Data Augmentation

第**15**組 具凱真優秀
資工所 簡辰哲 R09922146
電機所 楊智凱 R09921104
電機所 張景翰 R09921095
地理所 杜承軒 R07228005

Motivation

The prediction accuracy of the Supervised Deep Learning models is largely reliant on the amount and the diversity of data available during training. DL models trained to achieve high performance on complex tasks generally have a large number of hidden neurons. As the number of hidden neurons increases, the number of trainable parameters also increases. Developing an efficient method to collect more data automatically will be significantly helpful in model training, since it reduces the burden on manually augmenting the data by programmers. Thus, programmers can put more time and focus on the model itself.

Description

Data augmentation in data analysis are techniques used to increase the amount of data by adding slightly modified copies of already existing data or newly created synthetic data from existing data. It acts as a regularizer and helps reduce overfitting when training a machine learning model. It is closely related to oversampling in data analysis.

Objectives

In the final project, we would like to present “MyAUG” , which combines database technology (MySQL) and data augmentation techniques, that is able to produce image data with efficiency based on slightly modified SQL commands.

Implementation

We wish to implement a new syntax for the MySQL language, and improve the accessibility of data augmentation techniques for ML engineers.

Image data or its path must be stored in MySQL table, then implement data augmentation methods through MySQL syntax.

The list of the data augmentation methods to be implemented should include but not limited to the followings:

1. horizontal/vertical flipping, rotation, scaling, cropping, translation
2. elastic transform, perspective transform, piecewise affine transforms, pincushion distortion
3. noise injection
4. blurring

5. color to gray
6. channel shuffle
7. PCA jittering
8. dropout of regions

Outcome

Steps:

1. Insert the image data, which will be applied with the augmentation methods later, to a table.
2. MySQL terminal takes a modified SQL command, which contains the data augmentation and the table information.
3. MySQL loads the images and does the specified augmentation methods.
4. MySQL inserts the generated images into the current table or a newly created table.

MyAUG should be able to accomplish tasks like the followings:

	Image	Affine	
Original Input		Crop + Pad	
Gauss. Noise + Contrast + Sharpen		Fliplr + Perspective	

Reference

Data augmentation - wikipedia
https://en.wikipedia.org/wiki/Data_augmentation

資料預處理--資料擴增/Data Augmentation/影象增強
https://www.itread01.com/content/1543809605.html

imgaug - GitHub
https://github.com/aleju/imgaug?fbclid=IwAR3e7nWUG_IZbKmJLlizdpnqK1I5nPyIlrKRXQke9jvVDUVmN2F-ffE16-E

imgaug
https://imgaug.readthedocs.io/en/latest/