CHAPTER V
CLOSING

5.1 Conclusion

Based on all stages of the research that have been conducted, including dataset analysis, deduplication, model training, evaluation, and website interface implementation, the following conclusions can be drawn.

1. This study successfully designed an ordinal classification model for liver fibrosis based on ConvNeXt V2 Tiny with the CORAL approach to predict F0–F4 fibrosis stages from B-mode ultrasound images. The model was designed with ConvNeXt V2 Tiny as the feature extraction backbone, a CORAL ordinal head as the main output for F0–F4 staging, and an advanced fibrosis head as an additional output to distinguish between non-advanced and advanced fibrosis. Before model training, the dataset also underwent MD5-based auditing and deduplication, resulting in cleaner data and reducing the risk of identical images being distributed across different subsets.
2. The CORAL ordinal approach proved to be more suitable than the nominal softmax classification baseline for the F0-F4 fibrosis staging task. In the comparison between Scenario 1 and Scenario 2, CORAL increased QWK from 0.7991 to 0.8133, reduced MAE from 0.4026 to 0.3680, and improved adjacent accuracy from 0.8788 to 0.9091. These results show that CORAL not only improved overall performance, but also helped keep prediction errors closer to the true classes, in accordance with the ordinal nature of fibrosis stages.
3. The main model using a 5-fold ensemble and 12-way TTA showed strong internal performance on the holdout set. For the F0–F4 staging task, the model achieved a QWK of 0.8520, macro-F1 of 0.7584, MAE of 0.3333, and adjacent accuracy of 0.9221. For the main binary focus of advanced fibrosis (F≥3), the model achieved an AUC of 0.9258 with a sensitivity of 0.9565. These results indicate that the model was able to perform ordinal fibrosis staging while also providing advanced and non-advanced fibrosis separation as an additional interpretive output.

4. The implementation of the FibrosisRisk website successfully demonstrated how the model outputs could be presented through a simple interface prototype. The website displays the input image, F0-F4 fibrosis class prediction, staging map, class probability distribution, and advanced/non-advanced fibrosis status. In addition, the website also provides initial validation when the uploaded image is not recognized as a liver ultrasound image, preventing inference from being directly performed on inputs outside the research domain. Therefore, this implementation shows that the model outputs can be presented in a more structured and understandable manner, while still being positioned as a research prototype rather than a standalone clinical diagnostic system.

5.2 Recommendations

Based on the results and limitations of this study, several recommendations for future research are as follows.

1. Future studies should conduct external validation using datasets from different sources, different ultrasound devices, and more varied acquisition conditions. External validation is important to determine whether the model can truly generalize beyond the dataset used in this study. Without this stage, the model results should still be interpreted as internal performance, not as evidence of clinical readiness.
2. Future studies may consider using datasets that include patient identity or examination identity information. Such information is important to ensure that data splitting can be performed more carefully at the patient or examination level.