Classification of *Rastrelliger kanagurta* and *Rastrelliger brachysoma* using Convolutional Neural Network (CNN)

K Kurniawan, B B Sedayu, A R Hakim and I M S Erawan
Loka Riset Mekanisasi Pengolahan Hasil Perikanan
Jl Imogiri Barat Km 11.5, Jetis, Bantul, Yogyakarta 55781

koernia26@gmail.com

Abstract. Mackerel is an important commercial caught fish for local fishermen, including *Rastrelliger kanagurta* and *Rastrelliger brachysoma*. However, to distinguish these two species is rather difficult because of their similar appearance. Convolutional Neural Network (CNN) is a deep learning method that can be used to classify images. One of parameters contributing to the level of accuracy is the layers number applied in CNN architecture. This study aims to classify those two species using CNN with a range of two to five convolutional layers architectures i.e CNN1, CNN2, CNN3 and CNN4, respectively. In this study, 434 images were used as a training group with 217 images for each class. The validation group consisted of 21 images for each class and the test group consisted of 19 images. The results showed that the CNN3 provided the best training and validation accuracy of respectively 100% and 92.6%. The lowest value of training loss and validation loss of 0.000057 and 0.49. The accuracy values of the CNN models using different testing images reached 94.7%.

1. Introduction
Fishing of pelagic fish in the Java Sea shows important economic value for Indonesian fisheries. In the 2003-2013 periods, it shared for a 20% from the total capture of national small pelagic fishing [1]. This amount is still persistant along the 2019 in the northern coastal sea of Central Java. The capture fisheries production in Batang Regency (Central Java) reaches more than 10,000 tons of small pelagic fish with an economic value of IDR 22 billion [2]. Likewise, the production of mackerel (*Rastrelliger sp.*) which increased in that periods reached for 300 tons with an economic value of more than IDR 1 billion (BPS, 2020). With the increasing production, it shows that mackerel has high economic value [3]. These fish stock is usually to be used for boiled salted fish (pindang) industry, fish canning and sold fresh.

Two types of mackerel that are commonly caught in the Batang sea are *Rastrelliger kanagurta* and *Rastrelliger brachysoma* [4] which are usually called as “kembung lelaki” and “kembung perempuan” for their local trading names. Both types of fish have the same genus and have almost the same morphology. However, to distinguish these two species would be quite difficult. Distinguishing characteristics are the presence of black spots close to the pectoral fins and a smaller belly width in male mackerel fish [5]. Up to date, the method of distinguishing the two types of species is carried out by manuals or done by experienced persons. This bring about difficulties for consumers and academics who are not familiar with these fishes.

Deep learning with Convolutional Neural Network (CNN) method applied for image classification [6] may be potential to distinguish these mackerel species. CNN is one of the deep learning algorithms
and a development of multilayer perceptron devoted to process data in two-dimensional form in the form of image data or voice data [7]. The difference between CNN and multilayer perceptron are due to the feature extraction layer and fully-connected layer. The feature extraction layer is located at the beginning of the architecture and composed of several layers. Each layer is composed of several neurons that are connected to the previous layer of neurons. The feature extraction layer consists of two layers, namely the convolutional layer and pooling layer. This first receives input of an image file and produces an output of vectors. The second layer is the classification layer consisting of neurons that are fully connected to other layers. This second layer receives input of the vectors from previous layer which is then transformed through a Multi Neural Network with several hidden layers [8].

In image classification, several factors effecting the level of accuracy i.e., image quality, and CNN architecture. Image quality includes the image exposure. While accuracy in CNN architecture is attributed to the number of convolution layers used, the selected optimization function and the selected dense layer function. Moreover, the number of batches and epoch applied during developing the model can also influence the level of accuracy and loss values of the model [9].

Several studies have reported that CNN is able to be used to distinguish several types of fish and resulted in good outcomes, including the study conducted by Prasmatio et al. [10] which distinguish 9 different freshwater ornamental fish species; Fauzi et al. [11] that classify 9 different freshwater fish species; and Pratiwi [12] that obtained high accuracy in distinguishing gourami and tilapia using the CNN method whereby those three fishes show similar morphology. In this study, fish samples were selected from a group of similar species, namely male mackerel (Rastrelliger kanagurta) and female mackerel (Rastrelliger brachysoma). Where the two types of fish have almost the same morphology [3]. Considering that the number of batches and CNN architecture can affect the prediction accuracy level, this study aims to determine the optimal number of batches and CNN architecture to obtain the highest accuracy value from the classification of Rastrelliger kanagurta and Rastrelliger brachysoma.

2. Methodology

Fresh mackerels for the samples were taken from Batang Fish Landing, Central Java. The mackerels were cold-transported using ice and insulated container to Laboratory in Bantul-Yogyakarta, which took about 4 hr transportation. Two types of mackerel used were male mackerel (R. kanagurta) and female mackerel (R. brachysoma). Fish image acquisition was carried out in a designed chamber (25 cm x 20 cm x 40 cm) with white color surface. The lighting was adjust using LED lights equipped with brightness controller. The webcam Logitech Type C922 Pro was used to take the images (see Fig. 1).

Python programming language was used to build the CNN architecture for fish classification with several packages related to deep learning such as TensorFlow library, hard library, matplotlib, open cv, NumPy.
A total of 434 fish images was used for the training dataset, which were divided equally into two classes i.e. male mackerel (*R. kanagurta*) and female mackerel (*R. brachysoma*).

In this study, the CNN input file is an RGB image file with a size of 200 x 200. As for the treatment, four CNN architectures are used. CNN1 is a CNN architecture consisting of two convolution layers with 16 filters in the first layer and 32 in the second layer with the activation function is relu. From each convolution layer will be followed by a 2D MaxPool layer measuring 2x2. CNN2 is a CNN architecture consisting of three convolution layers with 16 filters in the first layer, 32 in the second layer and 64 in the third layer with the activation function is relu.
the third layer with the activation function is RELU. From each convolution layer will be followed by a 2D MaxPool layer measuring 2x2. CNN3 is a CNN architecture consisting of four convolution layers with the number of filters 16 in the first layer, 32 in the second layer, 64 in the third layer and 128 in the fourth layer with the activation function is relu. From each convolution layer will be followed by a 2D MaxPool layer measuring 2x2. CNN4 is a CNN architecture consisting of five convolution layers with the number of filters 16 in the first layer, 32 in the second layer, 64 in the third layer, 128 in the fourth layer and 256 in the fifth layer with the activation function relu. From each convolution layer will be followed by a 2D MaxPool layer measuring 2x2. After constructing the convolution layer of each CNN architecture, then the fully connected layer is built, followed by a fatten layer, a dense layer of 512 neurons with an activation function of RELU and ending with a dense layer of 1 and the activation function is sigmoid. In this study, 35 epochs were used and the learning rate per epoch was 35.

The validation was performed using 21 different images from the training group of each class. The four models obtained were analysed using the values of accuracy and loss training as well as accuracy and loss validation. Furthermore, each model was tested using different images out of the training and validation groups. Nineteen images dataset were used to test the models. Test result were displayed in the confusion matrix table (Table 1).

### Table 1. Confusion Matrix

| Predicted       | Actual   |          |          |
|-----------------|----------|----------|----------|
|                 | R kanagurta | R brachysoma |
| R kanagurta     | TP       | FN       |
| R brachysoma    | FP       | TN       |

\[
\text{accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \tag{1}
\]

\[
\text{sensitivity} = \frac{TP}{TP+FN} \times 100\% \tag{2}
\]

\[
\text{specificity} = \frac{TN}{TN+FP} \times 100\% \tag{3}
\]

The TP, FP, FN, and TN represent the true positive, false positive, false negative, and true negative.

### 3. Results and Discussion

#### 3.1. Effect of batch number

The results showed that variation on batch number used in training of the dataset imparted in the training accuracy value of the CNNs architectures, while the CNN3 showed a constant training accuracy for all the batch number used. Changes in the accuracy value can be seen in Figure 3. Validation accuracy value occurred on CNN1 was fluctuate, whereas from batch 3 to 9 was decreased but in batch 12 was increased. Other CNN architectures experienced a decrease in accuracy in batch 9 and an increase in batch 12 with an inconspicuous value (Table 2).

### Table 2. Accuracy, loss training and validation from various CNN architecture and batch

| Architecture | batch 3 training accuracy | batch 3 training loss | batch 3 validity accuracy | batch 3 validity loss | batch 9 training accuracy | batch 9 training loss | batch 9 validity accuracy | batch 9 validity loss | batch 12 training accuracy | batch 12 training loss | batch 12 validity accuracy | batch 12 validity loss |
|--------------|---------------------------|----------------------|---------------------------|----------------------|---------------------------|--------------------------|---------------------------|--------------------------|---------------------------|--------------------------|---------------------------|--------------------------|
| CNN1         | 0.93                      | 0.224                | 0.65                      | 0.63                 | 0.99                      | 0.083                    | 0.512                     | 3.34                     | 1                         | 0.001                   | 0.9                       | 0.5                      |
| CNN2         | 0.96                      | 0.128                | 0.9024                    | 0.147                | 1                         | 0.000                    | 0.926                     | 0.62                     | 0.98                      | 1.5                      | 0.9                       | 0.47                     |
| CNN3         | 1                         | 0.018                | 0.976                     | 0.06                 | 1                         | 0.000                    | 0.926                     | 0.498                    | 1                         | 0.000                   | 0.97                      | 0.18                     |
| CNN4         | 0.95                      | 0.18                 | 0.97                      | 0.15                 | 1                         | 0.000                    | 0.92                      | 0.98                     | 1                         | 0.000                   | 0.95                      | 0.31                     |
The increase in the accuracy value occurred in batch 9 for CNN1, 2 and 4 architectures which reached a value of 100%, but for the CNN2 architecture there was a decrease in the accuracy value in the 12th batch. From these results indicate that batch 9 is the maximum value to get the best accuracy. In their research, Rokhana et al. [9] stated that the number of 128 batches gave a maximum accuracy value of two classification classes with a total of 4,750 images of training data in each class.

To help understanding the built models for each architecture, it could be done using a confusion matrix table which could be seen in Table 3. From the table, it could be seen that the CNN 3 model only experienced one error prediction, namely one R kanagurta predicted by R brachysoma.

In Figure 4 it can be seen that the loss training value decreases with increasing batches but does not occur in CNN 2. The validation loss has a lower value when compared to the loss training value. This happened as reported by Fauzi et al. [11] which stated that the CNN model made was able to identify fish species with a training loss value of 0.0035 and a validation loss value of 0.445.
Figure 4. Loss training and validation changes across multiple batches and CNN architectures

Table 3. Confusion matrix from various CNN architecture

| Architecture | Predicted     | Actual         | R kanagurta | R brachysoma |
|--------------|---------------|----------------|-------------|--------------|
| CNN1         | R kanagurta   | 12             | 0           |
|              | R brachysoma  | 7              | 0           |
| CNN2         | R kanagurta   | 9              | 3           |
|              | R brachysoma  | 1              | 6           |
| CNN3         | R kanagurta   | 11             | 1           |
|              | R brachysoma  | 0              | 7           |
| CNN4         | R kanagurta   | 12             | 0           |
|              | R brachysoma  | 6              | 1           |

Accuracy, sensitivity and specificity values of several models can be seen in Table 4. From the table it could be seen that the accuracy value increases with the increase in the CNN architecture, but begins to decrease in the use of the CNN 4 architecture. So, the highest accuracy value is in the CNN 3 architecture. Value of sensitivity and specificity, CNN 3 has the best value. Rokhana et al. [9] stated that for binary classification of fractured and non-fractured bones, the CNN architecture that provides high accuracy consists of two sets of convolution layers, where each set consists of two convolution layers followed by the relay activation function.
Table 4. Accuracy, sensitivity and specificity from various CNN architecture

| Architecture | accuracy  | sensitivity | specificity |
|--------------|-----------|-------------|-------------|
| CNN1         | 0.632     | 0.632       | 0.000       |
| CNN2         | 0.789     | 0.900       | 0.667       |
| CNN3         | 0.947     | 1.000       | 0.875       |
| CNN4         | 0.684     | 0.667       | 1.000       |

4. Conclusion
The best number of batches used to classify *R kanagurta* and *R brachysoma* is 9 and the best CNN architecture used is CNN3 which consists of four convolution layers with the number of filters 16 in the first layer, 32 in the second layer, 64 in the third layer and 128 in the third layer. Four with activation function is relu. The values of accuracy, sensitivity and specificity on the CNN 3 architecture are 0.94; 1; 0.875.

Reference
[1] Purwanto 2015 *J. Lit. Perikanan*. Ind. 21 25-36
[2] Triarso I 2012 *Saintek Perikanan: Indonesian Journal of Fisheries Science and Technology* 8 65-73
[3] Utami M N F, Redjeki S and Supriyantini E 2014 *Journal of Marine Research* 2 99-106
[4] Triyono H, Muzakki S A and Mulyoto 2020 *Buletin JSJ* 2 69-81
[5] Nurjanah, Nurilmala M, Hidayat T and Azri R Y I 2016 *International Journal of Materials Chemistry and Physics* 2 54-61
[6] Suyatno 2018 *Machine Learning Tingkat Dasar dan Lanjut* (Bandung: Informatika)
[7] Suartika I W, Wijaya A Y and Soelaiman R 2016 *Jurnal Teknik ITS* 5 A65-A69
[8] Ulum M, Zakariya M, Fiqhi I A and Haryanto 2021 *Jurnal ilmiah Teknik Informatika, Elektronik dan Kontrol* 1 23-30
[9] Rokhana R, Priambodo J, Karlita T, Sunarya I M G, Yuniarno E M, Purnama I K E and Purnomo M H 2019 *Jurnal Nasional Teknik Elektro dan Teknologi Informasi* 8 59-67
[10] Prasmatio R M, Rahmat B and Yuniar I 2020 *Jurnal Informatika dan Sistem Informasi* 1 510-521
[11] Fauzi S, Eosina P and Laxmi G F 2019 *Semin. Nas. Teknol. Inf.* 2 163–167
[12] Pratiwi N F 2020 *Klasifikasi Spesies Ikan Air Tawar Menggunakan Convolutional Neural Network* (Medan: Universitas Sumatera Utara)
[13] Kasinathan T, Singaraju D and Uyyala SR 2020 *Information Processing in Agriculture* 8 446-457

Acknowledgments
We thank Mr. Lutfi Assadad as head of LRMPHP, Mr. Zaenal for the interesting discussion, and Caesar for assisting in image acquisition and all LRMPHP employees for their cooperation.