Arsha **Nagrani**

✉ arsha.nagrani@gmail.com | 🏠 www.robots.ox.ac.uk/ arsha/ | a-nagrani | 🔗 arsha-nagrani-601a726b/

Education ────────────────────────────────

University of Oxford *Oxford, UK*

DOCTOR OF PHILOSOPHY (PhD) IN COMPUTER VISION AND MACHINE LEARNING *Oct. 2016 - Oct. 2020*

- Thesis: 'Multimodal Video Understanding using Deep Learning'.
- Supervised by Prof. Andrew Zisserman in the Visual Geometery Group (VGG). Supported by an EPSRC grant and Google PhD Fellowship Award.

University of Cambridge *Cambridge, UK*

BA HONS AND MENG IN INFORMATION AND COMPUTER ENGINEERING, DISTINCTION (CLASS 1, 78.8%) *Oct. 2012 - Jun. 2016*

- Master's thesis supervised by Prof. Roberto Cipolla and Prof. Richard Turner in the Computer Vision group.

Honors & Awards ────────────────────────────

2021	**ICASSP Outstanding Paper Award**, For 'Slow-Fast Auditory Streams For Audio Recognition'		*IEEE*
2021	**Outstanding Reviewer Award**, CVPR		*IEEE*
2020	**Rising Star in EECS**, Awarded to EECS graduate and postdoctoral women		*UC Berkeley*
2018	**Google PhD Fellowship**, In Machine Perception, Speech Technology and Computer Vision		*Google LLC*
2017	**Interspeech Best Student Paper Award**, For 'VoxCeleb: A large-scale speaker identification dataset'		*ISCA*
2015,16	**Townsend Scholarship**, First Class degree of special merit, St. John's (twice)		*Cambridge, UK*
2016	**College Cargill Prize**, St. John's Cambridge		*Cambridge, UK*
2016	**Wright Prize**, St. John's Cambridge		*Cambridge, UK*
2016	**Shared Cunningham Year Prize**, St. John's Cambridge		*Cambridge, UK*
2012	**Award for a perfect score (45/45)**, International Baccalaureate (IB) examinations		*IBD*

Appointments ──────────────────────────────

Google Research *Cambridge, USA*

RESEARCH SCIENTIST *Oct. 2020 - Pres.*

- Video understanding using self-supervised and multi-modal machine learning.
- Manager: Cordelia Schmid.

Google Research *Mountain View, USA*

RESEARCH INTERN *May 2019 - Feb. 2020*

- Weakly supervised video understanding using deep learning.
- Supervised by Chen Sun, Cordelia Schmid and Rahul Sukthankar.

Cambridge, Computer Vision Lab *Cambridge, UK*

RESEARCH INTERN *Jul. 2015 - Sep. 2015*

- Implemented a Lettuce Image Classification tool using deep CNNs with the long term goal of automated lettuce harvesting.
- Funded by G's Global Growers and supervised by Roberto Cipolla.

Leadership/Outreach ──────────────────────────

2020	**Panelist**, Vaibhav Summit of Indian Scientists and Researchers organised by the Govt. of India		*India*
2020	**Visiting Researcher**, Wadhwani AI. Independent non-profit organisation developing AI solutions for social good. Research on detecting COVID-19 from cough sounds and 3D reconstruction of newborns.		*Mumbai, India*
2020	**Organiser**, Video Pentathlon, CVPR		*Seattle, USA*
2020,21	**Organiser**, Sight and Sound Workshop, CVPR		*Seattle, Virtual*
2019,20	**Founding Organiser**, VoxCeleb Speaker Recognition Challenge (VoxSRC19 and VoxSRC20) and Workshop		*Graz, Virtual*
2019,20	**Organiser**, Women in Computer Vision Workshop (WICV19 and WICV20), CVPR		*Long Beach, Virtual*
2018,19	**Mentor**, Black in AI Graduate Program		*Online*
2018-	**Reviewer**, ECCV, CVPR, ICCV, ICLR, BMVC, NeurIps, IEEE Access, numerous workshops		*Online*
2016	**Founder**, thefullapple technology publication for students		*Cambridge, UK*

Media

Teaching

Google Research India Summer School, 2020	Lecture on Video Understanding using Multimodality
International Conference on Data Mining (ICDM), 2020	Online Multimodal Knowledge Discovery Tutorial
Oxford University MPLS DTC on Statistics and Data Mining, 2020	Lectures on Applications of Machine Learning

Peer-Reviewed Publications

* Equal Contribution

Chen, H., Xie, W., Afouras, T., Nagrani, A., Vedaldi, A. and Zisserman, A. **"Localizing Visual Sounds the Hard Way"**. Conference on Computer Vision and Pattern Recognition (CVPR) 2021.

Seo, P.H., Nagrani, A. and Schmid, C. **"Look Before you Speak: Visually Contextualized Utterances"**. Conference on Computer Vision and Pattern Recognition (CVPR) 2021.

Kazakos, E., Nagrani, A., Zisserman, A. and Damen, D. **"Slow-Fast Auditory Streams For Audio Recognition"**. International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2021. **Outstanding Paper Award**

Brown, A.*, Huh, J.*, Nagrani, A.*, Brown, A., Chung, J. S., and Zisserman, A. **"Playing a Part: Speaker Verification at the Movies"**. International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2021.

Bain, M., Nagrani, A., Brown, A., and Zisserman, A. **"Condensed Movies: Story Based Retrieval with Contextual Embeddings"**. Asian Conference on Computer Vision (ACCV) 2020. **Oral Presentation**

Chung, J. S.*, Huh, J.*, Nagrani, A.*, Afouras, T., and Zisserman, A. **"Spot the conversation: speaker diarisation in the wild "**. INTERSPEECH 2020.

Arnab, A., Sun, C., Nagrani, A. and Schmid, C. **"Uncertainty-Aware Weakly Supervised Action Detection from Untrimmed Videos"**. European Conference on Computer Vision (ECCV) 2020.

Nagrani, A., Sun, C., Ross, D., Sukthankar, R., Schmid, C. and Zisserman, A. **"Speech2Action: Cross-modal Supervision for Action Recognition"**. Conference on Computer Vision and Pattern Recognition (CVPR) 2020.

Nagrani, A.*, Chung, J.S.*, Albanie, S.* and Zisserman, A. **"Disentangled Speech Embeddings using Cross-modal Self-supervision"**. International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020.

Bain, M., Nagrani, A., Schofield, D. and Zisserman, A. **"Count, Crop and Recognise: Fine-Grained Recognition in the Wild"**. International Conference on Computer Vision (ICCV) Workshop, 2019. **Oral Presentation**

Nagrani, A.*, Chung, J.S.*, Xie, W. and Zisserman A., **"VoxCeleb: Large-scale Speaker Verification in the Wild"**. Computer Speech and Language (CSL), 2019.

Kazakos, E., Nagrani, A., Zisserman, A. and Damen, D. **"EPIC-Fusion: Audio-Visual Temporal Binding for Egocentric Action Recognition"**. International Conference on Computer Vision (ICCV), 2019. **Second Highest Score at the Epic-Kitchens Action Recognition Competition at CVPR 2019**

Schofield, D.*, and Nagrani, A.*, Zisserman, A. Hayashi, M. Matsuzawa, T. Biro, D. Carvalho, S. **"Chimpanzee Face Recognition from videos in the wild using Deep Learning"**. Science Advances 5(9), eaaw0736, 2019

Liu, Y.* Albanie, S.* Nagrani, A.* and Zisserman, A., **"Use What You Have: Video retrieval using representations from collaborative experts"**. British Machine Vision Conference (BMVC), 2019

Xie, W. Nagrani, A. Chung, J. S. Zisserman A., **"Utterance-level Aggregation For Speaker Recognition In The Wild"**. International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2019 **Oral Presentation**

Albanie, S.* Nagrani, A.* and Zisserman, A., **"Emotion recognition in speech using Cross-modal Transfer in the wild"**. ACM multimedia, 2018

Nagrani, A.* Albanie, S.* and Zisserman, A., **"Learnable PINs: Cross-Modal Embeddings for Person Identity"**. European conference on computer vision (ECCV), 2018

Chung, J.S.* Nagrani, A.* and Zisserman, A., **"VoxCeleb2: Deep Speaker Recognition"** INTERSPEECH, 2018.

Nagrani, A. Albanie, S. and Zisserman, A., **"Seeing Voices and Hearing Faces: Cross-modal biometric matching"**. Conference on Computer Vision and Pattern Recognition (CVPR) 2018 **Spotlight**

Nagrani, A. and Zisserman, A., **"From Benedict Cumberbatch to Sherlock Holmes: Character Identification in TV series without a Script."**. British Machine Vision Conference (BMVC), 2017. **Oral Presentation**

Nagrani, A. Chung, J.S. and Zisserman, A., **"VoxCeleb: a large-scale speaker identification dataset."** INTERPSEECH, 2017. **Oral Presentation, Best Student Paper Award**