Reconstruction of Images Transmitted over LoRa Networks Using Fully Convolutional Neural Network

Jakub Nowak, Tomasz Nowak
Czestochowa University of Technology
Faculty of Computer Science and Artificial Intelligence
Czestochowa, Poland *{jakub.nowak, tomasz.nowak}@pcz.pl*

Szymon Połcik, Marcin Korytkowski
Czestochowa University of Technology
Faculty of Computer Science and Artificial Intelligence
Czestochowa, Poland *marcin.korytkowski@pcz.pl*

Pawel Drozda
University of Warmia and Mazury
Olsztyn, Poland *pdrozda@matman.uwm.edu.pl*

Rafał Scherer
Czestochowa University of Technology
Faculty of Computer Science and Artificial Intelligence
Czestochowa, Poland
AGH University of Krakow
Faculty of Computer Science
and Center of Excellence in Artificial Intelligence
Czestochowa, Poland *rafal.scherer@pcz.pl*

Abstract

LoRaWAN networks, which are extremely popular today, are based on the LoRa protocol and offer very long communication ranges, but they also come with significant limitations. These limitations stem primarily from two factors: duty cycle and maximum message size. During image transmission over LoRa-based networks, packet loss is a common problem resulting from limited bandwidth and transmission interference. During image transmission, it leads to missing data in the received content, most often visible as vertical or horizontal lines. We present a method to repair such corrupted images using a fully convolutional neural network inspired by the U-Net architecture. The experiments carried out show that the proposed approach effectively reconstructs missing parts of the image, achieving high structural similarity (SSIM) to the original. The proposed method can be applied to image transmission and reconstruction on low-power devices, typical of IoT systems that use LoRa for communication.

Keywords: LoRa, U-Net, Image Reconstruction, Image Transmission, Low-Power Networks.

1. Introduction

The LoRa (Long Range) technology [2] was designed for long-range communication of small data packets with minimal energy consumption. Due to its features, such as low power requirements, long battery life, and wide coverage, it is primarily used in Internet of Things (IoT) systems [10], including environmental sensors [7], utility meters [5], and localization systems [6]. However, using LoRa for image transmission poses a non-standard and technologically challenging use case. The standard is designed for small data frames, and transmitting images, even highly compressed and in grayscale, significantly exceeds the typical usage range of this technology. In practice, this leads to the fragmentation of data into smaller packets that are prone to transmission errors, interference, and loss, resulting in significant distortions in the received images. Typical artifacts caused by packet loss include horizontal or vertical lines, missing fragments, and zeroed-out pixel regions. To enable effective use of LoRa for image transmission, it becomes necessary to develop additional reconstruction mechanisms to recover lost data on the receiver side. The technological advancement in IoT systems has revolutionized the way we communicate, process data and make decisions. One of the key elements of this transformation is communication technologies such as LoRa, which enable long-range and energy-efficient data transmission in distributed IoT environments. LoRaWAN, as one of the most popular IoT standards, is dynamically evolving in applications such as smart cities, precision agriculture, and Industry 4.0 [8], where it is used for monitoring, management, and data analysis from machines and production systems. An example of a study addressing image transmission using LoRa technology is presented in [4]. The authors proposed a protocol based on sending multiple packets in batches and retransmitting only those that were not acknowledged by the receiver. However, this method does not implement any advanced error correction mechanisms (e.g., FEC) or techniques for reconstructing lost data, which means the image must be transmitted repeatedly until all packets are fully acknowledged. In contrast, the solution proposed in this paper eliminates the need to retransmit identical data. Instead, it introduces a receiver-side mechanism for reconstructing missing parts of the image using deep learning. This approach not only reduces the transmission time and network load, but also enables effective recovery of visual content under conditions with high error rates and packet loss. An alternative method is described in [13], where the authors focused on reducing the data volume transmitted over LoRa by applying WebP image compression and Base64 encoding. This reduced the number of packets needed to transmit a single image from 81 to 23, effectively shortening the transmission time by nearly 50%. However, the method does not include any mechanisms for recovering lost packets: there is no retransmission or error correction. As a result, any missing data results in an incomplete image, with no reconstruction performed on the receiver side. Packet loss in LoRa networks remains a significant challenge, especially in applications requiring precise data transmission, such as images. Existing methods for reconstructing corrupted data have mostly relied on simple linear interpolation, packet retransmission, or error-correcting codes. These methods often prove insufficient due to the non-linear nature of distortions and lack of contextual information. Therefore, it is necessary to develop new data reconstruction mechanisms on the receiver side that can effectively handle complex transmission errors in LoRa channels. In this context, modern image reconstruction methods, especially those based on the U-Net neural network architecture, offer new possibilities to improve data integrity. Such approaches enable the restoration of high-quality images even in cases of partial data loss. The Structural Similarity Index (SSIM) plays a key role in evaluating reconstruction quality.

Our approach to optimizing image transmission in IoT systems, which incorporates LoRa networks, image reconstruction mechanisms, and metrics such as SSIM, aims not only to minimize the impact of packet loss but also to improve the efficiency and quality of transmission in complex network environments. This opens new application possibilities for LoRaWAN in visual monitoring systems in hard-to-reach areas, where traditional video transmission technologies are impractical due to energy or cost constraints.

The remainder of this paper is organized as follows. In Section 2, we present challenges related to image transmission over LoRa networks, including packetization strategies and typical types of distortion. Section 3 describes the architecture of the proposed U-Net-based reconstruction model. Section 4 outlines the training procedure, including details of the data set and optimization parameters. The experimental results are presented and discussed in Section 5. Finally, conclusions and future research directions are provided in Section 6.

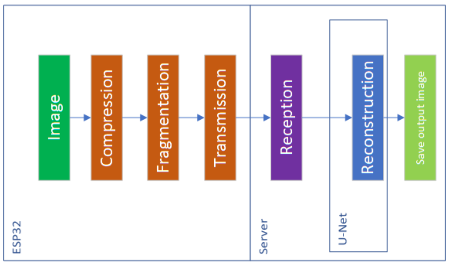

Fig. 1. Diagram of the overall image transmission and reconstruction process: from the source image, through compression and transmission, to reconstruction using the proposed neural model.

2. Problem description and data

In many water supply installations, traditional analog water meters are still in use and lack an output interface for pulse reading [1]. Replacing them with modern meters capable of generating digital signals involves very high costs, both in terms of hardware and the logistics of deployment. An alternative solution is to use cameras that capture images of the meter dials, which are then analyzed to extract the numerical readings. Applying LoRa technology to transmit such images makes it possible to build a remote reading system without the need to interfere with the existing water meter infrastructure. The data set presented in this work reflects such a scenario: images were generated based on data obtained from real-world LoRa-based water meter image transmission systems.

In its typical application, LoRa technology is used to transmit small amounts of data, such as sensor readings or location information. However, when it comes to transmitting images, even highly compressed and grayscale significant challenges emerge due to LoRa's limited bandwidth and susceptibility to transmission errors. Under standard conditions, the maximum size of a single data packet is restricted, which forces an image to be divided into many small fragments. If some of these packets are lost, the receiver is unable to reconstruct the full content, resulting in visible image distortions. In this work, we propose two approaches for transmitting images over a LoRa network: (1) Classical line-wise splitting – each packet corresponds to one full row of the image. If such a packet is lost, the row is replaced with pixels of zero value (black stripes). (2) Interleaved packet splitting – each packet contains only a portion of a line, with gaps every few pixels (e.g., only every other fragment is sent). If an error or packet loss occurs, the result is broken, discontinuous lines. All data used in the experiment were generated on the basis of real images of water meters transmitted through a LoRa-based communication system. The images were resized to 128 × 128 pixels in grayscale. From these, a training and test dataset was prepared, in which synthetic distortions were added to the original images to simulate typical transmission errors, following the two described variants: full-line and interleaved-line corruption. To better illustrate the impact of transmission errors, Figure 2 shows an example image in three variants: original (undistorted), corrupted with full black lines (simulating traditional packet loss), and corrupted with interleaved lines (simulating fragmented packet transmission). For training and evaluation purposes of the proposed model, a dataset of 100,000 synthetically corrupted 128 × 128 pixel images was prepared. The dataset was evenly divided into 50,000 samples that contained full-line distortions (variant 1) and 50,000 samples contained interleaved-line distortions (variant 2). Each sample consisted of a pair: an input (corrupted) image and a reference (original) image. The data were split 80:20 into training and validation sets while maintaining balance between the distortion types.

A custom transmission module based on the ESP32 microcontroller was developed, integrating an OV2640 camera and an RFM95W LoRa transceiver. The ESP32, equipped with Wi-Fi, Bluetooth, 520kB SRAM and 4MB Flash memory, was responsible for control and im-

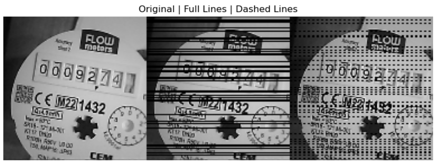

Fig. 2. Original water meter image (left), and its two corrupted versions: with full black lines (middle) and interleaved lines (right), corresponding to different types of LoRa transmission errors.

age processing [12]. The OV2640 sensor (up to 2MP) captured images of water meters, which were downscaled to 128 × 128 grayscale resolution. The image fragments were transmitted via the RFM95W module operating in the 868MHz band, which offers long-range communication with low energy consumption. The system operated in cycles: the camera captured an image, which was then processed and split into packets transmitted via LoRa to the receiver. Reception was handled by a MikroTik LoRa gateway configured as a data concentrator. The received packets were sent to a custom server-side service designed for the project. Its responsibilities included reassembling received packets, identifying missing segments, and storing and archiving the data. Each packet contained the following information: transmitter device identifier, image offset indicator, transmission metadata (block size, data format, encoding method), and binary payload – image fragment. This enabled asynchronous and distributed reconstruction of the image in server memory. Once all packets were received, the complete image transmitted from the remote device was reconstructed.

3. U-Net-inspired architecture for image reconstruction

To solve the task of reconstructing images transmitted in a non-continuous and potentially corrupted form, we employed a convolutional neural network based on the U-Net architecture. Originally designed for medical image segmentation, the U-Net structure has also proven effective in reconstruction tasks, thanks to its symmetric design and the use of so-called skip connections [14]. In our case, the U-Net model was used to restore the original grayscale image 128 × 128 from a distorted version, which contained random artifact lines or missing pixels resulting from packet loss. The structure of the network is illustrated in Figure 3. The input im-

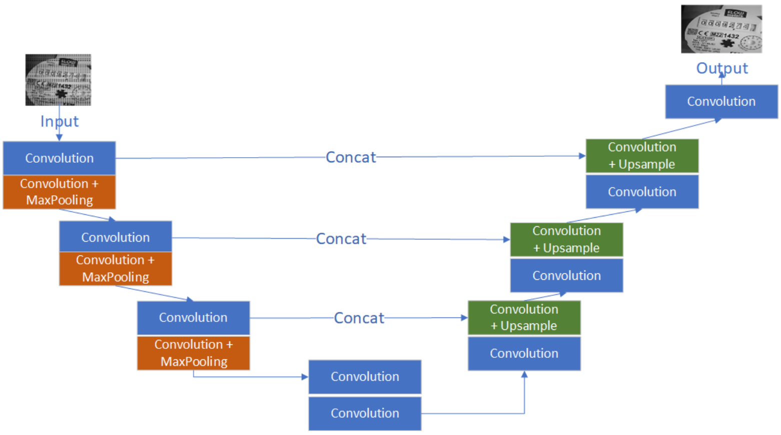

Fig. 3. Architecture diagram of the proposed model for grayscale image reconstruction. Left: encoder part, center: bottleneck, right: decoder part with skip connections.

age is processed through the encoder (blue-orange blocks), which performs feature extraction while reducing spatial resolution. This is followed by the bottleneck layer in the middle, and then the decoder (green-blue blocks) restores the full resolution of the image. Skip connections (*Concat*) link layers of the same resolution in the encoder and decoder, allowing the network to preserve local structural details. The model consists of three main components: the encoder, decoder, and skip connections. The encoder is responsible for extracting features from the input image. It includes convolutional layers with ReLU activation and max-pooling layers that reduce the resolution from 128×128 to 16×16 pixels, while increasing the number of channels to 512. The decoder reconstructs the full image resolution. It uses bilinear upsampling [9] followed by convolutional layers. This approach, aligned with the original U-Net architecture by Ronneberger et al. [11], ensures simplicity and model stability. Skip connections recover local details lost during downsampling and enhance reconstruction fidelity. The output of the model is a convolutional layer with a 1×1 filter and a sigmoid activation function, which maps pixel values to the [0, 1] range. Before saving the output image, the pixel values are multiplied by 255 to restore the full grayscale intensity. To evaluate the quality of image reconstruction, we use the SSIM metric [3], which compares two images in terms of their structure, luminance and contrast. Unlike traditional pixel-based metrics, such as MSE or PSNR, SSIM is designed to better reflect human visual perception. SSIM between images x and y is computed using

Table 1. Proposed neural model architecture for image reconstruction (input: 1 × 128 × 128)

Stage	Layer type	Parameters	Output size
Input	-	Grayscale image	1 × 128 × 128
Encoder	Conv2D + ReLU	64 filters, 3 × 3, padding=1	64 × 128 × 128
Encoder	Conv2D + ReLU	64 filters, 3 × 3, padding=1	64 × 128 × 128
Encoder	MaxPool2D	2 × 2	64 × 64 × 64
Encoder	Conv2D + ReLU	128 filters, 3 × 3, padding=1	128 × 64 × 64
Encoder	Conv2D + ReLU	128 filters, 3 × 3, padding=1	128 × 64 × 64
Encoder	MaxPool2D	2 × 2	128 × 32 × 32
Encoder	Conv2D + ReLU	256 filters, 3 × 3, padding=1	256 × 32 × 32
Encoder	Conv2D + ReLU	256 filters, 3 × 3, padding=1	256 × 32 × 32
Encoder	MaxPool2D	2 × 2	256 × 16 × 16
Bottleneck	Conv2D + ReLU	512 filters, 3 × 3, padding=1	512 × 16 × 16
Bottleneck	Conv2D + ReLU	512 filters, 3 × 3, padding=1	512 × 16 × 16
Decoder	Upsample (bilinear)	scale=2	512 × 32 × 32
Decoder	Concat (skip)	from 256 × 32 × 32	768 × 32 × 32
Decoder	Conv2D + ReLU	256 filters, 3 × 3	256 × 32 × 32
Decoder	Conv2D + ReLU	256 filters, 3 × 3	256 × 32 × 32
Decoder	Upsample (bilinear)	scale=2	256 × 64 × 64
Decoder	Concat (skip)	from 128 × 64 × 64	384 × 64 × 64
Decoder	Conv2D + ReLU	128 filters, 3 × 3	128 × 64 × 64
Decoder	Conv2D + ReLU	128 filters, 3 × 3	128 × 64 × 64
Decoder	Upsample (bilinear)	scale=2	128 × 128 × 128
Decoder	Concat (skip)	from 64 × 128 × 128	192 × 128 × 128
Decoder	Conv2D + ReLU	64 filters, 3 × 3	64 × 128 × 128
Decoder	Conv2D + ReLU	64 filters, 3 × 3	64 × 128 × 128
Output	Conv2D + Sigmoid	1 filter, 1 × 1	1 × 128 × 128

$$\mathrm{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{1}$$

where μ_x, μ_y is the mean luminance values of images x and y, σ_x^2, σ_y^2 is the luminance variance of the images, σ_{xy} is covariance between x and y, and C_1, C_2 are stabilization constants to avoid division by zero. The constants are defined as $C_1 = (K_1 L)^2$, $C_2 = (K_2 L)^2$, where L is the maximum pixel value (255 for 8-bit images), with $K_1 = 0.01$ and $K_2 = 0.03$. SSIM is calculated locally using small windows (e.g. 11×11), and the final score is obtained by averaging over the entire image, ensuring sensitivity to local structural distortions. The resulting value ranges from 0 to 1, where 1 indicates perfect structural similarity.

4. Network training

The model was implemented using the PyTorch library and trained on synthetically distorted data, including transformations such as image rotation, skewing, and shadowing. The optimization of the model parameters was performed using the Adam algorithm. with default momentum coefficients $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The learning rate was set to a constant value of 0.001. The training process was conducted over 50 epochs using a mini-batch size of 32. To assess the quality of image structure reconstruction, we used the SSIM as validation metric. SSIM takes into account contrast, luminance, and local structural consistency. For each epoch, we computed the loss value for the training set and the SSIM value for the validation set, enabling effective monitoring of training progress and detection of overfitting. MSE measures the average squared difference between the original and reconstructed images. Lower MSE values indicate a closer match between the two. The advantages of MSE include its simplicity, differentiability, and numerical stability, making it well suited for gradient-based optimization in neural networks. For grayscale images, pixel values are compared directly in the range of 0 to 1. The resulting loss value also lies between 0 and 1, where values closer to 0 indicate better reconstruction fidelity. Although MSE does not account for structural information, it remains a practical tool for quick and effective assessment of pixel-level accuracy in image reconstruction tasks.

5. Experimental results

Our model was trained on a dataset consisting of 100,000 images of water meters with synthetically simulated transmission artifacts, evenly divided into samples with full-line and fragmented-line distortions. The training process was repeated 20 times and was based on learning with training data in the form of 15 epochs using the MSE loss function and the structural SSIM metric to assess the quality of reconstruction. The suggested number of iterations did not result in further improvement in the quality of image reconstruction, measured by the SSIM coefficient. In contrast, it led to overfitting, which resulted in worsening results in the validation set. To determine the optimal depth of the model, experiments were conducted with different numbers of feature channels in the convolutional layers of the U-Net network. For each configuration, the model was trained using the same dataset and number of epochs. The results are presented in Table 2. The best reconstruction quality (highest SSIM and lowest MSE) was achieved for the scaling factor ×1.0, which corresponds to 64 feature channels in the first layer. Smaller models (×0.25, ×0.5) showed signs of underfitting (low SSIM, high MSE), while larger configurations (×1.5, ×2.0) led to overfitting, as indicated by the degradation of validation metrics. Therefore, the ×1.0 configuration was selected as a compromise between performance, accuracy, and

Table 2. Impact of U-Net width (scaling factor of feature map count) on reconstruction quality

Feature maps scaling factor	No. of features in layer 1	SSIM (%)	MSE error (%)
×0.25	16	90.1%	28.5%
×0.5	32	93.5%	24.7%
×1.0	64	**95.9%**	**14.2%**
×1.5	96	95.3%	16.1%
×2.0	128	94.6%	17.4%

computational complexity. Figure 4 shows the evolution of the loss values for the training and validation data sets, together with the SSIM scores on the validation set during the training process. As depicted, there is a steep decline in both training and validation loss within the first few epochs, indicating rapid model adaptation to the dataset structure. In later stages, the loss values stabilize at a low level (below 0.003). Figure 5 presents an example of image reconstruction for a water meter. The input image exhibits severe distortion in the form of black lines. After processing with the U-Net model, the reconstructed image recovers much of the original visual structure with readable digits. However, some residual artifacts remain where distortions were previously present, particularly in the case of wide or continuous lines. These observations suggest that the model is more effective in reconstructing images affected by fragmented or narrow lines, while thick, uninterrupted lines, especially those that overlap multiple digits, may locally obscure the structure, making full recovery more difficult.

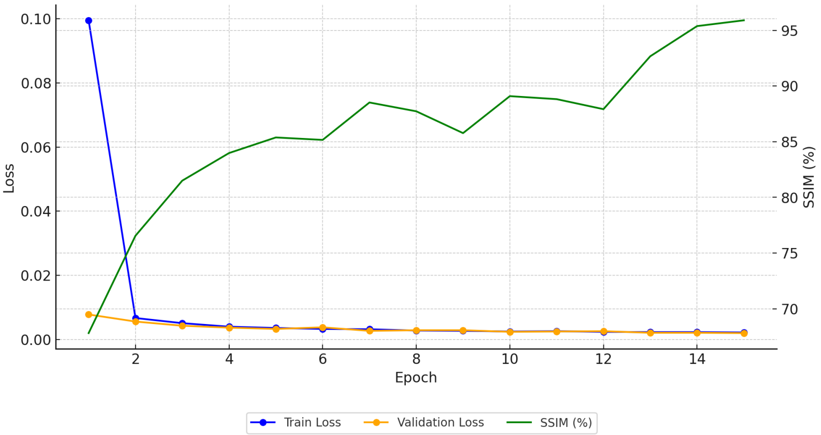

Fig. 4. Training performance of the model – loss values (Train Loss, Validation Loss) and SSIM metric (%) across epochs.

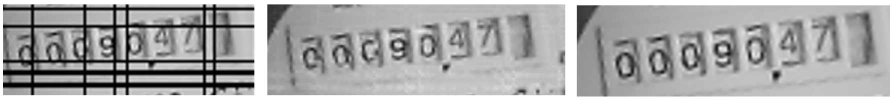

Fig. 5. Reconstruction example. From left: distorted input image, model output, reference image.

6. Conclusion and future work

This paper presents an approach for reconstructing images distorted during transmission using a convolutional neural network. Experimental results confirmed the model's effectiveness in removing artifacts such as lines and missing fragments while preserving structural features of the original image. High SSIM scores and low MSE values demonstrate the suitability of the method for low-bandwidth transmission scenarios, such as LoRa-based systems. To evaluate reconstruction quality, SSIM and MSE metrics were used, as shown in Table 2. The best results were achieved with a scaling factor of ×1.0, indicating a balance between network complexity and reconstruction quality. Smaller models led to underfitting, while larger ones caused overfitting. An SSIM of 95.9% and MSE of 14.2% confirm high reconstruction accuracy despite packet loss. However, the method remains vulnerable in cases where transmission artifacts affect critical parts of digits, particularly when key features are occluded, such as loops in 8 and 9 or vertical strokes in 1 and 7. Even minor damage to these regions can hinder the correct interpretation of the image.

References

[1] Bhoyar, D., Katey, B., Ingle, M.: Lora technology based low cost water meter reading system. In: Proceedings of 3rd International Conference on Internet of Things and Connected Technologies (ICIoTCT). pp. 26–27 (2018)

[2] Bor, M.C., Vidler, J., Roedig, U.: Lora for the internet of things. In: Ewsn. vol. 16, pp. 361–366 (2016)

[3] Brunet, D., Vrscay, E.R., Wang, Z.: On the mathematical properties of the structural similarity index. IEEE Transactions on Image Processing 21(4), pp. 1488–1499 (2011)

[4] Chen, T., Eager, D., Makaroff, D.: Efficient image transmission using lora technology in agricultural monitoring iot systems. In: 2019 International Conference on Internet of Things. pp. 937–944. IEEE (2019)

[5] Cheng, Y., Saputra, H., Goh, L.M., Wu, Y.: Secure smart metering based on lora technology. In: 2018 IEEE 4th International Conference on Identity, Security, and Behavior Analysis (ISBA). pp. 1–8. IEEE (2018)

[6] Delafontaine, V., Schiano, F., Cocco, G., Rusu, A., Floreano, D.: Drone-aided localization in lora iot networks. In: 2020 IEEE International Conference on Robotics and Automation (ICRA). pp. 286–292. IEEE (2020)

[7] González, E., Casanova-Chafer, J., Romero, A., Vilanova, X., Mitrovics, J., Llobet, E.: Lora sensor network development for air quality monitoring or detecting gas leakage events. Sensors 20(21), pp. 6225 (2020)

[8] Götz, M.: Przemysł czwartej generacji (przemysł 4.0) a międzynarodowa współpraca gospodarcza. Ekonomista 4, pp. 385–403 (2018)

[9] LeCun, Y., Kavukcuoglu, K., Farabet, C.: Convolutional networks and applications in vision. In: Proceedings of 2010 IEEE international symposium on circuits and systems. pp. 253–256. IEEE (2010)

[10] Madakam, S., Ramaswamy, R., Tripathi, S.: Internet of things (iot): A literature review. Journal of computer and communications 3(5), pp. 164–173 (2015)

[11] Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. arxiv 2015. arXiv preprint arXiv:1505.04597 (2015)

[12] Salikhov, R., Abdrakhmanov, V.K., Safargalin, I.: Internet of things (iot) security alarms on esp32-cam. In: Journal of Physics: Conference Series. vol. 2096, p. 012109. IOP Publishing (2021)

[13] Wei, C.C., Su, P.Y., Chang, C.C., Chang, K.C.: A study on lora dynamic image transmission. In: 2021 IEEE 4th International Conference on Knowledge Innovation and Invention (ICKII). pp. 10–13. IEEE (2021)

[14] Yu, Z., Yu, L., Zheng, W., Wang, S.: Eiu-net: Enhanced feature extraction and improved skip connections in u-net for skin lesion segmentation. Computers in Biology and Medicine 162, pp. 107081 (2023)