

BAB I

PENDAHULUAN

1.1 Latar Belakang

Negara Indonesia memiliki berbagai hasil produksi yang berasal dari jenis buah - buahan. Salah satu buah yang cukup diminati di Indonesia adalah buah anggur, dilihat dari data Badan Pusat Statistik dan Direktorat Jenderal Hortikultura, Indonesia mampu menghasilkan anggur sebesar 12.164 ton pada tahun 2021. Nilai tersebut mengalami kenaikan sebesar 2.18% dibandingkan tahun sebelumnya.

Buah anggur atau buah yang memiliki nama Ordo *Rhamnales*, Famili *Vitaceae*, Genus *Vitis*, dan Spesies *Vitis vinifera L* merupakan buah dari tanaman kering yang bisa hidup di dataran rendah hingga 300 meter di atas permukaan laut (Purwantiningsih, Leksono, & Yanuwidi, 2012). Buah anggur tergolong salah satu komoditas buah-buahan yang bergizi dan mempunyai nilai ekonomi tinggi dan dapat dibudidayakan di daerah yang beriklim tropis. Buah anggur dapat dikonsumsi dalam bentuk segar maupun olahan seperti *wine*, jus, sirup, permen, selai dan lain-lain (Krismawati & Prahardini, 2011). Selain dapat dikonsumsi dengan berbagai cara buah anggur juga kaya akan kandungan nutrisi seperti karbohidrat, mineral, protein, vitamin, serta *phytochemical* (Sasanthi, 2016). Di Indonesia buah anggur tergolong buah yang cukup umum untuk diperjual belikan bahkan relatif mudah untuk ditemukan, meskipun buah anggur mudah untuk ditemukan namun masyarakat Indonesia cukup sulit atau bahkan tidak dapat membedakan jenis buah anggur satu dengan yang lainnya, hal ini dapat membuat kesalahan dalam memilih jenis buah anggur yang diinginkan. Hal ini dapat diatasi dengan memanfaatkan teknik *Image Classification* yang terdapat pada bidang keilmuan *Machine Learning*.

Machine Learning merupakan sebuah bidang keilmuan yang mempelajari bagaimana cara untuk membuat dan mengembangkan sebuah program yang dapat berkembang sendiri dan menciptakan pengetahuan baru dari pengetahuan lama yang telah diberikan kepada program (Dinata & Hasdyna, 2020). Di dalam

Machine Learning terdapat sebuah bidang bernama *Image Classification* yang dapat menggantikan kemampuan visual manusia (Kohsasih, 2022).

Klasifikasi adalah sebuah proses untuk mengidentifikasi kelompok dari suatu data berdasarkan kesamaan yang dimiliki, dimana setiap kelompok telah terbentuk dan melalui suatu proses tertentu sebelumnya (Dillak, Pangestuty, & Bintiri, 2012). Terdapat berbagai macam metode yang dapat digunakan untuk melakukan proses pengklasifikasian, salah satu diantaranya adalah metode *K-Nearest Neighbor* (K-NN). K-NN merupakan sebuah metode pengklasifikasian yang mengklasifikasikan sebuah objek dengan cara mencari data pembelajaran yang jaraknya paling dekat dengan data objek yang ingin diklasifikasikan. Untuk melakukan perhitungan jarak antara objek pada umumnya K-NN akan menggunakan perhitungan *Euclidean distance* sebagai dasar perhitungannya (Liantoni, 2015). Keunggulan metode K-NN dalam melakukan klasifikasi yaitu algoritma yang sederhana dan mudah dipelajari, dapat melakukan pelatihan dengan cepat, tahan terhadap pelatihan yang memiliki derau, serta efektif untuk data pelatihan yang besar (Mutrofin, dkk, 2014).

Dalam melakukan proses pengklasifikasian terhadap suatu citra terdapat sebuah tahapan yang bernama ekstraksi fitur. Ekstraksi fitur merupakan sebuah tahapan yang berguna untuk memperoleh fitur yang terkandung pada suatu citra yang nantinya dapat digunakan sebagai dasar untuk pengklasifikasian (Arriawati, Santoso, & Christyono, 2011)

Grey Level Co-occurrence Matrix (GLCM) merupakan metode yang dapat digunakan untuk mendapatkan nilai fitur yang ada pada sebuah data dengan cara menghitung nilai probabilitas dari hasil perhitungan hubungan ketetanggaan antara dua piksel pada jarak dan orientasi sudut tertentu (B, Andriana, & Hidayat, 2017). Hasil dari pencarian fitur pada suatu citra menggunakan metode GLCM nantinya dapat digunakan sebagai masukan untuk mengklasifikasikan citra ke dalam kelompok yang telah disiapkan sebelumnya (Sukiman, 2020).

Hue, Saturation, Value (HSV) merupakan salah satu dari sistem warna yang digunakan oleh manusia dalam melakukan pemilihan warna yang ada pada suatu objek. Dengan mengolah HSV sebuah objek yang mempunyai warna

tertentu dapat diidentifikasi (Liantoni & Annisa, 2018). Ruang lingkup warna HSV terdiri dari 3 macam, yang pertama adalah *Hue* yang mewakili warna, lalu ada *Saturation* yang mewakili tingkat dominasi warna, dan yang terakhir ada *Value* yang mewakili tingkat kecerahan, Hal ini yang membuat metode HSV lebih cenderung untuk mendeteksi warna, tingkat kecerahan, serta tingkat dominasi suatu warna (Putranto, Hapsar, & Wijana, 2010) Nilai dari hasil ekstraksi fitur menggunakan HVS nantinya juga dapat digunakan sebagai dasar untuk klasifikasi buah anggur.

Pada penelitian sebelumnya telah dilakukan klasifikasi menggunakan metode K-NN dengan ekstraksi fitur HSV dan GLCM yang berjudul Klasifikasi Tingkat Kematangan Jambu Bol Berbasis Pengolahan Citra Digital Menggunakan Metode *K-Nearest Neighbor* (Syarifah, Riadi, & Susanto, 2022). Pada penelitian tersebut terdapat penerapan metode K-NN dengan ekstraksi warna HSV dan tekstur GLCM untuk mengklasifikasikan tingkat kematangan buah jambu bol ke kategori mentah, matang dan sangat matang. Hasil dari penelitian ini menunjukkan bahwa Metode K-NN dengan ekstraksi HSV dan GLCM dapat ditereapkan dengan baik dengan tingkat akurasi sebesar 93% dari data uji sejumlah 30, data latih sebanyak 60 data, dan menggunakan nilai ketetangaan $k=1$.

Pada penelitian lain membahas tentang Optimasi Ekstraksi Fitur Pada KNN Dalam Klasifikasi Penyakit Daun Jagung (Rachmawanto & Hadi, 2021). Penelitian ini menerapkan ekstraksi warna berupa HSV dan GLCM pada algoritma K-NN untuk mengklasifikasikan penyakit daun jagung ke dalam 4 kategori daun yaitu daun berpenyakit bercak, hawar, karat, dan daun yang sehat. Data yang digunakan pada penelitian tersebut berjumlah 200 dataset daun jagung yang dibagi menjadi 160 data latih dan 40 data uji. Pada penelitian tersebut dilakukan uji coba klasifikasi menggunakan metode KNN dengan ekstraksi fitur GLCM yang menghasilkan akurasi tertinggi sebesar 60% waktu melakukan klasifikasi menggunakan nilai $k = 5$ dan jarak piksel = 5. Sedangkan pada uji coba klasifikasi menggunakan metode KNN dengan ekstraksi fitur GLCM dan HSV diperoleh hasil akurasi paling tinggi sebesar 85% waktu melakukan klasifikasi menggunakan nilai komponen $K = 3$ serta jarak piksel = 1.

Berdasarkan pemaparan yang telah dielaskan sebelumnya, dapat diketahui bahwa metode K-NN, HSV, dan GLCM dapat menunjukkan hasil yang memuaskan dalam melakukan proses klasifikasi. Namun penelitian menggunakan metode yang serupa belum pernah dilakukan pada objek buah anggur, sehingga peneliti ingin mencoba melakukan sebuah penelitian untuk mengetahui tingkat efektivitas metode K-NN dengan kombinasi ekstraksi fitur metode GLCM dan HSV dalam klasifikasi buah anggur. Pada penelitian ini metode GLCM dan HSV akan digunakan untuk mendapatkan nilai fitur yang terdapat citra buah anggur, dimana nilai fitur tersebut akan dijadikan sebagai dasar untuk mengklasifikasikan citra buah anggur, sedangkan metode K-NN akan digunakan untuk mengklasifikasikan citra buah anggur dengan cara membandingkan nilai fitur pada citra buah anggur yang akan diklasifikasi dengan setiap nilai fitur yang telah dikelompokkan ke dalam 6 jenis kelompok (*Anggur Champagne, Cotton candy, Crimson Seedless, Gewurztraminer, Glenora, dan anggur Kyoho*) yang diperoleh dari website kaggle.com

1.1 Rumusan Masalah

Berdasarkan latar belakang diatas, maka dapat dirumuskan bahwa rumusan masalah yang akan dibahas sebagai berikut :

1. Bagaimana cara algoritma *Hue Saturation Value* dan algoritma *Grey Level Co-occurrence Matrix* dapat mengekstraksi fitur citra buah anggur ?
2. Bagaimana cara algoritma *K-Nearest Neighbor* dapat mengklasifikasikan citra buah anggur ?
3. Bagaimana performa algoritma *Hue Saturation Value, Grey Level Co-occurrence Matrix* dan *K-Nearest Neighbor* dalam mengklasifikasikan citra buah anggur ?

1.2 Tujuan

Berdasarkan rumusan masalah di atas, maka dapat ditentukan bahwa tujuan dari penelitian ini adalah :

1. Untuk mengetahui cara kerja algoritma *Hue Saturation Value* dan algoritma *Grey Level Co-occurrence Matrix* dalam mengekstraksi fitur citra buah anggur.

2. Untuk mengetahui cara kerja algoritma *K-Nearest Neighbor* dalam mengklasifikasikan citra buah anggur.
3. Untuk mengetahui performa algoritma *Hue Saturation Value*, *Grey Level Co-occurrence Matrix* dan *K-Nearest Neighbor* dalam mengklasifikasikan citra buah anggur.

1.3 Manfaat

Beberapa manfaat yang dapat diperoleh dari penelitian ini adalah sebagai berikut :

1. Dapat mengetahui akurasi dari kombinasi algoritma *Hue Saturation Value*, *Grey Level Co-occurrence Matrix* dan *K-Nearest Neighbor* dalam mengklasifikasi citra buah anggur.
2. Penelitian ini dapat digunakan sebagai bahan referensi untuk penelitian selanjutnya.

1.4 Batasan Masalah

Terdapat batasan masalah dalam melakukan penelitian yang diantaranya adalah sebagai berikut :

1. Data yang digunakan dalam penelitian ini merupakan data sekunder (data yang diambil secara tidak langsung oleh peneliti) yang diambil dari website kaggle.
2. Jenis buah anggur yang akan digunakan sudah ditentukan sebelumnya sebagai buah anggur *Champagne*, *Cotton candy*, *Crimson Seedless*, *Gewurztraminer*, *Glenora*, dan anggur *Kyoho*.
3. Jumlah data yang digunakan pada penelitian ini ada sebanyak 900 data yang tersebar kedalam 6 kelas yakni, 100 data kelas *Champagne* yang, 150 kelas *Cotton candy*, 100 kelas *Crimson Seedless*, 150 data kelas *Gewurztraminer*, 200 data kelas *Glenora*, dan 200 data kelas *Kyoho*.
4. Metode yang digunakan sebagai ekstraksi fitur adalah metode *Hue Saturation Value* dan metode *Grey Level Co-occurrence Matrix* sedangkan metode yang digunakan sebagai klasifikasi citra adalah metode *K-Nearest Neighbor*.