This is the review of ‘Mastering the game of Go with deep neural networks and tree search’:

Goals:

AlphaGo was a famous artificial intelligence invented by DeepMind team for playing Go game with human professional player. The goal of this paper is to train a deep neural network and implement new tree search algorithm to play Go game which can defeat human top Go players.

Techniques introduced:

In large games, especially Go (b≈250, d≈150) , exhaustive search is infeasible, so AlphaGo uses ‘value networks’ to evaluate board positions and ‘policy networks’ to select moves. The neural networks play Go at the level of state-of-the-art Monte Carlo tree search programs that simulate thousands of random games of self-play. Also a new search algorithm that combines Monte Carlo simulation with value and policy networks was introduced.

Network part:

The first stage of the training pipeline, AlphaGo built on prior work on predicting expert moves in the game of Go using supervised learning.

The second stage of the training pipeline aims at improving the policy network by policy gradient reinforcement learning. it helps to train the AlphaGo improves win rate from its past game experiences. The final stage of the training pipeline focuses on position evaluation.

Search part:

AlphaGo combines the policy and value networks in an MCTS algorithm that selects actions by lookahead search.

results:

The results of is that single machine AlphaGo is many dan ranks stronger than any previous Go program, winning 494 out of 495 games (99.8%) against other Go programs.