Biometric System 25/26

6D Segmentation-Free
Iris Localization For A Daugman
Pipeline

Dalla Ragione Ivan 2009768

Contents

1 Abstract

This project presents a biometric system based on iris recognition, inspired by recent segmentation-free approaches for iris localization. The proposed system addresses the iris recognition scenario by revisiting the classical Daugman pipeline and replacing the segmentation stage with a regression-based estimation of six geometric parameters describing the iris and pupil boundaries.

The pipeline relies on two deep learning models, namely an Iris Localization Network (ILN) and a Pupil Refinement Network (PRN), followed by a Daugman-style normalization and iris code generation. The resulting templates are evaluated in both verification and identification scenarios, including closed-set and open-set identification, using standard biometric performance metrics.

Experimental results show that the proposed segmentation-free approach is able to achieve reliable performance in verification and closed-set identification, while open-set identification highlights the expected trade-off between recognition accuracy and rejection capability.

Future improvements may focus on refining the iris boundary estimation in addition to pupil localization, as well as on improving robustness to challenging acquisition conditions, with the goal of enhancing verification and identification performance.

2 Introduction

Iris recognition is one of the most reliable biometric traits due to its high discriminative power and long-term stability. The iris texture exhibits complex and highly distinctive patterns that are formed early in life and remain largely unchanged over time. Compared to other biometric modalities, the iris offers a favorable balance between uniqueness, permanence, and compact template representation, making it suitable for both verification and identification tasks.

Despite these advantages, iris recognition presents several practical challenges, mainly related to image acquisition. The iris is a relatively small anatomical structure, and reliable acquisition typically requires short capture distances and sufficient spatial resolution. Variations in illumination, gaze direction, motion blur, specular reflections, as well as occlusions caused by eyelids and eyelashes, can significantly degrade image quality. For this reason, many operational systems rely on near-infrared illumination to enhance iris texture visibility and reduce noise.

The classical iris recognition pipeline was introduced by Daugman [1] and consists of four main stages: segmentation, normalization, feature extraction, and matching. In this framework, segmentation plays a critical role, as it aims to accurately localize the pupil and iris boundaries while excluding non-iris regions. However, iris segmentation is widely recognized as the most error-prone stage of the pipeline. Accurate segmentation often requires complex image processing techniques or pixel-wise annotated data, leading to high computational costs and limited robustness under unconstrained acquisition conditions. Moreover, segmentation errors tend to propagate through normalization and feature extraction, negatively affecting the final recognition performance.

Motivated by these limitations, recent research has explored segmentation-free approaches, where the geometric parameters of the iris and pupil are directly regressed from the input image, avoiding the generation of dense pixel-level segmentation maps and the associated post-processing stages In this project, the adopted approach is inspired by the segmentation-free framework proposed in [2], which replaces pixel-level segmentation with a regression-based estimation of iris geometry. This strategy aims to reduce computational complexity and annotation requirements while preserving compatibility with the classical Daugman normalization and encoding stages.

2

In particular, the system implemented in this work relies on the regression of a reduced set of six parameters describing the pupil and iris boundaries. Different from the full configuration proposed in [**?**], only the 6D representation is considered, and additional geometric cues such as eyelid landmarks are not used. This design choice simplifies the learning task and reduces model complexity, but also introduces more restrictive assumptions in the subsequent stages of the pipeline. As a consequence, specific constraints and filtering strategies are adopted during template generation and matching to compensate for the reduced geometric information.

The resulting system combines a segmentation-free localization stage with a Daugman-style normalization and iris code extraction, and is evaluated on verification and identification tasks, including both closed-set and open-set scenarios.

This setup allows for a direct analysis of the impact of segmentation-free geometry estimation on classical iris recognition pipelines, highlighting both its strengths and its current limitations.
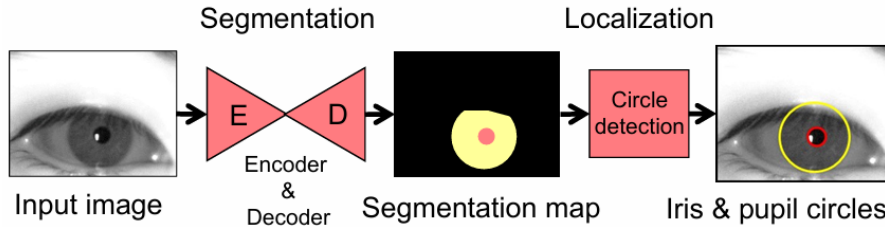

Figure 1: Comparison between segmentation-free and conventional iris localization pipelines.

3   Dataset

In this work, I designed and evaluated a segmentation-free iris recognition pipeline on the **CASIA-IrisV3 Interval**[3] dataset, (accessed in this work through the Kaggle mirror [4]), a well-known benchmark for iris biometrics acquired under near-infrared (NIR) illumination. The dataset is organized by subject identity and provides multiple samples per subject, allowing both training (for localization/refinement networks) and verification/identification experiments.

3.1   Dataset Structure

The downloaded dataset [4] contains 2639 NIR iris images collected from 249 unique subjects. Samples are organized by subject identity and separated by eye side (*Left / Right*), with a filename convention of the form:

```
S<subject><L/R><index>.jpg.
```

Overall, the dataset provides on average 10.6 images per subject (median: 10), with a minimum of 1 and a maximum of 40 samples per subject. Considering the eye side, the distribution is approximately balanced, with an average of about 5-6 images per subject for each eye.

Table 1: CASIA-IrisV3-Interval dataset statistics.

Statistic	Value
Total images	2639
Subjects (identities)	249
Average samples / subject	10.6
Median samples / subject	10
Min samples / subject	1
Max samples / subject	40

3.2   Acquisition Challenges and Data Difficulties

Although CASIA-Interval is acquired in controlled indoor conditions, some images still present several challenges that can affect both geometric estimation
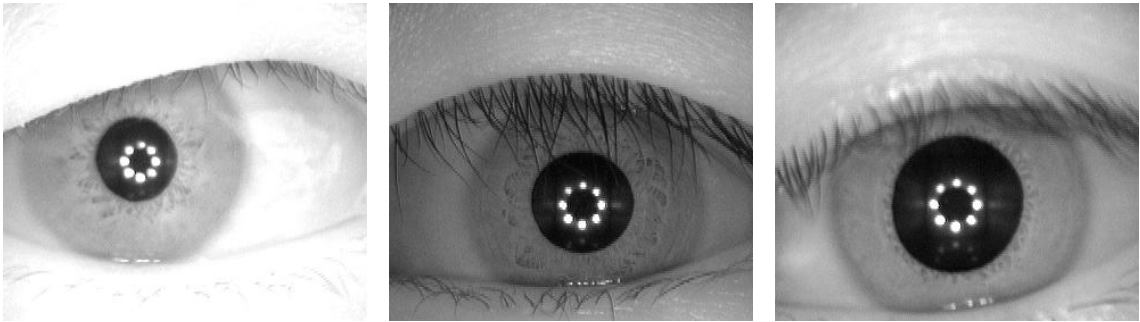
4

and recognition performance.

A common issue is the presence of specular highlights on the corneal surface, producing saturated regions that reduce the amount of reliable iris texture and may introduce noisy bits in the final template.(Fig. 2a)

Moreover, eyelid and eyelash occlusions frequently cover the upper/lower iris boundaries, reducing the effective visible area and making accurate circle fitting harder, especially in a segmentation-free setting where explicit eyelid keypoints are not available.(Fig. 2b)

Additional problems to face include blur and low-contrast boundaries, which degrade the precision of pupil/iris edge localization, which can introduce deviations from an ideal circular shape and negatively impact normalization. (Fig. 2c)

These factors motivate the use of robust preprocessing, masking strategies, and a refinement stage to improve stability under real acquisition noise.

(a) Strong specular highlight.   (b) Occlusion / eyelashes.   (c) Low contrast regions.

Figure 2: Examples of challenging samples from CASIA-IrisV3-Interval.

4 Models Architecture

The proposed system adopts a two-stage localization strategy based on cascaded deep learning models. The first network (ILN) performs a global estimation of pupil and iris geometry, while a second network (PRN) refines the pupil parameters through a localized regression step. At inference time, the models operate sequentially, enabling robust and accurate geometric estimation without relying on explicit segmentation masks.

4.1 Iris Localization Network ILN

The ILN is a convolutional neural network that aims to estimate the pupil and iris geometry without producing a segmentation mask.

The model predicts six parameters:

$$\hat{\mathbf{y}} = [\hat{x}_p, \hat{y}_p, \hat{r}_p, \ \hat{x}_i, \hat{y}_i, \hat{r}_i] \in R^6, \tag{1}$$

- (x_p, y_p, r_p) identify center and radius of the pupil

- (x_i, y_i, r_i) identify center and radius of the iris

Input. The ILN receives as input images *grayscale* with $H \times W = 480 \times 640$ resolution (represented as a tensor $(B, 1, H, W)$). The geometrical parameters are represented in normalized form:

$$\mathbf{y} = \left[\frac{x_p}{W}, \frac{y_p}{H}, \frac{r_p}{W}, \ \frac{x_i}{W}, \frac{y_i}{H}, \frac{r_i}{W} \right]. \tag{2}$$

The model will predict $\hat{\mathbf{y}}$ in the same normalized space.

Architecture. The ILN follows a standard CNN regressor design, composed of a convolutional backbone for multi-scale feature extraction and a final regression head. Feature maps are aggregated through Global Average Pooling and mapped by fully-connected layers to a 6D continuous output, corresponding to pupil and iris circle parameters $(x_p, y_p, r_p, x_i, y_i, r_i)$.

Loss function. The training of the ILN is set as a supervisioned regression using a Weighted L1 Loss defined on six parameters:

$$\mathcal{L}_{ILN}(\hat{\mathbf{y}}, \mathbf{y}) = \frac{1}{6} \sum_{k=1}^{6} w_k \ |\hat{y}_k - y_k|, \tag{3}$$

where w is a vector of weights that emphasizes pupil and iris radius.

$$\mathbf{w} = [1, 1, 2,\ 1, 1, 2]. \tag{4}$$

The weighs are based on the fact that errors performed on radius influence significantly the normalization step, so the general quality of the biometric template.

Training Setup. The ILN is trained using the Adam optimizer with a fixed initial learning rate. To automatically adapt the learning rate during training, we employ a ReduceLROnPlateau scheduler that monitors the *validation loss*: whenever the validation performance stagnates for a predefined number of epochs (*patience*), the learning rate is reduced by a constant factor to encourage further convergence. To mitigate overfitting and improve generalization, we also apply an early stopping strategy based on the validation loss, saving the model checkpoint corresponding to the best validation score and restoring it at the end of training.

4.2   Pupil Refinement Network PRN

The Pupil Refinement Network (PRN) is a lightweight convolutional neural network designed to refine the pupil geometry previously estimated by the ILN. While the ILN provides a global estimation of pupil and iris parameters on the full-resolution image, the PRN operates locally on a region of interest, centered around the pupil, enabling a more accurate and stable estimation of pupil parameters.

The model predicts three parameters:

$$\hat{\mathbf{z}} = [\hat{x}_p, \hat{y}_p, \hat{r}_p] \in R^3, \tag{5}$$

where (x_p, y_p, r_p) represent the center coordinates and radius of the pupil.

Input. The PRN receives as input a grayscale cropped ROI of fixed size 128×128, represented as a tensor $(B, 1, 128, 128)$. The pupil parameters are expressed in normalized form with respect to the ROI size:

$$\mathbf{z} = \left[\frac{x_p}{128}, \frac{y_p}{128}, \frac{r_p}{128} \right]. \tag{6}$$

The network predicts $\hat{z}$ in the same normalized space.

Architecture. The PRN follows a compact CNN regressor architecture, composed of a convolutional backbone with progressive spatial downsampling and a regression head based on Global Average Pooling followed by fully-connected layers. Compared to the ILN, the PRN is shallower and computationally lighter, as it focuses exclusively on local pupil refinement rather than global eye geometry estimation.

Loss function. The training of the PRN is formulated as a supervised regression task using a Weighted L1 Loss defined on the three pupil parameters:

$$\mathcal{L}_{PRN}(\hat{\mathbf{z}}, \mathbf{z}) = \frac{1}{3} \sum_{k=1}^{3} w_k \, |\hat{z}_k - z_k| \,. \tag{7}$$

The weighting scheme emphasizes the pupil radius, as inaccuracies in this parameter directly affect the subsequent iris normalization and template extraction stages.

Training Setup. The PRN is trained combining the Adam optimizer and a validation checkpoints strategy. It uses the pupil estimation provided by the ILN to define the ROI, while the supervision target remains the ground-truth pupil parameters remapped into the ROI coordinate system. A controlled geometric jitter is applied to the ROI extraction process to simulate ILN localization errors. The model with the lowest validation loss is saved and restored at the end of the optimization process.
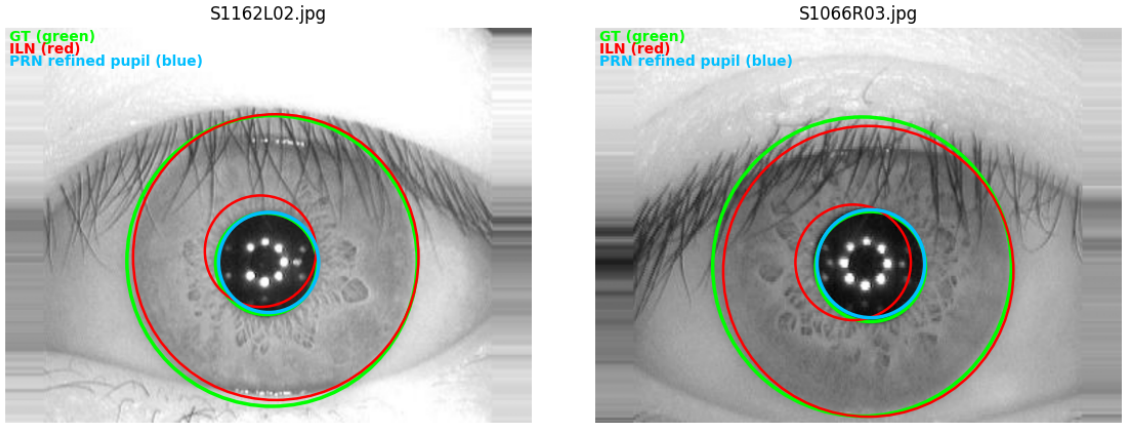

Figure 3: Qualitative examples of pupil refinement performed by PRN which improves alignment with the true pupil boundary (GT- circle) under challenging acquisition conditions.

5 Pipeline

The core idea of the project is to design and develop a Daugman-style pipeline, replacing the canonical "Segmentation" with the direct regression performed by the previously described model. The pipeline involves the following stages :

1. Input and Normalization

2. Localization (ILN + PRN)

3. Geometric Validation + Quality Gating

4. Rubber Sheet Normalization

5. IQR Masking + Column Gating

6. Iris Code

As explained before, due to the absence of the eyelid points described by the paper, my pipeline has to replace the geometric eyelid mask with a normalization-domain masking strategy based on robust statistics and angular consistency. This design choice preserves a fully segmentation-free pipeline while approximating the occlusion rejection capability of landmark-based eyelid modeling.

5.1 Input and Normalization

In iris recognition systems, especially when localization is performed via regression rather than segmentation, the input representation must be strictly controlled. Variations in resolution, intensity range, or color space introduce non-informative variance that propagates through the entire pipeline and degrades both localization accuracy and template stability.

Normalizing the input to a fixed spatial resolution and intensity range ensures that the downstream models (ILN/PRN) operate in the same domain used during training, improving robustness and reproducibility. Each input image is converted to grayscale, resized to a fixed resolution, and normalized to the [0,1] range. The output is a single-channel tensor compatible with the localization models.

5.2 Localization

Classical Daugman pipelines rely on accurate detection of pupil and iris boundaries. Errors at this stage directly affect the rubber sheet normalization, leading to misaligned texture sampling and unstable templates which can inflate the entire pipeline performances.

Instead of dense segmentation, this pipeline adopts a segmentation-free, regression-based localization strategy. The models used are previously described.

5.3 Geometric Validation and Quality Gating

The estimated iris and pupil are validated using some geometric constraints, verifying the radius range, pupil-iris size and distance between the two centers. Samples that violate these constraints are discarded and do not contribute to template generation. The rubber sheet model implicitly assumes a physically consistent iris geometry: the pupil must lie inside the iris, the two circles must be reasonably concentric. When these assumptions are violated, the resulting normalization produces distorted or incomplete iris strips, leading to unstable IrisCodes and unreliable similarity scores.

5.4 Rubber Sheet Normalization

The normalization stage follows a canonical Daugman rubber sheet model in which the annular iris region between the pupil and iris boundaries is re-

parameterized into a fixed-size polar representation by sampling image intensities along radial directions. Different from the naive normalization, an iris shrink factor of $\beta = 0.85$ is adopted to conservatively restrict the sampled outer iris region, increasing robustness to boundary noise and occlusions, particularly in the absence of explicit eyelid landmarks.
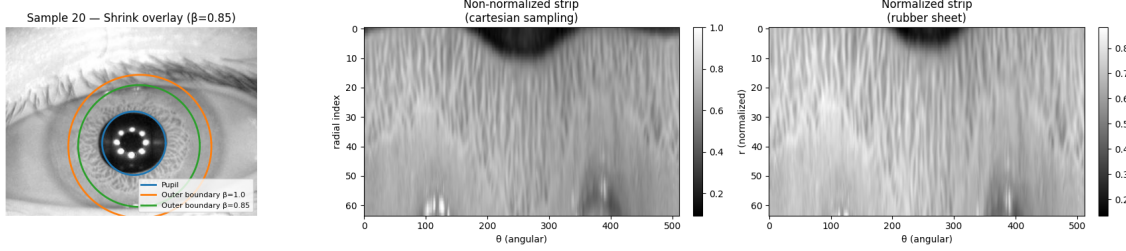

Figure 4: Rubber sheet normalization process with iris shrink. **Left**: estimated pupil boundary and effective outer sampling boundaries, showing the effect of the shrink factor ($\beta = 0.85$) with respect to the full iris contour. **Center**: non-normalized iris strip. **Right**: normalized iris strip, where the iris texture is re-parameterized into a fixed-size polar representation.

5.5 Interquartile Range Masking and Column Gating

After rubber sheet normalization, I built a segmentation-free quality mask directly in the normalized iris strip in order to suppress occlusions and unreliable texture before IrisCode extraction.

First, an IQR-based anomaly detection step is applied to the normalized iris strip. For each angular column, the first and third quartiles (Q1,Q3) of the intensity distribution along the radial direction are estimated to define a reliable intensity range. Pixels whose values fall outside this range are considered unreliable and are therefore masked out, effectively suppressing specular reflections, eyelashes, and other local occlusions.

In my implementation, IQR is computed in a *per-column* approach rather than a globally one, to better capture sector-wise degradations in the rubber sheet domain. This is motivated by the absence of eyelid landmarks (P1–P8) available in the paper, which would otherwise provide an explicit geometric eyelid mask. At the end, angular stability is improved via column gating: for each column we measure the valid-pixel ratio after masking and discard entire angular columns whose ratio is below a threshold; short contiguous runs of surviving columns are also removed to prevent intermittent, noisy sectors from contributing to the IrisCode.

5.6 Iris Code

After obtaining the mask, the goal is to obtain a binary representation of the Iris, that will be used in matching task. Following the Daugman style, the mask is built applying a Gabor filter. Each radial row is filtered along the angular direction using circular 1D convolutions, reflecting the periodic nature of the polar representation. The binary template is then obtained by quantizing the phase information through the sign of the real and imaginary filter responses, producing two bits per scale. To improve robustness, a multi-scale configuration is adopted by using multiple Gabor frequencies, and the quality mask is propagated to the bit level so that unreliable regions do not contribute to matching.
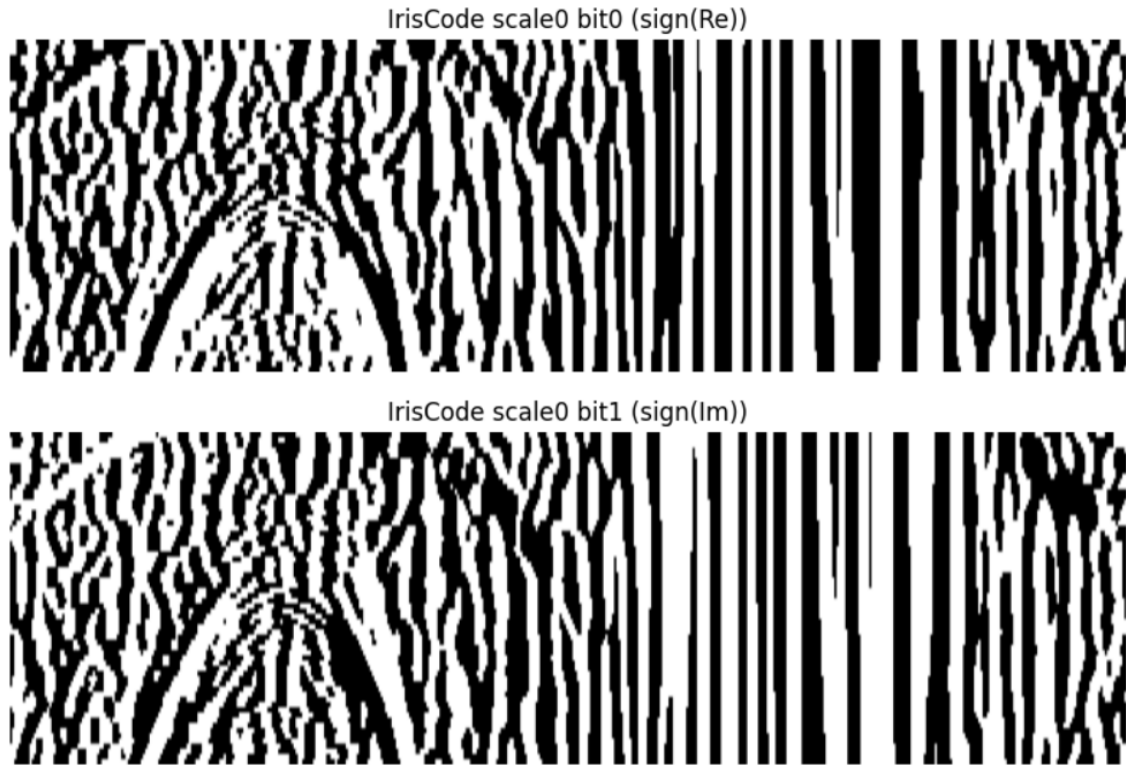

Figure 5: Example of IrisCode generation using a Daugman-like encoding with two frequencies

6 Evaluation

The work has been evaluated on three different tasks: verification, closed-set identification, and open-set identification. At the end of the previously explained pipeline, a list of templates is created, where each template contains the subject ID, the filename, the iris code, the corresponding mask, and the number of valid bits. Before starting the evaluation, some templates are discarded based on the number of valid bits and on geometric constraints limiting how far the iris center can drift from the pupil, in order to remove obvious failures without affecting valid samples. The resulting list of valid templates is the same for each task.

6.1 Verification

Verification is the first analyzed task and is based on a 1:1 comparison. As a first step, two different lists of pairs are built:

- **Genuine pairs**: templates are grouped by subject ID and only identities with $n \geq 2$ templates are considered, taking care not to pair left and right eyes together.

- **Impostor pairs**: templates are grouped by subject ID and pairs are formed by selecting samples belonging to different subjects.

The distance between the templates is now calculated using the Hamming Distance.

6.1.1 Masked Hamming distance.

Each iris template is represented by a binary IrisCode C and an associated binary validity mask M, both defined in the normalized polar domain. Template comparison is performed using a masked Hamming distance, which accounts for occlusions and unreliable regions by considering only the bits that are valid in both templates. Given two templates (C_1, M_1) and (C_2, M_2), the combined mask is computed as $M_{12} = M_1 \wedge M_2$, and the distance is defined as

$$HD = \frac{\sum (C_1 \oplus C_2) \wedge M_{12}}{\sum M_{12}}.$$

To compensate for eye rotation, the comparison is repeated for multiple circular shifts along the angular dimension, and the minimum distance across shifts is retained as the final matching score. Comparisons with an insufficient number of valid bits are discarded.

6.1.2 Metrics

After defining the matching distance, a verification decision is taken according to a distance-based rule: a comparison is accepted if $HD \leq t$, where t denotes the decision threshold otherwise, it is rejected. The evaluation is conducted on a set of 1871 iris templates, obtained by applying relatively strict quality constraints in the template generation stage, prioritizing template reliability over coverage. From this set, 2000 genuine pairs and 2000 impostor pairs are randomly generated for performance evaluation.
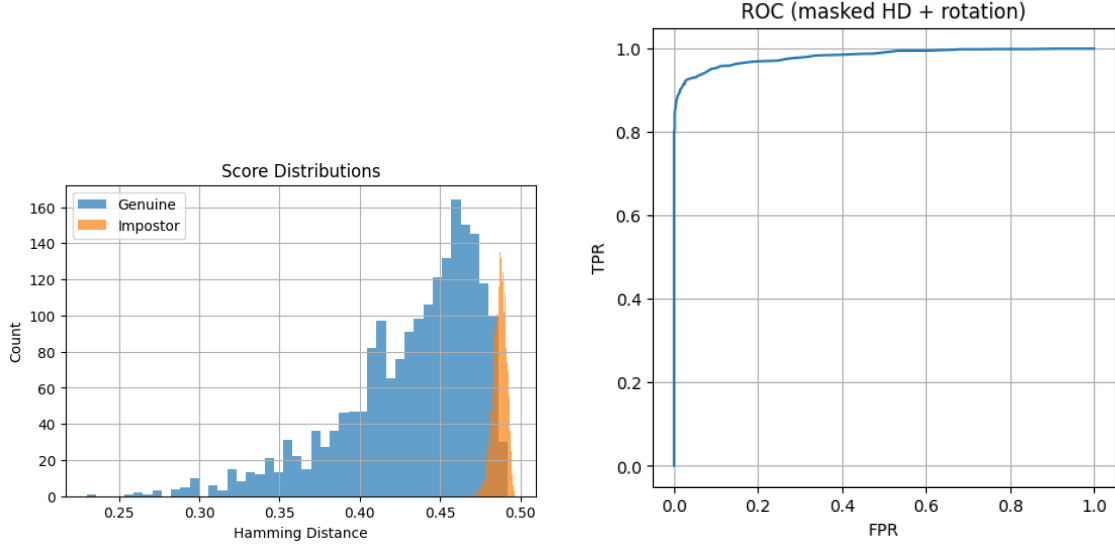
The decision threshold is set to $t = 0.4806$ according to the Equal Error Rate (EER) criterion, corresponding to the operating point at which the False Accept Rate (FAR) and the False Reject Rate (FRR) are approximately equal.

The resulting verification performance metrics are summarized in Table 2. In particular, the system achieves an Area Under the ROC Curve (AUC) of 0.982, indicating strong global separability between genuine and impostor comparisons. At the EER operating point, the system shows a FAR of 6.35% and an FRR of 6.25%. As shown in Fig. 6, these error rates are explained by a limited but non-negligible overlap between genuine and impostor score distributions in the critical decision region.

In addition to the EER operating point, verification performance is also evaluated at fixed FAR values to simulate more realistic security constraints, as reported in Table 3. When the FAR is constrained to 1%, the system reaches an FRR of 10.95%, indicating a moderate loss in usability in exchange for improved security. Under a stricter constraint (FAR = 0.1%), the FRR increases to 19.4%, showing a significant degradation in genuine acceptance. These operating points confirm the sensitivity of the system to threshold selection and highlight the trade-off between security and usability in the critical decision region.

Figure 6: Verification results

(a) Genuine and impostor score distributions.

(b) ROC curve obtained.

Table 2: Verification performances.

Metric	Value
Number of templates	1871
Identities (subject_eye)	371
Identities with ≥ 2 templates	336
Genuine pairs	2000
Impostor pairs	2000
AUC	0.982
EER	6.30%
EER threshold (t_{EER})	0.4806
FAR at t_{EER}	6.35%
FRR at t_{EER}	6.25%

Table 3: Verification operating points at fixed FAR values.

Target FAR	Threshold t	FRR	GAR
0.1%	0.468954	19.40%	80.60%
1.0%	0.475776	10.95%	89.05%
EER	0.480583	6.25%	93.75%

15

6.2 Identification: Closed Set

Closed-set identification is performed with a 1:N match, where the probe is entirely contained in the dataset; therefore, it is not possible to obtain an unknown result. The matching templates are ordered according to their distance and classified with a rank.

In particular, the templates are grouped by subject ID, and for each subject, the template with the highest number of valid bits is selected and placed in the gallery, while the remaining templates are used as probes. Then, due to the absence of a GPU and in order to reduce the computational cost of the 1:N matching process, a two-stage matching strategy is adopted:

Top-K Candidate Selection Given a probe template P, a fast pre-selection stage computes an approximate distance $d_{\text{cheap}}(P, G_j)$ against all gallery identities G_j, and selects the Top-K candidate identities:

$$\mathcal{C}_K(P) = \arg \operatorname*{topK}_{j \in \mathcal{G}} \ d_{\text{cheap}}(P, G_j) \tag{8}$$

The approximation relies on a reduced representation of the iris template, using a limited number of scales, rotation range, and a coarser angular resolution. The final identification is performed exclusively on the shortlisted candidates using the full masked Hamming distance with rotation compensation, which is identical to the distance function adopted in the verification experiments.

6.2.1 Metrics

The closed-set identification performances are reported using rank-based metrics, as standard in biometric identification. Results are summarized in Table 4, which reports Rank-k accuracies together with the Recall@K of the Top-K candidate selection stage. In particular, the system achieves a Rank-1 recognition rate of 75.13%, which increases to 79.60% and 82.20% at Rank-5 and Rank-10, respectively, indicating that the correct identity is often ranked among the top candidates even when not selected as the best match. Finally, Table 4 also reports a Recall@K of 87.93%, showing that in the majority of cases the true identity is retained within the Top-K shortlist. Probes whose true identity is not included in the candidate set cannot be correctly identified in the second stage; therefore, Recall@K represents an upper bound on

the attainable closed-set identification performance under the adopted Top-K configuration.

Table 4: Closed-set identification results

Metric	Value
Gallery identities	336
Probes	1500
Top-K shortlist size (K)	120
Rank-1 (RR)	75.13%
Rank-5	79.60%
Rank-10	82.20%
Recall@K (true in shortlist)	87.93% (1319/1500)

Figure 7 shows the corresponding CMC curve and a confidence proxy based on the *margin* between the best and second-best matches. The CMC curve exhibits a steep increase for small rank values (Rank-1 to Rank-10), consistently with the values reported in Table 4, and then progressively saturates, suggesting diminishing gains for larger ranks.

As shown in Fig. 7b, many probes have margins close to zero, meaning that the distances of the first and second-ranked candidates are almost identical and showing that the system is uncertain, since a very small change in the distance values could alter the ranking. Larger margins instead indicate a clear separation from the competitors, implying a more reliable identification

(a) Closed-set CMC curve.
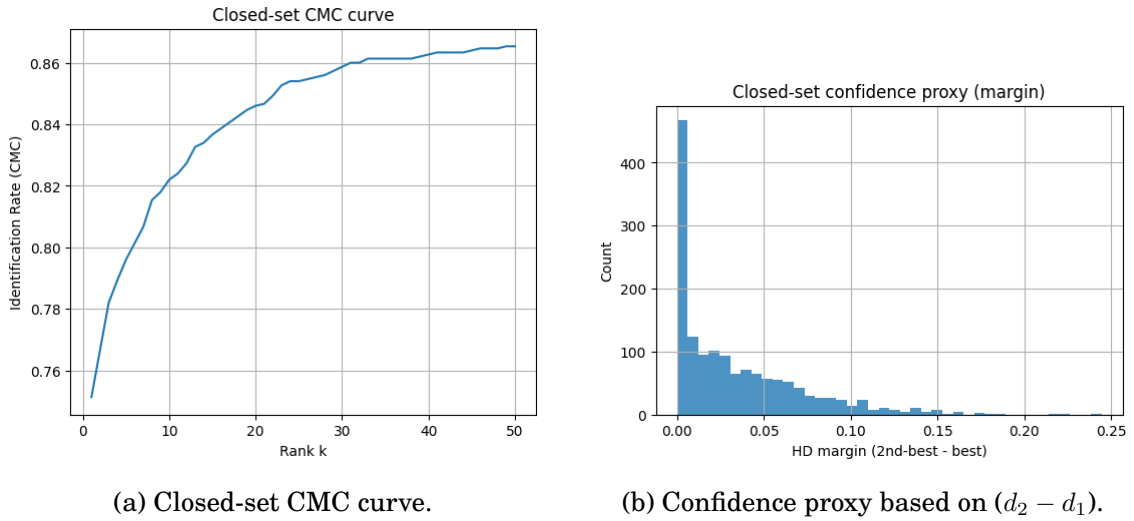
(b) Confidence proxy based on ($d_2 - d_1$).

Figure 7: Closed-set identification results: CMC curve (left) and distribution of the distance margin between the best and second-best candidates (right).

6.3 Identification: Open Set

In the Open Set identification scenario, the probe set includes both enrolled and non-enrolled identities. Unlike the closed-set case, the system must therefore perform two tasks: identifying known subjects and rejecting probes belonging to identities not present in the gallery. As a result, a rejection option is explicitly introduced, and identification is only performed when the probe is accepted.

The open-set protocol is defined by splitting the available identities into two disjoint groups: *known* is composed by identity enrolled in the gallery, while the remaining identities are grouped in *unknown* group and used exclusively as probe samples. For each known identity, one template is used for enrollment and the remaining ones act as known probes, while for unknown identities all available samples are treated as probes to be rejected.

As in the closed-set case, identification is carried out using a two-stage matching strategy. A fast Top-K candidate selection stage is first applied to reduce the computation, followed by the full masked Hamming distance with rotation compensation computed only on the shortlisted candidates. The same matching function adopted for verification and closed-set identification is therefore used, ensuring full consistency across tasks.

6.3.1 Metrics

Quantitative open-set identification results are reported in Table 5, which summarizes the system performance at two representative operating points. At an operating point of approximately 1% FPIR, the system achieves a DIR@Rank-1 of 51.50%, indicating that about half of the enrolled subjects are correctly identified while keeping the false identification of unknown subjects at a low level. By further tightening the security constraint to 0.1% FPIR, the DIR@Rank-1 slightly decreases to 50.34%, showing a stable behavior around the operating threshold.

Table 5: Open-set identification results at selected operating points.

Target FPIR	Threshold t	DIR@Rank-1	FNIR
1.0%	0.4666	51.50%	48.50%
0.1%	0.4645	50.34%	49.66%

Figure 8 reports the DIR–FPIR curve obtained by varying the decision thresh-

old. The curve shows a steep increase in DIR for low FPIR values, followed by a progressive saturation around a DIR of approximately 0.58. This behavior indicates that most of the performance gain is achieved by allowing a small fraction of false identifications of unknown subjects, while further relaxing the threshold yields diminishing improvements. As expected, open-set identification performance is significantly lower than the closed-set Rank-1 accuracy, reflecting the additional requirement of rejecting unknown identities.
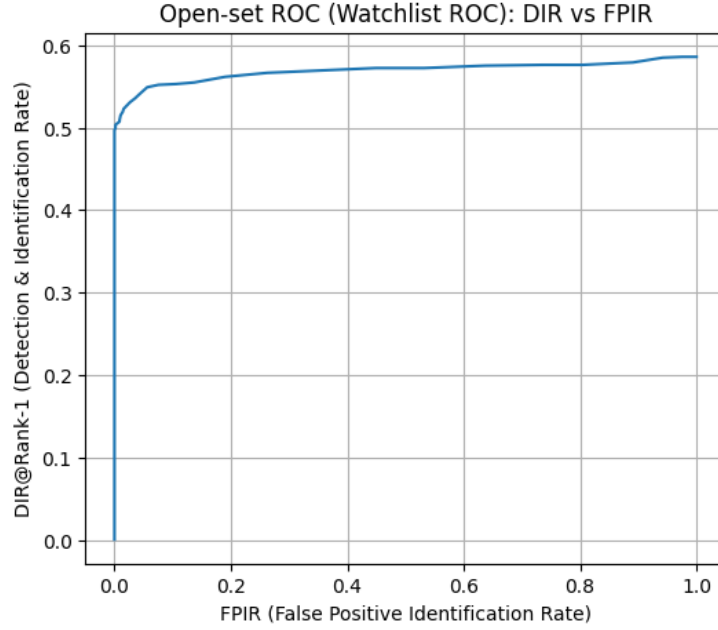

Figure 8: Open-set identification performance: DIR@Rank-1 versus FPIR

7   Conclusion

This work presented a segmentation-free iris recognition pipeline based on a regression-driven localization of pupil and iris geometry, combined with a Daugman-style normalization and encoding process. By relying on a reduced 6D representation and avoiding dense segmentation masks, the proposed approach simplifies the localization stage and reduces annotation and computational requirements.

The experimental evaluation shows that the system achieves quite good performance in verification and closed-set identification, while open-set identification remains the hardest task with weaker results. At the same time, the absence of explicit eyelid landmarks and pixel-level masks introduces additional challenges, particularly in the presence of occlusions and low-contrast iris boundaries.

Overall, the results confirm that a segmentation-free pipeline based on direct geometric regression can be effectively integrated into a classical iris recognition framework. While the lack of explicit occlusion modeling limits performance under challenging conditions, the proposed system provides a solid and interpretable baseline for further improvements, especially in scenarios where segmentation annotations are unavailable or impractical.

7.1   Future Work

Future work may focus on extending the current segmentation-free localization framework by incorporating additional geometric information. In particular, the use of a richer set of geometric points directly regressed from the input image, including eyelid landmarks (as performed in the paper [2]), could improve the robustness of iris localization.

Another promising direction concerns the refinement of the iris boundary itself. While the current pipeline includes a Pupil Refinement Network (PRN) to improve pupil localization accuracy, a similar refinement stage could be introduced for the iris, as a new model called Iris Refinement Network IRN. An Iris Refinement Network (IRN), operating on a localized iris region in a manner analogous to the PRN, could correct residual localization errors produced by the initial regression stage. This refinement could lead to a more stable normalization process and potentially improve both verification and identification performance, especially in open-set scenarios.

20

References

[1] J. Daugman, "How iris recognition works," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 21–30, 2004.

[2] T. Toizumi, K. Takahashi, and M. Tsukada, "Segmentation-free direct iris localization networks," *arXiv preprint*, 2022.

[3] Chinese Academy of Sciences, "Casia iris image database." `http://www.cbsr.ia.ac.cn/english/Irisdatabase.asp`, 2026.

[4] Swoyam2609, "Casia iris interval dataset." `https://www.kaggle.com/datasets/swoyam2609/casia-iris-interval`, 2025. Kaggle dataset.