Eduardo Fonseca

Research Scientist

92 7th Ave 8
11217 NYC, NY
+1 646 755 1511
efonseca@google.com
www.eduardofonseca.net
eduardofonsecamontero
*edfonseca_*

Current Position

March 2022 – present **Research Scientist**,
Google Research, NYC.
I am currently a Research Scientist working in the Sound Understanding Group on machine learning for audio processing. My research interests include learning algorithms for audio processing using different types of supervision, including self-supervised learning, multimodal learning and learning with noisy labels.

Work Experience

Oct 2016 – Dec 2021 **PhD Thesis: *Training Sound Event Classifiers Using Different Types of Supervision***,
Music Technology Group, Universitat Pompeu Fabra, Barcelona,
PhD Candidate.
Some highlights include the Best Audio Representation Learning Paper Award at WASPAA21, and the FSD50K paper and dataset, as well as approaches for CNN architectures, self-supervised learning and learning with noisy labels. It was partially funded by two Google Faculty Research Awards and included two internships at Google.

Oct 2020 – Feb 2021 **Research Internship**,
Sound Understanding, Google Research, NYC (virtual),
Returning Research Intern.
Exploration of self-supervised audio representation learning enabled by source separation pre-processing.

Sept – Dec 2019 **Research Internship**,
Sound Understanding, Google Research, NYC,
Research Intern.
Mitigating label noise impact when training sound event recognizers using large amounts of weakly labeled data.

April 2016 – Feb 2019 **AudioCommons project**,
Music Technology Group, Universitat Pompeu Fabra, Barcelona,
PhD Candidate.
Creation of the Freesound Annotator platform used to build and maintain several open audio datasets. Development of deep learning approaches for acoustic scene and sound event classification.

Sept 2014 – March 2016 **PHONEDRIVE project**,
Telefónica R&D & Technical University of Madrid,
R&D Engineer.
Design of a driving pattern recognition system using signals from accelerometers and gyroscopes in smartphones.

June 2015 **Forensic Acoustics**,
Spanish National Research Council (CSIC),
Freelance Acoustic Engineer.
Analysis of speech recordings carried out in CSIC's Phonetics Laboratory with forensic acoustic purposes.

Oct 2010 – Dec 2013 **HECCO (Communications Coverage Evaluation Tool) project**,
INECO & Technical University of Madrid,
R&D Engineer.
Machine learning and signal processing for objective speech quality assessment in air-ground communications.

July 2006 **Sound Technician Internship**,
Radiotelevisión Española,
Intern.
Audio production of news bulletin and basic tv programmes.

Sept 2003 – June 2004 **Radio station Coordinator**,
San Juan Evangelista University Residence.
Management tasks and maintenance of equipment in an amateur radio studio.

Selection of Publications (full list available at Google Scholar) - 1034 citations

[1] Mariana-Iuliana Georgescu, Eduardo Fonseca, Radu Tudor Ionescu, Mario Lucic, Cordelia Schmid, and Anurag Arnab. Audiovisual masked autoencoders. *arXiv preprint arXiv:2212.05922*, 2022.

[2] Eduardo Fonseca, Xavier Favory, Jordi Pons, Frederic Font, and Xavier Serra. FSD50K: an open dataset of human-labeled sound events. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 30:829–852, 2022.

[3] Eduardo Fonseca, Aren Jansen, Daniel PW Ellis, Scott Wisdom, Marco Tagliasacchi, John R Hershey, Manoj Plakal, Shawn Hershey, R Channing Moore, and Xavier Serra. Self-supervised learning from automatically separated sound scenes. In *2021 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 251–255. IEEE, 2021.

[4] Eduardo Fonseca, Andres Ferraro, and Xavier Serra. Improving sound event classification by increasing shift invariance in convolutional neural networks. *arXiv preprint arXiv:2107.00623*, 2021.

[5] Eduardo Fonseca, Diego Ortego, Kevin McGuinness, Noel E O'Connor, and Xavier Serra. Unsupervised contrastive learning of sound event representations. In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2021.

[6] Eduardo Fonseca, Shawn Hershey, Manoj Plakal, Daniel PW Ellis, Aren Jansen, and Moore R Channing. Addressing missing labels in large-scale sound event recognition using a teacher-student framework with loss masking. *IEEE Signal Processing Letters*, 27, 2020.

[7] Eduardo Fonseca, Frederic Font, and Xavier Serra. Model-agnostic approaches to handling noisy labels when training sound event classifiers. In *Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA19)*, New York, US, 2019.

[8] Eduardo Fonseca, Manoj Plakal, Frederic Font, Daniel P. W. Ellis, and Xavier Serra. Audio tagging with noisy labels and minimal supervision. In *Proceedings of the Detection and Classification of Acoustic Scenes and Events 2019 Workshop (DCASE2019)*, 2019.

[9] Eduardo Fonseca, Manoj Plakal, Daniel P. W. Ellis, Frederic Font, Xavier Favory, and Xavier Serra. Learning sound event classifiers from web audio with noisy labels. In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2019.

[10] Eduardo Fonseca, Rong Gong, Dmitry Bogdanov, Olga Slizovskaia, Emilia Gómez Gutiérrez, and Xavier Serra. Acoustic scene classification by ensembling gradient boosting machine and convolutional neural networks. In *Detection and Classification of Acoustic Scenes and Events 2017 Workshop (DCASE2017)*, 2017.

Teaching Experience

2019 – 2020 **Audio and Music Processing Lab - Intro to Sound Recognition**, *Master's degree in Sound and Music Computing*, Universitat Pompeu Fabra.

July 2018 **Introduction to Python and music**, *Girls in Data Science Seminar*, Universitat Pompeu Fabra.

2016 – 2018 **Sound and Music Processing**, *Bachelor's degree in Audiovisual Systems Engineering*, Universitat Pompeu Fabra.

2016 – 2017 **Audio Signal Processing for Music Applications**, *Coursera*, Stanford University & Universitat Pompeu Fabra.

2017 **Lab of Signals and Systems**, *Bachelor's degree in Audiovisual Systems Engineering*, Universitat Pompeu Fabra.

Awards

2021 **Best Audio Representation Learning Paper Award at WASPAA21**.
For the paper titled *Self-Supervised Learning from Automatically Separated Sound Scenes*.

2017 & 2018 **Google Faculty Research Award**.
Project proposals developed under the supervision of Dr. Xavier Serra focused on *(i)* the creation of FSD, an open dataset for sound event recognition based on Freesound content organized with the AudioSet Ontology, and *(ii)* its exploitation for machine listening competitions to foster sound event recognition research.

Co-Curricular

2019 **Detection and Classification of Acoustic Scenes and Events (DCASE) 2019 Task2 Organizer**, Music Technology Group & Google's Sound Understanding.
Co-organization of the *Audio tagging with noisy labels and minimal supervision* competition, which was run on the Kaggle platform with 880 teams participating.

2018 **Detection and Classification of Acoustic Scenes and Events (DCASE) 2018 Task2 Organizer**, Music Technology Group & Google's Sound Understanding.
Co-organization of the *General-purpose audio tagging of Freesound content with AudioSet labels* competition, which was run on the Kaggle platform with 558 teams participating.

2017 **Sónar Innovation Challenge Mentor**,
Sónar Festival & Music Technology Group.
Co-mentoring a team of technologists in the development of a tool for 3D exploration of Freesound based on sound similarity.

Education

2021 **PhD in Information and Communications Technology**,
Universitat Pompeu Fabra (Barcelona, Spain).
Dissertation titled *Training Sound Event Classifiers Using Different Types of Supervision*, carried out under the supervision of Dr. Xavier Serra and Dr. Frederic Font.

2015 **Master's Degree in Telecommunications Engineering**, *specialisation in Communications*,
Technical University of Madrid (Spain).
This qualification has been accredited by the Accreditation Board for Engineering and Technology (ABET).

2009 **Master of Science (MSc) in Engineering**, *specialisation in Acoustics*,
Aalborg University (Denmark).

2007 **Bachelor's Degree in Telecommunications Engineering**, *specialisation in Sound and Image*,
Technical University of Madrid (Spain).

2007 **Erasmus scholarship**,
Helsinki Polytechnic Stadia (Finland), (currently Helsinki Metropolia University of Applied Sciences).
Bachelor's final project and additional courses on digital signal processing and English communication.

Languages

Spanish Native Speaker

English High level

2010 Certificate in Advanced English (CAE), Grade B, University of Cambridge

2007 International English Language Testing System (IELTS) Academic Version, Grade 7, University of Cambridge

2000 Grade nine in Spoken English for Speakers of other Languages, Trinity College London

Computing & Technical Skills

Software *High*: Python, Keras, TensorFlow, MATLAB | *Basic*: PyTorch, JAX, Pax, Apache Beam, HTML

Acoustics *Simulation*: CATT Acoustics, EASE | *Measurement*: MLSSA, 01 dB Symphonie, Brüel & Kjær PULSE

Received Courses

2017 **Workshop on acoustic array settings for spectacles**, *1 week*,
Universitat Pompeu Fabra.

2015 **Course on *Machine Learning***, *3 months*,
Stanford University through Coursera.

2014 **Workshop on *Multichannel Audio***, *14 hours*,
Laboral Centro de Arte y Creacion Industrial (Gijon, Spain).

2010 **Workshop on *Experimental Audio with Sonotron***, *20 hours*,
Laboral Centro de Arte y Creacion Industrial (Gijon, Spain).

2005 **Course on *Telecommunication Infrastructure Projects***, *3 credits*,
Technical University of Madrid.