

Feature Level fusion based on Conical Correlation Analysis and Discriminant Correlation Analysis for Palm Print and Hand Vein

Shreyas Rangappa
Research Scholar
SJBIT, Bangalore

Naveena C.
Professor
SJBIT, Bangalore

H. K. Chethan
Professor
MIT, Mysuru

G. Hemantha Kumar
Professor
University of Mysore

ABSTRACT

Information fusion is a key step in multimodal biometric systems. Fusion of information can occur at different levels of a recognition system, i.e., at the feature level, matching-score level, or decision level. However, feature level fusion is believed to be more effective owing to the fact that a feature set contains richer information about the input biometric data than the matching score or the output decision of a classifier. The goal of feature fusion for recognition is to combine relevant information from two or more feature vectors into a single one with more discriminative power than any of the input feature vectors. In pattern recognition problems, we are also interested in separating the classes. In this paper, we present Discriminant Correlation Analysis (DCA), a feature level fusion technique that incorporates the class associations into the correlation analysis of the feature sets. DCA performs an effective feature fusion by maximizing the pairwise correlations across the two feature sets, and at the same time, eliminating the between-class correlations and restricting the correlations to be within the classes. Our proposed method can be used in pattern recognition applications for fusing features extracted from multiple modalities or combining different feature vectors extracted from a single modality. It is noteworthy that DCA is the first technique that considers class structure in feature fusion. Moreover, it has a very low computational complexity and it can be employed in real-time applications. Multiple sets of experiments performed on Palm print and Hand vein datasets, and using different feature extraction techniques.

Keywords

multimodal biometrics, feature level fusion, class structure, discriminant correlation analysis.

1. INTRODUCTION

From past few decades Biometrics which is pattern recognition technique used for recognizing humans which depends on their physiological and behavioral traits has been gained lot attention in today's research. Recent survey as shown that in today's real world most of the Biometric systems are unimodal systems which depends on one source of information for authentication [1]. Some of the widely known biometrics adopted for human identifications are single fingerprint, iris, face, voice, and DNApalmprint and hand vein. So on other hand multimodal biometric system fuses information's at various levels to obtain an accurate recognition system. Feature level, matching-score level and decision level are the various biometric information and stages used for an effective recognition system and particularly in feature level fusion the features which are extracted from different biometrics will be fused. In matching score level fusion scores which is generated from multiple classifiers

concerned to various modalities will be merged. Ultimately the decision level fusion combines the results from multiple classifiers [4].

Fusion of information through feature levels has been consider to be more effective than other levels of fusions as the feature sets contains a lot of rich information of input biometric data compared to matching score and decision making fusion levels. Serial feature fusion, parallel feature fusion and feature fusion method based on Canonical Correlation Analysis (CCA) are some of the popular feature fusion methods [3]. The working structure of serial feature fusion is by concatenating two sets of feature vectors into a one feature vector, here the first source feature vector, a , is x -dimensional and second source feature vector, b , is y -dimensional and the fused feature vector, c , is $(x + y)$ dimensional. On the other side the parallel feature vector fuses the two source feature vectors into a complex vector $(a + ib)$, where i is the imaginary unit. Important thing is that if the dimensions of the two input vectors are unequal than the one with lower dimensions is padded with zeros. CCA based feature fusion works on the correlation between two sets of feature to find two sets of transformations such that the transformed features have maximum correlation across two feature sets, while being uncorrelated within each feature set.

This paper is organized as: Section 2 describes feature level fusion and section 3 presents proposed methodology and section 4 shows experimental results and last section 5 gives conclusion and future scope.

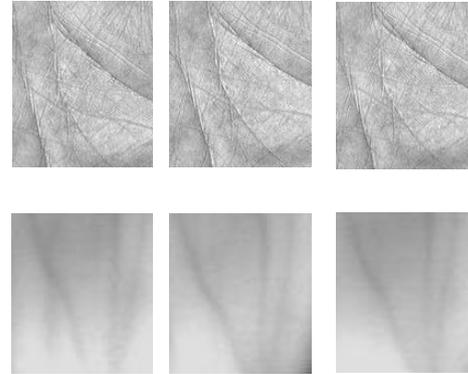
2. RELATED WORK

Veluchamy et al. [1] proposed a multimodal biometric system based on the fusion of finger knuckle and finger vein features by making use of novel technique fractional firefly (FFF) optimization method. To extract the features from finger knuckle and finger vein images line tracking method is adopted. New feature level fusion method developed FFF optimization method helps to find out the optimal weight score which is used for fusion of extracted features sets of finger knuckle and finger vein images and recognition of fused feature set make use of layered K-SVM i.e. by combining layered SVM classifier and k-neural network classifier. Gopal et al. [2] proposed a multimodal personal identification structure which uses fusion of three different biometric modalities such as palmprint, dorsal hand vein and new biometric modality "palm-phalanges print" is introduced. The proposed method adopts feature extraction, classification and fusion were as feature extraction is adopted using GMF based features, mean features and AAD features. K-nearest neighbor, support vector machine and random forest are the methods adopted for classification to authenticate the matching level and score level fusion which consists of

conventional operators, t-norm fusion operators and Yager's ordered weighted averaging are the technique applied for fusion. Haghghat et al. [3] proposed a feature fusion level approach which integrates the class associations in correlation analysis of the feature sets by adopting novel Discriminant Correlation Analysis. Maximizing the pair-wise correlations across the two feature sets and at the same time terminating between-class correlations which regulate the correlations to be within classes are the main characteristics which provides a constructive feature fusion model for real time applications in pattern recognition. Yang et al. [4] proposed a feature level fusion that incorporates the palmprint and palm vein features to provide an effective multimodal biometric based on canonical correlation analysis. The procedure of this work is applied by pre-processing stage and then extract the region of interest which also helps in enhancing the palmprint and palm vein images. To extract feature from palmprint and palm vein they adopt local binary pattern and then these features are integrated by canonical correlation analysis. Bharathi et al. [5] proposed a novel technique for biometric recognition by fusion of finger vein, palm vein and dorsal vein of hand features to provide an effective multimodal biometric system. To extract features from hand vein images they have incorporated Scale invariant and Shearlet feature transform and then these extracted features are stored in the form coefficients in the database. Finally matching of stored features and input test images are carried out by applying distance measures and fusion by maximum likelihood ratio technique. Vaidya et al. [6] proposed a multimodal biometric system which make use of contactless multispectral camera for capturing the visible and Near infrared images concurrently and by fusion of feature levels of palmprint and palm vein based on entropy technique to extract feature from palmprint and palm vein images and fusion of these features in this model make use of least features i.e. only 16 features. Li et al. [7] proposed a multimodal biometric system based on the fusion of palm print and palm vein features by adopting the Non-Halo Complex Matched Filtering (NH-CMF) for extracting features and also to enhance the system anti-spoof by the fusion technique of palm print and palm vein. The proposed method gives better matching accuracy and as well as verification processing time compared to other methods. Gaikwad et al. [8] proposed a multimodal biometric recognition system of palm print and palm vein images which fuses features of these two traits as biometric template. Edge information that incorporate texture features, principal line, wrinkles and ridges in palm print and palm vein images are detected using Contourlet Transform and for matching they have adopted Euclidean distance algorithm. Heenaye et al. [9] proposed a score level fusion technique for developing a multimodal hand vein biometric system by fusion of palmar hand vein and dorsal hand vein patterns features. To represent the features of vein Independent Component Analysis which is the most admired reduction technique has been applied. False acceptance rate and false rejection rate gives better results in multimodal biometric system than unimodal biometric system. Ashok Rao et al. [10] proposed a hand vein biometric for both condition of unimodal as well as multimodal with palmprint. To extract the hand vein features they have adopted novel method non-standard masks and gives better performance in all the conditions like multialgorithmic, unimodal and multimodal and classification is carried out using Kernel Direct Discriminant Analysis.

3. PROPOSED METHODOLOGY

Pre-processing is necessary before feature extraction. In our proposed approach, three steps are devised in the pre-processing module : 1) Image binaryzation; 2) Border tracing and segmentation; 3) ROI extracted ; finally, all images were converted to enhanced gray scale and cropped to 205×205 pixel. See Fig.1 for some palmprint and Hand vein images in the pre-processing.



A. Feature-Level Fusion Using Canonical Correlation Analysis

Canonical correlation analysis (CCA) is one of the valuable multi-data processing methods, which has been widely used to analyze the mutual relationships between two sets of variables. Suppose that $X \in \mathbb{R}^{p \times n}$ and $Y \in \mathbb{R}^{q \times n}$ denote two matrices, each contains n training feature vectors from two different modalities. That is, for each sample, two feature vectors with p and q dimensions are extracted from the first and second modalities, respectively.

Let $S_{xx} \in \mathbb{R}^{p \times p}$ and $S_{yy} \in \mathbb{R}^{q \times q}$ denote the within-sets covariance matrices of X and Y and $S_{xy} \in \mathbb{R}^{p \times q}$ denote the between-set covariance matrix (note that $S_{yx} = S_{xy}^T$). The overall $(p+q) \times (p+q)$ covariance matrix, S , contains all the information on associations between pairs of features:

$$S = \begin{pmatrix} cov(x) & cov(x,y) \\ cov(y,x) & cov(y) \end{pmatrix} = \begin{pmatrix} S_{xx} & S_{xy} \\ S_{yx} & S_{yy} \end{pmatrix} \quad (1)$$

B. Incorporating class structure in Multivariate correlation analysis

The feature fusion method described in the previous section has two disputable issues. The first issue is encountered in case of a small sample size problem. In many real world applications, the number of samples is usually less than the number of features ($n < p$ or $n < q$). This makes the covariance matrices singular and non-invertible. Therefore, we will face a major problem in inverting the S_{xx} and S_{yy} matrices used in Eq. (2). A solution to overcome this issue is to reduce the dimensionality of the feature vectors before applying CCA. Therefore, a two stage PCA + CCA approach can be considered [12].

The second issue in CCA-based approaches is their negligence of the class structure among samples. CCA decorrelates the features, but in pattern recognition problems, we are also interested in separating the classes. The dimensionality reduction approaches based on Linear Discriminant Analysis (LDA) [13] consider this matter by finding projections that best separate the classes. However, a two stage LDA + CCA will not be an effective solution due to the fact that the transformation applied by the second stage, i.e., CCA, will not preserve the properties achieved by the first stage, i.e., LDA.

Therefore, we need transformations to not only maximize the pair-wise correlations across the two feature sets, but also to simultaneously separate the classes within each feature set. In this section, we present a solution to achieve this goal. Our proposed approach, called Discriminant Correlation Analysis (DCA), is described next section.

C. Feature-Level Fusion Using Discriminant Correlation Analysis

Let's assume that the samples in the data matrix are collected from c separate classes. Accordingly, the n columns of the data matrix are divided into c separate groups, n_i columns comprising the i^{th} class ($n = \sum_{i=1}^c n_i$). Let $x_{ij} \in X$ denote the feature vector corresponding to the j^{th} sample in the i^{th} class. \bar{x}_i and \bar{x} denote the means of the x_{ij} vectors in the i^{th} class and overall, respectively, $\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}$ and $\bar{x} = \frac{1}{n} \sum_{j=1}^n x_{ij}$. The between-class scatter matrix is defined as

$$S_{bx(p \times p)} = \sum_{i=1}^c n_i (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T = \Phi_{bx} \Phi_{bx}^T \quad (2)$$

where

$$\Phi_{bx(p \times c)} = [\sqrt{n_1} (\bar{x}_1 - \bar{x}), \sqrt{n_2} (\bar{x}_2 - \bar{x}), \dots, \sqrt{n_c} (\bar{x}_c - \bar{x})]$$

If the number of features is higher than the number of classes ($p \gg c$) it is computationally easier to calculate the covariance matrix as $(\Phi_{bx}^T \Phi_{bx})_{c \times c}$ rather than $(\Phi_{bx} \Phi_{bx}^T)_{p \times p}$. As presented in [14], the most significant eigenvectors of $\Phi_{bx} \Phi_{bx}^T$ can be efficiently obtained by mapping the eigenvectors of $\Phi_{bx}^T \Phi_{bx}$. Therefore, we only need to find the eigenvectors of the $c \times c$ covariance matrix $\Phi_{bx}^T \Phi_{bx}$.

If the classes were well-separated, $\Phi_{bx}^T \Phi_{bx}$ would be a diagonal matrix. Since $\Phi_{bx}^T \Phi_{bx}$ is symmetric positive semi-definite, we can find transformations that diagonalize it:

$$P^T (\Phi_{bx}^T \Phi_{bx}) P = \hat{\Lambda},$$

Where P is the matrix of orthogonal eigenvectors and $\hat{\Lambda}$ is the diagonal matrix of real and non-negative eigenvalues sorted in decreasing order.

Let $Q_{(c \times r)}$ consist of the first r eigenvectors, which correspond to the r largest non-zero eigenvalues, from matrix P .

We have:

$$Q^T (\Phi_{bx}^T \Phi_{bx}) Q = \Lambda_{(r \times r)}$$

The r most significant eigenvectors of S_{bx} can be obtained with the mapping: $Q \rightarrow \Phi_{bx} Q$ [14]:

$$(\Phi_{bx} Q)^T S_{bx} (\Phi_{bx} Q) = \Lambda_{(r \times r)} \quad (3)$$

$W_{bx} = \Phi_{bx} Q \Lambda^{-1/2}$ the transformation that unitizes S_{bx} and reduces the dimensionality of the data matrix, X , from p to r . That is:

$$W_{bx}^T S_{bx} W_{bx} = I \quad (4)$$

$$X'_{(r \times n)} = W_{bx}^T X_{(p \times n)} \quad (5)$$

X' is the projection of X in a space, where the between-class scatter matrix is I and the classes are separated. Note that there are at most $c - 1$ nonzero generalized eigenvalues; therefore, an upper bound on r is $c - 1$ [15]. Other upper

bounds for r are the ranks of the data matrices, i.e., $r \leq \min\{c - 1, \text{rank}(X), \text{rank}(Y)\}$.

Similar to the above approach we solve for the second feature set, Y , and find a transformation matrix W_{by} , which unitizes the between-class scatter matrix for the second modality, S_{by} and reduces the dimensionality of Y from q to r :

$$W_{by}^T S_{by} W_{by} = I \quad (6)$$

$$Y'_{(r \times n)} = W_{by}^T Y_{(q \times n)} \quad (7)$$

The updated Φ_{bx} and Φ_{by} are non-square $r \times c$ orthonormal matrices. Although $S'_{bx} = S'_{by} = I$, the matrices $\Phi_{bx}^T \Phi_{bx}$ and $\Phi_{by}^T \Phi_{by}$ are strict diagonally dominant matrices, where the diagonal elements are close to one and the non-diagonal elements are close to zero. This makes the centroids of the classes have minimal correlation with each other, and thus, the classes are separated.

Now that we have transformed X and Y to X' and Y' , where the between-class scatter matrices are unitized, we need to make the features in one set have nonzero correlation only with their corresponding features in the other set. To achieve this, we need to diagonalize the between-set covariance matrix of the transformed feature sets, $S'_{xy} = X' Y'^T$. We use singular value decomposition (SVD) to diagonalize S_{xy} :

$$S'_{xy(r \times r)} = U \Sigma V^T \Rightarrow U^T S'_{xy} V = \Sigma \quad (8)$$

Note that X and Y are of rank r and $S'_{xy(r \times r)}$ is non-degenerate. Therefore, Σ is a diagonal matrix whose diagonal elements are non-zero. Let $W_{cx} = U \Sigma^{-1/2}$ and $W_{cy} = V \Sigma^{-1/2}$ we have:

$$(U \Sigma^{-1/2})^T S'_{xy} V \Sigma^{-1/2} = I$$

which unitizes the between-set covariance matrix, S_{xy} . Now, we transform the feature sets as follows:

$$X^* = W_{cx}^T X' = W_{cx}^T W_{bx}^T X = W_x X$$

$$Y^* = W_{cy}^T Y' = W_{cy}^T W_{by}^T Y = W_y Y$$

where $W_x = W_{cx}^T W_{bx}^T$ and $W_y = W_{cy}^T W_{by}^T$ are the final transformation matrices for X and Y , respectively.

It can be easily shown that the between-class scatter matrices of the transformed feature sets are still diagonal; hence, the classes are separated. The between-class scatter matrix for X^* is calculated as:

$$S_{bx}^* = W_{cx}^T W_{bx}^T S_{bx} W_{bx} W_{cx}$$

From Eq. (3), $W_{bx}^T S_{bx} W_{bx} = I$ and since U is an orthogonal matrix, we have:

$$S_{bx}^* = (U \Sigma^{-\frac{1}{2}})^T (U \Sigma^{-\frac{1}{2}}) = \Sigma^{-1}$$

Similarly, we can show that $S_{by}^* = \Sigma^{-1}$, which is diagonal. Similar to the CCA method, feature level fusion can be performed either by concatenation or summation of the transformed feature vectors. In our experiments, we use the summation method. In case of having more than two sets of features, we follow a cascade approach and apply DCA on two sets of features at a time.

D. Multiset Discriminant Correlation Analysis

Multiset Discriminant Correlation Analysis (MDCA) generalizes DCA to be applicable to more than two sets of features. Here, we assume that we have m sets of features, $X_i \in \mathbb{R}^{p_i \times n}$, $i = 1, 2, \dots, m$, which are sorted by their rank, that is $rank(X_1) \geq rank(X_2) \geq \dots \geq rank(X_m)$. MDCA applies DCA on two sets of features at a time. Based on the approach presented in the previous section, the maximum length of the fused feature vector is $\min\{c - 1, rank(X_i), rank(X_j)\}$.

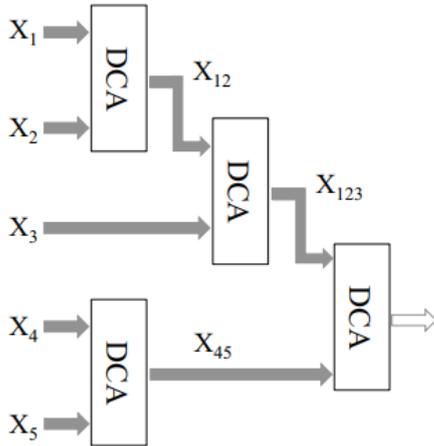


Fig. 2. Multiset discriminant correlation analysis techniques for five sample sets with $rank(X_1) > rank(X_2) > rank(X_3) > rank(X_4) = rank(X_5)$.

In order to maintain the maximum possible length of the fused feature vector, in each step, the two feature sets with the highest ranks will be fused together. For example, in the first step, X_1 and X_2 , which have the highest ranks, will be fused. The result of the fusion of X_1 and X_2 will be fused with the next highest rank feature set, i.e., X_3 , and so on. If there exists feature sets with equal ranks, they can be fused together at anytime. We choose the length of the fused feature vector, r , to be equal to $\min\{c - 1, rank(X_i), rank(X_j)\}$.

Fig. 2 shows an example framework of MDCA for five feature sets with $rank(X_1) > rank(X_2) > rank(X_3) > rank(X_4) = rank(X_5)$. In the first step of MDCA, we fuse X_1 and X_2 , which have the highest ranks. X_4 and X_5 , which have equal ranks, will be also fused together. The length of the X_{12} is expected to be greater than the length of the X_{45} . Therefore, in the next step, X_3 is fused with X_{12} . In this way, we keep the maximum possible length for the fused feature vector in every step. The expected, possibly shorter, feature vector length can be determined in the final step, $r \leq \min\{c - 1, rank(X_{123}), rank(X_{45})\}$.

4. EXPERIMENTS AND ANALYSIS

In this paper, we present several sets of experiments to demonstrate the performance of our proposed feature level fusion technique. We devise experiments for combining different features extracted from a single modality as well as combining feature vectors extracted from different biometric modalities. Three different features are extracted from these images. These features include Gabor wavelet features [16], Histogram of Oriented Gradients (HOG) [17], and Speeded-Up Robust Features (SURF) [18]. We employ forty Gabor filters in five scales and eight orientations. Since the adjacent pixels in an image are usually correlated, the information redundancy can be reduced by down sampling the feature

images that result from Gabor filters [16], [19]. In our experiments, the feature images are down sampled by a factor of five. HOG features, on the other hand, are extracted in 5×5 cells for nine orientations. We use the UOCTTI variant for the HOG presented in [20]. UOCTTI variant computes both directed and undirected gradients as well as a four dimensional texture energy feature, but projects the result down to 31

Table 1. Maximum Rank-1 Recognition Rates Using Individual and Fused Feature Vectors On Palmprint and Handvein (H: Hog, G: Gabor).

Method	H	G	HG
CCA/MCCA	85.52	82.31	95.78
DCA/MDCA	88.12	84.66	97.34

dimensions (27 dimensions corresponding to different orientation channels, 9 contrast insensitive and 18 contrast sensitive, and 4 dimensions capturing the overall gradient energy in square blocks of four adjacent cells). Finally, we extract SURF features from 68 keypoints in every image. These points are the facial landmarks detected by fitting an Active Appearance Model (AAM) to the face images. A 64-dimensional feature vector is extracted from each point and the final feature vector is constructed by concatenating the feature vectors of all keypoints. A simple minimum distance classifier is used for classification, in which one minus the sample linear correlation between observations is used as the distance.

5. CONCLUSION

In this paper, we presented a feature fusion technique based on correlation analysis of the feature sets. Our proposed method, called Discriminant Correlation Analysis, contemplates the class associations of the samples in its analysis. It aims to find transformations that maximize the pair-wise correlations across the two feature sets and at the same time separate the classes within each set. These characteristics make DCA an effective feature fusion tool for pattern recognition applications. Experimental results demonstrated the efficacy of our proposed approach in fusion of multimodal feature sets or different feature sets extracted from a single modality. In addition, the DCA method is computationally efficient and can be employed in real-time applications.

6. REFERENCES

- [1] S Veluchamy, L R Karlmarx,: System for Multimodal Biometric recognition based on Finger Knuckle and Finger Vein using Feature Level Fusion and K-support vector Machine Classifier, IEEE Biometrics, Vol 6 (2017).
- [2] Gopal, Smriti Srivastava, Saurabh Bhardwaj, Sandeep Bhargava,: Fusion of palm-phalanges print with palmprint and dorsal hand vein, Elsevier, Applied Soft Computing 47, 2016, 12-20.
- [3] M. Haghighat, M. Abdel-Mottaleb, W. Alhalabi,: Discriminant Correlation Analysis for Feature Level Fusion with Application to Multimodal Biometrics, IEEE, Speech and Signal Processing (ICASSP), 2016.
- [4] Xiaofeng, Yang, Dongmei,: Feature -Level Fusion of palmprint and palm vein Base on Canonical Correlation Analysis, IEEE, Signal Processing (ICSP), 2016.

- [5] Bharathi, S, R Sudhakar Valentina E Balas,: Hand Vein-based Multimodal Biometric Recognition, *ActaPolytechnicaHungarica*, (2015) 213-229.
- [6] D Vaidya, S Pawar, Madhuri A Joshi, A M Sapkal, S kar,: Feature level Fusion of Palm Print and Palm Vein for person Authentication Based on Entropy Technique, *International Journal of electronics and Communication Technology*,(2014) Vol.5.
- [7] Li Peidan, Miao Zhenjiang, Wang Zhiqun,: Fusion of Palm Print and Palm Vein Images for Person Recognition, *IEEE, Signal Processing(ICSP)*, 2014.
- [8] D P Gaikwad, S P Narote,: Multi-Modal Biometric System using Palm Print and Palm Vein Features, *Annual IEEE Indian Conference(INDICON)*, (2013).
- [9] M Heenaye, M Khan,: A Multimodal HandVein Biometric based on Score Level Fusion, *Internaional Symposium on Robotics and Intelligent Sensors(IRIS)*, Elsevier, (2012).
- [10] Ashok Rao, M Imran, R Ragavendra, Hemantha Kumar G,: Multibiometrics: Analysis and Robustness of Hand Vein and Palm Print Combination Used for Person Verification, *International Journal for Emerging Trends in Engineering and Technology (IJETET)*, (2011).
- [11] N.M. Correa, T. Adali, Y. Li, and V.D. Calhoun, "Canonical correlation analysis for data fusion and group inferences," *IEEE Signal Processing Magazine*, vol. 27, no. 4, pp. 39–50, 2010.
- [12] P.N. Belhumeur, J.P. Hespanha, and D. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711–720, 1997.
- [13] H. Li, K.A. Toh, and L. Li, *Advanced topics in biometrics*, World Scientific, 2012.
- [14] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.
- [15] R.O. Duda and P.E. Hart, *Pattern classification and scene analysis*, Wiley New York, 1973.
- [16] C. Liu and H. Wechsler, "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition," *IEEE Transactions on Image processing*, vol. 11, no. 4, pp. 467–476, 2002.
- [17] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, 2005, pp. 886–893.
- [18] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-Up Robust Features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.
- [19] M. Haghghat, S. Zonouz, and M. Abdel-Mottaleb, "Identification using encrypted biometrics," in *Computer Analysis of Images and Patterns (CAIP)*. Springer, 2013, pp. 440–448.
- [20] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1627–1645, 2010.