DEU FMD 27(80) (2025) 240-246

Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi
Dokuz Eylul University Faculty of Engineering Journal of Science and Engineering
Elektronik/Online ISSN: 2547-958X

RESEARCH ARTICLE / ARAŞTIRMA MAKALESI

Enhancing Finger Vein Classification through CLAHE and Sobel Filtering with Two Channel Hybrid Convolutional Machine Learning Algorithm

CLAHE ve Sobel Filtreleriyle Geliştirilmiş İki Kanallı Hibrit Evrişimli Makine Öğrenmesi ile Parmak Damar İzi Sınıflandırması

Berke Cansız [1*], Murat Taşkıran [2]

[1] Yıldız Teknik Üniversitesi, Elektrik-Elektronik Fakültesi, Biyomedikal Mühendisliği Bölümü, İstanbul, TÜRKİYE
[2] Yıldız Teknik Üniversitesi, Elektrik-Elektronik Fakültesi, Elektronik ve Haberleşme Mühendisliği Bölümü, İstanbul, TÜRKİYE

Corresponding Author / Sorumlu Yazar *: berke.cansiz@std.yildiz.edu.tr

Abstract

Advancements in digital technology have driven the rise of biometric security systems, notably in the field of finger vein detection. In most of the research on finger vein classification in the literature, achieving high accuracy is the main aim, while aspects such as generalization capacity and test distribution are mostly overlooked. In this study, two different datasets (MMCBNU_6000 and FV-USM) were tested with different test distributions, using a K-Fold structure for unbiased sampling in classification. In experiment part, two distinct image enhancement methods, namely Contrast Limited Adaptive Histogram Equalization (CLAHE) and Sobel filtering, were utilized on the datasets, and Convolutional Neural Networks (CNN) were used for feature extraction. Furthermore, machine learning algorithms were applied for classification, forming a Hybrid Convolutional Machine Learning algorithm. In this method, the model, which is fed through two different channels compared to conventional learning algorithms, combines classical machine learning classifiers with the CNN model. In the scope of this study, three tasks were outlined. The first two focused on implementing various machine learning algorithms for each dataset, while the third involved merging datasets and employing the Stacking Ensemble Classifier (SEC). For evaluating the models, accuracy and F1-score metrics were used. The results indicate that the highest accuracy rate was achieved in the third experiment, with a score of 98.94%. Additionally, it is also observed that increasing the amount of test data (the difference between 20% Test and 50% Test) has a minimal effect in reducing the model's accuracy metric compared to previous studies.

Keywords: *Convolutional Neural Networks, Machine Learning, Stacking Ensemble Learning, K-Fold Cross Validation, Finger Vein*

Öz

Dijital teknolojinin ilerlemesi, özellikle parmak damarı tespiti alanında biyometrik güvenlik sistemlerinin yükselişine sebep olmuştur. Literatürde parmak damarı sınıflandırması üzerine yapılan araştırmaların çoğunda yüksek doğruluk elde etmek ana amaç iken genelleme kapasitesi ve test dağılımı gibi konular genellikle göz ardı edilmektedir. Bu çalışmada, farklı test dağılımlarıyla iki farklı veri seti (MMCBNU_6000 ve FV-USM) K-Katlamalı yapı kullanılarak tarafsız örnekleme için test edilmiştir. Deney bölümünde, Kontrast Sınırlı Adaptif Histogram Eşitleme (KSAHE) ve Sobel filtreleme gibi iki farklı görüntü iyileştirme yöntemi veri setlerine uygulanmış ve özellik çıkarma için Evrişimli Sinir Ağları (ESA) kullanılmıştır. Ayrıca, sınıflandırma için makine öğrenimi algoritmaları uygulanmış ve Hibrit Evrişimli Makine Öğrenimi algoritması oluşturulmuştur. Bu yöntem, konvansiyonel öğrenme algoritmalarına kıyasla, iki farklı kanal ile beslenen model, klasik makine öğrenmesi sınıflandırıcıları ile ESA modelini birleştirmektedir. Bu doğrultuda çalışmada üç görev belirlenmiştir: ilk iki görevde her bir veri kümesi için çeşitli makine öğrenimi algoritmalarının uygulanması odaklanmışken, üçüncü görev veri kümelerinin birleştirilmesi ve Yığma Topluluk Sınıflandırıcısı (YTS) kullanımını içermiştir. Modellerin değerlendirilmesinde doğruluk ve F1-skoru metrikleri kullanılmıştır. Sonuçlar, en yüksek doğruluk skorunun %98.94 ile üçüncü deneyle elde edildiğini göstermektedir. Ayrıca test verisi sayısının artmasının (%20 Test ve %50 Test arasındaki fark) modelin doğruluk metriğinde önceki çalışmalara kıyasla minimal bir düşürme etkisine sahip olduğu gözlemlenmektedir.

Anahtar Kelimeler: *Evrişimli Sinir Ağları, Makine Öğrenmesi, Yığma Topluluk Öğrenmesi, K- Katlamalı Çapraz Doğrulama, Parmak Damar İzi*

1. Introduction

Today, in our rapidly digitizing world, security has emerged as a crucial concern. A system with simple login information like passwords is quite weak. In such a situation, biometric security systems are rapidly advancing, introducing novel aspects to the processes of personal identification and verification [1]. The word “Biometrics” is derived from the Greek words “Bios” and “Metron”, which means “life” and “measurement”, respectively. As the name suggests, it represents measurements related to living beings. It is expected to be unique to the individual and measurable, as well as repeatable. Biometric data can be essentially divided into two categories: physical (fingerprint, finger vein, face recognition, etc.) and behavioral (speech/speaker recognition, walking pattern, etc.). Among physical biometrics, finger vein can generally be examined under the vein detection category. The veins in our body have a highly complex structure and are uniquely developed for each

DOI: https://doi.org/10.21205/deufmd.2025278010

Geliş Tarihi / Received: 29.05.2024
Kabul Tarihi / Accepted: 22.08.2024

individual. This uniqueness makes it suitable for use in biometric systems. Veins that are not detectable in visible light waves (between 360 and 700 nm) become detectable in almost infrared wavelengths (800 to 1100 nm). This wavelength is known as Near Infrared (NIR). When this light wave enters our body, it is absorbed by the red blood cells, hemoglobin, in the blood. Therefore, if an image is obtained with a camera, the areas where veins are located appear black, while other parts (cell tissue, muscle, bone, etc.) are predominantly observed in white. Classification can be performed with these images; hence, identification and verification systems can be established. The labor-intensive process of obtaining this information puts this biometric at an advantage over systems that can be easily obtained and stolen, such as fingerprints.

With this study, a new perspective will be brought to finger vein classification problems through the developed model. Research results will be presented on various topics, including generalization capacity, the impact of different test distributions, and Feature/Decision level fusion.

The remaining parts of this paper will be structured as follows: In section 1.1 related works will be discussed. In section 2, the finger vein databases will be introduced, and the established model will be presented. Additionally, the tasks implemented in this study will be introduced. In section 3, the obtained results will be demonstrated. Lastly, in section 4, conclusion and discussion parts will be presented.

1.1. Related Works

The acquisition of fast and secure results through finger vein data is considered a significant aim. Consequently, computer-assisted algorithms are progressively becoming more prevalent and actively evolving [2], [3], [4], [5] in this field. As an example, in [6], the skeletal structure of finger vein images was formed by enhancing them with Gabor filters, followed by the segmentation of vascular shapes. Then three different feature types extracted from segmented vein images and fused for classification part. The first extracted feature type is the local moment feature, obtained by sliding windows (70x70) across the image. Through 20 steps of this sliding process, 5 new sub-images were generated, and a total of (7x5) features were obtained by extracting 7 moment features from each. Another feature type is the topological features, that derived from the connections between cross-points. The final feature type is the vein shape feature, obtained by extracting statistical features (mean, variance, skewness, and kurtosis). The extracted features were classified using a nearest cosine classifier. Although they achieved highly successful results (Accuracy rate 97.51%, 96.44% and 98.50% for each feature type, respectively) with their own experimental data, the classification method (nearest cosine classifier) may not perform as well for larger feature sets. Additionally, obtaining topological features for each test data is laborious. Utilizing more robust classification algorithms could enhance the model's generalization capacity. Therefore, Convolutional Neural Network (CNN) is frequently employed in such problems. It utilizes learnable filters in its convolutional layers for feature extraction. Subsequently, classification is performed based on the weights connected by neural networks within its structure. An example of studies employing this method is demonstrated in [7]. They use three image enhancement methods, which are CLAHE [8], Gabor Filter applied images and fused version two of them in Discrete Cosine Transform (DCT), then their models were fed with these three enhanced images. In their model, which construct from scratch, they tried various parameters and put optimum results (accuracy between 70.1% to 99.56% across various datasets) forward. Selecting the appropriate image enhancement method is very important, and various filters were applied in many studies [9], [10].

In another research [11], they use residual Gabor convolutional network (RGCN) which construct by residual Gabor filter in convolutional layers, for feature extraction. During their experiment they used finger vein mixture (FV-MIX) method for data augmentation. They present their result (range of between 75.83% to 100% accuracy) both augmented and not augmented version of each dataset. In the classification process of CNN architectures, the choice of the loss function is important. In [12], they introduced a new loss function named arccosine center loss, which they utilized in their work. Additionally, in this study, they integrated this loss function with the efficient channel attention residual network they devised and conducted experiments on four different sets. They achieved accuracy results ranging from 99.25% to 99.93%.

The construction of CNN from scratch can be laborious and inefficient in some cases. Also, to construct a deep model makes algorithm too complex and it causes forgetting the learnable things in first layer which called as vanishing gradient. As a solution to the vanishing gradient problem, in [13], observed that a CNN structure with two distinct sub-convolutional networks, each having different dimensions, increased the success compared to a single network of similar size (accuracy increase from 91.7% to 95%). However, these methods may still not be sufficient for achieving elevated results. Therefore, pre-trained deep learning models, that have proven their success on wide range of datasets, can be utilized through the transfer learning method. This approach not only overcomes the vanishing gradient problem but also yields robust results. In some studies, these networks can be retrained, or their already trained versions can be utilized. For instance, in [14], they employed the Densenet-161 architecture with modified version, along with matching using two different methods. They achieved an Equal Error Rate (EER) of 0.405% on the finger vein dataset.

Last layer of the CNN architecture, classification is performed with neural networks, and during training, it expects a substantial amount of data for each class. However, datasets like finger vein have a limited number of samples for each individual. Hybrid models can be solution for such problems. In these cases, the CNN architecture is utilized for the feature extraction process, and the final layer is fed into machine learning algorithms. For example, in [15] they employ VGG19 and ResNet50 architectures with transfer learning as feature extractors, followed by the use of a Support Vector Machine (SVM) to classify the extracted vectors.

Transformer models [16], which have become quite popular recently, are emerging as an alternative to CNN networks, especially after starting to be applied to image problems [17]. Numerous studies [18], [19], [20] using these model architectures can be found in the literature. However, it is important to note that achieving efficiency from such a network requires a very large dataset, which is not typically available for finger vein problems. [21].

In this proposed study, two image processing methods will be used. The first one is Contrast Limited Adaptive Histogram Equalization (CLAHE), and the other is the Sobel filtering image enhancement method. These two images will be fed into a CNN algorithm for feature extraction through two separate channels. The DenseNet201 architecture will be applied as the CNN model with pre-trained weight using Transfer Learning. The dense layer will not be added to the model; instead, the DenseNet 201 output vector was transferred to the classification algorithms. Machine learning methods, which are SVM, LDA and Multi-Layer Perceptron (MLP), have been chosen for classification. Various tasks will be identified based on the main motivations of the study. The first motivation is to increase the number of test samples compared to examples in the literature and achieve successful results. The second motivation is to merge both datasets and apply the Stacking Ensemble Classifier (SEC) algorithm, which is created by successful classifiers in the first two tasks.

2. Materials and Methods

In this section, detailed explanations will be provided sequentially, including a description of the utilized datasets, the image processing methods, an overview of the established models, and the classification phase.

2.1. Data Description

In this study utilizes two distinct datasets. The first dataset, MMCBNU-6000[22], which was published by the Multimedia Lab, Division of Electronic and Information Engineering, Chonbuk National University, comprises a total of 6000 images from a community of 100 individuals. For each person, there are 10 repeated images of 6 fingers, making a total of 600 classes. The ROI output, obtained in this [23] study, will be used directly for the subsequent image processing step. The second dataset, FV-USM[24], which was published from Universiti Sains Malaysia (Science University of Malaysia), is collected in two parts, consisting of 6 repeated vein pattern data for 4 fingers of each individual. In total, there are 5904 images from 123 different individuals across two sessions. In this study, each finger is individually investigated, resulting in a total of 492 classes. The ROI images is readily available along with this dataset. Images from both datasets are provided in Figure 1

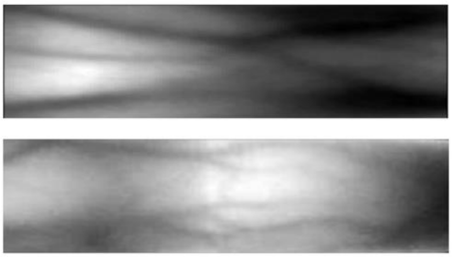

Figure 1. ROI images of Each Dataset (MMCBNU-6000[22], [23], and FV-USM[24], respectively).

2.2. Image Processing & Enchancement

The image processing and enhancement stage is a crucial step before the feature extraction, which is important for improving the outcome of classification. In this proposed study, two different image enhancement methods will be applied. The parameters and application specifics will be explained in the following sections.

2.2.1. Contrast Limited Adaptive Histogram Equalization (CLAHE)

Histogram equalization (HE) methods are visual enhancement techniques used to highlight details in an image. However, applying this process uniformly across all pixel values may yield suboptimal results when different regions exhibit distinct lighting conditions. To address this, the image needs to be divided into specific regions for localized equalization. This process is known as Adaptive Histogram Equalization (AHE). Despite its higher cost due to regional processing compared to HE, AHE provides a better solution. One significant drawback of AHE is that, depending on regional variations, it can lead to excessive contrast changes in some parts of the image compared to the overall contrast, disrupting the integrity of the image. Therefore, the CLAHE method, which has limited contrast value, ensures a more coherent image. Hence, in this study, this method will be used to create a base image.

CLAHE operates with two different parameters: "clipLimit", which determines the contrast limit, and "tileGridSize", which specifies the size of the masking grid. Based on preliminary work on the images, different parameters have been selected for both sets. For the MMCBNU-6000 dataset, these are determined as 20 and (5,5), respectively, and for the FV-USM dataset, they are set as 20 and (3,3). The CLAHE outputs are illustrated in Figure 2.

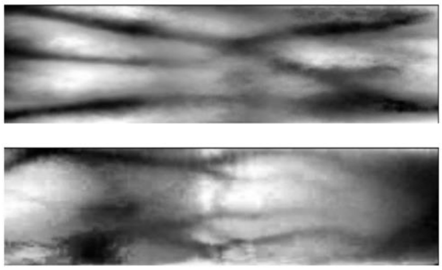

Figure 2. CLAHE Enhanced Images of Each Dataset (MMCBNU-6000 and FV-USM, respectively).

2.2.2. Sobel Filter Image Enhancement

The application of a Sobel filter to highlight the border regions of the image can enhance the visibility of vascular regions. In this proposed study, this process is planned to be conducted on images already enhanced with CLAHE. Thus, by initially applying CLAHE to enhance the visibility of veins and subsequently reinforcing it with Sobel, a more robust enhancement will be achieved. The Sobel operation will be performed in two dimensions, along the x and y axes, and then combined with a 50% weight contribution from each to facilitating the detection of diagonal veins. The kernel size, determined in preliminary experiments, is selected as 7 for both sets. The image enhanced with Sobel is illustrated in Figure 3.

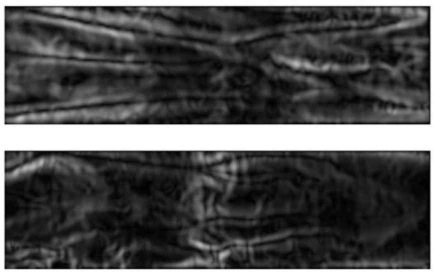

Figure 3. Sobel Enhanced Images of Each Dataset (MMCBNU-6000 and FV-USM, respectively)

Following this process, both images have been forwarded to the next stage.

2.3. Feature Extraction

In this section, features will be extracted from the images obtained from the previous stage. Convolutional layers of CNN networks will be utilized for this process. Various filters are applied in each layer by convolutional networks to detect features. While in the initial layer, it makes detections such as lines and corners, as the model becomes deeper, it begins to extract more complex shapes. In this study, it is planned to utilize robust CNN models that have been tested on large datasets through the transfer learning method. Accordingly, in preliminary work, models such as VGG19, EfficientNetB7, and

DenseNet201 were tested, and due to the efficient results obtained with the DenseNet201 architecture, it has been decided to employ it for this feature extraction process.

Unlike linearly evolving CNNs, DenseNet201 differs by feeding the outputs of its layers into subsequent layers, thereby preventing data loss in the flow and offering a solution to the vanishing gradient problem. This mechanism enables the establishment of a deep convolutional network. In this study, the proposed feature extraction method utilizes the DenseNet201 architecture as the feature extractor, and it will be fed with enhanced images through two different channels. To achieve this, two separate branches will be established, processing CLAHE images in one path and Sobel images in the other. The final layer outputs will be obtained using the Global Average Pooling method, and both outputs will be concatenated. Thus, the initial fusion process will be conducted at the feature level at this stage.

Train and test data were subjected to feature extraction separately. Once the features were extracted, they underwent a standardization process before proceeding to the classification step.

2.4. Classification

The classification process of the extracted features will be explained in this section. Various classification algorithms have been tried for this process. The first of these is the SVM, originally introduced by Cortes and Vapnik [25] for binary classification. Over time, it has become applicable for multi-class problems. SVMs detect the widest margins between classes to separate them and generate decision outputs. Both linear kernel and non-linear Radial Basis Kernel (RBF) can be employed in this study. The cost "C" parameter determines the penalty score to be applied when misclassification occurs, while the "gamma" parameter defines the area for the RBF kernel.

Another classification algorithm is the Multi-Layer Perceptron. Structurally resembling neural networks found in living organisms, this classification algorithm fires nodes based on weights and generates decision outputs. The number of layers determines the depth of the model, and as the number of layers increases, more complex shapes are learned. This classifier was included to compare proposed model with the conventional CNN.

The final classifier is the Linear Discriminant Analysis (LDA) method. LDA enhances the boundaries between features by emphasizing differences and similarities, with creating subspaces.

Parameter selections were made through trial in the preliminary study. Accordingly, for the SVM linear kernel, the C parameter is set to 0.01, while for the RBF kernel, the C parameter is set to 1, and the gamma parameter is set to "auto". In the MLP classifier, the number of hidden layers is set to 300, the solver parameter is set to "sgd", the learning rate is set to "adaptive", and the activation function is set to "relu". Finally, in the LDA algorithm, the solver parameter is set to "svd", and the shrinkage is set to "None".

2.5. Ensemble of Classifiers

The decision level fusion aiming to achieve more robust results by combining the models that yield the best outcomes. In proposed model, the utilization of the Stacking Ensemble Classifier (SEC)[26] method was selected. The SEC algorithm can be examined in two steps. The first step involves the selection of individual classifiers, known as base models, while the second step involves the creation of the decision output, known as the meta model. This method fundamentally combines the learning algorithms of base models, allowing weak learner classifiers to contribute to strong classifiers, leading to more robust results. By integrating the perspectives of different experts, learning is enhanced. Unlike other ensemble methods, this algorithm reclassifies the probability distributions of base model outputs using a new classifier, resulting in much more successful outcomes than a simple voting process.

The base models were selected with algorithms that yielded the highest results in previous task, then meta model performs classification using logistic regression (LR). The SEC algorithm initially trains the base models and generates decision outputs. Subsequently, based on these outputs, LR undergoes final training. In other words, meta model is trained not based on data but rather on the output probabilities of the base models.

All classification algorithms were implemented by using Scikit-learn library [27].

2.6. Experiment and Evaluation

Stratified K-fold structure has been employed at each stage of the experiment. K-fold involves dividing the data and sequentially setting each portion as a test and train set. This ensures that the analysis is performed not only on a specific subset of the data but across the entire dataset. In this study, various values of K ranging from 2 to 6 have been employed depending on the tasks and data distribution. The study involves three main tasks in working with the data:

Task A: Training MMCBNU_6000 dataset using individual classifiers. In the tests within the scope of this task, the K value for K-Fold was selected as 2 and 5, as shown in Figure 4.

Task B: Training FV_USM dataset using individual classifiers. For the tests conducted as part of this task, the K value for K-Fold was chosen as 6, 4, 3 and 2, as shown in Figure 4.

Task C: Combining the strongest individual classifiers in the SEC algorithm, creating a unified model, and feeding both datasets into the SEC algorithm. the K value for K-Fold was chosen as 6, 4, 3 and 2, as shown in Figure 5.

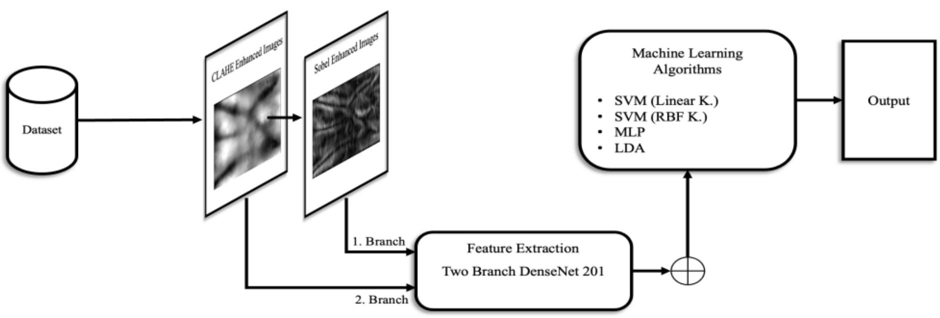

Figure 4. Proposed Model for Task A and B.

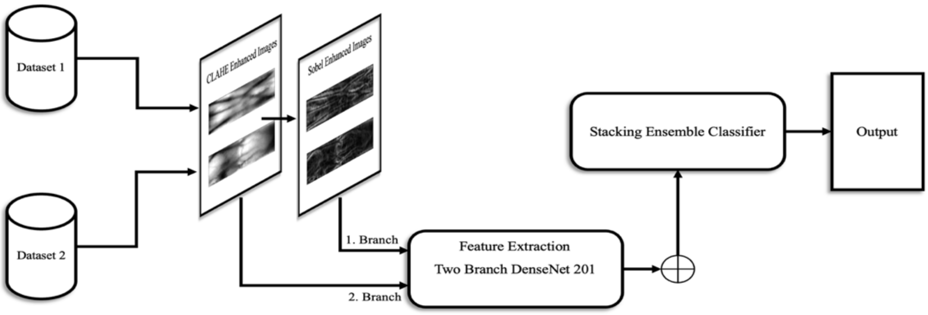

Figure 5. Proposed Model for Task C (Using the SEC Algorithm in the Combination of Two Different Datasets).

3. Results

In this section, the results obtained from the experiment were examined for each task. The results were represented with different splitting ratio using Accuracy and F1-Score, allowing for comparisons.

3.1. Task A and B Result Performance

The result table for Task A is as shown in Table 1, containing the accuracy and F1-Score results obtained for the MMCBNU_6000 dataset with the proposed model. Table 2 displays the results for Task B obtained with the FV-USM dataset. The outcomes vary depending on the different K-Fold applications on both datasets, leading to different training and test distributions. The findings from the table indicate the superiority of SVM (linear kernel) and LDA classifiers. Additionally, it is observed that as the training set increases and the test data decreases, higher success rates are achieved. In Task A, the highest result was obtained with SVM (linear kernel) at %80 Train - %20 Test data distribution, reaching 98.1%. In Task B, the highest achievement was obtained with the SVM classifier (linear kernel) at %83.3 Train %16.67 Test data distribution, reaching 98.8%.

3.2. Task C Results Performance

The results presented in Table 3 belong to Task C, representing the outcomes of the SEC algorithm created using successful algorithms (SVM and LDA) from the individual classification (Task A and B) section. In this task, the combination of two datasets formed a unified dataset, and 2, 3, and 5 K-Fold structures were employed. Upon examination of the experimental results, it was observed that an increase in the test data accompanied by a decrease in the training data negatively impacted the model's performance. While a 97.22% accuracy was achieved with the 2 K-Fold structure, the 5 K-Fold structure yielded results of 98.94%. Combining all available datasets for this study has increased the generalization capacity of the data, contributing to a more objective result.

Table 1. Accuracy and F1-Score for Task A (with Different Train Test Split Ratio).

	MMCBNU_6000 Dataset					
	5 K-Fold (%80 Train %20 Test)		2 K-Fold (%50 Train %50 Test)		5 K-Fold (%20 Train %80 Test)	
	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score
SVM_RBF	0.9645	0.9638	0.9537	0.9569	0.8601	0.8589
SVM_LIN	**0.9810**	**0.9794**	**0.9755**	**0.9748**	0.8911	0.8878
MLP	0.951	0.9475	0.9548	0.9536	0.8128	0.8053
LDA	0.9757	0.9741	0.9740	0.9750	**0.9279**	**0.9260**

Table 2. Accuracy and F1-Score for Task B (with Different Train Test Split Ratio).

	FV-USM Dataset					
	6 K-Fold (%83.3 Train %16.67 Test)		4 K-Fold (%75 Train %25 Test)		3 K-Fold (%66.67 Train %33.3 Test)	
	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score
SVM_RBF	0.9758	0.9741	0.9726	0.9717	0.9670	0.9668
SVM_LIN	**0.9880**	**0.9871**	**0.9834**	**0.9826**	**0.9807**	**0.9802**
MLP	0.9638	0.9605	0.9588	0.9564	0.9485	0.9468
LDA	0.9854	0.9845	0.9094	0.9063	0.9707	0.9709

	FV-USM Dataset			
	2 K-Fold (%50 Train %50 Test)		6 K-Fold (%16.67 Train %83.3 Test)	
	Accuracy	F1-Score	Accuracy	F1-Score
SVM_RBF	0.9492	0.9492	0.8024	0.8002
SVM_LIN	0.9687	0.9679	0.8364	0.8310
MLP	0.9289	0.9263	0.7235	0.7118
LDA	**0.9741**	**0.9741**	**0.8750**	**0.8706**

Table 3. Accuracy and F1-Score for Task C (with Different Train Test Split Ratio).

	MMCBNU_6000 & FV-USM Dataset					
	2 K-Fold (%50 Train %50 Test)		3 K-Fold (%66.67 Train %33.37 Test)		5 K-Fold (%80 Train % 20 Test)	
	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score
SEC	0.9722	0.9730	0.9861	0.9864	**0.9894**	**0.9891**

4. Conclusion and Discussion

In the digitally evolving era, biometric data has become highly valuable for security systems. Biometric data is characterized by its presence in every individual, distinctiveness, and measurability. Finger vein is among these biometric features and has become increasingly common research topic over time. In this study, the models constructed by using finger vein datasets which are using in the previous studies. Accordingly, two different image processing method, which are CLAHE and Sobel, were utilized to enhance images and benefited from CNN networks for feature extraction. At this point, DenseNet201 architecture was added into the model through transfer learning, and it was fed with two separate channels. Then, by utilizing the feature fusion method, a vector output was obtained. This output was fed into the classifiers to generate a decision output. Then the model was tested for three different tasks and obtained results were reported. The first task involved testing the MMCBNU-6000 dataset using SVM, MLP, and LDA methods. The SVM function was experimented with using both non-linear RBF kernel and Linear kernel separately. When the results are examined, it is observed that the SVM (linear kernel) function and LDA are characterized by quite high performance. Two different interpretations can be deduced from this. Firstly, it was revealed that the features are linearly separable; secondly, it implies that more advanced results can be obtained compared to traditional CNN applications with an MLP function in the final layer. These situations are also valid for Task B.

Task C constitutes the main motivation of this study. The merging of the dataset resulted in a more comprehensive solution, enhancing the model's capacity for generalization. To the best of our knowledge, this is the first study to incorporate the merging of these two datasets. This allowed for the examination of a more extensive dataset. Additionally, the application of the SEC algorithm resulted in the combination of the two classifiers that individually yielded the highest performance, leading to a more robust solution. Examining the results, the accuracy of 98.94% obtained for Task C, demonstrates the success of the model. Our second motivation was to try this model on a more extensive test set and achieve minimal impact on the accuracy rate decrease resulting from this. In this context, it can be observed that quite successful (%1.7) results were obtained.

One of the main goals of the study was to increase the average success rate through the fusion process performed in the two-branch CNN network. When examining Figures 6, 7, and 8, it can be observed that this goal has been successfully achieved.

The study demonstrates that feature fusion using two different channels contributes to the results (see Figures 6, 7, and 8). Therefore, it can be observed that improving the outcomes with such a system is possible. Although the results are relatively low in classification of using only Sobel or only CLAHE features, it has been observed that weak features can complement each other and improve the outcome when two channels are used.

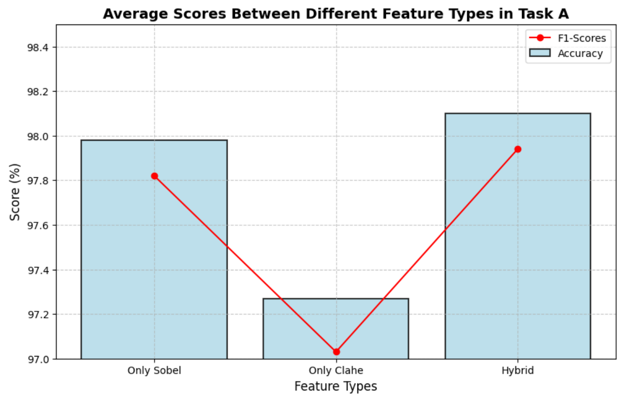

Figure 6. The Average Scores Between Different Feature Types in Task A.

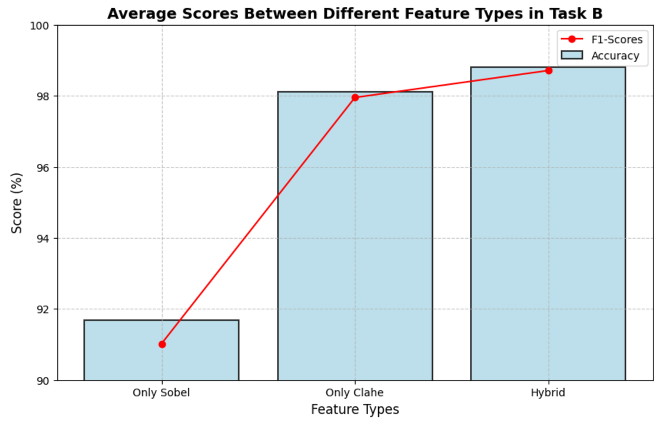

Figure 7. The Average Scores Between Different Feature Types in Task B.

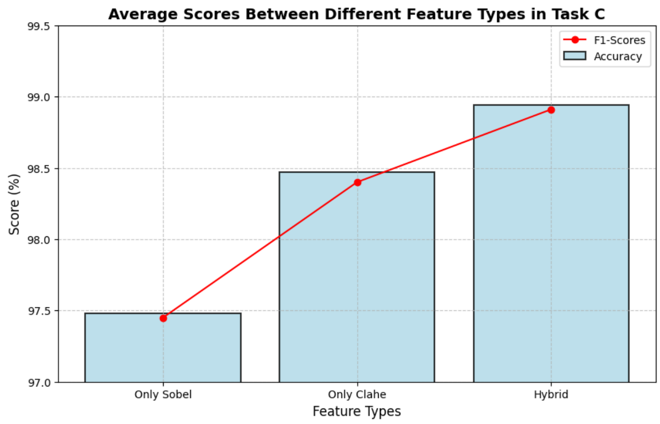

Figure 8. The Average Scores Between Different Feature Types in Task C.

Although the study was generally successful when compared with the state of the art, various limitations were encountered during its construction. The first one is the considerable amount of time spent training the model due to it running on the CPU. However, considering that the training process will be performed only once, this will not pose such a significant problem. The sequential application of image processing, feature extraction and classification also repetitively structure like K-fold also creates a considerable cost issue. In later systems, using a single model for all test operations would significantly reduce the time required, making the training cost negligible.

In further studies, models can be constructed to create a more comprehensive system and lighter the processing load. Additionally, models that ensure a high level of data security can be built by employing federated learning approaches for the combining of multiple datasets. Alternatively, synthetic data augmentation using generative networks could lead to much more successful results. Furthermore, the system can be transferred to an embedded structure and tested with real-time data, or it can be released as a final product.

Author Contribution Statement

Berke Cansız: Writing, Software, Methodology.

Murat Taşkıran: Writing, Supervision, Conceptualization.

References

[1] Shaheed, K., Mao, A., Qureshi, I., Kumar, M., Hussain, S., Zhang, X. 2022. Recent advancements in finger vein recognition technology: methodology, challenges and opportunities, Inf. Fusion, Vol. 79, pp. 84–109.

[2] Lian, F.-Z., Huang, J.-D., Liu, J.-X., Chen, G., Zhao, J.-H., Kang, W.-X. 2023. FedFV: A personalized federated learning framework for finger vein authentication, Mach. Intell. Res., Vol. 20, No. 5, pp. 683–696.

[3] Zhang, L., Li, W., Ning, X., Sun, L., Dong, X. 2021. A local descriptor with physiological characteristic for finger vein recognition, in: 2020 25th International Conference on Pattern Recognition (ICPR), pp. 4873–4878. DOI: 10.1109/ICPR48806.2021.9412203.

[4] Zhang, L., et al. 2022. A joint Bayesian framework based on partial least squares discriminant analysis for finger vein recognition, IEEE Sens. J., Vol. 22, No. 1, pp. 785–794. DOI: 10.1109/JSEN.2021.3130951.

[5] Lv, W., Ma, H., Li, Y. 2023. A finger vein authentication system based on pyramid histograms and binary pattern of phase congruency, Infrared Phys. Technol., Vol. 132, p. 104728. DOI: 10.1016/j.infrared.2023.104728.

[6] Yang, J., Shi, Y., Yang, J., Jiang, L. 2009. A novel finger-vein recognition method with feature combination, in: 2009 16th IEEE International Conference on Image Processing (ICIP), pp. 2709–2712.

[7] Boucherit, I., Zmirli, M.O., Hentabli, H., Rosdi, B.A. 2022. Finger vein identification using deeply-fused convolutional neural network, J. King Saud Univ. Comput. Inf. Sci., Vol. 34, No. 3, pp. 646–656. DOI: 10.1016/j.jksuci.2020.04.002.

[8] Pizer, S.M. 1990. Contrast-limited adaptive histogram equalization: speed and effectiveness, in: Proceedings of the First Conference on Visualization in Biomedical Computing, Atlanta, Georgia, p. 2.

[9] Zhang, L., Wang, X., Dong, X., Sun, L., Cai, W., Ning, X. 2021. Finger vein image enhancement based on guided tri-Gaussian filters, ASP Trans. Pattern Recognit. Intell. Syst., Vol. 1, No. 1, pp. 17–23.

[10] Kilic, U., Karabey Aksakalli, I., Tumuklu Ozyer, G., Aksakalli, T., Ozyer, B., Adanur, S. 2023. Exploring the effect of image enhancement techniques with deep neural networks on direct urinary system (DUSX) images for automated kidney stone detection, Int. J. Intell. Syst., Vol. 2023, No. 1, p. 3801485.

[11] Wang, Y., Lu, H., Qin, X., Guo, J. 2023. Residual Gabor convolutional network and FV-Mix exponential level data augmentation strategy for finger vein recognition, Expert Syst. Appl., Vol. 223, p. 119874.

[12] Hou, B., Yan, R. 2021. ArcVein-Arccosine center loss for finger vein verification, IEEE Trans. Instrum. Meas., Vol. 70, pp. 1–11. DOI: 10.1109/TIM.2021.3062164.

[13] Zhang, Y., Liu, Z. 2020. Research on finger vein recognition based on sub-convolutional neural network, in: 2020 International Conference on Computer Network, Electronic and Automation (ICCNEA), pp. 211–216. DOI: 10.1109/ICCNEA50255.2020.00051.

[14] Kuzu, R.S., Maiorana, E., Campisi, P. 2020. Vein-based biometric verification using transfer learning, in: 2020 43rd International Conference on Telecommunications and Signal Processing (TSP), pp. 403–409. DOI: 10.1109/TSP49548.2020.9163491.

[15] Tao, Z., Zhou, X., Xu, Z., Lin, S., Hu, Y., Wei, T. 2021. Finger-vein recognition using bidirectional feature extraction and transfer learning, Math. Probl. Eng., Vol. 2021, p. 6664809. DOI: 10.1155/2021/6664809.

[16] Vaswani, A., et al. 2017. Attention is all you need, Adv. Neural Inf. Process. Syst., Vol. 30.

[17] Dosovitskiy, A., et al. 2020. An image is worth 16x16 words: transformers for image recognition at scale, ArXiv Prepr. ArXiv201011929.

[18] Zhao, P., et al. 2024. VPCFormer: A transformer-based multi-view finger vein recognition model and a new benchmark, Pattern Recognit., Vol. 148, p. 110170. DOI: 10.1016/j.patcog.2023.110170.

[19] Li, X., Zhang, B.-B. 2023. FV-ViT: Vision transformer for finger vein recognition, IEEE Access.

[20] Garcia-Martin, R., Sanchez-Reillo, R. 2023. Vision transformers for vein biometric recognition, IEEE Access, Vol. 11, pp. 22060–22080.

[21] An, Z., Ren, X., Tao, Z. 2024. FV-DMHN: Dual multi-head network for finger vein recognition, IEEE Access.

[22] Lu, Y., Xie, S.J., Yoon, S., Wang, Z., Park, D.S. 2013. An available database for the research of finger vein recognition, in: 2013 6th International Congress on Image and Signal Processing (CISP), pp. 410–415. DOI: 10.1109/CISP.2013.6744030.

[23] Lu, Y., Xie, S.J., Yoon, S., Yang, J., Park, D.S. 2013. Robust finger vein ROI localization based on flexible segmentation, Sensors, Vol. 13, No. 11, pp. 14339–14366. DOI: 10.3390/s131114339.

[24] Mohd Asaari, M.S., Suandi, S.A., Rosdi, B.A. 2014. Fusion of band limited phase only correlation and width centroid contour distance for finger based biometrics, Expert Syst. Appl., Vol. 41, No. 7, pp. 3367–3382. DOI: 10.1016/j.eswa.2013.11.033.

[25] Cortes, C., Vapnik, V. 1995. Support-vector networks, Mach. Learn., Vol. 20, No. 3, pp. 273–297. DOI: 10.1007/BF00994018.

[26] Wolpert, D.H. 1992. Stacked generalization, Neural Netw., Vol. 5, No. 2, pp. 241–259. DOI: 10.1016/S0893-6080(05)80023-1.

[27] Pedregosa, F. 2011. Scikit-learn: Machine learning in Python, J. Mach. Learn. Res., Vol. 12, p. 2825.