

# Improving Convective Cloud Classification with Deep Learning: The CC-Unet Model

Humuntal Rumapea<sup>a,b,\*</sup>, Mohammad Zarlis<sup>c</sup>, Syahril Efendy<sup>a</sup>, Poltak Sihombing<sup>a</sup>

<sup>a</sup> Department of Computer Science, Universitas Sumatera Utara, Medan, 20155, Indonesia

<sup>b</sup> Department of Computer Science, Universitas Methodist Indonesia, Medan, 20155, Indonesia

<sup>c</sup> Department of System Information Management, Bina Nusantara University, Jakarta, 11480, Indonesia

Corresponding author: \*hrumapea1608@students.usu.ac.id

**Abstract**— Analyzing and mitigating natural disasters can be a challenging task, which is why the field of computer science, specifically artificial intelligence (AI) is necessary to aid in the complexity of disaster management. AI provides the tools and analytical models to help solve the intricacies of handling natural disasters. Convective clouds, closely related to rain and can lead to large-scale, prolonged hydrometeorological disasters, are a crucial component to consider. To improve the classification of these clouds, a predictive-analytical model based on deep learning, called the CC-Unet model, was developed. This model utilizes a U-Net architecture and is trained using a dataset of convective cloud images. The researchers used satellite image data from the Himawari 8 satellite collected in May and October 2021. The images were pre-processed and verified using observational data. The model was tested using a random train-test split method, showing that the CC-Unet model had a higher accuracy of 97.29% compared to the U-Net model, which had an accuracy of 94.17%. Additionally, the significance test using the Wilcoxon method showed that the CC-Unet model had significantly different performance results from the U-Net model. The ground truth image was also compared with the predicted image, showing a low root mean square error value of 0.0218, indicating a high level of similarity between the two. Overall, this research demonstrates the potential of AI and deep learning in classifying convective clouds to aid in natural disaster management.

**Keywords**— Natural disaster mitigation; artificial intelligence; convective cloud classification; deep learning.

Manuscript received 6 Feb. 2023; revised 14 Aug. 2023; accepted 9 Jan. 2024. Date of publication 29 Feb. 2024.  
IJASEIT is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.



## I. INTRODUCTION

Flooding is an urban problem that should be solved because it has a systematic adverse impact on human life [1]. This disaster is an event where the water level rises because the absorption capacity of the soil is not proportional to the volume. In low-lying urban areas, flooding can be caused by natural factors, such as rainfall and river overflow, negatively affecting human behavior toward the environment [2], [3]

Rain is one of the most significant key factors in the water balance on earth, which leads to flooding [4]. Medan is the economic center of the government in North Sumatra, which is constantly improving in creating a modern and livable city plan. Unfortunately, this city has been inseparable from the flooding problem due to climate change, often making rain prediction difficult [5]. Therefore, technology, such as Artificial intelligence, is expected to help decipher the complexity of predicting rain events in North Sumatra.

Its benefits can be felt in various sectors with an increase in numerous. This has led to a shift in learning from the previously used machinery field to deeper learning. Over the last few decades, deep learning has tackled several prediction problems [6]–[8]. One application of DL is detecting and mitigating the complexity of natural disasters today, especially in Indonesia. Therefore, a more precise learning method is needed to provide solutions to disaster detection problem-solving [9]–[12].

Deep learning is an essential element of data science, which includes statistics and predictive modeling [13]. It makes it easier for data scientists to collect, analyze, and interpret large amounts of data. Moreover, this technique can automate analytical and predictive processes [7]. In contrast to traditional machine learning, which uses linear hierarchy, deep learning algorithms are designed non-linearly, involving multiple layers [14]. Therefore, the convolution neural network (CNN) was used in this research to detect convective

clouds. The cloud image data was obtained through the Himawari 8 Satellite sensing [15]–[20].

There are two main approaches to CNN-based image segmentation, namely region-based semantic segmentation (RBSS) as in [21] and fully convolution network-based semantic segmentation (FCNBSS) as studied in [22]. In the first approach, image segmentation is performed by breaking the image into smaller parts. This method requires a complicated process, especially on complex images [23]. Meanwhile, in the second approach, segmentation is based on image pixels. It is simpler because it does not need to perform image feature extraction. Subsequent analyses in image segmentation are more directed to the application and development of FCNBSS-based models [24].

There are two main models for the FCNBSS approach: SegNet [25] and U-Net [26]. Both have been successfully applied to cloud image segmentation. Besides these two models, others related to cloud image segmentation are the development of the CD-Unet model [27], model Refined U-Net [28], [29], and SegCloud [30].

According to preliminary studies, the U-Net-based model is more effective than SegNet [31]. This is because U-net can distinguish the scale of image features even with small data. On the other hand, in cases with complex images, it often fails to produce satisfactory performance [32]. This is caused by background image factors capable of affecting the segmentation results [33]. Image background factors often arise due to interference from nature, buildings, and others. Then, the presence of these interfering objects affects the result of cloud image segmentation [34]. Due to these considerations, improving the U-Net model to capture the main image segment as much as possible is necessary. Improvements to the U-Net model are expected to enhance its ability to minimize residual or noise factors [35] in image data, especially those associated with the cloud [29], [37].

Several studies on U-Net have been conducted, such as the one performed by [38] using a deep-learning method named U-Net, which was applied to improve the skill in forecasting summer (June–August) precipitation for a one-month lead. The most improved areas were Northwest, Southwest, and Southeast China. Sensitivity experiments showed that soil moisture is the most crucial factor in predicting summer rainfall in China. Wang et al. [39] advances in deep learning enable complex spatial patterns such as urban development to be learned and simulated, wang used the U-Net deep learning algorithm to capture historical urban development and simulate future patterns for the North China Plain, the results showed that it can accurately predict urban land-use and mimic real-world spatial patterns. Marhamati et al. [40] propose a novel approach for segmenting depressed human tongues in photographic images called the learning-to-augment incorporated U-Net (LAIU-Net). The approach addresses the challenges of shape variability and flexibility of the tongue, and it uses a dataset of 333 tongue images from 111 depressed individuals. The LAIU-Net approach also addresses the problem of overfitting and increases the generalizability of the deep network by automatically choosing the best policies for data augmentation, it has been compared to other state-of-the-art U-Net configurations, and it achieved a mean boundary F1 score of 93.1%.

U-Net is a robust deep learning architecture effective in various image segmentation tasks [32], [41], [42]. It has been used to address the challenges of segmenting small vessels in retinal blood vessel segmentation [43], the frequent movement of the tongue due to its natural flexibility in the tongue diagnosis system, and the shape variability and flexibility of the tongue in depressed human tongue segmentation [40]. The U-Net architecture has also been modified and incorporated with techniques such as series deformable convolution [44], attention mechanism [45], and learning-to-augment to improve its performance and address overfitting issues [46], [47]. The results from these studies have shown that U-Net has strong capabilities in accurately segmenting images with high shape variability and flexibility, with mean boundary F1 scores reaching up to 93.1%.

This research used residual minimization by multiplying U-Net channels into two or dual channels. The first and second are applied to extract the primary and residual segments of the cloud image. This approach aligns with previous research on eliminating image residues, namely dual channel [48] and double branching [49]. Both studies have successfully applied the dual-channel principle in overcoming image residues. Therefore, based on the background described, this research built the CC-Unet model, a deep learning-based technique to detect convective clouds and improve the performance of the previous U-Net. The deep learning model aims to detect the type of CB cloud sourced from the satellite. The results of this prediction will be used as a flood early warning system.

## II. MATERIAL AND METHOD

### A. Datasets

Data collected from Himawari-8 satellite image from 2019 to 2021 were as follows:

- Himawari-8 satellite data IR-1 channel in pgm (Portable Gray Map) and .cn formats with a  $0.04^\circ \times 0.04^\circ$  per hour resolution. This data was downloaded from the GMS/GOES9/MTSAT Data Archive for Research and Education webpage at <http://weather.is.kochi-u.ac.jp/archive-e.html>.
- Hourly and daily rainfall observation data from weather observation stations and BMKG rainfall measurement posts in the Medan City area
- Upper air observation data (Radiosonde) from the Kualanamu Meteorological Station.

### B. Data Preparation Techniques

The data processing procedures carried out in this research are as follows:

- Meso-scale heavy rain timing, daily rainfall data from the BMKG Weather Observation Station in the North Sumatra region from 2019 to 2021 is used to view mesoscale heavy rain events in the Medan City area. The time of heavy rain that occurs evenly in two weather observation stations in the Medan area is selected as a case study.
- Pre-processing cloud image data in .cn format was converted into image data in .TIFF for easy processing by the CC-Unet model.

### C. Methods U-Net Based Deep Convolution

U-Net architecture introduced by Ronneberger [26] is one of the first convolutional networks specifically designed for biomedical image analysis. It is a network of derivative FCN that aims to address two domain-specific problems in medical image segmentation. This architecture aims to produce competitive segmentation results given the relatively small training data. Large data sets with a fully connected Feed-forward CNN layer are used to analyze many parameters. This model has the luxury of learning a bit of information through many examples. For instance, in the case of medical image

segmentation, the model needs to maximize what is learned from each sample.

Meanwhile, cloud image segmentation must maximize the information learned from each instance. Encoder-decoder architectures, such as UNet, have proven more effective even with small data sets. This is due to the ability of the up-convolution circuit to replace the fully connected layer on the decoder, which still has learnable parameters, which are far fewer than those of the fully connected layer. The second problem the U-Net architecture addresses is capturing context accurately and localizing the lesion at different scales and resolutions.

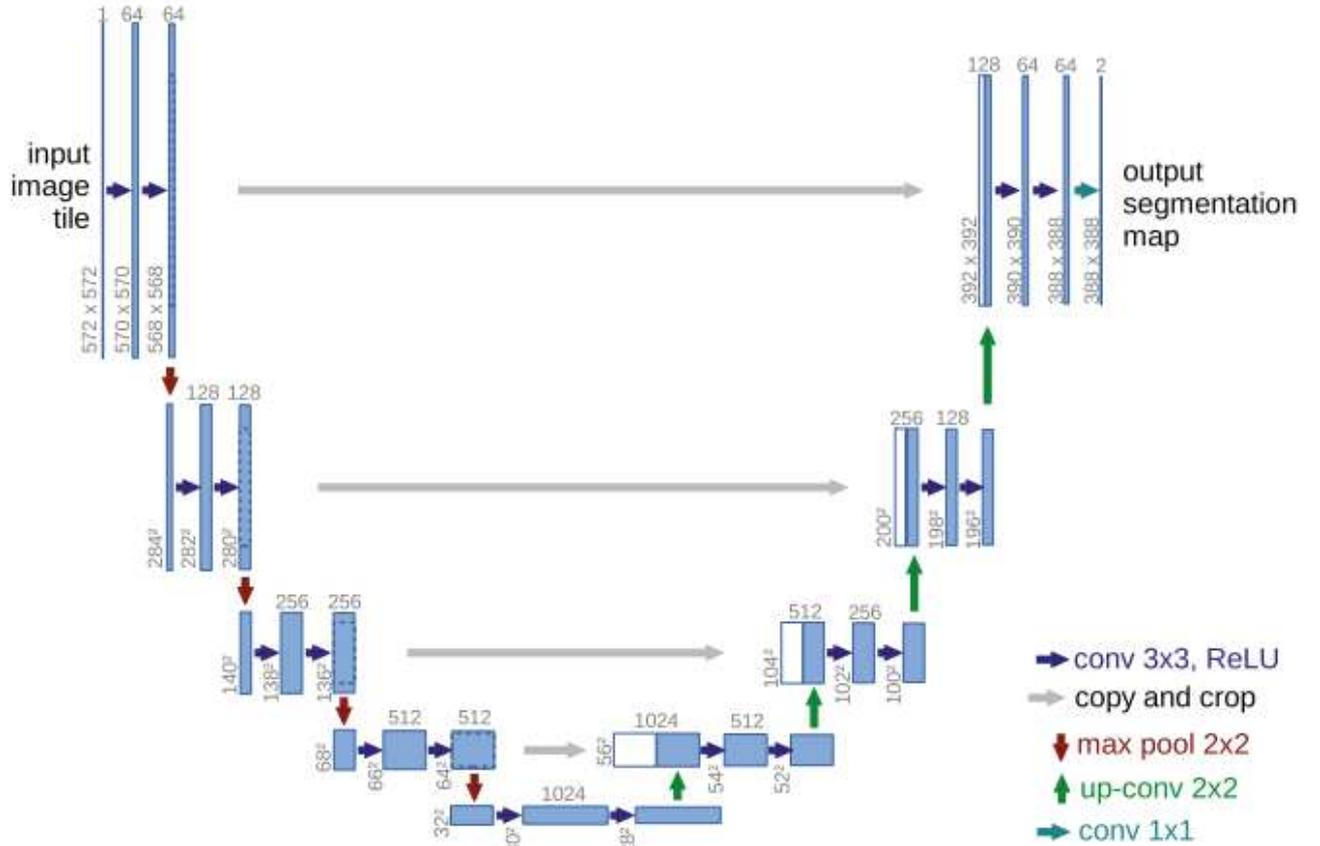


Fig. 1 U-Net Architecture

The structural representation of the U-Net architecture is illustrated in Figure 1. It is composed of two main components: the encoder and decoder. Each block in the encoder consists of two 3x3 convolutional layers, followed by a 2x2 pooling layer and a rectified linear unit (ReLU) activation function. Both the encoder and decoder comprise of convolution layers that form the contracting path. Each block in the decoder consists of two 3x3 convolutional layers and a 2x2 up-sampling layer. The encoder is responsible for

downsampling the image while increasing its resolution. In contrast, the decoder is an expanding path, consisting of upsampling operations. At the end, a 1x1 convolutional layer produces the segmentation output. In the U-Net architecture, the low-resolution features of the contracting path are combined with the upsampled output of the expanding path. In this research, a process for data processing is carried out using deep learning and general approaches. The details of the research procedure are presented in Fig. 2.

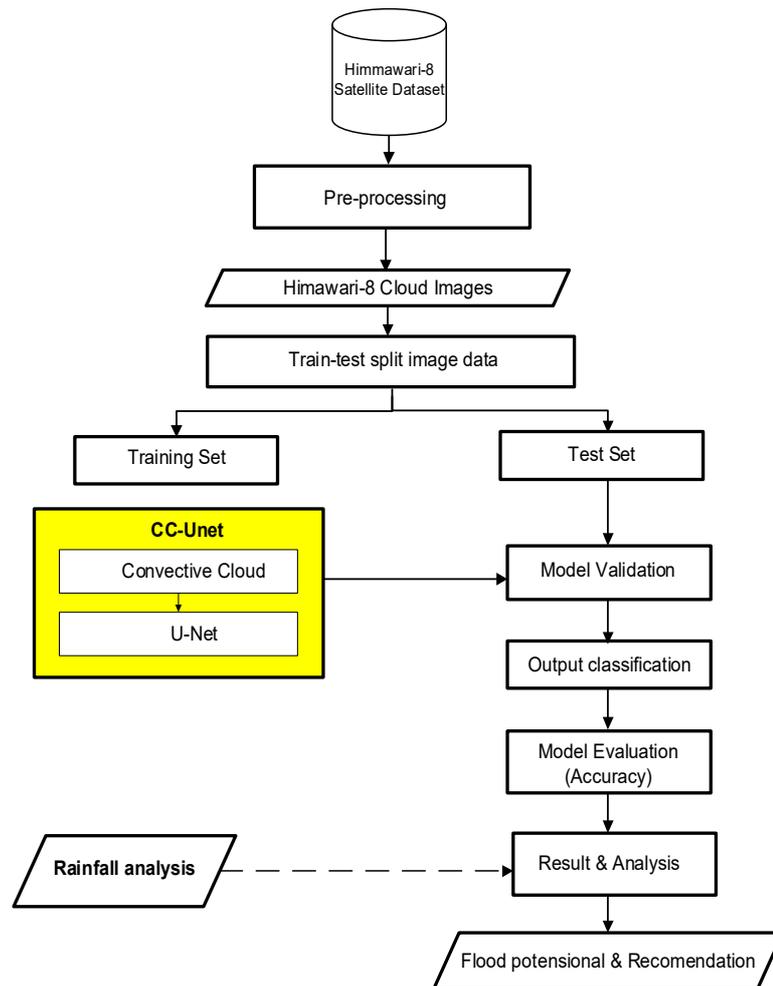


Fig. 2 Research Flowchart

### III. RESULT AND DISCUSSION

The CC-Unet architecture applied to the convective cloud image segmentation is illustrated in Figure 3. Each block in

the CC-Unet architecture is shown in Figure 4. Each block comprises of two parallel 3x3 convolutional layers, resulting in a total of four convolutional layers in each block.

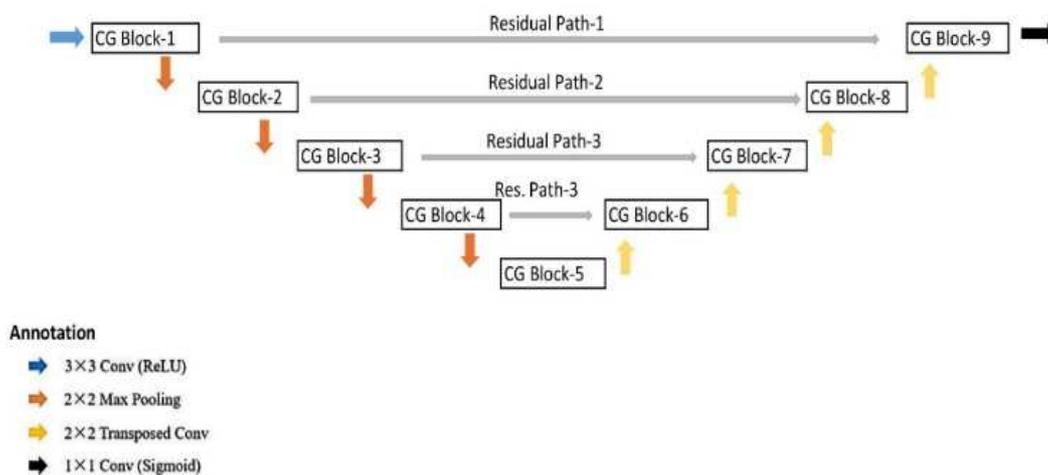


Fig. 3 CC-Unet Model

The CC-Unet model above can be described as a layered architecture, divided into encoder and decoder, as shown in Figure 4. The encoder section performs a downsampling function to reduce the image size. This section comprises of the parameter number of channels (n);  $n = \{64, 128, 256, 512\}$ , and the activation function (f) = ReLu. Meanwhile, the decoder section performs an up-sampling function to restore the image size. This section comprises the parameter number of channels (n);  $n = \{512, 256, 128, 64\}$ , and the activation function parameter (f) = ReLu. The last part is the output layer, a classified or segmented cloud image. Then, the output parameter is the number of output channels (N), = 1x1, with activation function (f) = softmax.

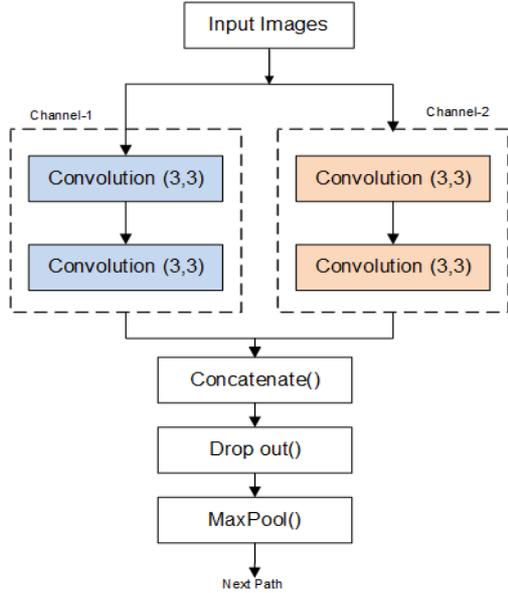


Fig. 4 CC-Unet Model

The CC-Unet model above can be described as a layer architecture, as shown in Table 1.

TABLE I  
CC-UNET DETAILS

Block	Channel-1	Channel-2	Path	Concrete Layer	Filter
CG Block-1	Conv(3,3)	Conv(3,3)		Conv(3,3)	64
CG Block-9	Conv(3,3)	Conv(3,3)		Conv(1,1)	64
CG Block-2	Conv(3,3)	Conv(3,3)	Residual Path-1	Conv(3,3)	64
CG Block-8	Conv(3,3)	Conv(3,3)		Conv(1,1)	64
CG Block-3	Conv(3,3)	Conv(3,3)	Residual Path-2	Conv(3,3)	128
CG Block-7	Conv(3,3)	Conv(3,3)		Conv(1,1)	128
CG Block-4	Conv(3,3)	Conv(3,3)	Residual Path-3	Conv(3,3)	256
CG Block-6	Conv(3,3)	Conv(3,3)		Conv(1,1)	256
CG Block-5	Conv(3,3)	Conv(3,3)	Residual Path-4	Conv(3,3)	512

Mathematically, the CC-Unet model in Figure 4 above can be modeled as an equation (1). It converts input convective cloud images into a vector or tensor-based prediction image

$$f(i) = \sigma(\sum_{j=1}^{H_2} u_j, i^\sigma (\sum_{k=1}^{H_1} u_k, j^\sigma (\sum_{m=1}^M x_m w_{m,k} + \beta_k) + \gamma_j) + \lambda_i) \quad (1)$$

where  $\beta_k$ ,  $\gamma_j$ , and  $\lambda_i$  are the image bias or noise factor, N is the number of layers, H is the hidden layer, M is the final or

output layer, and k is the number of image labels or classes. Determination of the pixel position of the output image on CC-Unet using probability theory.

The probability value is calculated using the SoftMax function to obtain the right value to produce good segmentation accuracy. The SoftMax equation is described in equation (2) as follows:

$$p_k(x) = \exp(a_k(x)) / (\sum_{k'=1}^K \exp(a_{k'}(x))) \quad (2)$$

Where  $p_k$  is the probability value of pixel x placement in the predicted image. The proposed model for the CC-U-Net architecture in the form of an algorithm is explained through the following stages:

**Algorithm 1 : CC-Unet-Classification of convective cloud**

**input :** image of cloud size 3600 x 2400

**for** i = 1 to 9 **do :**

{Step 1 Line of encoder(n)}

$c_i = \text{Convolution2D}(\text{number of channels}, (3, 3), f)$

$c_i = \text{Dropout}(0.2)(c_i) // 0.2 == \text{multiplier weight}$

$c_i = \text{Convolution2D}(16, (3, 3), f)$

$p_i = \text{MaxPooling2D}(2, 2)(c_i)$

{Step 2 : Cloud Image Decoder (n)}

$u_i = \text{Conv2DTranspose}(512, (2, 2)(c_i-1)$

$u_i = \text{concatenate}([u_i, c_i-2])$

$c_i = \text{Conv2D}(256, (3, 3), f)(u_i)$

$c_i = \text{Dropout}(0.2)(c_i)$

$c_i = \text{Conv2D}(128, (3, 3), f)(c_i)$

{Step 3 : Output

Softmaxx :

$p_k(x) = \exp(a_k(x)) / (\sum_{k'=1}^K \exp(a_{k'}(x)))$

}

**Output :**

$O_{\text{net}} = \text{Conv2D}(N, \text{dimention}, \text{activation})(c_i)$

The two types of activation functions applied in this research are ReLu and Softmax. In the CC-Unet architecture above, the Rectified Linear Unit (ReLu) activation function is applied to each convolution layer. The mathematical function of ReLu is described in equation (3), while Figure 5 is the graph of the ReLu function:

$$p_k(x) = \exp(a_k(x)) / (\sum_{k'=1}^K \exp(a_{k'}(x))) \quad (3)$$

where  $a_k(x)$  is an activation function in the k channel at the pixel x position, K parameter is the number of classes and functions,  $p_k(x)$  is the maximum approximation of the function.

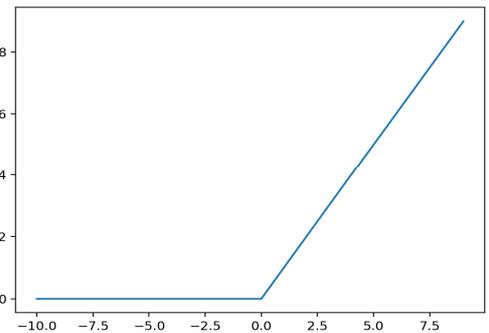


Fig. 5 ReLu function form

### A. Model Training

Cloud image segmentation aims to identify and classify the presence of cloud pixels. Additionally, the Binary Cross Entropy (BCE) value is the final determinant for evaluating the pixel position, as determined by equation (4):

$$L_{BCE} = -\frac{1}{N} \sum_{x \in \Omega} w(x) \log(p_{l(x)}(x)) \quad (4)$$

where  $N$  is the batch size, the value of  $\Omega$ , consisting of  $\{1, \dots, K\}$ , is the true image label at the time of segmentation, and  $w$  is a weight that determines the importance of an image. Normalization is essential in deep neural network-based models. Stabilizing the learning process and reducing the number of required epochs is necessary. The training phase of the network uses batch normalization based on the affine transformation:

$$z = g(W_u + b) \quad (5)$$

$$z = g(BN(W_u)) \quad (6)$$

With  $W$  and  $b$  being the parameters of the CC-Unet model,  $g(\cdot)$  represents the activation function used which is ReLu. The CC-Unet model is trained using the Adam optimizer [21] with a learning rate of 0.01,  $\beta_1=0.9$ ,  $\beta_2=0.999$ ,  $\epsilon=10^{-7}$ , and 50 epochs.

### B. Metric Performance Model

The performance of the model is measured based on the prediction accuracy, which is evaluated using the Intersection over Union (IoU), as described in equations (9) and (10). The IoU value is calculated for each class separately and globally determined based on the average value.

$$IoU = \frac{\text{Intersection}}{\text{Union}} \quad (7)$$

$$IoU = \frac{TP}{TP+FP+FN} \quad (8)$$

In equations (7) and (8) above, True Positive (TP), False Positive (FP), and False Negative (FN) are the number of image pixels ( $x$ ) that are correctly, incorrectly, and externally classified as  $x$ .

### C. Cross Validation

Cross-validation (CV) is generally used to test model performance. Therefore, the k-Fold cross-validation method was applied in this research where  $D$  image data is randomly split into  $k$ -sub-datasets of  $D_1, D_2, D_3, \dots, D_k$  with the same size between sub-datasets. The primary goal of this research is to enhance the performance of the CC-Unet model compared to the Segnet and Refined UNet models. The pre-processing stage involves converting the original satellite image data in .cn format to .TIFF.

The dataset used in this research is 100 Himawari 8 satellite images, divided into 10 subsets. During the training phase, each model is trained for 50 epochs. The original image resolution is 3600 X 2400, which is reduced to 384 x 384 due to memory limitations. In deep learning, the Sara network must minimize the loss value. The loss value is used to describe the extent of the difference between the actual and the predicted labels. Therefore, the neural network learns to adjust the weights and bias values to achieve its minimum value by minimizing loss.

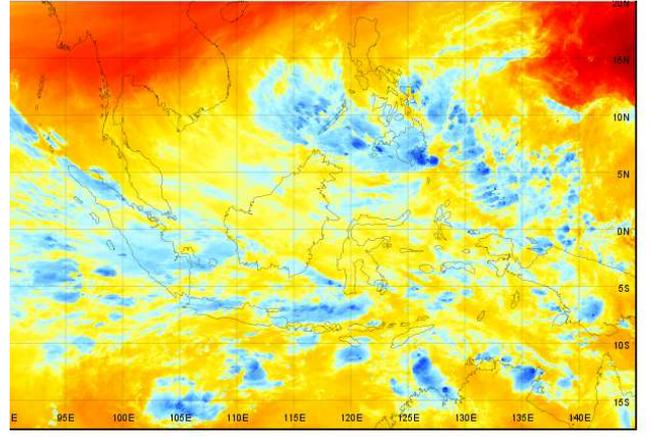


Fig. 6 Himawari-8 satellite cloud image plot results

The performance graph of the CC-Unet model with an epoch value of 50 iterations can be visualized in Figures 7 and 8. Figure 7 shows a loss condition during the training and model testing process. The loss of value decreases along with the increase in training data. Similarly, in Figure 7, the accuracy during the training process and model testing continue to rise as the epoch increases.

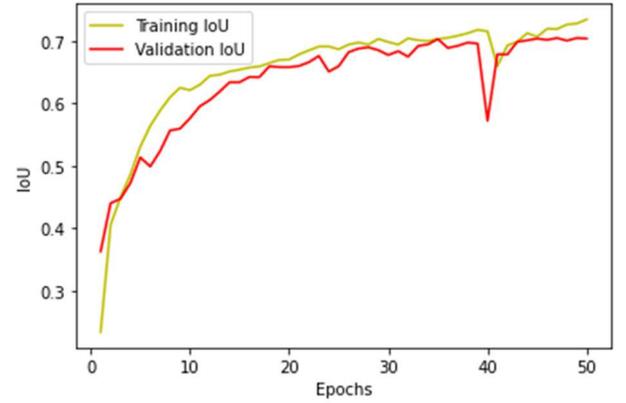


Fig. 7 CC-Unet model accuracy graph

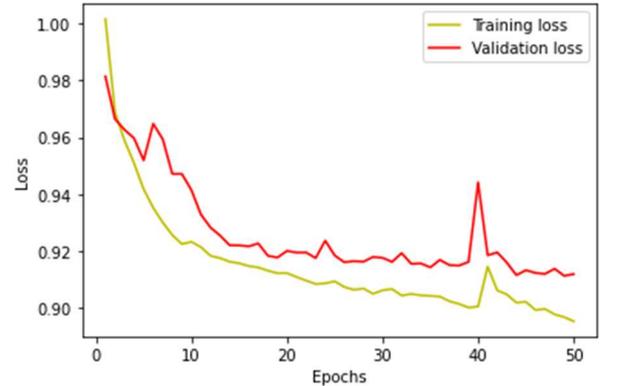


Fig. 8 Graph of the CC-Unet model loss.

The performance of the CC-Unet model is compared to its predecessor model. Table 2 shows the average accuracy of the U-Net and CC-Unet models on the classification of convective cloud images in this study. Performance testing was carried out in 10 batches by monitoring performance accuracy. CC-Unet and U-Net have average accuracies of 97.29% and 94.17%

TABLE II  
LOSS AND ACCURACY TESTING

	CC-Unet	SegNet	Refined Unet
Batch 1	<b>0,9705</b>	0,9305	0,9587
Batch 2	<b>0,9822</b>	0,9509	0,9665
Batch 3	<b>0,9745</b>	0,9698	0,9677
Batch 4	<b>0,9767</b>	0,9667	0,9567
Batch 5	<b>0,9817</b>	0,9795	0,9598
Batch 6	0,9620	<b>0,9800</b>	0,9512
Batch 7	<b>0,9870</b>	0,9710	0,9676
Batch 8	0,9590	0,9790	<b>0,9875</b>
Batch 9	<b>0,9877</b>	0,9398	0,9277
Batch 10	0,9834	0,9534	0,9534
<b>Rata-Rata</b>	<b>0,9765</b>	<b>0,9621</b>	<b>0,9597</b>

From the results in Table 2, CC-Unet provides more accurate segmentation compared with U-Net, SegNet, and RefinedNet. Visually, Figure 9 shows a simple case of image segmentation. From Figure 8, we can find that CC-Unet has a better result than another method. CC-Unet can capture cloud areas that U-net, SegNet, and RefinedNet missed.

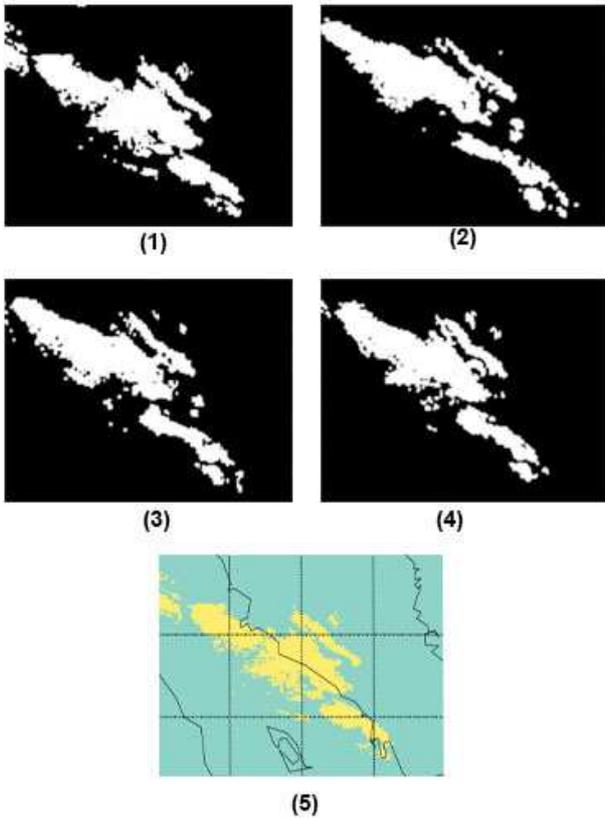


Fig. 9 Segmentation result of batch 1. (1) CC-Unet (2) U-Net (3) SegNet (4) Refined Net (5) Original image

From the results in Table 2, CC-Unet provides more accurate segmentation results. For example, Figure 9 shows a simple case. The segmentation accuracies of CC-Unet, U-net, SegNet, and Refined Net were 92.32 %, 87, 37 %, 90, 03%, and 89, 01 %, respectively. CC-Unet gives the best segmentation accuracy compared to the other models. A statistical significance test approach was used to test whether the performance results of the models in Figure 2 are significant. The Wilcoxon test [50] was applied at a significance level of 95% or % error rate of 5% (0.05).

Wilcoxon test is statistically different, assuming the p-value <0.05.

TABLE III  
THE RESULTS OF THE WILCOXON TEST FOR THE CC-UNET MODEL

VS	R <sup>+</sup>	R <sup>-</sup>	Exact P-value	Asymptotic P-value
SegNet	47.5	7.5	0.04296	0.034301
Refined U-Net	47.0	8.0	0.04882	0.041491

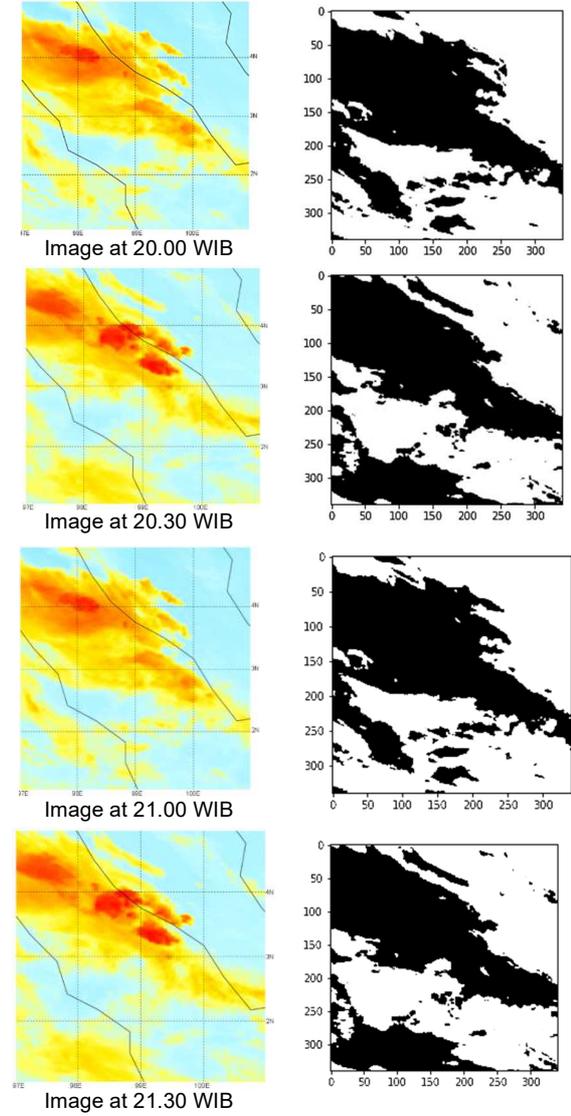


Fig 10 Himawari-8 satellite image and validation results as well as rain prediction in the Medan area using the CC-Unet model

To determine whether the predicted image is genuinely segmented to test the likeness or similarity level of the ground truth to the predicted images. The similarity test is evaluated based on the root mean square error (rmse) value. The rmse value between the ground truth and predicted images is 0.0218. It is classified as very small, meaning the image's similarity is high. This verification compares the actual event with the results of cloud image segmentation using CC-Unet. Verification was carried out on two different geographical structures, namely in Parapat and the urban areas, such as Medan City.

Based on extreme weather events that occurred on October 30, 2021, the community in the Aur Village area, Medan, North Sumatra, experienced significant flooding from Saturday, October 30 to Sunday, October 31, 2021. The flood height was about one meter, approximately an adult's waist. One of the worst spots occurred in Aur Village, on the banks of the Deli River, where floods have inundated since Saturday night.

Visually, the classification results showed that the CC-Unet model can display the results of convective cloud images close to the Himawari-8 satellite image. The verification and validation results of the convective cloud image with the CC-Unet model are used to predict rainfall and early warning of flood in Medan City.

#### IV. CONCLUSION

The CC-Unet model has better accuracy by 97.29% compared to the U-Net model by 94.17%. The ground truth image results are compared with the predicted image based on the root mean square error (rmse) value of 0.0218, which is very small. The significance test of the CC-Unet model with the Wilcoxon method at a significance level of 95% or an error level of 5% (0.05) gave different performance results to the U-Net model with the p-value of the U-Net VS CC-Unet model being  $0.0015258 < 0.05$ . Therefore, the accuracy of CC-Unet is significantly different from U-Net. These results indicate that the level of similarity between the ground truth image and the predicted image is very close. Additionally, the CC-Unet is one of the models on the deep neural network (DNN) successfully used to predict heavy rain based on the classification of convective cloud images.

#### ACKNOWLEDGMENT

This research is supported by Universitas Sumatera Utara (USU), Universitas Methodist Indonesia and Agency of Indonesia Meteorology, Climatology and Geophysics (BMKG).

#### REFERENCES

- [1] M. Tshuma, J. A. Belle, and A. Neube, "An Analysis of Factors Influencing Household Water, Sanitation, and Hygiene (WASH) Experiences during Flood Hazards in Tsholotsho District Using a Seemingly Unrelated Regression (SUR) Model," *Water*, vol. 15, no. 2, 2023. doi:10.3390/w15020371.
- [2] D. D. Harisdani and D. Lindarto, "Partisipasi Masyarakat Dalam Penggunaan Teknik Biopori Untuk Mengendalikan Banjir Kota (Studi Kasus : Kelurahan Tanjung Rejo – Medan)," *NALARS*, vol. 17, no. 2, p. 97, 2018, doi:10.24853/nalars.17.2.97-104.
- [3] B. I. Nasution, F. M. Saputra, R. Kurniawan, A. N. Ridwan, A. Fudholi, and B. Sumargo, "Urban vulnerability to floods investigation in Jakarta, Indonesia: A hybrid optimized fuzzy spatial clustering and news media analysis approach," *Int. J. Disaster Risk Reduct.*, vol. 83, p. 103407, 2022, doi:10.1016/j.ijdr.2022.103407.
- [4] D. R. Prabawadhani, B. Harsoyo, T. H. Seto, and B. R. Prayoga, "Karakteristik Temporal Dan Spasial Curah Hujan Penyebab Banjir Di Wilayah Dki Jakarta Dan Sekitarnya," *J. Sains Teknol. Modif. Cuaca*, vol. 17, no. 1, p. 21, 2016, doi:10.29122/jstmc.v17i1.957.
- [5] D. F. Saragih, "Problems and Solutions of the Flood Control Program in Medan City and its Surroundings," vol. 9, no. January, pp. 47–58, 2023, doi:10.22146/jcef.4784.
- [6] N. C. Dang, M. N. Moreno-García, and F. De la Prieta, "Sentiment Analysis Based on Deep Learning: A Comparative Study," *Electronics*, vol. 9, no. 3, 2020. doi:10.3390/electronics9030483.
- [7] M. I. Razzak, S. Naz, and A. Zaib, "Deep Learning for Medical Image Processing: Overview, Challenges and the Future BT - Classification

in BioApps: Automation of Decision Making," N. Dey, A. S. Ashour, and S. Borra, Eds., Cham: Springer International Publishing, 2018, pp. 323–350. doi:10.1007/978-3-319-65981-7\_12.

- [8] J. Chai, H. Zeng, A. Li, and E. W. T. Ngai, "Deep learning in computer vision: A critical review of emerging techniques and application scenarios," *Mach. Learn. with Appl.*, vol. 6, p. 100134, 2021, doi:10.1016/j.mlwa.2021.100134.
- [9] L. She, H. K. Zhang, Z. Li, G. de Leeuw, and B. Huang, "Himawari-8 Aerosol Optical Depth (AOD) Retrieval Using a Deep Neural Network Trained Using AERONET Observations," *Remote Sens.* 2020, Vol. 12, Page 4125, vol. 12, no. 24, p. 4125, Dec. 2020, doi:10.3390/RS12244125.
- [10] C. Liu *et al.*, "A Machine Learning-based Cloud Detection Algorithm for the Himawari-8 Spectral Image," *Adv. Atmos. Sci.* 2021, pp. 1–14, Jun. 2021, doi:10.1007/S00376-021-0366-X.
- [11] S. A. Ackerman, K. I. Strabala, W. P. Menzel, R. A. Frey, C. C. Moeller, and L. E. Gumley, "Discriminating clear sky from clouds with MODIS," *J. Geophys. Res.*, vol. 103, no. D24, pp. 32141–32157, 1998, doi:10.1029/1998jd200032.
- [12] M. B. Baker and T. Peter, "Small-scale cloud processes and climate," *Nature*, vol. 451, no. 7176, pp. 299–300, Jan. 2008, doi:10.1038/nature06594.
- [13] M. Mohtasham Moein *et al.*, "Predictive models for concrete properties using machine learning and deep learning approaches: A review," *J. Build. Eng.*, vol. 63, p. 105444, 2023, doi:10.1016/j.job.2022.105444.
- [14] J. Chen, G.-Q. Zeng, W. Zhou, W. Du, and K.-D. Lu, "Wind speed forecasting using nonlinear-learning ensemble of deep learning time series prediction and extremal optimization," *Energy Convers. Manag.*, vol. 165, pp. 681–695, 2018, doi:10.1016/j.enconman.2018.03.098.
- [15] T. Bai, D. Li, K. Sun, Y. Chen, and W. Li, "Cloud detection for high-resolution satellite imagery using machine learning and multi-feature fusion," *Remote Sens.*, vol. 8, no. 9, p. 715, Sep. 2016, doi:10.3390/rs8090715.
- [16] B. A. Baum *et al.*, "MODIS cloud-top property refinements for Collection 6," *J. Appl. Meteorol. Clim.*, vol. 51, no. 6, pp. 1145–1163, Jun. 2012, doi:10.1175/jamc-d-11-0203.1.
- [17] N. Chen *et al.*, "New neural network cloud mask algorithm based on radiative transfer simulations," *Remote Sens. Environ.*, vol. 219, pp. 62–71, Dec. 2018, doi:10.1016/j.rse.2018.09.029.
- [18] R. A. Frey *et al.*, "Cloud detection with MODIS. Part I: Improvements in the MODIS cloud mask for Collection 5," *J. Atmos. Ocean. Technol.*, vol. 25, no. 7, pp. 1057–1072, 2008, doi:10.1175/2008jtecha1052.1.
- [19] K. Bessho *et al.*, "An introduction to Himawari-8/9-Japan's new-generation geostationary meteorological satellites," *J. Meteor. Soc. Japan*, vol. 94, no. 2, pp. 151–183, 2016, doi:10.2151/jmsj.2016-009.
- [20] J. Gomis-Cebolla, J. C. Jimenez, and J. A. Sobrino, "MODIS probabilistic cloud masking over the Amazonian evergreen tropical forests: A comparison of machine learning-based methods," *Int. J. Remote Sens.*, vol. 11, no. 1, pp. 185–210, Jan. 2020, doi:10.1080/01431161.2019.1637963.
- [21] M. Rayat Intiaz Hossain, L. Sigal, and J. J. Little, "Semantically Enhanced Global Reasoning for Semantic Segmentation," *arXiv e-prints*, p. arXiv:2212.03338, Dec. 2022. doi:10.48550/arXiv.2212.03338.
- [22] Y. Guo, Y. Liu, T. Georgiou, and M. S. Lew, "A review of semantic segmentation using deep neural networks," *Int. J. Multimed. Inf. Retr.*, vol. 7, no. 2, pp. 87–93, 2018, doi:10.1007/s13735-017-0141-z.
- [23] X. Liu, D. Zhang, X. Zhu, and J. Tang, "VCT-NET: An Octa Retinal Vessel Segmentation Network Based on Convolution and Transformer," in *2022 IEEE International Conference on Image Processing (ICIP)*, 2022, pp. 2656–2660. doi:10.1109/ICIP46576.2022.9898038.
- [24] F. Lin, C. N. Gao, and K. D. Yamada, "An Effective Convolutional Neural Network for Visualized Understanding Transboundary Air Pollution Based on Himawari-8 Satellite Images," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, doi:10.1109/LGRS.2021.3102939.
- [25] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, 2017, doi:10.1109/TPAMI.2016.2644615.
- [26] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation BT - Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015," N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds., Cham: Springer International Publishing, 2015, pp. 234–241.

- [27] K. Hu, D. Zhang, and M. Xia, "Cdunet: Cloud detection unet for remote sensing imagery," *Remote Sens.*, vol. 13, no. 22, pp. 1–22, 2021, doi:10.3390/rs13224533.
- [28] L. Jiao, L. Huo, C. Hu, and P. Tang, "Refined UNet: UNet-based refinement network for cloud and shadow precise segmentation," *Remote Sens.*, vol. 12, no. 12, pp. 1–28, 2020, doi:10.3390/rs12122001.
- [29] M. S. Mekala, G. Srivastava, J. H. Park, and H.-Y. Jung, "An effective communication and computation model based on a hybrid graph-deeplearning approach for SIoT," *Digit. Commun. Networks*, vol. 8, no. 6, pp. 900–910, Dec. 2022, doi:10.1016/j.dean.2022.07.004.
- [30] W. Xie *et al.*, "SegCloud: A novel cloud image segmentation model using a deep convolutional neural network for ground-based all-sky-view camera observation," *Atmos. Meas. Tech.*, vol. 13, no. 4, pp. 1953–1961, 2020, doi:10.5194/amt-13-1953-2020.
- [31] A. Abdollahi, B. Pradhan, and A. M. Alamri, "An ensemble architecture of deep convolutional Segnet and Unet networks for building semantic segmentation from high-resolution aerial images," *Geocarto Int.*, vol. 37, no. 12, pp. 3355–3370, Jun. 2022, doi:10.1080/10106049.2020.1856199.
- [32] Z. Huang *et al.*, "A novel tongue segmentation method based on improved U-Net," *Neurocomputing*, vol. 500, pp. 73–89, Aug. 2022, doi:10.1016/j.neucom.2022.05.023.
- [33] L. Atika, S. Nurmaini, R. U. Partan, and E. Sukandi, "Image Segmentation for Mitral Regurgitation with Convolutional Neural Network Based on UNet, Resnet, Vnet, FractalNet and SegNet: A Preliminary Study," *Big Data and Cognitive Computing*, vol. 6, no. 4, 2022, doi:10.3390/bdcc6040141.
- [34] A. Iqbal, M. Sharif, M. A. Khan, W. Nisar, and M. Alhaisoni, "FF-UNet: a U-Shaped Deep Convolutional Neural Network for Multimodal Biomedical Image Segmentation," *Cognit. Comput.*, vol. 14, no. 4, pp. 1287–1302, 2022, doi:10.1007/s12559-022-10038-y.
- [35] S. Niu, Q. Chen, L. de Sisternes, Z. Ji, Z. Zhou, and D. L. Rubin, "Robust noise region-based active contour model via local similarity factor for image segmentation," *Pattern Recognit.*, vol. 61, pp. 104–119, 2017, doi:10.1016/j.patcog.2016.07.022.
- [36] Z. Shaukat, Q. ul A. Farooq, S. Tu, C. Xiao, and S. Ali, "A state-of-the-art technique to perform cloud-based semantic segmentation using deep learning 3D U-Net architecture," *BMC Bioinformatics*, vol. 23, no. 1, p. 251, 2022, doi:10.1186/s12859-022-04794-9.
- [37] Z. Zhang, A. Iwasaki, G. Xu, and J. Song, "Cloud detection on small satellites based on lightweight U-net and image compression," *J. Appl. Remote Sens.*, vol. 13, no. 2, p. 26502, Apr. 2019, doi:10.1117/1.JRS.13.026502.
- [38] Q. Deng, P. Lu, S. Zhao, and N. Yuan, "U-Net: A deep-learning method for improving summer precipitation forecasts in China," *Atmos. Ocean. Sci. Lett.*, p. 100322, Dec. 2022, doi:10.1016/J.AOSL.2022.100322.
- [39] J. Wang, M. Hadjikakou, R. J. Hewitt, and B. A. Bryan, "Simulating large-scale urban land-use patterns and dynamics using the U-Net deep learning architecture," *Comput. Environ. Urban Syst.*, vol. 97, p. 101855, Oct. 2022, doi:10.1016/j.compenvurbsys.2022.101855.
- [40] M. Marhamati *et al.*, "LAIU-Net: A learning-to-augment incorporated robust U-Net for depressed humans' tongue segmentation," *Displays*, vol. 76, p. 102371, Jan. 2023, doi: 10.1016/j.displa.2023.102371.
- [41] F. Shojaie and Y. Baleghi, "EFASPP U-Net for semantic segmentation of night traffic scenes using fusion of visible and thermal images," *Eng. Appl. Artif. Intell.*, vol. 117, p. 105627, Jan. 2023, doi:10.1016/j.engappai.2022.105627.
- [42] C. Beeche *et al.*, "Super U-Net: A modularized generalizable architecture," *Pattern Recognit.*, vol. 128, p. 108669, Aug. 2022, doi:10.1016/j.patcog.2022.108669.
- [43] S. V. and I. G., "Encoder Enhanced Atrous (EEA) Unet architecture for Retinal Blood vessel segmentation," *Cogn. Syst. Res.*, vol. 67, pp. 84–95, 2021, doi:10.1016/j.cogsys.2021.01.003.
- [44] X. Yang, Z. Li, Y. Guo, and D. Zhou, "DCU-net: a deformable convolutional neural network based on cascade U-net for retinal vessel segmentation," *Multimed. Tools Appl.*, vol. 81, no. 11, pp. 15593–15607, 2022, doi:10.1007/s11042-022-12418-w.
- [45] T. Zhou, S. Canu, and S. Ruan, "Automatic COVID-19 CT segmentation using U-Net integrated spatial and channel attention mechanism," *Int. J. Imaging Syst. Technol.*, vol. 31, no. 1, pp. 16–27, Mar. 2021, doi:10.1002/ima.22527.
- [46] F. Wang and Y. Zai, "Image segmentation and flow prediction of digital rock with U-net network," *Adv. Water Resour.*, p. 104384, Jan. 2023, doi: 10.1016/j.advwatres.2023.104384.
- [47] C. H. Chuang, K. Y. Chang, C. S. Huang, and T. P. Jung, "IC-U-Net: A U-Net-based Denoising Autoencoder Using Mixtures of Independent Components for Automatic EEG Artifact Removal," *Neuroimage*, vol. 263, p. 119586, Nov. 2022, doi:10.1016/j.neuroimage.2022.119586.
- [48] A. Lou, S. Guan, and M. H. Loew, "DC-UNet: rethinking the U-Net architecture with dual channel efficient CNN for medical image segmentation," p. 98, 2021, doi: 10.1117/12.2582338.
- [49] X. Liu, H. Du, J. Xu, and B. Qiu, "DBGAN: A dual-branch generative adversarial network for undersampled MRI reconstruction," *Magn. Reson. Imaging*, vol. 89, pp. 77–91, 2022, doi:10.1016/j.mri.2022.03.003.
- [50] V. A. Krylov, G. Moser, A. Voisin, S. B. Serpico, and J. Zerubia, "Change detection with synthetic aperture radar images by Wilcoxon statistic likelihood ratio test," *Proc. - Int. Conf. Image Process. ICIP*, pp. 2093–2096, 2012, doi: 10.1109/ICIP.2012.6467304.