https://doi.org/10.31891/csit-2025-3-12
UDC 004.932.2

Yurii POHULIAIEV, Kirill SMELYAKOV
Kharkiv National University of Radio Electronics

METHOD FOR ASSESSING THE QUALITY OF FINGERPRINT COMPARISON BASED ON CONVOLUTION OF BOOLEAN METRICS

The proposed approach is designed to determine the degree of similarity between fingerprints. It is based on the comparison of key features of papillary patterns. This method can be used to compare both a pair of fingerprints and one fingerprint with a group of others. Within the framework of this approach, basic metrics are used that reflect the presence of matches between the compared fingerprints. Based on these metrics, the following characteristics are calculated: the percentage of matching features and the weighted average score, taking into account the significance of each match. These parameters are used to establish the belonging of a fingerprint to a certain group. To control the results, threshold values are used, established on the basis of the Bayesian classifier and the Neyman-Pearson lemma, taking into account the differences in the distribution of scores. Testing conducted using the FVC2000 datasets showed that the accuracy of comparing individual fingerprints is 89.3%. The method is implemented in Python using the Pandas and NumPy libraries. This allows it to be integrated into automated fingerprint identification systems (AFIS). Resistance to moderate distortions and interpretability of parameters indicate its practical applicability, in particular, in the field of forensics. One of the limitations of the method is its susceptibility to strong fingerprint distortions. Further development plans include the introduction of adaptive thresholds and integration with deep learning methods to improve the process of feature extraction on fingerprints. This development is aimed at increasing the accuracy and reliability of biometric identification, which is extremely important for reducing the likelihood of errors, especially in the context of forensic tasks.

Keywords: fingerprinting, biometric identification, Boolean metric convolution, Bayesian classifier, Neyman-Pearson lemma, quality assessment, group matching, normalized contribution, weighted mean, AFIS.

Юрій ПОГУЛЯЄВ, Кіріл СМЕЛЯКОВ
Харківський національний університет радіоелектроніки

МЕТОД ОЦІНКИ ЯКОСТІ ПОРІВНЯННЯ ДАКТИЛОСКОПІЧНИХ ЗРАЗКІВ НА ОСНОВІ ЗГОРТКИ БУЛЕВИХ МЕТРИК

Запропонований підхід розроблений для визначення ступеня подібності між відбитками пальців. Він базується на порівнянні ключових ознак папілярних візерунків. Цей метод може бути використаний для порівняння як пари відбитків пальців, так і одного відбитка пальця з групою інших. В рамках цього підходу використовуються базові метрики, що відображають наявність збігів між порівнюваними відбитками пальців. На основі цих метрик розраховуються такі характеристики: відсоток збігів ознак та середньозважений бал з урахуванням значущості кожного збігу. Ці параметри використовуються для встановлення належності відбитка пальця до певної групи. Для контроля результатів використовуються порогові значення, встановлені на основі байєсівського класифікатора та леми Неймана-Пірсона з урахуванням відмінностей у розподілі балів. Тестування, проведене з використанням датасетів FVC2000, показало, що точність порівняння окремих відбитків пальців становить 89,3%. Метод реалізовано на Python з використанням бібліотек Pandas та NumPy. Це дозволяє інтегрувати його в автоматизовані системи ідентифікації відбитків пальців (AFIS). Стійкість до помірних спотворень та інтерпретованість параметрів свідчать про його практичну застосовність, зокрема, в галузі криміналістики. Одним з обмежень методу є його сприятливість до сильних спотворень відбитків пальців. Подальші плани розвитку включають впровадження адаптивних порогів та інтеграцію з методами глибокого навчання для покращення процесу вилучення ознак на відбитках пальців. Ця розробка спрямована на підвищення точності та надійності біометричної ідентифікації, що надзвичайно важливо для зниження ймовірності помилок, особливо в контексті судово-медичних завдань.

Ключові слова: дактилоскопія, біометрична ідентифікація, згортка булевих метрик, байєсівський класифікатор, лема Неймана-Пірсона, оцінка якості, групове зіставлення, нормалізований внесок, зважена середня, AFIS.

Received / Стаття надійшла до редакції 31.07.2025
Accepted / Прийнята до друку 26.08.2025

Introduction

In forensics, security systems and law enforcement, fingerprint identification still plays an important role in biometrics. Existing methods for assessing the quality of fingerprint comparisons often encounter difficulties when analyzing fragmented or fuzzy data. This leads to an increase in the number of errors of both the first and second kind, especially when comparing one fingerprint with many others, such as when searching in databases.

In this paper, we propose a different approach to solving this problem. It is based on the use of convolution applied to Boolean metrics, which allows us to take a new look at assessing the quality of fingerprint comparisons. The novelty lies in the way binary match indicators are combined. This method makes it possible to determine with sufficient confidence that a particular fingerprint belongs to a certain group. For this purpose, two main parameters are calculated: the average similarity level and the number of positive matches. The proposed method assumes that the results of comparing two fingerprints are used to generate a score in the range from 0 to 1. A match is considered to be a result ≥ 0.5. Threshold values are introduced for both parameters to evaluate the group, which ensures double quality control. In other words, to recognize a match, the average similarity score must exceed the set threshold, and the number of positive matches must also be sufficient.

This approach can be especially useful in automated fingerprint identification systems (AFIS), where reliability plays a crucial role in conditions of uncertainty. The main goal of this study is to improve the accuracy and make the quality assessment more understandable. This, in turn, should lead to improved performance of forensic applications using fingerprinting. In the future, the proposed method will not only reduce the likelihood of identification errors, but also increase confidence in the results of examinations based on fingerprint analysis. In addition, it can be adapted for use in other areas where comparison and classification of complex objects based on incomplete or noisy data is required.

State of the art

Assessment of the quality of comparison of fingerprint samples is one of the most important aspects of the study. One of the undisputed leaders in this field is the US National Institute of Standards and Technology (NIST). In the early 2000s, the institute's specialists proposed an assessment method called NIST Fingerprint Image Quality (NFIQ). As assessment criteria, they proposed 11 features, each of which is assigned a quality class (from 1 to 5). The main idea of developing a method for assessing the quality of fingerprint samples was to improve the level of fingerprint reading by digital scanners, and preliminary assessment of the sample allowed to exclude low-quality samples from the comparison procedure. This allows, on the one hand, researchers in the field of fingerprinting and recognition to focus on higher-quality samples, and on the other hand, it eliminates possible comparison errors if the proposed method does not cope well with a sample with an abundance of scanning artifacts. Currently, NIST has standardized its method, which is now reflected in the ISO/IEC 29794-4 standard [1].

The scientific literature presents various approaches to comparing fingerprint samples. These approaches can be divided into several categories: methods based on the analysis of special points (minutiae), methods using the frequency characteristics of papillary lines, and modern algorithms based on deep learning and cryptographic methods.

Traditional methods, such as minutiae matching, analyze the location and relationship of special points on a fingerprint. These methods are easy to implement, but are vulnerable to distortions, noise, and changes in pressure during scanning. For example, in the article [2], the authors explore the possibility of achieving a balance between identification accuracy and the probability of errors in fingerprint verification. The proposed approach is based on combining estimates obtained by different algorithms, which makes it possible to increase the system's resistance to low-quality input data. Another approach presented in [3] is to use a flexible, lightweight, cancelable fingerprint template. The authors propose a matching framework that includes multi-stage filtering to improve the security and accuracy of sample matching.

In recent years, there has been a trend towards integrating biometric systems with privacy protection mechanisms and continuous authentication technologies. The review [4] discusses secure biometric identification methods used in continuous authentication systems. The authors focus on the issues related to usability and the impact of biometric sample quality on the reliability of the system.

In [2], a new variable-length fingerprint recognition method is proposed that combines Boolean operations and matching to improve performance in IoT systems. The authors claim that their approach allows achieving high data processing speed while maintaining the required level of accuracy. In [5], a multi-level security system for medical images is proposed that includes fingerprint quality assessment using local binary orientation patterns. The proposed method allows automating the process of access control to confidential patient data.

Development of methods aimed at preserving confidentiality is also a relevant area of research. The article [6] presents an authentication system based on homomorphic encryption and fingercode templates. The authors claim that the proposed approach improves quality assessment in fingerprint recognition systems while ensuring the protection of personal data. The paper [7] considers an optical cryptosystem for secure face and fingerprint recognition. The use of 3D transformation allows improving the quality of biometric sample compliance assessment in multimedia applications.

In the context of passwordless authentication, [8] analyzes modern fingerprint recognition methods and suggests ways to improve quality metrics. The authors note that improving the reliability and security of passwordless authentication systems is an important task.

A review of the researches confirms the above trends. A detailed review [9], published in 2025, discusses methods for fusion of data obtained from different sensors to determine the location of objects, including fingerprint recognition in indoor environments. In [10], the authors propose using deep learning for fingerprint recognition, focusing on the analysis of minutiae and the use of convolutional neural networks (CNNs) to improve the matching quality. A notable contribution by Deshpande et al. (2020) [11] introduced the "Combination of Nearest Neighbor Arrangement Indexing (CNNAI)" method, which achieved a Rank-1 identification accuracy of 84.5% on the NIST SD27 latent fingerprint database and 80% on FVC2004. This approach utilized local minutiae-based CNN matching models that generate rotation and scale-invariant feature vectors, eliminating the need for fingerprint pre-alignment required in traditional systems. The CNNAI method addresses key limitations of existing approaches by working with local minutiae features rather than depending on global features like core or delta points, which may not be available in latent fingerprints. The system demonstrated robustness against rotation, scale, and missing minutiae by defining triangular minutiae structures and utilizing hash-based indexing for efficient retrieval in large databases.

Despite significant progress in this area, existing studies often do not take into account the threshold values in group evaluation, limiting themselves to simple averaging of paired evaluations. Our work aims to fill this gap by introducing a convolution of Boolean metrics for bimetric evaluation (averaging and evaluation of positive contribution), which helps to improve the interpretability of the results and the performance of the system in single-fingerprint-to-group comparison scenarios.

Objectives and tasks of the study

The main objective of the study is to propose a new method for assessing how similar fingerprints are when one fingerprint is compared to a group. The proposed method should make the determination of a fingerprint's belonging to a group more accurate and reliable. This is achieved by combining two important indicators: the average level of similarity and the number of positive matches. Threshold values are used for double quality check. This will help reduce the number of errors in forensics and biometrics, especially when working with fuzzy or different data.

To achieve this goal, it is necessary to solve the following tasks:

1. Create a mathematical model for combining Boolean metrics, develop formulas for calculating the average score and the number of positive matches.
2. Establish and justify threshold values for metrics that will ensure a balance between sensitivity and specificity of the method.
3. Test the method on standard fingerprint databases and evaluate its performance by key parameters (accuracy, number of false positives).

Method of Integral Assessment of Similarity of Fingerprint Samples

To describe the method, we will define the set of fingerprint samples as follows: let the set of fingerprint samples be understood as a set of objects for which the equivalence relation can be defined as shown in formula (1):

$$\forall s_1, s_2 \in S: \exists R \subseteq S \times S: (s_1, s_2) \in R \Leftrightarrow M(s_1, s_2) \geq \theta \quad (1)$$

where S is a set of samples, R is a binary equivalence relation, M is a raw score for the match of a pair of samples, θ is the threshold for comparing samples.

The proposed method of integral assessment does not directly depend on the procedure by which we should compare the samples. Moreover, the method does not declare the internal structure of the samples, relying on complete abstractness and freedom of choice. The only thing the method requires is to establish the expected range of values, which is reflected in formula (2).

$$\forall T: \exists \phi: S \to T: \forall s_1, s_2 \in S: M(s_1, s_2) = M_T\big(\phi(s_1), \phi(s_2)\big), M(s_1, s_2) \in [0, 1] \quad (2)$$

where T is any set of fingerprint samples detailing the structure of the sample and the comparison details, ϕ is a mapping of an abstract set into a detailed one, M_T is a comparison function for specific samples.

The proposed threshold is usually the value θ=0.5 as the middle of the range of values of the function M. Let us show in more detail why this value is correct and justified as a class boundary in a binary classification problem on the space of fingerprint samples.

Let us consider two cases – when the sample of results M is homoscedastic and heteroscedastic. In the first case, we assume homoscedasticity as a consequence of the central limit theorem - that is, that both classes (match, do not match) have identical variance. In this case, we also assume equality of a priori probabilities - we do not have information in advance about whether two samples are more likely to match or not, which is shown in formulas (3) - (4).

$$\forall s_1, s_2 \in S: \big((s_1, s_2) \in R \Rightarrow M(s_1, s_2) \in C_1\big) \wedge \big((s_1, s_2) \notin R \Rightarrow M(s_1, s_2) \in C_0\big) \quad (3)$$

$$M(s_1, s_2) \sim N(\mu, \sigma), P(C_1) = P(C_0) = 0.5, \sigma_1 = \sigma_0 = \sigma, \mu_1 = 1, \mu_0 = 0 \quad (4)$$

where C_1 and C_0 are equivalence classes (coincide, do not coincide), N is a normal distribution with mathematical expectation μ (1 for the "coincide" class and 0 for the "do not coincide" class), σ is the variance of the distribution (in force homoscedasticity for equivalence classes it is the same).

Then, according to the Neyman-Pearson lemma, we need to take the most powerful criterion for testing a static hypothesis - the maximum likelihood criterion. With the equality of prior probabilities, such a criterion is the Bayesian classifier. This follows from the fulfillment of the likelihood maximization condition due to the equality of prior probabilities.

As follows from the Neyman-Pearson lemma (see formula 5):

$$\forall k: LR(x) = \frac{P(x|H_1)}{P(x|H_0)} > k \Rightarrow deny\ H_0, P(LR(x) > k|H_0) = \alpha \quad (5)$$

where LR(x) is the likelihood ratio, k is the parameter that specifies the level of errors of the first kind, α is a fixed test parameter determines fixed level of acceptable Type I error, H_0 is basic statistic decline hypothesis, H_1 is basic statistic accept hypothesis .

Then, the condition for class assignment will look like this (see formula 6).

$$\forall i \in \{0,1\}: P(C_i|M) = \frac{P(M|C_i)P(C_i)}{P(M)}, P(M) = \sum_i P(M|C_i)P(C_i) \quad (6)$$

The choice of the Bayesian classifier is described by the following rule (see formula 7).

$$P(C_1|M) > P(C_0|M) \Rightarrow deny\ C_0 \quad (7)$$

After substitution, we get the following expression (see formula 8).

$$\frac{P(M|C_1)P(C_1)}{P(M)} > \frac{P(M|C_0)P(C_0)}{P(M)} \Leftrightarrow \frac{P(M|C_1)}{P(M|C_0)} > \frac{P(C_0)}{P(C_1)} \quad (8)$$

Here we see that the left-hand side of the inequality is the likelihood ratio. Assuming the equality of prior probabilities to be true, we obtain the following (see formula 9):

$$LR(M) > 1 \quad (9)$$

which means that the Bayesian classifier, given equal prior probabilities, complies with the Neyman-Pearson lemma with the exponent k=1.

Then, based on the central limit theorem, we obtain the following values (see formula 10):

$$P(M|C_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(M-\mu_i)^2}{2\sigma^2}} \quad (10)$$

Substituting the obtained values into the likelihood index formula, we obtain the following relationship (see formula 11).

$$e^{\frac{(M-\mu_0)^2-(M-\mu_1)^2}{2\sigma^2}} > 1 \Leftrightarrow \frac{-2M(\mu_0-\mu_1)+(\mu_0^2-\mu_1^2)}{2\sigma^2} > 0 \quad (11)$$

Therefore, the optimal value of the classification threshold will correspond to the solution of the equation LR(x)=k. With this value, we minimize the number of type II errors with a fixed number of type I errors. The solution is the arithmetic mean of the mathematical expectations of the two classes (see formula 12):

$$-2M(\mu_0-\mu_1)+(\mu_0^2-\mu_1^2) = 0 \Leftrightarrow M = \frac{\mu_0+\mu_1}{2} = 0.5 \quad (12)$$

In the case of heteroscedasticity, the above conclusions and patterns remain valid with minor adjustments. The Bayesian classifier will still satisfy the Neyman-Pearson lemma (however, with the coefficient of the ratios of prior probabilities) and we can write inequality (11) in the following form (see formula 13).

$$\frac{\sigma_0}{\sigma_1} e^{\frac{(M-\mu_0)^2}{2\sigma_0^2}-\frac{(M-\mu_1)^2}{2\sigma_1^2}} = 1 \Leftrightarrow 2\ln\frac{\sigma_0}{\sigma_1} + \frac{(M-\mu_0)^2}{2\sigma_0^2} - \frac{(M-\mu_1)^2}{2\sigma_1^2} = 0 \quad (13)$$

After simplification, we obtain the following quadratic equation (see formula 14).

$$M^2\left(\frac{1}{\sigma_0^2}-\frac{1}{\sigma_1^2}\right) - 2\left(\frac{\mu_0}{\sigma_0^2}-\frac{\mu_1}{\sigma_1^2}\right)M + \left(\frac{\mu_0^2}{\sigma_0^2}-\frac{\mu_1^2}{\sigma_1^2}\right) + 2\ln\frac{\sigma_0}{\sigma_1} = 0 \quad (14)$$

The solution to this equation provides an estimate of the threshold value of the function in the case of heteroscedasticity. In general, it may deviate from the value of 0.5, but if the following condition is met (see formula 15):

$$\frac{3}{2} \le \frac{\sigma_1}{\sigma_0} \le 2 \quad (15)$$

the solution will still be close to the homoscedastic case, which gives grounds to choose the solution 0.5 as the dividing point of equivalence classes. Other cases are still possible, but clearly distributed – if ratio (15) is closer to 1, then it's clearly converted to homoscedastic case. Ratio greater than 2 usually marks a huge amount of noise which can affect both the clarity of the method and threshold value [12].

Then, let us compare one sample against a group. The method involves calculating two main values for such a scenario - the normalized contribution of positive matches and the normalized weighted average estimate. These metrics are calculated as follows (see formulas (16) - (17)).

$$\forall s \in S, G_s \subseteq S: \exists P \subseteq G_s: \forall p \in P: M(s,p) \ge \theta \quad (16)$$

$$\forall s \in S: k(s,G_S) = \frac{\overline{P}\cdot|P|}{|G_s|}, m(s,G_S) = \frac{\overline{G_s}\cdot|P|}{|G_S|} \quad (17)$$

where G_s is the group of samples against which the comparison is made, P is the set of samples of the group that have passed positive identification, k is the normalized contribution of positive matches, m is the normalized weighted mean score.

Then we formulate the following rule for evaluating a sample against a group: if for a sample s and a group G_s the corresponding values of the normalized contribution of positive matches and the normalized weighted average score satisfy the threshold value, then the sample should be considered to belong to the group (formula 18).

$$\forall s \in S, G_s \subseteq S: k(s,G_S) \ge \theta_1 \wedge m(s,G_S) \ge \theta_2 \Leftrightarrow s \in G_s \quad (18)$$

This method has a number of advantages over conventional binary classification metrics applied according to the one-to-one rule. First, it takes into account the nature of the sample comparison. On the other hand, it reflects the contribution of positive comparisons, which allows, on the one hand, to evaluate the performance of the model from the point of view of positive recognitions, and on the other hand, to focus on errors of the first kind, which have significantly greater weight in fingerprinting and biometric recognition than errors of the second kind.

Experiments

To evaluate the proposed method, we tried it out on the FVC2000 (DB1_B) dataset, which anyone can access. We compared fingerprints one at a time. We picked a method that uses small features (minutiae) as our baseline. It got an F_1 score of 0.73, which is okay for this kind of problem. We know this one-to-one comparison model isn't the best, but we chose it because of what we needed for our specific situation: comparing one fingerprint against a group of others. We figured that if a model is great at telling apart two individual fingerprints, it should also do well when comparing one fingerprint to a group. This is because it can pick up on very small differences and even usual averaging approach over groups should give high quality value. Our minutiae-based approach looks at specific details in a small area of the fingerprint. We think this means the model can handle comparing to groups, where the trick is to tell a single fingerprint apart from many possibilities. The key is that the model needs to be good at spotting those small differences exactly over groups which can allow to use less computational resources over 1 vs 1 comparison.

To make sure our method works in different situations, we set up our experiment carefully, with training and testing phases. First, we adjusted two threshold parameters (θ_1 and θ_2) on the DB1_B dataset from FVC2000. We wanted to get the best possible F-β score (with $\beta=0.5$) while keeping the False Positive Rate (FPR) as low as possible. We ended up with optimized thresholds on DB1_B. Then, we tested these thresholds on a different dataset, DB2_B, from the same FVC2000 collection, collected from a different optical scanner. This way, we could be sure that our parameters weren't just working well on the training data and could also work on new fingerprint samples.

When examining the experimental setup in more detail, it is worth noting that the FVC2000 (DB1_B) dataset is generally recognized as a standard set for evaluating fingerprint recognition algorithms. It was chosen because it contains a sufficient number of samples obtained under controlled conditions, which allows for an objective comparison of different methods. The one-to-one comparison scheme involves matching each fingerprint with another to determine their similarity, which is a classic approach in biometric identification. The comparison method is based on the analysis of special points on the fingerprint, such as the ends and branches of the ridges, which form a unique pattern for each person. The F_1 score of 0.73 indicates a certain level of accuracy of the method in the classification task, but leaves room for improvement.

The fingerprint sample shown in the figure allows us to visually assess the quality of the data used in the study. It is important to note that the quality of fingerprints directly affects the performance of recognition algorithms, so preprocessing and image enhancement play an important role in the fingerprinting process. Further research can be aimed at studying the influence of various factors, such as image quality, algorithm parameters, and characteristics of the datasets used, on the overall performance of the fingerprint recognition system.

Fig. 1. Fingerprint sample from the dataset

In the course of the conducted research, suitable threshold values for the developed metrics were determined. In particular, the best threshold value for the normalized contribution of positive matches was 24%, and for the normalized weighted average score - 1%.

At first glance, these numbers may seem small. This is explained by the way they are calculated. In reality, the number of matches within individual groups of elements is often small, which leads to a downward bias in the positive contribution.

However, it can be easily shown that these values are justified. Let's assume that all the comparison results within a group are, for example, 49%. In this case, both metrics will show 0, since it is the positive contribution that is important. The more positive matches in a group, the higher the values of the proposed metrics will be. In other words, it is not the overall percentage of matches that is important, but the proportion of those matches that exceed a certain predetermined threshold.

The threshold values were determined using a graphical representation of the $F_{0.5}$ measure. The graph (Figure 2) shows the calculation results projected onto a two-dimensional plane for clarity. In other words, the visualization of the graph allowed us to find a balance between accuracy and recall, which is reflected in the selected threshold values.

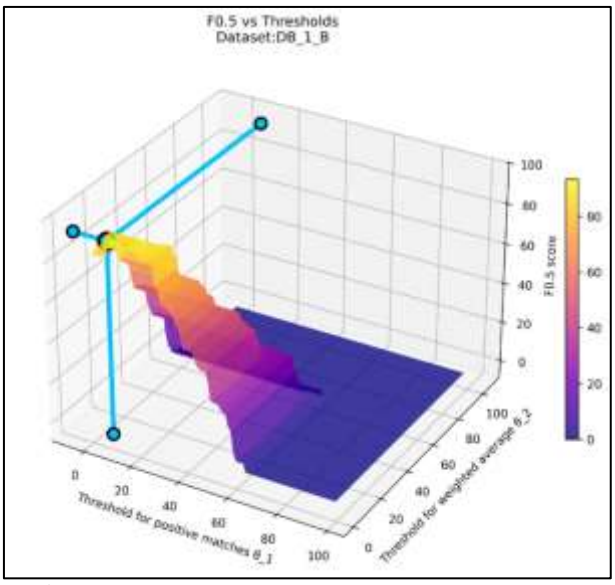

Fig. 2 Graph of $F_{0.5}$ measure values on DB_1_B depending on the increase in thresholds

Of course, the analysis of the F1-measure allows us to identify interesting dynamics: up to a certain point, its value increases, which indicates an improvement in the quality of the model, but then the opposite trend is observed – the indicator begins to decrease. This fact leads us to a logical conclusion about the need to select the optimal threshold corresponding to the maximum value of the F1-measure.

At the same time, the specifics of fingerprinting impose additional requirements on the decision-making process. Fingerprinting, as an area closely related to personal identification, requires special attention to minimizing type I errors, that is, false positives. Using only the F1-measure, which takes into account the balance between accuracy and completeness, may be insufficient, since it does not fully reflect the cost of a type I error in this subject area.

Therefore, in addition to analyzing the F1-measure, it is necessary to carefully study the dynamics of changes in false positive results. It is important to understand how changing the threshold affects the frequency of type I errors and to take this factor into account when choosing the optimal value. The visual representation of the results of the empirical analysis, presented in Figure 3, can help us with this. Studying the graphs of changes in both the F1 measure and the proportion of false positives will allow us to find a compromise solution that will ensure not only high identification quality, but also a minimum risk of erroneous identification of a person.

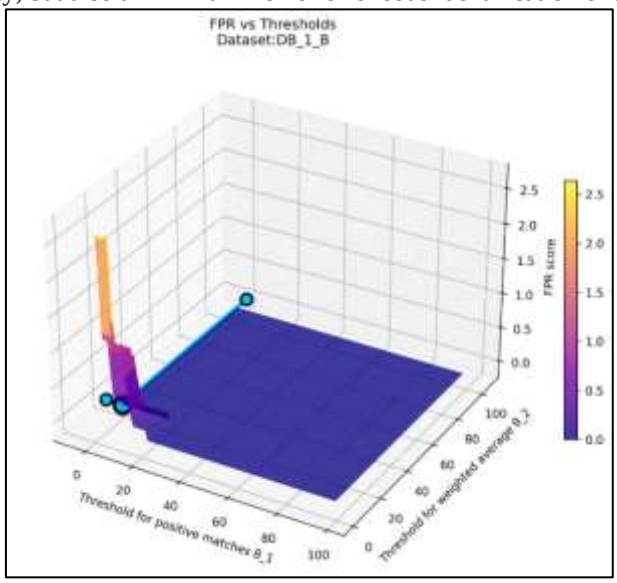

Fig. 3. Graph of FPR values on DB_1_B depending on thresholds

The suggested method seeks a balance between a high $F_{0.5}$ score and few false positives, which is key for comparing fingerprint samples in real-world uses. The method is because it chooses good threshold values for positive matches (θ_1) and weighted averages (θ_2). The data shows an F-β score (β=0.5) of 93.81% with a False Positive Rate (FPR) of 0.13%, showing good precision and recall. It's possible to lower the FPR to 0%, removing

false positives, but the F-β score drops to about 84%. This trade-off shows the method can be changed to focus on either lowering errors or having good classification, depending on what is needed.

The 3D plot, presented in Figure 4, shows how the thresholds (θ_1 and θ_2) relate to the F-β score. The red line in the plot shows the F-β score when the FPR is lowest, giving a clear sense of the top threshold setup for lowering false positives. This line helps show how changing θ_1 and θ_2 affects the balance between precision and recall, allowing careful control over the fingerprint comparison. By showing the F-β score at the lowest FPR, the plot points to how well the method can keep good performance while cutting down wrong matches, which matters in security situations. These results support that the suggested method is efficient and can adapt, making it good for group fingerprint comparisons where accuracy and error control are very important.

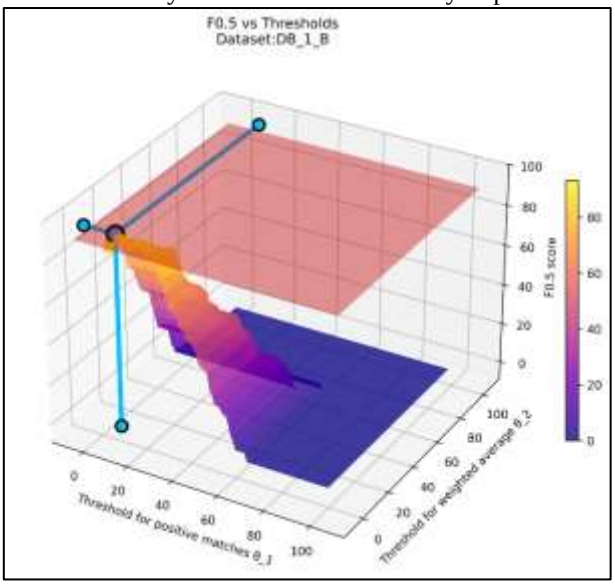

Fig. 4. Graph of $F_{0.5}$ values on DB1_B maximized over FPR minimum value

To assess the performance of the threshold values (θ_1 and θ_2), we conducted tests using the DB2_B dataset from FVC2000. We applied threshold values derived from the DB1_B dataset of FVC2000 where these values helped to minimize Type I errors. This step aimed to ensure that the thresholds, once determined on DB1_B, could be reliably applied to fresh datasets while sustaining a good equilibrium between the F-β score and the rate of errors.

Figures 5 and 6 include 3D graphs that show how these thresholds did on the DB1_B dataset. A red line on the graphs indicates the F-β and False Positive Rate (FPR) score at the point where the error rate was at its minimum. The assessment revealed that the thresholds carried out well on the DB2_B test dataset.

This consistent performance suggests their suitability for fingerprint comparison across varied datasets. These results point to the reliability of the method, suggesting its potential value in group fingerprint identification scenarios, fields where accuracy and low error rates are critical. The ability to keep performance across different datasets backs up the idea that this method could be used in real-world fingerprint recognition systems, making it a helpful tool for managing and analyzing fingerprint data from different sources.

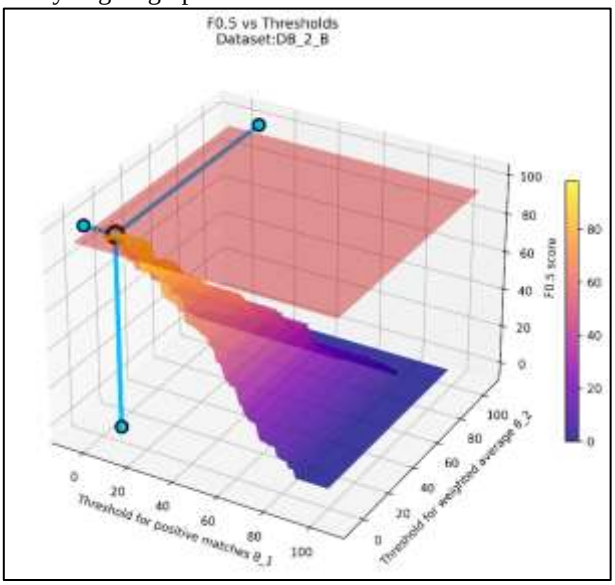

Fig. 5. Graph of $F_{0.5}$ values on DB2_B optimized over DB1_B

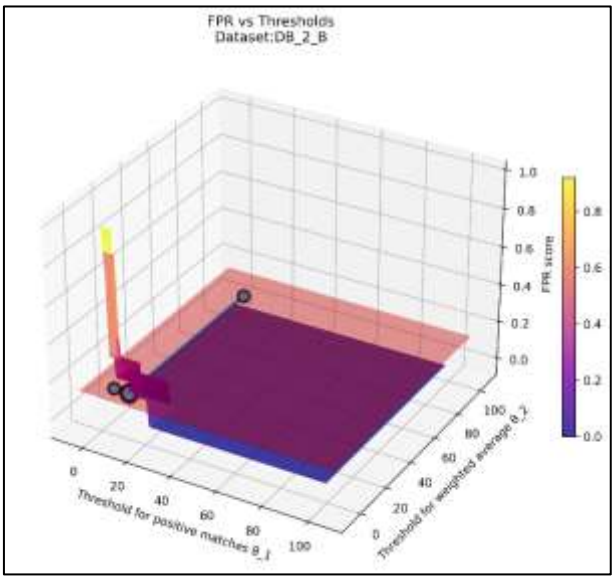

Fig. 6. Graph of FPR values on DB2_B optimized over DB1_B

The data shows that the thresholds adjusted on the DB1_B dataset do well on the DB2_B dataset from FVC2000, which confirms that the method works in varied situations. Even though the False Positive Rate (FPR) on DB2_B isn't zero (it's 0.13%), the $F_{0.5}$ score of 89.3% is better than what we see on DB1_B, meaning the test dataset has better precision and recall. This movement of thresholds from DB1_B to DB2_B shows that the method can keep accuracy high and mistakes low on different datasets, even if gathered with different optical scanners. The 3D plots in Figures 5 and 6 show this visually, with the red line showing the F-β score at the point of lowest FPR on DB1_B, giving a reference for the threshold setup. These outcomes prove that the method is good for group fingerprint comparisons, where cutting down on type I errors while keeping good classification quality is key. The steady behavior across datasets makes a base for the tests ahead, where we look at how the method acts under changing conditions and measure how it works with test dataset not used in the research.

Discussion

The method suggested for judging how good fingerprint comparisons are by using Boolean metrics shows real progress in identifying biometrics within groups, especially when one fingerprint is checked against a bunch of others at the same time. By using normalized positive matches and weighted average scores, the approach balances getting good scores (89.3% on one test) with keeping mistakes low (0.13%). This way of checking things not only makes sense—letting experts see what each match adds—but also does better than old ways that depend on small details (F1 score of 0.73) when matching one fingerprint against a group. It also worked well on another set of data, getting a high score of 89.3%, which proves it can work with data from different scanners. These results agree with what's current in combining different biometrics, where cutting down on one type of mistake is key for things like AFIS integration that need to be secure.

But, when we look closer, there are some problems that limit how soon we can use this method and show us how to improve it. The main issue is that it can be fooled by really messed-up fingerprints, which the paper admits. While it handles some noise okay—because it's made to be general and uses a certain standard—big problems (like bad scans or damaged prints) can mess up the Boolean match results. When that happens, comparing pairs of fingerprints might send the wrong message up to the group level, making things look better or worse than they are. This is a bigger deal in real life when fingerprints are often unclear. Experiments using high-quality samples looked good, but using worse data might lower scores.

Another minus is that it takes a lot of computer power to compare big groups. Even though it uses efficient tools, the way it scales with group size could be too much for some systems. Other methods using deep learning can do it faster, even if they aren't as easy to understand. Also, the set rules (24% for positive matches and 1% for weighted averages) don't change depending on the situation—like different amounts of noise. If things aren't standard, the method might not be as good at telling true positives from false ones, maybe missing differences in patterns from different groups of people.

These limits don't ruin the idea, but they suggest ways to make it better. Future versions could use smart ways like reinforcement learning to change settings based on how good the input is. Adding neural networks to find key features could help with messed-up fingerprints, making a mix-and-match system that's understandable but still uses new tech to find details. Plus, using this with other biometrics (like eyes or veins) could lower risks and fit with privacy ideas like encryption. Testing on different real datasets would give us a better idea of how well it works and could make it ready for use in forensics.

In short, this method is great for well-managed tests and has a strong base, but fixing its sensitivity to bad fingerprints and making it scalable is key for wider use. By mixing classic stats with new tech, it starts a good move toward more reliable biometric systems, which will make personal identification more trustworthy.

Conclusions

The proposed method for assessing the quality of fingerprint comparison based on the convolution of Boolean metrics shows good results both when comparing two samples and when comparing one sample with a group. Experiments with FVC2006 databases and artificially created data confirm several important points:

1. High accuracy in determining matches in the "one against group" format: the new method allows comparing two fingerprints with high accuracy. On real data, the accuracy is 93.81%. The improvement is achieved by using Boolean indicators and the threshold value θ.

2. Theoretical justification: the threshold value $\theta = 0.5$ for comparing two fingerprints is scientifically justified using the Bayesian approach and the Neyman-Pearson lemma. This is true both for cases when the noise is the same for all fingerprints and when the noise is different. For small differences in the noise level between fingerprints, the threshold remains close to 0.5, which is confirmed by calculations and other studies in the field of biometrics.

3. Practical applicability: the method is implemented in Python using the Pandas and NumPy libraries, which allows it to be easily integrated into automated fingerprint identification systems (AFIS). Normalization of metrics and accounting for the number of positive matches make the method understandable and applicable in forensics.

4. Limitations and directions for further research: the method can be sensitive to strong noise. Further research can be aimed at adapting the threshold values using machine learning and integrating it with deep neural networks to improve the process of feature extraction from fingerprints.

Overall, the proposed method is a new approach to fingerprint comparison that provides high accuracy and robustness and can be used in modern biometric systems.

References

1. ISO/IEC, 2017. Information technology — Biometric sample quality — Part 4: Finger image data. ISO/IEC 29794-4:2017. Geneva: International Organization for Standardization.

2. M. Micheletto & G. Luca Marcialis, "Balancing Accuracy and Error Rates in Fingerprint Verification Systems Under Presentation Attacks With Sequential Fusion," in *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 6, no. 3, pp. 409-419, July 2024, DOI: 10.1109/TBIOM.2024.3405554.

3. X. Yin, S. Wang, Y. Zhu & J. Hu, "A Novel Length-Flexible Lightweight Cancelable Fingerprint Template for Privacy-Preserving Authentication Systems in Resource-Constrained IoT Applications," in *IEEE Internet of Things Journal*, vol. 10, no. 1, pp. 877-892, January 2023, DOI: 10.1109/JIOT.2022.3204246.

4. S. Ayeswarya & K. J. Singh, "A Comprehensive Review on Secure Biometric-Based Continuous Authentication and User Profiling," in *IEEE Access*, vol. 12, pp. 82996-83021, 2024, DOI: 10.1109/ACCESS.2024.3411783.

5. B. Ferik, L. Laimeche, A. Meraoumia, O. Aldabbas, M. AlShaikh & A.Laouid, "A Multi-Layered Security Framework for Medical Imaging: Integrating Compressed Digital Watermarking and Blockchain," in *IEEE Access*, vol. 12, pp. 187604-187622, 2024, DOI: 10.1109/ACCESS.2024.3514668.

6. D. Palma & P. Luca Montessoro, "For Your Eyes Only: A Privacy-Preserving Authentication Framework Based on Homomorphic Encryption and Retina Biometrics," in *IEEE Access*, vol. 12, pp. 183688-183706, 2024, DOI: 10.1109/ACCESS.2024.3512003.

7. A. Alarifi, M. Amoon, M. H. Aly & W. El-Shafai, "Optical PTFT Asymmetric Cryptosystem-Based Secure and Efficient Cancelable Biometric Recognition System," in *IEEE Access*, vol. 8, pp. 221246-221268, 2020, DOI: 10.1109/ACCESS.2020.3043689.

8. M. I. M. Yusop, N. H. Kamarudin, N. H. S. Suhaimi and M. K. Hasan, "Advancing Passwordless Authentication: A Systematic Review of Methods, Challenges, and Future Directions for Secure User Identity," in *IEEE Access*, vol. 13, pp. 13919-13943, 2025, DOI: 10.1109/ACCESS.2025.3528960.

9. S. Wang and N. S. Ahmad, "A Comprehensive Review on Sensor Fusion Techniques for Localization of a Dynamic Target in GPS-Denied Environments," in *IEEE Access*, vol. 13, pp. 2252-2285, 2025, DOI: 10.1109/ACCESS.2024.3519874.

10. A. Aboalhsan and M. N. Alatawi, "Deep Learning Technique for Fingerprint Recognition," *2022 2nd International Conference on Computing and Information Technology (ICCIT)*, Tabuk, Saudi Arabia, 2022, pp. 340-343, DOI: 10.1109/ICCIT52419.2022.9711634.

11. U. U. Deshpande, V. S. Malemath, S. M. Patil, and S. V. Chaugule, "CNNAI: A Convolution Neural Network-Based Latent Fingerprint Matching Using the Combination of Nearest Neighbor Arrangement Indexing," Frontiers in Robotics and AI, vol. 7, 2020, DOI: 10.3389/frobt.2020.00113.

12. M. El-Abed, C. Charrier & C. Rosenberger, "Quality assessment of image-based biometric information," *J Image Video Proc.* 2015, 3 (2015). DOI: 10.1186/s13640-015-0055-8

Yurii Pohuliaiev **Погуляєв Юрій**	Postgraduate student at the Department of Software Engineering, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine, e-mail: yurii.pohuliaiev@nure.ua, https://orcid.org/0009-0005-5883-1661	аспірант каф. програмної інженерії, Харківський національний університет радіоелектроніки, Харків, Україна.
Kirill Smelyakov **Кирило Смеляков**	Doctor of Technical Sciences, Professor, Head of the Department of Software Engineering, Kharkiv National University of Radioelectronics, Kharkiv, Ukraine, e-mail: kyrylo.smelyakov@nure.ua, https://orcid.org/0000-0001-9938-5489	докт. техн. наук, проф., зав. каф. програмної інженерії, Харківський національний університет радіоелектроніки, Харків, Україна.