**Photo Annotation Task**

A multi-label classification problem

**Visual Information**

- Feature extraction:
  - ColorDescriptor software
  - Point detectors: Harris-Laplace, Dense Sampling
  - Descriptors: SIFT, HSV-SIFT, Huesift, OppenheimSIFT, C-SIFT, ngSIFT and RGB-SIFT
  - One multi-label dataset for each combination of detector-descriptor (14 in total)

- Learning method:
  - Binary Relevance (BR) coupled with Random Forest
  - Instance weighting to deal with class imbalance

**Textual Information**

- Feature extraction:
  - Flickr user tags of the training images were used as the initial vocabulary
  - Stemming and stop-word removal gave 27K stems approx.

- Learning method:
  - Ensemble of Classifier Chains (ECC) coupled with Random Forest

**Multi-modal**

- From scores to bipartitions: $m$ Jaccard Bag Descriptors: SIFT, Learning method:
  - Point detectors: Harris

**Concepts from scores to bipartitions:** $m$

**Learning method:**

- **Feature extraction:**
  - Harris-Laplace, Dense Sampling
  - Descriptors: SIFT, HSV-SIFT, HueSIFT, OppenemSIFT, C-SIFT, ngSIFT and RGB-SIFT

- **One multi-label dataset for each combination of detector-descriptor (14 in total)**

- **Learning method:**
  - Binary Relevance (BR) coupled with Random Forest

- **Instance weighting to deal with class imbalance**

**Photo annotation rankings**

<table>
<thead>
<tr>
<th>Approach</th>
<th>MIAP</th>
<th>F-ex</th>
<th>SR-Prec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual</td>
<td>9²/15 – 0.3114</td>
<td>5²/15 – 0.5595</td>
<td>9²/15 – 0.6981</td>
</tr>
<tr>
<td>Textual</td>
<td>3²/7 – 0.3256</td>
<td>2²/7 – 0.5061</td>
<td>3²/7 – 0.6257</td>
</tr>
<tr>
<td>Multi-modal</td>
<td>5²/10 – 0.4016</td>
<td>5²/10 – 0.5588</td>
<td>5²/10 – 0.6982</td>
</tr>
<tr>
<td>Overall</td>
<td>10²/18 – 0.4016</td>
<td>7²/18 – 0.5595</td>
<td>10²/18 – 0.6982</td>
</tr>
</tbody>
</table>

- Better in MIAP (model selection was based on Mean Average Precision)
- Averaging the multiple models worked better than arbitrary
- Good in textual – bad in visual – average overall

**Concept-based retrieval rankings**

- He are ranked 1st both in the automated and the manual retrieval approach
- Manual performs much better than automated on average
- Surprisingly automated performed better on 9 topics

**Software used**

- **Mulan** (http://mulan.sourceforge.net/)
  - Multi-label classification, feature selection and thresholding methods
  - Evaluation framework
- **ColorDescriptor** (http://koen.me/research/colordescriptors/)
- **Image feature extraction**
- **Weka** (http://www.cs.waikato.ac.nz/ml/weka/)
- **Text pre-processing – codebook generation (K-means clustering)**