This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2023.3234494

Annotated Pedestrians: A Dataset for Soft Biometrics Estimation for Varying Distances

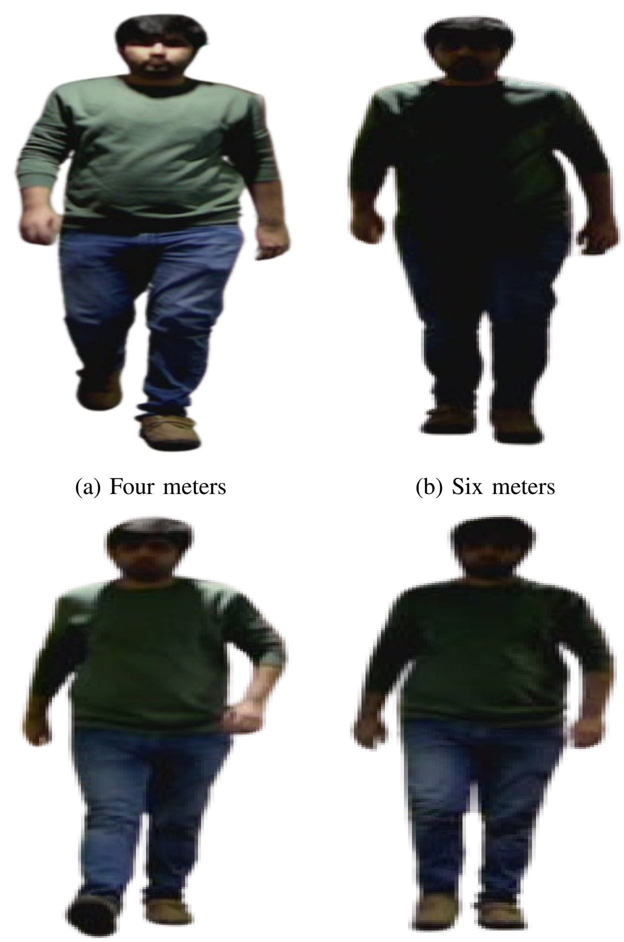
Bilal Hassan, *Member, IEEE*, Muhammad Fiaz, *Student Member, IEEE*, Hafiz Husnain Raza Sherazi, *Senior Member, IEEE* and Usman Javed Butt, *Member, IEEE*

***Abstract*—Following the significance of soft biometrics to facilitate seamless recognition or retrieval, the need for multi-modality annotated datasets is increasing - to evaluate any standalone soft biometrics system. Although, large-size datasets like PETA were annotated to evaluate soft biometrics systems, however, they were mainly annotated for global soft biometrics such as gender and age and for clothing modality. By looking at the usefulness of multiple modalities of the human body during recognition or retrieval, we designed, developed and annotated a new dataset called Annotated Pedestrians for the individuals. The images in the dataset were explicitly recorded for the individuals at four different distances from the camera and they incorporate annotations for four different modalities of the human body i.e., i) global soft biometrics, ii) extended facial region, iii) body including limbs, and iv) clothing with attachments. The annotation process was expert opinion and qualitative annotation types were used. There were a total of three global soft biometrics annotated and for remaining three modalities, categorical annotations for 46 soft biometrics were performed. In terms of comparative annotations, there were a total of 26 soft biometrics annotated for the same three modalities. To the best of our knowledge, Annotated Pedestrians is a unique dataset designed by incorporating the impact of distance during recognition or retrieval, where markers were placed on the surface at 4, 6, 8, and 10 m distances from the camera, and approximately 300 frames were recorded for 50 distinct individuals in a 20 m long corridor. Moreover, the usefulness of the dataset is annotation using four different modalities of the human body, and a total of 75 soft biometrics annotated using a qualitative approach – making Annotated Pedestrians a highly-diverse dataset to evaluate any soft biometrics system for recognition during short-term tracking and feature-based retrieval from the database.**

***Index Terms*—Soft biometrics, distance, dataset, annotations, categorical, comparative.**

I. Introduction

The major requirement for the evaluation of any benchmark soft biometric system is the availability of an annotated dataset specifically developed for this purpose [1]. Several existing facial and pedestrians datasets such as LFW [2], PETA [3], LFW Updates [4], ATVS Forensic DB [5] and MORPH [6], were used in the past for the training and evaluation of soft biometrics systems [7]. None of them was focusing on all the modalities [8] of the human body at once rather they cover one or few modalities simultaneously e.g., face or body. Moreover, a limited number of instances were recorded for different individuals and the repetition of the same individual over time was missing in most of the datasets [9]. Usually, the datasets introduce several different environmental constraints while recording e.g., lighting condition [10], subject angle from the camera [11], and distance [9], to name a few. Furthermore, the pedestrian datasets with higher diversity in terms of gender, age, and ethnicity are always considered better, whether using traditional or soft biometrics, for recognition in surveillance or retrieval from a large dataset [8]. Although, the datasets like Soton Gait [12] and Southampton Tunnel [13] were specifically developed and annotated for the soft biometrics-based recognition or retrieval, however, the number of distinct individuals is limited and only a few sessions were recorded in a controlled environment.



(a) Four meters (b) Six meters

(c) Eight meters (d) Ten meters

Fig. 1: Images from Annotated Pedestrians dataset at four different distances from the camera [14]

Currently, recognition or retrieval at distance [15] is one of the most critical problems in computer vision [16], due

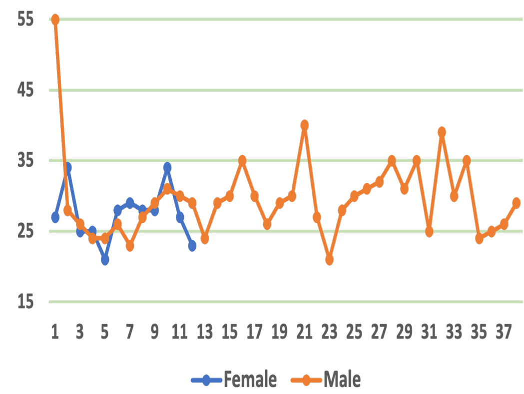
This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2023.3234494

to non-availability of such large size image datasets, where same individual was recorded at different distances [17] from the camera, introducing higher diversity and at the same time, the dataset annotation using multiple different soft biometrics modalities of the human body. To address those challenges, we proposed, developed and annotated a new dataset called Annotated Pedestrians. The dataset was originally recorded for 50 distinct individuals, including 38 males and 12 females, where the age of participants ranges from 15 to 60 years. There were five different ethnic groups who participated in the dataset recording from 19 different nationalities.

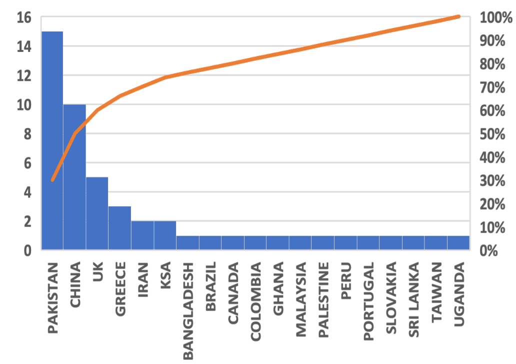
To the best of our knowledge, the Annotated Pedestrians is a unique and largest dataset of its kind which is annotated using three most widely used global soft biometrics. It introduces several factors related to diversity in the data like gender, age and ethnicity including country-specific information. One of the unique aspects related to Annotated Pedestrians dataset is the recording and annotation of a new ethnic group called 'Mix', to incorporate cross-ethnic information for better model training, beside existing common ethnic groups like African, Asian, Arab and White. On the other hand, people from a large age range participated in the recording process and during annotation, they were placed in four different age groups. These age categories known as youngsters (15 – 25), mature (25 – 36), experienced (36 – 45) and aged (45 – 60) are the most common ones observed during any recognition or retrieval tasks. Finally, gender is one of the most distinguishing global soft biometrics associated with human body; however, a balance between both gender's count always remains less in most of the datasets. In Annotated Pedestrians, the number of female participants is one third of the male participants which is slightly higher than the existing datasets of this type. Despite the fact, the diverse ranges of age and ethnicity for female participants in our Annotated Pedestrians dataset makes it flexible for applying techniques like data augmentation to balance the gender ratio and to overcome the limitation of least diversity.

The actual dataset was a stream of approximately 300 frames with a resolution of 1280 X 720 for each individual walking in a long corridor towards the camera. Later, the automated extraction of frames at four different distances from the camera i.e., 4, 6, 8, and 10 m was performed and verified. The Annotated Pedestrians is specially designed, developed, and annotated pedestrians dataset for the evaluation of soft biometrics based recognition or retrieval systems. It captures four different soft biometrics modalities of human body i.e., global, extended facial region, body including limbs and clothing with attachments. A summary of the annotated dataset using global soft biometrics like gender, age range and ethnic group for each individual is presented in Figure 2.

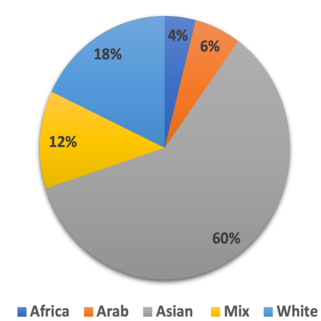
In our earlier research on soft biometrics, both images and human body annotations using soft biometrics were identified as meaningful source of information during recognition or retrieval. Sometimes, they were used as source and target for each other in the past; however, a better recognition or retrieval is possible if both work together. It mainly depends on the scenario, whether recognition or retrieval, like both image of an individual and whole-body annotations can better train a recognition model, while images stored with whole-body annotations may improve either feature or image-based retrieval process.



(a) gender based participant's age distribution



(b) Country-specific distribution of participants



(c) Ethnicity-specific distribution of participants

Fig. 2: Dataset distribution using global soft biometrics

Distance impacts soft biometrics based recognition or retrieval systems as in traditional biometrics [18] and open environments causes further complexity [19]. The authors in [1] compare the impact of distance using three different modalities of the human body and it was observed that soft biometrics from the body and clothing are easier to precisely estimate than of a face. A new dataset of soft biometrics from three different modalities of the human body is developed in [17], where 10 features from each modality were annotated. The same individuals were recorded at three distances from the camera e.g., close, medium and far. Pearson correlation [20] was used to measure sensitivity of the distance from the camera and it was determined that body and clothing are less sensitive to distance than face. In another experiment conducted by [21], to analyze the impact of continuously changing distance from the camera, 23 soft biometrics from head, body, and global domain were selected [22]. The aim of the experiment was to determine which of the modalities accurately and quickly predicts the gender from the far distance and it was reported that bodily soft biometrics are the stronger candidate. In both the analysis, the qualitative annotation types [23] such as categorical and comparative were used.

Annotating a dataset is an essential requirement before evaluation of any soft biometrics based recognition or retrieval system [24]. Usually, soft biometrics involve several permanent or temporary modalities of the human body, considering many different features. In [1], permanent modalities such as face and body while temporary modalities such as clothing with attachments were identified and a largest collection of more than 150 soft biometric features is presented. They also discussed several different methods for dataset annotation such as expert opinion, crowd sourcing, and measured, while the annotation types like absolute, categorical and comparative were introduced [25]. Generally, the dataset annotation is a highly sensitive process, despite selecting the most common annotation types like categorical [26] and comparative [27]. On the other hand, the annotation method like crowd sourcing is very expensive and non-availability of experienced annotators along with high training costs are the biggest problems. In this research, we employed expert opinion-based method to annotate 75 soft biometrics from four different modalities of the human body and the annotation types were categorical and comparative in case of all soft biometrics. Both categorical and comparative types of dataset annotation are qualitative, and they are useful in several application domains such as short-term tracking and feature-based retrieval from dataset, respectively [23].

The organization of this paper is as follows. The information about dataset acquisition process, critical discussion on annotation methods and types, and selection of Soft biometrics from each modality of the human body is presented in Section II. The dataset annotation process using categorical and comparative methods for soft biometrics annotation from each modality is demonstrated in Section III. The section IV provides a reflection on complete soft biometrics annotation process demonstrated earlier. Finally, the conclusion covers key contributions made and future course of action points as part of Section V.

II. About Annotated Pedestrians

As discussed earlier, Annotated Pedestrians is a dataset of people walking in a corridor towards the camera. The dataset was designed by keeping in mind the impact of distance [28] during recognition or retrieval. The walking corridor was marked with green color strips to verify the automated frame extraction process at different distances [29] from the camera. To record the dataset the design aspects [30] of the corridor were given significant importance - the details for which are provided ahead. Usually, temporal features are not considered good for recognition or retrieval process over a long period of time; however, this phenomenon in mostly associated with traditional biometrics, using a couple of features from a single modality. On the other hand, soft biometrics are multi-modality and a richer set of features than traditional biometrics with diverse annotation types. Moreover, several properties like multiple annotation types, permanence score and discrimination power are linked with each soft biometric and determine the temporal behaviour at both each individual modality and feature level.

A. Recording Environment

To record Annotated Pedestrians dataset, a 20 m long corridor was designed which was sufficient to capture minimum 300 frames at a rate of 30 FPS [24]. The corridor was providing a reflection of entry-point for any real-time infrastructure or building. The aim of designing such type of corridor was to give an impression close to the reality by presence of several other real-world objects. To incorporate the distance factor, markers were placed at 4, 6, 8, and 10 m distance from the camera. As in [17], the impact of distance was observed during person recognition process and it was noticed that face presents higher accuracy of recognition at only the close distance from the camera, which is not the case with body. Using these observations, we placed those green markers on the floor and extracted the frames at four different distances from the camera. These extracted frames are useful to train and test any soft biometrics based recognition or retrieval system.

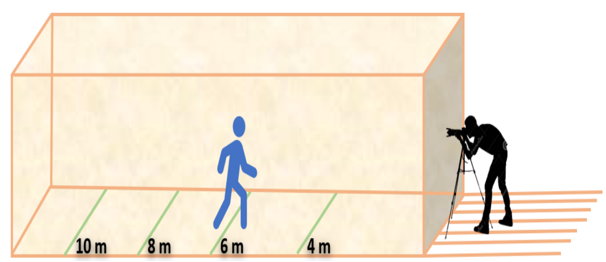


Fig. 3: Recording Environment for Annotated Pedestrians.

B. Qualitative Annotations

In [1], several different annotation types for soft biometrics were presented, while the aim of each annotation type was to increase the accuracy of overall recognition or retrieval process [31]. Like earlier research experiments, qualitative types of annotation known as categorical, and comparative are the most widely used semantics in soft biometrics [26]. The annotation methods like expert opinion are used to generate annotation values for soft biometric features. In our case, the expert was highly trained on similar annotation tasks [27] and performed categorical and comparative annotations for approximately 75 soft different soft biometrics from four different modalities of the human body. The qualitative types of annotation were also used extensively in different research experiments on large size image datasets. Like, a method known as superfine attributes was proposed to annotate several biometrics, e.g., gender, age and ethnicity using PETA dataset [32]. The crowd sourcing method was also used in some experiments, however, only a few soft biometrics were annotated because of large size of dataset [3] - the annotation type was categorical and a five scale criteria was defined. Although, crowd sourcing [33] is a successful approach for annotating the datasets, however, the feasibility of this approach is limited while annotating very large datasets. That is why, expert opinion was adopted as the best method for dataset annotation whether using categorical or comparative types of annotations. The expert opinion [34] method produced significant success in several different research experiments like in [9], to annotate the dataset using soft biometric features. In all these experiments, there were several hundred distinct individual images annotated using more than 50 soft biometrics. In another experiment [21], more than 1700 distinct individual images were annotated using 30 different soft biometrics from face, body and clothing etc. The expert opinion seems more feasible method in real-world scenario as compared to crowd sourcing for providing semantic description of a person's image using qualitative or quantitative [35] types of annotation. Usually, we have a very large number of distinct images in a single dataset with a large number of soft biometrics to annotate, however, there are very few recognition scenarios like [31], where a large LFW dataset was annotated only for 6 soft biometrics.

In our earlier research on soft biometrics [1], a bag of more than 150 soft biometrics is presented from three different modalities of the human body. Usually, qualitative methods for annotation like categorical or comparative were used for annotation, and they demonstrated significant success during several recognition or retrieval experiments. Usually, the image information is not sufficient enough during training of the recognition or retrieval model; however, the annotated soft biometric information stored with features directly extracted through images resulted in the form of an improved accuracy of the models in the past [36]. That is why, the application of annotated soft biometrics and their extended set of possible values produced better performance in both constrained and un-constrained environments. Based upon different research experiments, the qualitative annotation types yet presented high feasibility during recognition or retrieval [37].

C. Selecting Soft Biometrics

In order to annotate soft biometrics from all three modalities of an individual using qualitative annotation type, we have a large collection of more than 150 soft biometrics in [1]. Those features were annotated using different labels, where the objective was to select an appropriate number of soft biometrics from each modality and to define a set of possible labels for each soft biometrics [38]. Initially, there was redundancy found in annotations of different soft biometrics used in different research experiments. In our work, we addresses two main challenges, i) selecting an appropriate number of soft biometrics from each modality, and ii) defining a generalize set of annotations for each individual soft biometrics. The aim of building such collection was to support general recognition or retrieval process [39].

Fig. 4: An annotated image from the dataset using soft biometrics from global, extended facial, body including limbs and clothing with attachment modalities.

III. Dataset Annotation Process

In our earlier discussion, qualitative annotation [40] type was identified as a key success factors for recognition or retrieval and we planned to annotate our new Pedestrian dataset with recognition at distance. To achieve this goal, we annotated images of total 50 distinct individuals recorded in a corridor walking towards the camera [41]. For this purpose, we selected image of each individual recorded at 4 meter distance from the camera [30]. On the other hand, one of the major task in annotation was to select an appropriate number of soft biometrics from each modality of the human body [42]. In our work, we selected about 46 categorical soft biometrics from three different modalities of the human body, while 26 were selected for comparative annotation. For categorical annotation, there were 16 soft biometrics from extended facial modality, while eight and 22 from body and clothing modality respectively. Similarly, for comparative annotation, 12 soft biometrics were selected from extended facial modality, while eight from body and six from clothing.

A. Categorical Annotations

In terms of categorical annotation, selection of soft biometrics from three different modalities of the human body was a

Figure	Shoulder Width	Hips Width	Muscle Build	Shoulder Shape	Arm Length	Leg Length	Chest
Thin	Narrow	Narrow	Lean	Squared	Long	Long	Slim
Normal	Average	Average	Normal	Normal	Normal	Normal	Average
Thick	Wide	Wide	Build	Rounded	Short	Short	Large

Face Type	Hair Color	Forehead	Skin Exposure	Eyes	Skin Color	Nose Width	Chin Width	Moustache	Neck Thickness	Nose Shape	Nose Length	Chin and Jaw	Mouth Length	Beard	Neck Length
Bony	Black	Small	Low	Small	White	Small	Small	Yes	Thin	Flatter	Short	Angular	Average	Yes	Short
Normal	Brown	Large	Medium	Large	Tanned	Average	Large	No	Average	Same	Average	Same	Large	No	Medium
Fleshier	Blond		High	Round	Oriental	Large			Thick	Protruding	Large	Rounder	Small		Long
	Gray				Black										

Head Coverage	Neckline Size	Neckline Shape	Color Scheme Overall Clothing	Upper Body Clothing: Category	Upper Body Clothing: Brightness	Upper Body Clothing: Scheme	Upper Body Clothing: Dominant Color	Upper Body Clothing: Pattern	Sleeve Length	Gloves	Attached Object	Presence of Belt
Yes	Large	Shirt Collar	Warm	Hoodies	Light	Warm	Multiple	Complex	Long	Yes	Yes	Yes
No	Small	Round	Cool	Coat	Dark	Cool	Single	None	Short	No	No	No
	Medium	Strapless	Neutral	Sweater	Average	Neutral	Dual	Simple	Medium			
		V-Shap		Blouse								
				Shirt								
				Jumper								
				T-Shirt								

Lower Body Clothing: Category	Lower Body Clothing: Brightness	Lower Body Clothing: Scheme	Lower Body Clothing: Dominant Color	Lower Body Clothing: Pattern	Contrast	Footwear Toed	Heel Level	Shoes
Dress	Light	Warm	Multiple	Complex	High	Close	Very High	Shoe
Trouser	Dark	Cool	Single	None	Low	Open	High	Trainer
Skirt	Average	Neutral	Dual	Simple	Medium		Low	Boot
							Medium	Heels
								Flip FLop

Blue	Body
Yellow	Face
Peach	Clothing

Fig. 5: Selected soft biometrics from extended facial, body including limbs and clothing with attachments modalities for categorical annotations.

Face Coverage	Upper Body Clothing: Brightness	Upper Body Clothing: Scheme	Sleeve Length	Lower Body Clothing: Brightness	Lower Body Clothing: Scheme
Yes	Darker	Cooler	Shorter	Darker	Cooler
No	Similar	Warmer	Longer	Similar	Warmer
	Lighter			Lighter	

Figure	Legs Thickness	Muscle Build	Weight	Shoulder Shape	Chest Size	Arm Length	Leg Length
Thinner	Thinner	Leaner	Fatter	Squared	Bigger	Longer	Shorter
Same	Same	Same	Same	Average	Same	Same	Same
Thikcer	Thicker	Mascular	Thinner	Rounded	Smaller	Shorter	Longer

Face Type	Hair Color	Skin Exposure	E-Brows	Eyes	Ears	Nose Width	Nose Length	Mouth Width	Lips Thickness	Neck Thickness	Neck Length
Smaller	Lighter	Lower		Smaller	Samller	Narrower	Shorter	Narrower	Thinner	Thinner	Shorter
Larger	Same	Same	Thinner	Larger	Larger	Same	Larger	Wider	Same	Same	Same
Same	Darkar	Higher	Bushier	Same	Same	Wider	Same	Same	Thicker	Thicker	Longer

Blue	Face
Peach	Body
Green	Clothing

Fig. 6: Selected soft biometrics from extended facial, body including limbs and clothing with attachments modalities for comparative annotations.

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2023.3234494

bit challenging task [43]. It requires lot of experience to select only those soft biometrics which are highly relevant and supportive to each other for the recognition or retrieval task. An expert opinion method was applied based on the analysis performed in [1], including frequency of occurrence for each soft biometrics in multiple different recognition or retrieval tasks. There were 16 facial soft biometrics selected using expert opinion method and a set of possible categorical annotations was developed as shown in Figure 5. This process incorporates frequency of annotation from earlier research in a particular scenario and by using this factor, a minimal set of annotations was developed for each individual soft biometrics. The same experiment was carried out for 8 soft biometrics from body and 22 from clothing modality. Finally, the most complicated task of dataset annotation was performed on Annotated Pedestrians. More than 70 soft biometrics from face, body including limbs and clothing modality were annotated on 4-meter distance images of the individuals, including three from the global modality. One of the key parameters in selection of those 75 soft biometrics was their frequency in more than one research experiments along with other parameters like permeance and stability, which reflects their significance towards recognition or retrieval tasks [1]. To accomplish annotation task, an expert with experience spanning over a decade on similar annotation tasks was selected and the dataset annotation was performed using soft biometrics from three different modalities [44]. Over the years, expert opinion has shown more feasibility and resulted in more accurate annotations on such types of tasks as compared to other methods like crowdsourcing through image visualization. By using expert opinion method, a single label was generated for each soft biometric using image of an individual and it is verified by a different expert. To present a comprehensive view of the whole annotated dataset, a small set of annotations from the dataset for each modality are presented, like annotations performed for facial soft biometrics starting from person no. 1 to 10 are presented in Table I, while person no. 21 to 30 and person no. 41 to 50 were included as sample from the annotated dataset as shown in Table II and III, respectively. The selection of soft biometrics from each modality, set of annotations for each soft biometrics, and categorical annotations for 50 distinct individuals from Annotated Pedestrians dataset are highly relevant and supportive towards recognition in short-term tracking.

B. Comparative Annotations

Usually, feature based retrieval is one of the key task under the umbrella of security [45] and soft biometrics are significant set of features to accomplish this task. In terms of retrieval, a comparison using the features is very common approach followed in different research experiments. In order to support the retrieval process using visual features of an individual, we also performed comparative annotation of our Annotated Pedestrians dataset. Using collective expert opinion based method, there were 26 soft biometrics selected for comparative annotation of visual soft biometrics as shown in Figure 6. This number is less than the number of categorical soft biometrics selected in earlier annotation process [46]. There are two main reasons behind this e.g., i) increasing the number of soft biometrics always increase the number of comparisons during retrieval process and results in the form of increased algorithm complexity, and ii) several soft biometrics are not appropriate for comparison like categorical. In our experiment, 6 soft biometrics were selected for comparative annotations [47] from extended facial modality, while 8 from the body modality and 12 from the clothing modality. These soft biometrics are highly relevant and supportive for comparative annotation and again 4 meter images of all the 50 distinct individuals were used by the expert.

In this paper, we included one of major soft biometrics called Face Type from extended facial modality rather, total 6 used for comparative annotation and presented a comparison of Face Type by selecting images for person no. 1 to 10 and 41 to 50 in Table IV. In other words, person no. 1 was selected and compared with person number 41 to 50 using Face Type comparative soft biometrics and one out of three possible annotations e.g., shorter, similar or larger, was selected. The same experiments were carried out using comparative soft biometrics called Figure from body and overall clothing color scheme from clothing modality [21]. For examples, using body modality, Figure was used to compare person no. 11 with person number 31 to 40 as shown in Table V. Similarly, using overall clothing color scheme, the person no. 1 was compared with person number 21 to 30 as shown in Table VI. These comparative annotation experiments were performed for images of all 50 individuals, however, only a limited set is included for the purpose of whole dataset representation. Overall, the comparative soft biometrics selected from all three modalities of the human body are highly relevant and supportive towards retrieval task from the dataset.

IV. Key Reflections on Annotation Process

One of the major contributions made in this research is the development of a benchmark dataset for the pedestrians recognition or retrieval at distance and later, building an appropriate set of soft biometrics from each modality to perform qualitative annotations. This is one of the unique datasets in this domain, however, there are several challenges exist to develop more comprehensive annotations for the dataset like;

- The first phase of Annotated Pedestrians dataset is recorded in early 2021 including 50 distinct individuals at four different distances from the camera [14], however, the second phase is planned to be recorded in few months time with a gap of approximately two years. By this Annotated Pedestrians will be a novel dataset of its kind in which same individuals will be recorded after gap of two years. It will also incorporate different clothing [30] for the same individuals with change in appearance, after pandemic.
- The Annotated Pedestrians dataset contains a significant number of soft biometrics from three modalities of the human body using qualitative annotations, and they are more appropriate for short-term tracking [17] and retrieval [21] from the dataset respectively. By keeping in mind the usefulness of Bertillon system [35] and with

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2023.3234494

TABLE I: Categorical annotations for extended facial modality.

Attribute/Person	1	2	3	4	5	6	7	8	9	10
Face Type [48]	Fleshier	Normal	Bony	Bony	Normal	Bony	Fleshier	Normal	Normal	Bony
Hair Color [49]	Black	Black	Brown	Blond	Brown	Gray	Brown	Black	Black	Blond
Forehead [48]	Small	Small	Large	Large	Large	Small	Large	Small	Small	Large
Skin Exposure [50]	Medium	Low	Medium	Medium	Medium	Medium	Low	Low	Medium	Medium
Eyes [23]	Small	Small	Large	Small	Large	Small	Round	Small	Small	Small
Skin Color [47]	Tanned	Oriental	White	White	White	Tanned	White	Oriental	Tanned	White
Nose Width [9]	Average	Small	Average	Average	Small	Large	Small	Average	Average	Small
Chin Width [9]	Small	Small	Small	Small	Small	Small	Large	Small	Small	Small
Moustache [31]	No	Yes	No	No	No	No	No	No	No	No
Neck Thickness [51]	Thick	Thin	Thin	Thin	Average	Thin	Thick	Thin	Average	Average
Nose Shape [52]	Protruding	Same	Flatter	Protruding	Same	Protruding	Same	Same	Same	Flatter
Nose Length [52]	Average	Average	Short	Average	Average	Average	Large	Average	Average	Short
Chin and Jaw [9]	Rounder	Same	Angular	Angular	Rounder	Same	Same	Same	Same	Same
Mouth Length [23]	Average	Average	Average	Large	Average	Small	Average	Average	Average	Small
Beard [31]	No	No	No	No	No	No	No	No	No	No
Neck Length [44]	Medium	Short	Medium	Long	Medium	Long	Short	Short	Medium	Long

TABLE II: Categorical annotations for body including limbs modality

Attribute/Person	21	22	23	24	25	26	27	28	29	30
Figure [47]	Thick	Normal	Thin	Thick	Normal	Normal	Thin	Thin	Normal	Normal
Shoulder Width [49]	Wide	Average	Average	Wide	Average	Average	Narrow	Average	Wide	Wide
Hips Width [49]	Average	Average	Narrow	Wide	Average	Average	Narrow	Narrow	Wide	Average
Muscle Build [21]	Normal	Build	Normal	Normal	Normal	Build	Lean	Normal	Build	Normal
Shoulder Shape [47]	Rounded	Squared	Normal	Rounded	Normal	Squared	Normal	Normal	Squared	Squared
Arm Length [17]	Long	Normal	Normal	Normal	Normal	Normal	Normal	Normal	Normal	Long
Leg Length [21]	Long	Normal	Long	Normal	Normal	Normal	Long	Normal	Short	Long
Chest [44]	Large	Large	Average	Large	Average	Average	Slim	Slim	Large	Large

TABLE III: Categorical annotations for clothing with attachments modality

Attribute/Person	41	42	43	44	45	46	47	48	49	50
Head Coverage [40]	No	Yes	No	No	No	No	Yes	No	No	No
Neckline Size [40]	Medium	Small	Medium	Medium	Small	Medium	Medium	Small	Small	Medium
Neckline Shape [40]	Round	Shirt Collar	Shirt Collar	Round	Round	Round	Strapless	Shirt Collar	Shirt Collar	Shirt Collar
Overall Clothing										
Color Scheme [50]	Neutral	Cool	Cool	Cool	Cool	Cool	Cool	Cool	Cool	Cool
Upper Body Clothing [17]										
Category	Shirt	Coat	Coat	Hoodies	Sweater	Shirt	Coat	Coat	Coat	Jumper
Brightness	Average	Dark	Dark	Dark	Average	Light	Average	Dark	Dark	Average
Color Scheme	Neutral	Cool	Cool	Cool	Cool	Neutral	Cool	Cool	Cool	Cool
Dominant Color	Dual	Multiple	Single	Dual	Multiple	Single	Multiple	Single	Dual	Dual
Pattern	Simple	Simple	Simple	Simple	Complex	Simple	Complex	Simple	Simple	Simple
Sleeve Length [40]	Long	Long	Long	Long	Long	Medium	Long	Long	Long	Long
Gloves [40]	No	No	No	No	No	No	No	No	No	No
Attached Object [40]	No	Yes	No	No	No	No	Yes	No	Yes	No
Presence of Belt [17]	Yes	No	No	Yes	No	No	No	No	No	No
Lower Body Clothing [17]										
Category	Trouser	Trouser	Dress	Trouser	Dress	Dress	Trouser	Trouser	Dress	Dress
Brightness	Dark	Dark	Dark	Average	Average	Dark	Dark	Dark	Average	Dark
Color Scheme	Cool	Cool	Cool	Neutral	Neutral	Neutral	Cool	Cool	Neutral	Cool
Dominant Color	Single	Single	Single	Single	Single	Single	Single	Single	Single	Single
Pattern	Simple	Simple	Simple	Simple	Simple	Simple	Simple	Simple	Simple	Simple
Contrast [31]	Medium	Medium	Medium	Low	Medium	High	Medium	Low	Medium	Medium
Footwear Toed [23]	Close	Close	Close	Close	Close	Close	Close	Close	Close	Close
Heel Level [53]	Low	Low	Low	Medium	Low	Medium	Low	Low	Medium	Low
Shoes [53]	Boot	Boot	Boot	Boot	Boot	Boot	Boot	Boot	Boot	Boot

the availability of highly accurate human body landmark localization techniques like OpenPose [54], AlphaPose [55] and OpenPifPaf [56], we initially plan to extract key facial and body points and then to determine quantitative values for different soft biometrics of the human body. Those quantitative soft biometrics will be estimated directly from the images of the pedestrians and facilitate more accurate recognition in open environments.

- Though clothing and any material attached to the human body is a set of soft biometrics under the temporary modality [1] for recognition or retrieval, however, it has usefulness in short-term tracking and an auxiliary component to predict cultural or tradition. The cultural context [57] is not directly capable of recognition, however, it may become an auxiliary component for any soft biometric recognition system. For instance, for people arriving at an entry point, a cumulative clothing soft biometrics may be determined using their cultural context. One of the key future action point is to annotate the dataset for determining the clothing soft biometrics using techniques like Deep Fashion [58].

- In this work, we used expert opinion based annotation method to annotate the soft biometrics of the human body using qualitative annotation types [1]. They are all manual methods for annotation. In our earlier discussion, we plan to develop automated annotation techniques using landmarks localization tools and tools like deep fashion [58]. By this we have a very good opportunity to compare the outcome of automated annotation techniques with manual annotations for soft biometrics. In simple words, this comparative analysis will not only make the annotation task automated but also identify the areas of further improvement.
- Both traditional and soft biometrics are a very large set of features from different modalities of the human body [14], however, to use limited number of features from each modality results in better performance of overall recognition or retrieval system. This termed as feature selection in machine learning, however, in our approach feature selection should be application domain specific and only features related and supportive to each other should be used for recognition or retrieval. Like in [59], non-linear regression analysis was performed on soft biometrics from clothing modality to develop a set of highly relevant and supportive soft biometrics from clothing modality. In our future course of action, it is critical to develop multiple sets of related and supportive soft biometrics from different modalities of the human body in terms of different generalize application domains.
- In [1], an analysis was performed to determine the significance of two very important factors associated with individual soft biometrics called permanence score [60] and discrimination power [9]. The permanence score is measured across the images of the same individual over the time and distance, while discrimination power is calculated using the images of the different individuals. In our future work, we plan to define a weight for each individual soft biometrics from different modalities of the human body based on permanence score and discrimination power.

V. Conclusion

This paper introduces a novel annotated dataset for pedestrian recognition or retrieval using soft biometrics called Annotated Pedestrians, The dataset was recorded for 50 distinct individuals at four different distances from the camera. For each individual, there were 300 frames recorded at the rate of 30 FPS. To verify the automated extraction of the images at four different distance from the camera, markers were placed on the surface at 4 m, 6 m, 8 m and 10 m distance from the camera. The dataset comprised of both genders, age range from 15 to 60 years, and five different ethnic groups. The dataset was annotated using qualitative method of annotation by an expert in the field and 75 soft biometrics from four different modalities of the human body were annotated using categorical or comparative annotation types. To the best of our knowledge, Annotated Pedestrians is a unique dataset of its kind, which is recorded and annotated using multiple different modalities of the human body, including approximately 75 soft biometrics. Despite traditional biometrics, soft biometrics are more promising features of the human body during recognition in any constrained or un-constrained environment. Soft biometrics offer seamless mode for recognition or retrieval and a stepping-stone toward the development and evaluation of any standalone soft biometrics-based person verification system. On the other hand, the extended applications of soft biometrics over traditional biometrics expand from content-based image retrieval to continuous authentication in virtual environments, and fashion industry besides surveillance.

References

[1] B. Hassan, E. Izquierdo, and T. Piatrik, "Soft biometrics: a survey," *Multimedia Tools and Applications*, pp. 1–44, 2021.

[2] E. Learned-Miller, G. B. Huang, A. RoyChowdhury, H. Li, and G. Hua, "Labeled faces in the wild: A survey," in *Advances in face detection and facial image analysis*. Springer, 2016, pp. 189–248.

[3] Y. Deng, P. Luo, C. C. Loy, and X. Tang, "Pedestrian attribute recognition at far distance," in *Proceedings of the 22nd ACM international conference on Multimedia*. ACM, 2014, pp. 789–792.

[4] G. B. Huang and E. Learned-Miller, "Labeled faces in the wild: Updates and new reporting procedures," *Dept. Comput. Sci., Univ. Massachusetts Amherst, Amherst, MA, USA, Tech. Rep*, pp. 14–003, 2014.

[5] K. Ricanek and T. Tesafaye, "Morph: A longitudinal image database of normal adult age-progression," in *7th International Conference on Automatic Face and Gesture Recognition (FGR06)*. IEEE, 2006, pp. 341–345.

[6] R. Vera-Rodriguez, P. Tome, J. Fierrez, N. Expósito, and F. J. Vega, "Analysis of the variability of facial landmarks in a forensic scenario," in *2013 International Workshop on Biometrics and Forensics (IWBF)*. IEEE, 2013, pp. 1–4.

[7] Y. Pang, T. Wang, R. M. Anwer, F. S. Khan, and L. Shao, "Efficient featurized image pyramid network for single shot detector," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 7336–7344.

[8] Y. Pang, J. Cao, and X. Li, "Learning sampling distributions for efficient object detection," *IEEE transactions on cybernetics*, vol. 47, no. 1, pp. 117–129, 2016.

[9] P. Tome, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia, "Facial soft biometric features for forensic face recognition," *Forensic science international*, vol. 257, pp. 271–284, 2015.

[10] A. Rattani, R. Derakhshani, and A. Ross, *Selfie Biometrics: Advances and Challenges*. Springer Nature, 2019.

[11] Y. Zhang, L. Liu, C. Li *et al.*, "Quantifying facial age by posterior of age comparisons," *arXiv preprint arXiv:1708.09687*, 2017.

[12] J. D. Shutler, M. G. Grant, M. S. Nixon, and J. N. Carter, "On a large sequence-based human gait database," in *Applications and Science in Soft Computing*. Springer, 2004, pp. 339–346.

[13] P. Misra, N. Jain, and A. S. Mandal, "Search space reduction for person recognition using soft biometrics," in *2017 Conference on Information and Communication Technology (CICT)*. IEEE, 2017, pp. 1–5.

[14] B. Hassan and E. Izquierdo, "Onedetect: A federated learning architecture for global soft biometrics prediction," in *2022 International Conference on Intelligent Systems and Computer Vision (ISCV)*. IEEE, 2022, pp. 1–8.

[15] B. H. Guo, M. S. Nixon, and J. N. Carter, "Soft biometric fusion for subject recognition at a distance," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 1, no. 4, pp. 292–301, 2019.

[16] L. Wu, D. Liu, W. Zhang, D. Chen, Z. Ge, F. Boussaid, M. Bennamoun, and J. Shen, "Pseudo-pair based self-similarity learning for unsupervised person re-identification," *IEEE Transactions on Image Processing*, vol. 31, pp. 4803–4816, 2022.

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2023.3234494

TABLE IV: Comparative annotations for extended facial modality using soft biometrics (Face Type).

Person No.	41	42	43	44	45	46	47	48	49	50
1	Similar	Similar	Larger	Larger	Similar	Larger	Larger	Larger	Larger	Shorter
2	Shorter	Shorter	Similar	Similar	Shorter	Similar	Shorter	Similar	Similar	Shorter
3	Shorter	Similar	Similar	Similar	Shorter	Similar	Larger	Similar	Similar	Shorter
4	Shorter	Shorter	Shorter	Similar	Shorter	Similar	Shorter	Similar	Shorter	Shorter
5	Shorter	Similar	Shorter	Similar	Shorter	Similar	Similar	Similar	Similar	Shorter
6	Shorter	Shorter	Shorter	Similar	Shorter	Larger	Larger	Similar	Similar	Shorter
7	Similar	Larger	Larger	Larger	Similar	Larger	Larger	Larger	Larger	Similar
8	Shorter	Similar	Similar	Larger	Shorter	Similar	Similar	Larger	Similar	Shorter
9	Shorter	Similar	Similar	Larger	Shorter	Larger	Similar	Larger	Similar	Shorter
10	Shorter	Shorter	Shorter	Similar	Shorter	Larger	Similar	Similar	Larger	Shorter

TABLE V: Comparative annotations for body including limbs modality using soft biometrics (Figure).

Person No.	31	32	33	34	35	36	37	38	39	40
11	Thicker	Thicker	Thicker	Thicker	Same	Thicker	Thicker	Thicker	Thicker	Thicker
12	Thinner	Same	Thinner	Same	Thinner	Same	Same	Same	Thinner	Same
13	Same	Thicker	Same	Thicker	Same	Thicker	Thicker	Thicker	Thicker	Thicker
14	Thicker	Same	Same	Same	Same	Thicker	Same	Thicker	Same	Same
15	Same	Same	Thinner	Same	Same	Thicker	Thicker	Thicker	Same	Same
16	Same	Thicker	Thicker	Thicker	Same	Thicker	Thicker	Thicker	Thicker	Thicker
17	Thinner	Same	Thinner	Same	Thicker	Same	Thinner	Same	Same	Same
18	Same	Same	Thinner	Thinner	Thinner	Same	Thinner	Same	Thicker	Thinner
19	Same	Thicker	Thinner	Same	Same	Thicker	Same	Thicker	Same	Same
20	Thicker	Thicker	Thicker		Thicker	Thicker	Thicker	Thicker	Thicker	Thicker

TABLE VI: Comparative annotations for clothing with attachments modality using soft biometrics (Clothing Color Scheme).

Person No.	21	22	23	24	25	26	27	28	29	30
1	Cooler	similler	Cooler	Cooler	Cooler	Cooler	Cooler	Cooler	Cooler	similler
2	similler	similler	Cooler	Cooler	similler	Cooler	Cooler	similler	Cooler	similler
3	similler	Warmer	Cooler	Cooler	similler	Cooler	Cooler	similler	similler	similler
4	similler	similler	similler	Cooler	Cooler	Cooler	Cooler	Cooler	similler	similler
5	Cooler	similler	Cooler	Cooler	Cooler	Cooler	Cooler	Cooler	Cooler	similler
46	similler	Cooler	Cooler	Cooler	Cooler	Cooler	Cooler	Cooler	Cooler	similler
47	Warmer	Warmer	similler	similler	Warmer	Warmer	Warmer	Warmer	Warmer	Warmer
48	Warmer	Warmer	Warmer	similler	Warmer	similler	similler	Warmer	Warmer	Warmer
49	Warmer	similler	similler	similler	similler	similler	similler	similler	Warmer	Warmer
50	Warmer	similler	similler	Cooler	similler	Cooler	Cooler	similler	similler	similler

[17] B. H. Guo, M. S. Nixon, and J. N. Carter, "Fusion analysis of soft biometrics for recognition at a distance," in *2018 IEEE 4th International Conference on Identity, Security, and Behavior Analysis (ISBA)*. IEEE, 2018, pp. 1–8.

[18] H.-Y. Kwon and M.-K. Lee, "Comments on "passbio: Privacy-preserving user-centric biometric authentication"," *IEEE Transactions on Information Forensics and Security*, vol. 17, pp. 2816–2817, 2022.

[19] S. Lu, Z. Gao, Q. Xu, C. Jiang, A. Zhang, and X. Wang, "Class-imbalance privacy-preserving federated learning for decentralized fault diagnosis with biometric authentication," *IEEE Transactions on Industrial Informatics*, pp. 1–11, 2022.

[20] J. Benesty, J. Chen, Y. Huang, and I. Cohen, "Pearson correlation coefficient," in *Noise reduction in speech processing*. Springer, 2009, pp. 1–4.

[21] P. Tome, J. Fierrez, R. Vera-Rodriguez, and M. S. Nixon, "Soft biometrics and their application in person recognition at a distance," *IEEE Transactions on information forensics and security*, vol. 9, no. 3, pp. 464–475, 2014.

[22] E. Gonzalez-Sosa, A. Dantcheva, R. Vera-Rodriguez, J.-L. Dugelay, F. Brémond, and J. Fierrez, "Image-based gender estimation from body and face across distances," in *2016 23rd International Conference on Pattern Recognition (ICPR)*. IEEE, 2016, pp. 3061–3066.

[23] N. Y. Almudhahka, M. S. Nixon, and J. S. Hare, "Comparative face soft biometrics for human identification," in *Surveillance in Action*. Springer, 2018, pp. 25–50.

[24] D. Li, Z. Zhang, X. Chen, and K. Huang, "A richly annotated pedestrian dataset for person retrieval in real surveillance scenarios," *IEEE Transactions on Image Processing*, vol. 28, no. 4, pp. 1575–1590, 2019.

[25] D. Benini, "Biometric identification and verification," Oct. 8 2013, uS Patent 8,553,947.

[26] T. A. Lucas and M. Henneberg, "Comparing the face to the body, which is better for identification?" in *AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY*, vol. 156. WILEY-BLACKWELL 111 RIVER ST, HOBOKEN 07030-5774, NJ USA, 2015, pp. 207–208.

[27] Z. Zhou, G. H. T. Ong, and E. K. Teoh, "Soft-biometric detection based on supervised learning," in *2014 13th International Conference on Control Automation Robotics & Vision (ICARCV)*. IEEE, 2014, pp. 234–238.

[28] S. Ghosh, M. Vatsa, and R. Singh, "Suprear-net: Supervised resolution enhancement and recognition network," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 4, no. 2, pp. 185–196, 2022.

[29] C. Song, Y. Huang, W. Wang, and L. Wang, "Casia-e: A large comprehensive dataset for gait recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–16, 2022.

[30] B. Hassan and E. Izquierdo, "Person recognition across multi-session and multi-exemplar images using ensemble of classifiers," in *Thirteenth International Conference on Digital Image Processing (ICDIP 2021)*, vol. 11878. International Society for Optics and Photonics, 2021, p. 1187807.

[31] E. Gonzalez-Sosa, J. Fierrez, R. Vera-Rodriguez, and F. Alonso-Fernandez, "Facial soft biometrics for recognition in the wild: Recent works, annotation, and cots evaluation," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 8, pp. 2001–2014, 2018.

[32] D. Martinho-Corbishley, M. S. Nixon, and J. N. Carter, "Super-fine attributes with crowd prototyping," *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, no. 6, pp. 1486–1500, 2018.

[33] A. Jafarzadeh, M. L. Antequera, P. Gargallo, Y. Kuang, C. Toft, F. Kahl, and T. Sattler, "Crowddriven: A new challenging dataset for outdoor

visual localization," in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021, pp. 9825–9835.
[34] F. Sakketou, A. Lahnala, L. Vogel, and L. Flek, "Investigating user radicalization: A novel dataset for identifying fine-grained temporal shifts in opinion," *arXiv preprint arXiv:2204.10190*, 2022.
[35] R. B. Fosdick, "Passing of the bertillon system of identification," *J. Am. Inst. Crim. L. & Criminology*, vol. 6, p. 363, 1915.
[36] B. Hassan and E. Izquierdo, "Apparelnet: Person verification encompassing auxiliary attachments variation," in *2021 IEEE 23rd International Workshop on Multimedia Signal Processing (MMSP)*, 2021, pp. 1–6.
[37] M. Shoaib Farooq, B. Hassan, M. Naseer, A. Abid, Y. D. Khan, N. S. Khan, M. Usman Akram, and S. Ullah, "Studio applications and software development kits for microsoft kinect: A survey," 2014.
[38] S. Zhang, Y. Xie, J. Wan, H. Xia, S. Z. Li, and G. Guo, "Widerperson: A diverse dataset for dense pedestrian detection in the wild," *IEEE Transactions on Multimedia*, vol. 22, no. 2, pp. 380–393, 2020.
[39] D. Osorio-Roig, C. Rathgeb, P. Drozdowski, P. Terhörst, V. Štruc, and C. Busch, "An attack on facial soft-biometric privacy enhancement," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 4, no. 2, pp. 263–275, 2022.
[40] E. S. Jaha and M. S. Nixon, "Soft biometrics for subject identification using clothing attributes," in *IEEE International Joint Conference on Biometrics*. IEEE, 2014, pp. 1–6.
[41] B. Hassan, U. Akram, M. Naseer, F. Ali, S. Akhter, M. Ajmal *et al.*, "A publicly available rgb-d data set of muslim prayer postures recorded using microsoft kinect for windows," 2014.
[42] A. Dantcheva, C. Velardo, A. D'angelo, and J.-L. Dugelay, "Bag of soft biometrics for person identification," *Multimedia Tools and Applications*, vol. 51, no. 2, pp. 739–777, 2011.
[43] W. Yang, P. Luo, and L. Lin, "Clothing co-parsing by joint image segmentation and labeling," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 3182–3189.
[44] A. Nambiar, A. Bernardino, and J. Nascimento, "Shape context for soft biometrics in person re-identification and database retrieval," *Pattern Recognition Letters*, vol. 68, pp. 297–305, 2015.
[45] C. BenAbdelkader and Y. Yacoob, "Statistical body height estimation from a single image," in *2008 8th IEEE International Conference on Automatic Face & Gesture Recognition*. IEEE, 2008, pp. 1–7.
[46] M. R. Sayed, T. Sim, J.-H. Lim, and K. T. Ma, "Which body is mine?" in *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2019, pp. 829–838.
[47] D. Martinho-Corbishley, M. S. Nixon, and J. N. Carter, "Analysing comparative soft biometrics from crowdsourced annotations," *IET Biometrics*, vol. 5, no. 4, pp. 276–283, 2016.
[48] D. A. Reid and M. S. Nixon, "Human identification using facial comparative descriptions," in *2013 International Conference on Biometrics (ICB)*. IEEE, 2013, pp. 1–7.
[49] R. Vera-Rodriguez, P. Marin-Belinchon, E. Gonzalez-Sosa, P. Tome, and J. Ortega-Garcia, "Exploring automatic extraction of body-based soft biometrics," in *2017 International Carnahan Conference on Security Technology (ICCST)*. IEEE, 2017, pp. 1–6.
[50] E. S. Jaha and M. S. Nixon, "From clothing to identity: Manual and automatic soft biometrics," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 10, pp. 2377–2390, 2016.
[51] D. A. Reid, M. S. Nixon, and S. V. Stevenage, "Soft biometrics; human identification using comparative descriptions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 6, pp. 1216–1228, 2013.
[52] N. Almudhahka, M. Nixon, and J. Hare, "Human face identification via comparative soft biometrics," in *2016 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA)*. IEEE, 2016, pp. 1–6.
[53] N. Y. Almudhahka, M. S. Nixon, and J. S. Hare, "Semantic face signatures: Recognizing and retrieving faces by verbal descriptions," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 3, pp. 706–716, 2017.
[54] Z. Cao, G. Hidalgo Martinez, T. Simon, S. Wei, and Y. A. Sheikh, "Openpose: Realtime multi-person 2d pose estimation using part affinity fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.
[55] H.-S. Fang, S. Xie, Y.-W. Tai, and C. Lu, "Rmpe: Regional multi-person pose estimation," in *ICCV*, 2017.
[56] S. Kreiss, L. Bertoni, and A. Alahi, "Pifpaf: Composite fields for human pose estimation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.
[57] M. T. Teye, Y. M. Missah, E. Ahene, and T. Frimpong, "Evaluation of conversational agents: Understanding culture, context and environment in emotion detection," *IEEE Access*, vol. 10, pp. 24 976–24 984, 2022.
[58] Y. Ge, R. Zhang, X. Wang, X. Tang, and P. Luo, "Deepfashion2: A versatile benchmark for detection, pose estimation, segmentation and re-identification of clothing images," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 5337–5345.
[59] B. Hassan and E. Izquierdo, "Rsfs: A soft biometrics based relative support features set for person verification," in *Fourteenth International Conference on Digital Image Processing (ICDIP 2022)*. International Society for Optics and Photonics, 2022.
[60] W. R. Knight, "A computer method for calculating kendall's tau with ungrouped data," *Journal of the American Statistical Association*, vol. 61, no. 314, pp. 436–439, 1966.