Alexander Black

Research Scientist, Computer Vision

+447874709904 | alexander@black.com | alexblck.com | alexblck

Deep Learning Research Scientist with first author publications in CVPR, ICCV, AAAI, CVMP. Passionate about Computer Vision and Graphics with a specific focus on multimodal visual+text models for image and video explainability and generation. Seeking a Research Scientist role to apply my skills to an exciting projects within a talented and creative team.

Experience

Huawei Research & Development
London, UK

Research Scientist — *Jun 2024 - present*

- Designing a methodology for for enhancing LLMs with multi-task visual capabilities in a parameter efficient way.
- **PixArt:** Training a new iteration of photorealistic GenAI diffusion model.
- Supervising two interns on PEFT and Video Multimodal LLMs.
- Lead engineer for the LLM to VLM research direction, supporting multiple projects.
- Working with large scale multi-node distributed training, using accelerate, DeepSpeed, transformers and diffusers.

PhD - Computer Vision and Machine Learning
Guildford, UK

University of Surrey | Adobe Research — *Oct 2020 - Jun 2024*

- Robust image retrieval and manipulation localization. CVPR WMF 2021 [2]. US Patent approved [7].
- Multimodal (visual + audio) video deep metric learning, robust to augmentations and tampering. AAAI 2023 [8]
- Scalable (1M+) video retrieval with partial matching and localization. CVMP 2021 Best paper award [1]. US Patent approved [4].
- Image difference captioning with LLM using GenAI synthetic data. AAAI 2024 [6]. US Patent submitted for review.
- Order-aware edits captioning from a sequence on images and a novel sequence of manipulations dataset. US Patent submitted for review.

Gentian
London, UK

Computer Vision Engineer — *Nov 2022 - May 2024*

- **Habitat Mapping**: Predicting biodiversity index of areas with U-Nets using remote sensing data.
- Leading a team of 4 engineers to continuously train the models.
- Implementing a full multi-purpose machine learning pipeline for work with multispectral remote sensing data from the ground up.
- Delivered the project from ideation to MVP, achieving parity with expert performance.

Adobe Research
London, UK

Research Scientist Intern — *Jun 2022 - Dec 2022*

- Video retrieval, temporal alignment and manipulation detection. ICCV 2023[5].

FLOX
London, UK

Computer Vision Engineer — *Jan 2020 - Jun 2022*

- As a first technical hire, was responsible for growing and nurturing the ML team.
- Being responsible for all of the steps of ML process: data collection and labelling, pre-processing, model design and training, deployment.
- Designed and implemented a model based on YOLOv4 for detection of large number of chickens in commercial broiler sheds.
- Implemented semantic segmentation models using Detectron2.
- Project in collaboration with University College London: Object Detection with Synthetic Data in Unity.
- Multiple object tracking using DeepSORT.

Skills

Programming	Python, C/++, C#, Java, Lua, Go
ML/Dev Ops	Git, Unix, Docker, AWS/GCP, Wandb, DeepSpeed, FSDP, Hydra
Deep Learning	LLMs, Generative Models, Multimodal Fusion, Deep Metric Learning, Instance Retrieval, Image/Video Retrieval
Languages	English (C1), Estonian (native), Ukrainian (native), russian (native), Italian(B1), Chinese (A2)

Education

PhD - Computer Vision and Machine Learning *Guildford, UK*

University of Surrey | Adobe Research *Oct 2020 - Jun 2024*

- **Thesis:** Robust Video Fingerprinting and Comparison for Content Authenticity

MSc Computer Vision and Machine Learning (Distinction) *Guildford, UK*

University of Surrey *2019 - 2020*

- **Thesis:** Compositional Sketch-Based Image Retrieval. ICIP 2021 [3]
- Winner of MSc Advisory Board Prize for the best project in Electronic Engineering
- **Courses:** Computer Vision and Pattern Recognition, Image Processing and Deep Learning, Image and Video Compression, AI Programming, Speech Processing and Recognition, Robotics, Internet of Things, Satellite Remote Sensing

BSc Physics *Birmingham, UK*

University of Birmingham *2015 - 2019*

- **Courses:** Scientific Computing, Robotics and Machine Learning, Medical Imaging, Images and Communication, Physics of Music and Sound

Patents and Publications

[1] A. Black, T. Bui, S. Jenni, V. V. Swaminathan, and J. Collomosse. Vpn: Video provenance network for robust content attribution [best paper award]. *European Conference on Visual Media Production (CVMP)*, 2021.

[2] A. Black, J. Collomosse, T. Bui, H. Jin, and V. Swaminathan. Deep image comparator: Learning to visualize editorial change. *CVPR Workshop on Media Forensics (WMF)*, 2021.

[3] A. Black, J. Collomosse, T. Bui, L. Mai, and H. Jin. Compositional sketch search. *International Conference on Image Processing (ICIP)*, 2021.

[4] A. Black, J. Collomosse, S. Jenni, V. Swaminathan, et al. Determining video provenance utilizing deep learning, Feb. 29 2024. US Patent App. 17/822,573.

[5] A. Black, S. Jenni, T. Bui, M. M. Tanjim, S. Petrangeli, R. Sinha, V. Swaminathan, and J. Collomosse. Vader: Video alignment differencing and retrieval. *International Conference on Computer Vision (ICCV)*, 2023.

[6] A. Black, J. Shi, Y. Fan, T. Bui, and J. Collomosse. Vixen: Visual text comparison network for image difference captioning. *AAAI Conference on Artificial Intelligence*, 2024.

[7] J. Collomosse, A. Black, H. Jin, V. Swaminathan, et al. Identifying and localizing editorial changes to images utilizing deep learning, Nov. 30 2023. US Patent App. 17/804,376.

[8] S. Jenni, A. Black, and J. Collomosse. Audio-visual contrastive learning with temporal self-supervision. *AAAI Conference on Artificial Intelligence*, 2023.