Image and Vision Computing 143 (2024) 104917

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

Image-based human re-identification: Which covariates are actually (the most) important?☆

Kailash Hambarde *, Hugo Proença

IT - Instituto de Telecomunicações, University of Beira Interior, 6201-001 Covilhã, Portugal

ARTICLE INFO

Keywords:
Human re-identification
Performance covariates
Biometric menagerie

ABSTRACT

Human re-identification (re-ID) is nowadays among the most popular topics in computer vision, due to the increasing importance given to safety/security in modern societies. Being expected to sun in totally uncontrolled data acquisition settings (e.g., visual surveillance) automated re-ID not only depends on various factors that may occur in non-controlled data acquisition settings, but - most importantly - performance varies with respect to different *subject features* (e.g., gender, height, ethnicity, clothing, and action being performed), which may result in highly biased and undesirable automata. While many efforts have been putted in increase the robustness of identification to uncontrolled settings, a systematic assessment of the actual variations in performance with respect to each *subject feature* remains to be done. Accordingly, the contributions of this paper are threefold: 1) we report the correlation between the performance of three state-of-the-art re-ID models and different subject features; 2) we discuss the most concerning features and report valuable insights about the roles of the various features in re-ID performance, which can be used to develop more effective and unbiased re-ID systems; and 3) we leverage the concept of *biometric menagerie*, in order to identify the groups of individuals that typically fall into the most common menagerie families (e.g., goats, lambs, and wolves). Our findings not only contribute to a better understanding of the factors affecting re-ID performance, but also may offer practical guidance for researchers and practitioners concerned on human re-identification development.

1. Introduction

Image-based human re-identification (re-ID) is a critical and rapidly evolving area of research, with significant applications in surveillance, criminal investigations, and public safety scenarios [1,2]. The goal of re-ID is to match a probe image of an individual with images of the same identity from a gallery collected by multiple, non-overlapping cameras [3,4]. The complexity of real-world environments poses significant challenges for re-ID, including dynamic lighting, shadows, occlusions, atmospheric turbulence, and cluttered backgrounds. Crucially, re-ID performance is also impacted by subject features such as age, gender, and ethnicity, leading to potential biases in the system [5,6].

Soft biometric features like age, gender, clothing style, and body shape, when used as auxiliary information, can enhance re-ID performance but also increase the risk of biased recognition systems [6]. Fig. 1 illustrates how these attributes can be used in re-ID systems. Recognizing the crucial role of soft biometrics, our study explores the impact of 14 such attributes (including gender, height, ethnicity, and clothing) on three well-known state-of-the-art methods of re-ID performance. Having evaluated different possibilities, we decided to conduct the experiments on the P-DESTRE dataset [4], in which data were acquired in uncontrolled conditions and are annotated exactly for all our features of interest.

Our study aims to fill a major gap in current research by systematically examining how the different subject features may be correlated to the accuracy of well knwon re-ID methods. Our main objectives are:

- Investigate how different subject features, such as age, gender, ethnicity or clothing, typically affect the re-ID performance.
- Analyze the scores provided by three leading re-ID methods, with respect to the features considered and evaluate their correlation.
- Provide robust and relevant findings, by conducting experiments in the P-DESTRE dataset, which was chosen for its uncontrolled conditions and comprehensive annotation.
- The contributions of this research are threefold:

☆ This work is funded by FCT/MCTES through national funds and cofounded by EU funds under the project UIDB/50008/2020.
* Corresponding author.
E-mail address: kailas.srt@gmail.com (K. Hambarde).

https://doi.org/10.1016/j.imavis.2024.104917
Received 30 October 2023; Received in revised form 17 December 2023; Accepted 17 January 2024
Available online 20 January 2024

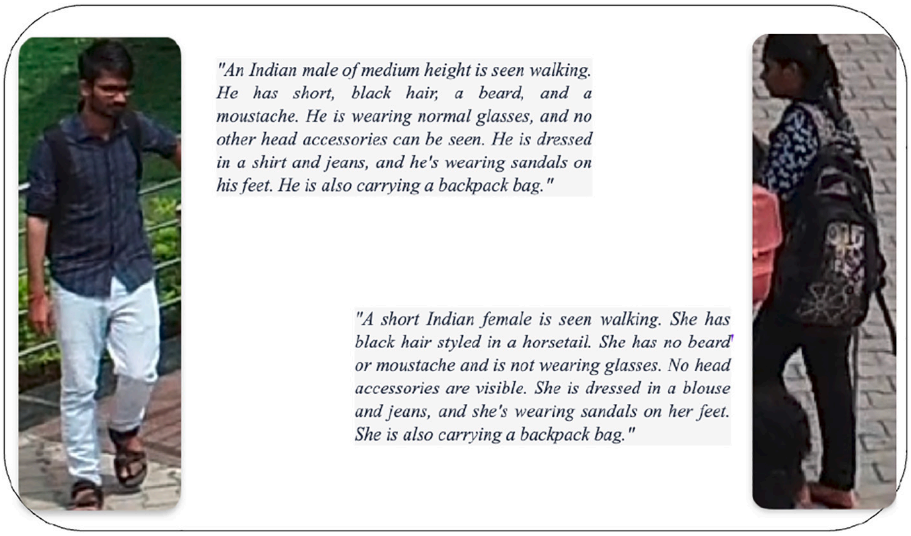


Fig. 1. Examples of feature values (attributes) that are typically used to describe pedestrians and can be used as auxiliary information in human re-identification (re-ID). This paper provides a systematic analysis about the typical effect of such feature values in re-ID performance.

- A systematic analysis of the impact of various subject features in human re-ID, identifying those that cause the largest performance variations.
- A comparison of the performance variations of three well-known re-ID methods concerning each of these features (covariates).
- An analysis of the relationship between each feature and the most important biometric families, regarding the well-known *menagerie* taxonomy.

This paper provides a comprehensive analysis of subject feature impacts on re-ID performance, offering insights for developing more effective and unbiased systems. Hence, this work fills a critical gap in understanding the role of the subject features in re-ID performance.

The remaining of this paper is organized as follows: Section 2 briefly summarizes the State-of-the-Art in re-ID. Section 3 provides a detailed description of the methodology steps that were followed. Section 4 provides all the details about our empirical framework, including details on the dataset and of the tested methods. Section 5 provides the results obtained and discusses their implications. Finally, Section 6 concludes the paper.

2. Re-ID state-of-the-art

Upon the proposal in [7], the existing re-ID methods were categorized into deep metric, local feature, generative adversarial, and video-based feature learning. Additionally, considering the most recent advancements, we considered graph-based, attention-based and multi-modal approaches.

2.1. Deep metric learning

It refers to the widely known Metric Learning topic that aims to learn a similarity function between two pedestrian images. The objective of Deep ML is to obtain a mapping from the original image to the feature embedding (FE) space, such that two samples of the same pedestrian have small distances on the FE space, while samples from different IDs should be far apart each other [8]. Learning is based on a loss function [9] that can have different forms (e.g., pairs, triplets, quadruplets). Anyway, such functions obtain features that are maximally invariant to different factors, such as pose, illumination, and occlusion, and can effectively distinguish between different IDs. verification loss [8], contrastive loss [10], triplet loss [11] quadruplet loss [12]. Overall, this family of methods enable models to learn discriminative features in an automatic way, which solves the problem of manually designing features, known to be particularly sensitive.

2.2. Local feature learning

Based on the spatial support of the features extracted, re-ID methods can be also be divided into global and local. The former family of methods extracts features that regard the whole pedestrian image [13], being particularly difficult to obtain minutia information about the pedestrian. Oppositely, local feature learning-based methods aim at learning pedestrian discriminative features and ensuring proper alignment of each local feature. Attention modules can be used to automatically focus on particularly important local regions. Commonly used local feature learning methods are stripe segmentation [14], multi-scale fusion [15], soft attention [16] pedestrian semantic extraction and global–local feature learning [17]. These methods also alleviate the problems of occlusion, boundary detection errors, and view and pose variations.

2.3. Generative adversarial learning

In 2014, Goodfellow et al. [18] first proposed the concept of generative adversarial networks (GAN), which has rapidly developed in recent years. Many variants and applications of GANs emerged [19], being used to synthesize pedestrian images with different poses, appearance, lighting, and resolution in order to expand the dataset and improve the generalization ability of the model [20] GANs have also been used to learn identity-related features that can improve the accuracy of feature-matching [19,21–23]. These methods can alleviate the small number of training samples, resolution, illumination, view, and pose variation.

2.4. Video-based feature learning

Various works have proposed also to account on the time-based information contained in video sequences for human re-ID purposes. These feature learning-based methods take short videos as input and use both spatial and temporal complementary cues, as an attempt to alleviate the limitations of appearance-based features. Most of these methods use optical flow information [24] 3-dimensional convolutional neural networks (3DCNNs) [25,26], recurrent neural networks(RNN) or long short term memory(LSTM) [24], spatial–temporal attention or graph convolutional networks (GCN) [27–30] to model the

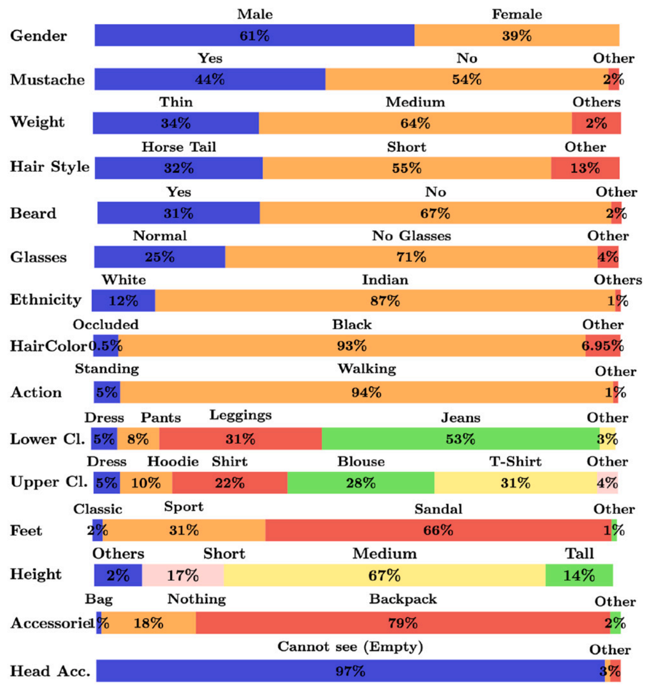


Fig. 2. Prior probabilities for each value of the 15 *attributes* manually annotated in the P-DESTRE dataset.

spatial–temporal information of video sequences. These methods can mitigate occlusions, resolution changes, illumination changes, and view and pose variations.

2.5. Graph-based re-ID

This family of re-ID methods relies in graph structures to analyze complex relationships between features, enhancing accuracy in diverse scenarios, including occlusions. Techniques like graph attention networks [31] emphasize the importance of contextual relationships. Also, graph convolution methods for re-ranking [32] are used to improve feature representations. Adaptable graph-based frameworks, such as [33,34], show promising results in handling complex data, including noisy labels [35]. Contributions like part-guided graph convolutions and probability predictions in networks [36,37] focus on enhancing identification accuracy while integrating deep learning with graph-based methods offers advanced solutions, as seen in [38].

2.6. Attention mechanisms for re-ID

As in many other computer vision fields, attention mechanisms are also critical for enhancing focus on relevant features in re-ID approaches. Recent advancements [39] conclude that attention mechanisms, when combined with global pooling methods, can significantly boost re-ID efficiency. Innovations addressing occlusions, such as semi-attention partition techniques [40], offer robust solutions for partially visible subjects. Studies exploring attribute-guided attention [41,42] highlight the importance of attributes in guiding attention mechanisms. Moreover, semantic-driven attention networks integrated with attribute learning [43] and fine-grained attribute-aware analysis [44] mark significant shifts towards more fine and context-aware approaches in re-ID technologies.

2.7. Multi-modal person re-ID

Recent research in multi-modal re-ID has led to the development of innovative models that significantly improve identification across various modalities. An approach that integrates soft bio-metrics with body figures effectively addresses long-term re-ID challenges, especially those related to clothing variations [45]. Furthermore, developing a modality-agnostic re-ID architecture marks a substantial advancement, enhancing retrieval accuracy and adaptability across different modal scenarios [46]. Additionally, strides in cross-modality re-ID, particularly in Visible-Infrared person re-ID through multi-task learning, have resulted in notable performance improvements on benchmark datasets [47]. These collective advancements represent significant progress in multi-modal re-ID, offering innovative solutions to the challenges posed by modality variations.

3. Methodology

A visual overview of the methodology followed in this work is provided in Fig. 3. We start by formalizing the re-ID problem, followed by the selection and exploration of the P-DESTRE dataset. A key aspect of our study involves examining the attribute covariates and their impact on the effectiveness of re-ID, highlighting their importance within the re-ID system. Next, we discuss the selected methods and the rationale behind their selection. Finally, we describe the evaluation metrics used in our study, ensuring a comprehensive assessment of the re-ID performance.

3.1. Problem definition

Formally, human re-ID aims at identifying a subject across different camera views, typically in a video surveillance setting, i.e., using images that are degraded under various factors. It is defined as: given a probe image, represented by a feature vector $\mathbf{x}$, and a gallery set, with elements represented by their feature vectors $\mathbf{Y} = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n\}$, the goal is to find the best matching feature vector $\mathbf{y} \in \mathbf{Y}$. This can be formalized as finding the gallery element $\hat{\mathbf{y}}$ that maximizes the similarity with respect to the query:

$$\hat{\mathbf{y}} = \underset{\mathbf{y}}{argmax} f(\mathbf{x}, \mathbf{y}), \tag{1}$$

with $f(.,.)$ being a similarity metric between any pair of elements, typically based in the Euclidean distance or cosine similarity, or any score yielding from a machine learning model. The re-ID problem can be further generalized to the case where multiple probe images and gallery sets correspond to different camera views. In this case, the goal is to find the best matching feature vectors for each probe image among all gallery sets.

3.2. Datasets selection

As above stated, in our experiments we considered the P-DESTRE[1] dataset [4]. It is a UAV-based data set, composed of video sequences acquired from DJI Phantom 4 drones flown by human operators over various outdoor urban environments at two university campuses. Data were recorded at 30fps, with 4 K spatial resolution and stored in “mp4” format with H.264 compression. Annotation is provided at the frame level, validated by human experts for different tasks, such as pedestrian detection, tracking and re-ID. With respect to the latter task, it should be noted that this set enables both short-term and long-term settings, depending on whether two samples of the same ID were acquired in the same day (i.e., with subjects wearing the same clothes - short-term, or not - long-term re-ID).

3.3. Attribute covariates in re-ID

When operating in large-scale scenarios, it is expected that subjects

[1] http://p-destre.di.ubi.pt/

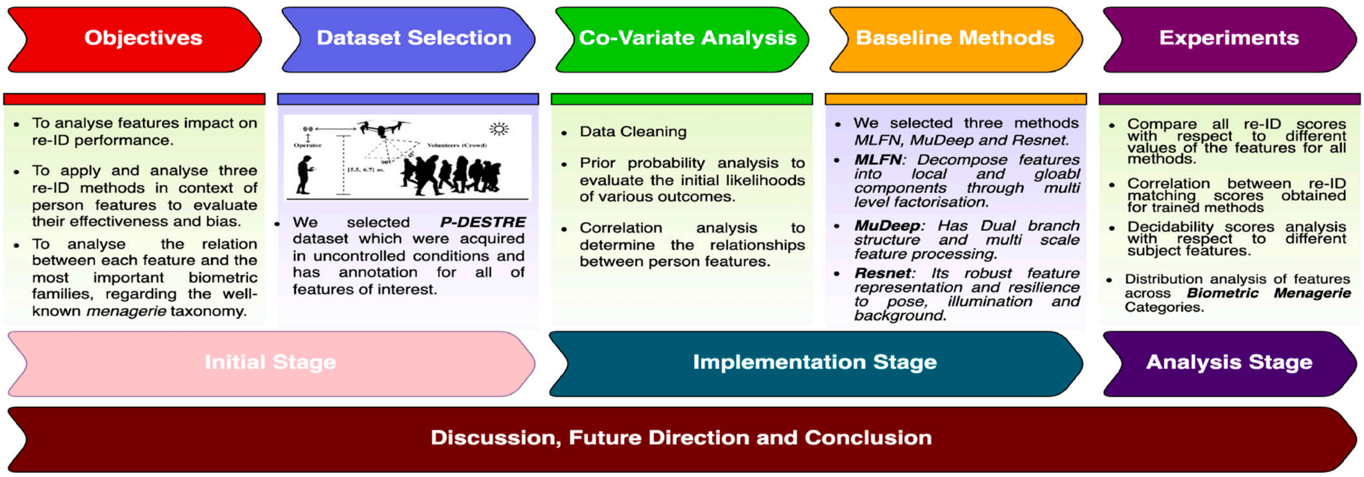
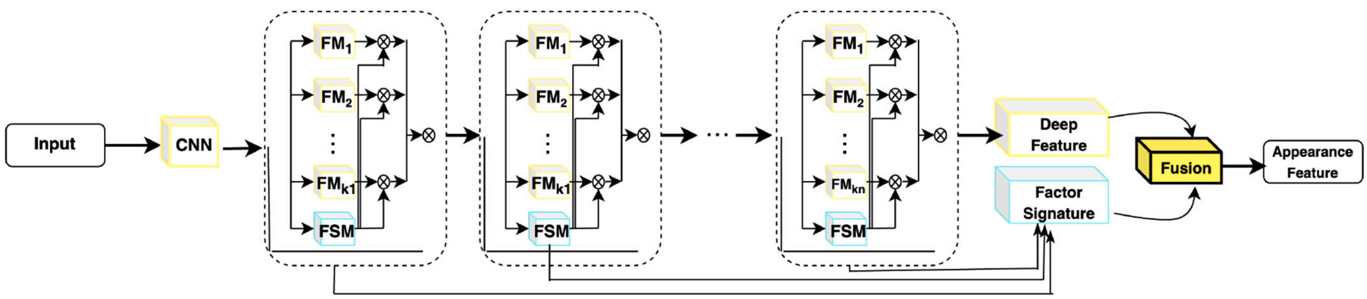
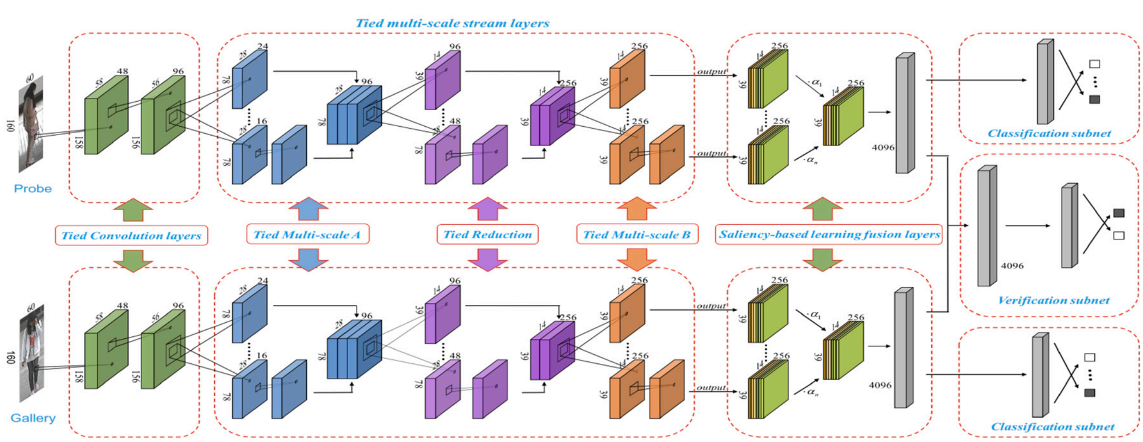


Fig. 3. Schematic overview of the research methodology used in this work.



(a) MLFN [49].



(b) MuDeep [13].

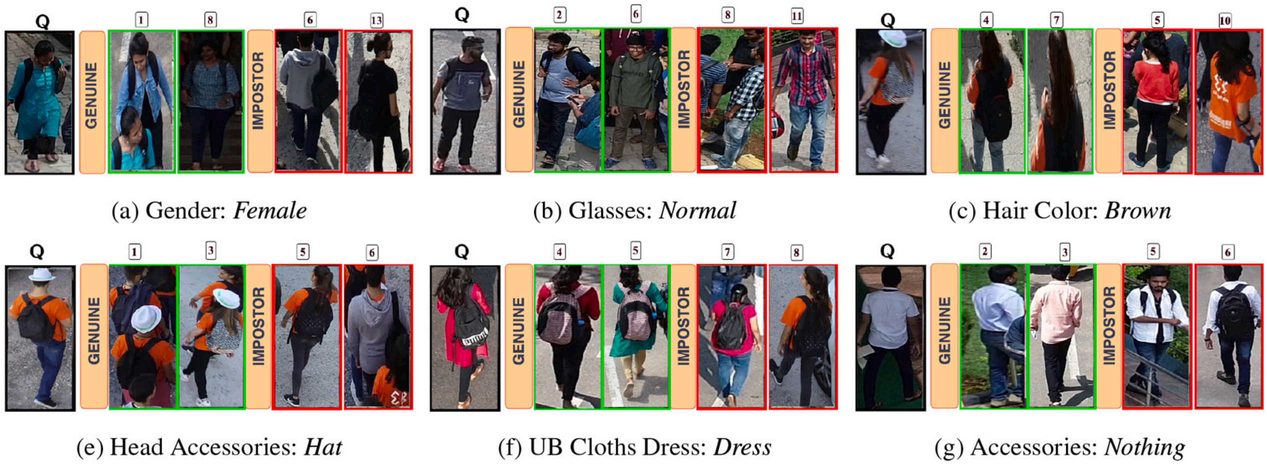
Fig. 4. Architecture of the deep learning-models (MLFN and MuDeep) considered in this work.

display many different factors of variability, such as *gender*, *age*, *height* or *clothing style*. Such characteristics - typically designated as *subject features* are known to determine the effectiveness of the re-ID task, and are designated as covariates.[2] Also, they can even constitute a significant bias for the effectiveness of the identification system, in particular with respect to some of the sub-sets of the population, which is considered socially and ethically unacceptable.

At the same time, using such attributes (as *soft information*) in the identification process has been reported to enhance the performance of re-ID by supplying supplementary information about an individual appearance, with numerous works focused in different sets of attributes. In these works, various attribute-annotated datasets, such as PA-100 K, PETA, Market-1501 Attribute, and DukeMTMC Attribute [48], have been used.

In this study, we consider the P-DESTRE dataset [4] due to a set of characteristics that we consider particularly important: 1) it is UAV-based; 2) it is fully annotated; 3) has multi-purposes (detection,

[2] https://www.statisticshowto.com/covariate/



centring

(a) Gender: *Female* (b) Glasses: *Normal* (c) Hair Color: *Brown*

(e) Head Accessories: *Hat* (f) UB Cloths Dress: *Dress* (g) Accessories: *Nothing*

Fig. 5. Examples of notable genuine/impostor results, using the MLFN model. For each query **Q**, we provide the corresponding ranks for samples of the same identity and illustrate some impostors that got particularly low ranks, due to gallery attributes that are common to the query.

(a) Sheep. (b) Goats. (c) Lambs/Wolves.

Fig. 6. Illustration of identities that are typically associated with the three biometric menagerie families considered in this work: *"Sheep"* (regular users of a biometric system), *"Goats"* (elements that are particularly hard to match, 6a, either for *genuine* and *impostor* comparisons) and *"Lambs/Wolves"*, that are particularly easy to impersonate 6c).

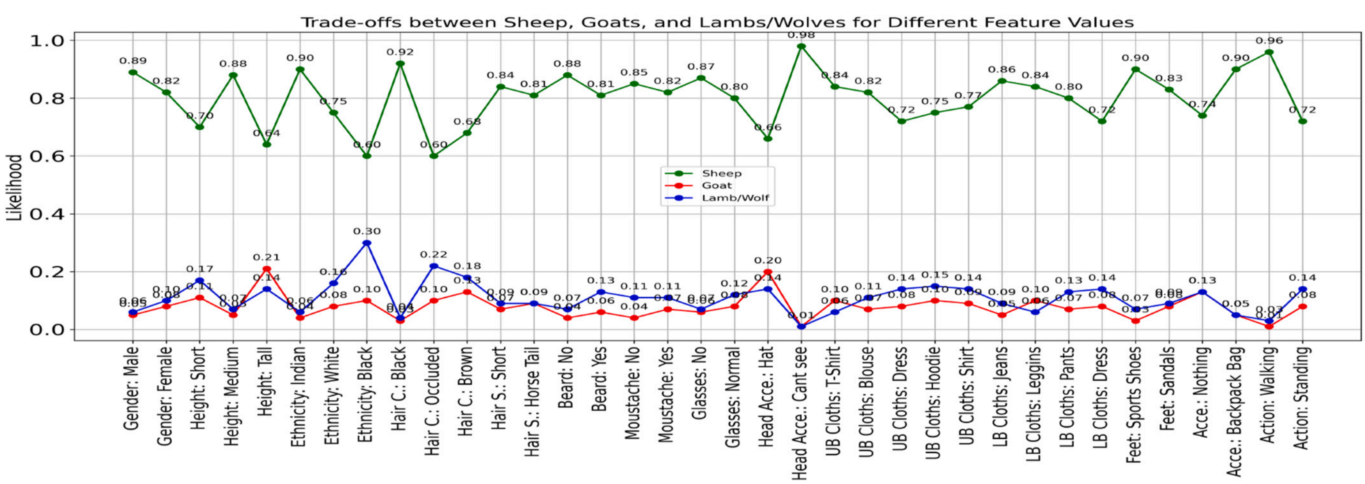


Fig. 7. Comparative view of the distribution and inherent trade-offs of classification likelihoods for different features across the *Biometric Menagerie* families *Sheep, Goats, Lambs/Wolves*.

tracking, and short/long-term re-ID); and 4) is freely available for research purposes. Fig. 2 provides the prior probabilities for each value of the different attributes (covariates). For instance, we can see that about 61% of the subjects in this dataset are males and 39% are females. Similarly, the prior probability of a subject having *medium* height is 67%, while the probability of being overweight is only 2%. In any case, it should be stressed that we attempted to keep the number of instances per feature/value in the learning and gallery sets as balanced as possible, having reflected such unbalanced priors mostly in the test (probe) sets.

3.4. Baseline methods

Having analyzed the state-of-the-art in human re-ID, we selected three methods based on different architectures that we considered to faithfully represen the terogeneity of the existing methods: the Multi-Scale Local Feature Extractor [49] (MLFN), MuDeep [13], and ResNet [50] methods as baselines.

MLFN [49] method comprises a backbone network and a multi-level factorization module, as shown in Fig. 4a. The base network extracts

Table 1
Biometric menagerie families considered in our study. While"sheep" make up the majority of a biometric system (high genuine matching scores and low impostor scores),"goats" are particularly difficult to match and lambs/wolves represent the opposite case, they are exceptionally good at impersonate.

Menagerie family	Genuine scores	Impostor scores
Sheep	High	Low
Goats	Low	Low
Lambs/Wolves	High	High

Table 2
Correlation matrix between the re-ID matching scores obtained for the three baseline methods considered: Resnet, MuDeep, and MLFN.

	ResNet	MuDeep	MLFN
ResNet	1.00	0.78	0.38
MuDeep	0.78	1.00	0.12
MLFN	0.38	0.12	1.00

image features, while the factorization module decomposes them into local and global components. The model stacks blocks with Factor Modules (FMs) and a Factor Selection Module (FSM), modeling multiple semantic levels. The FSM generates a selection vector for a subset of latent factors. The image representation at the nth level is expressed as a tuple M_n, S_n, where M_n is a tensor of feature maps, and S_n is the selection vector. A Factor Signature (FS) is obtained to enhance the final deep representation, combining output vectors from all levels. The deep features and FS are fused, and the final output representation R is obtained by averaging the two projected features.

MuDeep [13] has two branches to process image pairs and consists of five components: tied convolutional layers, multi-scale stream layers, a saliency-based learning fusion layer, a verification subnet, and a classification subnet as shown in Fig. 4b. The tied convolutional layers share weights between branches, and Inception architectures inspire the multi-scale stream layers. The fusion layer uses a saliency-based learning strategy to emphasize discriminative patterns. The verification subnet calculates the distance between images and predicts the likelihood of depicting the same person.

Finally, the well-known *ResNet* architecture has also been successfully applied to re-ID tasks, where they learn a feature representation robust to pose, illumination, and background variations. Typically, a ResNet architecture with multiple branches is used, each one responsible for learning different aspects of the input image, and fusing the features using a fusion layer, that obtains the difference between the output features of the two branches using element-wise subtraction and multiplication.

In order to perceive the effect of various subject features on re-ID accuracy, the selection of MLFN, MuDeep, and ResNet was strategic and purposeful:

- **MLFN** [49] is particularly relevant due to its capacity to decompose features into local and global components through its multi-level factorization module. This capability aligns with our goal to analyze how different subject features, such as age, gender, and clothing, impact re-ID performance.

Table 3
Human re-identification Average Precision (AP) with respect to different values of the features considered. As baseline, we provide the *overall* performance, for three different models: ResNet, MuDeep, and MLFN, along with the relative variability observed for each subset (Δ denotes performance *better* than the baseline, while ∇ denotes the opposite case. ΔΔ and ∇∇ denote variations with magnitude higher than 10%).

Feature	Value	ResNet	MuDeep	MLFN
Overall	–	0.80	0.83	0.86
Gender	Male	0.85 (+06.25%) Δ	0.87 (+04.82%)Δ	0.89 (+03.49%)Δ
	Female	0.75 (−06.25%) ∇	0.79 (−04.84%) ∇	0.82 (−03.65%) ∇
Height	Short	0.73 (−08.75%)∇	0.74 (−10.84%)∇∇	0.78 (−09.30%)∇
	Medium	0.85 (+06.25%)Δ	0.88 (+06.02%)Δ	0.89 (+03.49%)Δ
	Tall	0.70 (−12.50%)∇∇	0.69 (−16.87%)∇∇	0.74 (−13.95%)∇∇
Ethnicity	Indian	0.86 (+07.15%)Δ	0.88 (+06.02%)Δ	0.90 (+04.65%)Δ
	White	0.70 (−12.15%)∇∇	0.73 (−12.02%)∇∇	0.75 (−12.79%)∇∇
	Black	0.75 (−06.25%)∇	0.77 (−07.23%)∇	0.80 (−06.98%)∇
Hair Color	Black	0.90 (+12.50%)ΔΔ	0.92 (+10.84%)ΔΔ	0.93 (+08.84%)Δ
	Occluded	0.72 (−10.00%)∇∇	0.75 (−09.64%)∇	0.79 (−08.14%)∇
	Brown	0.63 (−21.25%)∇∇	0.70 (−15.66%)∇∇	0.68 (−20.93%)∇∇
Hair Style	Short	0.81 (+01.25%)Δ	0.83 (+00.00%) –	0.87 (+01.16%)Δ
	Horse Tail	0.74 (−07.50%)∇	0.77 (−07.23%)∇	0.82 (−04.65%)∇
Beard	No	0.82 (+02.50%)Δ	0.87 (+04.82%)Δ	0.89 (+03.49%)Δ
	Yes	0.76 (−05.00%)∇	0.78 (−06.02%)∇	0.82 (−04.65%)∇
Mustache	No	0.82 (+02.50%)Δ	0.83 (+00.00%) –	0.86 (+00.00%) –
	Yes	0.77 (−03.75%)∇	0.82 (−01.20%)∇	0.85 (−01.16%)∇
Glasses	No	0.80 (+00.00%) –	0.86 (+03.61%)Δ	0.88 (+02.33%)Δ
	Normal	0.73 (−08.75%)∇	0.77 (−07.23%)∇	0.79 (−08.14%) ∇
Head Accessories	Hat	0.66 (−16.50%)∇∇	0.69 (−16.87%)∇∇	0.71 (−17.44%)∇∇
	Cannot see	0.93 (+17.00%)ΔΔ	0.95 (+16.87%)ΔΔ	0.97 (+13.95%)ΔΔ
Upper Body Cloths	T-Shirt	0.81 (+01.23%)Δ	0.85 (+02.41%)Δ	0.87 (+01.16%)Δ
	Blouse	0.78 (−02.50%)∇	0.81 (−02.41%)∇	0.83 (−03.49%)∇
	Dress	0.63 (−21.25%)∇∇	0.69 (−16.87%)∇∇	0.73 (−15.12%)∇∇
	Hoodie	0.70 (−12.70%)∇∇	0.73 (−12.05%)∇∇	0.76 (−11.63%)∇∇
	Shirt	0.71 (−11.25%)∇∇	0.75 (−09.64%)∇	0.79 (−08.14%)Δ
Lower Body Cloths	Jeans	0.82 (+02.50%)Δ	0.84 (+01.20%)Δ	0.86 (+00.00%) –
	Leggins	0.77 (−03.75%)∇	0.81 (−02.41%)∇	0.84 (−02.33%)∇
	Pants	0.74 (−07.50%)∇	0.78 (−06.02%)∇	0.80 (−06.98%)∇
	Dress	0.67 (−16.25%)∇∇	0.70 (−15.66%)∇∇	0.73 (−15.21%)∇∇
Feet	Sports Shoes	0.76 (−05.00%)∇	0.80 (−03.61%)∇	0.84 (−02.33%)∇
	Sandals	0.81 (+01.25%)Δ	0.87 (+04.82%)Δ	0.89 (+03.49%)Δ
Accessories	Nothing	0.66 (−17.50%)∇∇	0.69 (−16.87%)∇∇	0.74 (−13.95%)∇∇
	Backpack Bag	0.82 (+02.50%)Δ	0.86 (+03.61%)Δ	0.90 (+04.65%)Δ
Action	Walking	0.90 (+12.50%)ΔΔ	0.93 (+12.05%)ΔΔ	0.95 (+10.47%)ΔΔ
	Standing	0.63 (−21.25%)∇∇	0.69 (−16.87%)∇∇	0.73 (−15.12%)∇∇

Table 4
Decidability scores with respect to different subject features. We provide the mean and standard deviation of the genuine and impostor scores, along with the decidability (d') for each value.

Feature	Values	Genuine	Impostor	d'
Gender	Male	0.88 ± 1.22	0.62 ± 2.06	14.35
	Female	0.79 ± 1.90	0.57 ± 2.37	10.64
Height	Short	0.76 ± 2.00	0.62 ± 2.90	06.32
	Medium	0.87 ± 1.89	0.58 ± 2.50	13.84
	Tall	0.72 ± 1.60	0.61 ± 2.89	05.03
Ethnicity	Indian	0.90 ± 1.22	0.56 ± 3.10	16.35
	White	0.77 ± 2.38	0.62 ± 2.90	06.09
	Black	0.72 ± 2.47	0.55 ± 3.96	06.52
Hair Color	Black	0.91 ± 1.80	0.52 ± 2.21	19.47
	Occluded	0.70 ± 1.20	0.57 ± 2.50	06.75
	Brown	0.70 ± 2.76	0.62 ± 3.92	03.09
Hair Style	Short	0.82 ± 2.15	0.67 ± 2.03	07.33
	Horse Tail	0.74 ± 2.90	0.57 ± 2.37	07.39
Beard	No	0.88 ± 2.00	0.59 ± 3.17	13.75
	Yes	0.81 ± 2.65	0.61 ± 4.17	07.65
Mustache	No	0.82 ± 2.36	0.60 ± 3.06	09.44
	Yes	0.84 ± 3.63	0.62 ± 4.13	07.89
Glasses	No	0.87 ± 2.25	0.54 ± 2.59	14.99
	Normal	0.80 ± 2.89	0.63 ± 3.49	06.73
Head Accessories	Hat	0.75 ± 2.17	0.63 ± 3.74	04.93
	Cannot See	0.93 ± 5.10	0.55 ± 3.84	12.70
Upper Body Cloths	T-Shirt	0.85 ± 2.78	0.67 ± 3.93	06.69
	Blouse	0.80 ± 2.48	0.62 ± 3.13	07.64
	Dress	0.74 ± 2.41	0.59 ± 2.88	06.52
	Hoodie	0.77 ± 2.90	0.60 ± 3.14	06.91
	Shirt	0.80 ± 2.79	0.61 ± 3.39	07.64
Lower Body Cloths	Jeans	0.85 ± 2.42	0.62 ± 3.04	09.84
	Leggings	0.82 ± 2.59	0.63 ± 3.78	07.52
	Pants	0.77 ± 3.93	0.61 ± 4.20	05.61
	Dress	0.74 ± 2.41	0.59 ± 2.88	06.52
Feet	Sports Shoes	0.86 ± 2.88	0.66 ± 3.93	07.66
	Sandals	0.71 ± 2.63	0.60 ± 3.10	04.59
Accessories	Nothing	0.71 ± 2.63	0.64 ± 4.10	02.69
	Backpack Bag	0.88 ± 2.31	0.53 ± 2.83	15.43
Action	Walking	0.93 ± 1.30	0.57 ± 2.55	18.34
	Standing	0.74 ± 2.41	0.65 ± 2.88	03.91

- **MuDeep** [13] brings a unique value with its dual-branch structure and multi-scale feature processing, which is crucial for our analysis of discriminative pattern recognition in the context of various subject features: the modelâ€™s architecture, especially its saliency-based learning fusion layer.
- **ResNet** [50] is included for its robust feature representation and resilience to pose, illumination, and background variations.

Finally, another important point is the open-source availability of official implementations of each method, which is important to assure the reproducibility of results.

3.5. Evaluation metrics

In assessing the re-ID performance, we considered evaluation metrics that are well knwon in the field: Average Precision (AP) and Mean Average Precision (mAP).

3.5.1. Average precision (AP)

It evaluates the average precision for a query. Precision, in this context, refers to the proportion of relevant instances among the retrieved instances. Essentially, AP reflects how many of the ID's retrieved are actually relevant to the query and how well the system ranks those relevant items.

3.5.2. Mean average precision (mAP)

While AP is calculated for a single query, mAP is the mean of the AP scores for a set of queries. It provides an overall effectiveness measure for the retrieval system across multiple queries. In our context, it measures how well our system retrieves relevant IDs across all queries in our test set.

Additionally, we have considered the main families described in the *biometric menagerie taxonomy*, as detailed in Section 3.6 and illustrated in Table 1. This taxonomy provides a framework for understanding different user types in biometric systems, which we consider important to contextualize our evaluation metrics within the broader scope of biometric re-identification.

3.6. Biometric menagerie taxonomy

The concept of *biometric menagerie* [51] is based in the assumption that the different subjects of a biometric system exhibit varying degrees of accuracy, with some individuals experiencing particular difficulties in genuine/impostor matching. According to the different performance levels per subject, various families are defined, including"Goats","Wolves","Lambs","Worms","Doves","Chameleons", and"Phantoms". Considering that some previous works argued about the actual existence and statistical stability of some of these families, we resorted to use the *main families* (groups), based exclusively in the average likelihood of each sample being matched against samples of the same class and of different classes, resulting in the three families listed in Table 1: while the"Sheep" family contains the large majority of the subjects, and corresponds to the average system performance, the remaining families represent the other extreme cases:"Goats" are particularly hard to match (either for *genuine* or *impostor* comparisons), and"Lambs/Wolves", on the other way, are very easy to impersonate.

Formally, let $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ be a dataset of n images, each one associated with a specific class (ID). We use the notation $d(\mathbf{x}_i, \mathbf{x}_j)$ to represent the distance between feature vectors extracted from images $\mathbf{x}_i/\mathbf{x}_j$. In an *all-against-all* paradigm, we can obtain the sets of genuine/impostor scores, which can be further divided per subject or feature. To obtain the discriminating power of each sample, we consider the decidability index:

$$d' = \frac{\mu_G - \mu_I}{\sqrt{\sigma_G^2 + \sigma_I^2}}, \quad (2)$$

where μ_G, μ_I represent the mean of the genuine and impostor scores, and σ_G, σ_I are the respective standard deviation values.

4. Experiments and results

This section presents the comprehensive experiments conducted and the results obtained in our study on human re-ID feature importance. We detail the implementation of the evaluated methods, the analysis of feature attributes and their correlations, the biometric menagerie family analysis, and the overall performance of the re-ID systems under various scenarios.

4.1. Implementation details

In our experiments, we used the implementations provided in [52].[3] Concerning the P-DESTRE dataset, we used its full set of 1894 tracks, which includes 608 individuals, averaging 67.4 frames per tracklet. The comprehensive specifications of the samples for learning, validation, and testing are available at.[4] For the MLFN model, we used the Adam optimizer with a mini-batch size of 64. The initial learning rate was set to 0.00035 and we incorporated a decay rate of 0.0005 across 307 training epochs, as described in the original paper. In the MuDeep configuration, we used the stochastic gradient descent algorithm with a mini-batch size of 32. The learning rate started at 0.001 and was

[3] https://github.com/KaiyangZhou/deep-person-reid
[4] http://p-destre.di.ubi.pt/pedestrian_reid_splits.zip

Table 5
Distribution of features across the *Biometric Menagerie* categories. Each entry presents the absolute proportion and the likelihood (λ) of an individual, characterized by a particular feature value, being classified into each of the"Sheep","Goats", or"Lambs/Wolves" menagerie families. The likelihood (λ) values correspond to the division of the absolute probability by the prior probability of that attribute.

Feature	Value	Sheep		Goat		Lambs/Wolves	
		Absolute	λ	Absolute	λ	Absolute	λ
Gender	Male	0.54	0.89	0.03	0.05	0.04	0.06
	Female	0.32	0.82	0.03	0.08	0.04	0.10
Height	Short	0.12	0.70	0.02	0.11	0.03	0.17
	Medium	0.59	0.88	0.04	0.05	0.05	0.07
	Tall	0.09	0.64	0.03	0.21	0.02	0.14
Ethnicity	Indian	0.78	0.90	0.04	0.04	0.05	0.06
	White	0.09	0.75	0.01	0.08	0.02	0.16
	Black	0.006	0.60	0.001	0.10	0.003	0.30
Hair Color	Black	0.86	0.92	0.02	0.03	0.04	0.04
	Occluded	0.002	0.60	0.001	0.10	0.002	0.22
	Brown	0.02	0.68	0.004	0.13	0.006	0.18
Hair Style	Short	0.46	0.84	0.04	0.07	0.05	0.09
	Horse Tail	0.26	0.81	0.03	0.09	0.03	0.09
Beard	No	0.59	0.88	0.03	0.04	0.05	0.07
	Yes	0.25	0.81	0.02	0.06	0.04	0.13
Mustache	No	0.46	0.85	0.02	0.04	0.06	0.11
	Yes	0.36	0.82	0.03	0.07	0.05	0.11
Glasses	No	0.62	0.87	0.04	0.06	0.05	0.07
	Normal	0.20	0.80	0.02	0.08	0.03	0.12
Head Accessories	Hat	0.01	0.66	0.003	0.20	0.002	0.14
	Cannot see	0.95	0.98	0.01	0.01	0.01	0.01
Upper Body Cloths	T-Shirt	0.26	0.84	0.03	0.10	0.02	0.06
	Blouse	0.23	0.82	0.02	0.07	0.03	0.11
	Dress	0.03	0.72	0.004	0.08	0.01	0.14
	Hoodie	0.07	0.75	0.01	0.10	0.015	0.15
	Shirt	0.17	0.77	0.02	0.09	0.03	0.14
Lower Body Cloths	Jeans	0.48	0.86	0.03	0.05	0.05	0.09
	Leggins	0.26	0.84	0.03	0.10	0.02	0.06
	Pants	0.065	0.80	0.005	0.07	0.01	0.13
	Dress	0.03	0.72	0.004	0.08	0.01	0.14
Feet	Sports Shoes	0.28	0.90	0.01	0.03	0.02	0.07
	Sandals	0.55	0.83	0.05	0.08	0.06	0.09
Accessories	Nothing	0.007	0.74	0.003	0.13	0.003	0.13
	Backpack Bag	0.71	0.90	0.04	0.05	0.04	0.05
Action	Walking	0.90	0.96	0.01	0.01	0.03	0.03
	Standing	0.03	0.72	0.004	0.08	0.01	0.14

decreased by a factor of 10 at specific intervals as described in the original paper. For the ResNet model, the default settings from the framework by [52][5] were maintained.

4.2. Feature analysis: Prior probabilities and correlations

The P-DESTRE dataset contains examples of various attributes, each with a distinct set of possible discrete values. Fig. 2 provides the prior probabilities of these attributes (features), for each possible value that can be assigned to a feature. For instance, the probability of head accessories not being visible is significantly high (97%), indicating that the large majority of the individuals in the dataset were not wearing any head accessories during data acquisition.

With respect to the main features, in terms of"gender", the dataset leans slightly towards males, with their prior probability being 61% compared to 39% for females. Furthermore, a majority of individuals 67% fall within the"medium height" category. The dataset displays a noticeable distinction in"mustache" occurrence, with a 44% probability of presence compared to 54% of absence. When it comes to hairstyles,"short hair" is the most prevalent, observed in 55% of individuals, followed by a"horse tail" style, mostly seen in females, with a 33% occurrence.

In terms of ethnicity, the dataset is predominantly composed of individuals of"Indian" origin 87%, with 93% having"Black hair". The majority of individuals (94%) are engaged in"walking" action. When it comes to clothing,"jeans" (53%) and"leggings" (31%) are the most common choices for lower body attire, while"t-shirt" (31%),"blouse" (28%), and"shirt" (22%) predominate for the upper part of the body. Footwear choices lean towards"sandals" (66%) and"sports shoes" (31%). As for accessories,"backpacks" are prevalent (79%), whereas 18% of individuals do not carry any accessories.

We also observed a significant influence of gender on the correlation with other features. Males exhibit high correlations with carrying backpacks, walking activities, wearing sports shoes, jeans, and t-shirts, not wearing glasses, having black hair, short hairstyles, beard, mustache, medium height, and Indian ethnicity. Conversely, they show a low correlation with having a short height, sporting a horsetail hairstyle, wearing a dress as upper body clothing, and certain footwear classifications. In opposition, females exhibit high correlations with features such as sporting a horse hairstyle, having a short height, wearing blouses as upper-body clothing, choosing leggings as lower-body attire, and carrying bags as accessories. These findings provide crucial insights into the relationships between gender and various features, contributing significantly to the development and enhancement of human re-identification methods.

4.3. Re-ID performance

For our empirical evaluation, we started by perceiving the contribution of each feature value on the re-ID performance.

At first, in terms of the covariance/agreement between the results provided by the different methods considered, Table 2 provides a correlation matrix showcasing the Pearson correlation between the

[5] https://github.com/KaiyangZhou/deep-person-reid

matching scores obtained for the three models considered.

Then, Table 3 can be understood as the main source for our discussion/findings and reports the re-ID performance (in terms of mAP values) when filtering the probe matching scores with respect to the features/values considered. It should be noted that - in all cases - the gallery set used remains fixed, with the variations in performance measured across 14 distinct features. As baseline, the first row of this table provides the values obtained when no filter was applied, i.e., when all the probe scores were considered, which was regarded as the baseline performance. Then, for each feature/value, the measured the relative performance with respect to the the baseline: upward green triangles represent feature/values observed to contribute to *increase* performance, indicating that the model performs particularly well when identifying individuals that have these feature values. Conversely, downward red triangles represent feature values that appear to contribute to decrease the system performance, suggesting that the model has particular difficulties in handling individuals that have such attribute value. For visualization and summarization purposes, we display a double up-down triangle for the *particularly evident* changes in performance, i.e., when the relative values changed over 10% in magnitude.

Among the methods tested, we observed that the MLFN outperformed its competitors in most cases, and resorted to this technique in the subsequent experiments carried out.

Further, Table 4 summarizes the overall re-ID performance, in terms of the decidability scores (2) for the different feature values. These results offer a fine understanding of the role that each feature value has in the re-identification performance of the best method (MLFN). One striking observation is the high d' value of 18.34 associated with the action of”walking”, suggesting that the distinct biomechanical attributes captured during walking might offer a unique signature, making it easier for the model to differentiate between genuine and impostor scores.

In opposition, gender and ethnicity appear to also play significant roles, though to a lesser extent than the”walking” actions. The d' value for males is 14.35, higher compared to females at 10.64, possibly indicating that the features captured for males are more distinctive or that the model has a training bias towards male attributes. Ethnicity also contributes, with Indian ethnicity having the highest d' value of 16.35, suggesting a better re-identification rate for this group. However, challenges in re-identification are evident. For instance, when no accessories are present, the d' value drops to a low of 2.69, indicating that the absence of additional attributes makes it more challenging for the model to distinguish between individuals. Similarly, static actions like standing result in a lower d' value of 3.91, which might be attributed to the limited movement and fewer distinctive features to capture. Other features show variable d' values; for example, black hair color has a high d' value of 19.47. In contrast, facial features like the presence of a beard or a mustache do not significantly improve the d', suggesting that these features alone are not highly distinctive. The variability in d' values across different features highlights the complex nature of human re-identification tasks. While some features offer high decidability, others pose challenges.

For subjective perception purposes, Fig. 5 illustrates some examples of genuine and impostor comparisons with respect to attributes like gender, glasses, hair color, head accessories, upper body (UB) clothes, and other accessories. Each category has a query image (Q) and several corresponding images that are ranked based on their similarity to the query.

4.4. Biometric menagerie results

In this section, we analyze the performance of re-ID in leans of *Biometric Menagerie*, with reference to Table 1 for various feature values. Again, values regard the MLFN method. Table 5 provides an overview of the distribution of person features across the *Biometric Menagerie* category. Each entry presents the *absolute proportion* and the *likelihood* (λ) of an feature, characterized by a particular value, being classified into each of the *”Sheep”, ”Goats”, or”Lambs/Wolves”* categories. The likelihood (λ) values were calculated by dividing the *absolute probability* by the *prior probability* of that features value.

These results reveal several interesting insights: for instance,”gender” appears to have a moderate impact, with males showing a higher likelihood ($\lambda = 0.89$) of being classified as Sheep compared to females ($\lambda = 0.82$).”Height” also influences re-ID: individuals of medium height are more likely to be typed as Sheep ($\lambda = 0.88$), while those who are tall are more likely to be Goats ($\lambda = 0.21$).”Ethnicity”-wise, the system performs best for individuals of Indian origin ($\lambda = 0.90$) for the Sheep category. In terms of”Hair” features, black-haired individuals have the highest likelihood of being typed as Sheep ($\lambda = 0.92$). Interestingly, while facial features like a beard or mustache don't significantly alter the likelihood of being a Sheep, they do affect the Goat and Lambs/Wolves categories. Actions also play a role; the act of walking increases the likelihood of an individual being classified as Sheep ($\lambda = 0.96$). These results collectively corroborate the reID varying performance across different physical and behavioral attributes.

As shown in Fig. 7, for each of these features, the likelihood in the *goats* and *lambs/wolves* categories generally increases as the likelihood in the *sheep* category decreases, suggesting a trade-off in the method's ability to accurately match these features.

Overall, the above observations highlight the variability in the performance of the re-ID method across different feature values. It suggests that some subsets of the population are more easily and accurately matched by the method than others, while some of them are also more likely to lead to false rejections (*goats* family) or prune to false acceptances (*lambs/wolves*). Understanding these patterns is regarded as very important, not only for improving the re-ID overall performance but also to obtain unbiased - and more fair - systems.

5. Discussion

It should be noted that, even if our results point for the existence of certain subgroups of the population that pose additional challenges in re-ID, there are other other effects that were considered irrelevant or apriori expected. For example, we found a strong correlation between the”male” gender and specific features such as”carrying backpacks”, engaging in”walking” activities, wearing”sports shoes”, and”t-shirts” as upper-body clothing and jeans as lower-body clothing. Oppositely,”female” element exhibited a high correlation with characteristics such as”horsetail hairstyle”, having a”short height”, wearing”blouses” as upper-body clothing, choosing”leggings” as lower-body attire, head accessories scarf, and carrying bags as accessories. These findings are consistent with the observations made by [53], who noted that the most significant portion of an image often contains the highest-rated attributes, while rarely occurring attributes can be most distinguishing for person re-id. Also, we observed that upper-body clothing, lower-body clothing, accessories, ethnicity, head accessories etc. feature values can play a major role in re-ID effectiveness. This observation matched the findings of [54,55], which reported that features such as hair, gender, upper body and lower body clothing influenced the performance of re-ID methods; however, performance was heavily degraded with respect to features such as sunglasses.

Also, considering the concept of the *Biometric Menagerie*, our results (Table 5) can be generally regarded as aligned to the findings of [56,57], which reported poor performance for features such as hair occluded, feet sandals, glasses sun/normal, accessories, hair color brown, hairstyle bald and long hair etc..

As a final illustration, Fig. 6 provides some of the examples that were the most typically classified into the biometric families types:”Goats”,”Lambs”, and”Wolves”. Subjectively, this has the potential to identify groups of individuals who may pose additional challenges to the re-ID process, due to their features. These conclusions may help researchers to focus on particular family types, and to prioritize and address the challenges specific to each group more effectively.

6. Conclusions

In this paper, we reported a comprehensive assessment about the effect of individual-specific subject features (e.g.,”gender”,”ethnicity” and”height”) on the performance of human re-identification (re-ID) methods. Our analysis underscored the significant influence of some features such as”head accessories”,”ethnicity” and”height” on the re-ID accuracy. In opposition, there are other features (such as”gender” and”beard”) that appear to have minor effects in the re-ID effectiveness. We consider that the results provided in this regard can be extremely valuable to support not only further developments in the re-ID technology, but also constitute an important source of information towards the availability of unbiased (and more fair) re-ID systems that minimize the probability of treating in an unfair way some of the subgroups of a population.

Additionally, we considered the main families of the *Biometric Menagerie*, and estimated the likelihood of individuals of a certain feature value (e.g.,”Male”,”Tall”, or”Indian”) being classified into “Sheep” (regular users of a biometric system), “Goats” (subjects that might be particularly hard to match) or “Lambs/Wolves” (subjects that might be easily impersonated). Overall, our findings corroborate previous studies that reported that, while recognition methods gain proficiency in recognizing certain sub-groups of the population, they also encounter serious difficulties with other sub-groups, which unavoidably will bias the effectiveness of recognition between different individuals, depending of the sub-groups where they belong.

According to our findings, we point for several directions that should be the focus of further work: 1) enhancing feature recognition is, among all other topics, the most crucial; future research should focus on improving the robustness of the resulting features, in particular to re-ID specific attributes such as hair occlusion, varying clothing styles (for long-term re-id); 2), there is a critical need for bias mitigation strategies, particularly those related to gender, ethnicity, and physical attributes that might be socially undetood as discriminating; and 3) expanding the diversity of the available datasets, while also keeping them unbiased remains as an open problem. Such sets would be the key to improve the generalization capabilities of re-ID systems.

CRediT authorship contribution statement

Kailash Hambarde: Data curation, Methodology, Software, Validation, Visualization, Writing – original draft. **Hugo Proença:** Conceptualization, Investigation, Methodology, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] Iwama Haruyuki, Muramatsu Daigo, Makihara Yasushi, Yagi Yasushi, Gait verification system for criminal investigation, Inform. Media Technol. 8 (4) (2013) 1187–1199.
[2] Puneet Kohli, Anjali Chadha, Enabling pedestrian safety using computer vision techniques: A case study of the 2018 uber inc. self-driving car crash, in: Advances in Information and Communication: Proceedings of the 2019 Future of information and communication conference (FICC) vol. 1, Springer, 2020, pp. 261–279.
[3] Hirzer Martin, Beleznai Csaba, M. Roth Peter, Bischof Horst, Person re-identification by descriptive and discriminative classification, in: Image Analysis: 17th Scandinavian Conference, SCIA 2011, Ystad, Sweden, May 2011. Proceedings 17, Springer, 2011, pp. 91–102.
[4] S.V. Aruna Kumar, Ehsan Yaghoubi, Abhijit Das, B.S. Harish, Proença Hugo, The p-destre: a fully annotated dataset for pedestrian detection, tracking, and short/long-term re-identification from aerial devices, IEEE Trans. Inf. Forensics Secur. 16 (2020) 1696–1708.
[5] Sameh Khamis, Cheng-Hao Kuo, Vivek K. Singh, Vinay D. Shet, Larry S. Davis, Joint learning for attribute-consistent person re-identification, in: Computer Vision-ECCV 2014 Workshops: Zurich, Switzerland, September 6–7 and 12, 2014, Proceedings, Part III 13, Springer, 2015, pp. 134–146.
[6] Yutian Lin, Liang Zheng, Yu Wu Zhedong Zheng, Hu Zhilan, Chenggang Yan, Yi Yang, Improving person re-identification by attribute and identity learning, Pattern Recogn. 95 (2019) 151–161.
[7] Zhangqiang Ming, Min Zhu, Xiangkun Wang, Jiamin Zhu, Junlong Cheng, Chengrui Gao, Yong Yang, Xiaoyong Wei, Deep learning-based person re-identification methods: a survey and outlook of recent works, Image Vis. Comput. 119 (2022) 104394.
[8] Mang Ye, Jianbing Shen, Gaojie Lin, Tao Xiang, Ling Shao, Steven C.H. Hoi, Deep learning for person re-identification: a survey and outlook, IEEE Trans. Pattern Anal. Mach. Intell. 44 (6) (2021) 2872–2893.
[9] Chuanchen Luo, Yuntao Chen, Naiyan Wang, Zhaoxiang Zhang, Spectral feature transformation for person re-identification, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 4976–4985.
[10] Yicheng Wang, Zhenzhong Chen, Feng Wu, Gang Wang, Person re-identification with cascaded pairwise convolutions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 1470–1478.
[11] Chunfeng Song, Yan Huang, Wanli Ouyang, Wang Liang, Mask-guided contrastive attention model for person re-identification, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 1179–1188.
[12] Weihua Chen, Xiaotang Chen, Jianguo Zhang, Kaiqi Huang, Beyond triplet loss: a deep quadruplet network for person re-identification, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 403–412.
[13] Xuelin Qian, Fu Yanwei, Yu-Gang Jiang, Tao Xiang, Xiangyang Xue, Multi-scale deep learning architectures for person re-identification, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 5399–5408.
[14] Yifan Sun, Qin Xu, Yali Li, Chi Zhang, Yikang Li, Shengjin Wang, Jian Sun, Perceive where to focus: learning visibility-aware part-level features for partial person re-identification, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 393–402.
[15] Wenjie Yang, Houjing Huang, Zhang Zhang, Xiaotang Chen, Kaiqi Huang, Shu Zhang, Towards rich feature discovery with class activation maps augmentation for person re-identification, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 1389–1398.
[16] Xin Ning, Ke Gong, Weijun Li, Liping Zhang, Jwsaa: joint weak saliency and attention aware for person re-identification, Neurocomputing 453 (2021) 801–811.
[17] Zhizheng Zhang, Cuiling Lan, Wenjun Zeng, Xin Jin, Zhibo Chen, Relation-aware global attention for person re-identification, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 3186–3195.
[18] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio, Generative adversarial networks, Commun. ACM 63 (11) (2020) 139–144.
[19] Yun-Chun Chen, Yu-Jhe Li, Du Xiaofei, Yu-Chiang Frank Wang, Learning resolution-invariant deep representations for person re-identification, in: Proceedings of the AAAI Conference on Artificial Intelligence vol. 33, 2019, pp. 8215–8222.
[20] Hao Chen, Yaohui Wang, Benoit Lagadec, Antitza Dantcheva, Francois Bremond, Joint generative and contrastive learning for unsupervised person re-identification, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 2004–2013.
[21] Yukun Huang, Zheng-Jun Zha, Fu Xueyang, Richang Hong, Li Liang, Real-world person re-identification via degradation invariance learning, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 14084–14094.
[22] Yixiao Ge, Zhuowan Li, Haiyu Zhao, Guojun Yin, Shuai Yi, Xiaogang Wang, et al., Fd-Gan: pose-guided feature distilling Gan for robust person re-identification, Adv. Neural Inf. Proces. Syst. 31 (2018).
[23] Yu-Jhe Li, Yun-Chun Chen, Yen-Yu Lin, Du Xiaofei, Yu-Chiang Frank Wang, Recover and identify: A generative dual model for cross-resolution person re-identification, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 8090–8099.
[24] Dapeng Chen, Hongsheng Li, Tong Xiao, Shuai Yi, Xiaogang Wang, Video person re-identification with competitive snippet-similarity aggregation and co-attentive snippet embedding, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 1169–1178.
[25] Xingyu Liao, Lingxiao He, Zhouwang Yang, Chi Zhang, Video-based person re-identification via 3d convolutional networks and non-local attention, in: Computer Vision–ACCV 2018: 14th Asian Conference on Computer Vision, Perth, Australia, December 2–6, 2018, Revised Selected Papers, Part VI 14, Springer, 2019, pp. 620–634.
[26] Jianing Li, Shiliang Zhang, Tiejun Huang, Multi-scale 3D convolution network for video-based person re-identification, in: Proceedings of the AAAI Conference on Artificial Intelligence 33, 2019, pp. 8618–8625.
[27] Jinrui Yang, Wei-Shi Zheng, Qize Yang, Ying-Cong Chen, Qi Tian, Spatial-temporal graph convolutional network for video-based person re-identification, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 3289–3299.

[28] Yichao Yan, Jie Qin, Jiaxin Chen, Li Liu, Fan Zhu, Ying Tai, Ling Shao, Learning multi-granular hypergraphs for video-based person re-identification, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 2899–2908.
[29] Wu Yiming, Omar El Farouk, Xi Li Bourahla, Wu Fei, Qi Tian, Xue Zhou, Adaptive graph representation learning for video person re-identification, IEEE Trans. Image Process. 29 (2020) 8821–8830.
[30] Jiawei Liu, Zheng-Jun Zha, Wei Wu, Kecheng Zheng, Qibin Sun, Spatial-temporal correlation and topology learning for person re-identification in videos, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 4370–4379.
[31] Meiyan Huang, Chunping Hou, Qingyuan Yang, Zhipeng Wang, Reasoning and tuning: graph attention network for occluded person re-identification, IEEE Trans. Image Process. 32 (2023) 1568–1582.
[32] Yuqi Zhang, Qi Qian, Hongsong Wang, Chong Liu, Weihua Chen, Fan Wan, Graph convolution based efficient re-ranking for visual retrieval, IEEE Trans. Multimed. 26 (2024) 1089–1101, https://doi.org/10.1109/TMM.2023.3276167.
[33] Yuqiao Xian, Jinrui Yang, Fufu Yu, Jun Zhang, Xing Sun, Graph-based self-learning for robust person re-identification, in: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2023, pp. 4789–4798.
[34] X. Nguyen Binh, Binh D. Nguyen, Tuong Do, Erman Tjiputra, Quang D. Tran, Anh Nguyen, Graph-based person signature for person re-identifications, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 3492–3501.
[35] Ahmet Iscen, Giorgos Tolias, Yannis Avrithis, Ondřej Chum, Cordelia Schmid, Graph convolutional networks for learning with few clean and many noisy labels, in: Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part X 16, Springer, 2020, pp. 286–302.
[36] Zhong Zhang, Haijia Zhang, Shuang Liu, Yuan Xie, Tariq S. Durrani, Part-guided graph convolution networks for person re-identification, Pattern Recogn. 120 (2021) 108155.
[37] Hongmin Liu, Zhenzhen Xiao, Bin Fan, Hui Zeng, Yifan Zhang, Guoquan Jiang, Prgcn: probability prediction with graph convolutional network for person re-identification, Neurocomputing 423 (2021) 57–70.
[38] Yantao Shen, Hongsheng Li, Shuai Yi, Dapeng Chen, Xiaogang Wang, Person re-identification with deep similarity-guided graph neural network, in: Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 486–504.
[39] Xu Ruyu, Yueyang Zheng, Xiaoming Wang, Dong Li, Person re-identification based on improved attention mechanism and global pooling method, J. Vis. Commun. Image Represent. 94 (2023) 103849.
[40] Mengxi Jia, Yifan Sun, Yunpeng Zhai, Xinhua Cheng, Yi Yang, Ying Li, Semi-attention partition for occluded person re-identification, in: Proceedings of the AAAI Conference on Artificial Intelligence 37, 2023, pp. 998–1006.
[41] Heyu Chang, Qu Dan, Kun Wang, Hongqi Zhang, Nianwen Si, Gengxiao Yan, Huazhong Li, Attribute-guided attention and dependency learning for improving person re-identification based on data analysis technology, Enterprise Inform. Syst. 17 (3) (2023) 1941274.
[42] Zhe Wang, Jun Wang, Junliang Xing, Attribute-guided transformer for robust person re-identification, IET Comput. Vis. (2023), https://doi.org/10.1049/cvi2.12215.
[43] Xu Simin, Lingkun Luo, Hu Jilin, Bin Yang, Hu. Shiqiang, Semantic driven attention network with attribute learning for unsupervised person re-identification, Knowl.-Based Syst. 252 (2022) 109354.
[44] Kunlong Bai, Fu Saiji, Linrui Yang, Dalian Liu, Fine-grained attribute-aware analysis for person re-identification, Proc. Comp. Sci. 199 (2022) 276–283.
[45] Nadeen Shoukry, Mohamed A. Abd, El Ghany, Mohammed A.-M. Salem, Multi-modal long-term person re-identification using physical soft bio-metrics and body figure, Appl. Sci. 12 (6) (2022) 2835.
[46] Cuiqun Chen, Mang Ye, Ding Jiang, Towards modality-agnostic person re-identification with descriptive query, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 15128–15137.
[47] Nianchang Huang, Kunlong Liu, Yang Liu, Qiang Zhang, Jungong Han, Cross-modality person re-identification via multi-task learning, Pattern Recogn. 128 (2022) 108653.
[48] Wu Duidi, Haiqing Huang, Qianyou Zhao, Shuo Zhang, Jin Qi, Hu. Jie, Overview of deep learning based pedestrian attribute recognition and re-identification, Heliyon (2022) e12086.
[49] Xiaobin Chang, Timothy M. Hospedales, Tao Xiang, Multi-level factorisation net for person re-identification, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 2109–2118.
[50] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
[51] Neil Yager, Ted Dunstone, The biometric menagerie, IEEE Trans. Pattern Anal. Mach. Intell. 32 (2) (2008) 220–230.
[52] Zhou Kaiyang, Xiang Tao, Torchreid: A library for deep learning person re-identification in pytorch, arXiv (2019), https://doi.org/10.48550/arXiv.1910.10093 preprint arXiv:1910.1009.
[53] Arne Schumann, Rainer Stiefelhagen, Person re-identification by deep learning attribute-complementary information, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2017, pp. 20–28.
[54] Yubin Deng, Ping Luo, Chen Change Loy, Xiaoou Tang, Pedestrian attribute recognition at far distance, in: Proceedings of the 22nd ACM International Conference on Multimedia, 2014, pp. 789–792.
[55] Su Chi, Shiliang Zhang, Junliang Xing, Wen Gao, Qi Tian, Deep attributes driven multi-camera person re-identification, in: Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part II 14, Springer, 2016, pp. 475–491.
[56] Su Chi, Shiliang Zhang, Junliang Xing, Wen Gao, Qi Tian, Multi-type attributes driven multi-camera person re-identification, Pattern Recogn. 75 (2018) 77–89.
[57] Huimin Li, Shuzhao and. Yu and Roland Hu., Attributes-aided part detection and refinement for person re-identification, Pattern Recogn. 97 (2020) 107016.