Predicting Segmentation Errors in an Iris Recognition System

Nitin K. Mahadeo, Gholamreza Haffari, Andrew P. Papliński
Clayton School of Information Technology
Monash University
{nitin.mahadeo, gholamreza.haffari, andrew.paplinski}@monash.edu

Abstract

Iris segmentation is defined as the isolation of the iris pattern in an eye image. A highly accurate segmented iris plays a key role in the overall performance of an iris recognition system, as shown in previous research. We present a fully automated method for classifying correctly and incorrectly segmented iris regions in eye images. In contrast with previous work where only iris boundary detection is considered (using a limited number of features), we introduce the following novelties which greatly enhance the performance of an iris recognition system. Firstly, we go beyond iris boundary detection and consider a more realistic and challenging task of complete segmentation which includes iris boundary detection and occlusion detection (due to eyelids and eyelashes). Secondly, an extended and rich feature set is investigated for this task. Thirdly, several non-linear learning algorithms are used to measure the prediction accuracy. Finally, we extend our model to iris videos, taking into account neighbouring frames for a better prediction. Both intrinsic and extrinsic evaluation are carried out to evaluate the performance of the proposed method. With these innovations, our method outperforms current state-of-the-art techniques and presents a reliable approach to the task of classifying segmented iris images in an iris recognition system.

1. Introduction

A typical iris recognition system consists of four stages, namely image acquisition, iris segmentation, feature extraction and matching. The accuracy of the segmentation stage plays a key role in the overall performance of an iris recognition system. The iris regions in eye images acquired in different environments need to be accurately segmented for better performance in the later stages of the system. However, accurate iris segmentation is a challenging task, especially in images captured in less constrained environments where imaging conditions are not always optimal. We present a fully automated and accurate method for classifying correctly and incorrectly segmented iris regions in eye images.

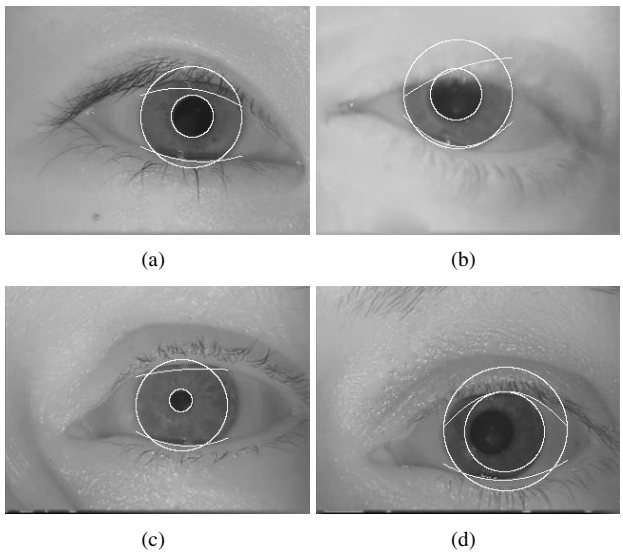


Figure 1: Iris segmentation assessment: (a) Properly segmented iris region (b) Incorrect upper eyelid detection (c) Inaccurate iris boundary detection (d) Failed pupil and iris border segmentation.

A correctly segmented iris region in an eye image involves isolating the iris pattern. This is achieved by (1) localizing the pupil boundary, (2) localizing the iris boundary, and (3) evaluating the amount of occlusion caused by the eyelids and eyelashes in the iris region. As an example, Figure 1(a) shows a properly segmented iris region where the aforementioned three criteria are satisfied. In contrast, Figure 1(b) is considered to be a segmentation failure as a result of inaccurate upper eyelid detection. Similarly, Figure 1(c) is also treated as a segmentation error due to inaccurate detection of the iris border, and Figure 1(d) has both the pupil and the iris regions incorrectly localized.

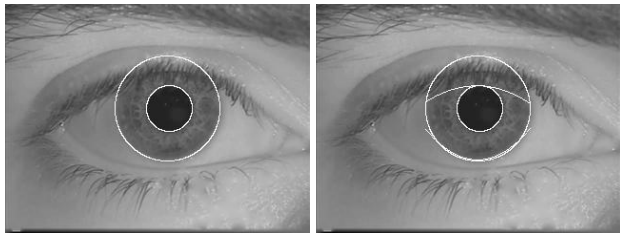
Recently, there have been two major developments in the iris biometric area. They include the adoption of a video-

based approach in the acquisition stage and the capture of iris images in less constrained environments. The success of these two approaches plays a key role in the adoption of the iris biometric by a larger community. However, this also presents new challenges. Due to limitations of segmentation algorithms and the presence of noise in eye images acquired in different settings, iris segmentation failures are more probable. A segmentation failure is an important event which needs to be accurately identified. This is even more relevant to the recent video-based iris recognition systems where identification of high quality iris-segmented frames and their neighbours can enhance the performance of the recognition pipeline [7],[14],[8].

Automatic prediction of (in)correctly segmented iris images can enhance the capability of an iris recognition system. An undetected incorrectly-segmented iris degrades the performance of the system and results in more processing time due to running of the later stages of the pipeline. An example of such a development is a subject being incorrectly classified as an impostor (false positive) as a consequence of poor segmentation evaluation. This in turn requires re-submission of an eye image for proper identification. Although several segmentation models have been proposed in the literature, there has been minimal work in the classification of segmented iris images [3]. Visual inspection appears to be the predominant way of evaluating the accuracy of the iris segmentation in the literature. This requires trained personnel to determine if an iris has been properly segmented. Furthermore, given the availability of large volume of frames in iris videos, a fast and more efficient way of classifying segmented iris images is necessary. The availability of such a model to the research community will also facilitate the publication of segmentation results in using the same standards. These arguments stress the crucial need for a fully automated model to predict iris segmentation failures.

In this paper, we propose and design a model for automatic classification of iris-segmented images. We introduce the following novelties which greatly enhance the performance of an iris recognition system. Firstly, we go beyond iris boundary detection and consider a more realistic and challenging task of complete segmentation which includes iris boundary detection and occlusion detection (due to eyelids and eyelashes). Secondly, an extended and rich feature set is investigated for this task. Thirdly, we examine the performance of the system using several non-linear learning algorithms. Finally, we extend our model to iris videos to leverage the temporal correspondence of frames in iris videos.

The rest of the paper is organized as follows. The next section discusses the related work. Next, we describe our method for constructing predictive features. This is followed by the presentation of our classification models for predicting the quality of iris segmentation in Section 4. Our results and discussions are then presented in Section 5 and Section 6. Finally, we conclude in Section 7.



(a) (b)

Figure 2: (a) shows the pupil and iris regions which have been localized (b) illustrates the full segmentation model — both boundary detection and occlusion estimation.

2. Related Work

We believe that there are only two studies which have examined automatic evaluation of iris segmentation in the literature. The first work takes into account the strength of the gradient at the inner and outer iris boundaries for estimating segmentation accuracy [18]. The second study [9] only considers iris boundary detection. Occlusion estimation is not considered in their work. A probabilistic model is implemented to evaluate pupil segmentation and the distance between the centre of the pupil and that of the iris is used to confirm if the boundaries have been properly localized. The pupil and iris scores are fed into a Naive-Bayes Tree (NBTree) for prediction [10]. Other techniques used to evaluate the precision of iris segmentation appeal to the ground truth information obtained by manually segmenting the iris in eye images [3].

We would like to highlight the difference between *iris boundary detection* and *iris segmentation* tasks. Iris boundary detection refers to the case where only the inner and outer boundaries of the iris have been detected. This is typically carried out by fitting a circle or an ellipse to the inner iris borders as shown in Figure 2(a). In contrast, iris segmentation includes both boundary detection and estimation of the amount of occlusion in the iris region. This is shown in Figure 2(b). The latter (the subject of this paper) is no doubt a more challenging and realistic task than the former (the subject of the previous work), as it also involves detection of eyelids and eyelashes which may cover the iris region.

3. Feature Extraction

In this section, we begin with an examination of the "basic feature" set considered in our model. The basic feature set, which consists of both geometric and texture based

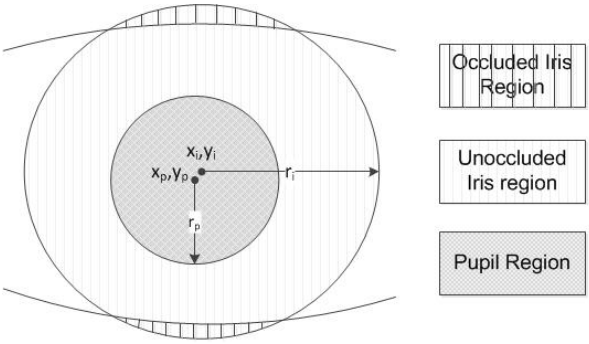


Figure 3: A schematic illustration of an eye representing the pupil, the unoccluded iris region and the occluded iris region caused by the upper and lower eyelids.

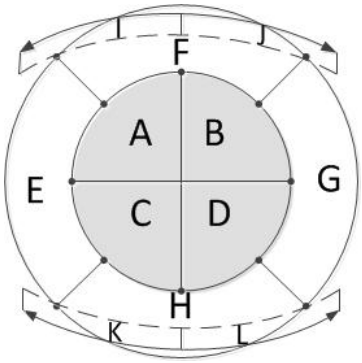


Figure 4: Texture-based features recorded as part of the basic feature set.

features, are collected during the segmentation stage in the pipeline. Afterwards, we outline the set of "composite features", each of which composed of a subset of basic features to capture higher-order interactions. We then perform feature selection and introduce a set of highly informative "important features".

3.1. Basic Features

Both the geometric and texture based features used in our algorithm were chosen based on the following two principles. First, they provide valuable information to determine segmentation accuracy. Second, they can easily be recorded by the addition of simple parameters to the segmentation algorithm.

Based on experimental observations and inherent heuristics of the eye, the following **geometric features** were considered as potential candidates in this research:

- The radius of the pupil, r_p, and that of the iris, r_i, are considered. The iris-pupil ratio, r_i/r_p, in normal eyes lies between 1.75 and 4.0 [4]. In cases of pupil dilation and constriction, this range may vary.

- The spatial coordinates of the centre of the pupil, (x_p,y_p) and that of the iris, (x_i,y_i) are recorded. These two centres typically lie close to each other [6]. This is shown in Figure 3.

- Severely occluded pupil or iris regions often results in poor segmentation. In our experiments, the following computations are performed to detect such cases. The ratio of the number of iris pixels in the un-occluded iris region to that of number of pixels in whole iris region (occluded and un-occluded) is computed. Similarly, the ratio for the amount of un-obstruction in the pupil region is computed. This is shown in Figure 3 where the amount of un-obstruction in the pupil, u_p is one and that in the iris region, u_i is less than one since occlusion is present in both the upper and lower areas of the iris.

The second set of features examined are **texture-based**. The objective here is to define a set of features which separates a good segmentation from a failed one by examining the texture in the pupil, iris and eyelid regions. The following features are investigated in this research:

- The mean intensity of the pupil and iris region are registered as μ_p and μ_i respectively. It can be observed in Figure 1(a) that the pupil region is darker than that of the iris region. [11].

- Similarly, the variance in the pupil and the iris region are listed as σ_p and σ_i. Based on our experiments on a subset of the dataset used in this work, it was observed that intensity variation within the pupil is low. This is not the case in the iris region due to occurrence of artefacts such as crypts, folds and collarettes [6].

- The entropy which is a measure of randomness is also computed. The entropy in the upper and lower half of the un-occluded iris region are denoted as H_{iu} and H_{il}. This approach ensures that the segmented iris is in both regions share the same properties.

- The localized pupil and iris region in the eye image are divided into four quadrants each and the mean intensity in each quadrant is computed excluding regions occluded by eyelid and eyelashes. This is illustrated in Figure 4 where the mean intensity in the four quadrants A, B, C and D which belong to the pupil are referred to as μ_{p1}, μ_{p2}, μ_{p3} and μ_{p4}. Similarly, the mean intensity in quadrants E, F, G and H which are part of the iris are recorded as μ_{i1}, μ_{i2}, μ_{i3} and μ_{i4}. Ideally, the intensity difference within each quadrant should be small for both the pupil and iris in a properly segmented eye image.

- Incorrect eyelid detection is detected as follows. The arc used to estimate the boundary between the iris and the

eyelid region is shown in Figure 4 as indicated by the arrows. A similar arc with a smaller radius is drawn parallel to the arcs as indicated by the dotted lines. Incorrect eyelid detection often results in a section of the eyelid boundary being missed. An example of such an occurrence is shown in Figure 1(b). In order to identify such errors, the eyelid region is divided into two sections, namely I and J for the upper eyelid and K and L for the lower eyelid. The mean intensity in the left and right region of the upper eyelid are recorded as μ_{ui} and μ_{uj} which would correspond to the unoccluded part below the arc. Likewise, the mean intensity above the lower eyelid are labelled μ_{lk} and μ_{ll}. Ideally, if the eyelids are properly segmented, the intensity in regions I, J, K and L should be close to that of the iris.

- We conclude this section with a description on how accuracy of the iris border is detected. Two circles, the inner and the outer circle, both with centres, (x_i, y_i), and radius $r_i - 5$ and $r_i + 5$ are drawn. In a properly segmented iris, the intensity outside the iris region (the sclera) is higher than within the iris region. The mean intensity difference between the iris region in the inner circle and that of the sclera in the outer circle for the right and left border are recoded as μ_{br} and μ_{bl} in our experiments. This implementation facilitates the detection of incorrect iris border detection as shown in Figure 1(c).

In all, 28 features were collected in this manner, i.e., 8 geometric features and 20 texture-based features.

3.2. Feature Selection

We now introduce the "composite features", each of which is a (possibly non-linear) function of multiple basic features seen in the previous section. Therefore, composite features capture higher-order interactions and contribute more information into the learning algorithm, e.g. non-linear high-order dependencies among the basic features. For example, the ratio of the radius of the iris to that of the pupil would indicate if the iris-pupil ratio makes sense according to the training images. Similarly, one would expect the intensity in the iris region to be higher than that in the pupil region. Overall, we propose 26 composite features to capture the higher-order dependencies, which are outlined in Table 1. The union of both basic and composite features amounts to 54 features, which we denote by "all features".

To reduce the chances of overfitting and possibly improve the performance (higher accuracy and lower training/prediction time), we apply a feature selection technique to all the features. The aim is to select the most important predictors and eliminate features which may be redundant i.e., features which do not provide any additional information to the learning algorithm. We make use of Lasso for feature selection. This technique was selected because it

Composite Features	Terms
c_1	r_i/r_p
c_2	r_p/r_i
c_3	$((x_p - x_i)^2 + (y_p - y_i)^2)^{1/2}$
c_4	μ_i/μ_p
c_5	$(\mu_{p1}, \mu_{p2}, \mu_{p3}, \mu_{p4})$
c_6	$(\mu_{p1}, \mu_{p2}, \mu_{p3}, \mu_{p4})/4$
c_7	$\sigma(\mu_{p1}, \mu_{p2}, \mu_{p3}, \mu_{p4})$
c_8	$(\mu_{p1}, \mu_{p2}, \mu_{p3}, \mu_{p4})$
c_9	$(\mu_{p1}, \mu_{p2}, \mu_{p3}, \mu_{p4})/4$
c_{10}	$\sigma(\mu_{i1}, \mu_{i2}, \mu_{i3}, \mu_{i4})$
c_{11}	$(\mu_{br} + \mu_{bl})$
c_{12}	$(\mu_{br} + \mu_{bl}/2$
c_{13}	$(\mu_{br} - \mu_{bl}/2$
c_{14}	$\lvert(\mu_{br} + \mu_{bl}/2\rvert$
c_{15}	$(H_{il} + H_{iu})$
c_{16}	$(H_{il} + H_{iu})/2$
c_{17}	$(H_{il} - H_{iu})$
c_{18}	$\lvert(H_{il} - H_{iu})\rvert$
c_{19}	$\mu_{ui} + \mu_{uj}$
c_{20}	$(\mu_{ui} + \mu_{uj})/2$
c_{21}	$(\mu_{ui} - \mu_{uj})$
c_{22}	$\lvert(\mu_{ui} - \mu_{uj})\rvert$
c_{23}	$\mu_{lk} + \mu_{ll}$
c_{24}	$(\mu_{lk} + \mu_{ll})/2$
c_{25}	$\mu_{lk} - \mu_{ll}$
c_{26}	$\lvert\mu_{lk} - \mu_{ll}\rvert$

Table 1: The 26 composite features and their respective terms.

is computationally feasible for high-dimensional data and is more stable than subset selection methods [16]. Lasso is an ℓ_1-regularization technique for least-square linear regression:

$$\sum_{i=1}^{N}(y_i - \mathbf{w} \cdot \mathbf{x}_i)^2 + \lambda||\mathbf{w}||_1 \quad (1)$$

where the response random variable $y \in \mathbb{R}$ is dependent on a d-dimensional covariate $\mathbf{x} \in \mathbb{R}^d$ (in our case $d = 54$; the number of all features), and $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^N$ is the training data. It is well known that the ℓ_1-regularization term shrinks many components of the solution to zero, and thus performs feature selection [17]. Lasso has identified 20 features which we refer to as "important features". They include 12 composite features and 8 basic features as follows:

$$I = [c_1, c_3, c_4, c_{10}, c_{11}, c_{14}, c_{17}, c_{18}, c_{21}, c_{22}, c_{25}, c_{26}, y_1, u_p, \sigma_i, \mu_{i3}, \mu_{br}, H_{il}, \mu_{ui}, \mu_{ll}]$$

The MSE (mean squared error) of a generalised linear model fitted based on the important features is 0.083543

which is quite close to that achieved by all features 0.080553. This confirms that the fit using only the selected features keeps most of the information contained in the fit using all features.

4. Models and Training

This section presents the models and learning algorithms. Firstly, we investigate the prediction problem using only individual eye images. The objective is to achieve high classification rate using features from an eye image. Secondly, in contrast to previous work where only still eye images have been used, we leverage neighbouring frames in an iris video for a better prediction. We make use of Support Vector Machines (SVMs) to build our classifiers for images, i.e. individual frames [5]. Figure 5.(a) shows the architecture of the model. The use of SVMs is motivated by the fact that the decision boundary may have non-linearity which can be captured by a suitable kernel in SVM. The training objective in SVM is:

$$\min_{\mathbf{w}} \frac{1}{2}||\mathbf{w}||_2^2 + C\sum_{i=1}^{N} \max_{\hat{y}\in\{-1,+1\}} \delta_{\hat{y},y_i} + \hat{y}\mathbf{x}_i \cdot \mathbf{w} - y_i\mathbf{x}_i \cdot \mathbf{w} \quad (2)$$

where $\delta_{\hat{y},y} = 1$ if $\hat{y} \neq y$, and zero otherwise. It can be shown that the optimal parameter can be written as:

$$\mathbf{w} = \sum_{i=1}^{N} \alpha_i y_i \mathbf{x}_i \quad (3)$$

which suggests that the only operation which involves data points when learning the parameters and prediction is the inner-product $\mathbf{x}_i \cdot \mathbf{x}_j$. This has led to a powerful machinery to learn non-inear decision boundaries by replacing the inner-product with a kernel function $k(\mathbf{x}_i, \mathbf{x}_j)$ [15]. The decision rule when classifying a new image $\mathbf{x}$ is:

$$f(\mathbf{x}) := \arg \max_{\hat{y}\in\{-1,+1\}} \hat{y} \sum_{i=1}^{N} \alpha_i y_i k(\mathbf{x}, \mathbf{x}_i) \quad (4)$$

Intuitively the kernel function captures the similarities between the two given data points. In our work, radial basis function (RBF) $\exp -\gamma||\mathbf{x}_i - \mathbf{x}_j||_2^2$ has resulted in excellent performance as the kernel between segmented images.

Secondly, we leverage the temporal information in iris videos for classification. The iris videos have been captured at a rate of 30 frames per second. This suggests that the change from one frame to another is gradual, hence the classification label should be *consistent* across neighbouring images. For example, in the case of a properly segmented frame, its neighbouring frames, i.e., the ones before and after it, are most probably successfully segmented as well. This has led us to take into consideration the neighbouring frames in iris videos when classifying a segmented iris image. Figure 5.(b) shows our *two-layer* model, i.e. the *deep architecture*. The output of the classifier in the first layer H_1 for the neighbouring images is fed as input to the classifier in the second layer denoted by H_2. The decision rule when classifying a new image $\mathbf{x}^{(t)}$ is:

$$g(\mathbf{x}^{(t)}) := \arg \max_{\hat{y}\in\{-1,+1\}} \hat{y} \sum_{i=1}^{N} \beta_i y_i k'(F(\mathbf{x}^{(t)}), F(\mathbf{x}_i^{(t_i)})) \quad (5)$$

where $F(\mathbf{x}^{(t)}) = [f(\mathbf{x}^{(t-\ell)}), \ldots, f(\mathbf{x}^{(t)}), \ldots, f(\mathbf{x}^{t+\ell})]$ for a neighbourhood window of size ℓ and f is defined in Eq (4). We make use of the greedy layer-wise training method [2]: To train H_1 using our first approach for images and then fix the learned H_1 while training H_2.

5. Experimental Results

Our proposed methods are evaluated using a dataset of near infrared (NIR) iris videos [1]. We first describe the experimental setup and implementation details. Afterwards we present the empirical results, comparing our methods with strong baselines.

5.1. Experimental Setup

We use the dataset provided for the Multiple Biometric Grand Challenge (MBGC) Version 2.0. It consists of 986 NIR iris video sequences acquired using an LG 2200 camera. The dataset consists of videos of both right and left eyes. In this study, the iris regions were segmented using the segmentation algorithm developed for iris videos in less constrained environments [13]. The ground truth dataset refers to correct (+ve) and incorrect (-ve) segmentations determined by visual inspection.

5.2. Evaluation

We compare SVM with several strong baselines. In line with previous work carried out in this area [9], the performance of the proposed model is evaluated against Naive-Bayes Tree (NBTree) classifier. We also compare against Decision Tree (DT) classifier, which is more powerful than NBTree in terms of the induced decision boundary. We start by presenting our results on learning classifiers using features from *individual* frames (images). Next, we evaluate our proposed deep architecture leveraging *neighbouring* frames in iris videos. Finally, we compare the performance of correct segmentation vs. incorrect segmentation in an end-to-end iris recognition system.

5.3. Intrinsic Evaluation: Individual Iris Images

This section investigates the prediction accuracy using individual iris images. We use 4775 individual segmented iris frames from the MBGC dataset which includes 3439 +ve samples (correct segmentation) and 1336 -ve samples (incorrect segmentation). As shown in Figure 5.(a), the features are fed to the learning algorithm to build the classifier H_1. The four types of features used in our experiments

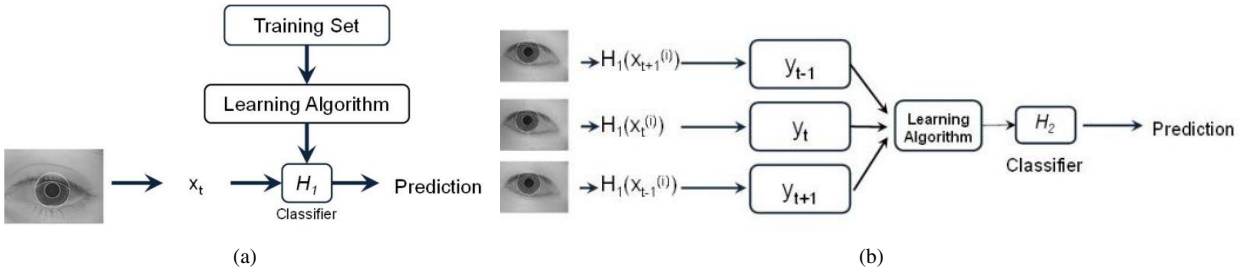


Figure 5: Prediction of segmentation error using: (a) Iris images (b) Neighbouring frames in an iris video.

are the basic features, composite features, all features (basic and composite), and important features (selected by the Lasso technique) as described in Section 3. We experiment with the Linear and RBF kernels in SVM.

Figure 6 and Table 2 represent the results of various learning models using different feature sets. SVM-Linear shows the lowest prediction accuracy compared to the other learning models whereas SVM-RBF performs better than all learning models. As shown in Table 2, the combination of SVM-RBF and "important features" achieves the highest average prediction accuracy of 97.7 ± 0.14 % using 10-fold cross validation. SVM-RBF is more powerful than SVM-Linear in terms of capturing non-linear decision boundaries

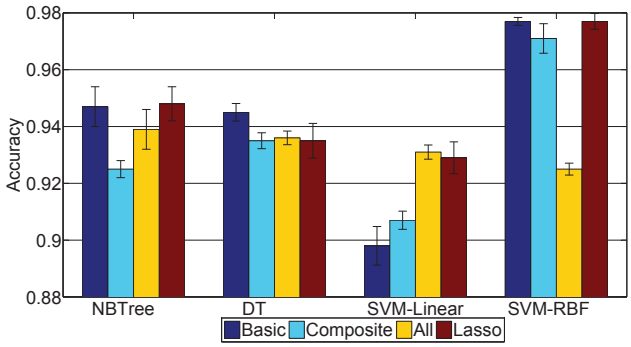


Figure 6: Prediction of segmentation accuracy using iris images. Each bar represent the types of features used with four learning algorithms.

Features	NBTree	DT	SVM-Linear	SVM-RBF
Basic	0.947	0.939	0.895	0.973
Composite	0.925	0.926	0.902	0.971
All	0.939	0.943	0.931	0.925
Important	0.948	0.936	0.929	**0.977**

Table 2: Average ten-fold cross-validation accuracies for iris images using different feature sets and learning algorithms.

exhibited by our data (Figure 7), however, this comes at the cost of more training time.

5.4. Intrinsic Evaluation: Neighbouring Iris Frames

This section examines the prediction accuracy using neighbouring iris frames in an iris video. We experiment with two choices for the window size for neighbouring frames: (i) The window size of 1 in which we only consider the previous f_{t-1} and successive f_{t+1} frames of a frame of interest f_t (denoted by *1-f-1*), and (ii) the window size of 2 in which we consider the frames $f_{t-2}, f_{t-1}, f_t, f_{t+1}, f_{t+2}$ (denoted by *2-f-2*). In the *1-f-1* setting, overall 11826 iris frames are used including 2785 -ve samples and 9041 +ve samples. In the *2-f-2* setting 17788 samples are considered including 13852 positive samples and 3936 -ve samples. Similar to the previous section, we evaluate the proposed settings using the 4 learning models for training H_2, while H_1 is fixed to the best model for the previous section (i.e. SVM-RBF with 'important features'). Figure 8 and Table 3 represent the results, i.e. the average of 10-fold cross-validation prediction accuracy and 95% confidence interval.

The neighbouring frames approach contributes a further

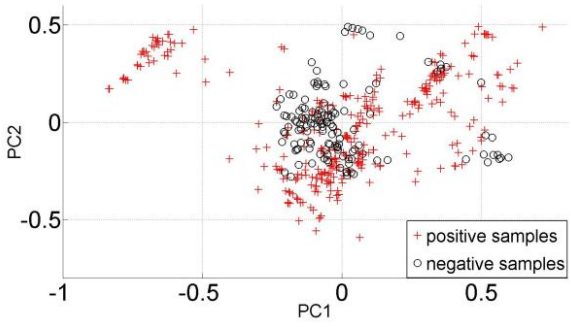


Figure 7: PCA plot of "important features", (344 +ve samples and 134 -ve samples) on two principal components PC1 and PC2. The projection into 2-dimensional space confirms the classes may not be linearly separable.

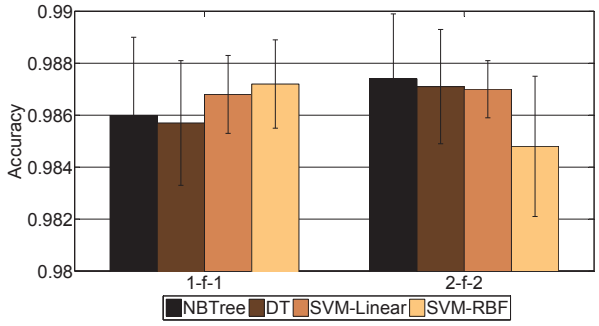


Figure 8: Prediction of segmentation accuracy using one (1-f-1) and two (2-f-2) neighbouring frames in iris videos. Each bar represent the types the learning algorithm used for prediction.

Algorithm	1-f-1	2-f-2
NBTree	0.986 ± 0.003	0.987 ± 0.003
DT	0.986 ± 0.002	0.987 ± 0.002
SVM-Linear	0.987 ± 0.002	0.987 ± 0.001
SVM-RBF	0.987 ± 0.002	0.985 ± 0.003
SVM-RBF (Images)	0.977 ± 0.001	

Table 3: Average ten-fold cross-validation accuracies using neighbouring frames. The best performance using images is also included for comparison.

1% improvement over the first method where iris frames are evaluated individually. No significant change in performance is observed among different learning model. In both the *1-f-1* and *2-f-2* combinations of neighbouring frames, SVM-Linear achieves the best prediction accuracies with the shortest 95% confidence interval, i.e., 98.7 $\pm$0.17% and 98.7 $\pm$ 0.11% respectively. Based on our experimental results, this would suggest that using the *1-f-1* arrangement is sufficient for the classification of iris segmentation errors.

5.5. Extrinsic Evaluation: The End-to-End Pipeline

We investigate the effect of correctly and incorrectly classified iris segmented images on the end-to-end recognition system. In order to ensure that the segmentation class is the main factor impacting the performance, we perform the experiment in a carefully controlled condition as follows. Firstly, only high quality frames are used in the experiment. This ensures that the frames used in our experiments are of high quality and consequently a drop in recognition performance cannot be attributed to poor image quality factors. The frames are obtained by implementing an optimal frame selection model as described in [12]. Secondly, the intra-class comparison is only carried out on frames which belong to the same video. This would involve comparisons which include neighbouring frames and frames captured in the same environment; therefore, environmental factors such as varying illumination are eliminated.

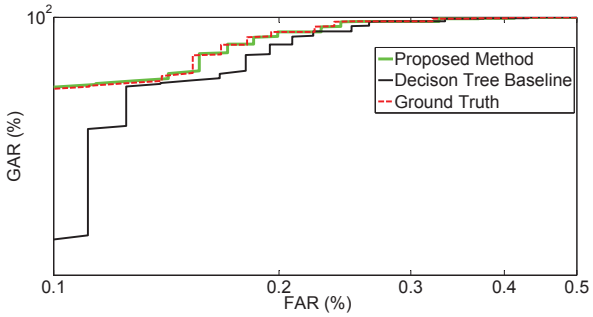


Figure 9: ROC curve on the log-log scale which compares the performance of +ve samples of the proposed method with DT and the baseline.

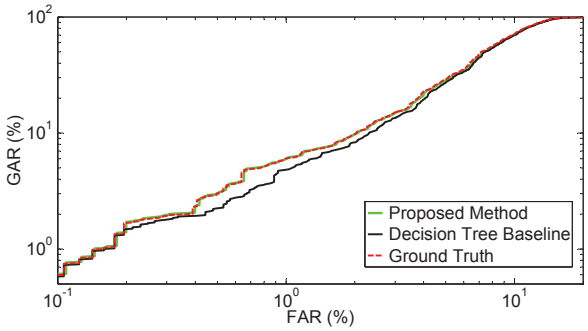


Figure 10: ROC curve on the log-log scale which compares the performance of -ve samples of the proposed method with DT and the baseline.

We use the same approach used in a conventional iris recognition system in our experiment, i.e., all segmented frames are normalized to a rubber-sheet model and the iris codes are obtained using Log-Gabor wavelets [6]. Hamming Distance (HD) is then used to evaluate the amount of dissimilarity between iris codes. Figure 9 displays the ROC curves for the correct (+ve samples) segmentation where GAR is the Genuine Accept Rate and FAR is the False Accept Rate. Here the "Ground Truth" refers to correctly segmented images which have been selected by visual inspection, which gives the upper-bound on the achievable performance. As expected, the Ground Truth curve achieves very high performance as illustrated by its high GAR and low FAR. The ROC curve where +ve samples are selected using the proposed method closely follows the Ground Truth and indicates that the prediction rate is accurate and close to the best achievable performance. On the other hand, the ROC curve using DT for selection is away from both the Ground

Truth and the proposed method showing lower prediction accuracy. Figure 10 displays the ROC curves for the incorrect (-ve samples) segmentation. It can be observed that segmentation accuracy affects recognition performance and plays a key role in an iris recognition system.

6. Discussion

In this section, we highlight some of the key points of the proposed model. Firstly, based on the recent adoption of video-based technology in iris recognition systems, the task of classifying segmentation errors is highly important. To the best of our knowledge, this is the first work where temporal correspondence in iris frames is exploited to enhance performance. In contrast, previous work has only made use of still-eye images. Secondly, the "important features" selected by Lasso play a significant role. They highlight the features which one should focus on when evaluating the segmentation. The classifier uses these features to determine if the iris is properly segmented before sending the iris image to other stages in the pipeline. The high prediction accuracy obtained using the important features improves the overall performance of the system and reduces the need of iris image re-capture for the user. Finally, the features can be collected during the segmentation task itself as opposed to previous methods which require segmentation to be carried out first.

7. Conclusion

In this work, we have presented two fully automated models for the classification of iris segmentation. An extended set of features consisting of basic and composite features collected during the segmentation stage have been investigated in our experiments. This is followed by the selection of an important set of features using the Lasso regularization technique. We have investigated various learning algorithms and models, taking into account both individual iris frames and neighbouring frames in videos for this task. As shown in our experiments, accurate segmentation heavily influences the performance of iris recognition systems. Our proposed models for individual images and videos achieve accuracy rates of 97.7% and 98.7% for predicting iris segmentation accuracy. Our technique should be useful for making iris recognition systems more reliable.

References

[1] NIST: Multiple Biometric Grand Challenge (MBGC) V2.0. http://face.nist.gov/mbgc, July 2009.

[2] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle. Greedy layer-wise training of deep networks. In *Advances in Neural Information Processing Systems 19, Proceedings of the Twentieth Annual Conference on Neural Information Processing Systems, Vancouver, British Columbia, Canada, December 4-7, 2006*, pages 153–160, 2006.

[3] K. W. Bowyer, K. P. Hollingsworth, and P. J. Flynn. A survey of iris biometrics research: 2008–2010. In *Handbook of iris recognition*, pages 15–54. Springer London, 2013.

[4] T. Camus and R. Wildes. Reliable and fast eye finding in close-up images. In *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, volume 1, pages 389–394 vol.1, 2002.

[5] C. Cortes and V. Vapnik. Support-vector networks. *Mach. Learn.*, 20(3):273–297, Sept. 1995.

[6] J. Daugman. How iris recognition works. *IEEE Transactions on Circuits and Systems for Video Technology*, 14:21–30, 2002.

[7] K. Hollingsworth, T. Peters, K. W. Bowyer, and P. J. Flynn. Iris recognition using signal-level fusion of frames from video. *Information Forensics and Security, IEEE Transactions on*, 4(4):837–848, 2009.

[8] R. Jillela, A. Ross, N. Boddeti, B. Vijaya Kumar, X. Hu, R. Plemmons, and P. Pauca. An evaluation of iris segmentation algorithms in challenging periocular images. *Handbook of Iris Recognition, Springer, Heidelberg (to appear, 2012)*, 2002.

[9] N. Kalka, N. Bartlow, and B. Cukic. An automated method for predicting iris segmentation failures. In *Biometrics: Theory, Applications, and Systems, 2009. BTAS'09. IEEE 3rd International Conference on*, pages 1–8. IEEE, 2009.

[10] R. Kohavi. Scaling up the accuracy of naive-bayes classifiers: A decision-tree hybrid. In *KDD*, pages 202–207, 1996.

[11] N. Mahadeo and N. Bhattacharjee. An efficient and accurate iris segmentation technique. In *Digital Image Computing: Techniques and Applications, 2009. DICTA '09.*, pages 347–352, Dec 2009.

[12] N. K. Mahadeo, A. P. Paplinski, and S. Ray. Automated selection of optimal frames in nir iris videos. In *Digital Image Computing: Techniques and Applications (DICTA), 2013 International Conference on*, pages 1–8. IEEE, 2013.

[13] N. K. Mahadeo, A. P. Paplinski, and S. Ray. Robust video based iris segmentation system in less constrained environments. In *Digital Image Computing: Techniques and Applications (DICTA), 2013 International Conference on*, pages 1–8. IEEE, 2013.

[14] N. K. Mahadeo, A. P. Paplinski, and S. Ray. Optimization of iris codes for improved recognition. In *Computer Vision and Pattern Recognition (CVPR) Workshops, 2014, IEEE Computer Society Conference on*, pages 1–8. IEEE, 2014.

[15] K.-R. Mller, S. Mika, G. Rtsch, S. Tsuda, and B. Schlkopf. An introduction to kernel-based learning algorithms. *IEEE Transactions on Neural Networks*, 12(2):181–202, 2001.

[16] R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society (Series B)*, 58:267–288, 1996.

[17] P. Zhao and B. Yu. On model selection consistency of lasso. *J. Mach. Learn. Res.*, 7:2541–2563, Dec. 2006.

[18] J. Zuo and N. A. Schmid. An automatic algorithm for evaluating the precision of iris segmentation. In *Biometrics: Theory, Applications and Systems, 2008. BTAS 2008. 2nd IEEE International Conference on*, pages 1–6. IEEE, 2008.