Journal

Sarah Faltaous*, Jonathan Liebers, Yomna Abdelrahman, Florian Alt, and Stefan Schneegass

VPID: Towards Vein Pattern Identification Using Thermal Imaging

https://doi.org/..., Received ...; accepted ...

Abstract: Biometric authentication received considerable attention lately. The vein pattern on the back of the hand is a unique biometric that can be measured through thermal imaging. Detecting this pattern provides an implicit approach that can authenticate users while interacting. In this paper, we present the Vein-Identification system, called VPID. It consists of a vein pattern recognition pipeline and an authentication part. We implemented six different vein-based authentication approaches by combining thermal imaging and computer vision algorithms. Through a study, we show that the approaches achieve a low false-acceptance rate ("FAR") and a low false-rejection rate ("FRR"). Our findings show that the best approach is the Hausdorff distance-difference applied in combination with a Convolutional Neural Networks (CNN) classification of stacked images.

Keywords: Thermal Imaging, Usable Security, Biometrics

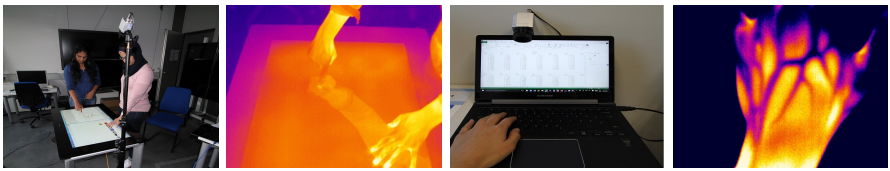
PACS: ...

Communicated by: ...
Dedicated to ...

1 Introduction

Previous work recognized the limitations of current authentication mechanisms. A survey of the password habits among American consumers, for example, indicates that the average consumer takes minimal action to secure their online accounts [12]. Prior work further showed that only few users use a PIN for their mobile device [28] and passwords are frequently re-used and forgotten [15]. Abdulwahid et al. conclude

***Corresponding author: Sarah Faltaous, Jonathan Liebers, Stefan Schneegass,** University of Duisburg-Essen
Yomna Abdelrahman, Florian Alt, Bundeswehr University Munich



(a) Thermal view from a top-mounted camera over an interactive table-top

(b) Thermal view from a camera attached to a laptop screen

Fig. 1: Examples of false-color images which interfere with the skeleton extraction.

that robust authentication mechanisms are needed that operate in a transparent, continuous, and user-friendly fashion [4]. In search for usable and secure authentication mechanisms, a large number of approaches have been proposed. A promising direction is biometric authentication [23]. Biometric authentication identifies a user using physiological or behavioral characteristics [19, 24, 32, 35]. One of its major advantages is that there is no need to remember a PIN, a password, or another secret.

In current systems, biometric authentication mechanisms typically require the user to perform an explicit action. Fingerprint authentication, for example, requires to put a finger on a fingerprint sensor and an iris scanner requires looking into a camera. Another biometric authentication approach is the use of vein patterns of the palm dorsal. Using vein patterns underneath the human skin has several advantages. Every person has a unique pattern of veins, which is stable from the age of ten and unique even for twins [1]. Unlike other biometrics (e.g., fingerprint), the extraction and duplication of vein patterns needs massive interventions. As a result of being less fraud-prone, such approaches are supposed to provide a higher security level for any system.

As veins have a different temperature than the surrounding skin [11], thermal cameras can capture these patterns from a distance without interfering with the user's current task. Figure 1 shows two examples of possible application scenarios with interactive system. The first example is for tabletop multi-user identification, as shown in Figure 1a, where users could be authenticated based on their captured vein patterns from a thermal camera mounted to cover the field of interaction. This could be applied to permit access and usage in a shared office setup or to customize settings for smart households. Another example is laptop authentication. Having a thermal camera that faces a keyboard, as shown in Figure 1b, the system would be able to track who is using a laptop and can reject access for unknown users or to specific information. In both cases, user identification and authentication happens continuously and implicitly and, thus, does not provide an additional burden to

the user. As soon as a user starts to interact, the system automatically recognizes the user and grants or denies access to it.

In accordance with Crawford [10] we define "implicit authentication" analogously to a "zero-interaction authentication" [8] that is transparent to the user. This kind of authentication bears the inherent advantage that a user is not bothered by the authentication process, which traditionally requires an explicit interaction with the authentication system (e.g., such explicit interaction additionally interrupts the interaction with the underlying systems thus negatively impacts usability). Our presented system eliminates the need for explicit interaction and is subsequently especially suited for the usage in public or open spaces, with public displays, in offices or in many other scenarios, where spontaneous interactions can occur.

We present Vein Pattern Identifier (*VPID*), a system that uses thermal images for vein pattern authentication. We implemented six different approaches including standard computer vision algorithms and deep learning methods which we compared against each other. We conducted a controlled lab study with 12 participants, where we collected data to evaluate the algorithms. The results show that the Hausdorff distance [21] as a shape comparison algorithm with skeleton stacking and a CNN for classification performs best.

The contributions of this paper are as follows:

1. A processing pipeline[1] that extracts the vein pattern from thermal images and identifies users based on it.
2. A study comparing the accuracy provided by six different computer vision algorithms along with Convolutional Neural Network (CNN).

2 Related Work

We identified two strands of related work to our research: (1) seamless user authentication and identification using biometric information and (2) thermal imaging as a sensing technology.

2.1 Biometric Authentication

The uniqueness of biometrical features has been widely explored to identify and authenticate users [9, 23, 32]. Biometrics are classified into behavioral and physiological biometrics [41]. Behavioral biometrics rely on behavioral cues to authenticate

1 The software can be found at http://research.hcigroup.de.

users, for instance, their touch and keystroke behavior [13, 16, 33]. Physiological biometrics are based on "something you are" and include a person's physiological information such as iris, face, voice, fingerprint, and hand geometry [37]. An example is Bodyprint [19] that uses body parts like ear, finger, fist, and palm prints, obtained via a mobile phone touchscreen as a biometric trait.

Veins patterns are another biometric feature that has been proposed for identification and authentication. Since vein patterns are unique for each users and stable for a lifetime [1], they are well suited as a biometric. For example, veins can be captured under IR-lights [5], using VGA [39], or near IR imaging [17]. While these technologies provide good input, they either require external sources of illumination [38], have high costs, or require multiple sensors [14].

Researchers also explored using thermal (far IR) imagining to extract the veins on the palm [27, 40]. As veins have different temperature than the surrounding skin [11], they are visible to thermal cameras without additional illumination. Vein triangulation and knuckle shapes are used to differentiate between users. From this we learn that veins are a powerful means for authentication. At the same time, approaches are required that allow authentication using vein patterns to be seamlessly integrated with interactive systems.

2.2 Thermal Imaging for Sensing Interaction

Due to the advantages that thermal imaging provides, including robustness to illumination and color changes, it has recently been proposed as a sensing technology for interactive systems [3, 29, 34]. This is achieved by integrating thermal imaging and existing computer vision techniques to improve touch or gestural interaction. The advantages of thermal imaging are utilized to overcome common RGB and depth cameras' limitations, such as light dependence and lack of visible differences between veins and skin. Larson et al. used heat traces caused by fingers touching a surface and detected pressure for interaction on surfaces [29]. Sahami Shirazi et al. proposed thermal reflection to expand the interaction space beyond the camera's direct field-of-view [34]. Abdelrahman et al. investigated the material properties supporting both touch and mid-air gestural interaction using thermal imaging as sensing technology [3]. Furthermore, Woo Cho et al. [7] introduced a new system to detect the thermal handprint for forensic identification.

Thermal imagers have been used in security contexts before. Abdelrahman et al. presented an approach to reveal PIN and lock patterns from thermal images of a user's mobile device [2]. Overall, thermal cameras are promising sensors for a range of interactive systems e.g., multiple user interaction, tabletop interactions [34] that would naturally benefit from seamless identification and authentication.

3 Vein Identification Approaches

In the following, we present the VPID system that uses thermal imaging to identify users. VPID consists of a recognition pipeline and an authentication part. The recognition pipeline consists of nine steps starting with the raw camera data and resulting in the user's vein pattern that can be fed into the authentication system. For the authentication part, we propose six different vein-based authentication algorithms.

3.1 Vein Pattern Construction Pipeline

For extracting the hand veins, we are using the OpenCV library[2] for image processing and features extraction. We apply in total nine steps to a thermal image stream to construct the vein pattern of each user that we later use to authenticate a user.

1. Image extraction: First, we extract frames of raw data from the live-stream of the thermal camera (cf., Figure 2a).

2. False-color image transformation: We transform the acquired raw thermal data, represented by an array of unsigned short integers, into a false-color image using an iron color palette scaled from $minimumtemperature - 7$ and $maximumtemperature + 1$. This palette boosts the color difference of temperatures across the shading palette without enhancing or omitting any potential temperature [18] (cf., Figure 2b). This insures that all small details are captured. Afterwards, we apply a color-based threshold to remove any background noise from the hand by setting it to black (cf., Figure 2c).

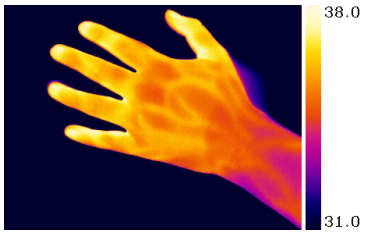
3. Hand-detection: Then, we use the data from step 2 to detect the warmest area in the image. A convex hull is then drawn to connect all the local maxima of the warm points. Next, we detect the hand informalities (i.e., fingers and knuckles) to verify a hand shape. We measure the defects, which are the points with the largest distance from the convex set (cf., Figure 2d). If no defects are detected, the shape is not identified as a hand.

4. Region of interest (ROI): As the vein pattern is located on the back of the hand, we define the ROI as the circle with the largest diameter enclosed in the hand's contour (cf., Figure 2e). For that, we try all possible combinations of circles across the hand area (i.e., warmest area). To reduce random noise and artifacts at

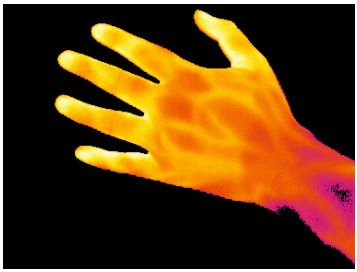
2 OpenCV: http://opencv.org/

```
jl@x250:~> hexdump raw_thermal_data.bin
0000000 18d2 194a 1952 1946 199b 1972 196e 19ee
0000010 199f 1957 19f9 196f 1a3e 19c3 1a5e 19b6
0000020 1a29 19c5 19e0 19dc 1a14 1a22 19df 19c1
0000030 19ea 199a 19db 19fa 1a6c 19df 1a1a 19c4
0000040 19f0 19aa 19d7 19b9 19bc 19de 19f5 19a3
0000050 19d9 19cd 1a15 19c2 1993 19f8 1a0a 19ec
0000060 19f0 1a00 19df 19b4 19e1 1a24 1a10 1a21
0000070 1a02 1a35 19af 1994 19a5 19c1 1a0c 19c7
0000080 1a01 19ac 19e4 19cb 19a6 19d4 19fd 1a24
0000090 19e1 19bb 19c9 19e1 19a6 19d4 19ec 1990
00000a0 1992 19a6 1a0a 1a0d 19a7 19c8 19d1 19fe
00000b0 199d 19f1 1a5f 19f5 1a68 19b3 1a0b 1aa5
00000c0 198f 1a55 19ce 1a04 1a19 1989 19d9 19b6
00000d0 1a28 19ae 19b0 19e8 19e6 19db 197e 196a
00000e0 19bf 19d5 1982 19d9 19c9 1934 19c0 194a
```

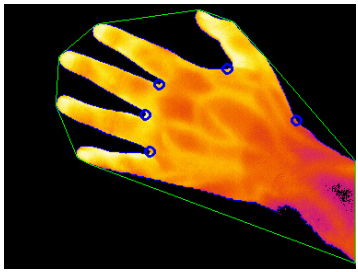
(a) Raw data received from the thermal camera in the form of unsigned short integers



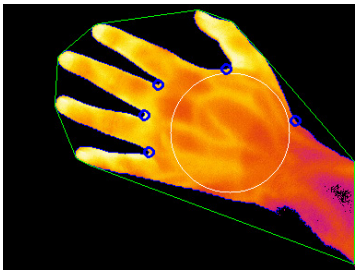
(b) The raw thermal data is then processed into a false color image. The colorbar maps the values of Celsius degrees into iron palette colors.



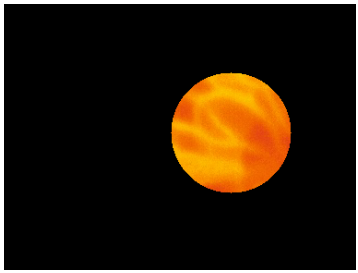
(c) A color-based thresholding removes the background of the hand by setting it to black.



(d) A convex hull is placed in green color around the hand. The blue smaller circles form up to five defects in the hull. Also, the hand is separated by a blue contour from the black background.

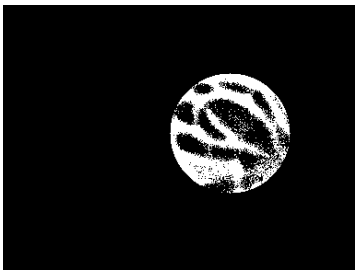


(e) The largest circle, depicted in white color, is found within the area of the hand. Its size is reduced by a factor of 0.95.

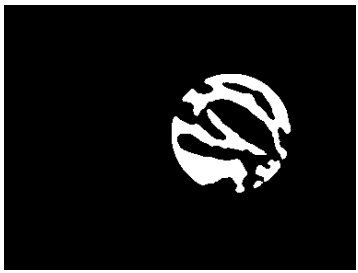


(f) The extraction of the region of interest that would be further used for processing

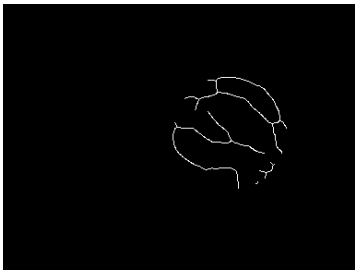
Fig. 2: The first half of the steps in our pipeline that would result in the region of interest



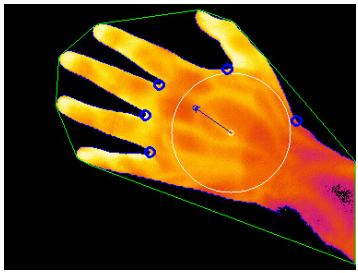
(a) An adaptive threshold is used to process the region of interest



(b) Noise filtering using morphological operations (i.e., open) followed by a median filter.

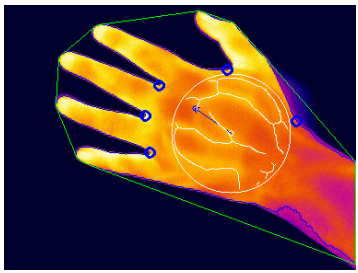


(c) Using the "thinning" function, the size of each white area is reduced till it is exactly one pixel thin.



(d) Using the defects (i.e., blue circles) along with the center of the ROI a vector is plotted. This vector is then rotated to overlap the postive y-axis.

(e) The skeleton after rotation. Furthermore the skeleton is zoomed until it fits the height of the 382×288 pixels resolution.



(f) Finally, all information are merged into a debug image, which shows the skeleton on the original image.

Fig. 3: Starting with the region of interest, the second half of the image processing pipeline is presented. Specially, the area from the region of interest is processed into a skeleton.

the border of the circle, the diameter is hence multiplied with a constant factor of 0.95 and we extract that mask (cf., Figure 2f).

5. Extraction of the vein pattern skeleton: We transfer the circular mask depicting the ROI to grayscale image with a normalized histogram. Then, we use an adaptive-threshold based approach to detect the local maxima and assigning them to be part of the vein skeleton (cf., Figure 3a).

6. Noise removal: Afterwards, we use median blurs followed by an opening morphological operation to eliminate noise and small, single artifacts from the image (cf., Figure 3b).

7. Topological skeleton: The image obtained from step 6 results in thick lines or broad areas (i.e., one or multiple connected lines with a width of more than one pixel). Hence, we used the thinning approach of Zhang Suen [42], which reduces the size of each area until it is exactly one pixel thin (cf., Figure 3c).

8. Skeleton affine transformations: Using the center of the ROI obtained in step 4, we obtain the center point of the defects in step 3 and plot a vector. This vector is used to rotate any captured image upwards, in the direction of a positive y-axis (cf., Figure 3d). To further unify the metrics of the logged data, we scale the ROI so that the diameter reaches the maximum height of the image (cf., Figure 3e).

9. Skeletons stacking: In this final step, the vein skeletons obtained from several images of the same video sequence (i.e., same hand) are stacked over each other to assure the best quality to the finest detail. We refined the first 14 frames of each video and stacked them as a final step.

3.2 Vein Pattern Authentication

The authentication process is done through comparing the similarities between vein pattern skeletons. We implemented six different approaches for shape comparison, namely, Hausdorff distance, Shape Context, Jaccard distance, and three different versions of Hu moments. In addition, we refine our results by machine learning algorithms.

Hausdorff and Shape Context distances based on the work of Huttenlocher, Klanderman and Rucklidge [21], and Belongie et al. [6], respectively. These two approaches depend mainly on comparing the distance-difference between two skeletons. Hence, a user is authenticated if the comparison of two skeletons, S_1 and S_2 results in $distance(S_1, S_2) \leq V_{threshold}$.

Jaccard distance. The Jaccard index [22] defines a metric for similarity between two sets (e.g., A and B). Where A consists of all white pixels of the first

skeleton, and B consists of all white pixels of the second skeleton. Each pixel is defined as a tuple of its X- and Y-coordinates. The Jaccard index is then defined as: $J(A,B) = \frac{|A \cap B|}{|A \cup B|}$, where $0 \leq J(A,B) \leq 1$. In other words, the size of the overlapping area between two images is divided by the area of the union of both images, where the area is defined by the presence of the white pixels in the binary image. To provide a comparable metric the Jaccard distance is used. It is defined as: $J_{Distance}(A,B) = 1 - J(A,B)$.

Hu moments based on the work of Hu [20]. These are numerical values retrieved from image characteristics (e.g., lie on the same plane, has non-zero values, etc.). These numerical values then can be compared to each other. Hu proposed three different ways, namely:

1. $I_1(A,B) = \sum_{i=1\ldots7} |\frac{1}{m_i^A} - \frac{1}{m_i^B}|$
2. $I_2(A,B) = \sum_{i=1\ldots7} |m_i^A - m_i^B|$
3. $I_3(A,B) = max_{i=1\ldots7} \frac{|m_i^A - m_i^B|}{|m_i^A|}$

where $m_i^A = sign(h_i^A) \cdot log(h_i^A)$ and $m_i^B = sign(h_i^B) \cdot log(h_i^B)$ are the Hu moments h_i^A, h_i^B of A and B respectively.

Convolutional Neural Network. Recent work using deep learning has shown that Convolutional Neural Network (CNN) can be used to classify images [25, 30]. A trained CNN can match one new, unknown image onto one of many learned classes of images. When used as standalone mechanism, the input data would always be identified and classified to any of the registered classes. Therefore, we deployed the CNN with the other comparison and matching algorithms. In our case, this can be used to train the system with the skeleton images of the user's vein pattern. For our work, we tested various CNNs (i.e., LeNet [31], AlexNet [26], and VGG-16 [36]) with sample data collected across several pilot studies. Finally, we decided to use "AlexNet" for our evaluation, as it does not require long training time and has high accuracy. The network was trained for 30 epochs, with a constant random seed, stochastic gradient descent as the solver and a base learning rate of 0.001.

4 Data Collection

For testing the approaches, we used an Optris PI450[3] contactless thermal camera. It has a 62° × 48° field-of-view, 382 px × 288 px optical resolution, and temperature

3 http://www.optris.com/thermal-imager-pi400

sensitivity of 40 mK. The camera was mounted on a tripod at a fixed height of 37 cm, facing towards a table to capture the participant's hand as depicted in Figure 4. The room temperature conditions were kept constant for all sessions.

For capturing thermal images, we have solely used the "iron" colouring palette scheme, as it provides an infliction point in its gradient over the relative colour intensity. This infliction point helps in separating the "warm" colour values per pixel from the "cold" colour values. To set the exposure parameters (i.e., maximal and minimal temperature to scale the colouring palette scheme) for the thermal imager, we have automatically determined the warmest point of temperature within the image, and added a constant offset of +1 °C for the maximum observed temperature and −5 °C for the minimum temperature. Our tests showed that this offset is ideal to capture the participant's hand as a whole with all of its contents.

4.1 Participants and Procedures

We recruited 12 participants (4 female, 8 male, $\mu = 28$, $\sigma = 8.5$) via university mailing lists. The participants have backgrounds in different majors. After welcoming the participants, we described the goal of the study and handed out consent forms as well as demographic questionnaires. We asked the participant before recording to wash his/her hands under normal tab water. This procedure was done at the beginning to unify the temperature of participant's hands. The study was conducted over two sessions on the same day, separated by $2 - 3\,mins$ break. In every session, each hand (i.e., left or right) was recorded for 20 $secs$. We defined the area and direction in which the hand should be placed on and pointing to, so the camera could start recording.

Collecting biometric information from participants raised the need for a secure procedure, in order to protect their privacy. As the always existing, possible risk of a data breach leaves participants with no way of changing or revoking their biometric information, a failure to comply with a secure procedure to protect their data is highly unethical. Subsequently, we gathered the informed consent of all our participants and assigned them an anonymous identifier. Directly upon finishing the acquisition of data per participant, we deleted the connecting attribute of their private information to the anonymous identifier, so that no connection can be re-established in the future.

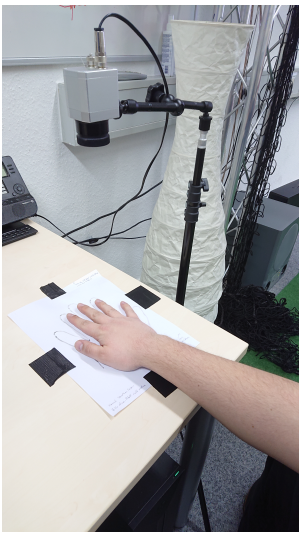


Fig. 4: The setup of the controlled study, which was used for data collection. The camera was mounted on a height of $37cm$.

4.2 Evaluation

To start with the evaluation process, we had to first categorise the received images. One major aspect that directly influences the obtained results is the quality of the images captured by the thermal camera. In general, we noticed that certain circumstances exist that result in a "good" hand thermal-image recording, which contain a very visible vein pattern. In contrast to other circumstances, a "bad" thermal recording with the vein pattern being only barely visible. The quality of a captured image is defined by the visibility of the vein pattern in the thermal image, which can be measured by its contrast towards the surrounding area of tissue. One main contributing factor is the angle of the hand towards the thermal imager. If the hand is twisted or curled while being captured by the thermal imager, the vein pattern might not be visible at all. One of the core assumptions of our vein-detection algorithm is that the vein pattern would be in general warmer (i.e., brighter) than the surrounding area. Hence, a quantifying quality metric had to be set to categorize any input. For that we used the manual Thermal Imagining Quality Rating (TIQR) algorithm, where we defined six main criteria. These are based on observations from several pilot studies, where we captured one or two hands and extracted the skeletons. Then we matched visually what we obtained compared to the original input. Each criterion is assigned a numerical value. Thus,

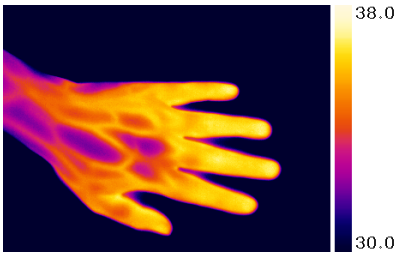
Nr.	Condition title	Condition description	Fig.	Rating
1	High contrast	The user's vein pattern is clearly visible and distinguishable from the surrounding area.	5a	1
2	Medium contrast	The user's vein pattern is barely visible. This condition can apply in conjunction with condition Nr. 1.	5b	1
3	Fingers visible	The user's fingers are clearly visible in the image and not cropped off or dark.	5e	1
4	Complete contour	The shape of the hand within the false-color image resembles the true form of a regular hand. No parts are cropped off. The hand is not twisted in any way.	5c	1
5	Dark vein pattern	The vein pattern appears dark instead of bright in the false-colour image.	5f	-4
6	Wrong ROI	Another warm object in the field of view of the thermal imager was falsely identified as a hand, thus the region of interest is not the user's back of the hand.	5d	-4

Tab. 1: The observed criteria that can characterize the content of a false-colour thermal image. The Thermal Image Quality Rating, ("TIQR") is the sum of its conditions, but at least zero.

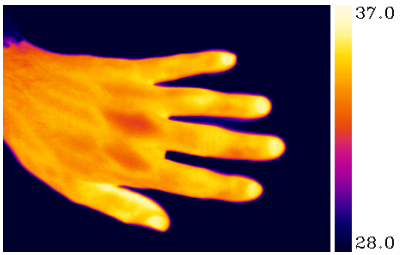
the TIQR value would be the summation of all the existing criteria ratings. Also, if the final TIQR value was found to be negative, it would be rounded up to 0.

By doing that, we ended up with four main categories of images, where only the categories having a TIQR of "3" or "4" is furthermore processed for vein pattern detection. All the data, that was marked for processing was subsequently classified into one of three groups. The first group is the "registered" data set. The data in this set was provided to the system and used to train an "AlexNet". It consists of, in total, twelve distinct skeletons, where each skeleton originates from a different person's hand. The second group is the "testing" data set. It consists of six elements, where each one has a counterpart in the "registered" data set. These are the skeletons extracted from the second recording session of the same registered hand. The third group is formed only of the skeletons of not-registered hands and is used also for testing, referred to as the "unknown-testing" data set.

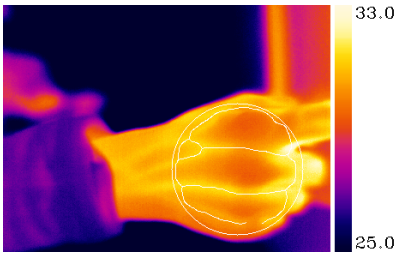
For the evaluation, the "testing"- and "unknown-testing"-skeletons are tested against the authentication system. Dependent on the input and the system's output, the true positive rate ("TPR"), true negative rate ("TNR"), false positive rate



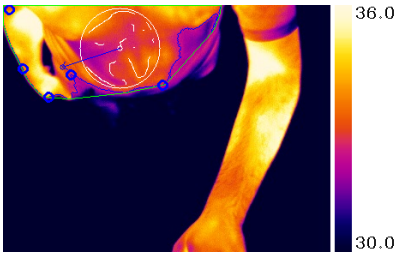
(a) An example of a vein pattern with high contrast.



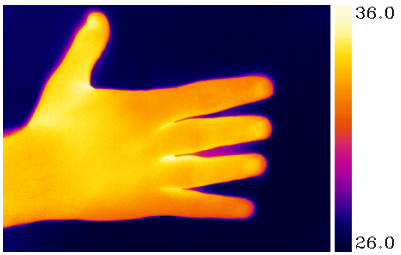
(b) An example of a vein pattern with medium contrast.



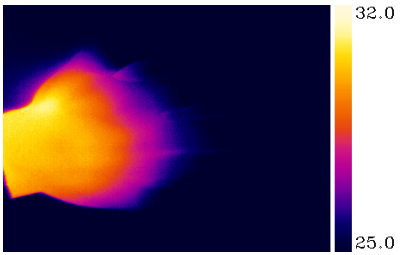
(c) An example of a slightly twisted hand. The contour is merged with a public display in the background.



(d) An example, where the vein pattern was detected in the wrong area within the captured image. Thus the region of interest is misplaced.



(e) An example of a constant temperature hand, where no vein pattern appears.



(f) The thermal condition is so imbalanced, that not even a dark vein pattern is visible, as the fingers are missing.

Fig. 5: Examples of the six image-quality rating criteria mentioned in Table 1.

("FPR") and false negative rate ("FNR") are determined. Ideally, all skeletons in the "unknown-testing" data set should be rejected by the system and become a true negative ("TN"), as they have no counterpart in the systems' "registered" data set. A false negative ("FN") is met, when an element of the "testing" data set is incorrectly rejected, since each element should have been accepted instead. A true positive ("TP") is counted, if an element of the "testing" data set is correctly accepted by the system. Last but not least, a false positive ("FP") is met, when an element of the "unknown-testing" data set is accepted by the system. As this data set does not have any counterparts in the "registered" data set, all of them should be rejected instead.

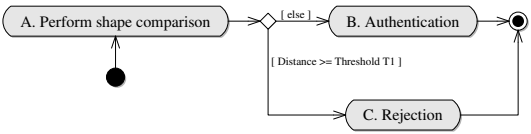
Figure 6 depicts the three evaluated variants of how the authentication system can be designed. The first variant (i.e., *Variant I*), depicted in Figure 6a, only uses a shape comparison algorithm as a measure of similarity between an input skeleton and any of its registered skeletons. The user's identity is assumed as the registered skeleton with the closest distance to the input. If the measure of similarity, the "distance", is below the threshold T1, the authentication attempt is successful, otherwise it is rejected.

The second variant (i.e., *Variant II*), which is illustrated in Figure 6b, additionally performs an image classification to detect the user's identity. For the classification, the pre-trained AlexNet is utilized, which states a top prediction and an associated certainty. If this certainty is below T2, the user is rejected. Otherwise, the user's identity is assumed to be the prediction of the neural network. The threshold T2 is set to a constant value of 75%.

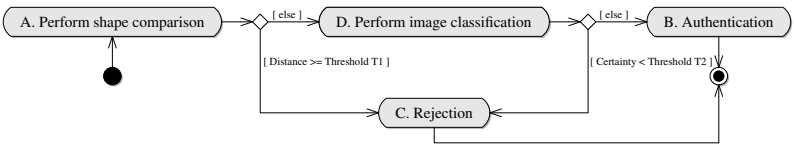
The third variant (i.e., *Variant III*), switches the process and exchanges the activities of image classification and shape comparison and is depicted in Figure 6c. Here, the neural network filters the registered data for a set of potential matches. Then, the shape comparison is applied to compare the input against the best matching image of this set only.

As a measure of comparison between the different approaches to utilize the system, we chose the accuracy ("ACC", see Equation 1a). To find the best suitable values for the threshold T1, we used an optimizer, which calculates the accuracy for every present shape comparison value. The threshold value, which lead to the highest accuracy per method, was then set as the constant value of T1.

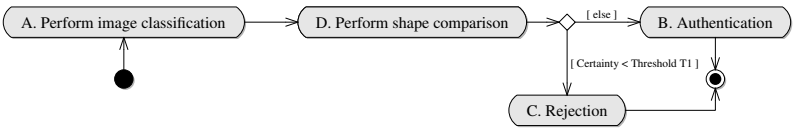
$$\mathrm{ACC} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{P} + \mathrm{N}} \tag{1a}$$



(a) Variant I.



(b) Variant II.



(c) Variant III.

Fig. 6: An activity chart depicting all three variants of the authentication system.

4.3 Results

All three variants, in which the system can be deployed, are depicted in Figure 7, when skeleton stacking is not applied. The bar chart reports the accuracy per shape comparison algorithm. In contrast, Figure 8 depicts the variants, when skeleton stacking is applied.

Skeleton stacking improved the performance, when the Hausdorff distance was used as a shape comparison algorithm. Here, combining skeleton stacking and Variant III. leads to an accuracy of 100%, i.e. the system did not make any errors, when the threshold T1 is set to a value between 9.48 to 10.63. In any other case, skeleton stacking did not improve the accuracy. When no skeleton stacking is applied, the shape comparison methods “Shape Context” and “Jaccard” worked best, yielding the highest accuracies in direct comparison to the Hu-methods and the Hausdorff distance. The Variant III, with the application of skeleton stacking and the Hausdorff distance as a shape comparison algorithm, was able to outperform any other combination, by reaching an equal error rate (“EER”) of zero. Out of the six skeletons in the “testing” data set, all were accepted, while all six skeletons in the “unknown-testing” data set were correctly rejected. When applying the same threshold boundaries of 9.48 to 10.63 to Variant I and II, utilizing skeleton stacking and the Hausdorff method, both reported between four and six false positives and no false negatives.

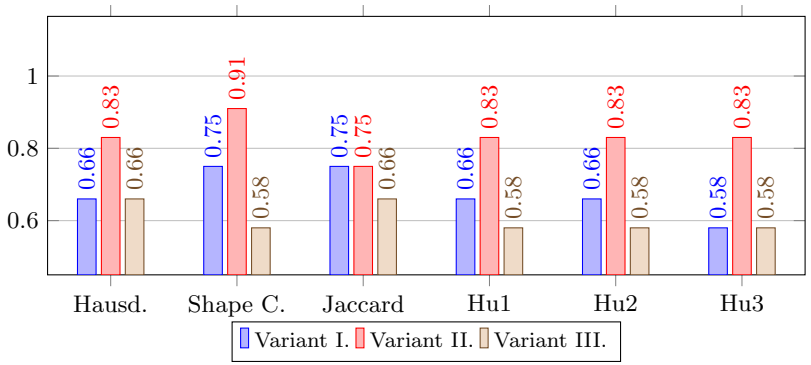


Fig. 7: Depiction of the maximum accuracy per variant, when skeleton stacking is not applied.

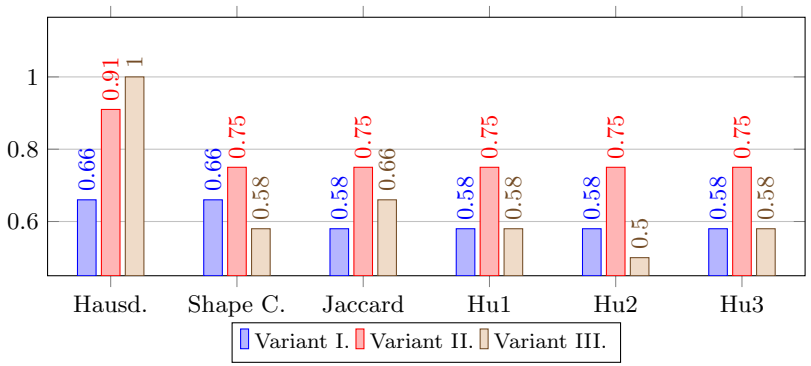


Fig. 8: Depiction of the maximum accuracy per variant, when skeleton stacking is applied.

Variant II, which utilizes the CNN besides the shape comparison algorithms, increased the accuracy in any case, except for the Jaccard-distance shape comparison algorithm without the application of skeleton stacking. A single false positive made the difference between Variant II and Variant III, when the Hausdorff distance was applied without the stacking of skeletons.

As a consequence, the combination of CNNs with the traditional shape comparison algorithms is beneficial. They work in general best, if the shape comparison algorithm first performs a rejection (Variant II), but can exceed their performance, if this process is switched (Variant III). In any case, shape comparison algorithms are useful to tweak the performance of deep learning methods. CNNs also states a certainty value besides its predictions, which can be utilized as a rejection criterion similarly to the shape comparisons. From our data, we saw, that for a learned

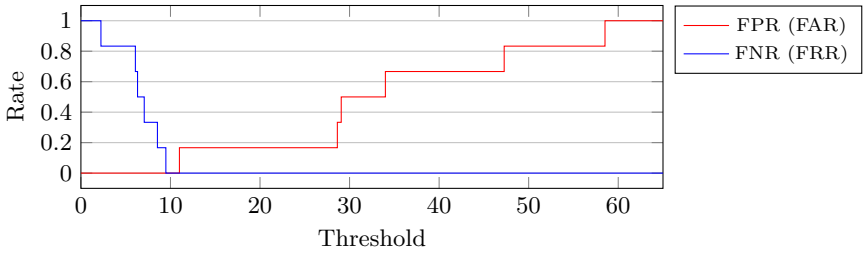


Fig. 9: The FAR/FNR-chart for the authentication system using using Variant III, skeleton stacking and the Hausdorff distance as a shape comparison algorithm.

skeleton as an input, the certainty very often was above 90% and many times close to 100%. In contrast, unknown data, often ranged in a certainty interval of 40% to 50%. The feasibility of certainty values needs to be investigated in future research.

5 Limitations and Future Work

An important factor influencing the accuracy of the VPID system is the image quality. If the image is blurry, the system is unable to extract the vein pattern. Given that the VPID system is designed to implicitly and continuously identify and authenticate users, it can discard images that do not provide proper vein patterns. This, however, requires automatic detection of such images. While we manually rated the images, future work needs to investigate how this can be automated.

External influences such as the room temperature or physical activity influence the clarity of the vein pattern. In cases in which the vein pattern is not fully visible, authentication becomes challenging. Understanding the influence of the context on the clarity of the vein pattern is an important aspect that can make vein-based authentication feasible. Using a machine learning model that takes these differences into account might further improve authentication accuracy.

During the controlled study, we have tested the system with only one thermal imager, which limited the position and orientation the participant's hand could have in the recorded frame. In order to allow a greater degree of freedom for the participant's hand, the system could be used with multiple thermal imagers, that cover more angles. An implementation and evaluation of an authentication system with more than one thermal imager needs to be performed in future work.

6 Conclusion

Relying on biometric features as replacement of passwords and tokens enhances the convenience of authentication. Users usually interact with devices like laptops and interactive tabletops without occluding the back of their hand. Thus, deploying a thermal camera, which currently becomes affordable and small, to capture their veins pattern allows user identification and authentication in a contactless and continuous manner.

Such implicit identification and authentication provides advantages, particularly from a usable security point of view. This is due to the fact that the additional identification or authentication step is not loaded onto the user and, thus, the user can interact with the system without such intermediate step.

To achieve this, we presented the VPID system consisting of a recognition pipeline and an authentication mechanism. We conducted a user study with 12 participants to evaluate the pipeline and different strategies for authentication. The results of the conducted study, highlight the potential of using veins patterns. We show that it is possible to distinguish users using their vein pattern. For that we used six different algorithms. Our results shows that the best used algorithm was the Hausdorff distance after applying CNN of stacked images.

References

[1] Vein recognition in europe. Biometric Technology Today (2004).

[2] Abdelrahman, Y., Khamis, M., Schneegass, S., and Alt, F. Stay cool! understanding thermal attacks on mobile-based user authentication. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (New York, NY, USA, 2017), CHI '17, ACM, pp. 3751–3763.

[3] Abdelrahman, Y., Sahami Shirazi, A., Henze, N., and Schmidt, A. Investigation of material properties for thermal imaging-based interaction. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (New York, NY, USA, 2015), CHI '15, ACM, pp. 15–18.

[4] Al Abdulwahid, A., Clarke, N., Furnell, S., Stengel, I., and Reich, C. The current use of authentication technologies: An investigative review. In Cloud Computing (ICCC), 2015 International Conference on (April 2015), pp. 1–8.

[5] Anand, J., Flora, T. A., and Philip, A. S. Finger-vein based biometric security system. International Journal of Research in Engineering and Technology eISSN (2013), 2319–1163.

[6] Belongie, S., Malik, J., and Puzicha, J. Shape matching and object recognition using shape contexts. IEEE Trans. Pattern Anal. Mach. Intell. 24, 4 (Apr. 2002), 509–522.

[7] Cho, K. W., Lin, F., Song, C., Xu, X., Gu, F., and Xu, W. Thermal handprint analysis for forensic identification using heat-earth mover's distance. In 2016 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA) (Feb 2016), pp. 1–8.

[8] Corner, M. D., and Noble, B. D. Zero-interaction authentication. In Proceedings of the 8th Annual International Conference on Mobile Computing and Networking (New York, NY, USA, 2002), MobiCom '02, ACM, pp. 1–11.

[9] Coventry, L., De Angeli, A., and Johnson, G. Usability and biometric verification at the atm interface. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (New York, NY, USA, 2003), CHI '03, ACM, pp. 153–160.

[10] Crawford, H. A. A framework for continuous, transparent authentication on mobile devices. PhD thesis, University of Glasgow, 2012.

[11] Cross, J., and Smith, C. Thermographic imaging of the subcutaneous vascular network of the back of the hand for biometric identification. In Security Technology, 1995. Proceedings. Institute of Electrical and Electronics Engineers 29th Annual 1995 International Carnahan Conference on (Oct 1995), pp. 20–35.

[12] CSID. Consumer survey: Password habits, a study among american consumers, 2012.

[13] De Luca, A., Hang, A., Brudy, F., Lindner, C., and Hussmann, H. Touch me once and i know it's you!: Implicit authentication based on touch screen patterns. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (New York, NY, USA, 2012), CHI '12, ACM, pp. 987–996.

[14] Fernández, R., and Armada, M. Multisensory system for the detection and localization of peripheral subcutaneous veins. Sensors 17, 4 (2017), 897.

[15] Florencio, D., and Herley, C. A large-scale study of web password habits. In Proceedings of the 16th International Conference on World Wide Web (New York, NY, USA, 2007), WWW '07, ACM, pp. 657–666.

[16] Francis, L., Mayes, K., Hancke, G., and Markantonakis, K. A location based security framework for authenticating mobile phones. In Proceedings of the 2Nd International Workshop on Middleware for Pervasive Mobile and Embedded Computing (New York, NY, USA, 2010), M-MPAC '10, ACM, pp. 5:1–5:8.

[17] Fujitsu, I. Palmsecure-sl, 2012.

[18] Gtattersall. Choosing the appropriate colour palette for thermal imaging in animals, Mar 2016.

[19] Holz, C., Buthpitiya, S., and Knaust, M. Bodyprint: Biometric user identification on mobile devices using the capacitive touchscreen to scan body parts. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (New York, NY, USA, 2015), CHI '15, ACM, pp. 3011–3014.

[20] Hu, M.-K. Visual pattern recognition by moment invariants. IRE transactions on information theory 8, 2 (1962), 179–187.

[21] Huttenlocher, D. P., Klanderman, G. A., and Rucklidge, W. A. Comparing images using the hausdorff distance. IEEE Trans. Pattern Anal. Mach. Intell. 15, 9 (Sept. 1993), 850–863.

[22] Jaccard, P. The distribution of the flora in the alpine zone. 1. New phytologist 11, 2 (1912), 37–50.

[23] Jain, A., Hong, L., and Pankanti, S. Biometric identification. Commun. ACM 43, 2 (Feb. 2000), 90–98.

[24] Jain, A. K., Ross, A., and Pankanti, S. Biometrics: a tool for information security. Information Forensics and Security, IEEE Transactions on 1, 2 (June 2006), 125–143.

[25] Krizhevsky, A., Sutskever, I., and Hinton, G. E. Imagenet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems 25, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105.

[26] Krizhevsky, A., Sutskever, I., and Hinton, G. E. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (2012), pp. 1097–1105.

[27] Kumar, A., and Prathyusha, K. V. Personal authentication using hand vein triangulation and knuckle shape. Image Processing, IEEE Transactions on 18, 9 (Sept 2009), 2127–2136.

[28] Kurkovsky, S., and Syta, E. Digital natives and mobile phones: A survey of practices and attitudes about privacy and security. In Technology and Society (ISTAS), 2010 IEEE International Symposium on (June 2010), pp. 441–449.

[29] Larson, E., Cohn, G., Gupta, S., Ren, X., Harrison, B., Fox, D., and Patel, S. Heatwave: Thermal imaging for surface user interaction. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (New York, NY, USA, 2011), CHI '11, ACM, pp. 2565–2574.

[30] LeCun, Y., Bengio, Y., and Hinton, G. Deep learning. nature 521, 7553 (2015), 436.

[31] LeCun, Y., Bottou, L., Bengio, Y., Haffner, P., et al. Gradient-based learning applied to document recognition. Proceedings of the IEEE 86, 11 (1998), 2278–2324.

[32] Pfeuffer, K., Geiger, M. J., Prange, S., Mecke, L., Buschek, D., and Alt, F. Behavioural biometrics in vr: Identifying people from body motion and relations in virtual reality. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (New York, NY, USA, 2019), CHI '19, ACM, pp. 110:1–110:12.

[33] Sae-Bae, N., Ahmed, K., Isbister, K., and Memon, N. Biometric-rich gestures: A novel approach to authentication on multi-touch devices. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (New York, NY, USA, 2012), CHI '12, ACM, pp. 977–986.

[34] Sahami Shirazi, A., Abdelrahman, Y., Henze, N., Schneegass, S., Khalilbeigi, M., and Schmidt, A. Exploiting thermal reflection for interactive systems. In Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (New York, NY, USA, 2014), CHI '14, ACM, pp. 3483–3492.

[35] Schneegass, S., Oualil, Y., and Bulling, A. Skullconduct: Biometric user identification on eyewear computers using bone conduction through the skull. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (New York, NY, USA, 2016), CHI '16, ACM, pp. 1379–1384.

[36] Simonyan, K., and Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014).

[37] Sonkamble, S., Thool, D. R., and Sonkamble, B. Survey of biometric recognition systems and their applications. Journal of Theoretical & Applied Information Technology 11 (2010).

[38] Van Tien, T., Mien, P. T., Dung, P. T., and Linh, H. Q. Using near-infrared technique for vein imaging. In 5th International Conference on Biomedical Engineering in Vietnam (2015), Springer, pp. 190–193.

[39] Wang, J.-G., Yau, W.-Y., Suwandy, A., and Sung, E. Person recognition by fusing palmprint and palm vein images based on "laplacianpalm" representation. Pattern Recognition 41, 5 (2008), 1514 – 1527.

[40] Wang, L., and Leedham, G. A thermal hand vein pattern verification system. In Pattern Recognition and Image Analysis. Springer, 2005, pp. 58–65.

[41] Wood, H. M. The use of passwords for controlled access to computer resources, vol. 500. US Department of Commerce, National Bureau of Standards, 1977.

[42] Zhang, T., and Suen, C. Y. A fast parallel algorithm for thinning digital patterns. Communications of the ACM 27, 3 (1984), 236–239.