

1. Figure 4 in this paper shows the evolution of the predicted label map. The initial label map is very noisy, but autocontext helps the final label map recover from some errors. Which kind of errors is autocontext able to recover from, and which can it not recover from?
2. For the t -th step, the algorithm trains a model that takes the image and the previous label map as input and produces the next label map. What convolutional network architecture might you use for this model?
3. The fundamental assumption behind autocontext (and other work on structured prediction) is that the label of a given pixel depends on the labels of its neighbors. However, some people argue that the label of a pixel is *conditionally independent* of all other pixels *given the image*. Why or why not?
4. How is the sequence of classifiers proposed by autocontext similar to a convolutional network? How is it different?