

Perbandingan Klasifikasi Jenis Apel Berkulit Merah Menggunakan Algoritma Linear Discriminant Analysis dan K-Nearest Neighbor

Ajib Susanto¹, Ibnu Utomo Wahyu Mulyono²

Program Studi Teknik Informatika, Universitas Dian Nuswantoro
Jl. Imam Bonjol 207 Semarang 50131

¹ajib.susanto@dsn.dinus.ac.id

²ibnu.utomo.wm@dsn.dinus.ac.id

Abstrak— Buah apel merupakan salah satu buah yang mempunyai rasa dominan manis segar dan memiliki vitamin C tinggi. Apel dibudidayakan untuk tujuan konsumsi, obat maupun industri. Dalam industri, apel digunakan sebagai bahan baku pembuatan berbagai macam bentuk makanan dan minuman misalnya sirup, jenang, wingko, dodol, manisan, asinan, keripik, dan sari apel. Jenis apel yang beragam dan kebutuhan waktu pendistribusian berdasarkan jenis apel memerlukan banyak waktu dan berhubungan dengan kemampuan mata manusia dalam proses sorting manual. Kebutuhan teknologi seperti computer vision melalui teknik pengolahan citra dapat diimplementasikan untuk proses sorting khususnya klasifikasi jenis apel. Dalam penelitian ini, digunakan apel dengan kulit berwarna merah sebagai dataset. Kesamaan warna kulit dan bentuk apel yang hampir sama, menjadi salah satu isu penting untuk proses klasifikasi citra. K Nearest Neighbor (KNN) dan Linear Discriminant Analysis (LDA) dipilih karena kemampuan klasifikasi citra dengan dataset kecil. Dalam penelitian ini telah dilakukan proses perbandingan hasil akurasi antara KNN dan LDA berdasarkan 400 dataset yang berasal dari 8 jenis apel merah antara lain Cameo, Honeycrisp, Pink Lady, Red Delicious, Royal Gala, Macintosh, Empire, Fuji. KNN dan LDA tanpa menggunakan ekstraksi fitur GLCM menghasilkan akurasi yang hampir sama yaitu 99,25% dan 99% sedangkan apabila tidak menggunakan fitur ekstraksi apapun dihasilkan akurasi 99,25% dan 99%. Dengan demikian diketahui bahwa KNN menghasilkan akurasi lebih tinggi dibanding PCA, meskipun hanya terdapat sedikit selisih akurasi.

Kata kunci— Klasifikasi, Apel, KNN, LDA, GLCM

I. PENDAHULUAN

Buah memegang peranan penting dalam perekonomian dunia khususnya pada bidang industri. Industri makanan dan minuman mempunyai kesinambungan dengan bahan baku, salah satunya berupa buah. Jenis buah yang mempunyai nilai ekonomi tinggi pada industri yaitu buah apel. Apel dengan berbagai macam bentuk, warna, jenis dan tekstur dapat digunakan sebagai bahan baku melalui proses sorting [1]. Sorting buah dapat dilakukan untuk mengklasifikasi mutu dan jenis. Pada proses klasifikasi jenis buah, khususnya apel mempunyai limitasi antara lain pada bentuk, warna dan tekstur buah [2]. Sorting berbagai

jenis apel dengan warna kulit buah yang sama memerlukan bantuan teknologi komputer. Computer vision melalui pengolahan citra digital dapat melakukan proses sorting dengan akurat. Hal ini akan meminimalkan ketergantungan sortir buah menggunakan metode konvensional melalui mata manusia [3]–[5]. Adanya penyakit buta warna dan tingkat kecerahan cahaya pada tempat sorting dapat mengaburkan hasil klasifikasi menjadi tidak akurat.

Machine learning, dalam bentuk supervise learning [6], unsupervised learning, dan reinforcement learning menjadi algoritma yang dapat dimanfaatkan untuk melakukan proses sorting melalui teknik pengolahan citra. Support Vector Machine (SVM) [7], Regresi Logistic, Regresi Linier, Decision Tree, Random Forest, K-Nearest Neighbor (KNN) [8], Naïve Bayes, Neural Network, Principal Component Analysis (PCA) [9], Linear Discriminant Analysis (LDA), Learning Vector Quantization (LVQ) [10], K-Means Clustering [11], Gradient Boosting merupakan bentuk supervised learning yang sudah pernah diimplementasikan dalam proses klasifikasi buah.

Proses klasifikasi pada K-Nearest Neighbor (KNN) sangat sederhana, namun memungkingkan terjadinya ketidakcocokan nilai fitur pada proses klasifikasi sehingga perlu adanya proses pengurangan dimensi. Linear Discriminant Analysis (LDA) mempunyai karakteristik untuk memaksimalkan keterpisahan nilai dengan membentuk sumbu baru pada visualisasi data. Baik KNN maupun LDA yang sama-sama merupakan algoritma supervised learning, dalam penelitian ini akan dilakukan analisa mengenai hasil akurasi yang dihasilkan dalam klasifikasi citra buah apel.

II. PENELITIAN TERKAIT

Beberapa penelitian yang relevan dengan objek berupa buah dan proses klasifikasi telah kami rangkum pada Tabel 1. Abdulhamid Haider tahun 2012 [12] melakukan klasifikasi buah kurma menggunakan fitur warna, bentuk dan tekstur. KNN, LDA dan ANN diuji dengan total 140 gambar kurma tersedia, dibagi rata antara tujuh jenis tanggal pasar umum – Sukkery, Mabroom, Khedri, Safawi

Al Madina, Madina Ajwa, Amber Al Madina, dan Safree. Gambar berukuran 480 x 640 dan berformat JPEG, serta diambil dari posisi yang telah ditentukan, sehingga semua gambar memiliki skala dimensi yang sama. Dari hasil percobaan ditemukan nilai akurasi KNN yaitu 90% sedangkan LDA yaitu 96%, sedangkan ANN menghasilkan akurasi tertinggi yaitu 98%.

Kadir Sabanci [9] pada tahun 2013, dalam penelitiannya menggunakan tiga varietas apel berbeda yang ditanam di provinsi Karaman diklasifikasikan menggunakan algoritma kNN dan MLP. Total 90 apel, 30 Golden Delicious, 30 Granny Smith dan 30 Starking Delicious telah digunakan dalam penelitian ini. Kamera industri DFK 23U445 USB 3.0 (dengan Fujinon C Mount Lens) telah digunakan untuk menangkap gambar apel. 4 properti ukuran (diameter, luas, keliling dan kepenuhan) dan 3 properti warna (merah, hijau, biru) telah diputuskan menggunakan teknik pemrosesan gambar melalui analisis setiap gambar apel. Kumpulan data yang berisi 7 fitur fisik untuk setiap apel telah diperoleh. Tingkat keberhasilan klasifikasi dan tingkat kesalahan telah diputuskan dengan mengubah jumlah neuron pada lapisan tersembunyi dalam klasifikasi menggunakan model MLP dan pada nilai tetangga yang berbeda dalam klasifikasi yang dibuat menggunakan algoritma kNN. Terlihat bahwa klasifikasi menggunakan model MLP jauh lebih tinggi. Sedangkan tingkat keberhasilan klasifikasi yang dibuat menurut jenis apel adalah 98,8889%.

Oktaviana Rena Indriani et al. [13] pada tahun 2017 telah melakukan proses klasifikasi buah tomat menggunakan algoritma KNN. Menggunakan 100 dataset yang terdiri dari 75 data latih dan 25 data uji yang menghasilkan tingkat akurasi tertinggi 100% dengan nilai p pada GLCM adalah 9 dan nilai keanggotaan (k) pada K-NN adalah 3 Berdasarkan hasil eksperimen, kami dapat menyimpulkan bahwa metode yang kami usulkan dapat mencapai akurasi tertinggi.

Penelitian yang dilakukan oleh Xi Yang et al. [14] pada tahun 2019, membandingkan antara dua metode pemilihan variabel dan enam metode pembelajaran mesin, SVM dalam kombinasi dengan SPA ditemukan memiliki akurasi tertinggi (90,7%) dalam set pengujian. Oleh karena itu, kombinasi ini dapat digunakan untuk mengidentifikasi kultivar aprikot secara otomatis dengan cara yang tidak merusak. Namun, penyelidikan tambahan mengenai fitur bentuk dan klasifikasi kultivar harus dilakukan. Peningkatan jumlah karakteristik terukur dan jumlah kultivar akan menjadi salah satu cara untuk meningkatkan kinerja klasifikasi.

Pada penelitian yang dilakukan oleh P.L. Chithra et al. [15] tahun 2019, komparasi algoritma juga telah dilakukan. Dataset yang digunakan terdiri dari 70 citra apel dan 70 citra pisang untuk pelatihan dan 25 citra apel dan 25 citra pisang untuk pengujian. Citra RGB terlebih dahulu diubah menjadi citra HSI. Kemudian dengan menggunakan metode thresholding Otsu, region of interest disegmentasi dengan hanya memperhitungkan gambar komponen HUE dari gambar HSI. Kemudian, setelah pengurangan latar

belakang, total 36 fitur statistik dan tekstur diekstraksi dengan bantuan koefisien yang diperoleh dengan menerapkan transformasi wavelet pada citra tersegmentasi menggunakan filter Haar. Fitur yang diekstraksi diberikan sebagai input ke pengklasifikasi SVM untuk mengklasifikasikan gambar uji sebagai apel dan pisang. Karena metode klasifikasi KNN tidak memberikan akurasi 100% sedangkan metode klasifikasi SVM digunakan. 140 sampel gambar apel dan pisang digunakan untuk pelatihan dan 25 gambar pisang dan 25 gambar apel digunakan untuk menguji algoritma yang diusulkan. Algoritma yang diusulkan memberikan tingkat akurasi 100%.

Tabel 1. Penelitian Terkait

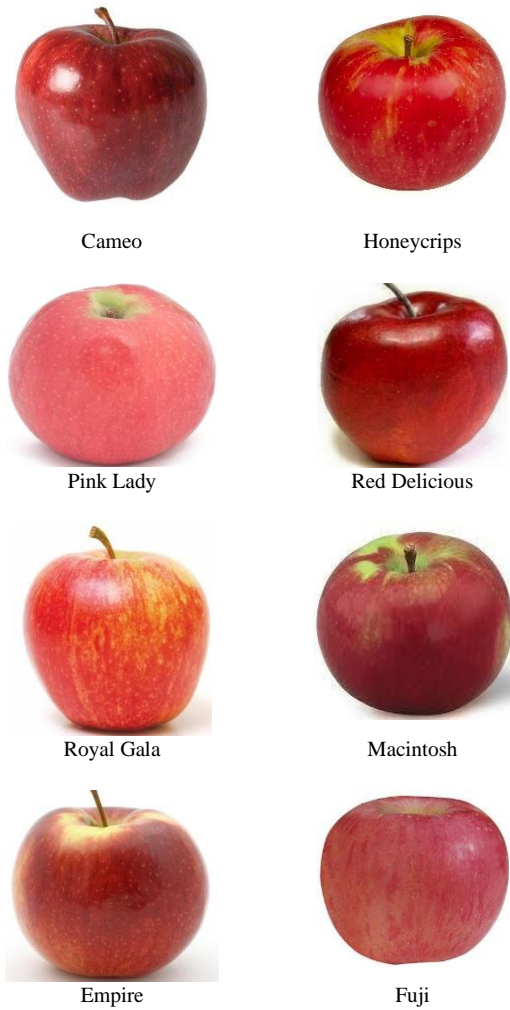
Author	Dataset	Fitur	Algoritma	Akurasi
Abdul Hamid Haider et al. [2012]	140	Warna, bentuk, tekstur	KNN, LDA, ANN	89%-99%
Kadir Sabanci [2013]	90	Bentuk dan warna	KNN, MLP	98,88%
Oktaviana Rena Indriani et al. [2017]	100	Tekstur dan warna	KNN	100%
Xi Yang et al. [2019]	800	bentuk	KNN, Naïve Bayes, SVM, LDA, Neural Network	SVM-90,7%
P.L. Chithra et al. [2019]	140	Statistikal dan tekstur	KNN, SVM	SVM-100

III. DATASET

Pada penelitian ini digunakan 400 citra buah apel dengan varian jenis buah yaitu Cameo, Honeycrisp, Pink Lady, Red Delicious, Royal Gala, Macintosh, Empire, Fuji. Masing-masing varian apel yaitu 80 citra dengan ukuran 256x256 piksel hingga 1024x1024 piksel dalam bentuk citra warna. Sampel dataset dapat dilihat pada Gambar 1. Sebagian data di dapat dari source internet dan sebagian dari hasil foto pribadi. Model pengambilan foto untuk dataset dilakukan pada siang hari yaitu jam 11 sampai jam 1 siang. Dataset tidak mengalami proses peningkatan kualitas citra dalam bentuk apapun, misalnya filtering, smoothing atau proses morfologi citra.

Pada penelitian ini digunakan apel berkulit merah dengan tujuan untuk menganalisa kemampuan algoritma karena kebanyakan proses sorting atau klasifikasi menggunakan objek dengan warna berbeda, sehingga dengan mudah proses deteksi dapat menghasilkan akurasi yang tinggi. Sedangkan dalam penelitian ini, warna dan bentuk apel serta tekstur buah yang sangat mirip menjadi

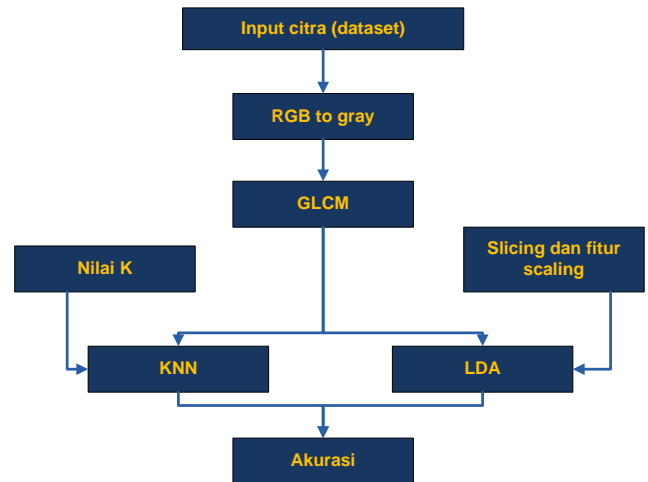
salah satu limitasi penelitian yang dapat menghasilkan hipotesa.



Gambar 1. Dataset Buah Apel

IV. USULAN METHODE

Pada penelitian ini akan dilakukan komparasi antara KNN dan LDA dalam akurasi klasifikasi citra buah apel seperti ditunjukkan Gambar 2.

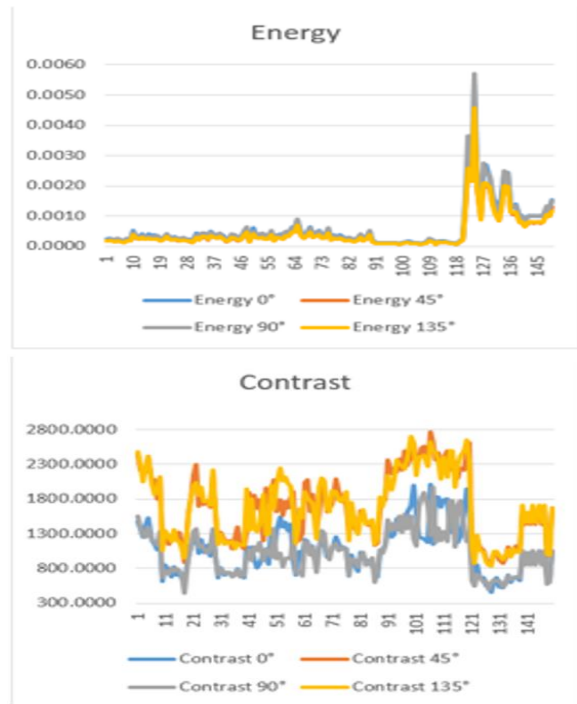


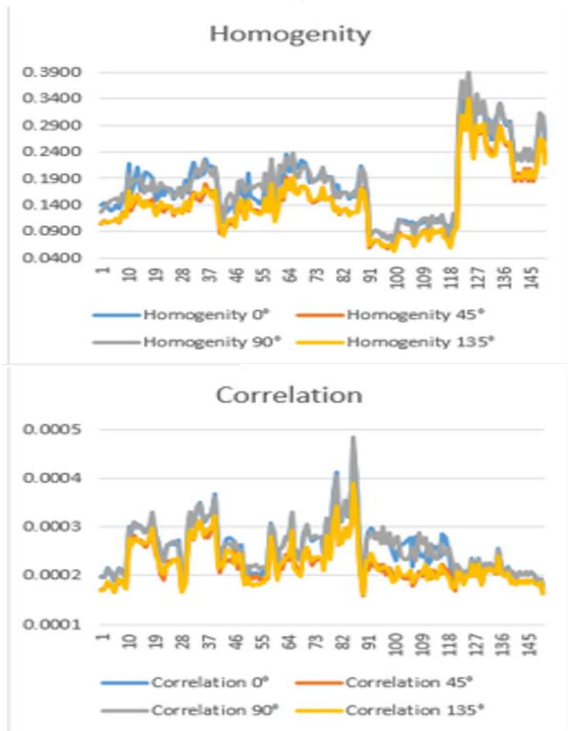
Gambar 2. Skema Usulan

Pada Gambar 2, klasifikasi pada KNN menggunakan perhitungan Euclidean Distance dengan nilai $K=1,3,5,7,9$. Klasifikasi LDA menggunakan slicing dan fitur scaling untuk menginputkan dataset dalam variable train dan test sehingga dapat dihasilkan variabel independent baru.

V. PEMBAHASAN HASIL

A. Fitur Ekstraksi GLCM





Gambar 3. Hasil perhitungan GLCM

Gambar 4 merupakan potongan coding yang digunakan untuk memproses GLCM.

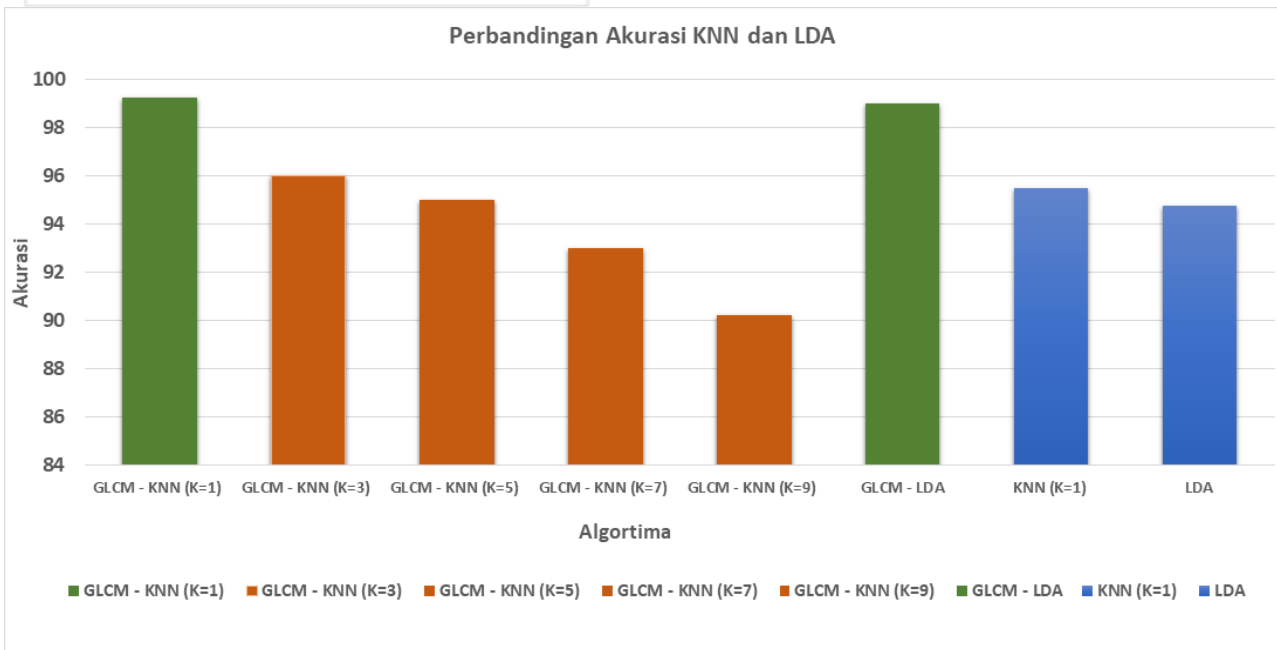
```
group = cell(total_images,1);
for n = 1:total_images
full_name = fullfile(image_folder, filenames(n).name);
our_images = rgb2gray(imread(full_name));
glcm_training_img =
graycomatrix(our_images,'Offset',[0 jrk; -jrk jrk; -jrk 0; -jrk -jrk]);
stats = graycoprops(glcm_training_img,
{'Contrast','Correlation','Energy','Homogeneity'});
Contrast = stats.Contrast;
Correlation = stats.Correlation;
Energy = stats.Energy;
Homogeneity = stats.Homogeneity;
```

Gambar 4. Coding GLCM

B. Klasifikasi

Tabel 2. Perbandingan ketepatan klasifikasi

Algoritma	Fitur	Terdeteksi benar	Terdeteksi salah
KNN (pada K=1)	GLCM	397	3
LDA	GLCM	396	4
KNN (pada K=1)	-	382	18
LDA	-	379	21



Gambar 5. Perbandingan akurasi KNN dan LDA

Berdasarkan Gambar 5, dapat diambil kesimpulan bahwa menggunakan fitur ekstraksi GLCM, KNN (pada K=1) menghasilkan akurasi sedikit lebih tinggi dibanding LDA, sedangkan tanpa menggunakan fitur ekstraksi KNN juga lebih unggul dibanding LDA. Untuk melihat performa KNN, kami juga melakukan uji KNN pada nilai K=3,5,7,9. Diketahui bahwa GLCM-KNN pada K=3 dan K=5 masih menghasilkan akurasi lebih tinggi dibanding LDA saja tanpa fitur ekstraksi.

VI. KESIMPULAN

Porses klasifikasi buah apel menggunakan KNN dan LDA telah dilakukan dengan hasil berupa akurasi. Akurasi terendah yaitu 90% sedangkan tertinggi yaitu 99,25%. Penggunaan GLCM terbukti meningkatkan akurasi, dimana KNN saja memperoleh 95% sedangkan GLCM-KNN mampu meningkatkan akurasi hingga 3,75% (dari 95,5 menuju 99,25%). Hal ini juga terjadi pada LDA, dimana penggunaan GLCM berhasil menaikkan akurasi sebesar 4% (dari 95% menuju 99%). Penggunaan GLCM pada KNN dengan K=3 menghasilkan akurasi hampir sama dengan LDA saja tanpa GLCM. Untuk nilai GLCM KNN pada K=7 dan K=9

sebetulnya masih menghasilkan akurasi cukup tinggi yaitu diatas 90%. Pada penelitian lebih lanjut dapat dilakukan uji perbandingan antara beberapa algoritma dalam supervised learning misalnya Support Vector Machine (SVM), dan Neural Network (NN) serta melakukan proses klasifikasi realtime menggunakan deep learning misalnya pada algoritma Convolutional Neural Network (CNN).

REFERENSI

- [1] Ferdian Wahyu Prianggara, Ahmad Bagus Setiawan, and Intan Nur Farida, "Indetifikasi Jenis Buah Apel Berdasarkan Ekstraksi Bentuk dan Warna," in Seminar Nasional Inovasi Teknologi, 2020, pp. 215–219.
- [2] Arun C, Akshatha Prabhu, Mohammed Zeeshan, and N. Shobha Rani, "A Study on Various Classifier Techniques Used in Image Processing," 2020.
- [3] P. Moallem, A. Serajoddin, and H. Pourghassem, "Computer vision-based apple grading for golden delicious apples based on surface features," *Information Processing in Agriculture*, vol. 4, no. 1, pp. 33–40, 2017, doi: 10.1016/j.inpa.2016.10.003.
- [4] X. Wu, B. Wu, J. Sun, M. Li, and H. Du, "Discrimination of Apples Using Near Infrared Spectroscopy and Sorting Discriminant Analysis," *International Journal of Food Properties*, vol. 19, no. 5, pp. 1016–1028, May 2016, doi: 10.1080/10942912.2014.971181.
- [5] W. Huang, J. Li, Q. Wang, and L. Chen, "Development of a multispectral imaging system for online detection of bruises on apples," *Journal of Food Engineering*, vol. 146, pp. 62–71, Feb. 2015, doi: 10.1016/j.jfoodeng.2014.09.002.
- [6] X. Yue and N. Tian, "Research on Apple Classification System Based on Hybrid Kernel Function and Multi-Feature Fusion," in *Proceedings - 2020 International Conference on Computer Engineering and Application, ICCEA 2020*, Mar. 2020, pp. 589–593. doi: 10.1109/ICCEA50009.2020.00129.
- [7] Y. Tian, E. Li, L. Yang, and Z. Liang, "An image processing method for green apple lesion detection in natural environment based on GA-BPNN and SVM," *Proceedings of 2018 IEEE International Conference on Mechatronics and Automation, ICMA 2018*, pp. 1210–1215, 2018, doi: 10.1109/ICMA.2018.8484631.
- [8] A. Bhargava and A. Bansal, "Classification and Grading of Multiple Varieties of Apple Fruit," *Food Analytical Methods*, 2021, doi: 10.1007/s12161-021-01970-0.
- [9] K. Sabanci and M. Fahri ÜNLERŞEN, "Different Apple Varieties Classification Using kNN and MLP Algorithms," *IJISAE (International Journal of Intelligent Systems and Applications in Engineering)*, no. 4, pp. 166–169, 2016, [Online]. Available: www.at-science.org/IJISAE
- [10] P. Boniecki et al., "Neural Identification of Chosen Apple Pests Using Algorithm LVQ," in *HAICTA*, 2017, pp. 293–303.
- [11] Y. Li, X. Feng, Y. Liu, and X. Han, "Apple quality identification and classification by image processing based on convolutional neural networks," *Scientific Reports*, vol. 11, no. 1, pp. 1–15, 2021, doi: 10.1038/s41598-021-96103-2.
- [12] Abdulhamid Haidar, Haiwei Dong, and Nikolaos Mavridis, "Image-Based Date Fruit Classification," in *IV International Congress on Ultra Modern Telecommunications and Control Systems 2012*, 2012, pp. 357–363.
- [13] O. Rena Indriani, C. Atika Sari, E. Jaya Kusuma, E. Hari Rachmawanto, and D. Rosal Ignatius Moses Setiadi, "Tomatoes Classification Using K-NN Based on GLCM and HSV Color Space," 2017.
- [14] X. Yang, R. Zhang, Z. Zhai, Y. Pang, and Z. Jin, "Machine learning for cultivar classification of apricots (*Prunus armeniaca* L.) based on shape features," *Scientia Horticulturae*, vol. 256, Oct. 2019, doi: 10.1016/j.scienta.2019.05.051.
- [15] H. Manickam, P. L. Chithra, and M. Henila, "Fruit Classification using image processing techniques Fruits Classification Using Image Processing Techniques," *International Journal of Computer Sciences and Engineering Open Access Research Paper*, no. 5, 2019, doi: 10.26438/ijcse/v7si5.131135.