Summary for AVEC 2017 - Real-life Depression and Affect Challenge and Workshop
Fabien Ringeval, Björn Schuller, Michel Valstar, Jonathan Gratch, Roddy Cowie, Maja Pantic

To cite this version:
Fabien Ringeval, Björn Schuller, Michel Valstar, Jonathan Gratch, Roddy Cowie, et al.. Summary for AVEC 2017 - Real-life Depression and Affect Challenge and Workshop. ACM Multimedia, Oct 2017, Mountain View, United States. 10.1145/3123266.3132049 . hal-02080833

HAL Id: hal-02080833
https://hal.archives-ouvertes.fr/hal-02080833
Submitted on 27 Mar 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Summary for AVEC 2017 – Real-life Depression and Affect Challenge and Workshop

Fabien Ringeval  
Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG  
F-38000 Grenoble, France  
fabien.ringeval@imag.fr

Björn Schuller∗  
University of Passau, Chair of Complex & Intelligent Systems  
Passau, Germany  
schuller@ieee.com

Michel Valstar  
University of Nottingham  
Mixed Reality Lab  
Nottingham, UK  
michel.valstar@nottingham.ac.uk

Jonathan Gratch  
University of Southern California  
Institute for Creative Technologies  
Los Angeles, CA, USA  
gratch@ict.usc.edu

Roddy Cowie  
Queen’s University Belfast  
Department of Psychology  
Dublin, UK  
r.cowie@qub.ac.uk

Maja Pantic†  
Imperial College London  
Department of Computing  
London, UK  
m.pantic@imperial.ac.uk

ABSTRACT

The seventh Audio-Visual Emotion Challenge and workshop AVEC 2017 was held in conjunction with ACM Multimedia’17. This year, the AVEC series addresses two distinct sub-challenges: emotion recognition and depression detection. The Affect Sub-Challenge is based on a novel dataset of human-human interactions recorded ‘in-the-wild’, whereas the Depression Sub-Challenge is based on the same dataset as the one used in AVEC 2016, with human-agent interactions. In this summary, we mainly describe participation and its conditions.

CCS CONCEPTS

• General and reference → Performance; • Computing methodologies → Biometrics;

KEYWORDS

Affective Computing; Social Signal Processing; Automatic Emotion/Depression Recognition

ACM Reference format:


DOI: https://doi.org/10.1145/3123266.3132049

1 INTRODUCTION

This year’s Audio-Visual Emotion Challenge and workshop (AVEC 2017) has been organised in conjunction with the 25th ACM International Conference on Multimedia, MM 2017, held in Mountain View (CA), USA, 23 – 27 October 2017.

The AVEC 2017’s theme is ‘Real-life Depression and Affect’ and it is the seventh competition event aimed at comparison of multimedia processing and machine learning methods for automatic audio, visual and audio-visual emotion analysis, with all of the participants competing under strictly the same conditions in this first of its kind series [4, 6–8, 10, 11]. However, further similar endeavours that have been undertaken since AVEC started in 2011 are to be noted, e.g., [1].

The AVEC 2017 edition shall help raise the bar for emotion and depression detection by challenging participants to estimate the level of depression and affect from audiovisual data captured in real-life conditions, and it will continue to bridge the gap between excellent research on emotion and depression recognition and low comparability of results.

As before, the goal of the Challenge is to compare the relative merits of the approaches for audiovisual emotion recognition and severity of depression estimation under well-defined and strictly comparable conditions, and establish to what extent fusion of the approaches is possible and beneficial. The main underlying motivation is the need to advance emotion recognition and depression estimation, and affective computing in general, for multimedia retrieval to a level where behaviors expressed during human-human, or human-agent interactions, can be reliably sensed in real-life conditions, as this is exactly the type of data that applications would have to face in the real world.

We called for participation in two Sub-Challenges, each focusing on specific dimensions of the spectrum of human behaviours; (i) Affect Sub-Challenge (ASC): participants were required to perform fully continuous affect recognition of three affective dimensions (arousal, valence, and likability), from audiovisual recordings of human-human interactions; (ii) Depression Sub-Challenge (DSC): participants were asked to assess the level of depression severity based on PHQ-8 [3] self-report scores, from audiovisual recordings of human-agent interactions.
As benchmarking database, a subset of a novel database of human-human interactions recorded ‘in-the-wild’ was used for the ASC; the Sentiment Analysis in the Wild (SEWA) database\(^1\). In this data set, audiovisual signals were acquired in various places, such as home and work place, and with arbitrary personal equipments. The German culture of the SEWA database was used for the ASC. The DSC is a refined re-run of the AVEC 2016 challenge \[^9\], based on the Distress Analysis Interview Corpus – A Wizard of Oz (DAIC-WOZ) database\(^2\) \[^2\]. Whereas the severity of depression was estimated as a binary task in AVEC 2016, we addressed this year the inference of the level of severity as a continuous value.

Besides participation in the Challenge we called for papers addressing the overall topics of this workshop, in particular works that address issues concerning fusion of audio-visual cues for affect recognition.

In the following sections, we will describe the participation in this year and outline the conditions for participation in particular in the competitive challenge event. We further acknowledge those that helped realise AVEC 2017.

2 CHALLENGE CONDITIONS

A baseline paper explaining the dataset, the challenge evaluation procedure, baseline features and baseline results, was made available while the challenge was ongoing \[^5\]. Features extraction methods and machine learning algorithms relied on open-source softwares, and scripts to reproduce the baseline performance were also provided for both Sub-Challenges. As in previous years, we required to sign an end user license agreement and for the DSC no raw video was made available, only video features, because of the sensitive nature of the data. After downloading the data, participants could directly start their own experiments with the train and development sets. Once they found their best method they had to write a paper for the workshop. At the same time, they could compute their results per instance of the test set. Participants’ results needed to be sent as a single packed file per Sub-Challenge to the organisers by email and scores were returned within 24 hours during typical working days.

Each participant had up to five submission attempts per Sub-Challenge (ASC/DSC). The organisers provided for each affective dimension the Concordance Correlation Coefficient (CCC), which is used to rank participants on the ASC, and the Root Mean Squared Error (RMSE) measure was used to rank participants on the DSC.

3 PARTICIPATION

The call for participation and papers attracted registrations of 41 teams from all over the world. 13 teams submitted results for the ASC, and 10 teams submitted results for the DSC. Finally, 17 paper submissions were received, which were assigned three reviewers, each, and reviewed independently. AVEC 2017 reviewing was double blind, and acceptance was based on relevance to the workshop, novelty, technical quality, and performance on the test partition. The program committee accepted 8 papers in addition to the independently reviewed baseline paper as oral presentation. Again, we hope that these proceedings will serve as a valuable reference for researchers and developers in the area of audio-visual emotion recognition and depression analysis in real-life settings.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Union’s 7th Framework Programme through the ERC Starting Grant No. 338164 (iHEARu), the Horizon 2020 Programme through the Innovative Action No. 645094 (SEWA), and the Research Innovative Action No. 645378 (ARIA-VALLUSPA), and No. 688835 (DE-ENIGMA). The data provided for the depression severity Sub-Challenge was sponsored by DARPA and the Army Research Laboratory under contracts W911NF-04-D-0005 and W911NF-14-D-0005. The authors further thank the sponsors of the challenge – audEERING GmbH and the aaac.

REFERENCES


\[^1\] http://sewaproject.eu

\[^2\] http://dcapswoz.ict.usc.edu