

Fine-grained Image Classification

Yasser Souris

Supervisor:

Prof. Shohreh Kasaei

Image Processing Lab
Computer Engineering Department
Sharif University of Technology



Outline

- **Introduction**
- **Review**
 - **Image Classification**
 - **Fine-grained Image Classification**
- **Baseline Methods**
- **Experimental Analysis**
- **Future Work**

Outline

- **Introduction**
- **Review**
 - **Image Classification**
 - **Fine-grained Image Classification**
- **Baseline Methods**
- **Experimental Analysis**
- **Future Work**

Image Classification



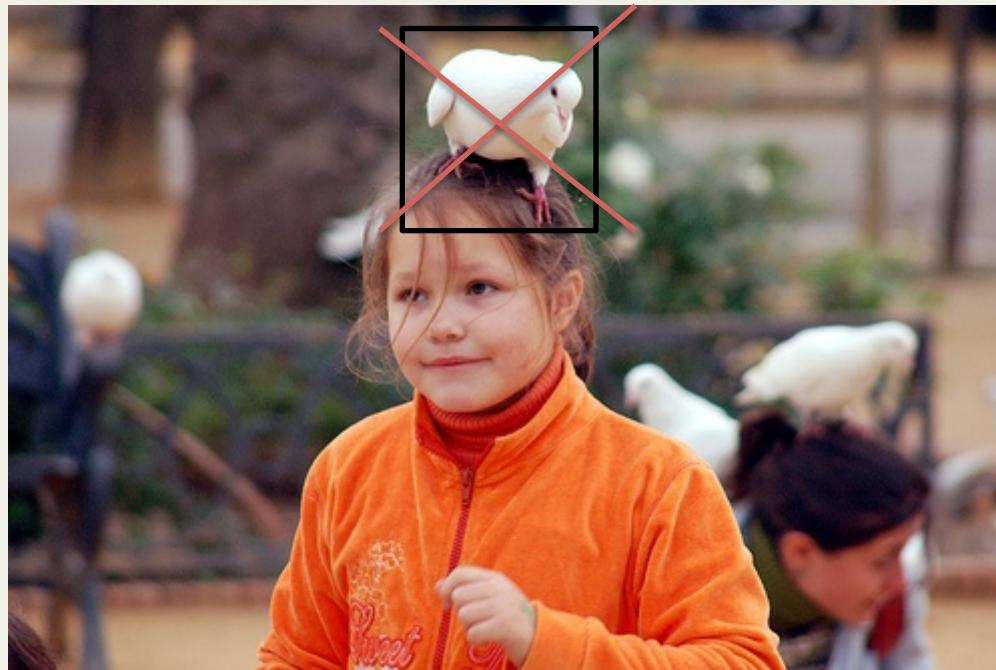
Image Classification



Image Classification



Image Classification



Fine-grained Image Classification

- Image classification, when considered classes are all **subclasses** of a certain class



- They are **similar** in appearance



Outline

- **Introduction**
- **Review**
 - **Image Classification**
 - **Fine-grained Image Classification**
- **Baseline Methods**
- **Experimental Analysis**
- **Future Work**

Image Classification Training/Testing

- Given
 - **Positive training images** containing an object class



- **Negative training images** not containing that object class



- **Classify a test image** whether it contains the object class

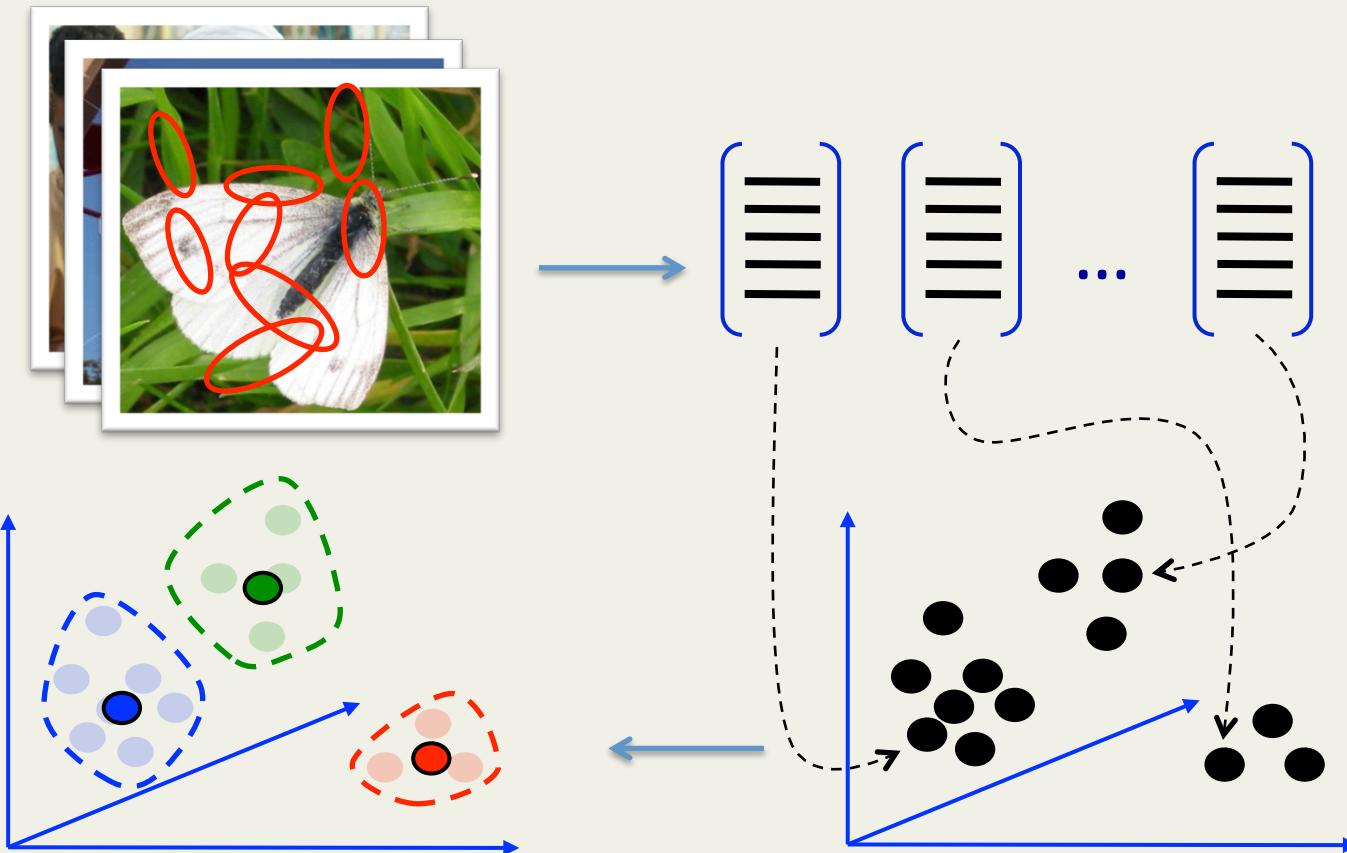


Image Classification Systems

- Vocabulary Based Methods
- Deep Learning Methods

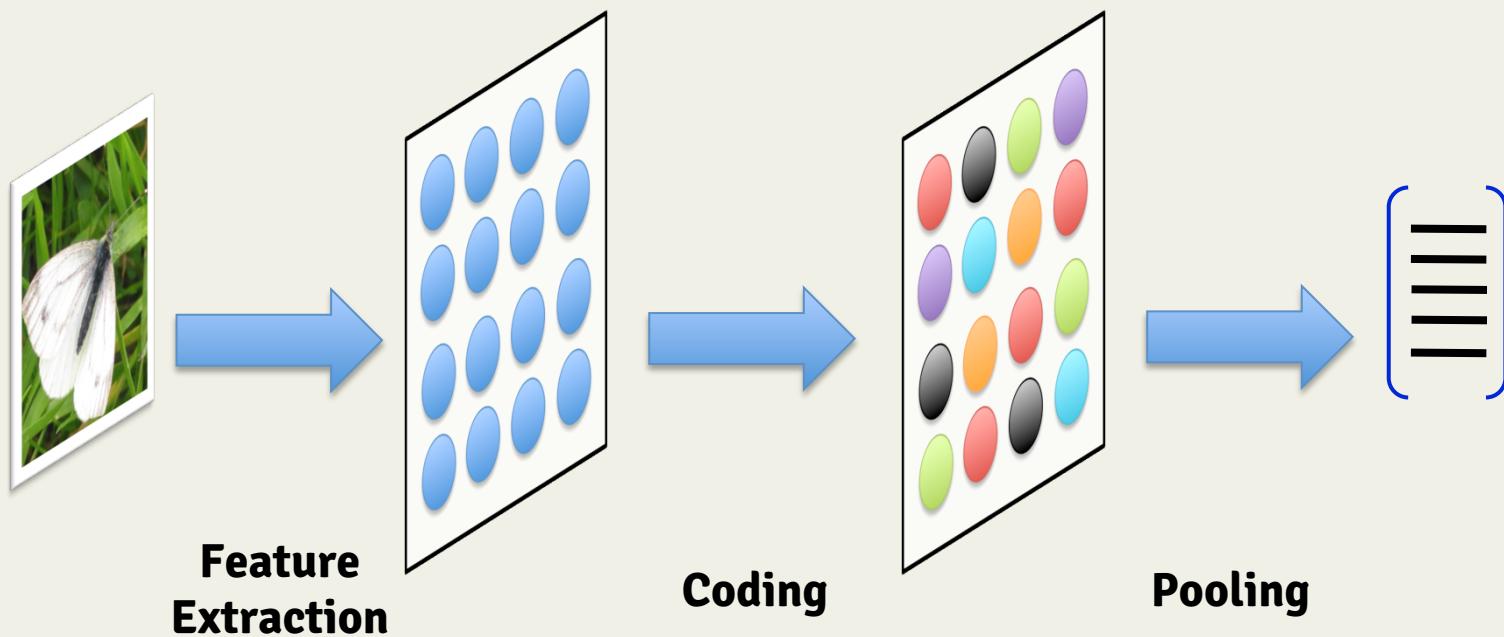
Vocabulary Based Methods

- Building the Vocabulary



Vocabulary Based Methods

- Bag of words

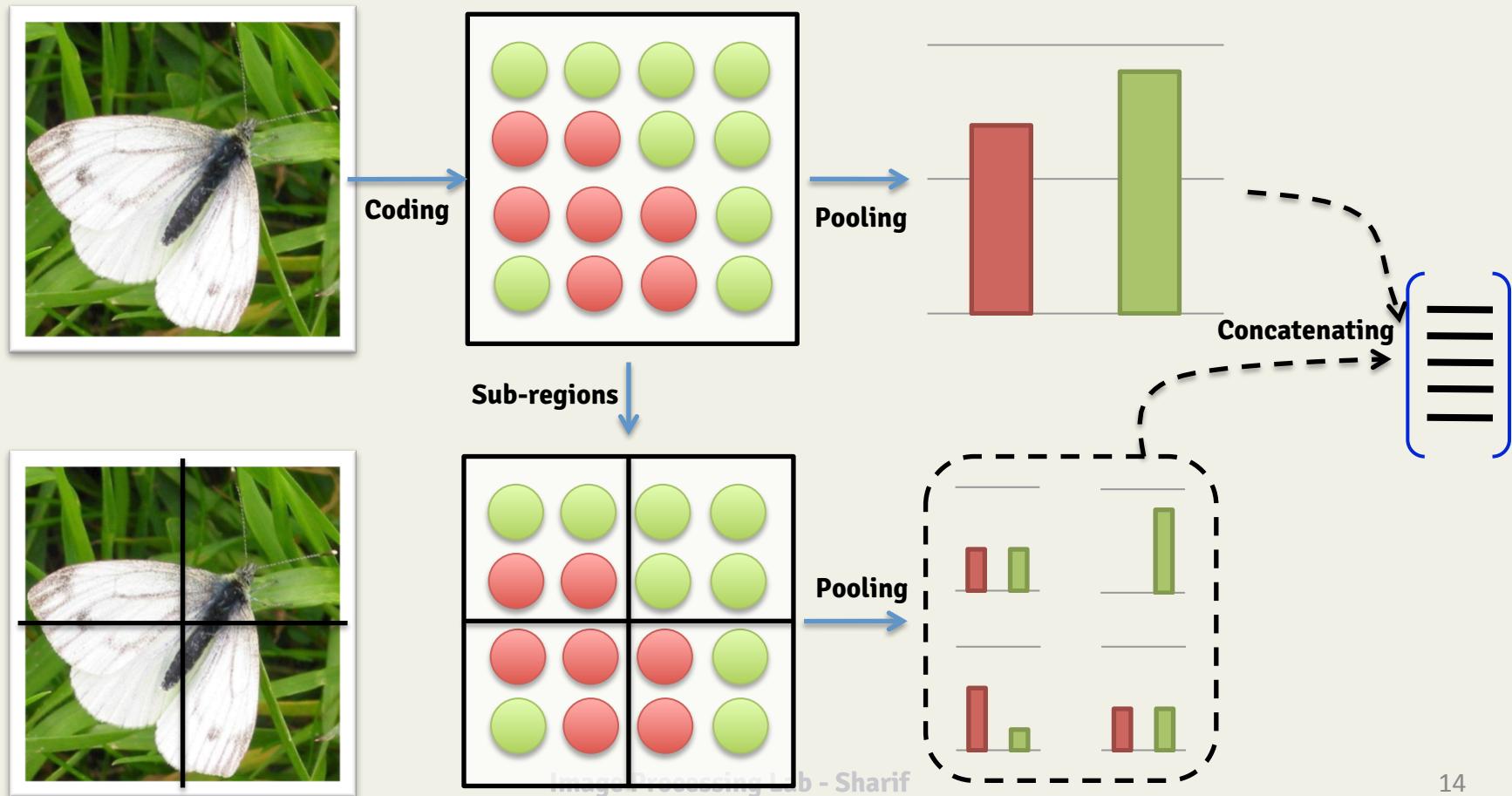


Using geometric invariant features like Harris, Affine, homography, Means

Using descriptors like **SIFT**

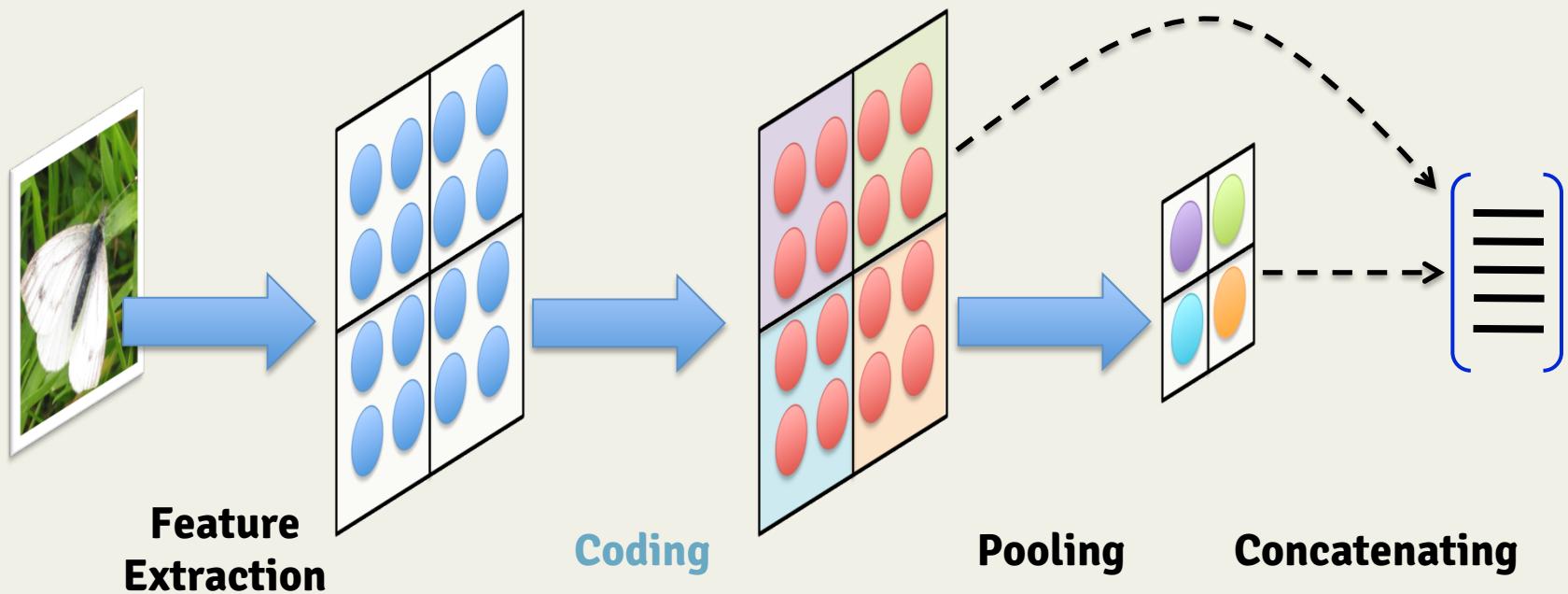
Vocabulary Based Methods

- Spatial Pyramid Matching



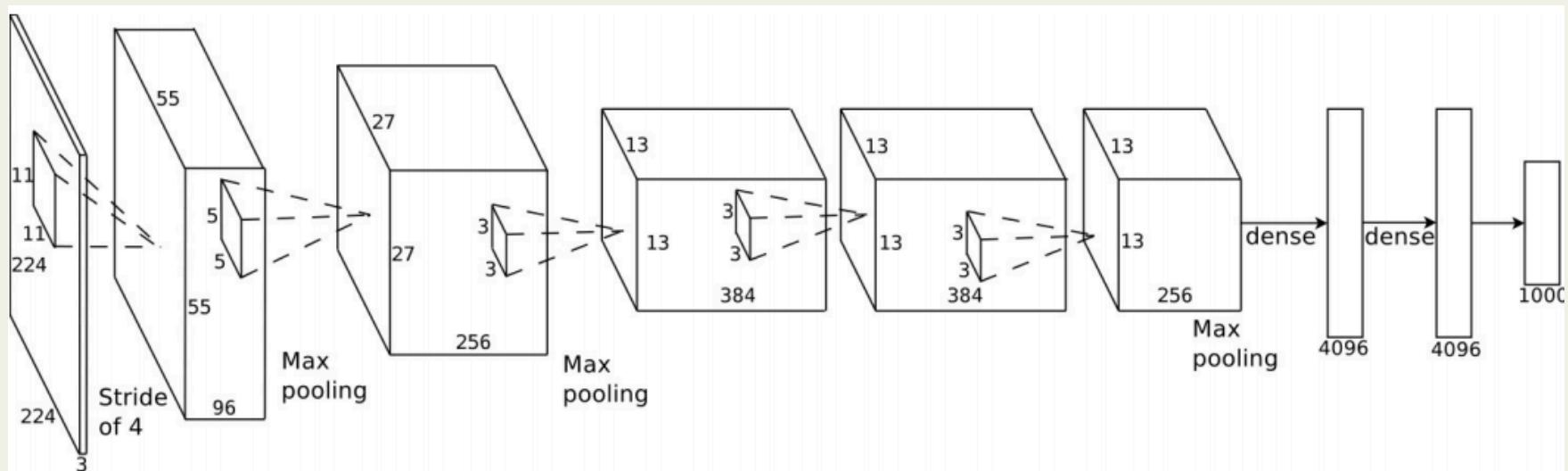
Vocabulary Based Methods

- Overview



Deep Learning Methods

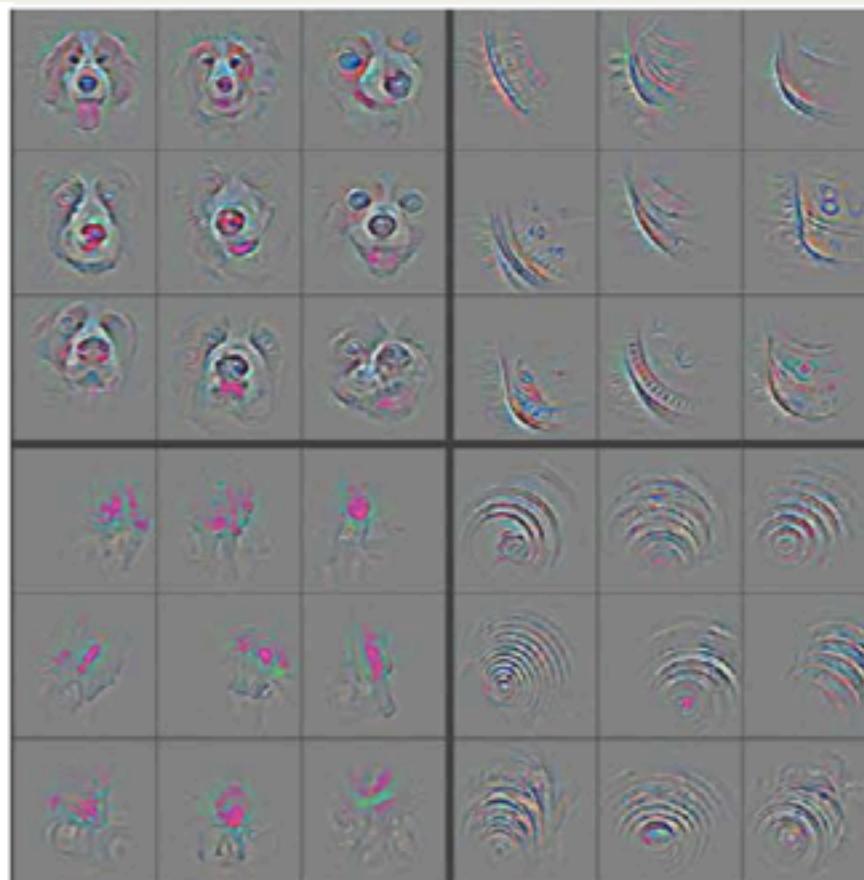
- Network architecture



- Trained with **1.2 million images**

Deep Learning Methods

- Learned Features



Deep Learning Methods

- Feature engineering vs. feature learning
- Results of Pascal VOC 2007 [47, 23]

Method	mAP
SIFT + VQ + SVM	46.54
SIFT + LLC + SVM	57.60
SIFT + FV + SVM	61.69
CNN + SVM	73.90

Outline

- **Introduction**
- **Review**
 - **Image Classification**
 - **Fine-grained Image Classification**
- **Baseline Methods**
- **Experimental Analysis**
- **Future Work**

Fine-grained Datasets



Oxford Airplanes Dataset

Fine-grained Datasets



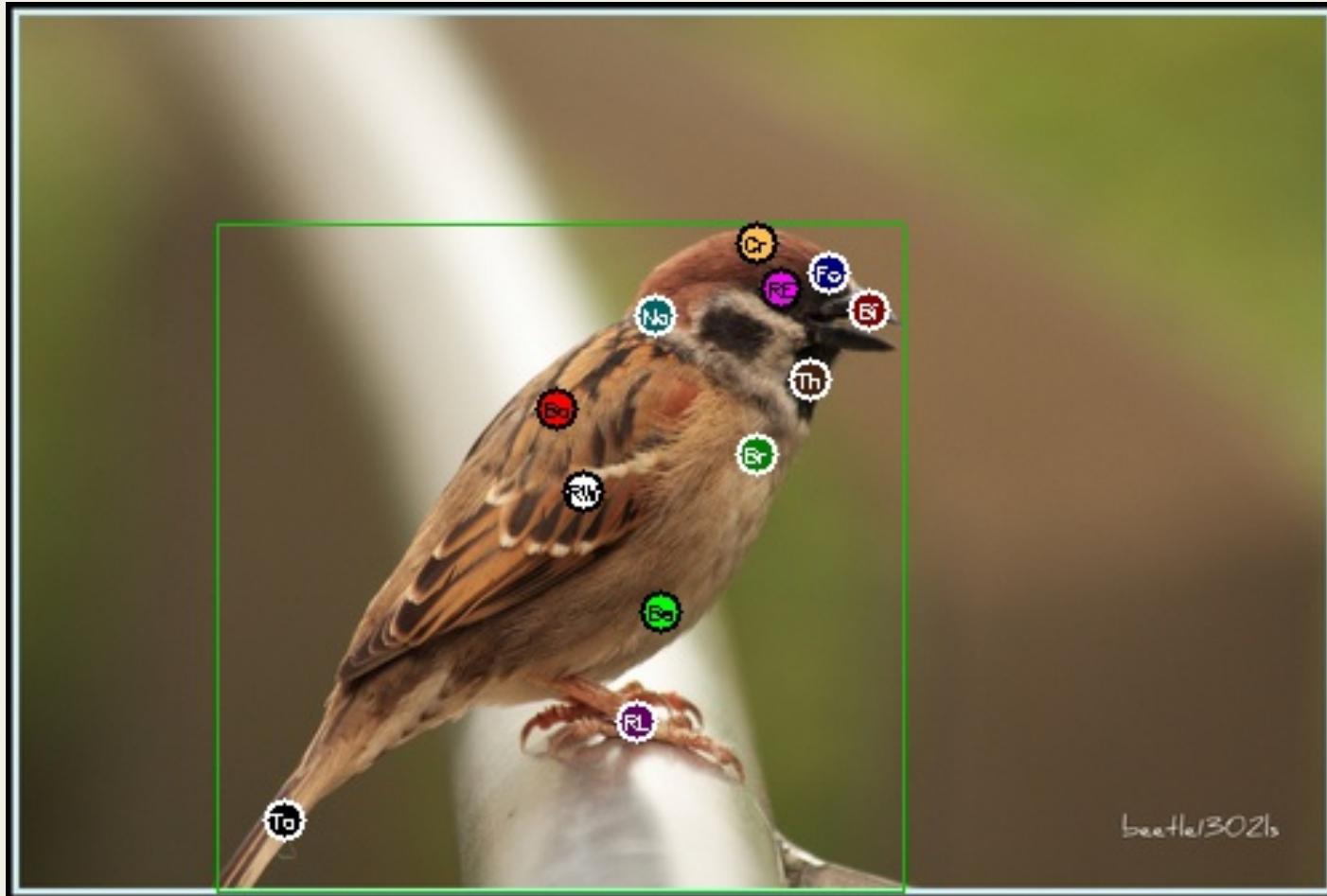
Stanford Cars Dataset

Fine-grained Datasets



Caltech UCSD Birds Dataset

CUB Dataset Annotations



Fine-grained Classification Methods

- Human In the Loop Methods

(A) Easy for Humans



Chair? Airplane? ...

(B) Hard for Humans



Finch? Bunting?...

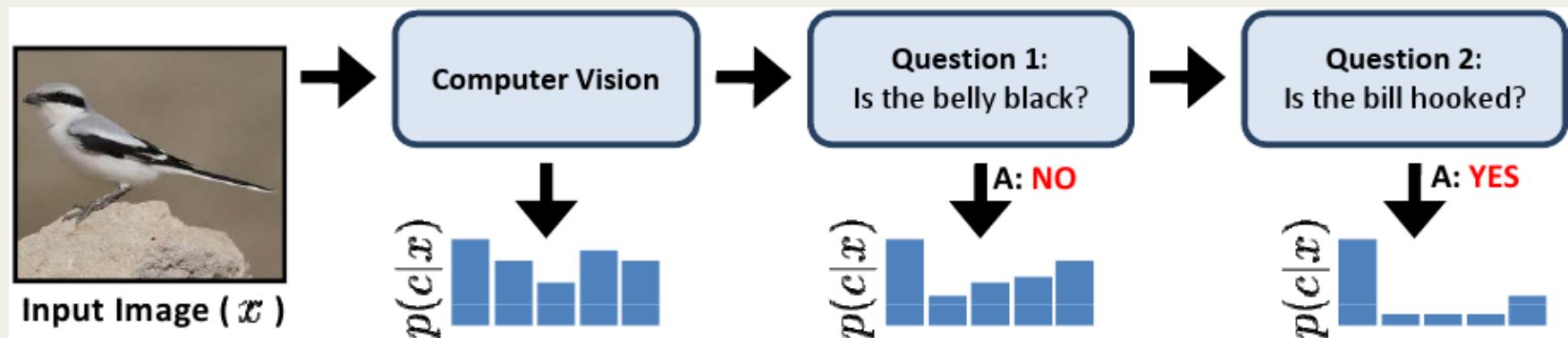
(C) Easy for Humans



Yellow Belly? Blue Belly? ...

Fine-grained Classification Methods

- Human In the Loop Methods



- CUB-2011 Accuracy [38]
 - 0 Q: 19%
 - 20 Q: 50%
- Not fully automatic

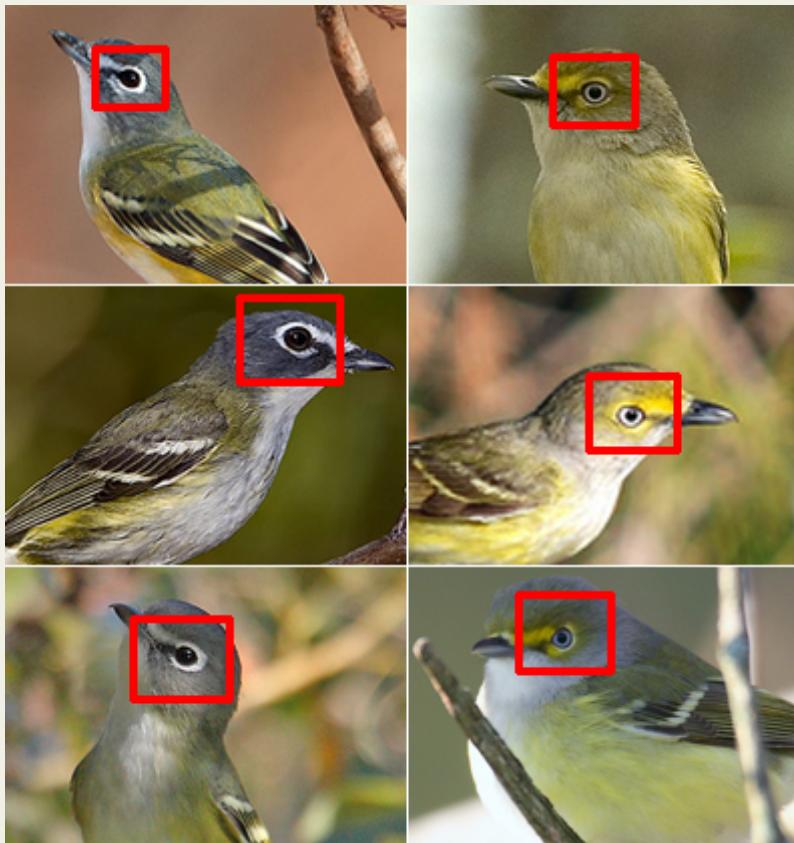
Fine-grained Classification Methods

- Discriminative parts are important



Fine-grained Classification Methods

- Discriminative parts are important



Finding Discriminative Parts

- Is Hard and **Computationally expensive**
- Results in **low accuracies**
- CUB-2011 Accuracy
 - 41.01 % [41]

Fine-grained Classification Methods

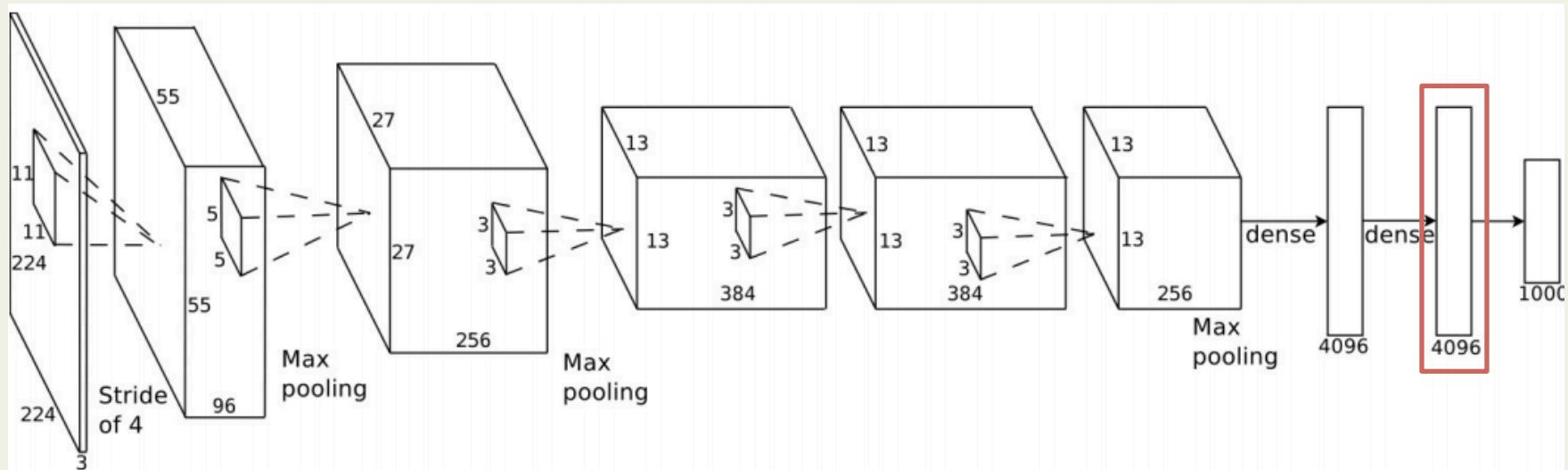
- Deep learning
- Achieving great results
 - 53.3 % [47]

Outline

- **Introduction**
- **Review**
 - **Image Classification**
 - **Fine-grained Image Classification**
- **Baseline Methods**
- **Experimental Analysis**
- **Future Work**

Baseline Methods

- Based on [47]
- Use the network for features
- Use **Linear SVM** for classification



Baseline Methods

- Ours-1
 - Not using the bounding boxes
 - 45.44 %
- Ours-2
 - Cropping with bounding boxes
 - 55.62 %
- Ours-3
 - Ours-1 + Ours-2
 - 57.83 %

Comparison

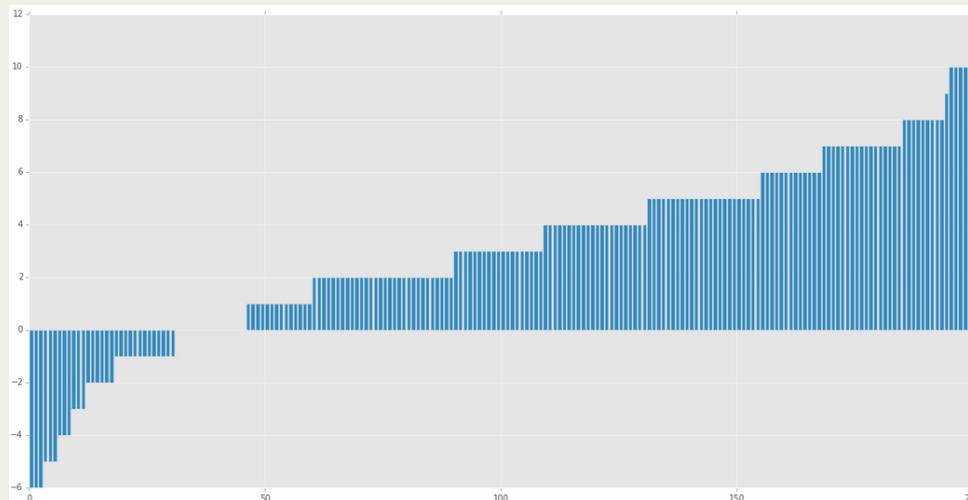
دقت %	استفاده از اطلاعات اجزا	استفاده از جعبه دور	سال ارائه	روش
۱۰,۳		✓	۲۰۱۱	[۴۳] CUB
۱۷,۳	✓	✓	۲۰۱۱	[۴۴] CUB
۲۸,۱۸		✓	۲۰۱۲	[۴۵] PPK
۴۱,۰۱		✓	۲۰۱۲	[۴۶] style
۵۶,۸		✓	۲۰۱۲	[۴۷] POOF
۶۲,۷		✓	۲۰۱۲	[۴۸] Alignments
۵۱,۰	✓	✓	۲۰۱۳	[۴۹] DPD
۶۵,۰		✓	۲۰۱۴	[۵۰] Decafe
۷۵,۷		✓	۲۰۱۴	[۵۱] PNDCN
۵۳,۳		✓	۲۰۱۴	[۵۲] CNN
۶۱,۰۸		✓	۲۰۱۴	[۵۳] CNN-aug
۴۵,۴۴			-	Ours-۱
۵۵,۶۲		✓	-	Ours-۲
۵۷,۸۳		✓	-	Ours-۳

Outline

- **Introduction**
- **Review**
 - **Image Classification**
 - **Fine-grained Image Classification**
- **Baseline Methods**
- **Experimental Analysis**
- **Future Work**

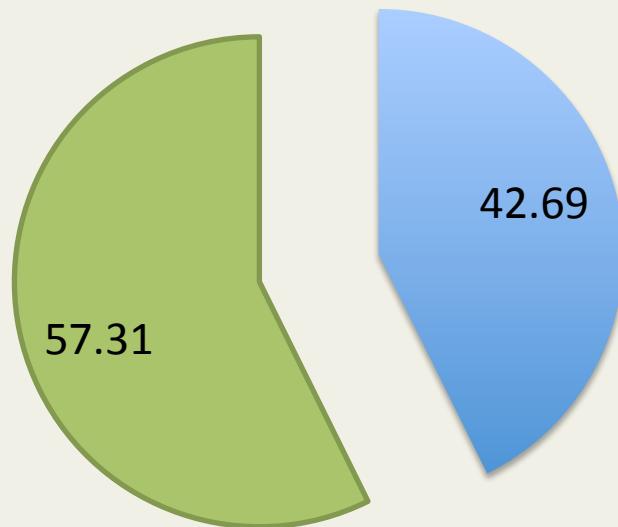
Experimental Analysis

- Ours-4
 - Model averaging
 - 58.00 %
- Ours-2 vs. Ours-1 (+ 10.18 %)
 - Removing background helps



Experimental Analysis

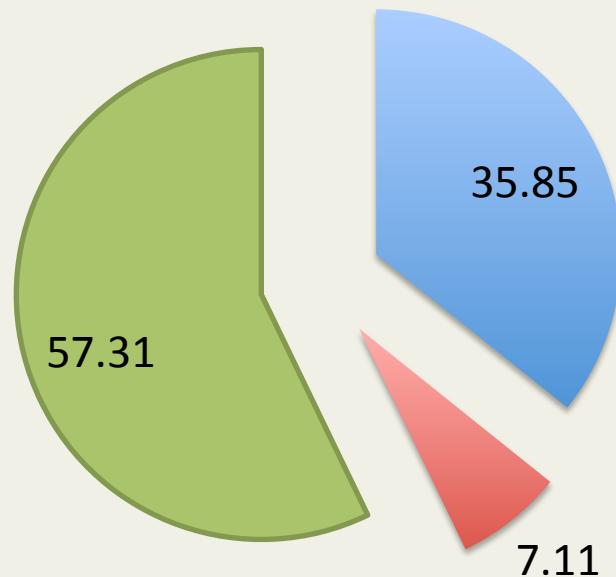
Total Agreement in Our-1, Our-2, Our-3



Experimental Analysis

Total Agreement in Our-1, Our-2, Our-3

Agreement is 84 % correct



Outline

- **Introduction**
- **Review**
 - **Image Classification**
 - **Fine-grained Image Classification**
- **Baseline Methods**
- **Experimental Analysis**
- **Future Work**

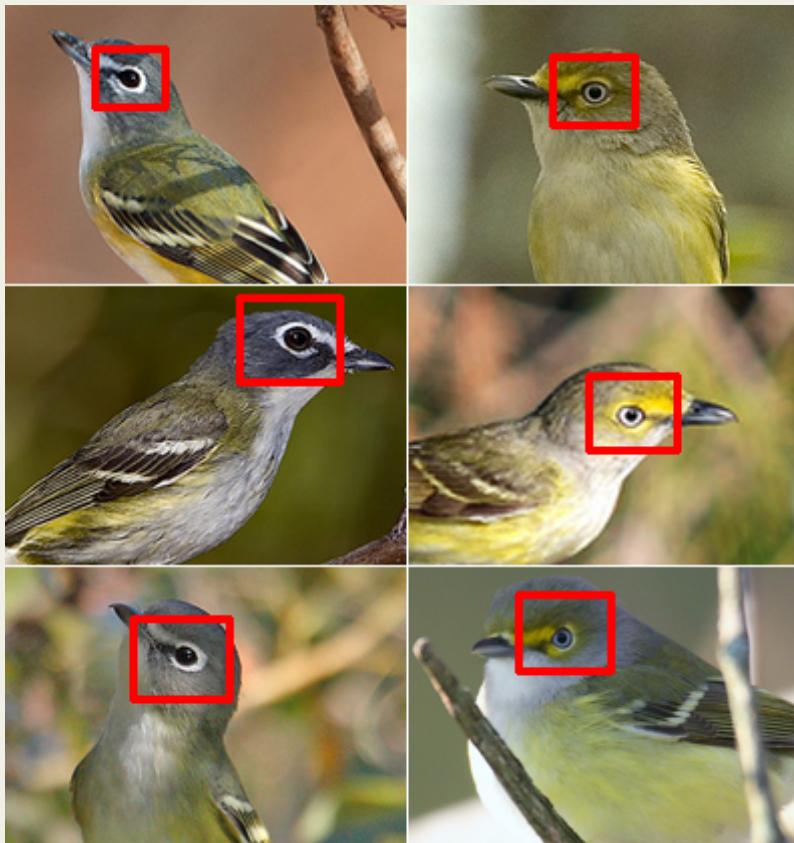
Future Work 1

- Getting the same baseline methods on other datasets



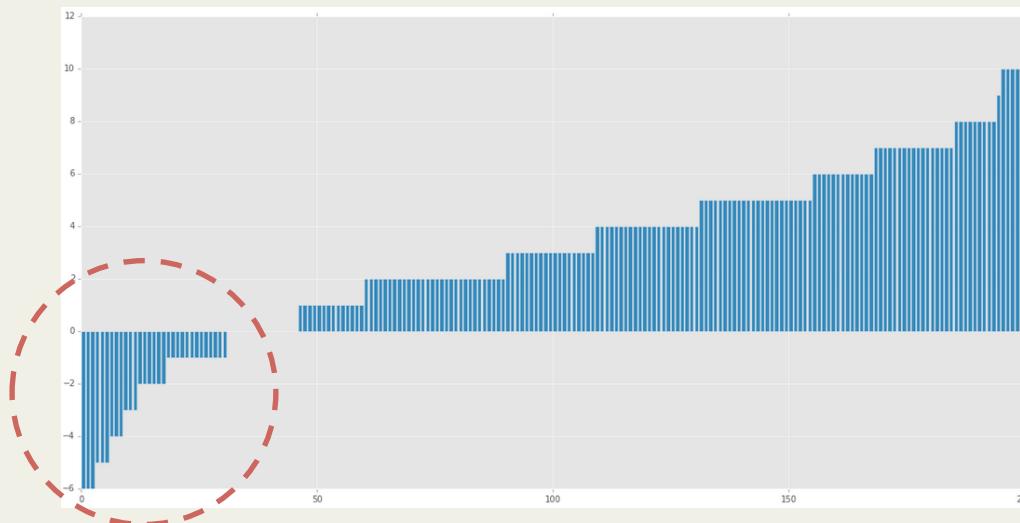
Future Work 2

- Using Part Annotation



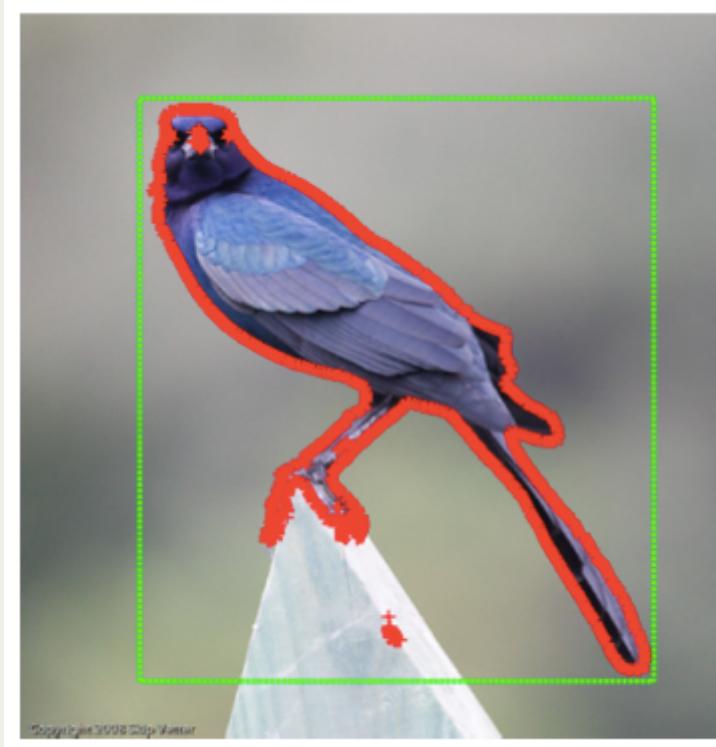
Future Work 3

- Finding the complementary model
- Using mixture of experts



Future Work 4

- Removing background even further with segmentation



Future Work

عنوان فعالیت	مدت زمان لازم	درصد پیشرفت	زمان اتمام
مطالعه روش‌های پیشین برای دسته‌بندی معمولی و دسته‌بندی ریزدانه‌ای	۰,۵ ماه	۱۰۰	شهریور ۹۳
پیاده‌سازی و اجرای روش‌های پایه	۱ ماه	۵۰	آبان ۹۳
پیاده‌سازی روش پیشنهادی جدید	۳ ماه	۲۰	بهمن ۹۳
نگارش پایان‌نامه	۱ ماه	۱۰	بهمن ۹۳
جمع‌بندی نتایج و نوشتتن مقاله	۰,۵ ماه	۰	اسفند ۹۳

Thank you

Summer 2014



Image Processing Lab - Sharif



Vocabulary Based Methods

- **Coding: VQ**

$$\arg \min_C \sum_{i=1}^N \|x_i - Bc_i\|^\gamma$$

s.t. $\|c_i\|_{\ell^1} = 1, \|c_i\|_{\ell^1} = 1, c_i \succeq 0, \forall i$

- **Coding: SC**

$$\arg \min_C \sum_{i=1}^N \|x_i - Bc_i\|^\gamma + \lambda \|c_i\|_{\ell^1}$$

Image Classification Methods

- Analysis

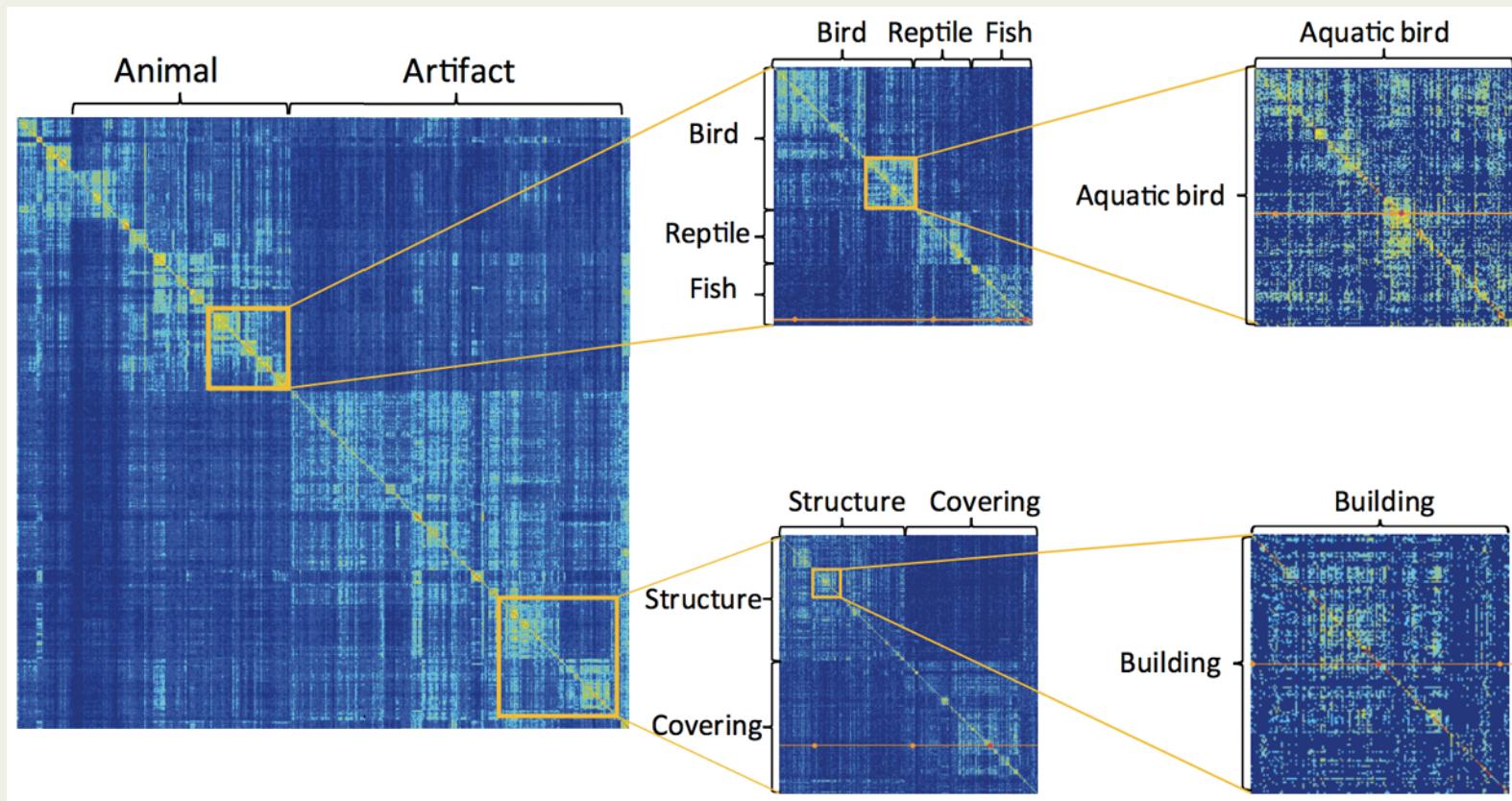


Image Classification



Image Classification

