

Dongyao Zhu

doz022@ucsd.edu | <https://dongyaozhu.github.io>

EDUCATION

Bachelor of Science, Computer Science	06/2016 – 06/2020
Bachelor of Arts, Economics	06/2016 – 06/2020
University of California, San Diego	

RESEARCH INTERESTS

I am interested in advancing and landing scalable, reliable, and efficient AI-assisted productivity:

- **Generative AI:** controllable, human-aligned, interactive generative models
- **Speech, Music, Audio:** generative and comprehensive tasks, such as vocoders and representation learning
- **Efficient and Reliable AI:** knowledge & data distillation, decision confidence calibration
- **Theoretical Foundations for Probabilistic Modeling:** sampling methods in high-dimensional domains

PUBLICATIONS

Rethinking Data Distillation: Do Not Overlook Calibration

Submitted to ICCV 2023

Dongyao Zhu, Bowen Lei, Jie Zhang, Yanbo Fang, Yiqun Xie, Ruqi Zhang, Dongkuan Xu

Efficient Informed Proposals for Discrete Distributions via Newton's Series Approximation

International Conference on Artificial Intelligence and Statistics (AISTATS), 2023

Yue Xiang, **Dongyao Zhu**, Bowen Lei, Dongkuan Xu, Ruqi Zhang

PROFESSIONAL EXPERIENCE

Huawei Technologies Co., Ltd. , Distributed Data Lab, 2012 Labs	04/2021 - now
Machine Learning Engineer & Software Engineer	
Phonetic AI, Inc.	06/2020 – 02/2021
Machine Learning Engineer	

ACADEMIC SERVICES

Reviewer

International Workshop on Resource-Efficient Learning for Knowledge Discovery	2023
---	------

TEACHING

University of California, San Diego

Tutor, CSE 130, principles and paradigms of programming languages	07/2019 - 09/2019
---	-------------------

PROJECTS

Faster FastSpeech2: co-design with MindSpore Static Graphs to speed up training from 90s to 8s per step
TensorFlow Speech Command Challenge: 17/1314 (top 1.29%) on Kaggle, speech recognition + language model
Real Time Foreign Accent Conversion: speech recognition + speaker verification + voice cloning + vocoder
MindSpore Audio: a ready-to-use development toolkit for audio research
Masked Temperature Scaling: SOTA calibration technique for neural networks trained on distilled datasets
MANA: efficient discrete sampler in high-dimensional space, applied to image retrieval & text summarization
MindSpore: an open-source deep learning framework for both daily and large-scale uses
Embedded Anomaly Detection: automatic detection & handling of traffic anomaly in network routers

SKILLS

Python, C, C++, Haskell, OCaml, HTML & CSS & JavaScript, Assembly, Shell, PyTorch, TensorFlow, MindSpore

AWARDS

Provost Honors, Revelle College	2018
---------------------------------	------