

IMPLEMENTASI KNN DENGAN FITUR EKSTRAKSI HARALICK DAN LOCAL BINARY PATTERN UNTUK KLASIFIKASI PENYAKIT PNEUMONIA PADA CITRA X-RAY

Rafi Anggara Nindiaputra, Yisti Vita Via , Hendra Maulana
Informatika, Universitas Pembangunan Nasional "Veteran" Jawa Timur
Jl. Raya Rungkut Madya, Gunung Anyar, Surabaya, Indonesia
yistivia.if@upnjatim.ac.id

ABSTRAK

Pneumonia adalah infeksi paru-paru yang dapat mengancam nyawa jika tidak terdeteksi pada tahap awal. Diagnosis tradisional oleh tenaga medis memerlukan waktu dan keterampilan khusus, sehingga sistem berbasis gambar digital diperlukan untuk mendukung identifikasi penyakit secara otomatis. Studi ini bertujuan untuk mengklasifikasikan gambar rontgen paru-paru guna mendeteksi pneumonia menggunakan algoritma K-Nearest Neighbor (KNN), dilengkapi dengan ekstraksi fitur melalui metode Haralick dan Local Binary Pattern (LBP). Tahapan penelitian meliputi prapemrosesan gambar, seperti mengubah ukuran menjadi 256×256 piksel, mengurangi noise, dan menormalisasi intensitas piksel. Selanjutnya, ekstraksi fitur dilakukan menggunakan metode Haralick yang dihasilkan dari Gray Level Co-Occurrence Matrix (GLCM) dan metode LBP yang mengekstrak pola tekstur lokal. Fitur yang diekstraksi kemudian dimasukkan ke dalam algoritma KNN untuk klasifikasi. Hasil pengujian menunjukkan bahwa kombinasi fitur Haralick dan LBP dengan algoritma KNN mencapai akurasi 80%, sementara penggunaan fitur Haralick saja menghasilkan akurasi 60%, dan fitur LBP saja memberikan akurasi 70%. Dari hasil ini, dapat disimpulkan bahwa kombinasi metode ekstraksi fitur Haralick dan LBP dapat meningkatkan kinerja klasifikasi gambar rontgen pneumonia menggunakan algoritma KNN, dibandingkan dengan penerapan salah satu metode saja.

Kata kunci : *Pneumonia, Citra X-Ray, Haralick, Local Binary Pattern, K-Nearest Neighbor*

1. PENDAHULUAN

Paru-paru memainkan peran penting dalam sistem pernapasan manusia, yang bertanggung jawab atas pertukaran oksigen dan karbon dioksida. Faktor lingkungan yang tidak sehat dapat memicu berbagai masalah kesehatan pada organ ini, termasuk pneumonia. Pneumonia, yang juga dikenal sebagai paru-paru basah, adalah infeksi pada alveoli paru-paru yang disebabkan oleh bakteri, virus, atau jamur, dan dapat menyerang satu atau kedua paru-paru secara bersamaan.

Pneumonia biasanya dideteksi melalui rontgen dada, yang menampilkan struktur paru-paru untuk membantu tenaga medis mengidentifikasi infeksi. Namun, proses diagnosis manual memiliki keterbatasan dan memakan waktu, sehingga memerlukan sistem berbasis komputer untuk mendukung klasifikasi gambar secara otomatis.

Penerapan teknik pembelajaran mesin dan pemrosesan gambar menawarkan solusi efisien untuk mendeteksi pneumonia melalui gambar rontgen. Salah satu algoritma klasifikasi yang populer adalah K-Nearest Neighbor (KNN), berkat kesederhanaannya dan efektivitasnya dalam mengenali pola. Efisiensi KNN sangat dipengaruhi oleh kualitas fitur yang diekstraksi dari gambar.

Studi ini menerapkan dua pendekatan ekstraksi fitur berbasis tekstur, yaitu Haralick dan Local Binary Pattern (LBP), untuk memperoleh informasi tekstur yang akurat dari gambar rontgen. Kombinasi kedua metode ini diharapkan dapat meningkatkan tingkat akurasi dalam klasifikasi penyakit.

Masalah utama dalam penelitian ini adalah klasifikasi pneumonia dari gambar X-ray yang belum optimal, karena metode ekstraksi fitur tekstur paru-paru masih terbatas. Fitur penting seperti pola co-occurrence dan variasi lokal belum dimanfaatkan sepenuhnya, sehingga akurasi klasifikasi tetap rendah. Algoritma K-Nearest Neighbors (KNN) untuk citra medis juga memerlukan desain fitur yang baik dan optimasi parameter seperti nilai k serta metrik jarak. Tanpa ekstraksi fitur yang efektif dan penyesuaian parameter, performa KNN sering tidak stabil dan kurang akurat. Oleh karena itu, penelitian ini akan menerapkan ekstraksi fitur Haralick dan Local Binary Pattern, serta mengoptimalkan parameter KNN untuk meningkatkan akurasi klasifikasi pneumonia dari X-ray.

Tujuan dari penelitian yaitu mengimplementasikan algoritma K-Nearest Neighbors (KNN) sebagai metode klasifikasi berdasarkan fitur yang diekstraksi, membandingkan performa klasifikasi antara penggunaan fitur Haralick saja, LBP saja, dan kombinasi Haralick + LBP untuk mengetahui fitur mana yang memberikan hasil terbaik, mengevaluasi performa sistem klasifikasi menggunakan metrik akurasi, precision, recall, F1-score, dan confusion matrix untuk menentukan efektivitas model yang diusulkan.

Rencana penyelesaian masalah dengan cara Ekstraksi fitur (Haralick dan LBP) dari citra X-ray, Implementasi KNN sebagai model klasifikasi, Optimasi parameter KNN, Evaluasi dan analisis hasil.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Penelitian mengenai deteksi penyakit menggunakan algoritma Transfer Learning dengan model ResNet50V2 yang sudah pernah dilakukan sebelumnya oleh [9] dengan judul “EFEKTIVITAS TRANSFER LEARNING DALAM PENDETEKSIAN PENYAKIT PNEUMONIA MELALUI CITRA X-RAY PARU MANUSIA” yang mengkaji penerapan metode transfer learning dengan model ResNet50V2 untuk klasifikasi pneumonia berdasarkan citra X-ray dada. Dataset yang digunakan terdiri dari 5863 citra, kemudian dilakukan augmentasi data untuk memperkaya variasi pelatihan. Model dilatih selama 30 epoch dengan evaluasi menggunakan metrik akurasi dan loss. Hasil penelitian menunjukkan bahwa model mampu mencapai akurasi maksimal sebesar 99,12% dengan nilai loss sebesar 0,0299. Penelitian ini menyimpulkan bahwa penggunaan ResNet50V2 efektif dalam mendeteksi pneumonia secara cepat dan akurat, serta menyarankan untuk melakukan perbandingan dengan arsitektur lain dan memastikan koneksi internet stabil saat menggunakan platform Google Colaboratory.

Pada penelitian ini oleh [10] berjudul “Improvement on KNN using genetic algorithm and combined feature extraction to identify COVID-19 sufferers based on CT scan image” yang mengkaji penerapan algoritma K-Nearest Neighbor (KNN) yang dikombinasikan dengan Algoritma Genetika (GA) untuk seleksi fitur dalam mendeteksi COVID-19 dari citra CT scan. Dataset terdiri dari 746 citra grayscale CT scan yang terbagi antara pasien positif dan negatif COVID-19. Fitur tekstur yang diekstraksi meliputi Haralick, histogram 32-bin, dan Local Binary Pattern (LBP), dengan berbagai kombinasi fitur diuji. Seleksi fitur dilakukan menggunakan GA untuk mengurangi dimensi dan memilih fitur paling relevan. Klasifikasi menggunakan KNN dengan nilai k bervariasi dari 2 hingga 17 dan validasi silang 10-fold. Evaluasi menggunakan akurasi dan Area Under Curve (AUC) menunjukkan hasil terbaik dengan kombinasi fitur Haralick dan LBP, mencapai akurasi 93,30% dan AUC 0,948. Metode ini juga menunjukkan keunggulan dibandingkan model CNN dengan peningkatan akurasi sekitar 4,3%. Penelitian ini menyimpulkan bahwa kombinasi ekstraksi fitur sederhana, seleksi fitur berbasis GA, dan KNN memberikan metode deteksi COVID-19 yang efisien dan akurat, menjadi alternatif yang menjanjikan dibandingkan model deep learning yang membutuhkan sumber daya lebih besar.

2.2. Chest X-Ray

Citra rontgen dada merupakan jenis pencitraan medis yang diperoleh melalui prosedur radiologi, yang memungkinkan pengamatan proyeksi struktur rongga dada dengan jelas. Proses ini krusial dalam evaluasi kondisi paru-paru, khususnya untuk mengidentifikasi penyakit seperti pneumonia berdasarkan citra yang dihasilkan. Dibandingkan dengan teknik pencitraan

lain seperti CT-Scan dan MRI, pemeriksaan X-ray umumnya lebih efisien, sederhana, dan ekonomis [5]. Namun, prosedur X-ray memiliki beberapa kekurangan, seperti kemungkinan hasil citra yang saling tumpang tindih atau overlap, sebagaimana terlihat pada Gambar 1. Fenomena ini terjadi ketika organ paru-paru dan tulang rusuk saling menutupi, sehingga bagian-bagian paru-paru tidak terlihat optimal akibat bayangan tulang rusuk (yang radiopak) menutupi jaringan paru (yang radiolusen) [1].

2.3. Pneumonia

Pneumonia adalah infeksi pada sistem pernapasan yang menyebabkan peradangan pada alveoli dan saluran napas bawah. Penyakit ini menyebabkan masalah kesehatan serius dengan tingkat morbiditas dan mortalitas yang tinggi, terutama pada kelompok rentan seperti balita, bayi, dan lansia. Pneumonia umumnya disebabkan oleh berbagai mikroorganisme, termasuk bakteri, virus, dan jamur. Patogen utama pneumonia yang didapat di masyarakat meliputi *Streptococcus pneumoniae*, *Haemophilus influenzae*, dan *Mycoplasma pneumoniae*, sementara pneumonia nosokomial sering dikaitkan dengan bakteri yang resisten terhadap obat seperti *Pseudomonas aeruginosa* dan *Staphylococcus aureus*. Faktor risiko utama pneumonia meliputi usia ekstrem (sangat muda atau tua), sistem kekebalan yang lemah, dan penyakit kronis seperti diabetes dan penyakit paru obstruktif kronik (PPOK) [2].

2.4. Pengolahan Citra

Pengolahan citra digital adalah cabang ilmu yang memfokuskan pada teknik-teknik untuk mengolah dan memanipulasi citra. Berbagai operasi yang dilakukan dalam pengolahan citra mencakup peningkatan kualitas citra (image enhancement), restorasi citra (image restoration), kompresi citra (image compression), segmentasi citra (image segmentation), dan analisis citra (image analysis). Tujuan utama dari pengolahan citra digital adalah untuk memperbaiki kualitas citra agar lebih mudah dipahami oleh manusia dan mesin, serta untuk mengekstrak informasi yang berguna dari citra tersebut [3].

2.5. Preprocessing

Preprocessing adalah langkah pertama dalam memproses gambar mentah sebelum ekstraksi fitur untuk pelatihan atau klasifikasi. Pada tahap ini, sejumlah operasi dasar diterapkan untuk mempersiapkan gambar agar siap dianalisis lebih lanjut. Langkah awal umumnya adalah pengubahan ukuran gambar, yang bertujuan untuk mengubah dimensi gambar asli menjadi ukuran yang sesuai guna memudahkan pemrosesan selanjutnya. Selanjutnya, gambar diubah menjadi gambar grayscale, di mana gambar asli dalam format RGB dikonversi menjadi gambar dengan saluran warna tunggal yang menunjukkan tingkat keabu-abuan. Terakhir, normalisasi diterapkan, yang mengubah rentang nilai

piksel dari 0-255 menjadi 0-1. Normalisasi ini bertujuan untuk menyelaraskan skala intensitas, sehingga memudahkan klasifikasi dan analisis lebih lanjut, serta meningkatkan akurasi model yang diterapkan [4].

2.6. Haralick

Ekstraksi fitur Haralick merupakan metode dalam pemrosesan citra yang bertujuan menganalisis serta merepresentasikan tekstur gambar dengan menggunakan pendekatan statistik. Metode ini memanfaatkan Gray Level Co-occurrence Matrix (GLCM), yakni matriks yang menghitung frekuensi kemunculan pasangan nilai keabuan piksel pada jarak dan arah tertentu dalam citra. Melalui pendekatan ini, sifat pola dan sebaran tekstur pada gambar dapat diekstrak serta digambarkan secara kuantitatif, sehingga detail tekstur yang sering kali tidak terlihat secara visual dapat diidentifikasi dengan lebih tepat [5].

2.7. Local Binary Pattern

Local Binary Pattern (LBP) pertama kali diperkenalkan oleh Ojala et al. sebagai metode untuk menganalisis tekstur dalam citra. LBP digunakan untuk menggambarkan pola lokal dalam citra dengan cara membandingkan nilai intensitas piksel pusat terhadap piksel-piksel tetangganya. Metode ini efektif dalam merepresentasikan karakteristik tekstur, memiliki kemampuan diskriminatif yang tinggi, dan cukup tahan terhadap perubahan intensitas keabuan yang bersifat monotonik. Proses LBP terdiri dari dua tahap utama, yaitu thresholding dan encoding. Pada tahap thresholding, intensitas piksel tetangga dibandingkan dengan nilai piksel pusat. Jika nilai piksel tetangga lebih besar atau sama dengan nilai piksel pusat, maka diberi nilai 1; jika lebih kecil, diberi nilai 0. Hasil dari proses ini adalah pola biner yang menggambarkan tekstur lokal di sekitar piksel pusat [6].

2.8. K-Nearest Neighbor

K-Nearest Neighbor (KNN) adalah salah satu algoritma dalam pembelajaran mesin yang digunakan untuk klasifikasi dan regresi. Metode ini bekerja dengan cara mengklasifikasikan data baru berdasarkan kedekatannya dengan data yang sudah ada di dalam dataset. KNN adalah algoritma yang sangat sederhana namun efektif, yang beroperasi berdasarkan prinsip "belajar dari tetangga terdekat". Ketika sebuah data baru ingin diklasifikasikan, KNN akan mencari k-tetangga terdekat dalam ruang fitur dan mengklasifikasikan data tersebut ke dalam kelas mayoritas dari tetangga tersebut. Dengan kata lain, kelas dari data baru ditentukan oleh kelas yang paling sering muncul di antara k-tetangga terdekat [7].

2.9. Confusion Matrix

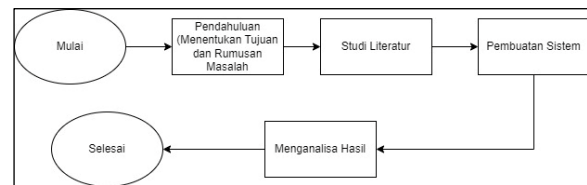
Confusion Matrix adalah tabel yang digunakan untuk menilai kinerja model klasifikasi. Perbandingan

antara nilai aktual (benar) dari data dan hasil prediksi model disajikan dalam tabel ini. Confusion Matrix menawarkan tinjauan menyeluruh tentang kinerja klasifikasi model termasuk kategori kesalahan yang dihasilkannya. Confusion Matrix adalah tabel berdimensi $N \times N$ (di mana N adalah jumlah kelas target) yang menggabungkan kinerja model klasifikasi pada data uji. Kelas aktual diwakili oleh setiap baris matriks, sedangkan kelas prediksi diwakili oleh setiap kolom untuk model. Hal ini memungkinkan visualisasi yang jelas dari akurasi keseluruhan model serta pola kesalahan yang dibuatnya (misalnya, kelas mana yang paling sering membingungkan) [8].

3. METODE PENELITIAN

3.1. Tahapan Metode

Penelitian ini dilakukan secara sistematis dengan mengikuti urutan yang telah direncanakan berdasarkan hasil studi literatur dari beberapa publikasi ilmiah terkait topik yang sama.



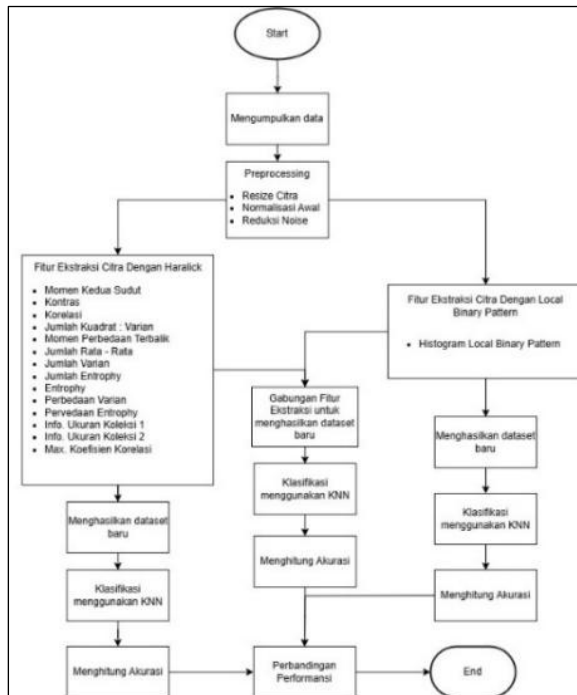
Gambar 1. Kerangka Kerja Penelitian

Pada Gambar 1 ini terdapat beberapa aktivitas yang bertujuan untuk mempersiapkan penelitian yang akan dilakukan agar proses penelitian dapat tersusun secara sistematis dan juga menghasilkan penelitian yang berkualitas, langkah-langkah yang dapat dilakukan di tahapan pendahuluan ini adalah Menentukan Tujuan, Merumuskan Masalah, Studi Literatur, Pembuatan Sistem, dan Menganalisa Hasil.

3.2. Desain Sistem

Pada Gambar 2 tahap awal, proses dimulai dengan pengumpulan data berupa citra X-Ray berformat JPEG/JPG. Selanjutnya, dilakukan preprocessing data melalui beberapa metode. Pertama, resize citra dilakukan untuk memastikan seluruh citra memiliki ukuran yang seragam sehingga konsistensi dalam pemrosesan dapat tercapai. Kedua, normalisasi awal diterapkan untuk meningkatkan konsistensi antar citra yang memiliki karakteristik berbeda. Ketiga, reduksi noise dilaksanakan untuk menghilangkan atau meminimalkan gangguan pada citra tersebut. Setelah preprocessing, tahap ekstraksi fitur dilakukan dengan dua pendekatan ekstraksi fitur Haralick dan ekstraksi fitur Local Binary Pattern (LBP). Ekstraksi fitur Haralick bertujuan untuk mengkuantifikasi karakteristik citra dan meningkatkan kinerja klasifikasi. Sementara itu, ekstraksi fitur LBP berfungsi mengekstrak pola tekstur lokal dengan cara membandingkan setiap piksel dengan piksel tetangganya, kemudian menghitung histogram LBP yang merepresentasikan distribusi pola tekstur dalam

citra. Pada tahap klasifikasi, tiga pendekatan diterapkan secara terpisah. Pertama, hasil ekstraksi fitur Haralick diklasifikasikan menggunakan algoritma KNN. Kedua, hasil ekstraksi fitur LBP juga diklasifikasikan dengan algoritma KNN. Ketiga, kombinasi kedua ekstraksi fitur tersebut diklasifikasikan dengan algoritma KNN. Pada tahap akhir, hasil dari ketiga metode klasifikasi tersebut dibandingkan untuk menentukan pendekatan mana yang menghasilkan tingkat akurasi tertinggi.



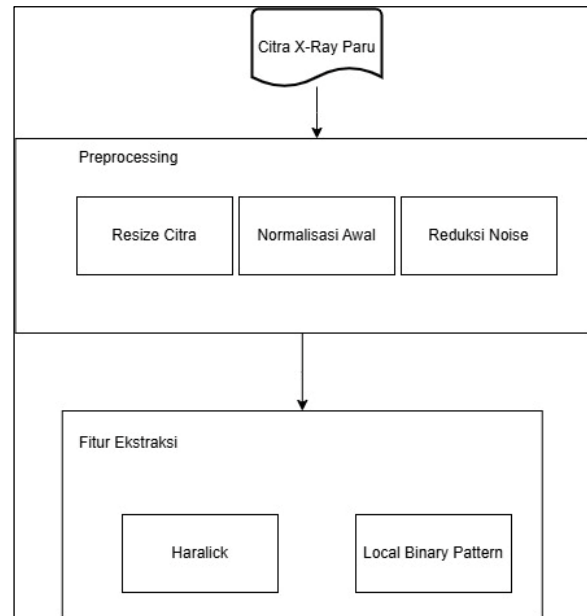
Gambar 2. Alur Desain Sistem

3.3. Pengumpulan Data

Pada tahap pengumpulan data pada bulan Januari – Juni 2021, didapatkan dari RS Husada Utama Surabaya yang relevan dengan penelitian ini. Dataset yang dikumpulkan merupakan citra X-Ray Paru terdiri dari dua label, masing-masing label mempresentasikan kategori citra berupa pneumonia dan tidak ada pneumonia. Pada penelitian ini, Dataset ini berjumlah 100 citra X-Ray Paru, dimana 50 citra X-Ray paru terdiagnosis pneumonia dan lainnya 50 citra X-Ray tidak terdiagnosis pneumonia.

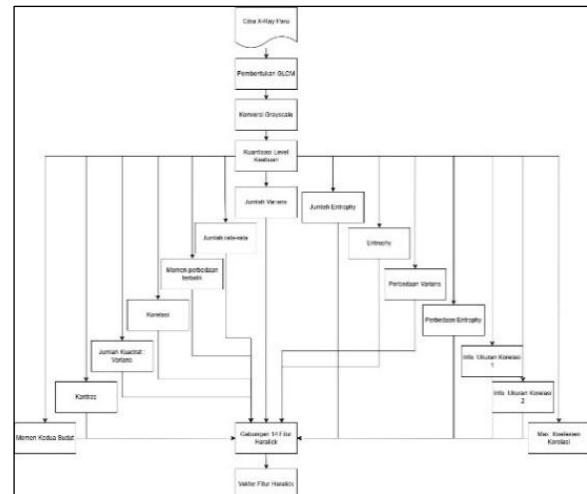
3.4. Preprocessing Data

Pada Gambar 3 tahap ini sebelum melakukan tahap klasifikasi pada citra, seluruh data citra harus melewati tahapan pre-processing terlebih dahulu. Pre-processing citra bertujuan untuk meningkatkan kualitas gambar dan menyederhanakan langkah proses selanjutnya dan dapat meningkatkan tingkat akurasi. Dalam penelitian ini penulis akan melakukan pre-processing dengan metode Resizing, Normalisasi Awal dan Reduksi Noise. Tahapan untuk pre-processing dapat dilihat pada Gambar 3.



Gambar 3. Tahapan Preprocessing

3.5. Haralick



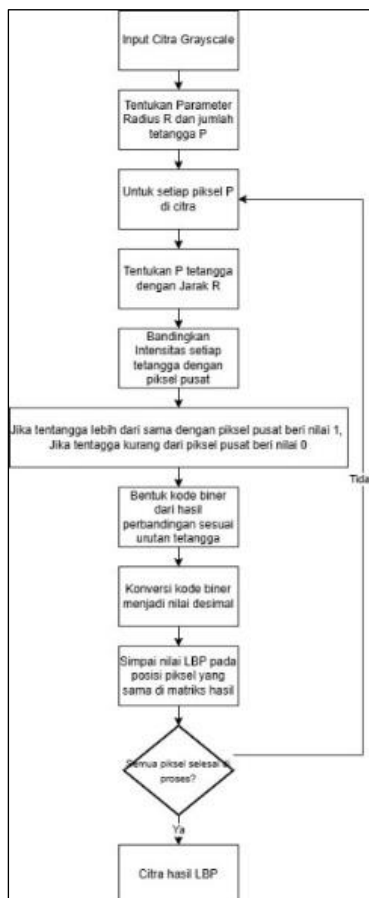
Gambar 4. Alur Kerja Haralick

Pada Gambar 4 Ekstraksi fitur metode Haralick ini berisi informasi tentang intensitas gambar dalam piksel dengan posisi tertentu dalam hubungannya satu sama lain yang terjadi secara bersamaan. Metode Haralick ini menghitung nilai fiturnya dari 4 sudut yaitu 0, 45, 90, 135. Dengan demikian, setiap gambar akan menghasilkan fitur. Kemudian dilakukan penyederhanaan dengan menghitung nilai rata-rata setiap fitur sudut pada setiap citra, sehingga jumlah fitur untuk setiap citra menjadi 14 fitur.

3.6. Local Binary Pattern

Pada Gambar 5 tahap pertama yang dilakukan adalah menentukan piksel pusat dan tetangganya. Dalam penelitian ini, metode LBP bekerja dengan membandingkan intensitas piksel pusat dengan intensitas piksel-piksel di sekitarnya. Pada kasus ini, dengan parameter $P=8$ dan $R=1$, setiap piksel pusat

akan dibandingkan dengan delapan piksel yang berada dalam radius satu piksel di sekelilingnya.



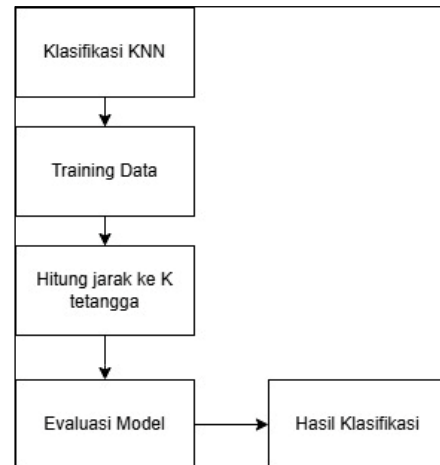
Gambar 5. Alur Kerja Local Binary Pattern

3.7. Penggabungan Ekstraksi Fitur (Hybrid)

Pada tahap penggabungan fitur, penulis mengkombinasikan vektor dari fitur Haralick dan Local Binary Pattern (LBP) menjadi satu vektor panjang. Pendekatan ini dipilih karena mampu mempertahankan seluruh informasi yang diperoleh dari kedua metode ekstraksi fitur tersebut. Pada tahap penggabungan fitur, akan dikombinasikan vektor dari fitur Haralick dan Local Binary Pattern (LBP) menjadi satu vektor panjang.

3.8. Klasifikasi K-Nearest Neighbor

Pada Gambar 6 implementasi algoritma KNN, tahapan pertama adalah persiapan data training, di mana peneliti mengumpulkan dan menyiapkan dataset yang akan digunakan. Selanjutnya, proses perhitungan jarak ke K tetangga dilakukan dengan menghitung jarak antara data baru dengan semua data training, kemudian mengurutkannya dari terdekat hingga terjauh untuk memilih K data training terdekat sebagai tetangga.



Gambar 6. Tahapan Klasifikasi KNN

3.9. Evaluasi

Pada tahap ini yang akan dilakukan pada bab 4 yang melibatkan beberapa percobaan yang menggunakan algoritma KNN. Penelitian ini akan menguji beberapa nilai seperti akurasi, presisi, Recall, dan F1-Score. Hasil dari evaluasi tersebut kemudian dimasukkan ke dalam table.

3.10. Perbandingan Performance

Pada tahap penelitian ini bertujuan untuk membandingkan performa tiga metode klasifikasi citra menggunakan algoritma K-Nearest Neighbors (KNN) dengan pengujian data split 70:30, 80:20, 90:10. Untuk tiga metode tersebut adalah:

- Haralick + KNN: Menggunakan fitur Haralick untuk ekstraksi ciri citra, kemudian diklasifikasikan dengan KNN.
- LBP + KNN: Menggunakan fitur Local Binary Pattern (LBP) untuk ekstraksi ciri citra, kemudian diklasifikasikan dengan KNN.
- Haralick + LBP + KNN: Menggabungkan fitur Haralick dan LBP untuk ekstraksi ciri citra, kemudian diklasifikasikan dengan KNN.

4. HASIL DAN PEMBAHASAN

Pada bab ini akan dibahas mengenai penerapan dan pembuatan model KNN dalam melakukan klasifikasi pada citra X-ray pneumonia dengan menggunakan ekstraksi fitur Haralick dan Local Binary Pattern (LBP)

4.1. Pengumpulan Data

Dataset yang digunakan pada penelitian ini merupakan dataset primer yang diperoleh dari data rekam medis pasien. Dataset ini terdiri dari 50 citra x-ray paru yang terdiagnosis pneumonia dan 50 citra x-ray paru yang tidak terdiagnosis pneumonia. Setelah dataset diunduh, citra-citra tersebut diekstrak dan disimpan dalam folder yang telah disiapkan di Google Drive yang terhubung dengan Google Colab.

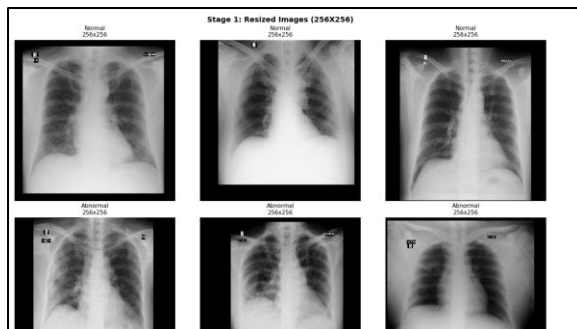
Normal: 50 images
Abnormal: 50 images
Total: 100 images

Gambar 7. Hasil Memuat Dataset

Output pada gambar 7 ini memberikan informasi awal mengenai distribusi gambar dalam dataset serta memastikan bahwa dataset telah dimuat dengan benar untuk langkah-langkah penelitian selanjutnya.

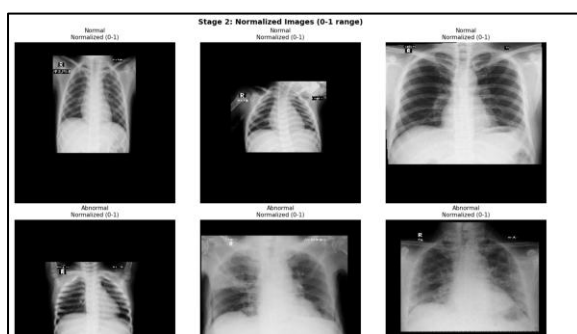
4.2. Preprocessing Data

Setelah menampilkan hasil dari dataset tersebut, tahap selanjutnya adalah preprocessing data yang merupakan langkah untuk mempersiapkan citra agar dapat di ekstrasi fiturnya dengan lebih optimal. Preprocessing bertujuan untuk menstandarisasi ukuran citra yang akan di gunakan dalam tahap ekstraksi dan klasifikasi setelah itu menormalisasikan citra tersebut dengan tahap terakhir yaitu tahap pengurangan noise pada citra.



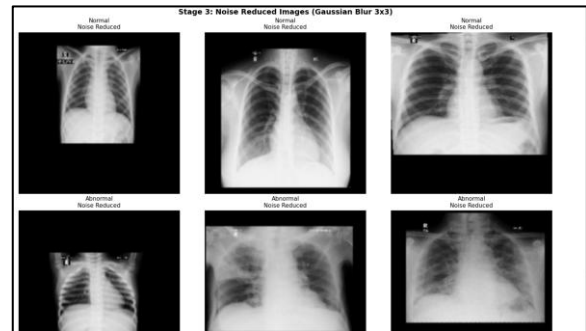
Gambar 8. Hasil Resizing Image

Hasil visualisasi ini seperti yang ditampilkan pada Gambar 8 memberikan gambaran visual tentang proses resizing dan memastikan gambar-gambar telah diubah ukurannya dengan benar menjadi 256x256 piksel.



Gambar 9. Hasil Normalisasi Image

Pada Gambar 9 yang ditampilkan memberikan gambaran visual tentang gambar yang telah dinormalisasi dalam rentang 0-1, dengan setiap gambar diberi label sesuai kategori dan ukuran normalisasi.

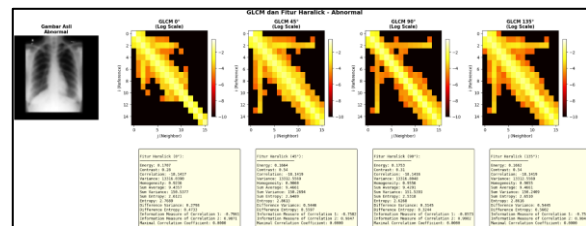


Gambar 10. Hasil Noise Reduction Image

Visualisasi ini memberikan gambaran mengenai hasil pengurangan noise pada gambar dengan kategori normal dan abnormal, seperti yang terlihat pada Gambar 10.

4.3. Ekstraksi Fitur Haralick

Setelah tahap preprocessing selesai, langkah berikutnya adalah ekstraksi fitur haralick sangat berguna untuk membedakan tekstur pada citra X-ray yang menunjukkan pneumonia dari citra normal. Fitur-fitur yang dihasilkan oleh Haralick akan memberikan informasi tekstural yang penting selanjutnya digunakan dalam proses klasifikasi menggunakan metode KNN untuk meningkatkan akurasi deteksi pneumonia pada citra X-ray.



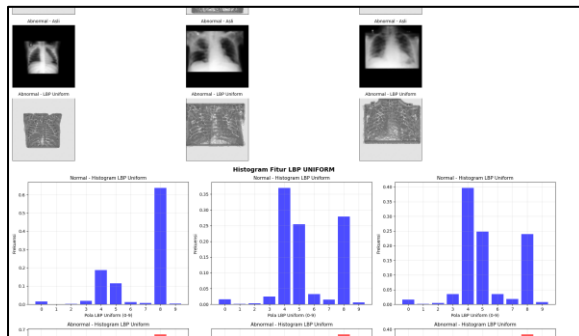
Gambar 11. Hasil Ekstraksi Haralick

Gambar 11 menunjukkan hasil ekstraksi fitur tekstur menggunakan metode Gray Level Co-occurrence Matrix (GLCM) pada dua citra satu dengan kategori Normal dan satu lagi dengan kategori Abnormal. Panel kiri pada setiap gambar menunjukkan citra asli dalam format grayscale, yang telah diproses untuk ekstraksi fitur tekstur. Di panel kanan, terdapat empat matriks GLCM yang dihitung pada empat sudut orientasi berbeda 0°, 45°, 90°, dan 135°, yang menunjukkan hubungan spasial antara piksel berdekatan pada citra.

4.4. Ekstraksi Fitur Local Binary Pattern

Tahap selanjutnya selain ekstraksi fitur Haralick penelitian ini juga menggunakan metode Local Binary Pattern (LBP) yang merupakan salah satu deskriptor tekstur yang efisien dan ringan dalam analisis citra medis, terutama untuk mendeteksi pola tekstur pada citra X-ray. LBP bekerja dengan cara membandingkan nilai intensitas piksel pusat dengan nilai intensitas piksel-piksel tetangga dalam area lokal tertentu. Jika intensitas piksel tetangga lebih besar atau sama dengan

piksel pusat, maka nilai biner 1 diberikan jika lebih kecil, diberikan nilai 0. Perbandingan ini menghasilkan representasi biner yang kemudian diubah menjadi nilai desimal yang mewakili identitas pola tekstur lokal di sekitar piksel tersebut.

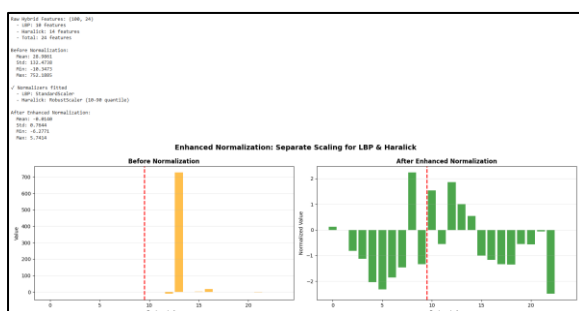


Gambar 12. Hasil Ekstraksi LBP

Pada Gambar 12 memperlihatkan hasil ekstraksi Local Binary Pattern (LBP) untuk citra normal dan abnormal dari dataset X-ray pneumonia. Panel atas menampilkan citra asli dan citra hasil transformasi LBP uniform untuk tiga sampel citra abnormal. Citra asli menunjukkan gambar X-ray yang telah diproses menjadi citra grayscale, sedangkan citra LBP uniform menggambarkan tekstur lokal berdasarkan pola biner yang dihasilkan oleh LBP. Di bagian bawah, terdapat histogram fitur LBP uniform untuk normal dan abnormal, yang menggambarkan distribusi frekuensi dari pola-pola LBP uniform (dari 0 hingga 9) pada masing-masing kelas citra. Pada citra normal, histogram menunjukkan bahwa sebagian besar pola LBP terpusat pada nilai 7, dengan frekuensi yang lebih tinggi di nilai ini, yang mencerminkan tekstur yang lebih seragam dan konsisten pada citra normal.

4.5. Penggabungan Ekstraksi Fitur (Hybrid)

Setelah mendapatkan fitur Haralick dan LBP secara terpisah, langkah selanjutnya adalah mengintegrasikan kedua jenis fitur tersebut untuk mendapatkan representasi yang lebih komprehensif dari citra. Penggabungan fitur Haralick dan LBP bertujuan untuk memaksimalkan kekuatan diskriminatif dari kedua metode dalam satu set fitur gabungan yang dapat meningkatkan performa model klasifikasi.



Gambar 13. Hybrid Haralick-LBP

Gambar 13 menggambarkan proses normalisasi fitur gabungan LBP (Local Binary Pattern) dan Haralick pada citra X-ray pneumonia menggunakan teknik normalisasi yang ditingkatkan. Panel kiri menunjukkan histogram dari fitur gabungan sebelum normalisasi, dengan nilai fitur yang sangat bervariasi.

4.6. Klasifikasi K-Nearest Neighbor

Pada tahap algoritma KNN, dataset yang digunakan merupakan hasil gabungan fitur LBP dan Haralick yang telah diekstraksi sebelumnya dan disimpan dalam format .csv. Setiap baris dalam dataset mewakili sebuah citra X-ray pneumonia yang berisi nilai-nilai numerik yang menggambarkan karakteristik tekstur dari citra tersebut. Setiap data dalam dataset telah dilabeli dengan kelas yang sesuai, yaitu normal atau abnormal, sehingga dapat digunakan sebagai referensi dalam proses klasifikasi. Algoritma KNN selanjutnya digunakan untuk menentukan kelas dari data uji berdasarkan kedekatannya dengan sejumlah tetangga terdekat dalam dataset latih, dengan jarak cosine digunakan sebagai ukuran kedekatan antar data. Kelas yang diprediksi untuk data uji ditentukan berdasarkan mayoritas kelas dari tetangga terdekat yang ada. Hasil evaluasi klasifikasi, seperti akurasi, digunakan untuk menilai performa model dalam mengenali dan mengidentifikasi kategori pneumonia berdasarkan fitur tekstur yang telah diekstraksi sebelumnya.

=====
HASIL EVALUASI
=====
Skenario: 1 (70/30)
K: 3
Train: 70, Test: 30
Accuracy : 0.7000
Precision : 0.8000
Recall : 0.5333
F1-Score : 0.6400
=====

Gambar 14. Hasil Klasifikasi KNN

Gambar 14 menunjukkan hasil evaluasi dari eksperimen K-Nearest Neighbors (KNN) pada dataset citra dengan pembagian data 70/30 (70% untuk pelatihan dan 30% untuk pengujian) serta nilai K=3. Evaluasi dilakukan menggunakan empat metrik utama: Akurasi (Accuracy) sebesar 70 yang menunjukkan bahwa model berhasil memprediksi dengan benar sebanyak 70% dari total data. Presisi (Precision) sebesar 80 menunjukkan bahwa dari semua prediksi yang dianggap benar (baik normal maupun abnormal) 80% di antaranya memang benar sesuai dengan kelas yang diprediksi. Recall sebesar 53,33 menunjukkan bahwa model hanya berhasil mendeteksi 53% dari seluruh data yang seharusnya termasuk dalam kategori abnormal. F1-Score sebesar 64 adalah metrik gabungan antara Precision dan Recall, yang mencerminkan keseimbangan antara keduanya.

4.7. Skenario Pengujian

Penelitian ini berfokus pada optimasi klasifikasi pneumonia pada citra X-ray menggunakan metode ekstraksi fitur Haralick-LBP dan algoritma K-Nearest Neighbors (KNN). Pengujian dilakukan melalui dua skenario utama, yaitu pembagian data dan parameter nilai K. Pada skenario pembagian data, evaluasi dilakukan untuk menentukan rasio terbaik dalam membagi data latih dan data uji dengan rasio 90:10, 80:20, dan 70:30. Tujuan pengujian ini adalah untuk menentukan rasio pembagian data yang memberikan hasil terbaik dalam pelatihan dan pengujian model klasifikasi serta mengevaluasi kemampuan model dalam melakukan generalisasi terhadap data baru.

4.8. Hasil Evaluasi Performa Model

Bagian ini memaparkan hasil evaluasi performa model yang dilakukan dengan menguji beberapa kombinasi model untuk menilai kinerja dalam mengklasifikasikan citra pneumonia pada X-ray. Model yang dievaluasi meliputi kombinasi Haralick-LBP dengan KNN, Haralick dengan KNN, dan LBP dengan KNN. Setiap model diuji menggunakan proporsi data 90:10, yang telah terbukti memberikan performa terbaik berdasarkan hasil pengujian sebelumnya. Selain itu, pengujian juga dilakukan menggunakan nilai K=5 yang merupakan nilai K optimal berdasarkan hasil pengujian nilai K. Tabel 4.3 menyajikan hasil evaluasi model berdasarkan pengujian ini.

Table 1. Hasil Evaluasi Model

No	Model	Akurasi
1	Haralick-LBP dan K-Nearest Neighbor	80
2	Haralick dan K-Nearest Neighbor	60
3	LBP dan K-Nearest Neighbor	70

Berdasarkan hasil evaluasi model yang ditampilkan pada Tabel 1, dapat disimpulkan bahwa kombinasi fitur Haralick-LBP dengan K-Nearest Neighbor (KNN) memberikan hasil klasifikasi yang lebih baik dibandingkan dengan model lainnya. Model ini memperoleh akurasi tertinggi sebesar 80% yang menunjukkan bahwa gabungan fitur Haralick dan LBP memberikan representasi tekstur yang lebih optimal dalam mendeteksi pneumonia pada citra X-ray.

5. KESIMPULAN DAN SARAN

Kesimpulan dari penelitian ini adalah penggunaan metode LBP-Haralick yang dikombinasikan dengan algoritma KNN telah terbukti efektif dalam mengklasifikasikan citra rontgen pneumonia, dengan LBP mendeteksi pola mikrotekstur lokal dan Haralick menggambarkan hubungan spasial piksel melalui GLCM untuk ekstraksi tekstur secara keseluruhan. Kombinasi fitur ini, yang dikategorikan oleh KNN dengan nilai k=5, menghasilkan model terbaik dengan akurasi hingga 80%, berkat keefektifan jarak Euclidean dalam membedakan antara gambar Normal dan Pneumonia.

Dibandingkan dengan penggunaan terpisah LBP atau Haralick, kombinasi ini meningkatkan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score, karena memberikan representasi gambar yang lebih lengkap dan bermakna.

Saran untuk meningkatkan penelitian masa depan dalam klasifikasi gambar sinar-X untuk pneumonia, disarankan untuk menggunakan dataset yang lebih besar dan beragam dalam hal resolusi dan kondisi pencahayaan guna meningkatkan generalisasi model dan mengurangi risiko overfitting. Selain itu, mengeksplorasi algoritma klasifikasi yang lebih canggih seperti Algoritma Genetika, Jaringan Saraf Konvolusional (CNN), atau metode ensemble seperti AdaBoost dan Gradient Boosting dapat diterapkan untuk meningkatkan akurasi dan ketahanan terhadap noise. Akhirnya, mengembangkan sistem ini menjadi aplikasi berbasis web atau mobile akan memudahkan implementasinya secara langsung oleh tenaga medis, laboratorium, atau masyarakat umum, sehingga memberikan manfaat praktis yang lebih signifikan dalam kehidupan sehari-hari.

DAFTAR PUSTAKA

- [1] D. Ibrahim and A. R. Bizri, "Chest computed tomography and chest X-ray in the diagnosis of community-acquired pneumonia: a retrospective observational study".
- [2] C. Lisa and M. Arini, "Duration of Hospitalization and Risk Factors of Readmission of Community-Acquired Pneumonia Incidence in Hospitalized Toddlers," *J. Respirologi Indones.*, vol. 44, no. 1, pp. 59–65, Jan. 2024, doi: 10.36497/jri.v44i1.560.
- [3] J. Jumadi and D. Sartika, "PENGOLAHAN CITRA DIGITAL UNTUK IDENTIFIKASI OBJEK MENGGUNAKAN METODE HIERARCHICAL AGGLOMERATIVE CLUSTERING," no. 2, 2021.
- [4] N. Wiliani, T. Khawa, and S. Ramli, "Peningkatan Kontras Pada Preprocessing Gambar Permukaan Solar Panel dengan Histogram," *J. Ilmu Komput.*, vol. 2, no. 1, 2025.
- [5] M. P. Juniarta, M. H. H. Arrahman, and N. Sulistianingsih, "Implementasi Gray Level Co-Occurrence Matrix (GLCM) Untuk Mendeteksi Penyakit Daun Pada Tanaman Hortikultura".
- [6] N. S. Fatimah and S. Agustin, "Klasifikasi Citra Batik Menggunakan Local Binary Pattern (LBP) dan Support Vector Machine (SVM)," *J. Algoritma*, vol. 22, no. 1, pp. 185–196, May 2025, doi: 10.33364/algoritma/v.22-1.2208.
- [7] N. T. Ujianto, Gunawan, H. Fadillah, A. P. Fanti, A. D. Saputra, and I. G. Ramadhan, "Penerapan algoritma K-Nearest Neighbors (KNN) untuk klasifikasi citra medis," *IT-Explore J. Penerapan Teknol. Inf. Dan Komun.*, vol. 4, no. 1, pp. 33–

- 43, Feb. 2025, doi: 10.24246/itexplore.v4i1.2025.pp33-43.
- [8] M. Fadli and R. A. Saputra, "KLASIFIKASI DAN EVALUASI PERFORMA MODEL RANDOM FOREST UNTUK PREDIKSI STROKE," vol. 12, no. 02, 2023
- [9] A. S. Wiratama, M. Rifqi, and S. Maesaroh, "EFEKTIVITAS TRANSFER LEARNING DALAM PENDETEKSIAN PENYAKIT PNEUMONIA MELALUI CITRA X-RAY PARU MANUSIA," *J. Ilm. Sains Dan Teknol.*, vol. 7, no. 1, pp. 43–52, Apr. 2023, doi: 10.47080/saintek.v7i1.2551.
- [10] A. Nugroho, A. S. Nugraha, A. A. Rasyid, and F. W. Rahayu, "Improvement on KNN using genetic algorithm and combined feature extraction to identify COVID-19 sufferers based on CT scan image," *TELKOMNIKA Telecommun. Comput. Electron. Control*, vol. 19, no. 5, p. 1581, Oct. 2021, doi: 10.12928/telkomnika.v19i5.18535.