MAKE PAIN ESTIMATION TRANSPARENT: A ROADMAP TO FUSE BAYESIAN DEEP LEARNING AND INDUCTIVE LOGIC PROGRAMMING

Ines Rieger[1], Bettina Finzel[2], Dominik Seuß[1], Thomas Wittenberg[1], Ute Schmid[2]
[1] Fraunhofer Institute for Integrated Circuits IIS, [2] Cognitive Systems Group of University of Bamberg

MOTIVATION

Reliable pain analysis of patients with limited communication skills – e.g. sedated patients in intensive care units, patients with neurocognitive disorders such as dementia, or patients suffering from the after-effects of general anesthesia.

GOAL

- Develop an image-based transparent medical expert companion for facial pain estimation.
- Combine human understandable logic rules and visualization to explain the results.

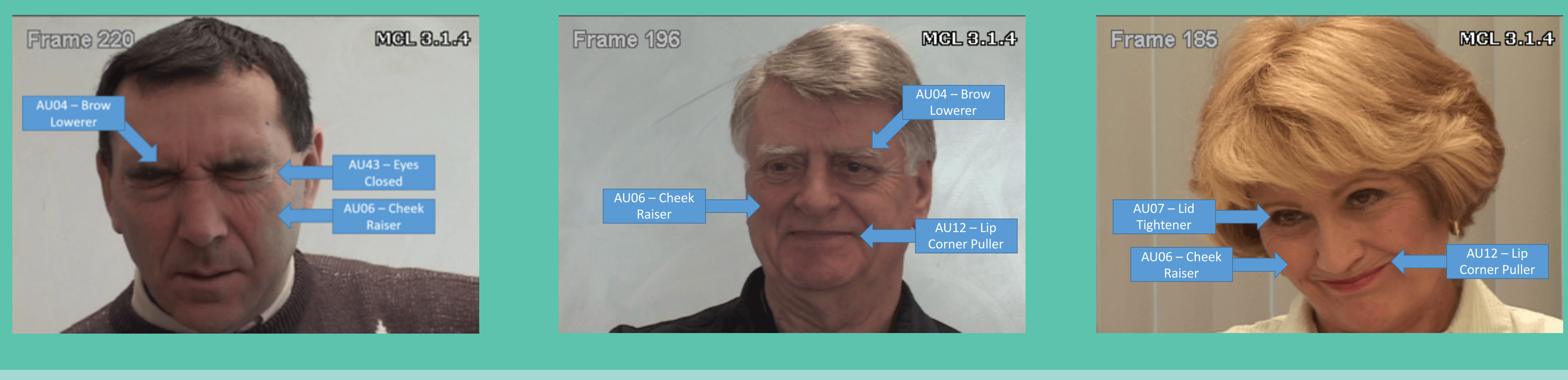
APPROACH

New interdisciplinary approach that combines the following fields (see Fig. 1):

- Bayesian methods in Convolutional Neural Networks (CNNs): They provide uncertainty information about the model and the input data and can enhance model accuracy.
- Inductive Logic Programming (ILP): They provide rule-based explanation understandable to humans.

ACTION UNIT CLASSIFICATION TO ESTIMATE PAIN

- The Facial Action Coding System [1] defines every reproducible facial movement as Action Unit (AU).
- Certain combinations of these AUs can indicate pain.



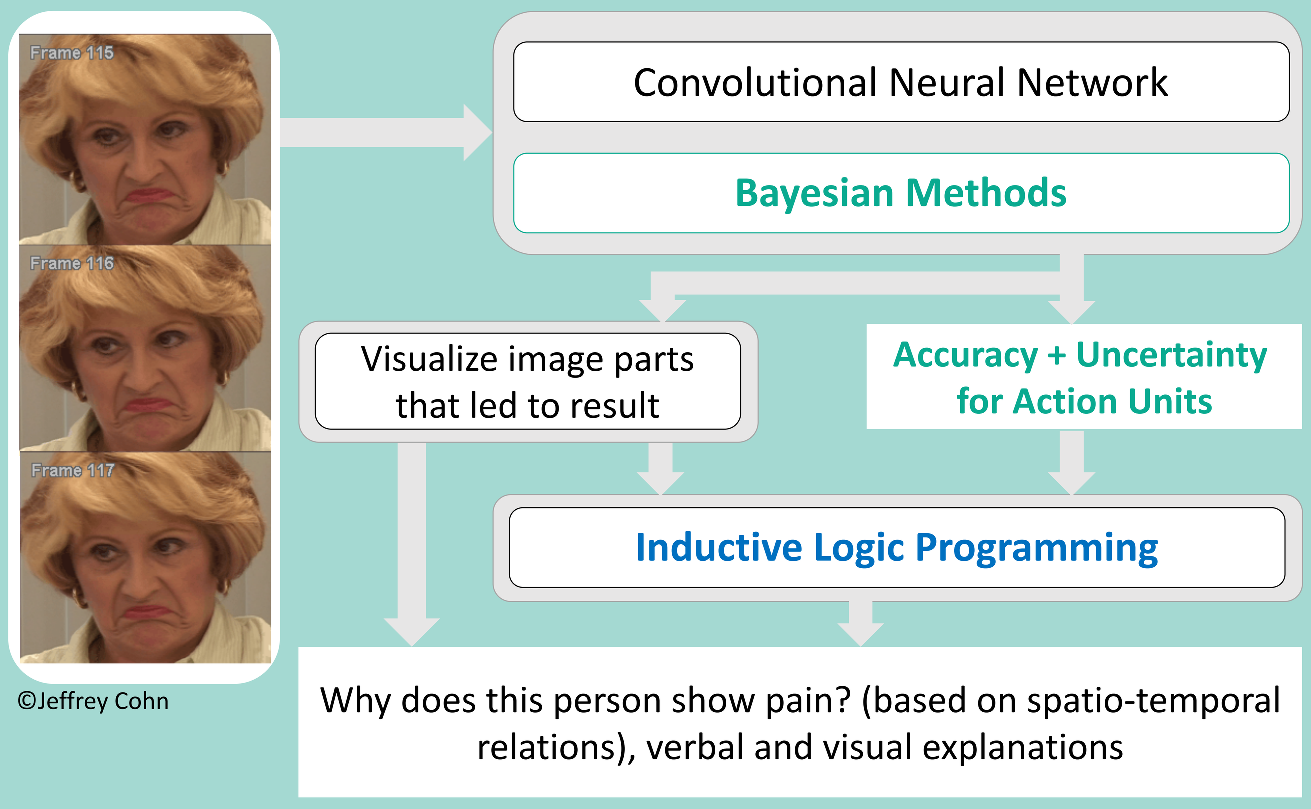


Figure 1. Technical steps for the transparent medical expert companion.

BAYESIAN DEEP LEARNING FOR TRANSPARENCY

- We plan to use Bayesian methods (see Fig. 2) for measuring uncertainty and for improving the accuracy of CNNs.
- Two types of uncertainty can be measured [2]:
  - Epistemic uncertainty displays uncertainty of the model parameters.
  - Aleatoric uncertainty refers to noise in the input data.

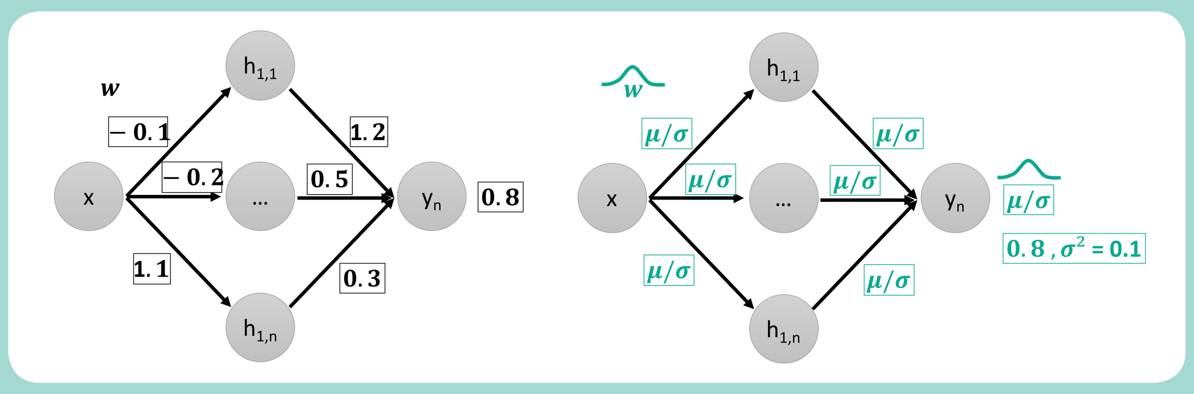


Figure 2. Left: Classic Neural Networks define weights as point estimates, Right: Bayesian Neural Networks define weights by probability distributions

INDUCTIVE LOGIC PROGRAMMING FOR COMPREHENSIBLE EXPLANATIONS

- We apply the inverse entailment algorithm Aleph [3] to induce complex and comprehensible explanations from relational background knowledge as well as positive and negative examples.
- We extract temporal information and spatial relations between relevant image parts, e.g. from heatmaps [4], to fill the back ground knowledge.
- Additionally, we consider the accuracy and uncertainty output of the Bayesian Neural Network.
- Aleph generates a set of rules that characterizes the concepts or classes learned by the CNN (see Fig. 3 and 4).

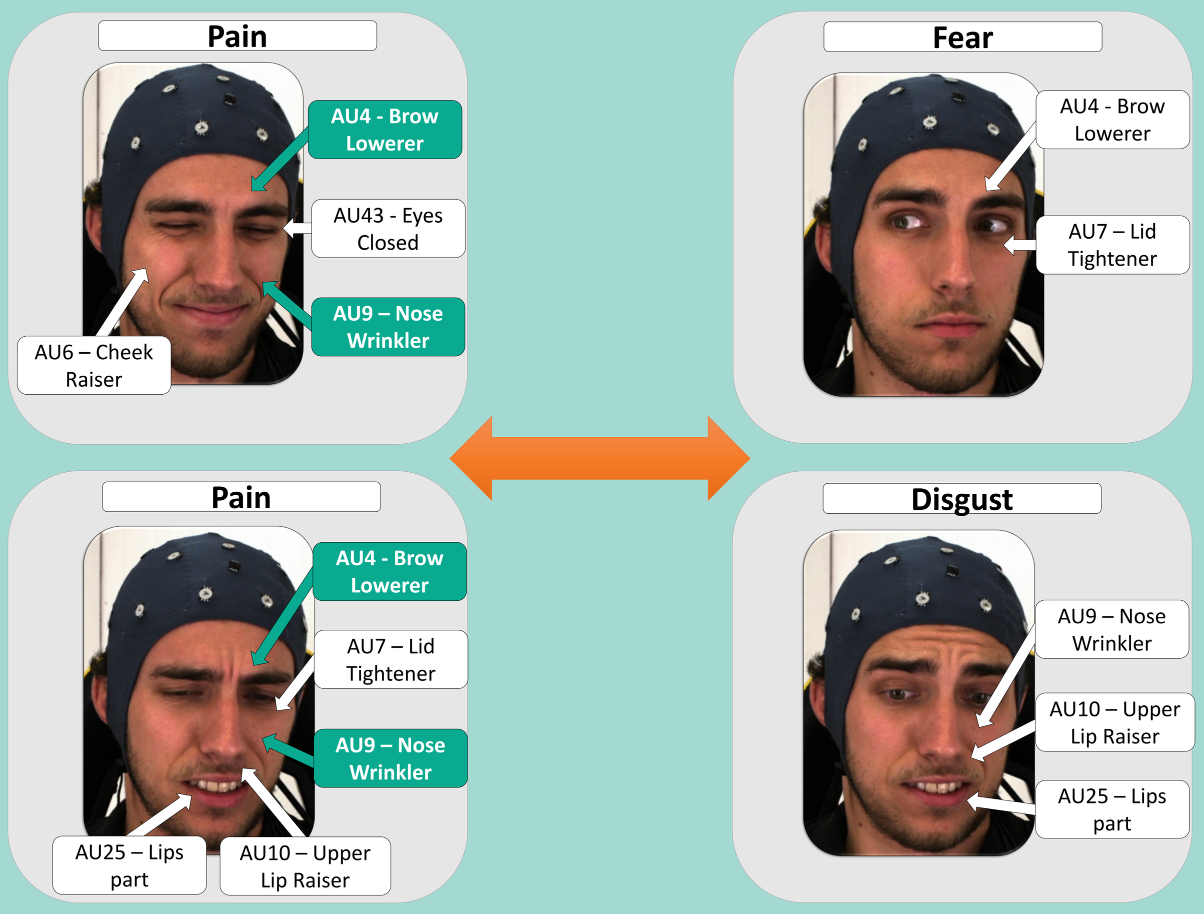


Figure 3. Exemplary images showing pain, fear and disgust. Pain and disgust are particularly difficult to distinguish.

Why Pain and not Fear?	Why Pain and not Disgust?	Why Pain and not Disgust or Fear?
[Rule 1] [Pos cover = 2 Neg cover = 0] **pain(A) :-** **contains(A,au4).**	[Rule 1] [Pos cover = 2 Neg cover = 0] **pain(A) :-** **contains(A,au9).**	[Rule 1] [Pos cover = 2 Neg cover = 0] **pain(A) :-** **contains(A,au4), contains(A,au9).**

Figure 4. Exemplary rules, which were generated by Aleph in order to distinguish pain from fear and disgust

[1] T. Hassan, D. Seuss, …, U. Schmid, "A Practical Approach to Fuse Shape and Appearance Information in a Gaussian Facial Action Estimation Framework," in Proc. ECAI, 2016, pp. 1812-1817.
[2] A. Kendall and Y. Gal, "What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?," in Proc. NIPS, 2017, pp.5574-5584.
[3] J. Rabold, M. Siebers, U. Schmid, "Explaining Black-box Classifiers with ILP-Empowering LIME with Aleph to Approximate Non-linear Decisions with Relational Rules," In Proc. ILP, 2018, pp. 105-117.
[4] S. Bach et al., "On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation," PLOS ONE, 10(7), e0130140, 2015.

Fraunhofer IIS

In collaboration with:

TraMeExCo

CSYS

Otto-Friedrich-Universität Bamberg – University of Bamberg

Research supported by the project "Transparent Medical Expert Companion" (TraMeExCo), funded by the BMBF; Reference: 01IS18056A/B