

Image Classification of Beef and Pork Using Convolutional Neural Network in Keras Framework

Salsabila^a, Anwar Fitrianto^{a*}, Bagus Sartono^a

^aDepartment of Statistics, IPB University, Bogor, Indonesia

ABSTRACT

Beef is a food ingredient that has a high selling value. Such high prices make some people manipulate sales in markets or other shopping venues, such as mixing beef and pork. The difference between pork and beef is actually from the color and texture of the meat. However, many people do not understand these differences yet. In addition to socialization related to understanding the differences between the two types of meat, another solution is to create a technology that can recognize and differentiate pork and beef. That is what underlies this research to build a system that can classify the two types of meat. Convolutional Neural Network (CNN) is one of the Deep Learning methods and the development of Artificial Intelligence science that can be applied to classify images. Several regularization techniques include Dropout, L2, and Max-Norm were applied to the model and compared to obtain the best classification results and may predict new data accurately. It has known that the highest accuracy of 97.56% obtained from the CNN model by applying the Dropout technique using 0.7 supported by hyperparameters such as Adam's optimizer, 128 neurons in the fully connected layer, ReLu activation function, and 3 fully connected layers. The reason that also underlies the selection of the model is the low error rate of the model, which is only 0.111.

Keywords: Beef and Pork, Model, Classification, CNN

Article History

Received 07 February 21

Received in revised form 30 March 21

Accepted 23 June 21

1. Introduction

The high selling price of beef has resulted in the presence of some seller to cheat by mixing pork that has a lower selling price with beef. This issue has been going on for several years in Indonesia. Hence, the government must consider the importance of safety and product quality assurance in the market. Food safety and assurance consist of several aspects, such as health, hygiene, and halal labelling [1]. Food products containing pork are strictly prohibited for consumption in countries with a majority Muslim and Jewish population [2].

The difference between pork and beef can be seen mainly in the color and texture. However, there are still many people who do not understand these differences. Nowadays, one of many solutions is to use a technique that might directly identify and distinguish the two types of meat. Image classification is one of the techniques in Deep Learning that can be used in distinguishing the two types of meat.

In the development of the image classification method, Convolutional Neural Network (CNN) has been used by researchers as a deep learning classification method with good performance compared to traditional machine learning methods. Several studies applied the traditional machine learning methods in classifying types of meat have been conducted such as image classification for types of goat, buffalo, cow, and horse meat types using SVM [3], and image classification for types of beef, goat, and pork using KNN [4]. However, as far as the authors knowledge, there is no research on using CNN for meat

classification.

One of the framework developed by Google for image classification using CNN is Keras. The advantages of Keras such as the tools are sufficient to build various deep learning algorithm, easy to understand, and work well on simple model construction. The development of CNN began with LeNet architecture built by LeCun [5] in the classification and handwriting recognition of bank accounts. LeNet network is a very simple deep learning network.

The Deep Learning uses experimentally principles in finding the best hyperparameter in various cases. Hence, LeNet architecture may also be inappropriate in certain cases which results in poor performance of the model. The model is also vulnerable to overfitting, that is, the model was able to work well on training data but not for test data [6]. There are several methods of handling overfitting to produce more optimal model performance. Hence, this study aims to find out the best image classification model of pork and beef images by comparing several regularization methods to overcome the symptoms of overfitting.

2. Literature Reviews

2.1 Convolutional Neural Network (CNN)

CNN is an implementation of a more special Artificial Neural Network (ANN) and is considered the best method for image recognition cases. CNN architecture consists of three layers, namely, the convolution layer,

* Corresponding author.

E-mail address: Anwar_stat@gmail.com.

the pooling layer, and the fully-connected layer [7]. The following is an architecture of CNN that uses two convolution and pooling layers, and two fully connected layers.

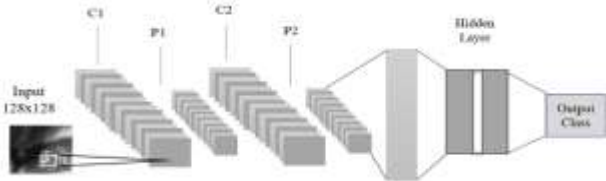


Figure 1. CNN Architecture

There are two main components of the CNN method, feature extraction and classifier. The feature extraction is at the convolution layer and the pooling layer, while the classifier is at the fully connected layer. Specific features of an object can be recognized in the feature extraction [8], while the classifier is used for model learning and find the corresponding label for every test image [9].

Every classification case in machine learning is vulnerable to overfitting issues. The implementation of the final CNN model only fits with train data so that it is unable to predict new data. The following is an illustration of overfitting which is showed in the loss classification metric plot.

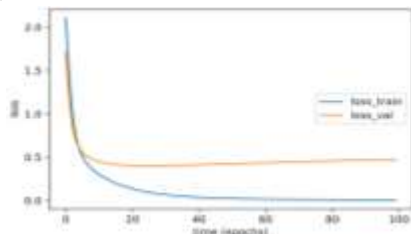


Figure 2. Illustration of Overfitting Shown in the Loss Classification Metric Plot [6]

Figure 2 shows that the loss from the train data decreases and then stabilizes, but the loss from the test data looks consistent away from the loss of train data. Furthermore, the regularization method is usually used in overfit model. There are several regularization methods, namely, dropout, L2, and max-norm.

2.2. Regularization Method

Regularization is one way of obtaining a robust model of the overfitting problem, in principle that limits the complexity of the parameters of the model and reduces the sensitivity of the effect on the training data. The most important concept of regularization is to find only useful features or characteristics and eliminate useless features from a network [10].

1. Dropout

The dropout technique can deal with overfitting [11] and is more effective for networks that use ReLU in the hidden layer. Dropout refers to eliminating neurons that are hidden or visible layers in the network. Each neuron will be given a p probability that is worth between 0 and 1. Dropout values in MLP hidden layers typically range from 0.5 to 0.8 [11].

2. L2

As the L2 regularization of weights, there is a penalty term in the loss function $\|w\|^2$ and λ is a value determined by the researcher to see how much regularization has occurred and can improve the loss function

without regularization. The following is an equation of the new loss function E .

$$E = E + \lambda \|w\|^2$$

The main feature of L2 regularization is that makes the weights evenly distributed, so the weight vector is smooth [12].

3. Max-norm

One of the main forms of regularization is by limiting the vector of weights that enter into each layer hidden by constant c . The two previous regularizations are applied by adding a decayed form to the loss function. However, max-norm assigns the role by limiting the weights $|w|$ to always be below a constant value c .

3. Materials and Methods

2.1 Dataset

This study uses image data of beef and pork. The data were obtained from direct observation by taking pictures of the meat using a smartphone camera. The required data is the image of each type of meat (i.e beef and pork). Both types of meat were obtained from a traditional market in Bogor, the Surya Kencana Bogor market. As much as 90% of the data from both types of meat is the result of the shooting of 250 grams of pork and fresh beef tenderloin, which is divided into several pieces.

From several pieces of meat, repeated shots were taken from different angles. While the other 10% of data comes from shooting for meat on the ribs and thighs without separating them into several pieces, but images were still taken repeatedly from different angles.

From the shooting, 3000 image data were obtained for the total of both types of meat. Based on data, the researcher used 15 % of the total as test data. therefore, as many as 2550 data are training data, and 450 other data as test data.

2.2 Analysis Procedure

The data analysis steps carried out in this study are as follows:

1. Preprocessing by resizing the image to 128x128 pixels and changing the color channel to greyscale with the following calculation.

$$G_{Average} = \frac{R+G+B}{3} \quad (1)$$

R, G, and B are the values for each component of the Red, Green, and Blue colors in the image.

2. Partitioning data: 85 % train data; 15 % test data.
3. In the CNN model training process, the following is the process:
 - a. Convolution, i.e multiplication between input matrix and filter kernel matrix. The following is the equation in calculating the value of the feature map resulting from the convolution process if given the image $\in \mathbb{R}^{W \times H}$, the convolution of the $f \in \mathbb{R}^{P \times Q}$ kernel.

$$(X * f)(i, j) = \sum_{m=0}^{P-1} \sum_{n=0}^{Q-1} X(i+m, j+n) f(m, n), \quad (2)$$

$$i = 0, \dots, H-1, j = 0, \dots, W-1$$

Where :

$(X * f)(i, j)$ = feature map
 X = input matrix

f = kernel matrix
 m, n = kernel size
 W, H = Width, Height (input size)

- b. Activating the results of the convolution layer by applying the non-linear function ReLu to the values in the feature map.

$$g(x) = \max(0, x) \quad (3)$$

If the input is negative, the output of the neuron is expressed as 0. Whereas if the input is positive, then the output of the neuron is the input value itself.

- c. Max-pooling to reduce input size.
- d. Flattening. Get the last max-pooling result into one vector.
- e. Apply backpropagation algorithm learning on the fully connected layer to get the best neurons that will be used in the classifier layer.
- f. Apply the sigmoid function to get a probability of the classification result.

$$p(y = 1|x; w) = g(\mathbf{w}\mathbf{x}^T) = \frac{1}{1 + e^{-\mathbf{w}\mathbf{x}^T}} \quad (4)$$

It is known that y is the class, x as the input image, w as weight, and g is the activation symbol of the function.

4. Build a CNN baseline model by inserting the required hyperparameter such as array batch size, optimizer type, number of iterations or epoch, initialization type, kernel filter type, and number of neurons in the fully connected layer.
5. Identifying overfitting.
6. Compare several models with regularization methods that can overcome the overfitting of the models.
 - a. The model uses Dropout regularization
 - b. The model uses L2 regularization
 - c. The model uses Dropout and L2 regularization
 - d. The model uses Max-Norm regularization
 - e. The model uses Dropout and Max-Norm regularization
 - f. The model uses L2 and Max-Norm regularization
 - g. The model uses Dropout, L2, and Max-Norm regularization

7. Evaluate the classification results of the baseline model and models at the sixth point.
 A model with high accuracy and small error will be selected as the image classification model of pork and beef. The accuracy, F1-Score, and AUC scores of each model will be compared to get the best results.

4. Result and Discussion

4.1 Baseline Model

The baseline model in this study is a classification model that only uses CNN's main architecture. The following are the classification results using the Baseline model.

Table 1 Confusion matrix of image classification of pork and beef using baseline model

Prediction	Reference	
	Pork	Beef
Pork	217	3
Beef	8	222

Table 1 shows the confusion matrix from the results of the

classification of pork and beef images using the CNN Baseline model. The value of 217 represents the number of True Positives (TP) which is the amount of data for a positive class namely '0' or pork and is correctly classified as pork. In contrast, the value of 222 represents the number of True Negatives (TN) which is the amount of data for the negative class namely '1' or beef and is correctly classified as beef. Meanwhile, the values 3 and 8 show the number of False Positive and False Negative, respectively.

Figure 3 shows the presence of overfitting symptoms characterized by the greater the epoch the greater the distance between plot loss in the training data and the test data. This may result in the constructed classification model being unable to generalize test data. Overall, the loss value generated by this model is 0.216.

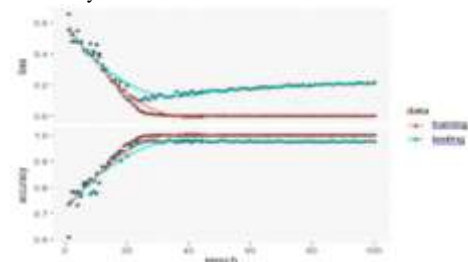


Figure 3. Plot Loss and Accuracy of Baseline Model

4.2 Summary of Classification Model Performance

This study compared several regularization methods consisting of several values in each method. Based on the results in Table 2, several models can be the best model selection candidates, namely the dropout model ($p = 0.5$) and the dropout model ($p = 0.7$). The following is a summary of model performance based on several model comparisons that have been done.

Table 2 Model performance based on comparison of regularization in the models

Regularization			Model Goodness Measure			
Dropout (p)	L2 (λ)	MaxNorm (c)	Accuracy	Loss	AUC	F1-Score
0.5			0.978	0.126	0.996	0.978
0.6			0.973	0.134	0.995	0.974
0.7			0.976	0.111	0.996	0.975
0.8			0.971	0.112	0.996	0.971
	0.001		0.969	0.164	0.989	0.968
	0.005		0.976	0.141	0.993	0.975
	0.01		0.942	0.257	0.989	0.939
	0.05		0.951	0.270	0.987	0.951
		3	0.973	0.136	0.994	0.973
		4	0.971	0.160	0.995	0.971
0.6	0.001		0.967	0.189	0.993	0.966
0.6	0.005		0.964	0.219	0.994	0.965
0.7	0.001		0.960	0.209	0.995	0.959
0.7	0.005		0.962	0.225	0.996	0.961
0.5		3	0.960	0.960	0.991	0.959
0.6		3	0.958	0.958	0.992	0.957
0.7		3	0.971	0.190	0.994	0.971
0.8		3	0.951	0.282	0.987	0.950
	0.001	3	0.971	0.142	0.994	0.972
	0.005	3	0.978	0.138	0.992	0.978
0.7	0.005	3	0.962	0.228	0.992	0.992

Table 2 shows that model dropout ($p = 0.7$) has the lowest error rate. However, the dropout model ($p = 0.5$) produced higher accuracy and F1-Score than the dropout model ($p = 0.7$). Based on Figure 4, the dropout model ($p = 0.7$) shows a smaller gap between training and test losses than the dropout model ($p = 0.5$).

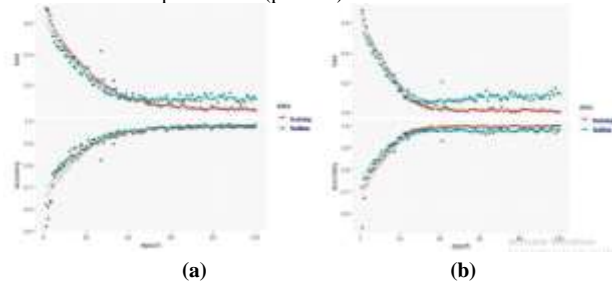


Figure 4. Classification Metric Plot, (a) Dropout Model ($p = 0.7$); (b) Dropout Model ($p = 0.5$)

The lowest error rate and ideal classification metric are the main factors to be achieved in the image classification modeling of pork and beef in this study. This fact is because the model will be able to generalize to new data. Therefore, in this study, the dropout model ($p = 0.7$) is the best way to classify pork and beef images. The following confusion matrix for the classification result of pork and beef images using the dropout model ($p = 0.7$).

Table 3 Confusion matrix of image classification of pork and beef using dropout model ($p = 0.7$)

Prediction	Reference	
	Pork	Beef
Pork	218	4
Beef	7	221

Table 3 shows that the number of images classified correctly by their class 'pork' and 'beef' is 218 and 221 images, respectively. The model performance can be seen in the 97.6% accuracy obtained from the dropout model ($p = 0.7$). At the same time, the AUC and F1-Score values of the dropout model ($p = 0.7$) were 99.96% and 97.5%, respectively. Although the accuracy of the dropout model ($p = 0.7$) and the benchmark model is not much different from F1-Score, it turns out that the error rate generated is small, 0.111.

5. Conclusion

The best classification result of pork and beef images were obtained from a model that used dropout regularization with $p = 0.7$ because it has the best model goodness measure and was able to overcome the overfitting symptoms that occurred in the baseline model. The CNN structure used is two convolution layers, three fully connected layers with hyperparameters such as Adam optimizer, Glorot Uniform Initializer, ReLu activation function, Max-pooling, and Binary Cross Entropy. Based on the model, from 450 test data, it is known that 218 images of pork were correctly classified and 221 images of beef were correctly classified.

REFERENCE

- [1] Y. Erwanto, A. Mohammad Z, E.Y.P, Muslim Sugiyono, and A. Rohman, "Identification of Pork Contamination in Meatballs of Indonesia Local Market Using Polymerase Chain Reaction-Restriction Fragment Length Polymorphism Analysis," *Asian Australasian Journal of Animal Science*, vol 27, no.10, 2014, pp. 1487-1492.
- [2] S. Rahmati, N.M. Julkapli, W.A. Yehye, and W.J Basirun, "Identification of Meat Origin in Food Products—A Review," *Food Control*, vol. 68, 2016, pp. 379-390.
- [3] Neneng and Y. Fernando, "Klasifikasi Jenis Daging Berdasarkan Analisis Citra Tekstur Gray Level Co-Occurrence Matrices dan Warna," *Semnastek* 2017.
- [4] N. Lihayati, R.E Pawening, and M. Furqon, "Klasifikasi Jenis Daging Berdasarkan Tekstur Menggunakan Metode Gray Level Coocurent Matrix," in *Proc. SENTIA 2016*, vol. 8 – ISSN: 2085-2347.
- [5] Y. LeCun, K. Kavukcuoglu, and C. Farabet, "Convolutional networks and applications in vision," *ISCAS*, 253–256, 2010.
- [6] S. Salman, and X. Liu, "Overfitting mechanism and avoidance in deep neural networks," *arXiv preprint arXiv:1901.06566*, 2019.
- [7] M. S. Junayed *et al.*, "AcneNet - A Deep CNN Based Classification Approach for Acne Classes," in *2019 12th Int. Conf. on Information & Comm. Tech.and System*, Surabaya, Indonesia, 2019, pp. 203-208.
- [8] J. Nagi, F. Ducatelle, G. A. Di Caro, D. Cireşan, U. Meier, A. Giusti, F. Nagi, J. Schmidhuber, and L. M. Gambardella, "Max-pooling convolutional neural networks for vision-based hand gesture recognition Signal and Image Processing Applications," presented at the 2011 IEEE International Conference, 2011, pp. 342-347.
- [9] M. Alaslani, and L. Elrefaici, "Convolutional Neural Network Based Feature Extraction for Iris Recognition," *Int. J. of Computer Science & Information Technology*, vol. 10, no.2, April 2018, pp. 65-78.
- [10] Ying, "An Overview of Overfitting and its Solutions," *Journal of Physics*, vol.1168, no.2, 2019, p.022022. doi: 10.1088/1742-6596/1168/2/022022.
- [11] N. Srivastava, G. Hinton, A. Krizhevsky, and I. Sutskever, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *Journal of Machine Learning Research*, vol 15, no.1, 2014, pp.1929-1958.
- [12] H.H. Aghdam, and E.J. Heravi, "Convolutional Neural Network," in *A Guide to Convolutional Neural Network (A Practical Application to Traffic-Sign Detection and Classification)*, Tarragona, ES, SER: Springer, 2017.