Desson
SHEN Wentian
201918020132

Recognition of Dog Breed Using Siamese Network

Department of Software Engineering, Oxford Brookes University- CDUT
Supervised by Dr. Grace Ugochi Nneji

Abstract

This project uses deep learning to create an automated system for recognizing dog breeds, which is difficult for experts and communities due to their large and diverse number and size. A convolutional neural network, specifically a dual residual network model, achieves a 95% identification accuracy by extracting distinct features. The system will benefit pet owners, pet care providers, and children by providing an automated solution for recognizing dog breeds.

Introduction

Dog breeds have a direct impact on dogs' behavior, so identifying them is important for pet growth and therapy, animal governance, and children's cognition. However, current image recognition technology is not well-suited for identifying dog breeds. A Siamese neural network, which can efficiently recognize and classify with limited data [1], is used in this project for dog breed recognition instead of a traditional CNN that requires more data.

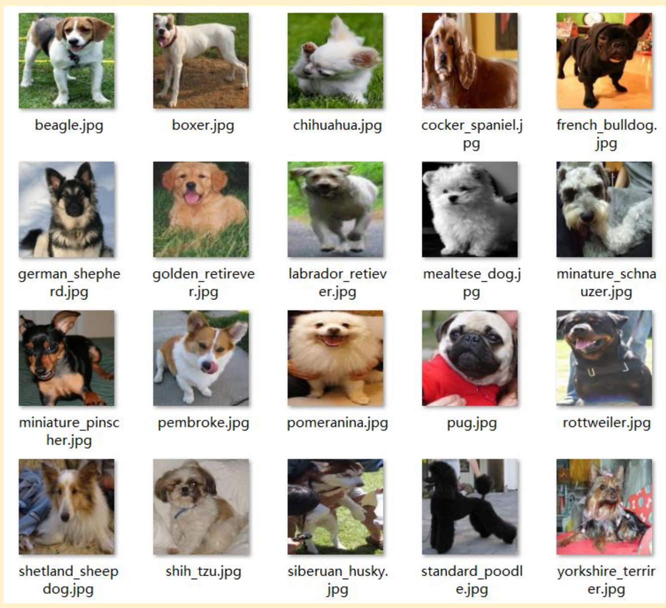

Figure 1：Dog Breed Example

Methodology

Aspect:	Approach:
Recognition model：	ResNet_V2_50 based Siamese Neural Network CNN ensures the accuracy of image recognition. Siamese network can effectively distinguish the type of data. The combination of CNN and Siamese neural network has higher data availability.
Dataset:	.Stanford Dogs Dataset(Train/test)
Train/test (evaluation method) :	Number of categories: 20(A subset of the original 120 varieties of the dataset was selected based on computer performance) Number of images: 3,717 (A subset of the original 120 varieties of the dataset was selected based on computer performance) Number of training set: 2,200 (11,999(59%) Number of test set: 1,517 (8,581(41%)
Evaluation Metrics	Accuracy, loss rate,

(a)

Hardware：	GPU	NVIDIA GeForce GTX 1660Ti
	CPU	Intel(R) Core (TM) i7-9750H CPU @ 2.60GHz
Software：	Software version management:	Relies on Python3.9's Miniconda
	Runtime Environment:	Python3.9
	Deep Learning Framework:	TensorFlow: A dataflow Programming-based symbolic mathematics system developed and maintained by Google Brain.
	IDE	PyCharm Community Edition version 2020.1.1
	Libraries	Keras, Numpy, tqdm, Pandas, OpenCV, Scipy

(b)

Table 1 (a) and (b): List of Methodology

Table 1 (a) and (b) describes the method of the key part of this project. Table 1(a) refers to the approach used to implement the project. Whereas Table 1(b) refers to the technologies used in the implementation process of the project.

Architecture

Creating the base network in Siamese neural network (ResNet, f in the Figure 2) by calling Subclassing Implementation. Euclidean distance is used to calculate the distance difference between two input samples, and RMSprop is used as the loss function of this network.

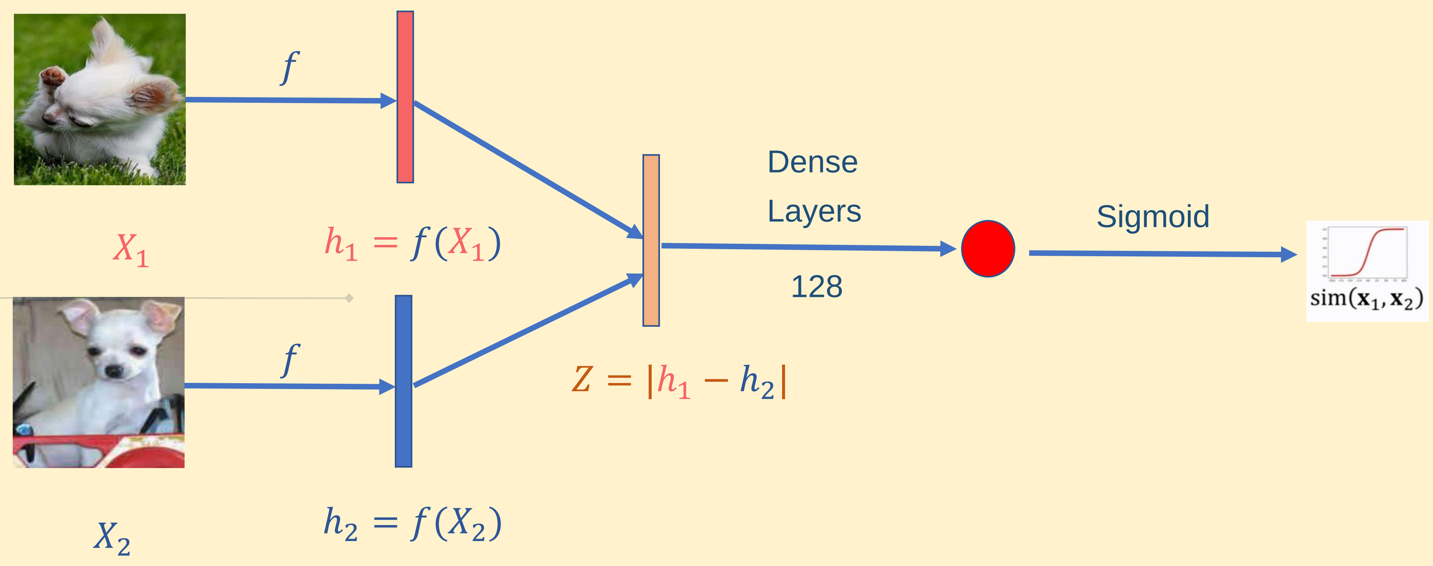

Figure 2: Siamese Network in Project

This project uses the loss function design of LeCun et al. [2], The contrastive loss is expressed as follows:

$$L = \frac{1}{2N}\sum_{n=1}^{N} y\, d^2 + (1 - y)\max(margin - d, 0)^2$$

Where d represents the Euclidean distance between the features of two samples, y is the label of whether the two samples match, y=1 means that the two samples are similar or matched, y=0 means that the two samples are not matched, margin is the set threshold.

Result

This project tested different batch sizes and learning rates to find the best accuracy for the model with 20 dog breeds. Finally, the best average recognition result of 95% is obtained when epoch is 48, batch size is 8, and learning rate is 0.0005.

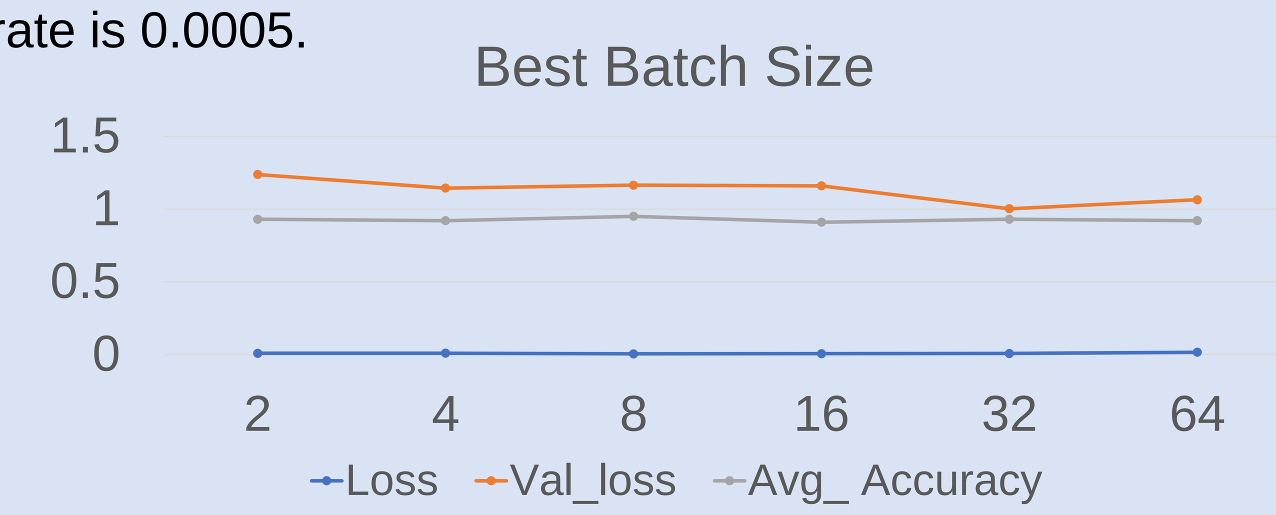

Figure 3: Results with different batch size

Figure 3 shows the test results under different batch sizes, and the test of learning rate uses the same method.

GUI

In this project, the recognition process of the model is visualized through Python Flask. After reading the picture through OpenCV, the list expression is used to meet the format or required for the entrance of the neural network, and finally the similarity with the image is judged.

Figure 4 shows recognizing the same photo of a dog (that is, the same breed) on the GUI, showing a result of 0.0. What's more, when two photos are of the same breed, their result probability is about 0 to 1, while the recognition result between different breeds will be large.

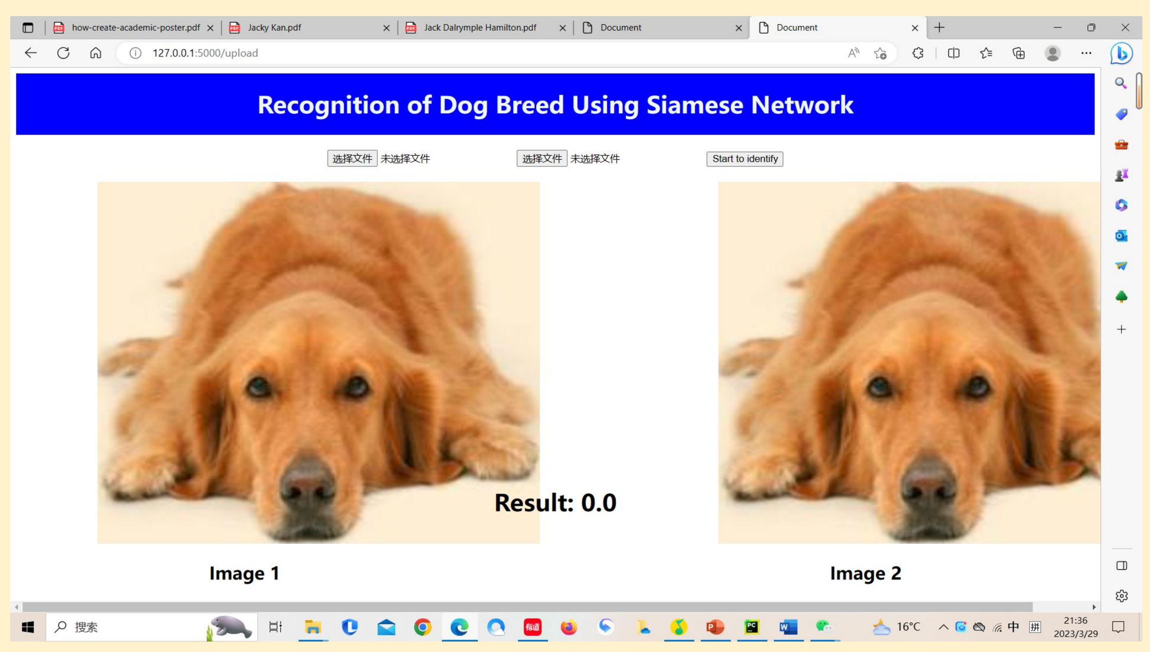

Figure 4: GUI for Siamese Network

Future Work

Future plans include further optimizing the structure and training method of the Siamese network to improve the accuracy and generalization ability of the model. The development of the project also plans to use a larger dataset with more powerful hardware support to train the model and try to apply the model to the dog species recognition task in real scenarios. In addition, we will explore the use of deep learning models for more advanced dog species identification tasks, such as identifying the dog's age, sex, etc., to optimize the identification interface and make the project more audience oriented.

Reference

[1]C. Zhang, W. Liu, H. Ma, and H. Fu, “Siamese neural network-based gait recognition for human identification,” in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, May 2016, vol. 2016-May, pp. 2832–2836. doi: 10.1109/ICASSP.2016.7472194.

[2]C. Szegedy *et al.*, “Going Deeper with Convolutions,” Sep. 2014, [Online]. Available: http://arxiv.org/abs/1409.4842