e-ISSN 2320 –7876 www.ijfans.org
Vol.11, S Iss 1, 2022
Research Paper UGC CARE Listed (Group -I) Journal

Combination Of Multispectral And Machine Learning Approach For Identifying Honey Adulteration

Barinderjit Singh[1*], Deepayan Padhy[2], Selva Ganapathy M[3], Yashi Srivastava[4],

N. Lenin Rakesh[5], Gurwinder Kaur[6], D R K Saikanth[7]

[1]Department of Food Science and Technology, I.K. Gujral Punjab Technical University, Kapurthala, Punjab, India

[2]Department of Entomology, M. S.Swaminathan School of Agriculture, Centurion University of Technology and Management, Odisha, India

[3]Department of Chemistry, Mount Carmel College, Autonomous, Bengaluru, Karnataka, India

[4]Department of Applied Agriculture, Central University of Punjab, Bathinda, Punjab, India

[5]Department of Mechanical Engineering, Bharath Institute Of Higher Education and Research, Chennai, Tamil Nadu, India

[6]Department of Food Science and Technology, I.K. Gujral Punjab Technical University, Kapurthala, Punjab, India

[7]Department of Agricultural Extension, College of Agriculture, Rajendranagar, Hyderabad, PJTSAU, Telangana, India

***Corresponding mail** id: barinderjitsaini@gmail.com

ABSTRACT

With the help of multispectral tomography data, this research develops a system for recognising honey polluted with jaggery water. A subsystem for plant origin identification is used to classify the floral source of a honey replica initially. An adulteration recognition subsystem determines the amount of the adulteration in the jaggery syrup once it has been identified. Each subsystem consists of two phases. The first stage is to extract important attributes from a sample of honey using direct analysis. In the additional stage, we use the K-Nearest Neighbours (KNN) typical to categorise the vegetal source of the honey in the first system and estimate the extent of adulteration in the second stage. We put the strategic approach to the test using a collection of publicly available honey multispectral photos. The findings show that the proposed system may effectively replace existing chemical-based detection techniques for identifying adulteration in honey, with an complete authentication accuracy score of 97.59%.

Keywords: KNN replica, machine learning, honey adulteration, multispectral image

1. Introduction

Food adulteration is a typical kind of fraud that attempts to make quick money by lowering food quality and having harmful financial and well-being values. Honey is one of the fluid foods that is prone to adulteration due to its larger price [1]. Furthermore, the obtainability of low-cost commercial jaggery syrups that may be blended with honey without affecting its flavour or colour has boosted the appeal of adulterating honey [2]. Honey may be polluted directly or indirectly by adding synthetic jaggery syrups or water to honey, or by feeding honeybees counterfeit jaggery.

Adulteration in honey has been discovered using well-developed diagnostic methods such as carbon isotope analysis, cinematography, and physiological parameter analysis[3]. Many studies have shown the efficiency of these events in identifying adulteration. However, expensive, laborious, damaging, and need investigation and sample preparation [4]. As a consequence, it is critical to augment current methods with rapid, non-destructive, and exact analytical approaches. Honey adulteration was previously discovered using nuclear magnetic resonance (NMR) spectroscopy. However, this procedure requires sample preparation, is time-consuming, and expensive. Multivariate analysis and terahertz spectroscopy were used to detect counterfeit honey. However, the detection replica was not particularly accurate [5]. In the identification of contaminated honey, NIR and VIS-NIR spectroscopy have shown promising finding. These spectroscopic approaches are more efficient, less costly, and need no sample preparation or specific training as compared to previous methods [6].

A three-dimensional picture is created using a multispectral imager. Geographical data is organised in the first two dimensions, whereas spectral data is organised in the third. A multispectral camera's typical spectral range is 400 to 1000 nm, including both the noticeable sections of the electromagnetic spectrum. Combining HSI with ML approaches like as Provision Course Mechanism and KNN resulted in the effective categorization of honey plant sources. Only one study has been conducted on the use of HSI to detect adulteration [7]. The replicas' performance was evaluated using a set of multispectral photographs of honey samples. 56 samples of pure and infected honey were used to collect data. The organization accuracy of the replicas utilising ANN, SVM, LDA, Fisher, and Parzen, respectively, was 95%, 92%, 90%, 89%, and 84%.

The different logical approaches utilised in preceding trainings were also meant to categorise honey's plant origins or to detect counterfeit honey. We propose a machine learning (ML)-based approach for categorising honey vegetal roots and identifying honey adulteration in this research. The spectral examples utilised in this investigation's data set were created by multispectral image segmentation of samples of clean and tainted honey from 11 distinct vegetal sources. The spectrum data for many samples of pure and contaminated Manuka plant honey are shown in Figure 1. The contaminated honey samples were created by adding 5%, 10%, 25%, and 50% jaggery syrup to pure honey samples. The collection contains 9778 rows of apparitional representations of honey examples from various flower bases. Each collection illustration includes 128 characteristics that represent spectral bands with wavelengths ranging from 400 to 1000 nm, rising by 5 nm [8].

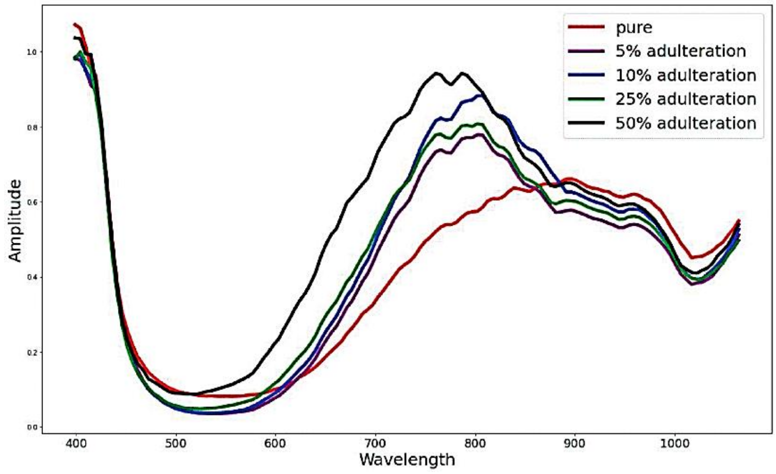

Fig. 1 The colour spectra of multiple specimens of genuine and falsified Manuka honey

2. Proposed system

Figure 2 depicts the two subsystems of the honey adulteration exposure structure suggested in this work: the adulteration detection subsystem and the subsystem for determining the plant origin of the honey. The first subsystem identifies the plant origins of honey. The second subsystem assesses the amount of adulteration and determines if the honey trial is pure or contaminated using a ML replica unique to a plant source. The goal of this subsystem is to categorise a honey sample into one of the collection's floral sources by determining the plant source of the honey by means of multispectral information. The plant origin identification technique is divided into two stages: feature extraction and categorization.

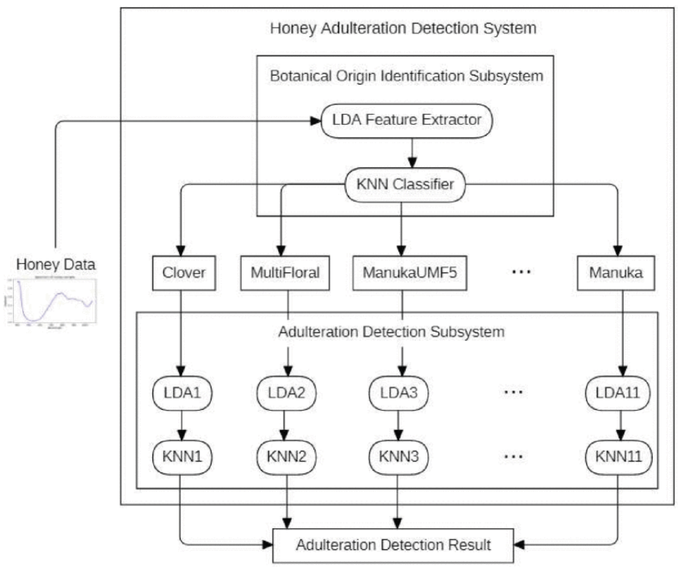

Fig. 2 Proposed architecture

The feature extraction technique is used to extract useful qualities from raw spectral data. Many classification issues need feature extraction since it improves organization replica concert by eliminating unnecessary landscapes from the dataset. The LDA technique was used to decrease dimensionality and extract features in this investigation. LDA is a supervised feature extraction and classification algorithm. To maximise class separation and enhance the relation of within-class variance to seperate-class variation, characteristics are translated into a lower-dimensional space using a linear approach. The LDA method was selected since our preceding effort on categorising honey plant heritages was effective. LDA is quick meanwhile it fair includes the resolution of a general eigenvalue problematic. It can to handle two and many-class problems, as well as nonlinear LDA utilising a quadratic equation kernel. Present investigate, we compare the efficacy of LDA to that of an unverified approach like as main constituent investigation (PCI).

During the organization stage, we employed the KNN replica to categorize the plant source of the honey test specimen. The presentation of the KNN classifier was assessed using the unique data, PCA-reduced topographies, and LDA structures. The act of the KNN perfect and the SVM classifier was also evaluated. Prior research suggested that the KNN and SVM classifiers were the most successful on a comparable dataset. The KNN classifier is an example of supervised nonparametric ML. It is a classifier that employs similarity metrics, with the distance between dataset samples acting as the resemblance measure. We used a value of 5 for the parameter k in the studies and the Euclidean as the distance metric since it produced the best results. The SVM technique determines the best in a interpreted large-dimensional interplanetary to discriminate between classes with the smallest faults. We employed two SVM classifiers with dissimilar kernel purposes to perform the experiments: a linear purpose and a centrifugal base function.

After the plant source proof of identity subsystem notices the floral basis of a honey sample, the adulteration recognition subsystem determines the degree of adulteration (jaggery content) in the honey trial. Jaggery was added to the honey samples used in this study in four different amounts: 5%, 10%, 25%, and 50%. The spectral data from these samples comprised the dataset utilised in this study. The strategy to detecting adulteration uses the same two processes as the previous subsystem's plant origin identification method using LDA and organization using KNN. This subsystem has 11 LDA replicas. The collection contains 11 distinct plant sources, and the features of each replica are taken from the honey spectral information of that particular plant source. The quantity of KNN replicas remains constant. To each replica categorises the amounts of adulteration in a certain kind of honey. The concert of the KNN classifiers was assessed using the unique features. The effectiveness of the KNN and SVM replicas was also examined.

3. Result and discussion

Figure 3 depicts the performance of the MI replica. The results reveal that using LDA-features and the KNN, the proposed method outperformed earlier strategies for identifying honey plant sources, with a truthful organization accuracy of 98.12%. The KNN classifier outclassed the SVM across all feature sets, with authentication precisions of 96.74%, 95.67%, and 98.01% by means of the unique structures, PCA and LCA -reduced, correspondingly. The correctness of the RBF SVM while utilising the unique features was 82.24%.

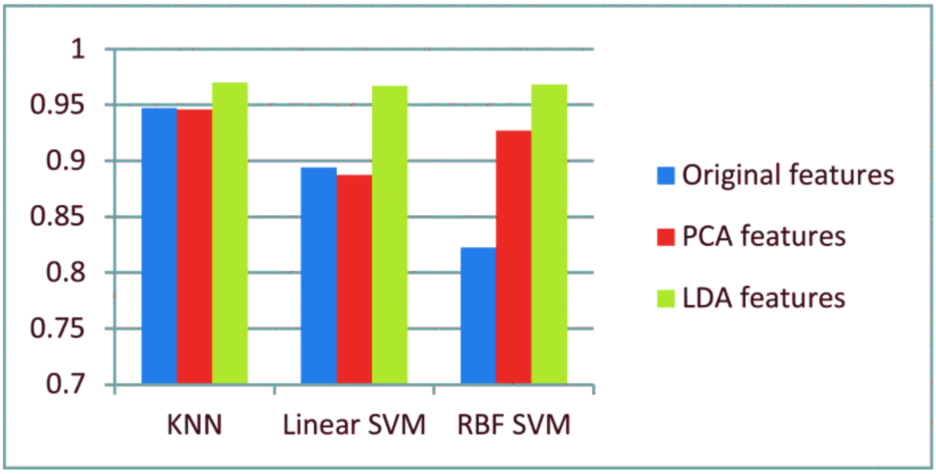

Fig. 3 Classifications of ML replica

The findings also show that, with the exception of the RBF SVM classifier, PCA had no effect on classifier performance. In comparison, the LDA technique meaningfully better the concert of all classifiers. Figure 4 demonstrations the efficacy of the classifiers for dissimilar LDA feature counts. The display demonstrations how using the top ten LDA features may improve classifier performance.

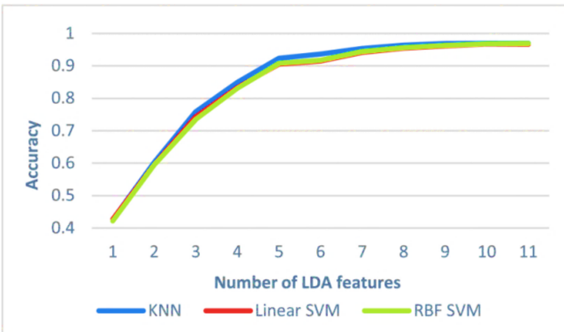

Fig. 4 Efficiency of ML replica

Figure 5 shows the average cross-validation accuracy of the three feature sets used in the ML replicas. When LDA-reduced and the KNN are used, the results show that the anticipated system notices adulteration in honey with an mean authentication precision of 97.49%. The efficiency of the classifiers in identifying adulteration varied depending on the kind of honey, with adulteration properly recognised in various honey types, including ManukaUMF5, ManukaUMF15, and ManukaUMF10.

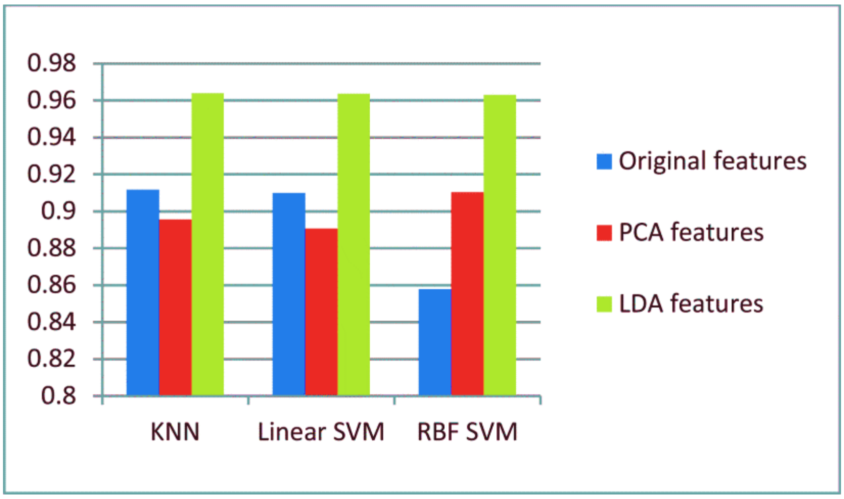

Fig. 5 Performance in finding the adulteration

The findings also show that when features were retrieved just using PCA, the RBF SVM classifier performed better. Dimensionality reduction using LDA, on the additional hand, significantly improved the concert of all classifiers. When PCA properties are used, RBF SVM outperforms KNN with an accuracy rating of 91.03%. Figure 6 portrays the visual operator communication of the adulteration uncovering structure formed in this research. On the user interface, the structure obtains and displays multispectral data from a honey. The plant source proof of identity recognises the fluorescent basis of the honey when the operator hits the Test Sample.

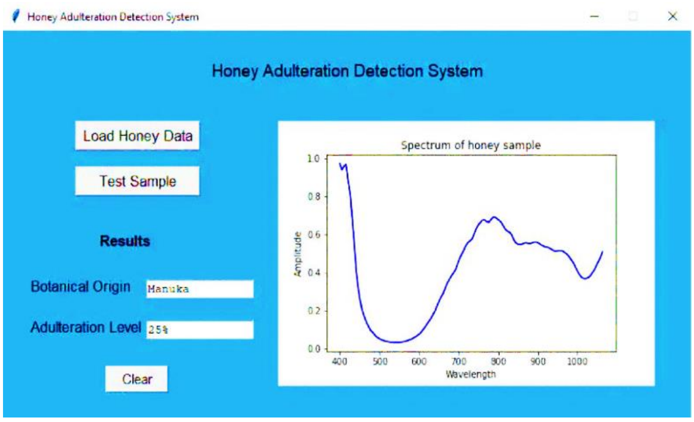

Fig. 6 GUI for adulteration

The purity of the honey test specimen is subsequently found by the adulteration detection. The UI shows the vegetal source as well as the quantity of adulteration. If the adulteration rate is 0, the honey sample is unadulterated. The graph portrays the spectral data plot, amount of adulteration, and plant source of the honey.

We developed a machine learning (ML)-based approach for distinguishing the plant origins of honey and detecting adulteration in this work. The LDA and KNN algorithms were utilised to achieve the highest accuracy in the current study, which was associated to preceding work utilising HSI honey information. The novel method identifies honey adulteration rapidly, automatically, and non-destructively by using honey multispectral imaging data. The KNN and SVM classifiers correctly differentiated between the numerous honey plant origins. Figures 7 and 8 indicate that LDA performed better than PCA because the classes employed by LDA were more distinct and the gaps between occurrences within the same class were less. The mean of the classifiers were modest, suggesting consistent concert. The occurrences of numerous classes overlapping in the dataset, as seen in Figure 8, explaining why the replicas were unable to categorise all forms of honey. The LDA decreased features increased the concert of the KNN and SVM replicas in identifying honey adulteration. The sample included a variety of honeys, all of which contained jaggery adulterants. As a consequence, the performance of the classifiers and the quantity of class labels change for each species of honey. The data supports the use of HSI and ML to detect tainted honey. The concert of the classifiers was affected by variations in adulterant kinds, honey floral bases, and adulteration attentions.

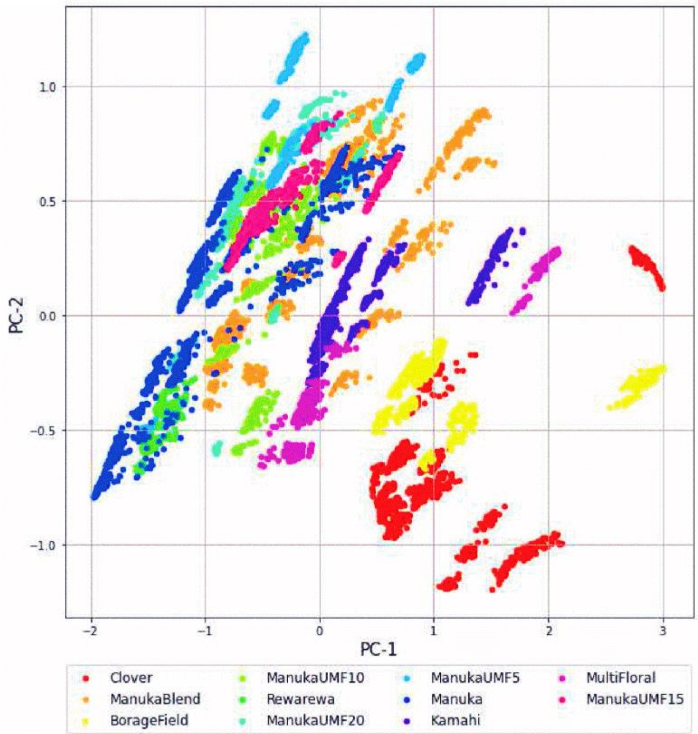

Fig. 7 Honey example matching on both the first and second critical elements

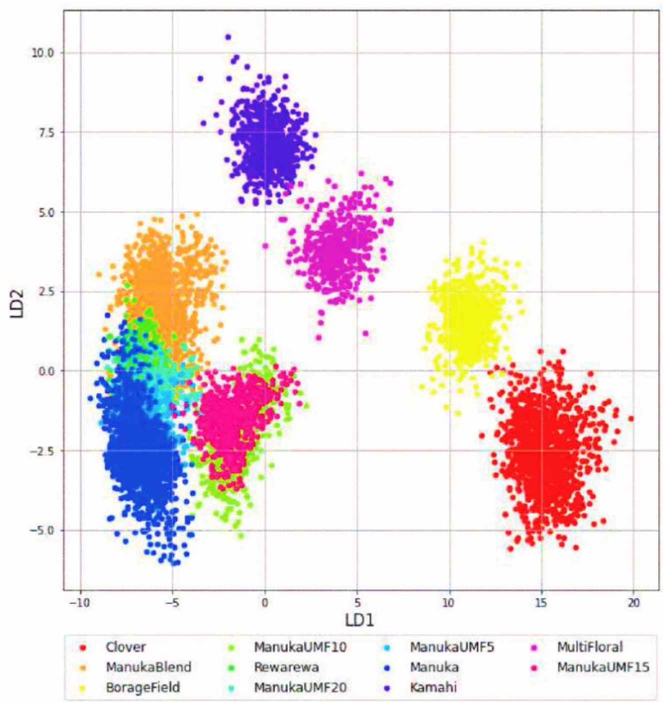

Fig. 8 Honey example mappings on the first and second linear classifier elements

Conclusion

The present study technologically advanced a technique for the robotic and non-destructive recognition of jaggery adulteration in honey using multispectral imaging and deep learning. The established adulteration detecting technique first establishes the plant source of a honey. The honey are the next examined for adulteration with jaggery syrup. A publicly accessible dataset of multispectral visualising data of unadulterated and polluted honey from numerous plant sources was used to assess the efficacy of the proposed technique. The established method positively detected adulteration in honey with an overall precision of 97.39%. The answers establish that a quick, low-cost, and non-invasive method for detection honey adulteration may be achieved by combining multispectral imaging with machine learning. Future research will apply regression replicas to boost our system's performance.

References

[1] A. A. Boateng, S. Sumaila, M. Lartey, M. B. Oppong, K. F. M. Opuni, and L. A. Adutwum, "Evaluation of chemometric classification and regression replicas for the detection of syrup adulteration in honey," *Lwt*, vol. 163, no. February, p. 113498, 2022, doi: 10.1016/j.lwt.2022.113498.

[2] K. Rachineni, V. M. Rao Kakita, N. P. Awasthi, V. S. Shirke, R. V. Hosur, and S. Chandra Shukla, "Identifying type of jaggery adulterants in honey: Combined application of NMR spectroscopy and supervised machine learning classification," *Current Research in Food Science*, vol. 5, pp. 272–277, 2022, doi: 10.1016/j.crfs.2022.01.008.

 ,

[3] S. Liu, D. Lang, G. Meng, J. Hu, M. Tang, and X. Zhou, "Tracing the origin of honey products based on metagenomics and machine learning," *Food Chemistry*, vol. 371, no. May 2021, p. 131066, 2022, doi: 10.1016/j.foodchem.2021.131066.

[4] Q. Li *et al.*, "Low risk of category misdiagnosis of rice syrup adulteration in three plant origin honey by ATR-FTIR and general replica," *Food Chemistry*, vol. 332, no. March, p. 127356, 2020, doi: 10.1016/j.foodchem.2020.127356.

[5] C. Maione, F. Barbosa, and R. M. Barbosa, "Predicting the plant and geographical origin of honey with multivariate data analysis and machine learning techniques: A review," *Computers and Electronics in Agriculture*, vol. 157, no. January, pp. 436–446, 2019, doi: 10.1016/j.compag.2019.01.020.

[6] A. Noviyanto and W. H. Abdulla, "Honey plant origin classification using multispectral imaging and machine learning," *Journal of Food Engineering*, vol. 265, no. January 2019, p. 109684, 2020, doi: 10.1016/j.jfoodeng.2019.109684.

[7] C. I. Mureșan, M. Cornea-Cipcigan, R. Suharoschi, S. Erler, and R. Mărgăoan, "Honey plant origin and honey-specific protein pattern: Characterization of some European honeys," *Lwt*, vol. 154, no. June 2021, 2022, doi: 10.1016/j.lwt.2021.112883.

[8] M. Lastra-Mejías, M. Izquierdo, E. González-Flores, J. C. Cancilla, J. G. Izquierdo, and J. S. Torrecilla, "Honey exposed to laser-induced breakdown spectroscopy for chaos-based plant classification and fraud assessment," *Chemometrics and Intelligent Laboratory Systems*, vol. 199, no. January, 2020, doi: 10.1016/j.chemolab.2020.103939.