

CHAPTER V

CLOSING

5.1 Conclusion

Based on the results of this study, it can be concluded that the Hybrid CNN–Swin Transformer architecture was successfully and effectively applied to classify malaria disease in microscopic red blood cell images. This model is able to accurately recognize and distinguish between infected and uninfected red blood cells by leveraging the strengths of CNNs in extracting local features and the Swin Transformer in modeling global spatial relationships through the window-based self-attention mechanism. Model evaluation indicates that all three tested architectures (CNN, Swin Transformer, and Hybrid CNN–Swin Transformer) are capable of achieving high accuracy with favorable clinical metrics, thereby demonstrating potential for use as an artificial intelligence-based malaria diagnosis support tool. The answers to the research questions posed in this study are presented as follows:

1. **How can a binary malaria classification model based on the Hybrid CNN–Swin Transformer be designed and implemented?**

The Hybrid CNN–Swin Transformer model was successfully designed and implemented by integrating the CNN as a local feature extractor and the Swin Transformer as a global context modeling component. The integration point between the two was implemented through a bridge layer that converts the CNN feature map into a token sequence of dimension 1×96 as input to the Swin Transformer, thereby enabling local feature representations, such as erythrocyte texture and Plasmodium parasite morphological patterns, to be enriched with an understanding of global spatial context unattainable by the CNN alone.

2. **How does the application of the Explainable AI method (Grad-CAM) enhance the interpretability of the classification results?**

The Grad-CAM method was successfully applied to generate heatmap visualizations demonstrating that the model focuses its attention on clinically relevant biological regions namely the Plasmodium parasite inclusion within the erythrocyte rather than on noise or staining artifacts. In True Positive cases, the heatmap exhibited red-yellow activation concentrated within the cell body containing the parasite, whereas in True Negative cases, the activation distribution appeared diffusely dispersed without a single focal point. This enhances the transparency and clinical trustworthiness of the model's prediction results, such that classification decisions can be assessed not only through quantitative metrics but also validated for their correspondence to relevant biological structures.

3. How does the performance of the Hybrid CNN–Swin Transformer model compare to that of the standalone models based on multi-metric clinical evaluation?

A comprehensive evaluation was conducted across the three models through 18 testing scenarios, varying split ratio, batch size, learning rate, and weight decay. The CNN achieved its best performance under a configuration comprising a 70:15:15 split ratio, batch size of 8, learning rate of $1e-4$, and weight decay of $1e-2$, yielding an accuracy of 0.9618, precision of 0.9658, recall of 0.9574, F1-score of 0.9616, and ROC-AUC of 0.9899. The Swin Transformer achieved the highest overall performance under a configuration comprising an 80:10:10 split ratio, batch size of 8, learning rate of $1e-4$, and weight decay of $1e-2$, yielding an accuracy of 0.9768, precision of 0.9768, recall of 0.9768, F1-score of 0.9768, and ROC-AUC of 0.9963. The Hybrid CNN–Swin Transformer ranked second, achieving an accuracy of 0.9731, the highest precision of 0.9808, recall of 0.9652, F1-score of 0.9729, and ROC-AUC of 0.9944 under a configuration comprising an 80:10:10 split ratio, batch size of 16, learning rate of $5e-5$, and weight decay of $1e-2$. Overall, the transformer-based approach proved more effective than the conventional CNN, although the CNN continued to demonstrate favorable performance, with all metrics exceeding 0.95.

4. **What are the results of implementing the model in the form of a web-based system prototype?**

The Hybrid CNN–Swin Transformer model, together with the Grad-CAM visualization, was successfully integrated into a web-based system prototype using the Streamlit framework. This system enables users to upload microscopic blood cell images, obtain class prediction results (infected or uninfected) along with their probability values, and view the Grad-CAM visualization interactively through a simple and informative web interface, thereby functioning as a rapid, accurate, and user-friendly malaria diagnostic support tool, particularly for healthcare facilities with limited access to trained microscopy personnel.

5.1 Recommendations

Although the developed model has demonstrated excellent performance, several aspects remain that could be further developed in future research, including:

1. **Expansion to Multiclass Classification**

This study focused exclusively on binary classification (infected vs. uninfected). Future research could develop a model for multiclass classification capable of distinguishing Plasmodium parasite species (such as *P. falciparum*, *P. vivax*, and *P. malariae*) as well as developmental stages (trophozoite, schizont, gametocyte), thereby enabling the detection system to provide more comprehensive and clinically relevant information.

2. **Advanced Hyperparameter Optimization**

This study employed the Grid Search method, which is exhaustive and computationally intensive. Future research could explore more efficient hyperparameter optimization techniques, such as Bayesian Optimization or Random Search, to identify more optimal model configurations within a shorter timeframe, thereby enabling a more efficient experimentation process.

3. **Use of Local Indonesian Datasets**

The dataset employed in this study is the international NIH Malaria dataset obtained from Kaggle. Future research is recommended to collect and utilize malaria microscopic image datasets from laboratories within Indonesia, in order to enhance the model's relevance and generalization

capability with respect to variations in preparation techniques, staining intensity, and slide quality commonly encountered in local healthcare facilities.

4. System Development and Integration into Healthcare Facilities

The web-based system prototype developed in this study remains experimental in nature and has not undergone clinical testing. Future research could further develop this system into a more robust application, integrated with hospital laboratory information systems, and clinically validated by medical personnel, thereby enabling its practical implementation as a malaria diagnosis support tool within healthcare facilities.