

# Ricardo Mokhtari

ricardomokhtari@gmail.com | +44 7508 813327 | [GitHub](#) | [LinkedIn](#) | [Google Scholar](#) | [ricardomokhtari.com](#)

## Summary

I am a highly self-motivated **AI Research Scientist at AstraZeneca** with a passion for driving impact in society using AI. At AstraZeneca I have worked on **applying and developing state of the art models to large datasets of millions of images**, as well as large scale **genomic data**. I have **3.5 years of experience in Machine Learning with 2.5 years in Computer Vision**, and I am highly proficient in **Python** and **PyTorch**. Prior to AstraZeneca, I graduated from **Imperial College London with a First-Class Master's Degree in Engineering**, where I worked on generative models and reinforcement learning. I have extensive experience of applying the fundamental architectures in **Computer Vision and Graph Machine Learning** such as **CNNs, GANs, VAEs, UNets and Transformers**.

## Experience

### AI Research Scientist | AstraZeneca

Sept. 2021 – Present

#### Project 1: *Applying self-supervised learning to medical imaging* [[Pre-Print](#)]

- Developed and applied state of the art open-source computer vision methods to 1000s of gigapixel images to inform translational medicine teams within AstraZeneca (work showcased to EVP level and presented research to Global Product Team) – pushed AUC from 0.6 to 0.87
- Proactively communicated research findings to unfamiliar and non-technical audiences
- Collaborated extensively with interdisciplinary teams including AI scientists, image analysts and pathologists, using Agile framework

#### Project 2: *Continual active learning platform for medical imaging*

- Developed robust infrastructure for deploying, continuously monitoring and improving computer vision models for 3 internal stakeholders
- Deployed models on server backend using MONAI, built a GUI using Dash to serve users
- Proactively gathered user requirements, refined solution using Agile methodology

#### Project 3: *Using graph machine learning to discover new cancer biomarkers*

- Leveraged multi-modal genomic dataset to identify novel cancer subtypes and associated biomarkers using graph machine learning and presented to translational teams to inform 2023 oncology R&D strategy
- Explored and demonstrated that graph approaches are powerful for multi-modal cancer datasets - pushed AUC from 0.83 to 0.89

#### *I have proactively championed a data-driven culture at AstraZeneca by:*

- Organising the first AZ Hack (organising team of 3) – a global-scale Data Science hackathon attended by 173 participants across 15 countries
- Writing (from scratch) and delivering a 3-hour computer vision workshop to 50 AZ employees
- Teaching an 8-week Python course to a class of 60 AZ employees

### Data Scientist | Imperial College London + Refinitiv

Nov. 2020 – May 2021

- Industrial Data Science research project with Refinitiv, a large financial services company
- Selected to be part of Imperial's Advanced Data Science Team - developed data-driven methods for autonomous web crawling using reinforcement learning
- Co-developed intelligent web crawling strategy from scratch, co-wrote technical reports and delivered presentations to managing directors at Refinitiv

### Research Assistant | Imperial College London

Oct. 2019 – Jun. 2021

- Explored and evaluated the utility of using generative models as a data augmentation technique for boosting the performance and robustness of computer vision models
- Devised a simple framework for quantitatively evaluating model robustness, showed that generative models are a successful approach
- Developed and applied SOTA models (StyleGAN, Pix2Pix, VAE) - [[Pre-Print](#)]

## Publications

- R Mokhtari et al.**, Interpretable histopathology-based prediction of disease relevant features in Inflammatory Bowel Disease biopsies using weakly-supervised deep learning. [[Accepted for MIDL 2023](#)]
- R Mokhtari et al.**, Predicting disease relevant features in Crohn's Disease and Ulcerative Colitis from Haematoxylin & Eosin stained whole slide images using self-supervised deep learning, *Journal of Crohn's and Colitis* 2023, <https://doi.org/10.1093/ecco-icc/ijac190.0407> (Impact factor >10)
- Attar, R., Hurault, G., Wang, Z., **Mokhtari, R.**, Pan, K., Olabi, B., Earp, E., Steele, L., Williams, H. and Tanaka, R.J., 2022. Reliable detection of eczema areas for fully automated assessment of eczema severity from digital camera images. *medRxiv*.
- Hurault, G., Pan, K., **Mokhtari, R.**, Olabi, B., Earp, E., Steele, L., Williams, H.C. and Tanaka, R.J., 2022. Detecting eczema areas in digital images: an impossible task? *JID Innovations*, p.100133.

References may be supplied on request

## Education

### Imperial College London

#### MEng Molecular Bioengineering

Oct. 2017 – Jun. 2021

- Grade: First Class Hons. (74.93%)**
- Dean's List 2021 – Prize for scoring in top 10% of students**
- Relevant modules:
  - Image Processing (1<sup>st</sup>)
  - Signal Processing (1<sup>st</sup>)
  - Reinforcement Learning (1<sup>st</sup>)
  - Mathematics (1<sup>st</sup>)
  - Modelling in Biology (1<sup>st</sup>)
  - Molecular & cellular biology (1<sup>st</sup>)
  - Synthetic Biology (1<sup>st</sup>)

### Charterhouse School

2012 – 2017

- A-Level:** A\* A\* A A
- GCSEs:** 11 A\*s
- Academic scholarship worth £1500/year

## Technical Skills

### Programming

**Highly proficient:** Python (3.5 years' experience)

**Familiar:** R, C/C++, MATLAB, JavaScript, ReactJS, HTML, CSS

### Machine Learning Tools

**Proficient:** PyTorch, PIL, OpenCV, pandas, numpy, sklearn

**Familiar:** TensorFlow/Keras, bokeh

### Machine Learning Theory

**Computer Vision** (vision transformer, SSL, WSL, CNN, GAN, VAE, UNet)

**Graph ML** (GCN, link prediction, knowledge graphs, graph embedding)

**Classic ML** (logistic regression, SVM, k-means, decision trees, random forests)

### Other Tools

Git, Bash scripting, HPC, LaTeX

### Soft Skills

Excellent presentation/communication skills

Agile working methodology (JIRA, MIRO)