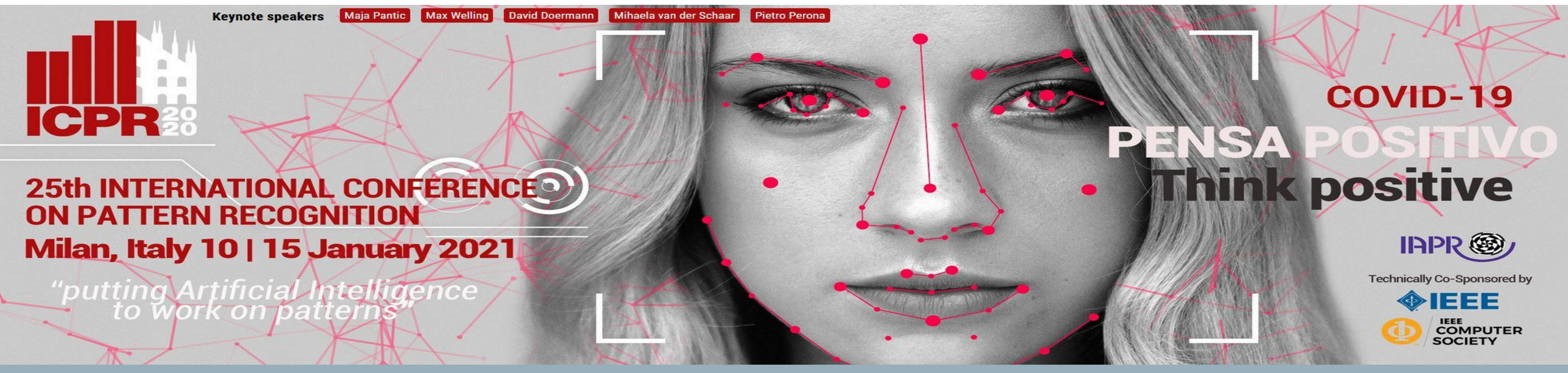
ulm university universität uulm



Using Meta Labels for the Training of Weighting Models in a Sample-Specific Late Fusion Classification Architecture

Peter Bellmann, Patrick Thiam, and Friedhelm Schwenker

ABSTRACT

Late fusion (LF) architectures are common approaches in classification tasks that consist of several feature subspaces, e.g. multi-modal sensor data, such as audio, video, and physiological signals. In the current study, we propose a trainable sample-specific LF architecture that combines classification and weighting models (CMs, WMs). The idea of our approach is to train each WM in combination with a set of CM-specific meta labels.

FORMALISATION

- $X \subset \mathbb{R}^d$, $d \in \mathbb{N}$: d-dimensional data set
- $m \in \mathbb{N}$: number of feature subsets
- CM_i: classification model that is trained on feature subset i
- WM_i: weighting model specific to classification model CM_i
- Each CM_i and WM_i is a **strong** model (*ensemble*)

PROPOSED APPROACH – Basic Idea

- Divide the training subsets $X_1, \dots, X_m$ into
  - $T_1, \dots, T_m$: Training Sets
  - $V_1, \dots, V_m$: Validation Sets
- The output of CM_i on V_i defines the labels for weighting model WM_i

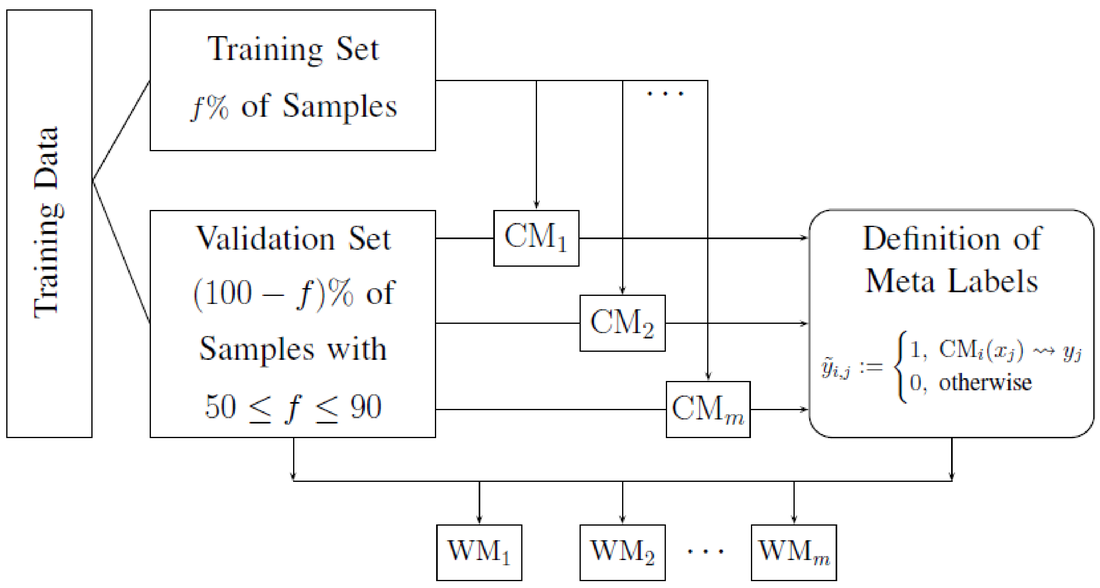
$$\tilde{y}_{i,j} := \begin{cases} 1, & \text{if } \mathrm{CM}_i(v_j) \rightsquigarrow y_j, \\ 0, & \text{otherwise.} \end{cases}$$

- The class-support vector for input $x \in \mathbb{R}^d$ is calculated as

$$\mu(x) = \sum_{i=1}^{m} s_i^{(1)}(x) \cdot \mathrm{CM}_i(x)$$

- $s_i^{(1)}$ denotes the decision confidence for classification model CM_i

PROPOSED APPROACH – System Diagram (Training)



RESULTS – Accuracy Performance (in %)

Data Set	Evaluation	Early Fusion	Late Mean	Our Method
BioVid	Leave-1-Out	81.90 ± 15.2	82.93 ± 16.0	$\mathbf{83.94 \pm 15.3}$
Mfeat	20-fold	96.02 ± 1.64	97.60 ± 1.47	$\mathbf{98.00 \pm 1.34}$
Arrhythmia	20-fold	74.62 ± 7.86	75.15 ± 10.6	$\mathbf{76.48 \pm 8.93}$
Fisher Iris	10-fold	94.39 ± 4.10	95.33 ± 5.49	$\mathbf{96.67 \pm 4.71}$

RESULTS – Operational Cost (Training & Testing Time in s)

Approach	BioVid	Mfeat	Arrhythmia	Fisher Iris
Late Mean	21.8 ± 0.3	$\mathbf{12.5 \pm 0.3}$	$\mathbf{2.53 \pm 0.2}$	$\mathbf{1.45 \pm 0.1}$
Our Method	$\mathbf{20.2 \pm 0.3}$	13.9 ± 0.3	4.29 ± 0.3	3.00 ± 0.2

CONCLUSION

- Proposed idea is a valid alternative for trainable LFs
- Proposed idea can be applied as a plain ensemble method
- In future, we aim to analyse the effectiveness of confidence:

$$\tilde{s}_i(x) = \begin{cases} +s_i^{(1)}(x), & \text{if } s_i^{(1)}(x) \geq \theta_2, \\ 0, & \text{if } s_i^{(1)}(x) \in (\theta_1, \theta_2), \\ -s_i^{(1)}(x), & \text{if } s_i^{(1)}(x) \leq \theta_1. \end{cases}$$

Contact

Peter Bellmann (peter.bellmann@uni-ulm.de)

Institute of Neural Information Processing

Ulm University, Germany

Reference for the BioVid Heat Pain Database

S. Walter, S. Gruss, H. Ehleiter, J. Tan, H. C. Traue, S. C. Crawcour, P. Werner, A. Al-Hamadi, and A. O. Andrade, „The BioVid Heat Pain Database Data for the Advancement and Systematic Validation of an Automated Pain Recognition System“, in CYBCONF. IEEE, 2013, pp. 128-131