

# MACHINE LISTENING – KI-BASIERTES HÖREN

Jakob Abeßer

Fraunhofer IDMT



2. Thüringer KI-Forum, 07.12.2020

---

# AGENDA

---

- Motivation
- KI-basierte Klangerkennung
- Anwendungsszenarien
  - Städtische Lärmüberwachung
  - Verkehrsmessung
  - Überwachung von Fertigungsstrecken

# Motivation

- Hören ist wichtiger Teil der menschlichen Wahrnehmung
  - Klänge tragen Information
  - Multi-modalität
- Geräuschwahrnehmung
  - Erkennung / Lokalisation von akustischen Ereignissen
  - Zuordnung zu Klangquellen
  - Verständnis komplexer akustischer Szenen



# Umweltgeräusche

- Klangquellen
  - Natur, Wetter, Mensch, Maschine, etc.
- Klangcharakteristika
  - Strukturiert / unstrukturiert
  - Stationär / unvorhersehbar
  - Zufällig / vorhersehbar & wiederholend
- Klangdauer
  - Sehr kurz (Pistolenschuss, Türklopfen, Schrei)
  - Sehr lang (laufende Maschinen, Wind, Regen)



# KI-basierte Klangerkennung

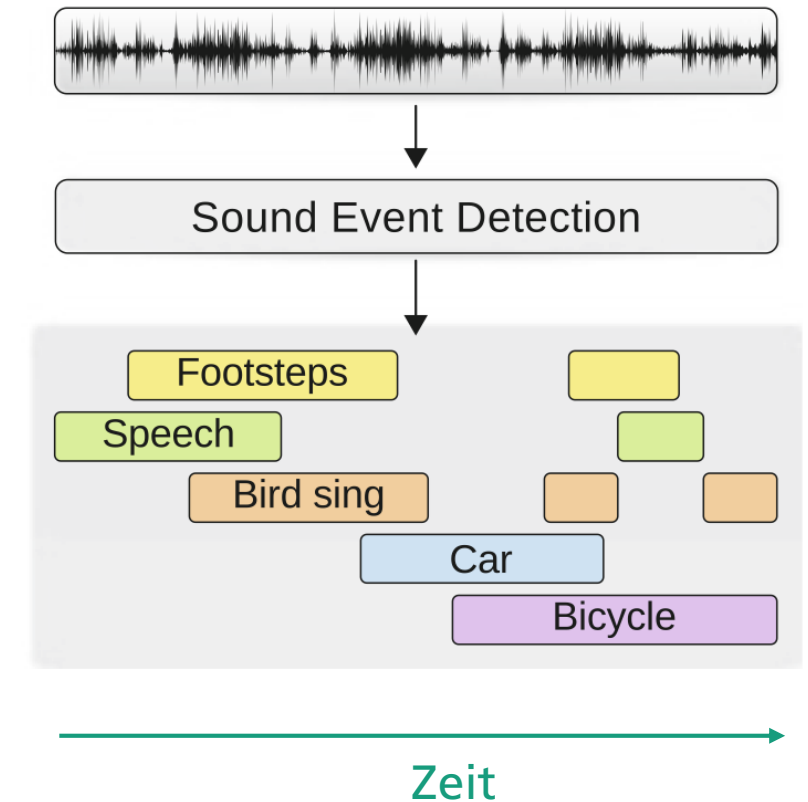
## Anwendungsszenarien

- Assistenzsysteme (Hörgerät, Handy)
  - Kontexterkenkung (indoor / outdoor / vehicle)
- Überwachungsinfrastruktur
  - Erkennung von Glasbruch, Pistolenschuss, Schreie, Hilferufe
- Verkehrsüberwachung
  - Erkennung & Klassifikation vorbeifahrender Fahrzeuge
- Automatische Qualitätskontrolle
  - End-of-line Testing, Predictive Maintenance

# KI-basierte Klangerkennung

## Aufgabenstellung

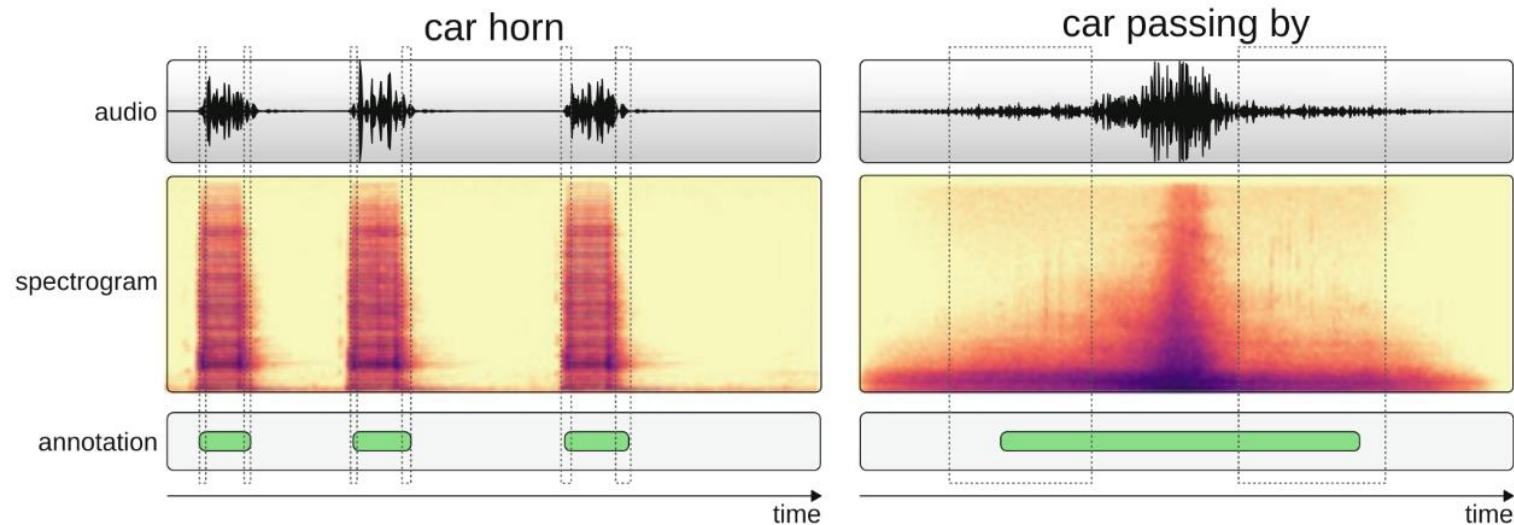
- Akustische Ereignisdetektion
  - Segmentierung
  - Klassifikation
- Unterschiedlicher „Polyphoniegrad“



# KI-basierte Klangerkennung

## Herausforderungen

- Vielzahl an
  - Klangcharakteristika
  - Klangdauern
- Start- und Endzeitpunkte von Geräuschen oft ungenau definiert

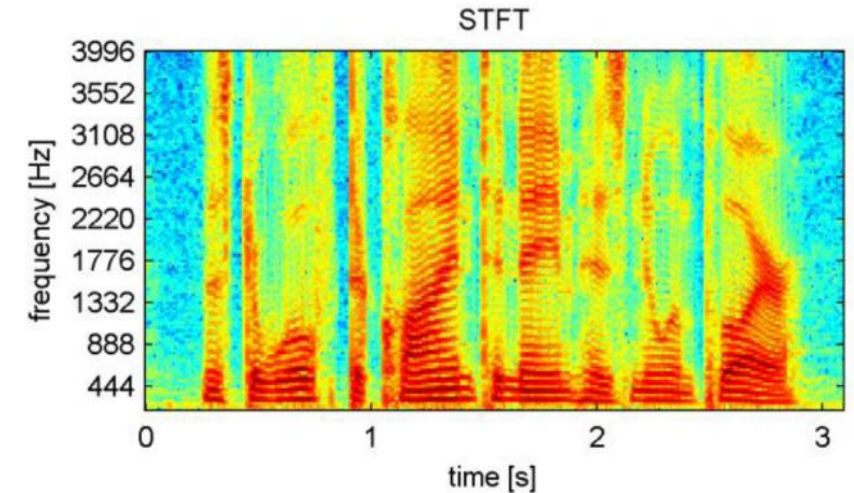




# KI-basierte Klangerkennung

## Herausforderungen

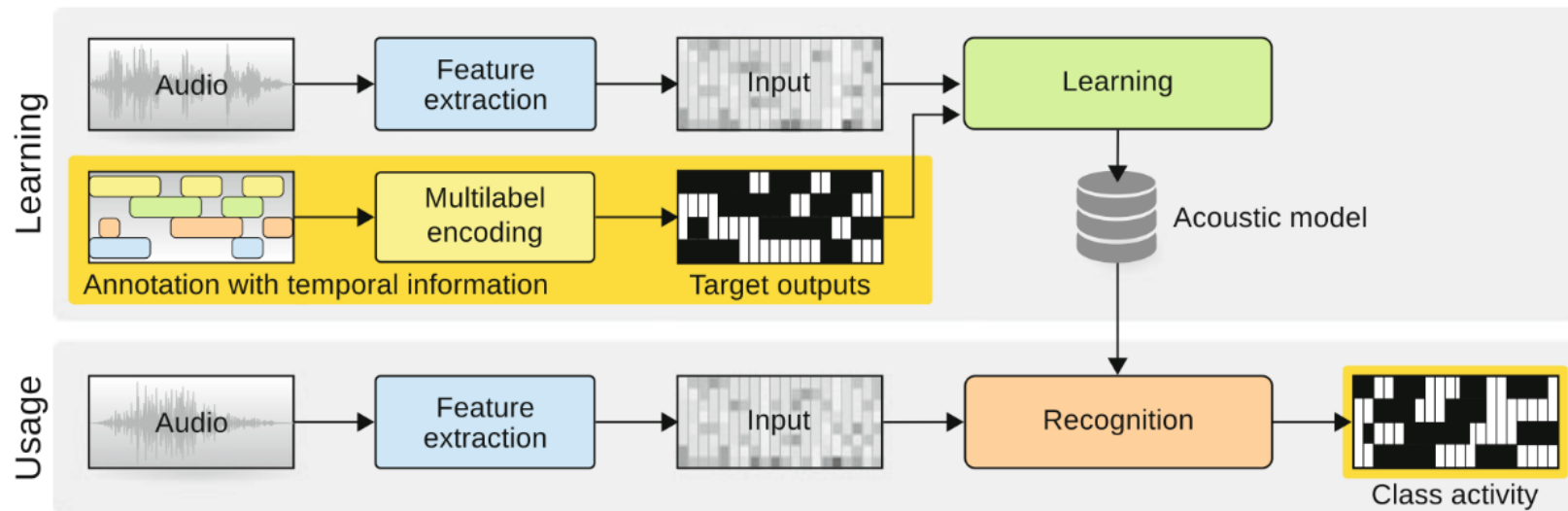
- Klänge im Vordergrund & Hintergrund
  - Abhängig von relativer Position der Quelle
- Klänge sind „transparent“
  - Phasen-abhängige Überlagerung
  - Mögliche Signalauslöschung
- Nicht-lokale Energieverteilung in Spektrogrammen





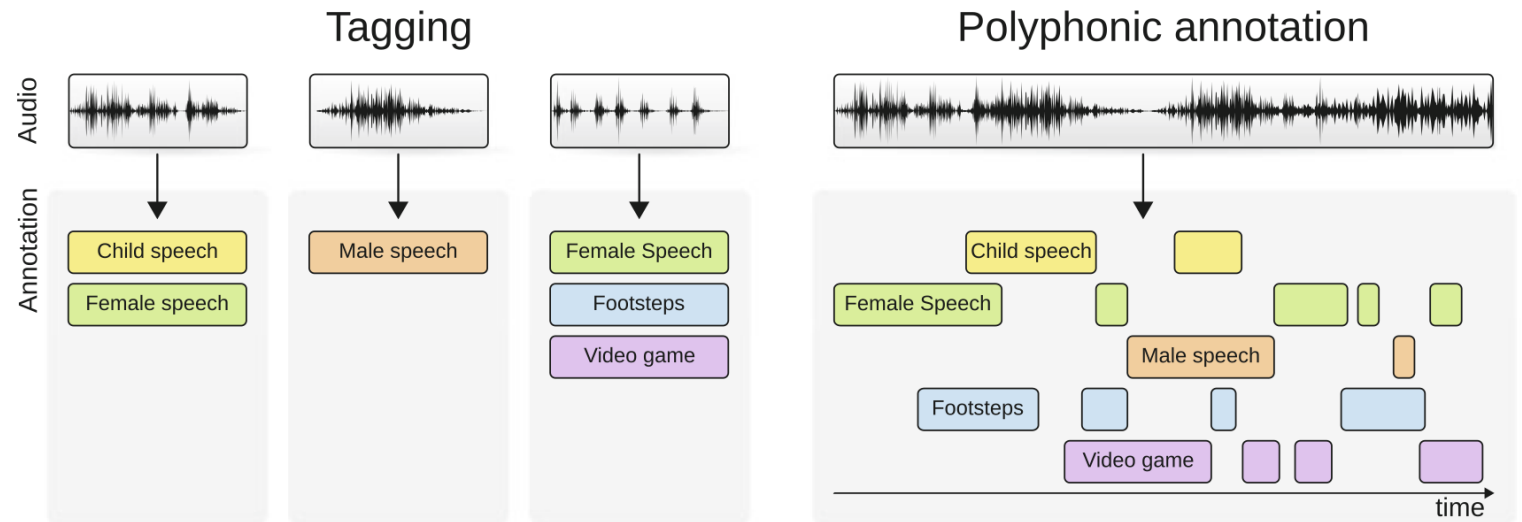
# KI-basierte Klangerkennung Pipeline

- Überwachtes Lernen
  - Merkmalsextraktion aus Audioaufnahmen (i.d.R. Spektrogramm)
  - Annotation / Labels
  - KI-Modellbildung



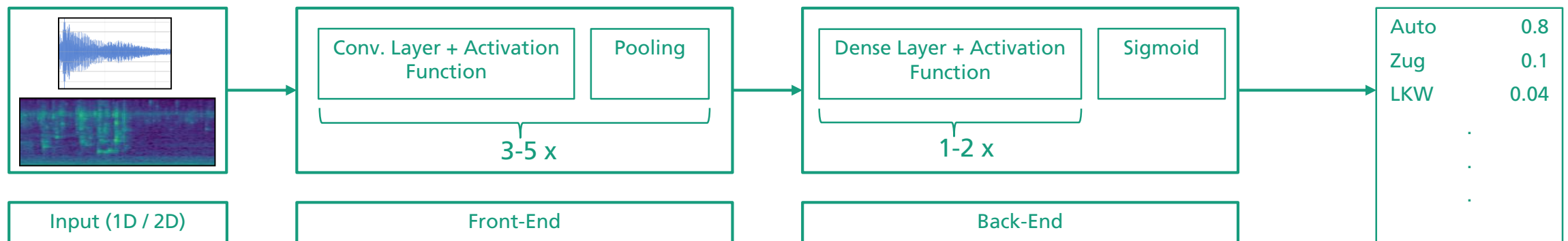
# KI-basierte Klangerkennung Pipeline

- Annotation
  - Granularität
    - Globale Annotation („weak“ labels) → günstig
    - Segmentannotation („strong“ labels) → teuer



# KI-basierte Klangerkennung Pipeline

- KI-Modellbildung
  - Convolutional Neural Networks (CNN) / Convolutional Recurrent Neural Networks (CRNN)
    - Front-End
      - Lernen von Merkmalen
    - Back-End
      - Zeitliche Modellierung & Klassifikation von Geräuschen

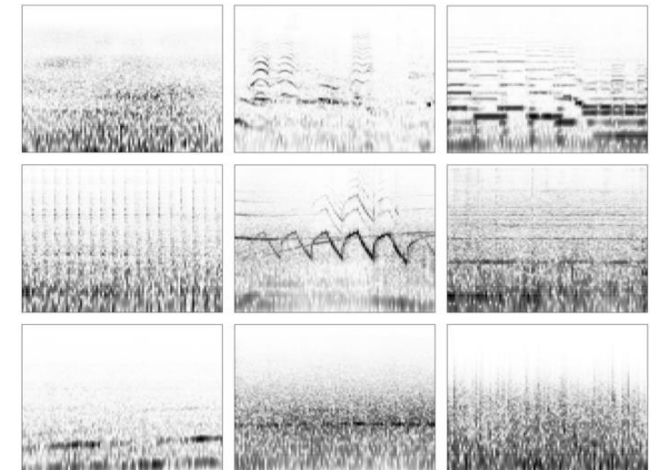


# Anwendungsszenarien

## Städtische Lärmüberwachung



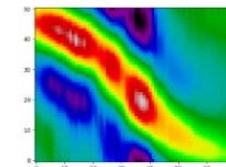
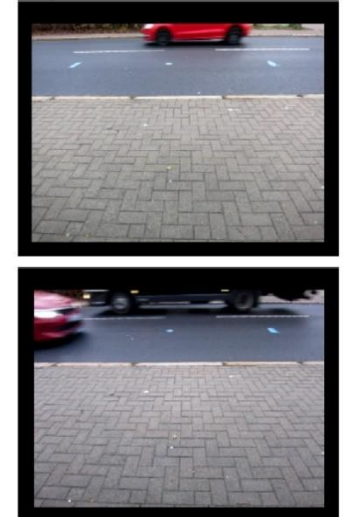
- ZIM-Projekt „Stadtlärm“ (2016-2018)
  - Fraunhofer IDMT, Institut für Mikroelektronik- und Mechatronik-Systeme gemeinnützige GmbH (IMMS), Software-Service John GmbH, Bischoff Elektronik GmbH
- Ziel
  - Verteiltes akustisches Sensornetzwerk in Jena
  - Messung von Lärmpegeln & Identifikation von Lärmquellen
  - Mobile Sensoreinheiten (Raspberry Pi 3 Plattform, MEMS-Mikrofon, Akku)
- Akustische Ereigniserkennung + Pegelmessung
  - 9 Klassen (Auto, Gespräch, Musik, Baustelle, Sirene, Zug, Straßenbahn, LKW, Wind)



# Anwendungsszenarien

## Verkehrsmessung

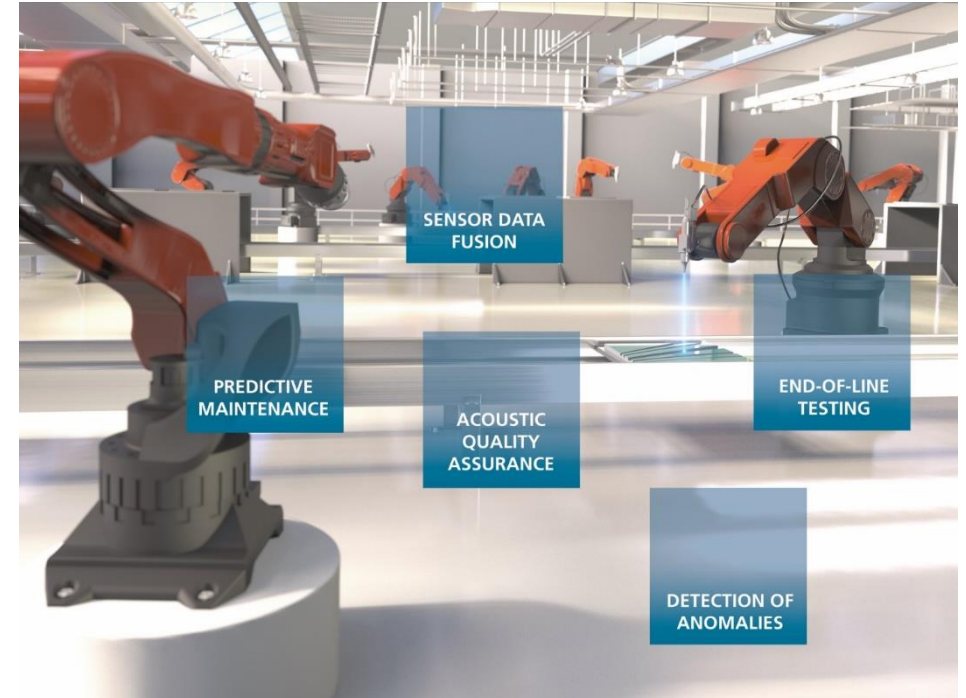
- Erweiterung des Stadtlärm-Sensors (stereo)
- Auswertung von Klang & Laufzeitunterschieden
- Multitask Learning
  - Detektion
  - Klassifikation (PKW, LKW, Bus, Motorrad)
  - Richtungs- & Geschwindigkeitsschätzung
- KI-Modelle
  - CNN-Modelle
  - Fokus auf energieeffiziente KI-Modelle (Miniaturisierung)



Korrelation zwischen Stereo-Kanälen für Bewegung links nach rechts

# Anwendungsszenarien Überwachung von Fertigungsstrecken

- Herausforderungen
  - Echtzeitfähige Klanganalyse im industriellen Umfeld
  - Energie-effiziente KI-Algorithmen
  - Klangvariation durch Maschinenzustände
  - Akustische Anomalien eher “subtil”





# Vielen Dank für Ihre Aufmerksamkeit!



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