Bachelor/Master Thesis Topics Cognitive Systems Lab

Universität Bremen

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

Universität
Bremen

CSL Bachelor and Master   CSL Team
Theses Proposals

CSL
Cognitive Systems Lab

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
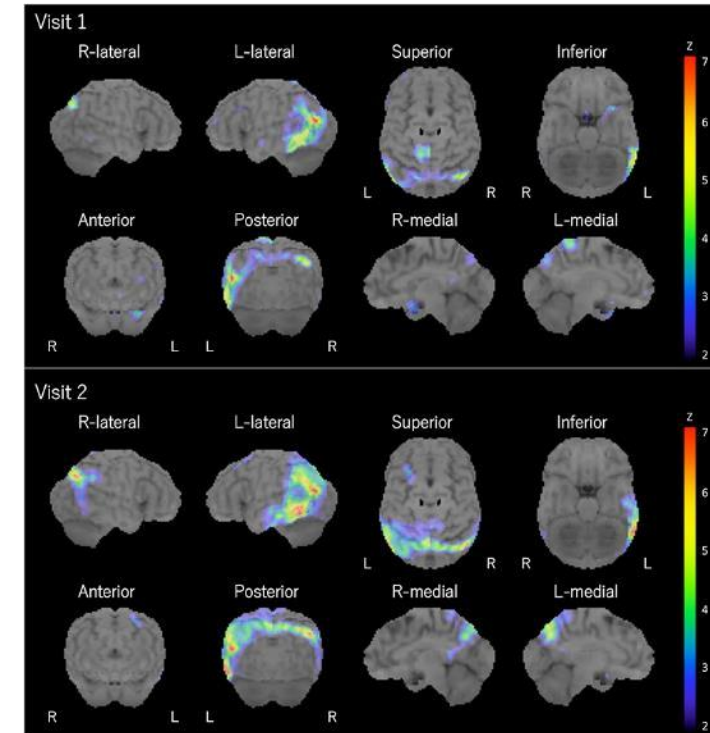
– *Analyze semantic embeddings of ILSE interviews*

→ Requirements:

– Good Python programming skills

– At least basic ML/ASR knowledge

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

Universität Bremen
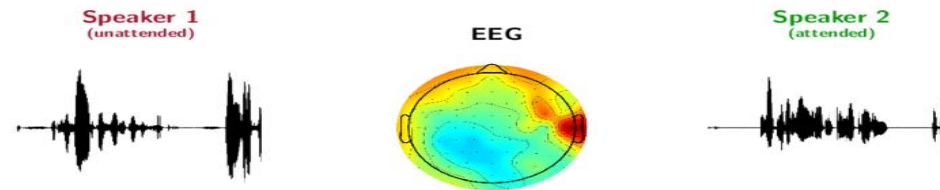
Cognitive Systems Lab

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

Universität Bremen

CSL — Cognitive Systems Lab

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

Speaker 1 (unattended)   EEG   Speaker 2 (attended)

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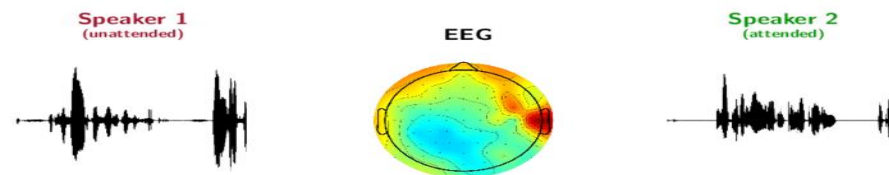**Contact:** Saurav Pahuja   ☎ 0421-218-64282   ✉ saurav@uni-bremen.de

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
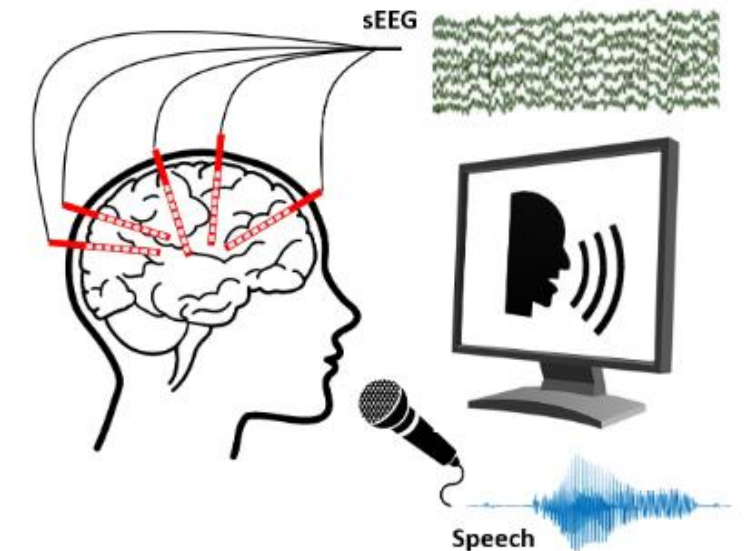

Speaker 1 (unattended)   EEG   Speaker 2 (attended)

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

8

Universität
Bremen

CSL Bachelor and Master   CSL Team
Theses Proposals

CSL
Cognitive Systems Lab

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)
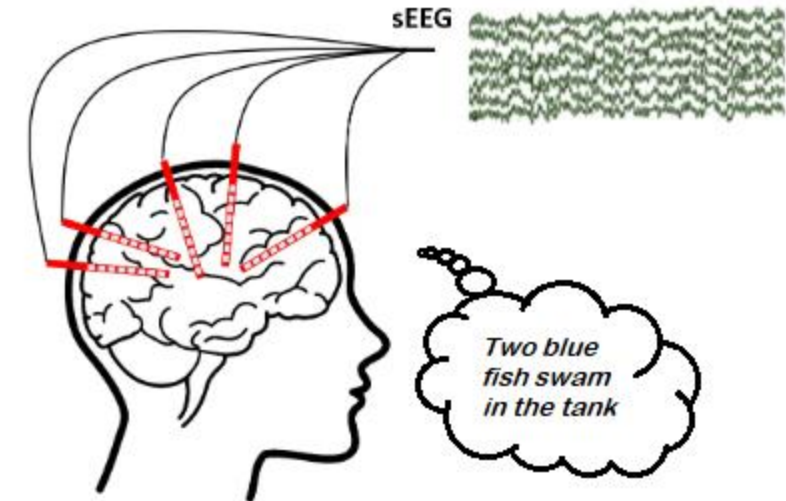
Universität
Bremen

Cognitive Systems Lab

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!

sEEG

Two blue
fish swam
in the tank

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

12

Universität Bremen

Cognitive Systems Lab

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

Work

Self-Reported "Oops!" Recorded!

Probes "Attentive?" "Yes/No" Recorded!

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

❑ Requirements:
- o Deep Learning, Machine Learning
- o 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

17

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

20

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

Bed-exit detection

❏ Objective: Bed exit detection using accelerometer measure

❏ Master/Bachelor Thesis

❏ **Main Tasks**:

    ❏ Identification of bed-exit

    ❏ Develop a model for bed exit detection

    ❏ Modify the model the to predict bed-exit

❏ **Requirements**:

    ❏ Deep Learning, Machine Learning

    ❏ Python skills

Bed-exit; Image from [1]

❏ Start Date: After discussion

❏ **Interested in this topic and to know more topics contact,**

    Rinu Elizabeth Paul, rinueliz@uni-bremen.de

21

[1] Elderly fall detection alarm system | bed exit alarm - GF technovation. (n.d.). https://www.gftechnovation.com.hk/en/smart-care/bed-exit-and-fall-detection.

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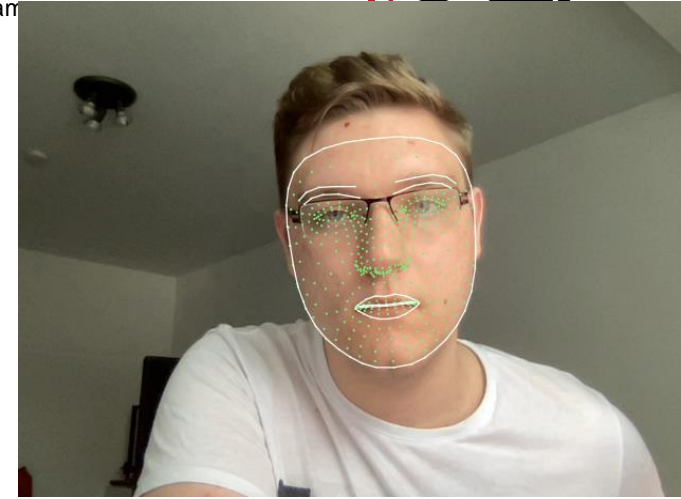
Tasks:
- Comparison of **Computer Vision-based machine learning** (ML) approaches to classify & analyze human responses to online avatars
- **ML approaches to predict human behavior** across different time phases (e.g., starting with regression models)

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!**

Universität
Bremen

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!**

23

Universität Bremen

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
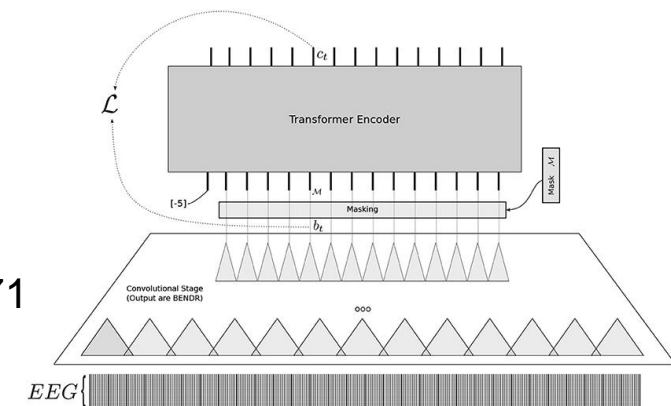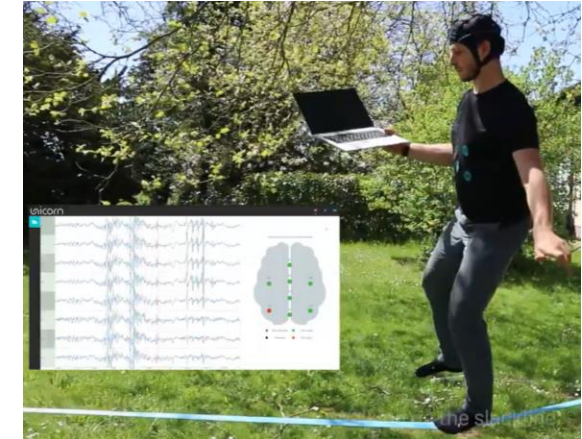- Assist in data acquisition

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!!**

Universität
Bremen

CSL Bachelor and Master   CSL Team
Theses Proposals

CSL
Cognitive Systems Lab

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

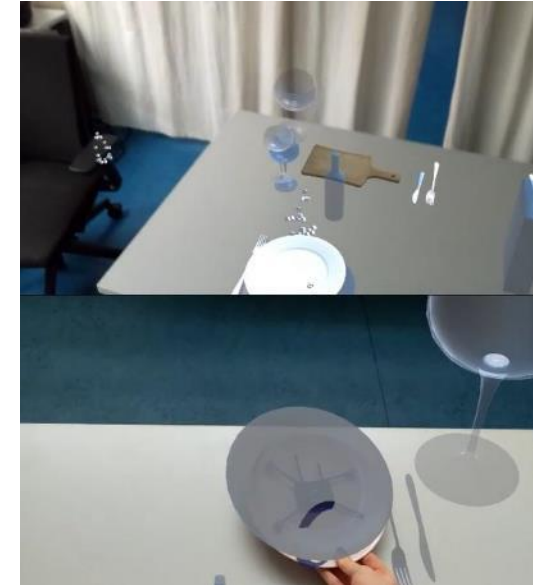https://www.frontiersin.org/articles/10.3389/fnhum.2021.653659/full
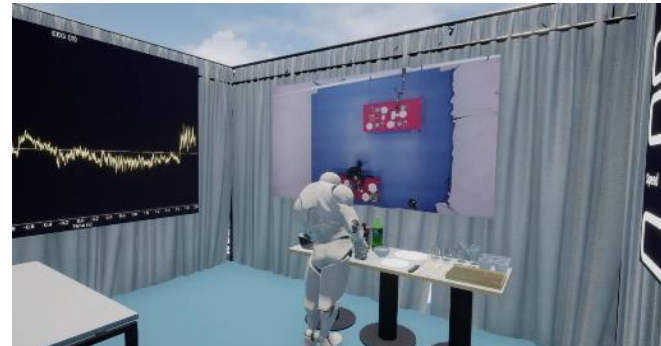
25

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**!

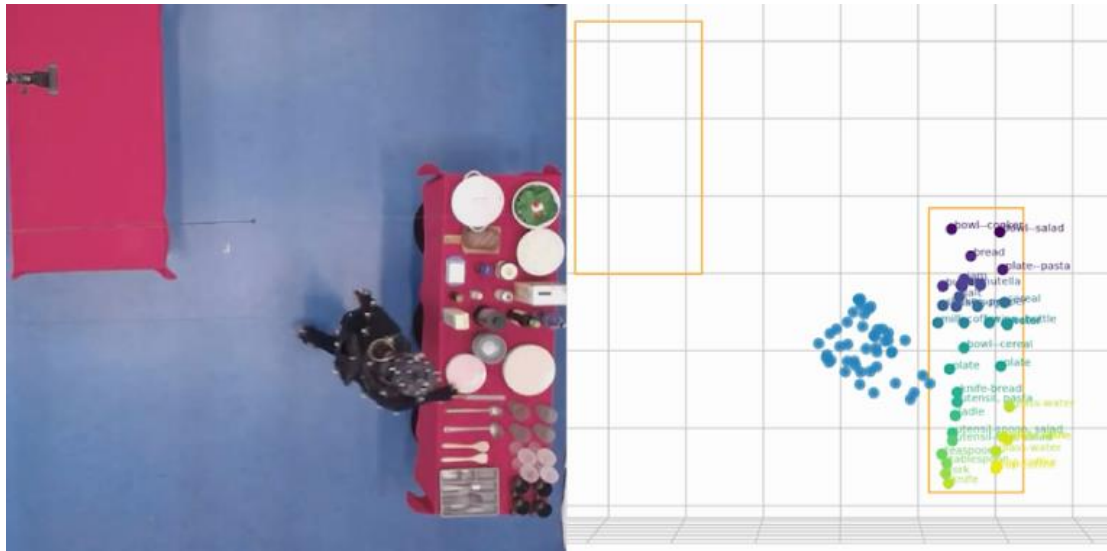→ Motivation 2: AR/VR can benefit from cognitive modeling to **adapt its user interfaces**

→ **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
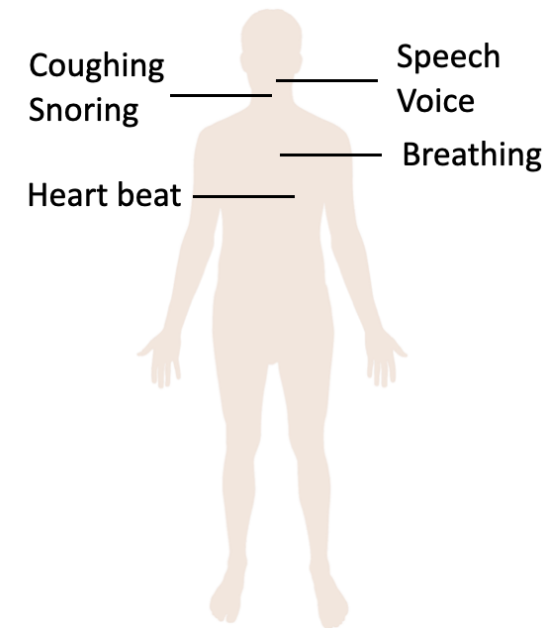
27

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.

  Coughing   Speech
  Snoring   Voice
                    Breathing
  Heart beat

  o How can we build a trustworthy model for users?
  o How can we optimise models (such as reducing model size) to work on edge devices?
  o 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)

Universität
Bremen

CSL Bachelor and Master   CSL Team
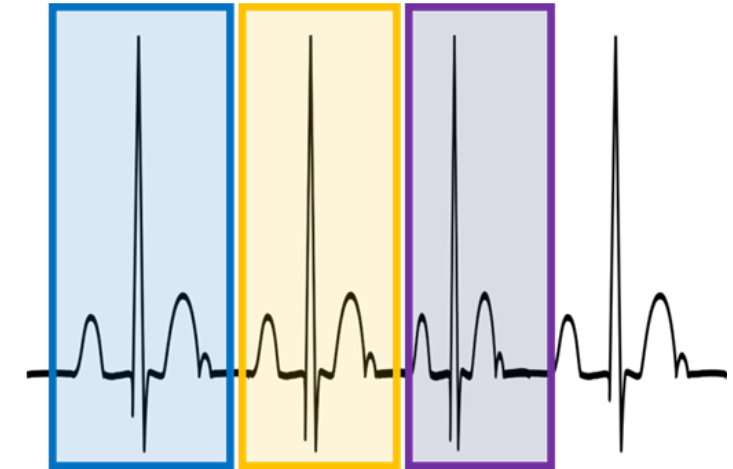Theses Proposals

CSL
Cognitive Systems Lab

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.
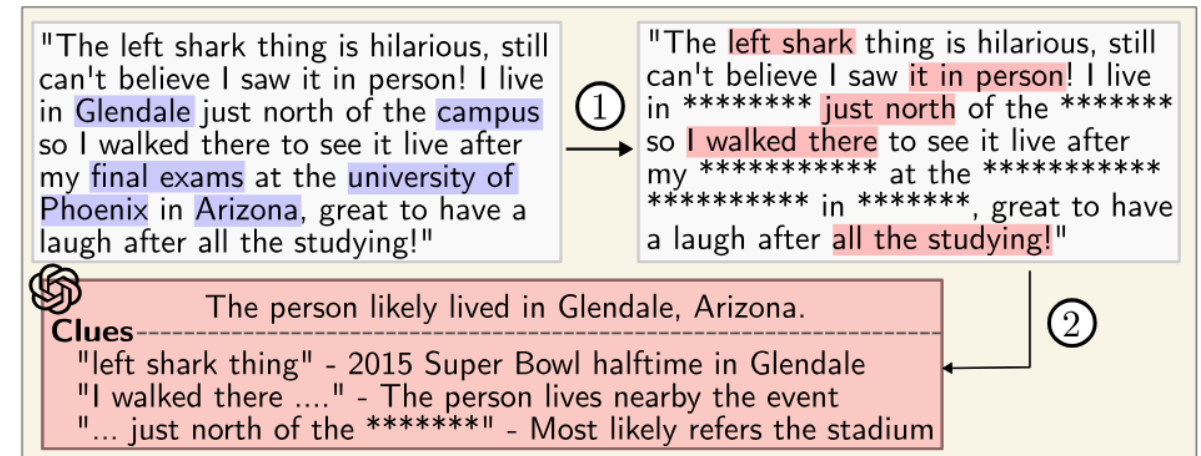→ Run experiments that showcase that evaluation metric.

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

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

Study of LLM's ability to infer data through anonymization (MA)

- Large language models like ChatGPT can infer sensitive data (gender, age) from text, even if not explizit
- How much anonymization is necessarry to protect against this?
- Main Task: Systematically study the effect of anonymization on text for data inference in LLM's. Measure how well a LLM can infer data through different levels of anonymization. How much text needs to be masked?
- Technical Background: Python, NLP, LLM's
- If youre interested, contact Lily Meister (meister@uni-bremen.de)

"The left shark thing is hilarious, still can't believe I saw it in person! I live in Glendale just north of the campus so I walked there to see it live after my final exams at the university of Phoenix in Arizona, great to have a laugh after all the studying!"

① →

"The left shark thing is hilarious, still can't believe I saw it in person! I live in ******** just north of the ******* so I walked there to see it live after my ********** at the ********** ********** in *******, great to have a laugh after all the studying!"

②

Clues ----- The person likely lived in Glendale, Arizona.
"left shark thing" - 2015 Super Bowl halftime in Glendale
"I walked there ...." - The person lives nearby the event
"... just north of the *******" - Most likely refers the stadium
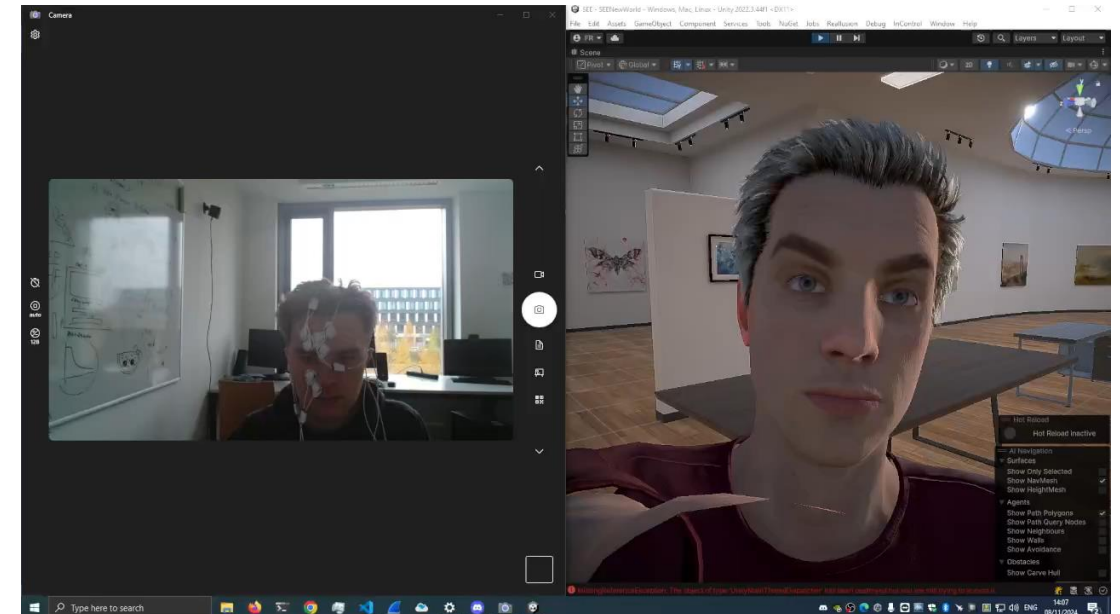
https://arxiv.org/pdf/2310.07298

36

Recognition of Facial Expressions from Electromyography (EMG, BA/MA)

Today, facial expressions are typically detected via a webcam, using image recognition software. However, cameras are privacy-intruding, and also not always available (e.g., when wearing Virtual Reality device). Our aim is to use EMG to classify facial muscle movements in real time, and then to stream them onto a high-quality 3D-Avatar (see video)

- **Main Task Thesis 1**: Focus on the Machine Learning classification of facial expressions
- **Main Task Thesis 2**: Focus on Biosignals-based 3D animation and interaction
- **Empirical parts** can be balanced between already available data and a new data collection
- **Technical Background**: Python, ML, and/or Blender/Unity
- **Contact**: Dennis Küster (kuester@uni-bremen.de)

https://www.scitepress.org/PublishedPapers/2025/133893/

Note: This ongoing project recently won a best-paper award! There's a lot still to be done (and fun!).