IET Journals
The Institution of Engineering and Technology

PROTECT: Pervasive and useR fOcused biomeTrics bordEr projeCT. A Case Study.

ISSN 1751-8644
doi: 0000000000
www.ietdl.org

Chiara Galdi[1], Jonathan Boyle[4], Lulu Chen[4], Valeria Chiesa[1], Luca Debiasi[3], Jean-Luc Dugelay[1], James Ferryman[4], Artur Grudzień[2], Christof Kauba[3], Simon Kirchgasser[3], Marcin Kowalski[2], Michael Linortner[3], Patryk Maik[5], Kacper Michoń[5], Luis Patino[4], Bernhard Prommegger[3], Ana F. Sequeira[6], Łukasz Szklarski[5], Andreas Uhl[3]*

[1] *Digital Security Department, EURECOM, 450 Route des Chappes, Sophia Antipolis, France*
[2] *Institute of Optoelectronics, Military University of Technology, Gen. S. Kaliskiego 2, Warsaw, Poland*
[3] *Department of Computer Sciences, University of Salzburg, Jakob-Haringer-Str. 2, Salzburg, Austria*
[4] *Department of Computer Science, University of Reading, Whiteknights, Reading RG6 6AY, United Kingdom*
[5] *Research and Development Department, ITTI Sp. z o.o., ul. Rubież 46, 61-612 Poznań, Poland*
[6] *INESC TEC, University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal*
** E-mail: chiara.galdi@eurecom.fr*

Abstract: PROTECT: Pervasive and useR fOcused biomeTrics bordEr projeCT is an EU project funded by the Horizon 2020 research and Innovation Programme. The main aim of PROTECT was to build an advanced biometric-based person identification system that works robustly across a range of border crossing types and that has strong user-centric features. This work presents the case study of the multibiometric verification system developed within PROTECT. The system has been developed to be suitable for different borders such as air, sea, and land borders. The system covers two use cases: the walk-through scenario, in which the traveller is on foot; the drive-through scenario, in which the traveller is in a vehicle. Each deployment includes a different set of biometric traits and this paper illustrates how to evaluate such multibiometric system in accordance with international standards and, in particular, how to overcome practical problems that may be encountered when dealing with multibiometric evaluation, such as different score distributions and missing scores.

1 Introduction

Biometric systems make use of the physiological and/or behavioral traits of individuals for recognition purposes [1]. The physiological traits include, among others, the so called "ICAO biometrics" from the International Civil Aviation Organization, which has defined the biometric file formats and communication protocols to be used in passports and selected the following biometric traits for this purpose: facial recognition, fingerprint recognition, and iris recognition [2]. Other non-ICAO physiological biometric traits are hand-geometry, retina, and ear. Among the behavioral biometric traits there are gait, signature, and voice. Biometric systems that use a single trait for recognition are referred to as *single-modality* or *unimodal* biometric systems. Compared to an unimodal system, a multimodal biometric system based on different traits is expected to be more robust to noise, address the problem of non-universality, improve the matching accuracy, and provide reasonable protection against presentation attacks by consolidating the evidence obtained from different sources [3].

PROTECT*: Pervasive and UseR Focused BiomeTrics BordEr ProjeCT is a H2020 security research project aimed at delivering to border authorities an enhanced biometric-based on-the-move traveller identification system. The system can be deployed at a wide range of Automated Border Control (ABC) areas – air, land, and sea – supporting border guards to facilitate smooth and non-intrusive rapid crossing by travellers. The ability for the system to efficiently process low-risk travellers, combined with increased levels of accuracy, security and privacy standards, are central ambitions of the project. PROTECT project aims to introduce a less intrusive approach to biometric data capture and verification with the use of contactless biometrics. For the purposes of multimodal biometric ID verification, project partners proposed to design and develop a Biometric Capture Area

* *http://projectprotect.eu/*

(BCA), which would incorporate a number of biometric sensors. Furthermore, an important feature of PROTECT is that it aimed at exploring solutions for ABC beyond current legislation, meaning that non-ICAO biometric traits and protocols have been tested in the expectation that in the future they will be selected for their use in ABC, possibly endorsed also by the results of this project.

This paper presents the case study of the PROTECT multimodal biometric ABC solution. The evaluation aims at assessing the performances of two deployments:

1. The **walk-through** system is a biometric capture area (BCA) through which the traveller can pass on foot, eventually carrying a luggage, rucksack, or bag, and get automatically authenticated (ABC) on the move. The walk-through BCA was designed to be accessible – i.e. to wheelchair users – with appropriate placement of biometric sensors for this purpose. The biometric sensors used for this scenario are all contactless, allowing the traveller to walk without stopping. The system is located indoor in presence of artificial and natural light. Fig. 1 illustrates the design of the walk-through biometric corridor.
2. The **drive-through** system is a BCA through which the traveller can pass in a vehicle. The variability here is given by the height of the vehicle and the distance and angle between the traveller and the sensors that change depending on where the vehicle stops. The biometric sensors used for this scenario include contact and contactless acquisition and allow the traveller to get authenticated from within the vehicle. The system is located outdoor in presence of mainly natural light during the day, and artificial light during night. Fig. 2 illustrates the design of the drive-through biometric corridor.

Data collection was performed in order to reflect the two scenarios described above. Table 1 summarizes the biometric traits in the PROTECT multimodal dataset v2 and indicates in which acquisition scenario they were collected. A more detailed description of the dataset is provided in Section 2.

Our key contributions in this work are:

- We provide an insight on how to evaluate a multibiometric systems in accordance with international standards and, in particular, on how to overcome practical problems that may be encountered when dealing with multibiometric evaluation. This is performed on a non-chimeric database acquired in environments simulating two real-case scenarios. This is very different from performing experiments on existing fully-controlled datasets.
- We describe the process and issues related to collecting a non-chimeric multimodal biometric database including 10 different modalities;
- We address practical problems, such as having different score distributions, with different size and shape, that arise from having a very heterogeneous dataset;
- We provided detailed, thus reproducible, solutions for data augmentation and missing score imputation in order to obtain a set of scores of sufficient size to perform reliable performance assessment according to the guidelines provided by international standards.

1.1 Technical performance testing of biometric systems and devices

International standards were adopted for consistency and comparable testing of the biometric traits. In particular, data collection and evaluation were guided by the ISO/IEC 19795 [4] standard series.

Technical performance testing seeks to determine error and throughput rates, with the goal of understanding and predicting the real-world error and throughput performance of biometric systems.

Metrics:

- False match rate (FMR): proportion of zero-effort impostor attempt samples falsely declared to match the compared non-self template;
- False non-match rate (FNMR): proportion of genuine attempt samples falsely declared not to match the template of the same characteristic from the same user supplying the sample;
- Equal error rate (EER): value at which FMR = FNMR;
- FMR1000: the lowest FNMR for FMR <= 1/1000;
- ZeroFMR: the lowest FNMR for FMR = 0;
- Detection error trade-off (DET) curve: modified ROC curve which plots error rates on both axes (false positives on the x-axis and false negatives on the y-axis).

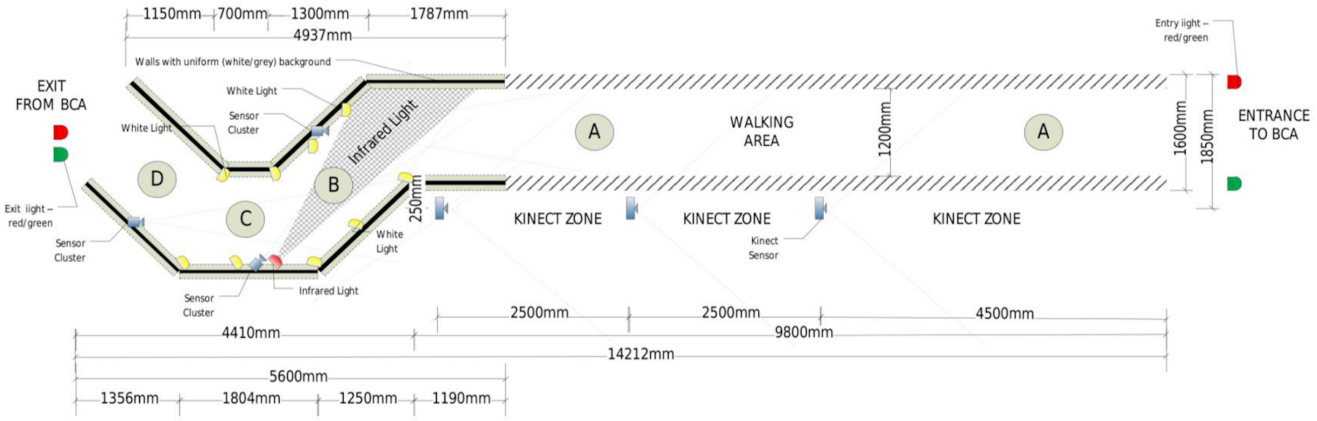

Fig. 1: Design of the Biometric Capture Area (BCA) for air/maritime borders (walk-through scenario). Section A is aimed at the capture of anthropometrics. Sections B, C, and D contain each a sensor cluster for the capture of 2D, 3D face, and periocular.

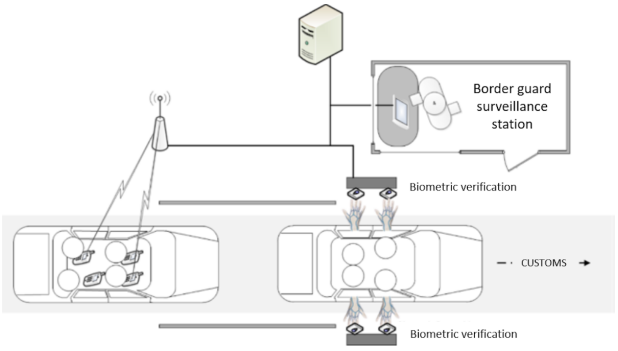

Fig. 2: Design of the Biometric Capture Area (BCA) for land borders (drive-through scenario).

2 PROTECT multimodal dataset version 2

A first version of the PROTECT dataset was collected in 2017 and the results of the experiments carried on it are reported in [5].

The target population for the PROTECT systems is the universe of travellers crossing the different types of border controls. It should be noted that, following the current laws that impose the users of eGates to be over 18 years old – in the majority of the European Countries – and due to the legal implications of collecting biometric data from minors, the under 18s were excluded from the targeted population. Considering that the target population is the universe of travellers, the corpus should include subjects with a wide range of variety in age, gender, ethnicity, and eye/skin types. The age of the subjects spans a large interval from 22 to 72. The age distribution in 5-year intervals can be observed in the graphic depicted in Fig. 3. Obtaining an operationally representative distribution of males and females was challenging as a result of the environment where the collection was made which is predominantly composed of male staff. The final distribution was 53% male, 44% female and 3% "prefer not to say" as shown in Fig. 4. Concerning the variety of ethnicities, the set of volunteers includes Caucasians, Asians, Africans, Indians, Middle-easterners and others. Naturally the variety of ethnicities results in a diversity of eye and skin types.

Data was recorded from 38 subjects. This collection was an extremely complex event and required a significant effort from all the partners involved. The collection events involved several logistically complex issues since all sensors had to be transported and installed in one place and the acquisition setup had to integrate all different sensors for each volunteer. In the collection of the final version of the PROTECT Multimodal DB dataset (v2) an effort was made to follow the recommendations given in the ISO/IEC standard documents regarding the multiplicity of sessions and other relevant aspects (such as a sufficient number of samples in each session and the repetition of the same sequence of collection across different sessions, among others) to ensure operationally representative variability in the data to be used for the performance evaluation of the recognition systems.

The collection spanned a period of one week and for some subjects it was possible to collect data in 2 different sessions, as much as possible one in the beginning of the week and another at the end. Fig. 5 shows two photos from the data collection campaign held at University of Reading in 2018. On the left, the data collection for the walk-through scenario: A corridor was assembled according to the design in Fig. 1. Clusters of sensors were attached to the corridor structure. In Fig. 5, on the right, the acquisition setup for the drive through corridor is

Table 1 List of biometric traits in the PROTECT multimodal dataset v2.

Biometric trait	Description	Acquisition scenario
2D Face	2D videos of the face (VIS and NIR)	Walk through, Drive through
3D Face	3D image of the face, composed by the 2D image plus the depth field information	Walk through, Drive through
Thermal Face	Body heat pattern of the face derived primarily from the pattern of superficial blood vessels under the skin	Walk through, Drive through
Iris	The coloured region of the eye enclosed between the pupil (dark region) and the sclera (white region)	Drive through
Periocular	Eye and surrounding area including eyelids, eyebrows (VIS and NIR)	Walk through
Finger Veins	The patterns of the human finger veins beneath the skin's surface	Drive through
Hand Veins	The patterns of the human hand veins beneath the skin's surface	Drive through
Anthropometrics	Skeleton joints positions	Walk through

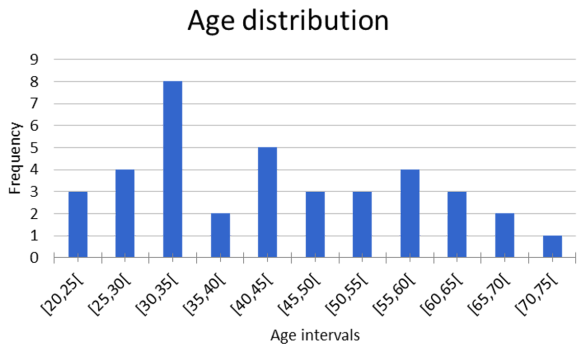

Fig. 3: Distribution of the age intervals of subjects in the corpus.

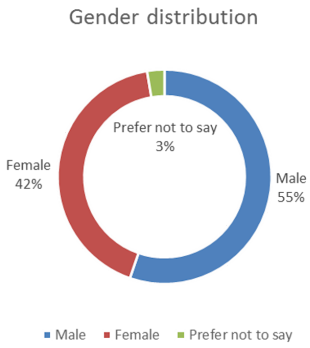

Fig. 4: Distribution of the gender of subjects in the corpus.

shown. The sensors were in turn positioned alongside the front-seat window. The subjects also used sensors on smartphones within the vehicle (e.g. for iris recognition).

In the following, the data collection as well as the verification algorithms for each biometric modality used in PROTECT are briefly described. More detailed descriptions of the methods can be found in the referenced articles in each biometric modality section below. A summary of the parameters used for the experiments in PROTECT is provided in Table 2.

2.1 2D face

As indicated in Table 1, as part of the PROTECT multimodal dataset v2, 2D face was captured in both 'walk-through' and 'drive-through' scenarios.

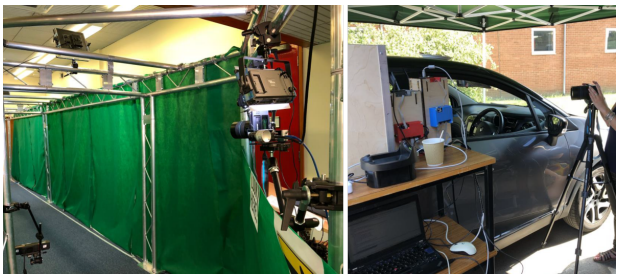

Fig. 5: Photos from the data collection event held at University of Reading. Left: simulated walk-through corridor. Right: drive through scenario data acquisition.

In the walk-through corridor scenario, two visible-light (VIS) digital cameras and one near infra-red (NIR) camera were used for collecting 2D faces. The NIR camera was placed between the two VIS cameras along the corridor. A video sequence was recorded by each camera while the user walked through the corridor. The user was instructed to perform multiple walks with three main types of variations: walking normally, walking with and without glasses on if the person has personal glasses, and pulling a wheeled travel bag and/or carrying a backpack. Some other variations include users holding a phone in front of the face, users making a phone call, and avoiding looking to the direction of the camera while walking. The variations aim to increase the range of challenges for face recognition such as non-straight head pose, partial occlusions, motion blur, sunglasses, and reflections across the dataset. Some examples from both VIS and NIR cameras are shown in Fig. 6.

One of the main objectives in PROTECT was to investigate the integration of biometric solutions across platforms. In particular, for 2D face, the verification needed to be implemented on both PCs and smartphones. Visage SDK is a mature product and optimised to support all major platforms including smartphone operating systems. As the task for 2D face was not a research topic, the implementation of 2D face verification was to use this mature commercial software. Also, in PROTECT, 2D face was not only used for biometric verification

Fig. 6: Example images of the 2D face extracted from the video streams in walk-through scenario.

but also for face tracking, thus requiring real-time face detection and recognition. This could not have been achieved with less mature open-source solutions such as OpenFace. OpenFace was instead used for 3D face verification (see Section 2.2) as the slow processing of light-field images with the current available libraries had already prevented the use of this modality in real time. The images selected for the evaluation were extracted automatically using the Feature Quality provided by Visage Technologies' software* for captured VIS video sequences, and manually confirmed to be of high quality. A minimum of 10 images per walk-through were extracted from across the two VIS cameras, with variations arising from quality of the overall sequence. Face verification was performed using the Visage face verification tool on each extracted face. The final score for a walk-through was the average of the scores against the gallery image. The setup for NIR face verification was similar to the VIS setup as described above. The selection of images was the same as for the VIS cameras as well as the method used in the calculation of the final scores per walk-through.

In the drive-through scenario, a modified Google Nexus 5X smartphone frontal camera was used for capturing selfie facial images while the user was sitting inside a vehicle. The NIR cut-off filter from the frontal camera was removed and the sensor refocused to enable NIR sensitivity imaging, with a custom attachable external NIR light, allowing the camera to capture NIR imagery. Fig. 7 shows some example images - the top row are examples of better quality images, and the bottom row contains examples of some of the possible challenges such as additional people in the background, glasses and closed eyes. Face verification on the acquired images was also achieved by using the Visage face verification tool.

Fig. 7: Example images of the 2D face captured on a smartphone in the drive-through scenario.

2.2 3D face

In the PROTECT multimodal dataset v2, the 3D face dataset contains face images captured with light-field technology, using a Lytro ILLUM plenoptic camera. Light-field imaging records not only the intensity of light, but also the intensity of light rays for any direction in space. This information allows to estimate the scene depth map and to reconstruct a 3D model of the observed objects.

Face images were acquired in the two scenarios that were simulated for the data collection. For the walk-through corridor, volunteers were acquired while walking in the corridor and approaching the 3D face capture area. The volunteers were asked to walk through the corridor 5 times per variation (with and without eyeglasses) and 4 pictures at different distances subject-camera were acquired each time (see Fig. 8). For the drive-through scenario, volunteers were acquired while they were sitting in a stationary car. Four test images per face variation, including neutral, wearing glasses, smile, sad, angry, surprise, and eyes closed, were acquired (see Fig. 9). For both scenarios, 4 enrolment pictures in more controlled conditions (uniform background, frontal face pose) were collected.

The 3D face verification algorithm was developed within PROTECT. It exploits one of the light-field image representations, namely the multi-view

*https://visagetechnologies.com/

Fig. 8: Example images of the 3D face PROTECT dataset. Images collected in the walk-through scenario.

Fig. 9: Example images of the 3D face PROTECT dataset. Images collected in the drive-through scenario.

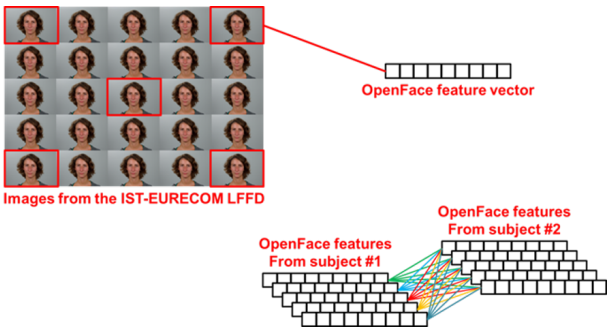

Fig. 10: Illustrated representation of the 3D face verification algorithm. OpenFace features are extracted from 5 face images selected from the multi-view representation. When two subjects are compared, cross comparison of the two set of features is performed.

or sub-aperture image (see Fig. 10). An array of 5×5 views has been extracted with regular angular sampling in the perspective range $[-0.5, 0.5]$. The tool used for processing the light-images, namely the Lytro power tools beta, allows sampling in the range $[-1.0, 1.0]$ but large perspective changes can result in artifacts. Five views, those with highest disparity between each other: corner images and central images, are selected and preprocessed separately. The face is located, aligned, and cropped with a pretrained model based on Histogram of Oriented Gradients (HOG) features available in the open source library DLIB*. OpenFace [6] features are then extracted from each view. When two faces are compared, cross distances among all views of the two light-field images are computed. The minimum distance is then selected as final comparison score. The reader is referred to [7] for a more detailed description of the technique and to [8] for a survey on light fields for face analysis.

2.3 Thermal face

A thermal infrared image of the human face has its unique heat-signature which can be used as a pattern for recognition. Thermal face verification does not require specific illumination and is robust to uneven lighting. This makes this modality particularly suitable for low-light acquisition. However, the use of infrared images for automatic face verification is not free of challenges. Thermal face imaging is sensitive to the emotional, physical and health condition of the subject. Moreover, the thermal features of the face depend on temperature of body, environment, and occlusions present on the face such as scarfs, hairs, facial hairs, glasses, or any disguise accessories that alter the emitted heat pattern.

Thermal face images were acquired in the drive-through scenario from volunteers sitting in a stationary car. The pictures feature different expressions to map the intra-class variability of the face. The collected dataset contains 28 subjects with 8 face images per subject. In case of a subject wearing eyeglasses, two sets of 8 images each (glasses/no glasses) have been collected for them. Examples of thermal face images are shown in Fig. 11. The resolution of the images is 640×512 pixels with a thermal resolution of 50 mK. The set of 8 images includes only images with frontal pose. Three images are with neutral facial expression, 4 images include emotions: smile, sad, angry and surprise; one picture with eyes closed. Several participants took part in a second image acquisition sessions following again the same protocol as described above.

Fig. 11: Examples thermal face images acquired in the drive-through scenario.

The verification algorithm relies on Support Vector Machine (SVM) and binary features. Face

*http://dlib.net/

detection is performed using customized Faster-RCNN algorithm. It is the Region Proposal Network (RPN) that shares full-image convolutional features with the detection network [9]. For face alignment Supervised Descent Method (SDM) with SIFT features [10] have been applied. SDM has been used for minimizing a Non-linear Least Squares (NLS) function and facial landmarks detection. The workflow of the thermal face verification system is illustrated in Fig. 12.

Since one of the biggest challenges for thermal face recognition systems is to reduce the fluctuations of temperature of the subject's face and environment, it is required to process the image to remove the fluctuations of pixel intensity. The Difference of Gaussians (DoG) filtering [11] has been applied, which is a common technique to remove illumination variations for visible face recognition. It is a band-pass filter that involves the subtraction of two Gaussians. As feature descriptor, multi-block local binary pattern (MB-LBP) was used [12]. MB-LBP achieved encouraging performance and due to modifications introduced to the original LBP method can capture large scale structures that may be the dominant features of images. The adopted DML scheme compares feature vectors using distance metrics. The extracted MB LBP feature vectors are classified using distance metrics to finally perform pairwise identity verification. The pairwise identity verification is proposed as a binary classification task performed with SVM. SVM, as a model employing kernel trick is well fitted for datasets with limited volumes. To provide a consistent distance learning comparison, binary SVM classifier was trained on values of Euclidean metrics between feature vectors.

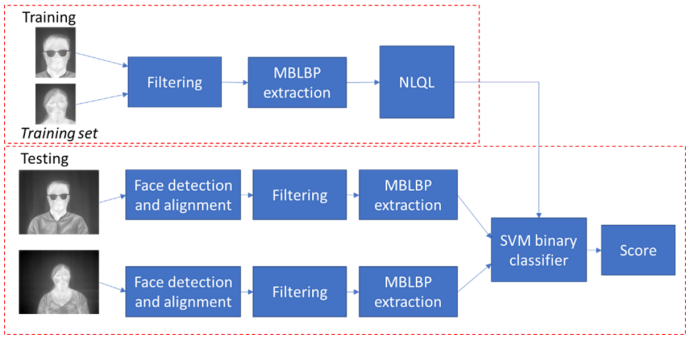

Fig. 12: Workflow of the thermal face verification system.

2.4 Periocular

Identifying people using the region around their eyes, known as periocular verification, has become an active research topic in biometrics. Background research has been conducted to prove the importance and strength of using the periocular region for biometric identification tasks [13], in particular for situations where the facial region is largely covered or occluded, and long-distance iris capture fails. For these reasons, as a novel biometric trait, periocular has primarily been considered for combination with face [14][15] and iris recognition [16][17] to enhance overall recognition accuracy.

Fig. 13: Example images of the periocular region extracted from the face frames in the walk-through scenario.

To obtain the periocular images, processing is first done on the face images. Thus, the same data collected in the walk-though scenario are used as for 2D face (see section 2.1). For each face image, cropping is implemented to get the two periocular regions: left eye and right eye regions. Fig. 13 shows a few example images from both VIS and NIR cameras of the extracted periocular regions, the top row showing better quality examples, and the bottom row showing more challenging cases where the image quality is affected by motion blur, relfection, eyewear and closed eye. Local Binary Patterns (LBP) [18] have been widely applied in biometric recognition. The literature describes variations of LBP used for face recognition, presentation attack detection, and iris recognition, due to their discriminative power for finding fine details on human skin, and their computational efficiency. Thus, LBP is chosen in this work to create a descriptor of the 2D periocular image. The same approach is used for both VIS and NIR periocular verification.

Table 2 List of algorithm parameters used in PROTECT evaluation.

Biometric trait	Algorithm parameters
2D Face	Version of the software: Visage SDK v8.6b1.
3D Face	Light-field image processing: Lytro power tools Beta; View extraction: array of 5x5 views extracted with regular angular sampling in the perspective range [-0.5, 0.5]; Number of views: 5 (corner images and central image); Face detection: DLIB (with DLIB's shape predictor with 68 face landmarks); Feature extraction: OpenFace (with DLIB's face recognition ResNet model v1); Distance: Euclidean; Verification threshold: 0.6.
Thermal Face	**Histograms of Multi-Block LBP:** Size of patch: 20; Number of blocks: 9; Scale: 3; Normalization: L2; Kernel in x and y directions for computing the gradient: 1. **Difference of Gaussian:** Standard deviation of the smaller Gaussian: 1; Standard deviation of the larger Gaussian: 2; Normalization: truncation of histograms ends; normalization to the 8-bit interval.
Iris	Software: IrisSDK by IriTech.
Periocular	**Preprocessing:** CLAHE: ClipLimit = 3.0. **LBP parameters:** NumNeighbours = 16; Radius = 2.
Finger & hand veins	**F1, F2, F3:** Preprocessing: Zhao09 (D0 =0.01, a = 0.5, b = 0.8, n = 2.0), CLAHE (ClipLimit = 0.001), Zhang09 (GaborBandwidth = 1.14, GaborSigma = 7), Resize (ResizeFactor = 0.5); Feature Extraction: MaximumCurvature (Sigma = 2.5); Comparison: MiuraMatcher (cr = 6, ch = 30, cw = 80). **F4:** Preprocessing: CLAHE (ClipLimit = 0.01), Zhang09 (GaborBandwidth = 1.12, GaborSigma = 3); Feature Extraction: MaximumCurvature (Sigma = 4); Comparison: MiuraMatcher (cr = 3, ch = 30, cw = 80). **G2, G7, G8:** Preprocessing: Resize (ResizeFactor = 0.5); Feature Extraction: MaximumCurvature (Sigma = 3.5); Comparison: MiuraMatcher (cr = 15, ch = 30, cw = 45). **G1, G3, G4, G5, G6:** Preprocessing: Zhao09 (D0 = 0.01, a = 0.5, b = 0.8, n = 2.0), Zhang09 (GaborBandwidth = 1.14, GaborSigma = 7), Resize (ResizeFactor = 0.5); Feature Extraction: SIFT (WindowSize = 2, EdgeThresh = 10, Levels = 3, FilterWidth = 15, MinNumberFeatures = 150); Comparison: SIFT (ScoreCalculation = Ratio, RatioThreshold = 2.0). All parameters and methods are according to the PLUS OpenVein Toolkit (http://wavelab.at/sources/OpenVein-Toolkit/), see reference: Kauba and Uhl (2020) [19]. The mapping for the finger vein sensor is: F1: Trans-illumination LASER, palmar; F2: Trans-illumination LASER, dorsal; F3: Trans-illumination LED, palmar; F4: Trans-illumination LED, dorsal. The mapping for the hand vein sensor is: G1: Trans-illumination, palmar; G2: Trans-illumination, dorsal; G3: Reflected Light 850 nm, palmar; G4: Reflected Light 850 nm, dorsal; G5: Reflected Light 950 nm, palmar; G6: Reflected Light 950 nm, dorsal; G7: Mobile Hand Vein AddOn (Nexus 5), palmar; G8: Mobile Hand Vein AddOn (Nexus 5), dorsal.
Anthropometrics	Software: ITTI's anthropometric verification system; Classifier: Artificial Neural Network; Number of cameras: 3; Distance between cameras: 2.5 meters.

2.5 Iris

Iris images were captured for the drive-through scenario using an IriShield^{TM}MK212OU device, which is a commercially available ultra-compact auto-capture iris scanner. The device is connected to a smartphone using a micro-USB cable, and is

operated via a dedicated app on the phone. The sensor works at an optimal distance of $5cm$ from the image sensor and has a focal depth of $6mm$, and includes a built-in NIR LED illuminator allowing it to work both in- and outdoors. The image format is 640 x 480 pixels in 8-bit grayscale, and is compliant to the ISO Standard 19794-6 (2005 & 2011). The iris images captured for the dataset were collected outdoor while the users were sitting in a vehicle, under a variety of conditions. Fig. 14 presents a few example iris images captured using the device. The top row demonstrates examples of good quality captures, whereas the bottom row contains more challenging factors such as contact lenses, partial occlusion, out-of-focus, and glasses with reflection. As for 2D face, in PROTECT iris verification was not a research topic. The selected software for iris verification was a commercial solution, namely the IrisSDK by IriTech*. However, iris was dropped from the evaluation after a biometric selection was performed to retain only the most suitable traits for the final demonstrators.

Fig. 14: Example iris images captured using the handheld device in the drive-through scenario.

2.6 Vascular Biometrics

Vascular pattern based biometric systems, commonly denoted as vein biometrics, offer several advantages over other well-established biometric recognition systems. In particular, hand and finger vein systems have become a serious alternative to fingerprint based ones for several applications. Vein based systems use the structure of the blood vessels inside the human body, which becomes visible under near-infrared (NIR) light. As the vein structure is located inside the human body, it is resistant to abrasion and external influences on the skin. Furthermore, a lifeness detection to detect presentation attacks can be performed easily [20].

In the course of the data acquisition, hand and finger veins were acquired using custom built capturing devices in the drive-through scenario. Finger veins have been acquired using the LED and laser version of the sensor presented in [21], hand veins using a stationary capturing device and a mobile phone, respectively. The stationary hand vein sensor [22] provides three different illumination scenarios: reflected light using 850nm and 950nm LEDs and a 16x16 NIR LED panel for transillumination. The mobile phone in use is a modified Nexus 5 smart phone where the NIR blocking filter was removed. For sufficient NIR light, a custom build illumination add-on [23] was added to the phone. Both, hand and finger veins, have been acquired from the palmar and dorsal view.

The acquisition was done in two timely separated sessions: (1) An outdoor session where the volunteers sit in a stationary car and (2) an indoor acquisition under controlled illumination conditions. The first session should simulate the drive-through scenario under real-world acquisition conditions, while the second one serves as a reference data set. In both sessions, all subjects have been acquired using all sensors and illumination scenarios. Both, hand and finger veins, have been acquired from the palmar and dorsal view. For each modality/sensor variation 5 samples have been taken. This sums up to 20 finger vein images (2 sensors x 2 perspectives x 5 samples) and 40 hand vein images (4 sensors x 2 perspectives x 5 samples) per subject per session.

The recognition tool-chain for evaluating the acquired vein images consists of the following components: (1a) For *finger region detection* and *finger alignment* an implementation that is based on [24] is used. (1b) The *ROI extraction* for hand vein images is done manually by fitting a rectangular ROI is fit inside the hand area. The ROI images have a size of $512x512$ pixels. (2) To improve the visibility of the vein pattern *High Frequency Emphasis Filtering* (HFE) [25], *Circular Gabor Filter* (CGF) [26] and simple *CLAHE* (local histogram equalisation) [27] are used during *preprocessing*. (3a) For the simple vein pattern based feature methods, MC and PC, the binary feature images are compared using a correlation measure, calculated between the input images and in x- and y-direction shifted and rotated versions of the reference image as described in [28]. and (3b) the SIFT based approach as described in [29], respectively.

*https://www.iritech.com/products/software/irissdk-eye-recognition-software

2.7 Anthropometrics

Anthropometry refers to the measurement of the human individual and involves the systematic measurement of the physical properties of the human body, primarily dimensional descriptors of body size and shape. Anthropometrics is one of the novel biometric traits investigated in PROTECT and consists in the measurement of the human body as well as the analysis of human gait. Of course, anthropometrics is very suitable for the walk-through scenario.

For anthropometric verification, the system developed by ITTI* was adopted. The method consists in collecting gait and anthropometric information by using a set of Microsoft Kinect devices. The Kinect SDK provides functionality for accurate person tracking from about 1.5m to up to 4.5m. Gait collected from 3m is vulnerable to random noise and presentation attacks. In order to extend the collection distance, several Kinects (3 for this collection) with preprocessing platforms have been used to construct a network of sensors that is able to cover the entire Biometric Capture Area.

The anthropometric verification system uses an artificial neural network as a classifier. The network can assess similarity between two recorded patterns (templates, feature vectors). As a first step of the training process, the collected data have been transformed into feature vectors (one vector for each subject and for each sample). Subsequently, the acquired feature vectors were paired with each other in order to produce training batches. The resulting dataset was divided into test and validation sets. The test dataset included about 60% of all subjects while the validation dataset contained the remaining 40%. The neural network model has been trained and validated based on these two datasets.

3 Multibiometric evaluation

In order to perform an evaluation able to correctly model the inter-class and intra-class variability, a data augmentation technique is tested and reported in this Section. The number of comparisons on which each biometric modality is tested can vary considerably (e.g. 128160 for 3D face in the walk-through scenario versus 7440 comparisons for hand veins). This is due to various factors, including the complexity of acquiring certain biometric traits compared to others, and the necessity of including more variations for certain biometric traits (e.g. face has a large intra-class variability that must be modelled). For multimodal evaluation, two problems had to be addressed: (i) different number of comparisons per modality; and (ii) missing scores.

Regarding (i), let s_1 and s_2 be two subjects to be compared. For both subjects, $E_h = 4$ and $T_h = 4$ hand vein samples and $E_f = 4$ and $T_f = 20$ face samples were collected, where E and T indicate the enrolment and test samples, respectively. For hand veins, there will be $E_h \times T_h = 4 \times 4 = 16$ comparisons, while for face $E_f \times T_f = 4 \times 20 = 80$ comparisons. To perform multibiometric evaluation, the 16 comparison scores from hand veins would be matched with only 16 comparison scores from face and the 64 remaining scores would not be used. Also, the total amount of comparisons would be limited by the lower bound defined by the biometric modality with less comparison scores, leading to less reliable results. Thus, it has been decided to overcome problem (i) by artificially creating independent observations for the multimodal system by randomly selecting comparison scores from the available pool of scores from the different single modalities. Thus: for each comparison of s_i with s_j, where $i = 1, 2, \ldots, I$ is the number of enrolled subjects, and $j = 1, 2, \ldots, J$ is the number of test subjects, the list of comparison scores $L_k = comparison(s_i, s_j)$ from each biometric modality k is obtained. To create a multimodal observation, that is a tuple composed by the scores coming from the different modalities, a score is randomly selected from each single-modality list to represent the output of that modality. In this way, it is possible to obtain any number of multimodal observations. An illustrated simplified example is provided in Fig. 15.

However, for some comparisons of s_i with s_j, the list of comparison scores from one or more modalities may be empty. In this case we run into problem (ii) of missing scores. Problem (ii) is solved by using a technique for missing score estimation from nearest neighbours. This is discussed in Section 3.2.

Minmax normalization followed by simple sum of scores has been observed to result in reasonable improvement in matching accuracy of a multimodal biometric system [30]. This scheme is tested along with weighted sum. The reason why it has been

*https://www.itti.com.pl/en/home/

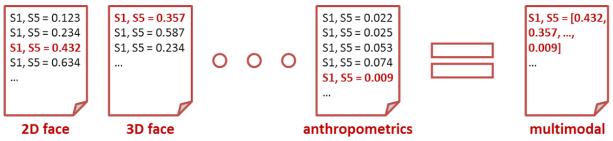

Fig. 15: Illustrated example of data augmentation.

decided not to use a training phase to learn the fusion weights is that the data are already limited in size (only 38 different subjects) and splitting the database in training, development, and test sets, would have led to poorly reliable performances.

Some of the biometric traits listed in Table 1 (e.g. iris, thermal face in the walking-through scenario) were not included in the following tests as, after data collection, a biometric selection was performed and only the most suitable traits for the final demonstrators were kept.

Table 3 Walk-through scenario: Comparison of the results obtained on the original data with the results obtained on the augmented comparison list.

3D face, walk-through scenario		
	Original	Augmented data
EER	0.0074	0.0075
FMR1000	0.0299	0.0353
ZeroFMR	0.1003	0.1102
Anthropometrics		
	Original	Augmented data
EER	0.0298	0.0247
FMR1000	0.1159	0.1037
ZeroFMR	0.3514	0.3375
2D face VIS		
	Original	Augmented data
EER	0.0179	0.0219
FMR1000	0.0518	0.0513
ZeroFMR	0.0956	0.0963
2D face NIR		
	Original	Augmented data
EER	0.0705	0.0752
FMR1000	N/A	N/A
ZeroFMR	N/A	N/A
Periocular VIS		
	Original	Augmented data
EER	0.4828	0.5450
FMR1000	N/A	N/A
ZeroFMR	N/A	N/A
Periocular NIR		
	Original	Augmented data
EER	0.3837	0.5613
FMR1000	N/A	N/A
ZeroFMR	N/A	N/A

3.1 Data augmentation

In this section, the performances obtained from data augmentation are first compared to the original non-augmented single-modality performances to demonstrate the reliability of the technique. The augmented distance matrices are then used for multimodal evaluation. The distance matrices are obtained by simulating a system with 10 enrolment samples and 40 test samples per subject, for a total of $(10 \times 38) \times (40 \times 38) = 380 \times 1520 = 577600$ comparisons. As can be seen from the values reported in Tables 3 and 5, the performances obtained on the augmented comparison lists are very similar to those obtained on the original number of comparisons. More details regarding the results are given in deliverable D7.4 of PROTECT*. Also, from Table 5 it can be noted that for finger and hand veins, two values are reported for the original comparison lists. This is due by the fact that left and right hand were evaluated separately. While computing the augmented comparison list for finger and hand veins, the scores were randomly selected from both left and right hand lists. The performance values obtained on the augmented list, as expected, are in the range defined by the two original score sets.

Since this technique implies random selection, for all the experiments employing data augmentation, performance assessment is carried out on 5 augmented matrices per modality and the results are then averaged.

3.1.1 Walk-through scenario performances: In this Section, the results obtained by multimodal fusion on the augmented data (577600 comparisons) are reported. For the walk-through scenario, the biometric traits involved are 2D face VIS, 2D face NIR, 3D face, periocular VIS, periocular NIR, and anthropometrics. For these tests, when a score is missing for one biometric trait, the entire tuple is removed from the score matrix, meaning that the corresponding (si, sj) comparison is removed. In this case, no scores were missing from the anthropometric system. After removing the tuples corresponding to the other traits missing scores, the final number of comparisons is 302400.

Fig. 16 reports the DET curves obtained from the different tests carried out and Table 4 reports the corresponding performance values. The first

* *http://projectprotect.eu/*

test consists in the fusion of all biometric traits ("ALL" in the figure). Among all the biometric traits in this scenario, periocular reported the worst performances in both visible and near-infrared illumination (as reported in Table 3). Thus, the walk-through system is further tested after removing periocular from multimodal fusion and the achieved performances are very good ("ANT, 3DF, 2DF VIS, 2DF NIR" in Table 4).

Then, all the combinations of three biometric traits are tested. The one achieving best performance is composed by ANT, 3DF, and 2DF VIS.

Finally, the combination of the two modules achieving best single-modality performances (ANT and 3DF), is tested by fusing the two systems via weighted sum. Three DET curves corresponding to the three weighted combinations achieving best performance are showed in Fig. 16. Several weight combinations were tested, with $w_1 = 0.1, 0.2, 0.3, ..., 0.9$ and $w_2 = 1 - w_1$.

Table 4 Multimodal fusion for the walk-through system: the results of different combinations of barometric traits are reported in terms of EER, FMR1000, and ZeroFMR.

	EER	FMR1000	ZeroFMR
ALL	0.0372	0.5069	0.9093
ANT, 3DF, 2DF VIS, 2DF NIR	0.0118	0.0248	0.0863
ANT, 2DF VIS, 2DF NIR	0.0151	0.0429	0.1676
3DF, 2DF VIS, 2DF NIR	0.0217	0.1195	0.2790
ANT, 3DF, 2DF NIR	0.0095	0.1079	0.4197
ANT, 3DF, 2DF VIS	0.0088	0.0224	**0.0339**
ANT*0.7 + 3DF*0.3	0.0076	**0.0117**	0.0932
ANT*0.9 + 3DF*0.1	0.0163	0.0204	0.0612
ANT*0.2 + 3DF*0.8	**0.0063**	0.0284	0.1127

3.1.2 Drive-through scenario performances: In this Section, the results obtained by multimodal fusion on the augmented data (577600 comparisons) are reported. For the drive-through scenario, the biometric traits involved are 3D face, finger veins, hand veins, and thermal face. Also for these tests, when a score is missing for one biometric trait, the entire tuple is removed from the score matrix. In this case, many scores were missing, in particular form finger and hand veins. This is due to the fact that part of the samples were collected outdoor in natural light and were then found to be not suitable for verification. In fact, the sunlight had a negative influence on the capturing device, leading to overexposed areas at outside regions of the image, where the index and ring finger are located (see Fig. 17 for a comparison between indoors and outdoors acquired images). The other problem was the unnatural and uncomfortable position which

Table 5 Drive-through scenario: Comparison of the results obtained on the original data with the results obtained on the augmented comparison list.

3D face, drive-through scenario		
	Original	Augmented data
EER	0.0328	0.0216
FMR1000	0.0384	0.0251
ZeroFMR	0.5830	0.5746
Thermal face		
	Original	Augmented data
EER	0.1052	0.0828
FMR1000	0.4752	0.3398
ZeroFMR	0.5555	0.4177
Finger veins		
	Original	Augmented data
EER	[0.0005, 0]	0.0003
FMR1000	[0.0042, 0]	0
ZeroFMR	[0.0085, 0]	0.0051
Hand veins		
	Original	Augmented data
EER	[0.1385, 0.1746]	0.1580
FMR1000	[0.2277, 0.3571]	0.2827
ZeroFMR	[0.2768, 0.4866]	0.5436

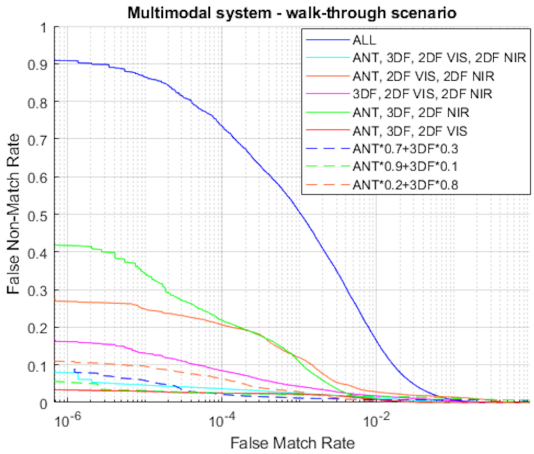

Fig. 16: DET curves for the walk-through system.

the data subjects had to hold during the capturing process. The back window of the car was situated rather high, so especially smaller people had to bend their arm and hand in order to reach the capturing device which was located outside the car window. This caused the hand not to lie flat on the sensor surface but tilted, which especially puts the index and ring finger in an unfavourable position for acquisition. Thus, the best performance as presented in Table 5 is given for both middle fingers only.

Regarding hand veins, the main cause of the missing scores is again the combined evaluation of the outdoor and indoor session. Same as for the finger vein sensor, the unfavourable position and

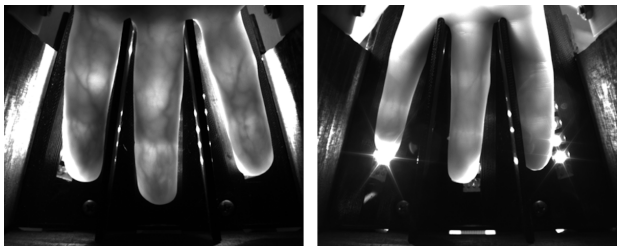

Fig. 17: Comparison of indoors (left) and outdoors (right) acquired finger vein images.

Fig. 18: Comparison of indoors (left) and outdoors (right) acquired hand vein images.

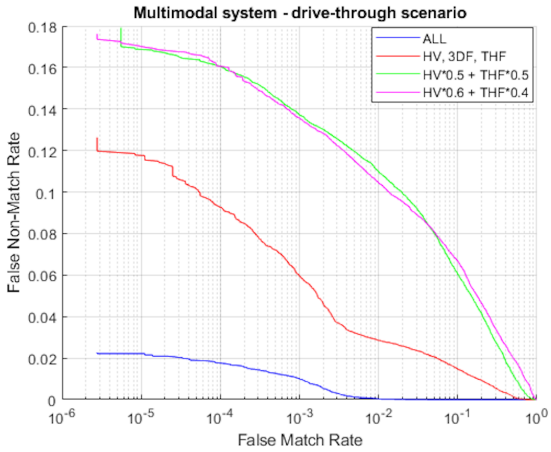

Fig. 19: DET curves for the drive-through system.

pose the data subjects had to maintain during the acquisition process leads to all kinds of pose variations rather than hands that are positioned flat on the sensor surface (see Fig. 18 for a comparison of an indoor and outdoor sample).

After removing the tuples corresponding to missing scores, the final number of comparisons is 75600. The DET curves obtained by fusing the four modalities using Minmax normalization and simple sum are illustrated in Fig. 19.

In this multibiometric system, finger veins have way better performances compared to the other biometric traits, thus the best performance are obtained by giving weighs equal to zero to all other biometric traits. For this reason, the case in which the answer from the finger vein module is missing is analysed in more detail. Fig. 19 reports the

Table 6 Multimodal fusion for the drive-through system: the results of different combinations of barometric traits are reported in terms of EER, FMR1000, and ZeroFMR.

	EER	FMR1000	ZeroFMR
ALL	**0.0030**	**0.0102**	**0.0227**
HV, 3DF, THF	0.0248	0.0596	0.1268
HV*0.5 + THF*0.5	0.0705	0.1378	0.1792
HV*0.6 + THF*0.4	0.0728	0.1357	0.1768

DET curve obtained by the fusion of hand veins (HV), 3D face (3DF), and thermal face (THF). The biometric trait obtaining best performance out of the three is further removed to assess the performance of the system when the answer from the 3D face module is missing too. The corresponding DET curves are showed in Fig. 19. The two DET curves corresponding to the two weighted combinations achieving best performance are showed in the figure. Several weight combinations were tested, with $w_1 = 0.1, 0.2, 0.3, ..., 0.9$ and $w_2 = 1 - w_1$.

The performance values corresponding to the DET curves in Fig. 19 are reported in Table 6.

3.2 Missing score imputation

As mentioned before, for some biometric traits in the PROTECT multimodal dataset v2 and for some comparisons of s_i with s_j, the score may be missing. Various factors can result in incomplete score vectors in multibiometric systems: (a) failure of a matcher to generate a score; (b) absence of a trait during acquisition; (c) sensor malfunction; or (d) during enrolment, all necessary biometric traits may not be available. In this section, the results obtained by using a technique for missing score replacement are reported. The technique is presented by Ding and Ross in [30]. Several techniques are tested by the authors of [30] on a multimodal database, but the one achieving best performance is the K nearest neighbour (KNN) imputation scheme. The idea is to predict the missing scores of a multimodal system from the k most similar (nearest) tuples. Let D be the score matrix:

$$D = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,p} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,p} \end{pmatrix}$$

Where $x_{i,j}$ denotes the comparison score from the j^{th} modality of the i^{th} user. Euclidean distance

Fig. 20: Illustrated example of missing score imputation.

d is considered to find the KNN:

$$d(x_i, x_j) = \sum_{h \in O_i \cap O_j} (x_{i,h} + x_{j,h})^2$$

Where $O_i = \{h|$ the h^{th} variable of the i^{th} observation is observed $\}$. In other words, only the mutually observed variables are used to calculate the distance between observations. The KNN imputation scheme is described as follows: (i) For each observation x_i, apply the distance function d to find the k nearest neighbour vectors in D; (ii) The missing variables x_i^{miss} are imputed by the average of the corresponding variables from those k-nearest neighbours. An illustrated simplified example is provided in Fig. 20.

KNN imputation does not require the creation of a predictive model for each variable, and so it can easily treat instances with multiple missing values. This is the case for the score matrices used for these experiments, for some instances, multiple values may be missing. The value of k is set to 5 as suggested in [30]. The choice of a small k may produce a deterioration in the performance of the classifier due to overemphasis in a few dominant instances in the estimation process of missing values. On the other hand, a neighbourhood of large size would include instances that are significantly different from the instance containing the missing values [30].

KKN imputation is then used to replace missing values in the data-augmented score matrices. Thus, all the following results are obtained on 577600 comparisons. Biometric fusion for both scenarios is performed using Minmax normalization and simple sum of the comparison scores.

3.2.1 Walk-through scenario performances: Two tests were carried out. The first test uses the original score distribution of the systems. As can be seen from Fig. 21, even if the six distributions are normalized in the range $[0, 1]$, the anthropometric scores are concentrated near 0. The performance obtained by fusing the original distributions is reported in Table 7.

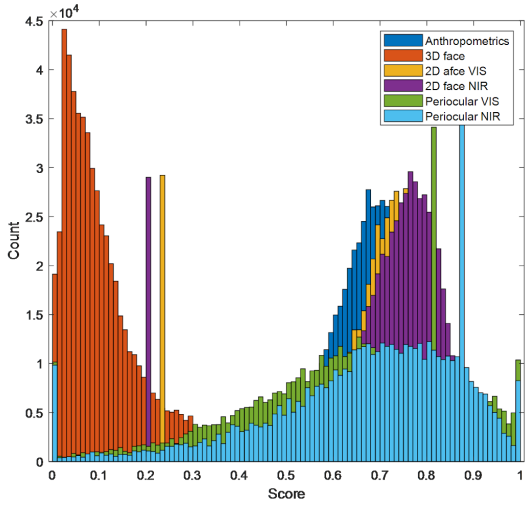

Fig. 21: Original score distributions for the modalities in the walk-through scenario.

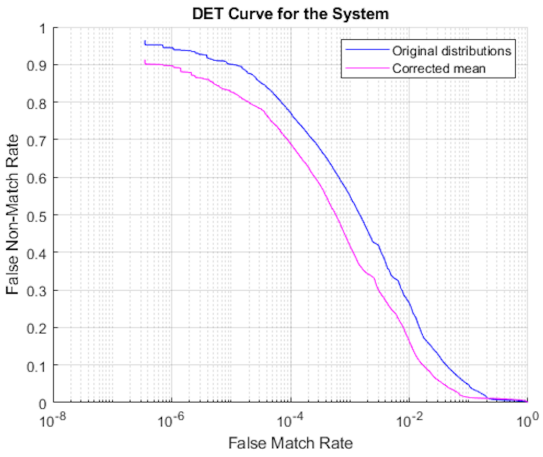

Fig. 22: DET curves for the walk-through scenario after missing score imputation.

An improvement in performance is obtained by the second test where the scores from each system are multiplied by a value m to make the distributions having mean equal to 0.5:

$$m = \frac{0.5}{\mu_i}$$

Where μ_i is the mean of the i^{th} score distribution and $i = 1, 2, 3, \cdots, I$, where I is the number of biometric modalities, 6 in this case. The improved performance values are reported in Table 7.

3.2.2 Drive-through scenario performances: The same tests are repeated for the drive-through scenario. The first test uses the original score distribution of the finger vein, hand vein, 3D face, and thermal face systems. Fig. 23 shows the original score distributions. For the second tests all distributions are

Table 7 Multimodal fusion for the walk-through system: the results are obtained after applying missing score imputation and for two cases: (i) original score distributions; (ii) all distributions have the same mean.

	EER	FMR1000	ZeroFMR
Original distributions	0.0678	0.5516	0.9651
Corrected mean	0.0433	0.4217	0.9148

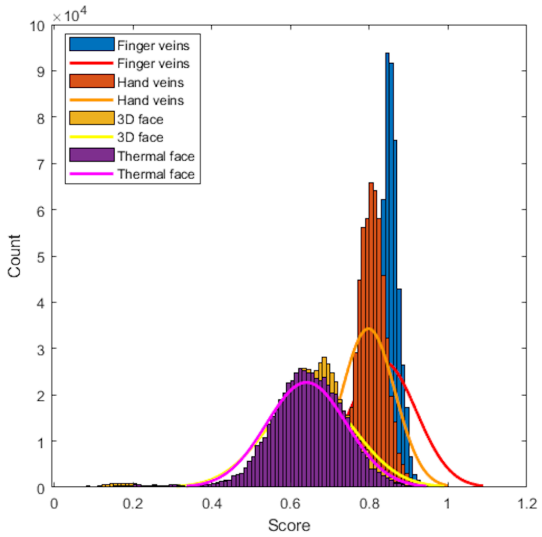

Fig. 23: Original score distribution for the modalities in the drive-through scenario.

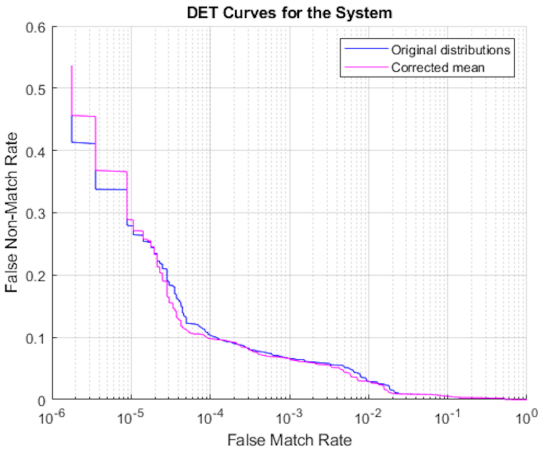

Fig. 24: DET curves for the drive-through scenario after missing score imputation.

multiplied for the value m, computed as showed before, so that all distributions have mean equal to 0.5.

Fig. 24 shows the DET curves for the two tests. The corresponding performance values are reported in Table 8.

Compared to the performances obtained in Section 3.1.2 on the augmented comparison list after removing the missing scores, the results obtained after missing score imputation are much worse. This is probably due to the fact that the percentage of missing scores in this case is too high to reliably use this technique.

Table 8 Multimodal fusion for the drive-through system: the results are obtained after applying missing score imputation and for two cases: (i) original score distributions; (ii) all distributions have the same mean.

	EER	FMR1000	ZeroFMR
Original distributions	0.0171	0.0664	0.4882
Corrected mean	0.0158	0.0664	0.5387

4 Conclusions

In this paper the case study of the PROTECT multibiometric system has been presented. Apart from presenting the system and its performance results, this article is intended to show what problems may be encountered when evaluating a multimodal system, from data collection to performance assessment, and how to face them. The issues addressed in PROTECT include (i) the collection of a non-chimeric multimodal biometric database including 10 different modalities (please note that for example 2D face in VIS and NIR light are considered as two different modalities); (ii) the problem of having different score distributions, with different size and shape, to be fused in a multibiometric system; and (iii) the problem of missing scores. The presented case study shows how to overcome these problems using techniques for data augmentation and missing score imputations. This is done in order to obtain a set of scores of sufficient size to perform performance assessment according to the guidelines provided by international standards. In particular, the technique used for data augmentation resulted in helping achieving more reliable performance assessment.

Acknowledgements

This research was funded by EU H2020 Research and Innovation Programme under grant agreement no. 700259 (PROTECT).

5 References

1 A. K. Jain, A. Ross, and S. Prabhakar. An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):4–20, Jan 2004.

2 International Civil Aviation Organization (ICAO). Machine readable travel documents. *Doc 9303*, 7th edition, 2015.

3 Anil K. Jain and Arun Ross. Multibiometric systems. *Commun. ACM*, 47(1):34–40, January 2004.

4 International Electrotechnical Commission (IEC) International Organization for Standardization (ISO). Iso/iec 19795 information technology — biometric performance testing and reporting.

5 A. F. Sequeira, L. Chen, and J. Ferryman et al. Protect multimodal db: fusion evaluation on a novel multimodal biometrics dataset envisaging border control. In *2018 International Conference of the Biometrics Special Interest Group (BIOSIG)*, pages 1–5, Sep. 2018.

6 Brandon Amos, Bartosz Ludwiczuk, and Mahadev Satyanarayanan. Openface: A general-purpose face recognition library with mobile applications. 2016.

7 V. Chiesa and J. Dugelay. On multi-view face recognition using lytro images. In *2018 26th European Signal Processing Conference (EUSIPCO)*, pages 2250–2254, Sep. 2018.

8 Chiara Galdi, Valeria Chiesa, Christoph Busch, Paulo Lobato Correia, Jean-Luc Dugelay, and Christine Guillemot. Light fields for face analysis. *Sensors*, 19(12), 2019.

9 Shaoqing Ren, Kaiming He, Ross B. Girshick, and Jian Sun. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), 2016.

10 Xuehan Xiong and Fernando De la Torre. Supervised descent method and its applications to face alignment. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 532–539, 2013.

11 Lucas Assirati, Núbia Rosa da Silva, Lilian Berton, Alneu de A Lopes, and Odemir M Bruno. Performing edge detection by difference of gaussians using q-gaussian kernels. In *Journal of Physics: Conference Series*, volume 490, page 012020. IOP Publishing, 2014.

12 Lun Zhang, Rufeng Chu, Shiming Xiang, Shengcai Liao, and Stan Z Li. Face detection based on multi-block lbp representation. In *International conference on biometrics*, pages 11–18. Springer, 2007.

13 Unsang Park, Arun Ross, and Anil K Jain. Periocular biometrics in the visible spectrum: A feasibility study. In *2009 IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems*, pages 1–6. IEEE, 2009.

14 Khary Popplewell, Aniesha Alford, Gerry V Dozier, Kelvin S Bryant, John C Kelly, Joshua Adams, Tamirat Abegaz, Kamilah Purrington, and Joseph Shelton. A comparison of genetic feature selection and weighting techniques for multibiometric recognition. In *ACM Southeast Regional Conference*, pages 205–208, 2011.

15 Ramachandra Raghavendra and Christoph Busch. Robust 2d/3d face mask presentation attack detection scheme by exploring multiple features and comparison score level fusion. In *17th International Conference on Information Fusion (FUSION)*, pages 1–7. IEEE, 2014.

16 Chun-Wei Tan and Ajay Kumar. Towards online iris and periocular recognition under relaxed imaging constraints. *IEEE Transactions on Image Processing*, 22(10):3751–3765, 2013.

17 Kiran B Raja, Ramachandra Raghavendra, and Christoph Busch. Binarized statistical features for improved iris and periocular recognition in visible spectrum. In *2nd International Workshop on Biometrics and Forensics*, pages 1–6. IEEE, 2014.

18 Timo Ojala, Matti Pietikäinen, and David Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, 29(1):51–59, 1996.

19 Christof Kauba and Andreas Uhl. *An Available Open-Source Vein Recognition Framework*, pages 113–142. Springer International Publishing, Cham, 2020.

20 Ajay Kumar and Yingbo Zhou. Human identification using finger images. *Image Processing, IEEE Transactions on*, 21(4):2228–2244, 2012.

21 Christof Kauba, Bernhard Prommegger, and Andreas Uhl. Openvein - an open-source modular multipurpose finger vein scanner design. In Andreas Uhl, Christoph Busch, Sebastien Marcel, and Raymond Veldhuis, editors, *Handbook of Vascular Biometrics*, chapter 3, pages 77–111. Springer Nature Switzerland AG, Cham, Switzerland, 2019.

22 Christof Kauba and Andreas Uhl. Shedding light on the veins - reflected light or transillumination in hand-vein recognition. In *Proceedings of the 11th IAPR/IEEE International Conference on Biometrics (ICB'18)*, pages 1–8, Gold Coast, Queensland, Australia, 2018.

23 Luca Debiasi, Christof Kauba, Bernhard Prommegger, and Andreas Uhl. Near-infrared illumination add-on for mobile hand-vein acquisition. In *2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pages 1–9, Los Angeles, California, USA, 2018.

24 Yu Lu, Shan Juan Xie, Sook Yoon, Jucheng Yang, and Dong Sun Park. Robust finger vein roi localization based on flexible segmentation. *Sensors*, 13(11):14339–14366, 2013.

25 Jianjun Zhao, Hogliang Tian, Weixing Xu, and Xin Li. A new approach to hand vein image enhancement. In *Intelligent Computation Technology and Automation, 2009. ICICTA'09. Second International Conference on*, volume 1, pages 499–501. IEEE, 2009.

26 Jing Zhang and Jinfeng Yang. Finger-vein image enhancement based on combination of gray-level grouping and circular gabor filter. In *Information Engineering and Computer Science, 2009. ICIECS 2009. International Conference on*, pages 1–4. IEEE, 2009.

27 K. Zuiderveld. Contrast limited adaptive histogram equalization. In Paul S. Heckbert, editor, *Graphics Gems IV*, pages 474–485. Morgan Kaufmann, 1994.

28 Naoto Miura, Akio Nagasaka, and Takafumi Miyatake. Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification. *Machine Vision and Applications*, 15(4):194–203, 2004.

29 Christof Kauba, Jakob Reissig, and Andreas Uhl. Pre-processing cascades and fusion in finger vein recognition. In *Proceedings of the International Conference of the Biometrics Special Interest Group (BIOSIG'14)*, Darmstadt, Germany, sep 2014.

30 Yaohui Ding and Arun Ross. A comparison of imputation methods for handling missing scores in biometric fusion. *Pattern Recognition*, 45(3):919–933, 2012.