

Shengyi “Costa” Huang

(864)501-6630
COSTA.HUANG@OUTLOOK.COM
GITHUB.COM/VWXYZJN
LINKEDIN.COM/IN/COSTA-HUANG
WWW.COSTA.SH

SUMMARY

I am a machine learning engineer at Hugging Face, with a doctorate degree from Drexel University specializing in deep reinforcement learning (RL) and creating efficient algorithms and reproducible research. My advisor is [Santiago Ontañón](#). I am also the creator of [CleanRL](#), a top 20 most-starred deep RL library on GitHub.

EDUCATION

Ph.D. in Computer Science **Sep 2023**
M.S. in Computer Science **Dec 2021**
Drexel University, Philadelphia, PA

B.S in Computer Science, B.S in Mathematics **May 2018**
Furman University, Greenville, SC

EXPERIENCE

Machine Learning Engineer **Jun 2023 – Present**
Allen Institute for Artificial Intelligence, Philadelphia, PA

Machine Learning Engineer **Jun 2023 – Present**
Hugging Face, Philadelphia, PA

- Drive research and engineering deep dive in reinforcement learning from human feedback (RLHF), documenting and validating over 21 useful RLHF implementation details.
- Improve the training stability in multi-GPU / multi-node settings for Hugging Face’s highly popular open-source RLHF library TRL (6.6 GitHub stars, used by 700+ open repositories).
- Lead the research on alignment techniques such as Anthropic’s Constitutional AI and support the development of Hugging Face’s Zephyr 7B model.

Data Scientist Intern **Jun 2022 – Sep 2022**
Riot Games, Los Angeles, CA

- Conducted research at the AI Accelerator team, with a focus on applying cutting-edge multi-agent deep reinforcement learning algorithms to the next generation of games.
- Led the technical discussions on algorithm design and iterated hundreds of research experiments to create playful and intelligent game-playing agents.
- Implemented an automated pipeline that saved 80% human labor spent on the evaluation process.

Deep Learning Algorithm Engineering Intern **Mar 2022 – Jun 2022**
NVIDIA, Santa Clara, CA

- Built better utilities in NVIDIA’s Isaac Gym robotic simulation environment, which helps researchers to develop, test, and manage AI-based robots.
- Analyzed and benchmarked the performance of state-of-the-art deep RL frameworks such as Stable-baselines 3, rl_games, RLLib, and CleanRL, creating solid baselines for the research team.
- Implemented a more efficient multi-GPU training paradigm for rl_games that increased up to 20% more system throughput for training agents.

Deep Learning Growth Engineer Intern **Jun 2021 – Sep 2021**
Weights and Biases, Palo Alto, CA

- Researched 37 implementation details of Proximal Policy Optimization. Published at ICLR 2022.
- Contributed the W&B visualization integration to popular RL projects, such as the MineRL competition (16k views) and Stable-baselines 3 library (used by 400+ packages).
- Created multiple ML educational content, including a blog post on AWS SageMaker and a webinar on experiment tracking and analysis workflow (200 + views on YouTube).

Machine Learning Engineer Intern

Curai Health, Palo Alto, CA

Apr 2021 – Jun 2021

- Analyzed 4 years of experiment management needs in an AI-first healthcare startup.
- Implemented an experiment management pipeline that was adopted by the team's workflow, which covered dataset versioning, experiment orchestration, tracking, analysis, and tuning.
- Led the design of a prototype project to generate medical conversation by leveraging DialoGPT.

Graduate Research / Teaching Assistant

Drexel University, Philadelphia, PA

Sep 2018 – Jun 2023

PROJECTS

CleanRL (github.com/vwxyzjn/cleanrl, 3.2k stars on GitHub)

High-quality single file implementation of Deep Reinforcement

Learning algorithms with research-friendly features

- Python PyTorch OpenAI Gym Tensorboard Docker AWS
Weights and Biases Deep Q-learning Policy Gradient Visualization

Gym-MicroRTS (github.com/vwxyzjn/gym-microrts, 100 stars on GitHub)

The OpenAI Gym wrapper of MicroRTS for deep RL research

- Python OpenAI Gym Policy Gradient Real-time Strategy Games
Docker AWS Learning through Self-play CI/CD Numpy

Portwarden (github.com/vwxyzjn/portwarden, 372 stars on GitHub)

Create Encrypted Backups of Your Bitwarden Vault with Attachments

- Go Docker Kubernetes AES Encryption

PUBLICATIONS

Huang, S., Noukhovitch, M., Hosseini, A., Rasul, K., Wang, W., and Tunstall, L. The N+ Implementation Details of RLHF with PPO: A Case Study on TL; DR Summarization. COLM 2024.

Tunstall, L., Beeching, E., Lambert, N., Rajani, N., Rasul, K., Belkada, **Huang, S.**, Werra, L., Fourrier, C., Habib, N., Sarrazin, N., Sanseviero, O., Rush, A., and Wolf, T. Zephyr: Direct distillation of lm alignment. COLM 2024.

Huang, S., Weng, J., Charakorn, R., Lin, M., Xu, Z., and Ontañón, S. Cleanba: A Reproducible and Efficient Distributed Reinforcement Learning Platform. ICLR 2023.

Sullivan, R., Kumar, A., **Huang, S.**, Dickerson P J., Suarez J., Reward Scale Robustness for Proximal Policy Optimization via DreamerV3 Tricks, NeurIPS 2023.

Weng, J., Lin, M., **Huang, S.**, Liu, B., Makoviichuk, D., Makoviychuk, V., Liu, Z., Song, Y., Luo, T., Jiang, Y. and Xu, Z., EnvPool: A Highly Parallel Reinforcement Learning Environment Execution Engine, NeurIPS 2022.

Huang, S., Kanervisto, A., Raffin, A., Wang, W., Ontan'on, S., & Dossa, R.F. A2C is a special case of PPO. *preprint, 2022*

Huang, S., Dossa, R., Raffin, A., Kanervisto, A., Wang, W. The 37 Implementation Details of Proximal Policy Optimization. *ICLR Blog Post Track, 2022*

Huang, S., Dossa, R., Ye, C., Braga, J., CleanRL: High-quality Single-file Implementations of Deep Reinforcement Learning Algorithms, *Journal of Machine Learning Research, 2022*

Huang, S., Ontañón, S., "A Closer Look at Invalid Action Masking in Policy Gradient Algorithms", *FLAIRS-35, 2022*

Compton, R., Valmianski, I., Deng, L., **Huang, C.**, Katariya, N., Amatriain, X., Kannan, A. MED-COD: A Medically-Accurate, Emotive, Diverse, and Controllable Dialog System. *Machine Learning for Health, 2021*.

Dossa, R., **Huang, S.**, Ontañón, S., Matsubara, T., “An Empirical Investigation of Early Stopping Optimizations in Proximal Policy Optimization”, *IEEE Access, 2021*

Huang, S., Ontañón, S., Bamford, C., Grela, L., “Gym- μ RTS: Toward Affordable Full Game Real-time Strategy Games Research with Deep Reinforcement Learning”, *IEEE Conference on Games 2021*

Huang, S., Ontañón, S., “Measuring Generalization of Deep Reinforcement Learning Applied to Real-time Strategy Games”, *AAAI 2021 Reinforcement Learning in Games Workshop*

Bamford, C., **Huang, S.**, Lucas, S., “Griddly: A platform for AI research in games.”, *AAAI 2021 Reinforcement Learning in Games Workshop*

Huang, S., Ontañón, S., “Action Guidance: Getting the Best of Training Agents with Sparse Rewards and Shaped Rewards”, *AIIDE 2020 Strategy Games Workshop*

Huang, S., Ontañón, S., “Comparing Observation and Action Representations for Reinforcement Learning in μ RTS”, *AAIIDE 2019 Strategy Games Workshop*

Huang, S., Healy, C., “StreetTraffic: a Library for Traffic Flow Data Collection and Analysis”, poster presentation in *ACMSE 2018 Conference*

SERVICES

I was a reviewer for the following journals, conferences and workshops: NeurIPS (2022), JMLR (2022), JOSS (2022), ICGA(2021), AIIDE-RTS (2020, 2021),

TALKS

12/2023, NeurIPS Neural MMO Workshop. Topics: CleanRL, Cleanba.

9/2024, InstaDeep. Topics: CleanRL, Open RL Benchmark.

9/2024, Cohere for AI. Topics: Cleanba.

1/2022, Riot Games, Topics: RL, Gym- μ RTS