

CHAPTER V

CONCLUSION

This chapter serves as the concluding section of the thesis, presenting the conclusions drawn from the research findings while also providing recommendations that may serve as references for future studies. The conclusions presented summarize the entire discussion described in the preceding chapters and provide answers to the research questions stated in the problem formulation. In addition, the recommendations are intended to provide general guidance for future research development, both in terms of methodological approaches and the expansion of topics that have the potential to be explored further.

5.1. Conclusion

Based on the results of this study comparing U-Net-ResNet34 and the Modified Deep Semantic Binarization Network (MDSBN) for the binarization of Indonesian ancient documents/manuscripts, the following conclusions can be drawn.

1. Based on the testing evaluation using the micro average, MDSBN achieved slightly higher performance than U-Net-ResNet34 on most evaluation metrics. MDSBN achieved a Dice performance of 0.8545, an IoU result standing at 0.7459, a precision value reaching 0.8600, a recall performance 0.8490, and an RMSE of 0.1529. Meanwhile, U-Net-ResNet34 achieved a Dice performance of 0.8535, an IoU result standing at 0.7445, a precision value reaching 0.8581, a recall performance 0.8490, and an RMSE of 0.1535. The differences between the two models were very small, indicating that both achieved competitive segmentation performance. In addition, the baseline model results showed that U-Net achieved the highest Dice and IoU scores, whereas ResNet34 produced the lowest performance. These findings indicate that integrating U-Net with the ResNet34 encoder does not necessarily improve performance on the dataset used in this study.
2. In terms of computational requirements, MDSBN had fewer parameters than U-Net-ResNet34, with 15,327,681 parameters compared to 24,349,441 parameters. However, U-Net-ResNet34 outperformed MDSBN in inference

time and GPU memory usage, with an inference time of 2.38 ms per image and GPU memory usage of 866.05 MB during inference, whereas MDSBN required 9.08 ms per image and 894.38 MB. Therefore, MDSBN is more lightweight in terms of the number of parameters, whereas U-Net-ResNet34 is more computationally efficient during inference.

3. MDSBN was shown to achieve performance that was comparable to, and even slightly surpassed, U-Net-ResNet34 on the testing data, particularly in terms of the Dice, IoU, precision, and RMSE metrics. This indicates that MDSBN can serve as a competitive alternative for ancient document segmentation while requiring fewer parameters. However, because its inference time and GPU memory usage remain higher, MDSBN is more suitable when the primary priority is model size, whereas U-Net-ResNet34 is more appropriate when inference efficiency is the primary concern.

5.2. Recommendations

Based on the results of this study, several recommendations can be proposed for future research.

1. Future studies may use larger and more diverse datasets. Indonesian ancient documents/manuscripts exhibit wide variations in script type, writing medium, degree of degradation, illumination, background texture, and ink thickness. Incorporating datasets from various types of manuscripts can help the model learn a broader range of document characteristics and improve its generalization capability.
2. The ground truth generation process can be improved through a more rigorous validation procedure. In this study, some masks were obtained through a semi-automatic process followed by manual refinement. Future studies may involve multiple annotators or implement a re-evaluation process to ensure greater consistency of the ground truth, particularly for text regions that are thin, faded, or difficult to distinguish from the background.
3. Future studies may incorporate model optimization techniques such as pruning, quantization, or knowledge distillation. These methods can be used to reduce model size, accelerate inference time, and decrease memory usage without significantly compromising segmentation performance.