

# Activity Recognition Using First-Person-View Cameras Based on Sparse Optical Flows

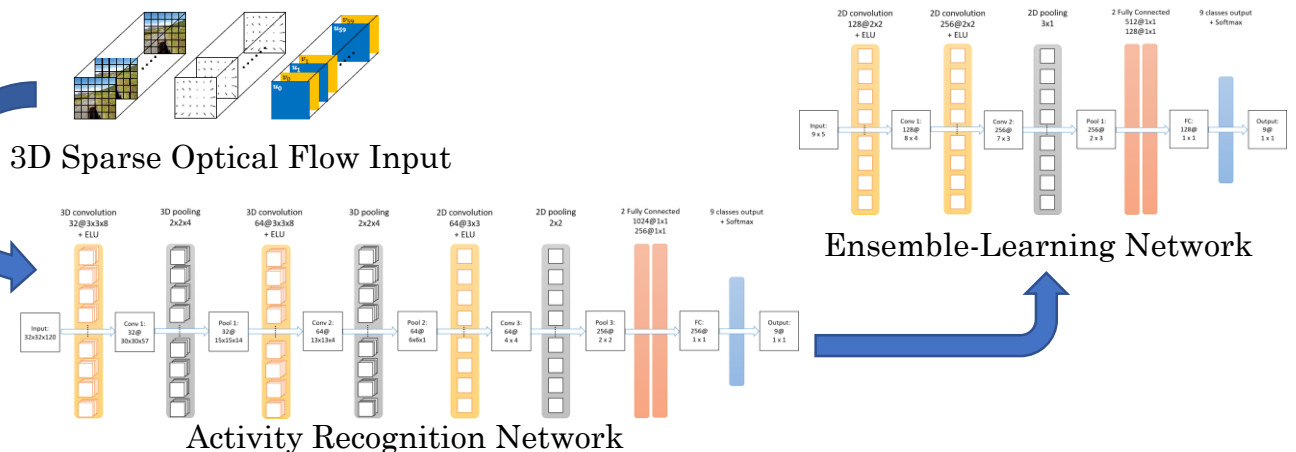
Peng-Yuan Kao, Yan-Jing Lei, Chia-Hao Chang, Chu-Song Chen,  
Ming-Sui Lee, and Yi-Ping Hung  
hung@csie.ntu.edu.tw



## Introduction

- Recent advances in wearable camera enable the feasibility of employing a first-person-view (FPV) camera to record the activities of a person in a whole day.
- We propose a **succinct and robust CNN** using **sparse optical flows**.

## Proposed Method



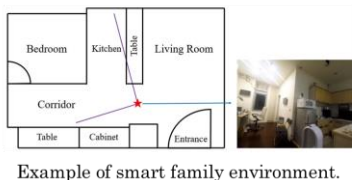
## Experiments

### Dataset

- We invited 21 people to film the daily living videos of 9 activities by using a chest mounted GoPro camera.
- Every subject recorded five minutes of each activity for a total of 45 minutes



Example of our chest mounted GoPro camera.



Example of smart family environment.



Example of Activities.

### Experimental Results

- In most of the activities, the recognition accuracy of our method is higher than that using the method in [Poleg 2016].
- The average accuracy of our method is 15% higher than that using the method in [Poleg 2016].

Activity	[Poleg 2016]	Ours
Walking	54%	<b>95%</b>
Eating	57%	<b>92%</b>
Watch TV	61%	<b>89%</b>
Sweeping	88%	<b>96%</b>
Use the phone	82%	77%
Reading a book	67%	<b>93%</b>
Use a pad	87%	74%
Take Medicine	91%	<b>95%</b>
Sleeping	91%	<b>100%</b>
Average Acc.	75%	<b>90%</b>

[1] Yair Poleg, et al., "Compact cnn for indexing egocentric videos," WACV, 2016.

## Conclusions

- The proposed method **achieves activity recognition accuracy of 90%**, which outperforms the previous method by 15%.