GoGoGo: Improving Deep Neural Network Based Go Playing AI with Residual Networks

Xingyu Liu

Introduction

- Go playing AIs using Traditional Search: GNU Go, Pachi, Fuego, Zen etc.
- Powered by Deep Learning: Zen → Deep Zen Go, darkforest, AlphaGo
- Goal: From by Vanilla CNN to ResNets

Training Methodology and Data

- SL on Policy Network → RL on Policy Network → RL on Value Network

  - Use Ing Chang-ki rule
    Board State + Ko is Game State, No need to remember the number of captured stones
  - From Kifu to Input Feature Maps
    Channels: 1) Space Positions; 2) Black Positions; 3) White Positions; 4) Current Player; 5) Ko Positions

- Dynamic Board State Expansion
  Ko fight performing. Saves disk space. Small Mem
- Two Levels of Batches (Kifus, moves)
  Random Shuffling. Mem usage small and locality.

Network Architecture

- Fig. 1. Ko fight explicitly expansion
- Fig. 2 (a) Policy Network
- Fig. 2 (b) Value Network
- Fig. 3 (c) Residual Module

  - Hyperparameters
    | Value |
    |-------|
    | Base learning rate | 2E-4 |
    | Decay Policy | Exp |
    | Decay Rate | 0.95 |
    | Decay Step (kifu) | 200 |
    | Loss Function | Softmax |

  - Monte Carlo Tree Search
    \[
    a_t = \text{argmax}_a (Q(s_t, a) + u(s_t, a))
    \]

  - Supervised Learning Training Loss

  - Reinforcement Learning of Value Network
  - Network Architecture Exploration
  - Real Match Testing against Human Players

Experiment Result

- Training Accuracy ~ 32%
- Testing Accuracy ~ 26%