International Journal of Electrical, Electronics and Computer Engineering **8**(1): 18-24(2019)

ISSN No. (Online): 2277-2626

Cross Spectral Face Recognition using Handcrafted and Deep Features

Akshaya Kumar Sharma[1] and Amit Shrivastava[2]

[1]Resaerch Scholar, Department of Electronic and Communication Engineering, VNS Group of Institutions, Bhopal (Madhya Pradesh), INDIA

[2]Assistant Professor, Department of Electronic and Communication Engineering, VNS Group of Institutions, Bhopal (Madhya Pradesh), INDIA

(Corresponding author: Akshaya Kumar Sharma)

(Received 06 February 2019 Accepted 15 March, 2019)

(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: Face recognition is very much popular in the present era and many researchers are working on face recognition and produce most promising results in term of recognition and human identification. It has many applications for the authentication and verification. Along with these advancements face recognition is still challenging the heterogamous environment such as near infrared and visible spectrum. Matching of face images capture in near infrared spectrum (NIR) to face images of the visible spectrum (VIS) is a very challenging task. In this research work, we propose a new approach for the face matching in the heterogeneous environment (cross-spectral scenario). The propose method uses deep learning based features and handcrafted features for matching. In the proposed method deep features are extracted from the pre-trained convolutional neural network AlexNet and shape and texture features (handcrafted features) are extracted by the histogram of oriented gradient (HOG) and Gabor, respectively. Followed by similarity scores between NIR image and the visible image are calculated and after that weighted score fusion is used for the distinguishing the genuine and imposter user. For the performance evaluation of propose cross spectral matching algorithm, experiments are performed on the CASIA 2.0 cross spectral database and proposed method achieves 94.60 % verification accuracy at the 0.1 % false accept rate (FAR).

Keywords: Cross Spectral Face Recognition, human identification, near infrared spectrum (NIR), discriminate feature extraction (CDFE), false accept rate (FAR).

I. INTRODUCTION

In this research work we have work on the face recognition system which is a very popular now days. It has very useful applications such as forensic, person identification, bank card identification [1], access control [2] and surveillance [3 4]. Face images are acquired by the camera, followed by features are extracted and stored in the database as the biometric template. For the recognition any user the similar process is repeated up to the features extraction then extracted features are matched with the stored features and decision is made as accept or reject. Face recognition has many challenges due to illumination variations, Noise removal is one of the very important aspect in the field of image processing. An image gets distorted with types of noise during the process of transmission and reception. Noise may be classified as substitutive noise speckle noise and additive white Gaussian noise [28] large dimensionality, uncontrolled environments, aging and pose variations. In the recent years, Face recognition get remarkable improvement and accuracy to overcome these challenges, but matching in the heterogamous environment such as near infrared and visible spectrum is very challenging task. Matching of face images capture in near infrared spectrum (NIR) to face images of the visible spectrum (VIS) is a very challenging task. Similarity based representation of different domain to common subspace where NIR images and VIS image have similar representation in subspace [5] proposed a common discriminate feature extraction (CDFE) in which intra-modality and intermodality local smoothing is done. Jun-Yong *et al.* [6] proposed transductive heterogeneous face matching (THFM) which learns the VIS-NIR matching from the VIS-NIR image. It also proposed feature representation based on Log-DoG filtering, local encoding, and uniform feature normalization. Yi et al. introduced canonicalcorrelation analysis (CCA) to find out the correlation between NIR and VIS faces from NIR-VIS face pairs [7]. Recently, Lei and Li [8] suggested determination the matter via coupled spectral regression (CSR). In their model, an occasional dimensional illustrationfor each face was initial computed victimization discriminative graph embedding methodology and so two associated projections were learned severally to project heterogeneous information into the discriminative common topological space for final classification. Our work conjointly mines a topological space, however our objective is for modeling domain adaptation for VIS-NIR matching in a very transductive way, whereas these connected works area unit non-transductive. Invariant features extraction can be in global based and Image is not a new technique while it was used from long year ago for the purpose of copyright protection and authentication.

This technique also makes an attempt to determine the problems associated with the management of property of media local feature based [29]. The objective of these methods to extract features which are invariant to lighting conditions. Tan and Triggs *et al.* [9] reduce the difference between NIR and VIS images by preprocessing based on Gamma correction.

Difference-of-Gaussian (DoG) filtering, Klare *et al.* [10] combine the histogram gradients (HOG) features with LBP to describe the face images. Light Source Invariant Features (LSIFs) is proposed to reduce the gap between VIS and NIR face image [11]. Goswami *et al.* introduced an efficient preprocessing chain to cut back the difference between VIS and NIR facial pictures supported Gamma correction, Difference-of-Gaussian (DoG) filtering and distinction deed [12]. Liao et al. advised encryption both VIS and NIR face pictures victimization Multi-block LBP (MBLBP) followed by DoG filtering [13]. Light AdaBoost and R-LDA were conducted for more feature choice. Following this work, Binary Laplacian of Gaussian (LoG) was also investigated in [14]. Recently, Liu et al. projected light-weight Source Invariant options (LSIFs) to fill the gap between VIS and NIR face pictures [15]. Yi *et al.* [16] used canonical correlation analysis primarily based learning in linear discriminate analysis (LDA) topological space for matching. Random subspaces primarily based ensemble of classifier is used alongside nearest neighbor (NN) and distributed illustration primarily based matching. Similarly, Maeng *et al.* [17] used HOG options for cross-spectrum and cross-distance face matching. Most of those algorithms are evaluated on tiny scale datasets, like heterogeneous face biometrics (HFB) dataset [18] and CARL [19] that comprises limited range of subjects and/or vague experimental protocols. Therefore, claims concerning generalize ability of performance may not be created with confidence and benchmarking will be difficult.

II. METHOD

The proposed approach is uses convolutional neural network and HOG gabor based feature extraction for cross spectral matching. For image matching we uses cosine similarity to generate similarity score between Visible and NIR features separately. At the final stage we have performed score fusion and threshold to separate the genuine and imposter user. Following section describe the detail description of each stage. Block diagram of the proposed cross spectral matching algorithm is shown in figure 1. Visible and NIR spectrum images are given as input then deep features and handcrafted features (texture, shape, and illumination invariant features) are extracted separately from the visible as well as from NIR images. Matching score of the different features are combined (fused) then apply threshold which differentiates between the genuine and imposter users.

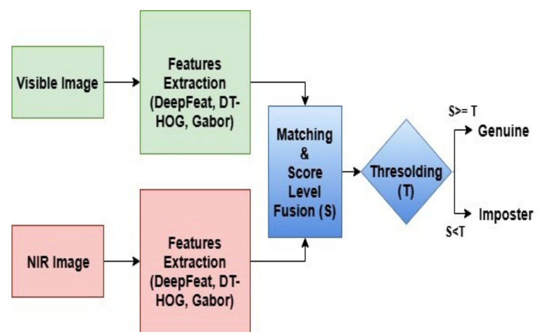

Fig. 1. Block diagram of proposed Approach.

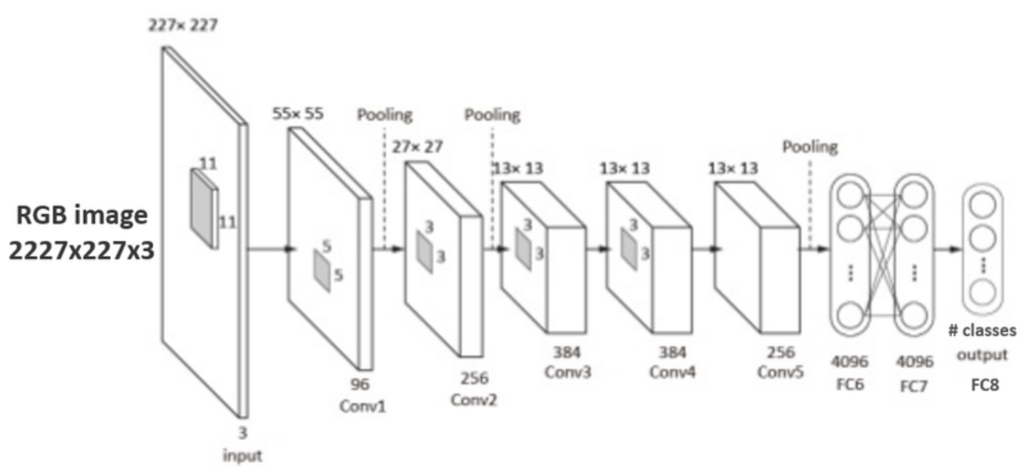

Fig. 2. AlexNet Layered architecture.

A. Features Extraction

After pre-processing of images, feature extraction step is applied. In the proposed approach features are extracted. Histogram of oriented gradient (HOG) [20] features which are invariant to changes in rotation, scale, translation, illumination, and noise. Gabor filter based features, in which magnitudes of the filtered image at each pixel are considered as features.

Deep learning based features shows best results compared to handcrafted features shows in the literature [21]. In this work AlexNet CNN is used for the extraction of deep features. Deep learning is very popular now days for the pattern recognition and computer vision. Due to the Image widespread use of internet connections leads to the vibrant accessing of digital content. The computer networks are more susceptible to penetration and thus steal or transform digital data [30]. Specially, convolutional neural network is become very important for image recognition. CNN can be used for recognition as well as for feature extraction; in this work we have used pre-trained AlexNet for the deep features extraction. AlexNet consist of 25 layers which trained on 1.2 million images of 1000 categories [22]. AlexNet architecture is shown in the figure 2. Once we get the feature vector from the features extration step then consine similarity is calculated on the features vector of visible image and the feature vectors of the NIR image. Cosine similiraty is calcuated seperatly for the each features like between HOG to HOG, Gabor to Gabor, and DeepFeat to DeepFeat featers. Cosine similarity is calculed by the cosine formula which mathematically represented as:

$$\cos\theta = \frac{Q_i \cdot T_i}{\sqrt{\sum_{i=1}^{n} Q_i^2}\sqrt{\sum_{i=1}^{n} T_i^2}}$$

Where Q_i and T_i represent query (NIR) image and template (Visible) image respectably.
After calculation of matching score fusion is performing in weighted manage to improve the matching efficiency of the system. Finally, the threshold is applied on the fused score to distinguish between the genuine and imposter users. Matching score of the images are greater than or equal to the threshold then the user is genuine otherwise it is imposter user. This image Face Recognition is-based approach can be called "active" since it embeds a unique Cross Spectral Face the encrypted flows by slightly adjusting the timing of selected packets and it does not make any limiting assumptions about the distribution or random process of the original inter-packet timing of the packet flow [31].

B. Cross Spectral Matching Algorithm
In this work first we have study different face recognition methods.

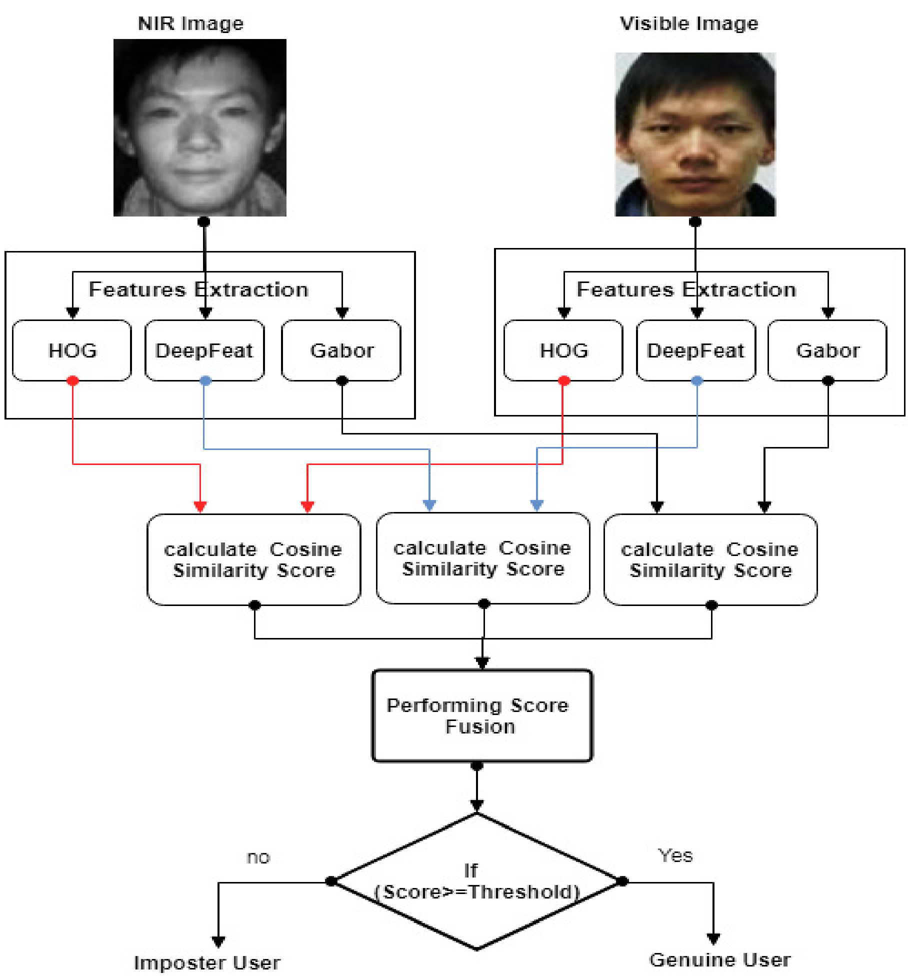

Fig. 3. Flowchart of proposed score level fusion based face matching algorithm.

We investigated the advantage and limitation of cross-spectral matching in a heterogeneous environment and proposing new approach, which extracted Histogram of oriented gradient (HOG), Gabor and deep features from the visible and NIR spectrum face images. After that we will fuse the matching score and at last step thresholding to separate the genuine and imposter user. Fig. 3 shows the data flow diagram of the algorithm. The cross-spectral matching algorithmis given in Fig. 3.

Algorithm: Face matching in cross spectral

Input: NIR and Visible spectrum image

Output: Genuine or imposter user

Step 1. Acquire image from the input and preprocessing the image by DOG normalization.

Step 2. Apply Feature extraction on the preprocessed image using Dalal and Triggs HOG (DT-HOG), Gabor filtering and deep Features.

Step 3. Calculate Cosine similarity between DT-HOG features extracted from visible image and near infrared image and generate similarity score. Similarly generate similarity score from the DeeFeat and Gabor feature.

Step 4. Get similarity score from the step 3 and perform weighted fusion of score.

Step 5. Apply Thresholding on the fused score and distinguish the genuine and imposter user as :

IF Score > = Threshold
THEN Genuine user
ELSE Imposter user
End IF

III. IMPLEMENTATION AND EXPERIMENTAL RESULTS ANALYSIS

A. Dataset Discription

In this research work we have used publically available face cross spectral database named as CASIA 2.0 [23]. CASIA is a bench mark and highly cited database many researcher are publish their results on this database. CASIA 2.0 has face images of 80 people of different age group. For the single person there are 9 images in visible spectrum and 9 images in the near infrared spectrum so total 720 images of size 128X128 pixels sample images of the visible and NIR are shown in the Fig. 6 (a-d respectively).

B. Protocol

To evaluate the proposed method we have create a verification system in which 1 to 1 matching is between the probe and query image.

Table 1: Performance comparison of various features extraction method and score fusion based method at the 0.1% FAR in the intra class environment (NIR to NIR).

Visible to Visible Matching		
Features	**EER (%)**	**Verification Accuracy (%)**
Gabor + DeepFeat	6.33	92.41
DeepFeat + HOG	4.78	95.28
HOG + Gabor	5.14	94.69
HOG + Gabor + DeepFeat (proposed)	**1.91**	**98.83**
HOG	5.22	94.63
DeepFeat	23.79	69.20
Gabor	6.54	92.28

In the verification system query image is match with the corresponding user`s gallery image. We have considered two cases for the matching. In the first case, matching is performed in the same spectrum as visible to visible and NIR to NIR. In second case, cross spectrum matching (NIR to VIS) is performed. Visible images are stored in the template as the gallery image and NIR image are matched with the gallery image.

IV. RESULTS DISCUSSION

A. Visible to Visible Matching

As we mentioned earlier, for experiment we have used CASIA cross spectral database. Here in we are going to present our experiments results in the visible to visible, NIR to NIR and NIR to visible matching. Table 1 shows the result of intra class modality matching. In the table all verification rate (genuine accept rate) is presented at the 0.1% false acceptance rate and their pictorial representation is given in the figure 4.

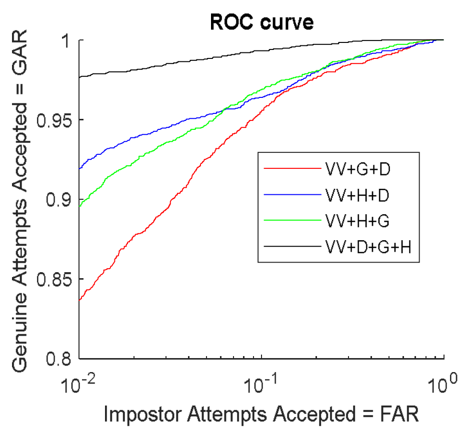

Fig. 4. ROC curve,after score fusion in visible spectram; Gabor+DeepFeat, HOG DeepFeat and Gabor+HOG+DeepFeat 2) NIR to NIR Matching.

Table 2 shows the results for the NIR to NIR matching. In NIR spectrum all feature extraction method are gives significantlly good results that can be oserved in the ROC curve shown in the Fig. 5.

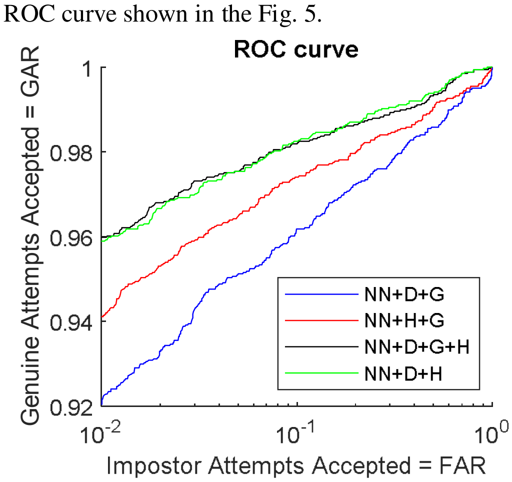

Fig. 5. ROC curve, after score fusion in NIR to NIR matching; Gabor + DeepFeat, HOG DeepFeat and Gabor + HOG + DeepFeat.

Table 2: Performance comparison of various features extraction method and score fusion based method at the 0.1% FAR in the intra class environment (NIR to NIR).

NIR to NIR Matching		
Features	**EER (%)**	**Verification Accuracy (%)**
Gabor + DeepFeat	4.92	95.09
DeepFeat+HOG	3.00	97.03
HOG+Gabor	3.77	96.45
HOG + Gabor + DeepFeat (proposed)	**2.85**	**97.53**
HOG	3.03	97.47
DeepFeat	6.08	93.06
Gabor	5.4	94.54

C. Visible to NIR Matching

Performance of proposed algorithm in cross spectral scenario is given in the table 3, as we can observed from the table fusion based approach is out performed in the all three cases and gives the significant improvement in terms of verification accuracy. In the figure 6 accuracy is 95.52 % at the 0.1 % FAR and Equal error rate is also very less.

Table 3: Performance comparison of various features extraction method and score fusion based method at the 0.1% FAR in the cross spectral environment (inter class NIR to VIS matching).

NIR to Visible Matching		
Features	**EER (%)**	**Verification Accuracy (%)**
Gabor + DeepFeat	19.04	70.06
DeepFeat+HOG	15.00	69.10
HOG+Gabor	27.00	27.70
HOG + Gabor + DeepFeat (proposed)	**5.30**	**94.60**
HOG	30.54	27.09
DeepFeat	44.00	17.08
Gabor	83.87	11.39

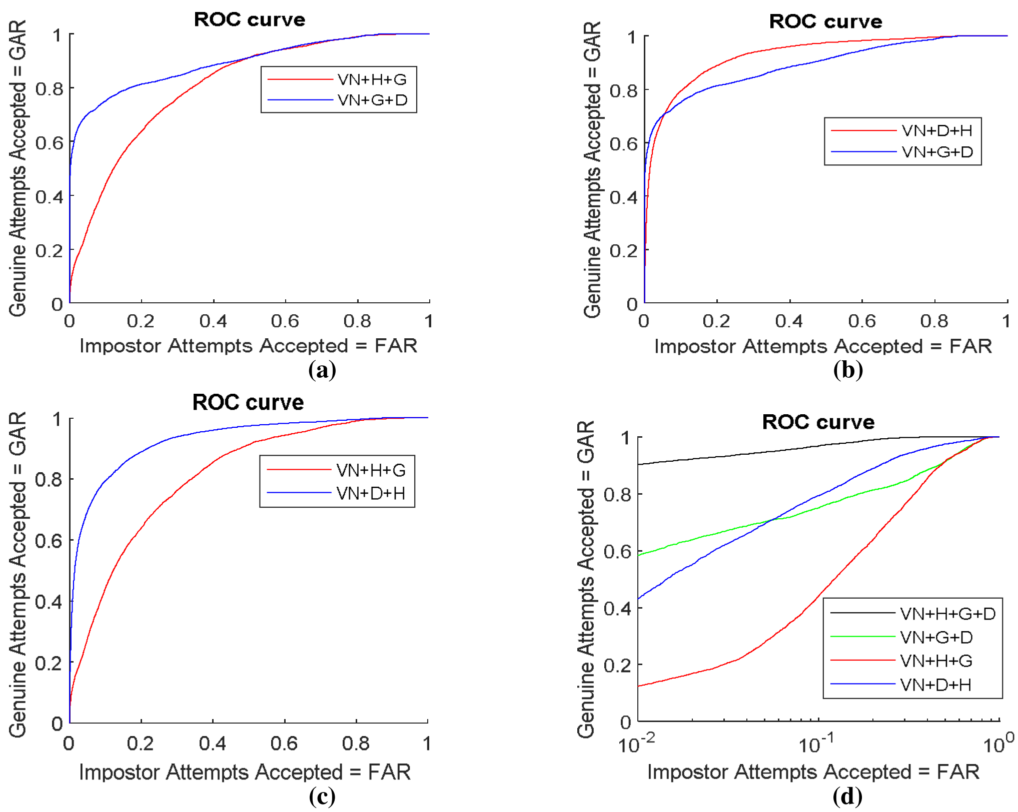

Fig. 6. ROC curve of proposed approach NIR to Visible Matching on CASIA database; a) Gabor + DeepFeat and HOG + Gabor; (b) Gabor + DeepFeat and HOG+DeepFeat; (c) HOG + DeepFeat and HOG + Gabor; (d) Gabor + DeepFeat, HOG DeepFeat and Gabor + HOG + DeepFeat.

C. Results comparison with exiting work

Comparison is shown in the table 4 with the recently published method.

As we can seen from the table, proposed method is out perform on the CASIA database compare to other matching methods.

Table 4: Comparison of proposed approach with existing method.

Method	Recognition Rate (%) @ 0.1% FAR
S. Z. Li *et al.,* [24]	19.27
F. Juefei-Xu *et al.,* [25]	85.80
X. Liu *et al.,* [26]	91.03
H. Shi *et al.,* [27]	89.91
Proposed method	**94.60**

IV. CONCLUSION

Face recognition in the heterogenous environment is very challenging task. Specifically when face images of the Near-infrared (NIR) to match with the visible face images. Recently many researchers have been work on this problem and provide the satisfactory solution. In this research work we have also proposed a matching algorithm for NIR to VIS cross spectral matching. Proposed method is based on the image preprocessing and score fusion in weighted manner. In this research work we have combined the deep features and hand crafted features and similarity scores between NIR image and the visible images have been calculated and followed by fusion of score has performed in a weighted manner. At final stage, thresholding is done, to distinguishing the genuine and imposter user. Our proposed method achieved 98.83 % verification accuracy for visible to visible matching, 97.53 % for NIR to NIR matching and 94.60 % accuracy for the heterogeneous environment. All results have taken at the 0.1 % FAR.

REFERENCES

[1]. C. Sanderson, (2008). Biometric Person Recognition: Face, Speech and Fusion. VDM Publishing, 2008.

[2]. S. Ouyang, T. Hospedales, Y. Song, and X. Li. (2014). A survey on heterogeneous face recognition: Sketch, infra-red, 3d and low-resolution. In arXiv preprint arXiv:1409.5114.

[3]. G. H. K. and S. T. (2012). Inter-modality face sketch recognition. In Multimedia and Expo, *IEEE International Conference on*, pages 224–229.

[4]. B. Klare and A. K. Jain (2010). Sketch-to-photo matching: a feature-based approach. In Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, volume 7667, page 1.

[5]. Jun-Yong Zhu, Wei-Shi Zheng and Jian-Huang Lai, (2014). "Matching NIR Face to VIS Face Using Transduction," *IEEE Trans on information forensics and security,* vol. **9** issue 3.

[6]. D. Yi, R. Liu, R. Chu, Z. Lei, and S. Li, (2007). "Face matching between near infrared and visible light images," in *Proc. Int. Conf. ICB*, pp. 523–530.

[7]. Z. Lei and S. Li, (2009). "Coupled spectral regression for matching heterogeneous faces," in *Proc. IEEE Conf. CVPR*, pp. 1123–1128.

[8]. L. Hong and A. K. Jain, (1998). "Integrating faces and fingerprints for personal identification," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. **20**, pp. 1295–1307.

[9]. X. Tan and B. Triggs, (2010). "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Trans. Image Process.,* vol. **19**, no. 6, pp. 1635–1650.

[10]. B. Klare and A. Jain, (2010). "Heterogeneous face recognition: Matching NIR to visible light images," *in Proc. 20th ICPR*, pp. 1513–1516.

[11]. S. Liu, D. Yi, Z. Lei, and S. Li, (2012). "Heterogeneous face image matching using multi-scale features," *in Proc. 5th IAPRICB*, pp. 79–84.

[12]. D. Goswami, C.H. Chan, D. Windridge, and J. Kittler, (2011). "Evaluation of face recognition system in heterogeneous environments (visible vs NIR)," in *Proc. IEEE ICCVW*, pp. 2160–2167, Nov. 2011.

[13]. S. Liao, D. Yi, Z. Lei, R. Qin, and S. Li, "Heterogeneous face recognition from local structures of normalized appearance," in *Proc. 3rd Int. Conf. ICB*, pp. 209–218.

[14]. D. Yi, S. Liao, Z. Lei, J. Sang, and S. Li, (2009). "Partial face matching between near infrared and visible images in MBGC portal challenge," in *Proc.3rd Int. Conf. ICB*, Jun. 2009, pp. 733–742.

[15]. T. Ojala, M. Pietikainen, and T. Maenpaa, (2002). "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. **24**, no. 7, pp. 971–987.

[16]. D. Yi, R. Liu, R. Chu, Z. Lei, and S. Z. Li, (2007). "Face matching between near infrared and visible light images," *in Advances in Biometrics*, 2007, pp. 523–530.

[17]. H. Maeng, S. Liao, D. Kang, S.W. Lee, and A. K. Jain, (2013). "Night time face recognition at long distance: cross-distance and cross-spectral matching," *in Proceedings of Asian Conference on Computer Vision*, pp. 708–721.

[18]. S. Z. Li, Z. Lei, and M. Ao, (2009). "The HFB face database for heterogeneous face biometrics research," *in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1–8.

[19]. V. Espinosa-Dur´ o, M. Faundez-Zanuy, and J. Mekyska, (2013). "A new face database simultaneously acquired in visible, near-infrared and thermal spectrums," *Cognitive Computation,* vol. 5, no. 1, pp. 119–135.

[20]. N. He, J. Cao and L. Song, (2008). "Scale Space Histogram of Oriented Gradients for Human Detection," *International Symposium on Information Science and Engineering*, Shanghai, pp. 167-170.

[21]. R. Roslan and N. Jamil, (2012). "Texture feature extraction using 2-D Gabor Filters," *International Symposium on Computer Applications and Industrial Electronics (ISCAIE)*, Kota Kinabalu, pp. 173-178.

[22]. Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. (2012). "Imagenet classification with deep convolutional neural networks." *In Advances in neural information processing systems*, pp. 1097-1105.

[23]. Stan Z. Li, Dong Yi, Zhen Lei and Shengcai Liao, (2013). "The CASIA NIR-VIS 2.0 Face Database", *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2013.

[24]. S. Z. Li, D. Yi, Z. Lei, and S. Liao. (2013). The casianir-vis 2.0 face database. In Computer Vision and Pattern Recognition Workshops, *IEEE International Conference* on, pages 348–353.

[25]. F. Juefei-Xu, D. K. Pal, and M. Savvides. (2015). Nir-vis heterogeneous face recognition via cross-spectral joint dictionary learning and reconstruction. In Computer Vision and Pattern Recognition Workshops, *IEEE International Conference on*, pages 141–150.

[26]. Hailin Shi, Xiaobo Wang, Dong Yi, Zhen Lei, Xiangyu Zhu, and Stan Z. Li, (2017). Cross-Modality Face Recognition via Heterogeneous Joint Bayesian, *IEEE Signal Processing Letters,* Vol. **24**, No. 1, January 2017.

[27]. Xiaoxiang Liu, Lingxiao Song, Xiang Wu, and TieniuTan, (2016). "Transferring Deep Representation for NIR-VIS Heterogeneous Face Recognition," *Int. Conf. on Biometrics (ICB)*, pp. 1-8.

[28]. M. Zahid Alam, Ravi Shankar Mishra and A.S. Zadgaonkar, (2015). Image Denoising using Common Vector Elimination by PCA and Wavelet Transform, *International Journal on Emerging Technologies* **6**(2): 157-164.

[29]. Nandan Kumar and Prof. Sneha Jain, (2018). Digital Water Marking Techniques and Uses intellectual property rights, *International Journal of Electrical, Electronics, and Computer Engineering* **7**(2): 42-50.

[30]. Nandan Kumar and Prof. Sneha Jain, (2018). A Review of Digital Water Marking in protect information copyright info privacy Techniques, *International Journal on Emerging Technologies,* **9**(2): 54-61.

[31]. Vivek Patil, (2015). Error-Free correlation in Encrypted Attack Traffic by Watermarking flow through Stepping Stones, *International Journal on Emerging Technologies,* **6**(2): 235-239.