C. V. Raman Global University
ODISHA BHUBANESWAR INDIA

Text Extraction from Image using Machine Learning

Under the Guidance of

Dr. Mamata P. Wagh

Anurag Kumar-2101020505

Taniya Anshu- 2101020507

Sajan Kumar- 2101020509

Outlines

- Introduction
- Literature Review
- Problem Statement
- Proposed solution
- Results
- Challenges
- Conclusion
- References

Introduction

- Scene text extraction involves detecting and extracting text from images like street signs, posters, etc.
- It is difficult due to complex backgrounds, lighting variations,, and noise in the images.
- There are some model designed to accurately find and highlight text in images and to overcome the challenges.
- Focuses on segmenting text from images using various methods of Machine learning to improve the accuracy of text.

Literature Review

Title	Author	Year	Advantages	Limitations
Accurate, Data-Efficient, Unconstrained Text Recognition with Convolutional Neural Networks	Mohamed Yousef, Khaled F. Hussain, Usama S. Mohammed	2020	It worked very well on both short and long lines of text and also can handle different handwriting styles and sizes.	Existing line recognition methods can't handle paragraphs or multiple lines without line segmentation algorithms.
EPAN: Effective parts attention network for scene text recognition	Yunlong Huang, Zenghui Sun, Lianwen Jin, Canjie Luo	2020	EPAN needs less detailed instructions during its training which is efficient.	The two-stage attention mechanism in EPAN takes more time to train.

Literature Review

Title	Author	Year	Advantages	Limitations
Intelligent character recognition using fully convolutional neural networks	Raymond Ptucha , Felipe Petroski Such , Suhas Pillai, Frank Brockler, Vatsala Singh, Paul Hutkowski	2019	It predicts both arbitrary symbols as well as words from a lexicon. And Avoids dependency on recurrent neural networks and CTC tuning.	The unconventional use of even convolution filters may limit adoption, and handling larger lexicons requires further investigation.
End-to-End Historical Handwritten Ethiopic Text Recognition Using Deep Learning	Ruchika Malhotra, Maru Tesfaye Addis	2023	The paper is one of the first to focus on recognizing historical handwritten Ethiopic text, helping preserve ancient documents. It uses deep learning to effectively handle complex scripts and enhance text recognition.	The model still struggles with high error rates in recognizing complex handwritten texts. It heavily relies on large datasets, which may not perform well on unseen or diverse text styles.

Literature Review

Title	Author	Year	Advantages	Limitations
Unambiguous Scene Text Segmentation With Referring Expression Comprehension	Zejian Yuan, Chunhua Shen, and Luc Van Gool	2020	The paper combines visual and language features to accurately detect and segment text in busy, cluttered images, improving text recognition.	It process the input text sequentially which slow down the training time as it enables the text to go under process one by one.
Distance Transform based Text-line Extraction from Unconstrained Hand-written Document Images	Suman Kumar Bera, Soumyadeep Kundu, Neeraj Kumar, Ram Sarkar	2021	Placing paragraphs in a one-page document has always yielded impressive results in most cases.	In multi-page documents, if there's a line between paragraphs, the method might mistake it for a new line.

Problem Statement

- Blurred and low-quality images make accurate text extraction difficult.
- Complex backgrounds and noise affect character segmentation and clarity.
- OCR(Optical Character Recognition) struggles with varied orientations, font styles, and layouts.
- Preprocessing limitations reduce accuracy in noisy, real-world conditions.

Proposed Solution

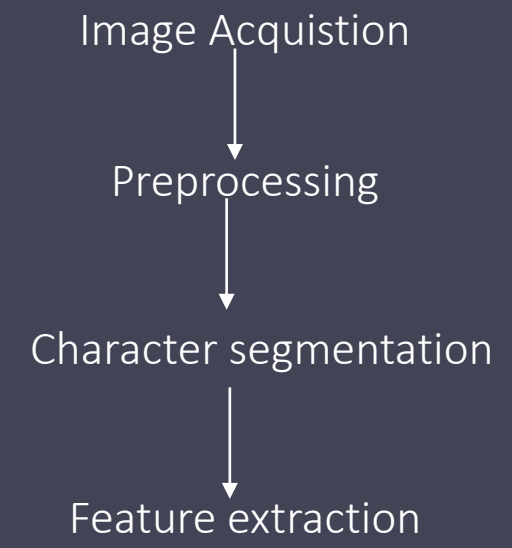
SVM (Support Vector Machine):

- SVM is trained to differentiate between text regions and non-text regions in an image.
- It classifies parts of the image based on features like texture, edges, and color patterns.
- SVM uses these features to identify which regions are likely to contain text.
- SVM can be used to classify individual characters after text regions are detected.

Dataset Description:

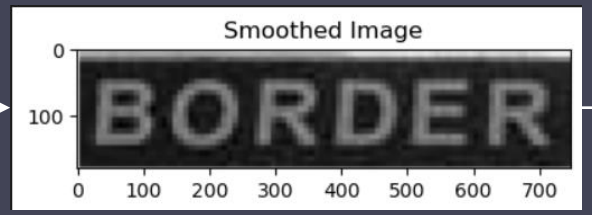
- Dataset Name: IIIT5K-Word
- Number of Samples: 5,000 images in total.

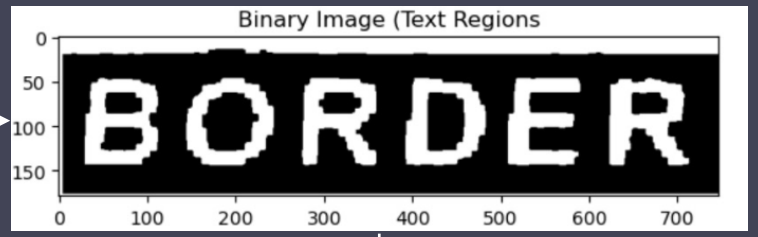
Implementation Sequence



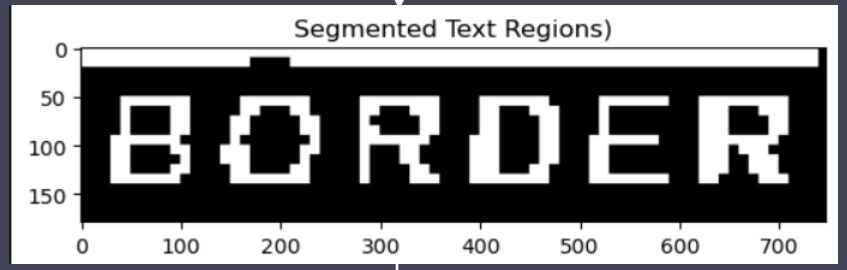
Result:

Flow Chart of model with an example





SVM(Support Vector Machine)



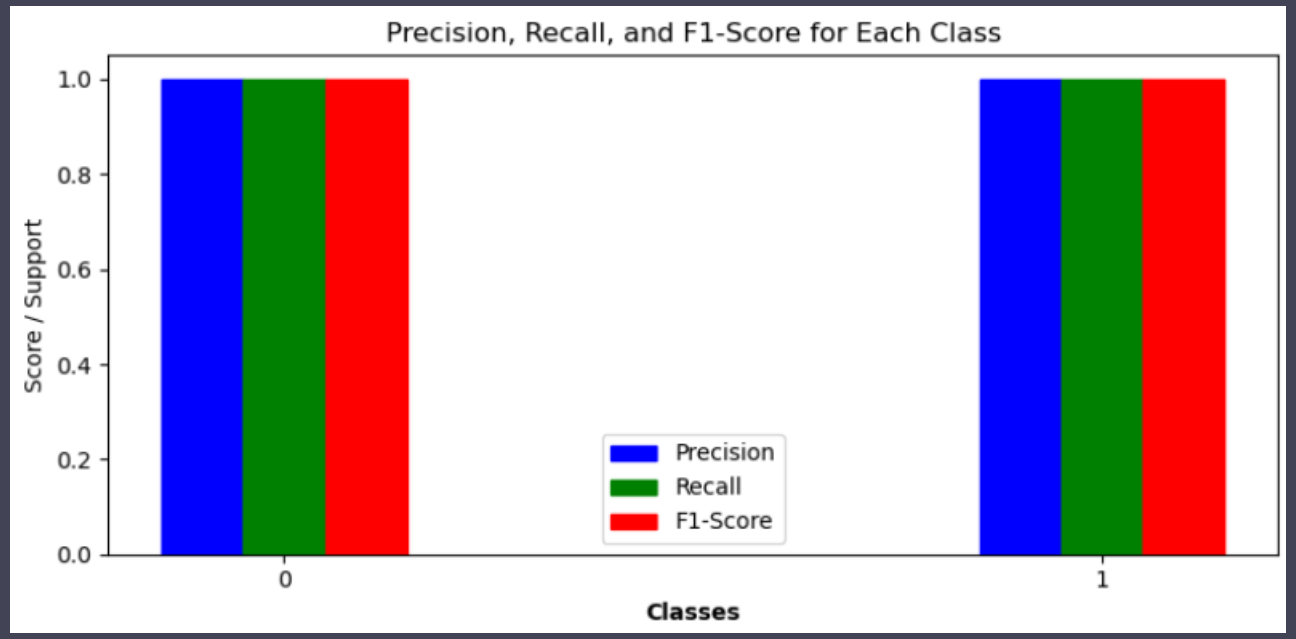
Extract text using Tesseract OCR(Optical Character Recognition)

```
Extracted Text:
 BORDER
```

Result:

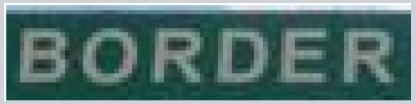
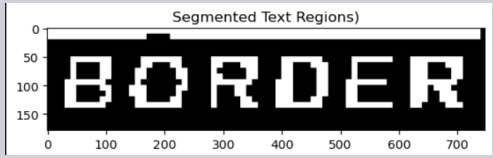
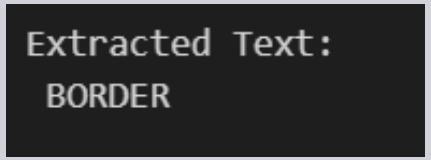
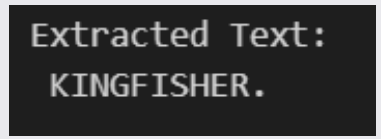
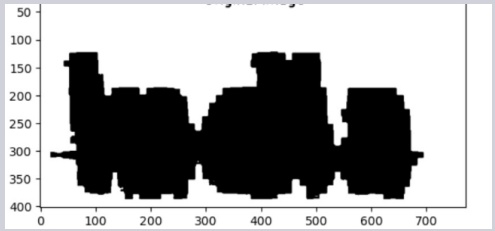
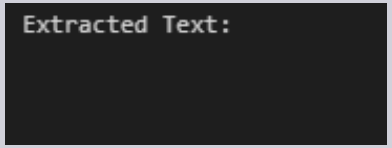
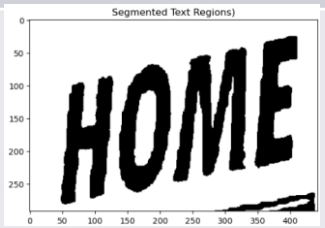
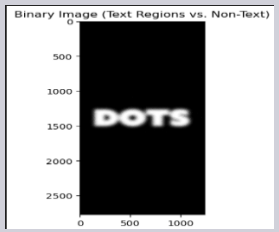
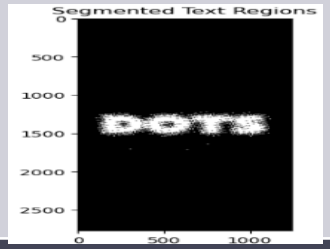
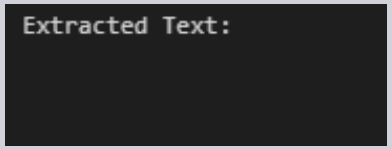
```
Accuracy Score: 1.0
Classification Report:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00       175
           1       1.00      1.00      1.00        77

    accuracy                           1.00       252
   macro avg       1.00      1.00      1.00       252
weighted avg       1.00      1.00      1.00       252
```

Evaluation of Model Accuracy



Bar Chart of Confusion Matrix

Predictions Obtained on dataset IIT5K-WORD

Input Image	Label	Segmented Image	Extracted Output
	BORDER		
	KINGFISHER		
	India		
	HOME		
	DOTS		

Challenges

•The model fails to precisely segment characters in some of the blurred images such as string “DOTS”.

•Inaccurate segmentation affects the ability to isolate individual characters.

•Inconsistent segmentation leads to merged or fragmented characters, making OCR recognition difficult.

•This impacts the extraction of specific strings, where characters are not accurately identified.

Conclusion

- Developed a text extraction solution integrating OCR with SVM-based classification.
- Improved text recognition through image preprocessing and segmentation techniques.
- Reliable performance on printed text across varied backgrounds and orientations.
- Expand the solution to handle multilingual or stylized text.

References

[1] Ptucha, R., Petroski Such, F., Pillai, S., Brockler, F., Singh, V., & Hutkowski, P. (2019). Intelligent character recognition using fully convolutional neural networks. Pattern Recognition, 88, 604–613. doi:10.1016/j.patcog.2018.12.017

[2] Suman Kumar Bera; Soumyadeep Kundu; Neeraj Kumar; Ram Sarkar; (2021). Distance transform based text-line extraction from unconstrained handwritten document images . Expert Systems with Applications, (), –. doi:10.1016/j.eswa.2021.115666

[3] Yousef, Mohamed; Hussain, Khaled F.; Mohammed, Usama S. (2020). Accurate, Data-Efficient, Unconstrained Text Recognition with Convolutional Neural Networks. Pattern Recognition, (), 107482–. doi:10.1016/j.patcog.2020.107482

[4] Y Huang, Z Sun, L Jin, C Luo EPAN: Effective parts attention network for scene text recognition 2020

[5] R Malhotra, MT Addis End-to-end historical handwritten ethiopic text recognition using deep learning - IEEE , 2023

[6] SX Zhang, X Zhu, JB Hou, C Yang Kernel proposal network for arbitrary shape text detection, 2022