

OPTIMIZATION OF THE DECISION THRESHOLD FOR APPLICATION IN BIOMETRIC SYSTEMS UNDER EXTREME CONDITIONS OF FACIAL IMAGE VERIFICATION

Nenad Badovinac¹

Abstract: The accuracy of biometric facial recognition systems depends on a well-defined decision-making threshold that controls the rate of false positive verification results of unregistered users (FAR) and false negative results of verification of registered users (FRR). The decision threshold defines the balance between the security of the system and a better user experience. Greater security means a reduction in the ability to accept unauthorized users (low FAR), while a better user experience means a reduction in the number of unauthorized user rejection errors (low FRR). In this paper, three adapted methods for determining the decision threshold in biometric systems are presented, the methods are adapted for application in extreme conditions, such as facial recognition of users with masks. Method 1 uses the Gaussian optimization process. Method 2 is used to create a decision threshold based on the defined relationship between the FAR and FRR errors. Method 3 is based on the Grid Search algorithm. The results of the research show the potential of these methods with the possibility to use one of three adapted methods of calculating the decision threshold, depending on the implementation.

Key words: Threshold decision-making, biometric facial recognition, Bayesian optimization, Grid Search.

1. INTRODUCTION

Biometric facial recognition is a biometric modality that can be used in the process of biometric verification, especially in the context of highly demanding secure systems such as electronic payment systems or a system for accessing a military-security complex. While facial recognition technology offers many benefits, challenges arise such as determining the correct decision threshold that affects system performance, Accuracy, False Acceptance Rate (FAR), and False Rejection Rate (FRR). In this paper, three adapted methods for the process of calculating the decision threshold for application in biometric systems are presented. A Sejong database of masked faces of 64 people was used to evaluate these methods [1]. The images are divided into an 80% training group and a 20% test group. The traditional approach to determining the decision threshold is the average value of the distance between the biometric template of images of the same user's face and the average distance for different users. The results presented in the paper demonstrate the limitations of the traditional calculation approach, especially when applied in complex situations such as masked faces. Existing methodologies, such as Gaussian Process Optimization [2], and Grid Search [3] were adapted to explore the threshold within a defined range, while performance was measured by the Accuracy metric. The development of innovative methods for calculating the decision threshold improves the performance of the biometric system. The paper is organized in such a way that the introductory chapter is followed by a chapter with an overview of previous scientific research. The third chapter describes the adapted methodologies. The results of the research are in the fourth chapter, while the fifth chapter is given a conclusion and the sixth chapter contains a list of literature.

2. AN OVERVIEW OF PREVIOUS RESEARCH

The scientific literature in the field of biometric verification and methods for facial recognition, research on geometric features of faces, as well as for optimization of decision thresholds, which achieves the improvement of the security and accuracy of the biometric system has been researched [5]. There is a lot of research that deals with these topics. For example, in paper [6], the authors use the firefly optimization algorithm in combination with deep neural networks to recognize and classify facial images, while optimizing decision thresholds increases the accuracy and reduces the error rate of

¹PhD student at University in Belgrade, Faculty of Organizational Sciences, email: nenad.badovinac@gmail.com

biometric recognition. In the paper [7], the authors use a genetic algorithm, which is inspired by natural selection and evolution, and which combines multiple deep learning algorithms for facial recognition. The authors of the paper suggest that the decision threshold should be automatically adjusted during the model training process. Paper [8] analyzes the fusion of biometric modalities in which the method of calculating the decision threshold is presented. Paper [9] explores the Reinforcement Learning algorithm, as a machine learning algorithm for optimizing the decision-making threshold in which the system learns through feedback in the form of rewarding and punishing. In paper [10], the authors implement deep neural network algorithms for facial recognition on mobile devices. In paper [11], the authors explore the possibilities of fusion of different biometric modalities, such as fingerprints and facial images. A review of the available literature has shown that the optimization of decision thresholds is beneficial due to the need to increase the accuracy and security of biometric systems.

3. RESEARCH METHODOLOGY

The implementation of a biometric verification system requires the definition of a decision threshold that achieves certain rates of false positive authentication results, i.e. False Acceptance Rate (FAR) and false negative authentication results, i.e. A registered user's rejection rate (FRR). In this sense, the decision threshold affects the performance of the biometric system. In this paper, three existing methods of calculating the decision threshold are adapted. The methods make it possible to achieve a certain rate of FAR and FRR errors, as well as the required level of accuracy (Accuracy). Method 1 uses the famous Gaussian Process Optimization [2] which is a well-known technique for optimizing hyperparameters. Method 2 uses a custom target function used to balance metrics in systems [4], this method combines FAR and FRR with weighting factors α and β , while Method 3 uses Grid Search hyperparameter optimization [3]. Customized methods for decision threshold optimizations consist of three basic steps:

- Generating a biometric template of masked images of faces
- Creating pairs of face images for same and different people
- Optimization of the decision threshold using one of three customized methodologies, based on existing techniques.

3.1. Generating a biometric template from a facial image

Facial images are processed by an algorithm to create a biometric template. The result of the processing is a 128-dimensional vector that describes the characteristics of the face. A biometric template is generated and stored in a .txt file. Python programming code has been developed for this image processing. First, a folder is processed from which the face images are divided into training and test subfolders, namely the training set (80% of the face images) and the test set (20% of the face images). Folders with biometric vectors have an equal number of vectors in both sets. Biometric templates are generated using the *face_recognition library*.

3.2. Formation of pairs of images of the same and different people's faces

In this step, the algorithm creates pairs of biometric vectors, combining vectors from the same folders belonging to the same people for the same person, while for different people, vectors from different folders belonging to different people are combined. Each pair of images is assigned a "1" for the same person, or a "0" for different people. In this step, pairs of biometric vectors are generated based on their folders. For example, for the same person, images of a face are created tagged with: (Face_1_1, Face_1_2, Face_1_3). Pairs of vector images of faces are organized into groups and are pairs of images of the same person, such as: (Face_1_1, Face_1_2), (Face_1_1, Face_1_3), (Face_1_2, Face_1_3). Or pairs of images of different people's faces, for example: (Face_1_1, Face_2_1), (Face_2_1, Face_2_2) from which the following pairs are created: (Face_1_1, Face_2_1), (Face_1_1, Face_2_2), (Face_1_2, Face_2_1), (Face_1_2, Face_2_2). Creating pairs of images of the same person's faces makes it possible to measure accuracy in recognizing faces when the images belong to

the same person, while images of different people's faces make it possible to assess the algorithm's ability to distinguish between people.

3.3. Research methodology

The methods presented in this paper make it possible to adjust the threshold in conditions when faces are covered with masks or various types of masking, and then faces are difficult to recognize. In such cases, the biometric system requires high security and it is necessary that the FAR error is as low as possible and that the accuracy is as high as possible. Three methods have been proposed, namely, the first method uses Bayesian optimization of the decision threshold, the second method optimizes the threshold by balancing the FAR and FRR errors, and the third method uses the Grid Search algorithm to precisely determine the range in which the threshold values are located.

3.3.1. Method 1 – Bayesian optimization for determining the decision threshold

The traditional threshold is determined based on the average value of the Euclidean distances between biometric vectors of images of the same person's face and images of the faces of different people. A method has been developed that uses a Bayesian algorithm to find the optimal threshold. The algorithm provides improved performance compared to the decision threshold that was created by the traditional method. This approach iteratively adjusts the threshold value to achieve maximum accuracy. A Bayesian approach with a Gaussian process algorithm was used to optimize the threshold, classifying the correctly verified cases (True Positives) and (True Negatives), and (False Positives) and (False Negatives) that were misclassified. The elements of the method are:

- Positive and negative pairs of biometric vectors used for calculation.
- The Gaussian Process algorithm iteratively investigates the space of possible allowable thresholds.
- A function that iterates the performance of the system: Accuracy, FAR, and FRR.

$$T_{opt} = \operatorname{argmax} f(T)_t \quad (1)$$

Where:

T_{opt} – It represents the decision threshold at which the performance of the $f(T)$ system is maximized.

$\operatorname{argmax} f(T)_t$ – threshold value, argument T for which $f(T)$ is the largest possible.

$f(T)_t$ – The function measures performance for the threshold T and $f(T)$ which indicates Accuracy

The recommendation to use this method is in high-security systems where it is important to minimize the FAR error to prevent unauthorized access. It can also be used to control access to infrastructure facilities, such as military bases, where transactions are high-risk. Due to the high FRR error, there are limitations to the application of this method.

3.3.2. Method 2 – FAR and FRR balancing to optimize the threshold

The method allows the weighting for FAR and FRR errors to be adjusted according to specific requirements, such as more safety or higher tolerance for errors. The goal is to minimize authentication errors, whereby a higher weight is given to a certain value (e.g., a reduction in FAR). Optimization is achieved by calculating the balance between security (low FAR) and user experience (low FRR).

$$f(T) = \alpha \cdot \text{FAR}(T) + \beta \cdot \text{FRR}(T) \quad (2)$$

Where:

$f(T)$ – The error minimization function measures system errors for a specific decision threshold (T).

α – The weighting gives relative importance to the reduction of false positive cases (FAR). A higher value α means greater importance of the security of the biometric system.

$FAR(T)$ – False positive verification decision rate for a specific threshold (T). Measures the percentage of cases where an unauthorized user is mistakenly accepted.

β – A weight that gives relative importance to the reduction of false negative decisions (FRR). A higher value β means a greater user experience.

$FRR(T)$ – Rate of false negative verification decisions for a given threshold (T). Measures the percentage of cases where an authorized user was incorrectly rejected. The method is suitable for systems where both security and user experience are important, for example in electronic payment systems, where high FAR and FRR errors are unacceptable, as well as in public service systems: e-government, identification in hospitals, etc.

$$T_{opt} = \operatorname{argmin} f(T)_t \quad (3)$$

Where:

T_{opt} – Optimal decision threshold.

$\operatorname{Argmin} f(t)$ – An operation that looks for the value of T for which $f(T)$ is the lowest.

3.3.3. Method 3 – Decision Threshold Optimization with Grid Search

An optimization method that uses the Grid Search algorithm searches a predefined range of threshold values to find an optimized threshold value. This method allows you to review the performance of the system for a wider range of threshold values. Although it requires more computing resources, the Grid Search algorithm is efficient for accurate analysis. Below is the threshold optimization function:

$$T_{opt} = \operatorname{argmax} f(T)_{T \in [Tmin, Tmax]} Accuracy(T) \quad T \text{ is limited to the interval } [Tmin, Tmax] \quad (4)$$

Where:

T – The optimal decision threshold found by the Grid Search method. This is the value of the threshold $_{opt}T$ that maximizes accuracy ($Accuracy(T)$) within a defined range of values.

$\operatorname{argmax} f(T)$ – An operation that finds the value of T for which the target function ($Accuracy(T)$) reaches its maximum.

$[Tmin, Tmax]$ – The lower (Tmin) and upper (Tmax) limits of the range of threshold values to be searched.

$Accuracy(T)$ – A function that measures the accuracy of the system for a specific threshold T. Accuracy is defined as:

$$Accuracy(T) = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

Where:

TP – True Positives (correctly accepted pairs of the same person).

TN – True Negatives (correctly rejected pairs of different people).

FP – False Positives (incorrectly accepted pairs of different people).

FN – False Negatives (incorrectly rejected pairs of the same person).

The method is suitable for mobile applications and for applications in IoT devices with limited processing capacity. Access control systems in low-risk environments, e.g. office buildings, universities, etc.

4. EVALUATION OF RESULTS

The results of the research show differences in the performance of the three proposed methods for optimizing the decision threshold in masked face recognition systems. Each of the solutions was examined through accuracy evaluation, through false positive verification case (FAR) and through false negative verification case (FRR) rates. The results presented in this paper were obtained using the Python program developed for the purposes of this research. The program is written using libraries for numerical computation and data processing (e.g. NumPy, Scikit-learn, Skopt), while the following code is used to calculate the results of the method:

Method 1 – Part of the Python code to calculate the optimized threshold using the Gaussian process:

```
def optimize_threshold(pairs, labels):
    def objective(threshold):
        predictions = [1 if np.linalg.norm(emb1 - emb2) < threshold else 0 for emb1,
            emb2 in pairs]
        return -accuracy_score(labels, predictions) # Maximizing Accuracy
    result = gp_minimize(objective, [(0.2, 1.5)], n_calls=20)
    return result.x[0]
```

Method 2 – Part of Python code to calculate the optimized decision threshold using the FAR and FRR weighting function:

```
def optimize_threshold_balanced(pairs, labels, alpha=0.5, beta=0.5):
    def objective_wrapper(threshold):
        return objective_custom(threshold[0], pairs, labels, alpha, beta)
    result = gp_minimize(objective_wrapper, [(0.1, 1.0)], n_calls=30)
    return result.x[0]
```

Method 3 – Part of Python code to calculate results using Grid Search Optimization:

```
def optimize_threshold_grid_search(pairs, labels, threshold_range, step=0.01):
    best_threshold = None
    best_accuracy = 0
    best_far = None
    best_frr = None
    for threshold in np.arange(threshold_range[0], threshold_range[1], step):
        FAR, FRR, accuracy = evaluate_threshold(pairs, labels, threshold)
        if accuracy > best_accuracy:
            best_accuracy = accuracy
            best_threshold = threshold
            best_far = FAR
            best_frr = FRR
    return best_threshold, best_far, best_frr, best_accuracy
```

All experiments were performed on a dataset that included extremely masked images of faces [1]. The calculation of the evaluation of the methods showed results, and the evaluation of the first method shows accuracy with an optimized threshold of 98.78%, while the FAR decreases to 0%, at an FRR error rate of 76.53%. In the second method, balancing the FAR and FRR errors increased the FAR error to 5.72% to reduce the FRR error to 36.63%. The third method, using the Grid Search algorithm, obtained a range of threshold values with an optimal threshold that gives an accuracy of 81.9%. FAR and FRR errors are relatively uniform in this method, at 18.1% and 15.6%, respectively. Table 1 shows the results of the evaluation of the three methods presented.

Table 1 – Evaluation of results

Methods	Accuracy	FARR	FRR
Traditional method	88,33%	10,95%	25,41%
Method 1	98,78%	0,1 %	76,53%
Method 2	93,83%	5,72%	36,63%
Method 3	81,90%	18,10%	15,60%

Table 1 shows how Method 1 achieves the highest accuracy thanks to Bayesian optimization, but since a high FRR value is obtained, some solutions will require the use of the second or third threshold calculation method. Method 2 allows the calculation of the threshold with a balanced FAR and FRR error, while Method 3 offers flexible performance but with lower accuracy.

5. CONCLUSION

In this paper, three methods for optimizing the decision threshold in biometric systems for the recognition of masked faces are presented. Research has shown that each method has specific advantages and applications, depending on the requirements of the implementation of the biometric application. Method 1, which uses Bayesian optimization, achieved the lowest FAR error value, making it ideal for high-security applications, such as access control on military bases. However, there is a high FRR error value and it limits the application in systems that require a high level of user experience. Method 2, focused on balancing FAR and FRR errors, allows for greater control over system performance. This method is particularly suitable for use in electronic payment systems and public services, where it is equally important to reduce both false positives and false negatives. Method 3, based on the Grid Search algorithm, provides stable performance and a balance between FAR and FRR errors. Lower resource requirements make this method suitable for mobile applications and IoT devices. In conclusion, the proposed adaptation of existing methods significantly improves the performance of biometric systems in extreme conditions, allowing adaptation to different scenarios. Future research should focus on integrating these methods with multimodal biometric systems to further increase accuracy and security in complex applications.

6. REFERENCES

- [1] Usman Cheema, Seungbin Moon, Sejong face database: *A multi-modal disguise face database*, *Computer Vision and Image Understanding*, Volumes 208–209, 2021, ISSN 1077-3142, <https://doi.org/10.1016/j.cviu.2021.103218>.
- [2] Bıyık E, Huynh N, Kochenderfer MJ, Sadigh D.: *Active preference-based Gaussian process regression for reward learning and optimization*. *The International Journal of Robotics Research*. 2024; 43(5):665-684. doi:10.1177/02783649231208729
- [3] Aratrika Ray-Dowling, Daqing Hou, Stephanie Schuckers, Abbie Barbir: *Evaluating multi-modal mobile behavioral biometrics using public datasets*, *Computers & Security*, Volume 121, 2022, 102868, ISSN 0167-4048, <https://doi.org/10.1016/j.cose.2022.102868>.
- [4] M. Nasr, A. Piórkowski, P. M. Mah and F. E. Abd El-Samie, *Cancelable Biometrics Based on Quaternion Adaptive Filter Optimization for Security Applications*, 2024 6th International Conference on Computing and Informatics (ICCI), New Cairo - Cairo, Egypt, 2024, pp. 490-498, doi: 10.1109/ICCI61671.2024.10485055.
- [5] Vasanthi, K. Seetharaman: *Facial image recognition for biometric authentication systems using a combination of geometrical feature points and low-level visual features*, *Journal of King Saud University - Computer and Information Sciences*, Volume 34, Issue 7, 2022, Pages 4109-4121, DOI:10.1016/j.jksuci.2020.11.028.

- [6] P. Annamalai: *Automatic Face Recognition Using Enhanced Firefly Optimization Algorithm and Deep Belief Network*, International Journal of Intelligent Engineering and Systems, 2020, 13, Intelligent Networks and Systems Society, DOI: 10.22266/ijies2020.1031.03
- [7] MS Sannidhan, Jason Elroy Martis, KN Pallavi, Vinayakumar Ravi, H. L. Gururaj, Tahani Jaser Alahmadi: *Genetic algorithms and deep learning for unique facial landmark-based key generation*, Computers and Electrical Engineering, Volume 118, 2024, DOI:10.1016/j.compeleceng.2024.109427.
- [8] T. Edwards and M. S. Hossain: *Effectiveness of Deep Learning on Serial Fusion Based Biometric Systems*, IEEE Transactions on Artificial Intelligence, vol. 2, no.1, pp. 28-41, 2021, doi: 10.1109/TAI.2021.3064003.
- [9] P. Wang, W. H. Lin, K. M. Chao and C. C. Lo: *A Face-Recognition Approach Using Deep Reinforcement Learning Approach for User Authentication*, IEEE 14th International Conference on e-Business Engineering (ICEBE), Shanghai, China, 2017, pp. 183-188, DOI: 10.1109/ICEBE.2017.36.
- [10] Ríos-Sánchez, B., Silva D.C., Martín-Yuste, N., Sánchez-Ávila, C.: *Deep learning for face recognition on mobile devices*. IET Biom., 109-117. 2019. DOI:10.1049/iet-bmt.2019.0093
- [11] Hadis Heidari, Abdollah Chalechale: *Biometric authentication using a deep learning approach based on different level fusion of finger knuckle print and fingernail*, Expert Systems with Applications, Vol. 191, 2022, ISSN 0957-4174, <https://doi.org/10.1016/j.eswa.2021.116278>.