Real-Time Privacy-Preserving Fall Risk Assessment with a Single Body-Worn Tracking Camera

Chiao-Yi Wang[1], Faranguisse Kakhi Sadrieh[1], Yi-Ting Shen[2], Giovanni Oppizzi[1], Li-Qun Zhang[1], Yang Tao[1]
[1]Department of Bioengineering, University of Maryland, College Park, USA
[1]Department of Electrical and Computer Engineering, University of Maryland, College Park, USA

Introduction

Motivation

- Falls are the leading cause of injuries among elderly adults.
- 30% of elderly adults in the US experience at least one fall each year.
- We need fall risk assessment !

Two biomechanical causes for slip-induced falls

- Unrecoverable collapse of limbs in the vertical direction
- Instability in the anterior-posterior (A-P), medial-lateral (M-L) or both directions.

(a) Normal (b) A-P Unstable (c) M-L Unstable

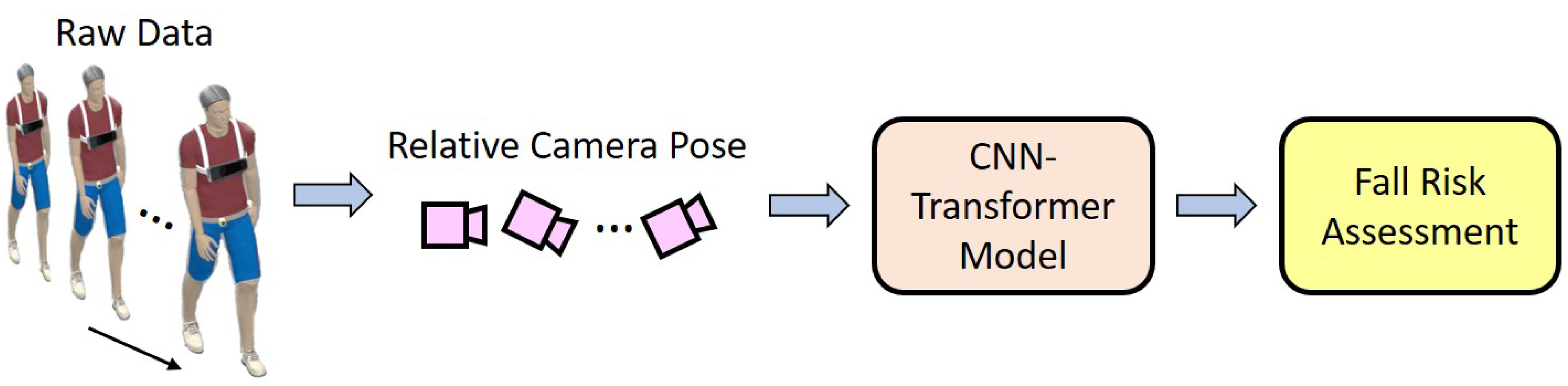
Existing methods & Challenge for fall risk assessment

- Sensors around the subject's surrounding environment (e.g. RGB cameras or radar sensors) → Privacy concerns
- Sensors on the subject's body (e.g. IMUs) → Requires more than one sensor, sensor drift

EgoFall System

Overview

- A chest-mounted tracking camera (Intel RealSense T265)
- A carefully designed pre-processing pipeline
- A lightweight CNN-Transformer model



Data Pre-processing

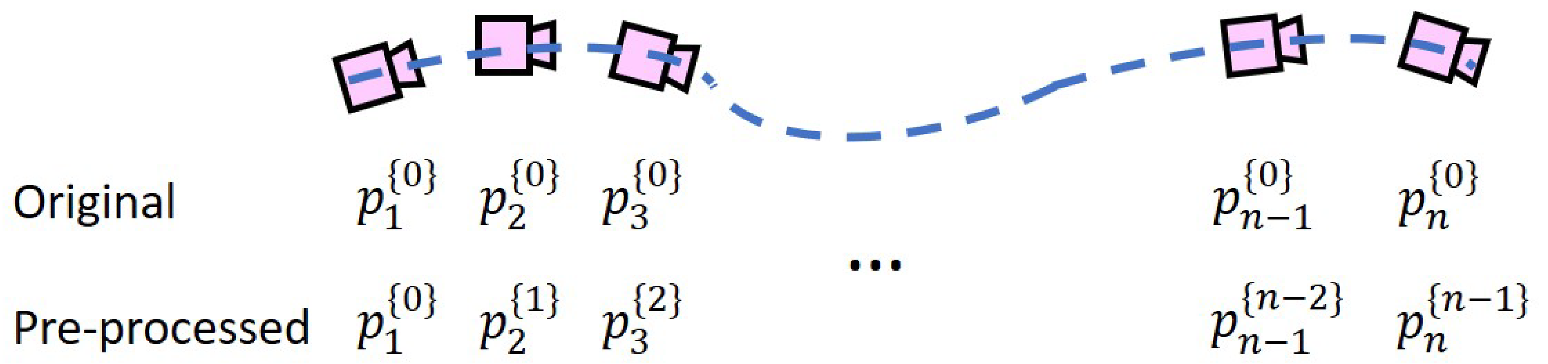
Intel RealSense T265 tracking camera generates the camera pose relative to the *initial* camera pose, which can be represented as a 7-dimensional vector.

$$p_i^{\{0\}} = (t_x, t_y, t_z, r_x, r_y, r_z, r_w)$$

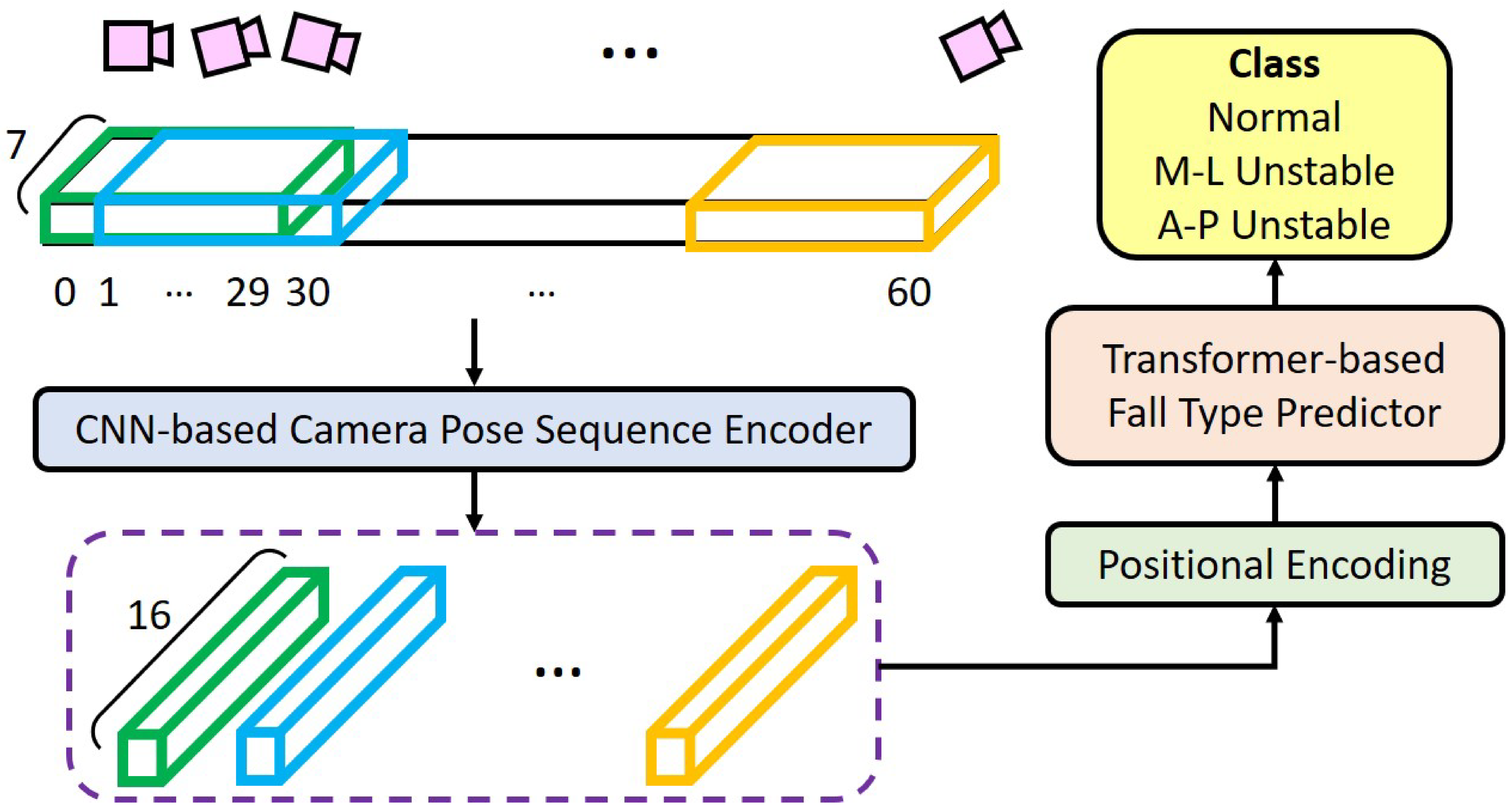
$$p_i^{\{0\}} \equiv P_i^{\{0\}} = \begin{bmatrix} R_i & T_i \\ 0 & 1 \end{bmatrix}$$

We used the camera pose relative to the *previous* camera pose (i.e., $p_i^{\{i-1\}}$).

$$p_i^{\{i-1\}} \equiv P_i^{\{i-1\}} = \begin{bmatrix} R_{i-1}^{-1} R_i & R_{i-1}^{-1}(T_{i-1} - T_i) \\ 0 & 1 \end{bmatrix}$$



Network Architecture

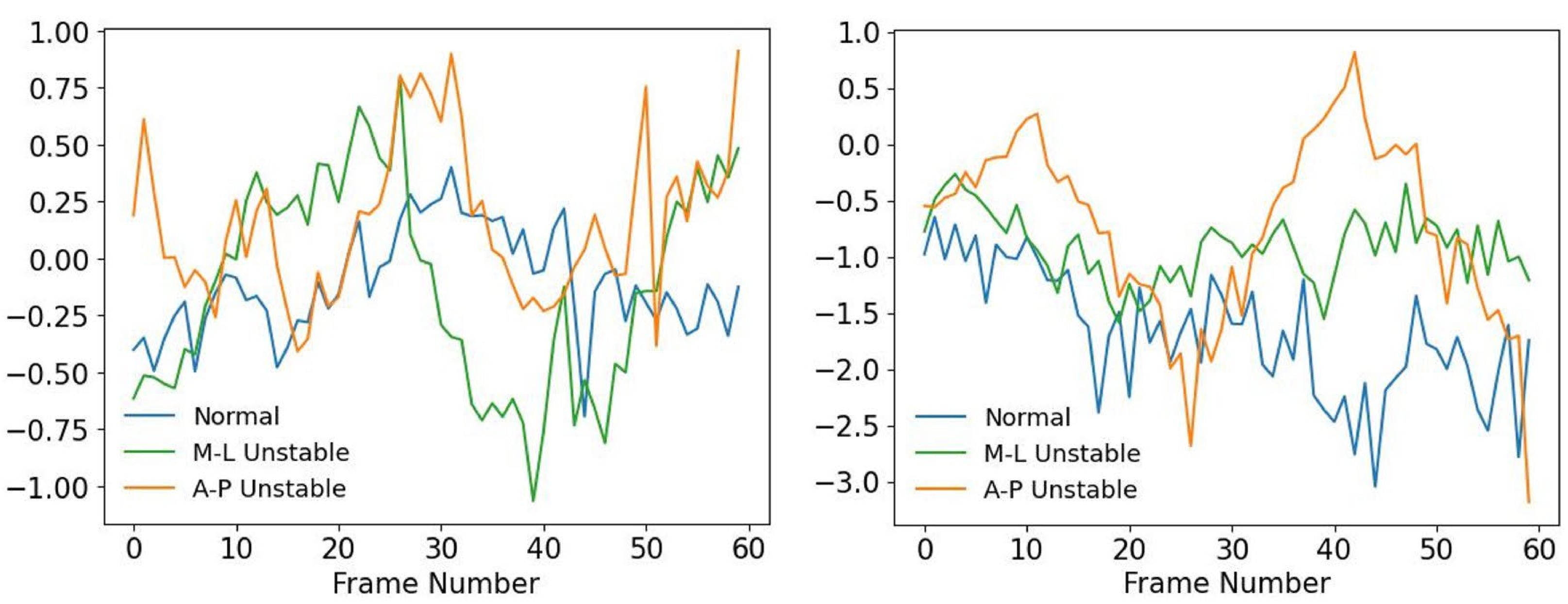


Experiments & Results

Dataset

- We collect 30 camera pose sequences for each of the 3 different walking types: normal, M-L unstable, A-P unstable.
- Total: 15 subjects. (Training: 10, test: 3, validation: 2.)

Walking Type	Training Set	Test Set	Validation Set
Normal	4567	1107	1299
M-L Unstable	8156	2618	1665
A-P Unstable	6764	2355	1661



(a) t_x (b) t_z

Main results

Method	Accuracy	Precision	Recall	F1	Parameters (k)	Runtime (ms)
SVM	0.46	0.30	0.44	0.35	-	-
CNN	0.90 ± 0.034	0.91 ± 0.037	0.91 ± 0.033	0.90 ± 0.034	5.5	**0.15**
LSTM	0.77 ± 0.041	0.77 ± 0.049	0.77 ± 0.049	0.76 ± 0.045	3.8	2
Transformer	0.87 ± 0.045	0.86 ± 0.049	0.89 ± 0.029	0.86 ± 0.048	**3.6**	0.63
CNN-LSTM	0.78 ± 0.133	0.77 ± 0.127	0.77 ± 0.148	0.76 ± 0.143	7.8	1.25
CNN-Transformer (Raw)	0.58 ± 0.031	0.57 ± 0.032	0.59 ± 0.044	0.57 ± 0.039	6.8	0.69
CNN-Transformer (Ours)	**0.93 ± 0.009**	**0.92 ± 0.01**	**0.93 ± 0.003**	**0.93 ± 0.007**	6.8	0.69