Ricardo Mokhtari

ricardomokhtari@gmail.com | **GitHub** | **LinkedIn** | **Google Scholar** | **ricardomokhtari.com**

Summary

I am a highly self-motivated **AI Research Scientist at AstraZeneca** with a passion for discovering disease insights with AI. At AstraZeneca I have worked on **building AI-powered tools to support and accelerate the drug discovery pipeline** by leveraging state of the art AI models on **large datasets of millions of images**. I have **3 years of experience in developing and applying Computer Vision models to medical imaging**, 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 in applying the fundamental **Computer Vision models** such as **CNNs, GANs, VAEs, UNets and Vision Transformers.**

Experience

AI Research Scientist | AstraZeneca   **Sept. 2021 – Present**

Project 1: *Applying Self-Supervised Learning to H&E images* **[Accepted MIDL '23] [Code]**

- Developed and applied state of the art Computer Vision models (MIL, ViT) to 1000s of Whole Slide Images to speed up the workflow of clinical pathologists and inform translational medicine teams within AstraZeneca (work showcased to EVP level and presented research to Global Product Team) – pushed AUC performance 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 in vivo (3D/4D MRI) imaging*

- Developed infrastructure for deploying, continuously monitoring and improving Computer Vision models in order to speed up the workflow of pre-clinical animal studies
- Deployed models on backend using MONAI, built custom GUIs using Streamlit 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:

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

Research Assistant| Imperial College London [Paper]   **Oct. 2019 – Jun. 2021**

- Explored and evaluated the utility of using SOTA generative models (StyleGAN, Pix2Pix, VAEs) as a data augmentation technique for boosting the performance and robustness of Computer Vision models
- Devised a simple framework for quantitatively evaluating model robustness, used this framework to demonstrate that generative models are a successful approach

Data Scientist | Imperial College London + Refinitiv   **Nov. 2020 – May 2021**

- 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

Publications

- **R Mokhtari et al.,** Interpretable histopathology-based prediction of disease relevant features in Inflammatory Bowel Disease biopsies using weakly-supervised deep learning. **[Accepted at MIDL 2023] [Code]**
- **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-jcc/jjac190.0407 **(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.

Education

Imperial College London
MEng Molecular Bioengineering
Oct. 2017 – Jun. 2021

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

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, Streamlit
Familiar: TensorFlow/Keras, bokeh

Machine Learning Theory
Computer Vision (Vision Transformer, SSL, WSL, MIL, CNN, GAN, VAE, UNet)
Classic ML (logistic regression, SVM, k-means, decision trees, random forests)
Graph ML (GCN, link prediction, knowledge graphs, graph embedding)

Other Tools
Git, Bash scripting, HPC, LaTeX

Soft Skills
Excellent presentation/communication skills
Agile working methodology (JIRA, MIRO)

References may be supplied on request