Deep Learning and its Applications

CS671

Course Instructor : Dr. Aditya Nigam

Assignment 2 : Question 3

Network Visualization

Group 13

Aman Khandelwal	B16007
Amrendra Singh	B16010
Bharat Lodhi	B16015

https://github.com/ashking13th/deepLearning

https://students.iitmandi.ac.in/~b16010/CS671_grp_13/

1. Problem Statement

Implement network visualization tchniques.

Part 1: Visualizing Intermediate Layer Activations

Intermediate layer activations are the outputs of intermediate layers of the neural network. You are required to plot the intermediate activations of the layers of the neural networks made in section 12 for at least 6 images.

Part 2: Visualizing Convnet Filters

In this part you have to visualize the lters in the convolutional neural network. This can be done by running Gradient Descent on the value of a convnet so as to maximize the response of a specic lter, starting from a blank input image. You are required to plot the lters of the layers of the neural networks made in section 12.

Part 3: Visualizing Heat maps of class activations

In this part you have to visualize the lters in the convolutional neural network. This can be done by running Gradient Descent on the value of a convnet so as to maximize the response of a specic lter, starting from a blank input image. You are required to plot the lters of the layers of the neural networks made in section 12.

2. Experiments and Results

1 Visualizing Intermediate Layer Activations

MNIST Dataset

For the sample image of 3

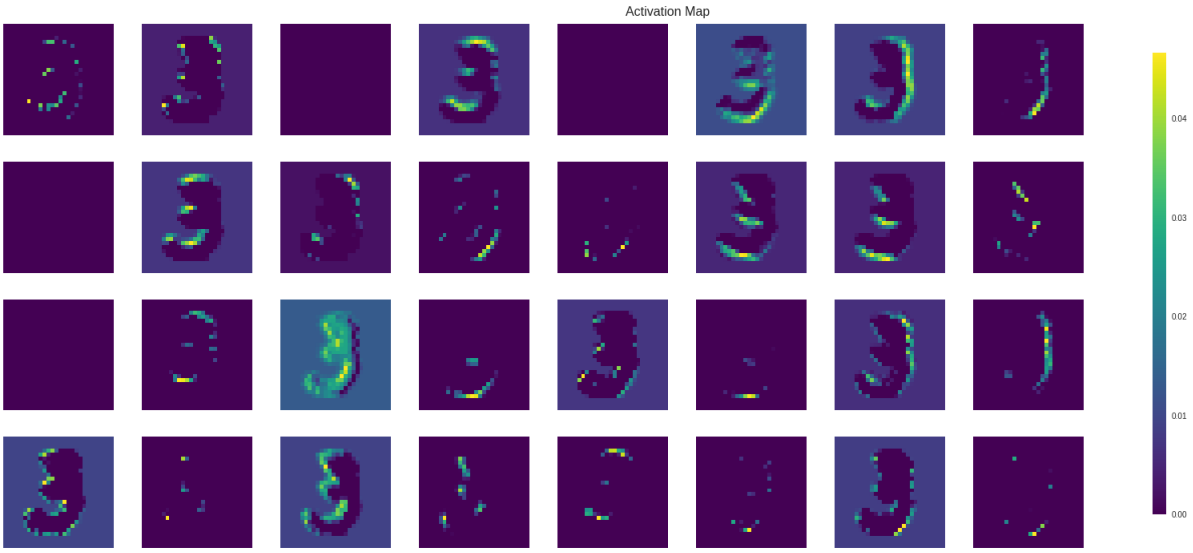


Figure 2..1

For the sample image of 2

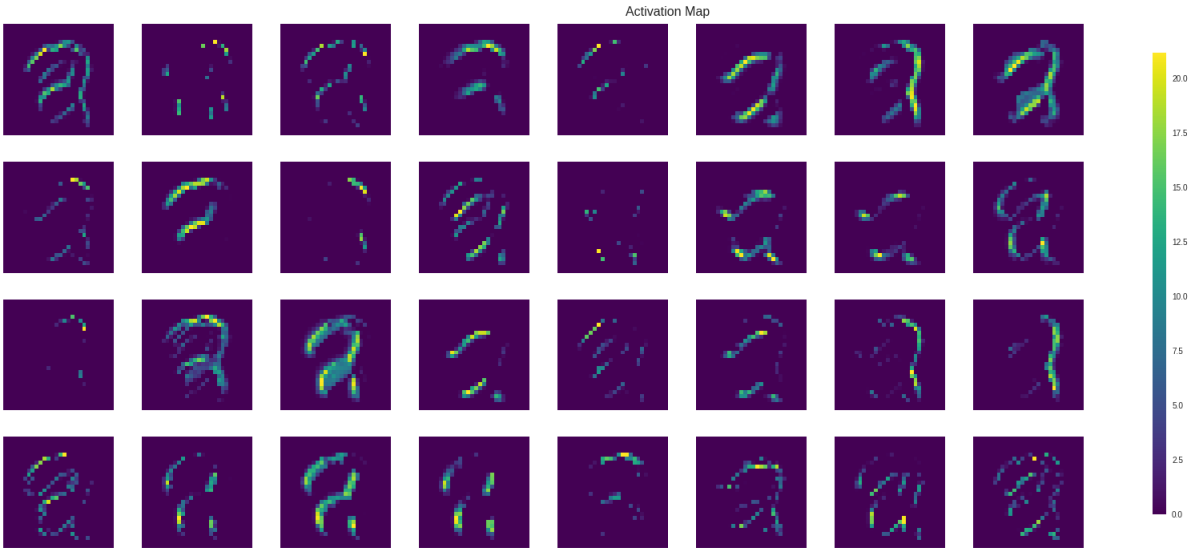


Figure 2..2

For the sample image of 9

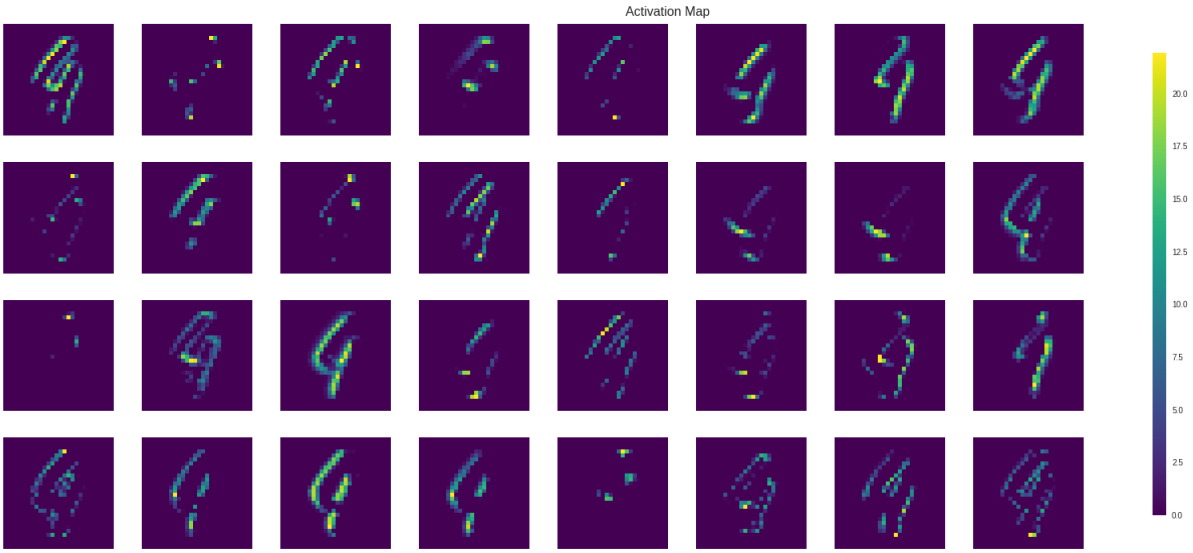


Figure 2..3

For the sample image of 6

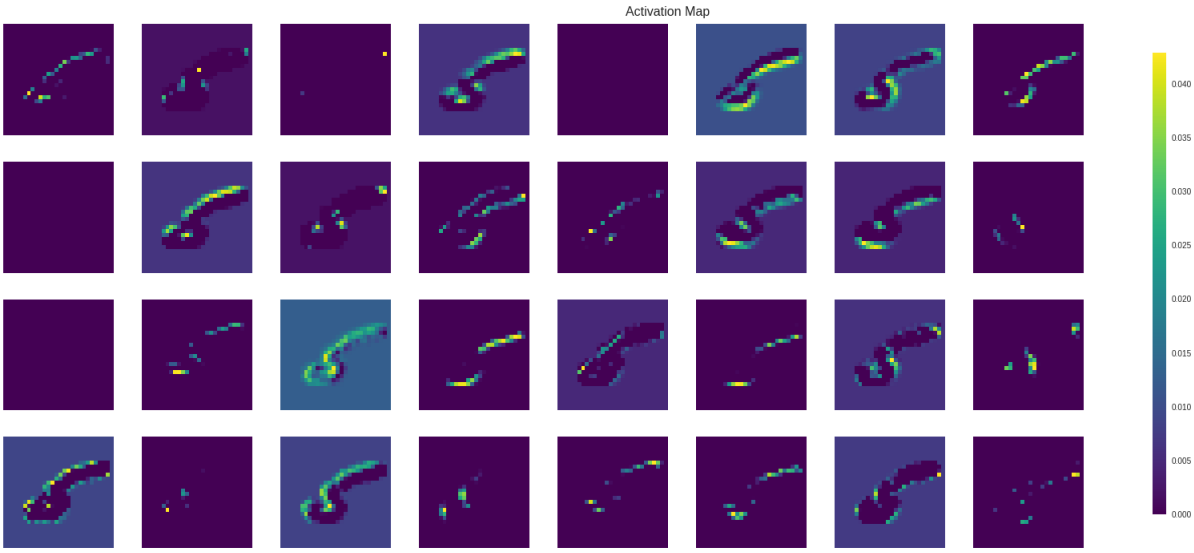


Figure 2..4

For the sample image of 7

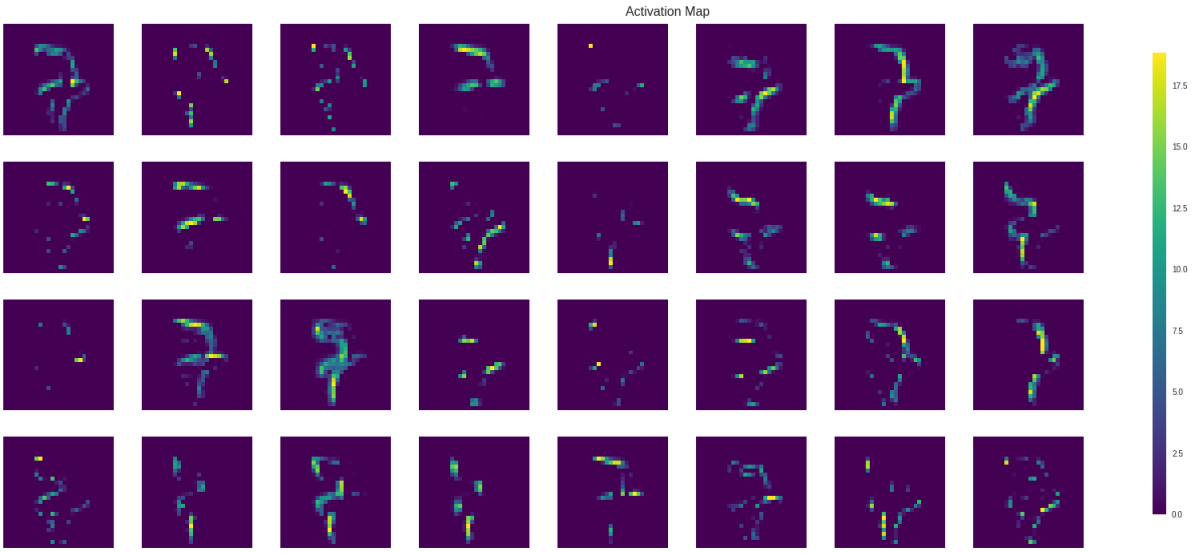


Figure 2..5

For the sample image of 9

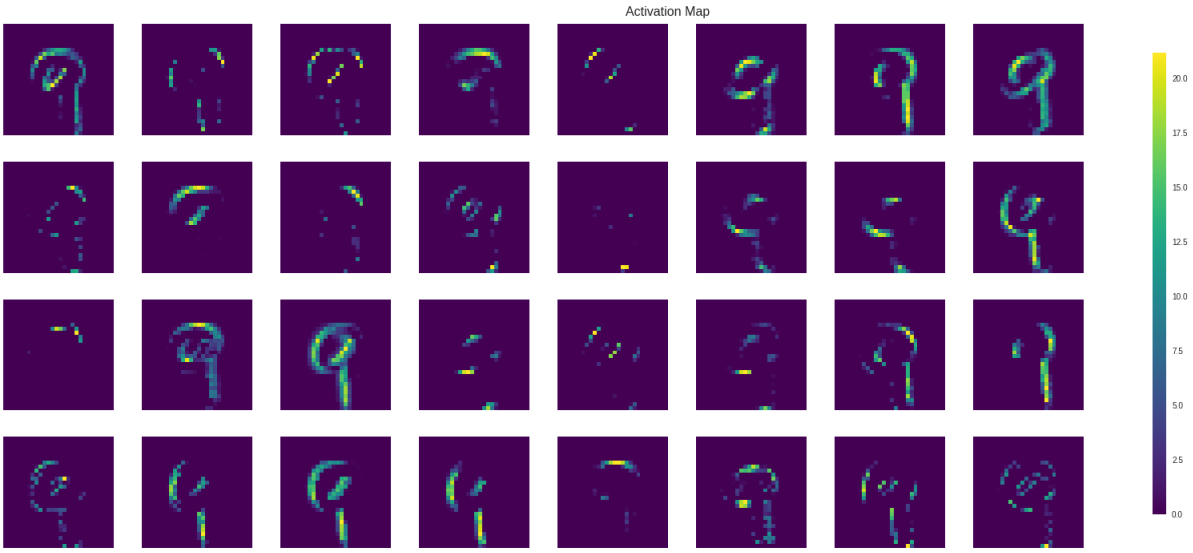
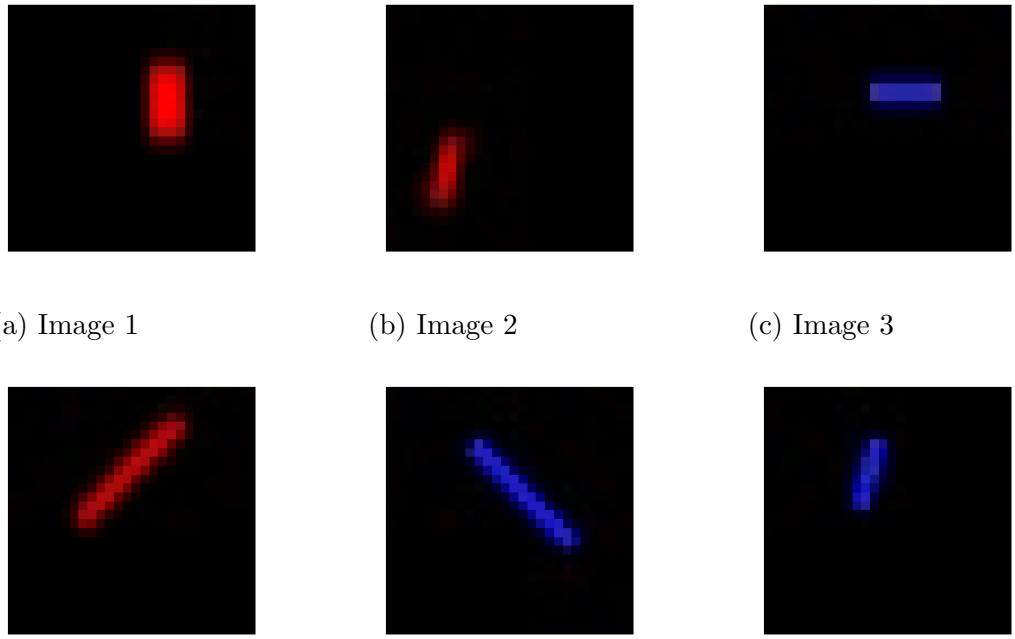


Figure 2..6

Line Dataset



(a) Image 1 (b) Image 2 (c) Image 3

(d) Image 4 (e) Image 5 (f) Image 6

Figure 2..7. Original Images

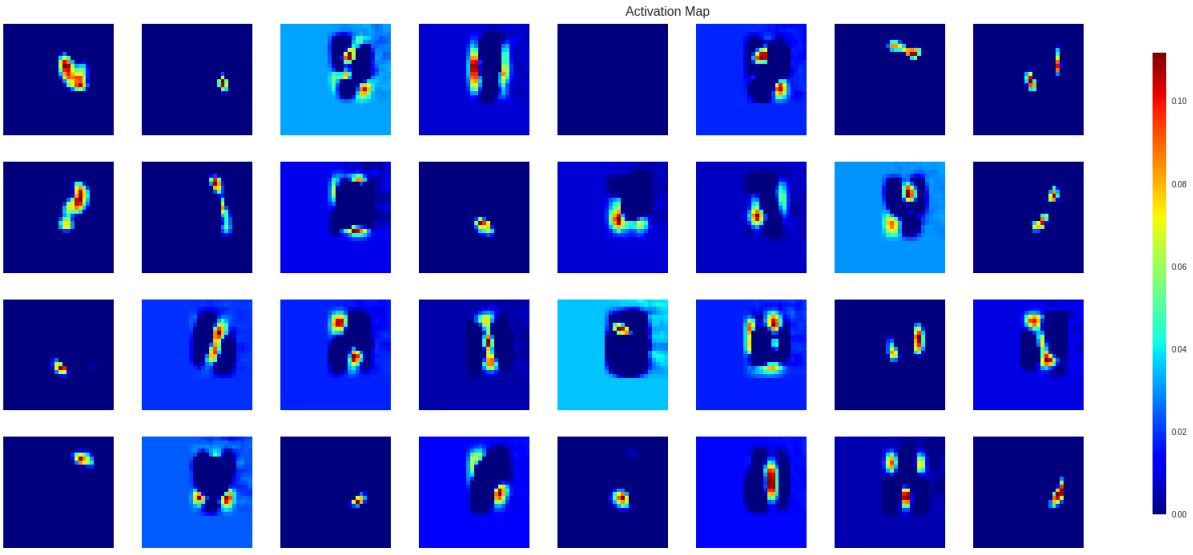


Figure 2..8. Activation visualization for image 1

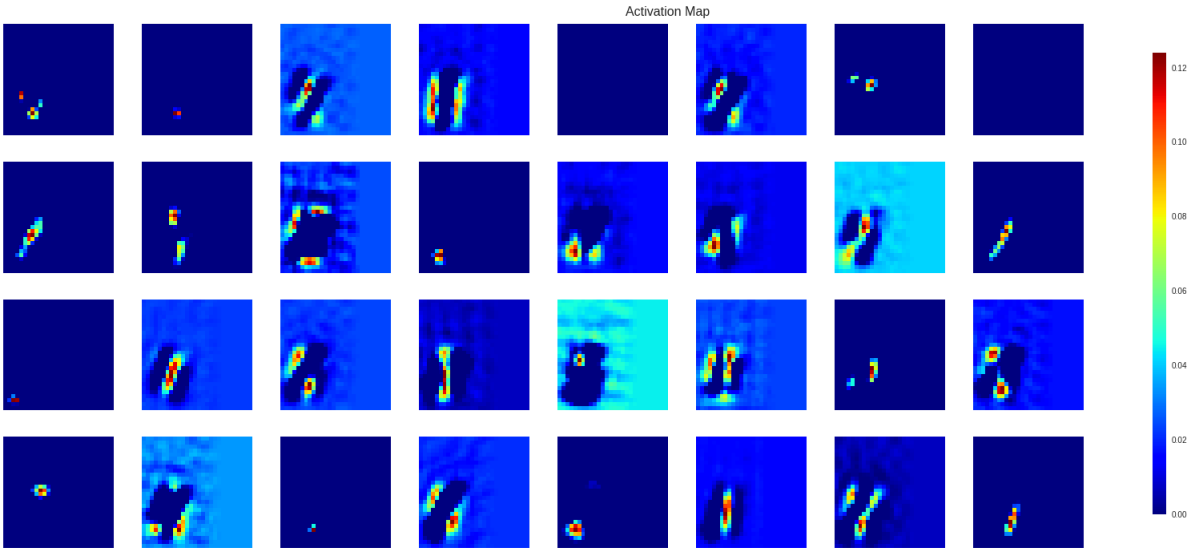


Figure 2..9. Activation visualization for image 2

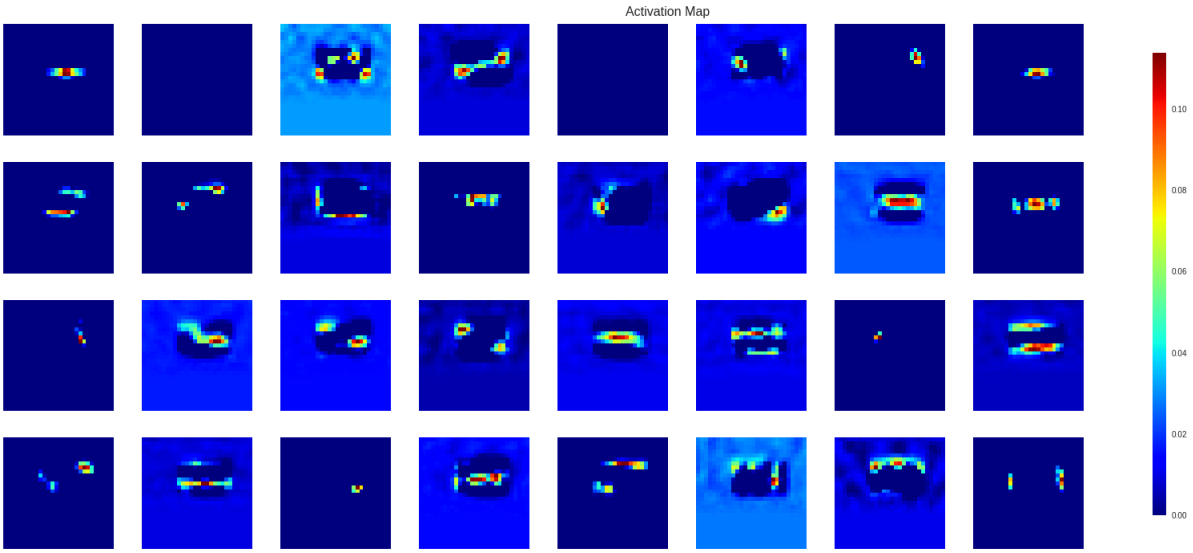


Figure 2..10. Activation visualization for image 3

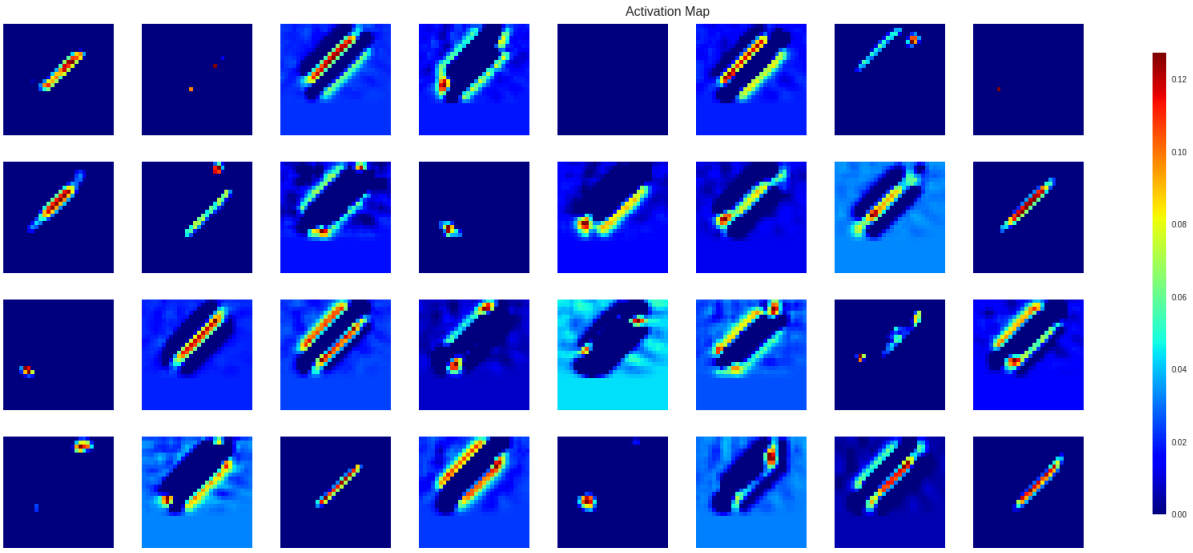


Figure 2..11. Activation visualization for image 4

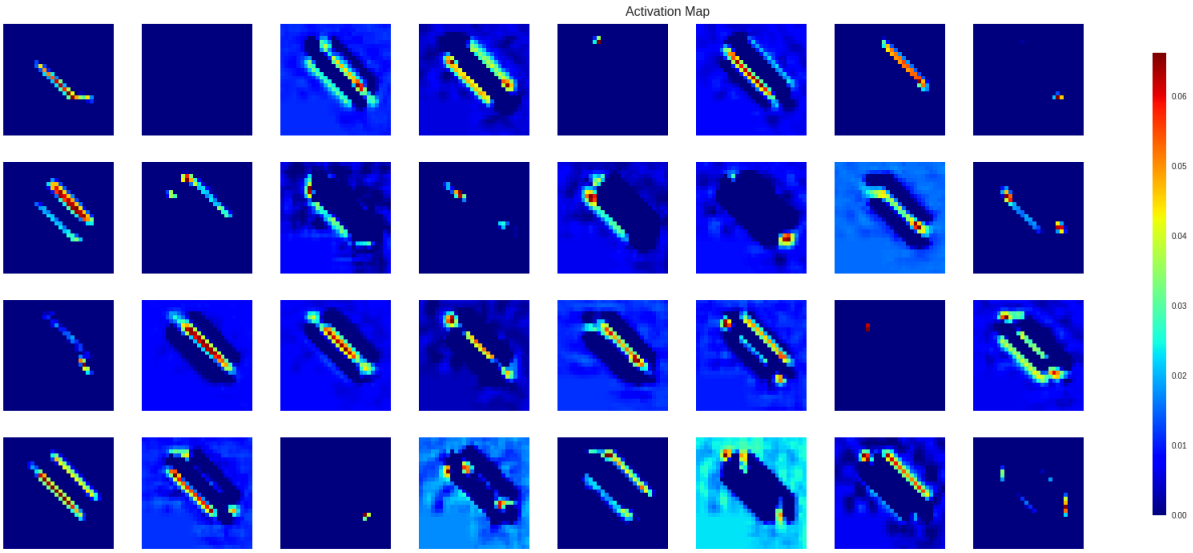


Figure 2..12. Activation visualization for image 4

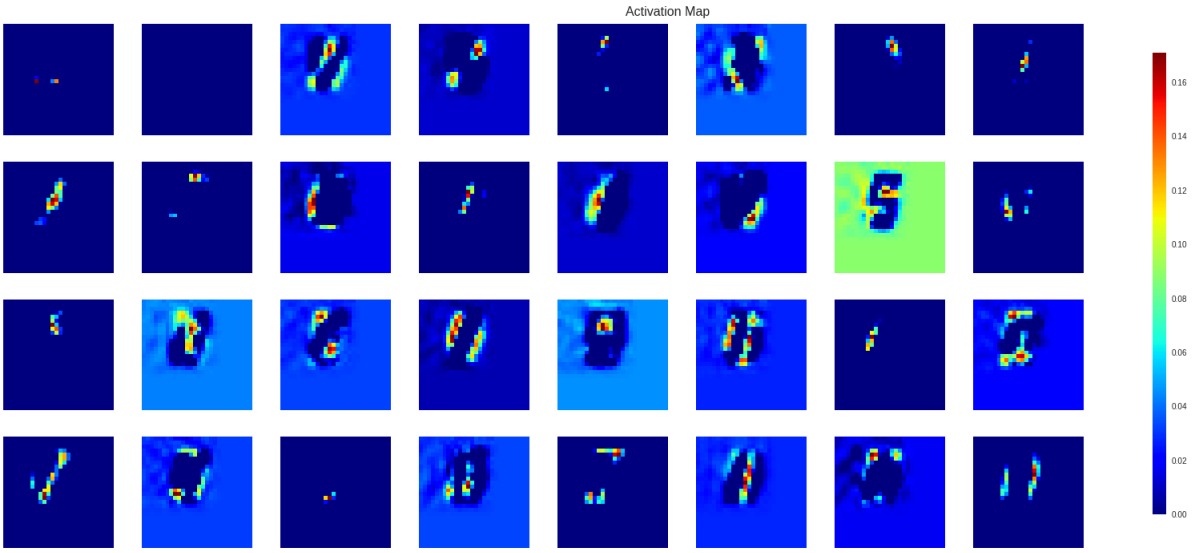


Figure 2..13. Activation visualization for image 6

1.1 Model 2 (Question 2)

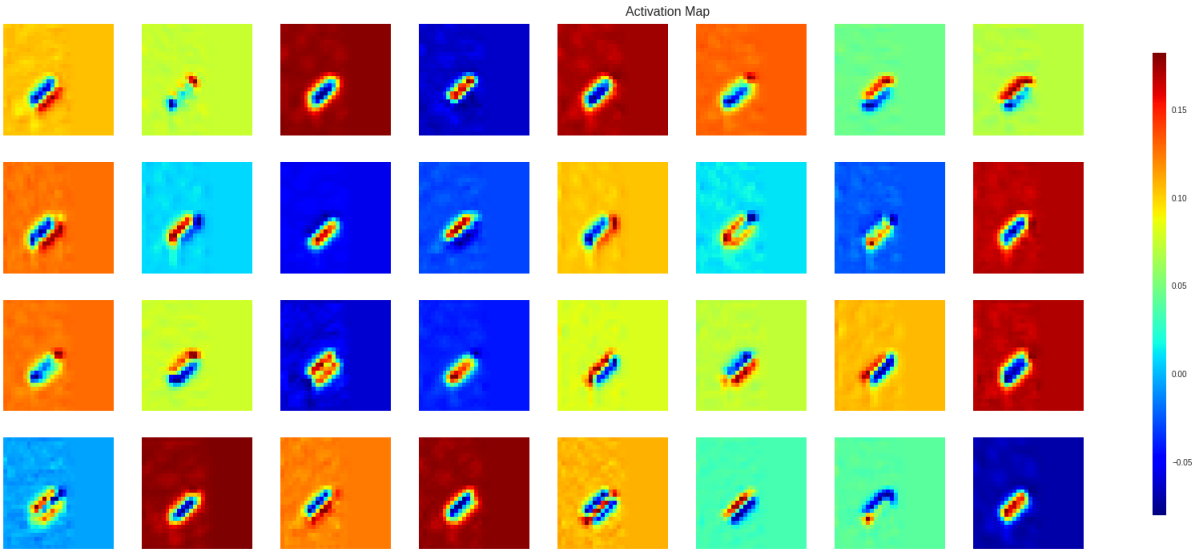


Figure 2..14. Activation visualization Conv Layer 1

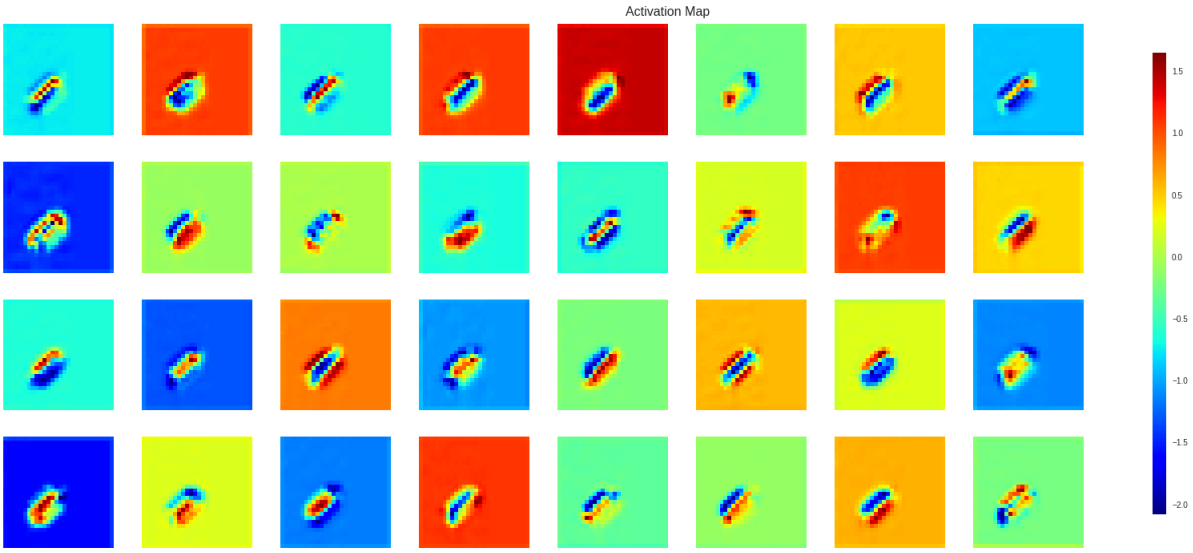


Figure 2..15. Activation visualization Conv Layer 2

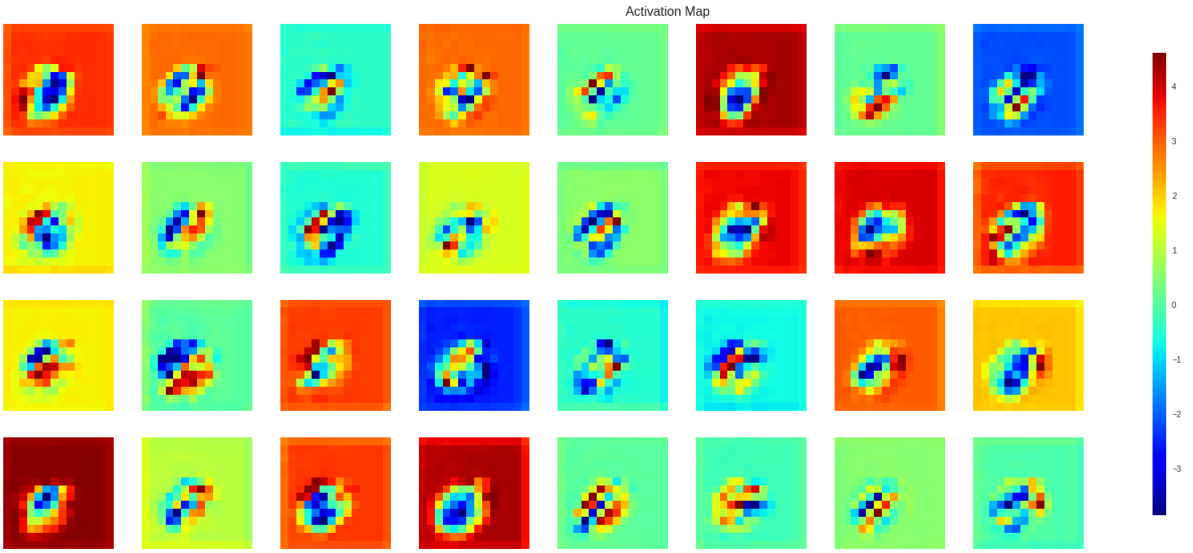


Figure 2..16. Activation visualization Conv Layer 3

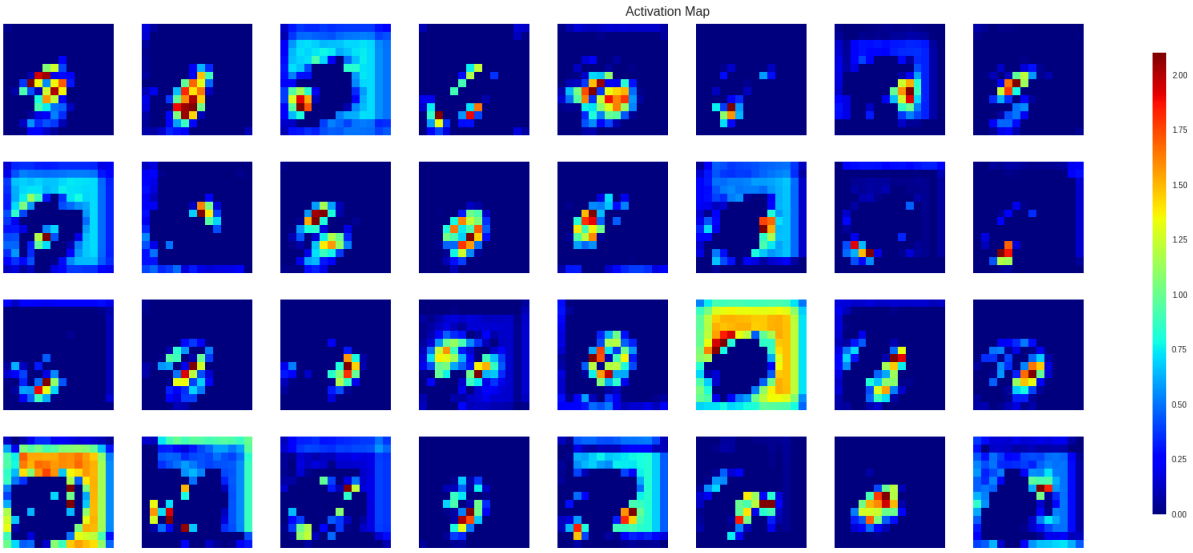


Figure 2..17. Activation visualization Conv Layer 4

2 Visualizing Convolution network Filters

Following are the visualizations of the filters of convolution layers of the models: For the following four images, in first 3 images only top 25 filters for which loss was maximum are plotted and last plot is for all the 64 filters.

2.1 Model 1 (Question 1

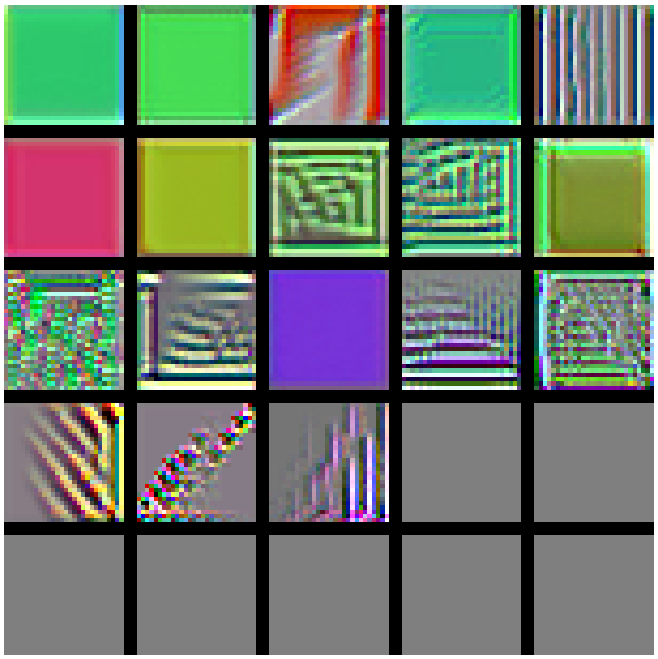


Figure 2..18

2.2 Model 2 (Question 2

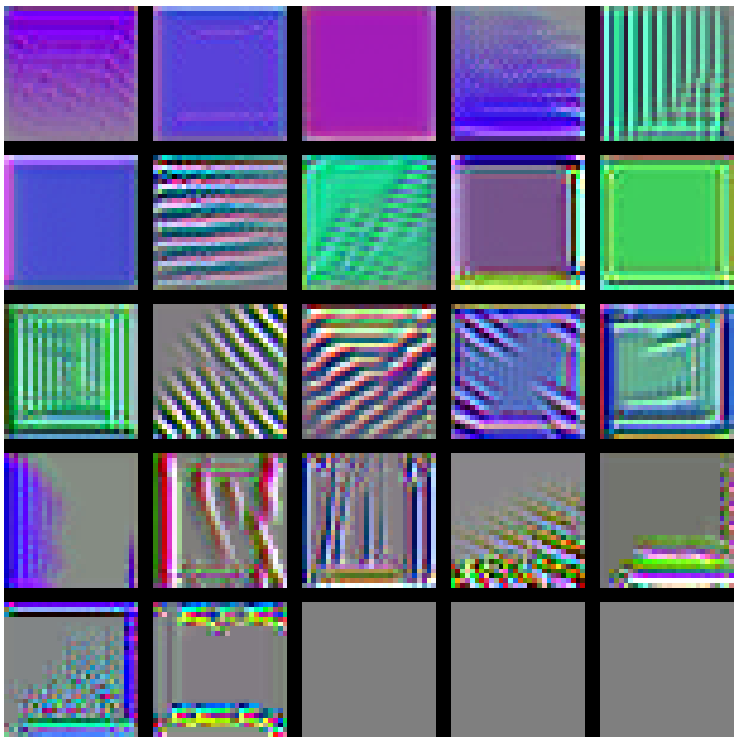


Figure 2..19

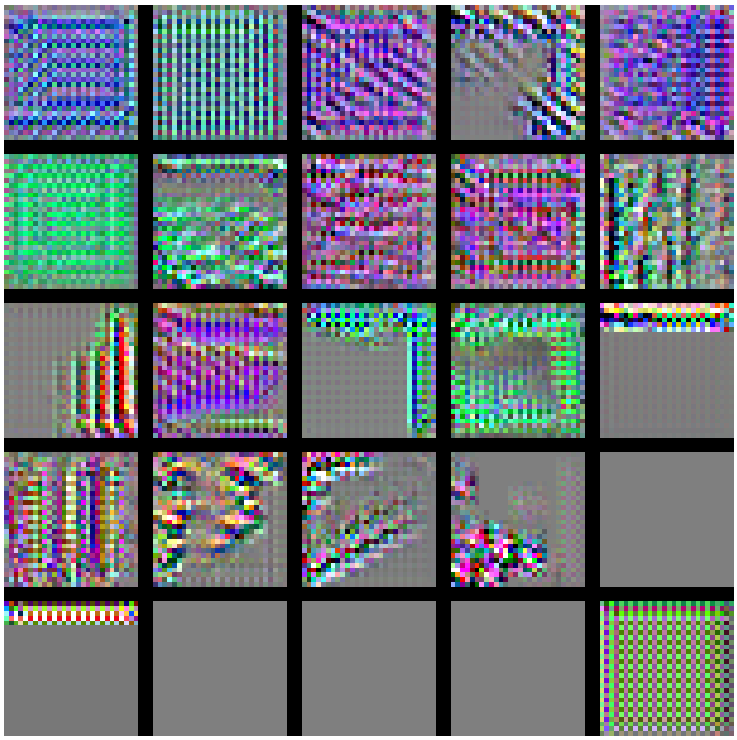


Figure 2..20

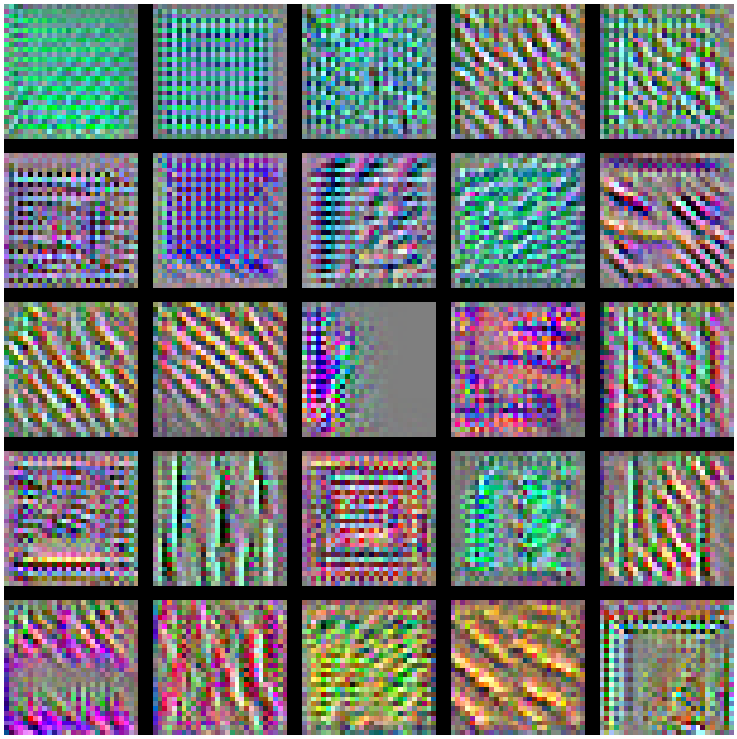


Figure 2..21

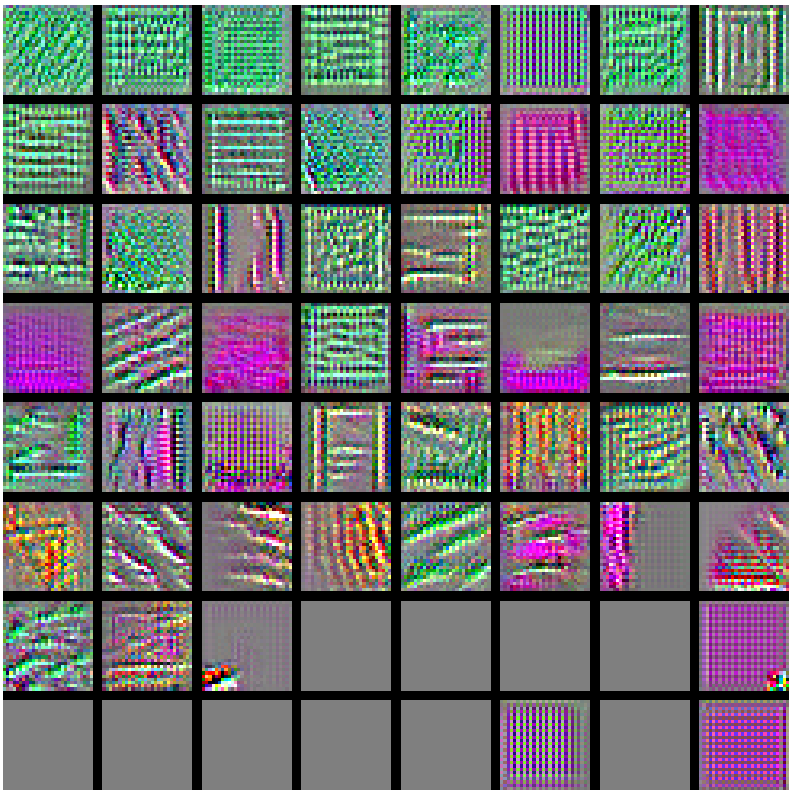


Figure 2..22

3 Visualizing Heat maps of class activations

3.1 CNN (Model for Question 1)

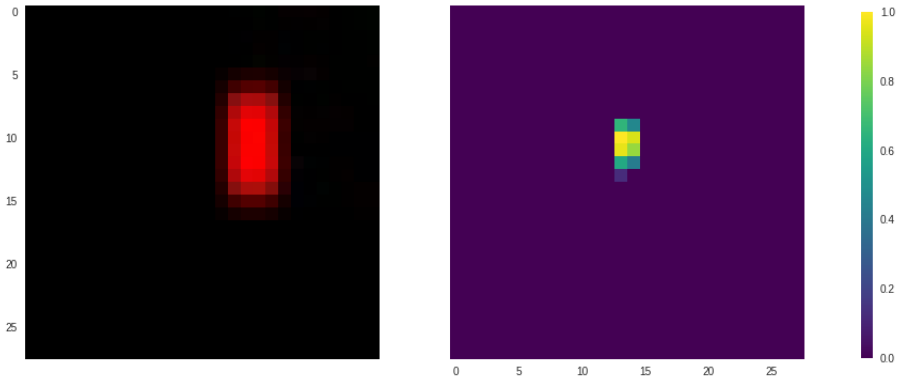


Figure 2..23

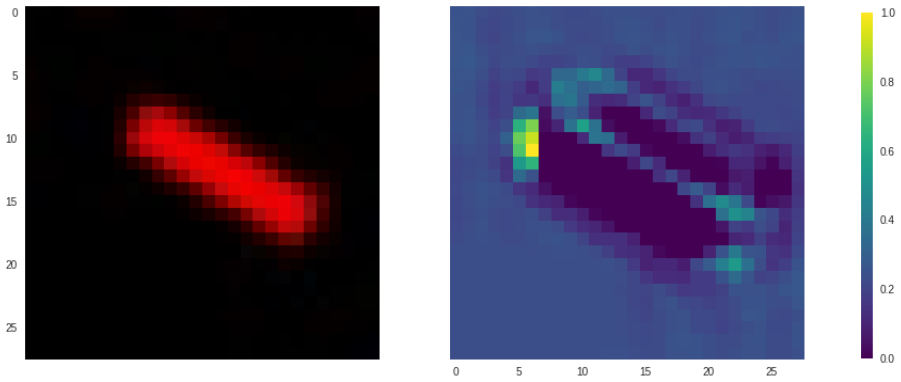


Figure 2..24

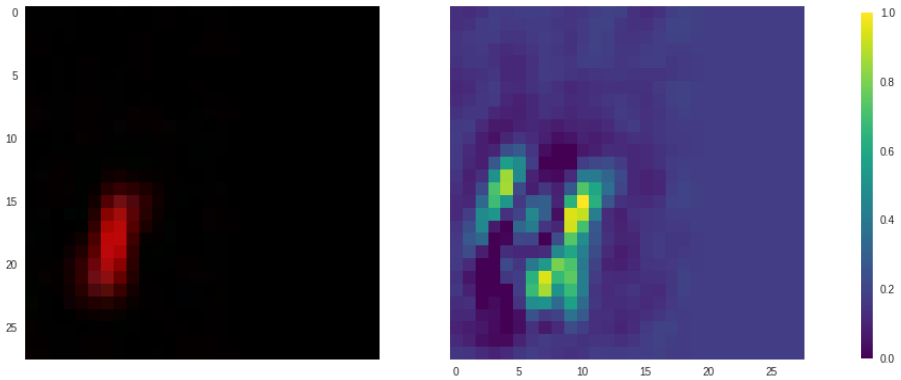


Figure 2..25

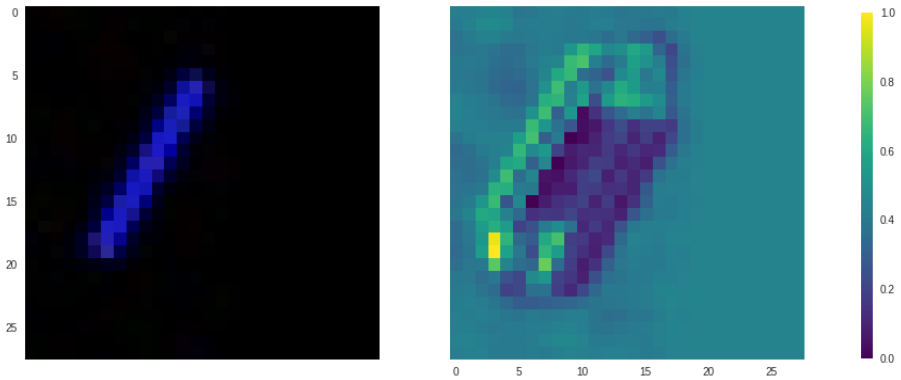


Figure 2..26

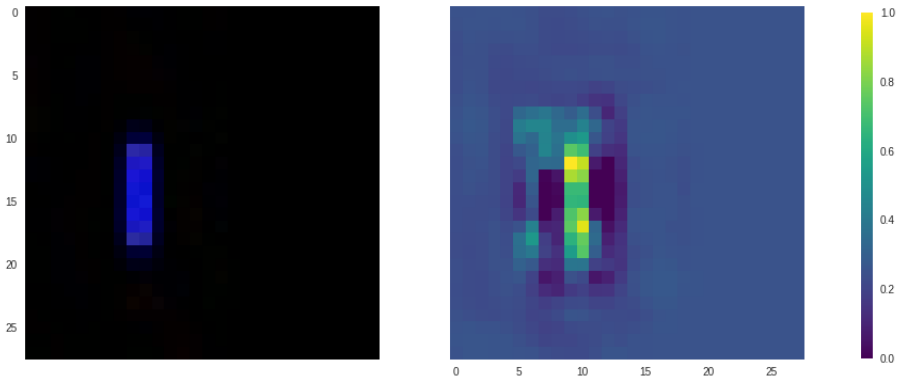


Figure 2..27

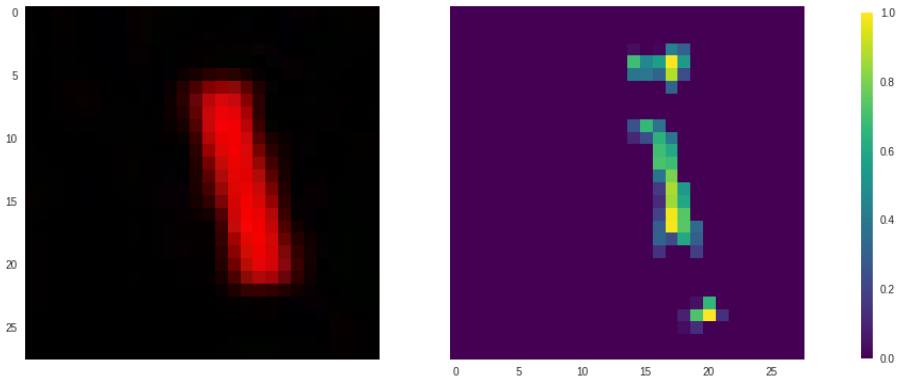


Figure 2..28

3.2 Multihead Classification (Model for Question 2)

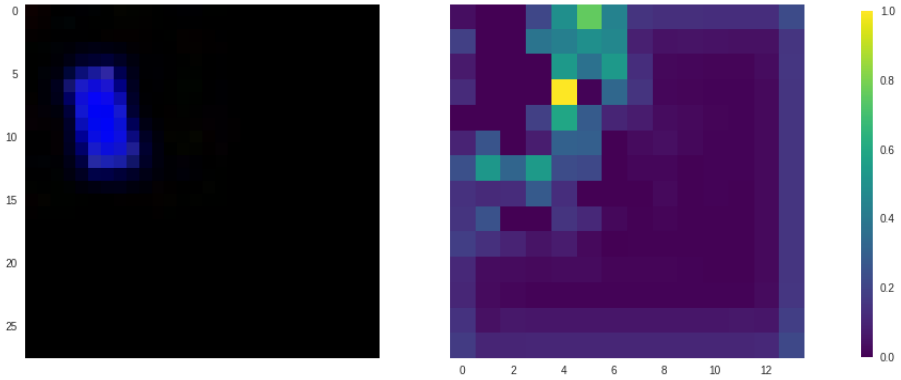


Figure 2..29

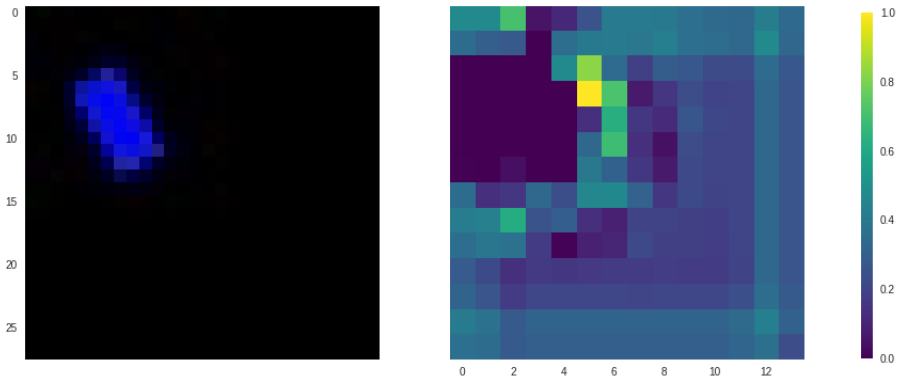


Figure 2..30

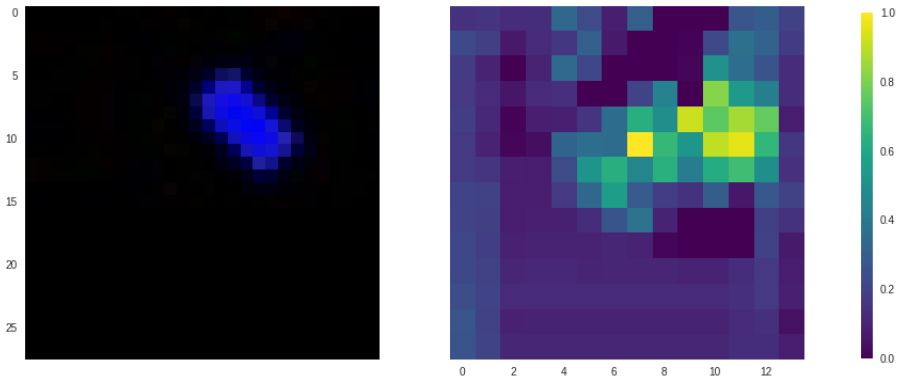


Figure 2..31

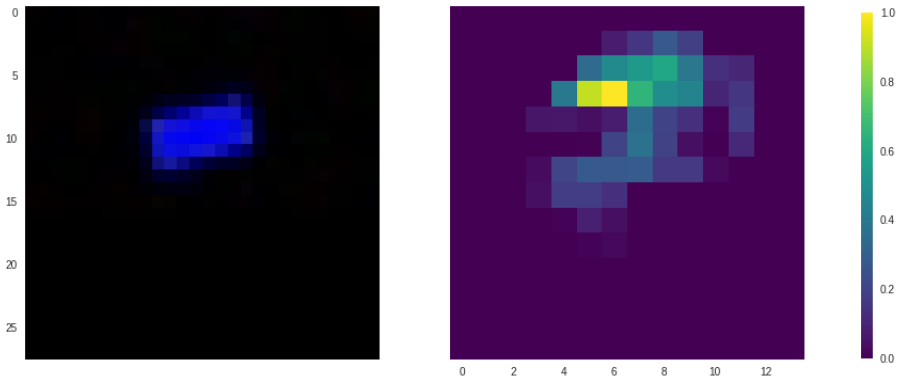


Figure 2..32

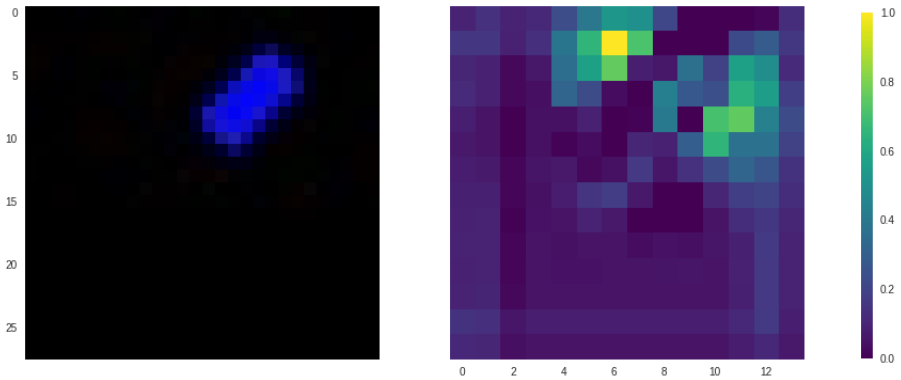


Figure 2..33

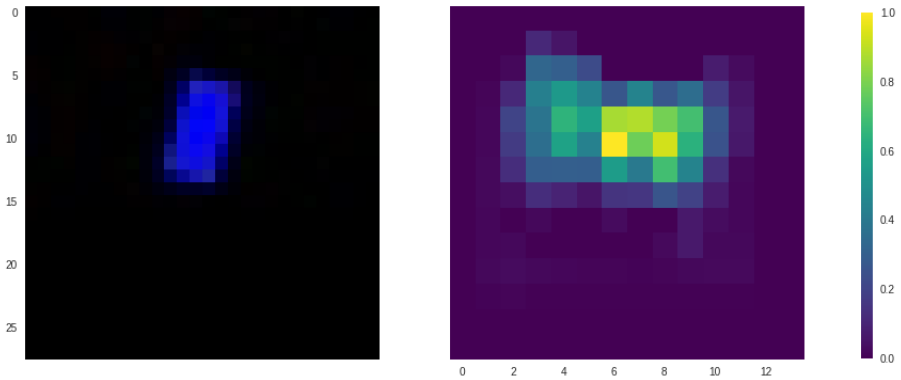


Figure 2..34

3. Inferences

- The first layer acts as a collection of various edge detectors
- As we go higher-up, the activations become increasingly abstract and less visually interpretable
- The sparsity of the activations is increasing with the depth of the layer: in the first layer, all filters are activated by the input image, but in the following layers more and more filters are blank. This means that the pattern encoded by the filter isn't found in the input image.
- The filters from the first layer in the model encode simple edges and colors
- The filters from further layers encode simple textures made from combinations of edges and colors
- Further layers seem to learn more of the entire textures of the images themselves
- The heatmaps highlight the regions the particular layer sees.
- We can observe that the heatmap is more prominent in the regions where there is a line present in case of the line dataset
- We can observe that the heatmaps show that regions with the object/line in the images have been highlighted. This also proves that the model has has been trained correctly.

Bibliography

[1] Stack Overflow
https://stackoverflow.com

[2] Understanding your Convolution network with Visualizations
https://towardsdatascience.com/understanding-your-convolution-\
network-with-visualizations-a4883441533b

[3] Tensorflow Tutorials
https://www.tensorflow.org/tutorials

[4] Data Science Blog
https://medium.com/datadriveninvestor/my-take-at-the-mnist-dataset\
-97304dff2057

[5] Analytics Vidha
https://www.analyticsvidhya.com/blog/2016/10/
an-introduction-to-implementing-neural-networks-using-tensorflow/