International Relations Team
291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea
Tel: +82-42-350-2441~9, Fax: +82-42-350-4930, E-mail: irt@kaist.ac.kr
http://io.kaist.ac.kr

2023 KISS Track 2: Research Report

Student Name	Jayan Patel	**Home University**	University of Leeds
Student ID No.		**Degree**	MS
Department	School of Computing	**Academic Advisor**	Prof. Sungho Jo
Participation Period	2023.07.03-2023.07.28	**Laboratory Name**	Neuro-Machine Augmented Intelligence Laboratory
Project Title	Investigation of Connected U-Net Architectures for Damaged Fingerprint Restoration		

Overview of Research

1. Research Abstract

Fingerprint alteration is a common method taken to evade identification. This paper investigates learning-based methods to restore these damaged fingerprints with the intention of improving how well they can be recognised. Three U-Net structures are developed and designed in this paper – a single U-Net, a double U-Net (two U-Nets in series) and a connected U-Net (connection of the feature maps of the first U-Net into the second U-Net). The models were trained and tested in restoring the damaged fingerprints in the SOCOFing dataset. All models showed strong performance in restoring the obliterated fingerprints, but they could not restore the centrally rotated or Z cut images well. The double U-Net generally showed the strongest performance across difficulties and damage types, but its increased performance was very slight. Obliteration damage is quite noisy in nature and since the performance was much better restoring that, it is possible that stable diffusion could be a better approach to improve the restoration of Z cut and centrally rotated fingerprints as well. More research is also required to evaluate the benefits of restoration in the identification of damaged fingerprints.

2. Research Participation Records

Week	Research Participation
1st week	Introduction to neural networks. Studied theory on fully connected and convolutional neural networks and made some neural networks to solve basic classification tasks.
2nd week	Started to investigate U-Net architectures, implementing existing code into the fingerprint restoration problem. Researched existing literature on the topic.
3rd week	Continued to develop the models for the project. This week involved lots of debugging to ensure the models were working and that the dataloader was giving fairly distributed data. Once correctly implemented, started to train the models and optimise them for the best results.
4th week	Data collection of the developed models. Ran several experiments to isolate information on the performance on each of the different models, damage levels and damage types. Started writing the report.

International Relations Team
291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea
Tel: +82-42-350-2441~9, Fax: +82-42-350-4930, E-mail: irt@kaist.ac.kr
http://io.kaist.ac.kr

Confirmation by Student, Academic Supervisor, and Mentor-Student		
By signing below, I hereby confirm that the information contained in this form is true, correct, and complete.		
Date : 2023.08.02.	**Name of the student :** Jayan Patel	**Signature:**
Date : 2023.08.03.	**Name of the advisor :** Sungho Jo	**Signature:**
By signing below, I hereby confirm that I provided support for the student's research participation.		
Date : 2023.08.03.	**Name of the mentor-student :** Hochang Lee	**Signature:**
This report is to be submitted no later than one week after the end of the program.		

Investigation of Connected U-Net Architectures For Damaged Fingerprint Restoration

1. Introduction

Fingerprint recognition is a common form of identification in modern society. They serve as unique biometric identifiers and maintain their form throughout a person's life (Fattahi & Mejri, 2020). For over 100 years fingerprint recognition has been used as a form of identification by law enforcement agencies (Jain & Kumar, 2012). Its use has further increased as technology has developed and fingerprint readers have become more widely used such as on mobile devices, in workplaces and when travelling. As a result of this, criminals are attempting to evade identification by damaging their fingerprints. They can damage their fingerprints in several different ways, so it is becoming increasingly important to have a way of effectively identifying damaged fingerprints. The aim of this research is to investigate learning-based methods to perform restoration of damaged fingerprints so that their original form can be reproduced, with the intention of improving the ability at which damaged fingerprints can be identified. This paper focuses on the restoration part of this process, developing and evaluating the performance of three different U-Net-based models and discussing the findings.

2. Related Works

The original U-Net architecture was presented by Ronneberger et al. (2015) and it was designed for biomedical image segmentation. It was developed to exceed the performance of convolutional neural networks (CNNs). CNNs were popular at the time for visual recognition tasks due to their ability to maintain the spatial information of images, as opposed to linear neural networks that flattened input image tensors, thus losing the relationship between neighbouring pixels. The model was trained such that when given an image it would be able to classify each of the pixels within the image to a certain class. In the biomedical domain this usually meant classification of cell type or biological structure, but the model will work in a similar way for any segmentation task, it just depends on the training material. It is important to note the difference in designated purpose of this architecture. As the U-Net was designed for segmentation, it was trained in a different way to conventional image restoration models. The output of the model was compared to a segmented image mask as opposed to the output image being compared to the original image.

U-Nets have been used in image restoration, however, much of this research revolves around denoising. Yan et al. (2022) presented an investigation of a more advanced U-Net model that uses transformers, and this too was designed for denoising and deraining of pictures. They point out the weaknesses of standard U-Net structures and other fully convolutional networks. Due the locality of the convolution processes, long range semantic information of the information can be lost. The U-Net structure does try to address this with the use of max pooling and skip connections, however, it is still prone to those problems. Thus, this could pose a problem in the restoration of the fingerprints in this tasks due to the importance of long-range information around the damaged area. Yan et al. (2018) presented a development of the U-Net structure made for image inpainting. Named shift net, they introduced a shift connection layer into the decoding path. This was able to recognise the long-range semantic structure of the input image very well, as well as the finer textures and details within the image, and so could inpaint blanked out regions of pictures. This shows that there are methods that enable the U-Net structure to recognise long range semantic structure of images. The complexity of the shift net implementation meant that it was not investigated in this paper. Furthermore, inpainting differs to the restoration of fingerprint. As fingerprint restoration involves overlaying existing structure and features within the image, it is unknown whether shift net would work. No research could be found using U-Nets for this kind of restoration.

3. Methods

3. 1 Dataset

The dataset used for training and testing of the model was the Sokoto Coventry Fingerprint, or SOCOFing (Shehu et al., 2018). The SOCOFing database has the ten fingerprints of 600 subjects scanned. Synthetic alterations were then made on each of the fingerprints using the STRANGE toolbox, thus creating 'damaged' fingerprints. The types of damages made were obliteration, Z cut and central rotation. Each of these were carried out to varying extents, thus the altered fingerprints are split into easy, medium and hard levels of damage. These can be seen in figure 1. Some of the damaged fingerprints were discarded as they were not deemed suitable for use, so the final size of the dataset is 55,273 (49,273 altered images). All pictures were labelled with the subject number, finger and damage type if damaged. The images were originally sized as

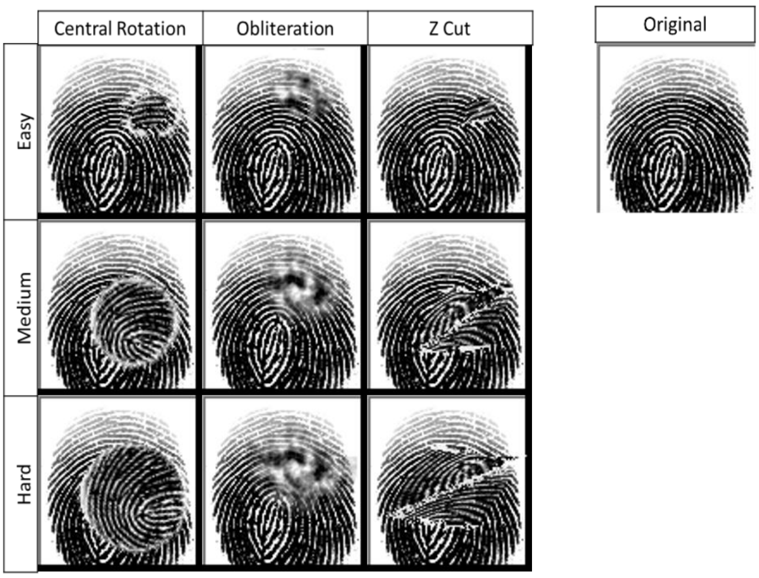

Figure 1. An example of the different types and levels of damages performed on the original fingerprints within the SOCOFing dataset.

1x96x103 bitmap files however these were transformed to 1x200x200 and then cropped to 1x180x180 to minimise the border that surrounds each of the images. The dataset was split with an approximate 90/10 training/test split, resulting in a training set of 44,344 and a training set of 4,929. Each of the easy, medium and hard images were also split with the same ratio.

3.2 Architectures

The structure of the U-Net in this paper is very similar to that of the original U-Net developed by Ronneberger et al. (2015). The code was taken from the repository of milesial (2021; 2023). There is a 3x3 convolution, a 2D batch normalisation followed by a rectified linear unit (ReLU). This is repeated twice and then a 2x2 max pooling is performed. This is done 4 times during the encoding. The differences between this structure against Ronneberger et al.'s is that the 3x3 convolutions were padded to ensure the output image was the same size as the input and the inclusion of a 2D batch normalisation. Each of the 4 stages of the decoding process involved an upscaling with scale factor 2, followed by the concatenation with the corresponding feature map from the encoding and a 2x2 convolution to half the number of channels. 2 3x3 convolutions, 2D batch normalisations and ReLUs were then performed.

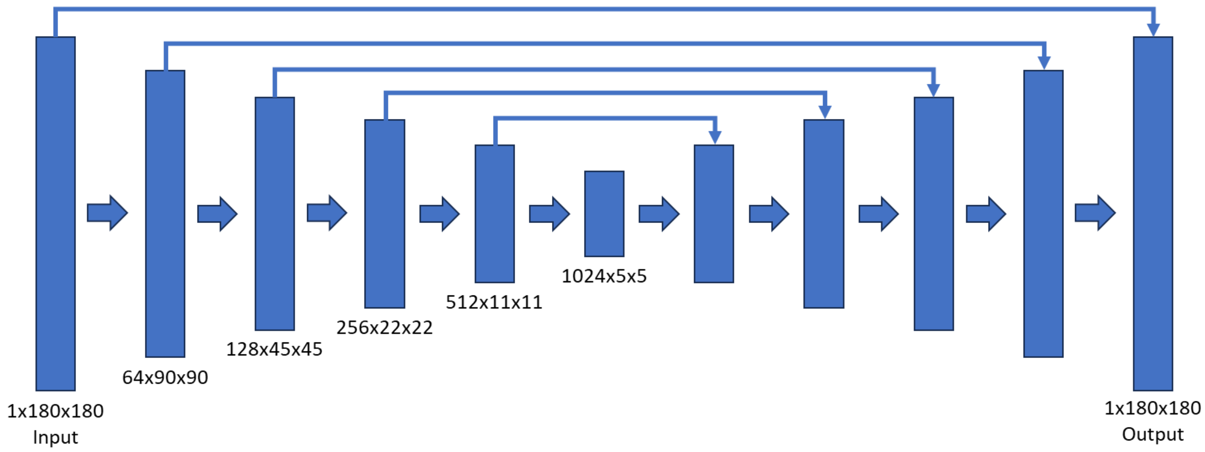

Figure 2. The single U-Net structure, labelled with the image sizes through each stage of the process. The connections by way of concatenation are depicted as arrows.

Three different U-Net-based architectures were evaluated. All models and the training and testing were implemented using the PyTorch library.

- The single U-Net, the structure of which has just been outlined. This can be seen in figure 2.
- The 'double U-Net'. This uses two single U-Net structures in series. Each of the U-Nets were no different in their architecture to the single U-Net but they were trained differently. This can be seen in figure 3.
- The 'connected U-Net'. This too uses two different U-Nets, but the feature maps of the first U-Net

are connected to the second U-Net and concatenated in the decoding phase. This can be seen in figure 4.

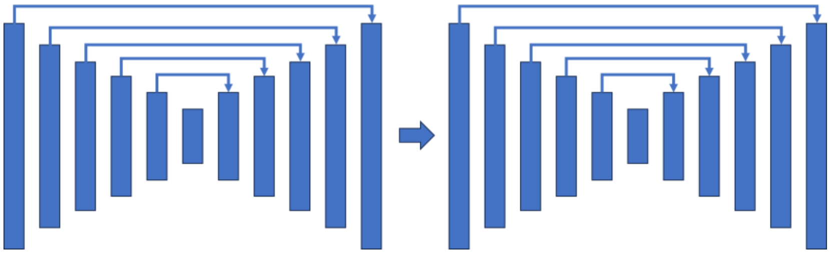

Figure 3. The double U-Net structure. This structure was made of two single U-Net structures in series.

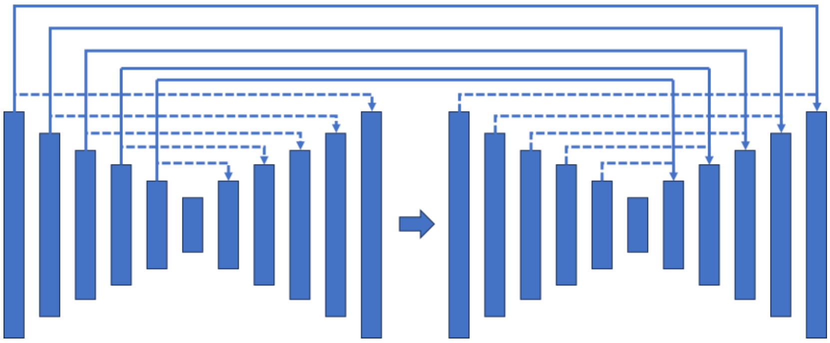

Figure 4. The connected U-Net structure. The second U-Net in this structure was modified to perform concatenate not only the feature maps of that U-Net, but also those of the first U-Net. The concatenation connections are depicted as arrows.

For the second U-Net within the connected U-Net structure, the model was changed to take an input of the feature maps from the contracting path of the first U-Net. In the second U-Net, during the concatenation stages of the decoding process, the corresponding tensor from the input tensors was also concatenated and the 2x2 convolution reduced the number of channels by 3 times.

3.3 Training

For training the mean squared error (MSE) loss was used. MSE loss calculates the mean squared difference between each of the pixels from an image when compared to a target image. As it uses mean squared difference, pixel values greater in difference are more largely penalised. Whilst Ronneberger et al. (2015) used cross entropy loss, their U-Net was made for segmentation and not restoration. In image processing, MSE loss is commonly used due to its ability to be differentiated making optimisation and backpropagation easier to implement. It is known not to translate well to the human visual system (HVS) (Zhao et al., 2016), however the image restoration task being investigated is also something the HVS is not designed for. Thus, due to its popularity and ease of use, it was decided as the loss function of choice for this research.

Using the training set of images, the single U-Net was trained for 20 epochs, each epoch being randomly shuffled. A scheduler was used to control the learning rate, starting at 0.001 and reducing by 0.0002 after every 5^{th} epoch. The Adam optimiser was used. The total loss after each batch was calculated and the model accordingly optimised. The single U-Net was first trained, and it was then used to train the second U-Nets in the double and connected architectures. In the double U-Net architecture, the input images were put through the trained single U-Net model. The output of this model was then used as the input to the second U-Net. For the connected architecture, the input images were put into the trained single U-Net model and the feature maps from the encoding were saved. These feature maps were then input to the second U-Net as well as the output images of the first U-Net. Both of the second U-Nets in the double and connected models used the same learning rate scheduler and optimiser as in the single U-Net training.

4. Results

Once trained, each of the models were tested on the testing dataset. The MSE loss of the final output images compared to the target images are presented throughout this section as well as the outputted images from the models. Table 1 displays the MSE loss values of each of the models for the whole test set as well as for each of the levels of damage. The performance between the various architectures appears to show no clear difference in performance. Across difficulty levels, there is an expected increase in the loss values as the level of damage is increased. The double u-net structure does show slightly lower values across all difficulty levels and for the entire dataset, although the difference is small.

Table 1. Table presenting the MSE loss values of each of the models, and the input images without having gone through a model, for each damage level.

Model	All levels	Easy	Medium	Hard
No model	0.0191	0.00797	0.0215	0.0304
Single	0.00625	0.00191	0.00728	0.0105
Double	**0.00591**	**0.00176**	**0.00687**	**0.00998**
Connected	0.00616	0.00188	0.00718	0.0103

The difference in network performance for each of the damage types is displayed in table 2. All models were able to reduce the loss of the obliterated fingerprints considerably more than for the Z cut and centrally rotated models. The reduction in loss for the obliterated models was roughly 5 times less, whereas the centrally rotated and Z cut fingerprints only saw a reduction of roughly 2 times. For each damage type, the losses achieved by each model were very similar.

Table 2. Table presenting the MSE loss values of each of the models, as well as the input images without having gone through a model, for each damage type.

Model	CR	Zcut	Obl
No model	0.0262	0.0132	0.0179
Single	0.0110	0.00630	0.00222
Double	**0.0104**	**0.00587**	0.00211

Connected	0.0109	0.00624	**0.00209**

Pictures of the hard damage restorations of each type can be seen in figure 5. The difference in results between each of the damage types matches the trends of the loss values measured. The ability of all u-net structures to repair the damage of the obliteration is very strong. Looking only at the outputted regions it is difficult to distinguish which regions had been damaged – there is slight 'smudging' in the middle area of damage, but the outer regions are well defined. Both the Z cut and centrally rotated fingerprints were unable to be restored and instead a grey region was produced over the damaged area. Within table 2, the loss values of the input images are displayed, without having gone through a model. This shows that the loss value of the centrally rotated prints was the highest, however the obliterated fingerprints had a higher loss than the Z cut fingerprints. Hence this proves that the improved restoration of the obliterated fingerprints was not because the obliteration required a smaller change in MSE value but due to other factors, likely the form of the damage.

Examining the z cut and central rotation outputs in figure 5, it is visible that there is slightly more fingerprint restoration done by the double U-Net architecture and less of a grey region when compared to the output images for the other architectures. Whilst, visually, the difference is small this does agree with the loss values recorded from the tests.

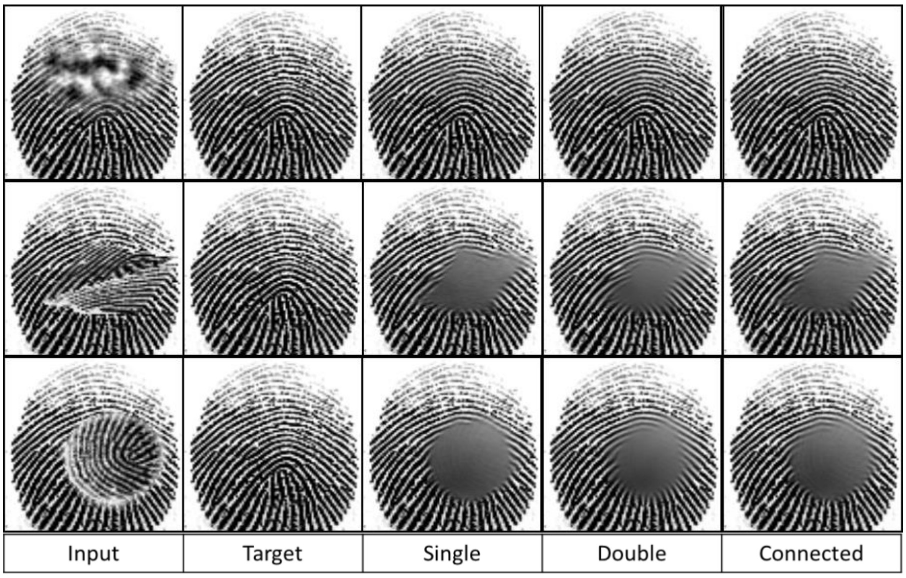

Figure 5. Images of a select fingerprint input, target, and hard level restoration for each damage type.

The data and images appear to present a threshold in the loss values - given that, visually, there is not much difference in the Z cut and centrally rotated results of figure 5, there may be some value of MSE loss between 0.006 and 0.002 below which the similarity between the outputted image and the target largely increases.

5. Discussion

The double U-Net showed the most promising results from testing. Resultant of it having lower loss values and visibly showing more restoration it is does raise the question whether a larger series of U-Net structures could be made and further improve the results, or whether there is a point at which the improvements start to slow down. Assuming that a larger series would improve performance, looking at the results of the obliterated damage, it appears that an MSE loss value of around 0.002 would need to be reached for true restoration to be achieved. It is unknown how the MSE would reduce with each additional U-Net, but it would take several more U-Nets in the series to reduce the loss of the double U-Net, 0.0058, to around 0.002. Whilst training each extra U-Net in the model is not too memory intensive as each model is trained individually, the use of a multi-U-Net series architecture once trained would be quite computationally heavy and potentially not the most efficient method of restoration.

The enhanced performance of the obliteration restoration over the centrally rotated and Z cut fingerprints could provide insight into how best to utilise U-Net structures for fingerprint restoration. It was proven in section 4 that the improved restoration of obliteration was not due to obliteration having a lower initial MSE values and that form of the damage must play a part. Obliteration is quite noisy in nature when compared to the Z cut and centrally rotated regions, which still had prominent ridges present. This points to U-Net structures being best for denoising specific tasks and not overlaying structured features. This agrees with much of the literature in the field regarding U-Nets. Research into U-Nets for restoration usually revolves around denoising as discussed in section 2. A U-Net-based architecture that could restore all types of damage could involve two U-Nets. The first U-Net could classify (segment) the damaged area and the damaged region could then be noised and then restored through the second U-Net. Another similar approach could be to have the damaged region blanked out and then inpainted using a network like the Shift-Net (Yan et al., 2018). However, these methods would be dependent on the ability of the first U-Net to correctly identify the damaged region of the z cut and centrally rotated regions. Furthermore, it would require more training material as the segmentation masks of the damaged regions of the SOCOFing dataset would need to be produced. If noising and denoising is the optimal way of restoring damage this could point to use of a stable diffusion-based architecture. This would also use the encoder decoder structure, but it is possible that the use of noising and denoising could allow the model to successfully restore the damaged region of the fingerprints. This method warrants further investigation.

From visual inspection it is possible that image restoration of this researched method may not be the best approach for restoring central rotation or Z cuts. In centrally rotated fingerprints there is minimal loss of information of the fingerprint, the difficulty in identification comes because of the change in orientation. Fingerprint restoration, whilst powerful at restoring lost information, requires extrapolation of the fingerprint ridges surrounding the damaged area. This need not be required for centrally rotated fingerprints if a different learning mechanism was used that could identify the damaged area and then rotate it until it best fits the surrounding fingerprint. This would again require segmentation masks for the segmentation training, but the resulting output of the model would be much more accurate if correctly implemented. A similar approach may be possible for the Z cut although the approach is not as obvious due to the irregular nature of the cuts and healing between different fingerprints.

If future models, of any nature and not just limited to U-Net-based architectures, are able to effectively restore the different types of damaged fingerprints research needs to be done as to how much this restoration helps when it comes to identifying damaged fingerprints. It could be that identification networks, such as Siamese networks, can accurately identify damaged fingerprints without the need for restoration thus rendering any restorative networks useless. However, it is doubtful that this would be the case as any identification network would have some error and if the restoration process is able to reduce that amount in any way it becomes useful, especially given the importance of fingerprint identification in modern society.

6. Conclusion

This report serves as preliminary research into the use of U-Net structures for fingerprint restoration. From the testing of three different U-Net-based architectures, the model showed strong performance in the restoration of obliterated fingerprints, but poor performance in the restoration of Z cut and centrally rotated fingerprints. Due to the noisy characteristics of obliterated damage, it warrants the use of stable diffusion to be integrated into the encoder decoder structure to improve the restorative performance for all damage types. Potentially a U-Net-based structure may not be the best approach for the restoration of centrally rotated and Z cut damaged fingerprints. Learning methods that can identify the damaged region and reshape it to fit into the surrounding fingerprint could reduce the loss of information of the fingerprint, hence making the model more accurate. Furthermore, research needs to be done to determine the extent to which restoration improves the performance of identification models and if the improved performance is worth the computational cost.

References

Fattahi, J. and Mejri, M. 2020. *Damaged Fingerprint Recognition by Convolutional Long Short-Term Memory Networks for Forensic Purposes.* [Online]. [Accessed 9th July 2023]. Available from: https://arxiv.org/pdf/2012.15041.pdf

Jain, A.K. and Kumar, A., 2012. *Biometric recognition: an overview.* [Online]. [Accessed 1st August 2023]. Available from: https://link.springer.com/chapter/10.1007/978-94-007-3892-8_3

Milesial. 2021. *Pytorch-UNet/unet/unet_parts.py* (2021). [Software]. [Accessed 11th July 2023]. Available from: https://github.com/milesial/Pytorch-UNet/tree/master/unet.

Milesial. 2023. *Pytorch-UNet/unet/unet_parts.py* (2023). [Software]. [Accessed 11th July 2023]. Available from: https://github.com/milesial/Pytorch-UNet/tree/master/unet.

Ronneberger, O., Fischer, P. and Brox, T. 2015. *U-Net: Convolutional Networks for Biomedical Image Segmentation.* [Online]. [Accessed 3rd July 2023]. Available from: https://arxiv.org/abs/1505.04597

Shehu, YI., Ruiz-Garcia, A., Palade, V. and James, A. 2018. Sokoto Coventry Fingerprint Dataset. *arXiv.* [Dataset]. [Accessed 9th July 2023]. Available from: https://arxiv.org/abs/1807.10609

Yan, L., Zhao, M., Liu, S., Shi, S. and Chen, J. 2023. *Cascaded transformer U-net for image restoration.* [Online]. [Accessed 26th July 2023]. Available from: https://www.sciencedirect.com/science/article/pii/S0165168422004418?casa_token=U3pwXpt-ji8AAAAA:TwixSkl09UMYzzAK92njnM9hsSro_26NBEOfOUSunmtfwcJlXumPNfuqeN-KIogaA8xVfp3JbyE

Yan, Z., Li, X., Mu, L., Zuo, W. and Shan, S. 2018. *Shift-Net: Image Inpainting via Deep Feature Rearrangement.* [Online]. [Accessed 4th July 2023]. Available from: https://openaccess.thecvf.com/content_ECCV_2018/papers/Zhaoyi_Yan_Shift-Net_Image_Inpainting_ECCV_2018_paper.pdf

Zhao, H., Gallo, O., Frosio, I. and Kautz, J. 2016. *Loss functions for image restoration with neural networks.* [Online]. [Accessed 26th July 2023]. Available from: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7797130