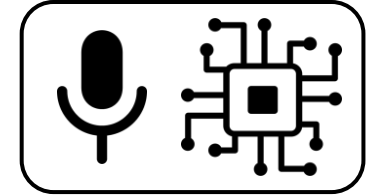


---

# Computational Analysis of Sound and Music

---

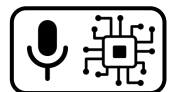


## Research Project Topics

Dr.-Ing. Jakob Abeßer

Fraunhofer IDMT

[jakob.abesser@idmt.fraunhofer.de](mailto:jakob.abesser@idmt.fraunhofer.de)



---

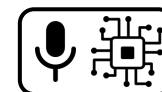
# Research Project



## Topic #1: Sound Event Classification

---

- Dataset
  - ESC-50 datasets (<https://github.com/karolpiczak/ESC-50> )
- Task
  - Classify isolated sound recordings into 50 sound classes
- Aspects to look deeper into
  - Compare different spectrogram representations (STFT, Mel Spectrogram etc.)
  - Data augmentation (<https://github.com/iver56/audiomentations> )



---

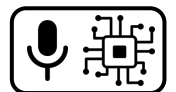
# Research Project

## Topic #2: Vehicle Type Classification

---



- Dataset
  - IDMT-TRAFFIC  
(<https://www.idmt.fraunhofer.de/en/publications/datasets/traffic.html> )
- Tasks
  - Vehicle type classification (bus, car, motorcycle, truck)
  - Movement direction estimation (left > right, right > left)
- Aspects to look deeper into
  - Classify between noisy sound classes
  - How to analyze stereo signals (time-of-arrival differences)



---

# Research Project



## Topic #3: Bird Activity Detection

---

- Dataset
  - warblrb10k dataset (<https://dcase.community/challenge2018/task-bird-audio-detection>) - 2,000 smartphone recordings
- Task
  - Classify a 10s audio recording for bird activity (active / not active)
- Aspects to look deeper into
  - How to deal with large variety of background sounds?
  - Convolutional Neural Networks to learn spectro-temporal patterns (bird vocalizations)



---

# Research Project

## Topic #4: Acoustic Scene Classification

---



- Dataset
  - DCASE-2013-Task1 (<https://dcase.community/challenge2013/task-acoustic-scene-classification>)
- Task
  - Classify the acoustic scene (10 classes) given a 30s binaural audio recording
- Aspects to look deeper into
  - How to summarize long-term characteristics of audio signals?
  - Convolutional Neural Networks



---

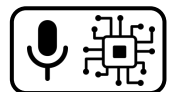
# Research Project

## Topic #5: Music Genre Classification

---



- Dataset
  - FMA-small (<https://github.com/mdeff/fma>) – 8000 30s tracks, 8 genres
- Task
  - Classify the music genre
- Aspects to look deeper into
  - Compare different audio features (rhythm, harmony, timbre)



---

# Research Project

## Topic #6: Music Instrument Classification

---



- Dataset
  - MedleyDB (<https://medleydb.weebly.com/>) – 196 multitracks
- Task
  - Instrument recognition in multitimbral mixtures or classifying individual stems (one instrument active per stem)
- Aspects to look deeper into
  - How robust is instrument recognition vs. #overlapping instruments?
  - How does instrumentation relate to music genre (also annotated)?
    - Co-occurrence matrix



---

# Research Project

## Topic #7: Chord Recognition

---



- Dataset
  - IDMT-SMT-CHORDS  
(<https://www.idmt.fraunhofer.de/en/publications/datasets/chords.html>)
- Task
  - Estimate chord type (3-voiced / 4-voiced chords) from keyboard instruments / guitars
- Aspects to look deeper into
  - Compare classical approach (template matching on chroma vectors) with deep learning based approach (CNN)





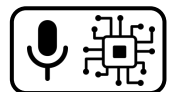
---

# Research Project

## Topic #8: Electric Engine Monitoring

---

- Dataset
  - IDMT-ISA-ELECTRIC-ENGINE (<https://zenodo.org/records/7551261>)
- Task
  - Estimate operational state from audio recording (“good”, “heavy load” and “broken”)
- Aspects to look deeper into
  - Data augmentation
  - Adding ambient background noise to increase robustness



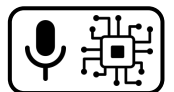
---

# Research Project

## Topic #9: Compressed Air Leakage Detection

---

- Dataset
  - IDMT-ISA-COMPRESSED-AIR (<https://zenodo.org/records/7551606>)
- Task
  - Estimate leak types (3 classes) and noise types (5 classes)
- Aspects to look deeper into
  - Influence of microphone recording setup
  - Background noises



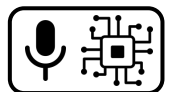
---

# Research Project

## Topic #10: Record-Your-Own-Soundscapes

---

- Dataset
  - Soundscape recordings
- Task
  - Sound Event Detection
  - Annotation using Sonic Visualiser
- Aspects to look deeper into
  - Annotator Agreement
  - Background Noises
  - Temporal long-term structure of audio recordings



---

# Research Project

## Sources for Audio Datasets

	MIR	Env. Sounds
▪ <a href="https://www.audiocontentanalysis.org/datasets.html">https://www.audiocontentanalysis.org/datasets.html</a>	✓	
▪ <a href="https://ismir.net/resources/datasets/">https://ismir.net/resources/datasets/</a>	✓	✓
▪ <a href="https://www.idmt.fraunhofer.de/en/publications/datasets.html">https://www.idmt.fraunhofer.de/en/publications/datasets.html</a>	✓	✓
▪ <a href="https://zenodo.org">https://zenodo.org</a>	✓	✓
▪ <a href="https://homepages.tuni.fi/toni.heittola/datasets">https://homepages.tuni.fi/toni.heittola/datasets</a>		✓
▪ <a href="https://towardsdatascience.com/40-open-source-audio-datasets-for-ml-59dc39d48f06">https://towardsdatascience.com/40-open-source-audio-datasets-for-ml-59dc39d48f06</a>	✓	

