

Available online at: <https://journal.lppmunindra.ac.id/index.php/JOTI>

## Jurnal Optimasi Teknik Industri

| ISSN (Print) 2656-3789 | ISSN (Online) 2657-0181|



# Klasifikasi Varietas Biji Kismis dengan Artificial Neural Network

Hanifah Nabila Wismadi<sup>1</sup>, Choirul Yof<sup>2</sup>, Thariq Faros Manumon<sup>3</sup>, Fauzi Arsyad Hendrawan<sup>4</sup>,  
Muhammad Raihan Hilmy<sup>5</sup>, Afifa Puspitasari<sup>6</sup>, Nur Mayke Eka Normasari<sup>7</sup>, Achmad P. Rifai<sup>8\*</sup>

<sup>1,2,6,7,8</sup>Industrial Engineering, Department of Mechanical and Industrial Engineering, Universitas Gadjah Mada, Jl. Grafika No. 2 Jalan Grafika No. 2, Yogyakarta 55281, Indonesia

<sup>3,4,5</sup>Mechanical Engineering, Department of Mechanical and Industrial Engineering, Universitas Gadjah Mada, Jl. Grafika No. 2 Jalan Grafika No. 2, Yogyakarta 55281, Indonesia

\*Corresponding author: achmad.p.rifai@ugm.ac.id

### ARTICLE INFORMATION

Received: 09 Agustus 2022  
Revised: 30 Januari 2023  
Accepted: 31 Januari 2023  
Available online: 25 Maret 2023

### KEYWORDS

klasifikasi,  
kismis,  
jaringan saraf tiruan,  
jumlah neuron,  
akurasi

### ABSTRAK

Penerapan kecerdasan buatan pada industri kismis banyak dikembangkan sebagai upaya mengatasi *human error* pada penyortiran secara manual. Pada studi ini jaringan saraf tiruan atau *artificial neural network (ANN)* diterapkan untuk mengklasifikasikan kismis besni dan kismis kecimen. Studi ini menggunakan dataset dari *UCI Machine Learning Repository*. Data tersebut kemudian dilakukan pemrosesan awal dengan metode *encoding categorical* dan *min max scaler*. Studi ini membandingkan akurasi tes model ANN yang memiliki jumlah neuron yang berbeda. Terdapat lima level jumlah neuron yang masing-masing ditinjau dengan lima kali trial. Jumlah neuron yang diterapkan pada model penelitian ini adalah 10, 20, 30, 40, dan 50. Pada penelitian ini menggunakan input berupa tujuh variabel yang menggambarkan karakteristik ukuran dan bentuk kismis untuk membedakan antara kismis berjenis besni dan kecimen yaitu jumlah piksel dalam batas kismis, panjang sumbu utama, panjang sumbu kecil, ukuran eksentrisitas elips, jumlah piksel kulit cembung tekecil, rasio wilayah antara kismis dan kotak pembatas, dan jarak antara batas kismis dan piksel sekitarnya. Penyelesaian dilakukan menggunakan aplikasi dari MATLAB dengan algoritma *scaled conjugate gradient*. Diperoleh bahwa terdapat *trend* peningkatan rata-rata *test accuracy* seiring dengan bertambahnya jumlah neuron. Nilai rata-rata akurasi tes tertinggi sebesar 86.7% diperoleh pada jumlah neuron 50. Akan tetapi, akurasi tes tidak bertambah lagi secara signifikan pada penambahan jumlah neuron dari 40 ke 50. Dengan demikian, studi ini dapat membuktikan adanya hubungan antara jumlah neuron dengan akurasi dari model ANN.

## I. PENDAHULUAN

Kismis merupakan produk makanan dari anggur yang dikeringkan. Selama pengeringan anggur, kandungan air akan berkurang 75% hingga di bawah 15% (Zoffoli & Latorre, 2011). Warna akhir dan kualitas dari kismis akan bergantung pada varietas dari anggur serta metode pengeringan yang digunakan (Thompson, 2000).

Terdapat berbagai metode untuk menilai kualitas dari kismis. Di antaranya adalah metode tradisional yang dilakukan secara manual oleh manusia. Akan tetapi, metode tradisional cenderung memakan waktu yang lama dan berpotensi terjadinya *human error*. Untuk mengatasi hal tersebut, berbagai peneliti kemudian mengembangkan metode alternatif untuk

menevaluasi kualitas kismis dengan *machine vision* system (Mollazade et al., 2012; Karimi et al., 2017; Cinar et al., 2020). Pengembangan metode ini diharapkan dapat menjadi dasar pengembangan sistem penyortiran kismis untuk meningkatkan efisiensi di industri.

Pada studi ini, permasalahan yang diangkat adalah penyortiran dua varietas biji kismis yakni kismis besni dan kismis kecimen. Data yang digunakan pada studi ini diambil dari dataset yang bersumber dari *UCI Machine Learning Repository* berdasarkan studi Cinar et al. (2020) yang menangkap gambar dengan *computer vision system (CVS)*. Terdapat sejumlah 900 data biji kismis dengan masing-masing varietas sejumlah 450 data.

Permasalahan yang ditinjau pada studi ini termasuk *classification problem* di mana output yang dihasilkan bersifat kategorikal. Pada studi ini dilakukan perbandingan akurasi dari model *artificial neural network (ANN)* dengan jumlah neuron yang berbeda-beda. Hal ini dilakukan untuk mengetahui jumlah neuron yang menghasilkan akurasi yang paling tinggi.

Terdapat berbagai studi yang melakukan klasifikasi biji kismis dengan *machine vision* dan *image processing*. Pada penelitian Khojastehnazhand & Ramezani (2020), kualitas biji kismis diklasifikasikan menggunakan metode *image processing*. Untuk melakukan hal tersebut, digunakan sebanyak 750 gambar biji kismis (50 gambar x 15 kelas) yang terdiri dari gambar biji kismis dengan kualitas baik atau buruk. Algoritma berbasis *texture* dikombinasikan dengan metode *modelling* yang berbeda yang digunakan untuk mengevaluasi kinerja sistem. Penelitian ini menunjukkan bahwa metode klasifikasi SVM yang menggunakan fitur *gray level run length matrix (GLRM)* menghasilkan akurasi klasifikasi yang lebih baik. Hasil klasifikasi yang ditunjukkan pada model 1 (terdiri dari 6 kelas kismis kualitas baik dan buruk) dan model 2 (terdiri dari 15 kelas) menghasilkan akurasi sebesar 85,55% dan 69,78%. Penelitian ini juga menunjukkan bahwa *machine vision system* dapat digunakan untuk mengklasifikasikan biji kismis dengan kualitas baik atau buruk tetapi untuk kasus model 2 menghasilkan akurasi yang lebih rendah.

Berbeda dengan penelitian Khojastehnazhand & Ramezani (2020) yang menggunakan SVM berbasis GLRM, penelitian Yu et al. (2012) menggunakan *least squares support vector machine (LSSVM)*

berbasis kombinasi fitur warna dan tekstur untuk mengklasifikasikan kualitas biji kismis. Total 480 gambar dari empat kelas biji kismis yang diambil menggunakan kamera digital Basler 401 fc IEEE1394, kemudian 200 gambar tersebut dipilih secara acak untuk digunakan sebagai training set, dan sisanya digunakan untuk melakukan verifikasi model (prediction set). Penelitian tersebut menunjukkan akurasi sebesar 95% dalam mengklasifikasikan kualitas biji kismis.

Selain itu, berbagai studi juga telah melakukan klasifikasi varietas biji kismis dengan *artificial neural network* dan algoritma pembanding lainnya. Pada studi Karimi et al. (2017) dilakukan klasifikasi kismis jenis *golden-bleached* menggunakan ANN dan *support vector machine (SVM)*. Diketahui bahwa dengan SVM dihasilkan akurasi yang lebih tinggi yakni rata-rata 92,71%. Sementara itu, pada studi Mollazade et al. (2012), dilakukan klasifikasi empat jenis kismis (*green, green with tail, black, dan black with tail*) menggunakan empat metode berbeda. Diketahui bahwa hasil akurasi tertinggi diperoleh dengan metode ANN (96,33%), diikuti dengan metode SVM (95,67%), *decision tree* (94,67%), dan *bayesian network* (94,33%). Dapat dilihat dari studi kedua bahwa performa dari ANN dapat lebih baik dari SVM tergantung permasalahan klasifikasi yang diselesaikan. Pada studi ini, metode ANN digunakan untuk menganalisis parameter jumlah neuron ANN yang menghasilkan akurasi yang tertinggi.

Selain itu, terdapat pula studi yang dilakukan oleh Cinar et al. (2020) yang mengklasifikasikan kismis kecimen dan besni. Pada studi tersebut, dilakukan perbandingan model dari tiga algoritma yang berbeda di mana diperoleh akurasi dari model logistic regression (LR) sebesar 85,22%, multilayer perceptron (MLP) sebesar 86,33%, dan akurasi tertinggi support vector machine (SVM) 86,44%. Pada studi ini, klasifikasi dilakukan berdasarkan data yang sama yang diambil oleh Cinar et al. (2020).

Selain untuk pengkategorian jenis raisin, metode *machine learning* juga diterapkan untuk permasalahan lain yang serupa. Guo et al. (2022) menerapkan SVM dan *multiscale fusion convolutional neural network (MCNN)* untuk identifikasi raisin palsu. Yajun et al. (2022) juga mengidentifikasi varietas raisin dengan jenis yang berbeda, yaitu Hongxiangfei, Manaiti, dan Munage. Studi tersebut membandingkan performa k-nearest neighbor (KNN), linear discriminant analysis (LDA),

dan SVM dalam mengidentifikasi jenis raisin. Beberapa penelitian terkini lainnya dalam aplikasi *machine learning* untuk identifikasi raisin antara lain KNN dan weighted KNN untuk klasifikasi raisin jenis Osmancik dan Cammeo (Tarakci dan Ozkan, 2021), SVM, KNN, dan Radial Basis Function Neural Network (RBFNN) untuk klasifikasi raisin jenis Wuhebai dan Xiangfei (Feng et al., 2018).

Studi sebelumnya menunjukkan bahwa machine learning merupakan salah satu metode yang menjanjikan untuk prediksi dan identifikasi raisin. Sebagai perkembangan dari studi sebelumnya, studi ini mengembangkan ANN untuk problem identifikasi raisin jenis Kecimen dan Besni. Pada studi ini dilakukan perbandingan akurasi dari beberapa model dengan jumlah neuron yang berbeda dan diidentifikasi pengaruh jumlah neuron terhadap akurasi model. Studi ini menggunakan aplikasi dari MATLAB untuk menyelesaikan *classification problem* tersebut dengan algoritma *scaled conjugate gradient*. Pada Tabel 1 ditunjukkan perbandingan studi ini dengan studi lainnya.

Tabel 1. Perbandingan Studi

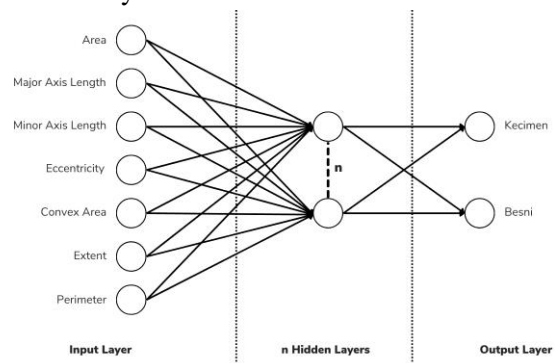
Peneliti	Jenis Kismis	Metode
Mollazade et al. (2012)	Golden-bleached	ANN, SVM, decision tree, bayesian network
Yu et al. (2012)	Bulk raisin	Least squares support vector machine (LSSVM)
Karimi et al. (2017)	Green, green with tail, black, dan black with tail	ANN dan SVM
Cinar et al. (2020)	Kecimen dan besni	LR, MLP, SVM
Khojastehnaz hand & Ramezani (2020)	Bulk raisin	SVM dengan gray level run length matrix (GLRM)
Guo et al. (2022)	Fake and real raisin	SVM, MCNN
Yajun et al. (2022)	Hongxiangfei, Manaiti, dan Munage	KNN, LDA, SVM
Tarakci & Ozkan, (2021)	Osmancik dan Cammeo	KNN, WKNN
Feng et al. (2018)	Wuhebai dan Xiangfei	SVM, KNN, RBFNN
Penelitian ini	Kecimen dan besni	ANN dengan variasi jumlah <i>hidden neuron</i>

## II. METODE

*Artificial Neural Network (ANN)* merupakan kecerdasan buatan yang merepresentasikan bagaimana cara kerja otak manusia dan mensimulasikannya ke dalam proses pembelajaran alat. ANN diinspirasi oleh adanya sistem syaraf dalam cara kerja impuls, sehingga ANN merupakan *interconnected* dari sistem syaraf buatan yang menggunakan model matematis dan komputasi.

Dalam penelitian Cinar et al. (2020) dilakukan proses identifikasi jenis kismis dari citra gambar yang didapat dari kamera kemudian diklasifikasikan menggunakan ANN *pattern recognition and classification* pada aplikasi MATLAB. *Network architecture* dari penelitian ini ditunjukkan pada Gambar 1. Pada *classification problem* ini, terdapat dua jenis kismis yakni kismis besni dan kecimen yang menjadi output, serta tujuh variabel sebagai input sebagai berikut.

1. Area: Jumlah piksel dalam batas kismis.
2. *Major axis length*: Panjang sumbu utama yaitu garis terpanjang yang dapat ditarik pada kismis.
3. *Minor axis length*: Panjang sumbu kecil yang merupakan garis terpendek yang dapat ditarik pada kismis.
4. *Eccentricity*: Ukuran *eksentrisitas* elips yang memiliki momen yang sama dengan kismis.
5. *Convex area*: Jumlah piksel kulit cembung terkecil dari wilayah yang dibentuk oleh kismis.
6. *Extent*: Rasio *wilayah* yang dibentuk oleh kismis dengan total piksel dalam kotak pembatas.
7. *Perimeter*: Jarak antara batas kismis dan piksel di sekitarnya.



Gambar 1. *Network architecture* dengan 7 input dan 2 output kategorikal.

Pada proses klasifikasi perlu dilakukan pengolahan data awal terlebih dahulu atau biasa disebut *preprocessing data*. Pada *preprocessing data* dilakukan beberapa tahap, yang pertama adalah

mendeteksi *missing value* dan *noise* dan diketahui pada dataset tidak terdapat data yang kosong dan tidak terdapat data yang diluar batas normal sehingga tidak ada *noise*. Kemudian tahap berikutnya adalah mendeteksi *outlier*. Meskipun terdeteksi adanya *outlier* namun model dianggap dapat mengakomodasinya sehingga *outlier* tidak dihilangkan. Kemudian pada tahap terakhir dilakukan transformasi data yaitu dengan mengubah data klasifikasi kismis ke dalam bentuk biner [0,1] menggunakan metode *encoding categorical* dan dilakukan data scaling menggunakan metode *min max scaler*.

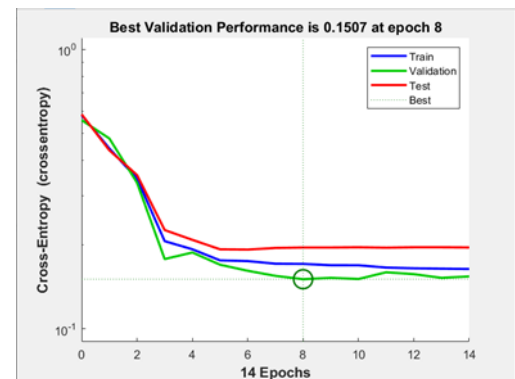
Setelah tahap *preprocessing data* selesai kemudian data akan digunakan sebagai input dan output di MATLAB. Langkah yang perlu dilakukan adalah sebagai berikut.

1. *Pattern Recognition and Classification App*  
Pada perintah *nstart* di MATLAB dipilih *pattern recognition and classification app* untuk mengklasifikasikan tiap data pada salah satu output kategori yang berbentuk biner.
2. Masukkan Data Input dan Output  
Data kemudian dimasukkan dalam bentuk file (.xlsx) atau excel sebagai input dan output. Data input dipilih sebagai input dan data biner output dipilih sebagai data output.
3. Pilih Jumlah *Hidden Neuron*  
Selanjutnya dilakukan penentuan jumlah *hidden neuron*. Pada penelitian ini digunakan lima level jumlah neuron yaitu 10, 20, 30, 40, dan 50.
4. *Train*  
Training pada setiap variasi *hidden neuron* dilakukan dengan algoritma *scaled conjugate gradient*. Dataset kemudian dibagi dengan pembagian 70% data digunakan sebagai *training data*, 15% data sebagai *validation data*, dan 15% data sebagai *test data*. Maka akan didapatkan hasil seperti jumlah iterasi *epoch*, *performance*, dan *gradient*.
5. *Plotting Hasil*  
Lalu dari *neural network training* dilakukan plot untuk *performance*, *error histogram*, *confusion matrix*, dan *receiver operating characteristic (ROC)*.

### III. HASIL DAN PEMBAHASAN

Proses training dalam pembangunan model prediksi dilakukan menggunakan cross-entropy sebagai loss function. Nilai cross-entropy yang

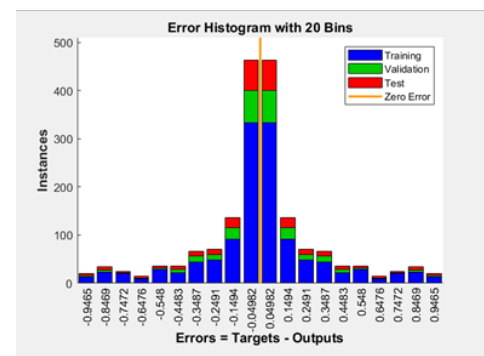
rendah menunjukkan bahwa model memiliki akurasi yang lebih baik. Gambar 2 menunjukkan progress dari proses training. Gambar tersebut menunjukkan penurunan nilai cross-entropy selama proses training. Penurunan secara signifikan terjadi pada epoch 1 sampai epoch 3. Setelah itu, nilai cross-entropy yang didapatkan baik untuk data training, test, dan validation cenderung stabil sampai proses iterasi training selesai pada epoch 14.



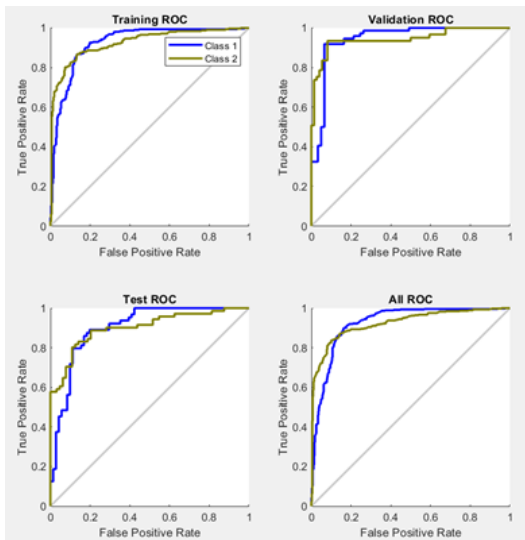
Gambar 2. Training progress



Gambar 3. Confusion matrix



Gambar 4. Error histogram



Gambar 5. Receiver operating characteristic

Setelah seluruh tahapan selesai, dapat diketahui bahwa untuk *neural network* dengan jumlah neuron 10 didapatkan nilai akurasi training sebesar 85,4%; akurasi validasi sebesar 90,4%; dan akurasi test

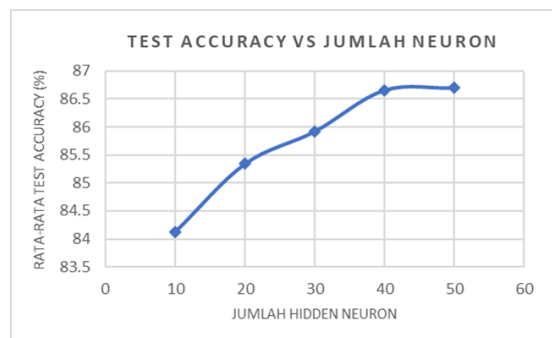
sebesar 84,4% sebagaimana yang ditunjukkan pada *test confusion matrix* pada gambar 3. Nilai akurasi yang akan menjadi pembanding tiap model pada studi ini adalah akurasi test sebagai nilai akurasi final dari *neural network* setelah dilakukan *training*. Detail dari hasil prediksi menggunakan test dataset berupa *error histogram* dan *receiver operating characteristic* ditunjukkan pada Gambar 4 dan 5.

Sebagaimana yang telah dijelaskan sebelumnya, data yang digunakan pada penelitian ini sebanyak 900 data biji kismis. Pada penelitian ini, ANN diimplementasikan untuk dataset tersebut dengan jumlah neuron dalam *network architecture* sebagai variabel bebasnya. Dengan menggunakan tujuh input, dilakukan klasifikasi data menjadi dua data output kategorikal yaitu kecimen dan besni. Terdapat lima level jumlah neuron yang ditinjau dengan tiap level dilakukan sebanyak lima kali trial. Hasil dari *tiap test accuracy* ditunjukkan pada Tabel 2.

Table 2. Nilai Test Accuracy (%) tiap Trial dan Rata-Ratanya

Jumlah Neuron	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Rata-Rata	Std. Dev
10	84.4	83.7	84.4	83.7	84.4	84.12	0.38
20	87.4	85.2	80	87.4	86.7	85.34	3.12
30	88.1	88.9	83	85.9	83.7	85.92	2.60
40	86.7	87.3	86.7	86.7	85.9	86.66	0.50
50	90.4	97.4	88.1	83.2	84.4	86.7	2.90

Berdasarkan hasil yang diperoleh diketahui bahwa peningkatan jumlah neuron diiringi dengan peningkatan rata-rata *test accuracy*. Namun demikian, dapat dilihat dari Tabel 2 bahwa hasil *test accuracy* dari tiap *trial* tidak selalu berbanding lurus dengan jumlah neuron. Hal ini kemungkinan terjadi karena pembagian dataset menjadi data *training*, *validation*, dan *test* dipilih secara *random* oleh aplikasi dari MATLAB. Meski begitu, hasil rata-rata *trial* tiap *level* menunjukkan adanya *trend* peningkatan sebagaimana yang ditunjukkan pada Gambar 6. Akan tetapi, pada jumlah neuron 40 ke 50, *test accuracy* tidak bertambah lagi secara signifikan. Hal ini selaras dengan teori yang ada bahwa pada jumlah neuron tertentu, *test accuracy* akan mencapai maksimum meskipun jumlah neuronnya bertambah atau yang disebut dengan *overfitting*.



Gambar 6. Kenaikan Rata-Rata Test Accuracy

#### IV. SIMPULAN

Berdasarkan analisis yang dilakukan pada penelitian ini dapat disimpulkan beberapa hal sebagai berikut:

1. Pada penelitian ini, proses klasifikasi kismis dilakukan menggunakan metode ANN dengan input berupa tujuh variabel yang menggambarkan karakteristik ukuran dan bentuk kismis untuk membedakan antara kismis berjenis besni dan kecimen. Proses klasifikasi ANN dilakukan dengan menggunakan lima level jumlah neuron,

mulai dari 10, 20, 30, 40, dan 50. Dari kelima pengujian yang dilakukan didapatkan hasil rerata akurasi tes tertinggi sebesar 86.7% dengan jumlah neuron sebanyak 50. Dari hasil penelitian, diketahui bahwa peningkatan jumlah neuron cenderung meningkatkan rata-rata hasil akurasi tes. Akan tetapi, pada jumlah neuron 50, peningkatan akurasi mencapai titik maksimum dan tidak lagi bertambah. Penerapan metode ANN pada proses klasifikasi kismis ini memberikan manfaat berupa peningkatan kualitas hasil klasifikasi, menurunkan biaya, serta mempersingkat proses pemilahan kismis.

2. Aplikasi klasifikasi kismis menggunakan ANN ini dapat diterapkan pada industri makanan maupun industri lainnya yang membutuhkan akurasi pemilahan kismis yang tinggi. Untuk penelitian lanjutan dengan metode klasifikasi ANN ini dapat dilakukan pengujian kembali agar dapat meningkatkan akurasi hasil klasifikasi dengan modifikasi arsitektur ataupun melakukan penambahan jumlah dataset.

## REFERENCES

- [1] Cinar, I., Koklu, M. and Tasdemir, S. 2020. Classification of Raisin Grains Using Machine Vision and Artificial Intelligence Methods, *Gazi Journal of Engineering Sciences*, 6(3): 200-209.
- [2] Feng, L., Zhu, S., Zhang, C., Bao, Y., Gao, P., and He, Y. 2018. Variety identification of raisins using near-infrared hyperspectral imaging. *Molecules*, 23(11), 2907.
- [3] Guo, J., Chen, C., Chen, C., Zuo, E., Dong, B., Lv, X., and Yang, W. 2022. Near-infrared spectroscopy combined with pattern recognition algorithms to quickly classify raisins. *Scientific Reports*, 12(1): 1-8.
- [4] Karimi, N., Kondrood, R.R., Alizadeh, T. 2017. An intelligent system for quality measurement of Golden Bleached raisins using two comparative machine learning algorithms. *Measurement*, 107: 68-76.
- [5] Khojastehnazhand, M., & Ramezani, H. 2020. Machine vision system for classification of bulk raisins using texture features. *Journal of Food Engineering*, 271, 109864.
- [6] Mollazade, K., Omid, M., Arefi, A. 2012. Comparing data mining classifiers for grading raisins based on visual features. *Computers and Electronics in Agriculture*, 84: 124-131.
- [7] UCI Machine Learning Repository. 2021 (April, 1st). Raisin Dataset Data Set. <https://archive.ics.uci.edu/ml/datasets/Raisin+Dataset>
- [8] Tarakci, F., and Ozkan, I. A. 2021. Comparison of classification performance of kNN and WKNN algorithms. *Selcuk University Journal of Engineering Sciences*, 20(2): 32-37.
- [9] Yu, X., Liu, K., Wu, D., and He, Y. 2012. Raisin quality classification using least squares support vector machine (LSSVM) based on combined color and texture features. *Food and Bioprocess Technology*, 5(5): 1552-1563.
- [10] Zhang, Y., Yang, Y., Ma, C., and Jiang, L. 2022. Identification of multiple raisins by feature fusion combined with NIR spectroscopy. *PloS one*, 17(7), e0268979.
- [11] Zoffoli, J.P. and Latorre B.A. 2011. Table Grape (*Vitis vinifera* L.). In E. Yahia (Ed.). *Postharvest Biology and Technology of Tropical and Subtropical Fruits: Cocona to Mango* (pp. 179-212). Woodhead Publishing Series in Food Science, Technology, and Nutrition. Schuur, "Mathematical model for warehouse design and product allocation," *International Journal of Production Research*. 2005, doi: 10.1080/00207540412331285841.