

Klasifikasi Kematangan Buah Apel Berdasarkan Warna Dan Tekstur Menggunakan Algoritma K-Nearest Neighbor

Classification of Apple Fruit Ripeness Based on Color and Texture Using the K-Nearest Neighbor Algorithm

Fandy Indra Pratama^{*1}, Akhmad Pandu Wijaya², Hasti Pratiwi³, Avira Budianita⁴

^{1,2,3,4} Program Studi Teknik Informatika, Fakultas Teknik, Universitas Wahid Hasyim

e-mail: ^{*}fandy@unwahas.ac.id, ²akhmadpandu@unwahas.ac.id, ³hastipratiwi@unwahas.ac.id,

⁴avirabudianita@unwahas.ac.id

Abstrak

Apel (*Malus Domestica*) adalah salah satu jenis buah yang unggul dan sangat digemari dan dikonsumsi masyarakat. Buah apel banyak digemari masyarakat karena rasanya yang bervariasi. Buah apel manalagi yang masih belum matang memiliki warna hijau dan untuk apel manalagi yang sudah matang memiliki warna hijau kekuningan. Hal tersebut tentu akan sedikit menyulitkan untuk membedakan antara buah apel manalagi yang belum matang maupun yang sudah matang apabila tidak terlalu memiliki ilmu mengenai dunia pertanian terutama pada buah apel. Sebab orang awam dirasa akan sulit membedakan. Sehingga akan diperoleh penilaian yang berbeda oleh setiap individu yang menyebabkan tingkat akurasi yang berbeda. Oleh karena itu dibutuhkan suatu teknologi yang dapat menentukan nilai tingkat akurasi dari pengklasifikasian kematangan buah apel manalagi sehingga memperoleh nilai yang konsisten. Penelitian ini bertujuan untuk menghasilkan prototype untuk mengklasifikasikan buah apel menggunakan Metode K-Nearest Neighbor berdasarkan tingkat kematangan yang direpresentasikan dari sisi warna dan teksturnya dan membantu memudahkan masyarakat dalam menentukan jenis buah yang matang, sedang, dan mentah. Diharapkan adanya teknologi ini dapat mempermudah manusia dalam mengklasifikasikan tingkat kematangan buah apel secara akurat. Metode pengujian ini menggunakan 12 buah apel yang terdiri dari matang, sedang, dan mentah yang menghasilkan tingkat kematangan buah apel sebesar 91,6667% dari total akurasi.

Kata kunci: *Apel, HSV, Klasifikasi, KNN*

Abstract

Apple (*Malus Domestica*) is one of the superior fruit types and is very popular and consumed by the community. Apples are much-loved by the public because of their varied tastes. Manalagi apples that are still immature have a green color and ripe Manalagi apples have a yellowish green color. This of course will be a little difficult to distinguish between unripe and ripe apples if you don't have too much knowledge about agriculture, especially apples. Because ordinary people feel it will be difficult to distinguish. So that a different assessment will be obtained by each individual which causes a different level of accuracy. Therefore we need a technology that can determine the level of accuracy of classifying the ripeness of manalagi apples so as to obtain a consistent value. This study aims to produce a prototype to classify apples using the K-Nearest Neighbor Method based on the maturity level represented in terms of color and texture and to help make it easier for people to determine the types of ripe, medium and unripe fruit. It is hoped that this technology will make it easier for humans to accurately classify the maturity level of apples. This test method uses 12 apples consisting of ripe, medium and unripe which produces an apple maturity level of 91.6667% of the total accuracy.

Keywords: *Apples, HSV, Classification, KNN*

Informasi Artikel:

Submitted: April 2023, **Accepted:** Mei 2023, **Published:** Mei 2023

ISSN: 2685-4902 (media online), **Website:** <http://jurnal.umus.ac.id/index.php/intech>

PENDAHULUAN

Apel (*Malus Domestica*) adalah salah satu jenis buah yang unggul dan sangat digemari dan dikonsumsi masyarakat. Buah apel banyak digemari masyarakat karena rasanya yang bervariasi. Buah apel sendiri memiliki banyak nutrisi dan berbagai macam vitamin diantaranya lemak baik serta karbohidrat, protein, vitamin C, vitamin A, vitamin B1, Vitamin B2 dan masih banyak lagi. Berdasarkan data yang dirilis oleh direktorat jenderal holtikultura kementerian pertanian tahun 2015 [1], produk pertanian apel memiliki kontribusi yang cukup tinggi di tahun 2014. Hal ini ditunjukkan bahwa buah apel di pulau Jawa telah diproduksi sebanyak 242.763 ton dari 249.915 ton produksi di Indonesia. Maka dapat disimpulkan bahwa buah apel menjadi salah satu produk holtikultura yang paling banyak diminati dan dikonsumsi oleh masyarakat, baik dikonsumsi secara langsung maupun yang telah diolah dalam bentuk seperti manisan, keripik, dodol dan minuman.

Salah satu jenis apel yaitu apel manalagi, memiliki ciri yang berbeda dari apel-apel lainnya yaitu terdapat pada tingkat warna kematangan buah tersebut. Buah apel manalagi yang masih belum matang memiliki warna hijau dan untuk apel manalagi yang sudah matang memiliki warna hijau kekuningan. Hal tersebut tentu akan sedikit menyulitkan untuk membedakan antara buah apel manalagi yang belum matang maupun yang sudah matang apabila tidak terlalu memiliki ilmu mengenai dunia pertanian terutama pada buah apel. Sebab orang awam dirasa akan sulit membedakan. Sehingga akan diperoleh penilaian yang berbeda oleh setiap individu yang menyebabkan tingkat akurasi yang berbeda. Sehingga akan diperoleh penilaian yang berbeda oleh setiap individu yang menyebabkan tingkat akurasi yang berbeda [2], [3]. Oleh karena itu dibutuhkan suatu teknologi yang dapat menentukan nilai tingkat akurasi dari pengklasifikasian kematangan buah apel manalagi sehingga memperoleh nilai yang konsisten.

Dalam mengklasifikasikan jenis buah apel itu terdapat algoritma klasifikasi citra yang dapat digunakan salah satunya adalah *K-Nearest Neighbor* (K-NN). Algoritma K-NN merupakan metode klasifikasi yang menentukan label dari sebuah objek baru berdasarkan mayoritas label dari jarak terdekat k dalam kelompok data latih [4], [5]. Secara umum tahapan dalam proses klasifikasi citra digital yaitu akuisisi citra, pra pengolahan citra, ekstraksi ciri/fitur, pelatihan, pengujian dan pengukuran akurasi. Tahapan mengekstrak ciri atau informasi dalam citra digital sangat mempengaruhi untuk mengenali objek yang ada dalam citra tersebut. Semakin banyak ciri yang diekstrak akan mempengaruhi tingkat akurasi klasifikasi citra. Terdapat bermacam-macam ekstraksi dalam ekstraksi ciri citra yaitu warna dan tekstur.

Penelitian terkait, tentang Klasifikasi Tingkat Kematangan Buah Berdasarkan Warna Dan Tekstur Menggunakan Metode K-Nearest Neighbor dan Nearest Mean Classifier (NMC) [6]. Penelitian ini bertujuan untuk menghasilkan Algoritma K-Nearest Neighbor dapat melakukan klasifikasi tingkat kematangan buah apel manalagi dengan tingkat nilai akurasi paling tinggi adalah dengan menggunakan fitur warna saja dan kombinasi fitur warna dan tekstur dengan masing-masing menggunakan 1-NN yang menghasilkan persentase kebenaran 73%. Hasil pengujian Akurasi yang diperoleh secara global menggunakan Nearest mean classifier menghasilkan tingkat akurasi sebanyak 70%. Berdasarkan latar belakang tersebut penulis tertarik penelitian membuat *prototype* yang dapat mengklasifikasi tingkat kematangan buah apel berdasarkan warna dan tekstur menggunakan metode K-Nearest Neighbor untuk mempermudah dalam membedakan antara buah apel manalagi yang belum matang maupun yang sudah matang apabila tidak terlalu memiliki ilmu mengenai dunia pertanian terutama pada buah apel.

METODE PENELITIAN

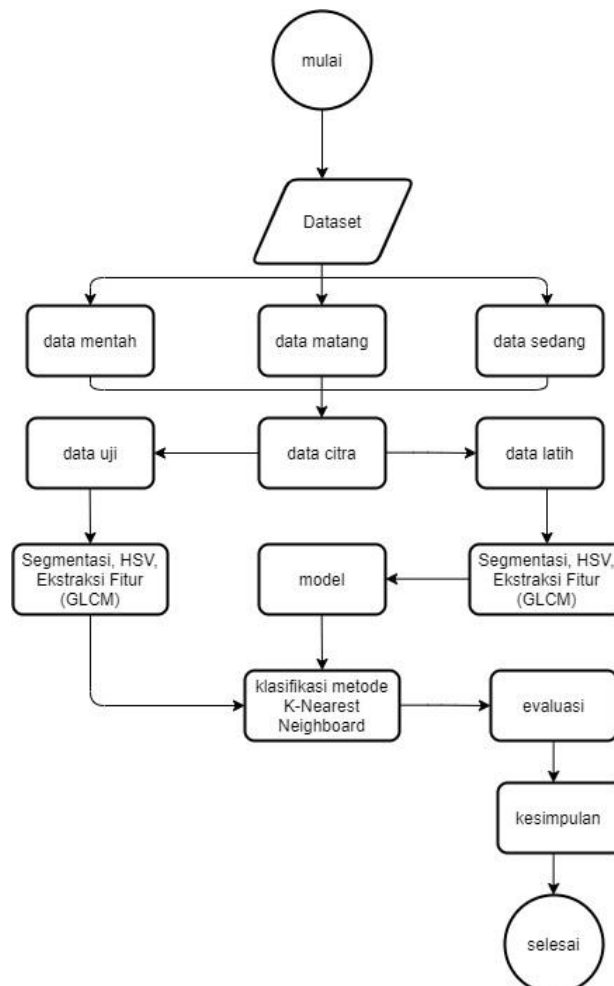
Permasalahan pada penelitian ini yaitu bagaimana cara merancang bangun prototype aplikasi berdasarkan warna dan tekstur untuk klasifikasi tingkat kematangan buah apel

manalagi. Untuk menyelesaikan permasalahan tersebut dimulai dengan mengumpulkan dataset 11 apel matang untuk training dan 4 untuk testing, 27 apel sedang untuk training dan 4 untuk testing, 37 apel mentah untuk training dan untuk testing. Sehingga total dataset gambar yang digunakan berjumlah 86 gambar. Yang mana gambar setiap kategori tersebut dibagi menjadi 2, yaitu untuk data training dan testing. Pembagiannya training (86%) dan testing (14%)



Gambar 1 Contoh dataset citra apel manalagi

Dari dataset yang dihasilkan dari penelitian sebelumnya kemudian diproses menggunakan alur penelitian pada Gambar 2. Sehingga dengan alur tersebut penelitian ini dapat menghasilkan kesimpulan tingkat keakuratan klasifikasi deteksi kematangan buah apel



Gambar 2 Alur Metode yang Diusulkan

1. Dataset
Dataset penelitian ini diperoleh berjumlah 86 gambar.
2. Data Citra
Data citra terdiri dari data buah apel mentah, sedang dan matang. Data citra buah apel yang digunakan yaitu 86 citra.
3. Dari data citra itu terbagi menjadi 2 yaitu data uji dan data latih
4. Data Uji dan data latih
Data uji dan data latih yang disegmentasikan terlebih dahulu kemudian untuk ekstraksi ciri warna mengkonversikan RGB ke HSV. Konversi ini mengacu pada penelitian sebelumnya yang mampu melakukan pengenalan buah dengan pendekatan konversi dari citra asli RGB ke bentuk warna HSV sebelum dilakukan histogram. Persamaan untuk melakukan konversi citra RGB menjadi citra HSV sebagai berikut. [7]

$$H = \tan \left[\frac{3(G-B)}{(G-B) + (R-B)} \right] \quad (2.3)$$

$$S = 1 - \left[\frac{\min(R,G,B)}{V} \right] \quad (2.4)$$

$$V = \frac{R+G+B}{3} \quad (2.5)$$

Pada rumus diatas dapat dilakukan proses normalisasi nilai RGB terlebih dahulu sebelum dilakukan konversi ke nilai HSV. Persamaan normalisasi RGB dapat dilihat sebagai berikut.

$$r = \frac{R}{R+G+B} \quad (2.6)$$

$$g = \frac{G}{R+G+B} \quad (2.7)$$

$$b = \frac{B}{R+G+B} \quad (2.8)$$

Keterangan rumus:

- R = nilai red belum normalisasi
- r = nilai red normalisasi
- G = nilai green belum normalisasi
- g = nilai green normalisasi
- B = nilai blue belum normalisasi
- b = nilai blue normalisasi

Setelah nilai normalisasi RGB selesai, maka citra akan dikonversi menjadi citra HSV. Persamaan yang digunakan untuk transformasi RGB ke HSV sebagai berikut.

$$v = \max(r, g, b) \quad (2.7)$$

$$S = \begin{cases} 0 & \text{jika } v=0 \\ v - \min(R,G,B) & \text{jika } v>0 \end{cases} \quad (2.8)$$

$$H = \begin{cases} 0 & \text{jika } s = 0 \\ 60 \times \left[0 + \frac{g-b}{v-s} \right] & \text{jika } v = r \\ 60 \times \left[2 + \frac{b-r}{v-s} \right] & \text{jika } v = g \\ 60 \times \left[4 + \frac{r-g}{v-s} \right] & \text{jika } v = b \end{cases} \quad (2.9)$$

$$H = H + 360 \quad \text{jika } H < 0 \quad (2.10)$$

Keterangan

- r = Nilai normalisasi dari R
- g = Nilai normalisasi dari G

- b = Nilai normalisasi dari B
- V = Nilai perhitungan value
- S = Nilai perhitungan saturation
- H = Nilai perhitungan hue

5. Extraksi tekstur menggunakan GLCM yaitu dengan mengkonversikan citra HSV ke grayscale dengan yang menghasilkan 4 nilai yaitu Contrast, Correlation, Energy, dan Homogeneity. GLCM memiliki $N \times N$ matriks persegi, di mana N mewakili jumlah tingkat abu-abu sebuah gambar. Sebuah elemen $p(i, j, d, \theta)$ dari GLCM dari gambar mewakili frekuensi relatif, di mana i merepresentasikan tingkat keabuan di lokasi (x,y), dan j merupakan tingkat keabuan piksel tetangga dengan jarak d dan orientasi θ dari lokasi (x,y). Jarak (d) yang digunakan biasanya 1 piksel dan orientasi sudut yang digunakan biasanya bernilai 0° , 45° , 90° , dan 135° . Formulasi matematika fitur tekstur dijelaskan pada Persamaan 4, Persamaan 5, Persamaan 6, dan Persamaan 7 [8].

1) Contrast

Merepresentasikan variasi antar derajat keabuan suatu daerah. Jika terletak jauh dari pusat diagonal utama, maka nilai kontras akan besar. Rumus menghitung kontras ditunjukkan Persamaan 2 [8].

$$\text{var} = \sum_i \sum_j (i - j)^2 p(i, j) \dots\dots\dots (2.13)$$

2) Correlation

Nilai korelasi merepresentasikan tingkat abu-abu ketergantungan linier antara piksel pada posisi tertentu terhadap piksel lain. Rumus menghitung korelasi ditunjukkan Persamaan 3 [8].

$$\text{Cor} = \sum_i \sum_j \frac{(i - \mu_i)(j - \mu_j) - p(i, j)}{\sigma_x \sigma_y} \dots\dots\dots (2.14)$$

3) Energy

Energy menyatakan tingkat keseragaman piksel-piksel suatu citra. Semakin tinggi nilai Energy, maka semakin seragam teksturnya. Rumus untuk menghitung Energy dapat dilihat pada persamaan 4 [8].

$$\text{Energy} = \sum_i \sum_j p(i, j)^2 \dots\dots\dots (2.15)$$

4) Homogeneity

Homogeneity menyatakan ukuran kedekatan setiap elemen dari cooccurrence matrix. Rumus untuk menghitung Homogeneity dapat dilihat pada persamaan 5 [8].

$$\text{Homogeneity} = \sum_i \sum_j \frac{p(i, j)}{1 + |i - j|} \dots\dots\dots (2.16)$$

6. Dari data uji dan data latih di klasifikasikan menggunakan metode K- Nearest Neighbor dengan nilai k=1. Algoritma K-Nearest Neighbor (KNN) adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Data pembelajaran diproyeksikan ke ruang berdimensi banyak, dimana masing – masing dimensi merepresentasikan fitur dari data. Algoritma KNN termasuk metode yang menggunakan algoritma supervised Perbedaan antara supervised learning dengan unsupervised learning adalah pada supervised learning bertujuan untuk menemukan pola baru dalam data dengan menghubungkan pola data yang

sudah ada dengan data yang baru. Sedangkan pada unsupervised learning, data belum memiliki pola apapun, dan tujuan unsupervised learning untuk menemukan pola dalam sebuah data Tujuan dari algoritma KNN adalah untuk mengklasifikasi objek baru berdasarkan atribut dan training samples Dimana hasil dari sampel uji yang baru diklasifikasikan [9].

$$dxy = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.12)$$

Keterangan:

- d : Jarak sample
- x : Data testing
- y : Data training
- n : Jumlah Sample

7. Setelah melakukan pengujian akan dilakukan evaluasi dengan menggunakan confusion matrik terhadap hasil dari pengujian tersebut. *Confusion matrix* memberikan keputusan yang diperoleh dalam *training* dan *testing*, *confusion matrix* memberikan penilaian *performance* klasifikasi berdasarkan objek dengan benar atau salah [10]. *Confusion matrix* berisi informasi aktual (*actual*) dan prediksi (*predicted*) pada sistem klasifikasi.

Table 1 Confusion Matrix untuk 2 Kelas

Classification		Predicted Class	
		Class = Yes	Class = No
Observed Class	Class = Yes	A (true positif – tp)	B (false negative – fn)
	Class = No	C (false positif – fp)	D (true negative – tn)

Keterangan:

- a. *True Positive* (tp) = proporsi positif dalam data set yang diklasifikasikan positif.
- b. *True Negative* (tn) = proporsi negative dalam data set yang diklasifikasikan negative.
- c. *False Positive* (fp) = proporsi negatif dalam data set yang diklasifikasikan positif.
- d. *FalseNegative*(fn) = proporsi negative dalam data set yang diklasifikasikan negative

Tabel 2. Tabel Confusion Matrix untuk 3 Kelas

		Predicted Class		
		Class 1	Class 2	Class 3
Actual Class	Class 1	Count11	Count12	Count13
	Class 2	Count21	Count22	Count23
	Class 3	Count31	Count32	Count33

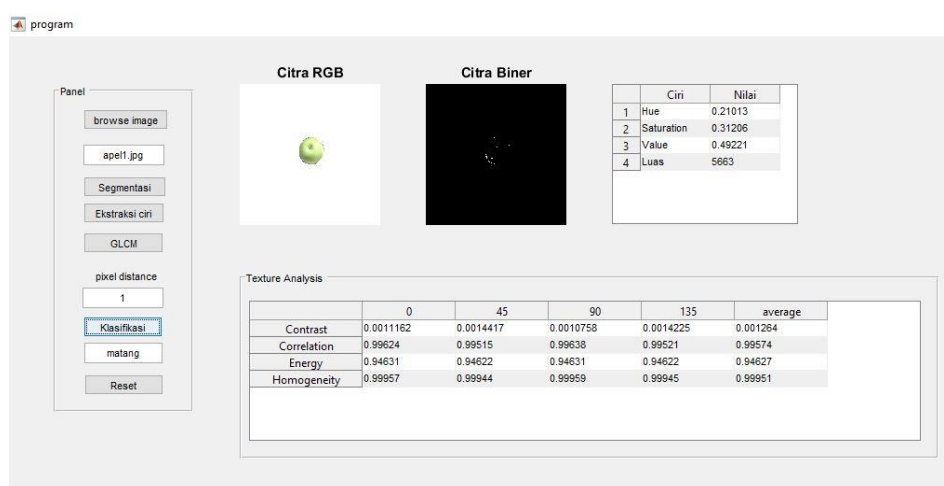
Berikut adalah persamaan model mencari akurasi pada *confusion matrix* untuk 3 kelas:

$$\text{Accuracy} = \frac{\text{Count11} + \text{Count22} + \text{Count33}}{\text{Count11} + \text{Count12} + \text{Count13} + \text{Count21} + \text{Count22} + \text{Count23} + \text{Count31} + \text{Count32} + \text{Count33}}$$

HASIL DAN PEMBAHASAN

Pada penelitian ini dataset dengan jumlah 11 apel matang untuk training dan 4 untuk testing, 26 apel sedang untuk training dan 4 untuk testing, 37 apel mentah untuk training dan 4 untuk testing. Sehingga total dataset gambar yang digunakan berjumlah 86 gambar. Dari dataset tersebut kemudian diproses untuk mejadi data latih dan data uji. Untuk mencapai tujuan penelitian ini, maka penelitian ini dibuatkan prototype dengan menggunakan software Matlab versi R2015B. Hasil dari perancangan prototype tersebut dapat dilihat di gambar 3.

Proses pada aplikasi yang sudah dibuat dimulai dengan memilih gambar dengan klik tombol "browse image" kemudian dilanjutkan dengan klik tombol "segmentasi" untuk merubah warna citra sebelum diekstraksi. Kemudian dilakukan ekstraksi dengan menekan tombol "ekstraksi ciri" dilanjut dengan klik tombol "GLCM" untuk mendapatkan nilai Contrast, Correlation, Energy dan Homogenity. Langkah terakhir yaitu dengan menekan tombol "klasifikasi" untuk mengklasifikasi citra yang dipilih merupakan tingkat kematangan yang matang, sedang, atau mentah.



Gambar 3 Prototype Klasifikasi Kematangan Buah Apel

Untuk evaluasi pada penelitian ini yaitu menggunakan confusion matrix untuk menghasilkan nilai akurasi. Berikut adalah hasil pengujian pada penelitian ini:

Tabel 3 Confusion Matrix

	Matang	Mentah	Sedang
Matang	3	0	1
Mentah	0	4	0
Sedang	0	0	4

Berdasarkan tabel diatas yang menunjukkan jumlah sampel yang bernilai benar, sehingga dapat dihitung nilai akurasi sebagai berikut

$$Akurasi = \frac{\text{Jumlah Data Yang Benar}}{\text{Total Data}} \times 100$$

$$Akurasi = \frac{3+4+4}{3+4+4+1} \times 100$$

$$Akurasi = \frac{11}{12} \times 100$$

Klasifikasi Kematangan Buah Apel Berdasarkan Warna dan Tekstur Menggunakan Algoritma K-Nearest Neighbor (Fandy Indra Pratama, Akhmad Pandu Wijaya, Hasti Pratiwi, Avira Budianita)

$$Akurasi = 91,6667\%$$

Berdasarkan hasil perhitungan tersebut maka diketahui nilai akurasi yang didapatkan dari 12 data testing buah apel sebesar 0,916667 atau 91,6667%.

KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan oleh Penulis dengan judul Klasifikasi Tingkat Kematangan Buah Apel Berdasarkan Warna Dan Tekstur Menggunakan Metode K-Nearest Neighbour dapat ditarik kesimpulan bahwa yang diuji sebanyak 12 buah apel yang terdiri dari matang, sedang, dan mentah yang menghasilkan tingkat kematangan buah apel sebesar 91,6667% dari total akurasi. Tingkat kematangan sebuah apel dapat diukur dengan menggunakan sebuah metode penghitungan berdasarkan melihat warna kulit dan tekstur yang dimiliki apel manalagi tersebut, dengan hasil lebih akurat. Sehingga memudahkan seseorang dalam mencari buah apel dengan tingkat kematangan yang cukup.

DAFTAR PUSTAKA

- [1] Y. Taufik, *Statistik Produksi Hortikultura 2014*. Direktorat jenderal hortikultura, kementerian pertanian, 2015.
- [2] T. Sutojo, D. R. I. M. Setiadi, P. S. Tirajani, C. A. Sari, and E. H. Rachmawanto, "CBIR for classification of cow types using GLCM and color features extraction," *Proc. - 2017 2nd Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng. ICITISEE 2017*, vol. 2018-January, pp. 182–187, Feb. 2018, doi: 10.1109/ICITISEE.2017.8285491.
- [3] A. Budianita and F. I. Pratama, "Penerapan Algoritma Klasifikasi Dengan Fitur Seleksi Weight By Information Gain Pada Pemodelan Prediksi Kelulusan Mahasiswa," *Infotekmesin*, vol. 11, no. 2, Aug. 2020, doi: 10.35970/INFOTEKMESIN.V11I2.255.
- [4] F. I. Pratama and A. Budianita, "Optimization of K-Nn classification in human gait recognition," *2020 5th Int. Conf. Informatics Comput. ICIC 2020*, Nov. 2020, doi: 10.1109/ICIC50835.2020.9288653.
- [5] S. Nuraisha, F. I. Pratama, A. Budianita, and M. A. Soeleman, "Implementation of K-NN based on histogram at image recognition for pornography detection," in *Proceedings - 2017 International Seminar on Application for Technology of Information and Communication: Empowering Technology for a Better Human Life, iSemantic 2017*, Institute of Electrical and Electronics Engineers Inc., Jul. 2017, pp. 5–10. doi: 10.1109/ISEMANTIC.2017.8251834.
- [6] I. Siswanto, E. Utami, S. Raharjo, and A. Yogyakarta, "Klasifikasi Tingkat Kematangan Buah Berdasarkan Warna dan Tekstur Menggunakan Metode K-Nearest Neighbor dan Nearest Mena Classifier," *Inspir. J. Teknol. Inf. dan Komun.*, vol. 10, no. 1, pp. 93–101, Jun. 2020, doi: 10.35585/INSPIR.V10I1.2559.
- [7] R. P. Rakhmawati, "SISTEM DETEKSI JENIS BUNGA MENGGUNAKAN NILAI HSV DARI CITRA MAHKOTA BUNGA," *Inf. Technol. Telemat.*, vol. 2, no. 1, 2013.
- [8] M. Dhewa Kusuma, U. Rosidin, and A. Suyatna, "The Development of Higher Order Thinking Skill (Hots) Instrument Assessment In Physics Study," vol. 7, no. 1, pp. 26–32, doi: 10.9790/7388-0701052632.
- [9] F. Liantoni, "Klasifikasi Daun Dengan Perbaikan Fitur Citra Menggunakan Metode K-Nearest Neighbor," *Ultim. J. Tek. Inform.*, vol. 7, no. 2, pp. 98–104, Aug. 2015, doi: 10.31937/TI.V7I2.356.
- [10] F. Gorunescu, "Data mining: Concepts, models and techniques," *Intell. Syst. Ref. Libr.*, vol. 12, 2011, doi: 10.1007/978-3-642-19721-5/COVER.

