

ABSTRACT

CONVOLUTIONAL NEURAL NETWORK WITH ADAPTIVE BOOSTING ALGORITHM FOR BREAST CANCER MAMMOGRAPHY DETECTION

Adam Rezky Tri Anugerah

20/457763/PA/19801

Breast cancer is a cancer that develops from breast tissue. Early symptoms of breast cancer include an existence of lump in the breast, a change in breast shape, or dimpling on the skin. Tremendous encouragement of breast cancer awareness and research funding has been beneficial in the advancement of diagnosis and medical treatment. Therefore, breast cancer survival rates have risen, and the number of fatalities attributed with the disease has continuously declined, attributable primarily to elements such as earlier identification, which has led to an increase in demand for diagnosis automation. This research explores the potential of ensemble learning, with a focus on the Adaptive Boosting algorithm, to enhance the performance of Convolutional Neural Networks (CNN) in image classification tasks. The architectures in focus were VGG-16, ResNet50, and Inception V4, three prevalent CNN models with proven efficiency in image recognition tasks. In individual tests, coupling these CNN models with Adaptive Boosting led to notable performance improvements. The study further pushed the envelope by constructing an ensemble model that combined the strengths of all three CNN models. This novel ensemble, further augmented with the AdaBoost algorithm, demonstrated impressive performance across various datasets. With precision and recall scores exceeding 0.94, an F1-Score above 0.96, and an overall accuracy oscillating between 0.95 to 0.99, the ensemble model remarkably outperformed the standalone CNN models. The significant performance boost can be attributed to the richer feature space generated by the ensemble of multiple CNN models and the iterative refinement of predictions provided by the Adaptive Boosting algorithm. Despite the complexity and increased computational demand of the ensemble model, the results provide a compelling justification for its use. The study paves the way for future explorations in machine learning and artificial intelligence, particularly in leveraging ensemble learning for complex classification tasks. Further research could delve into the optimisation of such ensemble models, exploration of other ensemble strategies, or testing the models on diverse datasets and tasks beyond image classification.

Index Terms: Ensemble Learning, AdaBoost, CNN, Breast Cancer

INTISARI

CONVOLUTIONAL NEURAL NETWORK DENGAN ALGORITMA ADAPTIVE BOOSTING UNTUK DETEKSI MAMMOGRFI KANKER PAYUDARA

Adam Rezky Tri Anugerah

20/457763/PA/19801

Kanker payudara adalah kanker yang berkembang dari jaringan payudara. Gejala awal kanker payudara mencakup adanya benjolan di payudara, perubahan bentuk payudara, atau kerutan pada kulit. Dorongan besar terhadap kesadaran kanker payudara dan pendanaan penelitian telah sangat bermanfaat dalam kemajuan diagnosis dan pengobatan medis. Oleh karena itu, tingkat bertahan hidup kanker payudara telah meningkat, dan jumlah kematian yang diatribusikan pada penyakit ini terus menurun, yang sebagian besar disebabkan oleh faktor-faktor seperti identifikasi yang lebih awal, yang telah menyebabkan peningkatan permintaan untuk otomatisasi diagnosis. Penelitian ini mengeksplorasi potensi metode *ensemble*, dengan fokus pada algoritma *Adaptive Boosting*, untuk meningkatkan performa *Convolutional Neural Network* (CNN) dalam proses klasifikasi gambar. Arsitektur yang menjadi fokus pembelajaran adalah *VGG-16*, *ResNet50*, dan *Inception V4*, tiga CNN yang populer dengan efisiensi yang terbukti dalam tugas pengenalan gambar. Dalam tes individual, penggabungan CNN ini dengan *Adaptive Boosting* menghasilkan peningkatan performa yang signifikan. Studi ini lebih lanjut mendorong batas dengan membuat model *ensemble* yang menggabungkan kekuatan ketiga CNN tersebut. *Ensemble* baru ini, yang lebih ditingkatkan dengan algoritma *Adaptive Boosting*, menunjukkan kinerja yang mengesankan di berbagai dataset. Dengan skor presisi dan recall yang melebihi 0,94, skor F1 di atas 0,96, dan akurasi keseluruhan berkisar antara 0,95 hingga 0,99, model *ensemble* secara luar biasa mengungguli CNN individu. Peningkatan kinerja yang signifikan dapat diatribusikan kepada ruang fitur yang lebih kaya yang dihasilkan oleh *ensemble* beberapa CNN dan penyempurnaan iteratif prediksi yang disediakan oleh algoritma *Adaptive Boosting*. Meskipun model *ensemble* membutuhkan tuntutan komputasi yang lebih besar dan kompleksitas, hasilnya memberikan pembenaran yang meyakinkan untuk penggunaannya. Studi ini membuka jalan untuk eksplorasi masa depan dalam pembelajaran mesin dan kecerdasan buatan, khususnya dalam memanfaatkan pembelajaran *ensemble* untuk proses klasifikasi yang kompleks. Penelitian lebih lanjut dapat menyelidiki optimasi model *ensemble* seperti eksplorasi strategi *ensemble* lainnya, atau menguji model pada dataset dan tugas yang beragam di luar klasifikasi gambar.

Kata Kunci: *Ensemble Learning*, *AdaBoost*, CNN, Kanker Payudara