Pose-based Body Language Recognition for Emotion and Psychiatric Symptom Interpretation

Zhengyuan Yang, Amanda Kay, Yuncheng Li, Wendi Cross, Jiebo Luo

Emotions from Body Language

Emotions from the body language.
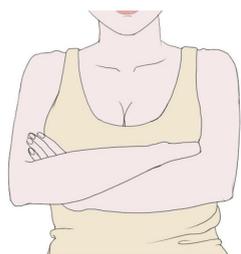
defensive, disagree   honesty, sincerity   ashamed, upset

Challenges

Body Language   Open palm facing upward

Emotion   Joy   1

- Body language recognition with limited samples
- Interpretable emotion and psychiatric symptom inference

Contributions

A framework to infer emotions from recognized body languages

- Interpretable and transferrable framework
- A URMC dataset with body language, emotion, and psychiatric symptoms annotations

Framework
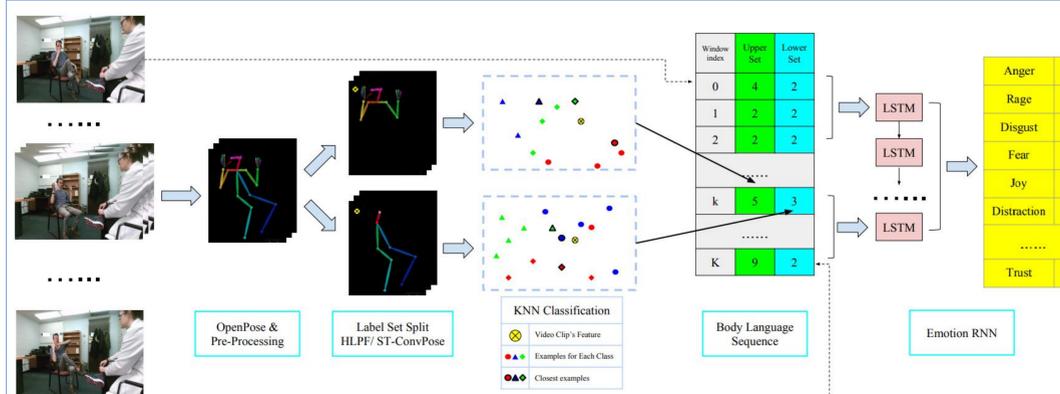
- Pose-based body language recognition: transferrable and data efficient
- KNN based classification: example based and interpretable
- Emotion RNN: emotion and psychiatric symptom interpretation based on the body language sequence

Pose-based Body Language Recognition

(a). Pose Frame Examples   (b). ST-ConvPose Framework
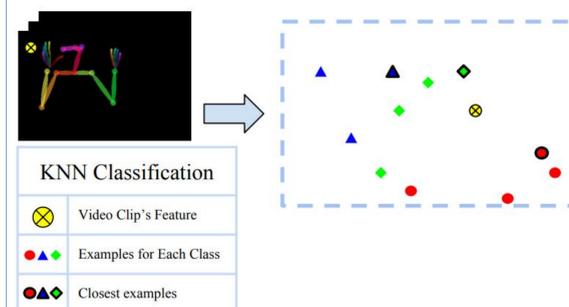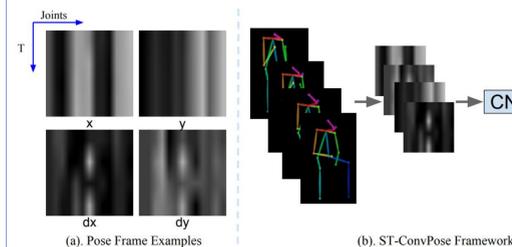
- Pose-based recognition: Requires less data and has better transferability
- Pose representation: CNN-based pose image pre-trained on action datasets
- KNN-based classification: only requires few example annotations. Interpretable and easily transferable to new body languages

KNN Classification
⊗ Video Clip's Feature
● ▲ ◆ Examples for Each Class
● ▲ ◆ Closest examples

URMC Dataset

a) Hands in air   b) Arm crosses   c) Biting nails   d) Touching nose

e) Locking ankles   f) Bouncing legs   g) Sitting with legs apart   h) Sitting with legs crossed

i) Examples of two body language in a frame   Upper Body Set   Lower Body Set
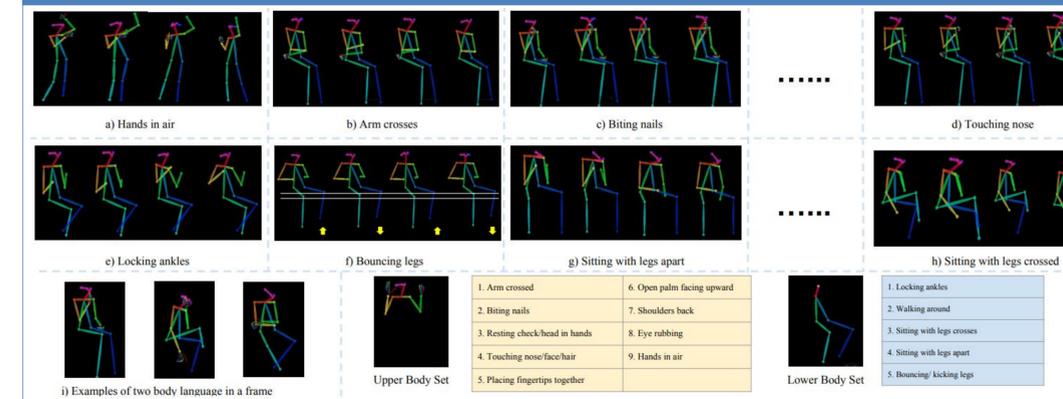
- 144 30-second video clips cropped from 12 20-minute videos
- 9 upper body body languages, 5 lower body body languages selected from 32 body languages in the initial annotations
- 24+1 emotion labels, and 24+1 symptom labels

Quantitative Results

Results on multi-label body language recognition

Lower Body Set	Acc.	Prec.	Recall	F1	Interpretability	transferability	Required Data Size
Two Stream	0.445	**0.581**	0.497	0.526			Large
$NTraj^+$+SVM	0.327	0.336	**0.690**	0.424			Medium
$NTraj^+$+KNN	0.483	0.516	0.616	0.538	✓		Small
ST-ConvPose+Dense	0.384	0.397	0.606	0.460		✓	Medium
ST-ConvPose+KNN	**0.488**	0.520	0.658	**0.554**	✓	✓	Small
Upper Body Set	Acc.	Prec.	Recall	F1	Interpretability	transferability	Required Data Size
Two Stream	0.341	0.472	0.567	0.492			Large
$NTraj^+$+SVM	0.346	0.388	**0.766**	0.485			Medium
$NTraj^+$+KNN	0.398	0.504	0.578	0.502	✓		Small
ST-ConvPose+Dense	0.374	**0.522**	0.486	0.473		✓	Medium
ST-ConvPose+KNN	**0.400**	0.497	0.641	**0.519**	✓	✓	Small

Results on emotion interpretation

LSTM+ST-ConvPose	Acc.	Prec.	Recall	F1
$L=1, S=1$	0.468	0.644	0.630	0.637
$L=7, S=3$	**0.564**	0.775	**0.674**	**0.721**
$L=48$	0.145	0.162	0.587	0.254
Other Methods	Acc.	Prec.	Recall	F1
LSTM+$NTraj^+$	0.510	**0.839**	0.565	0.675
Conv 1D+$NTraj^+$	0.490	0.788	0.565	0.658
Conv 1D+ST-ConvPose	0.556	0.789	0.652	0.714