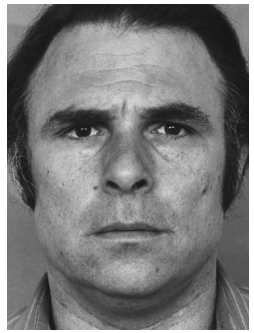
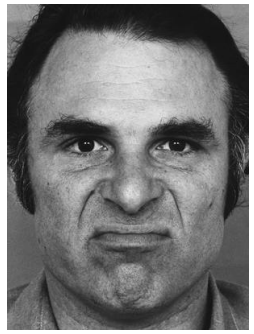
Research Domain

- Detection of Action Units, distinct facial movements of the face
- Used for objective, standardized description of facial expressions by psychologists
- Combinations of Action Units can describe the expression of emotions, pain, or stress in a face
- Application examples:
  - Monitoring drivers in driving assisting systems (stress, emotions)
  - Automatic monitoring of pain states in the medical field

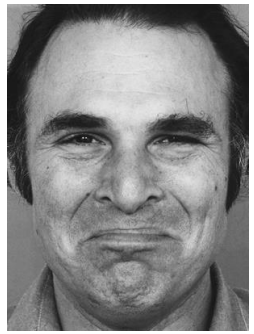


AU 04: Brow lowerer



AU 10: Upper lip raiser

AU 17: Chin raiser



AU 06: Cheek raiser

AU 07: Lid tightener

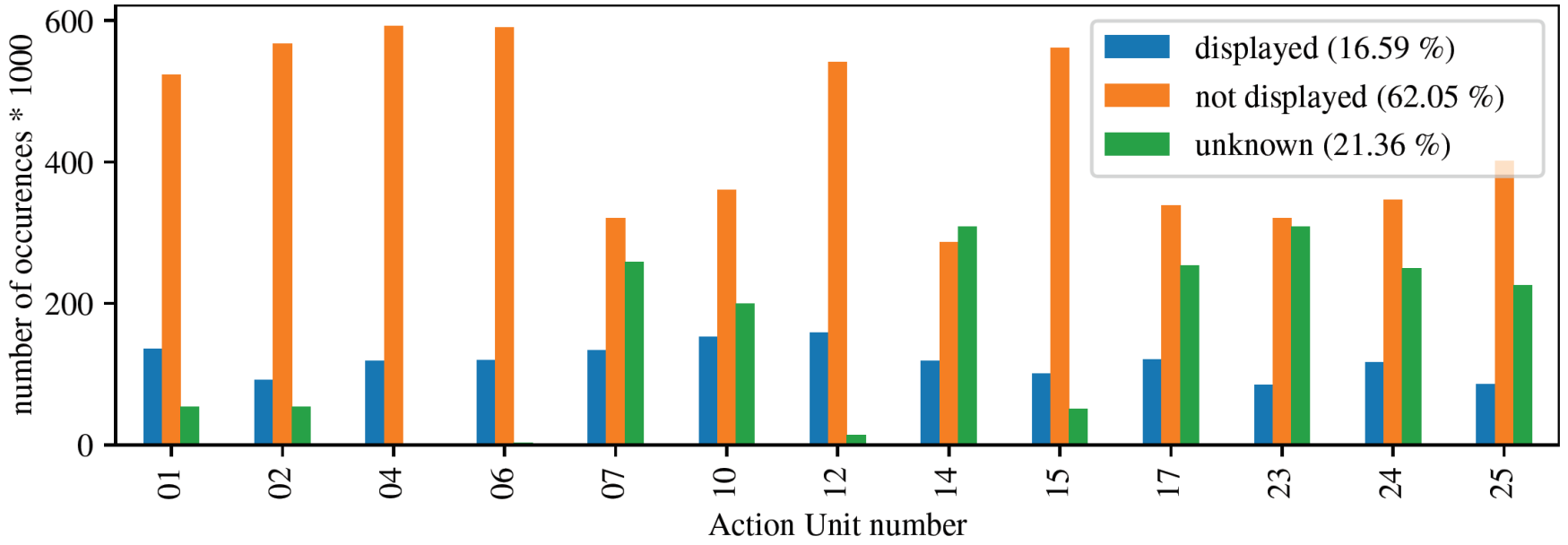
AU 12: Lip corner puller

AU 15: Lip corner depressor

AU 17: Chin raiser

Ekman, Paul, Wallace V. Friesen, and Joseph C. Hager. "Facial Action Coding System: The Manual on CD ROM." *A Human Face, Salt Lake City* (2002): 77-254.

Dataset	Setting	Expression	Format	Subjects	Images	Images / Person
Actor Study	lab	posed	video	21	100,232	4,773
Aff-Wild2	in-the-wild	natural	video	49	196,672	4,014
BP4D	lab	natural	video	41	366,954	8,950
CK+	lab	posed	video	123	10,733	87
EmotioNet manual	in-the-wild	natural	images	24,597	24,597	1
UNBC	lab	natural	video	25	48,397	1,936



Algorithm

- Action Unit labeling is very expensive and time-consuming, dropping images with missing labels is not feasible
- Combining databases allows to transfer deep learning's success factor of training with large databases to the domain of Action Units
- Data is missing due to varying labeling completeness and facial occlusions
- Data comes from different distributions (datasets), so pre-labeling missing values by inferring labels is not feasible
- Therefore, we aim to tweak our training algorithm to be able to handle missing labels while training instead of inferring values
- **When an unknown label occurs in the ground truth label vector, we delete this vector row in the predicted label vector and the ground truth label vector. This process is computed for each class separately in our loss function L:**

$$F1 = \frac{2 \cdot precision_{|\{0,1\}} \cdot recall_{|\{0,1\}}}{precision_{|\{0,1\}} + recall_{|\{0,1\}}} \qquad F1_{macro} = \frac{1}{N}\left(\sum_{i}^{N} F1_i\right) \qquad L = 1 - F1_{macro}$$

Results

- In comparison to two major recent challenges, our F1 macro score is competitively high:
- Affective Behavior Analysis in-the-wild Challenge at “Face and Gesture 2020“ conference:
  Top contestant reached F1 macro score of 0.31
- EmotioNet Challenge at “Computer Vision and Pattern Recognition 2020” conference:
  Top three participants reached F1 macro scores of 0.44 to 0.55

AUs	Train	Test	F1
AU01	107,007	20,944	0.69
AU02	78,461	5,528	0.42
AU04	94,986	14,150	0.59
AU06	95,539	21,212	0.63
AU07	107,283	20,124	0.73
AU10	121,313	22,056	0.80
AU12	129,915	22,550	0.77
AU14	94,422	15,010	0.62
AU15	83,580	6,099	0.33
AU17	98,892	11,388	0.53
AU23	65,769	5,970	0.35
AU24	94,006	7,781	0.54
AU25	71,048	5,727	0.59
F1 macro	-	-	0.58
Weighted F1 macro	-	✓	0.65
Weighted F1 macro	✓	-	0.61