Vascular Biometrics Experiments on Candy – A New Contactless Finger-Vein Dataset*

Sushil Bhattacharjee, David Geissbuehler, Guillaume Clivaz, Ketan Kotwal, and Sebastien Marcel

Idiap Research Institute Rue Marconi 19, 1920 Martigny, Switzerland

Abstract. The topic of finger-vein (FV) biometrics is an active and growing topic of research. Most FV systems available today rely on contact sensors that capture vein patterns of a single finger at a time. We have recently completed a project aimed at designing a contactless vein sensing platform, named *sweet*. In this paper we present a new FV dataset collected using *sweet*. The dataset includes multiple FV samples from 120 subjects, and 280 presentation attack instruments (PAI), captured in a contactless manner. Further, we present baseline FV authentication (FVA) results achieved for proposed dataset. The *sweet* platform is equipped to capture a sequence of images suitable for photometric-stereo (PS) reconstruction of 3D surfaces. We present a FV presentation attack detection (PAD) method based on PS reconstruction, and the corresponding baseline FV PAD results on the proposed dataset.[1]

1 Introduction

Vascular biometrics offers several advantages over biometrics, such as high accuracy, robustness to variations in ambient illumination and robustness to presentation attacks (PA). Finger-vein (FV) sensors are the most common type of vascular-biometrics devices. Most current FV sensors commercially available use transmissive near-infrared (NIR) illumination – the camera captures the light passing through the finger[2]. Currently available FV sensors mostly rely on contact-imaging. Contact-imaging for biometric sensors, however, may be undesirable in certain deployment scenarios where hygiene is critical, such as hospitals, or even commercial points of sale.

We have developed a platform for contactless hand-vascular biometrics. The open-source design of this hardware platform, named *sweet*, is presented in a separate paper [3]. The prototype *sweet* platform is shown in Figure 1(a). *sweet* captures biometrics samples under white LED light using a color (RGB) camera, and two NIR wavelengths, 850nm and 950nm, using two NIR cameras. The NIR cameras record vascular data. Each NIR camera also records data for Photometric-Stereo (PS) reconstruction of the target-surface, by capturing a set of images under a controlled illumination-sequence. This PS capability has been included for detecting presentation attacks (PA).

* This work was supported by the Innosuisse project Candy, and the Swiss Center for Biometrics Research and Testing.

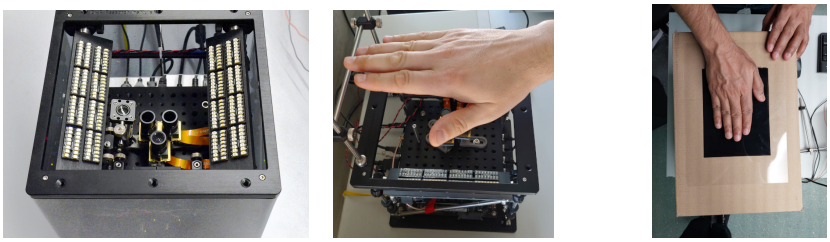
[1] The proposed dataset, python code for our baseline methods for FVA and FV PAD, and our experimental protocols will be shared publicly, under license, upon publication of this paper. Information for downloading these resources will appear in place of this footnote.

[2] Reflective illumination has been commonly adopted for palm-vein recognition devices, but not for FV sensors.

In this work we present a new dataset, named Candy_v3, collected using the *sweet* platform. This dataset includes several FV samples per subject from a gender-balanced test crew of 120 subjects. The dataset contains several kinds of FV imagery, and could support research on a variety of questions related to FV biometrics. In this work we have described a baseline FV authentication (FVA) method, and have provided results of this method on the new dataset.

The proposed dataset also includes PAs made using PA instruments (PAI) of 14 PAI species. We also present a FV PAD method that utilizes the surface-normal map reconstructed using PS.footnote 1 The main contributions of this work are (i) a new FV dataset containing 1200 genuine FV samples and 1400 PA samples from 14 PAI species; (ii) baseline FVA performance using a well-understood FVA pipeline; and (iii) a new baseline FV PAD method based on PS, accompanied by its performance on the new dataset.

We start, in Section 2, with a brief review of related research. The new dataset, including the data collection process, is described in Section 3. The baseline FVA method, and the PS based FV PAD method are described in Section 4. Then, the experimental methodology and results for both FVA and FV PAD are discussed in Section 5. A summary of this work and some perspectives on future work are presented in Section 6.



(a) *sweet* platform (b) *Bona fide* presentation (c) PA using a PAI of species B4

Fig. 1: (a) Prototype of the *sweet* platform. (b) Contactless *bona fide* presentation. (c) Apparatus for recording certain species of presentation attacks.

2 Related Research

For a comprehensive overview of FV biometrics, we point the interested reader to a recent review of the subject compiled by Hou *et al*. [4]. In this section we discuss selected works related to FVA and FV PAD that are relevant to the present study.

Finger-Vein Authentication Systems: Newly proposed FVA algorithms are compared to the state of the art using publicly available datasets such as SDUMLA-HMT [21], MMCBNU_6000 [8], VERA-finger [18], UTFVP [17], and SCUT-SFVD [12]. One common characteristic of these datasets is that the biometric samples represent only single fingers. In contrast, FV samples in the presented dataset show four fingers together, which enables finger-fusion for more robust FVA and FV PAD.

Mainstream vascular biometrics systems still rely on hand-crafted features such as Repeated Line-Tracking (RLT) [9], maximum curvature (MC) [10], wide-line detection (WLD) [5]. These algorithms extract binary pixel-maps representing the vein-network

in the biometric sample allowing to compare them. In this work we have used MC features.

Frequency-domain methods for FVA has also been proposed. Yang *et al*. [20] have used a bank of Gabor filters to enhance veins at different scales and then construct a set of FVCodes that are compared using a Cosine-similarity function. Yang *et al*. claim that their method performs better than MC features [10]. These results, however, have been estimated over a proprietary, unpublished dataset. More recently, Kovač and Marák [7] have used Gabor filters to detect feature-points in vein-images.

Publicly available FV datasets are not large enough to train a convolutional neural network (CNN) from scratch. Up to now, deep-learning based FVA approaches have adapted pre-trained CNNs through transfer-learning on FV datasets to construct feature-extractors. Besides actual FVA, deep-learning based methods have also been used for other purposes such as vein enhancement, vein segmentation, and even encryption (see [22]). Bros *et al.* [2] propose a Residual Convolutional Autoencoder (RCAE) for vein-enhancement that reduces the classification error on the UTVFP dataset from 2.1% to 1%. In the present work we have used this RCAE in our FVA pipeline as well.

Research on FV PAD: Since the first comprehensive review of FV PAD methods [13], several new FV PAD methods have been proposed. Qui *et al*. [12] use total variational (TV) decomposition to decompose a FV image into structure and noise components. The two components are then modeled separately using local binary patterns (LBP), and the LBP histograms are subsequently classified using a cascade-classifier. This method (TV-LBP) achieves nearly perfect PAD on two public datasets – IDIAP VERA-finger [18] and SCUT-FVD [12].

Singh *et al*. [16] have used the SfSNet [15] to reconstruct a surface-normal map as well as a diffusion-map from a single (transmissive) image. The key difference between their work and our proposed FV-PAD is that in our case the surface-normal map is computed from photometry-data collected explicitly for this purpose (PS), whereas Singh *et al*. decompose a single image into a surface-normal map and a diffusion map using a pre-trained SfSNet. Surface-normal maps reconstructed using SfSNet do not always reflect the reality.

Raghavendra *et al*. [14] have used transfer learning to adapt a pre-trained Alex-Net for FV PAD. They demonstrate that their method consistently achieves a BPCER of 0 on several public datasets. The APCER for their method varies from 0.5% to 3.5%, depending on the species of presentation attack instrument (PAI) used.

From the analysis of related works presented here we see that the idea of using Photometric-Stereo for FV PAD has not been explored previously (only for biometric recognition).

3 Data Collection

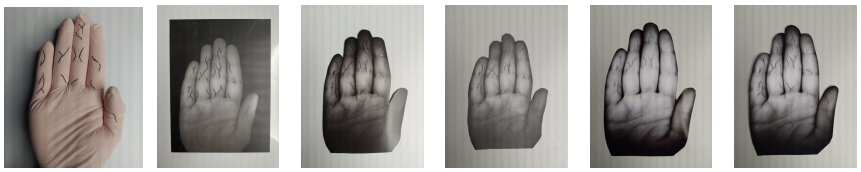
Using the *sweet* platform we have collected a new dataset for FVA and PAD experiments. This dataset, named Candy_v3,footnote 1 is organized in two parts. Part 1 contains data for FVA experiments, and Part 2 is used for FV PAD experiments. Figure 1(b) illustrate how *bona fide* presentations are made on *sweet*. The proposed dataset is described in this section.

(a) $850nm$ (b) $950nm$ (c) PS

Fig. 2: Finger-vein data captured in various channels using the *sweet* platform. (a) Fingers illuminated under 850nm NIR; (b) fingers illuminated under 950nm NIR; (c) pseudo-color map showing the 3D surface-normals reconstructed from the photometric stereo (PS) data.

***Bona Fide* FV Samples**: FV samples from 120 subjects (62 male and 58 female) comprise the Candy_v3 - Part 1 dataset. The two gender groups are approximately evenly distributed over three age-groups: 18-30 years, 31-50 years, and 51 and above. Five samples for each hand are recorded for each subject. The subject presents the hand with four fingers close together (named the *fingers-closed* modality), over the three cameras, at a distance of roughly 10 – 12cm from the cameras. (Note that, the presented hand is not in contact with the horizontal steel bar; it just appears so in the picture due to the perspective.) Thumbs are not used in our experiments.

Each sample includes 20 usable image-frames per camera (2 NIR cameras, and one color camera), captured under a variety of illuminations. For FVA, we use the FV images recorded by the two NIR cameras under 850nm and 950nm illumination. Each sample yields three images captured under NIR-850 illumination for each of the two (left, right) NIR cameras, and similarly, three images captured under NIR-950 illumination. The sample also includes four images captured specifically for PS reconstruction. These have been used in FV PAD experiments.



(a) B3 (b) B4 (c) B5 (d) B6 (e) B9 (f) B10

Fig. 3: Illustrations of some PAIs used in this study. The PAI species (described in Table 1) shown in each image is indicated in the caption.

Presentation Attacks: Here we have simulated the situation where a malevolent actor may gain unauthorized access to a database of FV biometrics samples. In other words, our PAIs have been constructed from the *bona fide* samples recorded previously using the *sweet* platform. To construct our PAIs we have arbitrarily selected 20 *target* subjects. For each target subject a single NIR (850nm) image of the right hand captured by the left NIR camera has been used to create PAIs of several species, grouped into two

levels, A and B:
Level A: PAIs can be created with minimal effort, within a day.
Level B: PAIs may take between one and three days to create, and require some expertise or specialized equipment.
As shown in Table 1, we have used PAIs of four Level A and 10 Level B species. The PAIs of Level A are made by printing the source-images on two kinds of photo-paper (glossy and matte) using two different kinds of printers – a laser printer and an inkjet printer. The toner used in laser printers typically has NIR-absorbent properties. The PAIs in the various Level B species have been made either by enhancing the source-images before printing them, or by creating artefacts in other ways that may confound the proposed FV-PAD method.

For each PAI also, five samples have been recorded. In most cases, the PAI is presented over the cameras in a way similar to a *bona fide* presentation. Care is taken to hold the PAI as flat and horizontal as possible. A different method has been used to capture PA samples of PAIs of the species B4. Here we have placed a cardboard box over the *sweet* platform. The box has a rectangular hole cut out on the top surface. A transparent plastic PAI is placed on the box such that the printed FV image is visible to the cameras through the hole in the box. A *bona fide* hand is then placed over the PAI, so that it is visible to the cameras through the PAI. This setup is illustrated in Figure 1(c). Examples of selected PAI artefacts are shown in Figure 3.

Table 1: Descriptions of the PAI species of levels A and B. The letters A and B in the first column indicate the level of the PAI species.

Species	Description
A1	Source image printed on glossy paper on a laser printer
A2	Source image printed on matte paper on a laser printer
A3	Source image printed on glossy paper on an inkjet printer
A4	Source image printed on matte paper on an inkjet printer
B1	Image printed on glossy paper on a laser printer, with veins enhanced manually using a black marker
B2	Image printed on matte paper on a laser printer, with veins enhanced manually using a black marker
B3	Vein pattern drawn with black marker on a latex glove
B4	Image printed on plastic transparency, with a *bona fide* hand presented over the transparency
B5	Hand-region cut out from image printed on glossy paper on a laser printer
B6	Hand-region cut out from image printed on matte paper on a laser printer
B7	Image printed on glossy paper on an inkjet printer, with veins enhanced manually using a black marker
B8	Image printed on matte paper on an inkjet printer, with veins enhanced manually using a black marker
B9	Hand-region cut out from image printed on glossy paper on an inkjet printer
B10	Hand-region cut out from image printed on matte paper on an inkjet printer

Fig. 4: Flowchart for constructing a FV-template from a FV-sample.

4 Baseline Algorithms

In this section we describe the finger-vein authentication (FVA) and FV PAD algorithms used in this study to generate baseline results for the Candy_v3 - Part 1 dataset.

4.1 FV Authentication Baseline

Finger-vein samples are compared based on *templates*. To enroll a new subject in the biometrics verification system, the subject first provides a biometric sample. A template, constructed from this sample, is stored in the biometrics system, associated with the subject's identity. During the probe phase, the subject claims a certain identity, and provides a new biometric probe sample. The system then compares the probe-template, derived from the probe-sample, with the template previously enrolled for the claimed identity. If the two templates are sufficiently similar (*i.e.*, the match-score is above a predetermined threshold), we consider that the claimed identity has been verified by the system. In this section we describe the template creation process used in our baseline FVA method, as well as the method used here for comparing FV templates.

4.1.1 Finger-vein Template Creation Figure 4 shows the flowchart of the FV template creation process. Each input FV-sample is an image corresponds to a presentation showing all fingers of the presented hand. First, the four fingers – index-, middle-, ring-, and little-finger – are segmented out from the input image. One template is constructed for each finger separately. The various steps highlighted in Figure 4 are described below.

Finger Segmentation: First we generate a *foreground mask* using adaptive thresholding (Otsu's method [11]) to detect the hand-region (foreground object) in the image. Small regions in the resulting binary image are deleted using morphological opening. We then scan the foreground mask along the horizontal axis for the first foreground pixel (assumed to belong to the hand in the image). The detected pixel is assumed to correspond to the tip of one finger (the tallest finger). The left and right boundaries of the finger are obtained by scanning horizontally for finger-edges on both sides. This scanning process is repeated for each row in the image, as long as the left and right finger-boundaries extracted correspond to a reasonable finger-width. The scanning process terminates when the estimated finger-width becomes too large at a certain row in the image. At this point we assume that we have identified all pixels representing a single finger. Then we remove this finger from the binary image (set all finger-pixels to background) and repeat the scanning process again, this time to find another finger. This procedure is repeated for up to four times, to detect four fingers in the image.

Fingers are detected in order of their height in the input image (the finger closest to the top-edge of the image is detected first, followed by the second-tallest, and so on). We use the relative coordinates of the center-of-gravity of each finger-mask to renumber the fingers in a natural order from index to little. This procedure of renumbering the fingers

works correctly only when all four fingers have been detected. (If, for example, only three fingers have been detected, then we cannot tell whether these are index, middle and ring fingers, or middle, ring, and little fingers.) For this reason, images where all four fingers are not detected, are excluded from further processing.

Finger Normalization: Next, a normalization step proposed by Huang *et al.* [5] is applied to each individual finger-image. This step simply rotates the finger-image to align the longitudinal axis of the finger to the vertical axis as best as possible.

Finger-vein Enhancement: We use a pre-trained autoencoder [2] to enhance the vascular structures in the input (normalized finger) image. Preliminary experiments showed that FV-enhancement improves the FV recognition accuracy significantly. Hence, we have included the FV-enhancement module in our processing pipeline.

Vascular Feature Extraction: FV patterns are compared based on a set of image-features extracted from the two vein-images being compared. In this work we have used the Maximum-Curvature (MC) features [10]. The MC feature-map extracted for a single finger is considered as the biometric template for the finger.

4.1.2 Finger-vein Matching We have used the method proposed by Miura *et al.*[10] to compare two MC-feature based templates. This method uses cross-correlation (computed in the frequency domain) to find the position of best match of the two input feature-maps. The cross-correlation coefficient at the best-match position is taken as the match-score between the two templates.

4.2 FV PAD Using Photometric Stereo

Our FV-PAD baseline method relies on a dense surface-normal map from the PS data recorded by each NIR camera. Before presenting our FV-PAD method, we briefly describe the processing of recovering the surface-normal map using PS.

4.2.1 Photometric Stereo Reconstruction The LEDs in the *sweet* platform are organized in four banks. The four LED banks, when fired in the sequence, illuminate the presented hand from different angles.The resulting set of four images can be used to recover the surface-normal map of the hand [19]. This PS reconstruction produces a dense surface-normal map (see Fig. 2(c)). The method also produces an albedo image of the hand being imaged.

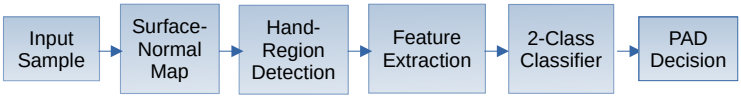


Fig. 5: Flowchart of the proposed FV-PAD algorithm.

4.2.2 Proposed FV-PAD Method The processing pipeline of the proposed FV PAD method is shown in Figure 5. The various steps in the proposed FV PAD method are described below.

Surface-Normal Map: The first step is to process these four images to generate the surface-normal map for the presented hand (as described in Section 4.2.1). One surface-normal map is reconstructed for each NIR camera. Each map consists of surface-normal vectors computed for each pixel-position.

Hand-Region Detection: K-means clustering [1] is used to group pixels into two clusters. The input to the clustering method consists of the three components of the surface-normal as well as the albedo value of each pixel. The cluster with the higher mean-albedo is considered as the foreground cluster. We assume that the largest foreground region corresponds to the presented hand. We use connected-component analysis to retain only the largest isolated region in the foreground cluster as the hand-region.

Feature Extraction: Each surface-normal vector (x, y, z) of the retained hand-region is transformed into an equivalent representation (θ_1, θ_2, M) where θ_1 represents the angle of the vector with respect to the x-y plane and θ_2 represents the angle of the vector with respect to the y-z plane. The angles θ_1 and θ_2 are bounded the range $[-\pi/2, \pi/2]$. The magnitude, M, of the surface-normal vector is ignored. For each surface-normal map we construct a two-dimensional (2D) histogram by quantizing θ_1 and θ_2 into N bins each, over the detected hand-region. ($N = 41$ in this study.) These 2D histograms are used to construct feature-vectors to distinguish between *bona fide* and PA classes.

For both classes, a large majority of the histogram-bins are empty (bin count = 0). Such empty bins act as confounders during the classification process. To mitigate this problem, we reduce the dimensionality of the feature-vectors using Principal Components Analysis (PCA) [1].

Feature Classification: We have used Gaussian Mixture Models (GMM) [1] to classify the feature-vectors generated in the previous step. One GMM is trained for each class (*bona fide*, PA). Given an input probe feature-vector, the likelihood that it comes from each of the two classes is evaluated using the GMMs corresponding to each class. Specifically, for a given feature-vector, let ll_{BF} and ll_{PA} be the log-likelihood values returned by the GMMs for the *bona fide* and PA classes, respectively. The difference $(ll_{BF} - ll_{PA})$ is taken as the classification-score for the probe feature-vector.

5 Baseline Experiments Using Candy Dataset

Baseline FVA and FV PAD results for the Candy_v3 dataset are discussed here. First we describe the methodology used in our baseline FVA experiments (Section 5.1), followed by the FVA performance shown by the baseline FVA method on the Candy_v3 - Part 1 dataset (Section 5.2). Then, in Section 5.3 we present the methodology adopted for our baseline FV PAD experiments, followed, in Section 5.4, by the FV PAD results for the Candy_v3 - Part 2 dataset .

5.1 Methodology For FVA Experiments

For the FVA experiments, first we group the subjects into two disjoint subsets, named the *development* ('Dev') set and the *evaluation* ('Eval') set. Data in the Dev set is used for tuning hyper-parameters of the finger-recognition system for the desired performance. The tuned recognition system is then applied to the data in the Eval set, to quantify the performance of the system. The Dev and Eval sets have been constructed arbitrarily – data for the first 60 subjects has been assigned to the Dev set and data for the remaining subjects has been assigned to the Eval set.

Recall that for the FVA experiments, only the NIR-850nm and NIR-950nm images from Candy_v3 - Part 1 dataset are used. From each image captured by the *sweet* platform, we extract three individual finger-vein images, corresponding to the index-,

middle- and ring-finger recorded in the image. Within each set (Dev or Eval), we have five FV samples for each subject. That is, for each camera we have 15 NIR-850 images and similarly 15 NIR-950 images, per subject, in each set. Considering three fingers (index, middle, and ring) per hand, for each of 60 subjects in each set, we have 90 single-finger vein images for each camera and each NIR-illumination.

In our single-finger FVA experiments, FV templates have been compared under eight different protocols, described in Table 2.[3] The table shows the number of enrolment and probe samples in each subset, for the eight protocols. Here, we have arbitrarily selected one sample for each modality of each subject for the enrolment sample. The remaining samples have been designated as probe-samples. For the single finger-vein recognition experiments, each enrolled sample is considered a unique identity. Each probe-sample has been used for four comparisons – one genuine comparison (with the correctly matched identity), and three *zero-effort-impostor* (ZEI) comparisons (with non-matched identities). In each ZEI comparison, the claimed-identity for a given probe template is selected randomly.

Table 2: List of protocols for finger-comparison experiments. For each protocol, the number of enrollment images and probe images in the development (Dev) set, as well as in the evaluation (Eval) set are also listed. Each protocol name consists of three strings indicating the hand ('LH' or 'RH'), the NIR camera used ('left' or 'right'), and the NIR illumination (850nm or 950nm).

Id.	**Protocol Name**	**Dev Set**		**Eval Set**	
		Num. Images		**Num. Images**	
		Enrol.	**Probe**	**Enrol.**	**Probe**
P1	LH_left_850	159	8064	159	7836
P2	LH_left_950	141	4908	138	4665
P3	LH_right_850	159	7968	156	7791
P4	LH_right_950	147	6372	144	6300
P5	RH_left_850	156	7200	156	6998
P6	RH_left_950	153	6138	136	4644
P7	RH_right_850	156	7116	156	7032
P8	RH_right_950	144	4842	123	3969

Next, we consider each hand of a subject as a unique identity. We have implemented a finger-score fusion based method to identify each hand based on three fingers of the hand: index-, middle-, and ring-finger. Thus, each hand-probe is represented by a 3-D feature-vector consisting of single-finger FVA scores. While constructing these feature-vectors, finger-FVA scores are selected either only from genuine-probes of a given hand identity, or only from ZEI-probes of the hand in question. In this way we obtain, for each hand-identity, a set of genuine-probe ('match') feature-vectors, and another set of ZEI ('non-match') feature-vectors.

A two-class classifier is then constructed using the feature-vectors in the Dev set. This classifier is used to label the hand feature-vectors of the Eval set. In this study, we have used Support Vector Machines (SVM) with RBF (radial basis function) kernels

[3] Each protocol name is composed of three items: <Hand>_<Camera>_<NIR>. The Hand may be 'LH' (Left hand) or 'RH' (Right hand). The Camera component ('left' or 'right') indicates the NIR camera from which the template has been derived, and the NIR component may be '850' or '950', indicating the illumination used to capture the image-sample. Thus, with two options for each of three variables, we have eight different experimental protocols.

[1], for the hand-identification experiments. In each protocol in Table 2, we fuse the FVA-scores of the three fingers of the hand. Thus, all probe feature-vectors used in a given experiment represent information from the same hand, captured by the same NIR camera, under the same NIR illumination.

5.2 FV Recognition At Fixed FMR

In this section we first present the results of single finger recognition. This is followed by the results of hand-recognition based finger-score fusion. In all our experiments, we have estimated the recognition performance at the operating point corresponding to a specific FMR of 0.1%.[4] In this analysis, the score-threshold is selected such that the FMR over the Dev set does not exceed the desired FMR. This score-threshold is then applied the Dev set and the Eval set, to determine the actual FMR and FNMR rates over each dataset.

5.2.1 Single Finger-vein Recognition In Table 3 we summarize the FMR and FNMR achieved for various evaluation protocols, for single finger recognition, for the FMR ceiling of 0.1%. The table shows that the FVA performance is significantly better for the right-hand fingers (protocols P5-P8) than for the left-hand fingers. We do not have any logical explanation for this phenomenon. We assume that right hand presentations by most subjects may have been more consistent (less variability) than left hand presentations, for one of two reasons: (1) increased familiarity with the data-capture procedure – subjects were consistently asked to present the left-hand first, or (2) simply due to right-handedness of most subjects.

Table 3: Finger-vein recognition performance (expressed as percentages) at False-Match rate (FMR) of 0.1%. The table shows results for single-finger recognition as well as hand-recognition based on finger-score fusion. The lowest HTER values are highlighted in bold characters.

	Single-Finger Recognition						**Multi-Finger Hand Recognition**					
Prot	**Dev Set**			**Eval Set**			**Dev Set**			**Eval Set**		
ocol	**FMR**	**FNMR**	**HTER**	**FMR**	**FNMR**	**HTER**	**FMR**	**FNMR**	**HTER**	**FMR**	**FNMR**	**HTER**
P1	0.1	3.82	1.96	0.17	7.3	3.74	0.1	2.68	1.39	0.20	4.29	2.25
P2	0.08	8.7	4.39	0.09	10.05	5.07	0.08	4.2	2.14	0.54	3.57	2.06
P3	0.08	5.77	2.93	0.0	6.46	3.23	0.05	2.71	1.38	0.05	3.93	1.99
P4	0.09	7.74	3.91	0.02	9.45	4.74	0.06	3.04	1.55	0.6	2.67	1.66
P5	0.09	0.66	0.38	0.79	0.61	0.7	0.06	0.0	0.03	0.46	0.0	0.23
P6	0.09	0.57	0.33	0.66	0.60	0.63	0.07	0.0	0.03	0.45	0.0	0.23
P7	0.09	0.33	0.21	0.68	0.45	**0.57**	0.06	0.0	0.03	0.11	0.0	**0.06**
P8	0.09	1.39	0.74	0.04	2.27	1.15	0.08	0.46	0.27	0.0	1.51	0.76

The single-finger FVA results in Table 3 that the recognition-rates achieved for protocols involving 850 nm NIR illumination are usually somewhat better than the corresponding (*i.e.*, same hand, same camera) protocols involving 950 nm illumination. This result is counterintuitive. In theory, we expect 950 nm illumination to provide better results than 850 nm, because 950 nm NIR penetrates the soft-tissue of the fingers to a deeper extent than 850 nm NIR. Also, 850 nm NIR tends to produce more speckle noise on the skin-surface. On the other hand, much more power is needed for the 950

[4] In some experiments, the actual FMR achieved over the Dev set may be slightly lower than the desired FMR of 0.1%.

nm illumination. Our conjecture is that in the *sweet* platform the 950 nm illumination may not be sufficiently powerful.

5.2.2 Hand-Recognition Based on Finger-Score Fusion The results of FVA-score fusion within each FVA protocol are also shown in Table 3. These numbers quantify the performance of the score-fusion system corresponding to the classification-score-threshold that limits the FMR over the Dev set to 0.1%.We note that FVA-score fusion improves the hand-recognition performance compared to single finger FVA. For the left hand, the single FVA error-rates (HTER in Table 3), range from 3.5% to 5% for each of the individual fingers. Finger-fusion reduces the left-hand recognition error-rates to about 2% or lower in all four left-hand protocols. For the right-hand, single-finger FVA performance is already very high (FVA in protocols P5 – P8 in Table 3). Multi-finger FVA performance for the right hand still reduces the classification error. The best performance for multi-finger FVA, with a HTER of 0.06% for the 'RH_right_850', is almost a 10-fold improvement over single-finger FVA in the same protocol.

5.3 Methodology for FV PAD Experiments

In this section we discuss the results achieved by the proposed photometry based FV PAD method (described in Section 4.2), on the Candy_v3 - Part 2 dataset. For the classification experiments reported here, the PCA model for dimensionality reduction has been trained to retain 99% of the total information in the input space (for the Training set). The trained PCA model is applied to transform the input histogram-features into a new 10-D feature-space. Thus, in the classification step, each presentation is represented by a feature-vector of length 10.

Data Partitions: We have randomly grouped the subjects into two non-overlapping sets: *Training* and *Test*[5]. Each set consists of data from 60 subjects – 10 target subjects, and 50 non-target subjects. In each set (training, test), the *bona fide* feature-vectors correspond to all 60 subjects comprising the set, whereas the PA feature-vectors come from the 10 target subjects in the set. For each subject we consider only the feature-vectors representing the surface-normal map of the right hand, computed for one camera – the 'left' camera. The Training set consists of 590 *bona fide* and 1390 PA vectors. The Test set consists of 608 *bona fide* and 1382 PA vectors.

PAD Classification: Two GMMs have been trained, one using the *bona fide* feature-vectors and the other using the PA feature-vectors (of the Training set). For each GMM we select the number of Gaussians such that the Akaike Information Criterion (AIC) is minimum [1], based on trials with different numbers of Gaussians. Thus, we obtain a GMM with two Gaussians for the *bona fide* class and a GMM with 26 Gaussians for the PA class.

PAD Performance Metrics: We report the PAD results using the following figures of merit defined in the ISO/IEC 30107-3 standard [6]:

APCER: Attack Presentation Classification Error Rate – the proportion of PAs that are misclassified, and

BPCER: *Bona fide* Presentation Classification Error Rate – proportion of *bona fide*

[5] Due to the small size of the Candy_v3 - Part 2 dataset, no separate *development* set has been used for tuning classifier hyper-parameters.

presentations that are misclassified.
The Average Classification Error Rate (ACER) – the mean of the APCER and BPCER – summarizes the two ISO figures of merit with a single number. APCER, BPCER, and ACER are expressed as percentages.

5.4 Baseline FV PAD Results

In this section we present the classification results produced by our baseline FV PAD algorithm at two classification-score-thresholds: (1) the threshold corresponding to the EER on the Training set, and (2) the threshold that limits the APCER to a maximum of 1% on the Training set. FV PAD results for both operating points are shown in Table 4. The table shows overall PAD results for each operating point, PAD results for each PA level (Level A and Level B), as well as for each individual PAI species.

5.4.1 PAD Classification at EER Here we have selected the classification-score-threshold, $\mathcal{T}_{EER} = -1.389$, corresponding to the EER (ACER = 2.4%) over the Training set. The APCER and BPCER values achieved for the Test set using $\mathcal{T}_{EER}$ are shown in the left half of Table 4. Following ISO convention [6], when several PAI species of the same Level are used, the performance metrics of the worst performing species are to be taken as the aggregate PAD performance for the Level. Therefore, PAD performance for Level A corresponds to the PAD performance obtained for PAI species A1, and the performance metrics stated for Level B correspond to the species B3. We note that, apart from PAI species B3, B5 and B6, the ACER for the other species are low.

5.4.2 PAD Classification at Fixed APCER (1%) Next we examine the performance of the proposed FV-PAD method when the APCER is limited to 1% (this operating point is also referred to as the $BPCER_{100}$ point). Note that APCER $<= 1\%$ is a very stringent constraint. The corresponding score-threshold, $\mathcal{T}_{1\%} = 0.464$, is determined from the Training set.

The PAD performance metrics for the Test set at this operating point are shown in the right half of Table 4. We note, again, that the APCER is high for three PAI species, namely B3, B5 and B6, but quite low for the remaining PAI species. We note that in this experiment $BPCER_{100}$ remains at an acceptable level for most applications. We also note that, except for the PAI species B3, B5 and B6, the APCER for the various PAI species remains at or below 3%.

5.4.3 Analysis of FV-PAD Results The results in Table 4 validate the PS based surface-shape reconstruction approach for FV PAD, in general. The proposed method is clearly not adequate for PAs of species B3 (PAs where the hand is covered with a latex glove on which vein-patterns have been drawn with dark (NIR-visible) ink). Surface-normal statistics alone are not sufficient to distinguish this class of PAs from *bona fide* presentations. Additional cues are required to detect PAIs of species B3.

PAI species B5, B6, B9, and B10 include PAIs made from printed vein-pattern images, where the hand region has been cut out following the outline of the fingers (see Fig. 3). Our experiments shows that the APCER is high for the two PAI species printed on laser printer (B5, B6) but significantly lower for B9 and B10, which have been printed on an inkjet printer. In general, we believe that for these four PAI species, additional data, in the form of more PAIs in the training set, may help improve the PAD results to some degree.

Table 4: Performance of the proposed FV-PAD method at two operating points: the EER for the Training set (*i.e.*, score-threshold $\mathcal{T}_{EER} = -1.839$), and $BPCER_{100}$ (*i.e.*, $\mathcal{T}_{1\%} = 0.464$, which limits the APCER on the Training set to 1%). Note that the APCER observed for the Test set may be higher from the desired limit.

Level / Species	**EER ($\mathcal{T}_{EER}$)**			**APCER <= 1% ($\mathcal{T}_{1\%}$)**		
	APCER	**BPCER**	**ACER**	**APCER**	**BPCER**	**ACER**
Overall	5.2	1.2	3.2	3.1	9.4	6.2
Level **A**	1.0	1.2	1.1	0.0	9.4	4.7
Level **B**	31.7	1.2	16.4	15.0	9.4	12.2
A1	1.0	1.2	1.1	0.0	9.4	4.7
A2	0.0	1.2	0.6	0.0	9.4	4.7
A3	0.0	1.2	0.6	0.0	9.4	4.7
A4	0.0	1.2	0.6	0.0	9.4	4.7
B1	1.0	1.2	1.1	0.0	9.4	4.7
B2	1.0	1.2	1.1	1.0	9.4	5.2
B3	31.7	1.2	16.4	12.2	9.4	10.8
B4	0.0	1.2	0.6	0.0	9.4	4.7
B5	18.0	1.2	9.6	15.0	9.4	12.2
B6	16.0	1.2	8.6	12.0	9.4	10.7
B7	0.0	1.2	0.6	0.0	9.4	4.7
B8	0.0	1.2	0.6	0.0	9.4	4.7
B9	5.0	1.2	3.1	2.0	9.4	5.7
B10	4.0	1.2	2.6	3.0	9.4	6.2

6 Conclusions

We present here a new finger-vein (FV) dataset, named Candy_v3, collected using a newly designed contactless finger-vein sensor named *sweet*. Our open-source sensor design is described in a separate paper. Unlike most FV sensors currently available, *sweet* can record FV samples of multiple fingers simultaneously, under two different NIR illuminations (850nm and 950nm). It also collects a sequence of four images where the presented hand is illuminated from a different angle for every image. Photometric Stereo (PS) reconstruction may be applied to this sequence of images to compute a surface-normal map of the hand.

The proposed Candy_v3 dataset is organized in two parts. Part 1 consists of *bona fide* FV samples from 120 subjects, and Part 2 includes PA samples using four species of PAIs of Level A, and 10 PAI species of Level B (20 PAIs per species). In this work we have also provided some baseline FVA and FV PAD results for the proposed dataset.footnote 1

We also provide the performance of baseline FVA and FV PAD algorithms on the new dataset. Our experiments show that FVA accuracy of 0.06% FNMR can be achieved at the operating point corresponding to a FMR of 0.1%. The baseline FV PAD method presented in this work analyzes the surface-normal map reconstructed using PS, to determine whether the corresponding presentation is *bona fide* or a PA. Overall PAD performance, summarized by the ACER for the test set, was 3.2% at EER and 6.2% when APCER was constrained to 1%. In future work, we plan to develop a CNN based method for FVA using this dataset, as well as to develop more effective FV PAD approaches.

References

1. Bishop, C.M.: Pattern Recognition and Machine Learning. Springer, 1 edn. (2007) 8, 10, 11
2. Bros, V., et al.: Vein enhancement with deep auto-encoders to improve finger vein recognition. In: Proc. IEEE Intl. Conf. BIOSIG (2021) 3, 7
3. Geissbühler, D., et al.: *sweet* – an open source modular platform for contactless hand vascular biometric experiments (2024), `https://arxiv.org/abs/2404.09376` 1
4. Hou, B., Zhang, H., Yan, R.: Finger-vein biometric recognition: A review. IEEE Transactions on Instrumentation and Measurement **PP**, 1–1 (01 2022) 2
5. Huang, B., et al.: Finger-vein authentication based on wide line detector and pattern normalization. In: 2010 20th International Conference on Pattern Recognition (2010) 2, 7
6. ISO/IEC 30107-3. Information technology — Biometric Presentation Attack Detection – part 3: Testing and Reporting. Standard, International Organization for Standardization, Geneva, CH (Mar 2017) 11, 12
7. Kovač, I., Marák, P.: Finger vein recognition: utilization of adaptive gabor filters in the enhancement stage combined with sift/surf-based feature extraction. Signal, Image and Video Processing **17**, 635 – 641 (2023) 3
8. Lu, Y., et al.: An available database for the research of finger vein recognition. vol. 1, pp. 410–415 (12 2013) 2
9. Miura, M., et al.: Feature Extraction of finger-vein patterns based on repeated line tracking and its Application to Personal Identification. In: Proceedings on IAPR conference on Machine Vision and Applications. vol. 15 (2004) 2
10. Miura, M., et al.: Extraction of Finger-Vein Pattern Using Maximum Curvature Points in Image Profiles. In: Proceedings on IAPR conference on Machine Vision Applications (2005) 2, 3, 7
11. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. on Systems, Man, and Cybernetics **9**(1) (1979) 6
12. Qiu, X., Kang, W., Tian, S., Jia, W., Huang, Z.: Finger vein presentation attack detection using total variation decomposition. IEEE Transactions on Information Forensics and Security **13**(2), 465–477 (2018) 2, 3
13. Raghavendra, R., Busch, C.: Presentation Attack Detection Algorithms for Finger Vein Biometrics: A Comprehensive Study. In: Proc. 11th Intl. Conf. on Signal-Image Tech. & Internet-Based Systems (SITIS) (2015) 3
14. Raghavendra, R., et al.: Transferable deep convolutional neural network features for fingervein presentation attack detection. In: 5th Intl. Wkshp. on Biometrics and Forensics (IWBF) (2017) 3
15. Sengupta, S., et al.: SfSNet: Learning Shape, Reflectance and Illuminance of Faces 'in the Wild'. In: Proc. IEEE Conf. on CVPR (2018) 3
16. Singh, J.M., et al.: Detecting Finger-Vein Presentation Attacks Using 3D Shape & Diffuse Reflectance Decomposition. In: 15th Intl. Conf. on Signal-Image Tech. & Internet-Based Systems (SITIS) (2019) 3
17. Ton, B.T., Veldhuis, R.N.J.: A high quality finger vascular pattern dataset collected using a custom designed capturing device. In: Proc. 2013 Intl. Conf. on Biometrics (ICB). pp. 1–5 (2013) 2
18. Vanoni, M., et al.: Cross-database evaluation using an open finger vein sensor. In: 2014 IEEE workshop on biometric measurements and systems for security and medical applications (BIOMS) proceedings. pp. 30–35. IEEE (2014) 2, 3
19. Woodham, R.J.: Photometric method for determining surface orientation from multiple images. Optical Engineering **19**(1) (1980) 7

20. Yang, J., Shi, Y., Wu, R.: Finger-vein recognition based on gabor features. In: Riaz, Z. (ed.) Biometric Systems, chap. 2. IntechOpen, Rijeka (2011) 3
21. Yin, Y., et al.: SDUMLA-HMT: A multimodal biometric database. In: Biometric Recognition. pp. 260–268. Springer Berlin Heidelberg (2011) 2
22. Zhang, R., et al.: Deep learning for finger vein recognition: A brief survey of recent trend (07 2022), https://arxiv.org/pdf/2207.02148.pdf 3