Lucas Hideki Ueda

Speech AI researcher
&
PhD student in speech synthesis

+55 (19) 9 97622045
lucashueda@gmail.com
Campinas (Brazil)

CAREER OBJECTIVE

To apply AI and data analysis concepts to solve diverse problems effectively. I thrive on challenges, such as Kaggle competitions, academic challenges, and hackathons, where I can leverage my theoretical knowledge. Currently, I specialize in generative AI, with a focus on speech generation, including text-to-speech and voice conversion. Additionally, I have interest in model deployment and MLOps mainly for torch-based models.

EDUCATION

State University of Campinas

PhD in Electrical Engineering
2021 – 2025
"*Disentangling speech attributes: a representation learning approach to achieve cross-speaker style transfer*"

State University of Campinas

MSc in Electrical Engineering
2020 – 2021
"*Expressive TTS using style representation*"

State University of Campinas

B.Sc. in Applied Mathematics
2014 – 2019
"*Extreme weather events: An analysis of Campinas' climate*"

SKILLS

- Python
  - ML: pytorch, tensorflow, sklearn
  - DS: pandas, numpy, matplotlib, etc
  - Audio: librosa, torchaudio
  - Other: onnx, gradio, fastAPI, wandb
- Cloud
  - Amazon EC2/Spot, S3, GCP
- Other: SQL, Docker, Github, Linux (shell)

PROFESSIONAL EXPERIENCE

CPQD

AI researcher | 2021 - Current

- **Speech models for Brazilian Portuguese:**
  - Training state-of-the-art Text-to-Speech models applied to PT-BR data, which includes: tacotron2, glowTTS, fastspeech2, fastpitch and VITS
  - Frameworks explored: ESPNET, coqui and amphion
  - Evaluate model performance, latency and onnx compatibility
  - Explore modification in models to allow streaming
  - Tools: pytorch, docker, EC2/SPOT, shell script, onnx
- **Cross-speaker style transfer for low-resource data:**
  - Adapt SOTA TTS models to be able to model expressiveness based on style embeddings (reference encoder, GST, VAE-based, diffusion based)
  - Perform disentanglement between speaker and style information using representation learning based techniques
  - Fine-tuning and style transfer for unseen speakers
  - Synthetic data augmentation using voice conversion
  - Tools: pytorch, docker, EC2/SPOT
- **Local accent synthetic data generation:**
  - Generate synthetic speech with different local accents for a target voice to further be used to fine-tune a TTS model
  - Adapt the normalizing flows layers of a VITS-based voice conversion model to be conditioned on style and speaker information
  - Create a gradio-based front end to ease the model inference
  - Tools: pytorch, docker, EC2, gradio

SOME PUBLICATIONS

- Exploring synthetic data for cross-speaker style transfer in style representation based TTS (Syndata4GenAI 2024) [**link**]
- Spectro-ViT: Vision transformer for GABA-edited reconstruction (Magnetic Resonance Imaging 2024) [**link**]
- Efficient Hyperspectral skin reconstruction challenge (ICASSP 2024) [**link**]
- Gesture Generation in GENEA challenge (ICMI 2023) [**link**]
- Diffusion-Based Approach to Style Modeling in Expressive TTS (BRACIS 2022) [**link**]
- Spanish TTS in Blizzard Challenge 2021 [**link**]

LINKS

Github: github.com/lucashueda
Scholar: bit.ly/4bOrYXQ
Linkedin: linkedin.com/in/lucashueda/

PROFESSIONAL EXPERIENCE

RECOD.ai

PhD student | 2023 - Current

- **Speech synthesis:**
  - Training TTS models: Gan-based (VITS), Flow-Matching-based (MatchaTTS) and internal-FastPitch
  - Training Voice Conversion models: sovits, softVC, KNN-VC, internal-RecursiveVC
  - Build TTS, VC and ASR demos using gradio to present in internal events
  - Tools: pytorch, multi-gpu/DDP, docker, gradio, wandb
- **GABA-edited reconstruction challenge:**
  - Experiments with different vision models (resnet, wideresnet) in an signal reconstruction task
  - Build a wandb based pipeline to track experiments
  - Tools: pytorch, timm, docker, wandb
- **Efficient hyperspectral skin reconstruction challenge:**
  - Train and modify baseline model incorporating pre-trained vision models
  - Explore loss function modifications to improve results (channel-weighted loss, focal loss)
  - Build the experimentation pipeline based on config oriented training script and experiment track recording on wandb
  - Tools: pytorch, docker, wandb
- **BirdCLEF 2024:**
  - Implement and explore domain adaptation techniques such as gradient reversal layer and wasserstein distance loss
  - Prepare training and inference notebooks to run on kaggle platform
  - Optimize inference using parallel processing on both data preprocessing and model inference
  - Export trained models using onnx to speedup inference time
  - Tools: pytorch, timm, onnx

Itaú Unibanco

Data scientist | 2018 a 2020

- **Model validation**
  - Perform statistical analysis to evaluate data consistency and credit risk model quality
  - Tools: jupyter, sql, SAS, sklearn, pandas, R
- **Document retrieval with text processing**
  - Build a internal document retrieval based on generating text embeddings and retrieving through cosine similarity
  - Tools: NLTK, sklearn, spacy, pandas