

EKTRAKSI FITUR GLCM PADA K-NN DALAM MENGGKASIFIKASI MOTIF BATIK

Cahaya Jatmoko¹, Daurat Sinaga²

¹Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro
Jl. Imam Bonjol 207 Semarang, 50131
e-mail: cahaya.jatmoko@dsn.dinus.ac.id, parendeborngin@gmail.com

ABSTRAK

Batik dipolakan menggunakan metode GLCM yang akan menghasilkan fitur *energy*, *contrast*, *correlation*, *homogeneity* dan *entropy*. Kemudian fitur-fitur tersebut dijadikan input untuk proses klasifikasi data training dan data testing menggunakan metode KNN dengan menggunakan pencarian jarak *eclidean*. Klasifikasi dilakukan sebanyak 2 kali, perbedaannya terletak pada penggunaan jumlah fitur pada saat ekstraksi fitur tekstur. Yang pertama menggunakan 4 fitur yaitu *energy*, *contrast*, *correlation*, dan *homogeneity*. Klasifikasi selanjutnya menggunakan 5 fitur yang memberikan informasi nilai-nilai *energy*, *contrast*, *correlation*, *homogeneity*, dan *entropy*. Dari kedua klasifikasi tersebut, akan dikomparasi mana yang menghasilkan akurasi terbaik. Data training dan data testing diuji coba menggunakan perhitungan *Recognition Rate* untuk evaluasi sistem

Kata Kunci: Klasifikasi citra, motif batik, K-NN, GLCM

1. PENDAHULUAN

Indonesia adalah negara yang memiliki semboyan “Berbeda-beda tetapi tetap satu jua”, baik perbedaan suku, agama, budaya, dan bahasa. Salah satu budaya yang telah diakui dunia yaitu batik. Batik Indonesia ditetapkan sebagai Warisan Kemanusiaan untuk Budaya dan Nonbendawi (*Masterpieces of the Oral and Intangible Heritage of Humanity*), yang telah diakui UNESCO sebagai hak intelektual bangsa Indonesia pada tanggal 2 Oktober 2009 [1]. Batik menurut Kamus Besar Bahasa Indonesia adalah kain bergambar yang pembuatannya secara khusus dengan menuliskan atau menerakan malam pada kain itu, kemudian pengolahannya melalui proses tertentu.

Tidak semua lapisan masyarakat mengetahui jenis dan motif batik, apalagi membedakan motif batik dalam bentuk citra digital. Model pengenalan batik sudah selayaknya menggunakan pemanfaatan teknologi, khususnya teknologi pengolahan citra digital. Pengolahan citra menawarkan model pengenalan awal hingga pengklasifikasi pola batik.

Pengolahan Citra Digital (PCD) [2] merupakan disiplin ilmu yang mempelajari tentang tranformasi gambar (translasi, rotasi, transformasi geometrik, skala), perbaikan kualitas gambar (transformasi warna, peningkatan kontras, restorasi citra), melakukan pemilihan citra ciri (*feature images*) yang optimal untuk tujuan analisis, melakukan reduksi atau kompresi data untuk tujuan penyimpanan data, waktu proses data, dan trasmisi data. Selain itu, PCD juga melakukan proses penarikan deskripsi objek atau informasi yang terkandung pada citra. Teknik pengolahan citra digital secara umum ada 3 tingkatan, pertama *low-level processing* (pengolahan tingkat rendah) yang merupakan operasi dasar dari PCD seperti perbaikan citra, restorasi citra, dan noise. Kemudian *mid-level processing* (pengolahan tingkat menengah) yang meliputi deskripsi objek, segmentasi citra, dan klasifikasi objek. Dan yang terakhir adalah *high-level processing* (pengolahan tingkat tinggi) yang meliputi analisis citra.

Model pengenalan yang sering digunakan disebut dengan teknik klasifikasi citra digital. Terdapat beberapa algoritma yang sering digunakan sebagai alat bantu antara lain K-Nearest Neighbor (K-NN) [3], Fuzzy C-Means, Naïve Bayes [4], Algoritma Genetika, Support Vector Machine [5], Jaringan Syaraf Tiruan seperti Leaning Vector Quantization [6] dan Multilayer Perceptron.

Pada penelitian yang telah dilakukan oleh Liantoni (Laintoni & Nugroho, 2015) yang membahas tentang klasifikasi daun berdasarkan bentuk tepi daun. Untuk mempermudah proses ekstraksi fitur dan klasifikasi maka dilakukan perbaikan citra terlebih dahulu menggunakan filter median dan erosi. Ekstraksi fitur pada penelitian ini menggunakan metode invariant moment, sedangkan klasifikasinya menggunakan algoritma KNN. Hasil akurasi rata-rata yang didapatkan yaitu sekitar 86,67%. Sedangkan Prahudaya telah melakukan klasifikasi mutu jambu biji yang dibagi menjadi mutu super, kelas A, kelas B dan luar mutu. Fitur yang digunakan untuk proses klasifikasi ada 7 fitur yaitu RGB (red, green, blue), luas cacat, dan GLCM dengan 3 buah fitur yaitu *energy*, *homogeneity*, dan *contrast*. Akurasi terbaik yang didapat pada penelitian tersebut yaitu 91,25% dengan nilai K=3.

Proses klasifikasi citra digital terdiri dari tahap *preprocessing* yang merubah citra batik berwarna menjadi citra grayscale dan merubah ukuran citra. Kemudian dilakukan tahap ekstraksi fitur untuk mendapatkan informasi pada citra, dan yang terakhir tahap klasifikasi untuk pengelompokan sesuai kelas. Dalam penelitian sebelumnya telah diusulkan beberapa teknik klasifikasi, salah satunya ada penelitian Indriani dkk [7] dalam mengklasifikasi tingkat

kematangan buah tomat berdasarkan teksturnya menggunakan GLCM dengan *output* 4 fitur yaitu contrast, homogeneity, correlation, dan energy yang menghasilkan akurasi 100% pada K=3 dan K=7.

2. TINJAUAN PUSTAKA

2.1 Tinjauan Studi

Beberapa penelitian yang berhubungan dengan topik klasifikasi citra digital dapat di lihat pada Tabel 1.

Tabel 1. Tinjauan Studi Klasifikasi Citra.

Tahun	Penulis	Algoritma klasifikasi				Ekstraksi Fitur			Hasil
		KNN	SVM	RGB	GLCM	Invarian moment	Color moment	Color histogram	
2017	Indrian,dk k	√			√				Akurasi mencapai 100% ketika p=9 dan k=3.
2017	Prahudaya dan Harjoko	√		√	√				Akurasi tertinggi 91,25% apda K=3
2017	Kusuma, dkk	√	√		√		√	√	Hasil terbaik pada klasifikasi SVM dan ekstraksi fitur warna
2015	Liantoni, Febri	√				√			Akurasi terbaik 86,67%.

Berdasarkan Tabel 1, pada penerapan algoritma klasifikasi KNN dan di dapatkan hasil akurasi yang didapatkan cukup baik, semuanya mempunyai tingkat akurasi diatas 85%. Maka pada penelitian ini akan digunakan metode KNN untuk mengklasifikasikan 5 motif batik tulis lasem yaitu motif aseman, motif kawung baganan, motif latohan, motif sidomukti, dan motif naga.

2.2 K-Nearest Neighbor

K-NN adalah metode untuk mengklasifikasikan objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek. K-NN adalah metode yang menggunakan algoritma supervised dimana bertujuan untuk menemukan pola baru dalam data dengan menghubungkan pola data yang sudah ada dengan data yang baru [3]. K-NN akan mengkonsolidasikan hasil perhitungan dengan data latih yang memiliki jumlah kerabat terbanyak di kisaran-kisaran (k) yang ditentukan. Jarak antara data latih dan data uji dapat dihitung dengan berbagai metode, termasuk menggunakan persamaan Euclidean.

$$E(i, j) = \sqrt{\sum (x_i - x_j)^2} \tag{1}$$

Dimana E (i, j) merupakan jarak euclidean antara vektor i dan vektor j, kemudian k adalah nilai keanggotaan, selain itu n adalah jumlah fitur pada vektor i dan j seperti ditunjukkan pada [8].

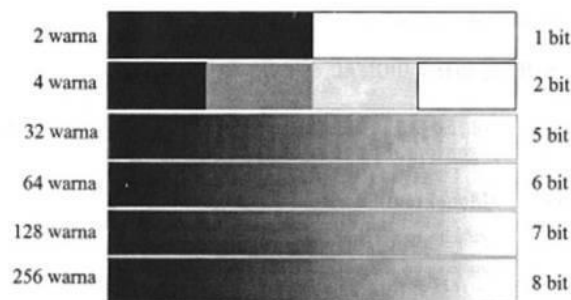
2.3 Batik

Batik adalah warisan nenek moyang yang memiliki nilai seni tinggi dan telah menjadi bagian dari budaya Indonesia (khususnya Jawa) sejak lama. Kata batik berasal dari bahasa Jawa, yaitu “tik” yang berarti titik/matik (kata kerja, membuat titik) yang kemudian berkembang menjadi batik. Salah satu bentuk seni kuno dalam pewarnaan kain batik yaitu teknik perintang pewarnaan menggunakan malam.

Di Indonesia, semua batik yang dihasilkan sebelum awal abad XX adalah batik tulis, yaitu batik tradisional yang diproduksi dengan teknik tulisan tangan menggunakan canting dan malam. Adanya globalisasi yang mengenalkan teknik otomatisasi, membuat adanya batik jenis baru yaitu batik cap dan batik cetak. Pada mulanya warna dan corak batik sangat terbatas, tetapi batik pesisir menyerap banyak pengaruh asing dari pedagang luar negeri dan penjajah.

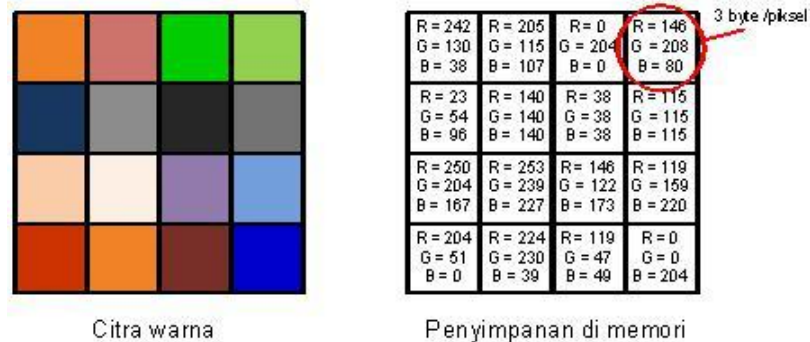
2.4 Citra Grayscale dan Warna

Banyak warna pada citra tergantung pada jumlah bit per piksel yang disediakan memori untuk menampung kebutuhan warna, semakin banyak bit yang disediakan maka gradasi warna yang didapatkan dari hitam ke putih (atau sebaliknya) akan semakin banyak sehingga berdampak pada warna citra yang semakin halus jika gradasi warna lebih banyak.



Gambar 1 Contoh perbandingan gradasi warna 1 bit, 2 bit, 5 bit, 6 bit, 7 bit, dan 8 bit.

Setiap piksel citra warna mempresentasikan 3 warna dasar yaitu red (merah), green (hijau), dan blue (biru) yang masing-masing warna menggunakan penyimpanan 8 bit (1 byte), jadi setiap warna mempunyai gradasi sebanyak 256 warna. Berarti satu piksel citra warna menggunakan penyimpanan 3 byte yang memungkinkan adanya kombinasi lebih dari 16 juta warna, sehingga bisa dikatakan hampir semua warna tercakup di dalamnya, sebab itulah dinamakan true color.

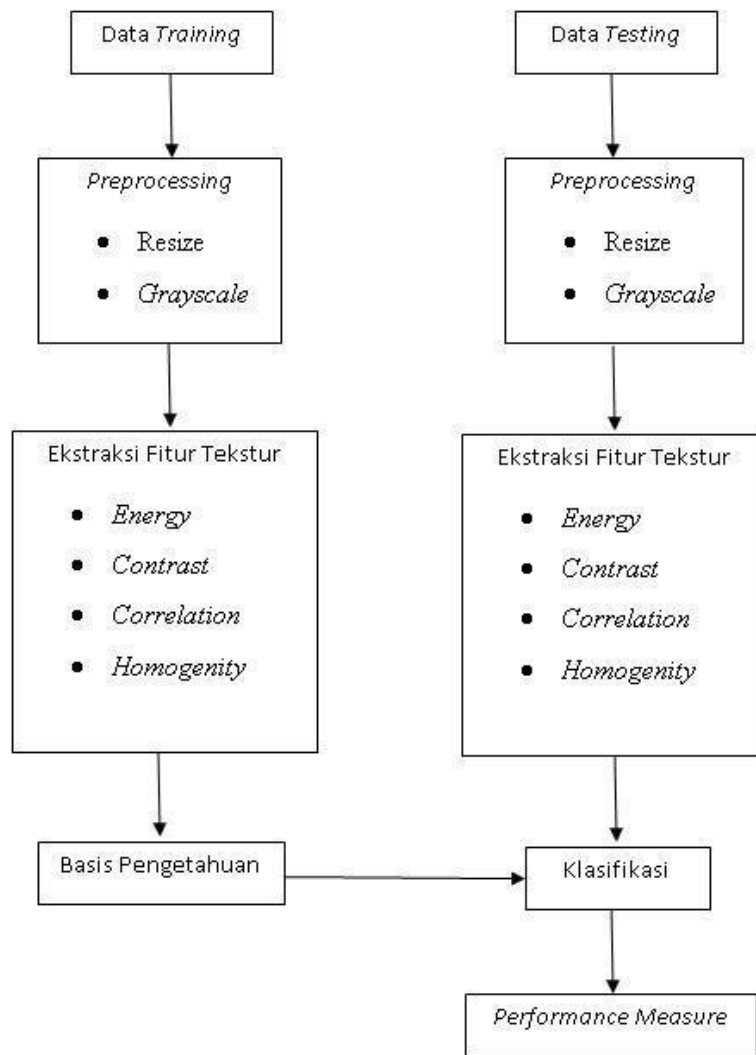


Gambar 2 Contoh penyimpanan citra warna di dalam memori.

Berbeda dari citra-citra sebelumnya, citra berindeks terdiri dari matrik citra dan matrik color map. Matrik map digunakan untuk menunjukkan alamat warna citra disimpan, sedangkan matrik color map menunjukkan peta warna yang digunakan sebagai acuan.

3. METODE PENELITIAN

Dalam makalah ini, skema eksperimen yang dilakukan seperti dalam Gambar 3.



Gambar 3 Skema eksperimen.

Pada Gambar 3, fitur yang digunakan untuk ekstraksi fitur tekstur yaitu energy, contrast, correlation, dan homogeneity. Fitur-fitur tersebut dipilih dengan mengacu pada penelitian Indriani dkk [8] yang mengklasifikasi kematangan buah tomat berdasarkan tekstur yang menggunakan 4 fitur GLCM dengan akurasi 100%. Evaluasi digunakan untuk mengukur kinerja sisten klasifikasi citra motif batik tulis lasem, berupa prosentase banyaknya data yang diklasifikasikan dengan benar. Tingkat akurasi dihitung menggunakan Recognition Rate.

4. HASIL DAN PEMBAHASAN

Pada klasifikasi yang menggunakan 4 fitur GLCM, telah dilakukan uji coba sebanyak 3 kali. Percobaan pertama menggunakan 135 data training dan 15 data testing, percobaan kedua menggunakan 120 data training dan 30 data testing, sedangkan percobaan yang ketiga menggunakan 100 data training dan 50 data testing.

Tabel 2 Hasil ekstraksi fitur GLCM pada sudut 0°.

Nomor	Motif Batik	Energy	Contrast	Homogeneity	Correlation
1	Aseman	0,000232	1469,63	0,139428	0,000199
2		0,000254	1366,946	0,144029	0,000197
3		0,000235	1247,132	0,135199	0,000215
:					
:					

28	Kawung Baganan	0,000452	813,2013	0,198025	0,000329
29		0,000314	710,6526	0,220355	0,000301
30		0,00039	750,0269	0,191741	0,000338
:					
:					
55	Latohan	0,00044	1211,476	0,196586	0,000253
56		0,000405	1188,809	0,182069	0,000257
57		0,000628	1049,954	0,235165	0,000276
:					
:					
82	Sidomukti	0,000127	1370,847	0,080822	0,000289
83		0,000113	1259,043	0,088844	0,000297
84		0,000124	1298,914	0,091933	0,000285
:					
:					
109	Naga	0,001396	787,7802	0,26729	0,000219
110		0,002959	547,3127	0,33972	0,000214
111		0,002487	848,895	0,31235	0,000191
:					
:					
136	Aseman	0,000235	1305,36	0,161849	0,000188
137		0,000235	981,9603	0,187163	0,0002
138		0,000175	662,695	0,161859	0,00031
139	Kawung Baganan	0,000362	710,366	0,213454	0,000306
140		0,000422	1039,553	0,156521	0,000274
141		0,0003	961,6573	0,15793	0,000206
142	Latohan	0,000365	1027,961	0,213517	0,000247
143		0,000498	1065,086	0,196421	0,000172
144		0,000206	1138,18	0,15167	0,000271
145	Sidomukti	7,31E-05	1597,893	0,079214	0,000248
146		0,000143	1950,31	0,112511	0,0002
147		0,00024	1872,91	0,116288	0,000191
148	Naga	0,00127	653,5124	0,301188	0,000188
149		0,001227	640,454	0,303611	0,000193
150		0,001545	1044,251	0,264581	0,000175

Data training dan data testing dihitung kedekatannya menggunakan *ecudiance distance*. Percobaan dilakukan dua kali, yang pertama data testing diambil dari data training, kemudian yang kedua data testing berbeda dengan data training. Supaya hasil klasifikasi dapat dikenali, maka diperlukan pengelompokan pada masing-masing motif batik tulis lasem sebagai berikut :

Tabel 3 Penamaan citra.

Nama Citra	Nama Motif	Kelompok
1_1.jpg	Aseman	1

1_2.jpg	Kawung Baganan	2
1_3.jpg		
1_4.jpg		
1_5.jpg	Latohan	3
1_6.jpg		
1_7.jpg		
1_8.jpg	Sidomukti	4
1_9.jpg		
1_10.jpg		
1_11.jpg	Naga	5
1_12.jpg		
1_13.jpg		
1_14.jpg		
1_15.jpg		

Tabel 4 Uji coba menggunakan data testing=data training.

Nama Citra	Target	Hasil Uji	Keterangan
1_1.jpg	1	1	TRUE
1_2.jpg	1	1	TRUE
1_3.jpg	1	1	TRUE
2_1.jpg	2	2	TRUE
2_2.jpg	2	2	TRUE
2_3.jpg	2	2	TRUE
3_1.jpg	3	3	TRUE
3_2.jpg	3	3	TRUE
3_3.jpg	3	3	TRUE
4_1.jpg	4	4	TRUE
4_2.jpg	4	4	TRUE
4_3.jpg	4	4	TRUE
5_1.jpg	5	5	TRUE
5_2.jpg	5	5	TRUE
5_3.jpg	5	5	TRUE

Langkah tersebut digunakan untuk menguji sistem, apakah dengan menggunakan data yang sama dapat menghasilkan akurasi klasifikasi 100%. Jika telah menghasilkan akurasi 100% seperti pada Tabel 4, maka sistem layak masuk pada tahap selanjutnya.

Tabel 5 Uji coba menggunakan data training 135 dan data testing 15 dengan k=1.

Nama Citra	Target	Hasil Uji	Keterangan
1_1.jpg	1	1	TRUE
1_2.jpg	1	1	TRUE
1_3.jpg	1	2	FALSE
2_1.jpg	2	2	TRUE
2_2.jpg	2	2	TRUE
2_3.jpg	2	1	FALSE
3_1.jpg	3	1	FALSE
3_2.jpg	3	3	TRUE
3_3.jpg	3	1	FALSE
4_1.jpg	4	4	TRUE
4_2.jpg	4	4	TRUE
4_3.jpg	4	4	TRUE
5_1.jpg	5	5	TRUE
5_2.jpg	5	5	TRUE
5_3.jpg	5	5	TRUE

Tabel 6 Uji coba menggunakan 135 data training dan 15 data testing dan k=3.

Nama Citra	Target	Hasil Uji	Keterangan
1_1.jpg	1	2	FALSE
1_2.jpg	1	1	TRUE
1_3.jpg	1	2	FALSE
2_1.jpg	2	2	TRUE
2_2.jpg	2	2	TRUE
2_3.jpg	2	1	FALSE
3_1.jpg	3	1	FALSE
3_2.jpg	3	1	FALSE
3_3.jpg	3	1	FALSE
4_1.jpg	4	4	TRUE
4_2.jpg	4	4	TRUE
4_3.jpg	4	4	TRUE
5_1.jpg	5	5	TRUE
5_2.jpg	5	5	TRUE
5_3.jpg	5	5	TRUE

Tabel 7 Uji coba menggunakan 135 data training dan 15 data testing dan k=5.

Nama Citra	Target	Hasil Uji	Keterangan
1_1.jpg	1	1	TRUE
1_2.jpg	1	1	TRUE
1_3.jpg	1	2	FALSE
2_1.jpg	2	2	TRUE
2_2.jpg	2	2	TRUE
2_3.jpg	2	1	FALSE
3_1.jpg	3	1	FALSE
3_2.jpg	3	1	FALSE
3_3.jpg	3	1	FALSE
4_1.jpg	4	4	TRUE
4_2.jpg	4	4	TRUE
4_3.jpg	4	4	TRUE
5_1.jpg	5	5	TRUE
5_2.jpg	5	5	TRUE
5_3.jpg	5	5	TRUE

Uji coba pertama, data testing diambil dari data training. Hasil perhitungannya sebagai berikut :

11 Σ
11 11+100%
1 = 100%

Pada pengujian selanjutnya, dilakukan pengujian sebanyak tiga kali dengan jumlah data training yang berbeda-beda. Uji coba kedua, menggunakan data testing 15 dan data training 135. Hasil perhitungannya dapat dilihat sebagai berikut :

11 Σ
11 11+100%
2 = 100%

Pada uji coba yang ketiga, data yang digunakan sebanyak 30 data testing dan 120 data training. Hasil perhitungannya dapat dilihat sebagai berikut :

11 Σ
11 11+100%
2 = 100%

Pada uji coba yang terakhir yaitu uji coba empat, menggunakan data testing sebanyak 50 citra dan data training sebanyak 100 citra. Hasil perhitungannya dapat dilihat sebagai berikut :

Uji coba yang ke empat memperoleh hasil akurasi 66%, dimana prosentase hasil klasifikasi untuk masing-masing motif sebagai berikut : motif aseman dapat dikenali sebanyak 40%, motif kawung dan sidomukti dapat dikenali sebanyak 90%, motif latohan dapat dikenali sebanyak 50%, dan motif naga dapat dikenali sebanyak 60%.5. kesimpulan

5. KESIMPULAN

Uji coba pertama menghasilkan akurasi tertinggi yaitu 100%. Karena hasil akurasi mencapai 100%, maka sistem layak dilanjutkan untuk menguji data testing yang berbeda dengan data training. Pada uji coba kedua memperoleh akurasi 73,33%, dimana prosentase hasil klasifikasi untuk masing-masing motif sebagai berikut : motif aseman, kawung baganan, dan latohan dapat dikenali sebanyak 66,67%, sedangkan motif sidomukti dan motif naga dapat dikenali sebanyak 100%. Uji coba ketiga memperoleh hasil akurasi 66,67%, dimana prosentase klasifikasi untuk masing-masing motif batik sebagai berikut : motif aseman dan latohan dapat dikenali sebanyak 50%, motif kawung dan sidomukti dapat dikenali sebanyak 83,33%, dan motif naga dapat dikenali sebanyak 66,67%. Hasil akurasi terbaik yaitu 100% pada percobaan pertama menggunakan 15 data *testing* dan 135 data training. Semakin banyak data *training*, hasil akurasi semakin baik.

DAFTAR PUSTAKA

- [1] "Indonesian Batik," 2018. [Online]. Available: <https://ich.unesco.org/en/RL/indonesian-batik-00170>.
- [2] "Pemrograman Matlab : Pengolahan Citra Digital, Pengolahan Video, Pengenalan Pola, dan Data Mining," 26 July 2017. [Online]. Available: <https://pemrogramanmatlab.com/2017/07/26/pengolahan-citra-digital/>. [Accessed 02 April 2019].
- [3] T. Y. Prahudaya and A. Harjoko, "Metode Klasifikasi Mutu Jambu Biji Menggunakan KNN Berdasar Fitur Warna dan Tekstur," *Jurnal Tekno Sains*, pp. 113-123, 2017.
- [4] L. Nugroho, "Klasifikasi Daun Dengan Perbaikan Fitur Citra Menggunakan Metode K-Nearest Neighbor," *ULTIMATICS*, 2015.
- [5] T. Y. Prahudaya and A. Harjoko, "METODE KLASIFIKASI MUTU JAMBU BIJI MENGGUNAKAN KNN BERDASARKAN FITUR WARNA DAN TEKSTUR," *Jurnal Tekno Sains*, 2017.
- [6] U. Sudiby, D. P. Kusumaningrum, E. H. Rachmawanto and C. A. Sari, "OPTIMASI ALGORITMA LEARNING VECTOR QUANTIZATION (LVQ) DALAM PENGKLASIFIKASIAN CITRA DAGING SAPI DAN DAGING BABI BERBASIS GLCM DAN HSV," *SIMTERIS*, 2018.
- [7] O. R. Indriani, E. J. Kusuma, C. A. Sari, E. H. Rachmawanto and D. Setiadi, "Tomatoes classification using K-NN based on GLCM and HSV color space," in *ICITech*, Salatiga, 2017.
- [8] 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)*, 1-6, 2017.