

# LUCAS CACCIA

lucas.page-caccia@mail.mcgill.ca | | webpage | github | google scholar

## EDUCATION

---

- PhD Computer Science** 2018 - 2023  
McGill University, Mila  
Advisor : Joelle Pineau  
Thesis: *Preventing Forgetting and Promoting Transfer in Continual Learning*
- BS Mathematics and Computer Science** 2014 - 2017  
McGill University  
GPA : 3.96/4.0

## WORK EXPERIENCE

---

- Microsoft Research** May 2023 - present  
*Post-Doctoral Researcher*
- Building scalable and modular systems for massively multitask adaptation. Applied to both language and behavioral cloning of individual chess players.
- Microsoft Research** March 2022 - May 2023  
*Student Researcher*
- Working on modular approaches for efficient forward transfer in Natural Language Processing
- Facebook AI Research** March 2020 - December 2021  
*Visiting Researcher*
- Developed new continual learning algorithms designed for realistic settings.
- McGill University** May 2017 - August 2017  
*Undergraduate Researcher*
- Worked on a differentiable simulator for self-driving cars. This includes generating both RGB images and corresponding LiDAR point clouds.
- Microsoft, Seattle** May 2016 - July 2016  
*Software Developer Intern*
- Worked with the Customer Relationship Management (CRM) team.
  - Added a new feature to the CRM interface called Approval Flow (C#, SQL)

## TEACHING EXPERIENCE

---

- African Institute for Mathematical Sciences (AIMS)** Jan 2019  
*Teaching Assistant*
- This was for a two week long Reinforcement Learning class held in Kigali, Rwanda. We prepared several tutorials covering the basics of RL (which can be found [here](#)).
- McGill University** Winter 2018 and Fall 2021
- COMP 551 - Applied Machine Learning Class. I gave a few tutorials on the basics of automatic differentiation in Pytorch.

## AWARDS & SCHOLARSHIPS

---

- |  |                |
|--|----------------|
| <b>Borealis AI Fellowship</b>  | 2020-2021      |
| · Selected alongside 10 fellows for research and academic achievements |                |
| <b>Dean’s Honour List</b>  | 2015-2016-2017 |
| · Awarded to the top 10% performing students                           |                |
| <b>CIBPA Foundation Bursary - Merit</b>                                | 2016-2019      |
| · Awarded to promising Canadian students of Italian descent            |                |

## PAPER IMPLEMENTATIONS

---

1. “*PixelCNN++: Improving the PixelCNN with discretized logistic mixture likelihood and other modifications*” **321 stars** [\[link\]](#)
2. “*Glow: Generative Flow with Invertible 1x1 Convolutions*” **89 stars** [\[link\]](#)
3. “*Improving Variational Inference with Inverse Autoregressive Flow*” **69 stars** [\[link\]](#)

## ORGANIZATION

---

- |   |             |
|---|-------------|
| <b>Continual Learning Reading Group Organizer</b>                                 | 2022 - 2023 |
| · Weekly meetings held at MILA attended by the local Continual Learning Community |             |

## REVIEWING

---

- |   |      |
|---|------|
| <b>ICLR</b>                                   | 2024 |
| <b>ICLR, ICML, CoLLAs, NeurIPS</b>            | 2023 |
| <b>ICML, ICLR, CoLLAs</b>                     | 2022 |
| <b>NeurIPS, ICLR, ICML CL Theory Workshop</b> | 2021 |
| <b>NeurIPS, MAIS</b>                          | 2020 |
| <b>ICRA</b>                                   | 2019 |

## SELECTED PUBLICATIONS

---

- [1] **Lucas Caccia**, Edoardo Ponti, Zhan Su, Matheus Pereira, Nicolas Le Roux, and Alessandro Sordani. Multi-head adapter routing for cross-task generalization. In *Advances in Neural Information Processing Systems*, 2023. [\[link\]](#).
- [2] Gwen Legate, Nicolas Bernier, **Lucas Caccia**, Edouard Oyallon, and Eugene Belilovsky. Guiding the last layer in federated learning with pre-trained models. In *Advances in Neural Information Processing Systems*, 2023. [\[link\]](#).
- [3] Jean-Baptiste Gaya, Thang Doan, **Lucas Caccia**, Laure Soulier, Ludovic Denoyer, and Roberta Raileanu. Building a subspace of policies for scalable continual learning. In *International Conference on Learning Representations*, 2023. [\[link\]](#).

- [4] **Lucas Caccia**, Jing Xu, Myle Ott, Marc’aurelio Ranzato, and Ludovic Denoyer. On anytime learning at macroscale. In *Proceedings of The 1st Conference on Lifelong Learning Agents*. PMLR, 2022. [\[link\]](#).
- [5] **Lucas Caccia** and Joelle Pineau. Special: Self-supervised pretraining for continual learning. In *Continual Semi-Supervised Learning*. Springer International Publishing, 2022. [\[link\]](#).
- [6] **Lucas Caccia**, Rahaf Aljundi, Nader Asadi, Tinne Tuytelaars, Joelle Pineau, and Eugene Belilovsky. New insights on reducing abrupt representation change in online continual learning. In *International Conference on Learning Representations*, 2022. [\[link\]](#).
- [7] **Lucas Caccia**, Eugene Belilovsky, Massimo Caccia, and Joelle Pineau. Online learned continual compression with adaptive quantization modules. In *Proceedings of the 37th International Conference on Machine Learning*, 2020. [\[link\]](#).
- [8] Massimo Caccia\*, **Lucas Caccia**\*, William Fedus, Hugo Larochelle, Joelle Pineau, and Laurent Charlin. Language gans falling short. In *International Conference on Learning Representations*, 2020. [\[link\]](#).
- [9] Massimo Caccia, Pau Rodriguez, Oleksiy Ostapenko, Fabrice Normandin, Min Lin, **Lucas Caccia**, Issam Hadj Laradji, Irina Rish, Alexandre Lacoste, David Vázquez, and Laurent Charlin. In *Advances in Neural Information Processing Systems*, volume 33, 2020. [\[link\]](#).
- [10] Pierre Thodoroff, Nishanth Anand, **Lucas Caccia**, Doina Precup, and Joelle Pineau. Recurrent value functions. *Conference on Reinforcement Learning and Decision Making*, 2019. [\[link\]](#).
- [11] **Lucas Caccia**, Herke van Hoof, Aaron Courville, and Joelle Pineau. Deep generative modeling of lidar data. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019. [\[link\]](#).
- [12] Rahaf Aljundi\*, **Lucas Caccia**\*, Eugene Belilovsky\*, Massimo Caccia\*, Min Lin, Laurent Charlin, and Tinne Tuytelaars. Online continual learning with maximal interfered retrieval. In *Advances in Neural Information Processing Systems*, 2019. [\[link\]](#).

## Preprints

- [1] Gwen Legate, **Lucas Caccia**, Eugene Belilovsky, Reducing Forgetting In Federated Learning with Truncated Cross-Entropy [\[link\]](#)