During the project, I became amazed by the possibilities that neural networks have to offer. While the first results of colorized images were reported a short period before we started the project, it surprised me that in the end we were able to generate the results that we have now. Working with Theano and Lasagne made it easy to experiment with different set-ups for the network. This makes it incredibly fun to work with. What has also stuck with me was the great difference between different network architectures and training methods/loss functions, in the output results. It made it very clear that neural networks, while very potent when managed well, are by no means techniques that work out of the box. After spending time with different options, I can now reason why some choices turned out to work well, but when confronted with a new problem, these choices may well be not the way to go.

Turning a regression problem (2 feature maps output) into a classification problem was for me one of the most interesting approaches. It turned out to work well, while also providing more information on the output color than just a single value as is the case with the the 2 feature map output. This information is a welcome addition for solving the averaging problem. It may also be interesting to look at if this information can be used in picking more uniform colors for objects. For example, in some images the color shifted half-way through a strawberry from red to green (probably because the network thought there were leaves), but that might be solved by looking more at the expected color for an object. A network architecture with concatenation (which intuitively should work very well since a combination of global and local information is available) might be greatly helped with this approach, producing less artifacts and less color shifts inside an object. This approach could also be used for video, where the expected color of an object should not shift too much over consecutive frames.

The five model architectures that were tested each produced different quality of results for different images. This has to do with the way the network reconstructs the color image, where one approach works best for a certain type of image, i.e. is more or less resistant to blur in the input image or having multiple objects instead of just one cherry. The VGG16 initialization of the network performed only a little bit better than the normal compact network. I think that for the current data set, containing only fruit, this is to be expected, but when a database like image net is used to train, the gains will be much more significant.

Concluding, I think that further research should include training the network on a much broader array of objects to start with. The averaging problem, and specifically the problem where multiple colors are possible for one object, is in my eyes the most interesting. For video, one of the key applications of the technology, it has to be solved first. All in all the project was a lot of fun to do, and it gave me the experience needed with theory and tools to apply neural networks to other problems.

Joost Meulenbeld
4103548