

Automated Personal Writing Pattern Replicator

Team Members:

Adarsh Ghimire(THA077BCT005)

Kishan Adhikari(THA077BCT021)

Nasir Hussain(THA077BCT027)

Prabuddha Jung Thapa(THA077BCT032)

Supervised By:

Er. Kiran Chandra Dahal

Department of Electronics and Computer Engineering
Thapathali Engineering Campus

August 2024

Presentation Outline

- Introduction
- Motivation
- Objective
- Scope
- Project Application
- Methodology
- Results & Discussion
- Remaining Task
- References

Introduction

- This project focuses on accurately cloning individual handwriting styles using advanced machine learning, enabling the generation of personalized handwritten text and digital fonts.
- It aims to preserve unique handwriting characteristics, to enhance digital personalization and accessibility for various applications, including custom fonts and educational materials.

Motivation

- Preserve unique handwriting styles affected by physical limitations
- Applications in customized fonts and educational materials.
- Aims to address limitations of traditional handwriting cloning methods by ensuring style consistency and handling arbitrary text lengths.

Objective

- Develop a machine learning model to replicate handwriting styles
- Generate consistent, high-fidelity handwritten font for various applications.

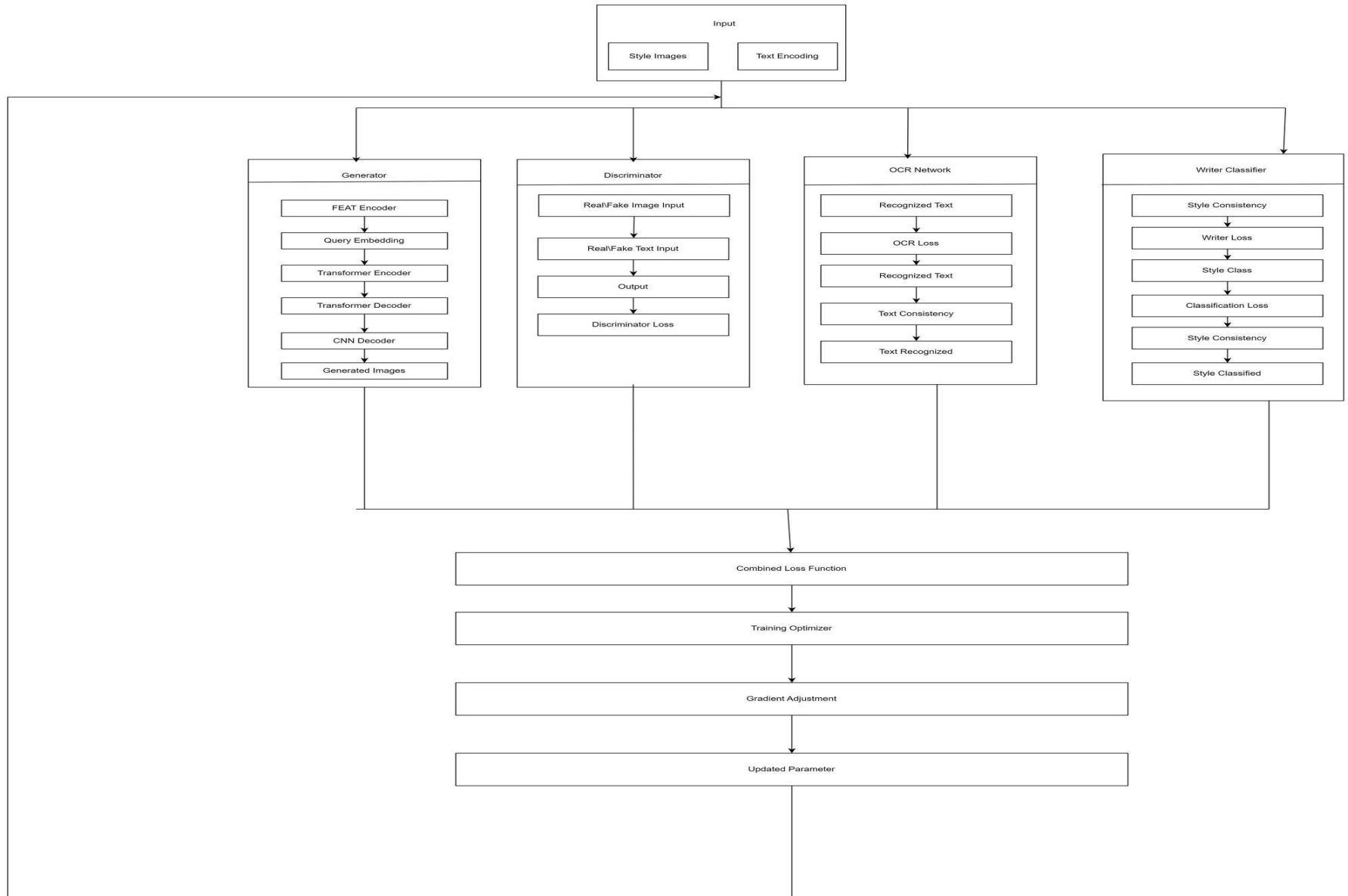
Scope

- Personalized digital communication
- Educational material creation
- Historical document preservation
- Branding and Marketing

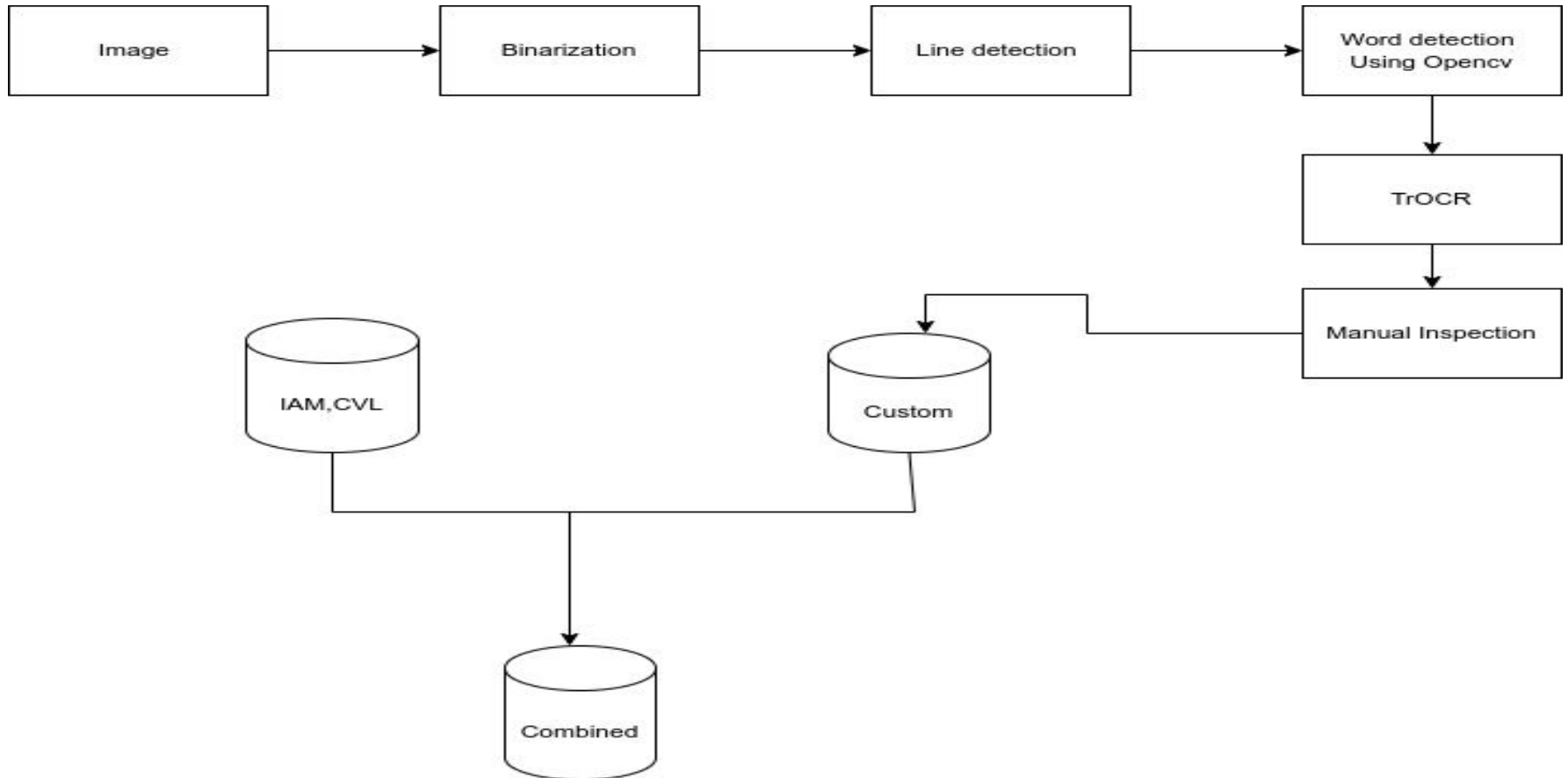
Project Application

- E-commerce Personalization
- Used in Calligraphy
- Custom Handwritten Digital Interfaces
- Personal Blogging and Digital Journals

Methodology(System Diagram)



Methodology (Dataset Preparation)



Methodology(Datasets)

S.N.	Datasets	Writers
1.	IAM Handwriting Dataset	657
2.	CVL Handwriting Dataset	350
3.	Custom Dataset	38

Though they may gather some Left-wing support, a large majority of Labour OM Bs are likely to turn down the Foot-Griffiths resolution. Mr. Foot's line will be that as Labour OM Bs opposed the Government Bill which brought life peers into existence, they should not now put forward nominees. He believes that the House of Lords should be abolished and that Labour should not take any steps which would appear to "prop up" an out-dated institution.

Imagine a vast sheet of paper on which straight Lines, Triangles, Squares, Pentagons, Hexagons, and other figures, instead of remaining fixed in their places, move freely about, on or in the surface, but without the power of rising above or sinking below it, very much like shadows - only hard and with luminous edges - and you will then have a pretty correct notion of my country and countrymen. Alas, a few years ago, I should have said "my universe": but now my mind has been opened to higher views of things.

Imagine a vast sheet of paper on which straight Lines, Triangles, Squares, Pentagons, Hexagons, and other figures, instead of remaining fixed in their places, move freely about, on or in the surface,

Micro kernels : Micro kernel incorporates only the essential set of OS services into the kernel. The rest of the OS services are implemented in program known as 'servers' which runs in user space. This provides a highly modular design and OS-neutral abstraction to the kernel. Memory management process, timer systems and interrupt handlers are the essential services. Examples: Mach, QNX

This architecture caters to the problem of ever growing size of kernel code which we could not control in the monolithic approach. This architecture allows some basic services like device driver management, protocol stack, file system to run in user space.

This reduces the kernel code size and also increases the security and stability of OS as we have the bare minimum code running in the kernel. So if suppose a basic service like network service,

DNF with suitable diagram. It is a step in the normalization process for database ensuring that all constraints on a relation are expressed as logical implications on the attributes of that relation. DNF requires that the database contains no constraints other than domain constraints and key constraints. Assume that the domain for wealthy person consists of the name of all wealthy people in a predefined sample of wealthy people. The domain for wealthy person for wealthy person type consists of the value 'Millionaire' and 'Billionaire' and the domain for net worth

Fig: Custom Dataset Sample

Methodology

(Data Processing)

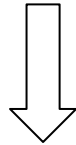
- Dataset present in IAM , CVL database consists of handwritten pages but for our model training we need to crop out the lines.
- We binarize the image to enhance the quality of the images.
- We extracted the lines by defining the bounding box.
- Further, bounding boxes were defined on the extracted lines to isolate the words present in those lines.

Methodology

(Data Processing)[Cont.]

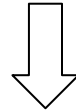
- Passed the extracted words images to TrOCR model to get the transcription.
- The words images were converted to grayscale to simplify processing and reduce computational load.
- Resized images to a consistent size of height 32 pixel and maximum width of 192 pixel to ensure uniformity in the dataset.

as logical implications on the attributes that
relation. DKNF requires that the database contains
no constraints other than domain constraints and
key constraints. Assume that the domain for
wealthy person consists of the name of all wealthy
people in a predefined sample of wealthy people

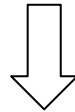


as logical implications on the attributes that
relation DKNF requires that the database contains
no constraints other than domain constraints and
key constraints Assume that the domain for
wealthy person consists of the name of all wealthy
people in a predefined sample of wealthy people

as logical implications on the attributes that
relation DKNF requires that the database contains
no constraints other than domain constraints and
key constraints Assume that the domain for
wealthy person consists of the name of all wealthy
people in a predefined sample of wealthy people



the normalization process for database ensuring



database
database
for
for
ensuring
ensuring
process
process
the
the
normalization

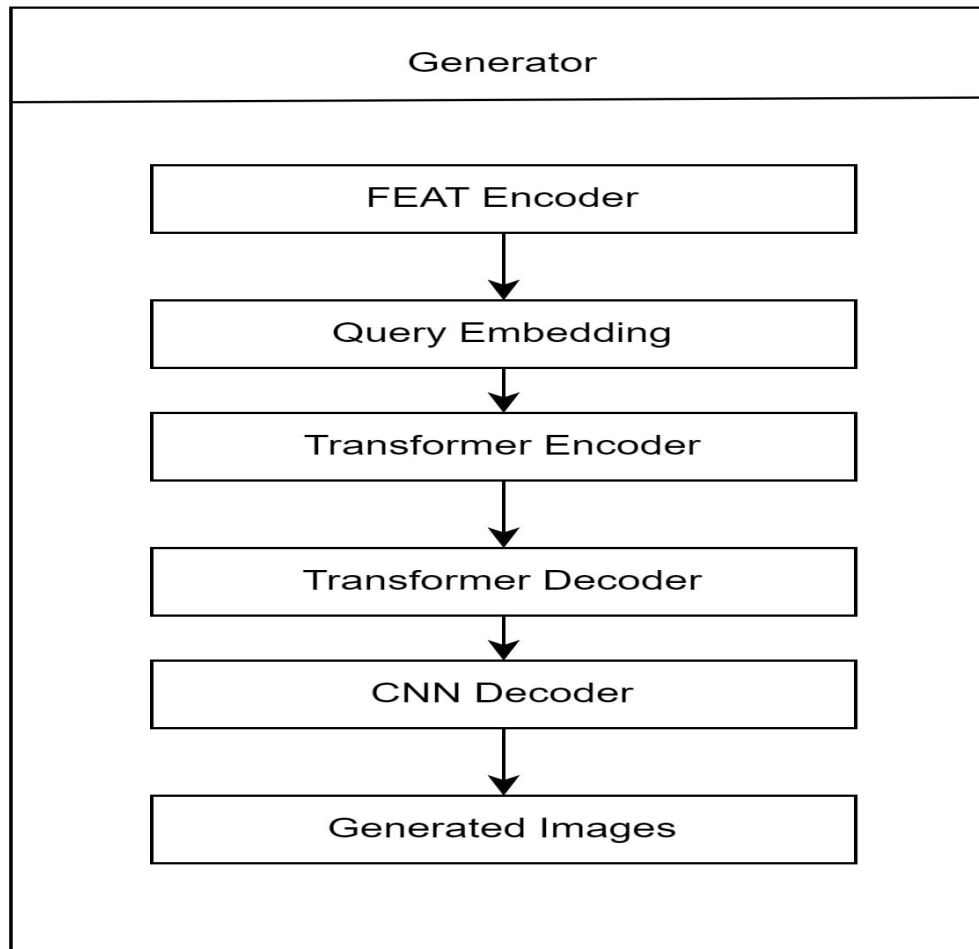
Fig: Custom Dataset Preprocessing

Dataset Overview

image	label	writer
image · width (px)	string · lengths	string · classes
		643 values
	in	245
	all	245
	parts	245
	of	245
	the	245

- **Image:** Contains images of individual handwritten words.
- **Label:** Includes the text transcription of each handwritten word image.
- **Writer:** Denotes the writer of each word, aiding in style and writer discrimination.

Methodology (Proposed Model Architecture)



Methodology (Proposed Model Architecture)- Generator

Style Encoding:

- **Feature Encoder:** CNN-based network extracts stylistic features from input handwriting examples, converting images into feature representations. The obtained features are flattened and passed to the transformer-based encoder layer.
- **Transformer Encoder:** Processes the style feature sequence to capture handwriting nuances, derived from the CNN based feature encoder.

Methodology (Proposed Model Architecture)- Generator

Content Encoding:

- **Query Embedding:** Transforms the words from the query into embeddings, representing the content that needs to be handwritten.

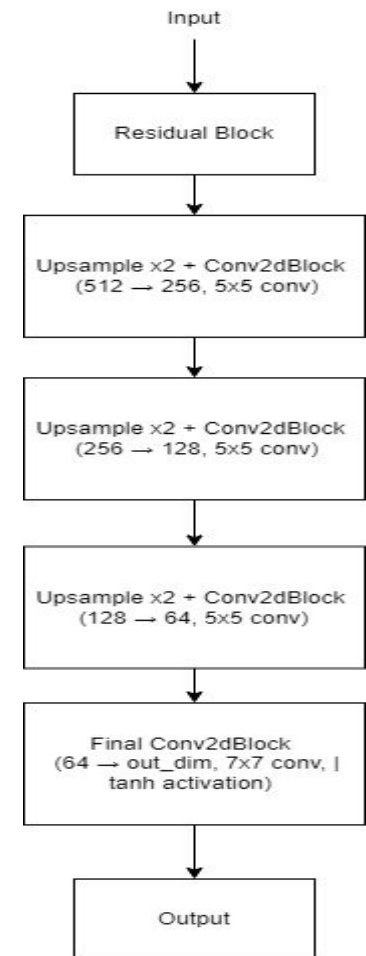
Decoding:

- **Transformer Decoder:** Merges the style features and the content embeddings to create a unified feature representation to ensures that the generated handwriting maintains the desired style and accurately reflects the content of the query words.

Methodology

(Proposed Model Architecture)

- **CNN Decoder:** Converts the combined feature representation into the final image of the handwritten text.

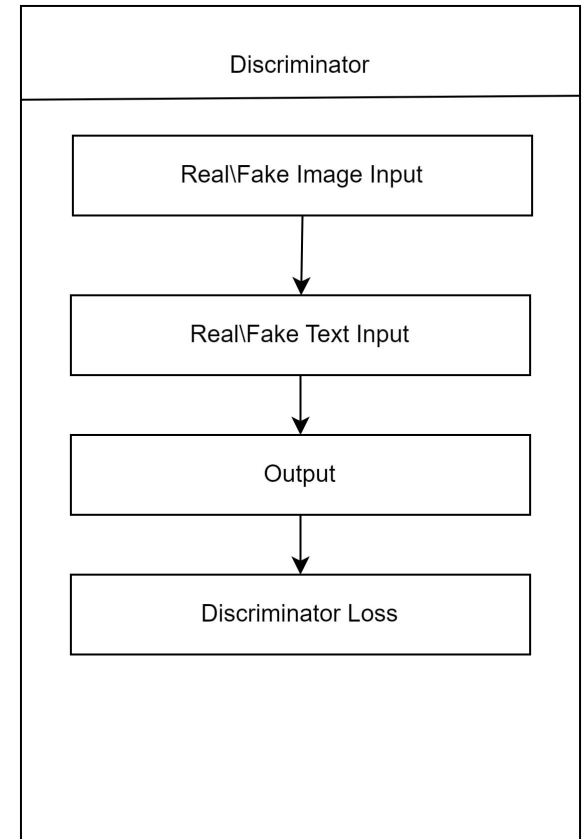


Methodology

(Proposed Model Architecture)

Discriminator:

- Differentiates between real and generated handwriting samples to improve the generator's output through adversarial training.

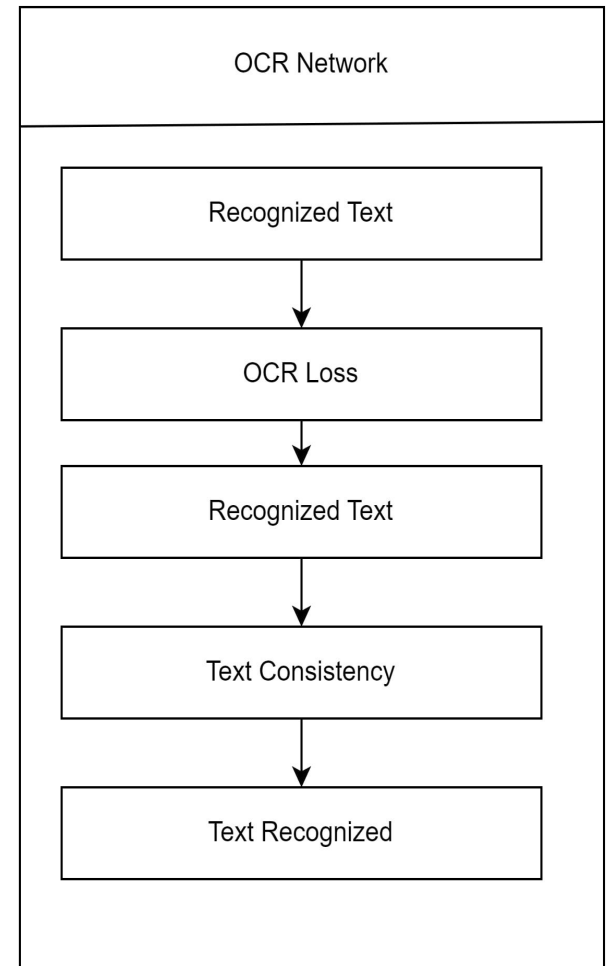


Methodology

(Proposed Model Architecture)

Recognizer:

- Ensures the generated handwriting is legible and correctly represents the input text by recognizing and validating the characters.

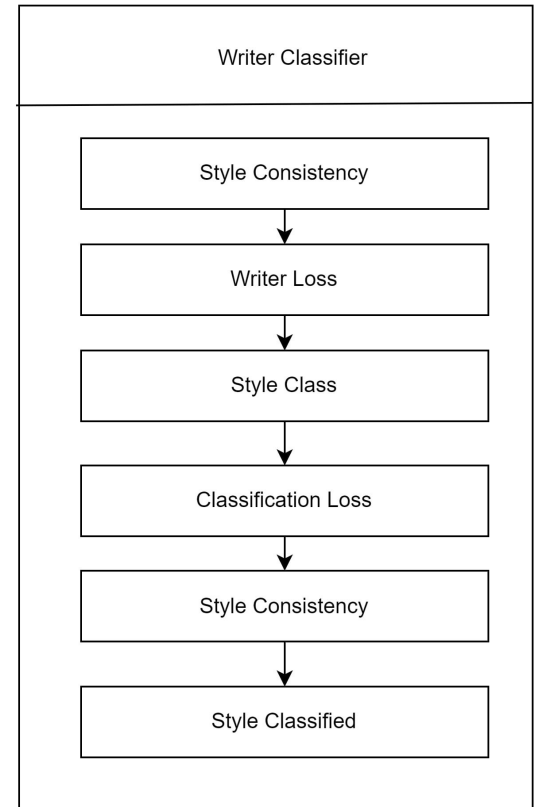


Methodology

(Proposed Model Architecture)

Style Classifier:

- Classifies and extracts individual handwriting styles, ensuring the generated handwriting matches the original style attributes of the writer.



Methodology (Evaluation)

FID(Fréchet Inception Distance)

- It is a metric used to measure the similarity between two sets of images

Human Evaluation

- Finally, We perform our model performance using human experience.

Results & Discussion

- So far we have trained our model on custom dataset that we have prepared and we ran around 963 epochs of training on that dataset and saw the result in validation set like this :

pre-defined trajectory for a particular includes select of case the number stands speed can be re-used and the ARCHITECTURE USE

and movement when his in in his himself last where up last things before

minds after then But said But But for to rock after Carine receive a trees

and which compute: best individuals were the most accurate of the 100 sound surfaces that

the hearing with flood police the one run no actually of from from with her

The Statue of Liberty is arguably one of New York City's most iconic symbols is a popular tourist attraction for first-time visitors. This 305-foot monument was gifted to the United States from France in order to celebrate 100 years of America's independence. When Claire Statue of Liberty for the first time, she instantly admired it as a symbol of freedom.

The Statue of Liberty is arguably one of New York City's most iconic symbols is a popular tourist attraction for first-time visitors. This 305-foot monument was gifted to the United States from France in order to celebrate 100 years of America's independence. When Claire Statue of Liberty for the first time, she instantly admired it as a symbol of freedom.

The Statue of Liberty is arguably one of New York City's most iconic symbols is a popular tourist attraction for first-time visitors. This 305-foot monument was gifted to the United States from France in order to celebrate 100 years of America's independence. When Claire Statue of Liberty for the first time, she instantly admired it as a symbol of freedom.

Result

✨ Handwriting Cloning and Font Generation

Input text


Computers are versatile electronic devices that can process, store, and retrieve data at high speeds. They consist of hardware and software components that work together to perform a wide range of tasks, from simple calculations to complex simulations, making them essential tools in various fields such as science, engineering, education, and entertainment.

Choose from provided writer styles

files/example_data/style-46

OR

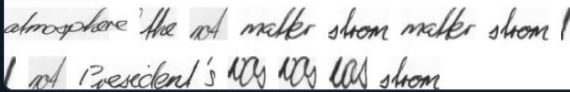
Upload multiple word images


Drop File Here
- or -
Click to Upload

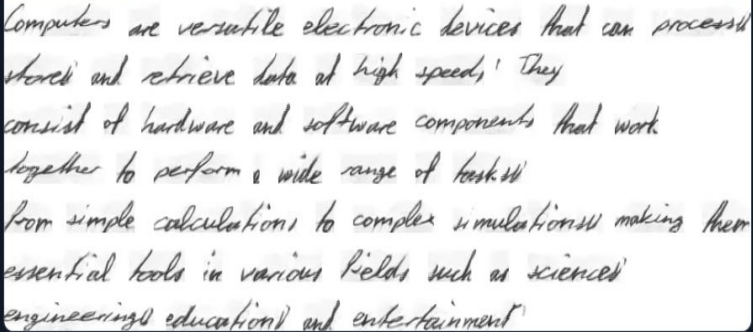
Clear

Submit



Style Image



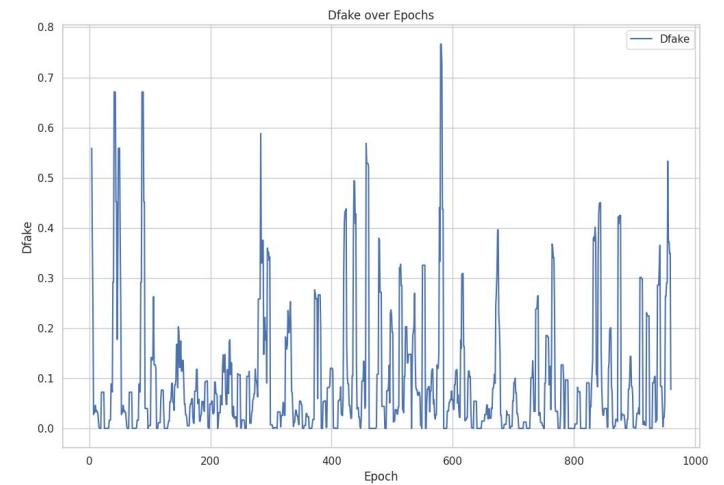
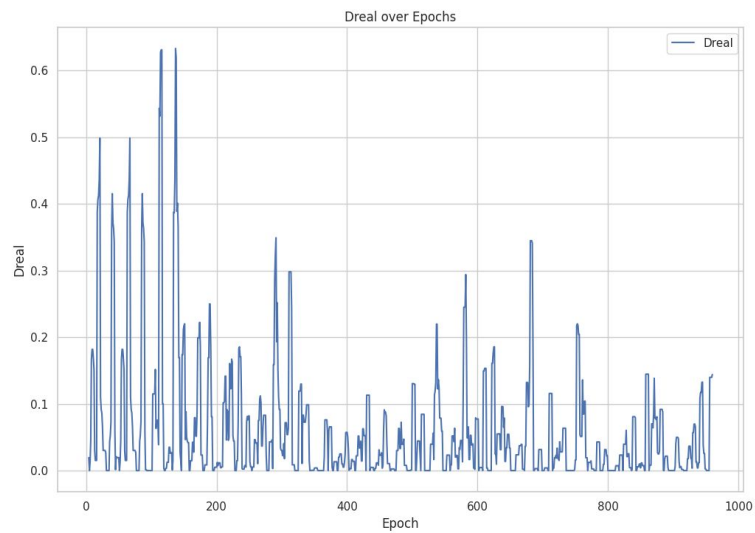
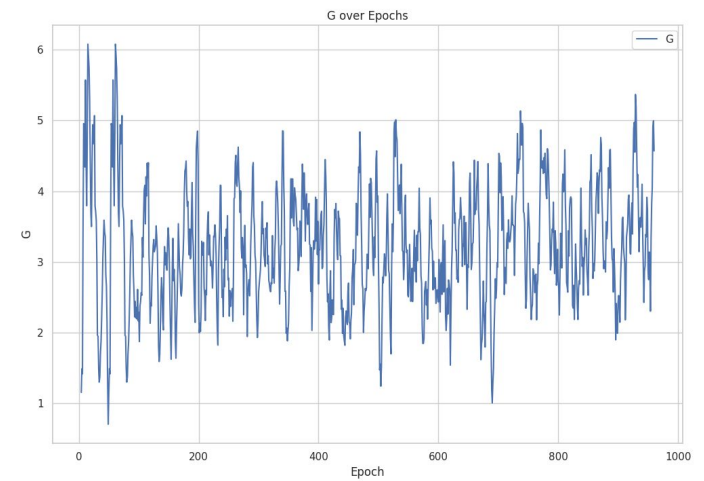
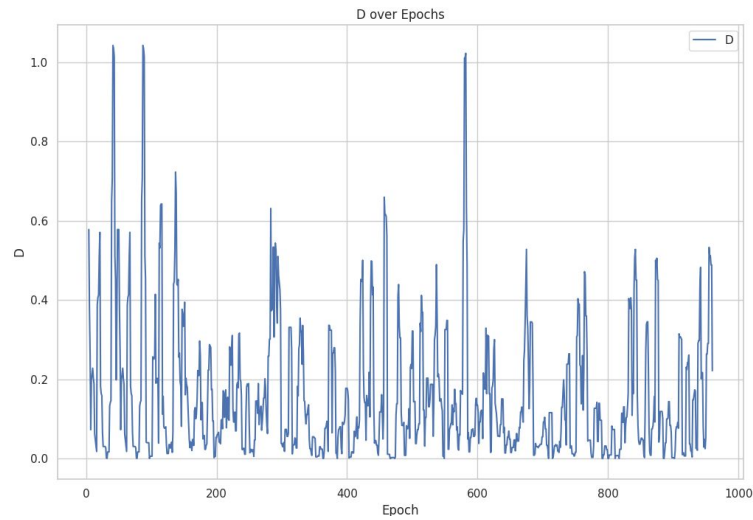
Generated Image

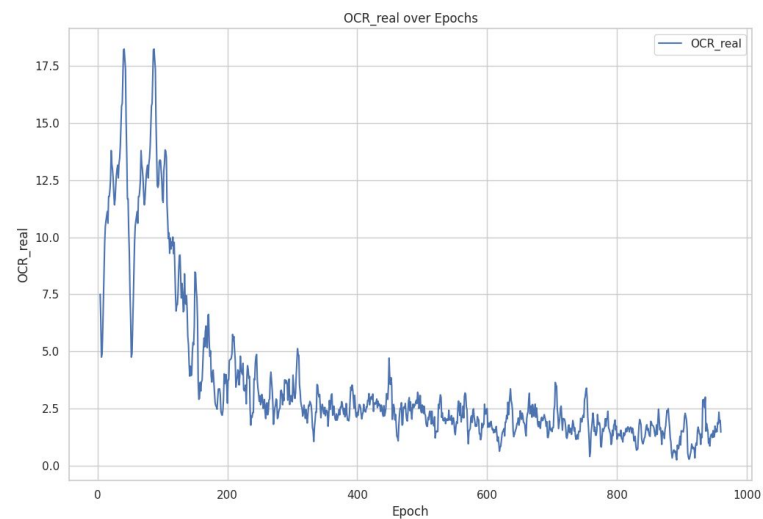
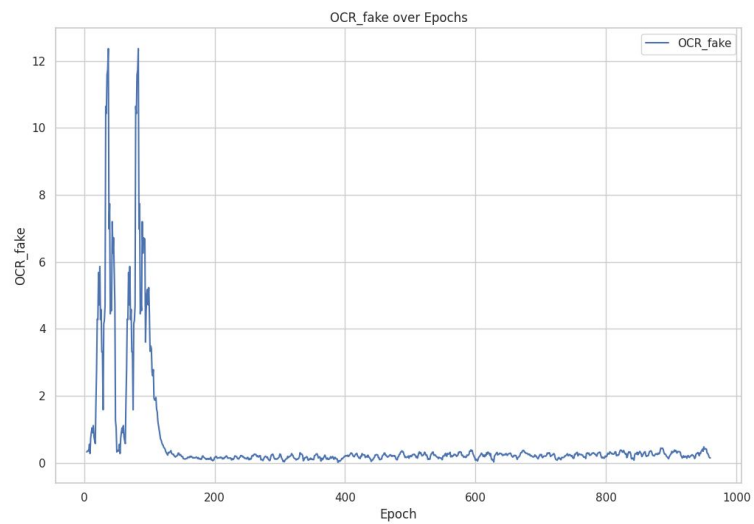
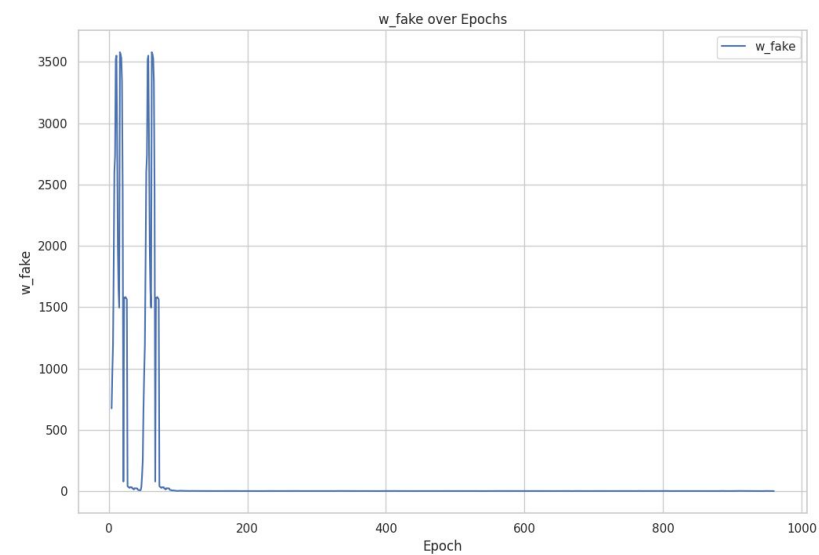
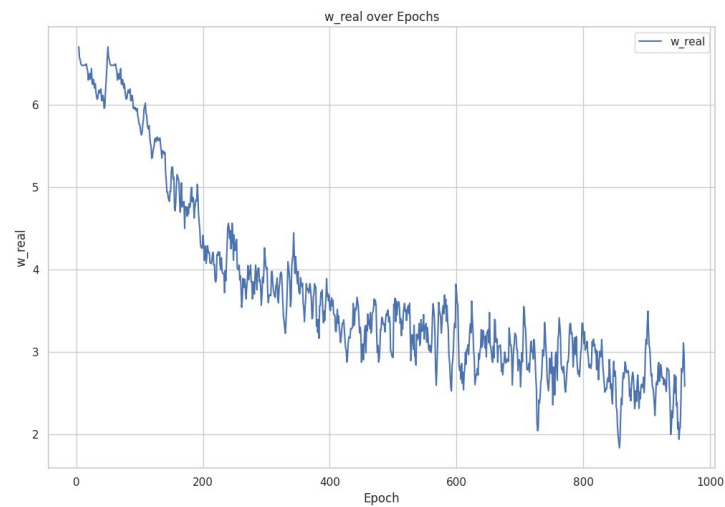


Flag

Use via API  · Built with Gradio 

Loss





Loss Analysis

- **loss-OCR_real: 0.9483**
 - OCR network loss on real data - measures the OCR network's ability to recognize real handwriting.
- **loss-w_fake: 1.0213**
 - Writer classification loss on fake data - indicates how well the classifier identifies the writer of generated handwriting.
- **loss-w_real: 2.3178**
 - Writer classification loss on real data - indicates how well the classifier identifies the writer of real handwriting.

Loss Analysis

- **loss-G: 4.4243**
 - Generator loss (G) - measures how well the generator produces realistic data.
- **loss-D: 0.0215**
 - Overall Discriminator loss (D) - measures the discriminator's ability to differentiate between real and fake data.
- **loss-D Fake: 0.0**
 - Discriminator loss on fake data (Dfake) - indicates how well the discriminator identifies generated data as fake.
- **loss-Dreal: 0.0215**
 - Discriminator loss on real data (Dreal) - shows how well the discriminator identifies real data as real.
- **loss-OCR_fake: 0.1757**
 - OCR network loss on fake data - measures the OCR network's ability to recognize generated handwriting.

Evaluation

While Evaluating the generated handwriting FID Score was found to be 112.55345241037465

Remaining Task

- To further train the model and check model accuracy.
- To diversify our dataset including alphanumeric dataset.
- To work on creating font of generated dataset.

References

- [1] E. Alonso, B. Moysset, and R. Messina, “Adversarial generation of handwritten text images conditioned on sequences,” in 2019 international conference on document analysis and recognition (ICDAR), IEEE, 2019, 481–486.
- [2] B. Davis, C. Tensmeyer, B. Price, C. Wigington, B. Morse, and R. Jain, “Text and style conditioned gan for generation of offline handwriting lines,” arXiv preprint arXiv:2009.00678, 2020.
- [3] A. K. Bhunia, S. Khan, H. Cholakkal, R. M. Anwer, F. S. Khan, and M. Shah, “Handwriting transformers,” in Proceedings of the IEEE/CVF international conference on computer vision, 2021, 1086–1094.

References

- [4] M. Zimmermann and H. Bunke, “Automatic segmentation of the iam off-line database for handwritten english text,” in 2002 International Conference on Pattern Recognition, IEEE, 4, 2002, 35–39.
- [5] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, “Generative adversarial networks: An overview,” IEEE signal processing magazine, vol. 35, no. 1, 53–65, 2018.
- [6] S. Fogel, H. Averbuch-Elor, S. Cohen, S. Mazor, and R. Litman, “Scrabblegan: Semi-supervised varying length handwritten text generation,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 4324-4333.