Вычислительные технологии, 2022, том 27, № 6, с. 88–99. © ФИЦ ИВТ, 2022 ISSN 1560-7534
Computational Technologies, 2022, vol. 27, no. 6, pp. 88–99. © FRC ICT, 2022 eISSN 2313-691X

COMPUTATIONAL TECHNOLOGIES

DOI:10.25743/ICT.2022.27.6.008

A comparative study of convolutional neural network and *k*-nearest neighbours algorithms for food image recognition

Arnita[1,*], M. Yani[2], F. Marpaung[1], M. Hidayat[1], A. Widianto[1]
[1]Department of Mathematics Universitas Negeri Medan, 20221, Medan, North Sumatera, Indonesia
[2]Department of Mechanical Engineering Universitas Muhammadiyah Sumatera Utara, 20238, Medan, North Sumatera, Indonesia
[*]Corresponding author: Arnita, e-mail: arnita@unimed.ac.id
Received June 22, 2022, revised September 20, 2022, accepted September 27, 2022.

Food plays a vital role in everyday life, and public awareness of food quality has increased. The availability of many types of food has made it difficult for people to choose the right type of healthy food for consumption. The Convolutional Neural Network (CNN) and k-nearest neighbours (KNN) algorithms can be used to create classification and identification models, including food identification. Therefore, we need a system that can quickly identify the type of food and calculate the caloric value contained in the food to be consumed to maintain a healthy diet. To create the best identification model based on the goodness of the model. Metrics for accuracy, prediction, recall, and F1-score will be used for food identification using the CNN and KNN algorithms. This research method extracts food image input using the hue, saturation, and value (HSV) color space. Then the extracted data is classified using the CNN and KNN algorithms. Simulation in this study is done using 900 food images. The data is divided into two categories, namely training and test data, with a ratio of 75 and 25 %, respectively. The KNN algorithm was tested with $k = 3$, 5, and 7, in simulation process and compared with the CNN. Based on the experiments conducted, it was found that the CNN method was better than the KNN Algorithm. There are two classes of food types that are resulted with wrong predictions, while the CNN method predicts only 1 class of food type as wrong. This is indicated by the accuracy of the CNN method, which is 5 % better than the KNN(3) method. The accuracy of the CNN method is 94 %, while the accuracy of the KNN(3) method is 89 %. The F1-score value for the CNN method is 0.94 and the KNN(3) method is 0.89. The CNN allows the model to produce an average precision of 87.7 %, the accuracy of 86.89 %, recall of 86.89 %, and F1-score of 86.33 %. The model formed using CNN is the best food identification model based on this simulation.

Keywords: food image recognition, convolutional neural network, k-nearest neighbours, HSV color space.

Citation: Arnita, Yani M., Marpaung F., Hidayat M., Widianto A. A comparative study of convolutional neural network and k-nearest neighbours algorithms for food image recognition. Computational Technologies. 2022; 27(6):88–99. DOI:10.25743/ICT.2022.27.6.008.

Introduction

Healthy food is one of the essential requirements for long life. However, globalization and urbanization have greatly affected people's habits of consuming fast food and luxury items with high-calorie content [1]. In recent years, food has played an essential role in everyday life because it is closely related to various diseases. The rise in various degenerative diseases such as obesity, heart disease, type 2 diabetes, hypertension, and cancer [2] has increased public awareness of the importance of food quality [3].

Image recognition identifies and detects objects or features in digital images or videos. This concept performs many machine-based visual tasks, including labelling image content with meta-tags, searching for image content, and recognizing food based on color, shape, and texture. Image recognition is a straightforward task for humans and animals, not computers. Therefore, it is necessary to do human learning in computer programming. Food image recognition is a promising visual object recognition application because it will help estimate food calories and analyze eating habits for health [4].

Feature extraction is the process of indexing an image database with its contents. Mathematically, each feature extraction is an encoded version of an n-dimensional vector called a "feature vector". The feature vector component is calculated by image processing and analysis techniques and is used to raise the features' significance and reduce the feature vectors' dimensionality [5]. Feature extraction is used for image classification. The region can be defined in a global or local environment and distinguished by shape, texture, size, intensity, and statistical properties. Several feature extractions in the image were used in this study. HSV is one of several color spaces used for feature extraction in computer vision. The HSV color space defines color in terms of hue, saturation, and value. The advantage of HSV is that there are colors that are the same as those captured by the human senses and separate the luminance color component from chrome [6].

There have been several studies on food object recognition, including [3], which conducted a study using a simplified convolutional neural network (CNN) for food recognition and proposed jumping convolution to extract food image features. CNN is one type of deep learning model to process data with a grid pattern like the image [7]. A CNN is a machine learning method developed from developing multi-layer perceptions, designed to process two-dimensional data. CNN also has a deep feed-forward architecture and excellent generalizability compared to a fully connected network [8]. The KNN (k-nearest neighbor) algorithm has been widely used to classify problems like classification, genetics, and forecasting. KNN is a non-parametric supervised learning method that is used for classifications and regression. KNN is a type of classification where the function is only approximated locally, and all the computation is deferred until function evaluation. KNN has several advantages, including being simple, easy, and more efficient, having quite competitive performance compared to similar methods, and being more robust to data with a lot of noise [9].

Being able to improve classification performance and reduce outlier effects, especially in small data sets [10, 11]. Based on the advantages of the KNN algorithm, we want to propose the KNN algorithm to model the food identification system in this study. Even so, KNN has some weaknesses. Some of the weaknesses of KNN are that KNN has poor run time performance during training and is very sensitive to irrelevant features and large numbers [9]. KNN is called a lazy learn algorithm because it does not build a model [12] and requires large memory to store training data [13]. Therefore, a comparison is made with the CNN method to assess the performance of two algorithms in identifying food.

CNN is a promising technique with high precision and accurate performance compared to other image processing techniques [14]. Also CNN can select features without supervision [15], and the preprocessing required is much less than other neural network techniques [16]. While the KNN algorithm has several advantages, it was chosen to be used in this study because of its ability to classify data accurately by, first, correctly selecting the k value of the nearest neighbour [7, 17]. Aside from the KNN algorithm being simple to use and intuitive to grasp, it can learn non-linear decision constraints and provide very flexible decisions based on the k values when used for classification. There is a single hyperparameter, k value, that is constantly evolving with new data. This makes fine-tuning hyperparameters simple. There are numerous distance metrics to choose from [18].

According to the findings of this study, the proposed CNN method can perform the task of recognizing food quickly and accurately, as presented in research [4], which showed a CNN-based food image segmentation that does not require pixel annotation. The study concluded that the deep CNN method (the proposed DCNN) outperforms the region-based CNN (RCNN) in detecting food regions. The study [19] also demonstrated a system for categorizing food images using the KNN algorithm. Compared to the Yahoo KNN, the system for identifying and classifying food using the Yahoo Kosakata Tree can improve accuracy. The article [20] also investigates food segmentation using the recipe learning module method (ReLeM). This study makes use of large amounts of data to segment food images. It was found that a more detailed model of food segmentation is needed. Therefore, the present study aims at developing the best food identification model based on model goodness metrics such as accuracy, prediction, recall, and F1-score using the CNN and KNN algorithms.

1. Research and methodology

1.1. HSV color space

The HSV color space defines color in terms of hue, saturation, and value. Hue represents true colors, such as red, violet, and yellow. Hue is used to distinguish between shades and determine light redness and greenness [21]. A hue value between 0 and 1 means a color between red passes through yellow, green, cyan, blue, magenta, and back to red. Saturation values ranging from 0 to 1 indicate that the color is unsaturated (gray) to fully saturated (not white) [22]. The 3-dimensional HSV vector is converted to a 1-dimensional vector while still considering the weight of each HSV component value [23]. The HSV image extraction process is carried out with the following steps [24]:

1. Input the image to be extracted.
2. Convert RGB images to HSV using the following steps [6].
   - R, G, and B values are divided by 255 to reduce the range from $[0;255]$ to $[0;1]$:

$$R' = R/255, \qquad G' = G/255, \qquad B' = B/255.$$
$$C_{\max} = \max(R', G', B'), \quad C_{\min} = \min(R', G', B'), \quad \Delta = C_{\max} - C_{\min}.$$

   - Calculate the hue:

$$H = \begin{cases} 0, & \Delta = 0 \\ 60^\circ \left([(G' - B')/\Delta] \bmod 6 \right), & C_{\max} = R', \\ 60^\circ \left((B' - R')/\Delta + 2 \right), & C_{\max} = G', \\ 60^\circ \left((R' - G')/\Delta + 4 \right), & C_{\max} = B'. \end{cases}$$

- Calculate the saturation:

$$S = \begin{cases} 0, & C_{\max} = 0, \\ \Delta / C_{\max}, & C_{\max} \neq 0. \end{cases}$$

- Calculate the value:

$$V = C_{\max}, \quad C \text{ is color.}$$

3. Separate the values of each component: hue, saturation, and value.
4. Identify H, S, and V according to the value for each feature.
5. Data will be segmented based on the H, S, and V criteria.
6. The data is ready to be processed with further analysis.

1.2. KNN algorithm

The steps of the KNN algorithm are [25]:

1. Determine the number of parameters k (number of nearest neighbours).
2. Using the following equation, calculate the distance (similarity) between all new objects

$$d(a_i, b_i) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2}$$

   where a_{il} — i-th test data on the l-th variable; b_{ij} — j-th training data on l-th variable; $d(a_i, b_i)$ — distance; N — dimension of independent variable data.
3. Sorting data by distance value from the smallest to the largest value.
4. Taking data from several values of k.
5. Determine the label that appears most frequently in the k training records closest to the object.

1.3. CNN

CNN is a deep learning method that gives significant results because it tries to imitate the image recognition system in the human visual cortex to process image information in an architecture that can be trained and consists of several stages. The CNN method consists of two stages. The first stage is image classification using feed-forward. The second stage is the learning stage with the backpropagation method [26]. CNNs mimic the way our nerve cells communicate with interconnected neurons, and CNNs share the same architecture. The convolutional operation makes it unique from other neural networks, which apply a filter to each part of the previous input to extract patterns and feature maps. Some of the main stages on CNN are described below.

Convolutional layers are the primary building blocks of CNN. Convolution is a mathematical operation that combines two sets of information. In this case, convolution is applied to the input data via a convolution filter to generate a feature map. Convolutional layers are the layers in which filters are applied to the original image or other feature maps in a deep CNN. The majority of the network's user-specified parameters are in this location. The most critical parameters are the number of kernels and the size of the kernels.

Pooling layers are used to reduce the number of parameters of the input tensor so that:

1. Helps reduce overfitting.
2. Identify representative features in the input tensor.
3. Reduce computation to improve efficiency.

Fully connected layer. The output from the final pooling or convolutional layer, which has been flattened, is then entered into the fully connected layer. The final pooling and convolutional layer results in a 3-dimensional matrix that needs to be flattened by converting all the values into 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 are used for each layer of the artificial neural network:

$$g(Wx + b),$$

where x — input vector dimension $[p_l, 1]$; W — weight matrix with dimensions $[p_l, n_l]$ where, p_l is the number of neurons in the previous layer and n_l is the number of neurons in the current layer; b — bias vector dimension $[p_l, 1]$; g — activation function.

Dropout is a neural network regularization technique where some neurons will be randomly selected and not used during training. These neurons are practically discarded randomly. This means that the contribution of discarded neurons will be stopped while the network and new weights are not applied to neurons during backpropagation.

1.4. Validation and evaluation

Cross-validation, often referred to as rotation estimation, is a model validation technique to assess the optimization of the analysis results. Besides, cross-validation is also a compositional technique in determining the amount of training data and testing data to be used. One of the most commonly used cross-validation methods is the holdout method. In this study, the holdout method is used, where the initial data that is partitioned into two random sets called training data and testing data. Data is divided into 75 % for training and 25 % for testing [27]. The evaluation aims to determine the level of success of the study. Evaluation in this study uses accuracy, precision, recall, and F1-score in the confusion matrix.

The higher the accuracy, precision, recall, and F1-score values, the better the system developed by [27, 28] to calculate the evaluation, using the following equation:

$$\text{F1-score} = \frac{1}{2}\left(\frac{1}{\text{precision}} + \frac{1}{\text{recall}}\right) \cdot 100\,\%, \quad \text{precision} = \frac{TP}{TP + FP} \cdot 100\,\%,$$

$$\text{accuracy} = \frac{TP + TN}{TP + FP + FN + TN}, \quad \text{sensitivity} = \frac{TP}{TP + FN}, \quad \text{specificity} = \frac{TN}{TN + FP},$$

where TP — number of true positives; TN — number of true negatives, P — number of positive records, N — number of negative tuples, FP — number of false positives.

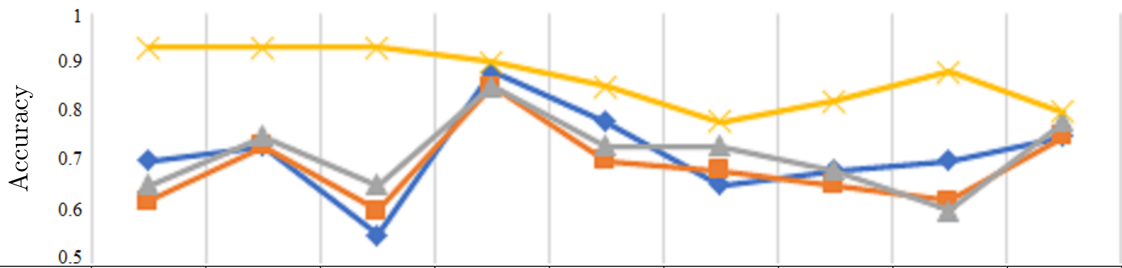
2. Result and discussion

Result. The data used in this paper are 900 food images, consisting of images of nine types of food: tempeh, steak, sausage, rendang, nuggets, rice, red rice, saut'eed water spinach, and green bean porridge (Fig. 1). Photos of the food were taken with a smartphone camera equipped with a 48 MP quad camera. Photos of food were also obtained from various sources to supplement learning and testing data collection. The data in the form of original images of food were divided into two parts for training and testing, with 75 % as training data and 25 % as test data in the formation of the model. Randomization of the training and test data was performed, considering the representation of each type of data.

Metrics are used in this study to examine various models of the feeding viewing system. The metrics used are accuracy, precision, recall, and F1-score. The food system model also employs the KNN algorithm with $k = 3$, 5, and 7, as well as the CNN. The value of k in the KNN algorithm is calculated based on the amount of existing data and the size of the dimensions formed by the data. The lower the number of k chosen, the more data there is. However, the greater the dimensionality of the data, the greater the number of k that should be chosen. As a result, the simulation is carried out by testing the values of $k = 3$, 5, and 7 to determine the best test.

The accuracy value of the food identification system using the KNN algorithms and CNN is obtained based on the simulation, as shown in Fig. 2. The CNN method produced the highest accuracy value for all types of food. The value of $k = 3$ means that the group is formed by the involvement of three closest neighbours, while $k = 5$ denotes that the group is formed by the participation of the five closest neighbours of the group. Similarly, if $k = 7$, the seven closest neighbours in the data set are used to form the group. The amount of existing data and the size of the dimensions formed by the data are used to determine the value of k. The highest accuracy value in the KNN method varies depending on the type of food. The highest accuracy value in the KNN(3) method is shown in the identification of green bean porridge, rice, and nuggets. While the identification of sausages and steaks had the highest accuracy at KNN(5), The highest accuracy value in the KNN(7) method was demonstrated in identifying food such as saut'eed water spinach, red rice, rendang, and tempeh.

Fig. 1. Examples of sample data



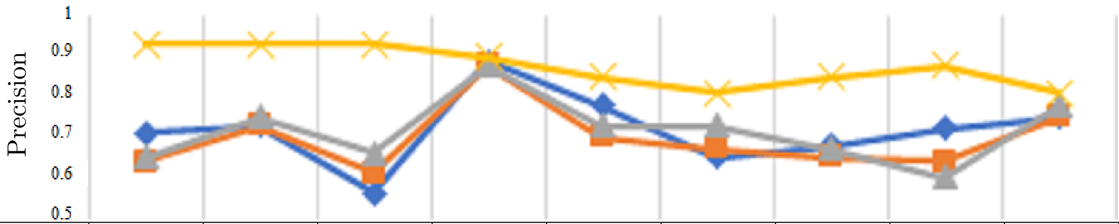
	Green bean porridge	Saut'eed water spinache	Red rice	Rice	Nuggets	Rendang	Sausage	Steak	Tempeh
$k = 3$	0.7	0.73	0.55	0.88	0.78	0.65	0.68	0.7	0.75
$k = 5$	0.62	0.73	0.6	0.85	0.7	0.68	0.65	0.62	0.75
$k = 7$	0.65	0.75	0.65	0.85	0.73	0.73	0.68	0.6	0.78
CNN	0.93	0.93	0.93	0.9	0.85	0.78	0.83	0.88	0.8

Fig. 2. The accuracy value of food identification using the KNN and CNN methods

According to Fig. 3, the CNN method has a higher precision value than the KNN method for all of the foods tested. The CNN method produces an average precision value of more than 87 %. While using the KNN method, the resulting precision value varies depending on the type of food and the value of k. When identifying green bean porridge, rice, nuggets, sausage, and steak, the KNN method $k = 3$ has the highest precision value among the KNN methods. Other KNN methods with the highest precision values were obtained at $k = 7$ when identifying saut'eed water spinach, red rice, rendang, and tempeh.

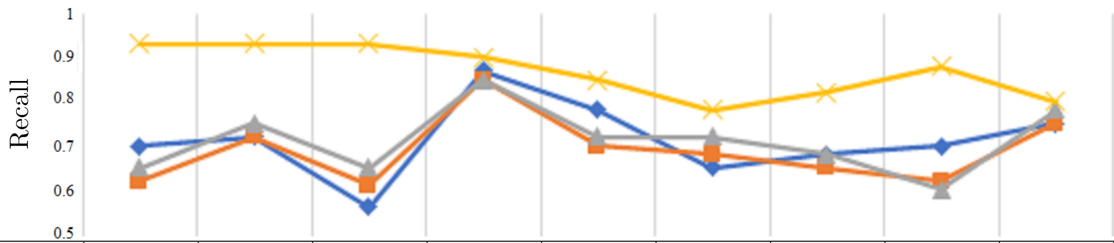
According to Fig. 4, the CNN method also produces the highest recall value of all types of food, with an average recall value of more than 86 %. While the KNN method produces the highest recall value, which varies depending on the type of food. In the identification system of green bean porridge, rice, sausage, and steak, the KNN algorithms with $k = 3$ produces the highest recall value. The KNN(7) method produced the highest recall value for the saut'eed water spinach, red rice, rendang, sausage, and tempeh types.

According to Fig. 5, the CNN method continues to provide the highest metric value. The F1-score of the CNN method is higher than that of the KNN method, with an average of more than 86 %. When comparing KNN algorithms, the highest F1-score produced will vary depending on the type of food. When identifying green bean porridge, rice, nuggets, sausage, and steak, the highest F1-score, $k = 3$ was obtained. While the KNN algorithms with $k = 7$ produced the best F1-score for identifying saut'eed water spinach, red rice, rendang, sausage, and tempeh types.



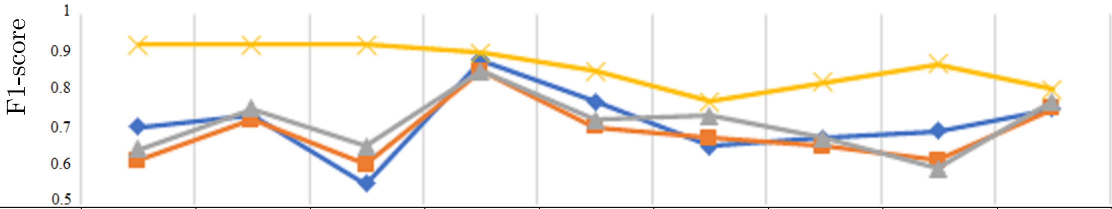
	Green bean porridge	Saut'eed water spinache	Red rice	Rice	Nuggets	Rendang	Sausage	Steak	Tempeh
$k = 3$	0.71	0.73	0.56	0.89	0.78	0.65	0.68	0.72	0.75
$k = 5$	0.64	0.73	0.61	0.88	0.7	0.67	0.65	0.64	0.75
$k = 7$	0.65	0.75	0.66	0.88	0.73	0.73	0.67	0.6	0.78
CNN	0.93	0.93	0.93	0.9	0.85	0.81	0.85	0.88	0.81

Fig. 3. The precision value of food identification using the KNN and CNN methods



	Green bean porridge	Saut'eed water spinache	Red rice	Rice	Nuggets	Rendang	Sausage	Steak	Tempeh
$k = 3$	0.7	0.72	0.56	0.87	0.78	0.65	0.68	0.7	0.75
$k = 5$	0.62	0.72	0.61	0.85	0.7	0.68	0.65	0.62	0.75
$k = 7$	0.65	0.75	0.65	0.85	0.73	0.72	0.68	0.6	0.78
CNN	0.93	0.93	0.93	0.9	0.85	0.78	0.82	0.88	0.8

Fig. 4. The recall value of food identification using the KNN and CNN methods



	Green bean porridge	Saut'eed water spinache	Red rice	Rice	Nuggets	Rendang	Sausage	Steak	Tempeh
$k = 3$	0.7	0.73	0.55	0.88	0.77	0.65	0.67	0.69	0.75
$k = 5$	0.62	0.72	0.6	0.85	0.7	0.67	0.65	0.61	0.75
$k = 7$	0.64	0.75	0.65	0.85	0.72	0.73	0.67	0.59	0.77
CNN	0.92	0.92	0.92	0.9	0.85	0.77	0.82	0.87	0.8

Fig. 5. The F1-score value of food identification using the KNN and CNN methods

Discussion. The CNN approach provides a better results than the KNN algorithms, based on the results of food classification stated in the results section. The CNN method has a higher F1-score than the KNN algorithms, as evidenced by accuracy, precision, and recall metrics. This suggests that the CNN approach outperforms the KNN algorithms in detecting nine different types of food. This is in line with the research results of [29], which state that KNN and CNN appears competitive with their respective algorithms.

The CNN technique is a high-complexity artificial neural network method with many layers capable of modelling a considerably greater function. As a result, the CNN approach can create data with great accuracy. CNN, however, necessitates a vast amount of data and a significant amount of time to train. CNN's are made up of numerous layers, such as convolution layers, pooling layers, and fully connected layers, and are designed to learn the spatial hierarchies of features automatically and adaptively [30]. To reduce the number of parameters and complexity, CNN uses geographical information that other algorithms do not have. As a result of these factors, CNN gives better estimates of model quality metrics than KNN.

The KNN approach with $k = 3$ can produce better model metric values such as accuracy, precision, recall, and F1-score for numerous types of food such as green bean porridge, rice, nuggets, sausage, and steak, as shown in the findings. Meanwhile, the KNN(7) approaches may correctly identify food, such as saut'eed water spinach, red rice, sausage, and tempeh. When viewed from the original image, the sorts of food that can be detected well by the KNN(3) are photos of food with a lighter brightness/color level. Meanwhile, an image of food with a darker color can be identified using [A1] the KNN(7) algorithms.

KNN is one method that accomplishes categorization based on training or learning data viewed from the object's closest distance using the k value [31]. The value of k has a significant impact on the level of classification accuracy when employing the KNN algorithm. The value of k represents the number of neighbours or data nearest to an object. The data determine the best k value in KNN. In general, a high value of k reduces the impact of noise on classification but blurs the distinctions between classifications. The classification findings of one object will most likely be influenced by the number of various neighbours [32].

Based on research [33], the number of k should be ideally an odd number, such as $k = 1$, 2, 3, and so on. A simulation of the KNN method was carried out in this study using a value of $k = 3, 5, 7$. Furthermore, according to [27], the value of k is determined empirically (trial and error), and the value of k that yields the lowest error rate can be chosen, considering the amount of data available and the size of the dimensions created by the data. The smaller the number of k picked, the more data there is.

Conclusion

The best food identification model is built with the CNN method based on the simulation results. The CNN performs better than the KNN algorithm. The CNN method outperforms the KNN algorithm on the accuracy metric by 13 % on average. Similarly, regarding precision metrics, CNN outperforms KNN by 15.7 % on average. While the average F1-score and recall increase, the CNN outperforms the KNN algorithm by 12.89 and 13.22 %, respectively.

Recommendation

Although the identification model generated by the CNN method is quite good, there are still system errors when processing the food images, especially if the food is nearly the same color. As a result, in addition to color features, other features such as shape, and texture. must be added so that the identification system can produce better model goodness metric values.

Acknowledgements. We thank Universitas Negeri Medan for providing funds to implement this research. The rector, LPPM, the dean of FMIPA Universitas Negeri Medan, and the head of the department have provided support and facilities for the implementation of this research series, especially the rector, LPPM. Although not perfect, hopefully, the results of this research will be helpful in the development of knowledge useful for people.

References

[1] **Ashakiran D.R., Deepthi R.** Fast foods and their impact on health. Journal of Krishna Institute of Medical Sciences University. 2012; 1(2):7–15.

[2] **Bahadoran Z., Mirmiran P., Azizi F.** Fast food pattern and cardiometabolic disorders: a review of current studies. Health Promotion Perspectives. 2015; 5(4):231. DOI:10.15171/hpp.2015.028.

[3] **Xiao L., Lan T., Xu D., Gao W., Li C.** A simplified CNNs visual perception learning network algorithm for foods recognition. Computers & Electrical Engineering. 2021; 92:107152. DOI:10.1016/j.compeleceng.2021.107152.

[4] **Shimoda W., Yanai K.** CNN-based food image segmentation without pixel-wise annotation. International Conference on Image Analysis and Processing. Springer, Cham.: 2015; 449–457. DOI:10.1007/978-3-319-23222-5_55.

[5] **Sugiarth I.G., Sudarma M., Widyantara I.M.** Ekstraksi fitur warna, tekstur dan bentuk untuk clustered-based retrieval of images (CLUE). Majalah Ilmiah Teknologi Elektro. 2016; 16(1):85. DOI:10.24843/mite.1601.12.

[6] **Ibraheem C.M., Reddy G.U.** Content based image retrieval using HSV color, shape and GLCM texture. International Journal of Advanced Research in Computer and Communication Engineering. 2015; 4(10):378–383. DOI:10.17148/IJARCCE.2015.41082.

[7] **Anton A., Nissa N.F., Janiati A., Cahya N., Astuti P.** Application of deep learning using convolutional neural network (CNN) method for womens skin classification. Scientific Journal of Informatics. 2021; 8(1):144–153. DOI:10.15294/sji.v8i1.26888.

[8] **Indolia S., Goswami A.K., Mishra S.P., Asopa P.** Conceptual understanding of convolutional neural network — a deep learning approach. Procedia Computer Science. 2018; 132:679–688. DOI:10.1016/j.procs.2018.05.069.

[9] **Imandoust S.B., Bolandraftar M.** Application of k-nearest neighbor (knn) approach for predicting economic events: theoretical background. International Journal of Engineering Research and Applications. 2013; 3(5):605–610.

[10] **Syaliman K.U., Nababan E.B., Sitompul O.S.** Improving the accuracy of k-nearest neighbor using local mean based and distance weight. Journal of Physics: Conference Series. 2018; 978(1):012047. DOI:10.1088/17426596/978/1/012047.

[11] **Gou J., Xiong T., Kuang Y.** A novel weighted voting for k-nearest neighbor rule. Journal of Computers. 2011; 6(5):833–840. DOI:10.4304/jcp.6.5.833-840.

[12] **Bhatia N.** Survey of nearest neighbor techniques. 2010. Available at: `http://arxiv.org/abs/1007.0085`.

[13] **Pathanjali C., Salis V.E., Jalaja G., Latha A.** A comparative study of Indian food image classification using K-nearest-neighbour and support-vector-machines. International Journal of Engineering & Technology. 2018; 7(3.12):521–525. DOI:10.14419/ijet.v7i3.12.16171.

[14] **Kamilaris A., Prenafeta-Boldú F.X.** A review of the use of convolutional neural networks in agriculture. The Journal of Agricultural Science. 2018; 156(3):312–322. DOI:10.1017/S0021859618000436.

[15] **Alzubaidi L., Zhang J., Humaidi A.J., Al-Dujaili A., Duan Y., Al-Shamma O., Santamaría J., Fadhel M.A., Al-Amidie M., Farhan L.** Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. Journal of Big Data. 2021; 8(1):1–74.

[16] **Gupta V., Sachdeva S., Dohare N.** Deep similarity learning for disease prediction. Trends in Deep Learning Methodologies. 2021: 183–206.

[17] **Karyono G.** Analisis teknik data miningalgoritma c4. 5 dan k-nearest neighboruntuk mendiagnosa penyakit diabetes mellitus. STMIKPoliteknik PalComTech. 2016; 12. Available at: `http://news.palcomtech.com/wp-content/uploads/downloads/2016/06/IT13GiatKaryono.pdf`.

[18] **Vapnik V.N.** Statistics the elements of statistical learning. The Mathematical Intelligencer. 2009; 27(2):83–85. Available at: `http://www.springerlink.com/index/D7X7KX6772HQ2135.pdf`.

[19] **Wu X., Fu X., Liu Y., Lim E.P., Hoi S.C., Sun Q.** A large-scale benchmark for food image segmentation. Proceedings of the 29th ACM International Conference on Multimedia. 2021; 506–515.

[20] **He Y., Xu Ch., Khanna N., Boushey C.J., Delp E.J.** Analysis of food images: features and classification. 2014 IEEE International Conference on Image Processing (ICIP). IEEE; 2014; 2744–2748. DOI:10.1109/ICIP.2014.7025555. Available at: `https://ieeexplore.ieee.org/document/7025555`.

[21] **Gonzalez R.C.** Digital image processing. Pearson Education India; 2009.

[22] **Agaputra M.D., Wardani K.R., Siswanto E.** Pencarian citra digital berbasiskan konten dengan ekstraksi fitur HSV, ACD, dan GLCM. Jurnal Telematika. 2013; 8(2):8. Available at: `https://journal.ithb.ac.id/telematika/article/view/73`.

[23] **Kavitha C., Rao B.P., Govardhan A.** Image retrieval based on color and texture features of the image sub-blocks. International Journal of Computer Applications. 2011; 15(7):33–37. DOI:10.5120/1958-2619.

[24] **Hema D., Kannan D.S.** Interactive color image segmentation using HSV color space. Science and Technology Journal. 2020; 7(1):37-41. DOI:10.22232/stj.2019.07.01.05.

[25] **Handayani I.** Application of k-nearest neighbor algorithm on classification of Disk Hernia and Spondylolisthesis in Vertebral Column. Indonesian Journal of Information Systems. 2019; 2(1):57–66. DOI:10.24002/ijis.v2i1.2352.

[26] **Putra W.S.** Klasifikasi citra menggunakan convolutional neural network (CNN) pada Caltech 101. Jurnal Teknik ITS. 2016; 5(1). DOI:10.12962/j23373539.v5i1.15696. Available at: https://ejurnal.its.ac.id/index.php/teknik/article/view/15696.

[27] **Han J., Pei J., Kamber M.** Data mining: concepts and techniques. Elsevier; 2012.

[28] **Bramer M.** Data for data mining. Principles of Data Mining. London: Springer; 2016: 9–19.

[29] **Makkar T., Kumar Y., Dubey A.K., Rocha A., Goyal A.** Analogizing time complexity of KNN and CNN in recognizing handwritten digits. 2017 Fourth International Conference on Image Information Processing (ICIIP). IEEE; 2017; 1–6. DOI:10.1109/ICIIP.2017.8313707. Available at: https://ieeexplore.ieee.org/document/8313707.

[30] **Yamashita R., Nishio M., Do R.K., Togashi K.** Convolutional neural networks: an overview and application in radiology. Insights Into Imaging. 2018; 9(4):611–629.

[31] **Setianto Y.A., Kusrini K., Henderi H.** Penerapan algoritma k-nearest neighbour dalam menentukan pembinaan koperasi kabupaten kotawaringin timur. Creative Information Technology Journal. 2019; 5(3):232–241. DOI:10.24076/citec.2018v5i3.179.

[32] **Angreni I.A., Adisasmita S.A., Ramli M.I., Hamid S.** Pengaruh nilai k pada metode k-nearest neighbor (KNN) terhadap tingkat akurasi identifikasi Kerusakan Jalan. Rekayasa Sipil. 2018; 7(2):63–70. DOI:10.22441/jrs.2018.v07.i2.01. Available at: https://publikasi.mercubuana.ac.id/index.php/jrs/article/view/jrs.2018.v7.i2.01.

[33] **Thirunavukkarasu K., Singh A.S., Irfan M., Chowdhury A.** Prediction of liver disease using classification algorithms. 2018 4th International Conference on Computing, Communication and Automation (ICCCA). IEEE; 2018; 1–3. DOI:10.1109/CCAA.2018.8777655. Available at: https://ieeexplore.ieee.org/document/8777655/authors#authors.

[34] **Nafi Dzikrulloh N., Indriati B.D.** Penerapan metode k-nearest neighbor (KNN) dan metode weighted product (WP) dalam penerimaan calon guru dan karyawan tata usaha baru berwawasan teknologi (studi kasus: sekolah menengah kejuruan muhammadiyah 2 kediri). Journal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN. 2017; (2548):964X. Available at: http://repository.ub.ac.id/id/eprint/147404.

Вычислительные технологии, 2022, том 27, № 6, с. 88–99. © ФИЦ ИВТ, 2022 ISSN 1560-7534
Computational Technologies, 2022, vol. 27, no. 6, pp. 88–99. © FRC ICT, 2022 eISSN 2313-691X

ВЫЧИСЛИТЕЛЬНЫЕ ТЕХНОЛОГИИ

DOI:10.25743/ICT.2022.27.6.008

Сравнение применения сверточной нейронной сети и алгоритма *k*-ближайших соседей для распознавания продуктов питания на изображениях

Арнита[1,*], М. Яни[2], Ф. Марпаунг[1], М. Хидаят[1], А. Видианто[1]

[1]Департамент математики Университета Негери Медан, 20221, Медан, Северная Суматра, Индонезия

[2]Департамент машиностроения Университета Мухаммадия Суматра Утара, 20238, Медан, Северная Суматра, Индонезия

*Контактный автор: Арнита, e-mail: arnita@unimed.ac.id

Поступила 22 июня 2022 г., доработана 20 сентября 2022 г. принята в печать 27 сентября 2022 г.

Аннотация

Важную роль в повседневной жизни играет еда, и осведомленность населения о качестве продуктов питания повысилась. Доступность многих видов продуктов питания затрудняет выбор правильного типа здоровой пищи для потребления. Алгоритмы сверточной нейронной сети

(CNN) и k-ближайших соседей (KNN) можно использовать для создания моделей классификации и идентификации, включая идентификацию пищевых продуктов. Поэтому для поддержания здорового питания нужна система, которая может быстро определить тип потребляемой пищи и рассчитать ее калорийность. Необходимо создать наилучшую модель идентификации на основе показателей качества модели для точности, предсказания, отзыва и оценки F1, которые будут использоваться для идентификации пищевых продуктов с использованием алгоритмов CNN и KNN. Этот метод исследования извлекает входные данные изображения еды с использованием модели HSV (тон, насыщенность и значение цвета). Данные классифицируются с использованием алгоритмов CNN и KNN. Моделирование выполняется с использованием 900 изображений продуктов питания. Данные разделены на две категории, а именно обучающая и тестовая выборки, в пропорции 75 и 25 % соответственно. Алгоритм KNN тестировался с $k = 3$, 5 и 7 и сравнивался с CNN. На основании проведенных экспериментов установлено, что метод CNN лучше, чем алгоритм KNN. Есть два класса типов продуктов питания, прогноз по которым неверен, в то время как метод CNN предсказывает только один класс продуктов питания как неправильный. На это указывает точность метода CNN, которая на 5 % лучше, чем метода KNN(3). Точность метода CNN составляет 94 %, а метода KNN(3) — 89 %. Значение F1-оценки для метода CNN равно 0.94, а для метода KNN(3) — 0.89. CNN позволяет модели давать среднюю точность 87.7 %, точность 86.89 %, полноту (recall) 86.89 % и F1 86.33 %. По результатам исследования модель, сформированная с использованием CNN, является лучшей моделью идентификации пищевых продуктов.

Ключевые слова: распознавание еды на изображении, сверточная нейронная сеть, k-ближайших соседей, модель HSV.

Цитирование: Арнита, Яни М., Марпаунг Ф., Хидаят М., Видианто А. Сравнение применения сверточной нейронной сети и алгоритма $\boldsymbol{k}$-ближайших соседей для распознавания продуктов питания на изображениях. Вычислительные технологии. 2022; 27(6):88–99. DOI:10.25743/ICT.2022.27.6.008. (на английском)

Благодарности. Мы благодарим Университет Негери Медан за предоставление финансирования для реализации этого исследования. Ректора LPPM, декана FMIPA Университета Негери Медан и руководителя департамента за оказанную поддержку в реализации этой серии исследований, особенно ректора LPPM. Хотя исследование не идеально, надеемся, результаты помогут в развитии знаний, полезных для людей.