Creating wallpapers using mobile phone photos

A semester project by Seonwook Park

Supervised by Michael Gygli and Dengxin Dai

Outline

- Introduction
- Method
  - Suitability: Datasets, Training
  - Cropping: Saliency, Features, Datasets
- Results and Analysis
- Summary
- Further work

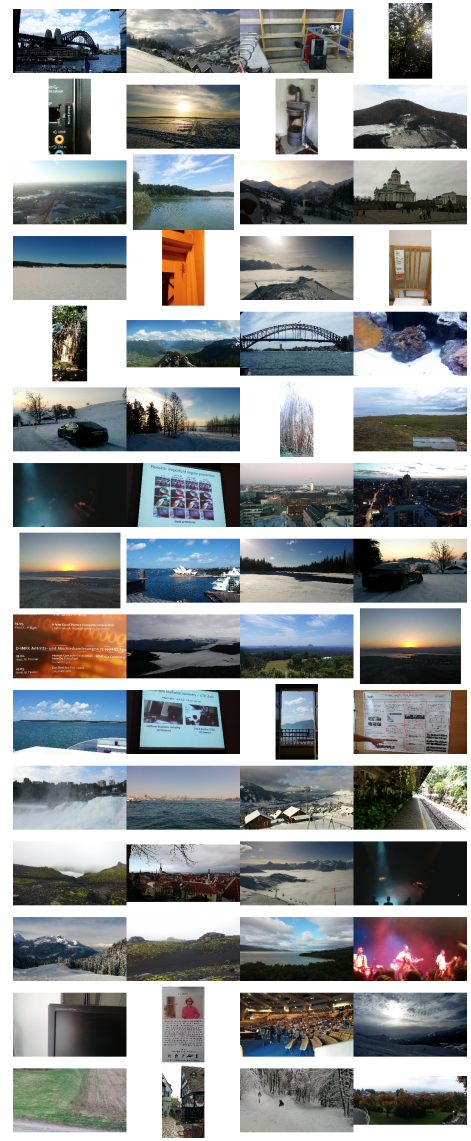
Introduction

- A mobile phone photo gallery contains a variety of images.
- Could use photos as wallpapers.
- Two challenges: Suitability and Cropping.

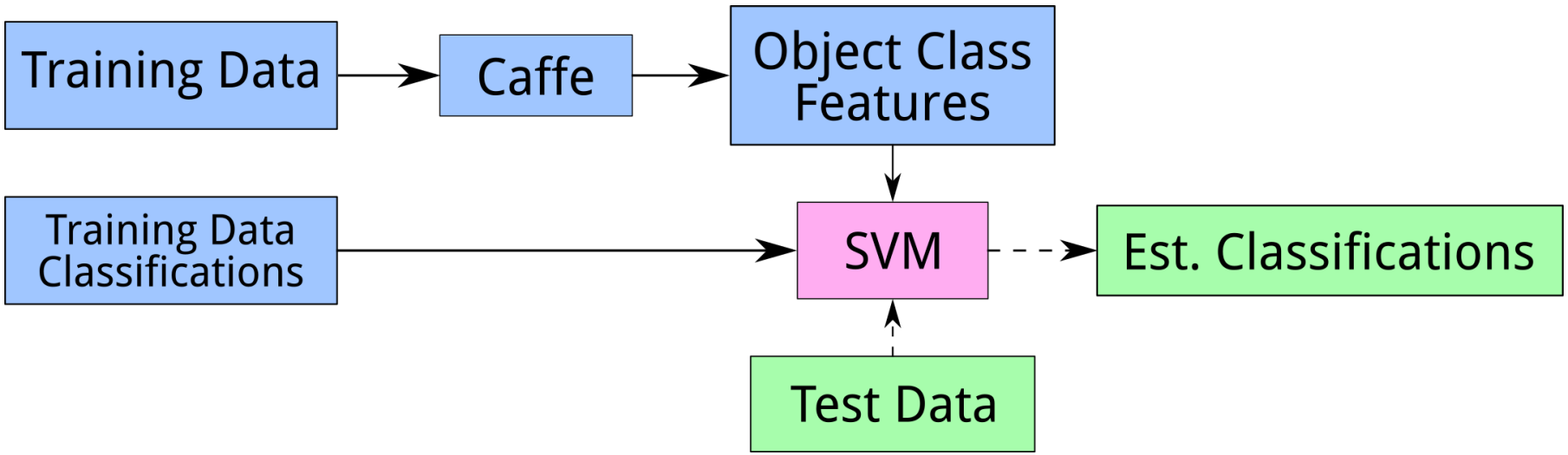
- Which image is suitable as a wallpaper?
- Which portion of a selected image should be visible?

Method: Suitability - Datasets

- Two datasets:
  - Michael (275 images)
  - Wookie (266 images)
- Includes:
  - Natural scenes
  - Cityscapes
  - Short-term memory
  - Quick photos for messaging

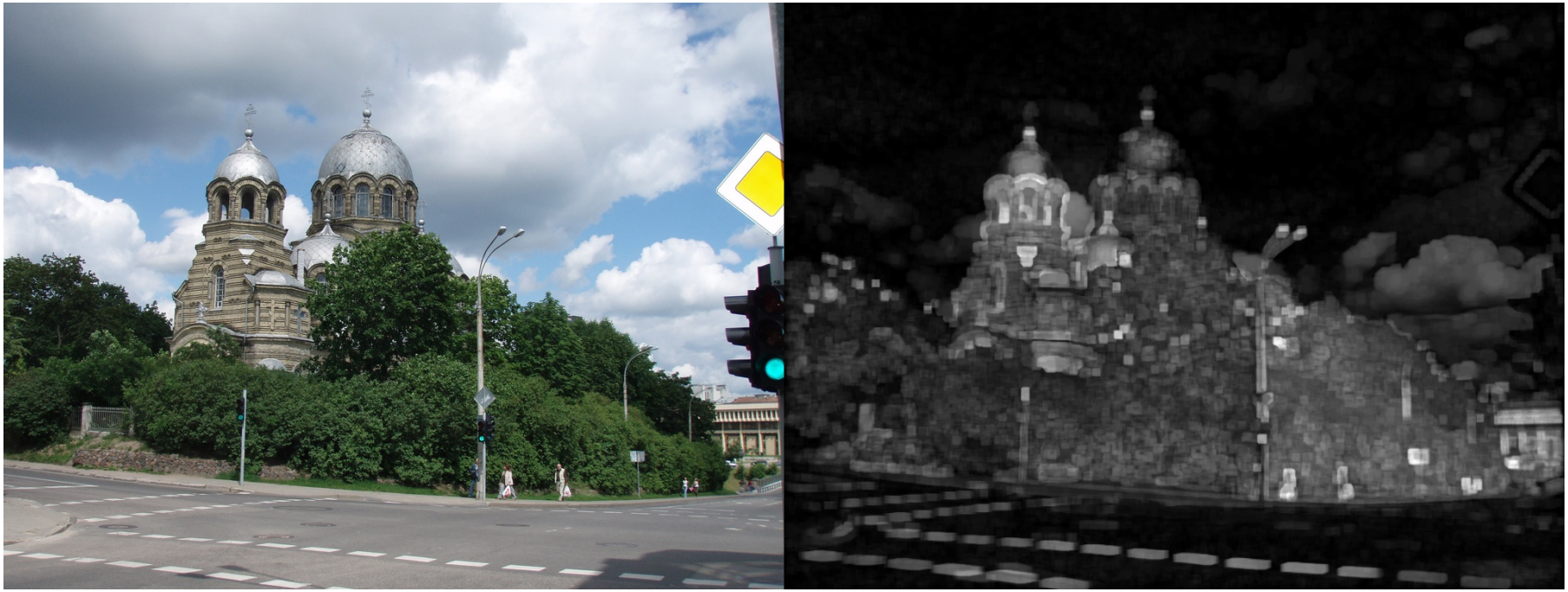


Method: Suitability - Training



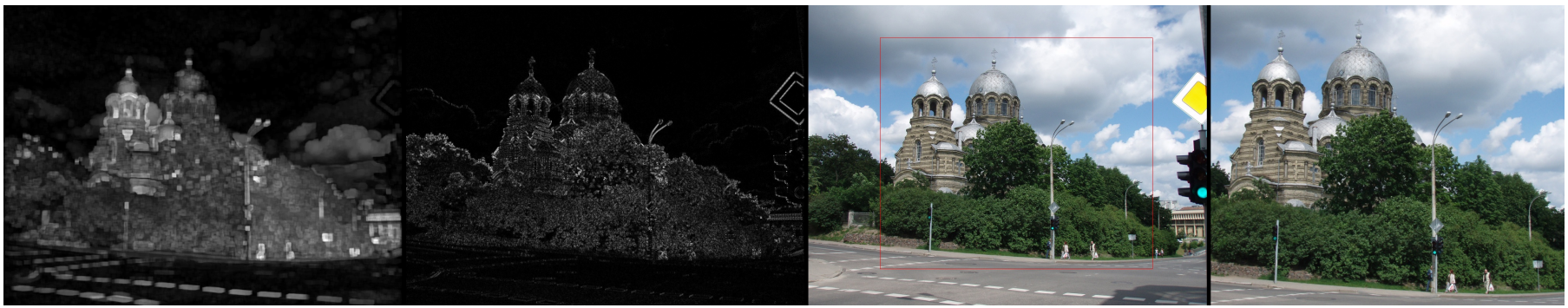
- Model used: BVLC Reference CaffeNet, 1000 object classes.

Method: Cropping - Saliency



- Saliency is distinctness compared to neighbouring regions.
- Boolean Map based Saliency (BMS) by Zhang et al.

Method: Cropping - Features



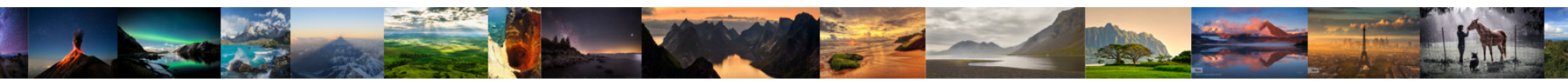
1. Saliency Composition
2. Boundary Simplicity
3. Content Preservation

Method: Cropping - Features

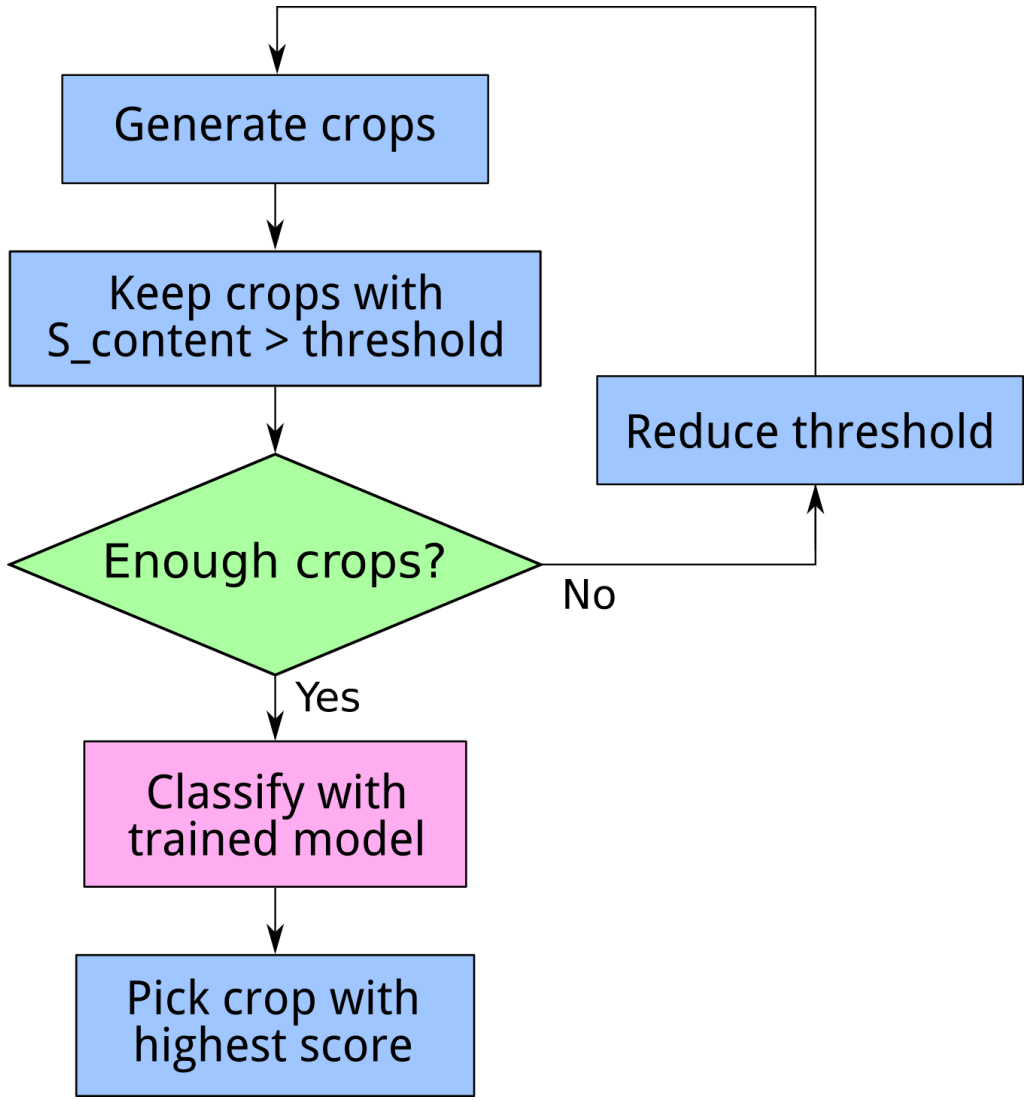
- Saliency composition
  - 3-level spatial pyramid of saliency map
  - $4^2 + 2^2 + 1 = 21$ features.
- Boundary simplicity
  - Gradient map generated using Sobel filter and `abs`.
  - Mean value of 2-pixel wide strips used.
  - 4 features, 1 per edge.
- Content preservation
  - Sum of all salient energy in crop

Method: Cropping - Datasets

- Human crop dataset
  - Collected from Amazon Mechanical Turks, Chen et al. (2014)
  - 500 images, 10 crops per image.
  - Used for evaluating model.
- Reddit dataset
  - Top 2000 images in the past year.
  - Subreddits used: CityPorn, EarthPorn, itookapicture, photocritique, WaterPorn, windowshots
  - Used for training model.



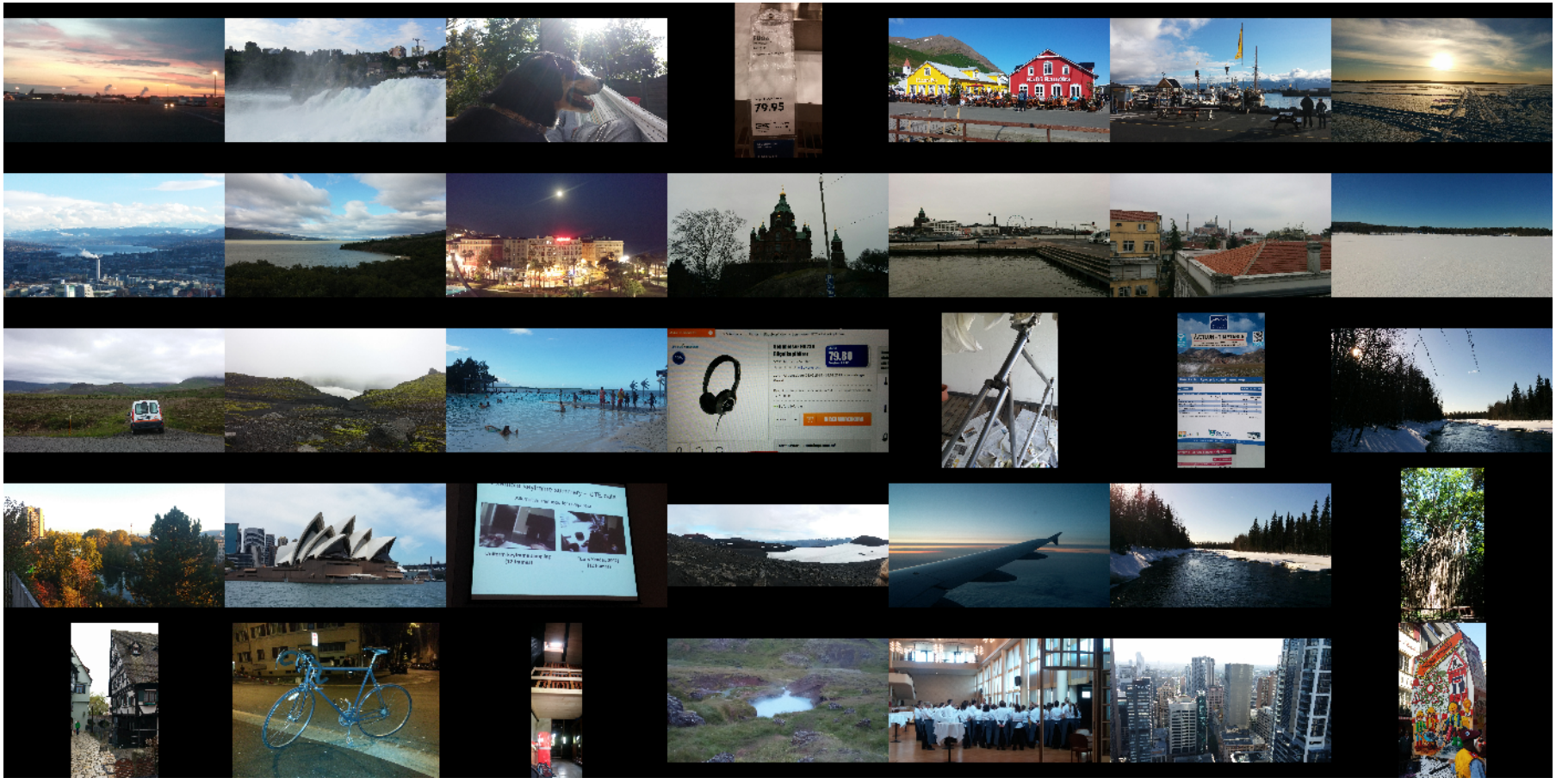
Method: Cropping - Final



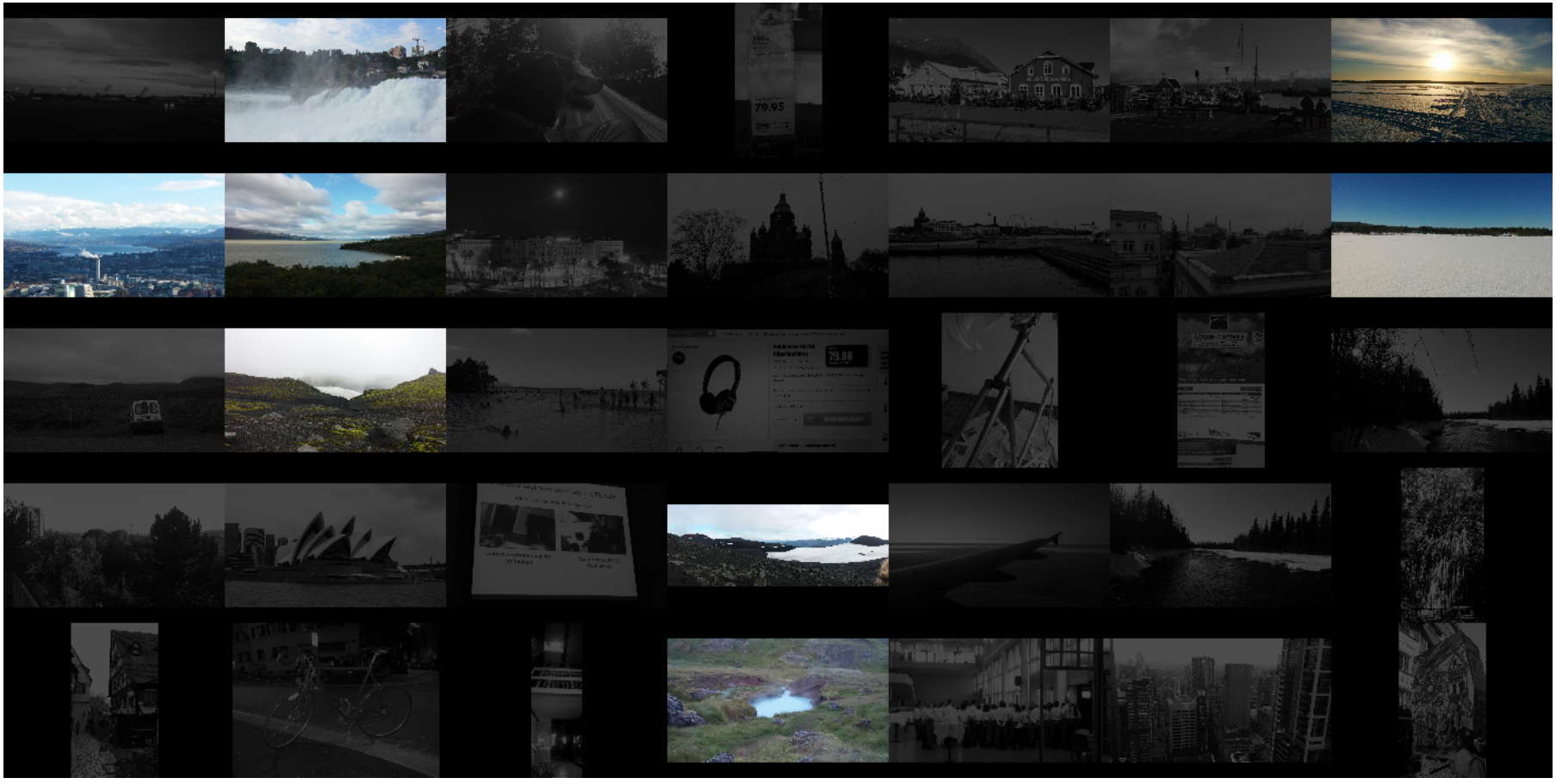
$$S_{content} = \frac{\text{sum(cropped saliency)}}{\text{sum(full saliency)}}$$

- Start from large crops.
- Find smaller crops.
- Stop when enough candidates.

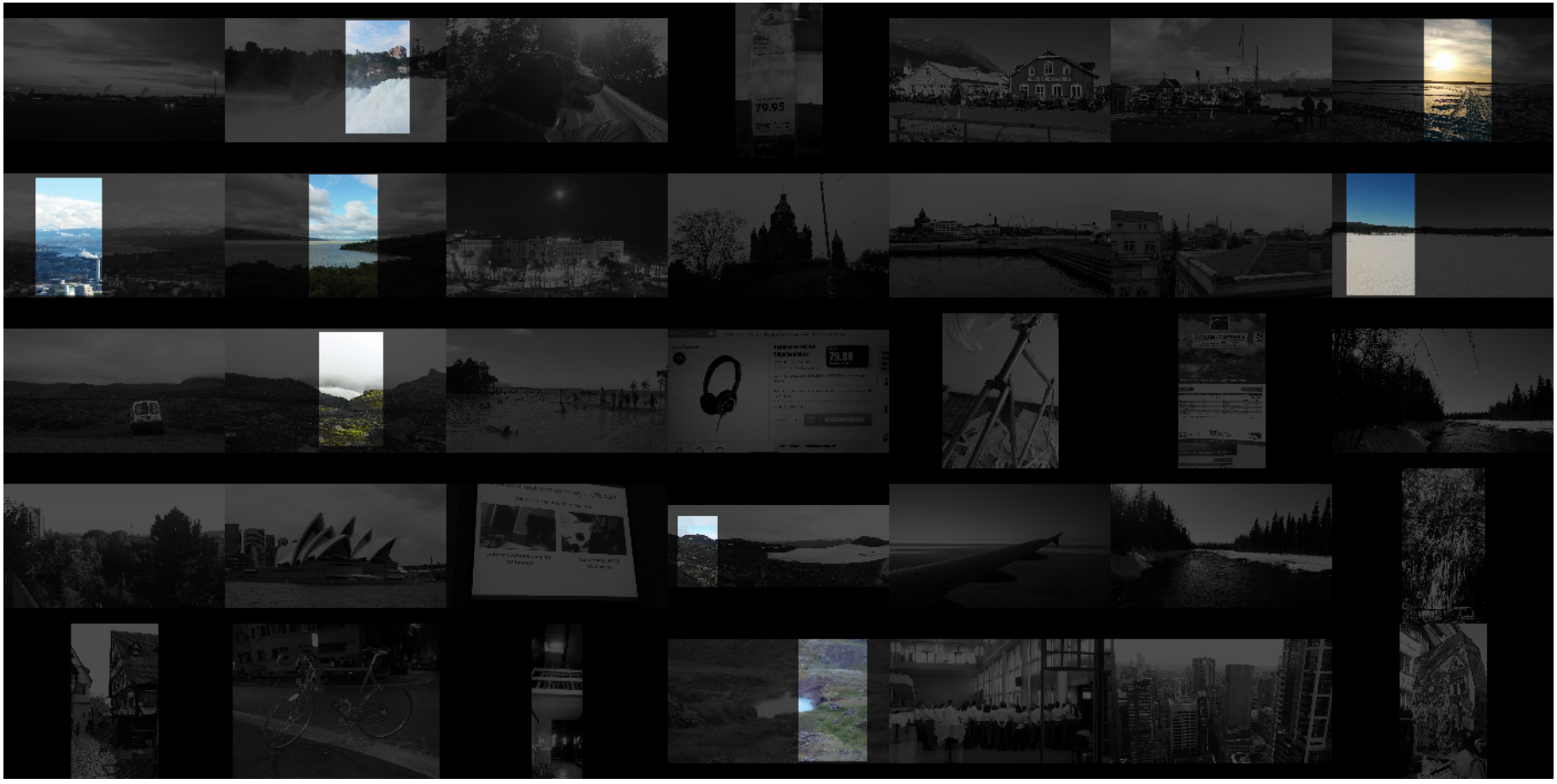
Final pipeline



Final pipeline

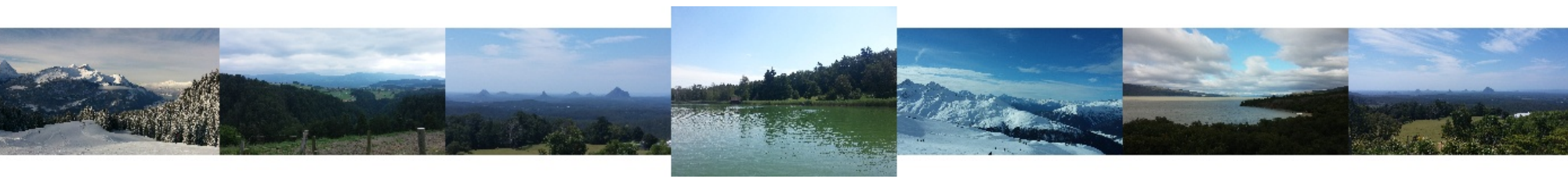


Final pipeline



Results: Suitability

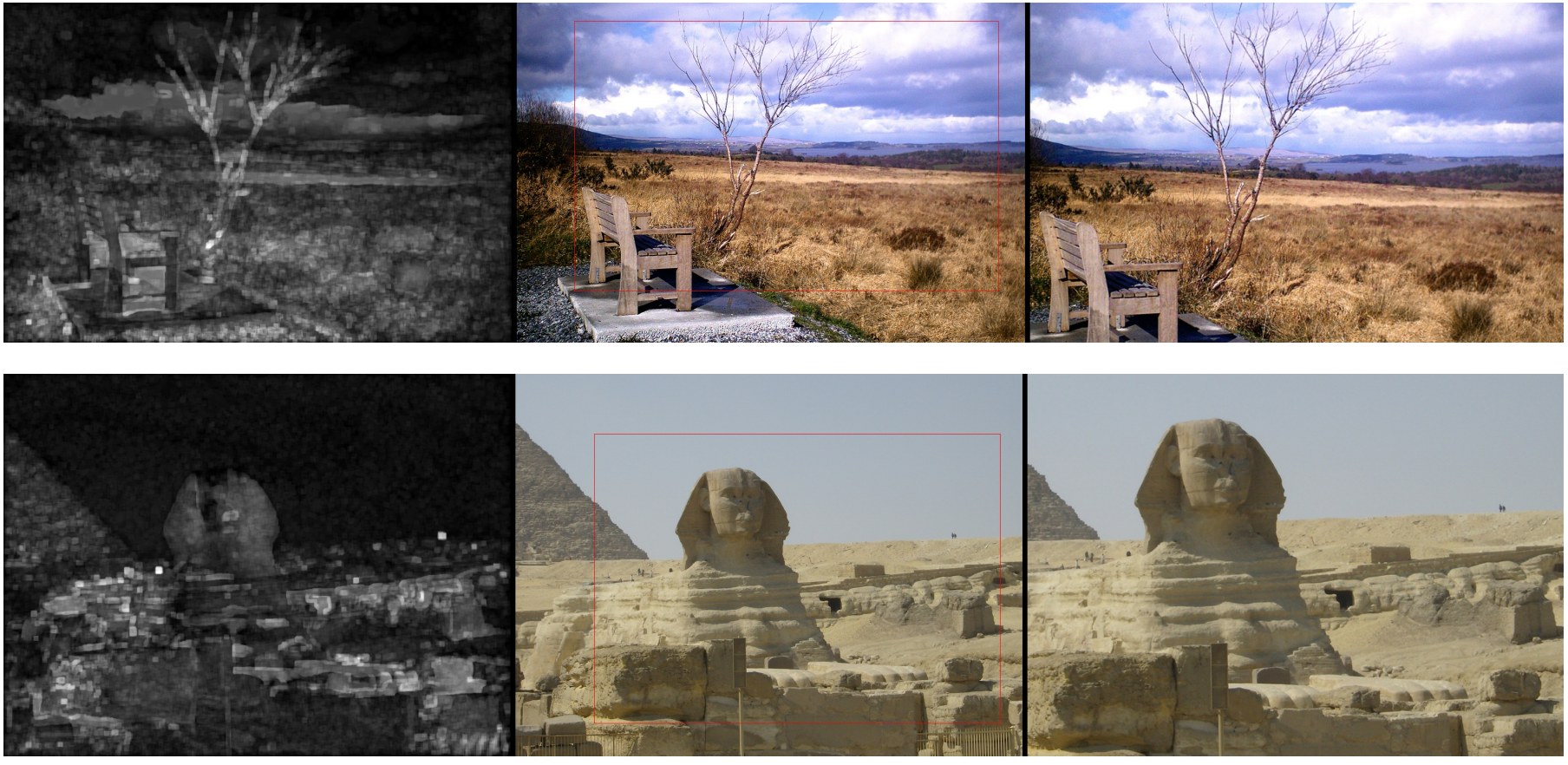
- Top 7 suitable images from Michael dataset.
- Suitability decreases towards right.
- 53 suitable images left out of 275.
- Favours distant natural scenes.



Analysis: Suitability

- Classifications gathered from four people. (% marked as suitable)
  - Michael set: 48%, 37%, 31%, 63%
  - Wookie set: 23%, 20%, 17%, 37%
- Pearson correlation coeff between two classification sets is 0.51 (Michael), 0.23 (Wookie)
- When considering images with classifications in consensus:
  - Trained on Wookie set: 9.8% error.
  - Trained on Michael set: 12.4% error.

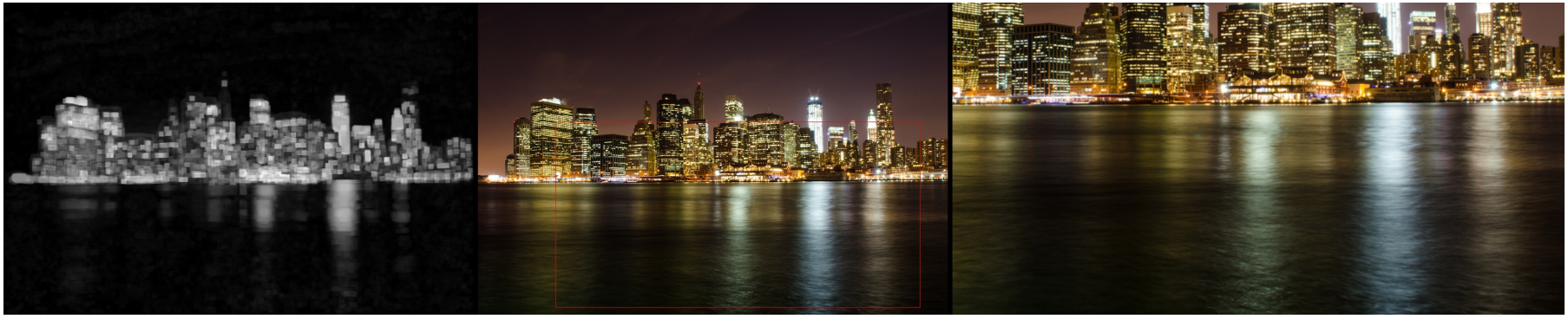
Results: Cropping (Good)

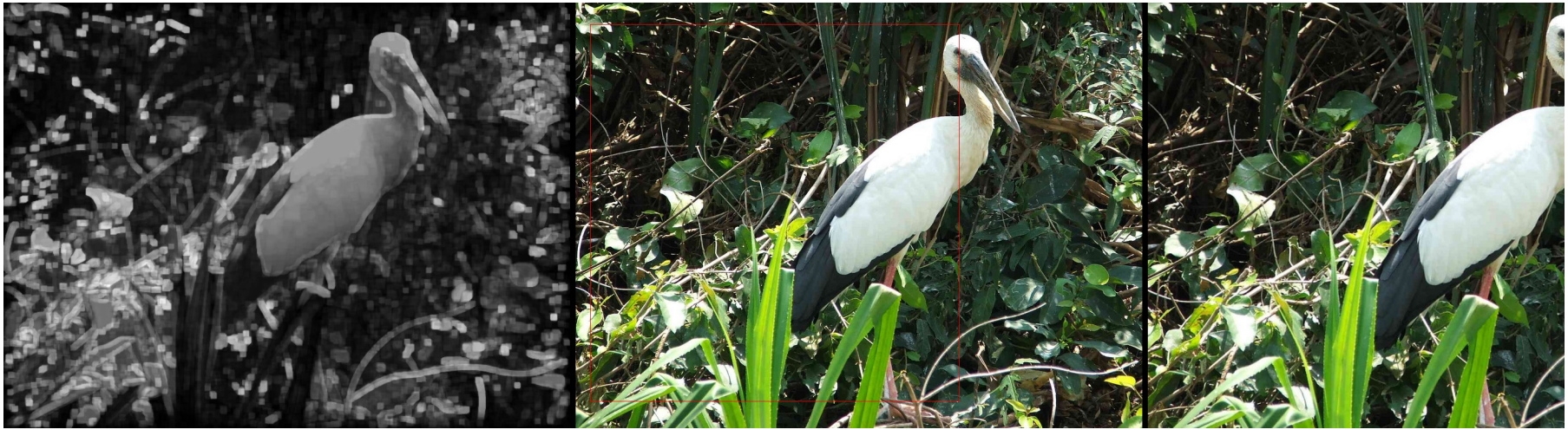


Results: Cropping (Good)

Results: Cropping (Good)

Results: Cropping (Bad)





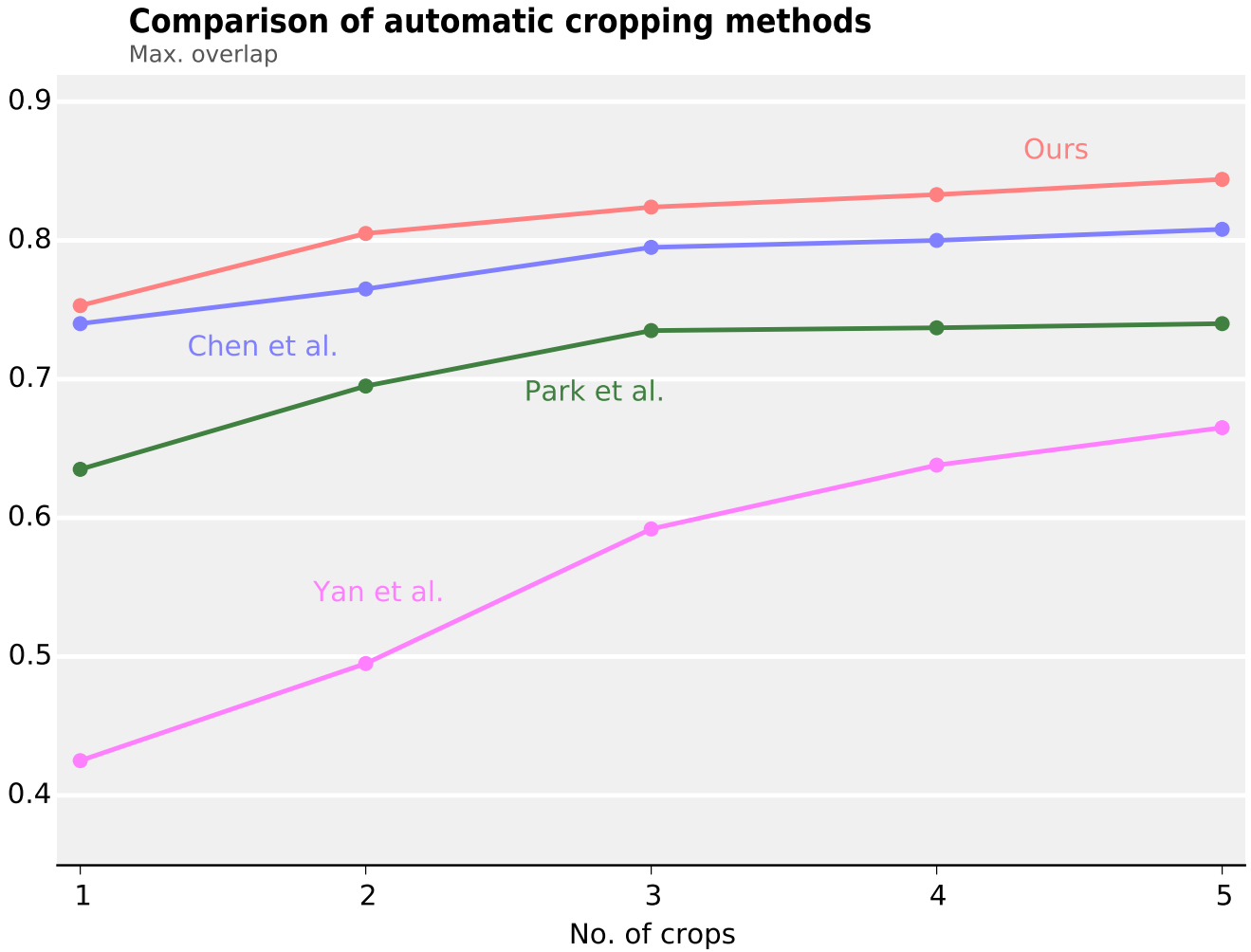
Analysis: Cropping

$$\text{Overlap}(C_i, H_j) = \frac{C_i \cap H_j}{C_i \cup H_j}$$

$$\text{MaxOverlap}(C_i, H) = \max_j \text{Overlap}(C_i, H_j)$$

- C_i is the i-th crop candidate, and
- H_j is the j-th human crop. There are 10 human crops per image.
- MaxOverlap is calculated for $i \in \{1, 2, 3, 4, 5\}$.
- C_0 is the highest scoring crop.

Analysis: Cropping



Improvements:

· 4 Boundary features

· Saliency sum feature

· Shrinking content threshold

Summary

- Trained models to select and crop an image given a set of images.
- Suitability classification error of 9.8%.
- Improved cropping algorithm.
  First 5 crop yields max. overlap of 0.84 over 0.81 from Chen et al.

- Implementation in `C++` (Cropping) and `Python` (Suitability) using OpenCV and Caffe.

Further Work

- Suitability
  - Train on larger datasets with wide range of objects.
  - Add features for image quality and colour distribution.
  - Reduce classification bias.
- Cropping
  - Try different saliency map implementations.
  - Find other visual cues such as variance of salient energy.