



Article

# Application of Machine Learning in Differentiating Broth Containing Pork Fat and Chicken Fat Using UV LED Fluorescence Imaging System

Widayanti<sup>1</sup>, Friesca Ayazya N F<sup>2</sup>, Frida Agung R<sup>3</sup>

<sup>1,2,3</sup>UIN Sunan Kalijaga, Physics Department, Faculty of Science and Technology Yogyakarta, Indonesia

## SUBMISSION TRACK

Received: 12, 26, 2023  
Final Revision: 01, 24, 2024  
Available Online: 02, 02, 2024

## KEYWORD

High-Power, Pork Fat, Chicken Fat, Machine Learning, K-NN

## CORRESPONDENCE

E-mail: [widayanti@uin-suka.ac.id](mailto:widayanti@uin-suka.ac.id)

## A B S T R A C T

In Indonesia, individuals have been found engaging in fraud for selling soupy dishes by adding pork fat to the broth. It is quite challenging to identify the pork fat contaminated soup from other halal broth. Using Machine learning, this study attempts to identify and differentiate between RGB (Red Green Blue) values in picture of broth tainted with chicken and pork fat. The successful detection and differentiation of RGB values in broth contaminated with pork fat and chicken fat have been achieved. The broth samples were detected using a high-power UV-LED (Ultra Violet-Light Emitting Diode) Fluorescence Imaging System, while differentiation was accomplished through the implementation of a machine learning system. The data were processed using RapidMiner software with the K-NN algorithm. Detection was successfully performed through the spectrum of RGB values generated, while differentiation achieved an accuracy of 100%, precision of 100%, recall of 100%, and an AUC of 1.0.

## I. INTRODUCTION

Based on the MUI (Majlis Ulama Indonesia) fatwa outlined in decision No. 4 of 2003 on the standardization of Halal food products, pork and its derivatives are considered forbidden (haram) food products. However, in reality, because that it enhances tenderness and imparts a pleasant aroma to dishes, many individuals still add pork fat into processed foods [1][2]. "Kue pia" can have its texture improved by adding lard or pork fat, which makes the cake crispier and more tasty [3]. One of dishes is adding pork fat into soup. Based on appearance, taste, smell, and color, it is challenging to distinguish soup broth contaminated with pork fat from broth without it. Several methods are needed to detect pork fat content in soup. *Real-time Polymerase Chain Reaction* (RT-PCR) is a tool that has been used to detect pork fat. RT-PCR has been widely employed for analyzing biological samples, including amplifying DNA (deoxyribonucleic acid) molecules [4]. It exhibits high sensitivity [5] and has been developed in recent years for testing food halalness [6]. However, RT-PCR requires professional expertise and involves substantial costs. Therefore, there is a need for more low cost technological development [5]. An alternative method that can be developed is the use of fluorescence spectroscopy. The High Power UV-LED Fluorescence Imaging System (HPFIS) is a detection system designed and built by [7]. It has been successfully used for detecting contaminants in liquid detergent solutions [8], formaldehyde-contaminated tofu [9], borax-contaminated cilok [10], and borax-contaminated wet noodles [11]. The High power UV-LED fluorescence imaging system make use of the principle of fluorescence correlation spectroscopy as a UV light-based detection technique with a camera device. During the tool's detection process, a fluorescence event happens as a result of the interaction between the tool's light source and the atoms present in the sample. These atoms are then detected and recorded in the form of an image [7]. While HPFIS detections have been tested for accuracy

and precision in those detecting, but the tool has shown limitations in differentiating samples effectively. This limitation has inspired researchers to apply HPFIS as a detection tool for pork and chicken fat content in soup, combined with machine learning using the K-Nearest Neighbors (K-NN) algorithm [12] to differentiate RGB (red-green-blue) values in contaminated broth images. The K-NN technique has been extensively utilized in classification issues because of its efficiency and simplicity [13]. Using the majority of the data from the closet data category, this algorithm is used to classify an object. Since it is said to be the most effective technique for categorizing data [14], it is should be able to handle the differentiation problems in the classification of pictures of broths tainted with chicken and pork fat. Equation (1) through (4) are used in order to compute accuracy, precision, recall, and AUC (Are Under Curve) wich indicate how well RGB values in tainted broth images can be distinguished. Accuracy reflects the closeness between the system's predicted values and human predictions [12], and measurement is performed using classification performance, generally utilizing a confusion matrix. The confusion matrix records the classification performance results [15]. In binary classification problems for two classes, such as classes 1 and 0, each cell  $f_{ij}$  in the matrix represents the number of data from class  $i$  whose predictions fall into class  $j$ . For instance, cell  $f_{11}$  is the number of correctly predicted class 1 data (TP/true positive), and  $f_{10}$  is the data in class 1 incorrectly predicted as class 0 (TN/true negative). All data are expressed as the sum of TP, TN, FP, and FN, where FP is false positive, and FN is false negative. Precision is the number of true positive data (correctly recognized positive data) divided by the number of data recognized as positive, and recall is the number of true positive data divided by the number of truly positive data (true positive + true negative) [15]. AUC calculates the area under the ROC (Receiver Operating

Characteristic) curve. It is used to measure discriminative performance by estimating the output probability of randomly selecting a sample from the positive or negative population. The larger the AUC, the stronger the classification used [16] and [17]). AUC values range from 0.0 to 1.0.

$$\begin{aligned} \text{Accuracy} &= \frac{\text{the number of data predicted correct}}{\text{the total number of prediction made}} \times 100\% \\ &= \frac{f_{11} + f_{00}}{f_{11} + f_{10} + f_{00} + f_{01}} \times 100\% \\ &= \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \end{aligned} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP} \times 100\% \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \times 100\% \quad (3)$$

$$\text{AUC} = \frac{1 + TP_{rate} - FP_{rate}}{2} \quad (4)$$

## II. LITERATURES REVIEW

The identification and classification of texture-based digital images through feature extraction using GLCM (Gray Level CO-occurrence matrix) with the K-NN algorithm is one of the prior works that are closely related to the application of the K-NN method [18]. Afterward, a meatball categorization system with a borax content was developed utilizing the K-NN algorithm and an Arduino Mega, a PH sensor and color sensor [19]. The system achieved an accuracy 93,33%. With an accuracy 80,95%, a system for classifying chicken flesh utilizing an Arduino Mega, TGS26020 sensor, and K-NN for dead chicken, spoiled chicken and formalin-treated chicken meat is developed [20]. Marlis [21] then developed a 93,33% accurate predictive algorithm for white copra quality utilizing RGB color attributes and form (area and perimeter). The studies that have utilized HPFIS include research from [8][9][10] and [11]. Meanwhile detection of oil and fat substances has been conducted by [22] where detection of suspected pork oil in canola oil has

also been done using an electronic nose based on linear discriminant analysis (LDA). Additionally, [23] analyzed the presence of pork oil contaminants in face moisturizer using Fourier Transform Infrared Spectroscopy (FTIR). K-NN is an algorithm for classifying an object based on the majority data of the nearest data category. [24] used K-NN to classify tofu contaminated with formaldehyde, and [25] employed K-NN to classify skin cancer types based on texture and color features of images.

## III. FRAMEWORK

The framework used in this research includes 4 stages, including Preparation Stage, Data Collection, Machine Learning Model, implementation for differentiating. More detailed information about the research method can be seen in Figure 1.

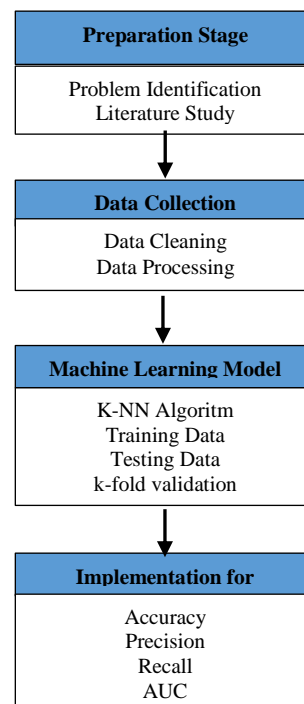


Figure 1. Research Flowchart

## IV. METHODS

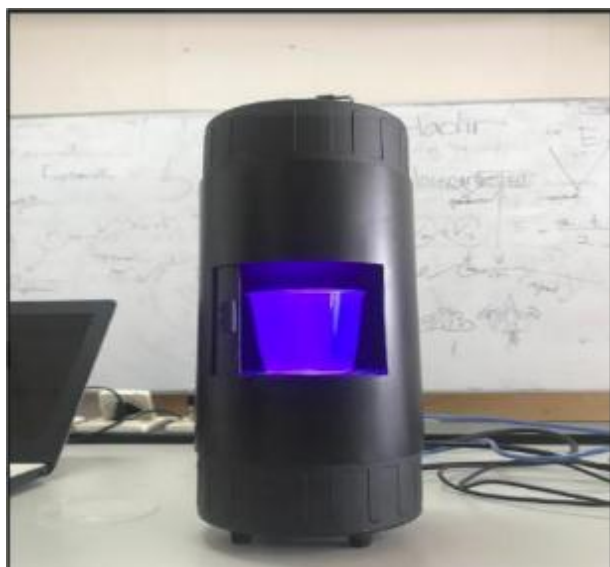
This research utilized the High Power UV-LED Fluorescence Imaging System (HPFIS) as illustrated in Figure 2(a). The system is composed of four high-power UV-LEDs as the light source with a wavelength of 395 nm [7]. It

also includes an image recording subsystem consisting of a M-Tech WB-100 webcam camera and a photocopier lens to enhance object clarity. The system records images generated from the fluorescence phenomenon of the samples and then displays the images along with RGB values on the HPFIS software. The differentiated samples are soup broths contaminated with pork fat and chicken fat. The image recorder is shown in Figure 2(b).

The samples are 10 cups of broth contaminated with pork fat and 10 cups of broth contaminated with chicken fat. Each cup of broth contains 0.5 grams of salt, 0.5 grams of finely chopped garlic, and water boiled to a temperature of 100°C. In 10 cups, pork fat is added with a mass variation from 0.5 grams to 5 grams, and in

another 10 cups, chicken fat is added with a mass variation from 0.5 grams to 5 grams to. All ingredients are stirred using a magnetic stirrer for three minutes at a speed of 300 rpm. Each cup (sample) is measured 10 times, resulting in a total of 100 data points for broths contaminated with pork fat and 100 data points for broths contaminated with chicken fat.

The research consists of three stages from preparation, data collection, and data processing. Data collection by HPFIS is conducted to obtain RGB values for each sample. The data collection process is illustrated in Figure 3.



**Figure 2(a). HPFIS**



**Figure 2(b). Software interface of UV Fluorescence Imaging System**

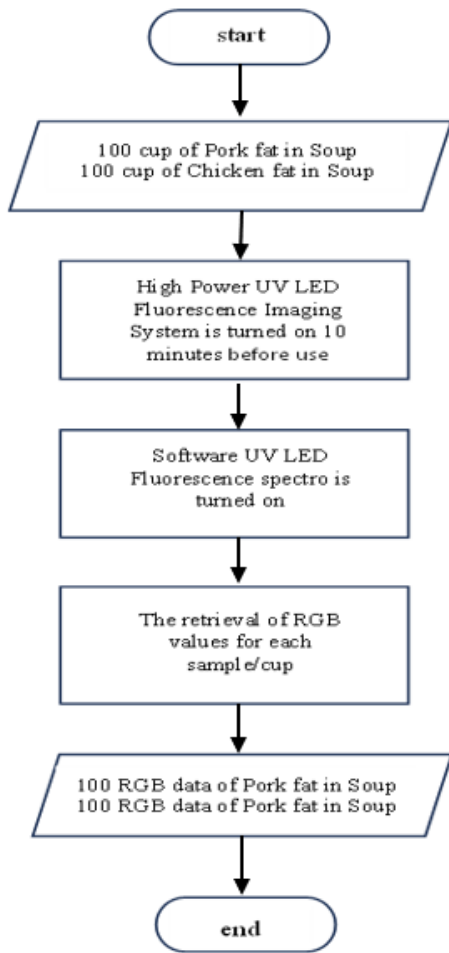


Figure 3. Flowchart of Data Collection

Whereas the RGB values of the 200 data obtained from the instrument are displayed through RapidMiner software, as shown in Figure 4. Data processing is conducted using machine learning with the K-NN algorithm,

which can differentiate the RGB values of broths contaminated with pork fat and chicken fat. The data processing involves three stages: preprocessing, processing, and evaluation, as depicted in Figure 5. Preprocessing involves creating new data before the actual processing, accomplished through sample imputation using bootstrapping with replacement. This ensures a consistent data quantity while introducing changes in the dataset to estimate standard errors. Processing is the data processing phase to obtain results from differentiating between pork fat and chicken fat using K-NN imputation. The algorithm's performance is validated through 10-fold cross-validation, enabling the determination of accuracy, precision, recall, and AUC based on the differentiation results. Evaluation is the process of calculating the optimal k-value by assessing accuracy through confusion matrices, precision, recall, and AUC. This step is crucial for determining the best k-value that maximizes accuracy while considering various performance metrics. The types of evaluation conducted in this study are accuracy, precision, and AUC (Area Under the Curve).

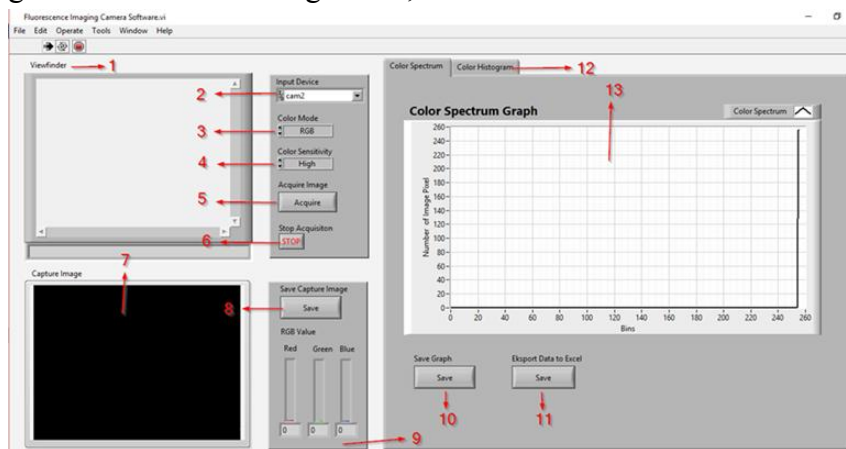


Figure 4. RGB Display

1. View finder
2. Active Webcam
3. Colour Mode
4. Colour Sensitivity Setting
5. Citra acquisition panel
6. Stop acquisition panel
7. Citra Display
8. Citra Restore panel
9. Average Histogram
10. Save Graph
11. Export Data to Excel
12. Color Spectrum
13. Color Histogram



10. Colour spectra Graph Panel 11. Export graph to excel panel 12. Colour hystogram Tab 13. Colour Spectra

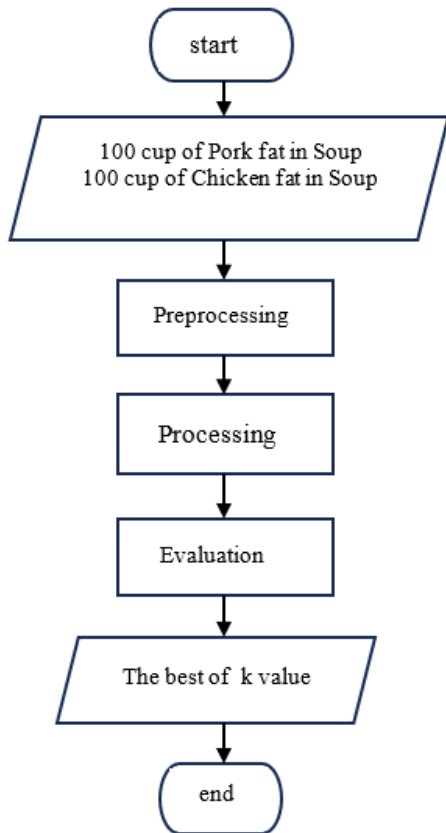


Figure 5. Data Processing

## V. RESULT

The results of detecting broths contaminated with pork fat, represented by RGB values, are shown in Figures 6(a) to 6(e), whereas chicken fat represented by RGB values, are shown in Figures 7(a) to 7(e). The RGB values obtained are numerical data that are subsequently input into *Microsoft Excel*.

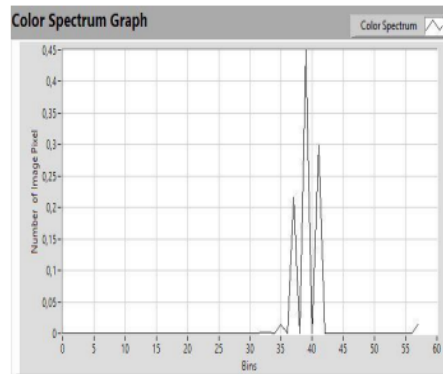


Figure 6 (a). RGB citra broth contaminated with pork fat at a concentration of 0.5 grams

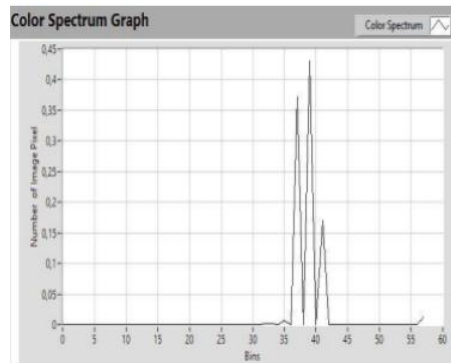


Figure 6 (b). RGB citra broth contaminated with pork fat at a concentration of 1,5 grams

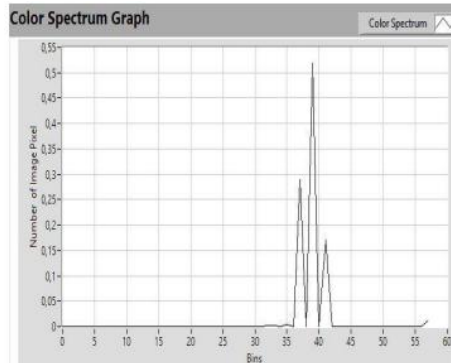
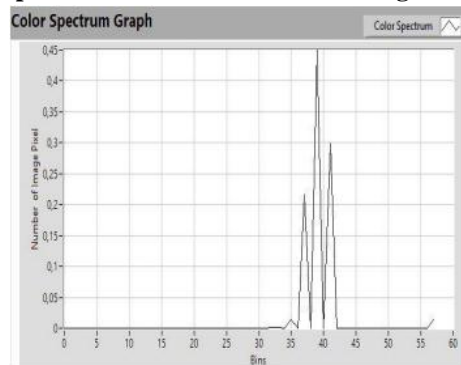
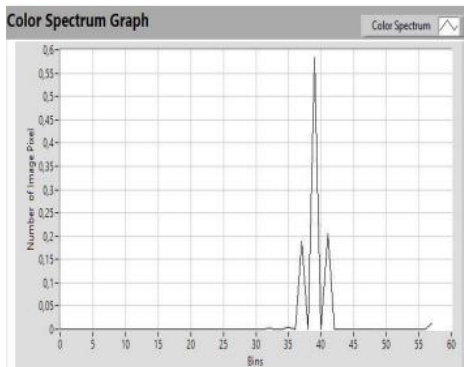


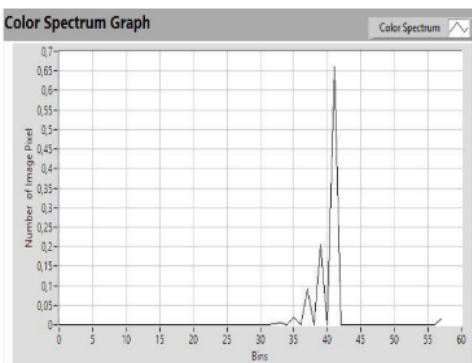
Figure 6 (c). RGB citra broth contaminated with pork fat at a concentration of 2.5 grams



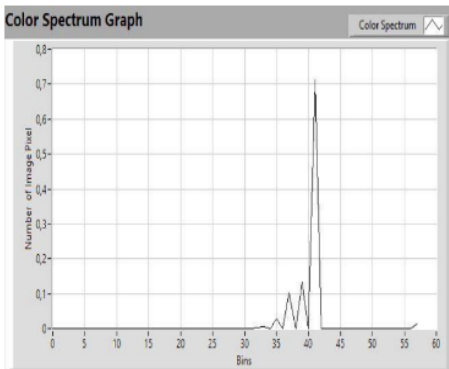
**Figure 6 (d). RGB citra broth contaminated with pork fat at a concentration of 3.5 grams**



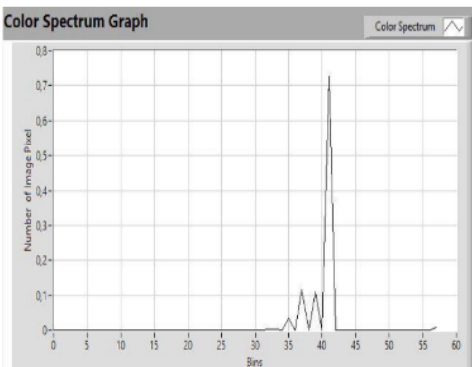
**Figure 6 (e). RGB citra broth contaminated with pork fat at a concentration of 5 grams**



**Figure 7 (a). RGB citra broth contaminated with chicken fat at a concentration of 0.5 grams**

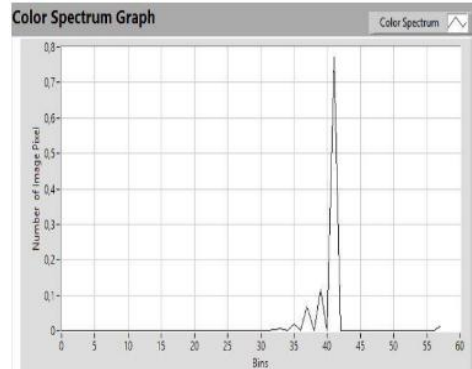


**Figure 7 (b). RGB citra broth contaminated with chicken fat at a concentration of 1,5 grams**

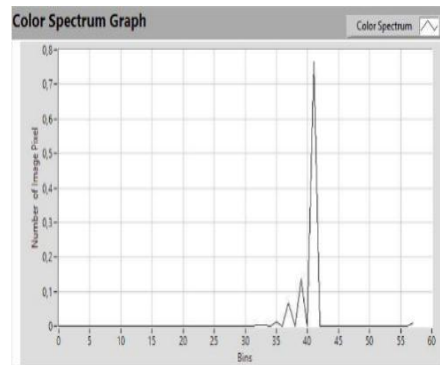


**Figure 7 (c). RGB citra broth contaminated with**

**chicken fat at a concentration of 2.5 grams**



**Figure 7 (d). RGB citra broth contaminated with chicken fat at a concentration of 3.5 grams**



**Figure 7 (e). RGB citra broth contaminated with chicken fat at a concentration of 5 grams**

The RGB values of the images were differentiated using machine learning with the K-NN algorithm. Bootstrap imputation, involving the imputation of the dataset, was employed to estimate standard errors in processing the RGB values of 200 sample data. The results of differentiating broths contaminated with pork fat and chicken fat are obtained from accuracy, precision, recall values, and AUC, as depicted in Figures 8(a) to 8(d).

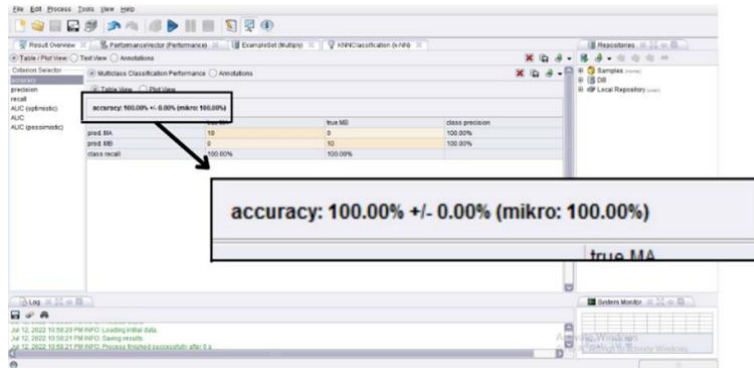


Figure 8 (a). Accuration of differentiation between broths contaminated with pork fat and chicken fat.

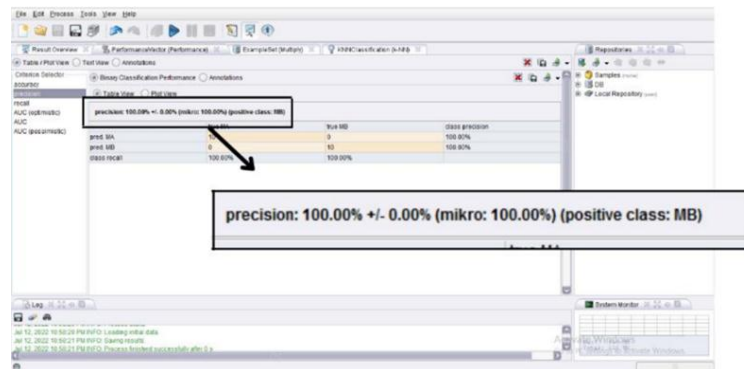


Figure 8 (b). Precision of differentiation between broths contaminated with pork fat and chicken fat

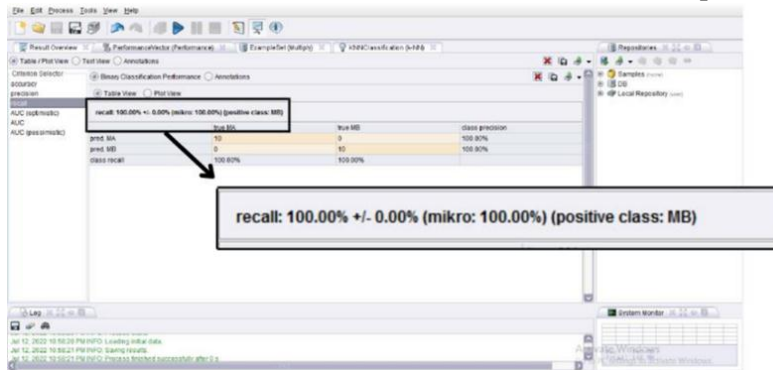


Figure 8 (c). Recall of differentiation between broths contaminated with pork fat and chicken fat

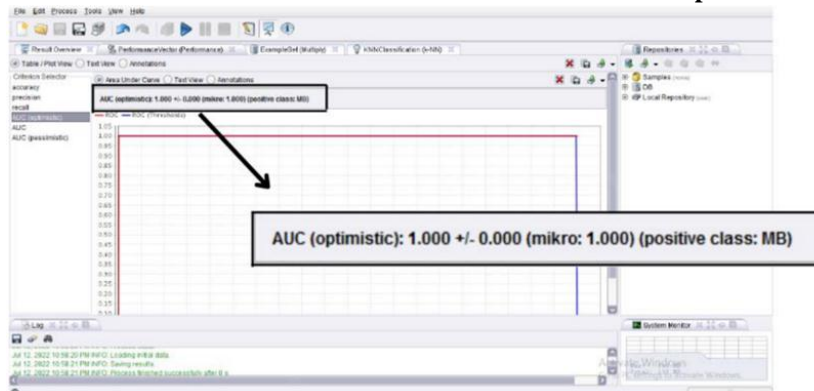


Figure 8 (d). AUC of differentiation between broths contaminated with pork fat and chicken fat



## V. DISCUSSION

Shown in Figure 6 (a) to 6 (e), the spectrum of each sample indicates the peak spectrum located in bin 39, and the color histogram is around the greenish color and shown in Figure 7 (a) to 7 (e), the spectrum of each sample indicates the peak spectrum located in bin 42, and the color histogram is around the bluish color. The difference in color spectrum shown in Figure 6 and Figure 7 is influenced by the number of bits in each pixel or different color ranges. The RGB color model used to display colors on monitors is based on the tri-stimulus vision theory, where the human retina is sensitive to the wavelength of light entering, with a wavelength of 630 nm for the color red, 530 nm for green, and 450 nm for blue [26]. Based on the obtained data, there is inconsistency in the data of broth contaminated with pig fat and chicken fat, which is influenced by the differences in contaminants found in the samples. The best  $k$  value is selected using parameter optimization, one of which is using *10-fold cross-validation*. In this study, the best  $k$  value used is  $k=1$ . As shown in Figure 8 (a), an accuracy result of 100% was obtained for differentiating broth contaminated with pig fat and chicken fat. This accuracy calculation uses a *confusion matrix*, where the number of data for each class can be determined. Based on the calculation results in Equation 1, an accuracy value of 100% is obtained. Figures 8 (b) and 8 (c) show the precision and recall results of differentiating broth contaminated with pig fat and chicken fat, which are both 100%. Precision and recall calculations use formulas in Equations 2 and 3.

The accuracy, precision, and recall values have exceeded the minimum Indonesian National Standard (SNI) value of  $\geq 97\%$ . Figure 8 (d) shows the AUC result of differentiating broth contaminated with pig fat and chicken fat, which is 1.0. This value is classified as *excellent classification*. Based on the overall results of accuracy, precision, recall, and AUC, it can be concluded that the high-power UV-LED fluorescence imaging system combined with machine learning and the K-NN algorithm can be used to differentiate broth contaminated with pig fat and chicken fat effectively.

## VI. CONCLUSION

Based on the obtained results, it can be concluded that broth contaminated with pork fat and chicken fat has been successfully detected using the previously designed High Power UV-LED Fluorescence Imaging System (HPFC). The RGB values of images of broth contaminated with pork fat and chicken fat were successfully differentiated using machine learning with the K-NN algorithm, achieving an accuracy of 100%, precision and recall of 100%, and an AUC value of 1.0.

## VII. ACKNOWLEDGEMENT

I would like to express our gratitude for the support provided by Faculty of Science and technology UIN Sunan Kalijaga Yogyakarta. The financial assistance from Research, Publication and Community Service Department (LPPM) UIN Sunan Kalijaga Yogyakarta is also greatly acknowledged.

## REFERENCES

- [1] D. A. Fikri, “Mengapa Mencampurkan Minyak Babi pada Masakan? Ini Alasannya,” *Okezone*, 2017. <https://lifestyle.okezone.com/read/2017/06/19/298/1720158/mengapa-mencampurkan-minyak-babi-pada-masakan-ini-alasannya>
- [2] A. Yulianto and B. Adji, “Waduh...Pemilik Warung Bakso Ini Akui Gunakan Kandungan Babi,” *republika*, 2022. <https://news.republika.co.id/berita/r54ya4396/waduhpemilik-warung-bakso-ini-akui-gunakan-kandungan-babi>
- [3] F. G. Winarno, *Kimia Pangan dan gizi*. P.T. Gramedia, 1984. [Online]. Available: [https://books.google.co.id/books?id=\\_P4StAEACAAJ](https://books.google.co.id/books?id=_P4StAEACAAJ)
- [4] H. Ma, K.-J. Shieh, G. Chen, X. Qiao, and M.-Y. Chuang, “Application of Real-time Polymerase Chain Reaction (RT-PCR),” vol. 2(3), 2006.
- [5] B. Huajun, C. Xiaolong, and Z. Xiaoyan, “Landscape Coronavirus Disease 2019 test (COVID-19 test) in vitro -- A comparison of PCR vs Immunoassay vs Crispr-Based test,” *Https://Osf.Io/6Eagn*. p. 13, 2020.
- [6] Y. L. Rahmania, T. W. Agustini, and M. Suzery, “Pengukuran Kandungan DNA Babi Dalam Berbagai Produk Pangan Dengan Metode Real Time- Polymerase Chain Reaction (RT-PCR),” *Indones. J. halal*, vol. 3, no. 2, pp. 129–133, 2021.
- [7] F. A. Rakhmadi, R. Rifai, and Khamidinal, “Design of First Generation of Sunan Kalijaga’s High Power UV-LED Fluorescence Spectroscopy System,” *Proceeding Int. Conf. Sci. Eng.*, vol. 3, no. April, pp. 17–19, 2020, doi: 10.14421/icse.v3.460.
- [8] A. F. Baraya, “Uji Presisi Fluorescence Imaging System Berbasis High Power UV-LED Generasi pada Sampel Larutan Detergen Cair,” 2020.
- [9] N. Rahmaningrum, F. A. Rakhmadi, and I. Fajriati, “Analisis Tahu Terkontaminasi Formalin Menggunakan Sistem Spektroskopi Fluoresensi Berbasis High Power Uv-Led,” *Sunan Kalijaga J. Phys.*, vol. 2, no. 1, pp. 29–33, 2020.
- [10] G. Haryarta, F. A. Rakhmadi, and I. Fajriati, “Analisis Cilok Terkontaminasi Boraks Menggunakan Sistem Spektroskopi Fluoresensi Berbasis High Power UV-LED,” *Sunan Kalijaga J. Phys.*, vol. 3, no. 3, pp. 28–35, 2021, [Online]. Available: <http://ejournal.uin-suka.ac.id/saintek/physics/article/view/2318/2387>
- [11] A. F. Hidayatulloh, “Analisis Mi Basah Terkontaminasi Boraks Menggunakan Sistem Spektroskopi Fluoresensi Berbasis High Power UV-LED Sunan Kalijaga Generasi Pertama.” 2021.
- [12] M. Azhari, Z. Situmorang, and R. Rosnelly, “Perbandingan Akurasi, Recall, dan Presisi Klasifikasi pada Algoritma C4.5, Random Forest, SVM dan Naive Bayes,” *J. Media Inform. Budidarma*, vol. 5, no. 2, p. 640, 2021, doi: 10.30865/mib.v5i2.2937.
- [13] T. A. Setiawan, R. S. Wahono, and A. Syukur, “Integrasi Metode Sample Bootstrapping dan Weighted Principal Component Analysis untuk Meningkatkan Performa K Nearest Neighbor pada Dataset Besar,” *J. Intell. Syst.*, vol. 1, no. 2, pp. 76–81, 2015.
- [14] A. Muhajirin, S. A. Sasosno, and T. Wangsalegawa, “Klasifikasi Algoritma K-Nearest Neighbor Untuk Memprediksi Barang Pada PT Enesis Group,” *J. Inf. Inf. Secur.*, vol. 2, no. 2, pp. 149–156, 2021, [Online]. Available: <http://ejurnal.ubharajaya.ac.id/index.php/jiforty>

- [15] T. Arifin, "Implementasi Metode K-Nearest Neighbor Untuk Klasifikasi Citra Sel Pap Smear Menggunakan Analisis Tekstur Nukleus," *J. Inform.*, vol. 2, no. 1, pp. 287–295, 2016, doi: 10.31311/ji.v2i1.83.
- [16] F. Gorunescu, *Data Mining: Concepts, models and techniques*. 2011.
- [17] A. P. Bradley, "The use of the area under the ROC curve in the evaluation of machine learning algorithms," *Pattern Recognit.*, vol. 30, no. 7, pp. 1145–1159, 1997, doi: [https://doi.org/10.1016/S0031-3203\(96\)00142-2](https://doi.org/10.1016/S0031-3203(96)00142-2).
- [18] D. Rohpandi, A. Sugiharto, and M. Y. S. Jati, "Klasifikasi Citra Digital Berbasis Ekstraksi Ciri Berdasarkan Tekstur Menggunakan GLCM Dengan Algoritma K-Nearest Neighbor," *J. VOI (Voice Informatics)*, vol. 7, no. 2, pp. 79–86, 2018.
- [19] D. D. Saputra, H. Fitriyah, and E. Setiawan, "Sistem Klasifikasi Bakso yang Mengandung Boraks dengan Sensor Warna Menggunakan Metode K-Nearest Neighbor Berbasis Arduino," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 10, pp. 10294–10300, 2019.
- [20] Zamaliq, F. Utaminingrum, and E. Setiawan, "Sistem Klasifikasi Kualitas Daging Ayam menggunakan Metode K-Nearest Neighbors berbasis Arduino," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. vol 4 no 5, no. 5, pp. 1455–1461, 2020.
- [21] R. R. Marlis, Abdullah, and F. Yunita, "Sistem Prediksi Kualitas Kopra Putih Menggunakan k-Nearest Neighbor (k-NN)," *Sist. J. Sist. Inf.*, vol. 10, pp. 290–299, 2021, [Online]. Available: <http://sistemasi.ftik.unisi.ac.id>
- [22] M. Muthmainnah, I. Tazi, S. Suyono, A. Ainur, F. Falah, and A. S. Santika, "Analisis Kandungan Minyak Babi Pada Minyak Kanola Melalui Klasifikasi Pola Hidung Elektronik (E-Nose) Berbasis Linear Diskriminan Analysis(LDA)," *J. Fis. Flux J. Ilm. Fis. FMIPA Univ. Lambung Mangkurat*, vol. 17, no. 1, p. 14, 2020, doi: 10.20527/flux.v17i1.5132.
- [23] Nurzakiyah, "ANALISIS KANDUNGAN MINYAK BABI PADA KRIM PELEMBAB WAJAH YANG MENGANDUNG MINYAK ZAITUN MENGGUNAKAN METODE SPEKTROSKOPI FTIR (Fourier Transform Infrared Spectroscopy)," UIN Alauddin Makassar, 2021.
- [24] D. Satriawan, H. Fitriyah, and A. S. Budi, "Sistem Klasifikasi Tahu Putih Murni dan Tahu Putih Mengandung Formalin Menggunakan Metode K-Nearest Neighbor," *J. Pengemb. Teknol. Inf. dan ilmu Komput.*, vol. 3, no. 10, pp. 10287–10293, 2019.
- [25] M. Faruk and N. Nafi'iyah, "Klasifikasi Kanker Kulit Berdasarkan Fitur Tekstur, Fitur Warna Citra Menggunakan SVM dan KNN," *Telematika*, vol. 13, no. 2, pp. 100–109, 2020.
- [26] Y. R. W. Utami, "K-Means Clustering Untuk Pengenalan Buah Berdasarkan Karakteristik Warna Citra," *J. Ilm. SINUS*, vol. 9, no. 2, pp. 11–20, 2011.

## BIOGRAPHY

**Widayanti** , currently working as a lecturer at Physics Departement, Faculty of Sciences and Technology Universitas islam Negeri Sunan Kalijaga Yogyakarta and has completed her doctoral studies in the field of Material Physics with GPA 3,80 at Universitas Gadjah Mada.

**Friesca Ayanza NH**, graduated in Physics Departement, Faculty of Sciences and Technology Universitas islam Negeri Sunan Kalijaga Yogyakarta and has completed her undergraduated studies in the field of Instrumentation Physics with at Universitas Islam Negeri Sunan Kalijaga Yogyakarta.

**Frida Agung Rakhmadi**, currently working as a lecturer at Physics Departement, Faculty of Sciences and Technology Universitas islam Negeri Sunan Kalijaga Yogyakarta and has completed her doctoral studies in the field of Instrumentation Physics at Universitas Gadjah Mada.