International Journal of Scientific Research in Science and Technology
Available online at : www.ijsrst.com
Print ISSN: 2395-6011 | Online ISSN: 2395-602X
doi : https://doi.org/10.32628/IJSRST52310631

Effect of noise on the robustness of MobileNetV2+U-Net semantic segmentation model for MRI images

Gunawan Nur Cahyo, Choirul Anam, Kusworo Adi
Department of Physics, Faculty of Sciences and Mathematics, Diponegoro University, Jl. Prof. Soedarto SH, Tembalang, Semarang 50275, Central Java, Indonesia
Corresponding author: gunawannc0@gmail.com

ARTICLEINFO

Article History:

Accepted: 10 Nov 2023
Published: 30 Nov 2023

Publication Issue
Volume 10, Issue 6
November-December-2023
Page Number
209-217

ABSTRACT

This study aims to evaluate effect of noise on the robustness of semantic segmentation models for Magnetic Resonance Imaging (MRI) head images with tumor. We implemented the MobileNetV2+U-Net architectural model. We tested the segmentation model with Gaussian and Poisson noises in various levels. The addition of noise was performed five iterations with a variance of 0.01 each iteration. We carried out evaluations by examining the segmentation results, loss function values, accuracy and dice score. Based on the results, increase in noise affects model performance. Evaluation using loss function shows that graph instability is influenced by the noise level. The accuracy results on the highest and lowest validation data were 99.47% and 98.99% for Gaussian noise and 99.64% and 99.5% for Poisson noise. Apart from that, the highest and lowest dice scores were 82.80% and 69.18% for Gaussian noise and 87.83% and 83.10% for Poisson noise. We recommend training the segmentation model using noisy data so that the model can adapt to noisy images.

Keywords: Semantic Segmentation, MobileNetV2+U-Net, U-Net, Noise

I. INTRODUCTION

Semantic segmentation is an algorithm technique based on the Deep Learning (DL) which is applied in computer vision. The advancement of this technique raises the ability to carry out classification at pixel level [1]. This algorithm has the ability to group parts of the same image by entering labels into the same object class [2]. The aim of semantic segmentation is to obtain smooth infers that are useful for predicting labels by assigning a class label that corresponds to the object or region in each image pixel [3]. Several architectural models have been successful in presenting segmentation semantically, including Fully Convolutional Network (FCN) [4], U-Net [5], SegNet [6], and U-Net++ [7].

U-Net is a semantic segmentation architectural model that is successful in segmenting biomedical images. Ronneberger et al. proposed this architectural model based on an encoder-decoder structure [5]. The main feature of U-Net is that it has a symmetrical structure model and a skip connection between the encoder and decoder. The function of this task is to combine low-level and high-level features to preserve image details and help to produce an accurate segmentation [8].

Segmentation using DL requires large data for training, persistent labelling, and good image quality. When the training process is performed, it is noted that noise is important for obtaining good image quality. The presence of noise can affect segmentation model training and can greatly influenced by poor labelling [9]. Disturbances in the training process can affect decision making by the model [10].

Techniques for reducing the effect of noise can be divided into two categories, namely: identifying noisy data and filtering the data to train the model, and training the model directly using noisy data so that the model is adapt to noisy data [11]. To the best of our knowledge, there is still little discussion regarding the effect of noise on the stability of segmentation models. Therefore, this study aims to examine the effect of noise on training architectural semantic segmentation models for Magnetic Resonance Imaging (MRI) head images containing tumor.

II. METHODS AND MATERIAL

A. Datasets

We used a public dataset of 330 images obtained from kaggle [12] and figshare [13]. A sample of the dataset used in this study is depicted in Figure 1. This dataset contains information on a collection of medical images using MRI modality with a diagnosis of brain tumor. The anatomical sections contained in the dataset are axial with a matrix size of 512 × 512 pixels, and also have -jpg and -mat extensions. We selected the data with the following criteria: head anatomy with clear tumor edges in the axial plane. This criterion is applied so that we obtain labels that tend to be consistent.

B. Model training

Training was carried out using a computer with a specification of Intel i7-6700HQ processor, Nvidia GeForce GTX 960M graphics card, and 16 GB memory. To run the DL platform and image processing, we used Python 3. The main library used in building this architectural model was Tensorflow-gpu 2.8.

C. Setup datasets

The purpose of determining setup datasets is to obtain consistent image conditions and lead to a more in-depth evaluation. We modified the images that have some artifact in the background. This process will improve the labelling performance in the training stage. The image background was uniformed to black (R:0, B:0, G:0). This image improvement was carried out using Adobe Photoshop. Meanwhile, to determine labels, we used the labelme framework [14].
Next, Gaussian and Poisson noise at a variance level of 0.01 were added to the image. The addition of noise was carried out in stages over five iterations. Hence, the noise level increased with each iteration. We used the scikit-image library to generate the noise. We also determined the distribution of the dataset in the proportion of 80:10:10 for training, validation and testing, respectively. Meanwhile, the hyperparameter design used was batch size 2, epoch 100, and Adam optimizer.

D. Model architecture

The model architecture used in this study was the MobileNetV2+U-Net. This model is based on the U-Net architectural model [5]. This model has a symmetrical "U" shaped structure consisting of an encoder and decoder. An important part in this model architecture is the presence of a skip connection between the two parts of the encoder and decoder. This section has a function to transmit low level features to order local information [15]. This section is useful in recovering unclear object details and functions to produce more detailed segmentation masks on complex backgrounds [16]. On the other hand, the encoder applied in developing this model was MobileNetV2 which has the advantage of the few parameters obtained. The trained model can help to

speed up the training process, and the trained encoder can help the model achieve better performance [17]. The architectural scheme that we developed adopted the model developed by [18], illustrated in Figure 2.

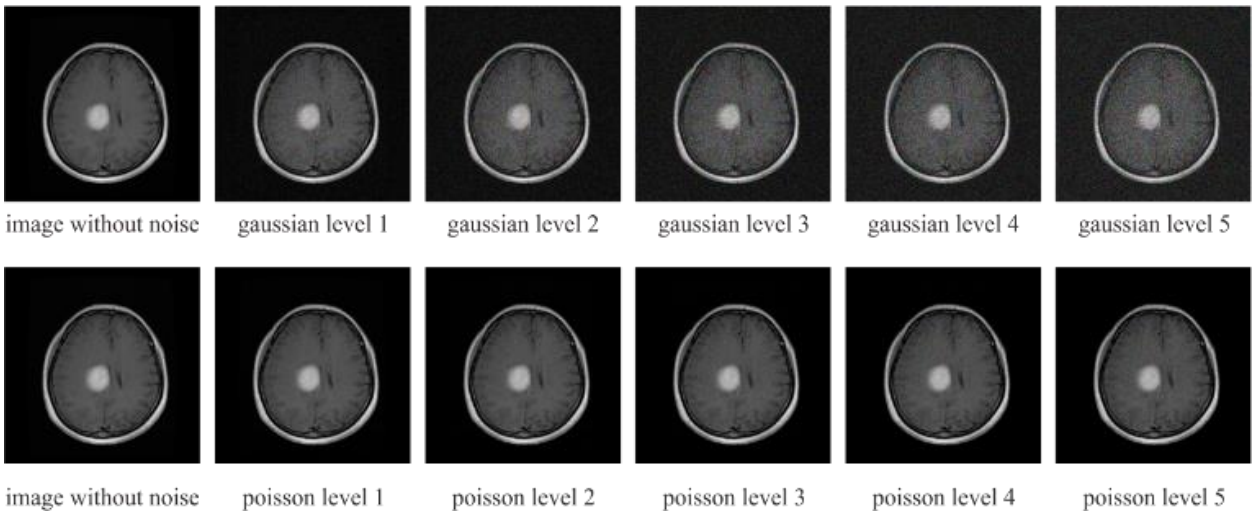


Figure 1. Sample of image dataset.

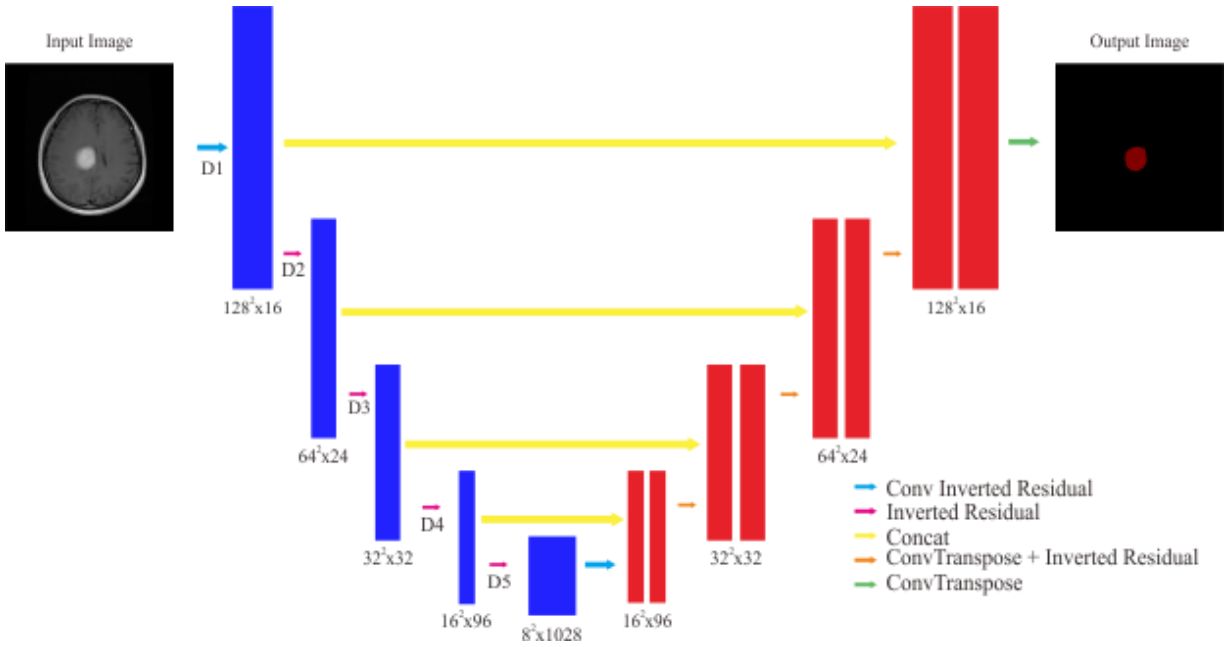


Figure 2. MobileNetV2+U-Net architecture model.

E. Matrix evaluation

The model performance was evaluated by observing the loss function, accuracy, and dice score from training and validation data. The loss function is a value that functions to calculate prediction uncertainty based on how much the prediction varies from the actual value. This function is the sum of the errors made for each sample in the training and validation dataset. Accuracy is a parameter to measure the performance of a classification model. This method is used to calculate predictions. In general, this value is presented in percentage units. The dice score is a parameter for measuring image similarity between ground truth images and images predicted by the model. This function can be calculated using Equation (1).

$$DCS = \frac{2TO}{2TP + FP + FN} = \frac{2|X \cap Y|}{|X| + |Y|} \quad (1)$$

where |X ∩ Y| represents an element between sets X and Y, |X| represents the number of elements in sets X, and |Y| represents the number of elements in set Y.

III. RESULTS AND DISCUSSION

A. Segmentation results

Figures 3 and 4 show the segmentation results of the model we developed for Gaussian and Poisson noises, respectively. T1 and T2 are examples of segmentation with training data input, while V1 and V2 are examples of segmentation with validation data input. In general, the segmentation results show correct segmentation for most cases. Several segmentation discrepancies were observed in V2 with Gaussian noise level 2, 3, and 4 (Figure 3). The discrepancies in the prediction results were influenced by the noise level, where the higher the noise level used, the farther the prediction results obtained were from the ground truth image. In Poisson noise, segmentation results generally show correct segmentation at all noise levels (Figure 4).

B. Loss function

Figure 5 and 6 display the loss function graphs for Gaussian and Poisson noises, respectively. It can be observed that for both types of noise, the higher noise level results in the more unstable the loss function. These results show that noise affects training performance. Higher noise makes it more challenging for the model to extract the features and recognize it into a correct category [19].

C. Accuracy

Tables 1 and 2 present the results of evaluating the accuracy of the developed model for Gaussian and Poisson noises. In general, accuracy for both noises is close to 100%, with the majority of values above 99%. However, the noise level is considered to still affect model accuracy. Adding noise at level 1 reduces the accuracy compared to without noise. However, for the Gaussian noise, the accuracy increases at level 3 and decreases gradually until level 5. Meanwhile, at the Poisson noise, the accuracy increased at level 2, decreases at levels 3 and 4, then increased at level 5.

The higher noise given to the training data lead to the lower the accuracy [20]. This was also reported by Jang et al. (2021) who used Deep Neural Networks (DNNs) for object classification to develop a visual system model. Similarly, our model shows a decrease in trend with noise level 1 to level 5, except at Gaussian noise level 2.

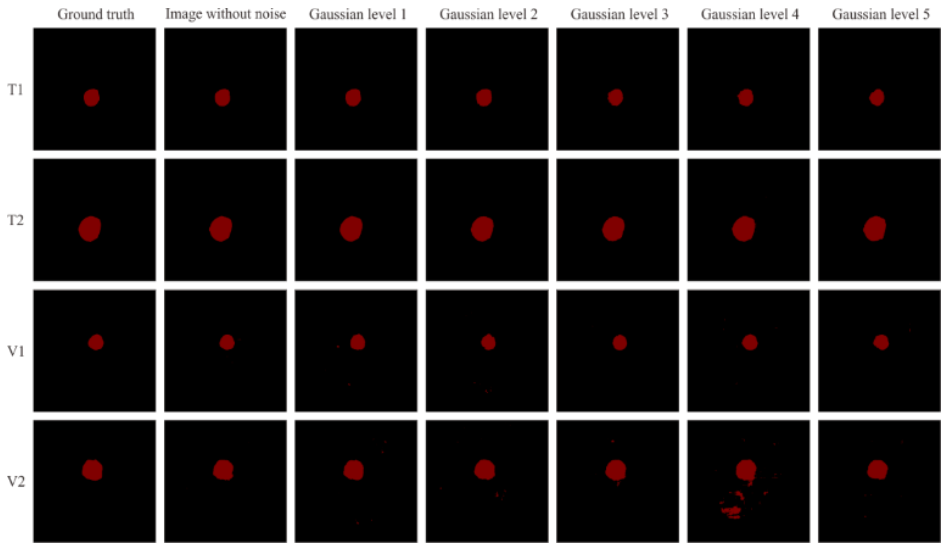


Figure 3. Segmentation results on images with Gaussian noise.

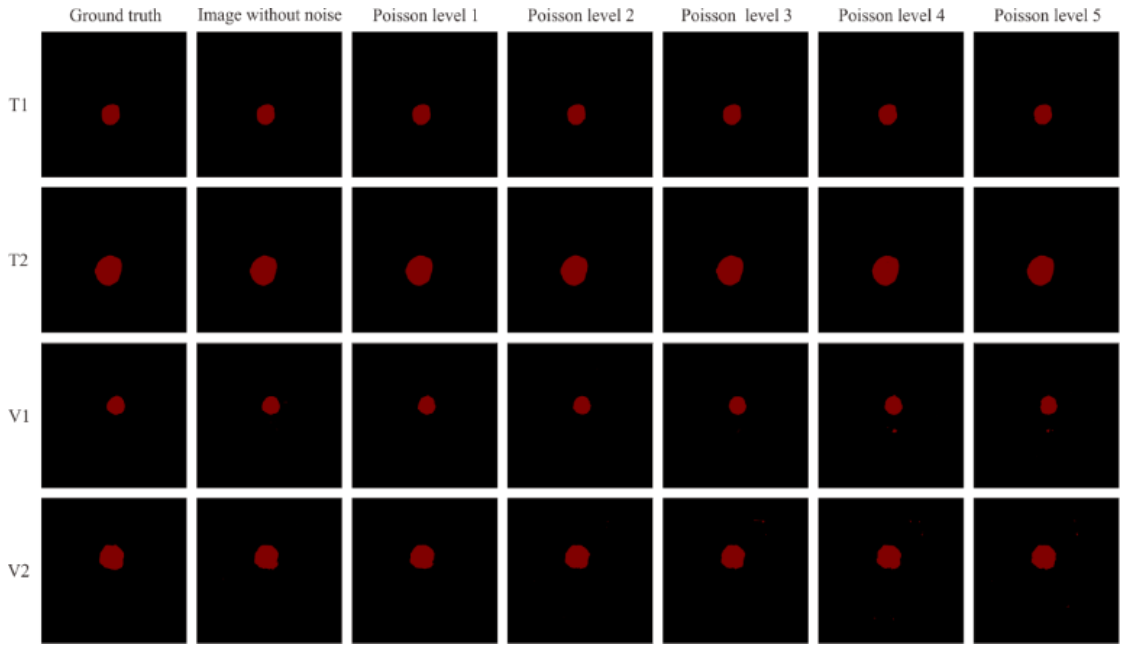
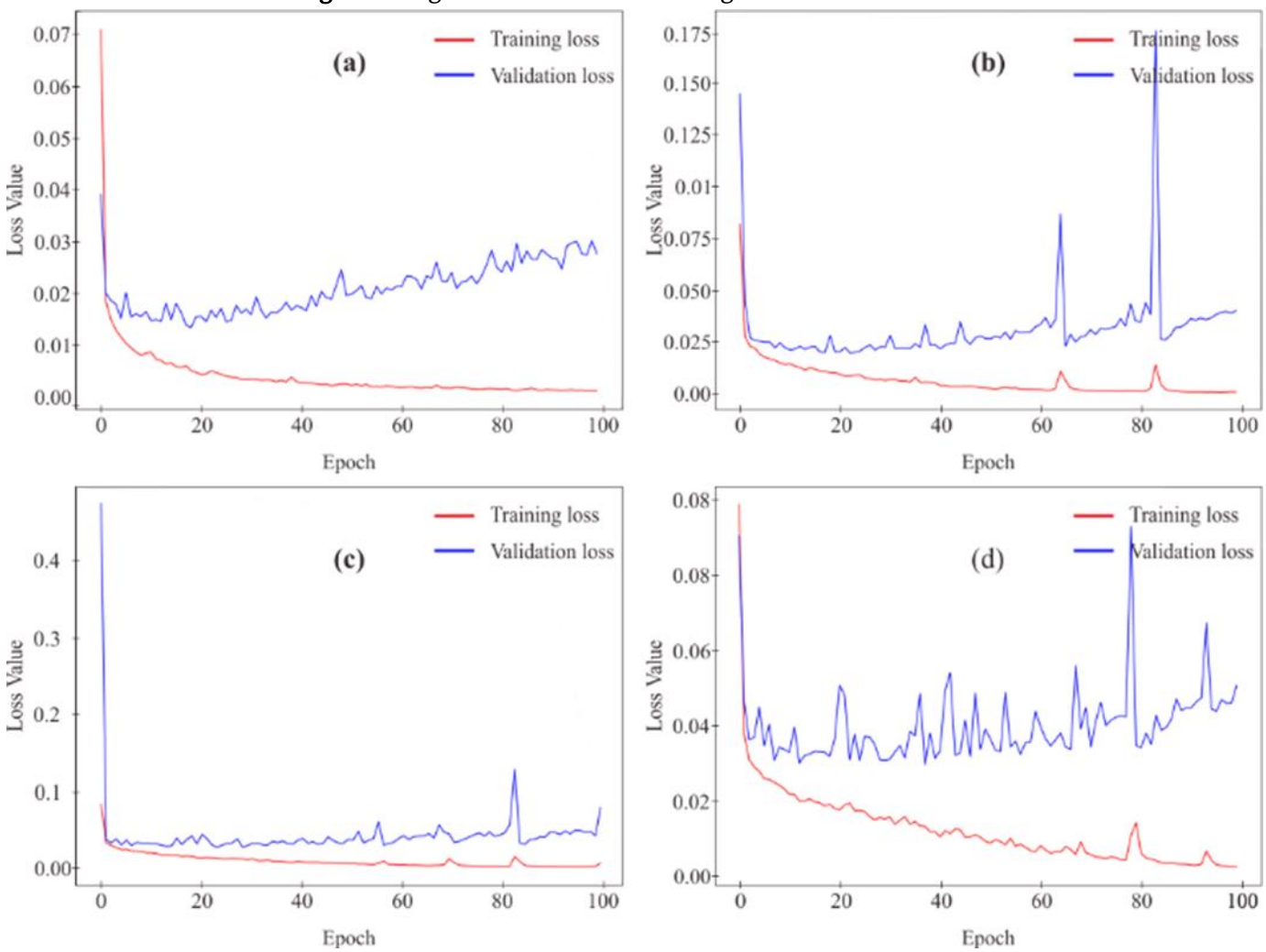


Figure 4. Segmentation results on images with Poisson noise



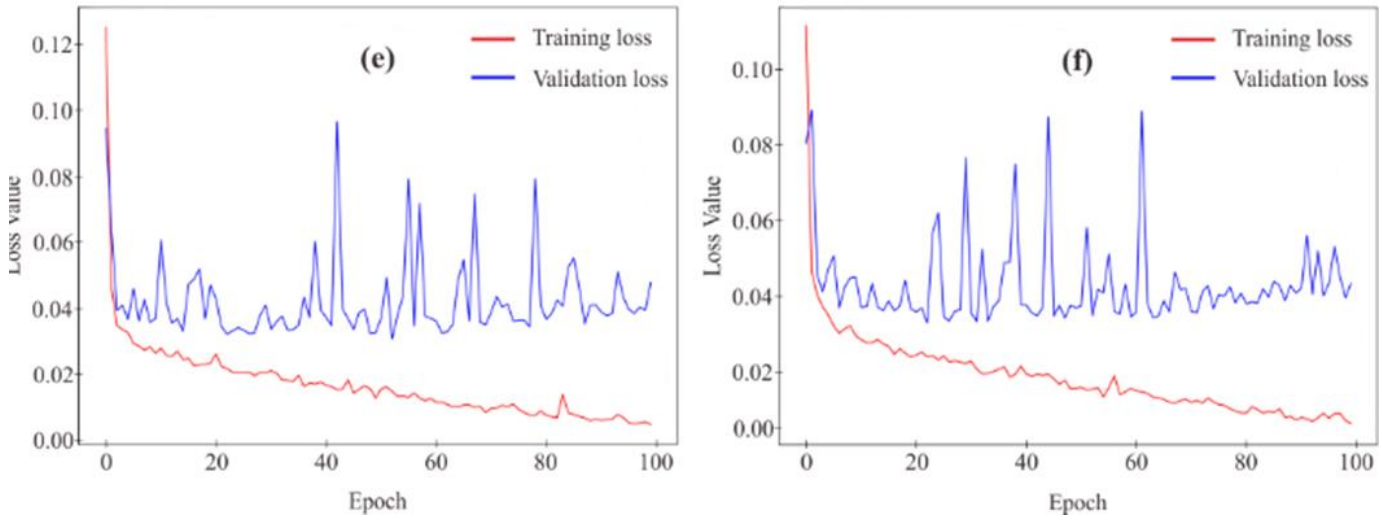


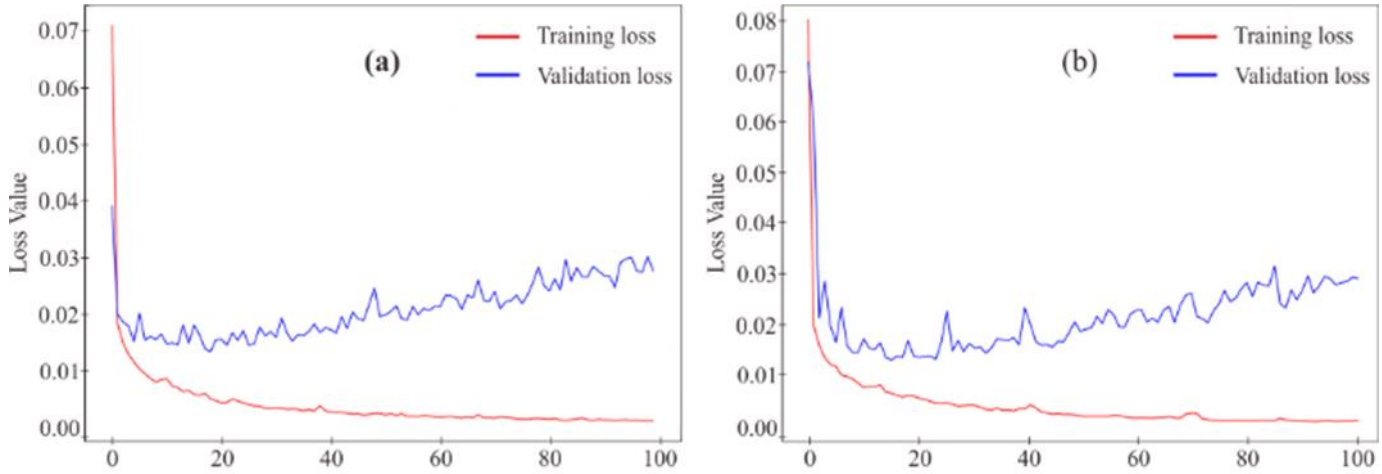
Figure 5. Loss function for image without and with Gaussian noise. (a) without noise, Gaussian noise level: (b) 1, (c) 2, (d) 3, (e) 4, and (f) 5.

Table 1. Accuracy results based on training.

Accuracy	Without noise	Level 1	Level 2	Level 3	Level 4	Level 5
Gaussian	99.96%	99.94%	99.78%	99.90%	99.82%	99.78%
Poisson	99.96%	99.95%	99.97%	99.93%	99.91%	99.96%

Table 2. Accuracy results based on validation.

Accuracy	Without noise	Level 1	Level 2	Level 3	Level 4	Level 5
Gaussian	99.62%	99.47%	99.08%	99.14%	99.14%	98.99%
Poisson	99.62%	99.64%	99.61%	99.5%	99.54%	99.57%



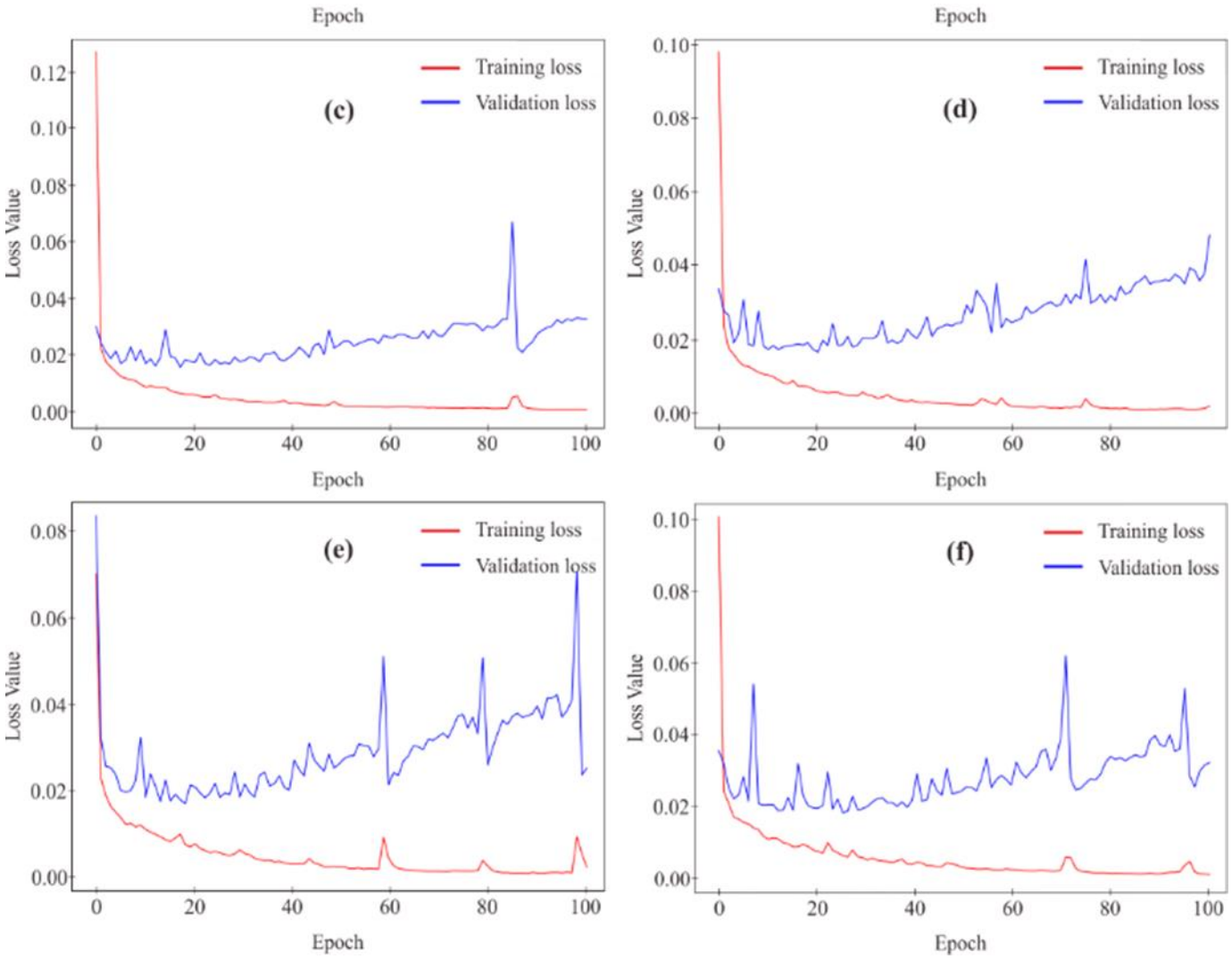


Figure 6. Loss function for image without and with Poisson noise. (a) without noise, Poisson noise level: (b) 1, (c) 2, (d) 3, (e) 4, and (f) 5.

D. Dice score

Dice score was calculated from the similarity of the ground truth image with the predicted image. The dice score was determined by taking into account both training data and validation data. Tables 3 and 4 present the dice scores for both types of noise.

The dice score results for the training and validation datasets show a decrease with each increasing noise level. However, at Gaussian noise level 2, the score is decreased drastically, then rises again at level 3, then decreases gradually at levels 4 and 5. Meanwhile, for Poisson noise, the score increases at level 2 in the training data, but decreases in validation. The score increases again at levels 4 and 5 for training and validation data.

Table 3. Dice score based on training data.

Dice score	Without noise	Level 1	Level 2	Level 3	Level 4	Level 5
Gaussian	98.71 %	97.75 %	79.19 %	95.29 %	91.59 %	89.65 %
Poisson	98.71 %	98.71 %	98.90 %	95.69 %	96.29 %	98.49 %

Table 4. Dice score based on validation data.

Dice score	Without noise	Level 1	Level 2	Level 3	Level 4	Level 5
Gaussian	88.03 %	82.80 %	69.18 %	78.87 %	75.14 %	73.10 %
Poisson	88.03 %	87.83 %	86.03 %	83.10 %	83.74 %	85.85 %

The dice score decreases as the noise level increases [22]. However, at a certain noise level, the dice score can increase. This phenomenon can be called an anomalous condition. In this study, an anomalous condition is observed when adding level 2 Gaussian noise, where the score decreases and increases again at level 3, then decreases gradually up to level 5. An anomaly was also reported by Kascenas et al. (2022) [23] when providing denoising autoencoder (DAE) noise (16×16 noise), which experienced an increase in dice score compared to giving noise at the DAE level (8×8 noise).

IV. CONCLUSION

In conclusion, noise affects the segmentation results, loss function, accuracy, and dice score of a semantic segmentation model. Segmentation results show correct predictions in most cases, both for Gaussian and Poisson noises. Evaluation using loss function shows that graph instability is influenced by the noise level. Accuracy results from validation data with Gaussian and Poisson noise were respectively obtained 99.47, 99.08, 99.14, 99.14, 98.99% and 99.64, 99.61, 99.5, 99.54, and 99.57%. Meanwhile, evaluation using dice scores for Gaussian and Poisson noise respectively obtained 82.80, 69.18, 78.87, 75.14, and 73.10% and 87.83, 86.03, 83.10, 83.74, and 85.85%. We recommend to use noisy data in the training stage so that the model can adapt to noisy images.

V. REFERENCES

[1]. Z. Luo, W. Yang, Y. Yuan, R. Gou, and X. Li, "Semantic segmentation of agricultural images: A survey," Inf. Process. Agric., pp. 1–15, 2023, doi: 10.1016/j.inpa.2023.02.001.

[2]. M. Thoma, "A Survey of Semantic Segmentation," pp. 1–16, 2016, [Online]. Available: http://arxiv.org/abs/1602.06541

[3]. A. Mueed, H. Ghulam, and M. Bhat, "A survey on instance segmentation : state of the art," Int. J. Multimed. Inf. Retr., vol. 9, no. 3, pp. 171–189, 2020, doi: 10.1007/s13735-020-00195-x.

[4]. J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," Nov. 2014, [Online]. Available: http://arxiv.org/abs/1411.4038

[5]. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," May 2015, [Online]. Available: http://arxiv.org/abs/1505.04597

[6]. V. Badrinarayanan, A. Kendall, R. Cipolla, and S. Member, "SegNet : A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," November 2015, [Online]. Available: https://arxiv.org/abs/1511.00561.

[7]. H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, and Y. Iwamoto, "UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation," April 2020, [Online]. Available: https://arxiv.org/abs/2004.08790.

[8]. R. Wang, S. Chen, C. Ji, J. Fan, and Y. Li, "Boundary-aware context neural network for medical image segmentation," Med. Image Anal., vol. 78, no. 8, pp. 1–10, 2022, doi: 10.1016/j.media.2022.102395.

[9]. Z. Lu, Z. Fu, T. Xiang, P. Han, L. Wang, and X. Gao, "Learning from weak and noisy labels for semantic segmentation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 3, pp. 486–500, 2017, doi: 10.1109/TPAMI.2016.2552172.

[10]. G. Cheng, H. Ji, and Y. Tian, "Walking on two legs: Learning image segmentation with noisy labels," Proc. 36th Conf. Uncertain. Artif. Intell. UAI 2020, vol. 124, pp. 330–339, 2020.

[11]. R. Yi, Y. Huang, Q. Guan, M. Pu, and R. Zhang, "Learning from Pixel-Level Label Noise: A New Perspective for Semi-Supervised Semantic Segmentation," IEEE Transactions on Image Processing, vol. 31, pp. 623–635, Jan. 2022, doi: 10.1109/tip.2021.3134142.

[12]. M. Nickparvar, "Brain Tumor MRI Dataset," 2021. https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset

[13]. J. Cheng, "Brain Tumor Dataset," 2016. https://figshare.com/articles/dataset/brain_tumor_dataset/1512427?file=7953679

[14]. K. Wada, "Labelme," 2022. https://github.com/wkentaro/labelme

[15]. H. Bi et al., "BPAT-UNet: Boundary Preserving Assembled Transformer UNet for Ultrasound Thyroid Nodule Segmentation," Comput. Methods Programs Biomed., p. 107614, 2023, doi: 10.1016/j.cmpb.2023.107614.

[16]. Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 11045 LNCS, pp. 3–11, 2018, doi: 10.1007/978-3-030-00889-5_1.

[17]. D. Lopes, L. Coelho, and M. F. Silva, "Development of a Collaborative Robotic Platform for Autonomous Auscultation," Appl. Sci., vol. 13, no. 3, 2023, doi: 10.3390/app13031604.

[18]. J. Jing, Z. Wang, M. Rätsch, and H. Zhang, "Mobile-Unet: An efficient convolutional neural network for fabric defect detection," Text. Res. J., vol. 92, no. 1–2, pp. 30–42, 2022, doi: 10.1177/0040517520928604.

[19]. Y. Lau, W. Sim, K. Chew, Y. Ng, Z. Arabee, and A. Salam, "Understanding how noise affects the accuracy of CNN image classification," J. Appl. Technol. Innov., vol. 5, no. 2, pp. 2600–7304, 2021.

[20]. T. H. Pranto, A. A. Noman, A. Noor, U. H. Deepty, and R. M. Rahman, "Effect of Label Noise on Multi-Class Semantic Segmentation: A Case Study on Bangladesh Marine Region," Appl. Artif. Intell., vol. 36, no. 1, 2022, doi: 10.1080/08839514.2022.2039348.

[21]. H. Jang, D. McCormack, and F. Tong, "Noise-trained deep neural networks effectively predict human vision and its neural responses to challenging images," PLoS Biol., vol. 19, no. 12, pp. 1–27, 2021, doi: 10.1371/journal.pbio.3001418.

[22]. A. Mortazi, N. Khosravan, D. A. Torigian, S. Kurugol, and U. Bağcı, "Weakly supervised segmentation by a deep geodesic prior," in Lecture Notes in Computer Science, 2019, pp. 238–246. doi: 10.1007/978-3-030-32692-0_28..

[23]. A. Kascenas, N. Pugeault, and A. Q. O'neil, "Denoising Autoencoders for Unsupervised Anomaly Detection in Brain MRI," Proc. Mach. Learn. Res. Rev., vol. 172, pp. 1–12, 2022.

Cite this article as :

Gunawan Nur Cahyo, Choirul Anam, Kusworo Adi, "Effect of noise on the robustness of MobileNetV2+U-Net semantic segmentation model for MRI images", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 10 Issue 6, pp. 209-217, November-December 2023. Available at doi : https://doi.org/10.32628/IJSRST52310631
Journal URL : https://ijsrst.com/IJSRST52310631