

BAB I

PENDAHULUAN

1.1 Latar Belakang

Penyakit paru-paru merupakan kondisi paru-paru yang tidak dapat berfungsi dengan normal. Penyakit paru-paru merupakan gangguan Kesehatan paling umum yang sering dijumpai. Menurut *World Health Organisation* (WHO), penderita penyakit paru-paru di seluruh dunia mencapai 235 juta orang dan sekitar tiga juta orang meninggal karena PPOK (Penyakit Paru Obstruktif Kronis). Penyebab penyakit paru-paru sebagian besar disebabkan oleh asap rokok. Saat ini pendeteksian dini untuk mengetahui diagnosa pada penyakit paru-paru dapat dilakukan melalui rontgen di dada oleh dokter ahli paru-paru. Namun proses rontgen tidak selalu akurat dan biasanya akan dianjurkan untuk melakukan serangkaian tes lainnya. Rontgen juga membutuhkan waktu yang lama agar terdiagnosa hasilnya dan menyesuaikan jadwal dokter untuk konsultasi hasilnya. Oleh karena itu dibutuhkan proses yang lebih cepat untuk mempermudah proses identifikasi hasil dari rontgen.

Saat ini, tidak dapat disangkal bahwa kemajuan teknologi informasi berkembang dengan cepat dan cenderung memberikan kontribusi signifikan dalam mempermudah manusia menyelesaikan tugas-tugas dengan lebih efisien dan dalam waktu yang lebih singkat. Proses diagnosa terbantu dengan penggunaan AI (*Artificial Intelligence*). Salah satu teknik pembelajaran dari AI (*Artificial Intelligence*) yaitu menggunakan *Machine Learning*.

Machine Learning merupakan cabang kecerdasan buatan di mana komputer mampu mempelajari data tertentu tanpa memerlukan instruksi langsung dari manusia. Meskipun demikian, teknik *Machine Learning* masih menunjukkan keterbatasan dalam menangani permasalahan klasifikasi citra yang melibatkan data yang kompleks dan beragam. Oleh karena itu perlu diterapkan model pembelajaran menggunakan teknik *Deep Learning*. *Deep Learning* merupakan salah satu sub bidang dari *Machine Learning* dimana algoritma yang dipakai terinspirasi dari bagaimana otak manusia bekerja. Beberapa orang mungkin lebih mengenalnya dengan jaringan saraf tiruan. Pada dasarnya *Deep Learning* adalah implementasi konsep dasar *Machine Learning* yang mengadaptasikan algoritma jaringan saraf tiruan dengan banyak lapisan tersembunyi, berupa lapisan masukan dan lapisan keluaran. Dalam beberapa tahun terakhir *Deep Learning* telah menunjukkan performa yang luar biasa. Hal ini sebagian besar dipengaruhi faktor komputasi yang lebih kuat, dataset yang besar dan teknik untuk melatih jaringan yang lebih dalam (Goodfellow et al., n.d.). *Deep Learning* telah menjadi salah satu topik hangat dalam dunia *machine learning* karena kapabilitasnya yang signifikan dalam memodelkan berbagai data kompleks seperti citra dan suara. Metode *Deep Learning* yang saat ini memiliki hasil paling signifikan dalam pengenalan citra adalah *Convolutional Neural Network* (CNN) (Fukushima, 1980). *Convolutional Neural Network* adalah salah satu algoritma dari *Deep Learning* yang merupakan pengembangan dari *Multi Layer Perceptron* (MLP) yang dirancang untuk mengolah data dalam bentuk grid, salah satunya citra dua dimensi, misalnya gambar atau suara. *Convolutional Neural Network* (CNN) memiliki lapisan khusus yang

disebut lapisan konvolusi, di mana citra masukan akan diproses menggunakan filter yang telah ditentukan. Dari setiap lapisan akan dihasilkan pola yang kemudian akan lebih mudah untuk diklasifikasikan. Selanjutnya, dilakukan pengujian terhadap kombinasi Hyperparameter (epoch , mini batch size , variasi dataset) dan Optimizer (Adam, SGDM, RMSProp) untuk menemukan model optimal dengan akurasi tertinggi dan nilai loss terendah. Epoch yang digunakan menentukan berapa kali model akan dilatih pada seluruh dataset, Sedangkan Mini Batch Size menentukan berapa banyak sampel data yang digunakan dalam setiap iterasi pelatihan. Penelitian ini menggunakan dataset sebanyak 1.050 citra yang terbagi dalam 3 kelas, yaitu Pneumonia, Tuberkulosis, dan Covid-19. Masing-masing kelas terdiri dari 350 data citra.

Beberapa penelitian sebelumnya telah dilakukan dalam bidang klasifikasi penyakit paru-paru menggunakan citra medis dan teknik *Machine Learning*. Penelitian oleh Nugroho dan Puspaningrum (2019) mengeksplorasi kinerja metode CNN dalam klasifikasi pneumonia dengan variasi ukuran citra input, menunjukkan potensi CNN dalam diagnosis penyakit paru-paru dengan hasil akurasi tertinggi 93.59%. Wijaya et al. (2020) mengusulkan penggunaan metode *K-Nearest Neighbor* (KNN) dengan ekstraksi fitur GLCM untuk klasifikasi pneumonia dengan akurasi terbaik dicapai adalah 66.20% , sementara Wati et al. (2020) menggunakan *Support Vector Machine* (SVM) dengan ekstraksi fitur GLCM untuk tujuan yang sama dengan akurasi terbaik yang dicapai adalah 62,66%. Penelitian lain yang relevan adalah penelitian oleh Fadlurrahman et al. (2023), yang berfokus pada klasifikasi COVID-19 menggunakan algoritma CNN dengan arsitektur VGG-16,

menunjukkan kemampuan CNN dalam membedakan COVID-19 dari pneumonia dengan tingkat akurasi yang dicapai adalah 90%. Wuryani dan Agustiani (2021) menggunakan *Random Forest Classifier* untuk mendeteksi COVID-19 pada citra CT Scan, memperluas penerapan pembelajaran mesin dalam diagnosis penyakit paru-paru dengan akurasi tertinggi yang dicapai adalah 96.9%. Ghani et al. (2021) mengimplementasikan arsitektur *Xception* untuk klasifikasi citra radiografi COVID-19 dengan akurasi pelatihan terbaik yang dicapai adalah 92%, dan akurasi validasi 92%, sementara Karno et al. (2021) menggunakan arsitektur AlexNet untuk tujuan yang sama, menunjukkan keragaman arsitektur CNN yang dapat diterapkan dengan akurasi yang dicapai adalah 85,5%. Penelitian oleh Nurdiansyah et al. (2020) tentang klasifikasi tuberkulosis menggunakan metode *Extreme Learning Machine* (ELM) dengan diperoleh nilai akurasi tertinggi sebesar 99,33%, dan penelitian oleh Muhdina et al. (2021) yang menggunakan *Compressive Sensing* dan SVM untuk klasifikasi tuberkulosis, memberikan wawasan tentang pendekatan alternatif dalam diagnosis penyakit paru-paru dengan akurasi sistem mencapai 92,593% . Terakhir, penelitian oleh Suharni et al. (2021) tentang implementasi model CNN untuk klasifikasi penyakit TBC berbasis desktop menunjukkan potensi penerapan CNN dalam lingkungan klinis dengan menggunakan 50 *epoch* dan *learning rate* 0.01, akurasi yang dicapai adalah 98,75% .

Berdasarkan uraian pada latar belakang, pada penelitian ini penulis memfokuskan klasifikasi pada citra rontgen paru-paru dengan menggunakan metode *Convolution Neural Network* untuk mengklasifikasi kondisi paru-paru berdasarkan citra rontgen dengan mencari kombinasi *hyperparameter* dan

Optimizer terbaik. *Hyperparameter* yang diujikan adalah nilai *Epoch*, *Mini Batch Size*, dan Variasi Dataset. *Optimizer* yang diujikan adalah *ADAM*, *SGDM*, dan *RMSPROP* Hasil klasifikasi data dibagi ke dalam 3 kelas penyakit, yaitu Pneumonia, Tuberkulosis, *Coronavirus Disease-19* (Covid-19).

1.2 Rumusan Masalah

Berdasarkan dari latar belakang yang telah diuraikan, maka dirumuskan masalah pada penelitian ini yaitu :

1. Bagaimana kinerja metode *Convolutional Neural Network* (CNN) dalam mengklasifikasikan jenis penyakit paru-paru berdasarkan citra rontgen?
2. Bagaimana tingkat akurasi metode *Convolutional Neural Network* (CNN) dalam mengklasifikasikan penyakit paru-paru pneumonia, tuberkulosis, dan COVID-19 berdasarkan citra rontgen?
3. Bagaimana pengaruh variasi *hyperparameter* (*epoch*, *mini batch size*, dan variasi dataset) dan *Optimizer* (Adam, SGDM, RMSProp) terhadap kinerja model CNN dalam klasifikasi penyakit paru-paru?

1.3 Batasan Masalah

Berikut batasan masalah pada penelitian ini :

1. Data yang digunakan adalah citra rontgen pada penyakit Pneumonia, Tuberkulosis, *Corona virus Disease-19* (Covid-19) dari situs <https://www.kaggle.com/datasets/omkarmanohardalvi/lungsdiseasedataset-4-types>.

2. Proses klasifikasi menggunakan metode *Convolution Neural Network*.
3. Simulasi program dibuat dengan menggunakan *software* matlab.

1.4 Tujuan Penelitian

Dari masalah yang telah dirumuskan, maka tujuan penelitian ini adalah :

1. Mengevaluasi kinerja metode *Convolutional Neural Network* (CNN) dalam mengklasifikasikan jenis penyakit paru-paru (pneumonia, tuberkulosis, dan COVID-19) berdasarkan citra rontgen.
2. Menentukan tingkat akurasi metode *Convolutional Neural Network* (CNN) dalam mengklasifikasikan penyakit paru-paru pneumonia, tuberkulosis, dan COVID-19 berdasarkan citra rontgen.
3. Menganalisis pengaruh variasi *hyperparameter* (*epoch*, *mini batch size*, dan variasi dataset) dan *Optimizer* (Adam, SGDM, RMSProp) terhadap kinerja model CNN dalam klasifikasi penyakit paru-paru.

1.5 Manfaat Penelitian

Manfaat dalam penelitian ini yaitu :

1. Dapat merealisasikan sebuah model yang dapat mengklasifikasikan jenis penyakit paru-paru dengan mendeteksi citra rontgen.
2. Untuk mengetahui bagaimana proses penerapan metode *Convolution Neural Network* dalam proses penerapan klasifikasi jenis penyakit paru-paru berdasarkan hasil rontgen.

3. Sebagai alternatif tenaga ahli/dokter spesialis paru-paru dalam mendiagnosa jenis penyakit pada citra rontgen paru-paru.

1.6 Struktur Penulisan Penelitian

Penulisan penelitian terbagi menjadi 5 bab yang diantaranya:

BAB I PENDAHULUAN

Berisi gambaran umum mengenai penelitian yang akan diteliti yang memaparkan tentang latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan

BAB II TINJAUAN PUSTAKA

Berisi dasar-dasar teori yang akan digunakan dalam penelitian sebagai acuan bagi penulis dalam melakukan penelitian.

BAB III METODOLOGI PENELITIAN

Berisi langkah-langkah perancangan dan pengolahan awal yang akan diteliti.

BAB IV HASIL DAN PEMBAHASAN

Berisi analisa-analisa masalah yang berperan pada tingkat akurasi dan hasil penelitian yang telah diteliti

BAB V KESIMPULAN DAN SARAN

Berisi kesimpulan pada seluruh isi penelitian dan saran-saran dari penulis yang berkaitan dengan penelitian agar lebih baik.