E-ISSN : 2540 - 8984

**JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)**
Volume 07, Nomor 04, Desember 2022 : 1080 - 1087

# CLASSIFICATION OF DOG BREEDS FROM SPORTING GROUPS USING CONVOLUTIONAL NEURAL NETWORK

**Naufal Harsa Pratama[1], Ema Rachmawati[2], and Gamma Kosala[3]**
[1, 2, 3]School of Computing, Telkom University
Jl. Telekomunikasi No. 1, Sukapura, Dayeuhkolot, Bandung, Jawa Barat, Indonesia
e-mail: naufalharsap@student.telkomuniversity.ac.id[1], emarachmawati@telkomuniversity.ac.id[2], gammakosala@telkomuniversity.ac.id[3]

**ABSTRAK**

*Penggunaan convolutional neural network sudah diterapkan terhadap berbagai macam aplikasi. Seperti dari pengklasifikasian citra, pendeteksi dan pengenalan objek, dan lainnya. Pengklasifikasian citra merupakan salah satu aplikasi neural network yang paling umum. Pengklasifikasian citra utamanya dilakukan untuk mengidentifikasi dan mengategorikan citra sesuai dengan kelompok yang ditetapkan. Salah satu penerapannya adalah untuk membedakan antara satu jenis anjing dengan lainnya. Pengklasifikasian jenis anjing memiliki tantangan tersendiri karena terapat beberapa jenis anjing yang memiliki kemiripan ciri fisik, terutama jenis anjing dalam satu grup tertentu. Penelitian ini menjelaskan bagaimana cara untuk mengembangkan sistem pengklasifikasian jenis anjing dari grup sporting dengan menggunakan residual neural network (Res-Net). Sistem ditujukan untuk lebih memudahkan manusia dalam membedakan jenis anjing tersebut. Digunakan lima jenis atau kelas anjing yang diambil dari dataset Tsinghua Dogs dataset. Dalam penerapannya, digunakan dua varian dari CNN untuk dibandingkan, yaitu ResNet 50 dan ResNet 101, dengan menggunakan konfigurasi yang sama. Berdasarkan hasil penelitian, ResNet 101 menunjukkan hasil rata-rata makro f1-score yang lebih baik dengan tetap mempertahankan akurasi yang tinggi. Varian ResNet 50 menghasilkan f1-score sebesar 84%, sedangkan ResNet 101 menciptakan hasil f1-score 86%.*

***Kata Kunci****: convolutional neural network, image classification, jenis anjing, residual network.*

**ABSTRACT**

*The use of convolutional neural networks has been applied to various applications. Such as image classification, object detection and recognition, and others. One of the most popular uses for neural networks is image classification. Image classification mainly identifies and categorizes images according to the specified group. One application is to distinguish between one type of dog to another. Classification of dog breeds has its challenges because several kinds of dogs have similar physical characteristics, especially those that belong to the same group. This study explains how to develop a dog breed classification system from a sporting group using a residual neural network (ResNet). The system's goal is to make it simpler for people to identify the dog breed. Five types of dog breeds were used, which were obtained from the Tsinghua Dogs dataset. In its implementation, two variants of CNN are used to be compared, ResNet 50 and ResNet 101, using the same configuration. Based on the research results, ResNet 101 shows better macro-average f1-score results while maintaining high accuracy. The ResNet 50 produces an f1-score of 84%, while ResNet 101 makes an f1-score of 86%.*

***Keywords****: convolutional neural network, dog breeds, image classification, residual network.*

## I. INTRODUCTION

LATELY, we have seen various usage of neural network applications, for example, a recommendation system, object recognition or detection, and image classification [1]. Using a neural network allows these applications to be carried out more quickly and gives better accuracy than traditional methods [2].

In image classification, Convolutional Neural Network (CNN) learns the features owned by the data during training. It has been applied in various fields. For example, it is used to classify medical images [3], determine indoor and outdoor scenes [4], and many more. However, one object that is quite difficult to differentiate is the type of dog, especially a group of dogs with similar size, shape, color, and other physical characteristics. It happens because CNN is difficult to study the specific features owned by the object to be classified [5].

Several studies have been conducted in classifying dog breeds. These studies used various methods and approaches to classify dog breeds. The process is made using a conventional algorithm and a neural network. The average accuracy value obtained from these studies is very high (above 85%) [5]–[7]. However, there is still no research on the classification of dog breeds against one particular group.

In using the conventional approach, classification is done by different feature extraction methods. One of them is to use part localization. By using this approach, the degree of classification accuracy can be increased with parts

E-ISSN : 2540 - 8984

**JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)**
Volume 07, Nomor 04, Desember 2022 : 1080 - 1087

that fit around the dog [8]. In addition, it is also carried out using landmark-based representations around the dog image. The accuracy performance of this approach can rival other classification methods with feature extraction [7].

Meanwhile, the use of neural networks is generally based on the CNN model. Researchers tend to use several different architectures of CNN, such as VGG-16 [9], LeNet, GoogLeNet [5], and yolov3 [6]. Using neural networks could make the accuracy value higher than other traditional methods [2].

The datasets used in the research also vary. Some collect or download dog images [8], and others use the Stanford Dogs dataset [5], which consists of 120 dog classes or breeds with a total of 20.580 images [10]. However, in 2020, Tsinghua University introduced a new dog breed image dataset consisting of 70.428 images with 130 dog breeds called the Tsinghua Dogs dataset [11].

In this work, we propose a sporting group dogs classification using CNN. We analyze two kinds of ResNet architectures to classify five types of retriever dogs that belong to the sports group [12].

## II. Research Methods

This section describes the research methods used in building the classification system, as shown in Figure 1.

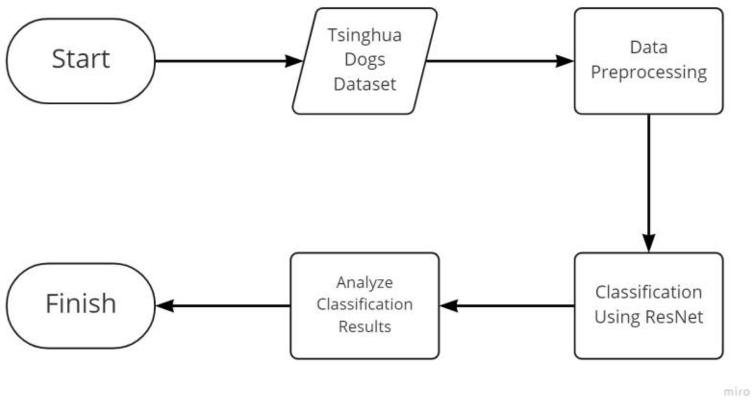


Fig. 1. Dog breed classification system workflow

### *A. Dataset*

We use the Tsinghua Dogs dataset to build our dogs classification system. As mentioned previously, it has 130 dog breeds with 70,428 images. In addition, annotations are also available in class labels and bounding boxes of the dog's head and body. This dataset provides images with high and low resolution.

We use five types of dog breeds in building the system: Golden Retriever, Labrador Retriever, Chesapeake Bay Retriever, Flat-coated Retriever, and Curly-coated Retriever. The number of images per class used can be seen in Table 1. Examples of images of each class can be seen in Figure 2 to Figure 6.

TABLE I
DATASET

| Class | Quantity |
|---|---|
| Chesapeake Bay Retriever | 215 |
| Curly Coated Retriever | 202 |
| Flat Coated Retriever | 206 |
| Golden Retriever | 5,355 |
| Labrador Retriever | 3,580 |
| Total Data | 9,558 |

E-ISSN : 2540 - 8984

**JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)**
Volume 07, Nomor 04, Desember 2022 : 1080 - 1087

Fig. 2. Golden Retriever

Fig. 3. Labrador Retriever

Fig. 4. Cheasapeake Bay Retriever

Fig. 5. Flat-coated Retriever

Fig. 6. Curly-coated Retriever

These five types of dogs are among the most common dog breed groups, namely the sporting group type [12]. In addition, these breeds have a similar appearance, distinguished by color, coat or skin type, and others. Sometimes it is hard to tell those differences, especially for ordinary people.

The dataset is divided into train, validation, and test at this stage. The datasets are split randomly with a ratio of 70%, 10%, and 20% for train, validation, and test data. The results of the division are listed in Table 2. In this study, no data preprocessing was carried out on the dataset.

TABLE II
DATASET DISTRIBUTION

| Class | Number of Data (Images) | | |
|---|---|---|---|
| | Train | Valid | Test |
| Chesapeake Bay Retriever | 150 | 21 | 44 |
| Curly Coated Retriever | 141 | 20 | 41 |
| Flat Coated Retriever | 144 | 20 | 42 |
| Golden Retriever | 3,748 | 535 | 1,072 |
| Labrador Retriever | 2,506 | 358 | 716 |
| Total Data | 6,689 | 954 | 1,915 |

## *B. Residual Neural Network Architecture*

Residual neural network (ResNet) is one of the architectures of the CNN model developed by Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. This development is motivated by the increasing difficulty of training deeper neural networks. ResNet aims to create a deeper network by using the residual function. This function is easier to optimize and can improve accuracy with greater depth [13].

E-ISSN : 2540 - 8984

**JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)**
Volume 07, Nomor 04, Desember 2022 : 1080 - 1087

ResNet is divided into several variants, one of the most famous being ResNet34, ResNet50 and ResNet101. These variants are distinguished by the number of neural network layers they have. For example, ResNet50 has 50 neural network layers, ResNet 101 has 101 neural network layers, and so on.

ResNet works by using residual or building blocks. Residual learning is carried out in several overlapping layers. The residual block is shown in Figure 7.

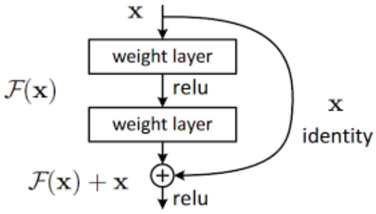


Fig. 7. Building Blocks [13]

The building block is defined as (1), where x and y are the input and output vectors of the layer. The function F (x, {Wi}) represents the residual mapping to be studied [13].

$$y = F(x, \{W_i\}) + x \tag{1}$$

This research uses ResNet 50 and ResNet 101. Using ResNet architecture, we could prevent the increasing value of training and testing loss when using a large number of layers, which usually happens using other CNN models. ResNet uses an identity shortcut connection to skip specific layers that could reduce the performance [13]. The difference between ResNet 50 and ResNet 101 is the number of layers. ResNet 101 has an additional 51 layers compared to ResNet 50, thus making it potentially slower to train, but it should also give better results than ResNet 50. Convolutional layers are used in both ResNet variants, and dense, dropout, and pooling layers were added.

The results of the classification of the two architectures are then compared using a performance measure. System development is carried out using the TensorFlow library. Both models use a pre-trained model on the ImageNet.

### *C. Performance Measure*

F1-score is used as the primary metric to measure the performance classification results. In addition, metrics that can be obtained from confusion metrics are also used: accuracy, precision, and recall [14]. The overall formula is obtained based on: TP (true positive), TN (true negative), FP (false positive), and FN (false negative) [15]. The following equations are the formula for the four metrics. Overall accuracy is shown in (2), precision in (3), recall in (4), and f1-score in (5).

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{2}$$

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

$$F1 = 2 * \frac{precision * recall}{precision + recall} \tag{5}$$

True indicates that the predicted data is correct in both positive and negative classes. In comparison, false indicates that the predicted data is wrong, both in the positive and negative classes. It also applies to multi-class classification with positive and negative values replaced based on existing classes.

## III. RESULT

This section explains the result during model training, using data train and validation, and testing using data tests. During training, validation accuracy and loss are the two main metrics to measure in each epoch. While in the testing phase, the performance measure mentioned previously is used as a measurement.

As mentioned previously, this research used two ResNet variants, ResNet 50 and Resnet 101. Both used the same configuration: image input is resized to 224x224 by width and height, with 32 batch size, twenty-five epochs, and four neural network layers consisting of two dense layers with 512 and two dropout layers with 0.3 rate used alternately.

E-ISSN : 2540 - 8984

**JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)**
Volume 07, Nomor 04, Desember 2022 : 1080 - 1087

The results from training the model ResNet 50 and ResNet 101 are visualized in Figure 8 and Figure 9. A more detailed version also can be seen in Table 3.

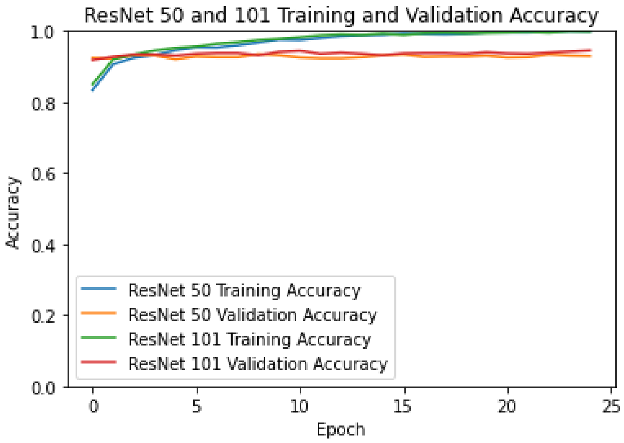


Fig. 8. Training and Validation Accuracy Comparison using ResNet 50 & ResNet 101

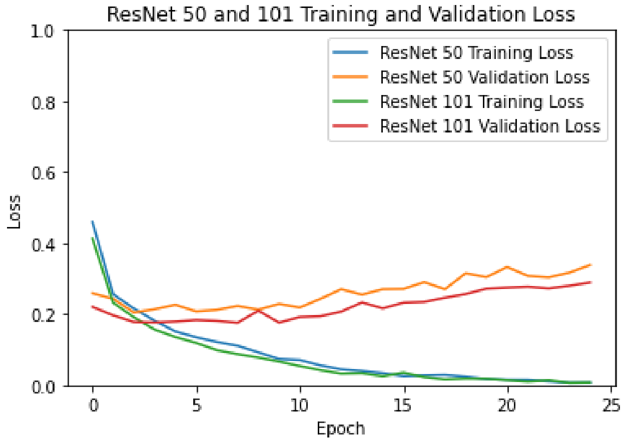


Fig. 9. Training and Validation Loss Comparison using ResNet 50 & ResNet 101

As can be seen from Figure 8, Figure 9, and Table 3, the performance from ResNet 101 is significantly better compared to ResNet 50 in both validation accuracy and validation loss. The validation accuracy in ResNet 50 averaged at 0.9283, with the highest of 0.9350. On the other hand, the ResNet 101 averaged at 0.9363 with the highest validation accuracy of 0.9455. Regarding validation loss, the ResNet 50 scored 0.2045 for the lowest value with an average of 0.2614. while the ResNet 101 averaged 0.2220 with the lowest of 0.1757. It is also to be noticed that when using the ResNet 101, the fluctuations in validation loss happened less compared to using ResNet 50.

TABLE III
TRAINING RESULTS

| Epoch | ResNet 50 | | ResNet 101 | |
|---|---|---|---|---|
| | val_acc | val_loss | val_acc | val_loss |
| 1 | 0.9245 | 0.2588 | 0.9182 | 0.2205 |
| 2 | 0.9224 | 0.2431 | 0.9277 | 0.1965 |
| 3 | 0.9319 | 0.2045 | 0.9340 | 0.1778 |
| 4 | 0.9319 | 0.2138 | 0.9340 | 0.1765 |
| 5 | 0.9203 | 0.2262 | 0.9308 | 0.1791 |
| 6 | 0.9287 | 0.2076 | 0.9361 | 0.1835 |
| 7 | 0.9266 | 0.2121 | 0.9382 | 0.1810 |
| 8 | 0.9266 | 0.2230 | 0.9382 | 0.1757 |
| 9 | **0.9350** | **0.2126** | 0.9319 | 0.2096 |
| 10 | 0.9319 | 0.2286 | 0.9413 | 0.1762 |
| 11 | 0.9256 | 0.2186 | 0.9444 | 0.1917 |
| 12 | 0.9235 | 0.2432 | 0.9361 | 0.1941 |
| 13 | 0.9235 | 0.2701 | 0.9392 | 0.2071 |
| 14 | 0.9266 | 0.2547 | 0.9361 | 0.2329 |
| 15 | 0.9319 | 0.2701 | 0.9319 | 0.2163 |
| 16 | 0.9340 | 0.2709 | 0.9371 | 0.2322 |
| 17 | 0.9277 | 0.2901 | 0.9382 | 0.2344 |
| 18 | 0.9287 | 0.2701 | 0.9382 | 0.2462 |
| 19 | 0.9287 | 0.3150 | 0.9361 | 0.2565 |
| 20 | 0.9308 | 0.3043 | 0.9403 | 0.2716 |
| 21 | 0.9256 | 0.3326 | 0.9371 | 0.2745 |
| 22 | 0.9266 | 0.3077 | 0.9361 | 0.2770 |
| 23 | 0.9340 | 0.3036 | 0.9392 | 0.2723 |
| 24 | 0.9308 | 0.3168 | 0.9423 | 0.2797 |
| 25 | 0.9298 | 0.3381 | **0.9455** | **0.2891** |
| average | 0.9283 | 0.2614 | 0.9363 | 0.2220 |

Based on the training result, the models with the highest validation accuracy from both variants are chosen to be used to test using test data. In this case, the 9$^{th}$ epoch model from ResNet 50 has a validation accuracy of 0.9350 and a validation loss of 0.2126. Also, the last epoch model from ResNet 101 with validation accuracy of 0.9455 and validation loss of 0.2891 (marked as bold in Table 3) is chosen.

Table 4 shows two confusion matrices in both models. The diagonal element from top left to bottom right represent the number of correctly classified images according to their class (marked as bold in Table 4). Class order from left to right or top to bottom is according to the order of classes in Table 2. From Table 4, it can be seen that the majority of the images are classified correctly. The ResNet 101 shows slightly better classification performance on the first three dog breeds with the lowest number of data tests: Chesapeake Bay Retriever, Curly Coated Retriever, and Flat Coated Retriever. While ResNet 50 shows better performance on the last two dog breeds: Golden Retriever and Labrador Retriever. This happens probably because of the data imbalance, making it hard for ResNet 50 to classify dog breeds with more minor data.

TABLE IV
CONFUSION MATRICES

| ResNet 50 | | | | | ResNet 101 | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **27** | 0 | 1 | 5 | 11 | **29** | 0 | 2 | 6 | 7 |
| 2 | **32** | 1 | 4 | 2 | 1 | **33** | 1 | 4 | 2 |
| 0 | 2 | **29** | 4 | 7 | 1 | 1 | **33** | 3 | 4 |
| 0 | 0 | 0 | **1,029** | 43 | 0 | 0 | 0 | **1,024** | 48 |
| 2 | 2 | 2 | 44 | **666** | 1 | 2 | 48 | 48 | **663** |

E-ISSN : 2540 - 8984

**JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)**
Volume 07, Nomor 04, Desember 2022 : 1080 - 1087

TABLE V
RESNET 50 CLASSIFICATION REPORT

| | Precision | Recall | F1-score | Support |
|---|---|---|---|---|
| Chesapeake Bay Retriever | 0.87 | 0.61 | 0.72 | 44 |
| Curly Coated Retriever | 0.89 | 0.78 | 0.83 | 41 |
| Flat Coated Retriever | 0.88 | 0.69 | 0.77 | 42 |
| Golden Retriever | 0.95 | 0.96 | 0.95 | 1,072 |
| Labrador Retriever | 0.91 | 0.93 | 0.92 | 716 |
| accuracy | | | 0.93 | 1,915 |
| macro avg | 0.90 | 0.79 | 0.84 | 1,915 |
| weighted avg | 0.93 | 0.93 | 0.93 | 1,915 |

TABLE VI
RESNET 101 CLASSIFICATION REPORT

| | Precision | Recall | F1-score | Support |
|---|---|---|---|---|
| Chesapeake Bay Retriever | 0.91 | 0.66 | 0.76 | 44 |
| Curly Coated Retriever | 0.92 | 0.80 | 0.86 | 41 |
| Flat Coated Retriever | 0.87 | 0.79 | 0.82 | 42 |
| Golden Retriever | 0.94 | 0.96 | 0.95 | 1,072 |
| Labrador Retriever | 0.92 | 0.93 | 0.92 | 716 |
| accuracy | | | 0.93 | 1,915 |
| macro avg | 0.91 | 0.83 | 0.86 | 1,915 |
| weighted avg | 0.93 | 0.93 | 0.93 | 1,915 |

Fig. 10. Chesapeake Bay Retriever Correctly Classified using ResNet 50

Fig. 11. Flat Coated Retriever Correctly Classified using ResNet 101

Fig. 12. Chesapeake Bay Retriever Misclassified as Golden Retriever using ResNet 50

Fig. 13. Flat Coated Retriever Misclassified as Labrador Retriever using ResNet 101

Fig. 14. Curly Coated Retriever Correctly Classified using ResNet 50

Fig. 15. Golden Retriever Correctly Classified using ResNet 101

Fig. 16. Curly Coated Retriever Misclassified as Golden Retriever using ResNet 50

Fig. 17. Golden Retriever Misclassified as Labrador Retriever using ResNet 101

Table 5 shows the report of the ResNet 50, while Table 6 shows the report of ResNet 101. It showed that ResNet 101 performs slightly better than the ResNet 50 in almost every metric in each class. The values of recall and f1 score showed a subtle increase in the first three classes and remained the same for the other two. It can also be seen from the increasing macro-average value of recall and f1-score. The value of weighted-average precision, recall, and f1-score also remains the same. The same thing also happens with the overall accuracy. An exception to precision is a slight decrease in precision values for three classes: Flat Coated Retriever, Golden Retriever, and Labrador Retriever. It can also be represented by the decreased value of the macro-average precision by 0.1.

Figure 10 and Figure 14 were correctly classified data using ResNet 50, while Figure 11 and Figure 15 were correctly classified data using ResNet 101. These images were correctly categorized, possibly because the pictures show all the dog's characteristics and have no additional object besides the dogs. These unwanted objects could potentially make the model falsely classify the data, as shown in Figure 12, Figure 13, and Figure 16. However, Figure 17 possibly happened because that specific image doesn't reflect the dog's general characteristics from data training.

## IV. CONCLUSION

Based on the experiments conducted, the performance of ResNet 101 is better than ResNet 50. It is shown from the f1 score value, which is superior by 0.2 compared to ResNet 50 while also maintaining the same high accuracy.

It can be said that using the ResNet architecture with more layers can improve overall performance, as indicated by the increase in the f1-score value. However, it should also be noted that there is a decrease in some metric values. This happens likely due to an imbalance in the data held in each class. Therefore, it is better to do further data preprocessing so that the difference in data frequency in each class is not too significant, preventing the model from leaning towards a specific class.

In addition, it is necessary to use ResNet with the correct number of layers. Because some metrics are not improving, it is possible that using the ResNet architecture with too many layers does not produce significant improvements, or even overfitting can occur.

## REFERENCES

[1] A. van den Oord, S. Dieleman, and B. Schrauwen, "Deep content-based music recommendation," *Advances in Neural Information Processing Systems 26 (2013)*, vol. 26, 2013, Accessed: Dec. 16, 2021. [Online]. Available: http://hdl.handle.net/1854/LU-4324554
[2] Z. Ráduly, C. Sulyok, Z. Vadászi, and A. Zölde, "Dog Breed Identification Using Deep Learning," in *SISY 2018 - IEEE 16th International Symposium on Intelligent Systems and Informatics, Proceedings*, Nov. 2018, pp. 271–275. doi: 10.1109/SISY.2018.8524715.
[3] K. Balaji and K. Lavanya, "Medical Image Analysis With Deep Neural Networks," in *Deep Learning and Parallel Computing Environment for Bioengineering Systems*, Elsevier, 2019, pp. 75–97. doi: 10.1016/b978-0-12-816718-2.00012-9.
[4] M. Szummer and R. W. Picard, "Indoor-outdoor image classification," in *Proceedings - 1998 IEEE International Workshop on Content-Based Access of Image and Video Database, CAIVD 1998*, 1998, pp. 42–51. doi: 10.1109/CAIVD.1998.646032.
[5] D. Hsu, "Using Convolutional Neural Networks to Classify Dog Breeds," 2015.
[6] C. Wang, J. Wang, Q. Du, and X. Yang, "Dog breed classification based on deep learning," in *Proceedings - 2020 13th International Symposium on Computational Intelligence and Design, ISCID 2020*, Dec. 2020, pp. 209–212. doi: 10.1109/ISCID51228.2020.00053.
[7] X. Wang, V. Ly, S. Sorensen, and C. Kambhamettu, "Dog breed classification via landmarks," in *2014 IEEE International Conference on Image Processing, ICIP 2014*, Jan. 2014, pp. 5237–5241. doi: 10.1109/ICIP.2014.7026060.
[8] J. Liu, A. Kanazawa, D. Jacobs, and P. Belhumeur, "Dog Breed Classification Using Part Localization," in *12th European Conference on Computer Vision*, 2012, pp. 172–185. doi: 10.1007/978-3-642-33718-5_13.
[9] M. V. S. Rishita and T. A. Harris, "Dog breed classifier using convolutional neural networks," Dec. 2018. doi: 10.1109/ICNEWS.2018.8903980.
[10] A. Khosla, N. Jayadevaprakash, B. Yao, and F.-F. Li, "Novel Dataset for Fine-Grained Image Categorization: Stanford Dogs," 2012. [Online]. Available: http://vision.stanford.edu/aditya86/StanfordDogs/
[11] D. N. Zou, S. H. Zhang, T. J. Mu, and M. Zhang, "A new dataset of dog breed images and a benchmark for finegrained classification," *Computational Visual Media 2020 6:4*, vol. 6, no. 4, pp. 477–487, Oct. 2020, doi: 10.1007/S41095-020-0184-6.
[12] B. McMillan, "Dog Breed Guide: Explore the 7 Major Dog Groups," Dec. 07, 2020. https://www.masterclass.com/articles/dog-breed-guide#what-are-the-7-major-dog-groups (accessed Dec. 16, 2021).
[13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Dec. 2015, vol. 2016-December, pp. 770–778. doi: 10.1109/CVPR.2016.90.
[14] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing & Management*, vol. 45, no. 4, pp. 427–437, Jul. 2009, doi: 10.1016/J.IPM.2009.03.002.
[15] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, Jun. 2006, doi: 10.1016/j.patrec.2005.10.010.