Benjamin **Biggs**

Generative AI | Diffusion Models | Multi-Modal LLMs | 3D Reconstruction

San Francisco, United States

(+1) 628-280-9167 | benjbiggs@outlook.com | www.benbiggs.co.uk | benjiebob

Introduction

I am a Senior Applied Scientist in the **AGI team at Amazon**, where I work on advancing research in **image/video diffusion models** and **multi-modal large language models**. Most recently, I worked on the **Amazon Nova family of state-of-the-art generative models**. I received my **PhD in Computer Vision & Deep Learning from the University of Cambridge** where I focused on **3D reconstruction** of human and animal categories. I have multiple first-author publications at top-tier conferences and hold an **EB1-A Green Card for Extraordinary Ability**.

Education

PhD in Computer Vision and Machine Learning — **The University of Cambridge**

Supervisors: Andrew Fitzgibbon & Roberto Cipolla — *2017 - 2021*

- My research focused on developing methods for 3D reconstruction and tracking for challenging categories such as humans and animals.

BSc in Discrete Mathematics — **The University of Warwick**

Awarded: First Class Honours (81.4%) with 83.6% final-year average — *2012 - 2016*

- ***Best Overall Graduating BSc student in Discrete Mathematics*** and ***Best Final Year Project Prize***.

Research

The Amazon Nova Family of Generative Models — **Amazon.Science**

Amazon AGI

- **Technical lead of cross-org team** working on key science and data strategies behind **Amazon Reel** (text-to-video) and **Amazon Canvas** (text-to-image). Currently developing Amazon's **Any-to-Any Large Language Model**, announced at re:Invent 2024. More in the technical report.

Diffusion Soup: Model Merging for Text-to-Image Diffusion Models — **ECCV 2024**

B. Biggs†, A. Seshadri†, Y. Zou, A. Jain, A. Golatkar, Y. Xie, A. Achille, A. Swaminathan, S. Soatto

- Compartmentalization method for **text-to-image diffusion models**, which enables efficient training-free continual learning and unlearning. Method achieves up to 30% improvement vs. a paragon model, and has applications in anti-memorization and zero-shot style mixing.

Titan Image Generator v1&2, Text-to-Image Model — **Amazon.Science**

Amazon Bedrock

- Key technical contributor to Amazon's flagship text-to-image model which launched at re:Invent 2023. I led multiple workstreams on **post-training (SFT)**, large-scale data acquisition & preparation and responsible AI. More details available in the launch article.

3D Multi-bodies: Fitting Sets of Plausible 3D Human Models to Ambiguous Image Data — **NeurIPS 2020, Spotlight**

B. Biggs, S. Ehrhadt, H. Joo, B. Graham, A. Vedaldi, D. Novotny

- Presented a multi-hypothesis neural CNN for network for plausible 3D human reconstructions from challenging views, utilizing a novel quantization scheme based on normalizing flows.

Who Left the Dogs Out? 3D Animal Reconstruction with Expectation Maximization in the Loop — **ECCV 2020**

B. Biggs, O. Boyne, J. Charles, A. Fitzgibbon, R. Cipolla

- First fully automatic, end-to-end system for 3D animal reconstruction, combining a novel deformable model (**SMBLD**) with expectation maximization. Released **StanfordExtra**, the largest 2D animal keypoint dataset, with over 1000 downloads.

All Creatures Great and SMAL: Recovering the Shape and Motion of Animals from Video — **ACCV 2018, Oral**

B. Biggs, T. Roddick, A. Fitzgibbon, R. Cipolla

- Developed the first automatic system to reconstruct 3D shape and pose for various animals by training onusing synthetic data to overcome dataset limitations.

Shape of You: Precise 3D Shape Estimations for Diverse Body Types — **CVPR-W 2023**

R. Sarkar, A. Dave, G. Medioni, **B. Biggs**

- Dense correspondence loss function & test-time optimization routine improve 3D reconstruction accuracy for diverse body types.

On the Road to Large-Scale 3D Monocular Scene Reconstruction using Deep Implicit Functions — **ICCV-W 2021**

T. Roddick, **B. Biggs**, D.O. Reino, R Cipolla

- Deep implicit functions for large-scale driving scenes, trained with LiDAR scans.

For a full list of publications and patents, please see **www.benbiggs.co.uk**.

Experience

Senior Applied Scientist, Amazon Nova — Amazon AGI

San Francisco, USA — *Apr. 2024 - Present*

- Key contributor to science and data strategies behind **Nova Reel** (text-to-video), **Nova Canvas** (text-to-image) and upcoming features related to **Nova Pro** (multi-modal LLM).
- Technical lead & primary developer of a high-visibility project within Amazon Nova with multiple team members across multiple organizations.
- Delivered multiple high-profile demos to executive leadership, including **Andy Jassy (CEO)** and **Rohit Prasad (SVP, Amazon AGI)**.

Applied Scientist II — Amazon AWS

San Francisco, USA — *Oct. 2021 - Apr. 2024*

- Key contributor to **Titan Image Generator**, Amazon's flagship text-to-image diffusion model. I was responsible for the post-training SFT strategy for improving model aesthetics, conducting data & architecture ablations and onboarding new datasets. Published **Diffusion Soup paper which appeared at ECCV 2024**.
- Submitted a **patent** for work on outfit virtual try-on, and delivered projects on 3D shape estimation and selfie-based fashion recommendation.
- I have regularly presented in VP-level review meetings, filed patents, supervised PhD interns and interviewed candidates.

AI Research Intern — Facebook AI Research

London, UK — *Jun. 2019 - Nov. 2019*

- Supervised by **Andrea Vedaldi**, I worked on an end-to-end deep learning system for reconstructing humans in 3D from ambiguous/challenging monocular images. Submitted paper achieved a **Spotlight at NeurIPS 2020**.

AI Research Consultant — GlaxoSmithKline

Stevenage, UK — *Jun. 2014 - Oct. 2021*

- Machine learning consultant, working as a **leadership team** member for a fast-paced innovation group at **GlaxoSmithKline**. Led projects on action recognition, machine setting optimization, and defect detection.
- Filed a **patent** and demonstrated work to executive staff, incl. **Emma Walmsley (CEO)** and **Sir Andrew Witty (ex-CEO)**.

Technical Skills

Deep Learning and Optimization

- **Generative AI**: Highly proficient with **diffusion models** conditioned image and video synthesis, and **multi-modal large language models** (LLMs) for text generation.
- **3D reconstruction**: Highly proficient, particularly for articulated categories such as **humans or animals**. Experienced with NERFs and methods using morphable models (e.g. **SMPL**). Confident with multiple view geometry.
- Highly proficient with approaches for object detection, 2D/3D keypoint estimation, segmentation and classification.

Programming Languages and Frameworks

- Highly proficient in Python, including **PyTorch** for deep learning and **PySpark** for distributed data processing. Experience writing custom kernels and extensions.
- Experience with C++/C, C# .NET (incl. WPF and Office Add-Ins), MATLAB, Java, JavaScript and PHP / HTML 5.0. Also confident with LaTeX markup language.
- Extensive experience in **Unix** (**Ubuntu**, RedHat and Fedora) and Windows operating systems.
- Experience with **Git** / TFS source control, **HPC environments** (incl. SLURM), **Agile methodologies** (incl. test driven development), database technologies (incl. Blockchain, Microsoft SQL Server and Oracle) and mobile app development.

Teaching and Presentations

- Experience **mentoring & supervising**, including research teams, PhD interns and multiple students at master's and undergraduate stage. I have run various tutorials and seminars in academic and industrial contexts.
- Wealth of **presentation experience**, incl. **NeurIPS 2020/ACCV 2018** conference orals, technical presentations incl. **Epic Games** [San Francisco, US], **UC Berkeley** [San Francisco, US], **National Institute of Health** [North Carolina, US]. Numerous live demos to executive stakeholders.
- Experience engaging young students in STEM, incl. representing Microsoft at the Summer Science fair at **The Royal Society, London**.

Other

- **EB1-A Green Card** – permanent resident in the USA granted for extraordinary ability.
- USA and UK Driver's License Holder.

References

References from supervisors and past employers are available on request.