Dorsal Hand Veins Based Biometric Identification System Using Deep Learning

Masoumeh ZEHTAB NAYEBİ[1], Zeynep TURGUT[1*]

[1]Computer Engineering Department, Halic University, Istanbul, Turkey
ORCID ID: 0000-0002-0288-6461, 0000-0002-9416-609X

Abstract

Identification systems have become biometric based, especially with the increase in the performance rates of machine learning methods. Biometric identification systems offer a high level of security by using reliable, difficult-to-change parameters. In this study, a biometric identification system is proposed using dorsal hand vein patterns. The relevant system has been tested on the sample dataset in the literature. The number of data were increased by adding noisy data to the data set used. Classification was made on the preprocessed images using SVM, ANN, LDA + KNN, and CNN methods. It has been determined that the highest identification accuracy is achieved when CNN is used, and CNN method provides higher performance compared to other methods. With the proposed identification system, after multiple runs, an average accuracy of 99.64% is achieved with the CNN machine learning method.

Keywords: deep learning, electronics, biometric identification, hand veins, electronic system.

Derin Öğrenme Kullanılarak El Sırtı Damar İzlerine Dayalı Biyometrik Tanımlama Sistemi

Öz

Kimlik tanımlama sistemleri özellikle makine öğrenme yöntemlerinin başarım oranlarının artması ile biyometrik tabanlı hale gelmiştir. Biyometrik kimlik tanımlama sistemleri güvenilir, değiştirilmesi güç parametrelerin kullanılması ile yüksek bir güvenlik seviyesi sunarlar. Bu çalışmada el sırtı damar izi kullanılarak bir biyometrik kimlik tespit sistemi önerilmiştir. Literatürde yer alan örnek veri kümesi üzerinde ilgili sistem test edilmiştir. Kullanılan veri kümesi içerisine gürültülü veriler eklenerek veri sayısı arttırılmıştır. Ön işleme tabi tutulan görüntüler üzerinde SVM, ANN, LDA+KNN, ve CNN yöntemleri kullanılarak sınıflandırma işlemi yapılmıştır. En yüksek kimlik tespit doğruluğuna CNN kullanıldığında erişildiği, CNN yönteminin diğer yöntemlere kıyasla daha yüksek performans sağladığı tespit edilmiştir. Önerilen tanımlama sistemin tekrarlı çalıştırmaları sonucunda CNN makine öğrenmesi yöntemi ile %99.64'lük başarı oranına ulaşılmıştır.

Anahtar Kelimeler: derin öğrenme, elektronik, biyometrik tanımlama, el damarları, elektronik system..

1. Introduction

Like fingerprints, the shape of each individual's veins is a unique feature. The dorsal hand vein pattern differ even in identical twins. Only the size of the veins changes over time. In the vein-based detection method, infrared light is applied to the palm or back of the hand. This helps to see the condition of the veins under the skin more

*Corresponding Author:zeynepturgut@halic.edu.tr

clearly.. One of the advantages of the vein-based detection technique is that it does not require physical contact and its accuracy is reasonably high. This method is particularly useful for business, spatial identification. This method, which is a direct and fast alternative to fingerprint, is cost effective, has high accuracy and is used for identification in various fields.

Ahmed M.Badawi in their studies "Hand Vein Biometric Verification Prototype: A Testing Performance and Patterns Similarity" proposed a system using patterns similarity (Badawi, 2006).

In the study Liveness Detection of Dorsal Hand Vein Based on the Analysis of Fourier Spectral by Fourier Spectral Analysis a system is proposed. First, by applying the Fourier Transform, a property is extracted as a statistical value of the spectral energy derived from each intercepted spectrum of single wavelength infrared images. In the classification stage, support vector machines one of the machine learning methods, were used as classifiers (Jain, 2007).

In the study Vein Recognition Across Different Devices the most important features are analyzed. After the grayscale images are normalized, a segmentation method based on gradient difference is used and uses SIFT to subtract and match features. The discrimination accuracy in this article is up to 90.17%. This method can effectively solve the problem of identification based on dorsal vein print in different devices (Wang, Zheng, & Wang, 2016)

In the study Dorsal hand-vein images recognition system based on grey level co-occurrence matrix and Tamura features vein images are analyzed and to design a system that has the ability to identify each person according to the vein by segmenting the veins. Experimental results showed that the MDC classifier achieved 92% accuracy when using wavelet transform and GLCM and Tamura features (Alasadi & Dawood, 2017)

In this study, a biometric-electronic identification system using hand dorsal vein print is proposed. This system consists of the approval and verification process using previously recorded dorsal hand veins.

In the process of collectiong dorsal hand vein data (Enrollment), the dataset collected by Badawi (Badawi, 2006) is used as the main example for different people. This data set includes 100 hand images (50 right hand images and 50 left hand images). In order to test the accuracy of the proposed method, noise is added to the images in the dataset and a total of 4000 image data is produced with an increased number of images. These images are processed during the pre-processing, feature extraction and modelling – classification stages to create customizable samples. It is aimed to increase the image quality and remove the dorsal hand vein part on the images with binary values in the pre-processing stage. The classification process is realized using machine learning methods on the dorsal hand vein skeleton images.

In this study, SVM, ANN, LDA + KNN and CNN techniques are used in order to classify the images in the dataset. The performances of the methods used are analyzed, an identification system using dorsal hand vein is proposed which is based on CNN, which offers the highest performance.

2. Material and Methods

In this study, the dataset which is known as "Badawi hand vein dataset" is used. There are 500 photographs of 100 hands in the relevant data set. 5 samples are taken for each person's hand sample. The age of the persons included in the dataset varies between 16-65, they have different occupational groups and do not have any health problems. Different images of the right and left hands are included in the dataset. Dr. The Badawi hand vein dataset is organized in a way that it can be used to identify the identity through dorsal hand veins, as well as to allow the analysis of the differences between the right and left hands. Data between 1-50 include left-hand images of individuals, while data between 51-100 include right-hand images of individuals. In this study, no distinction between right and left hand is made, and 100 different classes arecreated by assuming that the left and right hands reflect different people.

There are various techniques used to increase the images included in datasets in order to test the effectiveness of image processing approaches. The number of photographs, which is 500 in total, is increased to 4000 by adding noise at various rates to the data in the Badawi hand vein data set used in this study.

2.1. Pre-process Phase

Image processing is a technique to improve the quality of the image after extracting irrelevant data from the image in various areas. To remove the necessary parts of an image, some of the image preprocessing techniques are needed to better visualize the images before identification is performed. The process of increasing pixel density and image quality is also realized by pre-processing. Image quality increases after preprocessing.

Image pre-processing is often the first and most practical part of identification. The step-by-step image preprocessing includes the following steps:

- Subtracting the desired area and removing the background (Region of Interest)
- Enhancing image clarity
- Dividing the image into sections
- Skeleton extraction and processing operations

In order to remove the dorsal hand vein region in the hand image obtained, the photograph must first be processed. Adjacent pixels of equal intensity or close to each other are evaluated as zero when they are below a certain limit. This ensures that the target is subtracted from the background. It is important to enhance the image quality before doing any processing on the images. For this reason, a median filter is used. The advantage of this filter over other filters is that, in addition to reducing distortion, it can successfully detect the edges of the image. The median filter is based on the principle of sorting neighboring pixels and taking the value in the middle of the row. If there is an even number of pixels in the area under study (inside the template), the average of the two pixels in the middle is used as the middle value (Cross & Smith, 1995)

The median filter gives a much better result than the mean filter at the point of not losing the details on the picture. The median filter looks at its nearby neighbors to calculate the value of each pixel. In the median filter, instead of replacing the pixel value with the average of neighboring pixel values (mean filter), it sorts neighboring pixels and takes the value in the middle of the row. If the region under examination (inside the template) has

an even number of pixels, the average of the two pixels in the middle is used as the middle value. In the segmenting stage of the image, the adaptive thresholds method is used to extract the hand vein pattern. In this threshold type, threshold values change across the image as a function of the local image of the image. The method is shown in Equation 1.

$$I(i,j) = \begin{cases} 1 \\ 2 \end{cases} \quad I(i,j) \geq T_{ij} \qquad (1)$$

Mathematical morphology processes are used for image enhancement. There are two basic operations in mathematical morphology:

- Dilatation: Process expands the binary image and removes gaps in the image. In addition to expanding the image, it also removes broken edges.

- Erosion: The erosion process analyzes the image, actually narrows the image and removes extra parts.

There are also a number of combined processes available:

- Opening and Closing: These two processes are the most important processes in binary images. In the opening process, dilatation is first followed by erosion. Opening process eliminates the noise without narrowing the image and fills the gaps without enlarging the image. Closing proceeds in the reverse order of opening. In a closure procedure performed accordingly, first erosion is followed by dilatation.

The edge of the image is the point where the brightness in the image suddenly changes. Edge detection is one of the most important basic operations in image processing, so the efficiency of high-level operations such as object recognition and image segmentation directly depends on the efficiency of this low-level processing.

There are many algorithms that can be used in this field (depending on the type of image and the relevant process, each algorithm has its own advantages and disadvantages): Sobel , Canny, Zero Cross, Log, Prewitt, Roberts edge detection algorithms are some of them.

To skeleton a pattern from the edges means narrowing the pattern so that the overall shape of the pattern is not destroyed. Therefore, the skeleton of a pattern must be recognizable as the pattern itself and thus predict the overall shape of the pattern.

The resulting skeleton for a pattern should contain the following properties:

• It should be thin enough.

• It must be linked (ie the skeleton must not cause the skeleton to break).

2.2. Feature Extraction

Feature is the whole of the sub-components that make up the image. In order to create a high-performance biometric identification system, it is of great importance to remove these components in identification. Wavelet transform is used in this study to extract features.

The wavelet transform displays the image in the form of a series of wavelet functions with different positions and scales. Each image decomposition consists of a pair of waveforms, one representing the high frequencies associated with image details, the other representing low frequencies or smooth parts of the image. The result of the wavelet transformation is a wavelet expression that

measures the contribution of the wavelet to these locations and scales. The wavelet transform divides the image into images of different resolutions(Misiti, Misiti, Oppenheim, & Poggi, 2007).

In this study, the wavelet transform is used to derive the four coefficients of the coefficients in the horizontal, vertical and vertical directions and the product of these coefficients, such that each image is derived from a two-dimensional wavelet transform. The image is reproduced in these directions.

When it is compared with Fourier transform, Wavelet transform has some differences.(Benhar, Idri, & Fernández-Alemán, 2020). The main difference between them is that the Fourier transform decomposes the signal into sines and cosines. In other words, unlike functions localized in Fourier space, wavelet transform uses functions localized both in real and Fourier space Generally, the wavelet transform is expressed by Equation 2.

$$F(a, b) = \int_{-\infty}^{+\infty} f(x)\psi_{(a,b)}^{*}(x)dx \quad (2)$$

In information theory, entropy (or to be more precise, Shannon entropy) is the mathematical average of the information contained in each message - the data. Here, the data does not necessarily consist of text and can have any information flow. The entropy of each message is the amount of uncertainty. Using the equation in Equation 3, it is possible to calculate the entropy of H (x). Pi value is the probability of occurrence of a given symbol. The symbols here are pixels.

$$H(x) = -\sum_{i=1}^{n} P_i log_2 P_i \quad (3)$$

Variance is a measure of distribution. Compared to the mean, the mean is said to represent the place of distribution, while the variance is the measure of how the data is distributed over the mean. Less variance means that if a sample of the distribution is selected, it is expected to be near the mean. The variance of a square is one of its initial quantity. Variance shows how a data series is distributed around the mean value. Equation 4 shows how variance is calculated (where $\bar{x}$, is the arithmetic mean of x_i data).

$$H(x) = \frac{1}{n}[\sum_{i=1}^{n}(x_i - \bar{x})^2] \quad (4)$$

2.3 Classification

In this study, classification methods and feature reduction techniques, Support Vector Machine (SVM), Linear Discriminant Analysis (LDA) + K Nearest Neighbor (KNN), Artificial Neural Network - ANN) and the deep learning approach Convolutional Neural Network (CNN) are used.

Support Vector Machine (SVM) is one of the most common algorithms and methods in data classification. SVMs a set of points in n-dimensional space of data that can result in classification output that marks the boundaries of sets, delimits and classifies the data, moving one of them. In fact, a Support Vector Machine is a cluster or boundary that defines the best classification and separation of data by comparing vectors. It is not sensitive to other data points and aims to find the best boundary between data as close as possible to all clusters (support vectors) (Patel & Mishra, 2015). These algorithms have inherent limitations, such as how parameters can be specified for a mapping function. Decision support machines require complex and time-consuming computations and consume a lot of memory due to their complexity. Discrete and non-numeric data are also not compatible with this method and must be converted. (Vijayanand, Devaraj, & Kannapiran, 2018)

However, SVMs have a consistent theoretical foundation and the solutions they produce are global and unique. SVMs are the most common forecasting techniques in data mining today.

Linear Discriminant Analysis (LDA) is a technique based on size reduction. Used as a pre-processing step for machine learning and pattern classification applications. The purpose of LDA is to reflect the features from the higher dimensional space to a lower dimensional area in order to avoid the problems posed by the size problem and to reduce resources and dimensional costs. The main purpose is to reflect the data set to a lower dimensional area and remove unnecessary and dependent features. In simple terms, they reduce the dimensions (i.e. variables) in a given data et while preserving most of the data. Multidimensional data contain multiple features that correlate with each other. With dimensional reduction, multidimensional data can only be drawn in 2 or 3 dimensions (Turgut, Üstebay, Zeynep Gürkaş Aydın, & Sertbaş, 2019).

This approach ignores the useful information provided by the second feature. However, LDA can be used to reflect this. LDA uses information from the two properties to minimize variance and maximize the class distance of the two variables (Charte, Charte, García, del Jesus, & Herrera, 2018) .

LDA primarily focuses on projecting features in higher dimensional space to lower dimensions. This process can be summarized in three steps:

First, it is necessary to calculate the separability between classes, which is the distance between the mean of different classes. The related calculation is shown in Equation 5.

$$S_b = \sum_{i=1}^{g} N_i (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T \quad (5)$$

Second, the distance between the mean and sample of each class must be calculated. The related calculation is shown in Equation 6.

$$S_w = \sum_{i=1}^{g} (N_i - 1)S_i$$
$$= \sum_{i=1}^{g} \sum_{j=1}^{N_i} (x_{i,j} - \bar{x}_i)(x_{i,j} - \bar{x}_i)^T \quad (6)$$

Finally, a sub-dimension field that maximizes inter-class variance and minimizes intra-class variance should be created. P is considered to be low dimensional space projection, also called Fisher criterion. The related formula is presented in Equation 7.

$$P_{lda} = argmax_p \frac{|P^T S_b P|}{|P^T S_W P|} \quad (7)$$

The K-nearest neighbors (KNN) algorithm is a simple, easy-to-implement controlled machine learning algorithm that can be used to solve both classification and regression problems . The KNN algorithm assumes similar data is nearby. In other words, similar data are close to each other(Bagosi & Baruch, 2011). The performance of the KNN algorithm depends on the accuracy of this assumption. KNN captures the idea of similarity by calculating the distance between points on a graph. KNN distance function which is used in this study Euclidean distance formula is shown in Equation 8 .

$$\sqrt{\sum_{i=1}^{k} (x_i - y_i)^2} \quad (8)$$

The artificial neural network consists of three layers: input, output and hidden layer. Each layer contains a group of neurons that are generally associated with all other layer neurons, unless the user does not restrict communication between neurons, but the neurons of each layer do not interact with other neurons of the same layer. The input equation is presented in Equation 9.

$$X_i^T = \begin{bmatrix} x_{i1} \ldots \ldots \ldots x_{ip} & 1 \end{bmatrix} \qquad (9)$$

X is the input vector and x is the network input variable. The value of number 1 in the relation in Equation 4.7 is called the bias value. The primary role of bias is to provide a constant educational value for each neuron. In fact, the deviation value allows to change the activation function or the transfer function. Related equation is presented in Equation 10.

$$W_{ij}^T = \begin{bmatrix} w_{j1} \ldots \ldots \ldots w_{jp} & -\theta \end{bmatrix} \qquad (10)$$

W is the weight vector and w is the weight of each variable of the input vector. Activation occurs due to the sum of input weights and the threshold value of each neuron. The activated signal is extracted from the neuron due to a transfer function which is represented in Equation 11.

$$NET_i(x_i) = X_i^T W^T \quad \rightarrow$$

$$output_j = O_j = sf(NET_j) = \frac{1}{1-e^{-\beta(NET_j)}}$$

$$(11)$$

Deep learning algorithms are a subset of machine learning algorithms aimed at discovering a large number of distributed representations of input data.

Convolutional neural networks (CNN) are one of the most important deep learning methods in which multiple layers are trained strongly. This method is very effective and especially one of the most common methods used on image. Generally, a CNN network consists of three main layers: convolution layer, pooling layer, and fully connected layer (Li, Feng, & Kuo, 2018). Different layers perform different tasks. There are two phases of training in each neural network: forward and backward propagation. In the first step after propagation, calculations are made by feeding the input image to the network. Here, the output result is used to calculate the network error rate to adjust network parameters, for network training. To do this, it compares the network output with a loss function that gives the correct answer and the error rate is calculated (TTP et al., 2017). The next step starts with the back propagation phase based on the calculated error rate. At this point, the gradient of each parameter is calculated according to the chain rule and all parameters are changed according to their effect on the fault in the network (Varastehpour, 2020). After the parameters are updated, the next feed forward phase begins. Network training will end after repeating these steps

In convolution layer filters are used for producing a 2 dimensional activation map. By using related maps a deeper structure is presented. In Equation 4 m shows m core width and height, h convolution output, x input, w convolution core. (Ganegedara, n.d.)

$$h_{i,j} = \sum_{k=1}^{m} \sum_{l=1}^{m} w_{k,l} x_{i+k-1, j+l-1} \qquad (12)$$

Convolution layers can extract local features as they are locally constrained. Convolutional neural networks are the core of deep learning that uses a wide range of fields such as pattern recognition, speech recognition, natural language processing, and video analysis.

In pooling layer image features are reduced by changing the spatial size, i.e. depth, of the input volume. This change is realized reducing the amount of parameters or weights. Its mathematical formula is presented in Equation 13 (Ganegedara, n.d.)

$$h_{i,j} = \{x_{i+k-1, j+l-1}$$

$$\forall 1 \leq k \leq m \ and \ 1 \leq l \leq m \qquad (13)$$

In this study of 3x3 convolution layers and 8x8, 16x16, 32x32 filters are used respectively. The ReLu (Rectified Linear Unit) activation function is used after each convolution layer. ReLu layer changes the

negative values in input data to zero. It is presented mathematically in Equation 14.The maximum pooling layer is 2x2 in size and the range is 2.

$$f(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases} \quad (14)$$

2.4 Performance Metrics

The performance analysis of a classifier is calculated by using values obtained from confusion matrix (Veropoulos, Campbell, Cristianini, & Others, 1999).

• True Positive (TP): The situation in which the person being sought is approved correctly based on the dorsal hand vein images.

• True Negative (TN): The situation where the person being sought is not approved correctly based on the images of the dorsal hand vein images.

• False Positive (FP): The situation where the person sought is approved as wrong based on the images of the dorsal hand vein images.

• False Negative (FN): The situation in which the person sought is not approved as wrong based on the images of dorsal hand vein images.

Based on the confusion matrix, accuracy, error, specificity, sensitivity, F1-score, Matthews correlation coefficient, kappa performance metrics can be calculated. Accuracy calculation is presented in Equation 15. Error is presented in Equation 16. Specificity is presented in Equation 17. Sensitivity is presented in Equation 18, Sensitivity is presented in Equation 19, F1-Score is presented in Equation 20, Matthews Correlation Coefficient is presented in Equation 21 and Kappa is presented Equation in 22.

$$\frac{TP+TN}{TP+TN+FP+FN} \quad (15)$$

$$1 - \frac{TP+TN}{TP+TN+FP+FN} \quad (16)$$

$$\frac{TN}{TN+FP} \quad (17)$$

$$\frac{TP}{TP+FN} \quad (18)$$

$$\frac{TP}{TP+FP} \quad (19)$$

$$\frac{2(Precision \times Sensitivity)}{(Precision+Sensitivity)} \quad (20)$$

$$\frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TN+FN)(TP+FN)(TN+FP)}} \quad (21)$$

$$\frac{\frac{(TP+TN)}{TP+TN+FP+FN} - \left[\frac{(TN+FP)(TN+FN)+(FN+TP)(FP+TP)}{(TP+TN+FP+FN)^2}\right]}{1 - \left[\frac{(TN+FP)(TN+FN)+(FN+TP)(FP+TP)}{(TP+TN+FP+FN)^2}\right]} \quad (22)$$

3. Research Findings

In this study, a biometric identification system based on dorsal hand vein images is developed using a sample data set. Badawi dorsal hand vein dataset as sample data set. A total of 4000 data was obtained by adding noise at different rates to 100 data in the data set. Salt and pepper noise ranging from 0.01-0.07 are added to the data. An equal number of additions are made to the data from each noise ratio. Figure 1 shows the sample dorsal hand vein image in the data set (Shahin, Badawi, & Kamel, 2008).

Fig.1 Original image

In Figure 2, 0.01 (noise intensity) noise is added to the black and white dorsal hand vein image. In Figure 3, 0.05 (noise intensity) noise is added to the black and white dorsal hand vein image.

In the second step, extracting the desired area and removing the background is done. For this purpose, a zero value has been assigned to pixel values that have the same density or are
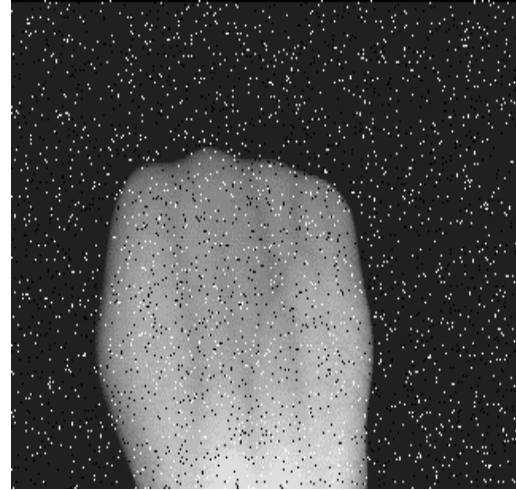

Fig 3. Image with 0.05 added noise

lower than the limit close to each other. Extended Maxima Transform is used for related operations. After this stage, the range of hands will be completely extracted from the background.

After removing the desired background, it is necessary to increase the clarity of the image first in order to process the image with high performance. For this reason, a 3×3 size median filter has been used to remove image distortion. The image obtained after the use of the median filter is given in Figure 4.

The hand vein pattern is removed on the sharpened image. In this step, adaptive

Fig 2. Image with 0.01 added noise

Fig 4. Image obtained after using median filter

thresholding procedure has been implemented to extract the dorsal hand vein pattern. In this type of thresholding, threshold values change across the image as a derivative of the local image of the image. The density of the gray image changes completely throughout the image according to the overall thumbnail and layout. The threshold value for the whole image cannot be obtained satisfactorily with the vein pattern.

Therefore, the built-in detection algorithm is followed. As explained earlier, this algorithm calculates threshold values for each pixel based on the analysis of pixels adjacent to the center pixel. Figure 5.a, extracting vein pattern from the image is shown.

After the adaptive threshold process is performed and the arrangement of the veins is removed, since the image has broken and separable points within the arteries, the closing algorithm is used together with the structural element of the disk in 3 radius for the integrity of the image. The result obtained after applying the closing operation is given in Figure 5.b.

Fig.5 Dorsal hand image in pre-process phase a) extracting the vein pattern from the image b) result obtained after applying the closing operation

In the following stages, the area of the hand circumference and the vascular trace are completely removed. In order to maintain the maximum appearance of the vein pattern, it is avoided to obtain the treated area in geometric shapes such as squares or rectangles. It is preferred that the image of all veins be fully enhanced in order to process and maximize distinctive features. Then, an edge detection algorithm is used to extract the hand dial in order to process the image.

In this study, Canny edge detection algorithm is used. When compared to other edge detection algorithms, for example, unlike the Sobel algorithm, it provides continuity in the vein image obtained and avoids creating break points in the image. Figure 6.a shows the image obtained after using the Canny edge detection algorithm.

A value should be assigned to the value between two edges for each vein on the image with edges. Thus, these fields can be used meaningfully in the classification stage. For this purpose, Canny edge detection algorithm is used. And the pixel values of the image is changed. Image obtained after using Canny edge detection algorithm is presented in Figure 6. Figure 6.a presents result obtained right after Canny edge detection algorithm is used and Figure 6. b presents image obtained after 0 pixel values transformed to 1 for remove ng vein area. Thus, it is ensured that the background and only the area belonging to the vein regions have a white color.

Fig.6 Image obtained after using Canny edge detection algorithm a) result obtained right after Canny edge detection algorithm is used b)image obtained after 0 pixel values transformed to 1 for removing vein area

In order to fully remove the skeleton, it is necessary to differentiate the hand skeleton and background colors. For this purpose, the values that make up the white background have been cleared from the image. The first step of turning the background values to black on the figure in Figure 7 in three steps.

Fig.7 Three steps of background removing from the image

The veins in the last image obtained in Figure 7 are quite thick. To remove the pattern of the veins, the veins must first be thick. The skeleton extraction algorithm has been applied on Figure 8. shows the result obtained by using the skeleton extraction algorithm.

Fig.8 Image obtained after skeleton extracted

There are a series of undesirable horizontal lines on the removed skeleton image in Figure 9. For this reason, the image is cleared of unwanted horizontal lines by using the erosion process again. During the erosion process, a 1 × 3 filter is used.

After this step, the expansion algorithm was implemented with a 5 × 1 rectangular filter to connect the breakpoints and remove the breakpoints. In the last stage, a 1 × 5 filter is opened. Thus, it is accepted that image values smaller than 5 pixels did not form a vein pattern and are cleared from the image. As a result of these successive operations, the skeleton i combined as a whole. The images obtained before and after the skeleton

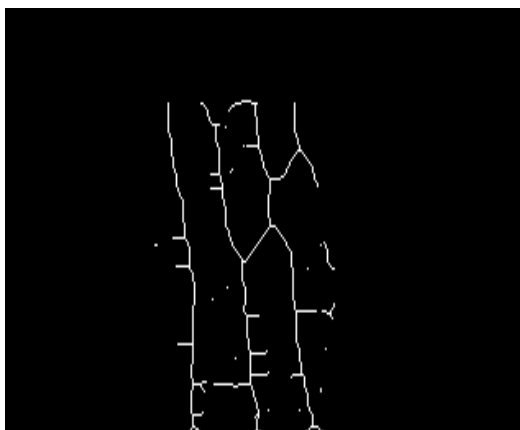assembly by filling the skeleton gaps are given in Figure 9 and Figure 10, respectively.

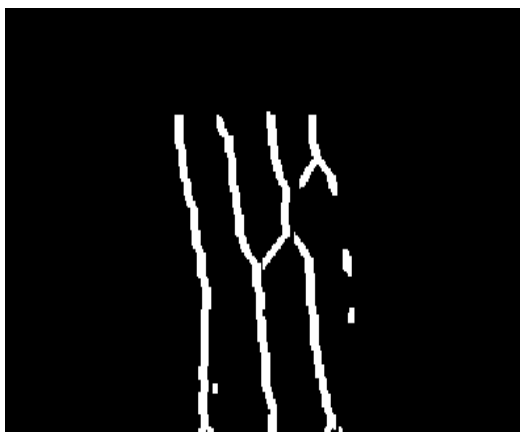**Fig.9** Image obtained before filling skeleton gaps

Fig.10 Image obtained after filling gaps on skeleton image

At this phase, only two features (entropy and variance) are obtained for each image. However, features need to be in order to increase the performance rate. Wavelet transform is used first to increase the features.

Using the wavelet transform, the image is reconstructed in four different directions. Figure 11 shows the four decomposable sub-pictures of this image represented by a wavelet transform.

Thus, using the wavelet transform, the feature number of each image is increased to eight features. By using wavelet transform for each hand image, a total of eight features are obtained.

Four of features are related to entropy and four of the features are related to variance, are determined. Single level 2-D wavelet decomposition is calculated and the compression technique is used for this operation.

Horizontal, vertical and diagonal matrices of the image are calculated in Table 11. The 2-dimensional wavelet transform function used is obtained by taking the one-dimensional wavelet and tensor of the scaling functions (Özkiper, 2019).
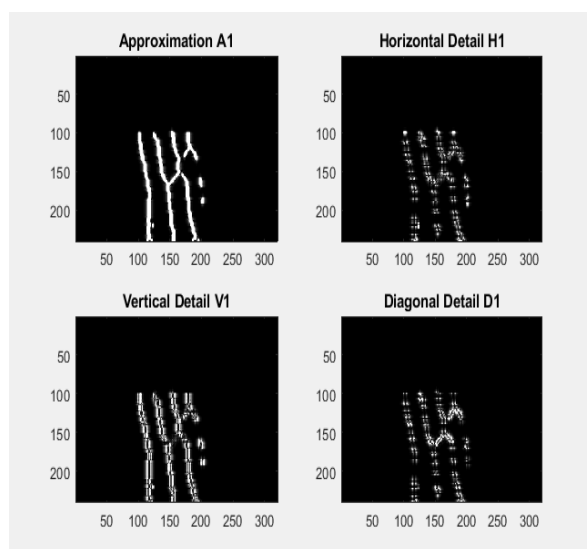
Fig.11 Images obtained after wavelet transform performed

The obtained features are classified with SVM classifier and performance measurement is performed. 80% of the dataset is used for training and 20% for testing purposes. Accuracy rate is obtained as 81.61% according to the results of multiple running experiments.

The features obtained are classified with LDA + KNN classifier and performance measurement is performed. 80% of the dataset is used for training and 20% for testing purposes. According to the results of multiple running experiments, the accuracy rate is 76.41.

In the use of ANN and CNN methods, entropy and variance values are not used since the feature extraction of images is calculated by the related methods.

The acquired qualities are classified with an ANN classifier and performance measurement is performed. 80% of the dataset is used for training and 20% for testing purposes. According to the results of multiple running experiments, the accuracy rate was 98.45.

The obtained qualities were classified with a double CNN classifier and performance measurement was performed. 80% of the dataset is used for training and 20% for testing purposes. According to the results of multiple running experiments, the accuracy rate was 99.64. Performance metrics using classifiers is presented in Table 1.

4. Results

In this study, a biometric identification system based on dorsal hand vein images was developed using a sample dataset. Badawi dorsal hand vein dataset was used as the sample data set. 4000 different images were obtained by adding noise to a total of 100 different images in the data set. Morphological operators, edge extraction algorithms etc. operations have been applied. Various machine learning methods have been used to identify images.

Artificial neural networks, nearest neighborhood, support vector machines and convolutional neural networks have been preferred as machine learning methods. In order to increase the performance of the k nearest neighbors method, the dimension was

Table 1. Performance metrics using different classifiers

Classifier	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)	F1 Score (%)	Matthews Correlation Coefficient (%)	Kappa (%)
SVM	0.8161	0.8404	0.8161	0.9981	0.8145	0.8196	0.8923
LDA+KNN	0.7641	0.7844	0.7641	0.9976	0.7604	0.7649	0.9161
ANN	0.9845	0.9872	0.9845	0.9998	0.9844	0.9850	0.2172
CNN	0.9964	0.9964	0.9964	1.000	0.9963	0.9965	0.8196

reduced by using linear discriminant analysis and then the classification process was carried out. After multiple runs and tests, the highest accuracy was achieved with the CNN method as 99.64%. CNN achieves higher accuracy compared to other machine learning methods with its multi-layered structure and strong representation of attributes on images.

In this study, it was observed that the identification performed using CNN achieved the highest accuracy. In future studies, performance measurements will be performed using CNN + SVM, CNN + Autoencoder, CNN + K Nearest Neighbor, CNN + Random Forest hybrid approaches in order to test the performance of CNN as a classifier in the field of image processing.

5. References

Alasadi, A. H. H., & Dawood, M. H. (2017). Dorsal hand-vein images recognition system based on grey level co-occurrence matrix and Tamura features. *International Journal of Applied Pattern Recognition*, *4*(3), 207. https://doi.org/10.1504/IJAPR.2017.086586

Badawi, A. M. (2006). Hand vein biometric verification prototype: A testing performance and patterns similarity. *Proceedings of the 2006 International Conference on Image Processing, Computer Vision, and Pattern Recognition, IPCV'06*, *1*, 3–9.

Bagosi, T., & Baruch, Z. (2011). Indoor localization by WiFi. *Proceedings - 2011 IEEE 7th International Conference on Intelligent Computer Communication and Processing, ICCP 2011*, 449–452. https://doi.org/10.1109/ICCP.2011.6047914

Benhar, H., Idri, A., & Fernández-Alemán, J. L. (2020). Data preprocessing for heart disease classification: A systematic literature review. *Computer Methods and Programs in Biomedicine*, *195*, 105635. https://doi.org/10.1016/j.cmpb.2020.105635

Charte, D., Charte, F., García, S., del Jesus, M. J., & Herrera, F. (2018). A practical tutorial on autoencoders for nonlinear feature fusion: Taxonomy, models, software and guidelines. *Information Fusion*, *44*(December 2017), 78–96. https://doi.org/10.1016/j.inffus.2017.12.007

Cross, J. M., & Smith, C. L. (1995). Thermographic imaging of the subcutaneous vascular network of the back of the hand for biometric identification. *IEEE Annual International Carnahan Conference on Security Technology, Proceedings*, (vi), 20–35. https://doi.org/10.1109/ccst.1995.524729

Ganegedara, T. (n.d.). Intuitive Guide to Convolution Neural Networks - Towards Data

Science. Retrieved June 28, 2020, from https://towardsdatascience.com/light-on-math-machine-learning-intuitive-guide-to-convolution-neural-networks-e3f054dd5daa

Jain, A. K. (2007). *Biometric recognition. Nature* (Vol. 449). https://doi.org/10.1038/449038a

Li, J., Feng, J., & Kuo, C. C. J. (2018). Deep convolutional neural network for latent fingerprint enhancement. *Signal Processing: Image Communication*, *60*, 52–63. https://doi.org/10.1016/j.image.2017.08.010

Misiti, M., Misiti, Y., Oppenheim, G., & Poggi, J.-M. (Eds.). (2007). *Wavelets and their Applications*. London, UK: ISTE. https://doi.org/10.1002/9780470612491

Özkiper, Z. İ. (2019). *Biyometrik kimlik tanimlama sistem tasarimi*. Halic University.

Patel, N., & Mishra, A. (2015). Automated Leukaemia Detection Using Microscopic Images. In *Second International Symposium on Computer Vision and the Internet* (Vol. 58, pp. 635–642). https://doi.org/10.1016/j.procs.2015.08.082

Shahin, M., Badawi, A., & Kamel, M. (2008). Biometric Authentication Using Fast Correlation of Near Infrared Hand Vein Patterns. *World Academy of Science, Engineering and Technology*, *2*(1), 756–763.

TTP, T., N. Pham, G., Park, J.-H., Moon, K.-S., Lee, S.-H., & Kwon, K.-R. (2017). Acute Leukemia Classification Using Convolution Neural Network in Clinical Decision Support System. *Computer Science & Information Technology (CS & IT)*, *7*, 49–53. https://doi.org/10.5121/csit.2017.71305

Turgut, Z., Üstebay, S., Zeynep Gürkaş Aydın, G., & Sertbaş, A. (2019). *Deep learning in indoor localization using WiFi. Lecture Notes in Electrical Engineering* (Vol. 504). https://doi.org/10.1007/978-981-13-0408-8_9

Varastehpour, S. (2020). *Visualising Vein Pattern Based on Sparse Auto-Encoder Algorithm*. Unitec Institute of Technology.

Veropoulos, K., Campbell, C., Cristianini, N., & Others. (1999). Controlling the sensitivity of support vector machines. *Proceedings of the International Joint Conference on Artificial Intelligence*, 55–60. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.42.7895%5Cnhttp://seis.bris.ac.uk/~enicgc/pubs/1999/ijcai_ss.pdf

Vijayanand, R., Devaraj, D., & Kannapiran, B. (2018). Intrusion detection system for wireless mesh network using multiple support vector machine classifiers with genetic-algorithm-based feature selection. *Computers and Security*, *77*, 304–314. https://doi.org/10.1016/j.cose.2018.04.010

Wang, Y., Zheng, X., & Wang, C. (2016). Dorsal Hand Vein Recognition Across Different Devices (pp. 307–316). https://doi.org/10.1007/978-3-319-46654-5_34