Visualchunking: A List Prediction Framework for Region-Based Object Detection

N. Rhinehart, J. Zhou, M. Hebert, J. A. Bagnell
Carnegie Mellon University, Pittsburgh, PA

Motivation

Task: produce list of class-specific detections

Common approach:
1. binary classification on windows or regions
2. thresholding (e.g. NMS)

Our approach:
1. Generate class-specific candidate regions via learned growing procedure
2. Build list from best candidates as scored by a list prediction learner

Differences:
- Common approach resorts to thresholding in attempt to reconcile multiple detections
- Our approach
  *incorporates information from previous predictions* to allow better detection of multiple instances
- directly targets a natural metric for scoring a list of detections by learning to emulate our training-time near-optimal greedy algorithm (Algorithm 1)

Approach

Class-specific growing procedure by learning to mimic Algorithm 2

Algorithm 1 Greedy List Generation with Ground-Truth Access

Input: Set of candidate chunks $C$, set of ground truth instances $G$, size of predicted list $k$

Output: A near-optimal list $L^P$ of chunks

for $i = 1$ to $k$
do
  $G_l = \arg\max_{G \in G, \pi \in \pi(G)} \pi(G, c, G_l) > c = \emptyset, G_l = \emptyset$
  \[ L^P = L^P \cup \{c\} \]
  \[ \text{append the chunk to the list} \]
  \[ G_m = G_m \setminus \{c\} \]
  \[ \text{remove the associated GT} \]
end for

Objective function rewards good coverage

Algorithm 2 Single Instance Chunk Growing Algorithm

Input: Set of superpixels $S$, grower predictor $G$

Output: A set of chunks, $C_g$

for $i = 1$ to $|S|$ do
  1. Sort elements in $S$ by decreasing order of $s_i = G(s_i)$
  2. for $i = 1$ to $|S|$ do
     1. $e = \cup \{s_i\}$, $C_g = C_g \cup \{e\}$
  end for
end for

Candidate quality and list prediction quality versus several baselines on LM+Sun multicars dataset (521 test images)

Performance of list prediction on PASCAL VOC2012. We find that for this dataset, training data lacks sufficient number of images with multiple instances.