Exercise 3.2 Report: Balloon Detection Pipeline

Talha Tariq

June 16, 2025

Objective

The goal of Task 3.2 is to implement a simplified object detection pipeline based on the method proposed by Uijlings et al. Using our custom selective search (from Task 3.1) to generate region proposals, we train an SVM classifier on CNN-extracted features to detect balloons in a small COCO-style dataset.

Pipeline Overview

The pipeline includes the following stages:

1. **Selective Search (from Task 3.1):** We implemented a simplified selective search using the Felzenszwalb algorithm via `skimage.segmentation`, extracting bounding boxes from the initial segmentation labels.
2. **Proposal Generation:** Region proposals were generated for each image in the train, validation, and test sets.
3. **Proposal Matching:** Ground truth annotations were matched to region proposals using the IoU metric. Proposals with IoU > 0.5 were labeled positive, and those with IoU < 0.2 as negative.
4. **Feature Extraction:** We used a pre-trained ResNet18 (from torchvision) to extract 512-D feature vectors from proposed regions.
5. **SVM Classification:** An SVM was trained using `sklearn.svm.SVC` on the extracted features.
6. **Inference:** The model was applied to test images to classify the proposed regions and draw boundary boxes. The results were displayed and saved.

Q5.2.1 Differences from Uijlings et al.

Uijlings et al. used SIFT features and combined multiple color spaces and similarity heuristics to guide merging in selective search. In contrast:

- We used only grayscale Felzenszwalb segmentation (no color histograms or merging similarities).

- Feature extraction was done using a CNN (ResNet18) instead of SIFT or VLAD.
- The final detection step uses a basic SVM rather than a more complex classification pipeline.

Q5.2.2 Threshold Design

We used two IoU thresholds: $tp = 0.5$ for positives and $tn = 0.2$ for negatives. This gap avoids ambiguous examples with moderate overlap. A single threshold would risk mislabeling borderline proposals, potentially degrading classifier performance.

Q5.2.3 Data Augmentation Suggestions

The dataset is small, limiting generalization. To improve training:

- Use geometric augmentations (flips, rotations, crops).
- Extract multiple positive proposals from each ground truth region using IoU ≥ 0.5.
- Use synthetic data or transfer learning to enhance model robustness.

Results

Detection bounding boxes were correctly predicted for several test images. All images were saved with overlay detections.

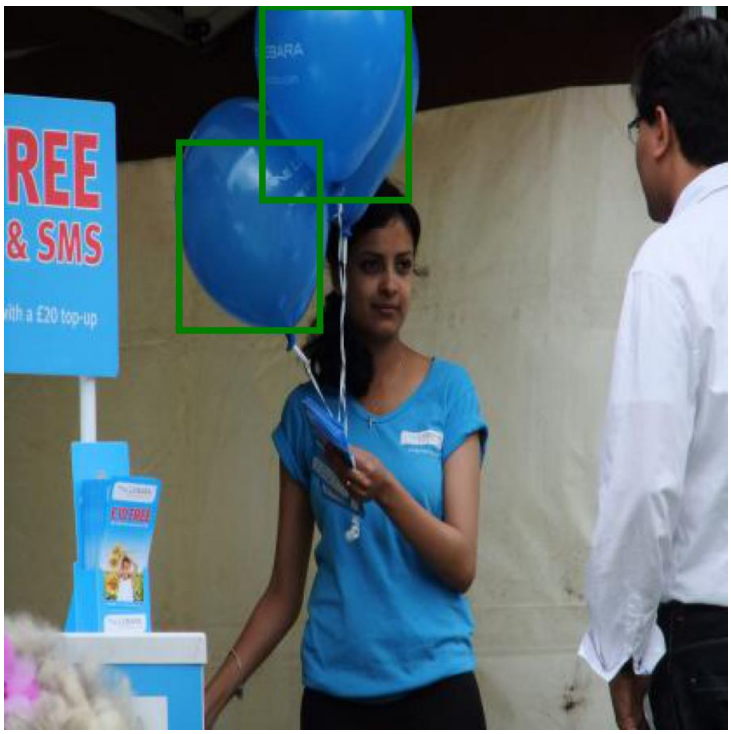

Figure 1: Balloon detection result using selective search and SVM