From Facial Identification to Clothing-Based Person Re-Identification: An Interpretable Framework for Identity Continuity in Multi-Camera Surveillance Systems

José Luis Carrillo-Medina[1] [0000-0001-5222-4039], Ivone Arias-Almeida[1] [0000-0001-7351-2013], Kevin Jair Chuquitarco-Velasco[1] [0009-0000-3877-7227], Cristian Santiago Viteri-Arias[1] [0000-0001-8877-4649], Nancy Jacho Guanoluisa[1] [0009-0004-9452-3810], and Marco Flores Calero[1] [0000-0003-1597-3385]

Universidad de las Fuerzas Armadas ESPE, Sangolquí, Ecuador {jlcarrillo, giarias, kjchuquitarco, csviteri1, npjacho, mjflores}@espe.edu.ec
http://www.espe.edu.ec

Abstract. Re-identifying individuals in multi-camera video surveillance requires maintaining identity continuity when facial evidence is no longer available, making it necessary to integrate biometric and soft-biometric signals within a single decision-making framework. A comprehensive approach is proposed that combines face and person detection, facial identification, and re-identification based on clothing color and texture, integrating HoG, HSV, LBP, and SVM into a unified pipeline designed to maintain identity under real-world capture conditions. The results show that the most robust module in the system is the facial component. In the re-identification stage, the descriptor based on clothing color performed best among the attributes analyzed, with an F1-score of 0.7410, surpassing both the combination of color and texture and the exclusive use of texture. Even though the integration of complementary sources is a methodologically sound strategy, the findings suggest that merging attributes does not always lead to better performance, especially when one of the individual descriptors has greater discriminatory power within the closed-circuit camera system. The study provides an interpretable framework for the identification and re-identification of individuals and uses experimental evidence to define the actual role of each visual source within the identification process.

Keywords: Person re-identification · facial recognition · soft biometrics · SVM · multi-camera systems

1 Introduction

In modern video surveillance systems, identifying and tracking individuals across multiple cameras poses a major challenge, especially in scenarios where there is no overlap between the fields of view [1]. In these environments, the ability to maintain identity continuity is limited by the loss of direct biometric information, such as facial features, which necessitates the incorporation of alternative recognition mechanisms based on complementary visual attributes [2]. Traditionally, facial recognition has proven to be one of the most effective techniques due to its high discriminatory power [3]; however, its performance deteriorates significantly when the face is not visible, is partially obscured, or exhibits variations in lighting and perspective [4]. Re-identification of individuals allows for continuous tracking when facial recognition is no longer feasible [5], using attributes such as the color and texture of clothing. Various studies have addressed this problem by using machine learning models and visual descriptors [6]. Nevertheless, treating identification and re-identification as separate processes can affect the consistency of the system, especially in real-world applications where it is necessary to maintain the same identity over time [7]. In this regard, an integrated architecture improves identity continuity by combining multiple information sources in multi-camera environments [8, 9].

This paper proposes a unified architecture that integrates the phases of detection, preprocessing, feature extraction, and supervised classification into a single processing workflow, linking initial facial identification with clothing-based re-identification. The system analyzes faces using HoG descriptors, and when facial recognition is no longer feasible, it falls back on re-identification by combining information about clothing color (HSV) and texture (LBP), relying on an SVM classifier to make decisions. The primary contribution of the study is to compare how these different sources of visual information function, both individually and collectively, in order to understand which one plays the greatest role in maintaining a person's identity over the course of the follow-up period. The results show that, while the integration of soft-biometric attributes is a methodologically sound strategy, its effectiveness depends on the degree to which the descriptors complement one another and on environmental conditions. These results indicate that the integration of soft-biometric attributes is valid, although its effectiveness depends on the complementarity of the descriptors and the environmental conditions.

2 Materials and Methods

The proposed methodology integrates the identification and re-identification of individuals within a unified, non-overlapping multi-camera video surveillance architecture. Table 1 summarizes the resources explicitly reported in the source document and their methodological role within the pipeline. Initial identification is performed using facial recognition, employing YOLOv8n detection, HoG feature extraction and SVM classification. When facial information is unavailable or becomes unreliable, identity continuity is maintained using soft-biometric attributes of clothing: texture using LBP and colour using HSV, fused and clas-

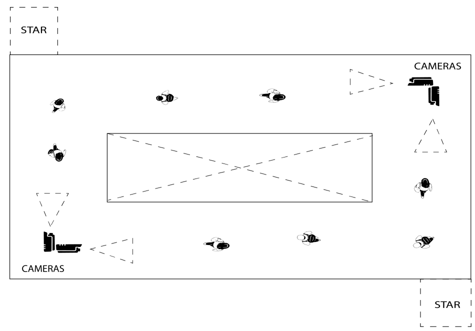

Fig. 1. Non-overlapping multi-camera video surveillance setup for re-identification of individuals

Table 1. System components and their role within the pipeline

Resource	**Methodological function**	**Reported observation**
Python	General system implementation	Primary development language
OpenCV	Image processing and manipulation	Applied with HoG and color histograms
scikit-learn	Training SVM models and handling descriptors	Applied using HoG and color histograms
PyTorch	Data increase	Used for rotation and brightness adjustment
YOLOv8	Face and people detection	System input module
Matplotlib and NumPy	Data visualization and analysis	Support for reviewing results
CRISP-DM	General methodological framework	Guide to data management and modeling

sified using SVM [10]. Thus, the pipeline integrates detection, pre-processing, feature extraction, descriptor combination and supervised classification into a coherent sequence aimed at preserving identity across cameras. Fig. 1 depicts the capture environment: a closed circuit of four cameras positioned at the corners of the scene, with non-overlapping fields of view and a central transit area. As the cameras observe separate areas, it is not possible to rely solely on trajectory; therefore, re-identification requires combining biometric and soft-biometric evidence to maintain the traceability of the individual.

2.1 Materials

The system was implemented in Python, incorporating libraries for computer vision, machine learning, data augmentation and visualisation. OpenCV was used for image processing; scikit-learn for SVM training and feature handling; PyTorch for data augmentation via rotations and brightness adjustments; YOLOv8n for detecting faces and bodies; Matplotlib and NumPy for analysis and visualisation; and CRISP-DM as a methodological guide from data preparation through to evaluation [11].

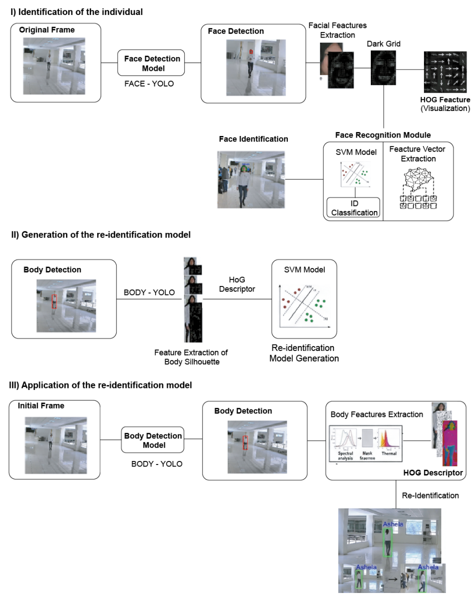

Fig. 2. Architecture of the person re-identification system: I) Individual identification, II) Generation of the re-identification model, and III) Application of the re-identification model

2.2 System Construction

The functional architecture is organised as a pipeline connecting three complementary workflows [12]. First, the system detects the face using YOLOv8n, extracts HoG descriptors and applies an SVM to identify the subject when the face is visible. Second, it detects the full body and generates the re-identification model based on clothing attributes, combining LBP texture and HSV colour into a single vector. Third, when the face is not available, it re-extracts the features of the detected body and uses the trained SVM to recognise the individual on another camera or at another location.

Fig. 2 shows the complete system architecture, comprising three complementary workflows. The first involves identifying the individual via facial recognition, where, based on a sequence of input images, facial detection is performed using YOLO; subsequently, facial features are extracted using the HoG descriptor and an SVM classifier is employed to assign the individual's identity when their face is visible. After establishing the identity in the initial stage, the system moves to a re-identification phase, in which the individual's body is detected and soft-biometric attributes of their clothing are extracted, primarily colour in HSV space and texture using LBP. These features are integrated into a single vector which is used to train a specific SVM model for each individual. At a later stage, this model is used when the face is no longer available, extracting the features from the detected body once again and applying the previously trained classifier to recognise the individual across different cameras or locations, thereby maintaining identity continuity within the multi-camera system.

Table 2. Explicit preprocessing parameters

Stage	**Configuration**	**Purpose**
Facial reshaping	80×100 px	Standardize facial input
Full-body resizing	64×160 px	Standardize body input
Color conversion	Grayscale, HSV	Support HoG/LBP and color features
Data augmentation	Rotation, brightness	Improve variability and generalization

Table 3. Composition of the dataset used for training

Module	Classes	Initial images by class	Initial total	Total after increase
Facial recognition	5	50	250	500
Re-identification	5	100	500	1000

2.3 Preprocessing

Pre-processing standardises visual inputs and reduces variability prior to feature extraction. Faces are resized to 80×100 pixels to stabilise HoG for facial recognition. Full-body images are resized to 64 x 160 pixels to maintain consistency in re-identification. Conversion to greyscale facilitates HoG and LBP, whilst HSV allows colour to be characterised with greater stability in the face of lighting changes. Furthermore, data augmentation through rotations and brightness adjustments increases training variability and promotes SVM generalisation.

2.4 Dataset

The training was structured into five classes corresponding to five identities. For facial identification, 50 initial images per class were used, equivalent to 250 samples, which were expanded to 500 through data augmentation. For re-identification, 100 initial images per class were used, equivalent to 500 samples, expanded to 1,000. This configuration preserves the correspondence between facial identity and physical appearance, and provides greater variability to the re-identification module.

Fig. 3. shows the dataset organised into DATA FACE and DATA BODY for the classes Esteban, Jean, Dalila, Gabriel and Luis. The facial samples include variations in expression, lighting and angle; the body samples incorporate changes in posture, movement and capture conditions. This organisation ensures consistency across modules and provides diversity for evaluating the system under operational conditions.

2.5 Feature extraction

Facial recognition was based on the Histogram of Oriented Gradients (HoG), a descriptor that represents the local structure of the face using histograms of gradient orientations calculated over cells and blocks [13, 14]. This representation emphasises contours and directional variations in intensity, making it suitable for distinguishing facial patterns in a supervised learning framework using SVMs. For re-identification, the colour of the clothing was represented using HSV histograms, as this colour space separates hue, saturation and value and

Fig. 3. Dataset presented in the source document.

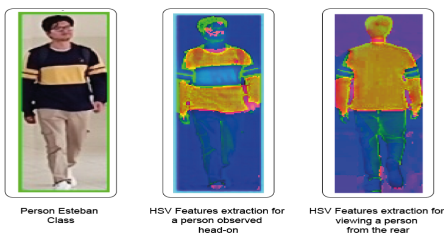

Fig. 4. HSV feature extraction for full body front and rear view of a person

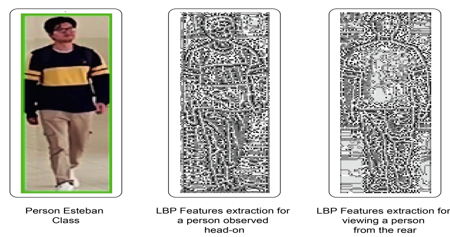

Fig. 5. LBP feature extraction for full body front and rear view of a person

offers greater stability in the face of certain lighting variations than RGB [15]. The images were converted from BGR to HSV, using the three channels with 8 bins per channel; the ranges were 0–180 for Hue and 0–256 for Saturation and Value. Fig. 4 illustrates the extraction of HSV features: the original image of the individual is transformed into chromatic representations of the front and rear views. These views highlight the colour distribution of the clothing, which helps to recognise the individual even when the face is not visible.

In addition, the texture of the clothing was modelled using the Local Binary Pattern Uniform (LBPU), a robust, low-complexity descriptor that compares each central pixel with its neighbours, encodes the resulting binary pattern, and constructs a histogram of textural frequencies [16]. Fig. 5 shows the LBP extraction for front and rear views of the body. Unlike HSV, which emphasises colour distribution, LBP highlights local texture patterns in the clothing. Therefore, both descriptors provide complementary information for re-identification in multi-camera scenarios. The visual representation of an individual is not formulated as a single full-body descriptor, but rather as an integration of complementary cues. The face serves as the primary biometric reference for identification, whilst the colour and texture of clothing maintain identity continuity during re-identification. At the heart of the re-identification module is the concatenation of HSV and LBP features into a single descriptive vector, which allows for the simultaneous incorporation of colour distribution and textural structure during training and inference [17].

2.6 Classification using machine learning algorithms

The final decision was implemented using support vector machines (SVMs), trained on the feature vectors and their identity labels. The classifier was applied to both facial identification and re-identification, using a linear kernel and regularisation $C = 0.1$, a value selected through exhaustive cross-validation after testing multiple configurations. The parameter C controls the balance between error penalty and model complexity; low values favour wider margins and better generalisation on unseen data. Taken together, the methodology operates as a sequential pipeline: YOLOv8n detects faces and bodies; pre-processing normalises the visual information; HoG, LBP and HSV construct discriminative representations; the HSV-LBP combination defines the appearance descriptor; and SVM assigns the identity. When the face is available, identity is established by HoG+SVM; when the face is intermittent or not visible, continuity is maintained by LBP+HSV+SVM.

2.7 Methodological integration of the system

The validation was organised using appropriate metrics for classification and retrieval in scenarios where there may be an imbalance between matches and non-matches. For this reason, accuracy is not considered sufficient on its own; precision, recall and F1-score are incorporated as key measures, alongside the precision-recall curve, to analyse the trade-off between false positives and correct detections [18]. In multi-camera re-identification tasks, this evaluation allows us to assess not only the classifier's point-in-time accuracy, but also the system's consistency in correctly retrieving identity when facial evidence is no longer available.

3 Results

This section presents the experimental performance of the proposed system for person identification and re-identification, distinguishing four levels of analysis: facial identification, re-identification based on clothing color, re-identification based on clothing texture, and evaluation of the combined color–texture representation. It highlights the response of the complete system under real-world capture conditions within the proposed camera network.

3.1 Overall results of the proposed system

The evaluation framework includes an initial identification stage based on facial features and a subsequent re-identification stage based on soft-biometric attributes of clothing. This structure allows us to examine, on the one hand, the robustness of the facial component and, on the other, the ability of appearance descriptors to maintain identity continuity when primary biometric evidence becomes unreliable. The system evaluation is based on an analysis that integrates confusion matrices, global metrics, and performance curves, providing a comprehensive view of its behavior; the facial identification module stands out for having the highest performance level, while the re-identification modules face more complex conditions, yielding moderate yet consistent results. Furthermore, the joint analysis of tables, matrices, and graphical representations facilitates the

Table 4. Face recognition confusion matrix

Confusion Matrix (Number of images = 936)					
	Dalila	**Esteban**	**Gabriel**	**Jean**	**Luis**
Dalila	180	0	0	0	0
Esteban	0	249	0	1	0
Gabriel	0	0	145	0	0
Jean	0	0	0	146	0
Luis	0	0	0	0	215

Table 5. Facial Recognition Model Metrics

Metrics	**Value**
Accuracy	0.9989
Precision	0.9986
Recall	0.9992
F1 Score	0.9989
mAP	0.9978

identification of confusion patterns and an understanding of the contribution of the various visual representations.

3.2 Results of facial feature-based identification

The facial recognition model was evaluated using 936 images across five classes, captured at various locations within the camera network and, therefore, under significant variations in lighting conditions. The test set includes images with excessive brightness, heavy shadows, and intermediate lighting conditions, making the evaluation a direct test of the model's robustness against variations in capture conditions. The confusion matrix shown in Table 4 indicates near-ideal performance, with all classes aligned along the main diagonal and a single confusion between Esteban and Jean. This pattern suggests that the facial representation retains a high discriminative capacity, even in the presence of lighting variations, and maintains low error rates.

The overall metrics in Table 5 reinforce this interpretation. An accuracy of 0.9989, along with precision, recall, and F1-score values close to one, indicate that the facial module is the most stable component of the entire system. In particular, the recall of 0.9992 suggests that most true identities were correctly retrieved, while the mAP of 0.9978 confirms a very high overall discrimination capability.

3.3 Results of re-identification based on clothing color

The color-based re-identification model was evaluated using 3,012 images captured at various locations along the camera circuit. Unlike facial recognition, this approach relies on the consistency of clothing colors under varying lighting conditions and on the similarity between garments, which increases the complexity of the problem. As shown in Table 6, the matrix retains a diagonal trend, although with the emergence of cross-class confusion. The greatest discrepancies are observed in Esteban, with errors toward Gabriel and Luis, and in Jean, with significant confusion toward Luis. This pattern indicates that color information

Table 6. Color-Based Re-Identification Confusion Matrix

Confusion Matrix Number of images = 3012					
	Dalila	Esteban	Gabriel	Jean	Luis
Dalila	402	10	24	0	5
Esteban	50	459	162	2	102
Gabriel	69	26	568	7	36
Jean	9	12	4	417	362
Luis	2	66	94	62	362

Table 7. Overall color-based re-identification metrics

Metrics	**Value**
Accuracy	0.7331
Precision	0.7433
Recall	0.7506
F1 Score	0.7410
mAP	0.81

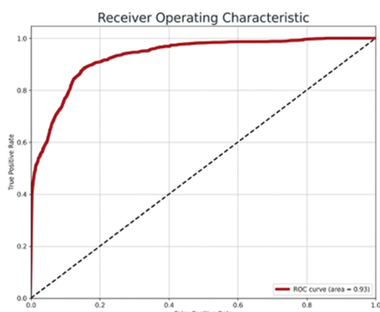

Fig. 6. ROC curve for the color-based re-identification model

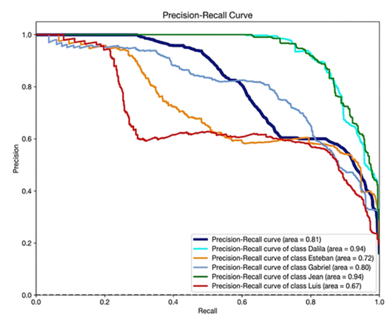

Fig. 7. Individual precision–recall curves for the color-based re-identification model

remains discriminative, but its performance is limited by changes in lighting and similarities in color distribution.

The overall metrics in Table 7 place this module in the middle range of performance, with an accuracy of 0.7331, a precision of 0.7433, a recall of 0.7506, and an F1-score of 0.7410. These values show that clothing color provides a useful basis for re-identification, although it does not achieve the stability of the facial module. The mAP of 0.81 also indicates a reasonable classification capability, consistent with a scenario in which chromatic appearance provides relevant information but is not entirely sufficient on its own.

3.4 Curves and visualizations of the color-based re-identification model

Performance curves complement the interpretation of the matrix and allow us to examine the model's discriminatory power from different perspectives. The ROC curve (Fig. 6) remains clearly above the random line, supporting the existence of class separation, even under a complex capture scenario.

Table 8. Confusion matrix for texture-based re-identification

Confusion Matrix Number of images = 3012					
	Dalila	Esteban	Gabriel	Jean	Luis
Dalila	497	15	138	19	12
Esteban	7	478	107	32	10
Gabriel	5	25	342	8	2
Jean	20	43	221	366	67
Luis	3	12	44	63	476

The overall Precision–Recall curve shown in Fig. 7, specifically the blue line, exhibits a downward trend as recall increases, indicating a rise in false positives as the decision threshold is improved. The individual curves reveal significant heterogeneity across classes, with some identities maintaining high performance while others show early declines, consistent with the observed confusion patterns. Color is a useful feature for re-identification, although it is sensitive to variations in lighting and similarities between garments.

3.5 Results of re-identification based on clothing texture

The texture-based re-identification model was evaluated on a dataset of 3,012 images. In this case, the decision is based on local patterns in the surface appearance of clothing; therefore, its performance depends directly on the effective resolution of the garments, the presence of repetitive details, and the stability of these patterns throughout the camera network. Table 8 presents the resulting confusion matrix, showing that, while texture retains its classification ability, it introduces more pronounced confusion in certain classes. In particular, the Dalila class shows significant errors toward Gabriel, while Jean concentrates confusion toward Gabriel and Luis. Likewise, the Gabriel class exhibits a less dominant diagonal compared to the color-based model. These patterns indicate that, although texture captures discriminative regularities, its performance is affected when garments exhibit structural similarities or when the quality of capture reduces the local contrast of details.

For its part, the overall metrics reported in Table 9 reflect this intermediate performance, with values of accuracy (0.7168), precision (0.7519), recall (0.7371), and F1-score (0.7181), along with an mAP of 0.81. Compared to the color-based model, the texture-based model exhibits slightly higher precision but lower accuracy, recall, and F1-score. This result suggests a more conservative behavior of the model, capable of maintaining correct decisions when identifying clear distinctive patterns, though less robust in fully recovering identities within the test set.

3.6 Curves and visualizations of the texture-based re-identification model

The graphical interpretation of the texture-based model confirms its intermediate performance within the system. Figure 8 shows that the ROC curve remains above the random line, indicating the presence of discriminative power, although with less separation between classes compared to the color-based model.

Table 9. Overall texture-based re-identification metrics

Metrics	Value
Accuracy	0.7168
Precision	0.7519
Recall	0.7371
F1 Score	0.7181
mAP	0.81

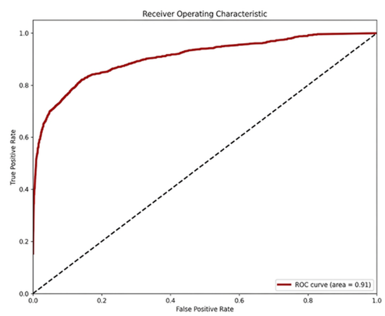

Fig. 8. ROC curve for the texture-based re-identification model

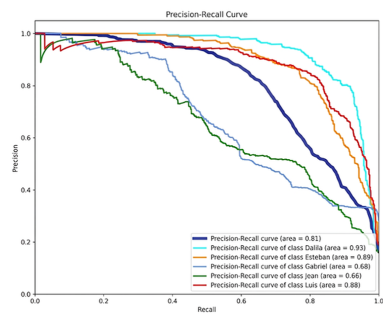

Fig. 9. Individual precision–recall curves for the texture-based re-identification model

For its part, the overall Precision–Recall curve (Fig. 9: blue line) shows a more pronounced decrease in precision as recall increases, reflecting the model's greater sensitivity to an increase in false positives when seeking greater completeness in identity retrieval. The following Precision–Recall curves show significant variability across classes, which is consistent with the patterns observed in the confusion matrix. This behavior indicates that the textural model does not completely lose its classification ability, but distributes its error unevenly across identities, depending on the distinctiveness of the visual patterns present in the clothing.

3.7 Results of re-identification based on color–texture matching

Evaluating the combined color–texture representation allows us to analyze the effect of integrating both sources of information on identity continuity, compared to the isolated use of each descriptor. The model was evaluated on the same set of 3,012 images, ensuring direct comparability among the three de-identification scenarios considered. Table 10 shows that the confusion matrix obtained maintains a predominant main diagonal, where significant confusion persists, particularly from Esteban to Gabriel and to Luis, and from Luis to Gabriel and to Jean. In the model based exclusively on texture, the error pattern is more balanced, indicating greater stability in classification. The pattern indicates that the combination of chromatic and textural information provides a more robust representation, although it does not completely eliminate the ambiguity between garments with similar characteristics.

Table 10. Re-identification confusion matrix based on color–texture combination

Confusion Matrix Number of images = 3012					
	Dalila	Esteban	Gabriel	Jean	Luis
Dalila	405	13	29	6	12
Esteban	48	460	161	2	103
Gabriel	68	26	549	6	31
Jean	9	12	4	410	52
Luis	2	62	109	64	369

Table 11. Overall re-identification metrics for color–texture matching

Metrics	**Value**
Accuracy	0.7281
Precision	0.7399
Recall	0.7447
F1 Score	0.7365
mAP	0.80

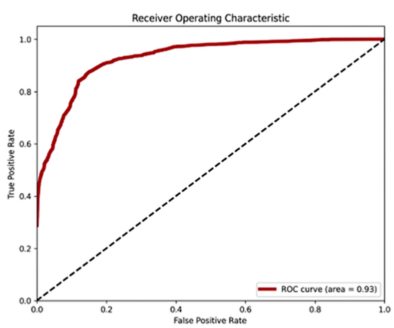

Fig. 10. ROC curve for the color–texture-based re-identification model

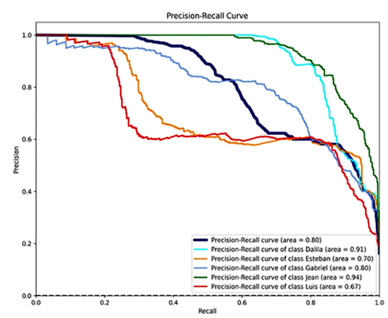

Fig. 11. Individual precision–recall curves for the color–texture-based re-identification model

For its part, the overall metrics reported in Table 11 reflect this intermediate performance, with values of accuracy (0.7281), precision (0.7399), recall (0.7447), and F1-score (0.7365), along with an mAP of 0.80. Compared to the individual models, the combination outperforms the texture-based model in terms of accuracy, recall, and F1-score, demonstrating the complementary contribution of both descriptors. However, performance remains slightly below that of the model based exclusively on color, indicating that, in this scenario, chromatic information remains the most decisive factor for distinguishing between identities.

3.8 Plots and visualizations of the combined model

The performance curves of the combined model support this interpretation. The ROC curve shown in Fig. 10 remains clearly separated from the random line, indicating that the model has adequate discriminatory power.

For his part, the global Precision–Recall curve (Fig. 11: blue line) shows a gradual decline in precision as recall increases, a behavior similar to that observed in the individual models and consistent with the increase in false positives that occurs when the decision threshold is relaxed. The other curves show heterogeneous trajectories across identities, indicating that feature fusion does not uniformly benefit all classes. This behavior suggests that the effectiveness of the combination depends on the degree of complementarity between the color information and the texture patterns available in each case.

3.9 Technical summary of the experimental findings

A comprehensive analysis of all the tests confirms a clear hierarchy of performance within the proposed system. Facial recognition is the strongest component, with a near-perfect confusion matrix and metrics close to ideal classification. Clothing-based re-identification faces a challenging environment, where color and texture provide descriptive information, but it is more vulnerable to lighting variations and similarities between garments. Among the approaches analyzed, the color descriptor performs best, followed by the color–texture combination, and finally, texture alone. The results demonstrate that the integration of clothing attributes is methodologically sound, making it more competitive, but its effectiveness depends on the interaction between descriptors in real-world scenarios.

4 Discussion

The results show that the proposed architecture effectively integrates initial facial identification with re-identification based on clothing attributes, rather than treating them as separate processes. This integration aligns with the system's methodological approach, in which identity continuity is maintained through a transition from primary biometric evidence to soft-biometric cues when facial information becomes unreliable. In general terms, the results obtained confirm the system's viability in non-overlapping multi-camera video surveillance scenarios, although significant performance differences are evident among its modules. From a methodological perspective, this behavior is directly related to the design decisions adopted. In particular, the incorporation of YOLOv8 for detection, along with the preprocessing stages and classification using SVM, allows for the structuring of a consistent pipeline. However, while facial identification operates under relatively controlled conditions, re-identification is influenced by factors inherent to the environment, such as lighting variations, changes in perspective, and similarities in clothing. Consequently, the observed differences should not be interpreted as limitations of the system, but rather as a manifestation of the inherent complexity of the problem addressed.

Facial recognition has established itself as the most robust component, achieving near-ideal performance levels even under varying lighting conditions, which reinforces its role as the primary means of identity verification. On the other hand, re-identification operates in a more challenging context, where, while it maintains discriminative power, it introduces higher levels of uncertainty by relying on the stability of clothing attributes. Within this framework, the comparative analysis reveals a clear hierarchy among the evaluated descriptors: color performs best, followed by the color–texture combination, and finally, texture alone. This result suggests that color constitutes the primary source of discrimination in the absence of the face, while texture acts as a complement. However, the integration of both attributes fails to surpass the performance of the individual color descriptor, indicating that complementarity does not always translate into an overall improvement.

The color-based model tends to produce errors between classes with similar color distributions, while the textural model shows greater dispersion in its errors. Although combining color and texture helps partially reduce these inconsistencies, some ambiguity between classes still persists, confirming the complex nature of the re-identification process in real-world settings. Regarding the study's contributions, the results validate the relevance of integrating multiple information sources into a single pipeline, but also clearly delineate the role of each: the face as the primary identification mechanism, color as the most effective support for re-identification, and texture as a complementary descriptor conditioned by capture conditions.

5 Conclusions

The proposed architecture demonstrates that it is possible to maintain identity continuity in a non-overlapping multi-camera video surveillance system by integrating facial identification and re-identification based on clothing attributes within a single pipeline. The results show that the proposed integration facilitates an effective transition from primary biometric information to soft-biometric attributes in the absence of a face. However, detailed analysis indicates that, while the combination of color and texture constitutes a methodologically valid alternative and improves performance compared to the exclusive use of texture, it fails to outperform the chromatic descriptor when considered individually. Consequently, the fusion of attributes does not result in significant increases in overall performance or a consistent reduction in cross-class confusion. Therefore, the integration of soft-biometric attributes does not in itself guarantee an improvement in re-identification. Under real-world capture conditions, clothing color remains the most stable and discriminative descriptor in the absence of the face.

Acknowledgment

The research project entitled "Development of video surveillance models and/or algorithms for Re-Identification of persons based on soft biometric features in a closed-circuit camera using computer vision and machine learning tech-niques" is the focus of this study. This project was supported by the Research Group on Intelligent Systems, Data Ana-lytics, Robotics, Cybersecurity and Software (SIARCyS) for their valuable assistance in the implementation of these models. In addition, we extend our thanks to the Universidad de las Fuerzas Armadas ESPE for their support during the realization of this research effort.

References

1. A. Nadeem et al.: Tracking Missing Person in Large Crowd Gathering Using Intelligent Video Surveillance. Sensors, vol. 22, no. 14, p. 5270 (2022). doi:

10.3390/s22145270
2. K. Rezaee, S. M. Rezakhani, M. R. Khosravi, M. K. Moghimi: A survey on deep learning-based real-time crowd anomaly detection for secure distributed video surveillance. Pers Ubiquit Comput, vol. 28, no. 1, pp. 135–151 (2024). doi: 10.1007/s00779-021-01586-5
3. S. Abdulmalek et al.: IoT-Based Healthcare-Monitoring System towards Improving Quality of Life: A Review. Healthcare, vol. 10, no. 10, p. 1993 (2022). doi: 10.3390/healthcare10101993
4. D. Nimma et al.: Object detection in real-time video surveillance using attention based transformer-YOLOv8 model. Alexandria Engineering Journal, vol. 118, pp. 482–495 (2025). doi: 10.1016/j.aej.2025.01.032
5. T. F. N. Bukht et al.: A review of video-based human activity recognition: theory, methods and applications. Multimed Tools Appl, vol. 84, no. 17, pp. 18499–18545 (2025). doi: 10.1007/s11042-024-19711-w
6. S. Wang et al.: Safety Helmet Monitoring on Construction Sites Using YOLOv10 and Advanced Transformer Architectures with Surveillance and Body-Worn Cameras. Journal of Construction Engineering and Management, vol. 151, no. 11, p. 04025186 (2025). doi: 10.1061/JCEMD4.COENG-16760
7. L. Laishram et al.: Toward a Privacy-Preserving Face Recognition System: A Survey of Leakages and Solutions. ACM Comput. Surv., vol. 57, no. 6, p. 147:1–147:38 (2025). doi: 10.1145/3673224
8. P. Gavilanes, J. Carrillo, E. Galarza: Person Re-identification System in a Controlled Environment Based on Soft Biometric Features. In: Proc. ICI2ST, pp. 14–20 (2022). doi: 10.1109/ICI2ST57350.2022.00010
9. J. Carrillo-Medina et al.: Person Re-identification using soft-biometric features: body silhouette and clothing texture in a multi-camera video surveillance environment. In: Proc. ICI2ST, pp. 1–8 (2022). doi: 10.1109/ICI2ST57350.2022.00013
10. Y. Huang et al.: CSFRNet: Integrating Clothing Status Awareness for Long-Term Person Re-identification. Int J Comput Vis, vol. 133, no. 6, pp. 3180–3202 (2025). doi: 10.1007/s11263-024-02315-0
11. M. A. Fouad et al.: A comprehensive survey on person Re-identification: methods, challenges, and future research directions. Multimed Tools Appl, vol. 85, no. 3, p. 235 (2026). doi: 10.1007/s11042-026-21333-3
12. G. Pandiya Rajan, D. Manimegalai: Person Re-Identification Using Deep Learning Based YOLO Network with Partially Occluded Faces. J. Electr. Eng. Technol., vol. 20, no. 4, pp. 2733–2745 (2025). doi: 10.1007/s42835-025-02159-9
13. K. Pal et al.: Machine Learning-Based Image Pattern Recognition Using Histogram of Oriented Gradient. IEEE Access, vol. 13, pp. 74396–74416 (2025). doi: 10.1109/ACCESS.2025.3564145
14. F. Ahmed: HOG-CNN: Integrating Histogram of Oriented Gradients with CNN for Retinal Image Classification. J Digit Imaging (2026). doi: 10.1007/s10278-026-01893-4
15. O. V. Brednev, A. V. Balobanov: Video Stabilization Quality Assessment Method Based on K-Means in HSV Color Space. In: Proc. IEEE Conf., pp. 1–7 (2025). doi: 10.1109/IEEECONF64229.2025.10948072
16. I. Pavăloi et al.: Experiments on Occluded Palmprint Images Recognition and Retrieval Using Texture and Shape Features. In: Proc. ISSCS, pp. 1–4 (2025). doi: 10.1109/ISSCS66034.2025.11105661
17. Y. V. Via et al.: Performance Analysis of Combined Feature Extraction Method on Chili Quality Classification using XGBoost. In: Proc. ITIS, pp. 517–522 (2025). doi: 10.1109/ITIS67966.2025.11308945

18. R. Dowerah, S. Patel: Comparative analysis of color histogram and LBP in CBIR systems. Multimed Tools Appl, vol. 83, no. 5, pp. 12467–12486 (2024). doi: 10.1007/s11042-023-15955-0