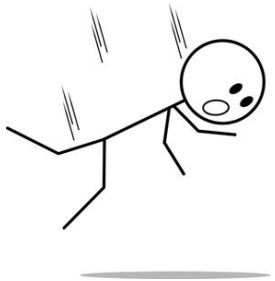
Motion and Region Aware Adversarial Learning for Fall Detection with Thermal Imaging



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

*Dept. of CSE, Indian Institute of Technology Ropar

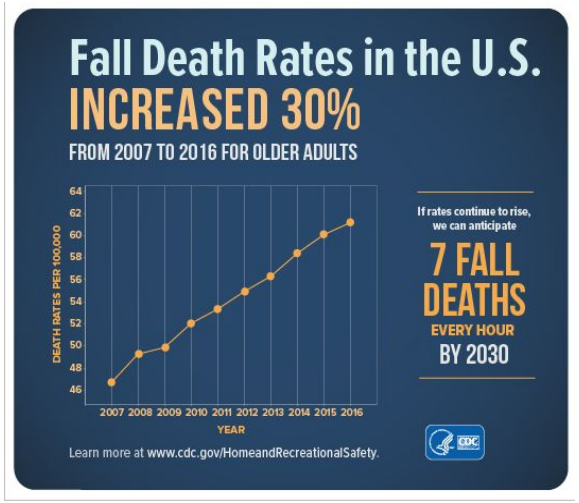
†Human-Centred Artificial Intelligence, Monash University

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

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

Introduction

- Falls can cause serious injuries
- Insufficient training data - **Anomaly Detection (One class Classification)**
- Privacy issues with RGB videos - **Thermal Imaging**



Fall Death rate over past years

Image Source: https://www.cdc.gov/homeandrecreationalsafety/falls/adultfalls.html

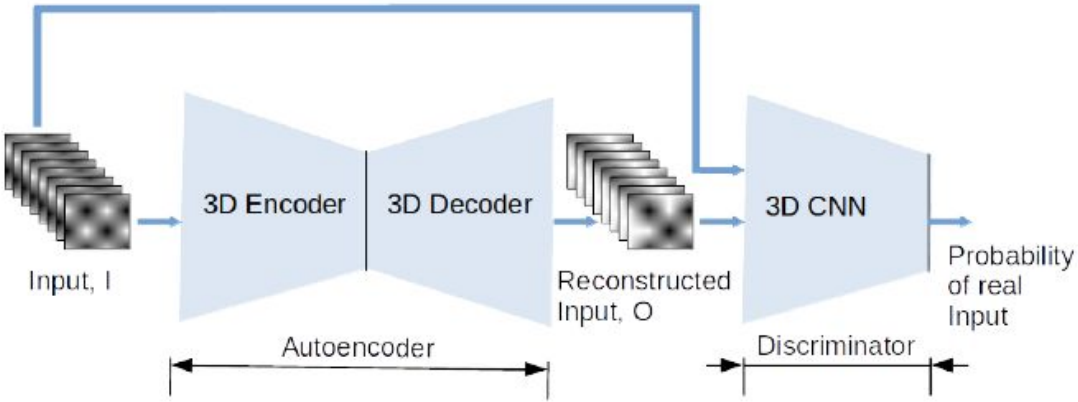
Related Works

- Traditional Techniques- One-class Support Vector Machines, Principal Component Analysis, Fisher Vectors, etc.
  - Thermal imaging based elderly fall detection (S. Vadivelu et al. ACCV 2016)
- Deep Learning methods- Autoencoders, Adversarial learning
  - Fall Detection from Thermal Camera Using Convolutional LSTM Autoencoder (J. Nogas et al. IJCAI Workshop 2018)
  - DeepFall - Non-invasive Fall Detection with Deep Spatio-Temporal Convolutional Autoencoders (J. Nogas et al. Journal of Healthcare Informatics Research 2020)
  - Spatio-temporal adversarial learning for detecting unseen falls (S. S. Khan et al. Pattern Analysis and Applications, 2020)

Adversarial Learning

- Discriminator Loss
  - Binary Cross Entropy
- Autoencoder Loss
  - Reconstruction Loss-MSE, $\mathbf{L_R}$
  - Binary Cross Entropy (at Discriminator) $\mathbf{L_{R+D}}$
  - Total Loss

$$\mathbf{L = L_{R+D} + \lambda L_R}$$



Spatio-temporal adversarial framework

Image Source: Spatio-temporal adversarial learning for detecting unseen falls (S. S. Khan et al. Pattern Analysis and Applications, 2020)

Region Extraction

- Person detection- R-FCN network
- Contour Box Localization
- Tracking- Kalman filtering

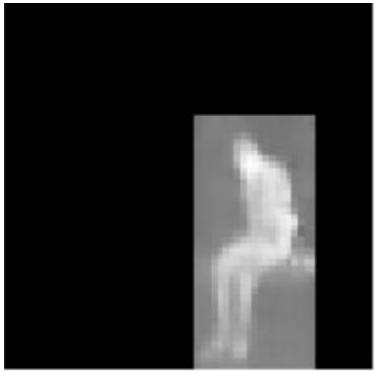
Contour Box Localization

Region based Reconstruction

- Frame masking by their region of interest (ROI)
- Region based Reconstruction Loss- $\mathbf{L_{ROI} = E_{ROI}\ [(ROI(I) - ROI(O))^2]}$
- Total Loss**(Thermal-ROI-3DCAE)**- $\mathbf{L = L_{R+D} + \lambda L_{ROI}}$

ROI-3DCAE input

ROI-3DCAE output



ROI-3DCAE masked output

Motion based Reconstruction

- **Difference constraint**
  - Difference frames
  - ROI masking- Union of the ROI of the individual frames
  - Difference frames reconstruction loss

    $$L_{Diff} = E_{ROI}\,[(ROI(DF(I)) - ROI(DF(O)))^2]$$
  - Total loss **(Thermal-Diff-ROI-3DCAE)**

    $$L = L_{R+D} + \lambda_S L_{ROI} + \lambda_D L_{Diff}$$
- **Optical Flow**
  - Dense optical flow for consecutive frames
  - Spatio-temporal network (**Flow-ROI-3DCAE**) for flow reconstruction

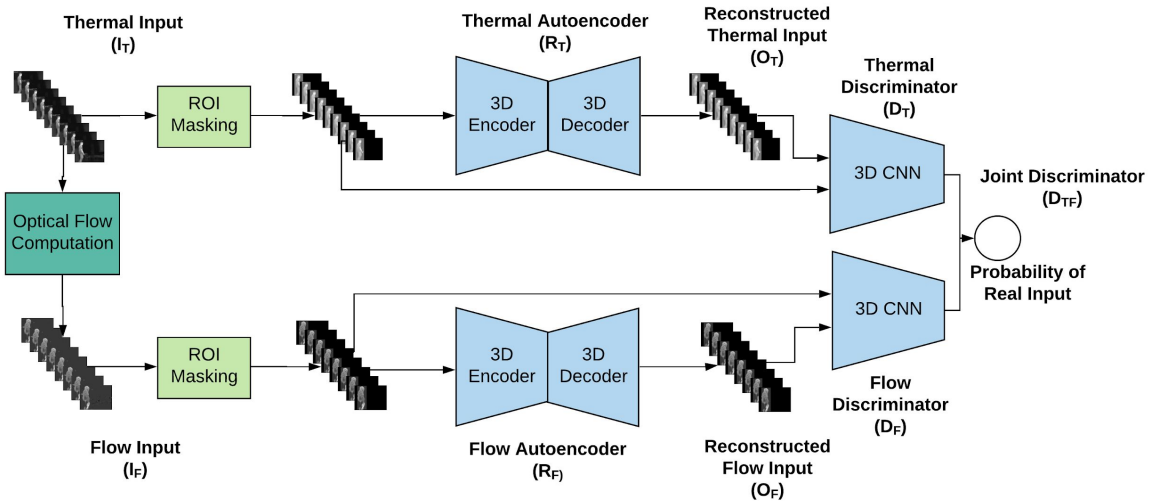
Thermal and Flow fusion

- Discriminator Loss
  - Binary Cross Entropy
- Autoencoder Loss
  - **Fusion-ROI-3DCAE**

$$L = L_{R+D} + \lambda_T L_{T\_ROI} + \lambda_F L_{F\_ROI}$$

  - **Fusion-Diff-ROI-3DCAE**

$$L = L_{R+D} + \lambda_{T\_S} L_{T\_ROI} + \lambda_{T\_D} L_{T\_Diff} + \lambda_F L_{F\ ROI}$$



Fusion 3DCAE network

Dataset and Evaluation

- TSF dataset by Vadivelu et al. 2016
  - 35 Fall videos (36,391 frames total, 828 fall frames)
  - 9 videos (22,116 frames) contain only ADL(Activities of daily life)
- Evaluation
  - AUC of ROC
  - AUC of PR curve

Top: Empty frames, **Middle**: ADL, **Bottom**:Fall

Ablation Studies

Method	ROC		PR	
	C_μ	C_σ	C_μ	C_σ
Thermal-3DCAE	0.88	0.90	0.47	0.48
Thermal-ROI-3DCAE	0.89	0.92	0.55	0.57
Thermal-Diff-ROI-3DCAE (ROI score)	0.90	0.92	0.57	0.56
Fusion-ROI-3DCAE (Thermal ROI score)	0.90	0.93	0.56	**0.58**
Fusion-Diff-ROI-3DCAE (Thermal ROI score)	**0.90**	**0.93**	**0.57**	0.57

AUC of ROC and PR curve using Mean($\mathbf{C_\mu}$) and Standard Deviation($\mathbf{C_\sigma}$) of frame reconstruction errors

Comparison

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	—	—	**0.90**	**0.93**	**0.57**	**0.57**

Comparison with the previous methods based on AUC of ROC and PR curve calculated on frame level anomaly scores

Prediction

Video Source: https://drive.google.com/file/d/13GXHlbwG8fbsM2aZ-Cbyo1G3hzIVGPVw/view?usp=sharing

Conclusion and Future works

- Region based learning for background agnostic models
- Motion constraints for discriminative spatio temporal learning

- Future directions
  - Depth camera or IP camera
  - Fusion with thermal imaging

Thank You

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