Summer Internship Presentation

Name- Raunak Nag
Er. No- 201B374
Course: BTECH
Branch: CSE
Semester: 7TH
Batch: 2020-2024

INTRODUCTION

- **Name of Institute:** International Institute of Information Technology, Hyderabad (IIIT-H)
- **Project:** Generating Synthetic Word Images for training Indian Scripts Scene Text OCR
- **Corporate Guide:** Prof CV Jawahar
- **Faculty Guide:** Ajoy Mondal
- **Period of Internship:** 2 Months (12-Jun-2023 - 7-Aug-2023)
- **Designation:** SRFP CVIT Intern

OBJECTIVE

* The Presentation is prepared based on my Two-month internship at CVIT Department in IIIT Hyderabad, Telangana, India.
* The project is based in the field of Computer Vision in Artificial Intelligence and Machine Learning.
* My task mainly included developing and modifying an existing python script to generate a synthetic word images dataset for Indian regional scripts especially for Bengali language.
* This presentation displays the daily progress of my weekly tasks that I have undergone during my internship duration.

ABOUT THE INSTITUTE

* **Institute Name: -** International Institute of Information Technology, Hyderabad, India
* **Founder**: -Narendra Ahuja
* **Institute Type:** - Engineering
* **Location**: - Gachibowli, Hyderabad, Telangana, India
* **Type:** -Public-Private Institute
* **Founded:** -1998
* **Mission: -** To contribute to transforming industry and society, by delivering research led education, promoting innovation, and fostering human values.

PROJECT INTRODUCTION

* OCR for Indian scripts scene text is a difficult task due to the complexity of the scripts and the lack of training data
* Synthetic data generation allows us to quickly generate large amounts of training data that can be used to train OCR models.
* With this, we can improve the accuracy of current OCR models used for scene text recognition and can lead to better performance in results.
* Moreover, we can also add different types of variations in synthetic data which may not be possible sometimes in real data

CHALLENGES

* The different shapes and sizes of characters, combined with the presence of ligatures and diacritics, make it difficult for OCR models to accurately recognize and transcribe the text in Indian languages
* There is also a lack of training data available for Indian scripts scene text OCR which makes it harder to develop accurate OCR models that can handle the diversity of the Indian scripts

IMPORTANCE

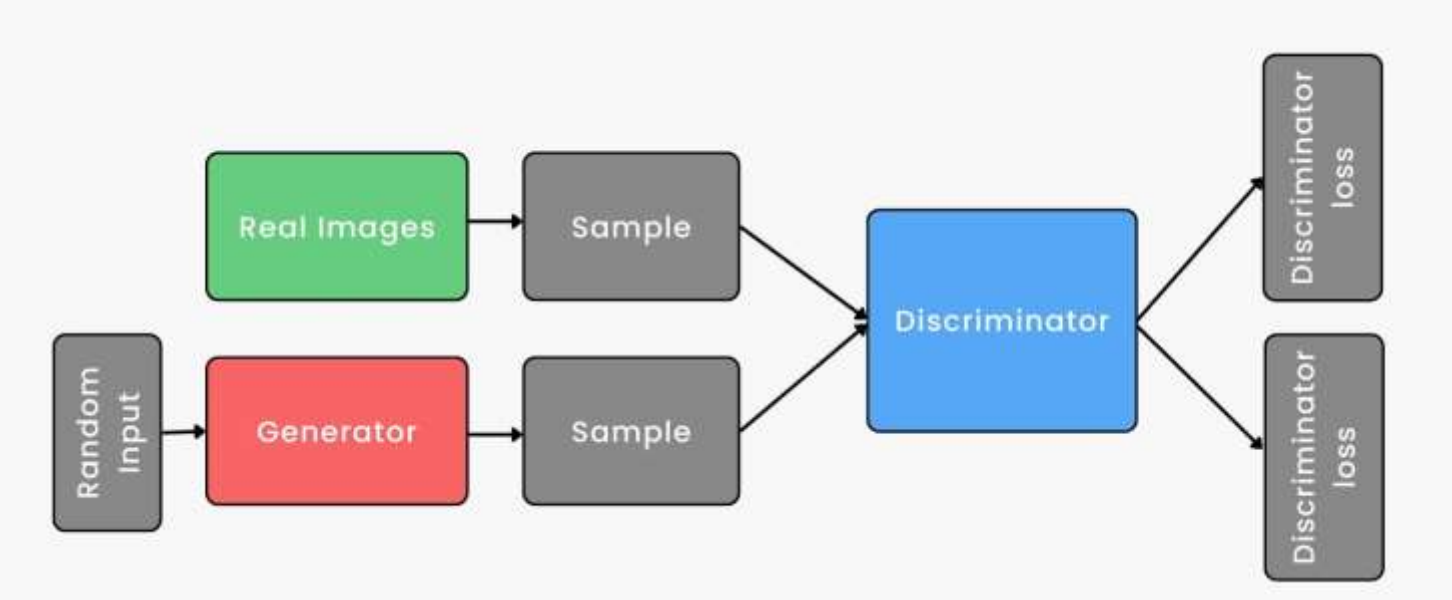
* Synthetic data plays a crucial role in training OCR models for Indian scripts scene text. It provides a cost-effective and scalable solution to overcome the challenges posed by the scarcity of real-world data
* By generating synthetic word images, we can create a diverse and comprehensive dataset that covers all possible variations of Indian scripts scene text
* Synthetic data allows us to augment our existing datasets and improve the performance of our OCR models

APPLICATIONS

EXAMPLES OF REAL SCENE TEXT

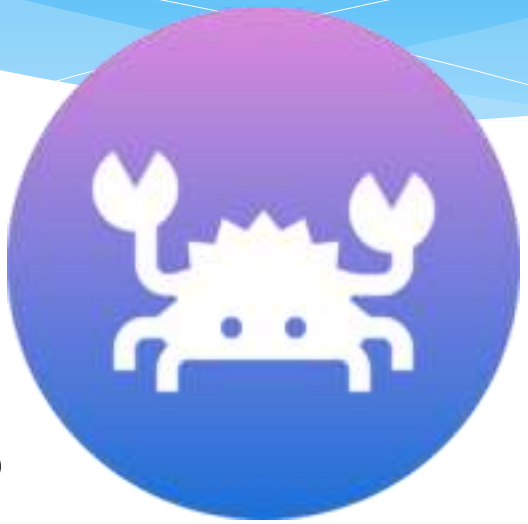
EXAMPLES OF SYNTHETIC TEXT

PROCESS


Real Images
Sample
Random Input
Generator
Sample
Discriminator
Discriminator loss
Discriminator loss

PRE-REQUISITES

ImageMagick



PangoCairo

Cairo

Pango

INPUTS

* Indian Language
* Vocabulary File
* Language Fonts

PROCEDURE

Steps involved in Generating Synthetic Data

* Data Preparation and Data Collection
* Reading the vocabulary file
* Choosing various rendering parameters
* Image processing
* Rendering Image

EXAMPLES OF GENERATED BENGALI SYNTHETIC SCENE TEXT

LEARNINGS

* Understanding what is synthetic data and why it is required
* Understanding about scene text and its different variations in Indian scripts
* To know and understand different ways of synthetic data generation.
* To learn how to read research papers and research articles

CONCLUSION

* Generating synthetic word images for training Indian scripts scene text OCR is a crucial step towards improving the accuracy of OCR models
* Data augmentation techniques can also be used to further enhance the synthetic dataset.
* Through 8 weeks of internship, I have grown rapidly in a professional explored research in computer science
* I was able to interact with my peers and fellows which also helped me enhance my communication skills

THANKYOU