Preprints.org

Article

U-Net Embeded Gabor Kernel and Coaxial Correction Methods to Dorsal Hand Vein Image Projection System

Liukui Chen , Monan Lv , Junfeng Cai , Zhongyuan Guo , Zuojin Li *

Article

U-Net Embeded Gabor Kernel and Coaxial Correction Methods to Dorsal Hand Vein Image Projection System

Liukui Chen [1], Monan Lv [1], Junfeng Cai [1], Zhongyuan Guo [2] and Zuojin Li [1],*

[1] Chongqing University of Science & Technology, Chongqing, 401331, China

[2] Southwest University, Chongqing, 400715, China

* Correspondence: cqustlzj@sina.cn

Abstract: Vein segmentation and projection correction constitute the core algorithms of auxiliary venipuncture device, responding to the accurate venous positioning to assist puncture and reduce the number of punctures and pain of patients. This paper proposes an improved U-Net for segmenting vein and a coaxial correction for image alignment in the self-build vein projection system. The proposed U-Net is embedded by Gabor convolution kernels in the shallow layers to enhance segmentation accuracy. Additionally, to mitigate the semantic information loss caused by channel reduction, the network model is lightweighted by the mean of replacing conventional convolutions with inverted residual blocks. During the visualization process, a method that combining coaxial correction and homography matrix is proposed to address the non-planarity of the dorsal hand in this paper. First, use a hot mirror to adjust the light paths of both projector and camera to be coaxial, and then align the projected image with dorsal hand using a homography matrix. Using this approach, the device requires only a single calibration before use. With the implementation of the improved segmentation method, an accuracy rate of 95.12% is achieved by the dataset. The intersection over union ratio between the segmented and original images is reached at 90.07%. The entire segmentation process is completed in a 0.09 second, and the largest distance error of vein projection onto dorsal hand is 0.53mm. Experiments show that the device has reached practical accuracy and has the value of research and application.

Keywords: auxiliary venipuncture; vein segmentation; improved U-Net; coaxial correction; vein projection system

1. Introduction

As a routine medical procedure, the success rate of venipuncture is not solely dependent on the medical staff's skill and experience [1]. Various objective factors can also affect the outcome, such as dark skin, a thick subepidermal fat layer, and poor lighting conditions, all of which can decrease the success rate [2]. Due to the high precision and sensitivity, venipuncture robots [3] are starting to emerge. With the continuous development of technology, the most important vein segmentation algorithm has been continuously updated. Currently, most vein projectors employ near-infrared imaging principles to capture vein images [4]. By using a near-infrared camera, the features of the veins in near-infrared images are more distinct compared to those appearing blue-green under natural light. The obtained image is transmitted to the processor for enhancement and segmentation processing. Finally, the segmented binary image is fed back to the dorsal of the hand by the projector, allowing medical staff to perform vein puncture through the projected image. Regarding the entire hand vein imaging process, this paper makes improvements to the vein segmentation and projection correction algorithms. Furthermore, a system is developed to validate the method, yielding a better result.

Since the development of the first assisted vein puncture device [5], there has been an increasing demand for accuracy in the projection results. Usually, the obtained NIR images are accompanied

with severe noise because veins are hidden underneath the skin and fat, but the NIR light penetrates the skin to a limited depth. Although the pixel's brightness in the vein and dorsal is different, the overall contrast of the image is low, and the edges of the veins are often blurry [6]. Furthermore, the camera and projector are not located in the same position, resulting in a misalignment of the projected image with dorsal hand. Therefore, this paper investigates two critical techniques in vein projector technology: image segmentation and projection correction. For the image segmentation problem, the core idea of traditional methods is to design a feature extractor. However, the extractor requires manual setting of certain parameters, which may not be suitable for different application scenarios. Additionally, manually designed feature extractors may fail to capture complex image features, resulting in poor segmentation accuracy and low robustness. Subsequently, regarding projection correction algorithms, most existing methods align the projected image with dorsal by using homography matrix. However, the hand has larger curvature towards the edges, resulting in less accurate projection results. Furthermore, as the supporting surface, the hand is not fixed, any change in its position will lead to changes in the homography matrix [7].

To address these issues, this paper utilizes corresponding solutions, which are also the main contributions as follows:

(1) The prior knowledge of Gabor kernel is integrated into U-net to improve the network. Through the embedding of the Gabor kernel into neural network, the feature extraction capability of shallow network is enhanced. Consequently, it leads to improved accuracy in image segmentation and lays a foundation for later precise projection.

(2) The proposed U-net achieves a lightweight design by replacing conventional convolutions with inverted residual blocks. This operation mitigate the semantic information loss caused by channel reduction, while decreasing the parameter size and is suitable for real-time projection.

(3) During the projection process, this paper proposes a method that combines coaxial correction with homography matrix. This approach enables the device requires only a single calibration before use. It enhances the accuracy of the vein projection system and simplifies the process of projection correction.

(4) To validate the improved algorithm, we have established a database, created corresponding labels manually, and constructed a dorsal hand vein projection system. The proposed methods for vein segmentation and projection have been trained and tested using the dataset. The accuracy of both segmentation and projection, along with the response time, meets the requirements of the application.

The rest of this survey is structured as follows. Section 2 presents the research background of image segmentation algorithm and projection correction algorithm. In Section 3, the modules of the imaging system are presented, along with the principles of the relevant algorithms used in this paper. Section 4 discusses the rationale behind parameter selection in the related algorithms, showcasing ultimate experimental outcomes. Lastly, section 5 summarizes the proposed algorithm, while delving into an analysis of the prevailing errors, elucidating the research trends and potential areas for improvement.

2. Related Research

The vein segmentation algorithm and projection correction algorithm are two crucial technologies in auxiliary venipuncture device. The former involves extracting vessels from dorsal hand images, while the latter addresses the issue of inaccurate projection caused by different equipment angles. In the field of image segmentation, Tong Liu [8] proposed an improved repeated line tracking algorithm for finger vein image segmentation in 2013. The algorithm utilizes vein width in the image to determine parameters, then uses the improved repeated line tracking method to figure out the locus space of the finger vein. Finally, the Otsu method is applied to extract the finger vein. To distinguish the vein region on the dorsal part of the hand from the non-vein region, Zhang J [9] utilized the kernel fuzzy C-means (KFCM) algorithm for initialization. Furthermore, they designed a defogging algorithm and an edge fitting term to improve the segmentation process, ultimately achieving better results. In 2012, with the ascent of convolutional networks, artificial intelligence [10]

has experienced rapid growth and advancement, and it gradually found its way into various aspects of human life. Long J [11] proposed an end-to-end fully convolutional network to handle image segmentation tasks in 2015. The approach utilized various classification networks for downsampling and ultimately yielded precise segmentation results. In the same year, Olaf Ronneberger et al. [12] proposed the U-Net for image segmentation. The U-Net consists of an encoder-decoder architecture that combines low- and high-level information, making it suitable for precise segmentation of medical images. M. Liu et al [13] proposed a deep convolutional neural network architecture with nested U-Net, by combining global and local loss information to optimize the network, achieving better segmentation results. In 2021, T. He et al [14] used U-Net with an attention mechanism for dorsal hand vein image segmentation. The algorithm highlighted the feature that had a significant impact on the results, and utilized a jump-connected structure for multi-scale feature recognition, thereby enhancing the accuracy of segmentation. Traditional image segmentation methods have limitations in practical applications, which can be greatly affected by grayscale distribution and poor at handling noise. Compared with traditional methods, neural networks can artificially divide the image contours to make labels [15], and specify the features to be extracted [16]. This approach can better handle noise, lighting variations, and other factors, making neural networks more robust than traditional methods. In this paper, we propose an approach to image segmentation using neural networks. Specifically, we improve the structure of existing segmentation networks and reduce the parameters for network training while ensuring segmentation accuracy.

Currently, although projection correction algorithms have yielded relatively good results, their application still has some shortcomings. The first venous projector was invented by American medical engineer Zeman [17] in 2004, which utilized a hot mirror to adjust the optical path of the camera and projector to coaxial, eliminating the influence of curved projection and enabling the projected image to coincide with the object. However, this method presents challenges in precision requirements for the equipment, and coaxial optical path alignment is difficult to achieve. In 2013, Dai X [18] designed a venous projector that treats the dorsum of the patient's hand as a planar surface with four labeled points. The projection image is adjusted through a perspective transformation, so that the four points in the projection image align with the four points labeled on the hand, facilitating accurate projection onto the dorsum of hand. However, this method requires the calibration of four points before each venipuncture, which is cumbersome. In 2016, Gan Q et al. [19] developed a projection navigation system that captures images at a rate of 60 frames per second. The projected image is calibrated through perspective transformation to guide the surgeon during the operation. In 2017, Tran Van Tien [20] used near-infrared light as the light source, fed near-infrared image into camera through the reflective function of beamsplitter. After passing through a median filter to reduce noise and undergoing Otsu segmentation, the near-infrared image was transmitted to a projector, and ultimately projected onto the patient's dorsum of hand. However, this method is limited by the high precision requirements of the instrument, as well as the information loss when the light passing through the beamsplitter. Hia Yee May [21] proposed a system for real-time visualization of veins based on near-infrared imaging, which displays the vein images in real time on the monitoring screen. However, the method does not project the segmented image on the dorsum of hand, and this method also lacks an effective design for the projection method. arkus Funk [22] proposed an in-situ projection-based prompter to assist workers with cognitive impairments. The product's location is obtained through a depth camera, and a homography matrix is used to project the related information onto the product. In 2018, Gunawa I [23,24] proposed a method of adjusting images using a distance sensor. After augmenting the acquired near-infrared images, the distance sensor is utilized to obtain the real-time distance from the hand's dorsum to the camera. With this distance, the projected image is finely tuned to synchronize with the object. This approach addresses the issue of changes in the homography matrix due to the hand's position change. Another method proposed by Liu P [25] uses fluorescence imaging with indocyanine green (ICG). After adjusting the optical paths of the camera and projector to coaxial, the images are projected back to the surgical site in real-time. Additionally, Li C [26] proposed a handheld projection device. In this device, the optical path of the camera in this device runs parallel to that of the projector. After cropping the non-

overlapping areas of the projection and the camera, the projected image is calibrated, and medical staff are guided during surgery based on the overlapping area image.

3. Our Method and System

To validate the improved algorithms, this paper devises a vein projection system comprising of three modules: an image acquisition module, an image processing module, and a projection module. In the image acquisition module, this paper designs a uniformly distributed near-infrared (NIR) light source by using a light guide plate. This design effectively reduces the impact of reflected flares on the images. The captured near-infrared images are transmitted to the image processing module, where they are enhanced and segmented by the improved U-Net algorithm. Finally, the light path of the projector and the NIR camera are corrected to be co-axial. Even if there is an error in the installation process that causes the light path of the two to be not completely co-axial, the projected image can also be adjusted to coincide with the real object through a homography matrix. Figure 1 depicts the physical components of the vein projector and illustrate each module in detail.
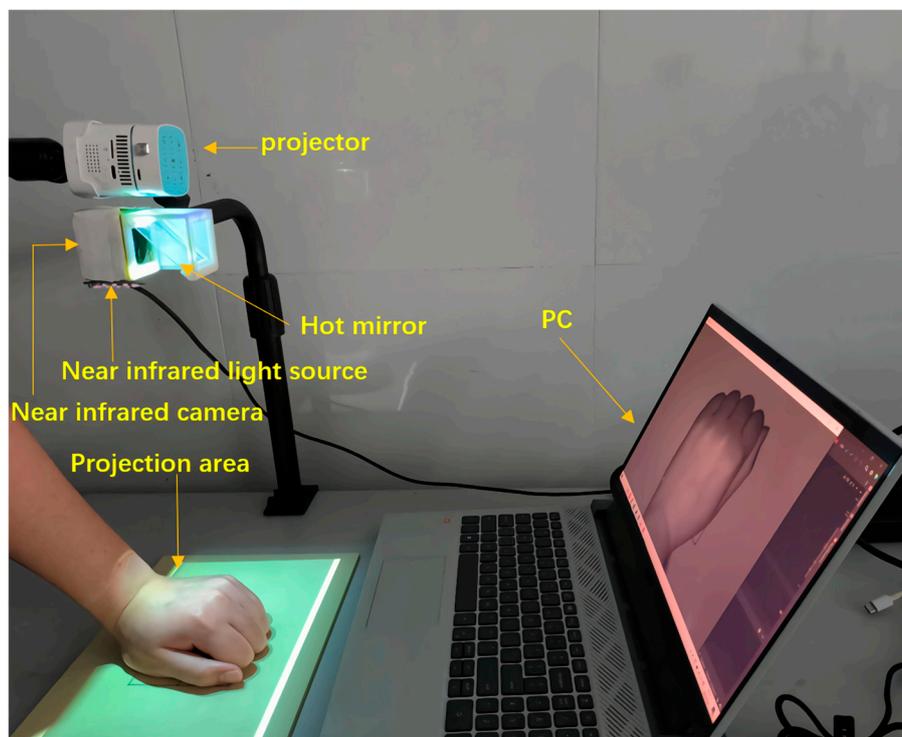

Figure 1. Venous projection system.

In this experiment, we employed the light-emitting diode (LED) with a wavelength of 850nm, placed it into the light guide plate, and utilized it to produce uniform near-infrared light. To capture the images, we used a distortion-free near-infrared camera with manual focus. The camera featured a frame rate of 30 frames per second and a viewing angle of 85 degrees. With regard to the projection module, we utilized a miniature projector with a model number of m100smart. This device achieved wireless connectivity with the image processing module via Parsec software, and the specific parameters can be viewed in the Table 1. The image processing module was configured with an AMD Ryzen 7 5800H and 16GB of RAM to ensure high performance. We used a hot mirror with an incidence angle of 45°, which had a strong reflection ability for near-infrared light with wavelengths ranging from 750nm to 1200nm. Additionally, this mirror allowed the light outside this range to pass through directly. The transmission rate and light wavelength relationship can be seen in Figure 2. To ensure stability and precision, we fixed the camera and hot mirror within a 3D printed box.

Table 1. Projector parameters.

Technical Index	Specification Parameters
Model Number	m100smart
Display Technique	DLP
Light Source Power	20W
Body Size	58x58x64mm
Range of Screen Placement	5-300 inches
Projector Brightness	400ANSI lumens
Resolution Ratio	1920x1080dpi
RAM	2GB
Support autofocus	Yes
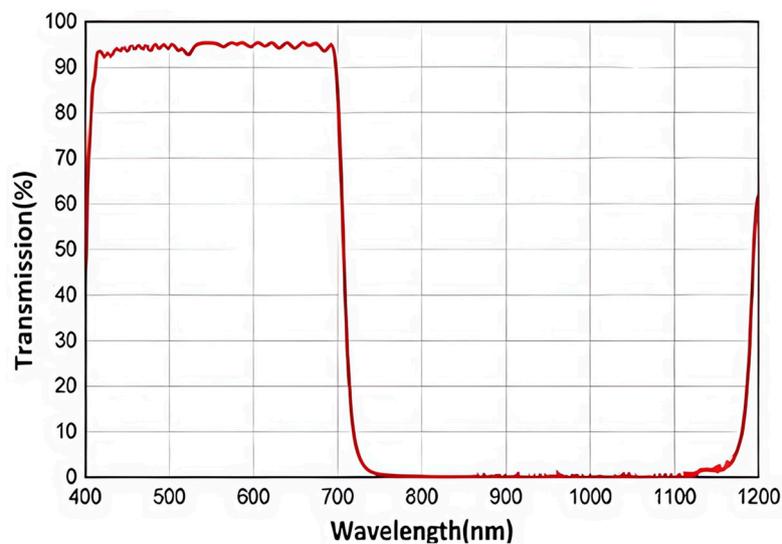
Figure 2. Relationship between transmittance and wavelength of the thermal mirror.

3.1. Image Acquisition

To acquire clear near-infrared images, this article utilizes 850nm near-infrared light as the source. The light-emitting diodes are embedded within a light guide plate to emit uniformly distributed near-infrared light. After NIR light is applied to the dorsal hand for venography, the imaging results are captured by a near-infrared camera after being reflected by a hot mirror. A narrowband filter with a wavelength of 850nm is installed in front of the camera, greatly minimizing the impact from visible light. Finally, the obtained image is transmitted to the processor for subsequent processing.

3.2. Lightweight U-Net Model Design with Embedded Gabor Kernel

3.2.1. Improved U-Net Architecture

The acquired image is subjected to binarization segmentation in the image processing module. This paper embeds prior knowledge into the segmentation network, fully utilizing the texture enhancement effect of Gabor filtering [27] to bolster the features in the shallow network. Specifically, we employ the U-Net for the segmentation network, which incorporate the multi-directional and multi-scale Gabor filter parameters into the encoder's shallow convolution kernels. This integration is shown in the red section of the network model depicted in Figure 3. After adjusting the overall channel number of the downsampling network, the convolutional operations are substituted with inverted residual blocks to minimize the semantic information loss, as depicted by the yellow segment of the network model. The improved U-Net's structure is shown in Figure 3.
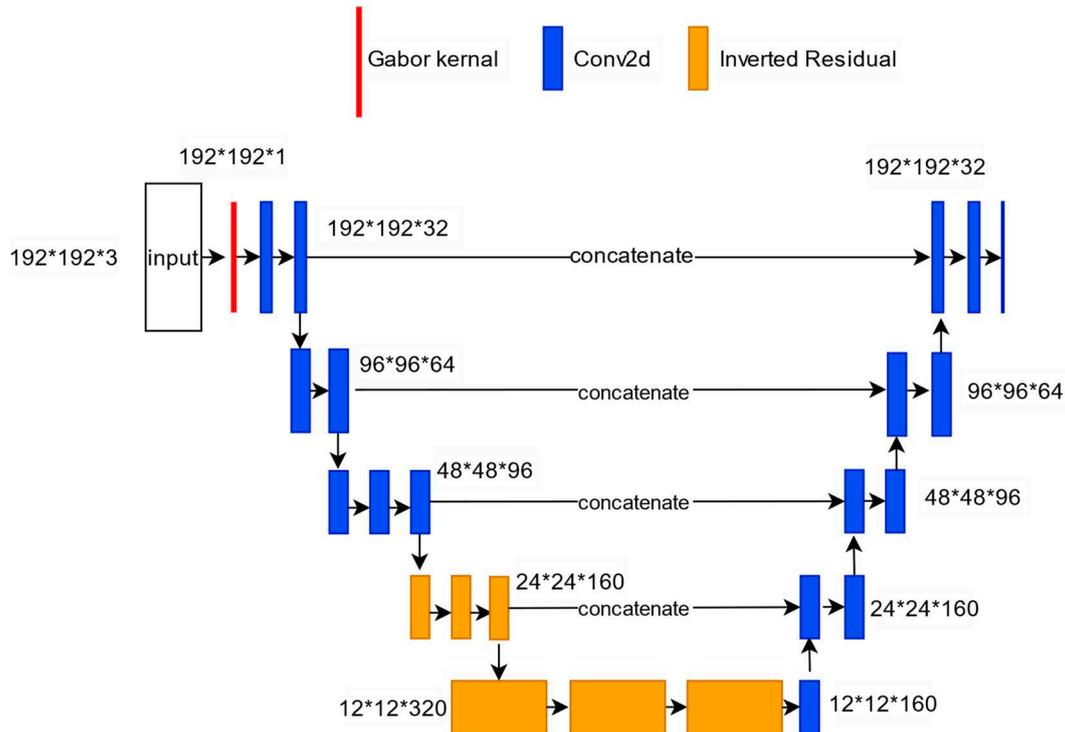
Figure 3. Improved U-Net architecture.

3.2.2. Embedding of Gabor Prior Knowledge

The process of convolving in a neural network is essentially a filtering process, and shallow layers of network primarily extract edge and texture information from the image. The Gabor filter, with its similarities in frequency and orientation to human vision, serves as an effective extractor of image features. Moreover, the two-dimensional Gabor transform has biological properties, and the image processed by its kernel function is highly similar to what is observed by higher animals [28]. Using the Gabor filter as a bandpass filter can eliminate noise while preserving the true structure information of ridges and valleys. Additionally, some trained shallow convolutional kernels in neural network are similar to the Gabor filter[29,30]. In order to extract richer image features, this paper embeds a Gabor layer in the first layer of the network. The Gabor kernel function is given by the following formula:

$$\text{Gabor} = \exp\left(-\frac{x' + \gamma^2 2y'^2}{2\sigma^2}\right)\exp\left(i * \left(2\pi\frac{x'}{\lambda} + \psi\right)\right) \tag{1}$$

$$x' = x\cos\theta + y\sin\theta \tag{2}$$

$$y' = -x\sin\theta + y\cos\theta \tag{3}$$

In this context, "i" denotes the imaginary unit, "ψ" signifies the phase parameter of the cosine function in the Gabor kernel, "θ" represents the orientation of the parallel stripes in the Gabor filter kernel, "λ" stands for the frequency of grayscale variation in the image, and "σ" denotes the standard deviation of the Gaussian factor in the Gabor function, which is related to "λ" as follows:

$$\frac{\sigma}{\lambda} = \frac{1}{\pi}\sqrt{\frac{ln2}{2}\left(\frac{2^b + 1}{2^b - 1}\right)} \tag{4}$$

The DC component represents the mean brightness value of the image and does not contain any frequency information. Removing the DC component in a Gabor filter can effectively reduce the influence of illumination or image background brightness. In the near-infrared light spectrum, there is little difference in the absorption rate of near-infrared light between the vein region and non-vein regions, resulting in severe contrast degradation and blurred vein edges. In the image, the non-vein

area acting as the background is relatively large. Subtracting the DC component is equivalent to performing a differential operation on the background, which enhances image contrast and makes vein texture more prominent. And for filters, it is best to keep the mean at 0 to maintain the original image intensity. This paper embeds the Gabor filter without DC component as a convolution kernel in shallow network, enhancing the features of image and helping the deeper network extract semantic information [31]. The formula for the Gabor convolution kernel without DC component is as follows:

$$Gabor = \exp\left(-\pi\frac{x' + \gamma^2 2y'^2}{2\sigma^2}\right)\left[\exp\left(i * \left(2\pi\frac{x'}{\lambda} + \psi\right)\right) - \exp\left(-\pi\frac{\sigma^2}{\lambda^2} + 2\psi\right)\right] \tag{5}$$

3.2.3. Lightweight Design of U-Net

In the course of downsampling in U-Net, different backbone networks may be employed as encoders, among which VGG16 is considered to be an exceptional model for feature extraction [32]. However, despite its efficacy in feature extraction, the VGG16 network's abundant channels may lead to an excessive number of parameters. It's necessary to reduce the network's weight while preserving its capacity for accurate image segmentation.

In this paper, channel numbers of each network layer have been reduced to decrease the number of parameters. However, reducing the channel numbers may lead to information loss, so we adopt the inverted residual block instead of the convolution method to mitigate this issue during downsampling. Unlike the residual block, the inverted residual block has a small head and a large middle structure. The ReLU activation function may cause significant loss of low-dimensional feature information [33], so the inverted residual block increases the dimension of the feature map before applying the ReLU activation function, thereby reducing information loss. The structure of inverted residual block and residual block [34] are shown in Figure 4 respectively.

The inverted residual block effectively addresses the issues of gradient disappearance and information loss. However, it is also more complex compared to the regular convolutional structure, which also elevates the computational burden on the model. The essence of this venous segmentation is semantic segmentation [35], and the processing of semantic information is mainly concentrated in the deep network. Hence, to minimize the loss of semantic information and reduce the processing time caused by complex structures, we exclusively employ inverted residual block in the last two blocks to substitute regular convolutions. This approach ensures that the model maintains its non-linear expressive capability while not increasing the image segmentation time.
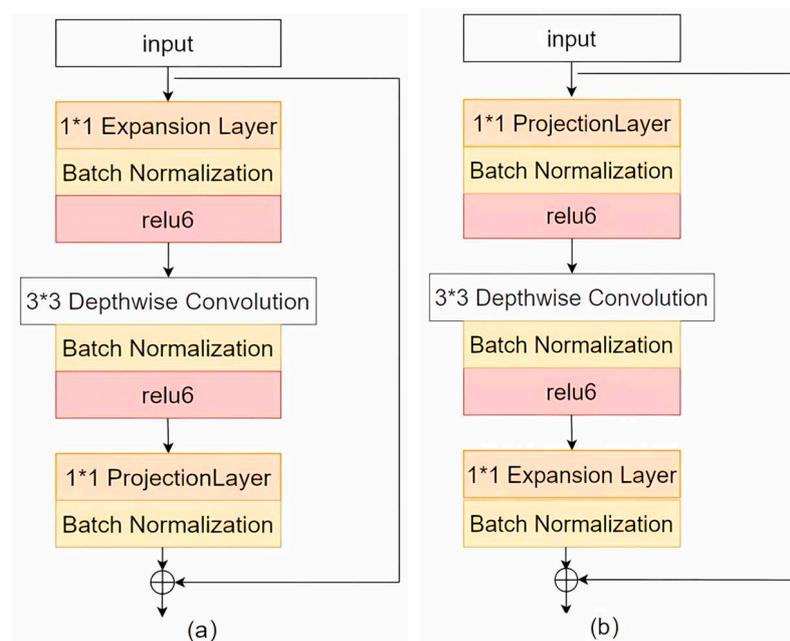

Figure 4. Differences in the two blocks. (a) The structure of inverted residual block and (b) residual block.

3.3. Coaxial Alignment Correction Algorithm Based on Optical Path

3.3.1. Coaxial Alignment of Optical Path

To align the projected image with the actual object, this paper proposes a method that combines coaxial correction and homography matrix. Through the adjustment of the optical path to mitigate the influence of the curved surface projection, and employing the homography matrix to refine the projected image, the resulting image is then perfectly aligned with the actual object.

The process of capturing images with a camera involves transforming the object from world coordinate system to pixel coordinate system, while the process of projection involves moving the image on pixel coordinate system to world coordinate system. Therefore, a projector can be viewed as an inverse camera [36]. To ensure that the projected image aligns perfectly with the object, the image in the inverse camera should be the same as the image captured by camera. According to the Zhang's camera calibration method, the image captured by the camera is essentially the projection of a three-dimensional object onto a two-dimensional plane.

When camera and inverse camera observe the same surface, the smaller the angle between their optical paths, the smaller discrepancy in the acquired surface information. When the optical paths of the two devices coincide, it is equivalent to projector and camera being in the same position. The image projected onto the surface is precisely the image captured by camera, and the impact caused by the projection on the curved surface is also reduced. Now, a string of information is printed on an A4 paper and attached to a bucket-shaped surface, and images captured by camera from angles A and B as shown in Figure 5.
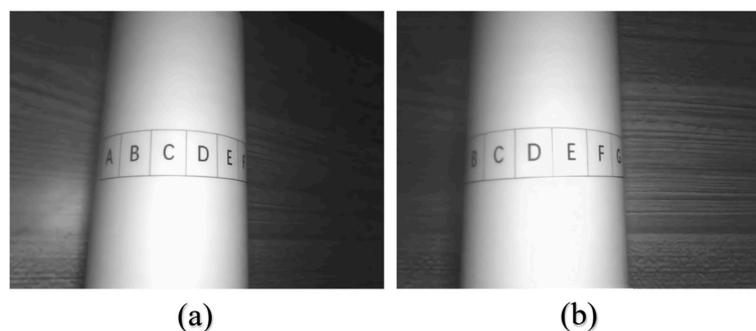

Figure 5. The surface information acquired by camera at different angles. (a) Image captured from angle A perspective and (b) which captured from angle B perspective.

From the above figure, it can be observed that due to the difference in position, the camera captures less information on the left side at angle B. Therefore, to reduce the impact of surface projection and minimize information loss, it is necessary to align the optical paths of both camera and projector. By using the property of the hot mirror, the near-infrared image can be reflected into the camera. Subsequently, the image is transmitted to the projector and projected as visible light, shining onto the dorsal hand through the hot mirror, as illustrated in Figure 6.
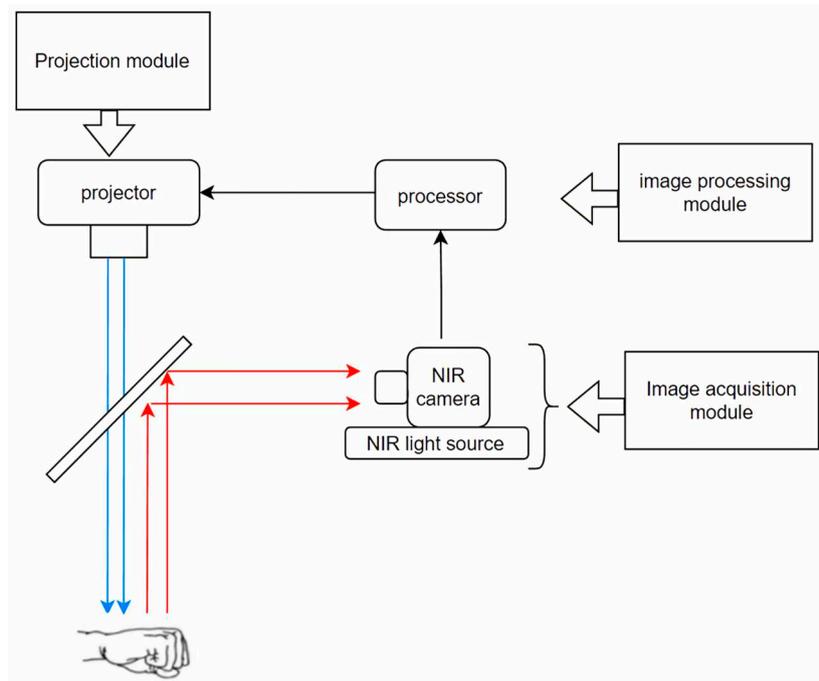
Figure 6. Schematic diagram of the light path.

3.3.2. Homography Matrix Calibration

During the installation and movement of the device, errors are easily generated. Therefore, adjusting the light path of the camera and projector using a hot mirror often results in an optical path near alignment but not coaxial. Although the error caused by surface projection can be reduced, the non-coaxial light path still results in a slight deviation between the projected image and actual object. To address this issue, this paper utilizes a plane with a printed grid pattern and homography matrix to calibrate the projection. The formula for the homography matrix is as follows:

$$\begin{bmatrix} x2 \\ y2 \\ z2 \end{bmatrix} = \begin{bmatrix} h1 & h2 & h3 \\ h4 & h5 & h6 \\ h7 & h8 & 1 \end{bmatrix} \begin{bmatrix} x1 \\ y1 \\ z1 \end{bmatrix} \tag{6}$$

The homography matrix is a mathematical tool used to describe the relationship between two planar transformations. By identifying four non-collinear corresponding points between the two planes, both linear and nonlinear transformations can be applied to the image. However, the projected image cannot be observed by the near-infrared camera, the four corners of the projection need to be manually marked on the flat surface. By using NIR camera, the positions of four corners on flat surface and projected image can be obtained. Subsequently, the homography matrix is computed as follows:

$$\begin{bmatrix} 1.09 & -1.62 * 10^{-2} & -7.92 * 10 \\ 3.03 * 10^{-2} & 1.07 & -5.05 * 10 \\ 3.72 * 10^{-5} & 1.86 * 10^{-5} & 1 \end{bmatrix} \tag{7}$$

The Frobenius norm of the homography matrix can be used to measure the size or scale of the matrix, and is calculated to be 93.96, this value is used to quantify the deviation between the projected image and the object, and assess the discrepancy between them. Then use the homography matrix to adjust the projected image to match the object. Images before and after adjustment are shown in Figure 7.
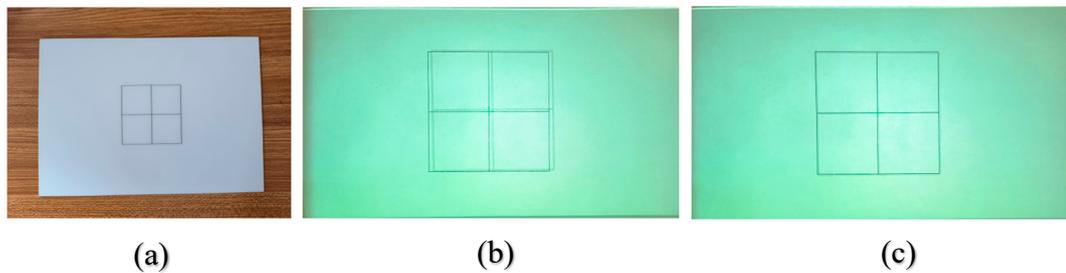
Figure 7. Projection of the image on the calibration plate after correction using homography matrix. (a) The calibration plate, (b) the original projection image and (c) the projection image after homography correction.

4. Experiment and Discussion

4.1. Image Segmentation.

4.1.1. Comparison with and without DC Component

This paper embeds fixed-parameter Gabor filters into shallow networks to enhance image features. This approach effectively alleviates the problem of semantic recognition difficulty in deep networks caused by insufficiently extracted features from shallow networks. Usually, ψ and γ are fixed at 0 and 1, respectively. To make the filter closest to the animal's vision system, the half-response spatial frequency bandwidth b is usually set to 1. The value of λ is determined by the width of the veins in the image [37], and can be used to calculate the value of σ. Based on vein features in the image, we created three different scales and eight different orientations of filters. By subtracting the DC component in the filter, the impact of lighting and image gray distribution is reduced, resulting in a frequency and directional visual representation closer to human perception. To better observe the results, we applied Gabor filters with and without DC components to process the near-infrared images of the dorsal hand. The original image and the filtered images with and without DC components are shown in Figure 8. It can be observed that the input image has better results after the Gabor filtering process without DC component.
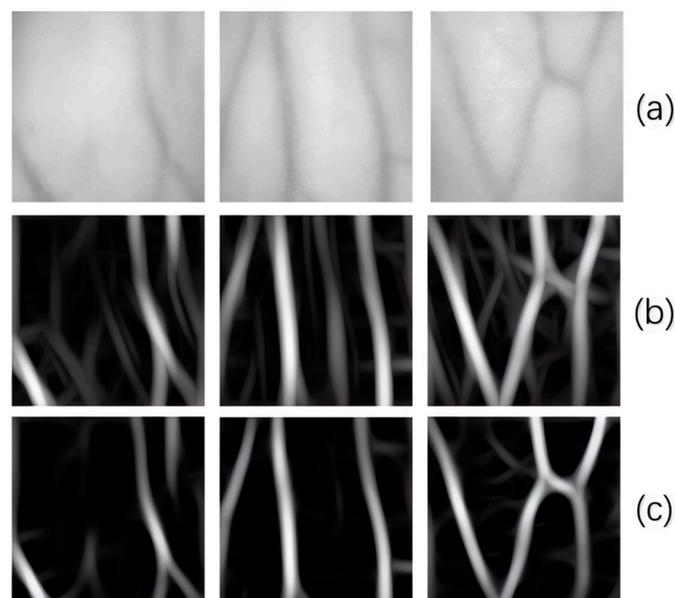

Figure 8. Comparison of infrared images after Gabor kernel filtering with and without DC components. (a) The NIR vein image, (b) the image filtered by Gabor, and (c) the image filtered by Gabor without DC component.

4.1.2. Comparison of Image Segmentation Results.

To train the neural network with the gathered images, it is imperative to generate suitable labels for the dataset. First, selecting different thresholds for each image to make coarse binary images. Then, the binary images are used as a mask for manually segmenting the veins, removing the veins that are too thin to be injected. The labels can be viewed in Figure 9.

Figure 9. Binary labels after segmentation.

We place the dataset and labels in the network to train, and observe the impact of Gabor kernel on the segmentation results. This paper conducts experiments using a network with Gabor kernel, Sobel kernel, and no kernels respectively. The obtained segmentation results are shown in Table 2. It can be observed that embedding Gabor kernels in the shallow convolutional networks significantly improves the segmentation results.

Table 2. Comparison of different convolutional kernel enhancement effects.

Convolution kernel	Miou	Precision
With Gabor convolution kernel	90.07 %	95.12 %
With Sobel convolution kernel	82.15%	90.64 %
Without convolution kernel	80.38%	87.55 %

To assess the performance of the improved network, comparative experiments were carried out using VGG16, Res-UNet, Dense-UNet, MobileNetv2 [38], and the proposed approach. During each network's training process, the loss function of both the training and validation sets were closely monitored to detect any signs of overfitting. After the completion of training, the test dataset was fed into the neural network for segmentation. The corresponding accuracy, size, and time taken for image segmentation by different networks were recorded as shown in Table 3.

Table 3. Different segmentation results of various networks comparison.

Downsampling Network	Miou	Precision	Size (M)	Consumption time (s)
VGG16	89.86 %	94.22 %	95	0.1963
Res-UNet50	88.64 %	92.79%	168	0.2147
Dense-UNet	90.08%	92.43%	111	0.3092
MobileNetv2	90.12 %	95.83 %	19	0.1513
Proposed method	**90.07 %**	**95.12 %**	**15**	**0.0910**

From the Table above, it can be observed that whether the IOU value or precision is used as the standard, the proposed method requires less time and takes up a smaller storage space while preserving the segmentation efficacy. The segmented results are shown in Figure 10. Examining the disparity between the segmented result and the labels, it is discernible that the error mainly exists at the ends of veins or at the connection points of two comparatively slender veins, which do not fall within the designated area for venipuncture.

Figure 10. Binary image after segmentation.

4.2. Results of Projection Correction

In this paper, we propose an improvement to the existing method, by using a hot mirror to adjust the optical path, which reduces the effect caused by curved projection. This is followed by an adjustment of the projection using a homography matrix, which reduces the requirement of the optical path coaxial to the experimental environment, and accurately projects the image onto the dorsal of the hand. Place hand on planes located at different distances from the camera to obtain the results. The images captured by the near-infrared camera, the pre-segmentation and post-segmentation images observed by the human eye are displayed in Figure 11.
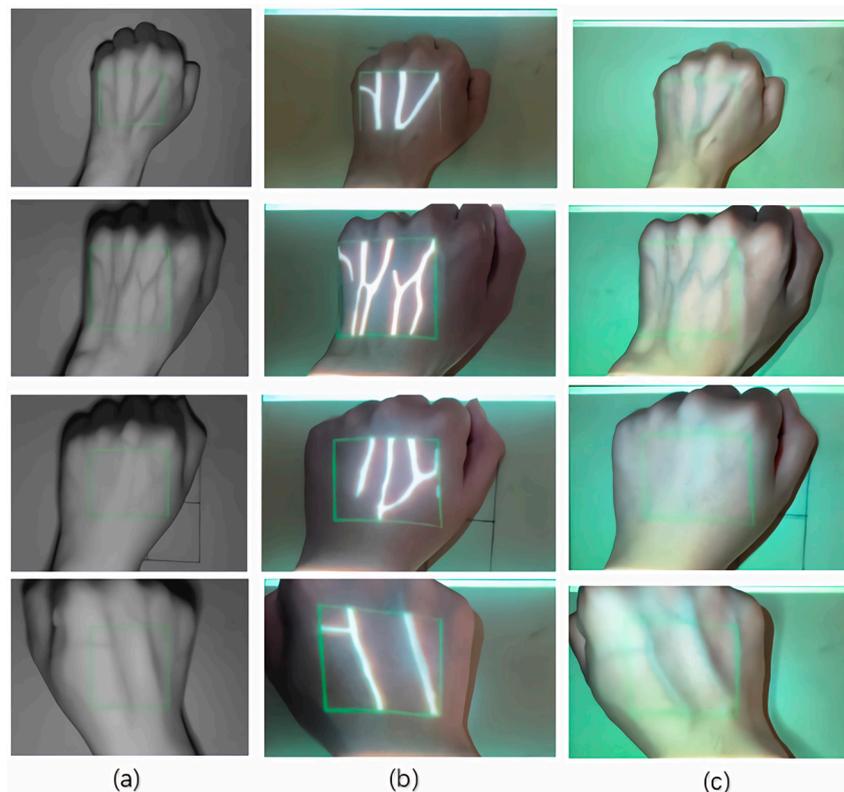

Figure 11. Imaging results of the dorsal hand on planes which located at different distances from the camera. (a) The image captured by a near-infrared camera, (b) the original image observed by mobile phone and (c) the segmented image projected onto the dorsal hand (c)

Above images show that the projection performs well both in the central and edge part of the hand back, and manifests no distortion caused by the curved surface. The projected binary vein image is consistent with the actual blood vessel. Although the homography matrix may alter with the plane's position, due to the prior proximal adjustment of the camera and projector's optical paths, the homography matrix will not change significantly.

As can be seen from Table 4, the offset errors of the projection in different planes are within 1mm, with the maximum average error being 0.53mm. Considering that the width of the venous blood is typically around 3-4mm, these errors can be considered negligible.

Table 4. Average error of adjusted projection.

L（mm）	0	370	340	310	280	250
x（mm）	0	0.12	0.26	0.36	0.45	0.53
y（mm）	0	0	0.07	0.08	0.11	0.14

5. Conclusions

Inspired by U-Net, this article has adjusted the structure of the encoder, but there is still room for improvement in the design of the image segmentation model. The experiments demonstrate that increasing the number of layers in the network has little impact on the segmentation results. Therefore, in the future, research can be conducted to lighten the network, coupled with specific network structures to reduce the parameter quantity and mitigate information loss. For projection correction, this article uses hot mirror and homography matrix to adjust the projection results, obtaining accurate results while reducing the precision requirements of the experimental environment. During the use of homography matrix correction, although the influence of hand position changes on the homography matrix has been reduced, there is still room for optimization in this method. When the distance between the plane and the camera changes, the nonlinear transformation of projected image far exceeds the linear transformation it undergoes. Therefore, by altering the distance and recording the position of the projector's focal point, it becomes feasible to establish a functional relationship between the two. When using the device, the distance can be obtained through binocular cameras or other devices, and based on the fitted function, the projection can be adjusted accordingly.

Conflicts of Interest: All authors declared that there are no conflicts of interest.

References

1. Indarwati, Ferika, Judy Munday, and Samantha Keogh. "Nurse knowledge and confidence on peripheral intravenous catheter insertion and maintenance in pediatric patients: A multicentre cross-sectional study." *Journal of Pediatric Nursing* 62 (**2022**): 10-16.
2. Jacobson, Ann F., and Elizabeth H. Winslow. "Variables influencing intravenous catheter insertion difficulty and failure: an analysis of 339 intravenous catheter insertions." *Heart & Lung* 34.5 (**2005**): 345-359.
3. He T, Guo C, Liu H, et al. Research on Robotic Humanoid Venipuncture Method Based on Biomechanical Model[J]. *Journal of Intelligent & Robotic Systems*, **2022**, 106(1): 31.
4. Sakudo, Akikazu. "Near-infrared spectroscopy for medical applications: Current status and future perspectives." *Clinica Chimica Acta* 455 (**2016**): 181-188.
5. Zeman, Herbert D, et al. "Prototype vein contrast enhancer." *Optical Engineering* 44.8 (**2005**): 086401-086401.
6. Huang, Zhenghua, et al. "Adaptive gamma correction based on cumulative histogram for enhancing near-infrared images." *Infrared Physics & Technology* 79 (**2016**): 205-215.

7.   Zhang, Zhengyou. "A flexible new technique for camera calibration." *IEEE Transactions on pattern analysis and machine intelligence* 22.11 (**2000**): 1330-1334.

8.   Liu, T., et al. "An algorithm for finger-vein segmentation based on modified repeated line tracking." *The Imaging Science Journal* 61.6 (**2013**): 491-502.

9.   Zhang, Jianfeng, Zhiying Lu, and Min Li. "Active contour-based method for finger-vein image segmentation." *IEEE Transactions on Instrumentation and Measurement* 69.11 (**2020**): 8656-8665.

10. Zhang H, He L, Wang D. Deep reinforcement learning for real-world quadrupedal locomotion: a comprehensive review. *Intelligence & Robotics.* 2022; 2(3):275-97.

11. Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." *Proceedings of the IEEE conference on computer vision and pattern recognition*. **2015**.

12. Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18.* Springer International Publishing, **2015**.

13. Liu, Manhua, and Peng Qian. "Automatic segmentation and enhancement of latent fingerprints using deep nested unets." *IEEE Transactions on Information Forensics and Security* 16 (**2020**): 1709-1719.

14. He, Tianbao, et al. "Automatic venous segmentation in venipuncture robot using deep learning." *2021 IEEE International Conference on Real-time Computing and Robotics (RCAR)*. IEEE, **2021**.

15. Chen Y, Ge P, Wang G, Weng G, Chen H. An overview of intelligent image segmentation using active contour models. *Intelligence & Robotics.* 2023; 3(1):23-55. http://dx.doi.org/10.20517/ir.**2023**.02

16. Jiang, Yun, et al. "Automatic retinal blood vessel segmentation based on fully convolutional neural networks." *Symmetry* 11.9 (**2019**): 1112.

17. Zeman, Herbert D, G. Lovhoiden , and C.Vrancken . "Prototype vein contrast enhancer." *Optical Engineering* (2004).

18. Dai, Xiaobin, et al. "A fast vein display device based on the camera-projector system." *2013 IEEE International Conference on Imaging Systems and Techniques (IST)*. IEEE, **2013**.

19. Gan, Qi, et al. "Benchtop and animal validation of a projective imaging system for potential use in intraoperative surgical guidance." *PloS one* 11.7 (**2016**): e0157794.

20. Van Tran, Tien, et al. "Design and enhance the vein recognition using near infrared light and projector." *VNUHCM Journal of Science and Technology Development* 20.K2 (**2017**): 91-95.

21. May, Hia Yee, and Ferda Ernawan. "Real Time Vein Visualization using Near-Infrared Imaging." *2020 International Conference on Computational Intelligence (ICCI)*. IEEE, **2020**.

22. Funk, Markus, Sven Mayer, and Albrecht Schmidt. "Using in-situ projection to support cognitively impaired workers at the workplace." *Proceedings of the 17th international ACM SIGACCESS conference on Computers & accessibility.* **2015**.

23. Gunawan, I. Putu Adi Surya, Riyanto Sigit, and Agus Indra Gunawan. "Vein visualization system using camera and projector based on distance sensor." *2018 International Electronics Symposium on Engineering Technology and Applications (IES-ETA)*. IEEE, **2018**.

24. Gunawan, I. Putu Adi Surya, Riyanto Sigit, and Agus Indra Gunawan. "Multi-Distance Veins Projection Based on Single Axis Camera and Projector System." *EMITTER International Journal of Engineering Technology* 7.2 (**2019**): 444-466.

25. Liu, Peng, et al. "A co-axial projection surgical navigation system for breast cancer sentinel lymph node mapping: system design and clinical trial." *Advanced Biomedical and Clinical Diagnostic and Surgical Guidance Systems XVII.* Vol. 10868. SPIE, **2019**.

26. Li, Chenmeng, et al. "Handheld projective imaging device for near-infrared fluorescence imaging and intraoperative guidance of sentinel lymph node resection." *Journal of Biomedical Optics* 24.8 (**2019**): 080503-080503.

27. Kumar, Ajay, and Grantham KH Pang. "Defect detection in textured materials using Gabor filters." *IEEE Transactions on industry applications* 38.2 (**2002**): 425-440.

28. [28] Yang, J., Zhou, H., Yang, B., et al. (**2018**). Improved algorithm and application of convolutional neural network based on Gabor kernel. *Journal of Yanshan University*, 42(5), 427-433.

29. Chen, Cheng, et al. "A learnable Gabor Convolution kernel for vessel segmentation." *Computers in Biology and Medicine* 158 (**2023**): 106892.

30. Luan, Shangzhen, et al. "Gabor convolutional networks." *IEEE Transactions on Image Processing* 27.9 (2018): 4357-4366.

31. Thoma, Martin. "A survey of semantic segmentation." arXiv preprint arXiv:1602.06541 (**2016**).

32. Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (**2014**).

33. Sandler, Mark, et al. "Mobilenetv2: Inverted residuals and linear bottlenecks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. **2018**.

34. He, Kaiming, et al. "Deep residual learning for image recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition.* **2016**.

35. Jalilian, Ehsaneddin, and Andreas Uhl. "Finger-vein recognition using deep fully convolutional neural semantic segmentation networks: The impact of training data." *2018 IEEE International Workshop on Information Forensics and Security (WIFS).* IEEE, **2018**.

36. Chaconas, Karen. "Range from triangulation using an inverse perspective method to determine relative camera pose." (**1990**).

37. Wang, Shupeng, et al. "Identification of structural parameters from free vibration data using Gabor wavelet transform." *Mechanical Systems and Signal Processing* 147 (**2021**): 107122.

38. Chazhoor AAP, Ho ESL, Gao B, Woo WL. Deep transfer learning benchmark for plastic waste classification. *Intelligence & Robotics.* 2022; 2(1):1-19. http://dx.doi.org/10.20517/ir.**2021**.15