



# Lovebird image classification based on convolutional neural network

Amelia Gizzela Sheehan Auni<sup>a</sup>, Christy Atika Sari<sup>a</sup>, Eko Hari Rachmawanto<sup>a</sup>, Mohamed Doheir<sup>b</sup>

<sup>a</sup>Universitas Dian Nuswantoro, Semarang 50131, Indonesia

<sup>b</sup>Universiti Teknikal Malaysia Melaka, Melaka 76100, Malaysia

<sup>1</sup>Corresponding author: [gizelaseehan@gmail.com](mailto:gizelaseehan@gmail.com)

## ARTICLE INFO

### Article history:

Submitted 8 September 2023

Received 11 September 2023

Received in revised form 26 September 2023

Accepted 27 October 2023

Available online on 30 November 2023

### Keywords:

Lovebird, agapomis, classification, CNN.

### Kata kunci:

Lovebird, agapomis, klasifikasi, CNN.

## ABSTRACT

Lovebird is a type of bird from the Psittacidae family, consisting of 90 generations. One of them is the genus Agapomis Selby or Lovebird, which has 9 species. In recognizing the differences of each species, you can use the Object Recognition system. One of them uses the popular CNN algorithm. The dataset was obtained from open sources totaling 8,992 datasets from 9 Agapomis species. It consists of 80% training images and 20% testing images from several datasets. After 10 accuracy tests, the results stated that the accuracy rate reached 89%. In addition, there are also extraction features extracted from images including color, shape, size, and texture characteristics. The things extracted in this study include the Mean, Standard Deviation, Kurtosis, Skewness, Variance, Entropy Value, Maximum Pixel, and Minimum Pixel.

## ABSTRAK

Lovebird merupakan jenis Burung yang berasal dari keluarga Psittacidae yang terdiri dari 90 genus. Salah satunya terdapat genus Agapomis Selby atau genus Lovebird yang memiliki 9 spesies. Dalam mengenali perbedaan dari masing-masing spesies, dapat dibantu menggunakan sistem dari Object Recognition. Salah satunya menggunakan algoritma CNN yang populer. Dataset diperoleh dari sumber terbuka yang berjumlah 8.992 dataset dari 9 spesies Agapomis. Dari sejumlah dataset terdiri dari 80% citra training dan 20% citra testing. Setelah 10 kali uji akurasi, hasil menyatakan bahwa tingkat akurasi nya mencapai 89%. Selain itu, ada juga fitur ekstraksi yang diekstraksi dari citra meliputi ciri warna, bentuk, ukuran, dan tekstur. Hal-hal yang di ekstraksi dalam penelitian ini meliputi Mean, Standar Deviasi, Kurtosis, Skewness, Variansi, Nilai Entropi, Pixel maksimum dan Pixel minimum.

Available online at <http://dx.doi.org/10.36055/tjst.v19i2.21946>

## 1. Introduction

Classification of lovebird species (Agapornis) using the Convolutional Neural Network (CNN) algorithm is a method that aims to recognize and differentiate lovebird species based on visual features in the image. The application of CNN in the classification of lovebirds involves convolution layers that can extract important patterns and features from bird images [1]. These layers help the model learn to recognize the unique characteristics of different types of lovebirds.

Object Recognition is a general science in the fields of computer vision and robot vision. In previous years in the field of neuroscience, CNNs have played a key role in solving many problems related to object identification and recognition [2]. As the visual system of our brain shares many features with CNN properties it is very easy to model and test the object classification and identification problem domains. CNN is usually a forward architecture feed, on the other hand, the visual system is based on CNN (CNN) to couple repeated connections to each convolutional layer. This method has the potential to recognize visual differences between lovebird species, as proven in this study. Globally, there are 330 species of the genus Psittacidae from the class Aves, one of which is the genus Agapornis Selby. This group consists of nine species, namely Peach-faced Lovebird (*A. roseicollis*), Black-winged Lovebird (*A. taranta*), Grey-headed Lovebird (*A. canus*), Masked Lovebird (*A. personatus*), Nyasa Lovebird (*A. lilianae*), Black-collared Lovebird (*A. swindemianus*), Fischer's Lovebird (*A. fischeri*), Red-faced Lovebird (*A. pullarius*) and Black-cheeked Lovebird (*A. nigrigenis*) [3].

This study implements the CNN algorithm to carry out semantic classification by assigning semantic labels to lovebird objects [4]. Lovebirds were chosen because of the characteristics of the species which have different color [3][5]. Research using CNN to classify lovebirds has produced a promising level of accuracy. The CNN model used consists of convolution layers, flattened layers, dense layers, and dropout layers, which help in producing good



classification results. By utilizing feature representation that is automatically learned by CNN, this algorithm can understand the differences between types of lovebirds based on their visual characteristics [6].

Deep learning, especially Convolutional Neural Networks (CNNs), excels at image classification tasks by autonomously extracting intricate features from raw pixel data. This ability eliminates the need for manual feature engineering, making the process more efficient and adaptable to various image types [7]. Moreover, deep learning models can recognize subtle patterns, textures, and relationships within images, enabling accurate differentiation between objects with similar appearances. This level of precision is crucial in diverse applications such as medical diagnosis, autonomous driving, security surveillance, e-commerce product recommendation, and more. The scalability of deep learning allows these models to process and classify vast amounts of images, leading to real-world solutions that are both accurate and efficient.

By harnessing the power of deep learning, image classification becomes a robust tool for automating tasks that require visual interpretation, reducing human intervention, and driving advancements in technology across multiple domains.

## 2. Proposed Method

This study uses the Deep-CNN model with 15 two-dimensional layers along with several parameters, which are visualized in the flowchart above. The training process is carried out in the Deep-CNN model consisting of Convolution, Batch Normalization, ReLU, and Max Pooling which are repeated 3 times. Then do the complete classification process. After repetition, a complete classification process is carried out, Fully Connected, Softmax, and Classification Output. Further explanation of the process is in Figure 1.



Figure 1. Flowchart of Proposed Diagram

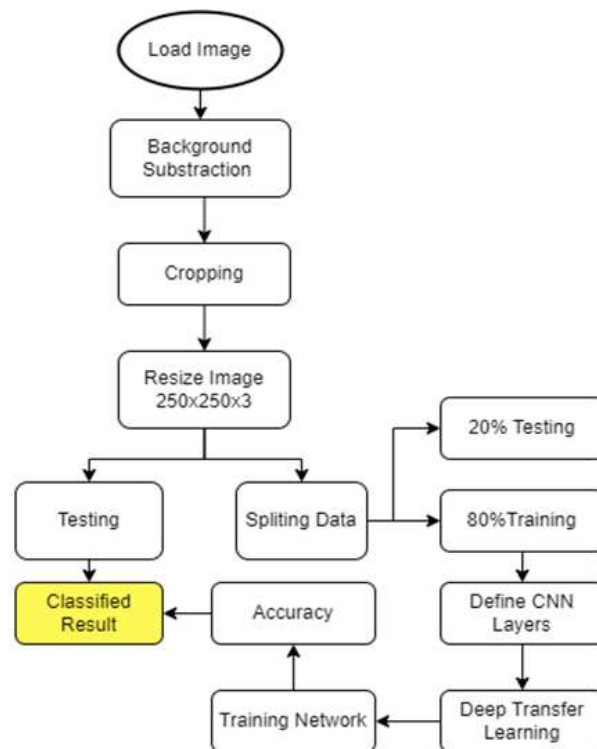


Figure 2. Classification Feature Architecture

Furthermore, the data is processed with Background Subtraction and image segmentation to obtain data with a homogeneous background [8]. From these data continue the process of cropping and resizing to get clear data. 80% of the training data and 20% of the testing data are processed for the training process further to processing the CNN (Define Transfer Layers) layer and defining the Architecture to be used [9]. Multiple convolutional layers ('convolution2dLayer'), batch normalization ('batchNormalizationLayer'), ReLU activation ('reluLayer'), and max-pooling ('maxPooling2dLayer') are defined to form a CNN. The last fully connected layer with softmax activation is defined for classification as in Figure 2.

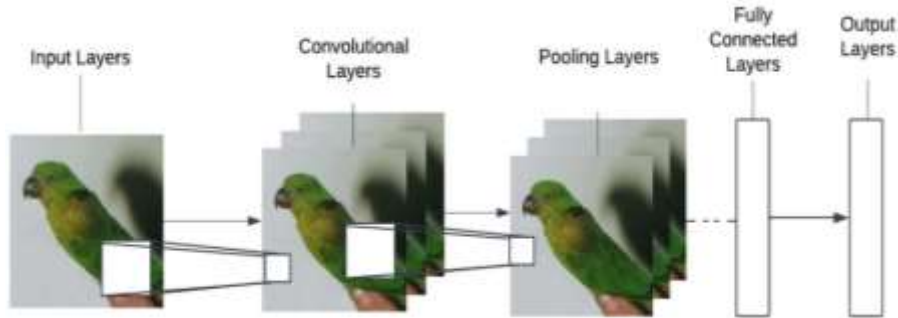
CNNs are trained using the 'train network' function with training data, layers, and options to process training images. In the Accuracy calculation step, the trained network is used to classify validation data using the classify function. Accuracy is calculated by comparing the predicted label (Pred) with the actual label (Valid). The percentage accuracy is then displayed in the MATLAB user interface using the set function [10][11]. The results of the training data are then exported for the classification process of the data entered. The new data that will be tested will also undergo an adjustment process with the cropping and resizing process.

### 2.1. CNN Fundamental

Convolutional Neural Network (CNN) is a special type of artificial neural network designed to process grid-like data, such as images or spatial data [12]. CNNs are very effective for tasks such as image classification and object detection because of their ability to learn feature hierarchies automatically from input data. CNNs use a convolutional layer to apply filters or kernels to the input data, allowing the network to detect different features at different scales. Learning this feature hierarchy helps CNN understand complex patterns in images [13]. Additionally, CNNs often include a joining layer to reduce spatial

dimension and prevent overfitting. By exploiting the spatial structure of the input data, CNN has advanced the fields of computer vision and pattern recognition significantly. They are widely used in a wide variety of applications, from image recognition to medical imaging [14]. The following is the structure of the CNN Algorithm [15].

Figure 3. Proposed CNN Architecture



The object detection model using regions with CNN is based on the following three processes:

1. Find areas in the image that may contain an object. These areas are called proposed areas.
2. Extract CNN features from region proposals.
3. Classify objects using extracted features

2.2. Confusion Matrix

A confusion matrix is a fundamental tool used in machine learning for evaluating the performance of a classification model. It provides a clear and organized way to summarize the results of a classification problem. Here's an explanation of the confusion matrix:

- True Positives (TP): These are the cases where the model correctly predicted the positive class (e.g., correctly identifying actual disease cases in a medical diagnosis).
- True Negatives (TN): These are the cases where the model correctly predicted the negative class (e.g., correctly identifying non-disease cases in a medical diagnosis).
- False Positives (FP): These are the cases where the model incorrectly predicted the positive class when it should have been negative (e.g., wrongly diagnosing a healthy person as having a disease).
- False Negatives (FN): These are the cases where the model incorrectly predicted the negative class when it should have been positive (e.g., failing to diagnose a person with a disease when they actually have it).

The confusion matrix is usually presented in a tabular form like this:

Table 1. Confusion Matrix

Actual/Predicted	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5
Class 0	TP (0,0)	FP (1,0)	FP (2,0)	FP (3,0)	FP (4,0)	FP (5,0)
Class 1	FN (0,1)	TP (1,1)	FP (2,1)	FP (3,1)	FP (4,1)	FP (5,1)
Class 2	FN (0,2)	FN (1,2)	TP (2,2)	FP (3,2)	FP (4,2)	FP (5,2)
Class 3	FN (0,3)	FN (1,3)	FN (2,3)	TP (3,3)	FP (4,3)	FP (5,3)
Class 4	FN (0,4)	FN (1,4)	FN (2,4)	FN (3,4)	TP (4,4)	FP (5,4)
Class 5	FN (0,5)	FN (1,5)	FN (2,5)	FN (3,5)	FN (4,5)	TP (5,5)

From the confusion matrix, several performance metrics can be calculated:

- Precision: The proportion of true positive predictions out of all positive predictions (TP / (TP + FP)).
- Recall (Sensitivity or True Positive Rate): The proportion of true positive predictions out of all actual positive instances (TP / (TP + FN)).
- F1-Score: The harmonic mean of precision and recall, useful when there is an imbalance between classes.

The choice of which metric to prioritize depends on the specific problem and the cost associated with different types of errors. For example, in a medical diagnosis, recall might be more critical to ensure that no positive cases are missed, even if it leads to some false alarms (lower precision). In summary, a confusion matrix is a crucial tool for assessing the performance of a classification model, allowing you to understand the balance between true positives, true negatives, false positives, and false negatives in your predictions.

2.3. Pixel Intensity Values of an Image

Pixel intensity values of an image refer to the numerical values assigned to each pixel in the image, representing the brightness or color of that pixel. When performing image processing or analysis, various operations can be applied to these intensity values to extract features, enhance certain aspects of the image, or prepare the data for further analysis. In this scope, mean, standard deviation, kurtosis, skewness, variance, entropy, minimum pixel value, and maximum pixel value are related to various statistical measures that can be computed from the pixel intensity values of an image. These measures are commonly used in image processing and computer vision for different purposes, including texture analysis, feature extraction, and image characterization.

2.3.1. Mean

Calculating the average (mean) of pixel values involves adding up all the pixel values in a data set and dividing by the total number of pixel values. It is often used in various contexts such as statistics, data analysis, and image processing. The general formula for calculating the mean is:

$$M = \frac{\sum_{i=1}^n x_i}{n} \quad (1)$$

$n$  = the amount of data in the sample.

$x_i$  = the data value at the  $i$ -th position.

### 2.3.2. Standard Deviation

The standard deviation measures how far the values in a data set are spread from the mean (mean). It gives an idea of the variability or dispersion of the data. The formula for calculating the standard deviation is as follows:

$$std = \sqrt{\frac{\sum_{i=1}^n (x_i - m)^2}{n}} \quad (2)$$

$n$  = the amount of data in the sample.

$x_i$  = the data value at the  $i$ -th position.

$m$  = the average of the data.

### 2.3.3. Kurtosis

Kurtosis measures the shape and sharpness of the peaks of the data distribution. The general formula for calculating kurtosis in image extraction is as follows:

$$Kurtosis = \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{n \cdot s^4} \quad (3)$$

$n$  = the amount of data in the sample.

$x_i$  = the data value at the  $i$ -th position.

$\bar{x}$  = the average of the data.

$s$  = the standard deviation of the data

### 2.3.4. Skewness

Skewness in image extraction refers to the asymmetry of the distribution of pixel intensities within an image. Skewness is a statistical measure that quantifies the lack of symmetry in a distribution. In the context of image processing, skewness can be used to detect and correct the tilting or rotation of text or objects in document images. The general formula for calculating the skewness is:

$$Skewness = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{n \cdot s^3} \quad (4)$$

$n$  = the amount of data in the sample.

$x_i$  = the data value at the  $i$ -th position.

$\bar{x}$  = the average of the data.

$s$  = the standard deviation of the data

### 2.3.5. Variance

In image analysis, variance is utilized to distinguish between different regions based on their pixel intensity distribution. It plays a role in tasks like object detection, edge detection, and texture analysis. The general formula for calculating the variance is:

$$variance = \frac{\sum_{i=1}^n (x_i - m)^2}{n} \quad (5)$$

$n$  = the amount of data in the sample.

$x_i$  = the data value at the  $i$ -th position.

$m$  = the average of the data.

### 2.3.6. Entropy

In image analysis, variance is utilized to distinguish between different regions based on their pixel intensity distribution. It plays a role in tasks like object detection, edge detection, and texture analysis. The general formula for calculating the Entropy is:

$$- \sum_{i=1}^N p(x_i) \cdot \log_2(p(x_i)) \quad (6)$$

$N$  = the number of possible different values in the dataset (in image contrast, the number of different pixel intensities).

$x_i$  = the intensity value of the pixel value in the  $i$ -th position.

$p x_i$  = the probability of occurrence of the value  $x_i$  in the dataset.

### 2.3.7. Minimum Pixel value

The minimum pixel value in image subtraction refers to the process of subtracting the minimum pixel value of one image from the corresponding pixels of another image. This operation can be used to emphasize or enhance the differences between images, highlight specific features, or normalize pixel values. The general formula for calculating the minimum value is:

$$Min = \min(x_1, x_2, \dots, x_n) \quad (7)$$

$n$  = the amount of data in the sample.

$X_i$  = the data value at the  $i$ -th position.

### 2.3.8. Maximum Pixel Value

When performing image subtraction, typically the pixel values of one image are subtracted from the corresponding pixel values of another image. This operation aims to highlight differences between images. The general formula for calculating the maximum pixel value is:

$$Max = \max(x_1, x_2, \dots, x_n) \quad (8)$$

$n$  = the amount of data in the sample.

$X_i$  = the data value at the  $i$ -th position.

## 2.4. Datasets

The dataset plays a critical role in image classification tasks. A dataset is a curated collection of digital images along with corresponding labels that define the classes or categories of the images. The success of an image classification model heavily depends on the quality, diversity, and size of the dataset used for training [16]. The dataset taken is open source with a total of 8,992 files with image extensions (.jpg) as shown in the sample in Figure 2. The dataset is processed to the classification stage with the first process grouping data based on the type of species. There are 9 folders containing ~800-1500 data that have been prepared for training sessions. Furthermore, the data is processed with Background Subtraction and image segmentation to obtain data with a homogeneous background [6]. 20% of the data has been used for training and 80% for testing.



Figure 4. Lovebirds Datasets

## 3. Result and Discussion

This study aims to classify images of lovebirds (Agapornis) based on their species using the Convolutional Neural Network (CNN) architecture. This method utilizes deep learning techniques that can extract important features from images automatically for classification purposes. CNN can recognize complex patterns in images, such as the unique characteristics of each type of lovebird [17].

This research involves implementing CNN with the appropriate architecture and configuration, as well as training the model using a lovebird image dataset that has been labeled with its species. The main goal is to create a model that can distinguish between types of lovebirds based on their visual characteristics. The results of this study can be used for various applications, including the introduction of bird species in the field of ornithology and natural observations. The CNN method in image classification is very useful in recognizing and classifying visual objects based on complex features that are difficult to identify manually. CNN is widely used in various applications of pattern recognition, image classification, object detection, image segmentation, and other image processing [18]. Its strengths lie in its ability to deal with differences in size, rotation, and other variations in image data, as well as its ability to extract hierarchical features useful in visual analysis [19]. Before being processed, we collect the dataset and prepare it to be processed using Matlab 2020. The following is a sample of the raw dataset as shown in Figure 4. Extraction has been carried out on 9 image categories of each Lovebird (Agapornis) species as shown in Table 2.

Table 2. Feature Extraction

Sample Image (.jpg)	Mean	Standard Deviation	Kurtosis	Skewness	Variance	Entropy	Pixel Min	Pixel Max
A. canus 1	151.285	43.8798	5.79894	-1.66791	1925.44	4.90373	0	253
A. fischeri 1	123.656	53.1678	2.38737	0.062108	2826.82	7.66734	1	238
A. lillianae	165.173	44.3342	4.82287	-1.72719	1965.52	4.7733	0	233
A. nigrigenis 1	120.251	57.6938	2.72542	0.933481	3328.57	7.32873	0	255
A. personatus 1	129.277	38.9767	3.55338	-1.23655	1519.18	6.37582	2	235
A. pullarius 1	147.626	33.7442	3.36227	-0.433357	1138.67	6.94849	1	245
A. roseicollis 1	113.043	30.548	5.40891	-0.42554	933.18	6.3855	2	254
A. swindermianus 1	158.485	69.302	1.44253	-0.307218	4802.76	7.396	5	255
A. taranta 1	123.151	51.0404	2.08603	-0.078518	2605.13	7.60317	0	255

The feature extraction function in image processing refers to the process of identifying, calculating and compiling representative features of an image. The goal of feature extraction is to simplify the image representation in such a way that the most relevant information can be represented by the features. These features can then be used as input for machine learning algorithms, classification, or other image analysis tasks.

**Table 3. Sample Result of Classification of Lovebird (Agapornis)**

Image Name	Real Name	Folder Name	True / False
1.jpg	A. canus	canus	F
19.jpg	A. canus	canus	F
3.jpg	A. fischeri	fischeri	T
35.jpg	A. fischeri	fischeri	T
8.jpg	A. liliane	liliane	T
23.jpg	A. liliane	liliane	T
4.jpg	A. nigrigenis	nigrigenis	T
19.jpg	A. nigrigenis	nigrigenis	T
20.jpg	A. personatus	personatus	T
26.jpg	A. personatus	personatus	T
17.jpg	A. pullaria	pullaria	T
5.jpg	A. pullaria	pullaria	T
7.jpg	A. roseicollis	roseicollis	T
120.jpg	A. roseicollis	roseicollis	T
59.jpg	A. swindernianus	swindernianus	T
32.jpg	A. swindernianus	swindernianus	T
22.jpg	A. taranta	taranta	T
20.jpg	A. taranta	taranta	T

Table 3 above contains 18 times Results of the Classification of Lovebird (Agapornis) from a different kind of picture using the CNN algorithm. Accuracy is a key metric in evaluating the performance of a classification model, and an 88% accuracy suggests that the model is making correct predictions for the majority of the cases. The classification results above are considered reasonable and it indicates that the model trained with the CNN algorithm has demonstrated a high level of success in distinguishing between different types of pictures of Lovebirds.

**Table 4. Result of Confusion Matrix**

Test	Precision	Recall	F1-Score
Test 1	0.80	0.95	0.87
Test 2	0.97	0.78	0.86
Test 3	0.91	0.94	0.92
Test 4	0.95	0.89	0.92
Test 5	0.89	0.87	0.88

CNNs are powerful algorithms for image processing that automatically learn relevant features from the input images. CNNs excel at feature extraction due to their ability to learn hierarchical representations directly from raw images [20]. This makes them highly effective for image classification tasks, allowing them to capture relevant information and patterns within the data. Besides extraction, there are also classification features that we also test directly. The results of the trial are as follows in Table 3. From 18 classification tests, the false result happened twice. This means the accuracy is equals 88.88%. The CNN Accuracy that are generated test also generates the same percentage of accuracy. Around 80% of CNN accuracy tests are recorded during the training system.

## 4. Conclusions

Based on the classification experiment using the CNN method which was carried out on 9 species of lovebird (Agapornis) on 8992 data. A total of 8,992 datasets, 80% consisting of training images and 20% are testing images. After testing the accuracy 10 times, the results showed an accuracy rate around 89%. In addition, there are extraction features extracted from images, including features such as color, shape, size, and texture. This research also involves the extraction of several things, such as Mean, Standard Deviation, Kurtosis, Skewness, Variance, Entropy Value, Maximum Pixel, and Minimum Pixel. The purpose of this research is to find the extracted value of pre-processed images. For future research, lovebirds might be real-time recognized using another CNN scheme such as implemented in YOLO.

## REFERENCES

- [1] Farou, Z. (2022). Automatic Fine-grained Classification of Bird Species Using Deep Learning Eötvös Loránd University DEPT. of Data Science and Engineering Automatic Fine-grained Classification of Bird Species Using Deep Learning Supervisor: Zakarya Farou PhD student Auth, "No. December, 2022. Doi: 10.13140/RG.2.2.18648.57601.
- [2] Naushad Ali, M. M., Abdullah-Al-Wadud, M., & Lee, S. L. (2013). Moving object detection and tracking using particle filter. *Appl. Mech. Mater.*, Vol. 321–324, No. 01, Pp. 1200–1204. Doi: 10.4028/www.scientific.net/AMM.321-324.1200.
- [3] Belynza, K. A., & Syaputra, H. (2022). Objek Deteksi Burung Lovebird Menggunakan Instance Segmentation Mask R-Cnn. *Bina Dama Conference on Computer Science (BDCCS)*, Vol. 4, No. 1, Pp. 245–254.

- [4] Sri, M. S., Naik, B. R., & Sankar, K. J. (2021). Object Detection Based on Faster R-Cnn. *Int. J. Eng. Adv. Technol.*, Vol. 10, No.3, Pp.72–76. Doi: 10.35940/ijeat.c2186.0210321.
- [5] van der Zwan, H., Visser, C., & van der Sluis, R. (2019). Plumage colour variations in the *Agapornis* genus: a review. *Ostrich*, Vol. 90, No.1, Pp. 1–10. Doi: 10.2989/00306525.2018.1540446.
- [6] Peng, C., Zhang, C., Xue, X., Gao, J., Liang, H., & Niu, Z. (2022). Cross-modal complementary network with hierarchical fusion for multimodal sentiment classification. *Tsinghua Sci. Technol.*, Vol. 27, No. 4, Pp. 664–679. Doi: 10.26599/TST.2021.9010055.
- [7] Zebari, R., Abdulazeez, A., Zeebaree, D., Zebari, D., & Saeed, J. (2020). A Comprehensive Review of Dimensionality Reduction Techniques for Feature Selection and Feature Extraction. *J. Appl. Sci. Technol. Trends*, Vol. 1, No. 2, Pp. 56–70. Doi: 10.38094/jastt1224.
- [8] Fung, K. L. Y. *et al.* (2020). Ultramicroscopy Accurate EELS background subtraction – an adaptable method in MATLAB. *Ultramicroscopy*, Vol. 217, No. June, P. 113052. Doi: 10.1016/j.ultramic.2020.113052.
- [9] Dharma, A. S., Tambunan, S., & Naibaho, P. K. (2022). Deteksi Objek Aksara Batak Toba Menggunakan Faster R-CNN dan YoloV3. [Online]. Available: [https://www.academia.edu/68753590/Deteksi\\_Objek\\_Aksara\\_Batak\\_Toba\\_Menggunakan\\_Faster\\_R\\_CNN\\_dan\\_YoloV3](https://www.academia.edu/68753590/Deteksi_Objek_Aksara_Batak_Toba_Menggunakan_Faster_R_CNN_dan_YoloV3)
- [10] Irawan, F. A., Sudarma, M., & Care Khrisne, D. (2021). Rancang Bangun Aplikasi Identifikasi Penyakit Tanaman Pepaya California Berbasis Android Menggunakan Metode Cnn Model Arsitektur SqueezeNet. *J. SPEKTRUM*, Vol. 8, No. 2, P. 18. Doi: 10.24843/spektrum.2021.v08.i02.p3.
- [11] Wang, X., Liu, Y., & Xin, H. (2021). Bond strength prediction of concrete-encased steel structures using hybrid machine learning method. *Structures*, Vol. 32, No. November 2020, Pp. 2279–2292. Doi: 10.1016/j.istruc.2021.04.018.
- [12] Andresini, G., Appice, A., & Malerba, D. (2021). Nearest cluster-based intrusion detection through convolutional neural networks. *Knowledge-Based Syst.*, Vol. 216, P. 106798. Doi: 10.1016/j.knosys.2021.106798.
- [13] Charli, F., Syaputra, H., Akbar, M., Sauda, S., & Panjaitan, F. (2020). Implementasi Metode Faster Region Convolutional Neural Network (Faster R-CNN) Untuk Pengenalan Jenis Burung Lovebird. *J. Inf. Technol. Ampera*, Vol. 1, No. 3, Pp. 185–197. Doi: 10.51519/journalita.volume1.issuue3.year2020.page185-197.
- [14] Zaman, K., & Maghdid, S. S. (2021). Medical Images Classification for Skin Cancer Using Convolutional Neural Network Algorithms. *Advances in Mechanics*, Vol. 9, No. 3, Pp. 526-541.
- [15] Jia, W., Tian, Y., Luo, R., Zhang, Z., Lian, J., & Zheng, Y. (2020). Detection and segmentation of overlapped fruits based on optimized mask R-CNN application in apple harvesting robot. *Comput. Electron. Agric.*, Vol. 172, No. March, p. 105380. Doi: 10.1016/j.compag.2020.105380.
- [16] Huang, R. *et al.* (2023). NLIP: Noise-Robust Language-Image Pre-training. *Proc. 37th AAAI Conf. Artif. Intell. AAAI 2023*, Vol. 37, Pp. 926–934.
- [17] Luo, T., Jiang, G., Yu, M., Zhong, C., Xu, H., & Pan, Z. (2019). Convolutional neural networks-based stereo image reversible data hiding method. *J. Vis. Commun. Image Represent.*, Vol. 61, Pp. 61–73. Doi: 10.1016/j.jvcir.2019.03.017.
- [18] Bai, Q., Li, S., Yang, J., Song, Q., Li, Z., & Zhang, X. (2020). Object detection recognition and robot grasping based on machine learning. *Asurvey. IEEE Access*, Vol. 8, Pp. 181855–181879. Doi: 10.1109/ACCESS.2020.3028740.
- [19] Usha Kingsly Devi, K., & Gomathi, V. (2022). Deep Convolutional Neural Networks with Transfer Learning for Visual Sentiment Analysis. *Neural Process. Lett.*, Vol. 55, No. 4, Pp. 5087–5120. Doi: 10.1007/s11063-022-11082-3.
- [20] Sabuj, H. H., *et al.* (2023). A Comparative Study of Machine Learning Classifiers for Speaker's Accent Recognition. *2023 IEEE World AIoT Congr. AIoT 2023*, Pp. 627–632. Doi: 10.1109/AIoT58121.2023.10174511.