View Planning for Object Recognition

Gabriel Oliveira and Volkan Isler
RSN Lab

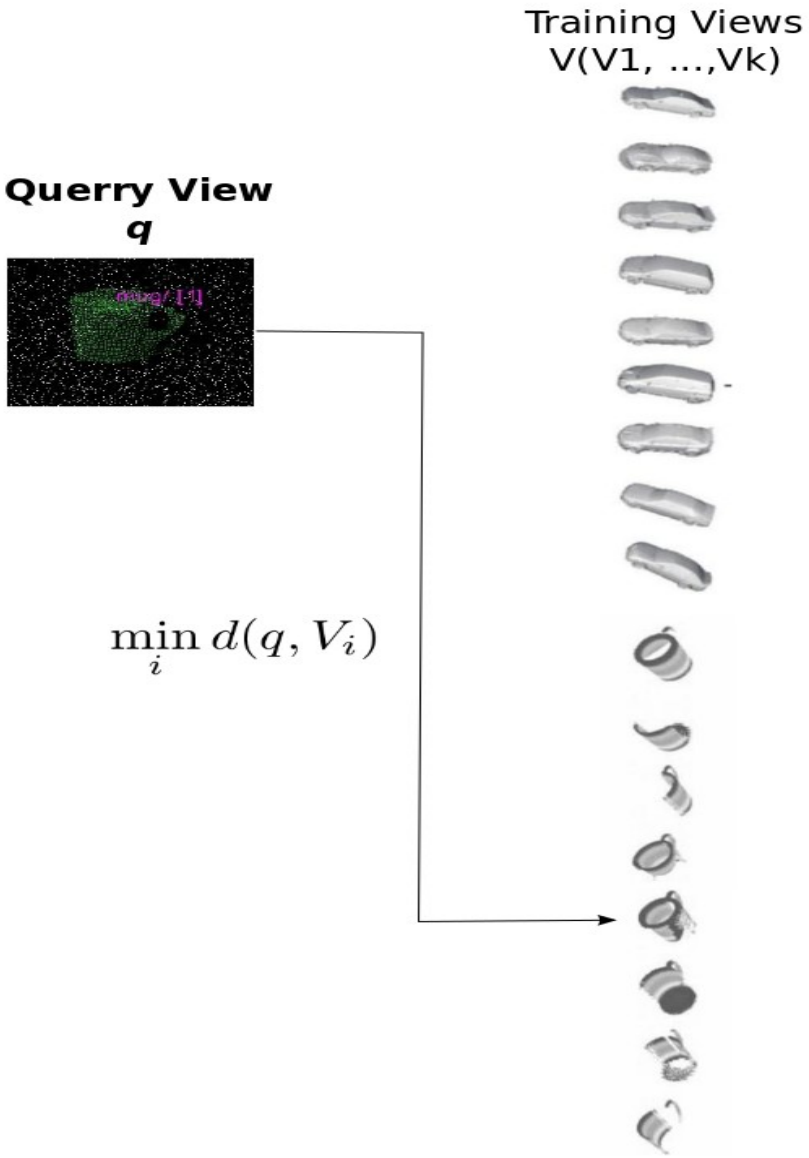
UNIVERSITY OF MINNESOTA
Driven to Discover℠

Motivation

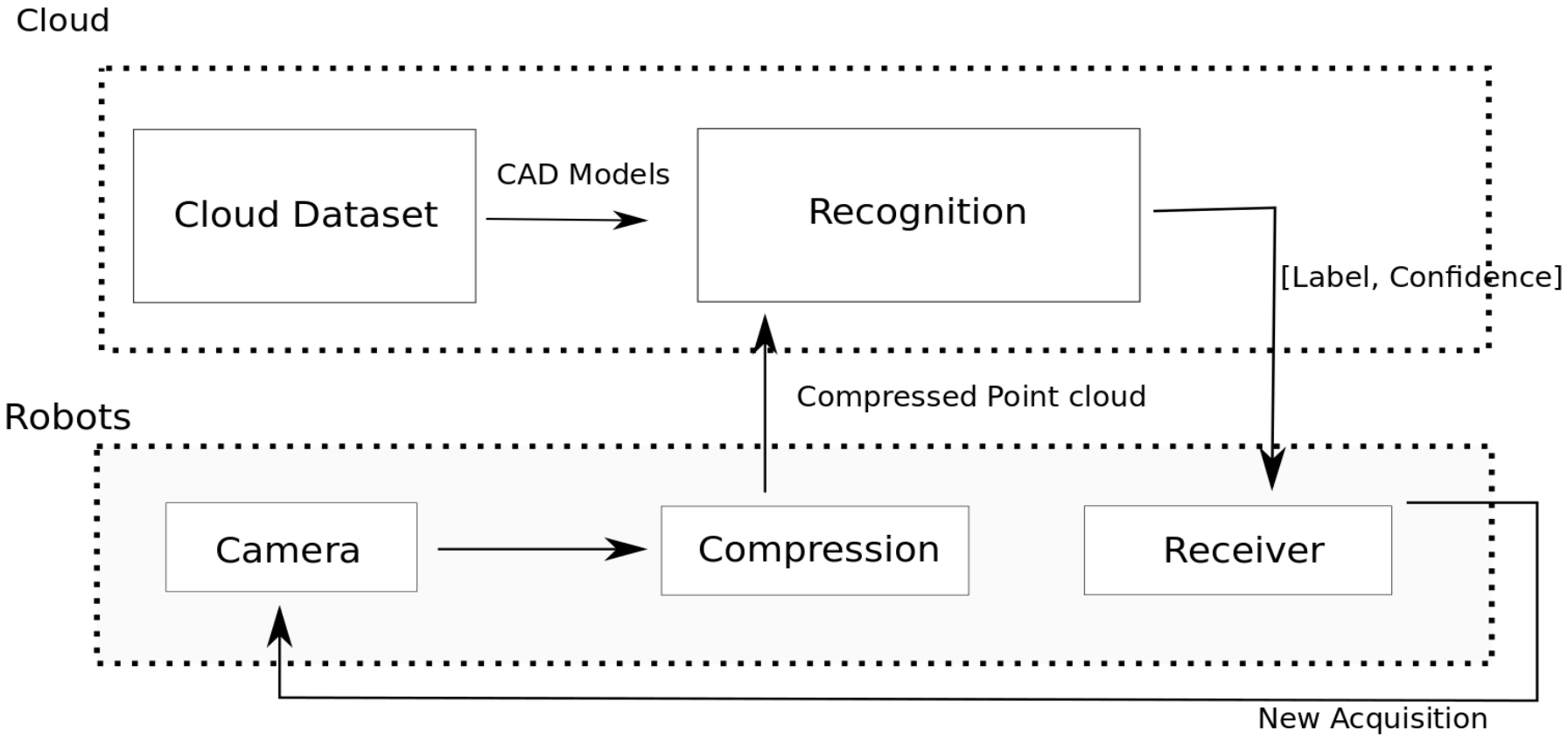
Objective

- **Cloud-Based (Active) Object recognition**
- **Goal: Find the minimum amount of views for recognition**

Problem Definition

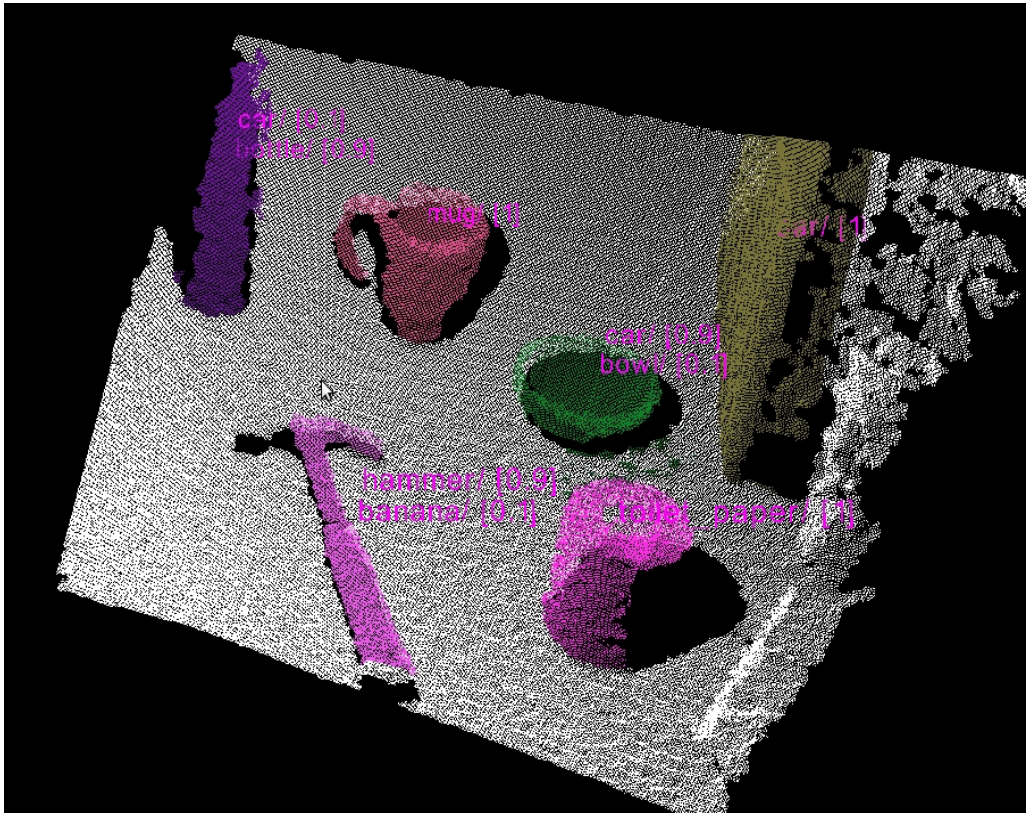


System Overview



Recognition

- **Recognition module**
- **[Vincze et al. 2012]:**
  - Segmentation (RANSAC)
  - Descriptor (ESF)
  - Matching (KNN)
  - Merging (Max all views)



Viewpoints

- **Open loop approach:**
  - No prior Knowledge about the next view
- **Approximation of Edge Based Best Next View approach [Abidi et al. 2000]:**
  - Explore areas of occlusion
  - Approximate the three first views to be pairwise orthogonal

Viewpoints

- **Empirical Upper Bound the number of views:**
  - 4 views in a plane:
  - All views are orthogonal to its 2 closest neighbors

Experiments

- Dataset
- Time performance
  - Communication
  - System Bottleneck (segmentation)
- Recognition Results
- Distribution of Viewpoints

Experiments

- **Dataset**
- Time performance
  - Communication
  - System Bottleneck (segmentation)
- Recognition Results
- Distribution of Viewpoints

Experiments - Setup

Experiments

- **Used Dataset**

Experiments

- Dataset
- **Time performance**
  - **Communication**
  - System Bottleneck (segmentation)
- Recognition Results
- Distribution of Viewpoints

Experiments

- **Communication Results**

Method	Mean	Standard Deviation	Size of cloud Sent
Transmission without Passthrough filter	10.35 fps	2.28	105 Kb from 4500 Kb original size
Transmission with Passthrough filter from 1.0 to 3.5 meters	6.55 fps	1.40	87 Kb from 4500 Kb original size

Experiments

- Dataset
- **Time performance**
  - Communication
  - **System Bottleneck (segmentation)**
- Recognition Results
- Distribution of Viewpoints

Experiments

- **Segmentation**

Segmentation # of objects (frame-rate)	Minimum (ms)	Maximum (ms)
1 object (~3.4 fps)	270	310
2 object (~2.6 fps)	355	400
3 object (~1.9 fps)	500	530

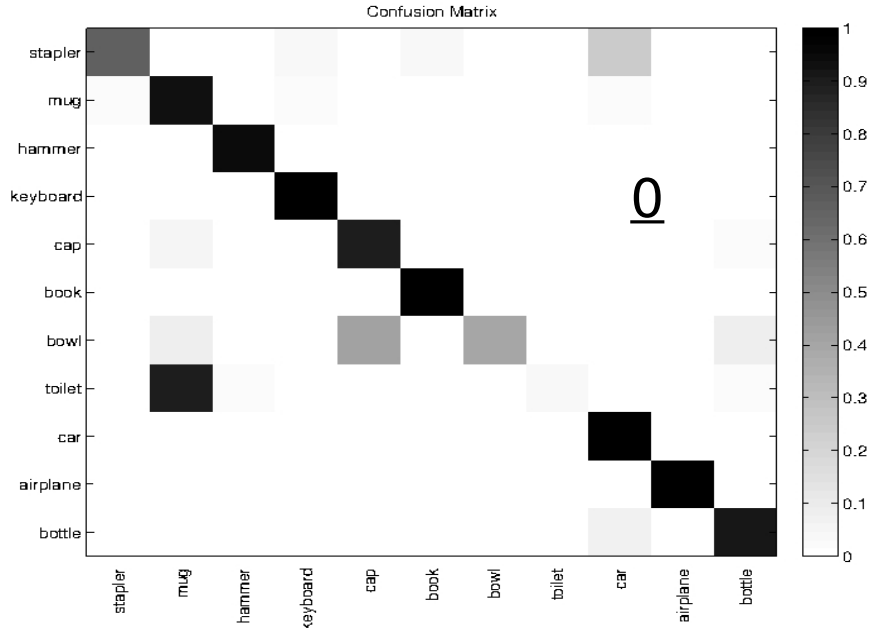
Experiments

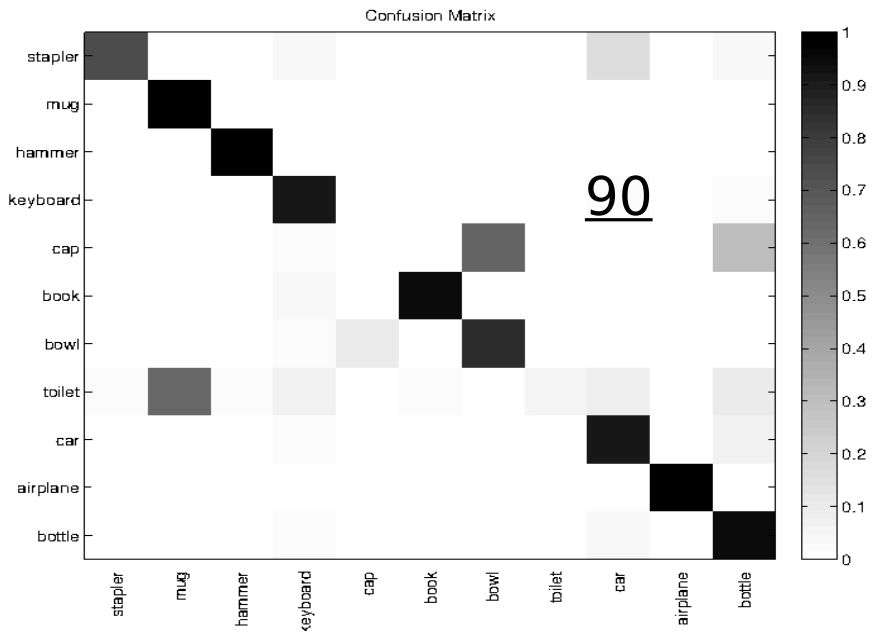
- Dataset
- Time performance
  - Communication
  - System Bottleneck (segmentation)
- **Recognition Results**
- Distribution of Viewpoints

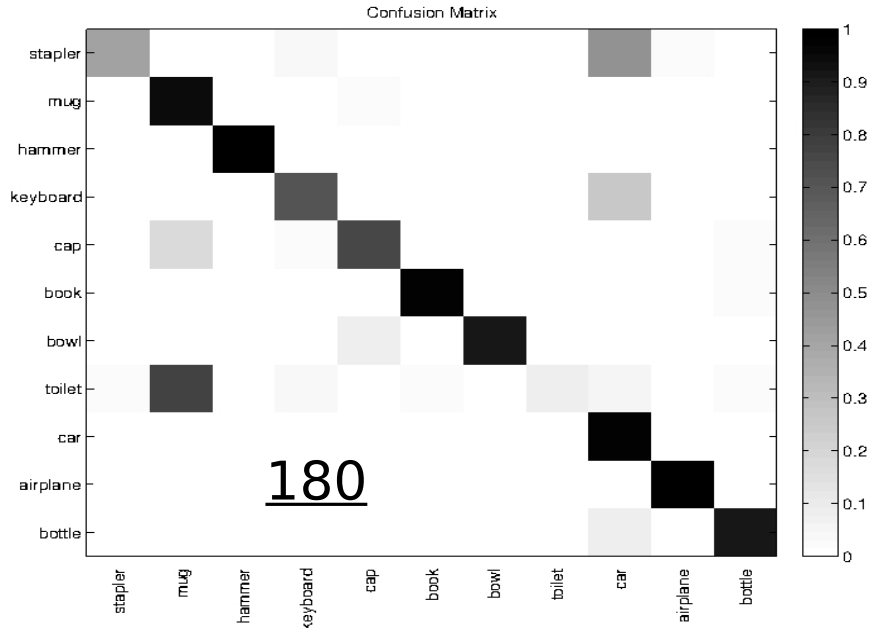
Experiments Recognition

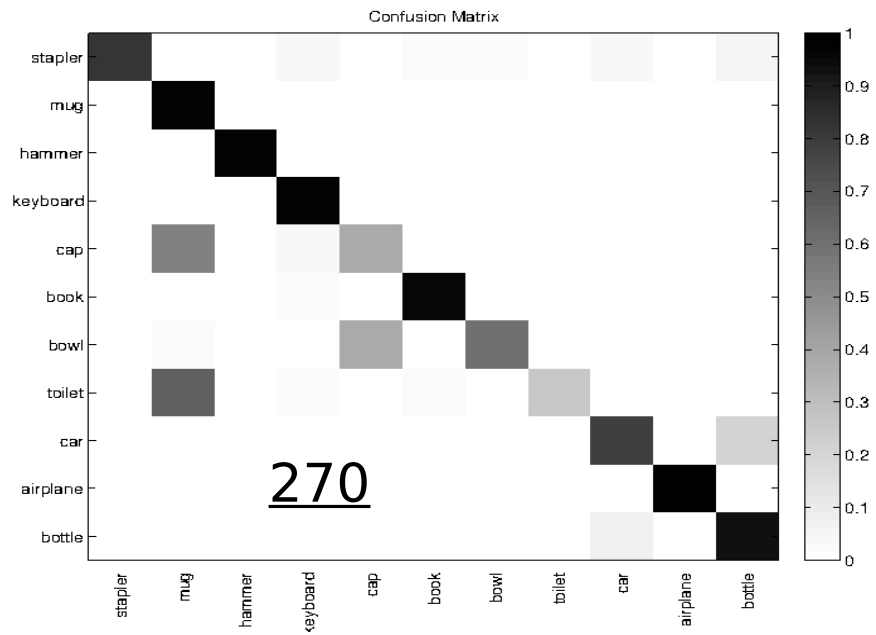
- Recognition from 0, 90, 180 and 270 degrees
- Fused recognition based on multiple views

Experiments Recognition



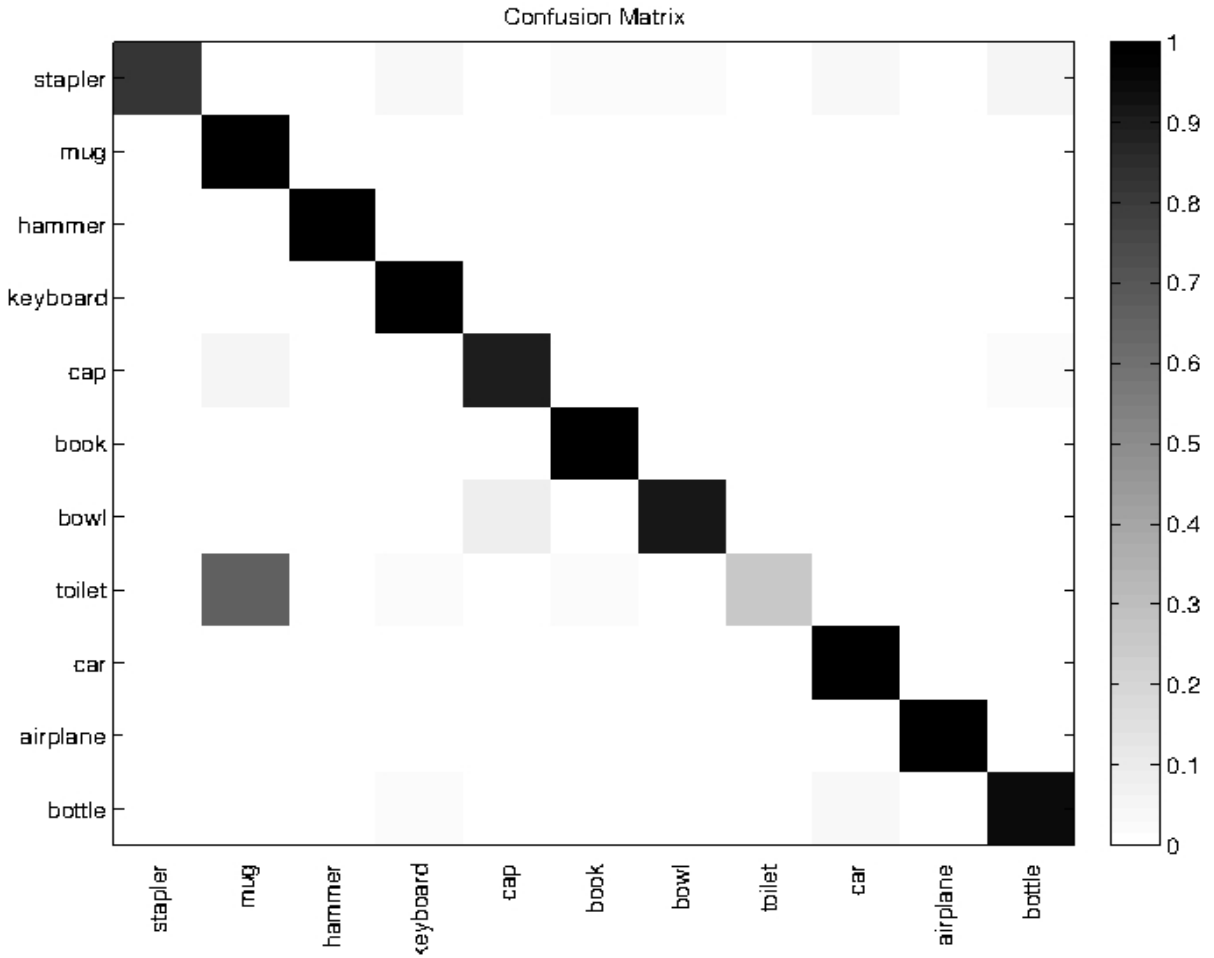






Experiments Recognition

- **Highest Values**



Experiments

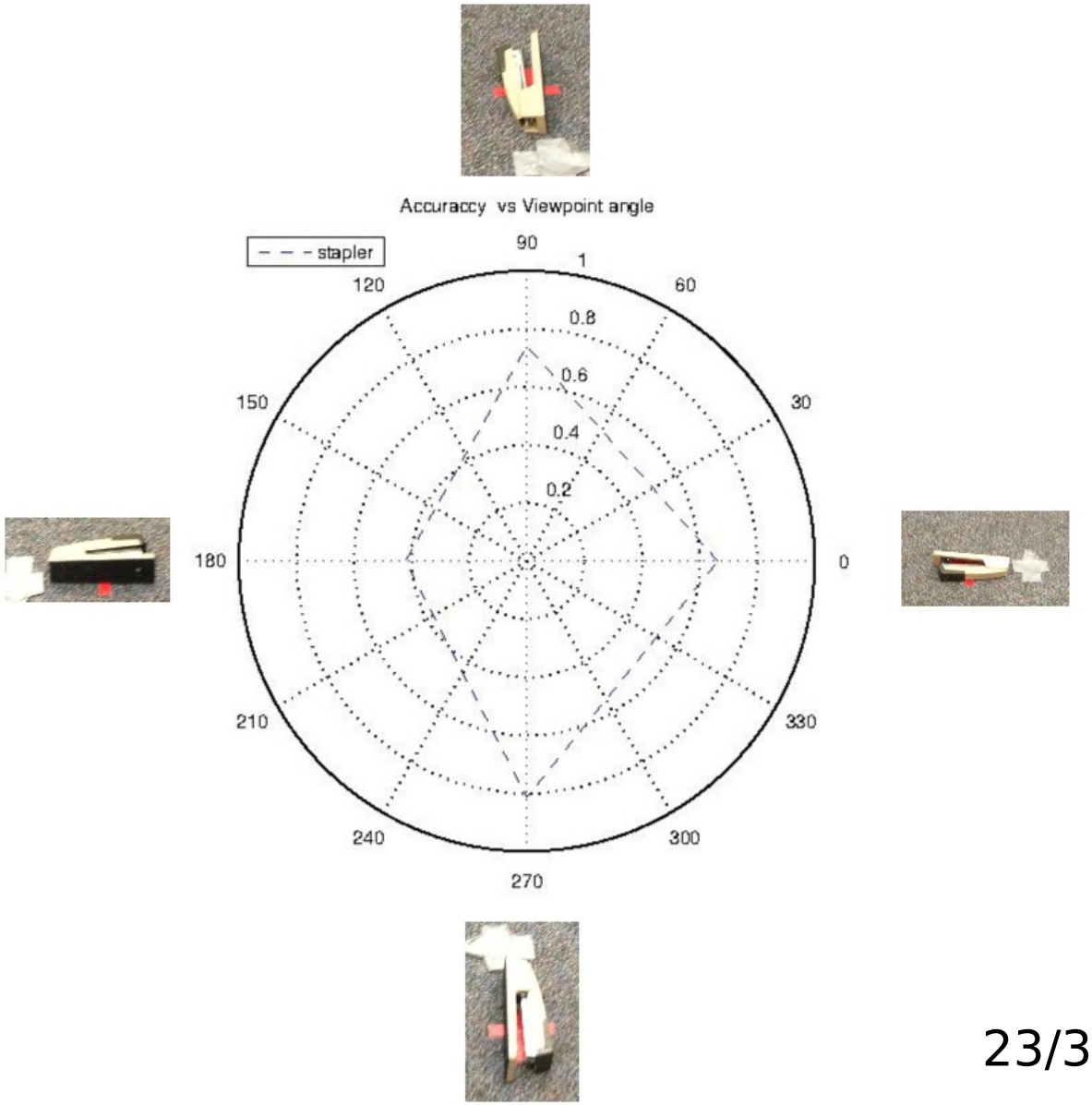
- Dataset
- Time performance
  - Communication
  - System Bottleneck (segmentation)
- Recognition Results
- **Distribution of Viewpoints**

Distribution of Viewpoints

- **Representative views of classes that present significant fluctuations:**
  - Stapler, Cap, Keyboard and Car

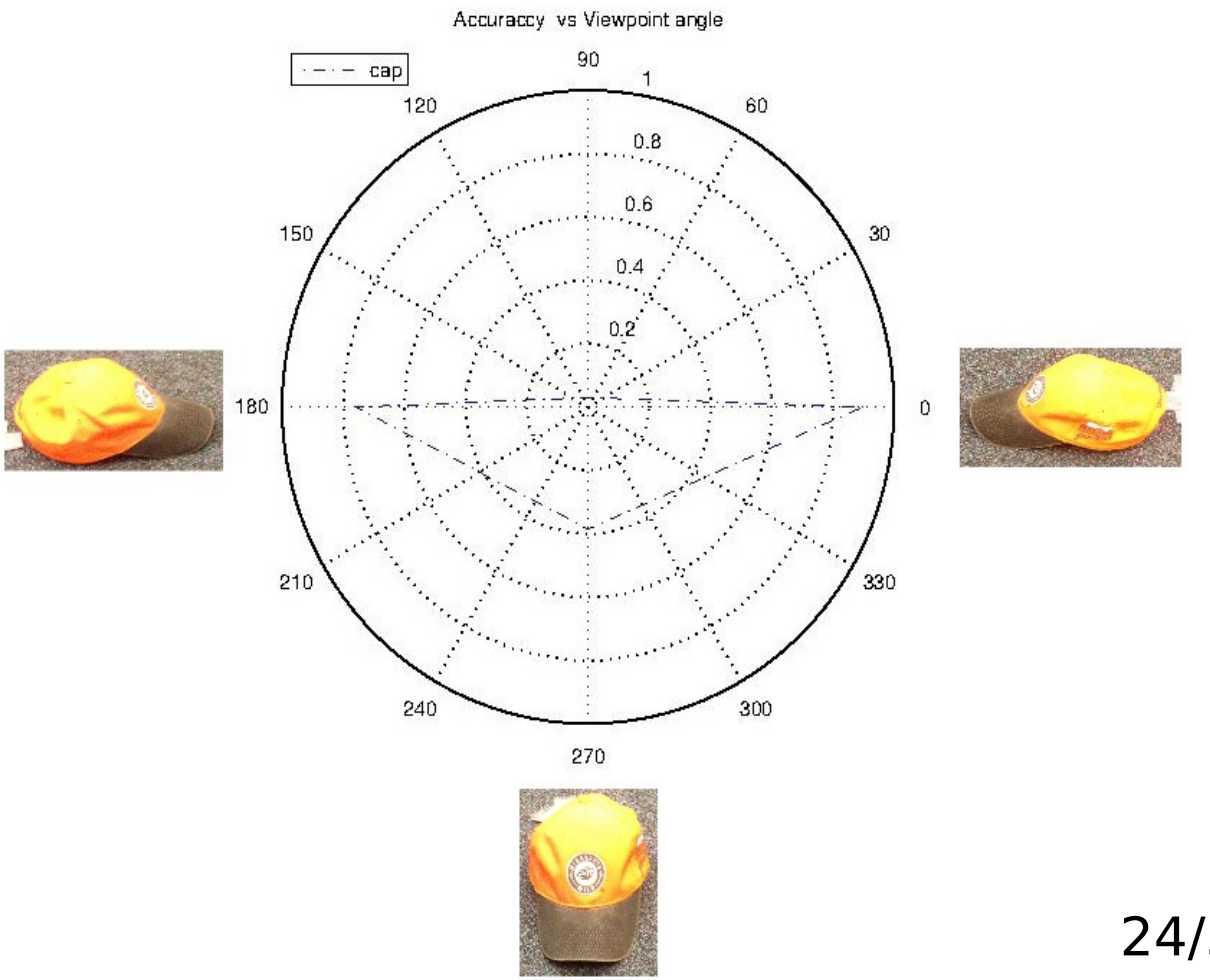
Viewpoints Distribution

- **Stapler**



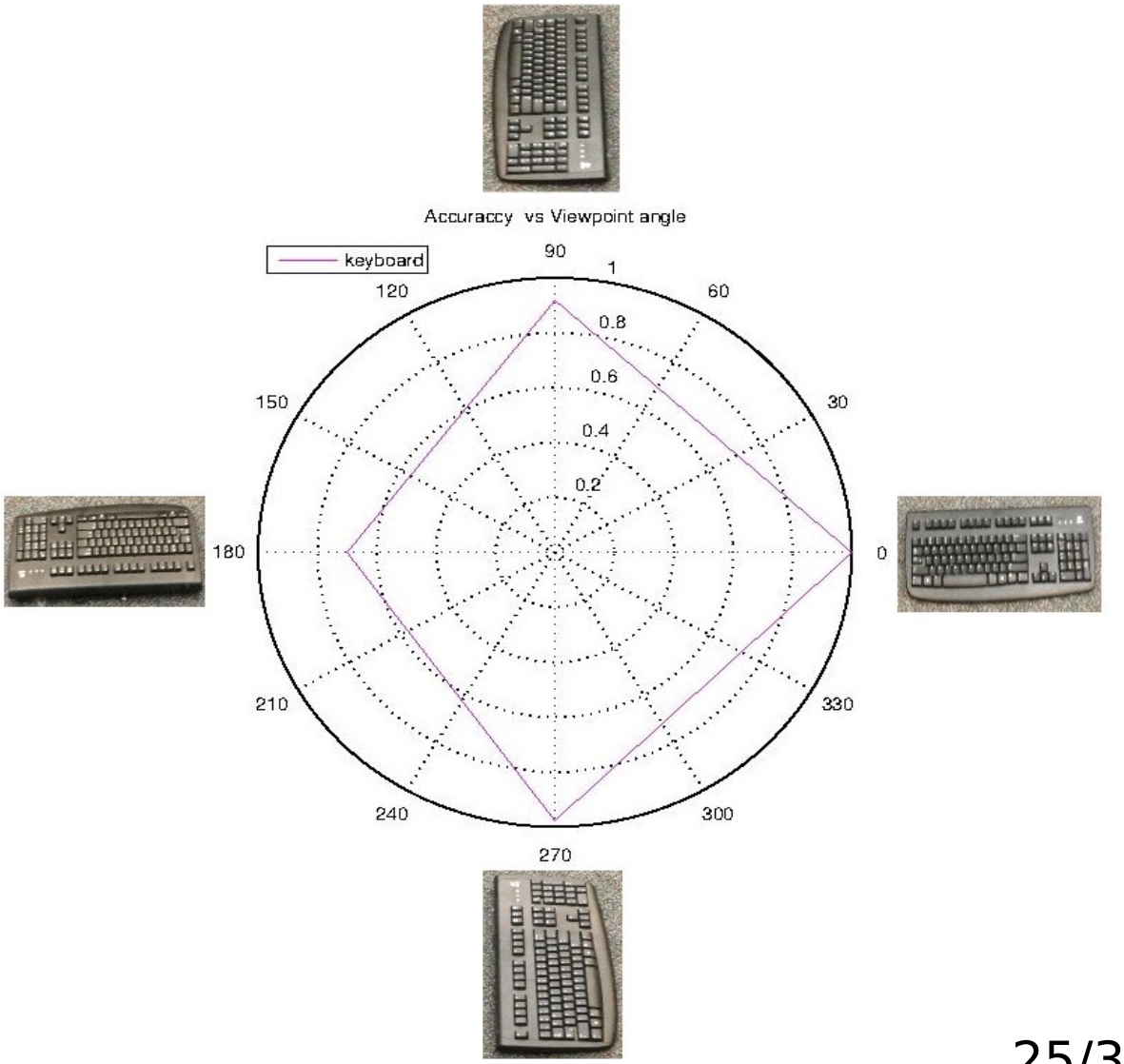
Viewpoints Distribution

- **Cap**



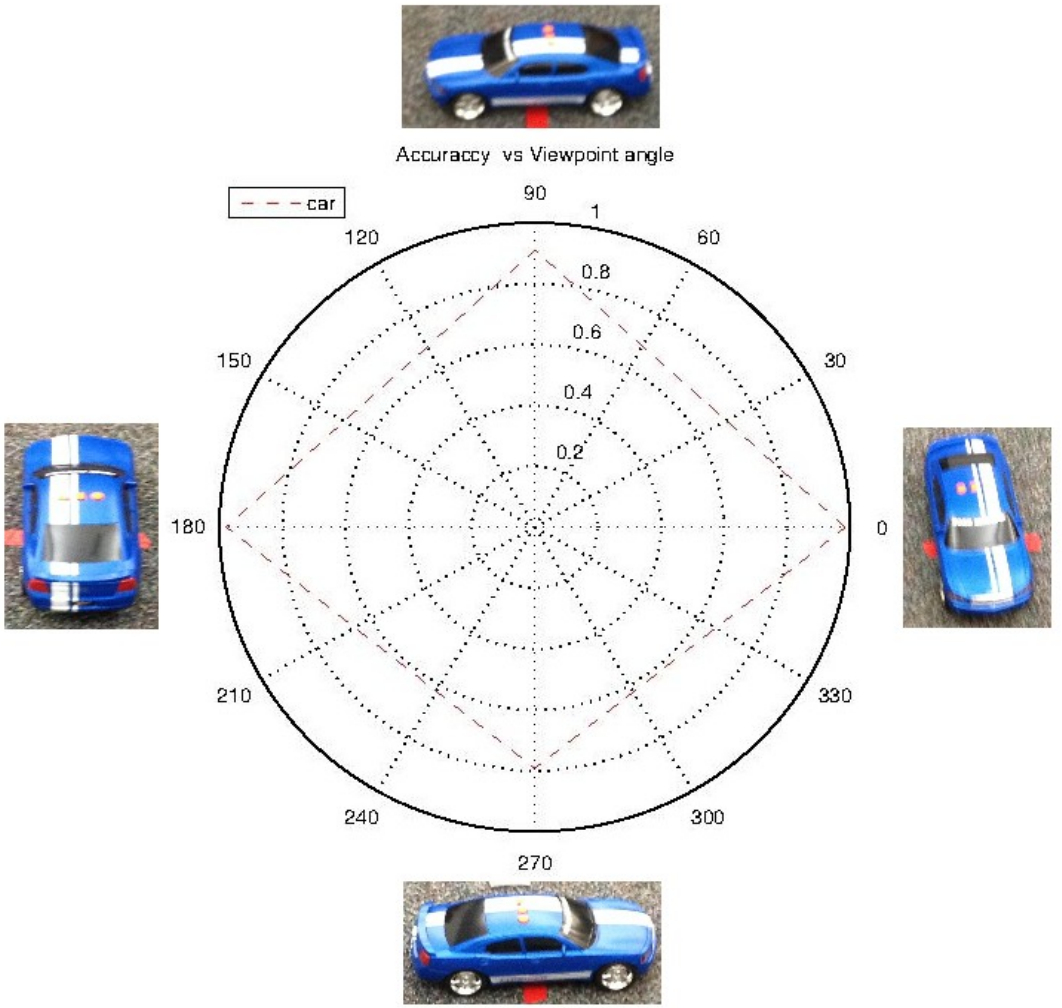
Viewpoints Distribution

- **Keyboard**



Viewpoints Distribution

- **Car**



Conclusions and Future Works

- **Four views show promising results**
- **Our goal is to prove this analytically**

- **System present high recognition rates to most of the objects**

Conclusions and Future Works

- **Test with larger datasets**
- **Refine or propose new approaches to:**
  - Segmentation
  - Partial Viewpoint generation for training

Thanks

- **Contact:**
  - **olvieira@cs.umn.edu**
  - **rsn.cs.umn.edu**

UNIVERSITY OF MINNESOTA
Driven to Discover℠