Generative AI for Music and Audio

Hao-Wen (Herman) Dong

Department of Performing Arts Technology
School of Music, Theatre & Dance
University of Michigan
hermandong.com

March 4, 2026

ILLIAC & ILLIAC Suite @ UIUC!

Photos taken on July 15, 2025

About Me

National Taiwan University
B.S. in Electrical Engineering

2013 – 2017

2017 – 2019 — ACADEMIA SINICA
Research Assistant

Summer 2019 — YAMAHA
Research Intern

UC San Diego
M.S. in Computer Science

2019 – 2021

Summer 2021 — Dolby
Deep Learning Audio Intern

Summer 2022 — SONY
Student Intern

Fall 2022 — amazon
Applied Scientist Intern

Winter 2023 — Dolby
Speech/Audio Deep Learning Intern

Summer 2023 — Adobe
Research Scientist/Engineer Intern

Fall 2023 — NVIDIA
Research Intern

UC San Diego
Ph.D. in Computer Science

2019 – 2024

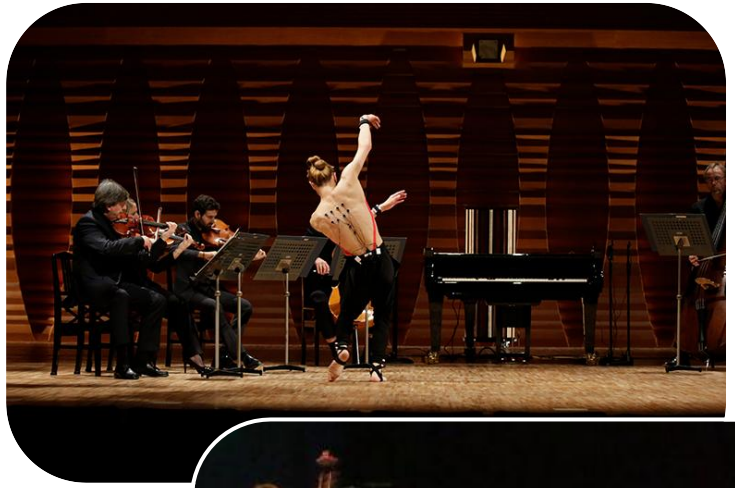
SCHOOL OF MUSIC, THEATRE & DANCE
PERFORMING ARTS TECHNOLOGY
UNIVERSITY OF MICHIGAN

Music & Technology Co-evolves

Hildegard Dodel, Public domain, via Wikimedia Commons.
Taken at Hamamatsu Museum of Musical Instruments, August 2019.
yan, CC BY-SA 4.0, via Wikimedia Commons.

Music & AI

(Source: Yamaha)



(Source: Sankei Shimbun)

(Source: Robot Gizmos)

(Source: NBC DFW)

yamaha.com/en/news_release/2018/18013101/
sankei.com/article/20240113-CQCOSQHJWFIYPJJKZDCITRTRVI/
roboticgizmos.com/shimon-musical-robot-deep-learning/
nbcdfw.com/entertainment/the-scene/how-verdigris-ensemble-is-using-ai-to-create-a-new-concert-experience/3366031/

Art challenges Technology

Creativity

Augmenting Human Creativity with AI

AI

Technology inspires the Art

Generative AI for Music, Audio & Video Creation

(Source: UploadVR)

(Source: The Denver Post)

(Source: Descript)

(Source: Daily Bruin)

(Source: Wikimedia Commons)

Universitaetsmedizin, CC BY-SA 4.0, via Wikimedia Commons
uploadvr.com/iron-man-vr-breaks-free-from-cords-load-screens-on-quest-2/
descript.com/blog/article/what-is-the-best-audio-interface-for-recording-a-podcast
denverpost.com/2019/08/02/colorado-symphony-movie-scores-harry-potter-star-wars/
dailybruin.com/2023/08/04/theater-review-the-musical-les-misrables-offers-stellar-displays-and-impassioned-vocals

Augmenting Human Creativity with AI

- **Novel Generative Models for New Domains**
  - **Multitrack music generation** (AAAI 2018, ISMIR 2018, ISMIR 2020, ICASSP 2023, ISMIR 2024), **text-to-music generation** (ISMIR 2025), **video-to-music generation** (ISMIR 2025), **symbolic music processing tools** (ISMIR LBD 2019, ISMIR 2020)
- **AI-assisted Tools for Content Creation**
  - **Violin performance synthesis** (ICASSP 2022, ICASSP 2025), **music instrumentation** (ISMIR 2021), **music arrangement** (AAAI 2018), **music harmonization** (JNMR 2020)
- **Multimodal Generative Models for Content Creation**
  - **Long-to-short video editing** (ICLR 2025, NeurIPS 2025), **text-queried sound separation** (ICLR 2023), **text-to-audio synthesis** (WASPAA 2023)

icassp 2022 Singapore

Deep Performer: Score-to-Audio Music Performance Synthesis

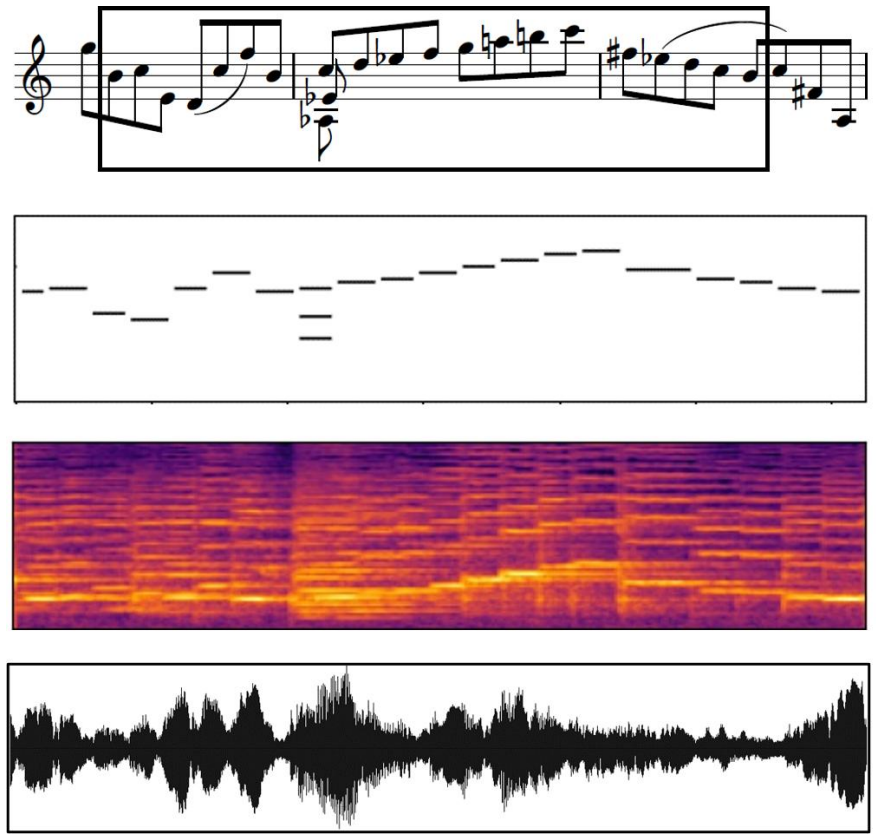
Hao-Wen Dong[1,2] Zhou Cong[1] Taylor Berg-Kirkpatrick[2] Julian McAuley[2]

[1] Dolby Laboratories [2] University of California San Diego




Dolby UC San Diego

Score-to-Audio Synthesis



Raw score

↓ *Alignment model*

Aligned score

↓ *Synthesis model*

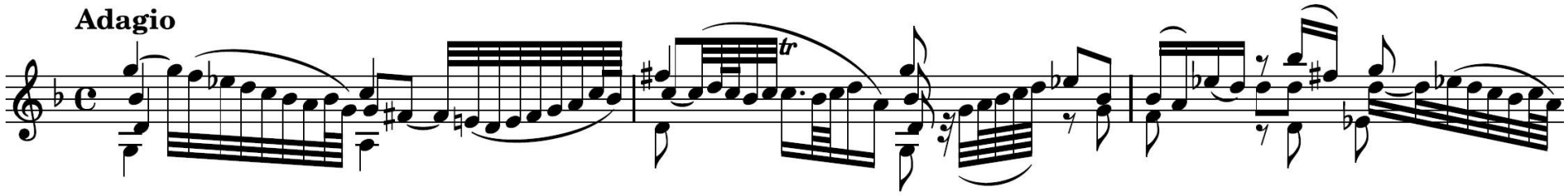
Mel spectrogram

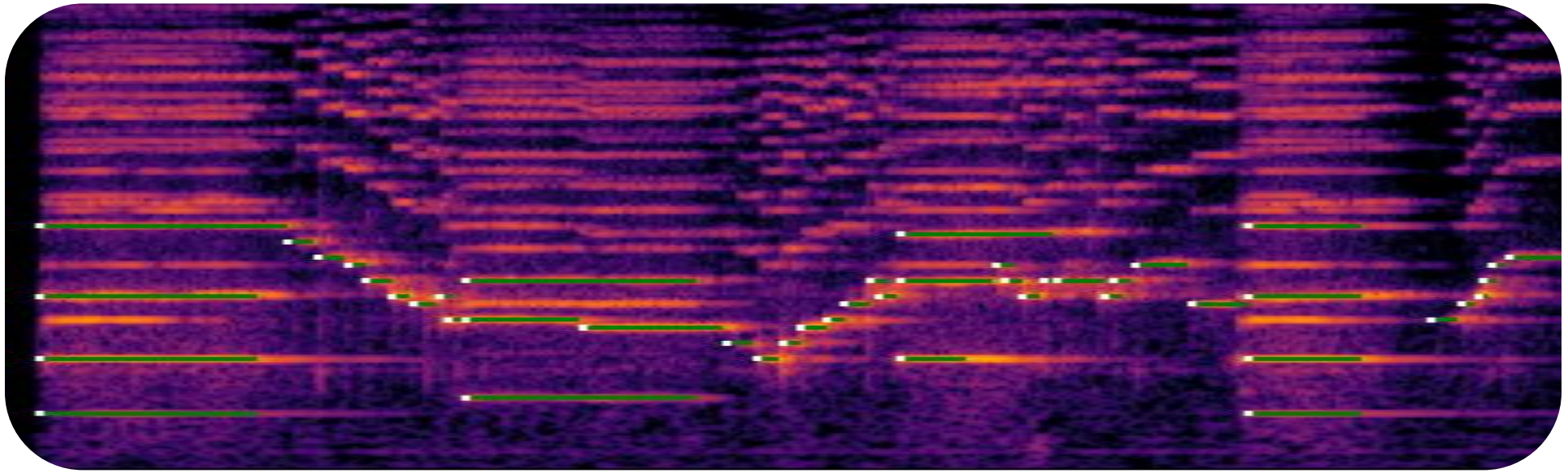
↓ *Inversion model*

Waveform

Hao-Wen Dong, Cong Zhou, Taylor Berg-Kirkpatrick, and Julian McAuley, "Deep Performer: Score-to-Audio Music Performance Synthesis," *ICASSP*, 2022.

Score-to-Audio Synthesis



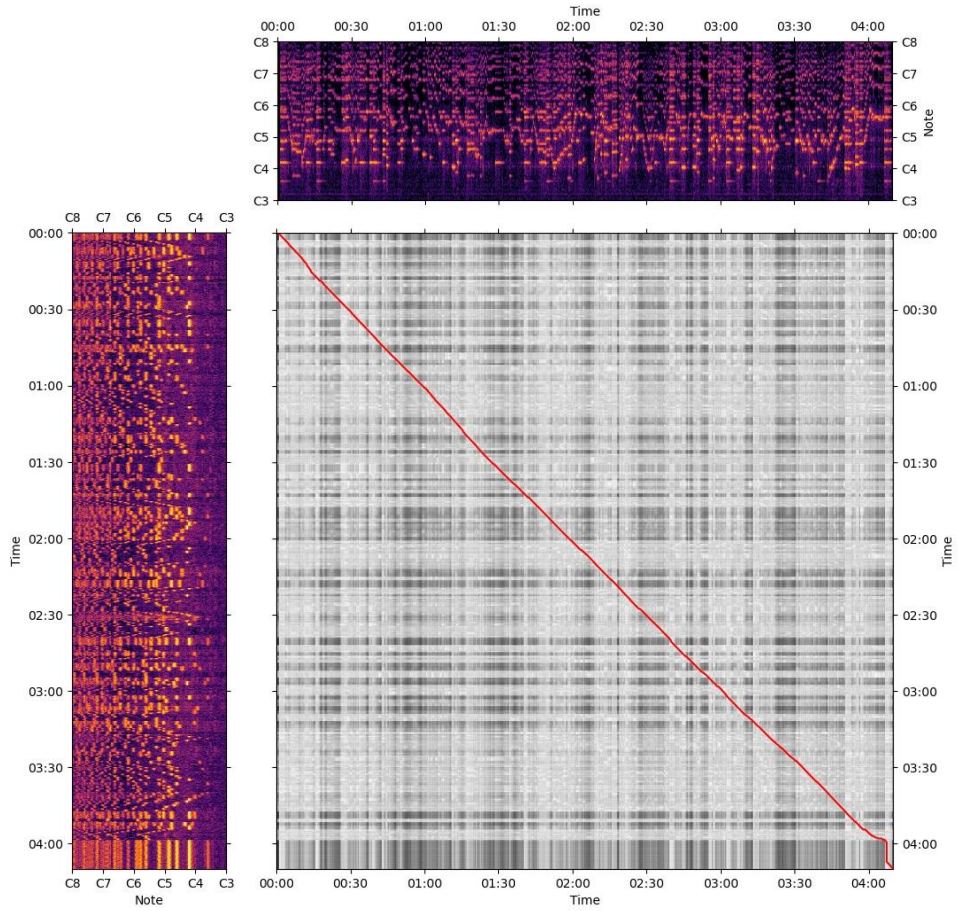


Hao-Wen Dong, Cong Zhou, Taylor Berg-Kirkpatrick, and Julian McAuley, "Deep Performer: Score-to-Audio Music Performance Synthesis," *ICASSP*, 2022.

Aligning Performances to Musical Scores

Performance recording

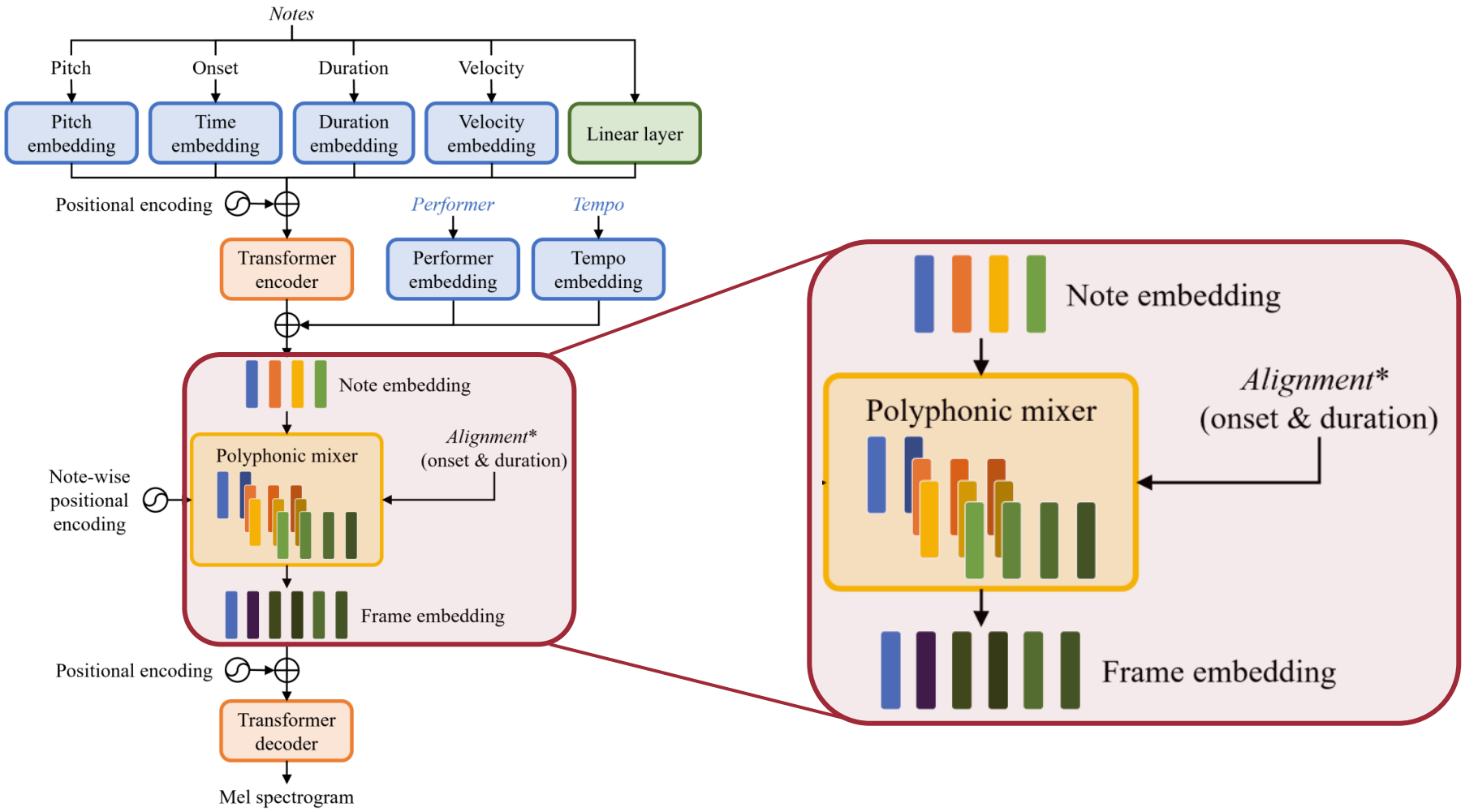
Synthesized audio



Dynamic time warping!

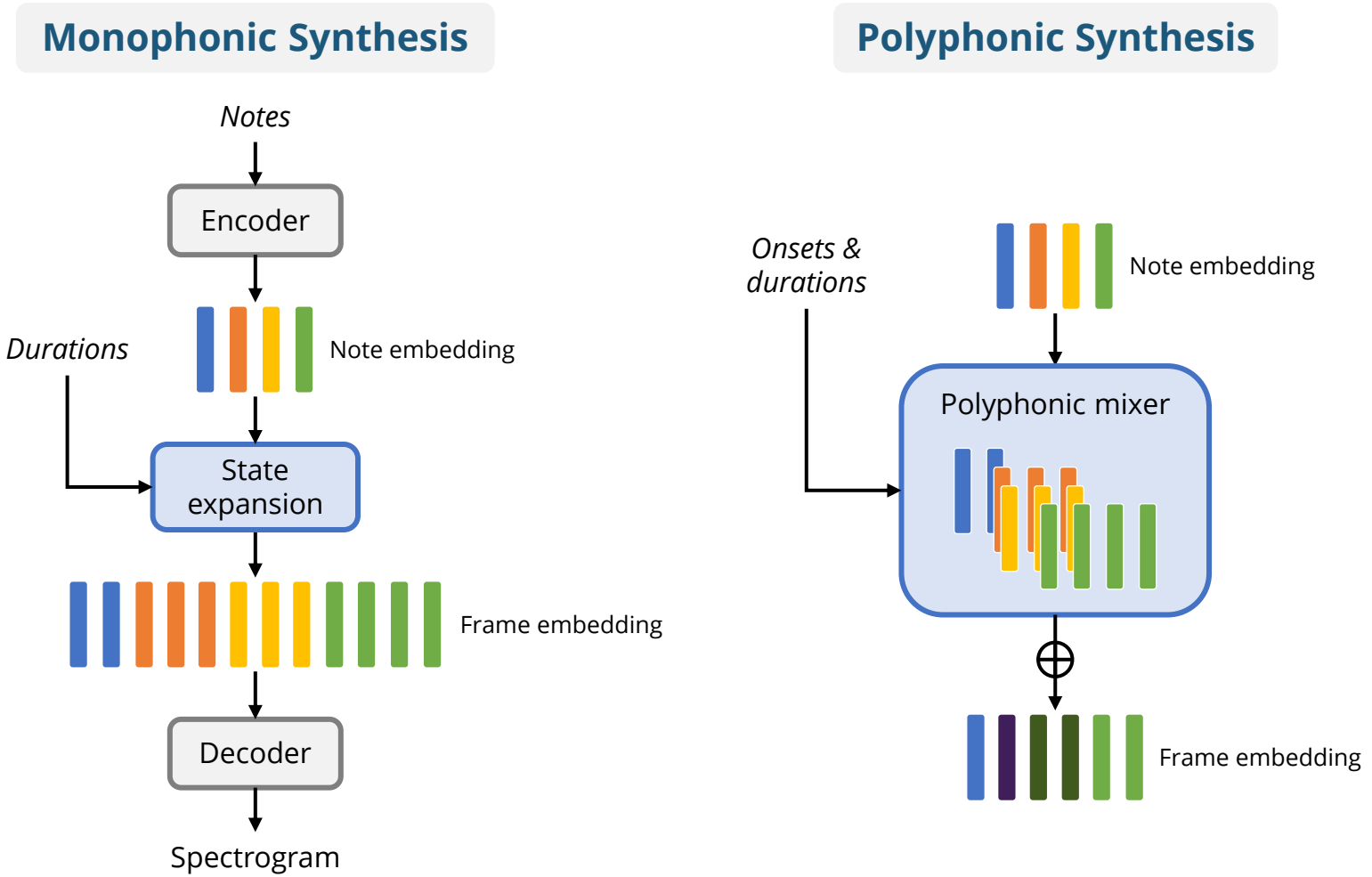
Hao-Wen Dong, Cong Zhou, Taylor Berg-Kirkpatrick, and Julian McAuley, "Deep Performer: Score-to-Audio Music Performance Synthesis," *ICASSP*, 2022.

A TTS-based Model for Score-to-Audio Synthesis



Hao-Wen Dong, Cong Zhou, Taylor Berg-Kirkpatrick, and Julian McAuley, "Deep Performer: Score-to-Audio Music Performance Synthesis," *ICASSP*, 2022.

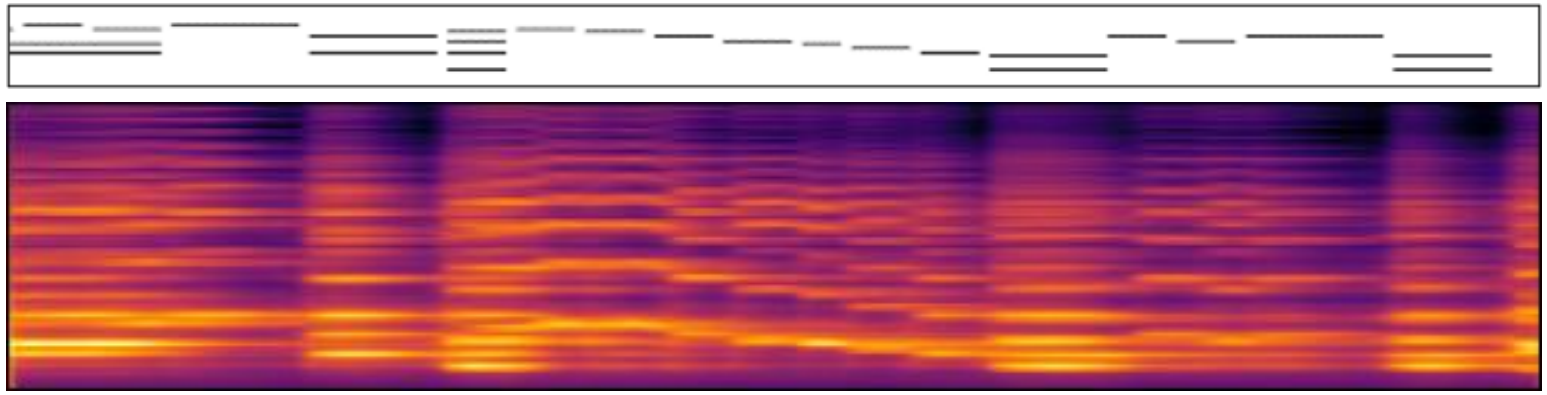
Handling Polyphony



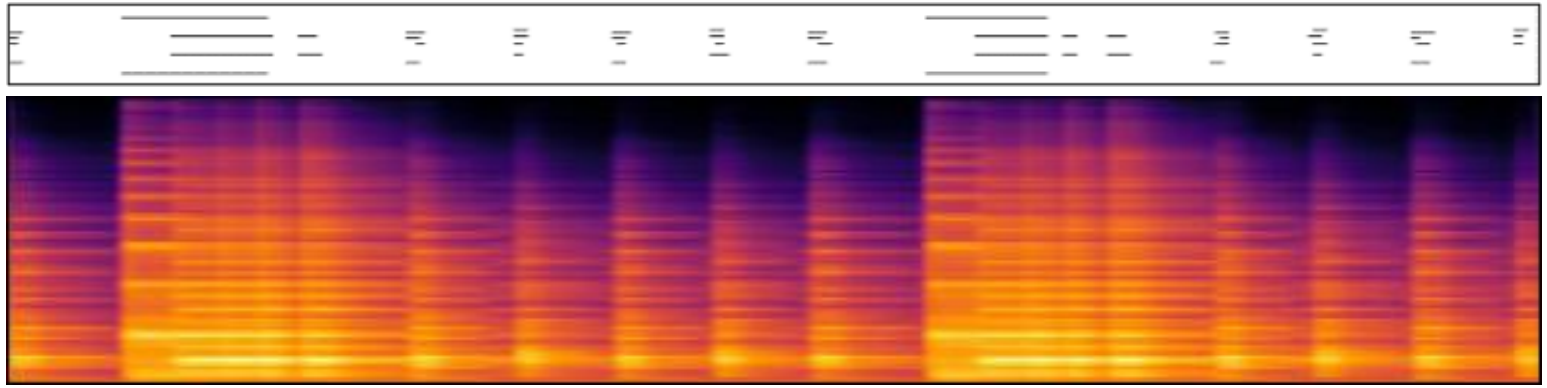
Hao-Wen Dong, Cong Zhou, Taylor Berg-Kirkpatrick, and Julian McAuley, "Deep Performer: Score-to-Audio Music Performance Synthesis," *ICASSP*, 2022.

Examples Results

Violin



Piano



hermandong.com/deepperformer

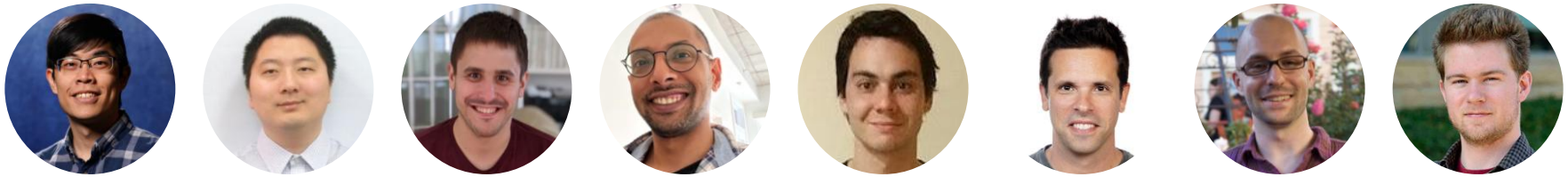
Hao-Wen Dong, Cong Zhou, Taylor Berg-Kirkpatrick, and Julian McAuley, "Deep Performer: Score-to-Audio Music Performance Synthesis," *ICASSP*, 2022.

IEEE WASPAA 2023

CLIPSonic: Text-to-Audio Synthesis with Unlabeled Videos and Pretrained Language-Vision Models

Hao-Wen Dong[1,2] Xiaoyu Liu[1] Jordi Pons[1] Gautam Bhattacharya[1]
Santiago Pascual[1] Joan Serrà[1] Taylor Berg-Kirkpatrick[2] Julian McAuley[2]

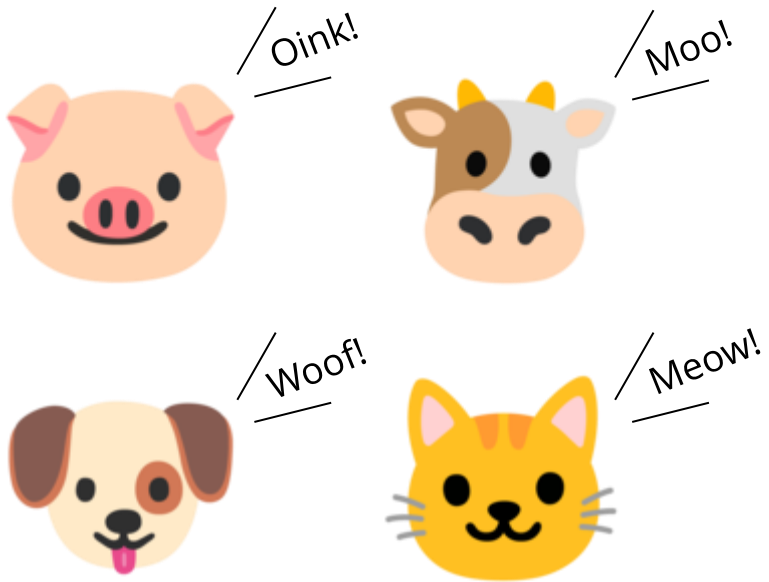
[1] Dolby Laboratories [2] University of California San Diego



Dolby UC San Diego

Learning Sounds from Observations

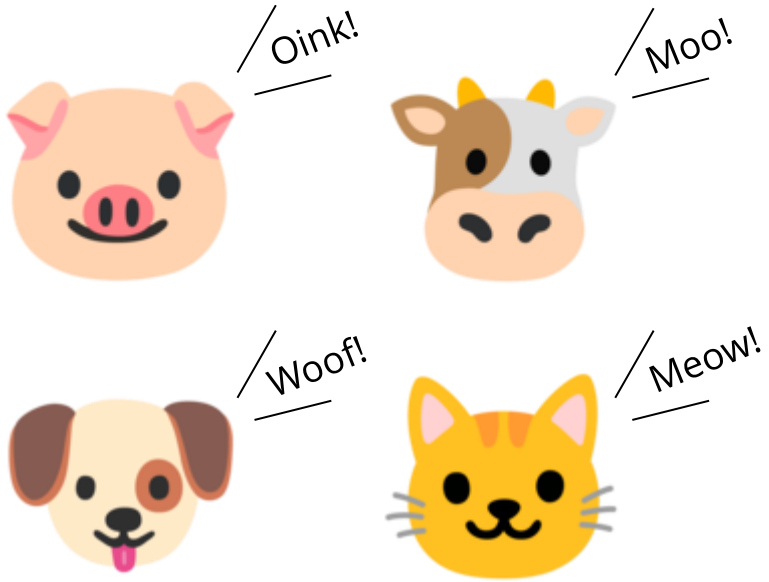
- Watching a dog barking, humans can easily **associate the barking sound to the dog**



What does the fox say?

Learning Sounds from Noisy Videos

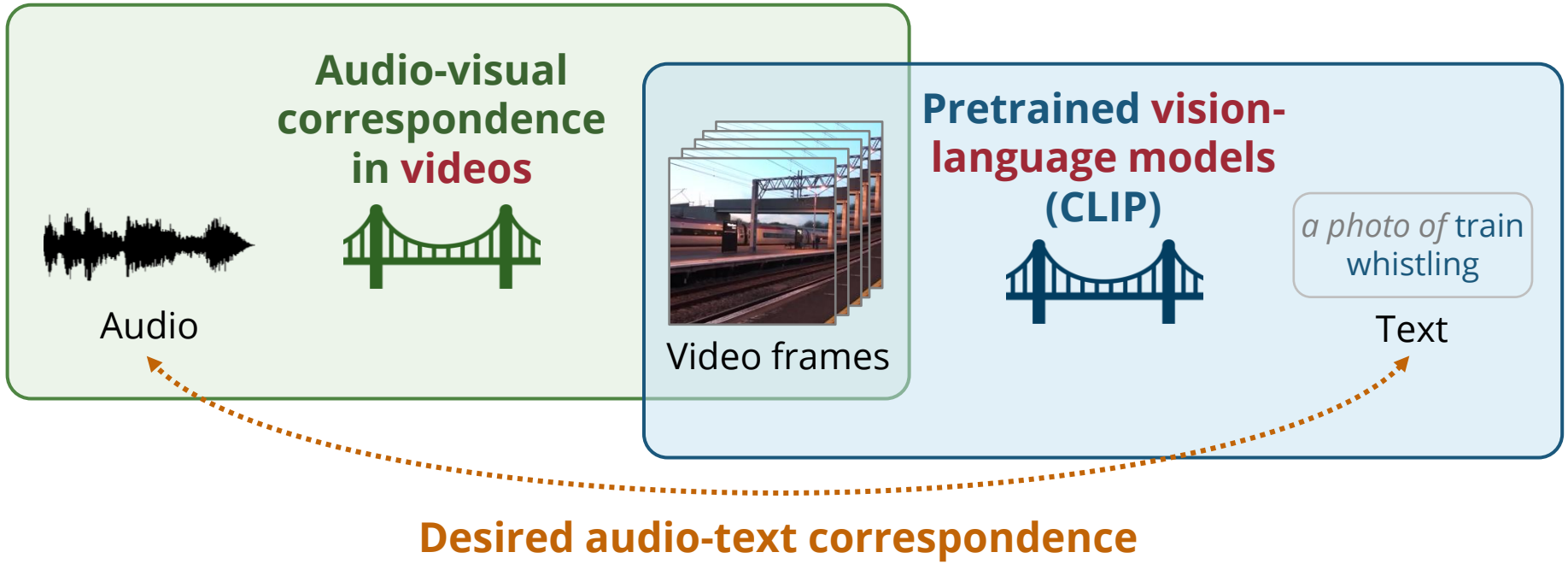
- Watching a dog barking, humans can easily **associate the barking sound to the dog**



Can machines learn to synthesize sounds from watching *noisy* videos?

URURU TV, "Excited Fox Noises - Fox Laughing and Squeaking," *YouTube*, 2019.

Leveraging the Visual Domain as a Bridge



No text-audio pairs required!

Scalable to large video datasets!

Data

VGGSound

(Chen et al., 2020)

Hedge trimmer running

Dog bow-wow

Bird chirping, tweeting

Noisy videos with diverse sounds

(172K videos, 310 classes)

MUSIC

(Zhao et al., 2018)

Violin

Acoustic guitar

Accordion

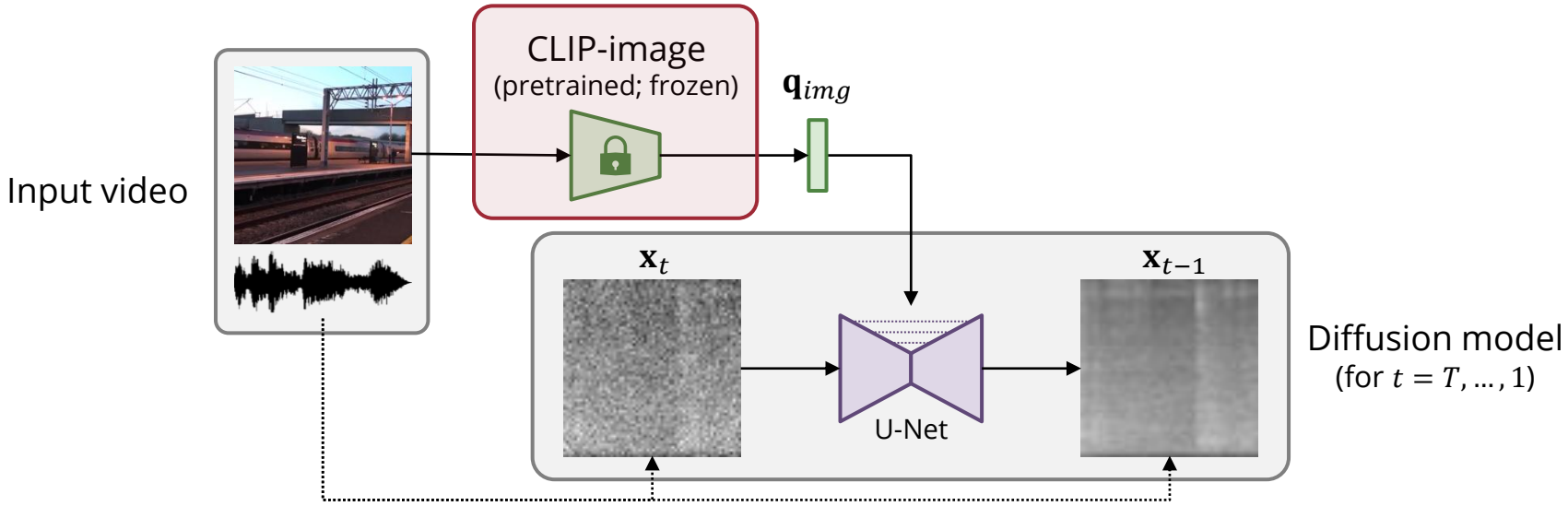
Music instrument playing videos

(1,055 videos, 21 instruments)

Hang Zhao, Chuang Gan, Andrew Rouditchenko, Carl Vondrick, Josh McDermott, and Antonio Torralba, "The Sound of Pixels," *ECCV*, 2018.
Honglie Chen, Weidi Xie, Andrea Vedaldi, and Andrew Zisserman, "VGGSound: A Large-Scale Audio-Visual Dataset," *ICASSP*, 2020.

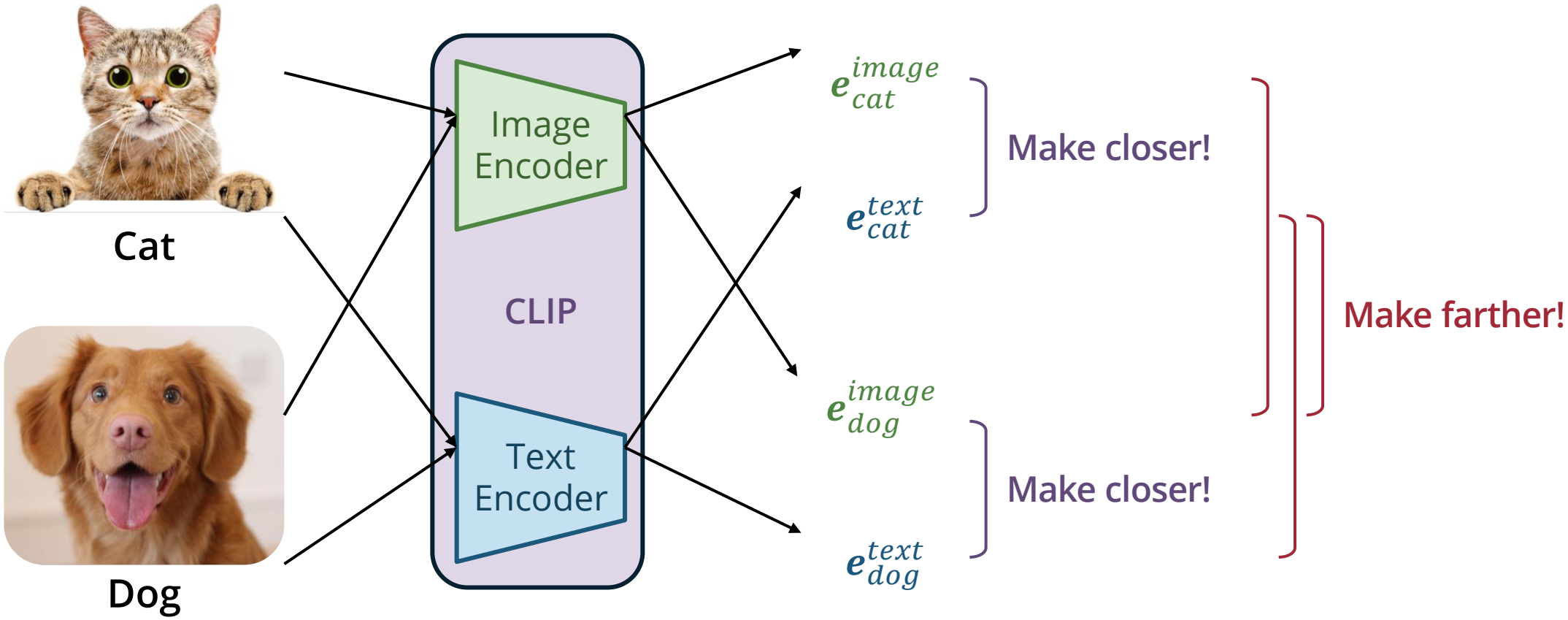
Training an Image-to-Audio Synthesis Model

- First train an **image-to-audio** synthesis model



CLIP (Contrastive Language-Image Pretraining)

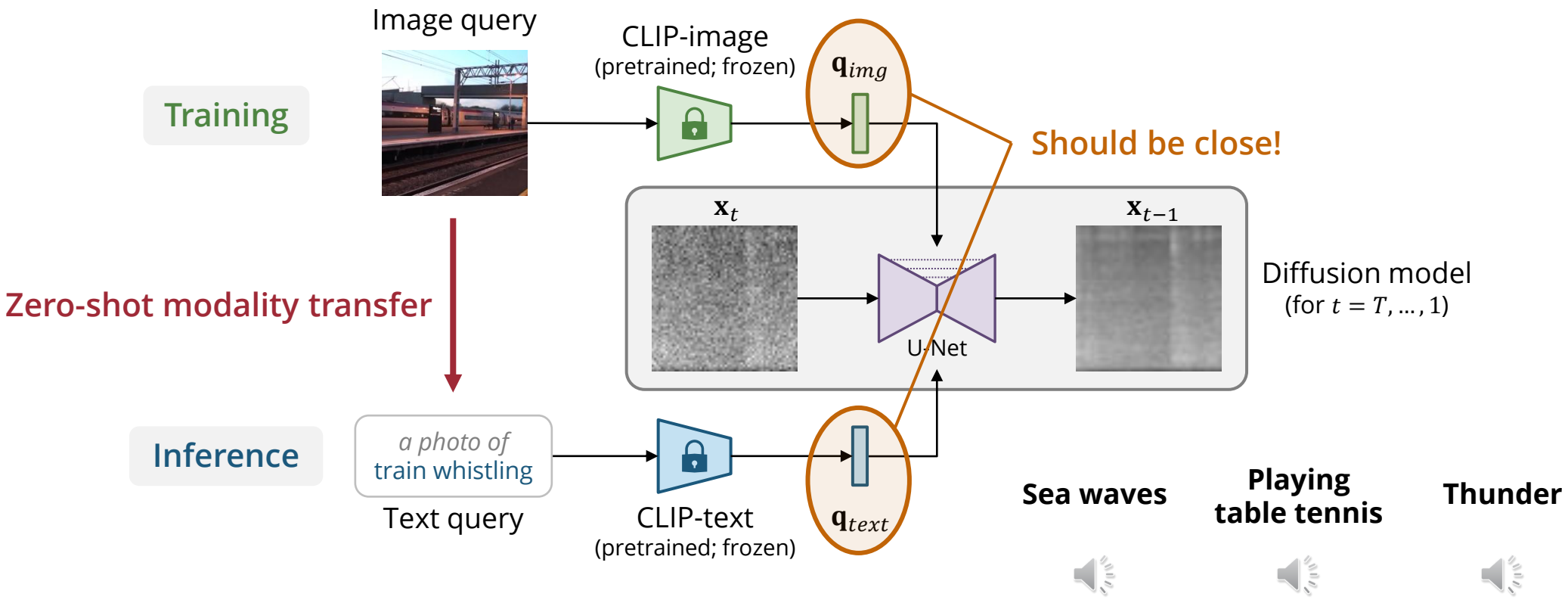
- Learn a shared embedding space for images and texts via contrastive learning



Radford et al., "Learning Transferable Visual Models From Natural Language Supervision," *ICML*, 2021.

Inference: Zero-shot Modality Transfer

- Switch to a **pretrained CLIP-text encoder** for text-to-sound synthesis



Inference: Zero-shot Modality Transfer

- Switch to a **pretrained CLIP-text encoder** for text-to-sound synthesis

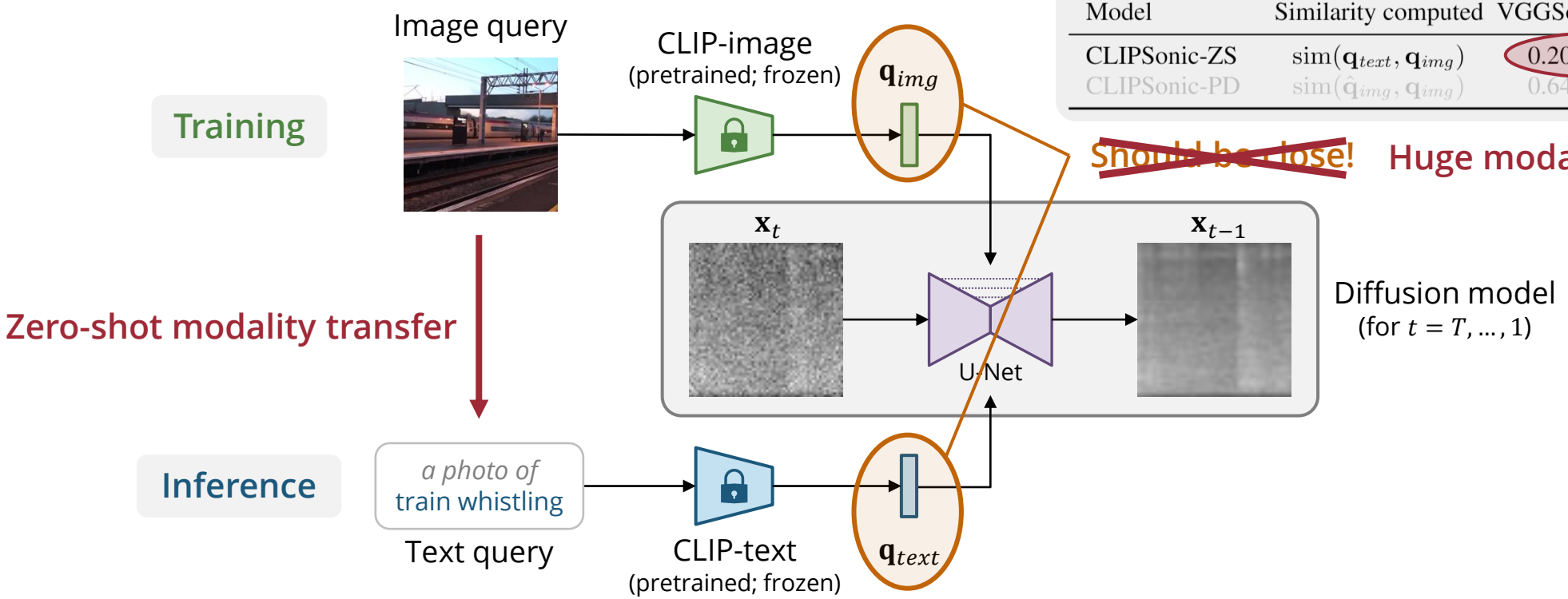


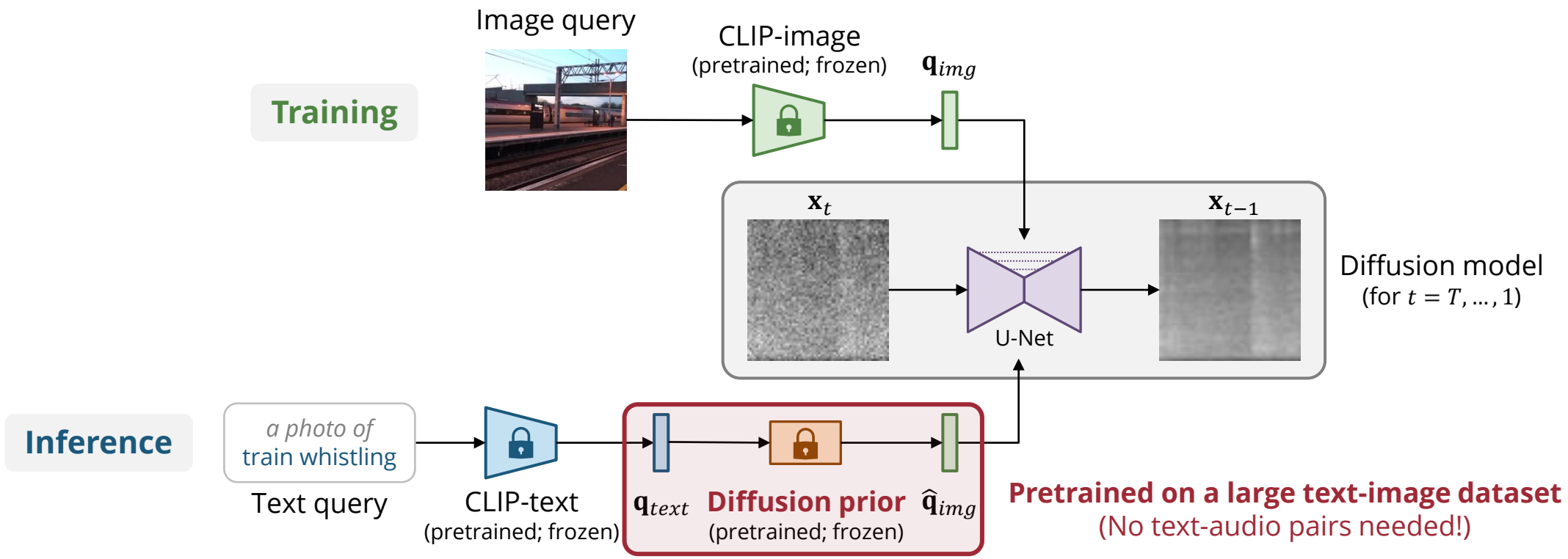
Table 2: Cosine similarities between various query embeddings.

Model	Similarity computed	VGGSound	MUSIC
CLIPSonic-ZS	$\text{sim}(\mathbf{q}_{text}, \mathbf{q}_{img})$	0.205	0.245
CLIPSonic-PD	$\text{sim}(\hat{\mathbf{q}}_{img}, \mathbf{q}_{img})$	0.647	0.720

~~Should be close!~~ Huge modality gap!

Leveraging Diffusion Prior to Close the Modality Gap

- Adopt a **pretrained diffusion prior model** to reduce the modality gap



Ramesh et al., "Hierarchical Text-Conditional Image Generation with CLIP Latents," *arXiv preprint arXiv:2204.06125*, 2022.

Example Text-to-Audio Synthesis Results

Rapping

Sea waves

Thunder

Smoke detector beeping

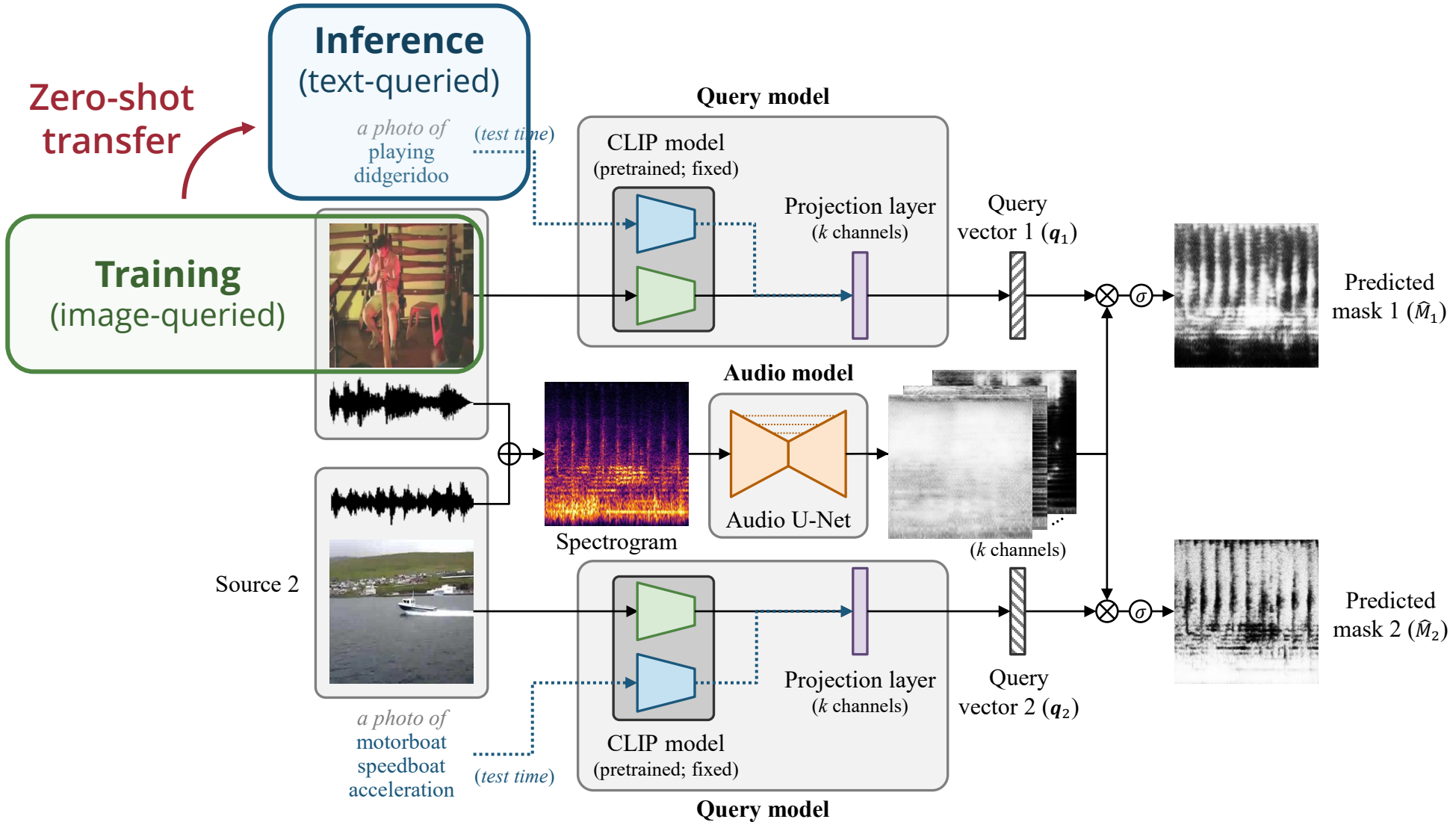
Playing table tennis

Playing violin fiddle

Example Image-to-Audio Synthesis Results (Out-of-distribution)

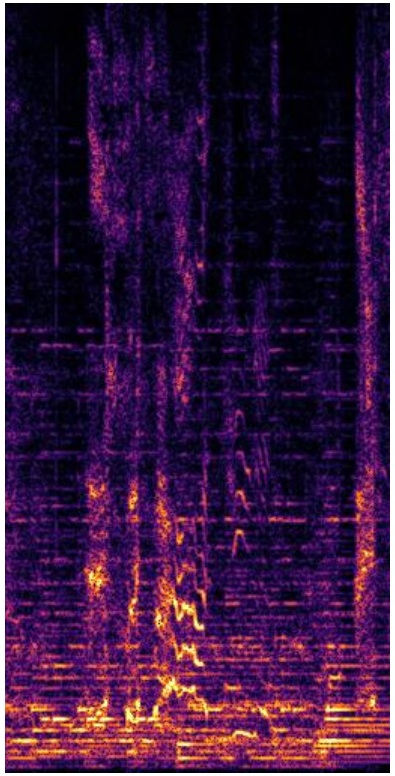
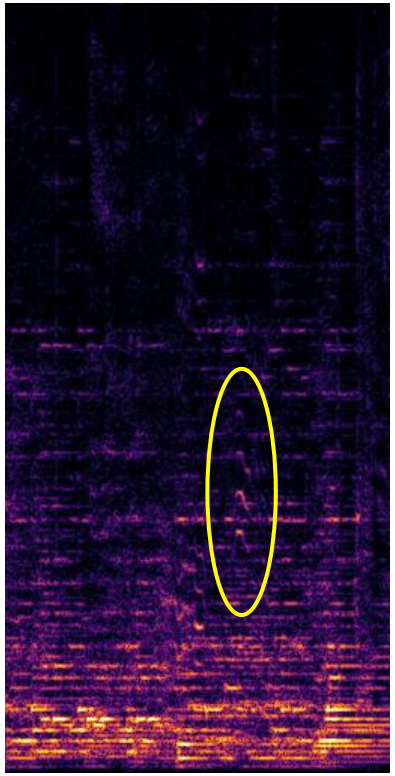
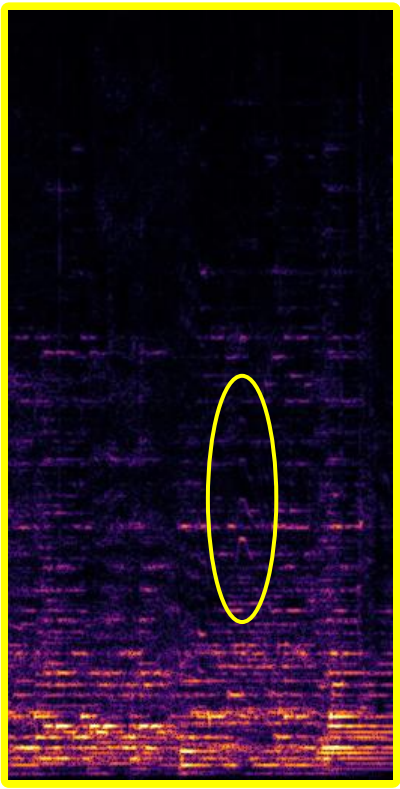
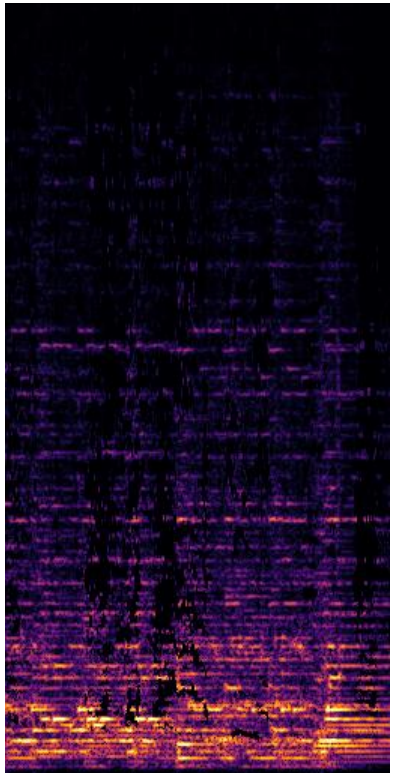
(Then!) State-of-the-art image-to-audio synthesis performance!

CLIPSep: Text-queried Sound Separation (ICLR 2023)



CLIPSep: Text-queried Sound Separation (**ICLR 2023**)

Query: "*playing harpsichord*"

Mixture	**CLIPSep**	**CLIPSep-NIT**	**Ground truth**
			

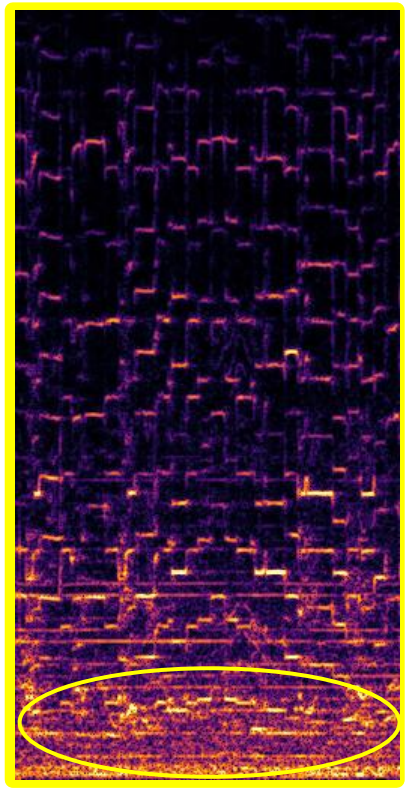
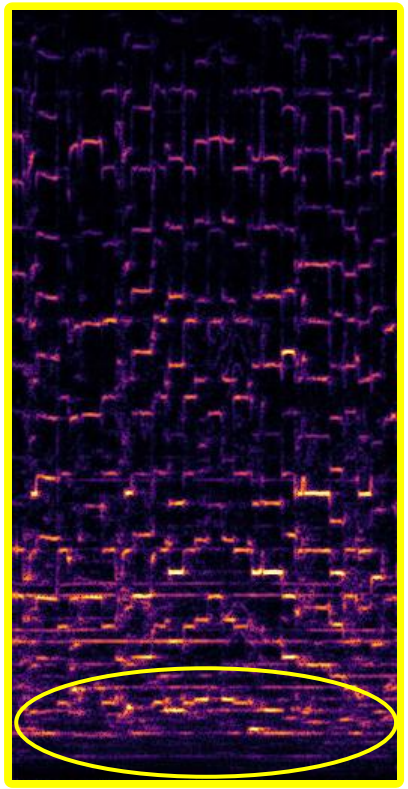
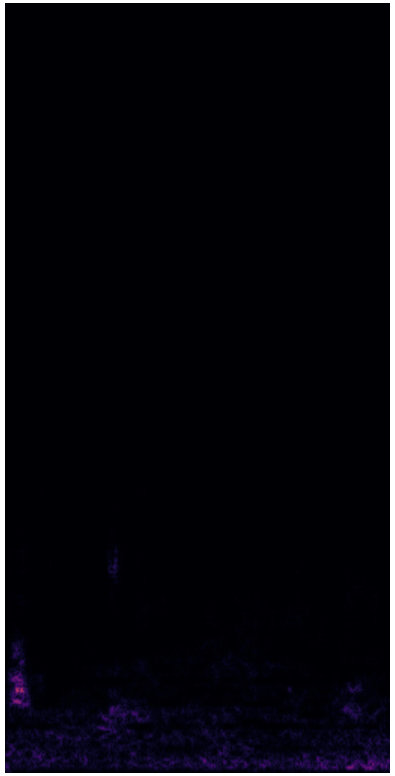
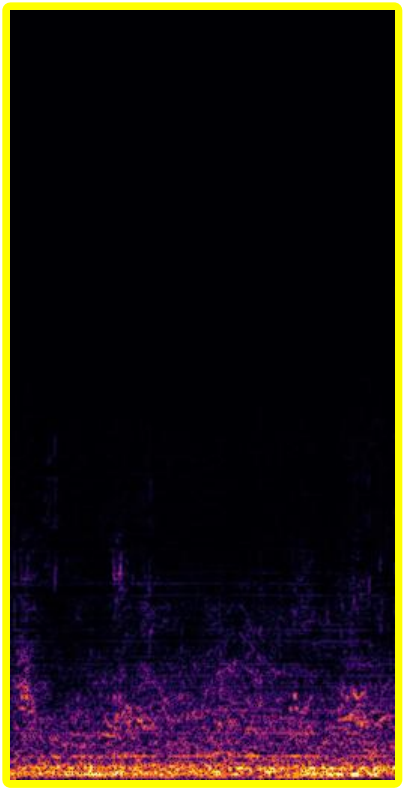
CLIPSep: Noise Removal (**ICLR 2023**)





SONY UC San Diego

Query: "*playing bagpipe*"

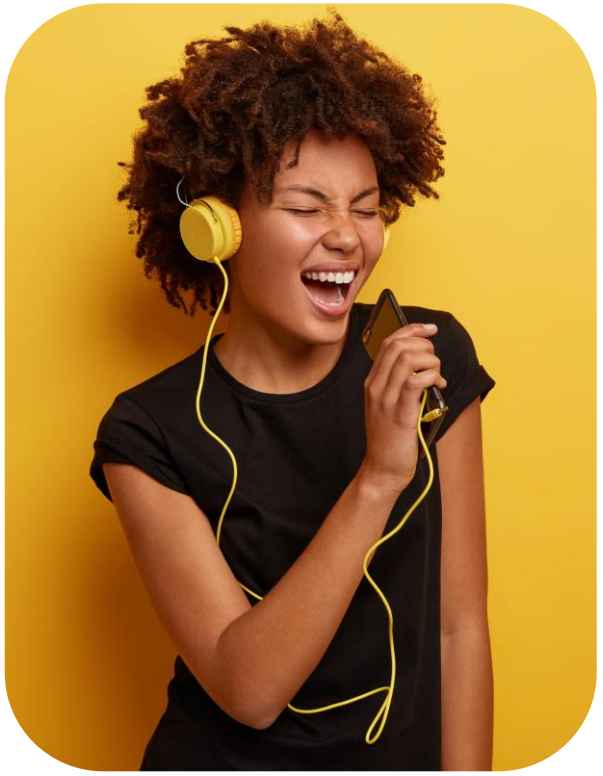
Mixture	Prediction	Noise head 1	Noise head 2
			

🔥 Ongoing Work: A Cappella Vocal Coach 🔥




How can AI Support A Cappella Singers?

Seagull-K from Hsinchu, Taiwan



How can we best support a novice a cappella singer in practicing their singing skills?

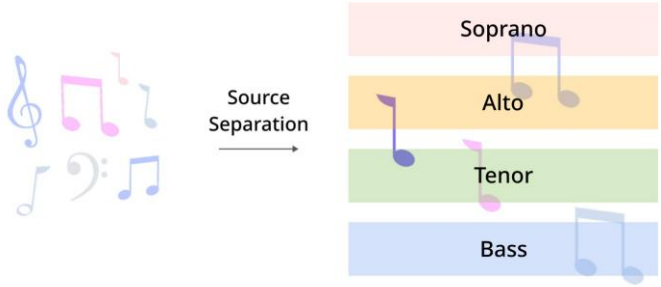
Formative Study

- Interviewed **12 a cappella singers**
  - Includes both beginners and professionals

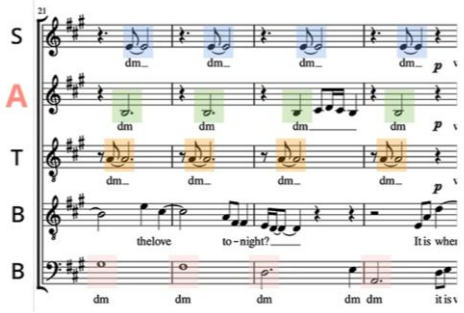
- **Four design goals** identified
  - **Simulating an authentic group singing context** during individual practice
  - **Offering intuitive, context-aware feedback** on the users' recordings
  - **Assisting a sensemaking of the "big picture"** (patterns and links between parts)
  - **Placing human creativity at the center** as an assistive tool

AcaMate Design

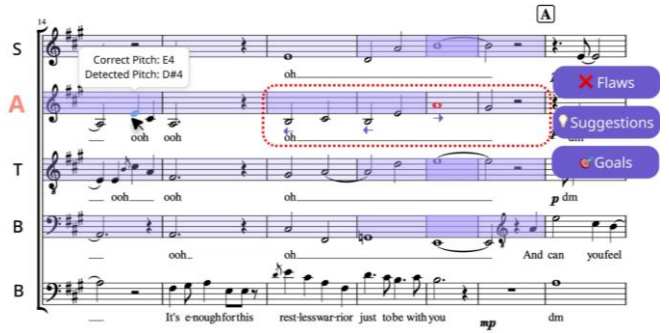

Separating voice parts



Highlighting musical patterns



Providing visual feedback
(on pitch, rhythm, dynamics)



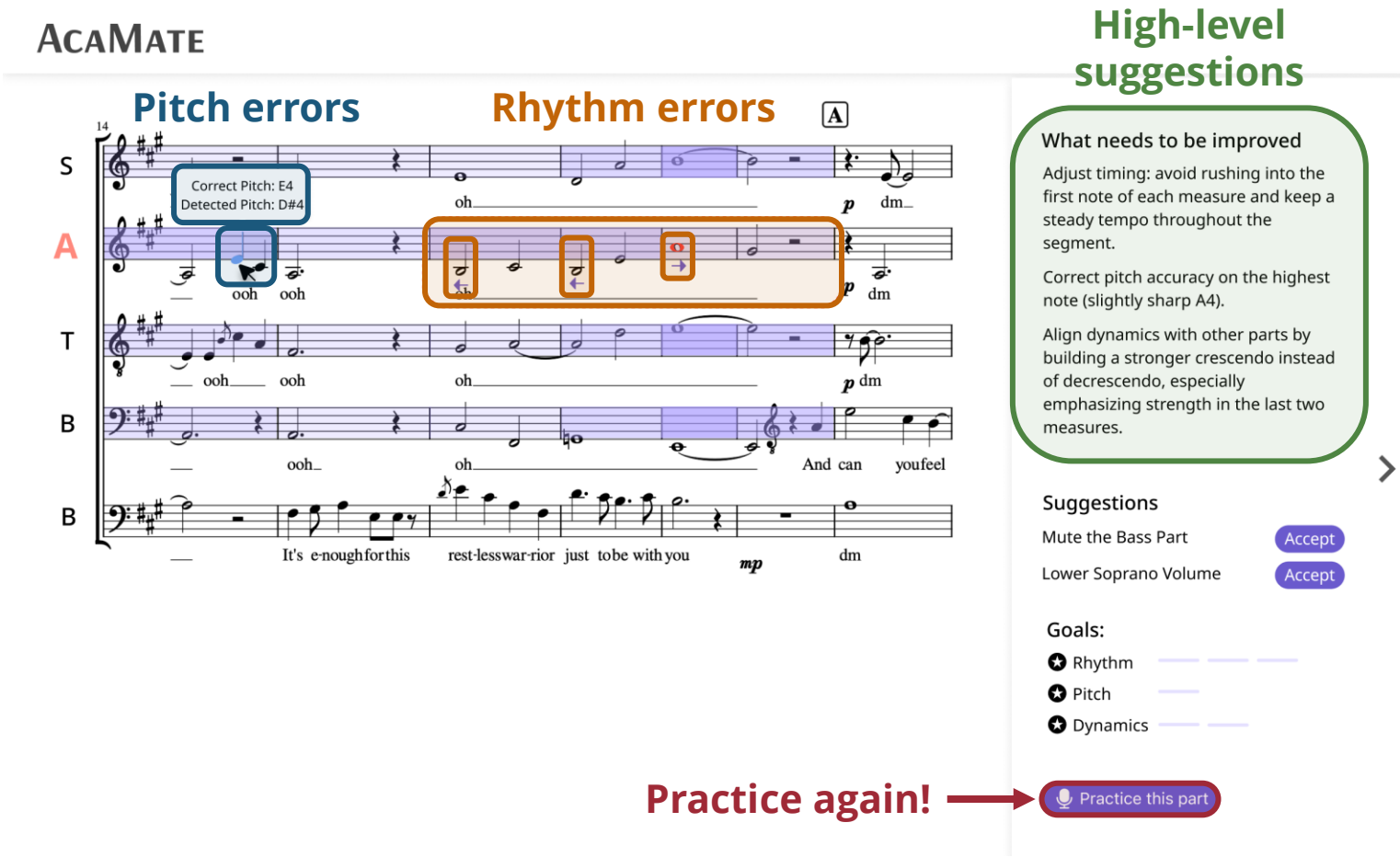
A Before Singing → B During Singing → C After Singing

System-Assisted Deliberate Practice

D Iterative Singing

Kexin Phyllis Ju, Ting-Yu Pan, and Hao-Wen Dong, "AcaMate: Supporting Novice A Cappella Singers in Iterative Individual Practice," *UIST Posters*, 2025.

AcaMate Design

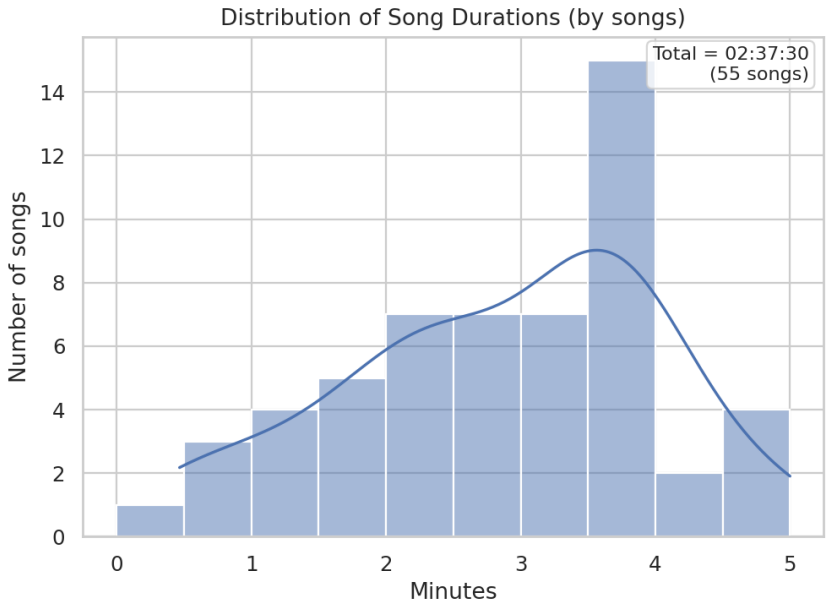



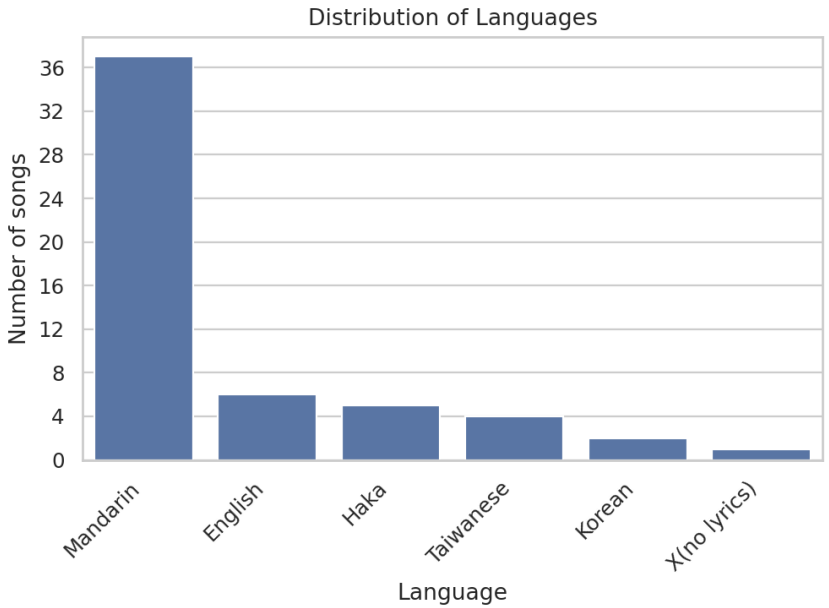
Kexin Phyllis Ju, Ting-Yu Pan, and Hao-Wen Dong, "AcaMate: Supporting Novice A Cappella Singers in Iterative Individual Practice," *UIST Posters*, 2025.

AcaMate Prototype

ACappellaSet: Studio Recordings with Stems

- **55** studio-quality a cappella songs **with stems** performed by 3 groups
- **2.6 hours** in total
- **Five languages**: Mandarin, English, Hakka, Taiwanese, and Korean





Ting-Yu Pan*, Kexin Phyllis Ju*, and Hao-Wen Dong, "Curating an A Cappella Dataset for Source Separation," *ISMIR Late-Breaking Demos*, 2025. (*equal contribution)

A Cappella Source Separation

Model	VP	Other	All
Pretrained (official)	5.22	10.66	7.94
Pretrained (drum)	3.66	9.24	6.45
Fine-tuned (ours)	**7.62**	**11.63**	**9.62**

+2.4 dB **+1 dB**

	Vocal percussion	SATB		Vocal percussion	SATB
Pretrained			**Pretrained**		
Finetuned			**Finetuned**		
Ground truth			**Ground truth**		

Ting-Yu Pan*, Kexin Phyllis Ju*, and Hao-Wen Dong, "Curating an A Cappella Dataset for Source Separation," *ISMIR Late-Breaking Demos*, 2025. (*equal contribution)

🔥 Ongoing Work: AI-assisted Film Scoring 🔥





Accessibility of Music-Video Datasets

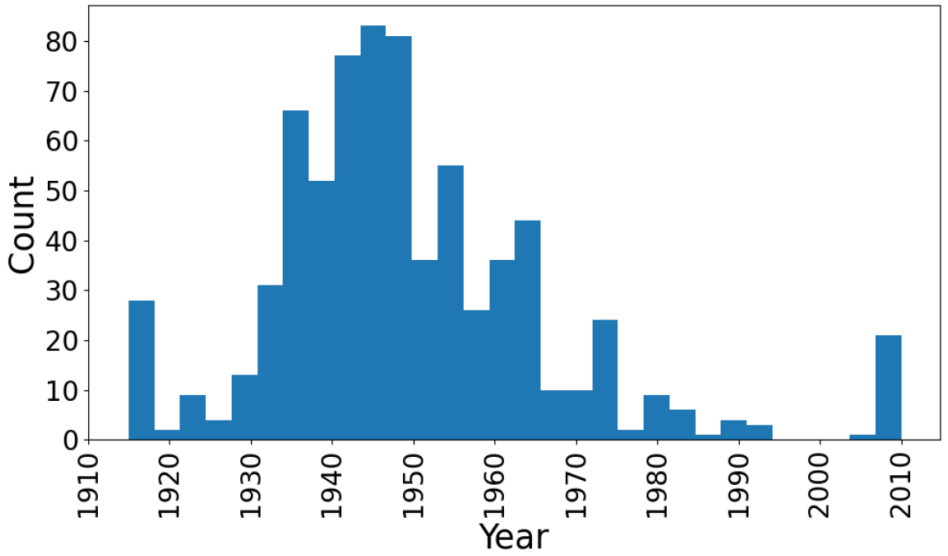


Dataset	Audio	MIDI	Self-Hosted	Mood	Video Content	Length (Hours)
HIMV-200K [48]	✓	✗	✗	✗	Music Video, User-Generated Video	-
URMP [49]	✓	✓	✗	✗	Music Performance	33.5
TikTok [50]	✓	✗	✗	✗	Dance Video	1.5
AIST++ [51]	✓	✗	✓	✗	3D Dance Motion	5.2
SymMV [52]	✓	✓	✗	✗	Music Video	76.5
MuVi-Sync [53]	✓	✓	✗	✗	Music Video	-
BGM909 [54]	✓	✓	✗	✗	Music Video	-
NES-VMDB [55]	✗	✓	✗	✗	Gameplay Video	474.0
OSSL (Ours)	✓	✗	✓	✓	Films	36.5

We aim to create an accessible, reproducible music-video datasets!

Open Screen Soundtrack Library (OSSL)

- **736 video clips** from **299 films** in **public domain** or **CC-licensed**
- **36.5 hours** in total
- **Mood annotations** as Russell's 4Q (arousal-valence model)



Haven Kim, Zachary Novack, Weihan Xu, Julian McAuley, and Hao-Wen Dong , "Video-Guided Text-to-Music Generation Using Public Domain Movie Collections," *ISMIR*, 2025.

Open Screen Soundtrack Library (OSSL)



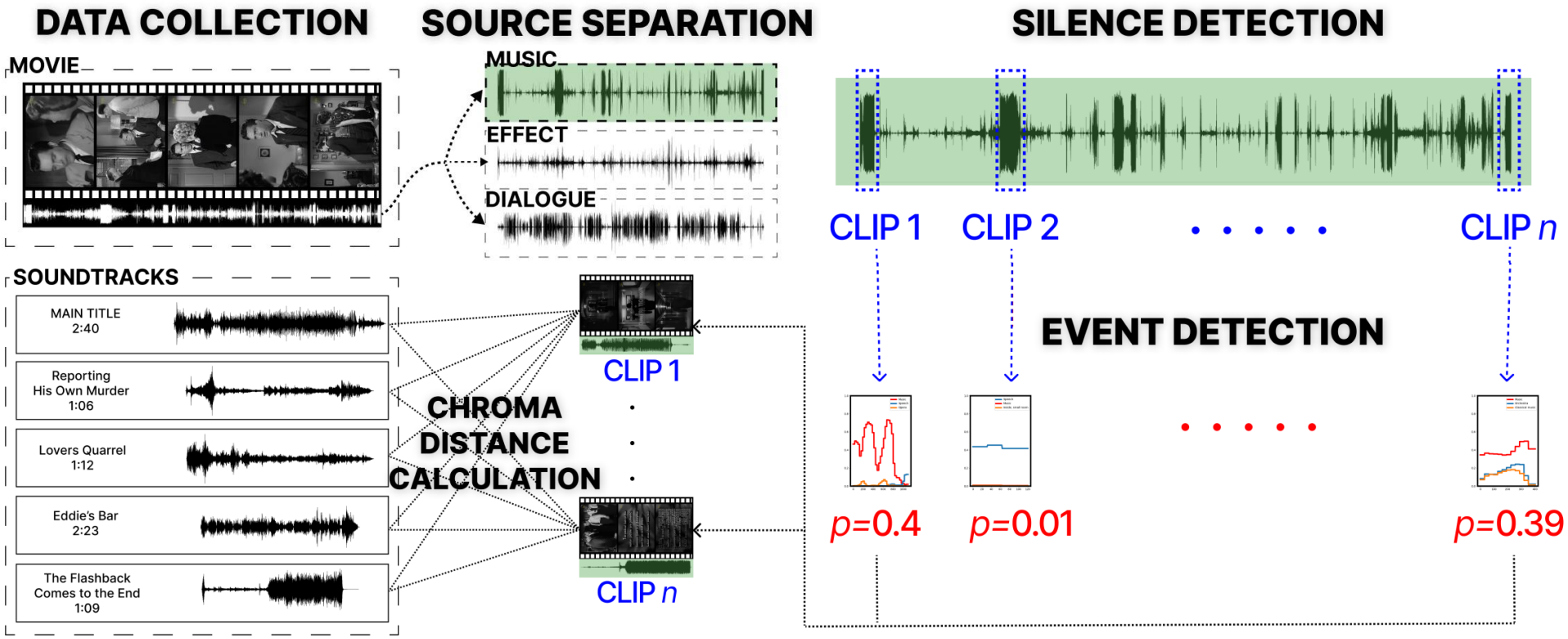





havenpersona.github.io/ossl-v1

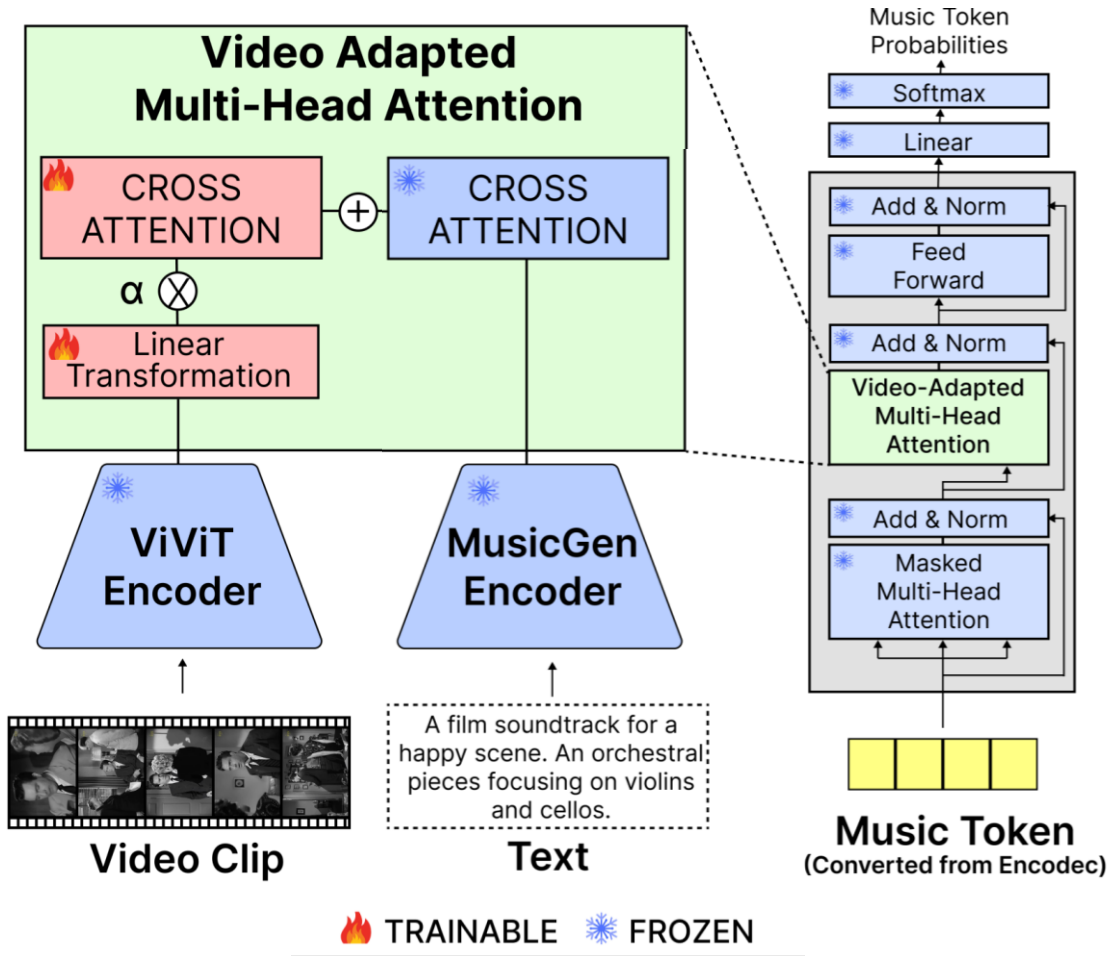
Haven Kim, Zachary Novack, Weihan Xu, Julian McAuley, and Hao-Wen Dong , "Video-Guided Text-to-Music Generation Using Public Domain Movie Collections," *ISMIR*, 2025.

Matching Soundtracks to Video Clips



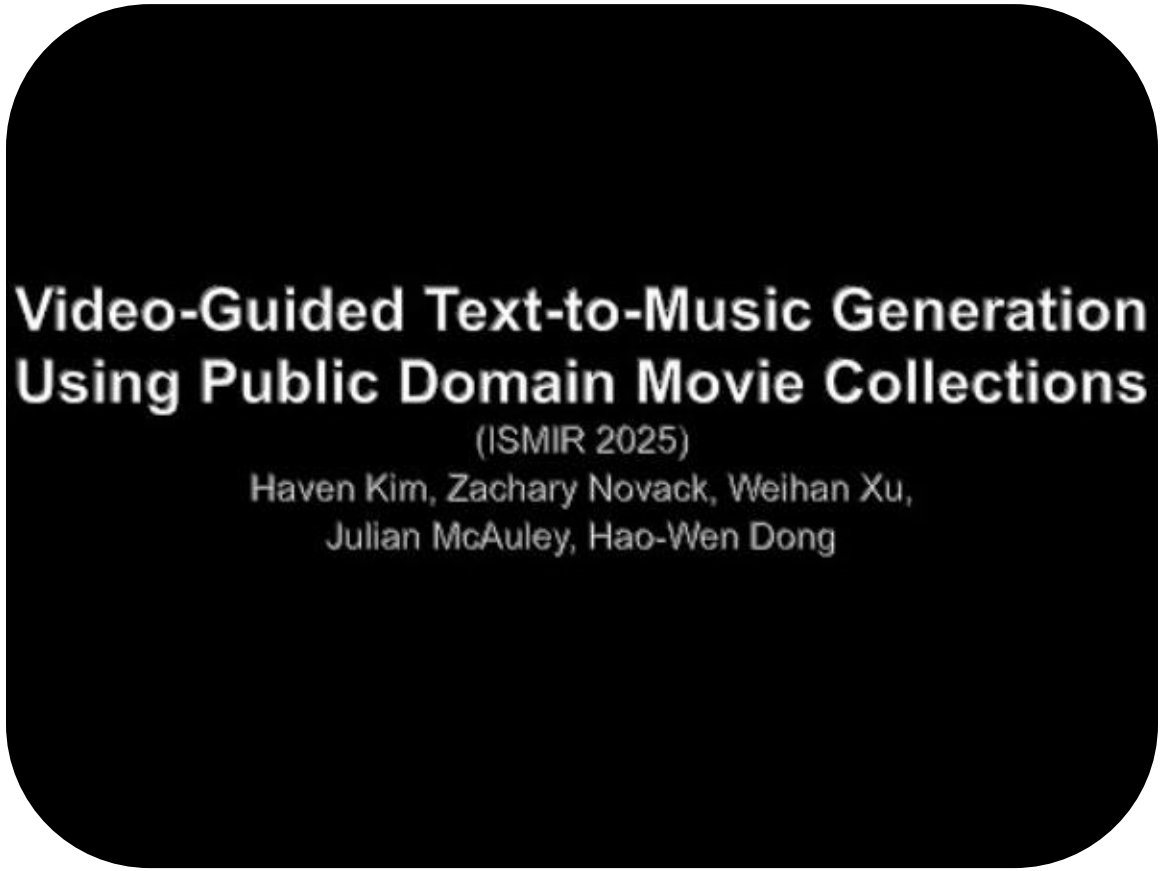
Haven Kim, Zachary Novack, Weihan Xu, Julian McAuley, and Hao-Wen Dong , "Video-Guided Text-to-Music Generation Using Public Domain Movie Collections," *ISMIR*, 2025.

Video-Guided Text-to-Music Generation



Video-Guided Text-to-Music Generation





youtu.be/S0BMicbdzmg

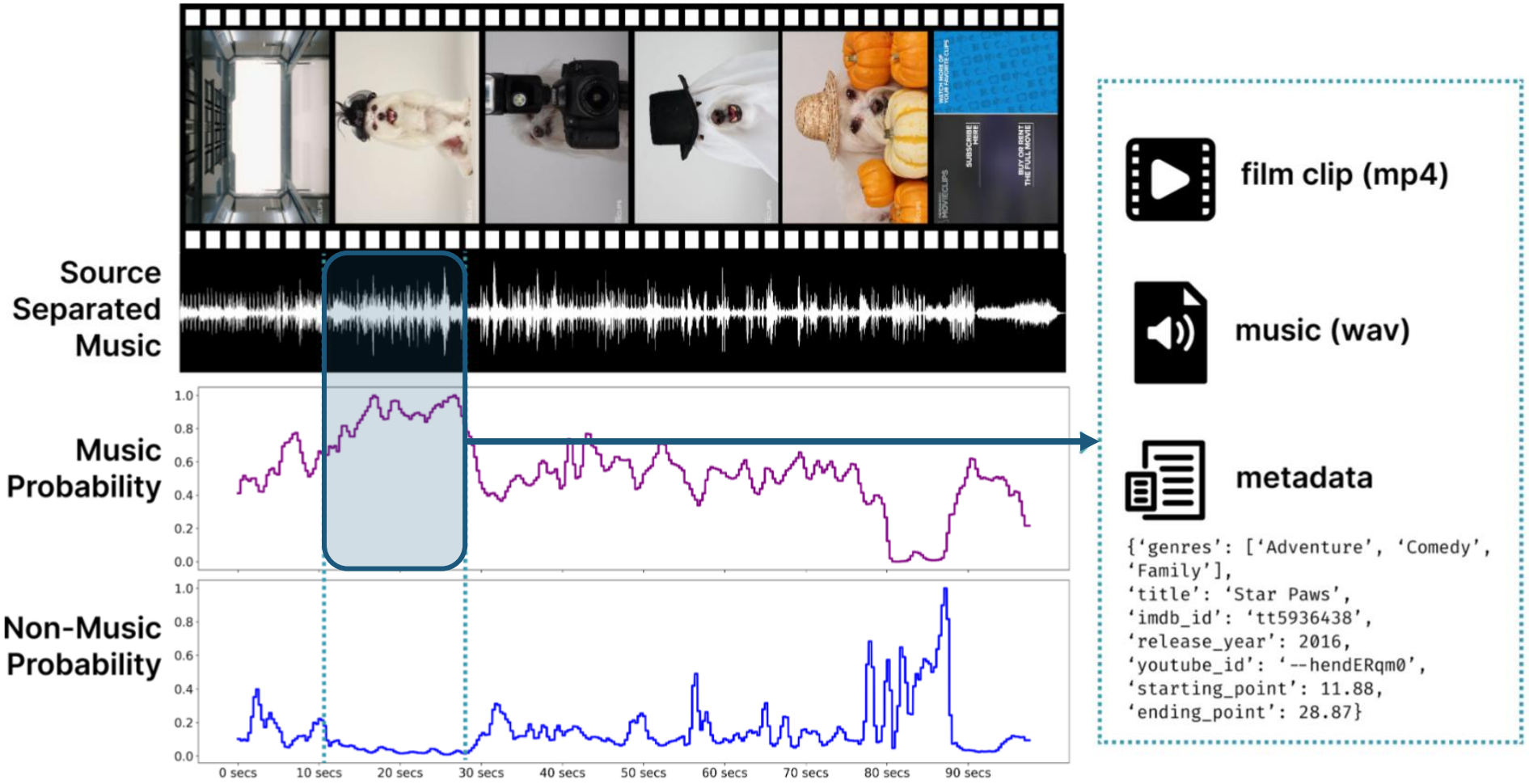
Haven Kim, Zachary Novack, Weihan Xu, Julian McAuley, and Hao-Wen Dong , "Video-Guided Text-to-Music Generation Using Public Domain Movie Collections," *ISMIR*, 2025.

Video-Guided Text-to-Music Generation Using Public Domain Movie Collections

(ISMIR 2025)

Haven Kim, Zachary Novack, Weihan Xu,
Julian McAuley, Hao-Wen Dong

Extending OSSL to OSSL v2



Haven Kim, Leduo Chen, Bill Wang, Hao-Wen Dong, and Julian McAuley, "Open Screen Soundtrack Library Version 2," *ISMIR Late-Breaking Demos*, 2025.

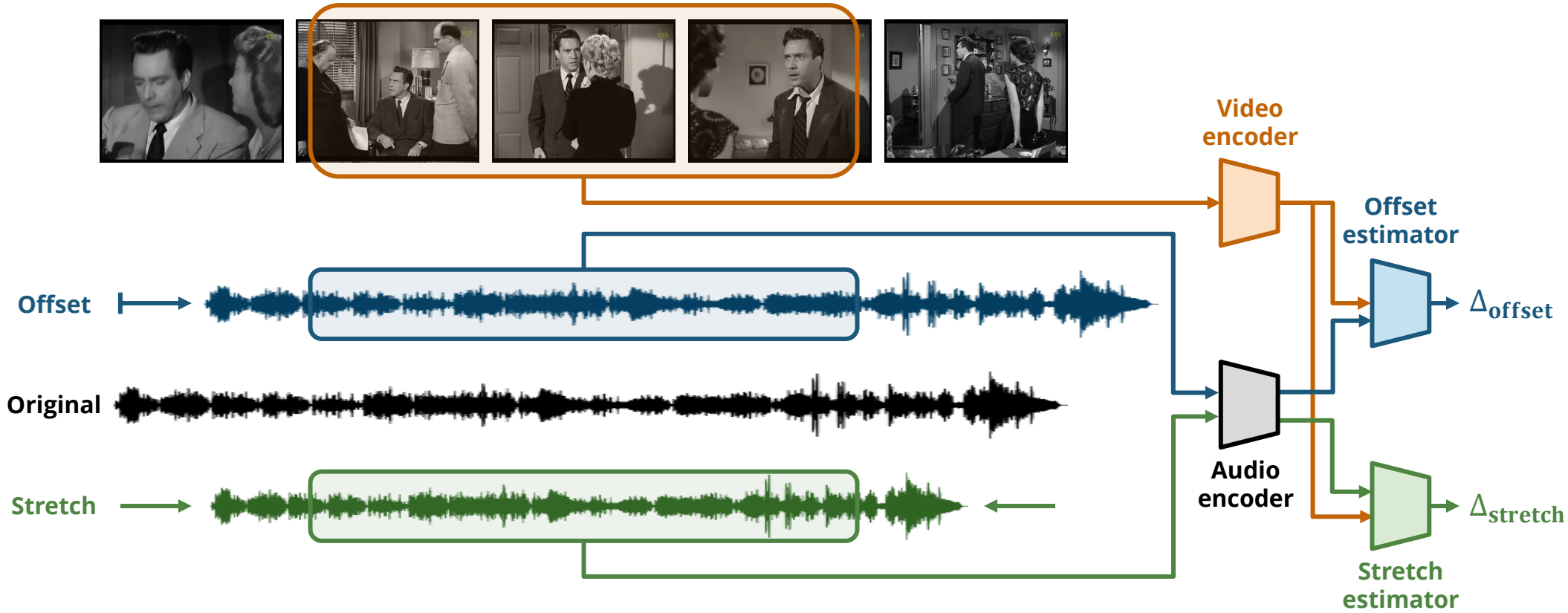
Extending OSSL to OSSL v2

	Directly downloadable	YouTube-based	
	Public Domain	**Commercial [2]**	**Total**
Number of Clips	35,705	40,703	76,408
Number of Unique Films	1,886	2,633	4,519
Average Length (seconds)	28.77	23.65	26.04
Total Length(hours)	285.31	267.39	552.70

7x larger than OSSL

Enables training foundational models for video-to-music generation from scratch!

Future Work: Measuring Video-Music Alignment



🔥 Ongoing Work: Playful Music GenAI 🔥

Maestro VR (2024)

youtu.be/OffnSNxidiY

SuperConductor

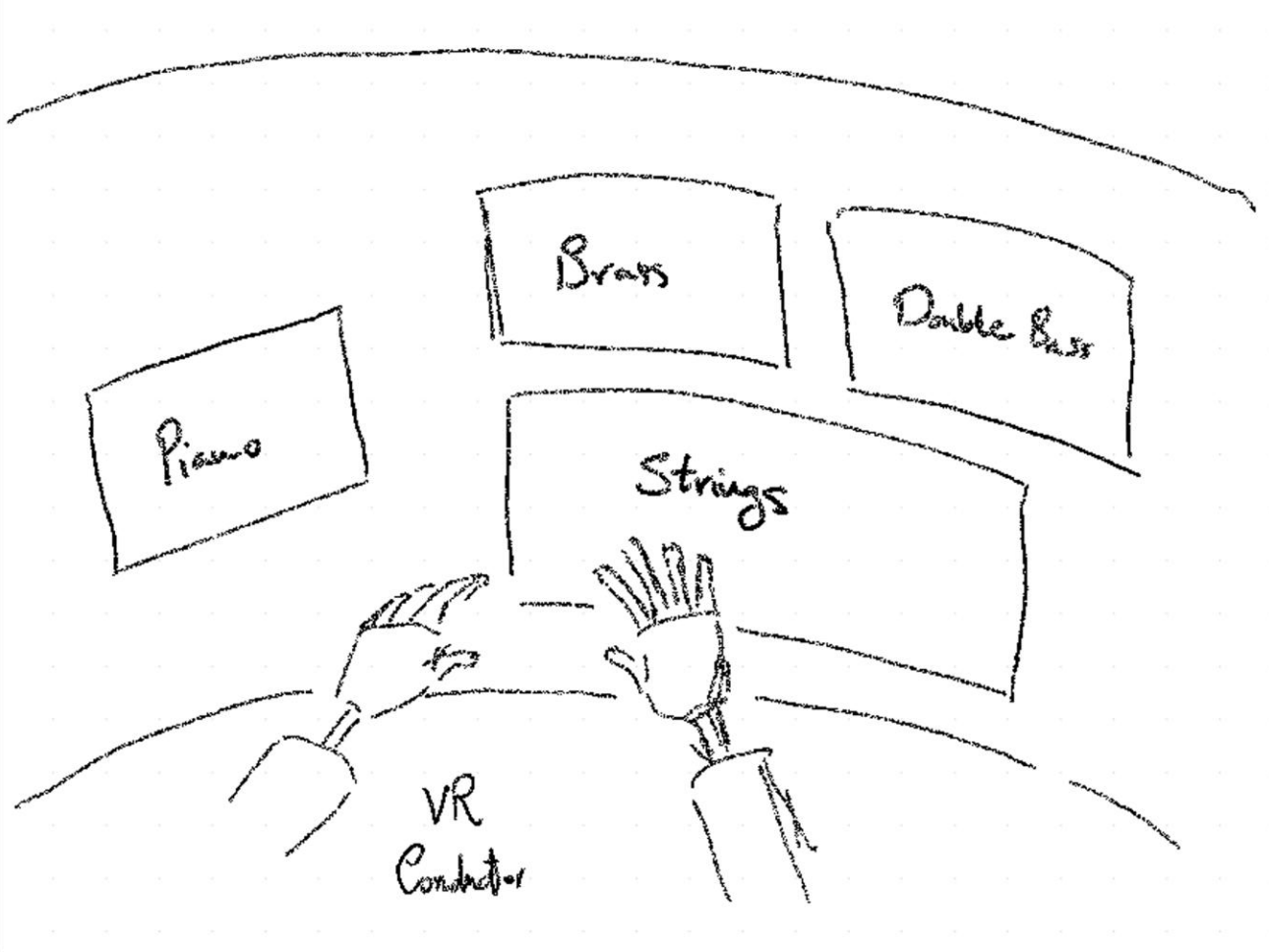
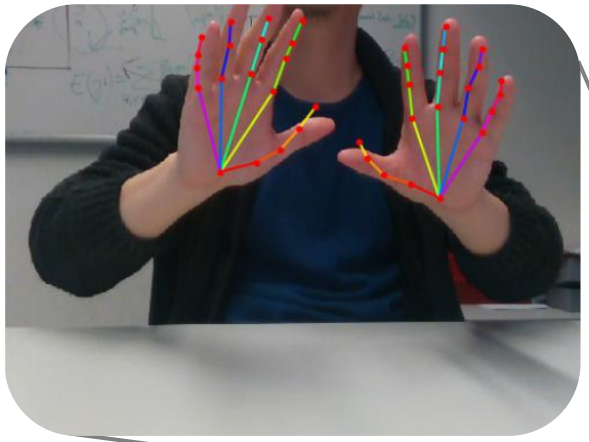


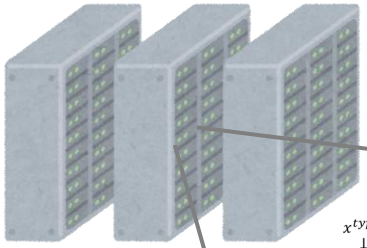
Illustration by Erfun Ackley

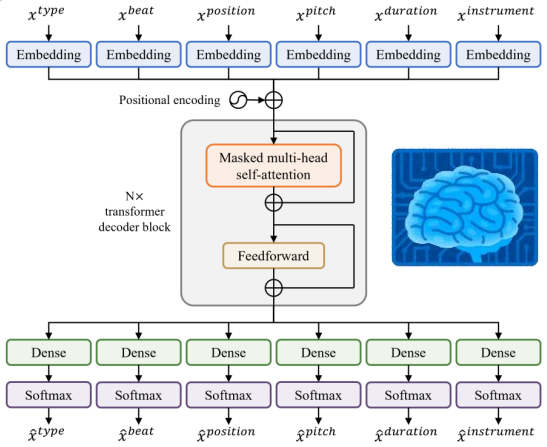
SuperConductor

(Source: Wang et al., 2020)



Low latency communication





(Source: Dong et al., 2023)

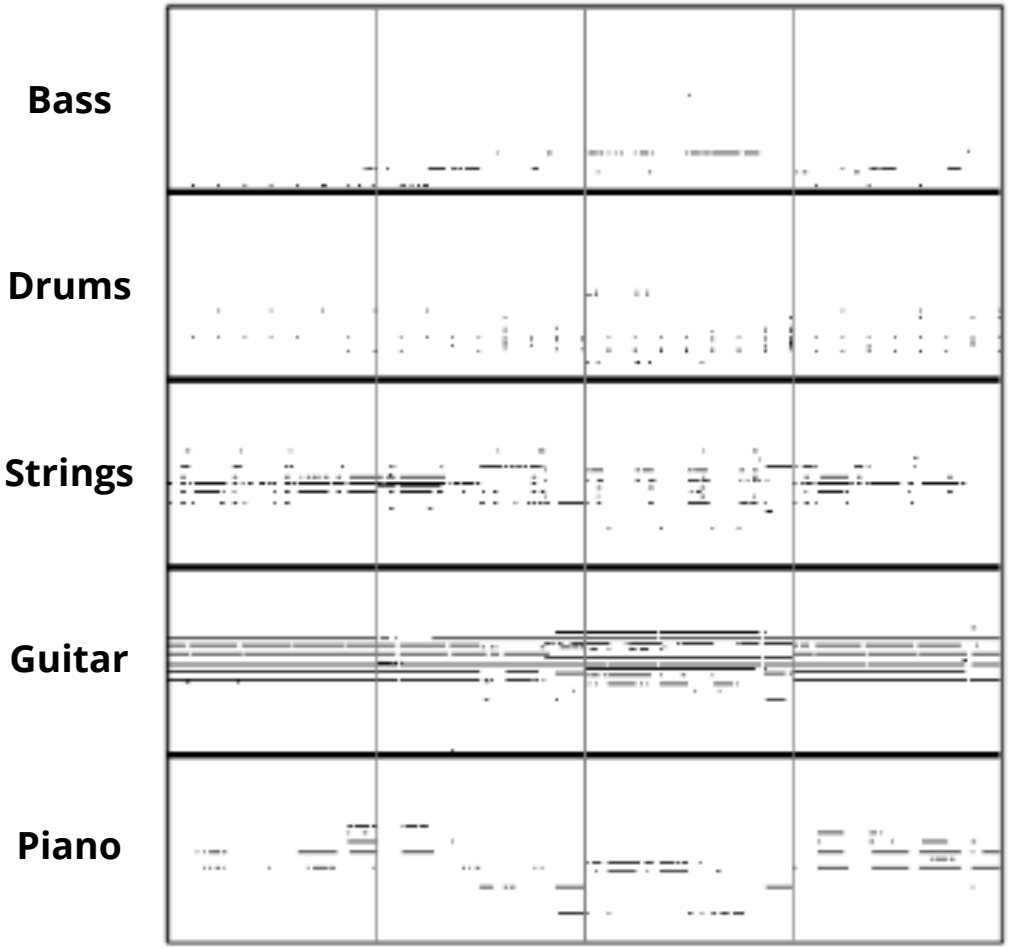
Jiayi Wang, Franziska Mueller, Florian Bernard, Suzanne Sorli, Oleksandr Sotnychenko, Neng Qian, Miguel A. Otaduy, Dan Casas, and Christian Theobalt, "RGB2Hands: Real-Time Tracking of 3D Hand Interactions from Monocular RGB Video," *SIGGRAPH Asia*, 2020.
Hao-Wen Dong, Ke Chen, Shlomo Dubnov, Julian McAuley and Taylor Berg-Kirkpatrick, "Multitrack Music Transformer," *ICASSP*, 2023.

Augmenting Human Creativity with AI

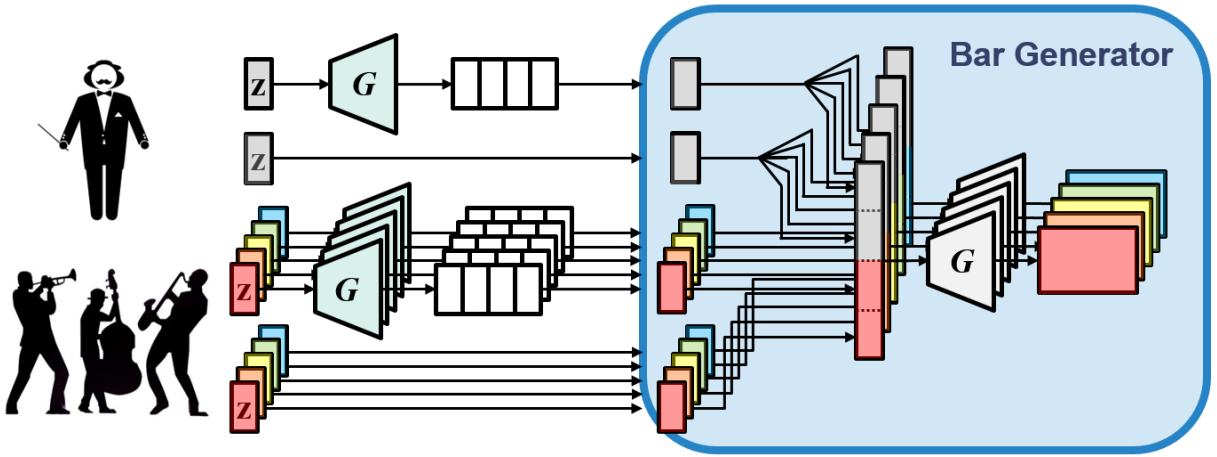
- **Novel Generative Models for New Domains**
  - **Multitrack music generation** (AAAI 2018, ISMIR 2018, ISMIR 2020, ICASSP 2023, ISMIR 2024), **text-to-music generation** (ISMIR 2025), **video-to-music generation** (ISMIR 2025), **symbolic music processing tools** (ISMIR LBD 2019, ISMIR 2020)
- **AI-assisted Tools for Content Creation**
  - **Violin performance synthesis** (ICASSP 2022, ICASSP 2025), **music instrumentation** (ISMIR 2021), **music arrangement** (AAAI 2018), **music harmonization** (JNMR 2020)
- **Multimodal Generative Models for Content Creation**
  - **Long-to-short video editing** (ICLR 2025, NeurIPS 2025), **text-queried sound separation** (ICLR 2023), **text-to-audio synthesis** (WASPAA 2023)

Generating Multi-instrument Music using GANs (**AAAI 2018**)

Multitrack Piano Roll



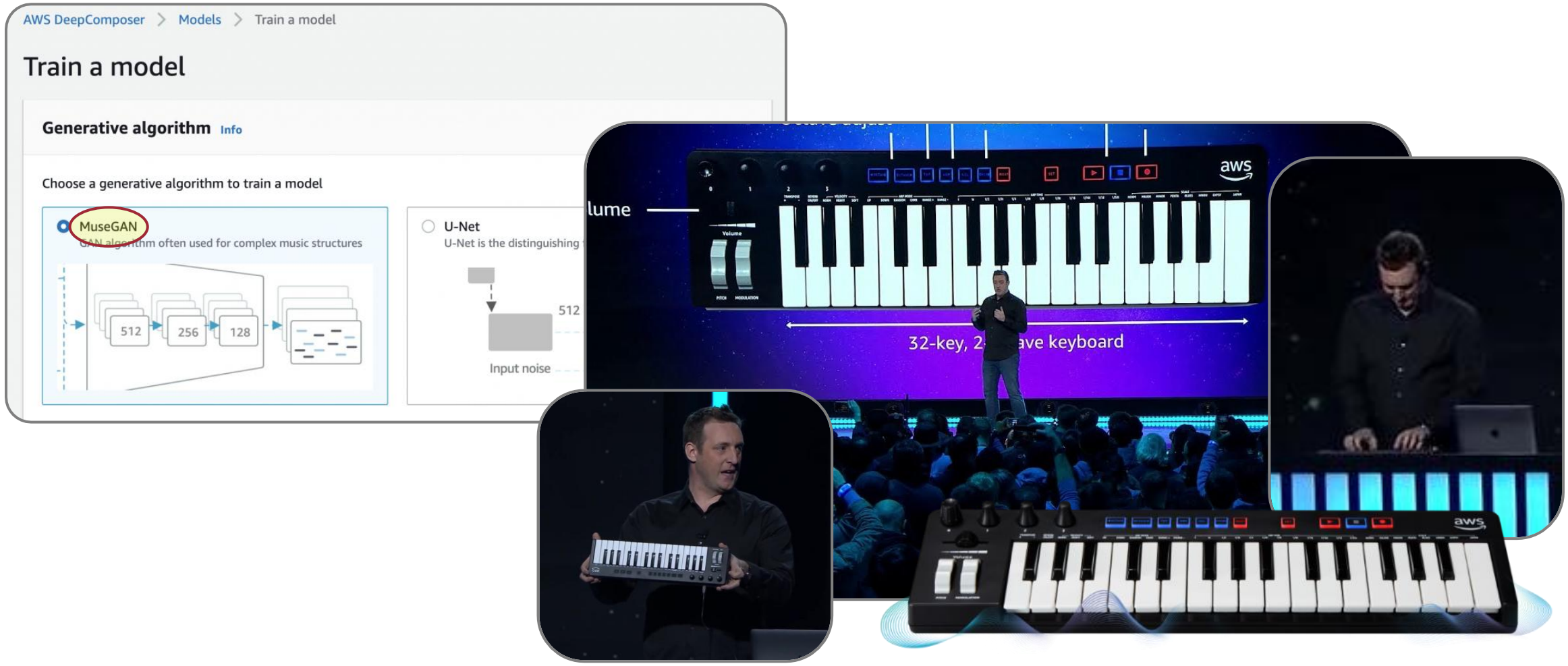
MuseGAN Generator



中央研究院
ACADEMIA SINICA

Hao-Wen Dong, Wen-Yi Hsiao, Li-Chia Yang, and Yi-Hsuan Yang, "MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment," *AAAI*, 2018.

MuseGAN Features in AWS DeepComposer (2020)



amazon.com/dp/B07YGZ4V5B/

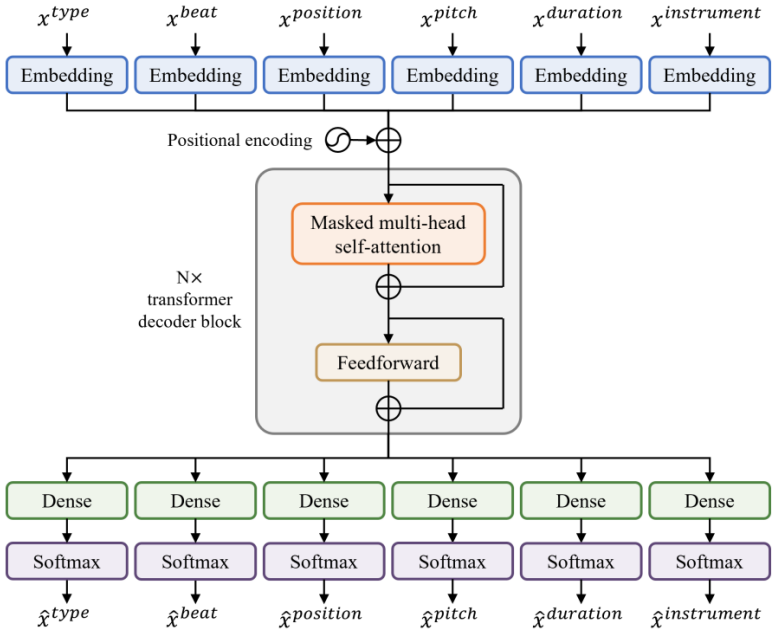
Julien Simon, "AWS DeepComposer – Now Generally Available With New Features," *AWS News Blog*, April 2, 2020.

Generating Multitrack Music with Transformers (ICASSP 2023)

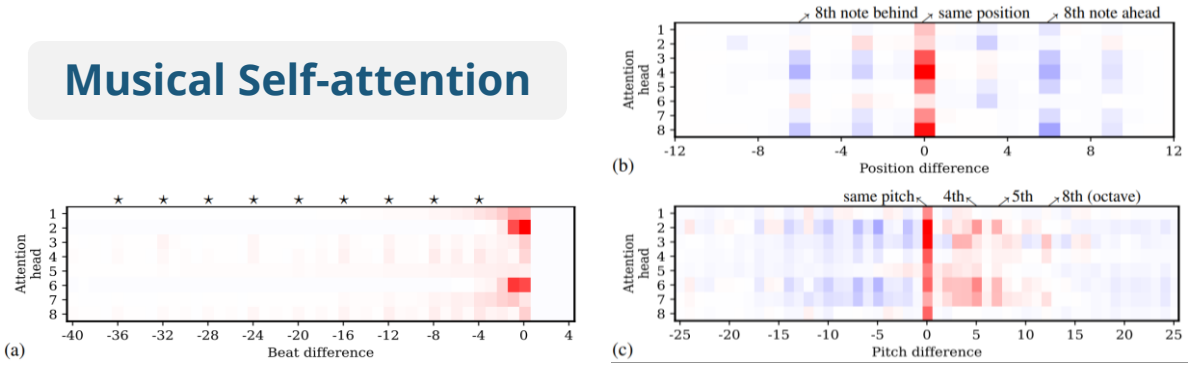
Multitrack Music Representation

```
(0,  0,  0,  0,  0,  0)   Start of song
(1,  0,  0,  0,  0, 15)   Instrument: accordion
(1,  0,  0,  0,  0, 36)   Instrument: trombone
(1,  0,  0,  0,  0, 39)   Instrument: brasses
(2,  0,  0,  0,  0,  0)   Start of notes
(3,  1,  1, 41, 15, 36)   Note: beat=1, position=1,  pitch=E2, duration=48, instrument=trombone
(3,  1,  1, 65,  4, 39)   Note: beat=1, position=1,  pitch=E4, duration=12, instrument=brasses
(3,  1,  1, 65, 17, 15)   Note: beat=1, position=1,  pitch=E4, duration=72, instrument=accordion
(3,  1,  1, 68,  4, 39)   Note: beat=1, position=1,  pitch=G4, duration=12, instrument=brasses
(3,  1,  1, 68, 17, 15)   Note: beat=1, position=1,  pitch=G4, duration=72, instrument=accordion
(3,  1,  1, 73, 17, 15)   Note: beat=1, position=1,  pitch=C5, duration=72, instrument=accordion
(3,  1, 13, 68,  4, 39)   Note: beat=1, position=13, pitch=G4, duration=12, instrument=brasses
(3,  1, 13, 73,  4, 39)   Note: beat=1, position=13, pitch=C5, duration=12, instrument=brasses
(3,  2,  1, 73, 12, 39)   Note: beat=2, position=1,  pitch=C5, duration=36, instrument=brasses
(3,  2,  1, 77, 12, 39)   Note: beat=2, position=1,  pitch=E5, duration=36, instrument=brasses
        ...               ...
(4,  0,  0,  0,  0,  0)   End of song
```

Multitrack Music Transformer



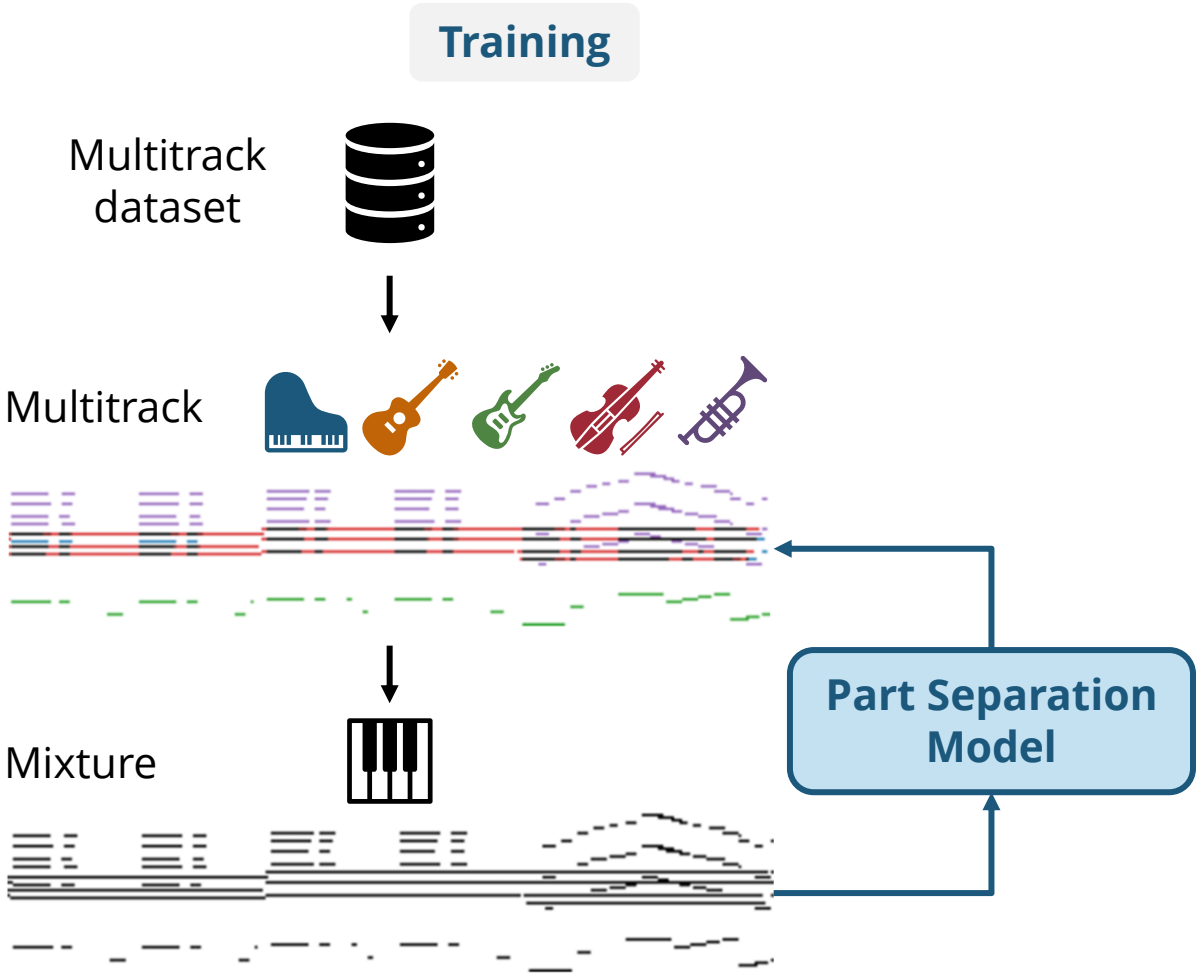
Musical Self-attention

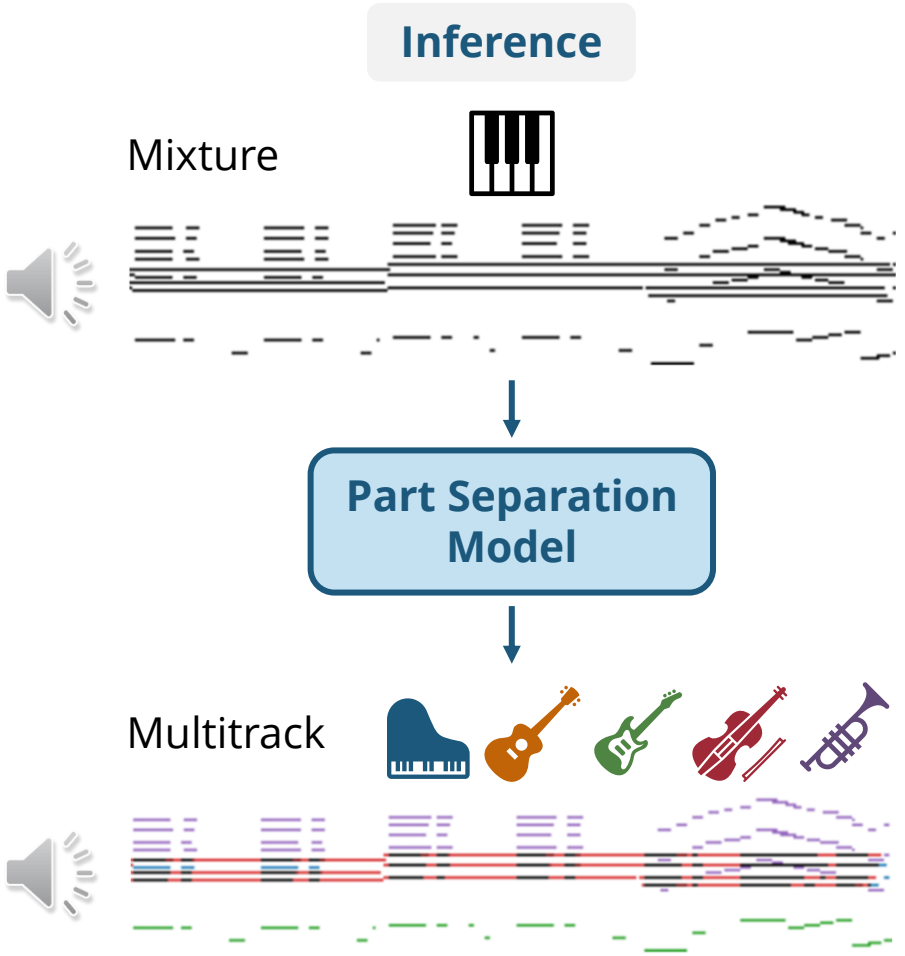


UC San Diego

Automatic Instrumentation (**ISMIR 2021**)

Stanford UC San Diego

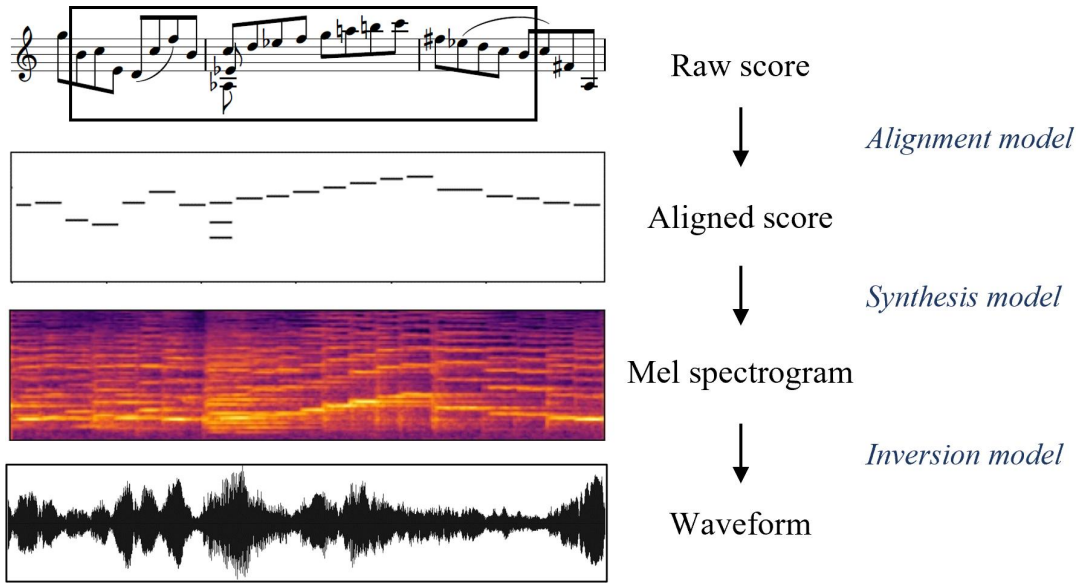




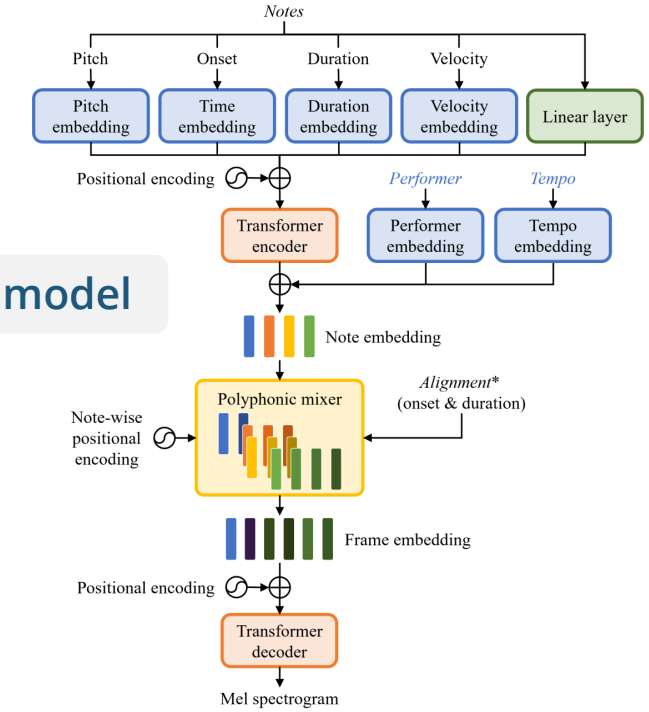
Hao-Wen Dong, Chris Donahue, Taylor Berg-Kirkpatrick and Julian McAuley, "Towards Automatic Instrumentation by Learning to Separate Parts in Symbolic Multitrack Music," *ISMIR*, 2021.

Synthesizing Expressive Violin Performance (ICASSP 2022)

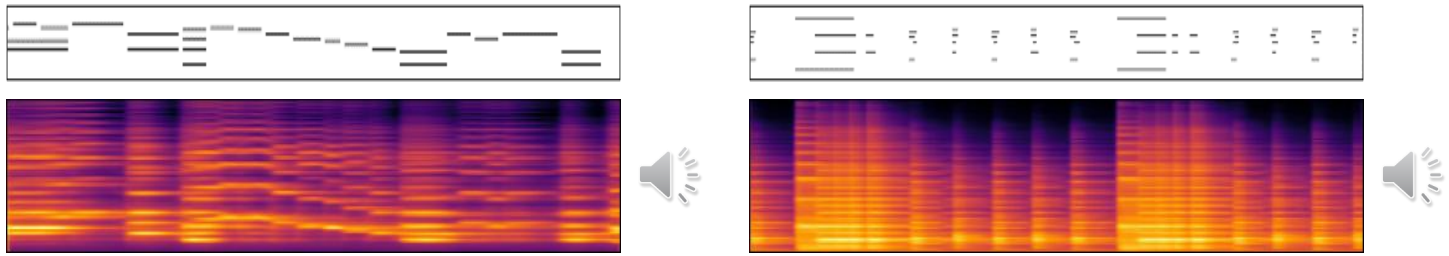
Performance synthesis



TTS-based model



Example results



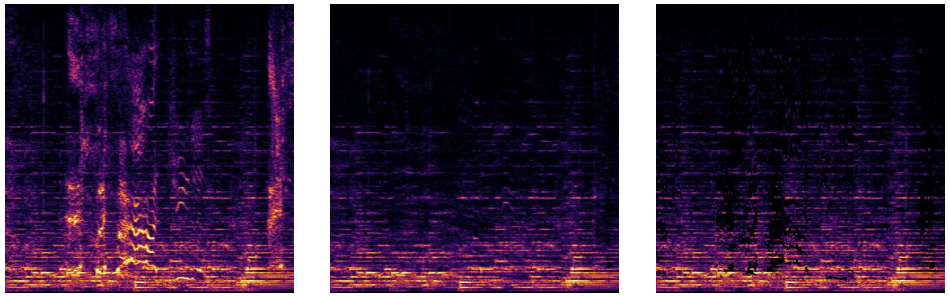
Dolby UC San Diego

Hao-Wen Dong, Cong Zhou, Taylor Berg-Kirkpatrick, and Julian McAuley, "Deep Performer: Score-to-Audio Music Performance Synthesis," *ICASSP*, 2022.

Text-queried Sound Separation (**ICLR 2023**)

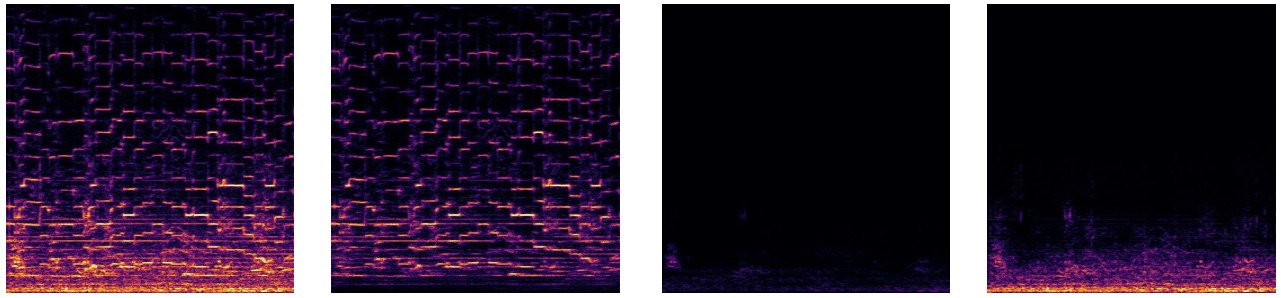
Query: "*playing harpsichord*"

Mixture **Prediction** **Ground truth**

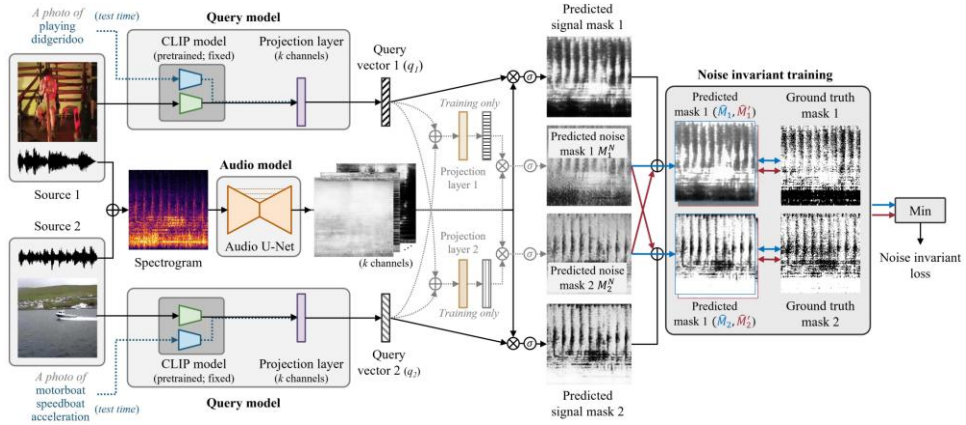


Query: "*playing bagpipe*"

Mixture **Prediction** **Noise head 1** **Noise head 2**

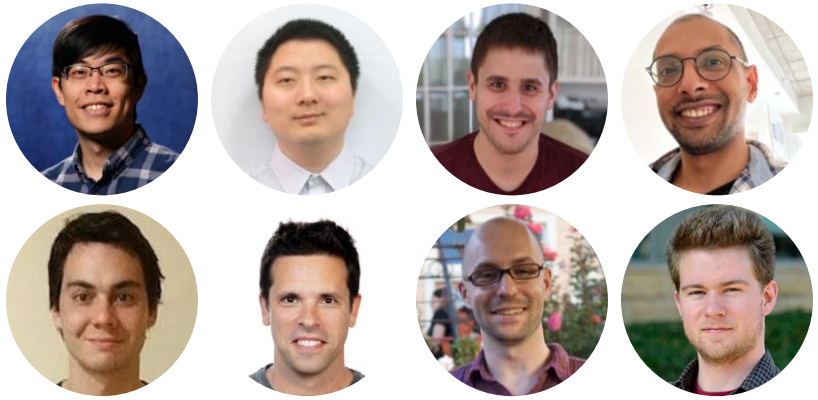


Text-queried sound separation model



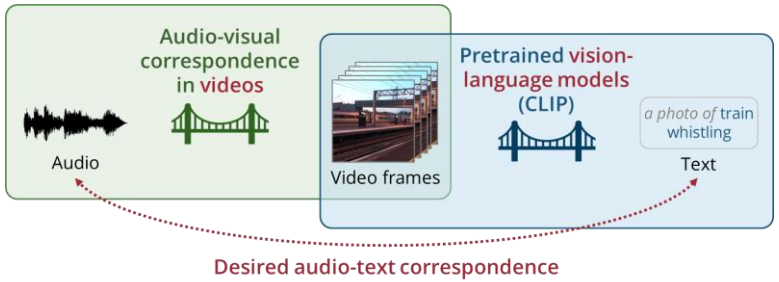
Hao-Wen Dong, Naoya Takahashi, Yuki Mitsufuji, Julian McAuley, and Taylor Berg-Kirkpatrick, "CLIPSep: Learning Text-queried Sound Separation with Noisy Unlabeled Videos," *ICLR*, 2023.

Text-to-Audio Synthesis (**WASPAA 2023**)

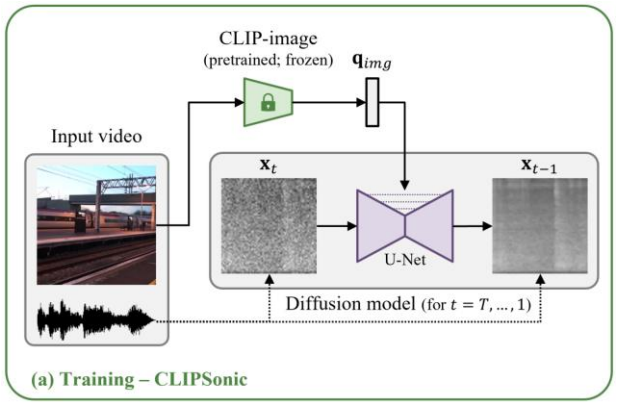


Dolby UC San Diego

Learning Sounds from Noisy Videos



Training



Inference

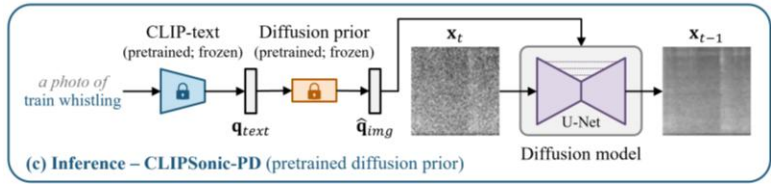


Image-to-sound results

Text-to-sound results

Hao-Wen Dong, Xiaoyu Liu, Jordi Pons, Gautam Bhattacharya, Santiago Pascual, Joan Serrà, Taylor Berg-Kirkpatrick, and Julian McAuley, "CLIPSonic: Text-to-Audio Synthesis with Unlabeled Videos and Pretrained Language-Vision Models," *WASPAA*, 2023.

Art challenges Technology

Creativity

Augmenting Human Creativity with AI

AI

Technology inspires the Art

Augmenting Human Creativity with AI

- **Novel Generative Models for New Domains**
  - **Multitrack music generation** (AAAI 2018, ISMIR 2018, ISMIR 2020, ICASSP 2023, ISMIR 2024), **text-to-music generation** (ISMIR 2025), **video-to-music generation** (ISMIR 2025), **symbolic music processing tools** (ISMIR LBD 2019, ISMIR 2020)
- **AI-assisted Tools for Content Creation**
  - **Violin performance synthesis** (ICASSP 2022, ICASSP 2025), **music instrumentation** (ISMIR 2021), **music arrangement** (AAAI 2018), **music harmonization** (JNMR 2020)
- **Multimodal Generative Models for Content Creation**
  - **Long-to-short video editing** (ICLR 2025, NeurIPS 2025), **text-queried sound separation** (ICLR 2023), **text-to-audio synthesis** (WASPAA 2023)

Generative AI for Music, Audio & Video Creation

(Source: The Denver Post)

(Source: UploadVR)

(Source: Descript)

(Source: Daily Bruin)

(Source: Wikimedia Commons)

Universitaetsmedizin, CC BY-SA 4.0, via Wikimedia Commons
uploadvr.com/iron-man-vr-breaks-free-from-cords-load-screens-on-quest-2/
descript.com/blog/article/what-is-the-best-audio-interface-for-recording-a-podcast
denverpost.com/2019/08/02/colorado-symphony-movie-scores-harry-potter-star-wars/
dailybruin.com/2023/08/04/theater-review-the-musical-les-misrables-offers-stellar-displays-and-impassioned-vocals

Augmenting Human Creativity with AI

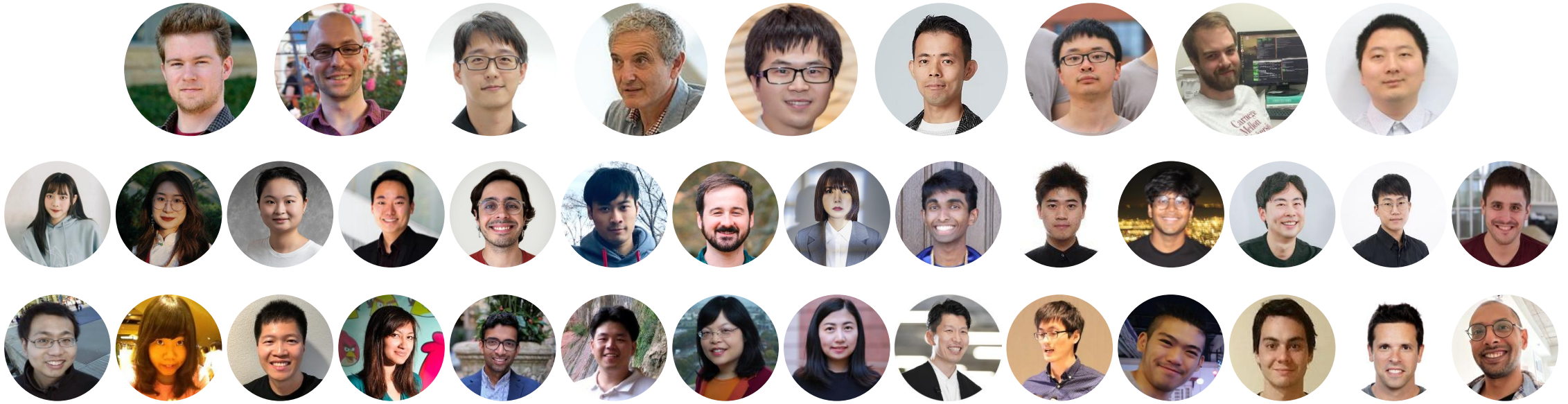
- **Multimodal generative AI for content creation**
- **Human-AI co-creative tools for music, audio and video creation**
- **Human-like machine leaning algorithms for music, movies and arts**

Works of art make rules;
rules do not make works of art.

– Claude Debussy

Generative AI for Music and Audio

Nothing would have been possible without all my fantastic collaborators!



UC San Diego · 中央研究院 ACADEMIA SINICA · Dolby · SONY · amazon · NVIDIA · Ministry of Education 教育部 · erc

hermandong.com / hwdong@umich.edu

UNIVERSITY OF MICHIGAN