Project Report: Text & Image Summarizer

---

1. Introduction

In today's fast-paced world, we are flooded with vast amounts of textual information daily. Processing and comprehending this information quickly is essential but often challenging. The "Text & Image Summarizer" application aims to tackle this problem by offering users a streamlined, AI-powered solution to condense large amounts of text—whether entered manually or extracted from images—into concise summaries. Leveraging the power of Natural Language Processing (NLP) and Optical Character Recognition (OCR), this tool enhances productivity, aids decision-making, and promotes efficient knowledge consumption.

2. Project Objective

- To develop a user-friendly web application for summarizing large text inputs.

- To accurately extract and summarize text from uploaded images.

- To provide real-time, high-quality summaries without compromising context.

- To assist students, researchers, and professionals in handling information overload efficiently.

3. Problem Statement

In many professional and academic fields, users face massive volumes of text, including scanned documents, reports, and articles. Manual summarization is labor-intensive and time-consuming. There is a growing need for a tool that:

- Understands and summarizes large text passages.

- Extracts meaningful text from images with high accuracy.

- Simplifies the information extraction and comprehension process.

4. Innovative Features

- **Dual Input Modes**: Accepts both plain text and images containing text.

- **Integrated OCR**: Extracts text from image files using Tesseract OCR.

- **AI-Driven Summarization**: Utilizes state-of-the-art NLP models for accurate summarization.

- **Real-Time Processing**: Provides quick summarization results.

- **User-Friendly Interface**: Intuitive design catering to all user types.

- **Short Text Optimization**: Skips summarization for texts under a threshold to mair relevance.

5. Technology Stack

- **Streamlit**: Rapid web application development.

- **Tesseract OCR**: Reliable optical character recognition for images.

- **HuggingFace Transformers**: Advanced NLP library powering summarization.

- **facebook/bart-large-cnn**: Pre-trained deep learning model used for high-quality summaries.

- **PyTorch**: Backend framework for model training and inference.

6. Application Architecture

- **User Input Layer**: Accepts text or image input.

- **Processing Layer**: Applies OCR if input is an image, followed by NLP summarization.

- **Output Layer**: Displays summarized content.

- **Optimization Layer**: Bypasses summarization for short texts.

7. Workflow Overview

1. User selects the type of input (text/image).

2. User inputs/pastes text or uploads an image.

3. Application processes input and triggers summarization.

4. Summarized content is displayed instantly.

8. Project Structure

.

├── app.py   # Main application logic and user interface

├── ocr_module.py   # Handles text extraction from images

├── summarizer_module.py # Executes text summarization

9. Installation Guide

1. **Clone the Repository**:

git clone https://github.com/yourusername/your-repo-name.git

cd your-repo-name

2. **Install Required Libraries**:

pip install -r requirements.txt

3. **Install Tesseract OCR**:

- Download from the [official site](#).

4. **Run the Application**:

streamlit run app.py

10. How Summarization Works

The summarizer uses the "facebook/bart-large-cnn" model, a transformer-based sequence-to-sequence architecture. The model reads the input text, captures contextual relationships, and generates a shorter, coherent output without losing the main message.

11. Advantages of the System

- **Time Efficiency**: Summarizes in seconds.

- **OCR Capability**: Transforms image text into digital format.

- **Context Retention**: Summaries are meaningful and focused.

- **Simple Deployment**: Easy to install and run on any system.

- **Expandable**: Future upgrades for multiple languages and formats.

12. Future Enhancements

- **Multilingual Summarization**: Support for global languages.

- **Flexible Summary Formats**: Paragraphs, bullet points, or abstracts.

- **Enhanced OCR**: Better accuracy for complex documents.

- **User Profiles**: Personalized settings and history tracking.

- **Voice-to-Text Summarization**: Enabling audio input.

13. Real-World Applications

- Academic research simplification.

- Legal document summarization.

- Healthcare record summarization.

- News aggregation and summarization.

- Corporate executive summaries.

- Social media post and article compression.

14. Acknowledgements

- HuggingFace for providing pre-trained transformer models.

- Tesseract OCR for their reliable text extraction technology.

- Streamlit for simplifying machine learning web app development.

15. Conclusion

The "Text & Image Summarizer" demonstrates how merging OCR and NLP technologies can lead to powerful, real-world applications. It addresses the modern challenges of information overload, offering a practical tool that enhances comprehension, saves time, and supports informed decision-making. The project showcases a forward-thinking solution in the evolving landscape of AI-driven automation.

End of Report