Technical Note

Forensic Iris Recognition: A Survey

Rasel Ahmed Bhuiyan [1]
Adam Czajka [1]

Abstract: Iris recognition is a biometric technology that utilizes the feature sets in an individual's iris for identification purposes. Iris recognition is a non-invasive technique that does not require physical contact with the identified individual. Post-mortem iris recognition refers to using iris images from a deceased person to identify or verify their identity. It has several potential applications, such as forensic investigations and disaster victim identification. However, biological changes after death can cause significant differences between the post-mortem and ante-mortem iris data, which presents challenges for iris-capturing sensors, feature extractors, and iris matchers. This paper surveys existing research on using iris images for post-mortem identification, including a comprehensive review of the state-of-the-art and a summary of the latest results and observations. This survey has several unique elements, which provide a valuable resource for researchers and practitioners seeking to understand the capabilities and limitations of post-mortem iris recognition technology. Firstly, we discuss the post-mortem iris recognition steps and biological changes in the iris texture after a person's death from a medical standpoint. We then present the approaches to address the post-mortem iris recognition problem, including traditional iris recognition techniques, deep learning-based strategies, and interpretable methods. Furthermore, we provide the potential confounding factors that might impact the recognition performance. We also offer a comprehensive review of the publicly available post-mortem iris databases and the evaluation metrics used to assess the performance of post-mortem iris recognition systems. Finally, we conclude the paper by providing a constructive discussion and emerging future research directions.

[1] Department of Computer Science and Engineering, University of Notre Dame, 384 Fitzpatrick Hall of Engineering, Notre Dame, 46556, Indiana, USA

Introduction

Biometric recognition is the science of identifying individuals by their unique biological and behavioral characteristics such as iris, fingerprints, face, voice, and gait [1]. Biometric systems generally use sensors to capture samples of these attributes, extract features from the samples, and compare them to determine if they match. This is done by generating a comparison score, which is used to decide if the samples come from the same individual or different individuals. While various biometric traits have been proposed in the literature, face, fingerprints, voice, and iris recognition have been extensively studied and used in many applications such as border control, disaster victim identification, access control, online payments, and forensic investigation [2, 3, 4, 5, 6]. The focus of this study is forensic iris recognition.

Iris recognition involves analyzing the texture of the iris, which is seen especially clearly in near-infrared light (NIR). Near-infrared light is agnostic to melanin pigment contained within the iris; this pigment can make it difficult to see iris texture in visible light. The human iris is an internal organ that regulates the amount of light that enters the retina. It is a colored, ring-like organ, with the pupil forming the inner border and the ciliary border forming the outer boundary. The texture of the iris is the main feature used for recognition, and it is highly variable from person to person. From the perspective of traditional iris recognition algorithms, two irises of the same person, or even the irises of identical twins, are as distinct from each other as the irises of unrelated individuals. The iris is well-protected by the cornea making it a desirable biometric identifier. Ongoing research efforts are focused on overcoming various challenges, including detecting presentation attacks (using someone else's biometric data for fraudulent purposes), dealing with eye diseases, using iris recognition for newborn identification, developing age-agnostic algorithms, and evaluating the reliability of forensic iris recognition [7].

It was long-believed in scientific and industrial communities that iris recognition is difficult or even not feasible after a person's death. As an example, Daugman stated in an interview with the BBC in 2001 that "shortly after death, the pupil dilates and the cornea becomes cloudy," making it difficult to use the iris for recognition purposes [8]. Others have made stronger claims about the decay of the iris after death, such as Szczepanski et al. stated that "the iris decays within a few

minutes after death" [9]. Some commercial materials also state that stealing someone's iris after death is scientifically impossible because the iris is a muscle that completely relaxes after death, resulting in a fully dilated pupil with no visible iris. Therefore, it is commonly believed that a dead person's iris cannot be used for recognition purposes due to the lack of usable iris area [10], or iris features vanishing along with the pupil's dilation [11]. However, recent studies indicate that the decomposition of the eye and iris is more complicated and slower than previously assumed. The rate of decay is heavily influenced by the surrounding conditions. At present, we know that iris patterns can still be usable for recognition purposes for a few days after death in a warm environment [12, 13] and even for several weeks in a mortuary setting [14, 15, 16].

Comparison of this survey with the existing forensic iris recognition surveys.

Currently, two notable surveys in the literature focus on forensic iris recognition. The first survey, conducted by Boyd et al. in 2020 [6], centered on introducing different methods for post-mortem iris recognition and delved into the reasons behind recognition failures. Additionally, the researchers conducted an experiment to understand the differences between pre-mortem and post-mortem iris recognition techniques. The second survey by Matey et al. [17], currently a Technical Note provided by the National Institute of Standards and Technology and not known to be published in a peer-reviewed venue, is an overview of the existing forensic iris recognition methods and focused on how optical wavelengths, image resolution, diseases, and illnesses impact iris recognition performance.

Both existing surveys, however, have certain gaps in their coverage. The surveys did not address the topic of interpretability, which is crucial in helping forensic examiners understand and trust the results provided by the forensic iris recognition systems. Furthermore, the previous studies did not explore potential model biases related to demographic factors like Post-mortem Interval (PMI), age, gender, and cause of death, which could affect the fairness and accuracy of the systems across different populations. Lastly, the absence of comprehensive evaluation metrics in the existing surveys is a significant limitation. Proper evaluation metrics are essential for objec-

tively assessing the performance, efficiency, and robustness of post-mortem iris recognition systems.

To fill these gaps, this survey paper aims to offer a comprehensive and well-structured overview of the process involved in designing a post-mortem iris recognition system. The paper does not only summarize existing works in the field but also makes significant contributions in several key areas:

1. Discussion of the post-mortem iris recognition steps and post-mortem changes that occur in the iris, including pupil dilation, corneal clouding, and iris color change from a medical standpoint.

2. Presentation of different approaches that have been proposed to address the post-mortem iris recognition problem, including traditional iris recognition techniques, deep learning-based approaches, and interpretable methods to assist forensic examiners.

3. Discussion of the potential confounding factors (e.g.PMI, age, gender, reason of death) that might impact recognition performance

4. Comprehensive review of the publicly available post-mortem iris databases and the evaluation metrics used to assess the performance of post-mortem iris recognition systems.

5. Discussion of emerging future research directions.

Steps of Post-Mortem Iris Recognition

The post-mortem iris recognition process is complex, and it requires expertise in various areas such as forensic pathology, biometrics, and computer vision. Figure 1 depicts the post-mortem iris recognition key steps from image acquisition through match decision. The process involves multiple steps, which are as follows:

Acquisition: The first step in post-mortem iris recognition is to acquire iris images from the deceased person. This can be done using various methods, such as using a digital camera, a mobile phone camera, or a specialized iris imaging device. Current state of the art in post-mortem iris recognition sensing is to follow the ISO/IEC 19794-6 recommendations developed for live iris scanning. This includes application on near infrared

illumination (700-900 nm), which allows for better visibility of iris texture.

Pre-processing: The acquired iris images need to be pre-processed to remove noise, annotate distortion, or artifacts that might have been introduced during image acquisition. This step includes image enhancement, segmentation, normalization, and quality assessment.

Feature extraction: The pre-processed iris images are then processed to extract the feature set of the iris that can be used for identification or verification. Classical approaches incorporate image filtering and creating a binary representation of the iris texture ("iris code"). Recent approaches, based on deep learning models, transform iris images into feature (a.k.a. embedding) vector in a feature space defined by a given deep learning model.

Template creation: The extracted iris features are used to create a template that represents the unique features of the iris. The template is a mathematical representation of the iris features that can be stored in a database for later use. Iris templates often are created to conform with the ISO/IEC 19794-6 for interoperability purposes.

Matching: The next step is to compare the post-mortem iris template with the templates of previously registered individuals to identify or verify the deceased person's identity. The matching algorithm compares the post-mortem iris template with the registered templates in the database and returns a comparison score.

Decision: The final step is to make a decision based on the comparison score obtained in the matching step. If the comparison involves calculating the similarity, and such score is above a certain threshold, the system concludes that the post-mortem iris template belongs to the registered individual, and their identity is confirmed. If the similarity score is below the threshold, the system concludes that the post-mortem iris template does not match any of the registered individuals, and their identities cannot be confirmed.

These steps involve several challenges due to biological changes that occur in the iris after death, such as post- mortem pupil dilation or corneal clouding. Consequently, post-mortem iris samples require specialized algorithms and techniques.
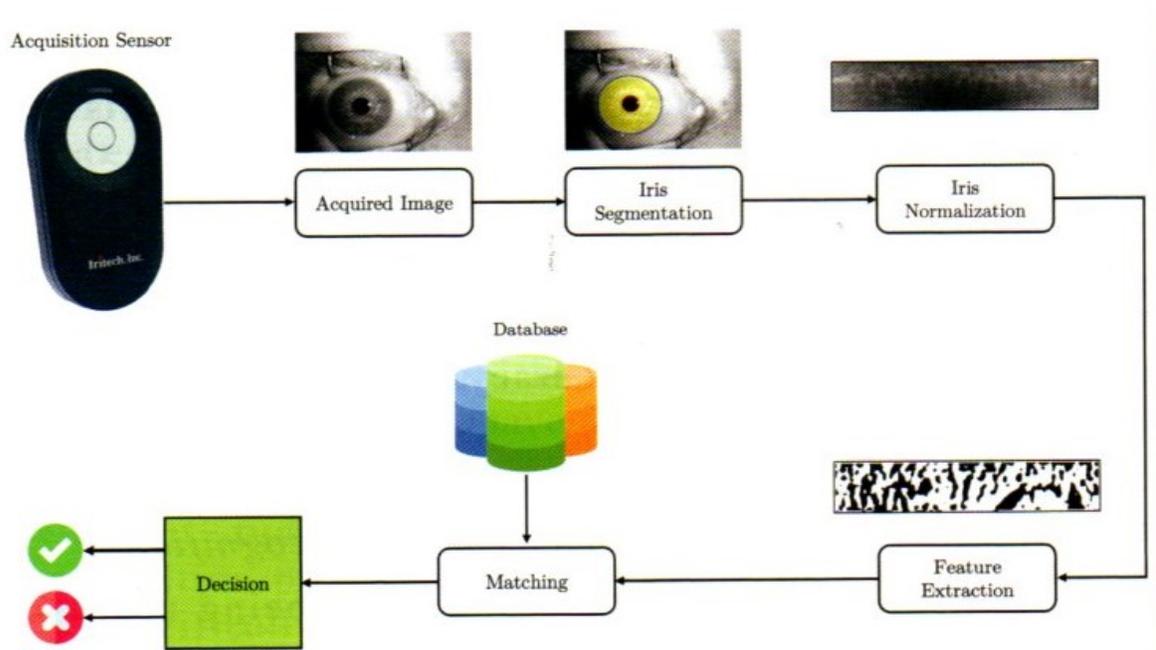
Figure 1

Overview of a typical automatic post-mortem iris recognition process. The image illustrates the key steps involved in post-mortem iris recognition: iris acquisition, image pre-processing, feature extraction, matching, and decision-making.

Selected Post-Mortem Changes in The Iris from The Medical Science Standpoint

Post-mortem changes that occur in the iris can significantly affect the performance of iris recognition systems. These changes can be classified into two categories: macroscopic changes and microscopic changes. Macroscopic changes refer to changes that are visible to the naked eye, such as pupil dilation, corneal clouding, and iris color change. Microscopic changes refer to changes that occur at the cellular or molecular level, such as cell death and tissue degradation [16, 18].

Pupil dilation is one of the most significant macroscopic changes that occur in the iris after death. Pupil dilation can occur due to the loss of control over the autonomic nervous system, which controls the size of the pupil. Pupil dilation can affect the recognition performance of iris recognition systems as it can change the iris texture and alter the appearance of the iris [19, 20, 21]. Corneal clouding is another macroscopic change that occurs in the iris after death. Corneal clouding refers to the opacity of the cornea, which can occur due to the accumulation of fluid or degradation of the corneal tissue [22, 23, 24]. Corneal clouding can affect the recognition performance of iris

recognition systems as it can make it difficult to extract the iris features accurately [25]. Iris color change is another macroscopic change that can occur in the iris after death. Iris color change can occur due to the accumulation of blood in the iris tissue, which can alter the appearance of the iris [26, 27]. While iris color is not currently the main identifying factor used in iris recognition (including forensic analysis), abnormal color may affect the contrast in near infrared image, which in consequence can affect the segmentation accuracy.

Feasibility Studies of Post-Mortem Iris Recognition

The notion of forensic iris recognition being unfeasible persisted for a considerable period owing to misconceptions about the dilation of the pupil following demise, the cornea becoming cloudy [8], or the iris decomposing completely within a few minutes after death [9]. Sansola et al. are credited as the first researchers to refute these beliefs by demonstrating that accurate comparison of iris images taken before and after death is possible [2]. Additionally, Sansola et al. found that when only post-mortem images were compared, the accuracy of the comparison was at least 70%, depending on the time passed after death. Due to the challenges involved in establishing a data collection process for post-mortem human subjects, some researchers opted to conduct experiments on animals. In one such experiment, Saripalle et al. studied the irises of domestic pigs and found that the biometric performance degrades gradually after the irises are removed from the body and that performance reached its lowest level 6 to 8 hours after death [28].

According to initial experiments done by Trokielewicz and his colleagues, the iris can continue to be used as a biometric identifier for up to 27 hours after death [14]. Trokielewicz et al. found that the iris can still be accurately encoded and recognized in over 90% of the cases if the image is captured between 5 to 7 hours after death. Further, Trokielewicz et al. experimented with a longer time-period, including samples collected up to 17 days after death from 17 deceased people [15]. The researchers found that the commercial IriCore method (iris recognition software developed by Iritech, Inc.) performed well with samples collected up to 60 hours after death, with an Equal Error Rate (EER) as low as 13% and claimed that correct matches could still even occur 17 days after death.

Bolme et al. investigated the biometric performance of face, fingerprint, and iris recognition during human decomposition in outdoor conditions [12]. Bolme et al. found that fingerprints and faces were moderately resilient to decomposition, but the irises degraded quickly, becoming useless for recognition only a few days after exposure to outdoor conditions. According to Blome et al., the correct verification rate was close to zero if the bodies were kept outside for 14 days. A later study by Sauerwein et al., however, showed that irises can still be readable for up to 34 days after death if the cadavers were exposed to low temperatures in winter [13].

A recent study on Trokielewicz et al. sheds light on the decomposition processes that occur in the human eye after death, specifically in the context of automatic identity recognition [29]. The paper presents an analysis of both local and global changes that occur during eye decay in one subject and provides recommendations for forensic examiners who use iris biometrics during their investigations.

Post-Mortem Iris Recognition Methods

Forensic iris recognition historically received less attention in the literature in part due to a belief that the iris possesses little biometric value after death. Recent discoveries in post-mortem iris biometrics, and the role of the iris as a forensic tool, have gained the attention of researchers. Several approaches have been proposed to address challenges seen in the forensic iris recognition. A recent study by Trokielewicz et al. demonstrated that it is still possible to make accurate matches up to 21 days after a person's death. The study identified segmentation errors as the primary reason for recognition failures and evaluated the risk of false matches when comparing post-mortem iris samples to databases of living iris samples [16].

Given this concern, Trokielewicz and Czajka [30] proposed the first known data-driven method for automatic segmentation of iris images obtained from deceased individuals using a deep convolutional neural network based on the SegNet architecture [31]. The model could recognize and accommodate specific distortions that exist in post-mortem samples but are absent in live irises. Further study demonstrated the ability to develop a dependable approach for detecting fabricated iris presentations from cadavers [32], reaching 99% accuracy.

Trokielewicz et al. proposed an end-to-end segmentation model to extract highly irregular iris texture areas in post-mortem iris images suited for detecting the rough iris and sclera boundaries [29], which is the extended model proposed in [30]. The proposed method achieves an ERR of less than 1% for samples collected up to 10 hours after death, and an EER of 21.45% for samples collected up to 369 hours post-mortem.

Though segmentation plays a crucial role in post-mortem iris recognition, a suitable feature extractor is needed to extract features from segmented irises for robust and accurate matching. Considering the limitation, Trokielewicz et al. introduced a novel iris feature representation specifically designed for post-mortem situations. This method used Siamese networks to learn data-driven filtering kernels that were later combined with a classical set of Gabor kernels that together could precisely describe iris characteristics regardless of the impact of certain post-mortem decay processes [33, 34]. This method improved accuracy rates for post-mortem intervals (12-369 hours) compared to a top commercial iris matching technology (IriCore).

Trokielewicz et al. investigated post-mortem iris recognition for forensic purposes and compared the differences in how human examiners and a convolutional neural network-based classifier perceived iris features for verification [4]. The study found that the machine classifier provides additional cues about matching and non-matching regions since the saliency maps from humans and machines rarely overlap. The study also found that both humans and the machine classifier prefer "sparse" attention or a "key-point-based" approach (focus on specific sub-sets of iris features).

There is still a requirement for an end-to-end model that can be understood and interpreted to assist forensic human examiners in their endeavors. This problem was approached by Kuehlkamp et al. who developed a deep learning-based end-to-end iris recognition system that is designed to provide interpretable results for forensic and medical examiners [5] in a form of a Class Activation Map (CAM) overlaid on post-mortem iris images to suggest to the human examiner regions that were salient to the model.

Boyd et al. addressed the issue of interpretability in post-mortem iris recognition by introducing a novel approach that utilized a human saliency-driven interpretable system based on patch-based matching [3]. This system presented human examiners with easily understandable visual cues to justify

their identification decision and provide a confidence score. The proposed method outperformed the commercial VeriEye approach and achieved state-of-the-art recognition performance. By providing interpretable results, this approach enhances the reliability and trustworthiness of post-mortem iris recognition and offers potential benefits for forensic investigations.

Table 1 provides an overview of papers that do not offer databases or new solutions in the form of source codes. Table 2 summarizes the available open-source solutions for post-mortem iris recognition proposed to date.

Authors	# Subjects	Potential Findings
Sansola [2]	43	False Non-Match Rate (FNMR) for post-mortem iris samples ranged from 19% to 30% based on time since death.
Saripalle et al. [28]	17 eyes	Biometric capabilities of ex-vivo pig eyes deteriorate after approximately 8 hours after removal from the cadaver.
Trokielewicz et al. [14]	17	A person's iris can be verified up to 27 hours after death (without noticeable pupil dilation) with a FNMR of 26.7%.
Bolme et al. [12]	12	The automatic verification method assessed that bodies kept outdoors for 14 days have a correct verification rate of 0.6%.
Sauerwein et al. [13]	12	According to a human examiner, some irises have interpretable feature sets after being outdoors during winter for 34 days.
Trokielewicz et al. [40]	1	Analyzed and compared iris decomposition processes and the dynamics of near-infrared and visible light. Guidance offered to forensic examiners.

Table 1

Main findings in research not offering databases or new solutions in a form of source codes.

Authors	# Subjects	Potential Findings	Offered
Trokielewicz et al. [30]	37	Proposed a new model for segmenting iris images after death based on CNN; model improved the IoU (Intersection over Union) metric from 73.6% to 83%.	CNN-based segmentation model weights
Trokielewicz et al. [32]	37	PAD (Presentation Attack Detection) method for post-mortem iris demonstrated a detection accuracy of 99%.	PAD solution
Trokielewicz et al. [4]	37	Determined that humans and trained CNN (Convolutional Neural Network) exploit different regions of the iris during classification tasks.	CNN-based classifier model weights
Trokielewicz et al. [29]	79	Introduced an iris recognition process that involves Gabor-based filtering and a CNN-based segmentation technique designed explicitly for post-mortem cases.	CNN-based model weights & source codes
Trokielewicz et al. [34]	79	Proposed a method for encoding post-mortem iris images using a combination of Gabor and CNN-generated kernels.	Segmentation model & filtering kernels
Kuehlkamp et al. [5]	430	Incorporated CAM (Class Activation Map) inside the proposed CNN model to suggest regions that were salient to the model to assist the forensic examiner.	Source codes
Boyd et al. [3]	259	Proposed a post-mortem iris comparison method directed at human interpretation and designed specifically for forensic application. EER reported as 12.8%.	CNN-based model weights & source codes

Table 2

Research offering source codes or model weights for post-mortem iris recognition methods.

Confounding Factors Impacting Post-mortem Iris Recognition Performance

Forensic iris is a new field that applies iris recognition technology to the investigation and prosecution of criminal acts [17]. Many factors might affect post-mortem iris recognition, such as time, temperature, and season [13]. The decomposition process causes deterioration in irises at different rates depending primarily on season and temperature. Irises are unusable after a few days in warmer temperatures due to rapid decomposition and insect colonization. However, irises may be usable in winter for a longer time [12].

Along with environmental factors, demographic dimensions such as Post-Mortem Interval (PMI), age, gender, and cause of death, might impact post-mortem iris recognition. PMI is the essential factor for post-mortem iris recognition. If the PMI is high, there is a high chance of losing some essential iris texture features, mostly due to cornea cloudiness and tissue degradation. A recent study demonstrates that increasing the PMI decreases recognition performance significantly [6]. Another critical factor is the subject's age. A recent study shows that live iris recognition performs worst for specific age groups and claimed that the age group between 30 to 60 performs better than the age group below 30 and above 60 [35]. The IREX IX part one [36] concluded that gender significantly impacts iris recognition performance, although the performance showed inconsistency in terms of favoring either of the group, males or females, towards better accuracy in general. These recent studies have not yet seen counterparts in post-mortem iris recognition, but chances are high that similar observations may be true in case of forensic iris.

Publicly Available Post-mortem Iris Recognition Datasets

Collection and curation of post-mortem biometric data belongs to the most challenging, complicated, and time-consuming biometric data collections. In addition to following the institutional review board and ethical committee's regulations, special respect is owed to all deceased subjects from whom researchers are given an opportunity to acquire knowledge ordinarily out of reach without such data. Table 3 summarizes the research datasets of post-mortem iris scans (datasets described further below), and Table 4 presents the main findings by research teams proposing these datasets.

Dataset Name	# Subjects (# Eyes)	# of Images (NIR/Visible)	Sensors Used	Min → Max PMI (hours)
Warsaw-BioBase-Post-Mortem-Iris v1.0 [15]	17 (34)	480 / 850	IriShield M2120U Olympus TG-3	5 → 814
Warsaw-BioBase-Post-Mortem-Iris v2.0 [16]	20 (40)	1,200 / 1,797	IriShield M2120U Olympus TG-3	7 → 453
Warsaw-BioBase-Post-Mortem-Iris v3.0 [29]	42 (84)	1,094 / 785	IriShield M2120U Olympus TG-3	7 → 369
NIJ-2018-DU-BX-0215 [37]	269 (522)	5,770 / 4,643	IriShield M2120U OmniVision OV8865	0.5 → 1,650

Table 3

Primary features of the currently available post-mortem iris datasets.

Authors	# Subjects	Potential Findings	Offered Database Size
Trokielewicz et al. [15]	17	According to the IriCore method, the ERR is 13% even 60 hours after a person's death. Some matches were found after 17 days, but with significant iris tissue damage.	17 subjects
Trokielewicz et al. [16]	37	Marches can occasionally occur up to 21 days after death, but comparing post-mortem samples to live samples may result in a higher FNMR.	20 subjects
Trokielewicz et al. [29]	79	Iris recognition process that involv ed Gabor-based filtering and a CNN-based segmentation technique designed explicitly for post-mortem cases.	42 subjects

Table 4

Research (with main findings) offering post-mortem iris databases.

Trokielewicz and colleagues have contributed significantly to the field of post-mortem iris recognition by releasing several public datasets. The first dataset, Warsaw-BioBase-Post-Mortem-Iris v1.0, includes 480 NIR and 850 color images of 17 subjects, along with age, gender, and cause of death metadata [14, 15]. The images were captured using IriShield M2120U and Olympus TG-3 cameras in 2 to 8 sessions, depending on tissue availability. The first session was conducted 5-7 hours after death, while subsequent sessions were taken at later intervals up to 17 days post-mortem.

The second dataset is the Warsaw-BioBase-Post-Mortem-Iris v2.0 [16]. This dataset contains 1,787 visible (VIS) and 1200 near-infrared (NIR) images of 20 cadavers captured by the same cameras used in [14, 15]. This dataset was collected in 2018 by conducting 1 to 13 sessions from 5 to 814 hours after death in the hospital mortuary. Most environmental conditions are unknown, such as where cadavers were kept before entering the cold storage, air pressure, and humidity. However, the mortuary room temperature was approximately 6° Celcius or 42.8° Fahrenheit.

The third dataset, which is available to the public, was presented in 2020 by Trokielewicz et al. [29] and is known as the Warsaw-BioBase-Post-Mortem-Iris v3.0. This dataset includes 1094 near-infrared (NIR) images and 785 visible (VIS) images from 42 cadavers, taken up to 369 hours after death.

To the best of our knowledge, the latest and most extensive dataset is NIJ-2018-DU-BX-0215 [37]. This dataset was gathered at the Dutchess County Medical Examiner's Office (DCMEO) from 269 individuals who had passed away. To compile the complete dataset, a total of 53 data-gathering sessions were conducted. Following meticulous selection and organization, this recently curated dataset now comprises 5,770 near-infrared images and 4,643 visible-light images. Table 3 summarizes all known to us publicly available forensic iris databases.

Figure 2 depicts a collection of iris images obtained at various time steps after demise. Based on the observations, it can be inferred that the live and peri-mortem irises exhibit clear distinctive features and textural patterns. However, during the period between the early and late post-mortem stages, deformations in the cornea and potentially the iris texture are evident. These post-mortem-related changes are mainly visible in the form of cornea opacification, which can be mitigated by applying near infrared illumination (in the ISO-compliant range,

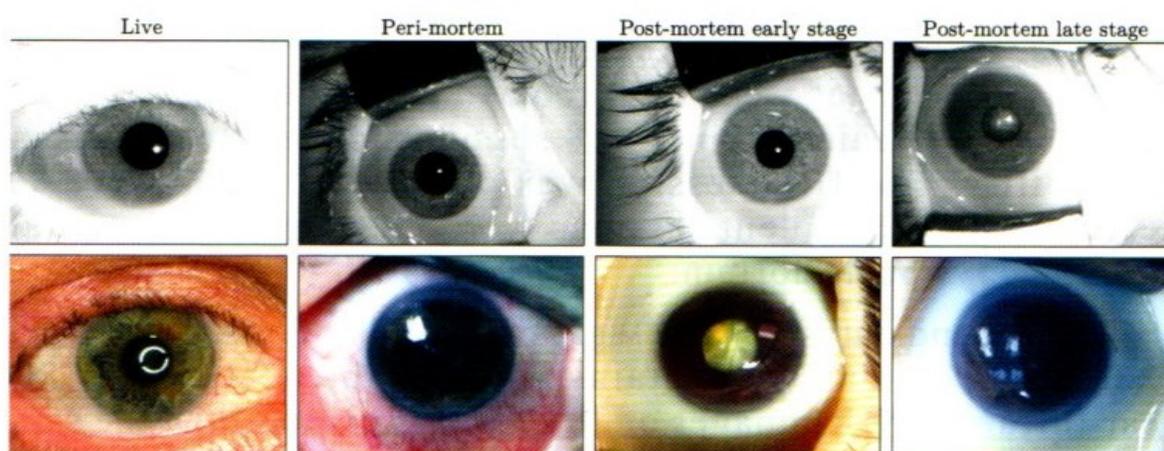from 700 through 900 nm) instead of visible light (400-700 nm) illumination.

Figure 2

The NIR and RGB samples collected at different time steps after death are represented in the first and second rows, respectively. For the live samples, data was collected from the Warsaw disease dataset [39], while post-mortem samples were obtained from the NIJ-2018-DU-BX-0215 dataset [37]. The live sample was acquired prior to death, the peri-mortem sample was collected around the time of death, and the post-mortem samples were obtained at early stage (18 hours after death) and late stage (389 hours after death). These samples provide insights into the changes in iris texture over time. It is worth noting that the RGB illumination is incapable of penetrating the cloudiness of the cornea, and thus capturing the iris texture information is difficult. This is to a high extend mitigated by applying NIR light, which is barely absorbed by the melanin pigment present in the eye tissue.

Performance Evaluation Metrics

The evaluation of post-mortem iris recognition systems is critical to determine their effectiveness in real-world applications. In this section, we discuss the most common performance evaluation metrics used in biometrics, including post-mortem iris recognition:

FMR: False Match Rate (FMR) is a commonly used metric for evaluating the performance of post-mortem iris recognition methods, excluding failures to acquire. FMR refers to the percentage of times that the method incorrectly matches an individual's biometric sample with a different identity. In other words, it measures the rate at which the algorithm produces a false positive result. FMR can be calculated as shown in Figure 3, where τ is the acceptance threshold:

$$FMR(\tau) = \frac{\text{\# of false matches for } \tau}{\text{\# of all impostor comparisons}}$$

Figure 3

Formula for calculating FMR score.

FNMR: False Non-Match Rate (FNMR) is another commonly used metric, which refers to the percentage of times that the algorithm fails to match a biometric sample with the correct identity. In other words, it measures the rate at which the system produces a false negative result. The formula for calculating the FNMR is shown in Figure 4:

$$FNMR(\tau) = \frac{\text{\# of false nonmatches for } \tau}{\text{\# of genuine comparisons}}$$

Figure 4

Formula for calculating FNMR score.

EER: The Equal Error Rate (EER) is one of many operational points and is calculated by setting a threshold for which FMR and FNMR are equal. A lower EER indicates better system performance, as it means that the system is able to accurately identify individuals with a lower rate of false matches and false non-matches.

Decidability score (d') measures the separation between two normal distributions and is often used in assessing the performance of iris recognition systems. It reflects the ability of the system to distinguish between the genuine and imposter identities, based on the overlap between the distributions of genuine and imposter scores. Figure 5 displays the formula for calculating the decidability score.

$$d' = \frac{|\mu_g - \mu_i|}{\sqrt{\frac{1}{2}(\sigma_g^2 + \sigma_i^2)}}$$

Figure 5

Formula for calculating decidability score.

In the decidability score calculation, μ_g and μ_i are means, and σ_g and σ_i are standard deviations of the comparison score distributions for genuine and impostor iris pairs. A higher d' value indicates better separation, with a d-prime value of 0 indicating a perfect overlap of both distributions (equivalent to a random chance classification).

ROC curve: The Receiver Operating Characteristic (ROC) curve is a parametric representation of FMR and FNMR, where the acceptance threshold is a parameter. A good post-mortem iris recognition system should have a low FMR and FNMR, resulting in a ROC curve that is close to the upper-left corner of the plot. To illustrate a state-of-the-art ROC curve for post-mortem iris recognition, we collected data from the Warsaw and NIJ datasets and sampled randomly 2,500 genuine and 2,500 impostor iris pairs. Using the post-mortem iris recognition matcher resulting in best performance across all matchers tested by the authors (HDBIF [38]), we determined the genuine and impostor comparison scores translated to a ROC curve shown in Figure 6. The ROC curve's proximity to the top-left corner indicates a high-performing model, achieving higher TMR and lower FMR across different thresholds. The gray dashed line represents the EER line, the orange line represents the ROC curve, and the red-green circle represents the EER where FMR and FNMR are equal. For the randomly chosen pairs, the HDBIF model achieved AUC=94% and EER=12%.
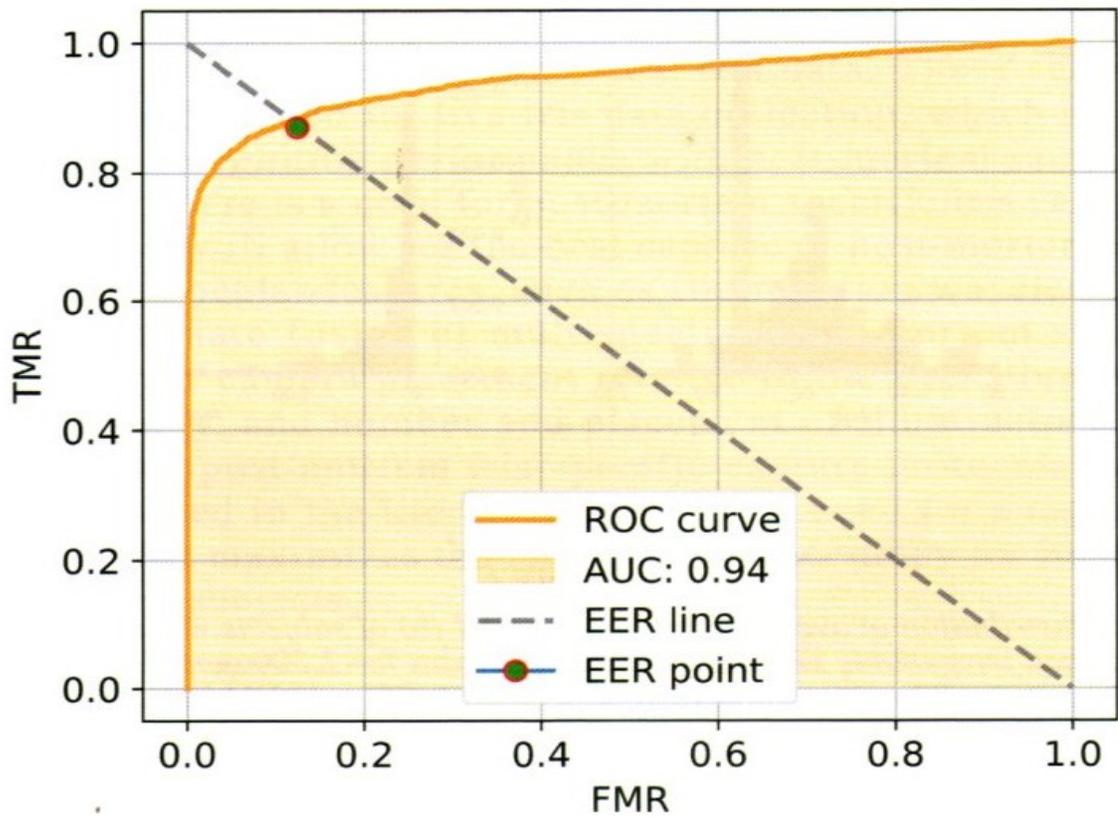
Figure 6

The ROC curve generated for the selected 5,000 genuine and impostor pairs drawn from Warsaw [15, 16, 29] and NIJ [37] combined datasets.

The genuine and impostor pairs for the selected 5,000 genuine and impostor pairs drawn from Warsaw [15, 16, 29] and NIJ [37] combined datasets were selected in a way to balance the age and PMI in both groups, as shown in Figure 7 (a) and in Figure 7 (b), respectively.
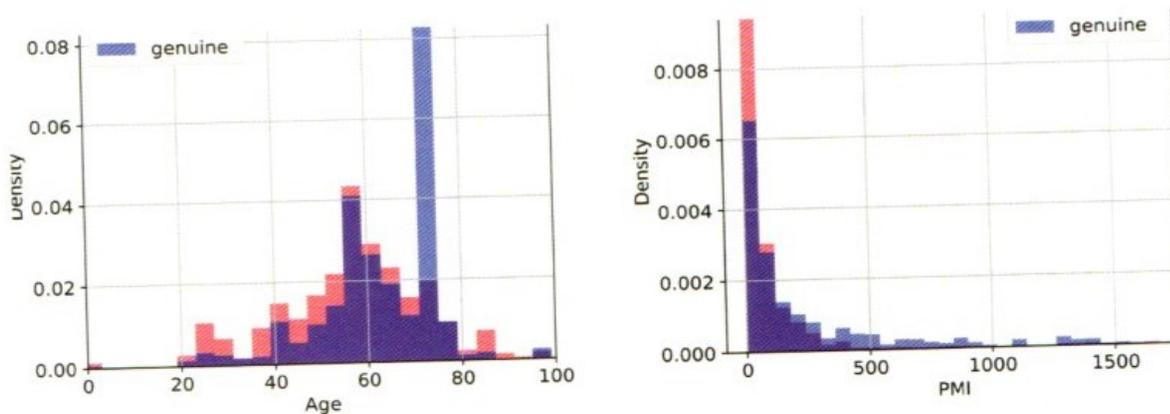
Figure 7

The distribution of average age (a) and PMI (b) of subjects whose eyes were selected for genuine and impostor comparisons for 2,500 genuine and 2,500 impostor pairs drawn from Warsaw [15, 16, 29] and NIJ [37] combined datasets.

AUC: The Area Under the ROC Curve (AUC) is commonly used in biometric recognition to evaluate the overall performance of a system. The AUC=1.0 corresponds to a perfect classifier, while AUC=0.5 denotes a random-chance classification.

It is noteworthy that calculating accuracy is not always the most appropriate metric since the distribution of genuine and impostor scores in post-mortem iris recognition is often different from that in live iris recognition. This can lead to artificially inflated accuracy scores that do not accurately reflect the system's ability to identify individuals correctly. In addition, the decidability d' assumes that the two distributions are normal. This assumption may not hold true in all cases, and in such situations, the decidability score may not accurately reflect the discriminability power of the system. Therefore, it is recommended to use metrics such as FMR (for a given FNMR), or FNMR (for a given FMR), EER, and AUC to provide a more comprehensive evaluation of the system's performance.

Future Directions

The current state-of-the-art on post-mortem iris recognition suggests that there has been significant progress in this area, but there are still many challenges that need to be addressed.

Firstly, all the studies used the IriShield M2120U and Olympus TG-3 sensors to collect post-mortem iris data. These sensors are mainly designed for live iris data collection, which might affect the recognition performance with post-mortem samples. Therefore, there is a need for post-mortem specific iris capture sensors that will allow for the best capture of post-mortem iris texture and ocular features, such as illumination wavelengths and appropriate fusion of multispectral post-mortem scans, polarization, camera placement relative to the eye, illumination intensity, and number and placement of illuminators. In addition, no post-mortem iris-specific capture protocols have been proposed in the literature. There is a need for a capture protocol that maximizes the identification accuracy for images of deceased subjects.

Secondly, several studies proposed small post-mortem iris recognition datasets with an insufficient number of subjects to support today's data-driven methods, which limits the generative power of those methods. Therefore, more initiatives are necessary to gather sizeable datasets and distribute these datasets, which can enable the development and testing of more accurate and reliable post-mortem iris recognition methods.

Thirdly, the performance of iris recognition systems is significantly inconsistent when dealing with post-mortem samples. In some cases, early scans do not match, while subsequent scans do match. This can be attributed to various factors, such as the degradation of the iris tissue, the presence of artifacts, and the variability in image quality. Therefore, more research is needed to better understand post-mortem tissue degradation's impact on iris recognition performance and to develop techniques that can compensate for this degradation.

Fourthly, many studies studied model biases for different confounding factors such as age, gender, and ethnicity in the context of live iris recognition. The few studies that checked the model's robustness for post-mortem interval still need an investigation into other confounding factors such as age, gender, and causes of death. It is assumed that these factors might affect recognition performance.

Finally, in the field of post-mortem iris recognition, the majority of the existing methods lack interpretability, which can be a significant limitation in forensic investigations. While some studies have introduced interpretability to assist forensic examiners, these methods are often hand-crafted and may not be suitable for large-scale automation. Hence, there is a need for the

development of automated systems that can provide interpretability with minimal human intervention. Such systems would increase the trustworthiness of post-mortem iris recognition results and improve the efficiency of the forensic investigation process. In addition, the integration of different modalities, such as iris and face or ocular regions, can be explored to develop more reliable and accurate post-mortem identification systems.

Conclusions

Post-mortem iris recognition is an evolving area of biometric recognition with numerous potential applications including forensic investigations and disaster victim identification. While the use of iris images for post-mortem identification presents several challenges, such as biological changes in the iris after death and limited availability of post-mortem iris databases, significant progress has been made in developing and evaluating post-mortem iris recognition systems.

The IREX IX report states that the FNMR for live irises should be less than 2% [36]. Post-mortem iris recognition is far from reaching such a low FNMR, and perhaps will never reach that accuracy levels due to inevitable and irreversible decomposition-related changes to the eye. Through a comprehensive survey of existing research, this paper has provided insights into the state-of-the-art in post-mortem iris recognition, including an overview of the different approaches proposed to address the problem, publicly available post-mortem iris databases, and evaluation metrics. Emerging trends and future directions have also been discussed, such as the need for more comprehensive and standardized evaluation protocols, increased focus on interpretable methods, and the exploration of multi-modal biometric approaches.

For further information, please contact:

Rasel Ahmed Bhuiyan
Department of Computer Science and Engineering
University of Notre Dame
384 Fitzpatrick
Hall of Engineering
Notre Dame, Indiana, 46556 USA
rbhuiyan@nd.edu