68th VDLM

Vienna Deep Learning Meetup

hosted by SBA Research Wien

Topic for today:
CommentSense

Send us announcements
& job openings!

contact@vdlm.at

Vienna Deep Learning Meetup

Welcome by the Meetup organizers:

Thomas Lidy
iGroove

René Donner
mva.ai

Jan Schlüter
JKU Linz

Alex Schindler
AIT & TU Wien

Pavol Harar
ISTA & ACAI.AI

contact@vdlm.at

Welcome to SBA Research

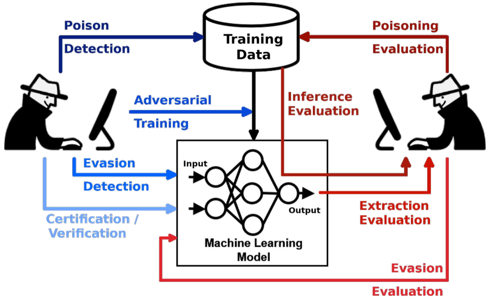
Research Group Security & Privacy, Uni Wien

Rudolf Mayer & the MLDM Team

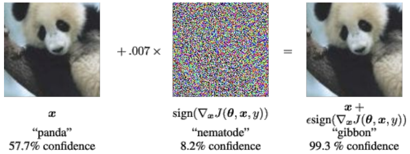
Security & Privacy @ SBA & Uni Wien

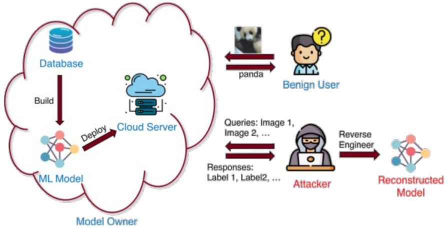
- ~140 employees combined
- Scientific Lead: Edgar Weippl
- Research topics
  - Mathematics for Testing, Reliability and Information Security
  - Networks and Critical Infrastructures Security
  - Systems and (I)IoT Security
  - Decentralized Systems & Distributed Ledgers
  - Complexity and Resilience
- More details:
  - https://www.sba-research.org/research, https://sec.cs.univie.ac.at/

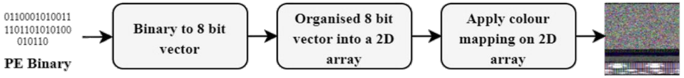
Why do we host the Vienna Deep Learning Meetup?



- Machine Learning and Data Management Group
  - Privacy-preserving Machine Learning
    - Federated Learning, Synthetic Data, Differential Privacy, …
  - Security of Machine Learning / Adversarial ML
    - Evasion attacks (adversarial examples), data poisoning, …
  - Protection of ML assets
    - (Training) Data, models, outputs (e.g. genAI), …
  - Machine Learning for security
    - Intrusion / malware / anomaly detection, …
- https://www.sba-research.org/mldm/







68th VDLM

Vienna Deep Learning Meetup

18:30 Welcome by the organizers
Welcome by our host: **SBA Research**

18:45 CommentSense: An On-Device AI Browser Extension for Real-Time YouTube Comment Understanding
Marc Kroll

19:10 Hot Topics
Announcements

19:30 Networking & Discussions

Hot Topics

Interesting recent research, feel free to contribute!
contact@vdlm.at

top papers from ISMIR 2025

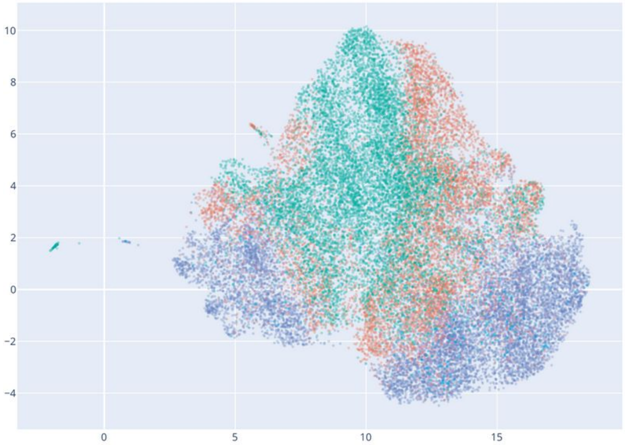
~~top papers from ISMIR 2025~~
some interesting ISMIR papers

The AI Music Arms Race: On the Detection of AI-Generated Music

Laura Cros Vila, Bob Sturm, Luca Casini, David Dalmazzo

- collected 10k human-made, 10k suno, 10k udio songs
- computed general audio + CLAP features

2D UMAP:

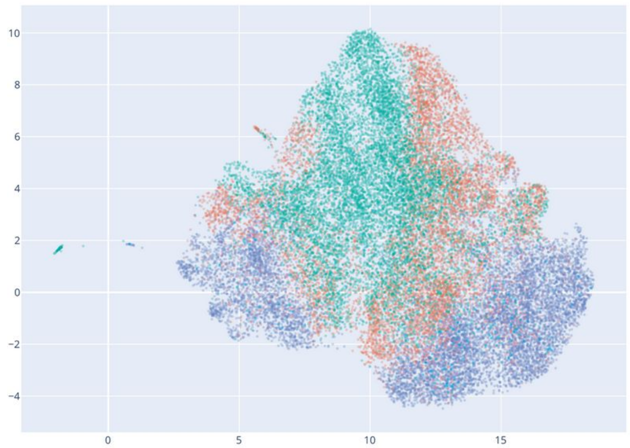


The AI Music Arms Race: On the Detection of AI-Generated Music

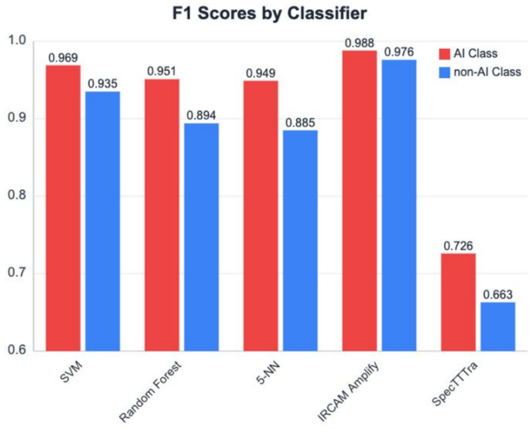
Laura Cros Vila, Bob Sturm, Luca Casini, David Dalmazzo

- collected 10k human-made, 10k suno, 10k udio songs
- computed general audio + CLAP features

2D UMAP:



Classifier:

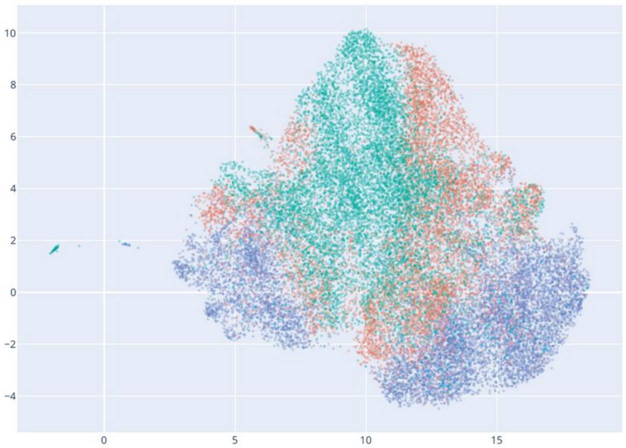


The AI Music Arms Race: On the Detection of AI-Generated Music

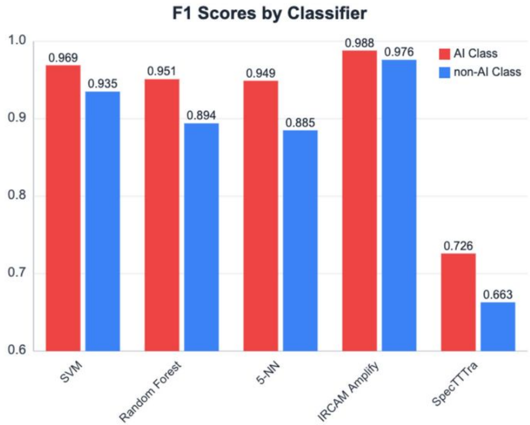
Laura Cros Vila, Bob Sturm, Luca Casini, David Dalmazzo

- collected 10k human-made, 10k suno, 10k udio songs
- computed general audio + CLAP features

2D UMAP:



Classifier:



But:

Performance breaks when downsampling audio to 22.05 kHz

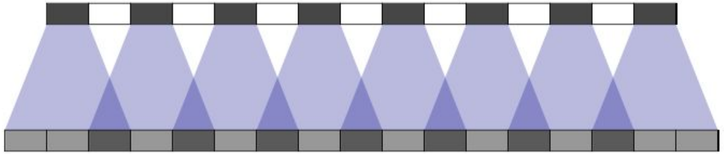
Does not generalize to unseen services

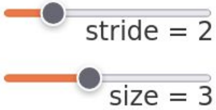
IRCAM Amplify updates often – arms race

A Fourier Explanation of AI-music Artifacts

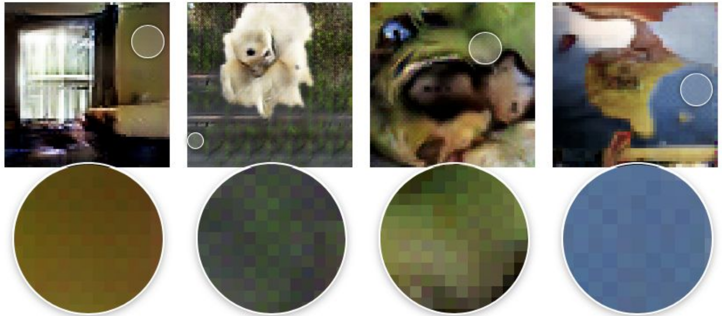
Darius Afchar, Gabriel Meseguer Brocal, Kamil Akesbi, Romain Hennequin

Neural audio codecs in music generation models use transposed strided convolutions





Transposed convolutions lead to checkerboard artifacts in images

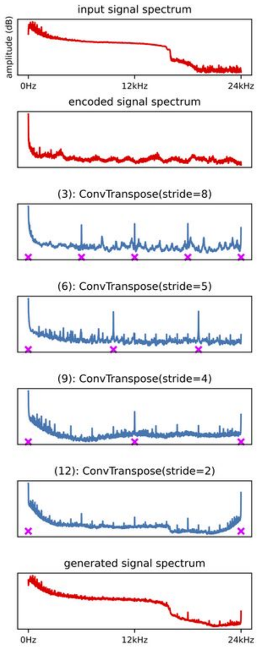


Radford, et al., 2015 [1] Salimans et al., 2016 [2] Donahue, et al., 2016 [3] Dumoulin, et al., 2016 [4]

A Fourier Explanation of AI-music Artifacts

Darius Afchar, Gabriel Meseguer Brocal, Kamil Akesbi, Romain Hennequin

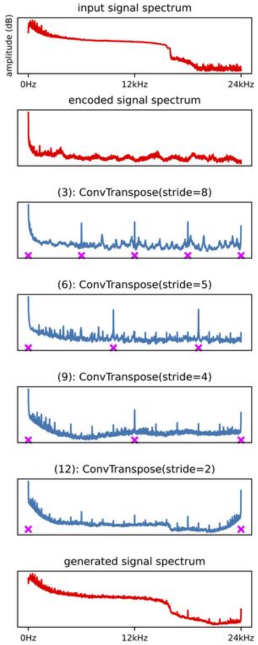
Finding: Transposed convolutions
lead to similar artifacts in spectra



A Fourier Explanation of AI-music Artifacts

Darius Afchar, Gabriel Meseguer Brocal, Kamil Akesbi, Romain Hennequin

Finding: Transposed convolutions lead to similar artifacts in spectra



Idea: Train a linear regressor to differentiate original audio vs. output of neural audio codec, works on par with existing, more complex models

Class	**Our**	*Reported from [17]*
Real	99.87	99.7
Synthetic		
↪ DAC (14kbps)	99.68	99.3
↪ Encodec (24kbps)	99.81	99.7
↪ Musika!	99.97	100.0

Class	Our	*Reported from [16]*
Real	99.97	99
Synthetic		
↪ Suno v3.5	100.00	100
↪ *Suno v3*†	100.00	96
↪ *Suno v2*†	99.90	78
↪ Udio 130	100.00	100
↪ *Udio 32*†	39.83	96

A Fourier Explanation of AI-music Artifacts

Darius Afchar, Gabriel Meseguer Brocal, Kamil Akesbi, Romain Hennequin

Work done by Deezer

Fun facts:

up to 25% of music uploaded to Deezer is generated by AI

A Fourier Explanation of AI-music Artifacts

Darius Afchar, Gabriel Meseguer Brocal, Kamil Akesbi, Romain Hennequin

Work done by Deezer

Fun facts:

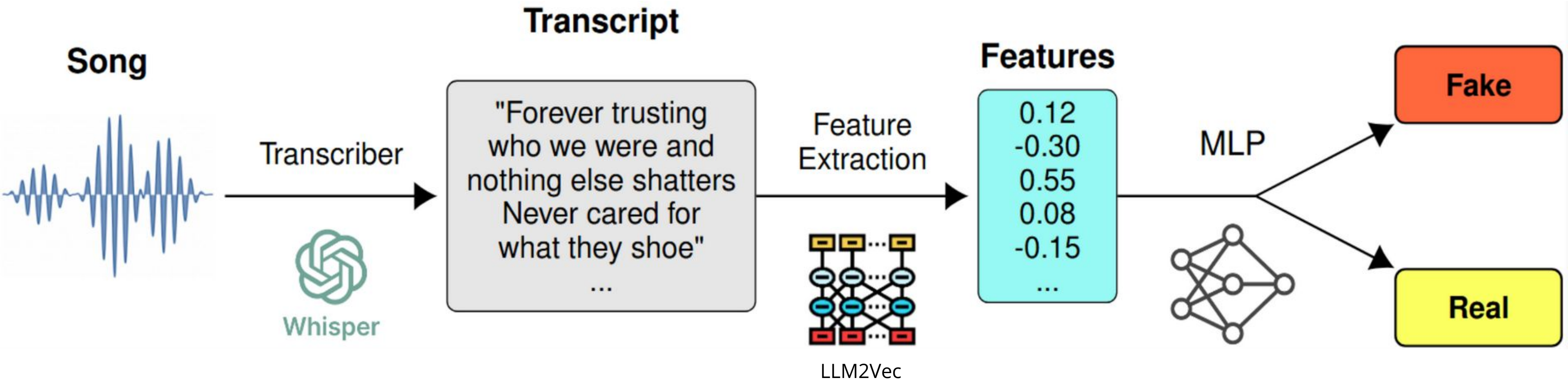
up to 25% of music uploaded to Deezer is generated by AI

up to 70% of streams of AI-generated music is by bots

AI-Generated Song Detection via Lyrics Transcripts

Markus Frohmann, Elena Epure, Gabriel Meseguer Brocal, Markus Schedl, Romain Hennequin

Work done by Deezer with JKU Linz

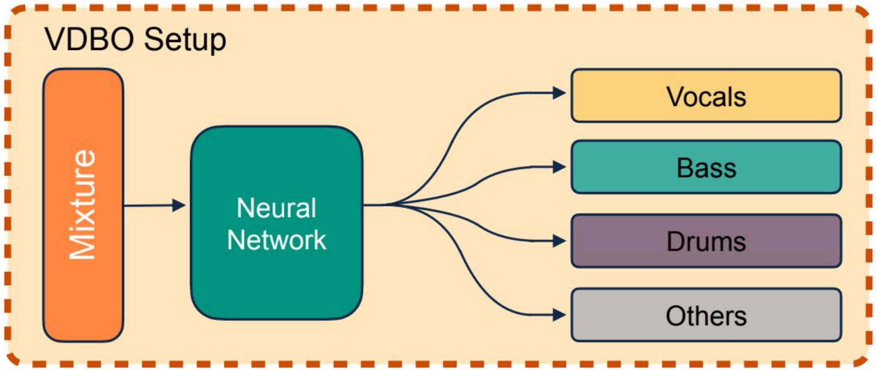


Somewhat worse than audio-based (90% vs. 97%), but generalizes across different services

User-Guided Generative Source Separation

Yutong Wen, Minje Kim, Paris Smaragdis

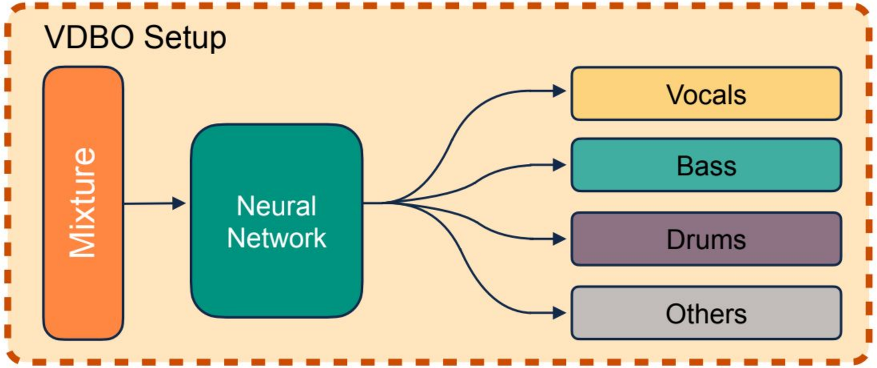
- Off-the-shelf systems: four predefined stems (e.g., Spleeter, Demucs)



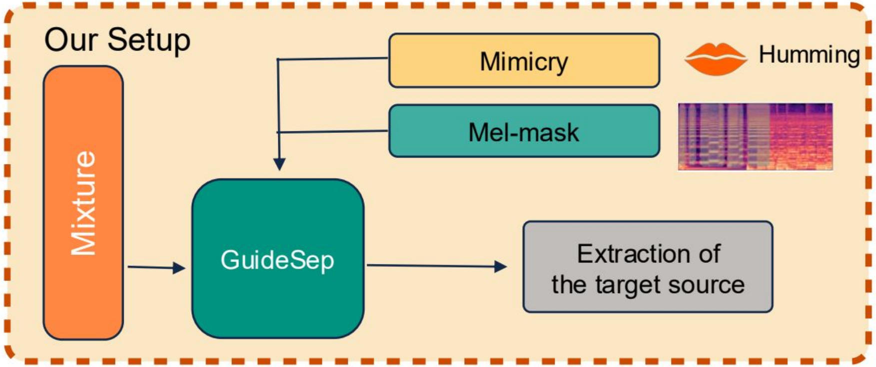
User-Guided Generative Source Separation

Yutong Wen, Minje Kim, Paris Smaragdis

Off-the-shelf systems: four predefined stems (e.g., Spleeter, Demucs)



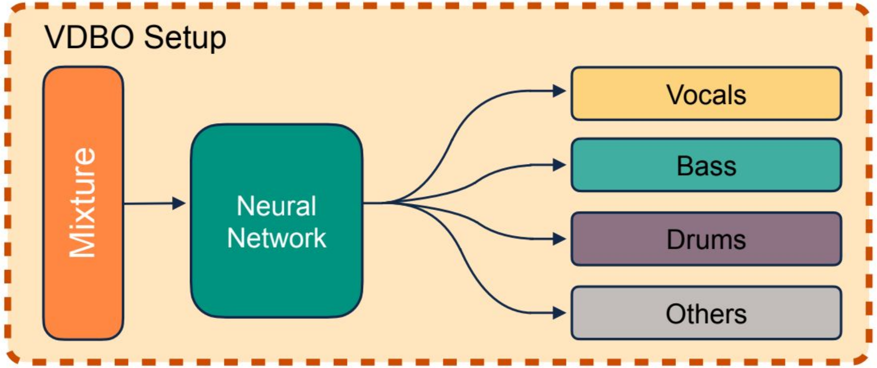
Their system: select part to extract by mimicking it (humming, other instrument)



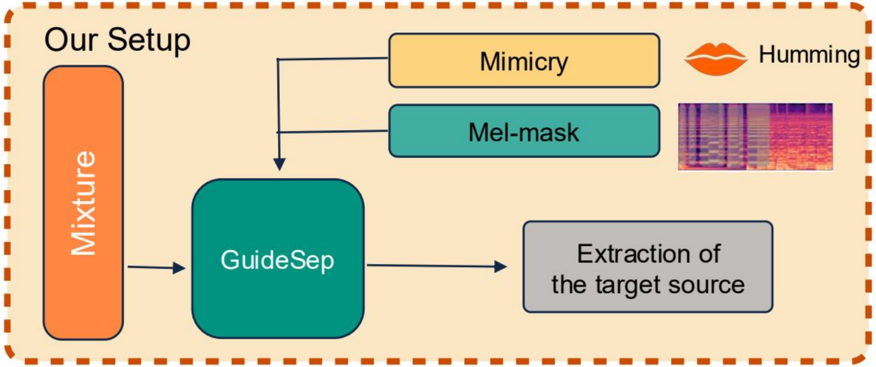
User-Guided Generative Source Separation

Yutong Wen, Minje Kim, Paris Smaragdis

Off-the-shelf systems: four predefined stems (e.g., Spleeter, Demucs)



Their system: select part to extract by mimicking it (humming, other instrument)



Model: Conditioned diffusion U-Net

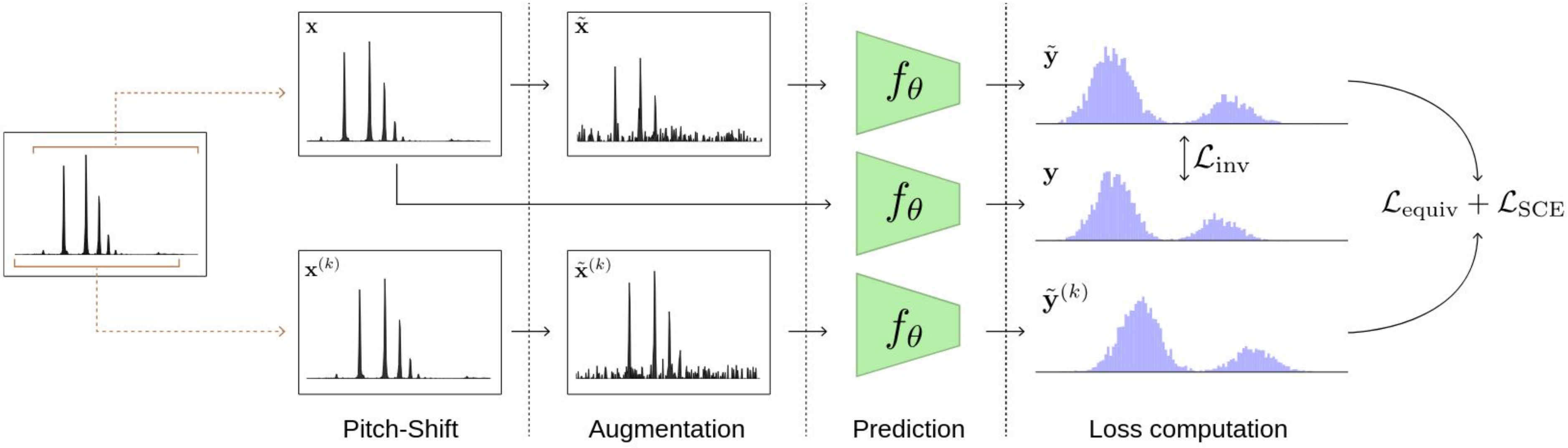
Demo:

Data: Synthesized mimicry from audio-aligned MIDI

Improving Neural Pitch Estimation with SWIPE Kernels

David Marttila, Joshua D. Reiss

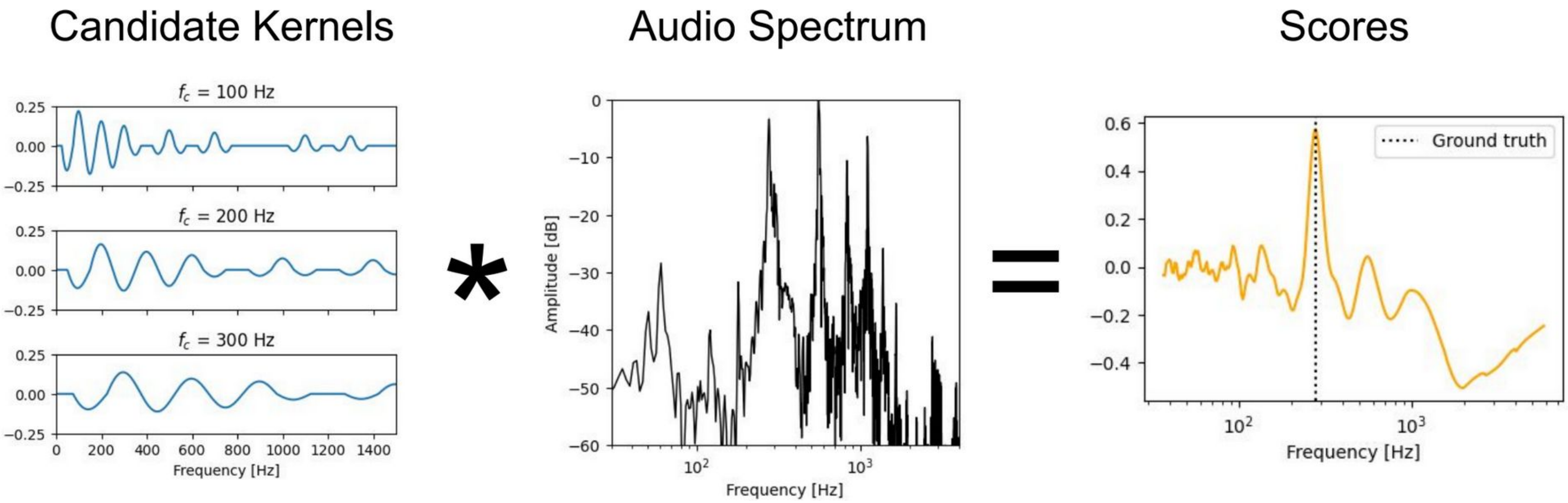
PESTO: self-supervised deep learning approach for pitch estimation



Improving Neural Pitch Estimation with SWIPE Kernels

David Marttila, Joshua D. Reiss

SWIPE: old spectral-candidate-based pitch estimation approach



Improving Neural Pitch Estimation with SWIPE Kernels

David Marttila, Joshua D. Reiss

Findings:

Replacing audio frontend in PESTO with SWIPE improves results

Method	# params	Trained on	Raw Pitch Accuracy MIR-1K	MDB-stem-synth
PESTO	28.9k	MIR-1K	96.1%	**94.6%**
		MDB-stem-synth	93.5%	95.5%
SWIPE-full	28.2k	MIR-1K	**97.0%**	89.7%
		MDB-stem-synth	96.1%	96.4%

Improving Neural Pitch Estimation with SWIPE Kernels

David Marttila, Joshua D. Reiss

Findings:

Replacing audio frontend in PESTO with SWIPE improves results

Shrinking model from 28k parameters to 647 parameters further improves results

Method	# params	Trained on	Raw Pitch Accuracy MIR-1K	Raw Pitch Accuracy MDB-stem-synth
PESTO	28.9k	MIR-1K	96.1%	**94.6%**
		MDB-stem-synth	93.5%	95.5%
SWIPE-full	28.2k	MIR-1K	**97.0%**	89.7%
		MDB-stem-synth	96.1%	96.4%
SWIPE-tiny	647	MIR-1K	96.6%	90.1%
		MDB-stem-synth	**96.4%**	**96.5%**

Improving Neural Pitch Estimation with SWIPE Kernels

David Marttila, Joshua D. Reiss

Findings:

Replacing audio frontend in PESTO with SWIPE improves results

Shrinking model from 28k parameters to 647 parameters further improves results

SWIPE alone works almost as well
→ Performance was misreported in many pitch estimation papers

Method	# params	Trained on	Raw Pitch Accuracy MIR-1K	Raw Pitch Accuracy MDB-stem-synth
SWIPE	-	-	96.2%	96.1%
PESTO	28.9k	MIR-1K	96.1%	**94.6%**
		MDB-stem-synth	93.5%	95.5%
SWIPE-full	28.2k	MIR-1K	**97.0%**	89.7%
		MDB-stem-synth	96.1%	96.4%
SWIPE-tiny	647	MIR-1K	96.6%	90.1%
		MDB-stem-synth	**96.4%**	**96.5%**

Announcements

Vienna Deep Learning Meetup

Resources - all past talks & slides:

github.com/vdlm

Send us SPEAKER suggestions,
job & event announcements:

contact@vdlm.at

Vienna Deep Learning Meetup

Call for Speakers!

Do you work on Deep Learning? In academia or industry?

We'd like to hear your story!

We are looking for speakers for **future meetups**.

Short talks (20-30 min) or **long talks** (40-60 min) welcome.

Talk to us in the break or send us an email!

contact@vdlm.at

Vienna Deep Learning Meetup

Next Meetup:

- October 23, 2025

Any suggestions for speakers or venues?

contact@vdlm.at

Send us SPEAKER suggestions, job & event announcements:

contact@vdlm.at