Latent Trees for Estimating Intensity of Facial Action Units
Sebastian Kaltwang, Sinisa Todorovic and Maja Pantic

**Problem**
Input Video Frame → Which FAUs are active and their intensity

**Approach**
- Latent Variables
- Landmarks
- FAU Intensities

We model landmarks and FAU intensities jointly as leave nodes of a Latent Tree (LT)

**Contributions**
- Joint estimation of multiple FAUs using LT
- Efficient model structure learning
- Evaluation under noisy conditions

**Model Structure Learning**
Repeat until $\Delta \mathcal{L} < \epsilon$
Try:
1. Add new edge:
   (a) select edge with max $\Delta \mathcal{L}$
   (b) require $\Delta \mathcal{L}_s \geq \epsilon$ for all siblings $s$

Otherwise:
2. Add new parent with the max $\Delta \mathcal{L}$

**Evaluation under noisy conditions**
Joint estimation of multiple FAUs using LT

**Qualitative Results**
Evaluation on the DISFA [1] and ShoulderPain [2] datasets under the varying amount of noise in features

**Quantitative Results**
Sampling of LT to generate landmark locations

**Acknowledgments**
H2020/2014-2020 645094 (SEWA)