ITU KALEIDOSCOPE
ATLANTA 2019

Elderly Health Monitoring System With Fall Detection Using Multi-Feature Based Person Tracking

Dhananjay Kumar &
Vivekanandan Dharmalingham

Anna University, MIT Campus, Chennai, India
dhananjay@annauniv.edu

4-6 December
Atlanta, Georgia, USA

Outlines

- Motivation
- System Architecture
- YOLO Model Based Object Detection
- Similarity Computation
- Fall Detection Approach
- Implementation Results

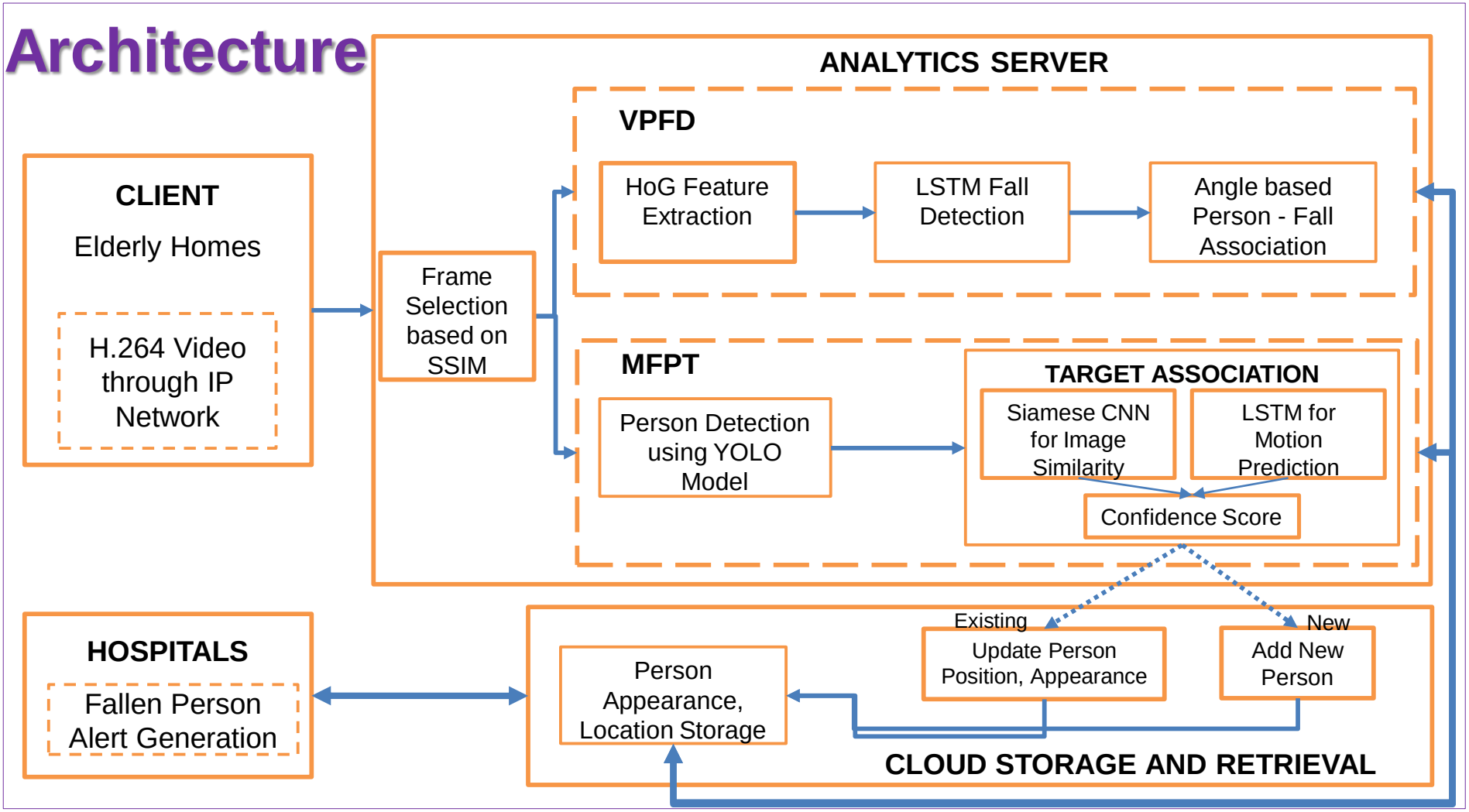
Motivation

- Need for an efficient health monitoring system for the rapidly increasing elderly population
- Usage of wearable technology for monitoring, entirely dependent on the wearer
- A non-invasive, hassle-free elderly location monitoring and fall detection system helps in timely treatment, avoiding severe effects
- Real-time data analysis using machine learning and cloud computing enables better and accurate health-care systems

Scope of the Work

- Improving object tracking and detecting falls using **deep learning**
- **Multi-Feature based Person Tracking (MFPT)**
  - Visual similarity
  - Motion similarity
- **Vision based Person Fall Detector (VPFD)**
  - Bounding box information from MFPT
  - Histogram of Oriented Gradients (HoG)
  - Long Short-Term Memory (LSTM)

Architecture



YOLO Model: Person Detector

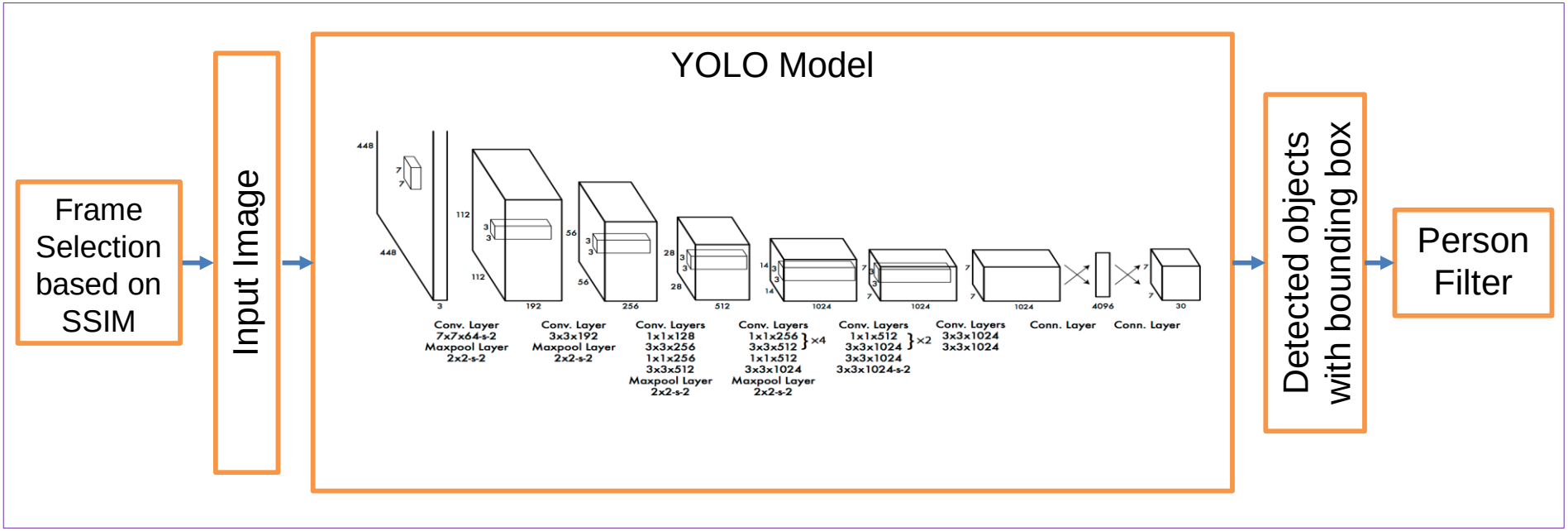


Figure 2 – YOLO Model based Object Detection

- Structural Similarity Index (SSIM) – to select key frame in improving processing speed

Similarity Computation

- **Target Association**: effective tracking of the same person in different frames
- **Image Similarity**: mapping visually similar persons
- **Motion Similarity**: mapping nearby persons with similar previous movements

Similarity Computation (cont.)

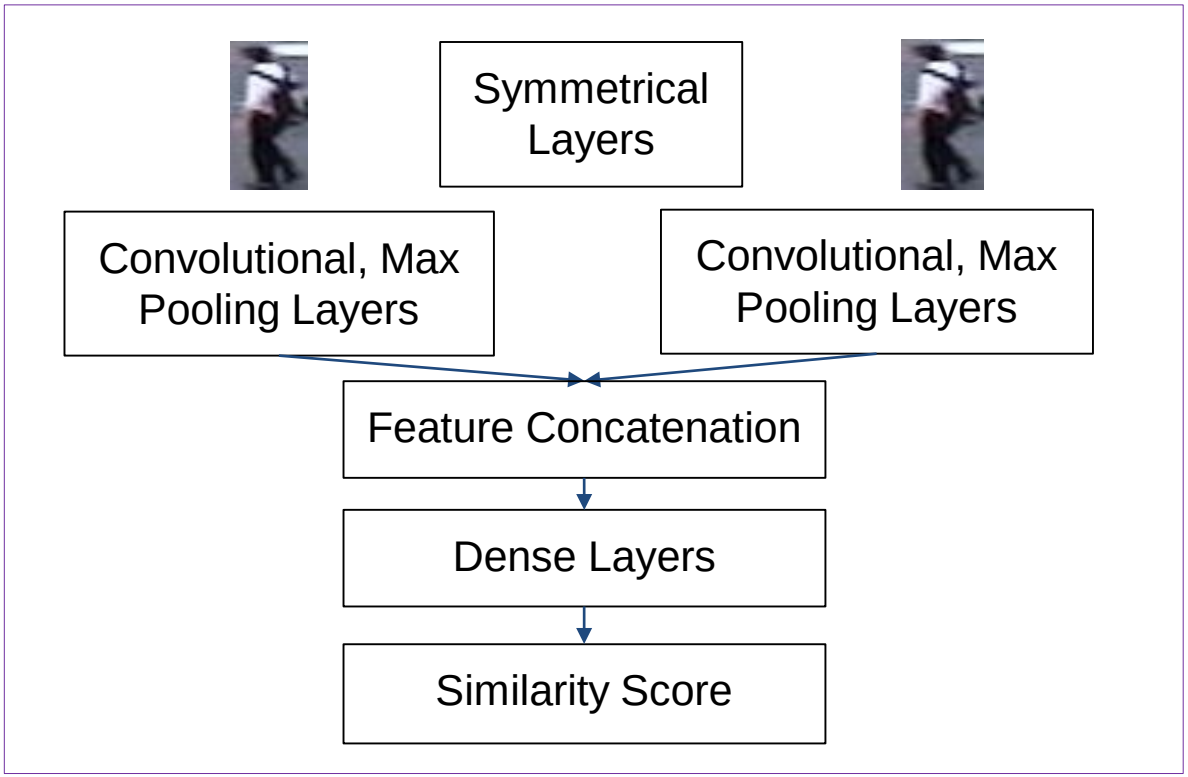


Figure 3 - Image Similarity using Siamese CNN

Similarity Computation (cont.)

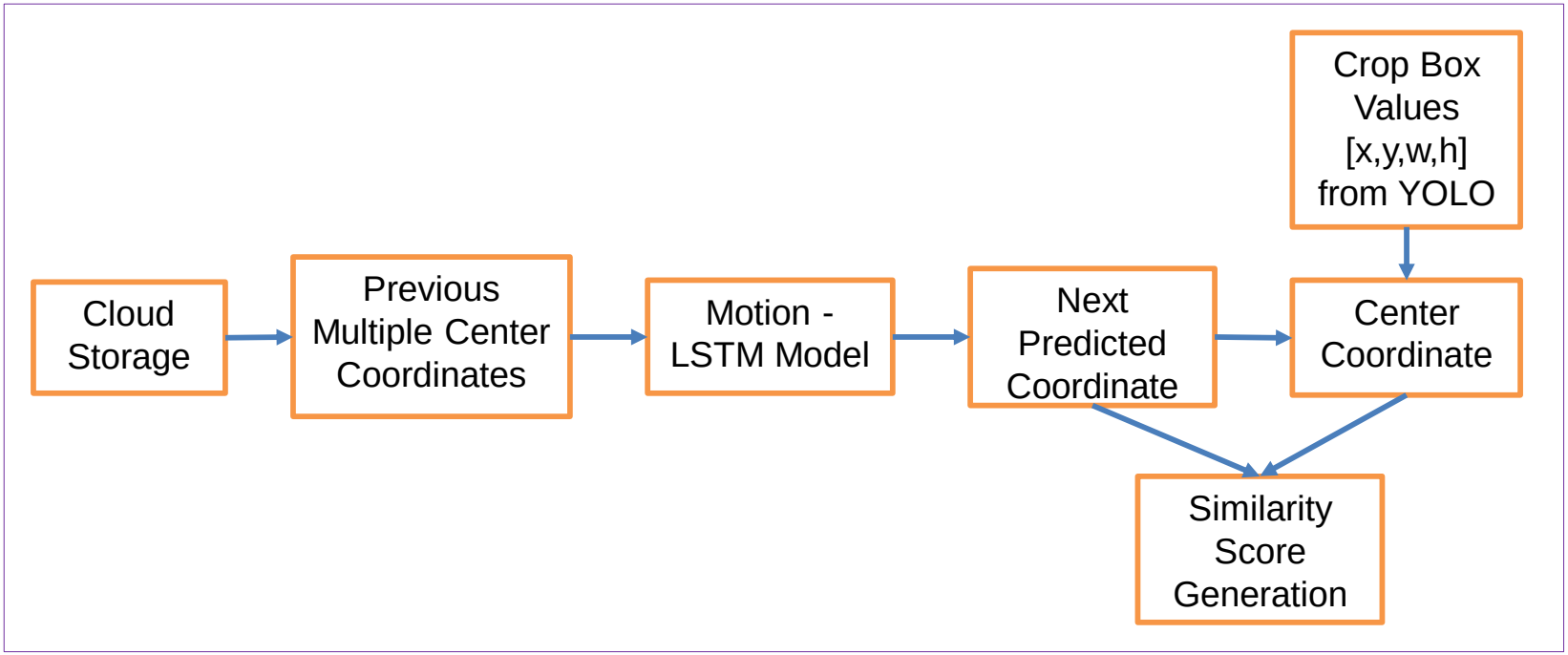


Figure 4 - Motion Similarity

Similarity Computation (cont.)

Two possible scenarios evaluated

- Selection of best candidate with highest image and motion similarity score
  - Score greater than specific threshold
  - The current person is mapped to target candidate
- If no such candidates are found, current person is considered as new person in the scene

Fall Detection Approach

VPFD Model

- Uses combination of:
  - Rate of change in angle obtained from bounding box
  - HoG - LSTM model for fall classification
- HoG – feature extraction, a representation of object's edge orientations and structure
- Sequence analysis using LSTM, to detect fall occurrence

Fall Detection Approach (cont.)

- Using Bounding Box

$$OA = (\mathrm{x1} - \mathrm{x0})i + (y1 - y0)j$$

$$OB = (\mathrm{x2} - \mathrm{x0})i + (y2 - y0)j$$

$$\cos\alpha = \frac{OA \,.\, OB}{|OA||OB|}$$

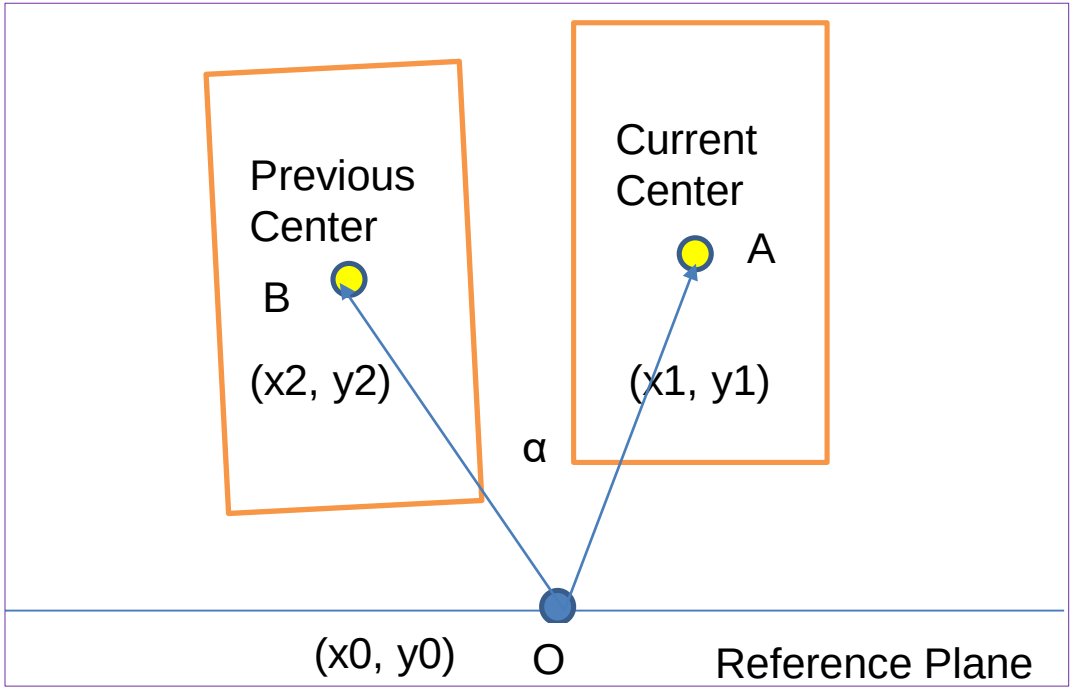


Figure 5 – Angle between two centers of same person

- Avoids false detection using angle threshold and information from HoG-LSTM model

Implementation Overview

- **Hardware**
  - Intel i5 Processor, Windows OS, 8GB RAM computing machine
  - Nvidia GeForce 940MX graphics card (for training in deep learning)
- **Software**
  - Anaconda distribution of Python programming language
  - Python flask web framework
  - Keras, Tensorflow library for deep learning model construction
  - Scikit-learn for frame processing
- **Dataset**
  - UR Fall dataset: 30 Fall event and 40 normal videos of daily life activities
  - Training - Testing split of data: 80 - 20

Results

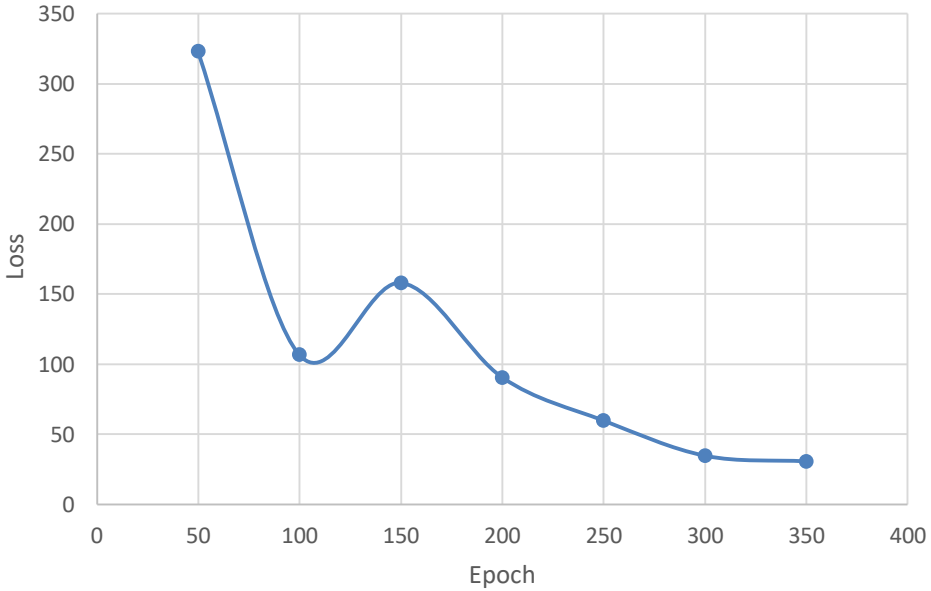


Figure 7 – Validation phase of LSTM

- The model convergence with respect to the input data after 350 epochs
- The percentage of average precision (OTB 100 dataset) at threshold value of 20 is 94.67%.

Figure 6 – Test on UR Fall Dataset

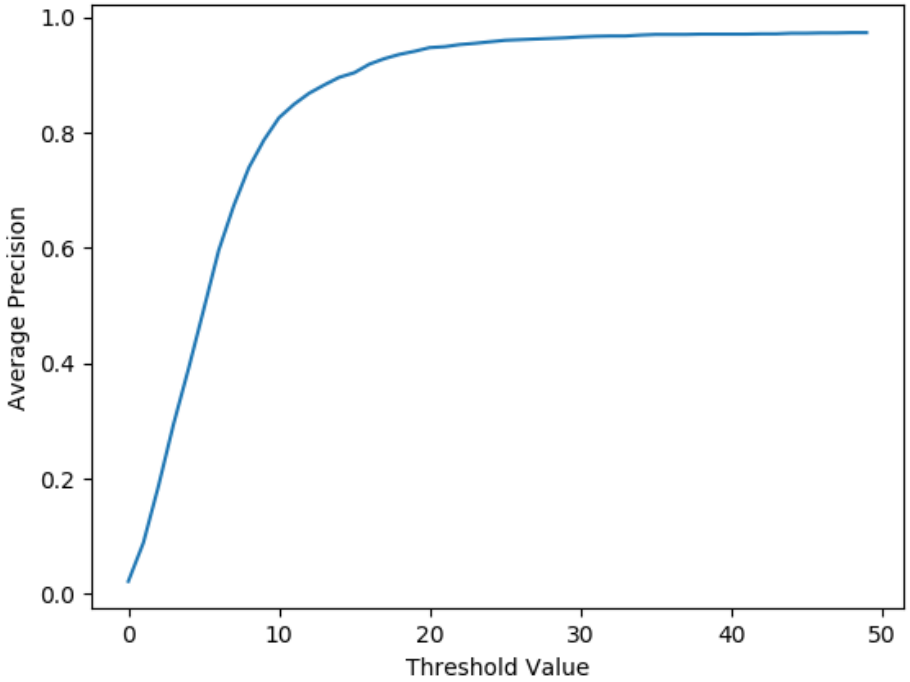


Figure 8 – Average precision in OTB 100 data set

Results (cont.)

The overall **Multiple Object Tracking Accuracy** was calculated using

$$MOTA = 1 - (M + WD + ID_{switch}) / (Obj_{gt})$$

where *M* → person misses, *WD* → wrong person detections, ID_{switch} → ID switches, Obj_{gt} → total persons in the entire video scene

Table 1 – Multiple Object Tracking Accuracy **(**MOTA)

Method	Correct Detects	Miss	Wrong Detects	ID switch	MOTA
CNN + LSTM	78.23%	12.2%	3.3%	7.5%	76.6%
CNN	77.1%	15.4%	7.01%	7.5%	70.1%
LSTM	78.96%	14%	8.1%	7.1%	70.8%

- The combination of appearance and motion similarity yields higher accuracy

Results (cont.)

Table 2 – Validation phase of VPFD

Epoch	Loss	Accuracy %
1	0.2937	87.42
2	0.1401	93.45
3	0.1051	96.52
4	0.0874	97.68
5	0.1211	95.20
6	0.0553	98.01

Table 3 – Comparison of methods based on accuracy

S. No.	Method	Accuracy %
1	Curvelets + HMM	96.88
2	Optical Flow + CNN	95.00
3	**HoG + LSTM (Proposed)**	**98.01**

- **VPFD Model Accuracy**: learning ability to **differentiate** between fall and non-fall sequences
- **Higher accuracy** by proposed method results due to **enhanced learning**

Summary

Conclusion

- A novel **multi-feature-based** person tracker, supported by an efficient **vision-based fall detection**
- The proposed system achieved 94.67% precision in **tracking** and 98.01% accuracy in **fall detection**
- The **fall detection module** (*HoG feature-based LSTM training network*) is relevant to the activities of **ITU-T Study Group 16** and Focus Group on Artificial Intelligence for Health (**FG-AI4H**)

Future work

- Detection of **different activities** apart from fall detection, and recognize and report in the cases of anomalies
- **Standardization** of the proposed scheme at **ITU-T SG16**

ITU KALEIDOSCOPE
ATLANTA 2019

Thank you