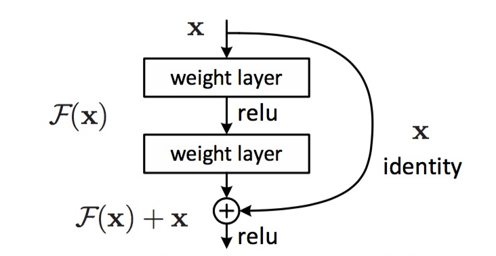
**Description**

Given an image, we are supposed to get the coordinates of the object in the file.

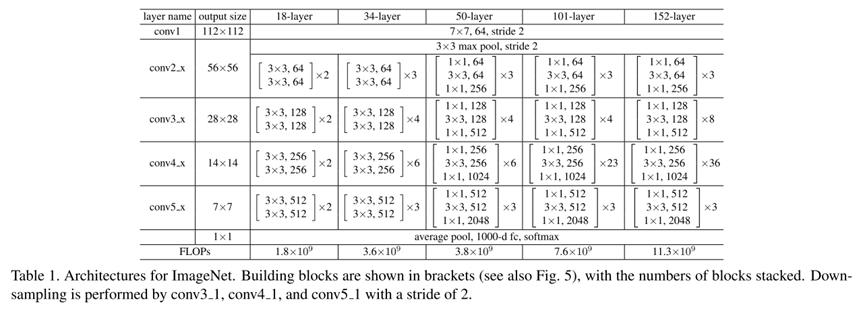
**Neural Network**

For traditional CNN and FC, it will meet degeneration problems when layers go deep.

In paper Deep Residual Learning for Image Recognition, they try to solve this problem by using a Residual Block:

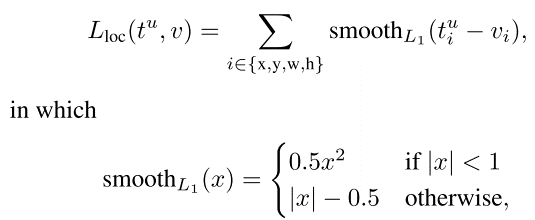
[](https://github.com/CKCZZJ/Image-Object-Localization/blob/master/img/ResidualBlock.png)

These blocks compose ResNet:

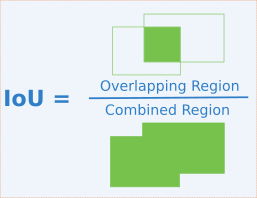


We use ResNet-18 in this project by adding a 4-dimension layer after ResNet-18 to predict box's x, y ,w and h.

Loss: smooth l1 loss

[](https://github.com/CKCZZJ/Image-Object-Localization/blob/master/img/loss.png)

Metric: IoU of ground truth and prediction, threshold=0.75

[](https://github.com/CKCZZJ/Image-Object-Localization/blob/master/img/IoU.png)

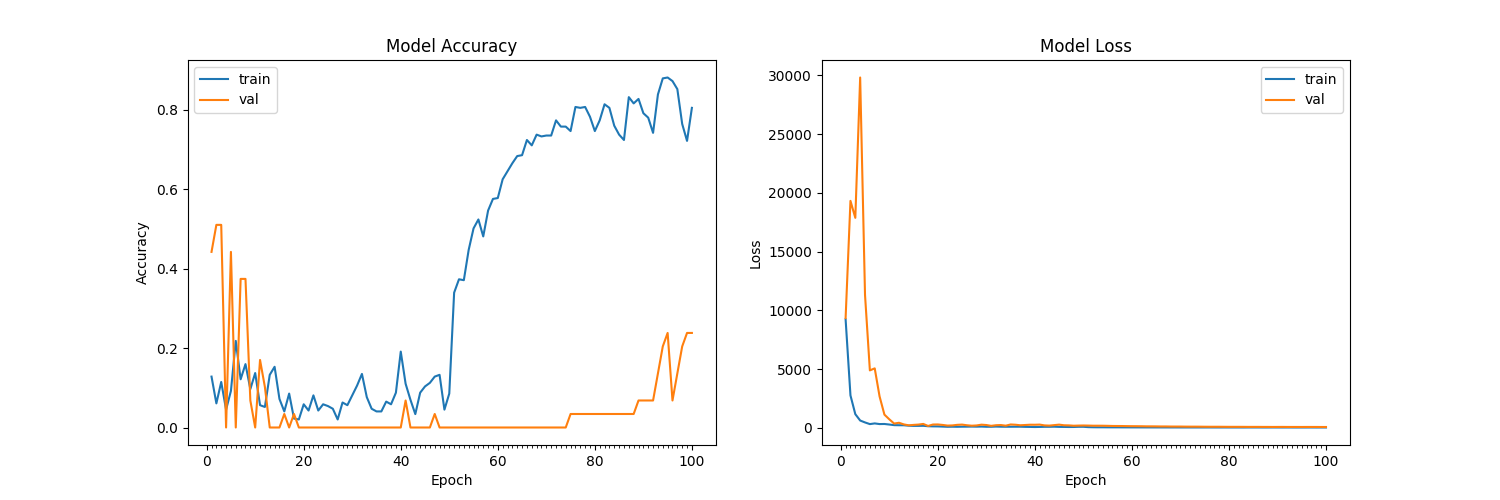
**Train**

Resize all images to 224\*224\*3

Then normalize and standardize all pixel channel.

Train network on training data using batch size=128, epoch=100 and validation split ratio=0.1

Training result:



**Dependency**

python 3.7

tensorflow

keras 2.1.0

numpy

PIL

pickle

matplotlib