ILKOM Jurnal Ilmiah Vol. 14, No. 2, August 2022, pp. 169-177
Accredited 2nd by RISTEKBRIN No. 200/M/KPT/2020;E-ISSN 2548-7779 | P-ISSN 2087-1716

Identification of sea urchins in melonguane coastal area using Multilayer Perceptron Neural Network

Andar Alwein Pinilas[a,1]; Luther Alexander Latumakulita[a,2*]; Djoni Hatidja[a,3]
[a] *Universitas Sam Ratulangi, Bahu, Kec. Malalayang , Manado and 95115, Indonesia*
[1] *18101106001@student.unsrat.ac.id;* [2] *latumakulitala@unsrat.ac.id;* [3] *dhatidja@unsrat.ac.id*
* *Corresponding author*

Article history: Received June 03, 2022; Revised June 03, 2022; Accepted July 29, 2022; Available online August 30, 2022

Abstract

Sea urchins (*Echinoidea*) are marine biota found in Indonesian oceans, and there are 950 types of sea urchins scattered throughout the world. This study aims to classify types of sea urchins based on the characteristics contained in sea urchin images using the Multilayer Perceptron Neural Network (MLP-NN) method with three classification classes. 120 of sea urchin image data were taken from Melonguane beach area, Talaud Islands Regency, North Sulawesi Province. In the MLP-NN stage, training, validation, and testing processes are carried out by applying 8-fold cross-validation. The system performance shows the lowest accuracy is 93.33%, the highest is 100%, and the average test data is 86.66%. The experimental results indicate that MLP-NN can classify sea urchins with reasonably good performance.

Keywords: Classification; Sea urchin; Multi-Layer Perceptron

Intoduction

Indonesia has a variety of marine life, and coral reefs which become a place for marine animal species to settle, one of which is *Echinoidea* species belonging to the Echinoderm class called sea urchins. In Melonguane beach area, Talaud Islands Regency, several types of sea urchins can be found around dead coral reefs at low tide. In [1] it mentions that Echinoderms come from the Greek *echin*, which means thorn, and *derma*, which means skin. So, *Echinoderms* are animals with thorny skin or also called thorn-skinned animals.

Sea urchins have various types based on different shapes and sizes. Because of the many differences in shape and size, not everyone can recognize the types of sea urchins. From these problems, an intelligent model can be built to identify the types of sea urchins.

Several studies have been carried out in the Talaud archipelago district related to the topics taken in this study, namely a study entitled Diversity of Echinoderms on Paranti Beach, Tabang Village, Rainis District, Talaud Islands Regency, North Sulawesi Province [2]. This study discusses the analysis of various species of the Echinoderm class on the coast of Paranti.

Several studies also apply the MLP-NN method in this study, including the application of Multi-Layer Perceptron in image annotation automatically [3]. This study discusses a model that will be implemented to predict an annotation of an image. In [4] and [5], use the MLP-NN to identify the eligibility of prospective recipients of the Bidik Misi scholarship. Another research related to the MLP-NN method is the identification system of local reef fish species in Bunaken National Park using the Backpropagation Neural Network method [6]. This study classifies reef fish species based on the characteristics contained in the fish image. The application of MLP-NN was used in [7] to distinguish types of puppet based on puppet images and in [8] as a method for grouping websites.

This study uses MLP-NN to identify sea urchins based on digital images of sea urchins. Build a smart model to identify sea urchin species using the MLP-NN method.

Sea urchins are round marine animals with spines on their skin that can be moved. These animals are divided into about 950 species and can be found in tidal areas to depths of 5,000 meters [9]. Some have benefits such as food, ecological, economic, and toxic properties. Others have been used as model organisms and ornamental animals in the health sector, especially for treating diseases in humans [10].

http://dx.doi.org/10.33096/ilkom.v14i2.1208.169-177

MLP-NN is an algorithm that adopts the workings of nerves in living things. MLP-NN as a classifier assisted by the proper feature selection method can increase the level of accuracy of the classification itself. This is due to the improvement process [11]. MLP-NN consists of three layers of neurons, namely the input layer, hidden layer, and output layer, where each layer contains several neurons that are interconnected with other neurons in the next layer. The input layer has the input values to be processed. The hidden layer serves as a liaison between the input and output layers [12].

The K-Fold Cross Validation is a validation method by dividing the data into k-subsets, then repeating k times for learning and testing [13].

The confusion matrix is one method that can be utilized to measure the performance of a classification method. The confusion matrix contains information that compares the results of the classification carried out by the system with the category results that should be [14].

The receiver Operating Characteristic (ROC) curve is a technique to visualize and test the classification performance based on its performance [15]. The ROC curve illustrates the prediction of classification based on the results of the true positive rate (TPR) and false positive rate (FPR) [16].

Method

Figure 1 illustrates the research stage, which is divided into three sections. The first stage is labeling all input data, followed by the evaluation stage of the model by conducting a training and testing process carried out in 8-fold cross-validation. The last stage is the identification stage, which uses the best intelligent model obtained to carry out the identification process of new data that has never been used in the training stage.

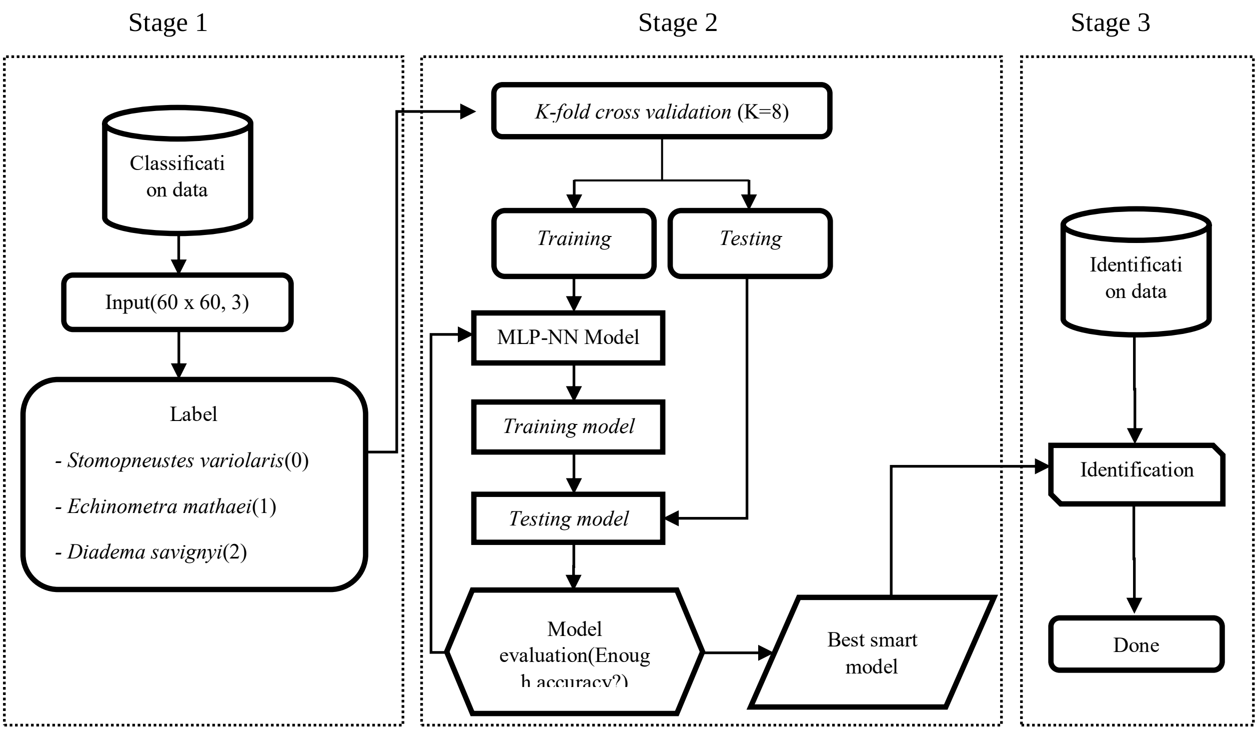


Figure 1. Research stage

A. Research Data

The total data obtained are 120 images, consisting of 3 types of sea urchins as shown in **Table 1**.

Table 1. Image of sea urchin

Species	Total
Stomopneustes variolaris	40
Echinometra mathaei	40
Diadema savignyi	40
Total	120

(1) (2) (3)

Figure 2. (1) *Stomopneustes variolaris,* (2) *Echinometra mathaei,* (3) *Diadema savignyi.*

B. Data Input

At this stage, data from the directory will be inputted with a size of 60 x 60 channel 3 (RGB) and labeled for each class as shown in **Table 2**.

Table 2. Data input label

Species	Label
Stomopneustes variolaris	0
Echinometra mathaei	1
Diadema savignyi	2

C. K-Fold Cross Validation

120 sea urchin images were used in the training and testing process with a composition of 80% of the data for training and 20% for the testing process. To evaluate the model used 8-fold cross-validation as shown in **Figure 3**.

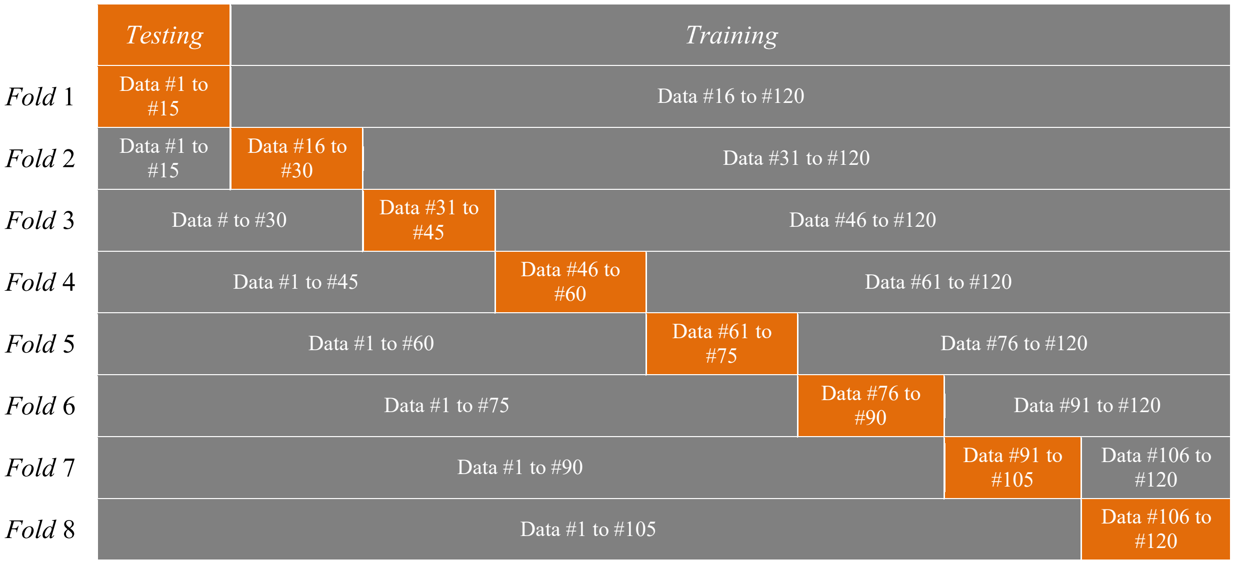


Figure 3. 8-*fold cross validation*

D. *Multilayer Perceptron Neural Network* (MLP-NN)

Figure 4 visualizes the architecture of the MLP-NN model that was built to identify sea urchins which consists of 1 input layer, 3 hidden layers and 1 output layer (1-3-1).

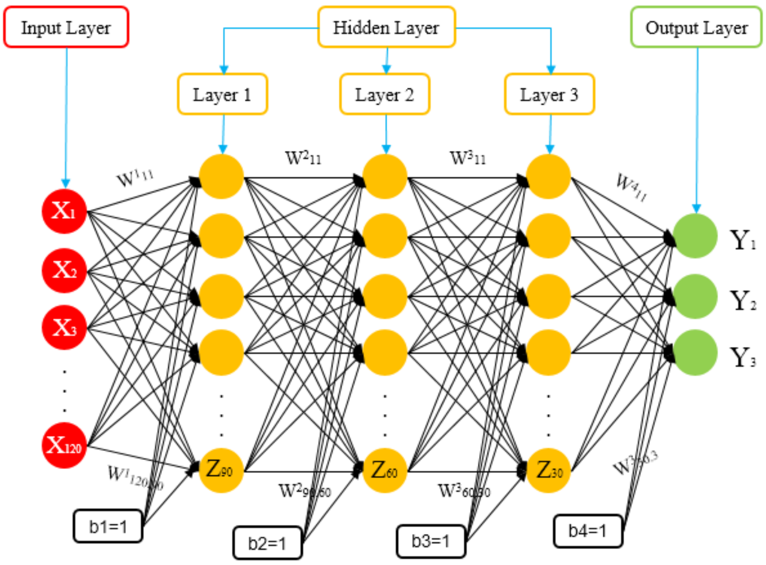


Figure 4. MLP-NN Architecture Identification of Sea Urchins.

X1 to X120 is the inputted sea urchin image. Hidden layer 1 consists of 90 neurons, 60 neurons in hidden layer 2, and 30 neurons in hidden layer 3, with ReLU as the activation function (1). In the Output layer, there are three neurons, Y1, Y2, and Y3, with sigmoid (3) as the activation function. Bias is shown as b1 to b4.

$$f(z) = \max(0, z) \tag{1}$$

Where obtained from

$$z = \sum_{i=1}^{m} w_i x_i + bias \tag{2}$$

$$\sigma(z) = \frac{1}{1+e^{-z}} \tag{3}$$

Information :

w = weight of random value
x = data input
f(z) = activation function
bias = constant value 1
max = maximum value
σ = sigmoid activation function
e = Euler's number (*e* = 2.718)

E. Model evaluation

At this stage, the intelligent model generated from the training process in each fold is tested using data testing. The confusion matrix is used to evaluate the test results and is visualized as a Receiver Operating Characteristic (ROC) curve. The horizontal and vertical axes of the ROC curve show the false positive rate (FPR) and true positive rate (TPR) values, respectively, calculated using **Equation 4** and **Equation 5**.

$$TPR = \frac{(True\ Positive)}{(True\ Positive + False\ Negative)} \tag{4}$$

$$FPR = \frac{(False\ Positive)}{(False\ Positive + True\ Negative)} \tag{5}$$

The level of identification accuracy for each class in the ROC curve is assessed based on the classification category to determine the best performance, as shown in **Table 3**.

Table 3. Classification categories

Score	Category
0.90 – 1.00	Excellent
0.80 – 0.90	Good
0.70 – 0.80	Fair

Score	Category
0.60 – 0.70	Poor
0.50 – 0.60	Very Bad

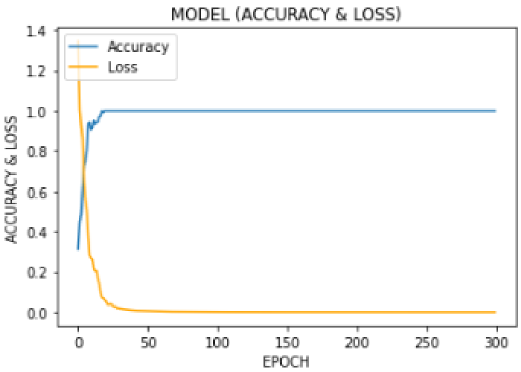
F. Identification

The identification stage is to test the model with the best performance generated from the previous model evaluation stage. The data used is new sea urchin image data that has never been used in the training and testing process in the model evaluation stage.

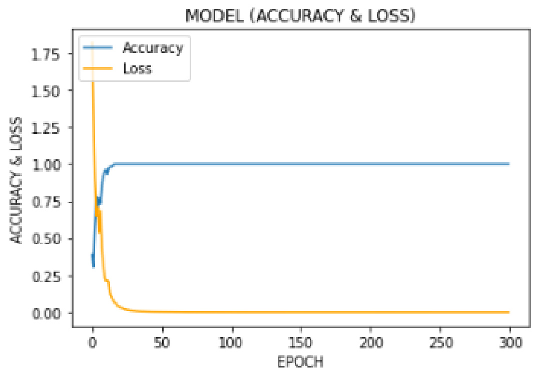
Results and Discussions

A. Training and testing models

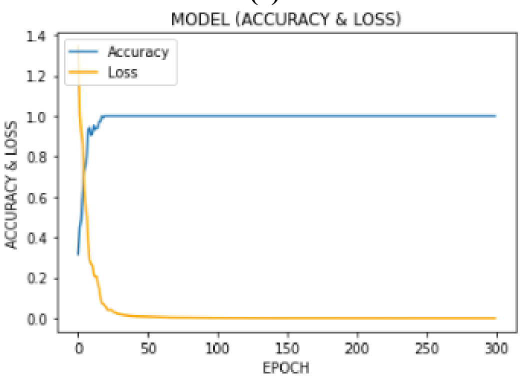
The training process is carried out by setting the epoch parameter value = 300, batch size = 30, and learning rate = 0.0001, and applying 8-fold cross-validation to conduct training eight times. The training results will be tested for performance on data testing in each fold by using a confusion matrix and visualizing it in the form of a ROC curve. The following results from fold 1 to fold eight are shown in **Figures 5** and **Figure 6** as follows:



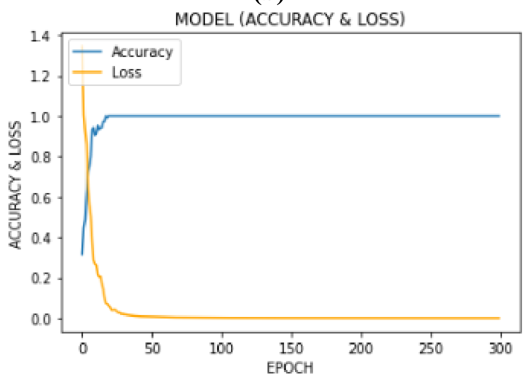
(a)



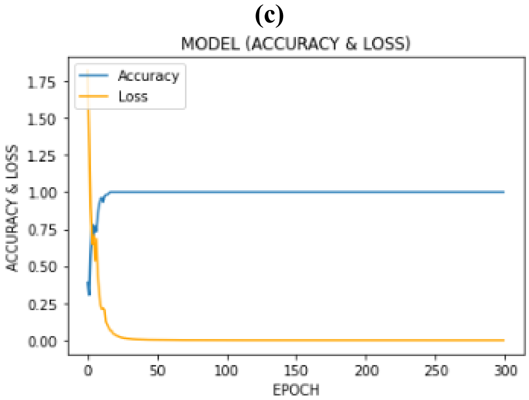
(b)



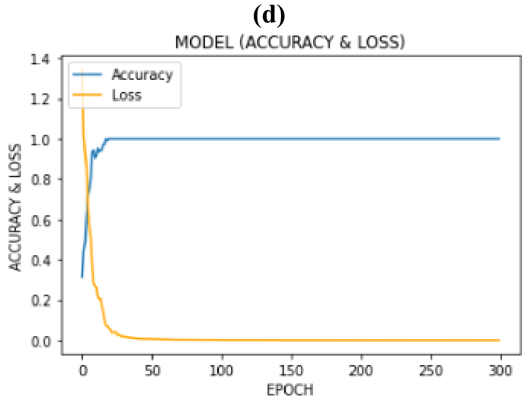
(c)



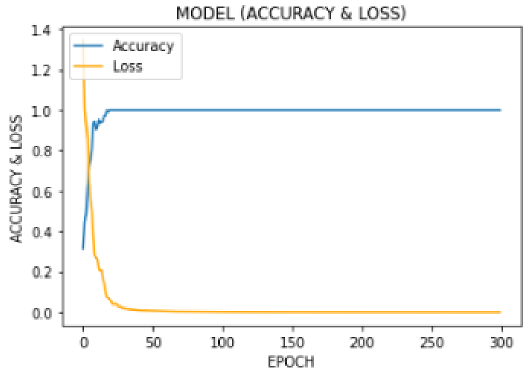
(d)



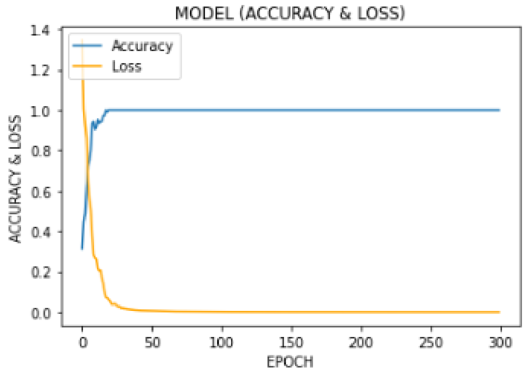
(e)



(f)

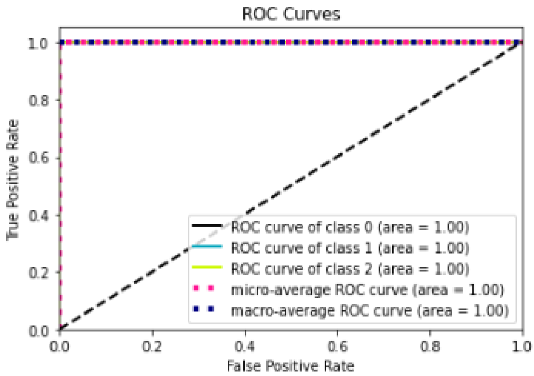


(g)

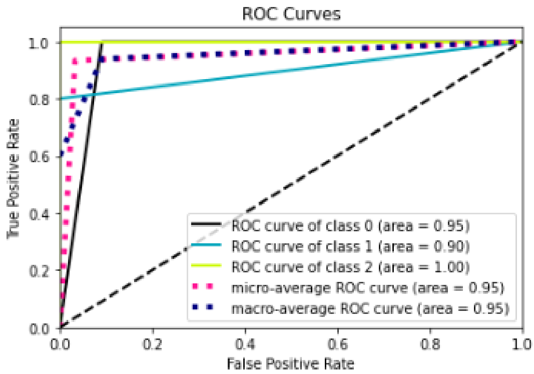


(h)

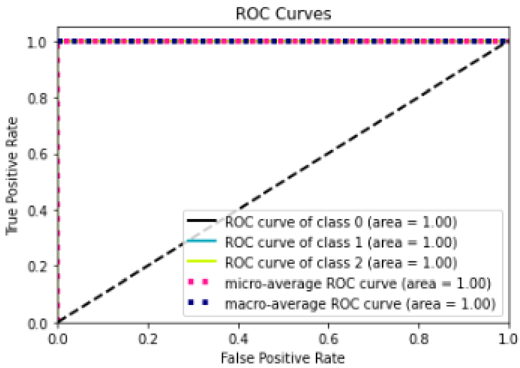
Figure 5. Plot training (a) *fold*-1, (b) *fold*-2, (c) *fold*-3, (d) *fold*-4, (e) *fold*-5, (f) *fold*-6,(g) *fold*-7, (h) *fold*-8.



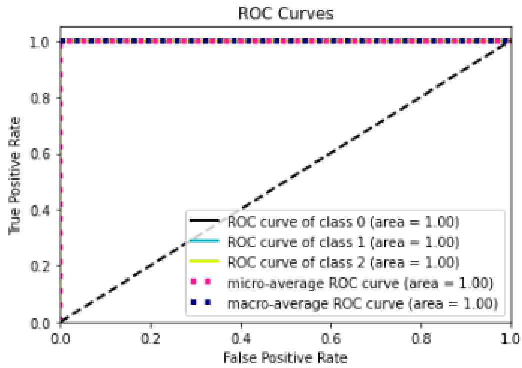
(a)



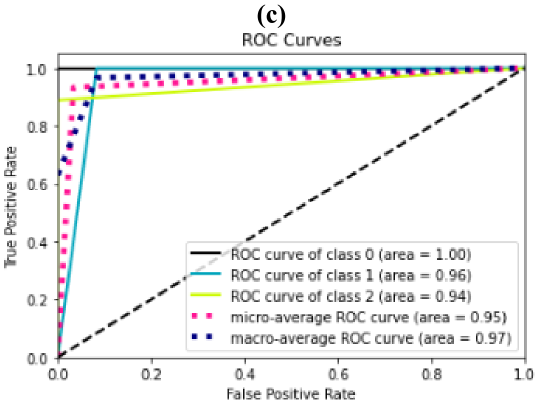
(b)



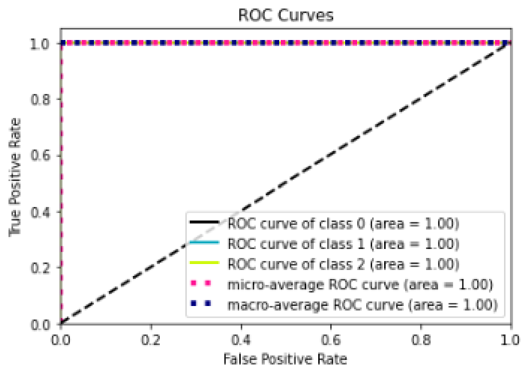
(c)



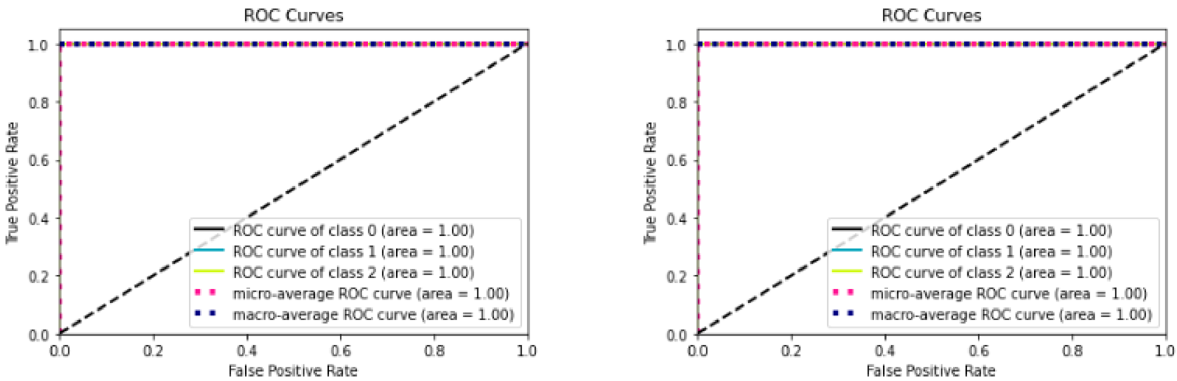
(d)



(e)



(f)



(g) **(h)**

Figure 6. ROC Curve (a) *fold*-1, (b) *fold*-2, (c) *fold*-3, (d) *fold*-4, (e) *fold*-5, (f) *fold*-6, (g) *fold*-7, (h) *fold*-8.

The ROC curve shows an 8-fold result with the best category between 0.90 to 1.00 in each class consisting of class 0 type stomopneustes variolaris, class 1 echinometra mathaei, and class 2 diadema savignyi.

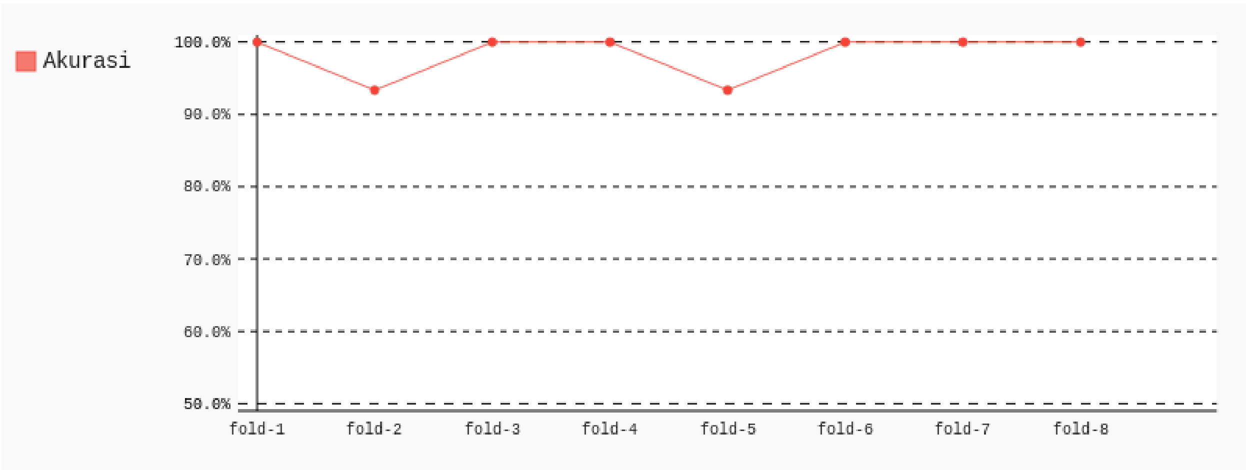


Figure 7. Accuracy of each fold.

The lowest accuracy results are 93.33% on the 2nd and 5th folds, and the highest accuracy of 100% on the 1st, 3rd, 4th, 6th, 7th, and 8th folds.

B. Identification

The identification results are visualized in **Table 4**. Of the 15 new data tested with five images for each class, 13 data were predicted to be correct, and 2 data were predicted to be incorrect. The resulting identification accuracy is 86.66%. The confusion matrix identification results are shown in **Table 5**. Two data that were predicted to be incorrect were from the Stomopneustes variolates species and the Diadema resavignyi species. In contrast, all data for the sea urchins from the Echinometra mathaei species were predicted to be correct.

Table 4. Identification results

15 data		**Stomopneustesvariolaris**	**Echinometramathaei**	**Diademasavignyi**
Prediction	Correct	4	5	4
	Incorrect	1	0	1
Accuracy		$\frac{prediksi\ benar}{jumlah\ data}\ x\ 100\% = \frac{13}{15}\ x\ 100\% = 86{,}66\%$		

Table 5. Confusion Matrix Identification results

15 data		Actual		
		Stomopneustes variolaris	***Echinometra mathaei***	***Diadema savignyi***
Prediction	***Stomopneustes variolaris***	4	0	1
	Echinometra mathaei	0	5	0
	Diadema savignyi	1	0	4

Conclusion

Identification of sea urchin species using the MLP-NN method can work well. The results show the lowest accuracy is 93.33%, the highest is 100%, and the average test data is 86.66%. The experimental results indicate that MLP-NN can classify sea urchins.

References

[1] Karmana, O., dan Fitriana, R., 2007. Cerdas Belajar Biologi. Grafindo Media Pratama, Bandung. 338 halaman.

[2] Lalombombuida, S., M. Langoy, dan D. Y. Katili, 2019. Keanekargaman *Echidodermata* di Pantai Paranti Desa Tabang, Kecamatan Rainis Kabupaten Kepulauan Talaud Provinsi Sulawesi Utara. *Jurnal Perikanan dan Kelautan Tropis*, 10(2), 39-50.

[3] Muliantara, A., dan I. M. Widiartha, 2011. Penerapan Multi Layer Perceptron dalam Anotasi *Image* secara Otomatis. *Jurnal Ilmu Komputer,* 4(2), 9-15.

[4] Latumakulita, L. A., and T. Usagawa, 2017. "A combination of backpropagation neural network on fuzzy inference system approach in Indonesia scholarship selection process: Case study: “Bidik misi” scholarship selection," *2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, pp. 1309-1314.

[5] Latumakulita, L. A., and T. Usagawa, 2018. indonesia scholarship selection model using a combination of Back-Propagation Neural Network and Fuzzy Inference System Approaches. *International Journal of Intelligent Engineering and Systems*, 11(3), 79-80.

[6] Mairi, V. G. N., A. L. Latumakulita, dan T. S. Deiby, 2021. Sistem identifikasi jenis ikan karang lokal taman nasional bunaken menggunakan metode *Backpropagation Neural Network* . *Proceeding KONIK (Konferensi Nasional Ilmu Komputer)*, 5(1), 307–311.

[7] Santoso, H. M,. D. A. Larasati, dan Muhathir, 2020. Wayang *Image Classification using* MLP *method* and GLCM *FeatureExtraction*. *Journal of Computer Science, Information Technology and Telecommunication Engineering (JCoSITTE)*, 1(2), 111-119.

[8]Purnama, N., I. K. G. D. Putra, dan P. A. Bayupati, 2014. klasifikasi website menggunakan algoritma *Multilayer Perceptron*. *Jurnal Teknologi Elektro*, 13(2), 8-15.

[9] Nurhadi, dan Yanti, F., 2016. Buku ajar taksonomi invertebrata/ oleh, Drs. Nurhadi, M.Si. & Febri Yanti, M.Pd. Deepublish, Yogyakarta. 115 halaman.

[10] Toha, A. H. A., 2006. Manfaat Bulu Babi (*Echinoidea*), dari Sumber Pangan Sampai Organisme Hias. *Jurnal Ilmu-ilmu Perairan dan Perikanan Indonesia*, 13(1), 77-82.

[11] Purnama, N., I. K. G. D. Putra, dan P. A. Bayupati, 2014. Klasifikasi website menggunakan algoritma *Multilayer Perceptron*. *Jurnal Teknologi Elektro*, 13(2), 8-15.

[12] Fachrie, M., dan A. P. Wibowo, 2018. Jaringan saraf tiruan untuk memprediksi kinerja SATPAM. *Jurnal Informatika dan Komputer* (JIKO), 3(1), 46-51.

[13] Widjaya, A., L., Hiryanto, dan T., Handhayani, 2017. Prediksi masa studi mahasiswa dengan voting feature interval 5 pada aplikasi konsultasi akademik online. Journal of Computer Science and Information Systems, 1(1), 25-33.

[14] Karsito, Susanti, S., 2019. Klasifikasi kelayakan peserta pengajuan kredit rumah dengan algoritma Naïve Bayes di perumahan azzura residencia. Jurnal Teknologi Pelita Bangsa, 9(3), 43-48.

[15] Yuniarto, S. M., dan A. E. Sarwoko, 2020. Implementasi metode K-Nearest Neighbor untuk diagnosis kanker kolorektal dengan Biomarker Micro-RNA, Jurnal Masyarakat Informatika, 11(1), pp. 35-48.

[16] Arini, L. K. Wardhani, dan D. Octaviano, 2020. Perbandingan seleksi fitur term frequency & tri-gram character menggunakan algoritma Naïve Bayes Classifier (NBC) pada tweet hashtag #2019gantipresiden. *KILAT*, *9*(1), 103 – 114.