Project #4

Huiyu Wang
604–592–364

1 bark

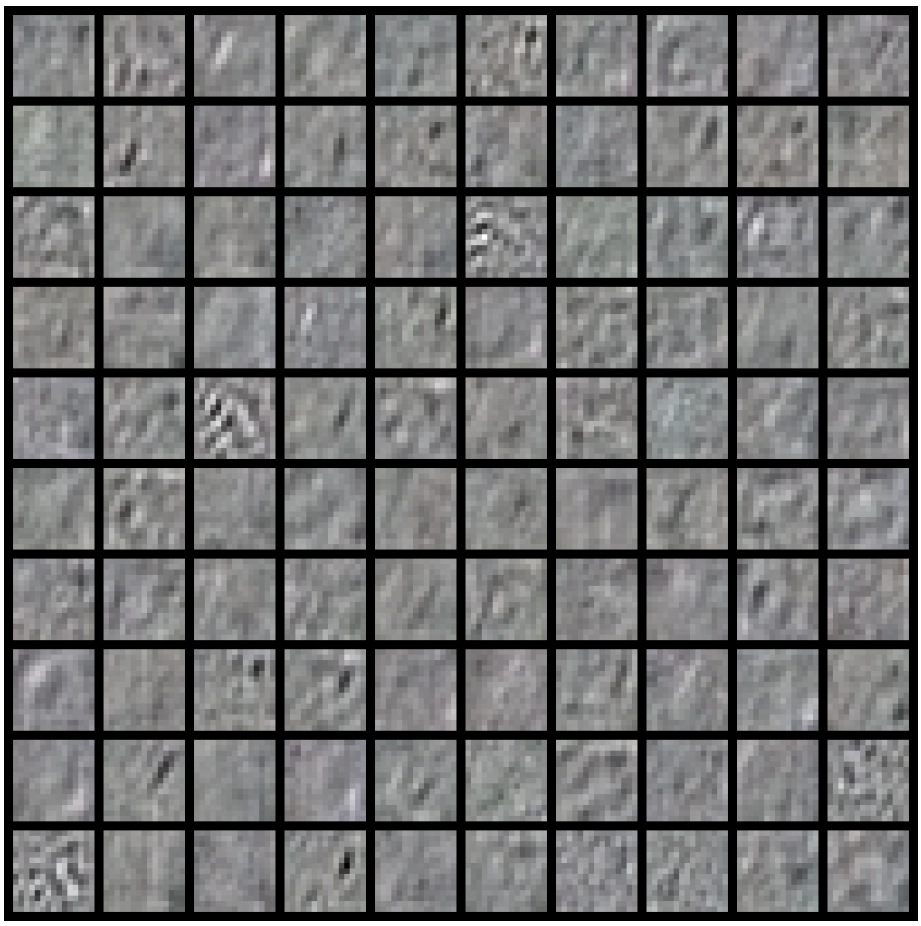


Figure 1: Normalized filters in the first layer

(a) layer 1 synthesis (b) layer 2 synthesis

(c) layer 3 synthesis (d) original image

Figure 2: Synthesized Images (bark)

2 beehive

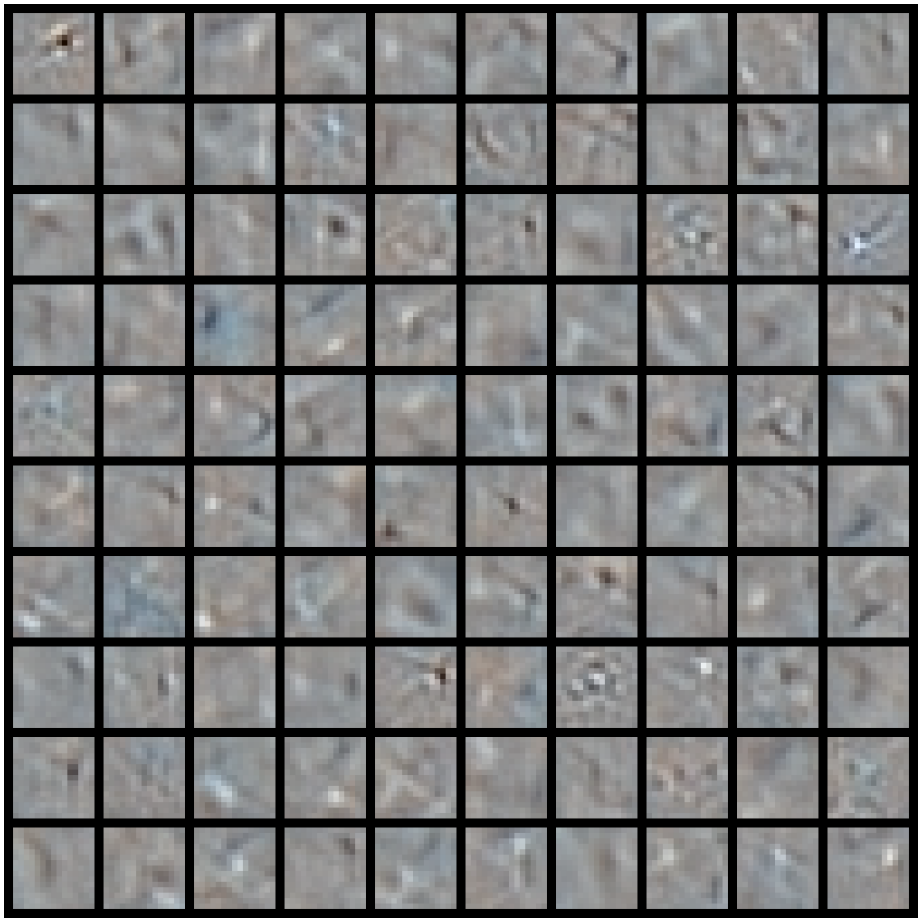
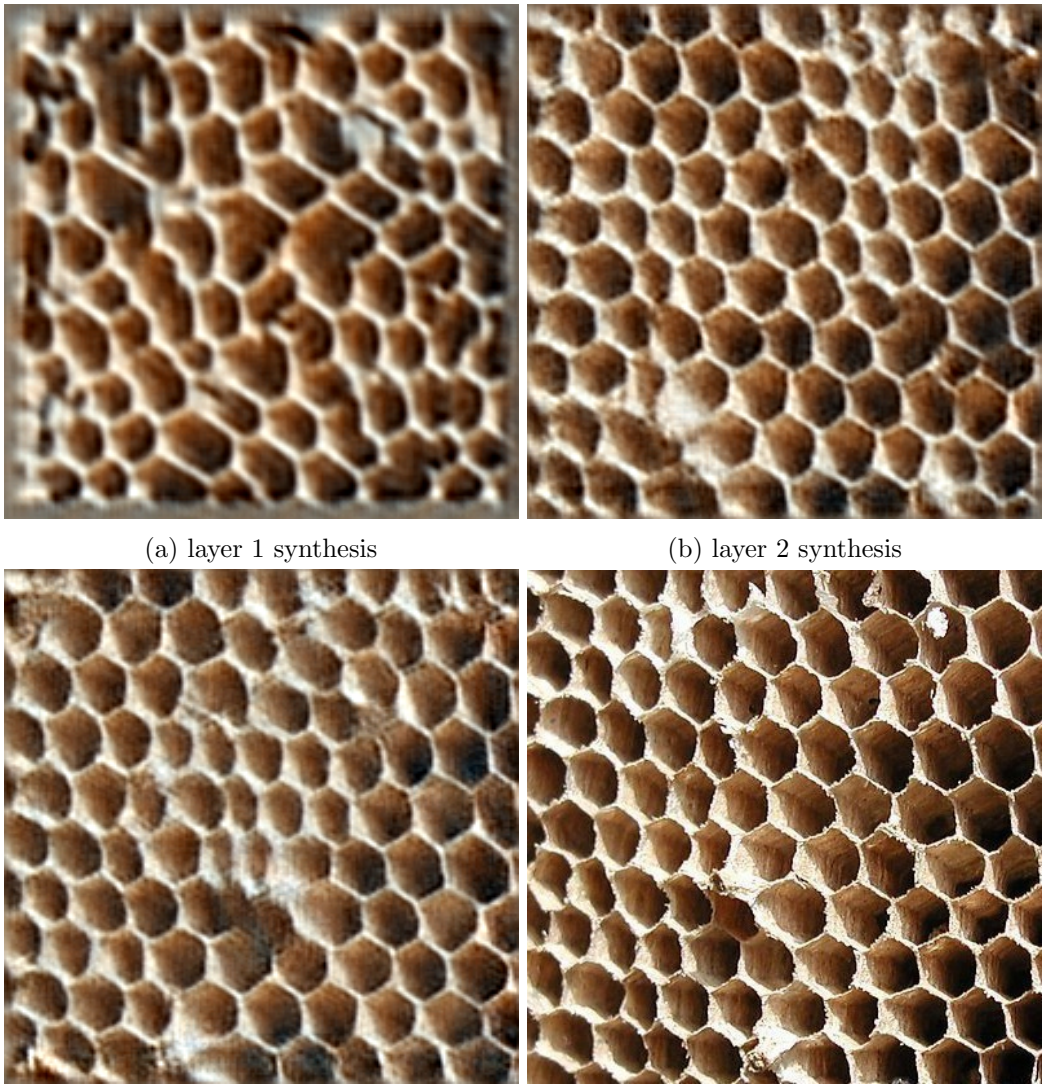


Figure 3: Normalized filters in the first layer



(a) layer 1 synthesis (b) layer 2 synthesis

(c) layer 3 synthesis (d) original image

Figure 4: Synthesized Images (beehive)

3 coffee

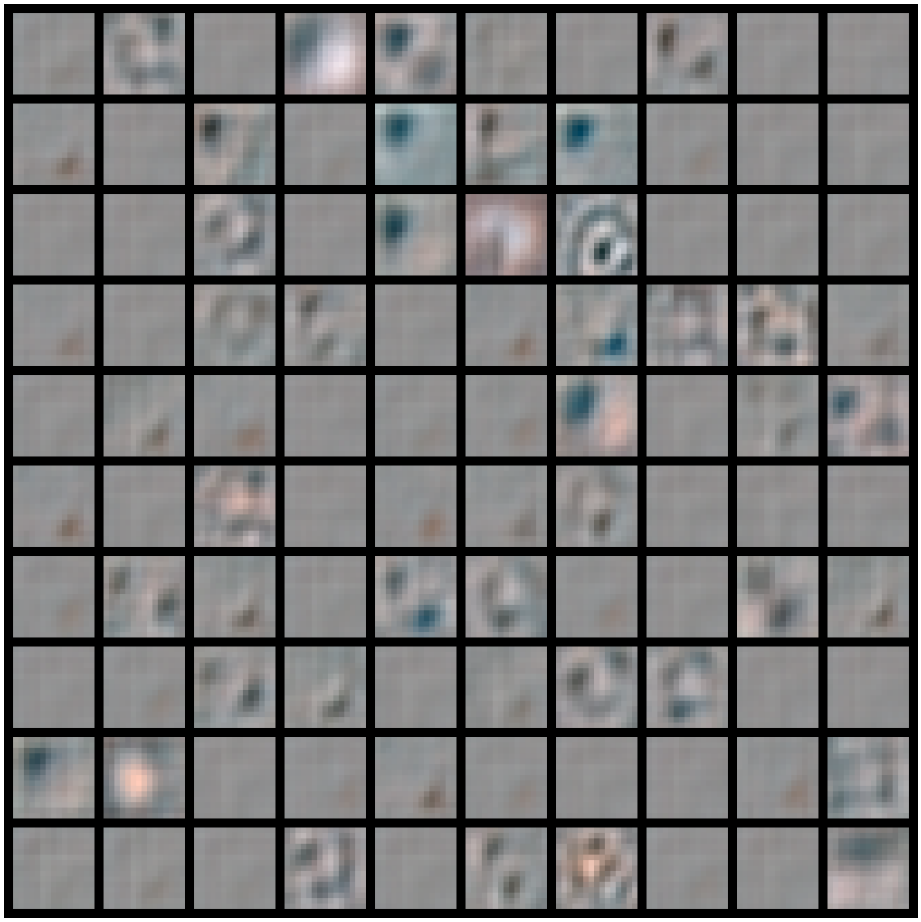
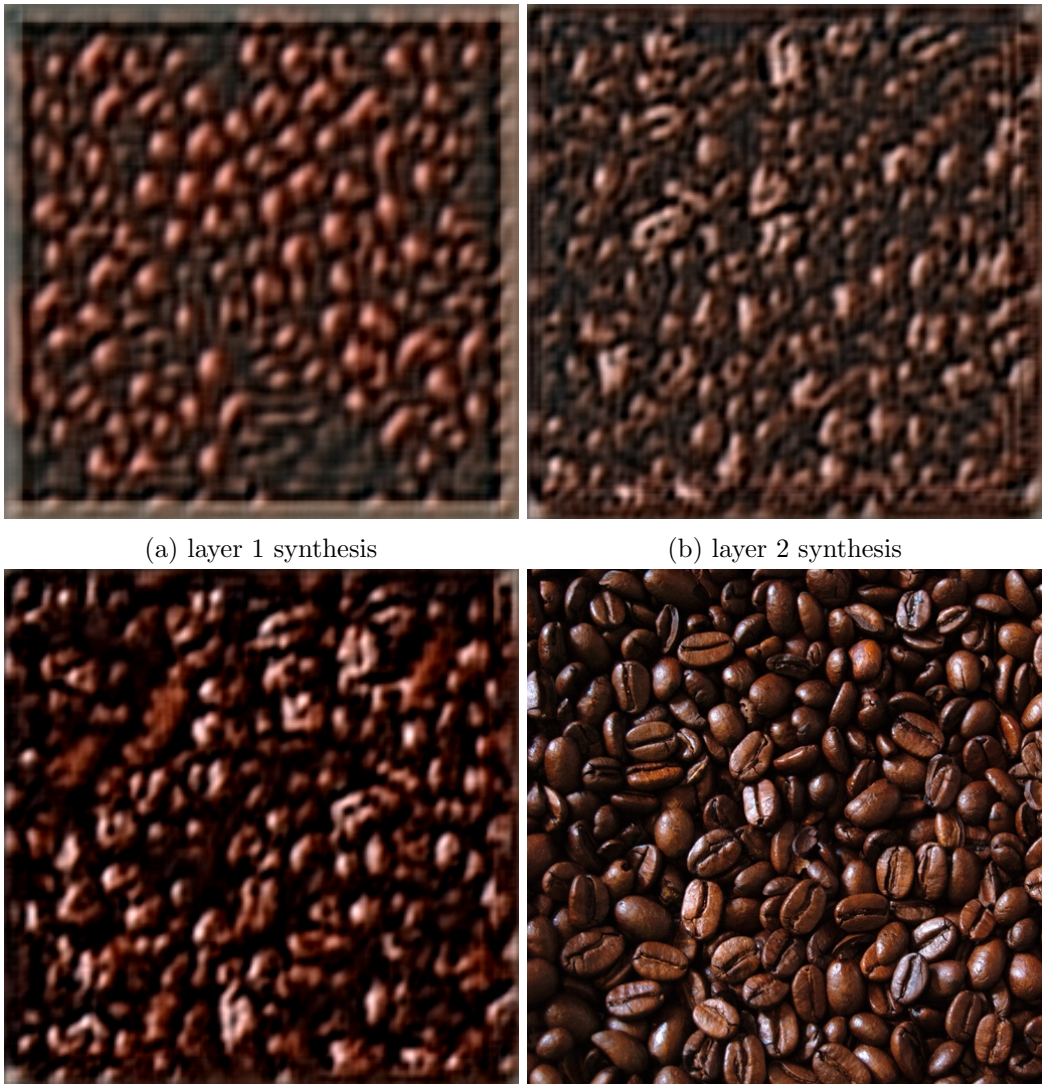


Figure 5: Normalized filters in the first layer



(a) layer 1 synthesis (b) layer 2 synthesis

(c) layer 3 synthesis (d) original image

Figure 6: Synthesized Images (coffee)

4 rose

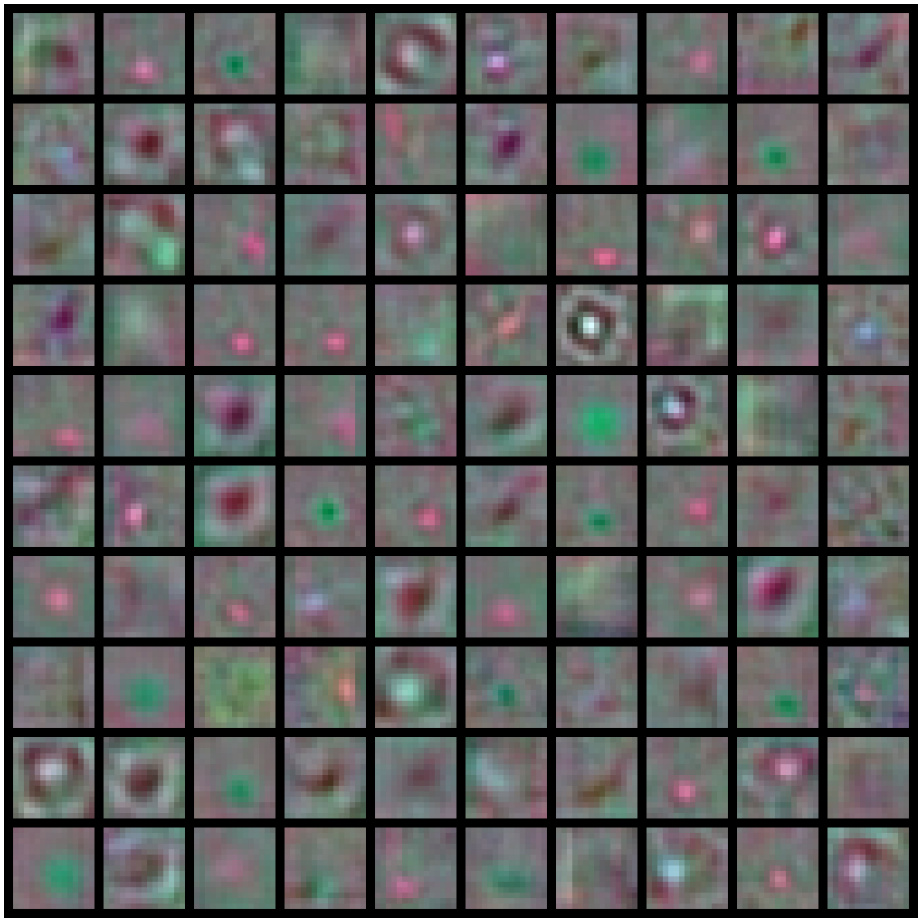
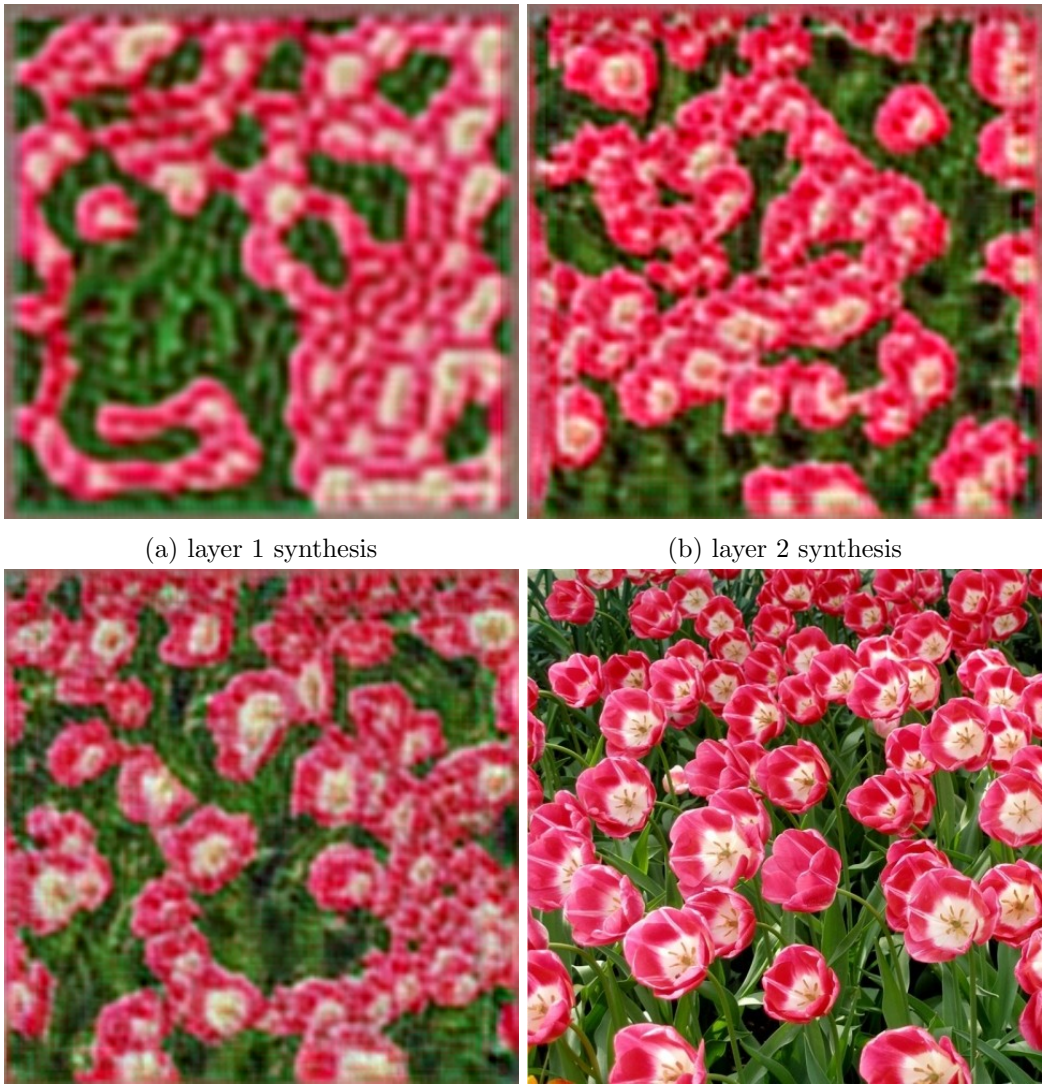


Figure 7: Normalized filters in the first layer



(a) layer 1 synthesis (b) layer 2 synthesis

(c) layer 3 synthesis (d) original image

Figure 8: Synthesized Images (rose)

5 stucco

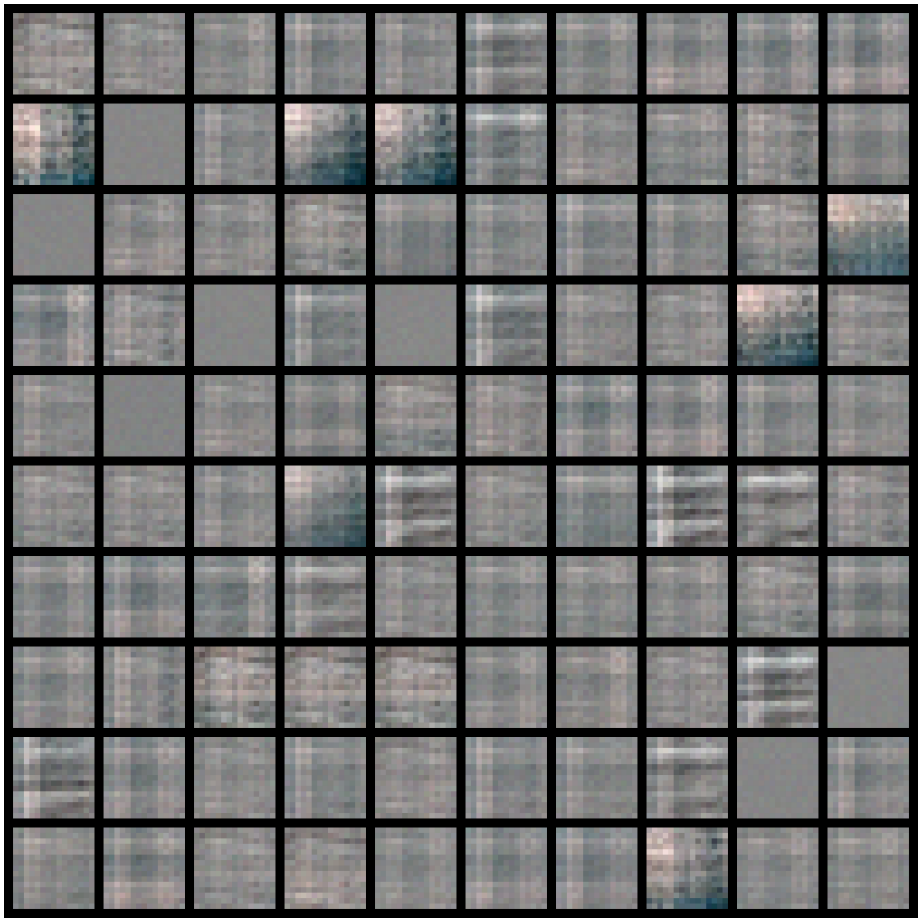


Figure 9: Normalized filters in the first layer

(a) layer 1 synthesis

(b) layer 2 synthesis

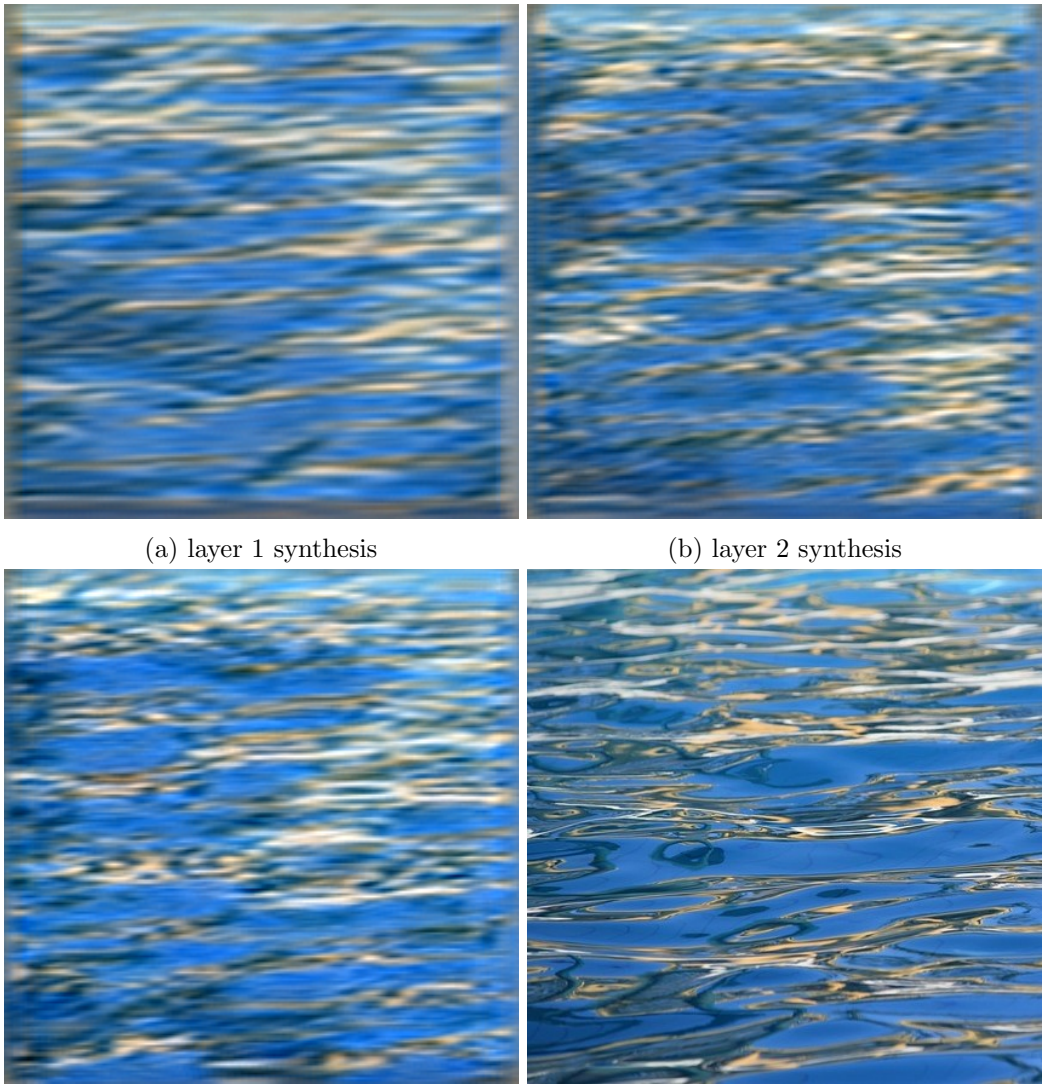
(c) layer 3 synthesis

(d) original image

Figure 10: Synthesized Images (stucco)

6 water

Figure 11: Normalized filters in the first layer



(a) layer 1 synthesis (b) layer 2 synthesis

(c) layer 3 synthesis (d) original image

Figure 12: Synthesized Images (water)

7 conclusion

After trying different parameters, such as one-layer-extra-large-filters, two-layer-default-setting, three-layer-7 $\times$ 7-filters, five-layer-3 $\times$ 3-filters, and different strides and number of filters. We find out the parameters used in [1] give the best synthesized results. The first layer has 100 15 $\times$ 15 filters with stride = 3 and padding size = 2. The second layer has 64 5 $\times$ 5 filters with stride = 1 and padding size = 2. The third layer has 30 3 $\times$ 3 filters

with stride = 1 and padding size = 2. Also the Langevin dynamics runs 20 steps for each sample. The normalized filters and the results are shown in Figure 1 - 12.

We can see that this model can fit the original image using the potential function of negative summation of total filtered response in the top layer, which will encourage the model to generate meaningful patterns. The synthesized images cannot be told apart from the original image if one is looking at the images from far away. However, the synthesized images don't have vivid details. They are like a little bit blurry especially when it comes to stucco, most of features of which are detailed textures rather than high level patterns. This may be because the first layer filters are too large to fit the local details or maybe we also need to put constraints to the filtered results like in project 3. Actually code in project 3 gives better result for stucco textures and details as shown in figure 13.

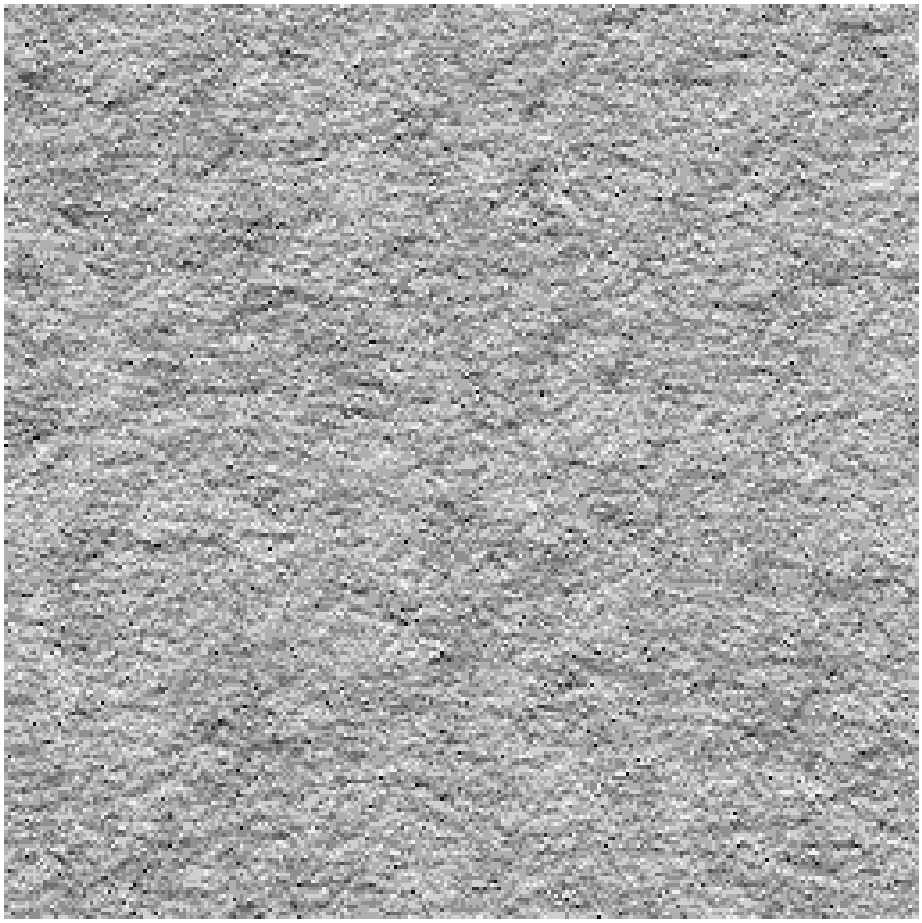


Figure 13: Stucco synthesized using code in project 3

References

[1] Jianwen Xie, Yang Lu, Song-Chun Zhu, and Ying Nian Wu. A theory of generative convnet. *arXiv preprint arXiv:1602.03264*, 2016.