

Knowing Your Dog Breed: Identifying a Dog Breed with Deep Learning

Punyanuch Borwarnginn

Worapan Kusakunniran

Sarattha Karnjanapreechakorn

Kittikhun Thongkanchorn

Faculty of Information and Communication Technology, Mahidol University, Nakhon Pathom 73170, Thailand

Abstract: Dog breed identification is essential for many reasons, particularly for understanding individual breeds' conditions, health concerns, interaction behavior, and natural instinct. This paper presents a solution for identifying dog breeds using their images of their faces. The proposed method applies a deep learning based approach in order to recognize their breeds. The method begins with a transfer learning by retraining existing pre-trained convolutional neural networks (CNNs) on the public dog breed dataset. Then, the image augmentation with various settings is also applied on the training dataset, in order to improve the classification performance. The proposed method is evaluated using three different CNNs with various augmentation settings and comprehensive experimental comparisons. The proposed model achieves a promising accuracy of 89.92% on the published dataset with 133 dog breeds.

Keywords: Computer vision, deep learning, dog breed classification, transfer learning, image augmentation.

1 Introduction

Image recognition and classification have successfully applied in various domains, such as face recognition^[1,2] and scene understanding for autonomous driving^[3]. At present, human face identification is successfully used for authentication and security purposes in many applications. Therefore, there are attempts to extend studies from human to animal recognition. In particular, dogs are one of the most common animals. Since there are more than 180 dog breeds, dog breed recognition can be an essential task in order to provide proper training and health treatment. Previously, dog breed recognition is done by human experts. However, some dog breeds might be challenging to evaluate due to the lack of experts and the difficulty of breeds' patterns themselves. It also takes time for each evaluation.

Besides, there are several studies on using dog images to identify their breeds. Chanvichitkul et al.^[4] proposed using coarse to fine classification by grouping similar face contours as a coarse classification and then applying a principle component analysis (PCA) classifier within the output group as fine classification. Prasong et al.^[5] extended the coarse to fine classification by adding local parts to reduce misclassification within the same group. This method used normalized cross correlation (NCC) to find each local part, such as ears and face. Then, the dog

breeds were classified in the PCA subspaces. It improved the runtime by four times and yielded an accuracy of 88% for 35 dog breeds.

Furthermore, a combination of shape and appearance features such as histogram of oriented gradient (HOG)^[6] and scale-invariant feature transform (SIFT)^[7] were used to classify breeds of cats and dogs^[8]. The model achieved 69% accuracy for identifying 37 breeds of cats and dogs. Similarly, Liu et al.^[9] reported an accuracy of 67% for 133 dog breeds from the Columbia Dogs Dataset by combining SIFT descriptors and color histograms on the SVM classifier with landmark data. Lai et al.^[10,11] introduced a deep learning method by transfer learning on convolutional neural networks (CNN) and achieved 86.63% accuracy on the same dataset in [9].

Most of the previous works used hand-crafted features that would find it difficult to discriminate between a large number of breeds. These selected features have been limited to certain types and may not contain sufficient information to increase the classification between breeds. Unlike conventional techniques, deep learning can create different features during training from the original images and achieved significant results that will be explored in this work. In summary, based on the public dataset, namely the Columbia Dogs Dataset^[9], our proposed method achieves the highest performance of 89.92% accuracy when compared to other existing methods in [9–11].

Our main contributions are twofold. First, we propose using a convolutional neural network (CNN) based model for dog breed identification. In order to prevent overfitting and imbalance classes, we also apply transfer learning and data augmentation techniques such as crop-

Research Article

Manuscript received June 4, 2020; accepted September 30, 2020; published online November 13, 2020

Recommended by Associate Editor Matjaz Gams

© Institute of Automation, Chinese Academy of Sciences and Springer-Verlag GmbH Germany, part of Springer Nature 2020

ping, translating, and rotating the original image. Therefore, three CNN architectures, including MobileNetV2, InceptionV3, and NASNet are evaluated and trained using dog face images. In our view, these results represent excellent baselines toward the further studies in dog identification using pre-trained model with some fine tuning. Second, the results from our model can demonstrate some preliminary findings of the selected areas on dog face images to use for classification. The key features which are significant to discriminate dog breeds are located in areas of eyes and nose.

The rest of this paper is organized as follows. Section 2 describes the applied technologies and approaches. Section 3 described the proposed model. The experiments are described and discussed in Section 4. Finally, the conclusion is drawn in Section 5.

2 Applied technologies and approaches

2.1 Convolutional neural networks

Over decades, computer vision techniques have been developed and achieved significant performance due to deep learning approaches^[12], e.g., convolutional neural networks (CNNs). According to the ImageNet Challenge^[13], InceptionV3^[14] is the first runner up for image classification in ImageNet 2015 and was published in CVPR 2016. It is the upgrade version of GoogleNet^[15], which reduces computational complexity. The main concept was to factorize convolutions into smaller convolutions. Fig. 1 shows an overview of InceptionV3 architecture. MobileNetV2^[16] aims to be a lightweight model. It is based on an inverted residual structure and depthwise separable convolutions to reduce the complexity and model size. NASNet^[17] is a current state-of-the-art on ImageNet classification with 82.70% top-1 accuracy. The architecture is based on neural architecture search (NAS) framework. The main concept is to search the best convolutional layer (cell) on the smaller dataset (i.e., CIFAR-10) and apply this layer to larger data such as ImageNet by stacking copies of this layer.

However, CNNs are known to have a high computation cost for training an entire model from scratch be-

cause they have several convolutional layers and connect with fully connected layers. It also requires a lot of data to get better accuracy and reduce overfitting. In order to reduce these limitations, there is a method called transfer learning. Transfer learning^[17–19] refers to the transfer of existing weights from pre-trained networks on a large dataset such as ImageNet^[20] and COCO^[21]. The main purpose is to reuse the parameters in the feature extraction layers to produce feature vectors instead of training them. The model has then replaced its full connected layers (classification layers) with the fully connected layers of the new dataset. Therefore, we can reduce computation costs by training the model on new classification layers. In this project, we apply dog breeds as the classification output and dog face images as the input images. We compare three pre-trained networks, including InceptionV3, MobileV2, and NASNet, to see the most suitable net for this particular task of our research question. This could be a good guideline for any other similar tasks.

2.2 Data augmentation

Since deep learning has outperformed in computer vision tasks, it requires a lot of training data to avoid overfitting. In the real world, data is limited due to various causes and might be an imbalance between classes. For instance, some breeds have less images than others because of their conservation. Therefore, several techniques have been attempted to overcome such limitation including dropout^[22], transfer learning^[18, 19, 23, 24], batch normalization^[25], and data augmentation^[26, 27]. In this paper, we apply transfer learning and data augmentation to reduce the overfitting problem.

Data augmentation is an approach to artificially increase the amount of training data by data wrapping or oversampling. Data wrapping is a technique that directly augments the existing images by performing geometric and color transformations such as cropping, translating, and rotating the image. Therefore, the augmented image preserved the same label as the input image, as showed in Fig. 2. Oversampling augmentation is another approach where an image is created by mixing images or using generative adversarial networks (GANs)^[28]. In this work, we demonstrate the data wrapping augmentation to increase

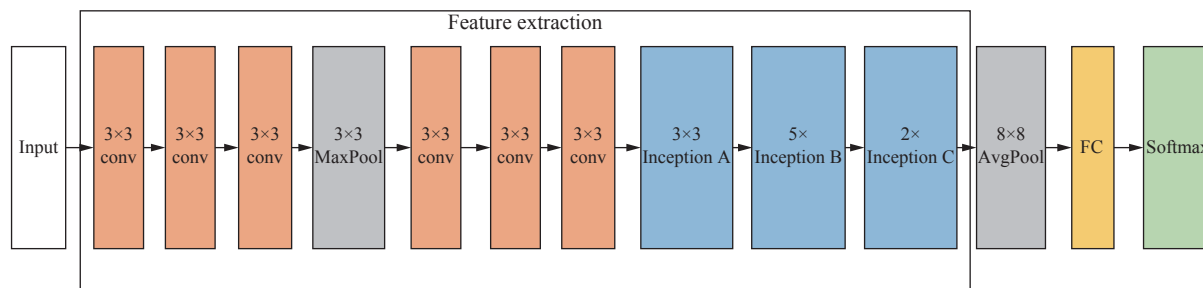


Fig. 1 Inception V3 architecture. Conv represents convolutional layer, FC represent the fully connected layer. Colored figures are available in the online version.



Fig. 2 Example of data wrapping augmentation: original images (left) and augmented images (right)

training images. Details of the setting will be explained in Section 3.

3 Proposed method

The proposed framework of dog breed classification is shown in Fig. 3. It consists of 3 main phases, which are data preparation, training and testing. Since we focus on dog face images, the data preparation step is required. Then, it is split for the training process and testing process. The output from the training model is a dog breed

model. The model is used for breed classification and model evaluation. Details are explained in the following subsections.

3.1 Data preparation

In this study, we use a public dataset to evaluate our method. The Stanford Dogs Dataset^[29] and Columbia Dogs Dataset^[9] are the public datasets for dog breed classification. We employ the Columbia Dogs Dataset as the data in this study. It contains 8351 dog images of 133 breeds by the American Kennel Club with 8-part locations annotated for each image. The sample images are shown in Fig. 4(a). Given the original images, it requires some pre-processing such as cropping and rescaling to extract dog faces as shown in Fig. 5. The pre-processed data is then split into a training set and testing set. The training set is augmented using data wrapping techniques such as rotation, flipping and adding noise.

3.2 Dog breed classification model

In this paper, the dog breed classification model is constructed by using transfer learning techniques. With transfer learning, we can train the model with a small dataset by using existing pre-trained CNNs from a large dataset such as ImageNet. Fig. 6 shows an overview of the dog breed classification model by using InceptionV3 as a pre-trained model. The model takes dog face images as the input and creates CNN features using an ImageNet

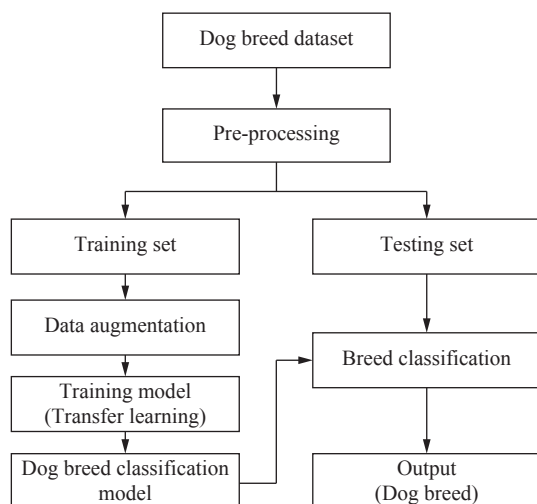


Fig. 3 Overview of the proposed framework

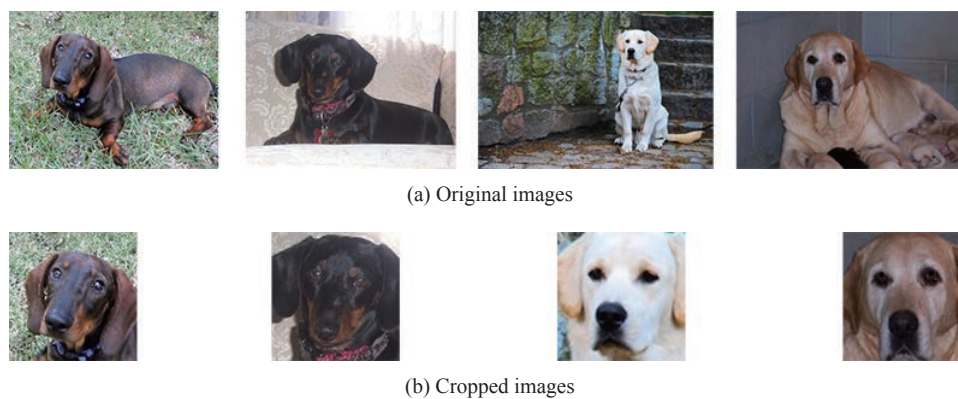


Fig. 4 Example of Columbia Dog Dataset

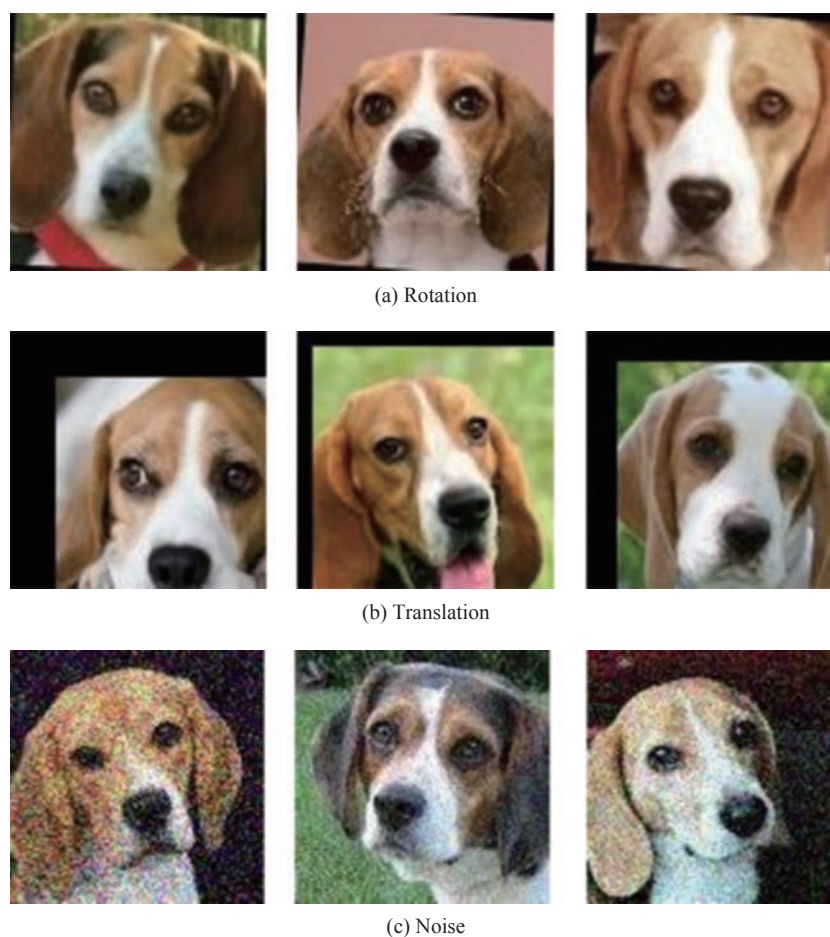


Fig. 5 Examples of augmented images

weight. Then it retrain the last fully connected layers with our dog breed data to build a new classifier.

In order to test the dog breed classification model, we use the testing set that is split from the data preparation phase. Dog face images in the testing set are fed into the dog breed model, which is trained from the training phase. Then the model output is a predicted dog breed. All experiment settings and results will be explained in Section 4.

4 Experiments

The proposed method is evaluated using two main scenarios in creating a training set: 1) apply without augmentation, 2) using various augmentation settings. In our experiments, we use Columbia Dogs Dataset^[9], the images are pre-processed and cropped faces. Therefore, 8111 images are selected and split into training and testing sets. The training set contains 6781 images, and the test-

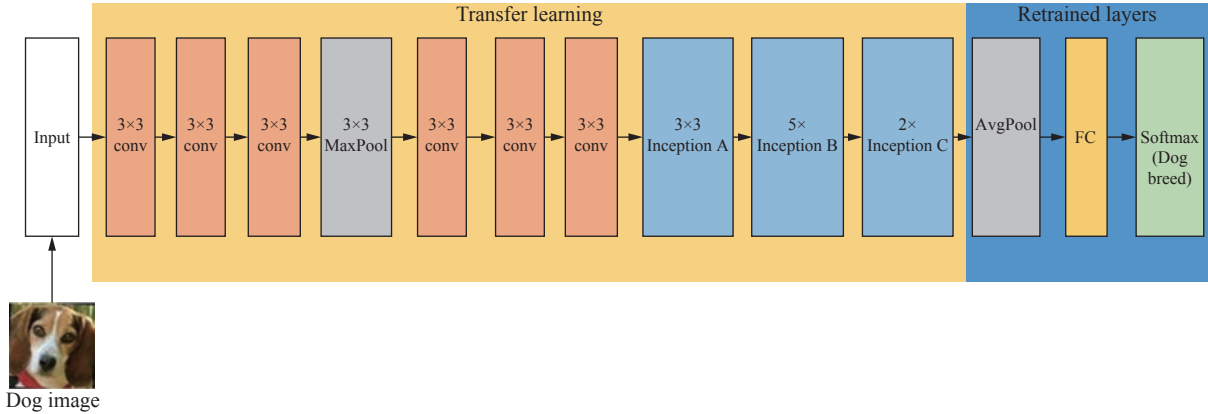


Fig. 6 An overview of transfer learning using the Inception V3 model. Transfer learning uses the feature extraction part from a trained model and retrain the new classification on the top layers

ing set consists of 10 images per breed at a total of 1330 images. Each setting is evaluated using three pre-trained models from the ImageNet dataset, including MobileNetV2, InceptionV3 and NASNet. We retrain the networks using the Tensorflow library.

Since our training set is small based on the number of classes that we have, we augment the training set to increase the number of images and to improve the performance. We apply data wrapping augmentation to the training set and compare performance between several settings, including rotation, translation, and adding noise. However, the number of degrees for transformation is decided based on the possible transformations that would occur in the real images. For example, the degree of rotation would not exceed 45 degrees, and the translation would not need to exceed half of the image, as shown in Fig. 5. Then we randomly select 200 images per breed as our training set.

As shown in Table 1, training sets containing rotation and translation achieve higher performance for dog breed classification than the baseline without augmentation. The reported results show that NASNet model achieves overall highest performances regardless of the training set used. We achieve an accuracy of 89.92% using the training set containing rotation images. Fig. 7 shows the confusion matrix of the model.

In addition, we evaluate our best setting using the 10-fold cross-validation, as reported in Table 2. We achieve an average classification rate of 89.74% with 1.07 standard deviations. Figs. 8 and 9 show the average classifica-

Table 2 Accuracy of 10-fold cross-validation using the NASNet model on the training set with rotation images

K fold	Accuracy (%)	SD
1	90.52	18.68
2	88.66	21.85
3	91.25	17.93
4	87.67	20.57
5	90.01	14.46
6	88.90	18.35
7	90.14	18.66
8	89.52	17.94
9	90.63	19.51
10	90.14	17.05
Avg±SD	89.74±1.07	

tion accuracy and standard deviation for each breed. Our results show that the models can recognize most breeds with overall accuracy above 80%.

5 Discussions

In the previous study on image classification, NASNet achieved the highest accuracy on image classification using the ImageNet dataset^[17]. Following our results in Table 1, we found that the results are in a similar orders, NASNet, InceptionV3, and MobileNetV2 regardless of training data. It confirms that the architecture of NASNet is current fit for our task. Although there are some improvements using augmentation techniques, we observe that adding noise reduces the performance. Noise can decrease the quality of images and could lead to model confusion. On the other hand, we found the rotating images can improve classification performance because testing images could be from various angles.

As shown in Table 3, we compare our proposed method and previous studies using the same dataset. Our result achieves the highest accuracy using NASNet with the augmentation training set, while previous methods used

Table 1 Accuracy of dog breed classification from different CNN models

Augmentation techniques	Accuracy(%)		
	MobileNetV2	InceptionV3	NASNet
Without	80.82	87.50	89.10
Rotation	81.65	88.42	89.92
Translate	81.65	89.02	88.87
Noise	80.30	85.94	88.80

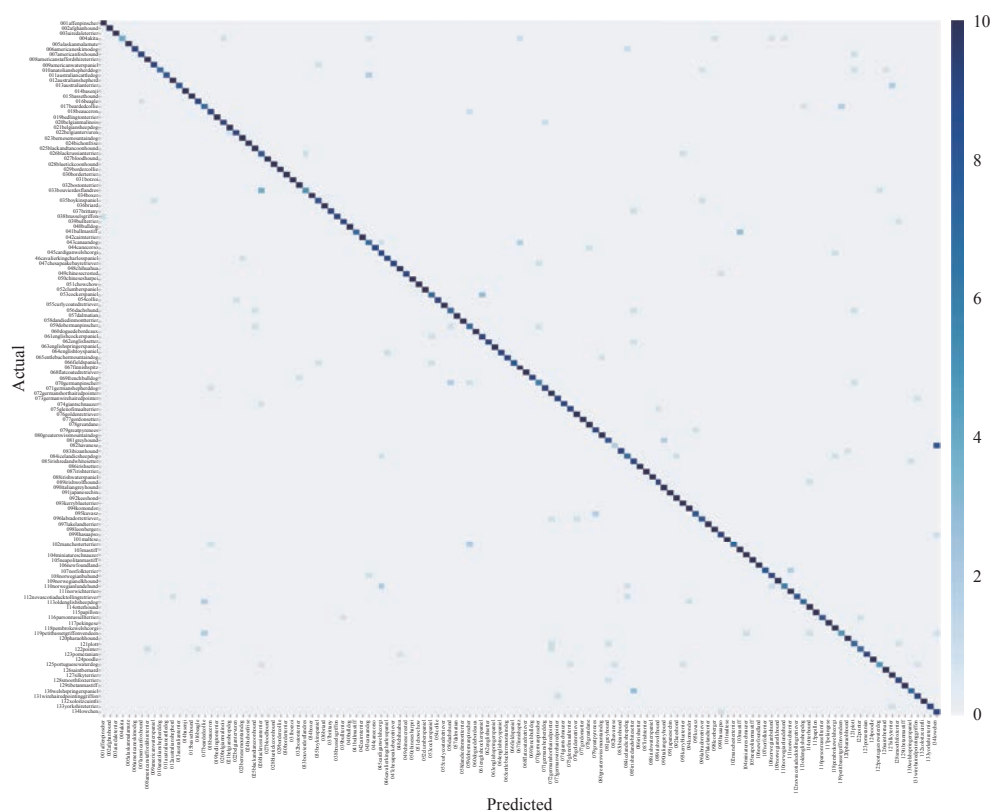


Fig. 7 Confusion matrix from the highest accuracy (89.92%) using NASNet models and training set containing some rotation images. The breed names listed from bottom to top in the y -axis and from left to right in the x -axis are ordered in the alphabet-ordering of breeds' names in the Columbia Dogs Dataset.

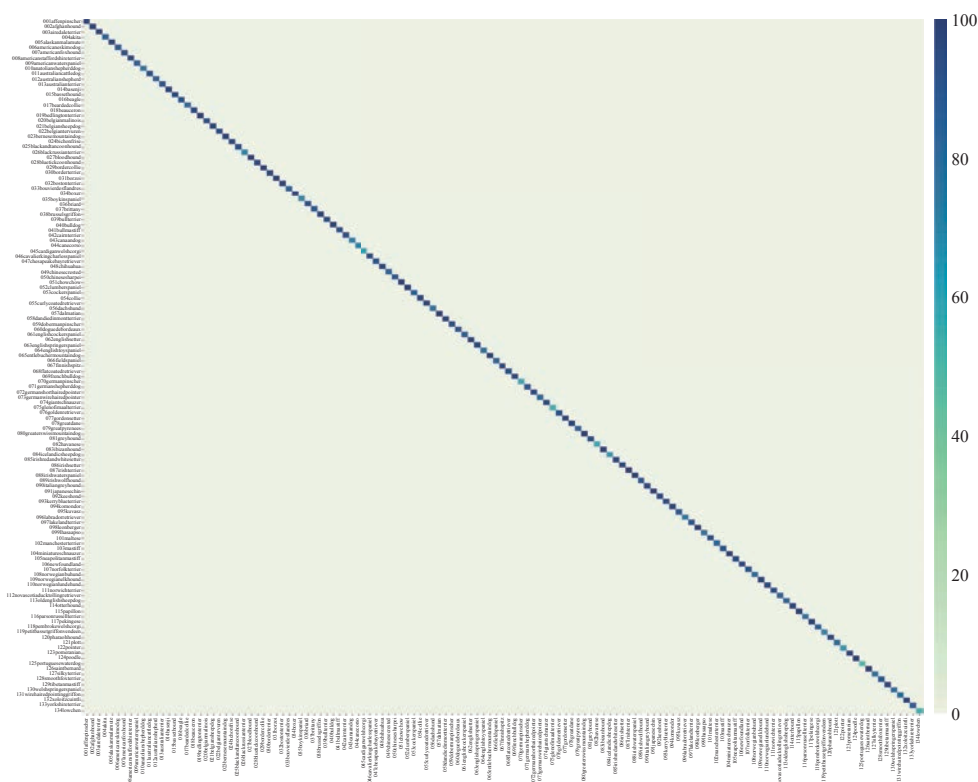


Fig. 8 Average accuracy of 10-fold cross-validation. The breed names listed from bottom to top in the y -axis and from left to right in the x -axis are ordered in the alphabet-ordering of breeds' names in the Columbia Dogs Dataset.

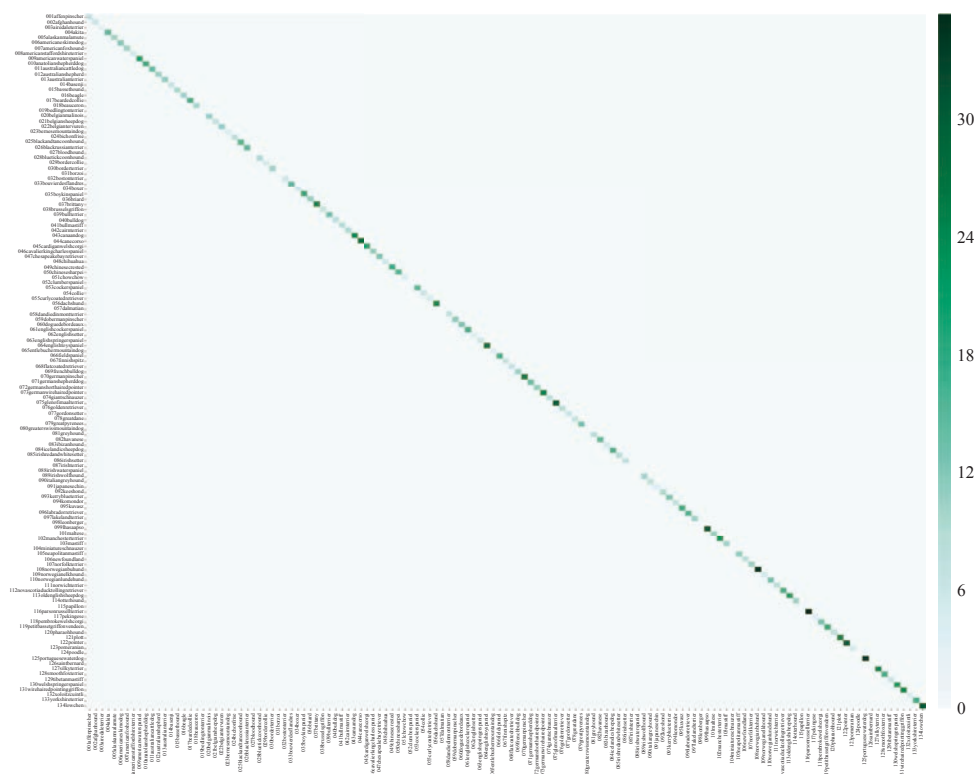


Fig. 9 Standard deviation of 10-fold cross-validation (133 breeds). The breed names listed from bottom to top in the y -axis and from left to right in the x -axis are ordered in the alphabet-ordering of breeds' names in the Columbia Dogs Dataset.

Table 3 Accuracy of dog breed classification from different CNN models

Model	Accuracy
Liu et al. [9]	67.00
BreedNet	86.63
NASNet with augmented data	89.92

traditional feature selection techniques that would find it difficult to discriminate between a large number of breeds.

While training, CNN layers in the network are updated by backpropagation from the optimizer and its loss function. Several features are generated from these iterations. In order to understand how the model distinguishes between dog breeds, we visualize a heatmap from the last feature extraction layer showing which parts of an image are used for classification based on higher weights, as shown in Fig. 10. The heatmap illustrates that the discriminative areas adopted in the classification are located in the center of an image, which contains the alignment between eyes and nose, their patterns, and textures, the rest are disregarded. Therefore, there are some confusions between breeds that have similar appearances and alignments of such faces' components, e.g., Lowchen/Havanese and Cardigan Welsh Corgi/Pembroke Welsh Corgi (Fig. 11). Since these breeds are similar or come

from the same origin, using only their faces might not be able to distinguish them.

This study has demonstrated that using deep learning can identify dog breeds from a dog face image. These results represent an initial step toward animal identification. However, it remains a challenge for us to improve the classification. To further our research, we will focus on cross-breed dogs by combining other parts, such as the body's shape, color, and texture, for training the breed classification model.

6 Conclusions

The paper proposes a method for identifying dog breeds using their face images with a deep learning-based approach. The proposed method applies the transfer learning technique using pre-trained CNNs and image augmentation to improve accuracy. The experiments examine three CNN models, which are MobilenetV2, InceptionV3 and NASNet. Each model is trained using training data containing image augmentation, including rotation, translation, and random noise. The NASNet model with a training set containing rotation images achieves the highest accuracy of 89.92%. Rotation can help with an alignment of images because the model mainly focuses on the center part of images. However, the proposed method could achieve a promising performance, with above 80% of classification accuracy on all settings. It



Fig. 10 Original image (left) and its heatmap (right) generated from the final feature layer of the model. The yellow spots mean the higher weights



Fig. 11 Examples of heatmap on misclassified breeds

could improve a high accuracy with augmented datasets such as rotation and translation.

Acknowledgments

This research was supported by the Royal Golden Jubilee (RGJ) Ph.D. Programme under the Thailand Research Fund (No. PHD/0053/2561)

References

- [1] S. G. Tong, Y. Y. Huang, Z. M. Tong. A robust face recognition method combining LBP with multi-mirror symmetry for images with various face interferences. *International Journal of Automation and Computing*, vol.16, no. 5, pp.671–682, 2019. DOI: 10.1007/s11633-018-1153-8.
- [2] F. K. Zaman, A. A. Shafie, Y. M. Mustafah. Robust face recognition against expressions and partial occlusions. *International Journal of Automation and Computing*, vol. 13, no.4, pp. 319–337, 2016. DOI: 10.1007/s11633-016-0974-6.
- [3] J. R. Xue, J. W. Fang, P. Zhang. A survey of scene understanding by event reasoning in autonomous driving. *International Journal of Automation and Computing*, vol.15, no.3, pp.249–266, 2018. DOI: 10.1007/s11633-018-1126-y.
- [4] M. Chanvichitkul, P. Kumhom, K. Chamnongthai. Face recognition based dog breed classification using coarse-to-fine concept and PCA. In *Proceedings of Asia-Pacific Conference on Communications*, IEEE, Bangkok, Thailand, pp.25–29, 2007. DOI: 10.1109/APCC.2007.4433495.
- [5] P. Prasong, K. Chamnongthai. Face-recognition-based dog-breed classification using size and position of each local part, and PCA. In *Proceedings of the 9th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, IEEE, Phetchaburi, Thailand, 2012. DOI: 10.1109/ECTCon.2012.6254212.
- [6] N. Dalal, B. Triggs. Histograms of oriented gradients for human detection. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE, San Diego, USA, pp.886–893, 2005. DOI: 10.1109/CVPR.2005.177.
- [7] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, vol.60, no.2, pp.91–110, 2004. DOI: 10.1023/B:VISI.0000029664.99615.94.
- [8] O. M. Parkhi, A. Vedaldi, A. Zisserman, C. V. Jawahar.

- Cats and dogs. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Providence, USA, pp.3498–3505, 2012. DOI: 10.1109/CVPR.2012.6248092.
- [9] J. X. Liu, A. Kanazawa, D. Jacobs, P. Belhumeur. Dog breed classification using part localization. In *Proceedings of the 12th European Conference on Computer Vision*, Springer, Florence, Italy, pp.172–185, 2002. DOI: 10.1007/978-3-642-33718-5_13.
- [10] K. Lai, X. Y. Tu, S. Yanushkevich. Dog identification using soft biometrics and neural networks. In *Proceedings of International Joint Conference on Neural Networks*, IEEE, Budapest, Hungary, pp.1–8, 2019. DOI: 10.1109/IJCNN.2019.8851971.
- [11] X. Y. Tu, K. Lai, S. Yanushkevich. Transfer learning on convolutional neural networks for dog identification. In *Proceedings of the 9th IEEE International Conference on Software Engineering and Service Science*, IEEE, Beijing, China, pp.357–360, 2018. DOI: 10.1109/ICSESS.2018.8663718.
- [12] B. Zhao, J. S. Feng, X. Wu, S. C. Yan. A survey on deep learning-based fine-grained object classification and semantic segmentation. *International Journal of Automation and Computing*, vol. 14, no. 2, pp.119–135, 2017. DOI: 10.1007/s11633-017-1053-3.
- [13] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. A. Ma, Z. H. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, F. F. Li. ImageNet large scale visual recognition challenge. *International Journal of Computer Vision*, vol.115, no.3, pp.211–252, 2015. DOI: 10.1007/s11263-015-0816-y.
- [14] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Las Vegas, USA, pp.2818–2826, 2016. DOI: 10.1109/CVPR.2016.308.
- [15] C. Szegedy, W. Liu, Y. Q. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich. Going deeper with convolutions. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Boston, MA, USA, pp.1–9, 2015. DOI: 10.1109/CVPR.2015.7298594.
- [16] M. Sandler, A. Howard, M. L. Zhu, A. Zhmoginov, L. C. Chen. Mobilenetv2: inverted residuals and linear bottlenecks. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Salt Lake City, USA, pp.4510–4520, 2018. DOI: 10.1109/CVPR.2018.00474.
- [17] B. Zoph, V. Vasudevan, J. Shlens, Q. V. Le. Learning transferable architectures for scalable image recognition. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Salt Lake City, USA, pp.8697–8710, 2018. DOI: 10.1109/CVPR.2018.00907.
- [18] J. Yosinski, J. Clune, Y. Bengio, H. Lipson. How transferable are features in deep neural networks? In *Proceedings of the 27th International Conference on Neural Information Processing Systems*, MIT Press, Montreal, Canada, pp.3320–3328, 2014.
- [19] L. Shao, F. Zhu, X. L. Li. Transfer learning for visual categorization: a survey. *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 5, pp.1019–1034, 2015. DOI: 10.1109/TNNLS.2014.2330900.
- [20] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, F. F. Li. ImageNet: a large-scale hierarchical image database. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Miami, USA, pp.248–255, 2009. DOI: 10.1109/CVPR.2009.5206848.
- [21] T. Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, C. L. Zitnick. Microsoft coco: common objects in context. In *Proceedings of the 13th European Conference on Computer Vision*, Springer, Zurich, Switzerland, pp.740–755, 2014. DOI: 10.1007/978-3-319-10602-1_48.
- [22] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, vol. 15, pp.1929–1958, 2014.
- [23] K. Weiss, T. M. Khoshgoftaar, D. D. Wang. A survey of transfer learning. *Journal of Big Data*, vol. 3, no. 1, Article number 9, 2016. DOI: 10.1186/s40537-016-0043-6.
- [24] A. R. Zamir, A. Sax, W. Shen, L. Guibas, J. Malik, S. Savarese. Taskonomy: disentangling task transfer learning. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Salt Lake City, USA, pp.3712–3722, 2018. DOI: 10.1109/CVPR.2018.00391.
- [25] S. Ioffe, C. Szegedy. Batch normalization: accelerating deep network training by reducing internal covariate shift. In *Proceedings of the 32nd International Conference on Machine Learning*, Lille, France, pp.448–456, 2015.
- [26] C. Shorten, T. M. Khoshgoftaar. A survey on image data augmentation for deep learning. *Journal of Big Data*, vol. 6, no. 1, Article number 60, 2019.
- [27] L. Perez, J. Wang. The effectiveness of data augmentation in image classification using deep learning. [online], Available: <https://arxiv.org/abs/1712.04621>, 2017.
- [28] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio. Generative adversarial nets. In *Proceedings of the 27th International Conference on Neural Information Processing Systems*, Montreal, Canada, pp.2672–2680, 2014.
- [29] A. Khosla, N. Jayadevaprakash, B. P. Yao, F. F. Li. Novel dataset for fine-grained image categorization: Stanford dogs. In *Proceedings of the 1st Workshop on Fine-Grained Visual Categorization*, IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Colorado Springs, USA, 2011.



Punyanuch Borwarnginn received the B.Sc. degree in information and communication technology from Mahidol University, Thailand in 2009, and the M.Sc. degree in informatics from the University of Edinburgh, UK in 2011. She is currently a Ph.D. degree candidate in computer science from Faculty of Information and Communication Technology, Mahidol

University, Thailand.

Her research interests include image processing, biometrics,

computer vision, pattern recognition and machine learning.

E-mail: punyanuch.bor@mahidol.edu

ORCID iD: 0000-0002-6309-5022



Worapan Kusakunniran received the B.Eng. degree in computer engineering from the University of New South Wales (UNSW), Australia in 2008, and the Ph.D. degree in computer science and engineering from UNSW, in cooperation with the Neville Roach Laboratory, National ICT Australia, Australia in 2013. He is currently a lecturer with Faculty of Information and Communication Technology, Mahidol University, Thailand.

He is the author of several papers in top international conferences and journals. Dr. Kusakunniran served as a program committee member for many international conferences and workshops. Also, he has served as a reviewer for several international conferences and journals, such as *International Conference on Pattern Recognition*, *IEEE International Conference on Image Processing*, *IEEE International Conference on Advanced Video and Signal Based Surveillance*, *Pattern Recognition*, *IEEE Transactions on Image Processing*, *IEEE Transactions on Information Forensics and Security*, and *IEEE Signal Processing Letters*.

His research interests include biometrics, pattern recognition, medical image processing, computer vision, multimedia, and machine learning.

E-mail: worapan.kun@mahidol.edu (Corresponding author)

ORCID iD: 0000-0002-2896-611X



Sarattha Karnjanapreechakorn received the B.Sc. degree in electrical-mechanical manufacturing engineering from Kasertsart University, Thailand in 2015, and the M.Sc. degree in game technology and gamification from Mahidol University, Thailand in 2017. He is currently a Ph.D. degree candidate in computer science of Faculty of Information and Communication Technology, Mahidol University, Thailand.

His research interests include image processing, biometrics, computer vision, pattern recognition and machine learning.

E-mail: sarattha.kar@student.mahidol.ac.th



Kittikhun Thongkanchorn received the B.Sc. degree in information and communication technology from University of Mahidol, Thailand in 2007, and the M.Sc degree in computer science from Faculty of ICT, Mahidol University, Thailand in 2012. He is currently a computer scientist, senior professional level with Faculty of ICT, Mahidol University, Thailand.

His research interests include computer system and network, elastic computing and distributed system, computer security and policy, image processing and machine learning.

E-mail: kittikhun.tho@mahidol.edu