Beyond Active Noun Tagging

Behjat Siddiquie  
behjat@cs.umd.edu

Abhinav Gupta  
abhinavg@cs.cmu.edu
What this Talk is About

Problem: Active Learning for Scene Understanding Tasks
What this Talk is About

Image Entropy

Region Entropy

Image Entropy

- sky
- airplane
- grass
- road
What this Talk is About

Introduce new ways of collecting knowledge which go beyond labeling regions
What this Talk is About

Introduce new ways of collecting knowledge which go beyond labeling regions

Linguistic Questions

- Use high confidence detections as anchors to ask questions about the unknown objects in a scene
What this Talk is About

Introduce new ways of collecting knowledge which go beyond labeling regions

Linguistic Questions

- Use high confidence detections as anchors to ask questions about the unknown objects in a scene
What this Talk is About

Introduce new ways of collecting knowledge which go beyond labeling regions

Linguistic Questions

- Use high confidence detections as anchors to ask questions about the unknown objects in a scene
What this Talk is About

Introduce new ways of collecting knowledge which go beyond labeling regions

Contextual Questions

- Importance of relationships
What this Talk is About

Introduce new ways of collecting knowledge which go beyond labeling regions

Contextual Questions

- Importance of relationships
- Learn general contextual relationships directly from the annotator
Introduction

Scene Understanding
- Learn $O(n)$ appearance models
- Learn $O(n^2)$ relationship models

Heavy Tailed Distribution
- The large majority of classes occur rarely
- Holds for relationships too

Randomly annotating training data is extremely inefficient

Russell et al. IJCV, 2008
Active Learning

Binary Classification Problems
- SVMs, Tong & Koller, 2001
Active Learning

Binary Classification Problems

- SVMs, Tong & Koller, 2001
Binary Classification Problems

- SVMs, Tong & Koller, 2001
Active Learning: Multi-Class

Binary Classification Problems
- SVMs, Tong & Koller, 2001

Multi-Class Classification Problems
- Jain & Kapoor, 2009

Caltech 101/256
Active Learning: Scene Understanding

Binary Classification Problems
- SVMs, Tong & Koller, 2001

Multi-Class Classification Problems
- Jain & Kapoor, 2009

Active Learning for Scene Understanding
Active Learning: Image Entropy

Uncertainty Based Approaches

Region Entropy
Active Learning: Image Entropy

Uncertainty Based Approaches

Region Entropy

Image Entropy

- airplane
- grass
- road
- sky
Active Learning: Image Entropy

Current Model

Image Dataset
Active Learning: Image Entropy

Region Entropy

Image Dataset

boat water tree on

Current Model
Active Learning: Image Entropy

- Boat
- Water
- Tree

Current Model

Image Dataset

Region Entropy

Image Entropy
Active Learning: Image Entropy

Image Dataset

Region Entropy

Image Entropy
Active Learning: Image Entropy

Image Dataset

Region Entropy

Image Entropy
Active Learning: Image Entropy

Entropy over all possible label assignments

Second order approximation
- Sum of the joint entropies of all pairs of regions
Active Learning Framework

Training Data

Unlabeled Set

Initially Labeled

Classifier
Active Learning Framework

Training Data

Unlabeled Set
Initially Labeled
Actively Labeled

Obtain answer

Generate set of questions
Classifier

What is above grass?

What is the relationship between car and road?

What is the relationship between cow and grass?

Obtain answer
Types of Questions

Region Labeling Questions

What is above the water?

Linguistic Questions

What is the relationship between boat and water?

Contextual Questions
Types of Questions

Region Labeling Questions

What is above the water?

What is the relationship between boat and water?
Region Labeling Questions

Appearance Model

boat  car
Region Labeling Questions

Appearance Model

boat
car
Region Labeling Questions

Appearance Model

boat
car

Appearance Model

boat
car
Region Labeling Questions

Appearance Model

**boat**

**car**

Appearance Model

**boat**

**car**

Importance of a Question

Initial Image Entropy

Final expected Image Entropy given labeling
Types of Questions

Linguistic Questions

What is above the water?

What is the relationship between boat and water?
Linguistic Questions

Appearance Model

boat  water

Contextual Model

sky above grass
sky bluer than grass
•
•
car above road
Linguistic Questions

Appearance Model

- boat
- water

Contextual Model

- sky above grass
- sky bluer than grass
- car above road
Linguistic Questions

Appearance Model
- boat
- water

Contextual Model
- sky above grass
- sky bluer than grass
- car above road

What is above the water?
Linguistic Questions

Appearance Model

boat
water

Contextual Model

sky above grass
sky bluer than grass

- 
- 
car above road

What is above the water?

1. Boat
2. Tree
Linguistic Questions

Appearance Model

boat  water

Contextual Model

sky above grass
sky bluer than grass

1. Boat
2. Tree

What is above the water?
Linguistic Questions

What is above the water?

1. Boat
2. Tree

Contextual Model:
- sky above grass
- sky bluer than grass
- car above road

Appearance Model:
- boat
- water

Appearance Model:
- boat
- water
- tree

Contextual Model:
- sky above grass
- sky bluer than grass
- car above road
- boat above water
- tree above water
What is below the sky?
Linguistic Questions: Expected Entropy Reduction

What is **below the sky**?
Linguistic Questions: Expected Entropy Reduction

What is below the sky?

Entropy Reduction

Expected Image Entropy given a weak labeling $C_q$ of the relevant regions
Types of Questions

What is above the water?

What is the relationship between boat and water?

Contextual Questions
Contextual Questions

Appearance Model

boat

| car |

Contextual Model

- **sky above grass**
- **sky bluer than grass**
- **car above road**
Contextual Questions

Appearance Model

boat

car

Contextual Model

sky above grass

sky bluer than grass

•

•

car above road

What is the relationship between boat and water?
Contextual Questions

What is the relationship between boat and water?

1. boat above water
2. water bluer than boat
Contextual Questions: Expected Entropy Reduction

Correct Relationship Prior

- Constrains the image labeling problem
- Results in a large entropy reduction

Sample object pairs based on their co-occurrence
Experiments

MSRC Dataset
- 532 images (276 training, 256 test)
- 21 categories and multi-label images

Ground Truth Segmentations
Results: MSRC Dataset – Ground Truth Segmentations
Results: MSRC Dataset – Ground Truth Segmentations
Results: MSRC Dataset – Ground Truth Segmentations

- Random
- Jain et al. (First Order Entropy)
- Our Approach (Region Labeling Questions)
- Our Approach (Linguistic Questions)
- Our Approach (Contextual Questions)
- Our Approach (Combined)

Combined

Region Labeling
Results: MSRC Dataset – Ground Truth Segmentations

![Graph showing performance comparison between different approaches across the number of questions.]

- Random
- Jain et al. (First Order Entropy)
- Our Approach (Region Labeling Questions)
- Our Approach (Linguistic Questions)
- Our Approach (Contextual Questions)
- Our Approach (Combined)

**Combined**

**Linguistic**
Results: MSRC Dataset – Ground Truth Segmentations

- Random
- Jain et al. (First Order Entropy)
- Our Approach (Region Labeling Questions)
- Our Approach (Linguistic Questions)
- Our Approach (Contextual Questions)
- Our Approach (Combined)

Combined

Contextual
Results: MSRC Dataset – Ground Truth Segmentations
Results: MSRC Dataset – Ground Truth Segmentations
Results: MSRC Dataset – Ground Truth Segmentations
Results: MSRC Dataset – Ground Truth Segmentations

Image Entropy

Region Entropy
Results: Sample Questions on the MSRC Dataset

Region Labeling Questions

Our Approach (Image Entropy)  [Jain & Kapoor, CVPR 2009] (Region Entropy)
Results: Sample Questions on the MSRC Dataset

Region Labeling Questions

Our Approach (Image Entropy)  [Jain & Kapoor, CVPR 2009] (Region Entropy)
Results: Sample Questions on the MSRC Dataset

Region Labeling Questions

Our Approach (Image Entropy) [Jain & Kapoor, CVPR 2009] (Region Entropy)
Results: Sample Questions on the MSRC Dataset

Region Labeling Questions

Our Approach (Image Entropy) [Jain & Kapoor, CVPR 2009] (Region Entropy)
Results: Sample Questions on the MSRC Dataset

Linguistic Questions

1. building - tree
2. cow - grass
3. tree - ground
4. sky - building
5. sky - tree
6. grass - building
7. sky - grass
8. tree - grass
9. boat - water

Contextual Questions

What is **above** the **ground**?

What is **above** the **grass**?

What is **more blue** than the **ground**?
Experiments

MSRC Dataset

- 532 images (276 training, 256 test)
- 21 categories and multi-label images

Automatic Segmentations
Results: MSRC Dataset – Automatic Segmentations

![Graph showing performance across different approaches and question numbers.](image-url)
Results: MSRC Dataset – Automatic Segmentations

- Random
- Jain et al. (First Order Entropy)
- Our Approach (Region Labeling Questions)
- Our Approach (Linguistic Questions)
- Our Approach (Contextual Questions)
- Our Approach (Combined)

Accuracy vs. Number of Questions

- Region Entropy
- Image Entropy
Results: MSRC Dataset – Automatic Segmentations
Results: MSRC Dataset – Automatic Segmentations

![Graph showing the accuracy of different approaches over the number of questions. The graph compares Random, Jain et al. (First Order Entropy), Our Approach (Region Labeling Questions), Our Approach (Linguistic Questions), Our Approach (Contextual Questions), and Our Approach (Combined).]
Experiments: Stanford Dataset

Stanford Dataset

- 715 images (415 training, 300 test)
- Annotated using Amazon Mechanical Turk
- 8 categories and multi-label images
Results: Stanford Dataset

- Combined
- Linguistic
- Region Labeling
- Contextual

![Graph showing the performance of different approaches over the number of questions.](attachment:graph.png)

- Random
- Jain et al. (First Order Entropy)
- Our Approach (Region Labeling Questions)
- Our Approach (Linguistic Questions)
- Our Approach (Contextual Questions)
- Our Approach (Combined)
Results: Stanford Dataset

The graph illustrates the accuracy of different approaches over the number of questions. The x-axis represents the number of questions, while the y-axis represents the accuracy. The legend indicates the following approaches:

- Random
- Jain et al. (First Order Entropy)
- Our Approach (Region Labeling Questions)
- Our Approach (Linguistic Questions)
- Our Approach (Contextual Questions)
- Our Approach (Combined)

The graph shows a comparison of these approaches, with the combined approach generally achieving the highest accuracy as the number of questions increases.
Results: Stanford Dataset

![Accuracy vs. Number of Questions for different approaches]

- Random
- Jain et al. (First Order Entropy)
- Our Approach (Region Labeling Questions)
- Our Approach (Linguistic Questions)
- Our Approach (Contextual Questions)
- Our Approach (Combined)
Summary

Active learning method for learning contextual object recognition models

Image Entropy

Linguistic Questions

Contextual Questions
Questions?

What is above the water?

What is the relationship between boat and water?

Please visit our poster located at B4