

Education

- September 2019–July 2023 **PhD in Machine Learning**, TÉLÉCOM PARIS, Institut Polytechnique de Paris, France
Thesis: A Flexible Framework for Interpretable Machine Learning: Application to Image and Audio Classification
Advisors: Prof. Florence d'Alché-Buc & Prof. Pavlo Mozharovskyi, Telecom Paris
Jury: Grégoire Montavon, Patrick Pérez, David Alvarez-Melis, Nicolas Thome, Stéphane Canu, Chloe Clavel
- Developed post-hoc/by-design interpretation methods using concept-based representations for deep neural networks, with applications to audio and image classification tasks
- July 2014–19 **Dual Degree (B.Tech + M.Tech)**, IIT Bombay, Electrical Engineering, CPI: 9.05/10
Thesis: Audio Style Transfer: Transformations between speech and singing
Advisors: Prof. Preeti Rao, IIT Bombay & Dr. Yi-Hsuan Yang, Academia Sinica, Taiwan
- Investigated style transfer techniques for audio signals
 - Focused on converting spoken audio into sung audio and vice-versa using deep learning

Awards and Achievements

- 2023 Awarded **Top reviewer (top 10%)** for NeurIPS'23.
- 2023 Awarded STIC Doctoral Prize (Saclay) 2023 **accessit (2nd place)** for NeurIPS'22 publication.
- 2015–18 Awarded **AP grade** for exceptional performance in the course EE763: Science of Information, Statistics and Learning (Spring 17-18), and ES200: Environmental Studies (Autumn 15-16)
- 2016 Awarded **travel grant** and **distinctive mention** for work at MediaEval 2016 Workshop held at Netherlands Institute of Sound and Vision, Hilversum, Netherlands
- 2013 Awarded **Kishore Vaigyanik Protsahan Yojana (KVPY) Fellowship** 2013

Selected Publications/Preprints

Complete list of papers on [Google Scholar](#). * denotes equal contribution below

- [1] P. Khayatan*, M. Shukor*, **J. Parekh***, and M. Cord. "Analyzing Fine-tuning Representation Shift for Multimodal LLMs Steering." **arXiv 2025** (under submission).
- [2] **J. Parekh***, Q. Bouniot*, P. Mozharovskyi, A. Newson and F. Buc. "Restyling Unsupervised Concept Based Interpretable Networks with Generative Models." **Accepted at ICLR 2025**
- [3] **J. Parekh**, P. Khayatan, M. Shukor, A. Newson and M. Cord. "A Concept-Based Explainability Framework for Large Multimodal Models." **NeurIPS 2024**.
- [4] **J. Parekh**, S. Parekh, P. Mozharovskyi, G. Richard and F. Buc. "Tackling Interpretability in Audio Classification Networks with Non-negative Matrix Factorization." **IEEE/ACM TASLP** (2023).
- [5] **J. Parekh**, S. Parekh, P. Mozharovskyi, F. Buc and G. Richard. "Listen to Interpret: Post-hoc Interpretability for Audio Networks with NMF." **NeurIPS 2022**.
- [6] **J. Parekh**, P. Mozharovskyi & F. Buc. "A Framework to Learn with Interpretation." **NeurIPS 2021**.
- [7] **J. Parekh**, P. Rao, and YH. Yang. "Speech-to-Singing Conversion in an Encoder-Decoder Framework." **IEEE ICASSP 2020 (Oral)**.

Work Experience and Projects

- December 2023–Present **Postdoctoral Researcher**, SORBONNE UNIVERSITÉ France
[Understanding Representations in Large Multimodal Models](#)
Advisor: Prof. Matthieu Cord, (Sorbonne Université and Scientific director @ valeo.ai)
- Developing methods to understand representations computed in large visual-language multimodal models.

Feb–April **Supervised Research Exposition**, IIT Bombay

2018 [Surface Defect Detection](#)

- Advisor:** Prof. Subhasis Chaudhuri, Department of Electrical Engineering, IIT Bombay
- Explored various techniques and studied relevant literature for surface defect detection
 - Applied transfer-learning based methods for plant disease detection

May–June **Summer Intern**, TECHNICOLOR R&D France

2017 [2D & 3D Human Pose Estimation Networks](#)

- Advisors:** Pierre Hellier (Principal Scientist) & Louis Chevallier, Technicolor R&D France
- Completed internship as part of an industrial project on Motion Synthesis in Animation
 - Implemented and tested stacked hourglass based deep CNNs for 2D and 3D human pose estimation

2016 & 2017 **MediaEval Benchmarking Initiative**

[Predicting Media Interestingness Task](#)

Organizers: Technicolor, France, ETH Zurich *et al.*

- Developed novel methods for ranking a set of images/video-shots extracted from movie trailers according to their interestingness to a common viewer
- Netherlands 2016: achieved a mean average precision (MAP) of 0.23 for images - **Team Rank 3/12**
- Ireland 2017 (web participation): achieved a MAP of 0.25 for images & 0.19 for videos

2016–18 **Selected Undergraduate Projects**

- [Image Style Transfer using Graph-CNN](#) – Implemented a unsupervised, graph signal processing based, random shallow CNN for image style transfer
- [Blind Audio Source Separation](#) – Implemented a NMF, LPC based error clustering criterion algorithm
- [Detection of Moving Objects in Videos](#) – Studied and implemented simplified version of a paper based on mean-shift and max-flow min-cut algorithms for the same
- [Artificial Synesthesia](#) – Implemented a CCA-based algorithm for image-audio cross modal retrieval

Services

Organizing Trustworthy and Frugal ML workshop, ELLIS Unconference 2023 Paris (Co-chair), ELLIS Member

Reviewing CVPR'25, ICLR'25, NeurIPS'24, ICML'24, NeurIPS'23, TPAMI, JMLR, TMM

Teaching Served as teaching assistant for courses: **Machine Learning** (Telecom Paris, 3X during 2020–22), **Analytical Signal Processing** (IIT Bombay, Spring 18-19), **Probability & Random Processes** (IIT Bombay, Autumn 18-19) and **Linear Algebra** (IIT Bombay, Autumn 17-18)

Skills

Programming Languages: Python (including PyTorch, Keras, Tensorflow), C/C++, VHDL

Courses Advanced courses in ML/AI, CV, Optimization, Signal Processing, Summer schools – MuSTeR 2016 (IISc Bangalore), MLSS 2021 (Virtual), OxML 2022 (University of Oxford)

Extra-Curricular

Sports

2014-18 **Silver medal** in Institute Squash League, **3X finalist** of Squash General Championship

2007-09 **Participated** in Junior National Squash Championship, Otters Open, CCI Open, HC Golcha Memorial Rajasthan Open Squash Championship, Umed Club District Open Squash Championship (Finalist)

Cultural

2008-12 Member of Choir group in DPS Jodhpur

2009 & 2005 Samvaad – A personality development program – 45 and 30 days respectively

Selected Talks

Apr 2024 Invited talk at ICASSP'24 workshop on XAI for Speech and Audio – “Tackling Interpretability Problems for Audio Classification Networks with NMF”

Sept 2023 Deezer Research Paris – “Listen to Interpret: Post-hoc interpretability for Audio Networks with NMF”

2021–2022 Talks at IDEMIA and Datacraft Paris – “A Framework to Learn with Interpretation”