

Prediksi Kanker Payudara Menggunakan Convolutional Neural Networks (CNN) dengan Analisis Citra Medis di Indonesia = Breast Cancer Prediction Using Convolutional Neural Networks (CNN) with Medical Image Analysis in Indonesia

Amalia Chusna Mohtar, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920571687&lokasi=lokal>

Abstrak

Penelitian ini bertujuan mengembangkan model prediksi kanker payudara dengan Convolutional Neural Networks (CNN) dan analisis citra medis untuk mendeteksi lesi yang bersifat jinak maupun ganas. Data yang digunakan mencakup citra mamogram dari CBIS-DDSM serta data primer dari rumah sakit. Tahap pertama melibatkan penggunaan algoritma YOLO untuk segmentasi breast tissue, guna menghapus noise latar belakang dan memastikan fokus pada area diagnostik yang relevan. Selanjutnya, diterapkan teknik Contrast Limited Adaptive Histogram Equalization (CLAHE) untuk meningkatkan kontras dan menonjolkan detail struktural. Ekstraksi fitur dari full mammogram dan Region of Interest (ROI) mask kemudian digabungkan dalam arsitektur CNN multi-input untuk klasifikasi biner kanker payudara. Hasil evaluasi menunjukkan bahwa meskipun model mencapai akurasi training yang tinggi (antara 95% hingga 98%), akurasi pada data validasi baru berkisar pada 56% hingga 63%, dengan nilai F1-score masing-masing 0,69 untuk kasus benign dan 0,55 untuk kasus malignant, serta AUC-ROC sebesar 0,57. Temuan ini mengindikasikan adanya tantangan overfitting dan kurang optimalnya pemisahan antara kelas benign dan malignant. Penelitian ini memberikan kontribusi penting sebagai langkah awal dalam pengembangan sistem diagnosis dini yang dapat mendukung proses pengambilan keputusan klinis.

.....This study aims to develop a breast cancer prediction model using Convolutional Neural Networks (CNN) and medical image analysis to detect both benign and malignant lesions. The data utilized includes mammogram images from CBIS-DDSM and primary data collected from a hospital. The initial stage employs the YOLO algorithm to segment breast tissue, removing background noise and ensuring focus on diagnostically relevant areas. Subsequently, Contrast Limited Adaptive Histogram Equalization (CLAHE) is applied to enhance image contrast and emphasize structural details. Feature extraction from full mammograms and Region of Interest (ROI) masks is then combined in a multi-input CNN architecture for binary breast cancer classification. Evaluation results indicate that although the model achieves high training accuracy (ranging from 95% to 98%), validation accuracy remains between 56% and 63%. The F1-scores are 0.69 for benign cases and 0.55 for malignant cases, with an AUC-ROC of 0.57. These findings highlight challenges related to overfitting and suboptimal class separability between benign and malignant categories. This research serves as an important initial step toward developing an early diagnostic support system to aid clinical decision-making.