DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
INDIAN INSTITUTE OF INFORMATION TECHNOLOGY, DESIGN AND MANUFACTURING KANCHEEPURAM
CHENNAI - 600127

Synopsis Of

Approaches for Homogeneous and Heterogeneous Face Recognition

A Thesis

To be submitted by

NILU R. SALIM

For the award of the degree

Of

DOCTOR OF PHILOSOPHY

Near infrared (NIR) face recognition has become increasingly important due to its ability to capture face images that are less sensitive to variations in illumination compared to visible light imaging (Li *et al.*, 2007). Traditional visible light imaging for face recognition can be affected by shadows, reflections, and other variations in lighting conditions, which can make it difficult to accurately recognize faces. However, NIR imaging can capture detailed facial features that are not visible to the naked eye ((Farokhi *et al.*, 2016; Peng *et al.*, 2016; Jo and Kim, 2019)). NIR imaging is also less sensitive to variations in lighting conditions, making it more robust and reliable for face recognition. NIR face recognition has many potential applications, including in security systems, access control, and law enforcement (Farokhi *et al.*, 2015). It is also being used in medical imaging, such as for detecting skin cancer or identifying facial abnormalities. Overall, the need for NIR face recognition is driven by the need for accurate and reliable face recognition technology that can operate in varying lighting conditions and capture detailed facial features. Face Recognition has often been influenced by various challenges (Jain *et al.*, 2011). It can either be intrinsic or extrinsic challenges.

Thermal spectrum imaging has gained popularity in recent years for face recognition due to its ability to capture the unique thermal patterns emitted by the human face (Lin *et al.*, 2021) mostly after the advent of COVID-19. Thermal imaging cameras can detect temperature variations on the surface of the face, which are caused by differences in blood flow, perspiration, and heat emissions (Weidlich, 2021). These patterns can be used to identify individuals, and thermal spectrum imaging is particularly useful in low light conditions where other imaging methods may struggle (Chatterjee and Chu, 2019). However, thermal spectrum imaging has its limitations, such as difficulties in capturing facial details such as skin texture and hair (Gao *et al.*, 2022). In addition, changes in body temperature, such as from fever or sweating, can also affect the accuracy of thermal imaging for face recognition. Nonetheless, thermal spectrum imaging is a promising technology that has the potential to be used in a range of applications, including security systems, access control, and medical diagnosis.

NIR- VIS (Near Infrared- Visible) cross- spectral face recognition is an emerging area of research that focuses on matching facial images captured in different spectral bands, specifically between near-infrared (NIR) and visible (VIS) domains (Wang *et al.*, 2009). NIR imaging provides advantages such as reduced sensitivity to lighting conditions and the ability to capture additional facial details, such as subsurface features (Song *et al.*, 2018). However, NIR images differ significantly from VIS images in terms of appearance, texture, and color information (Lezama *et al.*, 2017; Goswami *et al.*, 2011; Dhamecha *et al.*, 2014). The goal of NIR-VIS cross-spectral face recognition is to bridge the gap between these domains and develop algorithms that can accurately match faces captured in NIR and VIS spectra. This requires addressing challenges related to domain shift, feature extraction, and normalization techniques. Various approaches, including deep learning architectures, to learn robust feature representations that can generalize across NIR and VIS domains (Gong *et al.*, 2017; Lin and Tang, 2006). By effectively leveraging the complementary information present in NIR and VIS images, NIR-VIS cross-spectral face recognition holds promise for enhancing facial identification performance in scenarios with challenging lighting conditions and occlusions.

Th- VIS (Thermal- Visible) cross- spectral face recognition refers to the process of recognizing and verifying individuals based on their facial features captured in both thermal and visible light spectra. It combines the use of thermal imaging, which detects the heat signatures emitted by the face, and visible light imaging, which captures the visual appearance of the face. This hybrid approach leverages the unique advantages of both modalities to enhance the accuracy and robustness of face recognition systems, particularly in challenging conditions such as low lighting, variations in pose and expression, and partial occlusions (Zhang *et al.*, 2017). The implementation of thermal- visible cross- spectral face recognition has numerous practical applications. In security and surveillance systems, it can be used for access control, monitoring restricted areas, and identifying individuals in low light conditions (V. Espinosa-Duro, M. Faundez-Zanuy, and J. Mekyska, 2013). Law enforcement agencies can benefit from this technology in forensic investigations, where the combination of thermal and visible light images can provide valuable evidence for identification purposes (C. Lin, Z. Wang and G. Jong, 2020). Moreover, thermal- visible cross- spectral face recognition can be utilized in identity verification systems, border control, and attendance management, where accurate and robust identification is crucial (S. M. Iranmanesh et al., 2018).

Despite the advantages, there are also some challenges associated with thermal-visible cross- spectral face recognition. One challenge is the cost and complexity of deploying dual-modality imaging systems. Acquiring and maintaining both thermal and visible light cameras can be more expensive compared to traditional visible light-based systems. Additionally, the fusion process requires sophisticated algorithms and computational resources to align and extract features from the combined images effectively.

1 Abstract

Face recognition in the visible spectrum have been popular over the years. It is often influenced by various challenges. Out of several challenges, variations in the illumination affects face recognition in scenarios in which capturing the faces is of atmost importance. In such cases, face recognition fails due to lack of identifiable facial features which are mandatory for the task of face recognition. As a result, face images have been captured in the near infrared spectrum and deep learning methods have been applied for the face recognition. This issue has been addressed and analyzed using two methods: namely modified Resnet- 34 model with SVM classifier and an end to end light convolutional neural network model. The proposed methods have given higher accuracies on benchmark datasets CASIA NIR-VIS 2.0, Oulu-CASIA NIR-VIS, PolyU, CBSR and IIT Kh. Further, face images have been captured in the thermal spectrum and face recognition has been done. It is required in cases where there is no illumination available. This issue has been addressed and analyzed using two methods: namely an end to end light convolutional neural network model and shallow light convolutional neural network model. The proposed methods have given higher accuracies on benchmark datasets TUFTS, IRIS and CARL.

In various real- world scenarios, face images have been captured in a spectrum (NIR or thermal) and it has been matched with corresponding images in already collected in another spectrum (visible). So, face recognition happens between images captured in two different spectra. Such a scenario is called cross spectral face recognition. There are two different approaches for cross spectral face recognition. One is Direct Matching and

the other is Indirect Matching approach. In the case of direct matching approach, NIR query images have been directly matched to the corresponding visible images. Further, thermal query images have been directly matched to the corresponding visible images. However, thermal images have large difference in appearance when compared to visible images. Hence, indirect matching approach has been used, where visible images have been generated from thermal images. It has been further, matched to the corresponding visible images. Also, it has been analyzed for NIR images also by generating visible images from NIR images. It has been further, matched to the corresponding visible images. Finally, the analysis have been performed on the benchmark datasets and high accuracies have been obtained.

2 Objectives

- Near Infrared Face Recognition: Visible face images are often affected by varying illumination conditions. Near Infrared (NIR) spectrum enables to acquire high quality images, even with low illumination condition and hence it is a good method for solving the problem of illumination.

- Thermal Face Recognition: NIR images fail to capture the face temperature and requires presence of a light source to work effectively. These issues can be overcome by thermal images which is capable of capturing the face temperature (face thermograms), even in the absence of a light source.

- Cross Spectral Face Recognition Direct Matching Approach: In real world scenarios, during the time of enrolment/ registration, face images are captured in visible spectrum. Whereas, the query images have been captured in near- infrared spectrum. Since the images belong to two completely different spectrum it is called as Cross Spectral Face Recognition. So, the query face images captured in near- infrared spectrum have been matched with face databases consisting of visible face images. In this approach, query NIR images, NIR will be directly compared to that of ground truth visible images V. However, this method of direct matching fails to work with thermal vs visible face matching because most of the facial features are lost in the thermal spectrum. So, in order to handle this issue another approach called indirect face matching has been proposed.

- Cross Spectral Face Recognition Indirect Matching Approach: Matching query face images captured in thermal spectrum with face databases consisting of visible face images. During this approach, visible images have been generated from thermal images, T. Finally, the generated visible images V_t' will be compared to that of ground truth visible images V. Further, it has been analyzed on visible images have been generated from NIR images, NIR and the generated visible images V_n' will be compared to that of ground truth visible images V.

3 Existing Gaps Which Were Bridged

- Study of the effect of varying illumination in visible images and addressing it using near- infrared and thermal images.
- Comparative analysis of homogeneous (same spectral) face recognition in near-infrared and thermal spectrum.
- Comparative analysis of heterogeneous (cross spectral) near infrared- visible and thermal-visible face recognition using direct matching approach and indirect matching approach.
- Image Quality analysis of face images generated by near- infrared and thermal images.

4 Most Important Contributions

The contributions of the thesis have been highlighted in Figure 1.

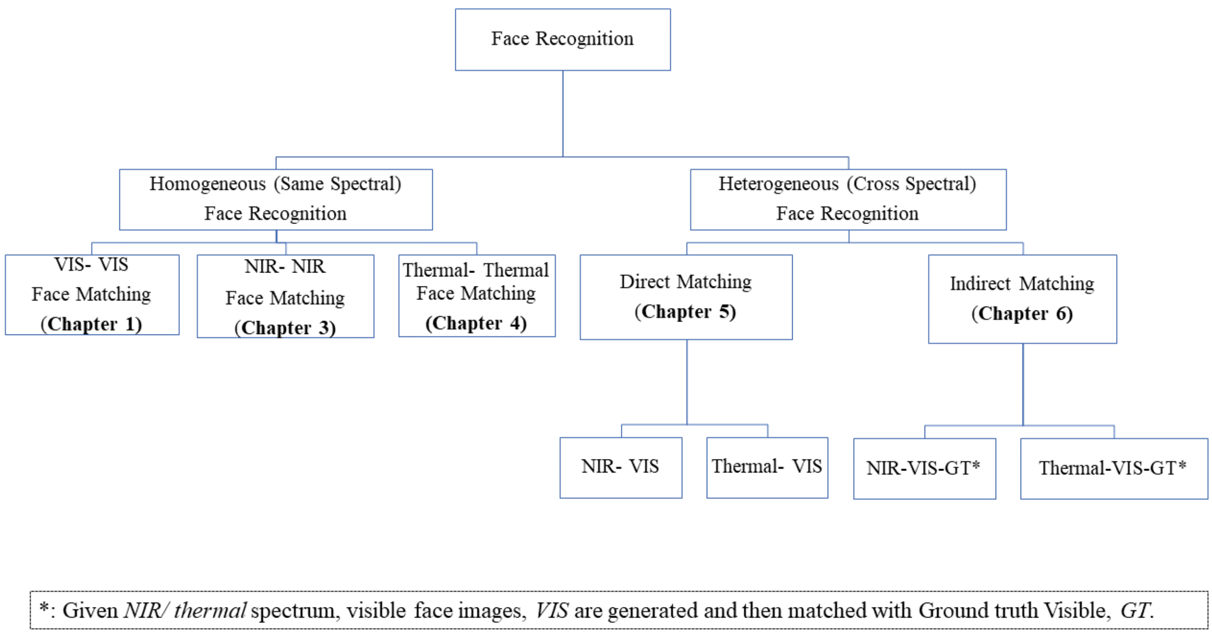

Figure 1: Contributions of the thesis

(a) Near Infrared Face Recognition has been proposed in order to address the issue of lack of proper illumination that retards the performance of visible face recognition. To solve this issue, two different models have been considered i.e. Modified Resnet- 34 model with SVM Classifier and Light Convolutional Neural Network Model.

(b) Thermal Face Recognition has been proposed in order to address the issue of complete absence of illumination or absence of light source. To solve this issue, two different models have been considered i.e. Light Convolutional Neural Network and a Shallow Light Convolutional Neural Network Model.

(c) Cross Spectral Face Recognition: Direct Matching Approach has been proposed to solve the problem of heterogeneity in the images that have been captured using different capturing modalities/spectrum. In the case of direct matching approach, the NIR query images have been directly matched to the corresponding visible images. For NIR- VIS Cross Spectral Face Recognition, two different models have been considered i.e. Modified Resnet- 34 model with SVM Classifier and Light Convolutional Neural Network. In the case of Thermal- Visible Cross Spectral Face Recognition, direct matching thermal query images to the corresponding visible images fails to give high performance accuracies due to the large variation in facial features in visible and thermal images. Hence, indirect matching approach has been used the overcome this issue.

(d) Cross Spectral Face Recognition: Indirect Matching Approach has been proposed to solve the problem of heterogeneity in the images that have been captured using different capturing modalities/spectrum. Since, facial features in thermal images vary largely when compared to visible images, indirect matching approach has been incorporated. During this approach, visible images have been generated from thermal images, T. Finally, the generated visible images V_t' will be compared to that of ground truth visible images V. Further, it has been analyzed on visible images have been generated from NIR images, NIR and the generated visible images V_n' will be compared to that of ground truth visible images V. In order to perform this task, pix2pix GAN has been used. Further, face recognition has been done using Light CNN and Shallow Light CNN models.

5 Conclusions

Face recognition in the visible spectrum fails to recognize faces due to the effects of varying illumination. In order to overcome this issue, robust images has been captured in the NIR spectrum. In homogeneous (same spectral) near- infrared face recognition, two methods that simultaneously performs face recognition, gender classification and facial expression recognition have been studied. The proposed method is based on transfer learning which is based on a modified version of Resnet- 34 model using small scale NIR images. It has been further compared to end to end Light CNN model. The proposed CNN model consists of 4 convolution layers each followed by a max pooling layer. It is trained and evaluated on the publicly available datasets. All the images in the dataset are considered during the experimentation and images from a few datasets are augmented to make the dataset more challenging. The proposed architecture has given higher accuracies with 0.12Million parameters.

In homogeneous (same spectral) thermal face recognition, two methods that performs face recognition have been studied. The proposed method is an end to end Light CNN model which has given superior performance on NIR images. The proposed architecture has given higher accuracies with 0.12Million parameters. Further, it has been compared to a Shallow Light CNN model which has given superior performance on thermal images. The proposed architecture has given higher accuracies with 0.28Million parameters.

Heterogeneous (cross spectral) direct matching approach has been proposed to solve the problem of heterogeneity in the images that have been captured using different capturing modalities/spectrum. In the case of NIR- Visible Cross Spectral Face Recognition direct matching approach, the NIR query images have been directly matched to the corresponding visible images. But, in the case of Thermal- Visible Cross Spectral Face Recognition, directly matching thermal query images matched to the corresponding visible images fails to give high performance accuracies due to the large variation in facial features in visible and thermal images. Hence, indirect matching approach has been used the overcome this issue. In heterogeneous (cross spectral) indirect matching approach, proposed architecture has given higher accuracies with 0.12Million parameters. Further, it has been compared to a Shallow Light CNN model which has given superior performance on generated visible images. The proposed architecture has given higher accuracies with 0.28Million parameters.

6 Organization of the Thesis

The proposed outline of the thesis is as follows:

(a) Chapter 1: Introduction
The generic overview of biometrics, face recognition. Face is considered as an appropriate biometric trait because of its non-obtrusiveness. It can be captured in a contactless manner, without the cooperation of the subject. In other words, the performance of the recognition process has been severely degraded by the variation in any one or all of the intrinsic and extrinsic factors. Out of all these challenges, varying illumination is one of most impactful. Further, the various challenges involved in face recognition have been discussed.

(b) Chapter 2: Literature Review
Various hand crafted feature based methods and deep learning based methods have been discussed in literature review. The implementation details and the corresponding accuracy has been discussed.

(c) Chapter 3: Homogeneous Face Recognition in Near Infrared Spectrum
The homogeneous face recognition in NIR spectrum tries to match faces on the same spectrum. The task of face recognition has to be performed in the NIR spectrum where the visible spectrum fails to work well. In this section, two methods have been proposed to extend the deep learning breakthrough for visible face recognition to the near- infrared spectrum.

(d) Chapter 4: Homogeneous Face Recognition in Thermal Spectrum
The homogeneous face recognition in thermal spectrum tries to match faces on the same spectrum. The task of face recognition has to be performed in the thermal spectrum where the NIR spectrum fails to work well.

(e) Chapter 5: Cross Spectral Face Recognition: Direct Matching Approach
Most real time face recognition scenarios involve matching a query face that has

been captured under near- infrared spectrum (NIR) against a previously acquired visible gallery database such as biometric images captured under controlled environment during enrolment/ registration. Such a scenario is called Near Infrared-Visible Cross Spectral Face Recognition (NIR- VIS CSFR). During direct matching approach, NIR query images are directly matched to the corresponding visible (VIS) images.

Further, this approach has been extended to thermal spectrum as well. So, query face that has been captured under thermal spectrum has been matched against a previously acquired visible gallery database such as biometric images captured under controlled environment. Such a scenario is called Thermal- Visible Cross Spectral Face Recognition (Th- VIS CSFR). During direct matching approach, NIR query images are directly matched to the corresponding visible (VIS) images.

(f) Chapter 6: Cross Spectral Face Recognition: Indirect Matching Approach
Due to the large variations in appearance of thermal and visible images directly matching them leads to low accuracy of the cross spectral face recognition. In such cases, visible images have been generated from NIR/ thermal images and then matched with ground truth visible images. Hence, during indirect matching approach, generated visible images, V' are matched to the corresponding ground truth visible, V images.

(g) Chapter 7: Conclusions and Future Scope
The effects of illumination in visible face recognition causes low performance due to lack of identifiable facial features. It has been overcome by near- infrared images. To perform the task of face recognition, two methods have been studied. One is based on transfer learning which is based on a modified version of Resnet- 34 model. The second one is an end to end Light CNN model. The proposed light CNN architecture has given higher accuracies with 0.12Million parameters. Further, thermal face recognition has been used to overcome the issues due to near- infrared face recognition which is not capable to capture images in complete darkness or to obtain the face temperature patterns. So, to perform the task of thermal face recognition, two methods have been studied. One is the Light CNN and Shallow Light CNN models. The second architecture has given higher accuracies with 0.28Million parameters. Finally, during Cross Spectral face recognition, direct approach has been analyzed for directly matching NIR images with visible images. Whereas, due to the large change in appearance, indirect matching approach has been followed to first convert thermal images to visible images and further, match to ground truth visible images.

7 List of Publications

I. REFEREED JOURNALS BASED ON THE THESIS

1. Nilu R. Salim, Srinath V., Umarani Jayaraman & Phalguni Gupta Recognition in the Near Infrared Spectrum for Face, Gender and Facial expressions *Multimedia Tools and Applications*, 81, 4143–4162, (2021).
2. Nilu R. Salim, Umarani Jayaraman Thermal- Visible and Near Infrared- Visible Heterogeneous Face Recognition in Small Scale and Noisy Data *Under Review*

II. PRESENTATIONS/PUBLICATIONS IN CONFERENCES BASED ON THE THESIS

1. Nilu R Salim, Umarani Jayaraman, Srinivasaraghavan Sundar, Tejas Sivan, Kongathi Mythri A Light Convolutional Neural Network Architecture for Cross-Spectral Face Recognition *IC2E3 2023*, 1-7, (2023). - Accepted
2. Nilu R Salim, Umarani Jayaraman, Srinivasaraghavan Sundar, Tejas Sivan, and Kongathi Mythri Near Infrared Face Recognition using End to End Light Convolution Neural Network *IEEE 4th Conference on Information & Communication Technology (CICT)*, 1-7, (2020).
3. Nilu R Salim, Umarani Jayaraman, and V Srinath Face Recognition in the Dark: A Unified Approach for NIR-VIS and VIS-NIR Face Matching *IEEE 4th Conference on Information & Communication Technology (CICT)*, 1-12, (2020).

III. PRESENTATIONS/PUBLICATIONS IN CONFERENCES (Others)

1. Nilu R Salim, N Sankaranarayanan, and Umarani Jayaraman Gender Classification beyond Visible Spectrum using Shallow Convolution Neural Network *IEEE Madras Section Conference (MASCON)*, 1-7, (2021).
2. Nilu R. Salim, M. Yasolakshmi sri, and Umarani Jayaraman Masked Face Detection for Effective COVID-19 Containment: A Light Convolution Neural Network based Model *9th International Conference on Pattern Recognition and Machine Intelligence (PREMI)*, (2021).

.

References

1. **C. Lin, Z. Wang and G. Jong** (2020). A de-identification face recognition using extracted thermal features based on deep learning. *IEEE Sensors Journal*, **20**(16), 9510–9517.
2. **Chatterjee, S.** and **W.-T. Chu** (2019). Thermal face recognition based on transformation by residual u-net and pixel shuffle upsampling. *In MultiMedia Modeling: 26th International Conference, MMM 2020*. Springer.

3. **Dhamecha, T. I.**, **P. Sharma**, **R. Singh**, and **M. Vatsa** (2014). On effectiveness of histogram of oriented gradient features for visible to near infrared face matching. *In 22nd IEEE International Conference on Pattern Recognition*.

4. **Farokhi, S.**, **J. Flusser**, and **U. Ullah Sheikh** (2016). Near infrared face recognition: A literature survey. *Computer Science Review*, **21**, 1–17.

5. **Farokhi, S.**, **U. U. Sheikh**, **J. Flusser**, and **B. Yang** (2015). Near infrared face recognition using zernike moments and hermite kernels. *Information Sciences*, **316**, 234–245.

6. **Gao, C.**, **X. Zhang**, **H. Wang**, **L. Song**, **B. Hu**, and **Q. Wang** (2022). Two-directional two-dimensional pca: An efficient face recognition method for thermal infrared images. *In 5th IEEE International Conference on Information Communication and Signal Processing (ICICSP)*.

7. **Gong, D.**, **Z. Li**, **W. Huang**, **X. Li**, and **D. Tao** (2017). Heterogeneous face recognition: A common encoding feature discriminant approach. *IEEE Transactions on Image Processing*, **26**(5), 2079–2089.

8. **Goswami, D.**, **C. H. Chan**, **D. Windridge**, and **J. Kittler** (2011). Evaluation of face recognition system in heterogeneous environments (visible vs nir). *In IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*.

9. **Jain, A. K.**, **B. Klare**, and **U. Park** (2011). Face recognition: Some challenges in forensics. *In IEEE International Conference on Automatic Face & Gesture Recognition (FG)*.

10. **Jo, H.** and **W.-Y. Kim** (2019). Nir reflection augmentation for deep learning-based nir face recognition. *Symmetry*, **11**(10), 1234.

11. **Lezama, J.**, **Q. Qiu**, and **G. Sapiro** (2017). Not afraid of the dark: Nir-vis face recognition via cross-spectral hallucination and low-rank embedding. *In IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

12. **Li, S. Z.**, **R. Chu**, **S. Liao**, and **L. Zhang** (2007). Illumination invariant face recognition using near-infrared images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **29**(4), 627–639.

13. **Lin, D.** and **X. Tang** (2006). Inter-modality face recognition. *In 9th European Conference on Computer Vision*.

14. **Lin, S. D.**, **L. Chen**, and **W. Chen** (2021). Thermal face recognition under different conditions. *BMC bioinformatics*, **22**, 1–17.

15. **Peng, M.**, **C. Wang**, **T. Chen**, and **G. Liu** (2016). Nirfacenet: A convolutional neural network for near-infrared face identification. *Information*, **7**, 61.

16. **S. M. Iranmanesh et al.** (2018). Deep cross polarimetric thermal-to-visible face recognition. *In International Conference on Biometrics (ICB)*.

17. **Song, L.**, **M. Zhang**, **X. Wu**, and **R. He** (2018). Adversarial discriminative heterogeneous face recognition. *In 32nd AAAI Conference on Artificial Intelligence*.

18. **V. Espinosa-Duro, M. Faundez-Zanuy, and J. Mekyska** (2013). A new face database simultaneously acquired in visible, near-infrared and thermal spectrums. *Cognitive Computation*, **5**, 119–135.

19. **Wang, R.**, **J. Yang**, **D. Yi**, and **S. Z. Li** (2009). An analysis-by-synthesis method for heterogeneous face biometrics. *In 3rd International Conference on Advances in Biometrics*.

20. **Weidlich, V. A.** (2021). Thermal infrared face recognition. *Cureus*, **13**(3).

21. **Zhang, H.**, **V. M. Patel**, **B. S. Riggan**, and **S. Hu** (2017). Generative adversarial network-based synthesis of visible faces from polarimetrie thermal faces. *In IEEE International Joint Conference on Biometrics (IJCB)*.