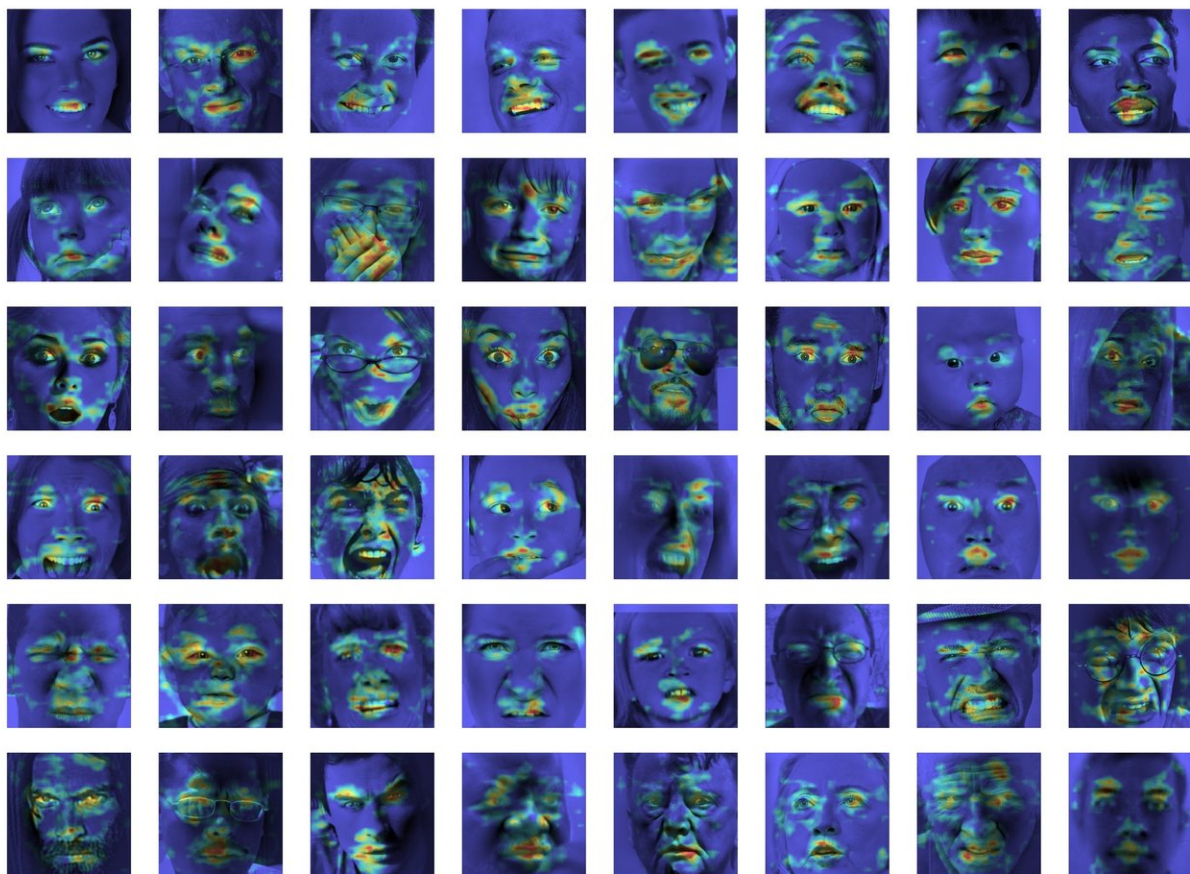# An emotional deep alignment network (DAN) to classify and visualize emotions

November 6 2018, by Ingrid Fadelli

Credit: Tautkutè & Trzcinski

Researchers at the Polish-Japanese Academy of Information Technology and Warsaw University of Technology have developed a deep alignment

network (DAN) model to classify and visualize emotions. Their method was found to outperform state-of-the-art emotion classification methods on two benchmark datasets.

Developing models that can recognize and classify human emotions is a key challenge in the field of machine learning and computer vision. Most existing emotion recognition tools use multi-layered convolutional networks, which do not explicitly infer facial features in the classification phase.

Ivona Tautkute and Tomasz Trzcinski, the researchers who carried out the recent study, were initially working on a system for a California-based startup that could be integrated into autonomous cars. This system was able to count passengers based on data extracted from a single video camera mounted inside the car.

At a later stage, the two researchers started exploring models that could do more than this, creating broader statistics about passengers by estimating their age and gender. An obvious extension of this system was for it to detect facial expressions and emotions, as well.

"As the system was to be used with elderly passengers, it was important to capture negative and positive emotions associated with driver disengagement," Tautkute explained. "Existing approaches for emotion recognition are far from perfect, so we started looking into interesting new ways to improve. An idea came to us after a discussion with a fellow computer vision researcher, Marek Kowalski, who was working on facial alignment with deep alignment network (DAN). Location of facial landmarks is straightforwardly related to expressed emotion, so we were curious about whether we could build a system that would combine those two tasks."

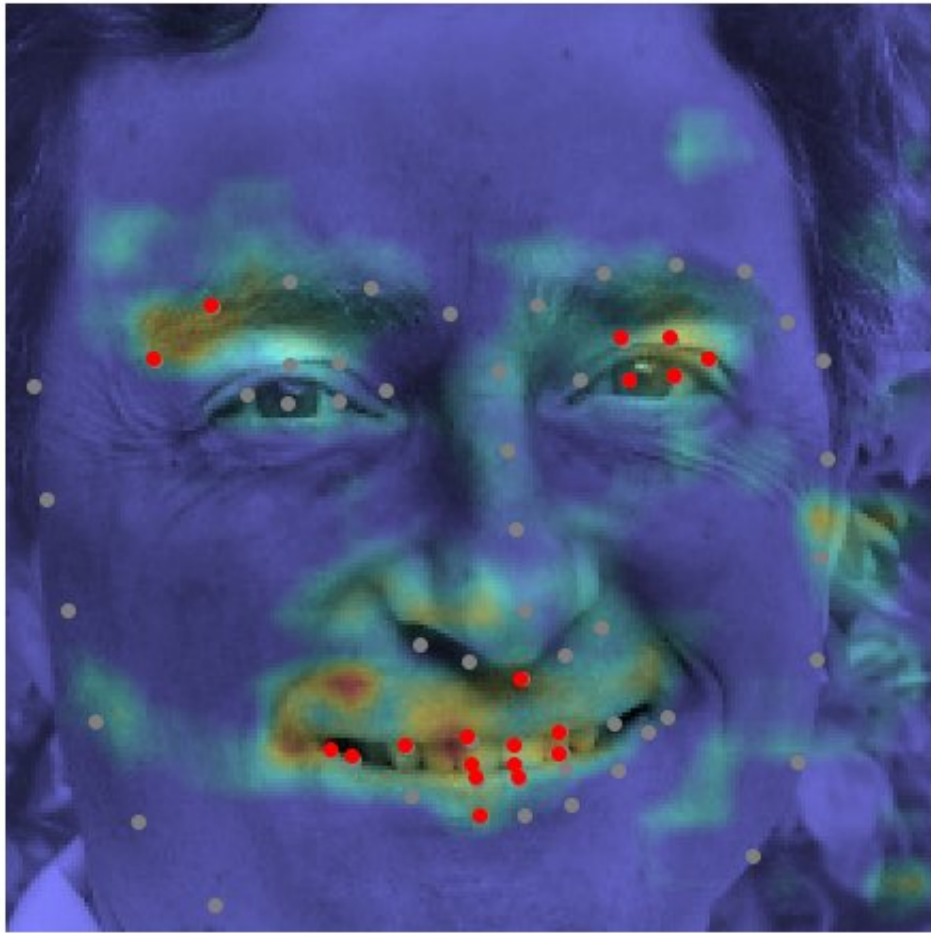EmotionalDAN, the model devised by Tautkute and Trzcinski, is an

adaptation of Kowalski's DAN model, which includes a term related to facial features. Thanks to this modification, their model simultaneously learns the location of both facial landmarks and expressed emotion.
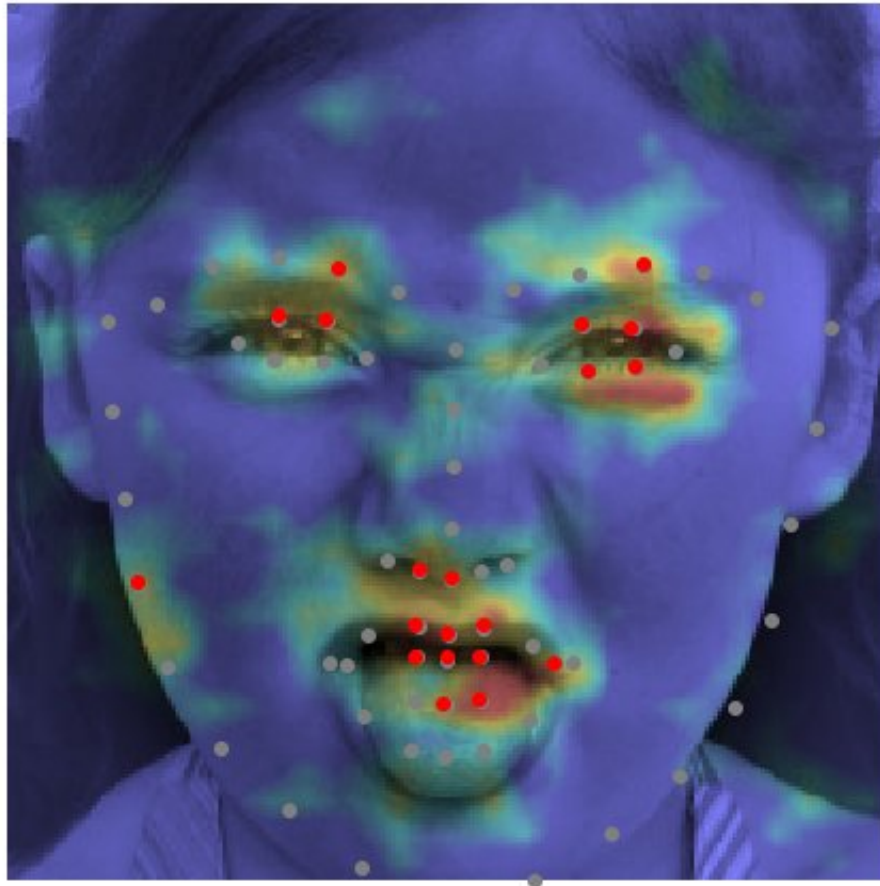
Credit: Tautkutè & Trzcinski
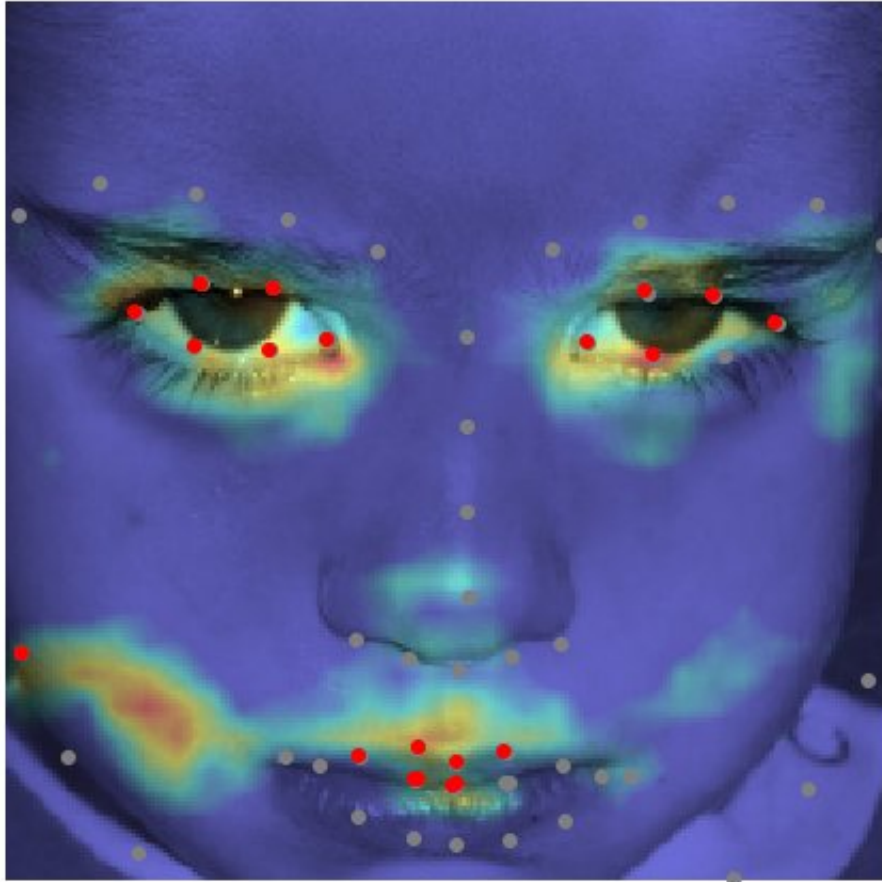
Credit: Tautkutè & Trzcinski
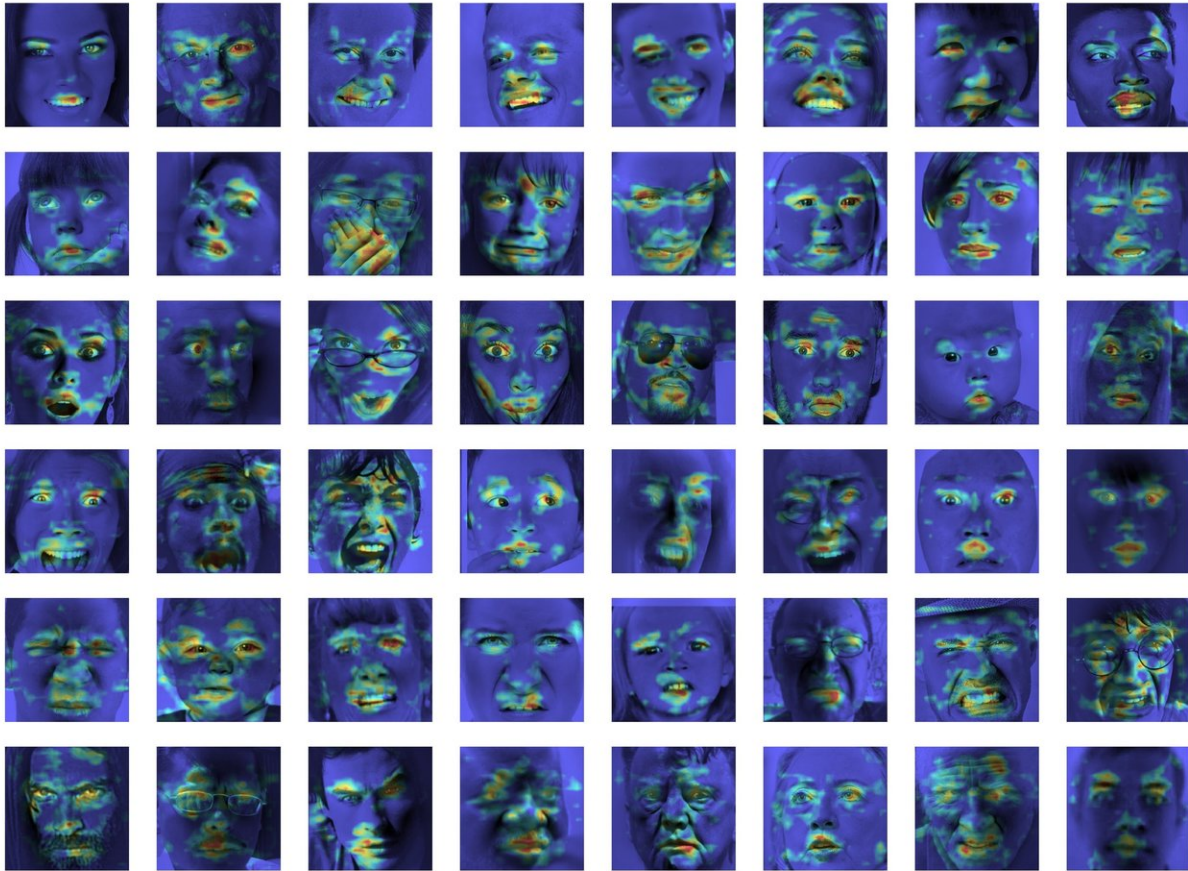
Credit: Tautkutè & Trzcinski

Credit: Tautkutè & Trzcinski

"We achieved this by extending the loss function of the original DAN by a term responsible for emotion classification," Tautkute said. "The neural network is trained in consecutive stages that allow for refinement of facial landmarks and learned emotions. There is also a transfer of

information between stages, which keeps track of the normalized face input, feature map and the landmarks heat map."

In an initial evaluation, EmotionalDAN outperformed state-of-the art [classification methods](#) by 5 percent on two benchmark datasets, namely CK+ and ISED. The researchers were also able to visualize image regions analyzed by their model when making a decision. Their observations revealed that EmotionalDAN could correctly identify facial landmarks associated with the expression of emotions in humans.

"What is really interesting about our study is that even though we do not feed any emotion-related spatial information to the network, the model is capable of learning by itself which face regions should be looked at when trying to understand [facial expressions](#)," Tautkute said. "We humans intuitively look at a person's eyes and mouth to notice smiles or sadness, but the neural network only sees a matrix of pixels. Verifying what image regions are activated for given classification decision brings us one step closer to understanding the model and how it makes decisions."

Credit: Tautkutè & Trzcinski.

Despite the very promising results achieved by EmotionalDAN and other emotion recognition tools, understanding human emotions remains a very complex task. While existing systems have achieved remarkable results, they have primarily been able to do so when emotions are expressed to a significant degree.

In real-life situations, however, the emotional cues expressed by humans are often far subtler. For instance, a person's happiness might not always be conveyed by showing all teeth in a broad smile, but might merely entail a slight movement of the lip corners.

"It would be really interesting to understand more subjective aspects of emotion and how their expression differs between individuals," Tautkute said. "To go further, one could try to distinguish fake emotions from genuine ones. For example, neurologists state that different facial muscles are involved in real and fake smiles. In particular, eye muscles don't contract in the forced expression. It would be interesting to discover similar relationships using information learnt from data."

More information: Classifying and visualizing emotions with emotional DAN. arXiv: 1810.10529v1 [cs.CV]. arxiv.org/pdf/1810.10529.pdf

Deep alignment network: a convolutional neural network for robust face alignment. arXiv:1706.01789 [cs.CV]. arxiv.org/abs/1706.01789