AN INTERPRETABLE, EFFICIENT, AND ACCURATE FRAMEWORK FOR FACIAL EXPRESSION RECOGNITION IN THE WILD

M. Lionello, E. Bucci, E. Sampaolo, L. Cecchetti

Social and Affective Neuroscience (SANe) group, MoMiLab - IMT School for Advanced Studies, Lucca

INTRODUCTION

Facial expressions are immediate means for communicating emotions. **F**acial **E**xpression **R**ecognition in the wild (FER) is crucial for social, commercial, and scientific tasks. Deep Learning models (DNN) excel in FER but generally lack self-explanatory features. Here, we propose a framework relying on **pairwise distances** face *landmarks*. First, we employ this method in a classification task using three publicly available datasets. Additionally, we present findings from a pilot study aimed at exploring **inter-subject correlations** (ISC).

STUDY I - Classification task

Materials and Methods

Three video datasets are employed: RAVDESS, including different modalities for each emotion; PEDFE, which comprises human ratings for genuine and posed expressions (hit-rates); FACEXP, a dataset with multiple repetitions of the same emotion for each actor. To model facial displacement, we detect face landmarks with **Medusa**. ONE-NN models and k-fold cross-validation are used to assess the overall performance. In RAVDESS and FACEXP, we also perform leave-1-subject-out, while in PEDFE we test varying training sizes and hit-rate levels. Analyses are restricted to Ekman's six basic emotions.

Results

The results are summarised in the **Table 1**. Moreover, in PEDFE, including low hit-rates showed a marginal impact on classification of genuine expressions (see **Fig. 1**), while the training curves for the increasing training set are shown in **Fig. 2a**. In **Fig. 2b**, the confusion matrix for the best model (acc = 0.71) is presented.

Dataset	Overall Accuracy	1SubOut	RepTr
RAVDESS	87%	~35%	--
FACEXP	96%	~35%	**1rep:** 90%
PEDFE	**Posed** 70% **Gen** 62% **Mix** 67%	—	--

Table 1

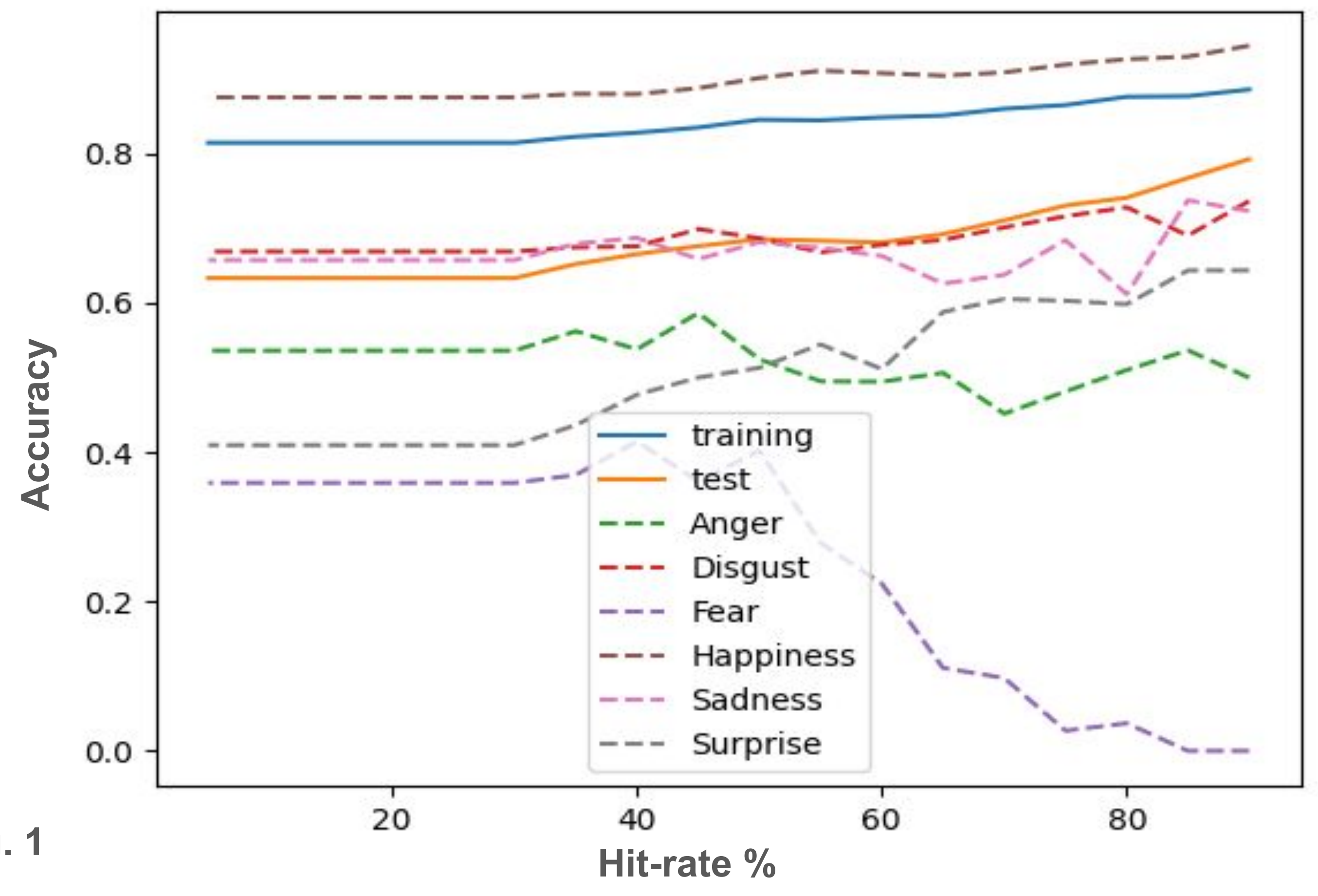

Fig. 1

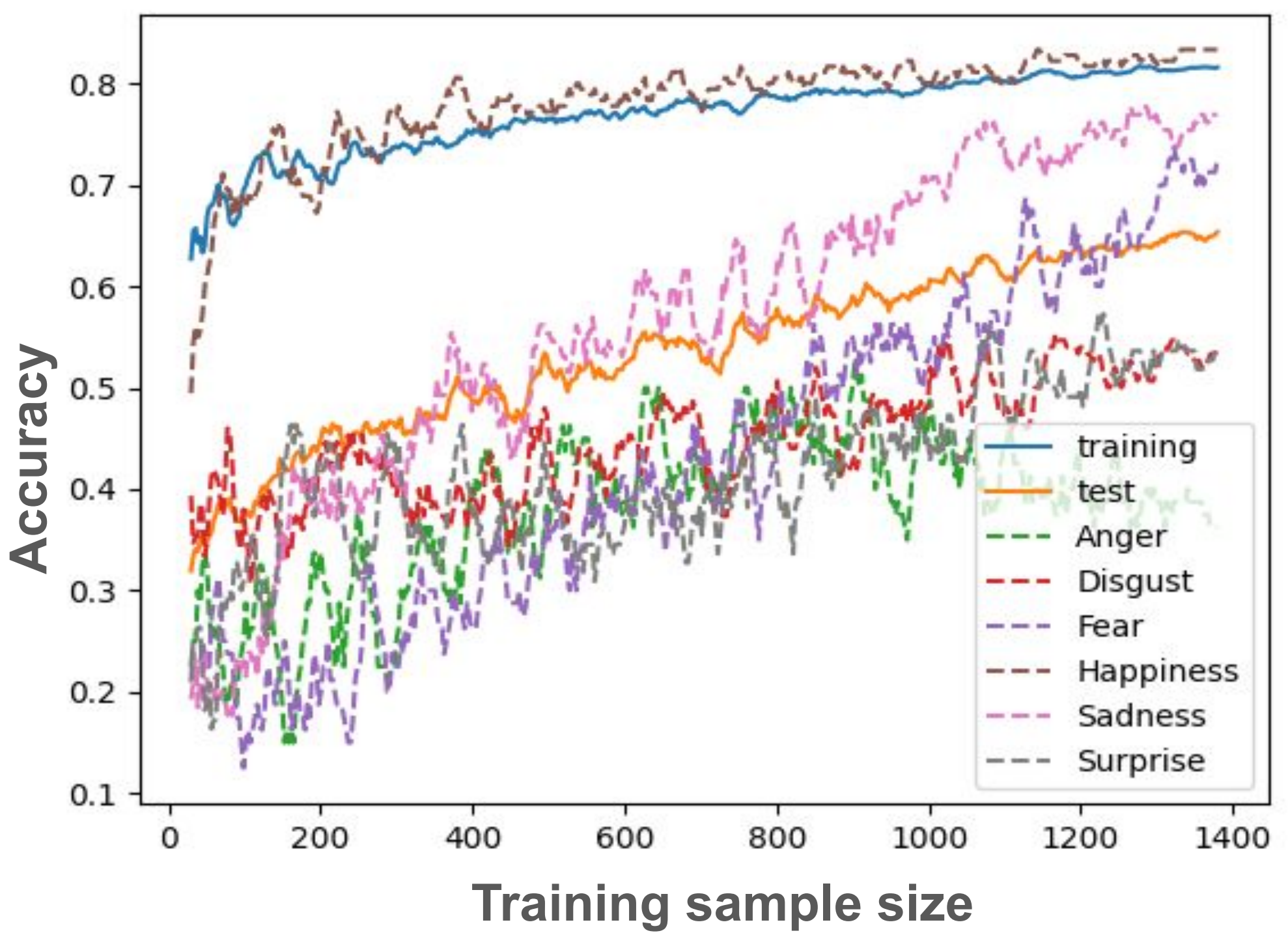

Fig. 2a

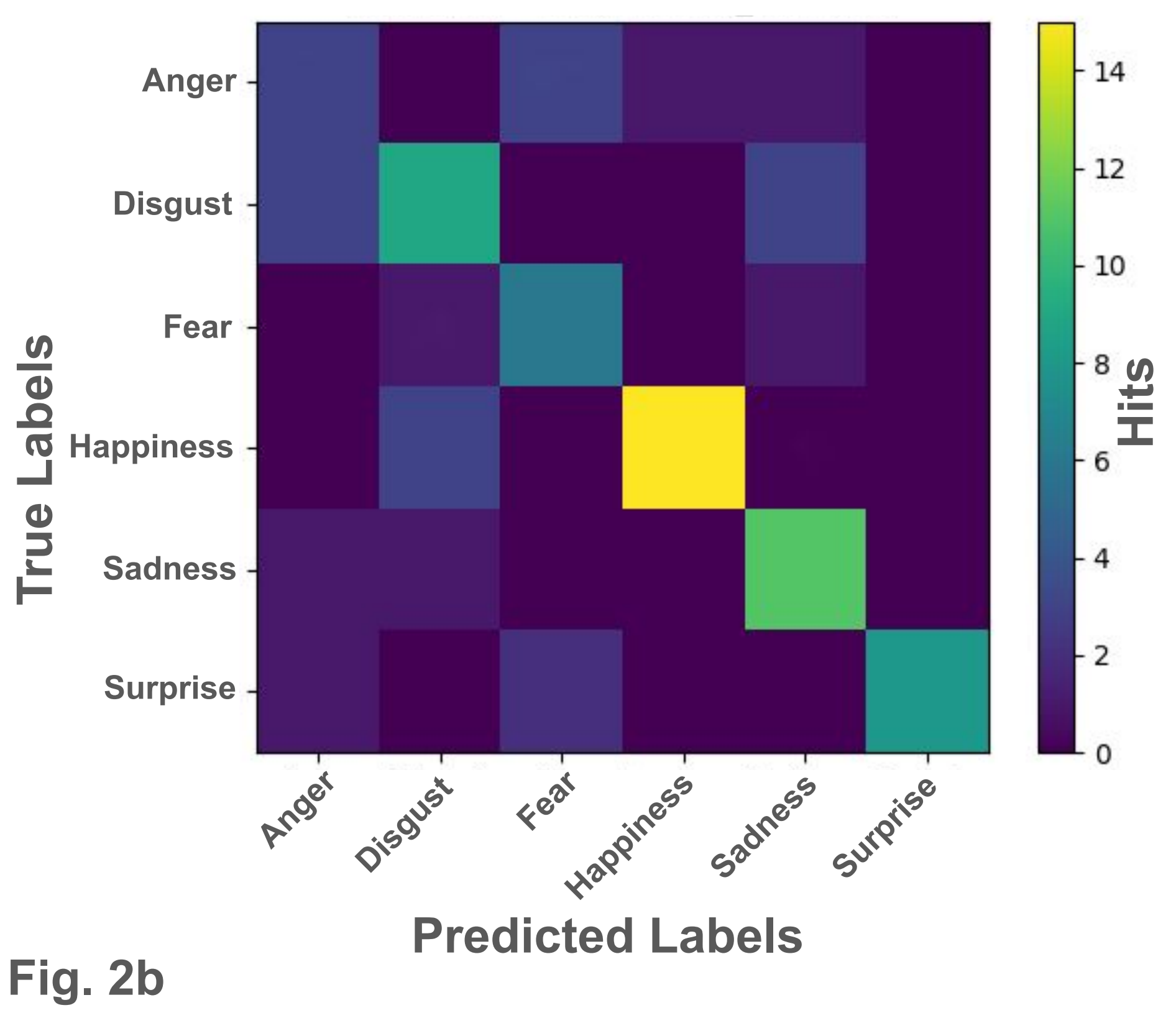

Fig. 2b

STUDY II - Intersubject synchronicity

Materials and Methods

In the second study, by using a similar framework to the first study, we measure how participants **synchronize** their facial expressions while watching a movie. 21 participants are asked to watch a 20-minutes movie while their **facial expressions** are recorded via a gopro camera. During the movie, the participants rate their emotional intensity with a keyboard. Similarly to Study I, their facial expressions are processed with Medusa and **single vertex trajectory** and **coupled vertices distances** are calculated. Temporal-ISC (**t-ISC**) is measured by means of a sliding window (L=5s, step=1.66s). Multiple comparison testing is used to assess statistical significance (α=0.05) with 200 shuffling of the time-points for each vertex by preserving the identity of the participant. The 95th and 5th perc. are calculated among the maxima and the minima of all the vertices in each shuffling. The t-ISC series obtained are correlated with the agreement of the emotional peaks reported by participants during the behavioural study (see **Fig. 3**).

Results

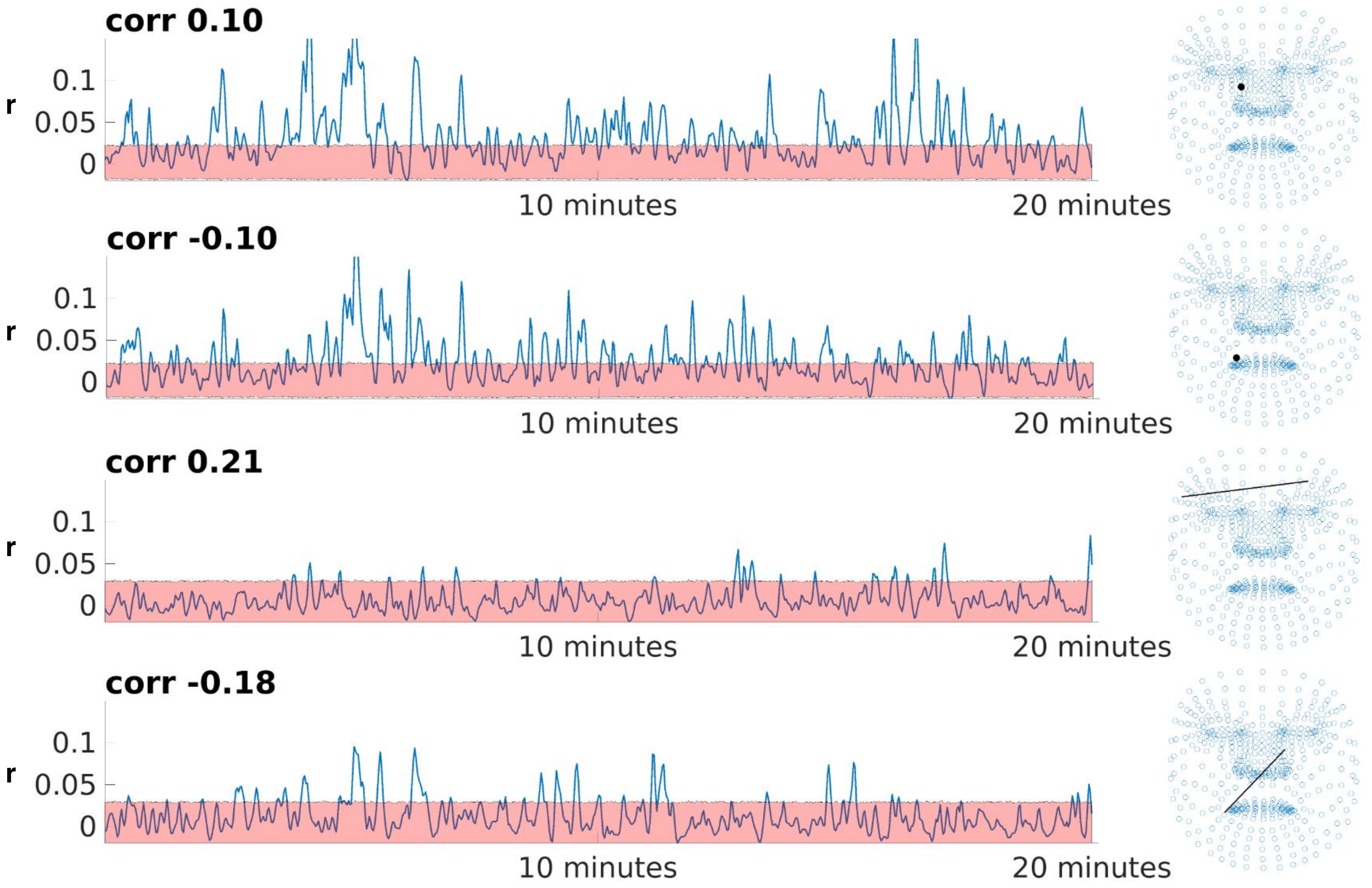

Fig. 3

CONCLUSIONS

In summary, performances of our model generally matches top DNN solutions. The t-ISC method is meaningful to extract both idiosyncratic and synchronized informations. The outcomes derived from the emotion classification and **inter-subject synchronicity** analyses demonstrate strong individual performance within the suggested framework, while also indicating potential for future integration to unlock additional application possibilities.

To conclude,

- Leave-1-subject-out reveals high variability between-subjects and proves that successful classification is driven by the presence of samples from the validated subject into the training-set
- Results suggest that facial expressions are highly consistent within-subject
- t-ISC offers a reliable approach which can be employed to infer synchronicity in emotional responses.

Contact: ✉ matteo.lionello@imtlucca.it github.com/mlionello/FACEXP