Teaching an Artificial Intelligence with an Artificial Intelligence: Image Completion Using Tensorflow

Strauss, Charles (School: Los Alamos High School)

I hypothesized that two completely uninformed artificial intelligences (AI) could learn by competing to restore obstructed images. Like two novice chess players, they might learn just from playing each other (to produce human-quality restoration). The hard part of image completion is the restored portion must “look right” within the surrounding image. A patient human could rank what “looks right” over millions of trials, but I’m lazy. To eliminate the human, a second AI scores the generated images. My AIs play a game, taking on roles of “Art Critic”, and “Forger”.[1] My critic wins by correctly discriminating real from generated images. My forger wins by fooling the critic. The players learn by themselves, starting from nothing and not even using an algorithm aware of what’s in the image. I studied the two player dynamics for both faces and handwritten digits. I wrote my own code (using Tensorflow gradient library). Facial completion required 4 and 5 layer deep convolutional neural networks trained on a GPU. Handwritten digits used a simpler, 2 layer, dense network. Both players went from random to excellent results. I found when one AI became too good, both players degraded. To prevent this, I paused learning in the better player. Surprisingly, increased obstruction allowed more creative results. Paradoxically, adding faulty neurons improved the robustness; I think it forced the neural nets to rely on more than one trick to win. Training my two-AI models required no human intervention, validating an AI can be taught by an AI. References: [1] Generative adversarial nets I Goodfellow, J Pouget-Abadie, M Mirza, B Xu… - Advances in neural information processing systems, 2014

Awards Won: