Human Fall Detection Based on ResNet and LSTM Network

Supervisor: Kin-Choong Yow
Xinyu Liu 200362878
Aug, 23, 2022

Problem

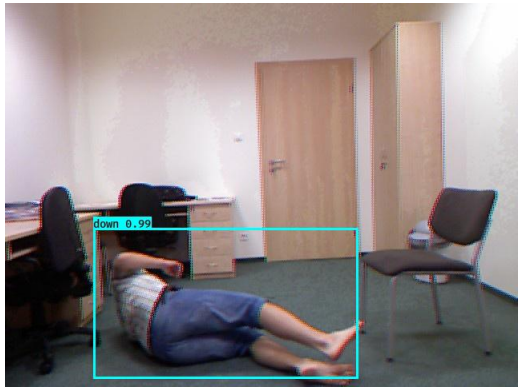
- With the development of population aging appearing in various countries, more and more elderly people appear in society
- The accidental fall is one of main factors that hurt their bodies
- It could lead to more serious injury or death

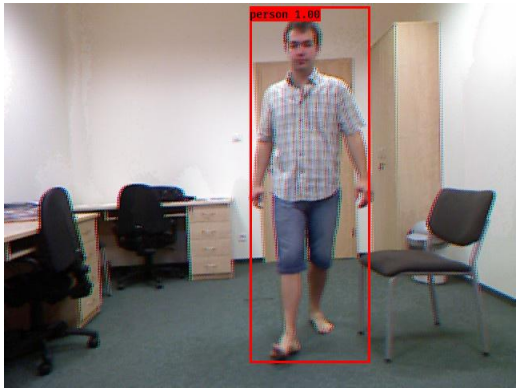
Motivation

- Lack of an efficient fall monitoring method - human observers would missed more than 60% information during the manual surveillance
- Internet of Thing (IoT) - false alarm, not convenient for elderly
- Object - Human fall detection model based on computer version
- Applying in home, nursing homes or hospitals. Providing immediate assistance

Related Work

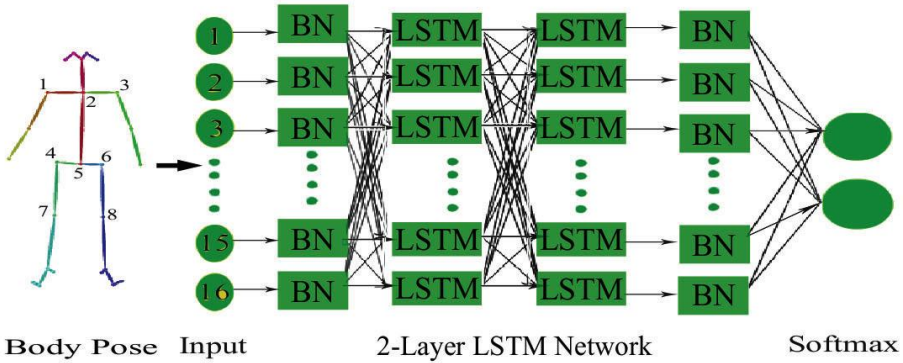
- In 2020, Wang and Jia[1] proposed an approach for human fall detection that is based on YOLOv3 algorithm with Darknet-53





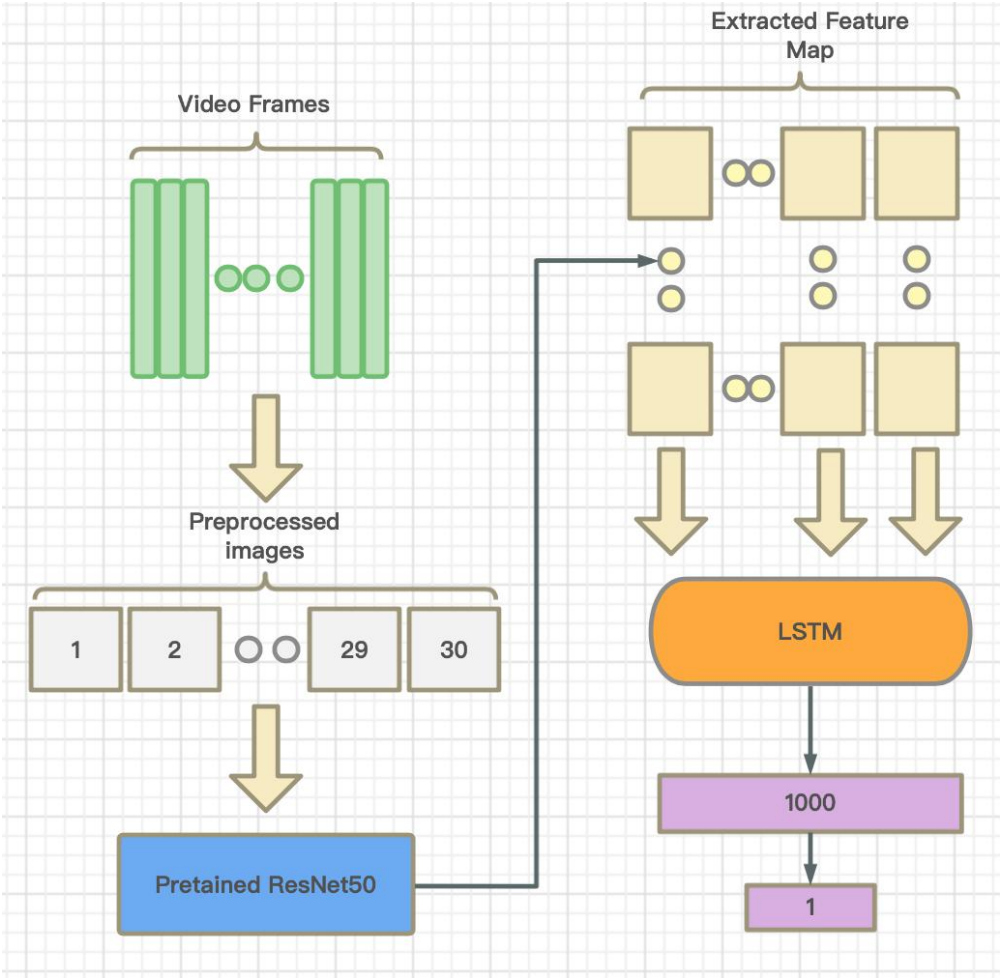
Related Work

- In 2019, Hasan et al[6] has proposed an approach to detect human fall action by developed Recurrent Neural Network **(**RNN) with human pose estimator.



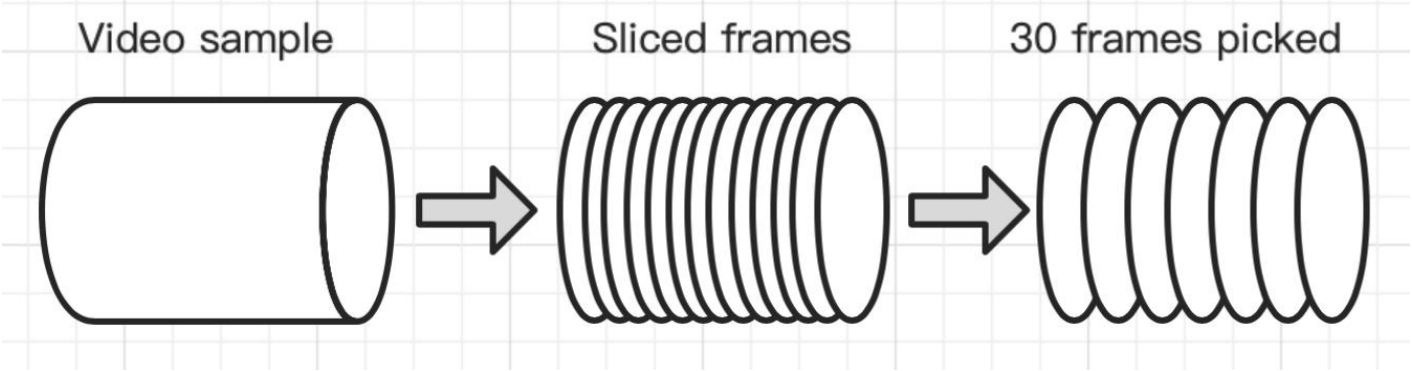
Architecture

- ResNet50 + LSTM
- Process video and frames
- ResNet50 extracted feature map
- Combine feature maps in time sequence, fed into LSTM network to train.
- Fully connected layer output result



Frame videos

- Divide the each video into frames format.
- 30 frames are averagely selected for the further data process.



Frame Preprocess

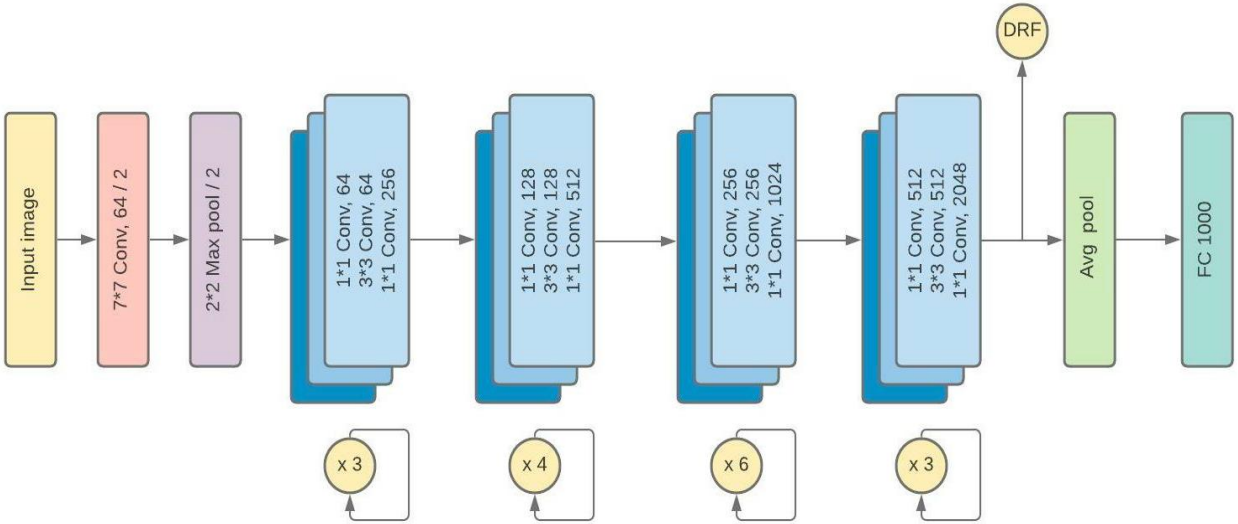
- Resized image into 224*224 for input trained Resnet50
- Keep important information
- Normalization, convert pixel range 0-255 into 0-1 within 3 channels

$$Pixel_{normal} = \frac{Pixel_{original}}{255}$$

Not resized image

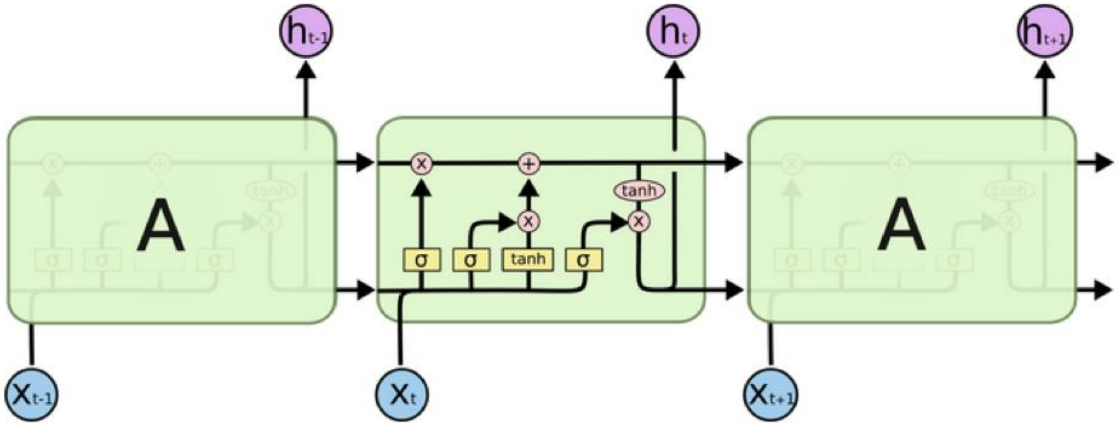
Resized image

ResNet50



- Residual neural network (ResNet)
- Pass 224 * 224 frame with 3 channel into ResNet50
- The feature maps will be generated before the average pooling layer
- Flatten the extracted data into 1-dimension

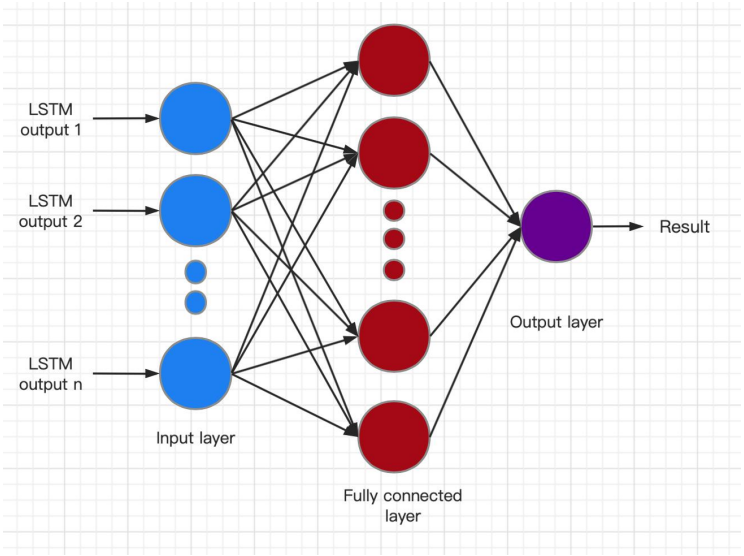
Long-short term memory (LSTM)



- Well-known Recurrent Neural Network(RNN)
- Outstanding performance on dealing with long-sequence training
- Extracted feature data from ResNet50 to LSTM in time series.

Fully connected layer

- The output of LSTM will compute with weight in fully connected layer
- Output of the fully connected layer will be used to classify fall
- <0.5 is non-fall, >0.5 is fall.



Dataset

- University of Rzeszow Fall Detection Dataset (URFD)
- Fall Detection Dataset (FDD)
- Videos samples Include fall action and other daily living
- Randomly selected training(80%) and testing set(20%).

Dataset/ Category	No. Fall Videos	No. Non–Fall Videos	No. Total Samples	No. Train samples	No. Test samples
URFD	30	40	70	56 (80%)	14 (20%)
FDD	121	70	191	153 (80%)	38 (20%)
Total	151	110	261	209	52

One sequences Fall frames from URFD

Frame 81

Frame 87

Frame 97

Frame 103

Frame 110

Frame 114

Frame 118

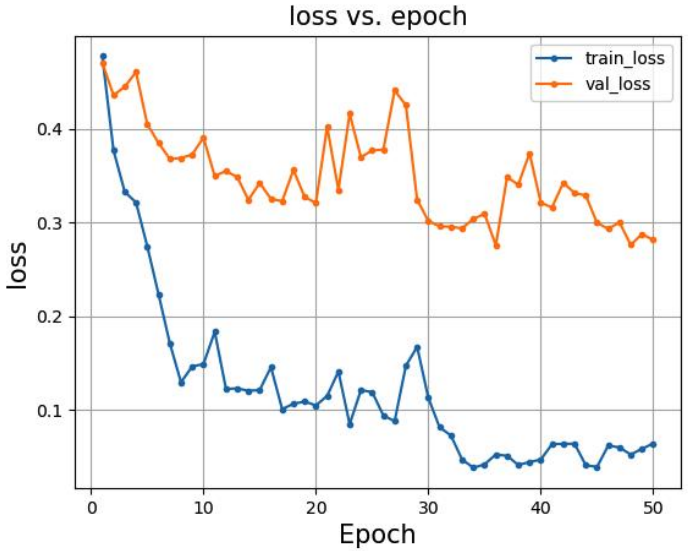
Frame 134

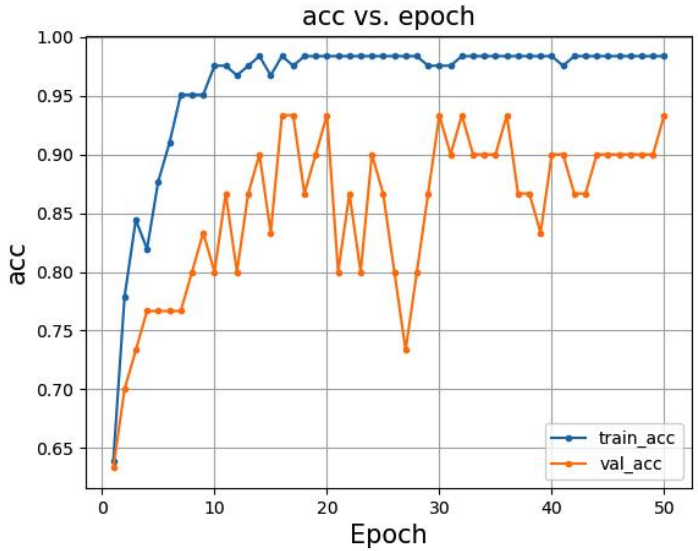
Experiment

- Python
- PyTorch
- Pre-trained ResNet50 from the PyTorch
- 80.8% accuracy on the ImageNet database

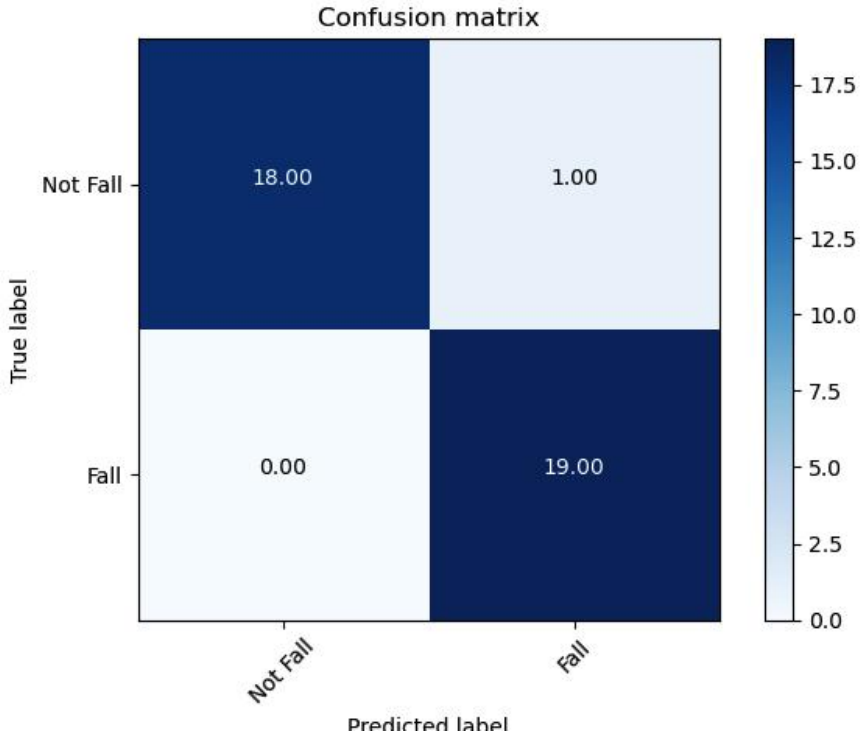
- OpenCV - pre-process the video sample which comvert them into frames.
- learning rate - 0.0001
- echos - 50 with 8 batch size

Evaluation (FDD)





Evaluation (FDD)



Evalution

(5) $$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

(6) $$Sensitivity = \frac{TP}{TP + FN}$$

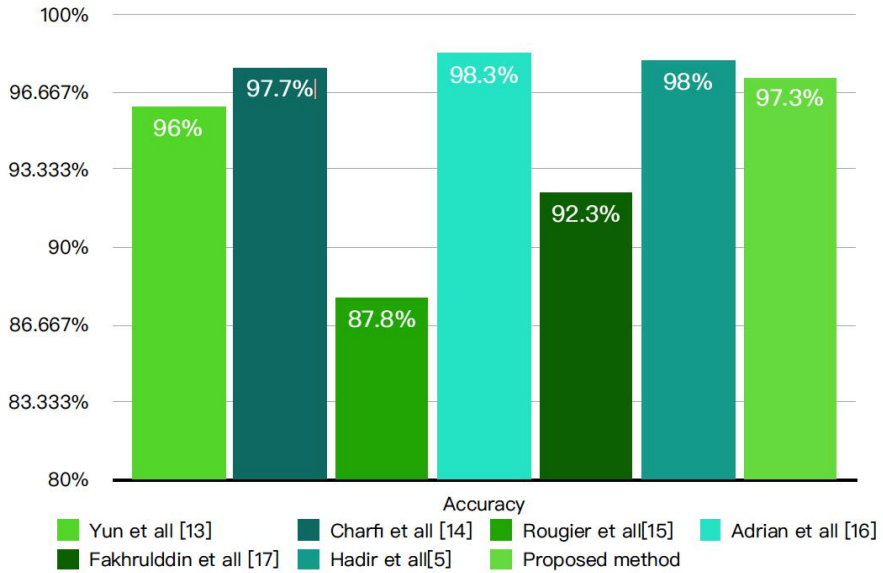
(7) $$Specificity = \frac{TN}{FP + TN}$$

	Output Label		
Actual Label	Class	Fall	Not Fall
	Fall	True Positive (TP)	False Negative (FN)
	Not Fall	Flase Positive (FP)	True Negative (TN)

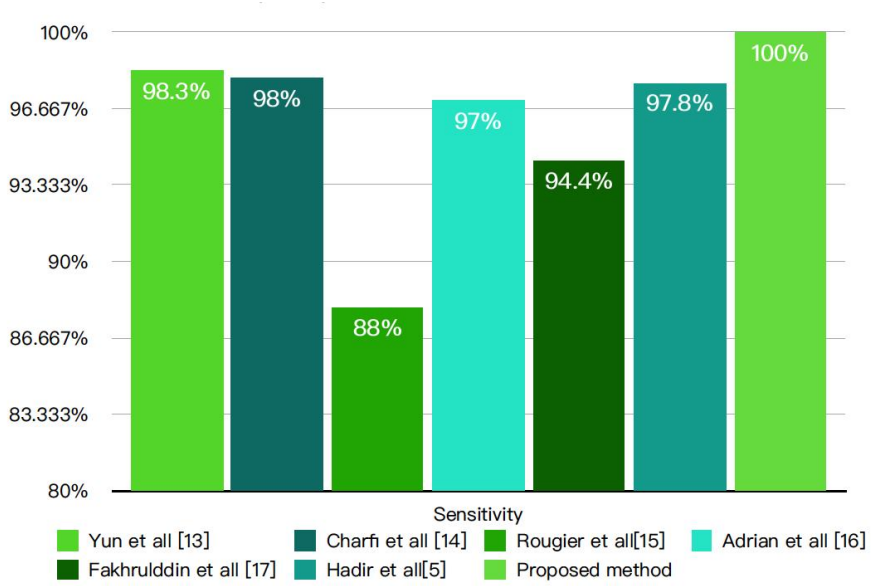
Evaluation Results

Evaluation Matrix	(Value%) for URFD	(Value%) for FDD
Accuracy	100%	97.3%
Sensitivity / Recall	100%	100%
Specificity	100%	94.7%

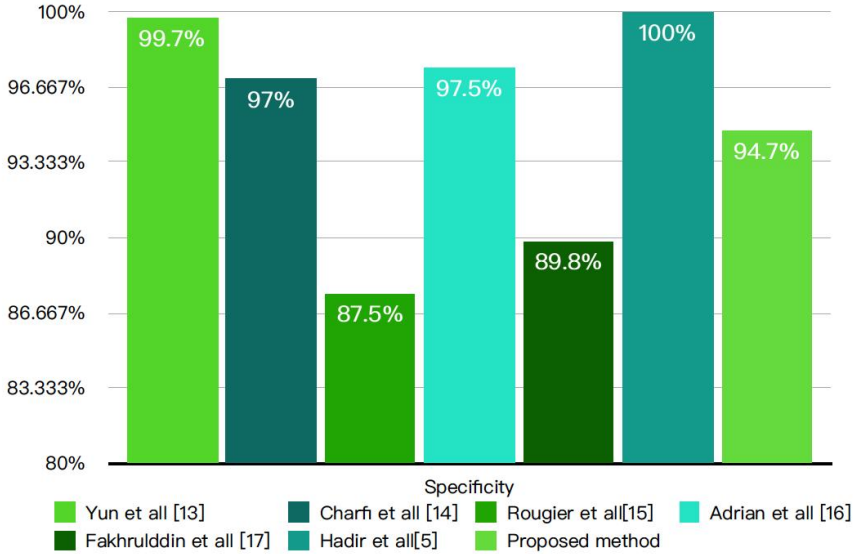
Performance comparison on FDD with other methods



Performance comparison on FDD with other methods



Performance comparison on FDD with other methods



Demo

Conclusion

- Human fall detection model
- ResNet50 + LSTM
- Two datasets, URFD, FDD
- Comparsion with human fall detection method.

References

[1]Wang, \& Jia, K. (2020). Human Fall Detection Algorithm Based on YOLOv3. 2020 IEEE 5th International Conference on Image, Vision and Computing (ICIVC), 50–54. https://doi.org/10.1109/ICIVC50857.2020.9177447

[2]M. M. Hasan, M. S. Islam and S. Abdullah, "Robust Pose-Based Human Fall Detection Using Recurrent Neural Network," 2019 IEEE International Conference on Robotics, Automation, Artificial-intelligence and Internet-of-Things (RAAICON), 2019, pp. 48-51, doi: 10.1109/RAAICON48939.2019.23.

[3]Soheil Vosta, \& Kin-Choong Yow. (2022). A CNN-RNN Combined Structure for Real- World Violence Detection in Surveillance Cameras. Applied Sciences, 12(3), 1021. https:// doi.org/10.3390/app12031021

[4]Bogdan Kwolek, Michal Kepski, Human fall detection on embedded platform using depth maps and wireless accelerometer, Computer Methods and Programs in Biomedicine, Volume 117, Issue 3, December 2014, Pages 489-501, ISSN 0169-2607 http://fenix.univ.rzeszow.pl/~mkepski/ds/uf.html

[5]fall detection Dataset .Available online “ http://le2i.cnrs.fr/Falldetection-Dataset “ Antoine Trapet - 27 February 2013.