Salah Zaiem

Website: salah-zaiem.github.io
Email: zaiemsalah@gmail.com
LinkedIn: salah-zaiem
GitHub: github.com/salah-zaiem

Education

Telecom Paris, Institut Polytechnique de Paris — Paris, France
Ph.D. in speech processing and Machine Learning, Advisors: Slim Essid and Titouan Parcollet — 2020–2024

- Thesis title: "Informed Training, Evaluation and Exploitation of Speech Self-Supervised Representations"

École Normale Supérieure Paris-Saclay — Paris, France
Master MVA, Highest Honors — 2018–2019

- Top Machine Learning Research Master in France. MVA stands for "Mathematics, Vision and Learning(Apprentissage)"
- Thesis title : "Better Representations for Unsupervised Spoken Term Discovery"

École polytechnique — Paris, France
Engineering Diploma, GPA: 3.91/4.00 — 2015–2018

- Applied mathematics and Computer Science majors.

Lycée Pierre de Fermat — Toulouse, France
Classe Préparatoire MP — 2013–2015

- Preparation for the competitive entrance exam to engineering schools. Focus on mathematics and physics.

Experience

Google Deepmind — Paris, France
Research Scientist — April 2024- Current

- Audio-Visual synthesis within the GenMedia team.
- Veo 2: State-of-the-Art video generation. Blogpost link
- Video-to-audio generation. Blogpost link

Ninth Frederick Jelinek Memorial Summer Workshop — Le Mans, France
Researcher — Summer 2023

- Project: Finite state methods with modern neural architectures for speech applications and beyond
- Working on FST-based inputs for Spoken Language Understanding in high Word-Error-Rates scenarios

Montreal Institute for Learning Algorithms (MILA) — Montréal, Canada
Visiting Researcher — Winter 2023

- Project: Better benchmarking for speech self-supervised representations
- Advisor: Mirco Ravanelli

Google Research — Zurich, Switzerland
Research Intern — Summer 2022

- Speech enhancement using discrete audio representations
- Advisors: Zalàn Borsos and Félix de Chaumont-Quitry

École Normale Supérieure, Cognitive Machine Learning Team — Paris, France
Research Engineer — 2020

– Working on multilingual Text-to-Speech Synthesis
– Advisor: Ewan Dunbar

Inria Paris, France
Research Intern Summer 2019

– Unsupervised learning techniques for spoken word discovery from raw speech data
– Advisor: Emmanuel Dupoux

National Bank of Canada Montréal, Canada
Research Intern Summer 2018

– Improving documents search using neural Query Expansion
– Part-time researcher at Université du Québec à Montréal, supervised by Fatiha Sadat

Teaching

- **Teaching Assistant** at École Polytechnique Spring 2021,2022
  Advanced Machine Learning
- **Teaching Assistant** at Telecom Paris Fall 2021
  Speech and Audio Processing
- **Teaching Assistant** at Telecom Paris Spring 2021,2022
  Machine Learning for Text Mining and NLP
- **Mathematics Oral Examiner (Colleur)** at Lycée Chaptal and Lycée de Vilgénis 2017,2019,2020
  Training in probabilities, analysis and linear algebra for undergraduates.

Skills

- **Deep Learning:** PyTorch, TensorFlow, JAX
- **Audio/Speech toolkits:** SpeechBrain (Core contributor), Librosa, k2
- **Other:** Git, Scikit-Learn

Languages

- **Code:** Python, bash, R, C++
- **Natural Language:** Arabic & French (Both native), English (Full proficiency), Spanish (Good commmand: C1 level)

Mentorship & Service

- **Reviewer** : Interspeech 2023/2024, TASLP 2023, ICASSP 2024
- **Organiser** : JJCAAS 2023: Journées Jeunes Chercheur · se · s en Audition, Acoustique musicale et Signal audio (French young audio researchers conference)
- **Core Contributor** : SpeechBrain Library. Developed the MP3S self-supervision benchmark.
- Research **internship supervisor** : Hugo Malard (Title: Dynamic model size selection for efficient automatic speech recognition)
- Research **internship supervisor** : Ahmed Ben Abdallah (Title: Pushing the boundaries of Tunisian code-switched automatic speech recognition)
- Research project supervisor : Nour Essayegh, Jianshu Zhu (Title: Self-supervised learning for speech recognition on low-resource languages)

Scholarships and Awards

- Nominated for **Best Student Paper Award** at **Interspeech** (13 Nominees out of $>$ 1100 accepted papers). 2023
- ISCA **Travel Grant** (registration + membership + travel funds) to attend InterSpeech 2021 2021

- Ecole polytechnique : Prix du stage de recherche. Prize given to the best research internships. 2018
- Quebec Hackathon on Environmental data: Rank **2** (out of 26 Teams).**2500$** prize. 2018
- **Excellence** Scholarship for the best performing Tunisian students in French engineering exams. 2015–2019

Publications

[1] H. Malard, **S. Zaiem**, and R. Algayres, "Big model only for hard audios: Sample dependent whisper model selection for efficient inferences", *Submitted to ICASSP 2024*,

[2] L. Della Libera, P. Mousavi, **S. Zaiem**, C. Subakan, and M. Ravanelli, "Cl-masr: A continual learning benchmark for multilingual asr", *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2024.

[3] S. Mdhaffar, F. Bougares, R. De Mori, **S. Zaiem**, M. Ravanelli, and Y. Estève, "Taric-slu: A tunisian benchmark dataset for spoken language understanding", in *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, 2024, pp. 15 606–15 616.

[4] P. Mousavi, J. Duret, **S. Zaiem**, L. Della Libera, A. Ploujnikov, C. Subakan, and M. Ravanelli, "How should we extract discrete audio tokens from self-supervised models?", *arXiv preprint arXiv:2406.10735*, 2024.

[5] M. Ravanelli, T. Parcollet, A. Moumen, S. de Langen, C. Subakan, P. Plantinga, Y. Wang, P. Mousavi, L. Della Libera, A. Ploujnikov, *et al.*, "Open-source conversational ai with speechbrain 1.0", *Journal of Machine Learning Research*, vol. 25, no. 333, pp. 1–11, 2024.

[6] **S. Zaiem**, T. Parcollet, and S. Essid, "Less forgetting for better generalization: Exploring continual-learning fine-tuning methods for speech self-supervised representations", *Preprint*, 2024.

[7] A. A. Ben Abdallah*, A. Kabboudi, A. Kanoun, and S. **Zaiem***, "Leveraging data collection and unsupervised learning for code-switched tunisian arabic automatic speech recognition", *ICASSP 2024*, vol. * : These two authors have contributed equally. 2023.

[8] G. Wright, U. Cappellazzo, **S. Zaiem**, D. Raj, L. Ondel Yang, D. Falavigna, and A. Brutti, "Training dynamic models using early exits for automatic speech recognition on resource-constrained devices", *ICASSP 2024*, 2023.

[9] **S. Zaiem**, R. Algayres, T. Parcollet, S. Essid, and M. Ravanelli, "Fine-tuning strategies for faster inference using speech self-supervised models: A comparative study", in *2023 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2023.

[10] **S. Zaiem**, Y. Kemiche, T. Parcollet, S. Essid, and M. Ravanelli, "Speech Self-Supervised Representation Benchmarking: Are We Doing it Right?", in *Proc. INTERSPEECH 2023*, 2023, pp. 2873–2877.

[11] **S. Zaiem**, Y. Kemiche, T. Parcollet, S. Essid, and M. Ravanelli, "Speech Self-Supervised Representations Benchmarking: a Case for Larger Probing Heads", *Computer, Speech and Language*, 2023. arXiv: 2308.14456.

[12] **S. Zaiem**, T. Parcollet, and S. Essid, "Automatic Data Augmentation for Domain Adapted Fine-Tuning of Self-Supervised Speech Representations", in *Proc. INTERSPEECH 2023*, 2023, pp. 67–71.

[13] R. Algayres, T. Ricoul, J. Karadayi, H. Laurençon, **S. Zaiem**, A. Mohamed, B. Sagot, and E. Dupoux, "Dp-parse: Finding word boundaries from raw speech with an instance lexicon", *Transactions of the Association for Computational Linguistics*, vol. 10, pp. 1051–1065, 2022.

[14] Y. Gao, T. Parcollet, **S. Zaiem**, J. Fernandez-Marques, P. P. de Gusmao, D. J. Beutel, and N. D. Lane, "End-to-end speech recognition from federated acoustic models", in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022, pp. 7227–7231.

[15] **S. Zaiem**, T. Parcollet, and S. Essid, "Automatic Data Augmentation Selection and Parametrization in Contrastive Self-Supervised Speech Representation Learning", in *Proc. Interspeech 2022*, 2022, pp. 669–673.

[16] **S. Zaiem**, T. Parcollet, S. Essid, and A. Heba, "Pretext tasks selection for multitask self-supervised audio representation learning", *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 6, pp. 1439–1453, 2022.

[17] **S. Zaiem**, T. Parcollet, and S. Essid, "Conditional Independence for Pretext Task Selection in Self-Supervised Speech Representation Learning", in *Proc. Interspeech 2021*, 2021, pp. 2851–2855.

[18] R. Algayres, **S. Zaiem**, B. Sagot, and E. Dupoux, "Evaluating the Reliability of Acoustic Speech Embeddings", in *Proc. Interspeech 2020*, 2020, pp. 4621–4625.

[19] **S. Zaiem** and E. Bennequin, "Learning to communicate in multi-agent reinforcement learning: A review", *arXiv preprint arXiv:1911.05438*, 2019.

[20] **S. Zaiem** and F. Sadat, "Sequence to sequence learning for query expansion", in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, 2019, pp. 10 075–10 076.

See Google Scholar profile for more details.