HELSINKI UNIVERSITY OF TECHNOLOGY
Department of Information and Computer Science

Techniques for Image Classification, Object detection and Object Segmentation

Ville Viitaniemi, 2008-9-12
paper co-authored with
Jorma Laaksonen

Ville.Viitaniemi@tkk.fi

Talk contents

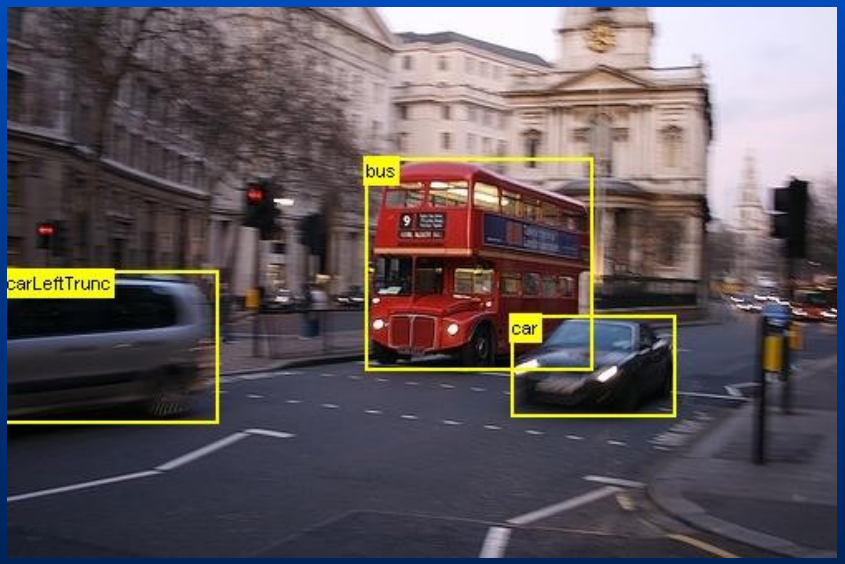
- aim: document and discuss the system we used for participating the VOC 2007 image analysis benchmark contest
  1) The VOC 2007 benchmark
  2) Our approach
     - image classification
     - object detection
     - object segmentation
  3) Results & conclusion

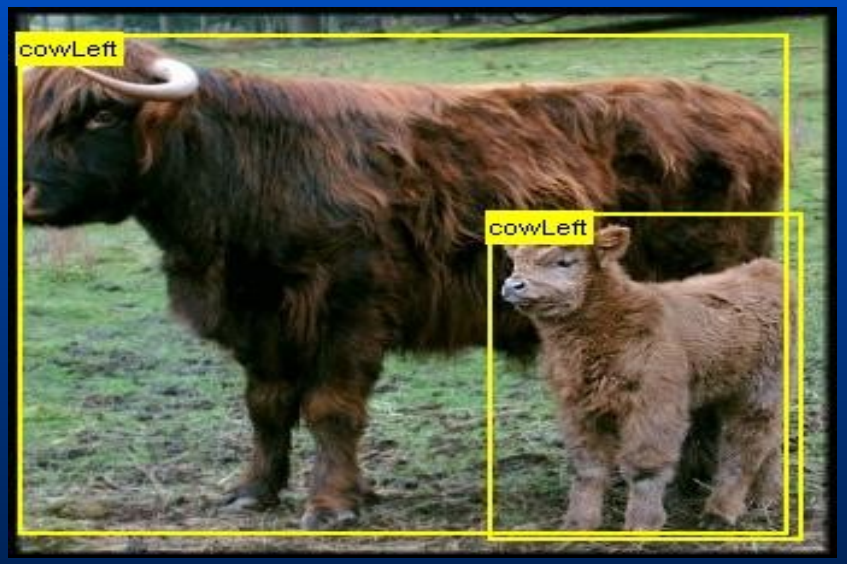
Ville.Viitaniemi@tkk.fi

VOC 2007 benchmark competition

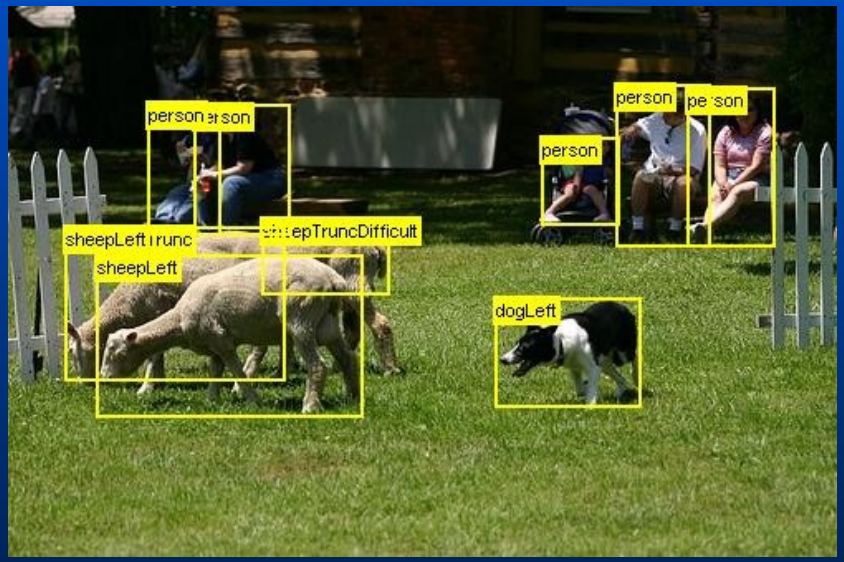
- Visual Object Classes challenge: a yearly benchmark competition organised by PASCAL Network of Excellence
  - nowadays ~ 20 participating groups
- ~ 10k photographic images
- annotated for 20 classes of objects
  - aeroplane, bicycle, bird, bottle, boat, bus, car, cat, chair, cow, dining table,...

Ville.Viitaniemi@tkk.fi





Ville.Viitaniemi@tkk.fi



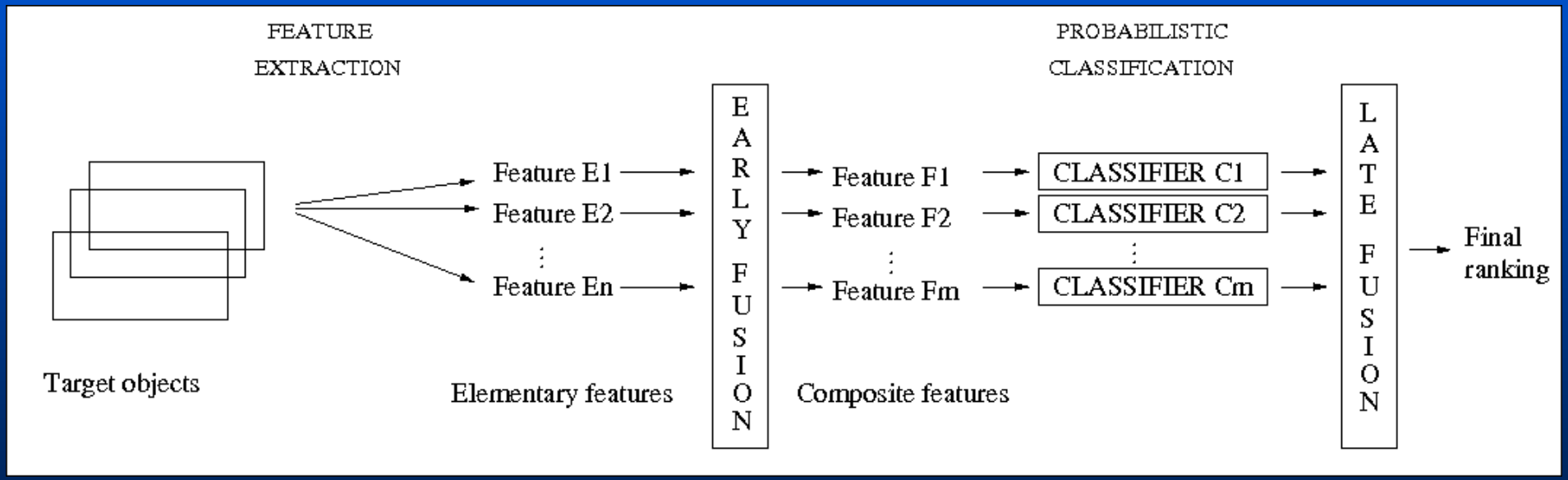
Ville.Viitaniemi@tkk.fi

VOC 2007 tasks

- we consider three (out of four defined) supervised tasks:
  1) image classification: determine whether or not a certain class (e.g. cow) of object is present in the images
  2) object detection: find bounding boxes of objects
  3) object segmentation: find pixel-level segmentation masks for partitioning the area of the images among the 20 object classes

Ville.Viitaniemi@tkk.fi

Framework for visual object classification



Ville.Viitaniemi@tkk.fi

VOC 2007 classification task

- framework used with images as target objects
- large set of global image features -> 141 composite features
- 141 separate SVM classifiers
- SVM-based fusion

Ville.Viitaniemi@tkk.fi

Detection task

- segmentation on images with a region merging type segmentation algorithm
- for each image segment s

$$p(I_\mathrm{s} = 1|c_\mathrm{i}, c_\mathrm{s}) \approx p(I_\mathrm{s} = 1|I_\mathrm{i} = 1, c_\mathrm{s})p(I_\mathrm{i} = 1|c_\mathrm{i}).$$

- latter probability directly from classification task
- the first probability calculated using the framework with image segments as target objects
  - feature set somewhat reduced but otherwise similar to classification task

Ville.Viitaniemi@tkk.fi

Segmentation task

- the bounding boxes from detection task automatically converted to segmentation masks

Ville.Viitaniemi@tkk.fi

Results and conclusions

- Classification: relatively good (4th out of 19)
  - our individual features probably not as good as the best ones,
  - the fusion aspect of the system more developed
- detection: quite modest, although not hopelessly bad
- segmentation: best among the participants
  - solutions of the participants obviously not yet well-developed
  - why our segmentations better than automatic segmentations converted from better detections?
  - segmentation has been markedly improved [3] after the challenge by re-segmenting our bounding boxes

Ville.Viitaniemi@tkk.fi

- A more detailed version of the paper available on-line as a technical report [4] at http://www.cis.hut.fi/projects/cbir/