Learning Domain Specific Features using Convolutional Autoencoder: A Vein Authentication Case Study using Siamese Triplet Loss Network

Manish Agnihotri[1], Aditya Rathod[1], Daksh Thapar[2], Gaurav Jaswal[2], Kamlesh Tiwari[3] and Aditya Nigam[2]

[1]*Department of Information & Communication Technology, Manipal Institute of Technology Manipal, India*

[2]*School of Computing and Electrical Engineering, Indian Institute of Technology Mandi, India*

[3]*Department of Computer Science and Information Systems, Birla Institute of Technology and Science Pilani, India*

Keywords: Siamese Network, Vascular Biometrics.

Abstract: Recently, deep hierarchically learned models (such as CNN) have achieved superior performance in various computer vision tasks but limited attention has been paid to biometrics till now. This is major because of the number of samples available in biometrics are limited and are not enough to train CNN efficiently. However, deep learning often requires a lot of training data because of the huge number of parameters to be tuned by the learning algorithm. How about designing an end-to-end deep learning network to match the biometric features when the number of training samples is limited? To address this problem, we propose a new way to design an end-to-end deep neural network that works in two major steps: first an auto-encoder has been trained for learning domain specific features followed by a Siamese network trained via. triplet loss function for matching. A publicly available vein image data set has been utilized as a case study to justify our proposal. We observed that transformations learned from such a network provide domain specific and most discriminative vascular features. Subsequently, the corresponding traits are matched using multimodal pipelined end-to-end network in which the convolutional layers are pre-trained in an unsupervised fashion as an autoencoder. Thorough experimental studies suggest that the proposed framework consistently outperforms several state-of-the-art vein recognition approaches.

1 INTRODUCTION

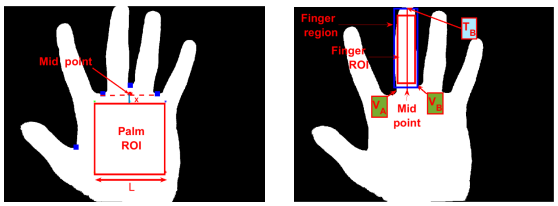
Authentication Systems based on biometrics have been increasingly used in a variety of security applications such as banking, immigration control, forensics and so on. Several biometric traits like iris, fingerprint, ear, face etc. have been studied for reliable security solutions. However, each trait has its own specific challenges that limit its usages in real time scenarios. Moreover, using a unimodal system makes it vulnerable to presentation attacks (Patil et al., 2016). As the multimodal systems blend the information from multiple biometric traits, it not only increases the performance of the system by complementing the lapses of each individual traits, it also makes the system secure from attacks like presentation attacks. Moreover, they also provide better population coverage, which are essential for performing recognition on large databases. Apart from such benefits, multiple sensors are required to capture the multi-biometric samples, but that increase the overall cost and require a higher degree of user cooperation (Kimura et al., 2015). However, the above-mentioned shortcomings can be improved if the biometric characteristics lie close to each other. For instance, the acquisition of frontal hand region offers this opportunity to capture the biometric characteristics together from finger and palm regions. Moreover, uniqueness of the blood vessel networks among the individuals and challenges to its reproduction make vein patterns a strong biometric identifier.

1.1 Related Work

Multimodal biometric systems involve the integration of evidence of information from multiple biometric traits to achieve a performance superior to any of the individual modalities. However, the choice of biometric traits plays a key role in the usability of the multimodal biometric systems. More specifically, re-

Agnihotri, M., Rathod, A., Thapar, D., Jaswal, G., Tiwari, K. and Nigam, A.
Learning Domain Specific Features using Convolutional Autoencoder: A Vein Authentication Case Study using Siamese Triplet Loss Network.
DOI: 10.5220/0007568007780785
In *Proceedings of the 8th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2019)*, pages 778-785
ISBN: 978-989-758-351-3

search in vascular biometrics has become very popular. Besides being unique, the subcutaneous vein structures have the added advantage of lying underneath the skin surface. This makes visibility to the eyes or general purpose cameras difficult and hence, this limits the ease of spoofing if not averting it completely (Patil et al., 2016). For instance, the authors in (Yang and Zhang, 2012) uses both fingerprint and finger-vein modalities as they are both extracted from the finger region, making the multimodal biometric collection process convenient. Similar researches have been carried on for palmprint and palm-vein (Wang et al., 2008), finger-vein and finger dorsal texture (Yang et al., 2014), hand dorsal and palm vein patterns (Ramalho et al., 2011), finger knuckle and finger vein (Veluchamy and Karlmarx, 2016) and so on. Although these approaches require lesser effort on the user's part, nevertheless leads to the employment of more complex hardware. The other studies in vascular biometrics explore the line based and curvature based information in the vessel structures present in the biometric samples (Zhou and Kumar, 2010). Also, it has been noted that much attention has been paid to the palm-vein and finger-vein modalities individually, but very few researchers attempted to address the problem of presentation attack for vascular technology (Choi et al., 2009). However, much recent work has been focused on employing the deep learning techniques in various domains and the field of biometrics is not an exception. These works have shown that the deep learning based features perform better than the handcrafted features for face (Taigman et al., 2014) and finger-vein (Qin and El-Yacoubi, 2017) verification systems. In (Xie and Kumar, 2017), authors demonstrated a new deep learning approach for the finger vein recognition using the CNN and supervised discrete hashing. In a very recent study (Fang et al., 2018), authors proposed a two-channel CNN network that has only three convolution layers for finger vein verification.



(a) (b)

Figure 1: Vein ROI Extraction (Bhilare et al., 2018): (a) Palm-vein ROI, (b) Finger Vein ROI.

1.2 Challenges and Contribution

Challenges: The vein patterns either collected from dorsal or palmer side of the hand, provide very vast textural information. The acquisition procedures are convenient and hygienic than other methods. But, substantial changes in hand positioning during contactless acquisition, make this task very difficult to achieve a satisfactory ROI segmentation since that affects the overall system performance. The other open issues in vein verification are the lack of robustness against image quality degradation, illumination, and point-wise matching. A few existing image transformation techniques (such as LBP (Ahonen et al., 2006), BOP(Jaswal et al., 2017), TCM (Umer et al., 2016), GOP(Nigam et al., 2016), IRT(Cummings et al., 2011) etc.) are well proposed in literature that create useful representation of image data and helps to improve the matching task (Jaswal et al., 2017). But no work has been proposed yet that encodes the image feature through a Deep Learning model. Earlier works in computer vision and big data have focused their attention on object detection, feature extraction and matching using deep learning models. Therefore, efforts have to be made to bridge the gap between deep learning and biometric recognition. To best of our knowledge, this is the first attempt in which, a convolutional autoencoder has been trained to learn the Texture Code Matrix (TCM) and Image Ray Transform (IRT) based encoding schemes to obtain the deep domain specific features for palm vein and finger vein modalities. By doing such network learned transformations, we use network learned features for matching and achieve speedy computations that surpasses the conventional TCM and IRT based matching.

Contribution: The major contribution of this work includes four folds. (i) A Deep Learning based vein recognition framework has been designed which consist of CAE and Siamese Network. (ii) ROI images are given as input to CAE capable of tuning image features into compact network codes. In practice, the proposed generalized hand vein transformation model has been trained by explicitly reformulate the layers as learning functions. We have trained a convolutional deep auto-encoder with merged connections for learning the TCM transformation and then trained another similar auto-encoder for learning second transformation IRT. In this way, we combine both models and train an end-to-end CAE model from the original image to final IRT transformed image. (iii) In the last part, a Siamese network with triplet loss has been trained and tested over the previously obtained network transformed images. Thus, our deep learning vein recognition framework is highly generalized

for operating on either of palm vein or finger vein databases. (iv) Deep networks usually require a lot of training data, otherwise, they tend to over-fit. To avoid that we pre-train the CNN in the Siamese network using an autoencoder. (vi) Finally, the feature level fusion of palm vein and finger vein modalities has been performed to compare the performance.

2 PROPOSED VASCULAR AUCTHENTICATION SYSTEM

At first, we extract the palm-vein and the four finger-vein (index, middle, ring and little) region of interest (ROI) seperatly from the given hand-vein images. The ROI images are enhanced using contrast limited adaptive histogram equalization (CLAHE). Following this, an Autoencoder has been trained to learn Texture Code Matrix (TCM) and Image Ray Transform (IRT) based encoding schemes, and finally, we use a Siamese network trained using triplet loss for efficient and accurate vein authentication.

2.1 Palm Vein and Finger Vein ROI Extraction

We use the state-of-art algorithm (Bhilare et al., 2018) for extracting the palm-vein ROI. Using this algorithm we have been able to extract centre region of hand as shown in figure 1 consistently.

We have modified (Bhilare et al., 2018) algorithm in order to extract finger-vein ROI of index, middle, ring and little fingers. Algorithm 1 and Algorithm 2 summarize the finger vein ROI extraction method.

2.2 Domain Specific Transformation Learning using Autoencoders

In this work, the auto-encoder is inspired by U-Net model which was used for segmentation task. We have modified this model for learning image transformations, providing output similar to what it has been trained on.

Network Architecture: The architecture of our deep framework consists of two consecutive 3×3 convolution layers with ReLU activation, each followed by batch normalization. Thereafter a 2x2 max pooling operation with stride 2 for down-sampling has been used. At each down-sampling step, we double the number of feature channels. Each step in up-sampling path consists of 2×2 up-sampling operation followed by a concatenation with the correspondingly

Algorithm 1: Finger-vein ROI extraction-Part 1.

Input: Acquired hand-vein image I, hand contour C, finger tips T_I, T_M, T_R and T_L identified in algorithm 1(palm-vein ROI extraction).

Output: Finger-vein ROIs corresponding to index $fROI_I$, middle $fROI_M$, ring $fROI_R$ and little $fROI_L$ fingers

for $i \in I, M, R, L$ **do**
  $T \leftarrow$ current finger f_i
  $T_{Left} \leftarrow$ adjacent finger on left side of T
  $T_{Right} \leftarrow$ adjacent finger on right side of T
  $[B_L, B_R]$ = FINDFINGERBASEPOINTS$(T, T_{Left}, T_{Right}, C)$ ▷ Identify base points for each finger f_i
  Join B_L and B_R
  Rotate I such that $B_L B_R$ is horizontal
  Crop rectangular region with top T_i and base $B_L B_R$
  Perform morphological erosion to obtain $fROI_i$
end for
function FINDFINGERBASEPOINTS$(T, T_{Left}, T_{Right}, C)$
  if $T_{Left} \neq \phi$ and $T_{Right} \neq \phi$ **then** ▷ Middle and ring finger
    $P_{Ref} = T_{Left}$ ▷ reference point
    $P_{End} = T$ ▷ end point
    B_L = FINDPEAK(P_{Ref}, P_{End}, C) ▷ find the point along the hand contour C between P_{Ref} and P_{End} with maximum distance from P_{Ref}
    $P_{Ref} = T_{Right}$
    $P_{End} = T$
    B_R = FINDPEAK(P_{Ref}, P_{End}, C)
  else if $T_{Left} = \phi$ **then** ▷ Little and ring finger of left and right hand respectively
    $P_{Ref} = T_{Right}$
    $P_{End} = T$
    B_R = FINDPEAK(P_{Ref}, P_{End}, C)
    B_L = FINDEQUIDISTANTBASEPOINT(T, B_R, C)
  else ▷ Little and ring finger of right and left hand respectively
    $P_{Ref} = T_{Left}$
    $P_{End} = T$
    B_L = FINDPEAK(P_{Ref}, P_{End}, C)
    B_R = FINDEQUIDISTANTBASEPOINT(T, B_L, C)
  end if
  return B_L, B_R
end function
function FINDPEAK(P_{Ref}, P_{End}, C)

$$p_{max} = \arg\max_{p} \text{DISTANCE}(P_{Ref}, C_p) \quad (1)$$

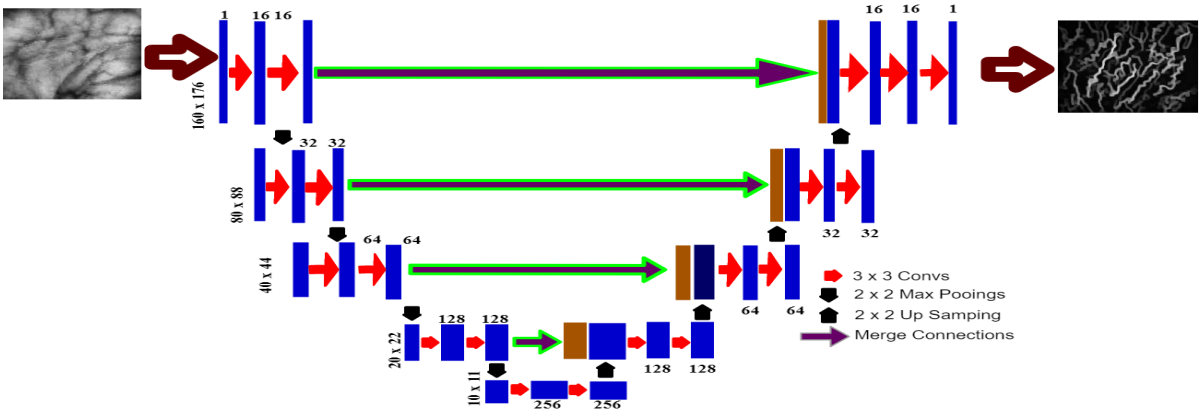


Figure 2: Autoencoder (CAE) Architecture.

Algorithm 2: Finger-vein ROI extraction-Part 2.

Input: Acquired hand-vein image I, hand contour C, finger tips T_I, T_M, T_R and T_L identified in algorithm 1(palm-vein ROI extraction).

Output: finger-vein ROIs corresponding to index $fROI_I$, middle $fROI_M$, ring $fROI_R$ and little $fROI_L$ fingers

34: **return** $C(p_{max})$
35: **end function**
36: **function** FINDEQUIDISTANTBASEPOINT(T, B_1, C)
37: x_1 = INDEXOF($C(B_1)$) ▷ index of finger base B_1 along the contour
38: x_2 = INDEXOF($C(T)$) ▷ index of finger tip T along the contour
39: **if** $x_1 > x_2$ **then**
40: $B_2 = C(x_2 - (x_1 - x_2))$ ▷ B_1 and B_2 are equidistant from T along the contour
41: **else**
42: $B_2 = C(x_2 + (x_2 - x_1)$
43: **end if**
44: **return** B_2
45: **end function**

feature map from the contracting path and reduction in the number of output channels by the factor of 2, and two 3×3 convolutions, each followed by batch normalization and a ReLU. At the final layer, a 1×1 convolution is used to match the output image channels. In total the network consists of 15 convolution's layers and also avoids gradient problem. To allow concatenation of layers it is important to select the input tile size such that all 2×2 max-pooling operations are applied to a layer with an even x- and y-size. Figure 2 shows the network architecture of image transformation model.

Network Training: For training the autoencoders, we first created ground truth by performing transformation operation namely TCM on 600 palm vein and 2400 finger vein samples and thereafter we applied another transformation namely IRT for getting ray tracing image features. We have used mean squared error as loss function and RMSPROP optimizer with default parameters. The detailed description of the network training is as follows:

An autoencoder is trained to learn the TCM operation on the original image (palm vein or finger vein). For that, the original image is shown to the network and asked to generate the TCM of that image. Then a second auto-encoder is trained which takes TCM image as input and the network has to learn IRT operation on TCM images. This autoencoder is also trained on the same number of images. Finally, we merged these two autoencoder models (end to end) from original to final transformation to create one deep autoencoder whose final task is to output the IRT image from the original image itself as shown in Figure 3. Thus, the combined autoencoder is then fine-tuned on 600 palm vein and 2400 finger vein images. The visual feature based appearance for palm vein and finger vein is depicted in Figure 4, which clearly highlights the curvilinear structures.

2.3 Siamese Matching Network

In order to match the multi-channel features, we train a Siamese network using a triplet loss function. The network consists of a Fused Feature Extractor (FFE) which gives us the feature embedding for all 5 traits of any single subject. Over these embedding, we apply triplet loss to train the FFE. Once the network is trained, we match the samples using L_2 distance between the fused feature vectors obtained from FFE.

Triplet Loss Function: The feature embedding from the FFE should be similar for a particular subject and dissimilar for all different subjects. To ensure the previous assumption, we train our network using triplet

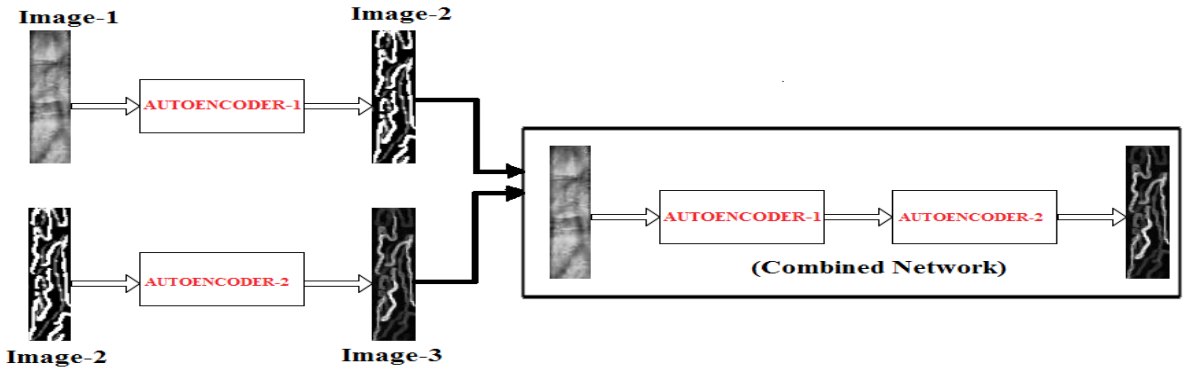


Figure 3: Combined Autoencoder Architecture.

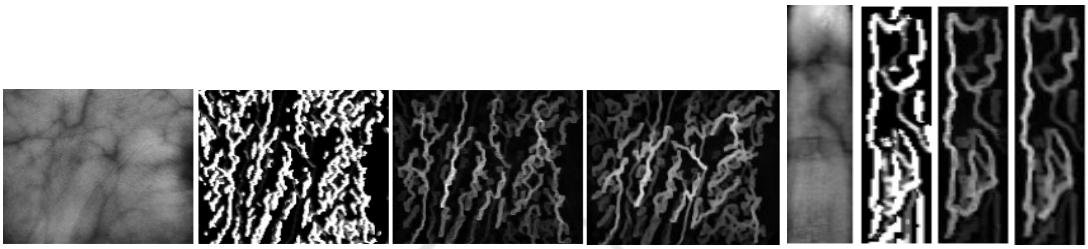


Figure 4: CNN based Image Transformation: (a) Palm Vein Samples, (b)Finger Vein Samples.

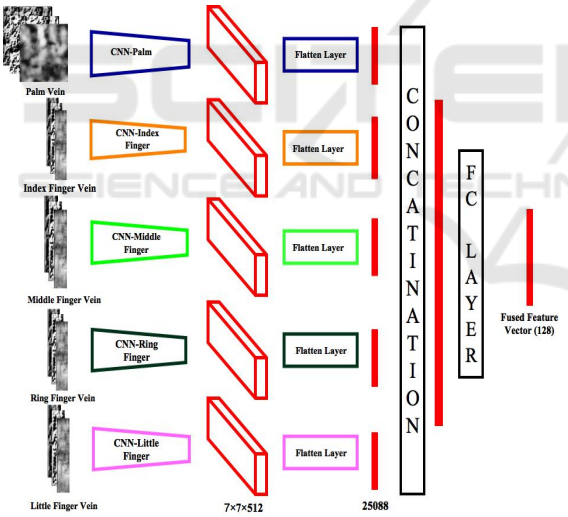


Figure 5: Fused Feature Network.

loss. The loss function ($D(a,p,n)$), based on the distances between anchor (a), positive (p) and negative (n) embedding have been used to train the network as defined below (where $M = 32$ from hinge loss):

$$D(a,p,n) = \frac{1}{2} max(0, M + L_2^2(a,p)) - L_2^2(a,n) \quad (2)$$

Fused Feature Extractor (FFE): For extracting the multi-modal fused embedding, we use a Fused Feature Extractor as shown in Figure 5. It contains 5 CNN models for five traits (palm, index finger, middle finger, ring finger and little finger). The architectures of palm CNN differs from that of 4 finger CNN's, as the size of palm images are different from that of finger images. But since the 4 fingers, can also be treated as different modalities, the weights are not shared between these networks. The architectures of the palm and finger CNN are shown in Table 1.

Each CNN gives us a $(7 * 7 * 512)$ dimensional feature vector that is flattened for fusion. These flattened vectors are then concatenated. A fully connected layer of 128 neurons is applied over the concatenated vector, giving us a 128 dimensional 5 trait fused feature vector. This has been done to achieve a feature level fusion using multi-channel network. This will ensure that each of the individual networks optimizes not only for their corresponding traits but for multiple traits and too simultaneously in one shot. It has been observed that such a network can achieve a better generalization ability.

Network Training: Deep networks require a lot of training data to learn the features. Since that luxury we do not have in biometrics, to ensure that the features learned are robust enough we first pre-train the CNNs' in the FFE using an autoencoder. This is an efficient procedure to train a network when not much data is available. Such an autoencoder can let network to initially understand how to regenerate the same input image by learning the holistic image generative features. Later we remove the decoder part and train the encoder to fulfil our objective to finally tune pa-

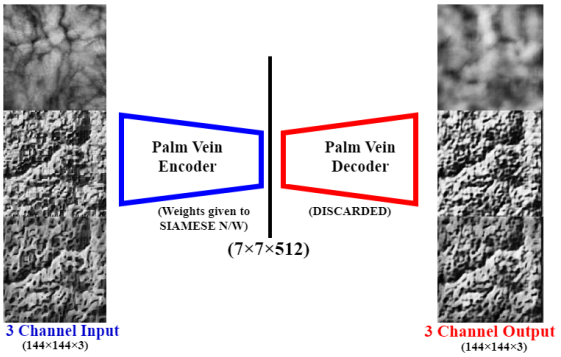


Figure 6: CNN Architecture for Palm Vein. Trained a 3-channel autoencoder to pre-train the encoder of Siamese network.

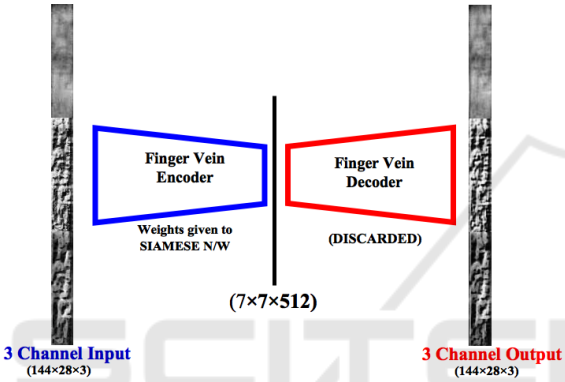


Figure 7: CNN Architecture for Finger Vein. Trained a 3-channel autoencoder to pre-train the encoder of Siamese network.

rameters to learn the best discriminative features.

We have designed an autoencoder whose encoder part is same as the CNN to be used in FFE. The input to encoder being the multi-channel feature image and output is a feature map of size $(7*7*512)$. The decoder has been fixed as a mirror of the CNN encoder, which will take the feature map from the encoder as an input and output the multi-channel feature image. We have trained this autoencoder over the gallery dataset ensuring us that the encoder has learned all the generative features present in the images. Once the autoencoder is trained, we discard the decoder and save the weights of the encoder and use it as a pre-trained weights in the FFE as shown in Figure 6 and Figure 7. Later we fine tune the FFE for the discrimination task using the triplet loss (as discussed above). Once the Siamese model is trained, we combine both Autoencoder and Siamese network and train the whole network in an end-to-end fashion.

Table 1: CNN Architecture for Palm Vein.

Palm Vein
Input Size $144 \times 144 \times 3$
CONV 32(9,9), Stride (1,1), Activation: Relu
MaxPool(2,2), Stride (2,2)
CONV 64(7,7), Activation: Relu, Padding(1,1)
MaxPool(2,2), Stride (2,2)
CONV 128(5,5), Activation: Relu, Padding(1,1)
MaxPool(2,2), Stride (2,2)
CONV 128(3,3), Activation: Relu, Padding(1,1)
MaxPool(2,2), Stride (2,2)
CONV 256(3,3), Activation: Relu, Padding(1,1)
CONV 512(3,3), Activation: Sigmoid, Padding(0,0)
Output Size $7 \times 7 \times 512$
Finger Vein
Input Finger Vein Size $144 \times 28 \times 3$
CONV 32(9,9), Activation: Relu, Padding(1,1)
MaxPool(2,1), Stride (2,1)
CONV 64(7,7), Activation: Relu, Padding(1,1)
MaxPool(2,1), Stride (2,1)
CONV 128(5,5), Activation: Relu, Padding(1,1)
MaxPool(2,2), Stride (2,2)
CONV 128(3,3), Activation: Relu, Padding(1,1)
MaxPool(2,2), Stride (2,2)
CONV 256(3,3), Activation: Relu, Padding(1,1)
CONV 512(3,3), Activation: Sigmoid, Padding(0,1)
Output Finger Vein Size $7 \times 7 \times 512$

3 EXPERIMENTAL ANALYSIS

In distinguishing experiments, the performance of the proposed method has been evaluated in terms of EER (Equal Error Rate), and DI (Decidability Index).

Database Specifications: The proposed system has been tested on publicly available Multispectral CASIA palm print database (43,), which consist of 6 images per subject for all the palm-vein and finger-vein traits. The left and right-hand samples from a subject are considered belonging to separate individuals, which resulted in 200 subjects. For each subject, the first three samples are considered as the gallery and the remaining as the probe.

Experiment 01: In the first experiment, the individual performance of finger vein and palm vein modalities have been computed. The corresponding ROC characteristics for palm vein and finger vein are shown in Figure 8. The matching between the gallery and probe samples are computed resulting in 1800 genuine and 358200 imposter scores. The values of EER and DI corresponding to the best performance for each modality have been highlighted in Table 2. To make a fair comparison, six state-of-the-art methods are reported. From table 2, we can make the following inferences, Firstly, palm vein samples achieved superior performance than any of the fin-

Table 2: Comparative Analysis (P: Palm; F: FingerVein; (+) : Fusion; 4F : All four fingervein).

	Evaluation Parameters				
Dataset	CASIA				
Approach	P/F	P+1F	4F	P+4F	DI
WLD (Huang et al., 2010)	6.08	6.11	-	5.23	-
MPC (Choi et al., 2009)	4.54	4.92	-	3.32	-
RLT (Miura et al., 2004)	17.17	18.56	9.00	13.98	-
NMRT (Zhou and Kumar, 2011)	9.58	7.88	-	6.06	-
Hessian phase (Zhou and Kumar, 2011)	9.39	9.09	-	7.24	-
Proposed Palm vein	4.12	4.02	-	-	**2.50**
Proposed Finger vein	8.50	-	6.45	-	2.15
Proposed Fused All	-	-	-	3.33	3.05

ger vein traits because of bigger and better ROI. Secondly, the proposed deep network clearly outperforms the state-of-the-art for palm-vein. The EER values of 4.125 % and 4.02% have been obtained over CASIA database. Among the individual performance of finger-vein traits, the best EER value of 8.50 % has been achieved by -middle finger. Therefore it is clear that the palm-vein is more discriminatory trait as compared to the finger-vein traits.

Experiment 02: The second experiment studies about the feature level fusion of palm vein and finger vein modalities. In addition to this, the combined performance of 4 fingers has also been evaluated. A low EER value of 3.33% has been achieved with feature level fusion of palm vein and four finger veins, which is superior to the EER values obtained from any of the 5 state of art methods as well as the proposed palm vein or finger vein individual networks. This justifies the strength of network learned fused feature representation. The detail description of other parameters is given in Table 2. The corresponding ROC characteristics and genuine Vs imposter score distribution graph for fused (palm vein, 4 finger vein) and fused finger vein are shown in Figure 9 and Figure 10.

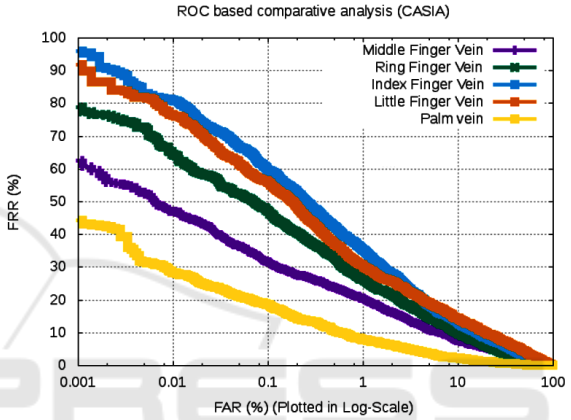


Figure 8: ROC based Performance Analysis for Palm Vein and Finger Vein.

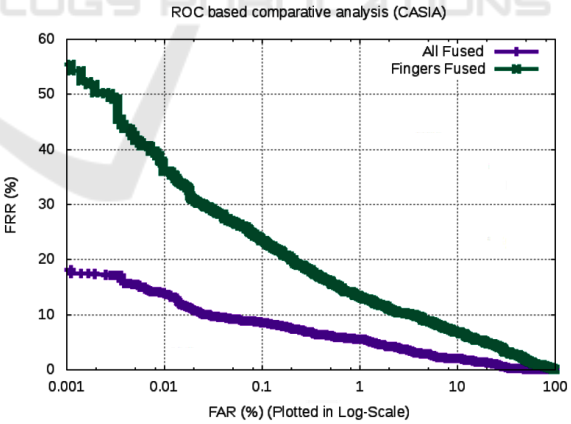


Figure 9: ROC based Performance Analysis for Fused Features.

4 CONCLUSIONS

In this work, we have proposed a novel end-to-end deep network design by combining domain specific knowledge and deep learning representation. The various challenging issues related to vein biometrics have been addressed suitably. The fixed size ROI images have been given to an end-to-end CAE augmented with Siamese network trained using triplet loss for vein recognition. Finally, in order to utilize the information present in the whole Palmer region of the hand, feature level fusion of palm vein and finger vein has been performed. As a part of future work, we will try to incorporate CNN based ROI segmentation and ROI enhancement network to further improve the recognition performance of the proposed system.

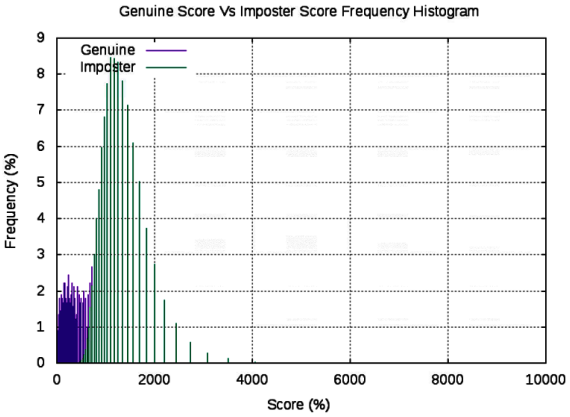


Figure 10: Genuine Vs Imposter Score Distribution (All traits fused together).

REFERENCES

Casia m.: Palmprint v1 database [online]. available:. http://www.cbsr.ia.ac.cn/ms palmprint.

Ahonen, T., Hadid, A., and Pietikainen, M. (2006). Face description with local binary patterns: Application to face recognition. *IEEE transactions on pattern analysis and machine intelligence*, 28(12):2037–2041.

Bhilare, S., Jaswal, G., Kanhangad, V., and Nigam, A. (2018). Single-sensor hand-vein multimodal biometric recognition using multiscale deep pyramidal approach. *Machine Vision and Applications*, 29(8):1269–1286.

Choi, J. H., Song, W., Kim, T., Lee, S.-R., and Kim, H. C. (2009). Finger vein extraction using gradient normalization and principal curvature. In *Image Processing: Machine Vision Applications II*, volume 7251, page 725111. International Society for Optics and Photonics.

Cummings, A. H., Nixon, M. S., and Carter, J. N. (2011). The image ray transform for structural feature detection. *Pattern Recognition Letters*, 32(15):2053–2060.

Fang, Y., Wu, Q., and Kang, W. (2018). A novel finger vein verification system based on two-stream convolutional network learning. *Neurocomputing*.

Huang, B., Dai, Y., Li, R., Tang, D., and Li, W. (2010). Finger-vein authentication based on wide line detector and pattern normalization. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 1269–1272. IEEE.

Jaswal, G., Nigam, A., and Nath, R. (2017). Deepknuckle: revealing the human identity. *Multimedia Tools and Applications*, 76(18):18955–18984.

Kimura, T., Makihara, Y., Muramatsu, D., and Yagi, Y. (2015). Single sensor-based multi-quality multimodal biometric score database and its performance evaluation. In *Biometrics (ICB), 2015 International Conference on*, pages 519–526. IEEE.

Miura, N., Nagasaka, A., and Miyatake, T. (2004). Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification. *Machine vision and applications*, 15(4):194–203.

Nigam, A., Tiwari, K., and Gupta, P. (2016). Multiple texture information fusion for finger-knuckle-print authentication system. *Neurocomputing*, 188:190–205.

Patil, I., Bhilare, S., and Kanhangad, V. (2016). Assessing vulnerability of dorsal hand-vein verification system to spoofing attacks using smartphone camera. In *Identity, Security and Behavior Analysis (ISBA), 2016 IEEE International Conference on*, pages 1–6. IEEE.

Qin, H. and El-Yacoubi, M. A. (2017). Deep representation-based feature extraction and recovering for finger-vein verification. *IEEE Transactions on Information Forensics and Security*, 12(8):1816–1829.

Ramalho, M., Correia, P. L., Soares, L. D., et al. (2011). Biometric identification through palm and dorsal hand vein patterns. In *EUROCON-International Conference on Computer as a Tool (EUROCON), 2011 IEEE*, pages 1–4. IEEE.

Taigman, Y., Yang, M., Ranzato, M., and Wolf, L. (2014). Deepface: Closing the gap to human-level performance in face verification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1701–1708.

Umer, S., Dhara, B. C., and Chanda, B. (2016). Texture code matrix-based multi-instance iris recognition. *Pattern Analysis and Applications*, 19(1):283–295.

Veluchamy, S. and Karlmarx, L. (2016). System for multimodal biometric recognition based on finger knuckle and finger vein using feature-level fusion and k-support vector machine classifier. *IET Biometrics*, 6(3):232–242.

Wang, J.-G., Yau, W.-Y., Suwandy, A., and Sung, E. (2008). Person recognition by fusing palmprint and palm vein images based on "laplacianpalm" representation. *Pattern Recognition*, 41(5):1514–1527.

Xie, C. and Kumar, A. (2017). Finger vein identification using convolutional neural network and supervised discrete hashing. In *Deep Learning for Biometrics*, pages 109–132. Springer.

Yang, J. and Zhang, X. (2012). Feature-level fusion of fingerprint and finger-vein for personal identification. *Pattern Recognition Letters*, 33(5):623–628.

Yang, W., Huang, X., Zhou, F., and Liao, Q. (2014). Comparative competitive coding for personal identification by using finger vein and finger dorsal texture fusion. *Information sciences*, 268:20–32.

Zhou, Y. and Kumar, A. (2010). Contactless palm vein identification using multiple representations. In *Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on*, pages 1–6. IEEE.

Zhou, Y. and Kumar, A. (2011). Human identification using palm-vein images. *IEEE transactions on information forensics and security*, 6(4):1259–1274.