GoogLeNet and InceptionNet

Paper 1: GoogLeNet: Going Deeper with Convolutions

Paper 2: Inception v3: Rethinking the Inception Architecture for Computer Vision

1. What are the drawbacks presented in paper 1 for uniformly increased network size?

2. Describe how an inception module handles multi-scale processing?

3. Paper 1 uses auxiliary classifiers (similar to deep supervision) connected to intermediate layers, what is the author's explanation for this?

4. 1x1 convolutions have a dual purpose. Explain.

5. Why it is preferred to use 1x1 convolutions inside inception modules? how does it help?

6. Explain in detail :), why did the author cite reference [1] in paper 1?

7. What changes were made inside an inception module in paper 2 as compared with the inception modules in paper 1?

8. What are the two problems described in paper 2, which happen are a result of the model being too confident about its classification prediction?

9. What are the general design principles presented in paper 2. Give at least 4 points.

10. What are the advantages of factorizing a convolution are presented in paper 2? does this factorization leads to any loss of expressiveness for the neural