ABSTRAK

Pemilihan arsitektur *deep learning* pada klasifikasi citra perlu mempertimbangkan performa, efisiensi komputasi, dan kelayakan implementasi. Penelitian ini bertujuan membandingkan kinerja ResNet-34 dan ResNet-50 menggunakan pendekatan *transfer learning*, dengan klasifikasi citra makanan lokal Indonesia sebagai studi kasus. Studi kasus ini dipilih karena memiliki variasi visual yang tinggi, kemiripan antar kelas, serta perbedaan pencahayaan, sudut pengambilan gambar, dan bentuk penyajian. Penelitian ini menggunakan metodologi *Cross-Industry Standard Process for Data Mining* (CRISP-DM), mulai dari business understanding hingga deployment. Dataset terdiri atas 10 kelas citra makanan Indonesia dari beberapa sumber publik. Data diproses melalui *cleaning*, standardisasi label, augmentasi untuk *balancing*, serta *grouped stratified split* untuk mencegah *data leakage*. Empat model diuji, yaitu ResNet-34 *baseline*, ResNet-34 *fine-tune*, ResNet-50 *baseline*, dan ResNet-50 *fine-tune*. Hasil penelitian menunjukkan bahwa ResNet-50 *fine-tune* menjadi model terbaik pada *clean benchmark* dengan *accuracy* 0,8300, *macro* F1-*score* 0,8308, dan *weighted* F1-*score* 0,8296. Model ini juga menunjukkan performa yang kuat pada evaluasi *robustness* dan skenario *long-tail* berat. Pada tahap *deployment*, *post-training quantization* memberikan efisiensi terbaik dari sisi ukuran model dan latensi CPU, tetapi menurunkan performa klasifikasi. Sementara itu, *structured pruning layer4* 10% menjadi konfigurasi pruning paling stabil, meskipun belum mengurangi ukuran checkpoint karena implementasinya masih berbasis *mask*. Dengan demikian, ResNet-50 fine-tune menjadi arsitektur terbaik dari sisi performa, sedangkan optimasi deployment perlu disesuaikan dengan prioritas antara akurasi, ukuran model, dan latensi inferensi.

Kata Kunci: *Deep Learning*, ResNet, *Transfer Learning*, Fine-Tuning, *Structured Pruning*.

ABSTRACT

The selection of a deep learning architecture for image classification should consider classification performance, computational efficiency, and implementation feasibility. This study compares the performance of ResNet-34 and ResNet-50 using a transfer learning approach, with Indonesian local food image classification as a case study. This case study was selected because it contains high visual variation, similarities between classes, and differences in lighting, image angle, and food presentation. This study applies the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology, from business understanding to deployment. The dataset consists of 10 Indonesian food image classes collected from several public sources. The data were processed through cleaning, label standardization, augmentation for balancing, and grouped stratified split to prevent data leakage. Four models were evaluated, namely ResNet-34 baseline, ResNet-34 fine-tune, ResNet-50 baseline, and ResNet-50 fine-tune. The results show that ResNet-50 fine-tune achieved the best performance on the clean benchmark, with an accuracy of 0.8300, a macro F1-score of 0.8308, and a weighted F1-score of 0.8296. This model also showed strong performance in robustness evaluation and the severe long-tail scenario. In the deployment stage, post-training quantization provided the best efficiency in terms of model size and CPU latency, but reduced classification performance. Meanwhile, structured pruning on layer4 at 10% became the most stable pruning configuration, although it did not reduce the checkpoint size because the implementation was still mask-based. Therefore, ResNet-50 fine-tune was the best architecture in terms of performance, while deployment optimization should be selected based on the priority between accuracy, model size, and inference latency.

Keywords: Deep Learning, ResNet, Transfer Learning, Fine-Tuning, Structured Pruning**.**