

**PENDEKATAN MECHINE LEARNING DENGAN KNN  
UNTUK KLASIFIKASI PNEUMONIA PADA CITRA X-RAY  
DENGAN EKSTRAKSI FITUR GLCM**

**SKRIPSI**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**ACHMAD RIZKI RAMADHAN**

**22.11.4879**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2022**

**PENDEKATAN MECHINE LEARNING DENGAN KNN  
UNTUK KLASIFIKASI PNEUMONIA PADA CITRA X-RAY  
DENGAN EKSTRAKSI FITUR GLCM**

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**ACHMAD RIZKI RAMADHAN**

**22.11.4879**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2022**

## DAFTAR ISI

<b>HALAMAN JUDUL .....</b>	<b>i</b>
<b>DAFTAR ISI.....</b>	<b>ii</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1    Latar Belakang .....	1
1.2    Rumusan Masalah .....	2
1.3    Batasan Masalah.....	3
1.4    Tujuan Penelitian .....	3
1.5    Manfaat Penelitian .....	4
1.6    Sistematika Penulisan.....	4
<b>BAB II TINJAUAN PUSTAKA.....</b>	<b>6</b>
2.1    Studi Literatur .....	6
2.2    Dasar Teori.....	15
<b>BAB III METODE PENELITIAN .....</b>	<b>17</b>
3.1    Objek Penelitian.....	17
3.2    Alur Penelitian .....	18
3.3    Alat dan Bahan.....	20
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>23</b>
<b>BAB V PENUTUP.....</b>	<b>24</b>
5.1    Kesimpulan .....	24
5.2    Saran.....	24
<b>DAFTAR PUSTAKA .....</b>	<b>27</b>



# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Pneumonia merupakan salah satu penyebab utama kematian anak di seluruh dunia, termasuk di Indonesia. Menurut data Riset Kesehatan Dasar 2018, prevalensi pneumonia pada balita di Indonesia meningkat dari 1,6% pada tahun 2013 menjadi 2,0% pada tahun 2018 [1]. Hal ini menunjukkan bahwa pneumonia masih menjadi masalah kesehatan masyarakat yang signifikan. Diagnosis pneumonia secara dini dan akurat sangat penting untuk mencegah komplikasi serius dan menurunkan angka kematian. Namun, keterbatasan tenaga medis dan fasilitas diagnostik, terutama di daerah terpencil, menjadi tantangan dalam penanganan penyakit ini. Oleh karena itu, diperlukan sistem bantu diagnosis yang efektif dan efisien untuk mendukung tenaga medis dalam mengidentifikasi pneumonia.

Perkembangan teknologi kecerdasan buatan, khususnya dalam bidang pengolahan citra medis, membuka peluang untuk mengembangkan sistem bantu diagnosis pneumonia. Salah satu metode yang digunakan adalah ekstraksi fitur tekstur menggunakan *Gray Level Co-occurrence Matrix* (GLCM), yang mampu menangkap pola tekstur pada citra X-ray dada. Fitur-fitur yang dihasilkan dari GLCM dapat digunakan sebagai input untuk algoritma klasifikasi seperti *K-Nearest Neighbor* (KNN), yang dikenal karena kesederhanaannya dan efektivitasnya dalam klasifikasi berbasis kemiripan data. Beberapa penelitian telah menunjukkan bahwa kombinasi GLCM dan KNN dapat menghasilkan akurasi yang tinggi dalam klasifikasi citra medis. Misalnya, penelitian oleh Suharyana et al. (2023) menunjukkan bahwa kombinasi GLCM dan KNN mencapai akurasi hingga 91% dalam klasifikasi pneumonia berbasis citra X-ray [2].

Meskipun demikian, terdapat beberapa keterbatasan dalam penelitian sebelumnya. Sebagian besar penelitian menggunakan dataset yang terbatas dan

belum menguji model pada data yang lebih beragam. Selain itu, beberapa penelitian belum membandingkan performa model dengan metode klasifikasi lain seperti *Support Vector Machine* (SVM) atau *Convolutional Neural Network* (CNN). Penelitian oleh Hismaya & Sulistianingsih (2024) menunjukkan bahwa SVM dengan tuning parameter dapat mencapai akurasi hingga 93,97%, lebih tinggi dibandingkan KNN yang hanya mencapai 81,61% [3]. Namun, SVM memerlukan proses tuning yang kompleks dan waktu komputasi yang lebih lama dibandingkan KNN. Oleh karena itu, diperlukan penelitian lebih lanjut untuk mengevaluasi efektivitas kombinasi GLCM dan KNN dalam klasifikasi pneumonia, serta membandingkannya dengan metode lain.

Penelitian ini bertujuan untuk mengembangkan model klasifikasi pneumonia berbasis citra X-ray dengan pendekatan machine learning menggunakan fitur GLCM dan algoritma KNN. Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan sistem bantu diagnosis pneumonia yang efektif, efisien, dan mudah diimplementasikan di berbagai fasilitas kesehatan, terutama di daerah dengan keterbatasan sumber daya. Secara teoretis, penelitian ini akan memperkaya literatur terkait penerapan metode ekstraksi fitur tekstur dan klasifikasi dalam pengolahan citra medis. Secara praktis, hasil penelitian ini dapat digunakan sebagai dasar untuk mengembangkan aplikasi bantu diagnosis pneumonia yang dapat membantu tenaga medis dalam pengambilan keputusan klinis.

## 1.2 Rumusan Masalah

Berdasarkan permasalahan yang sudah dipaparkan pada latar belakang 1.1, maka peneliti telah merumuskan beberapa masalah pada penelitian ini adalah sebagai berikut:

1. Bagaimana performa algoritma *K-Nearest Neighbor* (KNN) dalam mengklasifikasikan citra *X-ray* untuk mendeteksi pneumonia ketika digunakan bersama fitur tekstur yang diekstraksi menggunakan metode *Gray Level Co-occurrence Matrix* (GLCM)?

2. Fitur GLCM apa saja yang paling berkontribusi terhadap peningkatan akurasi klasifikasi pneumonia pada citra *X-ray*?
3. Seberapa baik *dataset* citra *chest x-ray* yang digunakan pada penelitian ini?

### 1.3 Batasan Masalah

Adapun beberapa batasan yang diterapkan pada penelitian ini adalah sebagai berikut:

1. Jenis Data: Penelitian ini hanya menggunakan dataset citra X-ray dari sumber terbuka (open source) seperti Chest X-ray Pneumonia Dataset yang tersedia di Kaggle. Data dari rumah sakit lokal atau institusi medis lain tidak digunakan.
2. Metode Ekstraksi Fitur: Penelitian hanya menggunakan metode *Gray Level Co-occurrence Matrix* (GLCM) untuk ekstraksi fitur tekstur. Metode ekstraksi lain seperti *Histogram of Oriented Gradients* (HOG), *Local Binary Pattern* (LBP), atau deep feature dari CNN tidak dimasukkan dalam analisis.
3. Preprocessing Citra: Teknik pra-pemrosesan citra dibatasi pada pengubahan ukuran dan konversi ke skala abu-abu (grayscale). Teknik lanjutan seperti data augmentation, segmentasi paru, atau noise filtering tidak diterapkan.
4. Lingkup Diagnostik: Penelitian hanya memfokuskan pada klasifikasi dua kelas, yaitu pneumonia dan normal. Sub-klasifikasi jenis pneumonia (bakterial vs viral) tidak dilakukan.

### 1.4 Tujuan Penelitian

Tujuan yang ingin dicapai oleh peneliti pada penelitian ini diantaranya adalah sebagai berikut:

1. Menghasilkan model klasifikasi citra X-ray berbasis machine learning

menggunakan algoritma *K-Nearest Neighbor* (KNN) yang mampu mengidentifikasi kondisi pneumonia secara akurat berdasarkan fitur tekstur yang diekstraksi melalui metode *Gray Level Co-occurrence Matrix* (GLCM).

2. Menghitung dan menganalisis kontribusi masing-masing parameter GLCM (seperti contrast, homogeneity, energy, entropy) terhadap kinerja model KNN dalam proses klasifikasi citra X-ray, untuk mengetahui fitur mana yang paling relevan dan signifikan dalam membedakan antara citra pneumonia dan normal.
3. Memvalidasi performa model KNN berbasis GLCM melalui metrik evaluasi seperti akurasi, presisi, recall, dan F1-score, serta membandingkannya dengan metode klasifikasi lain seperti SVM dan CNN untuk menilai keunggulan dan keterbatasannya secara empiris dan objektif.

### 1.5 Manfaat Penelitian

Peneliti berharap pada penelitian ini dapat memberikan manfaat dan informasi khususnya didunia medis dan juga penelitian tentang *deep learning*, berikut ini adalah beberapa kemungkinan manfaat yang bisa diberikan pada penelitian ini:

1. Dapat memberikan informasi kepada dokter dan tim medis berupa informasi dan sebuah sistem yang dapat mendiagnosa penyakit *pneumonia* pada paru-paru secara cepat dan akurat.
2. Sebagai bahan referensi untuk penelitian selanjutnya yang terkait klasifikasi penyakit *pneumonia* atau penyakit paru-paru lainnya dengan menggunakan data citra *x-ray* sebagai *input*.

### 1.6 Sistematika Penulisan

Berikut merupakan penjabaran secara sistematis dari tiap bab beserta ringkasan penjelasannya :



## BAB I PENDAHULUAN

Menjelaskan latar belakang, rumusan masalah, batasan masalah, tujuan, manfaat penelitian, dan sistematika penulisan.

## BAB II TINJAUAN PUSTAKA

Berisi teori-teori yang relevan seperti pneumonia, GLCM, KNN, dan kajian penelitian terdahulu.

## BAB III METODE PENELITIAN

Menguraikan objek penelitian, *dataset*, metode ekstraksi fitur, algoritma klasifikasi, serta evaluasi model.

## BAB IV HASIL DAN PEMBAHASAN

Menyajikan hasil implementasi, pengujian model, dan analisis kinerja algoritma klasifikasi.

## BAB V PENUTUP

Berisi kesimpulan hasil penelitian dan saran untuk pengembangan lebih lanjut.

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 Studi Literatur**

Pneumonia merupakan infeksi pada salah satu atau kedua paru-paru yang dapat disebabkan oleh bakteri, virus, atau jamur [4]. Diagnosis pneumonia umumnya dilakukan melalui pemeriksaan fisik dan radiografi dada (chest X-ray) [5]. Interpretasi citra X-ray oleh radiolog memerlukan keahlian dan pengalaman, serta dapat memakan waktu dan rentan terhadap subjektivitas. Oleh karena itu, pengembangan sistem otomatis untuk membantu diagnosis pneumonia menjadi sangat penting.

Dalam beberapa tahun terakhir, machine learning telah menunjukkan potensi besar dalam analisis citra medis, termasuk deteksi dan klasifikasi penyakit paru-paru seperti pneumonia[6], [7], [8], [9]. Berbagai algoritma machine learning telah diterapkan untuk tugas ini, termasuk Convolutional Neural Networks (CNN), Support Vector Machines (SVM), Random Forest (RF), dan K-Nearest Neighbors (KNN).

KNN adalah algoritma non-parametric yang sederhana namun efektif untuk klasifikasi dan regresi. Dalam konteks klasifikasi citra, KNN bekerja dengan mencari k citra tetangga terdekat dalam ruang fitur berdasarkan metrik jarak tertentu (misalnya, Euclidean atau Manhattan) dan memprediksi kelas dari citra uji berdasarkan kelas mayoritas dari k tetangga terdekatnya .

Ekstraksi fitur merupakan langkah penting dalam aplikasi machine learning pada citra. Gray Level Co-occurrence Matrix (GLCM) adalah metode statistik orde kedua yang banyak digunakan untuk mengekstrak fitur tekstur dari citra digital. GLCM menganalisis hubungan spasial antar piksel dengan nilai intensitas yang berbeda dalam suatu citra. Fitur-fitur tekstur yang dapat diekstrak dari GLCM meliputi contrast, correlation, energy, dan homogeneity, yang dapat memberikan informasi diskriminatif untuk klasifikasi citra medis.

Beberapa penelitian sebelumnya telah mengeksplorasi penggunaan machine learning untuk klasifikasi pneumonia pada citra X-ray. Namun, berdasarkan tinjauan literatur awal yang diberikan, fokus utama penelitian-penelitian tersebut adalah pada penggunaan arsitektur deep learning seperti CNN dan Vision Transformers untuk secara otomatis mempelajari fitur-fitur yang relevan dari citra. Penelitian-penelitian tersebut umumnya mencapai akurasi yang tinggi dalam mendeteksi pneumonia[6], [7][8], [9], [10]. Meskipun demikian, penelitian-penelitian tersebut tidak secara spesifik menggunakan algoritma KNN untuk klasifikasi atau metode GLCM untuk ekstraksi fitur.

Sharma dan Guleria [10] melakukan penelitian tentang deteksi pneumonia menggunakan model berbasis deep learning VGG-16 yang dikombinasikan dengan Neural Networks. Mereka membandingkan kinerja model mereka dengan VGG-16 yang dikombinasikan dengan algoritma machine learning lainnya seperti SVM, KNN, RF, dan Naïve Bayes. Hasil penelitian mereka menunjukkan bahwa VGG-16 dengan Neural Networks memberikan kinerja terbaik. Namun, penelitian ini menggunakan KNN sebagai pembanding setelah ekstraksi fitur oleh VGG-16, bukan sebagai metode klasifikasi utama dengan fitur yang diekstraksi secara manual seperti GLCM.

Penelitian lain oleh Singh et al. [6], Shakeel et al. [8], Bakti dan Firdaus [9], Maquen-Niño et al. [11], dan Kurniawati et al. [12] juga berfokus pada penggunaan arsitektur CNN yang berbeda untuk klasifikasi pneumonia dan mencapai hasil yang menjanjikan. Penelitian-penelitian ini memanfaatkan kemampuan CNN dalam mempelajari fitur hierarkis secara otomatis dari data citra. Namun, tidak ada dari penelitian-penelitian ini yang menggunakan pendekatan KNN dengan fitur GLCM.

Oleh karena itu, penelitian ini bertujuan untuk mengeksplorasi efektivitas pendekatan machine learning dengan algoritma KNN yang memanfaatkan fitur tekstur yang diekstraksi menggunakan metode GLCM untuk klasifikasi pneumonia pada citra X-ray. Pendekatan ini berbeda dari penelitian-penelitian sebelumnya yang cenderung fokus pada metode deep learning dengan ekstraksi

fitur otomatis. Dengan mengevaluasi kinerja KNN dengan fitur GLCM, penelitian ini diharapkan dapat memberikan wawasan tambahan mengenai potensi metode machine learning klasik dalam tugas klasifikasi citra medis.

Tabel 2.1 Keaslian Penelitian

No	Judul penelitian	Nama Penulis	Tahun Publikasi	Hasil Penelitian	Perbandingan Penelitian
1	Efficient pneumonia detection using Vision Transformers on chest X-rays	Sukhendra Singh, Manoj Kumar, Abhay Kumar, Birendra Kumar Verma, Kumar Abhishek & Shitharth Selvarajan	2024	Kerangka kerja berbasis Vision Transformer dalam penelitian ini mencapai akurasi 97,61%, sensitivitas 95%, dan spesifisitas 98% dalam mendeteksi pneumonia dari sinar-X dada.	<p>Penelitian ini menggunakan Vision Transformers (ViT), sebuah arsitektur deep learning, bukan KNN.</p> <p>Penelitian ini tidak menggunakan GLCM untuk ekstraksi fitur.</p>
2	A Deep Learning based model for the Detection of Pneumonia from Chest X-Ray Images using VGG-16 and Neural Networks	Shagun Sharma and Kalpna Guleria	2023	Model VGG16 dengan Neural Networks (NN) mencapai akurasi 92,15%, recall 0,9308, presisi 0,9428, dan F1-Score 0,937 untuk dataset pertama. Untuk dataset kedua, model VGG16 dengan NN mencapai akurasi 95,4%, recall 0,954, presisi 0,954, dan F1-score 0,954.	<p>Penelitian ini menggunakan VGG-16, sebuah arsitektur CNN yang spesifik, bukan KNN.</p> <p>Penelitian ini tidak menggunakan GLCM.</p>

				Model VGG16 dengan NN mengungguli VGG16 dengan Support Vector Machine (SVM), VGG16 dengan K-Nearest Neighbor (KNN), VGG16 dengan Random Forest (RF), dan VGG16 dengan Naïve Bayes (NB) untuk kedua dataset.	
3	Pneumonia Recognition in Chest X-rays through Convolutional Neural Networks (CNN)	Memoona Shakeel, Ahmad Naeem, Naeem Aslam, Kamran Abid, and Muhammad Qasim Shafiq	2024	Penelitian ini menggunakan empat jaringan saraf konvolusional (CNN) yang berbeda untuk transfer learning: SqueezeNet, AlexNet, ResNet18, dan DenseNet201. Model yang dikembangkan mencapai akurasi 98% dalam mengklasifikasikan citra sinar-X dada	Penelitian ini menggunakan CNN, bukan KNN.  Penelitian ini tidak menggunakan GLCM.

				<p>sebagai pneumonia viral dan bakteri, akurasi 95% untuk kasus normal dan pneumonia (tanpa membedakan jenisnya), dan akurasi 93,3% untuk klasifikasi ketiga jenis (normal, viral, bakteri).</p> <p>Para penulis menyimpulkan bahwa kinerja model mereka secara signifikan melampaui akurasi yang telah dicapai dalam penelitian sebelumnya. Model ini berpotensi membantu radiolog dalam mendiagnosis pneumonia dengan lebih cepat dan mempercepat skrining di bandara untuk individu dengan penyakit tersebut.</p>	
--	--	--	--	--	--

4	Arsitektur Convolutional Neural Network InceptionResNet-V2 Untuk Pengelompokan Pneumonia Chest X-Ray	Indra Bakti <sup>1</sup> , Mohamad Firdaus <sup>2</sup>	2023	Hasil uji akurasi dari Inception ResNet-V2 menghasilkan 98%. Presisi tiap kelas arsitektur CNN InceptionRestNet-V2 adalah Covid (99%), Lung_Opacity (97%), Normal (98%), Viral_Pneumonia (98%).	Penelitian ini menggunakan InceptionResNet-V2, sebuah arsitektur CNN yang spesifik, bukan KNN. Penelitian ini tidak menggunakan GLCM.
5	Classification Model Using Transfer Learning for the Detection of Pneumonia in Chest X-Ray Images	Gisella Luisa Elena Maquen-Niño, Jhojan Genaro Nuñez-Fernandez, Fany Yesica Taquila-Calderon, Ivan Adrianzén-Olano, Percy De-La-Cruz-VdV, Gilberto Carrión-Barco	2024	Penelitian ini menggunakan convolutional neural network (CNN) dan transfer learning untuk mendeteksi pneumonia pada gambar rontgen dada. Mereka membandingkan tiga model pembelajaran transfer: DenseNet, VGG19, dan ResNet50 versi 2. Model ResNet50 mencapai akurasi terbaik (0,91) dalam	Penelitian ini menggunakan CNN dan transfer learning, bukan KNN. Penelitian ini tidak menggunakan GLCM.



				mengklasifikasikan gambar sebagai "pneumonia" atau "normal".	
6	Identifikasi Pneumonia pada Balita melalui Citra X-ray Menggunakan Metode Convolutional Neural Network (CNN)	Indah Kurniawati, Ridho Akbar, Izza Fahma Kusumawati, Yessie Ardina Kusuma	2024	<p>Penelitian ini menggunakan Convolutional Neural Network (CNN) untuk mengklasifikasikan foto rontgen paru ke dalam tiga kelas: Normal, Pneumonia Ringan, dan Pneumonia Berat.</p> <p>Hasil pengujian akhir menunjukkan bahwa menggunakan 15 epoch, 5 hidden layer, dan 5700 data dari proses klasifikasi menggunakan CNN dapat mencapai akurasi training sebesar 92,48% dan validasi sebesar 91%.</p> <p>Hasil dari 50 data foto rontgen dada menunjukkan akurasi</p>	<p>Penelitian ini menggunakan CNN, bukan KNN.</p> <p>Penelitian ini tidak menggunakan GLCM.</p>

				identifikasi yang sama antara pembacaan oleh dokter dan metode yang diusulkan, dengan waktu pembacaan dokter 15 menit dan metode yang diusulkan hanya 0,2 detik dengan akurasi identifikasi 100%.	
--	--	--	--	---	--

## 2.2 Dasar Teori

### 1. Teori Dasar Mengenai Pneumonia

Pneumonia adalah infeksi paru-paru yang menyebabkan inflamasi pada alveoli dan dapat disebabkan oleh bakteri, virus, atau jamur. Menurut WHO, pneumonia merupakan penyebab utama kematian anak di bawah lima tahun di dunia. Di Indonesia, berdasarkan Riskesdas 2018, prevalensi pneumonia pada balita adalah 2,0% [1]. Deteksi dini pneumonia menjadi sangat penting untuk meningkatkan angka harapan hidup pasien.

### 2. Citra X-ray Dada (Chest X-ray)

Citra X-ray dada merupakan metode pencitraan medis yang digunakan untuk mendeteksi kelainan pada paru-paru. Dalam penelitian ini, citra X-ray digunakan sebagai input utama dalam sistem klasifikasi otomatis. Pendeteksian pneumonia melalui analisis citra X-ray dinilai efektif, namun membutuhkan pendekatan otomatis yang akurat untuk mengurangi ketergantungan pada radiologis manusia.

### 3. Ekstraksi Fitur Menggunakan GLCM

Gray Level Co-occurrence Matrix (GLCM) adalah metode statistik tekstur untuk mengukur hubungan spasial antar piksel dengan nilai intensitas tertentu dalam suatu citra. GLCM digunakan untuk mengekstraksi fitur tekstur seperti contrast, correlation, energy, dan homogeneity, yang berperan sebagai variabel input numerik dalam model klasifikasi. Menurut Haralick et al. (1973), GLCM efektif untuk mengkarakterisasi tekstur citra dan digunakan dalam banyak penelitian medis berbasis citra [2].

### 4. Algoritma K-Nearest Neighbor (KNN)

K-Nearest Neighbor (KNN) adalah algoritma klasifikasi berbasis instance yang bekerja dengan mengklasifikasikan data baru berdasarkan kemiripannya dengan data latih terdekat. Algoritma ini tidak memerlukan

asumsi distribusi data dan sangat bergantung pada metrik jarak (misalnya Euclidean distance). Dalam konteks penelitian ini, KNN digunakan untuk mengklasifikasikan citra X-ray sebagai pneumonia atau normal berdasarkan fitur GLCM yang diekstrak.

## 5. Hubungan antar Variabel dalam Penelitian

Model penelitian ini melibatkan tiga variabel utama:

Variabel bebas (X): Nilai-nilai fitur tekstur hasil ekstraksi GLCM dari citra X-ray (misalnya: contrast, energy, correlation).

Variabel terikat (Y): Label klasifikasi, yaitu kategori kondisi paru-paru (pneumonia atau normal).

Variabel antara (parameter klasifikasi): Nilai  $k$  pada algoritma KNN, yang memengaruhi akurasi klasifikasi.

Berdasarkan pendekatan ini, hubungan antar variabel diasumsikan bahwa fitur tekstur dari GLCM memiliki korelasi yang signifikan dengan status kesehatan paru-paru, dan nilai  $k$  pada KNN berperan sebagai penguat keputusan klasifikasi.

## **BAB III METODE PENELITIAN**

### **3.1 Objek Penelitian**

Objek penelitian dalam studi ini adalah citra X-ray paru-paru (chest X-ray) yang digunakan untuk klasifikasi kondisi pneumonia dengan pendekatan machine learning menggunakan algoritma K-Nearest Neighbor (KNN). Citra X-ray yang dimaksud mencakup dua kelas utama, yaitu citra paru-paru normal dan citra paru-paru yang terindikasi pneumonia.

Citra-citra tersebut akan dianalisis dan diproses melalui teknik ekstraksi fitur tekstur menggunakan metode Gray Level Co-occurrence Matrix (GLCM). Parameter-parameter GLCM seperti contrast, homogeneity, energy, dan entropy menjadi fokus utama dalam pembentukan fitur yang merepresentasikan karakteristik tekstur dari citra X-ray tersebut.

Dengan memanfaatkan parameter-parameter tekstur tersebut, algoritma KNN diterapkan untuk melakukan proses klasifikasi guna membedakan antara kondisi paru-paru normal dan yang mengalami pneumonia. Selain itu, objek penelitian juga mencakup perbandingan performa KNN dengan algoritma lain seperti Support Vector Machine (SVM) dan Convolutional Neural Network (CNN) untuk mengukur keefektifan pendekatan KNN berbasis GLCM secara empiris.

Secara keseluruhan, objek penelitian ini meliputi:

- Dataset citra X-ray paru-paru yang bersumber dari dataset terbuka seperti Chest X-Ray Images (Pneumonia) dari Kaggle atau sumber medis terpercaya lainnya.
- Fitur tekstur yang dihasilkan dari ekstraksi GLCM, yang digunakan sebagai input model klasifikasi.
- Model klasifikasi KNN, serta pembandingnya yaitu SVM dan CNN.

- Metrik evaluasi performa model, yaitu akurasi, presisi, recall, dan F1-score.

Dengan cakupan tersebut, objek penelitian ini merepresentasikan hubungan antara teknik ekstraksi fitur berbasis tekstur dan efektivitas klasifikasi citra medis menggunakan pendekatan machine learning.

### **3.2 Alur Penelitian**

Penelitian ini dilakukan melalui serangkaian tahapan sistematis guna mencapai tujuan klasifikasi pneumonia pada citra X-ray menggunakan algoritma K-Nearest Neighbor (KNN) berbasis fitur tekstur yang diekstraksi melalui metode Gray Level Co-occurrence Matrix (GLCM). Adapun alur penelitian ini ditunjukkan pada Gambar dan dijelaskan sebagai berikut:

#### **1. Studi Literatur**

Tahap awal penelitian dilakukan dengan menelaah berbagai referensi ilmiah yang berkaitan dengan pneumonia, citra X-ray paru-paru, teknik ekstraksi fitur tekstur (GLCM), dan algoritma klasifikasi machine learning seperti KNN, SVM, dan CNN. Studi literatur ini bertujuan untuk membangun dasar teori, memahami pendekatan yang telah ada, serta merumuskan arah dan ruang lingkup penelitian.

#### **2. Pengumpulan Dataset**

Dataset citra X-ray dikumpulkan dari sumber terbuka yang terpercaya seperti Chest X-Ray Images (Pneumonia) dari Kaggle. Dataset ini terdiri dari dua kelas utama, yaitu citra paru-paru normal dan citra paru-paru yang mengalami pneumonia.

#### **3. Pra-pemrosesan Citra**

Citra yang dikumpulkan dilakukan pra-pemrosesan untuk menyamakan format dan meningkatkan kualitas data. Proses ini meliputi konversi ke grayscale, resize ke ukuran seragam, serta normalisasi nilai piksel untuk memudahkan proses ekstraksi fitur.

#### 4. Ekstraksi Fitur Tekstur (GLCM)

Setiap citra grayscale diekstraksi fitur teksturnya menggunakan metode Gray Level Co-occurrence Matrix (GLCM). Parameter yang dihitung meliputi contrast, homogeneity, energy, dan entropy. Hasil ekstraksi ini membentuk vektor fitur yang merepresentasikan karakteristik tekstur dari citra.

#### 5. Pelabelan dan Pembagian Dataset

Setelah fitur diekstrak, data diberi label sesuai kelasnya (normal atau pneumonia) dan dibagi menjadi dua bagian: data pelatihan (training) dan data pengujian (testing) dengan rasio tertentu, misalnya 80:20.

#### 6. Klasifikasi dengan KNN

Data hasil ekstraksi fitur kemudian digunakan untuk membangun model klasifikasi menggunakan algoritma K-Nearest Neighbor (KNN). Nilai k akan ditentukan melalui eksperimen, dan model dilatih untuk mengenali pola-pola dari data pelatihan guna mengklasifikasikan data uji.

#### 7. Evaluasi Model

Kinerja model KNN diuji menggunakan data testing dengan menghitung metrik evaluasi seperti akurasi, presisi, recall, dan F1-score. Evaluasi ini digunakan untuk mengetahui seberapa efektif model dalam mengklasifikasikan pneumonia.

#### 8. Kesimpulan dan Saran

Tahap akhir adalah menyimpulkan hasil penelitian berdasarkan kinerja model, serta memberikan saran untuk pengembangan penelitian di masa depan. Saran dapat mencakup penggunaan metode klasifikasi lain, peningkatan jumlah data, atau eksplorasi fitur tambahan.



### 3.3 Alat dan Bahan

#### 3.3.1 Data Penelitian

Data yang digunakan dalam penelitian ini merupakan kumpulan citra X-ray dada yang digunakan untuk mendeteksi penyakit pneumonia. Dataset diperoleh dari situs penyedia data ilmiah terbuka, yaitu Kaggle, dengan nama dataset “Chest X-ray Dataset (Pneumonia)”. Dataset ini disusun oleh Assem ElQersh dan berisi dua kategori citra, yaitu Normal dan Pneumonia.

Sebelum dilakukan pemrosesan, citra X-ray terlebih dahulu diubah ke dalam format grayscale dan di-resize untuk menyeragamkan dimensi



guna memudahkan proses ekstraksi fitur GLCM dan klasifikasi menggunakan algoritma KNN.

### **3.3.2 Alat/Instrumen Penelitian**

Penelitian ini dilakukan dengan bantuan beberapa perangkat keras dan perangkat lunak yang digunakan untuk pengolahan citra dan penerapan algoritma pembelajaran mesin. Berikut adalah rincian alat dan instrumen yang digunakan:

#### **a. Perangkat Keras**

MacBook Pro (13-inch, M1, 2020)

Prosesor: Apple M1 Chip (8-core CPU, 8-core GPU)

RAM: 8 GB Unified Memory

Storage: 256 GB SSD

Sistem Operasi: macOS Sonoma

#### **b. Perangkat Lunak dan Pustaka**

Python: Bahasa pemrograman utama yang digunakan untuk seluruh implementasi algoritma.

Google Colab: Digunakan sebagai alternatif cloud computing dengan akses ke GPU untuk mempercepat proses pengolahan data saat dibutuhkan.

scikit-learn: Untuk implementasi algoritma K-Nearest Neighbor (KNN) dan evaluasi performa klasifikasi.

NumPy dan Pandas: Untuk manipulasi data dan komputasi numerik.

matplotlib dan seaborn: Untuk visualisasi hasil dan metrik evaluasi.

scikit-image: Digunakan untuk ekstraksi fitur menggunakan metode Gray Level Co-occurrence Matrix (GLCM).

Penggunaan kombinasi perangkat keras dan perangkat lunak tersebut memungkinkan penelitian ini dilakukan secara efisien dan sistematis dengan hasil yang dapat direproduksi.

## **BAB IV**

### **HASIL DAN PEMBAHASAN**

Pembahasan dalam bab ini akan berfokus pada interpretasi hasil yang diperoleh dari implementasi algoritma K-Nearest Neighbor (KNN) dalam mengklasifikasikan citra X-ray dada dengan fitur tekstur yang diekstraksi menggunakan metode Gray Level Co-occurrence Matrix (GLCM). Selain itu, hasil yang diperoleh akan dibandingkan dengan penelitian terdahulu yang relevan untuk memberikan konteks yang lebih luas terhadap temuan penelitian ini.

Seluruh pembahasan diarahkan untuk secara komprehensif menjawab rumusan dan tujuan penelitian yang telah dirumuskan pada Bab I, termasuk menilai efektivitas kombinasi metode GLCM dan KNN dalam mendeteksi pneumonia, mengevaluasi kontribusi masing-masing fitur GLCM terhadap akurasi klasifikasi, serta mengidentifikasi potensi implementasi metode ini dalam pengembangan sistem bantu diagnosis medis berbasis citra.

#### **4.1 Hasil Implementasi**

Penelitian ini telah berhasil mengimplementasikan klasifikasi pneumonia pada citra X-ray dada menggunakan kombinasi metode ekstraksi fitur tekstur Gray Level Co-occurrence Matrix (GLCM) dan algoritma klasifikasi K-Nearest Neighbor (KNN). Proses implementasi dilakukan pada lingkungan Google Colab dengan bantuan pustaka seperti scikit-learn, scikit-image, NumPy, dan matplotlib.

Dataset yang digunakan terdiri dari dua kelas: Normal dan Pneumonia. Seluruh citra diproses dengan konversi ke grayscale dan diubah ukurannya menjadi seragam. Ekstraksi fitur menggunakan GLCM dilakukan untuk mendapatkan nilai contrast, homogeneity, energy, dan correlation.

Model KNN dilatih dengan variasi nilai k untuk mencari performa terbaik. Pengujian dilakukan dengan membagi dataset menjadi 80% data latih dan 20% data uji. Hasil metrik evaluasi model terbaik ditampilkan pada gambar berikut:

Confusion Matrix (Validation Set):				
[[152 6]				
[137 290]]				
Classification Report (Validation Set):				
	precision	recall	f1-score	support
NORMAL	0.53	0.96	0.68	158
PNEUMONIA	0.98	0.68	0.80	427
accuracy			0.76	585
macro avg	0.75	0.82	0.74	585
weighted avg	0.86	0.76	0.77	585

## 4.2 Analisis Hasil

Dari hasil evaluasi, model klasifikasi menghasilkan akurasi sebesar 76% pada data validasi. Nilai ini menunjukkan bahwa model cukup efektif dalam membedakan antara citra normal dan pneumonia, meskipun masih terdapat tantangan dalam mengklasifikasikan kedua kelas secara seimbang.

Kelas Normal memiliki recall sangat tinggi (96%), yang berarti hampir semua citra normal berhasil dikenali. Namun, nilai precision-nya rendah (0.53), menunjukkan bahwa banyak citra pneumonia yang keliru diklasifikasikan sebagai normal (false positive rendah, false negative tinggi).

Sebaliknya, kelas Pneumonia menunjukkan precision sangat tinggi (0.98), yang berarti prediksi pneumonia oleh model sangat akurat. Namun, recall-nya hanya 68%, mengindikasikan bahwa masih banyak kasus pneumonia yang tidak terdeteksi (false negative tinggi).

Secara keseluruhan, model cenderung lebih berhati-hati dalam menyatakan pneumonia, yang bisa menjadi strategi aman secara klinis, namun harus diperhatikan agar tidak mengabaikan kasus sebenarnya. Nilai macro average f1-score sebesar 0.74 mengindikasikan keseimbangan performa antara dua kelas masih perlu ditingkatkan.

## 4.2 Kontribusi Fitur GLCM

Ekstraksi fitur tekstur menggunakan metode Gray Level Co-occurrence Matrix (GLCM) menghasilkan sejumlah parameter statistik yang mewakili karakteristik tekstur dari citra X-ray dada. Dalam penelitian ini, empat fitur utama yang diekstraksi dari masing-masing citra adalah:

- Contrast
- Homogeneity
- Energi
- Correlation

Analisis kontribusi masing-masing fitur dilakukan dengan mengamati korelasinya terhadap keberhasilan klasifikasi, baik secara individual maupun sebagai kombinasi. Hasil pengamatan menunjukkan bahwa fitur contrast dan homogeneity memiliki peran paling signifikan dalam membedakan citra pneumonia dari citra normal.

### 1. Contrast

Contrast mencerminkan tingkat variasi intensitas antara piksel yang berdekatan. Pada citra X-ray penderita pneumonia, jaringan paru yang mengalami inflamasi atau infeksi cenderung menampilkan tekstur yang lebih kasar dan kontras tinggi. Hal ini membuat nilai contrast cenderung lebih tinggi dibandingkan citra normal. Model menunjukkan peningkatan akurasi klasifikasi saat fitur contrast disertakan dalam vektor fitur.

### 2. Homogeneity

Homogeneity mengukur tingkat kesamaan nilai intensitas antara piksel-piksel yang berdekatan. Citra normal memiliki pola distribusi yang lebih seragam sehingga menghasilkan nilai homogeneity yang lebih tinggi. Sebaliknya, ketidakaturan pola pada paru yang terinfeksi menyebabkan nilai ini menurun.

### 3. Energy

Energy merepresentasikan keteraturan pola tekstur. Nilai energy yang tinggi mengindikasikan citra dengan struktur yang lebih berulang dan stabil. Citra paru normal cenderung menghasilkan nilai energy yang lebih tinggi. Namun, kontribusi fitur ini terhadap akurasi klasifikasi relatif moderat dibanding contrast dan homogeneity.

### 4. Correlation

Correlation mengukur hubungan linier antara nilai intensitas piksel yang berdekatan. Nilai ini dapat menggambarkan pola struktural dalam jaringan paru-paru. Meski memberikan informasi tambahan, fitur ini cenderung memiliki kontribusi yang lebih rendah terhadap klasifikasi dibandingkan contrast dan homogeneity, karena pola linier tidak selalu konsisten pada semua kondisi pneumonia.

Berdasarkan eksperimen, model dengan kombinasi fitur contrast + homogeneity memberikan hasil performa terbaik dibanding kombinasi lain. Hal ini menunjukkan bahwa perbedaan tekstur kasar dan tingkat ketidakseragaman struktur jaringan paru merupakan indikator utama dalam mendeteksi pneumonia melalui citra X-ray.

## DAFTAR PUSTAKA

- [1] K. K. R. Indonesia, “Potret Sehat Indonesia: Riskesdas 2018.”
- [2] S. Suharyana, F. Anwar, A. C. Dewi, M. Yunianto, U. Salamah, and R. Chai, “Pneumonia Classification Based on GLCM Features Extraction using K-Nearest Neighbor,” *Indones. J. Appl. Phys.*, vol. 13, no. 2, p. 325, 2023, doi: 10.13057/ijap.v13i2.77120.
- [3] N. F. Hismaya and N. Sulistianingsih, “Analisis Tekstur Berbasis GLCM dan Klasifikasi KNN-SVM untuk Identifikasi Pneumonia pada Citra Rontgen,” vol. 1, no. 1, pp. 68–78, 2025.
- [4] W. H. Organization, “Pneumonia.”
- [5] and B. I. National Heart, Lung, “How Is Pneumonia Diagnosed?” Accessed: May 22, 2025. [Online]. Available: <https://www.nhlbi.nih.gov/health/pneumonia/diagnosis>
- [6] S. Singh, M. Kumar, A. Kumar, B. K. Verma, K. Abhishek, and S. Selvarajan, “Efficient pneumonia detection using Vision Transformers on chest X-rays,” *Sci. Rep.*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-52703-2.
- [7] S. Sharma and K. Guleria, “A Deep Learning based model for the Detection of Pneumonia from Chest X-Ray Images using VGG-16 and Neural Networks,” in *Procedia Computer Science*, Elsevier B.V., 2022, pp. 357–366. doi: 10.1016/j.procs.2023.01.018.
- [8] M. Shakeel, A. Naeem, N. Aslam, K. Abid, and M. Q. Shafiq, “Pneumonia Recognition in Chest X-rays through Convolutional Neural Networks (CNN).”
- [9] I. Bakti, M. Firdaus, and S. Artikel, “Arsitektur Convolutional Neural Network InceptionResNet-V2 Untuk Pengelompokan Pneumonia Chest X-Ray INFO ARTIKEL ABSTRAK,” *Drh. Khusus Ibuk. Jakarta*, vol. 12530, no. 77, doi: 10.58290/jukomtek.
- [10] Q. An, W. Chen, and W. Shao, “A Deep Convolutional Neural Network for Pneumonia Detection in X-ray Images with Attention Ensemble,” *Diagnostics*, vol. 14, no. 4, Feb. 2024, doi: 10.3390/diagnostics14040390.
- [11] G. L. E. Maquen-Niño, J. G. Nuñez-Fernandez, F. Y. Taquila-Calderon, I. Adrianzén-Olano, P. De-La-cruz-vdv, and G. Carrión-Barco, “Classification Model Using Transfer Learning for the Detection of

Pneumonia in Chest X-Ray Images,” *Int. J. online Biomed. Eng.*, vol. 20, no. 5, pp. 150–161, Mar. 2024, doi: 10.3991/ijoe.v20i05.45277.

- [12] I. Kurniawati, R. Akbar, Y. Ardina Kusuma, and I. Fahma Kusumawati, “Identifikasi Pneumonia pada Balita melalui Citra X-ray Menggunakan Metode Convolutional Neural Network (CNN),” *J. Manuf. Ind. Eng. Technol.*, vol. 3, no. 1, pp. 10–23, Jun. 2024, doi: 10.30651/mine-tech.v3i1.23391.