Pattern Recognition, August 2018

Cross-Spectral Iris Recognition using CNN and Supervised Discrete Hashing

Kuo Wang, Ajay Kumar
Department of Computing, The Hong Kong Polytechnic University, Hung Hom, Hong Kong

Abstract: Completely automated iris recognition has emerged as an integral part of e-business and e-governance infrastructure which has acquired billions of iris images under near-infrared illumination to establish the identity of individuals. A range of e-business and surveillance applications can provide iris images that are acquired under visible illumination. Therefore, development of accurate cross-spectral iris matching capabilities is highly desirable. This paper investigates cross-spectral iris recognition using a range of deep learning architectures. Our experimental results on two publicly available cross-spectral iris databases, from 209 and 120 different subjects respectively, indicate outperforming results and validate our approach for the cross-spectral iris matching. Our observations indicate that the self-learned features generated from the convolution neural networks (CNN) are generally sparse and offer great potential for template compression. Therefore, this paper also introduces the iris recognition with supervised discrete hashing that can not only achieve more accurate performance but also offer a significant reduction in the size of iris templates. Most accurate cross-spectral matching performance is achieved by incorporating supervised discrete hashing on the features learned from the trained CNN with softmax cross-entropy loss. The proposed approach not only achieves outperforming results over other considered CNN architecture but also offers significantly reduced template size as compared with the other iris recognition methods available in the literature.

Keywords: Cross-spectral iris recognition, deep learning, convolutional neural network, Hashing.

1. Introduction

Wide spread use of biometrics identification has significantly increased the efficiency in e-governance and e-businesses for the benefit of citizens. Iris recognition has emerged as the most accurate, low-cost, convenient method of personal identification

that also enables better hygiene due to the contactless imaging [1]-[2],[40],[43]. Iris images from billions of citizens have been acquired worldwide and deployed for assisting in establishing the identity of individuals. Emerging dual imaging sensors can simultaneously provide near-infrared and visible illumination iris images at the same time and offer iris recognition capabilities that may not be possible with the use of only near-infrared images from popular iris sensors. Effective usage of emerging iris recognition capabilities, using the visible illumination images [10], and those acquired using smartphones [11], also requires accurate matching capabilities with corresponding images stored in national ID databases that are typically acquired under near-infrared illumination. Iris texture is known to appear differently under different spectral illumination. Earlier research work in this area has therefore underlined that the cross-spectral iris matching is far more challenging as compared to the same spectrum, i.e., near-infrared to near-infrared or visible to visible iris image matching.

Deep neural network architectures, like convolutional neural network (CNN), have shown remarkable capabilities in automatically learning the specific features for accurate classification of visual patterns that can generate outperforming results than those from the hand-crafted features. There are several examples in the literature for the successful use of CNN in hand-written character recognition [5],[42], large-scale image classification [19] and face recognition [16],[47] problems. Many of the recent advances in iris recognition have also investigated the capabilities of CNN for iris recognition [8], periocular recognition [14], and iris segmentation [7] and the detec-

tion of fake or spoof iris images [17]. The success of earlier work in the ocular recognition [14] has indicated capabilities of CNN in learning anatomical characteristics, like the gender or pose, associated with periocular images for more accurate user identification. Therefore, there is strong motivation to explore the capabilities of deep learning architectures in addressing more challenging biometrics identification problems, like cross-spectral iris matching investigated in this paper.

1.1 Related Work

Conventional iris recognition is generally considered as a matured area with successful deployments of such systems in national ID cards, border-crossings and e-governance. However, such deployments have also introduced new opportunities and challenges for the applicability of iris recognition technologies in new domains. In this section, we provide a brief overview of such related work on the cross-spectral iris recognition.

The cross-spectral iris matching requires the iris images acquired under the visible illumination (VIS) to be matched with the iris images that are conventionally acquired under near-infrared illumination (NIR). This problem is widely considered to be more challenging than the cross-sensor iris matching and has attracted the attention of several researchers. Early efforts [3] systematically investigated the information contents available from iris images acquired under different spectrum and proposed to combine such complementary information to improve the matching accuracy over those available from the conventional iris recognition under NIR. Several studies in the literature have shown that the performance of matching iris images acquired from two different sensors, or different resolutions [41], also degrades

seriously. Reference [28] lists such factors for the degradation in iris recognition performance. Reference [25] later introduced a framework to synthesize or predict NIR iris images from the corresponding VIS iris images for accurately matching them with NIR images. The experimental results were presented on the proprietary database from 68 subjects with only 232 instances, which can be considered as a limitation of this work. More recent or successful algorithm in [3], [30] is also based on the prediction of NIR images from VIS iris images, and then matching the predicted NIR instances with those NIR instances in the gallery database. Such a strategy incorporated Markov random fields (MRF) framework to generate more accurate alternative for the cross-spectral iris-matching problem. More recently Sequeira *et al*. [4] released a new cross-eyed and cross-spectral iris dataset and made the most promising effort to advance research on the challenging cross-spectral iris matching problem with equal error rate (EER) less than four percent. However, the work detailed in [4] should be considered preliminary, primary due to lack of comparison with most popular *IrisCode* approach [24] or usage of less challenging matching protocols or the lack of details on the selection of images or the availability of codes which is the major limitation for the reproducibility. Table 1 summarizes related work on the cross-spectral iris matching. The details on the number of impostor scores and template size are not explicitly stated in [25] and therefore these are estimated from numerical results provided in the original paper.

Table 1: Summary of Earlier Work on Cross-spectral Iris Recognition.

Ref.	Method	Database	Public	Iris Comparison	Template Size	Features	Matching Accuracy
[25]	A predictive NIR iris image is used from the color image	WVU Multi-spectral iris database	No	Genuine=280 *Impostor =20745	*128×720 ×8	Hand-crafted	95.2% (FAR=0.001)
[24]	*IrisCode* using 1D Log-Gabor filter	1) PolyU bi-spectral iris database 2) Cross-eyed-cross-spectral iris recognition database	Yes	Genuine=2800 Impostor =1953000 Genuine=2160 Impostors =516240	64×512×2	Hand-crafted	1) 52.6% (FAR=0.1) 2) 70.3% (FAR=0.1)
[3]	NIR to VIS texture synthesis using MRF model	1) PolyU bi-spectral iris database 2) Cross-eyed-cross-spectral iris recognition database	Yes	Genuine=2800 Impostor =1953000 Genuine=2160 Impostors =516240	64×512×2	Hand-crafted	1) 64.91% (FAR=0.1) 2) 78.13% (FAR=0.1)
Ours	CNN with softmax cross-entropy loss for feature extraction and SDH for compression and classification	1) PolyU bi-spectral iris database 2) Cross-eyed-cross-spectral iris recognition database	Yes	Genuine=2800 Impostor =1953000 Genuine=2160 Impostors =516240	1000×1	Self-learned	1) 90.71% (FAR=0.01) 2) 87.18% (FAR=0.01)

* estimated from details in [25].

Several other researchers have introduced promising algorithms to address performance degradation in cross-sensor iris matching. Xiao *et al*. [27] introduced an optimized model of coupled feature selection when comparing two iris templates acquired from different iris sensors. Pillai *et al*. [26] later investigated the adaptation of cross-sensor iris data which was represented by the binarized iris codes. However, more effective and robust learning for the corresponding features can be achieved with the real-valued feature representation, instead of binarized feature representation. Therefore reference [3] more recently proposed to incorporate real-valued features for the cross-spectral iris recognition. This approach incorporates the iris feature learning using the Naïve-Bayes Nearest-Neighbor (NBNN) and presents more accurate or promising (also reproducible) results to date and can serve as a

baseline for further work as also in this paper.

1.2 OUR WORK

Earlier work on cross-spectral iris recognition has made significant advances to improve the matching accuracy. The summary of related work in Table 1 indicates that the earlier efforts have been largely focused on extracting hand-crafted features from the corresponding spectrum during the learning phase. Although the matching performance in earlier work illustrates great improvement, more effort is required to enhance the matching accuracy for the real applications. The key contribution of our work can be summarized as in the following:

1) Earlier work on cross-spectral iris recognition has employed a range of hand-crafted features that are learned from the registration or training data. The key to success for accurately matching the cross-spectral iris images lies in automatically learning the corresponding features, instead of employing hand-crafted features which can only deliver limited improvement in the performance. This work therefore investigates the potential from the deep learning capabilities to for more accurate cross-spectral iris recognition. We present experimental results using a range of deep learning architectures, *e.g.,* CNN, Siamese networks, Triplet networks, VGG and deep residual networks (ResNet), to comparatively ascertain performance improvement over the state of the art or more recent approach in the literature. Our experimental results presented in section 3 achieve outperforming results on two publicly available cross-spectral iris image databases and validate our approach for more accurate cross-spectral iris matching.

2) Smaller template size and higher template matching speed is preferred in a variety of iris recognition applications. Our experiments to improve the cross-spectral iris matching accuracy indicate that the feature vectors recovered from the CNNs are generally sparse and therefore there is significant potential to achieve compact iris template representations. This paper therefore proposes to incorporate supervised discrete hashing (SDH) scheme to significantly reduce the iris template size and achieve improved template matching speed using the Hamming distance for the CNN based features. Our experimental results presented in section 3, on two publicly available databases indicate significantly reduced iris template size (1000 bits) while offering a noticeable improvement in the cross-spectral matching accuracy and validate our proposal. Our 1000 bits iris template representation scheme generates compact template size (125 bytes) while significantly improving the cross-spectral iris matching accuracy (e.g., 78% improvement in EER for PolyU bi-spectral iris dataset [3] from 140 subject and 86% improvement in EER for Cross-eyed-cross-spectral iris recognition database [4] from 120 subjects).

The rest of this paper is organized as follows. Section 2 details our framework and various CNN architectures that were investigated to evaluate their effectiveness in learning features. The SDH is also introduced in this section while the experimental protocols and results are presented in section 3. The discussion section that includes additional comparisons, and extended experimental results, appears in section 4. The key conclusions from this paper are summarized in section 5.

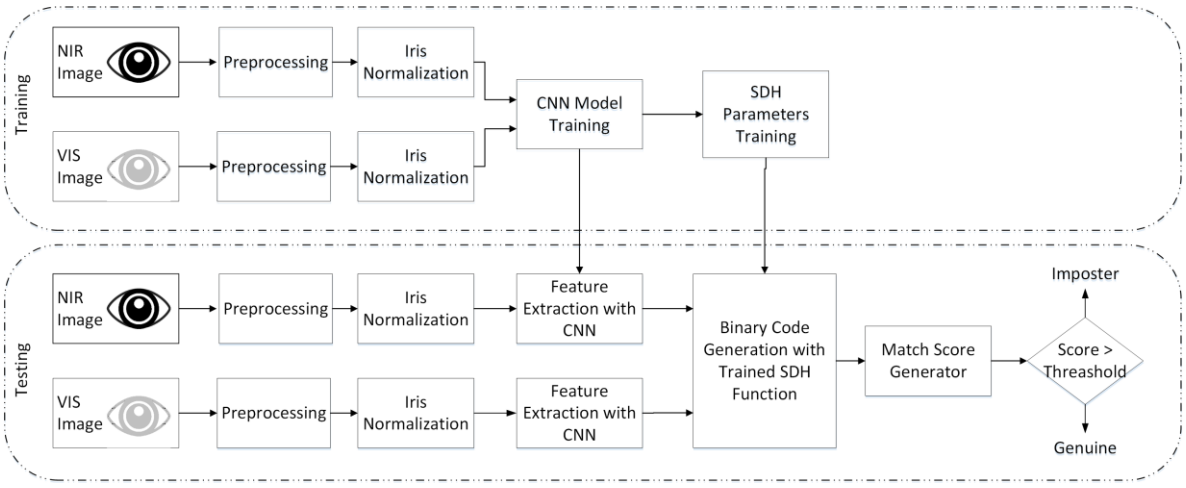


Figure 1: Block diagram of cross-spectral iris recognition framework using deep neural network and supervised discrete hashing.

2. Methodology

The framework for cross-spectral iris recognition investigated in this work is shown in Figure 1. This framework, like earlier work [3], [30], consists of offline training phase where the parameters are automatically learned. Each of the iris images from both spectra is firstly subjected to the segmentation and image enhance steps. The segmented images are used for the CNN based feature learning. The parameters for SDH are also learned during the training phase and employed to generate binarized features for the Hamming distance based matching during the test phase. These match scores are used to establish the identity of the unknown input iris image pairs.

We consider three CNN architectures to ascertain the learning for the cross-spectral iris features, *i.e.*, CNN with *softmax* cross-entropy loss, Siamese network and Triplet network. Normalized image samples are used to extract fixed length feature vectors which are used for the classification. Two approaches for the classification of feature vectors are investigated; the first one is based on the joint-Bayesian inference [18]

while the other one incorporates supervised discrete hashing (SDH) [9] which generates match scores using Hamming distance. We now briefly introduce the CNN architectures that are considered in our work to investigate cross-spectral iris matching.

2.1 Convolutional Neural Network

2.1.1 CNN with Softmax Cross-entropy Loss

CNN with *softmax* cross-entropy loss offers superior generalization capability and compact feature representations. The CNN employed in our framework adopts the architecture which is similar to AlexNet [19] and is illustrated in Figure 2. The network is composed of three convolutional layers, three pooling layers and two fully-connected (FC) layers. After each pooling layer and the first FC layer, there is a non-linear activation function, and we use Rectified Linear Unit (ReLU) function in this work. The last FC layer is used for label prediction and calculation of softmax loss. For every convolutional layer, it will calculate the i-th channel output y^i by using (1). In this equation, x^j is the j-th channel input from the previous layer, w^{ij} is the convolutional kernel between the x^j and y^i while b^{ij} is the neuron bias between input channel and the output channel.

$$y^i = \sum_j (b^{ij} + w^{ij} * x^j) \tag{1}$$

In our experiments, pooling layer extracts maximum value in the kernel. It aims to downsample the input size for the further processing so that features can be learned from a different scale. The rectified linear unit (ReLU) activation function activates all the output with non-negative values and can be defined as follows:

$$y^i = \max(y^i, 0) \tag{2}$$

The FC layer will connect all the nodes in the current layer with one node to generate the output vector as in the following.

$$y^i = b_i + \sum_j w^{ij} * x^j \tag{3}$$

The network is initialized randomly and therefore there are variations in how accurately the neurons learn. The *softmax* cross entropy loss function is employed to compute the normalized error between the ground truth label and our prediction label as follows:

$$E = \frac{-1}{N} \log(p_n, l_n) \tag{4}$$

where N is the number of classes in training samples, p_n is the predicted label and l_n is the ground truth label. The final output is a $1 * N$ vector with positive elements. Each element value represents the probability prediction of the class label for the input image. The training aims to minimize the loss by backpropagation so that the probability of predicting the ground-truth class can approach maximum.

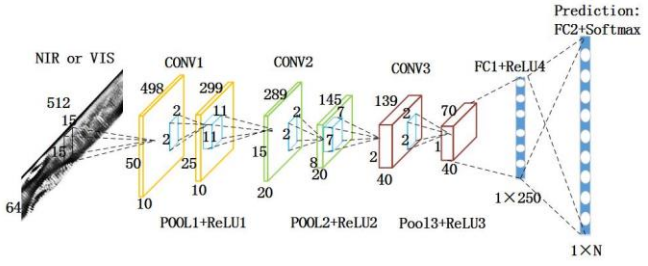


Figure 2: CNN with *softmax* cross-entropy loss**.**

2.1.2 SIAMESE NETWORK

The Siamese network [20] is essentially a twin branch structure that learns similarity between two comparable inputs. It contains two identical subnetworks whose all

the parameters are simultaneously updated. The motivation for employing this architecture in our experiment is to ascertain inner correspondences between the NIR image pixels and VIS image pixels. One NIR image and one VIS image form an input pair for Siamese network. Each input pair data has a binary label l indicating whether they are from same iris/eye. The architecture of Siamese network is illustrated in Figure 3 and each branch of this network has the same configuration as the CNN with *softmax* cross-entropy loss. In this network, the comparative loss from the last layer of twin branches is computed. This loss function can be written as follows:

$$E = \frac{1}{2N}\sum_{n=1}^{N}((l)d^2 + (1-l)\max(\alpha - d, 0)^2) \quad (5)$$

where N is the batch size set in forward propagation, l is the binary label indicating genuine pair and imposter pair, and d is the Euclidean distance between the two feature vectors generated from fully-connected layers while α is the margin set during the training process. This loss function aims to minimize the distance between genuine pairs and enlarge the distance between the imposter pairs so that the overlap can be reduced and the matching performance can be improved.

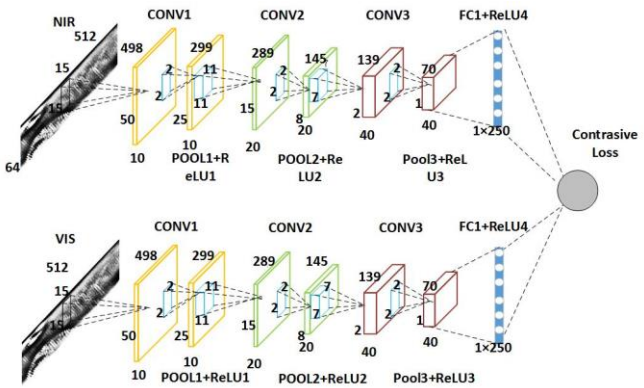


Figure 3: Siamese network for cross-spectral iris matching

2.1.3 TRIPLET NETWORK

Triplet network [21] is comprised of three branches as shown in Figure 4. The three-channel inputs include one anchor image, one positive image and one negative image. In this experiment, we select one NIR images as the anchor sample A, one VIS image from the same class as positive sample P and one VIS image from different class as negative sample N. The network architecture attempts to solve the 2-class classification problem, where the objective is to correctly classify which of P and N is of the same class as A. The configuration of every branch is same as for the CNN with *softmax* cross-entropy loss in Figure 2. The triplet loss function in this network can be defined as following:

$$E = \sum_{i=1}^{N}(||A_i - P_i||^2 - ||A_i - N_i||^2 + \alpha) \tag{6}$$

where N is the batch size in the training, A_i is the anchor image whereas P_i is positive one and N_i is the negative one while α is the minimum margin fixed during the training stage. This approach attempts to simultaneously increase the distance between the imposter pairs and reduce the distance between the genuine pairs.

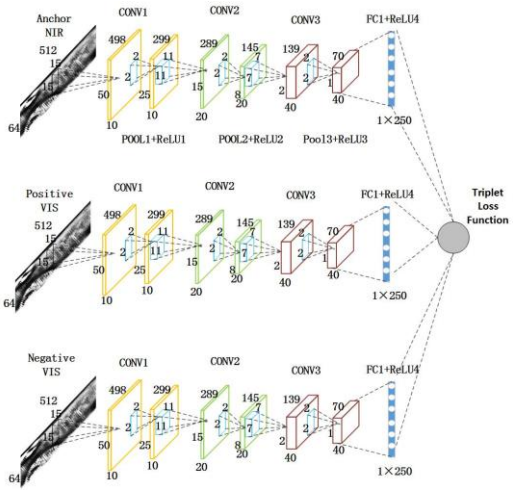


Figure 4: Triplet network for cross-spectral iris matching.

2.2 JOINT-BAYESIAN FORMULATION

The Joint-Bayesian formulation approach for the classification of features generated from the trained CNN has shown to offer more accurate performance for the classification of ocular images [14] and was therefore also attempted in this work. The fully connected layer outputs before prediction layer are utilized as the feature vector. Each of such templates from the CNN is represented by this real-value vector.

In theory of Joint-Bayesian, every feature vector x can be formulated as follows:

$$x = \mu + \varepsilon \tag{7}$$

where μ represents class identity and ε represents intra-class variation. Under the assumption of Gaussian distributions, *i.e.* $N(0, S_\mu)$, $N(0, S_\varepsilon)$, both variables are with zero mean. The covariance matrix of two observations $\{x_1, x_2\}$ can be written as follows:

$$cov(x_1, x_2) = cov(\mu_1, \mu_2) + cov(\varepsilon_1, \varepsilon_2) \tag{8}$$

Let H_I denote the intra-class hypothesis indicating that two observations are from the same class, and H_E the extra-class hypothesis. Under H_I, μ1 and μ2 are the same due to the same class identity, and ε_1 and ε_2 are independent. Therefore, the covariance matrix of the distribution $P(x_1, x_2|H_I)$ can be written as:

$$\sum_{Intra} = \begin{bmatrix} S_\mu + S_\varepsilon & S_\mu \\ S_\mu & S_\mu + S_\varepsilon \end{bmatrix} \tag{9}$$

On the other hand, under H_E, μ1 and μ2 are also independent, therefore the covariance matrix of $P(x_1, x_2|H_E)$ becomes:

$$\sum_{Inter} = \begin{bmatrix} S_\mu + S_\varepsilon & 0 \\ 0 & S_\mu + S_\varepsilon \end{bmatrix} \tag{10}$$

The log likelihood ratio $s(x1, x2)$ can be obtained in a closed form after simple algebraic operations:

$$s(x1, x2) = \frac{P(x_1, x_2|H_I)}{P(x_1, x_2|H_E)} = x_1^T A x_1 + x_2^T A x_2 - 2x_1^T G x_2 \tag{11}$$

where

$$A = \left(S_\mu + S_\varepsilon\right) - (F + G) \tag{12}$$

$$\begin{bmatrix} F+G & G \\ G & F+G \end{bmatrix} = \begin{bmatrix} S_\mu + S_\varepsilon & S_\mu \\ S_\mu & S_\mu + S_\varepsilon \end{bmatrix}^{-1} \tag{13}$$

The covariation matrix S_μ and S_ε can be estimated from expectation-maximization (EM) algorithm as described in [18]. In this work, the log likelihood ratio $s(x1, x2)$ is considered as the match score between two respective features $\{x_1, x_2\}$.

2.3 SUPERVISED DISCRETE HASHING

Shen *et al*. [9] have detailed a hashing algorithm to generate binary features for high dimensional data. One key advantage of such hashing is that it can reduce the storage requirements for the iris templates. The availability of such binarized templates can improve the matching speed as the Hamming distance operations required to compute the match score are significantly faster.

The match score generation with this hashing scheme incorporates the features vectors same as described in section 2.2. The training stage for the SDH aims to learn L bits binary code $B = \{b_i\}_{i=1}^n$ corresponding to the every input image using n training feature vectors $X = \{x_i\}_{i=1}^n$ and the ground truth label matrix $Y = \{y_i\}_{i=1}^n$. The following algorithm 1 summarizes the process for computing B.

Algorithm 1: Supervised Discrete Hashing

Input: training data $\{x_i, y_i\}_{i=1}^n$; code length L; Number of anchor points m; Maximum iteration number t; Penalty parameters λand v.

Output: Binary codes $\{b_i\}_{i=1}^n \in \{-1, 1\}^{L\times n}$; hash function $H(x) = sgn(F((x))$.

Step 1: Randomly select m samples $\{a_j\}_{j=1}^m$ from the training data and get the mapped training data $R(x)$ via the RBF kernel function.

Step 2: Initialize b_i as a {-1, 1}^{L} vector randomly, $\forall i$.

Step 3: Loop until converging or reach maximum iterations;

-projection step: Compute P using (18) to form F(x).

-mapping step: Calculate W using equation (21).

-Bits calculation step: For the l2 loss, iteratively learn $\{b_i\}_{i=1}^n$ bit by bit using the discrete cyclic coordinate descent (DCC) method with equation (26).

The expected binary code b should be ideal for classification, and the multi-class classification can be formulated as:

$$y = W^T b=[w_1^T b, \dots, w_C^T b] \tag{14}$$

where $w_k, k = 1, \dots, C$ is the class vector for class k if we have C classes, and y is the label vector whose maximum value indicates the predicted class.

To achieve high classification accuracy, we need to optimize the following problem:

$$\begin{aligned} &min_{B,W,F} \sum_{i=1}^n E(y_i, W^T b_i) + \lambda||W||^2 \\ &\text{s.t.} \quad b_i = sgn\big(F(x_i)\big), i = 1, \dots, n \end{aligned} \tag{15}$$

where $E(.)$ is the loss function. $F(.)$ is the hash function we try to learn. Here $sgn(.)$ is the sign function which produces +1 for the positive value and -1 for the negative value. λ is the regularization parameter. In order to achieve binary codes of better quality, we keep the binary constraints of b_i in the optimization. The problem described in (15) adds one more term modelling the fitting error of binary code b_i by the continuous embedding hashing function $F(x_i)$ and penalty term v as follows:

$$min_{B,W,F} \sum_{i=1}^{n} E(y_i, W^T b_i) + \lambda||W||^2 + v \sum_{i=1}^{n} ||b_i - F(x_i)||^2$$

$$\text{s.t.} \quad b_i \in \{-1, 1\}^L \tag{16}$$

However, the problem described in (16) is highly non-convex and difficult to solve. But it is tractable when we iteratively solve each variable one by one. Firstly, we adopt the following nonlinear form for the hash function $F(x)$:

$$F(x) = P^T R(x) \tag{17}$$

$R(.)$ outputs an m-dimensional column vector obtained by the RBF kernel mapping as described in [22]-[23]. The matrix P projects the mapped data $R(x)$ into the low dimensional space.

Projection-step: If the binary code b_i in problem (16) is fixed, the projection matrix P can be computed by regression:

$$P = (R(X)R(X)^T)^{-1} R(X)B^T \tag{18}$$

This step is independent of the loss function $E(.)$.

In this experiment, we select l_2 loss as $E(.)$ in our classification model. The problem (16) can be rewritten as:

$$min_{B,W,F} \sum_{i=1}^{n} ||y_i - W^T b_i||^2 + \lambda||W||^2 + v \sum_{i=1}^{n} ||b_i - F(x_i)||^2$$

$$\text{s.t.} \quad b_i \in \{-1, 1\}^L \tag{19}$$

And we can rewrite (19) in the form:

$$min_{B,W,F} ||Y - W^T B||^2 + \lambda||W||^2 + v||B - F(X)||^2$$

$$\text{s.t.} \quad B \in \{-1, 1\}^{L*n} \tag{20}$$

Mapping-step: If B is fixed for the optimization of equation (20), we can solve W by the regularized least squares problem using equation (21).

$$W = (BB^T + \alpha I)^{-1}BY^T \tag{21}$$

Bits-calculation-step: When all variables but B fixed, we can rewrite the problem (20) into form:

$$\min_B \left|\left|L - W^T B\right|\right|^2 + \nu \left|\left| B - F(X)\right|\right|^2$$
$$\text{s.t. } B \in \{-1, 1\}^{L\times n} \tag{22}$$

The equation above is still NP-hard. However, a closed-form solution can be achieved when we try to recover one-row B by fixing all the other rows. The problem (22) can be rewritten as:

$$min_B||Y||^2 - 2Tr(Y^T W^T B) + ||W^T\ B||^2 + \nu(||B||^2 - 2Tr\left(B^T F(X)\right) + ||F(X)||^2)$$
$$\text{s.t. } \ B \in \{-1, 1\}^{L\times n}\ , \tag{23}$$

which is equivalent to:

$$min_B\ ||W^T B||^2 - 2Tr(B^T Q)$$
$$\text{s.t. } \ B \in \{-1, 1\}^{L\times n}\ , \tag{24}$$

where Q represents $WY + \nu F(X)$ and $Tr(.)$ is the trace norm.

In order to minimize the l_2 loss, we need to iteratively learn every single bits in B using the discrete cyclic coordinate descent (DCC) method. Let z^T be the l^{th} row of B, $l = 1,2,3, \dots, L$, and B` is the matrix of $B$ excluding $z$ . Then z is one bit for all $n$ samples. Similarly, assume $q^T$ be $l^{th}$ row of $Q$ and $v^T$ is $l^{th}$ row of $W$, $Q$` is the matrix Q excluding q and W` is the matrix W excluding v. As analyzed in [9], we finally rewrite the equation (24) into form:

$$min_z\left(v^T {W'}^T B' - q^T\right) z$$
$$\text{s.t. } z \in \{-1, 1\}^n \tag{25}$$

It has the optimal solution:

$$z = sgn(q - B^{`T} W' v) \tag{26}$$

Each bits b in B is compute from the pre-trained $L-1$ bits vector B'. In our experiments, the whole L bits for X are iteratively learned from $5L$ times.

In our observation, the feature vector extracted from CNN is always sparse. Therefore there is huge bits compression capacity for image representation with SDH. The SDH code learning includes the training process which is expected to further enlarge the margin between different classes, to improve the overall matching accuracy.

3. Experiments and Results

We performed a range of experiments to ascertain the comparative performance of the cross-spectral iris recognition. Two publicly available cross spectral iris database, PolyU bi-spectral iris database [3] and cross-eyed-cross-spectral iris recognition database [4] are employed to ascertain performance improvement. In order to ascertain the usefulness of the cross-spectral iris recognition framework for the cross-sensor iris recognition problem, we also performed additional experiments using publicly available UND dataset, and these are also detailed in section 4.

We briefly introduce the publicly available databases employed in our experiments in the following subsection and discuss our experimental results in subsection 3.2.

3.1 DATABASE

Two recently available cross-spectral iris image datasets provide iris images acquired from 209 subjects and 120 subjects.

3.1.1 PolyU Bi-spectral Iris Database

PolyU bi-spectral iris database includes 418 classes bi-spectral images acquired from 209 subjects. There are 15 instances for every spectrum in each of the class. Images from two spectra were acquired simultaneously in this database. In total, there are 12540 iris images ($209 \times 2 \times 2 \times 15$). The origin images dimension is 640×480 pixels. We used publicly available implementation for iris segmentation algorithm in [13] to accurately segment iris images for the experiments. The dimension of each of the automatically segmented and normalized iris image is 512×64 pixels. Sample images of this database are shown in Figure 5(a).

3.1.2 Cross-Eyed-Cross-Spectral Iris Recognition database

Cross-eyed-cross-spectral iris recognition database provides 3840 iris images from 240 classes acquired from 120 subjects. Each of the eyes images from every subject for both spectra has eight image samples of 400×300 pixels. We use same iris segmentation algorithm [13] to automatically segment all the iris images. The dimension of all the segmented and normalized iris images from this dataset is 512×64 pixels. The sample images from the cross-eyed-cross-spectral database are shown in Figure 5(b).

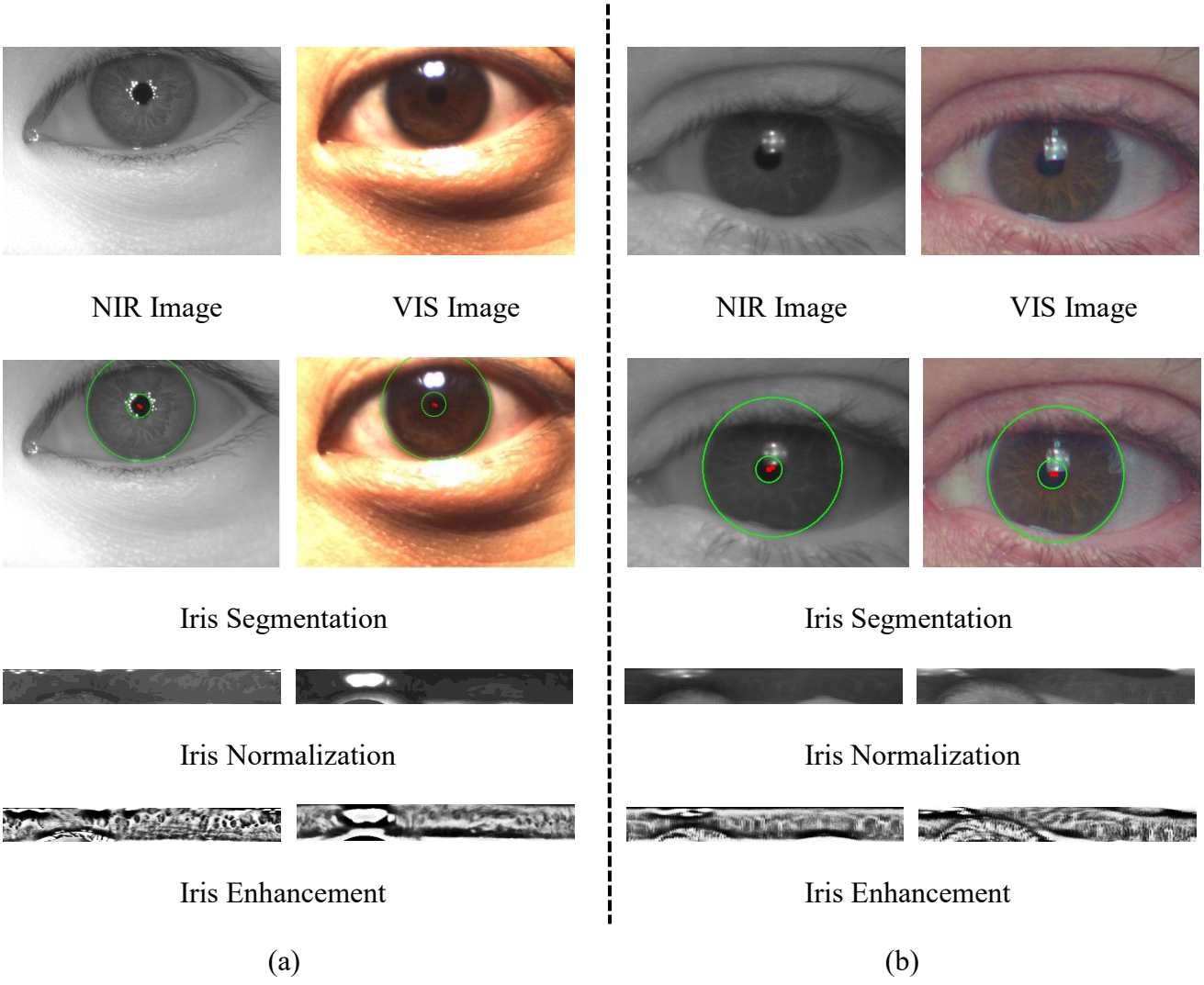


Figure 5: Sample iris image preprocessing for (a) PolyU bi-spectral iris database and (b) Cross-eye-cross-spectral iris recognition database.

Both the databases have low-quality image samples and some representative samples from the databases are shown in Figure 6. The key image degradation factors in PolyU database are iris occlusion and poor lighting conditions. In the cross-eyed database, the degradation in image quality mainly results from the reflection and pose variations.

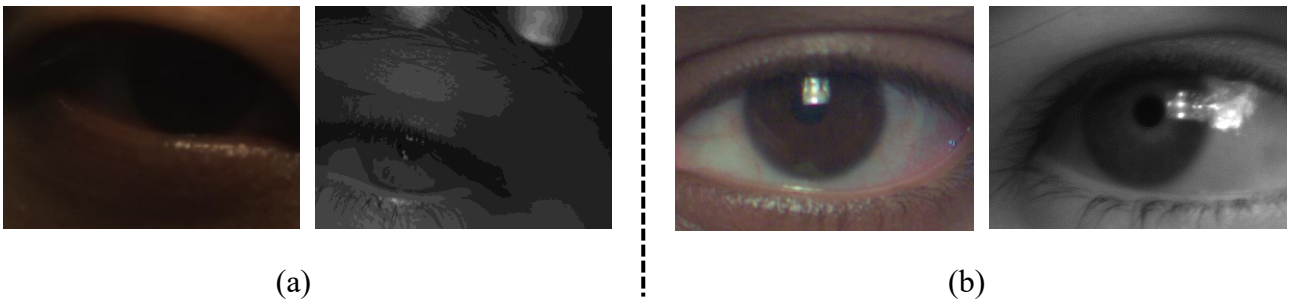


Figure 6: Low-quality image samples from (a) PolyU bi-spectral iris database and (b) Cross-eyed-cross-spectral iris recognition database.

3.2 MATCHING RESULTS FOR CROSS-SPECTRAL IRIS RECOGNITION

3.2.1 POLYU BI-SPECTRAL IRIS DATABASE

Our first set of experiments to ascertain the performance of cross-spectral iris recognition incorporated this publicly available dataset. In order to ensure fairness in the comparison using the approach in [3], we used iris images from 140 subjects for the experiment. The first ten instances were selected for the CNN training and the rest five instances were employed for the matching. The matching protocol was all-to-all which generated 2800 ($10 \times 140 \times 2$) genuine scores and 1953000 ($6975 \times 140 \times 2$) imposter scores.

Figure 7 depicts the visualization of kernels trained for the first convolutional layer of CNN with *softmax* cross entropy loss using PolyU bi-spectral iris dataset. It is generally observed [19] that the well-trained kernels present higher visual saliency reflecting its sensitivity to the edges with a different orientation. There are many kernels in this deep neural network architecture and Figure 7 is randomly chosen to visualize the effectiveness of the training and represents the self-learned features that are extracted from the edge, corners and other textures in the iris data.

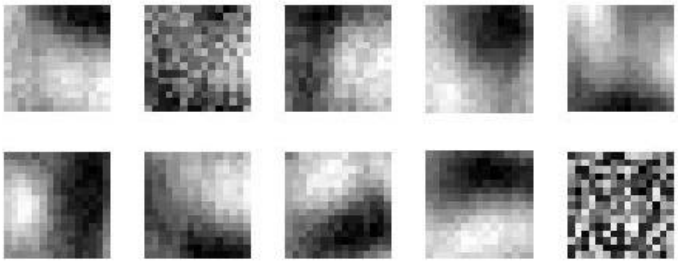


Figure 7 Visualization of CNN kernels from the first convolutional layer.

The CNN with *softmax* cross-entropy loss is used as the feature extractor in this work. The feature vectors generated from this network are further classified by the joint-Bayesian inference and the SDH as discussed in section 2. In SDH, all the feature vectors are hashed into a 1000 bits binary vector. The comparative matching results using the baseline MRF approach in [3] and using the popular *IrisCode* approach (similar to as in [24]) are used to ascertain the performance. The receiver operating characteristic (ROC) from this set of experiments is shown in Figure 8 while the equal error rate (EER) results are summarized in Table 2.

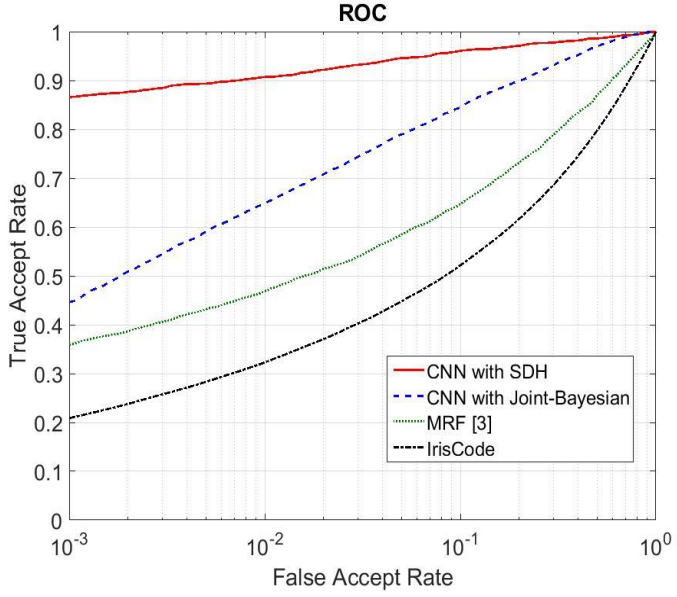


Figure 8: ROC from PolyU bi-spectral iris dataset (140 subjects as in [3]).

Table 2: Comparative results using PolyU bi-spectral iris database

Approach	EER
CNN-SDH (Ours)	5.39%
CNN-Joint Bayesian(Ours)	8.00%
MRF [3]	24.50%
IrisCode	30.81%

The experimental results illustrated in Figure 8 indicate that our approach for cross-spectral iris matching significantly improves the matching accuracy as compared with the algorithms with hand-crafted features presented in earlier work [3], [24]. It can be observed from these results that the usage of SDH generates outperforming results as compared with those from the joint-Bayesian inference. It can be observed from Figure 9 and Table 3 that the binary features generated from the SDH scheme increases the average separation between the genuine and imposter scores, as compared with those from the joint-Bayesian approach. The templates generated from SDH scheme are also significantly compact which can significantly reduce the storage requirements as Table 4 shows and also improve the matching speed between such compact templates.

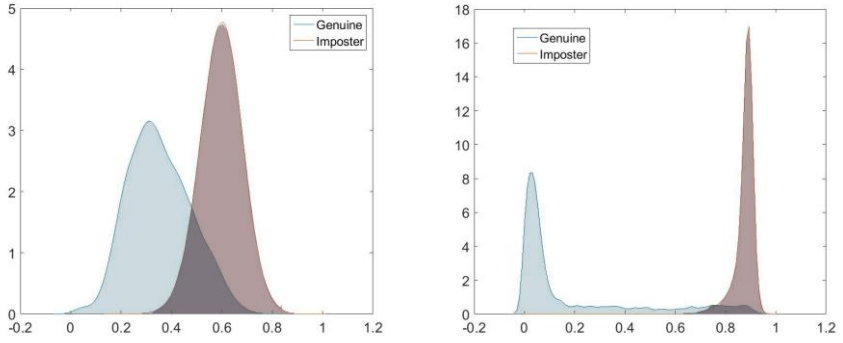
Table 3: Comparative DI from best-performing methods on PolyU bi-spectral iris dataset

Approach	Decidability Index (DI)
CNN-SDH (1000 bits)	2.1275
CNN-Joint Bayesian	1.1943

Table 4: Compression ratio with SDH on PolyU bi-spectral iris dataset

Origin templates (bits)	262144
Binary vector (bits)	1000
Compression ratio	262.144

The usage of supervised discrete hashing is further investigated to ascertain the impact of changes in the length of hashing bits on the performance. Therefore, the performance of the cross-spectral iris matching using the 500, 1000, and 1500 bits are also explored and is shown in Figure 10. These results indicate that the choice of 1000 bits can offer a reasonable tradeoff between the matching accuracy and the template storage size.



(a) (b)

Figure 9: Match score distribution from PolyU bi-spectral Iris database for (a) CNN with Joint Bayesian and (b) CNN with SDH 1000 bits.

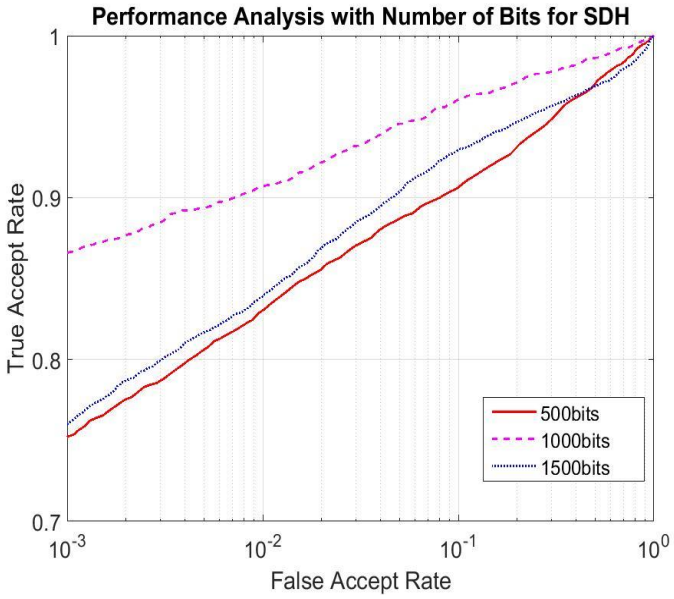


Figure 10: Performance analysis with number of bits for SDH on PolyU bi-spectral iris database.

It is worth noting that the PolyU bi-spectral iris database provides images from 209 different subjects, but only the images from first 140 different subjects were employed in [3]. Therefore we also investigated comparative cross-spectral iris matching accuracy using the images from all the 209 subjects in this database. These experiments generated 4180 (209 × 2 × 10) genuine and 4367650 (209 × 2 × 10425) imposter scores. The ROC from this set of experiments is shown in Figure 11. It should be noted that as compared to the results in Figure 8, the performance of CNN and SHD has comparatively degraded from EER of 5.39% to EER of 12.41% and this can be attributed to degraded image quality images from about 69 subjects. Figure 6 illustrates samples from such poor quality iris images that are also used in this set of experiments.

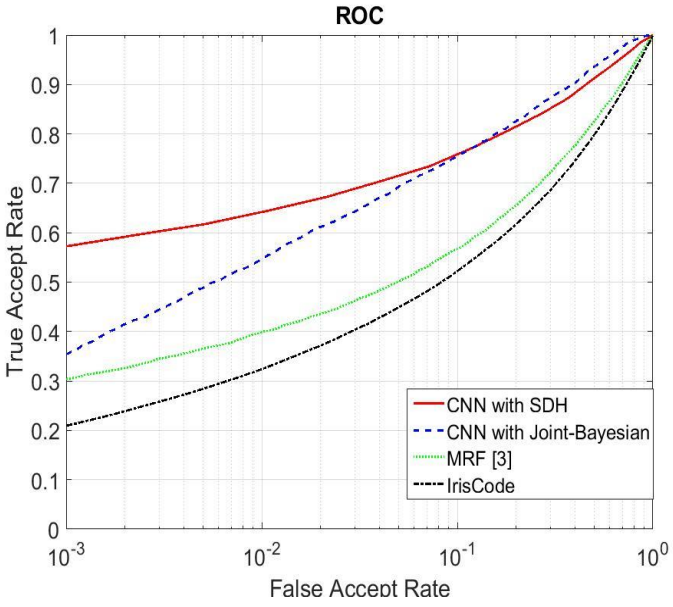


Figure 11: ROC from PolyU bi-spectral iris database (209 subjects).

3.2.2 Cross-Eyed-Cross-Spectral Iris Recognition Database

Another set of experiments are performed using the cross-eyed database to ascertain the performance. In this set of experiments, first five instances from every subject are used for the CNN training while the left three instances from all the subjects are employed to test cross-spectral iris matching performance. Such all-to-all matching protocol generates 2160 ($120 \times 2 \times 9$) genuine and 516,240 ($120 \times 2 \times 2151$) impostor match scores for the performance evaluation.

This set of experiments also incorporates CNN with *softmax* cross-entropy loss and the two classification algorithms. The length of feature vector generated from the hashing process is still the same as 1000 bits. Comparative performance using two baseline methods, *i.e.*, MRF and the *IrisCode*, is evaluated. The resulting ROCs are shown in Figure 12 and the EER are summarized in Table 5. These experimental results indicate that the combination of CNN with *softmax* cross-entropy loss and

SDH consistently achieves outperforming results for the cross-spectral iris matching. Figure 13 illustrates the distribution of genuine and impostor match scores to ascertain the influence of SDH scheme. The plots in this figure and the Decidability Index (DI) in Table 6 indicate that the usage of SDH improves the average distance between genuine and impostor scores.

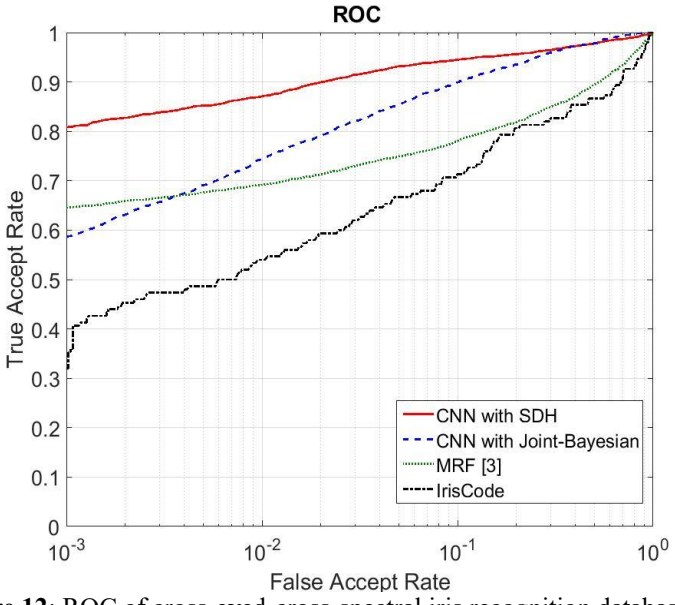


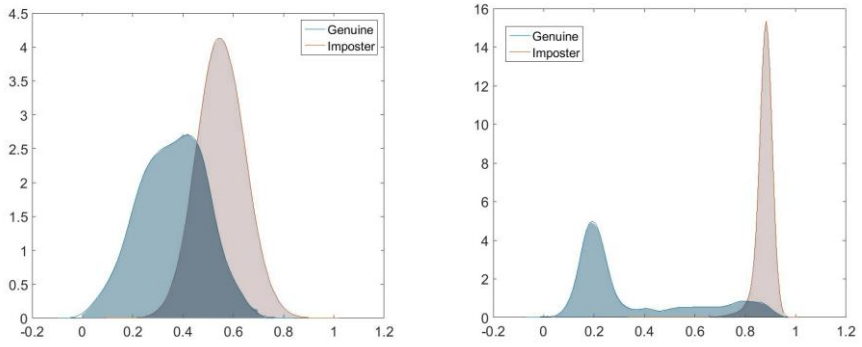
Figure 12: ROC of cross-eyed-cross-spectral iris recognition database.

Table 5: Comparative results using cross-eyed-cross-spectral iris recognition database.

Approach	EER
CNN-SDH (Ours)	6.34%
CNN-Joint Bayesian (Ours)	10.07%
MRF [3]	18.40%
IrisCode	19.48%

Table 6: Comparative DI from best-performing methods on cross-eyed-cross-spectral iris recognition dataset

Approach	Decidability Index (DI)
CNN-SDH (1000 bits)	2.5346
CNN-Bayesian	1.6257



(a) (b)

Figure 13: Match score distribution using cross-eyed-cross-spectral Iris recognition database for (a) CNN with Joint Bayesian and (b) CNN with SDH 1000 bits.

The influence from the number of bits for SDH on the matching accuracy for this dataset is also investigated. The test results illustrated in Figure 14 suggest that the choice of 1000 bits can still offer superior performance and reduction in template size. However, the difference in the performances from the template bit length changes is less significant compared with those for the PolyU bi-spectral iris database. Therefore when the template size is of key importance, 500-bits template size can be the judicious choice for faster template matching.

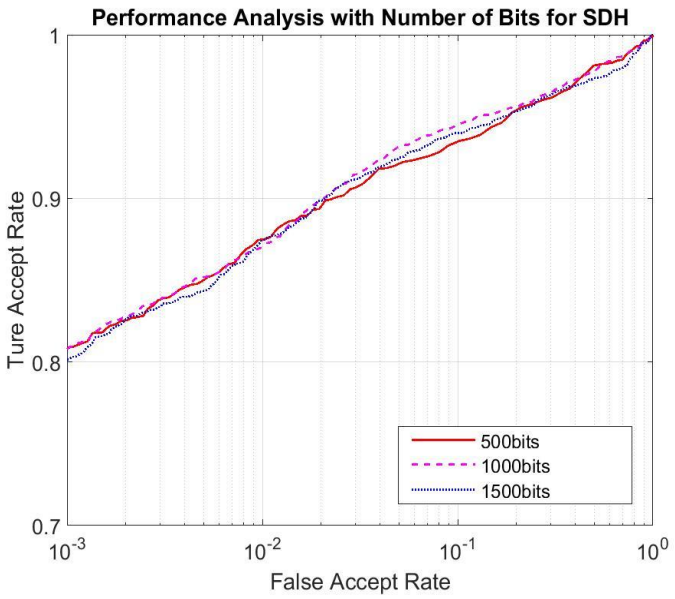
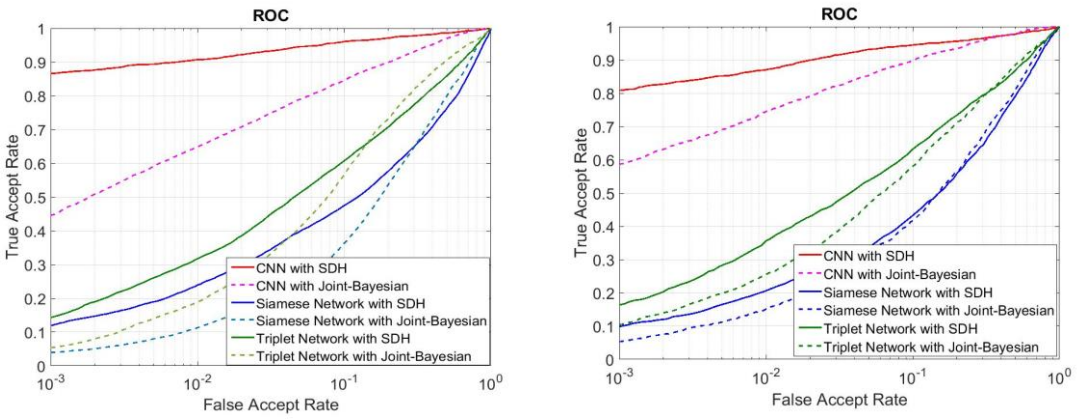


Figure 14: Performance analysis with number of bits for SDH using cross-eyed-cross-spectral iris recognition database.

4. Discussion

The experimental results presented in section 3.2 indicate superior results as compared with the earlier methods presented in the literature [3]-[24] for the cross-spectral iris recognition. We also perform experiments using other popular networks to comparatively ascertain the performance of the scheme discussed in section 2. The Siamese network and Triplet network are used to ascertain the performance using the same training and testing data/protocols as detailed in section 3.2. The results from these experiments using the ROCs are illustrated in Figure 15. Both of these CNNs compute the loss based on the binary labels that indicate genuine or imposter pair rather than the class label used in a typical *softmax* cross-entropy loss for the network in figure 2. It may be noted that the dataset used for the training generates a large number of imposter pairs than the number of genuine pairs. Therefore, the triplet loss function in used in these experiments can make it very difficult to span through all



(a) (b)

Figure 15: Comparative performance using other CNN architectures

the possible cases during the training process. This is a plausible reason for inferior performance from these two architectures, as compared to CNN with *softmax* cross-entropy loss employed during the experiments detailed in section 3.2.

More recent networks have shown impressive learning capability in large-scale image classification as compared to the AlexNet. Therefore, we performed new experiments to evaluate the 1000-bits code generated from the supervised discrete hashing using VGG-16 [31] and also using ResNet-50 [32]. In order to ensure fairness in the comparison, we employed same protocols as detailed in section 3.2. These experimental results are shown in Figure 16 and Table 7. The training data available to us for our problem is too small to converge the VGG-16 net and ResNet-50. Therefore, we performed the transfer learning by fine-tuning the *caffe* models trained from the *ImageNet* classification, despite the fact that the CNN model in our framework has been trained from scratch. The experimental results in Figure 16 and Table 7 indicate that the VGG16 can further improve the performance of the CNN in our framework.

However, there are two limitations that should be carefully considered with such alternative. Firstly, VGG16 needs much more parameters as compared with the CNN in our framework (98,596,544 versus 635,650) which can be a burden for many thin-client operations. Secondly, the VGG16 is fine-tuned from the pre-trained model and it has already learned diverse features from *ImageNet*, which means the training data is not exactly same and it cannot be considered as a fair comparison. We can also note

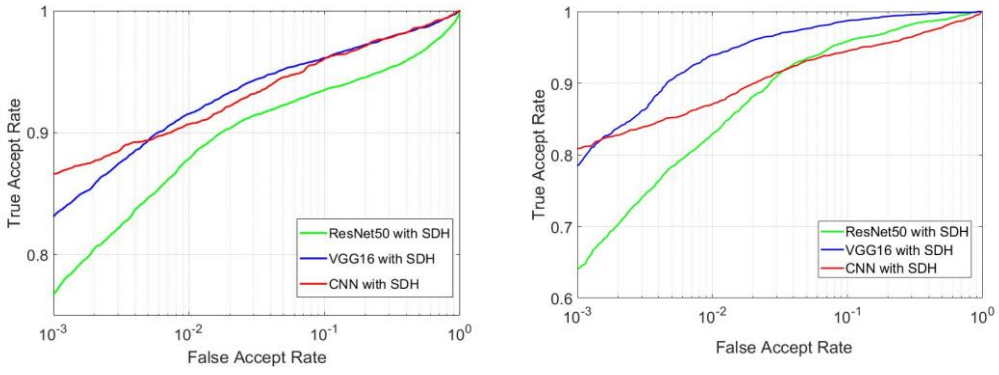


Figure 16: Different convolutional neural networks testing results from (a) PolyU bi-spectral iris database and (b) Cross-eye-cross-spectral iris recognition database.

that the *ResNet* performs poorly and this can be attributed due to the overfitting as our experiments indicated that the training loss gets minimized while the classification accuracy approaches to one after about one thousand iterations.

Table 7: Comparative results using the EER for the two databases.

	PolyU bi-spectral iris database	Cross-eye-cross-spectral iris recognition database
CNN with SDH	5.39%	6.34%
VGG16 with SDH	4.85%	3.13%
ResNet50 with SDH	7.17%	6.11%

Although our framework focuses on cross-spectral matching which is a more

challenging problem, we also performed the following ablation tests with same spectral data from both databases. Each of the cross-spectral datasets was partitioned into two same-spectral ones and performs the same-spectral matching. The baseline results from the *IrisCode* [3] using Log Gabor filter are also shown in Figure 17. It can be observed these results that our approach outperforms the baseline algorithm in most cases even for the same-spectrum matching. During the NIR matching for the cross-eye-cross-spectral matching, our approach indicates slightly better performance except at the lower end of the high false acceptance. Since our framework is designed to address the cross-spectral iris matching problem, these additional results also support the generalization capability from our framework for the same spectrum iris matching problem.

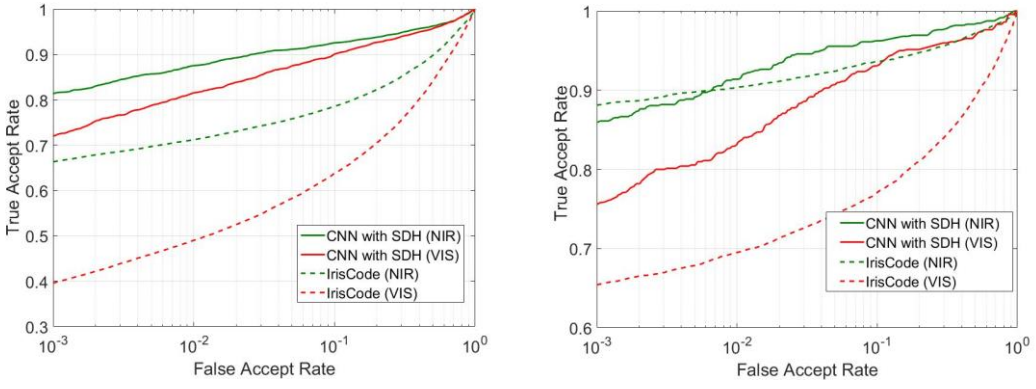
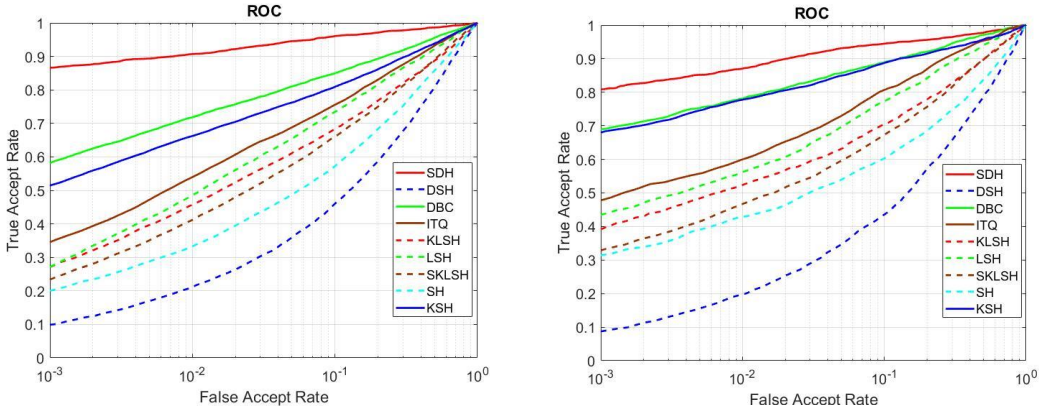


Figure 17: Same spectral iris matching results for (a) PolyU bi-spectral iris database and (b) Cross-eye-cross-spectral iris recognition database.

There are many promising hashing algorithms in the literature, *e.g.*, Locality Sensitivity Hashing (LSH) [33], Spectral Hashing (SPH) [38], Kernel LSH (KLSH) [34], Locality-sensitive binary codes from shift-invariant kernels (SKLSH) [35], Kernel-based Supervised Hashing (KSH) [36], Discriminative Binary Code (DBC) [37], and

Iterative Quantization (ITQ) [39], deep supervised hashing (DSH) [12]. It is worth investigating their effectiveness as compared with the SDH and therefore separate experiments were performed. The codes length in these experiments was set to 1,000 bits for the fair comparison. The ROC results are shown in Figure 19 and EER details are listed in Table 8. It can be observed from these results that the SDH significantly outperforms other methods for the hashing the feature vectors on both the datasets.



(a) **(b)**

Figure 18: Comparative results using different hashing algorithms for (a) PolyU bi-spectral iris database and (b) Cross-eye-cross-spectral iris recognition database.

Table 8: EER from two datasets using different hashing algorithms.

Approach	PolyU bi-spectral iris database	Cross-eye-cross-spectral iris recognition database
SDH	5.39%	6.34%
LSH	19.13%	17.73%
SPH	26.75%	27.18%
KLSH	26.74%	22.68%
KSH	15.82%	10.77%
DBC	13.24%	10.69%
ITQ	18.15%	15.69%
DSH	31.32%	32.33%

We also perform set of experiments to ascertain the effectiveness of the proposed cross-spectral iris matching approach for the cross-sensor iris recognition. This set of experiments utilize publicly available UND 2012 cross-sensor database [29]. The train-test protocol is consistent with those for the experiments on the PolyU bi-spectral iris database in section 3.2. We select iris images from the 214 different subjects in this database for the experiments. For every subject, ten instances are used for the training, and other five instances are used for the performance evaluation. Therefore, a total of 5350 (214 × 25) genuine and 1139550 (214 × 25 ×213) impostor scores are generated from the experiments and the resulting ROC curves are shown in Figure 19. Table 9 summarizes corresponding EER values from the experiments. These results indicate the effectiveness of the cross-spectral iris matching approach introduced in section 2 over other competing methods.

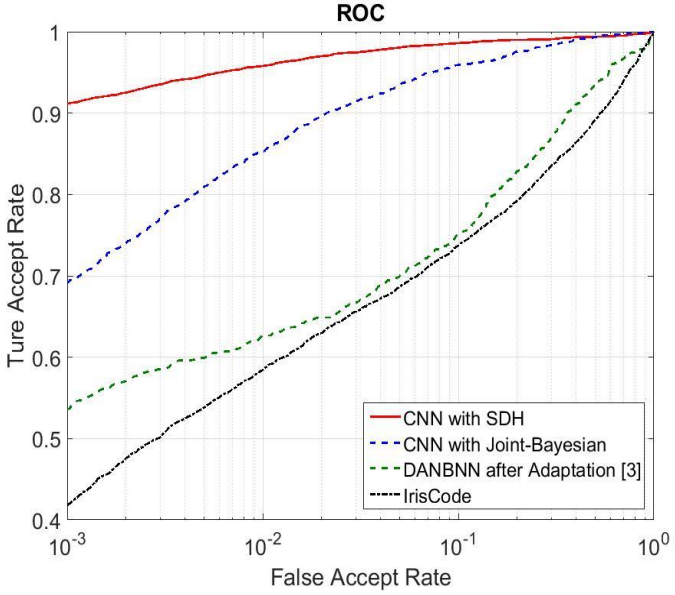


Figure 19: ROC from UND 2012 cross sensor database.

Table 9: Comparative results from UND 2012 cross sensor database.

Approach	EER
CNN-SDH [Ours]	4.50%
CNN-Joint Bayesian [Ours]	5.91%
DANBNN after adaptation [3]	18.20%
IrisCode	20.52%

5. Conclusions and Further Work

This paper introduces a new framework for the cross-spectral iris recognition using deep convolutional neural networks. We present experimental results using two publicly available cross-spectral iris databases. These results indicate superior performance over the earlier methods presented in the literature that incorporated handcrafted features for matching cross-spectral iris data. This paper also investigates the effectiveness of supervised discrete hashing algorithm for improving the cross-spectral iris matching accuracy and reducing the template size. Our results presented in section 3.2 indicate further improvement in the matching accuracy which is resulting from the improvement in the average distance between the genuine and imposter match scores generated from Hamming distance of the binary features. The usage of non-iris regions in [8] cannot be considered as the iris recognition but the ocular recognition [14] or as iris recognition with the usage of information in the vicinity. The key objective of this work is to evaluate cross-spectral iris recognition capabilities, using the iris regions defined in [12] or widely incorporated in the deployed systems. Addition of periocular information, along with those from the iris regions in this work, is expected to further improve [6] cross-spectral iris matching

capabilities and is suggested for further work. There have been more recent advances in data-dependent hashing algorithms [44]-[46] and effectiveness of such algorithms to improve cross-spectral iris recognition is also part of further extension of this work.

References

[1] K. W. Bowyer, K. Hollingsworth, and P. J. Flynn, "Image understanding for iris biometrics: A survey," *Computer Vision and Image Understanding*, vol. 110, no. 2, pp. 281–307, 2008.

[2] I. Nigam, M. Vatsa, and R. Singh, "Ocular biometrics: A survey of modalities and fusion approaches," *Information Fusion*, vol. 26, pp. 1–35, 2015.

[3] N. P. Ramaiah and A. Kumar, "Toward More Accurate Iris Recognition Using Cross-Spectral Matching," *IEEE Transactions on Image Processing*, vol. 26, no. 1, pp. 208–221, 2017.

[4] A. Sequeira, L. Chen, P. Wild, J. Ferryman, F. Alonso-Fernandez, K. B. Raja, R. Raghavendra, C. Busch, and J. Bigun, "Cross-Eyed - Cross-Spectral Iris/Periocular Recognition Database and Competition," *2016 International Conference of the Biometrics Special Interest Group (BIOSIG)*, 2016.

[5] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[6] C.-W. Tan and A. Kumar, "Towards Online Iris and Periocular Recognition under Relaxed Imaging Constraints," *IEEE Transactions on Image Processing,* vol. 22, no. 10, pp. 3751–3765, 2013.

[7] N. Liu, H. Li, M. Zhang, J. Liu, Z. Sun, and T. Tan, "Accurate iris segmentation in non-cooperative environments using fully convolutional networks," *2016 International Conference on Biometrics (ICB)*, 2016.

[8] A. Gangwar and A. Joshi, "DeepIrisNet: Deep iris representation with applications in iris recognition and cross-sensor iris recognition," *2016 IEEE International Conference on Image Processing (ICIP)*, 2016.

[9] F. Shen, C. Shen, W. Liu, and H. T. Shen, "Supervised Discrete Hashing," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

[10] J. Daugman, "How iris recognition works". *IEEE Transactions on circuits and systems for video technology*, 2004, 14(1): 21-30.

[11] A. K. Jain and A Kumar, “Biometrics of next generation: An overview,” *Second Generation Biometrics: The Ethical, Legal and Social Context*, E. Mordini and D. Tzovaras (Eds) , pp 49-79, Springer, 2012.

[12] ISO/IEC 19794-6:2005. Information technology -- Biometric data interchange formats -- Part 6: Iris image data.

[13] Z. Zhao and A. Kumar, “An Accurate Iris Segmentation Framework Under Relaxed Imaging Constraints Using Total Variation Model,” *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015.

[14] Z. Zhao and A. Kumar, "Accurate Periocular Recognition Under Less Constrained Environment Using Semantics-Assisted Convolutional Neural Network", *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 5, pp. 1017-1030, 2017.

[15] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," *Proc*. *IEEE* Conf. *Computer Vision and Pattern Recognition, CVPR 2015,* pp. 1-9, 2015.

[16] Y. Sun, X. Wang, and X. Tang, "Deep learning face representation from predicting 10,000 classes", *in Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, 2014, pp. 1891-1898.

[17] D. Menotti, G. Chiachia, A. Pinto, W. R. Schwartz, H. Pedrini, A. X. Falcão, and A. Rocha, “Deep representations for iris, face, and fingerprint spoofing detection,” *IEEE Trans. Info. Forensics & Security*, vol. 10, no. 4, pp. 864-879, 2015. 2016.

[18] D. Chen, X. Cao, L. Wang, F. Wen, and J. Sun, “Bayesian Face Revisited: A Joint Formulation,” *Computer Vision – ECCV 2012 Lecture Notes in Computer Science*, pp. 566–579, 2012.

[19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Advances in neural information processing systems*, pp. 1097-1105, 2012.

[20] S. Chopra, R. Hadsell, and Y. Lecun, “Learning a Similarity Metric Discriminatively, with Application to Face Verification,” *Proc. CVPR 2005,* 2005.

[21] E. Hoffer and N. Ailon, “Deep Metric Learning Using Triplet Network,” *Similarity-Based Pattern Recognition Lecture Notes in Computer Science*, pp. 84–92, 2015.

[22] B. Kulis and T. Darrell, “Learning to hash with binary reconstructive embeddings,” *Advances in neural information processing systems*, pp. 1042-1050, 2009.

[23] W. Liu, J. Wang, R. Ji, Y.-G. Jiang, and S.-F. Chang, "Supervised hashing with kernels," *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2012.

[24] L. Masek and P. Kovesi, "Matlab source code for a biometric identification system based on iris patterns," The School of Computer Science and Software Engineering, The University of Western Australia, 2(4), 2003.

[25] J. Zuo, F. Nicolo, and N. A. Schmid, "Cross spectral iris matching based on predictive image mapping," *2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, 2010.

[26] J. K. Pillai, M. Puertas, and R. Chellappa, "Cross-Sensor Iris Recognition through Kernel Learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 1, pp. 73–85, 2014.

[27] L. Xiao, Z. Sun, R. He, and T. Tan, "Coupled feature selection for cross-sensor iris recognition," *Proc. IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, 2013.

[28] K. W. Bowyer, S. E. Baker, A. Hentz, K. Hollingsworth, T. Peters, and P. J. Flynn, "Factors that degrade the match distribution in iris biometrics," *Identity in the Information Society*, vol. 2, no. 3, pp. 327–343, 2009.

[29] Y. Lee, R. Micheals, J. Filliben, and P. Phillips, "Robust iris recognition baseline for the grand challenge," *NIST-IR report*, vol. 7777, 2011.

[30] N. P. Ramaiah and A. Kumar, "Advancing Cross-Spectral Iris Recognition Research Using Bi-Spectral Imaging," *Machine Intelligence and Signal Processing Advances in Intelligent Systems and Computing*, pp. 1–10, 2016.

[31] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *Proc. Intl. Conf. Learning Representations*, *ICLR 2015.*

[32] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *IEEE Conference on Computer Vision and Pattern Recognition*, *Proc. CVPR 2016.*

[33] A. Gionis, P. Indyk, and R. Motwani, "Similarity search in high dimensions via hashing,*"25th Intl. Conf. Very Large Databases*, *VLDB 1999*.

[34] B. Kulis and K. Grauman. "Kernelized locality-sensitive hashing for scalable image search," *IEEE Intl. Conf. Computer Vision*, *ICCV 2009*.

[35] M. Raginsky and S. Lazebnik. "Locality-sensitive binary codes from shift-invariant kernels," *Advances in Neural Information Processing Systems*, *NIPS 2009*.

[36] W. Liu, J. Wang, R. Ji, Y.-G. Jiang, and S.-F. Chang. "Supervised hashing with kernels," *IEEE Conf. Computer Vision and Pattern Recognition*, *CVPR* 2012, *Providece*, Jun. 2012.

[37] M. Rastegari, A. Farhadi, and D. Forsyth, "Attribute discovery via predictable discriminative binary codes," *Computer Vision – ECCV 2012 Lecture Notes in Computer Science*, pp. 876–889, 2012.

[38] Y. Weiss, A. Torralba, and R. Fergus. "Spectral hashing," *Advances in Neural Information Processing Systems*, *NIPS*, 2008.

[39] Y. Gong, S. Lazebnik, A. Gordo, and F. Perronnin. "Iterative quantization: A procrustean approach to learning binary codes for large-scale image retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 35, no. 12, pp. 2916–2929, 2013.

[40] K. Nguyen, C. Fookes, R. Jillela, S. Sridharan, and A. Ross, "Long range iris recognition: A survey," *Pattern Recognition*, vol. 72, pp. 123–143, 2017

[41] N. Liu, M. Zhang, H. Li, Z. Sun, and T. Tan, "Deepiris: Learning pairwise filter bank for heterogeneous iris verification," *Pattern Recognition Letters*, vol. 82, pp. 154–161, 2016.

[42] J. Zhang, J. Du, S. Zhang, D. Liu, Y. Hu, J. Hu, S. Wei, and L. Dai, "Watch, attend and parse: An end-to-end neural network based approach to handwritten mathematical expression recognition," *Pattern Recognition,* vol. 71, pp. 196–206, 2017.

[43] Q, Zheng, A. Kumar, and G. Pang, "A 3D descriptor recovered from a single 2D palmprint image," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, pp. 1272–1279, June 2016.

[44] J. Gui, T. Liu, Z. Sun, D. Tao, and T. Tan, "Supervised discrete hashing with relaxation," *IEEE Transactions on Neural Networks and Learning Systems*, 2016.

[45] J. Gui, T. Liu, Z. Sun, D. Tao, and T. Tan, "Fast supervised discrete hashing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 2, pp. 490–496, 2018.

[46] Y. Cao, H. Qi, W. Zhou, J. Kato, K. Li, X. Liu, and J. Gui, "Binary hashing for approximate nearest neighbor search on big data: A survey," *IEEE Access*, vol. 6, pp. 2039–2054, 2018.

[47] Y. Li, G. Wang, L. Nie, Q. Wang, and W. Tan, "Distance metric optimization driven convolutional neural network for age invariant face recognition," *Pattern Recognition*, vol. 75, pp. 51–62, 2018.