City College Visual Computing Laboratory

Action Unit Detection with Region Adaptation, Multi-labeling Learning and Optimal Temporal Fusing

NSF Award #EFRI-1137172

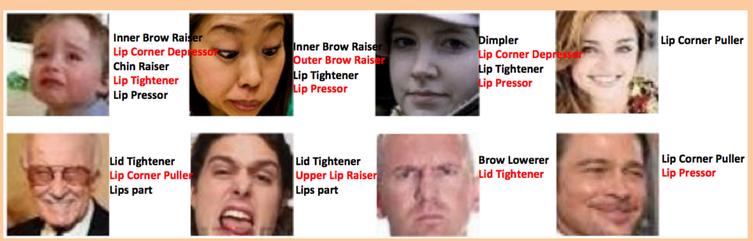
Wei Li[1,3], Farnaz Abtahi[2], Zhigang Zhu[1,2],
[1]Grove School of Engineering, The CUNY City College [2]Department of Computer Science, The CUNY Graduate Center
[3]Amazon AWS Rekognition Research

Overview

❖1) A set of adaptive ROI cropping nets (ROI Nets) are designed to learn regional features separately.
❖2) Multi-label and single AU based methods are compared. With additional AU correlations and richer global features, the multi-label learning approach shows slightly better performance.

❖3) An LSTM-based temporal fusion recurrent net (LSTM Net) is proposed to fuse static CNN features, which makes the AU predictions more accurate than with static images only. .

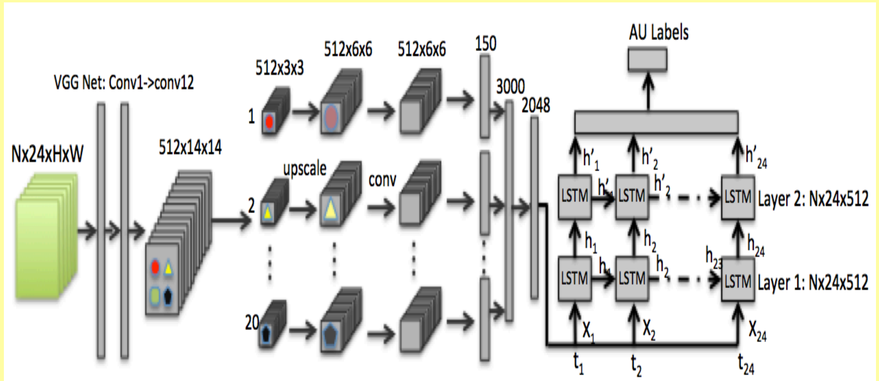
AU Detection



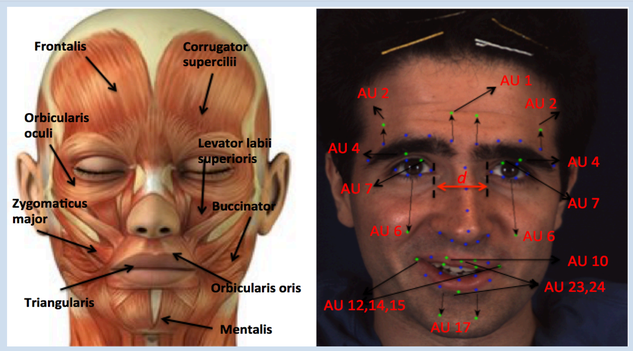
With deep pretrained models and a ``smarter'' way to focus on interest regions, the proposed approach shows its power in AU detection on multiple datasets. Our approach also shows the potential to deal with "wild" image AU detection in real time, which is our ongoing work.

Framework

Network for AU detection



Generation of Region Of Interest

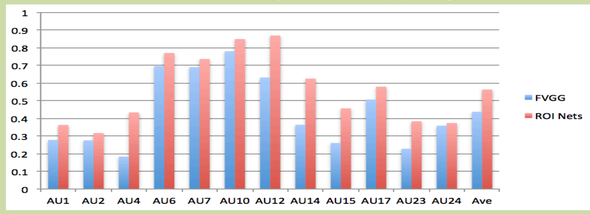


AU index	Au Name	AU Center
1	Inner Brow Raiser	1/2 scale above inner brow
2	Outer Brow Raiser	1/3 scale above outer brow
4	Brow Lowerer	1/3 scale below brow center
6	Cheek Raiser	1 scale below eye bottom
7	Lid Tightener	Eye center
10	Upper Lip Raiser	Upper lip center
12	Lip Corner Puller	Lip corner
14	Dimpler	Lip corner
15	Lip Corner Depressor	Lip corner
17	Chin Raiser	1/2 scale below lip
23	Lip Tightener	Lip center
24	Lip Pressor	Lip center

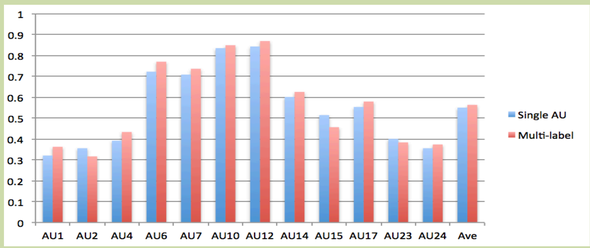
Results&Evaluation

AU Detection Comparison Experiment

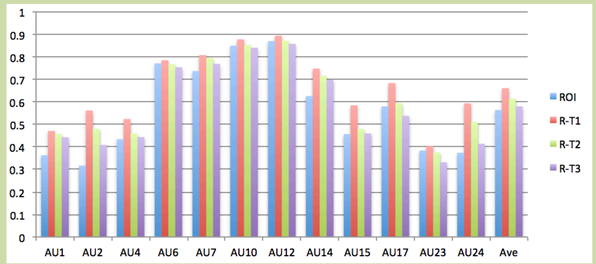
FVGG v.s. Region Cropping



Multi-Label v.s. Single Label



ROI Cropping v.s. ROI +T2 (LSTM)



Evaluation on BP4D AU Dataset

AU	LSVM	JPML[29]	DRML[30]	CPM[28]	CNN+LSTM[3]	FVGG	ROI	R-T1	R-T2	FERA[12]
1	23.2	32.6	36.4	43.4	31.4	27.8	36.2	**47.1**	45.8	*28*
2	22.8	25.6	41.8	40.7	31.1	27.6	31.6	**56.2**	48.0	*28*
4	23.1	37.4	43.0	43.4	**71.4**	18.3	43.4	52.4	45.9	*34*
6	27.2	42.3	55.0	59.2	63.3	69.7	77.1	**78.5**	76.7	*70*
7	47.1	50.5	67.0	61.3	77.1	69.1	73.7	**80.8**	79.6	*78*
10	77.2	72.2	66.3	62.1	45.0	78.1	85.0	**87.8**	85.3	*81*
12	63.7	74.1	65.8	68.5	82.6	63.2	87.0	**89.4**	87.2	*78*
14	64.3	65.7	54.1	52.5	72.9	36.4	62.6	**74.8**	71.6	*75*
15	18.4	38.1	36.7	34.0	33.2	26.1	45.7	**58.5**	48.0	*20*
17	33.0	40.0	48.0	54.3	53.9	50.7	58.0	**68.4**	59.5	*36*
23	19.4	30.4	31.7	39.5	38.6	22.8	38.3	**40.4**	37.5	*41*
24	20.7	42.3	30.0	37.8	37.0	35.9	37.4	**59.4**	51.1	-
Avg	35.3	45.9	48.3	50.0	53.2	43.8	56.4	**66.1**	61.4	*51.7*

Code Available at Github:
https://github.com/wiibrew/EAC-Net
Author contact: buaawelldon@gmail.com