

ABSTRAK

Penelitian ini menganalisis dan membandingkan kinerja dua arsitektur model *deep learning* dalam mengklasifikasikan tingkat kesegaran ikan berdasarkan citra mata ikan: Model Standar berbasis MobileNetV2 dan Model *Hybrid* yang mengintegrasikan MobileNetV2 dengan *Contrastive Language-Image Pre-training* (CLIP) melalui mekanisme *Feature-Level Fusion*. Kesegaran ikan merupakan faktor penting yang menentukan kualitas, keamanan pangan, dan nilai ekonomis produk perikanan. Metode penilaian konvensional berbasis pengamatan organoleptik bersifat subjektif dan tidak konsisten, sehingga diperlukan pendekatan otomatis yang lebih objektif.

Dataset yang digunakan adalah *Freshness of Fish Eyes* (FFE) yang terdiri dari 4.392 citra mata delapan spesies ikan, diklasifikasikan ke dalam tiga kategori kesegaran: *Highly Fresh*, *Fresh*, dan *Not Fresh*. Pra-pemrosesan mencakup *resize* citra ke resolusi 224×224 piksel, normalisasi piksel menggunakan parameter ImageNet, serta augmentasi data *on-the-fly*. Evaluasi dilakukan menggunakan skema *10-Fold Stratified Cross Validation* dengan metrik *Accuracy*, *Precision*, *Recall*, dan *F1-Score*.

Hasil pengujian menunjukkan bahwa Model Standar memperoleh *mean accuracy* sebesar 78,18% dengan standar deviasi 2,55%, sedangkan Model *Hybrid* memperoleh *mean accuracy* sebesar 78,47% dengan standar deviasi 1,91%. Selisih akurasi keseluruhan sebesar 0,29 poin persentase tidak signifikan secara statistik, namun perbandingan metrik per kelas mengungkap perbedaan yang lebih substantif. Model *Hybrid* memperoleh *macro average F1-Score* sebesar 74,82%, lebih tinggi dibandingkan Model Standar sebesar 73,35%. Peningkatan paling signifikan terjadi pada *Precision* kelas *Not Fresh*, di mana Model *Hybrid* mencapai 78,81% dibandingkan Model Standar yang hanya 68,49%. Visualisasi *Grad-CAM* menunjukkan bahwa Model *Hybrid* menghasilkan aktivasi yang lebih terfokus pada area kornea dan pupil mata ikan dibandingkan Model Standar.

Penelitian ini menyimpulkan bahwa pemilihan model optimal bergantung pada konteks penggunaan: Model Standar lebih unggul dalam efisiensi dan konsistensi, sedangkan Model *Hybrid* lebih sesuai untuk aplikasi keamanan pangan yang mengutamakan keseimbangan performa antar kelas.

Kata Kunci: kesegaran ikan, klasifikasi citra, *deep learning*, MobileNetV2, CLIP

ABSTRACT

This study analyzes and compares the performance of two deep learning model architectures in classifying fish freshness levels based on fish eye images: a Standard Model based on MobileNetV2 and a Hybrid Model that integrates MobileNetV2 with Contrastive Language-Image Pre-training (CLIP) via a feature-level fusion mechanism. Fish freshness is a critical factor determining the quality, food safety, and economic value of fishery products. Conventional assessment methods based on organoleptic evaluation are subjective and inconsistent, necessitating a more objective, automated approach.

The dataset used is the Freshness of Fish Eyes (FFE) dataset, which consists of 4,392 eye images from eight fish species, classified into three freshness categories: Highly Fresh, Fresh, and Not Fresh. Preprocessing includes resizing the images to a resolution of 224×224 pixels, pixel normalization using ImageNet parameters, and on-the-fly data augmentation. Evaluation was conducted using a 10-Fold Stratified Cross-Validation scheme with Accuracy, Precision, Recall, and F1-Score as evaluation metrics.

The experimental results show that the Standard Model achieved a mean accuracy of 78.18% with a standard deviation of 2.55%, while the Hybrid Model achieved a mean accuracy of 78.47% with a standard deviation of 1.91%. Although the overall accuracy difference of 0.29 percentage points is not statistically significant, the per-class metric comparison reveals more substantive differences. The Hybrid Model achieved a macro average F1-Score of 74.82%, which is higher than the Standard Model's score of 73.35%. The most significant improvement occurred in the Precision of the Not Fresh class, where the Hybrid Model reached 78.81% compared to the Standard Model's 68.49%. Grad-CAM visualization indicates that the Hybrid Model generates activations that are more focused on the cornea and pupil areas of the fish eye compared to the Standard Model.

This study concludes that the choice of the optimal model depends on the specific use case: the Standard Model excels in efficiency and consistency, whereas the Hybrid Model is better suited for food safety applications that prioritize balanced performance across classes.

Keywords: *Fish Freshness, Image Classification, Deep Learning, MobileNetV2, CLIP.*