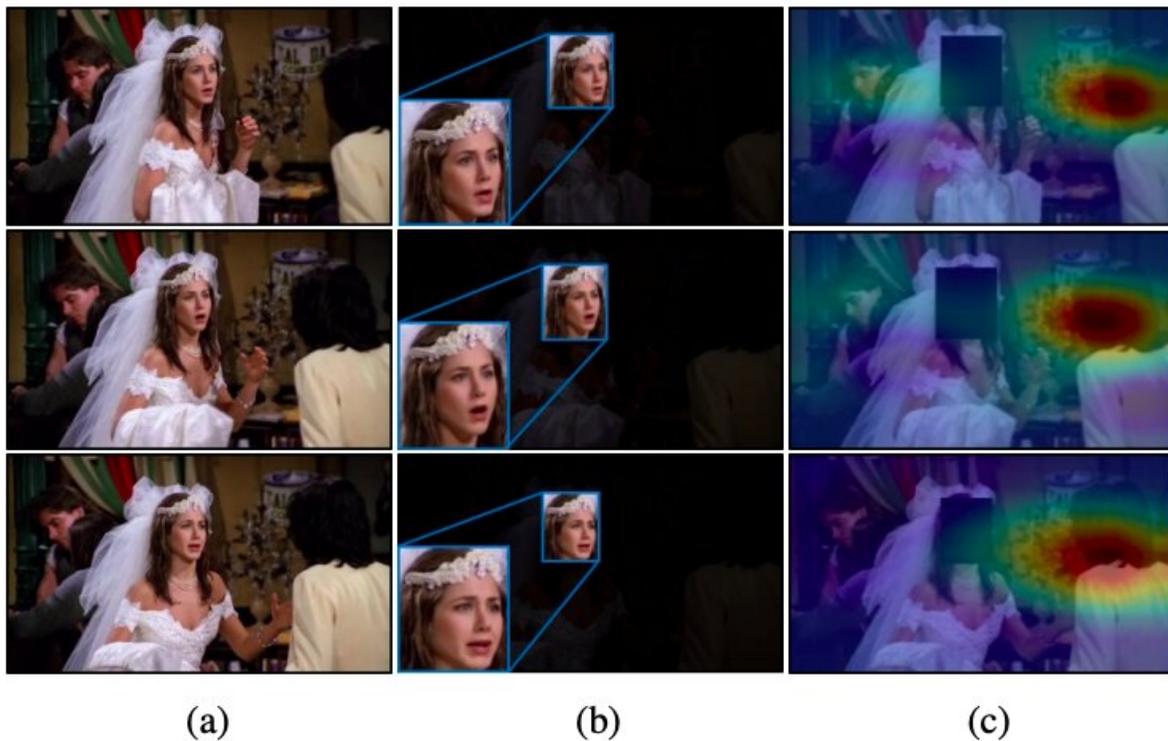# A deep learning technique for context-aware emotion recognition

August 28 2019, by Ingrid Fadelli

Intuition of CAER-Net for untrimmed videos, as in (a) conventional methods that leverage the facial regions only, as in (b), often fail to recognize emotion. Unlike these methods, CAER-Net focuses on both face and attentive context regions, as seen in (c). Credit: Lee et al.

A team of researchers at Yonsei University and École Polytechnique

Fédérale de Lausanne (EPFL) has recently developed a new technique that can recognize emotions by analyzing people's faces in images along with contextual features. They presented and outlined their deep learning-based architecture, called CAER-Net, in a paper pre-published on arXiv.

For several years, researchers worldwide have been trying to develop tools for automatically detecting human emotions by analyzing images, videos or audio clips. These tools could have numerous applications, for instance, improving robot-human interactions or helping doctors to identify signs of mental or neural disorders (e.g.,, based on atypical speech patterns, facial features, etc.).

So far, the majority of techniques for recognizing emotions in images have been based on the analysis of people's facial expressions, essentially assuming that these expressions best convey humans' emotional responses. As a result, most datasets for training and evaluating emotion recognition tools (e.g., the AFEW and FER2013 datasets) only contain cropped images of human faces.

A key limitation of conventional emotion recognition tools is that they fail to achieve satisfactory performance when emotional signals in people's faces are ambiguous or indistinguishable. In contrast with these approaches, human beings are able to recognize others' emotions based not only on their facial expressions, but also on contextual clues (e.g., the actions they are performing, their interactions with others, where they are, etc.).

Past studies suggests that analyzing both facial expressions and context-related features can significantly boost the performance of emotion recognition tools. Inspired by these findings, the researchers at Yonsei and EPFL set out to develop a deep learning-based architecture that can recognize people's emotions in images based on both their facial expressions and contextual information.
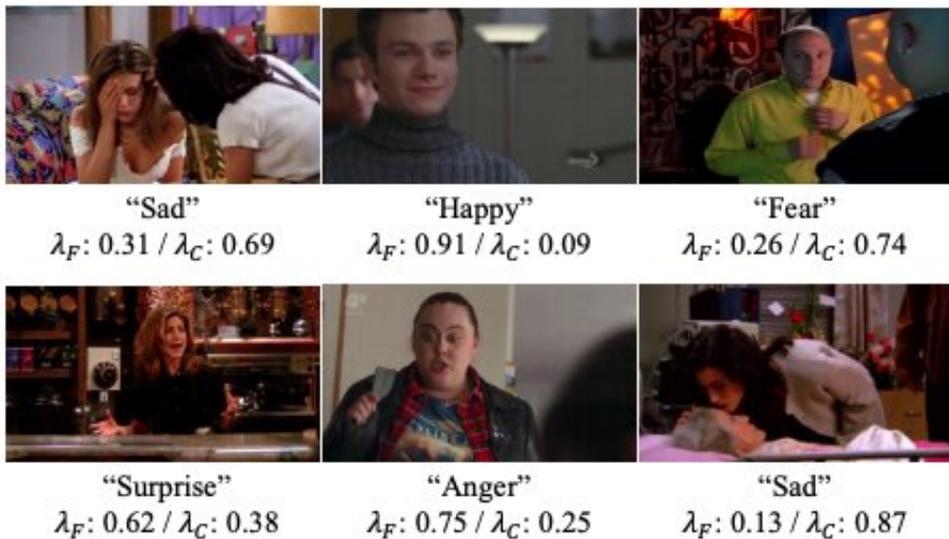
Examples of attention weights in the neural networks developed by the researchers. Credit: Lee et al.

"We present deep networks for context-aware emotion recognition, called CAER-Net, that exploit not only human facial expression, but also context information, in a joint and boosting manner," the researchers wrote in their paper. "The key idea is to hide human faces in a visual scene and seek other contexts based on an attention mechanism."

CAER-Net, the architecture developed by researchers, is composed of two key sub-networks and encoders that separately extract facial features and contextual regions in an image. These two types of features are then combined using adaptive fusion networks and analyzed together to predict the emotions of people in a given image.

In addition to CAER-Net, the researchers also introduced a new dataset for context-aware emotion recognition, which they refer to as CAER. Images in this dataset portray both people's faces and their

surroundings/context, hence it could serve as a more effective benchmark for training evaluating emotion recognition techniques.

The researchers evaluated their emotion recognition technique in a series of experiments, using both the dataset they compiled and the AFEW dataset. Their findings suggest that analyzing both [facial expressions](#) and contextual information can considerably boost the performance of emotion recognition tools, as indicated by previous studies.

"We hope that the results of this study will facilitate further advances in context-aware emotion recognition and its related tasks," the researchers wrote.

More information: Jiyoung Lee, et al. Context-aware emotion recognition networks. arXiv:1908.05913 [cs.CV]. arxiv.org/abs/1908.05913

Context-Aware Emotion Recognition: Large-Scale Static & Dynamic Emotion Recognition Benchmark
ICCV 2019 caer-dataset.github.io

© 2019 Science X Network

Citation: A deep learning technique for context-aware emotion recognition (2019, August 28) retrieved 17 June 2024 from https://techxplore.com/news/2019-08-deep-technique-context-aware-emotion-recognition.html