The research paper “Mastering the game of Go with deep neural networks and tree search”revolves around its goal of finding the most effective way to play the game Go. Unlike games like chess, checkers, Go is much more complex, and “recursively computing the optimal value function in a search tree containing approximately b^d possible sequences of moves” is not possible. Therefore, the researchers were trying to build a better method for playing Go, through Policy network, Value network, Policy and value networks search.

Policy network contains two stage, Supervised learning, and Reinforcement learning. Supervised learning is trained from 30 million positions from the KGS Go Server to predict human moves, with the accuracy of 57%. More positions trained mean more accuracy but more time required for this method. Reinforcement learning structure is similar to Supervised learning structure. The difference is the Reinforcement learning plays games between the current policy network and a randomly selected previous iteration of the policy network. Reinforcement leaning wins 80% against Superviced learning and 85% against Patchi.

Value network revolves around the method of generating dataset of Reinforcement learning against itself. This method has the same accuracy with Monte-Carlo rollouts (24%), however it uses 15000 times less computation.

Policy and value networks search consists of the combination between policy network and value network. The search goes through four steps, selection, expansion, evaluation, backup. Selection chooses the edge with maximum value and bonus prior probability. Expansion stage expanses the node left and store the probability given by the policy network. Evaluation evaluates nodes by the value network and rollout policy. Backup stage updates all the subtree values.

The result is that Alpha Go wins 99.8% against other Go programs, 77% against Crazy Stone, 86% against Zen and 99% against Patchi. Alpha Go evaluates thousands of times fewer than Deep Blue in its match against Fan Hui. Alpha Go is now believed to reach professional level against human.