Motion and Region Aware Adversarial Learning for Fall Detection with Thermal Imaging

Vineet Mehta*, Abhinav Dhall*†, Sujata Pal*, Shehroz S. Khan+

**Dept. of CSE, Indian Institute of Technology Ropar*

†Human-Centred Artificial Intelligence, Monash University

+KITE, Toronto Rehabilitation Institute, University Health Network, Canada

https://github.com/ivineetm007/Fall-detection

Background

Problem

- Falls can cause serious injuries.
- Insufficient training data.
- Privacy issues with the RGB videos.

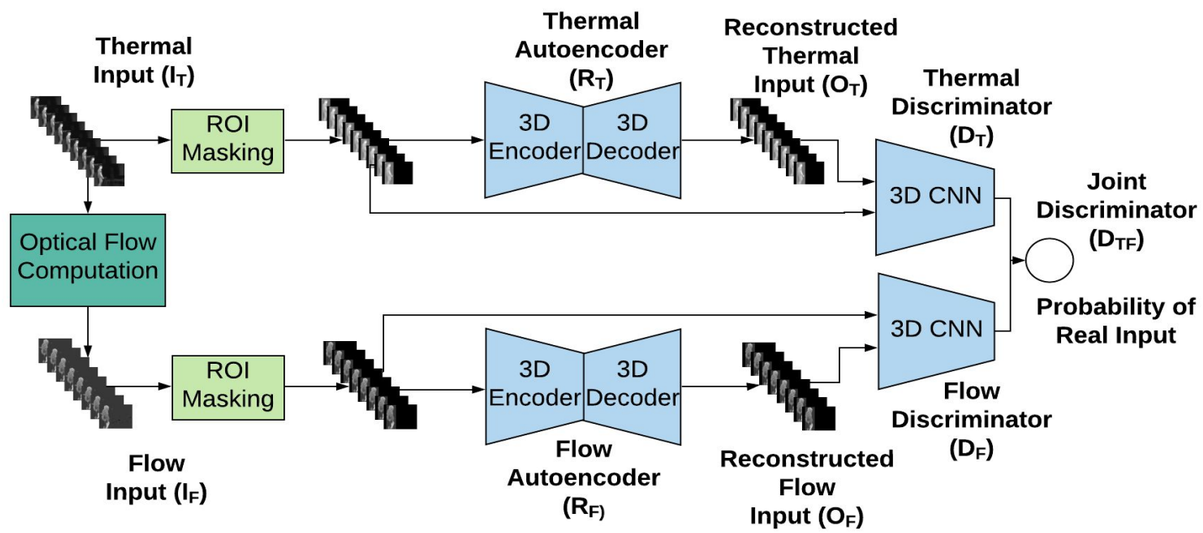
Method Hypothesis

- Formulated as anomaly detection using thermal imaging.
- One Class Classification Paradigm- Learning distribution of normal activities and distinctive patterns are detected as a fall event.

Previous Methods

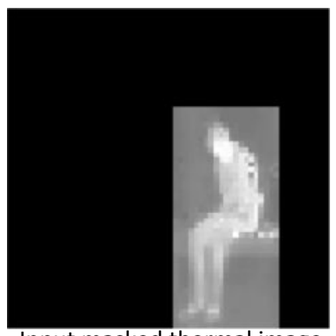
- Traditional ML techniques- One-class SVM, PCA, and Fisher Vectors, etc.
- Deep learning methods-
  - Using Convolutional LSTM Autoencoder (Nogas et al. 2018) or Spatio-Temporal Convolutional Autoencoders (Nogas et al. 2020)
  - Spatio-temporal adversarial learning (Khan et al. 2020) for detecting unseen falls in thermal and depth videos.

Method Part 1



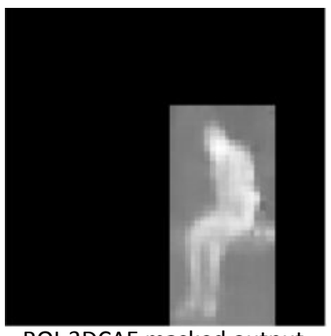
Overall pipeline of the framework.

Method Part 2



Input masked thermal image

ROI-3DCAE output



ROI-3DCAE masked output

Region based Reconstruction

- Person Tracking- R-FCN based detector, Contour Box Localization and Kalman Filtering.
- Frame masking by their region of interest (ROI) and region based losses.

Motion based Reconstruction

- Difference constraint- Difference of consecutive frames and their reconstruction.
- Optical Flow- Dense optical computation and spatio-temporal network for flow reconstruction.

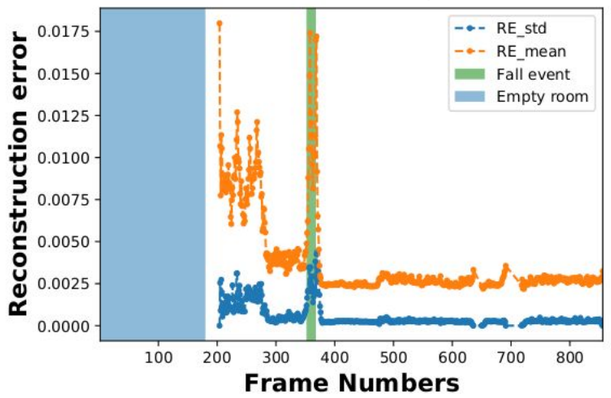
Fusion

- Two spatio-temporal autoencoders for reconstruction of thermal and optical flow sequences.
- Single Discriminator consists of two 3D CNN joined by a single sigmoid neuron.

Experiments and Results

Method	All frames		Tracked frames			
	ROC		ROC		PR	
	C_μ	C_σ	C_μ	C_σ	C_μ	C_σ
Conv-LSTM AE (J. Nogas et al.)	0.76	0.83	0.63	0.73	0.26	0.37
DSTCAE-C3D (J. Nogas et al.)	0.93	0.97	0.85	0.90	0.46	0.53
3DCAE-3DCNN (S. S. Khan et al.)	0.95	0.95	**0.90**	0.88	0.47	0.48
Fusion-Diff-ROI-3DCAE (Ours)	—	—	**0.90**	**0.93**	**0.57**	**0.57**

Comparison with the previous methods based on AUC of ROC and PR curve using Mean (C_μ) and Std. (C_σ) of frame reconstruction errors.



Conclusion and Future works

- Region based learning for background agnostic models.
- Motion constraints for discriminative spatio temporal learning.
- In future work, we will use Depth or IP cameras and its fusion with thermal imaging.

References

- Nogas et al., Fall detection from thermal camera using convolutional lstm autoencoder, IJCAI Workshop. 2018.
- Nogas et al., Deepfall: Non-invasive fall detection with deep spatio-temporal convolutional autoencoders, Journal of Healthcare Informatics Research 2020
- Khan et al., Spatio-temporal adversarial learning for detecting unseen falls, Pattern Analysis and Applications, 2020