Animal and Plant-Based Milk Identification System Using Hyperspectral Imaging and Convolutional Neural Network

Nugi Asmara[1,*] Adhi Harmoko Saputro[1,**]

[1] *Department of Physics, Faculty of Mathematics and Natural Sciences, Universitas Indonesia, Depok, West Java, 16424, Indonesia*
*Corresponding author. Email: * nugi.asmara@sci.ui.ac.id, **adhi@sci.ui.ac.id*

ABSTRACT

Milk is a beverage that completes human nutrition. It is produced by cows and goats and can be obtained by plants such as soy and coconut. The nutrition composition contained in kinds of milk is different from one another. The differences in nutrition composition have their identification potential, such as the processing, nutrition differences, purity, quality, etc. Hence, it is necessary to build a system that can identify milk types with a non-destructive method utilizing hyperspectral images and a Deep Learning algorithm. This research used a hyperspectral camera at a Visible and Near-Infrared (VNIR) range of light (400 - 1000 nm). We used Convolutional Neural Network (CNN) as its image classification algorithm. Milk sample was collected from cow, goat, soy, and coconut and obtained exactly 1920 data. After the data was collected, we created datasets based on the type of classification tested. The category includes milk types with classes of animal-based and plant-based milk, the organisms that produce the milk with classes of coconut, cow, goat, and soy, and the processing method with classes of fresh and Ultra High Temperature (UHT). The tested algorithms of CNN architecture are GoogleNet, AlexNet, and Proposed CNN. The highest accuracy for 480 data was 100% reached by processing method classification of soy milk, and the computation took only 20 seconds. Meanwhile, the highest accuracy for 1920 data was 99.9% achieved by Proposed CNN architecture, and the calculation took only 78 seconds. These results showed that hyperspectral imaging and CNN algorithm are suitable for classifying types of milk.

Keywords: hyperspectral, milk, animal-based, plant-based, UHT, CNN

1. INTRODUCTION

Milk is a beverage that completes human's nutrition and can be obtained from mammals such as cow, goat, camel, buffalo, etc. Besides it is produced by mammals, milk technically can also be obtained by several plants such as soy, almond, rice, and coconut. Those milks contain important main contents like protein, carbohydrate, fat, and so on [1], [2]. Their visual appearance is not significantly different with each other, hence it is necessary to build a system that can distinguish their nonvisual characteristics by its compounds composition.

Several methods have been developed to identify different types of milk, such as contact apparatus like electronic tongue [3], cationic polymer perylene probe [4], and spectroscopy-based instruments like mass spectroscopy [5], near-infrared spectroscopy [6], and FTIR spectroscopy [7]. However, those methods still have some disadvantages. The sample preparations are destructive, required plenty of reagent and other chemicals [3], [4], [8], and it only provided spectral information in certain range of wavelength [7].

Hyperspectral imaging can be alternatives for these issues, but it requires algorithm that capable to process its huge spatial and spectral information or data as an input. Therefore, we used deep learning as its processing algorithm. Convolutional Neural Network (CNN) has been commonly used for object identification, so that it is suitable to be used as the processing algorithm of hyperspectral images for classification and identification purposes. CNN also capable for automatically learn input's features so that it

doesn't require any additional feature manipulation [9], [10].

2. MATERIAL AND METHOD

The proposed measurement system consists of hyperspectral camera Specim FX-10 that suitable to capture lights at 400 – 1000 nm, aluminium holder for the camera, halogen lamp Philips QVF133 HAL-TDS, teflon, petri dish, and slider. The acquisition mode used line scanning method and captured reflectance of the light reflected from the milk sample on petri dish. Training and testing datasets were run at computer with Intel® Core™ i5-9400F CPU @ 2,90 GHz processor, 8 GB RAM, and NVIDIA GeForce GTX 1650 graphic card (Fig. 1).
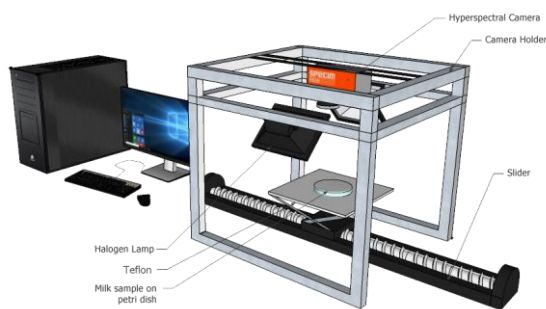
Figure 1 Measurement system for hyperspectral image acquisition of milk sample at petri dish

The milk samples are collected from supermarket, minimarket, stalls around Depok and Jakarta, and online stores a day before data acquisition and being stored at the referigreator. Hence all the samples were still fresh when they were being captured. The samples collected then divided and stored to each 5 200-mL bottles per brands and stores. The samples were poured only 100 mL each bottle. The volume of 100 mL had been set to fit the maximum capacity of petri dish. Then the bottles were labeled based on the initial of each brand and store. After the samples were completely divided, then image acquisition could be done. The samples at the bottles were poured onto 1.5-cm-height petri dish, and the distance to the lens of the camera were approximately 15 cm.

Image used in this study were consisted of 3 input hyperspectral image: white callibration, dark callibration, and raw image. These 3 hyperspectral images then used for correction purpose with relative reflectance correction. The equation for determining relative reflectance correction is shown at equation (1). Where R is the reflectance value, IR is raw image intensity, ID is dark callibration intensity, and IW is white callibration intensity. After the image had been corrected, then they would be segmented into 8 10×10-pixel ROIs using circle detection technique. Then the segmented images were annotated with classes required, and the annotated image then saved into datasets for

deep learning inputs. CNN algorithm architectures used in this study were GoogleNet, AlexNet, and AlexNet-based Proposed CNN. Evaluation parameters used for analyzing were accuracy, precision, sensitivity, specificity, and computation time. Equation (2) to (6) shows evaluation parameter used in this study.

$$R = \frac{I_R - I_D}{I_W - I_D} \tag{1}$$

$$\text{Accuracy} = \frac{TTP}{\text{total data}} \tag{2}$$

$$\text{Precision}_i = \frac{TTP}{TTP + TFP_i} \tag{3}$$

$$\text{Sensitivity}_i = \frac{TTP}{TTP + TFN_i} \tag{4}$$

$$\text{Specificity}_i = \frac{TTN_i}{TTN_i + TFP_i} \tag{5}$$

$$t_{\text{computation}} = \frac{t_{\text{train}} + t_{\text{test}}}{k} \tag{6}$$

The classifications were grouped into 3 main categories, they were based on types of milk with classes of animal-based and plant-based, based on organisms with classes of coconut, cow, goat, and soy, and based on processing procedure with classes of fresh and UHT (Ultra High Temperature). The classification based on processing procedure were also divided into 4 group: processing procedure for whole kind of milk, processing procedure of coconut milk, cow milk, and soy milk.

3. RESULT AND DISCUSSION

3.1. Image Correction and Segmentation

The acquired image shown at Fig. 2 (a) were corrected using relative reflectance correction to calibrate and obtain smoother reflectance value and lower noise, thus the following image processing performance can be run better. By this correction process we expected that the calibrated spectral information can be used properly for analysis purposes. The corrected images were then segmented using circle detection method to locate 8 10×10-pixel ROIs as shown at Fig. 2 (c) and the result of segmentation process is shown at Fig. 2 (d).

These 8 ROIs then saved into annotated datasets with classes corresponding to the sample's label and category, based on its types, organisms, processing procedure, and brand or store initials
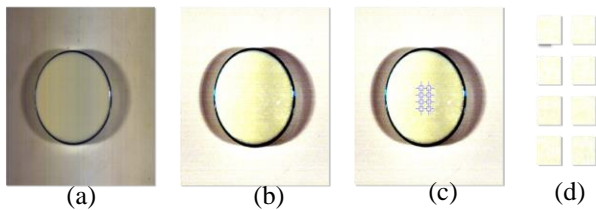
Figure 1. Results of image acquisition, correction, and segmentation of milk sample on petri dish.

3.2. Milk Spectral Result

We calculated the segmented images average reflectance for each wavelength from 400 – 1000 nm and the graph then plotted as shown at Fig. 3. Fig. 3 shows graph of reflectance from each sample after they are grouped by organisms, types, and processing procedures.
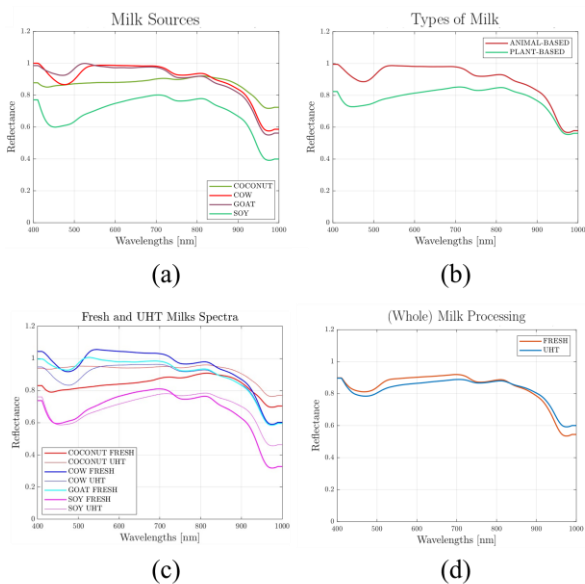
Figure 3. Reflectance pattern obtained of milk samples after being grouped by type, organism, and processing procedure.

Based on Fig. 3 (a) the reflectance pattern of cow and goat milk are not significantly differs, meanwhile we can see that reflectance of soy and coconut milk are significantly differs. It indicates that contents between cow and goat milk are quite similar, while the differences between soy and coconut milk contents causes differences in interaction with light from halogen lamp, so that the reflectance pattern differs as well. Fig. 3 (b) shows that there is significant difference between reflectance pattern of animal-based milk and plant-based milk at range 400 – 900 nm, while both curves are narrowing at 900 – 1000 nm. It indicates there are composition similarities that corresponds to reflectance at 900 – 1000 nm. Figs. 3 (c) and 3 (d) shows how different reflectance of processing procedures of milk from different organisms. Based on Fig. 3 (c) we can

see that for fresh cow milk, the reflectance is higher than the UHT one. Otherwise, coconut milk reflectance gives the opposite order of reflectance pattern where the fresh coconut milk is lower than the UHT one. Different from cow and coconut milk, soymilk's reflectance obtained intersection between fresh and UHT at about 730 nm where the UHT milk has higher reflectance than the fresh at 400 – 730 nm, and they flip at 730 – 1000 nm. These differences reflectance pattern for different kind of milk is caused by different treatment for each UHT procedure. The UHT procedure for cow milk causes loss at some contents [11], while UHT technique for soy and coconut milk requires some content addition so that it causes some contents percentage also increased [12], [13].

Based on graph at Fig. 3 (d) we can see that fresh and UHT milk reflectance are not significantly differs, they look narrow each other. It shows that reflectance value of UHT milk is lower compared to reflectance of fresh milk at range 400 – 850 nm. Meanwhile reflectance of UHT milk then appearantly higher than the reflectance of the fresh one at range 850 – 1000 nm. It is probably caused by different persentage of main content composition in either fresh milks or UHT milks. These insignificant differences also can be caused by dataset merging that has been done to whole sample of milk: fresh cow, UHT cow, fresh goat, fresh soy, UHT soy, fresh coconut, and UHT coconut milk. They have significant differences one another so that grouping to whole milk for fresh and UHT for 4 milks obtained a nonsignificant reflectance difference.

Table 1. Content percentage (%) of each milk sample based on direct measurement, references, and nutrition fact label.

Organism	Processing Procedure	Fat	Protein	Carbohydrat	Water
Cow	Fresh[a]	4.95	4.20	4.49	n.d.
	UHT[d]	3.41	3.09	5.01	n.d.
Goat	Fresh[a]	4.56	3.77	3.58	n.d.
Soy	Fresh[b]	1.91	2.92	4.92	91.90
	UHT[d]	2.48	2.70	8.53	88.12
Coconut	Fresh[c]	25.57	2.35	3.88	67.39
	UHT[d]	26.00	2.00	6.67	74.79

Description:
[a] taken from direct measurement with Lactoscan
[b] Rasika et al., 2021
[c] Joe et al., 2021
[d] taken from sample's nutrition fact label

Table 1 shows contents composition percentage that may be consisted in the tested samples. These values are collected from direct measurement for animal-based milk using Lactoscan, refers to some related references about fresh plant-based milk study, and refers to nutrition fact labels for UHT plant-based milk products.

The table doesn't provide measurement value directly by researcher, and it used only to show that there are content differences between some different milks. The table is also used as a supporting information for analysis purposes matched with available information that has been obtained. Based on Table 1, we can see that there are differences for each fresh samples from different organisms after being treated by UHT technique. UHT-treated cow milk has lower fat and protein content compared to the fresh one. Meanwhile for UHT-treated coconut milk, fat, carbohydrate, and water contents are higher than the fresh one. These different changes are consistent with the reflectance obtained as shown at Fig. 3 (c) where UHT-treated cow milk reflectance is lower than the fresh one, while the opposite pattern occurred at UHT and fresh coconut milk.

3.3. Classification Performance Result

The classification model architecture proposed in this study was called Proposed CNN where it was a modified version of AlexNet. The model was compared to two other classification architecture, they are GoogleNet and AlexNet. The data used for input was annotated segmented hyperspectral image datasets and the training and testing process would be validated using k-fold cross validation. The data for each training and testing process would be divided into k iterations, where we set the k = 5. Hence, the datasets were divided into 80% for training and 20% for testing. The evaluation measured for algorithm model performance analysis are accuracy, precision, sensitivity, specificity, and time comsumed for computation. Four of them utilize calculation using confusion matrix and then we calculated the score of total true positive, total true negative, total false positive, and total false negative. The scores were then used to calculate the percentage of accuracy, precision, sensitivity, and specificity.

Figs. 4, 5, and 6 shows the percentage of each performace percentage after the datasets were used for input of the classification models. The datasets were tested to 3 algorithm architectures of CNN, they are GoogleNet, AlexNet, and Proposed CNN. From the graphs with the following tables, we can see that the architecture performances reach percentage more than 90%. We can see that GoogleNet has the lowest minimum performance percentage where the evaluation score for accuracy, precision, sensitivity, and specificity varied at 90% - 100%. Followed by AlexNet with evaluation performance percentages are varied at 93% - 100%. Lastly, Proposed CNN got the highest result
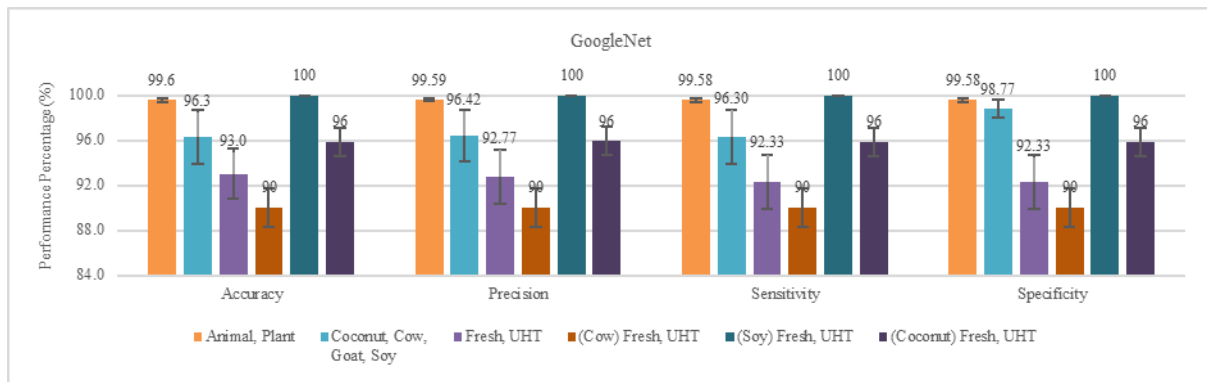

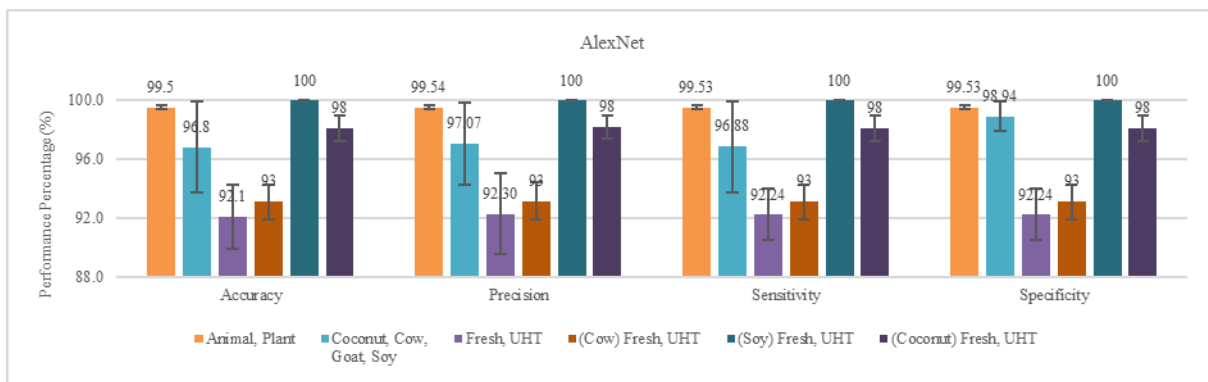**Figure 2.** Evaluation performance scores for GoogleNet architecture.

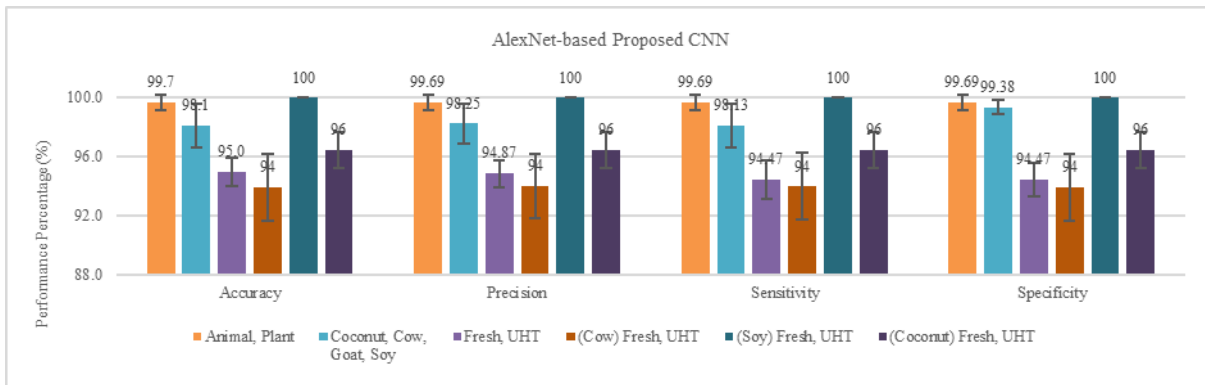**Figure 5.** Evaluation performance scores for AlexNet architecture.

Figure 6. Evaluation performance scores for Proposed CNN architecture.

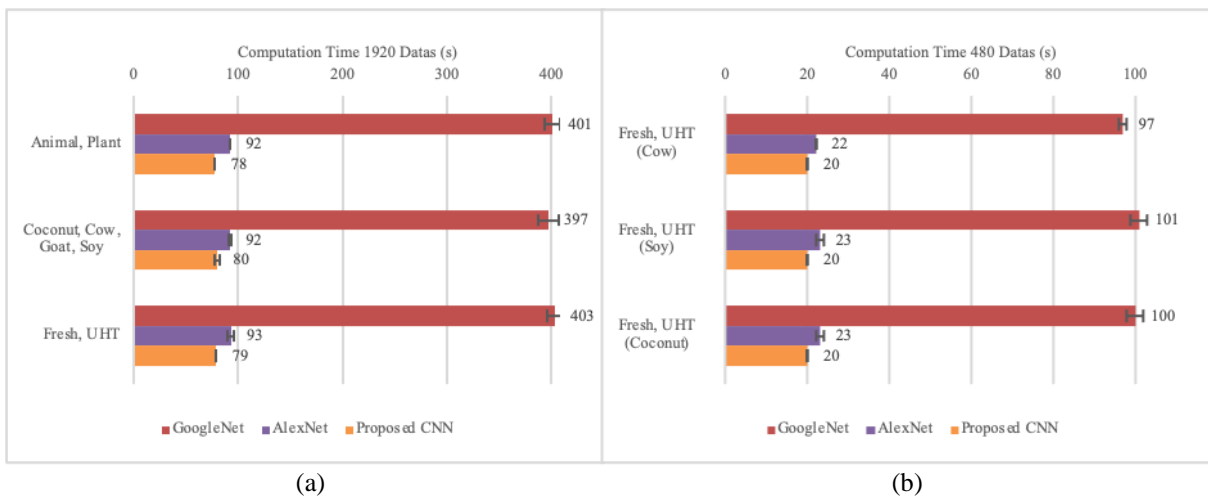| (a)   (b) |

Figure 7. Computation time for classification process of total (a) 1920 datas, and (b) 480 datas.

compared to 2 previous architecutre where the evaluation performance percentages are varied at 95% - 100%. The lowest classification group score was obtained by classification by processing procedure of cow milk with range of performance percentages are 90.00% to 94.01%, while the highest classification group score was obtained by processing procedure of soymilk with perfect score 100%.

Based on the graph, we can see that the evaluation parameters are different for each classification group. It shows that for the same structure of architecture for classification model, has different complexity for the learning models, so that for few numbers of classification, the related architecture can perform very well while for others it is just suitable enough does not perform as good as some others.

Other parameter used to analyze the performance of the algorithm was time consumed for computation or learning process included time for training and time for testing. The calculation was divided into 2 kinds of analysis, for total 1920 datas, and for total 480 datas.

Based on Fig. 7 (a) we can see that the longest computation time for total 1920 datas classification was obtained by GoogleNet with time consumed for learning

process was 397 – 401 seconds, while the shortest computation time was obtained by Proposed CNN with time consumed for learning process was 78 – 80 seconds. Meanwhile, for total 480 datas classification, based on Fig. 7 (b) we can see that the longest computation time was still obtained by GoogleNet with time consumed was 97 – 101 seconds, while the shorstest computation time was also obtained by Proposed CNN with time consumed only 20 seconds. It was because of kernels used for modifying the architecture from AlexNet to Proposed CNN was reduced. It causes the complexity of the architecture decreased then the computation could perform fasster. Bianco et al. (2018) and Shin et al. (2021) reported that the complexity of the alghorithm including layer depth and structure on CNN architecture might be one of factors that influence the time consumed for the learning models to process the computation [14], [15].

4. CONCLUSION

From this study, we conclude that deep learning model using hyperspectral imaging and Convolutional Neural Network (CNN) is suitable for identifying and classifying different kinds of milk either by the type, organism, and processing procedure with evaluation scores are more than 90%. The highest score of image

processing and sample classification algorithm was obtained by Proposed CNN architecture with accuracy, precision, sensitivity, and specificity percentage reached 95% - 100%, followed by AlexNet with accuracy, precision, sensitivity, and specificity percentage reached 93% - 100%, and GoogleNet with accuracy, precision, sensitivity, and specificity 90% - 100%. Proposed CNN was architecture model with the highest accuracy, precision, sensitivity, and specificity and also the most efficient for computing processes compared to AlexNet and GoogleNet for all classifications.

ACKNOWLEDGMENTS

REFERENCES

[1] Y. W. Park and G. F. W. Haenlein, Eds., Milk and Dairy Products in Human Nutrition. Wiley-Blackwell, 2013.

[2] D. M. Rasika et al., "Plant-based milk substitutes as emerging probiotic carriers," Curr. Opin. Food Sci., 38, pp. 8–20, 2021, doi: 10.1016/j.cofs.2020.10.025.

[3] L. A. Dias, A. M. Peres, A. C. A. Veloso, F. S. Reis, M. Vilas-Boas, and A. A. S. C. Machado, "An electronic tongue taste evaluation: Identification of goat milk adulteration with bovine milk," Sensors Actuators, B Chem. 136, no. 1, pp. 209–217, 2009, doi: 10.1016/j.snb.2008.09.025.

[4] L. Zhang et al., "Identification of milk adulteration by a sensor array based on cationic polymer induced aggregation of a perylene probe," Food Chem. 343, no. June 2020, p. 128492, 2021, doi: 10.1016/j.foodchem.2020.128492.

[5] P. England, W. Tang, M. Kostrzewa, V. Shahrezaei, and G. Larrouy-Maumus, "Discrimination of bovine milk from non-dairy milk by lipids fingerprinting using routine matrix-assisted laser desorption ionization mass spectrometry," Sci. Rep. 10, no. 1, pp. 1–7, 2020, doi: 10.1038/s41598-020-62113-9.

[6] E. V. dos S. Pereira, D. D. de S. Fernandes, M. C. U. de Araújo, P. H. G. D. Diniz, and M. I. S. Maciel, "Simultaneous determination of goat milk adulteration with cow milk and their fat and protein contents using NIR spectroscopy and PLS algorithms," Lwt. 127, no. November 2019, p. 109427, 2020, doi: 10.1016/j.lwt.2020.109427.

[7] P. Jaiswal, S. N. Jha, A. Borah, A. Gautam, M. K. Grewal, and G. Jindal, "Detection and quantification of soymilk in cow-buffalo milk using Attenuated Total Reflectance Fourier Transform Infrared spectroscopy (ATR-FTIR)," Food Chem. 168, pp. 41–47, 2015, doi: 10.1016/j.foodchem.2014.07.010.

[8] X. Zhu and F. Kang, "Frequency- and Temperature-Dependent Dielectric Properties of Goat's Milk Adulterated with Soy Protein," Food Bioprocess Technol. 8, no. 11, pp. 2341–2346, 2015, doi: 10.1007/s11947-015-1593-x.

[9] P. Wang, E. Fan, and P. Wang, "Comparative analysis of image classification algorithms based on traditional machine learning and deep learning," Pattern Recognit. Lett. 141, pp. 61–67, 2021, doi: 10.1016/j.patrec.2020.07.042.

[10] J. D. Kelleher, Deep Learning. Massachusets: MIT Press, 2019.

[11] H. C. Deeth and M. J. Lewis, High Temperature Processing of Milk and Milk Products. Oxford: John Wiley & Sons, 2017.

[12] Z. Abdullah, F. S. Taip, S. M. Mustapa Kamal, and R. Z. Abdul Rahman, "Effect of sodium caseinate concentration and sonication amplitude on the stability and physical characteristics of homogenized coconut milk," J. Food Process. Preserv. 42, no. 11, pp. 1–9, 2018, doi: 10.1111/jfpp.13773.

[13] E. M. Yahia, Postharvest Biology and Technology of Tropical and Subtropical Fruits, vol. 2. 2011.

[14] S. Bianco, R. Cadene, L. Celona, and P. Napoletano, "Benchmark analysis of representative deep neural network architectures," IEEE Access. 6, pp. 64270–64277, 2018, doi: 10.1109/ACCESS.2018.2877890.

[15] J. Shin, Y. K. Chang, B. Heung, T. Nguyen-Quang, G. W. Price, and A. Al-Mallahi, "A deep learning approach for RGB image-based powdery mildew disease detection on strawberry leaves," Comput. Electron. Agric. 183, no. June 2020, p. 106042, 2021, doi: 10.1016/j.compag.2021.106042.