Selective Transfer Machine for Personalized Facial Action Unit Detection

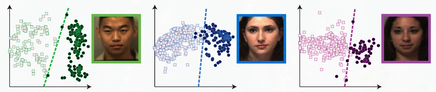
[1]Carnegie Mellon University [2]University of Pittsburgh

Wen-Sheng Chu[1], Fernando De la Torre[1] and Jeffery F. Cohn[1,2]

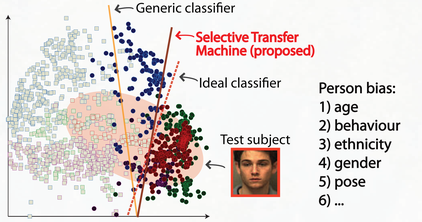
CVPR 2013 Portland, Oregon June 23-28

Problem

- Individual differences impair generalizability
  - △ Train and test on the same subject

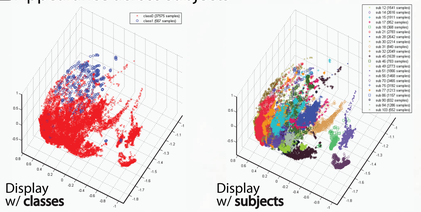


  - △ Test on a new subject

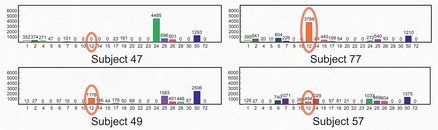


Person bias:
1) age
2) behaviour
3) ethnicity
4) gender
5) pose
6) ...

- Challenges
  - △ Appearance across subjects



  - △ Occurrence of AUs



Selective Transfer Machine (STM)

- Objective

$$(\mathbf{w}, \mathbf{s}) = \min_{\mathbf{w},\mathbf{s}} \underbrace{R_{\mathbf{w}}(\mathcal{D}^{\mathrm{tr}}, \mathbf{s})}_{\text{margin of penalized SVM}} + \lambda \overbrace{\Omega_{\mathbf{s}}(\mathbf{X}^{\mathrm{tr}}, \mathbf{X}^{\mathrm{te}})}^{\text{training/testing distribution mismatch}}$$

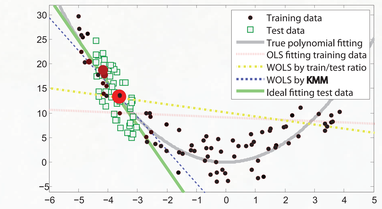
  - △ Penalized SVM:

$$R^{\mathrm{lin}}_{\mathbf{w}}(\mathcal{D}^{\mathrm{tr}}, \mathbf{s}) = \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{n_{\mathrm{tr}}} s_i L^p(y_i, \mathbf{w}^\top \mathbf{x}_i)$$

$$R^{\mathrm{nonlin}}_{\mathbf{w}}(\mathcal{D}^{\mathrm{tr}}, \mathbf{s}) = \frac{1}{2}\boldsymbol{\beta}^\top \mathbf{K}\boldsymbol{\beta} + C\sum_{i=1}^{n_{\mathrm{tr}}} s_i L_p(y_i, \mathbf{K}_i^\top \boldsymbol{\beta})$$

  - △ Domain mismatch (KMM [1]):

$$\Omega_{\mathbf{s}}(\mathbf{X}, \mathbf{X}) = \left\| \frac{1}{n}\sum_{i=1}^{n} s_i \varphi(\mathbf{x}_i) - \frac{1}{n}\sum_{j=1}^{n} \varphi(\mathbf{x}_j) \right\|^2_{\mathcal{H}}$$



- Comparison with existing methods

Methods	Instance Reweighting	Weight Refinement	Convergence
SVM			✓
T-SVM [2]			✓
DA-SVM [3]	✓	✓	
KMM [1]	✓		✓
STM	✓	✓	✓

[1] A. Gretton, A. Smola, J. Huang, M. Schmittfull, K. Borgwardt, and B. Schölkopf. Covariate shift by kernel mean matching. Dataset shift in machine learning, 2009.
[2] T. Joachims. Transductive inference for text classification using support vector machines. In ICML, 1999.
[3] L. Bruzzone and M. Marconcini. Domain adaptation problems: A DASVM classification technique and a circular validation strategy. PAMI, 32(5):770–787, 2010.

STM Optimization

- Alternative Convex Search [4]
  - ▲ Minimize over $\mathbf{w}$

$$\nabla_{\mathrm{lin}} = \mathbf{w} + 2C\sum_{i\in sv} s_i(\mathbf{w}^\top \mathbf{x}_i - y_i)\mathbf{x}_i$$

$$H_{\mathrm{lin}} = \mathbf{I}_d + 2C\sum_{i\in sv} s_i \mathbf{x}_i \mathbf{x}_i^\top$$

$$\nabla_{\mathrm{nonlin}} = \mathbf{K}\boldsymbol{\beta} + 2C\mathbf{K}\mathbf{S}\mathbf{I}^0(\mathbf{K}\boldsymbol{\beta} - \mathbf{y})$$

$$H_{\mathrm{nonlin}} = \mathbf{K} + 2C\mathbf{K}\mathbf{S}\mathbf{I}^0\mathbf{K}$$

  - ▲ Minimize over $\mathbf{s}$

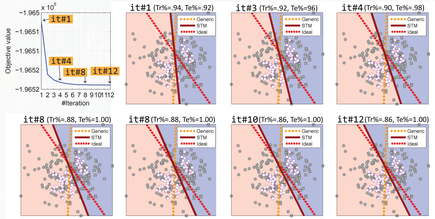
$$\min_{\mathbf{s}} \ \frac{1}{2}\mathbf{s}^\top \mathbf{K}\mathbf{s} + (\frac{C}{\lambda}\boldsymbol{\ell}_p - \boldsymbol{\kappa})^\top \mathbf{s}$$

$$\text{s.t.} \ \ 0 \le s_i \le B, n_{tr}(1-\epsilon) \le \sum_{i=1}^{n} s_i \le n_{tr}(1+\epsilon).$$

Algorithm 2: Alternate Convex Search Algorithm

1 **Step 1:** Choose a starting point $z \leftarrow (\mathbf{w}_0, \mathbf{s}_0) \in H$;
2 Set $t \leftarrow 0$;
3 **while** *not converged* **do**
4 **Step 2:** Solve for fixed $\mathbf{s}_t$ the convex optimization problem: $\mathbf{w}_{t+1} \leftarrow \min_{\mathbf{w}}\{f(\mathbf{w}, \mathbf{s}_t), \mathbf{w} \in H_{\mathbf{s}_t}\}$;
5 **Step 3:** Solve for fixed $\mathbf{w}_t$ the convex optimization problem: $\mathbf{s}_{t+1} \leftarrow \min_{\mathbf{s}}\{f(\mathbf{w}_{t+1}, \mathbf{s}), \mathbf{s} \in H_{\mathbf{w}_{t+1}}\}$;
6 **Step 4:** Set $z_{t+1} \leftarrow (\mathbf{w}_{t+1}, \mathbf{s}_{t+1})$;
7 Set $t \leftarrow t+1$;
8 **end**

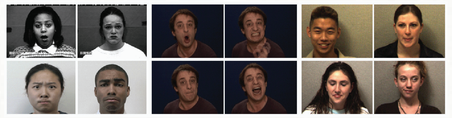
- Synthetic example



[4] J. Gorski, F. Pfeuffer, and K. Klamroth. Biconvex sets and optimization with biconvex functions: a survey and extensions. Mathematical Methods of Operations Research, 66(3):373–407, 2007.

Experiments

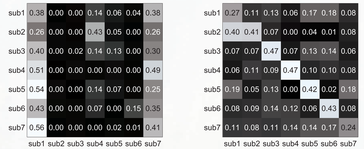
- Datasets



- Comparison with Person-Specific (PS) classifiers
  - △ **PS1:** train/test are separate data of the same subject
  - △ **PS2:** training subjects include the test subject (i.e., the same protocol as the quasi-PS [6])

	AUC			$F1$ Score		
AU	PS$_1$-SVM	PS$_2$-SVM	STM	PS$_1$-SVM	PS$_2$-SVM	STM
1	48.0	72.4	**79.2**	45.0	54.8	**61.9**
2	46.5	71.1	**80.2**	45.9	55.7	**64.3**
4	62.6	61.9	**66.5**	46.6	40.7	**60.4**
6	70.3	80.0	**86.4**	60.2	69.7	**78.5**
7	47.5	54.3	**72.4**	49.4	55.3	**58.4**
12	65.7	74.0	**72.3**	69.5	70.4	**72.6**
15	41.4	64.0	**70.5**	44.5	49.0	**56.0**
17	32.6	**70.3**	61.7	25.0	**40.3**	36.3
Avg	51.8	68.5	**73.6**	48.3	54.5	**61.0**

- Selection ability of STM
  - △ Experiment **PS2** protocol on GEMEP-FERA [6]



[5] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews. The extended Cohn-Kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression. In CVPR Workshops, 2010.
[6] M. F. Valstar, M. Mehu, B. Jiang, M. Pantic, and K. Scherer. Meta-analysis of the first facial expression recognition challenge. IEEE Trans. on Systems, Man, and Cybernetics, Part B, 42(4):966–979, 2012.
[7] M. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, and J. Movellan. Automatic recognition of facial actions in spontaneous expressions. Journal of Multimedia, 1(6):22–35, 2006.

Comparison on Benchmarks

- Settings
  - △ SIFT descriptors on predetermined facial landmarks
  - △ Leave-one-subject-out protocol
- CK+ [5]
  - △ 593 posed videos from 123 subjects
  - △ ~20 frames/video

	AUC					$F1$ Score				
AU	SVM	KMM	T-SVM	DA-SVM	STM	SVM	KMM	T-SVM	DA-SVM	STM
1	79.8	68.9	69.9	72.6	**88.9**	61.1	44.9	56.8	57.7	**62.2**
2	**90.8**	73.5	69.3	71.0	87.5	73.5	50.8	59.8	64.3	**76.2**
4	74.8	62.2	63.4	69.9	**81.1**	62.7	52.3	51.9	57.7	**69.1**
6	89.7	87.7	60.5	**94.7**	94.0	75.5	70.1	47.8	68.2	**79.6**
7	82.1	68.2	55.7	61.4	**91.6**	59.6	47.0	43.8	53.1	**79.1**
12	88.1	89.5	76.0	**95.5**	92.8	76.7	74.5	59.6	59.0	**77.2**
15	93.5	66.8	49.9	94.1	**98.2**	75.3	44.4	40.4	76.9	**84.8**
17	90.3	66.6	73.1	94.7	**96.0**	76.0	53.2	61.7	81.4	**84.3**
Avg	86.1	72.9	64.7	81.7	**91.3**	70.0	54.7	52.7	64.8	**76.6**

- GEMEP-FERA [6]
  - △ 87 videos from 7 trained actors
  - △ 40~110 frames/video

	AUC					$F1$ Score				
AU	SVM	KMM	T-SVM	DA-SVM	STM	SVM	KMM	T-SVM	DA-SVM	STM
1	71.5	43.3	72.2	83.3	**84.3**	56.5	48.5	60.3	59.1	**68.1**
2	73.9	51.0	74.3	**76.8**	73.3	56.9	50.2	58.5	57.1	**65.5**
4	58.5	53.5	42.8	**66.6**	60.0	43.5	39.8	36.9	**46.3**	43.3
6	80.4	60.2	81.1	**91.1**	87.7	63.7	58.7	63.8	**72.7**	71.6
7	66.9	59.4	70.8	**76.9**	75.4	63.1	63.5	63.7	**68.3**	66.2
12	77.7	58.8	74.8	74.5	**84.7**	79.1	68.4	77.6	75.5	**82.1**
15	55.5	58.7	67.2	67.5	**67.8**	33.4	35.2	35.2	**41.3**	39.3
17	59.8	51.8	63.8	**66.5**	63.3	32.0	27.8	36.2	**42.0**	35.9
Avg	68.0	54.6	68.4	**75.4**	74.5	53.5	49.0	54.0	57.8	**59.0**

- RU-FACS [7]
  - △ 29 interview videos from 29 subjects
  - △ ~6k frames/video

	AUC					$F1$ Score				
AU	SVM	KMM	T-SVM	DA-SVM	STM	SVM	KMM	T-SVM	DA-SVM	STM
1	72.0	74.0	72.0	77.0	**83.9**	40.8	37.7	37.4	35.5	**55.3**
2	66.6	58.6	71.1	76.5	**82.4**	35.7	32.2	36.2	34.1	**52.6**
4	74.8	62.2	50.0	76.4	**82.4**	25.2	14.5	11.2	**35.3**	30.4
6	89.1	88.8	61.6	60.3	**93.1**	58.3	39.2	33.1	42.9	**72.4**
7	86.7	87.0	86.7	84.4	**92.3**	61.9	63.0	62.6	71.4	**72.3**
12	71.8	67.8	74.4	70.4	**87.4**	31.3	25.8	25.8	40.9	**51.0**
15	72.5	68.8	73.5	58.1	**86.1**	32.3	29.5	32.3	34.9	**45.4**
17	78.5	76.7	79.5	75.7	**89.6**	39.5	35.6	44.0	46.5	**55.3**
Avg	76.5	72.3	71.1	72.3	**85.3**	40.6	37.3	40.6	42.7	**54.3**