Autonomous Agent for Beyond Visual Range Air Combat: A Deep Reinforcement Learning Approach

Joao P. A. Dantas, Marcos R. O. A. Maximo, Takashi Yoneyama

Institute for Advanced Studies | Aeronautics Institute of Technology

Introduction

- **Air Combat Types:** Within Visual Range (WVR) and Beyond Visual Range (BVR).
- **Beyond Visual Range Air Combat:** allows pilots to engage adversaries without direct visual contact (≈ 40 nautical miles).
- **Technological Advancements:** The face of modern warfare is changing rapidly due to the emergence of advanced sensors and weapons that provide superior battlefield awareness and unique capabilities.
- **Artificial Intelligence in Air Combat:** The advent of the AI era has led to the increasing involvement of Unmanned Combat Aerial Vehicles (UCAVs), adding a new dimension to air combat scenarios.

Contributions

- **Innovative AI Design:** We propose a novel Deep Reinforcement Learning (DRL) AI fighter specifically tailored for BVR air combat scenarios.
- **Continuous Improvement:** By incorporating an operational metric-based learning approach, our AI fighter represents a high-performance aircraft that evolves and enhances its capabilities over time.
- **New Air Combat Tactics:** The model enables the development of innovative air combat tactics through rigorous self-play experiments.
- **Shared Airspace:** Interactions between real pilots and trained AI agents within a high-fidelity virtual simulation environment.

Related Work

- Several recent studies have utilized DRL algorithms for autonomous decision-making in BVR air combat for various applications:
  - **Generation of air combat tactics** (Piao et al., 2020).
  - **Maneuver planning** (Zhang et al., 2022; Fan et al., 2022).
  - **Multi-UAV cooperative decision-making methods** (Liu et al., 2022; Hu et al., 2022).
- These works demonstrate the potential of DRL-based approaches in BVR air combat scenarios.
- Despite promising results, there is a need to develop more robust algorithms and assess the feasibility of these methods in real-world applications.
- Our work differs by applying DRL to BVR air combat using a **high-fidelity simulation environment**.
- To the best of our knowledge, no study has combined DRL with self-play techniques to train a high-performance agent that can **interact with real pilots in the same simulation environment**.

"This is the first work to propose a self-learning autonomous agent capable of mastering BVR air combat procedures and interacting with real fighter pilots in a high-fidelity simulation environment."

Aerospace Simulation Environment

- Aerospace Simulation Environment – *Ambiente de Simulação Aeroespacial in Portuguese* (Dantas et al., 2022).
- **Custom-made** in **C++** for advanced programming flexibility.
- **High-fidelity** representation for accurate scenario reproduction.
- Supported by the **Brazilian Air Force**.
- Dedicated to **modeling and simulation** of **military operational scenarios**.

Proposed Model

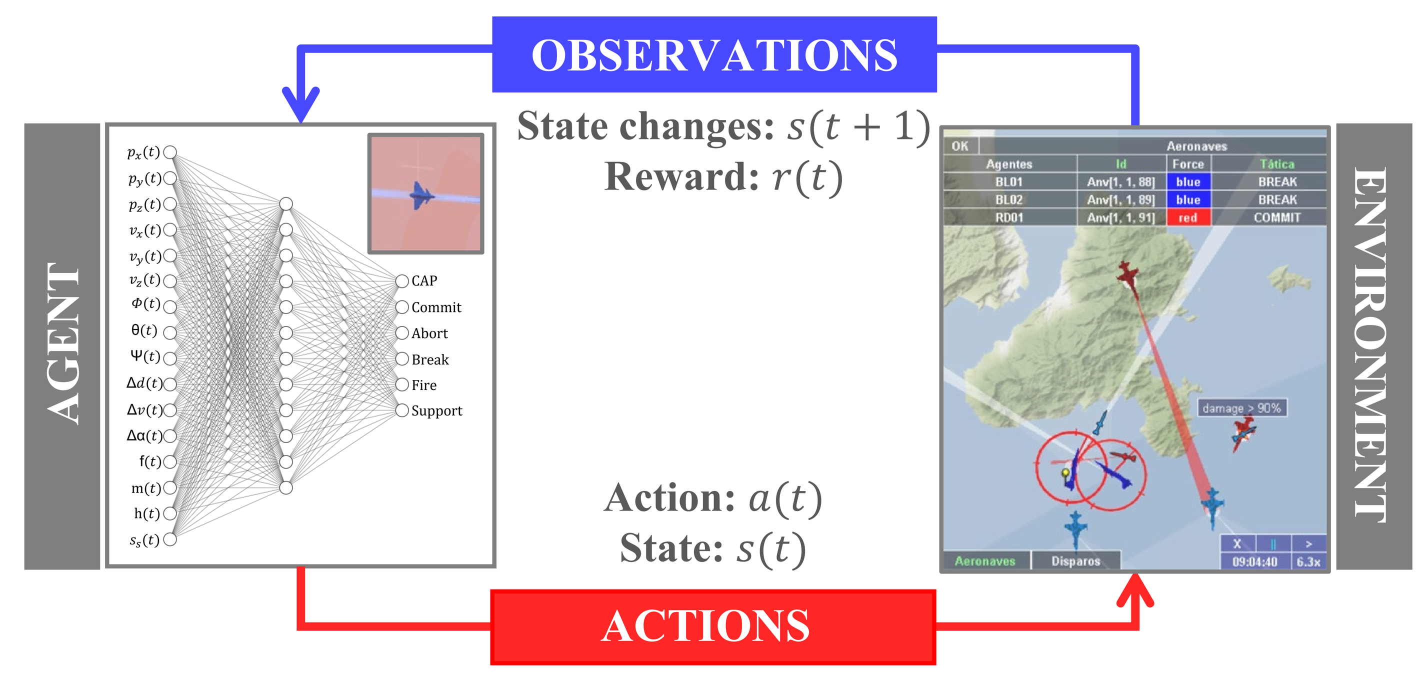

Figure 1. Agent-environment interaction: $a(t)$, $s(t)$, and $r(t)$ denote action, state, and reward at time step t, with $s(t+1)$ given by the environment for the next iteraction.

- Action $a(t)$
  - CAP
    - Keeping an orbit to monitor a specific area for enemy aircraft (Combat Air Patrol).
  - COMMIT
    - Switching to an offensive posture to threaten a key target.
  - ABORT
    - Shifting from offense to defense due to threats, loss of awareness, or task completion.
  - BREAK
    - Employing a final defense when threats exceed safety parameters.
  - FIRE
    - Meeting conditions to launch weapons against the enemy.
  - SUPPORT
    - Using radar to support the missile to the target, improving the success chances.
- State $s(t)$
  - Independent motion variables
    - Position $[p_x(t), p_y(t), p_z(t)]$.
    - Velocity $[v_x(t), v_y(t), v_z(t)]$.
    - Orientation [roll $\phi(t)$, pitch $\theta(t)$, yaw $\psi(t)$].
  - Comparative factors agent-target
    - Relative distance $\Delta d(t)$.
    - Relative speed $\Delta v(t)$.
    - Relative angle $\Delta\alpha(t)$.
  - Agent's conditions
    - Remaining fuel $f(t)$.
    - Remaining missiles $m(t)$.
    - Health condition $h(t)$.
    - Sensors' status $ss(t)$.
- Reward $r(t)$

$$d_i = \frac{(y_{99\%,i} - y_{1\%,i})}{(x_{99\%,i} - x_{1\%,i})} \cdot [D(i, CAP) - x_{1\%,i}] + y_{1\%,i}$$

$$I_{DCA} = w_1 \cdot \frac{m_{avail}}{m_{total}} + w_2 \cdot \frac{1}{1+\exp(-d_r)} + w_3 \cdot \frac{1}{N}\sum_{n=1}^{N} \frac{1}{1+\exp(-d_{e_n})}$$

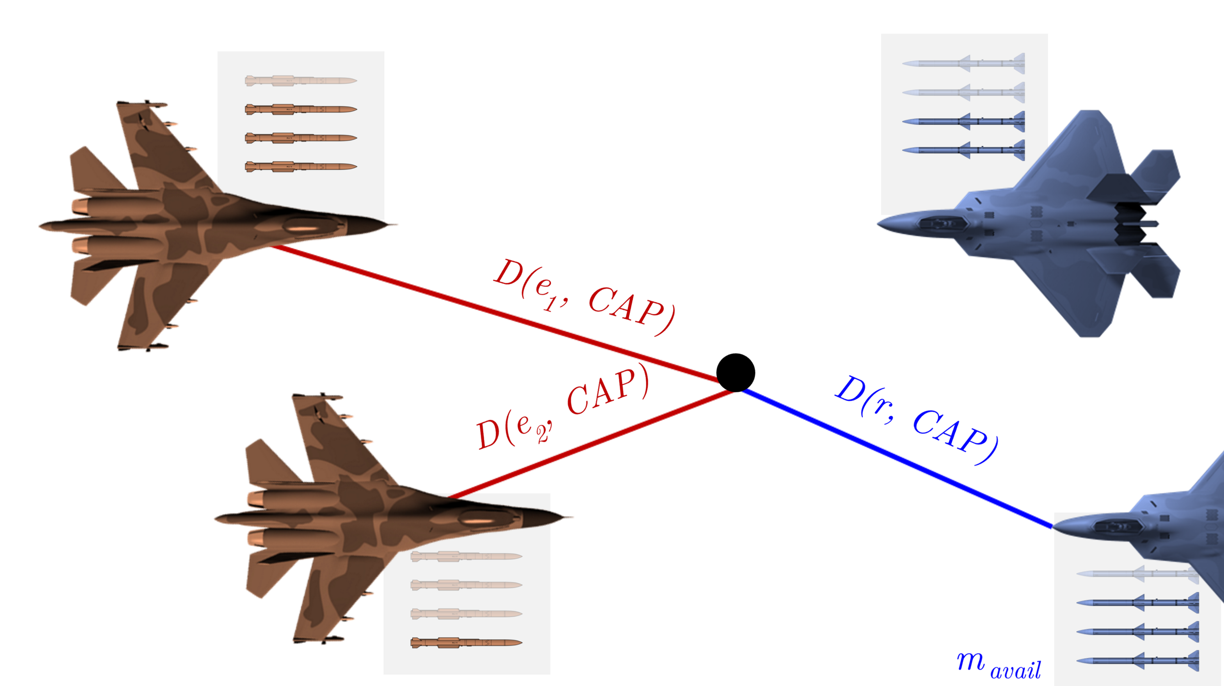

Figure 2. The factors of the Defensive Counter Air Index (I_{DCA}). Source: Dantas et al. (2021).

Cutting-Edge Algorithm

- Rainbow Algorithm
  - Provides state-of-the-art performance (Hessel et al., 2018).
  - Combines several improvements to the Deep Q-Network (DQN) algorithm (Mnih et al., 2015) for superior results.
- Subject Matter Expert Knowledge
  - Enhances both the exploration and exploitation steps.
  - Assists in shaping reward functions tailored for each application.

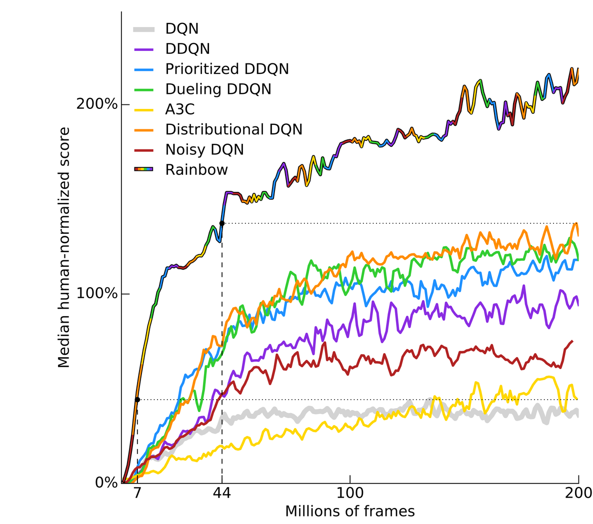

Figure 3. Median human-normalized performance across 57 Atari games, comparing the performance of the Rainbow algorithm (rainbow-colored) to DQN (grey) and six published baselines. Source: Hessel et al. (2018)

Conclusion

- **Training Enhancement:** We aim to enhance the quality of air combat training by developing UCAVs.
- **Advancing AI:** Our efforts are dedicated to advancing AI fighters that can support pilots as Wingmen and potentially even replace them in complex combat scenarios.
- **Unified Framework:** We aspire to establish a Simulation-as-a-Service (SimaaS) platform to meet the diverse simulation demands of the aerospace and defense sectors.
- Future Work
  - Execution of Turing Tests in collaboration with Brazilian fighter pilots to assess the competencies of the AI fighter.
  - Incorporate the feedback gathered from the human pilots to iteratively enhance the AI fighter's operational performance and adaptability in real-world scenarios.

References

Joao P. A. Dantas, Andre N. Costa, Diego Geraldo, Marcos R. O. A. Maximo, and Takashi Yoneyama. 2021. Engagement Decision Support for Beyond Visual Range Air Combat. In *Proceedings of the 2021 Latin American Robotics Symposium*. October 11th–15th, 96–101.

Joao P. A Dantas, Andre N. Costa, Vitor C. F. Gomes, Andre R. Kuroswiski, Felipe L. L. Medeiros, and Diego Geraldo. 2022. ASA: A Simulation Environment for Evaluating Military Operational Scenarios. In *Proceedings of the 20th International Conference on Scientific Computing*. 25th–28th, Las Vegas, NV, USA.

Zihao Fan, Yang Xu, Yuhang Kang, and Delin Luo. 2022. Air Combat Maneuver Decision Method Based on A3C Deep Reinforcement Learning. *Machines* 10, 11 (2022), 1033.

Matteo Hessel, Joseph Modayil, Hado van Hasselt, Tom Schaul, Georg Ostrovski, Will Dabney, Dan Horgan, Bilal Piot, Mohammad Gheshlaghi Azar, and David Silver. 2018. Rainbow: Combining Improvements in Deep Reinforcement Learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 32. 3150–3157.

Jinwen Hu, Luhe Wang, Tianmi Hu, Chubing Guo, and Yanxiong Wang. 2022. Autonomous maneuver decision making of dual-UAV cooperative air combat based on deep reinforcement learning. *Electronics* 11, 3 (2022), 467.

Xiaoxiong Liu, Yi Yin, Yuzhan Su, and Ruichen Ming. 2022. A Multi-UCAV Cooperative Decision-Making Method Based on an MAPPO Algorithm for Beyond-Visual-Range Air Combat. *Aerospace* 9, 10 (2022), 563.

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. 2015. Human-level control through deep reinforcement learning. *Nature* 518, 7540 (2015), 529–533.

Haiyin Piao, Zhixiao Sun, Guanglei Meng, Hechang Chen, Bohao Qu, Kuijun Lang, Yang Sun, Shengqi Yang, and Xuanqi Peng. 2020. Beyond-visual-range air combat tactics auto-generation by reinforcement learning. In *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 1–8.

Hongpeng Zhang, Yujie Wei, Huan Zhou, and Changqiang Huang. 2022. Maneuver Decision-Making for Autonomous Air Combat Based on FRE-PPO. *Applied Sciences* 12, 20 (2022), 10230.

dantasjpad@fab.mil.br | www.joaopadantas.com