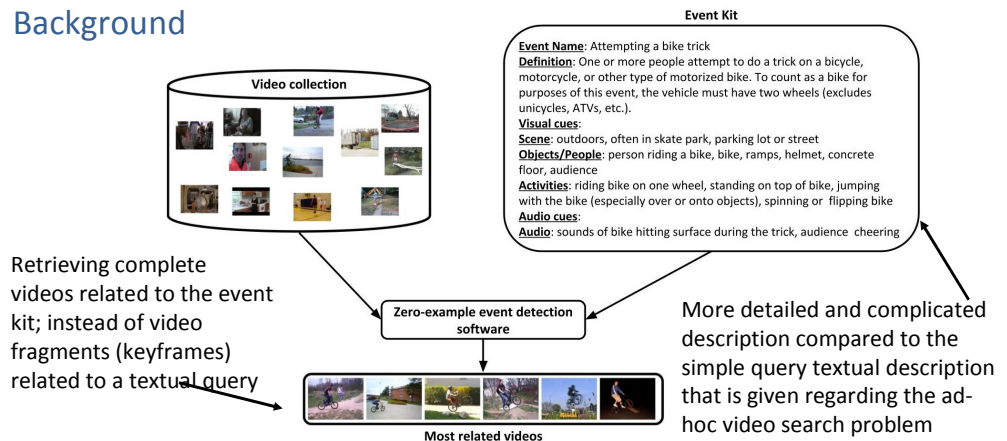


Problem and motivation

- Zero-example event detection:** retrieving, from a large video collection, videos that are related to a given event textual description, where training examples for this event are not given
- Typical solution:** transforming both the event textual description and the available videos into concept-based representations using simple concept terms retrieved from a pre-defined concept pool
- Motivation:**
 - Augmenting the simple concept terms with additional words and phrases
 - Exploiting the appropriate number of concepts per event

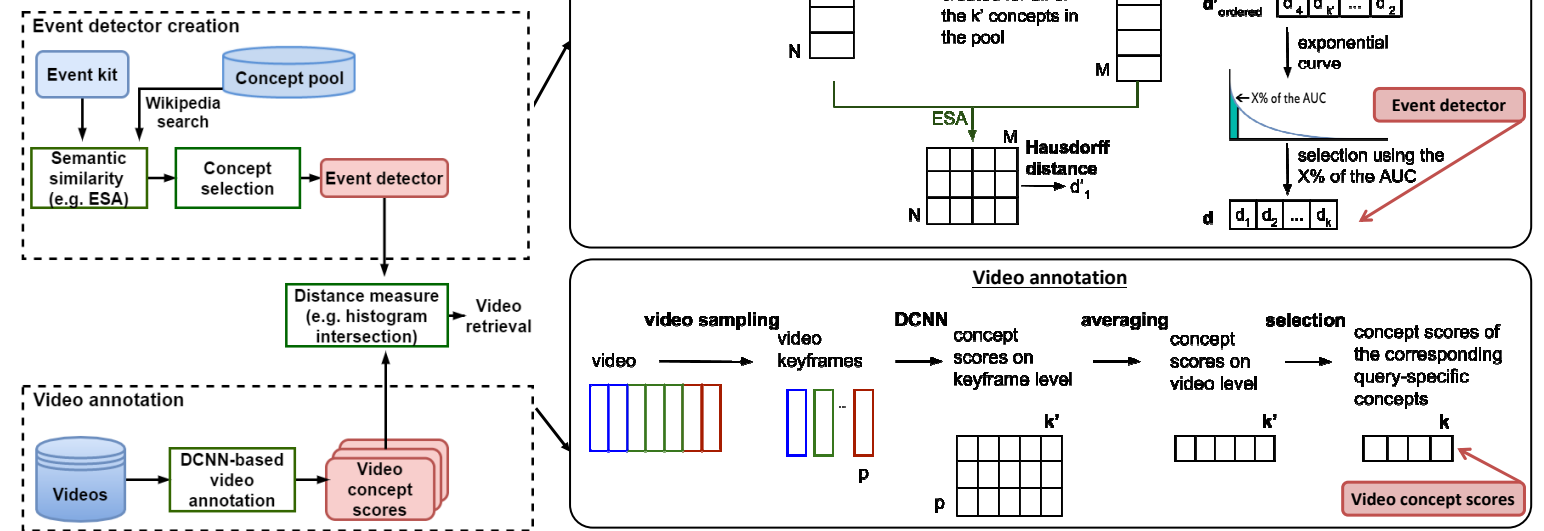
Background



Proposed method

Method outline

- Event detector creation: ELM, CLM, concept number selection
- Video annotation and retrieval: DCNNs and histogram intersection distance



Experimental results

- TRECVID MED14TEST dataset
- 25.000 videos
- 20 Pre-Specified Events (E021-E040)
- 13.488 semantic concepts
 - 12.988 ImageNet concepts
 - 500 event-related concepts (EventNet)
- Mean Average Precision (MAP)

A) Parameter selection for ELM, CLM

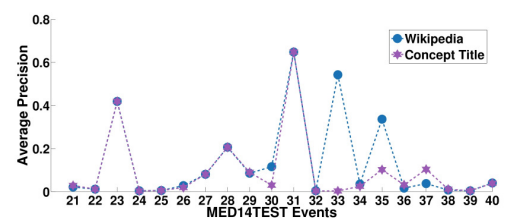
- (a) Comparison between different types of ELM

ELM type	Event Title	Visual	Audio-Visual
MAP	0.091	0.122	0.133

- (b) Comparison between different types of CLM

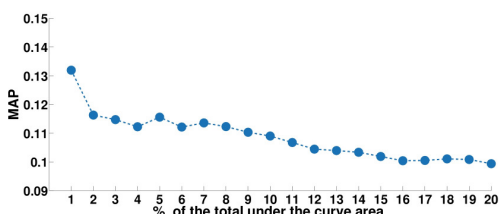
CLM type	Concept Title	Wikipedia
MAP	0.092	0.133

- (c) The performance for different types of CLM for the 20 MED14TEST events

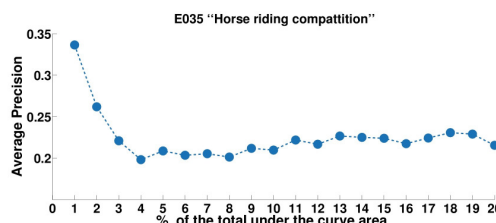


B) Concept number selection

- (a) The overall performance of different quota of under the curve area (AUC)



- (b) The performance of different quota of under the curve area (AUC) for the E035 event



C) Comparisons

- MAP between different zero-example event detection systems

Method	MAP
AutoSQGSys [13]	0.115
Concept Bank [1]	0.129
Tzelepis et al. [18]	0.119
Proposed System	0.133