Convolutional Neural Network:

Brief introduction to CNNs

Artificial neural networks are computing systems that are inspired by the biological neural networks in the human brain.

Convolutional neural network (CNN, or ConvNet) is a class of deep, feed-forward artificial neural networks that has successfully been applied to analysing visual imagery. In this project, we use CNNs to identify and mark airplanes from satellite imagery. We have also used few different CNN architectures and compared their accuracies.

A CNN consists of an input layer, output layer and multiple hidden layers. The hidden layers of a CNN typically consist of convolutional layers, pooling layers, fully connected layers and normalization layers. By varying the number of convolutional, pooling, fully connected and normalization layers and by combining these layers, we get different neural network configurations.

In this project we have tried the following network architectures:

1. CNN1

2. CNN2

3. AlexNet

4. VGGnNet

CNN1:

Convolutional -> MaxPool ->Convolutional -> MaxPool ->Convolutional -> MaxPool ->Convolutional -> MaxPool ->Convolutional -> MaxPool -> Fully connected -> Dropout -> Fully connected -> Regression

CNN2:

Convolutional -> MaxPool -> Convolutional -> Convolutional -> MaxPool -> Fully connected -> Dropout -> Fully connected -> Regression

AlexNet:

Convolutional -> MaxPool -> Normalization -> Convolutional -> MaxPool -> Normalization -> Convolutional -> Convolutional -> Convolutional -> MaxPool -> Normalization -> Fully connected -> Dropout -> Fully connected -> Dropout -> Fully connected -> Regression

VGGNet:

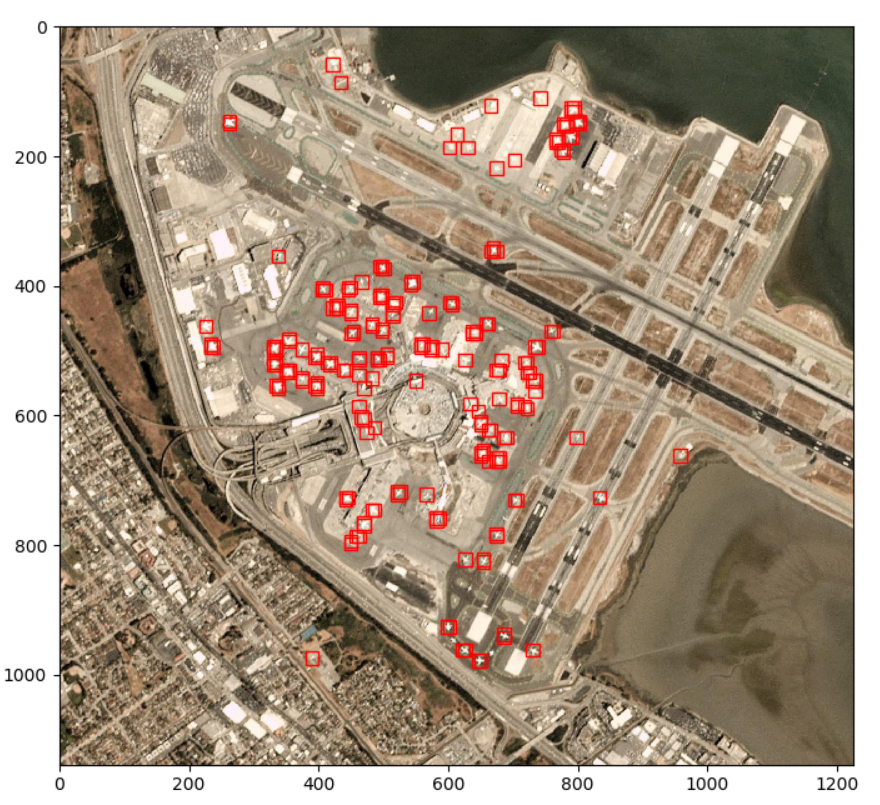
Convolutional -> Convolutional -> MaxPool -> Convolutional -> Convolutional -> MaxPool -> Convolutional -> Convolutional -> Convolutional -> MaxPool -> Convolutional -> Convolutional -> Convolutional -> MaxPool -> Convolutional -> Convolutional -> Convolutional -> MaxPool -> Fully connected -> Dropout -> Fully connected -> Dropout -> Fully connected -> Regression

Note: Although not mentioned, Every Convolutional layer is followed by a ReLU layer (Rectified Linear Units) which helps to increase non-linearity and trains the neural network faster.

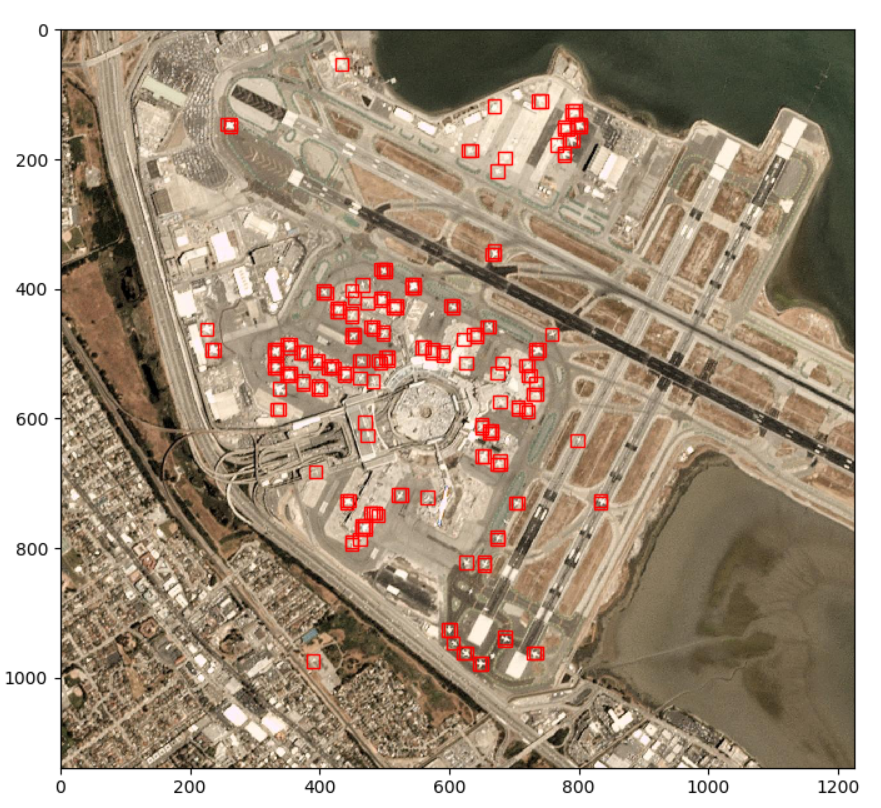
CNN1, CNN2 and VGGNet performed equally well on our test set. AlexNet was slightly lagging when compared to the other three.

Here are some of the results obtained by running each of these Convolutional neural networks on the test image using Sliding window detector:

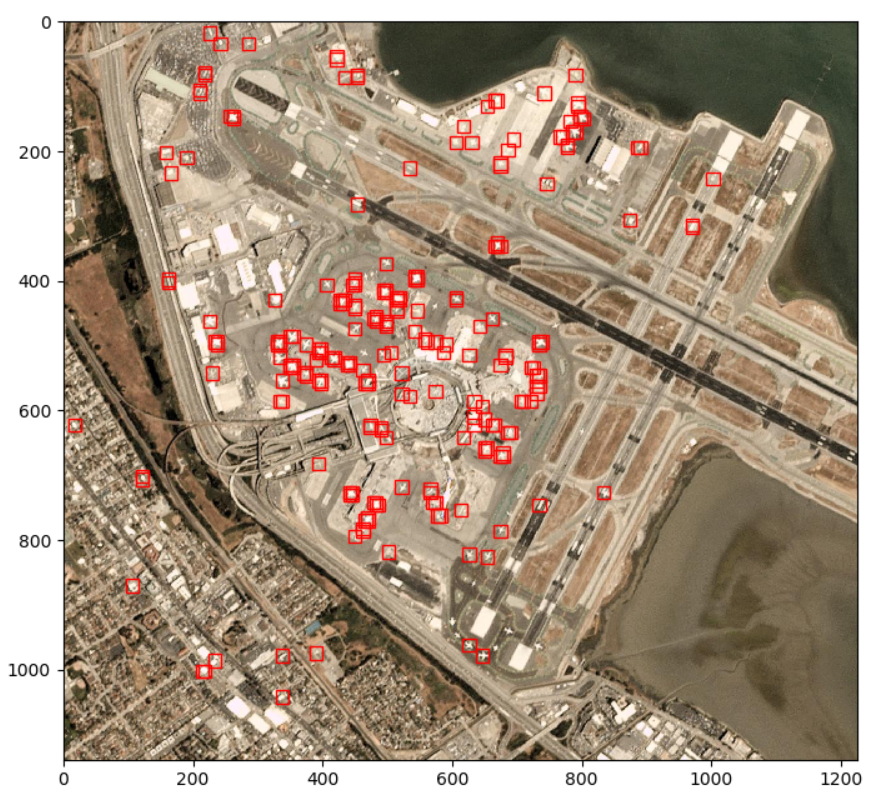
**CNN1**



**CNN2:**



**AlexNet:**



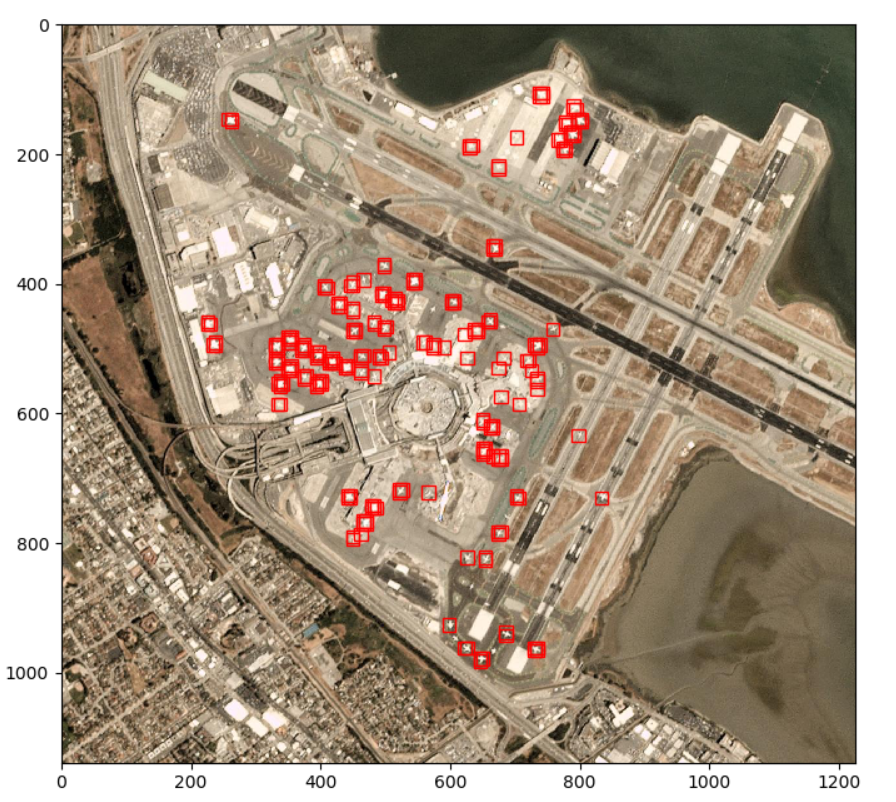


Missed to mark several planes in the above picture.



Also, mistook several white patches in the image to be aircrafts.

**VGGNet:**

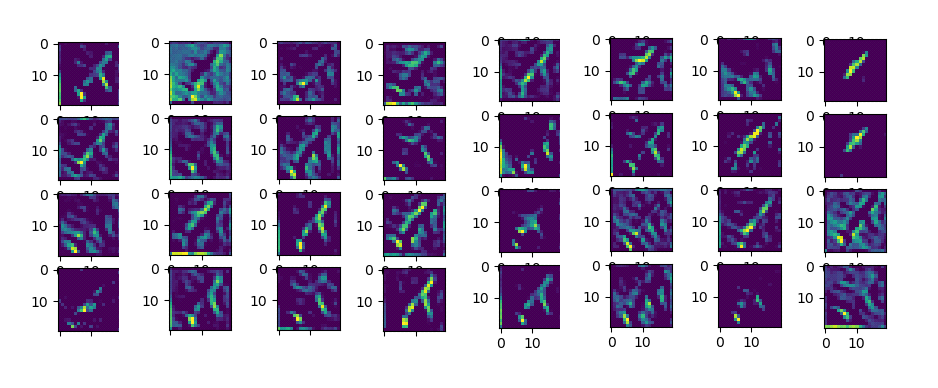


Let us see what's going on under the hood.

CNNs learn to identify features from the images without explicitly programming them to look for features. The features present in the images are learnt by the filters that scan through the image in the Convolutional layer. The weights learnt by the filters can be visualized to see the features learnt by the CNN.

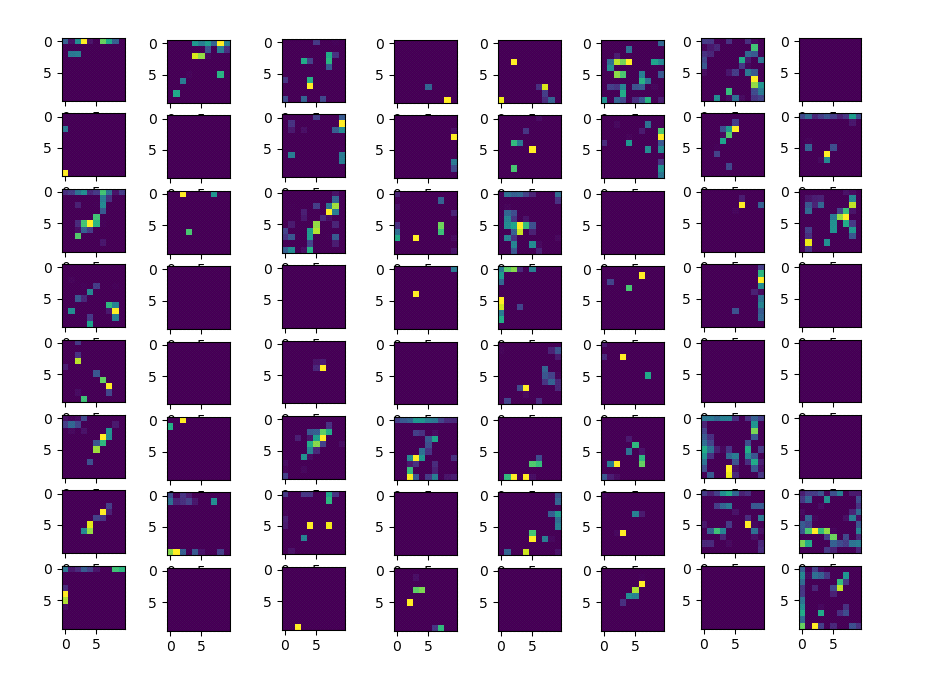
Here are the weights learnt by CNN2:

Weights learnt by the filters of the first convolutional layer with 32 filters.



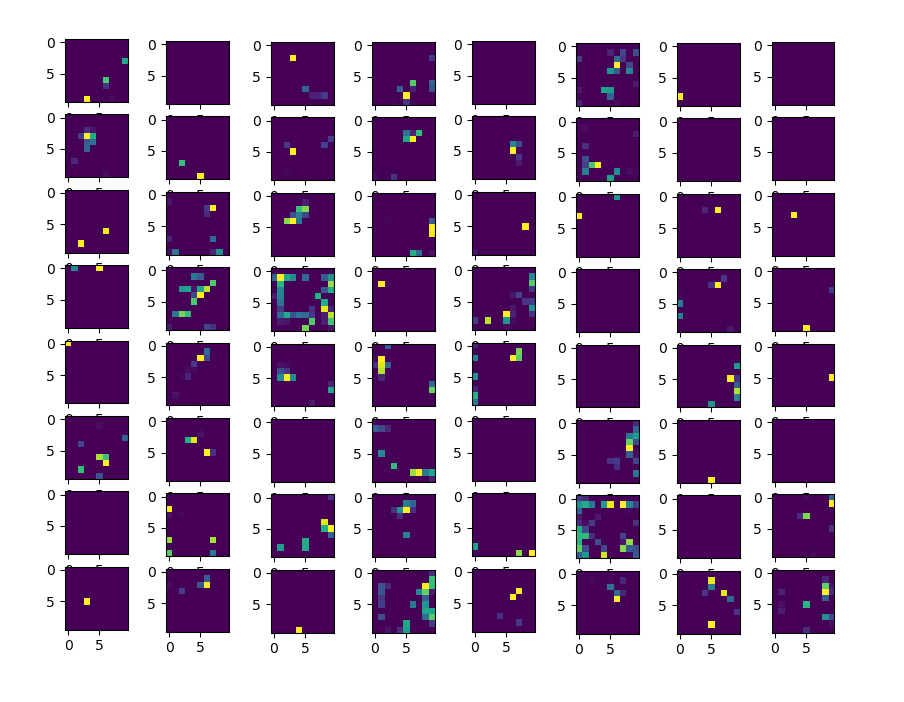
From the figure it can be seen that the filters have learnt to identify some features of the plane like the airplane's fuselage, left wing, right wing etc.

Weights learnt by the filters of the second convolutional layer with 64 filters.



These filers recognize lower level features when compared to the filters of the first convolutional layer.

Weights learnt by the filters of the third convolutional layer with 64 filters.



These filters identify the lowest level features in the convolutional neural network. Although the features look more like dots and less like features, they can represent important portion of an airplane that could be part of all airplane images. These filters are powerful enough to identify features that the human eye fails to observe.