STAT 578a Deep Learning Practice Project – Bird Recognition

Student: Yingjun Guan

Date: 03/27/2018

Project Description

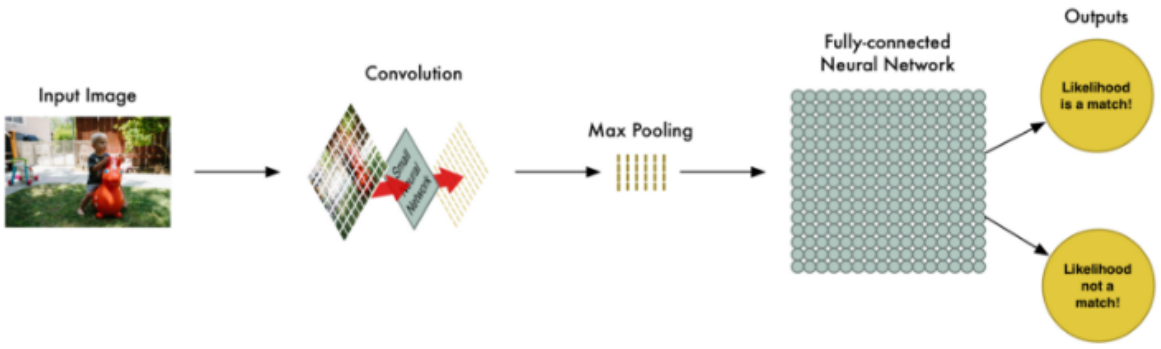
In the era of big data today, we have large amount of information as the training dataset, however, when facing similar but different entities, we are not confident enough to believe the machine can successfully make the judgement based on the training procedure. For example, in the world of birds, the problem came to whether the machine could recognize the bird, and furthermore, whether it can successfully label the species of the bird by training.

Human learn to recognize the bird out of other species by telling from the characteristics, feather, beak, number of legs, the ability of flying, and so on. However, when teaching the machine, instead of giving them the features mentioned above, we can offer them lots of specific examples (as shown in figure 1) for them to learn.

This project is to teach the machine to judge a bird, or to say, recognize whether there is a bird in the given picture after training with deep learning algorithms.

Algorithm

The algorithm of Convolution Neural Network (CNN) has been applied in this project. As instructed in class, CNN method consists of the procedure of breaking the image into overlapping image tiles; feeding each image tile into a small neural network; save the results from each tile into a new array; down sampling; (recursion if needed); and making the prediction. The flow process has been shown in the figure below.



Goals & Objectives

For this project, the objective is to use the dataset (introduced later) to train the machine to recognize the birds. The following goals are to be accomplished:

- Find the appropriate (as large as possible) dataset of birds, with good annotation (labelling).
- Parsing the data for training and testing (separate from each other).
- Transform image data to numeric data for computation and convolution.
- Train the data and save the classifier.
- Use the test data to validate and check the statistical results.

Dataset information

To find a large amount of appropriately annotated (labelled) bird images and non-bird images, our project turns to the CIFAR-10 and CIFAR-100 dataset. In the dataset, more and 80 million tiny images are labelled into 100 categories: airplane, automobile, bird, cat, deer, dog, frog, horse, ship, truck and so on. (Examples are shown in Figure 3)

For better realizing the purpose of our project all the categories other than birds are combined to non-birds. And to balance the number of birds and non-birds. Both polarity dataset have been normalized (nearly half to half). To further reduce the influence of dataset, we did the shuffle both categories before our training.

In the real dataset programming, we select 56,780 images (both of binary judgement) as the training dataset and 15,000 images (both positive and negative) as the testing dataset.

Results

Using the test dataset to validate the trained classifier. It can be noticed that our project separate the test dataset into 10 epochs, the accuracy of the classifier are between 85% to 93%, which is acceptable.

For each epoch, the time elapse is around 3-4 min. And the classifier can be trained better with more data.

Future work

If time allows, more detailed example of judgement will be provided, especially the false positive and false negative testing data.

The classifier can be improved also by adjusting the parameters more (so far most of the parameters are by default.)

References

- Learning Multiple Layers of Features from Tiny Images, Alex Krizhevsky, 2009.
- Goering, C., Rodner, E., Freytag, A., Denzler, J., "Nonparametric Part Transfer for Fine-grained Recognition", IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2014. **pdf**
- Wah, C., Van Horn, G., Branson, S., Maji, S., Perona, P., Belongie, S., "Similarity Comparisons for Interactive Fine-Grained Categorization", IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2014.
- http://www.vision.caltech.edu/visipedia/CUB-200-2011.html
- http://www.vision.caltech.edu/publications/WelinderPerona10.pdf
- https://www.cs.toronto.edu/~kriz/cifar.html
- https://medium.com/@ageitgey/machine-learning-is-fun-part-3-deep-learning-and-convolutional-neural-networks-f40359318721