

ABSTRAK

Perkembangan teknologi *deep learning* telah membuka peluang untuk mengotomatisasi penilaian kesegaran ikan yang selama ini dilakukan secara manual melalui inspeksi visual dan bersifat subjektif. Penelitian ini melakukan studi komparasi antara dua konfigurasi *classifier head* pada arsitektur Swin Transformer, yaitu Swin Transformer Baseline dengan *linear classifier* dan Swin Transformer dengan Multi-Layer Perceptron (MLP), pada domain klasifikasi kesegaran ikan berbasis citra multi-perspektif. Metodologi penelitian mengikuti kerangka CRISP-DM dengan strategi *fine-tuning* dua tahap pada *backbone* Swin-Tiny dan evaluasi menggunakan *5-Fold Cross Validation*. Dataset yang digunakan merupakan gabungan tiga sumber publik yang mencakup 6.930 citra dari 12 spesies ikan dengan dua perspektif utama yaitu *close-up* mata dan tubuh utuh, berlabel *highly fresh*, *fresh*, dan *not fresh*. Hasil evaluasi menunjukkan bahwa kedua model menghasilkan performa yang setara dengan *accuracy* sebesar 85,10% dan ROC-AUC di atas 0,96. Model MLP menunjukkan stabilitas pelatihan yang lebih konsisten namun tanpa peningkatan kemampuan generalisasi yang signifikan, mengindikasikan bahwa representasi fitur *backbone* Swin Transformer yang telah di-*fine-tune* sudah memadai tanpa memerlukan kompleksitas tambahan pada *classifier head*. Visualisasi Grad-CAM mengkonfirmasi bahwa model berfokus pada area mata dan permukaan tubuh ikan sesuai indikator organoleptik dalam SNI 2729:2021.

Kata Kunci: *classifier head*, komparasi model, kesegaran ikan, multi-perspektif, Swin Transformer.



ABSTRACT

The advancement of deep learning technology has opened opportunities to automate fish freshness assessment, which has traditionally relied on subjective manual visual inspection. This study conducts a comparative analysis of two classifier head configurations on the Swin Transformer architecture, namely the Swin Transformer Baseline with a linear classifier and the Swin Transformer with a Multi-Layer Perceptron (MLP), applied to multi-perspective image-based fish freshness classification. The research methodology follows the CRISP-DM framework with a two-stage fine-tuning strategy on the Swin-Tiny backbone and 5-Fold Cross Validation for evaluation. The dataset combines three public sources comprising 6,930 images from 12 fish species with two main perspectives — close-up eye and whole body images — labeled as highly fresh, fresh, and not fresh. Evaluation results indicate that both models achieved comparable performance with an accuracy of 85.10% and ROC-AUC exceeding 0.96. The MLP model demonstrated more consistent training stability but without significant improvement in generalization capability, suggesting that the feature representations produced by the fine-tuned Swin Transformer backbone are sufficient without requiring additional complexity in the classifier head. Grad-CAM visualization confirms that the model focuses on the eye region and body surface consistent with the organoleptic indicators in SNI 2729:2021.

Keywords: *classifier head, fish freshness, model comparison, multi-perspective, Swin Transformer.*

