

BAB I

PENDAHULUAN

Bab ini membahas mengenai latar belakang penelitian yang mencakup pentingnya klasifikasi penyakit daun teh serta peran teknologi dalam mendukung deteksi dini dan penanganan penyakit tersebut. Bab ini memberikan gambaran umum mengenai konteks penelitian dan urgensi topik yang diangkat, sehingga dapat menjadi landasan yang kuat untuk pembahasan lebih lanjut pada bab-bab selanjutnya.

1.1. Latar Belakang

Tanaman teh (*Camellia Sinensis*) adalah produk budidaya yang mempunyai peran dalam perekonomian Indonesia. Teh sendiri merupakan sumber penghasilan negara, seperti minyak dan gas [1]. Indonesia merupakan eksportir teh terbesar keenam di dunia, karena teh merupakan minuman paling populer di dunia dan banyak diminati karena manfaatnya untuk kesehatan. Namun, dalam satu dekade terakhir, produktivitas produksi teh di Indonesia seringkali terhambat yang disebabkan oleh penyakit yang dapat mengganggu kesehatan daun teh, sehingga terjadi penurunan hasil panen [2].

Menurut data yang dimuat oleh Badan Pusat Statistik (BPS), total produksi teh di Indonesia terjadi penurunan dari tahun 2020 sampai dengan 2023. Pada tahun 2020 produksi teh mencapai 94.156 ton, turun menjadi 87.568 ton pada tahun 2021 yang mengalami penurunan sebesar 7,0%. Tahun 2022 produksi teh turun menjadi 74.765 ton yang mengalami penurunan sebesar 14,6% dibandingkan tahun 2021, dan produksi teh turun lagi pada tahun 2023 menjadi 67.818 ton yang mengalami penurunan sebesar 9,29% dibandingkan pada tahun 2022 [3]. Hal tersebut menunjukkan bahwa produksi teh di Indonesia mengalami penurunan setiap tahunnya yang disebabkan oleh penyakit yang menyerang pada tanaman teh. Oleh karena itu, untuk menjaga kualitas produktivitas tanaman teh, perlu dilakukan identifikasi dan penanggulangan terhadap berbagai jenis penyakit yang dapat menyerang tanaman teh, sehingga memperoleh penanganan yang sesuai sehingga resiko gagal panen dapat diminimalkan [4].

Klasifikasi adalah suatu proses mengidentifikasi kelompok dari suatu data berdasarkan persamaannya, dimana masing-masing kelompok telah dibentuk sebelumnya dan melalui suatu proses tertentu [5]. Ada banyak metode untuk proses

klasifikasi. Namun, penelitian yang telah dilakukan sebelumnya ditemukan bahwa ekstraksi fitur GLCM dan LBP menggunakan klasifikasi K-NN mendapat *accuracy* yang cukup tinggi.

Gray Level Co-Occurrence Matrix (GLCM) merupakan sebuah *patch* yang memiliki representasi keterkaitan antar piksel citra berdasarkan tingkat keabuan, dengan mempertimbangkan sudut dan jarak tertentu. Melalui analisis hubungan spasial antar piksel, diperoleh matriks GLCM yang menggambarkan pola tekstur dalam citra [6]. Sedangkan *Local Binary Pattern* (LBP) adalah metode fitur tekstur untuk mengubah citra menjadi sekumpulan nilai biner berdasarkan hubungan intensitas antara piksel pusat dan piksel disekitarnya [7]. Algoritma klasifikasi *K-Nearest Neighbor* (K-NN) digunakan sebagai metode untuk melakukan klasifikasi. K-NN yang bekerja dengan cara menentukan kelas suatu data baru berdasarkan sejumlah tetangga terdekatnya dengan jarak paling dekat menggunakan data yang dilatih [8].

Penelitian membahas mengenai algoritma klasifikasi K-NN dengan ekstraksi fitur tekstur GLCM-LBP sudah pernah dilakukan sebelumnya oleh [9]. Dengan judul “Pendeteksian Penggunaan Masker Untuk Pencegahan Penyebaran Covid-19 Menggunakan Algoritma *K-Nearest Neighbor*”. Pada penelitian tersebut terdapat penerapan algoritma K-NN dengan ekstraksi ciri tekstur GLCM-LBP untuk mengklasifikasi penggunaan masker. Data yang digunakan berjumlah 5507 *dataset* pengguna masker yang terbagi menjadi 3 kelas data citra terdiri dari, menggunakan masker, tidak menggunakan masker, dan hanya menggunakan sebagian masker, data tersebut menggunakan pembagian 70:30 sebelum dilakukan proses. Hasil *accuracy* yang diperoleh sebesar 91,04%.

Penelitian lain mengenai klasifikasi menggunakan algoritma K-NN dengan ekstraksi fitur GLCM sudah pernah dilakukan sebelumnya oleh [10] dengan judul “*Classification Of Durian Types Using Features Extraction Gray Level Co-Occurrence Matrix (GLCM) And K-Nearest Neighbor (K-NN)*”. Pada penelitian tersebut terdapat penerapan metode K-NN dengan ekstraksi fitur GLCM untuk mengklasifikasi beberapa tipe durian seperti durian ajimah, bawor, candimulyo, matahari, miming, montong, petruk, dan musangking. Data yang digunakan pada penelitian tersebut berjumlah 1,602 *dataset* durian, akan menjadi 1281 data latih dan 321 data uji sehingga menghasilkan *accuracy* sebesar 93%.

Penelitian lain menggunakan ekstraksi fitur tekstur LBP dan klasifikasi K-NN sudah pernah dilakukan sebelumnya oleh [11] dengan judul “Ekstraksi Ciri Pada Telapak Tangan Dengan Metode *Local Binary Pattern* (LBP)”. Pada penelitian ini membahas tentang penerapan algoritma K-NN dengan ekstraksi ciri LBP untuk melakukan ekstraksi ciri pada telapak tangan. Dataset yang digunakan berjumlah 1344 *dataset* yang telah dilakukan cropping dengan ukuran 640 x 480. Hasil *accuracy* tertinggi yang diperoleh sebesar 92% dengan jarak $k=2$.

Perbedaan penelitian ini dengan penelitian sebelumnya terletak pada penggunaan *dataset* penyakit daun teh dan diklasifikasi dengan ekstraksi fitur tekstur GLCM-LBP dan algoritma klasifikasi K-NN, selain itu penelitian ini menggunakan beberapa parameter dari GLCM seperti *Contrast*, *Correlation*, *Energy*, *Homogeneity*. Digunakan juga ekstraksi fitur LBP untuk melengkapi kekurangan dari GLCM seperti untuk mendapatkan pola tekstur lokal. Kemudian untuk bagian pengujian dan evaluasi, penelitian ini menggunakan 3 proporsi pembagian data latih dan data uji yaitu 70:30, 80:20, 90:10, yang kemudian akan dilakukan pengujian menggunakan nilai $k=3$. Selanjutnya akan dilakukan klasifikasi menggunakan algoritma K-NN dengan variasi nilai $k=3$, $k=5$, $k=7$, $k=9$, $k=11$, pengujian dilakukan menggunakan hasil proporsi data terbaik yang telah dilakukan pengujian. Setelah itu menggunakan *confusion matrix* dengan empat metrik, yaitu *accuracy*, presisi, *Recall*, dan *F1-score*, sehingga menemukan kombinasi optimal antara nilai k dan proporsi data, guna menghasilkan *model* klasifikasi yang efisien.

Berdasarkan penelitian yang telah dilakukan sebelumnya, diketahui bahwa klasifikasi menggunakan kombinasi ekstraksi fitur GLCM-LBP dengan Algoritma K-NN menghasilkan *accuracy* yang baik. Oleh karena itu, penelitian ini mengambil judul penelitian “Klasifikasi Penyakit Daun Teh Menggunakan Metode K-NN Berdasarkan Fitur Tekstur GLCM-LBP”. Hasil dari penelitian ini diharapkan dapat membantu petani dalam mengidentifikasi penyakit yang menyerang daun teh, beserta penanggulangannya agar dapat meningkatkan produktivitas dan kualitas hasil panen.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, maka rumusan masalah yang didapat sebagai berikut:

1. Bagaimana penerapan ekstraksi fitur GLCM-LBP dan algoritma *K-Nearest Neighbor* dalam melakukan klasifikasi penyakit pada daun teh?

2. Bagaimana performa klasifikasi penyakit daun teh menggunakan ekstraksi fitur GLCM-LBP dan algoritma *K-Nearest Neighbor*?

1.3. Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dipaparkan, maka tujuan yang ingin dicapai dalam penelitian ini sebagai berikut:

1. Mengetahui Penerapan ekstraksi fitur GLCM-LBP dan algoritma *K-Nearest Neighbor* dalam melakukan klasifikasi penyakit pada daun teh
2. Mengetahui performa klasifikasi penyakit daun teh menggunakan ekstraksi fitur GLCM-LBP dan algoritma *K-Nearest Neighbor*.

1.4. Manfaat Penelitian

Dengan berjalannya penelitian ini, diharapkan penelitian dapat memberikan beberapa manfaat sebagai berikut:

1. Dapat menjadi acuan penelitian selanjutnya mengenai topik penelitian citra penyakit daun teh.
2. Dapat menjadi referensi bagi penelitian-penelitian selanjutnya terkait ekstraksi fitur GLCM-LBP dan algoritma K-NN dalam penggunaan *dataset* terbatas dan memperkaya literatur yang ada di bidang pemrosesan citra digital.
3. Dapat membantu petani dalam tindakan preventif secara lebih cepat dan akurat dalam penanganan penyakit pada daun teh, Sehingga produktivitas daun teh dapat meningkat.

1.5. Batasan Masalah

Penelitian ini perlu adanya batasan masalah supaya tidak keluar dari permasalahan yang dikaji, maka diperlukan batasan masalah sebagai berikut:

1. Dataset yang digunakan merupakan data sekunder yang diambil dari sumber website *open source* Kaggle.
2. Data yang digunakan berjumlah 1800 citra penyakit daun teh yang terdiri dari 6 kategori daun teh.
3. Penelitian ini menggunakan ekstraksi fitur *Gray Level Co-occurrence Matrix* (GLCM) dan *Local Binary Pattern* (LBP).
4. Penelitian ini menggunakan algoritma klasifikasi *K-Nearest Neighbor* (K-NN).
5. Penelitian ini berfokus pada pembangunan *model* klasifikasi penyakit pada daun teh.