

INTISARI

PERBANDINGAN PERFORMA MODEL CNN DAN RESNET PADA ALGORITMA DEEP REINFORCEMENT LEARNING ALPHA ZERO SEBAGAI AGEN PERMAINAN OTHELLO

Oleh

Rangga Rahman Ardiansyah

20/455452/PA/19667

Bidang *reinforcement learning* pada beberapa tahun terakhir berkembang dengan sangat pesat akibat munculnya teknik *deep reinforcement learning*. Salah satu algoritma yang menggunakan teknik tersebut adalah algoritma Alpha Zero. Namun, algoritma Alpha Zero memerlukan sumber daya komputasi tinggi dan waktu *training* yang lama. Selain itu, arsitektur *neural network* yang digunakan pada penelitian algoritma Alpha Zero tidak dijelaskan secara detail sehingga model yang optimal masih belum diketahui secara pasti.

Dalam penelitian ini dilakukan modifikasi pada implementasi algoritma Alpha Zero dan komparasi dua arsitektur. Arsitektur yang digunakan yaitu *convolutional neural network* sederhana dan *residual neural network*. Eksperimen dilakukan dengan melakukan *training* pada kedua model sebanyak 10, 20, dan 30 iterasi. Evaluasi kemudian dilakukan dengan metrik komparasi berupa *loss* model, waktu *training* model, dan performa model pada permainan.

Hasil penelitian menunjukkan bahwa modifikasi pada implementasi dapat mengurangi kebutuhan sumber daya komputasi namun masih belum berhasil mengurangi waktu *training*. Kemudian komparasi kedua model *neural network* menunjukkan nilai *loss* pada model ResNet selalu lebih rendah dari model CNN dengan *loss* total model ResNet terbaik bernilai 0,465 dibanding *loss* total model CNN yang bernilai 0,966. Namun *training* model ResNet memakan waktu jauh lebih lama dengan waktu training terlama mencapai 39,6 jam dibandingkan CNN yang hanya memerlukan maksimum 8,7 jam. Pada pertandingan evaluasi, model CNN yang dihasilkan setelah *training* 10 iterasi lebih unggul dibandingkan model ResNet yang dilatih sebanyak iterasi yang sama dengan tingkat kemenangan mencapai 80%. Namun model ResNet yang dihasilkan training 20 dan 30 iterasi lebih unggul dengan tingkat kemenangan 56% dan 68% melawan model CNN. Hal ini membuktikan bahwa model CNN bekerja lebih baik ketika dilatih pada waktu terbatas sementara model ResNet lebih unggul ketika tidak ada batasan waktu.

Kata kunci: Alpha Zero, Alpha Go Zero, kecerdasan artifisial, pembelajaran mesin, deep learning, deep reinforcement learning.

ABSTRACT

PERFORMANCE COMPARATION OF CNN AND RESNET MODEL IN ALPHA ZERO DEEP REINFORCEMENT LEARNING ALGORITHM AS AGENT IN OTHELLO

Oleh

Rangga Rahman Ardiansyah

20/455452/PA/19667

The development in the field of reinforcement learning in recent years has been very rapid, especially with the emergence of deep reinforcement learning technique. One of the algorithms that use such technique is the Alpha Zero algorithm. However, the Alpha Zero algorithm requires high computational resources and long training time. Moreover, the neural network architecture used in the Alpha Zero research is not explained in detail so the optimal architecture for the algorithm is still not definitely known.

In this study, modification will be done on an implementation of Alpha Zero algorithm and comparison will be done on two neural network architectures. The architectures that will be used are simple convolutional neural network and residual neural network. For the experiment, models were trained for 10, 20, and 30 iterations. The resulting models then evaluated based on total loss of the model, training time, and model performance in playing the game.

Results of the experiments indicate that while the modification managed to reduce computational resource usage, it still failed to reduce the time taken to train the model. Then the comparison of both models shown that the loss value in the ResNet model is always lower than the CNN model with the best ResNet model having total loss value of 0,465 compared to the best CNN model total loss value of 0,966. But, ResNet model took significantly longer time to train than CNN model with the longest training time of 39,6 hours compared to CNN model that only takes a maximum of 8,7 hours. In evaluation matches, CNN model that is trained for 10 iterations excels more than ResNet model that is trained for the same number of iteration with a win rate of 80%. But, ResNet models that are trained for 20 and 30 iterations excels more with a win rate of 56% and 68% against the CNN models. This prove that CNN model works better when trained in a limited time, while ResNet model works better when there are no time limit.

Keywords: Alpha Zero, Alpha Go Zero, artificial intelligent, machine learning, deep learning, deep reinforcement learning.