Novel Method for Handwritten Digit Recognition System

Project Mentor :
Suganthi. J

Team id: PNT2022TMID10802

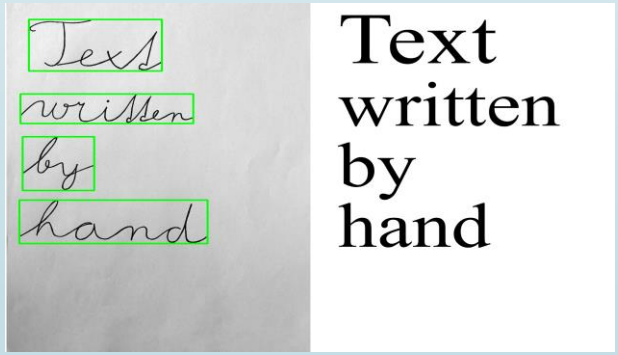
Team Members :
Gnanadesigan.V
Vimalraj.S
Sanjay.K
Praveen Raj.B

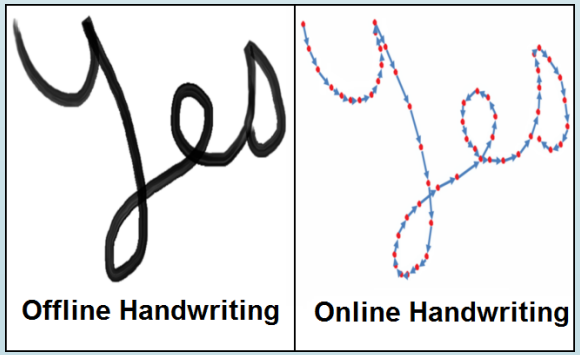
Parameters:

- Problem Statement
- Survey Papers
- Existing System
- Hardware & Software Requirements
- Objective

Problem Statement

- The process of converting handwritten text into machine-readable form is known as handwritten character recognition (HCR). The diversity of handwriting styles in the HCR system, which might be radically different for different writers, is the main issue.
- The goal of a handwritten character recognition system is to implement an easy-to-use computer assisted character representation that will enable the successful extraction of characters from handwritten documents as well as the digitalization and translation of handwritten text into machine readable text.





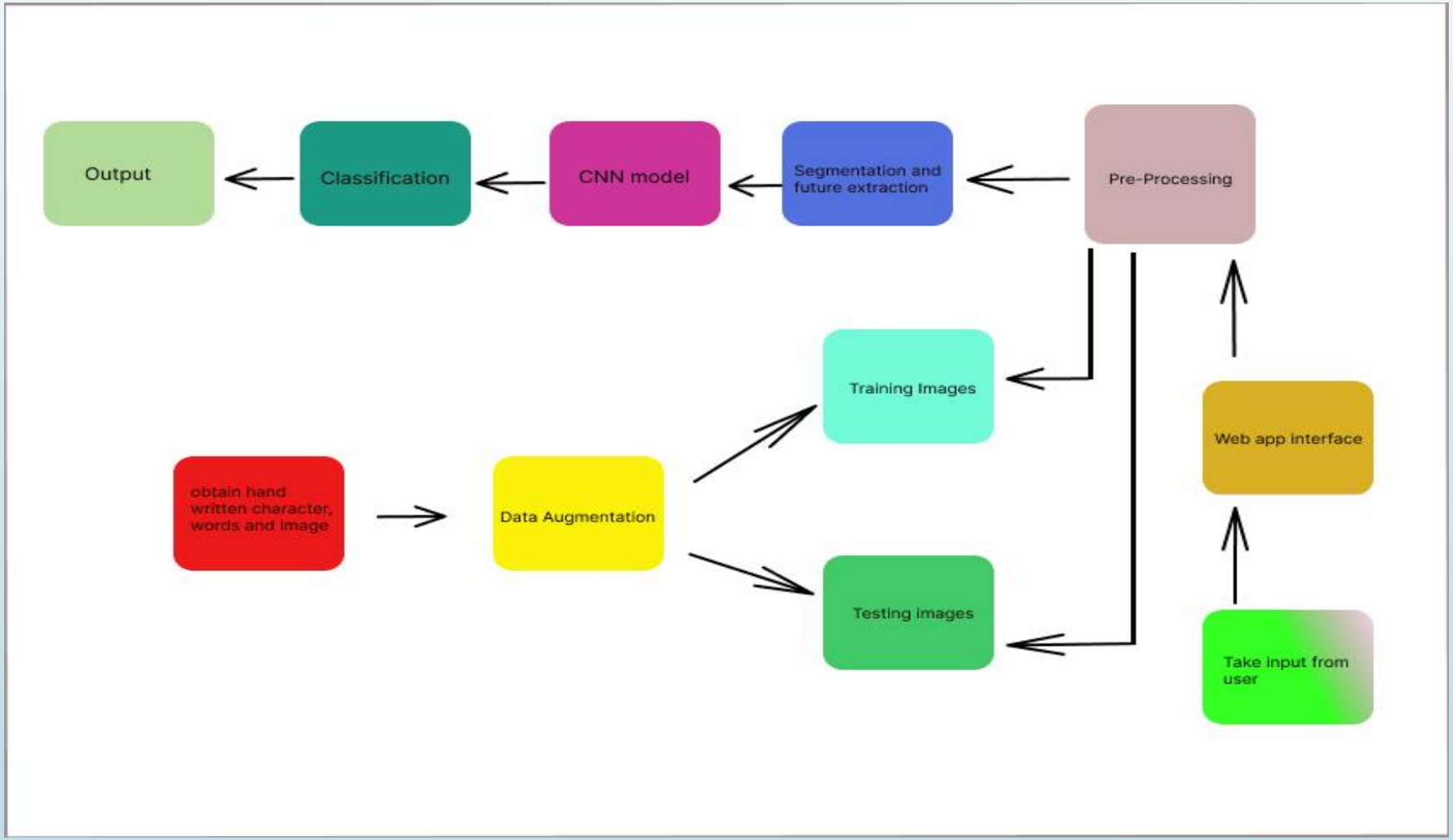
Problem Statement
Architecture of the System

- **Image Acquisition:** The act of obtaining an image from sources is known as image acquisition. Hardware systems like cameras, encoders, sensors, etc. can be used to do this. It is without a doubt the most important phase in the MV workflow because a bad image would make the workflow ineffective as a whole.
- **Data Augmentation:** Data augmentation is the technique of creating additional data points from current data in order to artificially increase the amount of data. In order to amplify the dataset, this may involve making small adjustments to the data or utilising machine learning models to produce new data points in the latent space of the original data.
- **Splitting of dataset:** The division of data into two or more subgroups is known as data splitting. In a two-part split, the model is often trained in the second portion while the first part is utilised to test or analyse the data. A crucial component of data science is data splitting, especially when building models from data.
- **Preprocessing:** Utilizing the method of data preprocessing, unclean data is transformed into clean data sets. In other words, when data is received from various sources, it is always gathered in raw format, making analysis impossible.

Problem Statement
Architecture of the System

- **Segmentation:** This is achieved by the strategy of dividing clients into several groups based on their characteristics or behaviour. Machine learning's use of customer segmentation can reduce waste and enable cost-effective marketing campaigns.
- **Feature Extraction:** The technique of turning raw data into numerical features that can be handled while keeping the information in the original data set is known as feature extraction. Compared to using machine learning on the raw data directly, it produces better outcomes.
- **Classification:** Cross-entropy loss is calculated for classification and weighted classification tasks with classes that are mutually exclusive by a classification layer. The output size of the preceding layer is used to infer the number of classes.
- **Taking real time user input:** The model is now used to identify characters that the user types in real time. To forecast the words that users have written, they must write on paper and upload a scanned copy of that writing.
- **Output:** You get the output detailing what was written in the input image.

Problem Statement Architecture



Survey Papers

- **Author Name :** Gaurav Surve
- **Title :** Handwritten Character Recognition
- **Publication website :** https://www.ijcrt.org/
- **Published Date :** April 4, 2022
- **Objective :** Identifying postal codes on envelopes, analysing extensive financial documents, and processing bank form input are just a few of the many uses for handwritten digit recognition. Many people believed that when using computers for business, whether they were typing or manually filling out forms with letters, they had to modify their input style to conform to what the computer anticipated.
- **Technology used: Tensor flow -** Tensor flow is the core open source library to help you develop and train Machine Learning models.

Existing System

- **Developed by :** Mobiledevca
- **Title :** Handwriting recognizer
- **Publication website :** https://play.google.com/store/apps/details?id=handwriting.to.text.convert
- **Published Date :** Sep 17, 2022
- **Objective :** Handwriting Recognizer uses an extensive and intelligent OCR algorithm to convert handwritten notes into text format. All you need to do is upload the photo or click an image from the camera to get the handwritten notes recognized. This handwritten recognition app “Handwriting Recognizer” easily converts handwritten words into digital form in very less time.

Hardware Requirements:

- Processor: Intel Core i5
- HDD: 1TB
- RAM: Minimum 4GB; Recommended 8GB

Software Requirements:

- Operating System: Windows 11
- Dataset: IAM Dataset
- Programming Language: Python
- OpenCV : OpenCV is a large open-source library for image processing, character recognition, and machine learning. It canscan handwritten images.
- Tensorflow : Tensorflow is the core open source library to help you develop and train Machine Learning models.

De-Merits:

- Multiple studies on the interface between computers and humans have been conducted since the development of computers. Analog to digital conversion is more necessary than ever in this age of modern technologies.
- The writing style of scripts like Devanagari and Gurmukhi makes it difficult to input data characters for them. The best ways to enter these characters are therefore orally and in handwritten documents.
- There are some drawbacks to speech recognition. Therefore, Handwritten Character Recognition is the main emphasis of this study. Handwritten character recognition can be useful in this situation. Recognition of handwritten characters.

Objective

- Due to its widespread use, handwriting recognition has attracted a lot of interest in the domains of pattern recognition and machine learning.
- The application domain for optical character recognition (OCR) and handwritten character recognition (HCR) is specific. For character recognition in a system for handwriting recognition, several strategies have been proposed.
- The methods for transforming the text of a paper document into a machine-readable format are still covered in enough studies and papers.
- Character recognition technology may be crucial in the near future in order to process and digitise existing paper documents and establish a paperless environment. This essay provides a thorough analysis of the field of handwritten character recognition.

Thank You!