Universität Bremen

Cognitive Systems Lab

Bachelor/Master Thesis Topics Cognitive Systems Lab

Summer Term 2024

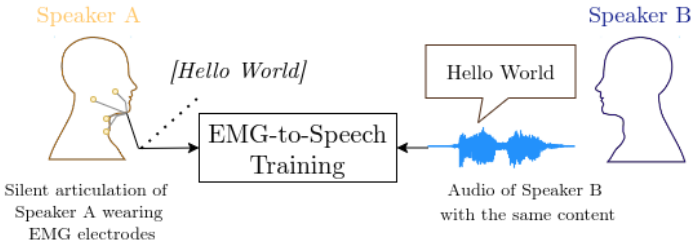
Speech Synthesis from Electromyography

Master‘s Thesis: Giving People back their Voice - Construction of an EMG-to-Speech Data Set for Voice Donation

Electromyography (EMG) signals capture the activation of muscles, such as the articulatory muscles during speech production. They are thus researched as a modality for a voice prosthesis with which subjects can speak „silently“ by moving their articulators. The main research focus are **EMG-to-Speech** models, which aim to predict acoustic speech from EMG signals

- → Current EMG-to-Speech models use the audio recordings of the person wearing EMG sensors as training target – this is inapplicable for speakers with speech impairments who cannot produce audible speech.
- → We want to use the audio of other speakers as targets - they act as "voice donor" for the person wearing the EMG sensors. For this, we aim to build a suitable data set.

- → **Main task of the thesis**: Setup a multi-speaker data set for EMG-to-Speech training with focus on using audio targets of different speakers.
- → Evaluate alignment techniques based on deep learning to align silent EMG of one speaker with the audio targets of another speaker for EMG-to-speech model training.



- → Requirements: Very good Python skills and deep learning knowledge
- → Nice to have: Basic knowledge of audio signal processing, completed the Advanced ML course
- → Contact: Kevin Scheck, mail to scheck@uni-bremen.de;
  Please add a small introduction about yourself and your CV or grades transcript.

Speech Synthesis from Electromyography

Master's Thesis: A Differentiable Frontend for Adaptable Speech Synthesis from Electromyography

Electromyography (EMG) signals capture the activation of muscles, such as the articulatory muscles during speech production. They are thus researched as a modality for a voice prosthesis with which subjects can speak „silently“ by moving their articulators.

- → The voice reconstruction is carried out by **EMG-to-Speech** models - Neural Networks (NNs) which map EMG signals to acoustic speech features.
- → Current methods pre-process the EMG signal using standard filters, e.g., a band-pass filter and a notch filter, before they are passed to the NNs.
- → This might make the EMG pre-processing sub-optimal for the given task and might make the system less adaptable.

- → **Main task of the thesis**: Implement and evaluate a differentiable EMG pre-processing frontend based on NNs. For instance, SincNets, data-driven normalization techniques etc.
- → Evaluation of learnable causal filters vs. non-causal filters for real-time vs. offline performance.
- → Evaluation on the adaptability of the frontend to new EMG recording setups which might change the signal properties.

Requirements: Good Python skills, signal processing knowledge and deep learning knowledge

- → Nice to have: Completed the Advanced ML course
- → Contact: Kevin Scheck, mail to scheck@uni-bremen.de; Please add a small introduction about yourself and your CV or grades transcript.

Machine Listening

Interesting machine learning topics on the disentanglement of speech and sounds (BA and MA theses)

1. The role of harmonic convolutions for the disentanglement of speech
2. Focus on what you dislike: Inverted selective auditory attention
3. Can video frame inpainting techniques enhance speech separation?
4. Can multilingual training data increase the model‘s robustness against noise?
5. Speaker embedding fusion for extracting multiple target speakers

Source: https://www.vocal.com/wp-content/uploads/2019/09/cocktail-party-problem.jpg

Other topics on machine listening → Just text me!

General requirements: Good knowledge in Python, at least basic knowledge in machine learning, motivation, reliability, personal responsibility, willingness to learn new topics

We offer: Intensive support, regular meetings, feedback discussions, nice work environment

Contact: Marvin Borsdorf ☏ 0421-218-64291 ✉ marvin.borsdorf@uni-bremen.de

HateSpeech: Automatic Data Collection, Annotation and Classification (BA/MA Thesis)

→ **Tasks**:

- *Automatic Hate Speech Data Collection*: Gather hate speech data from various social media platforms.
- *Building a Data Annotation Pipeline*: Develop an efficient pipeline for annotating collected data.
- *Classification of Hate Speech Data*: Apply machine learning techniques to classify hate speech content accurately

→ **Requirements**:

- Good knowledge in Python
- Basic understanding of Machine Learning
- High motivation, curiosity, and effective communication skills.
- Proficiency in both German and English languages.

If you find any of the above tasks interesting and meet the requirements, please contact

Rathi Adarshi Rammohan, rammohan@uni-bremen.de

For more details about the project: Hate Speech

Working with ILSE's conversational speech data

→ Possible BA or MA topics:

- *Source separation on ILSE interviews*,
  i.e. separate audio from two speakers for subsequent speaker-dependent ASR and further analyses
- *Analyze semantic embeddings of ILSE interviews*

→ Requirements:

- Good Python programming skills
- At least basic ML/ASR knowledge

→ Contact: Elisa Brauße (elisa.brausse@uni-bremen.de)
with short introduction and CV

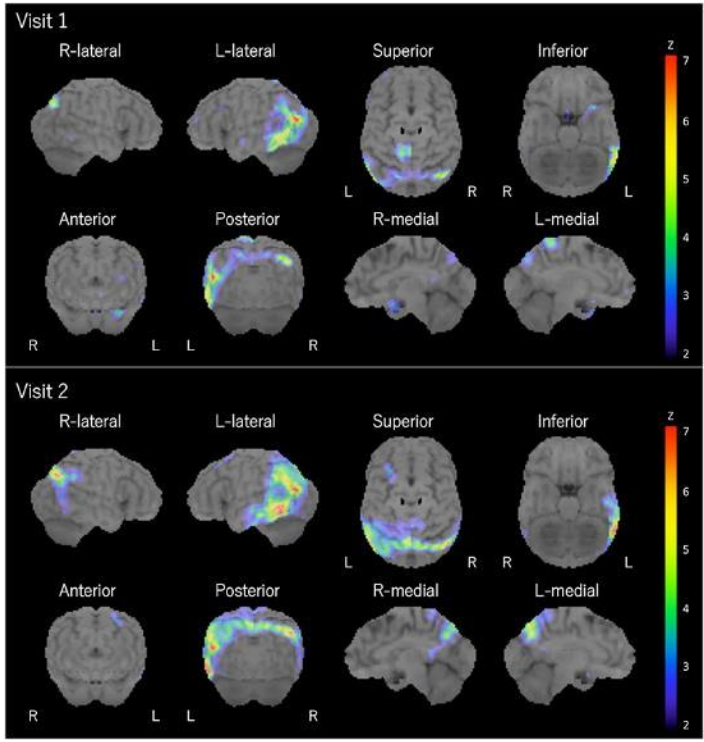
ML on ILSE's Imaging Data

→ Motto: ILSE does not only offer speech data!
→ Pretty loose topic with lots of freedom for brainy students, Bachelor's or Master's thesis possible

→ Your tasks:
- Work with Magnetic Resonance Imaging (MRI) data
- Implement ML models for MRI-based dementia classification (depression diagnoses also available)

→ Requirements:
- Good Python programming skills
- At least basic ML knowledge

→ Contact: Elisa Brauße (elisa.brausse@uni-bremen.de) with short introduction and CV

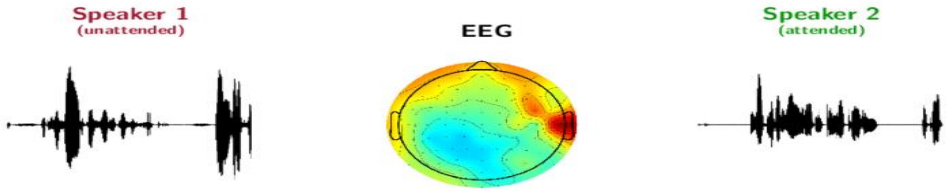


Selective Auditory Attention – Bachelor‘s Thesis

Interesting thesis topics on Solving the Cocktail Party‘s Problem using EEG

→ Topic: **Source Sound Decoding with Electroencephalogram** (Auditory Attention Decoding)

- Understanding of Auditory Attention Decoding and decoding source sound using eeg samples(with public datasets)
- Preprocess available data and fed it to machine learning/Neural Networks(Train different models) to classify the attended speaker
- Visualize results through various methods of applicable Neural Networks and compare results
- Classify performance based upon different validation measures & specify optimal way out



General requirements: Good knowledge in Python, at least basic knowledge in machine learning(Deep learning), motivation, reliability, personal responsibility, willingness to learn new topics

We offer: Intensive support, regular/weekly meetings, feedback discussions, nice work environment

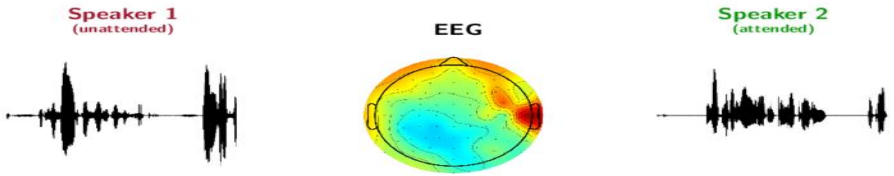
Contact: Saurav Pahuja ☏ 0421-218-64282 ✉ saurav@uni-bremen.de

Selective Auditory Attention – Master‘s Thesis

Interesting thesis topics on Solving the Cocktail Party‘s Problem using EEG

→ Topic: **Brain Tracking of Sound using Deep Learning** (Auditory Attention Decoding)

- Understanding of Auditory Attention Decoding and decoding source sound using eeg samples (public datasets) with Different Optimized Deep learning Models(CNN, GCN, Hybrid Models…)
- Generate Augmented Data using GAN… & check performance with other networks using original data and new generated data
- Try to interpret & Visualize results of different applicable Neural Networks and compare results based on EEG samples
- Do Ablation Analysis over different Networks and see through performance, various different visualizations the effects of Auditory Attention Decoding on Brain

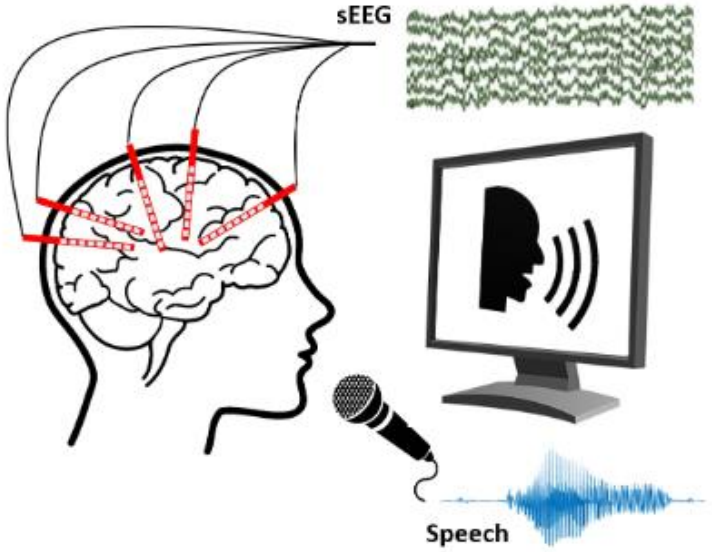


General requirements: Good knowledge in Python, at least basic knowledge in Deep learning, motivation, reliability, personal responsibility, willingness to learn new topics, some experience with EEG would be nice!
We offer: Intensive support, regular/weekly meetings, feedback discussions, nice work environment
Contact: Saurav Pahuja ☏ 0421-218-64282 ✉ saurav@uni-bremen.de

Speech Representations in the Brain

- stereotactical Electroencephalography (sEEG)
  → invasive method to measure neural signals

- Time-Aligned Speech and sEEG signals
  → Which representations of speech can be found in the neural signal? (Articulatory/Linguistic/Spectral…)
  → Build models for different representations and compare



- Requirements:
  → Decent Python programming skills
  → Basic knowledge in Machine Learning
- Nice to have:
  → Interest in Neuroscience/Speech
  → Experience in Machine Learning

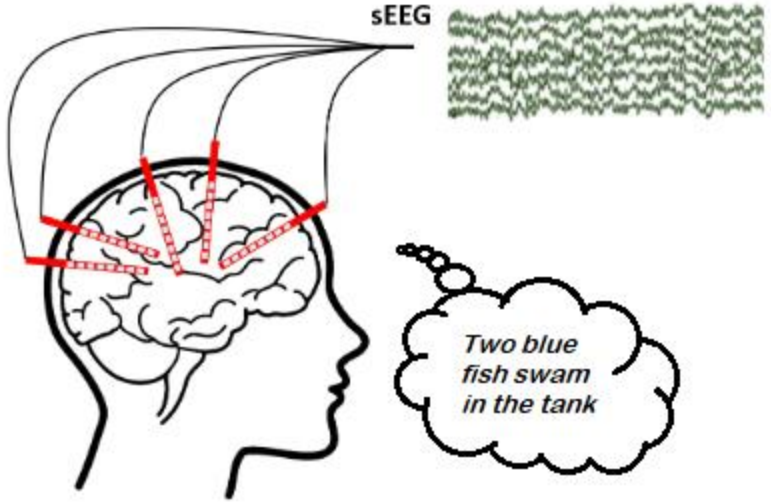
- We offer:
  →Bachelor/Master thesis
  →Extensive support, weekly meetings, infra-structure… and high interest in the results!

Contact: Darius Ivucic (darivuci@uni-bremen.de)

Imagined Speech decoding

- stereotactical Electroencephalography (sEEG)
  → invasive method to measure neural signals

- sEEG signals recorded during imagined speech
  → Decode speech features in real-time
  → Problem: No direct audio labels!
  → Try surrogate-labels (Text, vocalization,TTS...) for speech feature reconstruction



- Requirements:
  → Decent Python programming skills
  → Basic knowledge in Machine Learning
- Nice to have:
  → Interest in Neuroscience/Speech
  → Experience in Machine Learning

- We offer:
  →Bachelor/Master thesis
  →Extensive support, weekly meetings, infra-structure… and high interest in the results!

Contact: Darius Ivucic (darivuci@uni-bremen.de)

Self-report based Attention Modeling

Contact: Dr. Mazen Salous (salous@uni-bremen.de)

- **Project scope:** attention-sensitive smart helmet
- **Thesis target:** multimodal data attention modeling

- **Research Question:** How worthy can be the self-reported attention during cycling for powerful multimodal machine learning models?

- **Your tasks:** Contribute to the currently running self-report attention user study
  1. Questionnaire for attention distractions during the experiments
  2. Multimodal data analysis: To show potential correlations between the self-reported attention distractions and the recorded multimodal data
  3. Prepare an attention Ground-Truth from the recorded multimodal data (feature vectors) and the self-reported distractions (classes)
  - Optional (Bonus): Investigate machine learning (ML) models with your prepared attention ground-truth in (3).

- **We offer:**
  - Continuous friendly support: in field, analysis and ML models
  - Lab access, software and hardware for the user study.

Probes "Attentive?"
&
Self-Reported Distractions

Video reviewing with participant for self-reported distractions: internal mental, external visual/auditory

AR and ML based Attention Support in SmartHelm

- **Project scope:** attention-sensitive smart helmet
- **Thesis target:** Draw attention to critical moving objects

- **Research Question:** How valuable are AR and Trajectory Prediction for cyclists' safety?

- **Your tasks:** Contribute to the currently running self-report attention user study
  1. Questionnaire for critical moving objects during the experiments
  2. Video analysis: To show potential critical moving objects
  3. Trajectory Prediction: Show AR attention alerts!
  - Optional (Bonus): Investigate ML models for Eye-Tracking based attention distraction detection.

- **We offer:**
  - Continuous friendly support: in field, analysis and ML models
  - Lab access, software and hardware for the user study.

Contact: Dr. Dennis Küster (dkuester@uni-bremen.de)

Human activity recognition

- Objective: Activity recognition for elderly adults from accelerometer and gyroscope signals, collected using smart phone sensors.
- Bachelor Thesis, can be extended to Master Thesis topic
- **Main Tasks**:
  - Annotation of activities from continuous data
  - Training a model for activity recognition

- **Requirements**:
  - Deep Learning/Machine Learning
  - Python skills

- Start Date: After discussion
- **Interested in this topic and to know more topics contact,**

Rinu Elizabeth Paul, rinueliz@uni-bremen.de

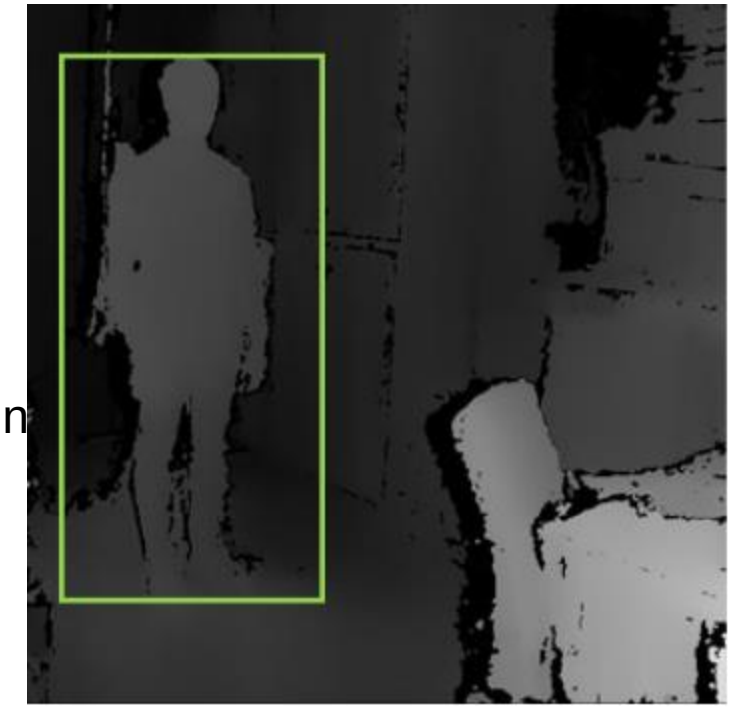
Skeleton extraction with joint prediction

- Objective is to extract skeleton from depth images and predict the missing joints in the skeleton using previous images
- Master Thesis

- Main Tasks:
  - Develop a model for skeleton joint prediction
  - Train model for skeleton joint prediction
  - Train a model for skeleton extraction with joint prediction

- Requirements:
  - Deep Learning, Machine Learning
  - Python skills

- Start Date: After discussion
- Interested in this topic and to know more topics contact,
  Rinu Elizabeth Paul, rinueliz@uni-bremen.de



Example image (from MP-3DHP dataset) with Human detected in room

Fall detection

- Objective: Fall detection on depth dataset and detect fall as outlier from other activities
- Master/Bachelor Thesis

- **Main Tasks**:
  - Identification fall as anomaly
  - Develop a model for fall detection

- **Requirements**:
  - Deep Learning, Machine Learning
  - Python skills

Fall; Image from [1]

- Start Date: After discussion
- **Interested in this topic and to know more topics contact,**

Rinu Elizabeth Paul, rinueliz@uni-bremen.de

[1] “Fall Detection Subscription,” *HandsFree Health*. https://shop.handsfreehealth.com/product/fall-detection-subscription/.

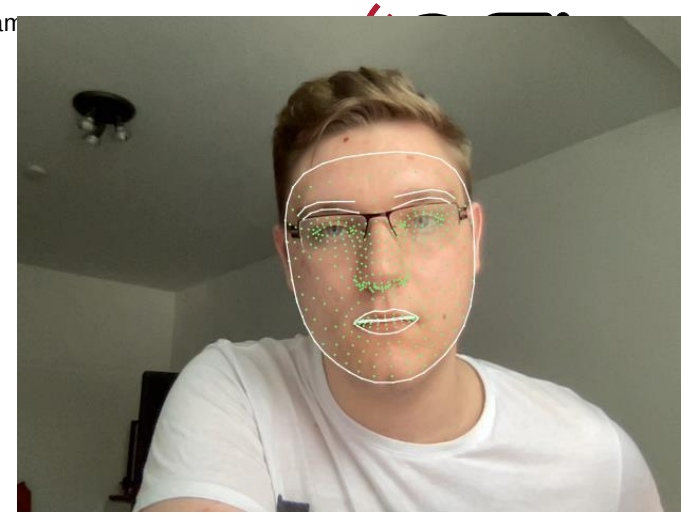
Bachelor Theses:
Computer Vision ML for Modeling Human-Avatar Interactions

Setting:
- We explore **social resonance** by studying how being "synchronized" with an avatars' face, influences human facial responses to avatars.
- Part of interdisciplinary DFG project '**The Active Self**' (https://www.uni-bremen.de/en/csl/projects/current-projects/ccsr & www.activeself.de)
- Data will be provided & feedback on a regular basis

Tasks:
- Comparison of **Computer Vision-based machine learning** (ML) approaches to classify & analyze human responses to online avatars
- **ML approaches to predict human behavior** across different time phases (e.g., starting with regression models)

Other:
- The focus of these BA theses is flexible depending on which of the above tasks are most interesting to you!
- **Interested? Contact: abdulhaq@uni-Bremen.de , verschoo@uni-bremen.de**
- **Please approach Stephan Verschoor!**

Master Theses: Multimodal Modeling of Human-Avatar Interactions

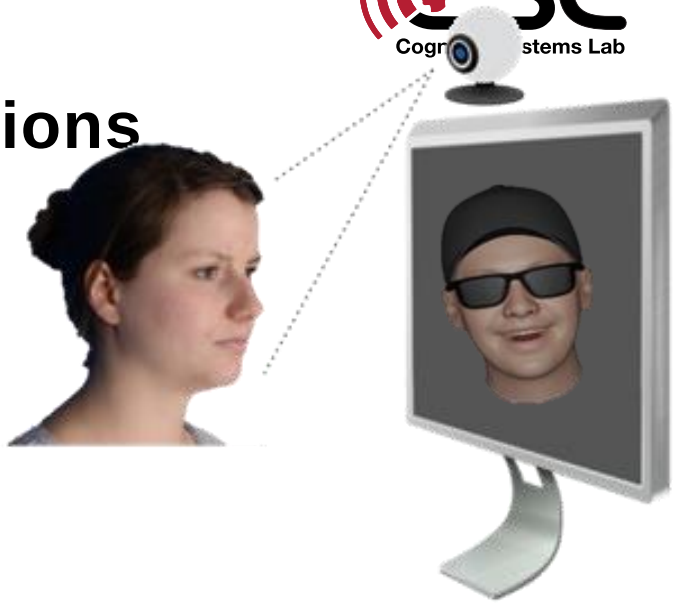


Setting:
- We explore **social resonance** by studying how being “synchronized” with an avatars’ face, influences human facial responses to avatars.
- Part of interdisciplinary DFG project ‘**The Active Self**’ (https://www.uni-bremen.de/en/csl/projects/current-projects/ccsr & www.activeself.de)
- Data will be provided & feedback on a regular basis

Tasks:
- Vision & speech-based machine learning (ML) to classify and analyze human responses to online avatars
- **Comparison of ML approaches** to predict human behavior across different time phases (starting with regression models up to more sophisticated ML approaches such as **neural networks** (LSTMs))

Other:
- The focus of these MA theses is flexible depending on which of the above tasks are most interesting to you!
- **Interested? Contact: abdulhaq@uni-Bremen.de , verschoo@uni-bremen.de**
- **Please approach Stephan Verschoor!**

Student Assistent:
Multi modal ML for Modeling Human-Avatar Interactions



Setting:
- We explore **social resonance** by studying how being “synchronized” with an avatars’ face, influences human facial responses to avatars.
- Part of interdisciplinary DFG project ‘**The Active Self**’ (https://www.uni-bremen.de/en/csl/projects/current-projects/ccsr & www.activeself.de)

Tasks:
- Implementation of machine learning (ML) approaches to classify and analyze and predict human responses to avatars
- Testing and further development of our current Interactive Online Experimentation Platform
- Assist in data acquisition

Other:
- **Interested? Contact: abdulhaq@uni-Bremen.de , verschoo@uni-bremen.de**
- **Please approach Stephan Verschoor!**
- **Win win situation possible! Combine Ba/Ma thesis with getting paid!!**

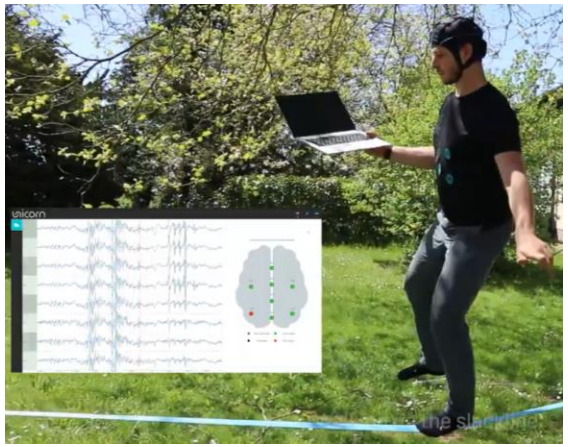
Casual Brain Activity Data Recording

→ Most recordings of brain activity data take place in the laboratory
- Highly controlled, but also unrealistic and not generalizable
- Small amounts of well-labeled data → does not scale

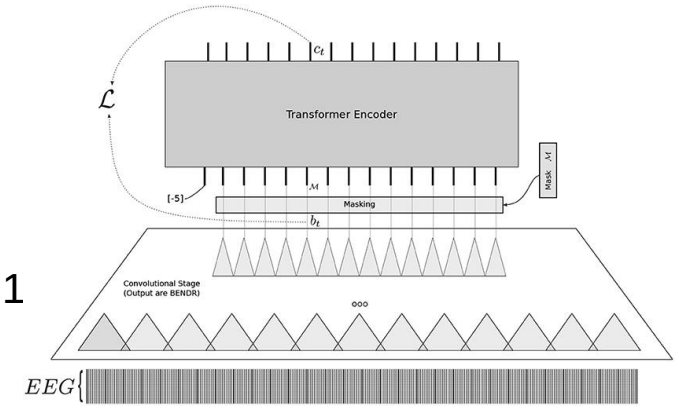
→ Alternative: Casual recording of brain activity data
- Completely uncontrolled, during everyday activities
- No or very weak labeling
- Leveraging self-supervised learning methods

→ **Your thesis**: Collect and evaluate such data
- Set up sensors, design data collection protocol „in the wild“
- Apply state-of-the-art machine learning for classification

→ Contact: Dr. Felix Putze, felix.putze@uni-bremen.de or Discord: felix.putze#871



https://www.youtube.com/watch?v=Tv5LyYuSDq4



https://www.frontiersin.org/articles/10.3389/fnhum.2021.653659/full

Representation Learning for Cognitive Modeling

→ Representation learning: Automatically finding a compact, latent representation of high-dimensional input
- Retaining important semantics and data characteristics
- Input: video, text, biosignals, graphs, …

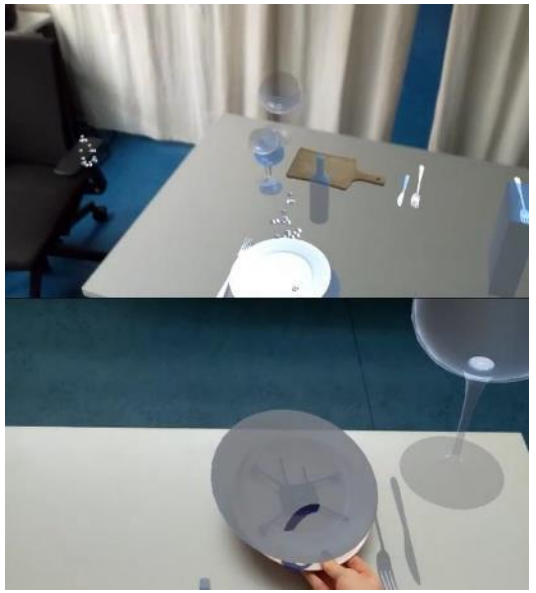
→ At the core of recent machine learning advances (Transformers, Large Language Models, DALL-E 2, …)

→ **Your thesis**: Apply representation learning for modeling context and cognition in Human-Computer Interaction:
- Representing context from video recordings of Augmented Reality device
- Representing neural data from EEG for Brain-Computer Interfaces
- Representing semantic concepts for creative idea generation

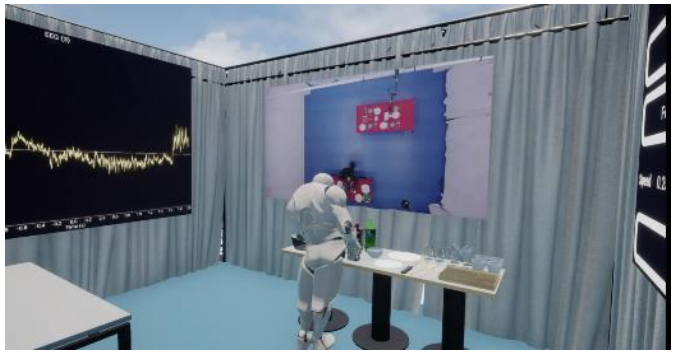
→ Contact: Dr. Felix Putze, felix.putze@uni-bremen.de or Discord: felix.putze#8717

Augmented/Virtual Reality for Cognitive Science

→ Motivation 1: In Augmented and Virtual Reality, we can create realistic, but controllable experiments to **study human behavior and cognition**!



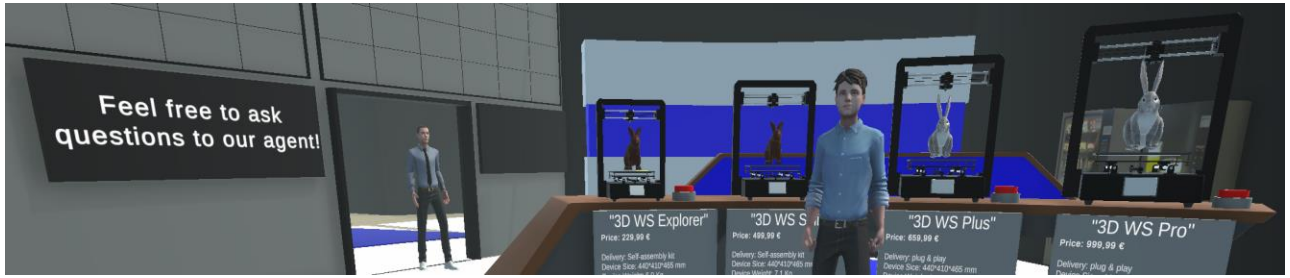
→ Motivation 2: AR/VR can benefit from cognitive modeling to **adapt its user interfaces**



→ **Your thesis**: Create AR/VR applications in the context of cognitive science, e.g.:
- Study in VR how different impairments influence how people solve a task
- Create an AR application that tracks the wearer‘s state of attention and adapts it

→ Contact: Dr. Felix Putze, felix.putze@uni-bremen.de or Discord: felix.putze#8717

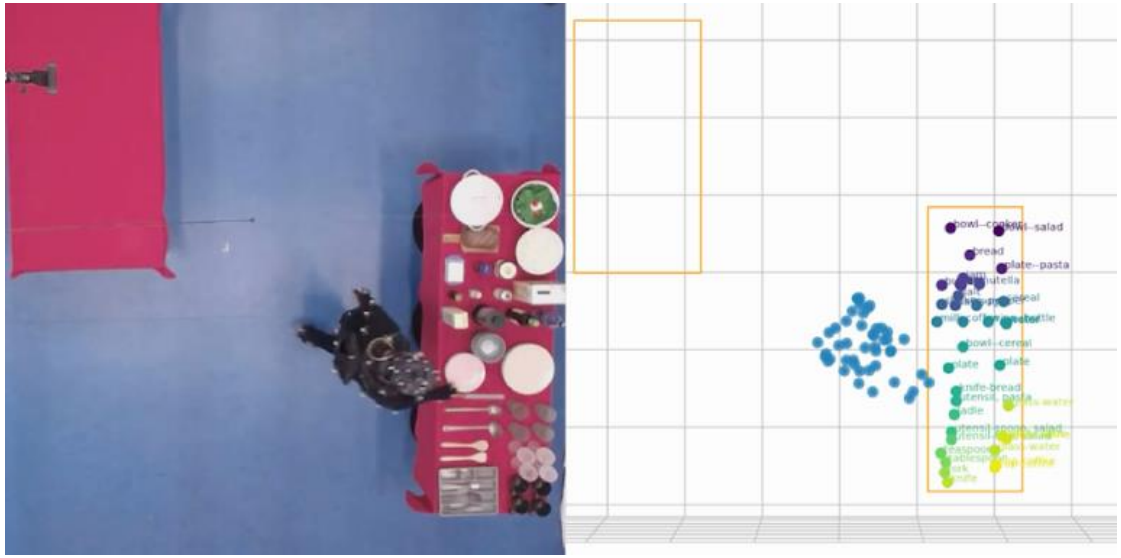
Procedural Animation for Avatars (MA)

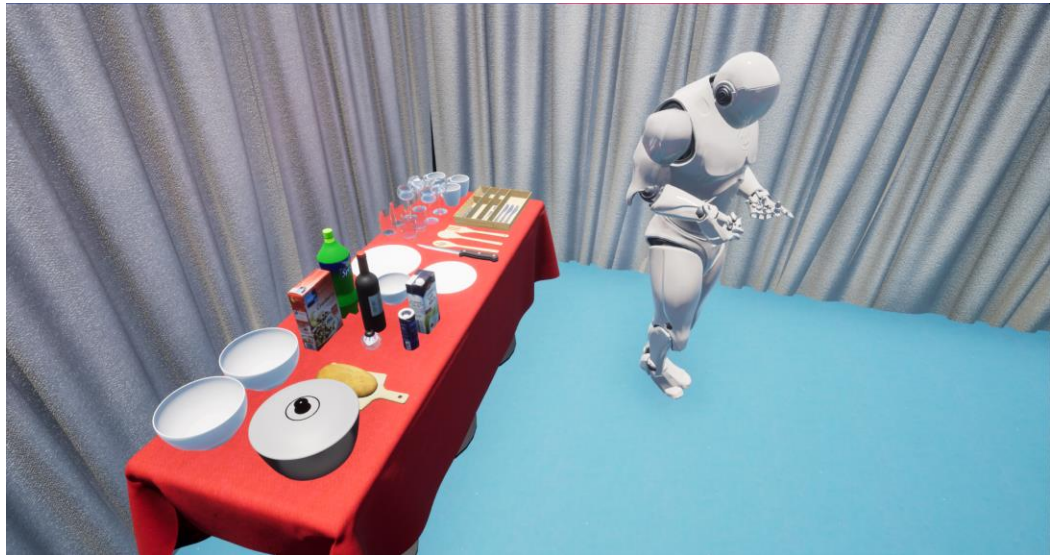


- The success of LLMs has opened a lot of new possibilities for Conversational Agents
- Embodied Conversational Agents (ECAs) require motion to not appear stiff and unrealistic
- Motion cant be easily pre-animated, as Text is based on LLM.
- Main Task: Implement plausible motions for an Avatar, either through Inverse Kinematics or Procedural Blending of Animation
- Technical Background: Games Engines (Unity or Unreal), VR
- If youre interested, contact Asmus Eilks (aeilks@uni-bremen.de)

Mapping Real-Life Recordings to VR

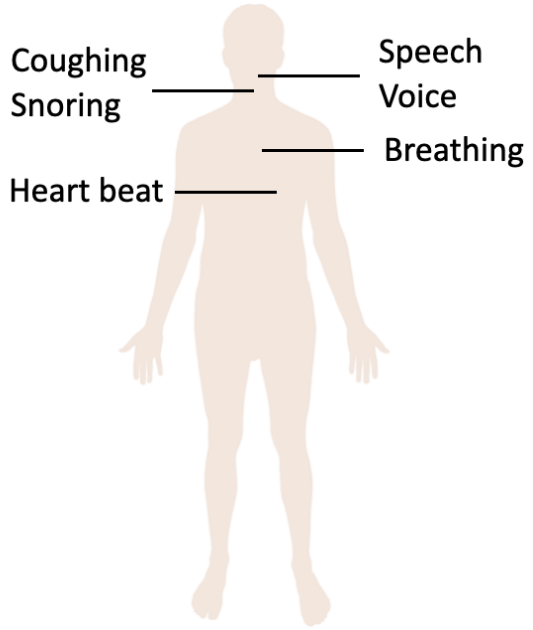
→ Create an automatic mapping from real-world sensor data (video, motion tracking, audio, …) to a Virtual Reality replay
→ Allows the rendering of new views of the same data (e.g., wie different lighting, from new camera angle, …)
→ Should include person performing everyday activity as well as manipulated objects





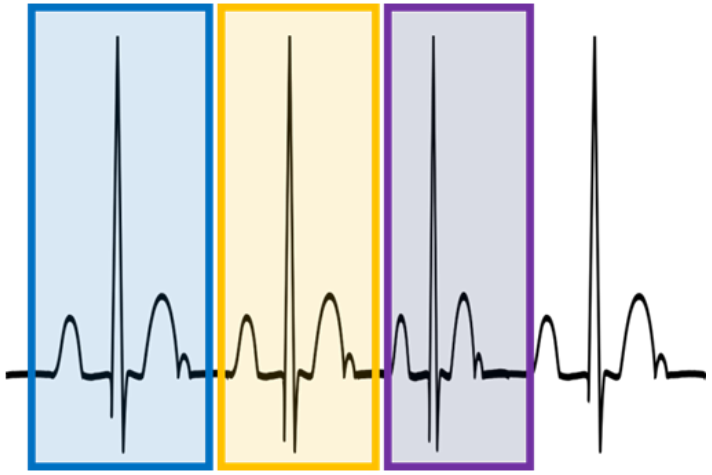
Computer Audition for Healthcare in Internet of Things (BA/MA theses)

- Computer Audition aims teach machines to listen to and understand the world. We mainly focuses on Computer Audition's applications in healthcare, like speech emotion recognition and bio-acoustic signal processing. Multiple research problems in the context of internet of things are waiting for us to address.
    - How can we build a trustworthy model for users?
    - How can we optimise models (such as reducing model size) to work on edge devices?
    - How can we protect users' privacy?
- Requirements: Strong motivation in this topic, interests in reading research papers, very good knowldge in deep learning, and skills of Python programming.
- If you are interested in working on any of the above research problems as your thesis topic, please send your **CV** and **transcripts** to Zhao Ren (zren@uni-bremen.de)



A Valid Data Augmentation Set for Biosignals (BA/MA)

Data augmentation is the base mechanism supporting image based unsupervised representation efforts with great success. While in image domain this is easy to implement, i.e. Proper transformations are naturaly available (cropping, flipping, scaling, bluring, etc.), the application of the same concepts to biosignals is more nuanced, as many easy to implement transformations can impact the information contained in them.



Your thesis:

→ Choose two biosignals (z.B. ECG and EEG) and understand how they are generated.
→ Search literature for already used data augmentation strategies for those signals.
→ Evaluate their validity based on your knowledge of the data generating process.
→ Find or propose an evaluation metric that measures the usefulness/validity of data augmentations.
→ Run experiments that showcase that evaluation metric.

Requirements: Good python skills, Signal processing knowledge, (Deep Learning knowledge is appreciated)

→ Contact: Lourenço Rodrigues (lourenco@uni-bremen.de)