FOUR WEEK REPORT

TITLE: - Generating Synthetic Scene Text Word Images for Indic Scripts

NAME: - Raunak Nag (ENGS4432)
NAME OF THE GUIDE: - Prof CV Jawahar

ABSTRACT:- Text recognition has been an active research area for real applications. Optical Character Recognition (OCR) is a technique to convert printed or handwritten text into machine-readable form. While text in natural scene images fall into the category of Scene Text Recognition. Scene text recognition has become an interesting field of research due to several complexities - (i) complex backgrounds, (ii) improper illumination, and (iii) distorted images with various noises, etc. Latin languages were found to be the center of attention till now and the field of scene text recognition has not been investigated for non-Latin languages. Scene text recognition in low-resource non-Latin languages is challenging due to the inherent complex scripts, multiple writing systems, various fonts, and orientations.

INTRODUCTION:-

Problem definition:- Current Scene Text Recognition models rely on datasets that mostly contain Latin/English text. But there has also been a prominence of non-Latin languages like Chinese and Indian languages, etc. Hence, it is important to construct recognition models for these non-Latin languages. Huge amount of language specific real scene text data is needed to build data centric OCR models which unfortunately is scarce, time-consuming and the annotation process is often prone to human error and only available abundantly for the English/Latin languages. One solution to the data scarcity issue is to generate high-quality realistic synthetic data which is cheap and scalable.

The objective is to generate high-quality synthetic word images that mimic the characteristics of real-world Indic scene text images. These images will serve as training for an OCR system, enabling to accurately recognize and transcribe text in Indian scripts from natural scene images.

EXPERIMENT:-

Our focus and approach would be to create a large diverse dataset of different regional Indian font scripts rendered on different natural scene images by developing an algorithm. The developed algorithm has the following functions -

- ➔ DistortArcOptions
- ➔ DensityOptions
- ➔ FontSizeOptions
- ➔ FontStretchOptions
- ➔ ShadowWidthOptions
- ➔ ShadowSigmaOptions
- ➔ ShadowOpacityOptions
- ➔ ShadowWidthSignOptions

Inputs:-

1. Indic language vocabulary word file
2. List of unique fonts for a language
3. Directory for the rendered images
4. File with the rendering language details
5. Iteration number

RESULTS:-

Sample synthetic images for Bengali language.



REFERENCES:-

- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9025185/pdf/jimaging-08-00086.pdf
- https://arxiv.org/pdf/1608.04224.pdf
- https://lear.inrialpes.fr/people/alahari/papers/mishra16-thesis.pdf