

Finger-UNet: A U-Net Based Multi-Task Architecture for Deep Fingerprint Enhancement

Ekta Gavas^a and Anoop Namboodiri^b

Center for Visual Information Technology, International Institute of Information Technology, Hyderabad, India

Keywords: Fingerprint Enhancement, Fingerprint Quality, Image Enhancement, Multi-Task Learning.


Abstract: For decades, fingerprint recognition has been prevalent for security, forensics, and other biometric applications. However, the availability of good-quality fingerprints is challenging, making recognition difficult. Fingerprint images might be degraded with a poor ridge structure and noisy or less contrasting backgrounds. Hence, fingerprint enhancement plays a vital role in the early stages of the fingerprint recognition/verification pipeline. In this paper, we investigate and improvise the encoder-decoder style architecture and suggest intuitive modifications to U-Net to enhance low-quality fingerprints effectively. We investigate the use of Discrete Wavelet Transform (DWT) for fingerprint enhancement and use a wavelet attention module instead of max pooling which proves advantageous for our task. Moreover, we replace regular convolutions with depthwise separable convolutions, which significantly reduces the memory footprint of the model without degrading the performance. We also demonstrate that incorporating domain knowledge with fingerprint minutiae prediction task can improve fingerprint reconstruction through multi-task learning. Furthermore, we also integrate the orientation estimation task to propagate the knowledge of ridge orientations to enhance the performance further. We present the experimental results and evaluate our model on FVC 2002 and NIST SD302 databases to show the effectiveness of our approach compared to previous works.


1 INTRODUCTION

Fingerprints are one of the most crucial biometric traits due to their characteristic of being unique and permanent (lifelong) to every individual (Jain et al., 2004). In addition, they are comparatively easy to acquire (Jain et al., 2004). So, fingerprints are very commonly used biometrics for identification. Also, the recognition of fingerprints collected from crime scenes is helpful in criminal investigations and forensic applications. However, the acquisition of 'good' quality fingerprints is not trivial. A fingerprint is identified based on its unique ridge-valley structure with a well-defined frequency, orientation, and the location of special points called minutia. In reality, a fingerprint may be degraded due to various reasons like sensor defects, noise, oily skin, finger cuts/wounds, or uneven pressure while acquiring. It may have poor ridge structure, overlapping backgrounds, and low contrast due to collection from crime scenes. Due to these factors, the performance of fingerprint sys-

tems gets affected. Fingerprint enhancement deals with this issue by uplifting the quality of fingerprints to, at the least, recover the fingerprint structure to the maximum extent possible. Whether it is matching or recognition, fingerprint enhancement thus becomes an essential step in pre-processing in cases where good-quality prints are rare.

Traditionally, filtering techniques and other classical image processing methods were used to improve the ridge clarity, and fingerprint quality (Chikkerur et al., 2005; Greenberg et al., 2002; Hong et al., 1998; Kim et al., 2002; Yang et al., 2002). Then, with the advent of deep learning, neural networks, particularly convolutional neural networks (CNNs), are being employed to tackle this problem (Li et al., 2018; Qian et al., 2019; Joshi et al., 2019). In this paper, to recover good quality fingerprints, we explore a popular architecture, U-Net (Ronneberger et al., 2015), essentially an encoder-decoder-style architecture with skip connections. This paper focuses on improving the basic U-Net architecture in order to improve fingerprint quality in a robust and intuitive manner. We suggest that these changes improve the fingerprint quality in addition to reducing the network parameters.

^a  <https://orcid.org/0000-0001-6437-3357>

^b  <https://orcid.org/0000-0002-4638-0833>

Research studies show that CNNs, not being noise-robust, noise gets enlarged as data propagates through the layers of CNNs after several epochs of training which may significantly impact learning and can even lead to overfitting (Xie et al., 2019; Geirhos et al., 2018; Zhao et al., 2022; Li et al., 2021; Liu et al., 2019). The downsampling operation of CNNs is responsible for weak noise-robustness and loss of information. In the frequency domain, DWT can provide high-quality downsampling, significantly reducing this information loss. It decomposes the 2D image into four frequency components, and the noise-containing component is filtered and dropped; hence it is not forwarded into the network layers, avoiding noise propagation. In this paper, we use a Wavelet-Attention block proposed in (Zhao et al., 2022) in our U-Net architecture as a downsampling layer, which constrains the noise from high-frequency components obtained with DWT to propagate further, whereas information in low-frequency components is unaffected.

Further, we incorporate domain knowledge into our learning for the model to better understand the fingerprint structure. We achieve this by adding minutia-prediction and orientation estimation branches with a multi-task learning approach, which further helps to improve performance. Moreover, we replace standard convolution layers with depthwise separable convolutions (Chollet, 2017), which does not impact the performance considerably but helps to reduce the model parameters drastically.

1.1 Related Work

In past decades, several works have been proposed for fingerprint enhancement that involves applying traditional or classical image processing techniques. The research focused on improving the ridge structure and increasing the contrast in images (Greenberg et al., 2002). It started with most basic techniques like histogram equalization (Ezhilmaran and Adhiyaman, 2014) to using Fourier transforms (Sherlock et al., 1992; Chikkerur et al., 2005; Rahman et al., 2008) to remove noise from images. (Chikkerur et al., 2005) extended the use of short-term Fourier Transform (STFT) analysis to 2D fingerprint images and estimated the intrinsic properties of fingerprints. Gabor filtering was used to estimate the orientation and frequency fields which were later used to enhance fingerprints (Hong et al., 1998; Kim et al., 2002; Yang et al., 2002). (Liu et al., 2014) proposed a dictionary-based approach where dictionaries were created with a set of Gabor filters, and the multi-scale representation is iteratively applied to recover the enhanced

image. (Feng et al., 2012) proposed a path-based dictionary approach for orientation estimation.

For many years, much research has been carried out on neural networks, especially CNNs. With this, many past works proposed deep architectures for fingerprint enhancement. (Qian et al., 2019) introduced a deep network with dense blocks called DenseUNET to improve image quality in pixel-to-pixel and end-to-end manner. (Joshi et al., 2019) incorporated adversarial training using GANs for fingerprint enhancement. In another interesting work, (V and Sivaswamy, 2018) posed the fingerprint denoising problem as a segmentation (task) using M-net-based architecture. Few previous works had considered adding orientation knowledge to guide the enhancement task. (Li et al., 2018) proposed a deep architecture called FingerNet with enhancement and orientation deconvolution branches to enhance images using multi-task learning in two-stage training. They posed coarse orientation estimation as a classification problem with a quantized orientation field.

Generally, CNNs work with the matrix of pixels in the spatial domain. In contrast, the frequency domain deals with how these pixel values change in the spatial domain. Several mathematical transforms exist in the frequency domain, including Fourier, Laplace, Z, and wavelet transform. Wavelet transform is widely used in signal processing applications, image denoising (Kimlyk and Umnyashkin, 2018; Ismael et al., 2016), and compression (Chowdhury and Khatun, 2012; Kanagaraj and Muneeswaran, 2020). It decomposes the image information into signal details and approximations, commonly known as high-frequency and low-frequency components. Previous studies show several attempts have been made to incorporate DWT into CNNs (Bae et al., 2017; Liu et al., 2019; Li et al., 2021; Zhao et al., 2022). (Bae et al., 2017) showed that CNNs could benefit from learning about wavelet subbands and proposed a wavelet residual network (WavResNet). (Duan et al., 2017) applied dual-tree complex wavelet transform (DT-CWT) and designed a Convolutional-Wavelet Neural Network (CWNN) to suppress noise and extract features robustly from SAR images. (Liu et al., 2019) proposed a multi-level wavelet CNN (MWCNN) model which integrates wavelet transform into CNN to reduce feature map resolution and increase receptive field. (Li et al., 2021) designed DWT/IDWT layer for integration into deep networks. Later, (Zhao et al., 2022) modified these layers to include a wavelet attention module to retain detailed information in high-frequency components in DWT. In the domain of fingerprint biometrics also, research has been carried out to show the effectiveness of wavelet transform for fingerprint

enhancement task (Zhang et al., 2002; Hsieh et al., 2003). Hence, this paper combines the advantages of deep networks over traditional filtering techniques and wavelet attention module to design an architecture for fingerprint enhancement.

1.2 Contributions

We made the following contributions to this paper:

1. We propose a U-Net-based architecture with the DWT-based wavelet attention block for fingerprint enhancement.
2. We demonstrate that using minutia detection and orientation estimation branches in a multi-task manner can guide enhancement task to use domain knowledge to improve performance further.
3. We evaluate our model with publicly available datasets FVC 2002 and NIST SD302, and show that our model performs well in the context of structural similarity, fingerprint quality, and matching.

2 METHODOLOGY

The fingerprint enhancement task can be related to image denoising. However, we need to consider the inherent properties of biometric data and that it should be handled differently than normal real-world images. Autoencoder-style architectures are pretty popular for image-denoising tasks. We design an architecture called FingerUNET, based on the popular U-Net (Ronneberger et al., 2015) originally proposed for biomedical image segmentation tasks but has been shown to perform well on fingerprint enhancement task (Qian et al., 2019; Liu and Qian, 2020). In this paper, we offer a modified U-Net targeted to handle fingerprint data for enhancement task.

2.1 Wavelet Transform as Pooling

Wavelet transform can be said as a transformation that maps the signal to a multi-resolution representation. Wavelet has been combined with neural network for function approximation (Zhang and Benveniste, 1992), signal representation and classification (Szu et al., 1992). (Li et al., 2021) proposed DWT layers that decompose a 2D image into its frequency components. Here, the pooling is performed with down-sampling operation instead of max pooling or average pooling to avoid information loss and aliasing effect. It also increases the noise robustness of CNNs.

Given 2D data X , the DWT usually does 1D DWT on every row and column, resulting in four frequency components X_{LL} , X_{LH} , X_{HL} , and X_{HH} . X_{LL} is the low-frequency component of input X , representing the main information, including the basic structure in the image; X_{LH} , X_{HL} , and X_{HH} are three high-frequency components that save the horizontal, vertical, and diagonal details of X , respectively. (Li et al., 2021) designed these DWT/IDWT layers in Pytorch and made DWT/IDWT operations differentiable and compatible with CNNs.

$$\begin{aligned} X_{LL} &= LXL^T, X_{LH} = HXL^T, \\ X_{HL} &= HXL^T, X_{HH} = HXH^T, \end{aligned} \quad (1)$$

where matrix L and H are the cyclic matrix composed of wavelet low-pass filter $\{l_k\}_{k \in \mathbb{Z}}$ and high-pass filter $\{h_k\}_{k \in \mathbb{Z}}$ respectively. Here L and H are as in (Li et al., 2021). In FingerUNET, we make use of wavelet attention block (WA Block) proposed in (Zhao et al., 2022) built by modifying the above DWT layer. The wavelet attention block can be defined as

$$x_g = \sigma(f(X_{LH}, X_{HL})) \quad (2)$$

$$A_m = X_{LL} * x_g \quad (3)$$

$$Z = f(X_{LL}, A_m) \quad (4)$$

where f represents the feature aggregation function, σ denotes the softmax function. Once the four components are obtained using DWT, the WA block takes the horizontal feature X_{LH} and vertical feature X_{HL} and aggregates them by element-wise addition as a global detail feature. Then this feature is normalized using the softmax function. The normalized feature x_g and the low-frequency component X_{LL} are used to generate the attention map A_m through element-wise multiplication. Finally, the original low-frequency component X_{LL} is added to the attention map A_m by element-wise operation and given as output Z . As the component X_{HH} does not contain additional information for fingerprint ridges, it is not used in WA block. We replace the max-pooling layers in vanilla U-Net with this WA block and notice an improvement in the fingerprint enhancement task. IDWT layer is used to reconstruct the output back to spatial domain (Li et al., 2021).

2.2 Reconstruction Loss

The choice of the loss function is crucial in any neural network training. Earlier works (Burger et al., 2012; Dong et al., 2014) used mean squared error (MSE) or l_2 loss for reconstruction in image enhancement tasks. (Zhao et al., 2016) pointed out several limitations to using l_2 for image restoration tasks. l_2 does not correlate well with human perception of image

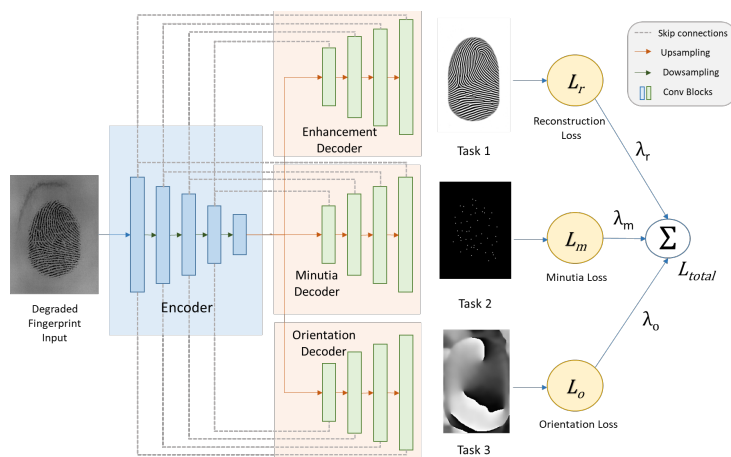


Figure 1: Architecture of our proposed approach with enhancement, minutia detection and orientation estimation branches in a multi-task learning setting.

quality (Zhang et al., 2012), due to the assumption that the impact of noise is independent of the local characteristics of the image. Moreover, l_2 penalizes larger errors but tolerates small ones without considering the underlying structure of the image. In contrast, the human visual system (HVS) is more sensitive to luminance and color variations in texture-less regions (Zhao et al., 2016; Winkler and Susstrunk, 2004). (Zhao et al., 2016) suggested using l_1 loss instead of l_2 as it does not over-penalize large errors and it also helps reduce artifacts introduced by l_2 . We utilize the effectiveness of using l_1 as a reconstruction loss for fingerprint enhancement. Even though the images are grayscale, we witness a performance improvement compared to the l_2 counterpart, as demonstrated in later sections.

2.3 Minutia Detection and Orientation Estimation

Every fingerprint has a well-defined ridge-valley structure and a frequency associated with it. The presence of specific patterns like whorl, loop, and arch in a fingerprint at different locations makes it unique. Ridge endings and bifurcations (minutiae) are important in fingerprint matching. Hence, fingerprint images prove to be very different from normal real-world images and hence they need to be dealt with differently. These properties of fingerprint data should not be neglected in learning. For this, adding domain knowledge to our network training becomes useful to model the data better. Moreover, in the case of low-quality images where the structure is not intuitive, the knowledge of fingerprint properties can help the network to learn and predict the structure better.

Previous works utilize ridge orientations when en-

hancing fingerprints (Hong et al., 1998; Kim et al., 2002; Yang et al., 2002; Li et al., 2018). Orientation estimation gives the direction of the gradient of the fingerprint segment which can be important in fingerprint enhancement. (Li et al., 2018) trained enhancement and orientation estimation branches in an end-to-end manner and demonstrated the effectiveness of orientation knowledge to aid the enhancement. The orientation estimation branch in our work predicts vectorized orientation fields, specifically sine and cosine values. This eliminates the need to segregate orientation patches into 20 fixed sets of classes (Li et al., 2018) and allows for the estimation of precise ridge orientation.

We also incorporate the minutia detection branch for using domain knowledge. (Darlow and Rosman, 2017) proposed a minutia extraction network called MENet, which outputs the minutiae probabilities map, which is later processed to determine precise minutia locations. To propagate the information about fingerprint properties to the enhancement branch, we use a gray-scale image with minutiae locations marked with a white dot on a black background as ground-truth as shown in the minutia branch (Task 2) image in Figure 1.

We approach the problem at hand using multi-task learning by optimizing fingerprint enhancement, minutia detection, and orientation tasks simultaneously. We use a shared encoder that learns the feature representations from noisy fingerprint images and one decoder for each of the three tasks, as shown in Figure 1. We use the l_1 loss mentioned in Section 2.2 for the enhancement task denoted by L_r . As we are solving orientation estimation as a regression problem, we use l_2 loss or MSE for this branch. We denote this loss by L_o . For the minutia detection branch, the values in the

ground-truth map are zero or one based on the presence of minutia locations, hence we use binary cross entropy as the loss function and it is denoted by L_m . The final loss L is the summation of L_r , L_m , and L_o , weighted by scalars λ_r , λ_m and λ_o respectively.

$$L_{total} = \lambda_r L_r + \lambda_m L_m + \lambda_o L_o \quad (5)$$

2.4 Operation Approximation with Depthwise Separable Convolution

Standard convolution operation applies spatial and channel interactions by multiplying values over several spatial pixels and all the channels. The idea of depthwise separable convolution (Chollet, 2017) is to disentangle the two by using each filter channel only at one input channel. Then, we use a 1×1 filter to cover the depth dimension. Though this method is an approximation to the usual convolution operation, it does not incur a considerable drop in performance. In addition, it reduces the number of model parameters by 30% in our experiments which in turn helps to avoid the over-fitting problem. Hence, in our work, we replace the convolution layers with depthwise separable convolutions.

3 EXPERIMENTS

3.1 Dataset

The fingerprint datasets available in the public domain are either not huge enough to train a deep network or do not contain enhanced/clean impressions for ground-truth. So, for all the experiments in this paper, we are using synthetically generated fingerprints from SFinGe (Cappelli et al., 2002). SFinGe can generate fingerprints with the required number of impressions and varying ridge structures and patterns. A wide range of noise and degradation like skin elasticity, noise, pressure, scratches, etc., can also be added. We generated 10,000 fingerprint pairs (degraded and ground-truth image pairs) for training, 1,000 for validation, and 3,000 for testing. The generated data contain varying types and degrees of available noises and backgrounds (optical, capacitive sensors, and no background). Figure 2 shows sample input and ground-truth pairs from the dataset in the first two rows. For evaluation, we use publicly available datasets FVC 2002 (Maio et al., 2002) and NIST Special Database (SD) 302 (Fiumara et al., 2019). FVC 2002 consists of fingerprints from optical, capacitive sensors, and synthetic fingerprints available in four

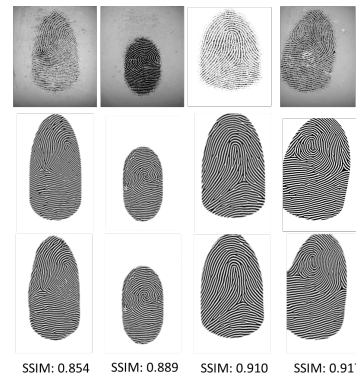


Figure 2: Enhanced images (bottom row) with corresponding degraded input (top row) and ground-truth (middle row) with our approach on SFinGe synthetic test set. SSIM is reported between each enhanced and ground-truth pair.

sets, of which 80 images are publicly available in each set. NIST SD302 has plain, rolled and touch-free impressions captured from various devices. We use the subset 302d containing 5141 fingerprint images acquired from 4 different auxiliary devices.

3.2 Training

We use the PyTorch framework for all experiments in this paper. We apply various data augmentations like random translation, rotation, flip (horizontal/vertical), and shear, for the network to generalize well. We apply the same degree of augmentations to each input-ground truth pair for consistency. The images are resized to fixed dimensions of 400×256 . The hyperparameters are chosen using grid search. The model is trained with Adam optimizer with a learning rate of 0.001. The batch size is set as 32. The loss weights λ_r , λ_m , and λ_o are 0.8, 0.1, 0.1 respectively. The network is trained on four GPUs in data parallel mode, and each GPU is NVIDIA GeForce TITAN X with 16 GB RAM. As the training is end-to-end, the losses from minutia and orientation guide the enhancement branch for better fingerprint output.

4 RESULTS AND ANALYSIS

For evaluation, we make use of standard metrics like SSIM, RMSE, and PSNR (V and Sivaswamy, 2018). Additionally, we utilize NFIQ2 (Tabassi et al., 2021) package from NIST's NBIS (Ko, 2007) to measure the quality of fingerprint images. Quality scores can range from 1 to 100. Moreover, we present the average matching scores of genuine pairs on all four subsets of FVC2002 using BOZORTH3 (Ko, 2007). SD302 does not contain multiple impressions of a



Figure 3: Illustration of enhancement results with our approach on samples from test datasets.

finger, so matching performance can not be obtained.

Ridge Structure Preservation: In Figure 2, we show a few sample images from our SFinge test set with corresponding ground-truth and enhanced images. We see the SSIM values are higher which suggests that our approach tries to preserve the ridge structure while performing enhancement on degraded input. In addition to this, Table 1 reports SSIM values for various combinations of techniques suggested in this paper. In Figure 3, we show the results of our approach on both datasets.

Fingerprint Quality Analysis: We report the average NFIQ2 scores on the test sets FVC 2002 and NIST SD302 in Table 2. From the results, we say that the fingerprint quality improved by a significant amount of 58% in the case of FVC 2002, whereas it improved by 23% in the SD302 dataset after enhancing the raw images with our approach. Further, our approach gives comparable results with the previous works. Moreover, we also present the NFIQ2 scores on the SFinGe test set in Table 1 which supports our approach to use WA Block, depthwise separable convolutions, and domain knowledge.

Matching Performance: We report and compare the average matching scores of genuine pairs of our approach with raw images and previous works in Table 3. The results suggest our approach with the inclusion of domain knowledge from the minutia and orientation branch is able to retain the minutiae from the degraded images which increases the matching score and performs well in comparison to earlier works. This suggests the effectiveness of our approach in the pipeline of fingerprint matching.

Ablation Study: We report the evaluation metrics for the techniques discussed in the paper as a part of the ablation study in Table 1. We show



Figure 4: Illustration of failure scenarios observed with our approach. Top row represents degraded input and bottom row denotes the enhanced images with our model.

how each discussed method in this paper helps to improve the enhancement further. The depthwise separable convolution do not significantly impact the performance but help reduce the model parameters to a large extent. Multi-task learning with minutia detection and orientation branches gives a clear performance improvement. Moreover, the use of wavelet attention block further improves the SSIM and NFIQ2. Overall, the combination of these techniques results in the best performing model.

Observed Challenges: We observed that our approach Finger-UNet performs well from our experimental results. However, we also observed scenarios where our approach did not perform well. A few of these cases are demonstrated in Figure 4. We saw that in cases where the input has severe artifacts, the model mistakes it for a portion of the fingerprint and tries to enhance it. Moreover, in cases where the input is too dark or too light to correctly figure out the ridge structure, the model fails to predict good fingerprints in those areas. In addition, if the input contains a nail, it fails to discriminate the nail from the fingerprint. We believe these issues arise as the network has not seen such data during training, as the SFinGe dataset does not contain samples with nails or artifacts.

Table 1: Ablation study: Evaluation performance on SFinGe test dataset with different modifications to U-Net suggested in this paper.

Approach	SSIM	RMSE	PSNR	NFIQ2
Raw Images	0.605	117.31	6.89	36.42
(l_2 loss)	0.863	43.67	12.78	47.72
(l_1 loss)	0.883	41.83	14.12	49.45
Depth. Sep. (DS)	0.890	41.25	14.17	49.78
WA Block (WA)	0.919	40.31	14.43	51.01
Minutia (M)	0.934	37.92	17.15	52.86
Orientation (O)	0.928	38.58	16.81	51.32
M+O	0.943	37.43	17.62	53.11
M+O+WA	0.954	36.91	18.47	55.32
M+O+WA+DS	0.955	36.78	18.81	55.34

Table 2: Average NFIQ2 scores of the images from FVC 2002 and NIST SD302 datasets. Higher scores represent higher fingerprint image quality.

Dataset	FVC 2002	NIST SD302
Raw Images	35.10	46.97
Joshi et al.	54.11	56.84
Hong et al.	56.01	58.23
Ours	56.26	58.49

Table 3: Average matching scores from BOZORTH3 on different subsets (DB1, DB2, DB3 and DB4) of FVC 2002 dataset. Higher the scores, better is the approach. Feature extraction was performed using MINDTCT.

Database	DB1	DB2	DB3	DB4
Raw Images	52.77	48.62	45.21	50.26
Joshi et al.	71.34	71.06	67.08	68.20
Hong et al.	73.52	72.31	69.12	70.29
Ours	74.01	72.61	69.10	71.08

5 CONCLUSION

In this paper, we modified vanilla U-Net, combining multiple techniques to improve the fingerprint quality with synthetic data from SFinGe. We evaluated our model on two public fingerprint datasets FVC 2002 and NIST SD302. The network is robust enough to recover fingerprints even with various degrees of degradation. From the experimental results, we say that l_1 loss performed well for this task, along with domain knowledge from minutia and orientation branches, which improved performance above baselines. We also discussed the challenging cases in our experi-

ments and the possible solutions. Moreover, using the wavelet attention block helped improve the performance. In future works, we plan to explore and design new quality metrics better suited for fingerprint enhancement along with other deep architectures.

REFERENCES

- Bae, W., Yoo, J., and Chul Ye, J. (2017). Beyond deep residual learning for image restoration: Persistent homology-guided manifold simplification. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*.
- Burger, H. C., Schuler, C. J., and Harmeling, S. (2012). Image denoising: Can plain neural networks compete with bm3d? In *2012 IEEE conference on computer vision and pattern recognition*. IEEE.
- Cappelli, R., Maio, D., and Maltoni, D. (2002). Synthetic fingerprint-database generation. In *2002 International Conference on Pattern Recognition*, volume 3. IEEE.
- Chikkerur, S., Govindaraju, V., and Cartwright, A. N. (2005). Fingerprint image enhancement using stft analysis. In *International Conference on Pattern Recognition and Image Analysis*, pages 20–29. Springer.
- Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*.
- Chowdhury, M. M. H. and Khatun, A. (2012). Image compression using discrete wavelet transform. *International Journal of Computer Science Issues (IJCSI)*, 9(4):327.
- Darlow, L. N. and Rosman, B. (2017). Fingerprint minutiae extraction using deep learning. In *2017 IEEE International Joint Conference on Biometrics (IJCB)*. IEEE.
- Dong, C., Loy, C. C., He, K., and Tang, X. (2014). Learning a deep convolutional network for image super-resolution. In *European conference on computer vision*. Springer.
- Duan, Y., Liu, F., Jiao, L., Zhao, P., and Zhang, L. (2017). Sar image segmentation based on convolutional-wavelet neural network and markov random field. *Pattern Recognition*, 64:255–267.
- Ezhilmaran, D. and Adhiyaman, M. (2014). A review study on fingerprint image enhancement techniques. *International Journal of Computer Science & Engineering Technology (IJCSET) ISSN*, pages 2229–3345.
- Feng, J., Zhou, J., and Jain, A. K. (2012). Orientation field estimation for latent fingerprint enhancement. *IEEE transactions on pattern analysis and machine intelligence*, 35(4):925–940.
- Fiumara, G. P., Flanagan, P. A., Grantham, J. D., Ko, K., Marshall, K., Schwarz, M., Tabassi, E., Woodgate, B., Boehnen, C., et al. (2019). Nist special database 302: Nail to nail fingerprint challenge.
- Geirhos, R., Rubisch, P., Michaelis, C., Bethge, M., Wichmann, F. A., and Brendel, W. (2018). Imagenet-trained cnns are biased towards texture; increasing

- shape bias improves accuracy and robustness. *arXiv preprint arXiv:1811.12231*.
- Greenberg, S., Aladjem, M., and Kogan, D. (2002). Fingerprint image enhancement using filtering techniques. *Real-Time Imaging*, 8(3):227–236.
- Hong, L., Wan, Y., and Jain, A. (1998). Fingerprint image enhancement: Algorithm and performance evaluation. *IEEE transactions on pattern analysis and machine intelligence*, 20(8):777–789.
- Hsieh, C.-T., Lai, E., and Wang, Y.-C. (2003). An effective algorithm for fingerprint image enhancement based on wavelet transform. *Pattern Recognition*, 36(2):303–312.
- Ismael, S. H., Mustafa, F. M., and Okümüs, I. T. (2016). A new approach of image denoising based on discrete wavelet transform. In *2016 World Symposium on Computer Applications & Research (WSCAR)*, pages 36–40. IEEE.
- Jain, A. K., Ross, A., and Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on circuits and systems for video technology*, 14(1):4–20.
- Joshi, I., Anand, A., Vatsa, M., Singh, R., Roy, S. D., and Kalra, P. (2019). Latent fingerprint enhancement using generative adversarial networks. In *2019 IEEE winter conference on applications of computer vision (WACV)*, pages 895–903. IEEE.
- Kanagaraj, H. and Muneeswaran, V. (2020). Image compression using haar discrete wavelet transform. In *2020 5th International Conference on Devices, Circuits and Systems (ICDCS)*. IEEE.
- Kim, B.-G., Kim, H.-J., and Park, D.-J. (2002). New enhancement algorithm for fingerprint images. In *2002 International Conference on Pattern Recognition*, volume 3, pages 879–882. IEEE.
- Kimlyk, M. and Umnyashkin, S. (2018). Image denoising using discrete wavelet transform and edge information. In *2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EICoRus)*. IEEE.
- Ko, K. (2007). User’s guide to nist biometric image software (nbis).
- Li, J., Feng, J., and Kuo, C.-C. J. (2018). Deep convolutional neural network for latent fingerprint enhancement. *Signal Processing: Image Communication*.
- Li, Q., Shen, L., Guo, S., and Lai, Z. (2021). Wavecnet: Wavelet integrated cnns to suppress aliasing effect for noise-robust image classification. *IEEE Transactions on Image Processing*, 30:7074–7089.
- Liu, M., Chen, X., and Wang, X. (2014). Latent fingerprint enhancement via multi-scale patch based sparse representation. *IEEE Transactions on Information Forensics and Security*, 10(1):6–15.
- Liu, M. and Qian, P. (2020). Automatic segmentation and enhancement of latent fingerprints using deep nested unets. *IEEE Transactions on Information Forensics and Security*, 16:1709–1719.
- Liu, P., Zhang, H., Lian, W., and Zuo, W. (2019). Multi-level wavelet convolutional neural networks. *IEEE Access*, 7:74973–74985.
- Maio, D., Maltoni, D., Cappelli, R., Wayman, J. L., and Jain, A. K. (2002). Fvc2002: Second fingerprint verification competition. In *2002 International Conference on Pattern Recognition*, volume 3. IEEE.
- Qian, P., Li, A., and Liu, M. (2019). Latent fingerprint enhancement based on denseunet. In *2019 International Conference on Biometrics (ICB)*, pages 1–6. IEEE.
- Rahman, S. M., Ahmad, M. O., and Swamy, M. (2008). Improved image restoration using wavelet-based denoising and fourier-based deconvolution. In *2008 51st Midwest Symposium on Circuits and Systems*, pages 249–252. IEEE.
- Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer.
- Sherlock, B., Monro, D., and Millard, K. (1992). Algorithm for enhancing fingerprint images. *Electronics letters*, 18(28):1720–1721.
- Szu, H. H., Telfer, B. A., and Kadambe, S. L. (1992). Neural network adaptive wavelets for signal representation and classification. *Optical Engineering*, 31(9):1907–1916.
- Tabassi, E., Olsen, M., Bausinger, O., Busch, C., Figlarz, A., Fiumara, G., Henniger, O., Merkle, J., Ruhland, T., Schiel, C., and Schwaiger, M. (2021). Nist fingerprint image quality 2.
- V, S. A. and Sivaswamy, J. (2018). Fpd-m-net: Fingerprint image denoising and inpainting using m-net based convolutional neural networks. *CoRR*, abs/1812.10191.
- Winkler, S. and Susstrunk, S. (2004). Visibility of noise in natural images. In *Human Vision and Electronic Imaging IX*, volume 5292, pages 121–129. SPIE.
- Xie, C., Wu, Y., Maaten, L. v. d., Yuille, A. L., and He, K. (2019). Feature denoising for improving adversarial robustness. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*.
- Yang, J., Liu, L., and Jiang, T. (2002). Improved method for extraction of fingerprint features. In *Second International Conference on Image and Graphics*, volume 4875, pages 552–558. SPIE.
- Zhang, L., Zhang, L., Mou, X., and Zhang, D. (2012). A comprehensive evaluation of full reference image quality assessment algorithms. In *2012 19th IEEE International Conference on Image Processing*. IEEE.
- Zhang, Q. and Benveniste, A. (1992). Wavelet networks. *IEEE transactions on Neural Networks*, 3(6).
- Zhang, W.-P., Wang, Q.-R., and Tang, Y. Y. (2002). A wavelet-based method for fingerprint image enhancement. In *Proceedings. International Conference on Machine Learning and Cybernetics*, volume 4. IEEE.
- Zhao, H., Gallo, O., Frosio, I., and Kautz, J. (2016). Loss functions for image restoration with neural networks. *IEEE Transactions on computational imaging*, 3(1):47–57.
- Zhao, X., Huang, P., and Shu, X. (2022). Wavelet-attention cnn for image classification. *Multimedia Systems*, 28(3):915–924.