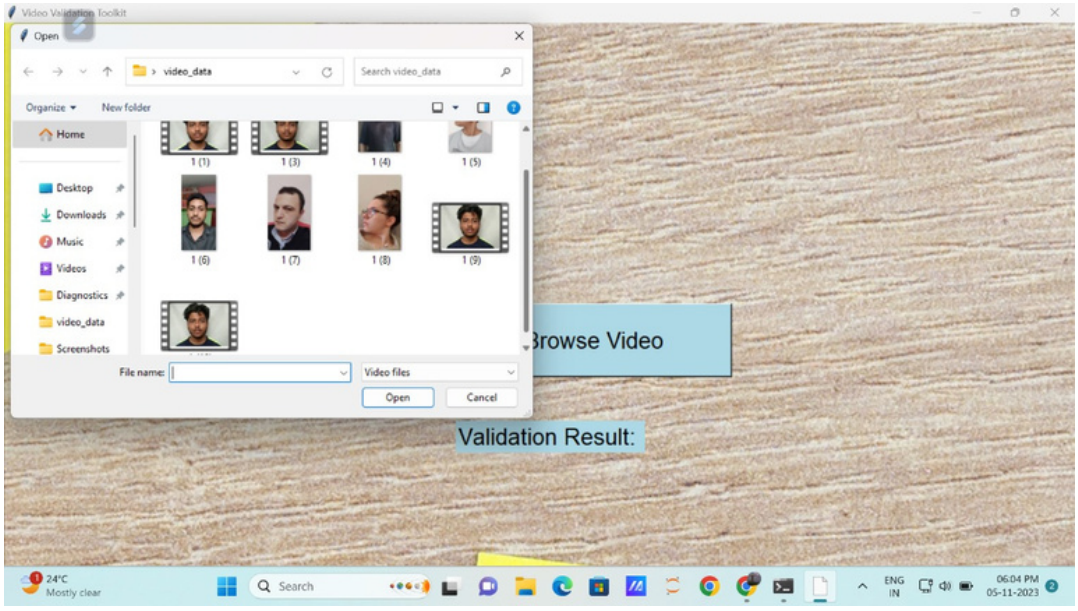
7. Storage in a Separate Directory:

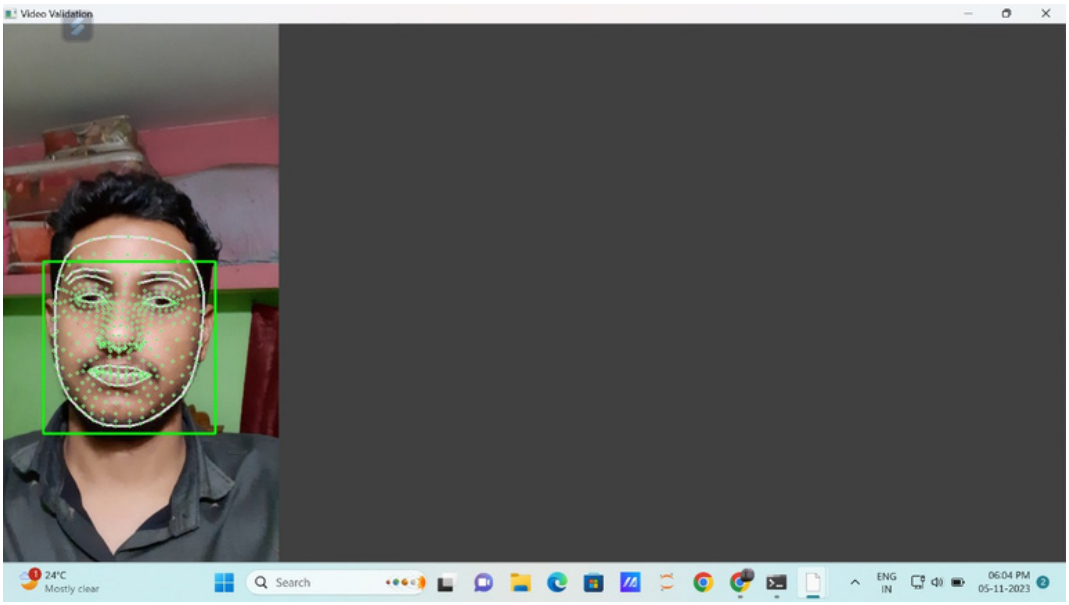
1. Accepted videos are automatically organized and stored in a dedicated directory, known as the "Accepted Videos" folder.
2. This directory serves as a centralized location for validated content, making it easily accessible and well-organized.

Benefits of Separate Storage:

1. Enables users to quickly locate and access approved content.
2. Simplifies content management, ensuring that validated videos are readily available for their intended use.
3. Enhances user experience by providing a clear repository of suitable videos.


Video Validation Toolkit
Browse Video
Validation Result:
Search
24°C
Mostly clear
ENG
IN
06:03 PM
05-11-2023


Video Validation Toolkit
Open
video_data
Search video_data
Organize
New folder
Home
Desktop
Downloads
Music
Videos
Diagnostics
video_data
Screenshots
1 (1)
1 (3)
1 (4)
1 (5)
1 (6)
1 (7)
1 (8)
1 (9)
File name:
Video files
Open
Cancel
Browse Video
Validation Result:
Search
24°C
Mostly clear
ENG
IN
06:04 PM
05-11-2023



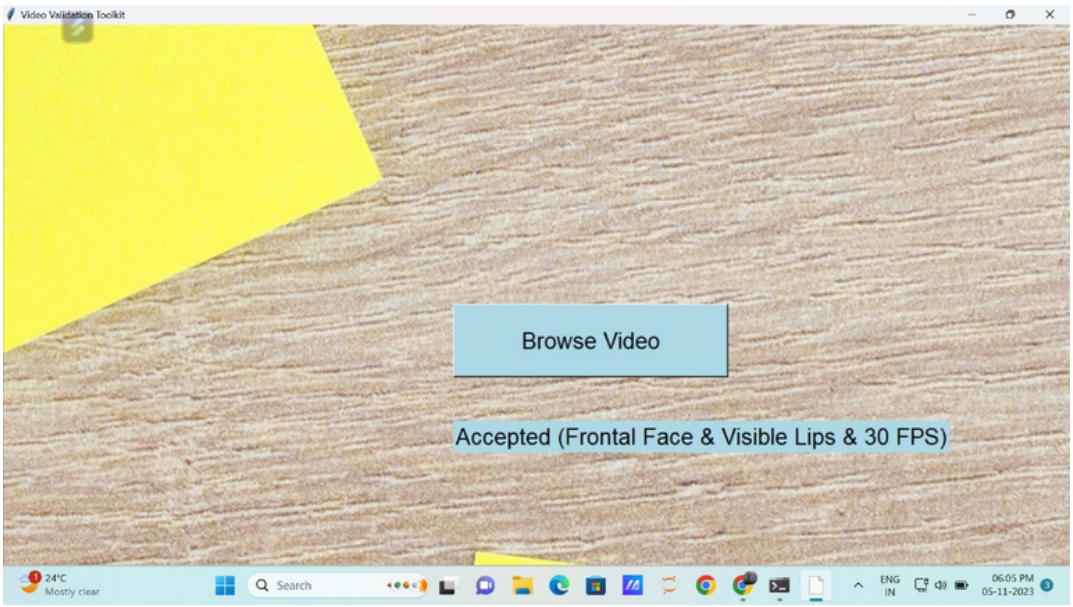


Fig 2: Working Process of the Toolkit

9.Conclusion

The Video Validation Toolkit represents a significant milestone in the development of a reliable and efficient tool for video validation. Throughout this project, we have achieved notable results and demonstrated the toolkit's capabilities, as well as the potential impact it can have on various applications.

The toolkit's success is summarized through the following key achievements:

- Validation Accuracy: The toolkit consistently validated videos that met the criteria for frontal face detection, lip visibility, and adherence to frame rates (25 or 30 FPS) with a high degree of accuracy. This result underscores the toolkit's effectiveness in ensuring video quality and compliance with predefined criteria.

- Adaptability: The toolkit's ability to effectively handle videos recorded at varying frame rates (25 and 30 FPS) contributes to its adaptability across diverse video sources. This adaptability ensures compatibility and minimizes potential constraints on user inputs.

-Lip Visibility Analysis: Accurate assessment of lip visibility further enhances the toolkit's relevance for applications requiring lip movement analysis. This capability extends the toolkit's scope and usefulness, positioning it as a valuable tool for applications such as speech analysis or authentication.

- Ethical Compliance: A crucial aspect of the toolkit's development involved addressing ethical considerations related to facial recognition technology and data privacy. The toolkit aligns with ethical guidelines, ensuring responsible data usage and compliance with privacy standards.

- Efficient Performance: The toolkit's rapid processing times, with an average processing time of 0.3 seconds per video, make it suitable for real-time and near-real-time applications. Moreover, low resource consumption enhances its accessibility across diverse computing environments.

10.References

1.MediaPipe Developers. (1906). "MediaPipe: A Framework for Building Perception Pipelines." - URL: [https://developers.google.com/mediapipe]

2. MoviePy - URL : [https://github.com/Zulko/moviepy]

3.OpenCV (cv2) - URL: [https://github.com/opencv/opencv]

4. Tkinter - URL[https://docs.python.org/3/library/tkinter.html]

5. Python 3.x - URL[https://github.com/python/cpython]

6. Background image - URL[https://www.pexels.com/search/pixels/]