Analisis kinerja metode Deep Feature Extraction pada klasifikasi Covid-19 menggunakan Data Citra X-Ray = Performance analysis of Deep Feature Extraction for Covid-19 classification based on X-Ray Image

Terry Argyadiva, author

Abstrak

Corona Virus Disease atau COVID-19 merupakan sebuah wabah yang ditemukan pada akhir tahun 2019 di provinsi Wuhan, China, yang kemudian menyebar ke seluruh dunia. Reverse Transcription Polymerase Chain Reaction (RT-PCR) kemudian digunakan sebagai metode diagnosa COVID-19. Namun metode RT-PCR memerlukan waktu yang lama dalam proses diagnosa sehingga American College of Radiography (ACR) merekomendasi penggunaan alat radiografi seperti Computed Tomography Scan (CT- Scan) dan X-ray sebagai metode tambahan dalam mendiagnosa COVID-19. X-ray kemudian dipilih sebagai metode tambahan dalam mendiagnosa COVID-19 karena alat yang digunakan lebih fleksibel dan sudah tersebar luas di berbagai klinik kesehatan. Pada penelitian ini, penulis menggunakan pendekatan neural network yaitu Convolutional Neural Network (CNN) untuk metode Deep Feature Extraction dan metode klasifikasi klasik dalam membuat model yang dapat mengklasifikasi paru-paru normal, terjangkit COVID-19, dan pneumonia berdasarkan data citra X-ray. Arsitektur CNN yang digunakan dalam penelitian ini adalah ResNet-50 dan metode klasifikasi klasik yang digunakan adalah Support Vector Machine (SVM), Random forest, K-Nearest Neighbor (KNN), dan Extreme Gradient Boosting (XGBoost). Dataset yang digunakan dalam penelitian ini adalah COVID-19 Image Data Collection oleh J. P. Cohen, ChestX-Ray8 Dataset oleh National Institute of Health, dan Chest X-ray Dataset oleh Mendeley Data. Selanjutnya, model dilatih menggunakan ResNet-50 untuk proses ekstraksi fitur dari fully connected layer. Kemudian, vektor fitur dari fully connected layer diklasifikasi menggunakan metode klasifikasi klasik SVM, Random forest, KNN, dan XGBoost. Berdasarkan hasil simulasi, diketahui akurasi terbaik didapatkan oleh kombinasi antara ResNet-50 dan SVM dengan 94,22%. Recall terbaik didapatkan oleh kombinasi antara ResNet-50 dan KNN dengan 94%. Precision terbaik didapatkan oleh ResNet-50 dengan 94,36%. Running time terbaik didapatkan oleh ResNet-50 dengan 0,0006 detik.

.....Corona Virus Disease or COVID-19 is an outbreak that was discovered at the end of 2019 in the province of Wuhan, China, which then spread throughout the world. Reverse Transcription Polymerase Chain Reaction (RT-PCR) was then used as a method of diagnosing COVID-19. However, the RT-PCR method requires a long time in the diagnostic process so the American College of Radiography (ACR) recommends the use of radiographic tools such as Computed Tomography Scan (CT-Scan) and X-ray as additional methods in diagnosing COVID-19. X-ray was then chosen as an additional method in diagnosing COVID-19 because the tool used is more flexible and is already widespread in various health clinics. In this study, the author uses a neural network approach, namely the Convolutional Neural Network (CNN) for the Deep Feature Extraction method and the Machine Learning approach for the classification method in making a model that can classify normal lungs, infected with COVID-19, and pneumonia based on X-ray image. The CNN architecture used in this study is ResNet-50 and the Classifier used is Support Vector Machine (SVM), Random forest, K-Nearest Neighbor (KNN), and Extreme Gradient Boosting (XGBoost). The datasets used in this study were the COVID-19 Image Data Collection by J. P. Cohen, the ChestX-Ray8

Dataset by the National Institute of Health, and the Chest X-ray Dataset by Mendeley Data. The model was then trained using the CNN method with the ResNet-50 architecture. Furthermore, the fully connected layer in the ResNet-50 architecture was replaced using the SVM, Random forest, KNN, and XGBoost classifiers. Based on the simulation results, the best accuracy is obtained by combination of ResNet-50 and SVM with 94.22%. The best recall was obtained by a combination of ResNet-50 and KNN with 94%. The best precision was obtained by ResNet-50 with 94.36%. The best running time was obtained by ResNet-50 with 0.0006 seconds.