OXFORD BROOKES UNIVERSITY

Josephine Qiao

An Exploration of The Potential of Ensemble Models In Dog Breed Identification

Oxford Brookes University in collaboration with Chengdu University of Technology
Supervised by Dr Grace Ugochi Nneji

Abstract

With the number of dog breeds reaching 20,580 in 2021[1], it is challenging and time-consuming to identify dog breeds by traditional canine differentiation methods. To provide a better solution for the increasingly complex dog breed classification area, this project will use an ensemble model to provide a more accurate and robust network for dog breed recognition. The model will utilize the stacking ensemble algorithm to merge three pre-trained CNN models obtained through transfer learning: VGG16, ResNet-50, and InceptionV3, into a stacked ensemble model. During this process, fine-tuning, data augmentation, and early stopping techniques will be utilized to improve the models' performance.

Figure 1: An example of dog breed identification

Dataset

"Dog Breed Identification Dataset" is an imbalance dataset and has 120 dog breeds. It has a total of 10222 images which the training dataset is 7218, the validation dataset is 1804 and the testing dataset is 1200.

"70 Dog Breeds-Image Data Set" has a total of 9346 images with 70 dog breeds, which the training dataset is 7946, the validation dataset is 700 and the testing dataset is 700.

Image processing

Resizing the images:
Resizing the first dataset into two version:
224*224*3: VGG16, ResNet50
299*299*3: InceptionV3

Resizing the second dataset:
299*299*3.
The original image size: 224*224*3

Data Augumentation:
Rotation: 5 degrees, Translation: 10%, Zoom: 10%, Shear: 5%, Brightness: 0.8-1.2, Horizontal Flipping.

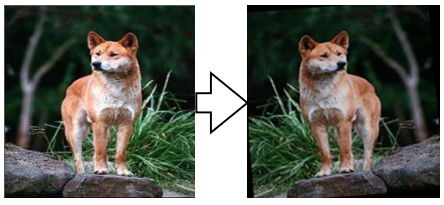

Figure 2(a): The original one

Figure 2(b): After the data augumentation

Figure 2: An example of doing the data augumentation

Ensemble models

CNN is popular in the field of image classification due to its excellent feature extraction ability. But a CNN whose output is ideal possibly will have an overfitting situation because of high complexity[2]. Individual CNN models are prone to the drawback of overfitting. In contrast, using the ensemble network can combine the advantages of different CNN models while compensating for each other's drawbacks to achieve a relatively desirable result for species classification and recognition. The excellent performance of ensemble learning has led to an increasing tendency to combine various CNN models[3]. The comparisons in Figure 4 demonstrate the benefits of using ensemble models.

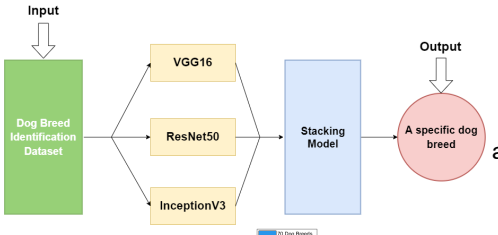

Figure 3: The proposed architecture of this project

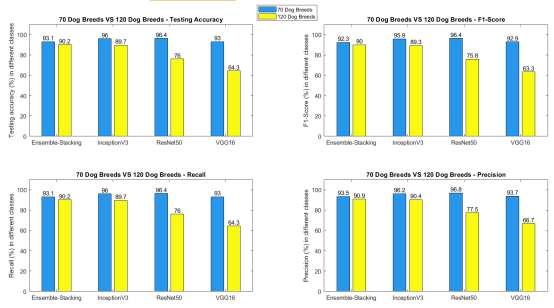

Figure 4: 70 Dog breeds VS 120 Dog breeds

A mini application for dog breed identification

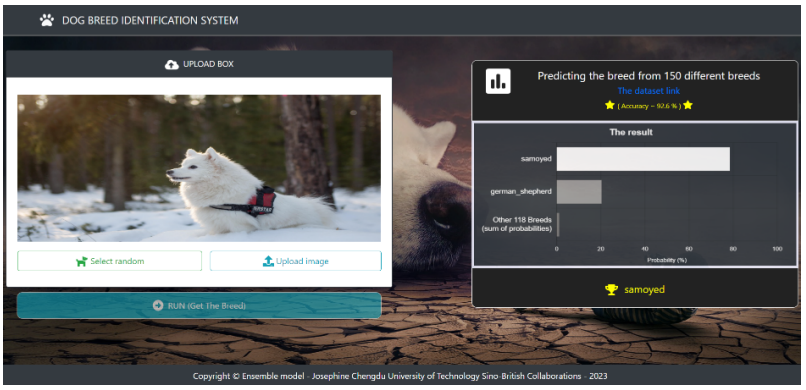

Figure 5: The deployment of the stacking ensemble model by Flask

- Increased usage and testing -> Collecting more data -> Improve the effectiveness and reliability of the model.
- Model performance become more intuitive.
- Gathering user feedback to improve the model's performance -> Identifying and solving problems with the model in practice -> Optimising the model and algorithms

Future work

- Implement the object detection algorithm before feeding the images into the model
- Combine with other better CNN models
- Implementation of other ensemble algorithms, such as boosting

References

[1] A. Varshney, A. Katiyar, A. K. Singh, and S. S. Chauhan, "Dog Breed Classification Using Deep Learning," in 2021 International Conference on Intelligent Technologies, CONIT 2021, Jun. 2021. doi: 10.1109/CONIT51480.2021.9498338.
[2] X. Cheng and H. Lei, "Remote Sensing Scene Image Classification Based on mmsCNN–HMM with Stacking Ensemble Model," Remote Sens (Basel), vol. 14, no. 17, Sep. 2022, doi: 10.3390/rs14174423.
[3]D. Vasan, M. Alazab, S. Wassan, B. Safaei, and Q. Zheng, "Image-Based malware classification using ensemble of CNN architectures (IMCEC)," Comput Secur, vol. 92, May 2020, doi: 10.1016/j.cose.2020.101748.