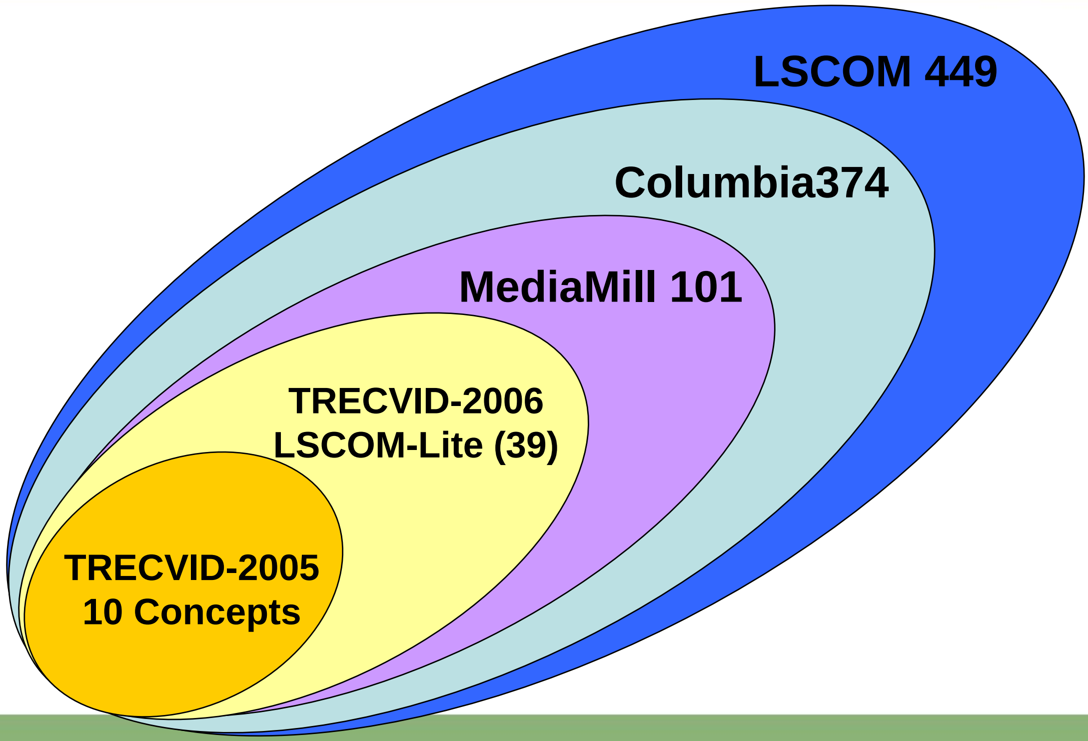
New Challenges in Semantic Concept Detection

M.-F. Weng, C.-K. Chen, Y.-H. Yang, R.-E. Fan,
Y.-T. Hsieh, Y.-Y. Chuang, W. H. Hsu, and C.-J. Lin
National Taiwan University

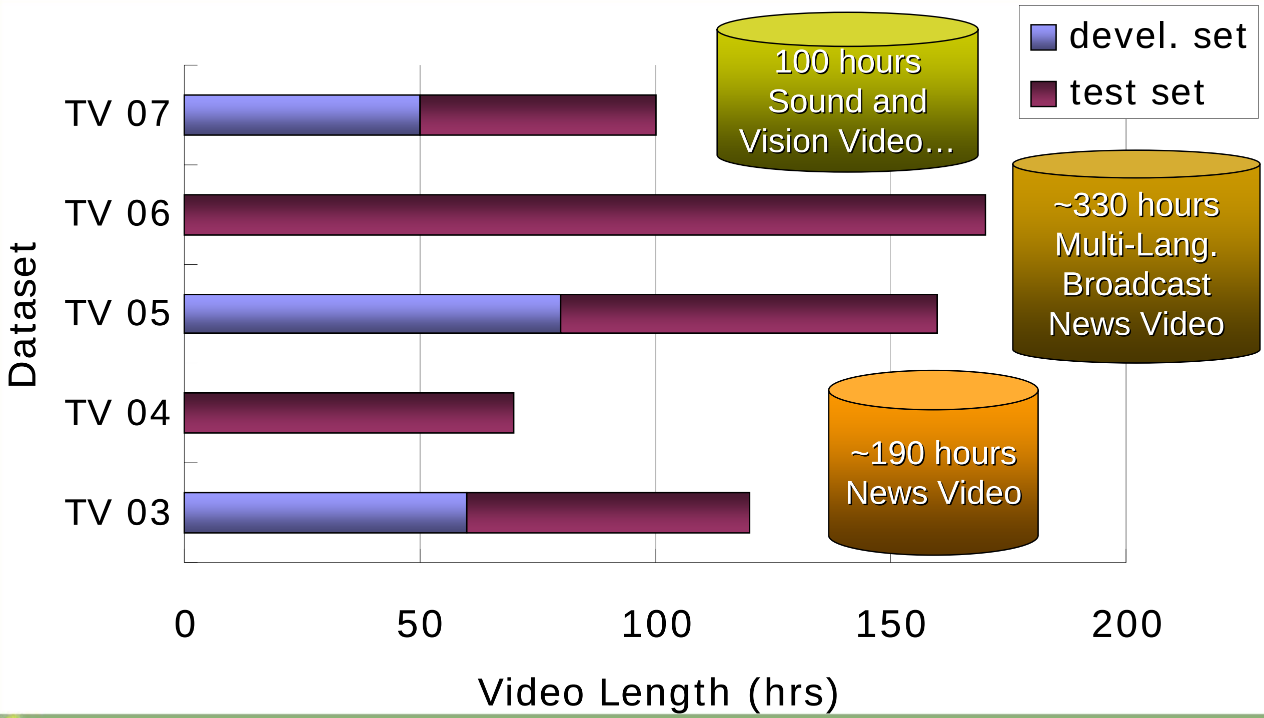
Preliminaries for Semantic Concept Detection

- What are preliminaries for building a semantic concept detection system?
  - A lexicon of well-defined concepts
  - Training resources
    - Video data
    - Annotations
    - Features
  - Tools
    - Tagging or labeling tools, (e.g., CMU and IBM tools)
    - Feature extractors
    - Machine learning tools, (e.g., LIBSVM)
    - Semantic concept detection tailored tools

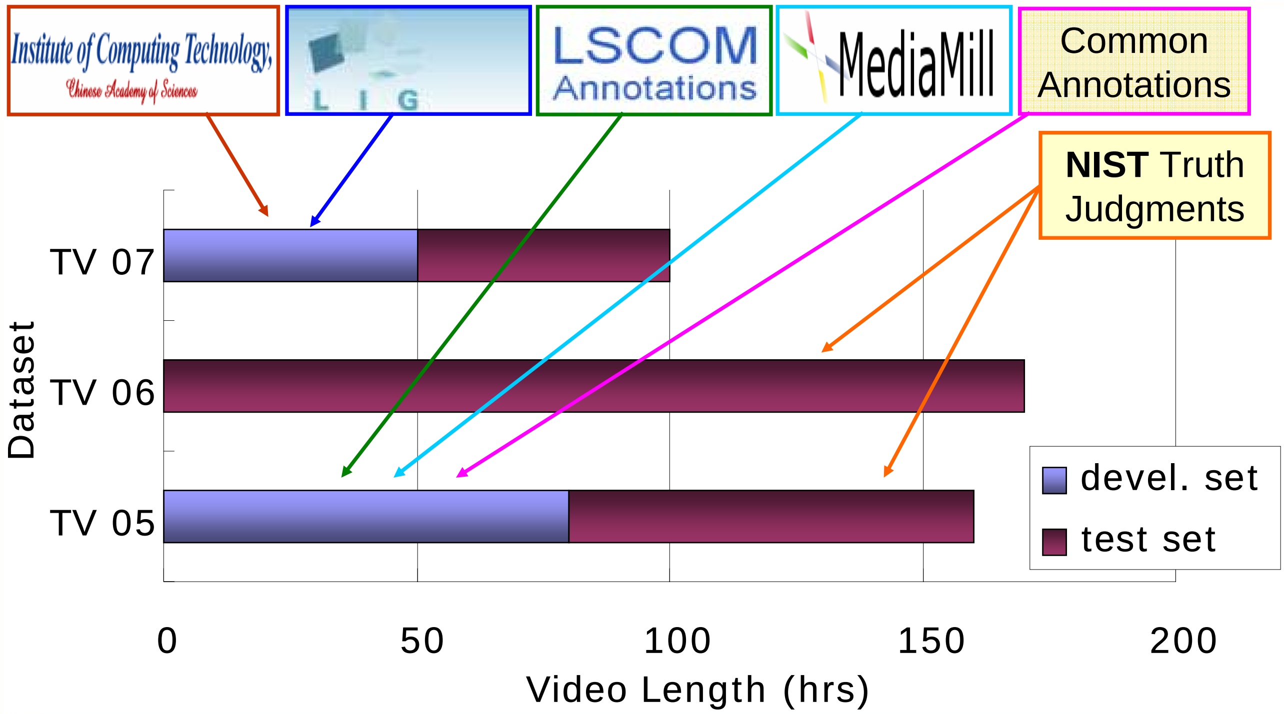
Semantic Concepts



Video Data Sets



Annotations



Features, Detectors, Scores

	Features	Detectors	Scores
MediaMill Baseline	•Visual feature •Text feature	5 sets of 101 classifiers	Scores of TV 05/06 dataset
Columbia374	•EDH •GBR •GCM	Columbia 374 detectors	Columbia374 scores of TV 06/07 dataset
VIREO-374	•Color moment •Wavelet texture •Keypoint feature	VIREO-374 detectors	Scores of TV 07 dataset

Available Resources

Well-defined concepts		○
Training Resources	Video data	○
	Annotations	○
	Features	○
Tools	Tagging Tools	○
	Feature Extractors	×
	Machine Learning Tools	○
	Tailored Tools	×

- Concept definition is sufficient
- Training resources are plentiful
- No feature extractors and tailored tools available

The New Challenges

- **Challenge 1 : Easy and Efficient Tools**
  - ***L*** datasets, ***M*** concepts, ***N*** features, imply ***L*M*N*** classifiers
  - Each classifier has to consider many parameters
  - **Time seems very limited** to validate each parameter and to train all classifiers
- **Challenge 2: Resource Exploitation or Reuse**
  - Resources are precious
  - Existent resources are **potentially useful** for new dataset
  - Plentiful resources have **not been fully utilized**

Facing the New Challenges

- **Challenge 1:**
  - Extended LIBSVM to improve training efficiency
  - Developed an efficient and easy-to-use toolkit tailored for semantic concept detection
- **Challenge 2:**
  - Reused classifiers of past data to improve accuracy by late aggregation
  - Exploited contextual relationship and temporal dependency from annotations to boost accuracy

Facing the New Challenges

- **Challenge 1 : Easy and Efficient Tools**
  - Extended LIBSVM to improve training efficiency
  - Developed an efficient and easy-to-use toolkit tailored for semantic concept detection
- **Challenge 2**
  - Reused classifiers of past data to improve accuracy by late aggregation
  - Exploited contextual relationship and temporal dependency from annotations to boost accuracy

A Tailored Toolkit

- We extended LIBSVM in three aspects for semantic concept detection:
  - Using dense representations
  - Exploiting parallelism of independent concepts, features, and SVM model parameters
  - Narrowing down parameter search to a safe range
- Overall, training time of our baseline was approximately reduced **from 14 days to about 3 days**

Facing the New Challenges

- Challenge 1
  - Extended LIBSVM to improve training efficiency
  - Developed an efficient and easy-to-use toolkit tailored for semantic concept detection
- **Challenge 2 : Resource Exploitation or Reuse**
  - Reused classifiers of past data to improve accuracy by late aggregation
  - Exploited contextual relationship and temporal dependency from annotations to boost accuracy

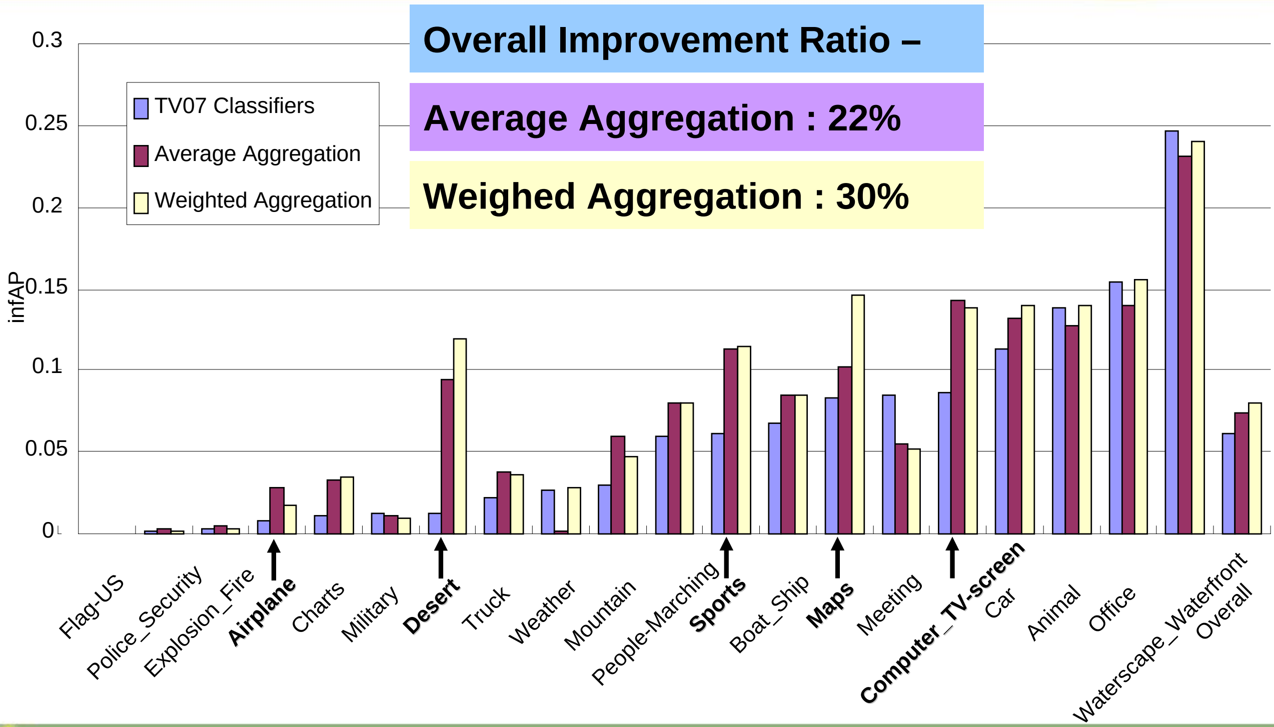
Reuse Past Data

- **Early aggregation**
  - Must re-train classifiers
  - Cause considerable training time
- **Late aggregation**
  - Simple and direct
  - May be biased

Late Aggregation

- We adopt late aggregation to reuse existent classifiers by two strategies:
  - Equally Average Aggregation
    - Simply average the scores of past and newly trained classifiers
  - Concept-dependent Weighted Aggregation
    - Use concept-dependent weights to aggregate classifiers

Aggregation Benefits



Facing the New Challenges

- Challenge 1
  - Extended LIBSVM to improve training efficiency
  - Developed an efficient and easy-to-use toolkit tailored for semantic concept detection
- **Challenge 2: Resource Exploitation or Reuse**
  - Reused classifiers of past data to improve accuracy by late aggregation
  - Exploited contextual relationship and temporal dependency from annotations to boost accuracy

Observation in Annotations

A sequence of video shots

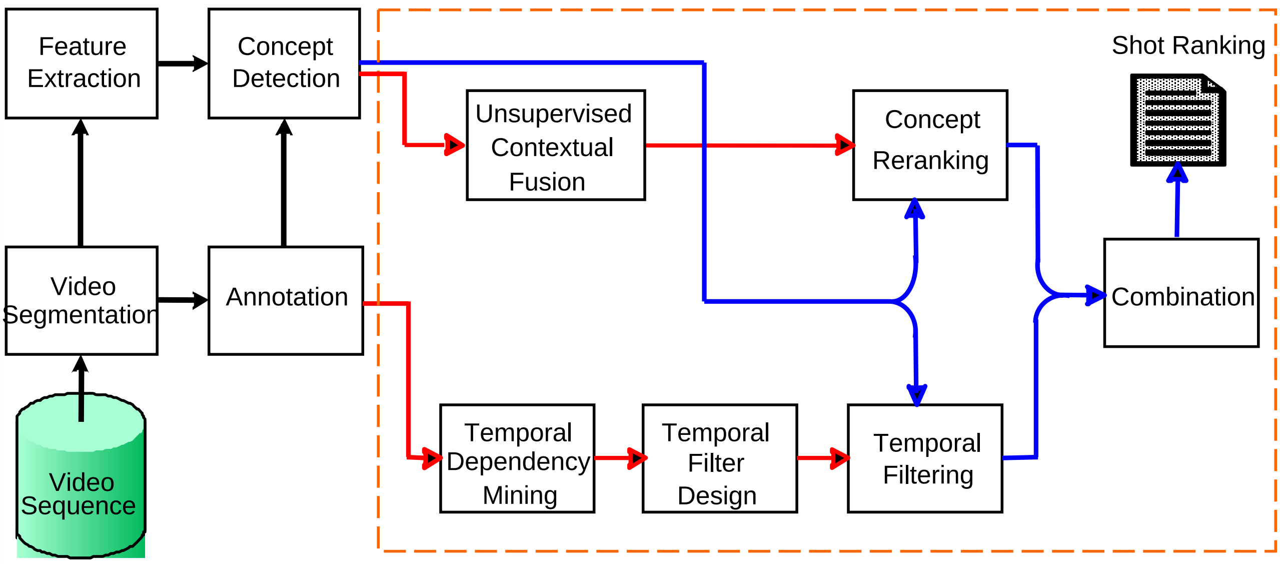
car	1	1	1	1	1	1	1	1
outdoor	1	1	1	1	1	1	1	1
urban	1	1	0	1	1	1	1	1
building	1	0	0	1	1	1	1	0
sky	0	1	1	1	0	1	1	0
people	0	0	0	1	0	1	1	0

A lexicon of concepts

Temporal Dependency

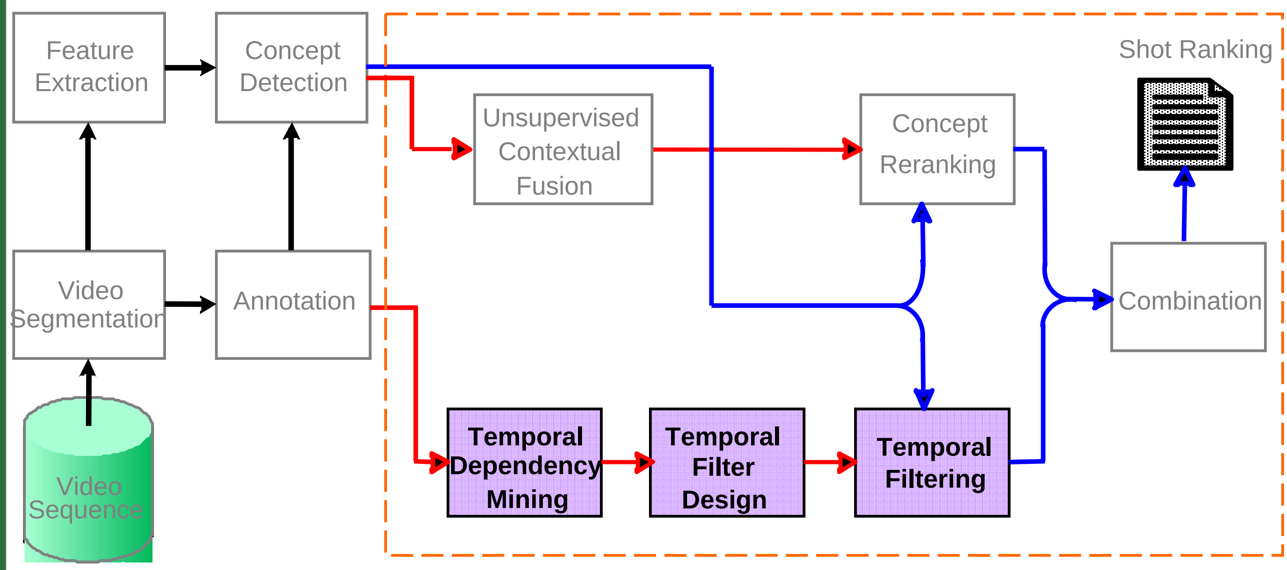
Contextual relationship

Post-processing Framework



Detecting phase: **Mining phase:** **Processing phase:**

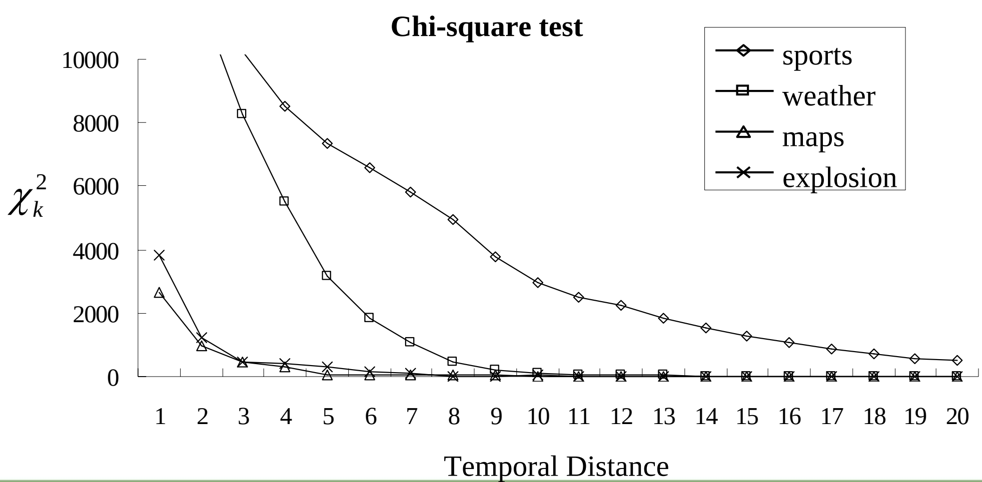
Temporal Filtering



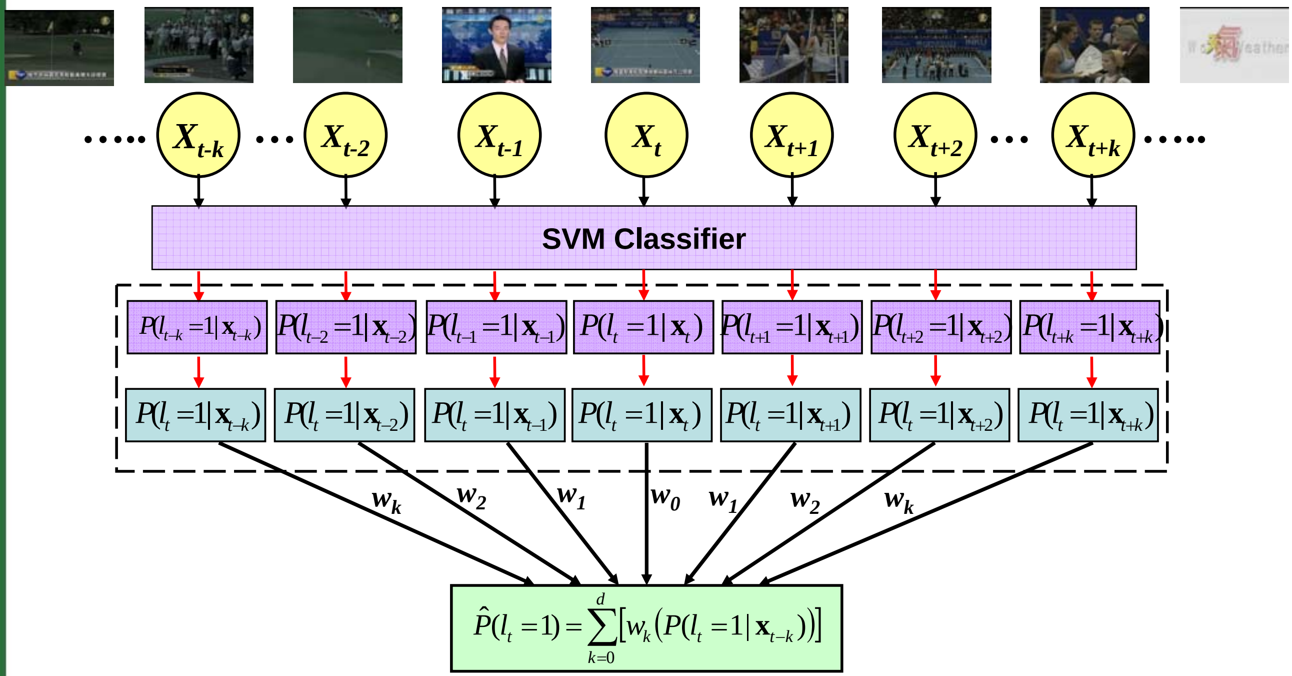
Detecting phase: Mining phase: Processing phase:

Temporal Dependency

- Different concepts have different levels of dependency at different temporal distance
  - E.g., *sports, weather, maps, explosion*



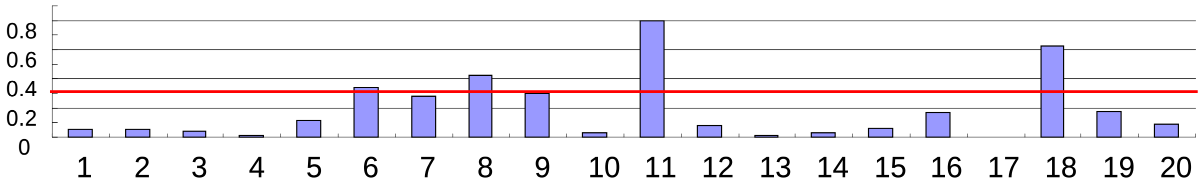
Temporal Filter



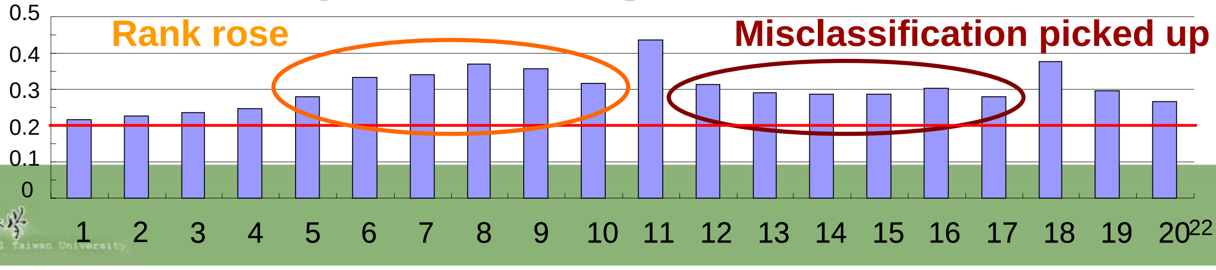
Filtering Prediction

- A sequence of shots for predicting *sports*

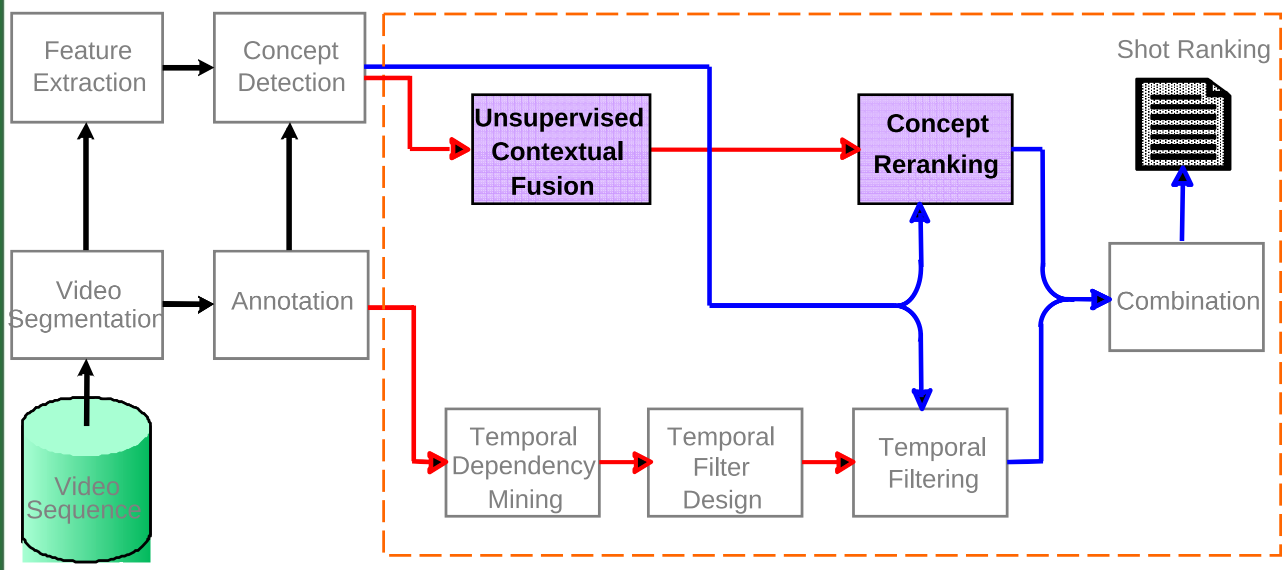
– **Classifier prediction results**



– **After temporal filtering**



Concept Reranking

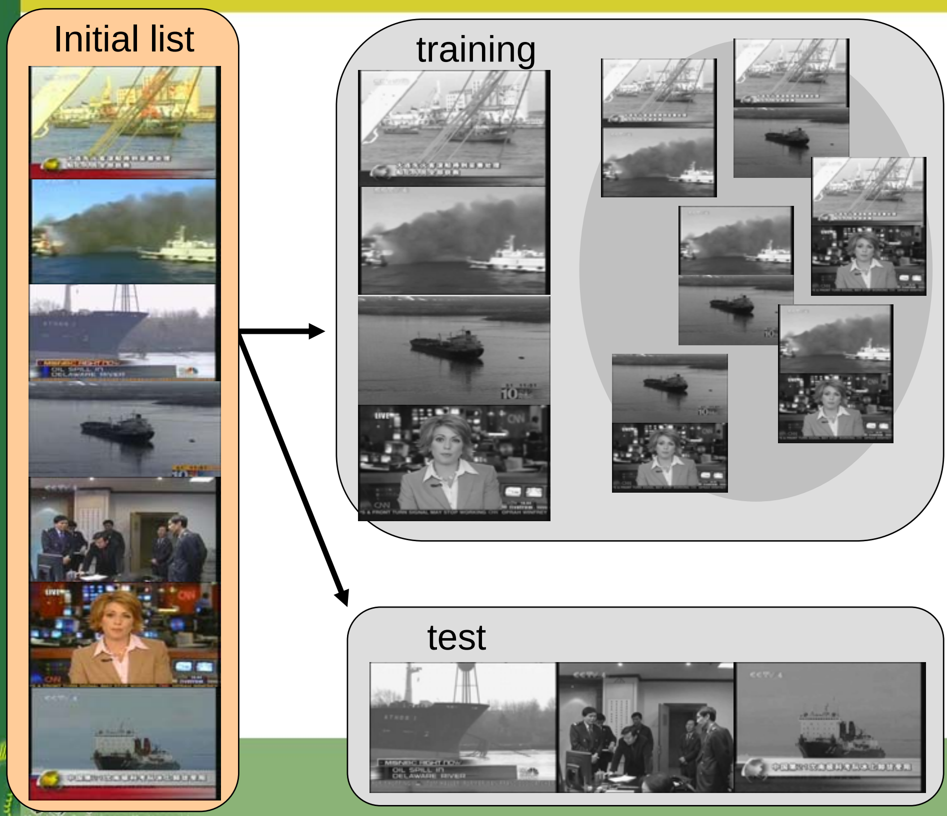


Detecting phase: **Mining phase:** **Processing phase:**

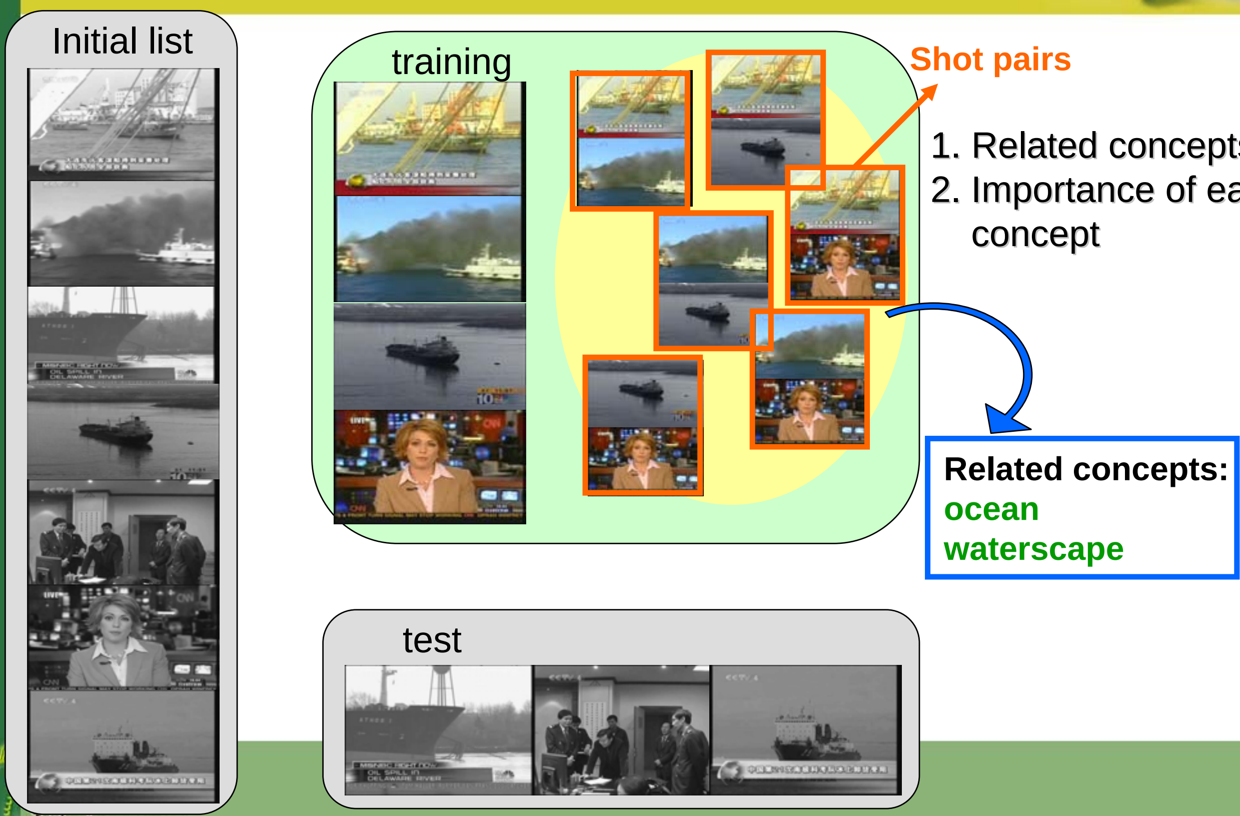
Initial ranking list produced by a baseline method
Target concept: 'boat' (search or detection)

Initial list

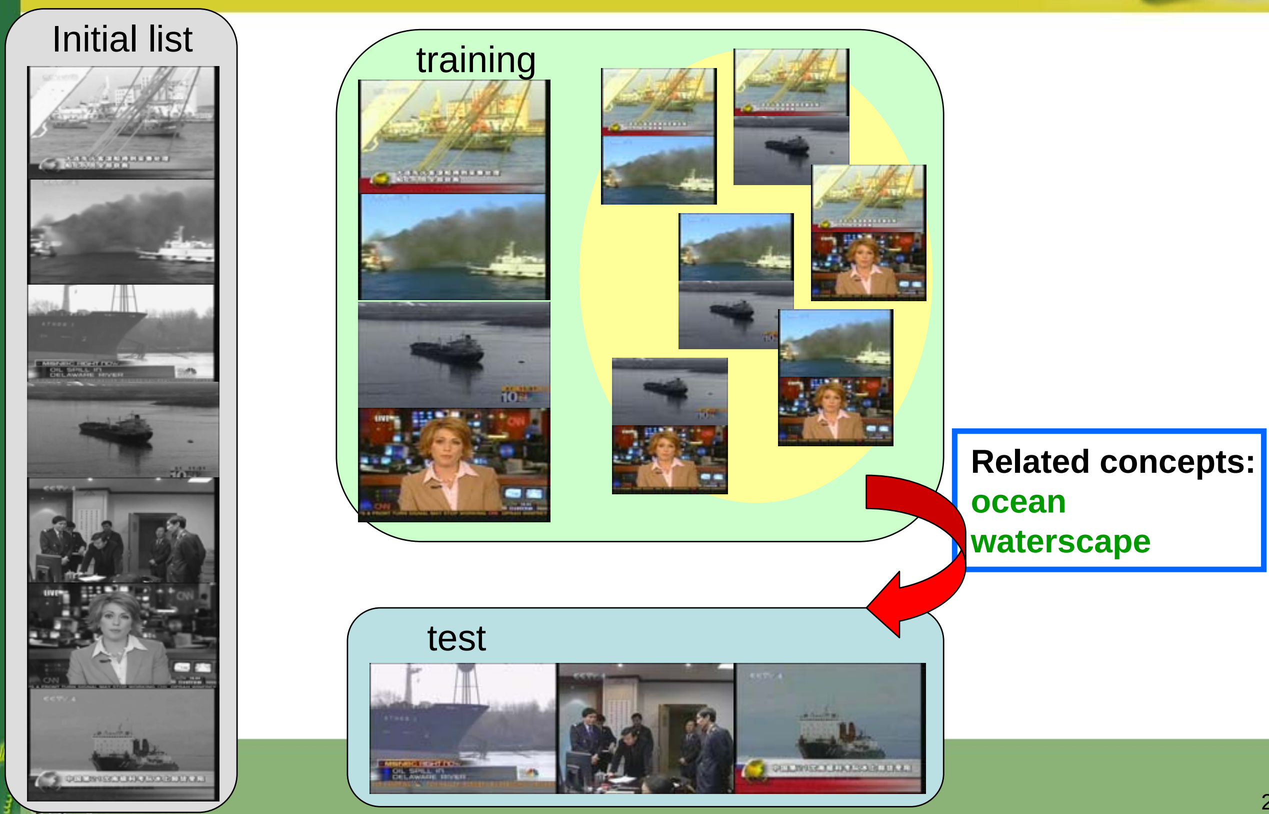
Step 1. Randomly split to training and test sets



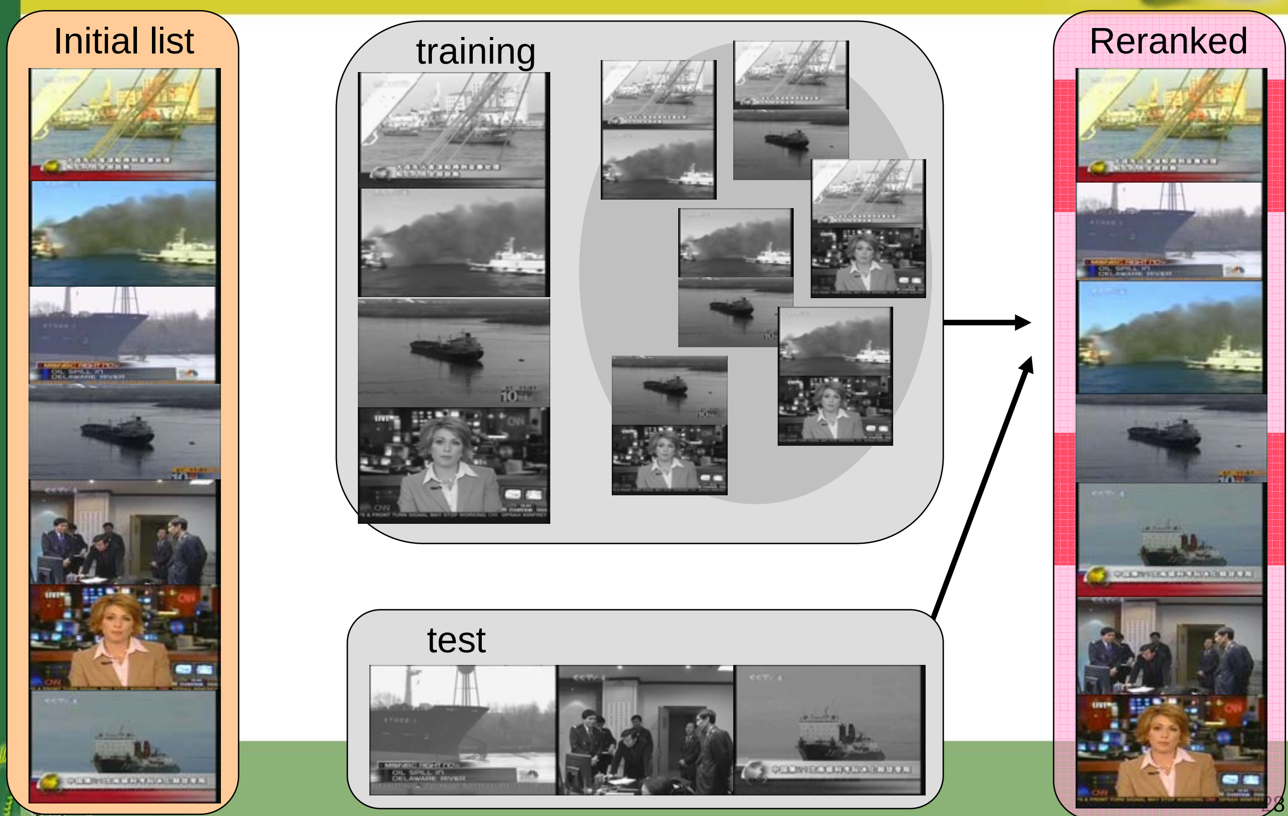
Step 2. Learn to maintain the ranking orders



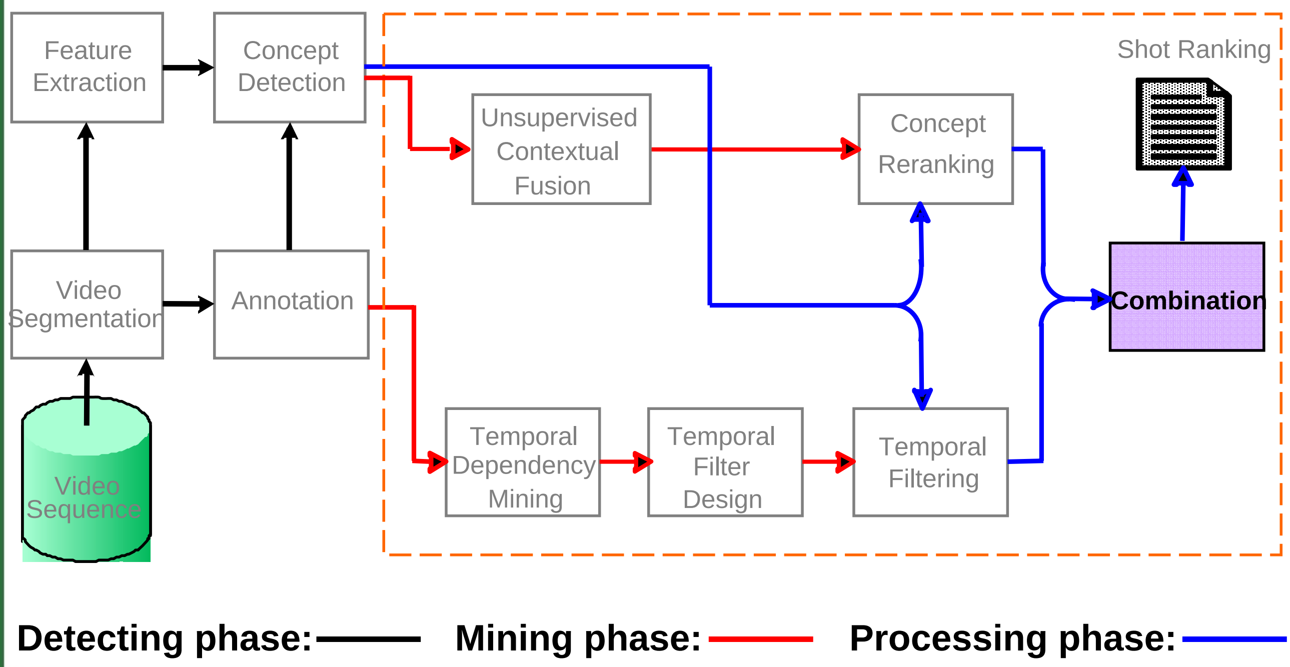
Step 3. Context fusion on the test data



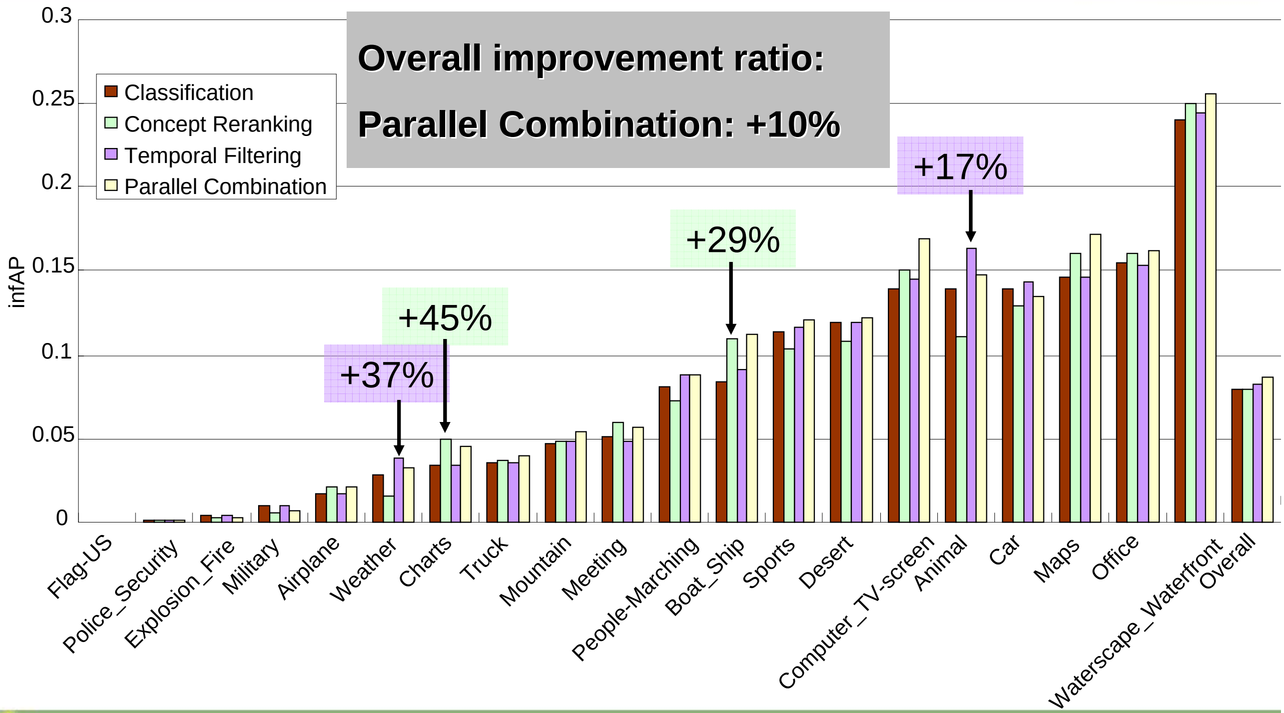
Step 4. Merge



Combination



Post-processing Benefits



Conclusion

- We reduce the training time of detectors by using a tailored toolkit for semantic concept detection
- The proposed aggregation methods reuse the classifiers of past data and can boost the detection accuracy
- Our post-processing approaches exploit existent resource and can further improve detection results

Thank You For Your Attention