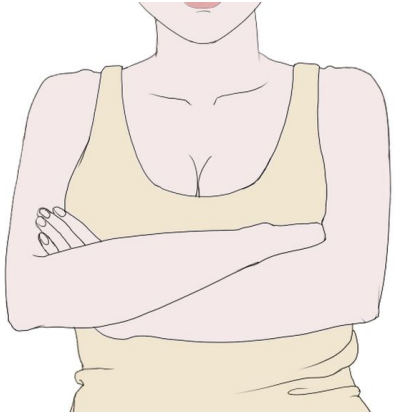
Pose-based Body Language Recognition for Emotion and Psychiatric Symptom Interpretation

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

Emotions from Body Language

- Given an image, predict the semantic segmentation mask for each body part



defensive, disagree

honesty, sincerity

ashamed, upset

Examples of body language and their meanings can be found in this website. https://www.enkiverywell.com/body-language-examples.html

Motivation

Body Language

Open palm facing upward

Emotion

Joy	1

Challenge

Body Language

Open palm facing upward

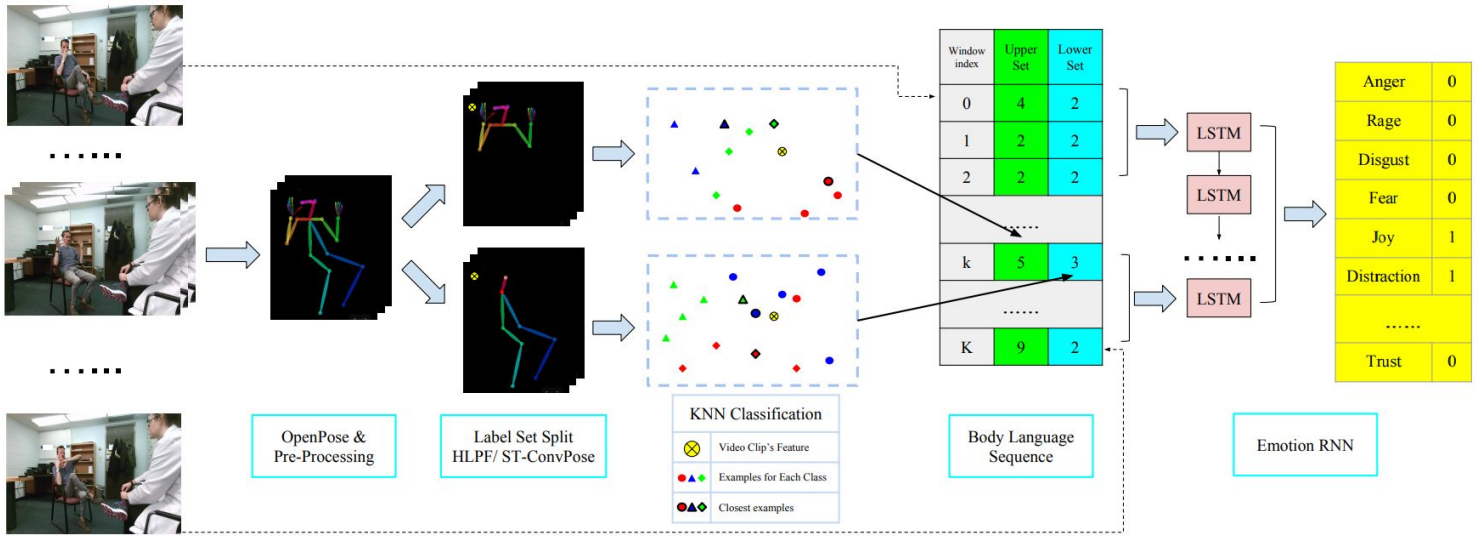
Emotion

Joy	1

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

Method

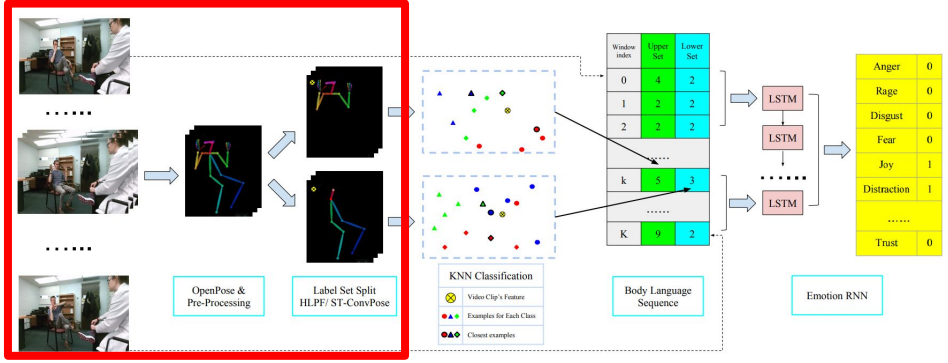
- **Framework overview**

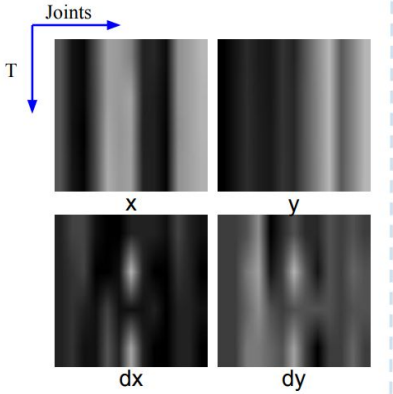


- Body language recognition
- Emotion recognition/ psychiatric symptom interpretation

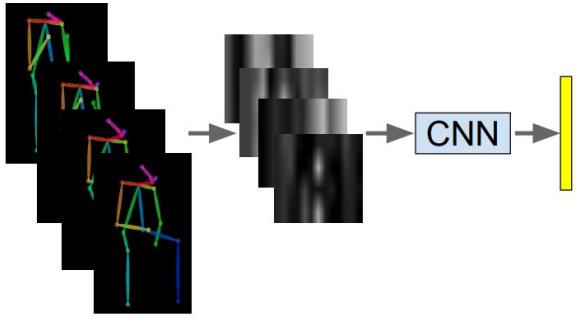
Method

- **Pose-based body language recognition**





(a). Pose Frame Examples

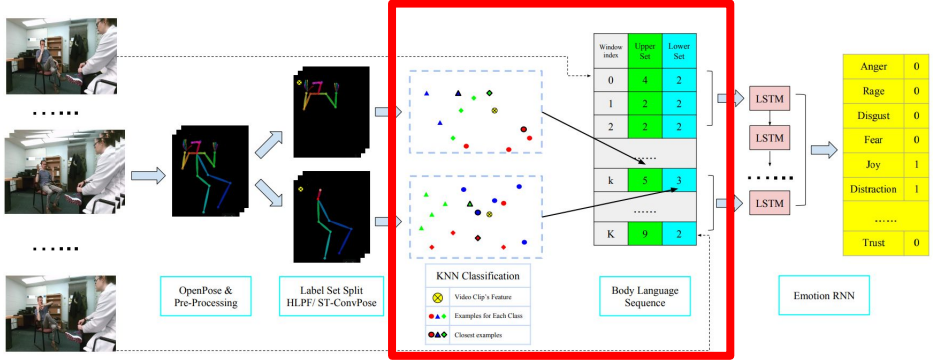


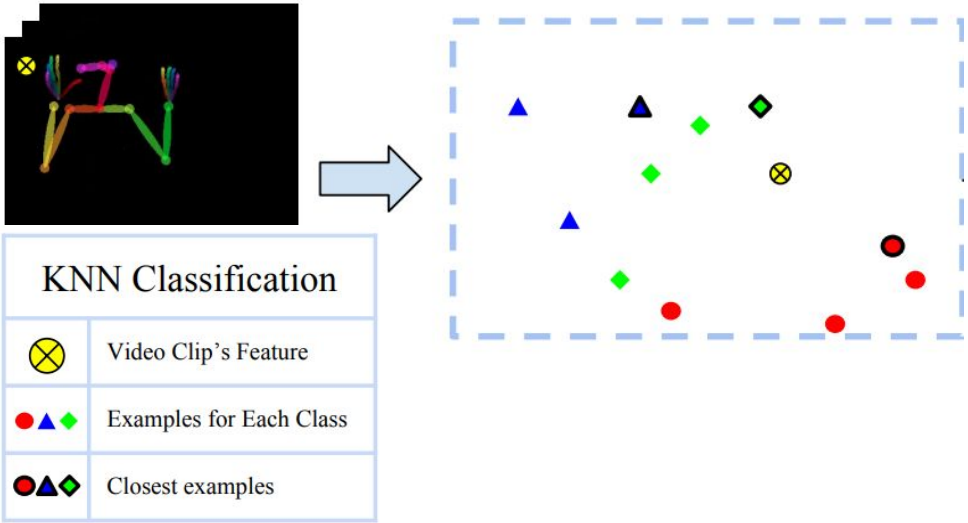
(b). ST-ConvPose Framework

- Pose-based recognition: requires less data and has better transferability
- Pose representation: CNN-based pose image pre-trained on action datasets

Method

- **KNN-based classification**

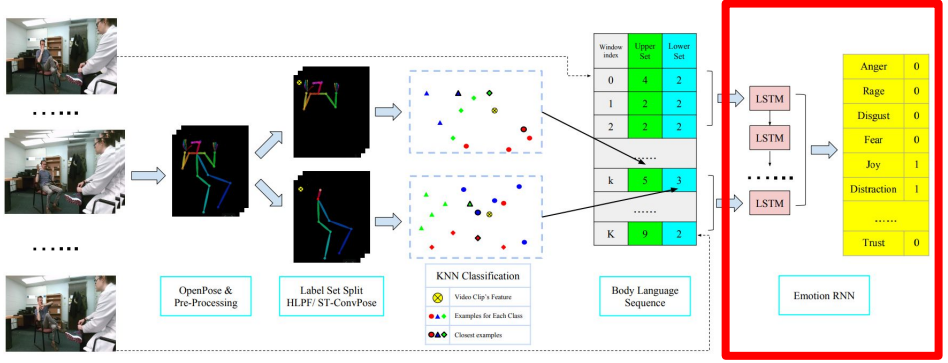




- KNN-based classification: only requires few example annotations. Interpretable and easily transferable to new body languages

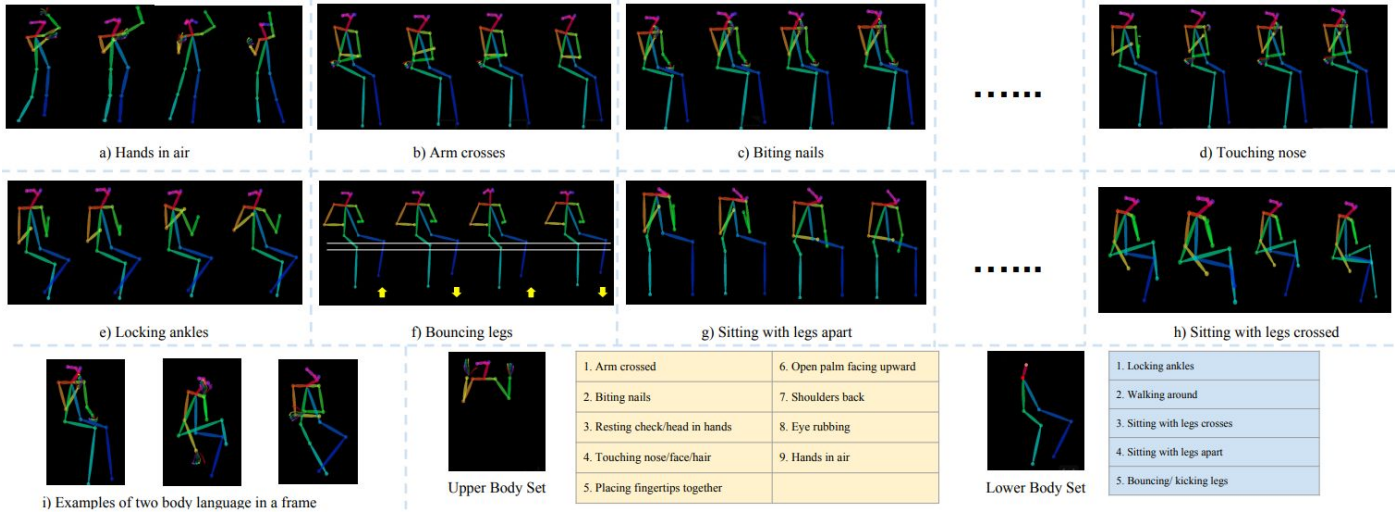
Method

- **Emotion interpretation from body language**



- Emotion interpretation

URMC Dataset



- 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

Experiment Results

- **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

- Interpretable and transferrable framework

Experiment Results

- **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

- 79.9% accuracy on Major Depressive Disorder v.s. with Manic Episode prediction

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

Poster: #79

Contact:
zyang39@cs.rochester.edu