

# Mastering the game of Go with deep neural networks and tree search

David Silver, Aja Huang, Chris J. Maddison, et al.

---

Kyungsu Kang

This paper study the deep neural networks to use a game playing agent of Go. It was not conquered easily because there were too much number of cases( $10^{360}$ ) and  $b \approx 250, d \approx 150$  considering Go' rules. So, search is infeasible. Therefore, To conquer Go game needs to reduce search space. AlphaGo reduced depth-search space and breadth-search space to use basic game play algorithm called MCTS(monte carlo tree search) and deep neural networks.

AlphaGo Algorithms to reduce of game tree of depth-search space

1. Predicting expert moves in Go using SL(supervised learning) policy
  - a. CNN(convolutional neural networks) training about 30 million samples
  - b. Maximizing  $P(a|s)$  to train CNN
2. To prevent optimization of only training set, reinforcement learning to self-play
  - a. RL(reinforcement learning) policy evaluate self-play outcomes of current state of the game
  - b. Applying policy gradient reinforcement learning algorithm
3. Simple pattern recognition called as Rollout policy
  - a. Calculating  $P(a|s)$  considering rules

And to reduce depth-search space training CNN to calculate using RL policy. This paper is used in combination with MCTS and deep neural networks(SL, RL, Rollout policy). And AlphaGo deep neural network learning pipeline process classification using human expert positions, SL policy networks, self-play, RL policy network, self-play and create self-play data to predict value network.

This paper shows that each algorithms has not always performed good solutions but we use another combination of algorithms properly, if we will obtain more efficient search space and reduce time to search.