

## INTISARI

Tingginya mortalitas kanker paru seringkali disebabkan oleh keterlambatan diagnosis. Meskipun model *deep learning* seperti arsitektur U-Net menawarkan potensi besar untuk deteksi dini, sifatnya sebagai "kotak hitam" menjadi penghalang utama adopsi klinis karena kurangnya transparansi dalam proses pengambilan keputusan. Penelitian ini mengimplementasikan arsitektur *Attention U-Net*, yang unggul dalam segmentasi objek kecil dan mempertahankan detail, serta menggunakan metode *Explainable AI (XAI)* untuk mengatasi masalah kepercayaan.

Di antara berbagai metode XAI seperti SHAP, LIME, dan Grad-CAM, *Integrated Gradients (IG)* dipilih karena landasan teoritisnya yang kuat, kemampuannya menghasilkan atribusi tingkat piksel yang detail, dan efisiensi komputasi yang seimbang untuk model kompleks. Model *Attention U-Net* yang dikembangkan mencapai kinerja yang cukup baik dengan *Dice Coefficient* sebesar 0,73, *IoU* sebesar 0,69, *Precision* sebesar 0,73, dan *Recall* sebesar 0,75.

Analisis menggunakan IG menunjukkan bahwa metode ini mampu memvalidasi penalaran model pada kasus-kasus yang berhasil. Namun, temuan paling krusial adalah terungkapnya anomali penalaran, di mana model dengan metrik kuantitatif tinggi ternyata mengandalkannya fitur yang tidak relevan seperti derau, lesi, dan pembuluh berwarna terang untuk membuat prediksi. Penelitian ini menegaskan bahwa XAI merupakan komponen esensial untuk mengaudit dan memastikan keandalan model AI medis sebelum implementasi klinis.

**Kata Kunci:** *Explainable AI, Integrated Gradients, Attention U-Net, Kanker Paru, Deep Learning.*

## ABSTRACT

*The high mortality rate of lung cancer is often caused by delayed diagnosis. Although deep learning models like the U-Net architecture offer great potential for early detection, their "black box" nature serves as a primary barrier to clinical adoption due to a lack of transparency in the decision-making process. This research implements an Attention U-Net architecture, which excels in segmenting small objects and preserving details, and utilizes the Explainable AI (XAI) method to address the issue of trust.*

*Among various XAI methods such as SHAP, LIME, and Grad-CAM, Integrated Gradients (IG) was chosen for its strong theoretical foundation, its ability to generate detailed pixel-level attributions, and its balanced computational efficiency for complex models. The developed Attention U-Net model achieved competent performance with a Dice Coefficient of 0.73, an IoU of 0.69, a Precision of 0.73, and a Recall of 0.75.*

*Analysis using IG demonstrated that the method is capable of validating the model's reasoning in successful cases. However, the most crucial finding was the uncovering of a reasoning anomaly, where a model with high quantitative metrics was found to be relying on irrelevant features such as lesion, noises, bright contours to make its prediction. This research confirms that XAI is an essential component for auditing and ensuring the reliability of medical AI models prior to clinical implementation.*

**Keywords:** *Explainable AI, Integrated Gradients, Attention U-Net, Lung Cancer, Deep Learning.*