



## ABSTRACT

According to the Global Burden of *Cancer* Study (Globocan) from the World Health Organization (WHO), lung cancer in Indonesia is ranked third with 34,783 cases (8.8% of total cases) in 2020. The mortality rate from lung cancer can be reduced through early detection and diagnosis. Convolutional neural networks (CNNs) have shown significant success in lung cancer detection and classification. The limitation are their high computational requirements and parameter storage costs which limit its deployment in healthcare facilities with limited infrastructure. Recent advancements to overcome challenges involve compression techniques, such as pruning the weights of filters in convolutional layers, aiming to maintain performance while reducing model complexity.

Two CNN architectures, VGG16 and ResNet50 were employed using transfer learning. The preprocessing steps included Contrast Limited Adaptive Histogram Equalization (CLAHE), resizing, and normalization. The pruning technique utilized the L2-norm to calculate filter importance score based on feature maps and prune the filters which have small importance score. The pruned models were fine-tuned using the original Chest CT-Scan dataset and an augmented dataset to recover the accuracy loss.

The classification results show that VGG16 and ResNet50 before *pruning* achieve an accuracy of 87.62%, and 92.17%. Fine tuning the model using original dataset results in significant accuracy loss. We perform fine tuning using 10% dataset augmentation on both VGG16 and ResNet50 show the best performance at 10% pruning ratio. For VGG16 and ResNet50 pruned model, we achieve accuracies of 89.04% and 92.49% respectively, outperforming earlier methods while reducing training time and model size. The experimental results reveal that the pre-trained model's performance combined with pruning enhances performance and efficiency compared to pre-trained models.

Keywords: CNN, classification, *pruning*



## INTISARI

Menurut *Global Burden of Cancer Study* (Globocan) dari *World Health Organization* (WHO), kanker paru di Indonesia menempati peringkat ketiga dengan 34.783 kasus (8,8% dari total kasus) pada tahun 2020. Tingkat kematian akibat kanker paru dapat dikurangi melalui deteksi dan diagnosis dini. CNN telah menunjukkan keberhasilan yang signifikan dalam mendeteksi dan mengklasifikasikan kanker paru. Namun, keterbatasannya terletak pada kebutuhan komputasi yang tinggi dan biaya penyimpanan parameter, yang membatasi penerapannya di fasilitas kesehatan dengan infrastruktur terbatas. Kemajuan terbaru untuk mengatasi tantangan ini melibatkan teknik kompresi, seperti *pruning* bobot *filter* pada lapisan konvolusional, dengan tujuan mempertahankan performa sambil mengurangi kompleksitas model.

Dua arsitektur CNN yaitu VGG16 dan ResNet50 digunakan dengan pendekatan transfer learning. Tahapan *pre-processing* meliputi CLAHE, pengubahan ukuran citra, dan normalisasi citra. Teknik *pruning* menggunakan metode *L2-norm* untuk menghitung skor kepentingan *filter* berdasarkan *feature maps* dan memangkas *filter* dengan skor kepentingan yang kecil. Model yang telah dipangkas kemudian dilatih ulang (*fine-tuned*) menggunakan dataset asli CT-Scan dada dan dataset yang telah diaugmentasi untuk memulihkan akurasi yang hilang.

Hasil klasifikasi menunjukkan bahwa VGG16 dan ResNet50 sebelum dilakukan *pruning* mencapai akurasi sebesar 87,62% dan 92,17%. *Fine-tuned* model menggunakan dataset asli menghasilkan penurunan akurasi yang signifikan. Namun, *Fine-tuned* model dengan augmentasi dataset sebesar 10% pada VGG16 dan ResNet50 menghasilkan performa terbaik pada rasio *pruning* 10%. Model VGG16 dan ResNet50 yang telah dipangkas mencapai akurasi sebesar 89,04% dan 92,49%, melampaui metode sebelumnya sekaligus mengurangi waktu pelatihan dan ukuran model. Hasil eksperimen menunjukkan bahwa kinerja *pre-trained* model dengan *pruning*, dapat meningkatkan performa dan efisiensi dibandingkan model yang dilatih sebelumnya.

Kata kunci – CNN, Klasifikasi, *Pruning*