ECTI Transactions on Computer and Information Technology
Journal homepage: https://ph01.tci-thaijo.org/index.php/ecticit/
Published by the ECTI Association, Thailand, ISSN: 2286-9131

Dog Breed Classification and Identification Using Convolutional Neural Networks

Nattakan Towpunwong[1] and Napa Sae-Bae[2]

ABSTRACT

This study aimed to assess the effectiveness of using pre-trained models to extract biometric information, specifically the dog breed and dog identity, from images of dogs. The study employed pre-trained models to extract feature vectors from the dog images. Multi-Layer Perceptron (MLP) models then used these vectors as input to train dog breed and identity classifiers. The dog breeds used in this study comprised two Thai breeds, Bangkaew and Ridgeback, and 120 foreign breeds. For dog breed classification, the results showed that, among the ImageNet classification models, the pre-trained NasNetLarge model has the highest dog breed classification accuracy (91%). The newly trained MLP model, which used feature vectors obtained by NasNetLarge, achieved higher accuracy at 93%. For dog identification, the results showed that, without data augmentation, the pre-trained ResNet50 model had the highest dog identification accuracy (75%). However, with data augmentation, MobileNetV2 could achieve a higher accuracy of 77%. When evaluating the identification performance of each breed, it is important to note that pugs achieved the lowest identification rate at 57.4%. Conversely, Bangkaew dogs demonstrated outstanding performance, with the highest identification rate at 98.6%.

Article information:

Keywords: Dog Breed Classification, Dog Identification, Dog Face Detection, Convolutional Neural Networks, CNNs

Article history:
Received: August 4, 2022
Revised: September 21, 2023
Accepted: November 30, 2023
Published: December 16, 2023 (Online)

DOI: 10.37936/ecti-cit.2023174.253728

1. INTRODUCTION

Stray dogs are a widespread and ongoing issue in Thailand and other countries. This issue stems from the rising number of dogs raised by owners who lack adequate knowledge of responsible dog care. As a result, dogs end up abandoned in public areas or get lost without a way to reunite them with their owners due to missing information about the dog.

Usually, identifying a dog's identity-related information relies on invasive methods such as collars, microchips, or GPS tags. However, these methods can be relatively expensive and are susceptible to loss or damage. A more efficient and non-invasive solution is to utilize an image classification system that can extract this information directly from dog images.

As such, we conducted this study to evaluate the effectiveness of utilizing pre-trained CNN models for identifying dog breeds and individual dogs from images. In particular, these pre-trained models were employed as feature extractors to derive feature vectors from dog images. Then, these feature vectors served as input to the Multi-Layer Perceptron (MLP) models, enabling the classification of dog breeds and identities. The dog breeds used in this study comprised two Thai breeds, Bangkaew and Ridgeback, as well as 120 other foreign breeds. Specifically, we conducted the first study to examine the classification performance of dog breeds (considered as one type of soft biometric information) from dog images. The dataset used in this study comprised 20,949 images with 122 breeds from the Stanford Dogs dataset and two Thai breeds locally collected from four Thai dog farms. In this study, the pre-trained CNNs models used to extract feature vectors from dog images were: Xception [1], VGG16 [2], ResNet50 [2], InceptionV3 [3], MobileNetV2 [4], and NasNetLarge [5]. In particular, we studied the efficacy of the pre-trained ImageNet classification models. In addition, we also stud-

[1] The author is with Information and Communication Technology Center, Department of Livestock Development, Bangkok, Thailand, E-mail: nattakan.t@dld.go.th
[2] The author is with Department of Computer Science, Srinakharinwirot University, Bangkok, Thailand, E-mail: napasa@g.swu.ac.th
[2] The corresponding author: napasa@g.swu.ac.th

ied the breed classification performance of a newly trained MLP model using feature vectors from the pre-trained CNNs models.

In addition, we conducted the second study to examine the classification performance of dog identity (considered as hard biometrics information). The dataset used in this study comprised 500 images of 62 dogs from 4 breeds (Husky, Pug, Bangkaew, and Ridgeback). In this experiment, a dog face was first detected from dog images using the DLLIB Library. VGG16, ResNet50, InceptionV3, MobileNetV2, and NasNetLarge, as well as two VGGFace models[6] (the pre-trained models trained for the human facial recognition task were then used to extract feature vectors from these dog face images. MLP model then used these feature vectors as input to classify the dog identities. In particular, we also investigated the benefits of using data augmentation for training the dog identification model.

In particular, our study provided the following distinctive contributions:

1) Inclusivity of Thai Local Breeds: We introduced Thai local dog breeds into this study, bringing a novel dimension to the research. These breeds may possess unique characteristics that previous studies have not investigated.

2) Exploration of CNN pre-trained Models for Face Identification: This study investigated the use of CNN pre-trained models trained for the human face identification task on the dog face identification task. This insight can shed some light on their suitability and performance in this specific context.

3) Breed-Specific Identification Performance: In contrast to earlier research that has not offered comprehensive insights, our study presented detailed identification performance of each dog breed through several performance metrics.

Finally, we propose combining a high-performance model for identifying dog breeds with a diverse dog identification model designed to recognize different dog breeds. This strategy aimed to enhance the accuracy and reliability of dog identification.

Figure 1 depicts the overview of the proposed system. In the initial step, a dog image was acquired from frontal and side angles, ensuring clear visibility of the dog's face and body without obstructions (or shadows) and in optimal condition. Images captured from the side angle are employed for dog breed classification, while those taken from the frontal angle undergo dog identification. For the latter, the system first detected the face from a dog image, to use as the input for subsequent processing.

The organization for the rest of this paper was as follows. Section 2 provides related literature on breed classification using deep learning models. Then, the dataset used in this study, including data collection, preprocessing, and augmentation techniques, was described in Section 3. Section 4 described the experimental setup, evaluation metrics, and the results of the dog breed classification and identification experiments using different pre-trained models. Finally, Section 5 discussed the findings and future directions for enhancing the performance of dog breed classification. In addition, we suggested dog identification models using deep learning techniques.

2. LITERATURE REVIEW

Using CNN models for image classification has become popular due to its effectiveness and efficiency. Moreover, with various pre-trained models such as VGG, ResNet, and Inception, one can develop an image classification system without an extensive dataset to train the network. As such, numerous studies have proposed using pre-trained models for various image classification tasks, including the classification of dog

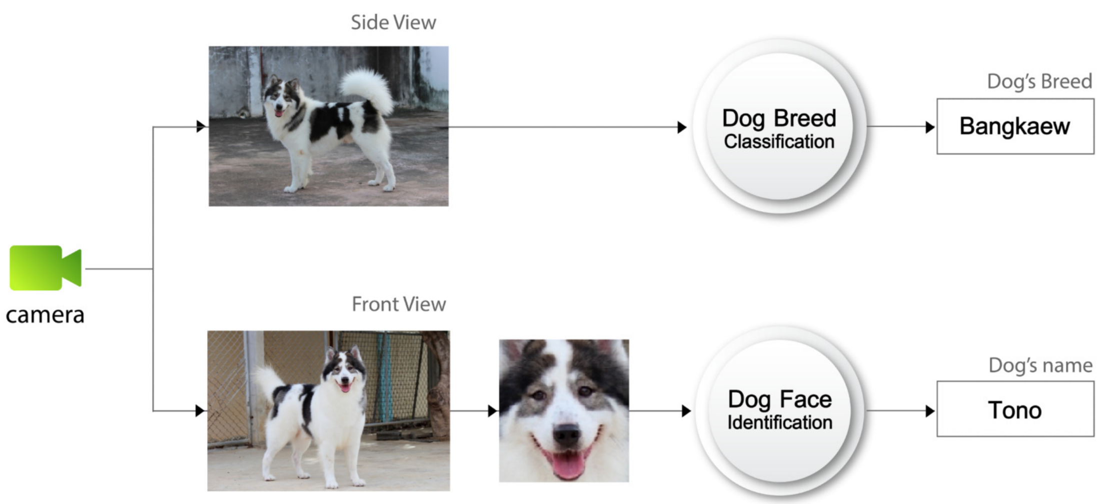

Fig.1: *Structure of an LSTM unit.*

images.

For dog identification, Kenneth et al.[1] have proposed a system that utilizes soft biometrics (breed, height, and gender) in combination with hard biometrics" (facial features) to identify the dog's identity. Their work studied the efficacy of various pre-trained CNNs for this task. The results revealed the highest accuracy in identifying breeds was 90.80% and 91.29% when utilizing the Xception model to separate dogs into two breeds on two different datasets. When using the dog identification network with soft biometrics, the system reached 84.94% identification accuracy.

For dog breed classification, previous work conducted by Weerasekara et al.[7] aimed to classify dog breeds using dog faces as input to convolutional neural networks (CNNs) and a shift-invariant neural network (SIANN) to extract image features. Their study gathered dog images from various sources, including Google and other websites, resulting in 2,500 images from 320 dogs. The results showed that the breed classification performance was over 90%. In addition, previous work conducted by Varshney et al. [7] proposed a model for classifying dog breeds on the Stanford dataset, which consists of 20,580 images of 120 dog breeds. In this work, pre-trained models used to extract prominent features included InceptionV3 and VGG16. The results showed an accuracy of 85% and 69%, respectively, with InceptionV3 outperforming VGG16 due to differences in their deep learning layers. Another research by Venkata et al. [8] compares three dog breed classifiers based on three types of Convolutional Neural Networks (CNNs): the first one was the model trained from scratch, the second one was the pre-trained VGG16 dog breed classifier, and the third one was to use pre-trained InceptionV3 to extract feature vectors and built a model to classify dog breed classification model from a dataset of 8,351 images of 133 breeds. The results revealed that the CNNs created from scratch achieved an accuracy of 10.89%, the VGG16 attained an accuracy of 36.60%, and the newly built model using feature extraction from InceptionV3 attained an accuracy of 87.42%.

3. DATASET

Three datasets used in this research were foreign dog breeds from the Stanford Dogs and Flickr-dog datasets and Thai dog images from four farms in Thailand. Details of datasets used in each experiment were as follows. Figure 2 depicts dataset usage.

3.1 Dog Breed Classification

We performed this experiment on 20,949 dog images. In particular, there were 122 breeds, with a minimum of 140 images for each breed. We gathered all the images from two sources. Details were as follows:

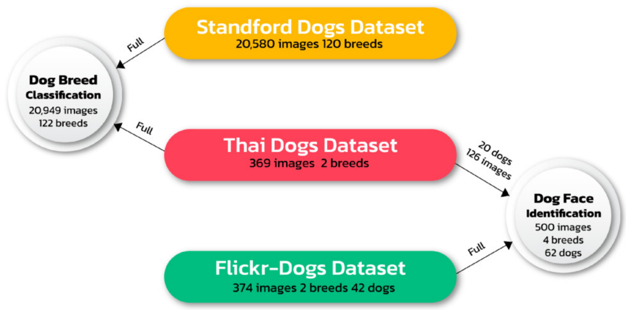

Fig.2: *Three datasets and their usage in dog breed classification and dog identification experiments.*

1) Foreign dog breeds: These images were from the Stanford Dogs Dataset [9], consisting of 20,580 images with 120 breeds. Each breed had at least 140 images. Notably, all images in the dataset showed the whole body of the dogs.

2) Thai dog breeds (Thai's Dog Dataset): These images were collected from four farms in Thailand, totaling 369 images, representing two distinct breeds: Bangkaew (186 images) and Ridgeback (183 images). Notably, all images in the dataset display the whole body of the dogs. Table 1 reports a summary of these details

3.2 Dog Identification and Detection of Dog Faces

We performed this experiment on 500 images from 62 dogs. We gathered all the images from two sources. Details were as follows:

1) Foreign dog breeds: These images were from the Flickr-Dog Dataset [10] consisting of 374 images of 42 dogs (at least five images for each dog) of the Husky and Pug breeds. Note that, in this dataset, dog images labeled by their names showed only the faces of the dogs.

2) Thai dog breeds (Thai's Dog dataset): These images were from 4 farms in Thailand consisting of 126 images of 20 dogs (at least five images for each dog). These 20 dogs were 10 Bangkaew and 10 Ridgeback. In this dataset, dog images labeled by their names showed the whole body of the dogs.

As such, in this dataset, we first applied face detection to crop faces from dog images. This operation is to retain only the facial region for subsequent processing. Additionally, considering the limited number of images available per dog, we have applied augmentation techniques to generate a more diverse set of images.

Table 1: *Summary of Thai Dog Dataset.*

Breed	# images	Reference
Bangkaew	186	Kamolchaibangkaew
Ridgeback	183	Barommasuk Farm, Damrongthai, and Thai Ridgeback Muang Non-TRD

4. EXPERIMENTAL RESULTS

The experiments to assess the effectiveness of using pre-trained models to extract biometric information from dog images included two classification tasks: 1) dog breed classification and 2) dog identification. Details of each experiment and respective results are as follows.

4.1 Dog-breed classification

Dog-breed classification is the task of predicting the dog-breed based on its image. The datasets used in this experiment are from the Stanford dogs and Thai datasets without modification. The datasets contained images of the whole body of dogs taken from the front or the side view without obstructions or shadows.

4.1.1 Evaluation matrix

Evaluation metrics used in assessing the performance of the models for dog breed classification in this study are the following.

1) Accuracy is a classification performance metric for measuring the overall accuracy or performance of the model. The formula is as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

where TP denotes the number of true positive samples, FP denotes the number of false positive samples, TN denotes the number of true negative samples, and FN denotes the number of false negative samples.

2) Precision is a classification performance metric for measuring the proportion of correctly predicted positive samples among the instances predicted as positive. The formula is as follows:

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

3) Recall is a classification performance metric for measuring how well a model can identify all positive samples out of the total number of positive samples. The formula is as follows:

$$Recall = \frac{TP}{Tp + FN} \tag{3}$$

4) F1-Score is a classification performance metric in predicting a particular class (Positive or "True"). The formula is as follows:

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{4}$$

Fig.3: *Examples of images from each dataset.*

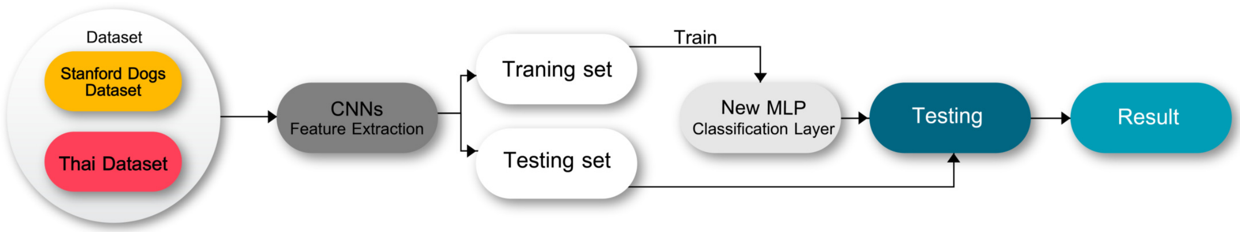

Fig.4: *Dog breed classification process with a pre-trained CNN model as the feature extraction and a newly trained classifier.*

4.1.2 The proposed model

The methodology involved the following steps:

1) Feature extraction was performed on each dog image using six selected CNN models trained from the ImageNet dataset with for object recognition task. These are Xception, VGG16, ResNet50, InceptionV3, MobileNetV2, and NASNetLarge.

2) A classifier was trained using an MLP model. In particular, the model consists of 10 layers. This model takes the feature vector as input to calculate the probability of a dog in the image being each dog breed. We reported details of these layers in Table 2.

Table 2: *Classifier for dog breed classification.*

Layer	Description
1	GlobalAveragePooling2D
2	Flatten
3	Dense 512, activation = Relu
4	Dropout 0.4
5	Dense 256, activation = Relu
6	Dropout 0.4
7	Dense 128, activation = Relu
8	BatchNormalization
9	Dropout 0.4
10	Dense n_classes, activation = Sigmoid

We divided the dataset into training and testing sets, with 85% allocated to the training set and 15% to the test set. In particular, we set the loss function as *sparse categorical_crossentropy*, the optimizer as *adam*, and the evaluation metric as accuracy. The training continued for 40 epochs, with each round of training using a batch size of 128.

4.1.3 Experiment results

In this experiment, we evaluated the performance of pre-trained models for extracting a feature vector in conjunction with a newly trained MLP model for the dog breed classification task. In addition, we also evaluated the performance of the pre-trained classifiers from the ImageNet dataset for the object recognition task (denoted as full layer model), which also included 120 dog breed classes. Table 3 reported the results. In particular, there were two settings for dataset usage: SF only, which included only the Stanford Dogs Dataset, and SF + Thai, which included both Stanford Dogs and the Thai Dataset, divided into four experimental protocols. The analysis of these results was as follows:

1) Comparing the pre-trained breed classification models (full layer) with the Stanford dataset, the model with the lowest accuracy was VGG16 (at 74% accuracy), whereas the model with the highest accuracy was NasNetLarge (at 92% accuracy).

2) Comparing the original pre-train models (full layer) for breed classification with both the Stanford dataset and the Thai dogs, the model with the lowest accuracy was VGG16 (at 72% accuracy), and the model with the highest accuracy was NasNetLarge (at 91% accuracy).

Table 3: *Dog breed classification performance.*

Dataset	Model	Accuracy		Precision	Recall	F1-Score
		Train	Test			
Full Layers						
Sf only	VGG16		0.74	0.74	0.72	0.73
	ResNet50		0.78	0.78	0.76	0.77
	MobileNetV2		0.83	0.82	0.81	0.82
	InceptionV3		0.89	0.88	0.87	0.87
	Xception		0.89	0.88	0.87	0.87
	NasNetLarge		0.92	0.9	0.9	0.9
Sf + Th	VGG16		0.72	0.72	0.71	0.71
	ResNet50		0.77	0.75	0.75	0.75
	MobileNetV2		0.82	0.8	0.8	0.8
	InceptionV3		0.88	0.85	0.86	0.85
	Xception		0.87	0.85	0.86	0.85
	NasNetLarge		**0.91**	0.88	0.89	0.88
Pre-train Models with a newly trained classifier						
Sf only	VGG16	0.87	0.60	0.62	0.60	0.59
	ResNet50	0.92	0.67	0.69	0.67	0.66
	MobileNetV2	0.96	0.77	0.78	0.77	0.77
	InceptionV3	0.94	0.89	0.89	0.89	0.89
	Xception	0.96	0.88	0.89	0.88	0.88
	NasNetLarge	0.98	0.93	0.94	0.93	0.93
Sf + Th	VGG16	0.88	0.61	0.63	0.60	0.60
	ResNet50	0.93	0.66	0.68	0.66	0.65
	MobileNetV2	0.96	0.77	0.78	0.76	0.76
	InceptionV3	0.94	0.89	0.89	0.88	0.88
	Xception	0.97	0.87	0.88	0.87	0.87
	NasNetLarge	0.98	**0.93**	0.93	0.93	0.93

3) Comparing pre-trained models for feature extraction on the Stanford dataset along with the new classification, we obsered that the model with the lowest accuracy was VGG16 (at 61% accuracy), and the model with the highest accuracy was NasNetLarge (at 93% accuracy).

4) Comparing pre-trained models for feature extraction on both the Stanford dataset and Thai Dogs, in conjunction with a newly trained classifier, the model with the lowest accuracy was VGG16 (60%), and the model with the highest accuracy was NasNetLarge (93

In addition, we observed that a newly trained classifier with NasNetLarge achieved the highest accuracy for classifying Thai and foreign dog breeds with the highest accuracy. This NasNetLarge model predicted all classes with 100% accuracy, and the class with the lowest accuracy was Class 80 or collie, with only 16 out of 23 images predicted correctly. For Thai dog breeds, NasNetLarge correctly predicted 50 out of 52 images (94%). In comparison, the VGG16 model had the lowest accuracy in all experiments.

4.2 Dog identification

The dataset used in this experiment is divided based on individual dog names and is sourced from the Flickr-Dog Dataset and the Thai Dataset. Due to the limited data available in the dataset, we also performed image augmentation on these images to increase diversity and improve the model identification performance. Table 5 reported augmentation parameters used in this experiment. Note that after the augmentation, the dataset increases from 500 images to 3,500 images in total.

Table 4: *The prediction characteristics of each model..*

Models	Prediction	
	The best prediction results	**The common mistakes**
VGG16	The model predicted Class 1 (Japanese_spaniel) correct 26 images out of 28.	In Class 50 (silky_terrier), out of 27 images, the model predicted three as Class 36 and 13 as Yorkshire_terrier.
ResNet50	The model predicted Class 120 (Thai_Bangkaew) correct 27 images out of 28.	In Class 29 (American_Staffordshire_terrier), out of 25 images, the model predicted five as Class 28 and nine as Staffordshire_bullterrier.
MobileNetV2	The model predicted Class 9 (Afghan_hound) correct 34 images out of 36. The model also predicted Class 23 (Norwegian_elkhound) correct 26 images out of 29.	In Class 80 (collie), out of 23 images, the model predicted three as Class 79 and predicted six as Shetland_sheepdog.
InceptionV3	The model predicted 18 classes correctly.	In Class 97 (Eskimo_dog), out of 22 images, the model predicted seven as Class 99 and predicted 12 as Siberian_husky.
Xception	The model predicted 23 classes correctly.	In Class 80 (collie), out of 23 images, the model predicted nine as Class 81 and seven as Border_collie.
NasNetLarge	The model predicted 40 classes correctly.	In Class 97 (Eskimo_dog), out of 23 images, the model predicted six as Class 99 and 16 as Siberian_husky.

Table 5: *Summary of Data Augmentation Parameters.*

Parameters	Value
rotation_range	45
shear_range	0.3
zoom_range	0.2
horizontal_flip	True
horizontal_flip	(0.3, 1.8)

4.2.1 Evaluation matrix

Evaluating its performance using the same metrics as for classifying dog breeds, with a focus on precision and recall to prevent misidentification, which could cause harm to the dog or society.

In addition, we also measured top-k accuracy (the identification accuracy when the classifier returns the top-k classes with the highest probabilities) in this experiment.

4.2.2 The proposed dog identification process

The process for dog identification begins with dog face detection to crop face image from whole body dog image. We utilize each dog face image as input for the pre-trained model to extract the feature vector. We then use these feature vectors to train a multi-layer perceptron classifier for dog identification tasks. Details for each step are as follows:

1) For detecting dog face area from dog images, we used Dlib Library v19.4 and OpenCV 3.3.0 [11]. Then, we cropped each dog image according to the detected bounding box and resized it to 250×250 pixels.

2) We extracted the feature vector of each face image using the seven pre-trained models. The training dataset and task for five models was the ImageNet dataset for the object recognition task. These are the ResNet50, VGG16, MobileNetV2, InceptionV3, and NasNetLarge models. In addition, the training dataset and task for two additional models were the VGG face dataset for the human face recognition task. We denoted these models as VGG-face (VGG) and VGG-face (ResNet50) according to their model architectures. The use of these pre-trained models trained from two distinct classification tasks was to compare the dog face identification performance of these models.

3) We then trained an MLP for a dog identification task. In particular, an MLP model consists of 6 layers that take a feature vector extracted from a pre-trained model as input and output the dog identity class. Table 6 reports details of these layers.

Table 6: *Dog identity recognition model.*

Layer	Process
1	GlobalAveragePooling2D
2	flatten
3	dropout 0.8
4	Dense100, activation = Relu
5	dropout 0.4
6	dense n_classes, activation = sigmoid

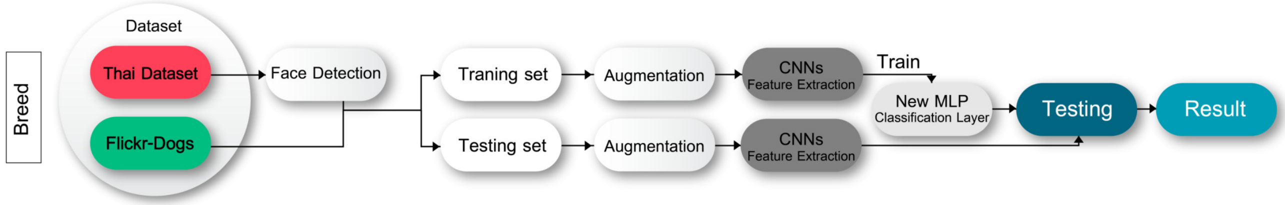

Fig.5: *The process of dog identification experiment. Here, face detection was exclusively applied to the dog images from the Thai dataset, as the images from the Flickr-Dogs dataset inherently focused solely on the facial region..*

For training the model, the dataset was divided into a training set consisting of 85% of all samples (with 80% training and 20% validation set), and a test set consisting of the remaining 15%. Then, to increase image variety, we augmented images in each set separately to prevent data leakage from one set to another. We set the loss function as *sparse_categorical_crossentropy*, the optimizer as *adam*, and the evaluation metric as accuracy. The training continued for 200 epochs, with each round of training using a batch size of 258.

4.2.3 Experiment result

In this experiment, we compared the performance of CNN models to extract image features in conjunction with a multi-layer perceptron classifier for the dog identification task. In addition, we investigated the effectiveness of using data augmentation to increase the quantity and diversity of training samples. Table 7 reports the results of this experiment. The analysis of these results was as follows:

1) The comparison of pre-trained models for feature extraction and new classification using the dataset without data augmentation showed that the model with the lowest accuracy was VGG16 (0.5%) and the model with the highest accuracy was ResNet50 (75%).

2) The comparison of pre-trained models for feature extraction and new classification using the dataset with data augmentation showed that the model with the lowest accuracy was VGG16 (2.3%), and the model with the highest accuracy was InceptionV3 (78%). This performance was higher than the accuracy obtained from the dataset without data augmentation. Upon further examination, we observed that, for each pre-trained model, the top-3 and top-5 accuracies were higher than the top 1 accuracy. This observation implied that the correct classification for each sample could fall within the top 3 or top 5 predicted classes, highlighting the potential usefulness of considering multiple predictions when analyzing the results of these models.

In addition, one can observe that VGG16 and ResNet50 models trained on the VGGFace dataset for the human face recognition task result in lower accuracy for dog face recognition than models trained on ImageNet for object recognition tasks. This observation implied that the features learned by the models from the VGGFace dataset are more optimized for human face recognition rather than for recognizing other objects, such as dogs. This discrepancy may arise due to the nature of the datasets; the VGGFace dataset contains only human faces, while the ImageNet dataset contains a wide variety of objects, including animals. As a result, when we used the VGG16 and ResNet50 models trained on the VGGFace dataset for dog face recognition, they may not be able to extract relevant features from the images of dog faces as effectively as the models trained on the ImageNet dataset.

Table 7: *Dog identification for performance Each CNN Model.*

Model	Train Acc	Test Acc	Top3 Acc	Top5 Acc	Precision	Recall	F1-Score
Raw Data 500 Images, 62 Dogs, 4 Breeds							
MobileNetV2	0.8	0.71	86.7%	94.7%	0.62	0.68	0.63
InceptionV3	0.78	0.65	90.7%	97.3%	0.61	0.65	0.61
NASNetLarge	0.87	0.72	92%	97.3%	0.64	0.69	0.65
VGG16	0.04	0.05	13.3%	21.3%	0.01	0.03	0.01
ResNet50	0.8	0.75	86.7%	93.3%	0.66	0.73	0.68
VGGFace (VGG16)	0.08	0.15	26.7%	41.3%	0.04	0.09	0.05
VGGFace (ResNet50)	0.83	0.61	89.3%	93.3%	0.54	0.59	0.55
Raw Data 500 Images, 62 Dogs, 4 Breeds							
MobileNetV2	0.82	0.77	92.2%	96.8%	0.77	0.77	0.74
InceptionV3	0.79	0.78	93.1%	97.7%	0.77	0.77	0.76
NASNetLarge	0.89	0.74	88.6%	94.%	0.71	0.72	0.7
VGG16	0.16	0.23	48.3%	65.-%	0.13	0.19	0.13
ResNet50	0.86	0.76	85.7%	93.1%	0.73	0.75	0.73
VGGFace (VGG16)	0.17	0.28	52.1%	67.6%	0.19	0.22	0.17
VGGFace (ResNet50)	0.76	0.54	81.9%	88.4%	0.56	0.52	0.51

Then, we examined the identification performance for each dog breed, and reported the results in Table 8. Notably, we observed that the Pug breed consistently exhibits the lowest identification performance in both experiments for all models, except for VGG16, with average F1 scores ranging from 0.07 to 0.59. This observation underscored the challenges associated with accurately identifying this particular breed. Furthermore, it's worth highlighting that MobileNetV2 is the top-performing model for the Pug and Husky breeds. On the contrary, InceptionV3 excels in identifying the Bangkaew and Ridgeback breeds, especially when we trained the models using augmented images in addition to the original ones.

Table 8: *Dog identification performance for each CNN model and dog breed.*

Breeds	Measure	MobileNetV2	InceptionV3	NASNet Large	VGG16	Resnet 50	VGG face (VGG16)	VGG face (ResNet50)
Raw 500 images from 62 Dogs - 4 Breeds								
Husky	F1score	0.7	0.72	0.7	0	0.83	0.072	0.56
	Recall	0.76	0.76	0.74	0	0.85	0.14	0.62
	Precision	0.68	0.71	0.68	0	0.83	0.06	0.54
Pug	F1score	0.54	0.37	0.51	0.03	0.48	0.07	0.39
	Recall	0.57	0.41	0.55	0.08	0.55	0.13	0.4
	Precision	0.58	0.4	0.51	0.03	0.44	0.05	0.4
Bang kaew	F1score	0.51	1	0.87	0	0.62	0	0.87
	Recall	0.6	1	0.9	0	0.7	0	0.9
	Precision	0.48	1	0.85	0	0.58	0	0.85
Ridge back	F1score	0.77	0.52	0.64	0	0.87	0	0.53
	Recall	0.8	0.6	0.7	0	0.9	0	0.6
	Precision	0.75	0.48	0.6	0	0.85	0	0.5
Total 500 images	**Avg F1-score**	0.63	0.65	0.68	0.01	0.7	0.04	0.59
	Avg Recall	0.68	0.69	0.72	0.02	0.75	0.07	0.63
	Avg Precision	0.62	0.65	0.66	0.01	0.68	0.03	0.57
Raw and augmented 3,494 Images from 62 Dogs, 4 Breeds								
Husky	F1score	0.87	0.84	0.79	0.09	0.85	0.2	0.56
	Recall	0.9	0.85	0.81	0.15	0.87	0.27	0.59
	Precision	0.88	0.85	0.8	0.08	0.86	0.2	0.62
Pug	F1score	0.59	0.53	0.47	0.14	0.45	0.11	0.4
	Recall	0.61	0.55	0.51	0.18	0.51	0.16	0.41
	Precision	0.61	0.53	0.47	0.17	0.43	0.1	0.42
Bang kaew	F1score	0.85	1	0.97	0.22	0.94	0.21	0.72
	Recall	0.87	1	0.97	0.31	0.94	0.2	0.71
	Precision	0.92	1	0.98	0.18	0.96	0.35	0.78
Ridge back	F1score	0.67	0.83	0.72	0.08	0.82	0.19	0.4
	Recall	0.73	0.83	0.73	0.14	0.83	0.27	0.41
	Precision	0.72	0.86	0.78	0.08	0.86	0.21	0.49
Total 3,494 images	**Avg F1-score**	0.75	0.80	0.74	0.13	0.77	0.18	0.52
	Avg Recall	0.78	0.81	0.76	0.20	0.79	0.23	0.53
	Avg Precision	0.78	0.81	0.76	0.13	0.78	0.22	0.58

Table 9: *Dog identification confusion matrix for Pug breed using MobileNetV2.*

True label	Prediction ID																					cross breed
	2	6	7	8	9	12	13	15	22	23	24	25	26	27	29	31	32	33	37	38	39	
2	1.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	1.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0.29	0	0	0	0	0	0	0	0.29	0	0	0	0.29	0	0.14	0	0	0	0	0
8	0	0	0	0	0	0.29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.71	0
9	0	0	0	0	0.29	0	0	0	0	0	0	0.29	0	0	0.43	0	0	0	0	0	0	0
12	0	0	0	0	0	0.14	0	0.29	0	0	0	0	0	0	0	0	0	0	0	0	0.57	0
13	0.07	0.29	0	0	0	0.07	0.21	0.14	0	0	0	0	0	0	0	0	0	0	0	0	0.21	0
15	0	0.29	0	0	0	0	0	0.71	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0	1.00	0	0	0	0	0	0	0	0	0	0	0	0	0
23	0.14	0	0	0	0	0	0	0	0	0.86	0	0	0	0	0	0	0	0	0	0	0	0
24	0	0	0	0	0	0	0	0	0	0	1.00	0	0	0	0	0	0	0	0	0	0	0
25	0	0	0	0	0.43	0	0	0	0	0	0	0.57	0	0	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0.14	0	0	0	0	0.86	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	1.00	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.00	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0	0.14	0	0	0	0.14	0	0	0.71	0	0	0	0	0	0
32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.00	0	0	0	0	0
33	0	0	0	0	0.07	0	0	0	0	0	0.14	0	0	0	0	0	0	0.57	0.21	0	0	0
37	0	0	0	0	0.71	0	0	0	0	0	0	0.14	0	0	0.14	0	0	0	0	0	0	0
38	0	0	0	0	0	0	0	0.43	0	0	0	0	0	0	0	0	0	0	0	0.29	0.29	0
39	0	0	0	0	0	0	0	0	0	0	0	0	0	0.43	0	0	0	0	0	0	0.57	0
Crossbreed	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 10: *Confusion Matrix for Husky breed from dog identification using MobileNetV2.*

True label	Prediction ID																					cross breed
	0	1	3	4	5	10	11	14	16	17	18	19	20	21	28	30	34	35	36	40	41	
0	0.57	0	0	0	0	0	0	0	0	0	0	0	0	0	0.43	0	0	0	0	0	0	0
1	0	0.86	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.14	0	0	0	0	0
3	0	0	1.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	1.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	1.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0.07	0	0.93	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	1.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	1.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	1.00	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	1.00	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	1.00	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	1.00	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	1.00	0	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0.07	0	0.93	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.00	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0.14	0	0	0	0	0.86	0	0	0	0	0	0
34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.00	0	0	0	0	0
35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.00	0	0	0	0
36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.00	0	0	0
40	0	0	0	0	0	0	0	0	0.71	0	0	0	0	0	0	0	0	0	0	0	0.29	0
41	0	0	0	0	0	0	0	0	0.14	0	0	0	0	0	0	0	0	0	0	0	0.86	0
Crossbreed	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 11: *Dog identification's confusion matrix for Bangkaew breed using InceptionV3.*

True label	Prediction ID										cross breed
	42	46	52	53	54	55	56	59	60	61	
42	1.00	0	0	0	0	0	0	0	0	0	0
46	0	1.00	0	0	0	0	0	0	0	0	0
52	0	0	1.00	0	0	0	0	0	0	0	0
53	0	0	0	1.00	0	0	0	0	0	0	0
54	0	0	0	0	1.00	0	0	0	0	0	0
55	0	0	0	0	0	1.00	0	0	0	0	0
56	0	0	0	0	0	0	1.00	0	0	0	0
59	0	0	0	0	0	0	0	1.00	0	0	0
60	0	0	0	0	0	0	0	0	1.00	0	0
61	0	0	0	0	0	0.14	0	0	0	0.86	0
Crossbreed	0	0	0	0	0	0	0	0	0	0	0

Table 12: *Dog identification's confusion matrix for Ridgeback breed using InceptionV3.*

True label	Prediction ID										cross breeds
	43	44	45	47	48	49	50	51	57	58	
43	0.71	0	0	0	0	0	0	0	0.29	0	0
44	0	1.00	0	0	0	0	0	0	0	0	0
45	0	0	1.00	0	0	0	0	0	0	0	0
47	0	0	0	0.86	0	0	0	0	0	0.14	0
48	0	0	0	0	0.86	0.14	0	0	0	0	0
49	0	0	0	0	0	0.57	0.43	0	0	0	0
50	0	0	0	0	0	0	0.71	0	0.29	0	0
51	0	0	0	0	0	0	0	0.86	0	0.14	0
57	0	0	0	0.14	0	0	0	0	0.86	0	0
58	0	0	0	0	0	0	0	0	0	1.00	0
Crossbreed	0	0	0	0	0	0	0	0	0	0	0

The confusion matrices for identifying dogs in each dog bread of each model are also presented in Tables 9-12. The results indicate that the Pug breed presents the most significant challenge for the *MobileNetV2* model. In particular, the models can identify only six dogs with a 100% identification rate. In addition, we noticed that the inaccurate identification results spread among other dogs within the same breed. Regarding the other three breeds, the models' accuracy in identifying dogs is over 80%. Notably, the Bangkaew breed achieved a 98.6% identification rate, followed by the Ridgeback and Husky breeds.

5. CONCLUSION

This study assessed the effectiveness of using pretrained models to extract biometric information, specifically the dog breed and dog identity, from images of dogs.

For dog breed classification, the results suggested that creating a new classification layer specific to the task led to higher accuracy rates. In particular, the NasNetLarge model with the newly trained classifier achieved the highest accuracy at 93%.

In the process of dog identification, the result shows that using data augmentation to increase the quantity and diversity of the data can slightly improve the model's performance. Therefore, increasing the number of images and adding more augmentation techniques, such as more than six different types, could lead to higher recognition performance of the model. In addition, the study observed that VGG16 and ResNet50 models trained on the VGGFace dataset for human face recognition resulted in lower accuracy for dog face recognition compared to models trained on the ImageNet dataset for object recognition. This finding highlights the importance of using datasets specifically tailored to the task at hand when training machine learning models, as the features learned by the models can significantly affect their ability to classify data accurately.

When considering Thai dogs for dog breed classification and identification, the study showed that the model could provide a relatively high accuracy rate of over 90% for both Thai dog breed classification and Thai dog breed identification.

Furthermore, since various dog breeds possess unique physical characteristics that influence their visual appearance, employing a model tailored to recognize these distinct features can significantly enhance the accuracy of dog identification, as demonstrated in the experiment. Therefore, combining a high-performance dog breed classification model with a heterogeneous dog identification model specifically designed to handle different dog breeds can lead to more accurate and reliable dog identification. This model would be helpful in applications such as pet registration, animal shelters, and veterinary medicine.

ACKNOWLEDGEMENT

The researchers would like to thank Barommasuk Farm, Nakhon Ratchasima Province, Thai Ridgeback Muang Non-TRD, Damrongthai Province, Kamlochai Bangkaew Farm, Phitsanulok Province, and Muang Non-TRD Thai Dog Farm, Nonthaburi Province, for providing photographs of Thai breed dogs, Ridgeback breed dogs, and Bangkaew breed dogs, which were essential data for conducting this research.

The researchers would also like to thank Mr. Chin Lertvipada for his assistance and advice in the breed classification of dogs

References

[1] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, pp. 1800-1807, 2017.

[2] K. Saraubon, *AI: Deep Learning by Python*. Bangkok, Thailand:Inter media, 2022.

[3] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, pp. 2818-2826, 2016.

[4] P.-L. Pröve. "MobileNetV2: Inverted Residuals and Linear Bottlenecks," Retrieved from https://towardsdatascience.com/mobilenetv2-inverted-residuals-and-linear-bottlenecks-8a4362f4ffd5 (accessed.)

[5] B. Zoph, V. Vasudevan, J. Shlens and Q. V. Le, "Learning Transferable Architectures for Scalable Image Recognition," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, pp. 8697-8710, 2018.

[6] W. Khan, A. Daud, F. Alotaibi, N. Aljohani, and S. Arafat, "Deep recurrent neural networks with word embeddings for Urdu named entity recognition," *ETRI Journal*, vol. 42, no. 1, pp. 90–100, 2020.

[7] D. T. Weerasekara, M. P. A. W. Gamage and K. S. A. F. Kulasooriya, "Combined Approach of Supervised and Unsupervised learning for Dog Face Recognition," *2021 6th International Conference for Convergence in Technology (I2CT)*, Maharashtra, India, pp. 1-5, 2021.

[8] M. V. Sai Rishita and T. Ahmed Harris, "Dog Breed Classifier using Convolutional Neural Networks," *2018 International Conference on Networking, Embedded and Wireless Systems (ICNEWS)*, Bangalore, India, pp. 1-7, 2018.

[9] A. Khosla, N. Jayadevaprakash, B. Yao, and F.-F. Li, "Novel dataset for fine-grained image categorization: Stanford dogs," in *Proc. CVPR*

workshop on fine-grained visual categorization (FGVC), vol. 2, no. 1, 2011.

[10] T. Pinheiro Moreira, M. Lisboa Perez, R. de Oliveira Werneck, and E. Valle, "Where Is My Puppy? Retrieving Lost Dogs by Facial Features," *arXiv e-prints*, p. arXiv: 1510.02781, 2015.

[11] A. Vlachynska, Z. K. Oplatkova, and T. Turecek, "Dogface Detection and Localization of Dogface's Landmarks," *Artificial Intelligence and Algorithms in Intelligent Systems. CSOC2018 2018. Advances in Intelligent Systems and Computing*,Springer, Cham, vol. 764, pp. 465-476, 2019.

Nattakan Towpunwong was born on January 5, 1990, in Phrae Province, Thailand. She obtained her Bachelor's degree in Computer Science from the Faculty of Information Technology and Communication at Phayao University in 2012. In 2016, she joined Global Wireless Co., Ltd. as a Flash Animator/Designer and Application Developer. Subsequently, in 2017, she took on the role of Help Desk at NECTEC NSTDA. In 2018, she transitioned to the position of Computer System Officer at the Information and Communication Technology Center, Department of Livestock Development. Successfully passing the civil service examination, she was appointed as a Computer Technical Officer, Practitioner Level in 2019, a position she currently holds.

She has a strong passion for data analysis and creating dashboards. In the year 2022 (B.E. 2565), she teamed up with colleagues to participate in the DIGI Data Camp Season 1 competition organized by the Data Innovation and Governance Institute of the Digital Government Development Agency of Thailand. Her team received the first prize, surpassing over 60 competing teams from various government agencies in Thailand. The competition focused on analyzing and designing reports or dashboards for use in work or for the public, aiming to provide overall benefits. In this competition, she and her team worked on the topic of the egg industry in Thailand, a subject closely related to their organization's work.

Moreover, in the article on "Dog Breed Classification and Identification Using Convolutional Neural Networks," she addressed a real issue in Thailand, and her organization is actively involved in disease prevention from animals. She sought solutions to tackle this specific problem.

Napa Sae-Bae received the Bachelor of Engineering degree in telecommunication engineering and the Master of Engineering degree in information system engineering from the King Mongkut's Institute of Technology Ladkrabang, Thailand, and the Master and Ph.D. in computer science from the NYU Tandon School of Engineering, New York University. She is currently an assistant professor at the Computer Science Department, Faculty of Science, Srinakharinwirot University, Bangkok, Thailand. She has published 30 articles in journals and conference proceedings in different venues and holds a US patent on multitouch gesture for user authentication on touch interface. Her research interests lie in the area of biometric, authentication, signal processing, pattern recognition, and consumer security.