



ABSTRAK

Kanker paru merupakan salah satu penyebab utama kematian di dunia dengan tingkat mortalitas yang tinggi. Upaya deteksi dini berbasis citra medis, khususnya *CT-scan*, menjadi sangat penting untuk meningkatkan angka harapan hidup pasien. Namun, proses klasifikasi citra *CT-scan* menghadapi tantangan besar berupa ketidakseimbangan data (*class imbalance*) yang dapat menyebabkan bias model terhadap kelas mayoritas. Dalam skenario *baseline*, ketidakseimbangan ini menyebabkan penurunan akurasi klasifikasi pada kelas minoritas (*Normal* dan *Benign*) sebesar 3–5%, serta menghasilkan bias prediksi hingga 20% terhadap kelas mayoritas (*Malignant*), yang dapat berdampak serius dalam konteks diagnosis klinis.

Penelitian ini menerapkan pendekatan *deep learning* dengan membandingkan dua arsitektur *transformer*, yaitu *Vision Transformer* (ViT) dan *Swin Transformer*, untuk klasifikasi citra *CT-scan* paru-paru. Dataset yang digunakan adalah IQ-OTH/NCCD dengan total 1.097 citra dari tiga kelas (*Normal*, *Benign*, *Malignant*). Teknik *balancing* diterapkan secara selektif pada data train menggunakan augmentasi citra berbasis *ImageDataGenerator* untuk mengatasi ketidakseimbangan jumlah data antar kelas. Evaluasi dilakukan dengan metrik akurasi, *Precision*, *Recall*, *F1-score*, serta uji signifikansi *Wilcoxon* untuk menilai perbedaan performa antara skenario *data balanced* dan *imbalanced*.

Hasil penelitian menunjukkan bahwa teknik *balancing* secara signifikan meningkatkan performa klasifikasi, terutama pada model ViT yang mencatatkan peningkatan *F1-score* dari 0,95 menjadi 0,98, tanpa menghasilkan *false positive* maupun *false negative* pada kelas *Malignant*. Model *Swin Transformer* juga menunjukkan peningkatan dari *F1-score* 0,96 menjadi 0,97. Uji *Wilcoxon* membuktikan bahwa peningkatan tersebut signifikan secara statistik ($p < 0,05$). Temuan ini menegaskan bahwa *balancing* berbasis augmentasi citra efektif dalam meningkatkan keandalan sistem klasifikasi citra medis berbasis *deep learning*, dan model ViT dengan data seimbang layak dipertimbangkan sebagai solusi diagnosis berbantuan AI untuk kanker paru.

Kata kunci— *Lung Cancer*, *Balancing Data*, *Vision Transformer* dan *Swim Transformer*.



ABSTRACT

Lung cancer is one of the leading causes of death worldwide, with a high mortality rate. Early detection efforts based on medical imaging, particularly CT scans, are crucial for improving patient survival rates. However, automatically classifying CT scan images faces a significant challenge in data imbalance (class imbalance), which can cause model bias toward the majority class. In the baseline scenario, this imbalance results in a 3–5% decrease in classification *Accuracy* for the minority class (*Normal* and *Benign*). It generates a prediction bias of up to 20% toward the majority class (*Malignant*), which can have serious implications in clinical diagnosis.

This study applies a deep learning approach by comparing two transformer architectures, namely Vision Transformer (ViT) and Swin Transformer, to classify lung CT-scan images. The dataset used is IQ-OTH/NCCD with 1,097 images from three classes (*Normal*, *Benign*, *Malignant*). Balancing techniques were selectively applied to the training data using image augmentation based on ImageDataGenerator to address the imbalance in the number of data points across classes. Evaluation was conducted using *Accuracy*, *Precision*, *Recall*, F1-score metrics, and the Wilcoxon significance test to assess performance differences between balanced and unbalanced data scenarios.

The study results indicate that the balancing technique significantly improves classification performance, particularly in the ViT model, which increased the F1-score from 0.95 to 0.98, without producing false positives or false negatives in the *Malignant* class. The Swin Transformer model also improved from an F1-score of 0.96 to 0.97. The Wilcoxon test confirmed this improvement was statistically significant ($p < 0.05$). These findings confirm that image augmentation-based balancing effectively enhances the reliability of deep learning-based medical image classification systems, and the ViT model with balanced data is worthy of consideration as an AI-assisted diagnostic solution for lung cancer.

Keywords— Lung Cancer, Balancing Data, Vision Transformer and Swin Transformer.