**Knightingale: A Block-Recurrent Vision Transformer Chess Engine**

**Braden Stonehill**

[bstoneh1@students.kennesaw.edu](mailto:bstoneh1@students.kennesaw.edu)

https://github.com/bkstonehill/Knightingale

Department of Computer Science, Kennesaw State University

Marietta, GA

**Abstract**

There have been many developments for utilizing neural networks to play and evaluate chess positions such as Stockfish, Leela Zero, and many other engines that utilize various neural network architectures including CNNs and text Transformers. As Transformers have been increasingly used for vision tasks, I propose a hybrid architecture of a Block-Recurrent Transformer and a Vision Transformer for evaluating chess positions in an engine called Knightingale that trains purely through recurrence rather than the common approach of reinforcement learning. While unable to beat the current state-of-the-art engine, Stockfish, its potential to learn the game of chess can be drastically improved with sufficient computational resources and training data.

1. **Introduction**

Chess is a complex game that has undergone much research in the realm of computer science due to its very high value of unique reachable positions, the approaches in tree searching algorithms to explore the game tree, and in terms of how to evaluate how good a position is for a player. Traditionally, hand-crafted heuristic functions, created based on knowledge of the game, were used to evaluate any given position. A simple heuristic function could be to count the material for each player.

In recent years, many approaches for evaluating a position have been developed with the use of neural networks, the current state-of-the-art being Stockfish.

1. **Related Work**
2. **Methodology**
3. **Experiments**
4. **Conclusions**

**References**

Hutchins, D., Schlag, I., Wu, Y., Dyer, E., & Neyshabur, B. (2022). *Block-Recurrent Transformers.* ArXiv. <https://arxiv.org/abs/1901.02860>

Yang, J., Dong, X., Liu, L., Zhang, C., Shen, J., & Yu, D. (2022). *Recurring the Transformer for Video Action Recognition*. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 14063–14073.

Dai, Z., Yang, Z., Yang, Y., Carbonell, J., Le, Q. V., & Salakhutdinov, R. (2019). *Transformer-XL: Attentive Language Models Beyond a Fixed-Length Context.* ArXiv. <https://arxiv.org/abs/1901.02860>

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2021). *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale*. International Conference on Learning Representations. <https://openreview.net/forum?id=YicbFdNTTy>

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention Is All You Need*. ArXiv. <https://arxiv.org/abs/1706.03762>

Romstad, T., Costalba, M., Kiiski, J., Linscott, G., Nasu, Y., Isozaki, M., Noda, H., Aditya, Petrescu, A., Jose, A., Savard, A., Feh, A., Kure, A., Pagel, A., Menezes, A., AlZhrani, A., Vetrov, A., Grant, A., Neporada, A., … Negri, V. *Stockfish*. <https://github.com/official-stockfish/Stockfish>