Triplet Loss : Offline Handwritten Signature Verification

A report submitted in partial fulfilment of the requirements
for the degree of B.Tech. Computer Science Engineering
with specialization in Design and Manufacturing

by

Utkarsh Tiwari (Roll No: 119CS0022)
Shubham Maurya (Roll No: 119CS0034)

Under the guidance of

Dr. K Nagaraju
(Assistant Professor CSE Department IIITDM Kurnool)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
INDIAN INSTITUTE OF INFORMATION TECHNOLOGY
DESIGN AND MANUFACTURING KURNOOL

April 2023

Evaluation Sheet

Title of the Project:

Name of the Students:

Examiner(s):

_____

_____

Supervisor/Advisor:

_____

Head of the Department:

_____

Date:

Place:

Certificate

We, **Utkarsh Tiwari** with Roll no: **119CS0022** and **Shubham Maurya** with Roll no: **119CS0034** hereby declare that the material presented in the Project Report titled **Offline Handwritten Signature Verification** represents original work carried out by us in the

With my signature, I certify that:

- I have not manipulated any of the data or results.
- I have not committed any plagiarism of intellectual property.I have clearly indicated and referenced the contributions of others.
- I have explicitly acknowledged all collaborative research and discussions.
- I have understood that any false claim will result in severe disciplinary action.
- I have understood that the work may be screened for any form of academic misconduct.

We have done this under Department of Computer Science And Engineering at the **Indian Institute of Information Technology Design and Manufacturing Kurnool** during the years **2022–2023**.

Utkarsh Tiwari (119CS0022) Shubham Maurya (119CS0034)

Date: 16/04/2023

In my capacity as supervisor of the above-mentioned work, I certify that the work presented in this Report is carried out under my supervision, and is worthy of consideration for the requirements of B.Tech. Project work.

Advisor's Name : **Dr. K Nagaraju** Advisor's Signature

Copyright Transfer Certificate

Title of the Project:

Name of the Students:

Copyright Transfer

The undersigned hereby assigns to the Indian Institute of Information Technology Design and Manufacturing Kurnool all rights under copyright that may exist in and for the above project report submitted for the award of the B.Tech. degree.

Student's Signature Student's Signature

Note: However, the author may reproduce or authorize others to reproduce material extracted in exactly the same words from the project report or a derivative of the project report for the personal use of the author given that the source and the Institute's copyright notice are indicated.

Plagiarism Report

This is to certify that B.Tech project report entitled **Triplet Loss: Offline Handwritten Signature verification** submitted by **Utkarsh Tiwari, Shubham Maurya:** holding Roll No. **119CS0022, 119CS0034** respectively under the supervision of **Dr K Nagaraju** in the **Department of Computer Science and Engineering** is an original research work done by the student.

The project report has been checked for the Plagiarism and the Plagiarism report is submitted along with the project report for further processing.

- Originality content (including the contents from own publications): _____%
- Similarity Reproduction of the content from other sources: _____%

We are aware that any issue related to plagiarism in future will have to be addressed by the student and the concerned supervisor.

Name of the Student: Student's Signature

Date:

Name of the Student: Student's Signature

Date:

Name of the Supervisor: Signature

Date:

Abstract

- Signature is the description of identity of the person can be used as identification. So there can always be chances that it can be forged. Hence, in order to determine whether a given signature is genuine or forged we need to be verify the signature. In the previous paper for verifying the Offine Handwritten Signature using, deep CNN connected to fully connected layer is used. They have used this as network of two different configurations: one of the part is used as feature extractor with hybrid classifier and another as an end to end classifier in Siamese Network. Hybrid classifier consist of support vector machine which is used for verifying the genuineness of the signature. Siamese network consist of two or more identical networks or sub-networks joined by one or more fully connected layers. Before this they have used simple deep CNN and constrasitive loss function and extract feature using different method to classify original/forgery signatures.

- In this project we have done work on Offline Handwritten Signature Verification problem. In this project we have discussed various approach using deep learning used for signature verification purposes. Here in this project we are using Triplet loss netwrok. It is one the best method for one shot learning. In earlier cases researches is done using CNN(Convolutional Neural Network) VGG-16 architecture using contrastive loss. Other approach in this field is using hybrid model of siamese network archuitecture using SVM(support vector machine) as classifier. But problem which come is related to large amount of dataset. Although tremendous accuracy have achieved but still problem is their of large dataet. On large dataset both model get failed, Which is one of the motivation of us. Second main problem is accuracy in case of dataset is writer independent. Siamese Netwrok paper achieve massive accuracy in writer dependent dataset which is of in 99.78.But in case of writer independent dataset it is about 80 percentage, while the approach we proposed which is of one shot learning no need for large dataset and have accuracy of 85 percentagte in writer dependent model.

Acknowledgements

We would like to express our sincere gratitude to **Dr. K Nagaraju**, Assistant professor, Department of Computer Science Engineering for motivating and supervising me all through this project. We would like to thank for his timely guidance and support in completion of the project.

The opportunity to join the **IIITDM Kurnool** is an achievement for us. We are thankful to **beloved god and our parents** for the grace and motivation they gave to join the college and complete the degree. Sincerest gratitude to all the faculty members of **IIITDM Kurnool** those who taught us academics and life lessons.

We are grateful to our institute, **Indian Institute of Information Technology Design and Manufacturing Kurnool**

Contents

List of Figures

Abbreviations

SNN	**S**iamese **N**eural **N**etwork
CNN	**C**onvolutional **N**eural **N**etworks
VGG	**V**isual **G**eometery **G**roup
LSTM	**L**ong **S**hort **T**erm **M**emory
TNA	**T**riplet **N**eural **A**rchitecture
WI	**W**riter **I**ndependent
WD	**W**riter **D**ependent
Relu	**R**ectified **L**inear **U**nit
SN	**S**iamese **N**etwork
SVM	**S**upport **V**ector **M**achine

Dedicated to parents, all the IIITDM professors and students who guided and helped us during the course of this project ...

Chapter 1

Introduction

One of the realistic authentication methods that is widely used for personal identity is verifying signatures. The signature verification procedure in many business settings, such as the payment of a bank cheque, is predicated on a single known sample being human examined.
Basically till now in Offline Handwritten Signature Verification work is done using two approaches which are using normal CNN using VGG-16 architecture, second is Siamese network. We are using is Triplet network, so regarding all under consideration we have worked for best model. Offline handwritten signature verification using CNN, Siamese, and triplet networks is a popular approach to verifying signatures that have been written on paper or another physical medium. These networks are powerful machine learning algorithms that can learn to recognize the unique features of an individual's signature and compare them to a reference signature to determine if they match.

CNN is a type of deep learning algorithm commonly used in image recognition tasks. For verification of signatures, CNN can be trained on a dataset of signatures to learn to recognize the unique features of each signature, such as the shape of the letters and the stroke patterns. Once the CNN is trained, it can be used to classify new signatures as either genuine or fraudulent.

A SN is one of neural network architecture, which is used for signature verification by differentiating two signatures and determining how similar they are. The network takes in two images of signatures and learns to produce a similarity score, which can be used to determine whether the two signatures match or not.

A variation of the Siamese network that can compare three signatures at once is Triplet Network. In a triplet network, the network takes in three images: an anchor, a positive (a genuine signature from the same individual as the anchor image), and a negative image (a fraudulent signature from a different individual). The network learns to produce a similarity score between the originalanchor and positive images that is higher than similarity score between the originalanchor and negative images.

Overall, these networks can be effective for offline handwritten signature verification because they can learn to recognize the unique features of a signature and compare it to a reference signature to determine if it is genuine or not. However, the performance of these networks can depend mainly on the length,quality and size of the training dataset, as well as the specific architecture and parameters of network used.

1.1 Handwritten Signature

Since ancient times, handwritten signatures have served as a means of identification and authentication. The following are some benefits of using handwritten signatures:

- Unique: It is possible to recognise each person by their individual signature.
- Personal: Because they serve as a trustworthy means of identification verification and are regarded as personal, signatures.
- Authenticity: Handwritten signatures are trusted forms of verification since they are challenging to counterfeit.
- Legally bound: Signatures can be used to ratify contracts, agreements, and other documents since they are legally binding.
- Cost-effective: Signatures are a simple and inexpensive way to verify someone's identity.
- Handwritten signatures are used for verification in a variety of disciplines, including:
- Banking and finance: Checks, loans, and other financial transactions are verified by signatures.
- Legal: In order to legitimise contracts, agreements, and other legal papers, signatures are employed in legal documents.

- Government: Passports, driver's licences, and voter registration forms all require signatures for verification.
- Medical: Consent forms and medical records, among other papers, are verified using signatures.
- Business: Contracts, agreements, and other documents are verified by signatures.
- Overall, handwritten signatures have been used for centuries and will continue to be utilised in many different industries in the future as a trusted and dependable form of verification[1].

1.2 Approaches for Handwritten Signature Verification

There are various approaches used for signature verification. Here are some of the most common ones:

- Graphology: Graphology is the study of handwriting and can be used to analyze a person's signature to verify their identity. This approach involves analyzing the shape, size, and pressure of the signature to determine whether it is authentic.
- Expert verification: This approach involves having an expert compare the signature to a known signature of the individual to verify its authenticity. The expert may use various tools, such as a microscope or magnifying glass, to examine the signature closely.
- Biometric signature verification: This approach involves using biometric technology to verify a signature. Biometric signatures capture data such as pressure, speed, and stroke order to create a unique signature profile for an individual. This profile can then be used to verify the individual's signature in the future.
- Machine learning-based approaches: This involves training a computer algorithm on a set of signatures to recognize the unique characteristics of an individual's signature. These algorithms can then be used to verify signatures automatically.
- Feature-based approaches: Feature-based approaches involve extracting features from the signature, such as stroke width and curvature, and comparing them to a reference signature to determine whether they match.

In order to assure accuracy and dependability, signature verification is a complicated procedure that calls for a mix of many strategies. Depending on the particular needs of the application, the resources at hand, and the needed level of security, many strategies could be employed.

1.3 Signature Verification Using Machine Learning

The practise of confirming a handwritten signature's legitimacy is known as signature verification. In order to identify forgeries or forge signatures, various number of signature verification techniques have been developed since invention of machine learning. These are a few typical machine learning approaches for verifying signatures:

1. Method based on feature extraction: In this approach, machine learning techniques are employed to featuresfrom a signature is extracted, like its size, shape, speed, and pressure. The authenticity of the signature is then determined by comparing these qualities to a database of real signatures.

2. Artificial neural networks are used in the neural network-based approach to identify if a signature is real or forged. In order to identify the patterns and features that are typical of real signatures, a neural network is trained on a dataset of real signatures. The network can be used to categorise fresh signatures as legitimate or false after it has been trained.

3. Method using a (SVM): It is a well-liked machine learning algorithm for signature verification. It is the way of supervised learning system which learns how to distinguish between real and fake signatures using a training dataset of real ones[2].

4. Method based on deep learning: To analyse signature photos, Recurrent neural networks and other deep learning techniques are used in this strategy. Since they were developed using a dataset of real fingerprints, these algorithms are capable of identifying even subtle differences between genuine and fake signatures. Method using an ensemble of machine learning techniques to increase the precision of signature verification. To categorise signatures, ensemble methods combine algorithms like SVM, neural networks, and feature extraction-based techniques.

Overall, machine learning has greatly increased the precision of signature verification techniques, enabling the detection of even minute variations between real and false signatures. These techniques can be used in a number of fields where signature verification is required, including banking, law, and government.

1.4 Convolutional Neural Network(CNN)

CNN stands for Convolutional Neural Network, which is a algorithm in deep learning, it is particularly effective for image classification, object detection, and recognition. It is based on the principles of convolution, pooling, and non-linear activation functions to extract and analyze the features of an image.

Here is a brief overview of the key concepts and formulas used in CNNs:

Convolution: It involves mathematical operation. In order to create a third function, two functions must be combined in this. It is a technique which is used in CNNs to extract features from a picture. A tiny matrix, referred as as a filter or kernel, is dragged across the input picture to execute the process, and the dot product between the filter and the corresponding pixels in the image is computed. A feature map that emphasises the image's key elements is the result of this procedure. The formula for 2D convolution is:

$$Y(i,j) = \sum\sum X(m,n) * K(i-m, j-n) \tag{1.1}$$

where: Y(i,j) is output feature map at position (i,j) X(m,n) is input image pixel at position (m,n) K(i-m, j-n) is the filter value at position (i-m, j-n)

Pooling: A downsampling method called pooling aids in removing the most important details from the image while minimising the size of the feature maps. Max pooling and Average pooling are the two different types of pooling. In Average pooling determines the average value, max pooling keeps the highest value found in each area of the feature map. The maximum pooling formula is:

$$f(i,j) = max(L(2i,2j), L(2i,2j+1), L(2i+1,2j), L(2i+1,2j+1)) \tag{1.2}$$

where:
f(i,j) is the output of the max pooling operation at position (i,j)
L(2*i,2*j), L(2*i,2*j+1), L(2i+1,2*j), L(2*i+1,2*j+1) are the input feature map values within the pooling region
Non-linear activation functions: They are used to introduce non-linearities into the CNN and enable it to learn complex features. The most commonly used activation functions in CNNs are ReLU (Rectified Linear Unit), sigmoid, and tanh.
RELU is the activation function defined as:

$$f(x) = max(0, x) \tag{1.3}$$

where: f(x) is the output of the ReLU function
x is the input to the function.
Backpropagation: The backpropagation technique is an algorithm employed to modify the biases and weights in a convolutional neural network (CNN) to decrease the loss function. For the calculation of the discrepancy we use loss function between the CNN's expected and actual output. The CNN's weights and biases are updated in accordance with the output of the loss of gradient function as determined by the backpropagation method.

The formula for the gradient/slope of the loss function with respect to the weights is:

$$dL/dW = dL/dY * dY/dW \tag{1.4}$$

Terms used:

- Here term dL/dW is the slope of the loss function with respect to the weights
- Here term dL/dY is the slope of the loss function with respect to the output of the CNN
- Here term dY/dW is the slope of the output of the CNN with respect to the weights

1.5 Signature Verification Using CNN

Convolutional Neural Networks (CNNs) have excelled at recognising and classifying images, including verifying signatures.
Here are a few typical methods for employing CNNs to verify signatures:

1.Image-based method: In this method, the CNN is trained on the unprocessed picture of the signature to identify the characteristics that separate authentic from fake signatures. Several convolutional and pooling layers are often followed by fully linked layers in the CNN architecture. A binary classification result indicating whether the signature is real or fake is the CNN's output.

2.Approach based on individual strokes rather than the complete image of the signature: In this method, the CNN is trained on the individual strokes that make up the signature. LSTM network is usually followed by a series of convolutional and pooling layers in the CNN design. The CNN generates a series of classification results that show whether each stroke is real or fake.

3.Hybrid technique: In this approach, both image-based and stroke-based features are used to improve accuracy of signature verification. The CNN commonly combines convolutional and pooling layers for feature extraction from images, and LSTM layers for feature extraction from strokes. A binary classification result indicating whether the signature is real or fake is the CNN's output.

4.Multi-modal approach: In this method, the CNN is fed multiple modalities, including pressure, velocity, and acceleration, to capture the distinctive properties of the signature. Convolutional and pooling layers are frequently combined in the CNN design, which is then followed by fully connected layers. A binary classification result indicating whether the signature is real or fake is the CNN's output.

By extracting and examining the most important aspects of the signature, CNNs provide an efficient method for verifying signatures. CNNs can be customised to specific applications and datasets using a variety of techniques and architectures, resulting in excellent efficiency and accuracy in the verification of digital signatures.

1.6 Previous Work Over Offline Signature Verification

VGG-16 architecture in Offline Signature verification:

The VGG-16 architecture is a deep CNN that has been primarily used for image recognition tasks, including recognition of object and image classification. However, it is also be used for signature verification[1].

Signature verification is a process of authenticating a person's signature by comparing it to a reference signature. It involves extracting features from the signature and comparing them with those from the reference signature. The VGG-16 architecture can be used to extract features from signatures that can then be used for verification.

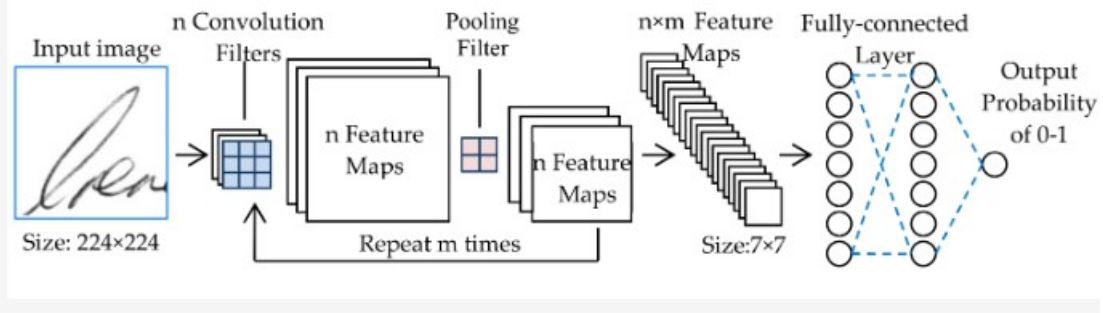

FIGURE 1.1: Signature Verification using CNN. [1]

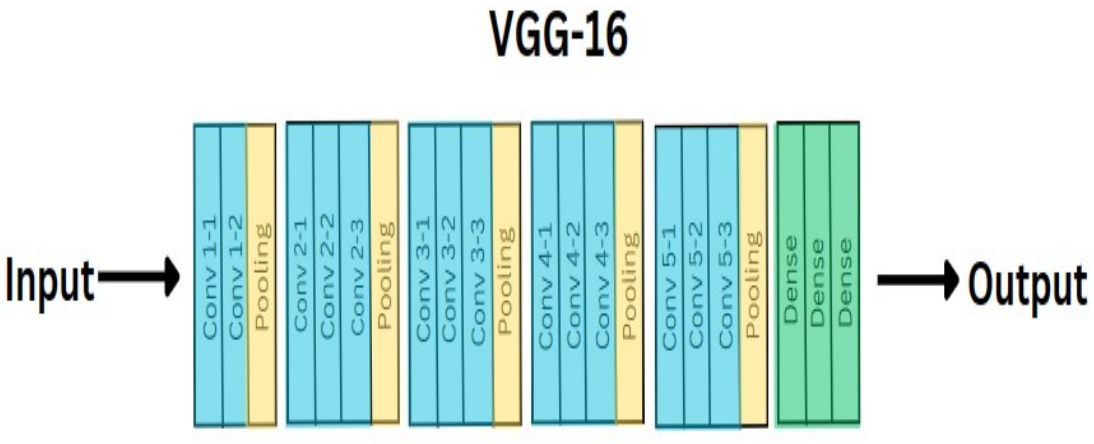

FIGURE 1.2: VGG-16

To use VGG-16 for signature verification, the input signature image is first preprocessed to enhance its quality and remove any noise. The preprocessed image is then fed into the VGG-16 network, it is consists of 3 fully connected layers and 16 convolutional layers.

In figure 1.1 it shows how CNN work for signature verification purposes so that we can understand internal working of VGG-16.

By using a number of convolutional filters, the convolutional layers of the VGG-16 network extract features from the input signature image. The fully connected layers then execute classification using the extracted features after being provided the extracted features.

In figure 1.2,For signature verification, the final layer of the VGG-16 network is changed with a softmax layer, which outputs the probability of the input signature being genuine or forged. During training, the VGG-16 network is trained on a dataset of genuine and forged signatures and the softmax layer is trained to minimise the discrepancy between the predicted and actual labels using a dataset of real and fake signatures[1].

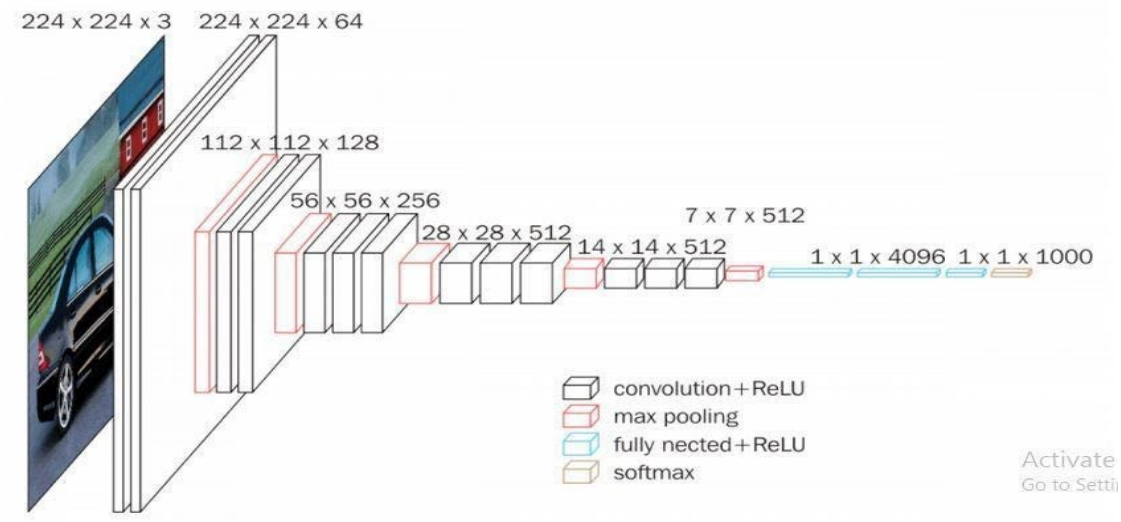

FIGURE 1.3: VGG16 architecture
ref:- ResearchGate VGG-16

In summary, the VGG-16 architecture in figure 1.3 can be used for signature verification by extracting features from the input signature image and classifying it as genuine or forged using a softmax layer.

WRITER DEPENDENT MODEL:

This approach is built on writer-dependent features that are determined based on a score calculated for each feature of the associated writer, leading to the selection of various

sets of features for various writers[3].

Feature Extraction:

- **Ratio** It is the ratio of the signature image's surface area to the signature's surface area contained within a bounding box. The area of a signature is the sum of all of its individual pixels.
- **Centroid** You must compute the average coordinate in order to determine the centroid of a collection of pixels in a binary image. To determine the weighted average position for a grayscale image, use the grey values of the pixel.
- **Eccentricity** The smallest distance between a particular vertex (v) and any other vertex (w) in a linked graph is known as eccentricity. When computed for each vertex, it transforms the graph's connection structure into a set of values. A linked region of a digital image's of nearby connected graph and the specified metric have been used for explaining it.
- **Solidity** In relation to the region's convex hull, solidity is a proportion of the area. . Solidity is what fraction of the actual area your region is.
- **Skew** Skewness is a measure of (lack of) symmetry .
- **Kurtosis** For kurtosis first we have to find mean(or expected value) of the normalized data and then increase it to the fourth power.

- Considering S as Summation and A as Acitvation

$$S = x1 * W1 + x2 * W2 \tag{1.5}$$

$$A = exp(x)/1 + exp(x) \tag{1.6}$$

- Using chain Rule:

$$dE/dW1 = dE/dA * dA/dS * dS/dW1 \tag{1.7}$$

$$dE/dW2 = dE/dA * dA/dS * dS/dW2 \tag{1.8}$$

1.7 Neural Networks Architecture for Image Similiarity

In Siamese Network two identical sub-networks with the same weights and architecture is their. It takes two inputs and computes a similarity score between them. Siamese networks are often used for tasks such as image or face verification, where the goal is to determine whether two images or faces belong to the same person or not.

A triplet network, on the other hand, takes in three inputs - first is an anchor image or text, second is a positive image or text that is similar to the anchor, and a negative image or text which is disimiliar to the anchor image.

When compared to the distance between the orignalanchor and the positive samples, the distance between the originalanchor and the negative samples is increased. When retrieving related objects from a sizable database for tasks like picture or text retrieval, triplet networks are frequently employed.

Both Siamese and triplet networks are examples of deep learning architectures that can learn complex representations of data and perform tasks such as similarity matching or clustering[2].

- **Siamese Neural Network:**

  It is a type of neural network architecture which can be used for signature verification tasks. In this architecture, two identical neural networks are trained on pairs of images, where each image is a signature.

  In figure 1.4, The two networks have the same weights and are connected by a similarity metric that compares the features extracted from the two images. This similarity metric can be defined as the Euclidean distance or the cosine similarity between the features.

  During the training process, the Siamese Neural Network learns to discriminate between original and manipulated signatures by using or learning the similarity metric between pairs of signatures. Once the network is trained, it can be used for signature verification by comparing the features extracted from two signatures and determining their similarity.

  In summary, the SNN is a powerful tool for signature verification, as it can accurately classify signatures and detect forgeries.

  A Triplet Neural Network is one of the type of neural network commonly used for face recognition, but can also be applied to signature classification. The basic idea behind

a triplet network is to learn a feature space where the distance between embeddings of different classes is maximized, while keeping the distance between embeddings of the same class minimized.

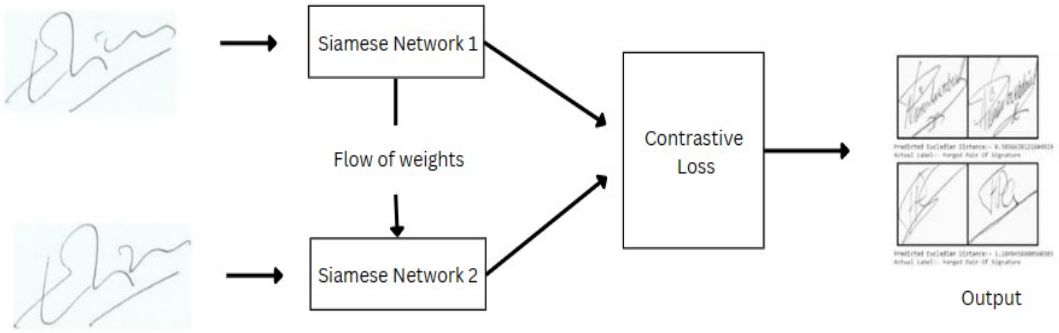

FIGURE 1.4: Siamese Network

In the context of signature classification, a triplet network takes three inputs: an anchor signature, a positive signature (from the same person as the anchor), and a negative signature (from a different person). The network gains the ability to map each signature to a feature space where the distance between the original anchor and positive to anchor and negative signatures is minimised and maximised, respectively[4].

- During training, the network adjusts its weights to reduce a triplet loss function, it is designed to ensure that the distances between the originalanchor and positive signature is lesser than the distance between the orignial anchor and negative signature by a predefined margin. Once trained, the triplet network can be used to classify new signatures by comparing their feature representations to those of known signatures.
- Overall, the triplet neural network is a powerful approach for signature classification, as it allows the network to learn robust features that are specific to individual signatures and can better discriminate between genuine and forged signatures.

– function f find similarity in i,j by taking (i,j) as query images
filter is g_z

$$f_z(x) = 1/(n^2) \sum \sum g_z(x, i, j) \tag{1.9}$$

- To Formalise the requirement of the loss will be defined over triplet(a,p,n) is:

$$L = max(dis(a, p) - dis(a, n) + margin, 0) \tag{1.10}$$

- Here we have to minimize the loss which pushes dis(a,p) to 0 to be greater than dis(a,p)+margin.As soon as n becomes an easy negative,the loss function become zero.

- **Literature Survey:** Here we mainly have 5 papers in Literature survey which we are using as our base papers.These papers help us lot in understanding the previous work which is performed previously in Offline Signature Verification and what are drawbacks are their in this papers which we were taken as motivation of our work.By taking that we proceed further in this.

- **An Offline Signature Verification and Forgery Detection Method Based on a Single Known Sample and an Explainable Deep Learning Approach. Applied Sciences[1]:**

  One of the commonly utilized biometric methods for personal identification is signature verification. In various business settings, including payment in bank using check, the verification process of signature heavily relies on human scrutiny of a sole reference sample.

  Despite the abundance of studies on signature verification, there have been limited efforts to conduct verification solely relying on a single signature sample. In this approach, the authors employ an offline handwritten signature verification technique using deep learning in explainable approach (deep CNN, DCNN) and a distinct localised feature extraction method.

  The owner uses the open-source ICDAR 2011 SIGCOMP dataset to train their algorithm and determine whether a query signature is authentic or fake. They use samples from an unknown author, whose signatures were not included in the training or any other steps, to do their tests[1].

  According to the experiments, they got the accuracy of 94.37% to 99.96%, on false rejection rate(FRR) between 5.88% and 0%, false acceptance rate (FAR) between 0.22% and 5.34% in their testing dataset. But the limitation of this paper is they are failed on using large datasets it takes huge time to train or even sometimes model get crashed.

- **Offline signature verification using a region based deep metric learning network [3]**

  In document forensics, handwritten verification of signature is a popular realistic authentication for confirming a proof of identity. Despite the enormous research efforts, offline signature verification is still difficult, especially when distinguishing skilled forgeries from genuine signatures because there may be less of a visual difference between a skilled forgery and a genuine signature. In a situation where each writer has extremely few training samples, this problem is considerably more pressing.

  The Deep Convolutional Siamese Network proposed in this paper, uses metric learning and is suitable to both (WD) and (WI) scenarios.

  In order to express minuscule yet discriminative information, a Mutual Signature DenseNet (MSDN) is designed to extract features and learn the similarity measure from small sections instead of whole signature photos.

  The proposed approach merges the similarity scores of various regions, based on local region comparison, to make the ultimate verification decision. The experimental outcomes on the publicly available datasets CEDAR and GPDS display that the suggested method surpass the previous models which has state-of-art results with an EER of 6.74 percent and 8.24 percent in the WI case, respectively, and 1.67 percent and 1.65 percent in the WD case, respectively. However, it should be noted that the model still has limitations and fails to perform well on larger datasets. Despite attempting to enhance the model architecture, it faces similar constraints as the prior models.

- **IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)[2]:**
  A signature serves as a representation of an individual's name and serves as a form of identification, however, it is susceptible to forgery. As a result, the verification of a signature's authenticity is necessary. In this study, a 3-layer deep CNN or (DCNN) with fully connected layer are used to verify offline signatures. The number of parameters used for training in this proposed method is considerably lower than what has been previously reported in the literature.

  In two alternative configurations this network has been applied: first, it is used for extracting features in a hybrid classifier, and second, as an end-to-end classifier in

a SNsiamese network. Support vector machines are employed in hybrid classifier schemes to validate the validity of the signature.

The neural network employed in this research, which is a combination of fully connected layers and identical subnetworks, is incorporated as a subnetwork in a Siamese network. The study evaluates both the hybrid classifier and Siamese network in the context of Writer independent (WI) and dependent (WD) verification using three datasets: CEDAR, GPDS Synthetic Signature, and BHSig260.

With the Hybrid classifier scheme for signature verification, the suggested technique in this study achieved excellent accuracy rates of 99.91%, 92.28%, 86.88%, and 90.58%, and with the Siamese network for writer-dependent verification, it reached 99.87%, 92.14%, 86.33%, and 90.80%. On the CEDAR, BENGALI, HINDI, and GPDS datasets, the Siamese network attained accuracy rates of 80.26%, 75.06%, 89.33%, and 82.42%, respectively, in terms of writer-independent verification.

The results of the experiments demonstrate that the suggested approach for handwriting identification in HINDI outperforms the current cutting-edge techniques in WI. The method has also shown equivalent results in WD verification for the CEDAR and BENGALI datasets.

The paper attains a remarkable 99.91% precision for writer-dependent scenarios, however, it falls short at 80.26% for the WI dataset. Taking (WI) case as motivation, we work in this field and achieve better than them.

- **Collaborative human-ai (chai):evidence-based interpretable melanoma classification in dermoscopic images[4]:**

The diagnostic performance of automated dermoscopic image analysis has rapidly progressed. However, the resistance to its adoption persists partly due to the lack of evidence supporting the decision-making process. This study introduces an evidence-based classification approach, which utilizes a triplet network and a CNN. For classification purposes, the KNN method is employed. To ensure the accuracy and human like similarity or human visual similarity of the outcomes for all skill levels, a unique hierarchical triplet logic is employed to jointly learn an embedding based only on disease labels. The findings indicate an enhancement in results compared to baselines trained solely on disease labels and the standard multiclass loss.

1.8 Organisation of the Project Report

Chapter-1

In this chapter it is explained about Introduction to Offline Signature Verification and previous work perform in the Offline Signature Verification.Later in this chapter we were discuss about the Literature Survey of our papers. After then we were disuss about how VGG-16 structure,Siamese Network,Tripet Network used for signature forgery detection and how they are improved with change in Neural Network Architecture.

Chapter-2

Data sets and Libraries present in the project are explained in detailed by briefing all the concepts are involved.

Chapter-3

Model Architecture which we are used are described in detail, as well as the proposed model.

Chapter-4

Results of the project and the outcomes of the method are shown in detail.

Chapter-5

Conclusion and Future work are discussed in this chapter so that the further improvements can be carried in this project.

Chapter 2

Datasets, Libraries and Previous Works

2.1 Datasets:

2.1.1 CEDAR Signature Dataset

We have used CEDAR Signature Dataset.In VGG-16 architecture as well as simaese we have divide it into two type of datasets which are below. There are two type of model of data which is as Writer Dependent Model and another is Writer Independent Model.

Writer Dependent Model: Here training dataset is on both forgery and genuine signatures. Testing is also done on both forgery and genuine signatures.

Writer Independent Model: Here Training is Done on only genuine signature. While testing is done on both genuine and forgery signature.

In Triplet loss we have used same CEDAR dataset but it is only contain two type of data one is genuine image folder and another is forgery image folder. Both genuine and forgery image has 55 signers and each persons 24 signature in both forgery and genuine.So in total we have 1320 signatures in our dataset both folders.

2.2 Libraries:

2.2.1 Keras

A Python-based high-level neural network building API called Keras can be used to build neural networks on top of TensorFlow, CNTK, or Theano. It is made to make rapid experimentation and deep learning model prototyping possible.

In various tasks such as detection of object, picture classification, and natural language processing, Keras offers a number of pre-trained models available. Examples are ResNet50, InceptionV3, VGG16, and BERT. Using the applications module of Keras, these models may be loaded quickly. Cross Entropy:

$$Loss = -1/(output) \sum_{i=1}^{output_{size}} y_i log y_i + 1 - y_i log 1 - y_i (2.1)$$

2.2.2 Open-CV

A free and open-source computer vision and machine learning software library is called OpenCV. It is frequently used for image and video processing, object detection, face recognition, and other computer vision applications. It was first created by Intel in 1999, and eventually evolved into a cross-platform library that was compatible with Windows and Linux.

2.2.3 Tensorflow

It is a library that is used for the numerical computation using data flow graphs where:

- The nodes in the graph represents mathematical operations.
- The edges in the graph represent the tensors that are communicated between them. The tensor is the central unit of the data in TensorFlow.

  Here are some of the commonly used formulas in TensorFlow:

- Gradient descent: This formula is used to minimize the loss function in a neural network model. The basic formula for gradient descent is:

$$x = x - alpha * J(x) \tag{2.2}$$

where x is the weight parameter, is the learning rate, and J(x) is the gradient of the loss function J(x) with respect to x.

- Cross-entropy loss:In categorization issues, this loss function is frequently employed. The cross-entropy loss formula is:

$$H(p, q) = -\sum[p(x) * log(q(x))] \tag{2.3}$$

where p is the true probability distribution, q is the predicted probability distribution, and x is the set of all possible outcomes.

- Softmax function: This is a popular activation function used in the output layer of a neural network for multi-class classification problems. The formula for the softmax function is:

$$y_i = e_i^z / (\sum_1 k)(e_j^z) \tag{2.4}$$

where K is the number of classes, zi is the output of the ith neuron in the output layer, and yi is the predicted probability of the i-th class.

- Convolution operation: This is a commonly used operation in convolutional neural networks for image processing. The formula for the convolution operation is:

$$f(i, j) = \sum(m, n)x(i - m, j - n) * h(m, n) \tag{2.5}$$

where x is the input image, h is the filter or kernel, and y is the output feature map.

These are just a few examples of the formulas used in TensorFlow. The library provides many other functions and operations for building and training machine learning models.

2.3 Previous Works:

Other than this We have also used siamese and VGG-16 strucutre in begining for understanding about how signature verification work

2.4 Siamese network Architecture

Siamese Network Architecture consists of two or more sub-networks that have similar topologies, the same weights, and the same parameters. The Siamese twins, who are frequently identical and cooperate well, gave their name to the architecture that connects these sub-networks using one or more completely linked layers.

This network architecture is frequently applied in tasks requiring similarity or dissimilarity measurements, such as signature verification, where it may be used to compare two signature photos and determine whether they belong to the same person. Individual inputs are given to the sub-networks, which subsequently produce embeddings. These embeddings are then compared using metrics like cosine similarity or Euclidean distance..

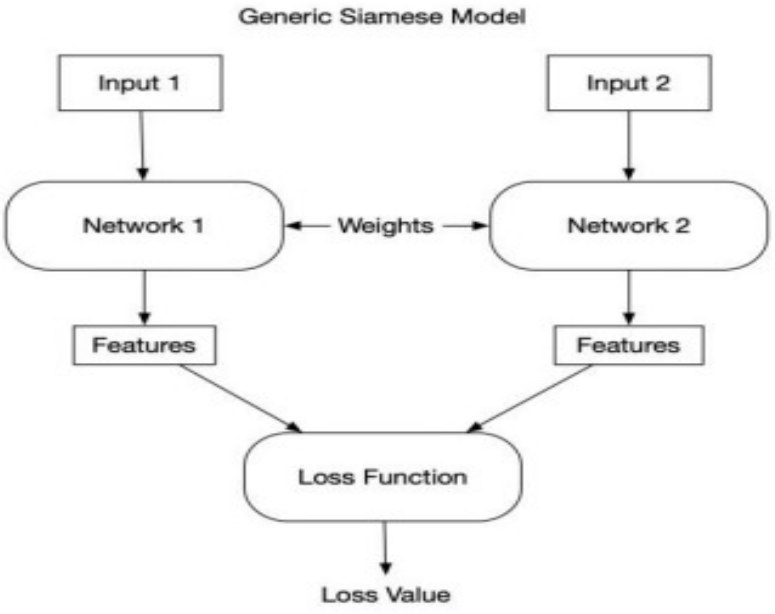

FIGURE 2.1: Siamese Network

Siamese networks have several advantages for signature forgery detection:

They can learn complex patterns: Siamese networks can learn to detect subtle differences between genuine and forged signatures, even if they are not immediately apparent to the human eye. This is because they are trained to identify similarities and differences between pairs of signatures, and can therefore capture complex patterns that may be difficult to articulate or quantify explicitly.

Chapter 3

Proposed Method

The proposed model consists of a Triplet loss using Neural Network which is most famous in image dissimilarity classification.So we used this as fro finding ditance of test image to samples of genuine person's signature. If it's distance is less than than original image then it is original image otherwise person's image is not of genuine person's image.

Also it takes small training datasets: Triplet networks require only a few examples of each class during training. This is because they use a metric learning approach that learns to compare the similarity between images. Therefore, even if there are only a few images per class, the network can still learn to classify them accurately.

Note : One main purpose of using triplet loss is, it is better than other loss functions for tasks such as image and face recognition because it is able to capture more information about the relationships between different samples, which allows it to learn a more discriminative representation of the input data.

It is one shot learning because it compares distance between original(anchor) to given image sample and then we compare distance of this to other original images of sample of original person.

Here we are using resnet-50 model for training instead of VGG-16 because of Vanishing Gradient Problem.

Vanishing Gradient: As we know gradients are multiply in each layer whenever it is transfer.So in case of linearly connected models such as VGG-16 it is get vanished at end. Because it's value at starting layer itself in decimal which is less than 1. So at last

layer which is 16th layer it is almost 0.

But in case of resnet-50 architectures every other layer are directly connected so gradients are directly pass from first to last layer without get vanished.

Allows for fine-grained classification: Triplet networks are designed to learn subtle differences between images. This makes them ideal for fine-grained classification tasks, where there may be small differences between images of the same class.

Embedding-based approach: Triplet networks learn an embedding for each image, which is a compact representation of its features. This allows for efficient storage and retrieval of images, making them useful for tasks like image search and retrieval.

Transfer learning: Triplet networks can be used as feature extractors, allowing them to be fine-tuned for other related tasks. This makes them useful for transfer learning, where a pre-trained model can be used as a starting point for a new task.

We are using threshold as Otsu threshold in our paper. Otsu's thresholding is a widely used image processing technique for automatic threshold selection. It is a form of image segmentation, which is the process of dividing an image into multiple regions or segments with similar characteristics.

The Otsu's thresholding method works by finding the threshold that minimizes the intra-class variance of the pixels in the image. The basic idea is to divide the image into two classes, foreground and background, based on the threshold value. The threshold is chosen to maximize the difference between the inter-class variance and the intra-class variance. In other words, the threshold is chosen such that the foreground pixels are as different as possible from the background pixels.

3.1 Model Architecture

3.1.1 Triplet Network

It is Deep Metric Learning for Signature Verification.

This architecture's one-shot learning is its greatest asset.which is, as you are aware, facial recognition and signature verification for employees who work for corporations or other companies.That's why it is one shot learning asset.Every time we add new photographs to the database, we must retrain the model.The advantage of this is that you no longer need to train after adding an image to the dataset in the database.Large dataset training is a challenge for both the Siamese network and the VGG-16.Thus, we attempt to apply this to our project.

Most often, triplet networks are used for similarity learning. In order to understand about similarity, one must first determine how similar or different two different things or samples are. Three samples must be entered into the triplet network: an anchor sample, a positive sample, and a negative sample. The positive sample is a sample that resembles the anchor sample, whereas the anchor sample acts as the reference sample against which

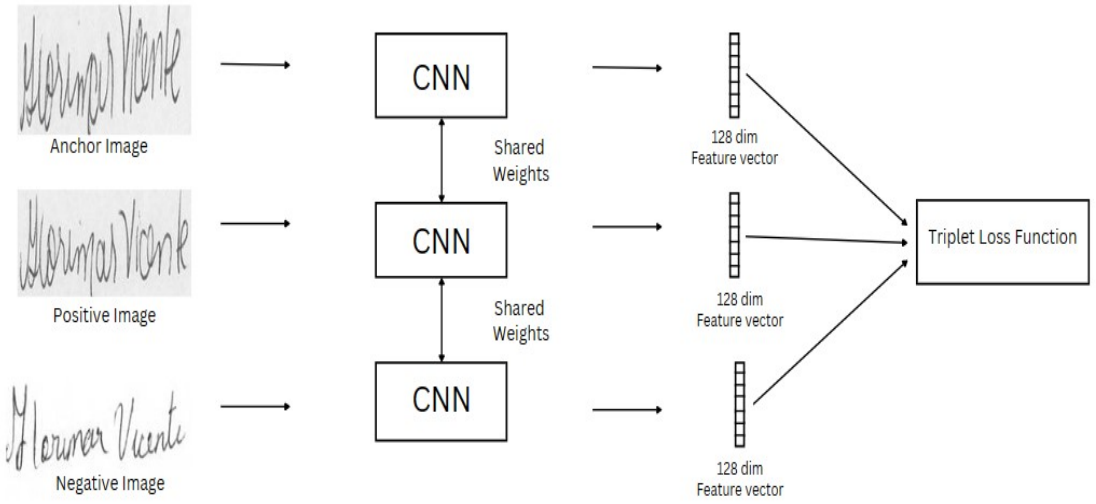

FIGURE 3.1: Our Triplet Network Architecture

the similarity is assessed. The negative sample, on the other hand, is a sample that differs from the anchor sample.
The goal of the triplet network is to develop a mapping function that can map the input samples into a high-dimensional feature space where the anchor and positive samples are located near to one another and the negative sample is located further away.

$$Loss(A, P, N) = max(|f(A) - f(P)|^2 - |f(A) - f(N)|^2 + del, 0) \quad (3.1)$$

The term 'del' refers to the margin used to increase the distance between similar and dissimilar pairs in a triplet. The feature embedding for anchor, positive, and negative images are denoted as Fa, Fp, and Fn, respectively. Fine-tuning the model can be beneficial to adjust the pretrained weights according to our data distribution. This is important to overcome the differences between the training data of the pretrained model (natural photos) and the data related to our task (signature images). Additionally, the pretrained model seeks to maximise the separation between various classes by learning linearly separable features prior to the softmax fully connected layer because it was initially trained for a classification task using cross-entropy loss.

The goal is to establish features that may effectively differentiate between two data points by highlighting the similarities and contrasts between them. This can be done by learning a distance metric that emphasises how important it is to minimise the distance between similar items and maximise the distance between objects that are distinct in a certain metric space. It is intended to provide representations that are widely dispersed for dissimilar objects and strongly clustered for related objects.

3.1.2 Metric learning

What is Metric Learning?

- Metric learning involves three type of images which are :
- Original or Anchor image,
- a Positive image
- a Negative image.

Metric learning is a type of machine learning approach that is specifically designed to learn a distance function between samples. The fundamental idea is that by learning a distance measure between similar and dissimilar objects, the model can be applied to tasks such as signature verification, where this distance measure is critical. In other words, metric learning involves training a model to differentiate between similar and dissimilar objects by learning a distance metric that can be used to measure the similarity between two objects.

The formulation of the metric learning objective may vary depending on the structure of the training dataset.

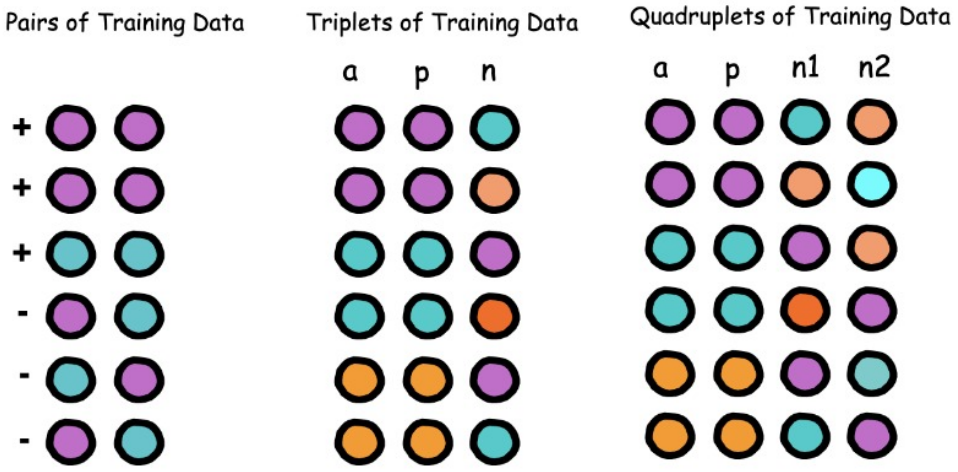

FIGURE 3.2: Model Working

Chapter 4

Results and Discussion

4.1 Seperated data by Labeling:

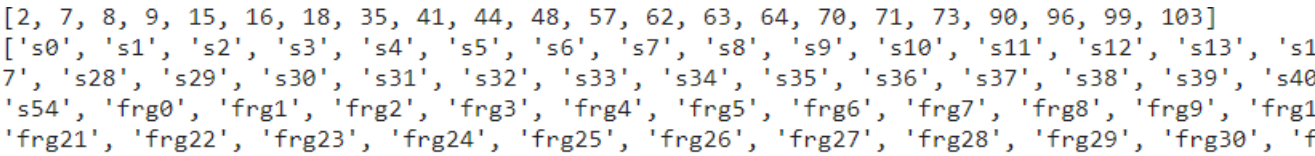

FIGURE 4.1: Labelling

4.2 Case 02:

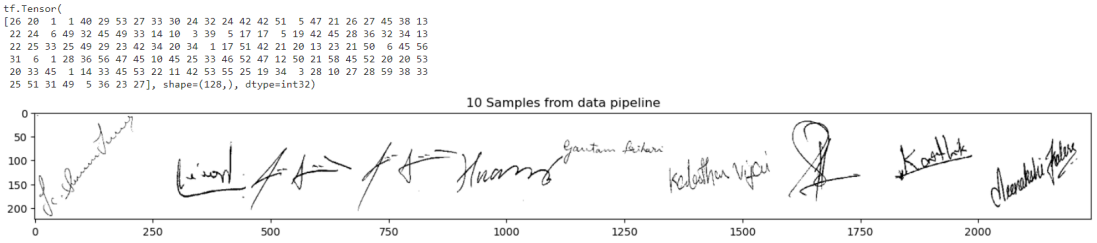

FIGURE 4.2: checking samples

From the Test Case 2, model performs moderately due to unsufficient data related spraying noise model confuses with sizzling food.

4.3 Case 03:

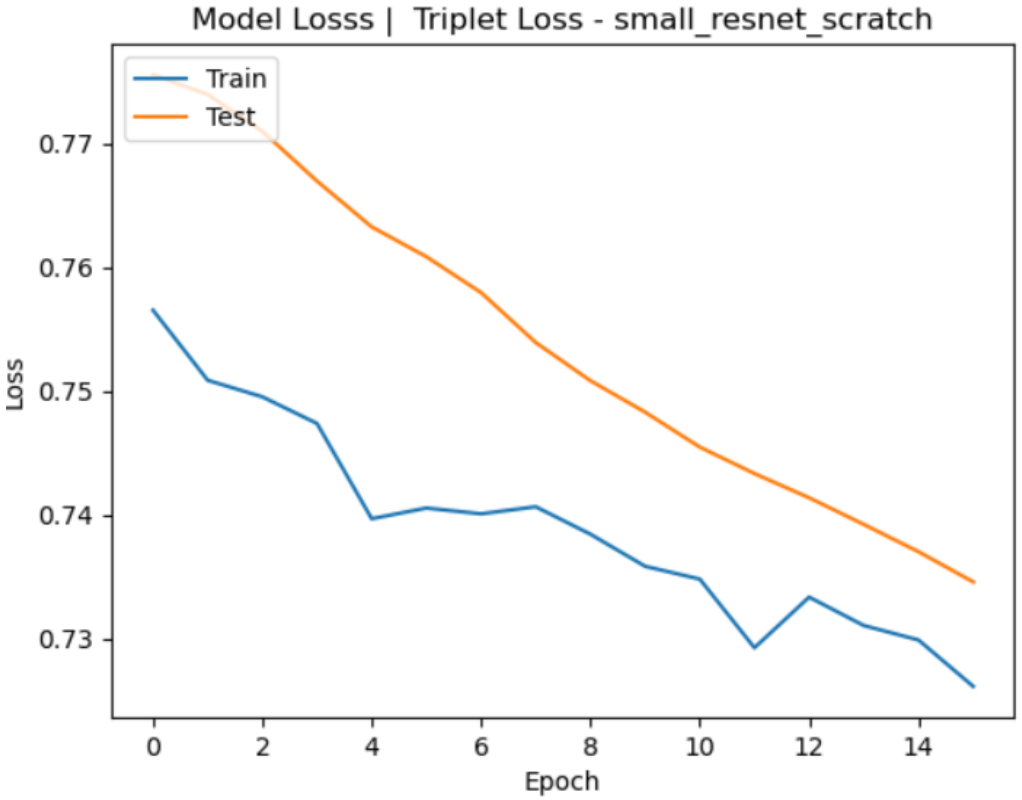

FIGURE 4.3: Loss Function triplet loss

When training a model using triplet loss, it is important to monitor the performance on both the training and test sets. The triplet loss function curve between train and test can give insights into how well the model is generalizing to new data.

Ideally, we would like to see the triplet loss decreasing for both the training and test sets over time, indicating that the model is learning to better discriminate between images of different classes.

4.4 Case 04:

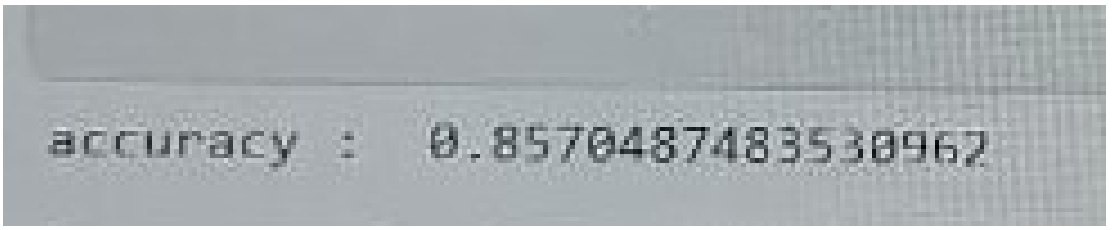

FIGURE 4.4: Test Case 04

From Test Case 4, Accuracy

From the Loss function, model predicts how much correctly about training. Here we are finding accuracy by using clossest positive image/total image

4.5 Case 05:

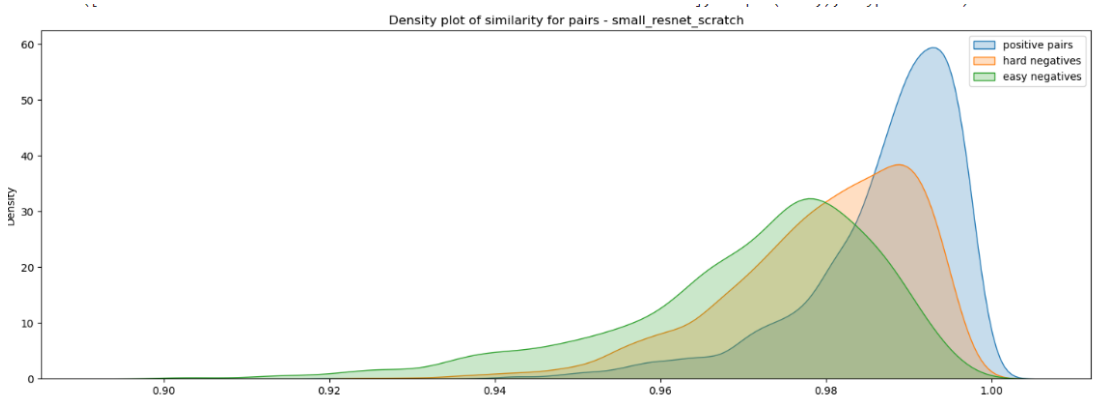

FIGURE 4.5

From Test Case 5, Here Diagramm compare how much skilled forgery(hard-negative) and normal forgery(easy-negative) deviate from original image(positive) curve.

4.6 Case 06:

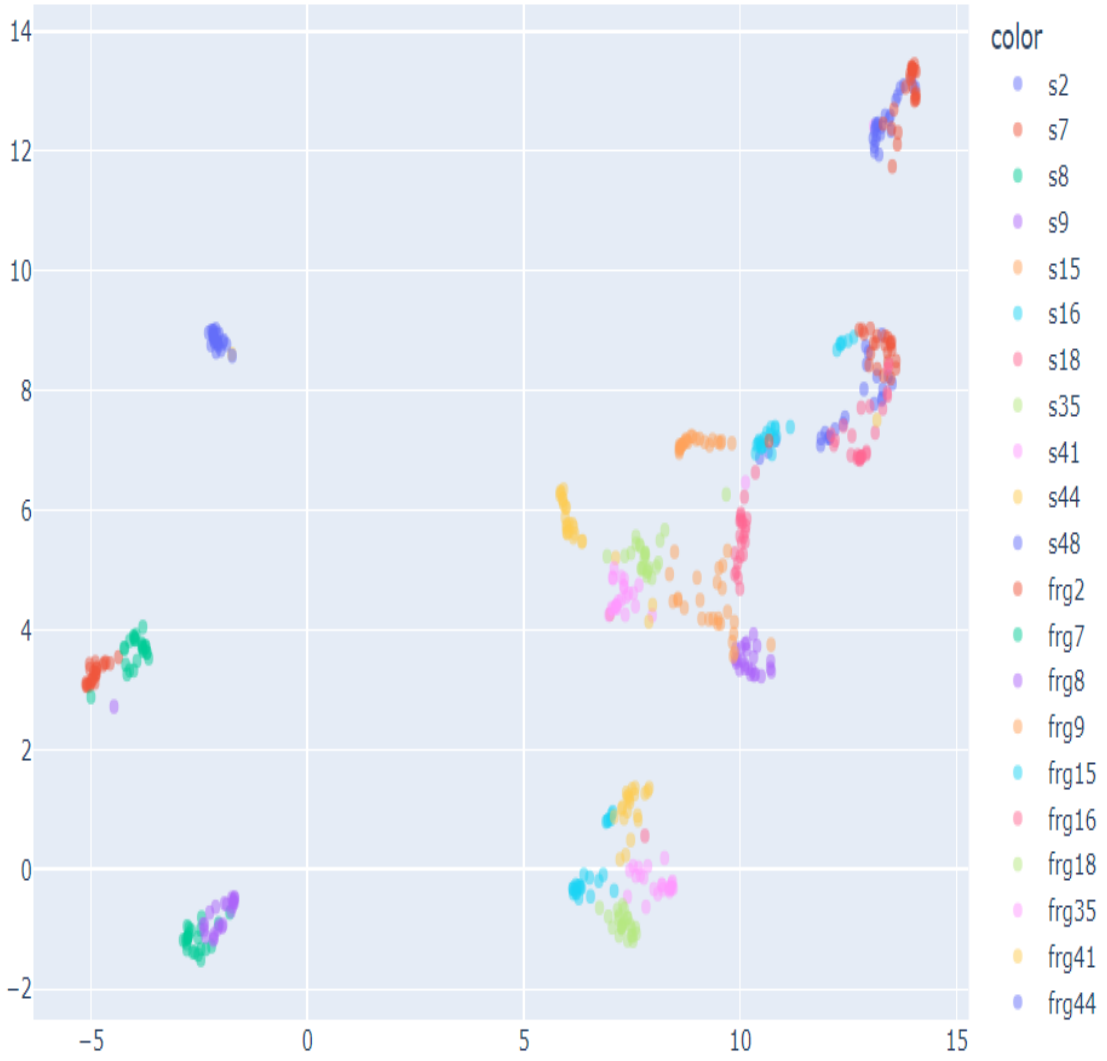

FIGURE 4.6

From Case 06 We plot using seaborn that how originial and forgery image are seperated by distance from each other.

Chapter 5

Summary

5.1 Conclusion

In this paper we have used triplet network for finding human like similarity and dissimalirity. So here we have find three type of images which are positive images, hard-negative images, easy-negative images. Here positive images are images which are closer to genuine person, hard negative are skilled forgery images and easy negative unskilled forgery images. Next we are comparing distance of positive image with these hard-negative and easy-negative with anchor image and those positive we find most closest to anchor we will count those as under genuine.

The threshold is chosen here is otsu to maximize the difference between the inter-class variance and the intra-class variance. In other words, the threshold is chosen such that the foreground pixels are as different as possible from the background pixels. Then genuine with respect to total images gives us how accurate our system predict correct images.

Earlier work is done using VGG-16 CNN classification and using Siamese hybrid network which has best accuracy in writer dependent model. But in case of writer independent model it has accuracy some about 80 percentage. So by taking this as our motivation we have used triplet network and get accuracy of 85.70 percentage in triplet network on writer independent model.

Bibliography

[1] H.-H. Kao and C.-Y. Wen, "An offline signature verification and forgery detection method based on a single known sample and an explainable deep learning approach," *Applied Sciences*, vol. 10, no. 11, p. 3716, 2020.

[2] A. Rateria and S. Agarwal, "Offline signature verification through machine learning," pp. 1–7, 2018.

[3] L. Liu, L. Huang, F. Yin, and Y. Chen, "Offline signature verification using a region based deep metric learning network," *Pattern Recognition*, vol. 118, p. 108009, 2021.

[4] N. C. Codella, C.-C. Lin, A. Halpern, M. Hind, R. Feris, and J. R. Smith, "Collaborative human-ai (chai): Evidence-based interpretable melanoma classification in dermoscopic images," pp. 97–105, 2018.

[5] R. Tolosana, R. Vera-Rodriguez, C. Gonzalez-Garcia, J. Fierrez, A. Morales, J. Ortega-Garcia, J. C. Ruiz-Garcia, S. Romero-Tapiador, S. Rengifo, M. Caruana *et al.*, "Svc-ongoing: Signature verification competition," *Pattern Recognition*, vol. 127, p. 108609, 2022.

[6] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," pp. 815–823, 2015.

[7] W. Chen, X. Chen, J. Zhang, and K. Huang, "Beyond triplet loss: a deep quadruplet network for person re-identification," pp. 403–412, 2017.

[8] G. Wang, Y. Guo, Z. Xu, Y. Wong, and M. Kankanhalli, "Semantic-aware triplet loss for image classification," *IEEE Transactions on Multimedia*, 2022.

[9] H. Kaur and M. Kumar, "Signature identification and verification techniques: state-of-the-art work," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 2, pp. 1027–1045, 2023.

[10] M. M. Hameed, R. Ahmad, M. L. M. Kiah, and G. Murtaza, "Machine learning-based offline signature verification systems: A systematic review," *Signal Processing: Image Communication*, vol. 93, p. 116139, 2021.

[11] S. Lai and L. Jin, "Recurrent adaptation networks for online signature verification," *IEEE Transactions on information forensics and security*, vol. 14, no. 6, pp. 1624–1637, 2018.

[12] M. Ye, J. Shen, G. Lin, T. Xiang, L. Shao, and S. C. Hoi, "Deep learning for person re-identification: A survey and outlook," *IEEE transactions on pattern analysis and machine intelligence*, vol. 44, no. 6, pp. 2872–2893, 2021.

[13] I. Melekhov, J. Kannala, and E. Rahtu, "Siamese network features for image matching," pp. 378–383, 2016.

[14] F. Sultana, A. Sufian, and P. Dutta, "Advancements in image classification using convolutional neural network," pp. 122–129, 2018.