

Submission Summary

Conference Name

International Conference on Knowledge Engineering and Artificial Intelligence

Paper ID

301

Paper Title

Optical character recognizer Web Application using Tesseract

Abstract

The abstract introduces a web application for optical character recognition (OCR) that has been developed on Tesseract. Optical Character Recognition (OCR) is a technological advancement that facilitates the transformation of printed or handwritten text into a digital format, hence enabling efficient processing by computer systems. The online application under consideration employs Tesseract, an open-source optical character recognition (OCR) engine, to execute text extraction and recognition operations. The application offers users a user-friendly interface via which they may submit photographs or documents that contain textual content. The files that have been uploaded are thereafter subjected to processing by the Tesseract OCR engine. This engine utilizes advanced algorithms to evaluate and identify the characters that are contained within the photos. The retrieved text is subsequently identified and presented to the user. In addition to its core features, the online application has supplementary capabilities including the option to modify and store the identified text, choose the language for text identification, and fine-tune the image processing parameters to enhance precision. The primary objective of this project is to enhance the efficiency and precision of transforming physical text documents into a digital format. This will result in a reduction in the reliance on manual data entry and enable expedited retrieval of information. This web application, which utilizes Tesseract OCR technology, holds the potential to serve as a valuable tool across several businesses and areas that necessitate the management of substantial quantities of text-based data.

Created on

1/29/2024, 7:02:08 PM

Last Modified

1/29/2024, 7:02:08 PM

Authors

BHARATH SAI LINGAMPALLI (SATHYABAMA INSTITUTE OF SCIENCE AND TECHNOLOGY) < bharathsai2525@gmail.com > ✓

Lokesh Mandarapu (Sathyabama Institute of Science and Technology) < mandarapulokesh958@gmail.com > ✓

Submission Files

Optical character recognizer Web Application using Tesseract (1).pdf (153.1 Kb, 1/29/2024,

7:01:50 PM)
