# Image Summarization Tool Detailed Documentation

## Introduction

This document provides an in-depth overview of the Image Summarization Tool, which integrates Optical Character Recognition (OCR) and Natural Language Processing (NLP) technologies to process images and extract meaningful summaries.

## Approach and Techniques

The Image Summarization Tool utilizes a dual-technology approach, combining OCR and NLP to extract text from images and summarize it effectively. This is implemented through a FastAPI backend for API interactions and a Streamlit frontend for a user-friendly experience.

### Optical Character Recognition (OCR)

Tesseract OCR is employed to convert images containing printed or handwritten text into machine-encoded text. This is essential for digitizing text content in images.

Techniques and Configurations:

* - Image Preprocessing: Images are converted to grayscale and binarized using a threshold to enhance text clarity, improving OCR accuracy.
* - Tesseract Configuration: Configured to use both LSTM and legacy OCR engine modes (--oem 3) and set to Page Segmentation Mode 6.

### Natural Language Processing (NLP)

The BART model from Facebook, specifically trained for summarization tasks, is used to generate concise summaries from the extracted text.

Techniques:

* - Summarization Pipeline: Utilizes the Hugging Face transformers library to process the extracted text and produce a summary.

## Implementation Details

This section covers the functional aspects of both the FastAPI endpoint and the Streamlit frontend.

### FastAPI Endpoint

Functionality includes receiving image uploads, processing them to extract text, and generating text summaries. The endpoint offers robust error handling capabilities.

### Streamlit Frontend

Provides an interactive interface allowing users to upload images, view extracted text, and the generated summaries in real-time.

## Future Enhancements

Planned improvements aim to enhance OCR accuracy, summarization quality, user experience, and the tool's scalability and performance.

## Conclusion

The Image Summarization Tool demonstrates a practical application of combining OCR and NLP technologies to assist users in quickly understanding image content.