Team name	ADSC
Team leader name	Yong Pei
Team leader address, phone number and email	1 Fusionopolis Way, 08-10 Connexis North Tower Advanced Digital Sciences Center Singapore 138632 Phone: +65 85253212 Email: Pei.yong@adsc.com.sg
Rest of team members	Bingbing Ni Indriyati Atmosukarto
Team website URL (if any)	No

Title of the contribution	A Mixture of Heterogeneous Analyser for Action Detection
General method description	We proposed a mixture of heterogeneous analysers for action recognition and localization. Human key poses are first learned by middle-level discriminative part detection. Humans are detected and tracked into short sequences. Then, the detected discriminative key poses are integrated with the popular dense motion trajectories for target action detection. The combined detectors and the mutual contexts are learned by structural SVM.
References	[1] N. Dalal and B. Triggs. Histograms of Oriented Gradients for Human Detection. In CVPR, pp. 886-893, 2005. [2] H. Wang, A. Klaser, C. Schmid, and C. Liu. Action Recognition by Dense Trajectories. In CVPR, pp. 3169-3176, 2011. [3] T. Joachims, T. Finley and C. Yu. Cutting-plane training of structural SVMs. ML, Vol. 77, No 1, pp. 27 – 59, 1999 [4] B. Ni, Y. Pei, P. Moulin and S. Yan. Multilevel Depth and Image Fusion for Human Activity Detection. T-SMC-B, Vol. 43, No. 5, pp.

Describe data preprocessing techniques applied (if any)	Human tracking
Describe features used or data representation model (if any)	1. HOG (Histograms of Oriented Gradients). 2. Dense Trajectory feature.
Dimensionality reduction technique applied (if any)	No
Temporal clustering approach (if any)	No

Temporal segmentation approach (if any)	**N/A**
Gesture representation approach (if any)	HOG (Histograms of Oriented Gradients) + mined key poses
Classifier used (if any)	SVM (Support Vector Machine) and Structural SVM
Large scale strategy (if any)	N/A

Transfer learning strategy (if any)	N/A
Temporal coherence and/or tracking approach considered (if any)	Our human tracking is performed by simply connectting nearest human detection results (HOG-SVM) between neighbour frames into sequences.
Compositional model used, i.e. pictorial structure (if any)	Yes. The composition between key pose and the pooled dense motion trajectory feature is used.
Other technique/strategy used not included in previous items (if any)	N/A
Method complexity analysis	The major computational complexity includes: HOG key pose detection, dense trajectory extraction, and sliding window combinatorial detection

Qualitative advantages of the proposed solution	**The proposed method outperforms both dense trajectory method and key pose method, since it inherit advantages from both representations.**
Results of the comparison to other approaches (if any)	N/A
Novelty degree of the solution and if is has been previously published	Our solution uses a mixture of heterogeneous analyser for action recognition and detection, employing both key pose information and dense trajectory representation as well as their context. It hasn't been published previously.

Language and implementation details (including platform, memory, parallelization requirements)	Matlab, Visual Studio C++, OpenCV
Human effort required for implementation, training and validation?	1 software engineer, 4 months
Training/testing expended time?	Training: 3-5 days for labelling samples and training models. Testing: 900-1000 frames sequence will take approximately 30 minutes to run.
General comments and impressions of the challenge	A good testbed for researchers to evaluate their action detection algorithms.