

Komparasi Model Convolutional Neural Network-Random Forest dan Convolutional Neural Network-XGBoost untuk mendeteksi tuberkulosis paru berdasarkan data radiografi toraks = Comparison of Convolutional Neural Network-Random Forest and Convolutional Neural Network-XGBoost Models for detecting pulmonary tuberculosis based on thorax radiography data

Afifah Rofi Laeli, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920555176&lokasi=lokal>

Abstrak

Tuberkulosis (TB) merupakan suatu penyakit menular yang sebagian besar menyerang paru-paru manusia. Penularan penyakit ini terjadi ketika pasien tuberkulosis paru mengeluarkan percikan dahak yang mengandung kuman tuberkulosis ke udara. Penularannya yang mudah menjadikan tuberkulosis sebagai masalah kesehatan masyarakat, baik di Indonesia maupun internasional. Deteksi dini tuberkulosis paru dapat mencegah penularan serta menyembuhkan pasien. Namun, adanya pandemi COVID-19 saat ini dapat menurunkan angka kasus tuberkulosis yang berhasil terdeteksi. Hal ini menunjukkan perlu adanya kemajuan dalam metode pendeteksian penyakit tuberkulosis paru. Kini, perkembangan teknologi dapat dimanfaatkan untuk membantu bidang kesehatan, salah satunya dengan machine learning. Machine learning dapat digunakan untuk mendeteksi adanya suatu penyakit berdasarkan data citra. Dalam penelitian ini, model machine learning, Convolutional Neural Network–Random Forest (CNN– Random Forest) dan Convolutional Neural Network–XGBoost (CNN–XGBoost), diimplementasikan untuk mendeteksi tuberkulosis paru berdasarkan citra radiografi toraks. Selanjutnya, kedua model tersebut dievaluasi dan dibandingkan kinerjanya berdasarkan nilai akurasi dan nilai luas wilayah di bawah kurva ROC, atau biasa disebut dengan area under the curve (AUC). Data yang digunakan sebanyak 6000 yang terdiri dari 3000 citra radiografi toraks tuberkulosis paru dan 3000 citra radiografi toraks normal. Berdasarkan hasil yang diperoleh, model CNN-Random Forest dan CNN-XGBoost memberikan kinerja yang baik dan dapat diterapkan untuk mendeteksi tuberkulosis paru, dimana CNN digunakan untuk mengekstraksi fitur pada citra, kemudian hasil ekstraksi fitur tersebut menjadi input bagi pengklasifikasi Random Forest dan XGBoost. Evaluasi kinerja berdasarkan rata-rata nilai akurasi dan rata-rata nilai AUC pada model CNN-Random Forest memberikan hasil terbaik masing-masing sebesar 98.667% dan 99.933%, sementara pada model CNN-XGBoost memberikan hasil terbaik masing-masing sebesar 98.367% dan 99.866%. Kemudian berdasarkan perbandingan kinerja yang dilakukan, model CNN-Random Forest memberikan kinerja yang lebih baik dalam mendeteksi tuberkulosis paru dibandingkan dengan model CNN-XGBoost.

.....Tuberculosis (TB) is an infectious disease that in most cases attacks the human lungs. Transmission of this disease occurs when a patient with pulmonary tuberculosis expels phlegm containing tuberculosis germs into the air. Its easy transmission makes tuberculosis a public health problem, both in Indonesia and internationally. Early detection of pulmonary tuberculosis can prevent transmission and cure patients. However, the current COVID-19 pandemic can reduce the number of successfully detected tuberculosis cases. This shows the need for progress in the detection method of pulmonary tuberculosis. Now, technological developments can be used to help the health sector, one of which is machine learning. Machine learning can be used to detect the presence of a disease based on image data. In this study, machine

learning models, Convolutional Neural Network–Random Forest (CNN–Random Forest) and Convolutional Neural Network–XGBoost (CNN–XGBoost), were implemented to detect pulmonary tuberculosis based on thorax radiography images. Furthermore, the performances of the two models were evaluated and compared based on the values of accuracy and area under the ROC curve, or commonly called the area under the curve (AUC). The data used were 6000 consisting of 3000 thorax radiography images of pulmonary tuberculosis and 3000 normal thorax radiography images. Based on the results obtained, the CNN-Random Forest and CNN-XGBoost models provided good performances and can be applied to detect pulmonary tuberculosis, where CNN was used to extract features in the image, then the results of the feature extraction became input for the Random Forest and XGBoost classifiers. Performance evaluation based on the average values of accuracy and AUC in the CNN-Random Forest model gave the best results of 98.667% and 99.933%, respectively, while the CNN-XGBoost model gave the best results of 98.367% and 99.866, respectively. Then based on the performance comparison, the CNN-Random Forest model provided a better performance in detecting pulmonary tuberculosis compared to the CNN-XGBoost model.