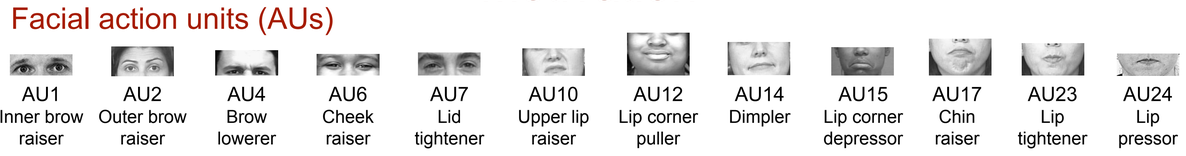
Cross-domain AU Detection: Domains, Learning Approaches, and Measures

Itır Önal Ertuğrul[1], Jeffrey F. Cohn[2], László A. Jeni[1], Zheng Zhang[3], Lijun Yin[3], Qiang Ji[4]

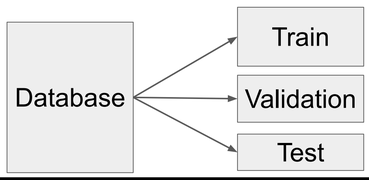
[1]Carnegie Mellon University, [2]University of Pittsburgh, [3]State University of New York at Binghamton, [4]Rensselaer Polytechnic Institute

Motivation

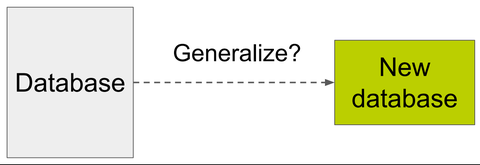
Facial action units (AUs)



Within-domain



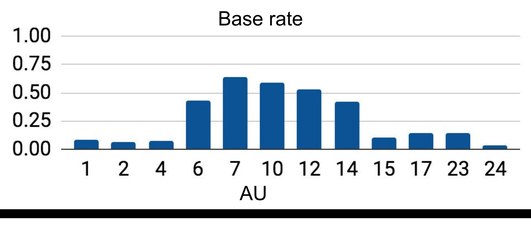
Cross-domain



Databases

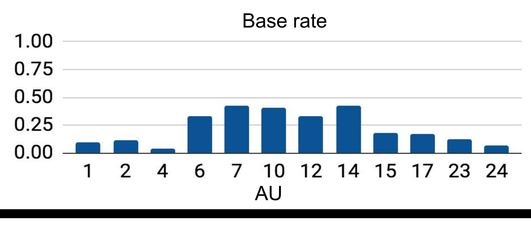
Expanded BP4D+ (EB+) [1]

- 200 subjects (375K frames)
- Emotion inductions
- Nearly frontal
- 1040 x 1392 pixels



Sayette GFT

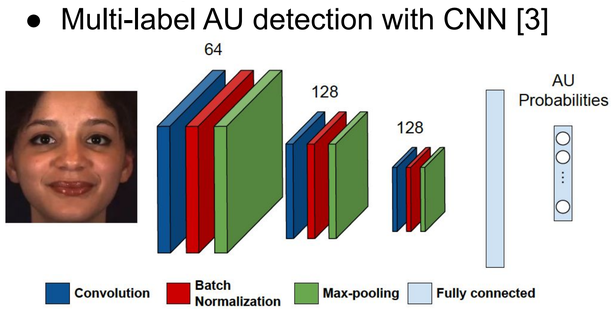
- 150 subjects (517K frames)
- Social interaction
- Out of plane head motion
- 720 x 480 pixels



Methods

Deep approach

- Face registration with ZFace [2]
- Person-specific normalization
- Multi-label AU detection with CNN [3]



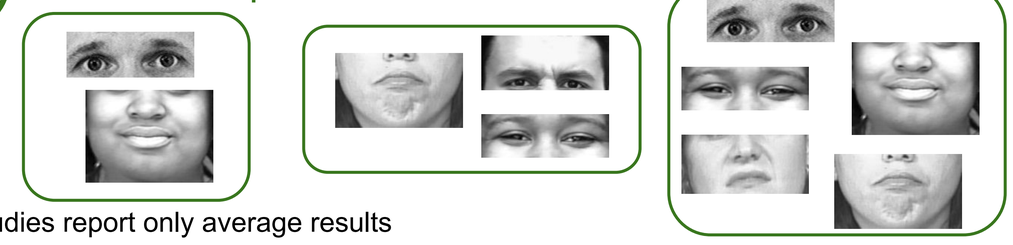
Shallow approach

OpenFace [4]: an open source facial behavior analysis toolkit

- Face alignment
- Appearance (HoG) and geometric (landmarks) features
- Person-specific normalization
- Train SVM

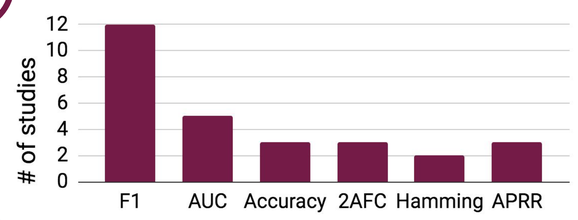
Problems in Comparing Generalizability

P1 Lack of AU specific Results

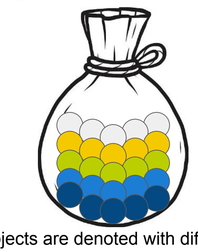


6 studies report only average results

P2 Inconsistent Measures

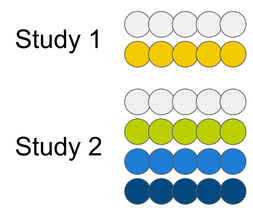


P3 Subsampling



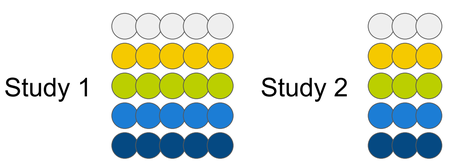
Subjects are denoted with different color

Case1: Subject subsampling



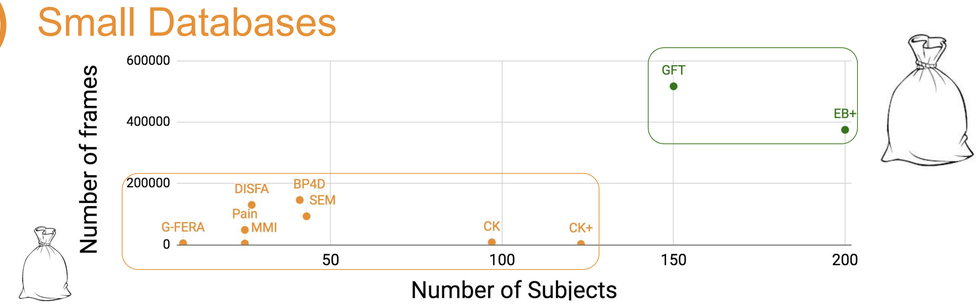
MMI dataset: 10, 11, 15, 70 subjects

Case2: Image subsampling



DISFA dataset: 2400, 4000, 4845 frames
CK+ dataset: 327, 582, 593 sequences

P4 Small Databases



Solutions

- S1 Report AU-specific results of 12 AUs
- S2 Report multiple measures: S score, AUC, F1 and NA
- S3 Use all available frames (375K in EB+, 517K in GFT)
- S4 Use two large databases EB+ and GFT

Results

Within-domain comparison (EB+ and GFT)

Within-domain: EB+

-	S score	AUC	F1 (PA)	NA
AU1	0.79	0.81	0.47	0.94
AU2	0.86	0.82	0.44	0.96
AU4	0.87	0.88	0.53	0.97
AU6	0.68	0.93	0.82	0.86
AU7	0.65	0.89	0.86	0.75
AU10	0.71	0.93	0.88	0.82
AU12	0.74	0.95	0.88	0.86
AU14	0.57	0.85	0.75	0.81
AU15	0.78	0.81	0.41	0.94
AU17	0.64	0.79	0.34	0.90
AU23	0.72	0.85	0.57	0.92
AU24	0.94	0.90	0.24	0.99
Avg 12	**0.75**	**0.87**	**0.60**	**0.89**
Avg 11	**0.73**	**0.86**	**0.63**	**0.88**

Cross-domain: GFT →EB+

-	S score	AUC	F1 (PA)	NA
AU1	0.74	0.72	0.31	0.93
AU2	0.84	0.67	0.22	0.96
AU4	0.85	0.75	0.20	0.96
AU6	0.37	0.75	0.58	0.75
AU7	0.27	0.69	0.68	0.58
AU10	0.47	0.81	0.77	0.69
AU12	0.53	0.85	0.76	0.77
AU14	0.23	0.71	0.63	0.60
AU15	0.65	0.66	0.27	0.90
AU17	0.38	0.65	0.30	0.80
AU23	0.25	0.66	0.32	0.74
AU24	0.73	0.72	0.14	0.93
Avg 12	**0.53**	**0.72**	**0.43**	**0.80**
Avg 11	**0.51**	**0.72**	**0.46**	**0.79**

Within-domain: GFT

-	S score	AUC	F1 (PA)	NA
AU1	0.83	0.83	0.44	0.95
AU2	0.77	0.81	0.45	0.94
AU4	0.93	0.75	0.20	0.98
AU6	0.68	0.91	0.75	0.88
AU7	0.53	0.84	0.72	0.79
AU10	0.62	0.89	0.77	0.84
AU12	0.74	0.93	0.80	0.91
AU14	0.25	0.66	0.50	0.69
AU15	0.58	0.70	0.34	0.87
AU17	0.64	0.67	0.17	0.90
AU23	0.74	0.69	0.17	0.93
AU24	0.85	0.81	0.13	0.96
Avg 12	**0.68**	**0.79**	**0.45**	**0.89**
Avg 11	**0.66**	**0.79**	**0.48**	**0.88**

Cross-domain: EB+ →GFT

-	S score	AUC	F1 (PA)	NA
AU1	0.74	0.73	0.26	0.93
AU2	0.60	0.72	0.34	0.88
AU4	0.82	0.71	0.18	0.95
AU6	0.56	0.85	0.69	0.83
AU7	0.25	0.79	0.67	0.57
AU10	0.49	0.85	0.73	0.76
AU12	0.54	0.87	0.70	0.81
AU14	0.08	0.65	0.62	0.42
AU15	0.31	0.62	0.32	0.77
AU17	0.22	0.65	0.33	0.72
AU23	0.67	0.66	0.25	0.91
AU24	0.53	0.73	0.23	0.86
Avg 12	**0.48**	**0.74**	**0.44**	**0.79**
Avg 11	**0.48**	**0.74**	**0.46**	**0.78**

Cross-domain: OpenFace →GFT

-	S score	AUC	F1 (PA)	NA
AU1	0.66	0.70	0.37	0.90
AU2	0.58	0.69	0.39	0.87
AU4	0.64	0.57	0.10	0.90
AU6	0.49	0.76	0.68	0.79
AU7	0.31	0.64	0.59	0.70
AU10	0.51	0.77	0.74	0.77
AU12	0.47	0.78	0.69	0.77
AU14	0.04	0.56	0.61	0.38
AU15	0.41	0.58	0.32	0.81
AU17	0.41	0.61	0.35	0.81
AU23	0.30	0.52	0.20	0.78
AU24	-	-	-	-
Avg 12	-	-	-	-
Avg 11	**0.44**	**0.65**	**0.46**	**0.77**

Within-domain cross-domain comparison (EB+)

Within-domain cross-domain comparison (GFT)

Deep shallow comparison

Analysis

Significance

	Within EB+ > Within GFT		Within GFT > Cross on GFT		Within EB+ > Cross on EB+		Deep > Shallow	
AU	S	AUC	S	AUC	S	AUC	S	AUC
AU1	n.s.	n.s.	***	***	***	***	***	n.s.
AU2	***	n.s.	***	***	*	***	n.s.	*
AU4	n.s.	***	***	n.s.	n.s	***	***	***
AU6	n.s.	n.s.	***	***	***	***	**	***
AU7	***	**	***	***	***	***	n.s.	***
AU10	***	*	***	***	***	***	n.s.	***
AU12	n.s.	n.s.	***	***	***	***	n.s.	***
AU14	***	***	***	***	***	***	*	***
AU15	***	***	***	***	***	***	***	***
AU17	***	***	***	***	***	***	***	**
AU23	***	***	***	***	***	***	***	***
AU24	***	n.s.	***	***	***	***	-	-

- Within-domain results are better for EB+ than for GFT.
- For AU 6 and AU 12, within-domain results of both databases are similarly good.

- Within-domain results are better than cross-domain results for both domains.
- The model trained on EB+ generalizes well to GFT database. But, reverse is not true.

- Deep model gives slightly better S score and similar F1 on average, while average AUC of deep approach is much higher than Openface.
  - **AUC:** Deep approach ranks instances with AUs present or absent much better
  - **F1:** Both deep and shallow approaches perform similar on positive instances
  - **S score:** When the effect of chance is discarded, deep approach performs slightly better.
- Reporting multiple measures is important.

Take home message

- Use all available frames of the datasets and report AU-specific results with multiple measures to ensure comparability for cross-domain generalizability.
- Exercise caution in applying available systems to new domains.
- Consider domain adaptation approaches for AU detection.

References

[1] For EB+ please visit: http://www.cs.binghamton.edu/~lijun/Research/3DFE/3DFE_Analysis.html

[2] L. A. Jeni, J. F. Cohn, & T. Kanade, Dense 3D face alignment from 2D videos in real-time. FG, 2015

[3] For AFAR please visit: http://www.jeffcohn.net/Resources/AFAR

[4] T. Baltrušaitis, A. Zadeh, Y. C. Lim, and L. P. Morency, OpenFace 2.0: Facial Behavior Analysis Toolkit, FG, 2018

This research was supported in part by NIH awards NS100549 and MH096951 and NSF awards CNS-1629716, CNS-1629898, and CNS-1629856.