StratAI - Foundation Model-based Strategy Advisor Agent

Andrew Jeon
University of Washington ECE
Paul Allen Center, 185 E Stevens Way NE, Seattle, WA 98195
ajjeon2@uw.edu

Abstract

I introduce a novel tactical advisor agent that uses foundation models to analyze StarCraft minimap images to provide real-time strategic insights. Unlike other Reinforcement learning approaches which focus more on solving the game, I focus more on the general problem of strategic, spatial-semantic understanding. I hope to fine-tune a ViT or foundation model on a dataset of annotated minimaps that I will annotate. I then plan to evaluate the advisor agent on real match data as well as data from some of my own games. For simplicity I will likely have recommended action classes such as "retreat, expand, defend, all-in" paired with "red advantage, blue advantage, yellow advantage and so on."

1. Introduction

Real-time strategy (RTS) games such as StarCraft were one of Artificial Intelligence's first significant challenges after the "Deep Learning Explosion of the last decade. These multi-agent, partially observable, fast-paced, complex environments represent unique challenges for agents to navigate. Research in this field had focused on achieving superhuman performance through reinforcement learning and imitation learning from expert replays. A significant milestone in this field is AlphaStar [2], which achieves grandmaster-level play by combining supervised learning and a coliseum of reinforcement learning agents competing against and learning from each other. CICERO [3], achieved human-level strategic negotiation in the board game Diplomacy, being one of the first works to integrate strategic reasoning with semantic understanding.

Recent advances in foundation models have sparked a new wave of interest in applying vision-language models to these games. The work *Large Language Models Play StarCraft II* [1] showed that pre-trained foundation models, guided with prompts and context, can generate competent strategic summaries and suggest tactical actions, showing the potential foundation models have as strategic advisor agents.

Inspired by all of these works, I propose a lightweight tactical advisor that operates directly on StarCraft minimap images. Instead of requiring full game-state access or structured inputs, my system applies VLMs to extract spatial-semantic patterns from raw minimaps to generate actionable strategic recommendations. These include identifying who has more bases/territory, identifying high density of units which could represent attacks, and suggesting tactical actions such as "retreat, expand, defend, all-in".

My approach bridges the gap between visual perception and high-level strategy in RTS games. This works towards fast-deployable, lightweight advisors that can assist humans or augment autonomous agents with real-time battlefield awareness.

1.1. Subsection Placeholder

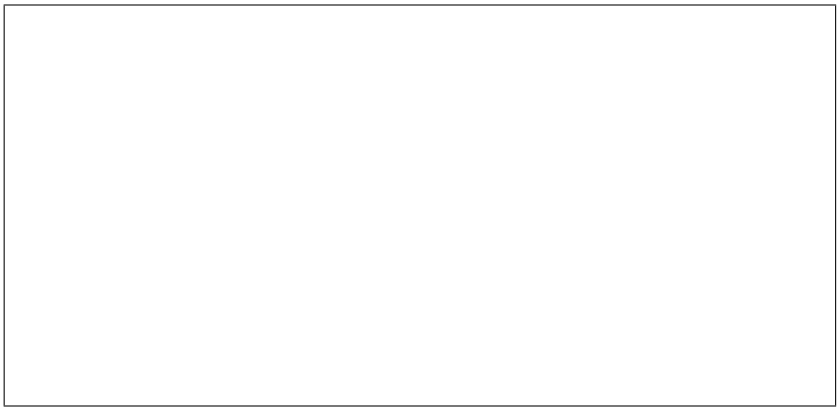
Placeholder

References

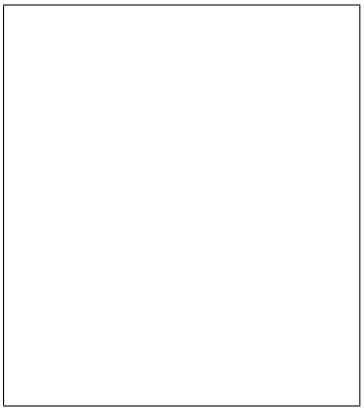
[1] Weiyu Ma, Qirui Mi, Yongcheng Zeng, Xue Yan, Yuqiao Wu, Runji Lin, Haifeng Zhang, and Jun Wang. Large language models play starcraft ii:benchmarks and a chain of summarization approach. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang, editors, *Advances in Neural Information Processing Systems*, volume 37, pages 133386–133442. Curran Associates, Inc., 2024. 1

[2] Michaël Mathieu, Sherjil Ozair, Srivatsan Srinivasan, Caglar Gulcehre, Shangtong Zhang, Ray Jiang, Tom Le Paine, Richard Powell, Konrad Żołna, Julian Schrittwieser, David Choi, Petko Georgiev, Daniel Toyama, Aja Huang, Roman Ring, Igor Babuschkin, Timo Ewalds, Mahyar Bordbar, Sarah Henderson, Sergio Gómez Colmenarejo, Aäron van den Oord, Wojciech Marian Czarnecki, Nando de Freitas, and Oriol Vinyals. Alphastar unplugged: Large-scale offline reinforcement learning, 2023. 1

[3] Meta Fundamental AI Research Diplomacy Team (FAIR), Anton Bakhtin, Noam Brown, Emily Dinan, Adam Fisch, Naman Goyal, Mike Lewis, Julian Michael, Tim Rocktäschel, Da Ju, et al. Human-level play in the game of diplomacy by combining language models with strategic reasoning. *Science*, 378(6624):1067–1074, 2022. 1



(a) An example of a subfigure.



(b) Another example of a subfigure.

Figure 1. Example of a short caption, which should be centered.