Abstract:

Current vision systems are trained on huge datasets, and these datasets come with costs: curation is expensive, they inherit human biases, and there are concerns over privacy and usage rights. To counter these costs, interest has surged in learning from cheaper data sources, such as unlabeled images. In this talk I will consider if we can go a step further and do away with real images entirely, instead learning from synthetic images sampled from procedural image programs. I will present our work on several ways of doing this: 1) learning from simple statistical models of images such as dead leaves, and 2) learning from thousands of video programs collected from the "demoscene" community of algorithmic art. I will talk about the lessons we have learned as to what kinds of synthetic processes make for the effective training data, and will touch on related work in adjacent communities such as NLP, where there is growing evidence that pretraining on random processes can lead to strong representations.