

BAB 3

METODOLOGI PENELITIAN

3.1 Studi Literatur

Pada tahap awal yang dilakukan dalam penelitian ini adalah melakukan studi literatur dengan menelaah serta memahami berbagai sumber referensi, termasuk artikel ilmiah, jurnal, dan buku yang memiliki relevansi dengan topik penelitian. Studi literatur ini bertujuan untuk memperoleh pemahaman mendalam terkait metode dan algoritma yang digunakan dalam pengembangan sistem, seperti algoritma Discrete Cosine Transform (DCT), metode ekstraksi fitur Gray Level Co-occurrence Matrix (GLCM), serta Convolutional Neural Network (CNN).

Selain itu, dilakukan pula analisis terhadap penelitian-penelitian sebelumnya yang memiliki keterkaitan dengan topik yang diangkat, guna memperoleh wawasan tambahan dan memperkuat dasar teori yang digunakan.

3.2 Pengumpulan Data

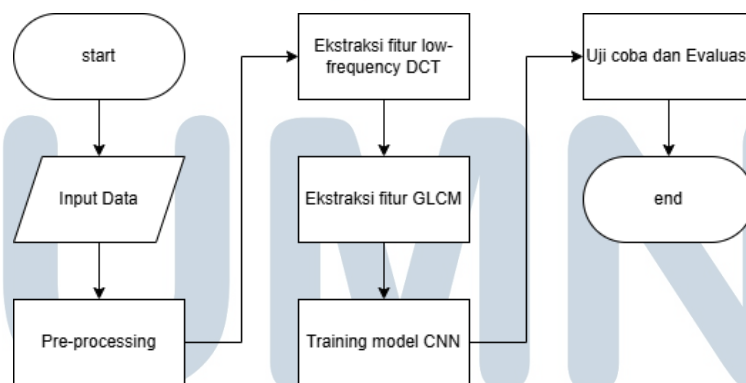
Pada tahap ini, dilakukan proses pencarian dataset yang sesuai dan memiliki tingkat kredibilitas tinggi untuk mendukung penelitian. Dataset yang digunakan dalam penelitian ini diperoleh melalui situs web Kaggle, sebuah platform berbasis komunitas yang menyediakan berbagai kumpulan data untuk analisis dan pengembangan model kecerdasan buatan. Dataset yang dipilih berjudul Chest X-Ray Images (Pneumonia), yang dibuat oleh Paul Mooney dengan tujuan utama membedakan antara citra rontgen paru-paru manusia yang normal dan citra rontgen paru-paru yang mengalami infeksi Pneumonia [14].

Dataset ini terdiri dari ribuan gambar rontgen dada yang telah dikategorikan berdasarkan kondisi medis pasien, sehingga memungkinkan penelitian dalam bidang deteksi dan klasifikasi penyakit paru-paru. Selain itu, dataset ini banyak digunakan dalam berbagai penelitian di bidang *computer vision* dan *deep learning*, terutama dalam pengembangan model untuk diagnosis berbasis citra medis. Dengan kualitas data yang baik serta keberagaman sampel yang tersedia, dataset ini menjadi pilihan yang tepat untuk mendukung penelitian ini dalam mengoptimalkan kualitas citra X-ray paru-paru sebelum dilakukan klasifikasi menggunakan model Convolutional Neural Network (CNN).

3.3 Perancangan

Setelah memperoleh dataset yang relevan dan sesuai, tahap berikutnya adalah merancang alur kerja model. Perancangan ini divisualisasikan dalam bentuk representasi grafis untuk mempermudah pemahaman. Salah satu metode yang umum digunakan dalam menggambarkan alur kerja secara sistematis adalah flowchart. Flowchart merupakan representasi visual yang menyajikan tahapan-tahapan dalam suatu algoritma atau sistem secara terstruktur dan sistematis [15].

Pada tahap ini, dilakukan perancangan Flowchart untuk model yang akan dikembangkan. Gambar 3.1 menyajikan flowchart rancangan model yang digunakan dalam penelitian ini. Proses dimulai dengan memasukkan data gambar dari dataset yang telah dipilih, kemudian dilanjutkan dengan tahap pre-processing data. Selanjutnya, dilakukan ekstraksi fitur menggunakan metode Discrete Cosine Transform (DCT) pada komponen low-frequency serta ekstraksi fitur menggunakan Gray-Level Co-occurrence Matrix (GLCM). Fitur-fitur yang telah diperoleh kemudian digunakan untuk melatih model Convolutional Neural Network (CNN). Tahapan akhir meliputi pengujian dan evaluasi kinerja model guna menilai efektivitas sistem yang dikembangkan.



Gambar 3.1. Flowchart perancangan sistem

3.4 Pre-processing

Pada tahap ini, citra yang akan diproses terlebih dahulu melalui proses normalisasi. Normalisasi dilakukan dengan cara membagi setiap nilai piksel pada citra dengan angka 256 atau 255, tergantung pada rentang nilai intensitas piksel yang digunakan (biasanya dalam format 8-bit, yaitu dari 0 hingga 255). Tujuan utama dari langkah ini adalah untuk mengubah skala nilai piksel agar berada dalam

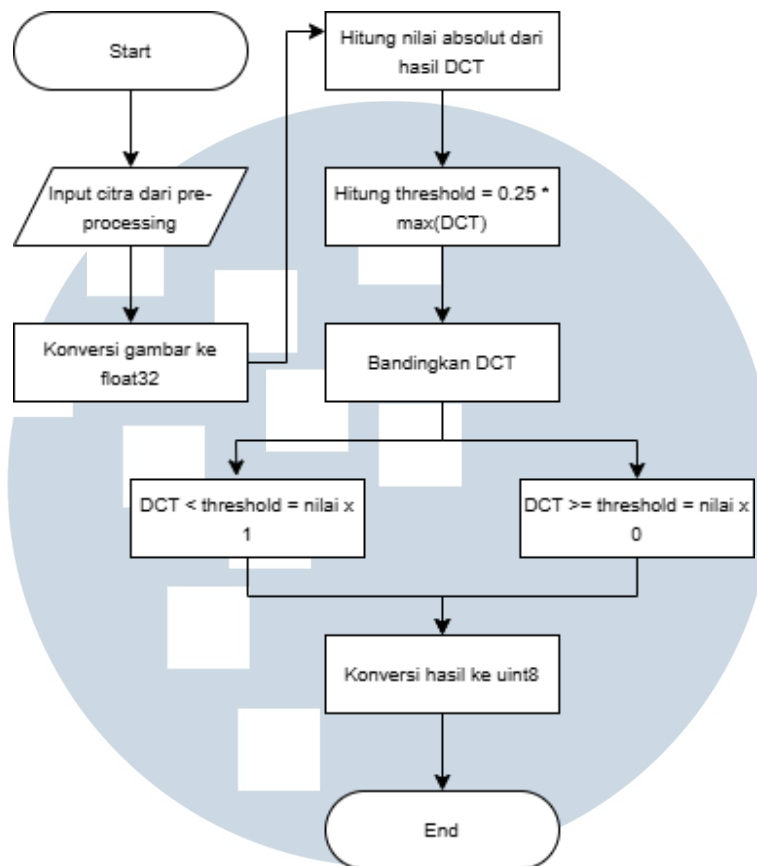
rentang $[0, 1]$, sehingga proses komputasi yang melibatkan data citra dapat berjalan lebih stabil dan efisien. Dalam konteks komputasi numerik, penggunaan nilai yang lebih kecil dan terstandarisasi dapat mengurangi risiko ketidakseimbangan angka dalam perhitungan matriks maupun algoritma numerik lainnya, terutama dalam algoritma optimasi dan pembelajaran mesin yang sensitif terhadap skala data.

Selain itu, normalisasi juga berperan penting dalam meningkatkan kualitas citra secara umum sebelum dilakukan proses-proses lanjutan seperti ekstraksi fitur, transformasi, atau klasifikasi. Dengan nilai piksel yang telah dinormalisasi, citra menjadi lebih seragam dan perbedaan kontras antar bagian citra dapat terlihat lebih jelas, sehingga fitur-fitur penting dalam citra lebih mudah dikenali oleh sistem. Hal ini berdampak langsung pada peningkatan akurasi dan konsistensi hasil analisis yang dilakukan di tahap-tahap berikutnya. Oleh karena itu, proses normalisasi menjadi salah satu tahap awal yang sangat krusial dalam alur pemrosesan citra digital, khususnya dalam penelitian ini yang bertujuan untuk melakukan optimasi kualitas citra x-ray paru-paru menggunakan algoritma tertentu.

3.5 Ekstraksi fitur low-frequency DCT

Pada tahap ini, dilakukan proses ekstraksi fitur pada komponen frekuensi rendah menggunakan transformasi Discrete Cosine Transform (DCT). Proses ini diawali dengan memasukkan citra yang telah dilalui oleh tahap pre-processing. Selanjutnya, perhitungan transformasi DCT dilakukan dengan memanfaatkan fungsi DCT yang tersedia dalam library OpenCV.

Hasil transformasi DCT tersebut digunakan untuk menghitung nilai absolut DCT serta menentukan nilai ambang batas (threshold). Nilai threshold yang digunakan ditetapkan sebesar 25% dari koefisien DCT tertinggi, yang diperoleh dengan mengalikan nilai maksimum DCT dengan faktor 0,25. Setelah nilai threshold diperoleh, dilakukan perbandingan antara nilai absolut DCT dengan nilai threshold tersebut. Jika nilai absolut DCT lebih kecil dari threshold, maka nilainya tetap dipertahankan dengan dikalikan 1. Namun, jika nilai absolut DCT lebih besar atau sama dengan threshold, maka nilainya diubah dengan dikalikan 0. Hasil akhir dari proses ini kemudian dikonversi ke dalam tipe data uint8 dan disimpan dalam bentuk list untuk digunakan dalam tahap pemrosesan selanjutnya.

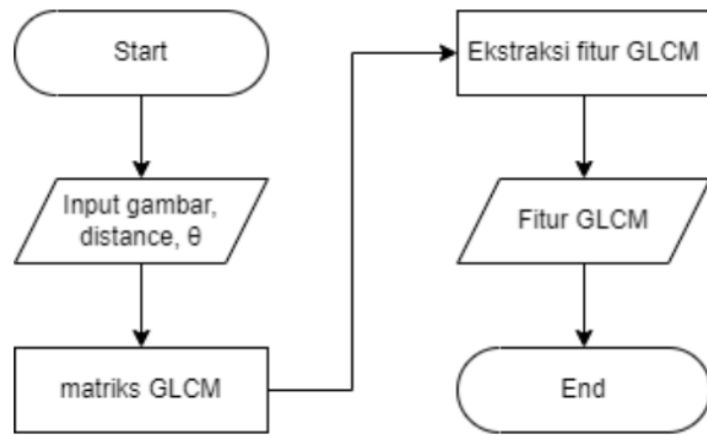


Gambar 3.2. Flowchart Ekstraksi fitur low-frequency DCT

3.6 Ekstraksi fitur GLCM

Pada tahap ini, dilakukan ekstraksi fitur menggunakan metode Gray-Level Co-Occurrence Matrix (GLCM) pada data yang telah melalui tahap ekstraksi fitur Discrete Cosine Transform (DCT) frekuensi rendah. Proses diawali dengan memasukkan citra bersama dengan parameter jarak antar piksel serta nilai sudut yang menentukan arah piksel yang akan dianalisis. Dalam proses ini, digunakan empat nilai sudut, yaitu 0° , 45° , 90° , dan 135° .

Selanjutnya, perhitungan matriks GLCM dilakukan dengan memanfaatkan fungsi `graycomatrix()`, `graycoprops()`, dan `shannon_entropy()` dari library `scikit-image`. Setelah itu, dilakukan ekstraksi berbagai fitur GLCM, di antaranya energy, contrast, correlation, homogeneity, dissimilarity, entropy, dan autocorrelation. Pada tahap akhir, dilakukan normalisasi terhadap fitur-fitur GLCM guna memastikan bahwa seluruh variabel berada dalam skala yang serupa, dengan memanfaatkan fungsi `StandardScaler` dari library `scikit-learn`.



Gambar 3.3. Gambar 3.3 Flowchart Ekstraksi GLCM

3.7 Ekstraksi fitur CNN

Pada tahap ini, dilakukan proses pelatihan dan pengujian data menggunakan model Convolutional Neural Network (CNN). Model CNN yang diterapkan merupakan varian LeNet yang telah dimodifikasi agar sesuai dengan karakteristik dataset yang digunakan. Struktur model diawali dengan lapisan input yang berfungsi menerima citra sebagai masukan. Lapisan selanjutnya terdiri dari convolutional layer dan max pooling layer, yang bertugas mengekstraksi fitur-fitur penting dari citra input. Model LeNet yang digunakan memiliki dua lapisan konvolusi yang berperan dalam mempelajari pola dan karakteristik dari gambar yang diolah. Selain itu, diterapkan batch normalization untuk meningkatkan stabilitas pelatihan serta mengurangi risiko overfitting.

Setelah proses ekstraksi fitur melalui lapisan konvolusi selesai, representasi data diubah menggunakan flatten layer agar data berbentuk dua dimensi (matriks) dapat dikonversi menjadi satu dimensi (vektor). Selanjutnya, hasil ekstraksi fitur diproses melalui dua dense layers atau fully connected layers, yang bertanggung jawab dalam melakukan klasifikasi berdasarkan fitur yang telah diekstraksi dari metode Discrete Cosine Transform (DCT) dan Gray-Level Co-Occurrence Matrix (GLCM). Pada lapisan terakhir, diterapkan dropout layer sebagai mekanisme regulasi untuk mengurangi kemungkinan overfitting dan meningkatkan generalisasi model.

3.8 Uji Coba Evaluasi

Pada tahap ini, model yang telah dilatih akan diuji menggunakan data testing untuk melihat seberapa baik performanya. Hasil pengujian biasanya ditampilkan dalam bentuk akurasi dan loss. Selain itu, confusion matrix juga digunakan sebagai alat evaluasi untuk menilai kinerja model secara lebih mendalam. Dari proses pengujian ini, kita bisa mendapatkan beberapa metrik penting seperti accuracy, loss, recall, precision, dan F1 score, yang secara bersama-sama memberikan gambaran lengkap tentang seberapa efektif model tersebut dalam melakukan prediksi.

