

Acoustic Scene and Room Classification for Real-Time Applications

Silvio Emmenegger

Final Presentation, Master Thesis FS20

July 1, 2020

Acoustic environments in daily life



Column 2

Column 2

Livingroom/Music

Lucerne, Home



Sources: <https://bit.ly/31p4k32>

Intention

- Creation of acoustic classifier system for scenes and soundscapes
- Classification at the edge in real-time → hearing aids
- Use of latest AI techniques → Convolutional Neural Networks (CNN)

Intention

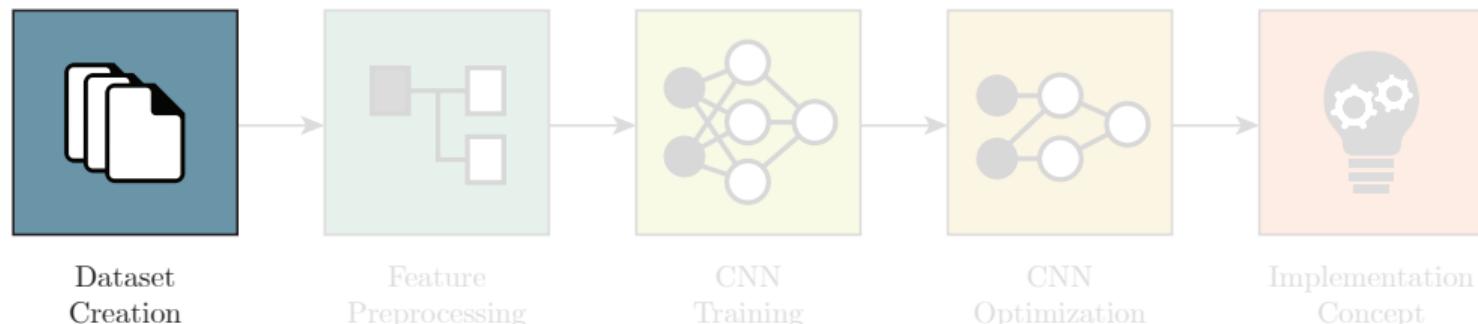
- Creation of acoustic classifier system for scenes and soundscapes
- Classification at the edge in real-time → hearing aids
- Use of latest AI techniques → Convolutional Neural Networks (CNN)

Intention

- Creation of acoustic classifier system for scenes and soundscapes
- Classification at the edge in real-time → hearing aids
- Use of latest AI techniques → Convolutional Neural Networks (CNN)

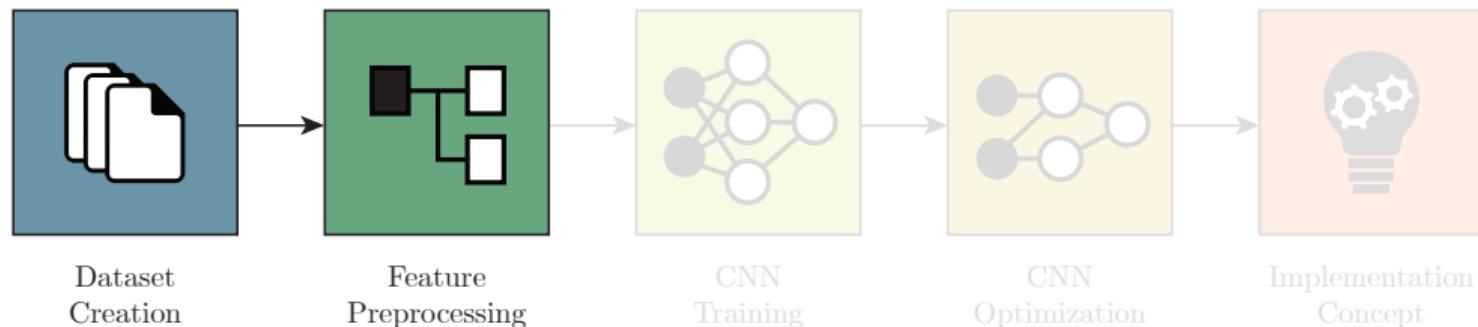
Tasks

- Record new audio dataset and labels with high-quality equipment
- Preprocessing of audio into image-like features
- Train CNN model on created feature dataset
- Optimize CNN model with simple grid search method
- Propose real-time implementation concept for dedicated hardware



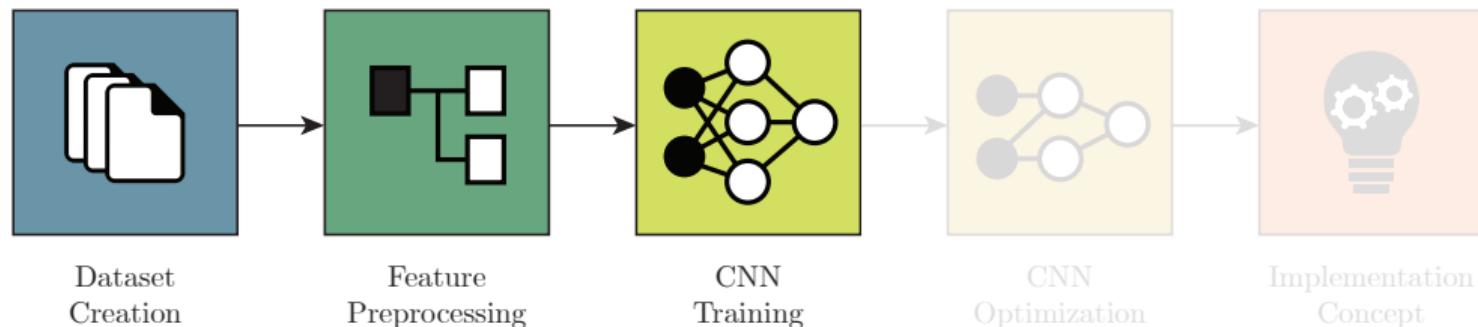
Tasks

- Record new audio dataset and labels with high-quality equipment
- Preprocessing of audio into image-like features
- Train CNN model on created feature dataset
- Optimize CNN model with simple grid search method
- Propose real-time implementation concept for dedicated hardware



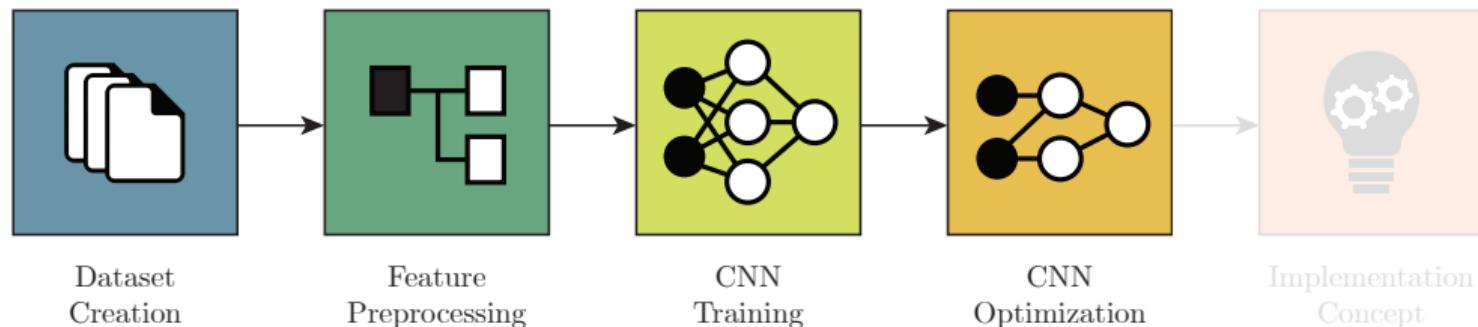
Tasks

- Record new audio dataset and labels with high-quality equipment
- Preprocessing of audio into image-like features
- Train CNN model on created feature dataset
- Optimize CNN model with simple grid search method
- Propose real-time implementation concept for dedicated hardware



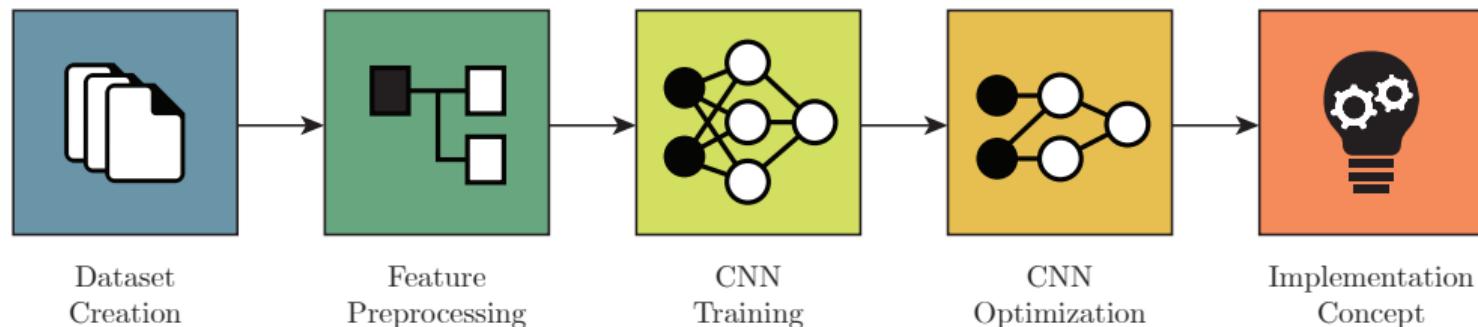
Tasks

- Record new audio dataset and labels with high-quality equipment
- Preprocessing of audio into image-like features
- Train CNN model on created feature dataset
- Optimize CNN model with simple grid search method
- Propose real-time implementation concept for dedicated hardware



Tasks

- Record new audio dataset and labels with high-quality equipment
- Preprocessing of audio into image-like features
- Train CNN model on created feature dataset
- Optimize CNN model with simple grid search method
- Propose real-time implementation concept for dedicated hardware



Dataset Creation

- Two labels with five classes each („2D“ label combinations)
- Multi-class multi-output classification problem, supervised
- Binaural recordings with $f_s = 48$ kHz and 16-bit quantization learning
- At least 60 min for each label combination → 25h in total

Scenes	Soundscapes
Cafeteria/Canteen	Music
Car (inside)	Noise
Livingroom	Speech
Nature	Noisy Speech
Street	Quiet

Dataset Creation

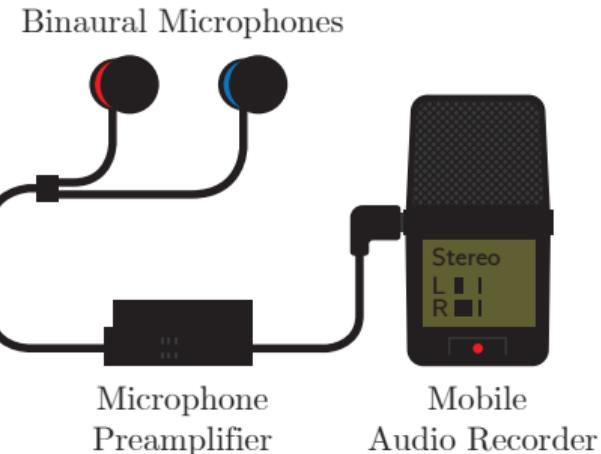
- Two labels with five classes each („2D“ label combinations)
- Multi-class multi-output classification problem, supervised
- Binaural recordings with $f_s = 48$ kHz and 16-bit quantization learning
- At least 60 min for each label combination → 25h in total

Scenes	Soundscapes
Cafeteria/Canteen	Music
Car (inside)	Noise
Livingroom	Speech
Nature	Noisy Speech
Street	Quiet

Dataset Creation

- Two labels with five classes each („2D“ label combinations)
- Multi-class multi-output classification problem, supervised
- Binaural recordings with $f_s = 48$ kHz and 16-bit quantization learning
- At least 60 min for each label combination → 25h in total

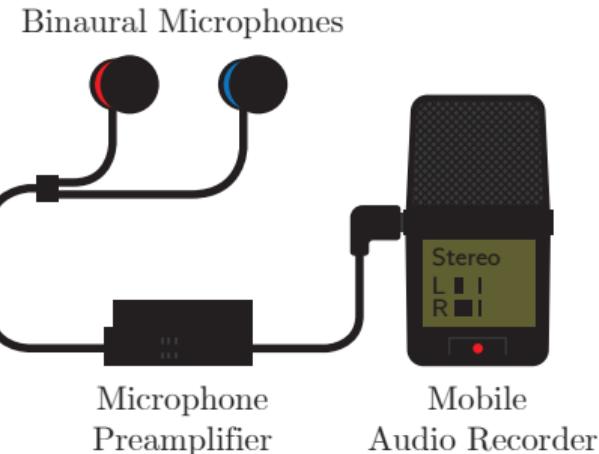
Scenes	Soundscapes
Cafeteria/Canteen	Music
Car (inside)	Noise
Livingroom	Speech
Nature	Noisy Speech
Street	Quiet



Dataset Creation

- Two labels with five classes each („2D“ label combinations)
- Multi-class multi-output classification problem, supervised
- Binaural recordings with $f_s = 48$ kHz and 16-bit quantization learning
- At least 60 min for each label combination → 25h in total

Scenes	Soundscapes
Cafeteria/Canteen	Music
Car (inside)	Noise
Livingroom	Speech
Nature	Noisy Speech
Street	Quiet



Demonstrator

Questions & Discussion



Backup Slides



Bibliography



Backup Slides 1 Ref. [?]