Scientific Journal of Informatics
Vol. 10, No. 2, May 2023

p-ISSN 2407-7658 http://journal.unnes.ac.id/nju/index.php/sji e-ISSN 2460-0040

Selection of Food Identification System Features Using Convolutional Neural Network (CNN) Method

Arnita[1*], Faridawaty Marpaung[2] , Zainal Abidin Koemadji[3], Mhd Hidayat[4], Aziz Widianto[5], Fitrahuda Aulia[6]

[1,4,5,6]Department of Computer Science, Universitas Negeri Medan, Indonesia
[2]Department of Mathematics, Universitas Negeri Medan, Indonesia
[3]School of Education, University of Leicester, United Kingdom

Abstract.
Purpose: The identification and selection of food to be consumed are critical in determining the health quality of human life. Our diet and the illnesses we develop are closely linked. Public awareness of the significance of food quality has increased due to the rising prevalence of degenerative diseases such as obesity, heart disease, type 2 diabetes, hypertension, and cancer. This study aims to develop a model for food identification and identify aspects that can aid in food identification.
Methods: This study employs the convolutional neural network (CNN) approach, which is used to identify food objects or images based on the detected features. The images of thirty-five different types of traditional, processed, and western foods were gathered as the study's input data. The image data for each type of food was repeated 100 times to produce a total of 3500 images. Using the color, shape, and texture information, the food image is retrieved. The hue, saturation, and value (HSV) extraction method for color features, the Canny extraction method for shape features, and the gray level co-occurrence matrix (GLCM) method for texture features, in that sequence, were used to evaluate the data in addition to the CNN classification method.
Results: The simulation results show that the classification model's accuracy and precision are 76% and 78%, respectively, when the CNN approach is used alone without the extraction method. The CNN classification model and HSV color extraction yielded an accuracy and precision of 51% and 55%, respectively. The CNN classification model with the Canny texture extraction method has an accuracy and precision of 20% and 20%, respectively, while the combined CNN and GLCM extraction methods have 67% and 69% success rates, respectively. According to the simulation results, the food classification and identification model that uses the CNN approach without the HSV, Canny, and GLCM feature extraction methods produces better results in terms of accuracy and precision model.
Novelty: This research has the potential to be used in a variety of food identification applications, such as food and nutrition service systems, as well as to improve product quality in the food and beverage industry.

Keywords: CNN, HSV, Canny, GLCM, Food identification
Received April 2023 / **Revised** May 2023 / **Accepted** May 2023

INTRODUCTION

Computer vision is an emerging field and its advancement is necessary for the overall revolution in computer science. Object detection involves some main elements like datasets, algorithms, and techniques. In object detection, normally have a dataset of some objects and this study wants the computer to detect those objects without telling the computer about objects. The whole process is divided into many parts and learning can be of many types. The types of learning include supervised and unsupervised learning [1], [2]. The object is seen with the human sense of sight and the image of the object is analyzed and interpreted using artificial intelligence so that it can be used for decision-making. Computer vision is used in the development of self-driving cars, Google translation software, facial recognition, and many other applications. It also entails using QR scans on a computer or smartphone to identify objects in order to make payments or other transactions [3],[4].

*Corresponding author.
Email addresses: arnita@unimed.ac.id
DOI: 10.15294/sji.v10i2.

Computer vision is one of the emerging technologies that may be applied to data science by combining it with deep learning. Deep learning is a branch of machine learning in which the algorithm is influenced by the human brain network and can recognize patterns and information from unstructured or unlabeled data without supervision [5]. Depending on the purpose of the analysis, several deep learning algorithms evolve. The convolutional neural network (CNN) is a deep learning algorithm that can analyze unstructured data such as video or image data. CNN is a deep learning technique created to compensate for the drawbacks of earlier approaches [6]. There are some shortcomings in previous conventional methods, but with the CNN model, several independent parameters can be reduced and input image deformations such as translation, rotation, and scale can be handled [7]. CNN is a more effective artificial neural network architecture for image categorization. The main idea behind CNN is its convolution operation, which extracts an image for each feature to form several patterns that are easier to classify [8]. Learning by using low-resolution images, image quality can be improved by changing the parameters in the CNN method so that image quality can be improved [9].

CNN is an example of a neural network that is frequently used in the processing of image data. Convolution is a matrix that is used to filter an image. Each process is filtered through a number of CNN layers. This procedure is known as the training procedure. The training procedure is divided into three phases: the convolutional layer, the pooling layer, and the fully connected layer [10], [11]. In general, the CNN model adheres to the architecture depicted in Figure 1.

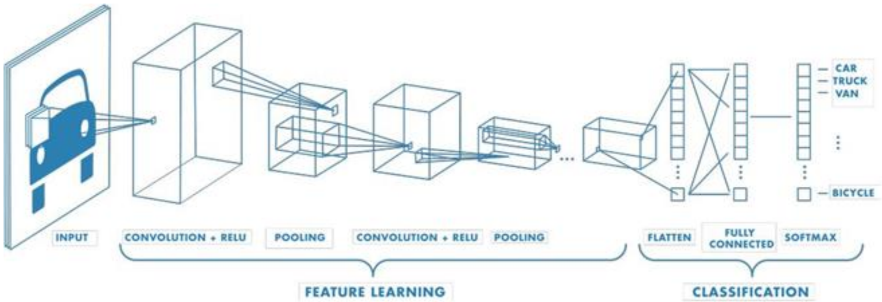

Figure 1. Operation of a convolutional neural network

According to various studies, the advantages of CNN include the capacity of the CNN method to increase the speed of the face detection process based on research findings [11], [12]. the CNN method was used to perform a face detection algorithm representation using multilayer as a feature extractor to automatically obtain special features [13]. According to the results of these studies, the CNN approach is capable of solving highly complicated issues with a very quick system. CNN outperforms the KNN algorithm as well. According to [14] CNN has a 13% higher accuracy than KNN.

Protein, fat, vitamins, and minerals are examples of nutrients found in healthy food. Eating well is very helpful for enhancing organ function and general wellness. Modern lifestyles frequently cause people to disregard their health, failing to maintain a healthy diet, and consume nutritious foods. Diet is strongly related to health, but many individuals are unaware of the calorie content of the foods they eat. Most people do not want to bother remembering and measuring how many calories they consume. A sizable amount of study has recently been conducted in relation to the development of applications for categorizing food types and determining the calorie composition of foods [15]. [16] with the title "The Food Recognition Benchmark: Using Deep Learning to Recognize Food in Images" is one of the researchers who have created food identification applications. The objective of this study is to classify 273 different types of food. The recall average is 0.885. The following study is titled "Machine Learning Based Approach on Food Recognition and Nutrition Estimation"[17]. The prototype system is made up of three software components, specifically a classification pretraining modulean attribute estimation model, to create a food classification model, as well as a server-side module. The experiment was repeated until the categorization model produced the maximum accuracy value.

Feature extraction is the process of extracting features from an object that can be used to describe the object's characteristics [18]. The process of indexing an image database together with its contents is referred to as feature extraction. Mathematically, each feature extraction is an encoding of an n-dimensional vector called a feature vector. Image processing and analysis techniques determine the feature vector component, which is subsequently applied to image comparison. The three types of feature extraction include Low-level, middle-level, and high-level feature extraction. Low-level feature extraction is based on visual content such as color and texture, middle-level extraction is based on the image region identified by segmentation, and high-level feature extraction is based on semantic information present in the image. In digital data processing, there are several types of feature extraction: color, shape, and texture are among them[19]. This study employs the following color, shape, and texture feature extraction methods: hue, saturation, and value (HSV), Canny, and gray level co-occurrence matrix (GLCM).

Color segmentation is the separation of image segments based on the colors in the image. Color segmentation has been carried out using a variety of techniques in the development of computer vision systems, including the clustering method and the index method. The HSV color scope is made up of three components: hue, which represents color; saturation, which represents the level of color dominance; and value, which represents the level of brightness. As a result, this technique recognizes color as well as the level of dominance and brightness [20]. One of the more recent edge-detection algorithms is the Canny method. A professional operator is the best at detecting edges. The operator filters noise from the initial image with a Gaussian derivative kernel to achieve smooth edge detection results [21]. The Canny edge detector is designed to address a variety of problems in three stages: a) image blurring with a Gaussian function to remove image noise and spots; b) gradient detection using one of the filters listed above, which results in two images, one with the gradient magnitude (G) and one with the orientation (G); and c) the gradient magnitude, which must be greater than a particular minimum threshold value to detect only the main edge [22].

The GLCM is a technique for texture analysis and feature extraction. The GLCM is a frequency of occurrence matrix in an image of pairs of two pixels with a specified intensity at a certain angle and location [23]. The GLCM method has been demonstrated to be a useful texture descriptor when compared to other texture extraction methods in terms of accuracy and computational time [24]. As a result, food recognition modeling was carried out throughout this work by utilizing the CNN method in conjunction with preprocessing techniques for HSV, Canny, and GLCM feature extraction imagery.

This paper is divided into three sections, the first of which is the background, which explains the research background, the issues raised, and the research objectives. Section 2 discusses the data, data sources, and features employed, as well as the data analysis process until the findings of food identification using the CNN approach and HSV, Canny, and GLMC feature extraction are received.

METHODS

In this study, we apply the CNN algorithm using preprocessing of HSV color extraction, Canny shape extraction, and GLCM texture extraction to identify the type of food. This study used 35,000 data points from 35 different food groups. The information is derived from free Internet search results. According to [25], the model's performance as they increase in number. The method of separating training and testing data has been very influential in increasing the algorithm's predicted performance. The data-splitting algorithm has had less influence on prediction results when the data size is large. However, if the amount of data used is small, data separation can be performed to improve the algorithm's prediction performance. A good balance of training size, validation, and testing, on the other hand, can result in a stable model. Due to this, two components of the data set were utilized in this study: training data for developing the model and testing data for model validation, with a 70%:30% split selected at random. We can achieve the greatest outcomes with valid and accurate estimates of accuracy by using a 30:70 data composition for training and testing [26].

Convolutional Neural Network (CNN) Algorithm

CNN is a multilayer perceptron created to process two-dimensional data in the form of images [6]. CNN is one of the deep learning techniques that produce notable results because it tries to emulate the image recognition system in the human visual cortex in order to interpret image data. CNN is an architecture that can be trained and consists of several stages. The CNN approach is broken down into two phases. The first stage involves feedforward image classification and the second stage is the learning phase using the backpropagation method [10]. CNNs communicate in the same way that human nerve cells do, with interconnected neurons that share a common architecture. The convolutional operation that applies filters to each component of the prior input to extract patterns and feature maps is what distinguishes CNN. CNN's main stages are as follows [27]:

1. Convolution Layers
   Convolution is a mathematical operation that combines two types of information. Convolution is applied to the input data via a convolution filter to generate a feature map.
2. Layer of Pooling
   The pooling layer reduces the number of parameters in the input tensor, which a) aids in the reduction of overfitting, b) removes representative features from the input tensor, and c) reduces computation and thus aids in efficiency.
3. Layer Completely Connected
   The output of the final flattened pooling or convolutional layer is then fed into the fully connected layer. The final pooling and convolutional layer produces a three-dimensional matrix, which must be flattened by changing all the values to vectors. These flattened vectors are then connected to the same number of fully connected layers as the neural network and perform the same mathematical operations. The following calculations in Equation (1) are performed for each layer of the artificial neural network:
   $$g\ (W_x + b) \tag{1}$$
   where x is a [p l, 1] input vector.
   W: [p l, n l] weight matrix where p l represents the previous layer's number of neurons and n l represents the current layer's number of neurons
   b: bias vector of size [p l, 1]
   g: activation function, ReLU
4. Dropout
   Dropout is a neural network regularization technique in which some neurons are chosen at random and not used during training. These neurons are essentially discarded randomly. This means that the discarded neurons will no longer contribute to the network, and new weights will not be applied to neurons during backpropagation.

Goodness of Model Test

The confusion matrix is a technique for measuring the performance of a model, particularly in the classification case (supervised learning) in machine learning. The confusion matrix is sometimes referred to as the error matrix. The confusion matrix essentially compares the classification results of the system (model) to the actual classification results. The confusion matrix is a matrix table that describes the performance of the classification model on a set of test data with known actual values [24]. The confusion matrix is also used to calculate various performance metrics to assess the model's performance. Accuracy, precision, and recall are three popular performance indicators that are widely and often utilized [28]. Table 1 differentiates accuracy, precision, and recall in general:

Table 1. Confusion matrix

		True Value	
		TRUE	FALSE
Predictive value	TRUE	TP (True Positive) Correct outcome	FP (False Positive) Unexpected outcome
	FALSE	FN (False Negative) Missing outcome	TN (True Negative) Correct the lack of a result

$$\text{accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{2}$$

$$\text{precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{3}$$

$$\text{recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{4}$$

$$\text{F1} = \frac{\text{2 x recall x precision}}{\text{recall} + \text{precision}} \tag{5}$$

The following formulas describe the distinction between accuracy, precision, and recall [29]:
1. Accuracy is defined as the degree of similarity between the predicted and actual value.
2. Precision is the degree of agreement between the user's request and the system's response.
3. Recall is the system's success rate in retrieving information.
4. The F1-score represents the harmonic average of precision and recall.

Color Extraction Using HSV

To form the desired segment, a pixel sample is chosen as a color reference in the segmentation method with the HSV color detection described in [30]. The RGB color model serves as the color reference standard in digital images, so the initial stage in this process is to translate the RGB color model to HSV. According to [31], the segmentation process is broadly defined as follows:

a) Identify the RGB image that is the target of detection, the HSV color value used as a reference (the outcome of the data training process), and the HSV tolerance value used.
b) Convert the RGB image to HSV.
c) Apply a color filter to the image based on the reference value (T) and the tolerance value (toll).
d) The existing pixels.
e) Revert the image to RGB and show the filter results.

Extraction of Texture Using GLCM

According to [32], there are a number of methodologies frequently employed for texture analysis, including statistical methods, structural methods, model-based methods, and transformation-based methods. The GLCM is one "one such statistical techniquetechnique. GLCM is a popular statistical technique in texture analysis. [33] defines co-occurrence as an event in which one pixel's gray level is adjacent to another pixel's gray level. Suppose that L is the distance between two adjacent pixels, which is the angle between pixels in degrees, and N is the number of pixel intensity levels in an image. Then GLCM is a P[i,j] square matrix with the dimension N^2. Image representations of different texture features are generated from the value of the co-occurrence matrix. These characteristics are as follows: [34] cluster advantage, cluster shade, dissimilarity, energy, entropy, homogeneity, maximum probability, a sum of squares (variance), and the sum of averages are some of the GLCM features that are frequently used. Mean, variance, entropy, variance difference, entropy difference, correlation-measure information, normalized inverse difference, and inverse difference when normalized.

Extraction of Shape Using Canny Edge

Shape features are the characteristics of an object that are defined by lines and contours. Shape attributes are categorized depending on the technique used. There are two types of categories: those based on borders and those based on regions (region-based). Boundary-based approaches define the geometry of a region by using external properties, such as pixels along the object boundary. The clever edge detection method can be used to locate an object's edges. Edge detection is the process of finding sections of an image where the color intensity has been dramatically altered. The Canny edge detection algorithm works as follows[35]:
1. Edge detection is smoothed to lessen the effect of noise.
2. Determine the image's gradient potential.
3. Image gradient non-maximal suppression for precise edge localization.
4. Hysteresis thresholding for final classification.

Canny's algorithm can satisfy some of the best edge detection requirements [36]:

a) Accurate detection
b) Accurate localization
c) Accurate response

The Gaussian value's standard deviation and the threshold value have a significant influence on the extraction outcomes. The convolution of the image function with the Gaussian operator and its derivatives is used in the Canny algorithm approach. The first derivative of the image function is convolved with the Gaussian function, can be shown in Equation (6).

$$g(x, y) = D[gauss(x, y) * f(x, y)] \tag{6}$$

$g(x, y)$	=	Gaussian convolutional function
$D[gauss(x, y) * f(x, y)]$	=	First derivative of Gaussian convolutional function

$$= \left(\frac{x^2 + y^2 - \sigma^2}{\sigma^4}\right) e^{\frac{x^4+y^4}{2\sigma^4}}$$

Equivalent to the image function convoluted with the Gaussian function's first derivative. Hence, the level of smoothness and edge detection can be combined into a convolution in one dimension with two distinct orientations (vertical and horizontal).

RESULTS AND DISCUSSIONS

Results

In this study, 35 different food groups were used, with each food group using 100 different images. Hence, the data set used is 3500. Table 2 shows the food samples.

Table 2. Food types tested in the identification model

No.	Type	No.	Type
1	Red Rice	19	Capcay
2	White Rice	20	Curry
3	Water Spinach	21	Pempek
4	Tempeh	22	Potato Sauce
5	Steak	23	Fried Noddles
6	Rendang	24	Fried Fish
7	Sausages	25	Yellow Rice
8	Nuggets	26	Black Glutinous Porridge
9	Green Bean Porridge	27	Corn Porridge
10	Mashed Sweet Potato Leaves	28	Pumpkin Compote
11	Vegetable	29	Gomak Noddles
12	Tofu	30	Anyang
13	Perkedel	31	Gado-gado
14	Egg Fried	32	Sour Vegetable
15	Fried Chicken	33	Jengkol
16	Urap	34	Vegetable Lodeh
17	Meatball Soup	35	Cabbage Stir-Fry
18	Satay		

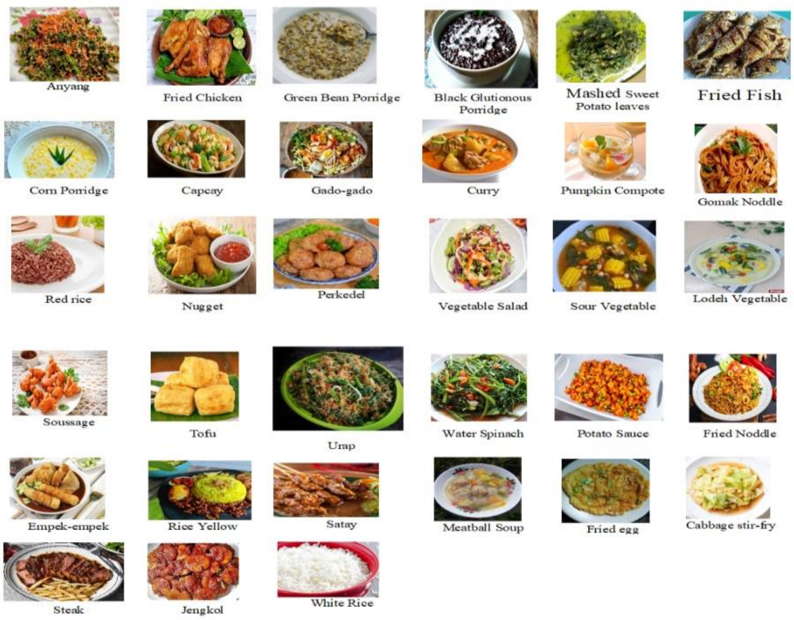

Figure 2. Examples of the type of food

Figure 2 shows a selection of food images used in this study. Rice, porridge, processed vegetables, and meats are among the foods eaten. Based on the collected food image data, the image is extracted using three predetermined features: the HSV color feature, Canny edge detection, and GLCM. The extraction of the three features yields the following information.

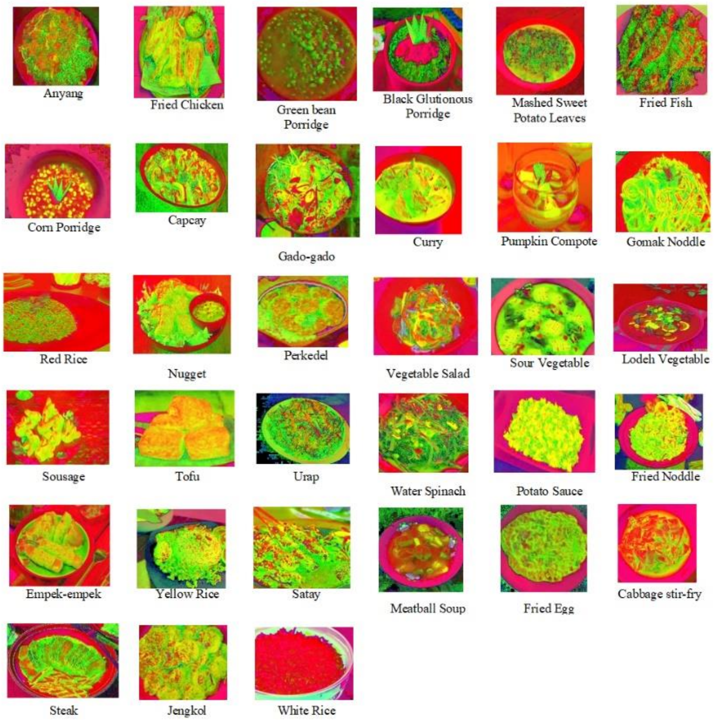

Figure 3. Example of color feature extraction using the HSV method

Figure 3 shows the outcome of the HSV color extraction. The HSV color defines the color in HSV terminology. The true colors of the hue are red, yellow, and violet, with the goal of being able to define and differentiate colors such as red, green, and others, to assess the authenticity of a color, and use saturation color, while the value "color" is used to gage how much light the eye takes independent of hue. The HSV extraction results are then used as input for food classification using CNN.

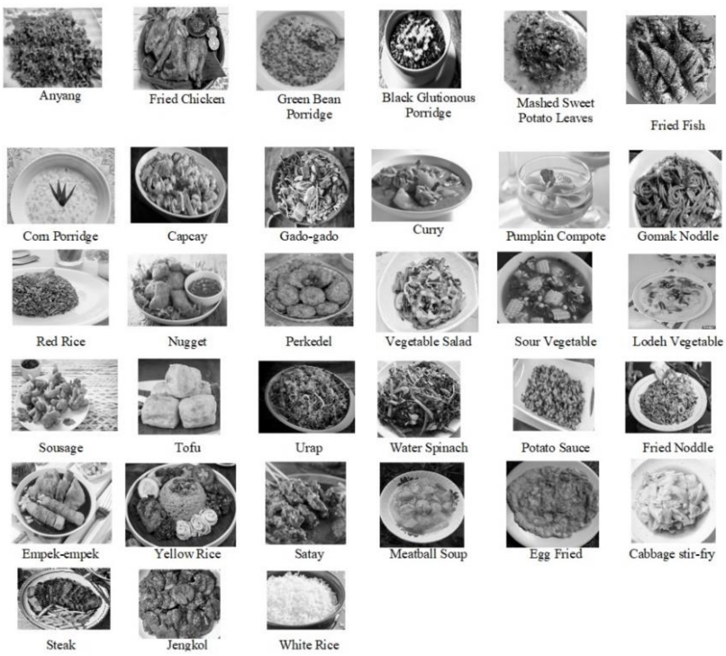

Figure 4. GLCM shape feature extraction results

Figure 4 shows the result of the GLCM extraction method. The co-occurrence matrix employs a grayscale degree matrix to demonstrate how one degree of gray occurs in conjunction with another degree of gray. The gray degree matrix is a matrix whose elements are the frequencies of relative co-occurrence of a specific combination of inter-gray level pixel pairs [37].

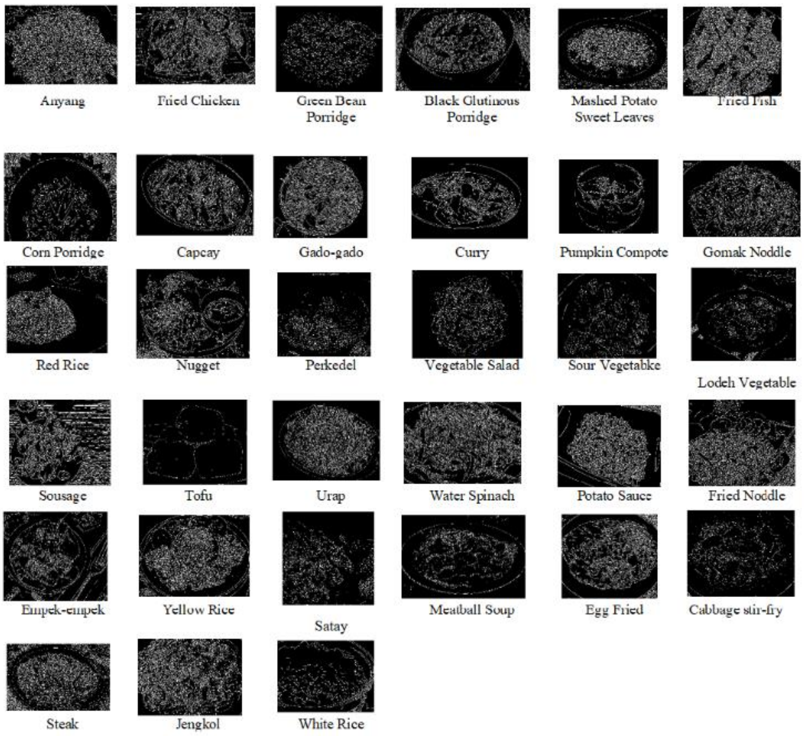

Figure 5. Texture feature extraction sample using the Canny edge detection method

Figure 5 depicts the outcome of edge extraction using the Canny method. The Canny approach is the result of separating light and dark shadows, which gives the method the advantage of being able to lower the noise before performing edge detection calculations. This method can reveal the contents of objects in each image [38]. Food images, on the other hand, have less relevance in the research of object separation because the objects in them are so diverse and complex.

Discussion

The food identification model is developed in two stages: the first stage involves using the model developed on the original data using the CNN method, followed by feature extraction. This is done to evaluate the CNN algorithm's capability to perform automatic extraction, which is one of CNN's advantages in identifying image data. Color, shape, and texture features are used to classify the extracted features, and a food identification model is created using the CNN algorithm. Based on the model's goodness evaluation outcomes, the precision, recall, F1-score, and accuracy metric values of each feature extraction method and CNN algorithm are provided in Table 3.

Table 3. Good food image classification model results

Data	Precision	Recall	F1-score	Accuracy
Normal (CNN)	0.78	0.76	0.76	0.76
HSV + CNN	0.55	0.51	0.50	0.51
Canny + CNN	0.20	0.20	0.18	0.20
GLCM + CNN	0.69	0.67	0.67	0.67

Based on the results of the classification and identification analysis using the CNN algorithm, the best value for the goodness of the model was obtained when using the original data (without extraction). This is reflected in the highest precision value when compared to the data retrieved using the HSV, Canny, and GLCM features. When using the normal data, the precision value obtained is 0.78. Meanwhile, using the extracted data, the values for the HSV, Canny, and GLCM features were obtained as 0.55, 0.20, and 0.69, respectively. This corresponds to the findings of a study [39] which demonstrated that CNN was capable of performing data preprocessing techniques using an extraction process comprised of several hidden layers, namely, the convolution layer, ReLU, and pooling. CNN operates in a hierarchical way, with the output of the first convolution layer serving as the input for the subsequent convolution layer.

Similarly, with the recall value in the original data (without extraction), the CNN method can produce the highest value when compared to feature extraction data. Table 3 shows that the recall value in the original data (without extraction) is 0.76, whereas the recall values in the data recovered using the HSV, Canny, and GLCM features are 0.51, 0.20, and 0.69, respectively. The same was observed with the F1-score parameter; normal data produced the greatest result with a large F1-score value of 0.76, while data extracted from HSV, Canny, and GLCM yielded 0.50, 0.18, and 0.67, respectively. On the accuracy parameter, the original data also has the highest accuracy value (0.76) when compared to the data extracted from the HSV, Canny, and GLCM features, which have accuracy values of 0.51, 0.20, and 0.67, respectively. CNN algorithm works well because CNN performs data transformation before examining the data for classification purposes [25].

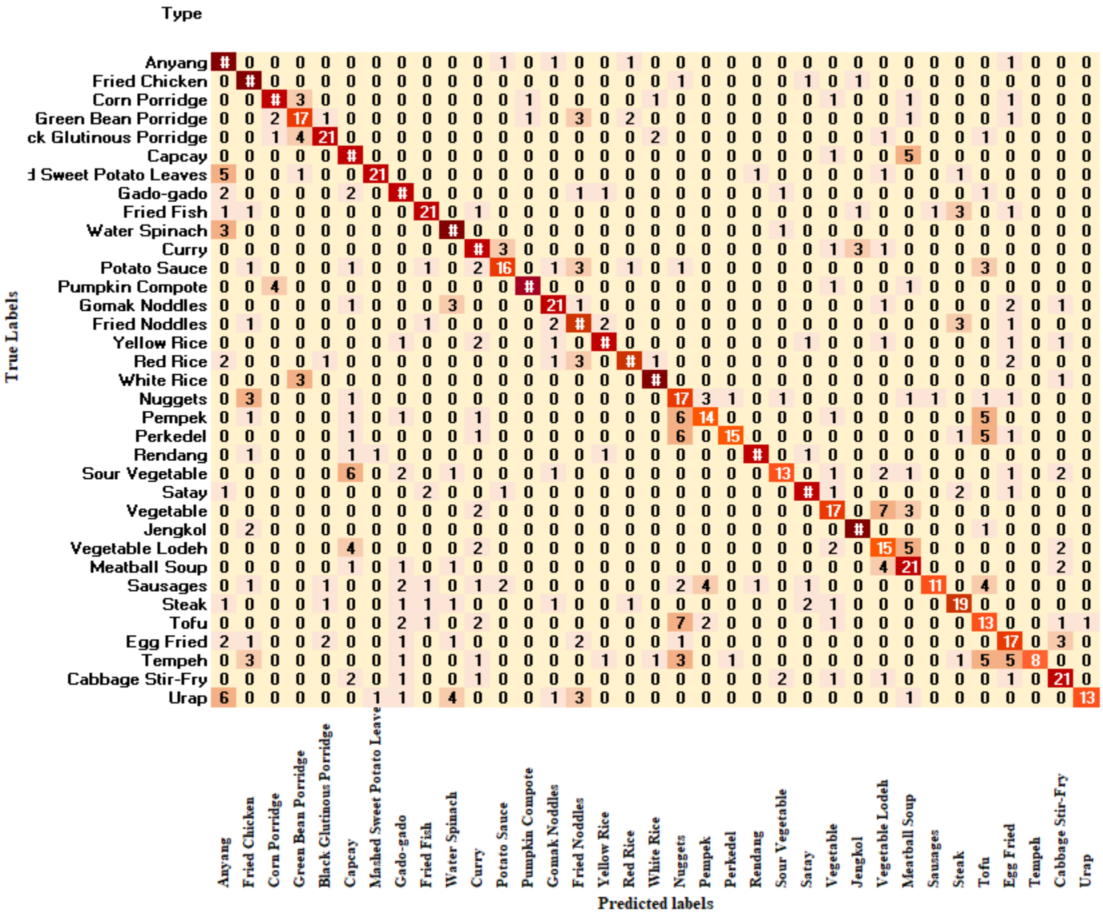

Figure 6. Confusion matrix food identification model without extraction technique using CNN

Figure 6 depicts a confusion matrix of the training data. The confusion matrix in Fig. 6 shows the miss class of numerous related objects in each food class. 3500 available data points, 25% are used as training data. Figure 6 shows some miss class items, despite the fact that "all data can *be* predicted with improved accuracy overall. Only the tofu object has the lowest accuracy; only 9 of the 25 data sets used as training data can be accurately predicted, while the remaining are predicted to be misclassified as empek-empek, tempeh, and cakes. Table 4 shows the prediction accuracy for each food item.

Table 4. The value of prediction accuracy (precision) on testing data

Food type	Precision	Recall	F1-Score	Support
Anyang	0.75	0.72	0.73	25
Fried Rice	0.54	0.66	0.59	29
Green Bean Porridge	0.82	0.85	0.84	33
Corn Porridge	0.92	0.68	0.78	34
Black Glutinous Porridge	0.85	1.00	0.92	33
Capcay	0.73	0.73	0.73	33
Mashed Sweet Potato Leaves	0.94	0.94	0.94	29
Gado-gado	0.61	0.71	0.66	33
Fried Fish	0.71	0.92	0.80	34
Water Spinach	0.92	0.89	0.91	33
Curry	0.60	0.55	0.57	33
Potato Sauce	0.90	0.74	0.81	31
Pumpkin Compote	0.83	0.87	0.85	31
Fried Noddle	0.73	0.59	0.65	24
Gomak Noddle	0.71	0.86	0.77	27
Yellow Rice	0.77	0.80	0.78	28
Red Rice	0.95	0.81	0.88	25
Nuggets	0.79	0.82	0.81	30
Pempek	0.76	0.90	0.82	33
Perkedel	0.86	0.67	0.75	31
Rendang	0.95	0.78	0.86	27
Salad Vegetable	0.55	0.71	0.62	27
Satay	0.85	0.81	0.83	24
Sour Vegetable	0.77	0.73	0.75	36
Jengkol	0.71	0.94	0.81	33
Lodeh Vegetable	0.76	0.41	0.53	32
Meatball Soup	0.60	0.88	0.72	33
Sausages	0.74	0.65	0.69	26
Steak	0.74	0.85	0.79	27
Tofu	0.90	0.38	0.53	24
Fried Egg	0.81	0.53	0.64	32
Tempeh	0.76	0.81	0.78	36
Cabbage Stir Fry	0.96	0.76	0.85	29
Urap	0.50	0.75	0.60	28
Accuracy			0.76	1040
Macro Average	0.78	0.76	0.76	1040
Weighted Average	0.78	0.76	0.76	1040

CONCLUSION

According to the results of the classification and identification study using the CNN algorithm, the best value for the goodness of the model was obtained when using the original data (without extraction). This is reflected in the highest precision value when compared to data extracted using the HSV, Canny, and GLCM features. When utilizing normal data, the precision value is 0.78, while using extracted data yielded precision values of 0.55, 0.20, and 0.69 for the HSV, Canny, and GLCM features, respectively.

ACKNOWLEDGMENT

We would like to thank Universitas Negeri Medan for funding this study. Especially the rector, LPPM, the dean of FMIPA Universitas Negeri Medan, and the department head who have provided support and

facilities for the implementation of this series of research. Although not perfect, the findings of this study should be beneficial to the advancement of knowledge and the welfare of many people.

REFERENCES

[1] U. Arshad, “Object Detection in Last Decade - A Survey,” *Sci. J. Informatics*, vol. 8, no. 1, pp. 60–70, 2021, doi: 10.15294/sji.v8i1.28956.

[2] M. A. Muslim *et al.*, “New model combination meta-learner to improve accuracy prediction P2P lending with stacking ensemble learning *,” *Intell. Syst. with Appl.*, vol. 18, no. December 2022, p. 200204, 2023, doi: 10.1016/j.iswa.2023.200204.

[3] K. Umam and B. S. Negara, “Deteksi Obyek Manusia Pada Basis Data Video Menggunakan Metode Background Subtraction Dan Operasi Morfologi,” *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 2, no. 2, p. 31, 2016, doi: 10.24014/coreit.v2i2.2391.

[4] G. Chung, B. Gesing, G. Steinhauer, M. Heck, and K. Dierkx, “Important_Artificial Intelligence in Logistics,” *DHL IBM*, p. 45, 2018.

[5] I. G. and Y. B. and A. Courville, *Deep Learning*. 2016.

[6] A. Anton, N. F. Nissa, A. Janiati, N. Cahya, and P. Astuti, “Application of Deep Learning Using Convolutional Neural Network (CNN) Method For Women’s Skin Classification,” *Sci. J. Informatics*, vol. 8, no. 1, pp. 144–153, 2021, doi: 10.15294/sji.v8i1.26888.

[7] Y. Lecun, L. Bottou, Y. Bengio, and P. Ha, “LeNet,” in *Proceedings of the IEEE*, 1998, no. November, pp. 1–46.

[8] X. Luo, R. Shen, J. Hu, J. Deng, L. Hu, and Q. Guan, “A Deep Convolution Neural Network Model for Vehicle Recognition and Face Recognition,” *Procedia Comput. Sci.*, vol. 107, no. Icict, pp. 715–720, 2017, doi: 10.1016/j.procs.2017.03.153.

[9] A. K. Nugroho, I. Permadi, and M. Faturrahim, “Improvement Of Image Quality Using Convolutional Neural Networks Method,” *Sci. J. Informatics*, vol. 9, no. 1, pp. 95–103, May 2022, doi: 10.15294/sji.v9i1.30892.

[10] A. Patil and M. Rane, “Convolutional Neural Networks: An Overview and Its Applications in Pattern Recognition,” *Smart Innov. Syst. Technol.*, vol. 195, pp. 21–30, 2021, doi: 10.1007/978-981-15-7078-0_3.

[11] F. Maulana, T. Sri, and N. Asih, “Hoax Classification in Indonesian Language with Bidirectional Temporal Convolutional Network Architecture,” *J. Soft Comput. Explor.*, vol. 4, no. 1, pp. 39–52, 2023.

[12] G. Guo, H. Wang, Y. Yan, J. Zheng, and B. Li, “A fast face detection method via convolutional neural network,” *Neurocomputing*, vol. 395, no. Bo Li, pp. 128–137, 2020, doi: 10.1016/j.neucom.2018.02.110.

[13] F. Shu Zhan, Qin-Qin Tao, Xiao-Hong Li, “Face detection using representation learning,” *Neurocomputing*, vol. 3, no. 2, pp. 80–91, 2016.

[14] Arnita, M. Yani, F. Marpaung, M. Hidayat, and A. Widianto, “A comparative study of convolutional neural network and k-nearest neighbours algorithms for food image recognition,” *J. Comput. Technol.*, vol. 27, no. 6, pp. 88–99, 2022, doi: 10.25743/ICT.2022.27.6.008.

[15] H. Cena and P. C. Calder, “Defining a Healthy Diet : Evidence for the Role of,” *Nutrients*, vol. 12, no. 334, pp. 1–15, 2020.

[16] S. P. Mohanty *et al.*, “The Food Recognition Benchmark: Using Deep Learning to Recognize Food in Images,” *Front. Nutr.*, vol. 9, no. May, pp. 1–13, 2022, doi: 10.3389/fnut.2022.875143.

[17] Z. Shen, A. Shehzad, S. Chen, H. Sun, and J. Liu, “Machine Learning Based Approach on Food Recognition and Nutrition Estimation,” *Procedia Comput. Sci.*, vol. 174, pp. 448–453, 2020, doi: 10.1016/j.procs.2020.06.113.

[18] L. A. Ruiz, A. Recio, A. Fernández-Sarría, and T. Hermosilla, “A tool for object descriptive feature extraction: Application to image classification and map updating,” *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.*, vol. 38, no. 4C7, 2010.

[19] D. R. Tobergte and S. Curtis, *Algorithms for image prcessing and computer vision*, vol. 53, no. 9. 2013.

[20] J. Mantik, R. Antika, and M. Iqbal, “Using the Matlab App to Detect Objects By Color With Hsv Color Segmentation,” *J. Mantik*, vol. 5, no. 4, pp. 2702–2708, 2022.

[21] N. D. Lynn, A. I. Sourav, and A. J. Santoso, “Implementation of Real-Time Edge Detection Using Canny and Sobel Algorithms,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1096, no. 1, p. 012079, 2021, doi: 10.1088/1757-899x/1096/1/012079.

[22] E. A. Sekehravani, E. Babulak, and M. Masoodi, “Implementing canny edge detection algorithm

for noisy image," *Bull. Electr. Eng. Informatics*, vol. 9, no. 4, pp. 1404–1410, 2020, doi: 10.11591/eei.v9i4.1837.
[23] M. O'Byrne, B. Ghosh, V. Pakrashi, and F. Schoefs, "Texture Analysis Based Detection And Classification Of Structure Features On Ageing Infrastructure Elements," *Hal*, pp. 1–6, 2018.
[24] F. Roberti de Siqueira, W. Robson Schwartz, and H. Pedrini, "Multi-scale gray level co-occurrence matrices for texture description," *Neurocomputing*, vol. 120, pp. 336–345, 2013, doi: 10.1016/j.neucom.2012.09.042.
[25] D. E. Birba, "A Comparative study of data splitting algorithms for machine learning model selection," *Degree Proj. Comput. Sci. Eng.*, no. December, 2020.
[26] A. Gholamy, V. Kreinovich, and O. Kosheleva, "Why 70/30 or 80/20 Relation Between Training and Testing Sets : A Pedagogical Explanation," *Dep. Tech. Reports*, vol. 1209, pp. 1–6, 2018.
[27] R. Sarki, K. Ahmed, H. Wang, Y. Zhang, and K. Wang, "Convolutional Neural Network for Multi-class Classification of Diabetic Eye Disease," *EAI Endorsed Trans. Scalable Inf. Syst.*, vol. 9, no. 4, 2022, doi: 10.4108/eai.16-12-2021.172436.
[28] C. O. A. Freitas, J. M. De Carvalho, J. J. Oliveira, S. B. K. Aires, and R. Sabourin, "Confusion matrix disagreement for multiple classifiers," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 4756 LNCS, no. May 2014, pp. 387–396, 2007, doi: 10.1007/978-3-540-76725-1_41.
[29] E. R. Davies, *Computer Vision: Principles, Algorithms, Applications, Learning: Fifth Edition*. 2017.
[30] J. Ilmiah and P. Teknik, "Implementation Segmentation of Color Image With," *VOLT (Jurnal Ilm. Pendidik. Tek. Elektro)*, vol. 2, no. 2, pp. 157–166, 2017.
[31] V. B. Vladimir Chernov, Jarmo Alander, "Integer-based accurate conversion between RGB and HSV color spaces," *Comput. Electr. Eng.*, vol. 46, pp. 328–337, 2015.
[32] A. Ramola, A. K. Shakya, and D. Van Pham, "Study of statistical methods for texture analysis and their modern evolutions," *Eng. Reports*, vol. 2, no. 4, pp. 1–24, 2020, doi: 10.1002/eng2.12149.
[33] T. I. Simanjuntak, S. Suwilo, and R. W. Sembiring, "Analysis of Detection of Drow and Entire Co-Occurrence Matrix GLCM Method on the Classification of Image," *J. Phys. Conf. Ser.*, vol. 1361, no. 1, 2019, doi: 10.1088/1742-6596/1361/1/012028.
[34] Ş. Öztürk and B. Akdemir, "Application of Feature Extraction and Classification Methods for Histopathological Image using GLCM, LBP, LBGLCM, GLRLM and SFTA," *Procedia Comput. Sci.*, vol. 132, no. Iccids, pp. 40–46, 2018, doi: 10.1016/j.procs.2018.05.057.
[35] B. Ramprakash, D. Indumathi, and C. Science, "Text detection from traffic regulatory sign boards," vol. 11, no. 5, pp. 11–14, 2020.
[36] Q. Xu, S. Varadarajan, C. Chakrabarti, and L. J. Karam, "A distributed canny edge detector: Algorithm and FPGA implementation," *IEEE Trans. Image Process.*, vol. 23, no. 7, pp. 2944–2960, 2014, doi: 10.1109/TIP.2014.2311656.
[37] A. Septiarini and R. Wardoyo, "Kompleksitas Algoritma GLCM untuk Ekstraksi Ciri Tekstur pada Penyakit Glaucoma," *Pros. Semin. Nas. Komun. dan Inform.*, no. 978-602-72127-1–8, pp. 98–102, 2015.
[38] G. F. Laxmi, P. Eosina, and F. Fatimah, "Analisis perbandingan metode prewitt dan canny untuk identifikasi ikan air tawar," *Pros. SINTAK*, pp. 201–206, 2017.
[39] A. L. Katole, K. P. Yellapragada, A. K. Bedi, S. S. Kalra, and M. Siva Chaitanya, "Hierarchical Deep Learning Architecture for 10K Objects Classification," pp. 77–93, 2015, doi: 10.5121/csit.2015.51408.
[40] W. S. Eka Putra, "Klasifikasi Citra Menggunakan Convolutional Neural Network (CNN) pada Caltech 101," *J. Tek. ITS*, vol. 5, no. 1, 2016, doi: 10.12962/j23373539.v5i1.15696.