Making Noisy Videos Clear - with Machine Learning

Have you ever tried filming a video during nighttime? If you did, the video presumably ended up grainy and rather unattractive. We have through machine learning developed a noise filter for video, which immensely improves the quality of low-light videos. Algorithms for noise removal have existed for quite some time. However, it is only in the last few years that machine learning have been experimented with to replace these older methods. Machine learning have so far mostly been applied on single images. Our approach was to take it one step further by noise filtering videos, where more information can be utilized.

Our eyes capture images of the world around us, while our brain process them in order for them to make sense. For a machine, this same process is done through a camera sensor followed by some image processing hardware and software. It is however an incredibly hard task to mimic human eyes and brain. Current knowledge suggests that there are certain segments an image must pass through, in order to appear in the human mind. An attempt to model this in a computer has resulted in *convolutional neural networks*.

The basic building block for any neural network is the neuron. A neuron is rather simple in itself, given some amount of stimulus it may fire, otherwise it does not. A neuron firing means that it passes some information onward. By having layers upon layers of multiple parallel neurons, one can create incredibly complex information networks. This is exactly what neural networks are.

Convolutional neural networks have in recent years been shown to successfully filter noise in single images. In order to function properly, neural networks have to learn. Training is executed by feeding a network millions of pairs of corresponding clean and noisy images. The network learns by attempting to cleanse the noise and comparing the result to the clean images, which are treated as answer keys. A successfully trained network may then be applied as a noise filter for video on a frame by frame basis.

We did things differently. Instead of feeding our networks with single noisy images, we used noisy video sequences. This way, by feeding the network with more relevant information, we hypothesized that denoising results should improve remarkably. We built several convolutional neural networks with differences in structure and configuration.

We compared our best network with an existing noise filter in a modern camera. The network performs better in almost all situations but produces a somewhat fluttery video. In comparison to a frame by frame noise filter, our network produces better results except for quickly moving objects.

The ability to produce neatly looking videos is not only important for your average hobby filmmaker. In surveillance, nighttime cameras are completely dependent on noise filters to function well. To perform object tracking and detection, a precondition is often a clear video. These applications are significant as both of these fields have gained much focus in recent years.

There is a considerable amount of work left to be done in order to apply neural networks as noise filters in live settings. Research on neural networks have opened up exciting alternatives to classical noise filtering methods, and the future certainly seem to be in favor of artificial intelligence.

By Tor Nilsson and Tim Borglund

Title of our master's thesis: *Spatio-Temporal Noise Filtering using Convolutional Neural Networks with a Realistic Noise Model under Low-Light Conditions*