

BAB I PENDAHULUAN

1.1 Latar Belakang

COVID-19 (*Coronavirus Disease 2019*) telah dinyatakan sebagai pandemi global dan kabarnya berasal dari kota Wuhan di China sejak Desember 2019. WHO (World Health Organization) menyatakan wabah COVID-19 ini penyebarannya sangatlah cepat dan sebagian besar dunia terjangkit virus ini [1]. Pada tanggal 12 Februari 2020, WHO secara resmi menyatakan penyakit itu sebagai penyakit *Coronavirus Disease 2019 (COVID-19)*. *Coronavirus Disease-2019 (COVID-19)* adalah penyakit menular akut terutama melibatkan sistem pernapasan, yang baru-baru ini ditemukan pada manusia [2]. Per tanggal 05 Mei 2022, terdapat 512 juta kasus positif COVID-19 dan 6,2 juta orang telah dilaporkan meninggal dunia di seluruh dunia [3].

Tenaga medis mendeteksi adanya COVID-19 dengan pemeriksaan medis seperti rapid test, swab, dan PCR. Orang yang memiliki gejala-gejala terindikasi COVID-19 akan di minta melakukan cek darah, pemeriksaan radiologi rontgent menggunakan sinar-x (x-ray), lalu dilanjutkan swab [4]. Rontgent merupakan tes diagnostik berbasis pencitraan dengan sinar-x (x-ray) yang bertujuan untuk melihat bagian dalam tubuh seperti paru-paru dan saluran pernapasan dalam bentuk citra [4]. Rontgen digunakan untuk memindai keadaan paru-paru pasien. Hasil pemindaian gambar paru-paru pasien diperlukan untuk analisa apakah paru-paru pasien terdeteksi telah terinfeksi oleh COVID-19 atau tidak, tindakan ini penting untuk diagnosis klinis COVID-19 bagi penderita menjalankan isolasi dini dan pengobatan serta mencegah penyebaran penyakit [5]. Dari permasalahan tersebut, maka pada penelitian akan dilakukan pengidentifikasian penyakit COVID-19 berdasarkan citra x-ray paru-paru.

Pemrosesan citra gambar atau *image processing* mengalami peningkatan di berbagai bidang seperti pemrosesan gambar industri, pencitraan medis, pencitraan *real time*, klasifikasi tekstur, pengenalan objek, dsb [6]. Beberapa penelitian yang mengimplementasikan citra medis dilakukan oleh [7] mengklasifikasikan pneumonia menggunakan metode klasifikasi *K-Nearest Neighbor (K-NN)* dan

ekstraksi fitur *gray level co-occurrence matrix* (GLCM) untuk metode ekstraksi. Tahap *preprocessing* sebelum dilakukan klasifikasi yaitu *cropping*, *resizing*, *contrast stretching*, dan *thresholding*. Hasil penelitian yang diperoleh menunjukkan akurasi terbaik pada $k=5$ adalah 66,20%. Penelitian selanjutnya [8] mengkaji perbandingan antara algoritma *Naive Bayes* dan *Decision Tree Classification* menggunakan dataset COVID-19, tahap *preprocessing* penyeleksian data dibersihkan dari *missing value* atau *noise*, selanjutnya membuang atribut yang tidak diperlukan, ekstraksi fitur dengan teknik *correlation matrix* dan *feature importance*. Hasil ujicoba dan perhitungan akurasi menggunakan *Confusion Matrix* mendapatkan hasil untuk algoritma *Naive Bayes* dengan akurasi sebesar 22,2%, sedangkan untuk *Decision Tree* menghasilkan akurasi lebih tinggi yakni 92,59%. Penelitian lain terkait deteksi COVID-19 pada citra x-ray [9], mengkaji tingkat efisiensi dan persentase keakurasian dengan metode *Convolutional Neural Network* (CNN) dalam mengklasifikasi citra X-ray *positive* dan *negative* COVID-19 yang dihasilkan oleh *Computed Tomography* (CT), melakukan pengujian terhadap kumpulan data dibagi menjadi 352 gambar latih, 88 gambar validasi dan 110 gambar uji. Hasil penjujian akurasi klasifikasi, *precision*, *recall*, *F1-score*, *support* masing-masing adalah 95% 96%, 97,31%, 96,60%.

Pada penelitian ini, pengidentifikasian penyakit COVID-19 berdasarkan citra x-ray paru-paru dan metode klasifikasi berbasis *machine learning*. Dengan demikian maka penulis mengangkat permasalahan ini sebagai penelitian dengan judul "**Perbandingan K-Nearest Neighbor Dan Naive Bayes Untuk Klasifikasi COVID-19 Pada Citra X-Ray**". Data akan dibagi berdasarkan data latih dan data uji, metode *gray level co-occurrence matrix* (GLCM) membantu dalam mengidentifikasi dan mengklasifikasikan citra x-ray ke dalam kelas *positive* dan *negative* COVID-19.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dikemukakan, maka permasalahan yang dapat dirumuskan adalah sebagai berikut:

- a. Bagaimana algoritma *K-Nearest Neighbor* (K-NN) dan *Naive Bayes* mampu mengklasifikasikan citra x-ray paru-paru dengan baik?

- b. Berapa akurasi citra x-ray paru-paru dengan algoritma *K-Nearest Neighbor* (K-NN) dan *Naive Bayes*?
- c. Bagaimanan perbandingan performa algoritma *K-Nearest Neighbor* (K-NN) dan *Naive Bayes*?

1.3 Batasan Masalah

Beberapa batasan masalah yang digunakan dalam penelitian ini adalah sebagai berikut:

- a. Citra medis berupa hasil rongent sinar-x (x-ray) paru-paru dari kaggle.
- b. Input berupa image sebuah citra x-ray paru-paru.
- c. Input berupa image berformat PNG.
- d. Satu image berisi satu hasil x-ray paru-paru.
- e. Analisis pengolahan citra dan klasifikasi citra dikembangkan menggunakan bahasa pemrograman Matlab.
- f. Metode GLCM digunakan untuk ekstraksi citra medis serta klasifikasi algoritma *K-Nearest Neighbor* (K-NN) dan *Naive Bayes*.

1.4 Maksud dan Tujuan Penelitian

Adapun maksud dan tujuan yang akan dicapai dalam penelitian ini adalah Mengklasifikasi COVID-19 dalam kelas positive dan negative pada citra x-ray paru-paru serta perbandingan akurasi dari algoritma *K-Nearest Neighbor* (K-NN) dan *Naive Bayes*.

1.5 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah :

- a. Membantu meningkatkan keakuratan dalam mengklasifikasi citra x-ray yang ada.
- b. Penelitian ini diharapkan bisa menjadi dasar bagi penelitian lain terkait klasifikasi.

1.6 Metode Penelitian

Peneliti Metode penelitian yang dilakukan dalam penelitian ini adalah sebagai berikut:

1.6.1 Metode Pengumpulan Data

Mengumpulkan data atau materi penelitian, berupa :

- a. Observasi, suatu metode untuk memperoleh data dengan cara melakukan pengamatan data citra, dataset diperoleh dari situs [kaggle.com](https://www.kaggle.com) untuk dianalisis citra x-ray paru-paru yang sesuai dengan penelitian, mengenai hasil pengamatan terbagi sebagai citra data latih dan data uji .
- b. Studi Pustaka atau Literatur, suatu metode dimana sumber data berdasarkan teori-teori dan literatur yang berhubungan dengan pembahasan. Sumber referensi berupa buku, skripsi, jurnal, proceeding, google book, dan internet official dari penelitian sebelumnya.

1.6.2 Metode Pengembangan Sistem

Metode yang digunakan adalah *waterfall*, karena tahapan yang dilalui harus menunggu tahap sebelumnya rampung dan berjalan secara berurutan untuk menghindari terjadinya pengulangan tahapan. Metode ini merupakan metode yang sering digunakan untuk pengembangan pada umumnya. Program *prototype* akan dikembangkan melalui platform pengembangan matlab. Metode dikembangkan untuk mengklasifikasi citra x-ray paru-paru yang ada.

1.6.3 Analisis dan Perancangan Sistem

- a. Merancang citra x-ray paru-paru dapat dibedakan meliputi konversi citra *rgb* menjadi citra *grayscale*.
- b. Merancang metode segmentasi citra menggunakan *Thresholding* dan *adjust image intensity*.
- c. Analisis ekstraksi ciri citra x-ray menggunakan GLCM yaitu nilai *contrast*, *correlation*, *energy*, dan *homogeneity*.
- d. Merancang sistem klasifikasi citra menggunakan algoritma *K-Nearest Neighbor* (K-NN) dan *Naive Bayes*.

1.6.4 Implementasi

Melakukan implementasi hasil perancangan melalui program aplikasi. Pemrograman dilakukan pada program pengembangan matlab untuk pemrosesan data dengan ekstraksi fitur GLCM di komputer atau laptop personal menggunakan sistem operasi Microsoft Windows 10.

1.6.5 Pengujian dan Analisis Hasil

Pengujian sistem untuk mengetahui hasil *output* atau keluaran dari sistem pengenalan. Data citra x-ray paru-paru diperoleh dari dataset kaggle.com, data citra masukan tersebut disimpan dalam folder yang memisahkan citra x-ray paru-paru dalam data latih dan data uji, sehingga mempermudah proses pengujian. Pengujian dilakukan untuk menjawab tujuan dan permasalahan dalam penelitian ini adalah dilakukannya pengujian citra data uji dan klasifikasi dengan algoritma *K-Nearest Neighbor* (K-NN) dan *Naive Bayes* serta tingkat akurasi yang dihasilkan. Peneliti melakukan penarikan kesimpulan setelah pengujian sistem.

1.7 Sistematika Penulisan

BAB I : Pendahuluan.

Bab ini membahas mengenai latar belakang masalah, perumusan masalah, batasan masalah, tujuan dan manfaat penelitian, metode penelitian, dan sistematika penulisan laporan penelitian.

BAB II : Landasan Teori

Bab ini menjelaskan dasar teori dari klasifikasi sayur-mayur dengan teknik image processing metode *Naive Bayes* serta *software* yang digunakan atau dibutuhkan dari beberapa pustaka kemudian menjadikannya sebagai acuan dan panduan dalam penulisan laporan penelitian.

BAB III : Metode Penelitian

Bab ini menjelaskan metode penelitian yang dilakukan dalam menyelesaikan penelitian.

BAB IV : Hasil Dan Pembahasan

Bab ini menjelaskan mengenai hasil pengujian yang telah dilakukan pada citra data uji. Hasil pengujian kemudian dibahas untuk mengetahui akurasi metode *Naive Bayes* dalam mengklasifikasi sayur-mayur.

BAB V : Penutup

Berisi kesimpulan dan saran dari pembahasan untuk melengkapi penyusunan laporan penelitian dan saran yang selanjutnya dilakukan bila penelitian ini dilanjutkan.

