International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 14, No.1 (Jun-24)

http://dx.doi.org/10.12785/ijcds/1501129

Milk Spoilage Classification through Integration of RGB and Thermal Data Analysis

Puteri Nur Farzanah Faghira Kamarudin[1], Nik Mohd Zarifie Hashim[1*], Masrullizam Mat Ibrahim[1] and Mahmud Dwi Sulistiyo[2]

[1]*Fakulti Teknologi dan Kejuruteraan Elektronik dan Komputer, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia*
[2]*School of Computing, Telkom University, West Java, Indonesia*
[*]*Corresponding author*

Received 20 Mar. 2024, Revised 20 Jun. 2024, Accepted 22 Jun. 2024, Published 26 Jun. 2024

Abstract: In Malaysia, milk consumption is commonly associated with family households, specifically children. The nutrition in milk is fundamental for children's growth which is why the parents will ensure their children have adequate milk intake from an early age. Various kinds of milk are available on the market, but pasteurized milk and UHT milk are the most consumed. Without proper storage and packaging conditions, milk could spoil quickly; hence an early detection method is needed to detect milk staleness and spoilage. Much research and study has been done regarding the classification of milk spoilage. However, factors such as the unreliability of data and time-consuming methods prove that a better working model with high accuracy needs to be developed. Efficient detection methods are crucial for ensuring milk quality. This proposed paper is targeted to develop and introduce image based analysis to detect the spoiled milk in various packaging and storage conditions using Deep Learning and Python programming language to cater for the problem stated above. A dataset containing both RGB and thermal images of milk was self-obtained. Both datasets undergo image processing method to enlarge the datasets. They were then divided into training dataset, validation dataset and testing dataset. The proposed model developed in this proposed paper is based on Convolutional Neural Network (CNN) which were modified to produce a high percentage accuracy result. The proposed model in this paper has achieved the accuracy of 99% for classification of RGB images of milk and 98% for the thermal images of milk.

Keywords: classification, Convolutional Neural Network (CNN), deep learning, milk spoilage, RGB image, thermal image.

1. Introduction

The study of food in deep learning has been available for a long time such as studying fruits and vegetables. Nowadays, despite these earlier two foods, other types of food and even ingredients are being involved in deep learning such as milk. In Malaysia, milk consumption is commonly associated with family households, specifically children. The nutrition in milk is fundamental for children's growth so the parents will ensure their children have adequate milk intake from an early age, making milk a staple food. For the first six months of life in a child, they were advised to exclusively being fed breastmilk.

After the age of six months, non-human milk commonly cow milk is consumed to aid in the child's growth. The nutrition in milk such as calcium, proteins, high level of energy, micronutrients and macronutrients and insulin like growth factor-1 (IGF-1) are beneficial in a child's development. In 1928, a study conducted by Boyd Orr had shown that Scottish children ranging from the age of 5 to 14 years old that had consumed milk other than their normal diet had a 20% increase in their height and weight [1].

Fig. 1, Fig. 2 and Fig. 3 show the production, imports and exports of milk in Malaysia respectively from the year 2020 until 2023 [2]. In Malaysia, the production is rather slow compared to the consumption [3]. In average from the year 2020 until 2021, milk production was found to be at 48 thousand metric tons and the count slightly decreased to 48 thousand metric tons in 2022 and it stays with the same amount in the year 2023 [2].

Meanwhile, milk import was found to be at 2,379 thousand metric tons in the average of the year 2020 until 2021. In 2022, the number increases to 2,423 thousand metric tons and slightly decrease to 2,159 thousand metric tons in 2023 [2]. As for exports of milk, 526 thousand metric tons were exported in the average year of 2020

E-mail address: m122310065@student.utem.edu.my, nikzarifie@utem.edu.my, masrullizam@utem.edu.my, mahmuddwis@telkomuniversity.ac.id

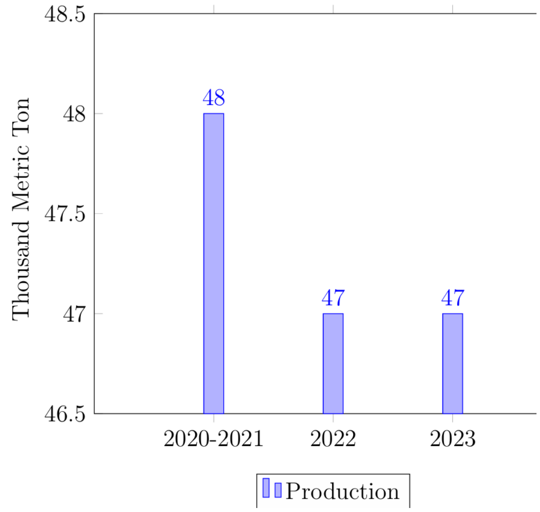

Figure 1. Production of milk in Malaysia from the year 2020 until 2023 [2].

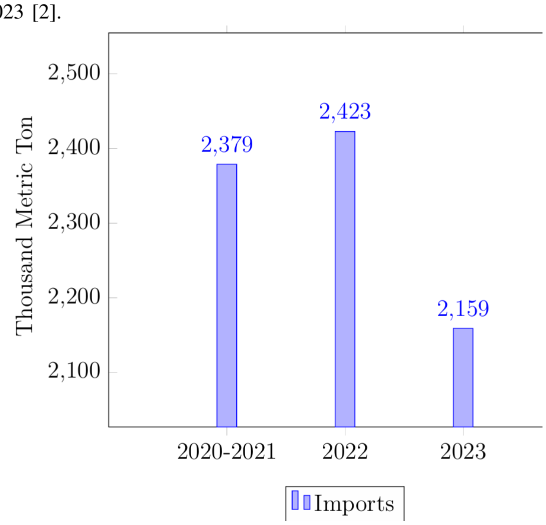

Figure 2. Imports of milk in Malaysia from the year 2020 until 2023 [2].

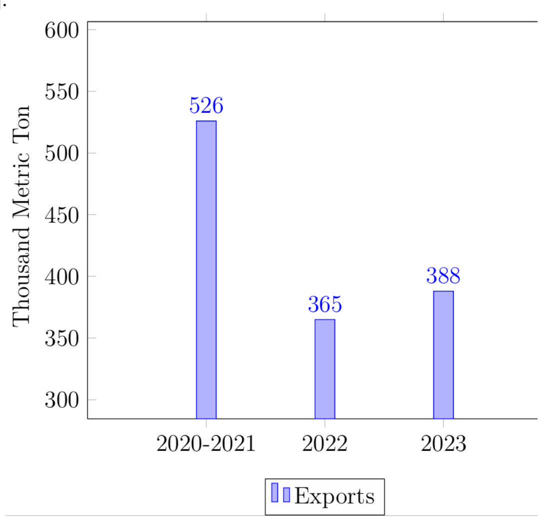

Figure 3. Exports of milk in Malaysia from the year 2020 until 2023 [2].

until 2021. The amount decreased to 365 thousand metric tons in 2022 and increases in 2023 at around 388 thousand metric tons [2]. Now, the expanding demand for milk and other milk-related products in Malaysia has changed due to changes in customer tastes, rising incomes, and increased knowledge of the products' nutritional advantages [3].

In recent years, numerous methods have been used to detect spoiled milk which are considered viable for research and studies. One of the methods discussed is by using chemical sensors or biosensors. The growth of microorganisms in milk will cause modification to happen to the medium's composition through their metabolic products, as well as the pH, conductivity, ionic content, color, odor, viscosity, and other milk's physicochemical characteristics. Hence, chemical sensors and biosensors come into place as a great deal of research has been done on chemical and biosensors as viable substitutes for the current standard techniques for detecting milk quality [4]. As mentioned by International Union of Pure and Applied Chemistry (IUPAC), the chemical sensors are often made up of two fundamental parts that are connected in series, which are a physico-chemical transducer and a chemical (molecular) recognition mechanism, or receptor [4].

Meanwhile, biosensors are the same as chemical sensors except a biochemical mechanism is used as the recognition system. The biological recognition system, often called a bioreceptor, uses data from the biochemical domain, frequently an analyte concentration, to produce a physical or chemical output signal with a defined sensitivity. Chemical sensors and biosensors can be categorized based on the types of transducers employed, the biochemical recognition element used for detection, or the analytes to be detected. The beneficial traits of these sensors such as fast response time, high sensitivity, small in size and with little to no sample preparation make them suitable for the use of detecting milk spoilage [4].

Besides chemical sensors and biosensors, pH value is a famous parameter for detecting milk spoilage. The growth of bacteria differs from one species to another. One type of bacteria may thrive in some conditions while others may weaken. At approximately 6.7 in pH level, milk is considered to be unspoiled or still in good condition. Some bacteria may also thrive at this pH level. However, as the pH level goes lower to a range of 4.0 to5.0, lactic acid may be produced due to its bacteria growing [5]. With this information, any appropriate pH meter can be used to determine the condition of milk.

Next, is the use of methylene blue reduction test which is an electrical method that can be used to detect bacteria. In this test, an appropriate amount of methylene blue solution was added to a milk sample and the sample was then placed in water bath at certain temperatures. Time taken for the methylene blue color to reduce is recorded and whichever milk samples take a longer time for reduction means the milk is of good quality. In other words, the slower the time taken for reduction, the better the milk condition is [6]. Although methylene blue reduction test can be used to determine the quality of milk, it is however very time consuming and repetitive procedures [5].

Other than that, radiographic imaging technique were also used in detecting milk spoilage. By using an X-ray technology, milk spoilage can be detected by observing any abnormalities in the gray levels. There would be a noticeable drop in X-ray intensity at various density regions as the radiation traveled through the milk solution. At the matching locations on the image, the resulting gray level would be different. Gray levels for any unspoiled milk would remain uniform while the gray levels for an unspoiled milk would show an abnormality [7].

Finally, image processing technique were used to detect spoiled milk. Image processing is digitally converting an image and carrying out the necessary operations to obtain helpful information. Combining image processing with artificial intelligence or deep learning can be one method used to detect spoiled milk [8]. A dataset of images is created by taking an appropriate number of images of milk bottles that contain both good and spoiled milk. A model is created, and the dataset of images will then be used to train and test the model. The model will then predict the condition of milk to be either good or spoiled, based on the images given [9].

The emerging use of food and ingredients in research nowadays calls for the study of milk. Depending on the storage conditions, milk can stay fresh for a few days or months. An enormous amount of milk is wasted annually as a result of spoiling either before the milk reaches the shop or after it reaches the consumer [10]. Even after the milk reaches the consumer, the given expiry date provided by milk manufacturers or producers only indicates the milk's peak quality and not the final day it may be safely consumed, resulting in a considerable amount of milk waste [11].

Given that in Malaysia, there are many techniques that can be applied in the study of classification of milk spoilage. One of the possible techniques that can be applied as discussed above is the use of image data analysis. Both RGB image data and thermal image data can be used in this study.

The objectives of this study are to prepare dataset containing RGB images of milk and thermal images of milk, to develop a classification model for detecting spoiled milk in various storage conditions and to analyze and detect spoiled milk via multiple models to compare accuracy with proposed model.

Correlating from the objectives, our paper proposed a classification model to detect spoiled milk in various storage condition by using RGB images of milk and thermal images of milk. Our paper is the first to investigate the use of thermal imaging in milk spoiling detection, and it offers a novel approach to the field. We designed a quick and non-invasive classification system by taking RGB and thermal images of milk and using powerful image processing and deep learning to analyse the features of milk. This innovative method greatly improves food safety and quality control, and it may be integrated into supply chains and dairy production to lower waste and increase quality assurance. Our research lays the foundation for the use of thermal imaging in the future to identify deterioration in a variety of perishable foods.

2. Literature Review

This literature review focuses on the classification of milk spoilage utilizing deep learning methods and the use of RGB image and thermal image as the datasets. Researchers have explored various deep learning architectures to accurately identify and classify spoilage patterns in milk samples. Studies have focused on pre-processing techniques to enhance the quality of input data, including image enhancement and feature extraction algorithms tailored to milk spoilage characteristics.

Furthermore, researchers have examined the transferability of deep learning models trained on one dataset to unseen datasets.

The pH values of milk in both good and spoiled condition are also discussed to ensure that the images collected in this proposed paper are labeled correctly according to their condition. Overall, the literature underscores how deep learning techniques can transform the way that milk deterioration is detected and classified, creating a path for more dependable and effective quality control measures in the dairy industry.

A. pH Values of Good and Spoiled Milk

In this proposed paper, the RGB images and thermal images captured must be labeled correctly whether there are in good condition or spoiled. One of the most reliable method to determine the condition of milk is to use their pH value. However, the threshold of the pH value of milk varies from one researchers' group to another. The threshold value used in this proposed paper is set based on the literature review done on several research groups.

Based on a study done by Max Weston et al. (2020), pH value's reaction towards bacterial growth is least dependent on the type of bacteria. Hence, making it one of the best approaches to monitor milk spoilage [12]. It was also mentioned by Yanlan Maet al. (2020) that when milk deteriorates, the quantity of bacteria in the milk increases which results in decreasing pH values [13]. As instance, a study done by Ruichang Gao et al. (2022) states that the pH value of milk decreases when milk is spoiling because lactic acid bacteria break down lactose which in turn produces lactic acid [14]. There is still no definitive value of pH where the milk is not safe to consume which in turn limits the designing of a freshness detection method using pH value [12].

Table I is constructed to summarize the pH value threshold determined by these researches.

B. Classification and Detection of Milk Spoilage

Based on research done by Daniel Rodriguez et al. (2020), WPT/NFC coil is used to record data from fresh milk and spoiled milk. In this research, it is believed that when raw milk expires, its electrical properties change. As the beverage bottle moves towards the coil, it generates induced eddy currents and electromotive force within it, causing a shift in the coil's impedance. These alterations occur from differences in the resistivities and dielectric constants between the surrounding air and the beverage bottle. Hence, the detection method can be done based on the dielectric constant correlation [15]. However, since the data collected is not directly from the milk itself but from the milk carton, the accuracy of the testing may not represent the real condition or freshness of the milk. Other than that, more research is required to determine the impacts of various milk storage conditions and packaging types in order to develop a more comprehensive classification system.

A study on Miniaturized Milk Adulteration Detection System by Suryasnata Tripathy et al. (2018) is conducted to develop a compact and cost-effective platform designed for monitoring the natural physical characteristics of milk. to detect milk adulteration. Other than adulteration, microbial spoilage is also proven to increase the acidity of milk which will give a lower reading of pH value. In this study, pH sensor strip is used to identify milk spoilage. Every milk sample underwent individual testing utilizing a minimum of 15 sensor strips. Precisely, each strip was immersed into the milk sample, allowed to adequately dry, and then captured using a smartphone camera. Subsequently, the areas on the sensors showing color changes were cropped from the images and associated with respective pH values. The samples are then labelled with values in the range 6.6–6.9 as pure (class label 1), less than 6.6 as acidic (class label 0), and more than 6.9 as basic (class label 2) [16]. However, collecting data of pH value using image of each sensor strip is time consuming considering each sensor strip needs to be captured. Using pH value instead of pH strip colour is more time efficient.

One of newest method in detecting spoiled milk has been done by Wu et al. (2024) where VibMilk application is introduced. VibMilk is a vibration-based milk spoilage detection method where commercially available smartphones were used to utilize their Inertial Measurement Unit (IMU) and ubiquitous vibration motor. Distinct vibration responses at different stages of deterioration will cause changes in the physical characteristics of the milk which can be use to detect spoiled milk. To use the VibMilk app, user needs to open the VibMilk app in their phone and attach their phone to the front of the milk carton. To keep the phone from slipping as it vibrates, regular transparent tape is used to firmly adhere the top and bottom borders of the device to the milk carton. By clicking a start button from the app, vibration motor will start. As the vibration occurs, the IMU in the user's phone will record the acceleration of the carton in response to the vibration. Using a trained deep learning model downloaded by the app from the phone's manufacturer, IMU data is analysed and the spoilage level of milk is predicted. An appropriate pH value appears on the screen in app, giving Alice a heads-up regarding the freshness of the milk. The VibMilk app is implemented and demonstrated on four different smartphone models which are Google Pixel 5, Google Pixel 6, Samsung Galaxy S21 FE and Samsung Galaxy S22. These models were chosen because they come equipped with the two most popular vibration motor types which are x-axis and z-axis. Google Pixel 6 and Samsung Galaxy S22 were used for the x-axis motor while Google Pixel 5 and Samsung Galaxy S21 FE were used for the z-axis motor. These vibration motors which range from 140 Hz to 160 Hz are running at their fixed frequencies. The IMU's sample frequency is 400 Hz in order to capture the Nyquist rates of various vibration frequencies. In this study, there are a total of 23 pH values class being tested and the highest accuracy is achieved by Google Pixel 5 with an accuracy percentage of 100%. This proves that all pH values tested is correctly classified by VibMilk with Google Pixel 5. The second highest, scoring an accuracy of 99.13% are the Google Pixel 6 and Samsung Galaxy s21 FE. Finally, scoring last place at an accuracy of 95.13% is Samsung Galaxy S22 [11].

C. Classification Approach with RGB Images

A study on Fruit Disease Classification and Identification using Image Processing is conducted by Ayyub et al. (2019) to identify apple fruit disease. The dataset consists of a total of 280 apple fruit images of different conditions. The images vary such as rot, scab, blotch, and normal apple where each class contains 70 apple fruit images. Image segmentation, feature extraction and feature combination are conducted before classification. The classifier chosen for the study is Multi- Class Support Vector Machine (MSVM). Features that were used to identify and classify the diseased and normal apple fruit were color coherence vector (CCV), zernike moments (ZM), improved sum and difference histogram (ISADH), completed local binary pattern (CLBP) and gray level co-occurrence matrix (GLCM). After accuracy percentage is calculated, 96.07% average classification accuracy was achieved by using ISADH+GLCM and 96.29% average classification accuracy achieved by using ISADH+CLBP+ZM features combination. Considering that RGB images are the only type of data collected, using MSVM only may not be satisfying. Multiple models can be used to have various accuracy results. Hence, the best model can be chosen based on the accuracy percentage [17].

Kaya and Gürsoy (2023) have developed a novel deep learning-based architecture that feeds two visual input which are RGB images and segmented RGB images.

TABLE I. TABLE SUMMARY FOR PH VALUE THRESHOLD OF MILK.

Author	pH value of milk		
	Fresh Milk	Spoiling	Spoiled
Yanlan Maet al. (2020) [13]	6.6 – 6.8	5.5 – 6.0	4.5 – 5.5
Max Weston et al. (2020) [12]	±6.8	-	less than 4.5
Ruichang Gao et al. (2022) [14]	6.6 – 6.8	-	4.0 – 5.0

PlantVillage dataset was obtained which includes 54183 images under 38 different classes. Since the image resolution of PlantVillage datasets comes in two different pixels which are 256 ×256 pixels and 514 ×514 pixels, preprocessing technique needs to be applied to both datasets. All images are reduced to 224 ×224 pixels to match the input shape of tested models. As for segmented RGB images, the background in all images are removed. Some most known pre-trained models are tested which are VGG16, ResNet50, InceptionV3, Xception, MobileNetV2 and DenseNet121. An 80%-20% hold-out validation method is used for both RGB images and segmented RGB images and among all the stated pre-trained models, the highest percentage of accuracy is achieved by DenseNet121 with 96.10% for RGB images and 95.30% for segmented RGB images. Hence, the design of the developed proposed model is based on DenseNet and their study is to classify plant diseases into 38 different classes. Important hyperparameters used are set such as learning rate (0.001), batch size (16), optimizer (adam optimizer) and number of epochs (100). A fivefold cross-validation technique is used to assess the developed proposed model and there are three models developed which are RGB model, segmented RGB model and RGB and segmented model. The RGB and segmented model are fusion of both RGB image and segmented RGB image and is the proposed model. Image fusion integrates complementary information from multiple sources of the same scene into a single image. The resulting fused image contains more detailed data than any single source image, as it incorporates diverse types of information from various image sources. The study managed to achieve high accuracy percentage for all 5 folds across all developed model except for fold 4. The average accuracy except the 4th fold for RGB model is 97.49%, for segmented RGB model is 95.89% and for the proposed model, RGB and segmented model is 98.91%. At the end, the developed proposed model managed to obtained an accuracy of 98.17% including 4th fold and 98.90% excluding the 4th fold [18].

D. Classification Approach with Thermal Images

A study on Deep Learning-Based Plant Classification and Crop Disease Classification by Thermal Camera by Batchuluun, G et al. (2022) is conducted, and the proposed method used is Convolutional Neural Network (CNN) where plant diseases and crop diseases is classified by using thermal images as data, accompanied by Explainable Artificial Intelligence (XAI), the proposed method is called Plant Deep Explainable Artificial Intelligence (PlantDXAI). Two datasets were used in this study which were their thermal plant image dataset which was self-obtained and an open database of crop diseases which were the paddy crop dataset. To enlarge the paddy crop open database, they applied image augmentation to the 447 training images [19].

The augmentation was done by flipping the images horizontally and rotating the images by 90° three times. In total, they managed to obtain 3,576 images for the open database. As for their self-obtained thermal plant image dataset, they used a Tau® 2 FLIR thermal camera and captured various images of roses and rose leaf. To enlarge their self-obtained thermal plant image dataset, they applied image augmentation to 3,314 training images. The augmentation was done by flipping the images horizontally and rotating the images by 90° three times. In total, they managed to obtain 26,512 images for the selfobtained thermal plant image dataset which has 28 classes in total. For plant image classification using self-obtained thermal plant image dataset, CNN-16 obtained the average accuracy of 98.55%. As for crop disease image classificationusing paddy crop open database, CNN-16 obtained the accuracy of 88.63% without Class Activation Map (CAM) and 90.04% with CAM [19].

Based on a study done by Mammoottil et al. (2022), thermal images of breast were used to detect breast cancer. The study uniquely uses five different views of the breast to capture the thermal images. The dataset of thermal images consists of images that are captured from the front, 45° to the left, 90° to the left, 45° to the right and 90° to the right and is from the Database for Mastology Research (DMR). The initial images obtained had the size of 640×480 pixels and are enhanced to 640×640 pixels following data preprocessing to enable the data as an input for Convolutional Neural Network (CNN) since CNN are trained to received square images. Any blurry image or image with the absence of all 5 views and any abnormal data found during preprocessing are immediately deleted. To test of any changes in the accuracy, the patient's clinical data with information such as age, symptoms and conditions is obtained to be used along thermal image. There are two types of model being built in this study which is Model 1 and Model 2. Model 1 will be use to train thermal images of frontal view, 90° to the left and 90° to the right. The channel size of CNN Model 1 is 32, 64 and 128. For Model 2, the same view of thermal images is used but the front view uses a different CNN model with channel size of 50, 100, 150, and 200. Meanwhile, the right and left view uses the same channel size of 45, 90, 135, and 180. After the three views in respective Model 1 and Modal 2 are trained, and after the clinical data is trained on a different model, the outputs from Model 1 and outputs from clinical data will be combined using another model and same goes for Model 2. Model 1 managed to achieve 85.4% without clinical data and 93.8% with clinical data. For Model 2, it obtained an accuracy of 81.2% without clinical data and 89.6% with clinical data [20].

Table II shows the comparison method used by the studies above.

E. Convolutional Neural Network (CNN) as classification model

At present, the use of deep learning has been massively applied in research [21]. Two significant field that uses deep learning are the biological and ecological field [22]. Deep learning is a big help to perform traditional applications and has managed to surpass machine learning techniques in various fields. Lately, deep learning performance has surpassed human capabilities in jobs like image classification. This technique has affected almost every scientific field. The application of deep learning has already disrupted and revolutionized the majority of sectors and businesses. The world's top technology and economy-focused businesses are competing to advance deep learning [21]. The healthcare industry is one area where deep learning is being applied. Creating prediction models for illness diagnosis, prognosis, and therapy suggestions is one task where deep learning is applied. It is also used in medical imaging methods such as image reconstruction from CT and MRI scans.

One of the pros provided by deep learning that attracts researchers in using deep learning is its ability to learn large values of data. It utilizes the large data provided to learn and then map the given data to corresponding labels [21]. One of the most used deep learning techniques specifically in supervised techniques and among the best for image processing techniques [23], is Convolutional Neural Network (CNN). Convolutional Neural Network, or CNN, have flourished and achieved notable advancements in the field of computer vision for the past few years up until now. An increasing number of studies have utilized CNN in the domain of biological image processing, yielding positive outcomes [22]. Similar to a traditional neural network, the structure of CNN was modeled by neurons found in the brains of humans and other animals. More precisely, the CNN simulates the intricate cell sequence that makes up the visual cortex in a cat's brain [21].

CNN in particular are a unique kind of multilayer neural network that were first introduced by Yann LeCun in 1998 and have a variety of real-world uses. The initial architecture of the first CNN model, known as LeNet-5, is shown in Fig. 4. After the AlexNet model triumphed in the 2012 ImageNet (ILSVRC) competition, CNN became extremely popular [24].

The most important component in CNN is convolutional

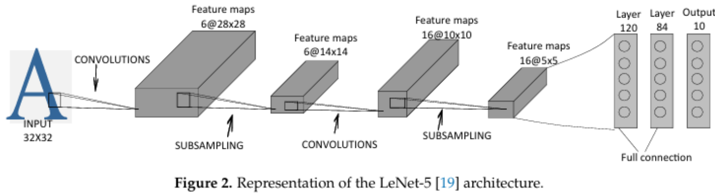

Figure 4. The architecture of LeNet-5 [24].

layer. It is comprised of a set of convolutional filters, also referred to as kernels. These filters are used to convolve the input image, which is expressed as N-dimensional metrics, to produce the feature map that is the output [21]. The 2D matrix representing the image (I) is convolved with the smaller 2D kernel matrix (K) in the convolution operation, which is frequently used in digital image processing. The resulting mathematical formulation (1) with zero padding is as below [24].

$$S_{i,j} = (1 * K)_{i,j} = \sum_m \sum_n I_{i,j} .K_{i-m,j-n} \qquad (1)$$

The next component is pooling layer, where it creates a smaller feature maps by reducing the size of the initial feature maps. Simultaneously, it preserves most of the prominent data (or attributes) at each stage of the pooling process [21]. In other words, it decreases the spatial dimension of the convolutional outputs, which lowers the number of network parameters. The two most used pooling layers are max pooling and average pooling. As for max pooling, it determines the highest value for every input patch. By swiping the filter over the feature map, the max-pooling layer maintains the highest value of every patch. The equation for max pooling is represented as follows (2) [24]:

$$f_{max}(A) = max_{n\times m}(A_{n\times m}) \qquad (2)$$

For average pooling, it calculates the average value for each patch of the input. By dividing the input into pooling regions and calculating the average values of those regions, the average pooling layer downsamples the convolutional activation. The mathematical representation is as follows (3) [24]:

$$f_{max}(A) = \frac{1}{n+m} \sum_{i=1}^{n} \sum_{k=1}^{m} (A_{i,k}) \qquad (3)$$

Dropout layer is as a form of regularization layer that removes network neuron units at random to stop the units from overly co-adapting. This technique enhances the network's performance while enabling the overfitting issue to be addressed. It is applicable to all network layers [24].

3. Methodology

Fig. 5 shows the overall flowchart of the methodology that is carried out in this proposed paper. Firstly, sample preparation is done to create two datasets for this study. RGB images of milk and thermal images of milk will be used as data for the model. The datasets are RGB image

TABLE II. TABLE OF COMPARISON FOR RELEVANT STUDIES.

Author	Dataset	Proposed Method	Performance Measure Accuracy Percentage
Daniel Rodriguez et al. (2020) [15]	VNA dataset Low frequency dataset	20 machine learning classifiers under 6 categories (Decision Tree, Naive Bayes, K-Nearest Neighbor, Discriminant Analysis, Support Vector Machine, Ensemble Learning)	100% for all classifiers Cross validation: 99.2% (4 best classifiers), standard deviation: 2.8 (all classifiers)
Suryasnata Tripathy et al. (2018) [16]	Labeled images of pH sensor strip	Support Vector Machine (SVM)	99.71%
Wu et al. (2024) [11]	Inertial Measurement Unit (IMU)	VibMilk	Google Pixel 5: 100% Google Pixel 6: 99.13% Samsung Galaxy s21 FE: 99.13% Samsung Galaxy S22: 95.13%
Ayyub et al. (2019) [17]	Apple fruit images	Multi- Class Support Vector Machine (MSVM)	ISADH+GLCM: 96.07% ISADH+CLBP+ZM: 96.29%
Kaya et al. (2023) [18]	PlantVillage	DenseNet121 (RGB + segmented)	include 4th fold: 98.17% exclude 4th fold: 98.90%
Batchuluun, G et al. (2022) [19]	Thermal plant image Paddy crop dataset (open database)	CNN-16 CNN-16 with Class Activation Map	98.55% 90.04%
Mammoottil et al. (2022) [20]	Thermal breast image	Model 1 Model 2	with clinical data: 93.8% without clinical data: 85.4% with clinical data: 89.6% without clinical data: 81.2%

dataset which contains RGB images of milk and thermal image dataset which contains thermal images of milk.

To prepare dataset for this study, sample needs to be prepared for thermal image, RGB image and pH value collection. To ensure reliability and consistency of data, the same brand of milk was used throughout this study. The sample was prepared according to Table III.

TABLE III. PREPARATION OF SAMPLES.

Carton Condition	Storage Condition	Time Left (Hours)
Perfectly sealed carton	Refrigerated Room Temperature	24, 48, 72, 96
Holed carton	Refrigerated Room Temperature	24, 48, 72, 96
Opened carton	Refrigerated Room Temperature	24, 48, 72, 96

In this sample preparation, one milk carton was used per category. For example, one perfectly sealed carton is placed in the refrigerator for 24 hours. Another one perfectly sealed carton is placed in the refrigerator for 48 hours and so on. There was a total of 24 milk cartons used in this study. For perfectly sealed carton, the cartons were placed in their respective category as it is. While for holed carton, a hole was made at the foil protector on top of the milk carton where the straw was supposed to go in. Lastly, for opened carton, the top of the carton was cut open entirely to be left exposed.

When the sample had reached its time mark, thermal image and RGB image of the sample were captured. A thermal camera was used to capture thermal image while an Android phone camera, Samsung J6 model was used to capture RGB image. To maintain quality of data and to avoid overexposure of light from surrounding area, the image for all sample were captured in the same room and a lightbox was utilized. The sample was first poured into a glass cup, then the glass cup is placed inside the lightbox. The pH values of milk from each sample will be recorded every time RGB and thermal image are captured with the size of 3096x4128 pixel and 240x240 pixel respectively. The pH values will not be included in the dataset because the pH values only serve the purpose of determining whether the milk is still good or already spoiled.

Fig. 6 and Fig. 7 show the example of RGB images of milk and thermal images of milk captured.

To enlarge both datasets, image processing method will be applied to both RGB and thermal images captured. Both RGB and thermal images will then be grouped into two different classes which are ‘Spoiled’ and ‘Good’. A random

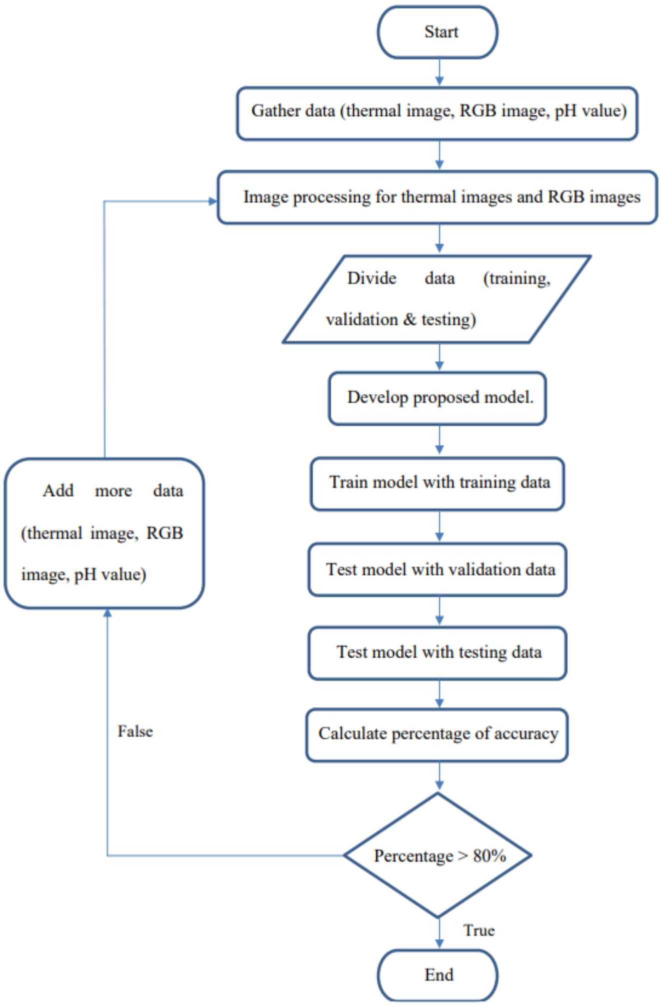

Figure 5. Flowchart of methodology.

set of images for both RGB and thermal image are collected to create a third class called 'Unclassified'. In total, each dataset contains three classes which are 'Spoiled', 'Good' and 'Unclassified'. The number of images in RGB and thermal image dataset is the same which were 963 images. The splitting percentage used were 70% for training dataset, 15% for validation dataset and 15% for testing dataset. Table IV and Table V show the number of images for each class in each dataset for both RGB image dataset and thermal image dataset and the number of images in each dataset respectively.

The proposed model used in this study is Convolutional Neural Network (CNN) and the layers in this model is shown in the model summary in Table VI.

Multiple layers were added to the model to create the proposed classification model. The first two layers added to the model were a 2D convolution layer which usually abbreviated as conv2D. Next, max pooling 2D layer or

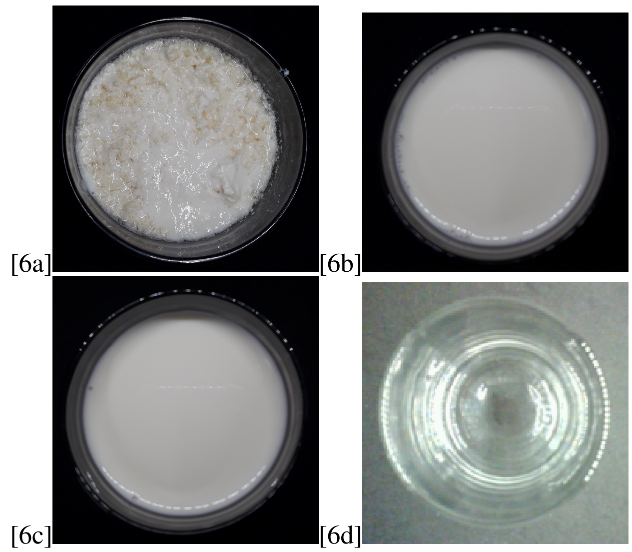

Figure 6. RGB images of milk: a) spoiled milk at room temperature, b) good milk at room temperature, c) milk in refrigerator, d) unclassified

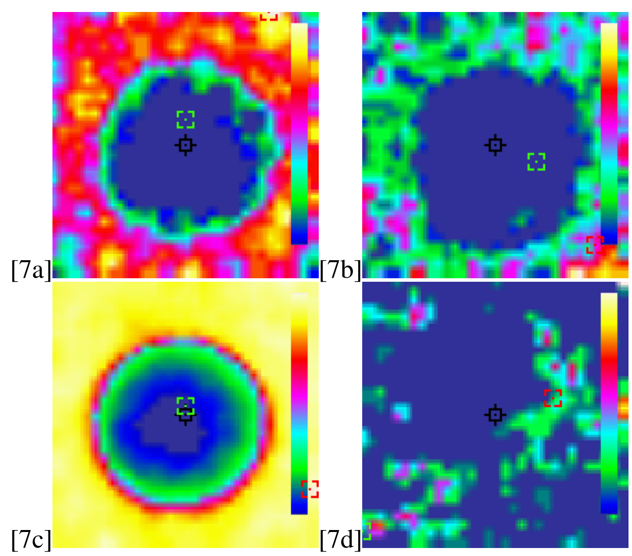

Figure 7. Thermal images of milk: a) spoiled milk at room temperature, b) good milk at room temperature, c) milk in refrigerator, d) unclassified

TABLE IV. THE NUMBER OF IMAGES FOR EACH CLASS IN EACH DATASET FOR BOTH RGB IMAGE DATASET AND THERMAL IMAGE DATASET.

Class	Training	Validation	Testing	TOTAL
Spoiled	225	48	48	321
Good	225	48	48	321
Unclassified	225	48	48	321
TOTAL	**675**	**144**	**144**	**963**

TABLE V. NUMBER OF IMAGES IN EACH DATASET.

Dataset	RGB Image	Thermal Image
Training (70%)	675	675
Validation (15%)	144	144
Testing (15%)	144	144
TOTAL (100%)	**963**	**963**

TABLE VI. MODEL SUMMARY.

Layer (type)	Output Shape
conv2d (Conv2D)	(None, 238, 53, 32)
conv2d_1 (Conv2D)	(None, 236, 51, 64)
max_pooling2d (MaxPooling2D)	(None, 118, 25, 64)
dropout (Dropout)	(None, 118, 25, 64)
conv2d_2 (Conv2D)	(None, 116, 23, 128)
max_pooling2d_1 (MaxPooling2D)	(None, 58, 11, 128)
conv2d_3 (Conv2D)	(None, 56, 9, 128)
max_pooling2d_2 (MaxPooling2D)	(None, 28, 4, 128)
dropout_1 (Dropout)	(None, 28, 4, 128)
flatten (Flatten)	(None, 14336)
dense (Dense)	(None, 1024)
dropout_2 (Dropout)	(None, 1024)
dense_1 (Dense)	(None, 3)

most known as MaxPooling2D was added where it will choose the highest value from every pool. Then, dropout layer was added where it will randomly ignore some nodes in the layer during training. Conv2D and MaxPooling2D were added two more times followed by a dropout layer. Afterwards, a flatten layer was introduced to convert the pooled feature map into a singular column, which was then passed to the fully linked layer. The fully connected layer was integrated into the neural network using a dense layer. Finally, an additional dropout layer and dense layer were appended to the model. The summary of the model was then printed out.

A few important hyper parameters were tuned to ensure the proposed model is able to obtained the highest percentage of accuracy for this proposed paper. The learning rate were set to be at 0.0001, batch size was 9, the number of epochs were 50 and the optimizer used in this model was Adam optimizer. The RGB image dataset specifically the training dataset dan validation dataset of RGB was first used with the proposed model for training and validation respectively. After training and validation was done and no error was found, testing was carried out with the RGB image testing dataset. To determine the performance measure of the proposed model, the percentage of accuracy was calculated. Accuracy percentage indicates how well the proposed model works in classifying. Higher accuracy proves the model is accurate and precise. Anything greater than 80% is a great model performance. An accuracy measure of anything between 70%-90% is ideal. This is also consistent with industry standards. If the accuracy percentage of the proposed model has achieved more than 80%, the classification of RGB image dataset is considered successful. However, if the accuracy percentage of the proposed model falls below 80%, more image data needs to be captured and more image augmentation needs to be done to enlarge the number of images in the dataset. The same was done to the thermal image dataset for training, validation and testing.

The time spent for training of proposed model was recorded to be 40 minutes on average while running on graphic processing unit (GPU) which is T4 GPU. GPU was chosen for this study since a study has proven that GPU has faster training time compared to central processing unit (CPU) [25].

4. Results, Analysis and Discussion

To evaluate the performance of the developed proposed model, have a look at its percentage of accuracy. Fig. 8 and Fig. 9 show the comparison of accuracy percentages between the developed proposed model and with VGG16, VGG19 and ResNet for RGB and thermal image dataset respectively. It can be seen that for both RGB and thermal image dataset, the developed proposed model is able to achieve the highest accuracy percentage compared to all the other models involved in this study.

Other than percentage of accuracy, precision, recall and F1-score which can be obtained from the classification report can also be used to evaluate the performance of the developed proposed model. Support in the classification result is the number of samples that belong to each three classes.

A. *Results for RGB Image Dataset*

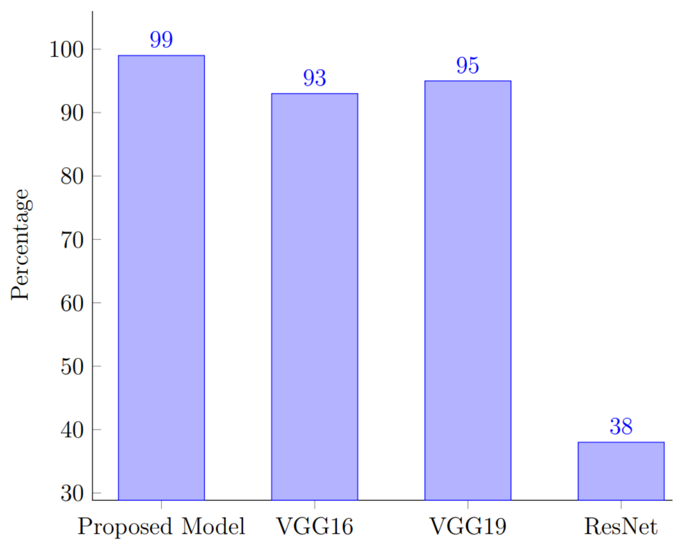

Figure 8. Accuracy percentages between the developed proposed model and with VGG16, VGG19 and ResNet for RGB.

The model proposed in this research has attained a 99% accuracy rate, surpassing the VGG16 model at 93% and the VGG19 model at 95%. In contrast, the ResNet model achieved an accuracy of 38%. The proposed model in this study outperformed the VGG16 and VGG19 models. This proves that, for the application at hand, the proposed model is precise and effective. Compared to the performance of the proposed model and VGG models, the performance of ResNet is notably lower. Table VII shows the classification report for the proposed model.

TABLE VII. CLASSIFICATION REPORT FOR PROPOSED MODEL (RGB).

Class	Precision	Recall	F1-score	Support
Class 1 (Spoiled)	0.96	1.00	0.98	48
Class 2 (Good)	1.00	0.96	0.98	48
Class 3 (Unclassified)	1.00	1.00	1.00	48

For proposed model, Class 2 and Class 3 both have the same precision and were higher than Class 1, which proves that the model correctly identified all positive outcomes for Class 2 and Class 3. While for recall, Class 1 and Class 3 have the highest recall, leaving Class 2 with 0.96. The majority of the dataset's positive occurrences for Class 1 and Class 3 are captured by the model. F1-score are the combination of precision and recall. As Class 3 has a perfect score of 1.0 for precision and recall, it automatically gained the F1-score of 1.0 while Class 1 and Class 2 gained F1-score of 0.98.

Table VIII, IX and X show the classification report for the VGG16, VGG19 and ResNet respectively.

TABLE VIII. CLASSIFICATION REPORT FOR VGG16 (RGB).

Class	Precision	Recall	F1-score	Support
Class 1 (Spoiled)	1.00	0.79	0.88	48
Class 2 (Good)	0.84	1.00	0.91	48
Class 3 (Unclassified)	0.98	1.00	0.99	48

TABLE IX. CLASSIFICATION REPORT FOR VGG19 (RGB).

Class	Precision	Recall	F1-score	Support
Class 1 (Spoiled)	1.00	0.85	0.92	48
Class 2 (Good)	0.87	1.00	0.93	48
Class 3 (Unclassified)	1.00	1.00	1.00	48

TABLE X. CLASSIFICATION REPORT FOR RESNET (RGB).

Class	Precision	Recall	F1-score	Support
Class 1 (Spoiled)	0.35	0.15	0.21	48
Class 2 (Good)	0.00	0.00	0.00	48
Class 3 (Unclassified)	0.39	1.00	0.56	48

B. Results for Thermal Image Dataset

The model suggested in this research has reached an accuracy of 98%, outperforming the VGG16 model at 83% and the VGG19 model at 86%. In comparison, the ResNet model achieved an accuracy of 51%. The proposed model appears to be very effective based on the notable performance difference between the proposed model and the VGG models. The proposed model's promise for practical applications is demonstrated by its ability to outperform the VGG models, which are wellknown and often used deep learning architectures. By comparison, the accuracy of the ResNet model was just 51%, a far lower result than the performance of the proposed model. Table XI shows the classification report for the proposed model.

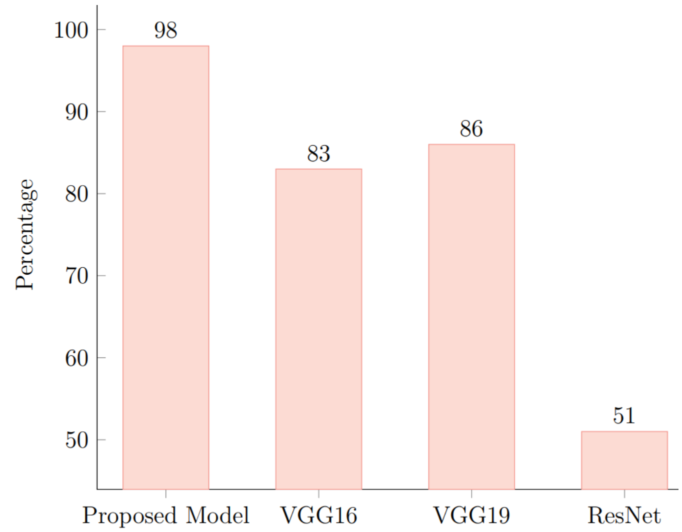

Figure 9. Accuracy percentages between the developed proposed model and with VGG16, VGG19 and ResNet for thermal.

TABLE XI. CLASSIFICATION REPORT FOR PROPOSED MODEL (THERMAL).

Class	Precision	Recall	F1-score	Support
Class 1 (Spoiled)	0.94	1.00	0.97	48
Class 2 (Good)	1.00	0.94	0.97	48
Class 3 (Unclassified)	1.00	1.00	1.00	48

For proposed model, Class 2 and Class 3 both have the same precision and were higher than Class 1, which proves that the model correctly identified all positive outcomes for Class 2 and Class 3. While for recall, Class 1 and Class 3 have the highest recall, leaving Class 2 with 0.94. The majority of the dataset's positive occurrences for Class 1 and Class 3 are captured by the model. F1-score are the combination of precision and recall. As Class 3 has perfect score of 1.0 for precision and recall, it automatically gained the F1-score of 1.0 while Class 1 and Class 2 gained F1-score of 0.97.

Table XII, XIII and XIV show the classification report for the VGG16, VGG19 and ResNet respectively.

C. Discussion

From the classification results of both RGB image dataset and thermal image dataset, it is proven that the simple and new change implemented to the existing CNN

TABLE XII. CLASSIFICATION REPORT FOR VGG16 (THERMAL).

Class	Precision	Recall	F1-score	Support
Class 1 (Spoiled)	1.00	0.50	0.67	48
Class 2 (Good)	0.69	1.00	0.81	48
Class 3 (Unclassified)	0.96	1.00	0.98	48

TABLE XIII. CLASSIFICATION REPORT FOR VGG19 (THERMAL).

Class	Precision	Recall	F1-score	Support
Class 1 (Spoiled)	0.71	1.00	0.83	48
Class 2 (Good)	1.00	0.79	0.88	48
Class 3 (Unclassified)	1.00	0.79	0.88	48

managed to successfully gain a high percentage of accuracy. This could bring our new proposed idea to a wider application in image processing and deep learning. Since we introduced the use of thermal imaging in milk spoilage detection, this technique can be explored further in the use of other food. In the future, the proposed model and the thermal imaging technique can be used to detect other food spoilage which could lower waste and increase quality assurance either in the supply chain or production. The complexity of some model architecture such as ResNet makes it more prone to overfitting and more difficult for the model to generalize well when trained on limited datasets, which might result in decreased accuracy [26]. Since ResNet model are tuned for larger images, low quality or smaller images makes it difficult for ResNet to extract significant features and patterns that are essential for accuracy [27]. Consequently, the performance of ResNet models is largely dependent on the size of dataset and the quality of the photos, emphasizing the significance of having enough data and high-quality images for the best outcomes.

5. Limitations

There were a few limitations observed during the run of this research which were room temperature or weather. Since the collection of data involves different number of days for the sample to be collected, the current weather could affect the room temperature which in turn could affect the spoiling rate of the milk sample. Next, lighting challenge during data collection which is capturing RGB image and thermal image. The surrounding area may emit different lighting at times which can cause shadows and colour discrepancies. Hence, a lightbox was used as additional equipment to control the emition of light. Other than that, the use of pH values along with images of milk can give a more credible classification as one of the methods used to detect milk spoilage are by measuring its pH value. For our future work, we can develop a multimodal deep learning model to combine both image and numerical data, in this case pH value, to obtained a more accurate classification. A larger dataset will also improve the classification model since a deep learning model will learn better with larger dataset.

TABLE XIV. CLASSIFICATION REPORT FOR RESNET (THERMAL).

Class	Precision	Recall	F1-score	Support
Class 1 (Spoiled)	0.00	0.00	0.00	48
Class 2 (Good)	1.00	0.58	0.74	48
Class 3 (Unclassified)	0.40	0.96	0.57	48

6. Conclusion

This study proposed a classification model in detecting the freshness of milk in various storage conditions. Two datasets which are RGB image dataset which contains RGB images of milk and thermal image dataset which contains thermal images of milk were selfobtained throughout this study. The highest percentage of accuracy has been successfully achieved by the proposed model in comparison to other chosen models which are VGG16, VGG19 and ResNet. For RGB image classification, the proposed model has an accuracy percentage of 99% while for thermal image classification, the proposed model has an accuracy percentage of 98%. The results above prove that by only a few new changes implemented to the existing CNN model, a high percentage of accuracy can be obtained. In conclusion, all the objectives in this study are successfully achieved which are to prepare dataset containing RGB images of milk and thermal images of milk, to develop a classification model for detecting spoiled milk in various storage conditions and to analyze and detect spoiled milk via multiple models to compare accuracy with proposed model. There are many aspects in this study that can be improved such as the quality of images captured, size of dataset and variables used in classification. To obtain a better working classification model, images may be captured in various lighting conditions, not in controlled environments. Other than that, more images can be captured and apply image augmentation to the captured images to enlarge the dataset. Larger dataset will improve the classification model because the model will have large amount of data to learn from. Last but not least, milk does not only come in full cream type only, but it also comes in different flavors and varieties such as strawberry, chocolate, and low fat. This means more milk flavors and varieties can be added as variables for future work.

7. Acknowledgement

Heartfelt thanks to Bahagian Keselamatan dan Kualiti Makanan, Jabatan Kesihatan Negeri Perak for their valuable insights and cooperation via this research collaboration, and Fakulti Teknologi dan Kejuruteraan Elektronik dan Komputer, Universiti Teknikal Malaysia Melaka for providing the necessary resources and conducive environment for academic growth and learning.

References

[1] S. V. S. S. V. Christine Herber, Lisa Bogler, "Association between milk consumption and child growth for children aged 6–59 months," *Scientific Reports*, 2020.

[2] FAO, *Food Outlook – Biannual report on global food markets*. FAO, 2023.

[3] C. Suntharalingam, *Food Value Chain in ASEAN: Case Studies Focusing on Local Producers*. ERIA, 2019, ch. Chapter 5.

[4] M. J. S. Arshak Poghossian, Hanno Geissler, "Rapid methods and sensors for milk quality monitoring and spoilage detection," *Biosensors and Bioelectronics*, 2019.

[5] J. W. R. L. H. K. T. B. T. P. T. J. A. C. J. C. E. O. N. S. N. S. W. Michael Lu, Yvonne Shiau, "Milk spoilage: Methods and practices of detecting milk quality," *Food and Nutrition Sciences*, 2013.

[6] N. S. Pooja Sharmaa, Mona Singh Narukab, "Microbiological analysis of raw and pasteurized milk by qualitative methods," *International Journal of Creative Research Thoughts*, 2023.

[7] H. B. N. T. T. K. T. T. N. D. T. Thuy, H. B. Tien, "Detection of spoilage in canned pasteurized milk using the radiographic imaging technique," *Atom Indonesia*, 2022.

[8] A. J. M. M. M. S. A. S. M. K. A. K. N. A. I. M. N. Nik Mohd Zarifie Hashim, Salizawati Mohd Yusoff, "A review: Milk spoilage and staleness detection approaches, technique, and technology trends," *International Journal of Scientific and Research Publications*, 2022.

[9] M. Kumari, "Ai-based milk deterioration detection," 2021.

[10] A. K. S. K. P. P. B. R. Shrikant Kapse, Priya Kedia, "A non-invasive method for detection of freshness of packaged milk," *Journal of Food Engineering*, 2023.

[11] Y. W. D. M. Yuezhong Wu, Wei Song, "Vibmilk: Non-intrusive milk spoilage detection via smartphone vibration," *IEEE Internet of Things Journal*, 2024.

[12] J. A. R. C. Max Weston, Minh Anh Thu Phan, "Anthocyanin-based sensors derived from food waste as an active use-by date indicator for milk," *Food Chemistry*, 2020.

[13] T. J. W. W. D. E. S. S. A. W. Q. J. D. Y. L. Yanlan Ma, Siying Li, "Development and optimization of dynamic gelatin/chitosan nanoparticles incorporated with blueberry anthocyanins for milk freshness monitoring," *Carbohydrate Polymers*, 2020.

[14] T. S. Y. B. Q. S. L. W. Y. R. W. J. L. Y. Ruichang Gao, Huiling Hu, "Incorporation of gelatin and fe2+ increases the ph-sensitivity of zein-anthocyanin complex films used for milk spoilage detection," *Current Research in Food Science*, 2022.

[15] C. L. Daniel Rodriguez, Mohammad A. Saed, "A wpt/nfc-based sensing approach for beverage freshness detection using supervised machine learning," *IEEE Sensors Journal*, 2021.

[16] S. R. K. V. S. J. S. G. S. Suryasnata Tripathy, Manne Shanmukh Reddy, "A step towards miniaturized milk adulteration detection system: Smartphone-based accurate ph sensing using electrospun halochromic nanofibers," *Food Analytical Methods*, 2018.

[17] "Fruit disease classification and identification using image processing," 2019.

[18] E. G. Yasin Kaya, "A novel multi-head cnn design to identify plant diseases using the fusion of rgb images," *Ecological Informatics*, 2023.

[19] K. R. P. Ganbayar Batchuluun, Se Hyun Nam, "Deep learning-based plant classification and crop disease classification by thermal camera," *Journal of King Saud University - Computer and Information Sciences*, 2022.

[20] A. S. C. P. M. K. H. M. M. Mathew Jose Mammoottil, Lloyd J. Kulangara, "Detection of breast cancer from five-view thermal images using convolutional neural networks," *Journal of Healthcare Engineering*, 2022.

[21] A. J. H. A. A.-d. Y. D. O. A.-S. J. I. S. M. F. M. A.-A. L. F. Laith Alzubaidi, Jinglan Zhang, "Review of deep learning: concepts, cnn architectures, challenges, applications, future directions," *Journal of Big Data*, 2021.

[22] X. X. Y. T. G. H. Jiaohua Qin, Wenyan Pan, "A biological image classification method based on improved cnn," *Ecological Informatics*, 2020.

[23] C. X. L. L. J. Z. R. L. F. W.-M. Z. W. L. Dengshan Li, Rujing Wang, "A recognition method for rice plant diseases and pests video detection based on deep convolutional neural network," *Sensors*, 2020.

[24] R. H.-G. R. J. B. C. F. A. V. José Naranjo-Torres, Marco Mora, "A review of convolutional neural network applied to fruit image processing," 2020.

[25] D. Gyawali, "Comparative analysis of cpu and gpu profiling for deep learning models," 2023.

[26] D. Z. Wanni Xu, You-Lei Fu, "Resnet and its application to medical image processing: Research progress and challenges," *Computer Methods and Programs in Biomedicine*, 2023.

[27] B. C. Michał Koziarski, "Impact of low resolution on image recognition with deep neural networks: An experimental study," *International Journal of Applied Mathematics and Computer Science*, 2019.

Puteri Nur Farzanah Faghira Kamarudin a research scholar who previously obtained her degree, Bsc (Hons) in Computer Engineering in November 2023 at Universiti Teknikal Malaysia Melaka. She is currently pursuing her master's degree in electronic engineering at the same university. Her research interests are in Artificial Intelligence, Deep Learning and Image Processing.

Nik Mohd Zarifie Hashim received theBEng and MEng degrees in Electronics Electrical Engineering and from Univeristy of Fukui, Japan in 2006 and 2008. He received the Ph.D in Computer Science from Nagoya University in 2022. In 2007, he joined Universiti Teknikal Malaysia Melaka as a Tutor. In 2008, he became Lecturer in the same university. Since 2013, he has been a Senior Lecturer there. His research focuses on computer vision, object identification, robot vision, and signal processing. He received the best paper award from ICTEC2022, Gold Research Award from ITEX2023, and Bronze Research Award feom MTE2023. He is member of IEEE, ASEAN Chartered Professional Engineer (ACPE), Professional Engineer of Board of Engineering Malaysia (BEM), Professional Technologists of Malaysia Board of Technologists (MBOT), and Member of The Institution of Engineers, Malaysia (IEM).

Masrullizam Mat Ibrahim was born in Perak, Malaysia. He received the B.Eng.(Hons.) and M.Eng in Electrical and Electronics Engineering from Universiti Teknologi Malaysia, Johor, Malaysia, in 2001 and 2006 respectively. Then, he received PhD in 2014 from University of Strathclyde, United Kingdom. He has work experience in research and development for telecommunication services company (2002-2003). Currently, he has been working as academic staff in Universiti Teknikal Malaysia Melaka since 2003. He currently holds the Faculty Dean position since 2020. His research interests are Artificial Intelligence, Signal Processing, Image and Video processing, Internet of Things (IoT), and embeddded system application.

Mahmud Dwi Sulistiyo received his Bachelor and Master's degrees in Informatics Engineering of Telkom Institute of Technology, Indonesia, in 2010 and 2012, respectively. In 2017 he was awarded the Indonesia's LPDP Scholarship and started a doctoral program in Department of Intelligent Systems at Graduate School of Informatics, Nagoya University, Japan, with a research focusing on Attribute-aware Semantic Segmentation. He received his Doctor of Informatics in 2021. Since 2013, Dr. Mahmud has been a Lecturer at School of Computing of Telkom University, where he is currently the Artificial Intelligence Laboratory Supervisor. His research topics are related to artificial intelligence, machine learning, and computer vision. His main activities include lecturing, conducting research and community services. He has received numerous funding grants, both from internal of Telkom University sources, as well as from external sources, including the Ministry of Research, Technology, and Higher Education and the Indonesia Endowment Fund for Education (LPDP) under the Ministry of Finance. Currently, he becomes the members of IEEE and ACM.