Project Report

Webapp for Cognifront

Book Detection

by

Vinit Nitin Hingmire

(TE-IT 2022-23)

College Name : KK Wagh , College of Engineering , Nashik , Mahashtra

Mentor : Mr. Suchit Tiwari

Cognifront Software Private Limited

1 UMA APPARTMENTS, NEAR SATHE HOSPITAL OPP VED MANDIR, TRIMBAK ROAD, NASHIK Nashik MH 422002 IN.

कॉग्निफ्रंट
Cognifront

Contents

Internship Project Report

Problem Statement : Create a web application that will run in a mobile browser and uses a mobile phone camera and is able to identify Cognifront books if they are held in front of the camera. Books are listed on Cognifront website.

Introduction: During my one-month internship at Cognifront as a software developer intern, I had the opportunity to work on an exciting project focused on book recognition using cutting-edge technologies. Specifically, I worked on developing a web application that leverages object detection and text detection techniques to recognize and extract information from books. In addition to developing the backend functionality of the application, I also gained valuable experience in frontend development and database handling. Throughout the internship, I had the opportunity to work closely with experienced professionals in the field, learning new skills and techniques that will undoubtedly prove valuable in my future career as a software developer. Overall, my experience at Cognifront was both rewarding and educational, and I am grateful for the opportunity to have worked on such an exciting project.

Modules Used:

1. **OpenCV :** OpenCV (Open Source Computer Vision Library) is a powerful open-source computer vision library that provides a wide range of tools and functions for image and video processing, object detection, and machine learning.

   OpenCV supports a wide range of programming languages including C++, Python, and Java, and runs on multiple operating systems such as Windows, Linux, and MacOS. It provides a simple and user-friendly API for loading and manipulating images and videos, as well as advanced algorithms for feature detection, image filtering, and object recognition

2. **YOLO** : YOLO (You Only Look Once) is a deep learning algorithm that is widely used for object detection in computer vision. Unlike other object detection algorithms, YOLO looks at the entire image once and predicts the locations and classes of objects within the image. This makes YOLO faster and more efficient than other algorithms that require multiple passes through the image.

   YOLO works by dividing the input image into a grid and predicting bounding boxes around objects within each grid cell. Each bounding box is associated with a confidence score and a class probability distribution, which represents the likelihood of the box containing an object of a particular class.

3. **Pytesseract** : Pytesseract is a Python wrapper for Tesseract-OCR, an open-source optical character recognition engine. Pytesseract makes it easy to extract text from images and PDFs using Tesseract-OCR, which is one of the most accurate and widely used OCR engines.

Pytesseract supports multiple image formats such as JPEG, PNG, BMP, TIFF, and GIF, and can be used to extract text from images with different languages and scripts. Pytesseract also provides advanced options for preprocessing the image such as resizing, cropping, and applying various filters to improve OCR accuracy.

4. **MySQL** : MySQL is a powerful and versatile RDBMS that offers a wide range of features for managing and manipulating data. Its flexibility, scalability, and open-source nature have made it a popular choice for developers and organizations of all sizes.

5. **HTML:** Hypertext Markup Language is the standard markup language used for creating web pages and web applications. HTML provides a way to structure content and define its meaning, making it easier for web browsers to display content in a consistent and understandable way.

6. **CSS** : CSS (Cascading Style Sheets) is a stylesheet language used to describe the presentation of web pages and web applications written in HTML and XML. CSS allows developers to separate the presentation of a document from its content, making it easier to maintain and update the styling of a website.

7. **Flask Web Framework**: Flask is a micro web framework written in Python that is used for building web applications. Flask provides a lightweight and minimalistic approach to web development, allowing developers to focus on building the core functionality of their application without getting bogged down by boilerplate code. Flask also provides a wide range of extensions and plugins that can be used to add additional functionality to the framework

Application Development: The application is developed using Flask Web Framework. The Frontend of the web app is designed using the HTML and CSS. The basic requirement of web app is to recognize the book using the camera frames and after that extracting the text of the frame, for this purpose there was need of object detection and text detection module. As an object detection module I used OpenCV library with YOLO algorithm trained on COCO dataset and for text recognition purpose I used Tesseract-OCR and matched the text with the books registered in the database of the Cognifront website and if the text get matched I redirect to the books location in Cognifront website and also extracted the details from database.

Working : Working of the Application.

The Flask web application you developed as a software developer intern allows users to upload an image of a book cover, detect the book using object detection with OpenCV and YOLO v3, extract text from the book cover using Pytesseract, and search a SQL database of book titles to find a matching book. If a match is found, the user is redirected to the location of the book online.

The web application is built using the Flask framework, which provides the structure and functionality needed to create a web application. The frontend of the application is designed using HTML and CSS, with Bootstrap used to style the page and make it responsive. The web application allows users to upload an image of a book cover, which is then processed using OpenCV and YOLO v3 to detect the book within the image. The detected book is then cropped and the title text is extracted using Pytesseract.

Once the title text has been extracted, the application searches a SQL database of book titles for a matching book. If a match is found, the user is redirected to the location of the book online. If no match is found, the user is notified that the book was not found.

The SQL database is created and managed using MySQL, with Python's Flask-MySQLdb library used to interact with the database. The database contains a list of book titles and their corresponding online locations. When a user uploads an image of a book cover and the title text is extracted, a SQL query is used to search the database for a matching book title. If a match is found, the user is redirected to the online location of the book and details from database is extracted.

Project Management : During the project, I followed the Agile methodology, which helped me to manage the project efficiently. I broke down the project into smaller tasks and set deadlines for each task. I also created a backlog of tasks and prioritized them based on their importance.

I used to do google meetings with my mentor Mr.Suchit Tiwari sir to ensure that the code was stable and met the requirements.

Training and Development : During the internship, I received adequate training and support from my mentor. My mentor provided me with the required resources and guidance to develop the application. I also received feedback on my work, which helped me to improve my coding skills.

The internship provided me with an opportunity to work with various Python modules and to understand their functionalities. Throughout the internship, I had the opportunity to work closely with professionals in the field, learning new skills and techniques that will undoubtedly prove valuable in my future career as a software developer. Overall, my experience at Cognifront was both rewarding and educational, and I am grateful for the opportunity to have worked on such an exciting project.I learned how to use modules such as flask, pytesseract, mysql and OpenCV,YOLO which were new to me.

Impact of Project : By utilizing object detection and text detection technologies, the application can accurately identify and extract text from images of books, making it easier to digitize and manage large collections of books and documents.

One of the main impacts of this project is increased efficiency and accuracy in the digitization of books and documents. The use of object detection and text detection technologies reduces the amount of manual effort required to digitize books and documents, leading to faster turnaround times and more accurate results. This can be particularly beneficial for organizations with large collections of books and documents, such as libraries and archives.

Another impact of this project is improved accessibility to books and documents. By digitizing books and making them available online, people who might not have access to physical copies can still benefit from the information contained within them. This can be particularly beneficial for researchers and students who need access to a wide variety of books and documents for their work.

The frontend development aspect of the project also has an impact on user experience. By creating a user-friendly interface, users can easily upload and process images of books, making the application more accessible to users who may not have a technical background.

Finally, the database handling aspect of the project has an impact on data management and storage. By creating a database to store the digitized books and associated metadata, users can easily search for and access the information they need, leading to more efficient and effective document management. This can be particularly beneficial for organizations with large collections of books and documents that need to be organized and searchable.

Overall, the book recognition web application has the potential to make a significant impact in the field of document digitization and management, improving efficiency, accessibility,

and user experience.

Limitations : Although the project has been developed successfully and works efficiently, there are still some limitations to the application. The application requires a pytesseract and yolo files in the system, Also good internet connection and Camera Quality is also required for successful implementation.

Future Scope :

1.Expansion to other languages: Currently, the text recognition feature of the application only works with English language text. A future scope of the project could be to expand this feature to support other languages, making the application more versatile and accessible to a wider range of users.

2.Improved accuracy: While the current implementation of the application is accurate, there is always room for improvement. The accuracy of the book detection and text recognition features could be further improved through fine-tuning of the models and algorithms used, and by incorporating other technologies such as natural language processing

3.Support for different file formats: The current implementation of the application only supports image file formats. A future scope of the project could be to support other file formats such as PDF, EPUB, or Microsoft Word documents, making the application more versatile and useful for document digitization and management.

Conclusion:

In conclusion, as a software developer intern, I developed a Flask web application for book recognition, utilizing object detection, text recognition, frontend development, and database handling technologies. The application allows users to upload an image of a book cover, detect the book using OpenCV and YOLO v3, extract text from the book cover using Pytesseract, and search a SQL database of book titles to find a matching book. If a match is found, the user is redirected to the location of the book online.

The project has a significant impact, as it can streamline the process of finding online versions of books by digitizing book covers and titles. Moreover, the project has significant potential for future improvements, such as expanding the application to support other languages, improving accuracy, integrating with e-book platforms, supporting different file formats, and developing a mobile application version.

Overall, the project has been an excellent opportunity for me to develop my skills as a software developer, learn new technologies, and gain experience in developing a web application from start to finish. The project has also taught me the importance of designing an intuitive and user-friendly interface and optimizing the application for speed and efficiency.