

OPTIMASI EKSTRAKSI FITUR PADA KNN DALAM KLASIFIKASI PENYAKIT DAUN JAGUNG

Eko Hari Rachmawanto¹, Heru Pramono Hadi²

¹Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro

²Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro

e-mail: ¹eko.hari@dsn.dinus.ac.id, ²heru.pramono.hadi@dsn.dinus.ac.id

ABSTRAK

Di Indonesia jagung sering digunakan sebagai komoditas utama makanan pokok selain nasi. Tanaman jagung memiliki potensi terkena penyakit ataupun serangan hama kapan saja yang menyebabkan gagal panen. Penyakit yang dapat menyerang tanaman jagung bisa dilihat dari perubahan daun. Deteksi dini terhadap penyakit dapat mencegah penyakit menyebar lebih luas, salah satunya dengan perubahan yang terjadi pada daun jagung. Penelitian ini mencoba melakukan identifikasi daun yang tidak sehat dengan cara ekstraksi ciri dan warna pada citra untuk mendeteksi penyakit daun tanaman jagung yaitu hawar, bercak dan karat. Proses klasifikasi citra dilakukan melalui akusisi citra menjadi data latih dan uji, kemudian menghitung nilai hasil fitur ekstraksi warna dan ekstraksi ciri. GLCM (Gray-Level Cooccurrence Matrix) sebagai ekstraksi ciri dan HSV sebagai ekstraksi warna. KNN (K Nearest Neighbors) dengan jarak Euclidean untuk klasifikasi. Dari 160 data citra latih dan 40 citra uji menggunakan algoritma KNN-HSV-GLCM didapatkan hasil akurasi terbaik yaitu 85% dengan menggunakan dengan nilai k adalah 3 dan jarak piksel 1 dan akurasi terendah dengan nilai k adalah 3 dan jarak piksel 3 sebesar 70%.

KATA KUNCI: Daun Jagung, GLCM, HSV, KNN.

1. PENDAHULUAN

Jagung adalah salah satu pengganti makanan pokok yang terpenting di Indonesia setelah padi. Jagung juga ditanam dan diolah sebagai pakan ternak, minyak, tepung maizena, dan bahan baku industri. Dalam pertumbuhan jagung di Indonesia, Jawa Timur memberikan 40% dari total produksi nasional [1]. Namun, sering ada banyak kerugian pada tanaman karena penyakit pada tanaman, contohnya pada jagung. Secara umum penyakit pada tanaman awalnya disebabkan oleh perubahan morfologi pada daun. Deteksi dan klasifikasi penyakit yang timbul pada daun secara akurat akan mencegah penyebaran penyakit secara luas [2]. Kategori utama penyakit daun tanaman didasarkan pada virus, jamur dan bakteri. Pengamatan para ahli adalah pendekatan utama yang digunakan untuk mengidentifikasi penyakit pada tanaman. Tetapi jika ada lahan jagung yang sangat luas maka sulit dilakukan pemantauan secara terus menerus. Apalagi jika dalam negara berkembang seperti Indonesia, petani mungkin harus pergi jauh untuk menghubungi ahli. Ini membuat biaya konsultasi juga mahal dan memakan waktu apalagi jika petani tidak mengetahui daun yang terinfeksi.

Dalam teknik pengolahan citra, terdapat metode untuk digunakan untuk mendeteksi penyakit pada daun tanaman. Awalnya dibutuhkan ekstraksi ciri agar membedakan antara citra daun yang sehat dan yang tidak. Proses klasifikasi menggunakan data inputan yang berasal dari hasil perhitungan fitur [3]–[5]. Hasil ekstraksi yang valid akan memudahkan proses identifikasi penyakit pada daun jagung secara akurat. Penelitian sebelumnya [6], menunjukkan objek akan mudah teridentifikasi tekstur fiturnya dengan mengambil nilai derajat keabuan menggunakan metode GLCM. Penelitian milik [7] juga menggunakan ekstraksi ciri GLCM dan Ruang Warna serta menghasilkan akurasi sebesar 83%. Metode GLCM dapat menganalisa fitur suatu objek secara statistik dilihat dari tekstur objek. Tekstur objek dapat dihitung dari dasar distribusi statistik antara dua pixel dengan hubungan posisi spasial atau sering disebut intensitas pixel [8]. GLCM sering digunakan untuk objek citra yang mempunyai tekstur alami dengan sup-pola

dan himpunan yang tidak terstruktur. Sehingga GLCM merupakan metode yang efisien untuk ekstraksi fitur dan tekstur objek [9]. Tujuan dari penggunaan metode GLCM ini untuk mendapatkan set fitur yang berasal dari bagian daun citra tanaman jagung.

Dalam penelitian lain, proses klasifikasi citra penyakit daun menggunakan algoritma K-Nearest Neighbors (KNN) dan memberikan hasil akurasi sebesar 80%. Algoritma KNN adalah teknik klasifikasi yang diterapkan untuk mengklasifikasikan objek yang sama dan berbeda menjadi lebih dari satu kelas [10]–[12]. KNN digunakan untuk memberikan kembali prediksi bernilai asli untuk sampel yang tidak diketahui yang diberikan.

2. METODE PENELITIAN

2.1 Jenis Penyakit Jagung

Hawar daun disebabkan oleh jamur *Helminthosporium turcicum* dan akan terus hidup selama inang masih mendapat suplai makanan. Jenis penyakit ini dapat diminimalisir dengan menyeprotkan fungisida setiap 2 bulan sekali. Gejala penyakit pada hawar daun disebabkan oleh bercak kecil dan akan berkembang menyerupai bentuk oval di bawah daun sesuai ilustrasi pada Gambar 2. Bercak daun, umumnya berwarna hijau atau coklat pada bagian atas daun dan dapat menjalan ke bagian bawah menuju batang sehingga menyebabkan tanaman menjadi kering dan mati. Penyakit ini disebabkan oleh patogen *Bipolaris maydis* Syn. Gejala yang timbul dari penyakit ini berupa bercak berwarna hijau kekuningan atau coklat kemerahan pada daun sesuai ilustrasi pada Gambar 1. Tanaman yang sudah terinfeksi bercak daun dapat mengering kemudian mati dalam waktu kurang lebih 3 minggu. Infeksi pada tanaman ini bisa ditimbulkan oleh konidia yang terbawa angin atau air percian hujan. Karat daun disebabkan jamur *Puccinia polysora*. Karat daun dapat berupa bercak oval kecil yang menjalan di seluruh permukaan daun yang disebarkan oleh angin. Penyakit ini dapat dibasmi menggunakan fungisida sesuai ilustrasi pada Gambar 1.



Hawar Daun Bercak Daun Karat Daun
Gambar 1. Contoh citra penyakit daun yang menyerang tanaman jagung

2.2 Hue Saturation Valuse

Hue sebagai jenis pada warna seperti merah, hijau, kuning yang digunakan sebagai pembeda dari warna serta dapat menentukan warna kemerahan dan kehijauan dari cahaya [13]. *Saturation* merupakan kejernihan relatif dari warna. *Value* atau nilai merupakan tingkat kecerahan dari warna. Ruang warna HSV akan digunakan untuk menghasilkan area daun yang terinfeksi penyakit. Untuk memperoleh nilai HSV, harus mengkonversi citra RGB menjadi HSV sesuai persamaan (1) sampai persamaan (6).

$$r = \frac{R}{(R + G + B)} \quad (1)$$

$$g = \frac{G}{(R + G + B)} \quad (2)$$

$$b = \frac{B}{(R + G + B)} \quad (3)$$

$$V = \max (r, g, b) \quad (4)$$

$$S = \begin{cases} 0, & \text{jika } V = 0 \\ 1, & -\frac{\min(r, g, b)}{v}, V > 0 \end{cases} \quad (5)$$

$$f(x) = \begin{cases} 0, & \text{jika } S = 0 \\ \frac{60 * (g - b)}{S * V}, & \text{jika } V = r \\ 60 * \left[2 + \frac{b - r}{S * V} \right], & \text{jika } V = g \\ 60 * \left[4 + \frac{r - g}{S * V} \right], & \text{jika } V = b \end{cases} \quad (6)$$

2.3 Gray Level Co-occurrence Matrix (GLCM)

Dalam analisisnya, suatu objek secara statistik dapat dilihat dari tekstur objek tersebut, artinya tekstur objek dihitung dari dasar distribusi statistik antara dua pixel sesuai dengan hubungan posisi spasial derajat keabuan [9], [14], [15]. Konsep sederhana metode algoritma GLCM digunakan untuk menghitung kesamaan piksel misalnya kesamaan nilai piksel antara piksel I dan piksel j, dengan jarak (d) dan sudut tertentu. Beberapa fitur yang digunakan dalam penelitian ini sesuai persamaan (7) sampai persamaan (10).

1. Contrast adalah keberadaan variasi atas ke abuan piksel citra.

$$\text{Contrast} = \sum_i \sum_j (i - j)^2 p_{(i,j)} \quad (7)$$

2. Energy/energi merupakan representasi ukuran keseragaman pada citra.

$$\text{Energy} = \sum_i \sum_j p(i, j)^2 \quad (8)$$

3. Homogeneity merupakan representasi ukuran keserbasamaan citra.

$$\text{Homogeneity} = \sum_i \sum_j \frac{p(i,j)}{1+|i-j|} \quad (9)$$

4. Correlation merupakan representasi fitur tekstur terhadap ketergantungan linier derajat keabuan dari setiap piksel yang bertetangga dengan obyek citra.

$$\text{Correlation} = \sum_i \sum_j \frac{(i-\mu_i)(j-\mu_j)p_{(i,j)}}{\sigma_i \sigma_j} \quad (10)$$

2.4 K Nearest Neighbors (KNN)

Algoritma K-Nearest Neighbors (KNN) merupakan metode klasifikasi yang bekerja berdasarkan kedekatan jarak suatu data dengan data yang lain dengan menghitung kedekatan jarak objek. Menurut [12], KNN mencari jarak terdekat antara satu objek dengan objek lainnya. misalnya menggunakan euclidian distance seperti ditunjukkan persamaan (11).

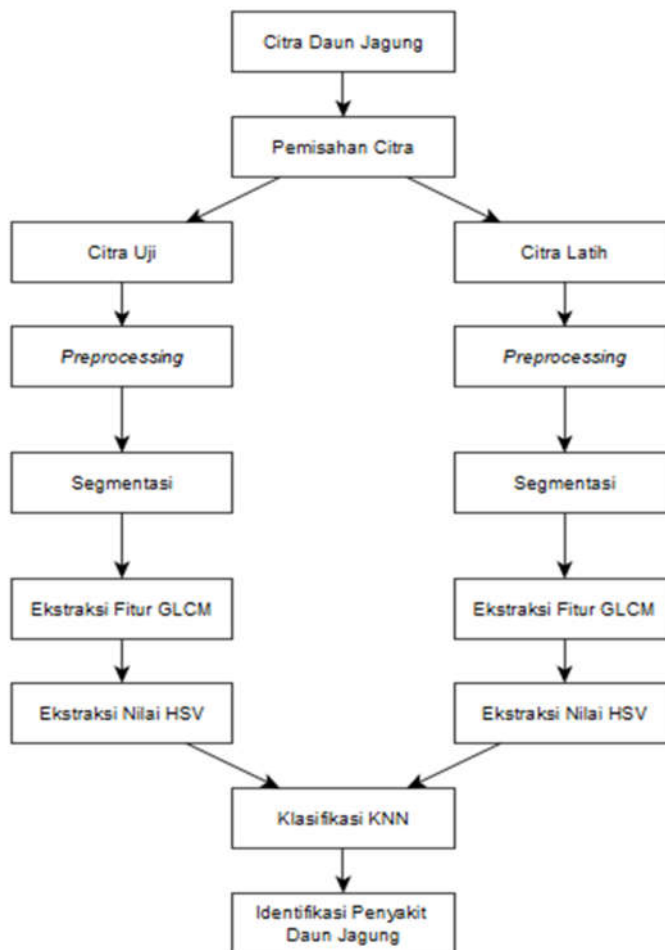
$$\text{Euclidean Distance} = d_{(x,y)} = \sqrt{\sum_{i=1}^p (x_i - y_i)^2} \quad (11)$$

2.5 Pengujian Metode

Pengujian dilakukan dengan menggunakan akurasi yang tujuannya memberikan informasi pada jumlah keberhasilan dari hasil proses klasifikasi terhadap suatu label/kelas yang telah ditentukan dengan seluruh jumlah kelas sesuai persamaan (12).

$$Akurasi = \frac{\text{jumlah citra benar atau dikenali}}{\text{jumlah total citra keseluruhan}} \times 100\% \quad (12)$$

2.6 Tahapan Penelitian



Gambar 2. Alur Penelitian

Berdasarkan Gambar 2, digunakan 200 citra daun jagung dengan sebaran 160 sebagai citra latih dan 40 sebagai citra uji. Setelah dikelompokkan, masing masing citra akan dilakukan *preprocessing* yaitu dengan mengubah background citra menjadi hitam dan mengubah ukuran citra menjadi 256x256 pixel. Setelah dilakukan *preprocessing*, citra dilakukan segmentasi warna untuk mendapatkan area daun yang terinfeksi. Setelah dilakukan segmentasi citra diubah ke *grayscale* untuk mendapatkan nilai matriks kookurensi dilanjutkan mencari nilai energi, kontras, korelasi, dan homogenitas untuk ekstraksi fitur GLCM. Citra segmentasi akan dicari nilai HSV yang digunakan sebagai fitur untuk klasifikasi berupa histogram nilai HSV. Ketika kedua ekstraksi fitur sudah dilakukan nilai tersebut akan dijadikan acuan untuk klasifikasi. Langkah selanjutnya melakukan klasifikasi dengan menggunakan algoritma klasifikasi KNN. Setelah mendapatkan hasil klasifikasi, kemudian hitung akurasi untuk mendapatkan pembuktian tingkat akurasi dari proses klasifikasi yang telah didapatkan. Hasil akurasi yang didapatkan tergantung dari objek citra, algoritma dan *preprocessing* yang digunakan.

3. METODE PENELITIAN

3.1 Pengumpulan Data

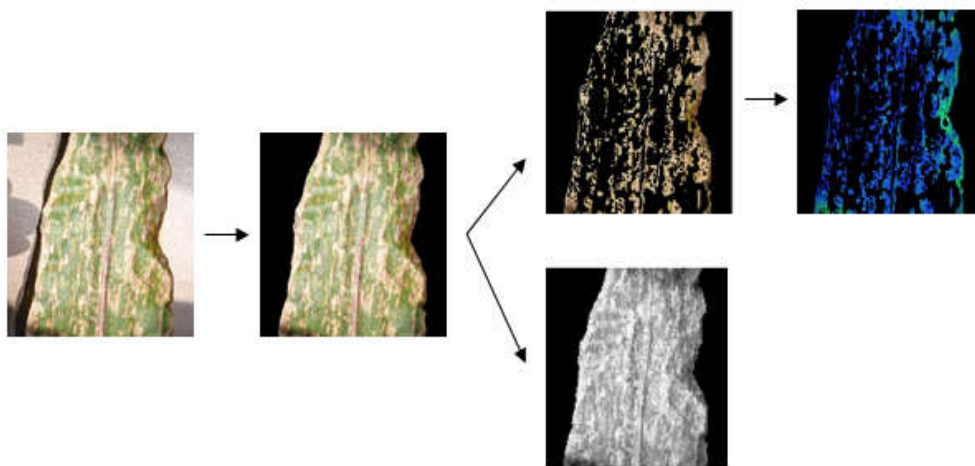
Sebanyak 200 citra secara keseluruhan yang terdiri 4 kelas, yaitu penyakit bercak daun, penyakit hawar daun, penyakit karat daun, dan daun sehat. Dari 200 citra tersebut akan dibagi menjadi 2 bagian yaitu 160 latih dan 40 uji, dimana setiap kelas digunakan 10 citra. Ukuran citra akan di resize menjadi 256 x 256 pixel. Citra tersebut akan diberi label seperti Tabel 1.

Tabel 1. Label Citra

Citra Latih			Citra Uji		
Baris Citra	Nama Kelas	Label	Baris Citra	Nama Kelas	Label
1-40	Penyakit Bercak Daun	Bercak	1-10	Penyakit Bercak Daun	Bercak
41-80	Penyakit Hawar Daun	Hawar	11-20	Penyakit Hawar Daun	Hawar
81-120	Penyakit Karat Daun	Karat	21-30	Penyakit Karat Daun	Karat
121-160	Daun Sehat	Sehat	31-40	Daun Sehat	Sehat

3.2 Pengolahan Citra

Tahap pertama pada pengolahan citra yaitu memotong background citra dan merubah menjadi warna hitam. Setelah itu citra dilakukan segmentasi untuk memperoleh warna HSV, dan mengubah citra *RGB to Gray* untuk membantu proses perhitungan ekstraksi fitur GLCM seperti pada Gambar 3. menggunakan fungsi yang ada didalam MATLAB.



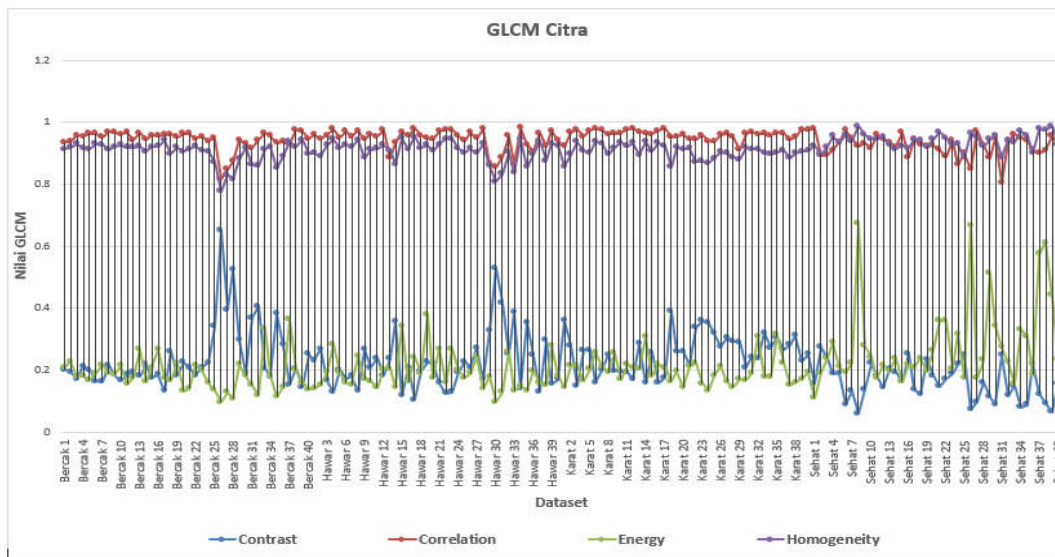
Gambar 3. Proses Konversi Citra

Berdasarkan Gambar 3, setelah melakukan pembagian data untuk latih dan uji, citra akan diolah dengan beberapa tahap, berikut merupakan proses dari pengolahan yang dilakukan. Perubahan background dilakukan pemotongan pada bagian background dan diganti warna background menjadi hitam menggunakan adobe potoshop agar mudah untuk dideteksi. Pada proses segmentasi, dilakukam hanya untuk area yang terinfeksi penyakit sehingga memudahkan dalam pengelompokkan nantinya. Selanjutnya mengubah citra menjadi HSV untuk mengambil nilai Hue Saturation dan Value nya sehingga dapat mengetahui perbedaan warna yang ada pada citra. Kemudian dilanjutkan dengan mengubah citra menjadi grayscale

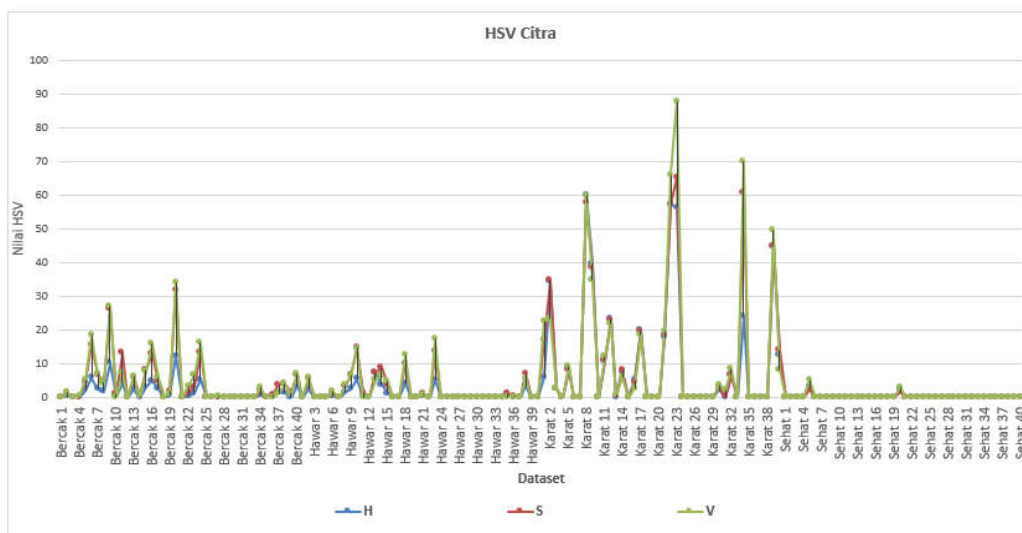
agar bisa dilakukan ekstraksi fitur glm dan mempermudah mengolah citra dengan warna yang sama.

3.3 Hasil Pengujian

Dilakukan ekstraksi fitur untuk setiap citra latih dan uji. Proses kalkulasi GLCM ini menghitung matrik dari tiap citra dengan jarak piksel (d) yang ditentukan dan orientasi dari ke-4 sudut. Citra RGB yang sudah dikonversi akan digunakan sebagai ekstraksi fitur tekstur. Di bawah ini adalah tabel yang sudah dinormalisasi y matriks untuk piksel 8×8 dengan jarak $d=1$ dan sudut 0^0 yang akan digunakan untuk mencari nilai Contrast Energy Correlation dan homogeneity seperti tampak pada Gambar 4. Untuk fitur ekstraksi warna menggunakan HSV sesuai Gambar 5.

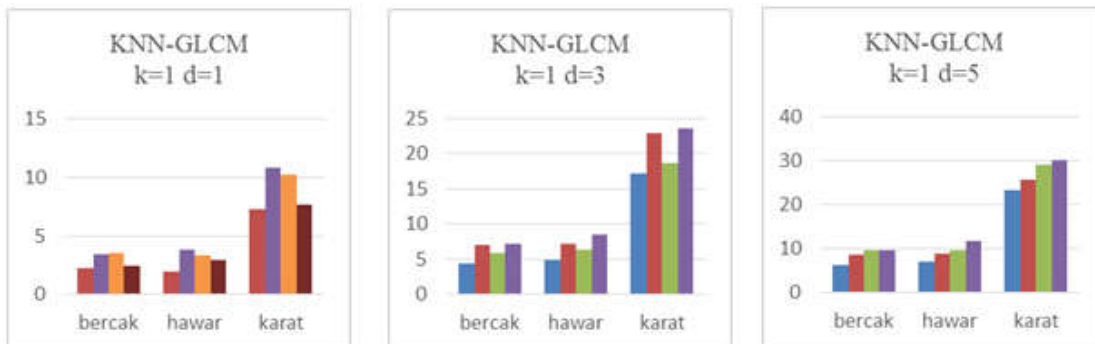


Gambar 4. GLCM pada data uji

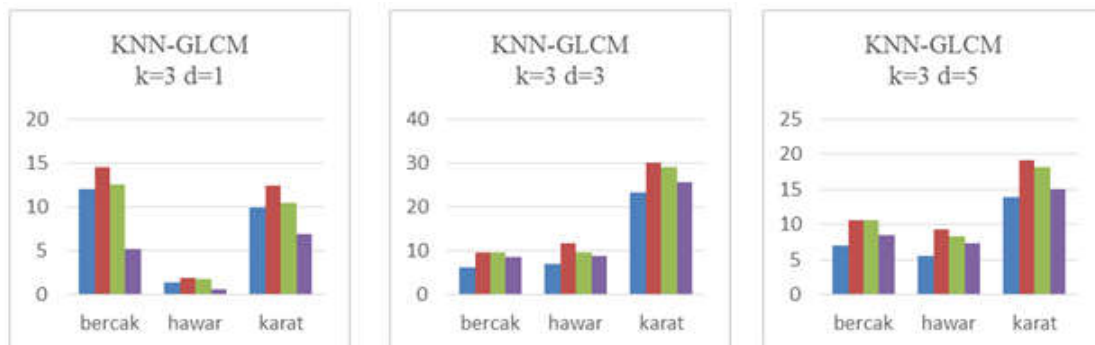


Gambar 5. HSV pada data uji

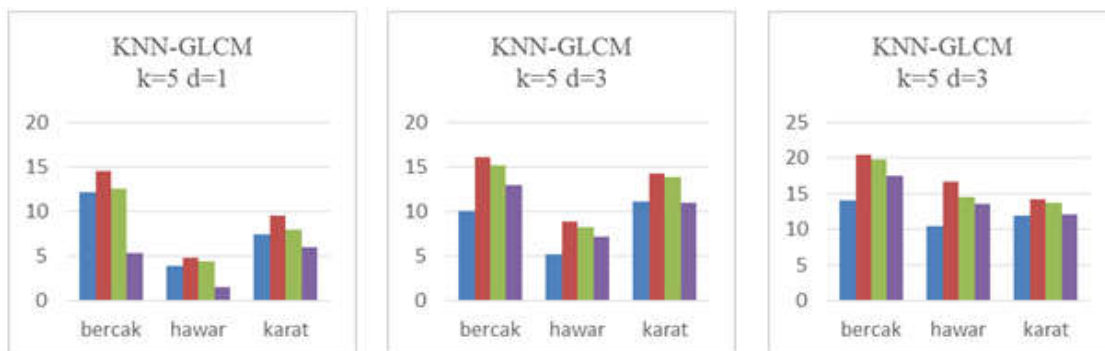
Penelitian kali ini dilakukan pengujian dengan menggunakan algoritma KNN-GLCM dan KNN-GLCM-HSV yang dapat ditunjukkan oleh Gmabar 6 hingga Gambar 11.



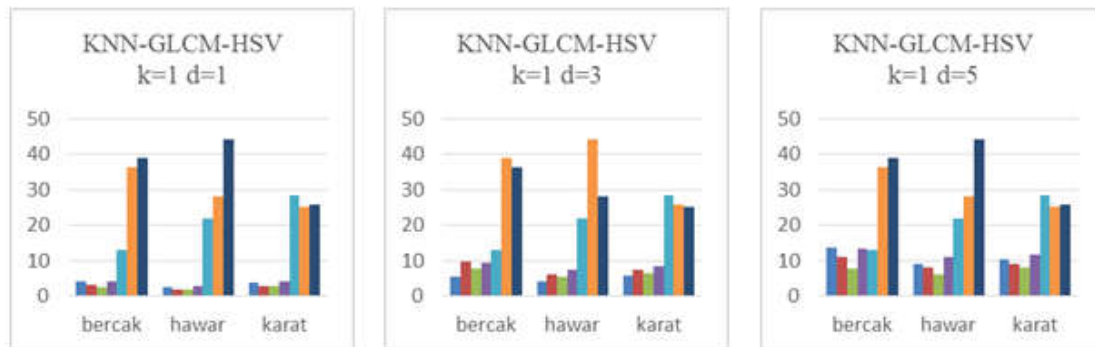
Gambar 6. Grafik KNN GLCM pada k adalah 1 dan jarak piksel 1



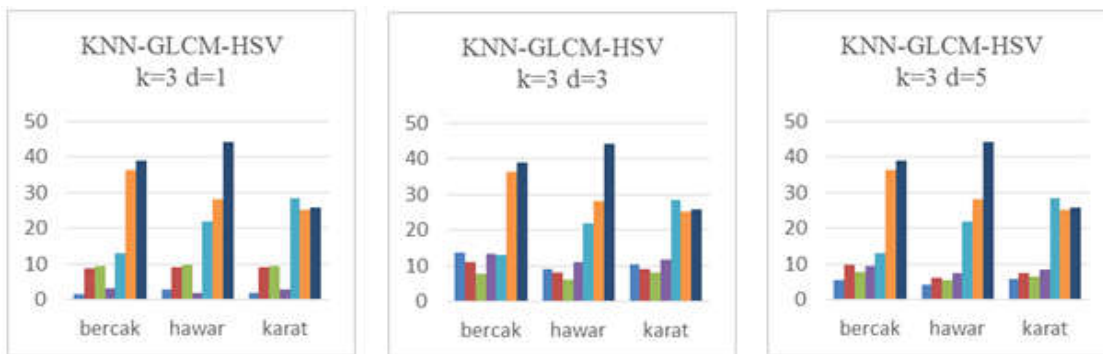
Gambar 7. Grafik KNN GLCM pada k adalah 3 dan jarak piksel 1



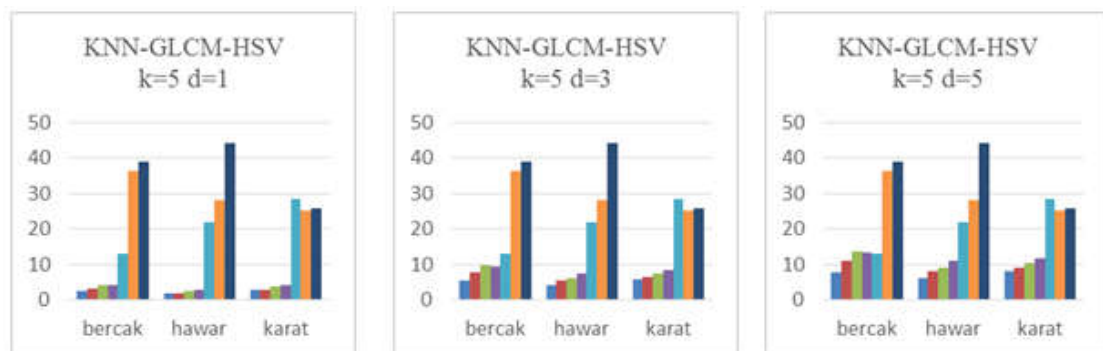
Gambar 8. Grafik KNN GLCM pada k adalah 5 dan jarak piksel 1



Gambar 9. Grafik KNN GLCM HSV pada k adalah 1 dan jarak piksel 1, 3, 5



Gambar 10. Grafik KNN GLCM HSV pada k adalah 3 dan jarak piksel 1, 3, 5



Gambar 11. Grafik KNN GLCM HSV pada k adalah 5 dan jarak piksel 1, 3, 5

Berdasarkan hasil grafik diatas didapatkan akurasi seperti pada Tabel 2, dimana di peroleh nilai akurasi tertinggi pada algoritma KNN GLCM k adalah 5 dan jarak piksel 5 dengan akurasi 60% dan pada algoritma KNN GLCM HSV k adalah 3 dan jarak piksel 1 dengan akurasi 85%. Dan akurasi terendah pada algoritma KNN GLCM k adalah 1 dan jarak piksel 3 dengan akurasi 45% dan pada algoritma KNN GLCM HSV k adalah 3 dan jarak piksel 5 dengan akurasi 70%.

Tabel 2. Perbandingan akurasi pada variasi nilai k dan d

Algoritma	K adalah 1			K adalah 3			K adalah 5		
	Jarak piksel								
	1	3	5	1	3	5	1	3	5
KNN-GLCM	55	45	47.5	52.5	55	57.5	55	52.5	60
KNN-GLCM-HSV	80	77.5	75	85	70	70	80	82.5	80

5. KESIMPULAN

Tujuan penelitian ini adalah mempermudah dalam mengklasifikasikan penyakit dalam daun jagung, dan yang kedua dapat menerapkan ekstraksi warna menggunakan HSV dan GLCM pada algoritma K-Nearest Neighbors (KNN). Dalam dunia industri penelitian kali ini dirasa mampu untuk mengatasi permasalahan yang ada pada saat ini yaitu sulitnya menentukan tingkat kematangan buah mengkudu karena dengan banyaknya data yang ada untuk diolah dibutuhkan suatu alat/program yang dapat meringankan energi manusia, selain itu juga dapat mempersingkat waktu dari kegiatan manusia yang awalnya memilah data secara manual ke dalam kelas yang telah ditentukan dibandingkan dengan menggunakan alat/program yang dapat membantu mengefisienkan sisa waktu untuk kegiatan lain. Dengan demikian, penelitian kali ini akan menggunakan dataset yang berjumlah 200 data dengan rincian 160 data pelatihan

dan 40 data pengujian sehingga didapatkan hasil akurasi dengan k adalah 3 dengan jarak piksel 1 didapatkan hasil akurasi tertinggi sebesar 85% pada algoritma KNN-GLCM-HSV dan pada algoritma KNN-GLCM dengan k adalah 5 dan jarak piksel 5 akurasi tertinggi sebesar 60%

6. SARAN

Untuk meningkatkan akurasi maupun mengefisiensi waktu terdapat beberapa saran yang diajukan penulis antara lain dengan menambah citra dataset dari data pelatihan atau data training maupun dari data sehingga dapat meminimalisir kesalahan saat dilakukan proses klasifikasi. Untuk proses preprocessing dapat menggunakan preprocessing lain yang memiliki peranan sama atau dapat menambahkan dengan proses-proses preprocessing lainnya agar citra data yang diolah dari citra data pelatihan atau data training maupun dari citra data pengujian atau data testing memiliki banyak informasi dan pengetahuan yang dapat dipelajari. Pada proses ekstraksi fitur dapat menggunakan ekstraksi fitur lain yang memiliki peranan sama atau dapat menggunakan ekstraksi fitur lebih dari satu untuk meningkatkan hasil klasifikasi yang maksimal.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada **LPPM Universitas Dian Nuswantoro** yang telah memberi “**dukungan financial**” terhadap penelitian ini khususnya pada skema penelitian Penelitian Dasar Perguruan Tinggi (PTPT) Tahun Anggaran Genap 2020/2021.

DAFTAR PUSTAKA

- [1] M. Syarief and W. Setiawan, “Convolutional neural network for maize leaf disease image classification,” *TELKOMNIKA (Telecommunication Comput. Electron. Control.*, vol. 18, no. 3, p. 1376, Jun. 2020.
- [2] F. Tupamahu, Christyowidiasmoro, and M. H. Purnomo, “Ekstraksi Fitur Tekstur dan Warna Citra Daun Jagung untuk Klasifikasi Penyakit,” in *PROSIDING SEMINAR NASIONAL ILMU KOMPUTER 2014*, 2014.
- [3] F. Liantoni, “Klasifikasi Daun Dengan Perbaikan Fitur Citra Menggunakan Metode K-Nearest Neighbor,” *Ultimatics*, vol. VII, no. 2, pp. 98–104, 2015.
- [4] D. Syahid, Jumadi, and D. Nursantika, “Sistem Klasifikasi Jenis Tanaman Hias Daun Philodendron Menggunakan Metode K-Nearest Neighbor (Knn) Berdasarkan Nilai Hue, Saturation, Value (Hsv),” *J. Online Inform.*, vol. 1, no. 1, pp. 20–23, 2016.
- [5] F. Liantoni and H. Nugroho, “Klasifikasi Daun Herbal Menggunakan Metode Naïve Bayes Classifier Dan K- Nearest Neighbor,” *Junral Simantec*, vol. 5, no. 1, pp. 9–16, 2015.
- [6] O. R. Indriani, E. J. Kusuma, C. A. Sari, E. H. Rachmawanto, and D. R. I. M. Setiadi, “Tomatoes Classification Using K-NN Based on GLCM and HSV Color Space,” in *International Conference on Innovative and Creative Information Technology (ICITech)*, 2017, pp. 1–6.
- [7] R. A. Surya, A. Fadlil, A. Yudhana, M. T. Informatika, P. T. Informatika, and U. A. Dahlan, “Ekstraksi Ciri Metode Gray Level Co-Occurrence Matrix (GLCM) dan Filter Gabor Untuk Klasifikasi Citra Batik Pekalongan,” vol. 02, no. 02, pp. 23–26, 2017.
- [8] I. U. W. Mulyono *et al.*, “Parijoto Fruits Classification using K-Nearest Neighbor Based on Gray Level Co-Occurrence Matrix Texture Extraction,” *J. Phys. Conf. Ser.*, vol. 1501, no. 1, 2020.
- [9] D. M. Mahalakshmi and S. Sumathi, “Brain Tumour Segmentation Strategies Utilizing Mean Shift Clustering and Content Based Active Contour Segmentation,” *ICTACT J. Image Video Process.*, vol. 9, no. 4, pp. 2002–2008, 2019.
- [10] C. A. Sari, M. W. Kuncoro, D. R. I. M. Setiadi, and E. H. Rachmawanto, “Roundness and eccentricity feature extraction for Javanese handwritten character recognition based on K-

- nearest neighbor,” *2018 Int. Semin. Res. Inf. Technol. Intell. Syst. ISRITI 2018*, pp. 5–10, 2018.
- [11] P. Moallem, A. Serajoddin, and H. Pourghassem, “Computer vision-based apple grading for golden delicious apples based on surface features,” *Inf. Process. Agric.*, vol. 4, no. 1, pp. 33–40, 2017.
- [12] H. Nguyen-Quoc and V. T. Hoang, “Rice seed image classification based on HOG descriptor with missing values imputation,” *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 18, no. 4, pp. 1897–1903, 2020.
- [13] E. Budianita, J. Jasril, and L. Handayani, “Implementasi Pengolahan Citra dan Klasifikasi K-Nearest Neighbour Untuk Membangun Aplikasi Pembeda Daging Sapi dan Babi Berbasis Web,” *J. Sains dan Teknol. Ind.*, vol. 12, no. Vol 12, No 2 (2015): Juni 2015, pp. 242–247, 2015.
- [14] R. Mehrotra, M. A. Ansari, and R. Agrawal, “A Novel Scheme for Detection Feature Extraction of Brain Tumor by Magnetic Resonance Modality Using DWT SVM,” *2020 Int. Conf. Contemp. Comput. Appl. IC3A 2020*, pp. 225–230, 2020.
- [15] I. Nurhaida, H. Wei, R. A. M. Zen, R. Manurung, and A. M. Arymurthy, “Texture fusion for batik motif retrieval system,” *Int. J. Electr. Comput. Eng.*, vol. 6, no. 6, pp. 3174–3187, 2016.