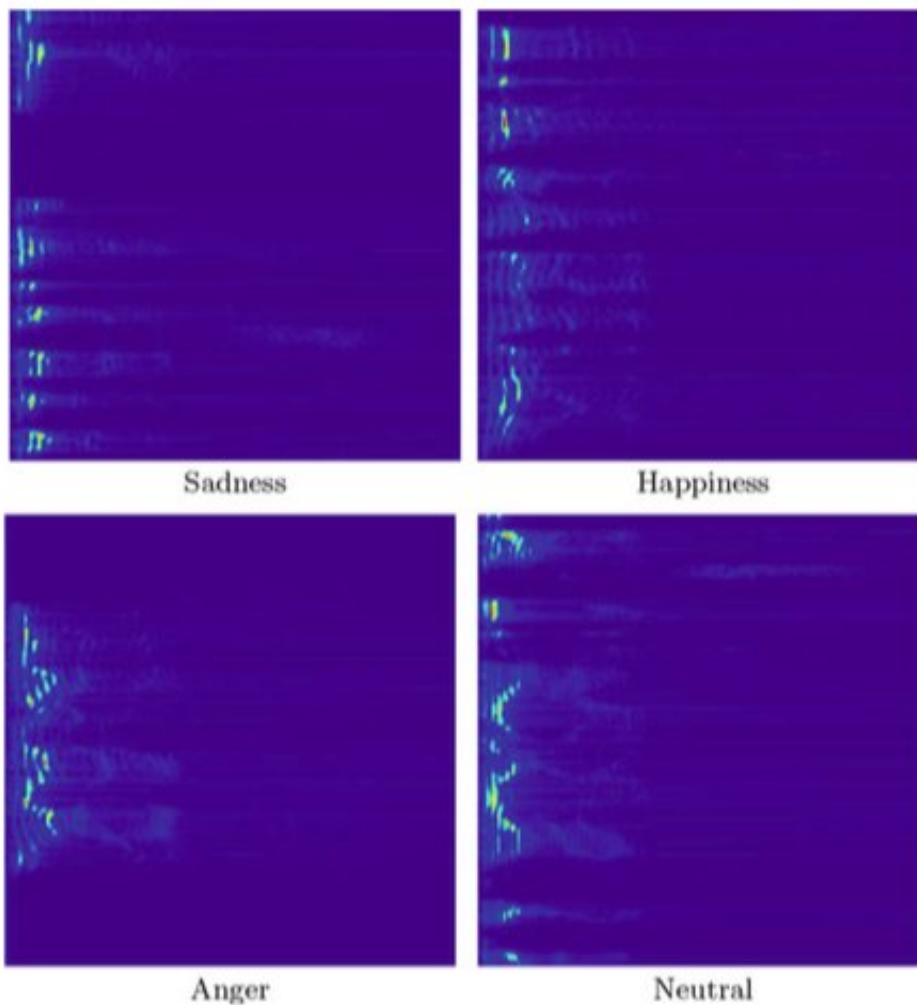


Emotion recognition based on paralinguistic information

December 11 2018, by Ingrid Fadelli



Example spectrograms from each of the 4 included classes. Credit: Papakostas et al.

Researchers at the University of Texas at Arlington have recently explored the use of machine learning for emotion recognition based solely on paralinguistic information. Paralinguistics are aspects of spoken communication that do not involve words, such as pitch, volume, intonation, etc.

Recent advances in [machine learning](#) have led to the development of tools that can recognize [emotional states](#) by analyzing images, voice recordings, electroencephalograms or electrocardiograms. These tools could have several interesting applications, for instance, enabling more efficient human-computer interactions in which a computer recognizes and responds to a human user's emotions.

"In general, one may argue that speech carries two distinct types of [information](#): explicit or linguistic information, which concerns articulated patterns by the speaker; and implicit or paralinguistic information, which concerns the variation in pronunciation of the linguistic patterns," the researchers wrote in their [paper](#), published in the *Advances in Experimental Medicine and Biology* book series. "Using either or both types of information, one may attempt to classify an audio segment that consists of speech, based on the emotion(s) it carries. However, [emotion recognition](#) from speech appears to be a significantly difficult task even for a human, no matter if he/she is an expert in this field (e.g. a psychologist)."

Many existing [automatic speech recognition](#) (ASR) approaches try to recognize emotions from speech by analyzing both linguistic and paralinguistic information. By partly focusing on linguistic properties, these models have several disadvantages, such as a strict language-dependency. The researchers hence decided to focus on emotion [recognition](#) based only on the analysis of paralinguistic information, with the hope of attaining multi-lingual emotion recognition.

"In this paper, we aim to analyze speakers' emotions based solely on paralinguistic information," the researchers wrote in their paper. "We compare two machine learning approaches, namely a convolutional neural network (CNN) and a support vector machine (SVM)."

The researchers trained a CNN model on raw spectrograms and an SVM model on a set of low-level features. Both models were trained and evaluated using three widely known emotional speech datasets: EMOVO, SAVEE, and EMO-DB. These datasets contain [emotional speech](#) recordings in different languages—Italian, English and German respectively.

The two machine learning models were trained to recognize four common emotion classes: happiness, sadness, anger and neutral. The researchers carried out three experiments for each machine learning approach, where a single [dataset](#) was used for testing and the remaining two for training.

"A major difficulty resulting from the choice of datasets is the great difference between languages, since besides the linguistic differences, there is also a big variability in the way each emotion is expressed," the researchers wrote in their paper.

Overall, they found that the SVM performed far better than the CNN, achieving the best results when trained on the SAVEE and EMOVO datasets, but tested on EMO-DB. These results were promising but not optimal, suggesting that we are still a long way from attaining consistently effective multi-lingual emotion recognition.

"Our plans for future work include the usage of more datasets for training and evaluation," the researchers wrote in their paper. "We also aim to investigate other pre-trained deep learning networks, since we feel that deep learning may significantly contribute to the problem at

hand. Finally, among our plans is to apply such approaches to real-life problems, e.g. emotion recognition within training and/or educational programs."

More information: Michalis Papakostas et al. Recognizing Emotional States Using Speech Information, *GeNeDis 2016* (2017). [DOI: 10.1007/978-3-319-57348-9_13](https://doi.org/10.1007/978-3-319-57348-9_13)

© 2018 Science X Network

Citation: Emotion recognition based on paralinguistic information (2018, December 11)
retrieved 14 April 2024 from
<https://techxplore.com/news/2018-12-emotion-recognition-based-paralinguistic.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.