JOHANNES KEPLER
UNIVERSITY LINZ

Author
Richárd Matuch
12349155

Submission
Institute of
Networks and Security

First Supervisor
Univ.-Prof. DI Dr.
René Mayrhofer

Second Supervisor
Dr. **Gergely Kovásznai**

Assistant Thesis
Supervisors
DI Dr. **Philipp Hofer**, BSc
Ádám Kovács, MSc

September 2025

Hand based vascular authentication on embedded devices

Master's Thesis

to confer the academic degree of

Diplom-Ingenieur

in the Master's Program

Computer Science

EGER 1774
ESZTERHÁZY KÁROLY CATHOLIC UNIVERSITY

JOHANNES KEPLER
UNIVERSITY LINZ
Altenberger Straße 69
4040 Linz, Austria
jku.at

Abstract

The human body has many unique biological identifiers, which can be used to differentiate an individual from the masses: face, fingerprints, or even the vein patterns under the skin. This property makes biometric authentication a highly viable and increasingly popular solution in modern security systems. Among these methods, vein recognition stands out as a particularly promising approach. Unlike traditional touch-based biometric methods such as fingerprint scanning, vein recognition does not necessarily require physical contact with the sensor. This makes it more hygienic, a factor that has gained even greater importance in the wake of global health concerns and the growing emphasis on reducing the spread of infectious diseases.

Vein recognition systems are more secure than some other biometric methods because vein patterns lie under the skin and are invisible to the naked eye. This makes them difficult to forge or replicate compared to surface-level identifiers like fingerprints or facial features, which can sometimes be copied or spoofed. A further advantage is that the method does not leave traces behind of the user, even when the capturing method is not fully contactless. In contrast, fingerprints can be lifted from surfaces and potentially misused, while facial images could be captured without consent.

Capturing vein patterns requires a specialized imaging device designed to highlight the unique vascular structures under the skin. By illuminating the targeted area and filtering out unwanted wavelengths, the device can isolate the vein patterns and capture them with high resolution. This level of detail is essential to ensure that the extracted vein maps are distinctive enough for reliable biometric authentication. Capturing high-quality images is only the first step, efficient processing and analysis are equally critical to build a practical biometric system.

An embedded device is utilized to capture the vein patterns, while keeping the form factor compact. Since these devices are small computers, they are capable of performing not only the high quality image acquisition but also the computationally intensive tasks of feature extraction and even the matching steps of the authentication process. This makes the whole system more self-contained, reducing the reliance on external servers or dedicated hardware. In practice, such integration enhances both portability and security, as sensitive biometric data does not need to leave the device for processing. These systems can achieve real-time performance, making them suitable for applications ranging from personal gadgets to access control in secure facilities. The compact design also enables seamless integration into everyday objects, thus improving usability and user adoption.

Acknowledgements

This work has been carried out within the scope of Digidow, the Christian Doppler Laboratory for Private Digital Authentication in the Physical World. We gratefully acknowledge financial support by the Austrian Federal Ministry of Economy, Energy and Tourism, the National Foundation for Research, Technology and Development, the Christian Doppler Research Association, 3 Banken IT GmbH, ekey biometric systems GmbH, Kepler Universitätsklinikum GmbH, NXP Semiconductors Austria GmbH & Co KG, and Österreichische Staatsdruckerei GmbH.

Contents

Chapter 1

Introduction

The introduction of authenticating users based on their biological features is a big stepping stone in authentication. Biometric recognition [39] is more user friendly, since it does not require remembering a sequence, where a prolonged period of not using the phrase can make the user forget it. Since biometric features, such as fingerprints or facial characteristics, are known to maintain their properties over time, they provide a more reliable and consistently available alternative to memory-based authentication.

For authentication, the user might choose a weak password, which can be easily brute forced. Biometric authentication does not inhibit this property, since the complex nature of the biological feature is almost impossible to brute force. Some biometric features leave traces behind (fingerprint on a surface), or is easy to capture without consent (face). These can be later used to forge a template to mimic the identity of that person. This could be mitigated by choosing an authentication vector, that does not leave traces behind.

Vein recognition [75] is a novel method in the topics of biometrics. It captures the pattern of the blood vessels, which is then used to make the template representing the user. This method is contactless, just like face recognition, making it more hygienic compared to methods requiring contact for scans. Due to the combination of the capture method and the biological properties, it is harder to forge vein patterns than face or fingerprints. The cost of manufacturing a vein capturing device might be more expensive than a camera used for face recognition or an optical sensor used for fingerprinting.

The properties of vein recognition like high spoof resistance, contactless form, and complicated unauthorized access to the vein features it is often used where a high security authentication form is required.

Goals of the thesis

This thesis aims to explore the current state of the vein biometrics. This involves an investigation into various state-of-the-art methods presented in academic literature and research. Based on these findings, various practical solutions will be implemented and evaluated, including both traditional and machine learning methodologies. For this purpose, a sufficient dataset is required.

An ultimate goal of the thesis is to create a self-contained authentication system for vein recognition. For this, a working prototype will be made. This system would be responsible for the acquisition and processing of the images, as well as the authentication process. The chosen platform for this system will be an embedded device, such as a Raspberry Pi, due to its versatility and cost-effectiveness.

Outline

To achieve the goals and to get a good understanding in the topic, we first have to go through the technical details of vein recognition. Chapter 2 discusses the the theoretical background of this topic, including the variations of vascular identification and its physical and biological properties. After which the metrics used for evaluating authentication systems will be discussed. This will follow the data preprocessing techniques typically used in these scenarios. Finally finishing the chapter with the traditional recognition methods used for these tasks.

This will follow Chapter 3, where we will review state-of-the-art capture devices, including proprietary solutions as well and their datasets, in order to explore different solutions to identify potential hardware and a sufficient dataset for this project.

Based on the findings Chapter 4 will go through the development and experimentation of the system, that is responsible for capturing, preprocessing and authentication. This includes the relevant hardware, the prototypes and the arising issues during the experiments as well as their possible solutions.

Chapter 5 analyzes and evaluates various authentication models, including traditional and machine learning methods. The problem arising during the training of the neural network, and a working solution for it will be discussed. The performance of each model will be assessed to determine the most suitable candidate for the embedded system.

Chapter 6 follows with the presentation of the final integrated authentication system. The system's architecture and the integration of the chosen authentication model and data processing techniques are detailed, showing how they result in a cohesive, functional unit.

Lastly, Chapter 7 concludes the thesis by summarizing the key findings, and proposing potential improvements for future work.

Chapter 2

Background

This Chapter provides an overview of vascular-based biometric recognition, covering the fundamental concepts, technical requirements and key processing techniques. It begins by categorizing the main types of vascular identification and discusses the unique properties of vascular patterns that make them suitable for reliable biometric recognition, such as their uniqueness and light absorption characteristics.

It then addresses the technical aspects of image acquisition, including the requirements for cameras and lighting, followed by the metrics commonly used to evaluate biometric systems. Next, it explores data preprocessing techniques which are important to enhance image quality, for example with the use of histogram equalization.

Then it examines keypoint detection and feature description methods, and discusses descriptor matching techniques, which form the backbone of accurate traditional vascular recognition systems. Together, these topics lay the basics to understand and implement a vascular biometric identification system.

2.1 Types of vascular identification

Vein recognition can be performed on multiple parts of the body, with the choice of region being largely influenced by criteria such as ease of access, usability, and practicality for biometric applications. Based on these considerations, vascular imaging is typically focused on three primary areas: the eye, the fingers, and the hand. The following subsections provide an overview of these approaches, highlighting typical usage in the context of vascular biometric systems.

2.1.1 Eye based

The eye has two popular features that can be used for authentication: the iris [7] and retina. [80] The iris is a muscle in the eye that controls the amount of light let on the retina. The iris's color depends on the amount of melanin pigment it contains, which also determines the color of the eye. When capturing in the infrared spectrum, the color difference disappears. To use the iris as the authentication method, we have to take an image or video of the eye from around 10 cm to a maximum of 1 m. This distance is mainly influenced by the resolution of the camera sensor.

To photograph the retina, we have to use a special scanner. When looking through the scanner's eye piece, the device shoots low energy infrared light to the retina. Since the retina consists of blood vessels, alongside with neural cells, the light is absorbed by the hemoglobin in the blood.

2.1.2 Finger based

The vascular patterns in our fingers could also be used for authentication just as fingerprints. Finger vein senors use infrared lighting in two different ways. [79] The light can be assembled on the same side as the camera, capturing the reflected light, or the opposite side of the hand, capturing the trans-illuminated picture of the finger. The latter method usually uses some kind of box apparatus, that the user puts their finger in.

2.1.3 Hand based

Hand biometrics typically employs the palm's vascular pattern as an authentication method. [63] The scanner device is using an infrared light source to enhance the visibility of the vascular patterns. Most scanners use the reflected light, instead of trans-illuminating the hand, as the thickness and underlying bone and muscle structure would block the light.

The back side of the hand can also be used for recognition. The dorsal vein pattern can be photographed the same way as the palm. [66] The most prominent vascular features are the dorsal veins, but during capturing the body hair on the dorsal side might interfere with the recognition process. Due to this factor dorsal capturing is less popular.

2.2 Vascular pattern properties

Vascular patterns are the structures formed by blood vessels, which transport blood cells, nutrients, and oxygen to the organs, while carrying waste products and carbon dioxide away from the tissues. The vascular system consists of three main types of blood vessels: arteries, veins, and capillaries. Arteries deliver oxygen-rich blood from the heart to the body, and veins return deoxygenated blood back to the heart. The exchange of oxygen and other substances occurs through the thin walls of the capillaries, which connect arteries and veins. [62] Beyond their physiological function, vascular patterns exhibit unique characteristics that make them suitable for biometric recognition.

2.2.1 Uniqueness

In 1935 Simon and Goldstein discovered that the vascular pattern with respect to the individuals are unique. [65] Later in 1955 Paul Tower confirmed this by studying the differences of these patterns between twins. [73] This uniqueness can also be found not only in the retina, but throughout the whole body. [5] This biological property can be used to differentiate individuals from one another, or to re-identificate users.

2.2.2 Light absorption

Most vascular biometric applications use infrared sensitive camera sensors with infrared lighting. [14] These devices are using the light absorbing property of the hemoglobin in our red blood cells. Hemoglobin is the primary transport molecule of oxygen. The blood absorbs different wavelength of infrared

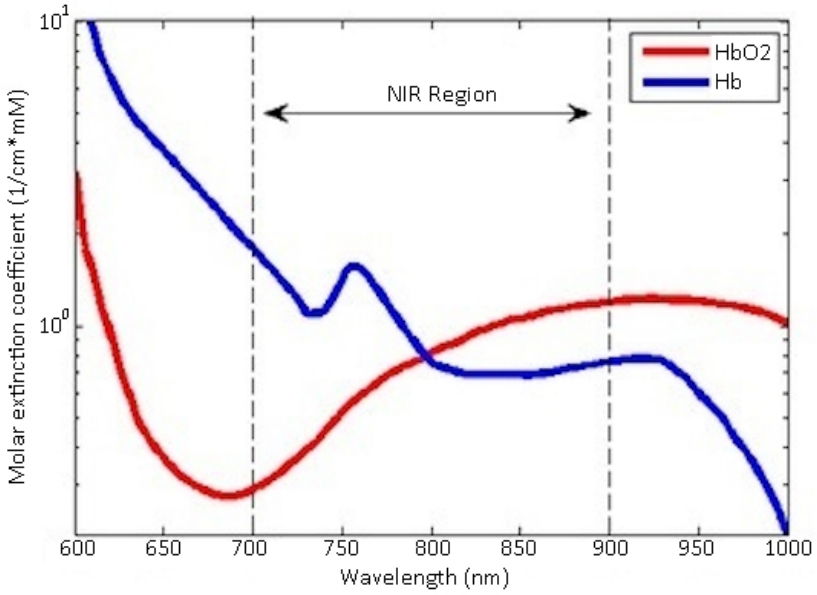


Figure 2.1: Absorption levels of hemoglobin [41]

light, given its oxygen level. The skin and other tissues are slightly translucent to near-infrared and infrared light, thus these types of light can penetrate the into the blood vessels, where the hemoglobin absorbs it.

On Figure 2.1 we can see that the deoxygenated blood can absorb more light on the lower wavelengths, while at 900 nm both type of blood could absorb around the same amount of light. Because of this property, images taken around 650 nm have darker areas only where veins are, while above 750 nm both veins and arteries appear as darkened areas on the skin.

2.3 Technical requirements in vein image capturing

Accurate vein recognition relies heavily on high-quality image acquisition. This section outlines the essential technical requirements for capturing vascular patterns, focusing on the critical roles of cameras and lighting. Proper configuration of these components ensure clear visualization of vein structures, which is fundamental for reliable preprocessing, feature extraction, and recognition.

2.3.1 Camera

Current camera sensors can capture in the near-infrared and infrared spectrum. Charge-Coupled Device (CCD) cameras are better to measure the infrared photons hitting the sensor surface, than Complementary Metal-Oxide-Semiconductor (CMOS) cameras. [67] Both monochrome and color cameras can be used for the image capturing, but since most cameras are designed to capture the visible spectrum, the infrared filter must be removed if it is present.

Infrared longpass or bandpass filters can also help reduce the interference from visible light. [67] These filters must have the properties of letting through the preferred lighting that are used in the capture device.

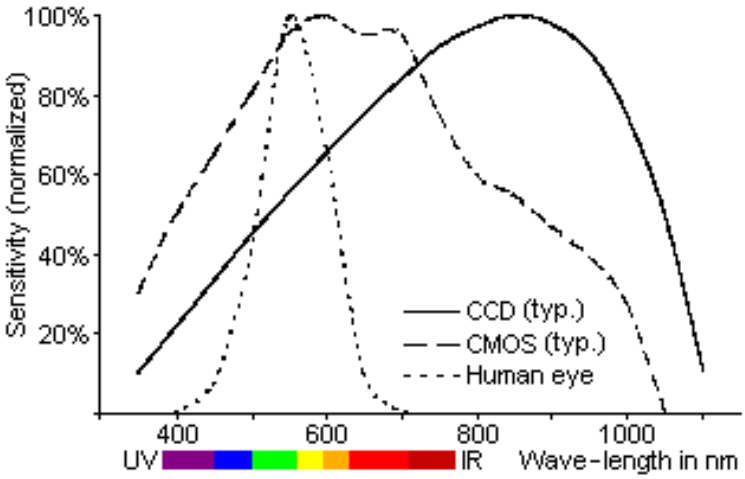


Figure 2.2: Spectral sensitivity curves [4]

2.3.2 Lighting

Most capturing devices use a infrared or near-infrared lighting source. Figure 2.1 shows which variation of light source can be used on a wavelength basis. A light source emitting around 700 nm gives us a good absorption level around the veins. [6] We can also seen on Figure 2.2 that the emitted light may still be visible to the human eye.

With a light source emitting between 750 and 925 nm, the light is completely invisible to the human eye. [9, 78] In this range, the emitted light is absorbed by the oxygen saturated and unsaturated blood.

When we use an optical filter, we must consider that it might block the light source or the reflected light. The wavelength emitted from the lighting must be around the same wavelength, which the filter is transparent to.

2.4 Metrics in biometric recognition

Evaluating the performance of a biometric system requires quantitative metrics that reflect its accuracy, reliability, and efficiency. This Section introduces the key measures commonly used in biometric recognition.

		Reality	
		Positive	Negative
Study Finding	Positive	**True Positive** (Power) (1-β)	False Positive **Type I Error** (α)
	Negative	False Negative **Type II Error** (β)	**True Negative**

Figure 2.3: Statistical Hypotheses and Error [76]

False acceptance rate [45] measures the likelihood of the system incorrectly accepting an unauthorized user. Lowering this metric lowers the security risk. False acceptance rate is a type I error (Figure 2.3). It measures the correctness when two non-matching templates are given to the system.

$$\text{False Acceptance Rate} = \frac{\text{Falsely Accepted Users}}{\text{Total number of attempts}} * 100$$

False rejection rate [45] measures the likelihood of the system incorrectly rejecting an authorized user. Lowering this metric improves the user experience. False rejection rate is a type II error (Figure 2.3). It measures the correctness when two matching templates are given to the system.

$$\text{False Rejection Rate} = \frac{\text{Falsely Rejected Users}}{\text{Total number of attempts}} * 100$$

Runtime metrics [74] are time based, which measure the execution time. These metrics are usually hardware dependent, while better optimized algorithms can also reduce the execution time. Lower runtime improves the user experience. Two variations of this metric is one-to-one and one-to-many.

2.5 Data preprocessing techniques

Effective vein recognition depends on high-quality images, which often require preprocessing [17] to enhance visibility and to highlight vascular structures. This section presents common preprocessing techniques, commonly used for images or image based recognition.

2.5.1 Histogram Equalization

The histogram of an image is a curve that describes the amount of pixels given the intensity. Histogram equalization is a process of uniformly distributing the image histogram over the entire axis by choosing a proper intensity transformation.

Let us consider that the intensity levels of the image is r. We limit the values that r can take between 0 and $L-1$, which is $0 \leq r \leq L-1$. Generally $L = 2^m$, where m is the number of bits required to represent the intensity levels. $r = 0$ represents black and $r = L-1$ represents white. Considering an arbitrary transformation function $s = T(r)$ where s denotes the intensity levels of the resultant image.

$$s_k = T(r_k) = (L-1)\sum_{j=0}^{k} p_r(r_j) = \frac{(L-1)}{N^2}\sum_{j=0}^{k} n_j \tag{2.1}$$

Equation 2.1 is the histogram equalization transformation function. [19]

2.5.2 Contrast Limited Adaptive Histogram Equalization

Contrast Limited Adaptive Histogram Equalization (CLAHE) [16] is used to improve the contrast of images. In traditional methods, contrast of whole image changes but CLAHE works by dividing the image into smaller parts and adjust the contrast in each part separately. CLAHE has two parameters:

- Clip limit: This parameter sets the threshold for contrast limiting.
- Tile grid size: This is used to divide the image into grids. It sets the number of rows and columns.

After processing the tiles, it combines them using bilinear interpolation to remove visible boundaries between the tiles.

2.5.3 Thresholding

Thresholding [61] is an image segmentation technique that creates a binary image from a color or grayscale image. It is usually performed on grayscale images. Each pixel of the image is compared with the threshold: if it is less than the threshold, the intensity of the pixel is set to 0, otherwise it is set to the maximum possible value.

2.5.4 Hessian Filter

The Hessian Filter [48] is applied at an image pixel level and is based on the eigenvalue decomposition of the local Hessian matrix of the image. By selecting specific eigenvalues and defining a vesselness function, the contrast of vessels is enhanced, while non-vascular structures and background noise are suppressed. The filter can be fine tuned by adjusting the alpha, beta and gamma parameters.

2.6 Keypoint detectors and Descriptors

Accurate feature extraction is crucial for reliable biometric recognition. This section introduces keypoint detection and description techniques.

2.6.1 Oriented FAST and Rotated BRIEF

Oriented FAST and Rotated BRIEF (ORB) [58] is a fusion of the Features from Accelerated Segment Test (FAST) keypoint detector and the Binary Robust Independent Elementary Features (BRIEF) descriptor, with modifications aimed at enhancing performance.

It uses the FAST algorithm to find keypoints, then applies a Harris corner measure to fint top N points among them. The FAST algorithm was modified to also compute orientation of the points. For improved rotation, invariance moments are calculated using the coordinates and the radius of the patch.

The modified BRIEF algorithm in ORB performs better with rotation. It has an important property, where each bit feature has a large variance and a mean near 0.5. But once it is oriented along keypoint direction, it loses this property and becomes more distributed. High variance makes a feature more discriminative, since it responds differently to inputs. BRIEF's descriptors are also encoded in binary making computations faster.

2.6.2 Scale-Invariant Feature Transform

Scale-Invariant Feature Transform (SIFT) [57] is a keypoint detector and descriptor. It can detect keypoints in a way, that their descriptors do not change with different scales. It uses Difference of Gaussians, which is an approximation of Laplacian of Gaussians, making it less costly. This help detecting scale-space extrema.

When the potential keypoints are found, it uses Taylor series expansion of scale space to get more accurate location of extrema. Any keypoint whose intensity is below a certain threshold is then rejected.

A keypoint descriptor is formed by first sampling a 16×16 neighborhood around the point. This area is then subdivided into 16 sub-blocks of 4×4 size, and an 8-bin orientation histogram is created for each sub-block. The histograms are then combined into a single vector, resulting in a descriptor that is robust to illumination and rotation changes. These descriptors are also encoded as floating-point numbers.

2.7 Descriptor Matching

Brute-Force matching [56] is one of the simplest feature matching algorithms used in image recognition. With this approach, the descriptor of each keypoint in the first set is compared with the descriptors of all keypoints in the second set. A distance metric is used to determine the similarity between descriptors, and the closest match is selected. Although it is simple and easy to implement, Brute-Force matching can become computationally expensive as the number of keypoints or the dimensionality of descriptors increases, which may lead to longer processing times for large datasets. Despite this, it remains widely used due to its deterministic behavior and high accuracy in matching distinctive features.

Fast Library for Approximate Nearest Neighbor (FLANN) [56] is a library designed to provide efficient solution for nearest neighbor search in large datasets with high-dimensional features. Unlike Brute-Force matching, which compares every descriptor pair, FLANN uses optimized search algorithms and tree-based structures to approximate the nearest neighbors, reducing computation time. By balancing speed and precision, FLANN provides an alternative to Brute-Force matching for real-time or resource-intensive tasks.

Chapter 3

Related work

This Chapter reviews the existing literature in the field of vascular biometric recognition. It first reviews the sensors and datasets that have been developed to support vein recognition research.

Section 3.1 details the data acquisition methods, imaging technology, and characteristics that make these datasets valuable for developing and evaluating vascular recognition algorithms.

Section 3.2 discusses the off-the-shelf vein recognition sensors and frameworks, developed by the two main companies prevalent in the field of vascular biometrics.

3.1 Hand vein sensors for dataset creation

The following datasets were found during the research for the thesis, from which the Dorsalhandveins dataset (Section 3.1.4) was used for training and evaluation, as it is open source and free to use.

3.1.1 VERA PalmVein

VERA PalmVein [36] is a dataset for palm vein recognition. The set consists of 2200 near-infrared images from 110 client captured with an open sensor. The clients presented the palm side of both hands to the sensor. 5 images were recorded per palm. The recordings were performed at two different locations in a building, with normal lighting conditions. The first 78 clients in the dataset were recorded in the first location, the other remaining 32 were in another.

Figure 3.1: An example image from the VERA PalmVein dataset [72]

The dataset consists of 110 individuals (40 women and 70 men) aged between 18 and 60, with an average age of 33. A separate metadata file is included, which saves this information for each image.

The database contains two different sets: raw and Region of Interest (ROI). The raw consists of images captured with the sensor, whereas ROI consists of images where the region is extracted during the image acquisition process.

The images are encoded in the PNG format, with an image size of 480×680, and every image is around 165 kb in size. The pictures are monochrome, as an example shows on Figure 3.2.

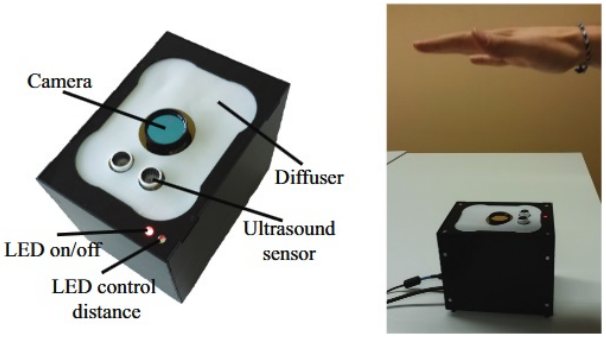


Figure 3.2: VERA PalmVein sensor prototype

During the capturing process, the individuals had to hover their hands over the device. The prototype was equipped with an ultrasound sensor, which gave feedback via an LED. The pictures were taken manually, when the client's palm was in the middle of the picture. The operator also controlled the brightness of the LEDs manually, to achieve a better contrast of the vein pattern. A total of 5 images were collected for each hand, starting with the left and then the right. The average duration for this entire procedure was approximately 5 minutes per participant.

The sensor prototype (Figure 3.2) contains an ImagingSource camera, containing a Sony ICX618 sensor. For lighting, it contains multiple 940 nm emitting wavelength LEDs. The distance between the client's hand and the sensor is measured by an HC-SR04 ultrasound sensor. The project used the Bob library [34] to evaluate the effectiveness of the sensor.

Bob is an open-source signal-processing and machine learning toolbox. It was developed by Biometrics Security and Privacy Group, the Biosignal Processing Group, and the Research and Development Engineers at Idiap. It contains biometric recognition packages, including a vein recognition library. It uses traditional matching algorithms (cross-correlation, hamming-distance and miura-match) and image preprocessing methods. The matching algorithms compare two binary images against each other.

Unfortunately due to legal and financial reasons, this dataset was not used during the development of the thesis.

3.1.2 FYO Multimodal Vein Database

The FYO Multimodal Vein Database [55] consists of palm, dorsal, and wrist images. The database is open for researchers. The images were taken in the Computer Engineering Department laboratories of Eastern Mediterranean University.

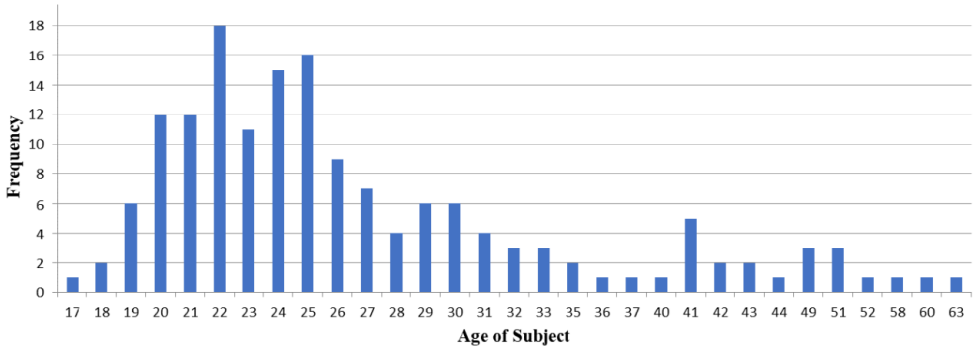


Figure 3.3: The distribution of subjects based on age [55]

The dataset consists of 1920 images acquired from 160 subjects. 2 dorsal, palm and wrist images were taken from both the right and left side. Out of the 160 subjects 49 were women and 111 were men, whose ages were between 17 and 63 (Figure 3.3). The volunteers are mainly from North Cyprus, Turkey, Nigeria, Iran and other parts of the Middle East and Africa.

The images were taken with a CMOS camera. The researchers constructed an LED infrared ring around it. The exact technical details are not discussed in the article. The sensor and the lighting were put into a holder. This holder can rotate at 360°, but for consistency, they chose a fixed downward position. The sensor was kept at a height of 35 cm. The researchers also constructed a wooden hand guide to reduce rotation of the hand. This guide was at a height of 7 cm, making the distance between the sensor and the target approximately 26-28 cm, depending on the captured part of the hand. We can see on Figure 3.4, how the hardware for the vein acquisition was set up.

Figure 3.4: Hardware Setup for Vein Acquisition [55]

The captured images were encoded in PNG format, with a resolution of 800×600. The researchers used both traditional (Binary Statistical Image Features, Gabor-filters, Histogram of Oriented Gradients) and machine learning methods, mainly Convolutional Neural Network (CNN), for evaluating the images.

When using the traditional matching methods, researchers tested different combination for re-identification. The recognition rates improved when using more biological traits.

For the CNN models the researchers used all three types of traits. Each network computed a sub-result. The final result was calculated from the three

sub-result with decision-level fusion.

Unfortunately we were not able to request access to the dataset because of its requirements.

3.1.3 CASIA Multi-Spectral Palmprint Database

The CASIA Multi-Spectral Palmprint Databasse [42] consists of 7200 palm images. Theses images were captured from 100 different subjects. For each subject there were 2 sessions. These sessions were conducted more than one month apart. During the image capture process, researchers took three samples per palm for each lighting option. The article does not specify the ages or sexes of the participants.

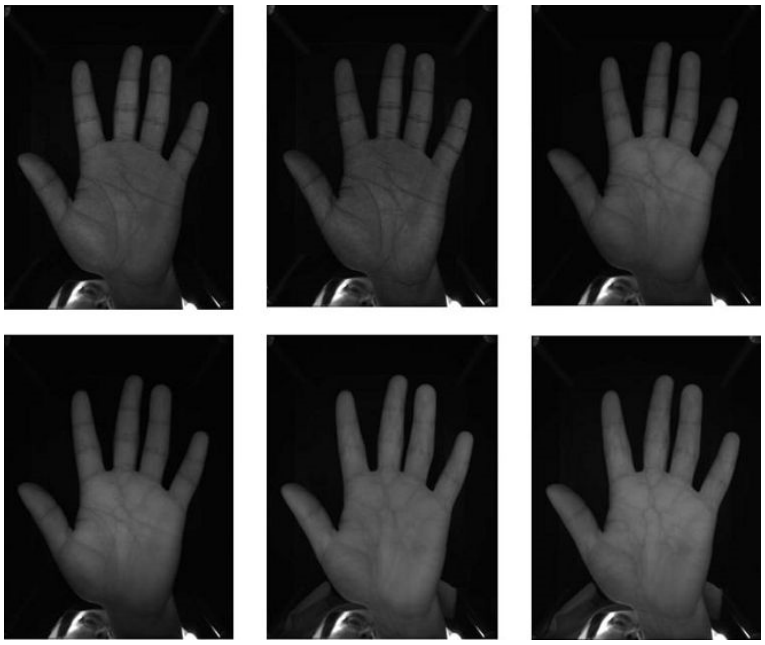


Figure 3.5: Example images from the CASIA database using the 6 light sources [35]

The images were created with a CCD camera. There is no technical information about the sensor in the article. For lighting, researchers used different sources: 460 nm, 630 nm, 700 nm 850 nm, 940 nm and a white light. During image capturing, there were no guides used for the hands. The subjects were required to put their hands into the device. On Figure 3.5 we can see examples of the images captured using the six light sources.

This dataset was not used during the development, because the request was rejected due to unknown reasons.

3.1.4 Dorsalhandveins datasets

For Dorsalhandveins, [11] researchers created two databases for dorsal hand veins in 2007. The participants were students and personnel attending the Pontificia Uviversidad Javeriana. The ages of the participants vary from 18 to 29 years old. The first database comprises 138 individuals. 4 pictures were taken from each hand during the creation process. This dataset contains 1104 images.

For the second dataset only 113 participant attended from the 138. During the image creation for the second database, only 3 pictures were taken from each hand. The second set consists of 678 images.

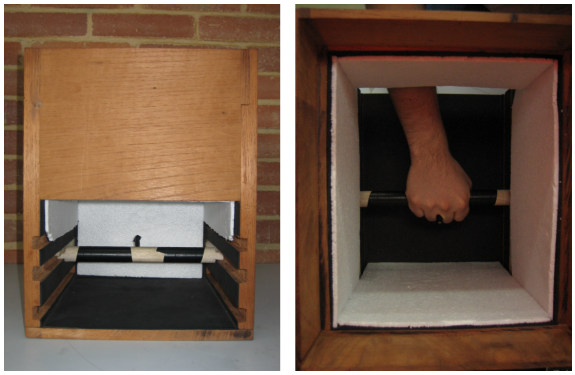


Figure 3.6: The box-like structure and its inside [11]

The participants had to put their hands into a box-like structure, which we can also see on Figure 3.6. This box contained a hand guide rod and the camera with the illumination. The box itself was 32 cm long, 29 cm wide and 35 cm in height. The bottom of the container was covered with black painted polystyrene foam. The distance between the target and the camera was around 20 cm.

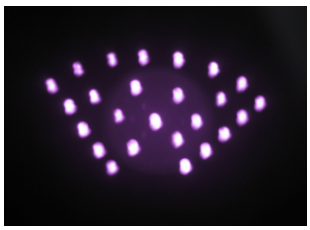


Figure 3.7: An LED module powered on [11]

For lighting, 4 custom made lighting modules were created (Figure 3.7). Each module contained 25 LEDs, emitting light with a centroid wavelength of 880. The range of the emitted wavelength was from 840 to 950. On top of the LED modules, a parchment paper was added as a diffuser.

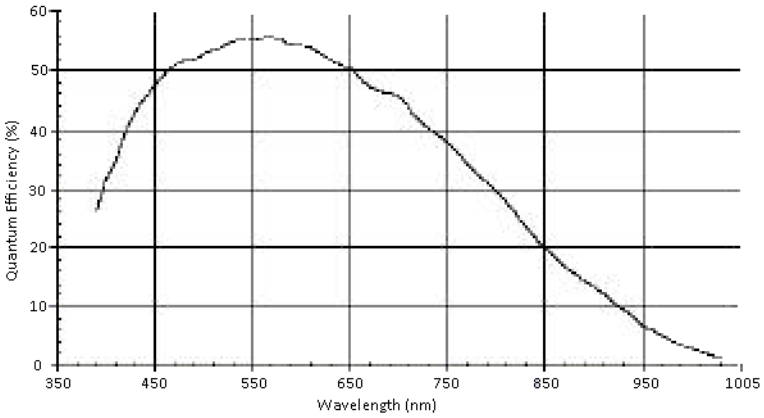


Figure 3.8: Quantum Efficiency of the LightWise 1.3-S camera [38]

The researchers used an ISG LW-1.3-S-1394 monochrome CMOS camera. For

infrared pass filter, a black processed photographic filter (negative) was used. The camera took 10 pictures of each hand, then the program averaged those into a single image. The resolution of the pictures are 752×560. The light sensitivity of the camera sensor can be seen on Figure 3.8.

This dataset was used during the training and evaluation phase of the thesis. It was preprocessed using multiple levels to assess the effectiveness at each step, which will be elaborated further in Section 5.1

3.2 Proprietary solutions

Several commercial organizations have developed off-the-shelf vein recognition systems for secure authentication. Among these, Hitachi and Fujitsu are two of the leading providers, each focusing on different types of vascular-based identification. These solutions are widely adopted in applications where reliable verification of an individual's identity is critical, such as banking, access control, and secure facilities. This section provides an overview of these proprietary technologies.

3.2.1 Hitachi

Hitachi is one of the leading companies in the commercial space of vein recognition. [31] They developed multiple finger vein sensors. Their products can be bought as a standalone sensor, or they could be built into a device – such as an ATM.

Figure 3.9: Hitachi H1 finger vein sensor [32]

The Hitachi H1 finger vein sensor, as seen in Figure 3.9 is an easy to use sensor itself. The sensor has a special case. The bottom part incorporates the infrared sensor and the ports for connecting it to a computer, while the top part has the illumination. The user has to put their finger between the lighting and the sensor plate.

Their software was developed on a closed dataset. It is also not publicly known, how many vein patterns have been used during development. The product has a false acceptance rate of 0.0001% and a false rejection rate of 0.01%. [31]

A bank in Poland has installed ATMs with finger vein sensors [28], which we can also see in Figure 3.10. Users can authenticate themselves with the help of their finger vein pattern.

This technology can not only be used to authenticate users, but it can also be used as a payment method. FingoPay [33] developed in partnership with Hitachi as a new way of paying. With only the finger vein sensor, registered users can pay without needing a separate physical card.

Figure 3.10: ATM with a Hitachi finger vein sensor [28]

3.2.2 Fujitsu

Fujitsu is an other significant company in the field of vein recognition. Their product focuses on palm veins. [49] The sensor itself can be purchased as a standalone device (Figure 3.11), or built into different electrical appliances like notebooks and computer mouse (Figure 3.12).

Figure 3.11: Standalone Fujitsu PalmSecure vein sensor [49]

Figure 3.12: Fujitsu Lifebook U939 (left) and Fujitsu PalmSecure F-Pro Mouse (right) [49]

The user has to hover their hand over the sensor plate. Both the illumination and image sensor is below the plate itself. For more convenience, a guide cone can be used to help position the hand over the device. [40]

Their implementation was also developed on a closed dataset. The amount of palm vein patterns is also not publicly known. They claim a 0.000001% false acceptance rate and a 1% false rejection rate. [40] Since the user does not have to touch the sensor, this authentication method is contactless.

It is mostly used as an authentication method for secured areas or devices. Their implementation can also be used as a payment method. Fulcrum biometrics – a subsidiary of Fujitsu – developed the technology, where users can pay with only their palm. [15] These sensors can be used in order for it to work.

Chapter 4

Sensor development process

In this Chapter, we will see the development of the prototypes, including the hardware. There were multiple iterations, where each prototype used the Raspberry Pi 4B as the image capturing device. A Raspberry Pi Camera V3 is attached to it. The lighting is connected to the Raspberry Pi's GPIO interface, which provides it with power and the control signal.

4.1 Hardware parts of the sensor

The device used is a Raspberry Pi 4B [52], which is a single board computer featuring General Purpose Input Output (GPIO) capability. It has a 64-bit ARM-Cortex A72 processor running at 1.5 GHz. The device has a built in wireless module supporting 2.4 GHz and 5 GHz Wi-Fi. The board also contains a Mobile Industry Processor Interface Camera Serial Interface (MIPI CSI) port, where official or third-party cameras can be mounted via a flat camera cable.

Raspberry Pi Camera Module v3 NO-IR [51] is Raspberry Pi's third generation camera. It has a Sony IMX708 CMOS sensor with 11.9 megapixel still resolution with a 66° field of view lens. It is capable of capturing 2304×1296 resolution video at 56 Fps. This version does not have a built in infrared filter, and is suitable for infrared image capturing. The sensor's light sensitivity can be seen on Figure 4.1.

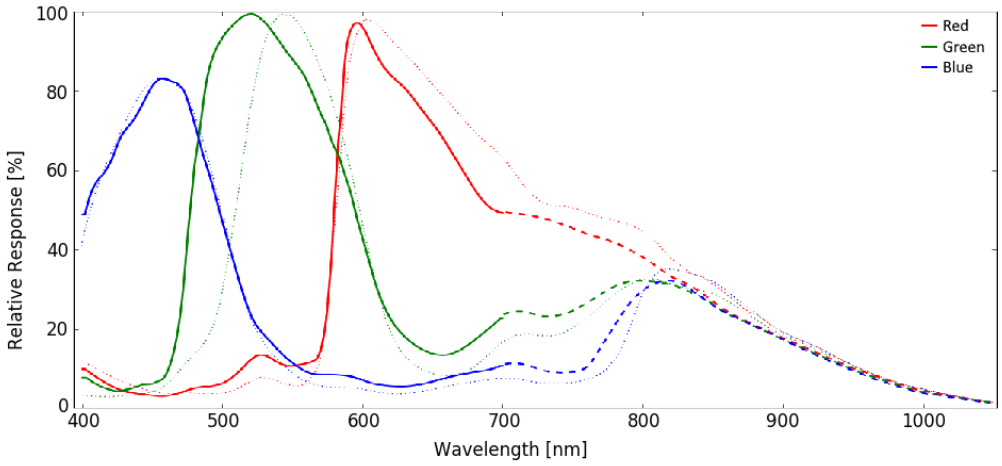


Figure 4.1: Spectral sensitivity curve of Sony IMX sensors [68]

Poly Methyl Methacrylate (PMMA) [29] is an infrared longpass filter, which means it is transparent to infrared, and longer wavelengths of light. At 650 nm, it absorbs all of the photons. At around 775 nm and at longer wavelengths, it lets through approximately 95% of the light. The transmittance of this filter can be seen on Figure 4.2.

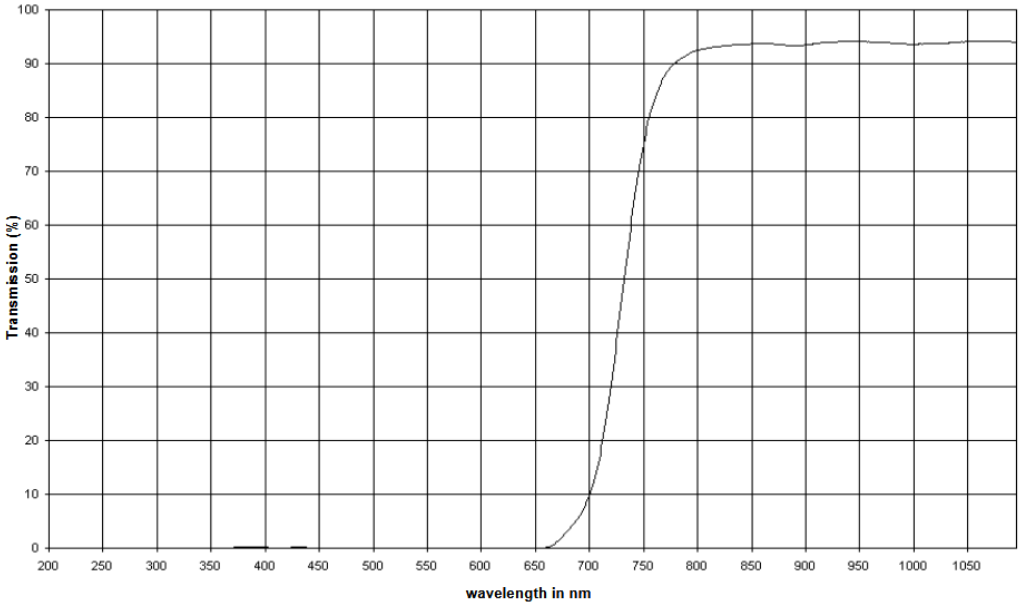


Figure 4.2: Transmittance of PMMA [29]

Oslon Black Infrared Starboard [60] is an infrared Light Emitting Diode (LED) made by OSRAM. Depending on the part specification, the LED's centroid wavelength can be at 735 nm, 850 nm or even 940 nm. The starboard that it is mounted on also functions as a heatsink. The LED can be powered by up to 1.0 A or the 850 nm version up to 1.5 A.

The design and prototyping of custom sensor housings required 3D modeling and slicing for fabrication. Blender [12] was used to create digital models of the sensor cases, allowing visualization and adjustments before production. These models were then converted into instructions for 3D printers using OrcaSlicer [59], which allowed for the full customization of the housings.

4.2 Conceptual prototype

The first prototype used four 940 nm LEDs[1]. Power was supplied to the LEDs by an output port of the Raspberry Pi GPIO interface, which has a maximum output of 3.3 V. Four 330Ω resistors were soldered in parallel to achieve a total resistance of 82.5Ω, supplying a current of approximately 26 mA. The prototype and its circuit diagram can be seen in Figure 4.3.

A single LED has a viewing angle of 30°. This angle later seemed too narrow to get a good amount of lighting emitted on the hand. When taking images with the combination of this LED, a light artifact was visible as a result. To mitigate these artifacts, the surface of these LED bulbs were sanded down, acting as a

[1]The LEDs were purchased on a Hungarian electronic online store. URL: https://www.hestore.hu/prod_10000626.html?lang=en

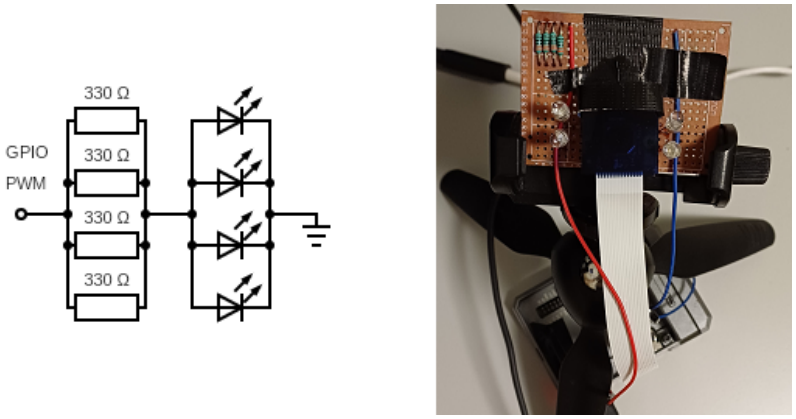


Figure 4.3: LED circuit diagram (left) and the implemented prototype (right)

diffuser. The camera and the LEDs were mounted on a solder board, and connected to the Raspberry Pi via jumper cables, and the camera cable. We can see the resulting images taken with this prototype on Figure 4.4.

Figure 4.4: Image of the palm, with (left) and without background lighting (right)

4.3 Lighting enhancements and sensor case

The second prototype incorporated a high power 850 nm LED mounted on a starboard. A variable resistor was utilized to fine tune the resistance for LED protection. The LED has a 90°viewing angle, and testing of the prototype showed that a diffuser was not needed due to this wide angle.

A 3D-printed case was made to more effectively integrate the camera, lighting, and PMAA filter. The optical filter was put into the detachable top. The lighting circuit, camera unit, and the Raspberry was still a separate unit. This was built on a breadboard for easier prototyping. The prototype and its circuit diagram can be seen on Figure 4.5.

This lighting circuit has a transistor, meaning other power sources could be connected to the LEDs. It is powered by the +5 V port of the GPIO interface, and the 2N2222 NPN transistor[2] is powered by the GPIO port. The Raspberry Pi and the lighting circuit are connected together with jumper cables. The LEDs are connected directly to the breadboard.

[2]https://www.mouser.at/ProductDetail/Diotec-Semiconductor/2N2222A?qs=OlC7AqGiEDmNlRSTGL2dYg%3D%3D

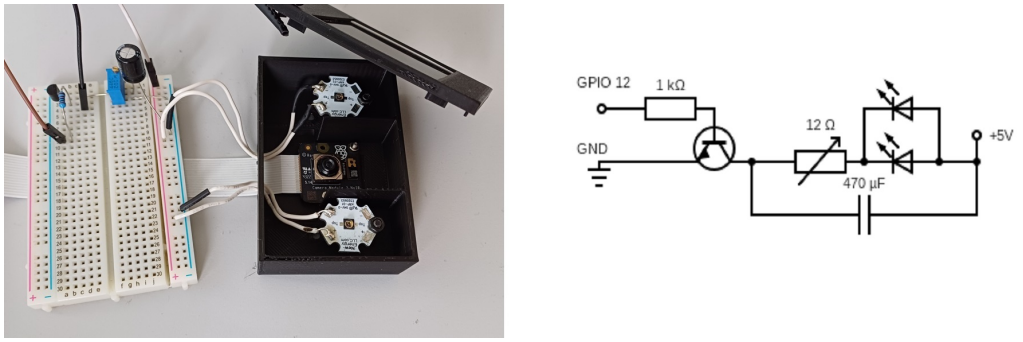


Figure 4.5: The second prototype (left) and the circuit diagram of its lighting (right)

The transistor supports up to 600 mA. The resistor is fine tuned to 12 Ω, which lets through a maximum of 160 mA, preventing the overheating and destruction of the transistor. A capacitor helps maintaining constant voltage through the LEDs, mitigating the flickering of the lights.

Figure 4.6: Image taken by the second prototype

There were two versions of this prototype: The first version had a problem with the LEDs reflecting from the filter itself, since it was in the same space. The second version includes walls between each LED and the camera. This prevents the reflection showing up, when taking a picture while hiding the inner components are hidden. On Figure 4.6, we can see an example image taken by the second prototype.

4.4 Improved case and compact LED circuit

The third prototype has a minimized light control circuit. A new case was also designed for it, which incorporates everything into a case. There are three smaller filters: two for each LED and one for the camera. The case top was mounted to the case with screws. This prototype can be seen on Figure 4.7.

The third prototype has the same lighting control circuit as the second one in Section 4.3. The holes at the bottom and the top of the case ensures that the Raspberry Pi has enough airflow, so it would not overheat over time. There is

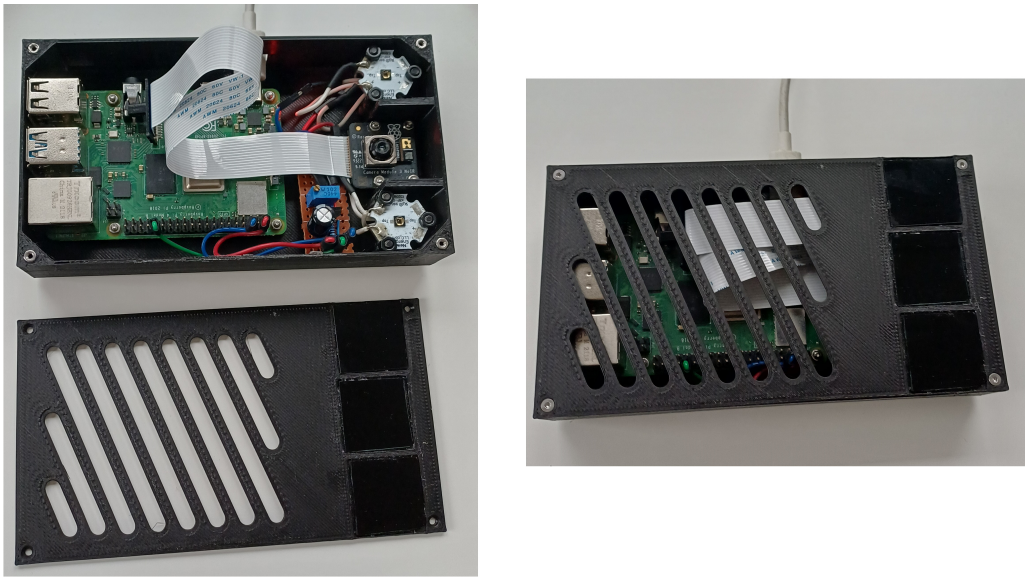


Figure 4.7: The third prototype without (left) and with the top (right)

only one hole on the side of the case, where the USB cable can be plugged into, in order to power the device. Everything is mounted on the case with screws and screw nuts.

Chapter 5

Development of the Authentication Model

This Chapter discusses the development and evaluation of models for vein-based authentication. It begins with a discussion of the datasets used for training and evaluation (Section 5.1), which serve as the basis to compare model performance under different preprocessing levels.

Section 5.2 introduces traditional models that act as baselines for comparison, highlighting their accuracy and performance across different settings. Then Section 5.3 discusses the machine learning-based approaches and the design and configuration of the neural network architectures.

During the training process, certain challenges were encountered that affected the performance of models. These issues were analyzed in Section 5.3.3 to better understand the limitations of the chosen approach. A solution to the identified problem is then presented in Section 5.3.4 in the form of an architectural change.

5.1 Dataset preprocessing for evaluation and training

The Dorsalhandveins dataset, previously mentioned in Section 3.1.4, was used for this project. A parallel can be drawn between dorsal and palm vein recognition techniques because of the similarities between their vein patterns and image acquisition methods. Access to other datasets mentioned in Section 3.1 was requested, however those requests were rejected due to legal requirements.

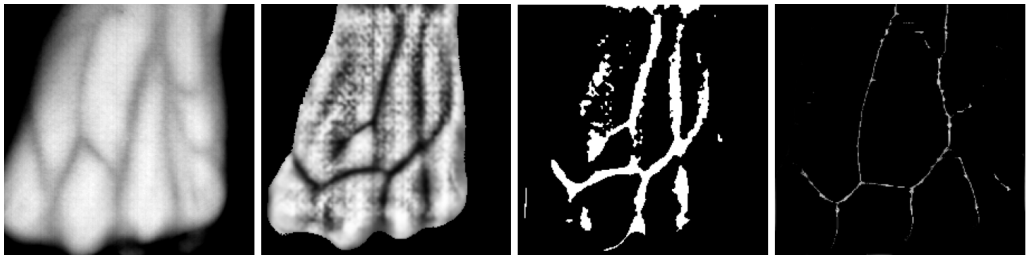


Figure 5.1: The same hand from each created dataset

For image processing, the OpenCV [37] library was used, since it includes many image processing algorithms. With this tool, four different datasets were created: The first dataset is the unprocessed. The second dataset is created by applying histogram equalization, then CLAHE with six different tile grid sizes. These sizes were 3, 7, 9, 11, 13 and 17. The clip limit was set to 2. The third dataset was created by applying thresholding on the second dataset. The threshold was

set to a dynamic value calculated by averaging the minimum and maximum intensity pixel. The fourth dataset was a precomputed one, that utilized valley and curvature detection methods. [8] I used the precomputed dataset in the project repository. These four types of preprocessing can be seen on Figure 5.1, where each method was applied on the same hand.

Each version of the datasets can point to the quality of the preprocessing level. They show how much each step can improve the effectiveness of the given algorithm, while the unprocessed set gives a baseline.

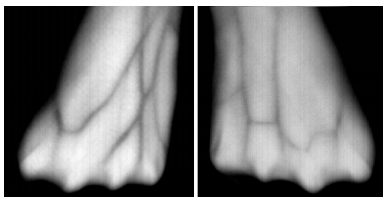


Figure 5.2: Left and right (mirrored) hand image of the same person

In the end, each dataset had 276 classes. Due to the distinct vein structures in the left and right hands, even when mirrored (Figure 5.2), the dataset of 138 participants resulted in 276 classes, with each class having four instances. The resulting datasets contained 1104 images each.

Different approaches can be found in literature which include vein extraction and segmentation. [64] The CLAHE contrast enhancing algorithm is a computationally effective alternative to the Hessian filter that is used for segmentation, according to researchers. [71] Hessian filters also require careful parameter tuning.

The CLAHE algorithm also performs well in a variety of lighting conditions. It is also easier to implement, since it does not require parameter tuning like the Hessian filter does. The processing time of CLAHE algorithm is also lower than the Hessian. [71]

5.2 Traditional models

ORB and SIFT were used for descriptor extraction. While ORB is a computationally more efficient algorithm, in some cases SIFT can create more precise descriptors. [57]

Algorithm	**720×1280 resolution**	**360×640 resolution**
ORB	140 ms	40 ms
BF Matcher (ORB)	4 ms	4 ms
FLANN Matcher (ORB)	3 ms	3 ms
SIFT	981 ms	201 ms
BF Matcher (SIFT)	22595 ms	522 ms
FLANN Matcher (SIFT)	1426 ms	184 ms

Table 5.1: Processing time with different image resolutions

Brute-Force was used for the ORB descriptors, as for with SIFT, FLANN matcher was used for quality metrics. On the Raspberry Pi, there was not a significant

difference in ORB descriptor matching performance at higher resolutions, but the runtime for SIFT was noticeably longer, as shown in Table 5.1.

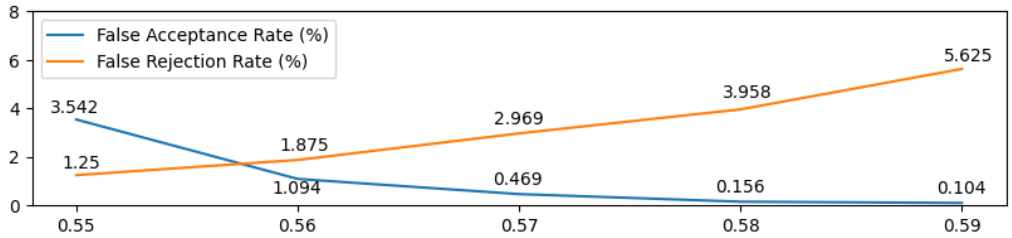


Figure 5.3: FRR and FAR for the ORB algorithm

The ORB model performed best on the dataset, which was preprocessed using CLAHE. The resulting false acceptance and rejection rates are shown on Figure 5.3. These threshold values also depend on the input data it is applied on. The quality and resolution of the sensor might pick up details that other type of sensors cannot.

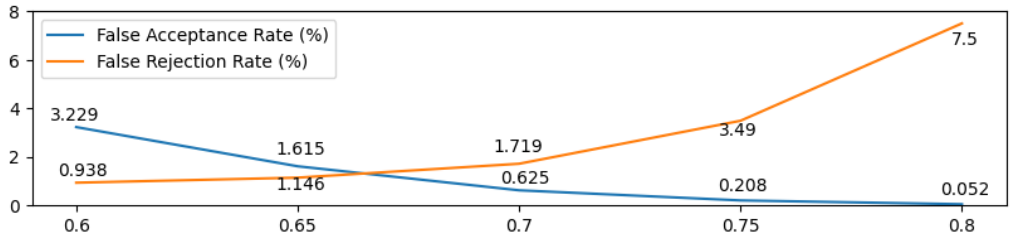


Figure 5.4: FRR and FAR for the SIFT algorithm

The SIFT model produced similar results, which performed best on the unprocessed set. The curves of the false rejection rate and acceptance rate is shown on Figure 5.4.

Both the ORB and SIFT model was tested on 1920 image pairs. The distribution of positive and negative matches was also 50%, to get a balanced set. The full set would have contained 1218816 pairs, with around 1.45% matching and 98.55% non-matching pairs, making it heavily unbalanced.

5.3 Machine learning models

The plan was to create an embedding model [24] by first simply training a Convolutional Neural Network (CNN) [20] with the classes found in the dataset. The images had a resolution of 224×224 and had only one channel. Each dataset was divided into a training set (600 images), a validation set (300 images), and a test set (204 images). Each dataset described in Section 5.1 was utilized to determine the effects of each preprocessing step.

In case the CNN model were to be successfully trained, its dense layers could be removed. The output from the flatten layer would then be used as the image embedding, which could in turn be used to train a Dense Neural Network [18] for image matching. This model would have a single output that represents the similarity between the inputs.

5.3.1 Software Tools and Frameworks

Python [13] was selected as the primary programming language for the development of the authentication models due to its flexibility, extensive library support, and wide adoption in the field of machine learning. In particular, TensorFlow [69] and Keras [54] libraries were used to implement, train and test various neural networks. These tools helped with rapid prototyping of new models, allowing experimentation with different network configurations and hyperparameters.

The performance of each trained model was evaluated on the loss metrics, where the best model was chosen. To compare the effectiveness of this neural network model, it had to be further evaluated using the metrics mentioned in Section 2.4. This gave a comprehensive difference between the created machine learning model and the traditional models.

5.3.2 Convolutional Neural Network

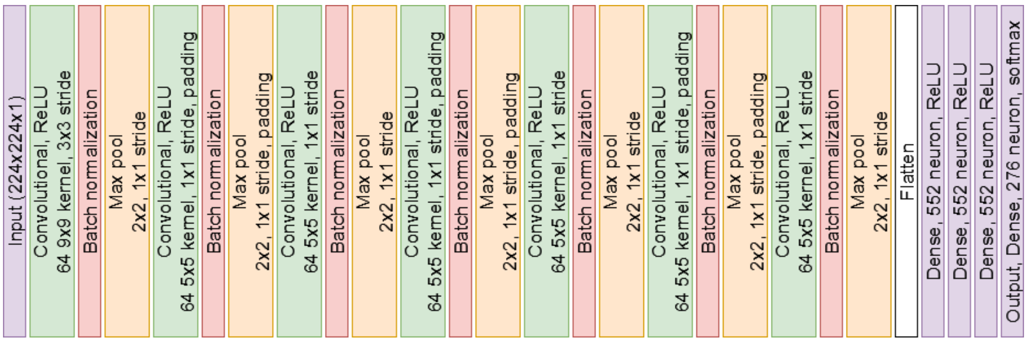


Figure 5.5: The final configuration Convolutional model

For a Convolutional Neural Network, multiple configuration were tried. The final architecture of the Convolutional model can be seen on Figure 5.5. Unfortunately, every configuration produced worse results during each step in the training on the validation set. ADAM optimizer [25] was used during the training, alongside with Stochastic Gradient Descent, [22] which did not change the results. The validation loss kept increasing with each step, and the validation accuracy stayed at near 0. For trials, both Cross-Entropy [26] and Mean Squared Error [21] were utilized as loss functions. At first I thought that the architecture of this model caused the problems. For that another approach was tried: transfer learning [27].

Transfer learning with ResNet-50 Model

ResNet-50 is a deep Convolutional Neural Network. ResNet stands for Residual Network, and the 50 refers to the number of layers it has. [43] It uses residual connections, which allow the network to learn residual functions that map the input to the desired output. Residual connections help deeper layers to learn much faster, without suffering from the problem of vanishing gradient. This network was used for training, where transfer learning was applied. The particular model was pretrained on the ImageNet dataset [44]. Unfortunately, this approach yielded the same outcome as the CNN model without transfer learning.

5.3.3 Possible problem

The models mentioned in the previous sections failed to produce meaningful results, as indicated by a validation loss that increased from the outset and a validation accuracy that remained near zero. A potential problem is that the class boundaries learned by the model were too indistinct. This means that a slightly varied input from the same class could be classified as two different classes.

To mitigate the challenges posed by the limited number of instances per class, two solutions were considered: data augmentation [77] and the use of an alternative dataset. While data augmentation was attempted to increase the number of instances, it proved insufficient in resolving the training problem. Additionally, using a different dataset with more instances per class was not a viable solution due to the reasons mentioned in Section 5.1.

The problem of having too few instances per class can be solved by reducing the number of classes for the model to learn from. This method however, does not improve real-world performance, as it creates a model that over-generalizes to the included classes. The benefit is that it makes hyperparameter tuning a much easier task. A total of 10 classes was chosen after trials with different numbers of classes. Even with this number of classes, the model still did not succeed to learn the classification task.

Another solution to this would be to use a different self-supervised learning method called similarity learning or contrastive learning [1, 10]. This method changes the model's output from a class label to an embedding space, where the training objective is to minimize the distance between embeddings of the same class while maximizing the distance between different classes. This process creates a well-separated embedding space and can be achieved by using a different learning approach, or by creating a model with a different architecture.

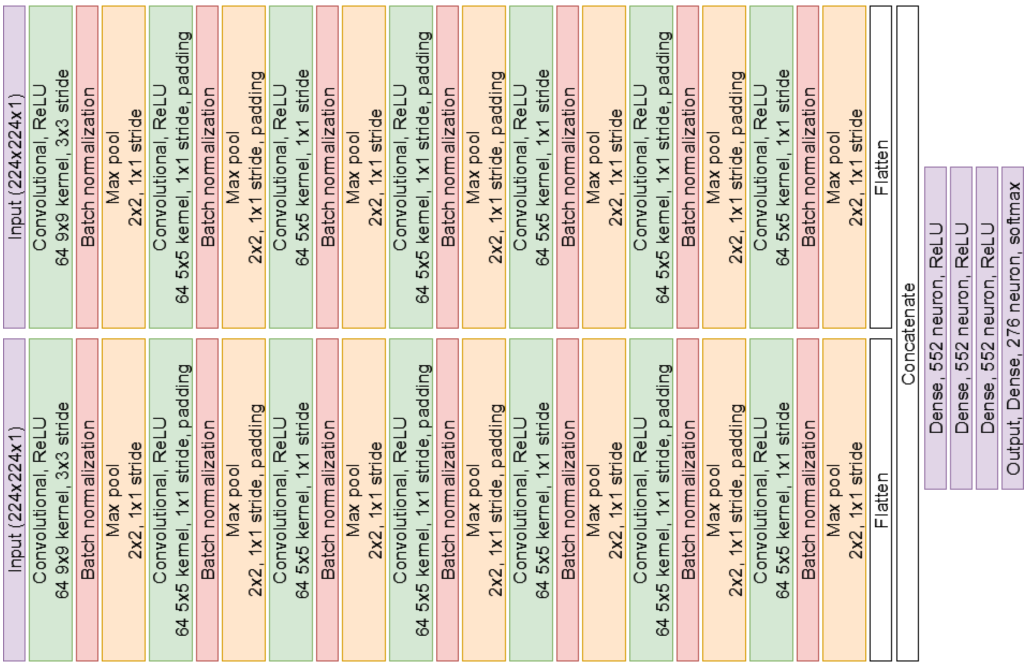


Figure 5.6: The architecture of the Half-Twin model

5.3.4 Half-Twin Neural Network Model

The Half-Twin model [3, 23] produced measurable results on the datasets. The network configuration is very similar to the Convolutional Neural Network mentioned in Section 5.3.2. The configuration of it can be seen on Figure 5.6.

The network was trained on 2000 instances of image pairs, and the validation set consisted of 320 pairs. The test set was the same used in Section 5.2 with 1920 instances of image pairs.

The end result of the twin model training is presented on Figure 5.7: as we can see, the model performed best on the unprocessed raw dataset. During training, each dataset has been normalized.

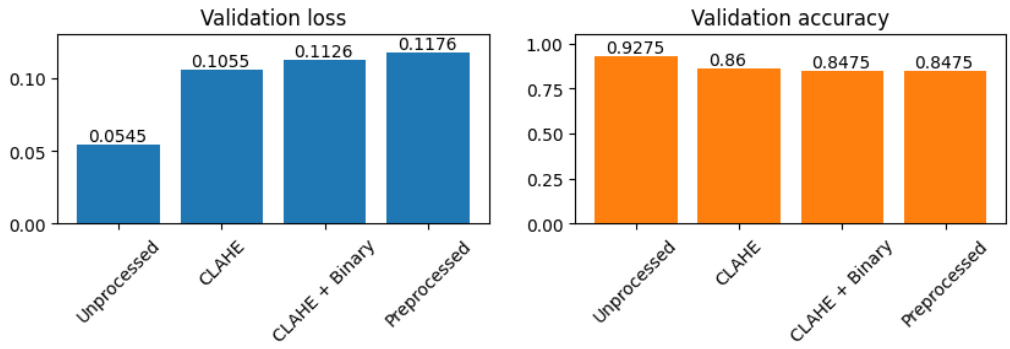


Figure 5.7: Loss and accuracy during training of the Half-Twin model

As we can see on Figure 5.8, the model performed best on the unprocessed dataset, where the threshold for recognizing the two patterns as „same" was set to 0.6. The time performance of its Convolutional layers on a 224×224 image is 320 ms, and the fully connected layers' is 493 ms.

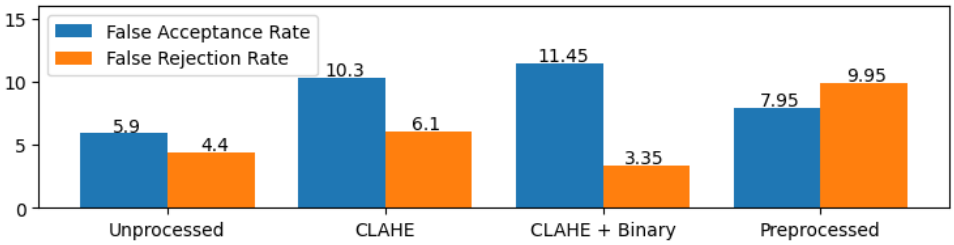


Figure 5.8: Authentication metrics of the Half-Twin model on each dataset

Chapter 6

Authentication pipeline

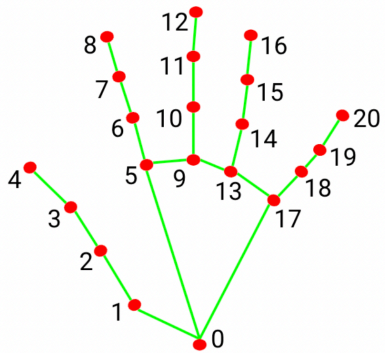
This Chapter discusses the authentication pipeline made to provide a realistic use case scenario on an embedded device using the traditional methods. Section 6.1 mentions the possible triggering mechanism for activating the preprocessing and feature extraction steps of the pipeline, while preserving resources during the inactive phase. Section 6.2 goes through the parts of the authentication algorithm, where each step is discussed in detail.

An experimental authentication pipeline was made to prove that hand vein recognition can be used on an embedded device. The second prototype version, described in Section 4.3, was used first and later a better case was designed, which can be seen in Section 4.4. The second prototype had a problem with lighting, because the filter itself was to narrow and half of the viewing angle's of the LEDs were blocked. This issue was fixed with the third prototype. The filter was separated into 4 equivalent segments, where 3 of them were used to cover the camera and the two LEDs.

6.1 Authentication triggers

Authentication triggers are important parts of the pipeline. We want to detect when to apply recognition, since we do not want to waste system resources on inputs not containing the target, in this case the palm of the user.

6.1.1 Region of Interest



0. WRIST
1. THUMB_CMC
2. THUMB_MCP
3. THUMB_IP
4. THUMB_TIP
5. INDEX_FINGER_MCP
6. INDEX_FINGER_PIP
7. INDEX_FINGER_DIP
8. INDEX_FINGER_TIP
9. MIDDLE_FINGER_MCP
10. MIDDLE_FINGER_PIP
11. MIDDLE_FINGER_DIP
12. MIDDLE_FINGER_TIP
13. RING_FINGER_MCP
14. RING_FINGER_PIP
15. RING_FINGER_DIP
16. RING_FINGER_TIP
17. PINKY_MCP
18. PINKY_PIP
19. PINKY_DIP
20. PINKY_TIP

Figure 6.1: Hand landmarks [46]

The Region of Interest (ROI) [2] is a specific area in the image or data that is identified for a particular purpose. This method is used to increase efficiency and reduce computational time by narrowing the scope of the analysis to the

palm of the hand. For this, Google's MediaPipe [47], an AI solution library was used. It is open-source and contains a suite of libraries and tools to quickly apply artificial intelligence and machine learning. These solutions include hand landmark detection, which was used to create the bounding box of the palm.

The algorithm can extract either the palm, or dorsal region of the picture when a hand is detected. It uses the resulting index finger metacarpophalangeal (MCP) and pinky MCP landmarks defined by the MediaPipe library – which we can see on Figure 6.1 – to create a square region on the hand, after which the detected region is rotated and cut. The resulting image is then resized to a 224×224 image.

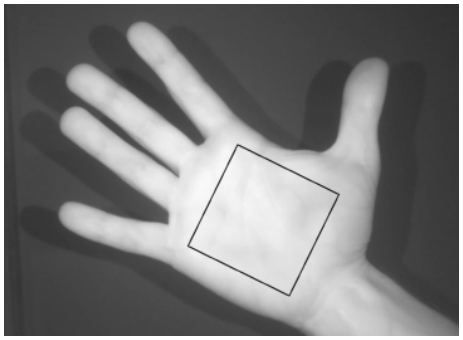


Figure 6.2: Region of Interest borders on the detected hand

It does not matter which side of the hand is shown on the image, the algorithm detects it, and calculates the right region boundaries (Figure 6.2). It also rotates the image, meaning the hand is always pointing in the upright position.

6.1.2 Closeness

The created detection algorithm is also capable of monitoring the distance of the closest object to the sensor. The used camera has an auto-focus feature. After it is performed, the algorithm queries the lens position of the camera, where the inverse of the lens position is the focal distance in meters. [50] If an object is close to the device – currently configured to 20 cm or less – the algorithm can trigger the authentication sequence.

6.2 Algorithm

As a proof of concept to demonstrate the models' application in a real-life scenario, a simple algorithm was created. This algorithm includes a timer for the LED that is initially set to zero and turned off to conserve energy. At each iteration, the timer is decremented if its value is greater than zero. Upon the detection of a hand, the LED is turned on and the timer is reset to the value configured in the settings.

We can set a minimum and a maximum value for the LED control, which represent the amount of power delivered to the LED. In a building where LED lighting is used, the lighting might not produce any infrared light, which would make the image completely dark. Setting the minimum value higher than zero ensures that visible light is always present during image capture. This is also useful for outdoor deployment or near a window, where during the night no natural infrared light source is present.

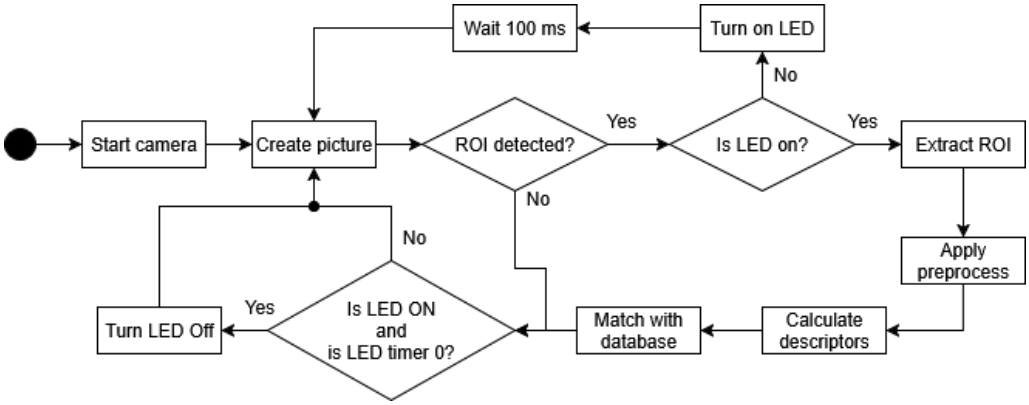


Figure 6.3: The authentication pipeline

After extracting the Region of Interest, the CLAHE preprocessing method mentioned in Section 2.5.2 is applied on the image. This preprocessing method is not sensitive to lighting conditions, which makes the resulting images easy to use for recognition.

The current pipeline uses the ORB descriptor with Brute-Force matching. Although the machine learning solutions need extra training data to get better performing result for a real-time scenario. When calculating the descriptors either ORB, SIFT or an embedding model could be used.

The resulting descriptors or feature vector then can be matched or compared with the database instances. The database itself can only store the precalculated values, instead of storing the images itself, to conserve space, save computational resource and to be more secure. The algorithm itself was designed to be modular, meaning different database solutions can be implemented.

The process of matching extracted data can be accomplished through several methods, each with its own advantages. Traditional methods and similarity metrics can be employed for this purpose. These techniques, such as Euclidean distance or correlation, directly compare the features of two templates to assess their likeness.

Alternatively, a more advanced approach could involve training a neural network to perform the matching. This method would be effective, because the network can be trained to learn a high-level embedding of the data. The network could learn to compare these embeddings, allowing it to determine the similarity between two inputs. This could be more effective than traditional metrics for handling complex, non-linear relationships within the data.

Chapter 7

Conclusion and Future Work

This concluding chapter reflects on the key findings and contributions of this thesis, and it outlines possible directions for future work.

Conclusion

For hand vein recognition task the need of specialized sensors makes it hard to capture by everyday cameras, making it hard to forge, and it's contactless form makes it more hygienic and harder to trace.

Using an embedded system is a viable solution for an authentication device, as it results in a compact form factor. This compact design makes the sensor easy to deploy and set up in various environments.

Traditional methods for vein recognition are a viable alternatives to machine learning models. Higher quality and bigger quantity of data needed for the machine learning models to perform better. This not only means the number of instances, but also the amount of classes that the dataset incorporates. Better quality of the instances – bigger image resolution, less noise on the image – also help with the training process

The developed authentication pipeline is a viable method for re-identification. The LED controls work perfectly under deployment, even without ambient infrared light. Both the ROI detection and closeness detection work, and produce usable images for the algorithm to process. The algorithm itself recognizes the hand, and if the database contains the specific target, a method is triggered.

The combination of the lighting and image sensor captures well the veins under the skin of the user. The infrared pass-through filter improved this feature acquisition, and with the help of the preprocessing algorithm – mainly the contrast limited adaptive histogram equalization – the target is successfully captured in different lighting conditions. This also means that the distance between the captured hand and the sensor itself can vary, a strict rule for distance is not needed.

While the developed solutions are not better than the commercially available sensors, with enough good quality data they can become close to them. It is also difficult to gave a comparison between the developed solutions, because they use different datasets for evaluation.

The experiments and results show that the goals of this thesis were successfully met. The thesis explored the field of vein biometrics and successfully built a working self-contained authentication system using an embedded device. This developed system works well, demonstrating that a small, low-cost device is a viable option for real-time vein recognition, as it successfully handles image capture, processing, and authentication. While its performance can be improved with higher quality data, this project provides a solid foundation for future work.

Future Work

Migrating the Python code to Rust [70] could improve the performance and increase security. Since Rust is a compiled language, it is much faster at executing the processing needed for recognition, feature extraction, preprocessing and matching. It is also a memory safe language, which further improves the security, after it has been deployed.

A solution other than a different architecture of the CNN model is similarity learning (Section 5.3.3). This helps to train the model to not recognize existing classes, but to differentiate non-matching vein patterns by increasing the distance in embedding level.

Trying out different cameras with the selected authentication algorithm: there are multiple cameras that have higher spectral sensitivity in the infrared range than the camera used in the thesis. Higher resolution cameras might capture fine details of the hand. A higher field-of-view can reduce energy consumption, since the hand can be closer to the camera, and this requires less intense lighting.

The Raspberry Pi Foundation produces smaller single board computers than the one used during development. This can further decrease the physical footprint of the device. A small screen displaying what the camera can see could also be included in this device, to give feedback to the user.

Bibliography

[1] Panagiotis Antoniadis; Michal Aibin. 2022. An Introduction to Contrastive Learning. (2022). https://www.baeldung.com/cs/contrastive-learning.

[2] AmazingAlgorithms.com. 2025. Region of Interest. (2025). https://amazingalgorithms.com/definitions/region-of-interest/.

[3] Poulami Bakshi. 2025. Siamese Networks. (2025). https://poulami98bakshi.medium.com/siamese-networks-d28ac0b0836d.

[4] Ravimal Bandara. 2011. *A Music Keyboard with Gesture Controlled Effects Based on Computer Vision*. PhD thesis. (January 2011).

[5] Sue Black. 2018. *All that remains: a life in death.* Doubleday, (April 2018). ISBN: 978-0-857-52492-8. DOI: 10.1007/978-0-857-52492-8.

[6] Franky Chandra, Aries Wahyudianto, and M Yasin. 2017. Design of vein finder with multi tuning wavelength using RGB LED. In *Journal of Physics: Conference Series* number 1. Volume 853. IOP Publishing, p. 012019.

[7] Tyler Choi. 2022. Iris recognition reaches the mainstream for identification, authentication. (June 2022). https://www.biometricupdate.com/202206/iris-recognition-reaches-the-mainstream-for-identification-authentication.

[8] Hai Chu. 2024. Dorsal hand vein recognition. (August 2024). https://github.com/hai-dchu/palm-vein-recognition/tree/main.

[9] Septimiu Crisan, Ioan Gavril Tarnovan, and Titus Eduard Crisan. 2007. Vein pattern recognition. Image enhancement and feature extraction algorithms. In *15th IMEKO TC4 Symposium, Iaşi, Romania*.

[10] Jiankang Deng, Jia Guo, Jing Yang, Niannan Xue, Irene Kotsia, and Stefanos Zafeiriou. 2022. ArcFace: Additive Angular Margin Loss for Deep Face Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44, 10, (October 2022), 5962–5979. ISSN: 1939-3539. DOI: 10.1109/tpami.2021.3087709. http://dx.doi.org/10.1109/TPAMI.2021.3087709.

[11] Bernardo Núñez-Álvarez Felipe Wilches-Bernal and Pedro Vizcaya. 2020. A Database of Dorsal Hand Vein Images. (2020). https://arxiv.org/pdf/2012.05383.

[12] Blender Foundation. 1994. Blender. (1994). https://www.blender.org/.

[13] Python Software Foundation. 2001. Python. (2001). https://www.python.org/.

[14] Rihards Fuksis, Modris Greitans, Olegs Nikisins, and M Pudzs. 2010. Infrared Imaging System for Analysis of Blood Vessel Structure. *Electron Electrical Eng*, 1, (January 2010).

[15] Inc. Fulcrum Biometrics. 2025. Palm vein identification payment solution powered by Fulcrum Biometrics. (2025). https://www.fulcrumbiometrics.com/palm-payments/.

[16] Sanchhaya Education Private Limited GeeksforGeeks. 2025. CLAHE Histogram Equalization - OpenCV. (2025). https://www.geeksforgeeks.org/python/clahe-histogram-eqalization-opencv/.

[17] Sanchhaya Education Private Limited GeeksforGeeks. 2025. Data Preprocessing in Data Mining. (2025). https://www.geeksforgeeks.org/dbms/data-preprocessing-in-data-mining/.

[18] Sanchhaya Education Private Limited GeeksforGeeks. 2025. DenseNet Explained. (2025). https://www.geeksforgeeks.org/computer-vision/densenet-explained/.

[19] Sanchhaya Education Private Limited GeeksforGeeks. 2025. Histogram Equalization in Digital Image Processing. (2025). https://www.geeksforgeeks.org/computer-graphics/histogram-equalization-in-digital-image-processing/.

[20] Sanchhaya Education Private Limited GeeksforGeeks. 2025. Introduction to Convolution Neural Network. (2025). https://www.geeksforgeeks.org/machine-learning/introduction-convolution-neural-network/.

[21] Sanchhaya Education Private Limited GeeksforGeeks. 2025. Mean Squared Error. (2025). https://www.geeksforgeeks.org/maths/mean-squared-error/.

[22] Sanchhaya Education Private Limited GeeksforGeeks. 2025. ML - Stochastic Gradient Descent (SGD). (2025). https://www.geeksforgeeks.org/machine-learning/ml-stochastic-gradient-descent-sgd/.

[23] Sanchhaya Education Private Limited GeeksforGeeks. 2025. Siamese Neural Network in Deep Learning. (2025). https://www.geeksforgeeks.org/nlp/siamese-neural-network-in-deep-learning/.

[24] Sanchhaya Education Private Limited GeeksforGeeks. 2025. What are embedding models. (2025). https://www.geeksforgeeks.org/nlp/what-are-embedding-models/.

[25] Sanchhaya Education Private Limited GeeksforGeeks. 2025. What is Adam Optimizer? (2025). https://www.geeksforgeeks.org/deep-learning/adam-optimizer/.

[26] Sanchhaya Education Private Limited GeeksforGeeks. 2025. What Is Cross-Entropy Loss Function? (2025). https://www.geeksforgeeks.org/machine-learning/what-is-cross-entropy-loss-function/.

[27] Sanchhaya Education Private Limited GeeksforGeeks. 2025. What is Transfer Learning? (2025). https://www.geeksforgeeks.org/machine-learning/ml-introduction-to-transfer-learning/.

[28] Edd Gent. 2014. Europe's first shared biometric ATM network to be created. (May 2014). https://eandt.theiet.org/2014/05/14/europes-first-shared-biometric-atm-network-be-created.

[29] TTV GmbH. 2009. LUXACRYL®-IR. (2009). https://www.go-ttv.com/materials/acrylic/luxacryl-ir/.

[30] Xiaofei He, Shuicheng Yan, Yuxiao Hu, Partha Niyogi, and Hong-Jiang Zhang. 2005. Face recognition using laplacianfaces. *IEEE transactions on pattern analysis and machine intelligence*, 27, 3, 328–340.

[31] Ltd Hitachi. 2025. Finger Vein Authentication Technology. (2025). https://www.hitachi.com/products/it/veinid/introduction/fingervein.html.

[32] Ltd Hitachi. 2021. Hitachi H1 finger vein sensor. (2021). https://www.m2sys.com/wp-content/uploads/custom-images/hitachi-finger-vein-h1.png.

[33] Ltd Hitachi. 2021. Pay by Finger. (2021). https://digitalsecurity.hitachi.eu/products/pay-by-finger/.

[34] Idiap Research Institute. 2024. Bob. (2024). https://www.idiap.ch/software/bob/.

[35] Idiap Research Institute. 2012. Note on CASIA Multi-Spectral Palmprint Database. (2012). http://biometrics.idealtest.org/dbDetailForUser.do?id=5#/datasetDetail/6.

[36] Idiap Research Institute. 2024. VERA PalmVein. (2024). https://www.idiap.ch/en/scientific-research/data/vera-palmvein.

[37] Itseez Intel Willow Garage. 1999. OpenCV. (1999). https://opencv.org/.

[38] Kerry Van Iseghem. 2012. LightWise 1.3-S Product Brief. (2012). https://www.spinel.com.tw/page/pd/isg/isg/LightWise_1.3-S_Product_Brief.pdf.

[39] A.K. Jain, A. Ross, and S. Prabhakar. 2004. An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14, 1, 4–20. DOI: 10.1109/TCSVT.2003.818349.

[40] Fujitsu Ltd; Leslie Kett. 2021. PalmSecure-F Pro. (2021). https://www.fujitsu.com/global/imagesgig5/F-Pro-data-sheet-8-2.pdf.

[41] Romain Kirszenblat and Paul Edouard. 2021. Validation of Withings ScanWatch as a Wrist-Worn Reflective Pulse Oximeter: Hypoxia Study (Preprint). *Journal of Medical Internet Research*, 23, (February 2021). DOI: 10.2196/27503.

[42] Dakshina Ranjan Kisku, Ajita Rattani, Phalguni Gupta, Jamuna Sing, and C Hwang. 2012. Human Identity Verification Using Multispectral Palmprint Fusion. *Journal of Signal and Information Processing*, 03, (January 2012). DOI: 10.4236/jsip.2012.32036.

[43] Nitish Kundu. 2023. Exploring ResNet50: An In-Depth Look at the Model Architecture and Code Implementation. (2023). https://medium.com/@nitishkundu1993/exploring-resnet50-an-in-depth-look-at-the-model-architecture-and-code-implementation-d8d8fa67e46f.

[44] Stanford Vision Lab. 2020. ImageNet image database. (2020). https://image-net.org/index.php.

[45] Bit Flip LLC. 2023. FAR vs. FRR. (2023). https://thisvsthat.io/far-vs-frr.

[46] GOOGLE LLC. 2020. MediaPipe Hands. (2020). https://chuoling.github.io/mediapipe/solutions/hands.html.

[47] GOOGLE LLC. 2025. MediaPipe Solutions guide. (2025). https://ai.google.dev/edge/mediapipe/solutions/guide.

[48] Antonia Longo, Stefan Morscher, Jaber Malekzadeh Najafababdi, Dominik Jüstel, Christian Zakian, and Vasilis Ntziachristos. 2020. Assessment of hessian-based Frangi vesselness filter in optoacoustic imaging. *Photoacoustics*, 20, 100200. ISSN: 2213-5979. DOI: https://doi.org/10.1016/j.pacs.2020.100200. https://www.sciencedirect.com/science/article/pii/S2213597920300409.

[49] Fujitsu Ltd. 2025. PalmSecure™ - Your business easily secured. (2025). https://www.fujitsu.com/global/services/security/offerings/biometrics/palmsecure/.

[50] Raspberry Pi Ltd. 2025. The Picamera2 Library. (2025). https://datasheets.raspberrypi.com/camera/picamera2-manual.pdf.

[51] Raspberry Pi (Trading) Ltd. 2023. About the Camera Modules. (2023). https://www.raspberrypi.com/documentation/accessories/camera.html.

[52] Raspberry Pi (Trading) Ltd. 2024. Raspberry Pi 4 Model B Datasheet. (2024). https://datasheets.raspberrypi.com/rpi4/raspberry-pi-4-datasheet.pdf.

[53] Davide Maltoni, Dario Maio, Anil K Jain, and Salil Prabhakar. 2009. *Handbook of fingerprint recognition*. Springer.

[54] ONEIROS. 2015. Keras. (2015). https://keras.io/.

[55] Yiltan Bitirim Önsen Toygar Felix O. Babalola. 2020. FYO: A Novel Multimodal Vein Database with Palmar, Dorsal and Wrist Biometrics. 8, 1, (December 2020), 82461–82470. DOI: 10.1109/ACCESS.2020.2991475. https://fyo.emu.edu.tr/en.

[56] OpenCV. 2025. Feature Matching. (2025). https://docs.opencv.org/4.x/dc/dc3/tutorial_py_matcher.html.

[57] OpenCV. 2025. Introduction to SIFT (Scale-Invariant Feature Transform). (2025). https://docs.opencv.org/3.4/da/df5/tutorial_py_sift_intro.html.

[58] OpenCV. 2025. ORB (Oriented FAST and Rotated BRIEF). (2025). https://docs.opencv.org/3.4/d1/d89/tutorial_py_orb.html.

[59] OrcaSlicer. 2023. OrcaSlicer - Powerful 3D printing slicer. (2023). https://www.orcaslicer.com/.

[60] OSRAM. 2024. OSRAM Oslon Black Infrared (IR) Starboards - Data Sheet. (2024). https://new-energyllc.com/wp-content/uploads/NewEnergy_StarBoard-IR-Osram_DataSheet.pdf.

[61] Tharinda Dilshan Piyadasa. 2022. Image Binarization in a Nutshell. (2022). https://medium.com/@tharindad7/image-binarization-in-a-nutshell-b40b63c0228e.

[62] Kenneth S. Saladin. 2011. *Human Anatomy*. McGraw-Hill Medical Publishing, (January 2011). ISBN: 978-0-071-22207-5.

[63] Ishani Sarkar, Farkhod Alisherov, Tai-hoon Kim, and Debnath Bhattacharyya. 2010. Palm Vein Authentication System: A Review. *Int J Control Autom Syst*, 3, (January 2010).

[64] Samiya Shakil, Deepak Arora, and Taskeen Zaidi. 2023. An optimal method for identification of finger vein using supervised learning. *Measurement: Sensors*, 25, 100583. ISSN: 2665-9174. DOI: https://doi.org/10.1016/j.measen.2022.100583. https://www.sciencedirect.com/science/article/pii/S2665917422002173.

[65] C. Simon. 1935. A New Scientific Methode of Identification. *The New York State Journal of Medicine*, 35, 18, 901–906. https://cir.nii.ac.jp/crid/1573668926135707776.

[66] Balaji Sontakke, Vikas Humbe, and Pravin Yannawar. 2017. Dorsal Hand Vein Authentication System: A Review. *Journal of Scientific Research and Development*, 6, (May 2017), 511–514.

[67] Balaji Sontakke, Vikas Humbe, and Pravin Yannawar. 2019. Vein Pattern Locating Technology for Cannulation: A Review of the Low-Cost Vein Finder Prototypes Utilizing near Infrared (NIR) Light to Improve Peripheral Subcutaneous Vein Selection for Phlebotomy. *Journal of Scientific Research and Development*, (August 2019).

[68] Brandon Swift. 2016. Raspberry Pi Camera spectral response curves. (June 2016). https://github.com/bluegreen-labs/raspberry_pi_camera_responses/blob/master/est-nir-response/SonyIMX219withNIRestimate_v_similarCMOS.png.

[69] Google Brain Team. 2015. TensorFlow. (2015). https://www.tensorflow.org/.

[70] Rust Team. 2025. Rust Programming Language. (2025). https://www.rust-lang.org/.

[71] 2020. *An Efficient Finger Vein Image Enhancement and Pattern Extraction Using CLAHE and Repeated Line Tracking Algorithm*. (January 2020), pp. 690–700. ISBN: 978-3-030-30464-5. DOI: 10.1007/978-3-030-30465-2_76.

[72] Pedro Tome and Sébastien Marcel. 2015. On the Vulnerability of Palm Vein Recognition to Spoofing Attacks. In *The 8th IAPR International Conference on Biometrics (ICB)*. (May 2015), pp. 319–325. DOI: 10.1109/ICB.2015.7139056. http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload=true&arnumber=7139056.

[73] PAUL TOWER. 1955. The Fundus Oculi in Monozygotic Twins: Report of Six Pairs of Identical Twins. *A.M.A. Archives of Ophthalmology*, 54, 2, (August 1955), 225–239. ISSN: 0096-6339. DOI: 10.1001/archopht.1955.00930020231010. eprint: https://jamanetwork.com/journals/jamaophthalmology/articlepdf/624622/archopht_54_2_010.pdf. https://doi.org/10.1001/archopht.1955.00930020231010.

[74] John Trader. 2015. The Difference Between 1:N, 1:1, and 1:Few and Why it Matters in Patient ID. (September 2015). https://www.rightpatient.com/blog/the-difference-between-1n-11-and-1few-and-why-it-matters-in-patient-id/.

[75] Andreas Uhl, Christoph Busch, Sébastien Marcel, and Raymond Veldhuis. 2020. *Handbook of Vascular Biometrics*. (January 2020). ISBN: 978-3-030-27730-7. DOI: 10.1007/978-3-030-27731-4.

[76] Julia Wainger. 2022. Statistical Hypotheses and Error, (April 2022).

[77] Mingle Xu, Sook Yoon, Alvaro Fuentes, and Dong Sun Park. 2023. A Comprehensive Survey of Image Augmentation Techniques for Deep Learning. *Pattern Recognition*, 137, 109347. ISSN: 0031-3203. DOI: https://doi.org/10.1016/j.patcog.2023.109347. https://www.sciencedirect.com/science/article/pii/S0031320323000481.

[78] Jin Yao, Jing Zhao, and Cui-ying Hu. 2012. Near-infrared imaging approach for in vivo detecting the distribution of human blood vessels. In *2012 International Conference on Biomedical Engineering and Biotechnology*. IEEE, pp. 841–844.

[79] Lizhen Zhou. 2023. Finger Vein Recognition Technology: Principles, Applications, and Future Prospects. *International Journal of Biology and Life Sciences*, 3, (July 2023), 45–48. DOI: 10.54097/ijbls.v3i2.10852.

[80] Andrew Zola. 2024. Retina scan. (March 2024). https://www.techtarget.com/whatis/definition/retina-scan.