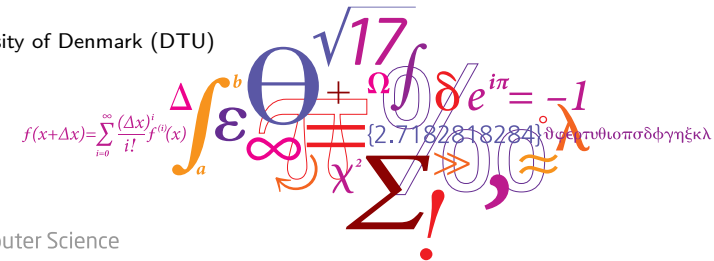
Multimodal Data Analysis for OCD Treatment and Management

Sneha Das

sned@dtu.dk | @dassneh

Statistics and Data Analysis, Technical University of Denmark (DTU)



DTU Compute
Department of Applied Mathematics and Computer Science

About me...

1. Bachelor of Engineering (India)
2. Masters of Science: Communications and Multimedia Engineering (FAU, Erlangen, Germany)
3. Doctor of Science (Tech) (Aalto University, Finland)
   Robust and Efficient Methods for Distributed Speech Processing–Perspectives on Coding, Enhancement and Privacy
4. Postdoctoral researcher (Technical University of Denmark)

WristAngel: Research for Intervention and Management of OCD

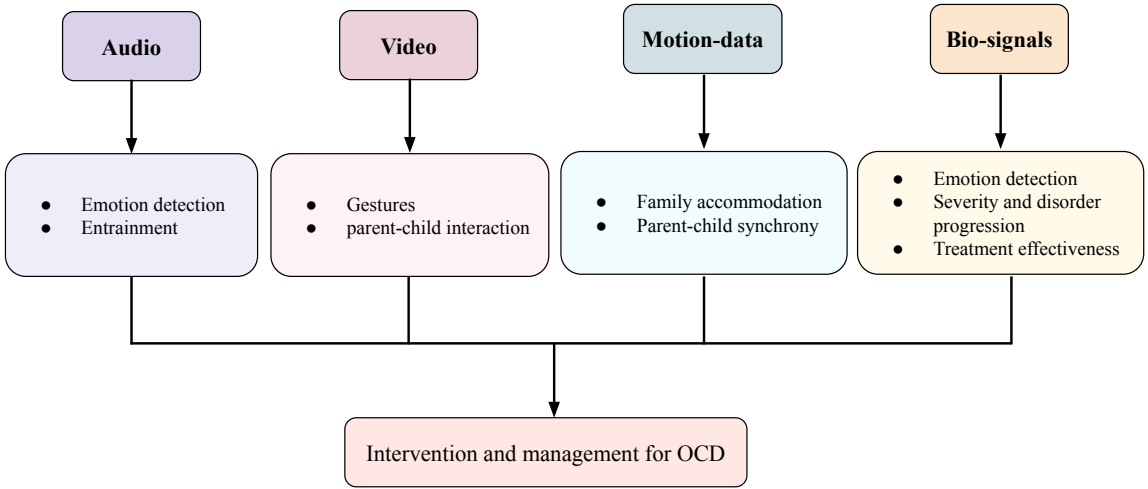
* Progression and severity of disorder.
* Improve efficiency in CIB (Coding Interactive Behavior)

¤ Identify and predict impending OCD events.

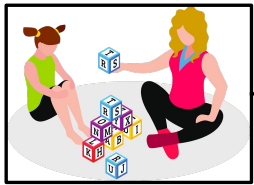
¤ Aid in delivering cognitive behavioral therapy to patients.

¤ Provide useful interventions for management.

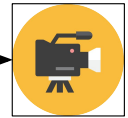
Mental Health meets Technology



Data and Signals

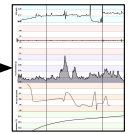


Therapy sessions

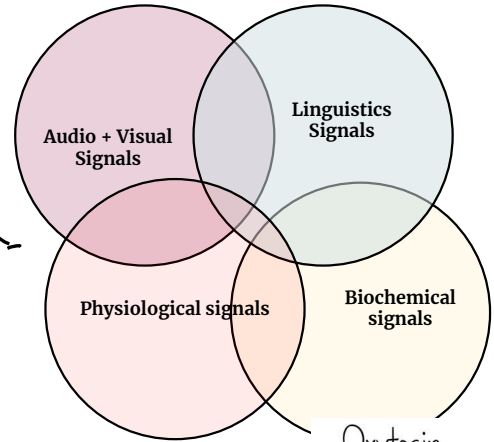


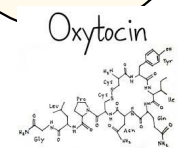
Videos

Wearable device
(Empatica E4)



Heart-rate, EDA,
Accelerometer



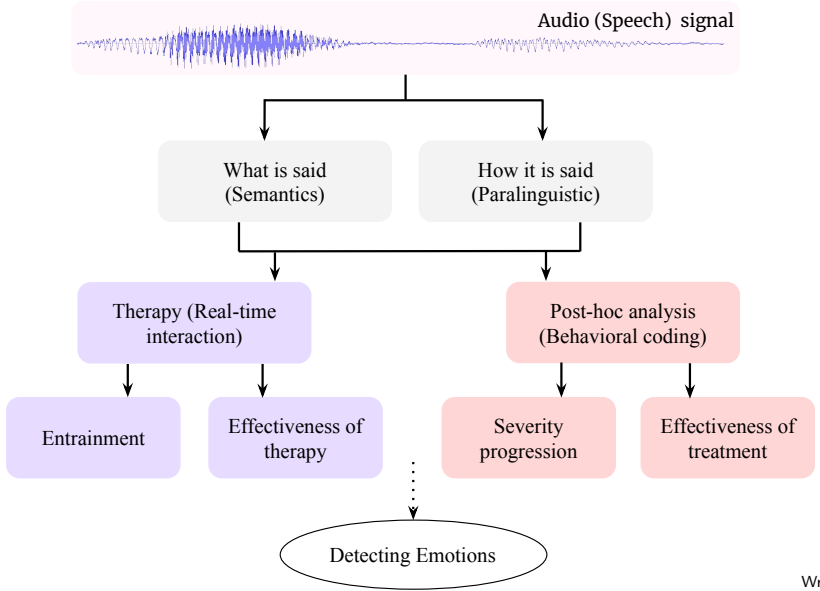


Some (Data) Constraints

1. Changes in data source (Two trials) $\rightarrow$ Generalization
2. Models built on baseline data only (Trials ongoing $\rightarrow$ Blinded)$\rightarrow$ Low-data resources.

AUDIO DATA

Role of Audio (Speech) in OCD Treatment



Speech Emotion Detection I

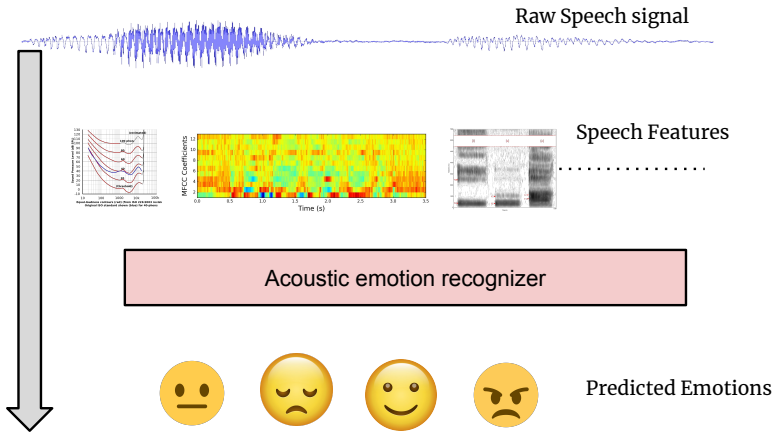


Figure: Image sources https://medium.com/prathena/the-dummys-guide-to-mfcc-aceab2450fd; https://commons.wikimedia.org/wiki/File:Lindos1.svg; https://commons.wikimedia.org/wiki/File:Spectrogram_-iua-.png

Speech Emotion Detection II

Conventional approaches

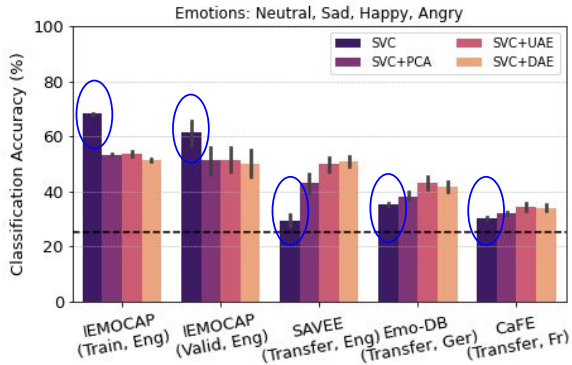
• Statistical ML and signal processing	HMM, GMM, SVM
• Deep learning (DL)	RNN, CNN, LSTM with deep architectures
• Hybrid	Eg., DL +SVM

Speech Emotion Detection III

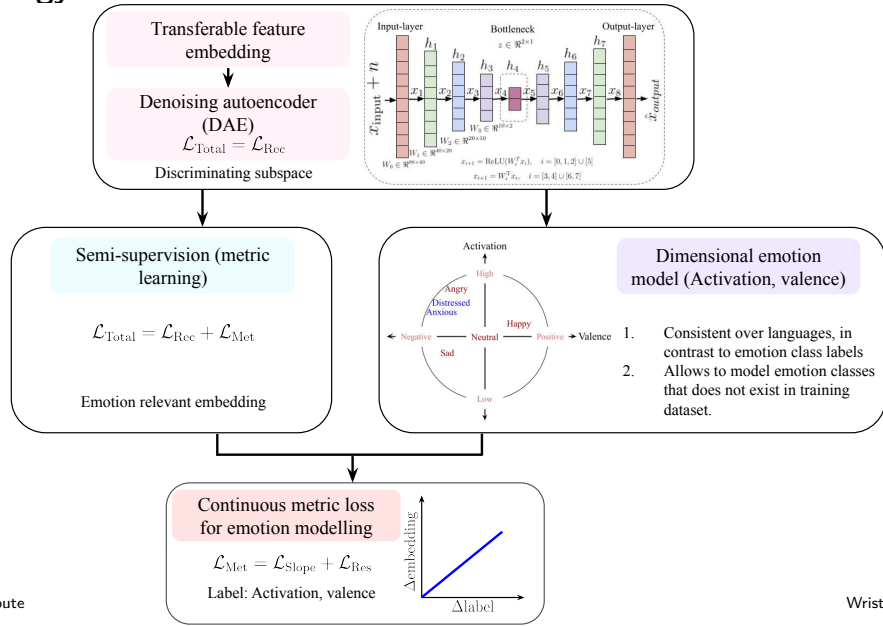
Persistent challenges

- Generalization
- Low-resource corpora
- Black-boxes

corpora, languages → cultural, phonetic differences (Danish, kids, clinical)
Small data set and lack of labels

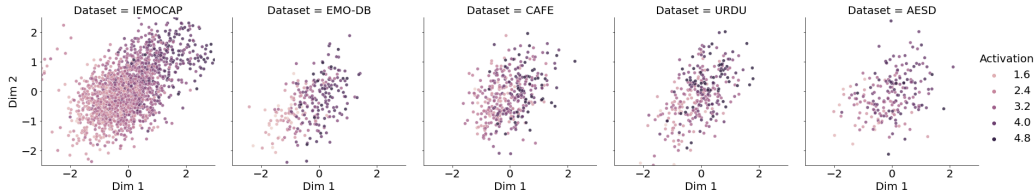


Methodology

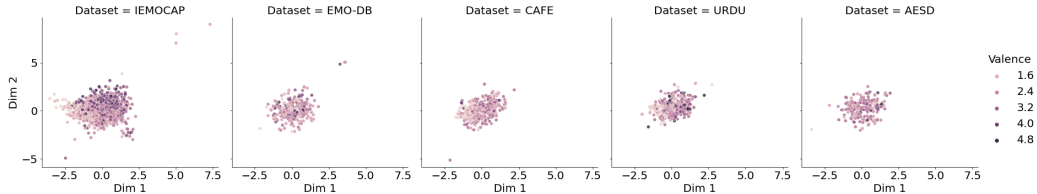


Generalization versus personalization

1. Activation:



2. Valence:



Currently submitted to: JMIR Research Protocols
Date Submitted: May 30, 2022
Open Peer Review Period: May 30, 2022 - Jul 25, 2022
(currently open for review)

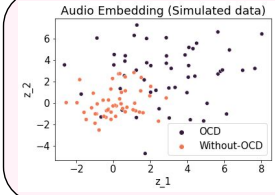
Peer-Review Me Tweet

Submitted Manuscript

Associations between OCD severity and vocal features in children and adolescents: A statistical and machine learning analysis plan

Line K. H. Clemmensen; Nicole Lønfeldt; Sneha Das; Nicklas Leander Lund; Valdemar Uhre; A.R. Cecilie Mora-Jensen; Linea Pretzman; Camilla Funch Uhre; Melanie Ritter; Nicoline Løcke Jepsen Korsbjerg; Julie Hagstrøm; Christine Lykke Thoustrup; Iben Clemmensen; Kersten Jessica Plessen; Anne Pagsberg

Step 1. Segmented speech signals



Step 2. Get audio embedding

H-1.a) Effect of diagnosis (OCD vs. no psychiatric diagnosis) on the vocal feature embedding.

OCD/no-OCD <-> vocal embedding

H-1.b) Effect of OCD severity on the vocal feature embedding.

OCD severity <-> vocal embedding

Model adaptation

Labels (Activation, valence)

Training

Labels (OCD-severity)

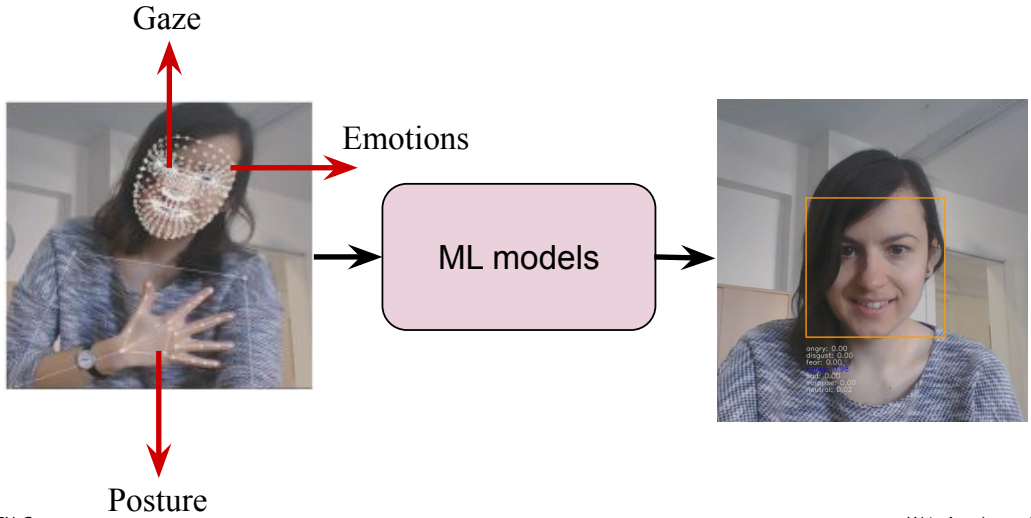
Improved prediction model

H-2) Improved models for OCD severity

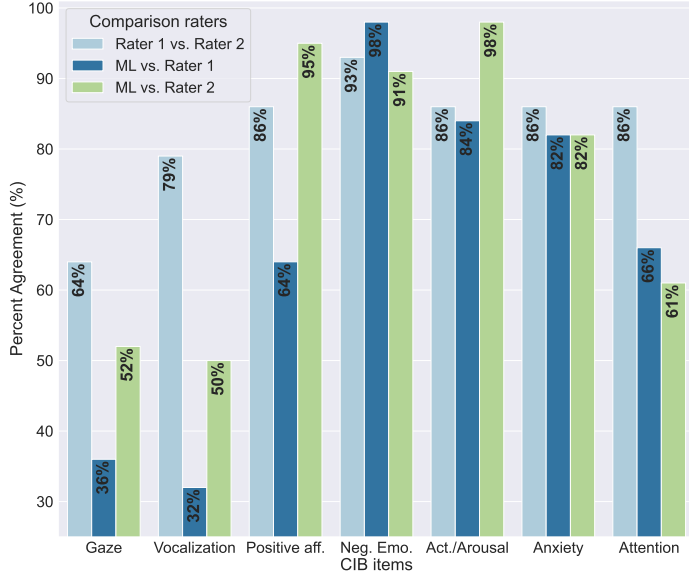
VIDEO DATA

From Subjective To Objective Units of Distress

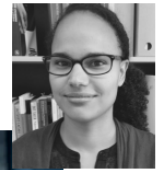
Gaze tracking + Facial expression recognition + Posture detection

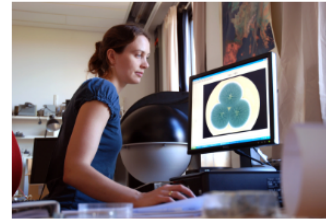


Experts (vs) Algorithms

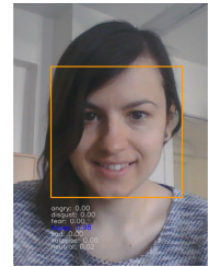


Thank you






TECTO

References I

Clemmensen, L. K. H., Lønfeldt, N., Das, S., Lund, N. L., Uhre, V., Mora-Jensen, A. C., Pretzman, L., Uhre, C. F., Ritter, M., Korsbjerg, N. L. J., et al. (2022).
Associations between OCD severity and vocal features in children and adolescents: A statistical and machine learning analysis plan.
In review at JMIR Research Protocols.

Das, S., Lønfeldt, N. N., Pagsberg, A. K., Clemmensen, L., et al. (2022a).
Speech detection for child-clinician conversations in danish for low-resource in-the-wild conditions: A case study.
arXiv preprint arXiv:2204.11550.

Das, S., Lønfeldt, N. N., Pagsberg, A. K., and Clemmensen, L. H. (2022b).
Towards transferable speech emotion representation: On loss functions for cross-lingual latent representations.
In *ICASSP International Conference on Acoustics, Speech, and Signal Processing*, volume 47.

References II

Das, S., Lund, N. L., Lønfeldt, N. N., Pagsberg, A. K., and Clemmensen, L. H. (2022c). Continuous metric learning for transferable speech emotion recognition and embedding across low-resource languages.
In *Proceedings of the Northern Lights Deep Learning Workshop*, volume 3.

Frumosu, F., Lønfeldt, N., Mora-Jensen, A., Das, S., Lund, N., Pagsberg, A., and Clemmensen, L. (2022).
Interpretability by design using computer vision for behavioral sensing in child and adolescent psychiatry.
2nd ICML Workshop on Interpretable Machine Learning in Healthcare.