Leonardo Boulitreau

AI Audio Research Engineer

leonardoboulitreau@gmail.com | leonardoboulitreau.netlify.app | leonardoboulitreau | leonardoboulitreau | leonardomarques

About Me

I am a researcher in the intersection of artificial intelligence and speech, with over 5 years of experience and a solid background in signal processing. Currently, I work as an AI Audio Researcher **@CPqD**. My research efforts are directed towards developing human-machine interaction systems that are more affective and human-empowering.

Experience

AI Researcher

São Paulo, Brazil

@CPqD - Research and Development Center in Telecommunications *2023 - Present*

- Improved the company's emotion recognition system in production, in terms of speed and performance, by developing a 2-step self-supervised based neural emotion recognition architecture.
- Enriched the company's call center customer profiling product by developing and sending to production age and gender classifier models.
- Currently upgrading the company's automatic speech recognition (ASR) system with external large language model fusion techniques.

Fellow Master

São Paulo, Brazil

@CPqD - Research and Development Center in Telecommunications *2021 - 2023*

- Developed neural customer-oriented expressive text-to-speech models for the Brazilian Portuguese language.
- Enabled customers to customize the synthesized audios' prosody in a fine-grained way by implementing character-level prosody control graph pipelines.
- Designed and conducted several subjective perceptual experiments to evaluate synthetic audio quality and naturalness.

Laboratory Intern

Paraíba, Brazil

@ LPS - Prototyping and Simulation Laboratory *2020 - 2021*

- Enhanced lab automation by designing neural speech commands recognition systems.
- Encapsulated the ASR system in a local private LoRa network for IoT purposes.
- Enabled long distance voice control by developing a wearable prototype with an embedded microphone.

Skills

Frameworks	Coqui, SpeechBrain, Amphion, ESPNET, NeMo
Deep Learning	PyTorch, Tensorflow, HuggingFace, ONNX, Gradio, MLFlow
Tools	Docker, Git, Kubernetes
Cloud	AWS, GCP
Programming	Python, C, C++, MATLAB, LaTeX
Languages	Portuguese, English, French
Competences	Paper Implementation, Experiment Design, Team Collaboration

Education

M.Sc. in Computer Engineering

São Paulo, Brazil

@UNICAMP (State University of Campinas) *2021 - 2024*

- GPA: 5/5
- Thesis: Cross-Speaker Style Transfer for TTS based on Singing Voice Conversion Data Augmentation, Style Filtering, and F0 Matching
- Study - On The Cross Modality of Pre-Training: Can models pre-trained in different domains adapt to a different one on a downstream task?
- Study - Meta-learning strategies to build a few-shot emotion classifier.

Exchange Student (Excellence Scholarship)

Paris, France

@Télécom Paris *2019 - 2020*

- Studying Machine Learning, Statistics, Optimization, Digital Signal Processing, IoT and Telecommunications.
- Cross-Cultural Social Experiences.

B.Sc. in Electronic Engineering

Paraíba, Brazil

@UFPB (State University of Paraíba) *2015 - 2021*

- GPA: 8.38/10
- Research - Ultrasonic time-of-flight estimation via the Extended Kalman Filter (DSP).
- Participated in the Electronics team of the national Aerodesign competition.
- Differential and Integral Calculus Tutor.

Projects

Cross-Speaker Style Transfer

Github

Co-Creator — *2021 - Current*

- Augmenting standard state-of-the art text-to-speech models to make a speaker with no emotional data speak emotionally.
- Implementation of generative architectures to model style: diffusion models, normalizing flows, VQ-VAEs, VAEs, etc.
- Exploit of representation learning techniques for disentangling style and speaker information: gradient reversal, information perturbation, etc.

The Sound of AI - Accelerator

Online

Participant — *2023 - Interrupted*

- Selected for the first edition. Taken start-up courses focused on AI+Audio businesses.

TTS Objective Metrics

GitHub

Creator — *2022*

- Repository containing a compilation of objective metrics used in several text-to-speech papers.

Publications

Exploring Synthetic Data for Cross-Speaker Style Transfer in Style Representation based TTS

Kos Island, Greece

SynData4GenAI Workshop 2024 — *2024*

- 0 Citations

Benchmarking Speech-Driven Gesture Generation Models for Generalization to Unseen Voices and Noisy Environments

Costa Rica

GENEA: Generation and Evaluation of Non-verbal Behaviour for Embodied Agents Workshop 2024 — *2024*

- 0 Citations

Gesture Area Coverage to Assess Gesture Expressiveness and Human-Likeness

Costa Rica

GENEA: Generation and Evaluation of Non-verbal Behaviour for Embodied Agents Workshop 2024 — *2024*

- 0 Citations

Gesture Generation with Diffusion Models Aided by Speech Activity Information

Paris, France

ACM ICMI 2023 Workshop GENEA Challenge — *2023*

- 3 Citations

Diffusion-Based Approach to Style Modeling in Expressive TTS

São Paulo, Brazil

Springer Lecture Notes in Computer Science - Brazilian Conference On Intelligent Systems (BRACIS) 2022 — *2022*

- 1 Citation
- Award - 2nd Best Paper

Cooperative Spectrum Sensing Based on Skewness Statistical Tests

Florianópolis, Brazil

XXXVIII Brazilian Symposium on Telecommunications and Signal Processing — *2020*

- 4 Citations

Competitions

Kaggle: Bangla AI Speech Recognition Challenge

Online

Participant — *2023*

- Finetuned a pretrained Wav2Vec2 model on the competition dataset.
- Used data augmentation techniques to improve model robustness.
- Timbre perturbation was employed to avoid modeling speaker information.

1st The Sound of AI Hackathon: A i t m o s p h e r i c / [Demo]

Online

Participant — *2022*

- Aitmospheric project: Generating soundscapes with VQ-VAEs for compositional use and inspiration.
- Constructed an environmental scene dataset by mixing individual sounds from the ESC-50 dataset with reverb and delay effects application.
- Leveraged a VQ-VAE to learn a distribution of environmental sound scenes and generate new ones with the latent space.
- Sent the model to production by hosting the model on a Gradio app on the HuggingFace spaces and developed a website based on React.js.

Honors & Awards

Olympiads

2014 **Gold Medal**, Brazilian National Olympiad in Astronomy *Brazil*

2014 **Silver Medal**, Paraíba State Olympiad in Chemestry *Brazil*

2013 **Outstanding State Performance Silver Medal** , Brazilian Olympiads in Physics *Brazil*