THERMAL IMAGING USING CNN AND KNN CLASSIFIERS WITH FWT, PCA AND LDA ALGORITHMS

Chigozie Orji , Evan Hurwitz and Ali Hasan

Department of Electrical and Electronic Engineering Science,
University of Johannesburg, Johannesburg, South Africa

ABSTRACT

This paper deals with the problem of errors in a biometric system that may arise from poor lighting and spoofing. To tackle this, images from the Terravic Facial Infrared Database have been used with Fast Wavelet Transform (FWT), an ensemble of classifiers and feature extractors, to reduce errors encountered in thermal facial recognition. By dividing the image set into a training set, comprising 1000 thermal images of 10 persons wearing glasses (X) and a test set comprising 100 image samples (y), of the same persons in glasses. A mean percentage error of 0.84% was achieved, when a Convolutional Neural Network (CNN) was used to classify the image set (y), after training with (X). However, when the images where pre-processed with Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) and k-Nearest Neighbors (KNN) classifier, a mean percentage error of 0.68% was achieved with the CNN classifier.

KEYWORDS

Thermal imaging, ensemble of classifiers, Deep Convolutional Neural Networks, K-Nearest Neighbors, Eigen Vectors, Principal Component Analysis, Linear Discriminant Analysis, biometrics, sensing, imaging, security

1. INTRODUCTION

The use of physical and behavioural attributes to recognize people is known as "biometric authentication" and this involves recognizing persons using attributes such as iris, finger print, voice sample, gait or facial features [1]. Among these attributes, the face possesses a unique set of features, for automatic detection and authentication of a person through computer vision. To successfully achieve this, under varying weather and lighting conditions, thermal imaging can be utilized to deal with some of the challenges associated with standard imaging methods. However, certain issues still exist with thermal images, such as the low quality of the images and the inability of humans to easily recognize persons in thermal images [2]. As a result feature extraction based on a robust approach, is required for person recognition using thermal images to reduce errors. In this paper, a comprehensive experimental study on facial recognition in thermal images is presented; FWT [3], PCA [4] and LDA [5] have been applied for facial feature Extraction and KNN [6] with CNN algorithms for human classification [7].

David C. Wyld et al. (Eds) : CCSEIT, AIAP, DMDB, ICBB, CNSA - 2017
pp. 133– 143, 2017. © CS & IT-CSCP 2017 DOI : 10.5121/csit.2017.70613

The focus here is to exploit the combined strengths of FWT, PCA, popularly known as Eigen faces and LDA, popularly known as Fisher faces [8]. To improve feature detection from 1000 facial thermal images, in the Terravic Facial Infrared Database [9] and achieve improved classification of the features, using KNN and CNN classifiers. Experiments show that the mean squared error for the CNN classifier is reduced through pre-processing and pre-classification with PCA, LDA and KNN algorithms respectively.

2. RELATED WORK

Debotosh et al. [10] experimented using the Terravic Facial IR Database, with two methods for face recognition: (1) Haar wavelet transform and (2) Local Binary Pattern (LBP). (1) The Haar wavelet transform method was applied with a cropping technique by binarizing the thermal faces, to separate them from their background, before feature extraction with Haar wavelet transform. (2) For the LBP method, each facial image was converted into 161 sub images, each comprising 8 by 8 pixels. These were concatenated into rows; which were processed to yield a more refined set of features using PCA. The reduced images, where then classified using a multi layer feed forward neural network and a minimum distance classifier. The weakness of this method is that, through cropping of the faces, some soft biometric information (physical and behavioural, such as material accessories worn) which can be used, to improve the matching accuracy of the classification system, are lost [11],[12],[13],[14]. Brais et al. propose recognition, in thermal images, through the detection of the eyes, nostrils and mouth, the subsequent decomposition into a feature vector with Haar wavelets, then classification using SVM and Gentle boost [15]. Christian et al. achieved an error rate reduction of up to 80% on LWIR images, on the AMROS and OTCBVS benchmark datasets, with detection using Maximally Stable Extremal Regions (MSER) and classification with a Convolutional Neural Network (CNN). With MSER acting as a hot spot detector and CNN as a classifier for detected hot spots [16].

The ensembling of classifiers, help algorithms achieve better results, however this has to be done with an understanding of the needed computing power and a knowledge of the strengths and weaknesses of individual classifiers [17], [18]. For deep CNN, Alex et al. [19] used non-saturating neurons and a GPU implementation of the CNN, to classify 1.2 million high resolution images. Despite their ability to learn complex patterns, CNNs are known to be easily fooled [20], [21]. Research work has been done, in the area of understanding CNNs, their flaws and ways to make them work better [22], [23], [24]. One is by using them with a less complex classifier, such as least squares regression and a linear classifier, this combination was used to achieve high training speed with low implementation complexity, on the MNIST test set, NORB-small and SVHN databases [25]. The same approach was used to improve the CNN algorithm's efficiency and reduce errors by [26]. Other variations of CNN-based classification are possible and would require additional research to discover. In an attempt to do this, we have analyzed the merits of ensembling a local learning algorithm, such as K-Nearest Neighbours (KNN) with Convolutional Neural Networks (CNN).

3. THE PROPOSED METHOD

We propose a combination-based approach for classifiers on the Terravic Infrared Database. By experimenting with a wavelet transform technique and known feature extraction algorithms. We found out that the combination of a spot classifier algorithm (KNN) and a deep learning classifier algorithm (CNN) reduces errors, when attempting to make predictions on samples not in the training set.

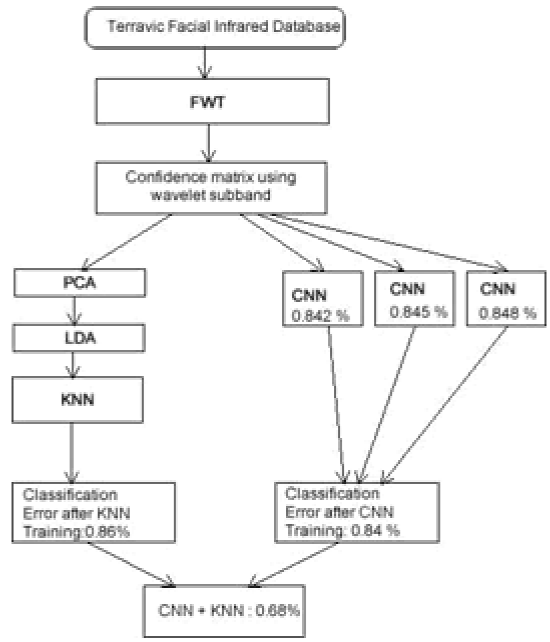

Figure 1. Proposed structure for CNN thermal face recognition combined with KNN.

3.1. The Terravic Infrared Database

The Terravic Facial Infrared Database is a sub set of the OTCBVS benchmark data set, a public database available, exclusively for educational and research purposes, for experiments in the area of computer vision algorithms and for exploring the advantages that derivable from using the invisible spectrum in real world applications. The benchmark comprises 12 separate categories of thermal images of which the Terravic Facial IR Databases is one [9].

Figure 2. Sample images from a glasses wearing group in the Terravic Facial Infrared Database

3.2. Fast Wavelet Transform (FWT)

Wavelet transform techniques are commonly used in image processing applications involving images with low signal-to-noise ratio such as thermal images [27].

It works by dividing an image into high and low frequency areas that can be represented graphically as sub bands. Several wavelet transform methods exist such as the Discrete Wavelet Transform (DWT) and Fast Wavelet Transform [3]. The Fast Wavelet Transform method, is used here, among the transform tools in Large Time Frequency Analysis Toolbox (LTFAT), in Octave, because of its simplicity and its application to thermal imaging [28].

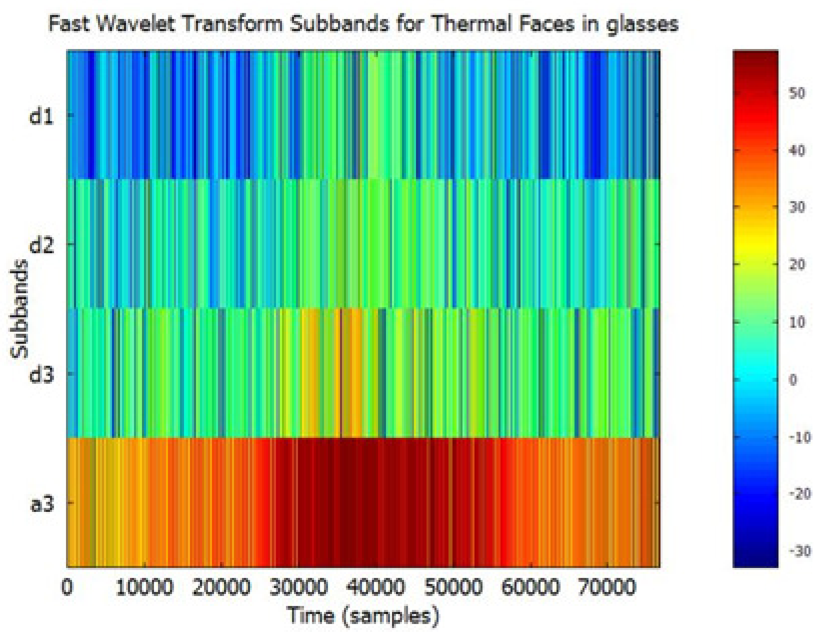

Figure 3. Visualized FWT of glass wearing images, from the Terravic Facial Infrared Database, showing spatial and frequency data represented using 4 sub bands

3.3. Principal Component Analysis (PCA)

PCA is a transform technique used for extracting points of highest variance or principal components, from data through an orthogonal projection of these points onto a new coordinate in a lower dimensional feature space. For on the spot classification using algorithms such as KNN, the fewer the number of components the easier it is to compute the most proximal example in the feature space [29]. Mathematically this is represented as:

$$AV = \lambda V \tag{1}$$

Where A=Matrix,V = Eigen vector matrix, λ = Diagonal matrix of corresponding Eigen values

The figure below shows Eigen faces from the thermal facial dataset, after principal component analysis.

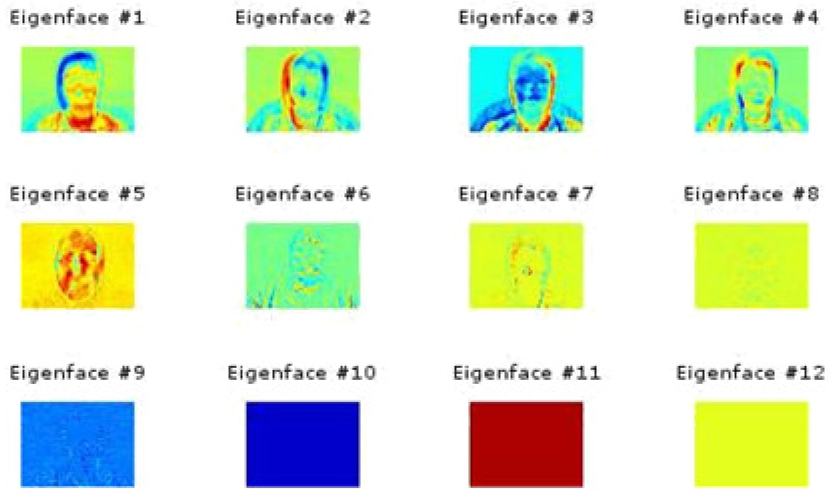

Figure 4. A Eigen face feature set, visualized after Principal Component Analysis (PCA), from the Terravic Facial Infrared Database, glass wearing images.

3.4. Linear Discriminant Analysis (LDA)

Fisher face is an enhancement of the Eigen face technique, achieving dimensionality reduction using Linear Discriminant Analysis (LDA). Unlike PCA, it seeks a projection axes between images, upon which to measure the variance of their data points, by maximizing between class differences and minimizing within class differences [30]. Mathematically this is represented as:

$$S_b V = \lambda S_w V \quad (2)$$

Where S_w = within class differences, S_b = between class differences, V = The eigen vector matrix, λ = *Eigen values*

The figure below shows Fisher faces from the thermal facial dataset, after Linear Discriminant analysis.

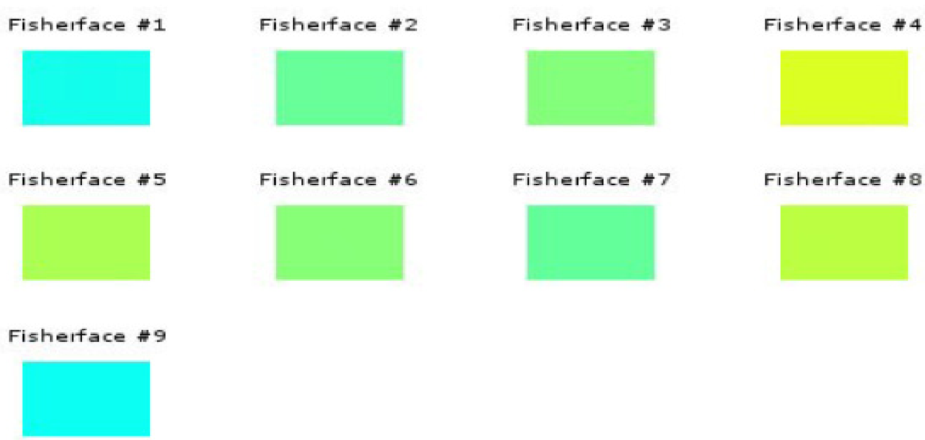

Figure 5. A Fisher face feature set, visualized after Linear Discriminant Analysis (LDA), from the Terravic Facial Infrared Database, glass wearing images.

3.5. K-Nearest Neighbors

For these experiments, involving PCA, LDA and K-Nearest Neighbours (KNN), we have used the Byte fish face recognition tool box [31]. KNN is an instance based learning classifier, which works without parameters but with the K closest training examples, in the feature space, as its input. Given a training set ranging from M_1 to M_K, it finds the best k means sample that can be used to represent the entire training set and classifies successive points Ni, based on their proximity P to the means.

$$C\ (N_i) = P\ (N_i,\ M_k) \tag{3}$$

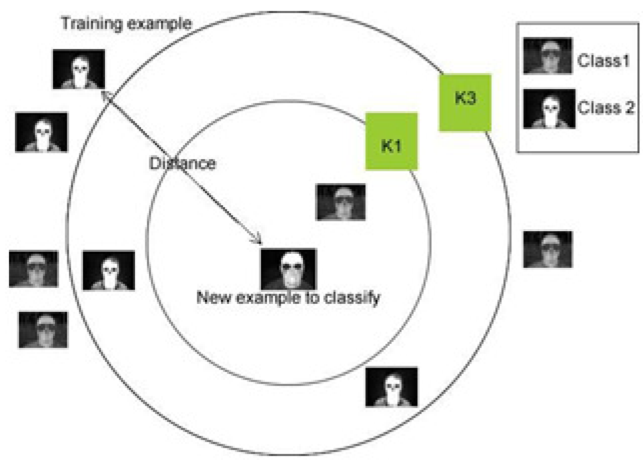

Figure 6. KNN local classification method based on distance from training example

3.6. Convolutional Neural Network (CNN)

We have used CNN tools from, Rasmus Berg Palm's, deep learning tool box [33]. The structure of our deep neural network, consists of 6 layers an input layer, 2 convolution layers, with 2 sub sampling layers and an output layer, this is shown in table I below. We trained the model several times with different initial seeds for each training session, the CNN network did a 100 epochs running 50 batches per epoch. After which the mean-squared error for the batch was plotted against the number of iterations. An error of 0.86% was achieved at the 100th iteration of the gradient descent.

Table 1. Convolusional Network Architecture

Activation function	Sigmoid
Input Layer	1
Number Convolusion Layers (CL)	2
Convolusion layer Output maps	6
Convolusion layer Kernel size	5
Number Sub sampling (pooling) layers	2
Sub sampling layers scale	2
Alpha	1
Batch size	50
Number of epochs	100

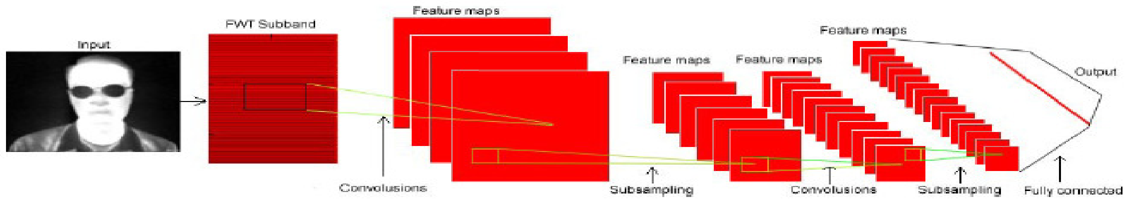

Figure 7. CNN deep classification method based on convolution and sub sampling

4. EXPERIMENTS

4.1. Fast Wavelet Transform

We experimented with Fast Wavelet Transform, using the Large Time-Frequency Analysis Toolbox (LTFAT) in Octave and produced 4 sub-bands from the 76800 by 144 data matrix, shown in Fig.3. The fourth sub band being the richest in time-frequency (TF) features was used to develop a confidence matrix (38400 by 144) which was subjected to Principal Component and Linear Discriminant Analysis.

4.2. Feature Extraction

Both PCA and LDA performed similarly on the TF features, when used with KNN for classification, producing a test set accuracy up to 16.67%, for 20 PCA and LDA components.

4.3. Classification

After training and classification with CNN and then re-training and classifying with KNN, PCA and LDA, it was observed that KNN's local prediction method improved CNN's deep and exhaustive approach to classification, and reduced errors on the TF feature set. Figures 6 and 7 below show this improvement.

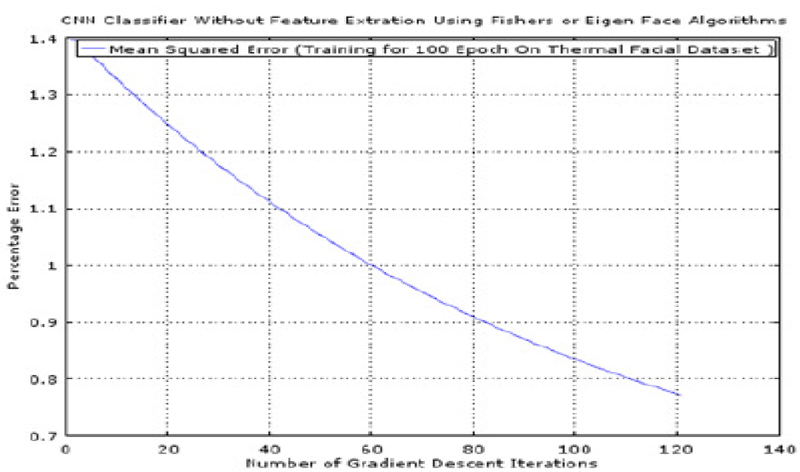

Figure. 8. CNN classification without PCA, LDA and KNN, using the time-frequency (TF) features, of images from the Terravic Facial Infrared Database, glass wearing images.

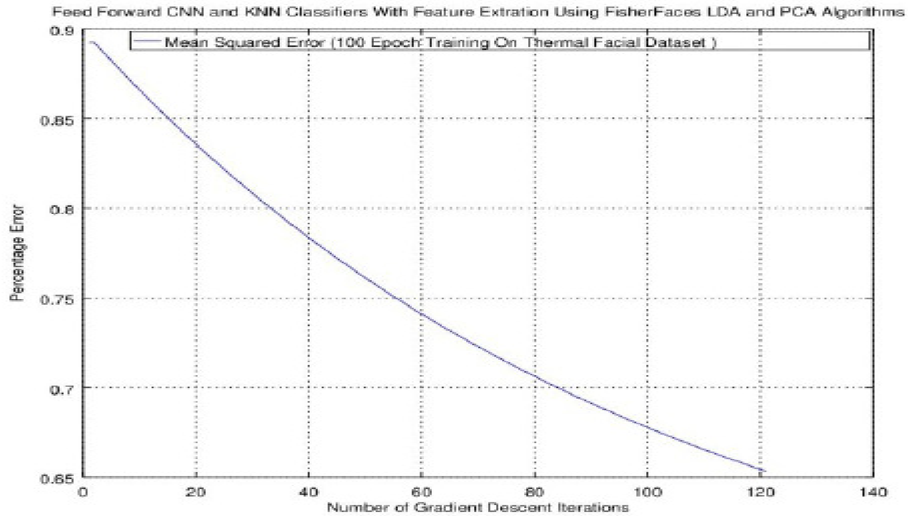

Figure. 9. CNN classification with PCA, LDA and KNN , using the time-frequency(TF) features, of images from the Terravic Facial Infrared Database, glass wearing images.

4.4. Classification Explained

The combination of FWT and CNN algorithm, achieves an average of 0.84% error, on the test set after training, by also recognizing 9 people out of 10 among 100. While the combination of FWT, PCA, LDA and KNN algorithms after training, achieves 0.86% error on the test set, by recognizing 9 people out of 10 among 100. The combined predictions show each algorithm making up for the shortcoming of the other, as the 1 person not recognized by the CNN classifier, was recognized by KNN and vice versa; making the union of results from both classifiers, yield a total decrease in the percentage error.

4.5. Results Analysis

Comparing the CNN and KNN algorithm approaches to classifying thermal facial images, we find that (1) Fast Wavelet Transform helped PCA and LDA with time-frequency (TF) features, for analysis on the facial thermal images. (2) KNN with PCA and LDA generalize well to new examples for a smaller test set of similar images, such as 40 images of 2 – 5 people wearing glasses, but do not for a larger test set, of up to 100 images of the same people. (3) CNN generalizes well for a large test set but doesn't do so for a smaller one comprising images of the same people. (4) The use of CNN with KNN, PCA and LDA on a large test set, improves classification, among images for both those with close distance in similarity, using soft biometric features such as glass wearing, and those without it.

5. CONCLUSION

The K-Nearest Neighbours (KNN) simplicity, based on distance geometrically, is demonstrated using the Terravic Facial Infrared Database. It moderates decisions made by the Convolutional Neural Networks (CNN), when classifying thermal facial images that are similar, in terms of soft-biometric features. Other soft biometric features should be experimented with, using CNN, FWT and KNN, and the results compared with the research we have presented in this paper.

Table 2. Experimental Results

Feature Extraction FWT	Result
FWT Sub-band 1	9600 features
FWT Sub-band 2	9600 features
FWT Sub-band 3	19200 features
FWT Sub-band 4	38400 features
Algorithm	Result
FWT Sub-band 4 +20 PCA+KNN	TPR 8.33%
FWT Sub-band 4 +20 PCA+KNN	Accuracy 16.67%
FWT Sub-band 4 +20 LDA+KNN	TPR 8.33%
FWT Sub-band 4 +20 LDA+KNN	Accuracy 16.67%
FWT Sub-band 4 +20PCA+120 LDA+KNN	0.86% Error (Training set)
FWT Sub-band 4 +20PCA+120 LDA+KNN	0.89% Error (Test set)
FWT Sub-band 4 +20PCA+120 LDA+KNN	0.89% Error (CV set)
FWT Sub-band 4 + CNN (session 1)	0.842% Error (Training, test and CV sets)
FWT Sub-band 4 + CNN (session 2)	0.845% Error (Training, test and CV sets)
FWT Sub-band 4 + CNN (session 3)	0.848% Error (Training, test and CV sets)
FWT Sub-band 4 + CNN	0.84% Error (Average)
FWT Sub-band 4 + CNN + FWT Sub-band 4 +20PCA+120 LDA+KNN	0.68% Error

ACKNOWLEDGEMENTS

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

REFERENCES

[1] A. K. Jain, P. J. Flynn, and A. Ross. Handbook of Biometrics. Springer, 2007.

[2] G. Bebis, A. Gyaourova, S. Singh, and I. Pavlidis, "Face Recognition by Fusing Thermal Infrared and Visible Imagery", Image and Vision Computing, vol. 24, no. 7, pp. 727-742, 2006.

[3] G. Beylkin, R. Coifman, and V. Rokhlin. Fast wavelet transforms and numerical algorithms I. Communications on Pure and Applied Mathematics, 44(2):141–183, March 1991.

[4] M. Turk and A. Pentland, "Face Recognition Using Eigenfaces," Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 1991, pp. 586-591.

[5] P.N. Belhumeur, J.P. Hespanha and D.J.Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection" IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 711-720, Jul. 1997.

[6] A. Coates and A. Y. Ng, "Learning feature representations with k-means," in Neural Networks: Tricks of the Trade. Springer, 2012, pp. 561–580.

[7] X. Chen, P. Flynn, and K. Bowyer, "PCA-based face recognition in infrared imagery : Baseline and comparative studies," IEEE International Workshop on Analysis and Modeling of Faces and Gestures, pp. 127–134, October 2003. Vision and Modeling Group, The Media Laboratory

[8] Lu, J., Plataniotis, K., Venetsanopoulos, A.: Face recognition using LDA-based algorithms. IEEE Transactions on Neural Networks 14 (2003) 195–200.

[9] IEEE OTCBVS WS Series Bench; Roland Miezianko, Terravic Research Infrared Database.

[10] D. Bhattacharjee, A. Seal, S. Ganguly, M. Nasipuri, and D.K. Basu, "A Comparative Study of Human Thermal Face Recognition Based on Haar Wavelet Transform and Local Binary Pattern," Computational Intelligence and Neuroscience, vol. 2012, Article ID 261089, 12 pages, 2012. oi:10.1155/2012/261089.

[11] D. Reid and M. Nixon. Using comparative human descriptions for soft biometrics. In Proc. of International Joint Conference on Biometrics, 2011.

[12] A. Dantcheva, C. Velardo, A. D'Angelo, and J.-L. Dugelay. Bag of soft biometrics for person identification. New trends and challenges. Multimedia Tools and Applications, 51:739–777, 2011.

[13] A.K. Jain, S.C. Dass, and K. Nandakumar. Soft biometric traits for personal recognition systems. In Proceedings of ICBA, pages 1–40. Springer, 2004.

[14] A.K. Jain, S. C. Dass, and K. Nandakumar. Can soft biometric traits assist user recognition? In Proceedings of SPIE, volume 5404, pages 561–572, 2004

[15] B. Martinez, X. Binefa, and M. Pantic. Facial component detection in thermal imagery. In Proc. IEEE Conference on Computer Vision and Pattern Recognition Workshops, pages 48–54, 2010.

[16] C. Herrmann, T. Müller, D. Willersinn, J. Beyerer, "Real-time person detection in low-resolution thermal infrared imagery with MSER and CNNs", in Proc. Electro-Optical and Infrared Systems, SPIE, 2016.

[17] T. G. Dietterich, "Ensemble methods in machine learning," International workshop on multiple Classifier systems, p. Springer.

[18] A. Rahman and S. Tasnim, "Ensemble Classifiers and Their Applications: A Review," International Journal of Computer Trends and Technology (IJCTT), vol. 10, no. 1, 2014.

[19] A. Krizhevsky, I. Sutskever and G. Hinton . ImageNet Classification with Deep Convolutional Neural Networks. In Advances in Neural information Processing Systems 25, pages 1106-1114, 2012.

[20] Nguyen A, Yosinski J, Clune J. Deep Neural Networks Are Easily Fooled: High Confidence Predictions for Unrecognizable Images. CVPR. 2015; 427-436.

[21] Szegedy C et al. Intriguing properties of neural networks. CoRR. 2013.

[22] Szegedy C et al. Going Deeper with Convolutions. CoRR. 2014.

[23] Oquab M et al. Learning and transferring mid-level image representations using convolutional neural networks. CVPR. 2014; 1717-1724.

[24] Szegedy C et al. Rethinking the Inception Architecture for Computer Vision. arXiv preprint. 2015.

[25] McDonnell MD, Vladusich T. Enhanced image classification with a fast-learning shallow convolutional neural network. In Proceedings of International joint conference on neural networks (IJCNN'2015), Killarney, Ireland, 12-17 July 2015.

[26] G.-B. Huang, Z. Bai, L. L. C. Kasun, and C. M. Vong, "Local receptive fields based extreme learning machine, " IEEE Computational Intelligence Magazine, vol. 10, pp. 18-29, 2015.

[27] P.L. Søndergaard, B. Torrésani, P.Balazs. The Linear Time-Frequency Analysis Toolbox. International Journal of Wavelets, Multiresolution Analysis and Information Processing, 10(4), 2012.

[28] L. Sirovitch and M. Kirby, “Low-Dimensional Procedure for the Characterization of Human Faces,” J. Optical Soc. of Am. A, vol. 2, pp. 519-524, 1987.

[29] M. Turk and A. Pentland, “Eigenfaces for Recognition,” J. Cognitive Neuroscience, vol. 3, no. 1, 1991.

[30] R.A. Fisher, “The Use of Multiple Measures in Taxonomic Problems,” Ann. Eugenics, vol. 7, pp. 179-188, 1936.

[31] P. Wagner, "Bytefish face recognition algorithms for MATLAB/GNU Octave and Python," 2015. [Online]. Available: https://github.com/bytefish/facerec

[32] L. Bottou and V. Vapnik. “Local learning algorithms,” Neural Computation, 4(6), pp. 888-901, 1992.

[33] "R. B. Palm,"Prediction as a candidate for learning deep hierarchical models of data," 2012. [Online]. Available: http://www2.imm.dtu.dk/pubdb/views/publication_details.php?id=6284