

Approach -

The given problem was a bounding box regression problem.

Therefore, we used a Convolutional Neural Network combined with fully connected layers to do regression and calculate the coordinates of the bounding box around the object.

We decided to use Resnet architecture as it is a state-of-the-art deep learning architecture.

The core idea of ResNet is introducing a so-called *“identity shortcut connection”* that skips one or more layers

‘We did not use the pre-trained weights of resnet and only used its architecture and trained it from scratch.’

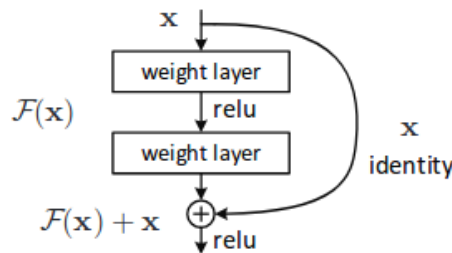


Figure 2. Residual learning: a building block.

Residual block

Therefore, ResNet Network Converges faster compared to plain counter part of it.

We trained two models from ResNet

1. ResNet18
2. ResNet50

We observed Resnet18 achieved higher accuracy than Resnet50 because the latter overfitted the data since the training dataset was small and hence more number of layers were not useful.

We used mini-batch size 10 and trained our model over 30 epochs.

We used 20% training data as validation dataset to keep a check on overfitting. Therefore our model was trained on 19200 images and rest were used for validation.

We normalised the images to bring the pixel values between the range of 0-1 in order to prevent the exploding gradients problem. Adam Optimisation algorithm was used as it has

1. **Adaptive Gradient Algorithm (AdaGrad)** - that maintains a per-parameter learning rate that improves performance on problems with sparse gradients (e.g. natural language and computer vision problems).
2. **Root Mean Square Propagation (RMSProp)** - that also maintains per-parameter learning rates that are adapted based on the average of recent magnitudes of the gradients for the weight (e.g. how quickly it is changing). This means the algorithm does well on online and non-stationary problems (e.g. noisy).

Our learning rate was 0.003

We trained our model on Google Colab as it provides free GPU support to train deep-learning models.

Tools used are -

1. Python3
2. Pytorch
3. OpenCV
4. Pandas
5. Numpy
6. Google Colab
7. Google Drive