RESEARCH PAPER

Swafv: Secure Web Authentication Using Multi-Finger Vein Biometrics With Spoof Detection And Betl

[1*]Vathsala V, [2]Pazhanikumar K

[1]Research Scholar, Register No: 23113152282008, Department of Computer Science, S.T. Hindu College, Nagercoil, India. Affiliated to Manonmanium Sundaranar University, Abishekapatti, Tirunelveli-627012, Tamil Nadu, India.*

[2]Head and Assistant Professor, Department of Computer Science, S.T. Hindu College, Nagercoil, India. Affiliated to Manonmanium Sundaranar University, Abishekapatti, Tirunelveli-627012, Tamil Nadu, India.

Corresponding author email id: vathsalav100@gmail.com

ABSTRACT

This research proposes a highly secure and efficient user authentication system for web applications using finger vein biometrics, named Secured Authentication for Web Applications using Finger Vein Biometrics (SAWFV). Unlike traditional password-based or unimodal biometric systems that suffer from spoofing risks, low accuracy, and slow response times, SAWFV leverages images from three fingers (index, middle, and ring) to enhance uniqueness and robustness. The system begins with Enhanced Contrast Limited Adaptive Histogram Equalization (ECLAHE) for vein enhancement, followed by a novel spoof detection approach using an ensemble of deep networks and a modified score fusion strategy to effectively identify fake biometric attempts. Feature extraction is performed using a Bidirectional Attention Module combined with Spatial Pyramid Pooling and a transfer learning-based architecture, ensuring both global and local pattern capture. The bidirectionally refined features are concatenated with general EfficientNet-derived features to form a strong representation, which is then classified using a Support Vector Machine (SVM) to accurately distinguish genuine users from imposters. Implemented in Python, the SAWFV framework achieves superior results with 98.65 accuracy, 97.86 precision, and low FPR and FNR values of 0.038 and 0.018, respectively making it a reliable solution for secure access in applications such as online banking, defense, and government portals.

Keywords: *Finger Vein Biometrics, Spoofing Detection, Ensemble Learning, VGG-16, XGBoost, Support Vector Machine, Web Application Security.*

How to cite this article: Vathsala V, Pazhanikumar K. SWAFV: Secure Web Authentication Using Multi-Finger Vein Biometrics with Spoof Detection and BETL. Int J Drug Deliv Technol. 2026;16(38s): 456-466. DOI: 10.25258/ijddt.16.38s.43

1. INTRODUCTION

Web applications are increasingly becoming core aspects of the modern digital world, allowing users to access services, conduct transactions, and manage their personal data [1][2]. As they become more widespread, ensuring secure and reliable access to web applications has become an essential factor in maintaining user trust, privacy, and data integrity. As a result, integrating seamless and reliable authentication methods has become an essential requirement for modern web-based applications [3].

Despite improvements, reliable user authentication in web applications remains a challenge due to changing digital threats such as credential theft and reuse [4]. Finger vein biometrics offer promising secure user authentication as thumb and finger vein biometric templates or biometric features are based on internal vascular patterns that are more difficult to forge than other biometric modalities [5][6]. However, challenges such as variable image quality, because of inconsistencies in applied light, imaging pressure during acquisition, and motion, complicate liveness detection and presentation attack resistance and thus their practical reliability in high-security environments [7].

Finger vein (FV) authentication depends on hardware and software optimization. NIR light sources and CCD cameras are widely used to capture vein patterns, whereas advanced image enhancement and feature extraction techniques for example, PCA, CMW filtering technique, or the Hierarchical Centroid Feature Method (HCM) provide added accuracy [8][9]. Image enhancement techniques applied in a specific ROI for aspect-ratio based feature detection and classification may provide even better identification performance in biometric systems [10].

In recent years, deep learning (DL) and machine learning (ML) have transformed biometric authentication through improvements in accuracy, robustness, and anti-spoofing capabilities across a range of domains, including online banking [11], defense [12], and secure web access [13]. Both techniques learn and derive complex representations from biometric data using deep and machine learning architectures which enable reliable recognition solutions [14][15]. Motivated by this development, this paper presents a framework for a finger vein-based high-security user authentication solution for web applications that incorporates both deep and machine learning methods with the goal of providing identity verification that is reliable and robust. The key contribution of the works is as follows:

- A robust preprocessing pipeline is proposed using image resizing, noise removal, ROI extraction, and a novel ECLAHE method with spatial weighting to enhance contrast and preserve vein clarity.
- An ensemble model combining CNN, ResNet-18, and LSTM is used for spoof detection, with a Modified Score Fusion method employing global-local inertia weight to improve decision accuracy.
- A novel Bidirectional Efficient Transfer Learning (BETL) architecture is proposed for deep feature extraction, integrating both bidirectional processing and pretrained EfficientNet modules to capture multi-scale spatial and contextual features, and the last classification is completed. through SVM, or support vector machine to achieve highly accurate and secure multi-finger vein-based user authentication.

The sequence of these sections is as follows: The suggested framework is introduced in Section 3, some

pertinent research and literature studies are covered in Section 2, a thorough analysis of the observed outcomes and responses is offered in Section 4, and the study's final evaluation is given in Section 5.

2 LITERATURE REVIEW

A few of the most recent studies about finger vein biometrics were reviewed in this section.

Shakil *et. al.,* presented a hybrid machine ML method and conventional methods are use SVM is used for classification, and principal component analysis (PCA) is used to extract vein patterns. The technique depends only on the various patterns' characteristics of the vein. The effective framework has been developed using the best optimal design [16]. Al-Obaidy *et.al.,* introduced the Deep Fingers Vein Learning (DFVL) deep model to improve recognition accuracy using training with Convolutional Neural Network. The resulting model consists of three layers: classification, pooling, fully connected, soft-max, convolutional, and ReLU. All this happens after the hand image has gone through the basic steps of preprocessing operations in order to identify the region of interest [17]. Liu *et. al.,* developed a DL model based on Loss function for Fourier-space Structural Similarity that was intended to optimize the education potential of the denoising mechanism and improve performance for the entire network. Also presented is the Transformer of Semantic Difference, which is combines Self-attention and cross-attention in Fourier space mechanisms, creating class embeddings that improve contextual understanding and are passed on to the segmentation module [18].

Wang *et. al.,* suggested a DL prototype for the preprocessing and template protection using integration of a complete cancellable network of finger veins (CFVNet) that could be employed in the design a reliable identification of finger veins (FVR) system. In CFVNet, determination has an easy-to-use BWR-ROI Align unit that has three smaller modules: transformation, compression, and localization. The location sub-component has automated location to the steady and distinct ROI for finger veins [19]. Al-Obaidy *et. al.,* investigated a DL model, a convolutional neural network is trained using the Deep Fingers Vein Learning (DFVL) and enhance the recognition rate. The final model has three convolutional layers & Rectified Linear Unit (ReLU), classification, soft-max, pooling, and completely linked [20]. Zhang *et. al.,* proposed a hybrid DL model that is a highly accurate and highly efficient FVR. The Faster Multi-Scale model FVR (FMFVNet) network. The accuracy of the model is assured with a MSCA model, or multi-scale convolutional attention using subcutaneous vein patterns while optimizing recognition efficiency from the use of the FasterNet Block module. In FVR utilizes biometric technology by employing using near-infrared imaging to identify vein patterns beneath the skin. Finger vein biometrics is very secure, stable, and resistant to spoofing [21]. Liu *et. al.,* developed an approach to deep ensemble learning that is suggested to resolve the finger vein recognition Single Sample per Person (SSPP). However, the several feature maps were produced from an input image of the finger vein by using different autonomous DL-based on models through a collection of each of these modified classifiers [22].

Kyeremeh *et. al.,* suggested a novel synthesis of multimodal system for fingerprint and FV recognition and advanced some novel improvements to traditional methods. However, it is the Fast Library for Approximate Nearest Neighbors (FLANN) and Scale-Invariant Feature Transform (SIFT), with improvements based on preprocessing CLAHE stands for contrast limited adaptive histogram equalization and strong alignment of descriptors for features as part of the extraction and matching processes [23]. Mustafa *et. al.,* developed ML methodology, high-frequency Attention Using filters (HFAF), photos using a Gabor Filter (GF), and some augmented image approaches that preserve edges and enhance an image's quality. Finally, the authors selected two authentication models, one that has HPML, or hyperparameter machine learning and operates had KNN, SVM, and Multi-Layer Perceptron's (MLP) as a random search, and the second takes the same, but this time as a simple search. Najeeb *et. al.,* presented a novel DL model known as the Reinforced (RDL). This method creates a new personal verification method using the FV. However, the RDL is created with multiple layers, with feedback is provided. For each person, two FV fingers are used. For first personal verification, the Index finger FV is used, and for reinforced confirmation, the Middle finger FV is used. Further, the Personal verifications turn becoming essential need for supplying security with financial transactions and personal accounts.

2.1 Research Gap

Although several advanced Techniques for identifying finger veins have been proposed. to hybrid ML approaches DCNN, transfer learning, and multimodal biometric systems, there are many advanced methods for FVR; however, they still encounter significant limitations. For example, many systems are not accurate enough, are computationally expensive, and are subject to spoofing attempts. Although DL architectures perform better in feature extraction, their deployment into applications is often not realistic, due to the fusion of the outputs requiring organization approaches with dominant working, and reliance on high-end hardware. Further, most of the current systems utilize a single finger for the authentication process; therefore, the reliability and security are compromised. Eventually, there exists no low-end, secure, and highly accurate multi-finger authentication vein systems with a high ability to deter attempts against spoofing for web-based environments. This indicates that there needs to be a unique multi-finger vein biometric authentication with strong spoof-resistance, and with the combination of advanced enhancement, ensemble classification, and secure features fusion for web-based applications.

3 PROPOSED METHODOLOGY

The Secured Authentication for Web Application using Finger Vein (SAWFV) framework provides an effective and secure biometric verification system in web settings. The proposed solution begins with some

preprocessing steps, which consist of image resizing, median filtering for the purpose of noise reduction, and ROI extraction of the finger vein area. The clarity of images is then enhanced through an Enhanced Contrast Limited Adaptive Histogram Equalization (ECLAHE) approach that allows for adaptively enhancing small vascular details. To reduce risk from spoofing attacks, a deep ensemble spoof detection module that utilizes CNN, ResNet-18, and LSTM is incorporated utilizing Modified Score Fusion to fuse their respective outputs assigning global-local best inertia weighting to improve reliability of the decision. The finger vein images are then sent through a Bidirectional Efficient Transfer Learning (BETL) model which layers representations based on EfficientNet, along with representations that increase bidirectional attention, representing both spatial dependency and contextual dependency. Finally, a support vector machine (SVM) will classify the features of the finger vein images, distinguishing between genuine users and impostors. Overall, this entire integrated pipeline proposed in the framework will provide high authentication accuracy, high resiliency to spoofing attacks, and potential application to secure web environments that require near features. Overall flow of the suggested framework is as shown in Figure 1.

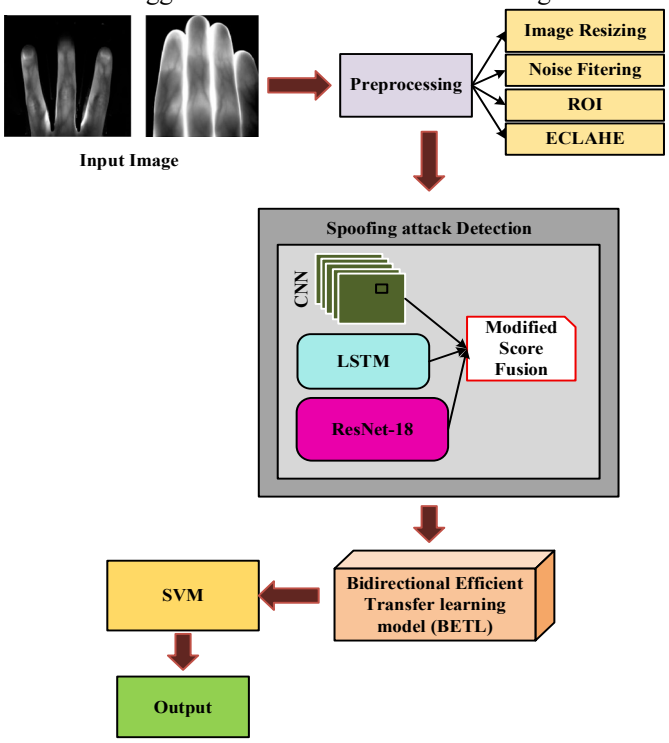


Figure 1 Overall flow of the proposed framework

3.1 Preprocessing

Preparation is an important step in the analysis of biometric images, as it serves to enhance the quality of the images and provides the features that makes extraction of features accurate and reliable. The work we proposed, includes stages of preprocessing which consist of resizing the image, denoising, Region of Interest (ROI) extraction, and Enhanced Contrast Limited Adaptive Histogram Equalization (ECLAHE) on the original image prior to extracting the features of engaged in a process that enhances the visibility and clarity of the finger vein structures.

3.1.1 Image Resizing

Image resizing is the process of altering a digital image's proportions by either making it larger (upscaling) or smaller (downscaling). In computer vision and image processing, this is a basic operation. A image dimensions are changed during the resizing process to satisfy particular specifications for image processing and computer vision jobs. Resizing an image entail altering its pixel count, which has an impact on the image's visual size as well as the quantity of data required to describe it. Resizing a digital image is known as image scaling. An image becomes smaller as it is scaled down, and larger when it is scaled up. Although they may both be scaled, raster and vector graphics yield distinct outcomes.

3.1.2 Noise Removal

Median filtering is a nonlinear noise-reduction method that replaces each pixel's median with immediate neighborhood for the pixel value. It performs especially good at removing impulsive noise, like pepper and salt noise, while preserving edge details. The filtered image is given by: $\hat{f}(x,y) = median\{g(x,y)\}$ (1)

where $\hat{f}(x,y)$ denotes the de-noised image and $g(x,y)$ the image contaminated by noise. Dynamic adjusting of filter window size will assist in overcoming variable noise density in effective elimination of the noise.

3.1.3 Region of Interest

A particular section of an image or dataset that has been chosen for in-depth examination is known as a "Region of Interest" (ROI). ROI helps draw attention to important aspects in a variety of domains, such as computer vision and medical imaging, increasing accuracy and efficiency.

3.1.4 Enhanced Contrast Limited Adaptive Histogram Equalization

A local contrast enhancement is called CLAHE. method created to increase the visibility of images, particularly in challenging conditions of low illumination or non-uniform illumination. To achieve this, the method divides the image into small overlapping tiles, equalized the histogram for each tile independently, and then interpolated between the equalized tiles using bilinear interpolation. Unlike AHE, CLAHE limits the degree of enhancement or over-enhancement through a clip limit. The clip limit is defined as:

$$\beta = \frac{M}{M\left(1+\frac{\alpha}{100}(S_{max}-1)\right)} \quad (2)$$

where β is the contrast limiting value, M represents the area size, α is the clip factor, S_{max} is the maximum allowed histogram slope that prevents sharp jumps in intensity.

To improve upon the CLAHE, we introduced a new modified weight function that takes into account the spatial relevance of pixels based on their distance from the center, as well as sharpness. The weight function is as follows:

$$\omega(i,j) = \frac{1}{Z(i)} exp\left(-\frac{d(i,j)}{h^2}\right)^n, (n \geq 2) \quad (3)$$

where $\omega(i,j)$ is the amount of weight attributed to the neighboring pixel concerning the center pixel i, $d(i,j)$ is the spatial distance between pixels $1\ and\ j$, n is the Sharpness parameter that controls the steepness of the exponential decay. $Z(i)$ Normalization factor ensuring that the sum of weights in the window is 1. The weighted equation of CLAHE equation becomes:

$$\beta' = \omega(i,j) \cdot \left[\frac{M}{M\left(1+\frac{\alpha}{100}(S_{max}-1)\right)}\right] \quad (4)$$

where β' is the updated contrast limiting value that incorporates the spatial weight $\omega(i,j)$, making the enhancement more adaptive to local pixel importance. Integrating the adjusted weight function to CLAHE enhances local contrast adaptively, giving priority to nearby and sharper pixels and allowing fine details such as finger vein patterns to be preserved. It will reduce noise amplification in uniform regions and edge preservation will be improved, while providing smoother transitions between tiles. The enhancement of the image would be more accurate, clearer, and visually consistent, which is particularly useful for biometric applications requiring detail and accuracy.

3.2 Spoofing Attack Detection

Spoofing attack detection is an important step in the finger vein biometric authentication system, where only true biometric input is accepted. The objective is to detect forgery or replay vein images that attackers typically use to defeat the security. The proposed architecture uses three strong deep learning models LSTM, CNN, and ResNet-18 are named as ensembling models, all with the same input image, will capture specific and different features from the input image in a specific order. The output of these different models is merged with a simplified version of Score Fusion, and yields a final decision on whether spoofing detection has occurred.

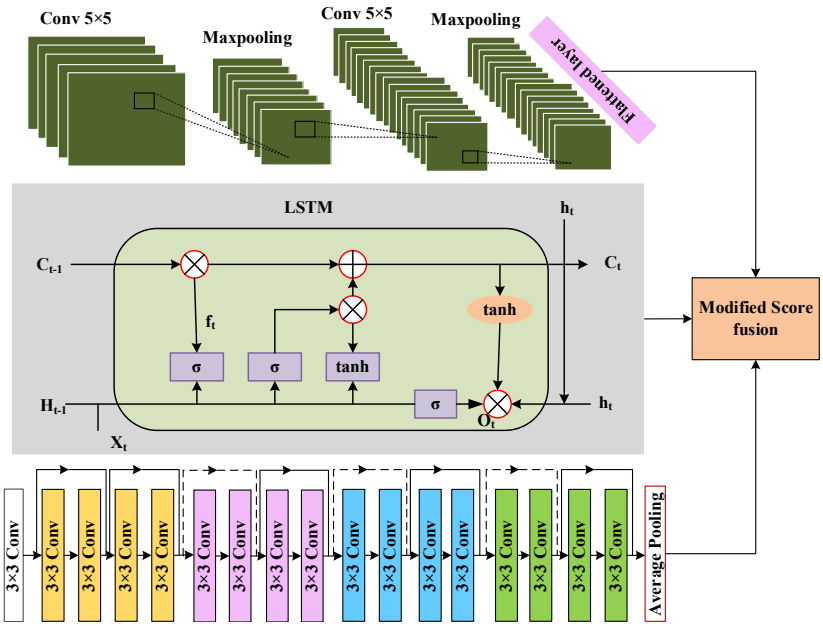


Figure 2 Ensembling architecture

Figure 2 shows the Ensembling architecture. In the architecture of the model, the first branch includes A Neural Network with Convolutions (CNN) to discover local spatial and texture features from pictures of veins in the fingers. The CNN includes two layers of convolution with 5×5 kernels, each with max pooling. The two convolutional layers learn to reduce spatial redundancy while preserving essential features; this way the model could learn fine structures of the veins and other surface level irregularities such as blurring or printing artefacts that can be present in spoofed samples. The convolution's output and max combining layers are pressed down and presented to a fully connected layer that produces a classification score for the spoof detection task.

The second ensemble model, the network known as Long Short-Term Memory (LSTM) is implemented in order

to learn temporal or consecutive dependencies between feature patterns. This is particularly relevant for biometric systems in which the input can be sequential (for example, line scans or time-varying imaging). LSTM networks are more sophisticated forms of neural networks that recur (RNNs) that can retain long-term reliance via internal memory units and gating. The LSTM takes a sequence of feature vectors from earlier CNN layers and learns how to remember useful patterns while ignoring non-informative durations in time. To implement it, the LSTM function is defined by a series of equations. The forget gate is defined as:

$$f_s = \sigma\left(W_f \cdot \left[h_{s-1}, x_s\right] + b_f\right) \quad (5)$$

$$i_s = \sigma\left(W_f \cdot \left[h_{s-1}, x_s\right] + b_i\right) \quad (6)$$

$$\widehat{C_s} = tanh\left(W_c \cdot \left[h_{s-1}, x_s\right] + b_o\right) \quad (7)$$

$$C_s = f_s \times C_{s-1} + i_s \times \widehat{C_t} \quad (8)$$

$$O_s = \sigma\left(W_o \cdot \left[h_{s-1}, x_s\right] + b_o\right) \quad (9)$$

$$h_s = O_t \times \tanh\left(C_s\right) \quad (10)$$

where h_{s-1} is the previous hidden state, x_s is the input at the time step s, and σ and $tanh$ are activation functions. These equations collectively allow the LSTM to focus on temporal behavior in the data, helping to detect subtle differences between real and fake vein images that evolve over time or across slices of the input.

The ResNet-18 is the third model used in the spoofing detection system and is a deep residual neural network with great performance for feature extraction. This moderate depth allows ResNet-18 to extract low-level spatial features in the earlier layers, and more abstract representations in deeper layers. The model comprises 16 convolutional layers, 2 down-sampling layers, and fully connected layers, with input images resized to 224964 pixels. The original convolution layer has a 7x7 kernel to extract macro features, in addition to multiple 3x3 convolution layers arranged in residual blocks. Each block increases to the input and takes two layers to capture maximum features, while reducing the occlusion of the vanishing gradient. Whenever the input and output of a layer have a different dimensionality, a dotted shortcut connection is used. After the last convolutional layer is a global average pooling layer that outputs a feature vector, and passes through a fully connected layer and Softmax activation to produce a probability distribution over classes indicating whether the image is 'real' or 'spoof'. ResNet-18 is confident in capturing subtle representation details such as continuity of vein patterns or texture consistency to optimize for improved spoofing detection. ResNet-18, when used in combination with the CNN and LSTM model, effectively explores spatial, temporal, and deep hierarchical feature representations. The outputs of each of the models are fused in the Modified Score Fusion module to produce one clear decision of spoof detection.

3.2.1 Modified Score Fusion

Modified Score fusion is used to fuse the outputs from CNN, LSTM, and a ResNet-18 classifier for spoofing attack detection. The score for individual classifiers shows the difference:

$$D_t = \frac{\sum_{q=1}^{C}(s[0] - s[q])}{C-1} \quad (11)$$

Where D_t is the decision score of the classifier, $s[q]$ is the score of the qth candidate, and C is the number of candidates. This is the average margin from the other scores, which indicates how confident the classifier is.

To improve this, a global-local best inertia weight is incorporated yielding the altered fusion score:

$$D'_t = \omega \cdot \left(\frac{\sum_{q=1}^{C}(s[0] - s[q])}{C-1}\right) \quad (12)$$

The inertia weight ω modulates the influence of each classifier's score based on global and local performance, and is defined as:

$$\omega = \left(1.1 - \frac{G_{best_i}}{P_{best_i}}\right) \quad (13)$$

where G_{best_i} is the global best score of the classifier i, and P_{best_i} is the local best score of the same classifier. This weight adaptively highlights classifiers that have both global and local trustworthy reliability. Using inertia weight helps to improve the adaptiveness of the fusion process and adds a higher priority to a classifier that continues to make accurate classifications, thus resulting in better spoof detection performance.

3.3 Feature Extraction and Finger Vein-Based User Authentication

After identifying a spoofing attack, the SAWFV framework enters the second stage to extract meaningful, deep, discriminative features from the finger vein image workflow to ultimately classify and verify legitimate users. During this stage, the goal is to learn a highly discriminative representation of images of finger veins images in the SAWFV system while ensuring the model is robust against environmental conditions. To accomplish this, a BETL model is adopted in our framework and built upon an EfficientNet architecture which has a multi-scale feature extractor design while utilizing fewer parameters for greater computing efficiency. Pruning and quantization are also performed to optimize performance while ensuring low computational and memory costs during model deployment.

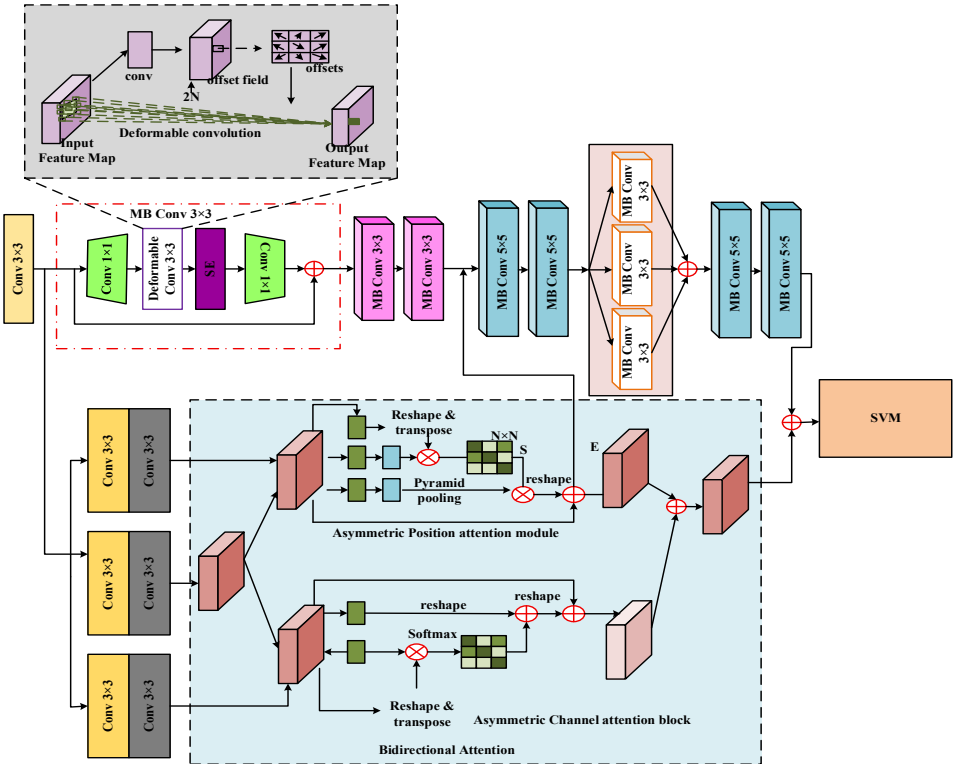


Figure 3 Architecture for Bidirectional Efficient Transfer learning model

The figure presents the architecture of the Bidirectional Efficient Transfer Learning (BETL) model. The structure is derived from EfficientNet, which utilizes Mobile Inverted Bottleneck Convolution (MBConv) blocks that comprise pointwise (1×1) convolution, depthwise (3×3) convolution, a squeeze and excitation (SE) module, and expansion layers. In this study, the original depthwise convolution is replaced with deformable convolution across all MBConv blocks, which allows the network to adaptively learn sampling positions and adapt to positional shifts, and distortions of finger vein patterns. Furthermore, the MBConv blocks are arranged in parallel branches as opposed to sequential stacking, retaining and improving feature diversity and speed of training and learning fine-grained vein texture.

Additionally, the model integrates a second feature extraction pathway after the first convolutional layer, which includes three convolution blocks, with BAMs, which stands for Bidirectional Attention Module, following each block. BAM includes two submodules, the Asymmetric Position Attention Module (APAM) incorporating pyramid pooling to learn long-range spatial dependencies, and the Asymmetric Channel Attention Block (ACAB) applying softmax weighting based on feature channel importance. The enhanced specific channel attention outputs are accumulated with the MBConv layers to improve feature representation. Finally, a single feature vector is formed by concatenating the EfficientNet-based and bi-directional attention features and classified with Support Vector Machine (SVM) for user authentication. The BETL model employs transfer learning, deformable convolutions, bidirectional attention, and channel-location aware fusion for highly accurate, spoof resistant, and environment robust finger vein-based authentication.

3.3.1 Deformable Convolution

The BETL model integrates Deformable Convolution (DConv) into every MBConv blocks, making the network more capable of learning vein patterns that may be non-rigid and curved, with distortions in the spatial domain. While ordinary convolution operates on a fixed sampling grid location (i.e., 3×3) for convolution, DConvolution can accept learnable offset parameters to allow shift to sampling points rather than have specific locations and the use of their filtering. Figure 4 shows the Deformable Convolution.

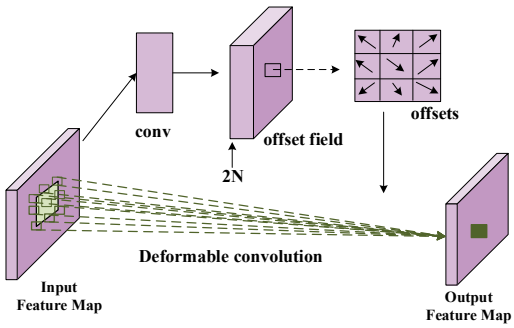


Figure 4 Deformable Convolution

In standard convolution, the output at a position p_0 is calculated as:

$$y(p_0) = \sum_{p_n \in \mathcal{R}} w(p_n) \cdot x(p_0 + p_n + \Delta p_n) \quad (14)$$

In DConv, this becomes:

$$y(p_0) = \sum_{p_n \in \mathcal{R}} w(p_n) \cdot x(p_0 + p_n + \Delta p_n) \quad (15)$$

where the input feature map is denoted by x. $w(p_n)$ is the convolution kernel, Δp_n are the learnable offsets for each location in the kernel.

Since Δp_n can be fractional, bilinear interpolation is used to compute values at non-integer positions:

$$x(p) = \sum_q G(q,p) \cdot x(q), \quad (16)$$

$$G(q,p) = g(q_x,p_x) \cdot g(q_y,p_y) \quad (17)$$

where $g(a,b) = \max(0,1-|a-b|)$

The deformable convolution layers are able to make the network actively change its sampling framework and increase the robustness of the feature extraction against misalignments where the finger is not positioned correctly, marginal rotations, and naturally occurring deformation in the vein textures. While incorporating deformable convolution as Depthwise convolution in our modified EfficientNet pipeline is not only leverageable for flexibility in changing the sampling framework of the deceived convolution operation, also it results in more accurate and reliable feature maps for high-security finger vein authentication.

3.3.2 Bi-directional Attention Module

To enhance the feature representation capability of the proposed Bi-directional Efficient Transfer Learning (BETL) model, we embed a Bidirectional Attention Module (BAM) comprising two essential elements: the Asymmetric Position Attention Module (APAM) and the Asymmetric Channel Module of Attention (ACAM). These modules are strategically designed to extract robust spatial and channel-wise contextual dependencies from pictures of the veins in the fingers, so improving the quality of extracted features and improving the performance of the model in spoof detection and authentication.

Asymmetric Position Attention Module (APAM)

Give the input feature map the following notation: $F \in \mathbb{R}^{C\times H\times W}$, where C, H, and W stand for the breadth, height, and number of channels, respectively. The input is first processed through pyramid pooling to capture multi-scale spatial contexts. Three convolutional layers are applied to produce query, key, and value maps:

$$Q = W_Q F, K = W_K F, V = W_V F \quad (18)$$

These are reshaped for matrix multiplication:

$$Q' \in \mathbb{R}^{G\times N}, K' \in \mathbb{R}^{G\times N}, V' \in \mathbb{R}^{G\times N},$$
$$\text{where } N = H \times W \quad (19)$$

The position attention map $S \in \mathbb{R}^{N\times N}$ is computed using dot-product attention:

$$S = Softmax(Q^T K') \quad (20)$$

This attention map captures long-range spatial dependencies. The output of the position attention is:

$$E_p = V'S \quad (21)$$

This is reshaped back to original dimensions and fused with the original feature map:

$$F_p = \gamma E_p + F \quad (22)$$

Here, γ is a learnable parameter that scales the attention output.

Asymmetric Channel Attention Block (ACAB)

Simultaneously, the channel-wise relationships are captured using the Asymmetric Channel Attention Block. Starting from the same input F, it is reshaped to $F_c \in \mathbb{R}^{G\times N}$. Then, channel-wise query and key matrices are formed:

$$Q_c = F_c, K_c = F_c^T \quad (23)$$

The channel affinity matrix is computed as:

$$S_c = Softmax(Q_c K_c) \quad (24)$$

This produces $S_c \in \mathbb{R}^{C\times C}$, representing channel dependencies. The attention-modulated feature is:

$$E_c = S_c F_c \quad (25)$$

Reshaping and residual fusion yield the output:

$$F_c' = \alpha \cdot reshape(E_c) + F \quad (26)$$

where α is another learnable scalar. The final output feature map F_{out} from the Bidirectional Attention Module is the combination of both attention-enhanced features:

$$F_{out} = \delta(F_p + F_c') \quad (27)$$

where δ denotes a non-linear transformation (e.g., ReLU or normalization), used to stabilize learning in downstream transfer learning tasks. Enhanced bidirectional attention with pyramid pooling reinforces both spatial and channel representations. This is especially important in multi-indicator shift learning as pretrained features can often be weak at identifying fine-grained patterns or contextual information. Including this module in a backbone model will help to refine features and improve generalization to downstream tasks, such as image segmentation, classification, or detection. In the proposed framework, the attention-enhanced output F_{out} is concatenated with the general EfficientNet feature maps, passed to the classification stage, thus allowing the model to leverage global efficiency and attention from its concatenated input in order to make informed decisions.

3.4 Support Vector Machine (SVM) for Finger Vein Classification

In the last stage of the proposed framework, a Support Vector Machine (SVM) classifier is used for verification of users, based on the highly discriminative features established from the BETL model, which are a combination of attention-refined and EfficientNet representations. The combination of these representations provides a robust characterization of user-specific vein pattern for classification. SVM is a powerful supervised learning algorithm that works exceptionally well in high-dimensional spaces, and is also appropriate for binary classification. It accurately classifies data that has a non-linear relationship by first transforming the data to a higher dimensional space, and then it computes an optimal hyperplane that provides maximum separation between the genuine and imposter classes providing reliable authentication for users:

$$\omega \cdot \varphi(x) + b = 0 \quad (27)$$

where x is the input feature vector, $\varphi(x)$converts the input into a higher-dimensional feature space, ω is the vector of weight, and b is the phrase for bias.

For non-linearly separable cases, the SV M introduces slack variables ξ_i and solves the following convex optimization problem:

$$\min_{w,b} \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{m}\xi_i \quad (28)$$

$$y_i(\omega \cdot \emptyset(x_i) + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, i = 1, \ldots, m \quad (29)$$

To handle non-linearity, SVM uses kernel functions $K(x_i, x_j) = \emptyset(x_i) \cdot \emptyset(x_j)$, allowing it to operate in the transformed feature space without explicitly computing $\emptyset(x)$. The dual problem becomes:

$$\max_{a} \sum_{i=1}^{m}\alpha_i - \frac{1}{2}\sum_{j=1}^{m}\sum_{k=1}^{m}\alpha_j\alpha_k\alpha_j\alpha_k \quad K(x_j, x_k) \quad (30)$$

$$\sum_{i=1}^{m} y_i\alpha_i = 0, 0 \leq \alpha_i \leq C \quad (31)$$

In our system, the refined feature vectors representing finger vain patterns are input into the SVM. It will then classify whether the input sample is from a registered (genuine) user, or whether it is an imposter. The SVM is inherently robust to small sample sizes and has high efficiency with high-dimensional data, which is why it is a good option for the biometric authentication task since it requires precise class separation.

4 RESULTS AND DISCUSSION

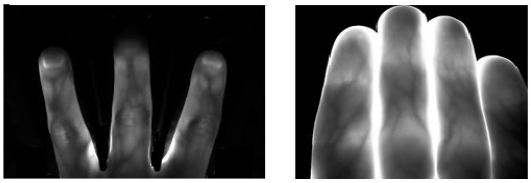
The reliability of the SAWFV system was determined using a variety of measurements to examine both accuracy and efficiency with biometric user authentication. Principal measurements of classification included False Positive Rate (FPR), False Negative Rate (FNR), sensitivity, accuracy, and precision, which illustrated how the model reliably authenticated an honest user and rejected imposters. The performance characteristics of image quality and processing were determined using Euclidean distance, Hamming distance, correlation coefficient, entropy, exposure time, and frame rate. All these metrics showed the quality of improved images for finger veins and performance characteristics of SAWFV to be robust, trustworthy, and capable of being a web-based user authentication solution.

4.1 Dataset Description

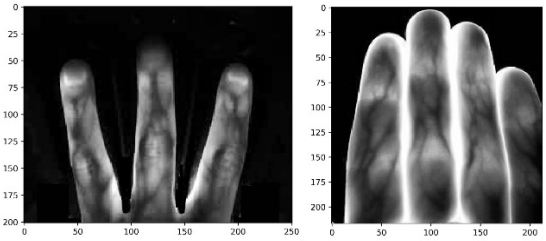
Publicly available finger vein datasets MMCBNU_6000 and PLUSVein-FV3 https://wavelab.at/sources/Drozdowski20a/ are popular datasets for evaluating finger vein recognition algorithms. The MMCBNU_6000 dataset provides images of six fingers (the middle, ring, and index fingers from both hands) of 100 volunteers and provides a sufficient dataset for processing as biometric data. The PLUSVein-FV3 dataset contains palmar and dorsal finger vein images of 360 fingers collected from 60 subjects, collected in the same session, with five samples collected for each finger. The PLUSVein-FV3 dataset uses two types of sensors: one that uses NIR laser modules, and the other one that uses NIR LEDs, that may equally help provide variation in imaging conditions across samples. These datasets will help assess the performance and fairness of vein recognition systems for fingers.

4.2 Performance Evaluation

The SAWFV framework was compared with existing models of CNN, RNN, LSTM, and ResNet-18 to assess its theoretical feasibility for using secure finger vein-based user authentication. The model was assessed based on available traditional experimental metrics, including Accuracy, Precision, Sensitivity, FPR, FNR, euclidean and Hamming Distance, Correlation Coefficient, Entropy, Exposure Time, and Frame Rate. The outcomes revealed that the SAWFV framework did better than all of the compared models, when the above metrics were reported, through higher accuracy results, image quality, lower error rates and faster processing times, and provide an opportunity for secure and web applications.



(a) Input Image



(b) Pre-processed Image

Figure 5 Representation of input and pre-processed image

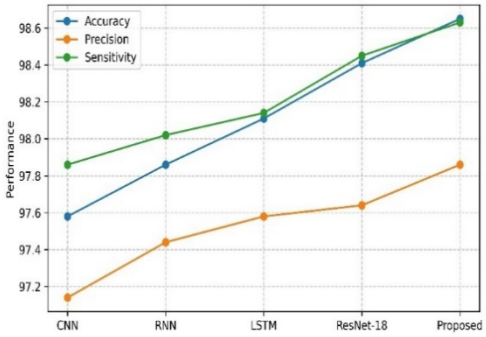
Figure 5 shows the effects of preprocessing on a picture of a finger vein. In subfigure (a), the raw input picture with low contrast and noise. In subfigure (b), a pre-processed image with enhancements made with ECLAHE - median filtering and extraction of the region of interest. This pre-processed image allows for clearer visible vein patterns leading to increased recognition accuracy.

Table 2 Evaluation of performance for proposed and existing methods

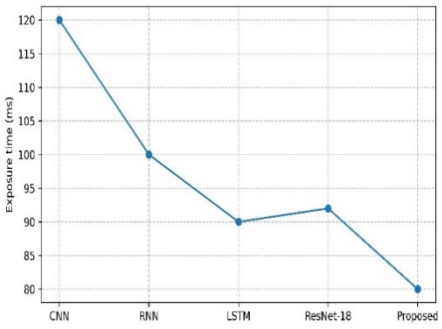
Metrices	CNN	RNN	LSTM	ResNet-18	Proposed
Exposure time (ms)	120	100	90	92	80
Frame rate (fps)	25	30	35	38	40
Euclidean distance	0.45	0.38	0.33	0.28	0.25
Hamming distance	0.36	0.29	0.22	0.18	0.15
Correlation coefficient	0.82	0.85	0.88	0.91	0.92
Entropy	5.1	5.4	5.7	6.1	6.2
Accuracy	97.58	97.86	98.11	98.41	98.65
Precision	97.14	97.44	97.58	97.64	97.86

Sensitivity	97.86	98.02	98.14	98.45	98.63
FNR	0.025	0.023	0.022	0.021	0.018
FPR	0.047	0.045	0.044	0.041	0.038

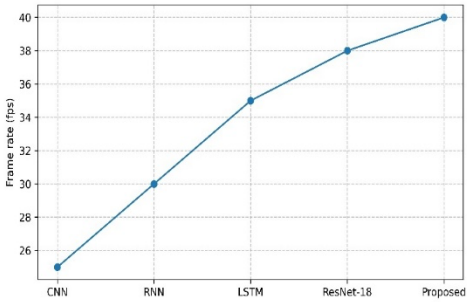
Table 2 provides a complete performance comparison against the CNN, RNN, LSTM, and ResNet-18 models with regard to metrics of interest. As observed, the proposed method has performed better, as it achieves the lowest exposure time (80ms) and the highest recognition frame rate (40fps) among all techniques that were studied. Additionally, the proposed method matched with greater accuracy, achieving the lowest Euclidean distance (0.25) and Hamming distance (0.15), and the highest correlation coefficient (0.92) and entropy (6.2), meaning that the proposed method represented richer features. The classification metrics further support this claim, demonstrating the highest overall accuracy (98.65%), precision (97.86%), and sensitivity (98.63%), with very low False Negative Rate (0.018) and False Positive Rate (0.038). In conclusion, the proposed technique has superior security, accuracy, and efficiency, and is an effective solution for user authentication based on finger vein patterns.



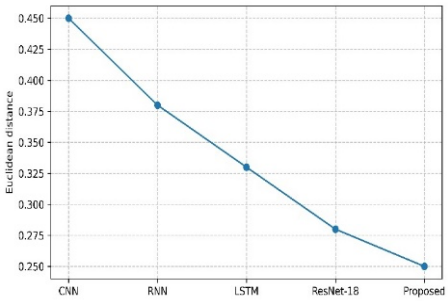
(a) accuracy, Precision, Sensitivity



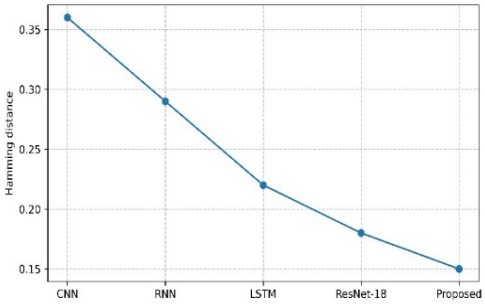
(b) Exposure Time



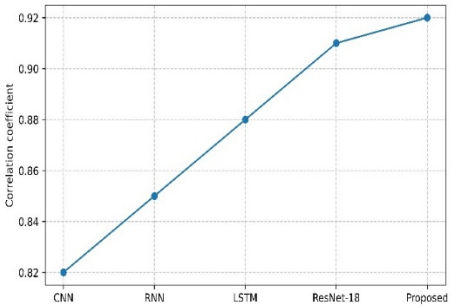
(c) Frame Rate



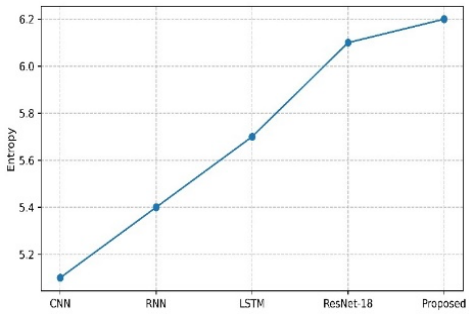
(d) Euclidean distance



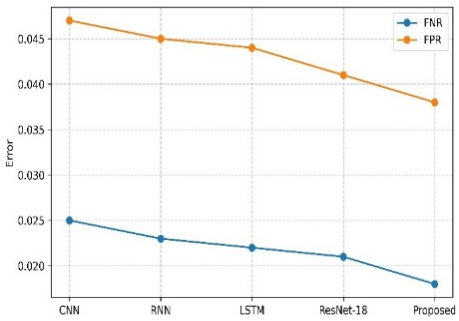
(e) Hamming distance



(f) Correlation Coefficient



(g) Entropy



(h) FPR, FNR

Figure 6 (a)-(h) Visual depiction of several performance indicators using both suggested and current methods

The performance indicators, which consist of (a) precision, accuracy, Sensitivity, (b) Exposure Time, (c) Frame Rate, (d) Euclidean distance, (e) Hamming distance, (f) Correlation Coefficient, (g) Entropy, and (h) FPR, FNR are represented graphically in Figures 6 (a) to (h).

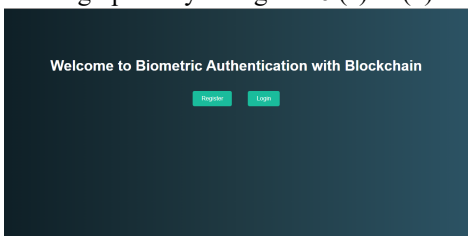


Figure 7 Home Page

The homepage of the biometric authentication system, portrayed in Figure 7, consists of a simple page presenting the opportunity to register or login to the system. The simple layout is clear and easy to navigate, presenting pathways to register or login directly to the users.

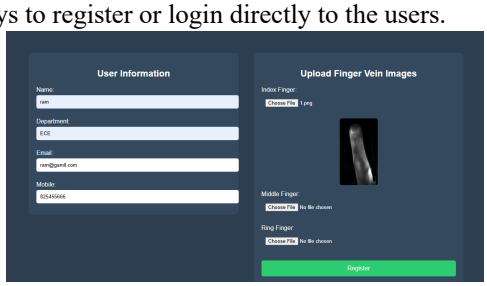


Figure 8 Register page

Figure 8 depicts the enrollment interface which allows users to input some basic information about themselves as well as some pictures of their finger veins (the middle, ring, and index fingers). Users are also allowed to preview their images before they submit before they can complete the enrollment. This is an important step in the enrollment and capture of biometric images.

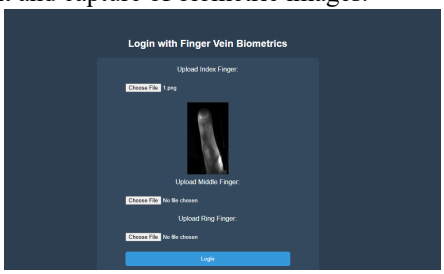


Figure 9 Login Page

Figure 9 demonstrates the login page where the registered user submits their three finger vein images for authentication. This input is verified through the SAWFV framework, which involves spoofing detection and feature classification.

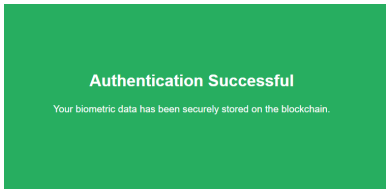


Figure 10 Authentication Page

Figure 10 shows the successful authentication and that the user's biometric data was satisfactorily verified and recorded to the blockchain. The message also emphasizes the transparency of the system, as well as the integrity of the biometric data.

5 CONCLUSION

The SAWFV framework design provides a secure and cost-efficient user authentication mechanism for web applications through the use of finger vein biometrics. The proposed SAWFV framework implements advanced preprocessing through an innovative ECLAHE, which enhances vein visibility while preserving finer details. Additionally, a spoof detection module combining CNN, ResNet-18, and LSTM with a Modified Score Fusion method based on global-local best inertia weights provides reliability. The inclusion of a BETL model increases spatial and channel-level feature extraction, which allows better generalization and adaptability to multiple finger submissions. The final classification is performed through SVM with a high accuracy (98.65), precision (97.86), and very low false positive (0.038) and false negative (0.018) rates. The system was implemented in Python and shows great promise for secure online authentication. Moving forward, there is work to be done on multi-modal biometric integration, as well as larger datasets for greater robustness and scalability.

DECLARATIONS:

Funding

On Behalf of all authors the corresponding author states that they did not receive any funds for this project.

Conflicts of Interest

The authors declare that we have no conflict of interest.

Competing Interests

The authors declare that we have no competing interest.

Data Availability Statement

The data used in this study were derived from simulation reports and publicly available finger vein datasets MMCBNU_6000 and PLUSVein-FV3 accessible at https://wavelab.at/sources/Drozdowski20a/. The authors are also working toward implementing the approach using real world data with necessary permissions.

REFERENCE

[1] Olanrewaju, R.F., Khan, B.U.I., Morshidi, M.A., Anwar, F. and Kiah, M.L.B.M., 2021. A frictionless and secure user authentication in web-based premium applications. *Ieee Access*, *9*, pp.129240-129255.

[2] Shethiya, A.S., 2025. Scalability and Performance Optimization in Web Application Development. Integrated Journal of Science and Technology, 2(1).

[3] Shethiya, A.S., 2025. Building Scalable and Secure Web Applications Using. NET and Microservices. Academia Nexus Journal, 4(1).

[4] Gugała, Ł., Łaba, K. and Dul, M., 2023. Protecting web applications from authentication attacks. Advances in Web Development Journal, 1.

[5] Kolluri, V., 2024. Cybersecurity Challenges in Telehealth Services: Addressing the security vulnerabilities and solutions in the expanding field of telehealth. International Journal of Advanced Research and Interdisciplinary Scientific Endeavours, 1(1), pp.23-33.

[6] Mekruksavanich, S. and Jitpattanakul, A., 2021. Biometric user identification based on human activity recognition using wearable sensors: An experiment using deep learning models. Electronics, 10(3), p.308.

[7] Maurya, S., Joseph, S., Asokan, A., Algethami, A.A., Hamdi, M. and Rauf, H.T., 2021. Federated transfer learning for authentication and privacy preservation using novel supportive twin delayed DDPG (S-TD3) algorithm for IIoT. Sensors, 21(23), p.7793.

[8] Malhotra, P., Singh, Y., Anand, P., Bangotra, D.K., Singh, P.K. and Hong, W.C., 2021. Internet of things: Evolution, concerns and security challenges. Sensors, 21(5), p.1809.

[9] Kyeremeh, G.K., Abdul-Al, M., Qahwaji, R. and Abd-Alhameed, R.A., 2024. Verification technology for finger vein biometric. arXiv preprint arXiv:2405.11540.

[10] Rehman, A., Harouni, M., Omidiravesh, M., Fati, S.M. and Bahaj, S.A., 2022. Finger Vein Authentication Based on Wavelet Scattering Networks. Computers, Materials & Continua, 72(2).

[11] Kolivand, H., Akintoye, K.A., Asadianfam, S. and Rahim, M.S., 2023. Improved methods for finger vein identification using composite Median-Wiener filter and hierarchical centroid features extraction. Multimedia Tools and Applications, 82(21), pp.31913-31944.

[12] Alay, N. and Al-Baity, H.H., 2020. Deep learning approach for multimodal biometric recognition system based on fusion of iris, face, and finger vein traits. Sensors, 20(19), p.5523.

[13] Haider, S.A., Ashraf, S., Larik, R.M., Husain, N., Muqeet, H.A., Humayun, U., Yahya, A., Arfeen, Z.A. and Khan, M.F., 2023. An improved multimodal biometric identification system employing score-level fuzzification of finger texture and finger vein biometrics. Sensors, 23(24), p.9706.

[14] Khodadoust, J., Medina-Pérez, M.A., Monroy, R., Khodadoust, A.M. and Mirkamali, S.S., 2021. A multibiometric system based on the fusion of fingerprint, finger-vein, and finger-knuckle-print. Expert Systems with Applications, 176, p.114687.

[15] Ma, H., Hu, N. and Fang, C., 2021. The biometric recognition system based on near-infrared finger vein image. Infrared Physics & Technology, 116, p.103734.

[16] Shakil, S., Arora, D. and Zaidi, T., 2023. An optimal method for identification of finger vein using supervised learning. Measurement: Sensors, 25, p.100583.

[17] Al-Obaidy, N.A.I., Mahmood, B.S. and Alkababji, A.M.F., 2022. Finger Veins Verification by Exploiting the Deep Learning Technique. *Ing. Syst. Inf*, *27*, pp.923-931.

[18] Liu, Y., Yang, W. and Liao, Q., 2024. Diffvein: A unified diffusion network for finger vein segmentation and authentication. IEEE Transactions on Circuits and Systems for Video Technology.

[19] Wang, Y., Gui, J., Tang, Y.Y. and Kwok, J.T., 2024. CFVNet: An end-to-end cancelable finger vein network for recognition. IEEE Transactions on Information Forensics and Security.

[20] Al-Obaidy, N.A.I., Mahmood, B.S. and Alkababji, A.M.F., 2022. Finger Veins Verification by Exploiting the Deep Learning Technique. Ing. Syst. Inf, 27, pp.923-931.

[21] Zhang, Z., Liu, P., Su, C. and Tong, S., 2025. A High-Speed Finger Vein Recognition Network with Multi-Scale Convolutional Attention. Applied Sciences, 15(5), p.2698.

[22] Liu, C., Qin, H., Song, Q., Yan, H. and Luo, F., 2023. A deep ensemble learning method for single finger-vein identification. Frontiers in Neurorobotics, 16, p.1065099.

[23] Kyeremeh, G.K., Abdul-Al, M., Qahwaji, R., Ali, N.T. and Abd-Alhameed, R.A., 2025. Fusion of Hand Biometrics for Border Control Involving Fingerprint and Finger Vein. IEEE Access.

[24] Mustafa, R.A. and Abbes, T., 2025. An Improvement Finger Vein Authentication System Based on PCANet Deep Learning and Hyper Parameter Machine Learning. Iraqi Journal for Computer Science and Mathematics, 6(2), p.6.

[25] Najeeb, S.M.M., Al-Nima, R.R.O. and Al-Dabag, M.L., 2021. Reinforced Deep Learning for Verifying Finger Veins. International Journal of Online & Biomedical Engineering, 17(7).