Palm vein Verification using siamese networks

A Project Report

submitted by

JAWAHAR A S , JAWAHAR ABISHEK K

(CS21B1031,CS21B1086)

in partial fulfilment of requirements

for the award of the degree of

BACHELOR OF TECHNOLOGY

Department of Computer Science Engineering

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY, DESIGN AND MANUFACTURING KANCHEEPURAM

MAY 2024

DECLARATION OF ORIGINALITY

I, **Jawahar A S , Jawahar Abishek K**, with Roll No: **CS21B1031,CS21B1086** hereby declare that the material presented in the Project Report titled **Palm vein Verification using siamese networks** represents original work carried out by me in the **Department of Computer Science Engineering** at the Indian Institute of Information Technology, Design and Manufacturing, Kancheepuram.

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.

Jawahar A S , Jawahar Abishek K

Place: Chennai

Date: 13.05.2024

CERTIFICATE

This is to certify that the report titled **Palm vein Verification using siamese networks**, submitted by **Jawahar A S , Jawahar Abishek K (CS21B1031,CS21B1086)**, to the Indian Institute of Information Technology, Design and Manufacturing Kancheepuram, for the award of the degree of **BACHELOR OF TECHNOLOGY** is a bona fide record of the work done by him/her under my supervision. The contents of this report, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Dr. Rahul Raman
Project Guide
Assistant Professor
Department of Computer Science Engineering
IIITDM Kancheepuram, 600 127

Place: Chennai

Date: 13.05.2024

ACKNOWLEDGEMENTS

We extend our sincere gratitude to Dr. Rahul Raman for his invaluable guidance and mentorship during the course "Introduction to Biometrics" at IIITDM Kancheepuram. His expertise and support have been indispensable throughout the duration of this project.

Special thanks to Felix Marattukalam et al. for their pioneering research in palm vein verification, which served as the main inspiration for our study. Their work from Auckland, New Zealand, provided valuable insights and directions for our research endeavors.

We are grateful to the team at CASIA for providing access to the dataset used in this project. Their contribution has been essential in conducting our experiments and validating our findings.

We would also like to acknowledge each other for our collaborative efforts, shared insights, and dedication to the project's success. Working together has been a rewarding experience, and we appreciate each other's contributions.

Finally, we express our appreciation to the faculty and staff of the Computer Science Engineering department at IIITDM Kancheepuram for their support and encouragement throughout our academic journey.

ABSTRACT

The need for contactless biometric systems have significantly increased due to the onset of the Covid-19 global pandemic. Although various extrinsic modalities like face, iris, and palmprint which are tangible part of the body are successfully being used, now there is a need for intrinsic systems like finger vein, hand vein, and palm vein which are subcutaneous and not visible to the naked eye. The features in these systems are the veins which helps liveliness detection, and is slightly more robust to spoof attacks that has been a challenge among researchers. Palm vein systems , which is the focus of the present investigation, is an intrinsic biometric system and preferred due to the ease of interaction with palm vein scanners.Our approach proposes an architecture using Siamese neural network structure for few shot palm vein verification. The proposed network uses images from both the palms and consists of two sub-nets that share weights to identify a person.

KEYWORDS: Palm vein; Siamese; Computer vision; Biometric

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

CASIA	The Institute of Automation, Chinese Academy of Sciences
ROI	Region Of Interest
CNN	Convolutional Neural Network
DBN	Deep Belief Network
AE	Auto Encoder
GAN	Generative Adversarial Network
JPEG	Joint Photographic Experts Group
SVM	Support Vector Machine
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative

NOTATION

λ	wavelength

CHAPTER 1

Introduction

1.1 Introduction

The need for contactless biometric systems have significantly increased due to the onset of the Covid-19 global pandemic. Although various extrinsic modalities like face, iris, and palmprint [1] which are tangible part of the body are successfully being used, now there is a need for intrinsic systems like finger vein, hand vein, and palm vein which are subcutaneous and not visible to the naked eye. The features in these systems are the veins which helps liveliness detection, and is slightly more robust to spoof attacks that has been a challenge among researchers. Palm vein systems, which is the focus of the present investigation, is anintrinsic biometric system and preferred due to the ease of interaction with palm vein scanners.

A palm vein biometric system has to go through several stages: acquisition, pre-processing, feature extraction, decision making. Portions of the research in this paper use the CASIA-MS-PalmprintV1 collected by the Chinese Academy of Sciences' Institute of Automation (CASIA)[2].

After acquiring the images, The incoming images are cropped to a region of interest (ROI), essentially the palm region having maximum vein information to perform recognition. Then the ROI image is processed using image processing methods, and passed on to matching algorithms for decision making. The methods for matching can essentially be classified as traditional and deep learning methods. Since palm vein recognition is a classification problem and deep learning methods have been successful on such tasks, researchers are inclined towards its use.

However, the limitation that deep learning has for palm vein recognition is the need for massive databases with high quality labeled images [3]. This is often scarce in the case of palm vein recognition systems. Therefore, our work showcases the use of deep learning networks using limited samples for vein verification. The architecture used is inspired from the Siamese neural network structure, and specifically addresses the verification setting in the recognition system.

CHAPTER 2

Related Works

2.1 Related Works

Numerous methods have been proposed to extract and match vein patterns from vein images. These patterns are used for biometric recognition using different approaches. In this section a few methods are presented for sake of completeness. The extraction methods can be categorised into subspace learning, local descriptor, vessel geometry, and deep learning. The recent inclination is in using Siamese networks to curb the need for large databases [3]. [4] elaborates how subspace learning uses obtained coefficients as unique features for recognition. Local descriptor approaches are better described in [5]. Finally, deep learning approaches such as convolutional neural network(CNN), deep belief network (DBN), and auto-encoders (AE) [6] are used for feature extraction and subject recognition. As highlighted already,deep learning in biometrics needs for large datasets .In one current public database, the images for each subject or class are limited to twelve [7] and not accounting dynamic class change. This led to the need for alternate approaches. Some researchers proposed to augment the available data. They explored generative adversarial networks (GANs) with data augmentation to improve classification performance [8]. This did achieve reasonable performance but did not solve the problem of system speed, data privacy and storage space associated with duplication of input data. Also, data augmentation can easily lead to overfitting. Researchers are inclined towards more effective methods like similarity learning and few-shot learning. [9] discusses these latest approaches and proposes a few shot learning approach for palmprint recognition, which displays good accuracy. We propose in this paper a combined Siamese structure for palm vein verification.

CHAPTER 3

Methodology

This section discusses the database used, the general Siamese neural network architecture, and the network structure implemented with the loss functions that have been used to evaluate the network performance.

3.1 Database

The research in this paper use the CASIA MS PalmprintV1 collected by the Chinese Academy of Sciences Institute of Automation (CASIA)[2].The CASIA dataset has been used for our research work. CASIA Multi-Spectral Palmprint Image Database contains 7,200 palm images captured from 100 different people using a self-designed multiple spectral imaging device.All palm images are 8 bit gray-level JPEG files. For each hand, we capture two sessions of palm images. The time interval between the two sessions is more than one month. In each session, there are three samples. Each sample contain six palm images which are captured at the same time with six different electromagnetic spectrums. Wavelengths of the illuminator corresponding to the six spectrum are 460nm,630nm,700nm,850nm,940nm and white light respectively. Between two samples, we allow a certain degree of variations of hand postures. Through that, we aim to increase diversity of intraclass samples and simulate practical use. From the dataset which contains various wavelengths, we are selecting the images which are captured at 850nm wavelength(λ) as it tends to produce better results for palm vein study as per Wu Wei et al[10]

3.2 Siamese Neural Networks

Even though deep learning algorithms have proven their ability to produce exceptional results, the performance of the designed algorithm is often dependant on the number of data samples available to train the network [9]. The performance of the network improves with the increasing number of data samples. In biometric systems, especially palm vein systems, suitably labelled datasets are not readily available. One-shot or few-shot learning is the appropriate approach when only a few training examples are available for the network to train. The few-shot learning approach uses Siamese neural network. As shown in Fig.3.1, a typical Siamese neural network has two paths and aims to find the similarity between its inputs. It has identical parallel networks which share the same architecture and weights. Siamese neural networks was first proposed by Bromley et al [11] for signature verification in biometrics and is widely used in face verification tasks [2]. Profound details about Siamese networks for image recognition are available in [12] [13] [14].

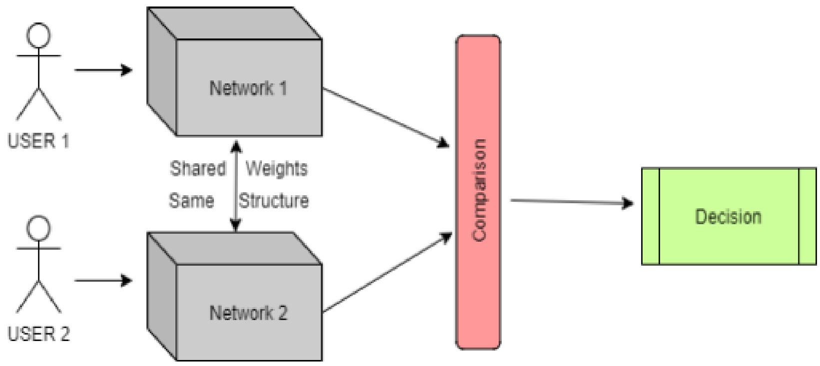

Figure 3.1: A typical siamese network

3.3 Network Structure

This paper proposes to develop two identical networks that process two images simultaneously and compute the similarity or difference between the two images. If the images are from two different candidates, the network essentially needs to compute the similarity function and increase the distance between them. Fig.3.2 shows the overview of the proposed network structure which is based on Siamese architecture. Here, there are two sub-networks, subnet 1 and subnet 2. Each of them has sub subnets within them to process the input image. The left and right palm images pass through a spatial feature extractor with the convolutional neural network network structure shown in Tab. 3.1.

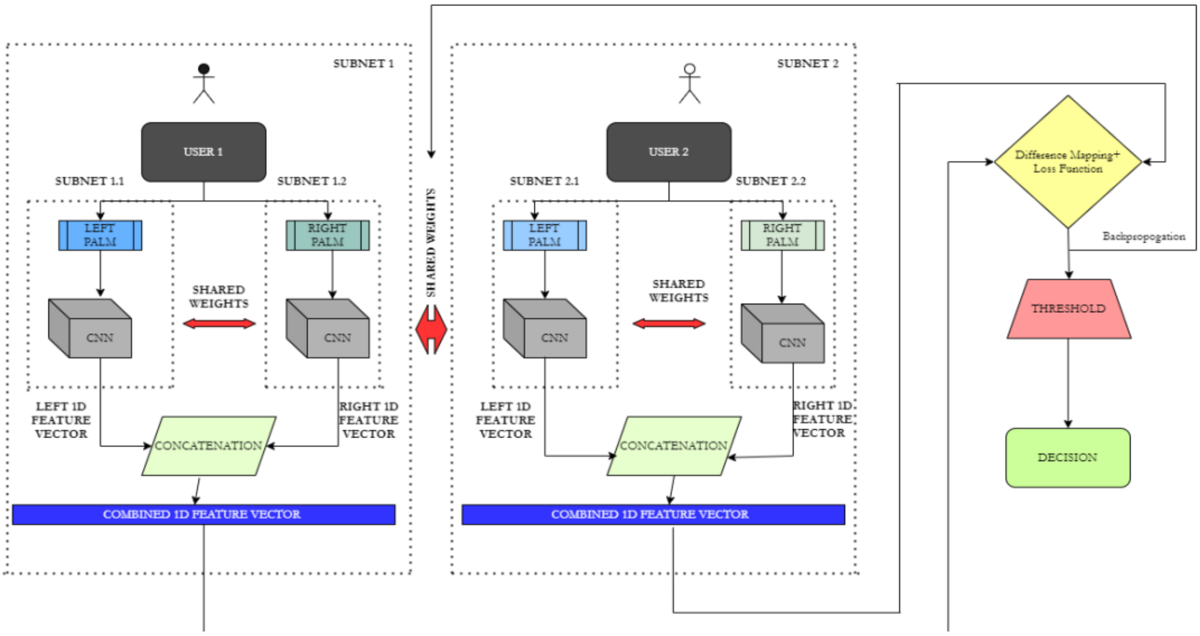

Figure 3.2: Overview of the proposed network structure. This network is based on the Siamese architecture. The red arrows indicate the weights being shared between sub-networks. The outputs of the sub-networks are 1-D feature vectors.

Consider two users, where each user submits the left and right palm image to the network. Hence, the network receives four images in total, namely, x1, x2 and y1, y2. The spatial feature extractor network shown in Tab. I generates the feature embeddings f(x1) and f(x2) respectively, which are one dimensional vectors of length 128×1. These vectors are then concatenated together to form F(X). Similar process is followed to obtain F(Y). Then the feature embeddings are subjected to a function E which computes the L1 distance.The function is given by eqn (3.1):

$$E(X,Y) = d(X,Y) = |||F(X) - F(Y)||| \tag{3.1}$$

The function E will be smaller if the concatenated feature vector F(X) is similar to F(Y). This distance value is used to finetune the network weights using back propagation. A sigmoid activation function is used to convert the distance to probability P.

Table 3.1: CNN feature extractor structure

Layer	Operation	Output Size	Kernel Size	Stride	Padding
1	Convolution 1 (RELU) Batch Norm + Max Pooling	$64 \times 3 \times 3$	3x3	1	0
2	Convolution 2 (RELU) Batch Norm + Max Pooling	$64 \times 3 \times 3$	3x3	1	0
3	Convolution 3 (RELU) Batch Norm + Max Pooling	$64 \times 3 \times 3$	3x3	1	1
4	Convolution 4 (RELU) Batch Norm + Max Pooling	$64 \times 3 \times 3$	3x3	1	1
5	Fully Connected	1000 hidden units	-	-	-
6	Fully Connected Sigmoid	128 hidden units	-	-	-
7	Extracted Features	-	-	-	-

3.4 Preprocessing

Original images from CASIA database cannot be directly used for training the model so ROI(Region Of Interest) extraction happens from the given palm image. The ROI extraction refers to carrying out a series of adjustment and key points location for different palm vein images, then the effective area of center is selected to extract features, and final matching is carried out for the recognition. This central region is usually called the region of interest (ROI),palm vein image of the same palm, the location of ROI should be the same. The purpose of ROI location and selection is to do the feature area normalization of the different palm vein, so the influence of adverse factors will be eliminated, and the sub image including rich information of palm vein is extracted, which is convenient for the subsequent feature extraction and matching.

3.4.1 ROI Extraction Method based on the Maximum Inscribed Circle

Figure 3.3: CASIA Palm vein image

Image clipping:

In all the palm vein images of CASIA palm vein database, the wrist portion's background is cluttered, and hence, it should be removed before ROI extraction with fxed parameter strategy [7] thus arriving at wrist eliminated palm vein image.

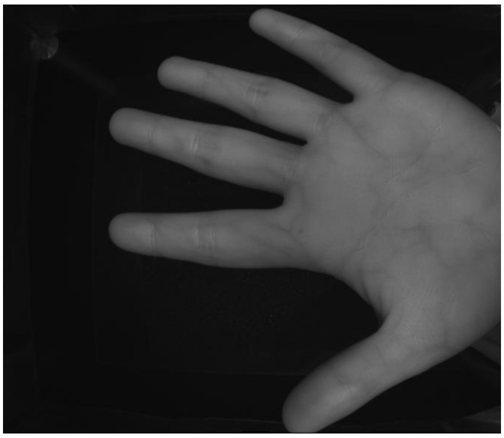

Figure 3.4: wrist eliminated palm vein image

Otsu thresholding:

To overcome the infuence of background, Otsu thresholding [15] is applied to the wrist eliminated palm vein image to get segmented parts of the vein image.

Figure 3.5: Segmented palm vein image

Largest connected component:

Upon obtaining the resultant binary image, a critical step in delineating the palm region from the background clutter involved identifying the largest connected region. This largest connected region encapsulated the boundary of the palm region along with its adjoining fingers.To accomplish this, contours were extracted from the binary image, revealing distinct shapes and structures. Among these contours, particular attention was directed towards isolating the largest contour, as it represented the dominant palm region within the image.Once the largest contour was identified, a strategy was employed to differentiate the palm region from the surrounding background. This was achieved through the creation of a mask specifically made to encompass the largest contour. Subsequently, this mask was applied to the original image.By employing this methodology, the resultant image retained only the region of interest, namely the palm area, while effectively eliminating extraneous background elements.

Figure 3.6: Effective palm binary image

Largest Inscribed circle:

The process began with the computation of the distance transform of the binary image, which served as a representation of the palm region. This transformative operation calculated the distance from every pixel residing within the palm region to the nearest boundary pixel. Such computation elucidated the spatial relationships within the palm region, facilitating subsequent analyses.Upon completion of the distance transform, a comprehensive analysis was conducted to discern pivotal characteristics. Notably, attention was directed towards identifying the maximum value within the transformed image. This maximum value, pinpointed through analytical scrutiny, served as a definitive marker denoting the center of the maximum inscribed circle within the palm region. It is important to note that the identified maximum value was associated with the radius of the maximum inscribed circle.Such delineation represents a fundamental step towards robust feature extraction and subsequent palm vein recognition methodologies.

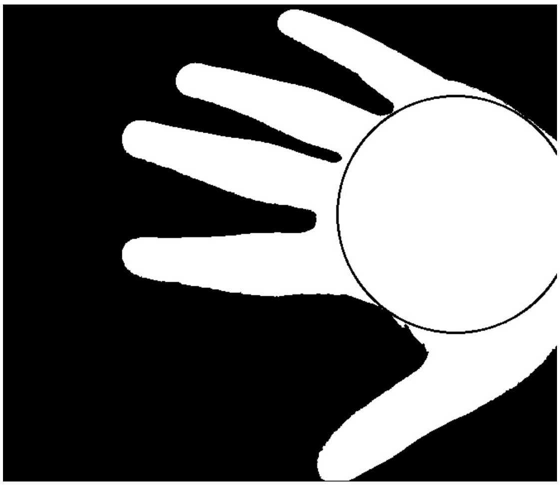

Figure 3.7: Largest inscribed circle

Largest Inscribed square:

Following the identification of the maximum inscribed circle within the palm region, further analysis was conducted to identify the dimensions of an inscribed square that encapsulated this circle.The first step involved the calculation of the side length of the square. This computation was performed by using the radius of the maximum inscribed circle, adhering to the geometric relationship between the diagonal of a square and its side length. Specifically, the side length was computed as the square root of two times the radius of the maximum inscribed circle.With the side length, precise calculations were made to ascertain the coordinates of the top-left corner of the inscribed square. By correctly positioning this corner, the square could be seamlessly aligned with the palm region's geometric center. Subsequently, the inscribed square was resized to ensure uniformity in dimensions, facilitating standardized feature extraction methodologies.

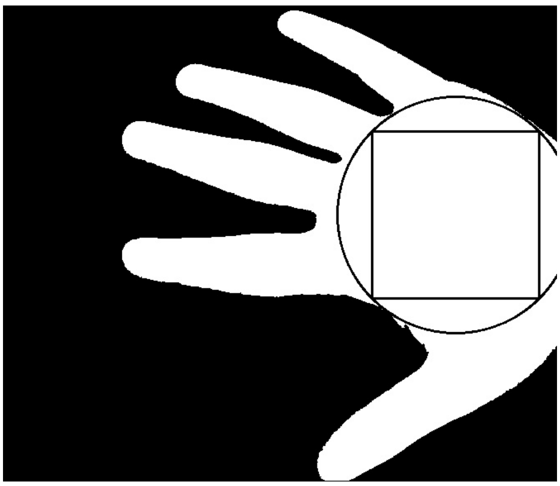

Figure 3.8: Largest inscribed square

Extraction of ROI:

Using the maximum inscribed square, we extract the ROI(Region Of Interest) from our binary image.

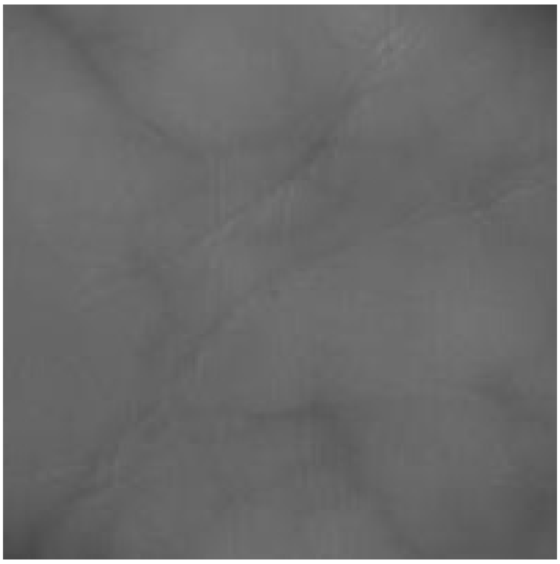

Figure 3.9: ROI image

Final Gabor ROI image:

To enhance the extracted square portion of the palm image and extract relevant features, a Gabor filter was applied.The initial step involved the creation of a Gabor filter kernel, tailored to the specific requirements of the image analysis task. Subsequently, the Gabor filter kernel was applied to the square portion of the palm image. This filtering operation facilitated the extraction of texture-based features inherent in the palm region, thereby enhancing the discriminative capabilities of subsequent analyses.To maintain consistency and facilitate further processing, the filtered image underwent normalization, ensuring its pixel intensity values ranged within the standardized interval of [0, 255]. This normalization procedure ensured uniformity in feature representation across varying palm images and alleviated potential scaling discrepancies.

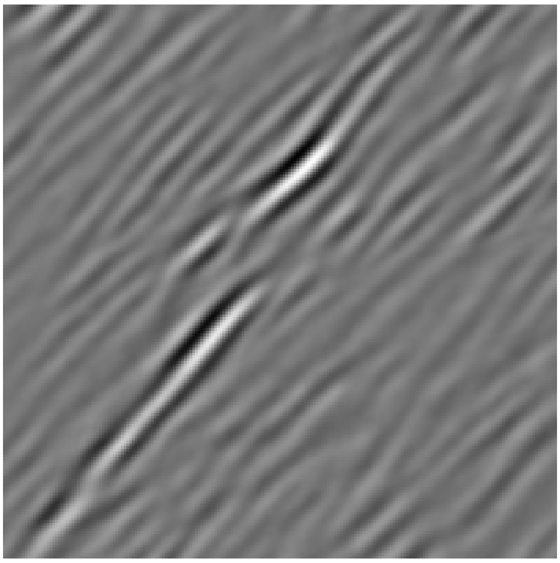

Figure 3.10: Final Gabor ROI image

3.5 Loss Function

Siamese networks classify the inputs into binary classes ie. ”1” being same inputs and ”0” being different. Contrastive loss and binary cross-entropy function loss are the two common loss options in binary classification

3.5.1 Contrastive loss:

It requires pairs of input samples. The encoder is penalized by the loss function based on the class of the input image. If the input images are from the same class, the model produces similar feature embeddings. Mathematically it is given by eqn (3.2):

$$\text{Loss} = (1-y)\cdot\frac{1}{2}(d)^2 + y\cdot\frac{1}{2}\left[\max(0, m-d)\right]^2 \tag{3.2}$$

Here y is the actual label and will be zero when the embeddings of combined input images (left palm and right palm) and one if they are not same, d is the distance measure between the feature embeddings and the input images, m is the hyper parameter margin which is maximized if the input images are similar. If the input pairs are dissimilar and the distance d is greater than the margin m, no loss is incurred.

3.5.2 Binary cross-entropy loss:

It is also known as log loss and is used to calculate the classifier performance which is in the range between 0 and 1. If the predicted probability varies from actual class, the loss increases. Mathematically it is given by eqn (3.3): The loss function is defined as:

$$\text{Loss} = -y \log(p) + (1 - y) \log(1 - p)^2 \tag{3.3}$$

Here y is the class label and p is the prediction probability. It is used to differentiate between similar and different images by providing the aggregate of positive and negative loss probability.

CHAPTER 4

EXPERIMENTAL RESULTS AND ANALYSIS

This section discusses about the results based on the experimental setting and the proposed architecture. The performance analysis of the results is done using the matrices namely, accuracy, precision, recall, specificity and F1-score.

4.1 Implementation

This is a k-way n shot classification problem. The dataset D with a data split of 60:20:20 was used. The training set contains n samples from k-classes adding upto k × n samples in the training dataset and a query set in the testing dataset. Predominantly there were only two classes i.e. same and different subjects, and hence, k=2 and n varying from one to five depending on the sample set considered from the existing database. The model was trained using batch of training tasks to ultimately categorize the image during the testing task. At the end of each epoch, the model parameters were updated through back-propagation as per the loss calculated.The CASIA database consists of uniform images. We used the region of interest (ROI) images of resolution by using the method in [16].This was subjected to the spatial feature extractor described section 3.3. The experiment was carried out using the Keras framework with Tensorflow on a GPU T4 X2 15GB with i5 2.5 GHz processor supported with 16 GBs of RAM on the kaggle virtual environment. The learning rate was set at 0.0001 and Adam Optimizer was used. The one dimensional feature vector extractor which uses the structure shown in Tab. 3.1 has fully connected layers having 128 hidden units followed by the sigmoid function.

The model was evaluated using the metrics: accuracy, precision, recall, specificity and F1- score. These parameters were preferred based on our previous study for comparison between SVM with CNN. These parameters are briefly summarised in Tab. 4.1. Here, the number of predictions where the classifier correctly predicts the positive class

Table 4.1: Performance metrics using confusion matrix

Measure	Formula
Accuracy (A)	$\frac{TP+TN}{TP+TN+FP+FN}$

as positive is True Positive (TP), the negative class as negative is True Negative (TN), the negative class as positive is False Positive (FP), and the positive class as negative is False Negative (FN).

4.2 Results and Discussion

The goal of this study specifically is to exploit the benefits of Siamese neural network in palm vein verification by discussing its performance parameters and establish its use in an end-to-end palm vein recognition system. Siamese neural networks uses the principle of similarity between image pairs.

While our achieved testing accuracy of 83% may seem modest, it underscores the potential of Siamese neural networks in palm vein verification. Despite the challenges posed by variations in lighting conditions, limited amount of dataset, and anatomical differences among individuals, our results demonstrate the feasibility of using Siamese networks for accurate palm vein recognition.

CHAPTER 5

CONCLUSION AND FUTURE SCOPE

5.1 Conclusion

This paper discusses the dynamics and benefits of palm vein verification and proposes a state-of-the-art deep learning Siamese neural network that can be used in contactless biometric systems. This is achieved by integrating a k-way n-shot learning network model with contrastive loss and an optimized CNN feature encoder. The results obtained are promising owing to the fact that this model is trained with only a limited amount of samples from each palm for a given training class/subset.

5.2 Future Scope

For the future works,we can alter the model in such a way that the images in the existing database will be stored as a feature encoding instead of the actual images in order to save storage and faster computation. We can also work on improving the model's accuracy.

5.3 Individual contribution

A S Jawahar(CS21B1031)

- Implemented image preprocessing techniques to prepare the data for model input. Developed a systematic approach to divide the preprocessed images into pairs of four, ensuring compatibility with the model.
- Designed and implemented the architecture of the base Siamese neural network, incorporating suitable layers and activation functions.
- Defined the contrastive loss function to measure the similarity or dissimilarity between pairs of input images.

- Divided the dataset into train and test portions.Ensured an appropriate distribution of data across training and testing sets to prevent overfitting and underfitting.
- Trained the model using the training dataset to optimize the parameters of the Siamese network. Employed validation datasets to fine-tune model hyperparameters and prevent overfitting. Utilized the trained model to make predictions on the test dataset, evaluating its performance and generalization capabilities.

Jawahar Abishek K(CS21B1086)

- Preprocessed of the CASIA database images which involved the following steps
  - clipping,
  - Otsu's thresholding,
  - retaining the largest connected component,
  - finding the maximum inscribed circle,
  - finding the maximum inscribed square,
  - extracting the square ROI from original grayscale image,
  - gabor applied final ROI image.
- Partitioned of training dataset into separate subsets, designated for training and validation purposes, ensuring a comprehensive assessment of model performance and generalization capabilities on unseen dataset.
- Fine-tuned the model hyperparameters, including adjustments to batch size, epoch count, and learning rate, optimizing the model's performance and adaptability to the specific characteristics of the dataset.
- Defined the accuracy function to calculate accuracy of the model's performance.

REFERENCES

[1] W. S. L.Fei, G.Lu and D.Zhang, "Feature extraction methods for palmprint recognition: A survey and evaluation," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 2, pp. 346–363, 2018.

[2] CASIA, "Casia-ms-palmprintv1 database," 2009-. [Online]. Available: http://www.cbsr.ia.ac.cn/MS_PalmprintDatabase.asp

[3] A. N. D. Thapar, G. Jaswal and V. Kanhangad, "Pvsnet: Palm vein authentication siamese network trained using triplet loss and adaptive hard mining by learning enforced domain specific features," *IEEE 5th International Conference on Identity, Security, and Behavior Analysis (ISBA)*, pp. 1–8, 2019.

[4] S. Elnasir and S. M. Shamsuddin, "Proposed scheme for palm vein recognition based on linear discrimination analysis and nearest neighbour classifier,," *2014 International Symposium on Biometrics and Security Technologies (ISBAST). IEEE*, p. 67–72, 2014.

[5] F. Marattukalam and W. H. Abdulla, "Segmentation of palm vein images using u-net," *2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC). IEEE*, pp. 64–70, 2020.

[6] B. Hou and R. Yan, ""convolutional autoencoder model for finger-vein verification,," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 5, pp. 2067–2074, 2019.

[7] K. W. D. F. W. Q. Yan, X., "Palm vein recognition based on multi-sampling and feature-level fusion." *Neurocomputing*, no. 157, pp. 798–807, 2015.

[8] Y. L. H. Qin, M. A. El-Yacoubi and C. Liu, "Multi-scale and multidirection gan for cnn-based single palm-vein identification," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 2652–266, 2021.

[9] X. D. S. D. H. Shao, D. Zhong and R. N. Veldhuis, "Few-shot learning for palmprint recognition via meta-siamese network," *IEEE Transactions on Instrumentation and Measurement*, 2021.

[10] S. M. A. Uhl, C. Busch and R. Veldhuis, *Handbook of vascular biometrics*, 2020.

[11] Y. L. E. S. J. Bromley, I. Guyon and R. Shah, "Signature ¨ verification using a" siamese" time delay neural network," *Advances in neural information processing systems*, vol. 6, no. 5, pp. 737–744, 1993.

[12] R. Z. G. Koch and R. Salakhutdinov, "Siamese neural networks for one-shot image recognition," *ICML deep learning workshop*, vol. 2, 2015.

[13] D. Chicco, "Siamese neural networks: An overview," *Artificial Neural Networks*, pp. 73–94, 2021.

[14] M. Y. X. S. H. W. Y. Zhan, K. Fu and X. Qiu, "Change detection based on deep siamese convolutional network for optical aerial images," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 10, pp. 1845–1849, 2017.

[15] N. Otsu, "A threshold selection method from gray-level histogram," *IEEE Trans. Syst Man Cybern.*, no. 9, pp. 62–66, 1979.

[16] T. X. S. Lin and X. Yi, "Region of interest extraction for palmprint and palm vein recognition," *2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISPBMEI). IEEE*, pp. 538–542, 2016.