

BAB 2

LANDASAN TEORI

2.1 Tinjauan Teori

Peningkatan kualitas citra X-ray paru-paru menjadi fokus utama dalam penelitian ini dengan memanfaatkan algoritma ekstraksi fitur DCT dan GLCM pada sistem klasifikasi berbasis CNN. Pendekatan ini dipilih karena kombinasi dari transformasi domain frekuensi dan analisis tekstur dinilai mampu menangkap karakteristik penting dari citra medis, khususnya dalam mengidentifikasi struktur dan pola yang relevan terhadap kondisi paru-paru. Untuk menunjang proses tersebut, dilakukan tinjauan pustaka yang mendalam terhadap literatur relevan guna memperoleh pemahaman menyeluruh tentang perkembangan terbaru dalam pengolahan citra medis, baik dari segi metode peningkatan kualitas citra maupun sistem klasifikasi berbasis kecerdasan buatan.

Integrasi algoritma lain dievaluasi sebagai langkah strategis dalam memperkaya hasil pengolahan citra. Evaluasi ini mencakup analisis terhadap algoritma modern yang memiliki performa tinggi dalam domain serupa, termasuk kemungkinan penggunaan teknik pre-processing, augmentasi data, atau optimisasi parameter yang dapat meningkatkan efisiensi sistem secara keseluruhan. Dengan demikian, penelitian ini tidak hanya berfokus pada pengembangan metode baru, tetapi juga mengadopsi pendekatan holistik dalam menyusun sistem klasifikasi citra medis yang lebih akurat.

2.2 Citra Digital

Citra merupakan representasi dari objek nyata dalam bentuk fungsi intensitas dua dimensi yang biasa dituliskan sebagai $f(x, y)$, di mana variabel x dan y merepresentasikan koordinat spasial pada bidang gambar, dan $f(x, y)$ menunjukkan tingkat kecerahan (brightness) atau nilai intensitas cahaya pada titik tertentu dalam citra tersebut. Dengan kata lain, setiap titik dalam bidang citra memiliki nilai intensitas yang menyatakan seberapa terang atau gelap suatu titik berdasarkan pantulan cahaya atau sensor pencitraan. Citra digital sendiri merupakan versi terdiskretisasi dari citra kontinu, di mana representasi visual dua dimensi ini direpresentasikan sebagai kumpulan nilai numerik dalam bentuk digital yang terdiri dari elemen-elemen terkecil yang disebut sebagai piksel (pixel).

Setiap piksel menyimpan informasi tentang intensitas atau warna, tergantung pada jenis citra, dan bersama-sama membentuk keseluruhan gambar digital. Piksel ini menjadi komponen fundamental dalam pemrosesan citra digital karena seluruh proses transformasi, analisis, hingga klasifikasi bergantung pada informasi yang terkandung di dalamnya. Umumnya, bentuk dari citra digital adalah persegi panjang atau bujur sangkar karena efisiensi dalam penyimpanan dan pengolahan data digital, meskipun dalam beberapa sistem pencitraan, seperti pada sensor mosaik tertentu, bentuk elemen penyusunnya bisa berupa segi enam.

Dalam praktiknya, citra digital dapat diklasifikasikan ke dalam beberapa format tampilan berdasarkan jumlah kanal warna dan informasi yang dimilikinya. Tiga jenis citra digital yang paling umum digunakan dalam berbagai aplikasi pengolahan citra adalah: citra biner, yaitu citra yang hanya memiliki dua tingkat warna (hitam dan putih) dan digunakan pada aplikasi segmentasi atau pemisahan objek latar; citra grayscale, yang merepresentasikan intensitas cahaya dalam skala abu-abu dari hitam ke putih tanpa informasi warna; serta citra warna (RGB), yang menyimpan informasi visual dalam tiga kanal warna utama: merah, hijau, dan biru. Ketiga jenis citra ini memiliki karakteristik masing-masing dan digunakan sesuai dengan kebutuhan aplikasi yang dihadapi, baik dalam bidang medis, industri, keamanan, maupun teknologi informasi. Dengan memahami struktur dan jenis citra digital, maka pemrosesan citra lebih lanjut dapat dilakukan secara optimal sesuai dengan tujuan analisis dan transformasi yang diinginkan [5].

2.3 Foto Rontgen

Sinar X atau foto rontgen adalah pemeriksaan menggunakan radiasi elektromagnetik dalam jumlah kecil untuk mendapatkan gambaran struktur tubuh bagian dalam tanpa melakukan pembedahan [6]. Pemeriksaan sinar X akan menghasilkan gambar yang bisa dicetak atau dilihat dalam bentuk gambar digital. Teknik ini memungkinkan visualisasi bagian-bagian tubuh seperti tulang, paru-paru, dan organ dalam lainnya secara non-invasif, sehingga sangat membantu dalam proses diagnosis dan pemantauan kondisi kesehatan pasien.

Prosedur pencitraan ini bekerja dengan cara memancarkan sinar-X ke tubuh, di mana jaringan dengan kepadatan yang berbeda akan menyerap sinar tersebut dalam kadar yang berbeda. Foto Rontgen banyak digunakan dalam berbagai bidang kedokteran, terutama untuk mendeteksi patah tulang, infeksi paru-paru seperti pneumonia atau tuberkulosis, kelainan jantung, dan kondisi medis lainnya.

Meskipun melibatkan paparan terhadap radiasi, jumlah dosis yang digunakan dalam pemeriksaan ini sangat kecil dan umumnya dianggap aman jika dilakukan sesuai dengan standar keselamatan radiasi.

2.4 Ekstraksi Fitur

Ekstraksi fitur merupakan tahap fundamental dalam proses pengolahan data, khususnya pada citra digital, yang bertujuan untuk memperoleh ciri-ciri atau atribut penting dari suatu objek yang bertujuan untuk membedakannya dengan objek lain [7].

Proses ini memiliki peranan krusial karena hasil fitur yang dihasilkan akan menjadi input utama dalam proses klasifikasi, sehingga kualitas fitur sangat berpengaruh terhadap performa dan akurasi model klasifikasi. Selain fungsi utamanya dalam mereduksi dimensi data agar pengolahan menjadi lebih efisien, ekstraksi fitur juga meningkatkan kemampuan model untuk mengenali dan menginterpretasikan pola-pola yang signifikan dalam data. Maka dari itu, ekstraksi fitur merupakan langkah awal yang sangat menentukan keberhasilan dalam berbagai aplikasi klasifikasi, termasuk dalam bidang pengolahan citra medis dan sistem kecerdasan buatan.

2.5 Discrete Cosine Transform (DCT)

Discrete Cosine Transform (DCT) merupakan metode yang banyak digunakan dalam proses ekstraksi fitur karena kemampuannya dalam mengonversi citra dari domain spasial ke domain frekuensi [4]. Keunggulan utama DCT terletak pada kemampuannya dalam memampatkan energi dan mengurangi korelasi antar piksel, sehingga transformasi ini dianggap optimal dalam mempertahankan informasi penting dari citra sekaligus mengurangi ketergantungan statistik antar koefisien transformasi [8]. Dalam proses ekstraksi fitur, DCT lebih efektif jika hanya mengambil komponen *low-frequency*, karena bagian *mid-frequency* dan *high-frequency* cenderung mengandung lebih banyak *noise* yang dapat mengganggu kualitas informasi yang diperoleh [9]. Prinsip dasar DCT yang digunakan dalam penelitian ini didasarkan pada rumus 2.1.

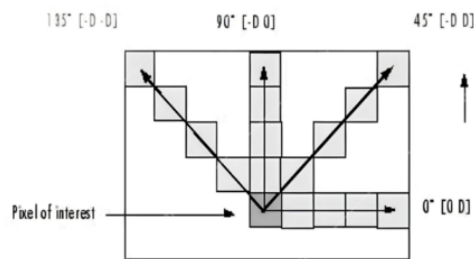
$$C(u) = \sigma(u) \sum_{x=0}^{N-1} f(x) \cos \left[\frac{\pi(2x+1)u}{2N} \right], \quad u = 0, 1, \dots, N-1 \quad (2.1)$$

Untuk DCT berbasis 2D dapat dilihat pada rumus 2.2.

$$C(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \cos \left[\frac{\pi(2x+1)u}{2N} \right] \left\{ \sum_{y=0}^{N-1} f(x, y) \cos \left[\frac{\pi(2y+1)v}{2N} \right] \right\} \quad (2.2)$$

2.6 Gray Level Co-occurrence Matrix (GLCM)

Metode fitur ekstraksi *Gray level co-occurrence matrix* (GLCM) diterapkan dalam pemrosesan citra untuk menyelesaikan berbagai permasalahan klasifikasi gambar. GLCM merepresentasikan informasi statistik tingkat kedua dengan menganalisis hubungan keabuan antara piksel-piksel yang berdekatan dalam sebuah citra [10]. GLCM menggambarkan hubungan statistik orde kedua dengan menganalisis distribusi tingkat keabuan antara piksel yang berdekatan dalam suatu citra [4]. Untuk menganalisis pola tekstur dalam suatu citra, GLCM sering diterapkan dalam empat orientasi sudut, yaitu 0° , 45° , 90° , dan 135° [3]. Gambar sudut dapat dilihat sebagai berikut



Gambar 2.1. Gambar 2.1 Sudut GLCM

Ekstraksi fitur menggunakan GLCM bertujuan untuk memperoleh karakteristik tekstur suatu citra seperti *Contrast*, *Correlation*, *Energy*, dan *Homogeneity* [3].

- *Contrast* (CON) adalah fitur yang mengukur tingkat perbedaan intensitas warna atau skala keabuan antara piksel dalam suatu citra [3]. *Contrast* dapat dihitung menggunakan rumus 2.3.

$$Contrast = \sum_{i=1}^N \sum_{j=1}^N (i - j)^2 GLCM(i, j) \quad (2.3)$$

- *Correlation* (COR) adalah nilai yang menunjukkan tingkat keterkaitan linear antara tingkat keabuan piksel dalam suatu citra. Fitur ini berfungsi sebagai indikator untuk mendeteksi adanya pola linier dalam struktur gambar [11]. *Correlation* ini dapat dihitung menggunakan persamaan 2.4.

$$Correlation = \frac{\sum_{i=1}^N \sum_{j=1}^N (i - \mu_i)(j - \mu_j) GLCM(i, j)}{\sigma_i \sigma_j} \quad (2.4)$$

- *Energy* adalah ukuran yang menggambarkan tingkat keseragaman keseluruhan dalam suatu citra [12]. *Energy* dapat dihitung menggunakan persamaan 2.5.

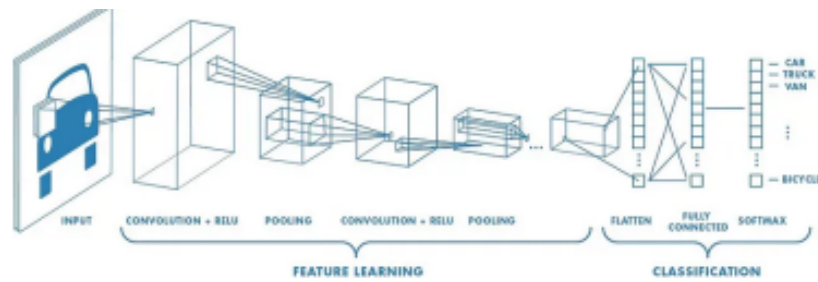
$$Energy = \sum_{i=1}^N \sum_{j=1}^N GLCM(i, j)^2 \quad (2.5)$$

- *Homogeneity* digunakan untuk mengukur tingkat keseragaman tekstur pada area lokal dalam suatu citra. Semakin tinggi nilai *homogeneity*, semakin seragam distribusi teksturnya [11]. *Homogeneity* dapat dihitung menggunakan persamaan 2.6.

$$Homogeneity = \sum_{i=1}^N \sum_{j=1}^N \frac{GLCM(i, j)}{1 + |i - j|} \quad (2.6)$$

2.7 Convolutional Neural Network (CNN)

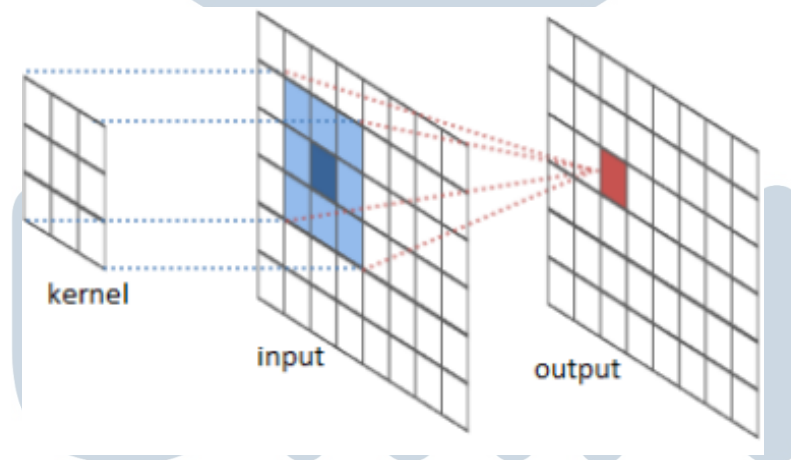
Convolutional Neural Network (CNN) merupakan sebuah jaringan saraf yang berfungsi untuk memproses data yang telah diketahui dalam bentuk topologi grid. Pada CNN terdapat beberapa *layer*, seperti *convolutional layer*, *pooling layer*, dan *fully connected layer* [13]. Berikut gambar struktur umum dari CNN yang dapat dilihat pada gambar 2.2.



Gambar 2.2. Struktur Umum CNN

2.7.1 Convolutional layer

Convolutional dalam CNN merupakan lapisan yang bertugas mengekstraksi fitur dari suatu citra dengan menggeser piksel demi piksel. Proses ini dilakukan menggunakan kernel, yaitu filter konvolusi yang membagi gambar menjadi bagian-bagian kecil untuk menghasilkan *feature map* yang merepresentasikan karakteristik penting dari citra tersebut [5]. Berikut gambar lapisan *convolutional* yang dapat dilihat pada gambar 2.3.



Gambar 2.3. *Convolutional Layer* pada CNN

2.7.2 Pooling Layer

Pooling layer merupakan kunci dalam arsitektur *Convolutional Neural Network*(CNN) yang berperan penting dalam mengurangi dimensi data hasil konvolusi dari *convolutional layer* sebelumnya. Fungsinya tidak hanya untuk menyederhanakan representasi data tetapi juga untuk meningkatkan efisiensi komputasi. Dengan adanya pengurangan dimensi maka akan membuat model

menjadi memahami fitur-fitur yang esensial dan dapat memangkas beban pada setiap lapisan.

Pooling terdiri dari dua jenis utama, yaitu *max pooling* dan *average pooling*. *Max pooling* berfungsi mengambil nilai tertinggi dari area gambar yang diproses oleh kernel, sedangkan *average pooling* menghitung rata-rata nilai piksel dalam area tersebut untuk menghasilkan fitur yang lebih halus [5].

2.7.3 Fully Connected Layer

Fully connected layer adalah lapisan akhir dalam arsitektur *Convolutional Neural Network* (CNN) yang berfungsi untuk menggabungkan semua fitur yang telah diekstraksi pada lapisan sebelumnya. Pada tahap ini, setiap neuron dalam *fully connected layer* memiliki hubungan langsung dengan seluruh neuron di lapisan sebelumnya maupun lapisan berikutnya. Lapisan ini berperan dalam proses klasifikasi dengan mengolah informasi dari fitur yang telah diproses sebelumnya dan meneruskannya ke output akhir [5].

2.7.4 Confusion Matrix

Confusion matrix adalah sebuah matriks berukuran $N \times N$ yang digunakan sebagai alat untuk mengevaluasi performa model machine learning dalam melakukan prediksi. Nilai N dalam matriks ini menunjukkan jumlah kelas yang menjadi target klasifikasi. *Confusion matrix* menyajikan representasi dalam bentuk tabel yang merangkum jumlah prediksi yang dikategorikan dengan benar maupun yang salah oleh model klasifikasi. Sebagai ilustrasi, Tabel 2.1 menampilkan contoh *confusion matrix* berukuran 2×2 yang digunakan dalam klasifikasi biner.

Tabel 2.1. Confusion Matrix

	Positif (Aktual)	Negatif (Aktual)
Positif (Prediksi)	TP	FP
Negatif (Prediksi)	FN	TN

- TP: *True Positive* yaitu nilai positif diprediksi model sebagai nilai positif
- FP: *False Positive* yaitu prediksi model adalah nilai positif, dan hasil prediksinya salah.

- TN: *True Negative* yaitu prediksi model adalah nilai negatif, dan hasil prediksinya salah.
- FN: *True Positive* yaitu nilai negatif diprediksi model sebagai nilai negatif.

Terdapat beberapa metrik yang dapat disimpulkan dari confusion matrix, yaitu:

- *Accuracy* merupakan matrik evaluasi yang mengukur persentase prediksi yang benar dari keseluruhan data yang diuji. Semakin tinggi nilai akurasi, semakin baik performa model dalam melakukan klasifikasi. Rumus *Accuracy* dapat dilihat pada rumus 2.7.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (2.7)$$

- *Precision* merupakan perbandingan antara jumlah *true positive* (TP) dengan total seluruh prediksi yang diklasifikasikan sebagai positif. Rumus *Precision* dapat dilihat pada rumus 2.8.

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (2.8)$$

- *Recall* adalah matrik yang mengukur sejauh mana model mampu mengidentifikasi kasus positif secara akurat dari keseluruhan data yang sebenarnya termasuk dalam kategori positif. Rumus *Recall* dapat dilihat pada rumus 2.9.

$$Precision = \frac{TP}{P + FN} \quad (2.9)$$

- *F1-Score* adalah nilai rata-rata harmonik antara precision dan recall yang memberikan representasi seimbang dari kedua matriks tersebut. Rumus *F1-Score* dapat dilihat pada rumus 2.10.

$$F1-Score = \frac{1}{\frac{1}{recall} + \frac{1}{precision}} \quad (2.10)$$