https://doi.org/10.1186/s13634-023-01026-x

EURASIP Journal on Advances in Signal Processing

RESEARCH Open Access

A multi-channel soft biometrics framework for seamless border crossings

Bilal Hassan[1], Hafiz Husnain Raza Sherazi[2], Mubashir Ali[3*] and Ali K. Bashir[4]

*Correspondence: m.ali.16@bham.ac.uk

[1] Faculty of Engineering and Environment, Northumbria University UK, London, UK
[2] School of Computing and Engineering, University of West London, London, UK
[3] School of Computer Science, University of Birmingham, Birmingham, UK
[4] Department of Computing and Mathematics, Manchester Metropolitan University, Manchester, UK

Abstract

As the number of passengers at border entry points such as airports and rail stations increases, so does the demand for seamless, secure, and fast biometric technologies for verification purposes. Although fingerprints are currently useful biometric technologies, they are intrusive and slow down the end-to-end verification process, increasing the chances of tampering. Emerging as an alternative technology, soft biometrics have proven successful for non-intrusive and rapid verification. Soft biometrics consists of a large set of features from three different modalities of the human body, including the face, body, and essential & auxiliary attachments. This paper proposes a multi-channel soft biometrics framework that leverages soft biometrics technology over traditional biometrics. The framework encapsulates four distinct components: ApparelNet, which verifies essential and auxiliary attachments; A-Net, which measures anthropometric soft biometrics; OneDetect, which predicts global soft biometrics; and RSFS, which develops a set of highly relevant and supportive soft biometrics for verification. The proposed framework addresses several critical limitations of existing biometrics technologies during the verification process at border entry points, such as intrusive behavior, response time, biometric tampering, and privacy issues. The proposed multi-channel soft biometrics framework has been evaluated using several benchmark datasets in the field, such as Front-view Gait (FVG), Pedestrian Attribute Recognition At Far Distance (PETA), and Multimedia and Vision (MMV) Pedestrian. Using heterogeneous datasets enables the testing of each framework component or channel against numerous constrained and unconstrained scenarios. The outcome of the envisioned multi-channel soft biometrics framework is presented based on distinct outcomes from each channel, but it remains focused on determining a single cumulative verification score for verification at border control. In addition, this multi-channel soft biometrics framework has extended applications in several fields, including crowd surveillance, the fashion industry, and e-learning.

Keywords: Anthropometric soft biometrics, Biometric tampering, Border control, Crowd surveillance, Multi-channel soft biometrics framework

1 Introduction

Due to the increase in threats and terrorism reported in the last two decades, security technologies have become increasingly important [1]. New security technologies are capable of generating a unique identity, replacing the need for manual verification. There are several systems for identity verification and matching around the world, typically developed by immigration authorities to fulfill border verification requirements [2]. While security is a wide-ranging subject with applications in various settings, border control is one of its key applications [3]. Currently, border control and verification processes are rapidly adopting biometric technologies, such as fingerprint and facial feature-based verification [4]. These are intrusive modes of verification at border control that use biometric data produced by the respective country or region using their own mechanisms [5]. However, these technologies require intentional involvement of individuals in the verification process, which can be time-consuming and result in long queues at international airports. One of the main reasons for delays in border control is the heterogeneous documents produced at border verification points, such as passports [6]. These verification documents are issued by different countries worldwide, using various biometric technologies, as shown in Fig. 1. As a result, country-specific verification checks are necessary and must be conducted manually or using the same technology.

After examining the limitations of traditional biometric technologies, we sought to identify the key factors responsible for these limitations: verification speed, security,

Fig. 1 Taxonomy of biometrics technologies for security—image courtesy [7]

and seamless behavior in the context of border control. These limitations also pose significant research challenges in the field of soft biometrics, which encompass the following:

- Quickness
  - Verification speed depends on the chosen mechanism.
  - A non-intrusive mode, such as walking through corridors, can speed up the process in a seamless manner.
- Security
  - Traditional biometrics are secure, but there is a higher probability of intentional tampering.
  - The pandemic has affected people's overall appearance and physical structure, and traditional biometrics may have limited adaptability.
  - Cosmetic changes can also affect verification using traditional biometrics.
- Seamlessness
  - Traditional biometrics are intrusive, and soft biometrics may provide an alternative.
  - Soft biometrics refer to multi-modality human body features.
  - Ultimately, soft biometrics can serve as a foundation for a transition toward soft biometrics-based **SMART Borders**.

Based on the earlier analysis presented, it appears that soft biometrics is a leading alternative solution for achieving fast, secure, and seamless border crossing [5]. The use of soft biometrics enables quick and secure border verification, which can be accomplished through the implementation of walking corridors [8], serving as an alternative to fast-track counters. These walking corridors provide a non-intrusive method for verification, thus making the overall verification process quick. Therefore, soft biometrics seems to be a suitable choice for this purpose.

A multi-channel soft biometrics framework is proposed for this purpose, which combines various soft biometrics technologies. The framework is designed specifically for installation in an airport-like corridor [9]. To evaluate the proposed framework, we used our developed MMV Pedestrian dataset along with several existing datasets. The purpose of developing the MMV Pedestrian dataset was to record in a walking corridor style and create an impression similar to future border crossing points [10]. The proposed multi-channel soft biometrics framework has multiple channels that were evaluated against specific criteria. Keeping in mind the limitations of existing border verification processes, this framework is being proposed with the following set of objectives [4]. First, to analyze the effectiveness of non-intrusive soft biometrics during the border verification process. Second, to work toward the development of a diverse set of soft biometric datasets [11]. Third, to propose a generic framework to predict soft biometrics such as gender, age, and ethnicity, known as global soft biometrics. Fourth, to understand the usefulness of essential and auxiliary clothing during the border verification process. Fifth, to work on human body anthropometrics for recognition. Sixth, to develop a set of supportive and relevant soft biometric features. Finally, to propose, develop and evaluate a multi-channel soft biometrics framework in a corridor-like environment.

The main contribution of this research is the development of a multi-channel soft biometrics framework. The proposed framework will comprise four channels to support the overall verification process at border entry points. Initially, the framework is an experimental system, but the plan is to test it in real-world scenarios by developing a prototype application. The framework operates on images from three different datasets and features from the PETA dataset. The overall aim of this research was to verify people at border control while tracking and retrieving them using CCTV. This paper made the following scientific contributions in the field:

- Development of the MMV Pedestrian dataset by following the guidelines established for evaluating soft biometrics systems at border control.
- Development of a federated learning architecture to predict global soft biometrics, such as gender, age, and ethnicity.
- Implementation and evaluation of a transfer learning architecture for auxiliary attachment-based verification over a long period for the same individual.
- Implementation and evaluation of human body anthropometrics for recognition in a walking corridor.
- Development of a highly relevant and supportive set of soft biometrics to facilitate the recognition process.
- Evaluation of the multi-channel framework using its distinct channels under the single objective of border verification.

Following the discussion, this paper is organized as follows: Section 2 presents the related work on alternative biometric verification technologies. Section 3 covers the discussion on the dataset selection process, dataset annotation, and building a new dataset. Section 4 presents the proposed multi-channel soft biometrics framework, its four distinct channels, and their implementation details. Section 5 validates the proposed framework using each distinct channel and provides an analysis of the cumulative outcome of the framework under a single aim. Finally, Section 6 outlines the contributions made in this research along with highlighting the future research directions.

2 Related work on alternative border verification technologies

In general, biometrics have gained a lot of attention due to their application for efficient and secure verification [12]. Several different types of biometrics are associated with the human body; however, only a few common biometrics like facial scans and fingerprints have been commercially used, either at a border crossing, kiosk [13], or authentication during an online session. Such biometric verification technologies require intentional inclusion of the human throughout the whole process of verification and they are commonly known as intrusive biometrics [14].

Following recent developments in the field of biometrics, the concept of seamless borders is emerging using non-intrusive biometric technologies [5]. They are an absolute alternative to existing or traditional biometrics. Usually, this fast and secure mode of seamless verification is deployed as Smart Borders [15], as shown in Fig. 2. To transform the idea of Smart Borders into reality, soft biometrics, known as non-intrusive biometrics [5], are the fundamental technology. In our work, we thoroughly investigated and applied the concept of soft biometrics toward the development of Smart Borders. Nowadays, soft biometrics are emerging as the core of recognition systems instead of an auxiliary component to traditional biometrics. Moreover, they have applications from recognition to retrieval and tracking using CCTV.

One of the main challenges in the verification process performed at border control, such as air, sea, or land, is to recognize the same individual after a long gap. Usually, the process of recognizing the same individual is known as person re-identification [16]. The majority of the existing person re-identification applications are developed to recognize the same individual in multiple views, like a person reappearing in multiple cameras during the same time interval [17]. It is still a hard problem for computer vision models to recognize people after a gap of a certain time period, especially if the gap increases between two consecutive appearances of the same individual while the change in essential and auxiliary attachment is an additional factor [18].

Usually, the verification process at border control is a complex task. Generally, individuals are registered for recognition once, and there is a chance of immediate or delayed appearance at different border crossing points. The same problem exists in traditional biometrics where face, fingerprints, and retina scans should be recorded again after a specific time period [19]. As we are working on soft biometrics, finding and testing multiple types of features and human body modalities is an initial step in our experimental phase. To accomplish this task, we have already conducted a comprehensive study on soft biometric feature estimation and classification [5]. A comprehensive analysis of soft biometrics reveals their advantages over traditional biometrics, such as non-intrusive nature [20] and a fast method for seamless recognition.

Fig. 2 Smart tunnel—a concept of future borders installed at Dubai airport

Soft biometrics [5, 21] are gaining attention for their applications ranging from border control to continuous authentication during online sessions and for re-identification in unconstrained scenarios [16]. Soft biometrics provide a seamless method for recognition, which is why they are emerging as an essential component of verification and re-identification systems. Although soft biometrics worked as an ancillary component of traditional biometric systems in the past, which includes fingerprints and retina scans, they are now taking the lead [22] over these non-intrusive ways of recognition. To this end, it is very important to further explore and understand the emergence of soft biometrics and their fusion with traditional biometrics, ultimately leading to an independent system [23].

A comprehensive view of soft biometric technology is critical to understanding the field, determining its success stories, and finding open challenges. In our work, we analyzed the way soft biometrics provide a seamless recognition process [24], which is an advantage over traditional feature-based approaches such as retina scans. One of the key benefits of soft biometrics is their non-intrusive nature in the recognition process. Moreover, compared to traditional biometrics, soft biometrics contain a huge set of features from several different modalities of the human body, such as the face, body, including limbs, and material attached [25, 26]. Another set of features known as global soft biometrics exists, including gender, age, and ethnicity, which can be estimated using face, body, or both modalities. Some of the modalities are permanent, while others are temporary. Similar to traditional biometrics, soft biometrics also have meaningful descriptions, known as semantic descriptions [27].

In the recent era, a complete understanding of the discipline of soft biometrics and its application in several domains entails many different aspects. One of the key aspects is to find the dependence of standalone soft biometrics systems on training and test datasets [28], which are specifically designed or developed for this purpose. Initially, human gait datasets were used for the evaluation of traditional biometrics systems. Such datasets were first annotated [29] for soft biometrics systems evaluation. For this purpose, annotation processes and types were defined in line with soft features. The soft biometrics systems were then evaluated using these annotated datasets in several different kinds of experiments. In an earlier work [5], a summary of different datasets used in various soft biometrics experiments is presented. This summary includes the annotation process, type, volume, and number of attributes annotated. Similarly, an analysis of different soft features used in various research experiments is also summarized. To achieve this goal, the largest collection of soft traits with more than 170 features [30] was presented. The annotation type used for each of these soft features is also part of the research.

Developing a standalone soft biometric system for recognition or retrieval is currently one of the biggest challenges due to certain critical factors directly affecting accuracy. The overall accuracy can be highly improved if these factors are well addressed [31]. In earlier research on standalone soft biometric systems, a comparison of benchmark recognition and retrieval systems was performed using a multi-scale criterion. It also provides information about their success rate using global soft biometrics [32–34]. Finally, the evaluation of the development of standalone soft biometrics systems generates a list of open challenges in the field [5].

Biometric verification technologies have been extensively used for efficient and secure verification. However, the emergence of non-intrusive biometric technologies, also known as soft biometrics, has transformed the concept of Smart Borders into reality. Soft biometrics are now the fundamental technology behind recognition systems and have applications from recognition to retrieval and tracking using CCTV. Despite their potential, recognizing the same individual after a long gap remains a challenge in the verification process performed at border control, and person re-identification is still a hard problem for computer vision models, especially if the gap increases between two consecutive appearances of the same individual while the change in essential and auxiliary attachment is an additional factor. Therefore, the development of a standalone soft biometric system for recognition or retrieval is one of the biggest challenges. To improve the overall accuracy, critical factors affecting accuracy must be addressed. The research gap in this area is significant, and a novel multi-channel framework using soft biometrics is proposed to address these challenges. The proposed framework has the potential to improve recognition accuracy in scenarios where the gap between two consecutive appearances of an individual is long and essential, and auxiliary attachment changes are present.

3 Datasets: analysis, manipulation, and development

For real-time recognition or retrieval, the diversification in the dataset is a critical factor [35], while the accuracy of the annotation method and type of annotation is another important factor [36]. To accomplish this, the dataset design or development process can be of two types such as (i) using an existing dataset alone or in combination with other datasets while annotation is already performed on it, or (ii) development of a completely new dataset [37].

This paper investigates and presents several types of datasets that were recorded in a single session and using multiple cameras. These datasets are very useful for the evaluation of multi-channel soft biometrics framework proposed in the paper. Although the list of these datasets is quite long, however, there are several distinguishing ones such as CASIA [35]. One famous version of the CASIA dataset is CASIA-B, which has been recorded from 11 different angles for 124 distinct individuals. Besides this, there are a few more datasets that are comprised of the sufficient gap between multiple sessions of different individuals such as Frontal-View Gait (FVG) [38, 39], ILRW [40], and Motion Re-ID [41]. They are not available for commercial research except Front-View Gait (FVG) dataset which is open for academics and research. This paper plans to evaluate different channels of the proposed multi-channel framework using FVG datasets.

To extend the evaluation scope of the proposed framework in this paper, a new dataset known as MMV Pedestrian was recorded. Several objectives were in mind to record a new dataset, like recording the same individual at changing distances from the camera, including all the modalities such as the face, body including limbs, and clothing. Moreover, recording an individual with a certain time gap of days, months and years was one of the biggest challenges and it was incorporating structural, physical, and appearance changes. The MMV Pedestrians dataset recorded 50 individuals and it will also be used to evaluate the proposed framework.

In an earlier research on soft biometrics [5], it was revealed that there exists a strong direct or inverse relationship among soft biometrics. To test this hypothesis, a large annotated dataset like PETA [42] was annotated for its limited number of attributes. This paper used PETA too for the evaluation of the proposed multi-channel soft biometrics framework. The purpose was to develop a relative and supportive set of soft biometrics.

3.1 Using FVG dataset

Developing a dataset that consists of repeated images or videos of the same individuals with a certain time gap is a challenging task. One of the main reasons for this is the non-availability of the same group of people again and again. However, this is critical for evaluating any border verification application that is developed. To the best of our knowledge, the only dataset that was recorded in the years 2017 and 2018 respectively is FVG for the same 12 individuals [38]. Sample images from the FVG datasets are shown in Fig. 3.

Although the FVG dataset contains images of a total of 124 distinct individuals, only 12 individuals were part of the image acquisition process in consecutive years. The FVG dataset also includes variations in image acquisition from three different angles, changes in three different walking speeds, changes in appearance such as clothing, and carrying of any auxiliary items [39]. Additionally, the dataset represents all modalities of the human body, including the face, body including limbs, and essential and auxiliary items of clothing. The FVG dataset was originally developed for gait recognition. For our proposed multi-channel soft biometrics framework, we selected a subset of the FVG dataset to evaluate the performance of the transfer learning-based ApparelNet [43] architecture—this is one of the most influential channels in our proposed framework and was presented as part of one of our earlier research studies [44]. We used approximately 450 images per individual each year, with a total of 5563 images selected from 2017 and 4973 images from 2018.

3.2 Using PETA dataset

In previous research on soft biometrics, one of the biggest challenges identified was the use of a highly relevant and supportive set of soft biometrics for recognition and retrieval tasks [5]. Soft biometrics can consist of multiple modalities of the human body, and a large dataset with high diversity and annotations is always required for their development. The PETA dataset, which stands for "Pedestrian Attribute Recognition At Far

(a) Fast walking (b) Changing Accessories

Fig. 3 Images used from FVG [39]

Distance," provides annotations for different human body soft biometrics such as age, gender, and several related to essential and auxiliary clothing [42]. It consists of a total of 19K images for 8705 distinct individuals, with only one soft biometric, gender, being binary, while the remaining are multi-class.

In the multi-channel framework proposed in this paper, the aim is to identify gender and age from the patterns of essential and auxiliary clothing. To accomplish this, the style information was utilized to determine gender and age using the Relative Support Feature Set (RSFS) channel [45].

During the process of collecting annotated data, this research has selected seven different soft biometrics that are related to clothing patterns or styles of the human body, in addition to age and gender. These patterns usually distinguish different age groups and both genders in real-world observation. The image annotations from 5 different datasets of PETA were used, and the following soft biometrics were selected along with their annotations, as shown in Fig. 4. Clothing pattern information is usually significant in determining age and gender, and can successfully be acquired using tools

Fig. 4 Selected features from PETA dataset [45]

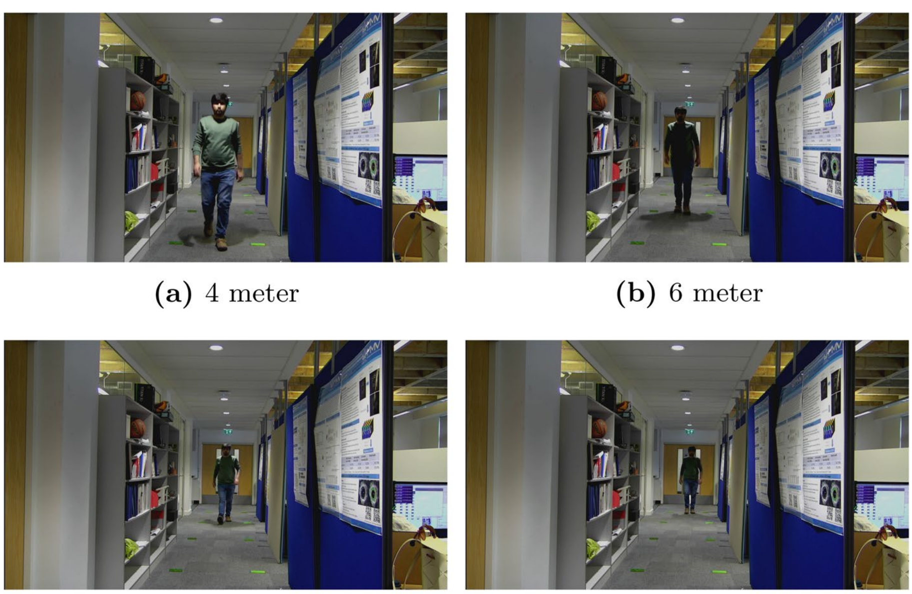
like DeepFashion [46]. There generally exists a strong co-occurrence between two soft biometrics of the human body, known as mustache and male, which are highly correlated. This has been confirmed by various research experiments. Similarly, gender and age are strongly correlated with several other essential and auxiliary attachments of the human body.

3.3 Developing MMV pedestrian dataset

In earlier research on soft biometrics [5], various challenges were identified, such as the need to have repeating sessions for individuals with a certain time gap and images captured at changing distances from the camera in a long corridor.

To address these challenges, we recorded a new dataset known as MMV Pedestrian dataset [44], as shown in Fig. 5. It contains recordings of 50 people walking in a corridor and there were markers placed on the floor at four distinct distances from the camera. The markings were put at four different distances such as 4, 6, 8, and 10-m from the camera. The MMV Pedestrian dataset includes 38 males and 12 females, and they belong to 19 different nationalities. In addition to this, the age range of people was large enough to range from 21 years and up to 55. The images were captured with a resolution of 1280 × 720 and approximately 300 frames were captured for each individual at the rate of 30 FPS. The multi-channel soft biometrics framework proposed in this research will use MMV Pedestrian dataset for the prediction of global soft biometrics [47].

Moreover, an extended facial region includes the hair, forehead, and neck, and it is essential to predict several global soft biometrics such as age, gender, and ethnicity. To achieve this goal, this paper used an existing model for face detection known as Multi-task Cascaded Convolutional Network (MTCNN) [48]. Using MTCNN, the extended face region was detected and segmented by applying geometrical scaling [47]. The model



(a) 4 meter **(b)** 6 meter **(c)** 8 meter **(d)** 10 meter

Fig. 5 Images at four different distances from MMV Pedestrians dataset was applied to the images captured at four distances from the camera and reported an accuracy of approximately 96%.

4 Our proposed multi-channel soft biometrics framework

In earlier research on soft biometrics several limitations were identified [5] while the aggregation of concepts learned from different research experiments, this paper proposes a multi-channel soft biometrics framework and its architecture in Fig. 6. This framework includes four different channels as follows, (i) ApparelNet [43], (ii) A-Net, (iii) OneDetect [47], and (iv) RSFS [45]. These four different channels were designed to work on images or features collected from multiple individuals. Each of these four channels requires a specific type of input data for the training and evaluation—later multiple different types of outputs were produced as part of the verification process. Several different AI models [49] were proposed in this research as distinct channels.

The multi-channel soft biometrics framework proposed in this paper takes two main types of inputs: images and annotations. To acquire images, two different datasets, FVG [38] and MMV Pedestrian dataset, were used. To develop a set of relative and supportive features, the well-known PETA [42] dataset was used. The FVG dataset provides images of 12 individuals walking at three different speeds and three different angles. It was recorded with a 1-year gap and includes people with changing appearances, such as variations in essential and auxiliary clothing. On the other hand, the MMV Pedestrian dataset was recorded specifically for border

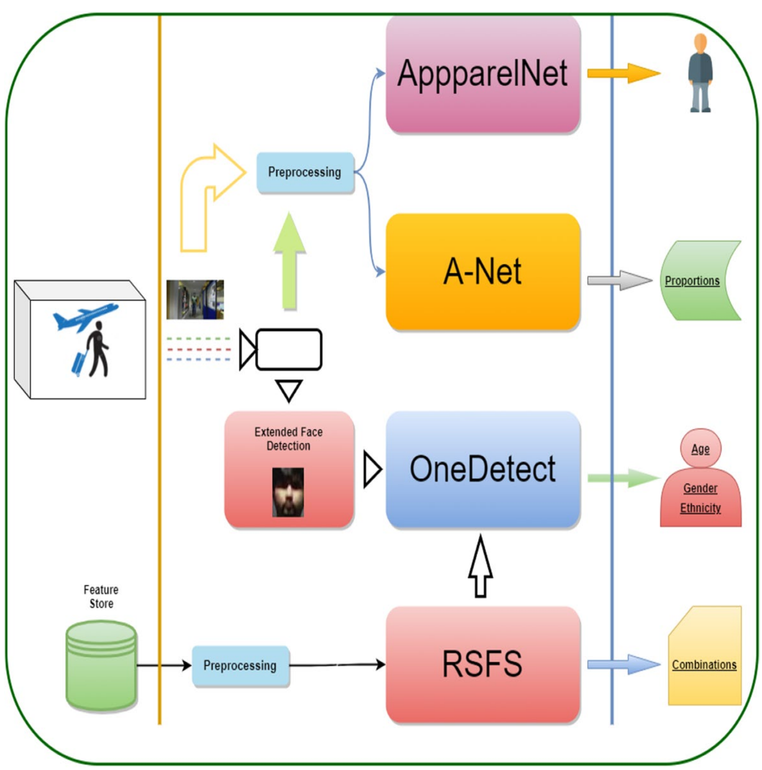


Fig. 6 Our proposed multi-channel soft biometrics framework

verification scenarios. It contains image sequences of people walking in a corridor with markers placed on the surface at multiple distances from the camera to track their exact distance.

For the relative and supportive feature set, the PETA dataset was used, which is composed of 10 different image datasets. It provides annotations for 61 binary and 4 multi-class attributes. More than 1K images from five different datasets were selected based on expert opinion to develop a limited number of relevant and supportive features. Overall, the proposed multi-channel soft biometrics framework and its channels perform the following functions: (i) ApparelNet for essential and auxiliary attachment-based verification, (ii) A-Net for anthropometric features estimation, (iii) OneDetect for gender, age, and ethnicity prediction, and (iv) RSFS to develop a highly relative and supportive set of soft biometrics, establishing a real-world relationship among sets of soft biometrics. Therefore, the output format for each channel is different. However, the ultimate goal was to facilitate the overall verification process and determine a composite matching score.

4.1 Channel one: ApparelNet

Person verification is quite a broad discipline and it is mainly studied with person re-identification [50], however, they tend to have a little different implementation, output format, and applications. Usually, verification is one-to-one while re-identification is one-to-many, like person-to-document matching and query from database respectively [51]. Most of the time, verification is performed using facial soft biometrics, including multiple modalities from the human body [19, 44] and it may include fusion of multiple modalities too [51]. In an earlier research, an ApparelNet architecture was proposed [43], which uses essential and auxiliary attachments based soft biometrics. This is a temporary modality, however it is a reflection of cultural norms and it has genuine application in short-term tracking. On the other hand, it can be an ancillary component for a verification system [5].

Hence, the proposed ApparelNet architecture contains several components, such as OpenPose [52] and a pre-trained EfficientNetB0 [53] with additional layers, and it was trained and evaluated on a subset of original and augmented images from FVG [38], as shown in Fig. 7 and 8, respectively. The pre-trained EfficientB0 is also fine-tuned for improved verification, and it is a highly scalable and efficient transfer learning model. This research, reported an overall training and validation accuracy of about 98%. In terms of the border verification replication scenario, 12 images, one from each class of FVG, was selected, and the probability of matching with all 12 classes was computed and it was approximately 96%. Different images from FVG were used for training, validation, and testing, including images from the years 2017 and 2018, while they contain a diverse range of variations present in the FVG dataset. Based on the verification scenario, the model may be useful for one-to-many scenarios as well, such as re-identification. Moreover, the purpose of using OpenPose was to use only those images for training that have detection confidence higher than the defined threshold. Later on, OpenPose localized 25 landmarks of the human body on a 2D image with detection confidence for each distinct landmark to compute a composite detection score for a single image [53].

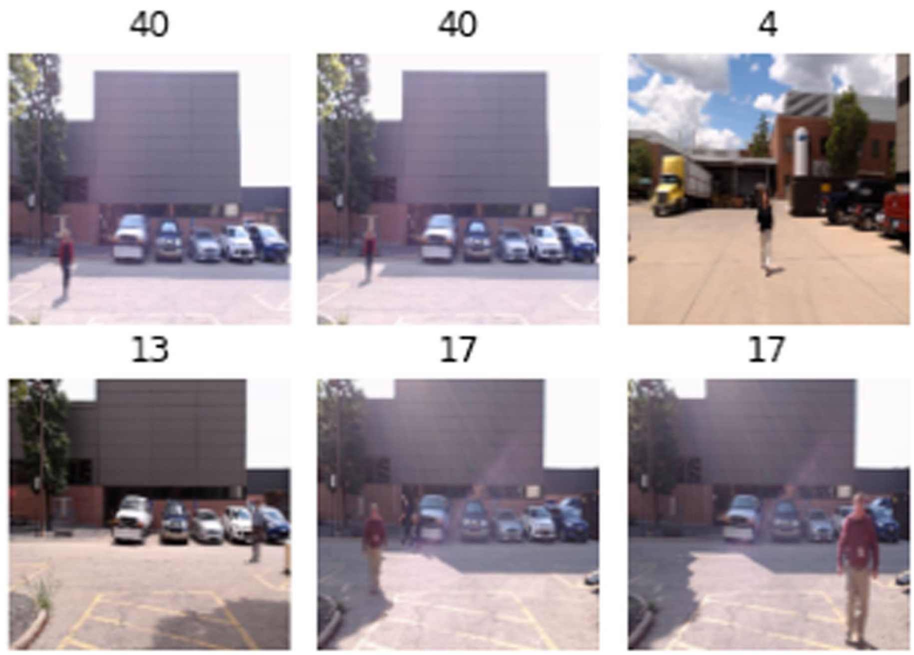


Fig. 7 Original images used from FVG dataset [38]

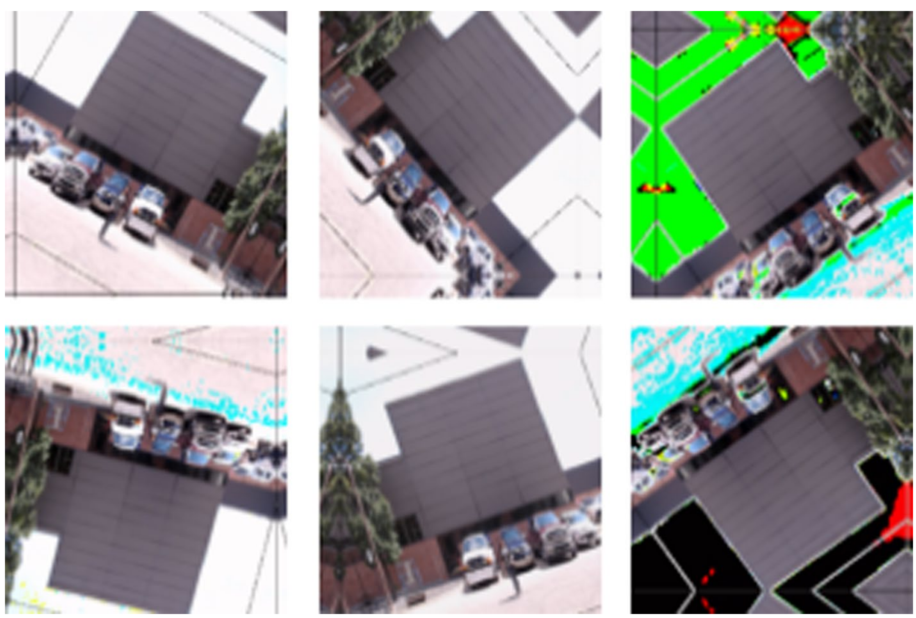


Fig. 8 Images after augmentation [43]

4.1.1 Proposed architecture and validation

The ApparelNet architecture [43] uses CMU-OpenPose [54] to first detect the whole body with its confidence scores for each landmark in a given image $\boldsymbol{I}$. The detection confidence range is from 0 to 1, and a threshold $\boldsymbol{T}$ of 0.5 or higher is set. This generates a list of confidence scores $\boldsymbol{A}$ from the image $\boldsymbol{I}$. If each value in the list satisfies the threshold, it is included in set $\boldsymbol{B}$; otherwise, it is ignored. In this way, all images with threshold levels equal to or higher than 0.5 are used for training the proposed transfer learning architecture.

$$B \subset A : I \in A \geq T \tag{1}$$

The second step in the ApparelNet architecture involved using EfficientNetB0 [55]. A few additional layers were added on top of the existing model to compute high-level features. To extract these high-level features, several layers were added on top of the model, such as dense with softmax activation, global average pooling, and dropout with a ratio of 0.2. The training of the model using the FVG dataset [38] was performed by fine-tuning the overall model. As the data was limited, augmentation steps such as translation, rotation, flipping, and contrast [56] were performed on the original data to avoid any over-fitting problems. Therefore, the proposed model had 15K trainable parameters. Finally, a total of 8K images from 12 distinct image classes of FVG were used for training and validation, with two sets of 80% and 20%, respectively [56]. The model reported a training and validation accuracy of approximately 98%.

4.2 Channel two: A-Net

As mentioned earlier, the quantitative mode of annotating the dataset is a critical factor for achieving further accurate recognition [57, 58]. However, accurately detecting anthropometric features remains a significant challenge, as landmark localization is the baseline for estimating anthropometric features from a 2D image, as identified in earlier research [5].

To overcome this challenge, OpenPifPaf [59] is a widely used pose estimation tool to extract anthropometric soft biometrics from 2D images, particularly suitable for delivery robots and self-driving cars. It outperforms its competitors in low-resolution, crowded, occluded, and cluttered scenes, and provides locations of 17 body joints in the 2D image along with their detection confidence. As the proposed features-based recognition system will be used in public places that are often crowded, OpenPifPaf proved to be highly effective in extracting human body landmarks from an image, with its detection confidence for each landmark [60]. OpenPifpaf takes an image as input, processes it, and outputs an image with localized landmarks.

Our developed A-Net architecture as shown in Fig. 9, customizes the output of OpenPifpaf. It access and store the required landmarks information from the face, body, and foot. The landmark information was stored in a data frame of size 17 X 3, and each row is a landmark as shown in Algorithm 1.

Algorithm 1: Customizing OpenPifPaf - accessing landmarks for each individual

Input: OpenPifpaf JSON O/P File
Output: Data_Frame: 17 X 3
$n \leftarrow NoOfPeopleIn.jsonfile$;
for $People\_Counter \leftarrow 0$ ***to*** $n-1$ **do**
  $Key\_Points \leftarrow people[People\_Counter]$;
  $Data\_Frame[0] \leftarrow Key\_Points$;

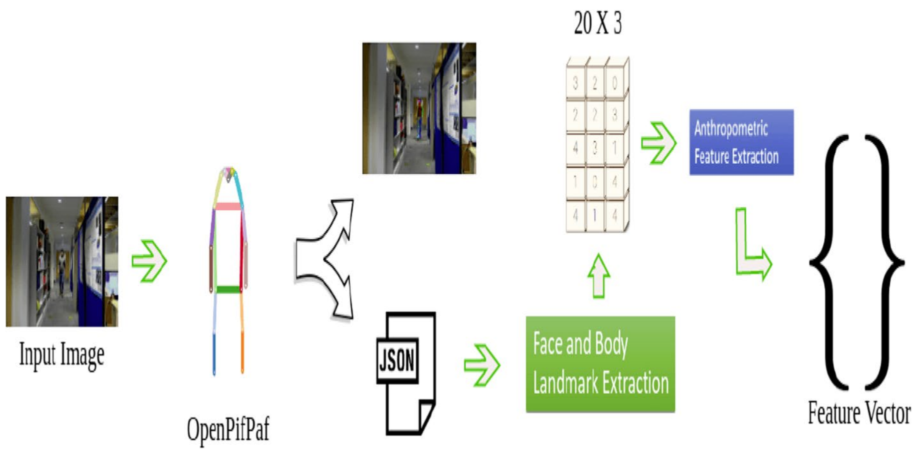


Fig. 9 OpenPifPaf framework for anthropometric feature estimation

The estimation process is computing anthropometric features which includes ratios and the measurements of the human body to be stored in a feature vector. For example, the Euclidean distance can be used to measure the ratio between upper and lower body [61], using their respective landmarks. Like the lower body length can be measured by computing the Euclidean distance between the hip center and ankle center. Similarly, the upper body length can be measured by computing the Euclidean distance between the neck and the hip center while to compute the hip center, it requires half of Euclidean distance between the right and left hip. As a result, the A-Net architecture provides a feature vector with ratios between different human body parts.

Algorithm 2: Computing ratio between Upper and Lower body

Input: DataFrame[Index]
Output: AttributeValue
$HDC \leftarrow HeadDetectionConfidence$;
$RHDC \leftarrow RightHipDetectionConfidence$;
$LHDC \leftarrow LeftHipDetectionConfidence$;
$LADC \leftarrow LeftAnkleDetectionConfidence$;
$RADC \leftarrow RightAnkleDetectionConfidence$;
$A \leftarrow Head$;
$B \leftarrow LeftHip$;
$C \leftarrow RightHip$;
$D \leftarrow LeftAnkle$;
$E \leftarrow RightAnkle$;
if $HDC \geq 0.7 \& RHDC \geq 0.7 \& LHDC \geq 0.7 \& LADC \geq 0.7 \& RADC \geq 0.7$ **then**
  $HipCenter \leftarrow EuclideanDistance(B, C)$;
  $AnkleCenter \leftarrow EuclideanDistance(D, E)$;
  $UpperBody \leftarrow EuclideanDistance(A, HipCenter)$;
  $LowerBody \leftarrow EuclideanDistance(AnkleCenter, HipCenter)$;
  $AttributeValue \leftarrow Ratio(UpperBody, LowerBody)$;
else
  $AttributeValue \leftarrow -1$;

4.3 Channel three: OneDetect

Predicting age, gender and ethnicity is the biggest challenge during the verification process generally and at border control more specifically [16]. Usually, many different standalone applications were used to pedict global soft biometrics in the past using images. In am earlier research on soft biometrics [5], several single and hybrid models for the prediction of global soft biometrics were summarized as shown in Fig. 10. Most of the time, face, body, or clothing modalities were used for this purpose [23]. Moreover, in many cases, distinct datasets were used for the training and evaluation of such models. They demonstrated very good performance, however they were not generalized enough. One of the biggest reasons for lacking generalization in global soft biometrics prediction is the non-availability of whole data at one point [19]. A common reason for non-availability is the application of privacy and ethics laws on image data.

To avoid ethics and privacy concerns and to make whole data available at one point, a federated learning architecture known as OneDetect [47] was proposed for the prediction of global soft biometrics [62]. In this client–server architecture, the model visits each client for the purpose of training instead of data visiting the model. The concept of federated averaging was applied to update model weights [63] and later on, the model weights were updated at the server end. This process continues until each client updates

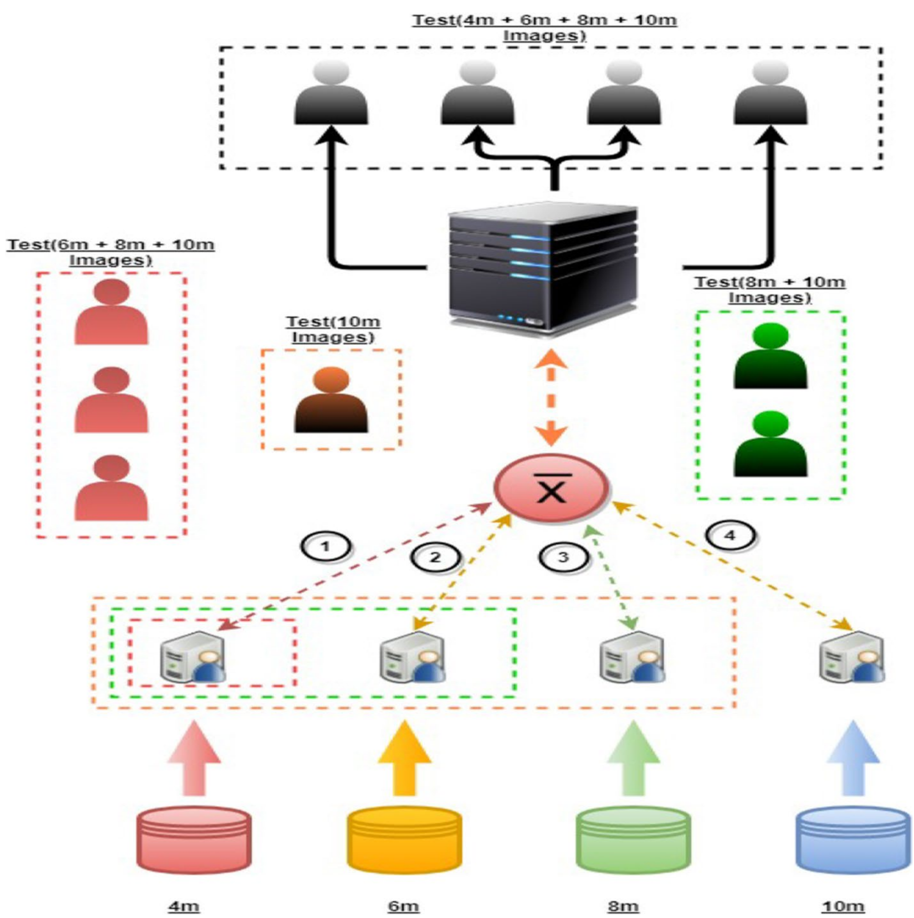


Fig. 10 OneDetect model [47]

its weights to the server. Another advantage of this model is to include and update its weights through any client, whenever a new instance of any individual is available after verification. It is usually acquired after a certain gap of time, from different orientations, changing distances from the camera, and varying ambient conditions [57]. More importantly, OeDetect [47] architecture addresses several issues related to data privacy and ethics by sending the model to the data using a federated learning approach. In order to develop a federated learning model for the prediction of global soft biometrics, EfficientNetB3 [53], was selected as a base model. The EfficientNetB3 was trained using UTK Face dataset [64] originally. The images from MMV Pedestrian dataset were selected to the Base model. The selected set of images contains four different subsets of images for each individual that were recorded at four changing distances from the camera, while people walking in a corridor similar to an airport. As discussed earlier, extended face region was segmented already for all 50 people using four subsets of selected images [65]. By this, four experiments were performed using OneDetect to predict of global soft biometrics as follows, (i) evaluation of Base model using images acquired at 4, 6, 8, and 10-m distances, (ii) training of Base model using 4-m images (Base04) and evaluation using 6, 8 and 10-m images, (iii) training of Base04 model using 6-m images (Base46) and evaluation using 8 and 10-m images, and finally, (iv) training of Base46 model using 8-m images (Base468) and evaluation using 10-m images.

4.4 Channel four: RSFS

In earlier research on soft biometrics [5], the development of a highly relative and supportive set of soft biometrics for improved recognition was investigated as a critical area under soft biometrics research. It is evident from different research experiments [66], that soft biometrics are related to each other and they support each other which always results in better recognition. though it depends on the application domain too. One of the key methods to find a relation between soft biometrics is correlation, [31], however it works only for pairs of soft biometrics and results in the form of direct or inverse correlation.

Usually recognition [67] process, may have more than two soft biometrics while correlation has no application in such type of scenario. That is why, to perform regression [68] analysis among a given set of soft biometrics is a useful technique. Not only it determine the relationship among soft biometrics but the level of support they provide each other. To develop a set of highly relevant and supportive set of soft biometrics, a set of essential and auxiliary clothing soft biometrics was built from the PETA [42] dataset. The total 1K annotated images were used. Seven annotated soft biometrics were termed as independent variables while age and gender were declared dependent variables. As part of RSFS architecture [45], two different experiments were conducted, the level of relationship and support using independent variables were determined for age and gender separately. To perform the regression analysis,several non-linear layers were used for the prediction of age and gender. A set of clothing pattern-based soft biometrics was used for this purpose.

Normalization was a preliminary step to perform regression analysis using non-linear layers on the numerical data [69]. In deep learning individual features are usually multiplied by the model weights and the output scales and gradients are affected by the inputs. The normalization process actually stabilizes the overall model training process. In this research, normalization layers from KERAS [70] were used and they computed mean and variance and stored in the layer through which normalized features are returned. Finally, a deep learning model comprised of multiple non-linear layers was used to predict age and gender on a given set of soft biometrics. This model actually performs regression analysis for age and gender with other 7 soft biometrics. It is of utmost important to discuss here that the total dataset was divided into three sets, known as training, validation, and test sets. The complete model includes several dense layers besides the normalization layer and it had 4737 trainable parameters.

5 Evaluating our proposed multi-channel soft biometrics framework

To validate the proposed framework, this paper evaluates each distinct channel of the proposed multi-channel soft biometrics framework independently, and finally, uses them to compute a cumulative output value. Several different formatted inputs were provided to each channel and each of them provided different outputs. However, an aggregated analysis and outcome of each channel supports the overall verification process at border control and recognition at large.

5.1 Outcome of ApparelNet

In the proposed multi-channel soft biometrics framework, ApparelNet is the first architecture designed to verify individuals with several different variations, including images recorded over several months, different walking speeds with and without essential and auxiliary attachments, and images acquired from different camera views. For the purpose of training and validating this proposed architecture, 12 images of individuals from the FVG dataset [38] were used without any auxiliary attachment. Later, more than 2K images of the same individuals were used for the purpose of evaluation, although they were different from the training and validation images. This proposed architecture reported an overall accuracy of approximately 96%, as shown in Fig. 11.

In addition, a one-vs-rest strategy was formulated, where one image per person was selected randomly and compared with the rest of the images from each class. The matching probability for each class was presented using the one-vs-rest schema. The highest matching probability was the true outcome of this proposed ApparelNet architecture for query images against a given set of images. This process of matching was actually a verification at border control, where the identity document of an individual contains the image of a person, alongside other personal information, and is matched against a dataset of images.

5.2 Outcome of A-Net

The accuracy of anthropometric soft biometrics, while extracting them from image sequences, is usually affected by several factors such as the distance from the camera, the camera view, and occlusion. For example, the length of limbs or legs is affected by the distance from the camera, and this remains the same for images of the same person as well. Although managing the distance factor is extremely difficult, setting parameters like camera calibration can be useful to interpolate different soft biometrics over distance. On the other hand, mathematical ratios and proportions may result in

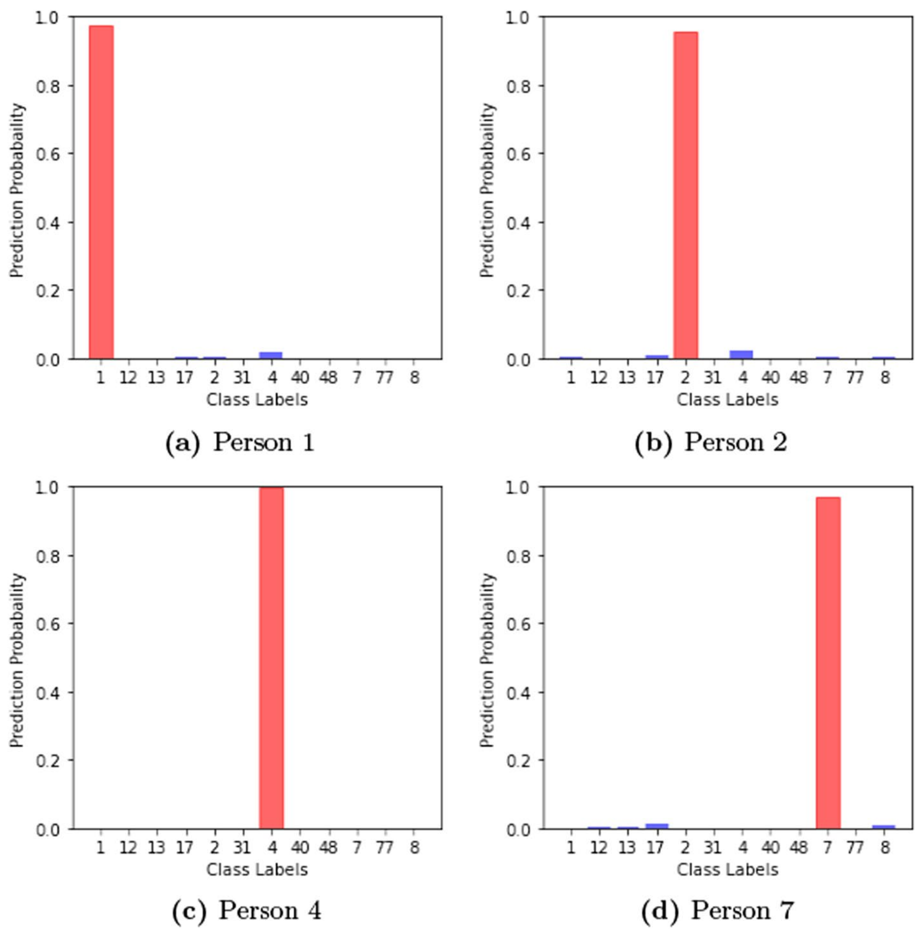


Fig. 11 One-Vs-Rest prediction score using proposed ApparelNet [47]

Table 1 Outcome of A-Net for shoulder–elbow length using three different images

Person no.	Right shoulder–elbow length		
	4 m	**6 m**	**8 m**
1	27.17958103061924	25.01403813872865	22.71992014271561
2	42.04004226924609	39.94966057329605	35.31176439941793
3	35.134080918675004	31.98958400402402	29.193923922261686

better estimation of anthropometric soft biometrics, as they are even more accurate than straight-line measurements of soft biometrics. In any case, whether straight-line measurements or proportions are used, human body landmark localization tools such as OpenPifPaf are preliminary components [71].

In the first experiment, the straight-line distance between the right shoulder and elbow was measured using the A-Net architecture, and the Euclidean distance was used for this purpose, as shown in Table 1. Three images of each of the three persons were taken to carry out the experiments, with the images acquired at distances of 4, 6, 8, and 10 m from the camera. It was observed that increasing the distance from the camera decreased the length between the shoulder and elbow for each person. Based on this experiment, it is evident that straight-line length measurements are highly affected by the distance from the camera, and other factors like the camera's calibration parameters and distance from the camera are critical in determining the value of anthropometric soft biometrics.

5.3 Outcome of OneDetect

The multi-channel soft biometrics framework proposed in this paper, OneDetect, is a key component based on federated learning for predicting age, gender, and ethnicity. It was evaluated using images from the MMV Pedestrian dataset, which were of the same individuals recorded at four different distances from the camera. To predict these global soft biometrics, an aggregated average is computed. Our proposed OneDetect architecture begins with the training of the Base model, which is EfficientNetB3, and all four different distance images were used for evaluation.

To evaluate the outcome of our proposed OneDetect architecture, the four different distance images of all 50 individuals were provided to all four models under the architecture, i.e., Base, Base04, Base46, and Base468 models. In Fig. 12, the difference in age approximation is presented for all 50 people using all four types of models. It has been observed that most of the time, age prediction was in the close age range of ±5, while there were only 9 instances for each model when it remained between the age range of ±10. There was only one instance for each model where a completely wrong age group approximation seemed to occur.

Similar to the age prediction experiment, we used the same four different distance images of individuals from the MMV Pedestrian dataset and provided them to four different models under our proposed federated learning-based OneDetect architecture, as depicted in Fig. 13. We selected images of 5 individuals as samples for predicting their gender and ethnicity. It was observed that the prediction rate was higher when the Base model was trained on 4, 6, and 8-m images. However, the prediction rate was very low when using the Base model alone.

5.4 Outcome of RSFS

The last component in our proposed multi-channel soft biometrics framework was RSFS. It was actually the development of a highly supportive and relative set of soft biometrics

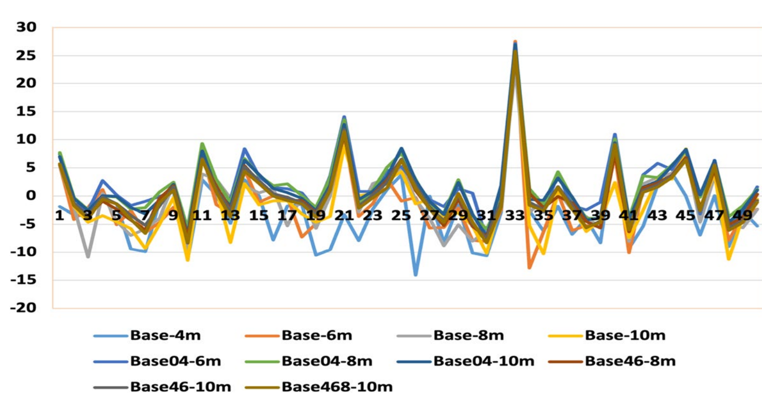


Fig. 12 Different between actual and predicted age

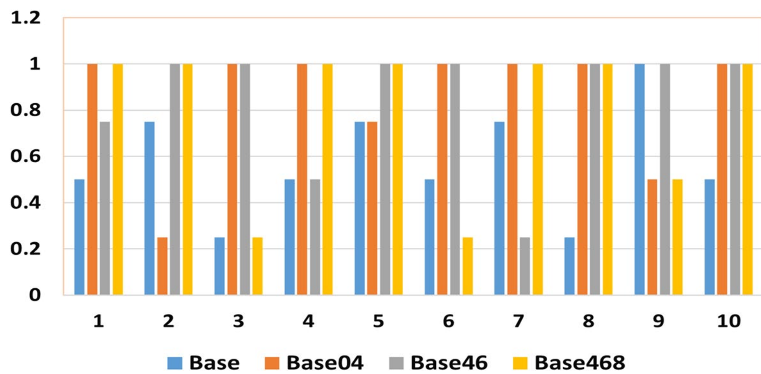


Fig. 13 Gender prediction using different models

Table 2 MAE for age and gender

Model	Age	Gender
DNN	0.538575	0.345418

by which age and gender can be determined. To accomplish this task, a subset of the PETA dataset was used and 7 different clothing pattern soft biometrics with annotations were selected. The following soft biometrics age and gender were predicted using DNN regression and performance was measured using Mean Absolute Error (MAE) as shown in Table 2.

5.5 Analyzing cumulative outcome of proposed framework

Following the detailed discussion on the design and evaluation of our proposed multi-channel soft biometrics framework, it is evident that soft biometrics are the future of border verification specifically and surveillance in general. Soft biometrics incorporate several modalities from the human body, whether temporary or permanent, and there is a huge amount of soft biometrics present in each modality. Our proposed multi-channel soft biometrics framework takes several different types of inputs, such as images and annotations, and performs verification of individuals using their essential and auxiliary attachments, anthropometrics, extended face region, and annotations. The output is presented in the form of person verification, including measurements of anthropometric soft biometrics, the prediction of global soft biometrics, and the development of a highly relevant and supportive set of soft biometrics.

Each channel included in our proposed multi-channel soft biometrics framework has performed well overall. However, there are certain limitations that need to be addressed to make it more generalized and supportive. For instance, recording people with more variance in appearance depicting modern culture and trends, and using an extended collection of anthropometrics from the human body. Furthermore, it should be trained by more clients and more instances of individuals for global soft biometrics prediction. Finally, the top three channels of our proposed multi-channel soft biometrics framework should use only those soft biometrics that are relevant and supportive, as proposed by the fourth channel RSFS, and from each modality of the human body.

6 Conclusion

The principal contribution of this research is the development of a multi-channel soft biometrics framework that works on images from two different datasets and features from the PETA dataset. The overall objective of this research was to verify people at border control, especially for recognition during surveillance. The research made four contributions in a hybrid framework: (i) evaluation of transfer learning architecture for auxiliary attachment-based verification over a long period of time for the same individual, (ii) evaluation of federated learning architecture for predicting global soft biometrics such as gender, age, and ethnicity, (iii) development and evaluation of human body anthropometrics for recognition in a walking corridor, and (iv) proposal for a highly relevant and supportive set of soft biometrics to facilitate the overall recognition process. Moreover, in this research, several potential extensions were analyzed as part of future research to make the proposed multi-channel soft biometrics framework more robust for border verification. Future work can be categorized into three main areas. First, to further improve the prediction of global soft biometrics, a federated learning architecture can be proposed, which should be re-introduced as the One-Identity system where the same instances of the same people are used for training the current model after verification each time. Second, anthropometric soft biometrics from the human body should be included to enhance the overall recognition process. Finally, to include more annotated images and annotations for the development of a highly relevant and supportive set of soft biometrics. These extensions can make our proposed multi-channel soft biometrics framework more robust for border verification, ultimately leading to a more seamless, secure, and fast border verification process.

Abbreviations

FVG	Front-view Gait
PETA	Pedestrian attribute recognition at far distance
MMV	Multimedia and vision
RFSF	Relative support feature set
CNN	Convolutional neural netowrk

Acknowledgements
Not applicable.

Author Contributions
All authors equally contributed. All authors approved the final manuscript.

Funding
Not applicable.

Availability of data and materials
The data used in this research is already available publicly.

Declarations

Ethics approval and consent to participate
The research carried out in this paper is in accordance with the ethical standards of the research community

Consent for publication
All authors agree to publish this manuscript.

Competing interests
The authors declare that they have no competing interests

Received: 31 January 2023 Accepted: 25 May 2023
Published online: 17 June 2023

References

1. J.-Y. Yu, Y. Kim, Y.-G. Kim, Intelligent video data security: a survey and open challenges. IEEE Access **9**, 26948–26967 (2021)
2. A. James, Government mass surveillance and law in the five eyes countries. Ph.D. thesis (2018)
3. S. Milivojevic, *Border Policing and Security Technologies: Mobility and Proliferation of Borders in the Western Balkans* (Routledge, London, 2019)
4. M. Leese, Fixing state vision: interoperability, biometrics, and identity management in the EU. Geopolitics **27**(1), 113–133 (2022)
5. B. Hassan, E. Izquierdo, T. Piatrik, Soft biometrics: a survey. Multimedia Tools and Applications, 1–44 (2021)
6. D. Semedo, D. Carmo, R. Padnevych, J. Magalhaes, Contact-free airport borders with biometrics-on-the-move, in: 2021 IEEE International Workshop on Biometrics and Forensics (IWBF), pp. 1–2 (2021). IEEE
7. What is biometrics? https://www.techtarget.com/searchsecurity/definition/biometrics. (Accessed on 04/08/2022)
8. C. Galdi, J. Boyle, L. Chen, V. Chiesa, L. Debiasi, J.-L. Dugelay, J. Ferryman, A. Grudzień, C. Kauba, S. Kirchgasser, Protect: pervasive and user focused biometrics border project-a case study. IET Biom **9**(6), 297–308 (2020)
9. S. Noori, Suspicious infrastructures: automating border control and the multiplication of mistrust through biometric e-gates. Geopolitics **27**(4), 1117–1139 (2022)
10. S. Zawadzka et al., Biometric technology in European union border management after 2015. Prz Geopolit **39**, 93–113 (2022)
11. C. Charrier, C. Rosenberger, A. Nait-Ali, Biometrics and applications (2022)
12. Q.N. Tran, B.P. Turnbull, J. Hu, Biometrics and privacy-preservation: How do they evolve? IEEE Open J. Comput. Soc. **2**, 179–191 (2021)
13. I. Stylios, S. Kokolakis, O. Thanou, S. Chatzis, Behavioral biometrics & continuous user authentication on mobile devices: a survey. Inf. Fus. **66**, 76–99 (2021)
14. B. Hassan, E. Izquierdo, K. Chandramouli, Performance assessment of soft biometrics technologies for border crossing. Secur. Technol. Soc. Implic 82–121 (2022)
15. M. Aizeki, G. Boyce, T. Miller, J. Nevins, M. Ticktin, Smart borders or a humane world? (2021)
16. M. Ye, J. Shen, G. Lin, T. Xiang, L. Shao, S.C. Hoi, Deep learning for person re-identification: a survey and outlook. arXiv preprint arXiv:2001.04193 (2020)
17. Y. Lin, Y. Wu, C. Yan, M. Xu, Y. Yang, Unsupervised person re-identification via cross-camera similarity exploration. IEEE Trans. Image Process. **29**, 5481–5490 (2020)
18. F. Wan, Y. Wu, X. Qian, Y. Chen, Y. Fu, When person re-identification meets changing clothes, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp. 830–831 (2020)
19. A. Dantcheva, P. Elia, A. Ross, What else does your biometric data reveal? A survey on soft biometrics. IEEE Trans. Inf. Forensics Secur. **11**(3), 441–467 (2015)
20. A. Sundararajan, A.I. Sarwat, A. Pons, A survey on modality characteristics, performance evaluation metrics, and security for traditional and wearable biometric systems. ACM Comput. Surv. **52**(2), 39 (2019)
21. R.M. Bolle, J.H. Connell, S. Pankanti, N.K. Ratha, A.W. Senior, *Guide to Biometrics* (Springer, Berlin, 2013)
22. J. Chen, H. Zhao, Z. Cao, W. Zhao, L. Pang, Successive minutia-free mosaicking for small-sized fingerprint recognition. IET Image Proc. **13**(7), 1146–1151 (2019)
23. M.S. Nixon, P.L. Correia, K. Nasrollahi, T.B. Moeslund, A. Hadid, M. Tistarelli, On soft biometrics. Pattern Recogn. Lett. **68**, 218–230 (2015)
24. A. Dantcheva, C. Velardo, A. D'angelo, J.-L. Dugelay, Bag of soft biometrics for person identification. Multimed. Tools Appl. **51**(2), 739–777 (2011)
25. H. Zhang, J.R. Beveridge, B.A. Draper, P.J. Phillips, On the effectiveness of soft biometrics for increasing face verification rates. Comput. Vis. Image Underst. **137**, 50–62 (2015)
26. M.S. Nixon, B.H. Guo, S.V. Stevenage, E.S. Jaha, N. Almudhahka, D. Martinho-Corbishley, Towards automated eyewitness descriptions: describing the face, body and clothing for recognition. Vis. Cogn. **25**(4–6), 524–538 (2017)
27. S. Denman, M. Halstead, C. Fookes, S. Sridharan, Searching for people using semantic soft biometric descriptions. Pattern Recogn. Lett. **68**, 306–315 (2015)
28. D. Li, Z. Zhang, X. Chen, H. Ling, K. Huang, A richly annotated dataset for pedestrian attribute recognition. arXiv preprint arXiv:1603.07054 (2016)
29. E. Gonzalez-Sosa, J. Fierrez, R. Vera-Rodriguez, F. Alonso-Fernandez, Facial soft biometrics for recognition in the wild: recent works, annotation, and cots evaluation. IEEE Trans. Inf. Forensics Secur. **13**(8), 2001–2014 (2018)
30. O.A. Arigbabu, S.M.S. Ahmad, W.A.W. Adnan, S. Yussof, Recent advances in facial soft biometrics. Vis. Comput. **31**(5), 513–525 (2015)
31. J. Benesty, J. Chen, Y. Huang, I. Cohen, Pearson correlation coefficient, in: Noise reduction in speech processing, pp. 1–4. Springer, Berlin (2009)
32. Y. Pang, T. Wang, R.M. Anwer, F.S. Khan, L. Shao, Efficient featurized image pyramid network for single shot detector, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7336–7344 (2019)
33. Y. Pang, J. Cao, X. Li, Learning sampling distributions for efficient object detection. IEEE Trans. Cybern. **47**(1), 117–129 (2016)
34. Y. Pang, J. Cao, X. Li, Cascade learning by optimally partitioning. IEEE Trans. Cybern. **47**(12), 4148–4161 (2016)
35. S. Zheng, J. Zhang, K. Huang, R. He, T. Tan, Robust view transformation model for gait recognition, in: 2011 18th IEEE International Conference on Image Processing, pp. 2073–2076. IEEE (2011)
36. R. He, T. Tan, L. Wang, Robust recovery of corrupted low-rankmatrix by implicit regularizers. IEEE Trans. Pattern Anal. Mach. Intell. **36**(4), 770–783 (2013)
37. S. Yu, D. Tan, T. Tan, A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition, in: 18th International Conference on Pattern Recognition (ICPR'06), vol. 4, pp. 441–444. IEEE (2006)
38. Z. Zhang, L. Tran, X. Yin, Y. Atoum, X. Liu, J. Wan, N. Wang, Gait recognition via disentangled representation learning, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4710–4719 (2019)

39. Z. Zhang, L. Tran, F. Liu, X. Liu, On learning disentangled representations for gait recognition. arXiv preprint arXiv:1909.03051 (2019)
40. K.W. Lee, N. Sankaran, S. Setlur, N. Napp, V. Govindaraju, Wardrobe model for long term re-identification and appearance prediction, in: 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), pp. 1–6. IEEE (2018)
41. P. Zhang, Q. Wu, J. Xu, J. Zhang, Long-term person re-identification using true motion from videos, in: 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 494–502. IEEE (2018)
42. Y. Deng, P. Luo, C.C. Loy, X. Tang, Pedestrian attribute recognition at far distance, in: Proceedings of the 22nd ACM International Conference on Multimedia, pp. 789–792 (2014)
43. B. Hassan, E. Izquierdo, Apparelnet: Person verification encompassing auxiliary attachments variation, in: 2021 IEEE 23rd International Workshop on Multimedia Signal Processing (MMSP), pp. 1–6. IEEE (2021)
44. B. Hassan, 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, p. 1187807. International Society for Optics and Photonics (2021)
45. B. Hassan, I. Izquierdo, RSFS: a soft biometrics based relative support features set for person verification (in press)
46. Y. Ge, R. Zhang, X. Wang, X. Tang, 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, pp. 5337–5345 (2019)
47. B. Hassan, I. Izquierdo, OneDetect: a federated learning architecture for global soft biometrics prediction (in press)
48. J. Xiang, G. Zhu, Joint face detection and facial expression recognition with mtcnn, in: 2017 4th International Conference on Information Science and Control Engineering (ICISCE), pp. 424–427. IEEE (2017)
49. I. Goodfellow, Y. Bengio, A. Courville, Y. Bengio, *Deep Learning*, vol. 1 (MIT Press, Cambridge, 2016)
50. M. Ye, J. Shen, G. Lin, T. Xiang, L. Shao, S.C. Hoi, Deep learning for person re-identification: a survey and outlook. IEEE Trans. Pattern Anal. Mach. Intell. **44**, 2872–2893 (2021)
51. A. Khatun, S. Denman, S. Sridharan, C. Fookes, Joint identification-verification for person re-identification: a four stream deep learning approach with improved quartet loss function. Comput. Vis. Image Underst. **197**, 102989 (2020)
52. Z. Cao, G. Hidalgo Martinez, T. Simon, S. Wei, Y.A. Sheikh, Openpose: Realtime multi-person 2d pose estimation using part affinity fields. IEEE Trans. Pattern Anal. Mach. Intell. **43**, 172–186 (2019)
53. M. Tan, Q.V. Le, Efficientnet: rethinking model scaling for convolutional neural networks. arXiv preprint arXiv:1905.11946 (2019)
54. Z. Cao, T. Simon, S.-E. Wei, Y. Sheikh, Realtime multi-person 2d pose estimation using part affinity fields, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7291–7299 (2017)
55. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, Imagenet: a large-scale hierarchical image database, in: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255. IEEE (2009)
56. S. Pal, S. Gulli, Deep learning with keras: Implementing deep learning models and neural networks with the power of python. Packt Publishing, Birmingham (2017)
57. D. Sadhya, P. Pahariya, R. Yadav, A. Rastogi, A. Kumar, L. Sharma, S.K. Singh, Biosoft-a multimodal biometric database incorporating soft traits, in: 2017 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA), pp. 1–6. IEEE (2017)
58. R.B. Fosdick, Passing of the Bertillon system of identification. J. Am. Inst. Crim. L. Criminol. **6**, 363 (1915)
59. S. Kreiss, L. Bertoni, A. Alahi, Pifpaf: Composite fields for human pose estimation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2019)
60. S.-E. Wei, V. Ramakrishna, T. Kanade, Y. Sheikh, Convolutional pose machines, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4724–4732 (2016)
61. P.-E. Danielsson, Euclidean distance mapping. Comput. Gr. Image Process. **14**(3), 227–248 (1980)
62. Q. Yang, Y. Liu, T. Chen, Y. Tong, Federated machine learning: concept and applications. ACM Trans. Intell. Syst. Technol. **10**(2), 1–19 (2019)
63. Y. Xia, D. Yang, W. Li, A. Myronenko, D. Xu, H. Obinata, H. Mori, P. An, S. Harmon, E. Turkbey, et al.: Auto-fedavg: learnable federated averaging for multi-institutional medical image segmentation. arXiv preprint arXiv:2104.10195 (2021)
64. Z. Zhang, Y. Song, H. Qi, Age progression/regression by conditional adversarial autoencoder, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5810–5818 (2017)
65. K. Zhang, Z. Zhang, Z. Li, Y. Qiao, Joint face detection and alignment using multitask cascaded convolutional networks. IEEE Signal Process. Lett. **23**(10), 1499–1503 (2016)
66. D. Reid, S. Samangooei, C. Chen, M. Nixon, A. Ross, Soft biometrics for surveillance: an overview, in: Handbook of Statistics, vol. 31, pp. 327–352. Elsevier, Amsterdam (2013)
67. N. Reddy, A. Rattani, R. Derakhshani, Comparison of deep learning models for biometric-based mobile user authentication, in: 2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS), pp. 1–6. IEEE (2018)
68. N.R. Draper, H. Smith, Applied Regression Analysis, vol. 326. Wiley, New York (1998)
69. A. Gulli, S. Pal, Deep Learning with Keras. Packt Publishing Ltd, Birmingham (2017)
70. N. Ketkar, Introduction to keras, in: Deep Learning with Python, pp. 97–111. Springer, Berlin (2017)
71. Y. Xiu, J. Li, H. Wang, Y. Fang, C. Lu, Pose flow: efficient online pose tracking. arXiv preprint arXiv:1802.00977 (2018)