Image Segmentation & Classification

Outline
- Image segmentation
  - Clustering
  - Supervised
  - Semi-supervised & active learning
- Image classification
  - “Holistic”
  - Combining classifiers
  - Multiple instance learning

Image Segmentation
- Label pixels as coming from two or more classes / divide images in two or more regions
- Core technique in image analysis
  - Radiological images, computer aided diagnosis and detection, satellite images, automated sorting machines, image guided surgery, geological analyses, and so on...

Image Segmentation via Clustering
- Possible preprocessing
  - E.g. for patch-based supervised labeling
  - Or interactive labeling
- Divide image in “homogeneous” patches
- Other unsupervised approaches may be rephrased in similar terms, or are at least related
  - E.g. normalized cuts, [fuzzy] connectedness, region growing, watershed segmentation, etc.

E.g. Color Image Segmentation
- PCA projection from RGB to 2D subspace

Clustering Result
- Colors are cluster means
- How to attach meaning to clusters?

Other unsupervised approaches may be rephrased in similar terms, or are at least related
- E.g. normalized cuts, [fuzzy] connectedness, region growing, watershed segmentation, etc.
**From Clusters to Classes**

- Class names [meaning] may be assigned to clusters based on our interpretation of classifier outputs
- Raw per-cluster outputs of trained nmc model

**Change in Number of Clusters**

<table>
<thead>
<tr>
<th>Fields</th>
<th>Road</th>
<th>Sky</th>
<th>Objects</th>
</tr>
</thead>
</table>

**Some Remarks**

- To find right objects they have to be rather homogeneous
  - Other features can be used, but is often problematic
  - Generally, over or under segmentations will occur
  - Unsupervised segmentation may be useful if
    - Expert labeling is very expensive, only coarse segmentation is necessary, too many possible objects, too much variability, as preprocessing...

**E.g. Interactive Segmentation**

- Liver segmentation

**Supervised Segmentation**

- Potentially, considerably more powerful than unsupervised methods
- Obviously, need for examples / training data
- Many possibilities, many approaches, many published methods; all variations on few themes?
  - Active shape models, active appearance models, random field, pixel-based classification, patch classification

**Pixel-Based Methods**

- Direct application of supervised pattern recognition techniques
- "Avoid" difficult modeling
- Simple idea
  - Extract features per pixel [filter bank, raw intensities, etc.]
  - Train classifier and classify pixels based on these features
- Performs state of the art on several problems
E.g. Retinal Vessel Segmentation

Pixel Classification Setup

- Experiments on 40 images from DRIVE database
  [www.isi.uu.nl/Research/Databases]
- Features: Gaussian filters
  - Derivatives up to order 2 at scales 1, 2, 4, 8, 16 and original green plane value,
  i.e., 31 features per pixel
- 30NN classifier [also tested LDC and QDC]

Pixel Classification Setup

Results / ROC

Another E.g.

- T->B, L->R: orig., man., ASM, AAM, PC, PC+

Another E.g.

- PC performs very well
- PC+ outperforms 2nd human observer on lungs
- Classifier combination performs best
- Which approach should be used where?

Some Remarks

- Supervised methods can perform very well
- How complex should segmentation approach be?
- Choice of features important and could [still] be difficult
  - Solve through feature extraction and selection, dimensionality reduction
- Unclear how much training data needed / level of further improvement possible
- Training data remains serious bottleneck
Opportunities [and Challenges]

• Semi-supervised learning
  • Possibly not hard to collect loads of image data
  • Mostly used, however, in interactive setting

• Active learning
  • Intelligent, adaptive querying
  • May need significantly fewer labeled examples for training

Semi-Supervised Segmentation

• E.g. interactive, via graph cuts
  \[ E(f) = \frac{1}{2} \sum_{i,j} w_{ij} |f(i) - f(j)| + \frac{1}{4} \sum_{i,j} w_{ij} (f(i) - f(j))^2 \]

Remarks

• Semi-supervised learning employed in image segmentation, but not part of learning process
  • Segments are done on per-image basis; no “bootstrapping” takes place
  • Semi-supervised training / testing loop seems good idea
  • How do we know problem is due to current labeling and not, say, choice of features?
  • More problems?
  • However, have not seen this fully functional yet...

Active Learning

• Equip learner with possibility to present one or more objects that operator should [re]label
  • Preferably, objects picked such that they are the most informative to the learner
  • Can e.g. be measured through uncertainty or stability
  • Exploration and exploitation may play a role
  • Seems very valuable technique, but needs even more research than semi-supervised...

Image Classification

• Every single image is being labeled as belonging to a certain class

• Possible approaches considered here
  • “Holistic”
  • Combining classifiers
  • Multiple instance learning

“Holistic” Image Classification

• Consider every image as a feature vector on which classifiers can be trained

• Possible preprocessing
  • Subsampling
  • Histogram equalization
  • [De]blurring
  • Cropping
"Holistic" Image Classification

- Obviously, feature dimensionality very large, so often use feature extraction / reduction strategies
  - Mostly linear
  - LDA, PCA
- Other option: Regularization
  - Little study into image-specific regularization techniques

Face Recognition

- Holistic approach popular and rather successful
  - Lot of research in feature extraction / dimensionality reduction techniques
- Error on notorious "ORL database" [40 subjects with 10 92×112 face images each] close to zero

Face Recognition

- FERET database
- Combining LDA or PCA with nearest neighbor
- 93.2% vs. 80.0%

Combining Classifiers Approach

- Useful if possible to identify object or person based on local [image] features $f$
  - $I$ being the image, $c$ the class label, and assuming independence given $c$, we have
    $$ p(c|I) = p(c|f_1, \ldots, f_N) $$
    $$ \propto p(f_1, \ldots, f_N|c)p(c) $$
    $$ = p(c) \prod p(f_i|c) $$
- Identification of $I$ goes through $f$s
- So, if our classifier can provide posteriors for these local features...

Some Remarks

- Approach is product classifier combination
- May be problems with product
  - Multiplication of a lot of small numbers
  - E.g. if every pixel represented by feature vector $f$
- Obviously, we might
  - Substitute product for other combiners
  - Deal directly with local class labels, not probs
Detecting Changes in Mammograms

- Aims:
  - Separation of hormone replacement therapy (HRT) group from placebo
  - Detection of aging effects
  - [Ultimately: utilization in cancer risk quantization]

Main Results [for Completeness]
- Method can quantify both age-related effects and effects caused by HRT
- Age effects are significantly detected
  - Standard methodologies fail
- Separation of HRT subpopulations is comparable to best methodology
  - Latter is interactive

Also for Completeness
- I mention multiple instance learning (MIL)
- Also used for the classification of images
  - Main assumption: image label is + iff there is at least 1 feature vector labeled +
    - Seems reasonable, for example in detection of diseases
    - "Combining with max operator"
    - Seems sensitive to mistraining / mislabeling

Axis-Parallel Rectangles
- Classical approach to MIL [not necessarily for images]
- Problem on the right
  - Closed are +
  - Open are –
- Construct box that at least includes one + from every positive instance and excludes as many –

Also for Completeness
- I mention multiple instance learning (MIL)
- Also used for the classification of images
  - Main assumption: image label is + iff there is at least 1 feature vector labeled +
    - Seems reasonable, for example in detection of diseases
    - "Combining with max operator"
    - Seems sensitive to mistraining / mislabeling

Axis-Parallel Rectangles
- Construct box that at least includes one + from every positive instance and excludes as many –
  - Difficult problem...
  - Generally, don’t think it has been convincingly demonstrated that MIL is the way to approach image classification...

OK solution?
References

- Olsen and Nielsen, Multi-Scale Gradient Magnitude Watershed Segmentation, 1997
- Niemeijer et al, Comparative study of retinal vessel segmentation methods on a new publicly available database, 2004
- Loog and Van Ginneken, Segmentation of the posterior ribs in chest radiographs using iterated contextual pixel classification, 2004

References

- Boykov and Funka-Lea, Graph Cuts and Efficient N-D Image Segmentation, 2006
- Roth, Face Recognition
- Raundahl et al, Automated Effect-Specific Mammographic Pattern Measures, 2008
- Loog and Van Ginneken, Static posterior probability fusion for signal detection, 2004
- Dietterich et al, Solving the multiple instance problem with axis-parallel rectangles, 1997