



Motivation/ Problemstellung

Ziel :

- automatisierte Geräuscherkennung in Video-Tonspuren
- Ermöglicht Untertitel um Geräusch-Untertitel zu ergänzen

Kontext :

- Kooperation des SWR mit dem IAAI

Motivation :

- Untertiteln von Geräuschen für den SWR weitest gehend automatisieren
- Einsparen von Zeit und Aufwand im Tagesgeschäft
- Unterhaltungsprogramm des SWR möglichst barrierearm zugänglich machen

Umsetzung/ Lösung

- Python
- Jupyter Lab
- PyTorch
- Sklearn
- Pandas
- Etc.



Team



Anh Thu Bui

Medieninformatik

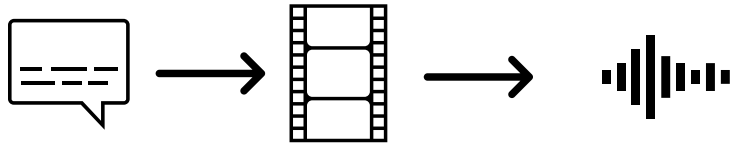
6. Semester

Natalie Hußfeldt

Medieninformatik

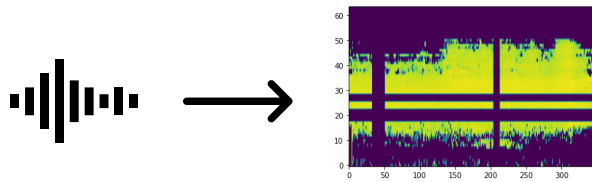
6. Semester

Umsetzung



1. Extrahieren der Geräusche aus Videos und Labeln der Daten

```
</tt:p>, <tt:p begin="00:00:12.120" end="00:00:14.120" region="bottom" style="textCenter" xml:id="sub5">  
<tt:span style="textWhite">(Lebhafte Musik)</tt:span>  
</tt:p>, <tt:p begin="00:00:12.120" end="00:00:14.120" region="bottom" style="textCenter" xml:id="sub5">  
<tt:span style="textWhite">(Lebhafte Musik)</tt:span>  
</tt:p>, <tt:p begin="00:00:43.160" end="00:00:44.720" region="bottom" style="textCenter" xml:id="sub12">  
<tt:span style="textWhite">(Treibende Musik)</tt:span>  
</tt:p>, <tt:p begin="00:00:43.160" end="00:00:44.720" region="bottom" style="textCenter" xml:id="sub12">
```



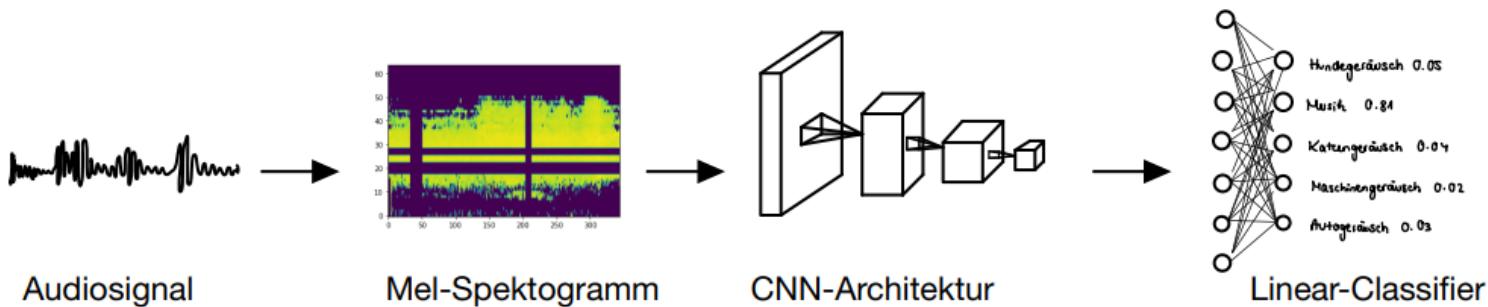
2. Preprocessing und Augmentation der Daten

Raw Audio

- Resampling
- Rechanneling
- Anpassung der Dauer
- Timeshift

Spektrogramm

- Erstellen Mel-Spektrogramme
- SpecAugment



3. Aufbau und Training eines Classifiers