Research Review
-- Mastering the game of Go with deep neural networks and tree search

Goals：

The goal of this article is to find an efficient way to build an artificial intelligence on the highly complex Go game and to improve the state-of-art performance of the Go playing machine.

Techniques:

This article composes of three deep neural networks, which can be seen as a pipeline when training.

The first one is the supervised learning(SL) policy networks. It takes the human experts' Go playing records as input, using convolutional neural networks as structure, and a softmax output layer to output a vector of probability of next move, to predict the experts' next move. It results in a 55.7% of prediction accuracy.

The second one is the reinforcement learning(RL) policy networks. It has the exact same structure as the SL policy networks, and initialized as the same as SL policy networks. But its objective is reset to the goal of winning game, not only predicting the next expert move. Also, it is training on self-play with previous version of itself (in order not to overfit), not using experts' records. At this stage, the game with another Go playing AI Pachi, results in an 85% of winning rate.

The third one is the reinforcement learning value networks. It is also trained on self-playing with itself, and the objective is to minimize the mean square error between the predicted value of the value network and the corresponding outcome. This value networks will be combined with previous two policy networks by Monte Carlo Tree Search.

Results:

Combining all three networks, the AlphaGo becomes undefeatable. Gathering most of the state-of-art commercial or open source Go player machine, the AlphaGo wins 494 out of 495 games. Even when playing game with four handicap stones, it can perform higher than 75% of winning rate. The most exciting result is the distributed version of AlphaGo defeat Fan Hui, a professional 2 *dan*. This is the first time that an artificial intelligence beat a human professional in a full game of Go without handicap stone.