ECTI Transactions on Computer and Information Technology
Journal homepage: https://ph01.tci-thaijo.org/index.php/ecticit/
Published by the ECTI Association, Thailand, ISSN: 2286-9131

Enhanced Hand Vein Segmentation Using Generative Adversarial Network Integrated with Modified ECA Module

Marlina Yakno[1], Mohd Zamri Ibrahim[2], Muhammad Salihin Saealal[3], Norasyikin Fadilah[4] and Wan Nur Azhani W. Samsudin[5]

ABSTRACT

Hand vein image segmentation is crucial for diverse applications such as precise biometric identification and facilitating medical intravenous procedures. This paper introduces an enhanced hand vein image segmentation method utilizing deep learning, specifically a conditional generative adversarial network (cGAN). The cGAN is trained adversarially and augmented with a modified efficient channel attention (ECA) mechanism module. The efficiency of the proposed technique was evaluated using four hand vein datasets: self-acquired dataset, SUAS, WILCHES, and BOSPHORUS. Performance comparison reveals that the proposed method outperforms alternative approaches, achieving the best results across all datasets with an average sensitivity of 0.8878, average accuracy of 0.9639, and average dice coefficient of 0.7904 for vein patterns. Our experimental findings demonstrate that the proposed segmentation technique significantly enhances hand vein patterns and improves dorsal hand vein detection accuracy.

Article information:

Keywords: Attention Mechanism Module, Efficient Channel Attention, Generative Adversarial Network, Vein Segmentation

Article history:
Received: October 28, 2024
Revised: January 16, 2024
Accepted: February 20, 2025
Published: March 8, 2025 (Online)

DOI: 10.37936/ecti-cit.2025192.259390

1. INTRODUCTION

Vein viewing technology is an advanced imaging technique used to visualize blood vessels beneath the skin, primarily for medical applications such as intravenous injections and blood draws. Vein identification involves image acquisition, pre-processing, enhancement, segmentation, and pattern extraction. Among these, vein segmentation is a critical step that isolates vein structures from the surrounding tissue in the acquired image [1]–[3].

Researchers have developed various vein segmentation techniques to address this challenge. The vein identification process involves several stages: image acquisition, ROI extraction, image enhancement, and vein segmentation [4], [5]. Vein segmentation aims to distinguish vein images from non-vein regions and researchers employ various techniques to address this challenge, including thresholding [6], kernel filtering [7], model-based [8], and curvature-based methods [9]. Despite their simplicity and popularity, these techniques frequently produce noisy vein images. Furthermore, these segmentation techniques require various combinations of pre- and post-processing methods.

Recently, Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs) have gained prominence in vein segmentation, offering notable advantages over traditional methods. These deep-learning approaches have transformed medical image analysis, particularly in vein segmentation tasks. CNNs excel in automatically learning hierarchical features from raw image data, eliminating the need for manual feature extraction. Meanwhile, GANs help overcome the challenge of limited labeled training data by generating realistic synthetic vein images that expand the training dataset, enabling more robust model training without requiring additional manual annotations [10]–[15].

In their work, Marattukalam & Abdulla [10] proposed a U-Net architecture for palm vein segmentation, evaluating its performance on the HK PolyU Multispectral Palm Print database. Their approach

[1,2,4,5]The authors are with the Faculty of Electrical and Electronics Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, Pahang, Malaysia, Email: marlinayakno@umpsa.edu.my, zamri@umpsa.edu.my, norasyikin@umpsa.edu.my and nurazhani@umpsa.edu.my
[3]The author is with the Faculty of Electrical Technology Engineering, Universiti Teknikal Malaysia Melaka, Durian Tunggal, Melaka, Malaysia, Email: salihin@utem.edu.my
[1]Corresponding author: marlinayakno@umpsa.edu.my

achieved high dice coefficients, highlighting its effectiveness in segmenting palm veins. Similarly, Lefkovits & Emerich [11] introduced a modified U-Net architecture for dorsal hand vein segmentation, employing a two-stage approach. In the first stage, nine variants of the U-Net model were utilized, with outputs combined through a majority voting mechanism. The system subsequently feeds these outputs into a second stage modified U-Net for refined segmentation. The method, evaluated on the NCUT dorsal hand vein database, demonstrated a high Dice coefficient, highlighting its segmentation accuracy.

In another study, Kuthiala *et al.* [12] proposed a U-Net architecture with additional skip connections for arm vein segmentation, validating their approach on a custom dataset they created specifically for this purpose. Applying basic pre-processing techniques and image augmentation improved segmentation performance, which the researchers evaluated using the dice coefficient and intersection over union (IoU) metrics. Similarly, He *et al.* [13] developed an adapted U-Net model for forearm vein segmentation. To address the challenge of limited data availability, the researchers implemented extensive dataset augmentation, including rotation, flipping, contrast adjustment, and elastic deformations. The resulting model exhibited exceptional performance across various evaluation metrics, including sensitivity, specificity, accuracy, F1-score, and area under the curve (AUC).

Additionally, Gao *et al.* [16] introduced a U-Net architecture enhanced with residual connections specifically designed for dorsal hand vein segmentation, validating their approach on a custom dataset of near-infrared hand images. This improved model achieved higher accuracy and F1 scores in dorsal hand vein segmentation compared to traditional methods. Subsequent refinement of this U-Net model achieved an exceptionally high Dice coefficient, further validating the effectiveness of U-Net-based frameworks for vein segmentation applications.

While convolutional neural network (CNN) architectures, particularly U-Net, and its derivatives, have been widely applied for vein segmentation, generative adversarial networks (GANs) are increasingly utilized in this domain. For instance, Yang *et al.* [17] proposed a GAN-based approach for segmenting finger veins, evaluating their method on the THU-FVFDT2 and SDU finger vein datasets. Notably, their GAN-based technique outperformed traditional methods and CNN models in finger vein recognition, achieving the lowest equal error rate (EER).

Despite their effectiveness, both CNN and GAN architectures consist of multiple layers, which can result in information loss during down-sampling and up-sampling processes. This loss of spatial information and image details can negatively impact the network's performance. Researchers commonly integrate attention mechanism modules into these models to address this challenge.

These attention mechanisms enhance model performance by focusing on key features within the input data and de-emphasizing less important ones. This capability allows the network to maintain essential image details, even with multiple convolutional and pooling layers. Integrating attention modules into CNN or GAN models readily improves their effectiveness. Some widely used attention mechanisms include Squeeze-and-Excitation (SE), Convolutional Block Attention Module (CBAM), and Efficient Channel Attention (ECA).

In their study, Abramova *et al.* proposed incorporating SE modules into U-Net for segmenting hemorrhagic stroke lesions. This resulted in significant performance enhancements. However, this improvement came at the cost of additional computational resources. Researchers then sought to improve performance while minimizing computational overhead. Therefore, to address this challenge, Luan *et al.* [19] introduced CBAM for liver tumor segmentation, further enhancing results. Despite the improvements, the dual attention mechanism in CBAM increased computational demands. In an effort to optimize the balance between computational complexity and performance, Wang *et al.* [20] developed the Efficient Channel Attention (ECA) module. Unlike SE, the ECA module replaces the fully connected layer with a more efficient convolutional layer. The ECA module employs a 1D convolution operation to calculate channel-wise attention weights, which are then applied element-wise to the feature maps to produce the final output. Fig. 1 illustrates the operational process of the ECA module.

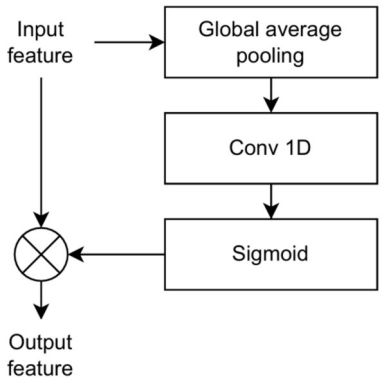

Fig.1: *Block diagram of ECA Module.*

Researchers subsequently employed ECAU-Net [21] for the precise identification and segmentation of the cerebellum. This approach integrates the advantages of the ECA module and U-Net architecture. Researchers conducted comprehensive experiments using the Dice loss function on the JSUAH-Cerebellum dataset, and the results demonstrated superior performance compared to other U-Net variants.

Furthermore, researchers propose the ECA-TFUnet [22] model as a U-shaped CNN-Transformer network with Efficient Channel Attention for automated organ segmentation in anatomical sectional images of canines. The model integrates the U-Net architecture for capturing fine local details, the Transformer block for learning global dependencies, and a mixed loss strategy to address class imbalance. Experimental results indicate that ECA-TFUnet outperforms 11 state-of-the-art methods, achieving an Intersection over Union (IoU) of 92.63% on a custom dataset and 87.93% on a public dataset.

Despite the substantial enhancements in segmentation results provided by the ECA module, it primarily encodes channel relevance features using global average pooling (GAP) signals. While GAP reduces feature map dimensionality by condensing it into a single value, this process may lead to the loss of discriminative information. As a result, incorporating an additional component that preserves the maximum value of each feature map could help retain more discriminative information.

Furthermore, the conventional convolution used in current ECA modules is limited to a fixed receptive field size, which may inadequately capture features across various scales. This limitation hinders detecting long-range dependencies within the input image and overlooks relationships between distant pixels. Such constraints can be particularly problematic in tasks like segmentation, where capturing spatial dependencies among objects is crucial. Therefore, a solution is urgently needed to address these specific limitations.

This work proposes several key contributions to overcome the challenges outlined above: (1) an enhancement to the existing Efficient Channel Attention (ECA) module by incorporating global maximum pooling (GMP) and dilated convolution, which facilitates the acquisition of critical information while preserving discriminative features; (2) the integration of these modified ECA modules into a Generative Adversarial Network (GAN) architecture to improve vein segmentation performance; (3) a comparative evaluation of the best modified ECA module against three widely used attention mechanisms: Squeeze-and-Excitation (SE), Convolutional Block Attention Module (CBAM), and the original ECA module; and (4) a comprehensive assessment of the proposed method using hand vein databases, measuring key performance metrics.

This paper organizes its remaining sections as follows. Section 2 provides a detailed introduction to the proposed technique, including the hand vein databases utilized and the performance metrics employed. In conclusion, Sections 3 and 4 present the experimental results and concluding remarks.

2. METHODOLOGY

This section covers data preparation, the cGAN segmentation model, and the three modified ECA modules integrated into the architecture.

2.1 Pre-processing and Dataset Preparation

This study utilized low-cost infrared image acquisition methods, as previously proposed in [23]. We gathered a total of 114 near-infrared hand vein images from four databases: self-acquired, SUAS [24], WILCHES [25], and BOSPHORUS [26]. The self-acquired database consisted of images captured under controlled conditions, representing two hand poses (open and clenched) and two different hand locations (dorsal and wrist).

To acquire images, volunteers positioned their hands in front of an infrared camera. For the open-hand pose, researchers captured each right and left hand for three seconds. For the clenched hand pose, researchers followed the same three-second imaging procedure for both hands. Finally, for the wrist pose, volunteers made a fist with their right and left hands, with images captured for three seconds. Researchers captured three grayscale images with a resolution of 640 $\times$ 480 pixels for each hand position and saved them in JPG format.

To enhance vein detection accuracy and efficiency, Faster RCNN was used for vein localization, as in prior work [27], to identify the most prominent vein features in the hands. The size of the region of interest (ROI), representing the vein, varied based on the subject's hand size. To enhance contrast after ROI extraction, a fusion of CLAHE and adaptive gamma correction, as proposed in [28], was applied. Then, the improved ROI was resized to 192 $\times$ 192, which was then used as input to the cGAN segmentation model.

The training set consisted of 60 images from the self-acquisition and SUAS datasets, out of a total of 114 hand vein images. Since hand vein databases do not provide ground truth images for direct comparison, we created a crucial set of ground truth images for this work. Due to the limited number of training images, we applied an online data augmentation technique, which expanded the training dataset to a total of 1482 images. For validation, we set aside 10% of the augmented training set.

The test dataset comprised 24 images from the self-acquisition and SUAS datasets. Additionally, we obtained 30 images from the WILCHES and BOSPHORUS databases for inference. These test and inference datasets were also augmented using the same technique, resulting in a total of 432 augmented images across both datasets. This augmentation approach helped ensure the generalization of the proposed method despite the limited size of the dataset.

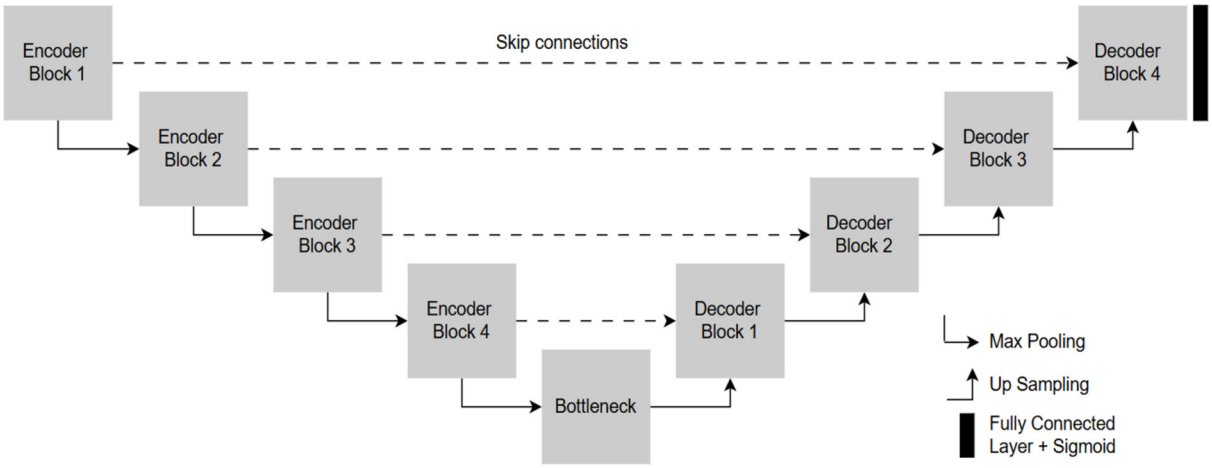

Fig.2: *Generator architecture.*

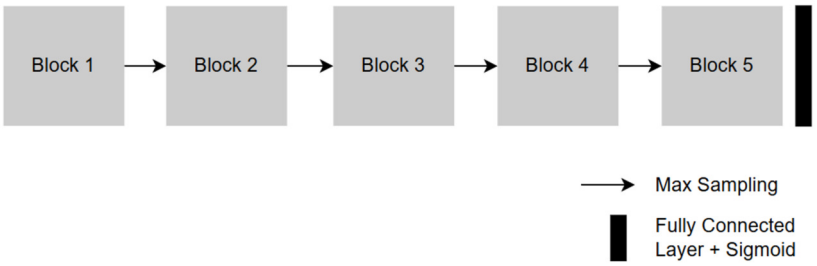

Fig.3: *Discriminator architecture.*

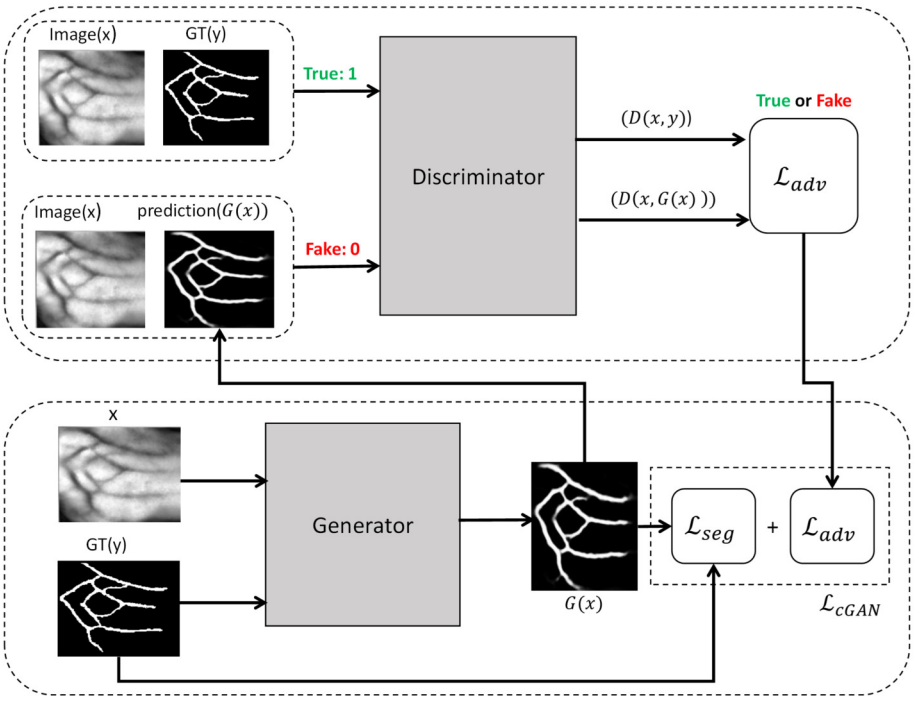

Fig.4: *The overall architecture of cGAN for vein segmentation.*

2.2 Vein Segmentation Model

2.2.1 Developing of cGAN Vein Segmentation Model

The generator architecture employed in this work follows a U-Net structure, which comprises three main parts: the encoder, central, and decoder paths, as illustrated in Fig. 2. The encoder path is composed of four blocks, each consisting of two 3×3 convolutional layers, batch normalization, and a rectified linear unit (ReLU) activation function. Max-pooling down-samples the input image, enhancing the representation of vein features. Conversely, the decoder path mirrors the encoder, utilizing up-sampling instead of max-pooling to restore the spatial dimensions of the features. The central path shares blocks with the decoder. A fully connected layer scales the output values between 0 and 1 per pixel.

The discriminator employed in this work is constructed upon a conventional CNN, as illustrated in Fig. 3. It comprises five convolution blocks, each mirroring the structure of the encoder block in the generator. The input image is down-sampled via a 3×3 kernel max-pooling layer within each convolutional block. A final layer, consisting of a 1×1 convolution followed by a sigmoid activation layer, scales the output of the discriminator between 0 and 1 for each pixel.

The cGAN model utilized in this work leverages an enhanced vein image x and its corresponding annotated vein image $GT(y)$ as conditional inputs for the generator, aiming to produce a segmented vein image $G(x)$ as depicted in Fig. 4. Unlike conventional cGANs, this model omits the addition of noise z as input to expedite training and reduce the likelihood of generating fake images.

Instead, the model relies solely on vein image x and its annotated image $GT(y)$ to guide the segmentation process, aligning with established methodologies in the field. Moreover, the discriminator in the cGAN model distinguishes between the annotated and generated segmented vein images, thereby contributing to the adversarial training process and improving the quality of the generated segmentation outputs.

The generated segmented vein image $G(x)$, alongside the conditional input $GT(y)$ and the enhanced vein image x, are input into the discriminator to assess their authenticity. The performance of the cGAN is evaluated based on its total loss, comprising the segmentation loss and the adversarial loss. The calculation of the total loss is as follows [29]:

$$\mathcal{L}_{cGAN} = \mathcal{L}_{seg} + \lambda \mathcal{L}_{adv} \tag{1}$$

Where $\mathcal{L}_{seg}$ represents the loss between the segmentation result $G(x)$ and the annotated vein image $GT(y)$, while $\mathcal{L}_{adv}$ denotes the adversarial loss of the generator against the discriminator. For optimal performance, set the balancing parameter λ t to 10.

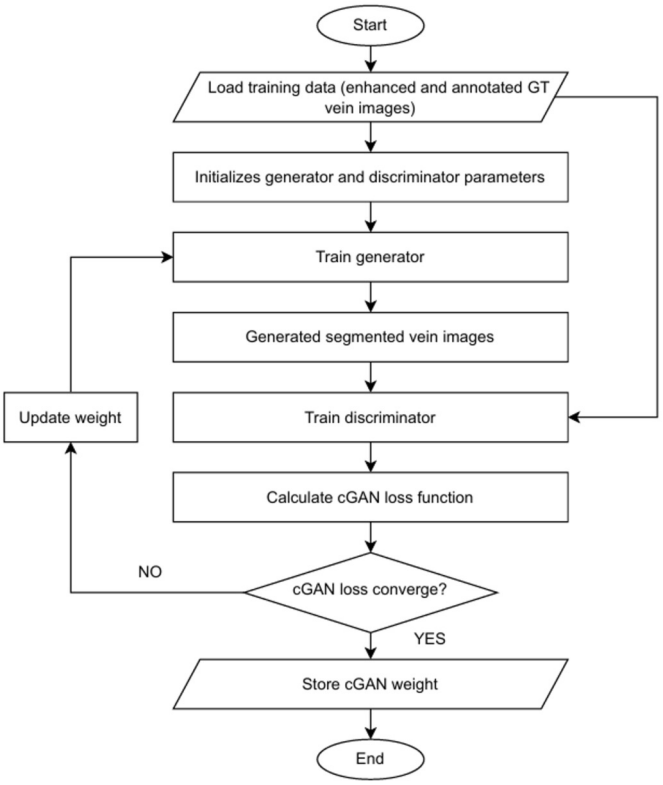

Fig.5: *Flowchart of the training process to develop a cGAN vein segmentation.*

2.2.2 Adversarial Training of cGAN Model

The proposed cGAN model was implemented using the Keras-TensorFlow deep learning library on Google Colab, leveraging NVIDIA GPU K80s acceleration. After constructing the generator and discriminator, the cGAN vein segmentation model undergoes adversarial training, alternating between optimizing the generator and the discriminator. The ADAM optimizer, with a learning rate 0.0002 for both components, is employed to iteratively update their parameters. The training process, as illustrated in Fig. 5, consists of three main steps:

i. **Generation:** The generator produces a segmented vein image $G(x)$ using the enhanced vein image x and the annotated vein image, $GT(y)$ as the input.
ii. **Discriminator:** The generated vein images $G(x)$ and the enhanced vein images x are fed into the discriminator for authenticity identification. The discriminator's parameters are updated using the loss function in Equation (1).
iii. **Fine-tuning:** The discriminator becomes untrainable, freezing its parameters. Only the generator's parameters are adjusted to enhance the realism of segmented vein images. Once a cycle completes, the training process recommences from Step i.

The generator and discriminator alternate optimization cycles until they reach a Nash equilibrium. At this point, the generator produces realistic data, and the discriminator struggles to distinguish between fake and real data. The training process terminates when no further decrease in total loss is observed over successive iterations, indicating that the model has converged and additional training is unnecessary.

This adversarial framework inherently addresses the overfitting problem by encouraging the generator and discriminator to generalize better. The generator is discouraged from memorizing the training data, while the discriminator learns meaningful representations of the data rather than overfitting to specific examples. Additionally, cGANs augment the training dataset by generating synthetic samples that closely resemble real data, thereby increasing dataset diversity and reducing the risk of overfitting.

Furthermore, the adversarial process forces both networks to focus on capturing the underlying data distribution rather than finding trivial solutions. If the generator produces repetitive or unrealistic outputs, the discriminator quickly identifies them, prompting it to explore the data space more effectively. This dynamic interaction ensures that the generator produces generalized outputs, which reduces model complexity. Together, these mechanisms make cGAN training a powerful approach to mitigating overfitting.

2.3 Modified Attention Module

2.3.1 Modified ECA Design 1

Fig.6 shows the block diagram of the first design of the modified ECA module, with the highlighted block representing the altered element. In the previous ECA module, channel-wise attention weights are computed after the GAP operation using a 1-D standard convolution with an adaptive kernel size of k_1. This work replaces the 1-D standard convolution with a 1-D dilated convolution, where the dilation rate r adapts based on the kernel size k_1. The advantages of dilated convolution inspired this modification, as it allows for a larger receptive field compared to standard convolution. Consequently, it enables better capture of more significant spatial features in the input image, enhancing the module's performance. The adaptive kernel size, k_1, is calculated following the method outlined in [16]:

$$r = k_1 = \left[\frac{\log_2 C}{\gamma} + \frac{b}{\gamma} \right]_{odd} \tag{2}$$

Where $|t|_{odd}$ denotes the nearest odd number of t, and constants γ and b are set to 2 and 1, respectively. The output feature map of this module is computed as follows:

$$F_C = \sigma\left(f_1^{k1}(GAP(F))\right) \otimes F \tag{3}$$

Where f_r^{k1} represents the 1D dilated convolution operation, F is the merged feature map inserted into the modified ECA design 1 and $\otimes$ denotes element-wise multiplication. Unlike prior work by Wang *et al.* [20], which manually fixes the dilation rate, r, this work employs an adaptive value of r dependent on k_1. The sigmoid activation function $\sigma(\cdot)$ maps the output between 0 and 1. The primary difference from the previous ECA module lies in adopting dilated convolution with an adaptive dilation rate, which enhances the module's ability to capture more significant spatial features in the input image compared to standard convolution.

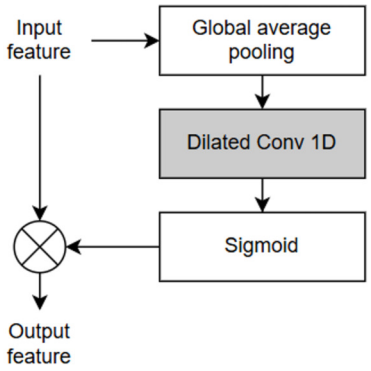

Fig.6: *Block diagram of modified ECA Design 1.*

2.3.2 Modified ECA Design 2

Fig. 7 illustrates the block diagram of the second design of the modified ECA module, with highlighted blocks indicating modifications. To leverage the ability to discern crucial regions and channels in the input image, as demonstrated by the spatial and channel attention mechanisms of the previous CBAM module, we introduced a spatial module into the prior modified ECA design 1, resulting in modified ECA design 2. The spatial module comprises average pooling, followed by two 2D dilated convolutions with a 3×3 kernel size and varying dilation rates of 2 and 3. The computation of the additional spatial module is expressed as follows:

$$F_S = \sigma\left(f_{r=3}^{3\times3}(f_{r=2}^{3\times3}(AP(F)))\right) \otimes F \quad (4)$$

Where $f_{r=2}^{3\times3}$ and $f_{r=3}^{3\times3}$ denote two 2D dilated convolutions with dilation rates of 2 and 3, respectively. $AP(F)$ represents the average pooling applied to the input feature map F, and $\sigma(\cdot)$ denotes the sigmoid activation function. Finally, the outputs from modified ECA design 1, denoted as F_C (refer to Eq. 3), and the added spatial module is, denoted as F_S (refer to Eq. 4), are combined using element-wise multiplication to generate the feature map of modified ECA design 2:

$$F_O = F_C \otimes F_S \quad (5)$$

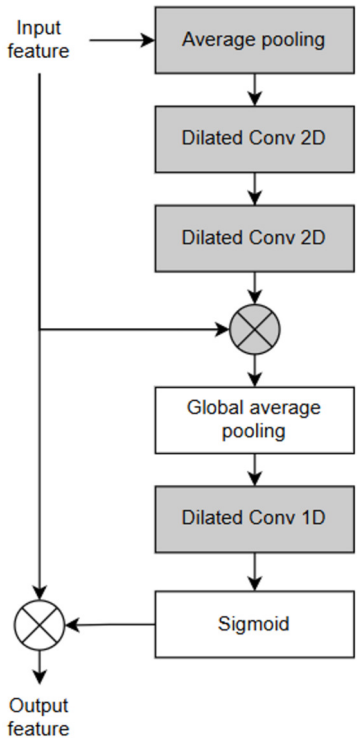

Fig.7: *Block diagram of modified ECA Design 2.*

2.3.3 Modified ECA Design 3

Fig. 8 depicts modified ECA design 3, highlighted modifications made to the existing ECA module. The original ECA module summarizes spatial information into feature maps using only the GAP operation. However, combining GAP with global maximum pooling (GMP) can capture more comprehensive object-specific features, leading to improved channel-wise attention. Thus, both operations are used concurrently in this work. Their outputs are then passed through a joint dilated 1D convolutional layer with an adaptive kernel size k_1 and dilation rates r to generate a channel attention map. The process can be summarized mathematically as follows:

$$F_C = \sigma\left(f_r^{k1}(GAP(F)) \oplus f_r^{k1}(GMP(F))\right) \otimes F \quad (6)$$

Where f_r^{k1} denotes the 1D dilated convolution operation with adaptive kernel size k_1 and dilation rates r. $GAP(F)$ and $GMP(F)$ represent the global average pooling and global maximum pooling operations, respectively, applied to the input feature map F. The sigmoid activation function $\sigma(\cdot)$ maps the output between 0 and 1.

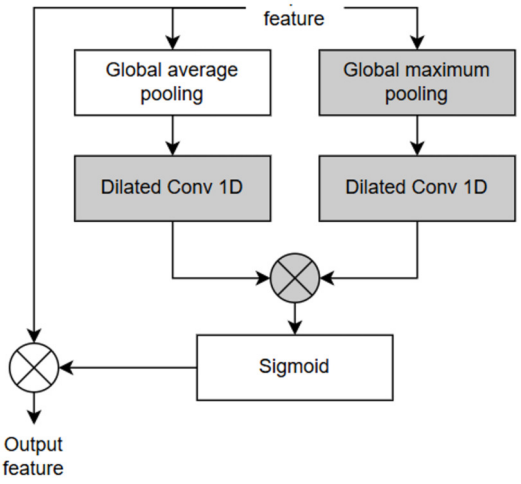

Fig.8: *Block diagram of modified ECA Design 3.*

2.4 Performance Measures

The proposed vein segmentation is quantitatively evaluated using sensitivity, accuracy, and dice coefficient by comparing each segmented image with its ground truth (GT) image. Sensitivity measures the ratio of correctly identified foreground (vein pixels) to the total vein pixels in the ground truth image. It indicates the algorithm's ability to detect veins accurately. It is calculated using the formula [30]:

$$sensitivity = \frac{TP}{TP + FN} \quad (7)$$

Where TP (true positive) is the number of correctly

segmented vein pixels, while FN (false negative) is the number of vein pixels wrongly classified as background. Sensitivity values range from 0 to 1, with higher values indicating more accurate vein segmentation.

Accuracy evaluates the overall performance of the modified segmentation technique by measuring the ratio of correctly segmented veins and non-vein structures in a vein image. It confirms the technique's effectiveness in accurately distinguishing between these components. Accuracy is calculated as [31]:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

Where TN (true negative) is the number of correctly segmented background pixels. FP (False positive) is the number of background pixels incorrectly classified as vein pixels. Higher accuracy values indicate effective segmentation of both veins and non-vein pixels.

The Dice coefficient is a performance metric used to quantify the overlap between segmented veins and their corresponding ground truth images. It assesses the degree of agreement and overlap between the two, measuring how well the segmented veins align with the actual vein structures. The dice coefficient is calculated as [32]:

$$dice\ coefficient = \frac{2 * TP}{2 * TP + FP + FN} \tag{9}$$

The dice coefficient ranges between 0 and 1, with a higher value indicating a greater level of similarity between the segmented and ground truth vein images.

Two performance evaluations are conducted to validate the proposed techniques. The first evaluation assesses the performance of the three best modified ECA designs (1, 2, and 3) in cGAN vein segmentation. The second evaluation evaluates the performance of the best modified ECA module against commonly used attention mechanism modules (SE, CBAM, and ECA modules).

3. RESULTS

The evaluation process involved utilizing 432 images from four hand vein datasets. Qualitatively, segmentation results were visually inspected, with sample outputs for self-acquisition, SUAS, WILCHES, and BOSPHORUS datasets shown in Fig.9. Upon inspection, it was observed that while all techniques preserved original vein patterns, designs 1 and 2 exhibited coarse vein edges, and disconnected parts of the patterns. These designs also misclassified some vein pixels, potentially impacting clinical diagnosis accuracy. In contrast, design 3's proposed modified ECA module produced more precise vein patterns in terms of connectivity and smoothness compared to other designs. This indicates that Design 3 accurately segments hand vein images, demonstrating its practicality for reliable clinical diagnosis and analysis.

The quantitative analysis presented in Tables 1

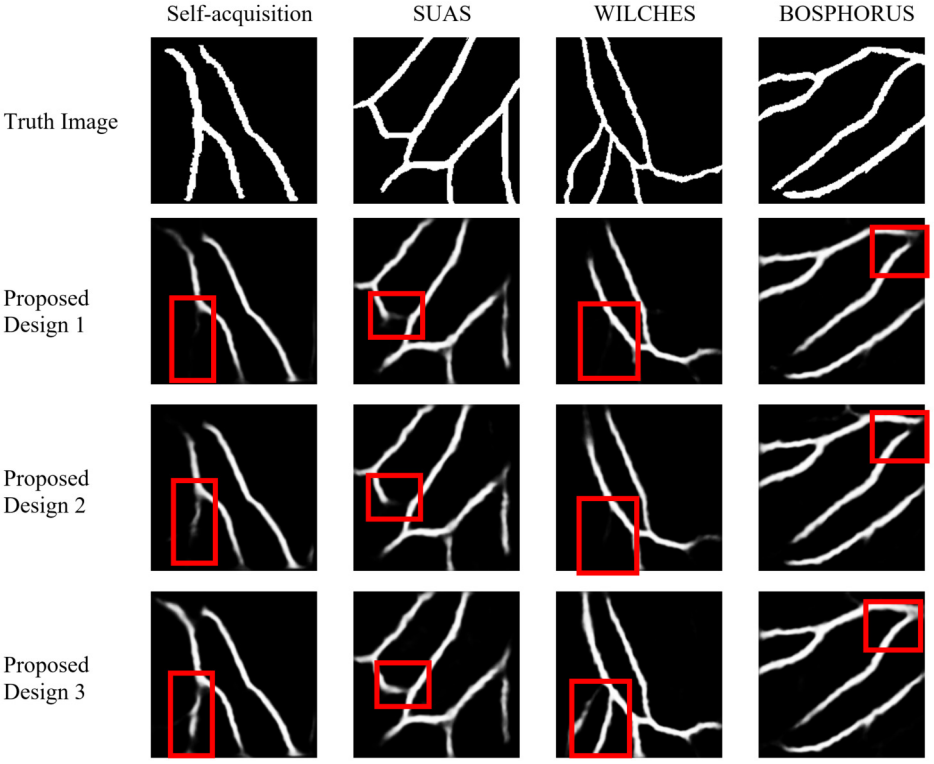

Fig.9: *Sample of segmented vein images using cGAN with improved attention modules for different datasets.*

Table 1: *Result of modified attention module in cGAN architecture for self-acquisition and SUAS datasets.*

Design No	Self-acquisition			SUAS		
	Sen	Acc	Dice	Sen	Acc	Dice
1	0.8397	0.9597	0.7604	0.8534	0.9575	0.7599
2	0.8397	0.9606	0.7671	0.8557	0.9574	0.7548
3	**0.8482**	**0.9675**	**0.7725**	**0.8644**	**0.9598**	**0.7632**

Table 2: *Result of modified attention module in cGAN architecture for WILCHES and BOSPHORUS datasets.*

Design No	WILCHES			BOSPHORUS		
	Sen	Acc	Dice	Sen	Acc	Dice
1	0.8944	0.9595	0.8063	0.8705	0.9578	0.8059
2	0.9083	0.9604	0.8081	0.8798	0.9584	0.8071
3	**0.9125**	**0.9674**	**0.8162**	**0.8860**	**0.9608**	**0.8096**

Table 3: *Result of average performance of modified attention module in cGAN architecture based on images from all four datasets.*

Design No	Average Performance		
	Sen	Acc	Dice
1	0.8645	0.9586	0.7831
2	0.8722	0.9591	0.7834
3	**0.8778**	**0.9639**	**0.7904**

and 2 complements the findings from the qualitative evaluation. Across all four hand vein datasets (self-acquisition, SUAS, WILCHES, and BOSPHORUS), Design 3 of the modified ECA module demonstrated higher sensitivity, accuracy, and dice coefficient than Design 1 and Design 2.

Furthermore, the average metrics for all datasets in Table 3 further confirm Design 3's superior performance, showing higher average sensitivity, accuracy, and dice coefficient. Compared to Design 1 and Design 2, Design 3 exhibited notable improvements in average sensitivity, accuracy, and dice coefficient, indicating its robustness across different datasets and its ability to produce accurate segmentation results consistently. These quantitative results validate the qualitative observations and highlight the effectiveness of Design 3 in vein segmentation tasks.

The proposed technique underwent qualitative and quantitative evaluation using 432 images from four hand vein datasets. The results are visually depicted in Fig. 10, showcasing the segmentation outcomes for self-acquisition, SUAS, WILCHES, and BOSPHORUS datasets. In the analysis, all techniques could maintain the original vein patterns to some extent. However, SE, CBAM, and ECA methods showed coarse vein edges and disconnected segments in their segmentation results. The proposed attention module showed a significant difference upon comparison with these techniques. SE, CBAM, and ECA modules also displayed misclassification of vein pixels, as indicated by the red box in the figures. In contrast, the proposed module produced more precise vein patterns with improved connectivity and smoothness, surpassing other segmentation methods regarding segmentation quality.

The quantitative analysis in Tables 4 and 5 supports the earlier segmentation results. The modified ECA module demonstrates superior sensitivity, accuracy, and dice coefficient compared to existing attention modules (SE, CBAM, and ECA) across all four hand vein datasets: self-acquisition, SUAS, WILCHES, and BOSPHORUS.

Furthermore, Table 6 summarizes the average metrics for all datasets, showing the proposed attention module consistently outperforming SE, CBAM, and ECA in terms of sensitivity, accuracy, and dice coefficient. Specifically, compared to SE, CBAM, and ECA, the proposed module improved average sensitivity, accuracy, and Dice coefficient, demonstrating its robustness across various datasets. These quantitative findings validate and reinforce the qualitative observations, establishing the efficacy and superiority of the proposed modified ECA module, particularly Design 3, in accurately segmenting hand vein images and outperforming existing attention mechanisms in vein segmentation tasks.

4. CONCLUSION

In conclusion, this study leveraged the capabilities of a cGAN deep learning model to overcome limitations commonly associated with conventional techniques, particularly issues related to noise in generated images. Incorporating a modified ECA module

into the model architecture to address information loss and redundant layers significantly improved vein segmentation tasks. The results demonstrated that the modified ECA module outperformed existing attention modules (SE, CBAM, and ECA) across all evaluated metrics, underscoring both the robustness and efficacy of the proposed method in enhancing vein segmentation within the cGAN framework. As the field of medical image analysis continues to evolve, this study provides valuable insights to guide the development of more accurate and reliable segmentation techniques for clinical applications. The robustness and effectiveness of this approach highlight its potential for advancing medical image analysis and diagnosis in various healthcare settings.

Although the cGAN integrated with the improved ECA model demonstrates promising segmentation performance for hand vein images, there is still considerable room for improvement. In the future, we aim to prune the existing models and reduce network parameters without affecting performance, enabling real-time medical segmentation and enhancing practical utility.

ACKNOWLEDGEMENT

This research was funded by Universiti Malaysia Pahang Al-Sultan Abdullah, grant number RDU220371.

AUTHOR CONTRIBUTIONS

Conceptualization, M.S.S and N.F.; methodology, M.Y.; software, M.Y.; validation, M.Z.I., W.N.A.W.S and M.S.S.; formal analysis, M.Y.; investigation, M.Y.; data curation, M.Y.; writing—original draft preparation, M.Y.; writing—review and editing, M.Z.I., N.F. and W.N.A.W.S.; visualization, M.Y. and M.S.S.; supervision, M.Z.I.; funding acquisition, M.Y. All authors have read and agreed to the published version of the manuscript.

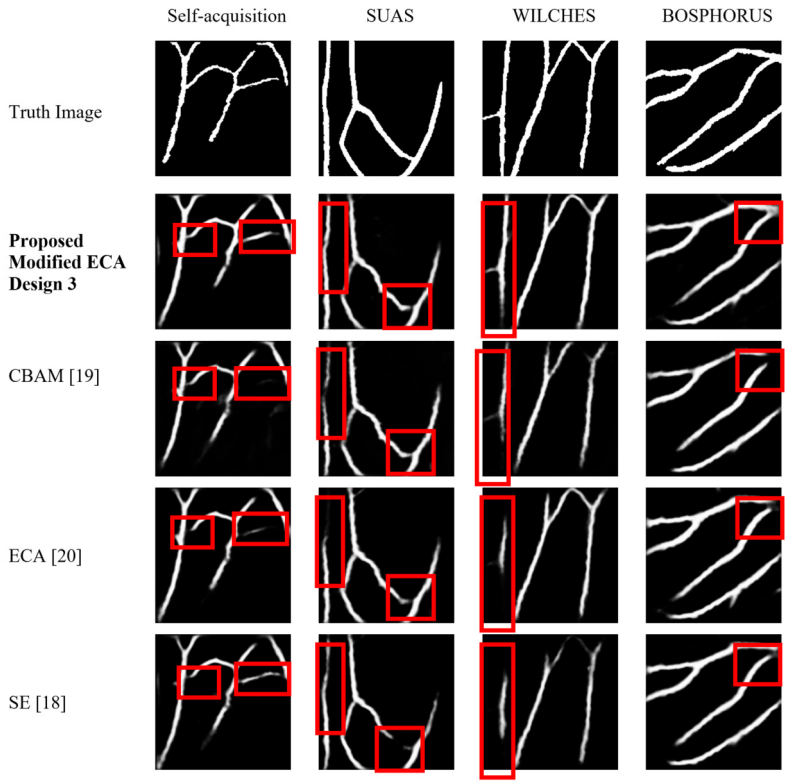

Fig.10: *Sample of segmented images integrating various attention modules in cGAN for different datasets.*

Table 4: *Result of proposed attention module comparison to existing attention modules in cGAN architecture for self-acquisition and SUAS datasets.*

Attention modules	Self-acquisition			SUAS		
	Sen	Acc	Dice	Sen	Acc	Dice
Proposed	**0.8482**	**0.9675**	**0.7725**	**0.8644**	**0.9598**	**0.7632**
CBAM [19]	0.8383	0.9579	0.7476	0.8469	0.9564	0.7528
ECA[20]	0.8386	0.9598	0.7619	0.8555	0.9574	0.7560
SE[18]	0.8365	0.9595	0.7600	0.8528	0.9570	0.7549

Table 5: *Result of proposed attention module comparison to existing attention modules in cGAN architecture for WILCHES and BOSPHORUS datasets.*

Attention modules	WILCHES			BOSPHORUS		
	Sen	Acc	Dice	Sen	Acc	Dice
Proposed	**0.9125**	**0.9674**	**0.8162**	**0.8860**	**0.9608**	**0.8096**
CBAM [19]	0.9011	0.9587	0.8033	0.8678	0.9537	0.7796
ECA[20]	0.9024	0.9598	0.8043	0.8762	0.9579	0.8049
SE[18]	0.8972	0.9578	0.7893	0.8753	0.9568	0.7987

Table 6: *Result of proposed attention module comparison to existing attention modules in cGAN architecture for four hand vein datasets.*

Attention modules	Average Performance		
	Sen	Acc	Dice
Proposed	**0.8778**	**0.9639**	**0.7904**
CBAM [19]	0.8635	0.9566	0.7708
ECA[20]	0.8681	0.9587	0.7817
SE[18]	0.8654	0.9577	0.7757

References

[1] N. T. Han, N. Mukahar and B. A. Rosdi, "Finger Vein Recognition Using Principle Component Analysis and Adaptive k-Nearest Centroid Neighbour Classifier," *Int. J. Integr. Eng.*, vol. 13, no. 1, pp. 177–187, 2021.

[2] C. X. Jing, G. C. Meng, M. T. Chai, S. A. Z. S. Aluwee and S. A. A. Shah, "Subcutaneous Vein Recognition System Using Deep Learning for Intravenous (IV) Access Procedure," *Int. J. Integr. Eng.*, vol. 15, no. 3, pp. 73–83, 2023.

[3] M. Yakno, J. Mohamad-Saleh and B. A. Rosdi, "Grayscale and binary enhancement of dorsal hand vein images," *Int. J. Circuits, Syst. Signal Process.*, vol. 11, pp. 129–141, 2017.

[4] S. C. Soh, M. Z. Ibrahim and M. B. Yakno, "A review: Personal identification based on palm vein infrared pattern," *J. Telecommun. Electron. Comput. Eng.*, vol. 10, no. 1–4, pp. 175–180, 2018.

[5] N. Laopracha and K. Bunluewong, "Improving the Performance of CNN by Using Dominant Patterns of CNN for Hand Detection," *ECTI Transactions on Computer and Information Technology*, vol. 17, no. 2, pp. 265–277, 2023.

[6] Y. Wang, X. Cao and X. Miao, "Cross-device recognition of dorsal hand vein images by two-stage coarse-to-fine matching," *Vis. Comput.*, vol. 38, no. 11, pp. 3595–3610, 2022.

[7] Q. Yan *et al.*, "Attention-Guided Deep Neural Network with Multi-Scale Feature Fusion for Liver Vessel Segmentation," *IEEE J. Biomed. Heal. Informatics*, vol. 25, no. 7, pp. 2629–2642, 2021.

[8] X. Li, Z. Li, D. Yang, L. Zhong, L. Huang and J. Lin, "Research on finger vein image segmentation and blood sampling point location in automatic blood collection," *Sensors*, vol. 21, no. 1, pp. 1–14, 2021.

[9] D. Zhong, H. Shao and S. Liu, "Towards application of dorsal hand vein recognition under uncontrolled environment based on biometric graph matching," *IET Biometrics*, vol. 8, no. 2, pp. 159–167, 2019.

[10] F. Marattukalam and W. H. Abdulla, "Segmentation of palm vein images using U-Net," in *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, no. December, pp. 64–70, 2020.

[11] S. Lefkovits and S. Emerich, "Boosting unsupervised dorsal hand vein segmentation with U-Net variants," *Mathematics*, vol. 10, no. 2620, pp. 1–31, 2022.

[12] A. Kuthiala *et al.*, "U-DAVIS-Deep learning based arm venous image segmentation technique for venipuncture," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–9, 2022.

[13] T. He, C. Guo, L. Jiang and H. Liu, "Automatic Venous Segmentation in Venipuncture Robot Using Deep Learning," *2021 IEEE International Conference on Real-time Computing and Robotics (RCAR)*, Xining, China, pp. 614-619, 2021.

[14] A. Z. Alali and K. Hussein Ali, "Segmentation of Human Brain Gliomas Tumour Images using U-Net Architecture with Transfer Learning," *Diyala J. Eng. Sci.*, vol. 8716, no. 1, pp. 17–29, 2022.

[15] M. Yakno, J. Mohamad-Saleh and M. Z. Ibrahim, *Dorsal Hand Vein Segmentation Using Vein-Generative Adversarial Network (V-GAN)*

Model, vol. 829 LNEE. Springer Singapore, 2022.

[16] X. Gao, G. Zhang and K. Wang, "Segmentation model of dorsal hand vein based on improved U-Net," in *Proceedings of the 2021 5th International Conference on Electronic Information Technology and Computer Engineering*, pp. 1361–1365, 2021.

[17] W. Yang, C. Hui, Z. Chen, J. -H. Xue and Q. Liao, "FV-GAN: Finger Vein Representation Using Generative Adversarial Networks," in *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 9, pp. 2512-2524, Sept. 2019.

[18] V. Abramova *et al.*, "Hemorrhagic stroke lesion segmentation using a 3D U-Net with squeeze-and-excitation blocks," *Comput. Med. Imaging Graph.*, vol. 90, no. 101908, pp. 1–12, 2021.

[19] S. Luan, X. Xue, Y. Ding, W. Wei and B. Zhu, "Adaptive attention convolutional neural network for liver tumor segmentation," *Front. Oncol.*, vol. 11, pp. 1–12, 2021.

[20] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo and Q. Hu, "ECA-Net: Efficient channel attention for deep convolutional neural networks," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2020, pp. 11531–11539.

[21] X. Shu, F. Chang, X. Zhang, C. Shao and X. Yang, "ECAU-Net: Efficient channel attention U-Net for fetal ultrasound cerebellum segmentation," *Biomed. Signal Process. Control*, vol. 75, no. December 2021, p. 103528, 2022.

[22] Y. Liu *et al.*, "ECA-TFUnet: A U-shaped CNN-Transformer network with efficient channel attention for organ segmentation in anatomical sectional images of canines," *Math. Biosci. Eng.*, vol. 20, no. 10, pp. 18650–18669, 2023.

[23] M. Yakno, J. Mohamad-Saleh, M. Z. Ibrahim and W. N. A. W. Samsudin, "Camera-projector calibration for near-infrared imaging system," *Bulletin of Electrical Engineering and Informatics*, vol. 9, no. 1, pp. 160–170, 2020.

[24] M. E. Cimen, O. F. Boyraz, M. Z. Yildiz and A. F. Boz, "A new dorsal hand vein authentication system based on fractal dimension box-counting method," *Optik (Stuttg).*, vol. 226, no. P1, p. 165438, 2021.

[25] F. Wilches-Bernal, B. Núñez-Álvares and P. Vizcaya, "A Database of Dorsal Hand Vein Images," pp. 1–4, 2020.

[26] O. Toygar, F. O. Babalola, and Y. Bitirim, "FYO: A novel multimodal vein database with palmar, dorsal and wrist biometrics," *IEEE Access*, vol. 8, pp. 82461–82470, 2020.

[27] M. Yakno, J. Mohamad-Saleh, M. Z. Ibrahim, S. M. Shaharum and R. Abdul-Karim, *Hand Vein Region of Interest (ROI) Extraction Using Faster Region-Based Convolutional Neural Network (R-CNN)*, vol. 1213 LNEE. Springer Nature Singapore, 2024.

[28] M. Yakno, J. Mohamad-Saleh and M. Z. Ibrahim, "Dorsal hand vein image enhancement using fusion of clahe and fuzzy adaptive gamma," *Sensors*, vol. 21, no. 19, pp. 1–19, 2021.

[29] X. Guo *et al.*, "Retinal vessel segmentation combined with generative adversarial networks and Dense U-Net," *IEEE Access*, vol. 8, pp. 194551–194560, 2020.

[30] H. A. Phan *et al.*, "Development of an Autonomous Component Testing System with Reliability Improvement Using Computer Vision and Machine Learning," *ECTI Transactions on Computer and Information Technology*, vol. 18, no. 1, pp. 64–75, 2024.

[31] D. Divya and M. Thilagu, "Region Growing based K-means Clustering and Optimal Weight Prior-Attention Residual Learning for Segmentation and Classification of COVID-19 CT Images," *ECTI Transactions on Computer and Information Technology*, vol. 18, no. 1, pp. 76–88, 2024.

[32] S. Das, M. Mishra and S. Majumder, "Identification of Glaucoma from Retinal Fundus Images using Deep Learning Model, MobileNet," *ECTI Transactions on Computer and Information Technology*, vol. 18, no. 3, pp. 371–380, 2024.

Marlina Yakno received her BEng in Electrical and Electronic Engineering from Universiti Tenaga Nasional (UNITEN) in 2007 and her MSc and PhD degrees from Universiti Sains Malaysia (USM) in 2013 and 2023, respectively. She is currently a Senior Lecturer at the Faculty of Electrical and Electronic Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA). Her research interests include computer vision, the Internet of Things (IoT), computational intelligence, and deep learning.

Mohd Zamri Ibrahim received his BEng and MEng in electrical engineering from the Universiti Teknologi Malaysia, Malaysia in 2004 and 2008 respectively. He earns his PhD from Loughborough University, UK in 2014. He is currently an Associate Professor in the Faculty of Electrical and Electronics Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah. His research interests include computer vision, plasma science, embedded system programming, brain computer interaction, image processing, intelligent system, deep learning and speech recognition.

Muhammad Salihin Saealal is a lecturer at Universiti Teknikal Malaysia Melaka, specializing in Optimization Algorithms, Neural Networks, Artificial Intelligence, Deep Learning, and Image Processing. He holds a Bachelor's (Hons) and Master's degree in Electrical Engineering from Universiti Teknologi Malaysia (UTM) and a PhD in Electrical and Electronic Engineering from Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA). His research focuses on developing advanced computational models and intelligent systems to enhance automation and decision-making processes. With a strong passion for AI and deep learning, Muhammad Salihin actively contributes to academic research, innovation, and the advancement of machine learning and optimization techniques.

Norasyikin Fadilah is a lecturer at Faculty of Electrical and Electronic Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah. She received her BEng in Electrical Engineering from Stevens Institute of Technology and MSc in Electrical and Electronic Engineering from Universiti Sains Malaysia. Her research interests include computer vision and deep learning.

Wan Nur Azhani W. Samsudin is currently a senior lecturer in the Faculty of Electrical and Electronics Engineering Technology, at Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA), Malaysia. She received a BEng in Electrical and Electronics Engineering from the Universiti Malaysia Pahang, Malaysia in 2007 and an MSc, majoring in Computer Science from Universiti Malaya, Malaysia in 2010. She earned her PhD from Universiti Malaysia Pahang in Electrical Engineering in 2017. Her research interest is in computer and machine vision, image processing, recognition, deep learning, artificial intelligence, behavioral analysis, and surveillance systems. Her latest research work has won in Creation, Innovation, Technology & Research Exposition, CITREX 2021, CITREX 2022 and CITREX 2024.