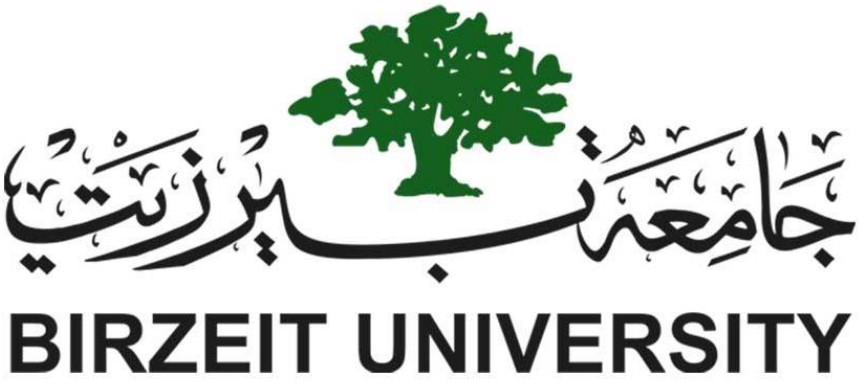


BERZIET UNIVERSITY
FACULTY OF ENGINEERING AND TECHNOLOGY
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

Secure Authentication through AI-Powered Contactless Palmprint and Palm Vein Recognition

---

Prepared by:
Layan Shoukri 1201225
Amal Butmah 1200623
Duaa Suliman 1200909

Supervised by:
Dr. Wasel Ghanem

A graduation Project submitted to the Department of Electrical
And Computer engineering in partial fulfillment of the requirements
For the degree of B.Sc. in Electrical Engineering

BIRZEIT
July 2024

Abstract

After the Coronavirus invaded the world, all the world's concepts and outlooks changed, so it became necessary for technology to keep pace with the aspirations of the post-Corona world. One of the effects of this pandemic is that people have become worried about the physical touch with public use devices, especially those who use fingerprint-based authentication devices. This project meets the desires of a large segment of people who have become obsessed with non-contact, as a contactless device will be developed to take images of the palm print and palm vein, which are unique features that are difficult to imitate. In contrast, the device does not require direct contact with the skin, which makes the experience comfortable for users. The image of the palm that an IR camera will take is applied to the algorithm to extract the veins through which each person will be distinguished from another which provides high accuracy in recognition and identity verification. The compatible size of the device, suitability for all authentication applications, ease of use, and low cost will allow it to be available at a low price in the future, making the device special.

Index Terms— Palmprint, Vein, Recognition, CNN, SNN, Raspberry Pi, IR.

المستخلص

بعد أن اجتاح فيروس كورونا العالم، تغيرت كل مفاهيم العالم ووجهات نظره، لذا أصبح من الضروري أن تواكب التكنولوجيا تطلعات عالم ما بعد كورونا. ومن آثار هذا الوباء أن الناس أصبحوا قلقين بشأن اللمس الجسدي مع أجهزة الاستخدام العام، وخاصة أولئك الذين يستخدمون أجهزة المصادقة القائمة على بصمات الأصابع. يلبي هذا المشروع رغبات شريحة كبيرة من الأشخاص الذين أصبحوا مهووسين بعدم التلامس، حيث سيتم تطوير جهاز بدون تلامس لالتقاط صور بصمة الكف وأوردة الكف، وهي ميزات فريدة يصعب تقليدها، في حين أن الجهاز لا يتطلب اتصالاً مباشرًا بالجلد، مما يجعل التجربة مريحة للمستخدمين. ويتم تطبيق صورة الكف التي سيتم التقاطها بواسطة كاميرا متعددة الأطياف على خوارزميات معينة لاستخراج الأوردة التي من خلالها سيتم تمييز كل شخص عن الآخر مما يوفر دقة عالية في التعرف والتحقق من الهوية. إنَّ حجم الجهاز المتوافق، وملاءمته لجميع تطبيقات التوثيق وسهولة الاستخدام والتكلفة المنخفضة ستسمح بتوفيره بسعر منخفض في المستقبل وهذا ما يجعل الجهاز مميزًا.

Contents

Acronyms and Abbreviations

IR	Infrared
SD	Secure Digital
CLAHE	Contrast-limited Adaptive Histogram Equalization
LED	Light Emitting Diode
PCA	Principal Component Analysis
NPE	Neighborhood-Preserving Embedding
SDSPCA-NPE	Supervised Discriminative Sparse Principal Component Analysis - Neighborhood-Preserving Embedding
CPU	Central Processing Unit
I/O	Input / Output
SDRAM	Synchronous Dynamic Random Access Memory
HDR	High Dynamic Range
ISP	Image Signal Processor
ROI	Region of Interest
CNN	Convolutional Neural Network
SNN	Siamese Neural Network
ORB	Oriented FAST and Rotated BRIEF
LoG	Laplacian of Gaussian

List of Figures

List of Tables

Chapter 1

Introduction

1.1 General Introduction

The subject of technology and systems in identification has remained a topic of discussion as technology advances and the need for a proper implementable security measure across the public grows. These systems have been adopted in areas such as access control, security, employee management, banking, health, and so on. Biometric aspects and unique features have garnered the most attention in these biometric authentication technologies such as fingerprints, face recognition, and iris scans, for their effectiveness and security as demonstrated in some research. More recently, the distinctiveness of palmprints has been noted and the concept of vein patterns as a biometric modality came to the foreground only towards the end of the 20th century and this is the topic that our project is based on [1].

To build a contactless system for palmprint and palm veins, we worked on researching the necessary components for the device and concluded with the basic components a high-resolution IR camera and infrared lights. We have undertaken a long search to provide these components, which has included many problems that we will talk about in the challenges section. Therefore, we initially decided to work on building a device that captures only the palmprint and not the veins, and during this, we will work on developing the algorithm to work on the palmprint and the veins together based on open-source datasets. We have ordered the basic components to build the device that will capture the palmprint: an IR-led strip of wavelength 850nm, 64GB micro-SD, Raspberry Pi, and Raspberry Pi noir camera. When the device is built, it will be able to take detailed pictures of the palm to take our data, while the program processes these images, extracts the unique region of interest, applies feature extraction, and feature matching, determines the result, and verifies the identity.

1.2 Problem statements

As mentioned before, the most popular user identification and verification systems are fingerprints, facial recognition, and iris scanning. However, it has been shown that these methods based on the

recently mentioned features suffer from some problems, the most important of which is the theft of biometric information such as the fingerprint, in addition to the lack of hygiene resulting from the finger touching the verification device. As another example of the problem, there is iris recognition technology, which is greatly affected by lenses. As for facial recognition, many people do not want to take pictures of their faces, in addition to the most important problem, which is the high cost of equipment for these methods and technologies, and thus the high costs of selling them. All of these contributed to the trend towards the palmprint and the palm vein recognition system, which solved most of the problems facing user identification and verification systems [1].

1.3 Motivation

Compared to identification systems, the palm is not considered sensitive to privacy compared to facial recognition. The absence of touch in this feature made the system that relies on it for identity recognition a healthy system that gained the admiration and interest of its users, especially after discovering the uniqueness of the palm features i.e. palmprint, which is rich in features such as lines and spots, and palm veins which are merged. Thus, a contactless identification system appeared with more safety, reliable, accurate, and highly comfortable due to its reliance on features that studies have proven to be unique and distinctive among people, in addition to the difficulty in capturing and stealing, also its presence under the skin tissue has increased the difficulty of forming it [2].

1.4 Device work description

After building our device, we will be able to take detailed pictures of the palm to assemble our data. After taking the palm picture, the processing of the captured image begins, through which the region of interest is extracted, which ensures the minimum loss of vital information and the highest recognition accuracy at the same time, which requires less processing time and less use of memory. Several algorithms are used in the processing stage to ensure that the region of interest is extracted with high accuracy, as algorithms based on deep learning are trained on a large set of general images of palmprints. Finally, the captured image is matched with the one stored in the database to complete identification. The most important feature of this system is its low total cost, which will allow it to be available in the future at very reasonable prices compared to currently existing systems, and with greater accuracy.

1.5 Organization of the report

The report is organized into five chapters and the remaining chapters of this report are organized as follows:

- Chapter 2 presents the overall background and prior works.
- Chapter 3 introduces datasets, hardware, and algorithms used, and the methodology followed.
- Chapter 4 describes the experiments and results analysis.
- Chapter 5 concludes the work presented, presents the challenges faced in this project and introduces some possible future works.

Chapter 2

Literature Review

Before any crucial step in graduation projects, such as this one, the topic of the proposed project is studied with a superficial and quick reading of previous similar projects. It is then read deeply and carefully to collect useful and comprehensive information about the project after it is agreed upon. This is one of the first and most important steps taken before starting anything related to the project, such as extracting features, building the system, and so on. There was great credit for previous relevant studies, research, and papers, which helped in understanding the project, knowing its characteristics, its importance, and how it is likely to be applied.

In this chapter, we will discuss the background and review related works related to the development of palmprint-based and palm-vein-based authentication systems.

2.1 Introduction to Contactless Biometric Authentication Using Palm Print and Palm Vein Recognition

As already mentioned in the chapters of this report, it is necessary to reaffirm the reliability of biometric authentication devices that take advantage of people's unique physiological or behavioural characteristics for identification and verification. Many of these characteristics have been mentioned that favour it over other traditional authentication methods because of its ability to provide high security due to the difficulty of theft and forgery. One of the most important of these biometric characteristics on which our project was based is the palmprint of the hand and its veins, which are considered distinctive characteristics.

One of them is visible, which is the palmprint and the other is invisible and can only be seen using near-infrared light, which are the palm veins. What makes the biometric authentication system based on the palmprint and its veins an interesting device and a focus for study and research is the lack of need for physical contact with the sensor of the authentication device, as these biometric features are contactless, which automatically leads to the user feeling comfortable, especially for those who have concerns regarding hygiene-related problems. This feature also allows for the long life of the authentication device, as it does not corrode due to lack of touching.

2.2 The Beginning of the World of palmprint-based identification systems

After the distinctive and unique characteristics of the palmprint were revealed among people, Zhang and others decided in 2003 to work on a palmprint-based identity verification device using low-quality images in which a two-dimensional Gabor coding scheme was used to extract features and represent the palmprint. The device was primitive and relatively large and its use requires hand contact with the device, as shown in Fig. 2.1. However, it achieved good results in terms of speed and accuracy, with a reasonable true acceptance rate of 97% [3].

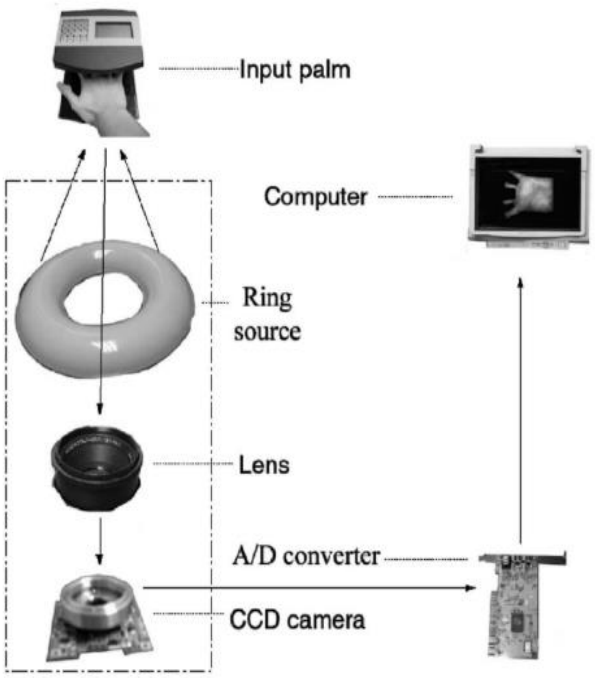


Fig. 2.1: Schematic diagram of the palmprint capture device [3]

In 2007, Han and others presented the first system that captures images of the palm without direct contact with the device in unrestricted scenes. They used two cameras, one of which photographed near-infrared rays to determine the shape of the hand, and the other to capture palmprints in visible light to extract features from them using a new algorithm based on using color and shape information of the palm [4].

2.3 Quick overview of the palm veins

The randomness of the vein patterns in the palm is genetic, and even twins each have a vein pattern different from the other, and this is what makes it a focus of interest. The most interesting aspect of it

is that it is invisible and hidden inside the body. It cannot be seen with the naked eye, so it cannot be impersonated or its representation can be left anywhere.

To take a picture of the veins, the palm or any part of the body to see its veins is exposed to near-infrared light, and then certain devices that sense this light are used, such as some cameras[5].

Many biometric authentication devices have used palm veins, and Gunjan Shah in 2015 implemented this in practice using mainly near-infrared light and a camera that takes images at that frequency. Several algorithms were used to process images, the best of which was the template matching technique, producing a system with a good accuracy of 93.54% [6].

Research is still achieving tremendous development and achievement with the advancement of technology and the accuracy of the tools used, achieving results with much higher accuracy.

2.4 Palmprint and palm veins recognition system development according to the reviewed papers

Developments and research continued in the field of devices based on biometric characteristics, as researchers were not satisfied with the handprint alone in the field of identity verification, especially since these palmprints are affected by external factors to which the hand is exposed, the most important of which is age and some works that erase these palmprints. They have worked hard to increase the reliability of such systems to become palm-based and veined.

In recent years, interest in palmprint and palm vein recognition technology has begun, and there have been many researches that discussed and followed the process of developing the design of the palmprint and palm vein recognition device to reach a device with high efficiency and low costs. Several hardware components and algorithms in the construction and design process were followed.

Table 2.1 shows some of the many contributions made to the topic of palmprints and palm veins recognition using various techniques and data sets, with the accuracy percentage for each contribution. The table itself shows the number of attempts and achievements in the same subject through the innovation of numerous techniques and the production of different data sets, and this indicates the importance of this system and the attempt to develop it to the extent that makes it very accurate.

Contribution	Technique	Dataset(s) used	Recognition Accuracy
Mirmohamadsadeghi and Drygajlo (2011) [7]	Local derivative pattern, histogram intersection	CASIA	98.30%
Lee (2012) [8]	2D Gabour filter, Hamming distance	Self-constructed	99.18%
Kang and Wu (2014) [9]	Local Binary pattern, Support Vector Machine	CASIA	100%
Elansir and Shamsuddin (2014) [10]	Linear Discriminative analysis, Cosine distance	PolyU	99.74%
Lu et al (2016) [11]	Multi-scale Local Binary pattern, Local derivative pattern, similarity measure	CASIA	99.99%
Cho and Kar-Ann (2018) [12]	Gabor filter, Hamming distance	PolyU	99.13%
Hernández-García et al (2019) [13]	CLAHE	CASIA Multi-Spectral Palmprint Images	99.28%
Pititheeraphab et al (2020) [14]	Geometric affine invariants	Self-constructed	99.76%

Table 2.1: Contributions to palmprint and palm vein recognition [2]

The following paper was the most recent paper we read that studied and applied in practice the subject of identification using the palmprint and palm vines features. This paper was produced in 2023 by Chinese researchers and it was the main reference for everything we needed during our research, so it has the largest share in this section to explain, what components and algorithms were used to build the system.

Wei Wu, Yunpeng Li, Yuan Zhang, and Chuanyang Li were interested in building an Identity Recognition System Based on Multi-Spectral Palm Vein Image, they built the contactless device that is shown in Fig. 2.2, from two Near-Infrared LED CST brand model BL-270-42-IR with 850nm wavelength which are equipped with an intensity adjustment controller and a camera model MV-VD120SM with multispectral imaging which uses multiple spectral ranges to capture palm veins with more detailed information [15].

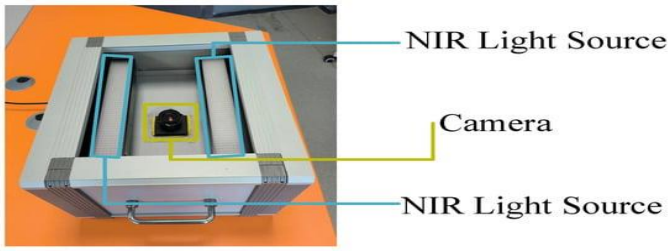


Fig. 2.2: Contactless device to capture multi-spectral palm vein image[15]

For the methodology, they followed four steps. First, they chose the region of interest in the image pre-processing step which was a square whose side length is equal to the distance between the valleys between the little finger and ring finger, the index finger and middle finger, and the square shape starts from these valleys towards the inside of the hand. Then, they adopt supervised discriminative sparse principal component analysis (SDSPCA-NPE), which combines supervised discriminative information and sparse constraints into the PCA model, in feature extraction which reduces the dimensionality and recognition and the projected palm vein data exhibits an improved distance distribution, enhancing the classification performance of palm vein data. After that, the next step was to test the feature matching and recognition which is based on Euclidean distance between feature vectors derived from the projected subspace with matching thresholds determined by the intersection of intra-class and inter-class matching curves [15].

This paper used several datasets, including the CASIA database, the Hong Kong Polytechnic University database, the Tongji University database, as well as several self-contained datasets. To increase their accurate research, they used several algorithms, where the results are shown in Table 2.2, which shows some of the algorithms used for each type of data set, with the results resulting from equal error rates and identification time [15].

Algorithms	Database	EER (%)	Time($\mathbf{10^{-4}}$s)
PCA	Self-built	0.28	19.59
	CASIA	2.38	19.77
	PolyU	1.5	19.45
	Tongji	6	19.58
NPE	Self-built	0.50	13.81
	CASIA	7.50	13.90
	PolyU	1	14.19

	Tongji	9.6	14.11
SDSPCA	Self-built	1.50	13.56
	CASIA	15.39	13.89
	PolyU	5.50	13.57
	Tongji	10.75	13.63
SDSPCA-NPE	Self-built	0.10	19.77
	CASIA	0.50	38.50
	PolyU	0.16	19.75
	Tongji	0.19	19.69

Table 2.2: Algorithms on different datasets [15]

After evaluating multiple algorithms and their respective results, as depicted in the table above, the combination of SDSPCA and NPE proved effective, showing a significant reduction in the Equal Error Rate.

Chapter 3

System Implementation and Design

This chapter introduces the system construction including hardware and software, where the components of our recognition device, the algorithms used in programming the system, especially those related to the machine and deep learning, the datasets, and the methodology will be discussed.

3.1 Datasets

Some of the related works discussed in Chapter 2 touched on the open-source datasets they used. Among these works, there were 2 common datasets used i.e. CASIA Version1.0 and PolyU-IITD V1.0 which were researched and used to build the model initially until assembling our dataset.

CASIA Multi-Spectral Palmprint Image Database Version 1.0 is a collection of 7200 palm vein images that were captured from 100 different hands by a multispectral camera, each image is an 8-bit gray-level JPEG file. The images of each hand were taken in two sessions, one month apart. In each session, images were taken at six different electromagnetic spectrums, i.e. 460nm, 630nm, 700nm, 850nm, 940nm, and white light while every six images were taken at a different spectrum [16]. Fig. 3.1 shows samples from this dataset.

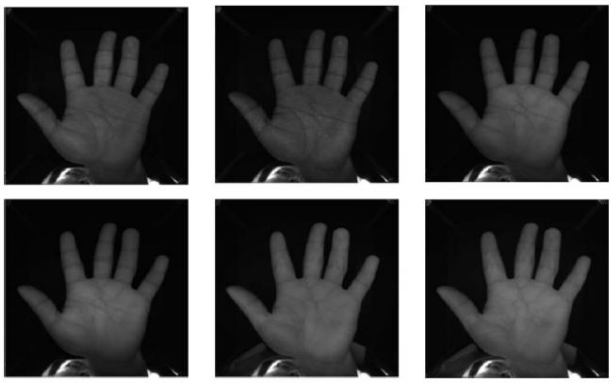


Fig. 3.1 : Samples from CASIA dataset [16]

IIT Delhi Touchless Palmprint Database (Version 1.0) consists of palmprint images that were taken from 230 different people from IIT Delhi University students and employees in the age group 12-57 using contactless imaging device during July 2006 - Jun 2007 [17].

Seven images were taken for each hand from each person in different hand positions. All were captured in a closed small room for the hand, using a camera with a circular fluorescent illumination around its lens [17].

The images were in bitmap (*.bmp) format with a resolution of 800 × 600 pixels, in addition to 150 × 150 pixels cropped and normalized palmprint images [17]. Fig. 3.2 shows samples from this dataset.

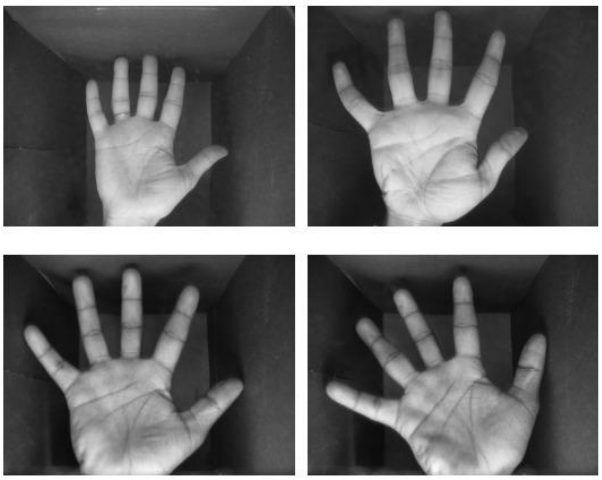


Fig. 3.2: Samples from IIT Delhi dataset [17]

3.2 Hardware Components

Our device will include the following components:

3.2.1 Raspberry Pi 5

Raspberry Pi is considered to be a minicomputer that uses different kinds of processors which makes it efficient in performing several tasks. It has Raspbian operating system based on Linux, and it supports installing only open-source operating systems and apps. Raspberry Pi also supports several programming languages such as Python, C, BASIC, Ruby, Perl, and C++ [18].

Raspberry Pi 5 is the latest version of a single-board computer, it is integrated with a 1.6GHz Processor. Also, it includes Wi-Fi, Bluetooth, Micro SD card storage, LPDDR4X-4267 SDRAM (4GB or 8GB), and many of computer features. It is 2 to 3 times faster in CPU performance and twice the memory and I/O bandwidth over the previous models [19].

The Raspberry Pi 5 which is shown in Fig. 3.3, was chosen as a processing unit. Because it has an improved Image Signal Processor (ISP) which processes the raw data from the camera sensor into a high-quality image. In addition to its HDR imaging feature which combines images to prevent the highlights from blowing out and dark areas from being too dark [20].

Fig. 3.3: Raspberry pi 5 [19]

3.2.2 Raspberry Pi camera module – pi noir 8MP Version2

The Raspberry Pi NoIR which is shown in Fig. 3.4, is a camera module with no infrared filters which makes it sensitive to infrared light at wavelengths from 700nm and up. It can take clear pictures in the dark by applying an infrared light on the object. It includes a Sony IMX219 sensor which can capture IR light in addition to visible light, to give images with high resolution, up to 3280 × 2464 pixels (8megapixels) [21].

The Raspberry Pi NoIR is compatible with all Raspberry Pi models and can be connected using Camera Serial Interface (CSI), and used in different applications such as security systems and scientific projects [22].

Fig. 3.4: Raspberry Pi Noir camera module [21]

This camera is used for palm vein imaging. According to Stefan-Boltzmann radiation law, which is a relation between the IR radiation captured by an IR camera and the body temperature, a thermal image

of the veins can be captured using an IR camera with spectrum sensitivity of 3-5μm or 8-14μm, due to the difference between the vein's temperature and the surrounding skin temperature [23], [24]. Also, Planck's emission law which is a function of temperature and wavelength, describes the heat radiation intensity of the black body[24].

Stefan-Boltzmann law declares that the amount of heat radiation that an object can emit has a maximum limit. Also, thermal energy is emitted from an object based on its material which has an emissivity constant between 0 and 1 based on the object's temperature. Equation 3.1 illustrates the Stefan-Boltzmann law [24].

$$\boldsymbol{P = \varepsilon(T) * \sigma * T^4} \qquad \textbf{(3.1)}$$

Where,

P = Power radiated,

ε(T)= emissivity constant,

σ = 5.67*10−8 W/m2K4 (Stefan-Boltzmann constant),

and T= Temperature in Kelvin

Planck's emission law describes electromagnetic wave energy as proportional to its frequency. Equation 3.2 illustrates this law [24].

$$\boldsymbol{E = h * f = h * c\lambda} \qquad \textbf{(3.2)}$$

Where,

E = Energy

H = 6.626*10−34 [Js], Planck's constant

f = frequency

c = speed of light in a vacuum

and λ = wavelength.

The diagram in Fig. 3.5 demonstrates that at human temperature, the peak value of radiation intensity is at a wavelength of around 10 μm. This radiation is in the infrared domain that the IR camera can detect.

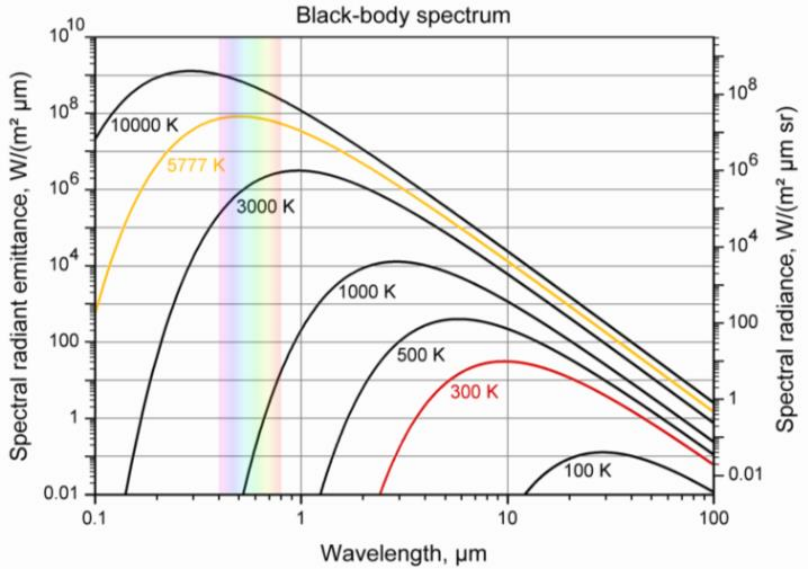


Fig. 3.5: Black body spectrum [24]

3.2.3 Micro SD card

Micro-SD is a removable flash memory card used as storage in several devices such as smartphones, digital cameras, USB flash devices [25] and Raspberry Pi's [19]The design consists of one or more flash memory dice and a microcontroller which acts as an interface between the card reader and the memory to supply the storage space abstraction [26]

As a benefit, the micro-SD provides a low-cost, fast writing speed and a large capacity of 400GB additional storage for electronic devices [26].

In our device, we use a 64GB micro-SD card as the one shown in Fig. 3.6, to save all the images that are captured by the camera. This card is put in a micro-SD card slot inside the Raspberry Pi to process the images.

Fig. 3.6: 64GB Micro-SD card

3.2.4 IR light

Infrared radiation is an invisible range of the electromagnetic spectrum, from the end of the red light of the visible light to the microwave range. It is divided into three regions with different wavelength ranges i.e. near-infrared with a range from 780nm to 2500nm wavelength, middle infrared with a range

from 2500nm to 50000nm wavelength, and far infrared with a range from 50000nm to 1mm wavelength[27].

Infrared waves are important in the greenhouse effect which led it to be used in several applications such as thermal imaging, data networking, and telecommunications [27].

In our project, an infrared light of 850nm wavelength, as the one shown in Fig. 3.7, was chosen to help in observing the veins. Because when this light is projected onto the palm, its absorption rate varies between different parts of the hand. The rate of absorption in blood hemoglobin (including oxygenated and deoxygenated hemoglobin) is high, which leads to a difference in the degree of shadow that is formed as a result of absorption. Thus, the locations of the veins are determined, and the image of them is generated [15].

Fig. 3.7: 850 nm IR led strip

3.3 Used algorithms and techniques

Our work is based on the following algorithms and techniques:

3.3.1 Traditional Algorithms and Techniques

The overall image processing process contains many traditional processing techniques, many of which have been used. Some techniques have been used to separate images from noise using filtering methods such as Gaussian low-pass filter to remove noise and smooth images. Laplacian of Gaussian (LoG) filter detects edges at different scales, and Gabor filters to detect textures and edges. ORB (Oriented FAST and Rotated BRIEF) plays a vital role in object detection by efficiently detecting and describing key points. Techniques like erosion which is a morphological operation remove small objects from the image. Rescaling which changes the dimensions of the image, noise reduction, Canny edge detection, thresholding, and converting the grayscale image to a binary image are all used to enhance the image processing for a wide range of applications.

3.3.2 Deep Learning Algorithms

Our system involves using deep learning to help improve the level of precision and effectiveness in producing the results. The CNN helps in learning and differentiating the spatial patterns concerning the inputs while the SNN helps to conclude the similarity of the inputs.

3.3.2.1 CNN

Convolutional Neural Networks (CNNs) are a prominent class of deep neural networks that extend the concept of multilayer perceptrons (MLPs), also known as feed-forward networks. The core component of CNNs is the convolution operation, which involves applying filters to input data matrices to produce feature maps. Unlike simple matrix multiplication, the convolution process captures local patterns and enhances important features in the data. This can help in reducing sensitivity to noise and improving data regularization [28]. Figure 3.8 illustrates the architectural layout of a typical CNN.

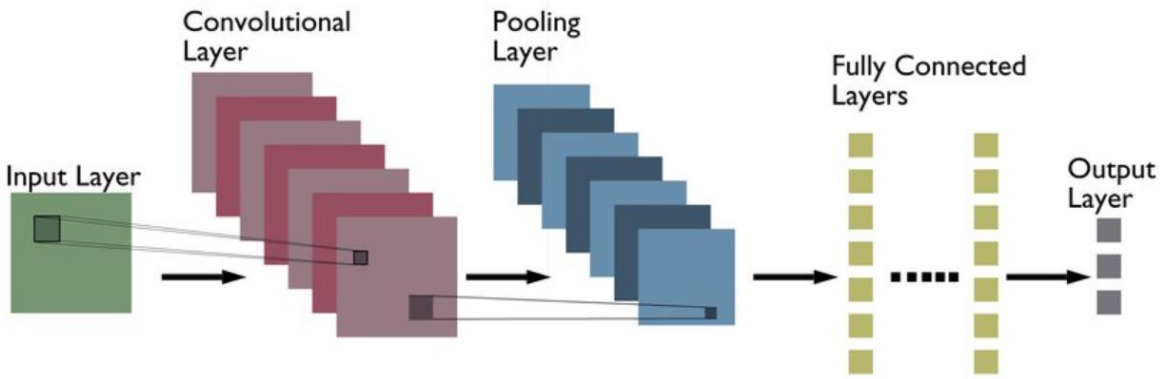


Fig. 3.8: CNN Architecture [28]

The convolution operation in CNN has equation 3.3:

$$(\boldsymbol{I} * \boldsymbol{K})(\boldsymbol{x}, \boldsymbol{y}) = \sum_i \sum_j \boldsymbol{I}(\boldsymbol{x}+\boldsymbol{i}, \boldsymbol{y}+\boldsymbol{j}) . \boldsymbol{K}(\boldsymbol{i}, \boldsymbol{j}) \qquad \textbf{(3.3)}$$

where I: Input image, K: Kernal (filter), x, y: Spatial coordinates in the output feature map, i, j: Offsets within the kernel window. And the output is the convolved value at (x, y) which is (I * K) (x, y) [29].

The architecture is divided into several layers. The input layer is the first, where the model obtains the input image. The second phase of CNN involves convolutional layers, which are composed of a set of learnable filters that are applied to a specific area of the input image in order to reduce the image's dimensions. without losing any of the image's key elements [30]. The convolution layer's operating concept is depicted in Fig. 3.9.

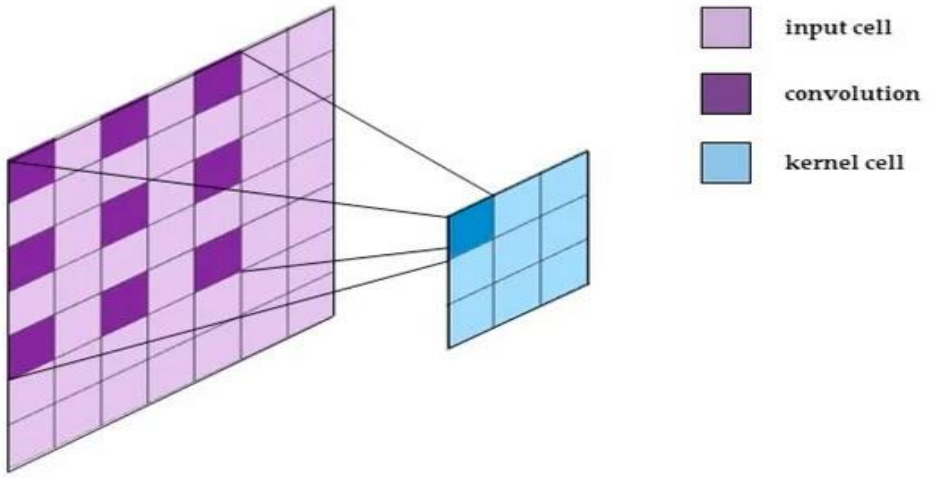


Fig. 3.9: Convolutional layer in CNN [31]

The third stage of the CNN architecture is the pooling layer, which reduces the dimensions (width and height) of the input while preserving the most important information. Max pooling is a widely used technique in this layer. It is a type of non-linear down-sampling that divides the input image into non-overlapping rectangles and selects the maximum value from each region. This process not only speeds up computation but also enhances the robustness of the identified features [32]. The operation of the pooling layer is illustrated in Fig. 3.10.

Equation 3.4 is the max pooling equation:

$$\boldsymbol{Y(i,j) = max\{x(s,t) \mid s \epsilon [i, i + n - 1], t \epsilon [j, j + m - 1]\}} \qquad \textbf{(3.4)}$$

where x: input feature map, (i, j): Coordinates in the output feature map, n, m: Size of the pooling window, and s, t: Coordinates within the pooling window. And the output which refers to the maximum value in the pooling window centered at (i, j).

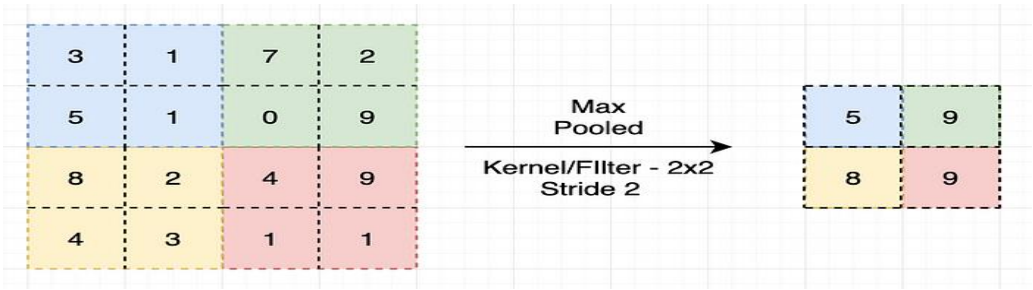


Fig. 3.10: Pooling layer in CNN [33]

The final stage of the CNN involves the fully connected layer and sometimes dropout layers. In the fully connected layer, each neuron is connected to every activation in the previous layer, producing class predictions. Dropout layers are used in CNNs to increase test accuracy and reduce overfitting by randomly turning off a fraction of neurons during training. Typically, a dropout rate of 0.5 (turning off 50% of the neurons) is used. Turning off more than 50% of the neurons can weaken the model's learning and lead to inaccurate predictions [34]. The operating principle of the fully connected layer is illustrated in Fig. 3.11.

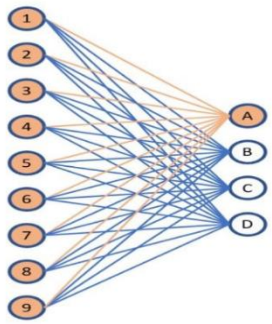


Fig. 3.11: Fully connected layer in CNN [35]

3.3.2.2 SNN

A Siamese network is a type of artificial neural network that consists of two or more identical sub-networks, meaning they share the same structure, parameters, and weights. Typically, only one of the sub-networks is trained to solve the chosen problem, while the others maintain the same configuration, parameters, and weights [36].

Siamese networks are used to determine the similarity level between inputs by matching their feature vectors. Unlike other neural networks in deep learning, a Siamese network requires fewer examples of each class to construct a good model. This means Siamese networks do not learn to recognize images as belonging to specific output classes. Instead, they use a similarity function that takes two images as input and returns a probability indicating how similar the images are [36].

If we have two images, we can cross-check them to determine if they belong to the same category by evaluating whether they form similar or dissimilar pairs.

The first subnetwork processes image A by applying convolutional and fully connected layers, resulting in a feature vector representing the image. Subsequently, the same network architecture, with identical weights and parameters, processes the second image (B) to generate its corresponding feature vector.

Having the encodings E(A) and E(B) for their respective images allows us to compare the similarity

between the two images. Given that the images are similar, their corresponding encodings should also be similar. The distance between these two vectors indicates their similarity: a smaller distance suggests that the images are closely related or belong to the same class, whereas a larger distance indicates greater dissimilarity [36]. Figure 3.12 illustrates the SNN architecture.

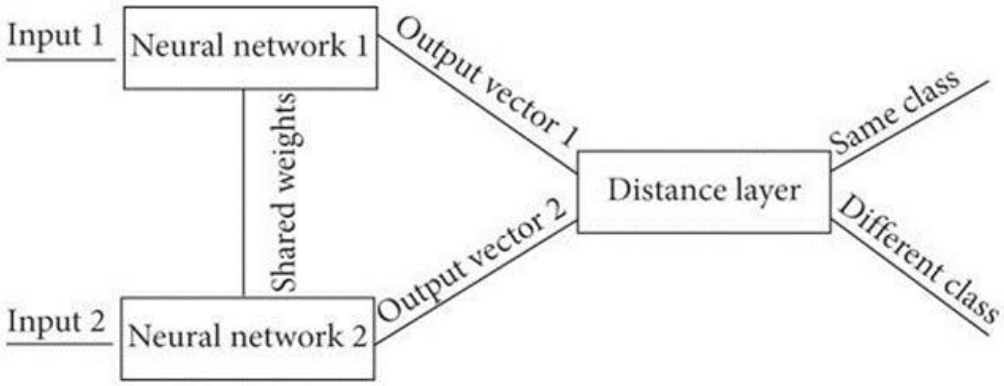


Fig. 3.12: SNN Architecture [36]

3.3.3 Machine learning algorithms

In our current work on contactless palmprint and palm vein detection, our system incorporates an authentication mechanism utilizing state-of-the-art machine learning algorithms. Specifically, we have employed K-Means Clustering as the most suitable method for clustering.

3.3.3.1 K-Means Clustering

K-Means is an unsupervised learning algorithm widely used in data science and machine learning to address clustering problems. It partitions the dataset into distinct clusters. The parameter K specifies the number of clusters to be created during the process; for example, setting K to 2 results in two clusters, while K set to 3 produces three clusters, and so forth [37].

A density plot is first constructed for the unlabelled dataset, followed by partitioning the dataset into exactly k clusters, where each cluster contains data points with similar features. This method efficiently identifies distinct groups within an unknown dataset without requiring prior exposure or training, categorizing the data into various groups [37].

In this technique, each cluster is associated with a centroid, adhering to centroid theory. The primary goal of the algorithm is to minimize the sum of distances between every data point and its corresponding cluster centroid. Initially applied to an unlabelled dataset, the algorithm iteratively divides the data into the specified number of clusters 'k', refining the clusters until optimal groupings are achieved. It's

important to note that the value of 'k' must be predetermined for this procedure [37].

The two primary functions of the K-Means clustering method are as follows: Firstly, it iteratively determines the optimal positions of 'K' centroids or center points within the dataset. Each data point is then assigned to the cluster whose centroid has the closest mean value to that point. Secondly, each cluster comprises data points located in proximity to a specific centroid [37].

Each cluster is delineated by boundaries and contains elements that share certain similarities while remaining distinct from other clusters. Fig. 3.13 shows a diagram that explains the working of the K-Means Clustering Algorithm.

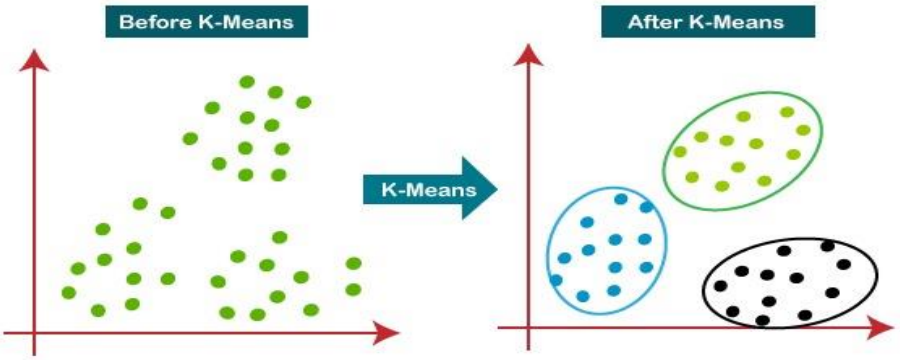


Fig. 3.13: K-means clustering [37]

3.4 Methodology

For the palmprint recognition mission in this project, we will detail four main steps based on the flowchart shown in Fig. 3.14.

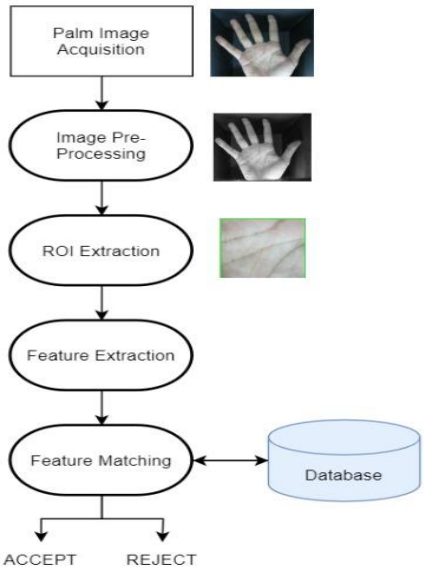


Fig. 3.14: System flow chart

3.4.1 Palm Image Acquisition

In this step, the acquisition of a clear image of the palm is crucial for subsequent data analysis. The image should ideally be of high resolution to enable effective auto-encoding.

3.4.2 Image pre-processing

In this step, preparations are made on the palm image to enhance its suitability for further analysis. Image pre-processing is crucial as it prepares the image for subsequent stages in the method, such as ROI extraction, feature extraction, and matching. Effective image pre-processing techniques include filtration, noise removal, binarization, segmentation, and conversion to grayscale.

3.4.3 ROI Extraction

In this step, the region of interest (ROI) is extracted from the palm representation image. This involves identifying and outlining segments of the image that are rich in features. Various techniques can be employed for ROI extraction to isolate the necessary area.

3.4.4 Feature Extraction

After identifying the ROIs, the next step involves extracting features from these areas. Algorithms analyze the ROIs to identify specific characteristics such as patterns, texture, and other discriminative features between palm images. Features such as texture, shape, histogram, angles, and others can be extracted to facilitate this process.

3.4.5 Feature Matching

The extracted features are compared with those stored in the database using matching algorithms after the model is trained. If the features extracted from the input image match with any of the features in the database, the palm image is accepted. Otherwise, the palm image is rejected.

Chapter 4

Simulation Results

This chapter is devoted to the presentation of the results of simulation of the system. We begin with the extraction of Regions of Interest (ROI) over the entire dataset, one the basic steps in data preprocessing. After this, we discuss the model training phase. Lastly, we go over the feature-matching procedure.

4.1 ROI Extraction

We worked on a code that sets up a fully functional image processing and machine and deep learning flow optimized for a palmprint recognition system based on SNN and CNN. The operation of the system can be represented by a set of steps which include pre-processing of inputs that are the images, feature extraction, as well as classification of the images based on the palmprint. Here is an in-depth explanation of the pipeline:

1. Data Preprocessing
    - Image Loading and Display: The pipeline starts with loading the image that the verification process will be applied on it. This image as shown in Fig. 4.1 is displayed with the help of the matplotlib package for the visualization of the results.

Fig. 4.1: Original image

- Image Padding and Conversion: This is done to standardize the size while working with the images as some of them may be of different sizes. After that, the image is transformed from RGB format to grayscale format as shown in Fig. 4.2 to reduce the complexity of the further processing steps.

Fig. 4.2: Grayscale image

- Thresholding: Otsu's thresholding technique is used to segment the image of the hand from the background and for this, the images are converted to binary format, the center of the hand is defined which is important for the subsequent steps as shown in Fig. 4.3.

Fig. 4.3: Center of the hand

2. Feature Extraction

- Hand Contour Detection: based on erosion and boundary detection techniques the contour of the hand is then detected.
- Valley Points Detection: the valley points are the points between the fingers as shown in Fig. 4.4, these points are essential for setting the selected ROI and they are critical to explaining the hand's configuration.

Fig. 4.4: Valley points

- Convex Hull and Angle Calculation: The convex hull of the hand is the outermost boundary of the hand. It is calculated to determine the fingertips, as shown in Fig. 4.5, Which in turn helps in aligning hand angles, computing rotation and extracting ROI.

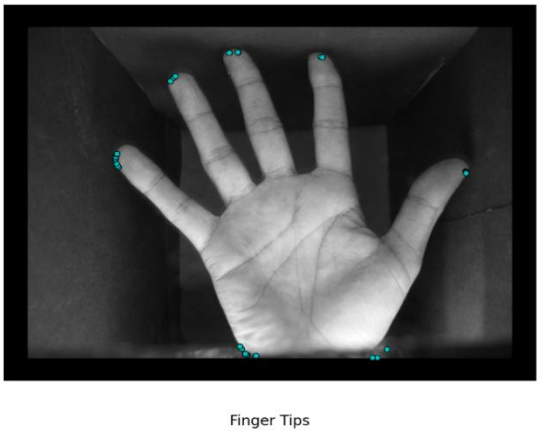


Fig. 4.5: Fingertips

The angle between some or all of the valley points and the midpoint of the hand is computed to properly orient the hand. The middle fingertip is singled and taken as a reference while calculating other angles from it as shown in Fig. 4.6.

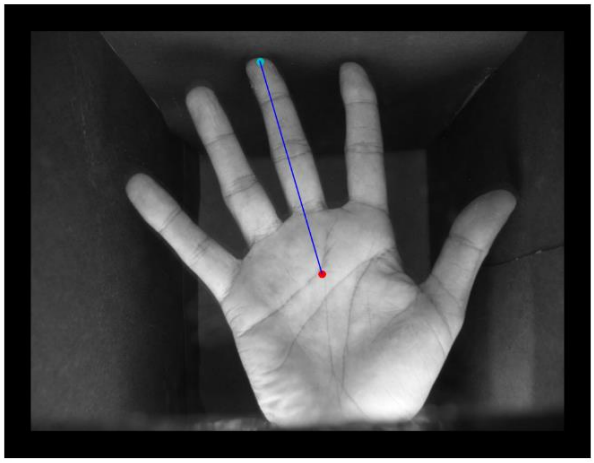


Fig. 4.6: Middle fingertip

The first valley point is aligned with the center of the hand and the middle of the fingertip, as shown in Fig. 4.7, in order to determine the designated number of rotations which is required for efficiency.

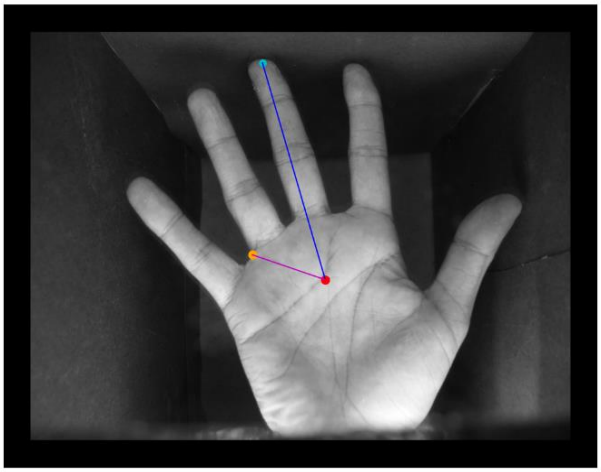


Fig. 4.7: First valley point

The second valley point is applied for rotation and proper positioning of the hand as well as shown in Fig. 4.8.

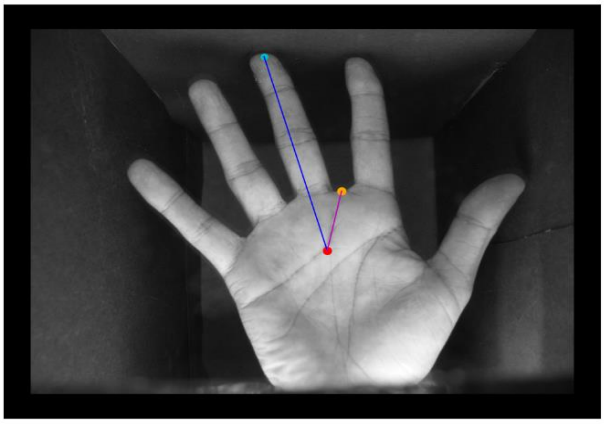


Fig. 4.8: Second valley point

Fig. 4.9 and Fig. 4.10 demonstrate the points that have been determined as the valley points after filtering and angle measurements which are used ROI determination.

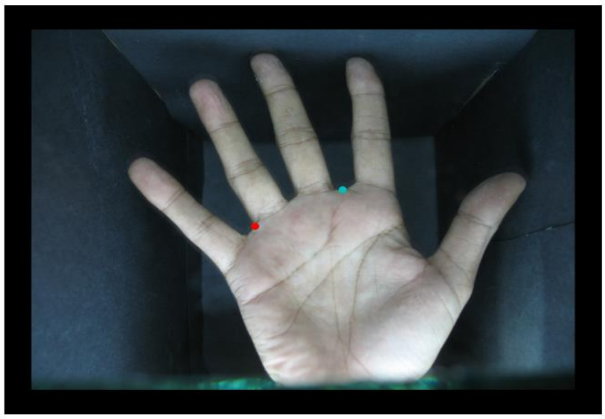


Fig. 4.9: Valley points

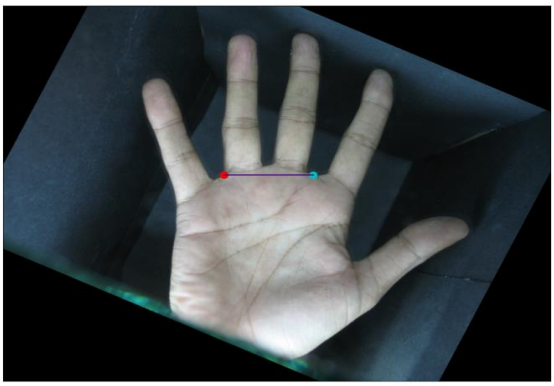


Fig. 4.10: Image rotation and valley point connection

- Region of Interest (ROI) Extraction: According to the valley points and convex hull traced here the ROI is obtained. This region is where most of the useful information in the hand is located – its primary elements are critical for gesture interpretation.

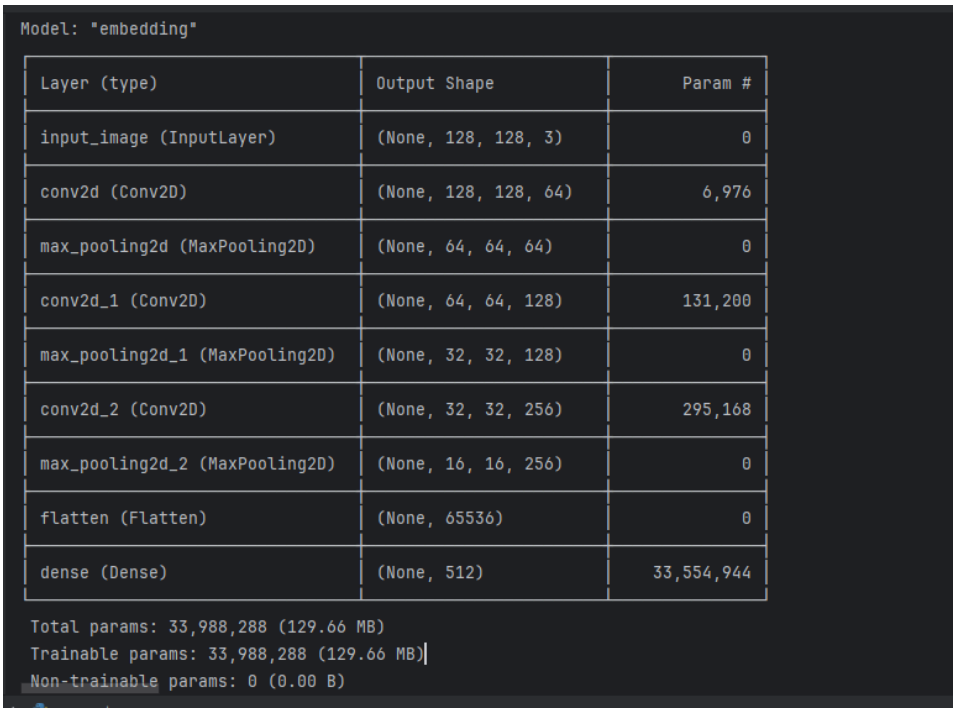
The ROI is extracted according to the aligned valley points with the maximum hand area included for palmprint recognition. Fig. 4.11 shows the extracted ROI of the palm.

Fig. 4.11: ROI image

4.2 Training model

Before training the model, the images are processed using the Laplacian of Gaussian (LOG) filter to enhance the images and reduce noise. After that, the feature extraction process is carried out using a Gabor filter bank which is used to capture texture information from the palmprint images.

Then, the model is trained using several parameters after a convolutional neural network model with its layers is built. The model is assembled using the Adam optimizer, which is preferred for deep learning models, which is considered to be the first parameter. It updates the model parameters based on the gradient of the loss function, which is the one that measures the difference between the expected probabilities and the real labels during the training process, and this is the second parameter. The model in this code was trained 100 times, which indicates the number of Epochs that was previously determined.



Model: "embedding"

Layer (type)	Output Shape	Param #
input_image (InputLayer)	(None, 128, 128, 3)	0
conv2d (Conv2D)	(None, 128, 128, 64)	6,976
max_pooling2d (MaxPooling2D)	(None, 64, 64, 64)	0
conv2d_1 (Conv2D)	(None, 64, 64, 128)	131,200
max_pooling2d_1 (MaxPooling2D)	(None, 32, 32, 128)	0
conv2d_2 (Conv2D)	(None, 32, 32, 256)	295,168
max_pooling2d_2 (MaxPooling2D)	(None, 16, 16, 256)	0
flatten (Flatten)	(None, 65536)	0
dense (Dense)	(None, 512)	33,554,944

Total params: 33,988,288 (129.66 MB)
Trainable params: 33,988,288 (129.66 MB)
Non-trainable params: 0 (0.00 B)

Fig. 4.12: CNN model summary

Fig. 4.12 shows a CNN model summary which represents details about the CNN layers including the output shape and number of parameters of each layer type. This shows the data flow from the input layer to the dense layer through this model.

The number of parameters for convolutional layers is calculated using equation 4.1:

$$\boldsymbol{Parameters} = (\boldsymbol{kw} \times \boldsymbol{kh} \times \boldsymbol{Cin} + \mathbf{1}) \times \boldsymbol{Cout} \qquad \textbf{(4.1)}$$

Where:

- kw = kernel width
- kh = kernel height
- Cin = number of input channels
- $Cout$ = number of filters (output channels)
- The "+ 1" accounts for the bias term for each filter [38].

And the number of parameters for the Dense layer is calculated using equation 4.2:

$$\boldsymbol{Parameters} = (\boldsymbol{input\ units} \times \boldsymbol{output\ units}) + \boldsymbol{output\ units} \qquad \textbf{(4.2)}$$

Where:

- Input units= the size of the flattened input
- output units= number of neurons in the dense layer

As for the input, max pooling, and flattened layers, they do not have parameters to learn.

The performed layers are:

- The input layer takes a three-channel (RGB) image of 128x128 size.
- Three convolutional blocks extract the features from the input images. Each block consists of a convolutional layer followed by a pooling layer.

Each convolutional layer applies a different number of filters of different sizes with a ReLU activation function according to the layer number. The size of the input and the output images of these layers are still the same, and the number of parameters depends on the number of filters, the size of them, and the number of channels.

Each max pooling layer performs a 2x2 pool which reduces the size of the image by a factor of 2.

- Flatten layer flattens the 3D tensor into a 1D tensor of size 65536
- Dense layer is a fully connected layer with a sigmoid activation function and 512 units and 33554944 parameters [39].

The total number of trainable parameters is 33988299 which includes biases that will be learned through training which indicates a complex and large model.

The trained model is then evaluated based on testing data using several metrics such as Accuracy, Recall, Precision, and F1Score. The values of this matrix are shown in Fig. 4.13. In general, these values indicate that the model works well on the test data.

```
Test Accuracy: 74.75%
10/10 ━━━━━━━━━━━━━━━━━━━━ 1s 51ms/step
Precision: 0.7609
Recall: 0.7475
F1 Score: 0.7228
```

Fig. 4.13: Performance measures

4.3 Matching Features

Fig. 4.14 provides a visual representation of matching feature points between two images. Two images, one from the training set and one from the test set, are linked to the background image. The characteristic points of the training image are represented by red dots, while the equivalent of the test image is represented by blue dots. On the other hand, the green lines connect the same distinct points. Each line connects two distinct points, one from each image, where these points correspond to their symmetry in their similarity.

The ORB algorithm was used to extract distinct points from both images and to match these points, a brute-force matcher with Hamming distance was used.

Fig. 4.14 in general shows a great match and similarity between the two images, and this was clear from the many green lines. The number of green lines can give a quick visual assessment of matching quality.

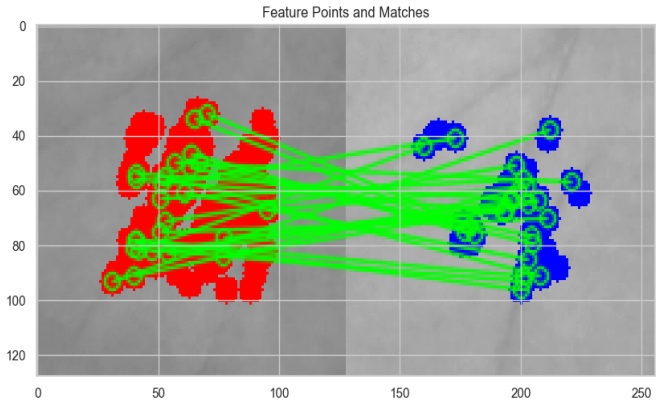


Fig. 4.14: Feature Points and Matches

Fig 4.15 contains two graphs, one of which is titled Truth, where the x-axis is titled Epochs, which indicates the number of repetitions of training.

It is very clear from the drawing that as the number of Epochs increases, the accuracy of training and verification increases.

The other drawing is titled Loss, which is the same title as the Y-axis, while the X-axis was titled with the same title as the previous drawing.

The training and validation loss decreases as the number of epochs increases, indicating that the model is learning and its predictions are becoming more accurate

The charts in Fig. 4.15 are essential for monitoring the training process of machine learning models,

as they provide insights into how well the model is learning. The main goal is to achieve high accuracy and low validation loss.

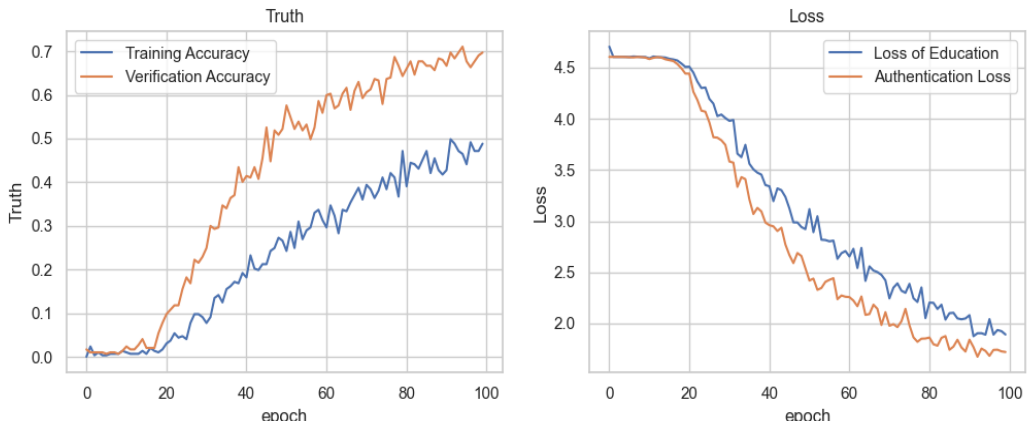


Fig. 4.15: Training Performance Metrics: Accuracy and Loss over Epochs

Fig. 4.16 shows a comparison between two image processing techniques, namely Otsu thresholding and K-means segmentation, which are widely used in the image processing process. These two techniques were used to demonstrate their effectiveness and the differences between them in dividing images into distinct regions.

Otsu Threshold Techniques: This technique, in its most basic form, generates a single intensity threshold that divides pixels into two classes: background and foreground. Minimizing the variance in intra-class intensity, or equivalently, maximizing the variance between classes, determines this threshold [40].

The equation of this technique:

$$\boldsymbol{\sigma_B^2(t) = \omega_0(t)\omega_1(t)[\mu_0(t) - \mu_1(t)]^2} \quad \textbf{4.3)}$$

where t: Threshold value, ω_0, ω_1: Weights of the two classes (foreground and background) and μ_0, μ_1: Means of the two classes, and these are the inputs. And $\sigma_B^2(t)$ is the output which is the variance for the threshold t between classes [40].

K means Segmentation Technique: One of the most popular techniques for accurately classifying an image's pixels in a decision-oriented application is image segmentation. To achieve high contrast and high similarity between regions, it separates an image into several discrete regions. This technique uses a K-means clustering algorithm to segment the interest area from the background [41].

These techniques were applied to two images, one from the training set and the other from the testing

set. The benefit of using Otsu thresholding on the training image was demonstrated in converting the grayscale image into a binary image while separating the background from the foreground. As for K-means segmentation, it is useful in dividing the image into regions effectively based on the intensity of the pixels in the image when applied to the same training image.

As for applying the two techniques to the testing image, it was for the purpose of seeing their effect on a different image, and this provided a visual understanding of how well the techniques generalized to new images.

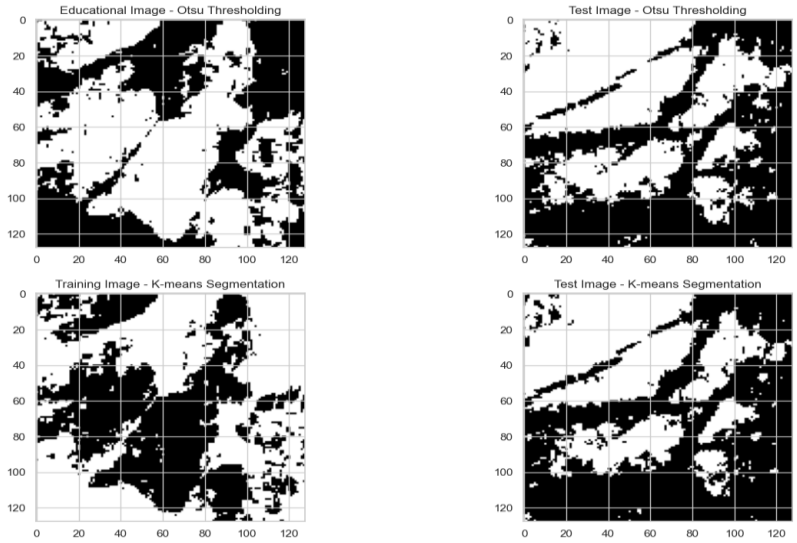


Fig. 4.16: Comparison of Otsu Thresholding and K-means Segmentation on Training and Testing Images

Chapter 5

Conclusion, Challenges, and Future Work

5.1 Conclusion

This research aims to create a device based on the palm print for identification. Through our systematic review of previous papers that dealt with the same topic in the first part of the study, we reached a clear vision regarding the techniques and algorithms that we will work on and the components that we will use to build the system.

We used some datasets to examine some algorithms and techniques, as one of the most important results was the accuracy, which was 74.75% after a model was trained for the examination using two sets of ROI images, one for training and the other for testing. As for the components that we will use, we have purchased them and gotten to know them closely, so the next stage will be to start creating the device.

In general, there are wide possibilities for applying this system in all places interested in identifying identities, as it is of great importance, has high accuracy, and provides unparalleled credibility. Based on many expert studies, it is an idea for the coming future.

5.2 Challenges

We faced many problems and challenges in the past period, during which our goal was to gain greater insight into the project and begin implementing it programmatically and practically using hardware components. Most of these problems focused on the issue of the system's camera, as it was supposed to be a multispectral camera or a camera that visualizes in the infrared range to show the veins in the hand since our project depends on both the palmprint and its veins. The first problem was the relatively high price of the camera, while we were trying to build a low-cost system compared to other systems. In addition to the high cost of the camera, due to the restrictions of the occupation, the possibility of the camera being rejected by them was high, even if the store containing the intended camera was

delivered to our occupied territories. The biggest problem was the war in the Gaza Strip, which affected the increase in the delivery period for components, including infrared LEDs, from the moment of ordering, which reached at least three weeks. These problems made us limit ourselves in principle to a palmprint scanning system to verify identity using a camera extracted from a home surveillance camera after losing hope of obtaining a multispectral camera, which we faced a problem in operating on the Raspberry Pi.

5.3 Future work

There will be major developments in the contactless palmprint and palm vein detection authentication system that will enhance both the software side and the hardware side. Concentrating on refining the software in the first period with the help of Python programming language, OpenCV, and TensorFlow for enhancing the algorithms of image processing, as well as constructing a multi-layered system for authentication. The actual hardware will be developed in the second period, creating a small device, a comfortable gadget with only an 8 GB micro-SD card to store data locally as well as an IR LED strip incorporated to capture the specifics of the palm vein pattern. Regarding the device operation, it will be based on Raspberry Pi and Raspberry Pi NoIR high-resolution camera. Finally, there is a demonstration of the system and the comparison of results with the data from the existing literature and trying to create our dataset. Our focus after that will be on enhancing this device to detect the palm vein in addition to the palmprint. This structured plan is thus meant to holistically improve the state of our authentication system using contactless palmprint and palm vein detection.

References

[1] Y. Yang, Y. Zhou, R. Huang, Q. Liu, H. He, and X. Li, "Contactless Palmprint and Palm Vein Identity Recognition via a Bimodal Network with Parameter-adaptive Log-Gabor Convolution," 2023. [Online]. Available: https://ssrn.com/abstract=4625638

[2] Rohit Khokher, Rajendra Kumar, and R C Singh, "Asystematic review of palm and dorsal hand vein recognition techniques," Dec. 2021, doi: 10.3306/AJHS.2022.37.01.100.

[3] D. Zhang, W.-K. Kong, J. You, and M. Wong, "Online Palmprint Identification," Sep. 2003. doi: 10.1109/TPAMI.2003.1227981.

[4] Y. Han, Z. Sun, F. Wang, and T. Tan, "Palmprint recognition under unconstrained scenes," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Springer Verlag, 2007, pp. 1–11. doi: 10.1007/978-3-540-76390-1_1.

[5] D. Hartung and C. Busch, "Why Vein Recognition Needs Privacy Protection," Dec. 2012.

[6] G. Shah, S. Shirke, S. Sawant, and Y. H. Dandawate, "Palm vein pattern-based biometric recognition system," *International Journal of Computer Applications in Technology*, vol. 51, no. 2, pp. 105–111, 2015, doi: 10.1504/IJCAT.2015.068921.

[7] L. Mirmohamadsadeghi and A. Drygajlo, "Palm vein recognition with Local Binary Patterns and Local Derivative Patterns," *2011 International Joint Conference on Biometrics, IJCB 2011*, Oct. 2011, doi: 10.1109/IJCB.2011.6117804.

[8] J.-C. Lee, "A novel biometric system based on palm vein image," *Pattern Recognit Lett*, vol. 33, pp. 1520–1528, Sep. 2012, doi: 10.1016/j.patrec.2012.04.007.

[9] W. Kang, Y. Liu, Q. Wu, and X. Yue, "Contact-Free Palm-Vein Recognition Based on Local Invariant Features," *PLoS One*, vol. 9, no. 5, pp. e97548-, May 2014, [Online]. Available: https://doi.org/10.1371/journal.pone.0097548

[10] S. Elnasir and S. M. Shamsuddin, "Proposed scheme for palm vein recognition based on Linear Discrimination Analysis and nearest neighbour classifier," in *2014 International Symposium on Biometrics and Security Technologies (ISBAST)*, 2014, pp. 67–72. doi: 10.1109/ISBAST.2014.7013096.

[11] W. Lu, M. Li, and L. Zhang, "Palm Vein Recognition Using Directional Features Derived from Local Binary Patterns," *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 9, no. 5, pp. 87–98, May 2016, doi: 10.14257/ijsip.2016.9.5.09.

[12] S. Cho and K. A. Toh, *Palm-Vein Recognition Using RGB Images*. 2018. doi: 10.1145/3278229.3278239.

[13] R. Hernández-García, R. J. Barrientos, C. Rojas, and M. Mora, “Individuals identification based on palm vein matching under a parallel environment,” *Applied Sciences (Switzerland)*, vol. 9, no. 14, Jul. 2019, doi: 10.3390/app9142805.

[14] Y. Pititheeraphab, N. Thongpance, H. Aoyama, and C. Pintavirooj, “Vein pattern verification and identification based on local geometric invariants constructed from minutia points and augmented with barcoded local feature,” *Applied Sciences (Switzerland)*, vol. 10, no. 9, May 2020, doi: 10.3390/app10093192.

[15] W. Wu, Y. Li, Y. Zhang, and C. Li, “Identity Recognition System Based on Multi-Spectral Palm Vein Image,” *Electronics (Switzerland)*, vol. 12, no. 16, Aug. 2023, doi: 10.3390/electronics12163503.

[16] Dr. Zhenan Sun, “Note on CASIA Multi-Spectral Palmprint Database.” Accessed: May 10, 2024. [Online]. Available: http://biometrics.idealtest.org/

[17] “IIT Delhi Touchless Palmprint Database,” Oct. 2007, Accessed: May 11, 2024. [Online]. Available: https://www4.comp.polyu.edu.hk/~csajaykr/IITD/Database_Palm.htm

[18] H. Ghael, H. Dipak Ghael, L. Solanki, G. Sahu, and A. Professor, “A Review Paper on Raspberry Pi and its Applications,” *International Journal of Advances in Engineering and Management (IJAEM*, vol. 2, p. 225, 2008, doi: 10.35629/5252-0212225227.

[19] M. Fezari, A. Al Dahoud, M. Fezari, and A. Al-Dahoud, “Raspberry Pi 5 : The new Raspberry Pi family with more computation power and AI integration,” Nov. 2023, doi: 10.13140/RG.2.2.13547.52009.

[20] “Raspberry Pi 5 features improved image processing - Geeky Gadgets,” Oct. 2023, Accessed: Jun. 10, 2024. [Online]. Available: https://www.geeky-gadgets.com/raspberry-pi-5-features-improved-image-processing/

[21] “Raspberry Pi NoIR camera marker tracking _ DreamOnward,” Oct. 2019, Accessed: Jun. 10, 2024. [Online]. Available: https://dreamonward.com/2019/10/16/picamera-exploration/

[22] “Raspberry Pi NoIR Camera V2 8MP - Bastelgarage Electronics Online Store”, Accessed: Jun. 10, 2024. [Online]. Available: https://www.bastelgarage.ch/raspberry-pi-noir-camera-v2-8mp

[23] C. L. Lin and K. C. Fan, “Biometric verification using thermal images of palm-dorsa vein patterns,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 2, pp. 199–213, Feb. 2004, doi: 10.1109/TCSVT.2003.821975.

[24] “Infrared Thermography - ZfP - BayernCollab,” Jun. 2015, Accessed: Jun. 30, 2024. [Online]. Available: https://collab.dvb.bayern/display/TUMzfp/Infrared+Thermography#:~:text=The%20Stefan%2DBoltzmann%20law%20gives,emission%20capacity%20of%20an%20object

[25] “Micro SD Cards Market Size, Trends and Forecast 2031 with Top Countries Data _ LinkedIn”, Accessed: Jun. 27, 2024. [Online]. Available: https://www.linkedin.com/pulse/micro-sd-cards-market-size-trends-forecast-2031-top-countries-2xxbe/

[26] "New Diagnostic Forensic Protocol for Damaged Secure Digital Memory Cards _ IEEE Journals & Magazine _ IEEE Xplore", Accessed: Jun. 27, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/9733364

[27] "Infrared Radiation _ Definition, Uses & Examples - Lesson _ Study.com," Nov. 2023, Accessed: Jun. 28, 2024. [Online]. Available: https://study.com/academy/lesson/infrared-radiation-definition-uses-effects.html

[28] "A Hands-On Guide to Document Image Classification _ by Babina Banjara _ Artificial Intelligence in Plain English," Nov. 2023, Accessed: Jun. 28, 2024. [Online]. Available: https://ai.plainenglish.io/a-hands-on-guide-to-document-image-classification-daed233ebeed

[29] "What is the mathematical formula of the convolution operation on a 2D image_ - EITCA Academy," May 2024, Accessed: Jun. 30, 2024. [Online]. Available: https://eitca.org/artificial-intelligence/eitc-ai-adl-advanced-deep-learning/advanced-computer-vision/convolutional-neural-networks-for-image-recognition/what-is-the-mathematical-formula-of-the-convolution-operation-on-a-2d-image/

[30] "Convolutional Layers vs Fully Connected Layers _ by Diego Unzueta _ Towards Data Science," Dec. 2018, Accessed: Jun. 29, 2024. [Online]. Available: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

[31] G. Yang, Q. Zhang, and G. Zhang, "EANet: Edge-aware network for the extraction of buildings from aerial images," *Remote Sens (Basel)*, vol. 12, no. 13, Jul. 2020, doi: 10.3390/rs12132161.

[32] "CNN _ Introduction to Pooling Layer - GeeksforGeeks," Apr. 2023, Accessed: Jun. 30, 2024. [Online]. Available: https://www.geeksforgeeks.org/cnn-introduction-to-pooling-layer/

[33] "Pooling-Layer (Albawi et al., 2017) _ Download Scientific Diagram", Accessed: Jun. 29, 2024. [Online]. Available: https://www.researchgate.net/figure/Figure-3-Pooling-Layer-Albawi-et-al-2017_fig1_349054980

[34] Y. Jia *et al.*, "Caffe: Convolutional Architecture for Fast Feature Embedding," Jun. 2014, [Online]. Available: http://arxiv.org/abs/1408.5093

[35] "Chapter 2_ Convolutional Neural Networks — Unveiling Hidden Patterns in Images _ by Issac kondreddy _ Medium," May 2023, Accessed: Jun. 30, 2024. [Online]. Available: https://medium.com/@issackondreddy/chapter-2-convolutional-neural-networks-unveiling-hidden-patterns-in-images-b4574d34f556

[36] Rinki Nag, "A Comprehensive Guide to Siamese Neural Networks _ by Rinki Nag _ Medium," Nov. 2019, Accessed: Jun. 30, 2024. [Online]. Available: https://medium.com/@rinkinag24/a-comprehensive-guide-to-siamese-neural-networks-3358658c0513

[37] "K-Means Clustering Algorithm - Javatpoint", Accessed: Jun. 28, 2024. [Online]. Available: https://www.javatpoint.com/k-means-clustering-algorithm-in-machine-learning

[38] “How to calculate the number of parameters in CNN_ - GeeksforGeeks,” May 2024, Accessed: Jun. 21, 2024. [Online]. Available: https://www.geeksforgeeks.org/how-to-calculate-the-number-of-parameters-in-cnn/

[39] Yagesh Verma, “What is Dense Layer in Neural Network_,” Jun. 2024, Accessed: Jun. 30, 2024. [Online]. Available: https://analyticsindiamag.com/topics/what-is-dense-layer-in-neural-network/

[40] “Otsu’s method - Wikipedia,” Apr. 2024, Accessed: Jun. 25, 2024. [Online]. Available: https://en.wikipedia.org/wiki/Otsu%27s_method

[41] N. Dhanachandra, K. Manglem, and Y. J. Chanu, “Image Segmentation Using K-means Clustering Algorithm and Subtractive Clustering Algorithm,” in *Procedia Computer Science*, Elsevier, 2015, pp. 764–771. doi: 10.1016/j.procs.2015.06.090.