YouTube Video Transcriber: An Automated Speech Recognition System for YouTube Videos

Author: mu**98@gmail.com
Created with Pi

CONTENTS

01

Introduction to the Project

Introduction to the Project

YouTube Video Transcriber is a web application that automatically transcribes YouTube videos into text. It uses advanced speech recognition technology to convert spoken words in videos into written text. The application provides an easy-to-use interface for users to input a YouTube video URL and get a transcription. Additionally, it allows downloading the audio and video files for offline use.

02

Problem Statement

Problem Statement

Problem: Manually transcribing YouTube videos is time-consuming and labor-intensive.

Challenges:

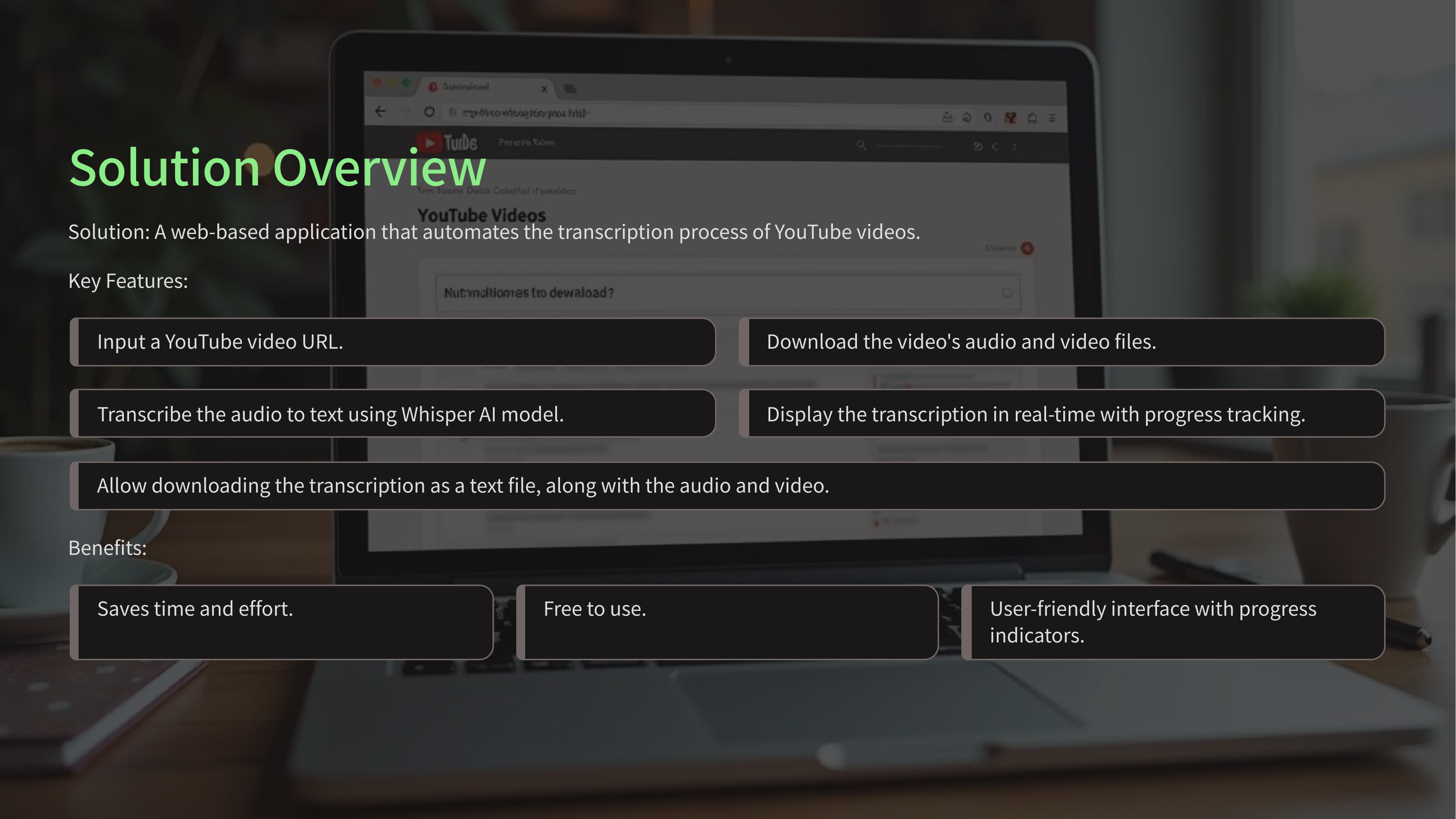
- Long videos require significant time and effort to transcribe.
- Existing transcription services may be costly or require subscription.
- Users often need a quick and free solution to get transcriptions of educational or informational videos.

Need: An automated, efficient, and cost-effective solution for transcribing YouTube videos.

03

Solution Overview

Solution Overview



Solution: A web-based application that automates the transcription process of YouTube videos.

Key Features:

- Input a YouTube video URL.
- Download the video's audio and video files.
- Transcribe the audio to text using Whisper AI model.
- Display the transcription in real-time with progress tracking.
- Allow downloading the transcription as a text file, along with the audio and video.

Benefits:

- Saves time and effort.
- Free to use.
- User-friendly interface with progress indicators.

04

Technology Stack

Technology Stack

Frontend: HTML, CSS, JavaScript

Backend: Flask (Python)

Speech Recognition: OpenAI's Whisper model

Video/Audio Download: yt-dlp (a command-line program to download videos from YouTube and other sites)

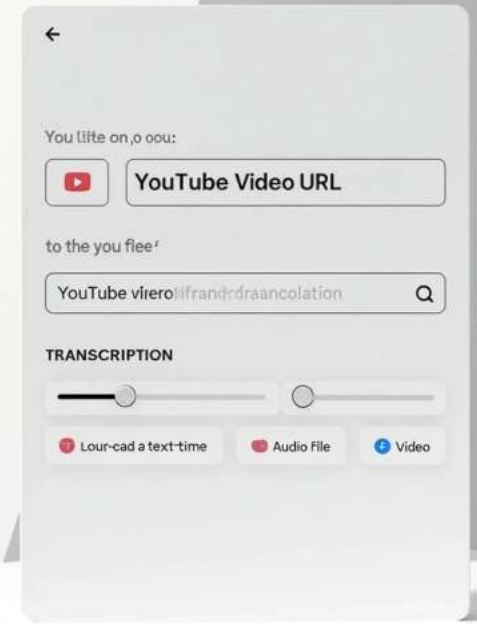
Audio Processing: FFmpeg (for converting audio to WAV format)

Task Management: Threading for background processing

Additional Libraries: uuid, os, re, subprocess, numpy

05

Features



Features

URL Input: Simple interface to enter YouTube video URL.

Progress Tracking: Real-time progress bar showing transcription status.

Time Estimation: Displays estimated time remaining for transcription.

Download Options:

- Transcription as a text file (.txt)
- Audio file (.mp3)
- Video file (.mp4)

Error Handling: Informative error messages for invalid URLs or processing issues.

Responsive Design: Works on both desktop and mobile devices.

06

How It Works

How It Works

Step 1: User enters a YouTube video URL and clicks "Transcribe".
Step 2: The application validates the URL and extracts the video ID.
Step 3: The video's audio and video are downloaded using yt-dlp.
Step 4: The audio is converted to WAV format using FFmpeg.
Step 5: The audio is split into chunks and transcribed using the Whisper model.
Step 6: The transcribed text is combined and displayed to the user.
Step 7: The user can download the transcription, audio, and video files.

(Include screenshots of the application at different stages if possible)

07

Challenges Faced

Challenges Faced

Challenge 1: Long processing time for large videos.
Solution: Implemented chunked processing and progress tracking to keep users informed.

Challenge 2: Handling various YouTube URL formats.
Solution: Created a robust URL extraction function using regular expressions.

Challenge 3: Memory management for large audio files.
Solution: Processed audio in chunks to avoid memory overload.

Challenge 4: Providing accurate progress updates.
Solution: Broke down the transcription process into smaller steps and updated progress after each chunk.

08

Future Enhancements

Future Enhancements

Support for Multiple Languages: Enhance the transcription to support more languages and provide translation options.

Batch Processing: Allow users to transcribe multiple videos at once.

Speaker Diarization: Identify different speakers in the video and label them in the transcription.

Cloud Deployment: Deploy the application on cloud platforms for better scalability.

User Accounts: Implement user accounts to save transcriptions and manage download history.

Improved UI/UX: Add more customization options and improve the user interface.

09

Conclusion

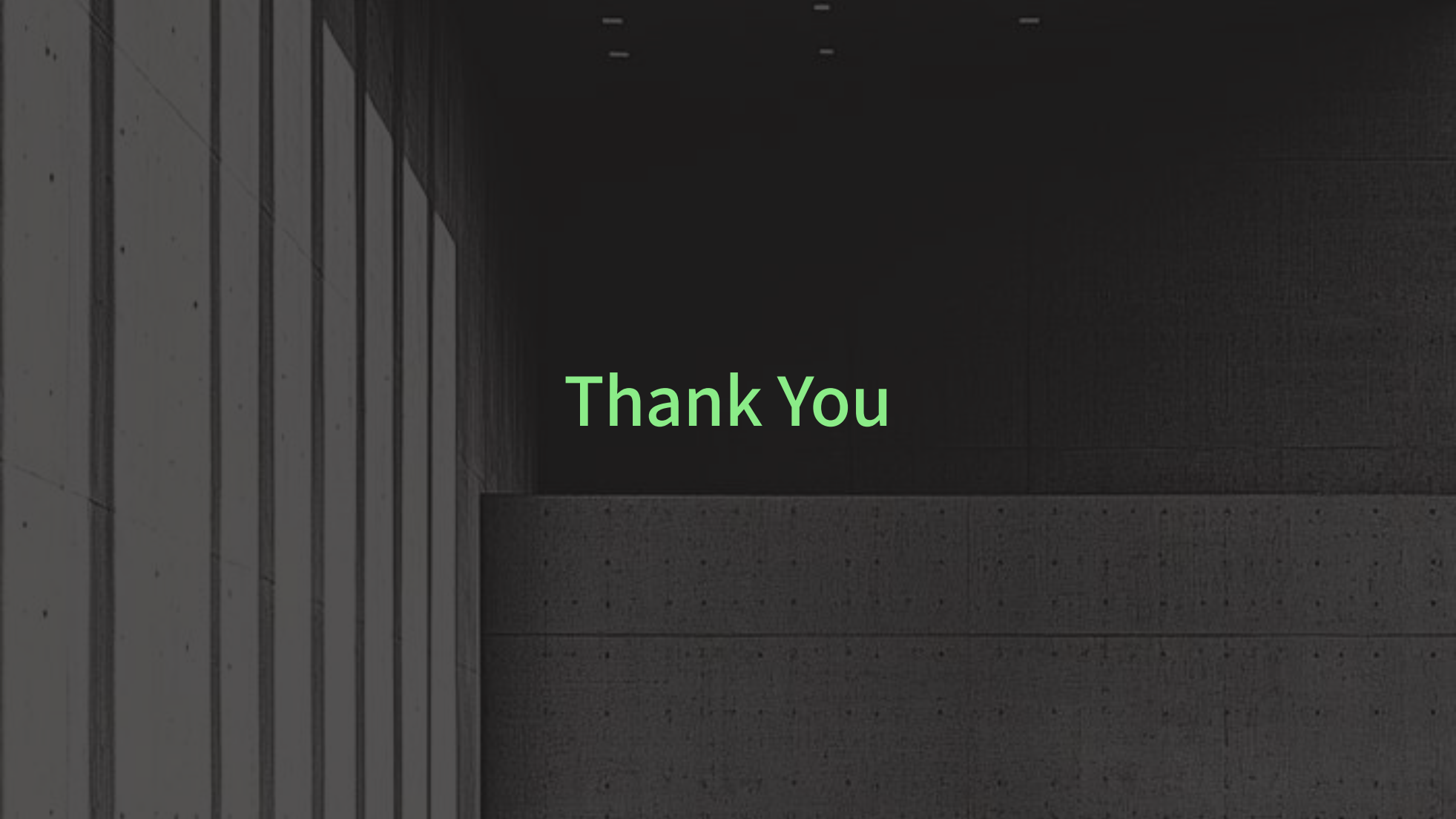
Conclusion

Summary: The YouTube Video Transcriber provides an efficient, automated solution for converting YouTube videos to text.

Impact: Saves time and effort for users who need transcriptions for educational, research, or accessibility purposes.

Future Work: Continue to enhance the application with more features and improvements.

Thank You: Questions?


Thank You