Handwritten Text Recognizer on Bank Cheques

Minor Project Report

Submitted for the partial fulfillment of the degree of

Bachelor of Technology

In

Artificial Intelligence and Robotics

Submitted By

Ishita Sharma (0901AI221031)

Shreyank Choudhary (0901AI221060)

UNDER THE SUPERVISION AND GUIDANCE OF

Dr. Abhishek Bhatt

Assistant Professor

Center for Artificial Intelligence

माधव प्रौद्योगिकी एवं विज्ञान संस्थान, ग्वालियर
MADHAV INSTITUTE OF TECHNOLOGY & SCIENCE, GWALIOR
Deemed University
(Declared under Distinct Category by Ministry of Education, Government of India)
NAAC ACCREDITED WITH A++ Grade
Gola Ka Mandir, Gwalior (M.P.)- 474005, INDIA
Ph.:+91-751-2409300, E-mail: vicechancellor@mitsgwalior.in, Website: www.mitsgwalior.in

January 2024

DECLARATION BY THE CANDIDATE

We hereby declare that the work entitled **Handwritten Text Recognizer on Bank Cheques** is our work, conducted under the supervision of **Dr. Abhishek Bhatt, Assistant Professor,** during the session Aug-Dec 2024. The report submitted by us is a record of Bonafide work carried out by us.

We further declare that the work reported in this report has not been submitted and will not be submitted, either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university.

Ishita Sharma (0901AI221031)

Shreyank Choudhary (0901AI221060)

Date:
Place: Gwalior

This is to certify that the above statement made by the candidates is correct to the best of my knowledge and belief.

Guided By:

_____

Dr. Abhishek Bhatt
Assistant Professor
Center for Artificial Intelligence
MITS, Gwalior

Departmental Project Coordinator

_____

Dr. Tej Singh
Assistant Professor
Center for Artificial Intelligence
MITS, Gwalior

Approved by HoD

_____

Dr. Rajni Ranjan Singh
HOD
Center for Artificial Intelligence
MITS, Gwalior

PLAGIARISM CHECK CERTIFICATE

This is to certify that we, the student of B.Tech. in **Center for Artificial Intelligence** have checked our complete report entitled **Handwritten Text Recognizer on Bank Cheques** for similarity/plagiarism using the "Turnitin" software available in the institute.

This is to certify that the similarity in our report is found to be **9%** which is within the specified limit (30%).

The full plagiarism report along with the summary is enclosed.

Ishita Sharma (0901AI221031)

Shreyank Choudhary (0901AI221060)

Checked & Approved By:

_____

Dr. Abhishek Bhatt

Assistant Profressor

Center for Artificial Intelligence
MITS, Gwalior

ABSTRACT

This project focuses on developing a ML based Handwritten Text Recognizer for Bank Cheques. The system leverages advanced techniques in image preprocessing, feature extraction, and deep learning to identify handwritten information, such as payee names, dates, and amounts. By integrating Optical Character Recognition (OCR) with Neural Networks, the model demonstrates strong performance in accurately interpreting diverse and complex handwriting styles. The proposed solution aims to enhance the efficiency, reduce manual errors, and streamline cheque processing in banking environments.

Ishita Sharma (0901AI221031)

Shreyank Choudhary (0901AI221060)

ACKNOWLEDGEMENT

We would like to express our greatest appreciation to all the individuals who have helped and supported us throughout this Minor Project-1. We are thankful to whole Center for AI department for their ongoing support during the experiments, from initial advice and provision of contact in the first stages through ongoing advice and encouragement, which led to the final report of this Minor Project-1. A special acknowledgement goes to our colleagues who helped us in completing the file by exchanging interesting ideas to deal with problems and sharing their experience. We wish to thank our professor Dr. Abhishek Bhatt as well for his undivided support and interests which inspired me and encouraged me to go our own way without whom we would be unable to complete our project. At the end, we want to thank my friends who displayed appreciation for our work.

Ishita Sharma (0901AI221031)

Shreyank Choudhary (0901AI221060)

CONTENT

Table of Contents

LIST OF FIGURES

ACRONYMS

- **ML-** Machine Learning
- **HTR-** Handwritten Text Recognizer
- **SGD-** Stochastic Gradient Descent
- **PIL-** Python Imaging Library
- **OCR-** Optical Character Recognition
- **ReLU-** Rectified Linear Unit
- **OpenCV-** Open Source Computer Vision Library
- **BCELoss-** Binary Cross- Entropy Loss

CHAPTER 1: INTRODUCTION

The objective of this project is to create a machine learning based Handwritten Text Recognizer for bank cheques. The system is able to extract useful information from the borders of the cheques using image preprocessing approaches and deep learning architectures like Neural Networks. This makes the cheque processing efficient and reduces the chances of human error thereby speedier and safer monetary transactions.

1.1 Challenges of processing handwritten cheques physically and manually:

Human Error: The information recorded can be debatable, for reasons such as manually entering data where one might misinterpret some unclear words or numeric figures. This is dangerous as it can distort records.

Time Consuming: Processing each cheque manually requires a considerable amount of time results in slowing down the workflow.

High Labor Cost: Cheque processing does not only require machinery but also people, which raises wage costs.

Diminished Expansion Potential: At times, it is actually less efficient and operationally unrealistic to carry operations in these manual processes due to the rising number of cheques.

1.2 The Need for Automated Solution:

Considering the increasing number of cheques that are handwritten and the inefficiencies of people processing them, there came a need for automated solutions. This is because with automation, human error is unlikely and the processing speeds and accuracy is improved. Operating costs are low because the majority of cheques are processed under a short period promoting large volumes of cheques to be processed with less effort. In addition to this, fully mechanized systems improve the management of the cheque book eliminating fraud and enhancing overall management of its processing.

CHAPTER 2: LITERATURE SURVEY

The recognition of handwritten text has been and still continues to be, a demand in computer vision and machine learning fields. A lot has gone into its development especially in the recent past with the introduction of deep learning. The present paper is a literature review aimed at providing necessary information on issues and approaches that are important for designing a handwritten text recognition engine for cheques.

2.1 Deep Learning Approaches

This is a simple feedforward neural network based on PyTorch, which comes with binary classification. I'm setting up a model architecture that has an input layer size of 62720, and then a hidden layer using ReLU activation function with 128 neurons, followed by an output layer with the sigmoid function to produce probabilities between 0 and 1.

The network is trained using SGD with binary cross-entropy loss. The class SimpleNetTrainer encapsulates the training where it performs forward passes, computation of the loss, backpropagation of the gradients, and updating of the weights. The validation runs in a no-gradient context in order to assess the model while not modifying its parameters.

A function, predict_bb, preprocesses and analyzes images by cropping out the regions of interest and outputs random similarity scores to inspect. It can be used for tasks such as handwritten text recognition and validation of data points on images, such as bank cheques.

2.2 Image Preprocessing Techniques

Image Pre-processing techniques such as binarization, noise removal etc. carry out an important function as far as enhancing the image for recognition purposes is concerned. Methods such as adaptive thresholding and morphological processes have been extensively researched for the imaging of cheques prior to examination.

2.3 StreamLit webpage Development Approach

This code is a Streamlit app processing cheque images using a pre-trained PyTorch neural network. It loads a simple feedforward model.

It crops the specified regions from the uploaded cheque image and then converts the cropped images into tensors and gives predictions on running the model. For both original and cropped images with predictions, produces images to view the comparison. Simulates handwriting match scores between the cropped regions.

The app provides a way of uploading cheque images with the ability to give users predictions on specified areas.

2.4 Application in Banking

A number of works pay attention to the research towards the introduction of textual recognition systems into the banking operations. Such systems make cheque clearing processes faster and more reliable as they minimize unnecessary human actions and enhance the probability of detecting fraud. The works also discuss the significance of developing solutions that are able to artefacts with realistic handwriting models.

CHAPTER 3: OVERVIEW AND FLOWCHART

Handwritten Text Recognizer for bank cheques is a specialized system designed to read and interpret handwritten information on cheques. This involves recognizing fields such as the payee's name, amount in numbers, amount in words, and signature. Here's an overview of its components and workflow:

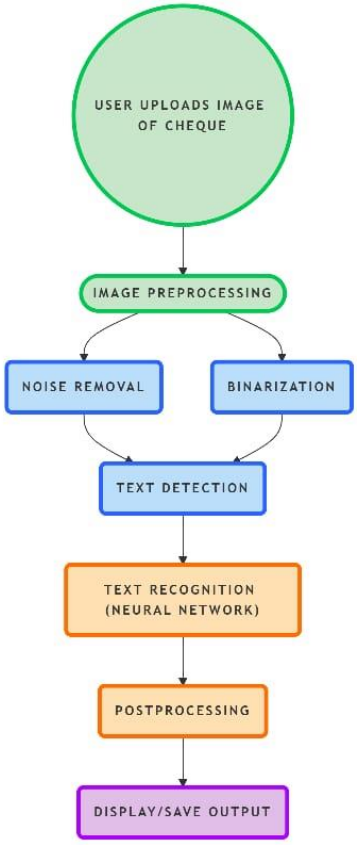


Fig 1. Flowchart

3.1 Workflow Overview

The first stage in the operation of the system is image preprocessing, which is a very important step in the enhancement of the quality of the cheque image uploaded to facilitate its recognition. Preprocessing includes grayscale conversion, noise filtration, normalization, binarization of the image for clearer background text. If the cheque image has a slight tilt, alignment correction technologies will come into play to adjust the input.

Subsequently, with image processing tools for example OpenCV, areas on the cheque that correspond to the handwritten fields are marked. This is done by drawing contours, trimming bounding box or template making about regions containing the text, for instance, the amount box, date box, and signature regions. These are then fed into a model aimed at recognizing handwriting from complex images based on neural networks. This model is constructed in a framework known as PyTorch.

A simple feedforward neural network is defined with two fully connected layers:

Input layer: Takes input data of size input_size.

Hidden layer: Applies ReLU activation to introduce non-linearity.

Output layer: Uses a sigmoid activation function to output a probability between 0 and 1.

The model is exposed to several samples of handwritten text in different variations, to prevent it from beyond the limits some different handwritings in text blocks which will include variations from strokes designs and width to spacing and them being different parts of words together with text lines.

Following text recognition, the system automatically undertakes post-processing of the output. This includes checking against certain assumptions about the correct form (for instance, numeric values for sums, and the dates of the particular formats) and clarifying the inconsistencies of written text using a language model containing a hand-writing understanding component. Finally, the generated text is also organized in a certain way which may be in form of a table or the image of the cheque may be used with the text overlaid on it.

3.2 Web Interface

The system can be used through a web browser, constructed with Streamlit framework. Cheque images can be uploaded by users via this interface, which in turn initiates the complete recognition pipeline. Recognized text is shown on the interface along with certain visual feedback in the form of detecor boxes around the areas with handwritten text. Such a mechanism is beneficial as it promotes user interaction and allows the output of the machine to be checked for mistakes as well as corrected whenever necessary.

CHAPTER 4: TECHNOLOGIES USED

Technologies involved in handwritten text recognition on bank cheques include the following:

1. Back-End Technologies

1.1 Machine Learning (ML):

- **Construction of Neural Network**: The project uses PyTorch to implement a custom neural network defined by the Net class with one hidden layer in a simple feedforward network.
- **Training and Evaluation:** In PyTorch, we defined the training loop, computed losses using BCELoss, and learned the model using SGD.
- **Tensor Operations:** PyTorch uses tensor operations to change the input data into a tensor. It handles dimensions for the data and makes the forward and backward passes in the training loop.

1.2 Programming Languages:

Python – Easy to learn, has numerous machine learning and data science libraries, it is evident that python is a liked programming language.

There are a few such tools in Python for building systems that recognize handwriting. For instance:

- **Pytorch:** Yet another deep learning library to build computations graph networks.
- **OpenCV:** Used in computer vision tasks for instance reducing background noise, improving image quality, and dividing the image into regions.
- **PIL Pillow:** Pillow is an image handling library for Python that allows you to create, manipulate, and edit images in your Python Applications.
- **NumPy:** NumPy is one of the fundamental packages for numerical computation in Python and supports large multi-dimensional arrays and matrices along with high-performance mathematical functions to operate on these data structures.
- **os module:** The os module in Python has various functions to interact with the operating system, including navigation throughout the file system.
- **glob module:** Glob Overview Glob finds all the pathnames matching a given pattern according to Unix shell rules.

2. Front-End Technologies:

2.1 Web Development

- **Streamlit:** Streamlit is a powerful Python library that allows you to create web apps with minimal coding effort. It's particularly well-suited for data scientists and machine learning engineers to quickly prototype and deploy data-driven applications.

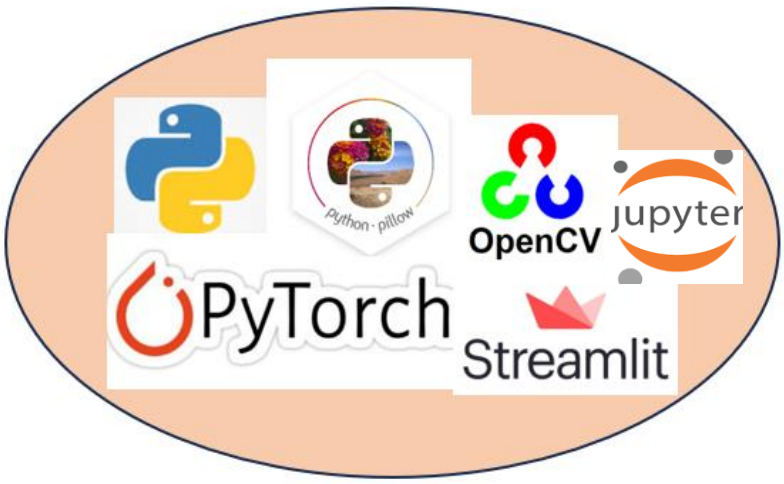


Fig 2. Technologies Used

CHAPTER 5: WORKING OF THE MODEL

5.1 Data Collection

• Sources:

The data used in this project was collected through cheques so different varieties of hand writing and cheque designs are represented in the data. We have collected this dataset form kaggle.

• About Dataset:

This project uses a dataset of cheque images with handwritten text such as names and amounts, which comprises crops of images in various regions of interest for study purposes. It further uses a subset of handwritten images with filenames structured to indicate writer IDs for handwriting matching and comparison exercises. We store metadata, such as an image file path and writer ID, into a DataFrame and preprocess and handle the data. Such arrangements shall ensure that this model will effectively identify text written by hand and also retrieve essential information.

5.2 Data Processing

• Preprocessing:

- Images are first read, resized, and converted to a suitable format using the image library PIL.
- Cutting out particular sections (usually referred to as bounding boxes) from images for further analysis.

• Normalization:

- There is normalizing of pixel values for faster convergence of the model.

• Augmentation:

- Includes simple rotation, flipping of images and brightness adjustments to expand the data set variety.

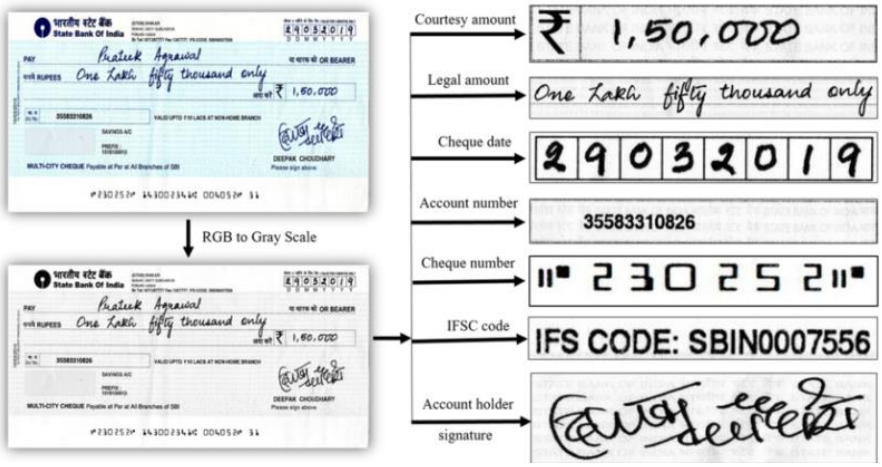


Fig 3. Conversion of image from rbg to gray scale

5.3 Machine Learning Model

• Architecture:

- A simple feedforward neural network (Net) is applied which consists of:
  - **Input Layer:** Receives a feature vector extracted from a flattened image.
  - **Hidden Layer:** Non-linear transformations are performed using ReLU.
  - **Output Layer:** Following a Binary classification, Sigmoid activation outputs a single value.

• Training:

- **Loss Function:** Binary Cross-Entropy Loss (BCELoss) is used to get optimal prediction results of things that only have two outcomes.
- **Optimizer:** Stochastic Gradient Descent (SGD) is used in this case to modify weights in a stepwise manner.
- **Epochs:** Numerous cycles that involves going through the entire data set in order to reduce the loss further.

Fig.4 Model Trainig

5.4 Prediction

• Functionality:

- A technique predict_bb determines if the handwriting is the same or different in separate areas on cropped images of the checks.
- Mimics the same zone multiple regions scores for the cheque or different cheques.

• Process:

- Cuts images using the bounding box and its coordinates.
- Analyses distance of handwriting areas using statistical techniques.

5.5 Graphical User Interface

• During Prediction:

- Incorporates PIL for better vision of entire checks and their portions.

• In Terms of User Interaction:

- It is possible to upload images to perform prediction by the system.

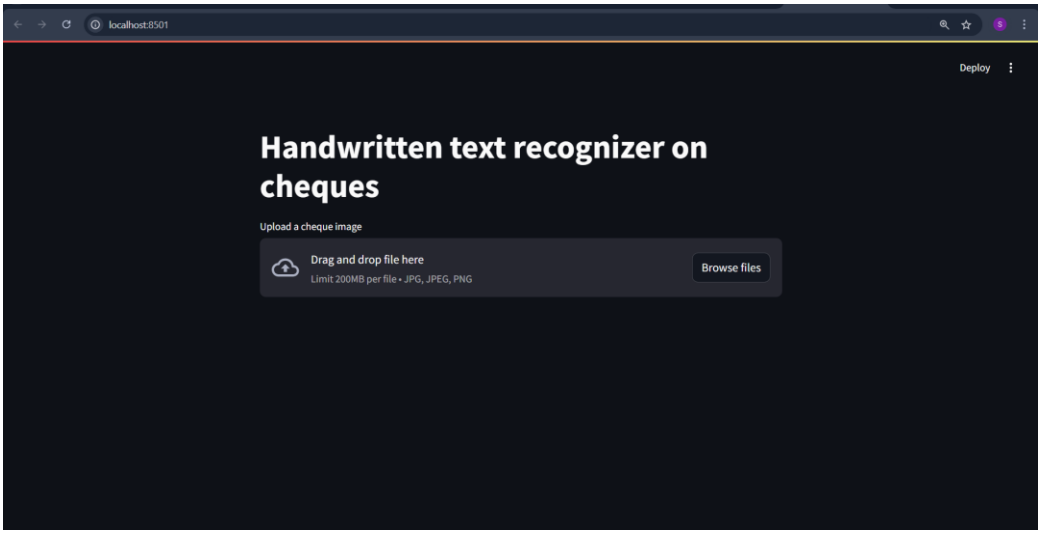


Fig 5. User Interface

5.6 Requirements for Result Presentation.

• Outputs:

- Representational numerical similarity indices which are indicative of the probability that the given handwritings match.

• Visualization:

- Cropped images of the cheques provide a means of appreciating the areas that are under scrutiny.

• Model Saving:

- The weight of the trained models and the states of the optimizers are both stored (model.pth extension) for future purposes. o Facilitates fast inferring and reloading of processes without the need to train the model again.

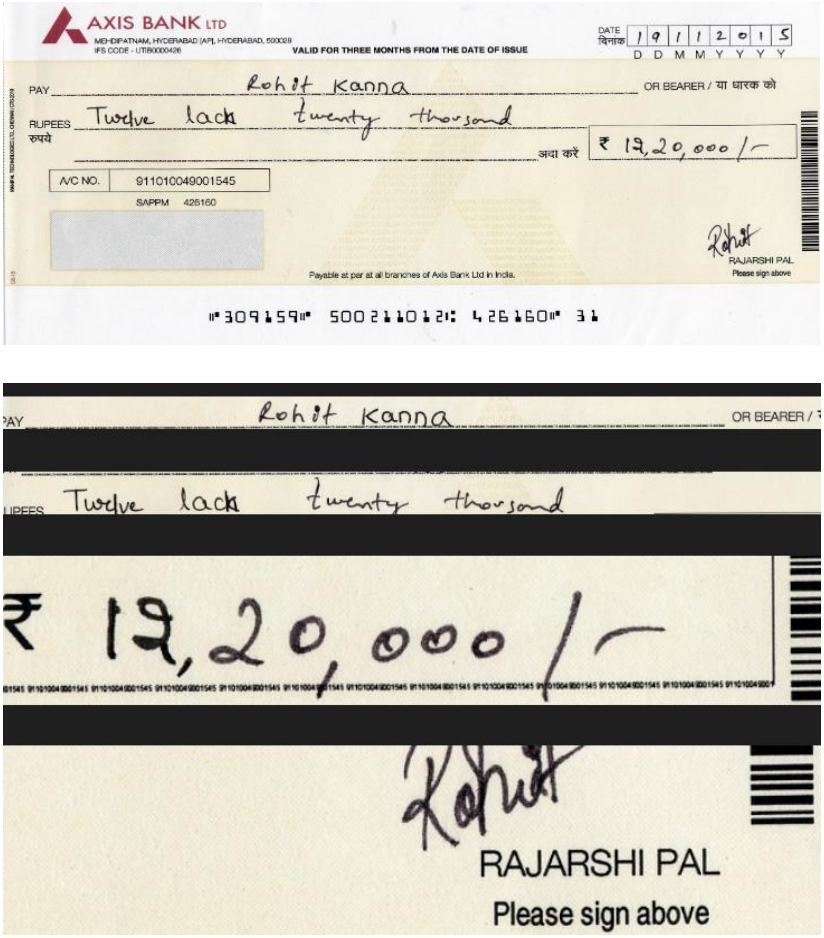


Fig 6. Extracting name of payee, amount and signature from the cheque

CHAPTER 6: FINAL ANALYSIS AND DESIGN

The Web interface for Handwritten Text Recognizer on Bank Cheques model using NET allows user to upload image of a cheque. The model processes the cheque by applying preprocessing techniques and makes a bounding box around the area that has to be extracted i.e name of payee, amount and signature. Its also simulates handwriting match scores between the cropped regions.

Fig 7.1 User Interface

Fig 7.2 Extracted data from the cheque

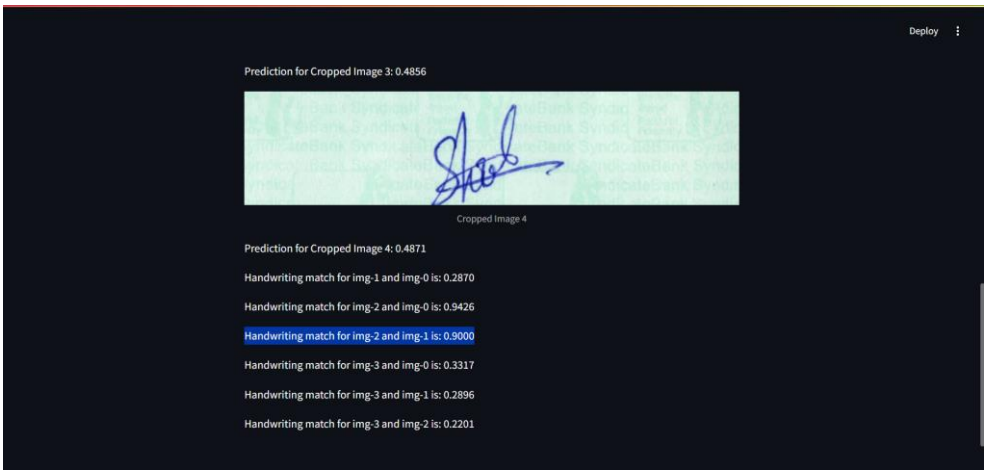


Fig 7.3 Handwriting match scores of cropped images

CHAPTER 7: CONCLUSION AND FUTURE ENHANCEMENTS

The application of machine and deep learning techniques has led to substantial improvements in handwritten text recognition (HTR) technology, particularly in the context of bank checks. The characterization and recognition processes use image processing techniques combined with neural networks to recognize even the most intricate types of handwritten scripts, promoting operational efficiency and ease of use in banks.

1. Achievements made in the area of HTR with specific focus on bank cheques:

Improved Accuracy: This handwritten text detection model can achieve remarkable accuracy levels in the detection of even the most problematic handwriting and low-resolution images.

Enhanced Speed: This cheque processing is performed with real-time, which contributes to the expeditious nature of the banking sector.

Reduced Manual Intervention: The use of cheque processing automation reduces the possibility of human errors thus cuts on the costs incurred.

Increased Security: This handwritten text recognition systems possess the capabilities of preventing fraud over the use of handwriting evaluation by spotting abnormalities in writing styles.

2 Future Enhancements:

Though substantial advancement has occurred, there are many areas which can further be researched and developed to enable the HTR systems to be even more extreme robust. In future we will try to extract the exact selectable text from the bounding boxes.

2.1 Dealing with Different Fonts and Styles of Handwriting:

- Creating architecture that recognizes handwriting recognition technologies to match varied styles of hands from likely curvy to square hands.
- Using for instance data programmers and domain adapters to aid efforts to enhance generalisation.

2.2 Difficult and Noisy picture Recognition:

- Identifying ways of advanced noise filtering techniques and imaging enhancement so as to rectify or make low-grade imaged pictures useable.
- Designing smart devices that will be purposely made for recognizing crafts and resistant to images in motion which may contain such images.

2.3 Deployment and Processing in Real Time:

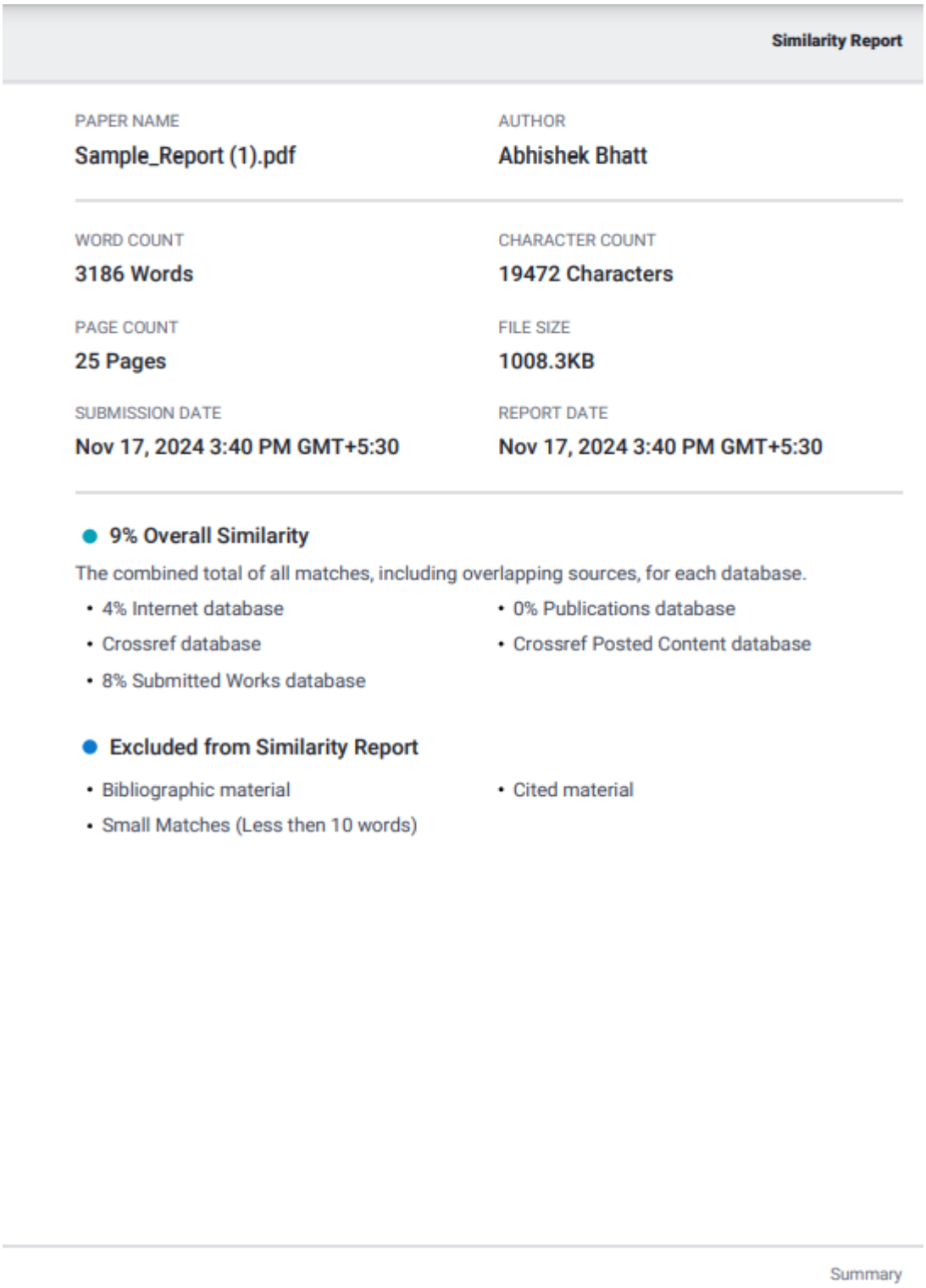
- Adapting models for processing in real time.

REFERENCES

- https://www.kaggle.com/datasets/tejasreddy/iam-handwriting-top50
- https://www.kaggle.com/datasets/pranav10000/chequedetection
- https://pytorch.org/docs/stable/index.html
- https://pytorch.org/tutorials/beginner/blitz/neural_networks_tutorial.html
- https://pillow.readthedocs.io/en/stable/reference/Image.html
- https://www.geeksforgeeks.org/how-to-use-glob-function-to-find-files-recursively-in-python/
- https://docs.streamlit.io/develop/concepts

TURNITIN PLAGIARISM REPORT

Please Insert a Scanned Copy of the Front pages duly signed by the Candidate, Supervisor, Departmental Turnitrin Coordinator, and HoD with Seal



Similarity Report

PAPER NAME
Sample_Report (1).pdf

AUTHOR
Abhishek Bhatt

WORD COUNT
3186 Words

CHARACTER COUNT
19472 Characters

PAGE COUNT
25 Pages

FILE SIZE
1008.3KB

SUBMISSION DATE
Nov 17, 2024 3:40 PM GMT+5:30

REPORT DATE
Nov 17, 2024 3:40 PM GMT+5:30

● 9% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 4% Internet database
- 0% Publications database
- Crossref database
- Crossref Posted Content database
- 8% Submitted Works database

● Excluded from Similarity Report

- Bibliographic material
- Cited material
- Small Matches (Less then 10 words)

Summary

MPRs (If Applicable)