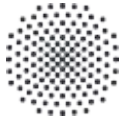
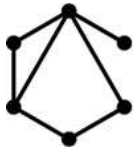
STUTTGART



University of Stuttgart
Germany



ANALYTIC
COMPUTING

Attention based Deep Learning and Affective Computing

Applied to Human Computer Interaction

Decky Aspandi

July 21, 2022

ellis-stuttgart.eu

e l l i s
European Laboratory for Learning and Intelligent Systems

Who Am I

Decky Aspandi

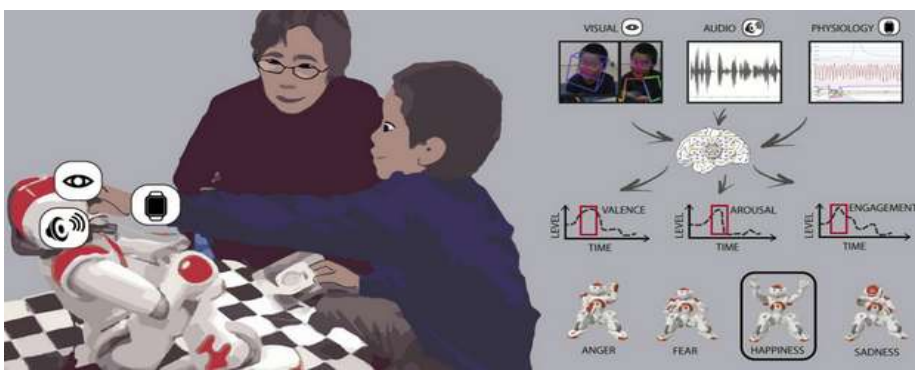
Department of Analytic Computing

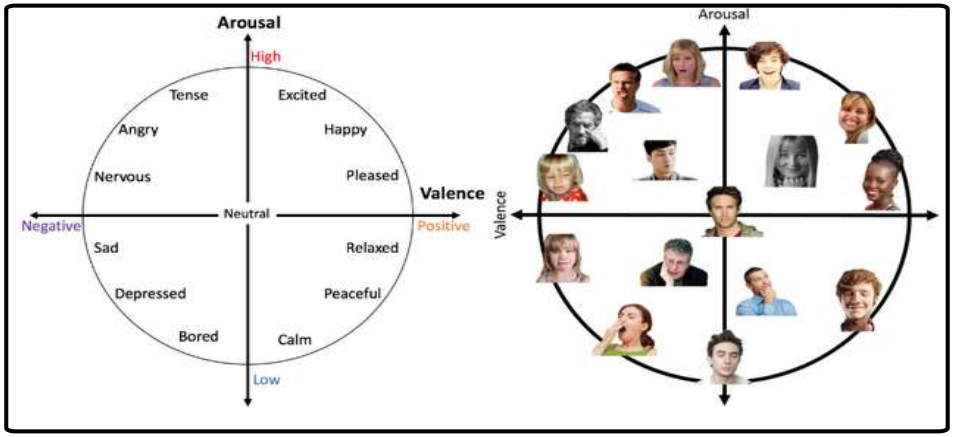
Deep Learning, Affective Computing

decky.aspandi-latif@ipvs.uni-stuttgart.de

Affective Computing

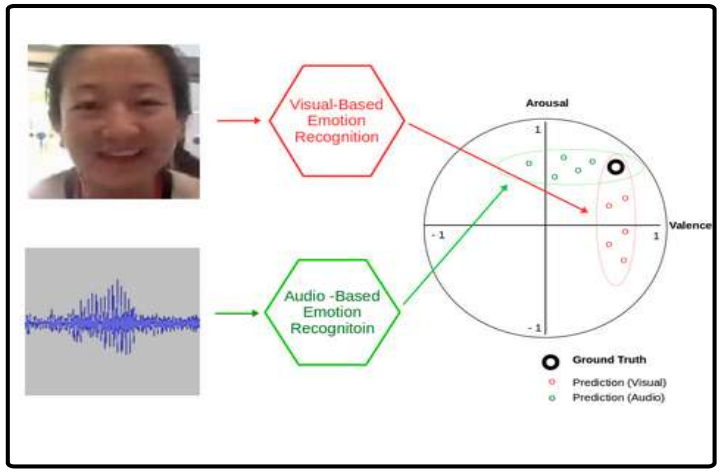
- There is a substantial demand to equip the computer with ability to understand human emotion states → useful for range of tasks such as health, education etc.
- In the current states, the valence arousal dimension (right figure) is commonly used to model wide array of emotion instances.

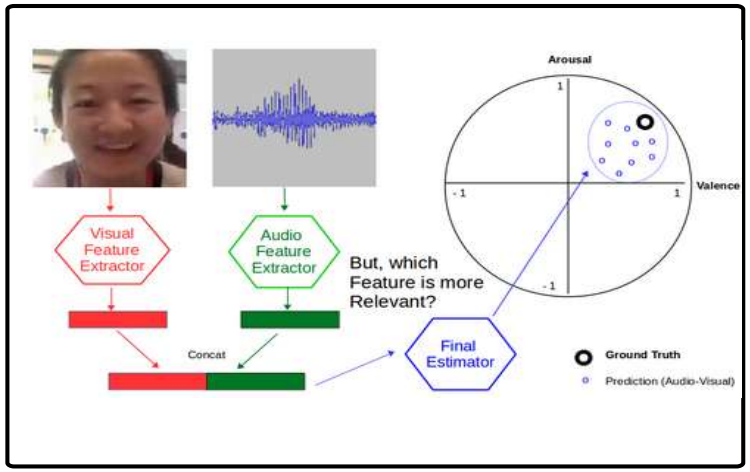




Visual vs Audio based affect estimation

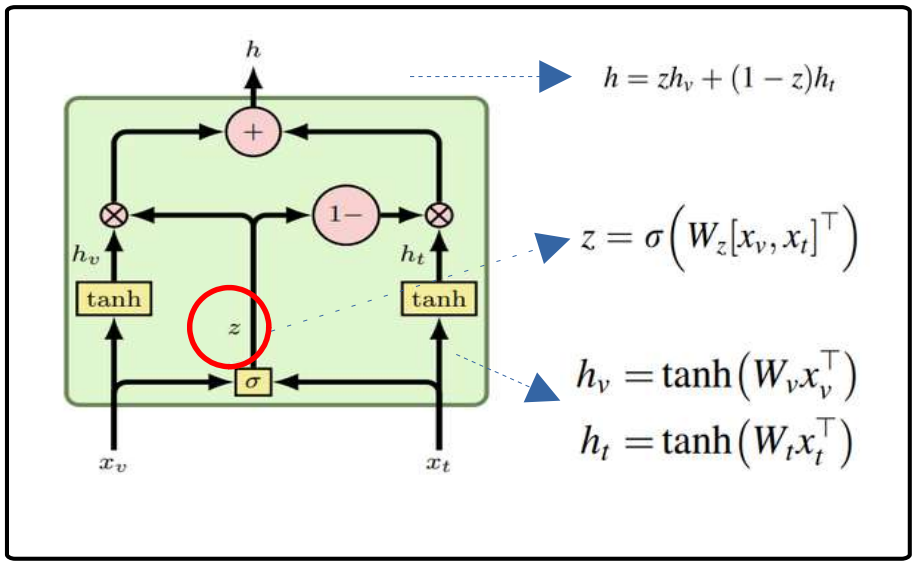
- Visual based (through facial input) affect estimation is prominent due to its intuitive nature.
- However, the estimation quality of facial based approach estimation usually suffers on arousal domain, as opposed to other signals (such as Audio).
- Current compromise is to use basic modality fusion mechanism (concatenation). The drawback is direct equalization of each modality.

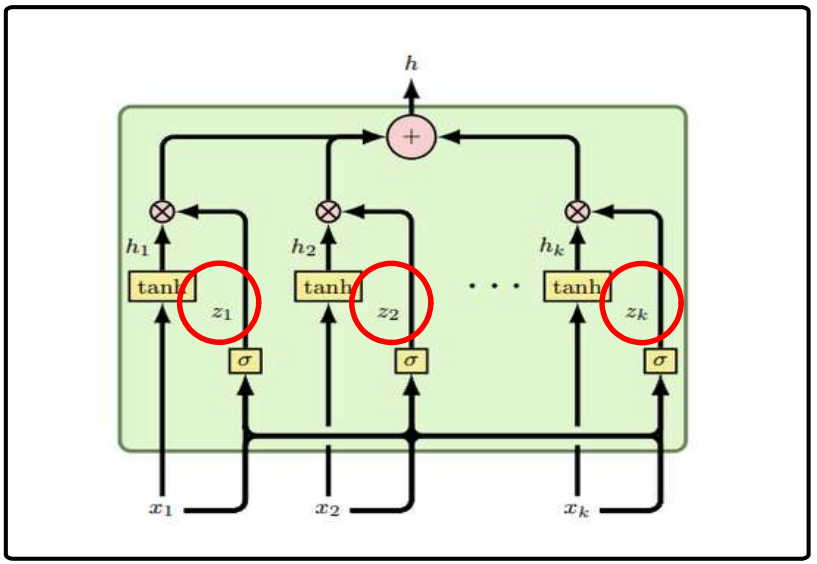




Gated Multimodal Units

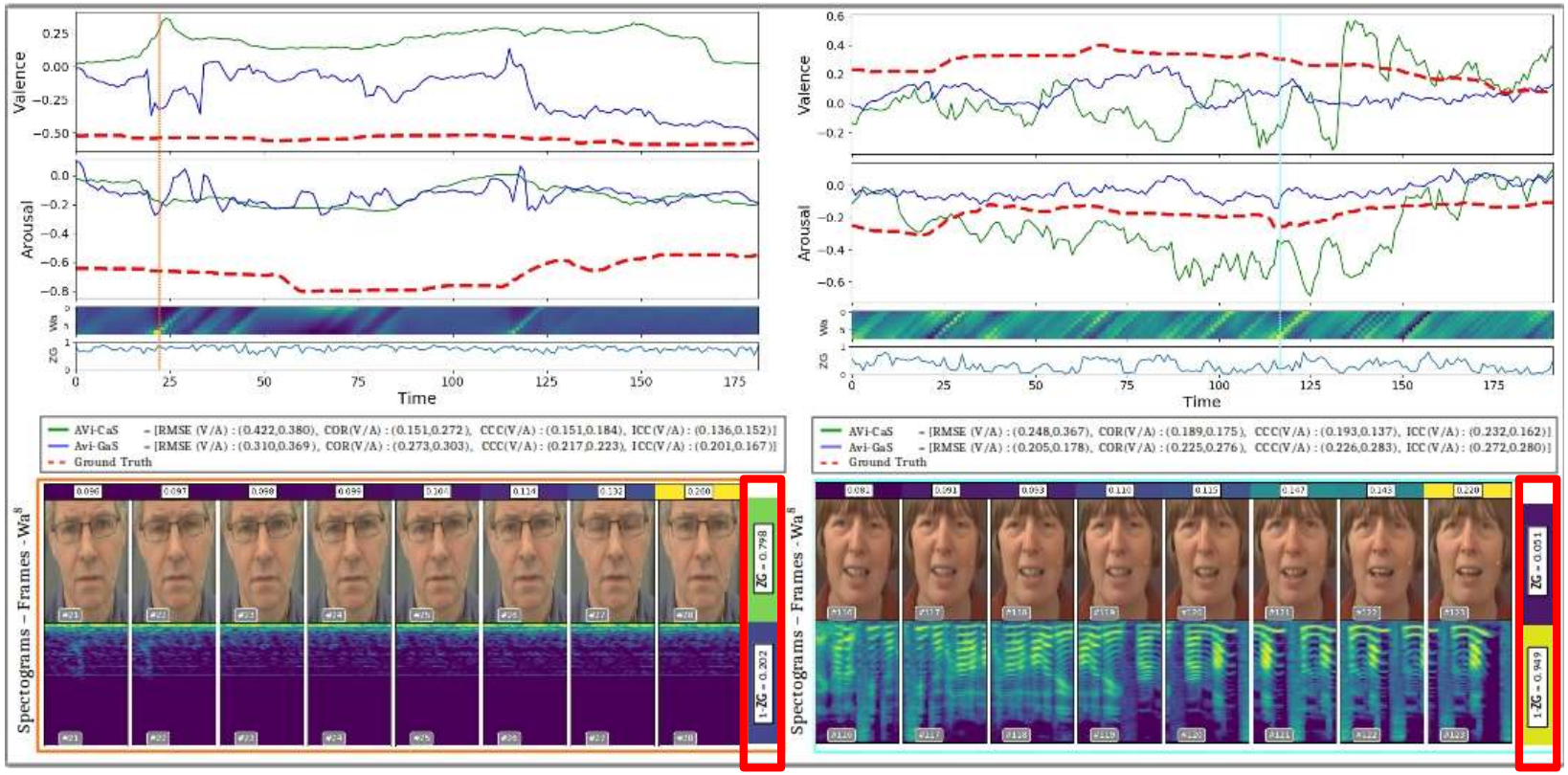
- The idea (John et al, 2020, Neurocomputing) is to use a gating coefficient (z) which evaluates the importance of modalities input during optimizations (loss evaluations).





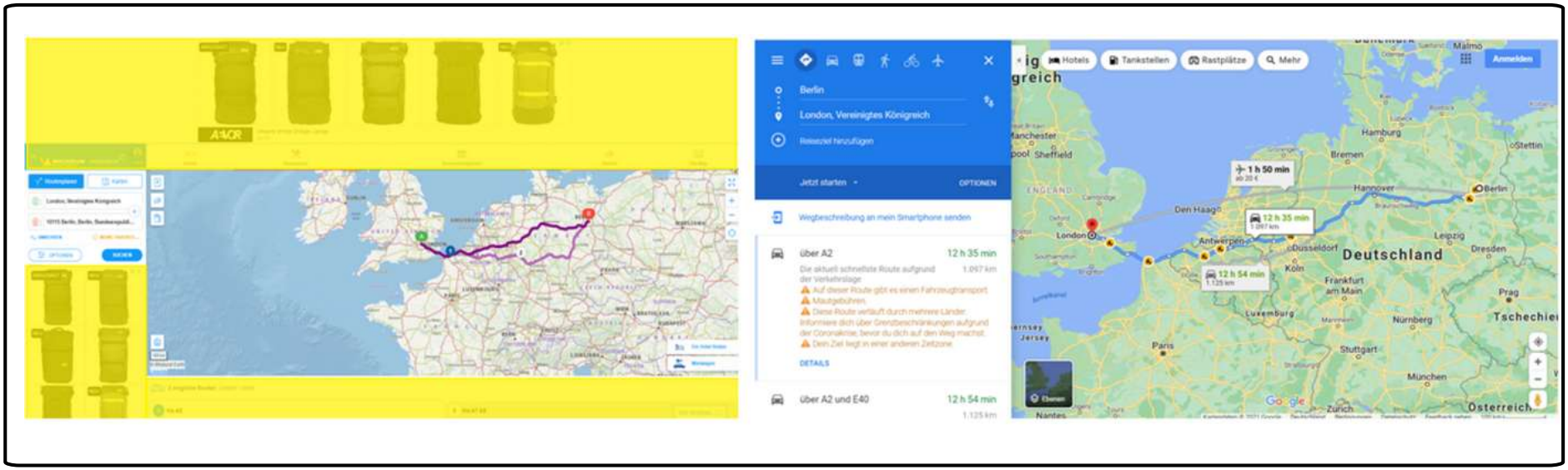
Results: Gated-Sequenced based Affect Recognition

- The gating coefficients (ZG) appears to regulate the importance of each modality inputs, where changes occurs on respective sequences (left changes perceived on visual inputs, right is on the audio) modality (Aspandi et al, 2022, IEEE Transaction of affective Computing).



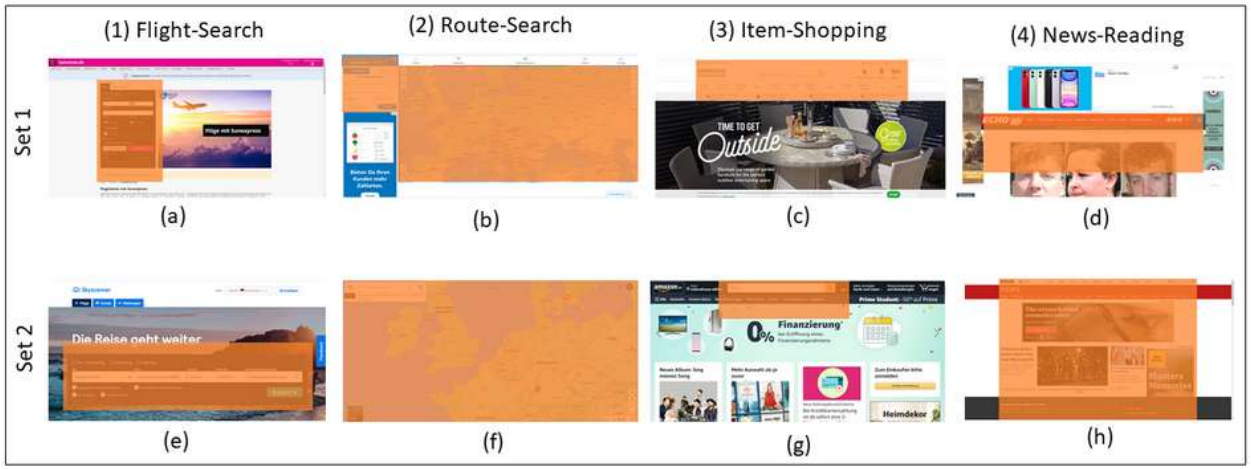
1st application: Affective Computing for Website Usability Analysis

- Challenge: would it be possible to quantify the different ‘usability level’ of each websites based on user interaction?
- Approach: to use facial based affect analysis to substantiate the usability analysis, based on the perceptive user affect during interactions.



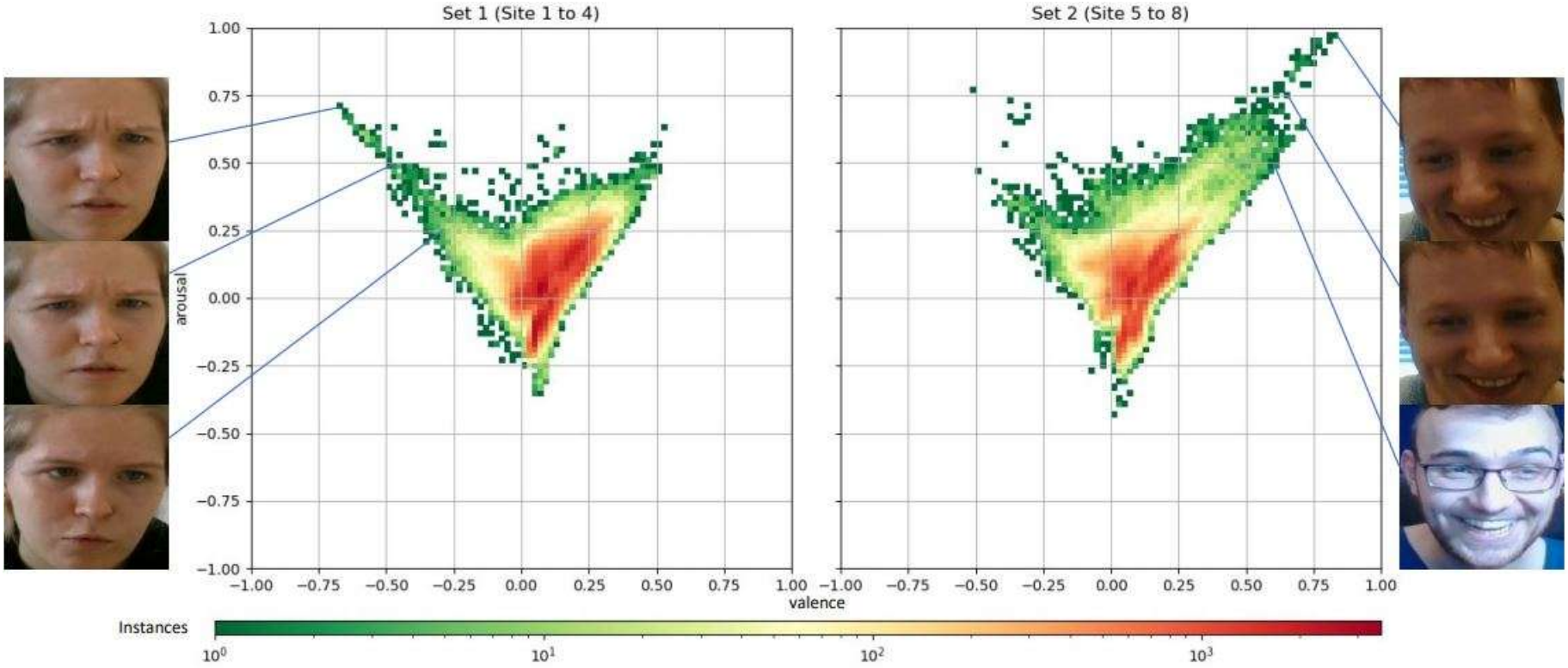
Contrastive Website Sets and Data Collections

- We chose two set of websites with different level of interactions:
  - Set 1 to contain websites with difficult to navigate functionallity and appearances.
  - Set 2 to exhibit ‘ease’ of use interaction.
- We record the facial area and user interactions (mouse) for analysis.



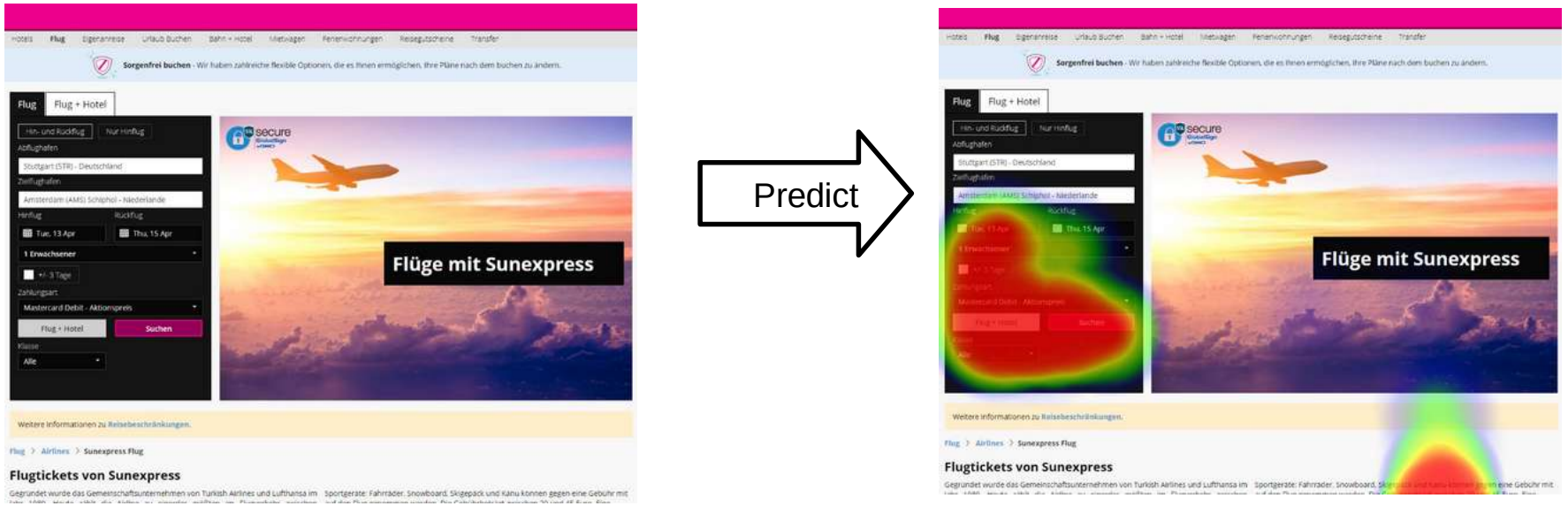
Results: difference of perceived user affect states between websites

- The use of affect recognition reveals that the users perceptions are more positive when dealing with ‘Good’ website (Set 2) compared to ‘Bad’ website (Set 1). (Aspandi et al, under submission).



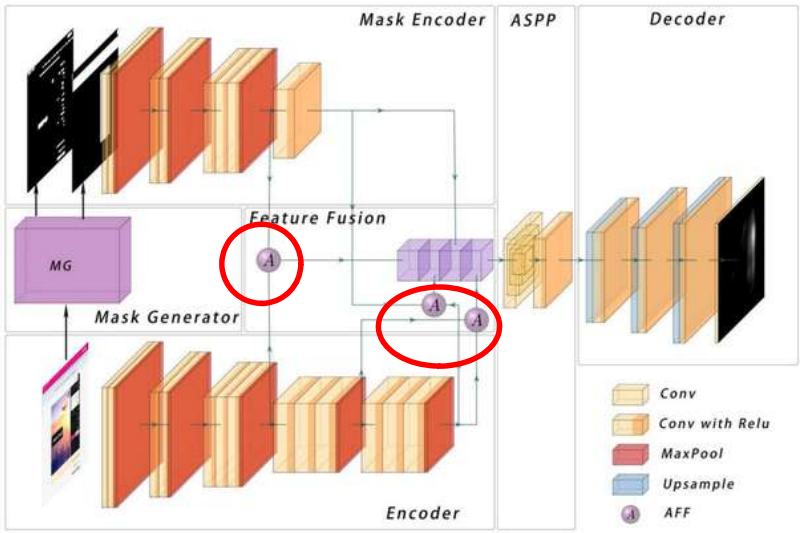
2nd application: Saliency prediction from Website Screenshot

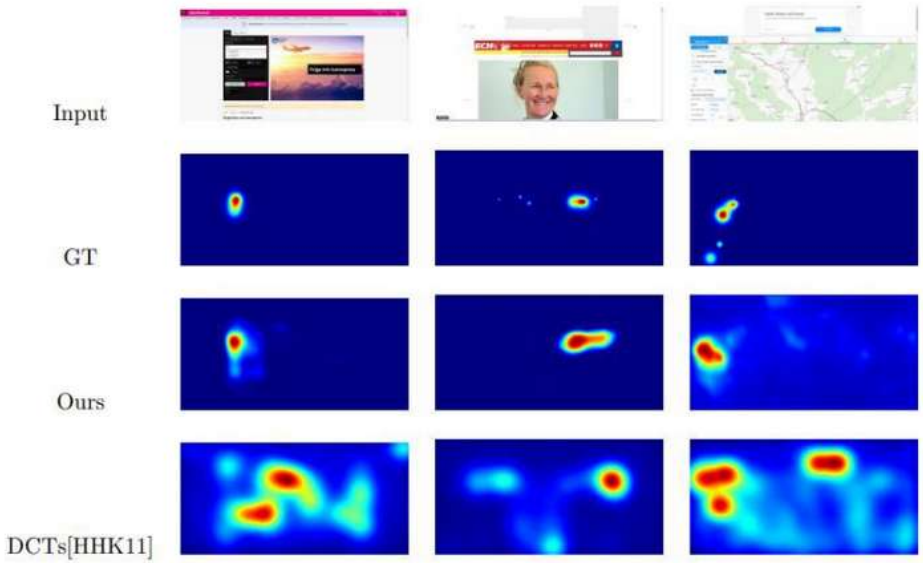
- Challenge: How we accurately detect ‘potential’ user gaze distribution only from the website screenshot input?
- Approach: To enhance existing gaze detections using attentional approach applied to visual and context modality.



Attending Visual and Mask for Improved Gaze Estimations

- The attentional mechanism is embedded to intermediate encoder of both raw visual inputs and mask inputs (text and images spatial locations).
- We found that our approaches produce more precise gaze predictions compared to existing approaches. (Zhang et al, under submission).





Conclusions & QA

- The use of multi-modal attentions improves affect recognitions tasks.
- The expansion of both multi-modal attentions application has been shown to benefit both visual and human computer interaction analysis.

- Thank you, and any questions?

Enjoy the kick-off event!