Make Pain Estimation Transparent: A Roadmap to Fuse Bayesian Deep Learning and Inductive Logic Programming

Ines Rieger, Bettina Finzel, Dominik Seuß, Thomas Wittenberg, Ute Schmid

Abstract— **This roadmap paper describes an interdisciplinary approach to develop an image-based transparent medical expert companion for facial pain estimation. The companion combines human understandable logic rules and visualization to explain the results. Using this information, healthcare professionals can then make informed decisions.**

I. Introduction

Reliable pain analysis is crucial in clinical diagnosis. Patients usually communicate their pain situation verbally to the medical staff. But what about 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?

To solve this problem, we analyze videos showing facial expressions. One of the most successful machine learning methods for image processing are *Convolutional Neural Networks* (CNNs) [1]. However, their drawback is that they do not reveal any explanatory details about the results. Consequently, to provide a basis for profound medical decisions, our goal is to enhance the transparency of CNNs.

For that, we propose a new interdisciplinary approach that combines the following fields (see Fig. 1): A) *Bayesian methods*: They provide uncertainty information about the model and the input data and can enhance model accuracy. B) *Inductive Logic Programming* (ILP): They provide rule-based explanations understandable to humans.

II. Methods

Action Unit Classification The *Facial Action Coding System* [2] defines every reproducible facial movement as *Action Unit* (AU). Certain combinations of these AUs can indicate emotions, stress or pain. The automatic and robust detection of AUs is still a technical challenge; thus current approaches do not meet the requirements for sensitivity and specificity in the medical field.

Bayesian Deep Learning In contrast to classic Neural Networks, the model parameters of Bayesian Neural Networks are not defined by point estimates, but by probability distributions. There are two types of uncertainty [3]: *Epistemic* uncertainty displays uncertainty of the model parameters and *aleatoric* uncertainty refers to noise in the input data. We plan to use Bayesian methods for introducing explainability, but also for improving the accuracy of CNNs.

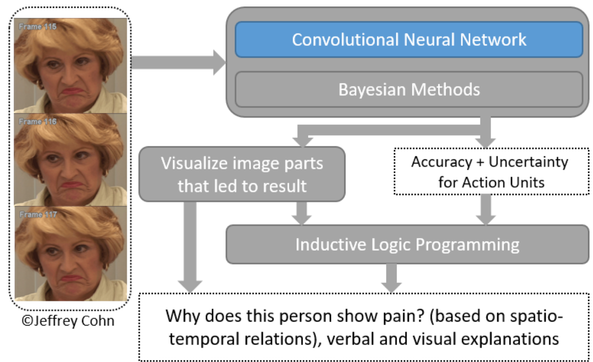

Figure 1. Technical steps for the transparent medical expert companion.

Inductive Logic Programming We apply the inverse entailment algorithm *Aleph* [4] 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 [5], to fill the background knowledge. Additionally, we consider the accuracy and uncertainty output of the Bayesian Neural Network. *Aleph* then generates a set of rules that characterizes the concepts or classes learned by the CNN.

III. Discussion & Conclusion

We envision that our companion will provide transparent estimation of facial pain – with comprehensible verbal and visual explanations and by considering the model uncertainty. We believe that our approach will not only support healthcare professionals with intelligent image analysis, but will also enhance their trust and control in automatic pain assessment.

References

[1] Y. LeCun, Y. Bengio, G. Hinton, "Deep learning," *Nature*, 521(7553), pp. 436-444, 2015.
[2] 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.
[3] A. Kendall and Y. Gal, "What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?," *in Proc. NIPS*, 2017, pp. 5574-5584.
[4] 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.
[5] S. Bach et al., "On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation," *PLOS ONE*, 10(7), e0130140, 2015.

*Research supported by the project "Transparent Medical Expert Companion" (TraMeExCo), funded by the BMBF; Reference: 011S18056A/B

I. Rieger, Fraunhofer Institute for Integrated Circuits IIS, 91058 Erlangen, Germany, phone: +49 91317765179; e-mail: ines.rieger@iis.fraunhofer.de.