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

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.

Method Part 2

Experiments and Results

<table>
<thead>
<tr>
<th>Method</th>
<th>All frames</th>
<th>Tracked frames</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROC</td>
<td>ROC</td>
<td>PR</td>
</tr>
<tr>
<td>Conv-LSTM AE (J. Nogas et al.)</td>
<td>0.76</td>
<td>0.83</td>
</tr>
<tr>
<td>DSTCAE-C3D (J. Nogas et al.)</td>
<td>0.93</td>
<td>0.97</td>
</tr>
<tr>
<td>3DCAE-3DCNN (S. S. Khan et al.)</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td>Fusion-Diff-ROI-3DCAE (Ours)</td>
<td>0.90</td>
<td>0.93</td>
</tr>
</tbody>
</table>

Comparison with the previous methods based on AUC of ROC and PR curve using Mean \(C_\mu\) and Std. \(C_\sigma\) 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.