Javier Fernandez-Marques

jafermarq@gmail.com
www.jafermarq.com

Linacre College
St. Cross Road
Oxford, OX1 3JA

Summary

I am a Researcher at Samsung AI. My research focuses on designing hardware-aware optimizations for deep learning workloads enabling efficient inference with images, audio and graphs. I investigate new ways of achieving this by intersecting of neural architecture search and approximate computting. I am also interested in federated learning. I hold a PhD form University of Oxford. I have two years of industry experience designing real-time computer vision systems for the sport and fashion-retail sectors.

Education

UNIVERSITY OF OXFORD — Oxford, UK
DPhil (PhD) Candidate in Computer Science — *Oct. 2017 – 2021*

- **Theme**: High Performing Deep Learning for Resource Constrained Platforms
- Supervised by Associate Prof. Nicholas D. Lane
- TA in Deep Neural Networks course for MSc students. (MT18, HL19, MT19)

QUEEN MARY, UNIVERSITY OF LONDON — London, UK
MSc in Computer Vision (Distinction) — *Sept. 2014 - Sept. 2015*

- **Thesis**: Pedestrian detection from varying viewing angles (89/100)
- Supervised by Professor Andrea Cavallaro

TECNUN, UNIVERSITY OF NAVARRA — San Sebastián, Spain
BSc in Telecommunications Engineering — *Sept. 2010 - May 2014*

- **Thesis**: Collagen Mesh Detection and Quantification in Reflection Microscopy (100/100)
- Co-supervised by Professor Carlos Ortiz-de-Solórzano and Associate Prof. Arrate Muñoz

Work Experience

SAMSUNG AI — Cambridge, UK
Researcher — *July 2021 - now*

- I design and implement new methods to accelerate Machine Learning workloads. My current research interest lie at the intersection of neural architecture search, approximate computting and federated learning.

ARM — Cambridge, UK
Machine Learning Research Intern — *May 2019 - Sept. 2019*

- Developed a novel approach of using Winograd convolutions, enabling the deployment of INT8 models with marginal accuracy drops, no impact on model size and, up to 2× speedup compared to optimized INT8 *im2row* on Arm Cortex-A73 and A53 mobile CPUs.
- Designed a framework that transforms a given model (a CNN) into its Winograd-aware counterpart. This framework jointly optimizes accuracy and latency via NAS.

HOLITION — London, UK
Computer Vision Engineer — *Oct. 2016 - Oct. 2017*

- Lead computer vision and machine learning engineer developing AR applications for the fashion retail sector. Research and implementation of low-level computer vision techniques suitable for real-time execution on hand-held devices (smartphones, tablets)
- Maintained and enhanced our in-house lightweight face tracking algorithm.
- Worked on full application cycle: design, prototyping, implementation and optimisation.

HAWK-EYE INNOVATIONS (SONY) — Basingstoke, UK

Software Engineer — *Feb. 2016 - Oct. 2016*

- Research and implementation of low-level image processing (demosaicing and colour correction) and multi-camera calibration and player-tracking algorithms. Improved UI and UX of a camera calibration software used in hundred of sport events through the year.

FRAUNHOFER IIS — Erlangen, Germany

Research Intern — *Summer 2014*

- Investigated Literature on stereo vision and implemented a state of the art stereo-vision algorithm to extract the depth-maps from a scene using a multi-camera array.

CIMA — Pamplona, Spain

Research Intern — *Jun. 2013 - Jun. 2014*

- Algorithm implementation to extract and analyse the 3D network architecture of collagen-based matrices in the presence of cancer cells. Developed GUI that engineers and biologists can use to setup their experiments. Publications in ICIP and EMBC.
- Worked in a multidisciplinary team of biologists, biochemists and engineers.

Publications

ICLR 2022 Xinchi Qiu*, **Javier Fernandez-Marques***, Pedro Porto Buarque de Gusmao, Yan Gao, Titouan Parcollet, and Nicholas D. Lane. *ZeroFL: Efficient On-Device Training for Federated Learning with Local Sparsity.* International Conference on Learning Representations (ICLR), 2022. (*equal contrib.)

Pre-print Xinchi Qiu, Titouan Parcollet, **Javier Fernandez-Marques**, Pedro Porto Buarque de Gusmao, Daniel J. Beutel, Taner Topal, Akhil Mathur and Nicholas D. Lane. *A First Look into the Carbon Footprint of Federated Learning.*

FCCM 2021 Stylianos I. Venieris*, **Javier Fernandez-Marques***, Nicholas D. Lane. *unzipFPGA: Enhancing FPGA-based CNN Engines with On-the-Fly Weights Generation.* IEEE International Symposium on Field-Programmable Custom Computing Machines (FCCM), 2021. (*equal contribution)

ICLR 2021 Shyam A. Tailor*, **Javier Fernandez-Marques***, Nicholas D. Lane. *Degree-Quant: Quantization-Aware Training for Graph Neural Networks.* International Conference on Learning Representations (ICLR), 2021. (*equal contribution)

MLSys 2020 **Javier Fernandez-Marques**, Paul N. Whatmough, Andrew Mundy and Matthew Mattina. *Searching for Winograd-aware Quantized Networks.* Conference on Machine Learning Systems (MLSys), 2020

ICLR 2019 Milad Alizadeh, **Javier Fernandez-Marques**, Nicholas D. Lane and Yarin Gal. *A Systematic Study of Binary Neural Networks' Optimisation.* International Conference on Learning Representations (ICLR), 2019

EMDL 2018 **Javier Fernandez-Marques**, Vincent W.-S. Tseng, Sourav Bhattacharya and Nicholas D. Lane. *On-the-fly deterministic binary filters for memory efficient keyword spotting applications on embedded devices.* International Workshop on Embedded and Mobile Deep Learning (EMDL-MobiSys), 2018

IJCAI 2018 Vincent W.-S. Tseng, Sourav Bhattacharya, **Javier Fernandez-Marques**, Milad Alizadeh, Catherine Tong and Nicholas D. Lane. *Deterministic Binary Filters for Convolutional Neural Networks.* International Joint Conference on Artificial Intelligence (IJCAI), 2018

SysML 2018 **Javier Fernandez-Marques**, Vincent W.-S. Tseng, Sourav Bhattacharya and Nicholas D. Lane. *BinaryCmd: Keyword Spotting with deterministic binary basis.* SysML, 2018

ICIP 2015 Marting Maška, Cristina Ederra, **Javier Fernandez-Marques**, Arrate Muñoz-Barrutia, Michal Kozubek and Carlos Ortiz-de-Solórzano. *Quantification of the 3D Collagen Network Geometry in Confocal Reflection Microscopy.* International Conference on Image Processing (ICIP), 2015

Software Skills

- **Languages**: Python, C++, MATLAB
- **Frameworks**: Pytorch, Tensorflow, NumPy and CUDA (basics)
- **Other**: Docker, git, slurm, LaTeX, 3D Studio Max

Awards

- EPSRC DPhil (PhD) scholarship at the University of Oxford (2017-2021)
- Best undergraduate thesis for the quality of my research.
- Honorific Mention award at UNIV congress (Rome, 2012) for a 2D computer graphics project.
- Scholarship award from Spanish Government: Leadership academy in Canada (summer 2009)

Other Activities

Secretary of the Spanish Society at the University of Oxford	2018-2019
College's photographer and graphic designer	2010-2013