ISSN 2089-8673 (Print) | ISSN 2548-4265 (Online)
Volume 13, Issue 2, July 2024

CLASSIFICATION OF DOG EMOTIONS USING CONVOLUTIONAL NEURAL NETWORK METHOD

Slamet Hermawan[1], Amril Mutoi Siregar[2], Sutan Faisal[3], Tohirin Al Mudzakir[4]

[1,2,3,4]Department of Informatics, Faculty of Computer Science, Buana Perjuangan University Karawang, Indonesia

email: if20.slamethermawan@mhs.ubpkarawang.ac.id[1], amrilmutoi@ubpkarawang.ac.id[2], sutan.faisal@ubpkarawang.ac.id[3], tohirin@ubpkarawang.ac.id[4]

The utilization of neural networks in dog emotion classification has great potential to improve the understanding of pet emotions. The goal is to develop a dog emotion classification system. This is important due to the lack of public ability to recognize and understand dog emotions. Neural networks able to create learning models can be used for decision-making, thus helping to reduce the risk of dangerous dog attacks. CNN itself is part of neural networks, where the CNN model has a higher accuracy rate of 74.75% compared to ResNet 65.10% and VGG 68.67%. Modeling using ROC-AUC shows the model's ability to distinguish emotion classes well. Angry has the highest AUC of 0.97, happy 0.93 and sad 0.96. While relaxed has the lowest AUC of 0.92. Classification report results show model has the highest precision and F1-Score values in angry class, while the highest recall value is in sad class.

Keywords : classification, convolutional neural network, dog emotion, ROC-AUC

Received: 17-01-2024 | **Revised:** 05-04-2024 | **Accepted:** 05-07-2024
DOI: https://doi.org/10.23887/janapati.v13i2.74340

INTRODUCTION

Understanding emotions in pets, especially dogs, has become an important aspect of human-animal relationships. Emotion recognition in dogs can provide owners with deeper insights into the needs and well-being of the pet [1][2]. Dogs have more expressions of emotion than cats, although both use a comparable set of behavioral cues to express the same emotion, different combinations tend to be associated with specific emotions in cats and dogs [3]. Dogs also have varied characteristics and personalities that are interesting to understand and research further [4][5] and can be interpreted visually [6].

Many of them have probably read a lot of research on facial expressions, so most people have done more research on things like human or animal behavior, emotions, breeds, and so on. In this study, the facial expression project was dog emotions. In a study on pet emotions where very different specific (cats -purr‖ and -hiss‖) and hetero specific (humans -happiness‖ and -anger‖) emotional stimuli were given to the tested population using the feeling paradigm [7].

The study of pets' emotions is important so that we can better understand their emotional needs. With this understanding, we can provide better care and build better relationships with our pets.

In the case of animals other than dogs, the object of this research is cats. Recent research in this regard focuses on the classification of cat emotions and uses various methods and applications in image-based emotion recognition, genome analysis, and cat breed classification. This research achieved a high accuracy rate of about 87.00% in recognizing basic cat emotions. This result shows that the method used in this study is effective in recognizing basic emotions in cats with a good level of accuracy [8].

For pet stimuli, namely cats, the vocalizations and facial expressions of cats vary depending on the stimuli they receive. In addition, cats have been shown to alter the acoustic features of their purrs to change the meaning of their vocalizations. This is similar to dogs where emotions towards dogs can vary depending on the individual and the experiences the individual has [9].

A serious problem faced today is the lack of human ability to recognize and understand dog emotions. Where cases of dog attacks cause serious injury and even death to humans. Most people do not have enough experience in reading dogs' expressions and body language [10]. People are often unable to distinguish whether a dog is angry, scared, or uncomfortable. This makes humans vulnerable to dog attacks that occur due to a lack of awareness of the warning signs given by dogs.

In addition, many people also lack exploration of animal faces, especially dog emotions. Dogs' faces contain many clues about the emotions they are feeling, but our lack of understanding of this makes it difficult for us to understand what dogs are feeling.

Each year, 60,000 deaths have been reported due to dog attacks, mostly in Asia and Africa [11]. For this problem, a broader education on dog behavior and body language is needed, and it can be supported by a specially adapted mechanism with the help of an artificial neural network system, which is a computation inspired by the structure and function of nerves in biological nervous systems. This model consists of information processing units called neurons, which are organized in layers. Neural networks can learn from data through a training process and then be used to make decisions. This can help reduce the risk of dangerous dog attacks.

Deep learning is a subfield of machine learning that focuses on using deep neural network architectures to model and solve complex tasks. Deep learning involves neural networks with multiple layers, which allows the model to learn a hierarchical feature representation of the data [12]. A well-known architecture in deep learning for image processing is Convolutional Neural Networks (CNN).

The application of this technology has great potential in better understanding the emotions of pets, giving dog owners a deeper insight into the emotional state of their beloved pets. In addition, the utilization of this technology also opens up opportunities for the development of practical applications, such as early warning systems for significant emotional changes in pets.

The use of deep learning in dog emotion classification opens up opportunities to provide a better understanding of pets, allowing owners to provide better care and support [13]. In other words, dog emotion classification using deep learning is an innovative step in the application of artificial intelligence technology to improve pet understanding and welfare [14].

The combination of deep learning and CNN is often used in image processing tasks that require a high level of understanding of visual data. CNNs can also be used as the primary algorithm for processing and analyzing images of dog facial expressions [15][16][17].

ResNet (Residual Network) and VGG16 are two well-known and frequently used CNN architectures in image processing tasks. ResNet, developed by Microsoft Research, introduced the concept of residue blocks to facilitate deeper network training. ResNet consists of residue blocks with shortcut connections or shortcut skips, which overcomes the problem of degradation during very deep network training.

The main advantage lies in the ability of residue blocks to allow information to skip one or more layers, easing the backward flow of information during training. In addition, ResNet uses batch normalization to accelerate convergence and prevent overfitting. Although architecturally complex, ResNet has proven effective in training deeper networks without compromising performance or signal integrity, especially beneficial for tasks such as object detection and image segmentation [18].

VGG16, developed by the Visual Geometry Group at Oxford University, has a simple structure with deep convolution layers, using 3x3 convolution filters on each of its 16 layers. It uses max pooling to reduce the spatial dimension and adopts the ReLU activation function throughout its network. Unlike ResNet, VGG16 does not involve the concept of residues or special blocks [19].

ResNet and VGG16 both have the potential to significantly contribute to emotional understanding in pets through deep learning approaches [20][21]. In the context of dog emotion classification, the use of ResNet and VGG16 architectures in CNNs provides a different yet effective approach. ResNet, with its residual block concept that allows for deeper network training, can provide an edge in the task of dog emotion classification.

Understanding the dog's emotional state allows the owner or person in charge to respond appropriately to the pet's feelings, and can make a positive contribution to pet understanding and care, building a deeper relationship between humans and pets [22].

In this research, CNN algorithm is used to classify emotions in dogs based on their facial expressions. CNN was chosen because of its ability to process visual data, such as images, and detect deep features. Based on this background, this research aims to develop an emotion classification system in dogs using the Convolutional Neural Network approach. This research is expected to provide information on the extent to which the model is able to identify emotions in dogs accurately, and make a positive contribution to the development of technology that can improve human understanding of pets.

METHOD

This research method contains a description of the steps that would be carried out from the beginning to the end of the research, and

ISSN 2089-8673 (Print) | ISSN 2548-4265 (Online)
Volume 13, Issue 2, July 2024

the flow is seen in the research diagram, as shown in Figure 1.

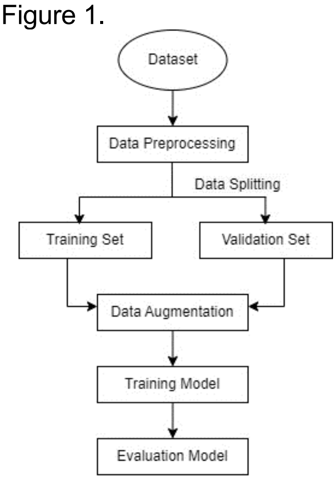

Figure 1. Research Method

Figure 1 describes the research flow starting with the collection of datasets. Next, data preprocessing is carried out in the form of data separation into training and validation data. After preprocessing, the next process is data augmentation. The next step in the augmentation process is model building and training process. The final result of this model is the evaluation or performance of the model created, namely the accuracy of the image given to the model created.

Dataset

Dataset is a collection of data that would be used as the object of research [23]. The use of diverse datasets is necessary to train and test the models built. The dataset used in this study consists of images of dogs with various expressions. Including, happy, sad, angry, and relaxed.

In the context of pets, two emotion measurement scales were created based on specific theories about pets. The validity of the two scales was then analyzed through factor analysis. The first scale is the panksepp emotion scale which includes emotions such as happiness, frustration, anger, fear, lust, enthusiasm, boredom, curiosity, affection, and sadness. The second scale is the Complex Social Emotions Scale which includes emotions such as jealousy, regret, pity, and shame [24].

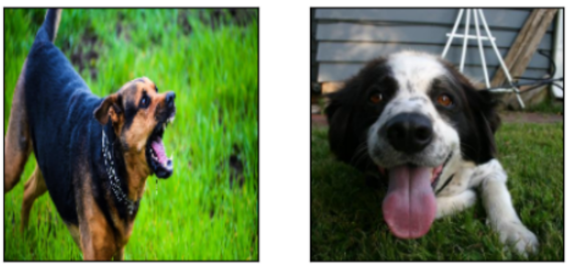

Figure 2. Sample of Dataset

In Figure 2 there is a sample dataset used in this study. This research data comes from kaggle with the link https://www.kaggle.com/datasets/danielshanbalico/dog-emotion where there are 4000 images with four classes, namely: Angry, Happy, Sad and Relaxed. Each class has 1000 images.

This research uses a dataset of pitbull dog breeds. It should be noted that each dog has a different personality and way of reacting. In this research only for pitbull dog breeds. Therefore, emotion classification in dogs is not always universally applicable to all dog breeds. However, several approaches to dog emotion classification have been conducted and some emotional behavior patterns may be recognizable in various dog breeds. Some of the studies and methods used involve analyzing a dog's facial expressions, behaviors, and body reactions to identify emotions such as happy, sad, relaxed, angry, etc., among others.

However, it is important to remember that differences in genetics, environment and past experiences as well as different situations would affect the uniqueness of each dog breed. Therefore, it is important to take a cautious approach and pay attention to context when classifying emotions in dogs.

Preprocessing

The preprocessing step improves the quality of the processed image and produces a better image than the original image. This research uses the Canny method which is one of the image analysis methods capable of producing accurate edge detection. By using the Canny method, edges in the image can be identified with high precision. [25].

The way to determine edge detection in the canny method is by using two threshold values, namely T1 and T2. The T1 value is the lower threshold, while T2 is the upper threshold. The edge detection process starts by comparing the pixel value with the threshold T1. If the pixel value is lower than T1, then the pixel is converted to black (0). If the pixel value is higher than T2, then the pixel is enlarged to white (255). Meanwhile, if the pixel value is between T1 and T2, then the pixel is converted to gray (128).

To determine the optimal combination of T1 and T2, experiments were conducted with various values. On visual observation, it was found that the combination of values (100, 200) resulted in clear and optimal edge detection. Thus, edge detection in the Canny method can be done by determining two threshold values, T1 and T2. Experiments were conducted to find the combination of values that resulted in clear and optimal edge detection. In visual observation, the combination of values (100, 200) was found to be the optimal threshold combination as shown in Figure 3.

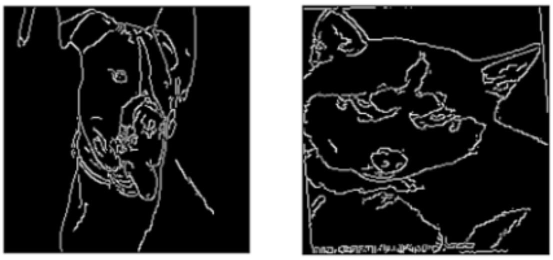

Figure 3. Sample Dataset using Thresholds (100, 200)

One of the main advantages of the Canny method is its ability to remove noise in images [19]. By removing noise, the Canny method is able to provide more accurate edge information, thus enabling a better image segmentation process. In addition to removing noise, the Canny method can also produce sharp edges by reducing the thickness of noise edges [26]. This makes the edge detection results clearer and easier to interpret. The edge sharpness produced by the Canny method makes it ideal for applications that require precise edge detection, such as in object recognition in images or in image processing.

Testing Scenario

In this research, there are 4 expression classes to be classified, namely Angry, Happy, Relaxed and Sad. Where the total data used is 4000 data. The amount of data used is 1000 in each class. In its division, the data is divided into 3 parts, namely training data, validation data, and test data, using a ratio of 80:10:10. This ratio means that 80% of the total data would be used to train the model, 10% would be used to validate the model, and another 10% to test the model. The selection of this ratio aims to provide enough data to the model in the training phase, so that the model can learn patterns from the data and make more accurate results.

Model Architecture

This research applies a deep learning approach using VGG and ResNet architecture in CNN method. The CNN architecture consists of convolution, pooling, and fully connected layers, where the convolution layer acts to identify relevant image features, while the pooling layer is used to reduce image dimensionality [27].

In the initial step, the input image passes through a convolution layer that extracts basic features such as edges and corners. Then, the image passes through an activation layer to perform activation on the output values. After that, the result passes through the pooling layer to reduce the image dimension and reduce the information density.

Next, the image passes through additional convolution and pooling layers to extract increasingly complex features. After the feature extraction stage, the result passes through a fully connected layer that performs classification and determines prediction probabilities by learning the complex relationships among the previously extracted features.

The output layer of the CNN model generates class predictions based on the probabilities that have been learned in the fully connected layer. A softmax activation function is used to generate prediction probabilities that help select the class with the highest probability [28].

Overall, the CNN model architecture has been proven effective in many image recognition tasks, including object recognition, face detection and segmentation [29]. Figure 4 shows the model architecture along with the proposed layers.

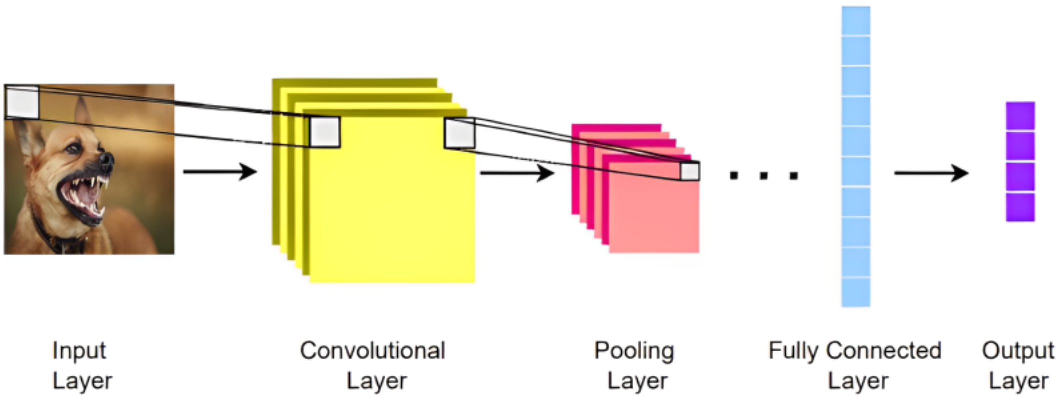

Figure 4. Convolutional Neural Network Architecture

Classification Report and ROC Curve

The developed convolutional neural network model was tested for performance using classification report and ROC Curve. Classification report contains a classification table from testing the correct data, with testing the wrong data. The mathematical equation for finding the value of accuracy, precision, recall and F1 Score. Can be seen in equations (1), (2), (3) and (4) [30]. The formula used to measure performance is:

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+FP+FN+TN}} \quad (1)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (2)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad (3)$$

$$\text{F1 Score} = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

It is known that TP stands for true positive, FP stands for false positive. While TN stands for true negative and FN stands for false negative. On the ROC curve there is a two-dimensional graph that has x and y axes. The "x" axis of the ROC represents the false positive rate (FPR) value. While the "y" axis represents the true positive rate (TPR). Meanwhile, AUC or area under curve is an area with a square shape that has a value of 0-1 [31]. In finding the sensitivity and specificity values, using equations (5) and (6).

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP+FN}} \quad (5)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN+FP}} \quad (6)$$

The AUC has a level in the value of diagnosis/classification as in Table 1 [32].

Table 1. Ranking of AUC Value

Acuracy	Classification Rate
0.90-1.00	Excellent Classification
0.80-0.90	Good Classification
0.70-0.80	Fair Classification
0.60-0.70	Poor Classification
0.50-0.60	Failure

RESULT AND DISCUSSION

After designing the CNN architecture, the next step is to train the model using the data that has been obtained in order to produce optimal results. This training involves the process of optimizing model parameters and weights to improve the accuracy and performance of the model in recognizing patterns and features in the given data.

Comparison of Loss and Training Accuracy

The training process at this stage is carried out to obtain a machine learning model. At the time of training the model was trained using 20 epochs on each of the architectural models used.

The use of 20 epochs was chosen due to the evaluation of model performance and computational efficiency. In the model performance evaluation if a very large number of epochs is used, it is possible that the model would have memorized the training data well and the performance on the training data would be very high. However, it does not guarantee that the performance on test data or data that has never been seen before would be that good. By using a sufficient number of epochs of 20, we can quickly evaluate the model's performance on the test data.

As for computational efficiency, the higher the number of epochs used, the longer it takes to train the model. This is because the model would go through more iterations to update the weights on each batch of data. By using 20 epochs, it can produce adequate results without spending too much computational time.

Table 2. Loss and Accuracy of CNN Model

Epoch	Loss	Accuracy
1/20	loss: 0.9782	accuracy: 0.5766
2/20	loss: 0.6916	accuracy: 0.7291
3/20	loss: 0.5826	accuracy: 0.7663
4/20	loss: 0.4925	accuracy: 0.8131
5/20	loss: 0.4395	accuracy: 0.8306
6/20	loss: 0.3937	accuracy: 0.8500
7/20	loss: 0.3474	accuracy: 0.8681
8/20	loss: 0.3009	accuracy: 0.8963
9/20	loss: 0.2695	accuracy: 0.9041
10/20	loss: 0.2325	accuracy: 0.9153
11/20	loss: 0.2166	accuracy: 0.9250
12/20	loss: 0.1852	accuracy: 0.9372
13/20	loss: 0.1748	accuracy: 0.9416
14/20	loss: 0.1511	accuracy: 0.9475
15/20	loss: 0.1401	accuracy: 0.9516
16/20	loss: 0.1482	accuracy: 0.9494
17/20	loss: 0.1279	accuracy: 0.9581
18/20	loss: 0.1200	accuracy: 0.9544
19/20	loss: 0.1178	accuracy: 0.9609
20/20	loss: 0.1109	accuracy: 0.9644

Based on Table 2, it can be seen for the evaluation results in the form of loss and accuracy values at each epoch of the CNN (Convolution Neural Network) model. Then in Figure 5 is intended for graphical visualization of each loss value and accuracy of each epoch.

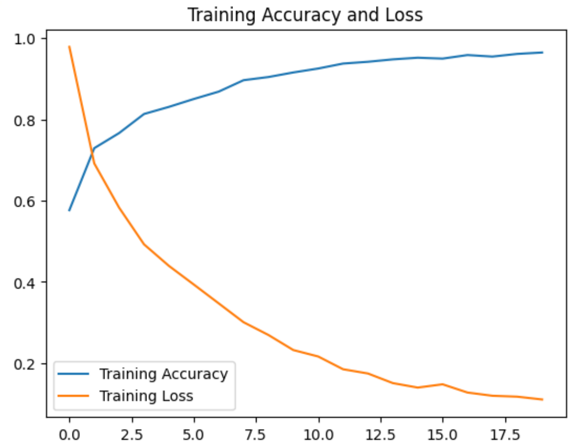

Figure 5. Visualization of CNN model accuracy and loss values

In the evaluation using the Convolution Neural Network (CNN) model, the result is an accuracy value on the test dataset, referred to as `test_accuracy`. The resulting accuracy value is 0.7475, or in percentage terms, the trained model is able to provide information with an accuracy rate of 74.75%.

In this evaluation, the CNN model was tested using a test dataset that was specifically used to test the performance of the trained model. The accuracy value obtained of 74.75% indicates that the CNN model is able to classify data with a high level of accuracy. In other words, 74.75% of the data in the test dataset was successfully classified correctly by the CNN model. In this case, the CNN model has achieved an accuracy of 74.75%, which shows good performance in classifying the data on the test dataset.

Table 3. Loss and Accuracy of ResNet Model

Epoch	Loss	Accuracy
1/20	loss: 2.6127	accuracy: 0.3392
2/20	loss: 2.2667	accuracy: 0.4187
3/20	loss: 1.2212	accuracy: 0.4367
4/20	loss: 1.2316	accuracy: 0.4500
5/20	loss: 1.0748	accuracy: 0.5408
6/20	loss: 0.8623	accuracy: 0.6395
7/20	loss: 0.7435	accuracy: 0.7005
8/20	loss: 0.6227	accuracy: 0.7487
9/20	loss: 0.5465	accuracy: 0.7870
10/20	loss: 0.6456	accuracy: 0.7398
11/20	loss: 0.8500	accuracy: 0.6497
12/20	loss: 0.6199	accuracy: 0.7613
13/20	loss: 0.4091	accuracy: 0.8355
14/20	loss: 0.3590	accuracy: 0.8633
15/20	loss: 0.3749	accuracy: 0.8615
16/20	loss: 0.3642	accuracy: 0.8593
17/20	loss: 0.2938	accuracy: 0.8938
18/20	loss: 0.3580	accuracy: 0.8618
19/20	loss: 0.2434	accuracy: 0.9078
20/20	loss: 0.3650	accuracy: 0.8665

Based on Table 3, it can be seen for the evaluation results in the form of loss and accuracy values at each epoch of the ResNet50 (Residual Network) model. Then in Figure 6 is intended for graphical visualization of each loss value and accuracy of each epoch.

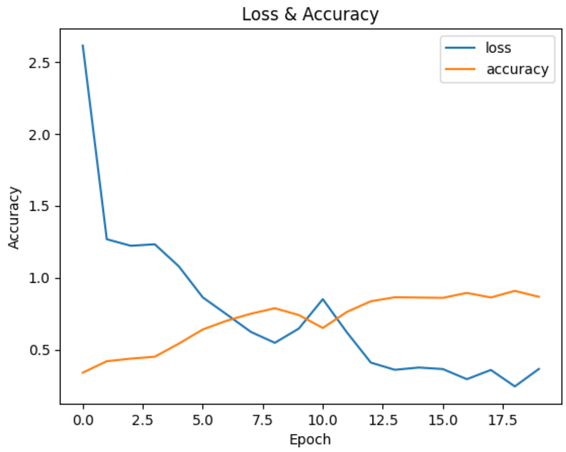

Figure 6. Visualization of ResNet model accuracy and loss values

The model was built, trained, and evaluated using the ResNet50 architecture, i.e. using TensorFlow and Keras. The model was set to not include a full layer at the end (`include_top=False`), and the input image dimensions were set to (192, 192, 3). Classification was performed on each output layer of the resnet model by adding Flatten, Dense layers with 128 units and ReLU activation, and a final Dense layer with the number of units corresponding to the number of classes in the dataset and softmax activation.

The results of the model evaluation using CNN architecture, ResNet50, obtained a loss value of 0.8482 and an accuracy value of 0.6510. Which shows that the accuracy information on the evaluation model that has been trained is 65.1%.

Table 4. Loss and Accuracy of VGG Model

Epoch	Loss	Accuracy
1/20	loss: 2.6512	accuracy: 0.2810
2/20	loss: 1.3080	accuracy: 0.3433
3/20	loss: 1.3094	accuracy: 0.3352
4/20	loss: 1.2689	accuracy: 0.3665
5/20	loss: 1.2765	accuracy: 0.3645
6/20	loss: 1.2531	accuracy: 0.3995

7/20	loss: 1.2393	accuracy: 0.3985
8/20	loss: 1.2281	accuracy: 0.4092
9/20	loss: 1.2181	accuracy: 0.4142
10/20	loss: 1.2045	accuracy: 0.4232
11/20	loss: 1.2009	accuracy: 0.4238
12/20	loss: 1.1768	accuracy: 0.4510
13/20	loss: 1.1525	accuracy: 0.4715
14/20	loss: 1.1293	accuracy: 0.4852
15/20	loss: 1.0130	accuracy: 0.5390
16/20	loss: 0.9750	accuracy: 0.5738
17/20	loss: 0.9190	accuracy: 0.6000
18/20	loss: 0.8973	accuracy: 0.6093
19/20	loss: 0.8563	accuracy: 0.6265
20/20	loss: 0.8316	accuracy: 0.6390

Based on the evaluation results in Table 4, it can be seen that changes in the loss value of the VGG16 model tend to be stable and consistent every epoch. This shows that the model is able to learn the patterns in the data well. In addition, the movement of the accuracy value also shows a similar pattern to the loss value, which rises consistently and stably. This indicates that the model has the ability to recognize and classify data with a high level of accuracy.

To better visualize the movement of loss and accuracy values, Figure 7 can be used as a reference. The graph shows the trend of decreasing loss value and increasing accuracy value as the epoch goes by. It also shows the correspondence between the loss and accuracy values obtained from Table 4.

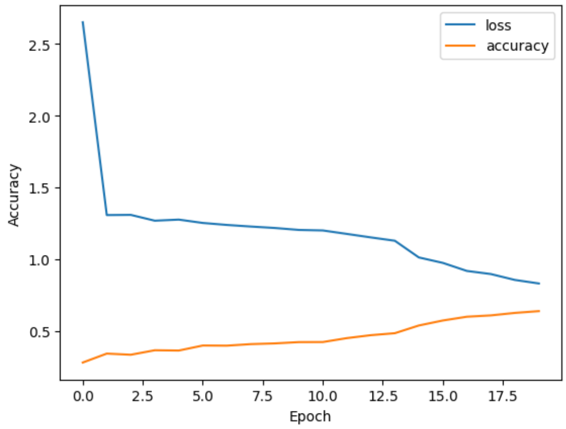

Figure 7. Visualization of VGG model accuracy and loss values

The same was done with the VGG-16 architecture. The evaluation model using CNN Architecture is VGG16, with the evaluation results achieving a loss of 0.7348 and an accuracy of 0.6867. The information shows that the trained model has an accuracy rate of 68.67%. This shows that the model using the VGG16 image recognition method provides quite good results, although there is still room for improvement.

Figure 8. Comparison Chart of Accuracy and Loss

ISSN 2089-8673 (Print) | ISSN 2548-4265 (Online)
Volume 13, Issue 2, July 2024

The results showed that the CNN model achieved a high level of accuracy in classifying emotions in dogs. Model performance was measured using standard evaluation metrics such as accuracy. Figure 8 compares the performance of CNN, ResNet and VGG models in terms of loss and accuracy. The results show that CNN performance has a higher accuracy value compared to ResNet and VGG.

ResNet and VGG are the basic methods chosen for image processing because they are related to deeper feature extraction and have good accuracy performance and good generalization ability. The architectures of VGG and ResNet have many deep convolution layers, which enable them to identify complex image features. These layers form a hierarchical representation of the image, from basic features such as lines and corners to more complex features such as objects and faces. This capability has proven to be effective in improving the performance of algorithms on image classification tasks.

In addition, VGG and ResNet also have good accuracy rates and generalization capabilities. As explained in the previous sentence, both have deep convolution layers. VGG has deep convolution and max pooling layers to extract complex features, while ResNet with shortcut connections is able to retain better visual information through the network layers, which theoretically can lead to more accurate classification results. ResNet and VGG have a good ability to generalize on image datasets. Generalization on image datasets is the process of transforming image data into a more general or abstract form. This is done by combining or reducing the features present in the image to form a simpler and more compact representation. By generalizing image data, we can obtain more information about the image, which can be used for various purposes such as pattern recognition and image classification.

They can recognize common patterns and features in images, thus classifying never-before-seen images with good accuracy. This makes them a good choice for systems that often have to classify images from different sources. Both ResNet and VGG have shown excellent results in various image processing tasks, including classification. The choice of using ResNet or VGG in a CNN algorithm may depend on the complexity of the dataset to be used and the specific context of the image processing task at hand.

The results also show the success of the CNN model in identifying dog emotions based on diverse image datasets. In this case the model is able to understand complex patterns in dog facial expressions. This can be seen from the significant level of accuracy in testing the model against the dataset. CNNs enable deep hierarchical feature extraction from images. The convolution layer in the CNN architecture plays an important role in identifying small details on the dog's face that characterize various emotions. Therefore, the model can understand the changes in the eyes, nose, and mouth that together form an emotional expression.

Variability in lighting conditions, position, and dog breed are factors that can affect emotion classification. To address this, further development of the model is required, as well as the addition of more diverse training data that represents real-world conditions.

The classification of dog emotions using CNN has positive implications for the understanding and interaction between humans and pets. Potential applications include pet mental health monitoring systems, behavioral understanding, and selection of more personalized care approaches. In addition, this technology can be integrated into mobile applications or smart devices that provide dog owners with deeper insights into the emotional state of pets.

Testing Results

In this testing, testing was carried out using 10% of the training data that had been provided. The training data consists of dog emotion images that have gone through augmentation and preprocessing processes. Taking test data is done with the aim of testing the extent to which the classification performed can predict dog emotions.

Table 5. Testing Dataset

Types of Emotion	Image Count
Angry	100
Happy	100
Relaxed	100
Sad	100
Total	**400**

For each class, 100 dog emotion images were tested. In total, there are several classes tested in this test. All tested images are images of dogs with various emotional states that have gone through augmentation and preprocessing.

In Table 5, the number of test data for each class is presented. It is important to see the distribution of the amount of test data and ensure that each class has comparable representation in this test. The amount of test data taken for each class is 10% of the total training data.

ISSN 2089-8673 (Print) | ISSN 2548-4265 (Online)
Volume 13, Issue 2, July 2024

Augmentation and preprocessing were done before the test data was taken. Augmentation is done with the aim of creating a greater variety of images so that the model can see a wide variety of images and can compare with images not seen before. Preprocessing is done to prepare the training data and test data to be ready for use in the classification process.

After the augmentation and preprocessing process is complete, the training data and test data are ready for use. The training data is given to the classification model to be trained so that the model can learn the patterns in the data to predict the dog's emotions. Once trained, the model is tested using previously prepared test data.

The results of this test would provide an overview of the performance and accuracy of the classification model that has been trained. By looking at the amount of test data for each class in Table 5, it can be seen whether the distribution of test data is evenly distributed so as to ensure that this test is carried out fairly.

Table 6. Classification Report

	Precision	Recall	F1-Score	Support
Angry	0.90	0.77	0.83	100
Happy	0.71	0.78	0.74	100
Relaxed	0.75	0.68	0.71	100
Sad	0.77	0.87	0.82	100
Accuracy	-	-	0.78	400
Macro Avg	0.78	0.78	0.77	400
Weighted Avg	0.78	0.78	0.77	400

In Table 6, the results of the tests conducted to evaluate the value or level of accuracy, precision, f1-score and recall.

Conducting this test is very important in this research, because the results would be an indicator of how well the model performs in classifying data. Accuracy in this case describes the extent to which the model can predict correctly. Precision measures how many positive predictions are correct, while recall shows the extent to which the model can identify correctly.

In Table 6, the test results would be recorded and analyzed to evaluate the performance of the model. By looking at the accuracy, precision, f1-score and recall values, researchers can see the extent to which the model can effectively perform the desired task. Good test results would provide confidence that the proposed model is effective in performing data classification.

Through this testing, it is hoped to provide a better understanding of the model's capabilities and possible improvements that can be made in the future. These improvements could include training the model with more diverse data, using more optimized algorithms, or more careful data processing. Thus, this research can make a meaningful contribution to the field of algorithm development and practical applications related to data classification.

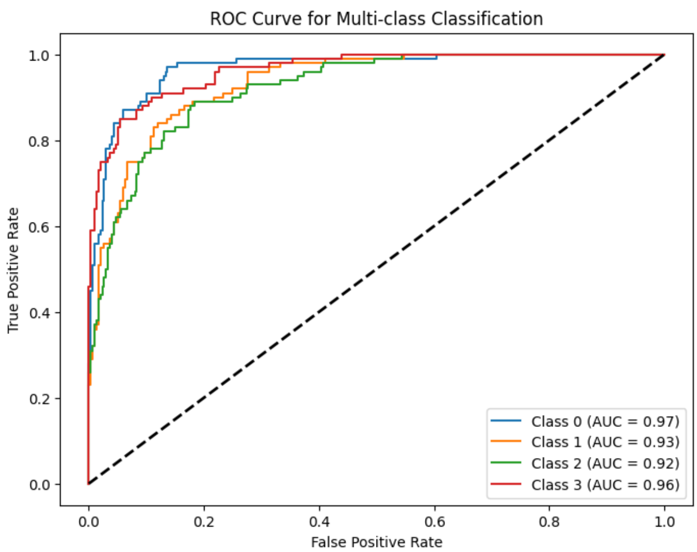

Figure 9. ROC and AUC Evaluation Results of each Class

The results in Figure 9 show the ROC (Receiver Operating Characteristic) and AUC (Area Under the Curve) plots. In the figure, there are four classes tested, namely angry (class 0), happy (class 1), relaxed (class 2), and sad (class 3). The results show that the highest AUC value is obtained by the angry class with a value of 0.97, while the lowest AUC value is obtained by the relaxed class with a value of 0.92. The AUC values can be seen in Table 7.

Table 7. AUC Value of each Class

Class	AUC value
Angry/Class 0	0.97
Happy/Class 1	0.93
Relaxed/Class 2	0.92
Sad/Class 3	0.96

ROC and AUC analysis are important in evaluating the model's performance in processing classification data. The higher the AUC value, the better the model's ability to distinguish between classes.

When looking at the results in Figure 9, it can be concluded that the model has a good ability to distinguish between the classes. The angry class (class 0) has the highest AUC value of 0.97, which indicates that the model is able to identify with high accuracy whether the emotion displayed in the image is anger. The happy class (class 1) has an AUC value of 0.93, which also shows good performance in distinguishing between happy facial expressions. The relaxed class (class 2) has the lowest AUC value of 0.92, which means the model has a slightly higher error rate in recognizing relaxed facial expressions. The sad class (class 3) has an AUC value of 0.96, which shows the model's ability to recognize sad facial expressions.

However, it is important to note that the results in Figure 9 only show the overall performance of the model in distinguishing emotions based on facial expressions. It does not include a detailed evaluation of the model performance for each evaluation metric such as accuracy, precision, F1-Score, and recall for each class.

In the previous test, it is known that the accuracy of the angry class is quite high, but the highest values for precision and F1-Score are obtained by the Angry class. Meanwhile, the highest value for recall was obtained by the Sad class. This shows that the model is able to recognize well the expression of anger, but for sad expressions the model's ability is slightly lower.

Overall, the results of this research show that the model has a good performance in distinguishing emotions based on facial expressions. The high AUC value indicates the model's ability to distinguish between the classes. However, further evaluation is needed to measure other evaluation metrics and conduct a more in-depth analysis of each class.

CONCLUSION

The use of CNN in dog emotion classification shows great potential in improving pet emotional understanding. The CNN model had higher accuracy (74.75%) compared to the ResNet (65.10%) and VGG (68.67%) models. The model also has a good ability to distinguish emotion classes using the ROC-AUC method. The angry class has the highest AUC value (0.97), indicating the model's ability to recognize anger with high accuracy. The happy (AUC 0.93), sad (AUC 0.96), and relaxed (AUC 0.92) classes also performed well in recognizing happy, sad, and relaxed facial expressions. In the previous test, the angry class had the highest precision and F1-Score values, while the sad class had the highest recall value. Overall, this research shows that the model has good performance in distinguishing emotions based on facial expressions.

The suggestions that researchers can give as further development are for further evaluation and analysis to measure the metrics of each class in depth. As well as looking at other variables such as lighting conditions, position, and dog breed to be factors that affect emotion classification. To overcome this, further development of the model is needed, as well as the addition of training data that is more diverse and represents real-world conditions.

REFERENCES

[1] R. D. Hawkins, B. D. M. Hatin, dan E. O. Revesz, “Accuracy of Canine vs. Human Emotion Identification: Impact of Dog Ownership and Belief in Animal Mind,” *Human-animal Interact. Bull.*, vol. 2021, no. 2, hal. 1–18, 2021, doi: 10.1079/hai.2021.0001.

[2] D. Mota-Rojas *et al.*, “Current advances in assessment of dog’s emotions, facial expressions, and their use for clinical recognition of pain,” *Animals*, vol. 11, no. 11, 2021, doi: 10.3390/ani11113334.

[3] M. Walker, “Dogs and Cats,” *Hitch. Motifs*, hal. 142–145, 2019, doi: 10.1515/9789048505456-016.

[4] L. Lazarowski *et al.*, “Selecting Dogs for Explosives Detection: Behavioral Characteristics,” *Front. Vet. Sci.*, vol. 7, no. September, 2020, doi: 10.3389/fvets.2020.00597.

[5] F. Amici, J. Waterman, C. M. Kellermann, K. Karimullah, dan J. Bräuer, “The ability to recognize dog emotions depends on the cultural milieu in which we grow up,” *Sci. Rep.*, vol. 9, no. 1, hal. 1–10, 2019, doi: 10.1038/s41598-019-52938-4.
[6] C. Correia-Caeiro, K. Guo, dan D. S. Mills, “Visual perception of emotion cues in dogs: a critical review of methodologies,” *Anim. Cogn.*, vol. 26, no. 3, hal. 727–754, 2023, doi: 10.1007/s10071-023-01762-5.
[7] P. Jain, A. K. Pandey, V. Manoj, K. Pandey, dan B. Yadav, “Facial Emotion Recognition of Cat Breeds by Using Convolution Neural Network,” vol. 7, no. 6, hal. 58–63, 2023.
[8] M. Vishwakarma, K. Jangid, K. Pal, dan L. Rodrigues, “Cat Breed & Emotion Detection Using Yolo , CNN & Canny Edge Detection,” vol. 7, no. 8, hal. 1–9, 2024.
[9] A. Quaranta, S. D’ingeo, R. Amoruso, dan M. Siniscalchi, “Emotion recognition in cats,” *Animals*, vol. 10, no. 7, hal. 1–13, 2020, doi: 10.3390/ani10071107.
[10] T. Boneh-Shitrit *et al.*, “Deep Learning Models for Automated Classification of Dog Emotional States from Facial Expressions,” hal. 1–6, 2022, [Daring]. Tersedia pada: http://arxiv.org/abs/2206.05619
[11] A. Khan *et al.*, “Knowledge, attitude & practices (KAPs) regarding rabies endemicity among the community members, Pakistan,” *Acta Trop.*, vol. 200, no. July, hal. 105156, 2019, doi: 10.1016/j.actatropica.2019.105156.
[12] W. Muldayani, “Implementasi Sistem Object Tracking Untuk Mendeteksi Dua Objek Berbasis Deep Learning,” *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 14, no. 1, hal. 1–14, 2023, doi: 10.24176/simet.v14i1.9236.
[13] S. Aich, S. Chakraborty, J. S. Sim, D. J. Jang, dan H. C. Kim, “The design of an automated system for the analysis of the activity and emotional patterns of dogs with wearable sensors using machine learning,” *Appl. Sci.*, vol. 9, no. 22, hal. 1–22, 2019, doi: 10.3390/APP9224938.
[14] A. Hussain, S. Ali, Abdullah, dan H. C. Kim, “Activity Detection for the Wellbeing of Dogs Using Wearable Sensors Based on Deep Learning,” *IEEE Access*, vol. 10, hal. 53153–53163, 2022, doi: 10.1109/ACCESS.2022.3174813.
[15] V. Franzoni, A. Milani, G. Biondi, dan F. Micheli, “A preliminary work on dog emotion recognition,” *Proc. - 2019 IEEE/WIC/ACM Int. Conf. Web Intell. Work. WI 2019 Companion*, hal. 91–96, 2019, doi: 10.1145/3358695.3361750.
[16] Y. Mao dan Y. Liu, “Pet dog facial expression recognition based on convolutional neural network and improved whale optimization algorithm,” *Sci. Rep.*, vol. 13, no. 1, hal. 1–20, 2023, doi: 10.1038/s41598-023-30442-0.
[17] N. H. Pratama, E. Rachmawati, dan G. Kosala, “Classification of Dog Breeds From Sporting Groups Using Convolutional Neural Network,” *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 7, no. 4, hal. 1080–1087, 2022, doi: 10.29100/jipi.v7i4.3208.
[18] J. Liang, “Image classification based on RESNET,” *J. Phys. Conf. Ser.*, vol. 1634, no. 1, 2020, doi: 10.1088/1742-6596/1634/1/012110.
[19] S. Tammina, “Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images,” *Int. J. Sci. Res. Publ.*, vol. 9, no. 10, hal. p9420, 2019, doi: 10.29322/ijsrp.9.10.2019.p9420.
[20] Y. Wu, “Emotion Detection of Dogs and Cats Using Classification Models and Object Detection Model Emotion Detection of Dogs and Cats Using Classification Models and Object Detection Model,” no. May, 2023.
[21] S. Hwang, H. K. Shin, J. M. Park, B. Kwon, dan M. G. Kang, “Classification of dog skin diseases using deep learning with images captured from multispectral imaging device,” *Mol. Cell. Toxicol.*, vol. 18, no. 3, hal. 299–309, 2022, doi: 10.1007/s13273-022-00249-7.
[22] H. Y. Chen, C. H. Lin, J. W. Lai, dan Y. K. Chan, “Convolutional Neural Network-Based Automated System for Dog Tracking and Emotion Recognition in Video Surveillance,” *Appl. Sci.*, vol. 13, no. 7, hal. 1–29, 2023, doi: 10.3390/app13074596.
[23] A. Mutoi Siregar, S. Faisal, H. H. Handayani, dan A. Jalaludin, “Classification Data for Direct Marketing using Deep Learning,” *Core.Ac.Uk*, vol. 7, no. 2, 2020, doi: 10.27512/sjppi-ukm/se/a15052020.
[24] E. M. C. Bouma, M. L. Reijgwart, P. Martens, dan A. Dijkstra, “Cat owners’ anthropomorphic perceptions of feline emotions and interpretation of photographs,” *Appl. Anim. Behav. Sci.*, vol. 270, no. November 2023, hal.

106150, 2024, doi: 10.1016/j.applanim.2023.106150.

[25] S. Pada Plat Kendaraan Dinas dengan Metode Deteksi Tepi Canny, R. Rizal Adi Saputra, dan F. Mimi Wahyuni, "Segmentasi Pada Plat Kendaraan Dinas Menggunakan deteksi Tepi (Reskal) |328," *J. Sains Komput. Inform. (J-SAKTI*, vol. 6, no. 1, hal. 328–339, 2022.

[26] R. Fakhri, I. R. Magdalena, R. Yunendah, dan N. Fu'adah, "Deteksi Status Gizi Balita Dengan Estimasi Antropometri Menggunakan Metode Deteksi Tepi Berbasis Pengolahan Citra Digital Child'S Nutritional Status Detection With Anthropometry Estimation Using Edge Detection Method Based on Digital Image Processing," vol. 6, no. 1, hal. 642–649, 2019.

[27] L. M. Prayogo dan S. Sarono, "Automated Extraction of Shoreline in Tuban Regency, East Java from Google Earth Imagery by Integrating Canny Edge Detector," *DoubleClick J. Comput. Inf. Technol.*, vol. 5, no. 2, hal. 81, 2022, doi: 10.25273/doubleclick.v5i2.11056.

[28] M. R. Wathani, N. Hidayati, U. Islam, K. Muhammad, dan A. Al, "Analisis Perbandingan Fungsi Aktivasi CNN Pada Beras," vol. 4, no. 2, hal. 144–153, 2023.

[29] A. Akram, S. A. Rachmadinasya, F. H. Melvandino, dan H. Ramza, "Klasifikasi Aktivitas Olahraga Berdasarkan Citra Foto Dengan Menggunakan Metode Convolutional Neural Network," *J. Inform. dan Tek. Elektro Terap.*, vol. 11, no. 3s1, 2023, doi: 10.23960/jitet.v11i3s1.3496.

[30] F. Z. Rahmanti, B. A. S. Aji, O. A. Permata, B. Amelia, dan M. H. Z. Al Faroby, "Leaf Health Identification on Melon Plants Using Convolutional Neural Network," *J. Nas. Pendidik. Tek. Inform.*, vol. 12, no. 1, hal. 127–134, 2023, doi: 10.23887/janapati.v12i1.58492.

[31] M. Dennis, R. Rahmaddeni, F. Zoromi, dan M. K. Anam, "Penerapan Algoritma Naïve Bayes Untuk Pengelompokkan Predikat Peserta Uji Kemahiran Berbahasa Indonesia," *J. Media Inform. Budidarma*, vol. 6, no. 2, hal. 1183, 2022, doi: 10.30865/mib.v6i2.3956.

[32] R. Prayogo, D. Anggraeni, dan A. F. Hadi, "Classification of Cardiovascular Disease Gene Data Using Discriminant Analysis and Support Vector Machine (SVM)," *Berk. Sainstek*, vol. 10, no. 3, hal. 124, 2022, doi: 10.19184/bst.v10i3.22259.